Supporting Online Material

Song, C., et al. 2009. Bioinformatic comparison of bacterial secretomes. Genomics Proteomics Bioinformatics 7: 37-46. DOI: 10.1016/S1672-0229(08)60031-5

Table S1 Predicted secretomes

Bacterium	Proteome (# of ORFs)	Secretome	% of Proteome	Accession #	Reference
Gram negative					
Acinetobacter sp. ADP1*	3325	869	26.1	NC_005966	1
Agrobacterium tumefaciens str. C58 chromosome circular*	2715	754	27.8	NC_003062	2
Anabaena variabilis ATCC 29413	5039	1177	23.4	NC_007413	3
Anaplasma marginale str. St. Maries*	949	277	29.2	NC_004842	4
Aquifex aeolicus VF5	1529	319	20.9	NC_000918	5
Azoarcus sp.	4133	1198	29.0	NC_006513	6
Bacillus clausii KSM-K16	4096	970	23.7	NC_006582	7
Bacteroides fragilis NCTC 9343*	4184	1504	35.9	NC_003228	8
Bacteroides thetaiotaomicron VPI-5482	4778	1797	37.6	NC_004663	9
Bartonella henselae str. Houston-1*	1488	371	24.9	NC_005956	10
Bartonella quintana str. Toulouse*	1142	287	25.1	NC_005955	10
Bdellovibrio bacteriovorus HD 100	3587	1520	42.4	NC_005363	11
Bordetella bronchiseptica RB 50*	4994	1807	36.2	NC_002927	12
Bordetella parapertussis 12822*	4185	1506	36.0	NC_002928	12
Borrelia burgdorferi B31*	851	199	23.4	NC_001318	13
Borrelia garini PBii*	832	192	23.1	NC_006156	14
Bradyrhizobium japonicum USDA 110	8317	2577	31.0	NC_004463	15
Brucella abortus biovar 1 str. 9-941 chromosome I*	2030	551	27.1	NC_006932	16
Brucella melitensis 16M chromosome I*	2059	489	23.7	NC_003317	17
Brucella suis 1330*	2123	583	27.5	NC_004310	17
Buchnera aphidicola str. APS	564	71	12.6		

(Acyrthosiphon pisum)				NC_002528	18
Burkholderia mallei ATCC 23344 chromosome 1*	2996	1017	33.9	NC_006348	19
Burkholderia pseudomallei 1710b chromosome I*	3736	1150	30.8	NC_007434	20
Burkholderia sp. 383*	1209	398	32.9	NC_007509	21
Campylobacter jejuni RM1221*	1838	437	23.8	NC_003912	22
Caulobacter crescentus CB15	3737	1196	32.0	NC_002696	23
Chlamydia muridarum Nigg*	904	255	28.2	NC_002620	24
Chlamydia trachomatis A/HAR-13*	911	259	28.4	NC_007429	25
Chlamydophila abortus S26/3*	932	253	27.1	NC_004552	26
Chlamydophila caviae GPIC*	998	272	27.3	NC_003361	27
Chlamydophila pneumoniae AR39*	1112	294	26.4	NC_002179	24
Chlorobium tepidum TLS	2252	519	23.0	NC_002932	28
Chromobacterium violaceum ATCC 12472	4407	1461	33.2	NC_005085	29
Coxiella burnetii RSA 493*	2016	435	21.6	NC_002971	30
Dechloromonas aromatica RCB	4171	1449	34.7	NC_007298	31
Dehalococcoides ethenogenes 195	1580	332	21.0	NC_002936	32
Desulfotalea psychrophila LSv54	3116	823	26.4	NC_006138	33
Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough	3379	915	27.1	NC_002937	34
Ehrlichia canis str. Jake*	925	243	26.3	NC_007354	35
Ehrlichia ruminantium str. Gardel*	950	227	23.9	NC_006831	36
Erwinia carotovora subsp. atroseptica SCRI1043*	4472	1339	29.9	NC_004547	37
Escherichia coli CFT073*	5379	1468	27.3	NC_004431	38
Francisella tularensis subsp. tularensis SCHU S4*	1603	420	26.2	NC_006570	39
Fusobacterium nucleatum subsp. nucleatum ATCC 25586	2067	361	17.5	NC_003454	40
Geobacter sulfurreducens PCA	3446	1047	30.4	NC_002939	41

