The signature protein of the PVC superphylum

Lagkouvardos *et al.*

Supplementary Material

Table S1: Signature proteins identified in genome sequences of members of the PVC superphylum and in metagenomic data. The SP was identified in the genome sequences of all PVC members; only one representative SP is shown if sequences were identical among the same species.
See file "Table S1.xlsx".

Table S2: Detection of the PVC signature protein in transcriptomic and proteomic studies. Due to the missing gene prediction in the genome of *Rhodopirellula baltica* SH1 the respective SP was never detected.

Organism	Method	Detection	Reference
Chlamydia trachomatis	microarray	–	Belland, 2003a (1)
Chlamydia trachomatis	microarray	+	Belland, 2003b (2)
Chlamydia trachomatis	RNA-Seq	+	Albrecht, 2010 (3)
Chlamydia trachomatis	proteomics	–	Shaw, 2002 (4)
Chlamydia trachomatis	proteomics	+	Skipp, 2005 (5)
Chlamydia pneumoniae	proteomics	–	Vandahl, 2001(6)
Chlamydia pneumoniae	proteomics	–	Molestina, 2002 (7)
Chlamydia pneumoniae	proteomics	–	Wehr, 2004 (8)
Chlamydia pneumoniae	proteomics	–	Mukhopadhyay, 2006 (9)
Chlamydia pneumoniae	microarray	+	Maurer, 2007 (10)
Chlamydia pneumoniae	RNA-Seq	+	Albrecht, 2011 (11)
Chlamydia pneumoniae	proteomics	–	Saka, 2011 (12)
Protochlamydia amoebophila	proteomics	–	Heinz, 2010 (13)
Protochlamydia amoebophila	proteomics	–	Sixt, 2011 (14)
Protochlamydia amoebophila	microarray	+	Haider, unpublished
Rhodopirellula baltica	proteomics	–	Gade, 2005a (15)
Rhodopirellula baltica	proteomics	–	Gade, 2005b (16)
Rhodopirellula baltica	proteomics	–	Hieu, 2008 (17)
Rhodopirellula baltica	RNA-Seq	–	Wecker, 2009 (18)
Rhodopirellula baltica	RNA-Seq	–	Wecker, 2010 (19)

