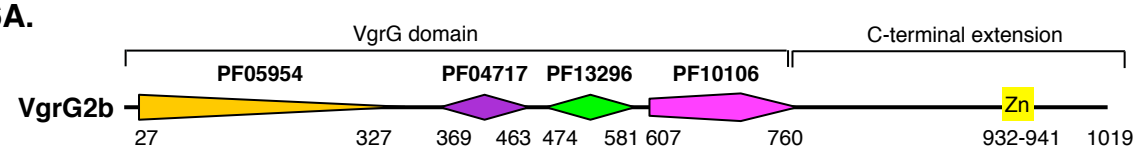
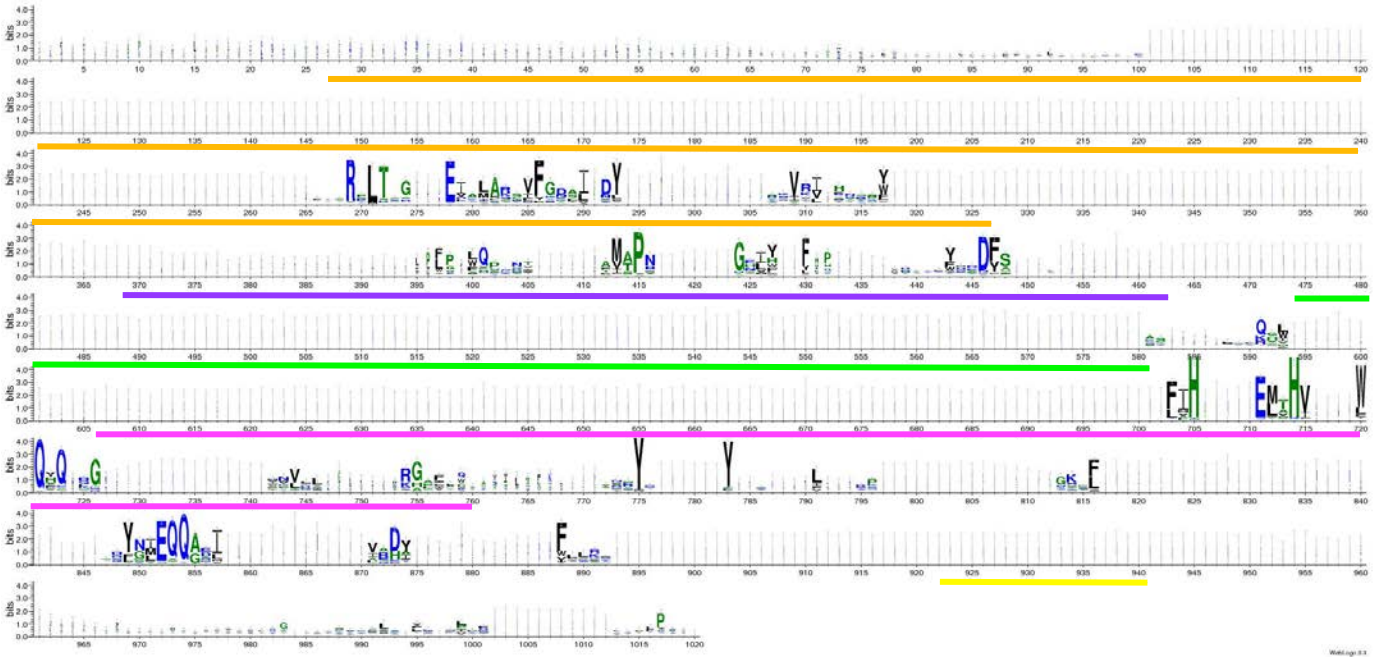
