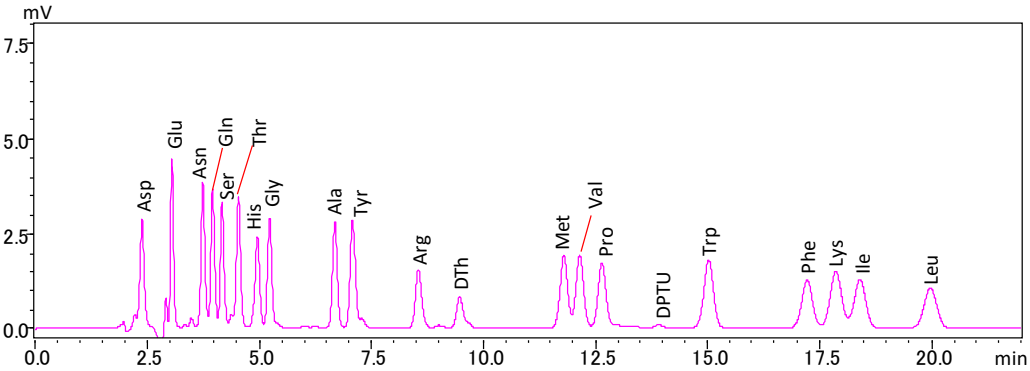
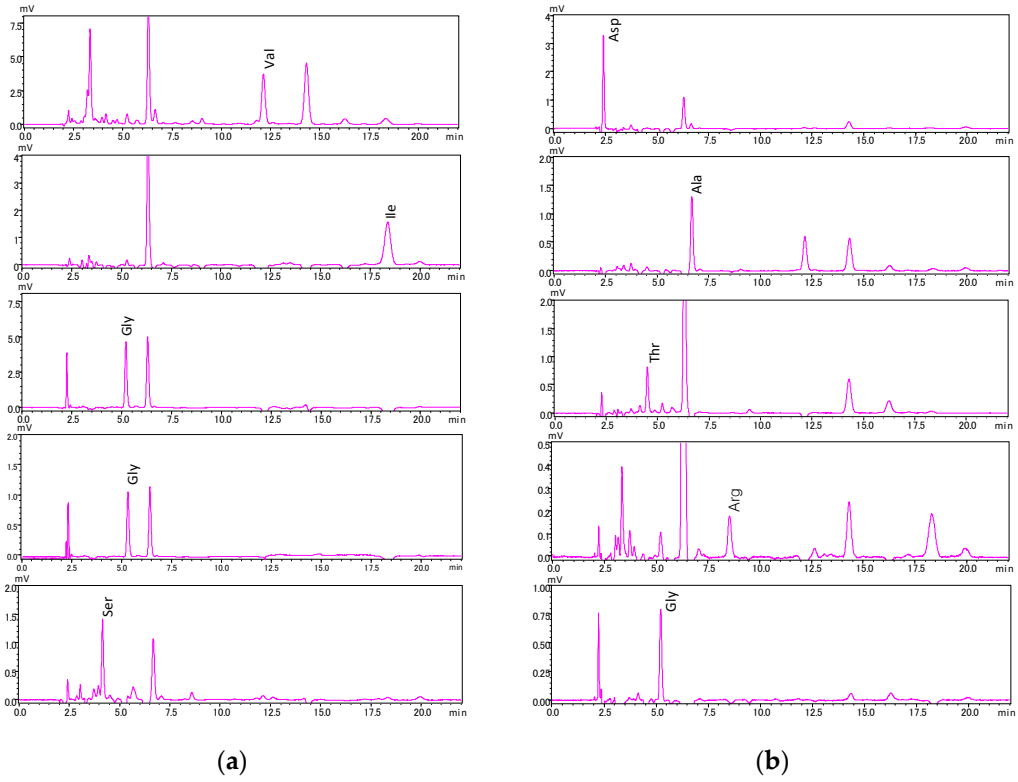
molecules

