Supplemental Material

DATA ACCESS

RNA-seq raw data files and a summary table of the results are accessible at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) through the accession GSE44564. Proteomics data associated with this manuscript can be downloaded from ProteomeXchange (http://www.proteomexchange.org/) through the accession PXD000153.

SUPPLEMENTAL FIGURES AND TABLES

Table S1. Plasmids and Strains used in this study.

Plasmid	Description	Source or reference
pDT024	*Kmr Tra− mob+ oriRSF1010 (IncQ), laqIq, Ptaclac-batR*	(Quebatte et al. 2010)
pTR1000	*oriT* ori_{ColE1} gfp_{mut2} $lacI^q$ *rpsL* Km^r; mutagenesis vector	(Schulein et al. 2005)
pMQ009	Derivative of pTR1000 used for *batR-batS* in-frame deletion	This work
pMQ012	Derivative of pPG612 used for chromosomal integration of P_{bepD}:gfp_{mut2}	This work
pIT011	gfp_{mut2} fused to P_{bepD} (−333 to +13) of RSE247	(Quebatte et al. 2010)
pPG161	Derivative of pTR1000 used for *bepD* in-frame deletion	(Scheidegger et al. 2009)
pPG612	Derivative of pPG161 with homology regions for insertion in *Bh* chromosome	(P. Guy, unpublished)
RSE247	Spontaneous Sm^r strain of *Bartonella henselae* ATCC 49882 T	(Schmid et al. 2004)
MQB242	*batR-batS* in-frame deletion mutant of RSE247	This work
MQB277	Derivative of MQB242 carrying chromosomal integration of P_{bepD}:gfp_{mut2}	This work
MQB307	MQB277 carrying pDT024	This work

Table S2. Oligonucleotides used in this study.

Name	Sequence
prAB007	GCTCTAGATTAAGCACGGTCAATTTCAGG
prMQ1088	TTATCGAGCGAAGGGAAGGGTGATTGTTGGATTATCTTTCAT
prMQ1091	CCTTCCCTTCGCTCGATAAACC
prMQ1092	GCTCTAGATAATATCGCCTCGGCGTTGATC
prMQ1191	AAGCGGCCGCTATTATTTGTATAGTTCATCCATGC
prMQ1192	TTGCCGCGGCTAGTAAAGGAGAAGAACTTTTCAC

Table S3. Overview of the number of RNA-seq reads in the four samples.

Sample	uninduced 1	uninduced 2	induced 1	induced 2
Physical reads	59,790,762	54,912,301	73,998,515	86,625,716
Uniquely mapping reads	16,059,124	10,676,768	26,427,712	23,541,042
Multiple mapping reads	43,731,638	44,235,533	47,570,803	63,084,674
16S, 23S rRNA genes (2 copies)	40,164,227	39,840,893	41,265,157	58,559,807
5S rRNA genes (3 copies)	2,358,053	3,773,825	4,171,233	2,188,380
percentage rRNA reads among multiple mapping reads	97.2%	98.6%	95.5%	96.3%
Filtered reads (uniquely mapped)	16,173,339	10,699,494	26,730,604	23,756,851

The total reads, divided into uniquely mapping and multiple mapping reads are shown (including the reads accounted for by the rRNA genes). The final number of uniquely mapped reads after our additional filtering steps (see Figure S12 and Methods) is shown and forms the basis for all further analyses. The successful rRNA depletion increased the mRNA percentage between 4 to 7-fold (from an estimated 5% in the total RNA to 20-35% after depletion; data not shown).

Figure S1: Correlation of RNA-seq transcription profiles among biological replicates.

RPKM values (Mortazavi et al. 2008) were calculated for 1,488 predicted *B. henselae* protein-coding ORFs based on unambiguously mapping, filtered reads and displayed in a scatterplot using a density function that assigns darker shades of grey to populated areas. The Spearman's rank correlation coefficient for the biological replicates uninduced1/uninduced2 and induced1/induced2 is shown in the lower right of each plot.

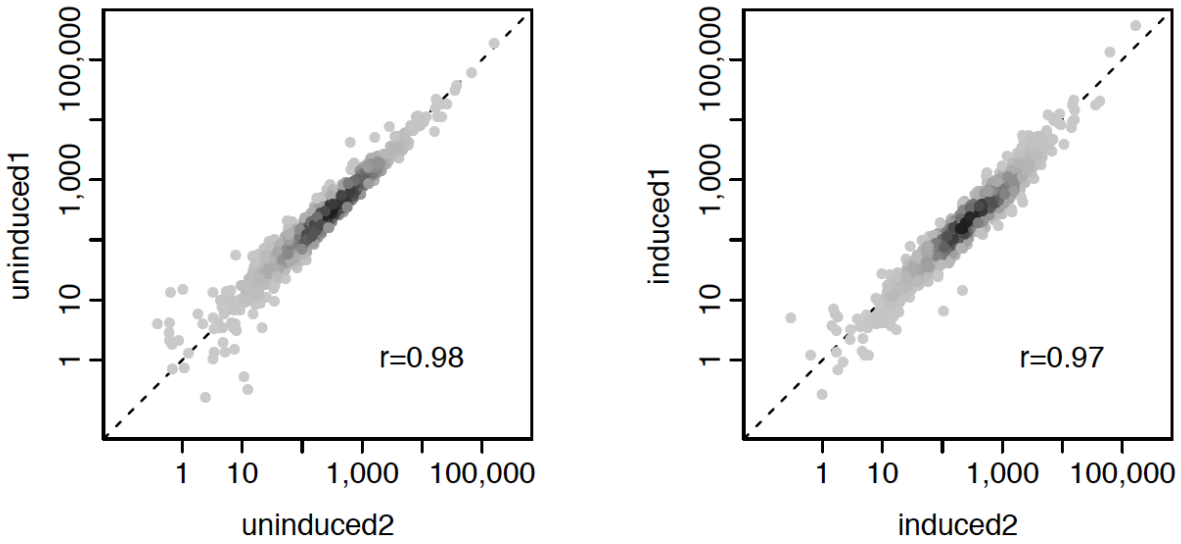


Figure S2: Coverage of protein-coding ORFs by RNA-seq reads.

The median value of reads per bin (with a total of 100 bins) averaged over the length of all 1,488 protein-coding ORFs is shown. In addition, we show the median value of RNA-Seq reads for 10% of base pairs both upstream and downstream relative to the protein-coding ORFs. A higher median of RNA-Seq reads at the 5' end of the protein-coding ORFs is apparent. This bias is different from the 3' end bias that is observed for oligo-dT-primed libraries (Nagalakshmi et al. 2008). It could be a consequence of using a protocol that relies on random hexamer-generated cDNA libraries (Roberts et al. 2011), or treating the samples with Invitrogen's riboMinus kit (Tariq et al. 2011) to remove highly abundant rRNA reads (see Methods). For both methods, higher coverage in the 5' end has been reported. Alternatively, it could also be a consequence of degradation of labile prokaryotic mRNA from the 3' end (Kennell 2002), or a combination of the above.

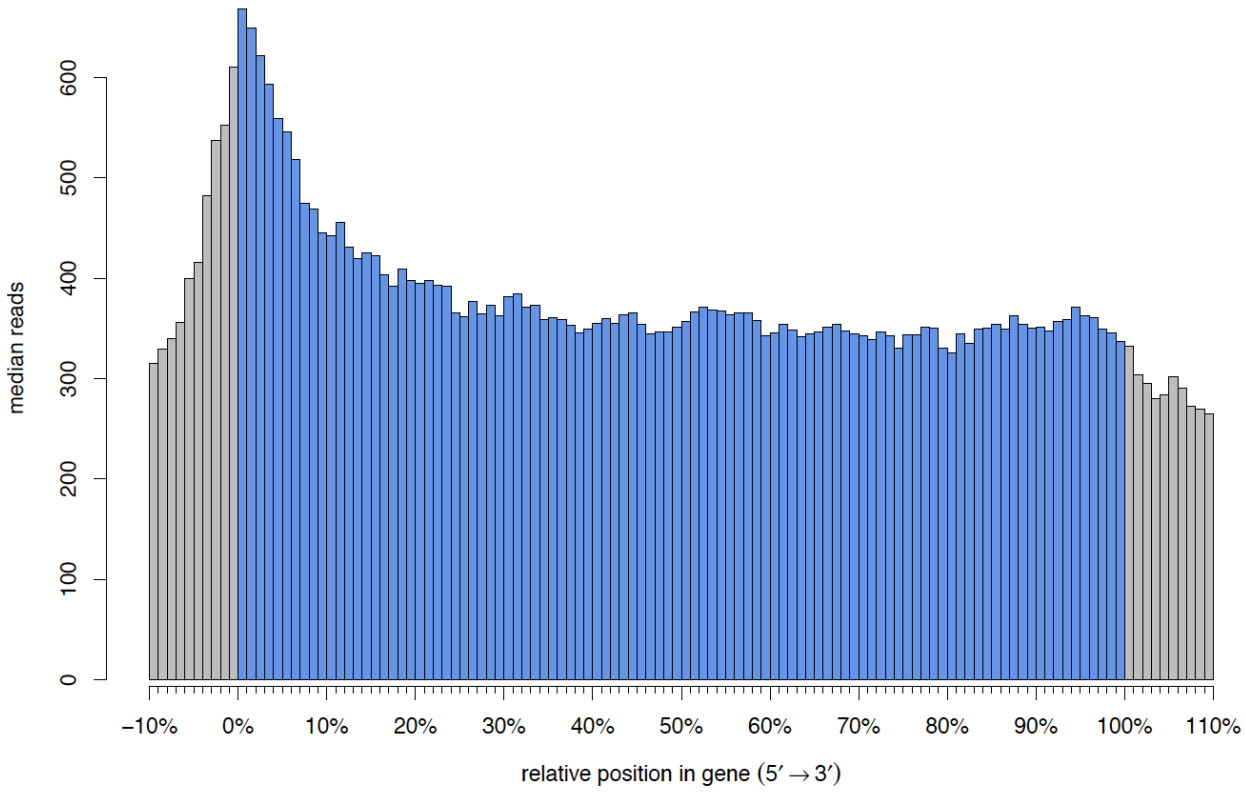


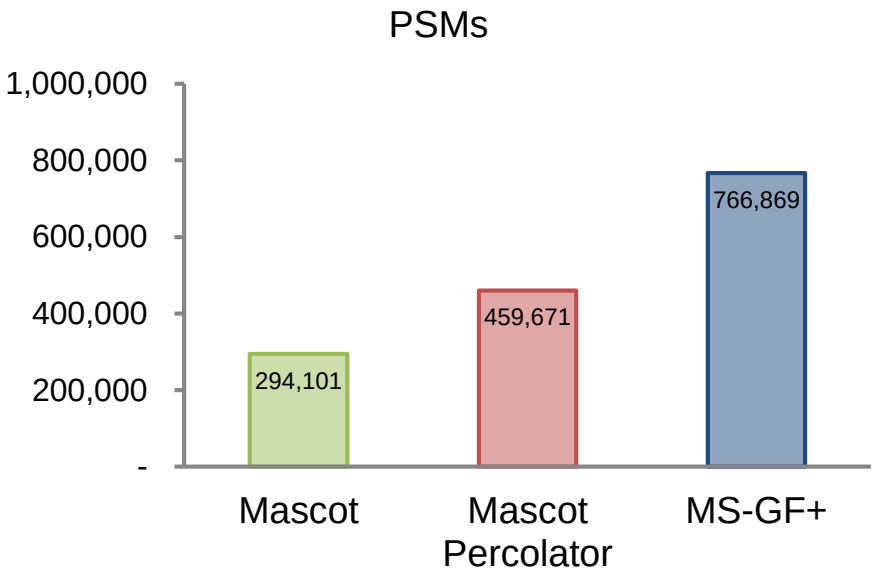
Table S4: Inter-replicate agreement concerning the number of actively expressed protein-coding ORFs.

The inter-replicate analysis identified a very high overlap of the protein-coding genes that are called actively expressed based on a combined threshold of greater than 10 RPKM and 5 or more distinct reads in the 5' end of the protein-coding ORF.

