Evolution and functionalization of *Vitellogenin* genes in the termite *Reticulitermes speratus*

Short running title: Evolution of *Vitellogenin* genes in termites

Hajime Yaguchi[1,2,3#*], Shogo Suzuki[1#], Naoto Kanasaki[1#], Yudai Masuoka[1,4], Ryutaro Suzuki[1], Ryouhei H. Suzuki[1], Yoshinobu Hayashi[5], Shuji Shigenobu[6] and Kiyoto Maekawa[7*]

1. Graduate School of Science and Engineering, University of Toyama, Gofuku, Toyama, 930-8555, Japan
2. Tropical Biosphere Research Center, University of the Ryukyus, Nishihara, 903-0213, Japan
3. Department of Bioscience, School of Science and Technology, Kwansei Gakuin University, Sanda, Hyogo 669-1337, Japan
4. Institute of Agrobiological Sciences, NARO (National Agriculture and Food Research Organization), Tsukuba 305-8634, Japan
5. Department of Biology, Keio University, Hiyoshi, Yokohama, 223-8521, Japan
6. NIBB Research Core Facilities, National Institute for Basic Biology, Okazaki, 444-8585 Japan
7. Faculty of Science, Academic Assembly, University of Toyama, Gofuku, Toyama, 930-8555, Japan

#: equally contributed to this work.

*: Corresponding authors: Kiyoto Maekawa (email: kmaekawa@sci.u-toyama.ac.jp), Hajime Yaguchi (email: hjiroo@gmail.com)

Abstract

Eusociality has been commonly observed in distinct animal lineages. The reproductive division of labour is a particular feature, achieved by the coordination between fertile and sterile castes within the same nest. The sociogenomic approach in social hymenopteran insects indicates that *vitellogenin* (*Vg*) has undergone neo-functionalisation in sterile castes. Here, to know whether *Vgs* have distinct roles in non-reproductive castes in termites, we investigated the unique characteristics of *Vgs* in the rhinotermitid termite *Reticulitermes speratus*. The four *Vg*s were identified from *R. speratus* (*RsVg1-4*), and *RsVg3* sequences were newly identified using the RACE method. Molecular phylogenetic analysis supported the monophyly of the four termite *Vg*s. Moreover, the termites *Vg1-3* and *Vg4* were positioned in two different clades. The *dN* / *dS* ratios indicated that the branch leading to the common ancestor of termite *Vg4* was under weak purifying selection. Expression analyses among castes (reproductives, workers, and soldiers) and females (nymphs, winged alates, and queens) showed that *RsVg1-3* were highly expressed in fertile queens. In contrast, *RsVg4* was highly expressed in workers and female non-reproductives (nymphs and winged adults). Localisation of *RsVg4* mRNA was confirmed in the fat body of worker heads and abdomens. These results suggest that *Vg* genes are functionalised after gene duplication during termite eusocial transition and that *Vg4* is involved in non-

reproductive roles in termites.

Keywords

eusocial insects, reproductive division of labour, vitellogenin, molecular evolution, gene expression

Introduction

Eusociality is primarily characterised by the reproductive division of labour between fertile and sterile castes within the same nest (Wilson, 2000). It has repeatedly evolved in several lineages, including those of ants, bees, wasps, and termites. Reproduction is generally monopolised by one or a few fertile individuals, whereas other tasks are performed by sterile individuals without activated gonads. The genetic background is fundamentally shared among individuals in a nest, and phenotypic plasticity leads to evolutionary success that accompanies their task specialisation for colony survival and growth (e.g. O'Shea-Wheller *et al.*, 2021). Importantly, sterile castes are assumed to be evolutionary novelties derived during the process of eusocial evolution from the solitary ancestor (Korb and Heinze, 2004). Therefore, understanding the molecular mechanisms of the altruistic roles of sterile castes is a fundamental research goal in evolutionary biology. In social Hymenoptera, important regulators involved in social behaviours and caste differentiation were identified based on genomic and transcriptomic analyses. These studies suggest that conserved genes have different functions from their original roles during social evolution (Toth and Rehan, 2017). One of the best-studied genes with these features is *vitellogenin* (Richards, 2019), which is a socially pleiotropic gene with different functions in the reproductive and non-reproductive castes (see below).

Vitellogenin (Vg) is a major yolk precursor in most oviparous animals and is involved in ovary

65 development (Wahli, 1988; Raikhel and Dhadialla, 1992). In insects, *Vg* genes are generally expressed
66 in the fat body of adult females, and the related proteins are transported into the oocytes for storage
67 (Tufail and Takeda, 2008). Another feature is that behavioural transitions (e.g. nurses to foragers) are
68 mediated by the difference in *Vg* expression patterns with worker age in the honey bee *Apis mellifera*
69 and the harvester ant *Pogonomyrmex barbatus* (Amdam *et al*., 2004; Nelson *et al*., 2007; Corona *et*
70 *al*., 2013). These studies suggest that Vg is a crucial regulator of worker-specific behaviours as well
71 as the reproductive maturity of queens among social hymenopteran insects. Based on the inferred
72 molecular phylogeny, Vg may be duplicated within formicid ant lineages (Corona *et al*., 2013). The
73 duplicated genes have been shown to be differentially expressed between queens and workers in *P.*
74 *barbatus* and the fire ant *Solenopsis invicta* (Wurm *et al*., 2011; Corona *et al*., 2013). In ants and honey
75 bees, three *Vg* homologs (*Vg-like-A*, *-B,* and *-C*), which may be important for worker-specific
76 behaviour and physiology, were also identified (Kohlmeier *et al*. 2018; Miyazaki *et al*. 2021; Morandin
77 *et al*., 2014; Morandin *et al*., 2019; Salmela *et al*., 2016). The expression patterns of *Vg-like-A* and *-B*
78 were different between castes in each of seven *Formica* species, but those of *Vg-like-C* were
79 consistently highly expressed in workers (Morandin *et al*., 2014). The existence of multiple *Vg*
80 homologs, some of which lost an ancestral function and acquired a new function, may be essential for
81 the evolution of highly sophisticated social organisation in ants (Friedman and Gordon, 2016). In other
82 eusocial lineages, including termites, however, little is known about whether the particular *Vg* gene
83 plays an important role in sterile castes.

84

85 Termites are hemimetabolous insects that evolved from a cockroach-like ancestor (e.g. Lo *et al*., 2000;
86 Inward *et al*., 2007). Contrary to social hymenopterans, termite primary reproductives (queen and king)
87 are adult individuals, but sterile castes, including soldiers and workers, are immature individuals
88 (Thorne, 1997). High *Vg* expression in female reproductives has been observed in several termite
89 species (reviewed in Korb, 2016). For example, *Vg* gene expression is positively correlated with
90 ovarian development in primary queens of the rhinotermitid termite *Reticulitermes speratus* (Ishitani
91 and Maekawa, 2010; Maekawa *et al*., 2010). The *Vg* gene was also highly expressed in secondary
92 reproductives (differentiated from immature stages and succeeded in the reproductive role from
93 primary reproductives; Fig. S1) of *R. speratus* and the dry-wood termite *Cryptotermes secundus* (Weil
94 *et al*., 2007; Saiki and Maekawa, 2011). The first termite species whose genome has been sequenced
95 is the damp-wood termite *Zootermopsis nevadensis*, in which the presence of four *Vg* genes have been
96 reported (Terrapon *et al*., 2014). Three *Vgs* were highly expressed in female reproductives, and one
97 was moderately expressed in non-reproductive individuals (Terrapon *et al*., 2014). This suggests the
98 different functions of *Vg* genes in different castes. Therefore, it is possible that termite *Vg*s have
99 alternative roles not only in reproductive regulation but also in non-reproductive traits. However,
100 further information is required to understand the molecular evolution and function of termite *Vg* genes;
101 for example, precise phylogenetic relationships and expression pattern/localisation should be
102 identified.

In the present study, to obtain additional information on termite *Vg* genes, we focused on *R. speratus*, for which whole genome sequences have recently been made available (NCBI Submission DRA010978; BioProject Accession PRJDB2984, Shigenobu *et al*., 2022). First, the *Vg* genes of *R. speratus* (*RsVgs*) were identified from the genome, and a precise molecular phylogenetic tree was constructed to determine the evolutionary relationships among multiple *RsVgs* with other termite/insect *Vgs*. Second, the expression levels of *RsVgs* were investigated using RNA-seq data from reproductive, worker, and soldier castes (NCBI BioProject Accession PRJDB5589, Shigenobu *et al*., 2022). Third, to determine whether there were any relationships between *RsVgs* expression and reproductive cycle in *R. speratus* queens, colony development was observed, and expression levels were examined by real-time quantitative PCR. Finally, to localise the mRNA of *RsVg1* and *RsVg4*, *in situ* hybridisation was performed in queens and workers. Based on the results obtained, we suggest that termite *Vgs* are functionalised after gene duplication and *RsVg4* is involved in non-reproductive role in *R. speratus*.

Materials and Methods

Termites

Mature colonies of *R. speratus* were collected in Toyama Prefecture, Japan, from 2015 to 2017. The collected colonies were maintained in plastic boxes and kept at 25 °C in constant darkness. This species has terminal workers (wingless immature), which occasionally differentiate into soldiers via presoldiers and secondary reproductives (Roisin, 2000). Secondary reproductives are differentiated from workers and nymphs (Fig. S1), and succeeded in the reproductive role when the primary reproductives die or become senescent (Thorne, 1997; Roisin, 2000). The nymphs (winged immature) can differentiate into winged alates and become the primary reproductives, and sometimes, secondary reproductives (Roisin, 2000; Saiki and Maekawa, 2011) (Fig. S1). It was impossible to collect all individuals from colonies with the constant social environment and resource availability. Workers, soldiers, and nymphs were sampled from mature colonies. Primary reproductives (queens and kings) were collected from the incipient colonies, which were established using alates (winged adults) derived from different mature colonies in accordance with previous studies (Maekawa *et al*., 2010; Shimada and Maekawa, 2010; Miyazaki *et al*., 2014) (see below). The sampled individuals were classified into W4 workers (6th instar, Fig. S1), W5 workers (7th instar), N6 nymphs (8th instar), alates, and primary reproductives, and separated into sex based on genital plates (Weesner, 1969).

Identification of vitellogenin (Vg) gene from *R. speratus* genome

The Vg amino acid sequences were obtained from *Z. nevadensis* (Terrapon *et al*., 2014) and other

141 insect protein databases. The similarity search was performed by blastP (evalue < 1.0E-05) using the
142 sequences obtained from the *R. speratus* protein database (Rspe OGS 1.0; Shigenobu *et al*., 2022). Of
143 the *Vg* sequences identified, *RsVg3* (gene ID: RS013977) mRNA was determined by 3′-rapid
144 amplification of cDNA ends (3′-RACE) because the fragment of *RsVg3* (RS013977; 2217 bp) was
145 incomplete at the 3′ site (see Results). The primer sequences were designed from the obtained *RsVg3*
146 sequence using the web interface Primer3Plus (Untergasser *et al*., 2007) to determine the sequence of
147 the full-length *RsVg3* (Table S1). Total RNA was extracted from the whole body of a female secondary
148 reproductive collected from mature colonies using ISOGEN II (Nippon Gene, Tokyo, Japan). The
149 extracted RNA was purified using DNase I treatment to remove genomic DNA. The quality and
150 quantity of purified RNA were measured using a NanoVue spectrophotometer (GE Healthcare Bio-
151 Sciences, Tokyo, Japan). The cDNA was synthesised using the SMARTer^{TM} RACE cDNA
152 Amplification kit (Clontech Laboratories, Mountain View, CA, USA) in accordance with a previous
153 study (Hojo *et al*., 2011). The obtained 3′ RACE products were purified using the QIAquick Gel
154 Extraction Kit (Qiagen, Tokyo, Japan), subcloned into a pGEM easy T-vector (Promega, Madison, WI,
155 USA), and transfected into *Escherichia coli* XL1-blue. Plasmids, including DNA fragments, were
156 extracted from a single colony of *E. coli* using the SIGMA GelElute^{TM} Plasmid Miniprep Kit (Sigma-
157 Aldrich, St Louis, MO, USA). The inserted DNA sequence was determined using the BigDye
158 Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA) with an automatic
159 DNA Sequencer 3130 Genetic Analyzer (Applied Biosystems). The obtained sequence was confirmed

160 by a BLAST search and deposited in the DDBJ/EMBL/GenBank databases (accession no. LC621340).

161

162 Molecular phylogenetic analyses of *Vg* genes

163

164 The amino acid sequences of Vg, Vg-like apolipocrustacein (ApoCr) and insect apolipophorin II/I
165 (apoLp-II/I) were obtained from NCBI, in accordance with previous studies (Avarre *et al*., 2007; Tufail
166 and Takeda, 2008; Hayward *et al*., 2010; Morandin *et al*., 2014; Roy-Zokan *et al*., 2015; Salmela *et*
167 *al*., 2016; Shang *et al*., 2017). According to these reports (Avarre *et al*., 2007; Hayward *et al*., 2010),
168 the ApoCr and apoLp-II/I sequences were used as outgroups. The DNA and/or amino acid sequences
169 were obtained from the appropriate databases and related references (Table S2). The obtained
170 sequences (286-3380 aa) were aligned using MAFFT with the option 'auto' (FFT-NS-i strategy) (Katoh
171 and Standley, 2013). The aligned sequence sets (Supplementary data) were processed using trimAl
172 v1.2 rev59 with the option 'automated1' (Capella-Gutierrez *et al*., 2009). The molecular phylogenetic
173 tree and branch support values were estimated using two different methods. The neighbor-joining
174 method (Saitou and Nei, 1987) was used to construct the phylogenetic tree using MEGA7 (bootstrap
175 values: 5000, JTT model, pairwise deletion) (Kumar *et al*., 2016). RAxML (Stamatakis, 2014) was
176 used for tree construction by the maximum likelihood method (bootstrap values: 1000). The amino
177 acid substitution models were estimated using MEGA7, and the LG + G model was selected (BIC,
178 26782.8213097696; lnL, -12549.156785332). The accession number or gene ID of each sequence is

179 shown in Table S2.

180

181 Molecular evolution of termite *Vg* genes

182

183 To investigate the selection pressure at play on *Vg* genes in termites, each *Vg* gene from all termite
184 species and *C. punctulatus*, serving as an outgroup, were independently aligned. Termite *Vg1* (see
185 Results) alignment was missing *C. formosanus* species as *Vg1* was not present in this species, and
186 termite *Vg3* (see Results) was completely excluded since it was only present in termites and the use of
187 a cockroach outgroup was therefore impossible. To estimate overall dN / dS (ω) ratios of each *Vg* genes
188 in the termite clade, the null model (M0) from codeml in PAML (v4.9j) was used (Yang 2007). To
189 assume an independent ω ratio for each branch and to evaluate variation of ω ratio between the branch
190 of the termite most recent common ancestor and the other branches, a branch test was used. The branch
191 test from codeml consists of a log-likelihood ratio test following a $\chi 2$ distribution of the null model
192 M0 with two models, a free-ratio model M1 and a two-ratio model M2.

193

194 *RsVg* gene expression analysis using RNA-seq data

195

196 RNA-seq data were used to compare the expression levels of *RsVgs* among three castes (workers,
197 soldiers, and reproductives), two body parts (heads and the remaining parts), and two sexes (biological

198 triplicates; Shigenobu *et al*., 2022). Briefly, older workers (W4–5) and soldiers were sampled from
199 natal colonies in accordance with their external morphology, and primary reproductives were sampled
200 from incipient colonies established by female–male pairs. RNA-seq analyses were performed with 12
201 categories (castes × sexes × body parts) with three biological replicates (in total 36 libraries). Detailed
202 procedures for the trimmed mean of M-values (TMM) normalization, reads per kilobase of exon per
203 million mapped sequence reads (RPKM) calculation were based on a previous analysis (Shigenobu *et*
204 *al*., 2022). To examine differential expression among castes and between sexes, we modelled TMM-
205 normalized counts per million (CPM) for each gene using a generalized linear mix model (GLMM) in
206 which caste, sex and their interaction are treated as fixed effects and individual sample as random
207 effects, using the R package glmmSeq v0.1.0 (https://github.com/KatrionaGoldmann/glmmSeq). A
208 false discovery rate (FDR) of 0.05 was used as a threshold for statistical significance.
209
210 Measurements of egg numbers during incipient colony development
211
212 Incipient colonies were identified according to previous studies (Maekawa *et al*., 2010; Shimada and
213 Maekawa, 2010; Miyazaki *et al*., 2014). Alates were collected from multiple mature colonies, and then
214 the individuals were separated into sexes based on abdominal sternites (Weesner, 1969). Female and
215 male de-winged individuals from different colonies were paired. Each pair was placed in a 20 mL glass
216 vial with mixed sawdust food (Mitani, Ibaraki, Japan) and maintained at 25 °C in constant darkness.

217 Colonies were sampled after 0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5 months (9–60 colonies at each
218 point). The number of eggs produced by queens was counted from the sampled colonies at each time
219 point. The number of eggs was compared among the periods during colony development using the
220 Kruskal-Wallis test followed by the Steel-Dwass test (based on the results of Levene's test) with R
221 v.3.4.0 (R Core Team, 2017). The number of eggs during colony development was described with the
222 polynomial regression equation, and the degree was statistically compared by the analysis of variance
223 (ANOVA) among the constructed polynomial regression models using R v.4.0.3 (R core team, 2020).
224
225 *RsVg* gene expression analysis in queens during incipient colony development
226
227 Gene-specific primers were designed against the obtained *RsVg* sequences using Primer3Plus
228 (Untergasser *et al*., 2007) for real-time quantitative PCR (Table S1). Total RNA was extracted from 3
229 different individuals (except for the head parts) and biological triplicates were prepared for each
230 developmental stage. Extracted RNA were purified by DNase treatment using the method described
231 previously. cDNA was synthesised from the purified RNA using a High-Capacity cDNA Reverse
232 Transcription Kit (Applied Biosystems). Expression analyses of *Vg* genes were performed using
233 Thunderbird SYBR qPCR Mix (Toyobo, Osaka, Japan) and Mx3005P Real-Time QPCR System
234 (Agilent Technologies, Santa Clara, CA, USA). To determine a sustainable internal control gene, the
235 suitability of six reference genes (Table S3), *EF1-alfa* (accession no. AB602838), *NADH-dh* (no.

AB602837), *β-actin* (no. AB520714), *GstD1* (gene ID: RS001168), *EIF-1* (RS005199), and *RPS18* (RS015150) were used and evaluated using GeNorm (Vandesompele *et al*., 2002) and NormFinder (Andersen *et al*., 2004) software. All gene-specific primers used in this analysis (Table S1) were confirmed to amplify a single PCR product using the dissociation curves and product sizes. The expression levels of *RsVgs* were statistically analysed using one-way ANOVA followed by Tukey's multiple comparison test (based on the results of Levene's test) with R v.3.4.0. The expression levels of *RsVgs* were described with the polynomial regression equation, and the degree was statistically compared by ANOVA among the constructed polynomial regression models using R v.4.0.3.

Localisation of *RsVg1* mRNA in queens and *RsVg4* mRNA in workers

Gene-specific primers were designed against the obtained *RsVg1* and *RsVg4* sequences using Primer3Plus (Untergasser *et al*., 2007) for RNA probe synthesis (Table S1). The probe for *in situ* hybridisation was prepared according to a previous study (Hojo *et al*. 2011). Total RNA was extracted from whole bodies of five workers using ISOGEN II, and purified as described previously. cDNA was synthesised as the method shown above and used for the amplification of *RsVg4* fragment. cDNA was synthesised from the total RNA extracted from female secondary reproductives (see above), and used for the amplification of *RsVg1* fragment. The amplified *RsVgs* fragment was subcloned, and inserted DNA sequence was determined as described previously. The obtained sequence was confirmed to be

255 the target sequence of *RsVgs* by BLAST search. The digoxygenin (DIG)-labelled RNA probes of
256 *RsVg1* and *RsVg4* were prepared from plasmids containing *RsVgs* fragments using a DIG RNA
257 Labeling Kit (Roche, Basel, Switzerland). Three workers were separated into heads and bodies (thorax
258 and abdomen) by dissection, and each part was fixed with 4% paraformaldehyde and cryoprotected in
259 sucrose solution. A queen was collected 5.5 months after the incipient colonies were constructed, and
260 the body separated from the head was fixed and cryoprotected. The fixed samples were embedded in
261 OCT compound (Sakura Finetek USA Inc., Torrance, CA, USA) and sliced to prepare 14 μm thick
262 parasagittal cryosections with a CM1510S cryostat (Leica, Tokyo, Japan). After the sections were
263 hybridised with DIG-labelled antisense probe (or sense RNA probe for control), the slides were washed
264 using an ISHR Starting Kit (Nippon Gene) in accordance with the manufacturer's instructions.
265 Detection of DIG-labelled RNA was performed with an alkaline phosphatase-conjugated anti-DIG
266 antibody with nitro-blue tetrazolium/ 5-bromo-4-chloro-3-indolyl-phosphate (NBT/BCIP) as a
267 substrate, using a DIG Nucleic Acid Detection Kit (Roche). Images were captured using a Biozero
268 microscope (Keyence, Tokyo, Japan).
269
270 Results
271
272 Identification of four *RsVg* genes
273

Similarity searches were conducted to determine the number of *Vg* genes in the *R. speratus* genome. There are four *Vg* genes in the genome, similar to those in *Z. nevadensis* (Terrapon *et al*., 2014). The lengths of the four *Vg* genes (*RsVg1*, *2*, *3*, and *4*) were predicted to be 6048 bp (*RS000616*), 6042 bp (*RS000610*), 2217 bp (*RS013977*), and 4269 bp (*RS011073*), respectively. The predicted CDS length of *RsVg3* (*RS013977*) was less than that of other *RsVg* sequences and incomplete at the 3′ site (with no stop codon). Moreover, the deduced amino acid sequences possessed a conserved Lipid transport protein, N-terminal (Lipid_transpt_N) (IPR001747), but the von Willebrand factor, type D domain (VWD) (IPR001846) could not be identified in *RsVg3* (*RS013977*). Consequently, we performed 3′-RACE and obtained the full length of *RsVg3* (5310 bp), which possessed these three domains. These results indicated that the four *RsVgs* identified were homologous genes with three conserved domains, Lipid_transpt_N, Vitellinogen, open beta-sheet, and VWD (Fig. S3).

Molecular phylogenetic analyses

To resolve the phylogenetic positions of *RsVgs*, molecular phylogenetic analyses were performed using maximum likelihood and neighbor-joining methods. Four monophyletic groups, including *Vg* genes obtained from termites (*R. speratus*, *Z. nevadensis*, *C. formosanus*, *C. secundus*, and *M. natalensis*) and cockroaches (*C. punctulatus* and/or *B. germanica*) were highly supported by both methods (99–100% bootstrap values) (Fig. 1). The extant copies were orthologous relationships among termite

293 species, and we named these clades the termites *Vg1-4*. The termite *Vg1-3* were closely related to
294 conventional *Vg* and *Vg-like-C* homologs, although the precise phylogenetic relationships among each
295 clade were not identified because of low bootstrap values (Fig. 1). In contrast, the termite *Vg4* was
296 closely related to the *Vg-like B* genes of social hymenopteran insects (Fig. 1). Note that we found 2
297 *Vg4* sequences in *C. secundus* (XP_023710404-5), but based on the position of their locus
298 (LOC111866074), we concluded that they were splicing variants, and one of them (OTU No. [63])
299 was used for molecular phylogenetic analyses. Although the precise numbers of *Vg* genes in
300 cockroaches and termites are still ambiguous [for example, no termite *Vg1* gene in the current gene
301 model of *C. formosanus* (Itakura et al. 2020)], no homologous termite *Vg3* genes were identified in
302 cockroaches. However, homologous termite *Vg4* genes were observed both in *C. punctulatus* (OTU
303 No. [64]) and *B. germanica* [65].
304
305 Molecular evolution of three termite *Vgs*
306
307 The estimated dN / dS (ω) ratios were much less than 1 in all focal branches observed in three *Vgs* (Fig.
308 S2, Table S4). The ω ratio observed in the branch leading to the common ancestor of termites was not
309 significantly different from those in other branches in the conventional termite *Vg1-2* (Table S4). In
310 termite *Vg4*, however, the ω ratio in the focal branch was significantly different from those in other
311 branches (Table S4), although an independent ω ratio obtained from a free-ratio model M1 was not

312 strongly lower or higher than those of the other branches (Fig. S2c).

313

314 Expression analyses among castes using RNA-seq data

315

316 A GLMM analysis for the RNA-seq expression data showed that the expression levels of *RsVg1-3*
317 genes were significantly different among castes and sexes (GLMM: FDR < 0.05, Fig. 2, TableS5) with
318 higher expression in female reproductives (queens) compared to males and sterile castes, both in heads
319 (Fig. 2a) and bodies (thoraxes + abdomens) (Fig. 2b), except for *RsVg3* in heads. The expression levels
320 of *RsVg4* genes were also significantly different among castes (FDR < 0.05) in heads (Fig. 2a) and
321 bodies (Fig. 2b). However, the *RsVg4* gene expression levels of workers were higher than those of
322 other castes, and no differences were observed between female and male individuals. We found no
323 significant interaction between caste and sex in all GLMM analyses. Note that because RNA-seq
324 libraries were prepared from pools of multiple tissues, and only three biological repications were
325 analyzed, the differences of gene expression levels may include confounded effects such as allometric
326 differences and/or other ecological factors (e.g. social context, colony size).

327

328 Measurements of egg production numbers during incipient colony development

329

330 The number of eggs was measured during incipient colony development, and the reproductive cycles

of queens were clearly observed (Fig. 3), as shown in a previous study (Maekawa *et al*., 2010). The number of eggs changed significantly during the period of colony development (Kruskal-Wallis test, $df = 7$, $x^2 = 84.44$, $p = 1.71\text{e-}15$; Steel-Dwass test, $p < 0.05$). The onset of egg production occurred during the initial periods of colony development (within 3.5 months of colony establishment). The diapause periods were observed from 3.5 to 6.5 months, and the restart of egg production occurred 7.5 months after colony establishment. The polynomial regression analyses indicated that the changes of egg numbers were significantly fitted by cubic equations during colony developments (Fig. S4a, Table S6).

Gene expression analyses of queens during incipient colony development

Expression analyses of *RsVg* genes in female individuals were performed using real-time qPCR. The *RsGstD1* gene was selected as an endogenous control based on a comparison with five other reference genes examined in this study (Table S3). The expression levels were compared among several time points during incipient colony development. The expression levels of *RsVg1* were higher in the female reproductives (queens) at 1.5 months after colony establishment but extremely low at 3.5 and 4.5 months, and then increased at 6.5 and 7.5 months (Fig. 4a; one-way ANOVA, $df = 10$, $F = 2.9585$, $p = 0.01624$; Tukey's test, $p < 0.05$). The expression patterns of *RsVg2* and *RsVg3* were essentially similar to those of *RsVg1*, whereas no significant differences were found between the periods (Fig. 4bc; one-

350 way ANOVA, $df = 10$, $F = 2.4151$, $p = 0.04069$ (*RsVg2*) and $df = 10$, $F = 2.3005$, $p = 0.04967$ (*RsVg3*);
351 Tukey's test, $p < 0.05$). Note that the lack of significance might be due to a lack of power from our
352 data as the number of biological replicates is low and no technical replicates have been performed. In
353 contrast, the expression levels of *RsVg4* were significantly higher in the periods before colony
354 establishment (N6 nymphs and alates) (Fig. 4d; one-way ANOVA, $df = 10$, $F = 4.4846$, $p = 0.00161$;
355 Tukey's test, $p < 0.05$). The polynomial regression analyses indicated that the expression patterns of
356 *RsVg1-3* were significantly fitted by cubic equations (Fig. S4b-d, Table S6), whereas those of *RsVg4*
357 were significantly fitted in quadratic equations (Fig. S4e, Table S6).
358
359 Localisation of *RsVg* mRNA
360
361 *RsVg4* mRNA localisation was examined by *in situ* hybridisation using worker heads and bodies from
362 mature colonies (Fig. 5). The results showed that *RsVg4* mRNA signals were localised in the fat body
363 of worker heads and bodies (Fig. 5ac). Specific signals were not observed in sections treated with the
364 sense probe (Fig. 5bd). Similarly, *RsVg1* mRNA signals were observed in the fat body of a queen
365 thorax and abdomen (Fig. 5e), and no signals were detected in the sections treated with the sense probe
366 (Fig. 5f).
367
368 Discussion

369

370 Phylogenetic relationships of termite *Vitellogenin* genes

371

372 In the present study, four *Vg* genes were identified in the genome of *R. speratus*. Generally, in
373 oviparous animals, Vg is distinguished from other large lipid transfer proteins (LLTP) by having three
374 conserved domains (Avarre *et al*., 2007). A conserved domain, Vitellinogen open β-sheet (IPR015255)
375 also known as a DUF1943 (PF09172), was lacking in the *Vg* genes within the clade containing *Vg-*
376 *like-B* of social hymenoptera insects (OTU No. [66]–[71] in Fig. 1) and those of Hemiptera (No. [72]–
377 [74]) (Shang *et al*., 2017), however, all termite *Vg4* genes except *C. formosanus* homolog (No. [62])
378 possessed this domain. Although the relevance of the Vitellinogen open β-sheet domain is currently
379 unclear, its loss may occur independently in several lineages. The molecular phylogenetic tree showed
380 that the termite *Vg1-2* were positioned at a conventional insect *Vg* clade, and termite *Vg3* was most
381 closely related to this clade, whereas the termite *Vg4* was closely related to the hymenopteran *Vg-like-*
382 *B* clade (Fig. 1). Phylogenetic relationships shown in Fig. 1 were essentially similar to those inferred
383 in Kohlmeier et al. (2018), for example *Vg-like-A* and *Vg-like-B* were closely related to each other. We
384 found that *RsVg3* possessed all three conserved domains (Fig. S3), whereas previous studies showed
385 that the hymenopteran *Vg-like-C* had incomplete domain structure (Morandin *et al*., 2014; Salmela *et*
386 *al*., 2016). There were no cockroach homologs for termite *Vg3*, suggesting that it had appeared only
387 in termites. However, there were cockroach homologs for termite *Vg4*, and it might have been already

388 present in the most recent common ancestor between termites and cockroaches. The *dN* / *dS* ratios
389 indicated that the branch leading to the common ancestor of termite *Vg4* was relatively under weak
390 purifying selection but not in termite *Vg1-2* clades (Fig. S2, Table S4). This is not contradicted with
391 the evidence in the bumble bee genus *Bombus*, which indicated that *Vg-like* genes showed a general
392 tendency of relaxation of purifying selection (Zhao *et al*., 2021). Considering the inferred phylogenetic
393 relationships of insect *Vg* genes, it is possible that termite *Vg4* has diverged less from the outgroups
394 compared to the other termite *Vg1-3* genes.
395
396 Different *Vg* gene expression patterns in females between reproductive and non-reproductive periods
397
398 Novel Vg roles may have been acquired to regulate caste-specific physiological and behavioural traits
399 during the course of hymenopteran social evolution (Robinson *et al*., 2005; Smith *et al*., 2008). The *P.*
400 *barbatus* and *S. invicta* genomes possess two and four *Vg* genes, respectively (Wurm *et al*., 2011;
401 Corona *et al*., 2013), suggesting that there were several duplication events that produced multiple *Vg*
402 genes in ants (Corona *et al*., 2013). Moreover, caste-specific *Vg* gene expression was associated with
403 the tasks performed by a queen and workers in honey bees and ants (Amdam *et al*., 2003; Corona *et*
404 *al*., 2013). The present study showed that the expression levels of *RsVg1-3* genes were particularly
405 high in female reproductives (queens) (Fig. 2). Furthermore, the expression patterns of *RsVg1-3* and
406 *RsVg4* were clearly different, and non-linear regressions observed in *RsVg1-3* were similar to the

407 reproductive cycles of queens (Fig. S4). In a previous study, the expression patterns of the *RsVg* gene
408 (*RsVg1* in this study) were related to the number of eggs in each colony and juvenile hormone (JH)
409 titre changes in queens at each period (i.e. the 'stages I–V' shown in Maekawa *et al*., 2010). JH-related
410 *RsVg1* expression has also been observed during the differentiation of secondary reproductives in *R.*
411 *speratus* (Saiki *et al*., 2015). However, the expression levels of *RsVg4* were higher in workers than in
412 reproductives (Fig. 2). In addition, *RsVg4* was highly expressed in nymphs and alates compared to
413 fertile queens (Fig. 4d). Since low JH levels are required for alate differentiation in termites (Cornette
414 *et al*., 2008; Oguchi *et al.*, 2021), similar JH-related *RsVg4* expression changes are unlikely. In
415 summary, the present results suggest that *RsVg1-3* are involved in the regulation of female
416 vitellogenesis, while *RsVg4* has a distinct role in non-reproductive individuals (see below). Further
417 analysis is required to determine whether the *RsVg1-3* genes have different subfunctions for
418 vitellogenesis in termites. Recently, Elsner et al. (2021) discovered that three queen-biased *Vgs* were
419 highly expressed in builders compared to foragers (all females) of the fungus-growing termite
420 *Macrotermes bellicosus*. To clarify the caste, sex and age specificity of termite *Vg1-3* gene expressions
421 should also be clarified in the future.
422
423 The potential roles of *RsVg4*
424
425 In insects, the *Vg* gene is mainly expressed in the female fat body for vitellogenesis (Tufail and Takeda,

426 2008). *RsVg1* mRNA was localised in the fat body of termite queens (Fig. 5). Similarly, *RsVg4* mRNA
427 was localised in the fat body of workers, but no positive signals were observed in other tissues (Fig.
428 5). Although the function of *RsVg4* is still unclear, there are several related reports on *Vgs* within the
429 *Vg-like-B* and closely related clades, which allow us to infer the function of *RsVg4*. For example, in *D.*
430 *melanogaster*, *Crossveinless d* mRNA (OTU No. [75] in Fig. 1) was expressed in the fat body of larval
431 stages in both sexes, and the product was transported into the wing tissues through haemolymph at
432 pupal stages to regulate the development of the posterior cross vein (Chen *et al*., 2012). However, it
433 should be noted that rhinotermitid workers are completely wingless, and *RsVg4* may not have any
434 developmental functions for wing formation. In the brown citrus aphid *A. citricidus*, *AcVg* (OTU No.
435 [73]) was also expressed in the fat body of both adults and nymphs, and the silencing of *AcVg*
436 expression resulted in a decrease in survival times and reproductive periods after alate emergence
437 (Shang *et al*., 2017). Moreover, the expression levels of the hymenopteran *Vg-like-B* were different
438 among *Formica* species with queen-biased expression in *F. aquilonia* and worker-biased expression in
439 *F. pratensis* (Morandin *et al*., 2014). Although no functional analyses have been performed in any
440 *Formica* species, *Vg-like-B* (OTU No. [71]) expression was upregulated in the fat body of honey bee
441 *A. mellifera* workers to protect against oxidative stress induced by artificial paraquat treatments
442 (Salmela *et al*., 2016). Termite workers are likely more exposed to oxidative stress compared to other
443 castes, because of the high oxygen concentration from the existence of gut symbionts (Brune *et al*.,
444 1995). Thus, one possible explanation for the role of *RsVg4* is protection from oxidative stress. To

445 further clarify the specific roles, RNA interference (RNAi) analysis should be performed in termites.

446

447 Conclusion

448

449 We clarified that different evolutionary trajectories could be found in *Vg* gene duplication and
450 functionalisation between termites and social hymenopterans. In social hymenopterans, particularly in
451 ants, *Vg* genes were duplicated and subsequently functionalised within a conventional clade. However,
452 in termites, *Vg* gene functionalisation after duplication might not have occurred within this clade (at
453 least between termite *Vg1* and *Vg2*). In contrast, termite *Vg4*, closely related to the *Vg-like-B* clade of
454 social hymenopterans, might have functionalised during the course of termite evolution. It might play
455 various crucial roles in termites, notwithstanding vitellogenesis of non-reproductive individuals.
456 Functional analyses should be conducted to determine whether hymenopteran *Vg-like-B* and termite
457 *Vg4* homologs contribute to social evolution in insects.

458

459 Acknowledgements

460

461 We are grateful to T. Tsuchida who helped with qPCR experiments, D. Watanabe and R. Saiki for
462 valuable discussions, S. Kawamura, T. Hatano, and M. Takizawa for their assistance during laboratory
463 work. We thank the University of the Ryukyus Center for Research Advancement and Collaboration

for the use of the computer server. Computational resources were also provided by the Data Integration and Analysis Facility, National Institute for Basic Biology. This study was partly supported by a grant for Scientific Research (Nos. 25128705, JP16K07511, and JP19H03273 to KM) from the Japan Society for the Promotion of Science, by research grants promoted by the University of the Ryukyus: Research Project Promotion Grant (Strategic Research Grant) (No. 16SP01302), "the Spatiotemporal Genomics Project", and the NIBB Collaborative Research Program (No. 20-323 and 21-220).

Author contributions

HY and KM planned and designed the experiments. HY, YM, RHS and RS collected the samples. HY, ShoS and NK performed the experiments. HY, YM, ShuS and YH analysed the data. HY, YM, ShuS and KM wrote and revised the manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that there are no conflict of interests.

Data availability statement

We deposited our raw sequence data to NCBI. Accession Numbers are stated in the manuscript.

References

Amdam, G.V., Norberg, K., Fondrk, M.K. and Page, R.E. (2004) Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proc Natl Acad Sci USA* **101**, 11350–11355.

Andersen, C., Jensen, J. and Orntoft, T. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and cancer data sets. *Cancer Res* **64**, 5245–5250.

Avarre, J.-C., Lubzens, E. and Babin, P.J. (2007) Apolipocrustacein, formerly vitellogenin, is the major egg yolk precursor protein in decapod crustaceans and is homologous to insect apolipophorin II/I and vertebrate apolipoprotein B. *BMC Evol Biol* **7**, 3.

Brune, A., Emerson, D. and Breznak, J.A. (1995) The termite gut microflora as an oxygen sink: Microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Appl Environ Microbiol* **61**, 2681–2687.

Capella-Gutierrez, S., Silla-Martinez, J.M. and Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973.

Chen, J., Honeyager, S.M., Schleede, J., Avanesov, A., Laughon, A. and Blair, S.S. (2012) Crossveinless d is a vitellogenin-like lipoprotein that binds BMPs and HSPGs, and is required for normal BMP signaling in the *Drosophila* wing. *Development* **139**, 2170–2176.

Cornette, R., Gotoh, H., Koshikawa, S. and Miura, T. (2008) Juvenile hormone titers and caste differentiation in the damp-wood termite *Hodotermopsis sjostedti* (Isoptera, Termopsidae). *J Insect Physiol* **54**, 922–930.

Corona, M., Libbrecht, R., Wurm, Y., Riba-Grognuz, O., Studer, R.A. and Keller, L. (2013) Vitellogenin underwent subfunctionalization to acquire caste and behavioral specific expression in the harvester ant *Pogonomyrmex barbatus*. *PLoS Genet* **9**, e1003730.

Elsner, D., Hartfelder, K. and Korb, J. (2021) Molecular underpinnings of division of labour among workers in a socially complex termite. *Sci Rep* **11**, 18269.

521 Friedman, D.A. and Gordon, D.M. (2016) Ant genetics: reproductive physiology, worker morphology,
522 and behavior. *Annu Rev Neurosci* **39**, 41–56.
523
524 Harrison, M.C., Jongepier, E., Robertson, H.M., Arning, N., Bitard-feildel, T., Chao, H. *et al.* (2018)
525 Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nat Ecol Evol* **2**, 557–566.
526
527 Hayashi, Y., Maekawa, K., Nalepa, C.A., Miura, T. and Shigenobu, S. (2017) Transcriptome
528 sequencing and estimation of DNA methylation level in the subsocial wood-feeding cockroach
529 *Cryptocercus punctulatus* (Blattodea: Cryptocercidae). *Appl Entomol Zool* **52**, 643–651.
530
531 Hayward, A., Takahashi, T., Bendena, W.G., Tobe, S.S. and Hui, J.H.L. (2010) Comparative genomic
532 and phylogenetic analysis of vitellogenin and other large lipid transfer proteins in metazoans. *FEBS*
533 *Lett* **584**, 1273–1278.
534
535 Hojo, M., Toga, K., Watanabe, D., Yamamoto, T. and Maekawa, K. (2011) High-level expression of
536 the geranylgeranyl diphosphate synthase gene in the frontal gland of soldiers in *Reticulitermes speratus*
537 (Isoptera: Rhinotermitidae). *Arch Insect Biochem Physiol* **77**, 17–31.
538
539 Inward, D., Beccaloni, G. and Eggleton, P. (2007) Death of an order: a comprehensive molecular

540 phylogenetic study confirms that termites are eusocial cockroaches. *Biol Lett* **3**, 331–335.

541

542 Ishitani, K. and Maekawa, K. (2010) Ovarian development of female-female pairs in the termite,

543 *Reticulitermes speratus*. *J Insect Sci* **10**, 194.

544

545 Itakura, S., Yoshikawa, Y., Togami, Y., and Umezawan K. (2020) Draft genome sequence of the termite,

546 *Coptotermes formosanus*: Genetic insights into the pyruvate dehydrogenase complex of the termite. *J*

547 *Asia-Pac Entomol* **23**, 666–674.

548

549 Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7:

550 improvements in performance and usability. *Mol Biol Evol* **30**, 772–780.

551

552 Kohlmeier, P., Feldmeyer, B. and Foitzik, S. (2018) *Vitellogenin-like A*-associated shifts in social cue

553 responsiveness regulate behavioral task specialization in an ant. *PLoS Biol* **16**, e2005747.

554

555 Korb, J. and Heinze, J. (2004) Multilevel selection and social evolution of insect societies.

556 *Naturwissenschaften* **91**, 291–304.

557

558 Korb, J. (2016) Genes underlying reproductive division of labor in termites, with comparisons to social

559 hymenoptera. *Front Ecol Evol* **4**, 45.

560

561 Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis
562 version 7.0 for bigger datasets. *Mol Biol Evol* **33**, 1870–1874.

563

564 Lo, N., Tokuda, G., Watanabe, H., Rose, H., Slaytor, M., Maekawa, K., Bandi, C. and Noda, H. (2000)
565 Evidence from multiple gene sequences indicates that termites evolved from wood-feeding
566 cockroaches. *Curr Biol* **10**, 801–804.

567

568 Maekawa, K., Ishitani, K., Gotoh, H., Cornette, R. and Miura, T. (2010) Juvenile hormone titre and
569 vitellogenin gene expression related to ovarian development in primary reproductives, as compared to
570 nymphs and nymphoid reproductives of the termite *Reticulitermes speratus*. *Physol Entomol* **35**, 52–
571 58.

572

573 Miyazaki, S., Yoshimura, M., Saiki, R., Hayashi, Y., Kitade, O., Tsuji, K. and Maekawa, K. (2014)
574 Intracolonial genetic variation affects reproductive skew and colony productivity during colony
575 foundation in a parthenogenetic termite. *BMC Evol Biol* **14**, 177.