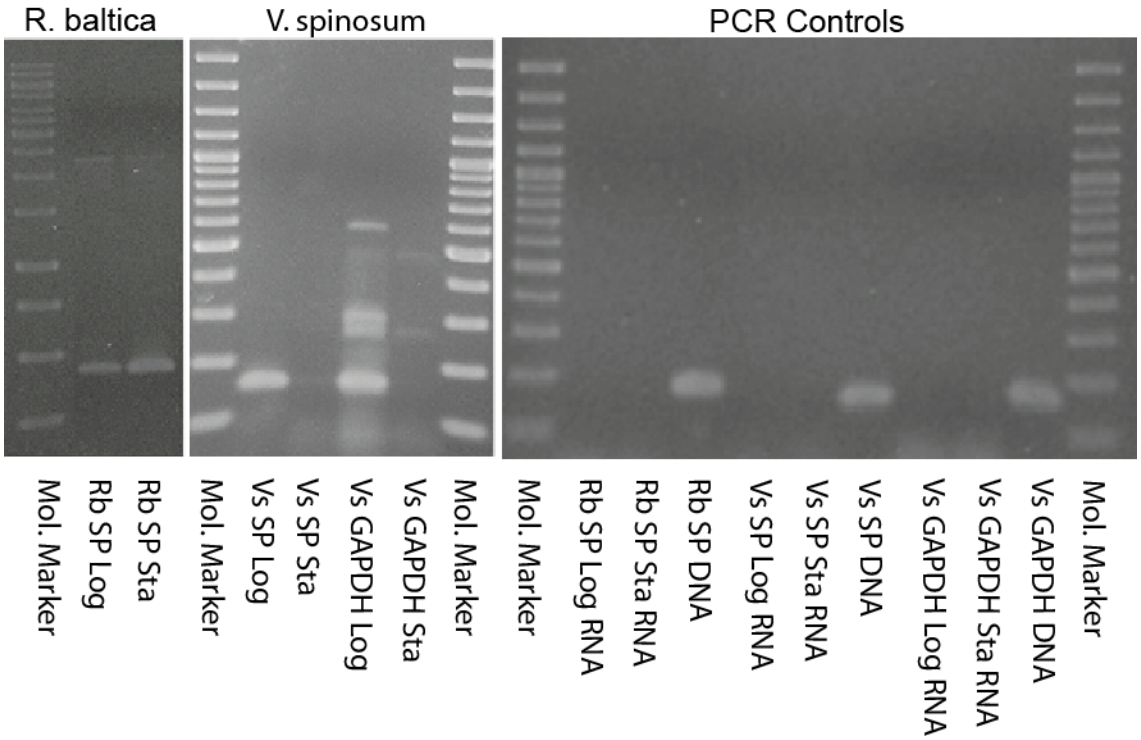

Figure S1: The SP of *Rhodopirellula baltica* (Rb) and *Verrucomicrobium spinosum* (Vs) is transcribed. Reverse transcriptase PCR using RNA isolated during logarithmic (3 days) and stationary growth (6 days). The GAPDH gene of *Verrucomicrobium spinosum* was used as positive control (left panel). A PCR using the same RNA samples demonstrates the absence of DNA in the RNA preparations (right panel). Note that the amount of RNA obtained from *V. spinosum* in the stationary phase was too low for successful detection of SP and the control. All RT-PCR products were cloned and verified by sequencing.

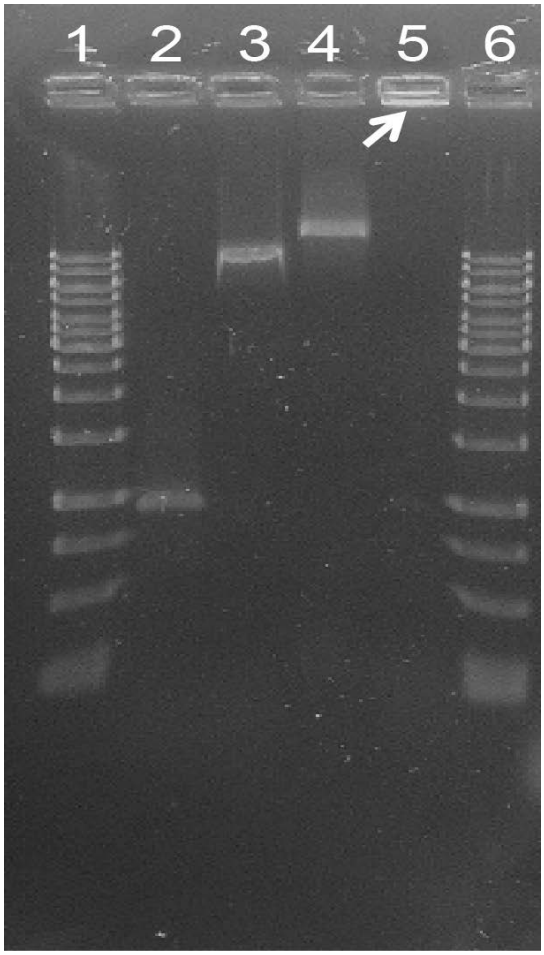

Figure S2: DNA mobility retardation by GST tagged and untagged SP of *R. baltica*. 1,6, molecular marker; 2, DNA (PCR product) only; 3-4 DNA with different doses of GST tagged SP of *R. baltica*; 5, DNA with the same amount of protein as in 4 but thrombin digested. The DNA incubated with the digested mixture of GST and Rb SP didn't enter the gel (arrow). The complete digestion of the GST tag was verified by SDS-PAGE.

