

Table 1: Sequences selected for the final dataset of PR-1

Accession	Entry name	Length	Weight	Species	Gene name	Protein name	PE	Date
O65157	B65157_CAPAN	179 AA	20094 Da	<i>Capsicum annuum</i>	BPRI	Basic PR-1 protein	2	05-JUL-2004
B2LW68	B2LW68_SOLLIC	179 AA	20423 Da	<i>Solanum lycopersicum</i>		PR1 protein	2	10-JUN-2008
P1L670	PRB1.TOBAC	177 AA	19729 Da	<i>Nicotiana tabacum</i>		Basic form of pathogenesis-related protein 1	3	01-OCT-1989
Q08697	PR1A.SOLLIC	175 AA	19660 Da	<i>Nicotiana glauca</i>		Pathogenesis-related protein 1A1	2	01-OCT-1996
Q00MX6	Q00MX6_MALDO	161 AA	17122 Da	<i>Malus domestica</i>		Pathogenesis-related protein 1a	2	14-NOV-2006
Q39188	Q39188_ARATH	169 AA	19330 Da	<i>Arabidopsis thaliana</i>	PR1	Pathogenesis-related protein 1	3	01-NOV-1996
B2CZ52	B2CZ52_CUCME	151 AA	16950 Da	<i>Cucumis melo</i>	Cuc m 3	Pathogen-related protein 1	2	20-MAY-2008
P08299	PR1A.TOBAC	168 AA	18574 Da	<i>Nicotiana tabacum</i>		Pathogenesis-related protein 1A	1	01-AUG-1988
Q40557	Q40557_TOBAC	168 AA	18590 Da	<i>Nicotiana tabacum</i>		PR1a protein	2	01-NOV-1996
P07053	PR1B.TOBAC	168 AA	18500 Da	<i>Nicotiana tabacum</i>		Pathogenesis-related protein 1B	2	01-APR-1988
P09042	PR1C.TOBAC	168 AA	18583 Da	<i>Nicotiana tabacum</i>		Pathogenesis-related protein 1C	2	01-APR-1990
Q40397	Q40397_NICGU	168 AA	18547 Da	<i>Nicotiana glutinosa</i>		Pathogenesis-related protein-1	2	01-NOV-1996
O82086	O82086_MAIZE	163 AA	17224 Da	<i>Zea mays</i>	PR-1	Pathogenesis-related protein-1	2	01-NOV-1998
B3TLW3	B3TLW3_ELAGV	162 AA	17599 Da	<i>Elaeis guineensis</i>		Pathogenesis-related protein	2	02-SEP-2008
C3UZE5	C3UZE5_WHEAT	164 AA	17635 Da	<i>Triticum aestivum</i>		Pathogenesis-related protein 1.1	2	16-JUN-2009
C3UZE4	C3UZE4_WHEAT	164 AA	17620 Da	<i>Triticum aestivum</i>		Pathogenesis-related protein 1.1	2	16-JUN-2009
O82714	O82714_WHEAT	164 AA	17651 Da	<i>Triticum aestivum</i>	PR-1.1	Pathogenesis-related protein 1.1	2	01-NOV-1998
Q3S413	Q3S413_TRIMO	164 AA	17645 Da	<i>Triticum monococcum</i>		Pathogenesis-related 1b	2	11-OCT-2005
Q43489	Q43489_HORVU	164 AA	17440 Da	<i>Hordeum vulgare</i>	Hv-1a	PR-1a pathogenesis related protein (Hv-1a)	2	01-NOV-1996
P35793	PR13_HORVU	164 AA	17697 Da	<i>Hordeum vulgare</i>	PRB1-3	Pathogenesis-related protein PRB1-3	2	01-JUN-1994
Q05968	PR1_HORVU	164 AA	17683 Da	<i>Hordeum vulgare</i>		Pathogenesis-related protein 1	2	01-JUN-1994