576

577 Miyazaki, S., Shimoji, H., Suzuki, R., Chinushi, I., Takayanagi, H., Yaguchi, H., Miura, T. and

578 Maekawa, K. (2021) Expressions of conventional *vitellogenin* and *vitellogenin-like A* in worker brains
579 are associated with a nursing task in a ponerine ant. *Insect Mol Biol* **30**, 113–121.
580
581 Morandin, C., Havukainen, H., Kulmuni, J., Dhaygude, K., Trontti, K. and Helanterä, H. (2014) Not
582 only for egg yolk-functional and evolutionary insights from expression, selection, and structural
583 analyses of formica ant vitellogenins. *Mol Biol Evol* **31**, 2181–2193.
584
585 Morandin, C., Hietala, A. and Helanterä, H. (2019) Vitellogenin and vitellogenin-like gene expression
586 patterns in relation to caste and task in the ant *Formica fusca*. *Insectes Soc* **66**, 519–531.
587
588 Nelson, C.M., Ihle, K.E., Fondrk, M.K., Page, R.E. and Amdam, G.V. (2007) The gene vitellogenin
589 has multiple coordinating effects on social organization. *PLoS Biol* **5**, e62.
590
591 Oguchi, K., Maekawa, K. and Miura, T. (2021) Regulatory mechanisms underlying the differentiation
592 of neotenic reproductives in termites: Partial release from arrested development. *Front Ecol Evol* **9**,
593 635552.
594
595 O'Shea-Wheller, T.A., Hunt, E.R. and Sasaki, T. (2021) Functional heterogeneity in superorganisms:
596 emerging trends and concepts. *Ann Entomol Soc Am* **114**, 562–574.

597

598 Poulsen, M., Hu, H., Li, C., Chen, Z., Xu, L., Otani, S. *et al.* (2014) Complementary symbiont
599 contributions to plant decomposition in a fungus-farming termite. *Proc Natl Acad Sci USA* **111**, 14500–
600 14505.

601

602 R Core Team. (2020) R: A language and environment for statistical computing. R Foundation for
603 Statistical Computing, Vienna, Austria.

604

605 Raikhel, A.S. and Dhadialla, T.S. (1992) Accumulation of yolk proteins in insect oocytes. *Annu Rev*
606 *Entomol* **37**, 217–251.

607

608 Robinson, M.D. and Oshlack A. (2010) A scaling normalization method for differential expression
609 analysis of RNA- seq data. *Genome Biol* **11**, R25.

610

611 Robinson, G.E., Grozinger, C.M. and Whitfield, C.W. (2005) Sociogenomics: social life in molecular
612 terms. *Nat Rev Genet* **6**, 257–270.

613

614 Roisin, Y. (2000) Diversity and evolution of caste patterns. In: *Termites: Evolution, Sociality,*
615 *Symbioses, Ecology* (Abe, T., Bignell, D.E. and Higashi, M. eds), pp. 95–119. Kluwer Academic Press,

616 Dordrecht.

617

618 Roy-Zokan, E.M., Cunningham, C.B., Hebb, L.E., McKinney, E.C. and Moore, A.J. (2015)
619 Vitellogenin and vitellogenin receptor gene expression is associated with male and female parenting
620 in a subsocial insect. *Proc. R. Soc. B* **282**, 20150787.

621

622 Richards, M.H. (2019) *Vitellogenin* and *vitellogenin-like* genes: not just for egg production. *Insectes*
623 *Soc* **66**, 505–506.

624

625 Saiki, R. and Maekawa, K. (2011) Imaginal organ development and vitellogenin gene expression
626 changes during the differentiation of nymphoids of the termite *Reticulitermes speratus*. *Sociobiology*
627 58: 499–511.

628

629 Saiki, R., Gotoh, H., Toga, K., Miura, T. and Maekawa, K. (2015) High juvenile hormone titre and
630 abdominal activation of JH signalling may induce reproduction of termite neotenics. *Insect Mol Biol*
631 **24**, 432–441.

632

633 Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing
634 phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

635

636 Salmela, H., Stark, T., Stucki, D., Fuchs, S., Freitak, D., Dey, A., Kent, C.F., Zayed, A., Dhaygude, K.,
637 Hokkanen, H. and Sundström, L. (2016) Ancient duplications have led to functional divergence of
638 vitellogenin-like genes potentially involved in inflammation and oxidative stress in honey bees.
639 *Genome Biol Evol* **8**, 495–506.

640

641 Shang, F., Niu, J.-Z., Ding, B.-Y., Zhang, Q., Ye, C., Zhang, W., Smagghe, G. and Wang, J.-J. (2017)
642 Vitellogenin and its receptor play essential roles in the development and reproduction of the brown
643 citrus aphid, *Aphis* (*Toxoptera*) *citricidus*. *Insect Mol Biol* **27**, 221–233.

644

645 Shigenobu, S., Hayashi, Y., Watanabe, D., Tokuda, G., Hojo, M., Toga, K., *et al.* (2022) Genomic and
646 transcriptomic analyses of the subterranean termite *Reticulitermes speratus*: gene duplication
647 facilitates social evolution. *Proc Natl Acad Sci USA* **119**, e2110361119.

648

649 Shimada, K. and Maekawa, K. (2010) Changes in endogenous cellulase gene expression levels and
650 reproductive characteristics of primary and secondary reproductives with colony development of the
651 termite *Reticulitermes speratus* (Isoptera: Rhinotermitidae). *J Insect Physiol* **56**, 1118–1124.

652

653 Smith, C.R., Toth, A.L., Suarez, A.V. and Robinson, G.E. (2008) Genetic and genomic analyses of the

654 division of labour in insect societies. *Nat Rev Genet* **9**, 735–748.

655

656 Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large

657 phylogenies. *Bioinformatics* **30**, 1312–1313.

658

659 Terrapon, N., Li, C., Robertson, H.M., Ji, L., Meng, X., Booth, W. *et al.* (2014) Molecular traces of

660 alternative social organization in a termite genome. *Nat Commun* **5**, 3636.

661

662 Thorne, B.L. (1997) Evolution of eusociality in termites. *Annu Rev Ecol Syst* **28**, 27–54.

663

664 Toth, A.L. and Rehan, S.M. (2017) Molecular evolution of insect sociality: an eco-evo-devo

665 perspective. *Annu Rev Entomol* **62**, 419–442.

666

667 Tufail, M. and Takeda, M. (2008) Molecular characteristics of insect vitellogenins. *J Insect Physiol* **54**,

668 1447–1458.

669

670 Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Greurts, R. and Leunissen, J.A.M. (2007) Primer

671 3 Plus, an enhanced web interface to Primer 3. *Nucleic Acids Res* **35**, 71–74.

672

673 Vandesompele, J., Preter, K.D., Pattyn, F., Poppe, B., Roy, N.V., Paepe, A.D. and Speleman, F. (2002)
674 Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple
675 internal control genes. *Genome Biol* **3**, 0034:1–0034:11.
676
677 Wahli, W. (1988) Evolution and expression of Vitellogenin genes. *Trends Genet* **4**, 227–232.
678
679 Waterhouse, R.M., Tegenfeldt, F., Li, J., Zdobnov, E.M. and Kriventseva, E.V. (2013) OrthoDB: a
680 hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Res* **41**, D358–D365.
681
682 Weesner, F.M. (1969) The reproductive system. In: *Biology of Termites, vol. I.* (Krishna, K. and
683 Weesner, F.M. eds.), pp. 125–160. Academic Press, New York.
684
685 Weil, T., Korb, J. and Rehli, M. (2007) Comparison of queen-specific gene expression in related lower
686 termite species. *Mol Biol Evol* **26**, 1841–1850.
687
688 Wilson, E.O. (2000) Sociobiology: the new synthesis, Twenty-fifth anniversary. The Belknap Press of
689 Harvard University Press, Cambridge.
690
691 Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B.G. *et al.* (2011) The genome

692 of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci USA* **108**, 5679–5684.

693

694 Yang, Z. (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**, 1586–

695 1591.

696

697 Zhao, F., Morandin, C., Jiang, K., Su, T., He, B., Lin, G. and Huang, Z. (2021) Molecular evolution of

698 bumble bee vitellogenin and vitellogenin-like genes. *Ecol Evol* **11**, 8983–8992.

699

700 Figure captions

701

702 Fig. 1. Molecular phylogenetic tree of *Vg* genes in oviparous animals.

703 Phylogenetic tree inferred from Vg amino acid sequences (228-1070 aa) based on the maximum

704 likelihood (ML) analysis. Apolipocrustacein (ApoCr) and insect apolipophorin II/I (apoLp-II/I) were

705 used as outgroups, according to previous reports (Avarre *et al*., 2007; Hayward *et al*., 2010). Bootstrap

706 values (above 50%) from ML (1000 times) and neighbor-joining (NJ, 5000 times) analyses are shown

707 above and below the branches to indicate the level of support for each node, respectively. The asterisk

708 indicates a branch for which high bootstrap values could not be obtained (below 50%). The numbers

709 in parentheses after species names indicate the OTU numbers shown in Table S2.

710

711 Fig. 2. Expression values (the reads per kilobase of exon per million mapped sequence reads; RPKM)
712 of four *Vg* genes among castes in *R. speratus*.
713 The graphs indicate the expression values of *RsVg* genes in (a) heads and (b) bodies (including thoraxes
714 and abdomens). PR: primary reproductive; S: soldier; W: worker. Asterisks indicate genes with
715 significant differential expression (GLMM analysis, *FDR<0.05, **FDR<0.01). Statistical results
716 (GLMM) are shown in Table S5.
717
718 Fig. 3. Numbers of eggs observed in the colony at the periods during incipient colony development.
719 The boxes and whiskers represent the median, quartiles, and ranges, respectively. Numbers in
720 parentheses indicate the number of colonies examined. Non-parametric tests were fitted to these data
721 based on the results of Levene's test (p = 4.438e-16). Different letters over the boxes denote significant
722 differences (Kruskal-Wallis test followed by Steel-Dwass test, $p < 0.05$).
723
724 Fig. 4. Relative expression levels of four *RsVg* genes using biological triplicates in female N6 nymphs,
725 alates, and reproductives (0.5–7.5 months after colony foundation).
726 Each value was normalised to the expression level of the *glutathione S transferase D1* homolog
727 (*RsGstD1*) (Table S3). The boxes and whiskers mean median, quartiles and range, respectively.
728 Parametric tests were fitted to these data based on the results of Levene's test (*RsVg1*, p = 0.1151;
729 *RsVg2*, p = 0.2338; *RsVg3*, p = 0.1316; *RsVg4*, p = 0.4887). Different letters above the bars indicate

730 significant differences (one-way ANOVA followed by Tukey's test, $p < 0.05$). N6: the 6th-stage nymph
731 (Fig. S1), Before moult: N6 just before the moult into alates; Alate: winged adults; Reproductive:
732 female reproductives collected each month after incipient colony establishment.
733
734 Fig. 5. Localisation of *RsVg* mRNA in *R. speratus*.
735 Parasagittal cryosections of a worker head (a, b) and abdomen (c, d) subjected to *in situ* hybridisation
736 with antisense (a, c) and sense (b, d) DIG-labelled *RsVg4* mRNA probe. Parasagittal cryosections of a
737 queen body subjected to *in situ* hybridisation with antisense (e) and sense (f) DIG-labelled *RsVg1*
738 mRNA probe. The signals were only detected with antisense probe treatments (arrowheads). Scale bars
739 indicate 100 μm (a-d) and 50 μm (e, f). Br: brain. Sp: sternal plate.
740
741 Supplementary Figure 1. Caste differentiation pathway of *R. speratus*.
742 They have bifurcated developmental pathways. There are at least five worker stages in the apterous
743 line. The 4th-stage (W4) and 5th-stage (W5) workers are able to differentiate into soldiers via the
744 presoldier stage. There were six nymphal stages on the imaginal line. The 6th-stage (N6) nymphs
745 differentiate into winged alates, and de-winged females and males become primary reproductives
746 (queen and king). In both lines, W2–W5 workers and N3–N6 nymphs occasionally moult into
747 secondary reproductives.
748

749 Supplementary Figure 2. Molecular evolutionary analysis of *Vg* homologs by the PAML branch test in
750 (a) *Vg1*, (b) *Vg2*, and (c) *Vg4* clades of the phylogeny shown in Fig. 1. The focal branches are indicated
751 with a bold line, and an asterisk shows the significant difference with other branches in each clade (p
752 < 0.05, Table S4). In the *Vg4* clade, the value above and below each node indicates an independent dN
753 / dS (ω) ratio obtained from a free-ratio model M1.
754
755 Supplementary Figure 3. The domain architecture of four Vgs in the termite *R. speratus*.
756 Black rods indicate full Vg amino acid sequences. Three domains; lipid transport protein, N-terminal
757 (IPR001747), vitellinogen, open beta-sheet (IPR015255), and von Willebrand factor, type D domain
758 (IPR001846), are highlighted. Scaffold numbers of the gene model Rspe OGS1.0 (Shigenobu *et al*.,
759 2022) are indicated.
760
761 Supplementary Figure 4. Polynomial regression (orange line) of egg number changes and *RsVg* gene
762 expression patterns.
763 The changes of egg numbers are significantly fitted by cubic equations during colony developments
764 (a). Expression patterns of *RsVg1-3* are significantly fitted by cubic equations (b-d), whereas those of
765 *RsVg4* are significantly fitted in quadratic equations (e). Statistical results (ANOVAs) are shown in
766 Table S6.
767

768 Supplementary data. Alignment of amino acid sequences used for molecular phylogenetic analysis.

769

Table S1. Primer sequences used in this study.

	Gene name	Forward sequence (5' - 3')	Reverse sequence (5' - 3')
3' - RACE	*RsVg3*	GTCTCAACGTGGTACGACCCCAGCTT	-
	RsVg3_1	GACGCCAACAACCTGTACTA	AGCGTATTCGTGACTGTTCA
	RsVg3_2	TGGACGTAGATATGCTGTAT	GGTAGTTGTCCAAGTTCAC
	RsVg3_3	CGTATTACTCAACTACACGAC	GTAGGGTAACAACGAATAAC
	RsVg3_4	CAGGAGCACATCAAAAACGA	CTATGAGCGAATATGTGAAG
real-time qPCR	*RsVg1*	CCTACATGCGTTGTTGATGG	TGACGACTATGCACTCCAGC
	RsVg2	GCGAAATGGAGAACCCTAAT	TCACTTCGTCGATTGCTGTA
	RsVg3	ATGAGCTGATCCAGAGTTCG	ATGAGCGTATCGCGTCTTAG
	RsVg4	AGAAGCCATGACACACCAAT	CTGCCTTCACAGACAGACCT
	EF1	GGTGATGCGGCTATTGTTAACC	GTGGTGGGAATTCTGAGAAAGATT
	β-Actin	AGCGGGAAATCGTCCGTGAC	CAATGGTGATGACCTGCCCAT
	NADH-dh	GCTGGGGGGGTTATTCATTCCAT	GGCATACCACAAAGGGCAAAA
	GstD1	GCTGTTGGTGTGGATTTGAA	GTATGCTGCGGGTTCATCTT
	EIF-1	ATGGTAGGCTTGAAGCGATG	TTTGCATCCTGGTAGTCACG
	RPS18	ACTCTCAGCTCACATCCAGT	CCTCAGGCCCCAATAATGTC
mRNA in situ hybridization	*RsVg1*	ATCCAAAACCGAACCAAC	TACGAGCACCGCTACATT
	RsVg4	TACGCTCACCAGATGCTGTA	GAGCCAGTGAGTGCAAAGTT

Table S2. The species used for the molecular phylogenetic analysis.

OTU nc	Species	Database and reference	Accession number or gene ID
[1]	*Cryptotermes secundus*	Csec_1.0 (Harrison et al., 2018)	XP_023706879
[2]	*Zootermopsis nevadensis*	Znev OGS v2.2 (Terrapon et al., 2014)	Znev_07681
[3]	*Reticulitermes speratus*	Rspe OGS1.0 (Shigenobu et al., 2021)	RS000616
[4]	*Macrotermes natalensis*	Mnat_gene_v1.2 (Poulsen et al., 2014)	MN002567
[5]	*Cryptocercus punctulatus*	ESM1 (DRA004598, Hayashi et al., 2017)	comp77743_c0_seq1
[6]	*Macrotermes natalensis*	Mnat_gene_v1.2 (Poulsen et al., 2014)	MN002568
[7]	*Coptotermes formosanus*	CopFor1.0 (Itakura et al., 2020)	GFG39872
[8]	*Reticulitermes speratus*	Rspe OGS1.0 (Shigenobu et al., 2021)	RS000610
[9]	*Cryptotermes secundus*	Csec_1.0 (Harrison et al., 2018)	XP_023706896
[10]	*Zootermopsis nevadensis*	Znev OGS v2.2 (Terrapon et al., 2014)	Znev_07682
[11]	*Cryptocercus punctulatus*	ESM1 (DRA004598, Hayashi et al., 2017)	comp194989_c0_seq1
[12]	*Blattella germanica*	OrthoDB (Waterhouse et al., 2013)	BGER020452
[13]	*Rhyparobia maderae*	NCBI database	BAB19327
[14]	*Rhyparobia maderae*	NCBI database	BAD72597
[15]	*Periplaneta americana*	NCBI database	BAB32673
[16]	*Periplaneta americana*	NCBI database	BAA86656
[17]	*Blattella germanica*	OrthoDB (Waterhouse et al., 2013)	BGER000125
[18]	*Graptopsaltria nigrofuscata*	NCBI database	BAA85987
[19]	*Pediculus humanus*	OrthoDB (Waterhouse et al., 2013)	PHUM524870
[20]	*Pediculus humanus*	OrthoDB (Waterhouse et al., 2013)	PHUM524850
[21]	*Anthonomus grandis*	NCBI database	AAA27740
[22]	*Nicrophorus vespilloides*	OrthoDB (Waterhouse et al., 2013)	XP_017778650
[23]	*Nicrophorus vespilloides*	OrthoDB (Waterhouse et al., 2013)	XP_017781017
[24]	*Riptortus clavatus*	NCBI database	AAB72001
[25]	*Plautia stali*	NCBI database	BAA88076
[26]	*Plautia stali*	NCBI database	BAA88075
[27]	*Plautia stali*	NCBI database	BAA88077
[28]	*Ephemera danica*	OrthoDB (Waterhouse et al., 2013)	EDAN008490
[29]	*Pimpla nipponica*	NCBI database	AAC32024
[30]	*Apis mellifera*	NCBI database	CAD56944
[31]	*Pogonomyrmex barbatus*	NCBI database	XP_025073621
[32]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	XP_025987667
[33]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	XP_025987667
[34]	*Harpegnathos saltator*	NCBI database	EFN86099
[35]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	NP_001291514
[36]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	NP_001291513
[37]	*Pogonomyrmex barbatus*	NCBI database	XP_025073621
[38]	*Formica exsecta*	NCBI database	AII96912
[39]	*Camponotus floridanus*	NCBI database	EFN64902
[40]	*Aedes aegypti*	NCBI database	AAA18221
[41]	*Lymantria dispar*	NCBI database	AAB03336
[42]	*Bombyx mori*	OrthoDB (Waterhouse et al., 2013)	NP_001037309
[43]	*Samia cynthia pryeri*	NCBI database	BAD91196
[44]	*Danaus plexippus*	NCBI database	OWR44310
[45]	*Coptotermes formosanus*	CopFor1.0 (Itakura et al., 2020)	GFG36045
[46]	*Macrotermes natalensis*	Mnat_gene_v1.2 (Poulsen et al., 2014)	MN000066
[47]	*Reticulitermes speratus*	Rspe OGS1.0 (Shigenobu et al., 2021)	RS013977
[48]	*Cryptotermes secundus*	Csec_1.0 (Harrison et al., 2018)	XP023712833
[49]	*Zootermopsis nevadensis*	Znev OGS v2.2 (Terrapon et al., 2014)	Znev_08605
[50]	*Apis mellifera*	NCBI database	XP_001122505
[51]	*Harpegnathos saltator*	NCBI database	EFN86751
[52]	*Camponotus floridanus*	NCBI database	EFN73216
[53]	*Formica exsecta*	NCBI database	AII96915
[54]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	XP_011172130
[55]	*Pogonomyrmex barbatus*	NCBI database	XP_011631341
[56]	*Penaeus monodon*	NCBI database	ABW77320
[57]	*Penaeus monodon*	NCBI database	AAF19002
[58]	*Daphnia magna*	NCBI database	BAD05137
[59]	*Zootermopsis nevadensis*	Znev OGS v2.2 (Terrapon et al., 2014)	Znev_06771
[60]	*Macrotermes natalensis*	Mnat_gene_v1.2 (Poulsen et al., 2014)	MN000928
[61]	*Reticulitermes speratus*	Rspe OGS1.0 (Shigenobu et al., 2021)	RS011073
[62]	*Coptotermes formosanus*	CopFor1.0 (Itakura et al., 2020)	GFG34789
[63]	*Cryptotermes secundus*	Csec_1.0 (Harrison et al., 2018)	XP_023710404
[64]	*Cryptocercus punctulatus*	ESM1 (DRA004598, Hayashi et al., 2017)	comp89049_c0_seq6
[65]	*Blattella germanica*	OrthoDB (Waterhouse et al., 2013)	BGER006422
[66]	*Harpegnathos saltator*	NCBI database	EFN87799
[67]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	XP_011162482

[68]	*Pogonomyrmex barbatus*	NCBI database	XP_011629719
[69]	*Camponotus floridanus*	NCBI database	EFN61573
[70]	*Formica exsecta*	NCBI database	AII96914
[71]	*Apis mellifera*	NCBI database	XP_006561178
[72]	*Diuraphis noxia*	NCBI database	XP_015366382
[73]	*Aphis citricidus*	NCBI database	AVP41182
[74]	*Aphis medicaginis*	NCBI database	AGC97434
[75]	*Drosophila melanogaster*	OrthoDB (Waterhouse et al., 2013)	NP_732076
[76]	*Nicrophorus vespilloides*	OrthoDB (Waterhouse et al., 2013)	XP_017777145
[77]	*Apis mellifera*	NCBI database	XP001121939
[78]	*Harpegnathos saltator*	NCBI database	EFN75537
[79]	*Formica exsecta*	NCBI database	AII96913
[80]	*Camponotus floridanus*	NCBI database	EFN69845
[81]	*Solenopsis invicta*	Ensembl Metazoa database based on the genome sequence (W	XP_011161539
[82]	*Pogonomyrmex barbatus*	NCBI database	XP_011644827
[83]	*Bombyx mori*	OrthoDB (Waterhouse et al., 2013)	XP_004926335
[84]	*Xenopus laevis*	NCBI database	P18709
[85]	*Gallus gallus*	NCBI database	P87498
[86]	*Caenorhabditis elegans*	NCBI database	CAA26849
[87]	*Caenorhabditis elegans*	NCBI database	CAA39669
[88]	*Caenorhabditis elegans*	NCBI database	CAA39670
[89]	*Locusta migratoria*	NCBI database	Q9U943
[90]	*Manduca sexta*	NCBI database	Q25490
[91]	*Anopheles gambiae*	NCBI database	XP_321226
[92]	*Penaeus monodon*	NCBI database	ABB89953
[93]	*Drosophila melanogaster*	NCBI database	Q9V496

Table S3. Stability values of reference genes in real-time qPCR analysis.

Gene symbol	Gene ID	Stability value	
		GeNorm	NormFinder
EF1-alfa	AB602838	0.883	0.340
NADH-dh	AB602837	1.059	0.534
beta-actin	AB520714	0.989	0.486
*GstD1**	RS001168	0.850	0.286
EIF-1	RS005199	0.912	0.418
RPS18	RS015150	1.336	0.828

* *GstD1* was selected by GeNorm and NormFinder due to the lowest stability values among six genes analyzed.

Table S4. The estimated dN / dS ratios based on Model 0 and the results of $\chi 2$ test.

		dN/dS ratio	2ΔlnL	d.f.	*p* value
Model 0 vs Model 1	Termite Vg1	0.3572	23.55872	7	0.001361654
	Termite Vg2	0.3910	31.607706	9	0.000232726
	Termite Vg4	0.3589	41.457698	9	4.12813E-06
Model 0 vs Model 2	Termite Vg1	0.3572	1.916162	1	0.166280377
	Termite Vg2	0.3910	0.382294	1	0.536378076
	Termite Vg4	0.3589	8.316270	1	0.003929152

Table S5. Summary of the results of the generalized linear mixed model (GLMM) analysis on RNA-seq data using the R package glmmSeq v0.1.0.

Body parts using RNA-seq	Gene name	Gene ID	Dispersion	AIC	logLik	X.Intercept.	CasteSoldier	CasteWorker	SexMale	CasteSoldier.SexMale	CasteWorker.SexMale	Chisq_Caste	Chisq_Sex	Chisq_Caste.Sex	P_Caste	P_Sex	P_Caste.Sex	q_Caste	q_Sex	q_Caste.Sex
Head	*Rs1g1*	RS000616	0.092440347	69.51950118	-26.75975059	4.965986905	-6.340888413	-5.929603223	-4.393977138	3.942202898	4.279081249	57.19834368	64.97947338	11.67740548	3.79785E-13	7.56823E-16	0.002912619	1.27813E-11	7.031E-12	0.999959964
	Rs1g2	RS000610	0.100514652	92.00559702	-38.00279851	5.771083004	-6.437928421	-7.292815217	-2.363601615	2.326764304	2.846771777	121.7027727	64.20680678	6.522405011	3.74E-27	1.12E-15	0.038342264	3.17E-25	7.03E-12	0.999959964
	Rs1g3	RS013977	0.2382322	38.43058185	-11.21529093	1.005628722	-3.085070264	-3.085070264	-2.571501349	2.571501349	2.571501349	4.701306416	2.093602091	1.430548039	0.095306887	0.147916949	0.489058082	0.300334409	0.999971164	0.999959964
	Rs1g4	RS011073	0.070476537	107.0676508	-45.5338254	1.95707444	0.341312556	0.825785472	-0.117493573	0.158784943	-0.030691925	11.19071702	0.158153901	0.190816692	0.003715067	0.690861893	0.909001672	0.023708815	0.999971164	0.999959964
Body (thorax and abdomen)	*Rs1g1*	RS000616	0.110010039	76.48448157	-30.24224078	5.41988602	-7.126803232	-6.67541230 7	-3.946609407	4.050069343	3.926246224	61.11515542	93.12199129	9.890539181	5.36E-14	4.92E-22	0.007116996	1.74E-12	2.12E-18	0.272783076
	Rs1g2	RS000610	0.131473795	105.0271683	-44.51358416	6.476783894	-7.451163839	-7.78011035	-2.078216518	2.52588054	2.699436784	158.9963222	40.0486997	7.090259579	2.98E-35	2.48E-10	0.028864876	7.43E-33	2.47E-07	0.617070078
	Rs1g3	RS013977	0.085968109	53.25117632	-18.62558816	3.069389033	-5.148830574	-4.491129747	-2.456276696	2.456276696	1.798575868	25.97812646	22.19486024	1.724635074	2.29E-06	2.46E-06	0.422182524	2.37E-05	0.001063745	0.999959435
	Rs1g4	RS011073	0.043817326	114.5213145	-49.26065724	2.338822098	0.189582119	0.72104712	0.1315356	0.145095626	-0.107433213	12.51501174	0.741346398	0.468774362	0.001916019	0.38922992	0.791055475	0.009029473	0.99982328	0.999959435

Table S6. Fitting polynomial models and comparison by ANOVAs.

	degree	Res.Df	RSS	Df	Sum of Sq	F	Pr (>F)
egg	1	135	626.98				
	2	134	560.67	1	66.318	24.2302	2.51E-06
	3	133	369.49	1	191.175	69.8485	7.79E-14
	4	132	361.28	1	8.207	2.9986	0.0857
RsVg1	1	31	2.3084				
	2	30	2.3083	1	0.00003	0.0005	0.9821
	3	29	1.958	1	0.3503	5.2397	0.0298
	4	28	1.8719	1	0.08612	1.2882	0.2660
RsVg2	1	31	1259.8				
	2	30	1214.85	1	44.948	1.2829	0.2670
	3	29	1017.23	1	197.625	5.6404	0.0246
	4	28	981.05	1	36.181	1.0326	0.3182
RsVg3	1	31	504.43				
	2	30	504.36	1	0.072	0.0049	0.9447
	3	29	421.37	1	82.99	5.6484	0.0246
	4	28	411.39	1	9.979	0.6792	0.4168
RsVg4	1	31	0.27563				
	2	30	0.22322	1	0.052408	7.1645	0.0121
	3	29	0.21214	1	0.011086	1.5154	0.2282

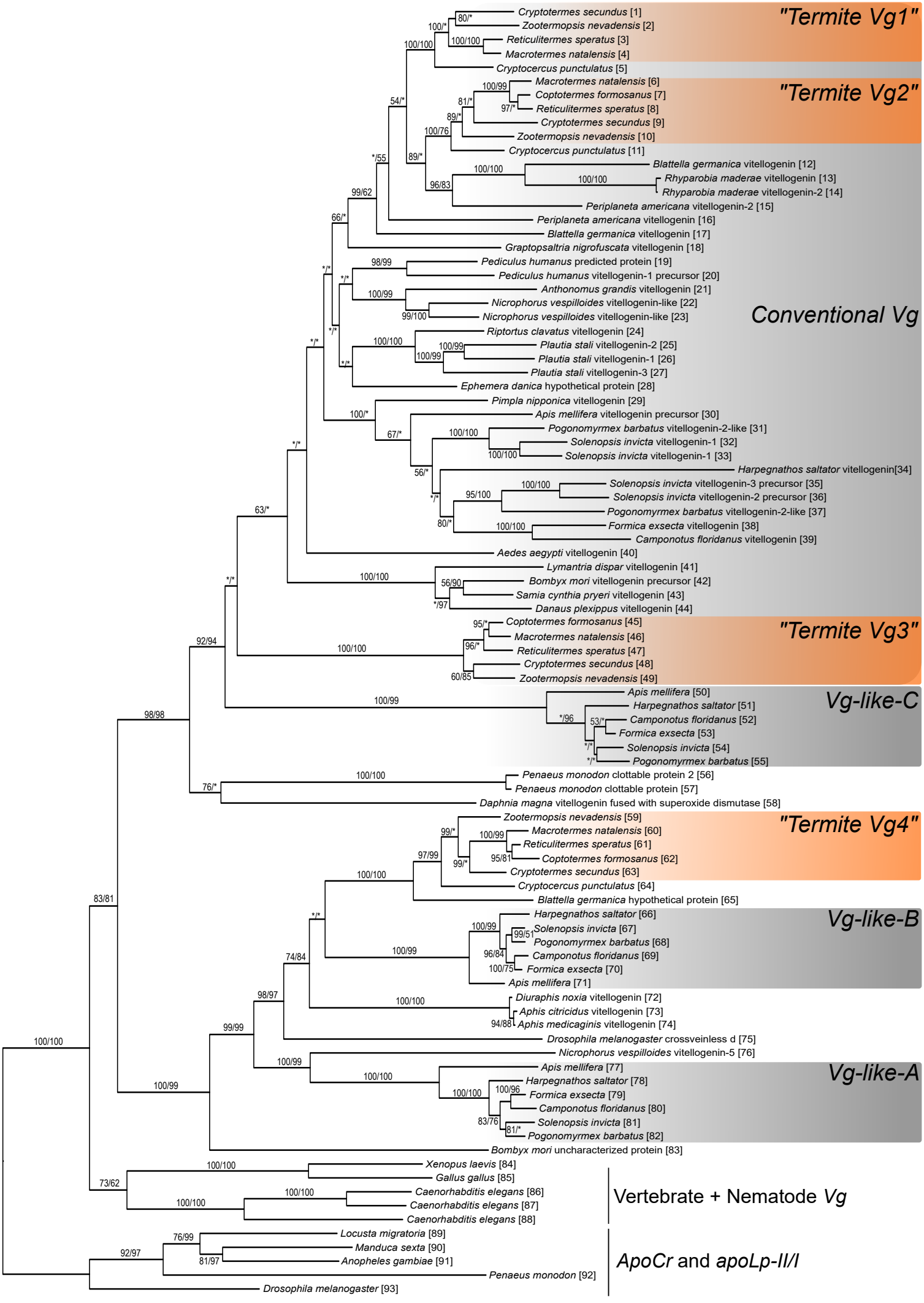

"Termite Vg1"
"Termite Vg2"
Conventional Vg
"Termite Vg3"
Vg-like-C
"Termite Vg4"
Vg-like-B
Vg-like-A
Vertebrate + Nematode Vg
ApoCr and apoLp-II/I
Cryptotermes secundus [1]
Zootermopsis nevadensis [2]
Reticulitermes speratus [3]
Macrotermes natalensis [4]
Cryptocercus punctulatus [5]
Macrotermes natalensis [6]
Coptotermes formosanus [7]
Reticulitermes speratus [8]
Cryptotermes secundus [9]
Zootermopsis nevadensis [10]
Cryptocercus punctulatus [11]
Blattella germanica vitellogenin [12]
Rhyparobia maderae vitellogenin [13]
Rhyparobia maderae vitellogenin-2 [14]
Periplaneta americana vitellogenin-2 [15]
Periplaneta americana vitellogenin [16]
Blattella germanica vitellogenin [17]
Graptopsaltria nigrofuscata vitellogenin [18]
Pediculus humanus predicted protein [19]
Pediculus humanus vitellogenin-1 precursor [20]
Anthonomus grandis vitellogenin [21]
Nicrophorus vespilloides vitellogenin-like [22]
Nicrophorus vespilloides vitellogenin-like [23]
Riptortus clavatus vitellogenin [24]
Plautia stali vitellogenin-2 [25]
Plautia stali vitellogenin-1 [26]
Plautia stali vitellogenin-3 [27]
Ephemera danica hypothetical protein [28]
Pimpla nipponica vitellogenin [29]
Apis mellifera vitellogenin precursor [30]
Pogonomyrmex barbatus vitellogenin-2-like [31]
Solenopsis invicta vitellogenin-1 [32]
Solenopsis invicta vitellogenin-1 [33]
Harpegnathos saltator vitellogenin[34]
Solenopsis invicta vitellogenin-3 precursor [35]
Solenopsis invicta vitellogenin-2 precursor [36]
Pogonomyrmex barbatus vitellogenin-2-like [37]
Formica exsecta vitellogenin [38]
Camponotus floridanus vitellogenin [39]
Aedes aegypti vitellogenin [40]
Lymantria dispar vitellogenin [41]
Bombyx mori vitellogenin precursor [42]
Samia cynthia pryeri vitellogenin [43]
Danaus plexippus vitellogenin [44]
Coptotermes formosanus [45]
Macrotermes natalensis [46]
Reticulitermes speratus [47]
Cryptotermes secundus [48]
Zootermopsis nevadensis [49]
Apis mellifera [50]
Harpegnathos saltator [51]
Camponotus floridanus [52]
Formica exsecta [53]
Solenopsis invicta [54]
Pogonomyrmex barbatus [55]
Penaeus monodon clottable protein 2 [56]
Penaeus monodon clottable protein [57]
Daphnia magna vitellogenin fused with superoxide dismutase [58]
Zootermopsis nevadensis [59]
Macrotermes natalensis [60]
Reticulitermes speratus [61]
Coptotermes formosanus [62]
Cryptotermes secundus [63]
Cryptocercus punctulatus [64]
Blattella germanica hypothetical protein [65]
Harpegnathos saltator [66]
Solenopsis invicta [67]
Pogonomyrmex barbatus [68]
Camponotus floridanus [69]
Formica exsecta [70]
Apis mellifera [71]
Diuraphis noxia vitellogenin [72]
Aphis citricidus vitellogenin [73]
Aphis medicaginis vitellogenin [74]
Drosophila melanogaster crossveinless d [75]
Nicrophorus vespilloides vitellogenin-5 [76]
Apis mellifera [77]
Harpegnathos saltator [78]
Formica exsecta [79]
Camponotus floridanus [80]
Solenopsis invicta [81]
Pogonomyrmex barbatus [82]
Bombyx mori uncharacterized protein [83]
Xenopus laevis [84]
Gallus gallus [85]
Caenorhabditis elegans [86]
Caenorhabditis elegans [87]
Caenorhabditis elegans [88]
Locusta migratoria [89]
Manduca sexta [90]
Anopheles gambiae [91]
Penaeus monodon [92]
Drosophila melanogaster [93]
0.5

Figure 1

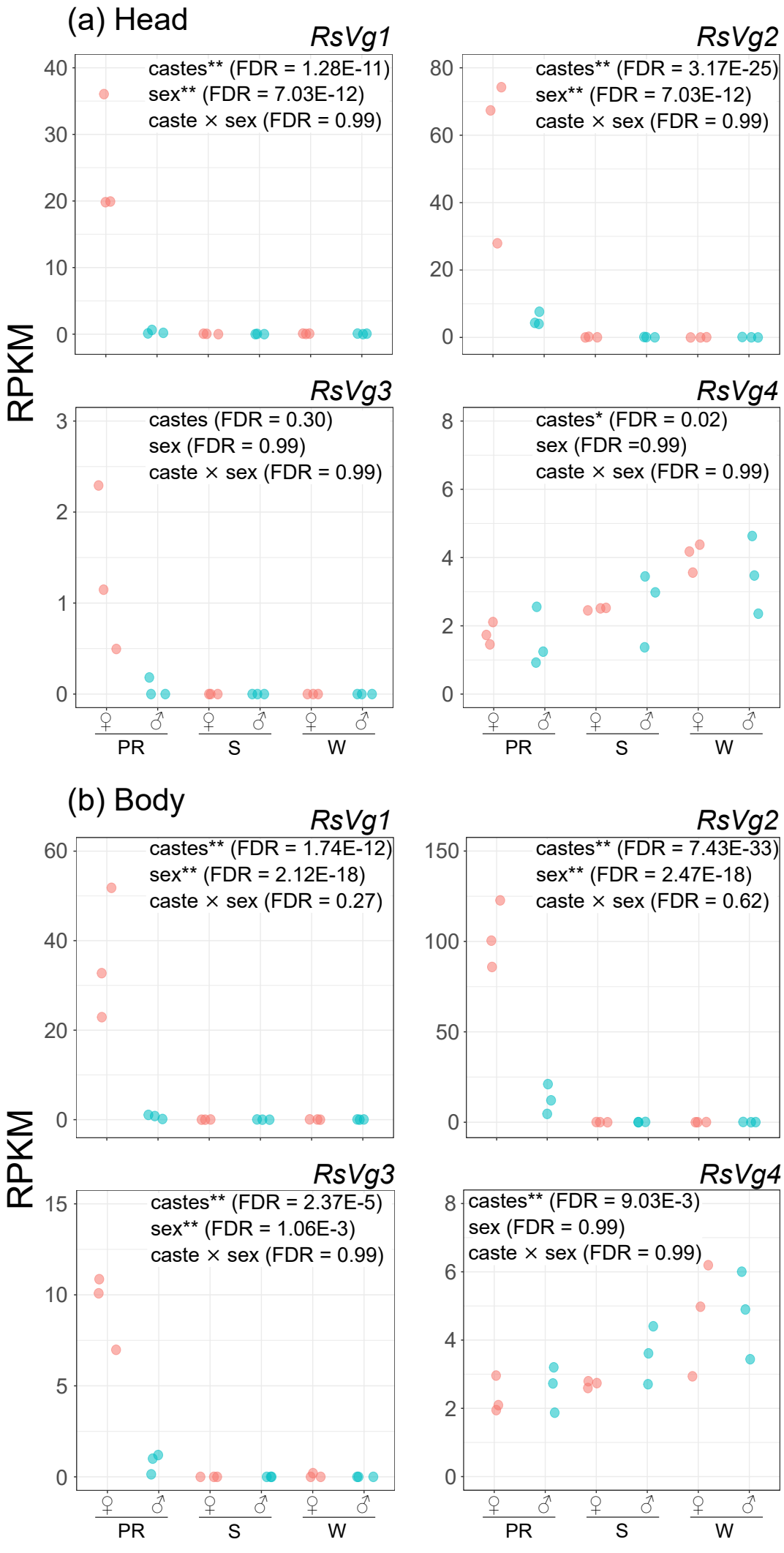

(a) Head
RsVg1
castes** (FDR = 1.28E-11)
sex** (FDR = 7.03E-12)
caste × sex (FDR = 0.99)
RsVg2
castes** (FDR = 3.17E-25)
sex** (FDR = 7.03E-12)
caste × sex (FDR = 0.99)
RsVg3
castes (FDR = 0.30)
sex (FDR = 0.99)
caste × sex (FDR = 0.99)
RsVg4
castes* (FDR = 0.02)
sex (FDR =0.99)
caste × sex (FDR = 0.99)
RPKM
PR
S
W
(b) Body
RsVg1
castes** (FDR = 1.74E-12)
sex** (FDR = 2.12E-18)
caste × sex (FDR = 0.27)
RsVg2
castes** (FDR = 7.43E-33)
sex** (FDR = 2.47E-18)
caste × sex (FDR = 0.62)
RsVg3
castes** (FDR = 2.37E-5)
sex** (FDR = 1.06E-3)
caste × sex (FDR = 0.99)
RsVg4
castes** (FDR = 9.03E-3)
sex (FDR = 0.99)
caste × sex (FDR = 0.99)
RPKM
PR
S
W

Figure 2

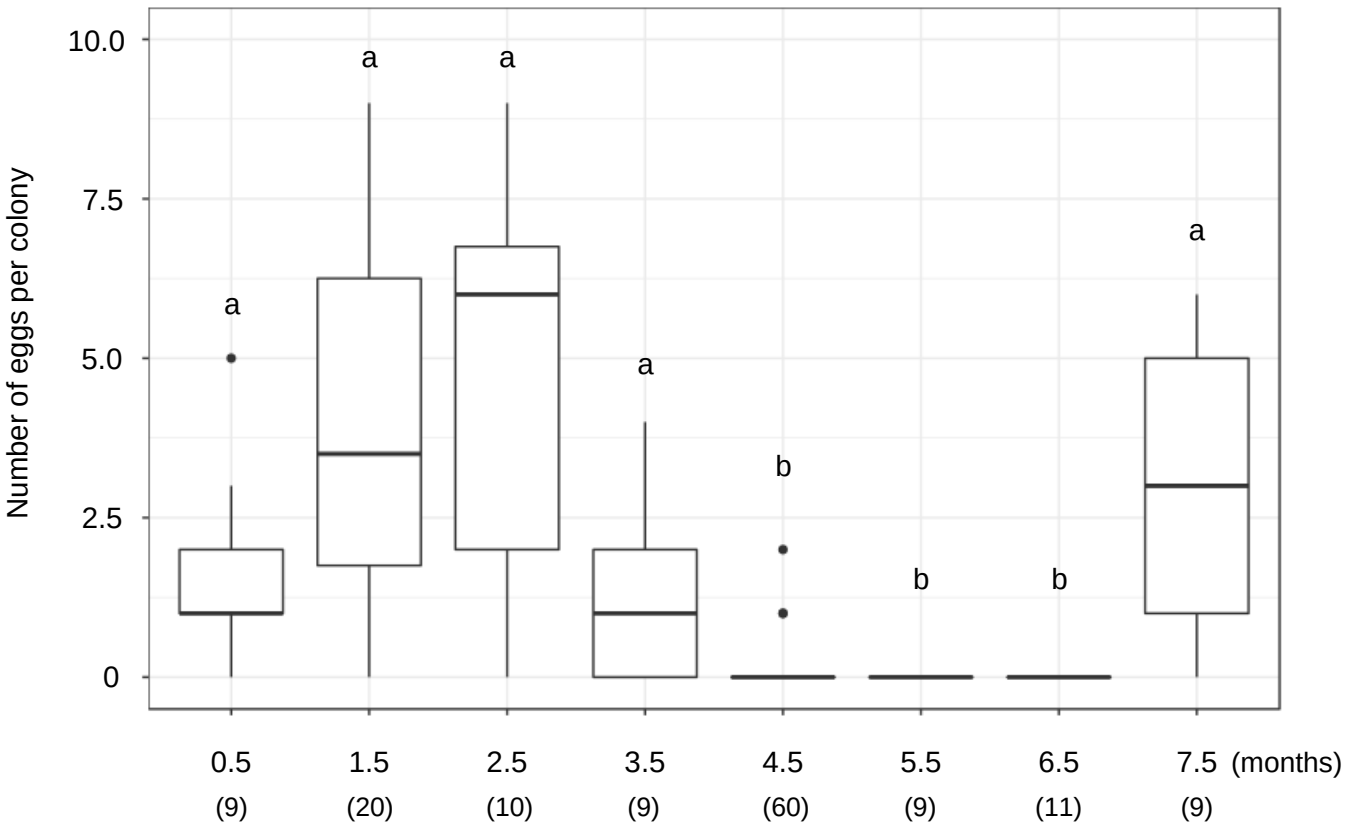

Number of eggs per colony
10.0
7.5
5.0
2.5
0
a
a
a
a
b
b
b
a
0.5
(9)
1.5
(20)
2.5
(10)
3.5
(9)
4.5
(60)
5.5
(9)
6.5
(11)
7.5
(9)
(months)