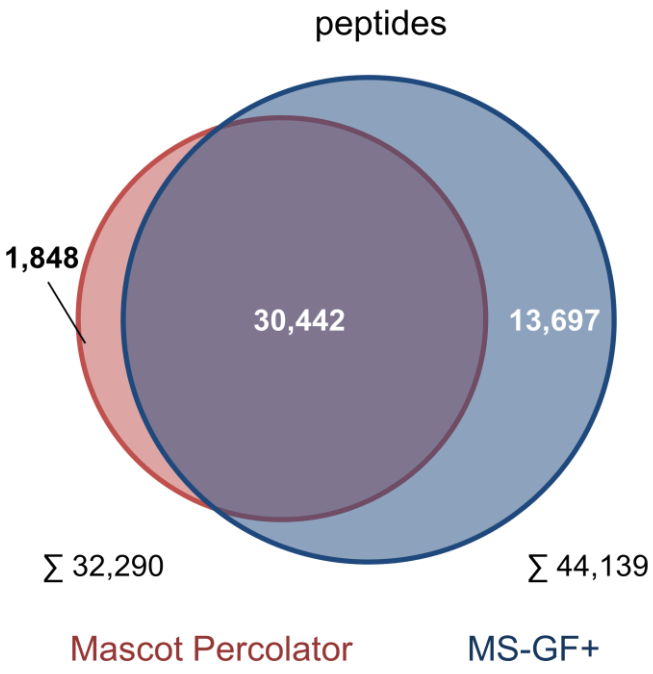
Condition	# protein-coding genes replicate 1	# protein-coding genes replicate 2	Overlap
uninduced	1,284	1,254	1,234/1,304 (95%)
induced	1,347	1,349	1,332/1,364 (98%)

Figure S3: Overlap of the proteomics search results of two different search engines.

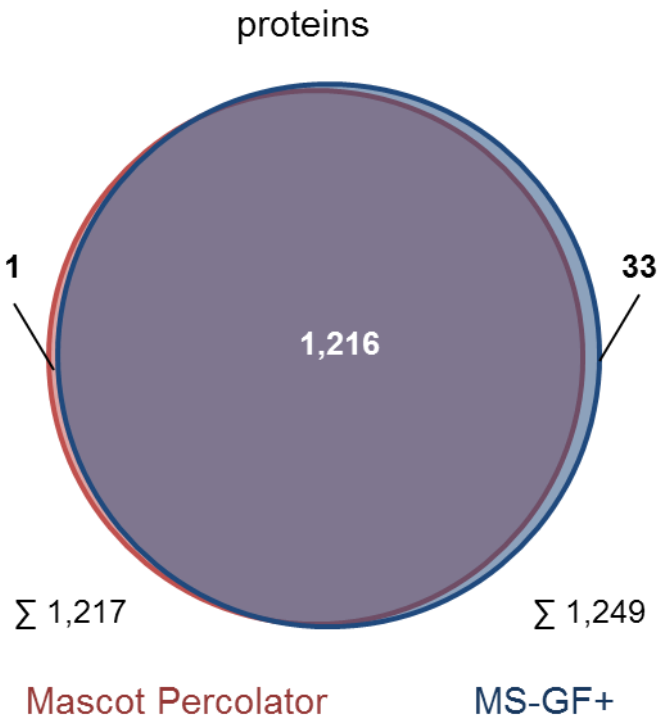
A. Comparison of the total number of PSMs reported at an FDR of 0.01%. Using Mascot in combination with Percolator, 56% more spectra can be assigned compared to Mascot alone. In turn, MS-GF+ assigns another 67% more spectra compared to Mascot-Percolator using the same stringent FDR cutoff.



B. Venn diagram showing the overlap at the level of the identified peptides. The number of distinct *B. henselae* peptides that were identified is shown, without considering modifications. Searching with MS-GF+ results in 37% more peptide identifications.



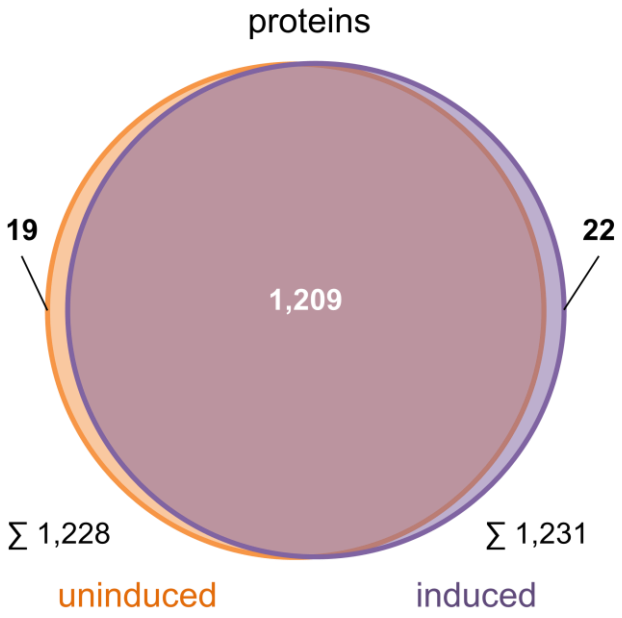
C. Venn diagram showing the overlap at the protein level. We only considered *B. henselae* proteins / protein groups that were unambiguously identified by class 1a or class 3a peptides, with a total of 2 or more PSMs over all experiments. Despite the vast amount of additional PSMs below the FDR threshold (> 300,000), MS-GF+ identifies only 33 additional protein groups.



We interpret this (together with several other lines of evidence) as indication that we have sequenced the protein extracts to saturation at the protein level using the applied discovery proteomics approach. This is supported by the fact that when searching all data with Sequest, we only found evidence for one additional protein not identified by Mascot-Percolator and MS-GF+.

Figure S4: Overlap of proteins detected in uninduced and induced state.

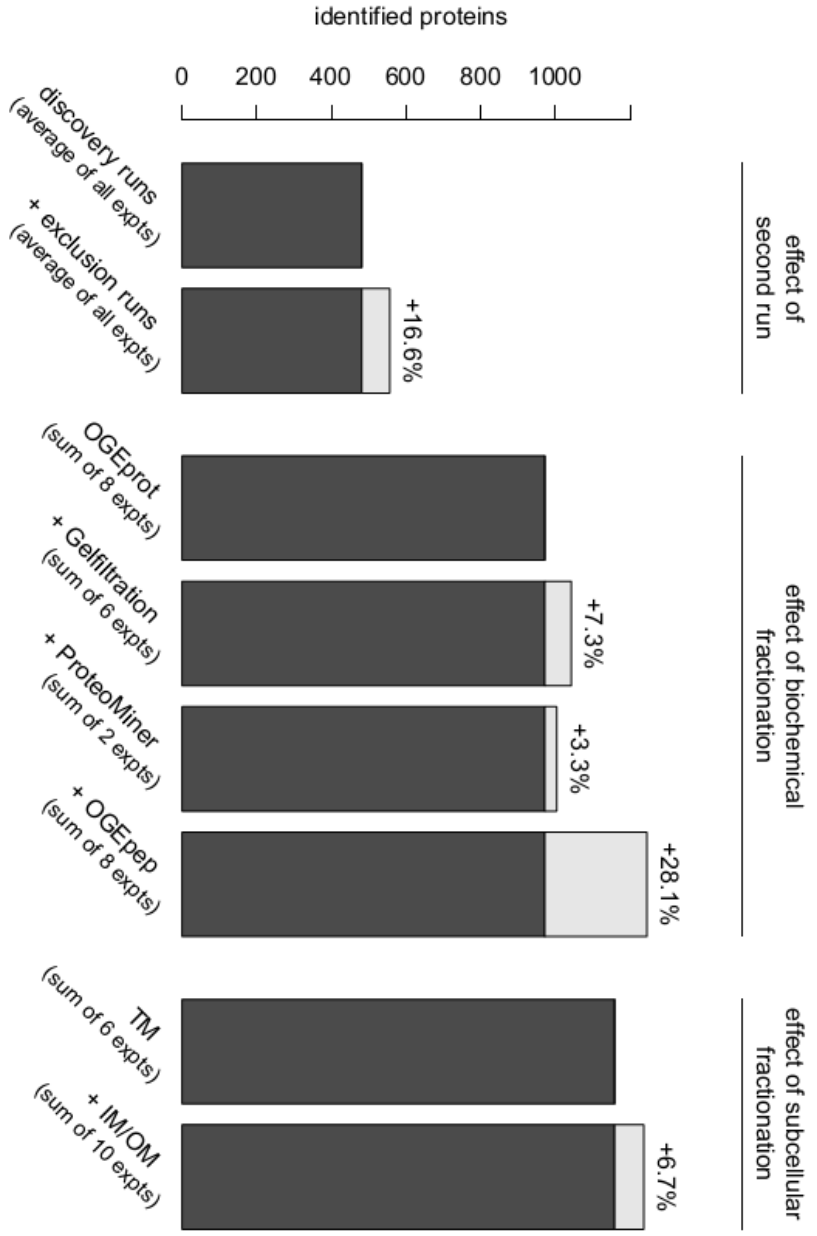
Venn diagram showing the overlap of proteins identified in the uninduced (orange) and induced (violet) condition. For proteins identified only in one condition, we list their accession number, gene name, and the spectral counts provided by MS-GF+ and Mascot.



protein	gene name	spectra MSGF+/ Mascot
BH01150	*ccmA*	9/2
BH03150		4/2
BH03170	*gp26*	4/0
BH03200		4/4
BH04260		5/1
BH06400		9/4
BH06830		2/0
BH07000		3/3
BH07230		7/0
BH07670		4/3
BH09290		3/2
BH11250	*ftsW*	5/4
BH13960		3/2
BH14030		3/0
BH14680	*bfsE*	3/3
BH15720	*trwH2*	6/1
BH16130		3/0
BH16240		4/4
BH16470		1/2

protein	gene name	spectra MSGF+/ Mascot
BH00150		5/1
BH00160		17/7
BH00810	*nth*	2/2
BH01730		3/3
BH02980;BH03420		3/2
BH03410		2/1
BH03640		2/2
BH03760		2/1
BH04500		4/4
BH04820		4/2
BH06170	*clpS*	4/1
BH07740		4/2
BH07890		5/2
BH11060		14/15
BH11570		4/4
BH12270		2/0
BH13250		77/46
BH13270	*virB3*	4/3
BH13290	*virB5*	52/30
BH13320	*virB8*	115/73
BH13390		12/8
BH15730	*trwG*	13/11

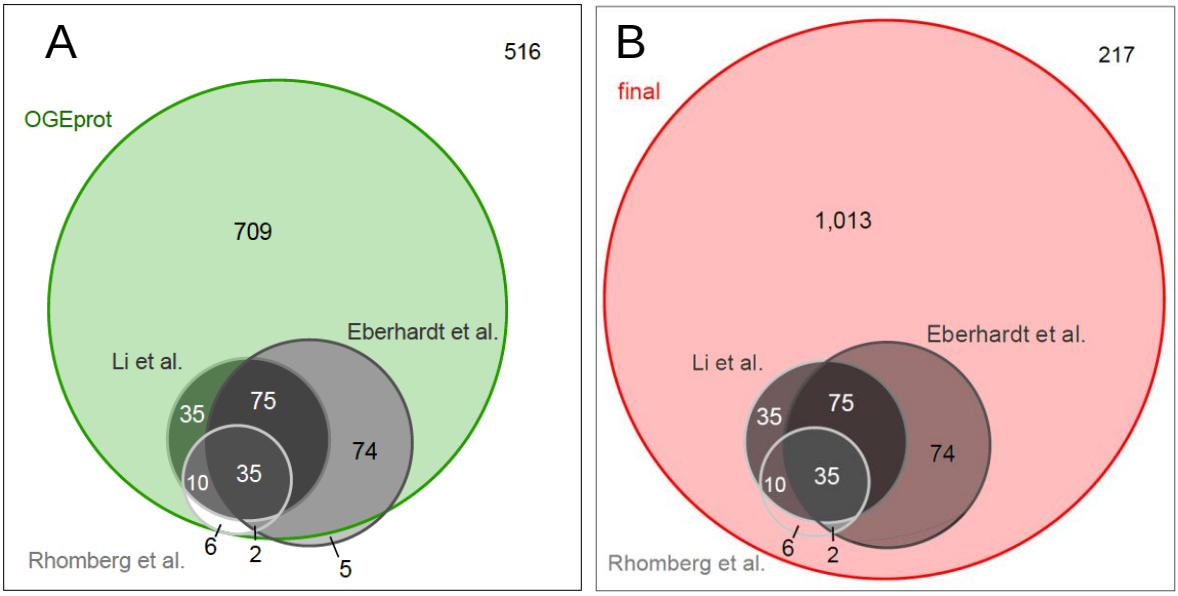
Figure S5: Success of different experimental and computational strategies



Each of the experimental and computational approaches that were used to maximize proteome coverage contributed unique protein identifications. As we did not use a full factorial experimental design to properly identify the contribution of each of the various techniques we can estimate the effect only for some individual techniques: Re-measuring a sample using the exclusion list approach (Kristensen et al. 2004) (i.e. precursor ion masses selected for fragmentation in a first run are excluded from fragmentation in a subsequent second run, which helps to identify new, low abundant proteins), added on average 16.6% additional protein identifications per sample. Compared to the pilot experiments (OGEprot), the gelfiltration approach added 7.3% protein identifications, the ProteoMiner approach added 3.3% protein identifications and the OGEpep approach added 28.1% protein identifications. Further fractionation of part of the TM sample into IM and OM allowed us to add 6.7% protein identifications compared to TM alone.

