

Table 2. Signature Proteins Specific for *Actinobacteria*
Gao, B., Parmanathan, R. and Gupta, R.S. (2006) *Antonie van Leeuwenhoek*, 90(1):69-91.

(a)

Protein	ML0257 [NP_301312]	ML0642 [NP_301530]	ML1009 [NP_301746]	ML1029 [NP_301762]	ML1306 [NP_301939]	ML0760 ⁴ [NP_301589]	ML0804 ⁴ [NP_301614]	ML0857 [NP_301645]
Length	167 aa	479 aa	326 aa	273 aa	274 aa	89 aa	84 aa	250 aa
Possible function	Unknown	Unknown	Unknown	Unknown	Unknown	whiB	whiB	Unknown
<i>Mycobacterium leprae</i>	1e-91 (167)	0 (479)	5e-177 (326)	8e-153 (273)	1e-147 (274)	6e-47 (89)	8e-45 (84)	3e-97 (250)
<i>Mycobacterium tuberculosis</i>	1e-75 (163)	0 (472)	9e-155 (324)	3e-97 (259)	3e-122 (292)	9e-41 (89)	2e-43 (84)	8e-80 (250)
<i>Mycobacterium avium</i>	3e-73 (162)	0 (459)	9e-152 (323)	2e-98 (257)	5e-134 (300)	5e-39 (121)	1e-43 (84)	2e-80 (250)
<i>Mycobacterium bovis</i>	5e-71 (155)	0 (472)	2e-156 (324)	1e-96 (259)	3e-122 (292)	9e-41 (89)	2e-43 (84)	8e-80 (250)
<i>Nocardia farcinica</i>	2e-57 (175)	4e-141 (470)	4e-109 (310)	1e-51 (299)	6e-109 (294)	8e-34 (92)	5e-38 (84)	4e-56 (247)
<i>Corynebacterium glutamicum</i>	4e-48 (184)	3e-93 (482)	3e-66 (319)	4e-38 (287)	7e-18 (319)	1e-29 (104)	2e-36 (86)	3e-36 (260)
<i>Corynebacterium efficiens</i>	2e-48 (184)	3e-92 (483)	6e-66 (319)	2e-35 (274)	2e-17 (319)	2e-29 (110)	4e-37 (86)	2e-36 (282)
<i>Corynebacterium diphtheriae</i>	2e-46 (188)	2e-89 (463)	9e-62 (313)	9e-35 (239)	4e-15 (313)	2e-29 (99)	2e-36 (86)	2e-36 (257)
<i>Corynebacterium jeikeium</i>	2e-44 (188)	1e-78 (494)	9e-65 (359)	7e-30 (233)	3e-16 (359)	1e-28 (121)	9e-35 (88)	6e-45 (259)
<i>Streptomyces avermitilis</i>	1e-47 (174)	3e-98 (483)	1e-60 (312)	6e-20 (254)	4e-78 (334)	1e-30 (87)	4e-32 (85)	6e-29 (233)
<i>Streptomyces coelicolor</i>	7e-48 (186)	6e-100 (487)	2e-58 (312)	8e-18 (250)	4e-77 (333)	1e-30 (87)	3e-32 (85)	3e-29 (234)
<i>Thermobifida fusca</i>	2e-38 (180)	1e-50 (453)	1e-41 (338)	5e-10 (244)	4e-35 (261)	1e-28 (85)	8e-31 (141)	4e-21 (256)
<i>Propionibacterium acnes</i>	3e-40 (188)	4e-92 (427)	1e-67 (313)	4e-15 (240)	2e-69 (281)	5e-30 (95)	1e-30 (90)	4e-24 (242)
<i>Nocardioides sp.</i>	4e-40 (200)	5e-84 (447)	1e-15 (323)	2e-17 (213)	2e-78 (323)	3e-29 (84)	5e-31 (83)	7e-25 (238)
<i>Frankia sp. Ccl3</i>	6e-45 (210)	3e-90 (455)	5e-74 (312)	7e-21 (237)	2e-79 (280)	1e-28 (210)	9e-30 (82)	4e-26 (288)
<i>Frankia sp. EAN1pec</i>	3e-46 (206)	4e-88 (645)	3e-73 (307)	6e-21 (218)	2e-79 (280)	1e-28 (213)	1e-28 (82)	1e-26 (322)
<i>Kineococcus radiotolerance</i>	2e-40 (182)	2e-88 (544)	3e-59 (311)	8e-18 (275)	8e-73 (294)	4e-28 (155)	1e-30 (82)	6e-14 (243)
<i>Brevibacterium linens</i>	4e-37 (176)	8e-63 (481)	2e-11 (265)	8e-15 (287)	9e-30 (265)	3e-29 (103)	1e-28 (82)	5e-08 (243)
<i>Arthrobacter sp.</i>	6e-39 (308)	6e-96 (483)	4e-59 (312)	1e-16 (234)	1e-60 (301)	2e-27 (181)	2e-29 (82)	2e-14 (254)
<i>Leifsonia xyli</i>	8e-36 (170)	4e-72 (447)	9e-54 (309)	2e-13 (167)	1e-50 (298)	5e-27 (88)	4e-26 (82)	2e-16 (232)
<i>Tropheryma whipplei</i> Twist	7e-30 (175)	7e-48 (424)	3e-08 (307)	3e-13 (207)	8e-41 (307)	2e-24 (115)	1e-22 (129)	2e-08 (232)
<i>Tropheryma whipplei</i> TW08/27	7e-30 (162)	2e-48 (424)	3e-08 (296)	3e-13 (191)	8e-41 (296)	2e-24 (102)	1e-22 (77)	2e-08 (232)
<i>Bifidobacterium longum</i>	1e-28 (188)	1e-33 (581)	9e-15 (285)	5e-17 (284)	1e-31 (285)	3e-27 (99)	4e-24 (92)	0.39 (260)
<i>Rubrobacter xylanophilus</i>	1e-08 (145)	4e-07 (340)	1e-24 (299)	0.071 (233)	7e-51 (299)	---	---	---
Non-Actinobacteria	See note 1	0.007 (391)	5.0 (863)	See note 2	See note 3	2.6 (377)	7.4 (520)	2.8 (399)
		<i>Chloroflexus aurantiacus</i>	Human enterovirus			<i>Rattus norvegicus</i>	<i>Rhodopirellula baltica</i>	<i>Novosphingobium aromaticivorans</i>