Figure 3

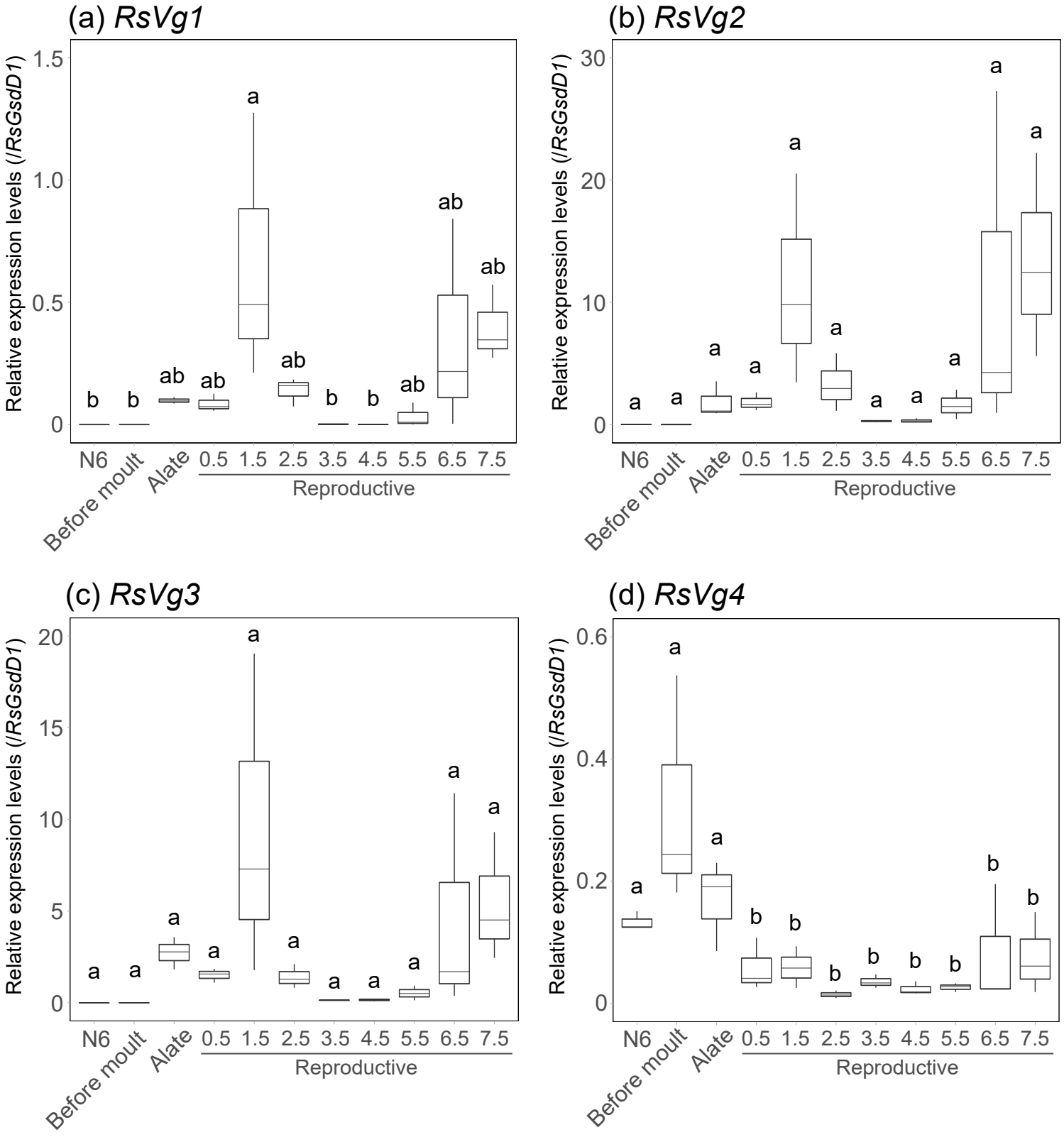

(a) RsVg1
(b) RsVg2
(c) RsVg3
(d) RsVg4
Relative expression levels (/RsGsdD1)
N6
Before moult
Alate
0.5
1.5
2.5
3.5
4.5
5.5
6.5
7.5
Reproductive

Figure 4

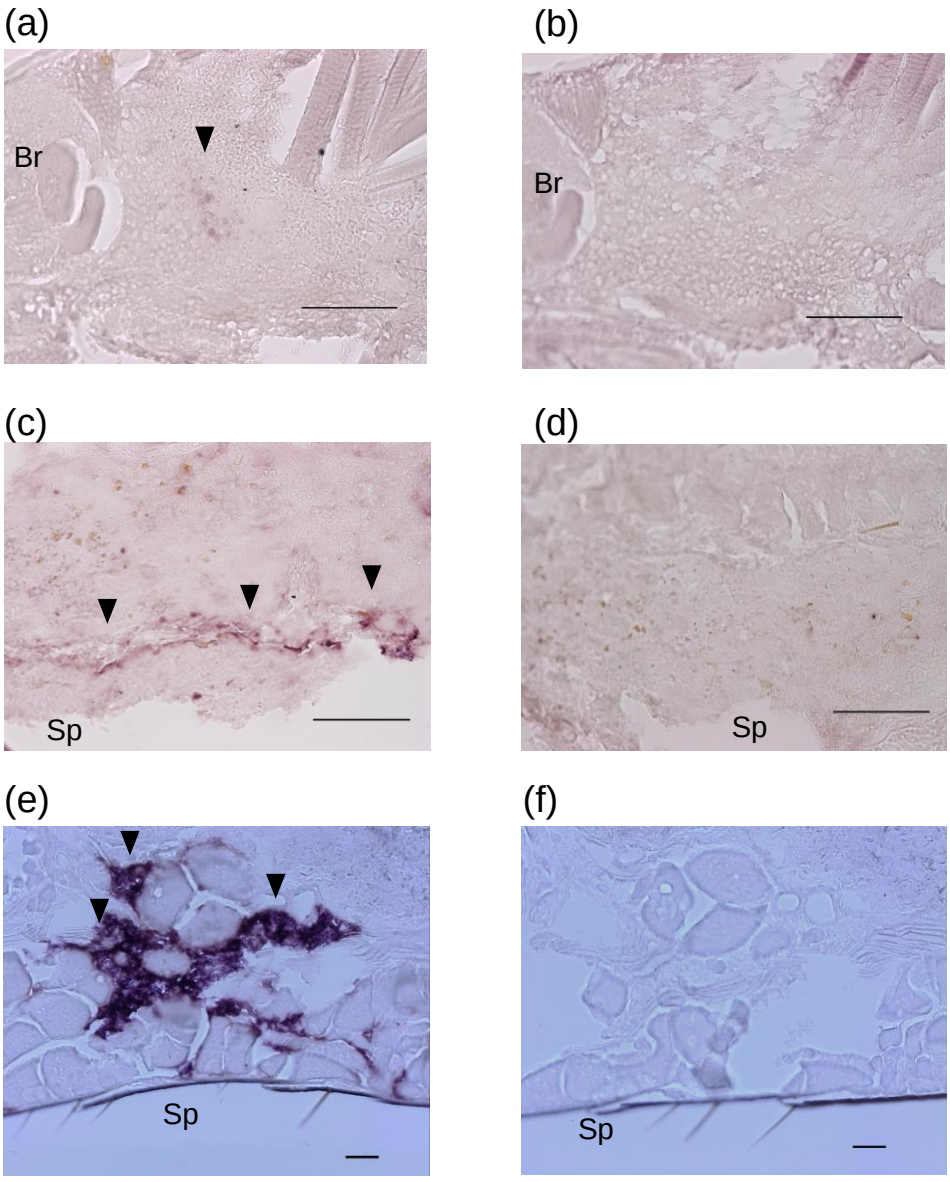
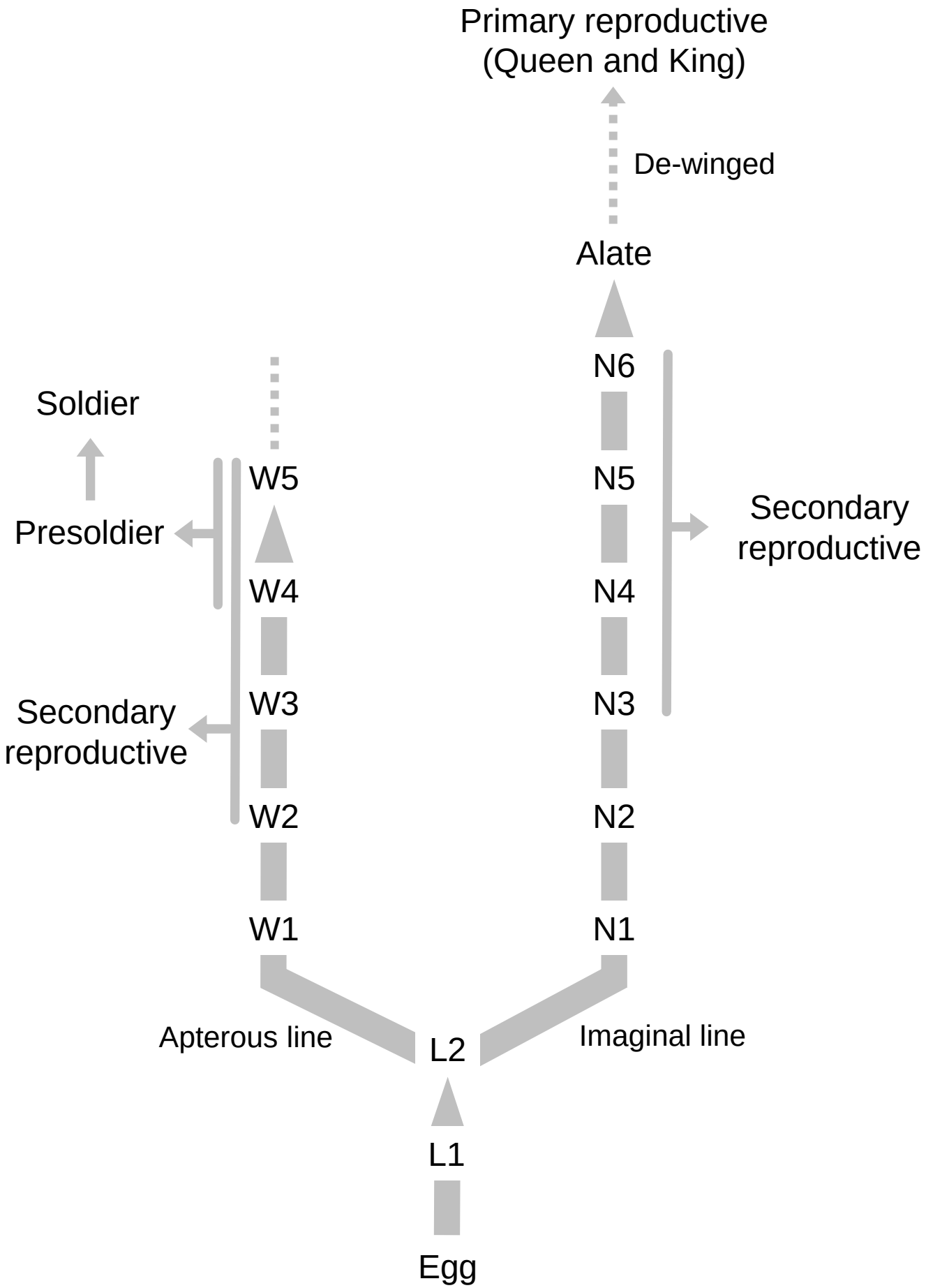
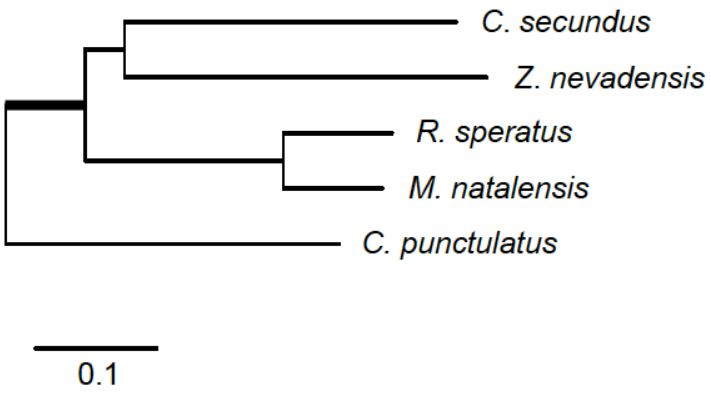


Figure 5

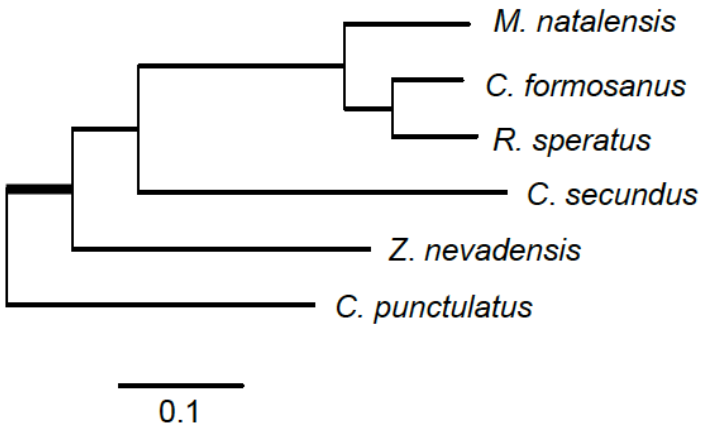


Supplementary Figure 1

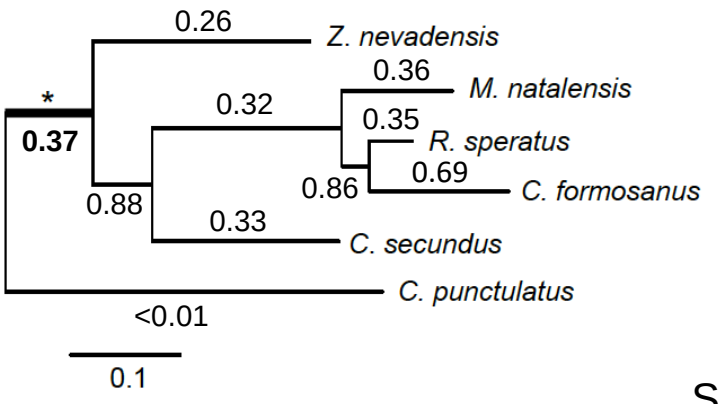
(a) *Vg1*



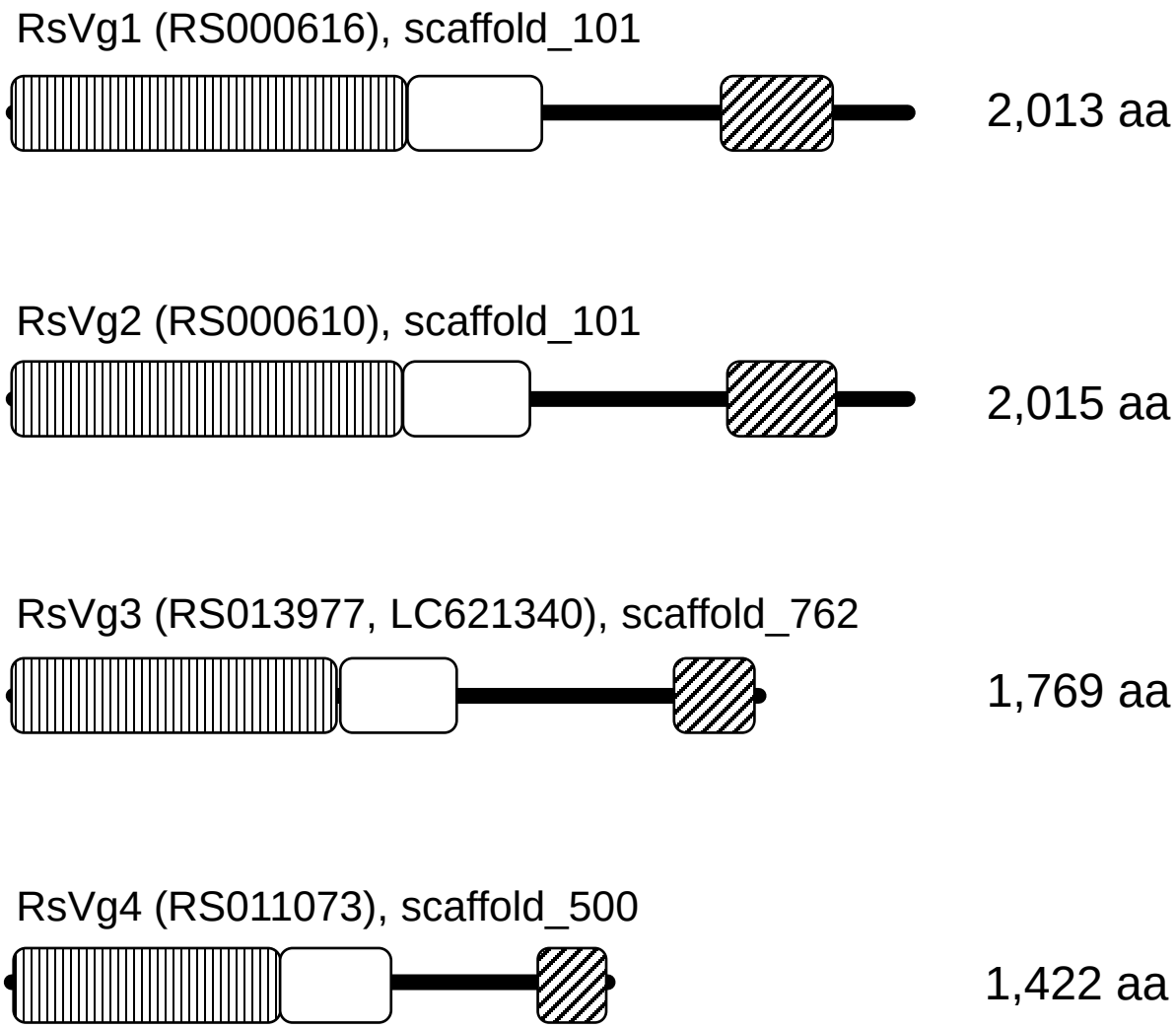
(b) *Vg2*



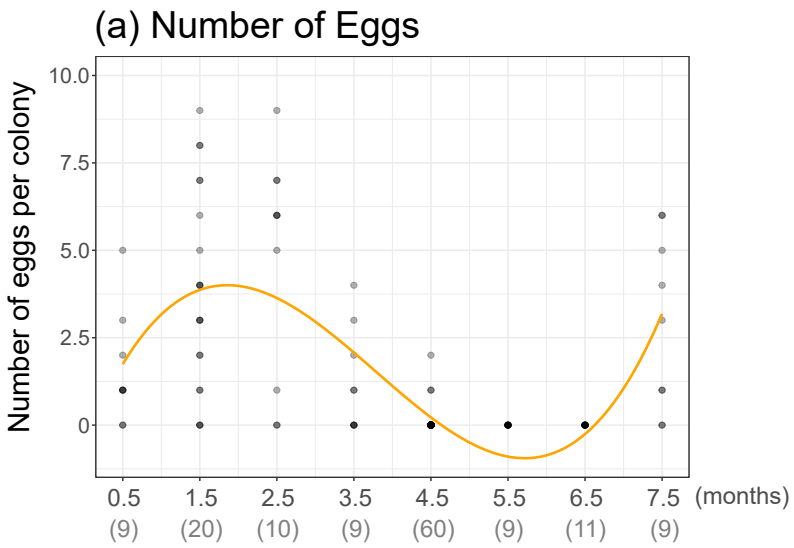
(c) *Vg4*

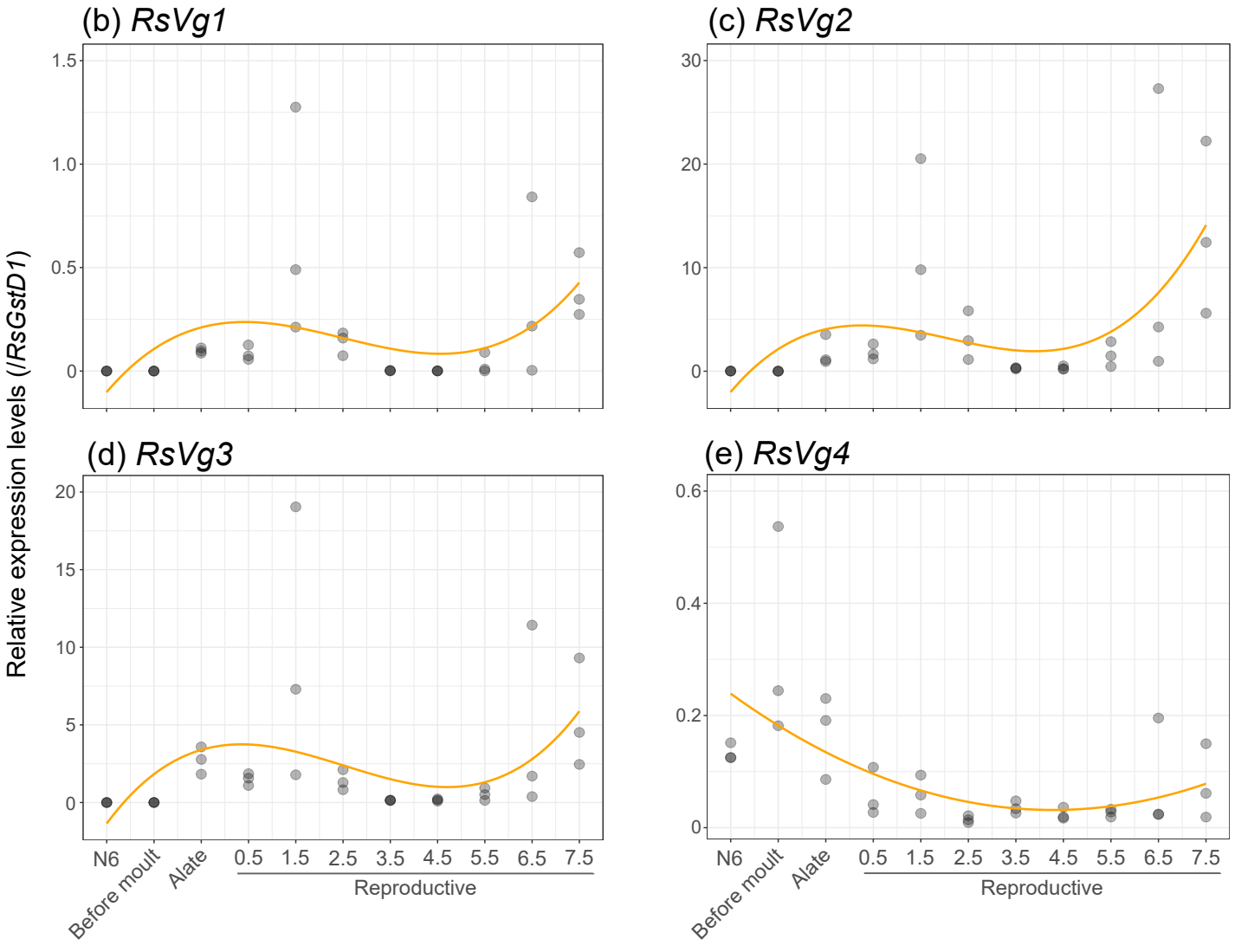


Supplementary Figure 2



Supplementary Figure 3





Supplementary Figure 4

>RsVg1_RS000616

AWEPKHRYVYKVQGRALTALHQLSNQYTGILIKADLALEPTRQELTGKLSDAEYAQVNMAPDLQINEPFQVQLKNGLVKNVIVSASTT
EWQRNMLKAILSPLQVNTQGARRAVFRAMESTITGECEVVYDISPSSSAFDVIKMQNFSNCEQRLHIGVSGLEDEPGDNKMGDFMSRSS
LNRAIINGSNFTIQSSVTTNKVLLSPEAYNNQQGMIVSVVNITIDSIERPKHPPENPQDAGDLVYVYPHEGNALNSSSQYADTTMKLANRI
ASDLQNASSIEQNTLSRMKRLVRLMDDKELDEVTEKLANIATWKVFRDAVAEAGTLPALAAITHWIRSDKLADAEAAEIVSALPRSAVTP
TPDYIRRFFELVEDPKKKQNSLNSTAILSFATLLRKVAVDSRFPPMTKPYFNVSLEKAIYEANSEKIQIYIRALGNIAHPKILQILEPYLEGEK
NVSRLLMVSTLDKVAHLYPKIARPVLYNLYQNTAESHEVRCVAVTLLMKTNPPLVMLQRMAEFTNIDPHQQVISAVQSAIKSAAALEDP
ANWELTEKAKAVVNLLNPTQDISFSKSLLNSLAMDDMELGYENYFSFVGSKDSLAPSSVLDHFIRNVGGRRDEHFSISGMVSSAKEVLNY
LKEKFENWNAEELENLLHIVRDQAKQVEANMQFSWLGIEKFLAMDNYTLDKIPLLVKAVAYLYSQPRYTKLYSPASVAFPNALGLPSLYT
LNVPTLLKSGGEVTADIDLVFSARKMSKMGFVAPFETATYIAALERDILINLPIRVAVNFDIENKHVDAALKPRPRKSYSTGMAFNLQNLV
RSSMSYENVTVSYNPDISTATTVKLTAYDDNKEPEFLAKAAVIDVSVKFEGQTVYSATAAYASSLVNASTTTLFLNAKPAKTLQAFLSVNV
SVPIMNFRQALKFNPNSSVSAQLDVIDAANNQAKALLKGRMEQSYAFKQFVASPMAKLCESQMREGNNLLSACRNVTEQANVLDLHNF
ALKYNNIPDSWKAYSVLRHYVTENKPNQLQVKVRLNNNITAVNVSILAAALLHNAEDSAADRIGNAMSPTCVVDGNAVDTFDNNTYPI
KLGSCWHVMMMEVPQTSNVAVLVRSNSGPQKEVQVILGENVITLSP--
ESVEVAVNDHPIRFSPNYELPSGAKLNFNSLAIAYDGERIMLEASNDYRDEVRGLCGTFNGDALSDFKKL

>RsVg2_RS000610

LLFLFAGTSA-----

SQGNLVYKVRSRTLTGLNQVANQYAGILINADLVVQPHENRKRAQLVNPVYAQVHSVDELPLNDPFEVKMDNGMIKEMIVSGKTTEV
EANMLKGIMSQLQLNTQAVNRAQYSVMESVVTGECEVLYDISPSQPIIQILKTQNFSNCHQRYNFGLEGFENQPGAN-
AGSFMSHSSMSRVVISGSNFIIQSSVTTNKVVLSPELYNNQKGMSVSTVNITIQS-
TRQQNPVSHPRSVGDLVYHYPIQGNDFGSNDDAVRVVTKLANKIGSEMENPNSLDENTLSKMTNIVQTLNRTAIDEVTRALDNIHAWE
VFRDSVAQAGSLPALSAISDWIHRDIISGNEAAEILSALARSVLTPTPEYIEVLFELVKAEAKQQPFLNSTAVLSFSEILREVAVNNNFRPLTE

PYLGKELQKAINEGSSPKIQVYIRALGKIGHPNILHYFEPYLEGQRTVSRLLMVTALKHLAKLHPEVATPILLSLYENLGEDRDIRTAAVFM
IMETNPSVNMLQRMAEFTNVDTSKQVIAAVVSAITSAANLEGPYTEHMATKAKAAVDLLTPEIEPSYSASIIRSYMVKNLDLGYNDYYSQI
ASNDHILPKALLHKLVTNIGGETTEPINIESIISSVRDTLDLVKEQFENFSNQDLSKLLGIVNEYRKQVEGSMAFNFAGGERFLPFDNHTL
QNIPSFVKIMAQLHKNPLYTKLYDRNTWVFPNAMGLPFVFTLKTPTLMRFSAKSSTSIQLVHASKIESVIGFFTPFDNMQYVAGLDKDF
MYNIPLNAKIDVDILNKNLSLTLMPRPTKQVDVGMVVRVDKAASQSIQYRNMSVTFEPQLSTAKVVKITAYDNNESNEFLSKAAVIDLSV
KFEGRTEFVMTAAWASSLVNGSQSVLFLKSIPAQPLQVCLVGKVSVPVTNFHEALKYDPNSSVRAQLSFGETCQHGATVSLKGEMEQSH
MRKESIASPMAKLCKSQMNMKNHLLPACRNVTQAANNLDHYHFKIMFNNIPDVLKVYSLARRYVTENVKDAVHIIVDINEESTALNVT
MCAQALFMNPAHTVVKRIGSTMSPTCVVDSTMANTFDNITYPMDLGKCWHVMMMVVPTLEEVSVLVK-
EVGNNKEVMVMTNSAHIHLKPALHSVLVTINDKPVPHKFYHELPEGAYLTFNSLVIIYDGVRVMLQASNEYRQQIRGLCGNMDGEPFN
DFAKI
>RsVg3_RACE
VLLLFFSTSFENGWRPRVEHTFALDGRTVAGIPSGQQQHTGIHYYGTLRVQRINGVLLMQFQNMTYTKFHENRTLPLSSTVAVKMGH
G-
IQTIKMAAGLNDWEMNFIQGVVNQLQIEFTGLLGTSYKMMENSVFGRCEVQYEIRRQMNFVEIFKTRNFTNCEYSFHFGLPVSKRHPG
GNACDDFWNRASVTRISGCGDDFRILQSETNSRVIVNLHLHKGTEASVSSYITMNLVSAKQTSRLPDNIVEQADLKFKFPLSRFK-----
QMTQIFKETMTDVVKDLEMYTAIEEMTLEKAVRICRAMTLFDLKRAVDSVNRQSESKLYRDVLVMCGTNPAFQILKTWIEERKLEGED
AIAVLATMPAHLQTPGPDIVREFYELIQSSARSDQLLKTTSIIAFSHLLRLACVDT---HQVSSWLATMQR-----
KDHPLQRVYIAALGNTGAVSALPLLQPYL-------
RATAVISMKYLALSVPEETAQILLSLYHDVGQPVAVRVAAVSMLLYTKPQPVLLQRIAVSTWYDPSLAVAAFVRSSMMSLANLEDPFFSEL
SRHVRMALHLVRPPLGMHTSHNVLWSRVSGELQIV----
MNQLSYEGSLDANNLYYRYTQRLGGLSRVPFQGEVTIANPDHFVTMLRELIHDRSTRWLREKLKIAQRNLPKFEGDIRIQFGFVEGIFPF
DNQLQVELAKAGRKLVEQMKTGKYLKV-
DKTKMQVVTELGIPINFELTVPWAVRLLGTAEMDVDMLYKSHLLSEISFFTAWDETVHMAASHGVTAVRLPTKVRASV-

PQSSAVQLAFEIRNTQEWNDRLTLELNRPEQQGAANDNLNMTSEQQNSTQEHIKNEKYDTSID-
YLLLRAHVMGLDVRVIGMQTFSTFWTYAAGRSGRLRRFVFLKED----
WQAGFTYSLITPILKLENALLANVTPSLQADLVYGKESRTEYEVSLQAVMEHSALK----
EQSLRAKCLDLTEGQTNINRQCVRAAMNAMFMDRYSATARYNEKCKSLKYFNYLRYYVKEDPTNWIKGVLEIAPSLDTVNAWIMIIEAIF
NPTVGYLRRLKGE--SHCGHGGSNFFTFDGAQFDYNLSSCWHLLAKDCSGQSRFAILTRSVQNNETEVEVNLDNYQVHLRPG---
PEVSVNKQPLRVTDDAEWPERLSINMAGLHLTYTGSNTLITADSRMRGRLCGICGDYDGDTVKEF---
>RsVg4_RS011073
VLSLVVNTAALELWLPGHEVVYQYNTSMVASTVFPEHMVSRWNMSGKLVVQGGEDLATVQLQDLKVLTYTG-
HDLPLLQPFIIHYYEGQVVGLSVKAGEEEWSINMKRGLVSLLQLDLLRLQS-
AFISTESSQYGHCSVQYSVQRSNGTVAVQKFVDFPSCVDFPVQKLSNVQKSCSSGPTATMSSGSMRAYVLEPSGHLLIHSVQADGVIHLQ
H----
PNSYFLQSWVGLYLVWQREVEVAMKGEGERVQLVHQVPSQGRSPTSQQDLLTMVTEILGELADSLENLDRLHEEHVVHALYLLSLLEQ
ASLHNLFLELRHETLRNLFLELLPFVGSRDSMLLVRNLVMRGAVKNDTAIRLLSSLPFNVHQPSEQLLTDLEVLLLIPE--
DTLEHSAALLSFATLVHKVCSSRCMPETVDRYVSQYLDLLTGSKDYAHQMLYLTALTNFELQQVVQLLTPLIRGDDGRYRVLAMWAT
MACSTLQPDQVYHLYWPLLTNCSEDLEVRVVALTMLVLGQPSVTHLYNLMLFMQHERNAHLRHFWYTTLHSLAHSQHPCHMHMRE
VAKKLVQFVVPEVGHWTGGNYVLDYADPDNGFSSQANLIMVSSPRTGLPTVFNLDIGSHTMGYTYNQFSVFIRLEGLGDVIK--
QQLMQVTQLVTMLQQLKIPTKPSEPVHVELLVKMQKVVLVHHANESTLNTLFKVLNSVQLLANGGI---
HLNAQKIMRVTDLGTNALLQMSAGGLVSVRASFSSGLELLSSLSYQMSLSQYNPVLNMWHGVQQSLNMHVNTAFNLQASFSATQEHV
RFTVQHTPAKEM--GLVTHVKGAAATKLCPVHVSVTRNVSEMS--GLRLALLEHSLSEKYLML---IKVEAIFPAGE-
YRAQLCMVQSLVNSSSSS-SITMELTTG-
KMCLNLTTSFLESDSVDSLINPRPHTAVLSVAWGPPTDGETSLQVALEGRVTTEQWSLVSAWPYDACRQDIQQGRHLIPACHEVAHKL
ASQRYYNISITSINLP----LQNFLPSKL-ASLSRNVSVHLLVPPEGRPV-
LTVMLGSAVTQLQFPPLLRLVSRLGGSCALTEHSVRTLDGVTLTTNVPTCYTVAMADCSSQPRFAVFVKMVESSPLAVKVYTNDNYVE

VLPTADQLSISVNGQIMRHEEGIHVGEGD------IWLEYYGHFITVSVLGLYRGGLCGLCGDHDGDPSNDAGKM
>Znev_07681

VESLNVSGSTNEWETNMLKAILSPLQVNTQGARSAVFTAMESTVTGECAVVYDISPDGDIFDVIKSQNFSNCDERLHFGVTGLDKEPGQ
NSMGEFMSHSSLTRVIINGTRFTIQSSVTTNKVLVNTELYNSLQGMIVSSVNLTIDSIEPSEQAPVNPRNAGDLVYTYPKEGSAINDSSDPV
REVMNLAGQIGSDLQNASTIDENTLARMTRLIRTMNTKELYQATKKLVSISTWKAFRDAVTEAGTLPALATITDWIKSNKVTGMDAAEI
VSALPRVAITPTPDYLKALFELSKDPKKREMFLNATAILSFANLLRKVAVDNKFPPLTKPYFGSLLNESIAEDDNQKTQIYIHALGNIGHPS
ILPIFEPYLEGEANVSRLHMVSALNKMVKLYPEEARAVLFKLYQNTGEPHEIRCVAVSLLMKTNPPANVLQRMAEFTNSDPDTQVISVVQ
SAIKSAASLKGPDTFELSSNAKTAVNLLNPEQDNAFSKTFLASKTVQEMNLGYEDDFSYVGSSDSLLPNVLFANLVRNVGGSTEEETKVS
WMVSSVKEALKYLREKIEKWMIGDLEKFLDIETDEPEKVEGNIEFRFAGAQRFFAFDNHTLDIIPKLVKEFARHYGRPQYTKLYSGHQVT
FPNALGLPSIYTLEKPILMKFGGEVTADIHIVSSARKMSKIGFIAPFESARYIASLERSSLLNIPLRVSFDYDIENNHVETVVKP----------
INNLFQFQCNYRRGICRFQVDTSDSGTSRLSIYDNNEES-----
ASVLDVAISFEGDTVYAATLACASSLVNGSTSTLFVSAKPAKSVQVIQKTESKSPRHFFLILIEI---SDIS----------
LQGSMEQSLEFREALASPMAKECEREMSEGNSVLSFCRNVTEAASQLDQHNLKLKYSNIPTSWRAYSVIRHYVSENKENQVEINLKLNSN
NSALNVSILAASLLINPKTSVLDRIGKRSSPTCVVDGNVANTFDNNTYPIRLGRCWHVMMMEVP--
GNVAVLVRGPSGQEKEVMVILGQTTVKMSP--
DSVKVSVNDKPLKLSPSYELHGGAKFVFDDLELAYDGERVLLQASNSYRDEIRGLCGTFNGDSLTDFKKL
>Znev_07682
AILLLVWTAALPAWENHKRYLYNVNGRTLTALHQVGNQYTGILVRAQLALEPSQNLLRAQLVSPEYVKVHSIDALPLNNPFEIKLNNGL
VKEIFVSGQTTEVEANMLKGILSELQLNTHAVDRAQFHVIESVVTGECEVLYDISPDEPLIQIIKSQNFSNCNQRFHFGFDGFANQPGGN-
IGDFLTRSSMSRVVISGSKFIIQSSVTTNKVLLSPEVYNNQKGMSVSRVNVTLESITKAQKPVESARSVGDLVYRYPIQGNKIGISKDVVKTV
VKLATKIGSDLENPGSLDENTLARMARIIRTMDRNQIAEVTNALDNITAWESFRDSVAQAGTIPALSAITDFIRRGLIRGEEAAEVVAVLT
LAVRTPTPDYIDLLFELVTDQRKQQMYLNCTAVLTFSEVLHEVAVNNEFPPLIEPYLGNEFRKAIEEGNSQKIQLYTRALGNLGHPMSLL

YFEPCLEGRINVSRLLMVSSLKKLANAYPDIASSVLLKLFENIGENNEIRIAALSLLMETKPSSFILQRIAEYTDMDTSKDVVSAVQSAIKSAA
NLQETFTLPMAKNSEAATNLLTPEMDLSYSTNLMTSFKMENVKIGYEDVFSYIGSKDSVFPKALFFKLFTSIGGDTTEQLNIVSMLSSVRD
TLNLMKDQFGNWTDEELDNLLGILKGNMEQIEGSILFNLAGGRRFLPFDDSTVENVPKFVKRIAQMYKHPWYTKIHDRYTLNFPNAM
GLPFVFTLEVPTIVKIAAETTTNLQFTYHFKHESKIGFLTPFDNTQYIAGLDNAVLYNIPLEVKIDVDLRKNNFSLTLKPGPTRQFETGLIF
RLNMFARRAIFYNNVALVYDTELSTADVVKISAYGNN---
EFLSRSAVIDLGVRFEGETEFVLTVAAASSVVNGNKSLVFLKTQPAQPFQICLASNISVPKINFHEALKFDPNSSVRAMLSIGKTCKDGGTI
LLKGEMEQTERRKKSIASPMAKLCESQMTKKNYLLPACRNVTELANNLDQYHFEITSEGVLETFKLYSTVRHYVTEDQEDSIDIVVNVNE
QSTALNVTMLTKTLLINPQSTTIDRLGKHALPTCVLDSTMANTFDNRTYPIELGDCWYVMMKLVPRLDQVTVLVK-
EVGRNKELRVLLSSASFELRPVSRSVELAMNGNPVRFTSSYGLPGGSHLKFNSLVIIYDGVRVLLQVSNEYRDQVKGLCGTFDGEPVTDF
KKM
>Znev_08605
VLLLILSKSYDHGWKPRKEHMFALDGRTVAGIPSGQQQHSGIRYRGILRVQTTSG-LFLKLQNMTYTKLHENDNLPISSTIAVRLSQG-
VETVSFKTGLDDWEMNLIQGIVNQLQIDLKFLKGASYKMMENSVFGRCEVLYEVTRNMNFVEIFRTRNLTNCEYSFHFGLPKSKEHPGG
NMCQDFWKRVSITRLSGCGDDFRILQTETNSRVLVNLHLHGESEASVSSYITMNLISMKETSRPPADPIVSNELKFKFPLQGSE-----
QQIQKLKETITDVATDFEESTLIEKMTLEKAVRICRTMTLFDLKRAVSSVNKQTESKVYRDVMVMCGTNPAFHILRTWIEQKDLRGEEA
NQVMVTMPAHLQTPGPDIVRQFYDMIQSMAREDPQLKTTSILAFSHLLRIACVDS---HHASSWFTSMLR-----
SDASLRRVYVAALGNTGAVAALGPLKPYL-------
RATAILAMKYQALSFPEQTSKILFSLYHDVGQPVAVRVAAVSLLFYTRPQLVLLQRIAVLTWYDPSLAVAAFVRSSLVSLANLKDPLLSVLS
RYARMVLSLVNPPLGLHKSHNVMWSTVSGSLQIV----
FNQLLHETSLDATNFYYRYTHRLEGVSRVAFQGDLFVSNPDQFLTLLRELIHERHTSWLRKNLNIAQRTLPKFEGDFRIQFGLMERIFPF
DKKLLVELSKAGRNLVEQMKTGKYQKI-
DKTTIRVVTELGFPINFDLSVPWAFRLVGNAEMHIDMTYTSHLLSELCFYTTWDETVHMAASHGVSVFRLPLTLTAVI-
YKNDIMQMAAKIRTTQKLRDGLNFDVN--GQQNVSFEDSTPTFSQNNATTDNIETIAYDTATT-