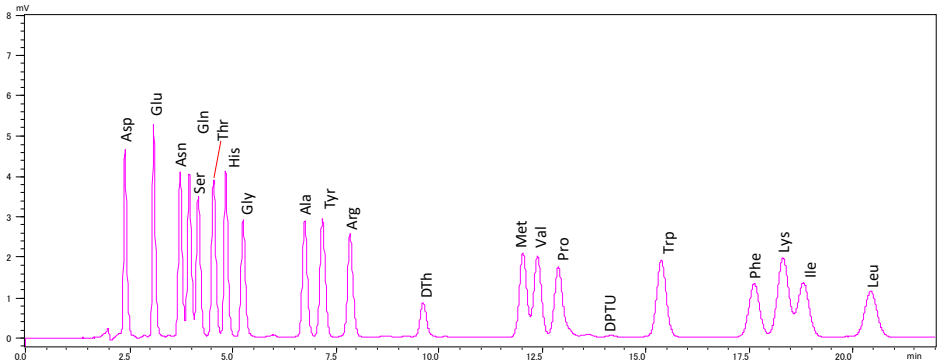
Supplementary Materials



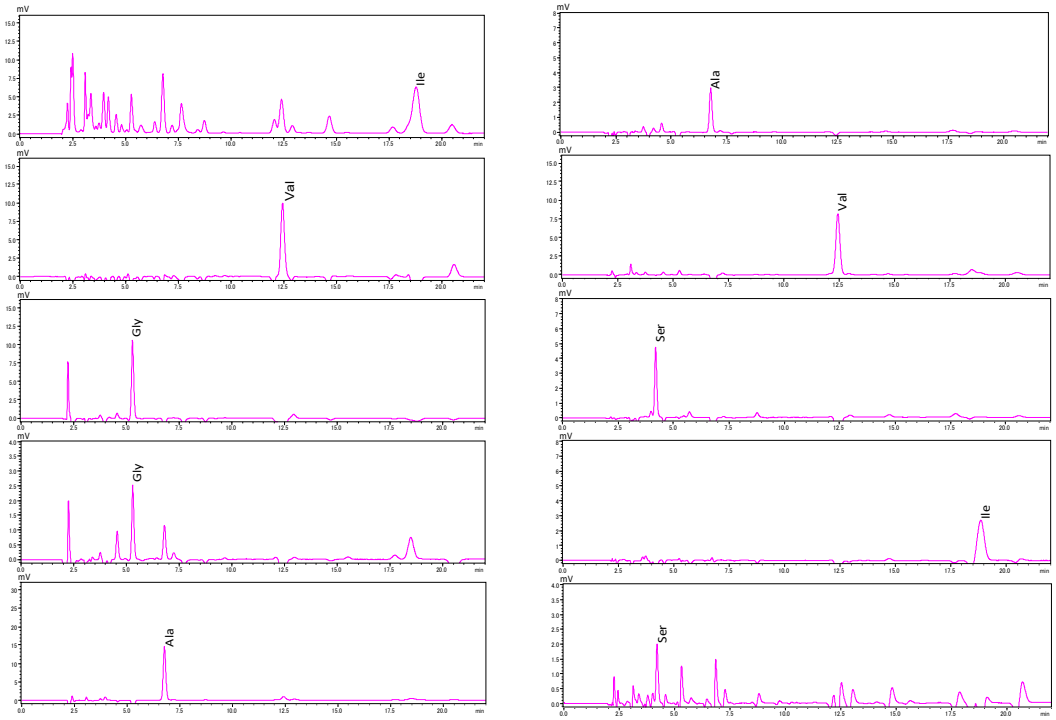
Suppl. Figure S1. Chromatography of 19 phenylthiohydantoin amino acid mixtures for *Pv*I.



Suppl. Figure S2. Chromatography of N-terminal amino acids of *Pv*I at positions 1-5 (a), 6-10 (b).