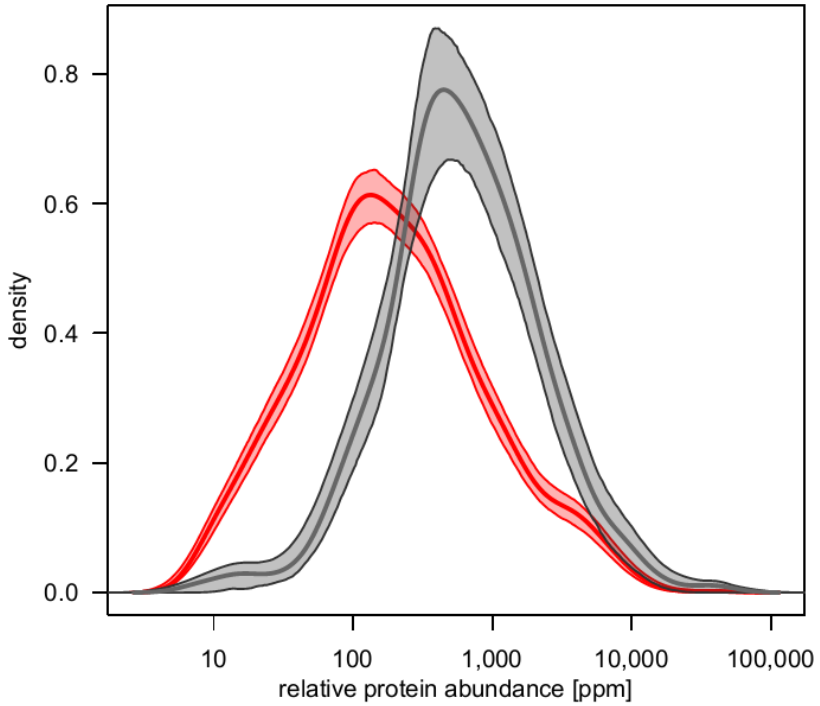
Figure S6: Overlap with previous proteomics studies.

A. Venn diagram showing the overlap of proteins identified in the pilot phase by OGEprot (946 proteins, green circle) with those described in three previous *B. henselae* proteomics studies. These include a study by Eberhardt et al., who had analyzed total cell extracts to identify proteins of potential sera-diagnostic value (Eberhardt et al. 2009) and identified 191 distinct proteins (two of the 192 accessions reported are 3a proteins, i.e. two distinct gene models encode an identical protein sequence), and two studies which had focused on the description of outer membrane proteins: the study by Rhomberg et al., which had identified 53 proteins (Rhomberg et al. 2004), and by Li et al., which had identified 155 proteins (Li et al. 2011) (we only considered proteins with a proper *B. henselae* identifier). The boxed area represents all proteins from *B. henselae* (NCBI's RefSeq) with a distinct protein sequence (1,467 protein groups); 516 proteins were not identified by these four studies.



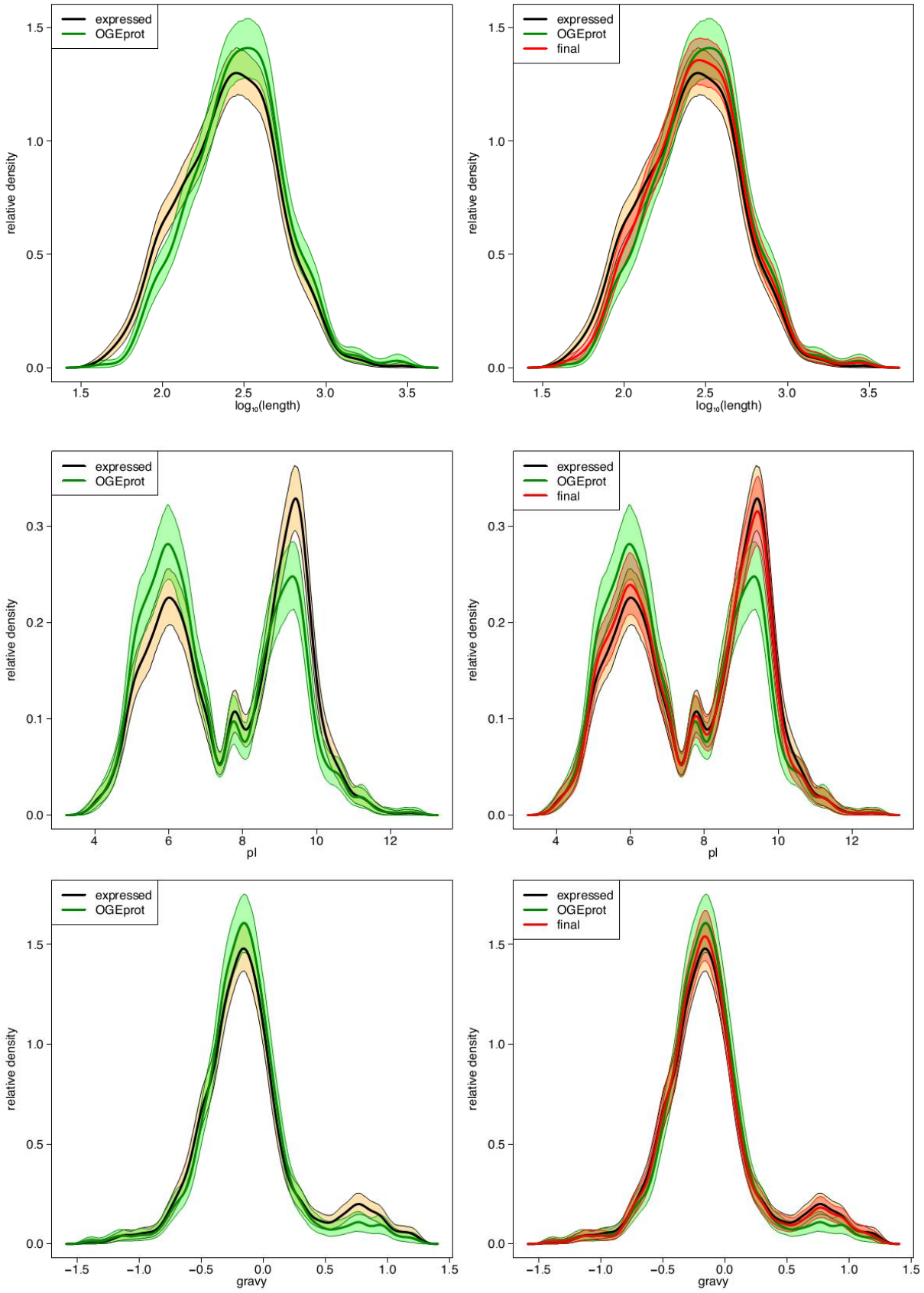
B. Venn diagram showing the overlap of the three previous studies (grey) with our final proteome dataset (red circle, 1,250 proteins). Again, the number of distinct proteins that were not identified is shown in the upper right hand corner.

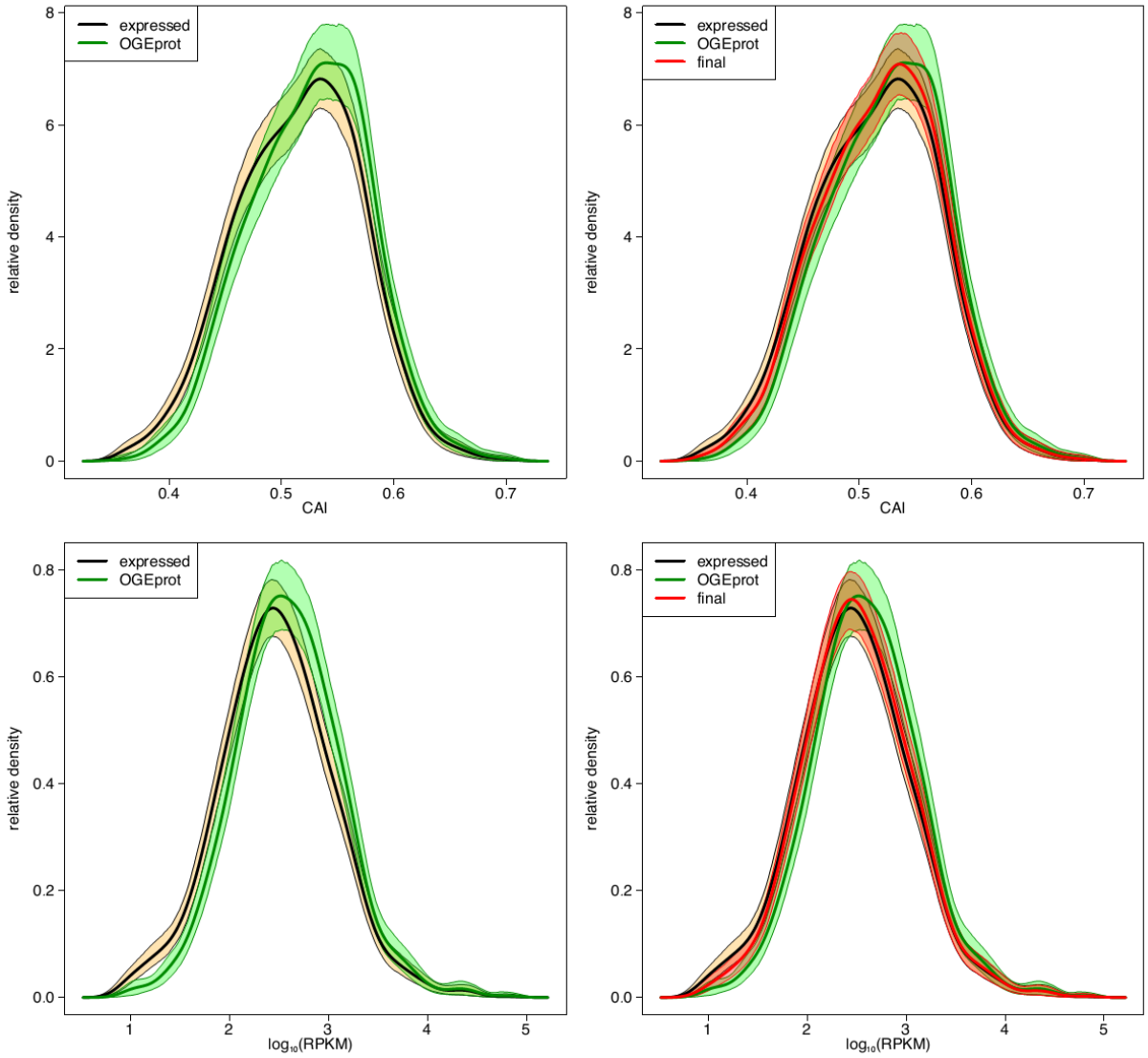
C. Comparison of protein abundance metrics. Kernel density estimates (wide center line) together with 95%-bootstrap-confidence bands (CBs, shaded area) are shown for normalized spectral count data from our final dataset (red) and the proteins identified in the previous



three studies (grey). A profound difference with respect to identification can be seen, showing a clear preference for higher abundant proteins in the previous studies.