YLLAQAYVVDINMRVNGTETFTTFWTYAAGYNGGVHRFFFLTEA----RQIRFTSTLITPVMNLKDIIRSNTTASIHADMVYNDDGR-
ECQLSLKTVMGHSTMKAESLEEHLTDKCIDNYEGKTDIDPRCLQAVMEAMFMDRYDVTVTYNDESKSFQYYDYLRYYATDDPANTIT
AVLEISPRLDNVNAWIMMTNILYNPAVSYARRIRGE--
SYCGHGGSKFLTFDGVEFDYNLTSCWHLLSKDCSGHSRIAILTRPDENNETEVEVNLDNYQIHLSPG---
LNVSVNGQGLQLAENAETPERLSVNMAGFRIIYTGSSTLIMV----------
>Znev_06771
LLPLVVSTASLELWLPGREVVYRYETSLEAGTIFPVATVSQWNMTGKLVVQGAGHMAVVQLQDLKVSMYNG-
HDLPLLQQFIIHYDAGQIIGLSVKAGEDEWAINMKRGLASLLQLDLSHLQS-
AFISAENSHYGHCDVEYSVQRSNEMLAVRKFVNFSSCMDFPVKKFGYVQK-
CPSGAQETVSSGSVRAYVLEPSGHLVIHSIQADGVTHLQHFGVQSDSYVLWSWVGLYLQQQGDMVVTMKGGGQRVGLTHQVPSQGR
SHPSQEDLLITVAKILGDLADSLEDRGMLHEEHVVHALHLLSLLNQASLQKLYSELQHETLRNLFLEMLPLVGSRDSILLVRDLMVRRTF
QNDTAVQLLSSLPFSVREPSEQLLTDLEVLLLLPE--
GTLEHQAAVLSFATLVNKVCSVRCAQETLDKYVGRYLDLFTGSSDHASQMLYLVALGNFEIRRVLQFLAPIIRGDDTEYRILAAWATMG
QAALQPDQVYHLYWPLLTNRSEYLEVRVAAFTMLVLSRPTLARFHNLMLILRQEQNPHLRHFWYTTLHSLAHTHCPCYSHLREVARNV
ARFIAPKVRRWATGNYILDYRDPDHGFGSLAHLIIVANPRTGLPASFFLDFSSHAMDFSYNQFSIYVRLDGLGDAMK--
RQLMHVTQLIHLLQQLEVPVNPSEPVHVELLVRLHKAVLVHHANQSTLSTLFRVLNTVQLLASGGV---
HLNAQNITRATDLGTNAILKVSSGWLVSVRANFSSGLHVRSLLNSRTALMQYNPVMNVWHGVEQSVSIQVNVPFSLKASFNTTQQSIRF
TVRQTPGEEL--GLVTHVKGVAATELCPLYVSVMRNAPELS--GLRLALLDHSYKDDYMIL---IEVAMSIPAGD-
SRLQVHVTRTPASSTSPY-KLTTEVTIN-KMCVDVT-AFIKSDFVVSIFCPRSHTVMLTSFSGHPTDG-SSLQVTVHGQAAAAEWSA--
AWPYSACRRDLQLNSHCLPACYEIVHEIATLRYYNISISSSNLP----LLDLGPSRL-VSLSSNITAHLAVPQESSPV-
LTVVLGSSVTRLQFSPVQRLLTHVRGSCVLTLHNVWTLDGANLTTDVPACYALAIADCSLQPQFALFIRKVEGTPLAVKMYAAGNQIEV
LPTADQVAISVNNRTVTSQEGIHLQEAH------VWLEYYGHLATVSVLGLYHGDLCGLCGDYNGDEGDDIAQV
>BLKM01000597_1

----------WRPRLEHTFSLDGRTVAGVPSGQGQHTGIHYYGSLRVQRINGVLLMKFQNMTYTKFHENHTLPLTSTVAVKLGHG-
IEMMKMKAGLHDWEMNLIQGMVSQLQVDYRALKGTTYKVMENSVFGRCEVTYEVTRHMNFFEIFKTRNFTNCEYSFHFGLPASKKHP
GGNACDDFWKRASVTRISGCGDDFRILQSETNSRVITNLHLHKGTEASVSSYVTMNLASTKQTSRLPDETVELGDLKFKFPLTEYE-----
QKSQKFKETIADIAKDFEMSKVIDEMSLEKAVRICRTMTLFDLKRAVENVKRQSESKLYRDILVMCGTNPAFQILKTWIEERKLEGEDASG
VLATLPAHLQTPGPDIVRQFYELIQTSARDDQQLKTTSILAFSHLVRAACVDK---HQVSAWLSSMLR-----
SEQSLQRVCIAALGNTGALSALNPLQPYL-------
RASAVISMKYQALRVPEQTSPILLSLYHDVGQPVAVRVAAASMLFYTKPELVLLQRIAVSTWYDPSMAVATFVRSSLTSFADLEDPIFSELS
RYARMVLPLVRPPLEMHTSHNVLWSRVSGDLQIV----MNQLSYEGSVDATNLYYRYTQRLGGLSRVPFQ----------
---

--VS----------
>BLKM01000504
LLSVVVNVAALELWLPGHEVVYQYNTSLVASTVFPEHALSQWNMSGKLVVQGGEDLAMGQLQDLKVTTYTG-
HDLPLLQPFIIHYYEGQVVGLSVKAGEEEWSINMKRGLVSLLQLDLSHLQS-
TFISNESSQYGHCSVQYSVQSSNGTVAVQKFVDIPSCVDFPVQKFSNVQKSCSSGPMATLSSGSMRAYVLEPSGHLLIHSIQADGVIHLQH

PNSYFLQSWVGLYMMWHHEVEVPMKGEDERVELVHQMPSQGHSPPSQQDLLTTATEILGELADSLENLDKLHEEHIVHALHFLSLLE
QASLQKLYLELRHETLRNLFLELLPLVGSRDSMLLVRDLVMRGAVKNDTAIRLLSSLPFNVHQPSEELLMDLEVVLLIPE--
DRLEHPAALLSFATLVHNVCSSRCMPETVDKYV--------
GSKDYAHQMLYLSALTNFELQQVVQLLMPLIRGDDSRYRVLTLWATMGCSTLQPDQVYHLYWPLLTNCSEDLEVRVISLTMLMLGQP
SVTHLYNLMLFMQHERSTHLRHFWYTTLHSLAHSQHPCYTHTREVAKQLVQFAVPEVGYWTSGNYVLDYADPDNGFSSQANLIMVSS
PRTGLPTVFNLDIGSHTMGYIYNQFSVFIRLEGLGEVIR--

QQLMQVTQLVTLLQQLKVPTKLSEPVHVELLVKVQKVMLVHYANESTLNALFKVLNSVQLLANGGI---
HLNAQKIMQVTDLGTNALLQMSAGGLVSVRANFSSGLELRSSLSYQMSLSQYDPVLNMWHGVQQSLDMHVNTAFNLQTSFSATHER
VRLTIQHTPAKEM--GLVRHVKGAAATKLCPVQASVTRNAP----------LL------IYKMA-------TLL----------SES----------
----------VFTEHESMIT----------GMRY----LRKCSGG----IRN---SHI-RGLNQ----------EPV----
MVGRRE--------LRLFGHL----------ISM-DSNRKP----------GQI----------EGQ----------
GQKRKSICGS----------
>BLKM01001253
-MFL----------
QLVNPVYAQVHSVDELPLNDPFEMKMDNGMIKEMFVSGKTTEVEANMLKGIVSQLQLNTQAVNRAQYSVMERVITGECEVLYDISPS
QPIIQILKSQNFSNCNQRYNFGLEGFDHQPGAN-LGSFMSHSSMSRVVISGSNFIIQSSVTTNKVILSPELYNNQK----------
----
GNNFGSNDDAVDVITKLANKIGSELENPNSLDENTLSKMTNIVQTLNRTAINEVTRALDNIHAWEVFRDSIAQAGTLPALSAISDWIHRG
IISGNEAAEILSTLARSVLTPTPEYLEVLFELVKDPKQMQPFLNSTAVLAFSDILREVAVNNHFRPLTEPYLGKELQKAINEGNSPKIQVYIR
ALGKVAHPNILQYLEPYLEGQINVSRLLMVTALKHLANLHPEVATPVLLSMYEDLGEDRDIRSAAVFMIVETNPSAHILQRMAEFTNVD
TSKQVIAAVVSAITSAANLEGPYTQEMATNAKAAVDLLTPEIEPSYSASNIKSYMVKNLDLGYNDYYSQIASSDHVLPKVLLNKLVTNIGG
ETTEPINIESILSSVRDTLDLVKEQFENFSNDDLSKLLGIVHEYTEQVEGSLAFNFAGGQRFLPFDNHTLQNIPSLVKIMAQLHKNPFYTK
LYDRNTWVFPNAMGLPFIFTLKTPTLMKLSAKSSTSIQMVHTYKVESKIGFITPFDNMKYVAGLDEDFLYNIPLNAKIDVDILNKNLSLTL
MPRPIKEVDTGMVVRVDKAASQSIHYQNISATFEPQLSTAKVVKITAYDNNESKEFVSKAEVLDLSVKFEGHTEFVMTAAWASSLVNGSG
SLLFLKSQPVRPLQACLVGKVSVPITNFHEALKFDRNSSVTAQLTFGETCQYGATVSLKGEMEQSQMRRESIASPMAKVCESQMKKKNH
LLPACRHVTQAANNLDHFHFKVKFNNIPDVLKVYSLARRYATENEENAVDITVDINEQSSALNVTMWTQALLMNPAQTVVERIGNTMS
PTCVVDSTMANTFDNITYPMQLGKCWHVMMMVVPTLDEVSVLVK-E-----
EVLVLMKSAYIHLKPALHSVLVTINDNPVPLTHYYELPDGAHVTLNSLVVIYDGMRVMLQARNEYRHQIRGLCGNMDGEPFNDFAKI
>XP023706879_Csecundus

FLCLFVATAAVPAWEPNKQYVYTVEGRALTGLHQLANQYTGILVKAKLAIEPTTQELRGKLFAAQFAHVNMVADLPWNKPFELKLNN
GLIQNLTVSDSSNEWQTNMLKGILSPLQVNVQGMNRAVFRAMEATITGKCEVVYDISPNDNIFDVTKNLNFLNCKERLHFGFTGLDDE
PGQNKMGEFLSHSSFNRAIISGTKFTIQSSVTINKVLVSPEMYNNQQGIIASRVNLTIVSIEEPENAPENPRNAGDLVYRYTLEWSAVNDPS
NPVESVIKLAYQINSDMQNASSIEQNTLAKMTRLVRIMNREELSELTKKLFNNATWKSFRDAVAEAGTLPALLTISEWIRDNKIAGTEAVA
VIAALPRASLVPTPEYLDAFFELSTDPKKKQNSLNSIVLLSFSNLLRQVAVDHRFPQLTKPHLGRLLNEAINEGDSRQIQVSIGALGNVGH
PSILPIFEPYLEGTRNASRFLMVSALDNLASVHSKEAAAVLLKLYQNTGESHEIRCLAVSLLMKTDPSASMLQRMAEFTNTDPDPQVTSM
VQSAILSAAKLQDPPNWQLANNAKAAVNLLKPIPDLSFSQTFMKSRMAEDLDLAYEEFISYIGSSDSLFPSSLFASVTRNMGGSRNEHSSIA
TMWSSVKEALNYFEQKFEQWNTDELESLLKIIRDKAEQVEGNMQFRFAGAQKFFPFDKNWMDRIPKLVKQFSQRFSQSNNVKLYNPVR
LTFPNALGLPSIYTLEAPSLLQSTGEFTTDLNLVFSMKKMSKIGFLAPFENARYIAGLERNLLVNIPLRASVNLDFENYHVDATLELRPSK
WFNTGIAFTLQSLASRSIKFTDMSVSHDPTLSTAKTVRFTAYDDNTEAEFLAKAAVLDVSVQFEGETVFTATAAKASSLFDGSSSSLFISAK
PAQSFQTFLSANVSVPSMDLLEALEFNPNSSMTARLDMNNGTNTQWTVILKGNMEQSLEFRQAVASPLAQECQYQMNHGNNLLSACR
NAIETASRLDDHTFTLKYKNIPTPVSVYTVIRHYFSENPQNEVQINVKMNENSNALSVNMLESSLFLNPNTKALERVRKQFSPSCVVDGT
AAVTFNNRTYPLNLGKCWHVMMMDVPQSSNLAVLVRNNFGLKKDVQVILGNNMIELSP--
QSVQVSVNDRPLRLWPNFRMQNGGQLNLYYLKLAYDSERVLLQASNKYRDEVRGLCGAFNGDPFSDFQVL
>XP023706896_Csecundus
VLFLFAATAAVPAWETNMRYLYNIQSRALTGLHQVADQYAGILMEARLAVEPSEELLRAQLLQPMYAKVHSIDDMPISNPFEIRLDDGL
VKELLVSRQTTESEANMLKGIVSQFQLNTQAVHRALYSVMESAVTGECEVLYDVSSKGSLIQIIKSQNFSNCNRRYHFGFSDSENQPGGN
-
VGNLVARSSLSRVVISGSKFIIQSSVTTNRVFLSLEVSNDQKGMSASRVNVTIESITKTQKSLRGQRSVGDLVYRFPIQESKLRASKDPVRTV
VKIATKIGSELEKPCWLEDNTLSKMTRVLRTMDSKQIDEATRTLGNIGAWKSLRDSVAQTGTLPALSVIADLIRKDKILGNDAAEVVALL
SRSALTPTPAYLDMFFELVKDPKREQPFLNLTAVISFSTILREVAVNTHFTALTEPYLGHELHKAITEGNSPKIQLHILALGNIAHPNILQY
LEPYLEGRIGVSRLLMVCALMKLTKMHPEVASPILLKLYENLGEAYEIRTAAVFLLMETKPSETLLQRMAEFSNFDTSKHVIAAVHSAIKSA
PNLEEPFT--FRTNGKAAVNLLAP----------VSSSRKTTNL--------

NNEETATIPDIANENMEQVEGNFMFFHAGSDRVFPFDYHTLQNIPVLMRDMARKYRNPEYTKIYYRDNIMFPNAMGLPFIFTLRTPTL
MKMAAETTADAQFVYNVQLDSMLGFVTPFDNMRYIAGLDRNMLFHIPVKANVDVDVLKRKLTINVMPRPTRKSNTGVVFHLDSAIAH
DILYEDTRLTYDPQRSTANVVKITAYDTD---
KFLSKASVIDLSITFEGQTVFLMTAASASSPVNGSGSLVFMTSQPAQPLQVCLFGDVSVPIINFSEALKFNPNSSVRALLNFGETCNNGAT
VMLNGQLELTNKRQKYIASPMAKLCESQMMEGNYLLPACRIATESANNLDQYRLKIALNNVTEFQILYNMARDYITENRDVTVHVTAN
LNEQSTALNMTMLARSLLINPAKTVMERIGRTVSPSCVLDTTEANTFDNRTYPLELGNCWYVMMMLVPRLDRVTVLVK-
AVGHKKKVKVLLSSATLDLQPVVEPVQLSINSKPVLLTSSYEFPSGAQLTLNNLVLVYDGARVMLQASNEYSHQVRGLCGTFDGEPFTD
FRNM
>XP023712833_Csecundus
----------MFALEGRTVAGFPSGQQQHTGIHYKGTMRVQKPTGILLLKFQNMTYTKFHENNTLPIGSTVAVKLGHG-
VETVSLKAGMEDWEMNFIQGMVNQLQIDFRPSKNTAYRLMENSVFGRCEVLYEVTRHMNFIEIFKTRNFSNCEYRLHFGLPTSKKQPG
GNACQDFWHRASVTRLSGCGDDFRILQSETNSRVIVNLHLHKGSEASVSSYLTMNLQSIKEVSRPHADPVVLQDLKFRFPLQGYE-----
LHVMKLKETVTEIAREFEASKMIERMTLEKAVRICRKMTLLDLQRAVVNVNKESESKVYRDVLVMCGTNPAFHILKKWIEQRDLRGEDA
SEVMAAVPSHLQTPGPDMIREFYELVRSWAREDQHLKITSLLAFSHLLRTACVEW---THASSWYASMLR-----
AEPSLLRVYIAALGNTGAVAALAPLQPYL-------
RATAVLAMKYQALRVPEKTSPILLSLYHDVGQPVAVRVAAAALLFYTKPELALLQRVAVLTWYDPSLAVAAFIRSSFMSLANSQDPVFSEL
SRYAGMALALVKPPLGLHMPHNFLWSRVVSDLQVV----
SIQLSHETSLDATNFYCRFTQRLGGVSWVAFQGDVFVSGPEKFLTMLRELIYKRSTKWLREALNIAPRSLPKLEGDIRVQFGMAELVFPF
DSELPAELAKAGRKLLEQMETGRYQKV-
NKNEMRVVTELGIPVNFQLSVPWAVRLEGRALMDVDMMYTCAMFSELSFHSAWDETVHMAASHGVTVIRIPTGIRAGL---
SETMQLALETRSTQKWNDGLNFELNRS-----SVQQSAVTTEAP-TPPRQLYF--
FDITETTYLLVKAYVVSLNMAINGTQKLSTLWTYAAGRSGKVRRFLFLKEE----
WQVGFTSILVIPVLKLENVIRANITSEIHADLLYSSDTA-

EHQFSLQTVMGRSALKAEALEEYLTDKCIDRNQGETDISPDCLRAAMDAMFLDRYNVTITYKGQSTSLQHFNSLRLYITDTPSNTITALL
EISPRLNTLNAWILVTEILFNPAVSYARRIRGE--
SYCSHGGSKFFTFDGVELDYNVTSCWHLLAKDCSGHSRFAVLMRSLNNQETELEVNMDNYLIRLRPG---
LNVSANEKPVELAGHRDTPEHVSVALAGFHIVYTGSSTLVMADRSMRGRLCGICGDFDGHAVKEF---
>XP023710404_Csecundus
LLTVVASTSALQLWLPGTEVVYQWETSLEAVTLFPAPTVSQWNMSGKVVVQGAEDLATVQLQDLEVSAYNG-
SDIPLLQPFIIHYHEGQVVSLTINTGEEQWAINMKRGLASLLQLDLSHMHT-
AFVSTESSQYGCCKVEYSVHNSNGTVAVQKFVDFPSCEDFPVQSFGNVQRLCPSGPQETVSSGSTRAYVLEPLGHLLIHSVQADGLTHL
QQ----
SNSYFLQSWITLHLQWQDKMVMTMKGGGERVGLTHQLPSQGRSPPSQKNLLTTVIEILGELADGLDNLEKLHEEHVVHALYLLSMLE
QASLWKLYLELRHETLRNLFLEMLPLVGSHDSIMLVRDLVMTGTVKNDTSIQLLSSIPFSVRQPSEELLTDLEVLLSLPE--
DTLENQAAVLSFATLVHKVCSAKCAPETLDKYVGQYMDLFTGSRDHARQMLYLEALSNFELQRVAQFLAPIISGDDTKYRILAMWTSV
GHATLQPDQAHYLYWPLLTNSSEHLEVRVAALTVLILSKPISTHLYNVMLFMMHERDPHLKHFWYTTLHSLVHTRHPCYRHLGEEAKK
LVQFIPPEVHHWATGNYMLDYTDPDYRFGFLAHLIMVANPRTGLPNVFYQDVGSHTKDFSYNQFSIYIRLDGVEDTIK--
RQLLQFTELVTLLHQLEVPIKPSEPVHVELLVRAQKAVVVHYANQSTLNSFFKVLRNVQLLAIGGI---
HLNAQNVMRVTDLGTNAFLQTSAGGLMSVRANFSSDLELRSMINSQTSFIQYNPVLNMWHGVQRSLNIQMNVPFSLRVSSNATEQSIK
FSFQQTPGEEV--GLVTHVKGVVATKLCPAHMSVTRNTPEMS--GLKFSLLEHSSTEDYLLL---TEVEVSIPAGE-
YRLQIGVIRAPVSSTLPD-RLTAELTIG-
KVCVNLTVNFTEPDTVISLTHPRLQSAMLTAVWGLPIDGSSSLQVLLDGQVTTEQLNLTSVWPYNMCQQDMQQGSHYMPACYEVAY
ELATLRYYNISISSRNHP----VRDFLPSKM-ASLSSNISIHLRVPPEVRPV-
LTVLLGSAVTKLQFPSLLQLVTNTGGSCVLTLHSVWTLDGAHLTTNVTTCYTLAMADCTTQPQFAVFIRKVEGTPLAMELYSDINRIEV
LPTADQLTISVNGQMVPNHEGIHLGEAN------VWLEYYGHFATVSVLGLYHGGLCGLCGDHNIDAGDDTAQV
>MN002568

EAFRDSVAQAGTLPTLSAISDWIHKDVIRGNEAAEILSILPRSALTPTPEYIELLFELVKDQNKMQPFLNSTAVLSFADILREVAVNNNFRP
LTEPYLGKELQKAINEGNSLKIQVYIRALGKTAHPNILQYFEPYLEGQKDVSRLLMVTALTHLANLHPQVVTPVLLSLYENLGEDREIRT
AAVFTIMETNPSVYILQRMAEFTNIDTSKQVNAAVVSAIKSAANLQGPYTQQMATNAKAAVELLTPEIEPSYSASIMRSYMMKHLDLGYE
EHFSQLASRDHVLPKALLTTLMTNIGGETTKPISIESILSSVRDTVDLVNRQFENLSNDDINKLLGIVNKYTEQVEGSIAFNFAGGQKFLPF
DNHTLQNIPSLAKIMAQLHKTPLYTKLYDRNTWVFPNAMGLPFVFTLKTPTLIKFSAMSSTSMQLVHTYKVESVIGFMTPFDDMEYVA
ALDKDVLYNIPLNAKIDVDILNKNLSLTLMPRPIKQVDTGMVIRLDKIASQSILHKNISIAFEPRLSTAKVVKITAYDNMESNEFLSKAAVF
DLSVKFEGQTEFVMTAALASSWMNGNRSVVFLKSHPAQPLQACLFGKISVPITNFHEALKFNGKSSVKALLSFGEKCQSGSTVSLKGEME
QSKMRQKSIANPMAKVCESQMNKKNYLLPACRNVTHAANNLDHFHFKIMFNNVPNVLKLYSMARRYVTENEENTIRVIANINEHSTAL
NVTMWTQALLINPAQSVIDRIGNAMAPTCVVDNTMANTFDNKTYPMQLSKCWHVMMKVVPTLEEVSVLVK-
EVGNNKEVLVLMNSTNIHLTPALHSVMVTINDKSVRLTHGYEMPDGAQVTLNSLVVVYDGVRVMLHASNEYRDQIRGLCGNMDGEP
FNDFAKI
>MN002567

LSDAQYAQINMAPDLQMNEPFQVKLKNGLVKNVIVPGSTHEWQRNMLKAILSPLQVNAQGVSRAVFRAMESTITGECEVVYDISPKSNI
FEVVKMQNFTNCEERIHFGVSGLEHEPGDNKMGNFMSRSSLNRAIINGTNFTIQSSVTTNKVLLSPEFYNNQHGMIVSVVNITIDSIERPK
HSPEDPEDAGDLVYRYPHEGNAANFSSQNVDIAMKLANRIASDLQNASSIEQNSLSRMTRLIRIMNDKELDEVTEKLANIATWRAFRDAV
AEAGTLPALSTIIRWIRTNKLAEAEAAEIVAALPRSVVTPTSDYIKSFFELSEDPKKKQSSLNSTAILSFATLLRKVAVDTGFPPMTKPYFNES
LKKAIHEGESEKIQVYIRALGNIGHPSILQILEPYLEGQKNVSRILMVLALDKLVHIHPQIARPVLYNLYQNTAESHKVRCAAVALLMKTDP
PLVMLQRMAEFTNSDPHQQVISAVQSAIKSAAALEDPYHYELAENAKAVVNLLNPTQDVSLSKSLLHSLVIDDMELGYENSFSFIGSKDS
FAPNSVLEKFIRNVGGRRNEHYSISAMVSSAKEAFNYLKEKFENWNAEDLETLLHIVRDQAQQVEGNMQFSLLGIEKLLAIDNHTLDKIP

QLVKAVAYMYSQPRYTKLYNPVGVTFPNALGLPSLYTLHVPTLVKSGGEVTADIDLVYSTRKMIKMGFVAPFESTRYTAALERDILVNLP
VRVSVNLDIENKRVDAALKPRPTKSFSVGMAFNLQNLVRSSMSYENVTVYHNPDISTATTVKITAYDDNEEPEFLAKAAVIDVSVKFEGH
SMYAATAAYASSLVNGSTSTLFLSAKPAKTLQAFLSVNVSVPIMNFRQALKFNPNSSVSAKLDVIDGYNNQAKAVLKGKMEQSLAFKQAI
ASPMAKYCEAQMREGNNLLSACRNVTEKANILDLHKFTLKYNNIPDSWKAYNVIRHYVTENQPNQLQFYLKLNKNLTAVNVSILAAALL
QYPDSSVADRFGHAMSPTCVVDGNAVNTFDNSTYPIRLGNCWHVMMMEVPQTSNVAVLVRSNSGPGKEVQVILEENTISLSP--
DSAQVAFNDRPIQFSPNYELPSGAQLNFNSLTIAYDGERVMLQASNEYRDEVRGLCGTFNGDVLSDFRKL
>MN000066
-----YGTS-ENGWRPRFEYTFSLEGRTVAGVPSGQQQHTGIQYYGTLRVQRINDVLLMKFQNMTYTKFHENNTLPLTSTVAVKLGHG-
IEMIKMKAGLHDWEMNLIQGMLNHLQVDFRALQGTTYRMFENSVFGRCEVLYEVTRHLNFVELFKTRNFTNCEYSFHFGLPVSKKHP
GGNACDNFWTRTSVTRISGCGDDFRILQSETNSRVITNLHLHKGTEASVSSYITMNLASMQQTTRLPEETVELRDLKLKFPIRGYD-----
QKIQKFKETVADIAKDFEMSKVIDEMSLEKAVRICHTMTLYELKRAVENVSRQSESKLYRDILAMCGTNPAFQILKTWIEERKLEGEDASG
VLATMPAHLQTPSPDIVRQFYELIRSPALDDQQLKITSILAFSHVLRAACVDK---HQASEWLVSMLQ-----
SDSASKRAYIAALGNTGAASALPHLQPYI-------
RATAVISMKYQALRVPEQTSPILLSLYHDVGQPVAVRVTAVAMLFYTRPQLALLQRIAVSTWYDPNMAVAGFVRSSLISLANLDDPAFSE
LCRYASAVLPLVRPPIEMHTSYNLLWSKVSGDLQLV----
MNQLSYEGSLGATNLYYRYTQRLGGTSRVPFQGDVVVSNPDHFLTIIREFIHKRSTRWLKEKLKIAQRTLPKFEGDFRVQFGFVERIFPF
DTELPVELARAGRKLVEDVKTGKYQKV-
DKTKMQVVTELGIPINFELTVPWAFRLAGRAEMDVDMMYSSHLLSELSFYTAWDETVHMAASHGVTVVRLPTKLMASA-
PQSSLVQLVFETRNTQEWNDGLALELNRPEQQGASTQDLNMTLEQQNSSQEHIQKESYDMTRDPYLLLRAYVLGLDVRVDGAQDFST
LWTYAAGRSGRILRFVFLKEA----WQGALTYSLITPVLKLQNAFRANFTPSVQADLVYGNEDRLEYQLSLQAVMEHSATK----
GDRLKGKCIDETEGGTDINAECIQAAKEATFMDRYNVTVTYNEKCKSFQYFNYLRYYVTEDPTNWIKGVLEIAPSLETVNAWIIMMDAL
FNPTVSYIRRLKGE--SSCGHGGSKFFTFDGATFDYNLSSCWHLLAKDCSGQSRFAILMRSVHNNETEVEVNLDNYQIHLRPG---
PDVSVNEKPLRLSEEAEWPEQFSLNMAGLHVIYTGSKTLIMAERRMRGRLCGICGDYDGDTVNEF---

>MN000928
ALSLVVNMAALELWLPGHEVVYQYKTSMVTSTLSPEHVVSQWNMSGKLVVQGGEDLATVQLQDLKVSTYTG-
RDLPLLQPFLIHYFEGQVVGLSVKPGEKEWSINMKRGLVSLLQLDLLHLQS-
AFISTEESQYGHCSVQYSVQSSNGTVTIQKFVDLPSCVNEPVQKFSNVQKFCSSGPTATVSSGSMRAYILEPSGHLVIHSIQAEGVTHLQH

PNSYFLQSRVSLYMLWQHKVEAPMREEVERVELVHQMPSQGRSPPSQEDLLTMVTEILGELADSLENLDKVHEEHVVHALYLLSLLELA
SLHKLYLDLQHETLRNLFLELLPLVGSHDSMLLVRDLVMRGTVRNDTAIRLLSSLPFNIQQPSEHLLTDLEVLMLIPE--
DTLEHPAAVLSFATLVHKVCSSGCLPETVDKYVGRYLDLFTGSKDYAHQILYLAALTNFDLQQVVQLLMPLIRGDDGRYRALAMWAT
MGCSALQPDQVYHLYWPLLTNCSEYLEVRVVALTMLVLGQPSGTHLYNLMLFMQHERSAHLRHFWRTTLHSLAHSQHPCHIHMREV
AKKLVQFVVPEVSYWTGGNYVLDYADPDNGFSSQASLIMVSSSRTGLPTVFNLDVESHTMGYSYNQFSIFIRLEGLGAVIK--
QQLMQVTQLVTVLQQLKMSTKPSEPVHLELLVRVQKVVLVHHANESTLNTFFRVLNSMQLLANDGI---
HLNAQKIMRVTDLGTNALLQVLSGGLVSVRANFSSGLELRSSLSYQMSLSQYNPILNMWHGVQRSLNMQVNIAFNLQISFDATQEGVRF
TIQHKPAKEM--GLVMHVKGVAATKLCPVHLSVTRNVPEIS--GLKFALLEQSFSENYLML---IKVEAIFPAGG-
YRAQLCMVQSLVSSTSSS-TITMELTTS-
KMCLNLTTSFLELDSAGSLITPRPHTAVLSVAWGSLISGETSLQVAMEGHVTTEQRNLVTAWPYDACKRNIQKGKHHIPACLEVARELT
TLRYYNISITSINLP----LRNFLPSKV-EALSRNTSVHLLVPPEGRPM-
LTVMLGSAVTQLKFPPLLQLVSHLGGSCLLTEHSVRTLDGLTLTTFVPACYTVTMADCSSQPRFAVFLKKEESTPLAVKVYSYDNYVEV
LPTADQLNISVNGQIMRSQEGIHLGDGD------IWLEYYGDFIIVSVLGLYHGGLCGLCGDHDGNPSNDAQQA
>comp77743_c0_seq1_Cpunctulatus
-
LCMLVTAVALPAWDPSSRYVYRVQGRALTALHQVSNQFTGIKIRAELAIEPSRQEFRGQLSRAEYGHTNAKPDLPINKPFEVKLKNGLVE
RVIVSGSTKEWEANILKGILSPLQVNAEGVHRAVFRTMESTITGECETVYDISPDHDFHDIIKSQNFSNCKERVHFGATGFEHEPGQNAV
GRFMSRSSLNRVIISGKKFTIQSSVITNKVLVSPEMYNNQQGMIVSRVNLTIKSIGQQEHAPENPRNVGDLVYRYPDGGNSIGTSRNPIEA

VTKIANRIGSDLQNASSIEQNTLARMTRFIRTMNRQQIDETTKKLDSITAWKAFRDAVAEAGTLPALANIAEWIRSNKIAEEEAAEIIATLP
HSAITPTTDYFKTFFELTKDHRTKQSHLNSTAVLSLAHLYRRVAVDSNFPHLTKPYLGRSFKEAIDEGDSYKAQVYIRALGYTGHPDILKY
FEPYLEGHRDSSRFFMVSALVHTAGLHPKPARAALFKLYQNIGESHEVRCSAVQLLMKTSPPANIIQRMAEFTNADPDPQVVSVVQSAIR
TAADLKEPHHYEFARNAKAAVNLLNPARDISFSRSFFKSYTIKDMELGYDQSFSYIGSRDHIAPKSLFAHFIRNMGGSRNENFRFVGMLAS
LKDAFKHVQEKFEHWSSEELEKIIGIIRDHAEQVEGNIDFGFVGSQRFTAFDNHTVDTFPRFVKEFARKYSQPHYTKFYNPYRLTFPNAL
GLPSIYTQKVPTIIKSGGGATADFQFVFSTRKMSKIGFISPFEGARYIAGLERNLLVNFPLRVSVDFDIKNTRVNAVLKPRPTKSFKTGIAFN
TESFASKSIHYDNISVSFDPDHSTAKAVKFSAYDDNDEPEFLAKAAVVDFSIKFEGDTVFIATAASASSLVNGSGSTFFFTAKPAKSFQTFFS
GNISVPILNFHEALKFNPESSGKARLDFGD-
VGTQTTVFIKGKLKQSSGFRKFVSDPVAKQCETHMNQKNHVTRACRNVTEKASRLDQAHFTIKYKDVSRYFKTYAIFRHYVTEDQRNR
VEIGFQLNNNATAFNISIYAAAAFFNPHYTAFDTVGRHAFPTCVVDGNIANTFDNRTYPIKFGSCWHVMMIHVPHEDNPVVLVRSSSGT
EKEGMAIVGKNVFSMRP---TVEISVNGQSVRFSPHYELPSGAHFTFNSLTIAYDGSRIYLEAGNRYRDEVRGLCGTFDADAFTDFIKL
>comp194989_c0_seq1_Cpunctulatus

MAEFSNHDPNNQLIAAVQSAIRSSANMKGPYTLKMARQAKAAVNLLNPEMDLSYSESFMRSYMVDDLDLGYEETFNYIGSKDSFAPKSV
CFKLASNVGGATHEKLNIAPMLSSVRDTFNLLKEQFVIFTDDEIEKLLHIVHEKAEQVEGNLLFDFAGSQRFLSFDNHTLEKIPKLVKELA
KTYKHPQYTKFYNRKAITFPNALGLPFLFTFRGTALVKISAETTTSFELVYYYKHQSQLSVVTPFDSMRYIAGLDRKYLHHIPAKAKLDVD
ILNTNISLTLSPRPTKAYNTGTFFRLKLAAADAIYYTKLTLDFDPQRSSNKLVKLTAYDNNDEDEFLSKASVFDFSAVFEGETEYILTAALA
SSLVNGSDSAVYIKAASTHEYQVCLAAKNLVPIINFAEAAKFDPTSYASALLNFGETCKSGATIRLKGEMEQTKKRQKFITSPMAKLCESQ
MRKKNYLLPACRNVTERANNFDRYHFKITFNDLAETFKLYATALQYVTENPRNAIGIIMHLNEHSTAFNVSIYAKALFVNPQHSVLDRFG
KTAFPTCVVDSTEANTFDNKTYPIELGNYWYTMMKLVPTLDNVTVLVK-

EVGEHKEVTVLLSSANFGLRPPLIQVELFMNGNSVDFTPSYELPSGAHITFNSLVLVYDGERVLLQASNEYRDQVRGLCGNFDGEPLTDF
RKI
>comp89049_c0_seq6_Cpunctulatus
LLALVGFTADKNLWVPDRELVYRYEASLGSGTELQGVTISRWNISGRVVLQGGERLATLQLQDLKVSTYNG-
HDIPLQQPFIIHYDADQVVGLSVKAGEEEWVTNMKRGLASLLQLDLPHLQE-
AFIAAENSQYGRCNVEYSVQSSNGTSAVRKFVDFSSCEGFPVQSWRNVQKPCLSQSQEAVSSGSVRSYLLTPTYRHEIYSVQANGVTHLQ
LPRDQEESYFLTSWVSLDLMRKGEVKKIVEGGGVRVGLTHVLPSQGRSPPSQKDLLAAVAKILGELVHSLEDPGKLHEERVAHALYFLSL
LDLASLKKLYTVLSRKTLRNLFLEFLPLVGSRDSILLVRDLVTGGVLKNDTAELMVSAFPFNVRQPSEQLLADMEVLLQLPE--
DMPVHDAAILSFATLVHQVCSHSCAALTVEKYVGHYLNLFTGSQNHEHQMLYLEALGNLESQRVTEFLAPIIRGEDPRYRFLAAWATQS
HASWQPKEVYRLYWPLLTNRSEHLEMRVAALTMLVLSQPTQARFYN--
LFMLRENDPHLRHFWYSTLRSLEHTKWPCYRHMRDVARNMVGLVLADEGHWATGNYILDYADPDHRFGSLAQLLTVANPRTGLPTV
FYLNVGSHAMDFSYDQFSVYVKLEGFGDEVK--
RLLLQVKRLVELLQRLGVPSETSVPLHFELVVKLQKVALVRHLNQTTIKSMYKVLASVQQMAVSGISGLHVNGQRVIRPTDLGTNALLR
ASSGWLVSQRANLGIGQDVRSFLSSQIALTLYNPVLDMWHGVQRSIDVHATVQFNLKTLIDPKEKTVKFTVLQTPGDEI--
GLVAHVRGIAATKLCSEHVFVTRNKLDLI--DSSLAPLEYSYDEDYLLL---TELVMTVPAAD-YSLQVDVVRTPASSASSN-RLTAEFNIG-
KACIDFTTKFPESVMAASLTDPENQTAALNAAWGLPTDETSNLHFTLEGDLSAEQRVLASTWPYDECRNDLLQGNHYAPACYKVAQE
LETPRHYAISITSNSLP---TLQGLLPTMLVVPLSNNITIDLEVPSEARPS-LTFALGSDTSDLKMRSMLQRGTSKWDSCLLTSSSVWT-------
----
>BAA86656_Pamericana
FLCLLVAGTSLQEWDPNRQYVYKVESRAFTALHQISNQYAGILMRAKLIIQPTREELHAQLVEVEHSQINKEEILSAKDAFAIILQNGHIVN
VKVKDSTPNWAVNILKGILSTLQVNTQADNRAVYSIRETTITGECEVEYDVSPNDNVIDIEKTQNFSNCKRRVHYGLAGIPDEPGQNQM
GDFLARSSVSRVVISGTKFTVQSSVTTNQVVMSPEMYNSQKGLIVSRVNVTIKDIEEARPLPGNLQDTGDLLYSYPNRGNAAFLLSDPAEI
VKSLAEEIKSDMKKPAYIERSTHAKMRDIVRTMTAKQLQKATSLI---

LGWIAYRDMVSESGTHPALEELSIWIISKKLSSEEGAELLATLPRAVIMPTPEYFEAFNKLVMDKRRNQPIVNSTGLLALATLHRQVH-
DANYNA-
TNDYLGKQLHAAMADGNRPKIQVIIRALGNTGNKRILNYLEPYLERKKNATRLLMVTSLDILAEINPELARQVLYNVYINIGENHELRCAS
VILLMRTQPPAAMLQRMAEFSNIDPVKQVVSAVQSAIRSAANLKEPGNLNLARAARSAVNILNPSMDIAYSNDILSSNMIQDMDLGYKD
NMAHVGSSDSIIPNTILRKFNRYAGGQAHSDINFSEMVSSVKQLLKALRNPLK--------
NIPPIIIDIAPPLEGNMMLRWLGNDRFFSYDKNDIKQLFRNYNAAALPLADMHDMKVYNQKSLAFPNALGLPSLFTIDVPTVLRANSTFR
SETSLTYSVKEMSKIAIVTPFNSMEHMAALERNILINIPIKLDVDFDLEAQNIALNMKLRPARMMRTGIALNRQELARRTINNHNVTLSID
EDQSTTNAVRIEGY------EFLNRA----
VRVAFTGSSTFDLTAALGLSNVNGSARAVSYISQPAHATEYDLFAALSPPIINFNQALEFDPDSNLDAGLSIFTNDKPSGNLRIKGELQQSE
ERRNAIRTPAALACMREMANNNNLLPSCRNATEMANRLDRIRLQAKFENLSDDLKAYTWIRYYVTENNPGRLNINVDVNNDGTALNAS
VWTRSLVPSPQDTALDRLAREALPTCVLDVSQAATFDNTTYPLTLSSAWHMMLQYQPIRPPVAILARETSNRRKEILMNLDHTIVAFKP-
-TSVNVEVNGRILIIEQSHRLPSRAYMVAQKLLMIHDGKRVLLQASNGYRDEVRGLCGTFDGEPTTDFNAW
>BAB32673_Pamericana
LLCLLAVSAALDGWEPGKRYEYHVRGRTLTALHEVANQYSGFRFKGKLVIEPTPSVLRGQLKDTYHMTVHRMEEVGMNKPFEVHVGE
FQFNKLIVTEDTPVWETNMIKGVLSQIQVNLK----
KIFKVHESSVTGRCEVLYDITPENNVLQVMKTQNFTDCKKLYVHGFYNFHNFPAQNKAG-
FMSRSQQTRTIVSRNRFTIRSSVTFHEVVLKPELFNSQQGISVSRMNVTLEEIKSQQHPPRPPKDVGDLVYRYPNQGSKIGEVD---
EKIVLLARTISSELQEPDTMKKNILSRLTNLVRAASFSQLEEATKRLDESDAWKAYRDSVAQAGTPAALKMVHTWIRKEYIKDEEAAKVV
AVIPHAADTPTDNYIAYFFEMVKDPVHGEKYLNSSAVLAFSKLLRLAAVDSNFSARVKEYFANKLKNAVKDKDSHKIQVYTRALGNTGH
ADIIRHFEPYLVGRESVSRVTMVFCLDEFVKTQPSVAQYILLRLFENVGETQEIRVAALYLLMKTDVSAELFQRLAEYTKFDKNHQVVSAV
QSAIRSAAKVEGPYKKETAKNAQAAVKILSSPYDDSYSKSFILNNYRREIDVGYSRLYNQIGSRDSFMPKSVFYKLVNIIDGDRDDQAKFGG
AVSSVRDVIDFIRQQFK-WDLREIANLLEMEEENVDPLEGNVHYDYFGAQRFFTLNKTSFE----
FREELKKYFKKPQINKLYNRMELKYPNVMGVPFFFTFKRPTLVKLTAKTTSDVSFVYSFDMHSHMGVVAPFNKKEYVTGIQRKHMIQIP

LNVSVHVNLDKNKVAADFKPRPSKAYEVGMVYHSNYFEAQPVFYYDFKLYLDSHNSPAKTVQLKAYDNRYTQELYDILYYASVSVAFKG
EEKYEVEGALASSQVNEHISTLRAHSNDAEK----
VRVNVTVPVIDYRKALEFDPTSKIQCEVHFGDTPEKKSKVYFQGKFERTDERKKFVASDMAQLCSAQQNNKNYLLPACRNVTEEASKLD
KYFFKVKYENLSEKCKAYSYLRHYITENKTDSVEVQVQFNEEINAVNVSVYARALFLNPRYPLLSQVAKTAFPTCVVDYSKVNTFDNRTY
EHDMLNDWVEVMFHKPIYKQVSVSAK-
QAHSKMVLKVLRGDEKFEMKQPRDSPELKMNDKVIEYSPAYALPSKAHLDLNSLVFVYDGERVMLHAGNHYRNQVRGLCGTFDGEPS
TDFLDR
>BGER000125_Bgermanica

