Supplementary Table 1. Examples of published proteogenomics studies for different organisms.

Organism	**Reference**
Pristionchus pacificus	1
Arabidopsis	2, 3
Deinococcus deserti	4, 5
Drosophila melanogaster	6-8
Zea mays	9
Anopheles gambiae	10
Mycobacterium tuberculosis	11, 12
Mycobacterium leprae	13
Saccharomyces cerevisiae	14, 15
Candida glabrata	16
M-smegmatis	17
Toxoplasma gondii	18
Stagonospora nodorum	19
C. elegans	20
Cryptococcus neoformans	21
Spiroplasma melliferum	22
Salmonella Typhimurium	23
Blumeria graminis f. sp hordei. (fungal pathogen)	24-26
Schizosaccharomyces pombe	27
Planarian Schmidtea mediterranea (parasitic worm)	28
Antarctic Archaeon Methanococcoides burtonii	29
Leptospira interrogans serovar Lai	30
Psychotria Leaf	31
Apis mellifera L. (honey bee)	32
Aspergillus flavus	33, 34
Ruegeria pomeroyi	35
Roseobacter (group of bacteria -comparative)	36
Echinococcus granulosus	37
Arhodomonas sp Strain Rozel	38
Flavobacterium columnare (fish pathogen)	39
Edwardsiella ictaluri	40
Methanosarcina acetivorans	41
Daphnia pulex and Daphnia longicephala	42
Riftia pachyptila and Tevnia jerichonana	43
Rice	44
Porphyromonas gingivalis	45
Mycoplasma mobile	46
Anopheles gambiae	47, 48
Pseudomonas fluorescens Pf0-1	49
Escherichia coli	50
Bradyrhizobium japonicum USDA110	51
Channel catfish	52

Acholeplasma laidlawii	53
Acidiphilium cryptum JF-5	54
Saccharopolyspora erythraea	55
Bacillus anthracis	56
Helicobacter pylori strain 26695	57
Streptococcus pyogenes SF370	58
Mycoplasma genitalium	59
Leishmania donovani	60
Yersinia pestis	61
Streptococcus suis	62
Streptococcus pneumoniae Virulent Phage Dp-1	63
Roseobacter Denitrificans Och114	64
Venturia pirina	65
Nicotiana occidentalis	66
Mycoplasma hyopneumoniae virulent strain 232	67
Clostridium stercorarium DSM8532	68
Dehalococcoides mccartyi strain CBDB1	69
Methylobacterium extorquens AM1 (ME-AM1)	70
Zebrafish	71

1. Borchert N, Dieterich C, Krug K, Schutz W, Jung S, Nordheim A, Sommer RJ, Macek B: **Proteogenomics of Pristionchus pacificus reveals distinct proteome structure of nematode models**. *Genome Research* 2010, **20**(6):837-846.
2. Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, Zimmermann P, Grossniklaus U, Gruissem W, Baginsky S: **Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics**. *Science* 2008, **320**(5878):938-941.
3. Castellana NE, Payne SH, Shen ZX, Stanke M, Bafna V, Briggs SP: **Discovery and revision of Arabidopsis genes by proteogenomics**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(52):21034-21038.
4. Baudet M, Ortet P, Gaillard JC, Fernandez B, Guerin P, Enjalbal C, Subra G, de Groot A, Barakat M, Dedieu A *et al*: **Proteomics-based Refinement of Deinococcus deserti Genome Annotation Reveals an Unwonted Use of Non-canonical Translation Initiation Codons**. *Molecular & Cellular Proteomics* 2010, **9**(2):415-426.
5. de Groot A, Dulermo R, Ortet P, Blanchard L, Guerin P, Fernandez B, Vacherie B, Dossat C, Jolivet E, Siguier P *et al*: **Alliance of Proteomics and Genomics to Unravel the Specificities of Sahara Bacterium Deinococcus deserti**. *Plos Genetics* 2009, **5**:e1000434.
6. Brunner E, Ahrens CH, Mohanty S, Baetschmann H, Loevenich S, Potthast F, Deutsch EW, Panse C, de Lichtenberg U, Rinner O *et al*: **A high-quality catalog of the Drosophila melanogaster proteome**. *Nature Biotechnology* 2007, **25**(5):576-583.
7. Findlay GD, MacCoss MJ, Swanson WJ: **Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in Drosophila**. *Genome Research* 2009, **19**(5):886-896.