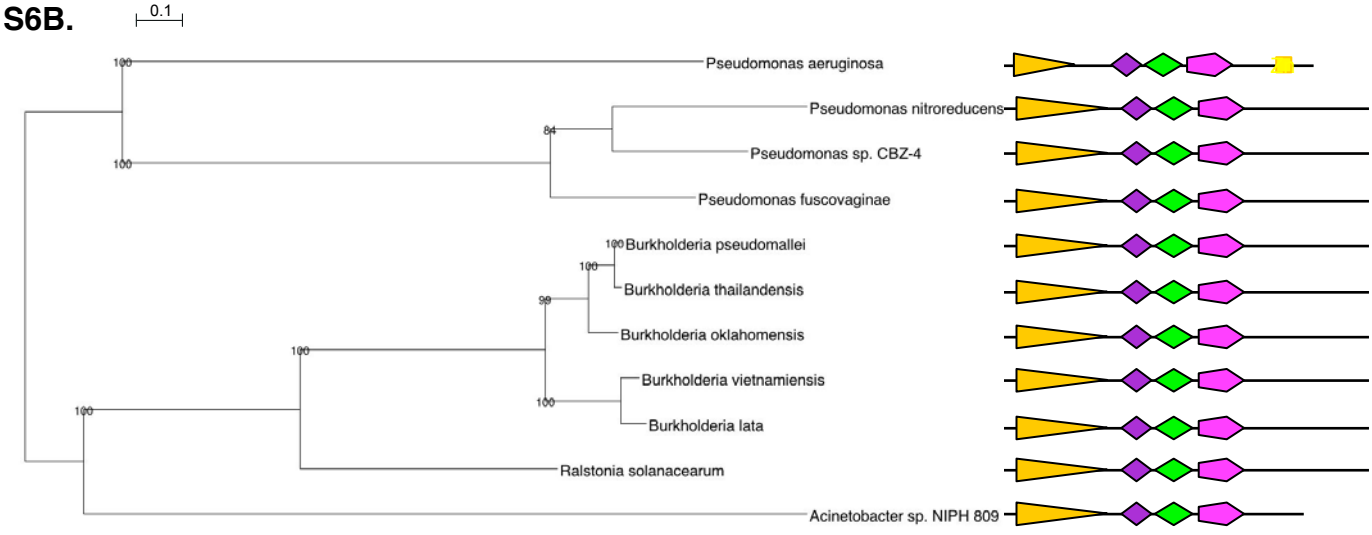
S6A.


