THE UNIVERSITY OF RHODE ISLAND

University of Rhode Island DigitalCommons@URI

Biological Sciences Faculty Publications

Biological Sciences

2019

Genome-wide annotation and comparative analysis of cuticular protein genes in the noctuid pest *Spodoptera litura*

Jianqiu Liu

Shenglong Li

See next page for additional authors

Follow this and additional works at: https://digitalcommons.uri.edu/bio_facpubs

The University of Rhode Island Faculty have made this article openly available. Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

Terms of Use

This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our Terms of Use.

Citation/Publisher Attribution

Liu, J., Li, S., Li, W., Peng, L., Chen, Z., Xiao, Y., Guo, H.,...Mita, K. (2019). Genome-wide annotation and comparative analysis of cuticular protein genes in the noctuid pest *Spodoptera litura*. *Insect Biochemistry and Molecular Biology*, *110*, 90-97. doi: 10.1016/j.ibmb.2019.04.012

Available at: https://doi.org/10.1016/j.ibmb.2019.04.012

This Article is brought to you for free and open access by the Biological Sciences at DigitalCommons@URI. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Authors

Jianqiu Liu, Shenglong Li, Wanshun Li, Li Peng, Zhiwei Chen, Yingdan Xiao, Huizhen Guo, Jiwei Zhng, Tingcai Cheng, Marian R. Goldsmith, Kallare P. Arunkumar, Qingyou Xia, and Kazuei Mita

Highlights

- In total, 287 cuticular protein genes are annotated in *Spodoptera litura* genome.
- Most RR-1 and His-rich RR-2 CPs form a major cluster in S. litura chromosome 9
- Histidine-rich RR-2 CPs are greatly expanded in *S. litura* compared to *Bombyx mori*.
- RR-1 CP genes are mostly conserved among *S. litura*, *B. mori*, and *Manduca sexta*.
- His-rich RR-2 CP genes form species-specific clades in phylogenetic analysis.

Major histidine-rich CP expansion in S. litura

Graphical Abstract



B.mori



1	Genome-wide annotation and comparative analysis of cuticular protein genes in
2	the noctuid pest Spodoptera litura
3	Jianqiu Liu ^{1,2} , Shenglong Li ^{1,2} , Wanshun Li ^{1,2} , Li Peng ¹ , Zhiwei Chen ^{1,2} Yingdan
4	Xiao ^{1,2} , Huizhen Guo ^{1,2} , Jiwei Zhang ^{1,2} , Tingcai Cheng ^{1,2} , Marian R. Goldsmith ^{1, 2, 3} ,
5	Kallare P. Arunkumar ^{1, 2, 4} , Qingyou Xia ^{1,2} , Kazuei Mita ^{1,2}
6	¹ State Key Laboratory of Silkworm Genome Biology, Biological Science Research
7	Center, Southwest University, Chongqing 400716, China.
8	² Chongqing Key Laboratory of Sericultural Science, Chongqing Engineering and
9	Technology Research Center for Novel Silk Materials, Southwest University,
10	Chongqing 400716, China.
11	³ University of Rhode Island, Kingston 02881, USA
12	⁴ Central Muga Eri Research and Training Institute, (CMER&TI), Central Silk Board,
13	Lahdoigarh, Jorhat 785700, India
14	Corresponding author
15	State Key Laboratory of Silkworm Genome Biology, Biological Science Research
16	Center, Southwest University, Chongqing 400715, China. E-mail:
17	mitakazuei@gmail.com
18	Keywords:
19	Cuticular protein
20	Comparative analysis

- 21 Histidine-rich
- 22 RNA-Seq

23 Spodoptera litura

24 Abstract

25 Insect cuticle is considered an adaptable and versatile building material with roles in 26 the construction and function of exoskeleton. Its physical properties are varied, as the 27 biological requirements differ among diverse structures and change during the life 28 cycle of the insect. Although the bulk of cuticle consists basically of cuticular proteins 29 (CPs) associated with chitin, the degree of cuticular sclerotization is an important 30 factor in determining its physical properties. Spodoptera litura, the tobacco cutworm, is 31 an important agricultural pest in Asia. Compared to the domestic silkworm, Bombyx mori, another lepidopteran whose CP genes have been well annotated, S. litura has a 32 shorter life cycle, hides in soil during daytime beginning in the 5th instar and is exposed 33 34 to soil in the pupal stage without the protection of a cocoon. In order to understand 35 how the CP genes may have been adapted to support the characteristic life style of S. 36 *litura*, we searched its genome and found 287 putative cuticular proteins that can be 37 classified into 9 CP families (CPR with three groups (RR-1, RR-2, RR-3), CPAP1, CPAP3, CPF, CPFL, CPT, CPG, CPCFC and CPLCA), and a collection of unclassified 38 39 CPs named CPH. There were also 112 cuticular proteins enriched in Histidine residues 40 with content varying from 6% to 30%, comprising many more His-rich cuticular