8. Loevenich SN, Brunner E, King NL, Deutsch EW, Stein SE, Aebersold R, Hafen E, FlyBase C: **The Drosophila melanogaster PeptideAtlas facilitates the use of peptide data for improved fly proteomics and genome annotation**. *Bmc Bioinformatics* 2009, **10**:59.
9. Castellana NE, Shen ZX, He YP, Walley JW, Cassidy CJ, Briggs SP, Bafna V: **An automated proteogenomic method uses mass spectrometry to reveal novel genes in Zea mays**. *Molecular & Cellular Proteomics* 2014, **13**(1):157-167.
10. Chaerkady R, Kelkar DS, Muthusamy B, Kandasamy K, Dwivedi SB, Sahasrabuddhe NA, Kim MS, Renuse S, Pinto SM, Sharma R *et al*: **A proteogenomic analysis of Anopheles gambiae using high-resolution Fourier transform mass spectrometry**. *Genome Research* 2011, **21**(11):1872-1881.
11. de Souza GA, Malen H, Softeland T, Saelensminde G, Prasad S, Jonassen I, Wiker HG: **High accuracy mass spectrometry analysis as a tool to verify and improve gene annotation using Mycobacterium tuberculosis as an example**. *BMC Genomics* 2008, **9**:316.
12. Kelkar DS, Kumar D, Kumar P, Balakrishnan L, Muthusamy B, Yadav AK, Shrivastava P, Marimuthu A, Anand S, Sundaram H *et al*: **Proteogenomic Analysis of Mycobacterium tuberculosis By High Resolution Mass Spectrometry**. *Molecular & Cellular Proteomics* 2011, **10**(12): M111.011627.
13. de Souza GA, Softeland T, Koehler CJ, Thiede B, Wiker HG: **Validating divergent ORF annotation of the Mycobacterium leprae genome through a full translation data set and peptide identification by tandem mass spectrometry**. *Proteomics* 2009, **9**(12):3233-3243.
14. Oshiro G, Wodicka LM, Washburn MP, Yates JR, Lockhart DJ, Winzeler EA: **Parallel identification of new genes in Saccharomyces cerevisiae**. *Genome Research* 2002, **12**(8):1210-1220.
15. Helsens K, Van Damme P, Degroeve S, Martens L, Arnesen T, Vandekerckhove J, Gevaert K: **Bioinformatics Analysis of a Saccharomyces cerevisiae N-Terminal Proteome Provides Evidence of Alternative Translation Initiation and Post-Translational N-Terminal Acetylation**. *Journal of Proteome Research* 2011, **10**(8):3578-3589.
16. Prasad TSK, Harsha HC, Keerthikumar S, Sekhar NR, Selvan LDN, Kumar P, Pinto SM, Muthusamy B, Subbannayya Y, Renuse S *et al*: **Proteogenomic Analysis of Candida glabrata using High Resolution Mass Spectrometry**. *Journal of Proteome Research* 2012, **11**(1):247-260.
17. Wang R, Prince JT, Marcotte EM: **Mass spectrometry of the M-smegmatis proteome: Protein expression levels correlate with function, operons, and codon bias**. *Genome Research* 2005, **15**(8):1118-1126.
18. Xia D, Sanderson SJ, Jones AR, Prieto JH, Yates JR, Bromley E, Tomley FM, Lal K, Sinden RE, Brunk BP *et al*: **The proteome of Toxoplasma gondii: integration with the genome provides novel insights into gene expression and annotation**. *Genome Biology* 2008, **9**(7): R116.
19. Bringans S, Hane JK, Casey T, Tan KC, Lipscombe R, Solomon PS, Oliver RP: **Deep proteogenomics; high throughput gene validation by multidimensional liquid chromatography and mass spectrometry of proteins from the fungal wheat pathogen Stagonospora nodorum**. *Bmc Bioinformatics* 2009, **10**:301.