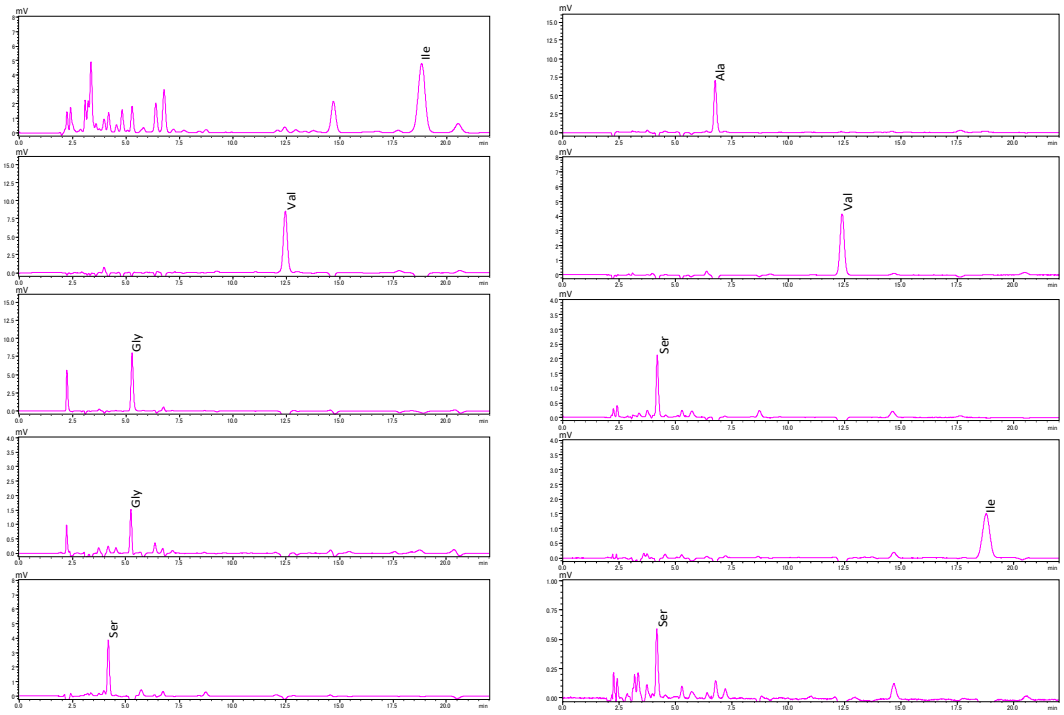


Suppl. Figure S3. Chromatography of 19 phenylthiohydantoin amino acid mixtures for *Pv*II-*Pv*VI.



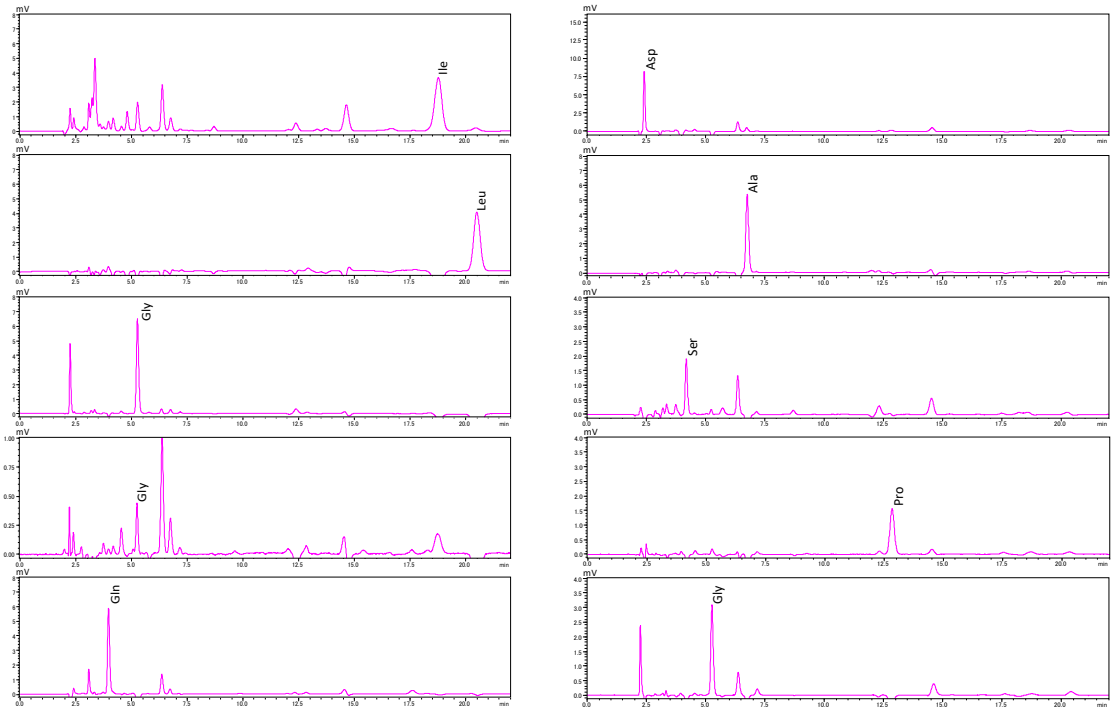
(**a**) (**b**)

