SUPPLEMENTAL METHODS

In-solution digestion

Cytosolic and membrane protein extracts (15 µg) present in 8M urea in 10 mM Tris-HCl pH 8.0 were subjected to disulfide reduction by adding 1 µL 10 mM dithiothreitol (DTT) and incubating for 30 min at RT prior to alkylation by adding 1 µL of 50 mM of iodoacetamide (IAA) in 50 mM Ammonium bicarbonate (ABC) and incubating for 20 min at RT. After digestion with Endoproteinase LysC (1:50 w/w) (Wako Chemicals, USA) for 3 h at RT, the solution was diluted to 2M urea with ABC and incubated overnight in the presence of sequencing grade modified trypsin (1:50 w/w) (Promega, Madison, WI) at 37^{o}C for complete digestion. Subsequently, the peptide mixture was concentrated and desalted using stop and go elution (STAGE) tips as described by Rappsilber and coworkers.[1] Briefly, the acidified peptides were bound on activated C18 reversed phase material fixed in a standard pipettor tip, washed with 0.5% acetic acid and eluted with 80% acetonitrile 0.5% acetic acid. Finally, the acetonitrile buffer was removed using a SpeedVac (Savant DNA 110 SpeedVac; Global Medical Instruments) after which the peptides were dissolved in 0.5% acetic acid for nLC-MS/MS analysis.

C18 Reversed phase LC-MS/MS analysis

The C18 reversed phase nanoflow LC-MS/MS analyses were performed using a Proxeon nanoflow liquid chromatograph (Easy-nLC, Thermo Fisher, Bremen, Germany) coupled on-line to a 7T linear quadrupole ion trap Fourier transform ion cyclotron resonance mass spectrometer (LTQ-FT Ultra, Thermo Fisher Scientific, Bremen, Germany) as described.[2] Briefly, chromatographic separations were performed using a 15-cm fused silica emitter (PicoTip emitter; tip, 8 ±1 µm; internal diameter, 100 µm; FS360-100-8-N-5-C15; New Objective) packed in-house with reverse phase ReproSil-Pur C18AQ 3 µm resin (Dr. Maisch GmbH). Peptides were eluted from the column using a linear gradient of 8-16% acetonitrile in

75 min and 16-32% acetonitrile in 55 min at a flow rate of 300 nL/min. The LTQ-FT ultra mass spectrometer was programmed to acquire a survey MS scan in the ICR cell (350-1600 m/z, R=100.000 FWHM @ 400 m/z, 1E6 ions) with parallel MS/MS spectra acquisition of the top 4 most abundant ions in the linear ion trap (30% normalized collision energy, 30ms activation time, activation Q=0.250, 3000 ions). Dynamic exclusion was enabled to prevent re-analyses of peptides during the analysis (exclusion duration: 180 seconds, early expiration enabled: 10 scans with S/N=2). Only peptide ions with charge states z=2+ and 3+ were considered for MS/MS spectra acquisition

Database searching and result validation

Database searches and validation of results was performed as described [2]. Briefly, the raw data files acquired by the nLC-MS/MS instrument were converted to mascot generic files using DTA supercharge.[3] Peptides and proteins were identified using Mascot software (Mascot 2.2; Matrix Science) to search a local copy of an in-house created pan-genomic *B. pertussis* protein database (based on the sequence data as described[4]) supplemented with known contaminant protein sequences (e.g. Trypsin, LysC and human skin proteins). The following parameters were used: 15 ppm precursor mass tolerance and a 0.5 Da fragment ion mass tolerance. Furthermore, one missed trypsin cleavage was tolerated and carbamidomethylation of cysteines was set as a fixed modification. Variable modifications included oxidation of methionine residues and N-terminal protein acetylation. The Mascot search results were subjected to heuristic iterative protein false discovery rate validation method of Weatherly to achieve a 1% false discovery rate or better.[5] A protein was considered identified when it was detected in at least 3 out of 4 biological replicates for the planktonic condition and 2 out of 3 biological replicates for the biofilm condition. Exponentially modified PAI (emPAI) scores were calculated as described. [6]

Label-Free Quantitative analysis

The IDEAL-Q (ID-based Elution time Alignment by Linear regression Quantification) software program was used for the label-free quantitative analysis of the nLC-MS/MS data. [7] The cytosolic and membrane protein fractions were processed independently. Semi-quantitative information was extracted from the LC-MS data and the Mascot search results by extracted ion currents. Details of the IDEAL-Q method are described by Tsou and coworkers. [7] Briefly, IDEAL-Q attempts to extract ion current information for every identified peptide (identified in any analysis) even in the absence of Mascot identification data for a peptide in some of the analyses. The following peptide ion information is used by IDEAL-Q to pinpoint ions across files where no identification information is available: mass to charge ratio (m/z), charge state (z), normalized retention time, and isotopic pattern. The combination of all of the above mentioned peptide ion characteristics allow IDEAL-Q to identify ions when no Mascot data is available for a specific analysis. Settings used by IDEAL-Q include 30 ppm mass tolerance, nondegenerate unique peptides only, and Dixon's outlier test to eliminate peptide ratio outliers for each proteins at 95% significance level. Protein ratios were calculated as the weighted average of all respective peptide ratios.