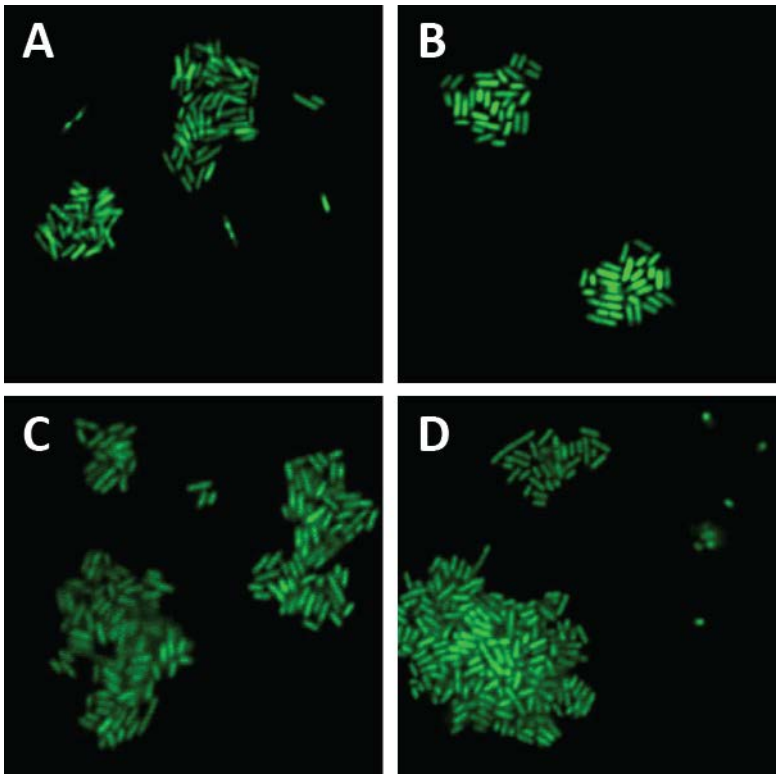

Figure S3: Effect of heterologous overexpression of signature proteins in *E. coli* BL21. A, GST-tagged SP of *P. amoebophila*; B, GST-tagged SP of *R. baltica*; C, GST only; D, non-induced *E.coli* cells. *E. coli* were induced for 2 hours using 1mM IPTG. Expression of the SP was verified by SDS-PAGE. No nucleation was observed after staining with SYBR Green.