Suppl. Figure S4. Chromatography of N-terminal amino acids of *Pv*II at positions 1-5 (a), 6-10 (b).



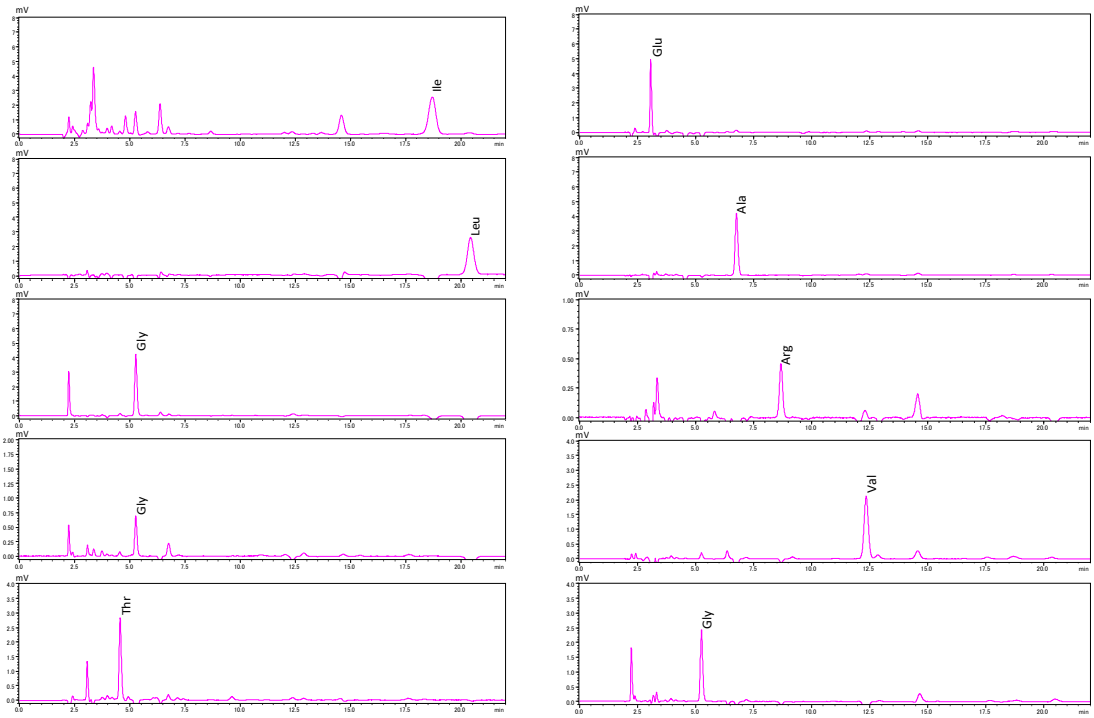
(**a**) (**b**)

Suppl. Figure S5. Chromatography of N-terminal amino acids of *Pv*III at positions 1-5 (a), 6-10 (b).