20. Merrihew GE, Davis C, Ewing B, Williams G, Kall L, Frewen BE, Noble WS, Green P, Thomas JH, MacCoss MJ: **Use of shotgun proteomics for the identification, confirmation, and correction of C. elegans gene annotations**. *Genome Research* 2008, **18**(10):1660-1669.
21. Nagarajha Selvan L, Kaviyil J, Nirujogi R, Muthusamy B, Puttamallesh V, Subbannayya T, Syed N, Radhakrishnan A, Kelkar D, Ahmad S *et al*: **Proteogenomic analysis of pathogenic yeast Cryptococcus neoformans using high resolution mass spectrometry**. *Clinical Proteomics* 2014, **11**(1):5.
22. Alexeev D, Kostrjukova E, Aliper A, Popenko A, Bazaleev N, Tyakht A, Selezneva O, Akopian T, Prichodko E, Kondratov I *et al*: **Application of Spiroplasma melliferum Proteogenomic Profiling for the Discovery of Virulence Factors and Pathogenicity Mechanisms in Host-associated Spiroplasmas**. *Journal of Proteome Research* 2012, **11**(1):224-236.
23. Ansong C, Tolic N, Purvine SO, Porwollik S, Jones M, Yoon H, Payne SH, Martin JL, Burnet MC, Monroe ME *et al*: **Experimental annotation of post-translational features and translated coding regions in the pathogen Salmonella Typhimurium**. *Bmc Genomics* 2011, **12**:433.
24. Bindschedler LV, Burgis TA, Mills DJS, Ho JTC, Cramer R, Spanu PD: **In Planta Proteomics and Proteogenomics of the Biotrophic Barley Fungal Pathogen Blumeria graminis f. sp hordei**. *Molecular & Cellular Proteomics* 2009, **8**(10):2368-2381.
25. Bindschedler LV, McGuffin LJ, Burgis TA, Spanu PD, Cramer R: **Proteogenomics and in silico structural and functional annotation of the barley powdery mildew Blumeria graminis f. sp hordei**. *Methods* 2011, **54**(4):432-441.
26. Pedersen C, van Themaat EVL, McGuffin LJ, Abbott JC, Burgis TA, Barton G, Bindschedler LV, Lu XL, Maekawa T, Wessling R *et al*: **Structure and evolution of barley powdery mildew effector candidates**. *Bmc Genomics* 2012, **13**:694.
27. Bitton DA, Wood V, Scutt PJ, Grallert A, Yates T, Smith DL, Hagan IM, Miller CJ: **Augmented Annotation of the Schizosaccharomyces pombe Genome Reveals Additional Genes Required for Growth and Viability**. *Genetics* 2011, **187**(4):1207-U1369.
28. Bocchinfuso DG, Taylor P, Ross E, Ignatchenko A, Ignatchenko V, Kislinger T, Pearson BJ, Moran MF: **Proteomic Profiling of the Planarian Schmidtea mediterranea and Its Mucous Reveals Similarities with Human Secretions and Those Predicted for Parasitic Flatworms**. *Molecular & Cellular Proteomics* 2012, **11**(9):681-691.
29. Burg DW, Lauro FM, Williams TJ, Raftery MJ, Guilhaus M, Cavicchioli R: **Analyzing the Hydrophobic Proteome of the Antarctic Archaeon Methanococcoides burtonii Using Differential Solubility Fractionation**. *Journal of Proteome Research* 2010, **9**(2):664-676.
30. Cao XJ, Dai J, Xu H, Nie S, Chang X, Hu BY, Sheng QH, Wang LS, Ning ZB, Li YX *et al*: **High-coverage proteome analysis reveals the first insight of protein modification systems in the pathogenic spirochete Leptospira interrogans**. *Cell Research* 2010, **20**(2):197-210.
31. Carlier AL, Omasits U, Ahrens CH, Eberl L: **Proteomics Analysis of Psychotria Leaf Nodule Symbiosis: Improved Genome Annotation and Metabolic Predictions**. *Molecular Plant-Microbe Interactions* 2013, **26**(11):1325-1333.