Protein	ML0869 [NP_301656]	ML1016 [NP_301752]	ML1026 [NP_301759]	ML2073 [NP_302382]	ML2137 [NP_302410]	ML2204 [NP_302445]	ML0013 [NP_301140]
Length	124 aa	107 aa	100 aa	231 aa	251 aa	62 aa	93 aa
Possible function	Unknown	Unknown	Unknown	MerR	Unknown	Unknown	Unknown
<i>Mycobacterium leprae</i>	1e-64 (124)	1e-58 (107)	5e-51 (100)	2e-128 (231)	1e-141 (251)	3e-29 (62)	2e-48 (93)
<i>Mycobacterium tuberculosis</i>	3e-50 (236)	3e-36 (82)	8e-49 (100)	7e-109 (225)	5e-109 (253)	3e-18 (60)	6e-45 (93)
<i>Mycobacterium avium</i>	2e-53 (230)	7e-35 (79)	1e-48 (100)	3e-102 (225)	6e-106 (254)	7e-17 (61)	3e-45 (93)
<i>Mycobacterium bovis</i>	3e-50 (236)	3e-36 (82)	8e-49 (100)	7e-109 (225)	5e-109 (253)	3e-18 (60)	6e-45 (93)
<i>Nocardia farcinica</i>	2e-29 (231)	1e-24 (81)	2e-41 (100)	2e-74 (185)	8e-59 (324)	5e-16 (68)	5e-24 (87)
<i>Corynebacterium glutamicum</i>	1e-15 (266)	8e-21 (79)	9e-22 (97)	1e-59 (191)	1e-29 (311)	1e-08 (71)	7e-13 (90)
<i>Corynebacterium efficiens</i>	5e-19 (273)	1e-20 (300)	2e-18 (106)	1e-61 (207)	3e-31 (350)	4e-09 (70)	3e-11 (90)
<i>Corynebacterium diphtheriae</i>	9e-22 (224)	6e-21 (79)	5e-22 (97)	2e-58 (186)	3e-30 (349)	7e-07 (68)	3e-12 (89)
<i>Corynebacterium jeikeium</i>	5e-22 (242)	1e-19 (80)	4e-23 (99)	4e-55 (170)	4e-17 (387)	3e-07 (85)	3e-12 (90)
<i>Streptomyces avermitilis</i>	1e-14 (208)	1e-06 (98)	7e-33 (98)	5e-54 (211)	1e-25 (348)	1e-07 (83)	2e-05 (84)
<i>Streptomyces coelicolor</i>	3e-15 (251)	1e-06 (97)	7e-33 (98)	3e-55 (228)	1e-24 (352)	5e-07 (84)	3e-05 (84)
<i>Thermobifida fusca</i>	8e-13 (214)	9e-06 (98)	1e-28 (98)	3e-50 (264)	3e-28 (330)	6e-07 (84)	---
<i>Propionibacterium acnes</i>	0.060 (251)	1e-08 (80)	1e-30 (99)	2e-46 (195)	1e-13 (359)	8e-07 (79)	0.002 (93)
<i>Nocardioides sp.</i>	2e-13 (194)	5e-09 (107)	4e-28 (96)	2e-54 (198)	1e-20 (318)	2e-06 (62)	---
<i>Frankia sp. Ccl3</i>	6e-12 (209)	---	3e-23 (97)	8e-55 (204)	4e-22 (593)	1e-08 (68)	3e-06 (87)
<i>Frankia sp. EAN1pec</i>	9e-14 (209)	0.003 (74)	1e-22 (97)	1e-54 (208)	4e-22 (755)	2e-05 (73)	2e-05 (88)
<i>Kineococcus radiotolerance</i>	2e-15 (225)	2e-07 (107)	4e-31 (99)	4e-55 (199)	4e-28 (?)	3e-07 (85)	8e-04 (110)
<i>Brevibacterium linens</i>	2e-15 (148)	7e-08 (95)	6e-33 (?)	2e-54 (172)	2e-18 (357)	6e-08 (83)	3e-07 (274)
<i>Arthrobacter sp.</i>	3e-10 (200)	3e-09 (118)	7e-32 (99)	3e-55 (198)	7e-17 (492)	8e-07 (106)	4e-07 (84)
<i>Leifsonia xyli</i>	4e-13 (213)	2e-05 (81)	3e-21 (99)	2e-44 (200)	2e-11 (365)	3e-04 (73)	4e-08 (86)
<i>Tropheryma whipplei</i> Twist	5e-06 (173)	1e-02 (81)	3e-16 (92)	2e-40 (158)	6e-05 (320)	2e-03 (41)	6e-04 (69)
<i>Tropheryma whipplei</i> TW08/27	5e-06 (173)	1e-02 (81)	3e-16 (92)	2e-40 (158)	6e-05 (307)	2e-03 (41)	6e-04 (69)
<i>Bifidobacterium longum</i>	3e-06 (171)	7e-09 (97)	1e-19 (120)	6e-33 (210)	2e-03 (352)	3e-05 (129)	4e-07 (156)
<i>Rubrobacter xylanophilus</i>	---	---	---	---	---	---	---
Non-Actinobacteria	1.9 (221)	7.4 (230)	0.009 (1306)	2e-04 (168)	0.73 (637)	7.4 (664)	1.5 (951)
	<i>Bacillus cereus</i>	<i>Cytophaga hutchinsonii</i>	<i>Arabidopsis thaliana</i>	<i>Nostoc sp.</i> 42/76 (55%)	<i>Drosophila melanogaster</i>	<i>Prochlorococcus usmarinus</i>	<i>Dechloromonas aromatica</i>