Processing semi-quantitative proteomics data

Protein abundance ratios were generated by IDEAL-Q for each biological replicate relative to mid-exponential sample one and normalized using the median ratio of all proteins quantified. The normalized ratios were Log2 transformed and used for a One-Way ANOVA with the maximum number of permutations (=34650) to identify proteins that were significantly different (*P*-value ≤ 0.05) between the three groups. Proteins that were found to be significantly different based on less than 3 out of 4 biological replicates (or 2 out of 3 biological replicates for the biofilm condition) in one of the groups, were excluded. The *P*-values of all the remaining proteins (both significant and non-significant) were corrected for multiple testing using the FDR method of Storey and Tibshirani to obtain q-values. [8] Proteins

with a *P*-value ≤ 0.05 and a fold change of ≥3 or ≤-3 were considered biofilm-regulated. Biofilm-regulated proteins were aggregated based on function (main role according to the TIGR *B. pertussis* genome and *B. bronchiseptica* database) and subcellular localization (predicted using PSORTb v3.0[9]) and significant enrichment in a certain class was determined by Fisher's exact test.

Validation proteomic datasets

Proteomics allowed us to identify a total of 749, 729, and 825 proteins in the mid-log, stationary, and biofilm cultures, respectively, representing 21-24% of the predicted total of 3449 protein-coding ORFs in the *B. pertussis* genome. Of the 887 detected proteins, 645 (73%) were identified under all growth conditions (Supplementary Table S2). Of the 749 proteins identified in the mid-log planktonic samples, 238 (32%) proteins were found only in the cytosolic fraction, 102 (14%) only in the membrane fraction, and 409 proteins (55%) were identified in both fractions. In the stationary planktonic samples, 355 of the 729 proteins (49%) were identified in both fractions whereas 220 (30%) and 154 (21%) proteins were found uniquely in the cytosolic and membrane protein fraction respectively. Of the 825 proteins identified in the biofilm samples, 222 (27%) proteins were found only in the cytosolic fraction, 129 (16%) only in the membrane fraction, and 474 proteins (57%) were identified in both fractions.

Correct fractionation of the cytosolic and membrane proteins was confirmed using the protein abundance emPAI values of proteins with a strongly predicted cytosolic and outer membrane localization (determined using PSORTb 3.0). This revealed clear enrichment (high emPAI scores) for cytosolic predicted proteins in the cytosol fractions and membrane-predicted proteins in the membrane fractions of all samples in both (data not shown). Furthermore, western blot analysis revealed that the outer membrane protein BP0840 was exclusively present in the membrane protein fraction of all samples (data not shown). This

data indicates correct protein fractionation of all samples.

Supplemental references

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

2. Wessels HJ, Gloerich J, van der Biezen E, Jetten MS, Kartal B. Liquid chromatography-mass spectrometry-based proteomics of Nitrosomonas. *Methods Enzymol* 2011; **486:** 465-482.

3. Pedrioli PG. Trans-proteomic pipeline: a pipeline for proteomic analysis. *Methods Mol Biol* 2010; **604:** 213-238.

4. Bart MJ, van Gent M, van der Heide HG *et al.* Comparative genomics of prevaccination and modern *Bordetella pertussis* strains. *BMC genomics* 2010; **11:** 627.

5. Weatherly DB, Atwood JA, 3rd, Minning TA, Cavola C, Tarleton RL, Orlando R. A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results. *Mol Cell Proteomics* 2005; **4:** 762-772.

6. Ishihama Y, Oda Y, Tabata T *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005; **4:** 1265-1272.

7. Tsou CC, Tsai CF, Tsui YH *et al.* IDEAL-Q, an automated tool for label-free quantitation analysis using an efficient peptide alignment approach and spectral data validation. *Mol Cell Proteomics* 2010; **9:** 131-144.

8. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* 2003; **100:** 9440-9445.

9. Yu NY, Wagner JR, Laird MR *et al.* PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 2010; **26:** 1608-1615.