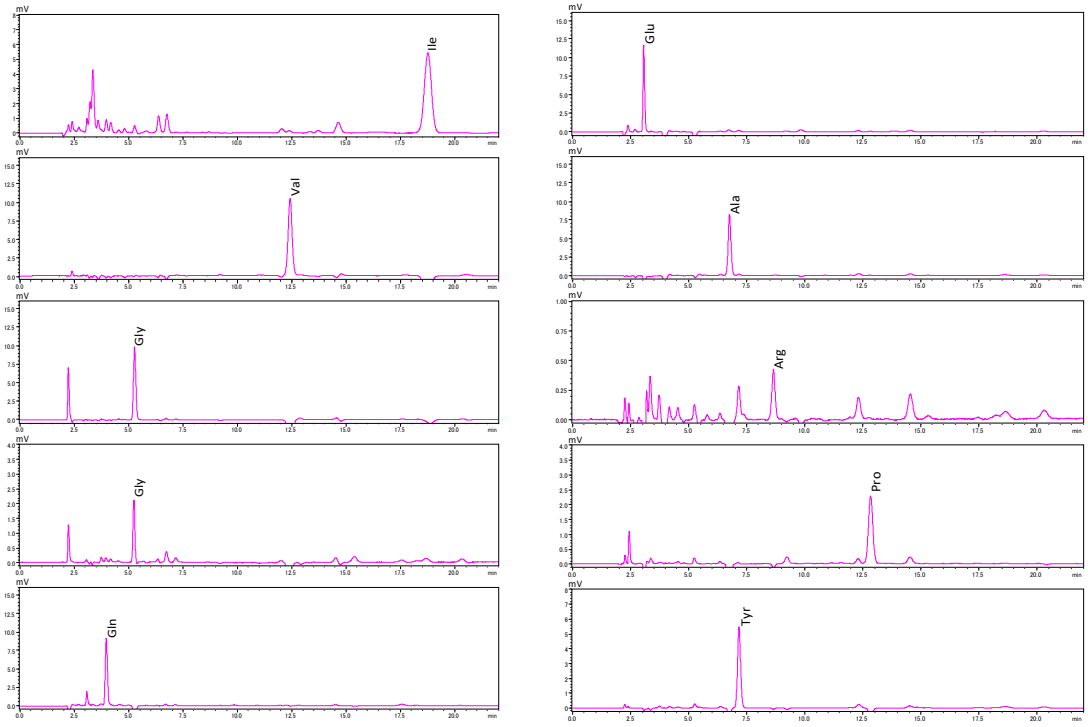
(**a**) (**b**)

Suppl. Figure S6. Chromatography of N-terminal amino acids of *Pv*IV at positions 1-5 (a), 6-10 (b).



(**a**) (**b**)

Suppl. Figure S7. Chromatography of N-terminal amino acids of *Pv*V at positions 1-5 (a), 6-10 (b).



(**a**) (**b**)

Suppl. Figure S8. Chromatography of N-terminal amino acids of *Pv*VI at positions 1-5 (a), 6-10 (b).

Suppl. Table S1. The amino acids sequence of *Pv*I.

Protein name	Amino acid sequence
*Pv*I	VIGGSDATRGEFPWQLSMTRSGSHSCGATLLSSTRALCAAH CVDGALTSSLAVIAGLHNRNDLTGTQS SDLSSYVMHSAYQ VGSASYANDIAILHFSSVISSGGNIQTATLPANNLDDYAGTT CEISGWGRTDGSNTLPNTLQ KASIPVISTAQCTVAIGGIGTIW DNHICVQDPRGNSGACNGDSGGPLNCPDGVTRVVGVTSW VVSSGLGDCLPSYPSVYTRVSAYLQWIADNSP

Suppl. Table S2. The amino acids sequence of *Pv*II.

Protein name	Amino acid sequence
*Pv*II	IVGGAAVSISNYPHQLSMRVAGSHSCGASLISNTRAVTAAH CSGSAIGAYSMLGGTTDRTVTNCGNCQLFDLTNINIHPNYD ENGNGYPNDVAVLSFISQAFNGDFQSIGLAQGSNDFAGDR CTITGWGLQQQGGNLPILLQGADMTVMTNAACASEWGAL SINDGHICLTDTNSGACNGDSGGPLVCGNVLAGATSWGH VTCSPSFPSVY TRISYFYSWIIIQ

Suppl. Table S3. The amino acids sequence of *Pv*III.