Gloeobacter violaceus PCC 7421	4430	1267	28.6	NC_005125	42
Gluconobacter oxydans 621H	2432	730	30.0	NC_006677	43
Haemophilus ducreyi 35000HP*	1717	460	26.8	NC_002940	44
Haemophilus influenzae 86-028NP*	1791	451	25.2	NC_007146	45
Helicobacter hepaticus ATCC 51449*	1875	454	24.2	NC_004917	46
Helicobacter pylori 26695*	1576	444	28.2	NC_000915	47
Idiomarina loihiensis L2TR	2628	843	32.1	NC_006512	48
Legionella pneumophila str. Lens*	2878	772	26.8	NC_006369	49
Leptospira interrogans serovar Copenhageni str. Fiocruz L1-130*	3394	796	23.5	NC_005823	50
Mannheimia succiniciproducens MBEL55E	2380	603	25.3	NC_006300	51
Mesorhizobium loti MAFF303099	6743	1911	28.3	NC_002678	3
Neisseria gonorrhoeae FA 1090*	2002	530	26.5	NC_002946	52
Neisseria meningitidis Z2491*	2065	550	26.6	NC_003116	53
Nitrobacter winogradskyi Nb-255	3122	881	28.2	NC_007406	54
Nitrosomonas europaea ATCC 19718	2461	638	25.9	NC_004757	55
Nostoc sp. PCC 7120	5366	1219	22.7	NC_003272	56
Oceanobacillus iheyensis HTE831	3500	752	21.5	NC_004193	57
Pasteurella multocida subsp. multocida str. Pm70*	2015	595	29.5	NC_002663	58
Pelobacter carbinolicus DSM 2380	3118	803	25.8	NC_007498	59
Pelodictyon luteolum DSM 273	2083	541	26.0	NC_007512	60
Photobacterium profundum SS9 chromosome 1	3416	941	27.5	NC_006370	61
Pirellula sp.	7325	2043	27.9	NC_005027	15
Porphyromonas gingivalis W83*	1909	509	26.7	NC_002950	62
Prochlorococcus marinus str. MIT 9313	2265	614	27.1	NC_005071	63
Pseudoalteromonas haloplanktis TAC125 chromosome I	2940	942	32.0	NC_007481	64
Pseudomonas aeruginosa PAO1*	5567	1979	35.5	NC_002516	65
Pseudomonas fluorescens Pf-5	6137	2089	34.0	NC_004129	66

Pseudomonas putida KT2440	5350	1744	32.6	NC_002947	67
Ralstonia eutropha JMP134 chromosome 1	3439	1119	32.5	NC_007347	68
*Ralstonia solanacearum**	3440	1166	33.9	NC_003295	69
Rhodobacter sphaeroides 2.4.1 chromosome 1	3022	932	30.8	NC_007493	70
Rhodopirellula baltica SH 1	7325	2043	27.9	NC_005027	15
Rhodopseudomonas palustris CGA009	4813	1576	32.7	NC_005296	71
Rickettsia conorii str. Malish 7*	1374	258	18.8	NC_003103	72
Rickettsia felis URRWXCal2*	1400	244	17.4	NC_007109	73
Rickettsia prowazekii str. Madrid Ei*	835	176	21.1	NC_000963	74
Rickettsia typhi str. Wilmingtoni*	838	172	20.5	NC_006142	75
Salmonella enterica subsp. enterica serovar Choleraesuis str.*	4441	1314	29.6	NC_006905	76
Salmonella typhimurium LT2*	4425	1358	30.7	NC_003197	77
Shewanella oneidensis MR-1	4324	1393	32.2	NC_004347	78
Shigella flexneri 2a str. 2457T*	4068	1052	25.9	NC_004741	79
Shigella sonnei Ss046*	4223	1086	25.7	NC_007384	80
Silicibacter pomeroyi DSS-3	3810	1155	30.3	NC_003911	81
Sinorhizobium meliloti 1021	3341	1021	30.6	NC_003047	82
Symbiobacterium thermophilum IAM 14863	3337	901	27.0	NC_006177	83
Synechococcus sp. CC9605	2638	688	26.1	NC_007516	84
Synechocystis sp. PCC 6803	3167	783	24.7	NC_000911	85
Thermoanaerobacter tengcongensis MB4	2588	524	20.2	NC_003869	86
Thermotoga maritima MSB8	1858	421	22.7	NC_000853	87
Thermus thermophilus HB8	1973	620	31.4	NC_006461	88
Thiobacillus denitrificans ATCC 25259	2827	1019	36.0	NC_007404	89
Thiomicrospira crunogena XCL-2	2192	597	27.2	NC_007520	90
Thiomicrospira denitrificans ATCC 33889	2097	556	26.5	NC_007575	91
Treponema denticola ATCC 35405*	2767	772	27.9	NC_002967	92
Treponema pallidum subsp. pallidum str. Nichols*	1036	348	33.6	NC_000919	93