Table 2. (cont.)

(b)

Protein	ML0007 [NP_301135]	ML0580 [NP_301492]	ML0921 [NP_301704]	ML1439 [NP_302017]	ML1610 [NP_302109]	ML2207 [NP_302448]	ML1439 [NP_302017]	ML0256 [NP_301311]	ML0775 [NP_301599]	Proteins showing similar specificity
Length	303 aa	265 aa	96 aa	111 aa	101 aa	131 aa	111 aa	227 aa	589 aa	
Possible function	Unknown	OpcA	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	LpqB	
<i>Mycobacterium leprae</i>	7e-161 (303)	2e-147 (265)	2e-35 (96)	1e-45 (111)	2e-52 (101)	3e-69 (131)	1e-45 (111)	1e-94 (227)	0 (589)	
<i>Mycobacterium tuberculosis</i>	3e-68 (304)	1e-87 (303)	2e-31 (96)	2e-44 (111)	5e-52 (101)	4e-50 (129)	2e-44 (111)	1e-60 (228)	0 (587)	ML0761 [NP_301590]
<i>Mycobacterium avium</i>	8e-72 (283)	2e-80 (303)	8e-31 (96)	6e-44 (111)	7e-51 (101)	9e-43 (131)	6e-44 (111)	2e-56 (225)	0 (585)	
<i>Mycobacterium bovis</i>	3e-68 (304)	1e-87 (303)	2e-31 (96)	2e-44 (111)	5e-52 (101)	4e-50 (129)	2e-44 (111)	1e-60 (228)	0 (583)	ML0814 [NP_301620]
<i>Nocardia farcinica</i>	6e-25 (389)	1e-53 (302)	3e-21 (101)	3e-37 (111)	4e-48 (101)	1e-32 (128)	3e-37 (111)	3e-14 (223)	7e-101 (604)	
<i>Corynebacterium glutamicum</i>	3e-05 (114)	7e-29 (319)	7e-11 (95)	1e-30 (132)	2e-38 (101)	3e-32 (125)	1e-30 (132)	4e-07 (180)	3e-50 (568)	
<i>Corynebacterium efficiens</i>	3e-05 (114)	3e-29 (321)	2e-10 (130)	3e-27 (120)	1e-38 (101)	2e-33 (142)	3e-27 (120)	3e-06 (150)	7e-49 (563)	ML1649 [NP_302131]
<i>Corynebacterium diphtheriae</i>	2e-08 (114)	7e-27 (319)	2e-08 (95)	4e-29 (129)	9e-39 (101)	2e-31 (131)	4e-29 (129)	4e-06 (177)	8e-45 (581)	
<i>Corynebacterium jeikeium</i>	2e-12 (217)	4e-31 (358)	2e-11 (96)	1e-31 (125)	3e-41 (101)	5e-18 (126)	1e-31 (125)	1e-08 (236)	1e-60 (583)	
<i>Streptomyces avermitilis</i>	4e-10 (204)	8e-28 (311)	5e-12 (94)	2e-19 (124)	4e-39 (102)	6e-21 (265)	2e-19 (124)	5e-08 (164)	2e-11 (610)	ML1666 [NP_302145]
<i>Streptomyces coelicolor</i>	2e-10 (185)	3e-27 (351)	5e-12 (94)	2e-19 (124)	4e-39 (102)	6e-21 (202)	2e-19 (124)	3e-09 (174)	2e-07 (615)	