PVLPMSQPFELQIKNGLVQELRVAQTVKEWEGNLLKGILSPLQVNDLVQNKAVFRAMESSVTGECEVLYNVAPKANLMNIVKTQDFNH
CKRRLNFFLKGLNHEPGQNQMGSFLSHSSQTQTVIEGSNFTIQSSVTTNEVIVSPEMHNKQKGIIASQVNLTIISIKPISETPVELHRLCNL
VYNFPNDGNAINSS-
EPVQAARNLCMQIALELQDANSIHKNTLAKLLDVLRTMNLQQLRNLVETVTQSMISIVLRDAVAQCGTHPALIIIIDWIEIQVLKGEIAAQ
VLSILPNAALLPTNEYMEDLFKLVKSEKQAQAILKRTALIVYTNLLHKVAVNHNHKVVRELYIAQELKKAIQESDSLSIQTLIRALGNIAHP
SILPHLKPYLEGVEKVTRLLMVASLDKMAFVYPKIVQEILFKLFENIGETHEIRCAAVFLLMQTKPSVVLLQRMAQLTNSDPHEHVIAAVQ
STIREAAALKGPTNIHLARNAQAVLHLLNPKFNIPRSRSFMRSQRIKDTQQETEHLLNFIGSTDHNIPNSIFAHINTIMGGAKKEHLHLAI
MMSSVKELIRLLQIQQQ--------
ENFKLVHEIAEEVEGNLHARLMGADRFLPFDRNTILTYIQELKTWLENPNRKTFTKLSEVQSMTFPNALGLPTIVSVKNPVLMNINLDAS
ANIQTTFSTWKQSKIGFVVPFNNMRYIAGVNRQAMVNIPIQMAVELDIQNQHMSTVLQPKPIQNVNTGMIFQLANLMNNIITHNKININ
FNSQDSTAKIVKFTFHQQQEVPEFLIKSSEIDIGIIFIGSDSFGLTVAHARSLLNPSFNIAFLRADPAKSFRAYLSIQNSTNFLDLQAALDSNP
NSTLKAQLNFGNKNSNLAAVHIESLLGRSLTHMENVLDPITAVCQAQMQNGNKFLPACRQAILQATVLDHIRAVIKYQGLPEPIKAYSL
LRHTVQENNPNQVEMDLQFNYEMTEFNALLLLAPLLLNPAPQNKISQLIPFLPLKSIDLPPA----------
HVRPFAMPPDRRYDFRLREENDLSVHVYVDLYIEVINLRPAEENIPVVIHDGEIVKEDKYEFGYEYNITLKALLVDYDSKNVVLKSHKHPH

DEVRGLCGTMNGEQITDFEEA
>BGER020452_Bgermanica
LLCLRTSV--
FSGWLPNSQLDYHVHGRTFSSLFQVANQYTGILYKARLSLDREDQLITGKVTEAQFSPVNTQPDVPMQQPFQIELNRGEVRKLRVNKF
VELWEINMIKAIISQLQVV------
KVYRVFESTVTGRCEALYEVDHHDTKLRIMKNHQFTNCRHNYKLHFNAFEYHLKQHKPETFLSNSAVSRVIADGDNFTFYSGETIHKIV
LNPEIYNKQKGMLVSHINVTVERKGRELTIDYELRNVGDLSYSTPFQGQDMSEEEDPVGVAVRLSKDIAADLQEPRVGDRHILPRLVRLL
KQLKVSQIMEAARKLNYMDTWRVYRDAVSQAGTWSALNSIQQFISSEMVEPKEASHLITVLPAAVSDKNKAYLHFLFEMTKDPVKNMT
YVNTSLVLAFSEVIHQVEMHQRDLKIKSPYLVQEFDDAVKENNSIKIQLYTHALGVTGNTHILHYLRPYIIQLKTITRLFMVQSLERVVEH
NPRKVIDLLLSLYLDQNEHADIRVEALFLLMKADPSIHVLKMVAELTHTESNNQVLSASQSAIKSAANVEGDIYSEMRRKAKAVEHLLST
NMDVSYSKSYLYGYKSKKINYDSLYNLNYIGSEDSIYPKSMLLNIFTNNLGRINTHVQKGYMVSSMTDLWEAFHTIYK----------
KFVEGNLKYFNMGVQKFWAFDNTTFSNAS----------AA----------IY----------
SLSIVFRFSQMYHAQLAFSTAFDNKEYISGLDRKVEVHVPVKFQINLDFKNHNGFIRIIPRKTAHFETGIVFKFTGLPTKDIFYNNLTVYYD
HEDTKNHAVSFTV-------
EYLSKAVVLDMTLKFEGPAYFTTTVSHATSLVNGSSNYLFYDQHYYEKNQFCLSWSVYVPIMNIYSAFEFDPNSKVHAIMNIGKECENGG
SAVANIDMLRLSEHLDYVKLTVSKLCDHEMRKRDHVLPACRNSTERASDLNRVHVDINYNLKHETFKVYDFVRTHVSEDSAN-----
FTLKDNTRAFNVSIWQSEMLMNPRTSFAKRFAKWALPTCVVDSSYINTFDNFTYSAHVQND-FYTILDIPEYFKVAF--
KPTSPVQREVLVFLRNAKIELKPNQGMPEVYVEGKRVDYNHHYALPTKAHIVFSEIEMFYDGSRIMIQASNMYRNFTKGLCGNMDGEFV
NDVRKI
>BGER006422_Bgermanica
LVVLAGTLDS--
FWVPGTEVVYNYYARLDANSVEPEESMSRWAMKGKLVVQVSGETAMVQLQDVTVSAYNGPYDPPLLQPFILHFENGKVVNLAVKAGE
DMWAINMKRGLAGILQLDPTVLYS-

ASVSTENGQWDNCTVEYTVHNTNGTVLVTKNFDFSKSLGSPKKTVTNTRKFCSVKPYNPLDVISLRHYKLEPTGELVISSVVADGGVYY
HPFQAQTEGLTIWTWMELHQETLGTIKESVKGGGEKVGLRYEMPSQGRSPPSQKQLISAVKKMLQELGESLENLTKLHEERVAHILYFL
SLLDHASFKQLYSDLQHETIRNIFMEVLRLVGSHDAVLLIRDLVMDGTLKNDTAIKLLMALPFTVRDPSEQLLKDFEVILTLPE--
DSVLHSSAVLTFATLVNKVHKTGSRSHIADKYVGQYFDLFNKSHSYEKQMLYLEALGNFELNRVAEYLAPIIAGTQSRYRYLASWTTFSY
VGAAPKEAHKLYWPIFSNYDEPLEMRVAALTVLLDSSPTASRMYQLAILMLEDPSVHLRHYWYTILLSFTQSKNPCNQDMVDLAKKVIM
YFPPSVRHWSTGSYIMDHADPKYGFGSQVLLNMVANPNTGMPSTVFLEFENYAMQYSYNQYSLYVRVDGVGDLLK--KSL--
IEKLVGLLQELKLPLKAVEPVHIELILKVQKAVLVHHANETSFTSLFEV--------GGGLSRDVLSLRRI----------
----------IHKN----------RNK----------

>BAB19327_Rmaderae
--
CLVVGGAGPRGWIPQRRYTYNVEGRTLSAIHEVSNKFTGIFLRGQLILEREPTLIRGQVREAKYAQVNQDPDLPLQDTFDAHIDSGEVEI
LYVNSQLQLWEVNVIKGLLSQIQLT------PVYRVKESIITGRCHTLYDFSPDNDNLQVTRTQNISHCNSHFHLKFSGFEH-----
NNGGFISNNVVTRMVVDSVNLTVIASNTVHKVILSPEYYNTQHAMTVSFMNV-LSRKSNQLSPVSDPRNVGDLVYLEP----------
DPVKAVVQLANDIGHDLIDPDSLDKDTITKMVRVLRNLQLSEILDIAQQLDSPEAWKSFRDAVSQTGTHAAVHSIIIFLSRRYISQSEAQD
LFNVLPAAVQHHDMQYINNMFDLIKDPVQQDRHVNETVVIAFSNAYRFIHARL-------
KYLFQEFENAYRRQNTTQMQVYVHTLGNTGDVRIIPYLEPYLLRQIHLSRAHMFKALERVVDANPHLLTRFFLKFLLDQTDHPDVRVQ
AVFLLMRSDPSVAVLRTMAELTHSEPVNQVVSAIQAAIRTAARLRGTRFYNLAFKAQTVVNLLSDNLDVSYSKNYMLDQEAREYNLDFQ
LFYEQIGSQDNLLPKSALLDIFSYVGGAKSDH-QTGYTVSSIDKVLNDIKNFTQ----------
KLVEGNIQYQVLGVQRFWPFDQDSIKSIPNVIQKFVKDYREVKLTKFFTTSTGIFPTVMGFPGVYTLHTPSLWKADGELQLRVRPLYAA
KLQSKLSVITPFNDMRYTAGVNRHFQLHVPVKIHAEMDDKNNNIRANVKQRPMKKYNVGQAVKLSPLMTSQIIYRKLGVTYATRQCT
NNVIELSAY------

QYLRASYVLDLGVNFKGQNYIVFTGAYAKSLVNGNSNHVFYNQQFLKEKQVCLSANIMMPLNNYDDALQSDPTSQVRMILNAGNKCQ
ESGQATVEESFKELRSMKNSSKTTSTAIISSE---------
HCQNVTYRADDLKDYTFRAIYDNKPDFVQAYALLRNHVSEDKANQLDLSVQLNNVSKVFNLTLWAGNMLVNPRTSIAERLAQYELPT
CALDNSAINTFDNKEYTLM-----QVKDQDTTRVRKIDV--
RKVQDSNKDVKIITEKATVQLKHNNDKPDVYFQDRKISYTNNYGLPKKSMVVLQNVAFVYENQRFLLQASNIYRNKTRGLCGNMDGE
EITDLYFP
>BAD72597_Rmaderae
--
CLVVGGAGPRGWIPQRRYTYNVEGRTLSAIHEVSNKFTGIFLRGQLILEREPTLIRGQVREAKYAQVNQDPDLPLQDTFDAHIDSGEVEI
LYVNSQLQLWEVNVIKGLLSQIQLT------PVYRVKESIITGRCHTLYDFSPDNDNLQVTRTQNISHCNSHFHLKFSGFEH-----
NNGGFISNNVVTRMVVDSVNLTVIASNTVHKVILSPEYYNTQHAMTVSFMNVTLEKKSNQLSPVSDPRNVGDLVYLEP----------
DPVKAVVQLANDIGHDLIDPDSLDKDTITKMVRVLRNLQLSEILDIAQQLDSPEAWKSFRDAVSQTGTHAAVHSIIIFLSRRYISQSEAQD
LFNVLPAAVQHHDMQYINNMFDLIKDPVQQDRHVNETVVIAFSNAYRFIHARL-------
KYLFQEFENAYRRQNTTQMQVYVHALGNTGDVRIIPYLEPYLLRQIHLSRAHMFKALERVVDANPHLLTRFFLKFLLDQTDHPDVRVQ
AVFLLMRSDPSVAVLRTMAELTHSEPVNQVVSAIQAAIRTAARLRGTRFYNLAFKAQTVVNLLSDNLDVSYSKNYMLDQEAREYNLDFQ
LFYEQIGSQDNLLPKSALLDIFSYVGGAKSDH-QTGYTVSSIDKVLNDIKNFTQ----------
KLVEGNIQYQVLGVQRFWPFDQDSIKSIPNVIQKFVKDYREVKLTKFFTTSTGIFPTVMGFPGVYTLHTPSLWKADGELQLRVRPLYAA
KLQSKLSVITPFNDMRYTAGVNRHFHSRLPVKIHAEMDDKNNNIRANVKQRPMKKYNVGQAVKLSPLMTSQIIYRKLDVTYATRQCTN
NVIELSAY------
QYLRASYVLDLGVNFKGQNYIVFTGAYAKSLVNGNSNHVFYNQQFLKEKQVCLSANIMMPLNNYDDALQSDPTSQVRMILNAGNKCQ
ESGQATVDGKLQRTKEYEKFIK-
ALARECQNDMDKRNYLLRDCQNVTYRADDLKDYTFRAIYDNKPDFVQAYALLRNHVSEDKANQLDLSVQLNNVSKVFNLTLWAGNM
LVNPRTSIAERLAQYELPTCALDNSAINTFDNKEYTLM-----QVKDQDTTRVRKIDV--

RKVQDSNKDVKIITEKATVQLKHNNDKPDVYFQDRKISYTNNYGLPKKSMVVLQNVAFVYENQRFLLQASNIYRNKTRGLCGNMDGE
EITDLYFP
>BAA85987_Gnigrofuscata
YMLLVAGAANNTGFEADKQYVYKLESRTLTSLNQSASQYTGIVIRATVLMHLEEGVAILELNQPEYGKIHGEPGLYLSSPVEVHYRNGVV
DSIRVSKNMPTYELNIIKSIVSQLQFDTEGENRVTFKTMEASVTGKCETLYDFSPNDGLFRVVKTINYDNCKLRYNFGFNQVPAKPGTNG
MENSPVRSTVTNAIMTGKRYNLESVETVNEVIVSPNLYDRQKGHVVSSVNMILTDVRPTPSVVGEAYTVNNLVYEYPNKGNSIVSSNKTY
DNVVKLANEMGEEIVDASSMGLNTVEKLSYIIGTLNARTIEYIGEQLDSVQAWVAYRDAVVDAGTGPALVALSSWIQNKKVYGMEAAKL
LSALPDTVKTPTNKYLKALYDLAFSPEINERFMNTSAILALGNMLRRTQVNK-
NNHVLHDKISGLLKEAVAQEDKPRIQVFIRAIGNLGYPKALARLSPYLEGKQPLTRLLIVTSLDKLVINYPKVARAMLFRIYEDMTENHAL
RCAAVHLLMKTNPSAVLLQRMAQFTNVDQSQHVISAVQSAIKSAANLNGET--
ELARNARAAVDLLNPDFGAQYTKNIFHNYISDEMYQSYDQYISYIGSEDSMYPSSAFYSINRQFGSSSKEKFSLEGMTSSIDSLMNFLGENF
S-
WTPEKIAQILNITASRAAPVEGNIQSKLFAGKDFFAFDSHTIQMAPKYLQNVLNELKSGKITKYSNRLHTVYPMASGLPFFHSYSIPTVIYL
GGKSSTGLEMTYASKARSRYGFTTPFDHTAYIAAVDKNFQLSVPVHVKVDANFEKNRVQIKVQPRQQFNFDTGFALRYKALTEADNDQ
KSFNISYVPDLSTAGDAKFNKYDAKGKSEFLTRAYVVDIAAEFNGQTHFIVTAAVASSPVSNTSRLFFFNKTPAQTVQVCLAAAAAFVPLVN
FIKALNANPNSSVNVQISTGHDC-
TGASINIDGKLQQSYERFQYVRQPLSRLCINQIEEGNQLLPACRNVTTQANLFDKYSFTATYKDISESMKGYAVIRNRVDENSNNRVELS
VNFNADHKFFNASLLFQSAVLHPYYSISERVAQVWLATCVIDNSEGTTFNNRSYPLRLQNGWYVMMHHTPKSGSASVLVQELPSHHKVI
KMLLGNDKIDLKPSDYAANVIV-GTHVQVSKTYALPSGVRLFAKGLEMLYDGSRVKLQADSGYMSKTHGLCGPYNGEYIDDFWSA
>AAC32024_Pnipponica
FLVLLAGAATLQAWKTDTEYQYAVRGRTLSALHDVADQYSGIIMRALLTIQPSDGTLEAKVTKARYAQIHTKPDFPMEKVFEIKMKHGV
VRDLIVDKDVPTWEVNVLKSIVSQLQVDTQGENKAMFKTMEDTVGGRCEVLYDINPDGEIIDIVKTKNYSNCEQRYHFGIPGTNKEPSS
GASGNFLSRSSVSRVIVTGNSFTIQSSVSTNKVILSPNFHENRKGMVSSRVNVTLVKIGAPSTA--

NPESTGNLVYNYPNHGKSIHKTDEPITMAKTLASQIGSDFQDPNSIDEQTLEKLVRIIRTMSTQQIAEAERDLQSVGTWAAFRDAVAQA
GTAPALVTITLWIKQKKIRGVEAASVVGVLAKTARTPTREYIDVFFELATMPEIHEPFLNTTALFTFSELVRYSQMDNINPEVFRPYLAEKL
KTSIEKNENAKAHTYIVSLGNIAHPKILAVFEPYLEGKMPASRLLMVISLNKLATLKPKLARGVFYRIYKNTGEAHQLRCAAVTALMSTNP
PASMLQRMAEFTNEDHSKHVNAAVKSAIESASELETPQWQELAENARNAKPLLNKSYGFEYSKLYLTDFIEREMNVAYQAQASFISSDDS
YVPEALFVKARAIFGGFTFPRTSAGAMVSSAKDLLSAFEDAFK-
FSPENVAKLLKIKSEPNEDVEGAFMLQSAYNNKFMSLDKNTLKNLPELSSKALKALKDGHNTARIESYEMTFPMESGFPFVYTMKVPSM
LKLAGSVHVHLRAVYSVQFQSKLGFVTPFEHQHYMAGINKNYQAYVPLRVNVNYEDSDKRVTLKIQPEEVSKNEHGVKFTFHLFVEQP-
VYHKVDIFMQPSKSSEQSIILSTFETLQDDKLLKEAFAVDIGLEYPGQAQAFATLAFASSEVDEKSRGVYWRLTDAGD-
ESCTSLEARRTSDFGKDMLDQ-
PEFKFDIDLRYGQTCEKGYKVEIQGSQEQTEKYRDTIKLPTAIECIRAAKKGMKSLPACHDVSVKMTMLDKTRITIKYDKDSREMDYLDS
MSDKKRLNYDDTVKIAVDLSPNDNNGRFSYGIGPLVVHPDLETAERLDMDMSRICVLDSHSATTFDESTYPINLGKCWHVVMTTFPED
MALSILVRDAGDKKKDIKVTLGDKELQFSHADSKNKVTLDGKKVDLSEKLQRPDGTGLVSKDIDAWYDGQRVQIKASGKYRSDIRGLC
GNFDGEPDNDFLEH
>AAB72001_Rclavatus
IICLLVGLASDNPWKAGTKYTYQIQGRTLTGLHQVADQYTGVMIQGQLTVRASDNKATAVIENAKYADLQANQELPIKDPFEIHFRNG
VIEKMVVNKDIPQWQLNFYKSIASQFQVDTQGKNKGVYKTMEDSVNGIYETLYDISVDGEFIEVVKTRNFSNSNQHYRFGSSGHEGEAG
TNQMGDFLSRSSQSRVILSGKRYTIQSAVTTEKVLVAPHVYNNQKGIVVSSMNITLQQATGSQGSVSNPREVKSLVYEYPGYYMGSMGQ
QNVPEQIKSLAKKVGQWVQNPGQIEENRLPLLTRMVQNANSQQLEQVTQSLERQQSWAAFRDAVAQAGTGPALVTIKQWIQSKKVQ
GEEAAQLLAVLPYTARYPNTEYMNYFFDMVTSEEQRQRSLNTTALFAFTELSRKSQVDYHRQAVPQPWLQERLERAVSRGDSIKAQVYI
RALGNTAHPKILAVFEPYLEGKSQISRLSIIASLDEMTRTHPNLARDVLYRIYQNTAEKSEIRAAAVRQLMRTNPPAQMLQRMADFTNYD
HSHQVNAAVKSAIESAAACTDSSDQQLVNNARAAANYLTPNYGVQYSKTHLKSHNVEGQNLQYQQHCSSIQSQDSSLPSSLMYEVQQS
VGGYHRDPHKFYFMTSSSEQAINLMRQTIAEQSYENILRSIHLHTDEPVQVEGNILVSGLGGKEFFTFDNHTIEQIPQFFSNINEKLQDKS
HTKYYNQYSLEFPSAMGLPFVYHLDVPTLASVSGKWHGDFEFVYSAQKSGLMGFVAPFSQNRYIAGLEKNIHAYVPVQAKVEYNPEENK

VQASLQPKAPRWYHTGYAFDLQC-
TPGTIQNDYISIKYNPQKSSPQEVNVALFISNEYGELLRNSHLLDASVTFGKNNGYSFTLAHANSPVSEKSRVFYYNAQPVSSFQAAFQLDI
KLPAFNYQDALNADSTSQVQGKINMGE--
QLESQITFKGKLERTQERKEYLQHPLARACESQMQRRNYIQPACLNATLQANFHNRYQFVFNYQNVPAGVSAYSMGRYNNEENQEGQI
QIGVQFEENLRSVNLSIMAVPFVYHPRYNTYQLLKSKLYATAVIDGNKATTFSNRTYPIDLGNCYHVFAMYAPIQQGYAVLVKETDSEHK
EVKVILGSDVIELKSKS-SLSIQANGQNIHASEKYARPDA-VLSLQGIELTYDGARVQLSASSSYKGNVRGLAGTFDGQEANDFVQR
>BAA88077_Pstali
TLVAFLTLASDNAWRPGKQYTYQVRGRTLTGIHQVRNQFSGVLLRALLQITVTETQLTLQLLDPEYADVNQHPEYPLNTPFQLLMKN
GVVDKMVVEKNLPAGDVNFYKAIASQLQVDTQGVNKGVYKTYEDSINGVYETIYEFYPTGEVYTVEKTRNFSSSDYHYRYGMNGHEGT
AGTNQMGNMLSRSSTSRIVITGAKYTIQYSDTIEKVLVSPHISNSQKGIAVSRVTLALQSVTPSPGTISQPRIIEKIVYEYPDNMVSVSSQKQ
VVEKMVNVAKRIGETILEPNRIDENILTSLARLVSGANSQQLQQAVDRMVYQQGWVAFRDAIANAGTGPALVTIRALIDSGKLKGEEAA
EVLATLPSAARYPTTEYMNYFFDLATSDTESQLYLNTSAAIAFSDLVRKSQVDRHRSVVAQPYYQQKLQDATDKGDSLKTQVYIRVLGN
MAHRQVLAVFEPYLEGKIPITRLLIVASLDEFTRVYPNTVRSVLYRIYNNQAEVPEIRVAAVMQLMKTNPPAQLLQHMAEQTNYDYSKH
VNAAVKSAIESAAR-GEYHYPEMLQNARSAANLLNPNYEAQYCRNFISSFIEKQQNMQH-
HHLSYIHALDGEYPSSIFYYLQKKVGGYLREPETVSVLTSSSQDLMSLLMQQLNDYTYQKIQDILQAKYDSQRQVEGNILLSLFGAKRFFT
FDNHTIEQWPKIALQLQQTLQKGVPTKYYNQDSVEFPMAMGLRFTYSYTAPSVQYLGVKGNGRVQYLQTNQRRAKIAFV--
FGQQRYVAGFERAAQVHIPLLAKLQYNGTQNTFNVSLQPETPVMFTTGYAFQLNYIFDDSVANNNYTVSYNPQESTSQETVIKVYN-
TQQGDLLSNARVLDASISFGKPNQYSITLGHSDSPVSEQSRYVYYRGTPLTTLEAAVEINVKVSVLNFEHALNAEPTSLIEGKINLGQ--
GSNTKITIQGKMERSPQRKQYVQHSLSTLCAQQMKQGHNLQYPCRNATLAAGYLDTYNLSINYQNVAPQVWLYSLARNNNVENPGG
VLSIGVRFTEDLTKANVSVVVQPLLQHPSYSQLHLISQQVFGTCTLDAVNTTTFNNKTYPVNLGDCYHVVAMYAPFQRGFAILVKQTD
SHKKVVKVVVGNEAVTLEPAGPGVSVTYNGAAIPLSEHYVQPGTTVLQFNSIQVVHDGQRVQIQIPNSYRNMVRGLCGTFDGEPVNDY
KEI
>BAA88076_Pstali

ALLTFVGLAADYGWKAGQQYTYKIEARTVTGLHQVADQYSGSLIQGILKVYASENQLTLQVQDAKYADINQRPEFPANQPFQLLLSNG
VVQKMVVDKDIPTWELNFLKGIASQFQVDTQGKNKGVFTTMEESTNGKYETIYEVSVDGEHYKVVKTRNFSNSDHHYRYGMSGHEGS
EGSNTMGNLLARTSESIIVLSGKRYTIQSSWTTEKVLVAPHLYNSQKGIVVSRLNLTLEDVSPLKGSVGNPRTVNTLVYEYPDDAASPNQ
KEQVVERMVRLARQMALSFNNPSEMKQDPLVQMSRTASIASYEQLQEATEKLSYQRAWQAYRDGVARAGTGPALVAMSSWIKSGKVH
GNEAAHLLSVLPRTARFPTSEYMDYFFELVKSKEQNQNYLNVSAILSFSDLLRRAHVDSHHVSVGEPYFQRQLENAINQGDSIKVQVYIR
ALGNTAHPKILSIFEPYLEGARKVTRFLMVAYLEEFIKVHPNTARAVLFRLYHNQGDIPEVRAAAAMQLMKTNPPPQILQHMAEQTNYD
HSKHVNAVVKSAIESAAN-
GEYSHPEFVENARAAANLLTTNYGLQYSQHFLKSFIGNDPDMYNEHYLTYIKGYDGAYPNSIRYNYQQKTGGYYNEPQTVSFMSSTSEG
FINHVFNQLQQYSYSKINDILNVEYDDPLQVEANFFLTVFGGARYYAFDNHTIEQLPEFTKYLSQKLRQGVESKYFSPHTVEFPMITGVQ
FTYSYEEPTVAYLAVQANGQYQAVWAQLKAGKLGFIAPCSEQRHLSIVKKRIQFYAPMEGQFEYYGHNNTFRANIRPRQNRTYETGYEF
QFKKQLGHGLYNNYYAISYNPQNSSPEHTQFSLYNSTIRGELVANASVFDAYVSFGNQDKYYFTAAHASSPVSEKSRFLYVSGRPYQSFQA
AAHVDFKVPIFNYKQAYEANPTSYFQGQVHFDL--
SGSSKIEFEGKMQQTEARKQYFKHAYSAVCEKQMQQGNYILPACHNASQFANVLDKYDVTFKYENVPLPVGVHDIARYYNTENPAGVL
KLSAKFAEKPFYVNYSVLFQPFVHSPAYNQLYLYSPHYYATCAIDNTEASTFNNRTYPIELGNCFHVLAMSIPNQDDIVIFAREPTPQKKE
VQVIFGYDVVSFEPTGASVTVKVNGESIPISASYALPGKNVAYFNSFQIYFDGHRAAITMPNSYRNKLRGLCGTFDGEPFNDFKEY
>BAA88075_Pstali
LLLAFACLAADYGWNTGYQYVYNFEARTVAGIHQVADQYSGNYIRGKLYAYVPDNKVTFQFQDAKYADFNEHHHFPINDVFQVYFNN
GLVDKLFVYKDIPHWELNFIKAIASQFQTDVQGQNKGIFKTWEDSVNGEYEVMNEVSVEGEFYELLKTRNFSSSNRHYRYGFNGHEGDA
GTNQMGNFLSRTSQSRIILSGKRFTIQSSVTTEKVLIAPHIYNSQKGMVGSRVNLTLDDYSNFKGTISNPKTVNNLVYEYPDQMVNFRSE
SHTVDNIVKVARQIGKFLEYPSNIKDNVLTYLPHLIAAANSKQLEQATEKLVYHRSWQVYRDAVAQAGTGPALVTIYSWLKSAKVRENE
GAQLLAVLPKAARYPTTEYMDYLFGMVKSKEQNQKYLNDSAVLSFADLVRRSQVDKHFAVVIEPYFQGQLRKAVQKGDSIQIQVYIRAL
ANMGHPTVWAGFEPYLEGKYPVTRLVMIAYLGDMIKVHPNTARSILYKIYQNEGDVPAIRVAAVMQLMKTNPPASLLQRMAEHTNYD
HSEHVNAAVKSAIESAAN-

GDFYYPELVRNARSAANLLTPNYGFQYSQNFIKSYIAKEHDLEYQHYFTYIKGLDGIYPSSVLYNLQRKIGGYVHEPQTMSFMSSSSEELL
NHFFGQLGDYTYQKIGEILNGQYDDPVQVEGNFFFSFLGSPRFFAFDNHTIEQIPQYIKHIHQKLQKGLETKYYNQYSVEFPLATGVPFIY
SYNEPTLFYADVEAEGTYKFTWSSQKWGKLGFIAPFNNQKYLAVVEKNTQFYVPFQGKLQFNGPEQYFQATIQPSSPQQFKTGYAFQF
KTFLVHSMANNYYAVSFDPRNSSPEKVVFRAYDNVEQGELFTNAQVFDASVTFGGQNQYAFTFAHAESPVSEKSRYAYFSASPYQSFKAA
FHLNFKVPVFNYKKALESDPTSYVEGQINFND--
NEKYKIAFQGKFEQTEERKNYVKSQYSAECLEQMQHDNYIQPVCLNASLNRNALDKYEFTFKFQNVPEYFKAYSFARYYNQEEEPGQIKF
GFNFSADFDYFNVSLHLRPFFQHPGYNQLYYFAPDYFATCGVDNGVVTTFNGQSYSIDYEDFTYVLVYALPYLGGFSVVAKDYGSNQKE
GRILLDDDKIEMRPSG-NVAVTANGEMVQVEEKFAMPGGGAFFFNGLEVYYDGKRLVVSVSNHHRDRVRGICGTMDGEPSYDFKPH
>EDAN008490_Edanica
FVCLLVCVAAEYGWKTGTEYRYQLKGRTLTALHQVSNQYSGILLKAIVTVQPNNDVISVKVSSPEYSPLQTVPELPISKPFEIKLKRGVVD
QLYISKNVPTWEVNILKSIASLFQVDTQGQNRTVYKTMEDTVTGECETLYDVSPSGEIIDVVKTKNFTNCKQRYHFGLTGMTDETGEN
RLGTFMARSSVSRVILTGTRYTIQSSVTTAKYIMSPVLYNNQKGMVVSRLNITLQAVRSSSGQVTNPRKISDLVYDYPYQGRAINYSSDPV
QAVKKLANQVGQELLQPSKIGYDTLGKIARLIRIMSTKQLEQVTRELS--
LAWKVFRDAVPVAGTGPALEIISEWISNRKIEKEEAAQVLAILPMTIRTPTPRLVDSYYKLVQKSEQHQQYLNVTALLSFTDLVRKSQVYN
DPPAITRPYLARQLKSAVRNADSHKIQVYIRALGNLAHPKILSVYEPYLEGKEDMSRLLIVISMDKLTRAYPVLARKVLYNIYQNTGEAHE
VRCAAVFQLMKTNPPVYMLQRMADFTNYDTSAQVNAAVKSSIESAASATDNWNIQLSKNAKAALHLLTPPYSLHYSRNFIKSYMVKEM
GLGYKQQVAYIGSTDSIYPNSLFYTVRRNLGGYKRQYLAFGAMVSSVDDLLDELYDQYDSLSPSKIADMLNIQPDDDDQIEGNFFINLG
GGKRFFTYENETIQQIPQAIHMAANAIRNGRFTKIYNQYQVSLPTPMGLPFWYSLKTPTLLSVNGEVTTEVHMVYSSKTAAKLSFVTPW
DHQRYVAGLNKYVQFNVPVKIQLGVDTAKSKVEVNVKPREPMKYDVGLSFRFDKWAERLIKRNNISVTFEGSRSQTRSVTMTFYDDSE
DSQFLKKGSVLDMAIEFNGPEYLTATVAYADSDVDEKSRLVYFYRSPM-
PLEMCMTTVTKVPILDFSKAMRSDPTSYLSTDLKFGEKCSQGAEVSLNLKMKQTDDRKDFVRHPVSKMCQRQMEEHNYLLPACQNAT
ISANLLDSYNLNLKFDRVPTNVKAYTALRYYLDENPKNKLQIKIEYNEDLRSKNITIWLRSVAQHPQFNPDELLGNELLPTCSIDVRAINT
FDNKTYPIRLGKCWHVAMQTTPDVAEIKILLG-----

DDDVVDLISTGSSQDDSSNERGKVILNGRELHISKTYATPNGVVLEAQDIDIIYDGERALLFADNDLRDDIRGLCGTFDGEHATDFKDL
>AAA27740_Agrandis
ALLLVGLSVSSPAWKDNTEYVYSVNGRTLTGLEETADQYSGVFLEAHLSIRP-
DGKLQGRISEPKFAQILSQPDLQLQKPFQLVLENGLIKRLIVEKDTLNWEANIIKSIVSQFQMDLQGENPAVFKTMEETVTGKTETIYDIH
REGDLIEVIKSKNYTNARDRYHYGFGEIEEEPTANKMGQFFIRQSNSRAILTGKRYIIQSTYTVNKIMVNPILKNKEMGSITSMVNVTLLEI
NNQQQELSNPLDIGNLVYTYPYHGKSIKENKDIRQNVENLVTEISDEIKQSEKTSKHTLDKLNTLVRLMDEDDIQFVAEQMNGQYTWSI
FRDSVAEAGTGPALLNIKKWIETKKIQKTEAAQVIGTLAQSTRFPTEEYMRKFFELATETQRQQETLNQTCILSYTNLVHKVYINRGREF
VKTPHLKQELEKAISNADNNKIHVMIRALGNIGHKSILNVFQPYFEGEKQVSRLMMVACMDRLADCYPHIARSVFYKIYQNTAELPEIRV
VAVHQLIRANPPVEMLQRMAQYTNTDSQEEVNAAVKSVIESSCKLESSKHAELRKAAQSARPLLTKQYGMEQSYINLRDYVAEQMGLEL
HVQRTSHSSAESSFPKIMKFQLHQHNHGMKQHILSTGGMISSIRELLNVLYRQTEEWSSANIARLMNYERDEREQLEAIIYAQVEDVQKL
WSFDNQTLEHLPEVIRQQEEIYRQGKYVKLKQLNEMAFPTEMGLPFLYTYDVPVLMKVEGKIEIKARVTCTGKTQSHLSFVTPFDHQIY
MAGYDKNMYVSIPVNARLEMDVKSKEAKIEFEVNNHRYFDTGLTFRWEQLDKCSTEYSEATISFIPSQSTTRKATFRIVDQKY--
EIKKHLSCSDISLEFEGDKEHVVGFAVAKSNADPKSRVFYYKNKNENQ---
ALEIRSETNGLNLDDSLDTEPSTKYNMRLQYGNSENDAFEISAQAQLSRSQERKQYLIQPLYHVCKEQMQQKNFQLPACQNMTIKANF
LDHIKYQVQYQKLNWKLGMFKGLRVPMTE-
GQNVVEGEVQFQPDFRQVNVTVLLRTLLPHPVFHAKCRFAGLMQPTCVIDQTTAQTFSNKTYSVNLDKEPTVVMQYVPSIENYVVLV
RVAANQ-
KEVIINLNHPKIEMKPSEDRAKVTIDGQEMHFDDKYALPNGEKLEVEAFYLIYDGQRVKVTAGNKLRDSVYGLCGRFSQDKHEDFQQW
>XP017781017_Nvespilloides
LLCLVGFAAADYGWKSNTEYLYKVQGRTLTGLPDVANQYAGIVMSANLIVRPGDGKLVAKISDAKYAQVHTQPELQLDKPFEIVMKNG
VVRDVIVSKEVHNWEANLLKSIISQIQLDTQGQHHAVFKTMEDTVTGKYETLYDITPDGEFIEIVKSKNFSRGEERYHFGLGANGKEPNT
NHMGEFFLRSSLSRAVISGNRYTIQSAVTTNKIILSPSLNSEQKGMVVSHLNLTLTHVQPASSKIQEPQQLGSLVYRYPNKGKSIKSAQNIV
EAAEKLAQQIGQEIQRPEKIKEQTLSKLTNLLRLMNEKEMQEVSRQIKGPDAWKSFRDAVATCGTGPALLTIKEWIKTKKIEQQEAAEVI

AVLAKSARHPTPEYMRTFYALVTDEDKSQQVLAETAILSFSKLVCRVYNDKGKQFVVQPYLKQKLNQAVEKADSRKVLVNIRALGYVAH
PKILAAFEPYLEGEKQVSRFNMILSLRKLVDVHPKIARSVLYKIYQNAGEEPVVRVAAVFLLMRTCPPASMLQRMAQYTNIDTSNQVNAA
VKSAIESATKLEGEQYAQLKENSEAAQPLLTKNFGIQYSRNFLRSFVNREMNLFYKQNIQIYGNEESSLPNGIKSMLRSELNGFNTEIFDIE
AMVSSIDELTNVLKQQTEQWSSEHIARLLNIQTEEREQLEGNFLFEFGIANRFMAFDNRTVELLPSVIRDIEQHLRDGKLCKFYDTKQVT
FPTEMSLPFVYTLHTPTLFHVKGQIETELHALFSTKVQTKIGFVAPFEHHSYVSGFDRNIQVNIPLKAQIEVDIKNKQMNAQLQPSEQTR
FQIGMAFRLEQLNDETIQYKKMEIEYAGAESTVRKVQFRMYESRYE-
QFAKMVRVVDASIEFDGQNEYTATAAIASSPVDQKARACYIRKQSANQFQFAVSAKAQTNGLDFAHALKFDPTSTVQIEVAAGEDLKSA
SKVHVQAKLRKSESRKQYLQLQRAHECRREMEEGNHQLPACANMTAEANLMDRVQVQLQFEHVSSEARAYDYLRYNLRENKQNEIEM
EARFEPNFESVNVTIWAKQVVVHPVFPMTNRVMGKLLPICVVDKTAANTFDNKTYPIDLGHAWTVMFQYVPQTEQYVVLVRTKSGQA
KEVRMTIRTPDLEIKPSENKAKVQVNGKQYEISDAFALPSGEKVTVRSFYIIYDGKRVKLAVSDKFRGSVRGLCGTFTGEQSDDFQQY
>XP017778650_Nvespilloides
LLCLVGLAAANHGWQNNKEYVYKVQGRTLTGLQDVANEFSGIVMKGNLVIRPSEGKLMAKIVDAKYSQVHSHPELPLDKPFEIQMQN
GVVRAIVVSKDVQTWESNIIKSFVSQIQLDVQGENHAVFKTMEETVTGKYETLYEISPDGEFIEVVKSKNFSRGEERYYYGLGSMGREPNT
NKMGDFFTRSANSRAVISGNRFTIQSSASTNKMILRPTLNSQEKGMVVSHLNLTLSRVQEASSQLHDGQKLNSLTYTFPYKGKNIKKSED
AVESAKKLAQQIGKELQKPEEIQEQSLAKLTSLNRIMTEKEMEQVAEQLQGPEAWFVFRDSIAQCGTGPALLTINQWIKTEKISYEEAAQI
ISTMAQSVRQPTEEYMRTFFRLATKEQKSQAILNDTALLSFTTLVRKVYVNKGKRFVEQPELKQQLQQAIQSADSHKILVFVRALGNVA
HESILAIFEPYLEGHKQVSRFHMVLALDKFVKVKLIKARSVLYKIYQNQGELEQIRVAAVYQLMRTCPPASMLQRMAQFTNVDTHEQVN
AAVQSAIRSAATLEGQDYKILADNARSAEPLLTKNYGAQYSKDNLHSFVQKEMSMFYIQNIQSFGSRDSIIPRGISYALRNTLGGAKHELIN
TQYMISNIDELINVCERLSQQWSSESISKLLNMKTSEREQLEASFYLQLSGATRFFTLNNRTVELIPRVLRDIEESLANGKYTKVWNSQEIA
FPTEMAFPFIYTLDAPTVIRIQGQVESDIHALVSTKIQARFGFITPFDHQQYSSGFDKNVQVNLPLKSKLNVDIKKQNINMELQPQPQTQ
FHTGMAFRYEQLNDESILYTDINVKYISNESSARKVQLRLF------
KFSKMICTCDMAVQFDGHNHYAATAAFAKSLVDDKSRVFFAKKESANQFEFAVAVKARTNGLNYQYALDFSPKVTAQFEFVFGKNLQS
SNQINVHAQYDKSESRRRFLKHPRSQECRREMEQENYQLPACANMTQEANLLDRVQVHVQFENMSQNMKLYNAIRYNLDENKENQIE

MEARFDADFQAVNVSFWVKKIVGHPVFTKKQRVMSEILPFCVVDKNAANTLDNKTYPINLGNSWTVMMQYVPQTEQHAVFVRDKSG
HGKEVKITLRNAEVDIQPQGQ-AKITINGKQIQYNIKYGLPNDEKIELENYYIIYDGERVKLTVDNKFRGSVRGLCGTFTGEKSDDFQEF
>XP017777145_Nvespilloides
AIAVLGYAGFESLIPDNKELVYHYHATIKAGTHVPAYYSSAMVLDGKLHVQNDRNSTFMQFSDLKYKMHNGKHELPLYKTFQVVHASG
IAVGLATEC-
EQEFVINFQKALGSVMQLDMQHLKKYAFQSEEYGLHGKTKQFYNVIPRKDFVRVEKMLETDSSRYMHEALTSNVIPMCDVKYSEPLTA
DSQRVYNIIKKDEHVMDKIEATGGIYYHPLEGESEAYHVFSNVSLHLLTVEPIKTDCGFKNFQHGFNYQTPTNGRRSWKQEDLVEKVYK
MLTEASQYLRKPVNHHGQLVNRIVRLMMHMDSTSFKKMHEKLSEEEVLQIYYNIVPLVGTAGAMHHIKELVKARQVKENVAIEMLEKM
GLSVKVPNKKLVDDMEDLMNLGDDCTLNTRKAALYCFANIVQKATYKECKHSLPEKYVNEIMKRLKATEDYNLRIHYMQVLGNMRM
DSVVEHLIPAVKGEMFE-
RLVAANAIMPMISRNPDLLFEVFLPIMSDTHLGMPLRIYAYYIIMESRPNLSRLINIYWQLQNERRNDLYQFHYNYLQKMSQTTDPCSEFF
AMRLNQILRYTEP-QFFGITGYNVMDYNNKLYNAGFKMDLVSIVSENTIFHK---FSTEYHFDRATVGDYEVMVRIEGLEKHLA--
DSLKIIDEFMKVIKDIQ------KDFHLNIVLSKNQFVMDYNYNCDNIKDISKFFSMW--------
EQTNLVKRLITMTSDIGFPIVFQKYIPSVKYFDFEIHVESTMKVWTYSRRGLSFYNPIADVWQGVNKYYALDVAFPTRFDLVFNPQQRSL
KLSWKRKAEDNI-IGLKSHVKD-
FETKLKNQNFEVVGLHIDDVNTGYKTQTFDRHMENEIMLLDLTIRLNSEFPGAKNWRATYVLKNKE--
KAVRTLKMQISRLIGFMICLDGSKK-----------WKNDELSLDVATSHSEETDVILKITGVGEKSEEQKS---KHQYKLCEEDVE-------
ECYEA---YTSLRKYTFNVKTHNIPSEYKIHNNYYKRPSKVEKGTTEVILEFPMGSTDSYITV---------
EGSRLLKSMYRLGHICLIDESSHENEHDEYVKYEMPTEWTMYIGDEETSCHHGVFTKRIEGTNIAVKVLHNSHVFEVYPKDGHYVFNID
GKEVSDQYSYLWAKESVVNITEFNMFYDGRHLEIVM-PKFNTKFYGKCFDKSH---------
>PHUM524850_Phumanus
ALVILAAVASDYGFKPGKEYHYEVTSTTLTSLHQVADRYAGVSMRNKLTLTPKEDEYLGTFTQSEVTEVHTTPLISLREPFVVKTKDGIV
DELVVSKDLPVWHVNVIRSIVSQFQIDTQGRRKAVYKVMEDTVEGECETLYDVTPKEDLIDIVKTANFSNCEQRYHFGLTGLTDKPTTN