Protein name	Amino acid sequence
*Pv*III	IVGGSAVSISDHPHQLSMRVAGSHSCGASLISNTRAVTAAH CSGSAIGAYSMLGGTTDRTVTNCGNCRLFDLTNINIHPNYD ENGNGYPNDVAVLSFFAQSTNGDFNTIALAQGSNDFAGDR CTITGWGLQQQGG NLPILLQGADVTVMTNTACEAEWGVV SVNDGHICVTDTNSGACNGDSGGPLECSGLLAGATSWGHV TCSPSFPSVYTRISYFYSWIIIQ

Suppl. Table S4. The amino acids sequence of *Pv*IV.

Protein name	Amino acid sequence
*Pv*IV	ILGGQDASPGEFPWQLSQLRGGSHSCGASLLHATAALSAAH CVDGALVSSVDVIAGLHQRSDLTGTQTSAAASFSIHANYDN GEGTFAHDISIIHLATAIDASPANIAFLTLPPD NSYQFTGDTC TLSGWGRTSASN ILPDTLQKVDIEVISTADCERLMVGVIGAD CTDNHIAVHDAANNEGSCNGDSGGPMNCQLNGQTVVAG ITSWGIQSGGACAPSYPSVYTRTSAYLQWISDNQ

Suppl. Table S5. The amino acids sequence of *Pv*V.

Protein name	Amino acid sequence
*Pv*V	ILGGTEARVGEIPWQLSQQRGGSHSCGASLLRPGSALSAAH CVDGAPPADVRIVAGLHLRSDESTAVASLAESFLIHPSYNV GEGTFPNDIAIIYLLTNINSAPVE NIDFALLPPDNVEQFVGF TCVLSGWGRTSASNVLPDALQKVSIDVITTAECDSRMAAV AGADCTDAHIAVFDPALQ KGSCNGDSGGPMNCPLSGEFV VAGVTSWGISGGGACLPEYPSVYTRTGFYRQWIIDNIR

Suppl. Table S6. The amino acids sequence of *Pv*VI.

Protein name	Amino acid sequence
*Pv*VI	IVGGQEARPYEFPWQVSVRRRSTDSHFCGGSIINELWVVT AAHCMDGESPAQVSIVAGEHDSNAISTVRESHDVAAIFV HESYSSIRLTNDVSVVKVQVAFGLSVNIQPVCAPDPANLY THQKSQCSGWGSLSSGGACCPPTLRYVTLNITTNAFCDQI YTTDDITDDMICATDNTGMTDRDSCQGDSGGPLTVKDGN GI FSLVGIVSWGIGCASGWPGVYARVGYQAGWITNIINNN

Suppl. Table S7. The gene sequence of *Pv*I.

Protein name	Gene sequence
*Pv*I	GTCATCGGTGGATCGGATGCCACCCGCGGAGAGTTCCC GTGGCAGCTGTCGATGACACGCAGCGGTTCACACAGCT GCGGAGCAACGCTGCTGAGTTCGACGCGCGCTCTCTGC GCCGCTCACTGCGTCGACGGAGCACTAACATCCAGCCT CGCAGTCATTGCTGGTCTTCATAATCGCAACGACTTAAC AGGAACTCAGAGCTCCGACCTTTCCAGCTATGTTATGCA CTCAGCCTACCAGGTTGGCAGCGCGTCCTATGCAAACG ACATTGCCATTCTGCATTTCTCTTCTGTCATCAGCAGCG GTGGAAATATTCAGACCGCCACTCTTCCAGCAAATAAC CTCGACGACTATGCCGGTACCACATGTGAAATATCTGG CTGGGGCCGCACAGACGGAAGCAACACTCTTCCAAAC ACCCTTCAGAAGGCCTCCATTCCAGTCATATCTACAGC TCAATGCACTGTAGCCATTGGTGGAATAGGAACGATCT GGGATAACCACATCTGCGTCCAAGACCCTAGAGGAAA TAGTGGAGCTTGCAATGGTGACAGTGGTGGTCCGCTCA ATTGCCCAGATGGCGTTACCCGAGTCGTGGGTGTTACT TCTTGGGTTGTTTCGAGCGGTCTCGGTGACTGCCTTCCG TCTTATCCCTCGGTCTACACGCGTGTGAGCGCCTACTTG CAATGGATCGCTGACAACAGTCCCTAA

Suppl. Table S8. The gene sequence of *Pv*II.