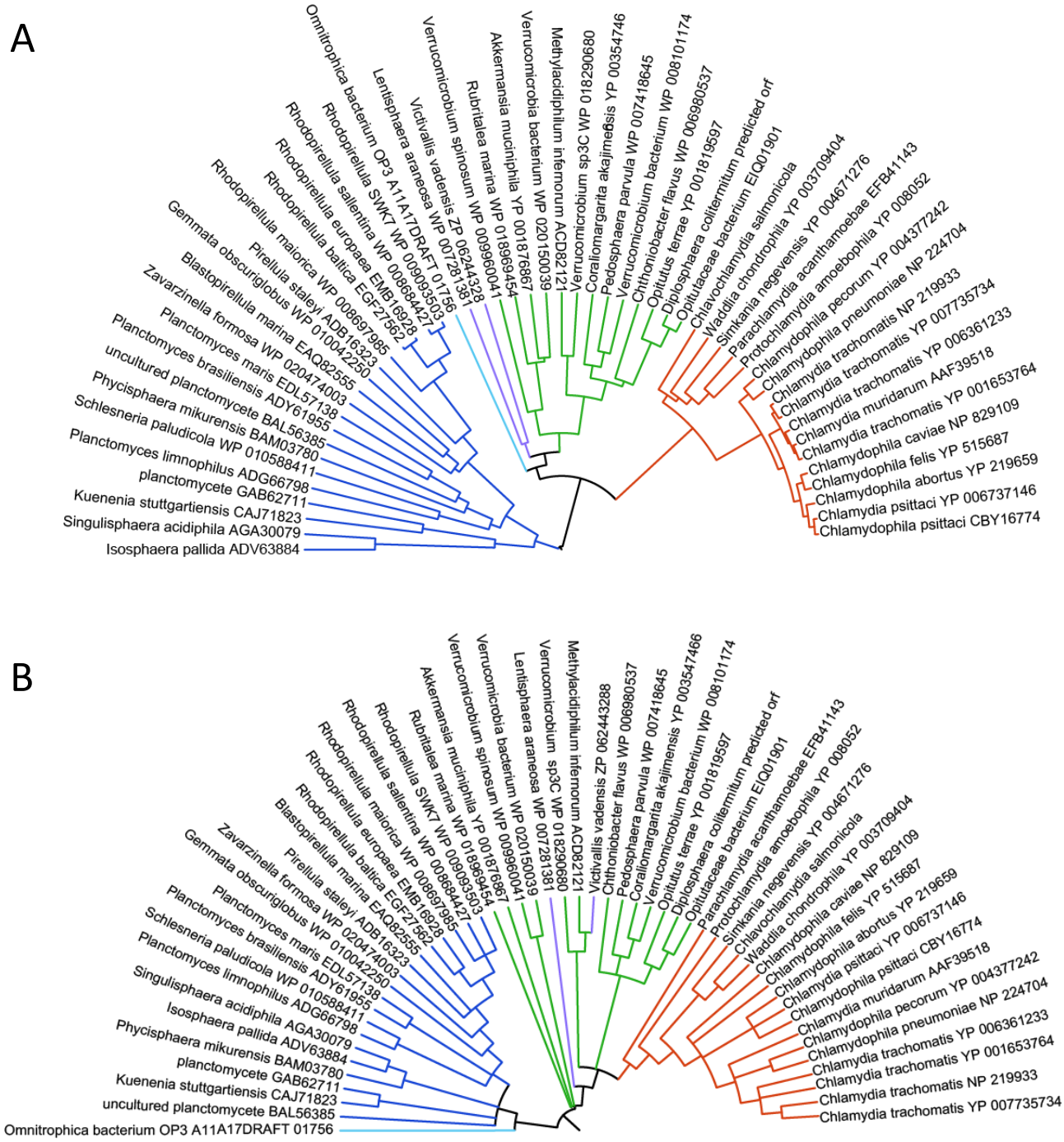

Figure S4. Evolutionary relationships of PVC superphylum signature proteins from published genome sequences. All SP sequences were aligned using MUSCLE (20) in MEGA5 (21) and their evolutionary history was inferred using (A) UPGMA or (B) FastTree (22). The evolutionary distances were computed using the JTT(23) for UPGMA and WAG model (23) for FastTree, while a gamma value of 20 was used for both. Nodes with less than 70% bootstrap support are collapsed in the maximum likelihood tree.