Figure S7: Comparison of physicochemical protein parameter and RPKM value densities for selected datasets.





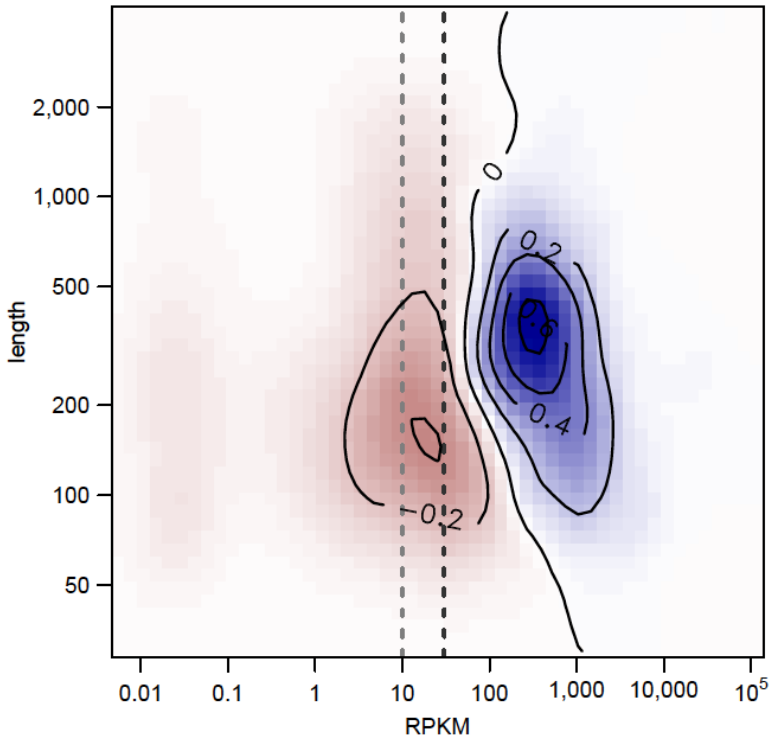
We compared the estimated densities of i) all proteins whose encoding gene models were expressed based on RNA-seq data from matched samples (n=1,353; expressed, black/orange) with ii) proteins identified by OGEprot in eight pilot experiments (n=946; OGEprot, green), and iii) the complete expressed proteome based on our ADE approach (n=1,250; final, red). The assessed protein parameters include isoelectric point, gravy (grand average hydropathicity), length, codon adaptation index (CAI) and the RPKM values for the respective protein-coding genes. The shaded areas illustrate the 95%-bootstrap-confidence bands (CBs) for the respective density curves. If a density curve overlaps with the CB of another dataset, then the parameter densities are statistically not distinct (on a 5% significance level). In the right-hand side panels it can be seen that the distribution of the physicochemical parameters in the complete expressed proteome (final) remedies all gaps of under- or overrepresentation which are distinguishable between the expressed protein-coding genes and the pilot experiment (left-hand side panels).

Figure S8: Proteome endpoint analysis based on gene expression levels and different physicochemical protein parameters.

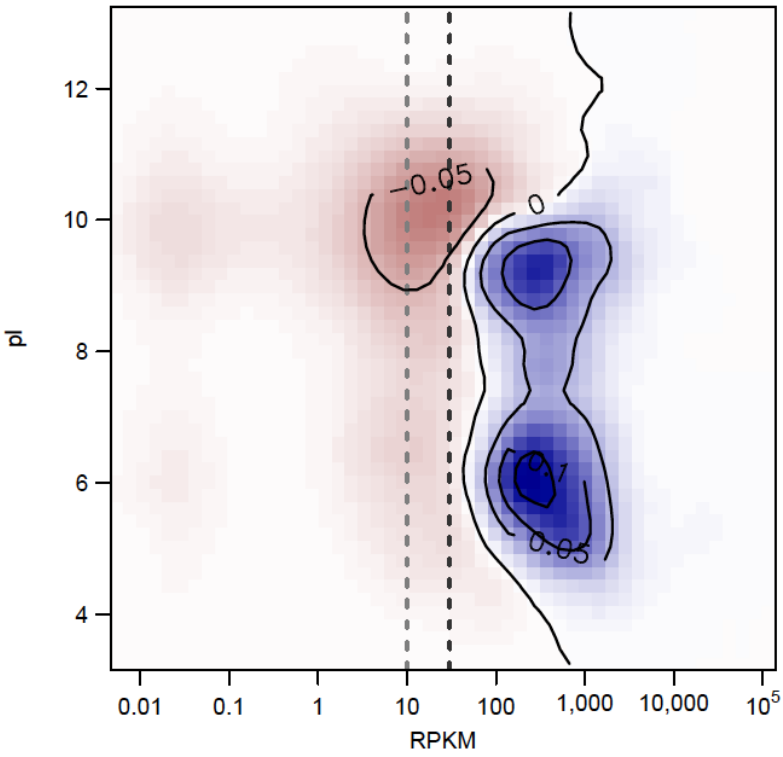
Several 2-dimensional density plots are shown below that compare the expression level of protein-coding genes (RPKM) and the physicochemical parameters protein length (A), isoelectric point (B) and gravy (C) for the datasets not seen (217; red) and expressed (1,250; blue). In addition, dashed lines indicate both the conservative, low RPKM cut-off of 10 (grey dashed line) and an RPKM value of 30 (black dashed line), which is based on the average RPKM values of the *virB/D4* operon in the uninduced condition, i.e. a more biologically motivated threshold.

The data indicate that there is a tendency for both for short (A.) and basic proteins (B.) to be particularly difficult to detect when their coding genes are expressed in the range of a potential lower gene expression level threshold. This is an inherent limitation of shotgun proteomics. In contrast, for proteins with higher positive gravy values (membrane proteins which contain one or more transmembrane domains) there is no detectable under-representation. Together with the coverage of predicted transmembrane and secreted proteins (see Figure S10A), and the data from the VirB/D4 T4SS coverage, we take this as evidence that we have identified a complete membrane proteome.

A. 2-dimensional density plot for protein length and gene expression level.



B. 2-dimensional density plot for protein isoelectric point and gene expression level.



C. 2-dimensional density plot for protein parameter gravy and gene expression level.

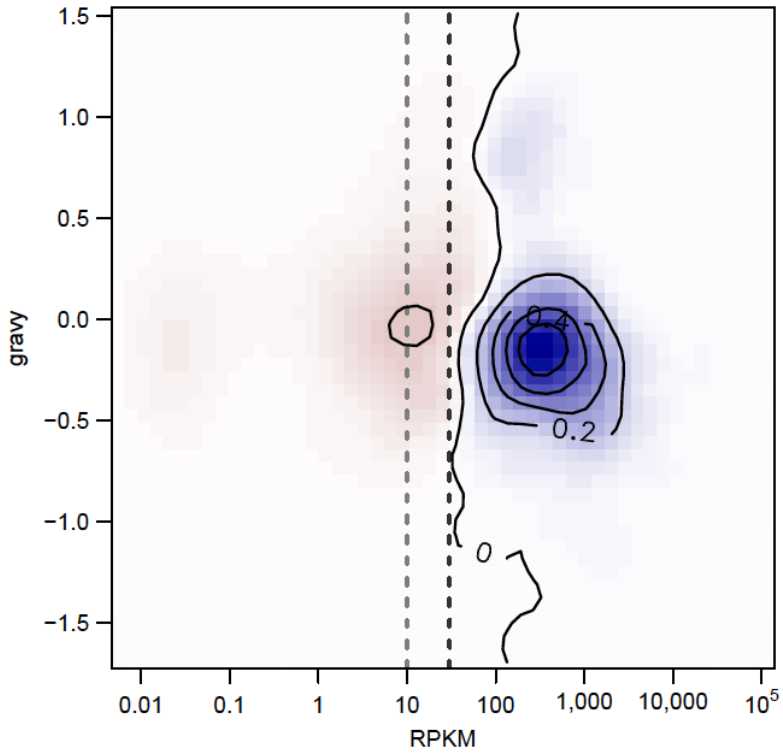
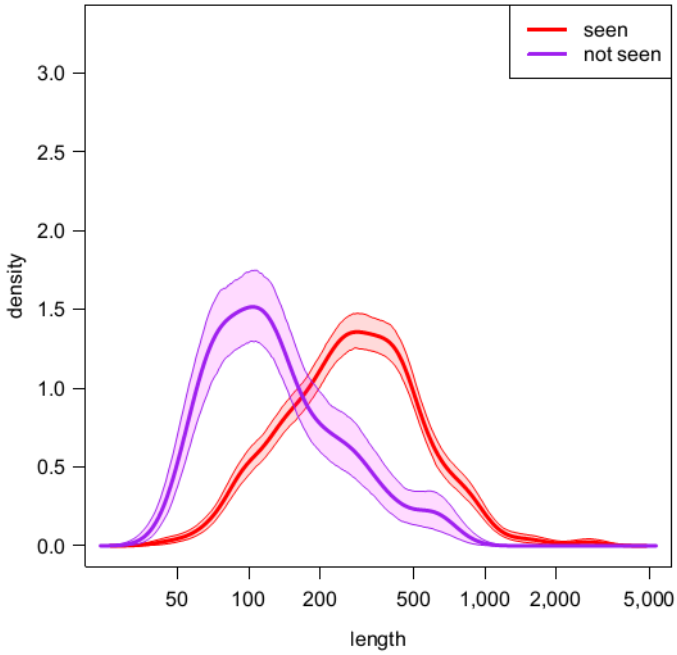
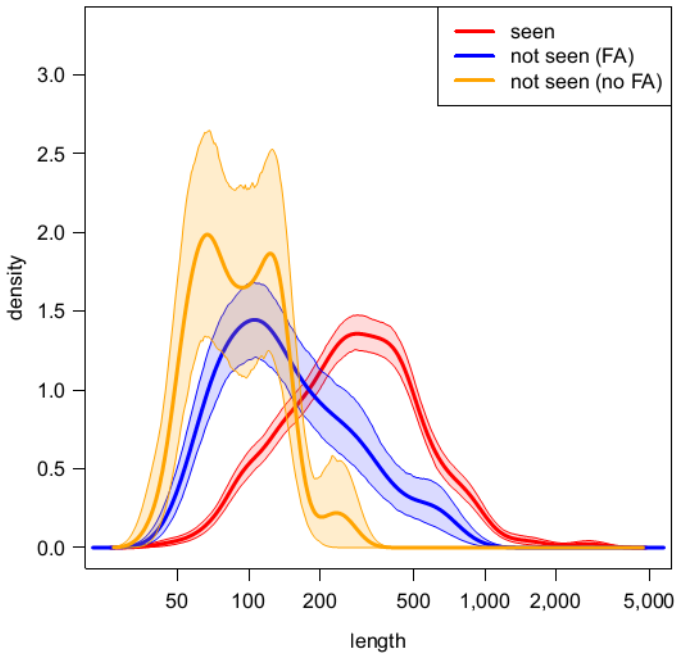


Figure S9: Not identified proteins are preferentially short and may include potential over-predicted ORFs.

A. The protein length densities are shown for the subsets “seen” (1,250 proteins, red) and “not seen” (217 proteins, violet). We illustrate the same 95%-bootstrap-confidence bands as in Figure S7.



B. Further distinction of the class not seen into proteins whose gene models have a functional annotation by EggNOG (not seen (FA), blue) and those lacking any functional annotation (not seen (no FA), orange), shows a clear separation of these two classes.



C. Importantly, we can identify short proteins in this length range. This is shown for the dataset seen in the plot below. We visualize the density of the shortest 150 proteins that were identified (seen (150 shortest), brown).

The findings of A-C taken together, suggest that protein-coding genes that lack any functional annotation may preferentially include genes that do not encode a functional protein and represent cases of over-prediction. This over-prediction of protein-coding genes has been noted previously both in prokaryotes and eukaryotes (Clamp et al. 2007; Skovgaard et al. 2001; Warren et al. 2010).