32. Chan QWT, Parker R, Sun Z, Deutsch EW, Foster LJ: **A honey bee (Apis mellifera L.) PeptideAtlas crossing castes and tissues**. *Bmc Genomics* 2011, **12**:290.
33. Chang KY, Georgianna DR, Heber S, Payne GA, Muddiman DC: **Detection of Alternative Splice Variants at the Proteome Level in Aspergillus flavus**. *Journal of Proteome Research* 2010, **9**(3):1209-1217.
34. Chang KY, Muddiman DC: **Identification of alternative splice variants in Aspergillus flavus through comparison of multiple tandem MS search algorithms**. *Bmc Genomics* 2011, **12**:358.
35. Christie-Oleza JA, Miotello G, Armengaud J: **High-throughput proteogenomics of Ruegeria pomeroyi: seeding a better genomic annotation for the whole marine Roseobacter clade**. *Bmc Genomics* 2012, **13**:73.
36. Christie-Oleza JA, Pina-Villalonga JM, Bosch R, Nogales B, Armengaud J: **Comparative Proteogenomics of Twelve Roseobacter Exoproteomes Reveals Different Adaptive Strategies Among These Marine Bacteria**. *Molecular & Cellular Proteomics* 2012, **11**(2):M111.013110.
37. Cui SJ, Xu LL, Zhang T, Xu M, Yao J, Fang CY, Feng Z, Yang PY, Hu W, Liu F: **Proteomic characterization of larval and adult developmental stages in Echinococcus granulosus reveals novel insight into host-parasite interactions**. *Journal of Proteomics* 2013, **84**:158-175.
38. Dalvi S, Azetsu S, Patrauchan MA, Aktas DF, Fathepure BZ: **Proteogenomic Elucidation of the Initial Steps in the Benzene Degradation Pathway of a Novel Halophile, Arhodomonas sp Strain Rozel, Isolated from a Hypersaline Environment**. *Applied and Environmental Microbiology* 2012, **78**(20):7309-7316.
39. Dumpala PR, Gulsoy N, Lawrence ML, Karsi A: **Proteomic analysis of the fish pathogen Flavobacterium columnare**. *Proteome Science* 2010, **8**:26.
40. Dumpala PR, Lawrence ML, Karsi A: **Proteome analysis of Edwardsiella ictaluri**. *Proteomics* 2009, **9**(5):1353-1363.
41. Ferguson JT, Wenger CD, Metcalf WW, Kelleher NL: **Top-Down Proteomics Reveals Novel Protein Forms Expressed in Methanosarcina acetivorans**. *Journal of the American Society for Mass Spectrometry* 2009, **20**(9):1743-1750.
42. Frohlich T, Arnold GJ, Fritsch R, Mayr T, Laforsch C: **LC-MS/MS-based proteome profiling in Daphnia pulex and Daphnia longicephala: the Daphnia pulex genome database as a key for high throughput proteomics in Daphnia**. *Bmc Genomics* 2009, **10**:171.
43. Gardebrecht A, Markert S, Sievert SM, Felbeck H, Thurmer A, Albrecht D, Wollherr A, Kabisch J, Le Bris N, Lehmann R *et al*: **Physiological homogeneity among the endosymbionts of Riftia pachyptila and Tevnia jerichonana revealed by proteogenomics**. *Isme Journal* 2012, **6**(4):766-776.
44. Helmy M, Tomita M, Ishihama Y: **OryzaPG-DB: Rice Proteome Database based on Shotgun Proteogenomics**. *Bmc Plant Biology* 2011, **11**:63.
45. Hendrickson EL, Xia QW, Wang TS, Lamont RJ, Hackett M: **Pathway analysis for intracellular Porphyromonas gingivalis using a strain ATCC 33277 specific database**. *Bmc Microbiology* 2009, **9**:185.
46. Jaffe JD, Stange-Thomann N, Smith C, DeCaprio D, Fisher S, Butler J, Calvo S, Elkins T, Fitzgerald MG, Hafez N *et al*: **The complete genome and proteome of Mycoplasma mobile**. *Genome Research* 2004, **14**(8):1447-1461.