KLGNFLSRSSVTRVVLSGNTYTVHSSVTTSKVIVSPEVYNNQKGEVISRVNVTLVSVGTGTSGPANPQTLKKLVYDFPYMGHSIQHSKNV
NQVVRKMAEQIGQTLQDQTHLKDNTLERLTEVVRTMNAKQLEQSVQEL---
NAWMVFRDVLAQAGTGPALTTIVSLITSKKLKGEEAAQVVSTLADTARYPTTEYLNTFFELVKHPEVSQPILRTTSVYTFTKLVRYALVN
PDDTTVTDPYLAQQLEQAVKEGDSTKVQTYVVGLGNVAHPKILQVFEPYLEGKQPMSRTLVVTSLGKLAEVHPKVARTVLYKVYTNT
GEVHQVRVAAVYNLMKTVPPVSMLQRMAQFTHEDPSTQVRSAVKSAIESAALLKTKTNFQLRKNAQAAVNMLVPVYGVQYSRLNLTD
YLVQGLHLGYTQTLQYVGSKDSLIPSSVFYKVLGNVNTYKTTLLKVSSMVSSVQAVTEFLQ----
AYTVEKITKLLNVEQLVKEQVEGNFNYRVGNLHRFFVYDNHTIETLTQTLNKYTQKTGQGLTQKFFTNYEITLPTTTGVPLVYTLSTPV
YIGIKGKVTTNVEVTVSTNTHGQISFIVPCEHKRFVSGYKKVVHVNVPLKGSVTVDLPKRNVKVEVKPKPLTVKTLGTVFRIHEVAEATI
DNTNVDVYYDSKLSTVQSLVLSAYEIKHETQFYEKAQVCDFVLNVNAQTELVGTFALARSPVDAPVNLGYFYKTAGN-
YEICYDSTFKSPIYDLEHAIQSESKSDLNVHVKFGNKCDTGAKVVLTGKLERTPERQQYVQLPQVQTCKTEVQQGNKYTANCRNATLQ
AGLFDKYTFNIKYEGVPEELRVVNIYTHYLTEGQSKEVEFDVNFTPDLKNVNLYLYTRTLLSHPVFNYVERYQWSAFAYCTIDKAQTKTF
DNQTYNTHLGNAWHVMSVSVP-VRDFSVLVREVENNQKEVQFTFNYAVVNVLP----
GTFTVNGQTVQVDEQYVLPTGEQFVSKGFNFLYDAQRVQLALSQSYRNRVGGLCGVYNGEHY-GYAEL
>PHUM524870_Phumanus
VLTLLVAGATEYGWKPNSEYRYEVTTKTFASLHQISNQYTGVLLNAQLKVVPTDSHLVVTLNQPRYSEVHTTPELTFTEPFEVKLNKGV
VKELIISRSVPVWEVNVIRSIVSQLQVDTQGVKKAVYKVMEDTVEGECEVLYDITPQGQIIDVMKSRNFSNCEHRYHFGLTGLTDEPTVN
QLGDFLARSSVSRVIISGSSYTIQSSVTTNKVILSPEVYNNQKGEIVSRVNLTLTSVHSATGTTQNPQSLKKLVYDYPMAGQSVATKPNIQ
QVAKKLAEEIGTQLQDGTQVKENTLARLTRVIRTMNAQQLEKTAQELTDPNAWMAFRDAVVQAGTGPALLTIVSLITTHKIKGEEAAQ
VVGSLANTARFPTTEYLQTFFDLVKHPETSQYNVYNSALFAFTNLVRMAQVNAHETMVTEPYLAKQLEQAVREGQSTKIQTFVRALGN
IGHPKIVEIFEPYLNGKTPVTRTLMVVSLNKLAVVHPELARSVLYKTYANTGEVYQVRVAAVYNLMKTVPPASMLQRMAQFTHEDPSN
QVRSAVKSAIASAAVLSEESDSELRKNAEAAVNLLTPTFGVQYSHHNLTDYIVKGLNLVYLQTVSNIGSEDSLIPQTIFFHLNRNLNGFKQ
QRVSAGGMVSSFKDLGDLLEEKIK-
FTVETITNLLNMEGGLTKQLEGNLLIGMTGFSRFFTFDNQTIQKVPTYLKKLTETLKNGRYQKLYNEYSVTLPTATGLPLVYTLNTPTL

LGAKGMFTTQMHITYTTRTQSKITFITPFNQKRYVAGYDKNTQFEVPIKKQFNLDLKNHHVKVEVAPHPLQKFQIGVVFRYQKFSTENI
EHTTLDLYYDGQKSTAKKVILTGYETTTSTELLQKAVVVDMGIQVEGQTEFLATVAVSTSPVDGQTRVGYGRKNVNS-
FEVCLDVASVVSVVDFTKILETDLKGDIYGKLLFGEQCQTGSKVSVVGELKRTQERKEYLEQPIVKQAKIEKSSGYNLQPSQTHASKEAGI
FDDYTFVVKYENVPETVKTYKALQYFLTENQEGQVELGVNFTPDLKALNLYILVTKGLIHPMYTTIERIANVAFASCTVDQTNFRTFDN
KSYSVNLEKVWHVMLFS--QTTDYVILVKNNADKQKVLKIVVKGMVLDVVP----
GTVKINGQEIPVSSQYGLPTGEKIVSVKFAILYYGETARITVNHDYMNQVKGLCGTNNGDNF-DFGEL
>AAA18221_Aaegypti
LLLALAGLAAFSSWMPNYEYVYNVTSKTMTALAELDDQWTGVFTRAYLVIRPSRDYVVAYVKQPEYAVFNERYHMPMSKPFGIRYHKG
AIKGLYVEKTIPNNEVNILKAWISQLQVDTRGANSGHYKVMEPLVTGECETHYDVNLDDQFIQVTKTQNFDRCDQRYHFGFTGYSDRP
NTNQMGNVASKSLVSYMYLTGNNFTIQSSSMINKVAIAPSLVNKEPALVYAQVNMTLNDVHPYDKPAEDLKVFVDLVYSYPYKGYNIFY
ARDAIRSVGKLVEEIASDLENPSNLKANTMSKLTRAIRAMGYEDIYELAQKYFSKDAWVTLRDAVAEAGTPSAFKLIFDFIKEKKLRGYEA
ATVIASLAQSIRYPTEHLLHEFFLLVTSDVLHQEYLNATALFAYSNFVNQAHVSNDYKIIEHPWFAHQLREAVNEGDSVKIQVYIRSLGNL
GHPQILSVFEPYLEGTIQITRLAIMVALDNLVIYYPSLARSVLYRAYQNTADVHEVRCAAVHLLMRTDPPADMLQRMAEFTHHDPSLYV
RAAVKSAIETAALADDYDDSKLALNAKAAINFLNPDVSIQYSFNHIRDYALENLELSYRLHYGEIASNDHRYPSGLFYHLRQNFGGFKK-
YTSFYYLVSSMEAFFDKYSKYYKKFTAPRIAKLLNIDAEEAEQLEGQLLFKLFNGYFFTAFDNQTIENLPHKMRHLFENLEDGYVTKFYQ
QQDVVWPLATGFPFIYTLKAPTVFKFEVDASVDVNLLYHRMVDAKVGFVTPFDHQRYIAGYQKKLHGYLPFNVELGLDFVKDEYEFEF
KFQSTKTMETGVALRWSLIAYQPYEYHQFNLFYDSQRTHAKSFKF--
YPSFEETQFFRNVRYYDFGFEFAPQYEFTFTTAFADSPVDKTSRQYYFYASPMFSFQFCATATSEVPYLKFSDFDKYYGDASQYFDFLYGE
SCQGGAHIAVKGKQKQTGKYREYLRSDVAKACKEQMANGYYQFEECQQAIDQAYYYDFYDYAIEYKDVGSVAKFYNYFQYYFESN--
NYIKAEFEFAPYGDYYNASFYMHRLPHKPELSNYEQFA----
PSCSFSNDYFYTFDNKKYFYDMGECWHAVMYTVKEYEQFAALARRGSDNQLYFKFLFGDNYIEVFPNNGGPFVKYNGRPYDISKSFAYP
NKDEVSFGKLKFATDGYRARFFSDYSFYNNFVGLCGTNNGEYFDEFYAY
>CAD56944_Amellifera

TLLLFAGTVAQHNWQVGNEYTYLVRSRTLTSLGDLSDVHTGILIKALLTVQADSNVLAAKVWNGQYARVQQSSDLPIGKPFQIRMKHG
LIRDLIVDRDVPTWEVNILKSIVGQLQVDTQGENNASFKAMEDSVGGKCEVLYDIAPDGRHMEVIKIKNFDNCDQRYHFGMTDNSREP
GTNKNGKFFSRSSTSRIVISESHFTIQSSVTTSKMMVSPRLYDRQNGLVLSRMNLTLAKMEKTSKMVDNPESTGNLVYIYPYKGKHIGKS
GDVINAAKELIFQIANELEDASNIVHATLEKLCNLMRTMNRKQISELESNMQVINTWTVFRDAITQTGTGPAFLTIKEWIERGTTKSMEA
ANIMSKLPKTVRTPTDSYIRSFFELLQNPKSNEQFLNTAATLSFCEMIHNAQVNKHDNSLYDPFLERELRKAHQEKDSPRIQTYIMALGM
IGEPKILSVFEPYLEGKQQMTRTLMVGSLGKLTETNPKLARSVLYKIYLNTMESHEVRCTAVFLLMKTNPPLSMLQRMAEFTKLDTNRQ
VNSAVKSTIQSLMKLKSPEWKDLAKKARSVNHLLTHEYDYELSRGYIDEKILENQNIITHMILNYVGSEDSVIPRILYLTWYSSNGDIKVPS
TKVLAMISSVKSFMELS---LR-SAAEKIAEELKIVPEELVPLEGNLMINNKYALKFFPFDKHILDKLPTLISNYIEAVKE-
GNVNMLDTYESVFPTETGLPFVYTFNVIKLTKTSGTVNSNLRLTFSKNVQGRVGFVTPFEHRHFISGIDSNLHVYAPLKISLDVNTPKGN
MQWKIWPLALPKNEFRLNVEIDTEMDNE-
RYVKVDMYMNLEGEQKDPVIFSTFDSKVMTEMMRESYVVDVRVHVPGESETVLTLAWSESNVESKGRLGFWRVEMPRN-
EVCIGSQIMETLLSYDEKMDQKPKMDFNVDIRYGKNCGKGERIDMNGKLRQSPRLKELVGTSIIKDCVEDMKRGNKILRTCQKAVVLS
MLLD--EVDISMEVPSDALQGLFSLSELDVSNGKKKIDVRAKLNEYLDKADVIVSTEDILTADEDLLINNVF--YE-
SCMLDKTRAQTFDGKDYPLRLGPCWHAVMTTYPKDKSVSVLSRENEAGQKEVKVLLGSDKIFVPGTTSQPEVFVNGEKIVVSRNYKMG
DRFGLTSKDVSLALDGERVMLKASEDYRYSVRGLCGNFDHDSTNDFLNV
>XP006561178_Amellifera
FLFLLYNNVTNSIFQSGKEYIYSYNALSSSGVLLPSGASSSWGFNGTLKIQAEENIATMQFESLKMTIWNG-
QDIPLLKPFQIIYKNGLIENFTTEA-
ISPWSVNIKRSIAGILQLDLSNLQKAAFHSTERNHYGQCNIEYVVNPQKNEWLIRKFFDPRTCIGHPHYTWSNVPNLCPNGDQNPILKSS
ERLYKIK-VNNDIL-FVNASGGIYVQPFQSFGEAHFHFTRQIFKLIMVKNTVNFTENLYFKV-
LQHELPTQGRGTLDKNTVFKSIGSLLDRLSQKLEETDNLHNTTITVLLYYLGMLDIVDLRNAYTKIKEETIRNMFLEALPQIGTKEAALFI
LELIQDKKVSDISAIQLLTQLPFHIRKPDVQLLVNLQIFLNLPEKISIEVQNTAILTYGTLIYKTCLLYCPYEMLDDYVRLYLDKFTETKEYE
KKMIWLEGLANIQLGRVVEFLEPIASGNNAESRVLAAWASLSTAPLRPDVIYPVYWPILINRTEHLEMRVAALTLLIVSSPTPSRLISLYWY

MQSEPSQHLYNYFYTILKSMEHTTHPCYVHIGMIAAQFTRVLHPKNQYLITGNYLFDYQDTYRKFGAMIHGIVIANPLTNIPEVLYVTVN
TYGSGININHVSLYIKAEGLLHSLS--THP--PTQVKDILKKFKLDQKQSGSVHLEIIARIQKTVLCLHLNETNIIKGFKYLSSL----
PDNIMEFHVNQQRIVQVTDLGANVRLAIIATSLFSMRGNFNNHVILRTSIHKSEIIENYNPFIDTWHSAERIHSIHGYLPINITLGF----
ENRPFISYNTPGEHLKMGITAHIKGLIETKLCPQLFKVTKSPTELAELGGQMKLFENIISRDYCTY---TQVKFEY-------
LSLTHRNAESLQILQQIKIKATKIIGLKFCLEVEKEIPWWDF---
LSIKPSEKIALNAVWGYSDTAGSSITLNLLGEISENQLENAKQWPYEECRIQSE-
GKRFIPACYEVSRELSTLRKYQIIAQYENIPENIAFYDLIGGNSTSDISKKFNVTAIFPKKSDTGELSLIIEYNYNLIRIHKYIDLSISKTGTCIVT
PDYIRSIHNVTYSFHLGHCY-------SENPKYAL---TARNDGININIYDEIDTVRIIPNQTGGTL-YNNTMYIPPQNMHWTQEQ------
ILLFFPTYLLEFS---------CGLCVLDFLDINNLYEKL
>XP001121939_Amellifera
VLGILTTVAVPAPFPHGKLVTYKYIADVKAG-
VDPSLSASNYGIECLLNVQHSKNTYYVNLTNVKYRLYNGFYEIPILNPFLVVYENGHLKGVKFVENEAGWSRNMKKGIVSMLQLDMTNI
RKHSFITHEETVHGTCQVAYNVHP-
NKEFVVTKIQESKNCTRFSYHETNHFDSKCHVDEWDDMTTASRRIFVVESQDEVLIKKLIGHGMINYLPWTAESEAHHLLTNQTLHLK
DVVSVSELSDVARLNPSIIENVPTHGRHIVKLDDLILKLRRMLDEVAGYLKKQSDWHDQTINRILHVMGYMDVASLERVYSEINEIAKRNI
FLGILSSVGTPAACLFTRNVVRTKSVPKWNAITMLGNLAMHVKVPSKKLLVQMEELLNLDDSVSVEVKEASIFCFATLIRKTFEHE---
PLLDKYLHHFMEHIRNEPTHHMKMVYMMAMKNVRLTQILKFLEPIIRGDEVVSRAQAIWIMRNVVFEHPRYCYNLLWPVLTDVILPTA
VRIAAFDVLMNETPQSERFINVYWLMVYEKNEHLYNYYVDTIKGLATSVDPCLMQARELAKKLLKIVRRNVTGPLSRKFYVDYVDNKYE
YNERVKGSLIVDHASGLPYIGSIEHVASVARKPVTKLGIHWNIDGLSEIVKDIENVFGNDNVKNVLISAAKIMPKIKDVNVNVILTMNNVV
SIFHCTKNNWQKILNELKQWKRLITEYM--
DVVNWQTVHVTTDLGVPAILSTKIPSLFSLVGSFKPKIKYQQWMHGEHTMSIYNPVADVWHSVHKTSSFDIVAPCEMTIGWNWKTKSL
KVTWPRLPVTDFSAGISIHVLN-QHDLLCHHYEVVTNNVNDSKDIGTRFQ-YDNDFIPDII--
TSVIDLIGKVRSKRDIRGNLNAKAASTNESIRQMRIMITRTIELKICIEAQNEYPETDDLLTV--

NGKKNTKMTITMGQTKEDEMDVTIIMKGEMLEEQKNQMSDSMYIECYNQSQHKSSKVPECIEEDILHSTLRKYTLDMTVRKVPQQIIV
HDKLRAHTWNHQPGNAKIVLEFPYLTSVLNAAV---------HYSTSILNDYFNERLCTIYPRLLITINNGVIQFIRSDEWILMSGD-
HIHHTYSIFVKLVQNKKLALRVYIGEHEMEIMPLEAKVSVKIDNKIIDDQKGMRLTTAYYVVISPVHIFYKVDSVTVSLHTDLQSRVTGIC
GNMDVMHKDEVPDV
>XP001122505_Amellifera
MLIILAARPSDSDWRRGPECTYDV--LVNMSLSNMDEDARICSVIAELKCRAGSDTLNCRFSNGRTARLEDG--APVEQPFEIRFNRGI-
ENLVVSRDIARWRLDAMRAIVSQLNVGFELGSGDRFVAMENSSVGYCEVEVKVSR--GSVRIEKVRRPKRCPNRYFFGNHR----------
----------FS----------DMITSVSRM----------
YSRREMNSSTGVMRTSNRPRTMNLHQRI----------
----------GLS----------
----------LRNI---------N----------PA------
----------TP----------
----------EIV----------
------NP----------STSLYAYT----------NLERIPE----------YK----------
>EFN64902_Cfloridanus
SLFLIVGVAVDYGWSKNKEYKYTVRSTTMTNLDDPHHEATGITMKGILTIQATPNTLYGTLRELQYAHKEEIQEVPIGKSFQINNKNGM
IQEMTVEREVPFWEVNILKSFMSQLQVDLSGQNKSSYKVMEDSVAGNCEVFYDILPDGNFIEIRKTKDYKNCKQRYIYNQTNYTKTPSD
TSV----SQISTTRMIISGNKFTTQSSVTHSTIFIKSDRKDRIIGNVRSMINLTLVEVNDKPV-AEAMVPTGNLMYMNPLSGKN--
SSEIAPEVAANLIMRVTKALEDPTVFSDEMIETLQDLLRTMNNDQLTEVERHADSTTVQKIIYKTIAQIGTGPSLIMFTNWIKNQAFTDV
DAAVLVARIPKTVRALTSEYLKHFHEMISTPEMDKDVFNVSAPIAFAELIRSAQERK--D--
LESYFAKQLKEAIEKNDRQRVHTYIVALSYTGHPKILSVFEPYLEGKQAITRTMMVTGLSSLARIYPALARSVLSKIYMNTREDDEIRVAAF
YSLIRTDPPLVTYMRMAQFTNFDRSPQVNSAVSTTIKSLASLKQARAQNIAIKARIARELLK-NSIAYYADTDLRSFHVETMG----------

RAKTDVNYMRVGINMVNDYLKLFNFRFGYGVSDIRQIMNVFNEYYNRSLIEELVNALNIKPLPVEPFEGYIFSDSKYETQFYPFDKNSFED
ISRTITNISRALNE--VNTMYNHDEV-FPTETGLPFVHTIQTPIVQKFIPK-
KTSIQMLFTSRVQNRWGFTVPFELQQYTAGVDRDFYTYLPAKVEVKKQRDQIQLQIGVDPQETSKRN---
KFQLNNIGQEDNRYKKIDVSVST---
AVKPVQLTIHKNNIDTQFLAENHVWDIDVGYTSQKHQVFTIGVGHSNADKRYHTFYWNIQPGTEYEICGNGRMAETPLDLQRMLDSE
QRDTFEYVLQYGKNYIDGRKVNIVGNTSQSSDLKNMLRNKIIKECLRDL-
KYNKESQACQKGSHLAQIKDRVNFSVATNSDDSIVNVFSHI--KIAKIRKNSIDVEVKMSQDRDQADVFVNIPS-
LLMPPLEAAQHLATGLLSICNVGKRQITTFDSKTYPMYGEKYWQVLV--IPEEERMSILSREKNGQ-
KELRFILNYKQIDIRRTNDGIQISIDGQPAKQHDRYELPDKSEIKSTGVDAIYNGETIAIRVADQYANEIRGLCGNYDSRPDNDFWIA
>EFN69845_Cfloridanus
VLLCLTAAAVPVHFPAQKTVIYNYHANVKTGIIEPAHYASQFVLVGQLHIKKLNNAYYVKLNNVKYGVHNGMDELEIQEPFVIVYDHGK
FQGIKILESESCWSKNIKQGIASMLQLDLTHIQQHNFVTHENTIHGTCQVAYDVHSINN-
FVVTKLHDPRNCSHFVQRVFDHVECKCHVESMDDMTTAARRVFQIEQQNDILIKHMVAHSVINYVPFNALSQAHYLLTNITLDLHSVV
PTMEMSTTNFQNVPMIRDIPTHGRHIVKLDVLIPKLKKMLVEAADYLEEEPEWHGQTINRLQYTMSYMDLGSLEQVYNTIKEVTMKNI
FLRMIPTVGTSAACYFTRNVIRKHKVTDAIAVAMLAKLSMHVKIPSERLLLEMEELLKLDNIVSSNVKKASILCFSILIRKTFMHQ---
PLLERYLNRFLDHVKNDQSYEMKMVYLMAMKNVQVGNIEKLLEPIIRGEIMVSRVQAIWAIKKAIADKVEYTHNLLWPILVDVTQPLPV
RIVTYDILMSQMPNMQRLMNIYWFMVHEQNNHLYNYHYTTLKGLANSVDPCLGPVREMARKVLRYTKRPETIELSSKHFVDYTDPVYE
HGESLNTAWVLDELTGLPHAGYMEHRVSVARRSVEKFGIYWNAYGLDEIIKIIEELMGDKHVEEIFIRAKHNIPVKKPVQIDICITLNHVIS
ANHYGQNTFRNMFTDISQLKQIM-------
VGNFQEIQVPIDMGLQGVFTTKMPYLFSLKFNNKAEIDTRIWRHGEYAMSIYNPIADVWHSIRRAIVQDVALPIEMNFSYNHEMKSLKIT
MPRLPVTKLSTGIRYHITE-EQDILCHHHTVVRTGEKDSKDIGLKYSVFENKITPDYI--MSQIDMTLRVSNKKNVRGTINAMG------
IRTVKIQMTRDTGLKICLDAQKSYPIKTDPLKLDIITKENTKMTITAGFTDDDEMLIVMTVKGEMTEEQKKQMKDSVHGFCIKDIQQSQ
GHIPNCIRETLQYTTMKKYTVNTSYKKVPVSVLIEDSIRVYITNRESGNAKVVIEFSSDHNDINATILFEGMTNKPRFSTASLYKMMKGEIC

TIYPKILIT-GKDVIPFVTSDKWTLVSGD-
YIDQTYAIFVKAVQDNALAVKIYIAGHEMNIVPNENHIVVTVDGNVVEQEKGIKLTENNHLVVSPIMMMWTPNSVTLMQDTSLQHRVT
GLCGNMD-IHKRTLSNI
>EFN61573_Cfloridanus
FILLLSNEAAQNLFRSGKEYIYSYNATSSTGVLVPSNAASSWNLNGKLVIQAEDDYVTVQLQSLKTNMWNG-
IDIYLLRPFRMSYNKGLIENFSIET-
EAIWATNIKRSIAGILQLDLSNLEKVAFHSSEINHYGKCNIDYVVSLWNNQRLVRKSIDPRICIGHPYRNWSNVPKQCTNEDQDPVLKNSE
RVYKLK-TDHDIL-SVNTSGSIYVQPFQSMGEAHYHIVQQAIELSSVKDITNVTNNLRTTI-
LQHELPTQGRDTPDKSFIFKSISRLLDRLSHRLEEIHNLHNTTISVLLYYLDMLDRDDLGKAYNRIKEETIRNMFLEALPQVGSKESAHFIL
NLIQNKKVSDISAIQLLMRLPLHLRRPDAQLLVSLQPFLTLPNEISAEVQNTAILAYGTLIYKTCLMHCPYEMLDDYVRLYLDKFTESTLY
EQKMVWLEGLANIQLGRVVEFLEPIASDNNIESRALAAWASIPTAPLRPDVIYRVYWPIVVNRTEHLEMRIAALTLLIVSSPTSNRLISLYW
YMQSEPNQHLYNYFYTMLKSMERTTYPCYKSIGRIAAQFSRVLRPNNEYLITGNYLVDYQDSSRRFGTMLHGAIIANPFTNIPEVIYITLN
NYGSGTHINHLSLYIKAEGVFHSLA--TSF--PTNIKDILKEFKLDERKRDPVHLEVIARIQKTVLCIHWNETKIVEGLKYLSSL----
WNDLMEFHINQQRIIQATDLGTNVRLAMMATSLFSMRGNFNNHVILRSSIYGIETIENYNPLVDLWHSAERVQSLHGYLPINITMGL----
EERPFISYNNLGEHLKTGITAHIKGAVKSKLCPISYT-VKSPSELPELGGRLKVFDTTMLYDYLTY---TQTKLEY-------
LSLTHHDLQSSQIVHQIKIKASKVVGLK----------
LPESLAFHDLIGGNSSNAKSNEVVIIATFPKESNVGELELAIEYNSRYIRIHKYMEMPLFRMSTCIVTSDSIKSAYNTVYTFALGQCY-------
DDNPRVAL---TATNSGININLTDESGIIRIVADMNKGTL-YNATSSIKESDIHWTQEQ------ILLSFPNHVLDFS---------
CGICTLDHSNVDRLYEKL
>EFN73216_Cfloridanus
----------MSLANVHENDNFCSIIAELKCRPGYDTLNCHFQNSRIKRPDPK--VPVEEPFEIRFNRGI-
ENLVVYRSIPRWRLDMIKVIVSQLNIGFEVQEHNRFIIMENSTMGHCEVDVKISR--ERFRVKKTRQLKKCPQKYFFGNHE----------
----------YS----------DMTTSASLI----------MSKEKFVSSEGVMKTANKSRIMRPHQNI--

----------SLK----------

-----LKRI----------DLAQSL----------
IP----------
----------EIQ----------NP----------STSLFAFS--------
----------NLKTIPKDIK----------NS----------
>AII96912_Fexsecta
ILLLIVGVAVEHAWDADNVYKYVIKSQTKTTMINKNQPAYGIAMKGVLVVHATPDKLQATVPKLQYVRVDE-
QQVPIGKPFEITLKHGVMRDVKVDKNVPVWEVNLLKSILSQLQVDTQGENRASYKVIEDSVGGNCEVHYYITPENNFIEIKKTKDYQRC
KQRYHYNQPNFTRSPNDKPI----
SHVSTNHMVISGNRFTVQSSVTTSNIFVKLGEKDSAIGSVDSKVNLTLVEKNKSKFSADQHVSTGNLVYMYPYNGKNIQTSDDAVQTAA
NLAIQITEGMENPNAQPGQTLERLQNLLRTMNTNQLAEVENAAKNIKVKNVLQKTIIQIGTGPALEFIKQVIKNKSIKGFDAAMSISRIPK
TVRTPTEEYIREFHKLISTPETNQNFLNVSAPLALAEMIRNVQYDK--
NSVLESYFAQQLQEAIRENNSPRIQTYILALCSTGHPKAISVAESYLEGNQGLTRLMMVNSLSPLARDYPKLVRPVVFKLYSDIHEPDDIRV
AAFYTLIKTNPPLPLMMHVARQTHYDKNYQVNSAVTTTINSLAKLNQWQARDMAVKARMVRNLLNYKSVGHYADSIVRRYYVETIG---
---------GNRYLPKYAHMKVDTKNDNFGSSTLDVEYALSSIQQLAKYY--
FKNGTLVEKIARALNMKPYHVEQLEGYLLSATPYETQLYPFDENVLQ---NLLTNWNNQENQ--MNLLYNRDKI-
FPTESGLPFTYTVKKPILSQVEAR-
KSAFNILFASRSQNRFGFVTPFEHQQYIAGIDRDFHVQLPLEFKGEQQEDQTNLELKLIPQEGASQEQGVSFE-
NNFEEEDVHYRKIDVFVDSN-
AANIGIRLNVHQSNTDNQFLTENHVWDIDVEYTLQKHAVLTFGVGHNKVDERYHIFYSNVQLAQGYEALGTGTVKRTPLNFVRTINRE
PQNNFEFLVQYGKNLKNAEKVHVHANMVQSNNLKHMIKYEIVKQCENEITQGNKGLYACQKATQLAHMHDQMKLSMRVESERDRV
NVFGDM--GIVDSGKDTVDMEARTLPDRDEVDVSVDTQSQLLSPLQELLSRLSSSTKATCGIDKNQIVTFDGKSYPLQATEDWQVVL--

IPEYMKMAILTRERDGS-
KQIRFVLNEKEVQVQKLNDDIQVLIDESQVKLQDKYELPDKSRIQSKGINAVYDGERIQIQVEDRYRHGIHGLCGTYDSQSDNDFYTA
>AII96913_Fexsecta
VLLCLAAAAVPVHFPAEKTVIYNYYADVKAGIIQPTQYASQFALVGQLHIKKLNNAYFVKLNNIKYGMYNGRHELGIQEPFVIVYDHGKF
QGIKMLESESVWSKNMKQGIASMLQLDLAHIQKHSFITHENTIHGTCQVAYDVHSTTN-
FVVTKLHDLKNCSHFVQRVFDHVECKCHVESMDDMTTDARRVFQVEQQDDILIKHMVAHSVINYFPFNAQSEAHYLLTNITLDLHSV
MPSMEVSSMNFQNVPIIRDIPTHGRHIVKLDTLIPKLKKMLVEAADYLEKEPEWHGQTINRLQHTMSYMDLGSLEQVYNAIKEVMMKN
IFLRMVPTVGTTAACHFTRNVIRKHKVTDATALAMLIKLPMHVKISSERLLLEMEELLKLDNTVSSDVRKASILCFSILIRKTFMHQ---
PLLERYLYRFLDHIKNDQSYKMKMVYLMAMKNVQVGNIEKLLEPIIRGEVMISRIEAIWAIKKAVANKVEYTHNLLWPILADVTQPLPIRL
VTYDVLISQLPNMQRLMNIYWLMVNEYDDHMYNYHYTTLKGLANSVDPCLTPVREMVRKILRYTKRPVTTELSSKHFVDYIDPVYKHG
ESFNTAWILDQLTGLPYAGYIEHRVSVARKPVEKIGIYWIAYGLDDIMKIVKTLMGNKNVQTILIRAAQDMPVKKPIQVDICITLNHTIVV
KHYDQNTFKNVFDDTFQMKQTM-------
VSNFQEIQVPIDMGLQGVFTTKMPQLMSLKFNNKAEVDARIWRHGEYVMSIYNPIADVWHSVRRATVQDVAVPIVMNFGYNHETKSLK
VTMPRLPITKLSTGIRYYITE-EQDILCHHHTVASTGEKDSKDMGLKYSIFENKVTPDYI--
TSHVDLNLRVNSKKNVRGTLDAKAASTSESVRTVKIQMTRETGLKICVDVQKNYPVETDPLKLD-
TTKENTKMAITAGFTDDDEMLIMMTVKGEMTEEQKKQMTDSVHGSCIKDIQQQQGHIPNCIRETVQYTTMRKYTINTLYKKVPVSVL
IEDSIRAYITDRESGNVKVVIEFPSDLNHVNATVLFEGMTDKPRFSVTSLHKMMQGKICTVYPQTLVTLDEDVIPFV-SNTWTLVSGD-
YVDQTYAVFVKPVQSNTLAIKMYVADHEMMIVPNEYRAVVTVDGNVVEQEKGIRLTENNHLVVSPIMVMWTPNSVTLMLDTVLQGR
VSGVCGHMDTMHEEKLPKI
>AII96914_Fexsecta

----------TDHDIL-FINASGSIYVQPFQSMGEAHYHFVQQTIELASVKDITNVTNNLRTTI-
LQHELPTQGRGTLDKSFIFKSISRLLDRLSHRLEEIHNLHNTTISVLLYYLDMLDRVDLGKAYNRIKEETIRNMFLEVLPQVGSKESALFIL

DLIQSNKVSDISAIQLLMRLPLHLRRPDAQLLVSLQSLLTLPSKISAEVQNTAILTYGTLIYKTCLVHCPYEMLDDYVRLYLDKFTESTRYE
RKMVWLEGLANIQLGRVVEFLEPIASGNNAESRALAAWASIPTAPLRPDVIYPVYWPILVNRTEHLEMRIAALTLLVVSSPTPNRLISLYWY
MQSEPNQHLYNYFYTMLKSMERTTYPCYKHIGRIAAQFSRVLRPNSKYLITGNYLVDYQDSSRRFGAILQGIIIANPSTNIPEVIYVTLNNY
GSGTHINHLSLYIKAEGVFHSLA--TSF--PTNIKDILKEFKLDEQKKNSVHLEIIARIQKTVLCVHWNETKIVEGLKYLSSL----
WNDLMEFHVNQQRIIQATDLGTNVRLAMTATSLFSMRGNFNNHVILRTSVHGIETIENYNPLVDLWHSAERVQSLHGYLPINITIGL----
EERPFISYNALGEHLKTGITAHIRGAVKSKLCPVSYTVLKSSSELPELGGRLNIFENTMLYDYLTY---SQTKLEY-------
LSLTHHNLQSSQIVHQVKIKASKVVGFKFCVEIERHMPWWEF---
LSNEPSDRIKLNIVWGLSDSVGSSISINLIGEISSEQLEESKNWPYGECKKESI-GKKFVPSCYEASRELSTLRKYTISAHYENV----------
KED----------

>AII96915_Fexsecta
MLTIL-----QSKWTMGPECTYDV--LVNMSLANVHENDNFCSMIAELKCRPGHDTLNCRFQNSRIKRPDPK--VPVEEPFEIRFNKGI-
ENLVVYRTIPRWRLDMIKVIVSQLNIGFEMQERNRFTIMENSTVGHCEVDVKITR--ERLQVEKTRQPKKCPRRYFFGNHE----------
----------YS----------DMTTSASLI----------TSKDKFVSSEGIMKTVNKSRIMRPYQNI--
----------SLK----------

-----LKRI----------DFAQSL----------
MP----------
----------EIP----------NP----------STSLFAFS-------
----------NLEVISEDVK----------NS----Y----------
>EFN86099_Hsaltator
VYCAIANAVMLLIWKPNHTYHYRVETQGLAGMNKVSQQFAGLQLRGELVIEVEPNVLLVHLQNVQHTQILKESDSWIQTTYKVDFQH
GNMQKIYTEPQSPRWQLNVMKGILSQLQLSIV----SEFIVDEQSVSGHCNTSYNIQPDEMYWDVMKTRNYRDCLSMYQFSIVK---

NMDHDIHERYLKRSTQVNMVISRDLTTIRSSVMKGTIDVRPKLNDFANSNILTKMKLTLTKLEPMTEPPLNLVPAGNLLYVH-
LNKNDGSKMLEMMKQVVILVNEMVSEVQSSDEVRQHTLEKLTSLLRHLNEKELATVENDLSKELRWKIWRDAVVQAGTGPAFLTIRG
WILEKKVRDFEAAQLLSRLPRVVHMPTAEYIKAFFELLKEPSKNEFLLNTTAPYAFAELVRYVMVWDEKKALTDPYLAQELMDAQQKHH
NSKILTYIIALGNIGHPAIVPVLEPYLEGKLKITRMTIVASLRKLMEHQPDVVRPIVYRIYQNDKEHQEVRIAAVQLLMMSHPPLDMLKRM
AERTHYEPNGEVNNAVVTGIRYLSKLSRERNRQDAHNAQAVRKLLNKKPNMRESYIFGQDYVNEQKEFEAAVYVNTIRGFKRFSPASAS
VELGTNRSHLSPPNLKIAVASTGSELLLRLLQLLFH---
VDAVAKNLQIPLQDEELLETSFVMDSEYAVNLIVLDKHEIFDLLESFLRVFKLAEE--
KSDLMLASMMAYPSELGLPVTYRLQKPQLGYMRSKLDVELDVLLSRKVMSRLDFVTVFDNQRYLSGFDVNMQLHLPAKLKLHVDEAA
RQLNVVYEPQPPSKTVITIHFRVEQIIDKDVHYNKIEVGTKFDPNTDNTISFTIYDDILPPKLVR----MDMKIPNHSEHHYKAILQTSHS----
----------WKFAAVLDLECPTLDLHSKHHN----
IWNALLHYESPDKTKQTIHLLAIQEQSEELNNFLVSEMHKKCMMEKERHTVGLKNCQEFCKFIDQKDEMHMTMKVGN-----
KPFHYMKYALAAHPIMHLELEPQPQESHRSISVMVYMHMDLDLP----------
RKCVLSADHLVSFNGLEYSVRLGQKPYVLLINPTDETNVAVLVQQLDNG-
RRSLIIMKKETVEILPPAPGVRILVDGQNLESPLESTLPSNVEILLSEITIWYDSTNVQIKASDKYNGELRGLCGAKANIREHQLEKA
>EFN75537_Hsaltator
ALLCLTAAAVPVHFPIEKTLIYKYHADVKAGIIEPAQYASQFALSGQLHIKKLNNAYYVKLQDVKYGMHNGMHQLDIHDPFLIVYENGKF
QGIKVPETESGWSRNIKQGIASMLQLDLSQVQKHAFVTTENTIHGTCRVAYDVHP--
NVFVVTKLHDLKNCSHFVQRVFDHDECKCHVESVDDMSAAVRRVFQMEQQDGILIKHLAAHSVINYLPYNARSEAHQLLTNTTLSLEN
VVPTTELPVVNFQSVPLVRDVPTHGRHIVKLDMLIPKLKKMLAEAADYLDKEPDWHGQTINRLQHTMSYMDLGSLEQVFNTIKEVTIK
NIFLQMIPTAGTTAACHFTRNVIRKHKVTDVTAVKMLAKLPMYVKVPSERLLLEMEDLLKLDNIVSPQIHKASILCFSTLIHKTFMHQ---
PLLSRYLQRFFDHVKNDQSYEMKIVYLMAIKNVQVGNIEKLLEPIIRGDIMVSRVQAIWAIKKAVADKVEYTHNLLSPILADITQPLTIRIA
AYDVLISQLPNMQRMMNMHWLMSYEQNSHLYNYHLTTMKGLANSVDPCLRPIREMARKIMRYTKRPITTPLSAKHFYDYVDPVYEHG
KSWTNALVLEEVTGLPQAGYVELRTSVARKPVEKIGIYWSAQGVDEIMKILKELLGNRNVQNILNRAAHDMPVKKNIHVDICITLHHVII

ANHYDRTHISKLPEELLHAVKQA-------
VGNLQYIQVPTDMGLQAVFSTKMPHLWSLRFNNKFEVDSRLWRHGEYVMSVYNPIVDVWHSIRKSTTQDVALPLEMSISYNHEAKSIKI
SMPRLPATKLSSGMRYHITE-EQDVLCQHHTVVTTGEKDSRDIGLKYSIFESGVTPDYI--
TSHVELNLRVNGKKNIRGTLDAKAASTNASVRSVKIQVTRVTGLKICVDAQKNYPDQIPYPKVS-
NSKVNTKMTISAGYTDEDEMLIIMTMKGEMTEEQKNQLADNVHGACVKDMTVKENHIPNCIHETVLYTTMRKYTINAFYKKLPVGIMI
DDHFRAYLTERESGNAKIVIEFPSDMNYINATVLFEGVLEKPHFSAPVLNKVTQEKICTVYPQVLITLDNGLIPFTVPDQWTLVSGD-
HVDQTYAVFIKNVQPNTIAVKVYVAGHEIEIIPSSPSHIVTVDGNMIDQEKGFKLTVHNHLVVSPVMVLWTPNSVSVMVETVLQGHVTG
VCGHMDDTHKERMPKI
>EFN87799_Hsaltator
LVLLLNAGVVQGLFRSGKEYVYSYNVVSSTGVLVPSNAAASWSLKGKLVIQAEDNVVTAQLQSLKTSMWNG-
IDIILLKPFQLSYNKGLIENFSIEA-EPAWATNIKRSIAGIFQLDLFNLVK-
TFHSSEINHYGKCNVDYIVSPGENQRLIRKSLDPRMCIGHPNRHWSNVPKQCTDEDQNPVLKSSERIYRLRFTDRDII-
FVNTSGGIYVQPFQSMGEAHYHFVEQTMELLSVEDTTNVANNLRTTV-
LQHELPTQGRGTPDKSFVFQSINRLLDRLSQRLEEILNLHNTTISVLLYYLGMLDRVDLGEAYDRLKEETIRNMFLEALPQVGSKESALFI
LDLIQNNKVSDISAIQLLMQLPFHIRKPDVSLLASLQPFLFMSNKIPREVQNTAILAYGTLIYKTCLAHCSHKILNDYVRLYVDKFIESTQY
EQKMIWLEGLANIQLGRFVEFLEPIVNGINNESRALAAWASLPIAPLRPDIIYPVYWPILVNRTEHLEMRVAALTLLVVSNPTPNRLISLYW
YMQSESCQHLYNYFYTMLKSMERSTFPCYKPIGRIAGQFGRLLRPNAKYLVTGNYLVDYQDSSRRFGAVLHGIIIANPSTNIPEVIYVTLN
NYGSGTHINHISLYVKAQGIFHSLA--TSF--PTNIRDILKEFKLDERKKGPAHLEVIIRIQKTVLCIHWNETSIVSGLKYLSSL----
SDNIMEFHVNQQRILQMTDFGTNVRLAMMTTTFFSMRGNFNNHVILRTSVHAIETIENYNPLVDAWHSAQRIQSFHGYLPVNIIFGF----
EERLFISYNASGENLKTGITAHIRGAVKSKLCPITYTVAKTPIKLPEIGGQMKMFESIVSYDYLTY---SETKLEY-------
VSLTHRDLESSQILHQVKIKAYKIVDSKFCVEAERRTPWWDF---
LSNEPSDRIKLNIAWGPSNTTGSSIIMDLIGEVSSEQLEESREGLYGECKKESA-
GKTFIPSCYEVSRELSTLRRYKIVARYENLPESLALYDLIGGNSSNANSDELVITAVFPKESNVGELELAIEYNSRYVRIHQYTEISLFKMSTCI

ITTDSIKTAYDMTYNFALGQCY-------DENPRLAL---TATNSGIDVNLTDELEVTRITADEDGGTL-YNATSFVQQSDMHWRQEQ------
ILLFFPNHMLDFS---------CGICTLNSPNIDRLYEKL
>EFN86751_Hsaltator
ETVLLTAVSTQSEWTAGAECTYDV--LVMMSLENVRPDDNFCSTIAELKCRPGQDTLSCRFQNSRIRHPRPE--VPVEEPFEIRFNRGI-
ENLVVHRTIPRWQLDMIKVIVSQLNIGFESWEHDRFAIMENSTMGHCEVDVKITR--GTFHVQKLRQPKRCPRRYFFGNEE----------
----------YS----------DMTDSVSHV----------
KNKHKFVSTAGVMKTANKSRVMRPQQVI----------
----------SLK----------
----------LKRI----------DLARHS----------
----------LP----------
----------EIQ----------
--------NP----------STSLFAFS----------HLETIPEDA----------