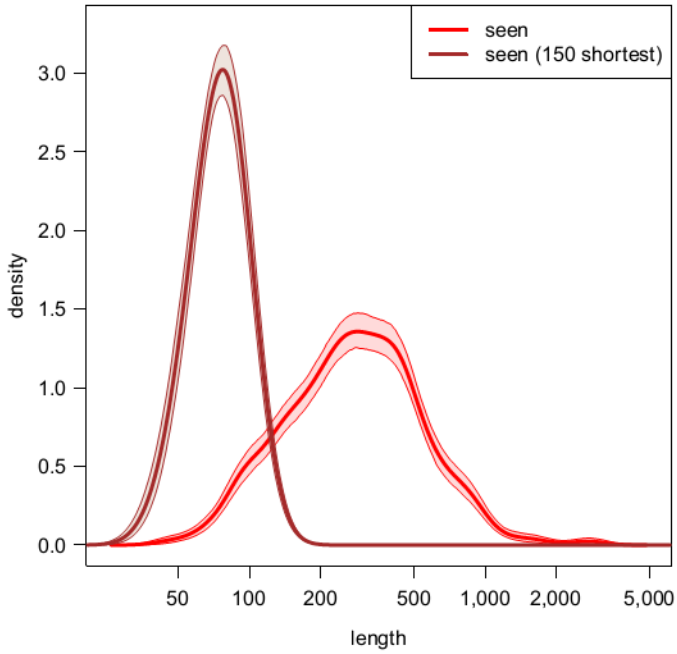
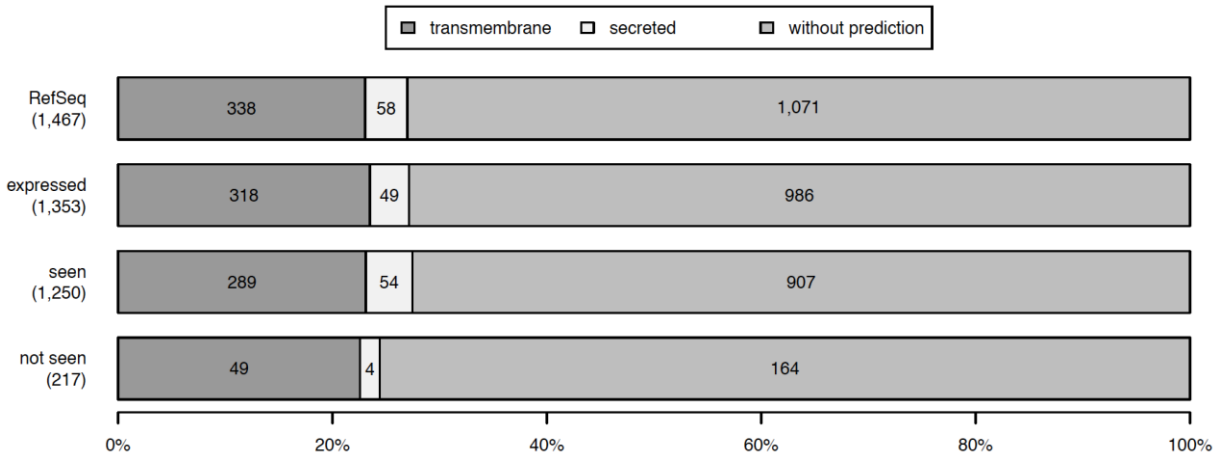


Figure S10: Membrane proteome coverage

A. Barplot of proteins in the datasets RefSeq (1,467), expressed (1,353), seen (1,250) and not seen (217), further classified as i) membrane proteins (containing one or more predicted transmembrane domains), ii) secreted proteins (proteins without further predicted transmembrane domain(s) after a predicted signal peptide cleavage site), or iii) proteins without a prediction by TMHMM (version 2.0) or SignalP (version 4.0). In the dataset seen, the percent-wise representation of transmembrane and secreted proteins is the highest.



B. Classification of 58 secreted proteins according to their predicted subcellular localization by PSORTb (version 3.0). 54 of 58 predicted secreted proteins were identified in the final dataset, several of which are predicted to localize to the membrane. Some examples are listed.

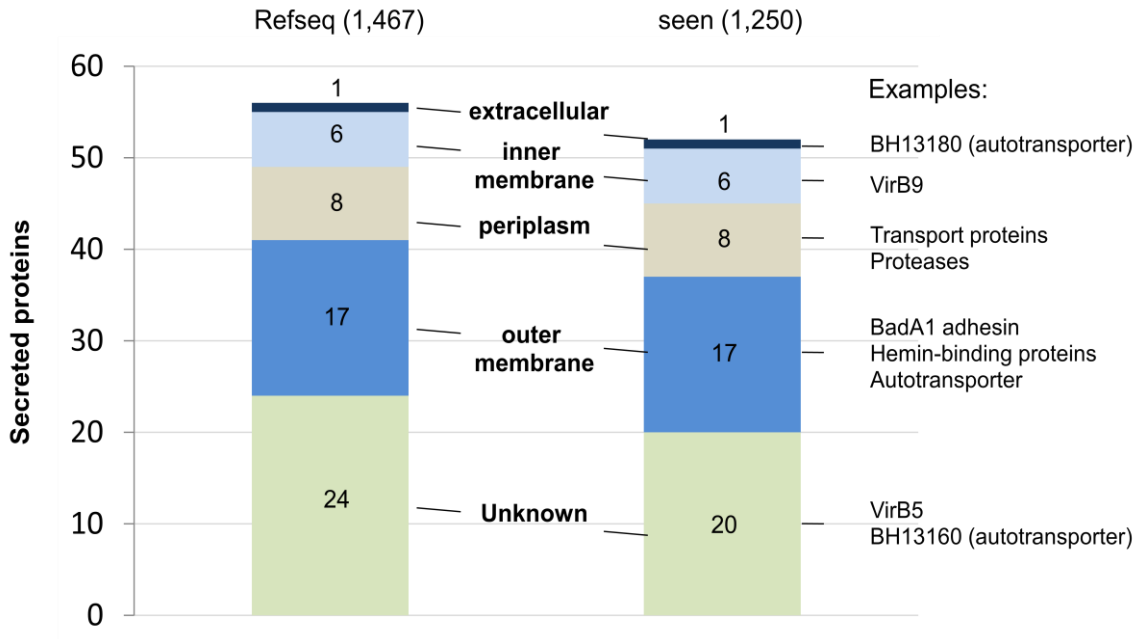
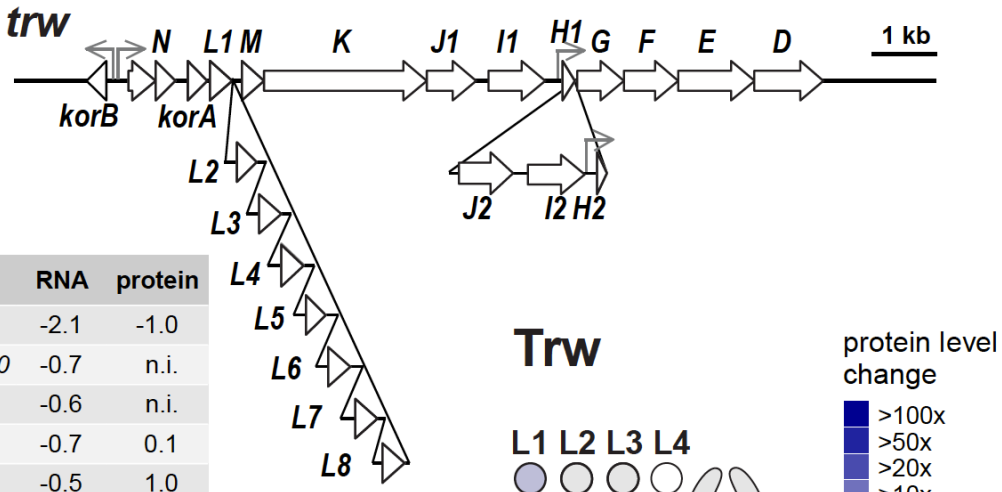


Figure S11: Transcript and protein expression changes of the *trw* operon.

The structure of the *trw* operon is shown on top. The list in the lower left panel shows the log2 fold changes at the transcript and protein level for the induced versus uninduced state. Fold changes and significance (Benjamini-Hochberg-adjusted p-values) were calculated in a DESeq analysis for both RNA-Seq reads and protein spectral counts ('n.i.' indicates that the protein was not identified, the '∞' indicates that the protein was only identified in the induced condition). The lower right panel visualizes the protein expression changes upon induction onto a schematic representation of the assembled Trw T4SS using a color scale. Under these two conditions none of the proteins is statistically significant differentially expressed.



gene	RNA	protein
korB	-2.1	-1.0
BH15540	-0.7	n.i.
trwN	-0.6	n.i.
korA	-0.7	0.1
trwL1	-0.5	1.0
trwL2	-0.6	-0.9
trwL3	-0.7	0.8
trwL4	0.4	n.i.
trwL5	-0.2	0.4
trwL6	-0.3	n.i.
trwL7	-0.2	-0.5
trwL8	0.6	n.i.
trwM	0.4	n.i.
trwK	1.0	1.1
trwJ1	0.2	n.i.
trwI1	1.6	n.i.
trwH1	0.4	0.2
trwJ2	0.4	0.2
trwI2	0.1	n.i.
trwH2	0.6	- ∞
trwG	1.4	∞
trwF	1.0	-1.3
trwE	1.7	-0.4
trwD	0.2	1.6

* p_{adj}<0.05

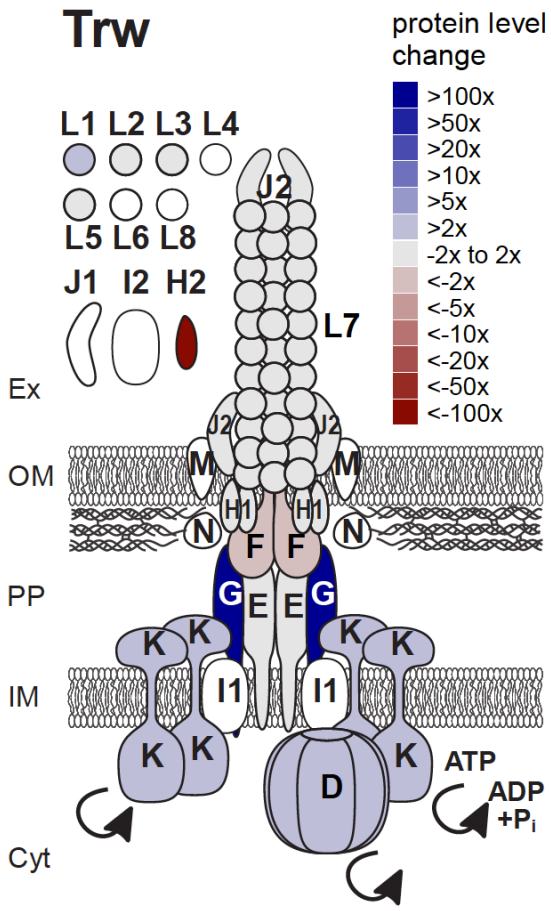


Table S5A. Information about the proteins identified in this study.

Table S5A is provided as separate Excel file. It contains information for the 1,488 annotated protein-coding genes along with numerous computational predictions, and the observed gene and protein expression values.

Table S5B. Information about the peptides identified in this study.

Table S5B is provided as separate Excel file. It provides more detail about the peptide evidence for each protein (all peptides, number of tryptic cleavage sites, number of PSMs further separated by subcellular localization) and allows user to select the best-suited PTPs for their protein set of interest.

Table S6. 125 differentially regulated proteins (top 10%) in the induced versus uninduced condition, ranked by DESeq.

Table S6 is provided as separate Excel file. It shows the *B. henselae* gene/protein identifier, the gene name (where available), overall spectral counts, spectral counts observed in the uninduced and the induced condition, followed by the normalized spectral count values calculated by DESeq for mean, uninduced, and induced condition, the p-value and a column with comments.