P35792	PR12_HORVU	164 AA	17679 Da	<i>Hordeum vulgare</i>	PRB1-2	Pathogenesis-related protein PRB1-2	2	01-JUN-1994
Q94F73	Q94F73_WHEAT	164 AA	17537 Da	<i>Triticum aestivum</i>		Pathogenesis-related protein 1	2	01-DEC-2001
C3UZE2	C3UZE2_WHEAT	164 AA	17511 Da	<i>Triticum aestivum</i>		Pathogenesis-related protein 1	2	16-JUN-2009
C3UZE3	C3UZE3_WHEAT	164 AA	17496 Da	<i>Triticum aestivum</i>		Pathogenesis-related protein 1	2	16-JUN-2009
A0N0C3	A0N0C3_ORYSI	167 AA	18184 Da	<i>Oryza sativa</i>	Pr1c	Pathogenesis-related protein PR1c	2	12-DEC-2006
Q00008	PRMS_MAIZE	167 AA	18390 Da	<i>Zea mays</i>	PRMS	Pathogenesis-related protein PRMS	2	01-DEC-1992
O82715	O82715_WHEAT	173 AA	18800 Da	<i>Triticum aestivum</i>	PR-1.1.2	Pathogenesis-related protein 1.2	2	01-NOV-1998
Q41359	PR1_SAMNI	167 AA	18411 Da	<i>Sambucus nigra</i>		Pathogenesis-related protein PR-1 type	2	01-NOV-1996
B5QTD3	B5QTD3_BETVU	168 AA	18083 Da	<i>Beta vulgaris</i>	pr1a	Pathogenesis-related protein 1a	2	04-NOV-2008
A0N0C2	A0N0C2_ORYSI	164 AA	17462 Da	<i>Oryza sativa</i>	Pr1b	Pathogenesis-related protein PR1b	2	12-DEC-2006
A0N0C1	A0N0C1_ORYSI	168 AA	17534 Da	<i>Oryza sativa</i>	Pr1a	Pathogenesis-related protein PR1a	2	12-DEC-2006
Q7FP72	Q7FP72_ORISA	168 AA	17560 Da	<i>Oryza sativa</i>	Pr1a	PR1a protein	2	05-JUL-2004
Q8LLU7	Q8LLU7_ORISA	165 AA	17016 Da	<i>Oryza sativa</i>		Pathogenesis-related protein 1	2	01-OCT-2002
Q3S414	Q3S414_TRIMO	167 AA	17803 Da	<i>Triticum monococcum</i>		Pathogenesis-related 1a	2	11-OCT-2005
Q0KIX9	Q0KIX9_EUTWA	161 AA	17668 Da	<i>Eutrema wasabi</i>	PR-1	Pathogenesis-related protein 1	2	03-OCT-2006
P33154	PR1_ARATH	161 AA	17677 Da	<i>Arabidopsis thaliana</i>	At2g14610	Pathogenesis-related protein 1	1	01-OCT-1993
Q2HZ50	Q2HZ50_BRAJU	161 AA	17532 Da	<i>Brassica juncea</i>		PR1	2	07-MAR-2006
Q6IUZ8	Q6IUZ8_BRACM	161 AA	17420 Da	<i>Brassica campestris</i>	PR-1	Pathogenesis-related protein 1	2	05-JUL-2004
BLN8M3	BLN8M3_BRAOG	162 AA	17772 Da	<i>Brassica oleracea</i>		Pathogenesis-related protein	2	29-APR-2008

Table 2: Sequences selected for the final dataset of PR01

Accession	Entry name	Length	Weight	Species	Gene name	Protein name	PE	Date
Q8L688	Q8L688.SOLN	159 AA	17314 Da	<i>Solanum phureja</i>	pr1b1	Pathogenesis related protein isoform b1	2	01-OCT-2002