>XP025073621_1_Pbarbatus
TLLILAGVAADCAWNSQTEYRYLVYSRTVAAIGQIKGQYTGLQVRGTLVVQASTNTLRARLENPQYANLHAPRDLPMGKPFEIKLKHGL
IQNVLMDNDVPIWEVNMIKAIISQLQVDTKGENKGTFKVMEDSVGGKCEVLYDVSPDGHFYEVKKTKNYKKCQQRYIYGLNDPTKSQL
DRQD-
DVVSELSSSEVVISGHEYTIQSSVTRNQIYVKPRRTEGVLGTVHCIMKLTLAKKSPITTLSDNLMSTGNLVYTFPFGGNNVQTSEN-
MESVMKIIMEISNEIQQPERIDNQTLEKLKNILRTMSYKQYSELNNRISNKDTWNIYRDAVAHAGTGPALLTITDWIKNKKIENLEAAQVL
AQIPKHVRTPTAEYIKTVFELIKHPVAQQKHVSSSAIVAFAELLRN--VN-
DNKDVEEPYMANQLKHAISRSDSPRIQSYIVALGVTSHPKIISVLEPYLEGTQPASRMLMVSSLTTLIQTQPKLVGPILYKIYMNTREAHEL
RCIAVHNFIKSDPPMLMLQRVAKFTNYETNEHVNSVVKTTLESLANTKRPGWEKLARKVRIALRQINPIATTWNSMGYINSWIINGFSM--
----QTVVSEDSYLPSFVHCALNTIFDRYSQPSLEEKYAVSSVKQLLNHLQNQGT-

TRVDKLVKTLKMQPAHLQKLEGSIFINSLIGFLFYPFDDEALDNVVKTMNEFMEKQQQ--
DVTYLRSYEATLPTENGLPFSYTVDSPTLVKIHASFTGLLNALIANKVQRRIGFATLFNHKTFVVGIDTNRHMQIPIKYEIETD---
RNIGVTIIPSRKDKTVLSLELELPKLPWEEMHYMKLNMKVVPQQWDSQEINVNIYDELEITQIMRSAYVLDISCNVPMLEQQVLTIGLVN
SNPERMSKAVYWNNRTPGEYEVCYLQELQSDTLNLMRALETTPQSHFKGNLYFGQTCQEGNQVEIIGLMQRSNELKEIIKTDVVKECQ
RGIEEGNQKMPACQRAVELANIED--QLDISITA-PSVVRVIEYVLDSTARTDIDHLEIKIILPSNDVMPKISVP-
HPSETPEIRKMRNEFPLAGEDMCTVDRNMVVTFDQVAYPVKMGYCKHVLMITNPEEQKVAVLAYET-
RESLNTILLLGNQEIKLLKKEGHVEVLIDERPIELSEHLKLADESEVFSKKLRALYDGNHIQLQASNKYRNAVRGLCGNYDMQSDNDF---
>XP025073621_2_Pbarbatus
TFLLIIGAVVQNGWETENEYHFRVQSRTLTALNKLAQQFSGILINGALIVQVSPDTLQAIISKTQYAPVHQVTDLSLGKSFEIKLKHGVIR
DVLVDRDVPTWEVNLLKSIVSQLQIDTQGENRGSFRAMEDSVGGKCEVLYDITPDGYHFDIIKTKNYTRCEQRYHFDIAGKANKPGSN
DG--
VLSRSSTSRILISGNRFTIQSSVTTDKILITSKIHDTYSGAVYSKVNVTLDHVQKISNASDNSVSTGNLVYIYPYKGKSIKESQNYSQVAVRLI
GQITGEIEDILHQFTETVERLNRLIRTMDTKQLAEVENMLDEYVSWNVFSSAVAHAGTGPALITIRNWITNGKLKDMRAVQAISKIPKVA
QAPTTEYVKAVFELIIDEQTKQKLLNTTAPLAFAELASNVYNNRHDNALLEPYMATQLKEAIKDGDSRRIQTYIMALGNFGHAKVFSIFE
PYLEGTLPLSRMMMVISFGKLAENFPRVTRSIAYKIYINVMETYEVRCAAVYIIMKTNPPLAMLQRMAEFTNQDASKQVNAAVKSSIEALA
NLKQPELQDLANKARIAKTLLNPNYSSEYSQTVIQEMIVASLNFAQREIFQIIGSNDGSWFKSVLLGEDLSYGGFNLPPTKLIYTISSLIEAI
QSMKDTFK-
VAIEEMIAQLNIKQDDLEHFEGSILMNTLFSSFFYPFDNNTIEEIMRVLKTDMISWREPGNINYLNNYEVMFPTESGLPFIYTLSLANLMRI
NGDATVAGHLLYKERIQSRIGVVTPFEHRHYIAGIDTNTQFFIPCSARVAM-IEAKRFELKIHPRIMHTAEHTINVAWNAFSGGA-
RYRNFSSIIFIR-----PAEVNILHAMIINQFLREAYIFDLGF---
ASEKQVYTLATSQSNVDEKYQAFYWNIQTPEGAEVCGAGQLQSNPLNFEKAIEQIPKDEFSAELRFG-
SCTNGERIKLKGNLTRTDRAKETAMSEVIAKCRQEMAQGNKGLLACQEAIELLEERD--
RLMVSMEMKSNRLNAIFSIKTVKMLSKENTIDMEINVPSKSNNGKISLSLADAVNIPEQKQKENFL------

CTLDKTKAATFDDNVYPLKLGKCWHVMMTTFPENSQVLIKVREMDDGSKEVKMIFGCKEIYLRKRDDRLEVTVDGKTEEFSQHYELH
DGSNVYSKKINAIYDGERIQVEAHRKYRNSIRGLCGNYDLRSDNDF---
>XP011644827_Pbarbatus_likeA
VFLCLAAAAVPVHFPIEKTVIYNYYADVKAGIIEPAQYISQFALNGQLHIKKLNNAYYVKLNNVKYGMYNGMHQLGIQEPFVIVYERGTF
QGIKTGTTESGWSKNIKQGIASMLQLDLAHIQKHSFVTHENTIHGTCQVAYDVHPNTNVFVVTKVHDLKNCSHFVQRVFDHVECKCH
VESVDDITTSARRVFEIKQQNNILIERLIADSVINYFPYNGRSEAHYLLTNTTLDLHSVVPMTEVLAVNVQNVPIIRDIPTHGRHIVKLDAL
MPKLKKMLAEAADYLDKEPDWHGQTINRLQHTMSYMDLGSLEQVYNALKEITMKNIFLRMIPTVGTTAACHFTRNVIRKHKVSDATAI
AMLVKLPMHVKVPSERLLLEMEDLLKFDNTVSPQVRKTSIFCFSILIRKVFIHQKRYPLLERYLHRFFDHVKNEQSYEMKIVYLMAIRNVQ
VGNIEKLLEPIIRGDVTISRVQAIWAIKKAVANKVEYTHNLLWPIFADVTQPLPIRIVAYDVLMSQMPNMQRLMNIHWLMVYEPNNHLY
NYHITTLKGLANSVDPCLAPIREMATKVLRFTKRPVTSQLASKQFIDYTDPIYEHGESWNTAWVLDETTGTPHAGFVEYHVSVARKPVE
KIGIYWSAYGMDEIMKVIKELVGNKNVQNILILAAQDMPVKRPIHVDICITLNQVIVANHYDQNTFKNVFDDTFQLKQTM-------
VGNFQEIQVPTDMGLQGVFSTKMPQLWSLKFNNNLNVDARIWRHGEYVMSIYNPIVDVWHSIRRATVQDVALPVEMSVGYNHEAKSL
KITMPRLPVTKVSTGIRSYITE-EQDILCHHHTVVTTGEKDSKDMGLKYSVFENRITPEYI--
TSHIDLNVRVSSHKNVRGTLDAKAASTNESVRSLKVQITRQTGLKICVDAQKNYPVETDPLKVD-
TTKENMKMTVTSGFTDDDDMMVIMTVKGELTEEQKKHMMDNVHGTCVKDTQQKQGHIANCLRETIYYTTMKKYTINTTYKKMPV
QVRIEDHIRAYITDRENGNVKIMIEFSSNLSHVNATVLLKGMASEPRFSLASLHKITQGKICTVYPQVLITLDEGVIPLNILDKWTLVSGD-
HIDRTYAIFVKSVQVNKLAVKMYIAGHEVEIIPNELNAAVTVDGNVIDKEKGFKLTQNNHLVISPVMVMWTSSSVTLMLDTTLQGHVT
GVCGHMDGIHKEKLPKV
>XP011629719_Pbarbatus_likeB
FILLLSTEVAQSLFRTGKEYIYFYNAISSTGVLVPSNAASSWSLKGKLIIQAEDDYVTLQLQSLRTNVHNG-
IDIYLLKPFRLSYNKGLIENFSTEA-
EPVWATNIKRSIAGILQLDLPNLEKVAFHSSEINHYGKCDIDYVVSPGENQRLIRKSLDPRMCTGHPSRDWSNVPKQCTDEDQNLILKSS
ERIYKLK-TDRDII-SVNASGSIYVQPFQSMGEAHYHFVQQTIELSSVRDIINTTNNSRTTV-

LQHELPTQGRSTMDKSFVFKSISRLLDRLSHRLEEIHNLHNTTISVLLYYLGMLDRVDLGKAYTRIKEETIRNMFLEALPQVGSKESALFIL
DLIQNNKVSDISAIQLLMRLPFHIRKPDVQLLVNLQPLLTLPSKTCAEVQNTAILTYGTLIYKTCLVYCPYEMLDDYVRLYLDKFTETTRY
EQKMVWLEGLANIQLGRVVEFLEPIASGNNAESRVLAAWASIPTAPLRPDVIYPVYWPILLNRTEHLEMRVAALTLLIVSSPTSNRLTSLY
WYMQDEPNQHLYNYFYTMLKSMERTTYPCYKRISKIAAQFSRILRPNSEYLTTGNYLIDYQDSLRRFGAMLHGIMIANPSTNIPEVIYVTL
NNYGSGVHINHVSLYIKAEGIFHSLA--ASF--PTNIKDILKEFKLDERKKGPMHLEIITRIQKTVLCVHWNETNIIEGIKYLSSL----
SDNIMEFHVNQQRIMQVSDFGTNVRLAMTGTSLFSMRGNFNNHIILRTSVHGIEIIENYNPLIDLWHSAERVQSLHGYLPINITLGL----
EERPFISYDNLGEHLKTGITVHIRGAVKSKLCSVSYTVVKLPSELPELGGRLKIFKNTVSQDYFTY---SQTRLEY-------
LSLTHHDLQSSQVLYQIKLKATKIVGFKFCVEAERHVPWWEI---
LSVEPSDRIKLNIVWGPSDTAGSSISINLIGDISSEQLEESKNWPYGECKKESM-
EKRVIPSCYEASKELSTLREYKIFAKYENLPQSLALYDLIGGNSSIAKSDELAITAIFPKDSNVGELQLAIEYNSRYIRIHKYLEIPLFRMSTCIIT
LDSIKSAYDTIYTFALGQCY-------EENPRLAL---TASNDGINVNLTDESGTTRITADKDRGIL-YNATSYIQQPDMHWTQEQ------
VLLFFPNHVLEFS---------CGICTLNHPNIDRLYEKF
>XP011631341_Pbarbatus_likeC
MLTIL-----QSEWAIGPECTYDV--LVNMSLANFQENDNFCTMIAELKCRPGDDSLSCRFQNSQIKRPDPT--VPVDEPFEIRFNKGI-
ENLVVHKSILRWRLDMIKVIVSQLNIGFETQERDRFTIMENSTVGHCEVDVKITR--YGLQVEKIRQPKRCPHRYFFGNHK----------
----------YS----------DMTTSESHI----------TSKSKFVTSEGTIKTANKTRIMRPTQNI---
----------SLK----------

----LKRI----------DIARSP----------LP--

----------EIQ----------NP----------STSLFAFS----------
----------NLEIVHEEII----------HNL--K----------
>NP001291514_3_Sinvicta

VLLLLVGVAVEHAWEPQNEYQYSVFVRTLTGVDTLKQQYTGIQLKGVLVIQVSEELLQAKYINPRYAHIHQEPEIPMGKPFEIKLKHGVI
RDLLFDRNVPTWEVNMLKGIVGQLQIDTQGENQATFKAMEDSVGGKCEVLYEITPKGHHYEVKKLKNYEKCQERYHYGMDVKMTEN
MMKRN-KVVSELSTTHIVITGTSFTIQSTEMKNEITVQPESSDSPIGTVYSITKLTLAKINKISNLS-
NVESTGNLVYIFPFKGKTIMKSDNVMQLAKDLLLQIAKEIQNPSEGEN-
TLEKLKNLIRTMDRKQYTELEQYVTVENAWYTLRDAVVHAGTGPAFVTIENWLKSGQVKGEEAAELLSKIPKSVHQPTPDYIKEFFKLIK
SSVTQQEYVNVSAPLAFAELLRNNYVV-GNEE-
IDSYLANQLQQGYLENNTQKIQTFIFALGVTAHPKIISVFEPYLEGKLPTTRMLMVAALYDLSRDIPKLVGPIFYKLYMNENEAHEVRCM
AVQQFILTDPPMITLQRVAKYTNYDQSDQVNSAVKSTLNSIINTKRPEWRNLANKARSVRYLVNPNYDTWYSKGYFENWVFKGLNV----
--
KMVASNDAVLPRYVYVGLDSIFNFLRKPTFEVGYAVSSYRQVYDLINELWNGSRVEKLAQELKIKSGQKNNLEGHVLFNSVYGSMVYPY
DKHRIREAVAALKKLLTSDSK--
KTTAFNNFEKIFPMEMGVPFVYSFELPVFVKSEINFSGVYETLFCNRVQKRFGFIAPFEYQNYIAGIDKNGIMRVPLKYETNIDKQ-
KNFALKIHPSEIHKMTFSIKIEISKSSNNE-
RYMKIDTTFESKQVAKCEIQIDMFDAVTIHEIVNVVFVADVSFSLPRLQTYVFTVGSVRSNIDQKLRHFYVNTNAA-
QYELCYSQEVQPTPLNFEYAINNEPKDKLKGVLRYGRTCNTGNEIVITGSSSQSPQLRDMIESSITKQCMEEIQKGKKSVRTCNKATDVA
QVRD--QLNFHIDAS-SEIQVIGLLN-YTNLSETNTIVVQMVPTVQEPWYRW--P-SESQQSEIDVLLD------
EPSCTLDNDKILTFDNQLYNVQLGKCKHVLLTTYPESSKVAVLAKDTDNDSRNVYVWLGNLEIELKKVGNDLKVAINGQNVEIPE-
VQLPDGSSVISKGITVIFDGKHVRLYANGAYRNAIRGLCGNYDSRRDNDFPEN
>NP_001291513_2_Sinvicta
TLLFLAGVAVEHAWETGNEYQYSVFGRTLAGVDKLKRQYTGIQYNGILTIQVSPELLQAKFDNQHYAHIHQEDPMPMEKPFEIKLKHGI
IRDLLFDRDVPTWEVNMMKAIVGQLQVDTQGENKATFKAMEDSVSGKCEVLYEITPDGNHYEVKKLKNYERCQERYHYGFDIKSAKK
WAGQD-KVISQLSVTEMVISGDRFTIQSTEMKNEIAVQPETSDSPIGNVYTITRLTLKKKNSISNIS-
NLESTGNLVYTFPYKGKTILKSDNVMQFAKNLISEIAEEVQITSEGEA-

TMEKLKKLLRTMNRKQYAELEQYVSDSNAWTTFRDAVLHAGTGPAYVTIENWIKSGQVKGAEAARLLSQLPKNVYLPTPNYVQAFFEL
IKSPMTQQEYVNVSAPIALAELLRNSYIG-
KNDEVVGNYLANQLQQGYQENNSRKIQTYIFALGVTAHPKIISVFEPYLEHSLPASRTLMVAALSDLAKVQPKLVGPIFYKLYLNENEAH
EVRAMAVHEFILTDPPMITLQRIAKNTNYDTSKQVNAVVKSTLESLVHTKRSEWRHLANKARNVRYLVTSNYGNWHSSGYFQDWLVN
GLSL------QTIAGDD-LIPKYVYVGVNSVFDFLDQPSVEAGYGLSSHRQFFNEISKQWY-
SRVEKLAQALQIKAKEQNNLEGHFLFNSVYDSAFYPYDRHRIREAVAALKQFLNGNNK--
EGSAFNNYENIFANEDGLPFVYTLDAPTFTKAKVNFSGTFQALIANNVQQQFGFVALFEHQKYIAGINNNRVLRVPVEYDVEFNNNAKN
FALKIRPSPVYKTTFSIKVEVLRLANDD-
HYTKISTILTSDQIQKSEGHITMYDTVTIGQIANNVYIFDVSYSVPMLHEYVFTFGGMKSNSNQKLRGFYWNSHAP-
QYEVCFSHEMQATPLNFKYALKNSPRDEYKAVLKYGKTCATGNKVVITGSSSQSQQLRDIIESSLPNNVWKRFQTGNKAVENCMKANDI
AQMRD--QIDVQFDLSPESVKIIEYLEKYKVCDEENTI--KFASPVNRMWPNWSGPSSESQQTMRLNVAD-
YPDEFEQSCTLDKDKVYTFDNQLYNVHLGKCKHVLLTTYPENSKVAILAEDADNDSRNVYIWLGKQEIKLKKAGNNVQAVVNGQNVEI
SD-LSLPDDSSVVSKGINAVYDGKRVVISASDAYRNAVRGLCGNFDSRPNTDFPEN
>XP_025987667_I_Sinvicta
AFLLIIGAYA--
AWETGNEYHYLIESRTLTVLDKLSQQFSGIVIKGGLTIQVSPDTLQAVVSKTQYAPVHKTADLSMGKSFEIKLKHGVIRDVLIDQDVLTW
EVNLIKSIVSQLQVDLQGENSGVYKAMEDSVGGKCEVLYSITPDGLNFFVTKTKNYNKCEQQYHSGITDKMNKLGSNDG--
VLSRSSTSSIIISGNHFTIQSSLTTSEIFVRSKLHDTYSSAVYDSVKLTLDRMNQISNASNNLVSTGNLVYIYPYKGESIQKSE-
ITSVAARFIAQIAWNLENAFNSSTEYIEPLIRLIRTMNVEQIAELENKLKSYTTWDIFFNAVVSAGTGPALISIKNMIKNGQLKDTQAAMIIS
KIPKTALTPTSEYVNKFFELITDEQTKQRFLNTSAPLAFAELVRYTQSNRQDNALLEPYMANQLTEAIKDGDSRRIQTYIMALGNFGHPK
VLSVFEPYLEGTLPASRLMMVVSLNRLSENFPRLARSVAFKIYMNIMEAYELRCAAVYVVMKTNPPLVMLQRMARFTNQDQDRHVNSVI
KTNIDALANLEQPEFQDLAAKARIARELLNPIDTESYSQGFLFKKIIPSLNMAEITILQIIGSQDTTVPKSSYLNIYQSYGGFNLPPSRMSYEI
SSFRALLDMW---YE-

LIIEETIEKLGIKGEDPVQFEGNIFVNTLYSSQFSPLDNNTIEETINAFKRIISSWQRSKNINYLHYYDMTFPTESGLPFIYTLTVPKLLRINIG
GTAAGYIMVHEKVQSRIGFVTPFEHRHYVAGIDTNTRLVTPLGLSISVNEENKKFKLTLQPLPVHTKEHEIHFYVNLFSRGA-
YYRRFDGLMYFG-----EVRIRAYDFAKLDQFLKEGYVFDMSI---
DQGVQVFTLAYSYSQIDHKSQAFYWNVQSVDPAELGAIGYVKSISLNPEKALEQIPNDEFKAEIRLG-
NNFNEEMIKLEGNWTRTDDVKDMAMSEIVKKCRQDMKQGNILLPACQKANKLINQKD--LLMMSIDTTSDILRGILWIQ-
VETMNSKNTIDMEIKMLPDNDDAKISLSLKDIINDSNVSMKDTFKEQLD-
VCVLDKTHAVTFDGKVYPLKLGKCWHVMMTIYPSDMRAIVMAQEMDDGSKQIKMILGDQEVHLQKSGDCLEASVDGETANFSDHYG
SNE-TTVFSTEITVEYDGEHILLMISDNYLNAVRGLCGNYDTQPNNDF---
>XP_025987667_II_Sinvicta
TFLLIIGVAINYAWETGNEYHYLIESRTLTALNKLAQQFSGIVIKANLTIQVSSDTLHAAVSKTQYAPVLKVTDLSMGKSFEIKLKRGVIQD
VLVDQDVPTWEVNLIKSIVSQLQIDLQGENEGSFKVMEDSVGGKCEVLYRIRPDGHHFSITKTKNYIKCEQRYHSGITGKMNRLGCSNG
--
LLSRSSTSSISISGQRFTIQTSYTMSEIIVKSKINDTYSGAVYSSVNLTLDRIGQISNASKNIVTTRNLVYTYPYKGKSIQKSENYASVAARLIG
QITGDIENVLNQSIEILEPLIKLIRTMNVKQIAEVENKLKLYTAWDVFCNAVANAGTGPALISIKNWIKNGKLKGTQAANIISKIPKVARIPT
TEYVRAFFELITDEQTKQKFLNISAPLAFAELVRYTQSNKHDNAFVEPYMANQLKEAIEDGDSRRIQTYIMALGNFGNAKILSVFEPYLEG
AIPTSRTMMVVSLIKLSENFPRIARSVAFKIYMNTVEAYELRCAAVYVIMRTNPPLAMLQRMARFTNQDQDRNVNSVVKTNIEALANLKQ
PELQDLANKARIAKELLNPIDIDNYSQGMFEESSTPSLNLVQIAIVQMIVSKVSNIPKSAYLDLLQSYGGFNLPTARISYEMSSIIDLIDML---
FR-
LFIEQTIEKLGIEAEDPIQLEGNFFVDSLFGSQFYPIDNHTIEEITNMVITYVTSLSKTQNTNYLNYYDMTFPTESGLPFIYTLTVPRLMRIS
NGGTVAGHVMSNEKIQSRVGFVTPFEHRHYIAGIDVNVQFFVPAGLSVSC-KDNKIRSIKIQPRIVHTKEHGIEFSCNLLYPGA-
HYRRINAIVHLS-----PAEVNISDFVTVNQFLMEVHVFDISV---
SVGIYVFTLASAYNPIGEQTYISYSNVQSSDGMEVCADIYMKSLPVIFEQIIEGVPPYELEADLRLG-
SCTNGQMIRIIGNWSRTDEVKEMFKSEVVKKCLQDIKQDNIWVPTCLAASKLINHKD--

LLTISMETDSDYFQVILMIKMTKHTDNKNTIHMEIKMPPDNNDAKISLSLSDVVNDSNISLKKSLEGDVS--
CILDKTRVVTFDGEVYPLKLGKCWHVMMTTYPSDMRAIIMVREIEDGSKQIWMLFGDQEVYVQKSGDHFEVILNGLKV-FPPESQSGQ-
TDVFAMGINLQRFNENILLQSSYKYYNAVRGLCGNYDTRSNNDF---
>XP011161539_Sinvicta_likeA
VLLCLAAAAVPVHFPAEKTVIYNYYADVRTGIIDPVHYVSQFSLTGQLHIQKLNNAYYIKLNDVKYGMYNGMHQLGIHEPFVIVYEYGKF
QGIKMPETESGWSRNMKQAIASMLQLDLAHIQKHAFVTHENTIHGTCQVAYDVHP--
NVFVVTKMHDMQNCSHFVQRVFNHDECKCHMESMDDTVASASRVFEIKQQDDILIERMVAHSTIVYSPYNSRSDSHYLLTNTTLELHS
VGPTIEVPAVNVQNVPIIRDIPTHGRHIVKLDALIPKLKKMLVEAADYLEKEPEWHGQTINRLQHTMNYMDLGSLEQVYNAIKEVTMRN
IFLRLIPTVGTTAASHFVRNVVRKNKVDETTAISMLMKLPMHVKVPSERLLLEMEELLKLDTTVSPQVRKASIFCFSILIRKTFIHQ---
PLIERYLHRFFDHVKNNQSYEMRMVYLMAMRNVQVGNMEKLLEPIIRGEIMISRTQAIWAIKKAVADKVEYTHNLLWPILADVTQLLSIR
IAAYDVLMSQMPNMQRIMNMYWLMMHERNNHLYNYHLTTLKGLANSVDPCLMPVREMTRMVLRFTKRPVTTHLSSKHFVDYTDPI
YEHGESMNTAWVMDEITGMPKAGYLEKYVSVARKPVNKVGIYWSAYGLDEVMKVVKELFGNKNVQNILLSAAQDMPMKRTLHVDICI
TVNLVVVVHHFDQNTFKSLLDEMYQIKQMM-------
TINFQEIQVPTDMGLQGVFSTKMPHLMSLRFNNKFDIDTRIWRHGEYAMSVYNPMVDVWHSVRRATVQDIALPVHMSVGYNPEAKSL
KITMPRLPVNKLSTGIRSYIIQ-EQDILCHHHTVVTTGEKDSKDVGLKYSIFENKMTPDYI--
TSHVDLSIRANSYKNVRGTLDAKAASTNESVRSFKVQMLRQTGLKICVDAQKNYPVGTDPLKID-
STKENSKMTITAGFTDDDEMMIVMTVKGEMTEEQKRQLTDNVHGACVKDTQQRQGHIANCVRETLQYTTLKKYTVNATYKKVPVQ
VMIEDHFRAYITDRESGNARVVLEYASDLSHVNASVLFDGMVSMPRFSIGAMYKMTQNKICTIYPQSLVTLDEGVIPFVTLDKWTLVSG
D-
HIDRTYAIFVKSVQNTKLAVKMYIGDHEMEIIPNDVNVHVSVDGKVVNQERGFRLTETFHLVVSPVMMMWTPNSVGLMLDTVLQGHV
TGVCGHMDGTHKEKLPKI
>XP011162482_Sinvicta_likeB
FILLLSTEVAQSLFRSGKEYIYSYNVTSNTGVLVPSNAASSWNLNGKLIIQAEHDYVTVQLQSLKTSMHNG-

IDIYLLKPFRLSYNKGLIENFSTEA-
EPIWTTNIKRSIAGILQLDLLNLEKIAFHSSEINHYGKCDIDYVVSPGDNQRLIRKSLDPRMCIGHPNCYWSNVPKQCTDEDQNPILKSSE
RIYQLK-TDHDII-SVNASGSIYVQPFQSMGEAHYHFVQQTIELSSVKDITNATNNLQTTA-
LQHELPTQGRGTPDKNFVFKSISRLLDRLSHRLEEIHNLHNSTISVLLYYLDMLDRVDLGKAYTRIKEETIRNMFLEALPQVGSKASALFIL
DLIQGNKVSDISAIQLLMRLPFYIRRPDVQLLVSLQPLLTLASKIRMEVQNTAILAYGTLIYKTCLVHCPYEMLDDYVRLYLDKFTESTRY
EQKMVWLEGLANIQLGRVVEFLEPIASGNNAESRVLAAWASIPTAPLRPDVIYPVYWPILLNRTEHLEMRVAALTLLIVSSPTPNRLISLY
WYMQNEHDQHLYNYFYTMVKSMERTTYPCYKRISKIAAQFSRVLRPNNDYLITGNYLVDYQDSLRRFGAMLHGIIIANPFTNIPEVIYVS
LNNYASGTHINHVSLYVKAEGIFHSLA--TSF--PTNIKDILKEFKLDERRKGPVHLEIITRIQKTVLCVHWNETNIIEGLKYLSSL----
SDNIMEFHINQQRIRQVTDFGTNVRLAMTGTSLFSMRGNFNNHIILRTSVHGIETIENYNPLVDLWHSAERVQSLHGYLPINVTLGL----
DERPFISYDTLEEHLKTGFTAHIRGAVRSKLCSVSHTVVKSPSELPELGGRLKIFENLMSHDYFTY---SQTRLEY-------
LSLTHRDLQSSQIFHQIKFKAIKIVGFKFCVEVERQVPWWEI---
LSNEPSGRIKLNIIWGPSDNAGSSISMNLIGEISNEQLVESKNWPYDECQRDSV-
GKRIVPSCYEVSKELSTLRRYEIIAQYENLPESLALYDLIGGNSSIAKSDELAITATFPKESNVGELQLAIEYNSHYIRIHKYMEIPLFRMSACIV
TSNSIRSAYNTIYSFALGQCY-------DENPRLAL---TVANGGIDVNLTDESGITRITAGKDRGTL-YNATSHIQEYDMHWTQEQ------
VLLFFPNHVLEFS---------CGICTLDHPNIDKLYEKL
>XP011172130_Sinvicta_likeC
MLTVL-----RSSWLTGPECTYDV--LVNMSLANVQETDNFCSMIAELKCRSGEDSLSCRFKNSRIKRPNPT--VPVDDPFEIRFNRGI-
ENLVVHRSISRWRLDMIKSIVSQLNIGFEIQEREKFTIMENSTVGYCEVDVKVIR--EKFQVEKVREPKRCPRRYFFGNHE----------
----------YS----------DMTTSASHI----------TSKDKFVSSEGVMKTANKSRIMRPYQKI-----
----------SLK----------

--LKRI----------DFAQSA----------LP---

----------EIQ----------DP----------STSLFAFS----------
----------NLRTIPEDE----------
>OWR44310_Dplexippus
LVATIVVVSSPWPWQVGKLYRYDVETHTLARYLD--
SFSSGNAFRAKFTVRASDGYLQARLENPEYAKVYQKMPVANDKPFEIYIEGGRILSVKLPFSVTLMQENLIKGLIGSLQVDLTSHRSGLFR
KMEIDVTGDCETLYTVSPDDEPIEITKSKNYGHCHHRYHFGVPE---TAHKTGKEQFINRATVSRMLVGK-
NGHIYKAETTSTVTAHPHLYGEQKAQVHGKVRFNLMSYEDDNEYPEGAREVTNLLYALPPDLHGDKKQNN----
AQKLLQEIAQQLQNPNNMKADTLSKLVRVIASMSYGQLGLTSRSIDVVNMWMIYRDAVAQAGTLPAFQQIKAWIESKKLEGEEAAEVIS
VLAVSLRYPTKVVMKQFFDLAMNPETKQMFLNDTALIAAAKLINMGQVNNEDAFVINPRLSQELQLAIENGDSRKSQVYIKAIGELGHP
AILDIFKPYLEGKIPASRTRIIEHLYVLAKGRDDYVRAVLFSVLKNTAEPYEVRVAAIDKIFMSRPSTAMMMAMAQMTKDDPSIQIRAALKS
AITSASELKNPRFHDLARTAAAVKDMLTSEFGLQYSGKNFLEHYDRDEQPSSMSVLSRLGSKDSLLPKYWRYSWKGRDGGWDQ-
ETVISGAASSWQELFDLFADQMFKYSAEKIADMLNVKKDDRESSEGSFYINLLNQRRYFAFNENDVKELGIKFREYLTNL-
KDVYTKVVNRNQVSFPIATGVPFIYKYKEPVLLHVRTVTNSELRIIYAENHDGNVGFLDTLGNQLASVGLVRKSQLNIPIKIDLEME-----
AKFHLSPRPEKVAQGSIQLQLTSIKQKDIAQTHFNLRYLGKQSKNKGVTITVYDTLYNQELIKRAHVVDLSAKFETEQEYTATLAIGASVV
DQKIQFLFAGRNSDQGNQLNAVGRVTSDSINFQKALEKELKMDFEADILYNQK----
ENIHILGSAERTKRYIEELQEPQVKRCLENYARGNYYQHDCHEAVVMAHAPDNFKFSVSYKDVSSGTYAYRILDGRS--
DPAGKLELNVDALYWTRNLNLIVTSRAVSYLP--SAYERILNYYTPYCSVDSNRVRTFSNREYDYTLSPSWHVVMHDD-
RNEDLVVLSRPQEM-KMQIYLSYRSYEMEVQPAPDTVQVKTNAKKV--SEGYSTGDNVMIKLERLRIVYDGERSVVLSRD-
NRKNIRGICGRMSGDPRDDYQEL
>NP001037309_Bmori
FVAIIAAVSSPSPWQVGKQYRYEVTSRTLAHLQE--
GPSSGSAFKAQFTIRVSPGRLQAKLENPQHGNFNEQLPTPNDKVFEIEIDGGRIVSLDFPTSVPVPQENLIKGLISALQLDTSAHRSGLFRK
METDVTGDCETLYTVSPEQDPVEVTKSTNYGHCHHRYHFGVPV---TAHKTQEQQLIGRATYSRILTGK-

EGPIYKAETTSTVHVHPHLYGKQKAEVYSHVHMELISVDQDSGRAGAMRPAQSILYSLPPQPHADKKQNN----
AQKILQDIAQQLQNPNNMKSDFLSKLVRLIASMSTEQLSQTSRSINIIDMWMIFRDGVTQAGTLPAFKQIQSWIENKKIQEEEAAQVVVA
LPRTLRYPTKQIMTQFFNFARSPAKDQMFLNSSALMAATKLINLGQVNNHDAFVLEPTLAADLKATVEYKDSTKAQVYIQAIGNLGHRE
ILKVFAPYLEGKVEISRTHIVKNLKTLAKLRDRHVRAVLFSILRNTAEPYPVRVAAIQSIFISHPTGEMMQAMAEMTHNDPSVEVRAVLKS
AILSAAELQHPRNFYLSRTAQAARYLVTNEFGYQHSFKFIDDSYDEDNDIGTF-VISHIGSEDSLLPKDFKIVTNSKGGAWE--
RNTIEASFSSAERFLDYLRDSVFKYSAEKIAKLLNIKNDEEEPLEASFYVDFMNNQRLFSFSESDLQQLSQYISEYMKKVESGAYTKVYNQD
QVSFPVASGMPFIFKYKEPAVIHFQSKLIKDVQFTYAINIDGNVGFMDTLSNQYSSVGVVNKLQFNIPFKFGIEIL-----
IKFRVEPRRSKVFSHAVYAQFTNLTQEDIALTHFNLKHLCKQSQSKALTITAYDEYYNQEMVKLVRVVDLSASFEGSQNYVLTGTWGDS
PVDSKVQGLFAGTKSATGQQINAVFATTIHSLSFSKPLQSDLRAPFGMHFKYGQS----
GEIRVSGSFDRTKKYTTELEHPLAKQCSQQTTLNNFYQDSCHKAIVMAHAPDHVEFSVSFQDMSPQYHTYRLYEYYT--
EQNGKMDFKIDFSYFDRTYTVDIAPGALSYQP--
KAYELVANYFTPYCSIDGTRIHTFSNRSYEYPLSRSWHVVMQDESNWHELAILSRQQRD-QQEIYISYKSEEIEIQPA-SGVKVTTNTKKI--
TDDHTHKDGVVINIDRIRAIYDGQRFVVFTQD-YRNSTRGICGRMSGEQRDDYQEL
>XP004926335_Bmori
--LILSATASQLLFPENRQYIYFIKANVSSTTLAPKRSGSYWTLEGKLYVRA---RLRLQLSDLKKSV--------YKEPWDLDYEGGLIKTISVGQ-
EPVWCTNLKRALSVNFQLK-----KGTYYRDEPCLDSVCMMIYT---KG-----NKIQKFTSYRLP----------
TTKSSRSWSSLSGDKPSVDTINTAERLYV----------VDLDLQGSYEFRTHDHILAVNSELSINYP--
GKSVEKLNSFLKNTQELLLKIARGIDDPSLINLDFIE-LLNTIAQLNYTTIIKLFEDL---
TSRNIFLEALPHARSEACARFIKYLVEKDKIEDAALLSLIRKLPFNVATFDQSLLEELEAFTKGP-DFRAEIRHAGILSFASLLHKTI---
EFSLIKQNVVVKYFRMYSDCPQYLDRMVWLQGLCNIGY-----SAEGYLSGR----RLWAAVAIDP---
KHEEAFKESTLSILTNQSEHIQLRIASLHMFLTSQQRENDFIFVHNYISTCNNNQLKRFWYTTVKNLMTSKSFTYRVASYYVPFIVNQVSP
PL-------YTNNYIISGEEVGPSLQVISVGDTLGVIPPFIAAKLNT--GGRRPYKVAVYLIIEGVTTTLKSL-----
VLNVEKLLLVLKNMKAWTEKVHIDIVIKTHKTIFATHLNQSRIDSVNAYIEDFLR------

FGSHINQQMVDVPSELGTPIRLQSLIVSFTSIRGNLRNDLHIRYQGTSVTSLSTDGPLIQTAHRARIQQSMVAHLPMKFNVSFDGS---
LQLTWPN-------GGVAMH----AMESVKKYNYTVTVG--DSSSGAVNQGIFDRPVTDDFLTSNDVLQAKIQARTEDKFQLT----------
-------FSL-------DVFLNMEAITNMEQNS--------GNT---DVLVSVRGHQAKNDSTKDWTL------CLRDVDQDISAFPKCKSSTSE----------
LKYKGVPQHHALGQLLRS-----ERDSLSVNIQFKSDNFAWLLGM----------RECRIKENEIKTLSGTTEDLPSIDC-
TLALADCSNSPRFAV--R------DGIKVFSGGY----------VSVKKNQSEIYVDDTFWISAKSKITSMALKVYYDLNETVILV---YLNNVCGLC----
----
>BAD91196_Spryeri
FVVLIAAVSSETPWEAGKVYRYEVKSHTLARLEE--
GPNSVSAFKARFVIRVPNGRMQAKLENPQHAQIHEQLPVQNDKPFEISYDGGRINTLSLPSSFSLPHENLLKGIISTLQADLSTHRSGLFR
KMETDITGDCETLYTVSPEEDPIEITKSKNFGHCHHRYHYGVPS---TAHKPEEVQFLRRASVSRILGGK-
QGPIYKAETTSTVQVNPHMYGKQKAEVHSHVHLELESAEQDSQKPEGNRHIKNLLYSMPPQTHTDKKQNN----
AQKLLQEIAQQIQNPNNMKADFLSKLVHLIASMTSEQLSQTSRTINNIDIWMVYRDAVTQAGTLPAFQQIKTWIQSKKIQGEEAAQVVA
SLSTTLRYPTKDAMLQFFKLAQSPEKDQQFLNTTAILAATRFINMGRVNNHDNFVLKPELAEALKNAIQRQDSVKSQVYIKAIGNLGHP
EILQVYAPYLEGDIKVPRVLMVSNLNVLCGQKNRQARAVLYSIVRNTAEPYEVRVAAIHNIFISHPTGAMMQAMAEMTHDDPSVHVRAA
LKSGIESAAELTGPRSWELSRSAQAAKWMLTKKFGLQYSLKQFDDGYDMEDDLEIFSALSHIGSDDSLVPKYLKYFVKTKNAGW----
NKVQASVSSYKRFVEILKEGMFKYPSTKISELLNIKRDQKDPLEASFYIDIVNHQRYFTFSEEDIKQLPHDISEYFKKLEKGVYTKIINQAQV
SFPVAMGVPFIYKYKEPTLIHIQGKATKEVQFTYARNIDGDVGFMDTISNQHVSVGVVSKLQLNIPVKLDIQVE-----
IKVKAEPRQKKSLSSTIHFQMVNMSQQDLAMTHFNLRYLGKQSQNKAVTFTAYNEYYNQAIAKRATVADFSASFEGSHDYVLTAAISKS
PVDPKVQYFFAGKNSAQGNQVNGVATIKTNALNFLQVLDNDLKTTLQVDIKFNHN----
GNINMEGEAERSKEYTEELQHPLAKQCAQDIARNNQYTHTCHKMIVQAHAPDRFKLTTNYKELSNTYHAYKLAKYYT--
EPEGKVEIEVAASYLDNTLDASMAAAAVAYQP--
QPQERVANYYTPYCSVDGNMIKTFSNRSYEYTLTSSWHVVMQDEPIGQEVVILARPKEN-QQEIYISYKSEEIEIQPAPEGVNVKTDSNKV-
-SEGYTQQDGVMLYIEKLRAMYDGQRLVVLASE-NRESARGICGSMSGEPRDDYQEL

>AAB03336_Ldispar
LLAFIAVVSSQSPWQNGKQYLYEVNSHTLSHFYE--
GTSSGNAFKANFIIRVAQGLLQAKLEEPRRAQFHQKIPVVNDQPFEISFDGGRILSLRLPEVLSVANENLLKGLISTLQVDLTNYRSGLFSK
FETDVTGDCETVYSVSPDSAPIQITKSKNYNRCHQCYNFGVPR---IAHGNTEKQFISHTTVSRILVGH-QG-
IYNTETTSSVNVHPHIHGRQKAEVHSYIQMRLKSIEENPGKPEATRHVKSLLHSMPTQPHVDKKQNS----
AQKLLQDLAQQLQNPSTMQNEFLTKLVRLISTMTTDQLAQFSRGIKIADMWFIFRDAVIQAGTQPAFQLIETWIKTKKIQGFLAAEVISS
LAKTLRHLTKDTMKQFFELALCPELEQKYLNSSALIAAAEFIRLGSIDYNDRFIIEPRLSQLLKQAVEQEDYGKTLVIVKAIGTLGHREILKV
FTPYLEGQIKVSRVQMILALKPLAEQKDRYVRSALFSILMNTAEPYEVRVAAAMILFLNEPTTDMLRVMAQLTNDDPSIHVRAVLKSSIET
AATLEGPKYWHLAKAAQSVKELVTSDFGYAYSAAFFAKNFDSEE-VTRYRLANYIGSDGSLIPNLFSLTWNNRFYGRGA-
ENTVGFTSSNVQDIFDYIKQLILQFSADKIAKLLNIKQDPLQPIESSIYYKIFDKEFFWPLDQPNIKQLVNDVFGFVKELENGVYTKVFYSK
QVSFPIATGEPFVFEYKEPVVVHLQTKLQTEIQFTFARNMEGSVGFIDTNTNQLANIGVVKKYQINLPVKINVIGK-----
FQVKLEPRLKKVAENLLKYHFSDF--
NDLAETHYEFTYSGKQSKTKAVTFTAYDVSHNLEIVKRARVIDFSASFEGLQEYAVTAAVAGSMVDLKTQFVFMGSQSD--
GQINAVFKLQMAPLDFHKALNSAVKYLIEADVTYGEN----
SNINFKGHTERSQEYAEQLKSLWANQCAQENAQQNKFQLGCHNVLIESHAPDRFKASITYKHIP---
HTALLDSYSKGFEPVGQIELEANASYVDQTANVAWTPALTTYQPIEDSWTHFANSYSPFCTVDGTKVKTFSNRDYNVTLPEIWTVLMH
AQTNWEELVVLAKPNEAKKLDLYISHKTAEVNI-PYSAAVKVETNAQLV--ADGSNHPDRVLLHLSGLHLLFDGKRGIFTTSQ-
YRNITRGICGQNSGDPLDDYQQW
>BAD05137_Dmagna
SISLISSLVSEQSWENGYTYEYDYAGWTSTGIMGISTKVSGGSIKGRLTIEPDESTAVVALLATKGKQFNEDPGQEHEKPFQVKFVSGKVE
TVAIGKEEPLWIVNFKRALAAQIQLQLDGVSQTVYHAMEGSVTGECQTWYHISRNFPVYEIVKNRDLDNCRILFNYNSNQ---
LVNGAGCENKISHSDTVRIIGCTSHFIVQRIKSIDKLVVKPFSYETEATEGLTVQHLTLRSATPTGYISSDVHIYQTLAYSYP----------
EILKKALRLLSEIISDVEDPST--