SUPPLEMENTAL METHODS

Construction of bacterial strains and plasmids

An in-frame Δ*batR-batS* 2CRS deletion mutant was generated from *B. henselae* strain RSE247 (a spontaneous streptomycin-resistant strain of *B. henselae* strain Houston-1 (ATCC 49822^T) (Schmid et al. 2004)) by a two-step gene replacement procedure, resulting in strain MQB242. For this, the mutagenesis plasmid pMQ009 (Table S1) was constructed as follows: Two homology regions (H1/H2) were amplified using oligonucleotide primers prAB007/ prMQ1088 (H1, 946 bp) and prMQ1091/ prMQ1092 (H2, 787 bp) (Table S2), and both fragments subsequently combined by megaprime PCR using prAB007 and prMQ1092. By using the flanking XbaI sites, the resulting 1714 bp fragment carrying an 2005 bp in-frame deletion within the coding sequence of BatR-BatS was inserted into the corresponding site of pTR1000, yielding pMQ009. Gene replacement in *Bh* RSE247 by use of pMQ009 resulted in the Δ*batR-batS* strain MQB242.

The chromosomal reporter consisting of the BatR inducible *bepD* promoter in front of *gfpmut2* (P_{bepD}:*gfp* (Quebatte et al. 2010)) was introduced by the same two-step procedure, allowing to monitor BatR mediated gene induction at the protein level in individual bacterial cells (strain MQB277). The mutagenesis plasmid pMQ012 (Table S1) was constructed as follows: a 736 bp fragment containing the P_{bepD}:*gfp* reporter was amplified from pIT011 using prMQ1191/prMQ1192 and ligated into the 9654 bp *Sac*II/*Not*I fragment of pPG612 after digestion with the same enzymes, yielding pMQ012. The use of pMQ012 for gene insertion in MQB242 resulted in the Δ*batR-batS* strain MQB277 (Table S1). Conjugation with plasmid p*batR* carrying the *batR* gene fused to an IPTG-inducible lacZ promoter (P_{taclac}) resulted in the final strain MQB307 in which *batR* gene expression can be induced by addition of IPTG.

RNA extraction and whole transcriptome sequencing

RNA was isolated from bacterial cells using a modified hot-phenol extraction, followed by DNase I digestion, RNA cleanup (RNeasy Mini Kit, Qiagen) and confirmation of RNA integrity as described (Quebatte et al. 2010). Whole Transcriptome libraries were produced using a protocol that preserves the origin of transcripts from either forward or reverse strand. Briefly, 5 µg of total RNA was depleted of rRNA and then fragmented using RNase III. Ligation of the adaptor mix and reverse transcription were performed following the manufacturer`s protocol. cDNA libraries were size selected for fragments between 150 and 250 bp, amplified for 18 cycles of PCR using barcoded adaptor primers and purified with the PureLink PCR micro kit

(Invitrogen). Library size and concentrations were assessed on a Bioanalyzer (Agilent) and on a Qubit fluorometer (Invitrogen). The whole transcriptome library was used for emulsion-PCR based on a concentration of 0.5 pM. Sequencing beads from four barcoded libraries (2 biological replicates each for uninduced and induced condition) were pooled and loaded on a full SOLiD™-4 slide (Applied Biosystems). SOLiD™ ToP Sequencing F3-Tag MM50 chemistry was used to produce 50 base sequencing reads. As suggested by one reviewer, we also assessed whether short genes were preferentially not detected due to the fragment size: Among the 1488 genes, 102 protein-coding genes are below 250 base pairs in size; of these 70 were detected at the transcript level, 32 not. Importantly, 30 of the 32 ORFs lack any ortholog in the Bartonella clade (Figure 5). Several of these could likely represent over-annotations and no bona-fide protein coding ORFs. When we looked for the 10 shortest proteins, (ranging in length from 41 to 51 amino acids), we find that 9/10 are expressed (with the one not expressed being a duplicated gene, i.e. we cannot sum up unambiguous reads). Two ribosomal protein genes are among these 10 proteins, both of which are strongly expressed (greater than 2400 and 6000 RPKM, i.e. present in expression bins 8 and 10 in Figure 4, and thus among the highest expressed genes). Together, there is no bias against short genes, but the approach would not be suited to identify short RNAs and potentially very short protein-coding ORFs.

RNA-seq data processing and transcriptome coverage analysis

The sequenced reads were mapped to the genome sequence of *B. henselae* Houston-1 strain (NCBI RefSeq acc. NC_005956.1, which was slightly modified to include the *bepD*:*gfp*$_{mut2}$ chromosomal reporter fusion, and retain the deleted *batR-batS* sequence) using the BioScope 1.3.1 mapping pipeline (mapreads software using local alignment strategy). SAM files were further processed to remove among all uniquely mapping reads those with more than two mismatches, a mean Phred read quality below 20, or a mapping quality below 10 (for more detail, see below and Figure S12).

We aimed to extract from the large number of multiple mapping reads all instances where based on metrics described below, one of the reads exhibited a better match with the reference genome sequence (Figure S12). For this, we proceeded as follows: For the class of doubly mapped reads, we blocked reads from duplicated 16S and 23S rRNA genes (accounting for 96-96%) and then filtered instances where, based on edit distance (ED) and mapping quality (MQ) as selection criteria, one of the two mappings exhibited a better match with the reference genome sequence (take read with lower ED; if ED identical, take read with higher MQ; if both identical, the read was not considered). This was repeated for triply

mapped reads (including the 5S rRNA genes) and up to 4-7 mapped genome positions per read. Since the MQB307 deletion strain does not contain the *batS* gene, we could assess the quality of our filtering step: we observed that after the quality control step not a single read remained erroneously assigned to the *batS* (BH00620) genome region.

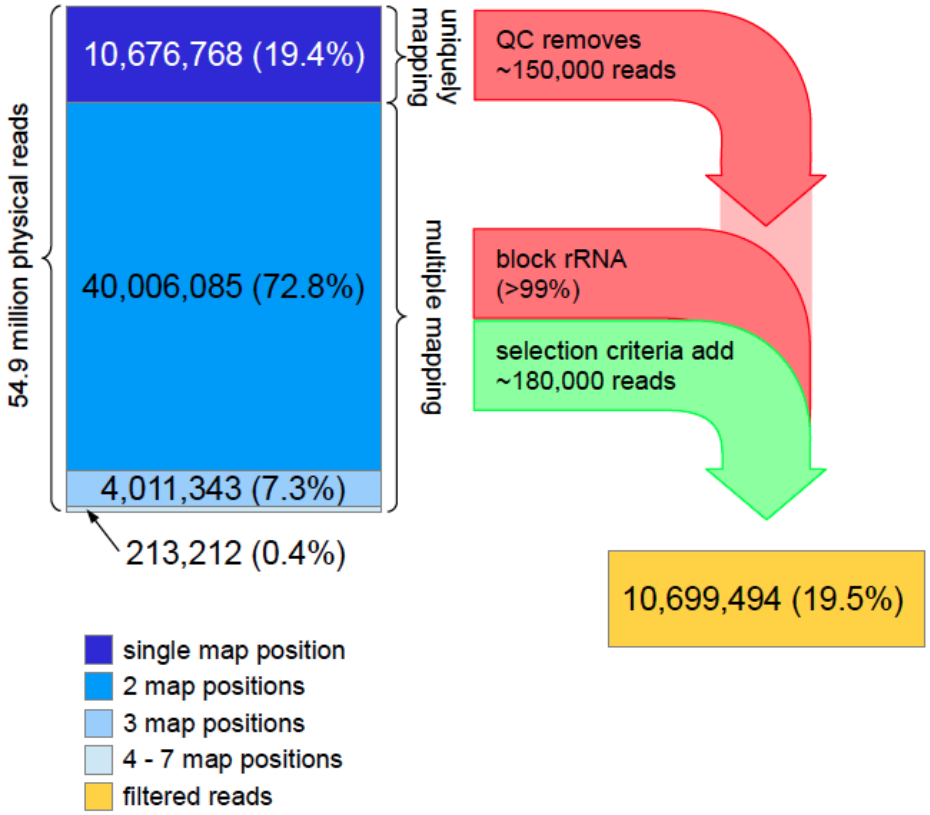


Figure S12: Steps for quality control of uniquely and multiple mapping reads exemplified for sample uninduced 2.

The count data summary for annotated *B. henselae* ORFs was generated using the HTSeq package. To create Figure 2A, the filtered reads were shuffled and sequentially mapped to the genome; a protein-coding ORF was classified as expressed when five or more distinct reads (having at least half of their length within the first 50 nucleotide region of the ORF) have accumulated in the 5' end of the ORF.

To estimate the number of expressed protein-coding genes that could be identified by doubling the number of RNA-seq reads, nonlinear regression models were constructed for

data from uninduced and induced experiments, respectively (Figure 2A). The model has the form:

$$G = \frac{a \cdot R^b}{c + R^b}$$

where *G* is the number of expressed protein-coding genes, *R* the number of filtered and mapped RNA-seq reads and *a, b, c* the shape parameters of the model. Robust M-estimators for the parameters were obtained using iterated reweighted least squares method (R package robustbase, version 0.9-4).

Subcellular fractionation

The subcellular fractionation by ultracentrifugation and subsequent differential lysis of total bacterial membranes with the ionic detergent lauroyl sarcosine was performed as previously described (Rhomberg et al. 2004), with the minor modification of using a French Press instead of sonication for the initial lysis. Bacteria were harvested from 20-40 CBA plates, washed in PBS, and pelleted twice for 15 min at 4,800 x g, 4°C. The pellet was resuspended in 4 mL hyperosmolar buffer (0.2 M Tris pH 8.0, 0.5 M sucrose, 250 µg/mL lysozyme, 1mM EDTA) and incubated on ice for 1 h. Subsequently, 0.5 ml protease inhibitor cocktail Complete (Roche) and Gentamycin to a final concentration of 200 mg/l was added to the sample prior to cell lysis by two passages in a French press. Cell debris was removed by centrifugation (30 min at 4,800 x g), and the supernatant (total cell lysate) containing membrane vesicles was subjected to ultracentrifugation (Centrikon T1075 with TFT 80.13 FA rotor) for 90 min at 40,000 rpm and 4°C. The supernatant, i.e. the cytoplasmic fraction (Cyt), was collected for analysis while the pellet, i.e. the total membrane fraction (TM), was resuspended in lysis buffer (10 mM HEPES pH 7.4, 1% w/v lauroyl sarcosine) and incubated at room temperature for another 20 min. One fourth of the TM fraction was kept for analysis, while the remaining sample was subjected to a second round of ultracentrifugation. The resulting supernatant (i.e. the inner membrane fraction (IM)) was collected for analysis and the lauroyl sarcosine-insoluble pellet, i.e. the outer membrane-peptidoglycan fraction (OM), was washed twice in 10 mM HEPES, pH 7.4, to remove residual detergent and pelleted again. Protein extracts from the cytoplasmic (Cyt), TM, IM, and OM fractions from bacteria grown under uninduced (e.g. Cyt_u) and induced (e.g. Cyt_i) condition were stored at -70°C until further analysis. Protein concentrations were determined with the Lowry method.

Computation of physicochemical parameters and other protein sequence features

Physicochemical protein parameters (length, theoretical isoelectric point (pI), grand average hydropathicity (gravy)) were retrieved from Expasy http://web.expasy.org/protparam/. Codon Adaptation Index (CAI) values were computed by the method of Carbone et al. (Carbone et al. 2003). Signal peptide predictions from SignalP (version 4.0), and transmembrane domain predictions from TMHMM (version 2.0; both from the CBS, Denmark), were used for a combined topology prediction: Proteins without a predicted transmembrane domain after a predicted signal peptide cleavage site are considered secreted. Proteins with one or more predicted transmembrane domains after a predicted signal peptide cleavage site or without a predicted signal peptide cleavage site are assumed to be transmembrane proteins.

Prediction of proteoypic peptides (PTPs)

Our data indicate that a comprehensive discovery proteomics approach adds clear value with respect to experimentally identified PTPs when compared to entirely relying on *in silico* prediction of PTPs using tools like PeptideSieve (Mallick et al. 2007). Among the 1,467 distinct proteins, PeptideSieve (version 0.51) predicts one or more PTPs above a probability cutoff of 0.8 for 1,263 proteins, including 1,105 proteins that we identified. However, our experimental dataset provides expression evidence for 145 of the proteins for which PeptideSieve predicted no PTP.