47. Kalume DE, Peri S, Reddy R, Zhong J, Okulate M, Kumar N, Pandey A: **Genome annotation of Anopheles gambiae using mass spectrometry-derived data**. *Bmc Genomics* 2005, **6**:128.

48. Li J, Riehle MM, Zhang Y, Xu JN, Oduol F, Gomez SM, Eiglmeier K, Beatrix MU, Shabanowitz J, Donald FH *et al*: **Anopheles gambiae genome reannotation through synthesis of ab initio and comparative gene prediction algorithms**. *Genome Biology* 2006, **7**(3):R24.

49. Kim W, Silby MW, Purvine SO, Nicoll JS, Hixson KK, Monroe M, Nicora CD, Lipton MS, Levy SB: **Proteomic Detection of Non-Annotated Protein-Coding Genes in Pseudomonas fluorescens Pf0-1**. *Plos One* 2009, **4**(12).

50. Krug K, Carpy A, Behrends G, Matic K, Soares NC, Macek B: **Deep coverage of the Escherichia coli proteome enables the assessment of false discovery rates in simple proteogenomic experiments**. *Molecular & Cellular Proteomics* 2013, **12**(11):3420-3430.

51. Kumar D, Yadav AK, Kadimi PK, Nagaraj SH, Grimmond SM, Dash D: **Proteogenomic Analysis of Bradyrhizobium japonicum USDA110 Using Genosuite, an Automated Multi-algorithmic Pipeline**. *Molecular & Cellular Proteomics* 2013, **12**(11):3388-3397.

52. Kunec D, Nanduri B, Burgess SC: **Experimental annotation of channel catfish virus by probabilistic proteogenomic mapping**. *Proteomics* 2009, **9**(10):2634-2647.

53. Lazarev VN, Levitskii SA, Basovskii YI, Chukin MM, Akopian TA, Vereshchagin VV, Kostrjukova ES, Kovaleva GY, Kazanov MD, Malko DB *et al*: **Complete Genome and Proteome of Acholeplasma laidlawii**. *Journal of Bacteriology* 2011, **193**(18):4943-4953.

54. Magnuson TS, Swenson MW, Paszczynski AJ, Deobald LA, Kerk D, Cummings DE: **Proteogenomic and functional analysis of chromate reduction in Acidiphilium cryptum JF-5, an Fe(III)-respiring acidophile**. *Biometals* 2010, **23**(6):1129-1138.

55. Marcellin E, Licona-Cassani C, Mercer TR, Palfreyman RW, Nielsen LK: **Re-annotation of the Saccharopolyspora erythraea genome using a systems biology approach**. *Bmc Genomics* 2013, **14**:699.

56. Misra RV, Ahmod NZ, Parker R, Fang M, Shah H, Gharbia S: **Developing an integrated proteo-genomic approach for the characterisation of biomarkers for the identification of Bacillus anthracis**. *Journal of Microbiological Methods* 2012, **88**(2):237-247.

57. Muller SA, Findeiss S, Pernitzsch SR, Wissenbach DK, Stadler PF, Hofacker IL, von Bergen M, Kalkhof S: **Identification of new protein coding sequences and signal peptidase cleavage sites of Helicobacter pylori strain 26695 by proteogenomics**. *Journal of Proteomics* 2013, **86**:27-42.

58. Okamoto A, Yamada K: **Proteome driven re-evaluation and functional annotation of the Streptococcus pyogenes SF370 genome**. *Bmc Microbiology* 2011, **11**:249.

59. Parraga-Nino N, Colome-Calls N, Canals F, Querol E, Ferrer-Navarro M: **A Comprehensive Proteome of Mycoplasma genitalium**. *Journal of Proteome Research* 2012, **11**(6):3305-3316.

