

Table 1: Sequences selected for the final dataset of PR-12

Accession	Entry name	Length	Weight	Species	Gene name	Protein name	PE	Date
P69241	DEF1.RAPSA	80 AA	8734 Da	<i>Raphanus sativus</i>	AFP1	Defensin-like protein 1	1	15-FEB-2005
Q94IN7	Q94IN7.BRAOL	80 AA	8740 Da	<i>Brassica oleracea</i>	def	Defensin	3	01-DEC-2001
P30230	DEF2.RAPSA	80 AA	8740 Da	<i>Raphanus sativus</i>	AFP2	Defensin-like protein 2	1	01-OCT-1996
P30224	DEF13.ARAH	80 AA	8875 Da	<i>Arabidopsis thaliana</i>	PDF1.1	Defensin-like protein 13	1	15-DEC-1998
Q39313	DEF3.BRANA	79 AA	8555 Da	<i>Brassica napus</i>	AFP3	Defensin-like protein 3	3	01-JAN-1998
Q9FS38	Q9FS38.EUTWA	80 AA	8761 Da	<i>Eutrema wasabi</i>		Gamma-thionin1	3	01-MAR-2001
Q5KU48	Q5KU48.EUTWA	80 AA	8784 Da	<i>Eutrema wasabi</i>		Gamma-thionin2	3	15-FEB-2005
Q24331	DEF4.RAPSA	80 AA	8873 Da	<i>Raphanus sativus</i>	AFP4	Defensin-like protein 4	3	01-JAN-1998
Q9FI23	DEF16.ARAH	80 AA	8518 Da	<i>Arabidopsis thaliana</i>	PDF1.2A	Defensin-like protein 16	2	01-MAR-2001
O80995	DEF14.ARAH	80 AA	8580 Da	<i>Arabidopsis thaliana</i>	PDF1.3	Defensin-like protein 14	2	01-NOV-1998
Q8H6Q0	DEF2.PETHY	101 AA	11049 Da	<i>Petunia hybrida</i>	D2	Floral defensin-like protein 2	1	01-MAR-2003
Q8GTMO	DEF.NICAL	105 AA	11722 Da	<i>Nicotiana glauca</i>	D1	Flower-specific defensin	1	01-MAR-2003
P32026	DEF.TOBAC	105 AA	11750 Da	<i>Nicotiana tabacum</i>	FST	Defensin-like protein	1	01-JUL-1993
O24115	DEF1.NICPA	106 AA	11722 Da	<i>Nicotiana glauca</i>	THI01	Defensin-like protein 1	2	01-JAN-1998
Q43413	DEF1.CAPAN	75 AA	8118 Da	<i>Capsicum annuum</i>		Defensin J1-1	1	01-NOV-1996
Q01784	DF39.PEA	74 AA	8255 Da	<i>Pisum sativum</i>	PI39	Defensin-like protein 39	2	01-JUN-1994
P81929	DEF1.PEA	46 AA	5209 Da	<i>Pisum sativum</i>		Defensin-1	1	05-MAR-2002
O8H6Q1	DEF1.PETHY	103 AA	11361 Da	<i>Petunia hybrida</i>	D1	Floral defensin-like protein 1	1	01-MAR-2003
Q8W4V6	Q8W4V6.CAPAN	78 AA	8689 Da	<i>Capsicum annuum</i>	DEF1	Defensin	3	01-MAR-2002
B2CM18	B2CM18.SOLCH	76 AA	8382 Da	<i>Solanum chacoense</i>		Gamma-thionin/defensin	3	20-MAY-2008
O65740	DEF2.CAPAN	74 AA	8249 Da	<i>Capsicum annuum</i>		Defensin J1-2	1	01-AUG-1998
Q40901	DEF.PETIN	78 AA	8799 Da	<i>Petunia integrifolia</i>		Defensin-like protein	2	01-NOV-1996
Q39182	DEF02.ARAH	77 AA	8524 Da	<i>Arabidopsis thaliana</i>	PDF2.2	Defensin-like protein 2	2	01-NOV-1996
Q9ZUL7	DEF01.ARAH	77 AA	8544 Da	<i>Arabidopsis thaliana</i>	PDF2.3	Defensin-like protein 1	2	01-MAY-1999
Q41914	DEF04.ARAH	77 AA	8578 Da	<i>Arabidopsis thaliana</i>	PDF2.1	Defensin-like protein 4	2	01-DEC-2000
C1K3M7	C1K3M7.VIGUN	78 AA	8655 Da	<i>Vigna unguiculata</i>	PDEF	Defensin	3	26-MAY-2009
P82659	DFS02.HELAN	78 AA	8627 Da	<i>Helianthus annuus</i>	SD2	Defensin SD2	2	01-OCT-2000
A3FPF2	DEF.NELNU	77 AA	8361 Da	<i>Nelumbo nucifera</i>		Defensin-like protein	3	20-MAR-2007
Q9C947	DEF05.ARAH	76 AA	8290 Da	<i>Arabidopsis thaliana</i>	PDF2.4	Defensin-like protein 5	2	13-APR-2004
Q01783	DF230.PEA	72 AA	7986 Da	<i>Pisum sativum</i>	PI230	Defensin-like protein 230	2	01-JUN-1994
Q8W434	Q8W434.9FABA	75 AA	8631 Da	<i>Vigna radiata</i>	PDF1	PDF1	1	01-MAR-2002
P18646	DEF.VIGUN	75 AA	8523 Da	<i>Vigna unguiculata</i>		Defensin-like protein	3	01-NOV-1990
Q40128	Q40128.SOLLC	105 AA	11914 Da	<i>Solanum lycopersicum</i>		Flower-specific gamma-thionin-like, acidic protein	3	01-NOV-1996
B3F051	DEF.LENCC	74 AA	8401 Da	<i>Lens culinaris</i>		Defensin Lc-def	1	22-JUL-2008
P82784	DEF07.ARAH	82 AA	9304 Da	<i>Arabidopsis thaliana</i>	LCR75	Defensin-like protein 7	2	28-JUL-2009
Q19JAL	Q19JAL.TRIKI	76 AA	8809 Da	<i>Trichosanthes kirilowii</i>		Defensin	3	11-JUL-2006
Q9ZUL8	DEF10.ARAH	73 AA	7718 Da	<i>Arabidopsis thaliana</i>	PDF2.6	Defensin-like protein 10	2	01-MAY-1999
Q670N7	Q670N7.GINBI	80 AA	8925 Da	<i>Ginkgo biloba</i>		Defensin	3	11-OCT-2004
P82782	DEF08.ARAH	80 AA	8986 Da	<i>Arabidopsis thaliana</i>	LCR73	Defensin-like protein 8	3	29-MAR-2004
P82787	DEF19.ARAH	78 AA	8840 Da	<i>Arabidopsis thaliana</i>	PDF1.4	Defensin-like protein 19	2	04-JAN-2005