Ureaplasma parvum serovar 3 str. ATCC 700970*	614	119	19.4	NC_002162	94
Vibrio cholerae O1 biovar eltor str. N16961 chromosome I*	2742	799	29.1	NC_002505	95
Vibrio fischeri ES114 chromosome I	2575	756	29.4	NC_006840	96
Vibrio parahaemolyticus RIMD 2210633 chromosome I*	3080	904	29.4	NC_004603	97
Vibrio vulnificus YJ016 chromosome I*	3259	944	29.0	NC_005139	98
Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis	611	79	12.9	NC_004344	99
Wolbachia endosymbiont of Drosophila melanogaster	1195	215	18.0	NC_002978	100
Wolinella succinogenes DSM 1740	2043	592	29.0	NC_005090	101
Xanthomonas axonopodis pv. citri str. 306*	4312	1606	37.2	NC_003919	102
Xanthomonas campestris pv. campestris str. 8004*	4273	1589	37.2	NC_007086	103
Xanthomonas oryzae pv. oryzae KACC10331*	4080	1254	30.7	NC_006834	104
Xylella fastidiosa 9a5c*	2766	742	26.8	NC_002488	105
Yersinia pestis CO92*	3885	1146	29.5	NC_003143	106
Yersinia pseudotuberculosis IP 32953*	3901	1150	29.5	NC_006155	107
Zymomonas mobilis subsp. mobilis ZM4	1998	515	25.8	NC_006526	108
Gram positive					
Bacillus anthracis Str. Ames*	5311	936	17.6	NC_003997	109
Bacillus cereus ATCC 14579*	5234	914	17.5	NC_004722	110
Bacillus halodurans C-125	4066	580	14.3	NC_002570	111
Bacillus licheniformis ATCC 14580	4196	745	17.8	NC_006322	112
Bacillus subtilis subsp. subtilis str. 168*	4112	737	17.9	NC_000964	113
Bacillus thuringiensis serovar konkukian	5117	865	16.9	NC_005957	

str. 97-27*					114
Bifidobacterium longum NCC2705	1727	364	21.1	NC_004307	115
Candidatus *Blochmannia floridanus*	583	80	13.7	NC_005061	116
Candidatus *Pelagibacter ubique* HTCC1062	1354	431	31.8	NC_007205	117
Carboxydothermus hydrogenoformans Z-2901	2620	383	14.6	NC_007503	118
Clostridium acetobutylicum ATCC 824	3672	878	23.9	NC_003030	119
*Clostridium perfringens**	2660	442	16.6	NC_003366	120
Clostridium tetani E88*	2373	445	18.8	NC_004557	121
Colwellia psychrerythraea 34H	4910	1237	25.2	NC_003910	122
Corynebacterium diphtheriae NCTC 13129*	2272	466	20.5	NC_002935	123
Corynebacterium efficiens YS-314	2950	574	19.5	NC_004369	124
Corynebacterium glutamicum ATCC 13032	2993	519	17.3	NC_003450	125
Corynebacterium jeikeium K411*	2137	478	22.4	NC_007164	126
Dehalococcoides sp. CBDB1	1458	254	17.4	NC_007356	127
Deinococcus radiodurans R1	2629	649	24.7	NC_001263	128
Enterococcus faecalis V583*	3113	620	19.9	NC_004668	129
Geobacillus kaustophilus HTA426	3498	497	14.2	NC_006510	130
Lactobacillus acidophilus NCFM	1864	497	26.7	NC_006814	131
Lactobacillus johnsonii NCC 533	1821	433	23.8	NC_005362	132
Lactobacillus plantarum WCFS1	3009	704	23.4	NC_004567	133
Lactococcus lactis subsp. cremoris SK11	2321	453	19.5	NC_008527	134
Leifsonia xyli subsp. xyli str. CTCB07*	2030	421	20.7	NC_006087	135
Listeria innocua Clip11262	2968	534	18.0	NC_003212	136
Listeria monocytogenes str. 4b F2365*	2821	491	17.4	NC_002973	137
Mesoplasma florum L1	682	127	18.6	NC_006055	138
Methylococcus capsulatus str. Bath	2960	567	19.2	NC_002977	139
Mycobacterium avium subsp. paratuberculosis K-10*	4350	849	19.5	NC_002944	140
Mycobacterium bovis AF2122/97*	3920	888	22.7	NC_002945	141