ADE analysis

The exponential model has the form:

$$P_j = b + a\left(1 - e^{-c \cdot S_j}\right)$$

for $j = 1, \ldots, k$, where k is the number of different biochemical fractionations, P_j the number of proteins detected, S_j the number of PSMs and a, b, c the shape parameters of the exponential curve. We estimated the parameters of the model using nonlinear least squares (Bates and Chambers 1992) and used the fit to predict the saturation beyond the point of experimentally observed PSMs for each biochemical fractionation regimen (Figure 3A, dashed lines). We approximated confidence bands for the fitted points based on linear models on the gradient of the exponential fit (Figure 3A). In case of the ProteoMiner setup only two experiments are available and thus a fit was not possible, nor could confidence bands be approximated. The grey, dashed line indicates the number of proteins additionally

detected in the second experiment (Cyt_u). For the density estimation of physicochemical protein parameters (Figure 3B), the following approach was used: density lines were estimated using a Gaussian kernel and the Sheather-Jones (SJ) plug-in bandwidth (Sheather and Jones 1991), which was computed from the 'expressed' dataset and also used for the OGEprot and the corresponding ADE protein subsets in each plot based on the smallest overlapping range of the physicochemical parameter. Values outside of this range were omitted from the density estimation. 95% bootstrap confidence bands were estimated using 1,000 bootstrap samples and density estimation based on the (constant) SJ plug-in bandwidth computed from the 'expressed' dataset.

Orthologs, and functional Protein classification

Orthologs are from (Engel et al. 2011) and were determined by using the "PhyloProfile Synteny" tool of MaGe (Vallenet et al. 2006) with a threshold of 60% protein identity over at least 80% of the length of proteins being directional best hits of each other. We relied on the group definitions from eggNOG (http://eggnog.embl.de), which contains non-supervised orthologous groups that were constructed from 1,133 organisms (including *B. henselae*). The depth and coverage of the functional protein annotations depends on the evolutionary level that is selected, e.g. the level proteobacteria (proNOG) has more annotations than the level bacteria (bactNOG). We considered COG, NOG, bactNOG, proNOG and aproNOG (for α-proteobacteria) classifications. Thereby, at least one functional annotation is assigned to 1,433 of the 1,488 *B. henselae* protein-coding genes (96%).

Database searching and data processing

We downloaded NCBI's *B. henselae* strain Houston-1 reference annotation (NC_005956.1; as of 22-Jan-2012) with its 1,488 protein-coding genes. Since the bacteria were grown on CBA plates, we added sequences of 3,336 sheep proteins (*Ovis aries*; downloaded from UniProtKB, February 2011) to minimize the chance for false positive assignments of spectra originating from sheep proteins to bacterial proteins, a positive control (myc-gfp), and sequences of 256 common contaminants (keratins, trypsin, etc.). Spectra were searched against the combined database.

Protein & peptide fractionation approaches

We used several biochemical fractionation approaches to further reduce sample complexity. Below, these are represented in an overview figure (Figure S13), and more detail for the different steps is provided in subsequent paragraphs. For the ProteoMiner approach, we only had enough protein sample from the cytoplasmic fractions ($Cyt_{u/i}$). The gelfiltration (size exclusion chromatography) approach was carried out for all subcellular fractions except the OM samples, from which the least amount of protein was available. The OM fraction was dissolved in 100 µl of 50 mM ammonium bicarbonate (ABC) pH 8; 1% sodium dodecylsulfate (SDS). Complete dissolution was achieved after 10 min incubation in an ultra-sonication bath. 50% of the OM fraction was precipitated with trichloroacetic acid (TCA) and used as described in the section OFFGEL protein fractionation (see below). The remaining protein solution was adjusted to 0.1% SDS with 50 mM ABC, pH 8 and used as described in the section in-solution digest of membrane fractions** but without TCA precipitation.

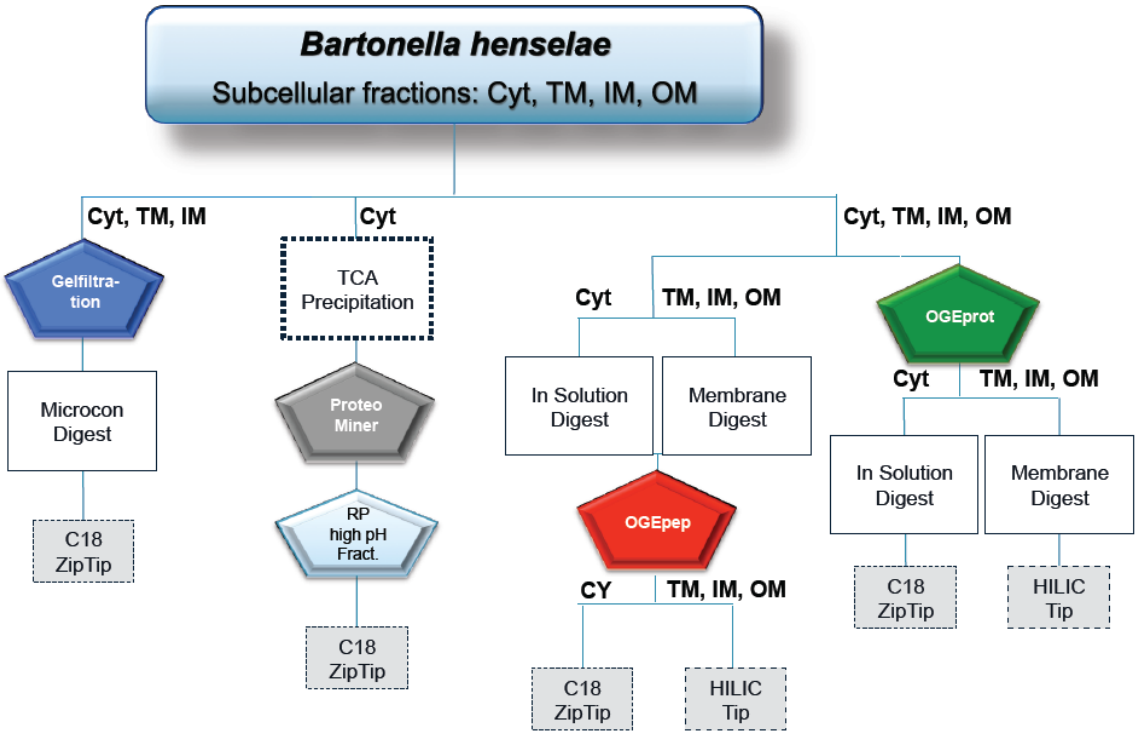


Figure S13: Overview over the different experimental steps for a respective workflow.

Individual steps are described in more detail in the sections below. OGEprot was used as biochemical fractionation method for the pilot phase.

OFFGEL Protein Fractionation

Isoelectric focusing was performed with the OFFGEL Fractionator 3100 using the low resolution Kit pH 3–pH 10 (Agilent). OFFGEL-electrophoresis (OGE) was performed according to the manufacturer's protocol. Briefly, two low-resolution IPG strips were used for 1 mg of protein extract. 500 µg of protein containing solution was dissolved in focusing buffer to a final volume of 1.9 mL (12% glycerol, 1.2% ampholytes, 6.7 M urea, 1.9 M thiourea, 62.4mM dithiothreitol (DTT)). 150 µL of this sample was loaded in each of the 12 wells. The temperature of the tray holder was set to 20°C and focusing was started with a maximum current of 50 µA per lane. The focusing stopped automatically after total voltage reaches 20 kVh. After focusing, all the fractions were recovered for each well and the corresponding fractions of the two lanes were pooled. Proteins of each fraction were precipitated by addition of trichloroacetic acid (TCA) to a final concentration of 12%. The precipitate was washed twice with pre-chilled acetone. The cytosolic samples were treated as described in the section in-solution digest of cytosolic fractions*. Membrane fractions were treated as described in the section in-solution digest of membrane fractions**. After digestion, the cytosolic samples were desalted using C_{18} ZipTips™ (Millipore) following the vendor's protocol. All membrane samples were concentrated using a vacuum concentrator and desalted with HyperSep™ spin tips (Thermo). The desalted peptide sample was further analyzed by nano-HPLC mass spectrometry.

* In-solution digest of cytosolic fractions

TCA precipitated protein was dissolved in 100 µL of 50 mM ammonium bicarbonate (ABC) pH 8 to complete dissolution. Protein reduction was carried out by adding dithiothreitol (DTT) to 5 mM final concentration and incubated for 30 min at 60°C. Reaction mixture was cooled on ice and alkylation was carried out for 45 min at room temperature and darkness by addition of iodoacetamide (IAA) at 15 mM final concentration. The access of IAA was quenched by addition of 15 mM DTT. Tryptic digest was performed in 60% methanol at 20:1 (w/w) protein to protease ratio for 6 h at 37°C. The enzymatic reaction was stopped by addition of trifluoroacetic acid to a final concentration of 0.5%. The sample was vacuum concentrated and dissolved in 0.1% formic acid (FA), 3% acetonitrile (ACN). Finally, the peptide containing solution was desalted using C_{18} ZipTips™ (Millipore) following the vendor's protocol. The desalted peptide sample was further analyzed by nano-HPLC mass spectrometry.

** In-solution digest of membrane fractions

TCA precipitated protein was dissolved in 100 µL of 50 mM ammonium bicarbonate (ABC) pH 8. Protein reduction was carried out by adding dithiothreitol (DTT) to 5 mM final concentration and incubated for 30 min at 60°C. Reaction mixture was cooled on ice and alkylation was carried out for 45 min at room temperature and darkness by addition of iodoacetamide (IAA) at 15 mM final concentration. The access of IAA was quenched by addition of 15 mM DTT. The first enzymatic digest was performed in 60% methanol at 20:1 (w/w) protein to protease ratio for 4 h at 37°C. The mixture was centrifuged for 10 min at 13000xg and the supernatant removed, acidified by addition of TFA to final concentration of 0.5% and stored at -20°C. The pellet was resuspended in 60% methanol, 50 mM ABC pH 8. After addition of the same amount of Chymotrypsin (Roche) and Trypsin (Promega) in a 1 to 100 (w/w) ratio the samples were incubated over night at 37°C. The enzymatic reaction was stopped by addition of trifluoroacetic acid to a final concentration of 0.5%. Both digests were pooled and vacuum concentrated. Dried samples were dissolved in 0.1% formic acid (FA), 3% acetonitrile (ACN) and desalted using C_{18} ZipTips™ (Millipore) following the vendor's protocol. The desalted peptide sample was further analyzed by nano-HPLC mass spectrometry.

OFFGEL Peptide Fractionation

Isoelectric focusing was performed with the OFFGEL Fractionator 3100 using the low resolution Kit pH 3–pH 10 (Agilent). OFFGEL-electrophoresis (OGE) was performed according to the manufacturer‘s protocol. Briefly, two low-resolution IPG strips were used for 1 mg of protein extract. 500 µg of peptide containing solution was dissolved in focusing buffer to a final volume of 1.9 mL (12% glycerol, 1.2% ampholytes). 150 µL of this sample was loaded in each of the 12 wells. The temperature of the tray holder was set to 20°C and focusing was started with a maximum current of 50 µA per lane. The focusing stopped automatically after total voltage reaches 20 kVh. After focusing, all the fractions were recovered for each well and the corresponding fractions of the two lanes were pooled. Cyt_u and Cyt_i samples were acidified by addition of trifluoroacetic acid (TFA) to a final concentration of 0.5% and desalted using C_{18} ZipTips™ (Millipore) following the vendor's protocol. All membrane samples were concentrated using a vacuum concentrator and desalted with HyperSep™ spin tips (Thermo). The desalted peptide sample was further analyzed by nano-HPLC mass spectrometry.

Size exclusion chromatography (SEC) / gelfiltration experiments

TCA precipitated protein samples were dissolved in 4 M guanidine hydrochloride (GdmHCl), 50 mM Tris/HCl pH 8.2. Protein extracts from Cyt_u and Cyt_i respectively were injected on a size exclusion column Superose™ 12 (10/30, GE Healthcare). The column was equilibrated with 20 column volumes of 4 M GdmHCl, 50 mM Tris/HCl pH 8.2. 50 µL protein extract was injected to the column with at a flow rate of 0.4 mL/min using an Agilent 1100 HPLC-system (Agilent). Forty 0.8 mL fractions were collected and the corresponding fractions of the two runs were pooled. These fractions were concentrated on Microcons® (Millipore YM-3kDA cut-off) and an on-membrane digest was performed subsequently.

