Table S1.
Sequence characteristics of members of the *pe_pgrs* family.

PE_PGRS Sequenced in this Work	Size (aa)	C-term unique domain (aa position)	Intercalated unique sequences (aa position)	Remarks
PE_PGRS 1	497	-	-	-
PE_PGRS 2	488	-	-	-
PE_PGRS 5	592	-	-	-
PE_PGRS 15	607	-		-
PE_PGRS 23	562	-	-	-
PE_PGRS 24	603	-	-	-
PE_PGRS 25	577	-	-	-
PE_PGRS 26	492	-	-	-
PE_PGRS 29	370	-	-	-
PE_PGRS 31	618	-	-	-
PE_PGRS 33	498	-	-	-
PE_PGRS 34	516	-	-	-
PE_PGRS 38	533	-	-	-
PE_PGRS 41	361	-	-	-
PE_PGRS 44	543	-	-	-
PE_PGRS 46	778	-	-	-
PE_PGRS 47	525		-	-
PE_PGRS 48	615	-	-	-
PE_PGRS 51	588	-	-	-
PE_PGRS 58	584	-	-	
PE_PGRS 59	439	-	-	-
Wag22	914	-	450 - 550	-
PE_PGRS 6	595	Yes (517 - 595)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 35	558	yes (250-558)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 39	414	yes (283 - 413)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 62	504	yes (110 - 504)	-	PE_Unique Domain
LipY	437	yes (110 - 437)	-	PE_Unique Domain

Other PE_PGRS	Size (aa)	C-term unique domain (aa position)	Intercalated unique sequences (aa position)	Remarks
PE_PGRS 4	838	-	460 - 510	-
PE_PGRS 7	1307	-	-	-
PE_PGRS 8	176	-	-	-
PE_PGRS 9	784	-	420 - 500	-
PE_PGRS 10	802	-	370 - 410	-
*PE_PGRS 12/13**	137/749	-	PE_PGRS 13: 360 - 421	-
PE_PGRS 14	882	-	-	-
PE_PGRS 19	667	-	-	-
PE_PGRS 20	463	-	-	-
PE_PGRS 21	767	-	-	-
PE_PGRS 22	853	-	-	-
PE_PGRS 27	1330	-	-	-
PE_PGRS 28	742	-	-	-
PE_PGRS 32	639	-	-	-
*PE_PGRS 36**	494	-		
*PE_PGRS 37**	257	-	-	-
PE_PGRS 40	61	-	-	-
PE_PGRS 42	695	-	351-359	
PE_PGRS 43	1660	-	-	-
PE_PGRS 45	461	-	-	-
PE_PGRS 52	731	-	-	-
PE_PGRS 53	1381	-	-	-
PE_PGRS 54	1901	-	-	-
PE_PGRS 57	1489	-	-	-
*PE_PGRS 60/61**	104/195	-	-	-
PE_PGRS 3	958	Yes (882 - 958)	490 - 546	PE_PGRS_C-terminal Unique Domain
PE_PGRS 11	584	yes (300 - 584)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 16	924	yes (650 - 924)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 17	331	yes (200 - 331)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 18	458	yes (300 - 458)	-	PE_PGRS_C-terminal Unique Domain
PE_PGRS 30	1011	yes (700 - 1011)	-	PE_PGRS_C-terminal Unique Domain
*PE_PGRS 50/49**	1538/484	PE_PGRS 50: Yes (1453 - 1538)	-	PE_PGRS 50: PE_PGRS_C-terminal Unique Domain
*PE_PGRS 55/56**	714/1079	PE_PGRS 55: Yes (700 - 714)	-	PE_PGRS 55: PE_PGRS_C-terminal Unique Domain

The pairs of PE_PGRS 12 and 13, 36 and 37, 59 and 50, 55 and 56, and 60 and 61 (highlighted with asterisks) are each annotated as two distinct pe_pgrs genes in the H37Rv genome. It is likely that they constitute a single PE_PGRS protein in an ancestral mycobacterium since they are separated by a frameshift mutation that prevents the synthesis of a functional protein in H37Rv.