60. Pawar H, Sahasrabuddhe NA, Renuse S, Keerthikumar S, Sharma J, Kumar GSS, Venugopal A, Sekhar NR, Kelkar DS, Nemade H *et al*: **A proteogenomic approach to map the proteome of an unsequenced pathogen - Leishmania donovani**. *Proteomics* 2012, **12**(6):832-844.

61. Payne SH, Huang ST, Pieper R: **A proteogenomic update to Yersinia: enhancing genome annotation**. *Bmc Genomics* 2010, **11**:460.

62. Rodriguez-Ortega MJ, Luque I, Tarradas C, Barcena JA: **Overcoming function annotation errors in the Gram-positive pathogen Streptococcus suis by a proteomics-driven approach**. *Bmc Genomics* 2008, **9**:588.

63. Sabri M, Hauser R, Ouellette M, Liu J, Dehbi M, Moeck G, Garcia E, Titz B, Uetz P, Moineau S: **Genome Annotation and Intraviral Interactome for the Streptococcus pneumoniae Virulent Phage Dp-1**. *Journal of Bacteriology* 2011, **193**(2):551-562.

64. Bland C, Hartmann EM, Christie-Oleza JA, Fernandez B, Armengaud J: **N-Terminal-oriented Proteogenomics of the Marine Bacterium Roseobacter Denitrificans Och114 using N-Succinimidyloxycarbonylmethyl) tris(2,4,6trimethoxyphenyl) phosphonium bromide (TMPP) Labeling and Diagonal Chromatography***. *Molecular & Cellular Proteomics* 2014, **13**(5):1369-1381.

65. Cooke IR, Jones D, Bowen JK, Deng C, Faou P, Hall NE, Jayachandran V, Liem M, Taranto AP, Plummer KM *et al*: **Proteogenomic Analysis of the Venturia pirina (Pear Scab Fungus) Secretome Reveals Potential Effectors**. *Journal of Proteome Research* 2014, **13**(8):3635-3644.

66. Luge T, Kube M, Freiwald A, Meierhofer D, Seemuller E, Sauer S: **Transcriptomics assisted proteomic analysis of Nicotiana occidentalis infected by Candidatus Phytoplasma mali strain AT**. *Proteomics* 2014, **14**(16):1882-1889.

67. Pendarvis K, Padula MP, Tacchi JL, Petersen AC, Djordjevic SP, Burgess SC, Minion FC: **Proteogenomic mapping of Mycoplasma hyopneumoniae virulent strain 232**. *BMC Genomics* 2014, **15**:576.

68. Schellenberg JJ, Verbeke TJ, McQueen P, Krokhin OV, Zhang XL, Alvare G, Fristensky B, Thallinger GG, Henrissat B, Wilkins JA *et al*: **Enhanced whole genome sequence and annotation of Clostridium stercorarium DSM8532(T) using RNA-seq transcriptomics and high-throughput proteomics**. *BMC Genomics* 2014, **15**:567.

69. Schiffmann CL, Jehmlich N, Otto W, Hansen R, Nielsen PH, Adrian L, Seifert J, von Bergen M: **Proteome profile and proteogenomics of the organohalide-respiring bacterium Dehalococcoides mccartyi strain CBDB1 grown on hexachlorobenzene as electron acceptor**. *Journal of proteomics* 2014, **98**:59-64.

70. Kumar D, Mondal AK, Yadav AK, Dash D: **Discovery of rare protein coding genes in model methylotroph Methylobacterium extorquens AM1**. *Proteomics* 2014. doi: 10.1002/pmic.201400153. [Epub ahead of print]

71. Kelkar DS, Provost E, Chaerkady R, Muthusamy B, Manda SS, Subbannayya T, Selvan LD, Wang CH, Datta KK, Woo S *et al*: **Annotation of the zebrafish genome through an integrated transcriptomic and proteomic analysis**. *Mol Cell Proteomics* 2014, **24**:038299.