Q9SC15	Q9SC15.SOLTU	159 AA	17323 Da	<i>Solanum tuberosum</i>	pr1-1	Pathogenesis related protein PR-1	2	01-MAY-2000
Q941G6	Q941G6.SOLTU	159 AA	17337 Da	<i>Solanum tuberosum</i>	pr1b	Pathogenesis-related protein 1b	2	01-DEC-2001
Q8L687	Q8L687.SOLN	159 AA	17623 Da	<i>Solanum phureja</i>	pr1b2	Pathogenesis related protein isoform b2	2	01-OCT-2002
P04284	PR06.SOLLC	159 AA	17520 Da	<i>Solanum lycopersicum</i>	PR1B1	Pathogenesis-related leaf protein 6	1	01-JUN-1994
Q04108	PR04.SOLLC	159 AA	17439 Da	<i>Solanum lycopersicum</i>		Pathogenesis-related leaf protein 4	2	01-JUN-1994
O24026	O24026.SOLLC	160 AA	17685 Da	<i>Solanum lycopersicum</i>	PR1d	Pathogenesis-related protein	3	01-JAN-1998
Q2VT59	Q2VT59.CAPAN	158 AA	17396 Da	<i>Capsicum annuum</i>		Cytoplasmic small heat shock protein class I	2	10-JAN-2006
Q6WHB9	Q6WHB9.CAPFR	158 AA	17369 Da	<i>Capsicum frutescens</i>		Cytoplasmic small heat shock protein class I	2	05-JUL-2004
O81889	O81889.ARATH	172 AA	18830 Da	<i>Arabidopsis thaliana</i>	Tl6L1.220	Pathogenesis-related protein-like	2	01-NOV-1998
O9LJM5	O9LJM5.ARATH	161 AA	17809 Da	<i>Arabidopsis thaliana</i>	At3g19690		2	01-OCT-2000
Q9XFB4	Q9XFB4.SOYBN	174 AA	18740 Da	<i>Glycine max</i>	PR1a		2	01-NOV-1999
Q40374	PR1.MEDTR	173 AA	19760 Da	<i>Medicago truncatula</i>	PR-1	Pathogenesis-related protein PR-1	2	01-NOV-1996
B6UG61	B6UG61.MAIZE	175 AA	18721 Da	<i>Zea mays</i>		Pathogenesis-related protein PR-1	2	16-DEC-2008
Q9M0C8	Q9M0C8.ARATH	161 AA	18028 Da	<i>Arabidopsis thaliana</i>	At4g30320	PR-1-like protein	2	01-OCT-2000
Q9SV22	Q9SV22.ARATH	185 AA	20931 Da	<i>Arabidopsis thaliana</i>	At4g31470	Pathogenesis-related protein homolog	2	01-MAY-2000
O6ID87	O6ID87.ARATH	241 AA	27396 Da	<i>Arabidopsis thaliana</i>	At1g01310	At1g01310	2	05-JUL-2004
Q2QLQ1	Q2QLQ1.ORYSJ	418 AA	43346 Da	<i>Oryza sativa</i>	Os12g0633400	Os12g0633400 protein	2	24-JAN-2006
Q39186	Q39186.ARATH	176 AA	19701 Da	<i>Arabidopsis thaliana</i>	At2g19990	At2g19990	2	01-NOV-1996
Q0DWY7	Q0DWY7.ORYSJ	199 AA	21567 Da	<i>Oryza sativa</i>	Os02g0786900	Os02g0786900 protein	2	17-OCT-2006
Q0DFE6	Q0DFE6.ORYSJ	332 AA	36324 Da	<i>Oryza sativa</i>	Os05g0595200	Os05g0595200 protein	2	17-OCT-2006
A6GVD5	A6GVD5.ARTVU	162 AA	18171 Da	<i>Artemisia vulgaris</i>	art v 2	Art v 2 allergen	2	24-JUL-2007
Q8S9B6	Q8S9B6.VOLCA	415 AA	43869 Da	<i>Volvox carteri f. nagariensis</i>	lsg1	PR-1 like protein	2	01-JUN-2002