Mycobacterium leprae TN*	1605	328	20.4	NC_002677	142
Mycobacterium tuberculosis CDC1551*	4189	1010	24.1	NC_002755	143
Mycoplasma gallisepticum R*	726	242	33.3	NC_004829	144
*Mycoplasma genitalium**	484	142	29.3	NC_000908	145
Mycoplasma hyopneumoniae 232*	691	192	27.8	NC_006360	146
Mycoplasma mobile 163K*	633	131	20.7	NC_006908	147
Mycoplasma mycoides subsp. mycoides SC str. PG1*	1016	366	36.0	NC_005364	148
Mycoplasma penetrans HF-2*	1037	387	37.3	NC_004432	149
*Mycoplasma pneumoniae**	689	252	36.6	NC_000912	150
*Mycoplasma pulmonis**	782	216	27.6	NC_002771	151
Mycoplasma synoviae 53*	672	291	43.3	NC_007294	152
Nitrosococcus oceani ATCC 19707	2974	491	16.5	NC_007484	153
Propionibacterium acnes KPA171202	2297	488	21.2	NC_006085	154
Psychrobacter arcticus 273-4	2120	428	20.2	NC_007204	155
Staphylococcus aureus subsp. aureus MW2*	2632	509	19.3	NC_003923	156
Staphylococcus epidermidis ATCC 12228	2419	442	18.3	NC_004461	157
Staphylococcus haemolyticus JCSC1435	2676	525	19.6	NC_007168	158
Staphylococcus saprophyticus subsp. saprophyticus*	2446	389	15.9	NC_007350	159
Streptococcus agalactiae 2603V/R*	2124	440	20.7	NC_004116	160
Streptococcus mutans UA159	1960	366	18.7	NC_004350	161
Streptococcus pneumoniae R6*	2043	368	18.0	NC_003098	162
Streptococcus pyogenes M1 GAS*	1697	337	19.9	NC_002737	163
Streptococcus thermophilus CNRZ1066	1915	441	23.0	NC_006449	164
Streptomyces avermitilis MA-4680	7575	1899	25.1	NC_003155	165
Streptomyces coelicolor A3(2)	7769	1914	24.6	NC_003888	166
Thermobifida fusca YX	3110	552	17.7	NC_007333	167
Thermosynechococcus elongatus BP-1	2475	361	14.6	NC_004113	168
Tropheryma whipplei TW08/27*	783	168	21.5	NC_004551	169

***** Considered pathogenic to generate the data for Figure 3.

References

1. Barbe, V., *et al*. 2004. Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res.* 32: 5766-5779.
2. Goodner, B., *et al*. 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science*. 294: 2323-2328.
3. Kaneko, T., *et al*. 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res.* 8: 205-213.
4. Brayton, K.A., *et al*. 2005. Complete genome sequencing of *Anaplasma marginale* reveals that the surface is skewed to two superfamilies of outer membrane proteins. *Proc. Natl. Acad. Sci.* USA. 102: 844-849.
5. Deckert, G., *et al*. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus. Nature* 392: 353-358.
6. Rabus, R., *et al.* 2005. The genome sequence of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. *Arch. Microbiol*. 183: 27-36.
7. Takaki, Y., *et al*. 2005. The complete genome sequence of the alkaliphilic *Bacillus clausii* KSM-K16. Direct Submission, National Centre for Biotechnology Information (NCBI).
8. Cerdeno-Tarraga, A.M., *et al*. 2005. Extensive DNA inversions in the *B. fragilis* genome control variable gene expression. *Science* 307: 1463-1465.
9. Xu, J., *et al*. 2003. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* 299: 2074-2076.
10. Alsmark, C.M., *et al*. 2004. The louse-borne human pathogen *Bartonella quintana*

is a genomic derivative of the zoonotic agent *Bartonella henselae*. *Proc. Natl. Acad. Sci.* USA. 101: 9716-9721.

11. Rendulic, S., *et al*. 2004. A predator unmasked: life cycle of *Bdellovibrio bacteriovorus* from a genomic perspective. *Science* 303: 689-692.

12. Parkhill, J., *et al*. 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat. Genet.* 35: 32-40.

13. Fraser, C.M., *et al*. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature*. 390: 580-586.

14. Glockner, G., *et al.* 2004. Comparative analysis of the *Borrelia garinii* genome. *Nucleic Acids Res.* 32: 6038-6046.

15. Kaneko, T., *et al*. 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res.* 9: 189-197.

16. Halling, S.M., *et al*. 2005. Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. *J. Bacteriol* 187: 2715-2726.

17. Paulsen, I.T., *et al*. 2002. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc. Natl. Acad. Sci.* USA. 99: 13148-13153.

18. Shigenobu, S., *et al*. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* 407: 81-86.

19. Nierman, W.C., *et al*. 2004. From the cover: structural flexibility in the *Burkholderia mallei* genome. *Proc. Natl. Acad. Sci.* USA. 101: 14246-14251.

20. Woods, D.E. and Nierman, W.C. 2005. Direct Submission, National Centre for Biotechnology Information (NCBI).

21. Copeland, A., *et al*. 2005a Direct Submission, National Centre for Biotechnology Information (NCBI).

22. Fouts, D.E., *et al*. 2005. Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. *PLOS Biol.* 3: 030-040.

23. Nierman, W.C., *et al*. 2001. Complete genome sequence of *Caulobacter crescentus. Proc. Natl. Acad. Sci.* USA. 98: 4136-4141.