Table 2: Sequences selected for the final dataset of PR12

Accession	Entry name	Length	Weight	Species	Gene name	Protein name	PE	Date
Q9FFP8	DEF06.ARATH	73 AA	8388 Da	<i>Arabidopsis thaliana</i>	PDF2.5	Defensin-like protein 6	2	01-MAR-2001
B2CNV2	B2CNV2.SACOF	80 AA	8605 Da	<i>Saccharum officinarum</i>	PDEF	Defensin	3	20-MAY-2008
Q8L698	Q8L698.WHEAT	82 AA	8973 Da	<i>Triticum aestivum</i>	Tadi	Defensin	3	01-OCT-2002
Q9FZ31	DEF18.ARATH	80 AA	9139 Da	<i>Arabidopsis thaliana</i>	PDF1.5	Defensin-like protein 18	2	01-MAR-2001
Q84ZX5	ALL1.ARTVU	132 AA	13404 Da	<i>Artemisia vulgaris</i>		Major pollen allergen Art v 1	1	01-JUN-2003
B6SQK6	B6SQK6.MAIZE	129 AA	13862 Da	<i>Zea mays</i>		Gamma-thionins family protein	2	16-DEC-2008
Q9SEMI	Q9SEMI.CAFAN	84 AA	9470 Da	<i>Capsicum annuum</i>	thion1	Gamma-thionin 1	3	01-MAY-2000