VgrG domain
C-terminal extension
PF05954
PF04717
PF13296
PF10106
VgrG2b
Zn
27
327
369
463
474
581
607
760
932-941
1019



S6B.


0.1
Pseudomonas aeruginosa
Pseudomonas nitroreducens
Pseudomonas sp. CBZ-4
Pseudomonas fuscovaginae
Burkholderia pseudomallei
Burkholderia thailandensis
Burkholderia oklahomensis
Burkholderia vietnamiensis
Burkholderia lata
Ralstonia solanacearum
Acinetobacter sp. NIPH 809

S6C.

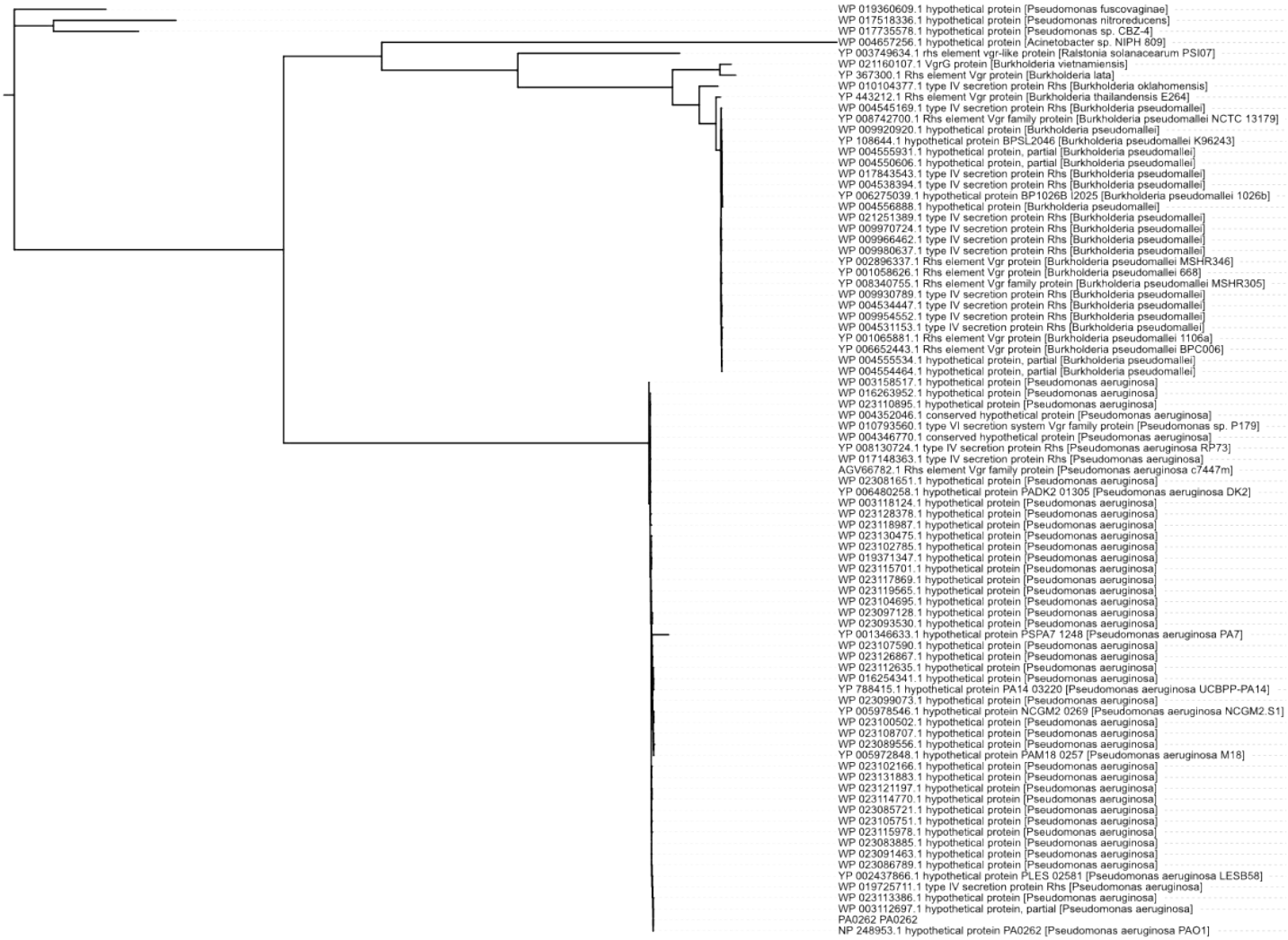


Figure S6: VgrG2b-like proteins phylogeny. Conserved features of VgG2b-like proteins (A). A sequence similarity search restricted to the last 259 residues of VgrG2b (residues 761 to 1019) retrieved 580 proteins (Step 1). Then a position specific model made from the multiple alignments of all matches (Hidden Markov Model) revealed several conserved motifs among these proteins (Step 2). The VgrG2b organization is presented with a colour code for the domains: PF05954 (orange), PF04717 (purple), PF13296 (pink), the putative Zn protease domain (yellow). WebLogo sequence alignment of VgrG2b homologues identified by the *in silico* screen. The VgrG2b domains are underlined with the same colour code. **A maximum likelihood rooted phylogenetic tree of VgrG2b homologues (B and C).** To construct a phylogenetic tree, the list of homologues has been condensed by keeping proteins harboring a VgrG domain (PF05954, PF04717, PFG13296) and the VgrG2b C-terminal extension (step 3). The phylogenetic tree was calculated using RAxML version 7.7.9. Scale bar represents the number of estimated changes per position for a unit of branch length.