24. Read, T., *et al*. 2000. Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. *Nucleic Acids Res.* 28: 1397-1406.

25. Carlson, J.H., *et al.* 2005. Comparative genomic analysis of *Chlamydia trachomatis* oculotropic and genitotropic strains. *Infect Immun.* 73: 6407-6418.

26. Thomson, N.R., *et al*. 2004. The *Chlamydophila abortus* genome sequence reveals an array of variable proteins that contribute to interspecies variation. *Genome Res.* 15: 629-640.

27. Read, T., *et al*. 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* 423: 81-86.

28. Eisen, J.A., *et al*. 2002. The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc. Natl. Acad. Sci.* USA. 99: 9509-14.

29. Vasconcelos, A.T.R., *et al*. 2003. The complete genome sequence of *Chromobacterium violaceum*, reveals remarkable and exploitable bacterial

adaptability. *Proc. Natl. Acad. Sci.* USA. 100: 11660-11665.

30. Seshadri, R., *et al*. 2003. Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. *Proc. Natl. Acad. Sci.* USA. 100: 5455-5460.

31. Copeland, A., *et al*. (2005b) Direct Submission, National Centre for Biotechnology Information (NCBI).

32. Seshadri, R., *et al*. 2005. Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes. Science* 307: 105-108.

33. Rabus, R., *et al*. 2004. The genome of *Desulfotalea psychrophila*, a sulfate-reducing bacterium from permanently cold arctic sediments. *Environ. Microbiol.* 6: 887-902.

34. Heidelberg, J.F., *et al*. 2004. The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nat. Biotechnol.* 22: 554-559.

35. Copeland, A., *et al*. 2005c. Direct Submission, National Centre for Biotechnology Information (NCBI).

36. Collins, N.E., *et al*. 2005. The genome of the heartwater agent *Ehrlichia ruminantium* contains multiple tandem repeats of actively variable copy number. *Proc. Natl. Acad. Sci.* 102: 838-843.

37. Bell, K.S., *et al*. 2004. Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. atroseptica and characterization of virulence factors. *Proc. Natl. Acad. Sci.* USA. 101: 11105-11110.

38. Welch, R.A., *et al*. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli. Proc. Natl. Acad. Sci.* USA.

99: 17020-17024.

39. Larsson, P., *et al*. 2005. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. *Nat. Genet.* 37: 153-159.

40. Kapatral, V., *et al*. 2002. Genome sequence and analysis of the oral bacterium *Fusobacterium nucleatum* strain ATCC 25586. *J. Bacteriol* 184: 2005-2018.

41. Methe, B.A., *et al*. 2003. Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments. *Science* 302: 1967-1969.

42. Nakamura, Y., *et al*. 2003. Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids (supplement). *DNA Res.* 10: 181-201.

43. Prust, C., *et al*. 2005. Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. *Nat. Biotechnol.* 23: 195-200.

44. Munson, R.S. Jr., *et al*. 2003. Direct Submission, National Centre for Biotechnology Information (NCBI).

45. Harrison, A., *et al*. 2005. Genomic sequence of an otitis media isolate of nontypeable *Haemophilus influenzae*: comparative study with *H. influenzae* serotype d, strain KW20. *J. Bacteriol* 187: 4627-4636.

46. Suerbaum, S., *et al*. 2003. The complete genome sequence of the carcinogenic bacterium *Helicobacter hepaticus*. *Proc. Natl. Acad. Sci.* USA. 100: 7901-7906.

47. Tomb, J.-F., *et al*. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388: 539-547.

48. Hou, S., *et al*. 2004. Genome sequence of the deep-sea {gamma}-proteobacterium *Idiomarina loihiensis* reveals amino acid fermentation as a source of carbon and

energy. *Proc. Natl. Acad. Sci.* USA. 101: 18036-18041.

49. Cazalet, C., *et al*. 2004. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat. Genet.* 36: 1165-1173.

50. Nascimento, A.L., *et al*. 2004. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *J. Bacteriol* 186: 2164-2172.

51. Hong, S.H., *et al*. 2004. The genome sequence of the capnophilic rumen bacterium *Mannheimia succiniciproducens*. *Nat. Biotechnol.* 22: 1275-1281.

52. Kaneko, T., *et al*. 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti. DNA Res.* 7: 331-338.

53. Lewis, L.A., *et al*. 2003. Direct Submission, National Centre for Biotechnology Information (NCBI).

54. Parkhill, J., *et al*. 2000. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature* 404: 502-506.

55. Starkenburg, S.R., *et al.* 2006. Genome sequence of the chemolithoautotrophic nitrite-oxidizing bacterium *Nitrobacter winogradskyi* Nb-255. *Appl. Environ. Microbiol*. 72: 2050-2063.