KHTSEKLRRALASLDYNELMGFVAQVSKETSNQIVVDALMLSGTNPSLMLVREYILQGKIAGEQAVQAISALVPTVETPTKELLTSLMEF
LKSEVQSHRQLKITTALSLSRLVYQACVNTSDSIVASPYLAEQAKIA---
KDAGERMAFLTALGNIGHEIIVPFVKPFIKKTLNLKRAKAIFALSTLAVQKKEIVGTLLMPIFFNKAEETEVRLAALTLLFVSNPPQAFWSR
VALSTWYEPNDQISHFIYTTIASRVANKNPLNREETVRAEAVIALMKPFWTSHAALNYQKAGFSEKTRLGYVTETVNFPGFESFVPSHHYS
SLSVAMGPWFTKLMEISIDSKHAEKFIDRLVGKPGSPELEKIHEELKIEARATGQPELYIYLNFLNYQRFFTINPNTIFMIEKQILQSGFREN
AKLFHKYWPLLDSFIPSAMGLAYTWASHHSVFVSLKSNVEGALKPVVVSKMSTRLMVETPFSRAYPTTGVDMEWAAALPGRFAIEGDM
KTGKIQTTWETEEPKENKLGMDFVVQPSMPSTIRQREWSLYLDNASSETKAIKTSISDSKFEKKF-
QHGYSLDFTAELPKSSARRIGSSIVYGMDGKSHRGLMVEKRDEVGFVLCAEVDAQNLVFKRKELIKDETERRSTIKIGFGKSCTDDRKITV
ATTWTRSEEDISPSLQWEKTQCQKQETLGRGMSDECIAARRLSSILNKAVMTIKYNEMPAVVKASNLVRHYMSDN-
ENQITVESVYYPLAGSMDVKV-------HPEVILPKRMAVPRSGVCLIGSETVTTFDGLFYNASFSGCDQVLTKDCSGRYKFAVLSR-
VEGDKKIVTVLLNKEKIEIFPAQ--
QKVNVNGMEISVSESKKTADNFEVDSIMIRVLTDAKEVVVLASPIHRGRLCGLCGSQNGDKVTDLAEL
>ABW77320_Pmonodon
LIMLLLGA--CQALQPGLEYQYRYSARVASGIPSINRQFAAAGIQADVTVQMADYSVVVQFSNIEVGDLNKV--
LELEKPFRVIMNESHVPSYMEGPEEPVWISNVRKGFVNIFRVPLFWSTTLNYGAVEETLVGRCQNWYTSMK--
TLWSVRRTTDFDMCENLLQIHSSK---------TEDIIQRSSVASYLIRGD--RIEQVVVEGTITVLTNKEQHEHVDTFTNQTLELKAVREVGQ-
---------ITVTYPPTGRPITDQIGIMEYIKKTVDDIAQRMAEPES------
LSLIEAVANLDYPQIQTLYSHFHGKLQKNIFAQLVIFAGTEPSVTFAIDNIKEPVLR----SDFYSIFVSNLRSPAP--IPKILNVLMNE--
EESMVHSIGLVNFATLAHDACLSEDPELIITPYLVRGLGN--
QETDVWQRLTYLQALSNLGTPQTINVLKPIILGVTETNRTNAIWSLSAYNMHKSALSQEILMPIIENKGEQFEIRNIAFLTLATWGPGHA
WWQQLAVSTWHDPSPQFANFVTTTIYSISD-----HTKLAETVSRIKHLCKAPASILHSTNFLHEYLFSEEH--
SRVNFAWFASVHGAIPEEVFLRIQREFFYGYTKVLKIDGQQRGLYNLMQLMKQKMTSGMWEELKEKIGFTSPESAS----------
DATLWLKLDRLIYV----------ALSGKPFYDKDLLFPTEIGIPFLYVAEDATWLNINGNPKVNVDFIFET---

SQVQFVVPWSKRPAVAGVYADTQLVLPLNVHASFDMYSQEIQLSIEPGEVPLFNVGL---
ITQLAHTSTHIWEQSIVFDPTLSTTKTVSFTFYDDGQISERIMRGHVRTISVEVQGTPEASLTWAISSSASTRSSKVVTLIKHPA--
YTICLNAQIVNPFFTTEDVLTANYHSTVKAELFEGETCEEAPVLLLEASFDVSSNVKERVK-AIEKSCDYSTYAGIDII--------
TSPLYDHSHITATWTPDPDNLYVDDIIKAHTSAHHASRVEID-----ATKCIETRTWANPIFPAPATTLLYDLAVGRT-
ACTIEETKIRTFDAKEYSYEPNVCWSAMSIDISHNQRGAILVR-----QWEIRIIWHMEVFDLTP----TNFLVNGEPAQGTDN------------
HTIVFESGTRIKV--SDKHRGNTAGLCGNFNGEPSNDMFNF
>AAF19002_Pmonodon
LILLLLGA--CQALQPGLEYQYRYSARVASGIPSINRQFAAAGIQADVTVQMADYSVVVQFSNIEVGDLNKV--
LELEKPFRVIMNESHVPSYMEGPEEPIWISNVRKGFVNIFRVPLFWSTTLNYGAVEETLVGRCQNWYTSMK--
TLWSVRRTTDFDMCENLLQIHSSK---------TEDIIQRSSVASYLIRGD--RIEQVVVEGTITVLTNKEQHEHVDTFTNQTLELKAVREVGE--
--------ITVTYPLTGRPVTDQISIMEYIKKTVDDLSHRMEKPEN------
LALVEAVANLDYPYIQTLYTHFDGKMQKYIFVQLVIHAGTEPAVNFAVDNLSDILFY----VTFYESIGVSLRTPAP--IPKIMHSLING--
EEGLHHSLGLMNFATLAHDACLSEDPELIITPYLVRELGN--
QGKDVWQRLIYLQALSNLGTPQTINVLKPIILGVTETNRTNAIWSLSAYNMHKTALSQEILMPIIENKGEQFEIRNIAFLTLATWGPGHA
WWQQLAVSTWHDPSPQFANFVTTTIYSISD-----HTKLAETVSRIKHLCKAPASILHSTNFLHEYLFSEEH--
SRVNFAWFASVHGAIPEEVFLRIQREFFYGYTKVLKIDGQQRGLYNLMQLMKQKMTSGMWEELKEKIGFTSPESAS----------
DATLWLKLDRLIYV---------------ALSGKPFYDKDLLFPTEIGIPFLYVAEDATWLNINGNPKVNVDFIFET---
SQVQFVVPWSKRPAVAGVYADTQLVLPLNVHASFDMYSQEIQLSIEPGEVPLFNVGL---
ITQLAHTSTHIWEQSIVFDPTLSTTKTVSFTFYDDGQISERIMRGHVRTISVEVQGTPEASLTWAISSSASTRSSKVVTLIKHPA--
YTICLNAQIVNPFFTTEDVLTANYHSTVKAELFEGETCEEAPVLLLEASFDVSSNVKERVK-AIEKGCDYSTYAGIDII--------
TSPLYDHSHITATWTPDPDNLYVDDIIKAHTSAHHASRVEID-----ATKCIETRTWANPIFPAPATTLLYGLTVGRT-
SCTIDETKVRTADGKVFAYEPSECWNAMTLDTTFDQRGAMLVR-----QWEVRIIWQKEVIDMSP----TNFLVNGEPAKGTDN------------
HTIVFESGTSVKV--SDKHRGHTTGICGNFDGEPSNDMFMY

>NP732076_Dmelanogaster
VYFLIFDAHAFSIIGLNKQMLYEYEGNVLVG-AKPQDETTGWIVRGKLTLQRSELVLAAALVIDDVTLNNSG--
LQPYKPFKIALKDGAISHVVFKEGDPIWSMNFKRAIASVLQFQMK---
SGAFVVDELGIHGTCRTEYFVSNRTNYISIRKTPEVKTCKPYVHTTRSNVPPTCEFDHQKSVIIGNEAIYGMSPHNYYLS-
MAHAKGTTLIHTFESTGEAQFINSELLLNFLNETPIDNDIETSMAAEPSNLELPTGGRSPQQQETLIAQAGTLLDSLAEALESEP--
YDSTLSDVIKLLSEMDFDSLTKLYREVRQETIRNIFHEIIPRIGTKASVFLTHHLVLNKLTKPQIAVQLLIPMPFHIFELSAELVQKCEDFLN
GPD--RPDVRQAAILSFATLIHNVYVAKIDKEKFEEYVQKYFNAYLSDRDFDQKMLYLQGLNNLQLGNVANYLEPIVQDPNE--
KFQAAWTTLALADRRAERIYEVYWPIFESRNASLELRVAAVTLLLISNPTAARLISIHRIIQSETDPHMINYYRTTVTSISETTYPCYQHLRR
LLSYMHRHLPKESRYWVTGNYIFDYRDSKFGIGAMLQVFLVGDPKSDMPVVAFFKFDTEALGKFTGQLALYIKARGLPDTIL--
NKMQSFKSIKALLAMLQAPIINSKDLHLEFILQMEKTVLSYYLNQRMFRQLTNILERMQQII----
TDSHINMQTVTVPTVLGTSSDVLLQTTVLTSLRGNITLEIDARYSSYASVRSRSYNPFLNLDHEINREQGFLIYIPFSSELHLNESSKCRRYS
FSR---QNLTSGLSFKTRG-----LAPFEEIMVPEGRPMTDLGVRLT--------
NFTMLEKNTRIEGNFKGKQTLEHTDHMNENHAADALHRFKLTGQVQRSWKACTKLTYE--------PLV-
FTKRTLNGDVVFGLATEEGSTVQFAARAGPSEHARAFLRLTDTDFCPKEVL-KFSPIPYCKRSNFEFTSITQYDMDLKFDNMPAWF-
LDHLVSAKVDSLMSQEINISMQTPQDQFRLAVEV---------
KFDASHELTFDSGRSCSVINGIVNTFDDYLINLRRPDCLTLLVADCSPLPQIAVFVSPVQGLNYGLRVHIGQNYFNFRARTDNSSLPTDEP
VLIYNQDVELNEQNLIVETSTIQFDLYNILNFEIYGVYKHQMCGLC-------SKPLNRM
>AVP41182_Acitricidus
VSYLVQASCNRDLFPEGVQLGYRWRATTHVGTVFPSEHLTNYSVEAYFLVQNDQNFTNFQIKEFKTNENQQ-----
ITLPFRCVYRDGEVKKFESESGDIAGSLNIKRALATMFQLKLDSLRR-SFATEEAGIYGKCNVQYLVTKNNNTN-
VKKIINFSACNNKLGQQWSNTPPICPSSYQDGSMSHSVRNYNLDEIN--VIRYLNIIGTVEFQPFQALAESHHIFVNQTFELDHIFPIGET----
-----
LVSKLDTMGFKPSNKSLLLAEILDALGQLTESLNGTTNLDNQTEFRALELMWWLDTSDWFTLYDTISQETIQHFFWDLVPQVGSASSV

EFIRDLIKTQKITSFLATGLLITFPYHVRYPNEKLLKESEILLYLDRDLENEVRKVAILSFASLIHKTCGKGCLDDTQNKYIKLFLDKFIESTS
HSRQMLFIEAFSNMKIDNILDFLQPIITDHNR--
RLVAIWAAKSATTAHSDKVAEVFWPILTNHSEPLEIRTSSLNMLMMSQPTVSRFLTLYWFMQAEPSQQIYNFYYTTINSMANSKYPCYKK
YSKIAAQIVRFVQ-KWHHWATGNYLLDYEDPMRGYGGIVQTILMANQRTGFPSVLQLTAEQHSFGTS-SKFQIHIKAEGLADNLK--
KELFNFKHLQDVLEKAKFTIREDENLHFETIIKINQTIYCHHLNRTTFKSFALVLKRL-----
SSLFQFSLNYQRLTQFTDFGTPILLQFRTASLLSLRGSIDAELDFRYSMDTVTSLKTFNPLSSTWRGADRFRCIHARIPFSTEIITHFFRSYIK
ASAYRVDGSKL--GIVWH-----
ATKLVPPENTLPAAKKESRDLGAKITVFQDEGKFNYFTFGNTVHITMNHADGNSFKLSHSNRERNITITNGY-RFEGSFILS----------
--SGALHPPPHGQATVSWGYHDSA-TTMLVSIKPENAHE------KWP-EQCVIDSP-------TCLQTASDISTSQI--ISLQFVNLP---
TMHSLFSDFVMDE-GNMTTVKFTFP----------KLPWTTR------GICAVNSVTALSLDNSTTHNFNQDCYTIALADCSPSSRFIVGI-
KMITITLSVRILTQNSVMDMVQENSNIKVYINGTEIEVAAGKKWDEEMDLEISGITVQHYGHTVIMLVESKLRGLTCGLCGNFNSENSDE
LNET
>AGC97434_Amedicaginis
VSYLVQTSCNRDLFPEGVQLGYRWKATTHVGTVFPSEHLTNYSVEAYFLVQNDQNFTNFQIKEFKTSENQQ-----
ITLPFRCVYRDGEVQKFESESGDVAGSLNIKRASATMFQLKLDSLRR-SFATEETGVYGKCNVQYLVTKNNNTN-
VKKIINFSACNNKLGQQWSNTPPLCPSSYQDGSMSHSVRNYNLDEIN--VIRYLNIIGTVEFQPFQALAESHHIFVNQTFELDHIFPIGEM--
----
LVSKLDTMGFKPSNKSLLLAEILDTLGQLTESLSGTTNLDNQTEFRALELMWWLDTSDWFTLYDTISQETIQHFFWDLVPQVGSASSVE
FIRDLIKTQKITSFLATGLLITFPYHVRYPNEKLLKESEILLYLDRDLENEVRKVAILSFASLIHKTCGKGCSDDTQNKYIKLFLDKFIESTSH
SRRMLFIEAFSNMKIDNILDFLQPIITDHNR--
RLVAIWAAKSATTAHSDKVAEVFWPILTNHSEPLEIRTSSLNMLMMSQPTVSRFLTLYWFMQAEPSQQIYNFYYTTINSMANSKYPCYDK
YSKIAAQIVRFVQ-KSHHWATGNYLLDYEDPMRGYGGIVQTILMANQRTGFPSVLQLTAEQHSFGTS-SEFQIHIKAEGLADSLK--
KELFNLKHLQDVLEKAKFTIREDENLHFETIIKINQTIYCHHLNRTTFKSFALVLKRL-----

SSLFQFSLNYQRLTHFTDFGTPILLQFRTASLLSLRGSIDAELDFRYSMDTVTSLKTFNPLSSTWRGADRFRCIHARIPFSTEIITHFFRSYIK
ASAYRVDGSKL--GIVWH-----ATKLVPPENTLPVAKKESRDLGAKITVFQDEGKF----------
----

>XP015366382_Dnoxia
VSYLVQTSCNRDLFPEGVQLGYRWKATTHVGTVFPSEHLTNYSVEAYFLVQNDQNFTNFQIKEFKTSETQQ-----
ITLPFRCVYRDGEVQKFESESGDVTGSLNIKRALATMFQLKLDSLKR-SFVAQETGVYGKCNVQYLVTKNNNTN-
VKKIINFSACDNKLGQQWSNTPATCPSSYQDGSMSHSVRSYNLDEIN--VIQYLNVIGTVEFQPFQALAESHHIFLNQTFELDDIFPIGET--
----
LVGKLETMGFKPSNKSLLLAEILDTLGQLTESLSGTTNLDNQTEFRALELMWWLDTSDWTTLYDTISQETIQHFFWDLVPQVGSASSV
EFIRELIKTQKITSFLATGLLITFPYHVRYPNEKLLKESEILLYLDRDLENEVRKVAILSFASLIHKTCGKGCSDDTQNKYIKLFLDKFIESTS
HSRQMLFIEAFSNMKIDNILDFLQPIITDHNR--
RLVAIWAAKSATTAHPDKVAEVFWPILTNHSEPMEIRTSALNMLMMSQPTVSRFLTLYWFMQAEPSQQIYNFYYTTINSMANSKYPCYD
KYSKIAAQIVRFVQ-KWHHWATGNYLLDYEDPMRGYGGIIQTILMANQRTGFPSVLRLTAEQHSFGTS-SEFQIHIKAEGLADSLK--
KELFNFKNLQDVLEKAKFAIREDENLHFETIIKINQTIYCHHLNRTTFKNFAHVLKRL-----
SSLFQFSLNYQRLTYFTDFGTPILLQFRTASLLSLRGSIDAELDFRYSMDTVTSLKTFNPLSSTWRGADRFRCIHARIPFSTEIITHFFRSYVK
ASAYRVDGSKL--GIVWH-----
ATKLVPPENTLPVVKKESRDLGAKITVFQDEGKFNYFTFGNTMHITMNHADGNSFKLSHSNRERNITITNGY-RFEGSFILS----------
--SGALHPPPHGQATVSWGYHDSA-TTMFVSIKPENARE------EWP-EQCVIDSP-------TCLQTASDISTSQI--ISLQFVNLP---
TMHSLFSDFVMDE-GNMTTVKFTFP----------KLPWTTR------GICAVNSVTALSLDNSTTHNFNQDCYTIALADCSLSSRFIVGI-
KMNAITLSVRILTQNSVMDMVQENSDIKVYINGTEIEVATGKKWDEEMDLEISGITVQHYGHTVIMLVESKLRGFTCGLCGNFNSENSD
ELNET
>CAA39670_Celegans

FIALLGAAASESSFRAGREYRYLFNGQLSAGLPVPSTPQGISRLQSQVTLQWDGNTVRMQLQKTRFATSQQ-
SMFELSMPVEFDYEHGLVREIRFAENDQPWSENIKRAVINMLQVNILKKEAFSFTNVERTLEGECEVLYTVEEKEDKQRWAKSINLDKCT
RRIHHVQTP---
DCQQTLEQDKMSSTVLNYNITGTSSFLINSVELRSQHLFAPISEKHQLVSAFTLNTMELIYAGEKKTKQVRNEKTSELVYNQP----------
NKVEMIKKMFSLMAKQIE----
LAPHTVARIVKVLRECNEEQLEQIYRHVDEKQLRSIYFNTLRLAGTRVTIQQFVDKVSRKNIAPLKASVAIKTL-
VDMRYPSLAIAEDIARLCESDVSSFPALRQSCWLTYGAIVNGVCGQTCPRDAKQRIVDKLVQQFESASTRYEKVLALKTLANAGLDLSVY
PLEKII-------
RTQAIESFRRLRTQMPTKIQRVLMPVYLNRQQPQHIRMSALHQIIYTQPEWSVLSQIGNQLRQERDQQVRAFTLSLLRSYANNESPCEQ
TFSSRVQSLLNNIPSEIDRFESVY-
KWSTYSRRHQSGFEANFASLFTTESVLPTEMMASIEGVLGEWNQYFAQIGFTQKNMEKIIKKLSNVQETEFLSNLLEKLRITRRQSQDPH
AFVYIRHRMDYAFLPIDADSIPEVVLEIGDIERVLAQGI----------
VVPTPMGLPVQFTSKMPTISSIRGQVRVQAGPRVASTHVLSLRVICPIAEVGT--KFLHQAVLNTPVDTEIRMN--
EDKIRAIYNTQLSQSLDQGT---VR------IVDSNVEVYYKPN-VEQYEAIFNLY------QITKKVYQHVAKFEI--------------------------
MEVEACTQIELKATIQTPR-----TVEELKEQQLVVEMNYG-----ENTIHININGQQSQEQKKFVK---------QIEQ----AP--
HETLLEASRLDQYQTVVEYEFEPK--RYWNMVQAWTSR-
RKNMIRATINVEPQRLTVNMTILPTAQIYQPQKPVMEKIAHHASANCVVKSTKINTFDQVAYRNQFTPCYSVLAKDCGSEPRFVVLMKK
INEKEKNVKVVYGENEIEMYKTEEGLICRVNGEEIEYPESIWLNKNT-
LKFSDVTVQFDGVNARIHLSALYRNQQCGLCGHYDNEKETEFREM
>CAA26849_Celegans
IISLVLAIASDRTFSPKSEYVYKFDGLLLSGLPTRSSDASQTLISCRTRLQADDRYIHLQLTDIQYSASHIPKILSLELPFRAQIRNGLISEIQFS
SEDAEWSKNAKRSILNLFSLRKSAPVQVFFNVHEKTMEGDCRSRLHI--EGEKTIYTKSVNFDKCITRTAYGLRF---
ECKECEKEGQFVKPQTVYTYTFKN---LQESEVHSVYTLNVNGQEVVKSETRAK--VTFVEESKINRIKKVSGPKEEIVYSMP----------

IEMEQKVEQLQEIFRQIQEQNTPTVHLIARAVRMFRMCTIEELKKVHTTITKAKVQLVIETSIAVAGTKNTIQHLIHHFEKKSITPLRAAEL
LKSV-
QETLYPSEHIADLLIQLAQSPLEKYEPLRQSAWLAAGSVVRGFASKTASRQTKEKYVRVFMQHFRNADSTYEKVLALKTLGNAGIDLSVY
ELVQII-------
RTEAVDALRLLKDVMPRKIQKVLLPVYKNRQNKPELRMAALWRMMHTIPEEPVLAHIVSQMENESNQHVAAFTYHVLRQFYKSTNPCY
QQLAVRCSKILLFTRPEQML-STYS--LPLFNSEWLSGVQFDFATIFEKNAFLPKEVQASLETVFGNWNKYFAQVGFSQQNFEQVLEKL-
SLYGIQMLQEIVKKMNIRPRVQQNAHAVFYLRYKMDYIVLPIDMETIDTLVFDIKSLLTFLTNDS----------
EIPTTIGMPLTISGKMPTILSINGKVVLDIVPTVATTHVTEMPLLYPVIEQGV--
KSLQSARLHTPLRFESTVELKKNTLEITHKFMSTEEIDFGLRIN--NVSQWTLTEQDFEYTLENKETDLSEIKFKIF------
QYFHKGFKNLITLKL----------MYWNTELRTCVEMTLRTELLMPS-----SLRQLREQEVQLNAKWGSS---
KSEITVNAQLEQSTEQKKFIR---------NIEREYKGIP--YELLIKAARLNQVNVVSEYKLTPQ--
RIFDLIKATVSEKVQNRVVLQLSVEPSRQSHEHDH-------
LPRSAMFQQTWEKTGATCKVDQSEVSTFHNVIYRAPLTTCYSLVAKDCSEQPRFAVLAKKINKNELLVKVVRREE--
EIKKSDDKFLVKVDGKKVNPTELEILGDNLVIRLQ--EVRFDGYTVKTNMPSVSQNQLCGLCGNNDGERDNEFDRI
>CAA39669_Celegans
IISLVLALSSERTFEPKTDYHYKFDGLVLSGLPSASSELSQSRISARARIQADDRYIHLQLVNIRMAASHLPSLLSMELPLRAQLRNGLISELQ
FDKEDAEWSKNMKRAVVNMISFNPIAPRKSFFT-NEKTLEGDCQVAYTV--EQKKTIITKSINFDKCTERIAYGLRY---
ECPECEKDTELIRPQTVYTYVLEN---LKESEVRSLYTVNVNGQEVMKTETRSK--LVLEENHSIKSIKKVNGEKESIIYSSP----------
FPLDKKMHLIKTITEQIQENNMPTSHFLARLVRIFRTTSTSQLKEIHETLVKAKIQSLMEHALAIAGTKNTIQHILVHMENEDILPL--
GQILKTI-
QETPFPSQSIAEALIKFAESRVKNNLVVRQAAWLAAGSVVRGIVDYKDKRELKEKFLRVFMQQYKDAETTYEKILALKTIGNAGLDISVN
QLNEII-------
RKEAIDALRLLKDTMPRKIQKVLLPIYKNRQYEPEIRMLALWRMMHTRPEESLLVQVVSQMEKETNQQVAALTHQMIRHFAMSTNPCY

QRVAIVCSKVLSFTRPEQMIASSYV--
LPRFLQNSFSGAQFDFAAIFEKNSFLPKDLHASLDAVFGNWNKYFAQIGFSQQHMDKYLEKLESLEKIKLLKELAQKMNIRACPAKDAFA
MVYLRYKMDYAFLPIDRQLVENLIVQFSEIRRLLNQEL----------
IFPTTLGLPLTISGKIPTVISAEGQFTVEARPSVAATHVYEMRMFTPLFEQGV--KSVQSVRAYTPIKIQAVAGMKRN-
FEIVYKVQKTQEIELGLEVS--NINQHTLAEQDFEVSVENKETELSQIKYKIF------EYSTRGYKSVISLKL----------
YTMNTELTTCWEVTVRSQLLMPS-----SLRQLRDQEVQLTSTWGSQ---KSEVTVNAQLQQSKEQKKYER---------NMDRQFNGMP--
YELLIKAARLNQINAVAEYKLTRE--RYFDLVKTTVSSRENNRVVVQLTVEPSRQYVNITM-------
LPQRSVKHQFTEASGSVWKVKRNQIRTFDDVLYNTPLTTCYSLIAKDCSEEPTFAVLSKKTEKNEMIIKVIRGEQ--
EIQLQNEEIRVKVDGKKIQSEDYERLGESAVIELE--EVRFDGYTIKTKIPYYRKNQLCGLCGNNDDEVHQRVERL
>P18709_Xlaevis
IVALLLALGSRTHIEPSKISVYNYEAVILNGFPE----RAGIKINCKVEISAAQRSYFLKIQSPEIKEYNGV----
QTKPARFEYSNGRVGDIFVADDVSDTVANIYRGILNLLQVTIKKSQ-
DVYDLQESSVGGICHTRYVIQERGDQIRIIKSTDFNNCQDKKTIGLEL---
FCHSCKQLNRVIQGAATYTYKLKGGTVIMEVTARQVLQVTPFAERHGAATMESR--QVLAWVGSKSGPQIQLKNRGNLHYQFP----------
QAVEVLQHLVQDTQQHIR------EDAPAKLVQLLRASNFENLQALWKQF---
QYRRCLLDALPMAGTVDCLKFIKQLIHNEELTTQEAAVLI-
TFAMRSARPGQRNFQISADLVQDSKQKYSTVHKAAILAYGTMVRRYCDQL-------
EPLHELAAEAANKGHYEDIALALKALGNAGQPESIKRIQKFLPGFSSSAQTDAVMALRNIAKEDPRKVQEILLQIFMDRDVRTEVRMMAC
LALFETRPGLATVTAIANVAARETNLQLASFTFSQMKALSKSSVPHLEPLAAACSVALKILNLNLGYRYSKVMRVDTFKYNLMAGAAAKV
FIMNSANTMFPVFILAKFREYTSLVENDDIEIGIRGEGIEEFLRKQ-----
RKKISQIVKSLGFKGLPSQVPLISGYIKLFQEIAFTELNKEVIQNTITMIRNVLNKLLNGV----------
RIPTTVGLPAELSLYQSAIVHAAVNSNGEVKPSVLVHTVATMGINSPLFQAGI--
EFHGKVHAHLPAKFTAFLDMKDRNFKIETPPTSAEGASAGVCIQMKPATEGAIEKLQLEITAGPKEMDETAVTKKM-

DESRKDSYSKRSSSQEQNKKRHGQKSSSSSSSSSSSSSSSSSSSSSSSSEENRH---
NKQAKMQKKNFSESSSSSSSSSSNKKKDLNFGSRLSSPPVLVVTFKAVRNDNTKQGYQ-
YQEYHSSKQIQAVMDISKACAQASLKGQNCQDYKINANWPKIPSKWVVGEYVPG-----
KRNQRQVKLVFALSPRTCDVVIHAKENVQTPTWNIFAEAPKLIMGECKVAQDQITTFNGVDLASAPENCYNVLAQDCSPEMKFMVLM
RNSKESNKDINVKLGEYDIDMYYSADAFKMKINNLEV--SEE-------
SLSAEIDSLDYDGLTFKFRPTIWMKGKTCGICGHNDDESEKELKHT
>P87498_Ggallus
LIALVLTLGSHLNYQPNKVYTYNYESILFSGIPE----RTGIRIRSEVEISGGPKLCLIRIHSIEAAEYNGI----
QSIPIKFEYSNGHVGNLMAPDSVSDDGLNIYRGILNILELSLKKMQ-
HSYSIQEAGIGGICNTTYAIQEKANLVDVTKSKDLNSCEEKVVTGSAY---
PCQTCQQRNKNSRATATYNYKIKYEAVITQAEVEEVHQFTPFHEITGGNAIVEAQKLALIEVQKQVAPPKEFQKRGSLQYQFP----------
QIEERLQDLVETTYEQLP------SDAPAKLMHLLRAANEENYESVWKQF---
AYRRYLLDLLPAAASHRSLRFLRHKMERQELTNWEIAQTV-LVALHSSSPTQEVMEEATLIVKKHCRSSSVLRKVCLLSYASLCHKRCSSP-
------
QVFHVFAGEALGKSNIEEVLLALKALGNVGHPASIKHIKKFLPGYAAGAHETAVMALKSIGMRDPQMVQAITLEIFLNHKIHPRIRMLAA
VVLLETKPGLPILMILVDAVLKEPSMQVASFIYSHLRALGRSTAPDLQMMASACRMAVRALSFRSGYQFSKVFRFSMFKEFLMSGLAAKY
FVLNNAGSLIPTMAVSQLRTHFLGRVADPIEVGIAAEGLQEMVRGY-----
NYDFREILKKLDWKALPRDKPFASGYLKMFQELLFGRLDKDTLQNVLPSIRRLISSLQTGI----------
RVPTCVGFPMETSFYYSSVTKVAGNVRSKMSLSMAKHMTFVIGINTNMIQAGL--
EAHTKVNAHVPVNVVATIQMKEKSIKAEIPPQSVEDVSFGVCIEFKPVSDVPIEKIQVTIQAGDQESSRKEVMKKI-
DNQATRSKSKSSNSKDSSSSSEGERQETQSSSSSSSRASSNSRSSSSTSSSSESWKQTKRVKSQTHWSGHRTSSSSSSSSNSSSDSSSGEKAAH
PPGITIVAQAVRSDNRNQGYQ-YVRSDAAKDVQLVVQLAEACAQARMRGKECRDYRIAVQWGIIPSWIALMRYVPG-----
QRNSRELIVRAVASPRSIDTVI---------AWNFLPEIASQIASTCEVSKGDFKTFDRMSFTCSNKSC-

VVVAQDCTEHPKFIITTRKVDHQSREVHINTSSANITICPAADSLLVTCNKESVLSSEY-------
IVEAILKNVNFDGEILKVTVASWMRGKTCGVCGNNDREKHNELQRR
>ABB89953_Pmonodon
FIAFV---SPKLAYQPDKTYAPPYSGKSRVHLKGVDNGDSEIEWTAGVDLTWSPCDMAISFRNTKMDGARGP----
LRHPLVVAVVDGRVQHVCAHPEDEPWAINMKKGVASAFQNSIPSLSVSGITVTETDVVGKCPTKYEIKTEGEKVIVVKEKNHRHCQER--
-YPTPAETAPWLKAPLPIEESKSQCRQEIANG--IYTAITCQDKNIVRP----AIGIYKYVEASISESSDTSAISHSGEVHIES-LLYNHP-----
APELDELMKEICAK--------TEAAAL----VAKALHVLRRVPETIVVETAQKVSDSRLESIFLDAVAFLHESGAVKVMVHEIENGRATGGR-
LALYMAAVYLTPRPNIEAVKALTPLFE-----ALKPSPSLLLAAASMVNNYCRHTCYEE--ATRVQSHCSPSAGAEDKEVALAIFKAVGNMG--
----VITPAVTRAAVACRVAAAHVFRQ--TKCYRPAVEKLVSIAVRPAFGTEVRIASY-
LAAIRCAEMEDLEEIFEKISVEENTQVRGFILGHLLNIQESTCPTKEHLRYLLTNFVIPIDERDFRKFSRNIDVAYHAPAFGMGAGLESNIIY
APGSFIPRAVNLKMKADVDEMHMDIAEIGARFEGIDSIIEELQGYLHMEDITGFAGEKGLKVMEHSHVHADLFARFMHEI-
TYADVAESLKGVTTLIETFFSFFENSLHMKDLNLNTAS-PDIQGTPLRLKLAGTAVAGVKMEGRVGIEMQNTISSATGAAIKIRTTENK---
--KIEMELEIP-ERMELLNIKAE-TYLVKAVKKITKISFGLKVCMRGYVTAAI-NKVIKMNVEHLSYPKEG------
DSSSIAKFIKKNLNIEVGADL----KFRILFPI-----------LRLTKGR---FPE-LQANIK---------LSALASENVVL----
GLEWKMKTTVGIAFILKTSSR-----REFSWD--ADDASLSSPANPE-YHAKLVTAANLVEHMEIALEKLPEEH--------ESYETGP-
TYTMQMRTPTRTMEGRAKLVFPQGLRSPRYPDTAEAGVYAA-
LKAVVMEVVEFPSSQLVSLVIGREMTELIADILDSPAVSFYTERAFAAGIAVQILIFAYVEYILPPYPRTATVVGGTEILTSGHAGSQQTFNV
RINKVVVSKAGEAIAVEASGWVFGRVAGLLGPNNGEIADDLAQA
>Q9U943_Lmigratoria
LIAIS---SSVFQYQKGQTYTYSFEGTTLTSLPGTQGEPVRLKLKATADLSVDDCNKVLRLRGVTVSGPDSK----
LAHPVLANFKGSSIKQLCSEDGDNQSSLNIKRAILSLLQTP----NKS--
TASEVDVFGICPTNVRHSQRGDVTVISKTRNLNRYASRIQETLSTRFTQSDLHA-TPFLDADLHVEQQIKGG--
LIVSATSRESYLFRPFSNQGNGAKTIVETKLTLTSQNAQPALASFTVPKS-IVFE-P-----LASVPGGSSAITAA-LHAAESSTDAAEK----

FRTLVSVLRQSSTTDILKVYNDVSNKSARNLLLDALFRTSTGDAVEVIARLLKTKEITANH----WYLSLAFIQHASLKSVVSISSLLD-----
QKNLPTEAFLGIGSFIGRYCREHCENV--AKHLSG--STTSKA-GENRAIAALKALGNIR------HLNNALGEKVKQLRVAALEVIQS--
DPCRKNIKQAALQILRDQVEDSELRIKAY-
LAVVECPCDNVVKTISNLLENEPIIQVGSFVVSHLKNLQASTDPSKAEAKEKLGQLKPKKISSDIRKYSQNYELSYAIDAINAGASVESNVIF
SQSSYLPRSVSLNLTADVFGHSYNVFEIAARTENLDHIIESFKGYIEVDEVEEKTKSLYNRITERRSLDVDLSLKTFSELAWFNYDGKH-
EQKSRVVDEIFDAIDEGLKSKKFN-YDFEYPTNLGFPLKLAIDGSIAARLKLNGEGGLKLDYNVHSSTG--INVAVHNLN---
DLGIDIKVGLPVKKQDIIDVKTDVLTTVKERHPETSTPSGLTFCVSEYFRAEI-
FKSAAVLFDTPEKPHQGITAKSIDNNEKSHFIKGDYELKFSAKIPKDDKYEFQVDVKLAAEDEYHTSYKLNRGSGNSFKYLTQADTKLHL
AEGRESAS--STQDLKLFLSGGYSPVATLELLLPFDKKLTGEAAGDEKETTWS----DIEF-SDNQP--
IDFELTTADQKKTDVKIELQWVPKFNSLHNILKVEQAS----GSGFTLTLPSRVIAAEVKLQFEV---STVLNKLKEISDEAG-
VSTLVNGILKTLNKKSLPALIAKSLTQSLVNILSSAAEILYLKKDADVEKAIDATVLPSFIQLIPPYDATAILLNSHHFFTDRRHAWREYNRV
GITKVTCETSIELCTFEINGFYFGKTRGLLGTINNEPWDDFAAA
>Q25490_Msexta
RLSVLACKSSSPSFAAGQKYNYGVEGTVSVYLTGADNQETSLKMLGQASVSASNCELELSVHNMVLSGPDGK----
IEKPVRFSYQDGRVPEICAAEDDSRRSLNIKRAIISLLQAE----QP-----
VQVDVFGVCPTEVSSSQEGGAVLLHRSRDLSRCAHRRNDFVNSIANDAGIKD-LQVLQSMLNVESKVNNG--
VPEKVSAIEEYLYKPFSVGENGARAKVHTKLTLSG-KGGAGNAHCTESRS-IIFD-P-----TSSASGNLNSVISAVKETARTVAKSAGQ----
FAQLVRIMRTSSKDDLMRIYSQV--HLEKRVYLDALLRAGTGESIEASIQILKSKDLSQLE-QHLVFLSLGNARHVNNPALKAAAGLLD-----
MPNLPKEVYLGAGALGGAYCREHCHNVKPESKLQN-CRPKNKP-DEDVVVAILKGIRNIR------HLEDSLIDKLVHCRAVALEAFHA--
DPCSAKIHKTAMDIMKNRQLDSEIRIKAY-
LAVIECPCSHSASEIKNLLDSEPVHQVGNFITSSLRHIRSSSNPDKQLAKKHYGQIRTPNKKVDERKYSFYREMSYKLDALGAGGSVDQTV
IYSQTSFLPRSVNFNLTVDLFGQSYNVMELGGRQGNLDRVVEHFKSFLRYDNLVKRFQESKKKVEDSNELDLDIYVKLFTDA-
VFLSFGDD-KGFDKMLDQILGGCNSGIKAKHFQ-QEIEYPTSVGLPLRLNLIGAATARLDVATSTGIKVITNLHSSTGVHVNAKVLENG---

-RGIDLQIGLPVDKQELIAASSDLVFVTAEKQKEKQKVLGLTMCLRGYIKAVVGRRSFELLLDTEEGVKLGLEAKV-
KPNELVQILKYDYSIKVKANLNAEKRYELS-----GFVLHKTKPLKIMRTA-NNYKV--DGAFK-DGYNGHMDMTI-
SDANLQSTFSLGFELAGTLTVRLPFEKKIDANGKVGPKDYNYK--SSD----SSYAPT-YSNSVVDADGRVYKIDVKVDNIKELGYVAES-
RVEVSDRPPRASGVTVRLPSRVMALEYTLMFET---TPLEDRIKNVGEKAS-IAKFIDEILPTLSKQIEPQV-
MKEYLDGVIDVVAHFAAIVYLEKMEESALAVRRASWSALRQLVPN-
KLRAVVVNGQHIFTDGRFAFRQTSRIGLSMAFCTSKFEVCYFEVNGFYLGKLPGLLGDGNNEPYDDFARG
>XP321226_Agambiae
SLAILSSNSSKYTFKPGFVYQYDVDSYVQLQQSDKENKQTTLKVDGKVEVYADNCQYTLKVVSLTSYAPDGK----
ISKPVQFTLSNDELPEICTEADDTDFSLNVKRGLISLFQVA----QK-----
TETDVFGVCQTSFSSYPSGDATVVEKVRDLGNCAYRSNSFVTRIVNKAGIKS-TPLLQSSYNAQQTIKGG--
LLSAVKLSEEYQYLPYLKDKVGVTAKVTTKLTLTGNKAGAALGAASEPRT-IIFE-P-------QQPAGNLAVIKQLKSTVESYTKTANL----
FVELIQLMRYSKKEDLLTLYNQVSNKLARKVYLDALFRVGTGDAVEAITQLYKNKELTGAQEQKLAFVSLTLVQSMTQDALKAVNKLL
D------GNPPREAYLSVGSLVSKYCQKHCQS---SAKLGK-CQSTSRA-QEDVIVAVLKGVRNSD------NLVAPLLDKVIQCRVAALQAYPA--
ASCNKKIVNAALSTLKDTNEDSEIRIHAY-
LSLVECPSANVANELKALLDAEKVYQVGSFITSHLASLRASVDPTRDAARQHFGKIRTSNKPFDVRRYSFNREFSYAVESLGVGASAETSVI
YSQKSFLPKSVGLNFTAELFGNGLNVFELEGRQDNLERLVEHYKGFFSYDLLAEQYQKLSGKAKERKDFNLDVTMKVFSEL-
FFLSTGEN-VPTDQFLDKALECFDKMIGAKKFE-
HTFDYTTALGLPLKLSAQGAGVARVDAALMTGMQVAVSGHSSGGAAVKFALHDATA-YDLTVDALQG----
KQELVSVNFREVIITRERNQLITLPIGVTVC-PKFFSGKF-HLQLSLSFDTPE----GVTA--VNNNDKEQYLKPDYDVQLGGSL-ANKKFELN-
----GAVTNRFTSFKLERGQ-DNLKS--DASTK-DERSGKMAASVYDKQDVVLDVTLGHGKPA-
VDVGAPSYKSVSTQAKGNYNEGSYTYDAKNFEFGTAKAPDIYNTKLQTADDKKFSVEAKVQNLPMLNVLHES-RVHTTD----
SAGASLILPTRTVAVEGVLTFGSYFGSPRYAKLAE------KMSALVNDV-
ETFNTKVLPALVYEETVKLVVAVFERTVKALYIKPVNDVNQAVERTSFVTFKFLVPMFHMHAQLGDGGHFFTDDKYVFRKYYTVTVTQ

VMCTTDLKICHFFVSGFYFGRLRGLLGNGNYEPYDDLARL
>Q9V496_Dmelanogaster
LIVLLIAVAELLKYIPGNYYDYSFDSILTIGASSVPNDDTSLKVSGSAKIFAGNCGYTLQLSSVKVTNTKES-----QKPVQFTL-
SGILPQICSDSSDLDYSLNIKRAVVSLLQSGIEAEH-------EVDVFGMCPTHTSTSK--NANIITKARNLNSCSHRINSGLVS---
KAGITLQANYIKESRIV-------NHLIENVQLTETYKFIGNTKRNSDISAKVV--LKLKNPSGTKASPGTGSTVRSLIF--P----------
NALKTLSDLVDSTGDYVK------KETAKKFIRLLRQSDSETLLELAAFP---LARKVYLDGLFRTSTAESARVILKQLEKEKLLAILSLNIVKSV-
-----DKETLNQAASQLLPAP--------KELYIAVGNLVAKYCLKNCQGPEIDKKFSDGLKHCKTKREEERIVYILKGLGNA----------
KSLSGNTVAARVAALHAFSK--VKCEETLQSKSLELLKNRNEDSELRIEAY-
LSAISCPNAEVANQISEIVNSETVNQVGGFISSNLKAIRDSTDVSRDQQKYHLANIRVTKTPDYRRYSFNN-
EVSYKLESLGVGASTDYQIIYSQHGFLPRSSRINVTTEFFGTNYNVFEASVRQENVEDVLEYYKGLVNVDDVSKISKKYKMYGVKNQDLN
LDVSLKLFSELAFLSLGDNIPSSLDDIINYFSTSFEKKQFEKQFSSHHLDYPTSIGVPLELVAQGFAATKVDLAVSTGLRVVSSAHSATGSDI
TAVISD-----
GEGFNVDLELPINFNVDTDLKNKKLNICIKSNKPKTFSGGIQYEINGYADGKINFENFQVSNSNNNSLTSELRIKLYDQSYALQVIKIDYSL
EASAKINGNVTSWGTLLSAKGEIGQRYSAIKLTSESQHQFRVNIDAQLSSISIGKLDRLNTKTKENSNLEISYGNYNNNQANIEIPKSTAQT
NNMESPLKGYHRKQTGNDLTQGAQYVKKVYVVTALKQTDKSYGLDASYVSISKKLNGLGTLKIYHLDSGNNSSTIKLALPKRQILFETKI
QFQYQISNAKVQILDKNPIEITGTAEVKGKEFQLFNQTNYFINLNGFWHIVESEINAGPDFSKIVKPINDANDKVKVHIDNEYIELYPGSSD
RGHVVDGKHIFTDGLHAWRRFYHLYSYGIVCTSDLKVCHININGFYTSKTRGLLGNGNAEPYDDFAKV