Protein name	Gene sequence
*Pv*II	ATCGTCGGTGGAGCAGCCGTCAGCATCAGCAACTATCC GCATCAGCTGTCGATGAGGGTAGCGGGATCTCATTCGT GCGGAGCGTCCCTCATTTCAAACACCAGAGCCGTCACC GCTGCTCACTGCAGTGGATCCGCAATCGGAGCTTATTC CATGCTGGGAGGCACTACCGACCGAACAGTCACCAAC TGCGGCAACTGTCAACTCTTTGACTTGACAAACATCAA TAT TCATCCTAACTACGATGAGAATGGAAATGGATAC CCCAACGATGTGGCTGTTCTCAGCTTTATTTCACAGGC CTTCAACGGCGACTTCCAGTCCATTGGTTTGGCTCAAG GGTCAAACGATTTTGCAGGAGACAGATGTACCATCAC TGGCTGGGGTCTGCAACAACAAGGTGGAAATCTGCCA ATCCTATTGCAGGGAGCAGATATGACTGTTATGACTA ACGCTGCTTGCGCGTCTGAATGGGGTGCTCTCAGTATA AACGATGGACACATTTGTCTCACAGATACCAACTCCG GCGCTTGCAACGGGGACAGTGGTGGTCCACTGGTGTG CGGTAATGTTCTTGCTGGAGCGACGTCCTGGGGTCAT GTGACATGCAGCCCGTCGTTTCCTTCCGTCTACACGAG AATCAGCTACTTCTATTCCTGGATCATCATTCAATAA

Suppl. Table S9. The gene sequence of *Pv*III.

Protein name	Gene sequence
*Pv*III	ATTGTTGGAGGATCGGCCGTCAGCATAAGCGACCATCC GCACCAGCTGTCGATGAGGGTAGCGGGATCTCATTCGT GCGGAGCGTCCCTCATTTCAAACACCAGAGCCGTCACC GCTGCTCACTGCAGTGGATCCGCAATCGGAGCTTATTC CATGCTGGGAGGCACTACCGACCGAACAGTCACCAAC TGCGGCAACTGTCGACTCTTTGACTTGACAAACATCAA TATTCATCCTAACTACGATGAGAATGGAAATGGATACC CCAACGATGTGGCTGTTCTTAGCTTTTTTGCACAGTCCA CCAACGGTGACTTCAATACCATTGCTTTGGCTCAAGGG TCAAACGACTTTGCAGGAGACAGATGTACCATCACTG GCTGGGGTCTGCAACAACAAGGTGGAAATCTGCCAAT CCTATTGCAGGGAGCAGATGTGACTGTTATGACTAACA CTGCTTGCGAGGCGGAATGGGGTGTTGTCAGTGTAAAC GATGGACACATTTGTGTCACAGATACCAACTCCGGCGC TTGCAACGGAGACAGTGGTGGTCCATTGGAGTGCAGTG GTCTTCTCGCTGGAGCGACGTCCTGGGGTCATGTGACAT GCAGCCCGTCGTTTCCATCCGTCTACACGAGAATCAGC TACTTCTATTCCTGGATCATCATTCAATAA

Suppl. Table S10. The gene sequence of *Pv*IV.

Protein name	Gene sequence
*Pv*IV	ATTCTCGGAGGACAGGATGCCTCTCCTGGAGAGTTCCC GTGGCAACTGTCCCAGTTGAGAGGTGGAAGCCACAGC TGCGGTGCCTCTCTCCTCCACGCCACAGCTGCTCTGAG CGCCGCCCACTGCGTCGACGGAGCACTCGTATCATCAG TCGATGTCATCGCCGGTCTTCATCAGCGCTCAGATCTTA