56. Chain, P., *et al*. 2003. Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea. J. Bacteriol* 185: 2759-2773.

57. Takami, H., *et al.* 2002. Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme

environments. *Nucleic Acids Res.* 30: 3927-3935.

58. May, B.J., *et al*. 2001. Complete genomic sequence of *Pasteurella multocida*, Pm70. *Proc. Natl. Acad. Sci.* USA. 98: 3460-3465.

59. Copeland, A., *et al*. 2005d. Direct Submission, National Centre for Biotechnology Information (NCBI).

60. Copeland, A., *et al*. 2005e. Direct Submission, National Centre for Biotechnology Information (NCBI).

61. Vezzi, A., *et al*. 2004. Direct Submission, National Centre for Biotechnology Information (NCBI).

62. Nelson, K., *et al*. 2003. Complete genome sequence of the oral pathogenic bacterium *Porphyromonas gingivalis* strain W83. *J. Bacteriol* 185: 5591-5601.

63. Medigue, C., *et al*. 2005. Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res.* 15: 1325-1335.

64. Stover, C.K., *et al*. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*. 406: 959-964.

65. Paulsen, I.T., *et al*. 2005. Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nature Biotech*. 23: 873-878.

66. Nelson, K., *et al*. 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.* 4: 799-808.

67. Copeland, A., *et al*. 2005f. Direct Submission, National Centre for Biotechnology Information (NCBI).

68. Salanoubat, M., *et al*. 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature*. 415: 497-502.

69. Copeland, A., *et al*. 2005g. Direct Submission, National Centre for Biotechnology Information (NCBI).

70. Glockner, F.O., *et al*. 2003 Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc. Natl. Acad. Sci.* U. S. A. 100: 8298-303.

71. Larimer, F.W., *et al*. 2004. Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris. Nat. Biotechnol.* 22: 55-61.

72. Ogata, H., *et al*. 2001. Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science*. 293: 2093-2098.

73. Ogata, H., *et al*. 2005. The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. *PLOS Biol.* 3: e248.

74. Andersson, S.G., *et al.* 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature*. 396: 133-140.

75. McLeod, M.P., *et al*. 2004. Complete genome sequence of *Rickettsia typhi* and comparison with sequences of other Rickettsiae. *J. Bacteriol* 186: 5842-5855.

76. Chiu, C.H., *et al*. 2005. The genome sequence of *Salmonella enterica* serovar Choleraesuis, a highly invasive and resistant zoonotic pathogen. *Nucleic Acids Res.* 33: 1690-1698.

77. McClelland, M., *et al*. 2001. Complete genome sequence of *Salmonella enterica*

serovar Typhimurium LT2. *Nature*. 413: 852-856.

78. Heidelberg, J., *et al*. 2002. Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis. Nat. Biotechnol.* 20: 1118-1123.

79. Wei, J., *et al*. 2003. Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect Immun.* 71: 2775-2786.

80. Jiang, Y., *et al*. 2005. The complete sequence and analysis of the large virulence plasmid pSS of *Shigella sonnei*. Plasmid. 54: 149-159.

81. Moran, M.A., *et al*. 2004. Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature*. 432: 910-913.

82. Capela, D., *et al*. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc. Natl. Acad. Sci.* USA. 98: 9877-9882.

83. Ueda, K., *et al*. 2004. Genome sequence of *Symbiobacterium thermophilum*, an uncultivable bacterium that depends on microbial commensalism. *Nucleic Acids Res.* 32: 4937-4944.

84. Copeland, A., *et al*. 2005h. Direct Submission, National Centre for Biotechnology Information (NCBI).

85. Kaneko, T., *et al*. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3: 109-136.

86. Bao, Q., *et al*. 2002. A complete sequence of the *T. tengcongensis* genome. *Genome Res.* 12: 689-700.

87. Nelson, K.E., *et al*. 1999. Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* 399: 323-329.

88. Masui, R., *et al.* 2004. Complete genome sequence of *Thermus thermophilus* HB8. Direct Submission, National Centre for Biotechnology Information (NCBI).

89. Copeland, A., *et al*. 2005i. Direct Submission, National Centre for Biotechnology Information (NCBI).

90. Copeland, A., *et al*. 2005j. Direct Submission, National Centre for Biotechnology Information (NCBI).

91. Copeland, A., *et al*. 2005k. Direct Submission, National Centre for Biotechnology Information (NCBI).

92. Seshadri, R., *et al*. 2004.Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc. Natl. Acad. Sci.* USA 101: 5646-5651.

93. Fraser, C.M., *et al*. 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* 281: 375-378.

94. Glass, J.I., *et al.* 2000. The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature* 407: 757-762.

95. Heidelberg, J.F., *et al*. 2004. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406: 477-483.

96. Ruby, E.G., *et al*. 2005. Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium with pathogenic congeners. *Proc. Natl. Acad. Sci.* USA. 102: 3004-3009.