References

1. **Belland RJ, Zhong GM, Crane DD, Hogan D, Sturdevant D, Sharma J, Beatty WL, Caldwell HD.** 2003. Genomic transcriptional profiling of the developmental cycle of Chlamydia trachomatis. P Natl Acad Sci USA **100:**8478-8483.
2. **Belland RJ, Nelson DE, Virok D, Crane DD, Hogan D, Sturdevant D, Beatty WL, Caldwell HD.** 2003. Transcriptome analysis of chlamydial growth during IFN-gamma-mediated persistence and reactivation. P Natl Acad Sci USA **100:**15971-15976.
3. **Albrecht M, Sharma CM, Reinhardt R, Vogel J, Rudel T.** 2010. Deep sequencing-based discovery of the Chlamydia trachomatis transcriptome. Nucleic Acids Res **38:**868-877.
4. **Shaw AC, Gevaert K, Demol H, Hoorelbeke B, Vandekerckhove J, Larsen MR, Roepstorff P, Holm A, Christiansen G, Birkelund S.** 2002. Comparative proteome analysis of Chlamydia trachomatis serovar A, D and L2. Proteomics **2:**164-186.
5. **Skipp P, Robinson J, O'Connor CD, Clarke IN.** 2005. Shotgun proteomic analysis of Chlamydia trachomatis. Proteomics **5:**1558-1573.
6. **Vandahl BB, Birkelund S, Demol H, Hoorelbeke B, Christiansen G, Vandekerckhove J, Gevaert K.** 2001. Proteome analysis of the Chlamydia pneumoniae elementary body. Electrophoresis **22:**1204-1223.
7. **Molestina RE, Klein JB, Miller RD, Pierce WH, Ramirez JA, Summersgill JT.** 2002. Proteomic analysis of differentially expressed Chlamydia pneumoniae genes during persistent infection of HEp-2 cells. Infect Immun **70:**2976-2981.
8. **Wehr W, Meyer TF, Jungblut PR, Muller EC, Szczepek AJ.** 2004. Action and reaction: Chlamydophila pneumoniae proteome alteration in a persistent infection induced by iron deficiency. Proteomics **4:**2969-2981.
9. **Mukhopadhyay S, Good D, Miller RD, Graham JE, Mathews SA, Timms P, Summersgill JT.** 2006. Identification of Chlamydia pneumoniae proteins in the transition from reticulate to elementary body formation. Mol Cell Proteomics **5:**2311-2318.
10. **Maurer AP, Mehlitz A, Mollenkopf HJ, Meyer TF.** 2007. Gene expression profiles of Chlamydophila pneumoniae during the developmental cycle and iron depletion-mediated persistence. Plos Pathog **3:**752-769.
11. **Albrecht M, Sharma CM, Dittrich MT, Muller T, Reinhardt R, Vogel J, Rudel T.** 2011. The transcriptional landscape of Chlamydia pneumoniae. Genome Biol **12**.
12. **Saka HA, Thompson JW, Chen YS, Kumar Y, Dubois LG, Moseley MA, Valdivia RH.** 2011. Quantitative proteomics reveals metabolic and pathogenic properties of Chlamydia trachomatis developmental forms. Mol Microbiol **82:**1185-1203.
13. **Heinz E, Pichler P, Heinz C, den Camp HJMO, Toenshoff ER, Ammerer G, Mechtler K, Wagner M, Horn M.** 2010. Proteomic analysis of the outer membrane of Protochlamydia amoebophila elementary bodies. Proteomics **10:**4363-4376.
14. **Sixt BS, Heinz C, Pichler P, Heinz E, Montanaro J, den Camp HJMO, Ammerer G, Mechtler K, Wagner M, Horn M.** 2011. Proteomic analysis reveals a virtually complete set of proteins for translation and energy generation in elementary bodies of the amoeba symbiont Protochlamydia amoebophila. Proteomics **11:**1868-1892.
15. **Gade D, Stuhrmann T, Reinhardt R, Rabus R.** 2005. Growth phase dependent regulation of protein composition in Rhodopirellula baltica. Environ Microbiol **7:**1074-1084.

16. **Gade D, Theiss D, Lange D, Mirgorodskaya E, Lombardot T, Glockner FO, Kube M, Reinhardt R, Amann R, Lehrach H, Rabus R, Gobom J.** 2005. Towards the proteome of the marine bacterium Rhodopirellula baltica: Mapping the soluble proteins. Proteomics **5:**3654-3671.
17. **Hieu CX, Voigt B, Albrecht D, Becher D, Lombardot T, Glockner FO, Amann R, Hecker M, Schweder T.** 2008. Detailed proteome analysis of growing cells of the planctomycete Rhodopirellula baltical SH1(T). Proteomics **8:**1608-1623.
18. **Wecker P, Klockow C, Ellrott A, Quast C, Langhammer P, Harder J, Glockner FO.** 2009. Transcriptional response of the model planctomycete Rhodopirellula baltica SH1(T) to changing environmental conditions. BMC Genomics **10:**410.
19. **Wecker P, Klockow C, Schuler M, Dabin J, Michel G, Glockner FO.** 2010. Life cycle analysis of the model organism Rhodopirellula baltica SH 1(T) by transcriptome studies. Microb Biotechnol **3:**583-594.
20. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792-1797.
21. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol **28:**2731-2739.
22. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. Plos One **5**.
23. **Whelan S, Goldman N.** 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol **18:**691-699.