On-membrane digest of SEC/gelfiltration fractions

Protein containing SEC/gelfiltration fractions were reduced by addition of dithiothreitol (DTT) to 5 mM final concentration and incubated for 30 min at 60°C. The reaction mixture was cooled on ice and alkylation was carried out for 45 min at room temperature and darkness by addition of iodoacetamide (IAA) at 15 mM final concentration. The access of IAA was quenched by addition of 15 mM DTT. These samples were loaded on Microcon® (Millipore, YM, 3 kDa cut-off) and washed twice with 200 µL 25 mM ammonium bicarbonate (ABC) pH 8, 5% acetonitrile (ACN) by repeating cycles of concentrating up to 10 µL. Tryptic digest was performed over night in 25 mM ABC pH 8, 5% ACN, 0.1% *Rapi*Gest™ (Waters) at 50:1 (w/w) protein to protease ratio at 37°C. The enzymatic reaction was stopped by addition of trifluoroacetic acid (TFA) to a final concentration of 0.1%. Peptide containing samples were collected by centrifugation. After vacuum concentration the samples were dissolved in 0.1% formic acid (FA), 3% ACN and desalted using C_{18} ZipTips™ (Millipore) following vendor's protocol. The desalted peptide sample was further analyzed by nano-HPLC mass spectrometry.

Desalting on HyperSep™ spin tips

Desalting with HyperSep™ spin tips (Thermo Scientific) was preformed according to manufacturer's instructions. Briefly, peptide-containing samples were dissolved in binding buffer containing 15 mM ammonium acetate (AmAc) pH 3.5, 85% acetonitrile (ACN). The spin tips were first conditioned with 3 x 50 µL elution buffer containing 15 mM AmAc pH 3.5, 3% ACN and afterwards equilibrated with 3 x 50 µL binding buffer. Sample was applied and

the spin tips washed twice with 50 µL binding buffer. Peptides were eluted with 40µL elution buffer. The desalted peptide sample was further analyzed by nano-HPLC mass spectrometry.

ProteoMiner™ protein enrichment (BioRad)

The ProteoMiner™-Kit (BioRad) was used in an effort to identify low abundance proteins. Due to the amount of protein required, we could only apply it to the cytosolic fractions ($Cyt_{u/i}$). 6 mg of TCA precipitated cytosolic protein samples were used to perform the enrichment following the vendor's instructions. Briefly, spin columns were conditioned using 3 x 200 µL wash buffer (150 mM NaCl, 10 mM sodium di-hydrogen phosphate, pH 7.4). 1 mL of cytosolic protein sample was incubated for 2 h at RT with gentle agitation. The column resin was washed 3 times for 5 min with 200 µL wash buffer. After complete removal of the wash buffer the columns were washed for 1 min in deionized water. The water was removed and 20 µL of 5% acetic acid were added to elute enriched proteins. This step was repeated 5 times. Eluted samples were combined, vacuum concentrated and in-solution digest was performed as described above.

Reverse Phase HPLC at high pH

The samples processed with the ProteoMiner™-Kit were further fractionated by RP-HPLC at high pH. Samples were dissolved in solvent A (25 mM potassium phosphate buffer pH 8.5, 5% acetonitrile (ACN)) and loaded on a RP column (YMC, Pack Pro C18 RS, 2.1 mm x 150 mm, 3 µm) using an Agilent HPLC 1100 system. After 10 min isocratic elution at a flow rate of 0.2 mL/min with 100% solvent A peptides were eluted using the following gradient of solvent B (25 mM potassium phosphate buffer pH 8.5, 85% ACN). In 5 min 0% to 5% solvent B, in 35 min 5% to 35% solvent B and in 10 min 35% to 100% solvent B. 12 fractions were collected. All fractions were vacuum concentrated and dissolved in 0.1% formic acid (FA), 3% acetonitrile (ACN). Finally, the peptide containing solution was desalted using C_{18} ZipTips™ (Millipore) and further analyzed by nano-HPLC mass spectrometry.

Nano-LC MS/MS analysis (common for all approaches)

Dissolved samples were injected into an Eksigent-nano-HPLC system (Eksigent Technologies) by an autosampler and separated on a self-made reverse-phase tip column (75µm × 80mm) packed with C_{18} material (3µm, 200Å, AQ, Bischoff GmbH). The column was

equilibrated with 97% solvent A (A: 1% acetonitrile; 0,2% formic acid in water) and 3% solvent B (B: 80% acetonitrile, 0,2% formic acid in water). Peptides were eluted using the following gradient: 0-47 min, 3-35% B; and 47-60 min, 35-97% B at a flow rate of 0.2 μL/min. High accuracy mass spectra were acquired at an LTQ-ICR-FT-Ultra or an LTQ-Orbitrap (Thermo Scientific) in the mass range of 300-2,000 m/z. Up to five data dependent MS/MS were recorded in parallel in the linear ion trap of the most intense ions with charge state 2+ or 3+ using collision induced dissociation. Target ions already selected for MS/MS were dynamically excluded for 120s. After data collection the peak lists were generated using Mascot Distiller software 2.3.2 (Matrix Science Ltd.).

Quantitative RT-PCR

Validation of the RNA-Seq data (concerning the induction of *batR* and several *virB*/*D4* and *bep* operon members) by qRT-PCR was performed on total RNA samples as previously described (Quebatte et al. 2010). In brief, 1 microgram of total RNA was reverse transcribed using random primers (Promega) and Superscript II reverse transcriptase (Invitrogen). SYBR green I quantitative RT-PCR was performed on a StepOnePlus instrument (Applied Biosystems) using Power SYBR Green Master Mix (Applied Biosystems) and data were normalized to *rpsL* transcript expression levels (BH10560).

Database searches

Mascot

For Mascot, carbamidomethylation was set as a fixed modification on all Cysteines and oxidation of Methionines as well as cyclization of N-terminal Glutamines was considered as optional modification. In OFFGEL fractionated samples Aspartic and Glutamic acids were considered optionally methylated. Precursor ion mass tolerance was set to 5 ppm, fragment ion mass tolerance was set to 0.8 Da, and the automatic decoy search option was enabled. Spectra were searched for a match to fully-tryptic and semi-tryptic peptides with up to two missed cleavage sites. The built-in version of Mascot Percolator was used to improve the peptide spectrum match (PSM) scores based on the target-decoy approach. A Percolator score cutoff was determined to result in a 0.01% FDR at PSM level per experiment.

MS-GF+

We used MS-GF+ as second database search engine (MS-GFDB v7747 as of 05/15/2012). The fragment spectra were extracted from Thermo RAW files using ProteoWizard's msconvert (version 3.0.3831) and combined in one MGF file per experiment, which served as input for MS-GF+. Carbamidomethylation was set as a fixed modification on all Cysteines and oxidation of Methionines as well as cyclization of N-terminal Glutamines was considered as optional modification. In OFFGEL fractionated samples Aspartic and Glutamic acids were considered optionally methylated. Precursor ion mass tolerance was set to 5 ppm and the decoy search option was enabled, which enabled us to determine a probability score cutoff to result in a 0.01% FDR at the PSM level per experiment.

REFERENCES

Bates, D.M. and Chambers, J.M. 1992. Nonlinear models. In *Statistical Models in S* (eds. J.M. Chambers and T.J. Hastie). Wadsworth & Brooks/Cole.

Carbone, A., Zinovyev, A., and Kepes, F. 2003. Codon adaptation index as a measure of dominating codon bias. *Bioinformatics* 19: 2005-2015.

Clamp, M., Fry, B., Kamal, M., Xie, X., Cuff, J., Lin, M.F., Kellis, M., Lindblad-Toh, K., and Lander, E.S. 2007. Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci U S A* 104: 19428-19433.

Eberhardt, C., Engelmann, S., Kusch, H., Albrecht, D., Hecker, M., Autenrieth, I.B., and Kempf, V.A. 2009. Proteomic analysis of the bacterial pathogen *Bartonella henselae* and identification of immunogenic proteins for serodiagnosis. *Proteomics* 9: 1967-1981.

Engel, P., Salzburger, W., Liesch, M., Chang, C.C., Maruyama, S., Lanz, C., Calteau, A., Lajus, A., Medigue, C., Schuster, S.C. et al. 2011. Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen *Bartonella*. *PLoS Genet* 7: e1001296.

Kennell, D. 2002. Processing endoribonucleases and mRNA degradation in bacteria. *J Bacteriol* 184: 4645-4657; discussion 4665.

Kristensen, D.B., Brond, J.C., Nielsen, P.A., Andersen, J.R., Sorensen, O.T., Jorgensen, V., Budin, K., Matthiesen, J., Veno, P., Jespersen, H.M. et al. 2004. Experimental Peptide Identification Repository (EPIR): an integrated peptide-centric platform for validation and mining of tandem mass spectrometry data. *Mol Cell Proteomics* 3: 1023-1038.

Li, D.M., Liu, Q.Y., Zhao, F., Hu, Y., Xiao, D., Gu, Y.X., Song, X.P., and Zhang, J.Z. 2011. Proteomic and bioinformatic analysis of outer membrane proteins of the protobacterium *Bartonella henselae* (Bartonellaceae). *Genet Mol Res* 10: 1789-1818.

Mallick, P., Schirle, M., Chen, S.S., Flory, M.R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T. et al. 2007. Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol* 25: 125-131.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5: 621-628.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320: 1344-1349.

Quebatte, M., Dehio, M., Tropel, D., Basler, A., Toller, I., Raddatz, G., Engel, P., Huser, S., Schein, H., Lindroos, H.L. et al. 2010. The BatR/BatS two-component regulatory system controls the adaptive response of *Bartonella henselae* during human endothelial cell infection. *J Bacteriol* 192: 3352-3367.

Rhomberg, T.A., Karlberg, O., Mini, T., Zimny-Arndt, U., Wickenberg, U., Rottgen, M., Jungblut, P.R., Jeno, P., Andersson, S.G., and Dehio, C. 2004. Proteomic analysis of the sarcosine-insoluble outer membrane fraction of the bacterial pathogen *Bartonella henselae*. *Proteomics* 4: 3021-3033.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., and Pachter, L. 2011. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 12: R22.

Scheidegger, F., Ellner, Y., Guye, P., Rhomberg, T.A., Weber, H., Augustin, H.G., and Dehio, C. 2009. Distinct activities of *Bartonella henselae* type IV secretion effector proteins modulate capillary-like sprout formation. *Cell Microbiol* 11: 1088-1101.

Schmid, M.C., Schulein, R., Dehio, M., Denecker, G., Carena, I., and Dehio, C. 2004. The VirB type IV secretion system of *Bartonella henselae* mediates invasion, proinflammatory activation and antiapoptotic protection of endothelial cells. *Mol Microbiol* 52: 81-92.

Schulein, R., Guye, P., Rhomberg, T.A., Schmid, M.C., Schroder, G., Vergunst, A.C., Carena, I., and Dehio, C. 2005. A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. *Proc Natl Acad Sci U S A* 102: 856-861.

Sheather, S.J. and Jones, M.C. 1991. A reliable data-based bandwidth selection method for kernel density estimation. . *Journal of the Royal Statistical Society. Series B* 53: 683-690.

Skovgaard, M., Jensen, L.J., Brunak, S., Ussery, D., and Krogh, A. 2001. On the total number of genes and their length distribution in complete microbial genomes. *Trends Genet* 17: 425-428.

Tariq, M.A., Kim, H.J., Jejelowo, O., and Pourmand, N. 2011. Whole-transcriptome RNAseq analysis from minute amount of total RNA. *Nucleic Acids Res* 39: e120.

Vallenet, D., Labarre, L., Rouy, Z., Barbe, V., Bocs, S., Cruveiller, S., Lajus, A., Pascal, G., Scarpelli, C., and Medigue, C. 2006. MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res* 34: 53-65.

Warren, A.S., Archuleta, J., Feng, W.C., and Setubal, J.C. 2010. Missing genes in the annotation of prokaryotic genomes. *BMC Bioinformatics* 11: 131.