97. Makino, K., *et al*. 2003. Genome sequence of *Vibrio parahaemolyticus*: a

pathogenic mechanism distinct from that of *V. cholerae*. *Lancet*. 361: 743-749.

98. Chen, C.Y., *et al*. 2003. Comparative genome analysis of *Vibrio vulnificus*, a marine pathogen. *Genome Res.* 13: 2577-2587.

99. Akman, L., *et al.* 2002. Genome sequ ence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nature Genetics* 32: 402-407.

100. Wu, M., *et al*. 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLOS Biol.* 2: E69.

101. Baar, C., *et al*. 2003. Complete genome sequence and analysis of *Wolinella succinogenes*. *Proc. Natl. Acad. Sci.* USA 100: 11690-11695.

102. da Silva, A.C.R., *et al*. 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417: 459-463.

103. Qian, W., *et al*. 2005. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. campestris. *Genome Res.* 15: 757-767.

104. Lee, B.M., *et al*. 2005. The genome sequence of *Xanthomonas oryzae* pathovar oryzae KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.* 33: 577-586.

105. Simpson, A.J.G., *et al*. 2001. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* consortium of the organization for nucleotide sequencing and analysis. *Nature* 406: 151-157.

106. Parkhill, J., *et al*. 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* 413: 523-527.

107. Chain, P.S., *et al*. 2004. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with Y*ersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci.* USA. 101: 13826-13831.

108. Seo, J.S., *et al*. 2005. The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat. Biotechnol.* 23: 63-68.

109. Read, T.D., *et al*. 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* 423:81-86.

110. Ivanova, N., *et al*. 2003. Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* 423: 87-91.

111. Takami, H., *et al*. 2000. Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res.* 28: 4317-4331.

112. Veith, B., *et al*. 2004. The Complete Genome Sequence of *Bacillus licheniformis* DSM13, an Organism with Great Industrial Potential. *J. Mol. Microbiol. Biotechnol.* 7: 204-211.

113. Kunst, F., *et al*. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390: 249-256.

114. Brettin, T.S., *et al*. 2004. Complete genome sequence of *Bacillus thuringiensis* 97-27. Direct Submission, National Centre for Biotechnology Information (NCBI).

115. Schell, M.A., *et al*. 2002. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc. Natl. Acad. Sci.* USA. 99: 14422-14427.

116. Gil, R., *et al*. 2003. The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc. Nat. Acad. Sci.* USA 100: 9388-9393.

117. Giovannoni, S.J., *et al*. 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309: 1242-1245.

118. Wu, M., *et al*. 2005. Life in hot carbon monoxide: the complete genome sequence of *Carboxydothermus hydrogenoformans* Z-29011. *PLoS Genet*. 1: e65.

119. Nolling, J., *et al*. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol* 183: 4823-4838.

120. Myers, G.S., *et al*. 2006. Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. *Genome Res.* 16: 1031-1040.

121. Bruggemann, H., *et al*. 2003. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proc. Natl. Acad. Sci.* USA. 100: 1316-1321.

122. Methe, B.A., *et al*. 2005. The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proc. Natl. Acad. Sci.* USA. 102: 10913-10918.

123. Cerdeno-Tarraga, A.M., *et al*. 2003. The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Res.* 31: 6516-6523.

124. Nishio, Y., *et al*. 2003. Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res.* 13: 1572-1579.

125. Nakagawa, S. 2002. Complete genomic sequence of *Corynebacterium*

glutamicum ATCC 13032. Direct Submission, National Centre for Biotechnology Information (NCBI).

126. Tauch, A., *et al*. 2005. Complete genome sequence and analysis of the multiresistant nosocomial pathogen *Corynebacterium jeikeium* K411, a lipid-requiring bacterium of the human skin flora. *J. Bacteriol* 187: 4671-4682.

127. Kube, M., *et al*. 2005. Genome sequence of the chlorinated compound-respiring bacterium *Dehalococcoides* species strain CBDB1. *Nat. Biotechnol.* 23: 1269-1273.

128. White, O., *et al*. 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* 286: 1571-1577.

129. Paulsen, I., *et al*. 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299: 2071-2074.

130. Takami, H., *et al*. 2004. Thermoadaptation trait revealed by the genome sequence of thermophilic *Geobacillus kaustophilus*. *Nucleic Acids Res.* 32: 6292-6303.

131. Altermann, E., *et al*. 2005. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc. Natl. Acad. Sci.* USA. 102: 3906-3912.

132. Pridmore, *et al*. 2004. The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. *Proc. Natl. Acad. Sci.* USA. 101: 2512-2517.

133. Kleerebezem, M., *et al*. 2004. Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc. Natl. Acad. Sci.* USA. 100: 1990-1995.