<i>Thermobifida fusca</i>	2e-08 (230)	2e-24 (308)	4e-12 (101)	3e-18 (129)	3e-38 (107)	1e-20 (169)	3e-18 (129)	1e-07 (179)	9e-17 (626)	ML2142 [NP_302413]
<i>Propionibacterium acnes</i>	1e-11 (210)	1e-19 (310)	3e-05 (96)	7e-13 (110)	5e-32 (103)	5e-10 (280)	7e-13 (110)	5e-06 (217)	5e-10 (591)	
<i>Nocardioides</i> sp.	4e-14 (172)	6e-23 (303)	2e-12 (96)	1e-15 (128)	4e-35 (102)	2e-17 (119)	1e-15 (128)	9e-05 (184)	1e-19 (582)	
<i>Frankia</i> sp. Ccl3	8e-09 (249)	3e-26 (370)	9e-09 (88)	0.001 (198)	7e-38 (98)	9e-19 (270)	0.001 (198)	---	---	
<i>Frankia</i> sp. EAN1pec	5e-13 (645)	6e-23 (340)	1e-07 (84)	7e-24 (214)	4e-38 (98)	2e-19 (336)	7e-24 (214)	3e-09 (134)	---	
<i>Kineococcus radiotolerance</i>	2e-11 (171)	1e-37 (406)	5e-12 (101)	4e-14 (115)	1e-38 (108)	1e-21 (130)	4e-14 (115)	1e-06 (212)	---	
<i>Brevibacterium linens</i>	9e-10 (153)	---	4e-11 (99)	1e-10 (113)	9e-33 (105)	1e-22 (133)	1e-10 (113)	---	1e-10 (562)	
<i>Arthrobacter</i> sp.	5e-10 (215)	1e-30 (313)	9e-12 (95)	5e-15 (115)	3e-37 (113)	3e-21 (137)	5e-15 (115)	6e-06 (229)	1e-15 (573)	
<i>Leifsonia xyli</i>	1e-05 (137)	5e-30 (320)	2e-10 (98)	6e-09 (123)	2e-30 (107)	9e-22 (118)	6e-09 (123)	3e-04 (177)	8e-18 (557)	
<i>Tropheryma whipplei</i> Twist	---	---	---	---	---	---	---	---	---	
<i>Tropheryma whipplei</i> TW08/27	---	---	---	---	---	---	---	---	---	
<i>Bifidobacterium longum</i>	7e-10 (188)	2e-26 (341)	2e-06 (100)	3e-10 (115)	1e-33 (104)	7e-08 (177)	3e-10 (115)	1e-07 (203)	1e-10 (576)	
Non-Actinobacteria	0.004 (2528)	0.002 (384)	2.5 (88)	2.2 (265)	0.059 (344)	1.5 (425)	2.2 (265)	0.60 (407)	1.1 (969)	For details, see supplemental table 1(a)
	<i>Magnaporthe grisea</i>	<i>Chloroflexus aurantiacus</i>	<i>Wolbachia endosymbiont</i>	<i>Leptospira interrogans</i>	<i>Gallus gallus</i>	<i>Thermoanaerobacter tengcongensis</i>	<i>Leptospira interrogans</i>	<i>Arabidopsis thaliana</i>	<i>Bacillus anthracis</i>	