	CTGGAACACAGACTTCAGCTGCTGCGTCATTCAGCATT CACGCGAACTACGATAACGGAGAGGGCACATTCGCAC ATGATATCTCCATCATCCACCTCGCCACAGCCATCGAC GCCTCGCCAGCCAACATCGCGTTTTTGACCCTTCCTCCA GACAACAGCTATCAGTTTACTGGTGACACCTGCACTCT CAGTGGATGGGGACGCACATCTGCCAGCAACATTCTTC CGGACACACTGCAGAAAGTTGACATCGAAGTCATCAG CACAGCCGATTGCGAACGCCTTATGGTTGGCGTCATTG GTGCGGACTGCACTGATAATCATATCGCAGTTCACGAT GCTGCCAATAACGAGGGATCGTGCAACGGTGACAGCG GTGGTCCGATGAACTGCCAACTCAATGGTCAGACTGTG GTTGCCGGAATTACGTCGTGGGGAATCCAGTCGGGCGG CGCCTGTGCACCATCCTACCCATCTGTCTACACCAGGAC TTCAGCCTACCTTCAATGGATCAGCGACAACCAATAG

Suppl. Table S11. The gene sequence of *Pv*V.

Protein name	Gene sequence
*Pv*V	ATCCTTGGAGGAACGGAAGCCAGAGTTGGAGAGATCC CATGGCAGCTGTCGCAGCAGAGAGGCGGAAGTCACAG CTGCGGAGCGTCTCTTCTCAGGCCCGGTTCGGCCCTCAG CGCCGCTCACTGCGTTGACGGAGCACCGCCAGCAGATG TGCGAATTGTCGCTGGACTTCATTTGCGCTCAGATGAAT CTACTGCAGTGGCTTCCCTTGCTGAGAGTTTCCTAATTC ACCCGAGTTACAACGTTGGAGAAGGAACTTTCCCAAAC GACATCGCCATCATCTACCTATTAACAAATATCAACTCT GCTCCAGTAGAAAACATTGATTTTGCTCTTCTACCTCCA GACAACGTCGAGCAATTCGTCGGATTTACTTGCGTGCTC AGTGGATGGGGACGCACATCGGCCAGCAATGTACTTCC CGATGCCCTGCAGAAGGTCAGCATCGACGTCATCACCA CAGCCGAATGCGACTCACGCATGGCCGCTGTTGCTGGA GCCGACTGCACTGATGCTCACATCGCCGTCTTCGATCCC GCTTTGCAGAAAGGATCGTGCAACGGTGACAGCGGTGG CCCAATGAACTGCCCTCTGAGCGGTGAATTTGTGGTTGC TGGTGTGACGTCATGGGGAATTTCGGGAGGCGGTGCCT GTCTGCCAGAATACCCATCAGTCTACACCAGAACTGGA TTCTACCGTCAATGGATCATTGACAACATTCGCTAG

Suppl. Table S12. The gene sequence of *Pv*VI.

Protein name	Gene sequence
*Pv*VI	ATCGTCGGAGGGCAGGAAGCCAGACCATACGAATTCC CGTGGCAGGTGTCCGTAAGGCGCAGGTCCACTGACAG CCATTTCTGCGGCGGCAGCATCATCAATGAACTCTGGG TTGTTACAGCTGCTCACTGCATGGACGGAGAGAGTCCC GCGCAGGTGTCGATCGTCGCAGGGGAACACGACAGCA ACGCCATCAGCACCGTTCGCGAGAGCCATGACGTCGC CGCCATCTTCGTGCACGAGTCCTACAGCAGCATCCGAC TGACAAACGACGTGTCCGTCGTCAAGGTGCAGGTGGC CTTCGGCTTGTCCGTGAACATCCAGCCGGTCTGCGCAC CGGACCCGGCCAACCTCTACACGCACCAGAAGAGCCA

	GTGCTCCGGTTGGGGAAGCCTCAGTTCAGGCGGAGCTT GCTGCCCACCCACCTTGCGTTACGTTACCCTCAACATC ACCACGAATGCCTTCTGTGACCAGATCTACACCACCGA TGACATCACGGACGACATGATCTGCGCCACAGATAAC ACTGGAATGACCGATAGAGATTCGTGCCAGGGTGATT CCGGTGGCCCCCTGACCGTCAAGGACGGAAACGGCAT CTTCTCCCTCGTCGGTATCGTCTCGTGGGGTATCGGCTG TGCCTCCGGATGGCCGGGAGTCTACGCCAGGGTGGGTT ATCAAGCAGGCTGGATCACGAACATCATTAACAACAAC TGA