134. Makarova, K., *et al*. 2006. Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci.*USA. 103: 15611-6.

135. Vitorello, C.B.M., *et al*. 2004. The genome sequence of the gram-positive sugarcane pathogen *Leifsonia xyli* subsp. xyli. *Mol. Plant Microbe Interact*. 17: 827-836.

136. Glaser, P., *et al*. 2001. Comparative genomics of *Listeria* species. *Science* 294: 849-852.

137. Nelson, K.E., *et al*. 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res.* 32: 2386-2395.

138. Birren, B.W., *et al*. 2004. Direct Submission, National Centre for Biotechnology Information (NCBI).

139. Ward, N., *et al*. 2004. Genomic insights into methanotrophy: the complete genome sequence of *Methylococcus capsulatus* (Bath). *PLOS Biol.* 2: E303.

140. Li, L., *et al*. 2005. The complete genome sequence of *Mycobacterium avium* subspecies paratuberculosis. *Proc. Natl. Acad. Sci.* USA. 102: 12344-12349.

141. Garnier, T., *et al*. 2003. The complete genome sequence of *Mycobacterium bovis. Proc. Natl. Acad. Sci. USA*. 100: 7877-7882.

142. Cole, S.T., *et al*. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409: 1007-1011.

143. Fleischmann, R.D., *et al*. 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* 184: 5479-5490.

144. Papazisi, L., *et al*. 2003. The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain R(low). *Microbiology* 149: 2307-2316.

145. Fraser, C.M., *et al*. 1995. The minimal genome complement of *Mycoplasma genitalium. Science* 270: 397-403.

146. Minion, F.C., *et al*. 2004. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis. *J. Bacteriol* 186: 7123-7133.

147. Jaffe, J.D., *et al*. 2004. The complete genome and proteome of *Mycoplasma mobile*. *Genome Res.* 14: 1447-1461.

148. Westberg, J., *et al*. 2004. The genome sequence of *Mycoplasma mycoides* subsp. mycoides SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). *Genome Res.* 14: 221-227.

149. Sasaki, Y., *et al*. 2002. The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. *Nucleic Acids Res.* 30: 5293-5300.

150. Dandekar, T., *et al*. 2000. Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res.* 28: 3278-3288.

151. Chambaud, I., *et al*. 2001. The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*. *Nucleic Acids Res.* 29: 2145-2153.

152. Vasconcelos, A.T., *et al*. 2005. Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae. J. Bacteriol.* 187: 5568-5577.

153. Copeland, A., *et al*. 2005l. Direct Submission, National Centre for Biotechnology Information (NCBI).

154. Bruggemann, H., *et al*. 2004. The complete genome sequence of

Propionibacterium acnes, a commensal of human skin. *Science* 305: 671-673.

155. Ayala-del-Rio, H.L., *et al*. 2005. Complete sequence of *Psychrobacter arcticum* 273-4. Direct Submission, National Centre for Biotechnology Information (NCBI).

156. Baba, T., *et al*. 2002. Genome and virulence determinants of high virulence community acquired MRSA. *Lancet.* 359: 1819-1827.

157. Zhang, Y.Q., *et al*. 2003. Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). *Mol. Microbiol.* 49: 1577-1593.

158. Takeuchi, F., *et al*. 2005. Whole-genome sequencing of *Staphylococcus haemolyticus* uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal specie. *J. Bacteriol.* 187: 7292-7308.

159. Kuroda, M., *et al*. 2005. Whole genome sequence of *Staphylococcus saprophyticus* reveals the pathogenesis of uncomplicated urinary tract infection. *Proc. Natl. Acad. Sci.* USA. 102: 13272-13277.

160. Tettelin, H., *et al*. 2002. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proc. Natl. Acad. Sci.* USA. 99: 12391-12396.

161. Ajdic, D., *et al*. 2002. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc. Natl. Acad. Sci.* USA. 99: 14434-14439.

162. Hoskins, J.A., *et al*. 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* 183: 5709-5717.

163. Ferretti, J.J., *et al*. 2001. Complete genome sequence of an M1 strain of

Streptococcus pyogenes. *Proc. Natl. Acad. Sci.* USA. 98: 4658-4663.

164. Bolotin, A., *et al*. 2004. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat. Biotechnol.* 22: 1554-1558.

165. Ikeda, H., *et al*. 2003. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* 21: 526-531.

166. Bentley, S.D., *et al*. 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417: 141-147.

167. Copeland, A., *et al*. 2005m. Direct Submission, National Centre for Biotechnology Information (NCBI).

168. Nakamura, Y., *et al*. 2002. Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1 *DNA. Res*. 9: 123-130.

169. Bentley, S.D., *et al*. 2003. Sequencing and analysis of the genome of the Whipple's disease bacterium *Tropheryma whipplei*. *Lancet* 361: 637-644.