These proteins were identified by BLASTP searches as detailed in the Methods section. The top line is the protein ID number in genome of *M. leprae* TN (ML), which was used as probe to perform the blast search. Accession numbers for these proteins are shown in square brackets. The second line and the third line describe the sequence length and possible function of each query protein. The left column lists the actinobacterial strains that have been completely sequenced or draft assembled. The expected (E) values for various actinobacterial species as well as the first non-actinobacterial species in the BLAST results are shown here. The values in brackets after the E-values represent the length of the hit protein. Proteins not found in a genome are indicated with dash (---). * indicates that the genome is incompletely sequenced so it is possible that the protein is present in the genome but not identified at the moment.

Note 1. The first 3 non-actinobacterial hits to ML0257 correspond to *Thermotoga maritima* MSB8 with E-value of 4e-14 (170 aa) [NP_228884]; *T. neapolitana* with E-value of 1e-12 (150 aa) [CAA07517]; and *Aquifex aeolicus* VF5 with E-value of 6e-04 (147 aa) [NP_214081]. The next non-actinobacterial hit is *Trypanosoma cruzi* with E-value of 0.035 (271 aa).

Note 2. The first non-actinobacterial hit for ML1029 is found in *M. magnetotacticum* with E-value of 2e-15 (235 aa) [ZP_00049023]; the next non-actinobacterial hit is *Microbulbifer degradans* with E-value of 0.30 (1245 aa).

Note 3. A low scoring homologue to ML1306 is also found in *Dehalococcoides ethenogenes* with E-value of 8e-09 (276 aa) [YP_181269]; the next non-actinobacterial hit is *Archaeoglobus fulgidus* with E-value of 0.17 (239 aa).

⁴Note 4. These 2 proteins are paralogous gene product recognized as WhiB. All sequenced actinobacterial species contain several copies of *whiB* gene. Some phages also have homologous gene as observed by their low E-values. These phage proteins include: protein [AAD17616] from Mycobacteriophage TM4 (76 aa); protein [NP_958255] from Bacteriophage VWB (81 aa); and protein [AAN01709] from Mycobacteriophage CJW1 (86 aa);