41	proteins than B. mori. A phylogenetic analysis between S. litura, M. sexta and B. mori
42	uncovered large expansions of RR-1 and RR-2 CPs, forming large gene clusters in
43	different regions of S. litura chromosome 9. We used RNA-seq analysis to document
44	the expression profiles of CPs in different developmental stages and tissues of S. litura.
45	The comparative genomic analysis of CPs between S. litura and B. mori integrated
46	with the unique behavior and life cycle of the two species offers new insights into their
47	contrasting ecological adaptations.

48 **1. Introduction**

49 Insect cuticle must provide an effective barrier from the natural environment. Consequently, its physical properties, such as thickness, stiffness, strength, elasticity 50 51 and color, show large variations at different metamorphic stages and in different 52 anatomical regions. Cuticles have a common fundamental structure, consisting of a 53 procuticle composed of a filamentous chitin structure within a protein matrix covered 54 by an epicuticle consisting of lipids and protein above which there is a dense cuticulin 55 lamina (Locke, 2001). The variation in physical properties of cuticle is partly due to 56 different degrees of cross-linking and hardening occurring during the process of 57 sclerotization, whereby phenolic material is incorporated into the CPs and/or other 58 cross-links are formed. Additionally, numerous CPs are identified in all insect species 59 studied, and their number and features are quite different among diverse species (Willis 60 et al., 2012).

61	Information about the CPs of various insects and their underlying genes has been
62	obtained at the transcriptome and protein levels in the past few decades (Andersen,
63	2000; Baton et al., 2009; Dittmer et al., 2015; Dong et al., 2016; Futahashi. et al., 2008;
64	Gu and Willis, 2003; He et al., 2007; Pan et al., 2018). With the improvement of
65	sequencing technology, more and more genomic information for insect CPs has
66	become available, with many thousands of CP coding genes accumulated in sequence
67	databases. Many CPs have been identified by their conserved protein sequence motifs.
68	Andersen et al. (1995), Willis (2010) and Willis et al. (2012) have defined several CP
69	families. The most abundant family of CPs contains the Rebers and Riddiford
70	Consensus (R&R Consensus), which in an extended form has been shown to bind
71	chitin (Dong et al., 2016; Rebers and Willis, 2001; Tang et al., 2010; Togawa et al.,
72	2007; Togawa et al., 2004). Three distinct forms of this consensus have been classified
73	as RR-1, RR-2 and RR-3 (Andersen, 1998; Andersen, 2000). The other families with
74	conserved motifs are CPs with a 44 amino acid motif (CPF), CPF-like in the conserved
75	C-terminal region (CPFL), the Tweedle motif (CPT), alanine-rich CPs of low
76	complexity (CPLCA), CPs of low complexity with two invariant glycine residues in
77	the conserved domain (CPLCG), CPs of low complexity with an invariant tryptophan
78	in the conserved domain (CPLCW), proline-rich CPs of low complexity (CPLCP), CPs
79	with well-conserved cysteine residues (CPCFC), a glycine-rich CP (CPG), and analogs
80	to peritrophins (CPAP1 and CPAP3).

l

81	Spodoptera litura (Lepidoptera, Noctuidae) is an important agricultural pest
82	distributed in the tropical and subtropical areas of Asia. Compared with the domestic
83	silkworm, Bombyx mori, S. litura has a shorter life cycle, although it has one more
84	larval instar (Fig. 1). In S. litura, instars L2-L5 are shorter than L1, whereas in B. mori,
85	larval stages following L1 take increasingly longer times. Moreover, from the 5th
86	instar S. litura hides in the soil during the daytime, comes out from the soil in the
87	evening to eat crops throughout the night and then goes back into the soil with
88	daybreak. Further, instead of residing in a protective cocoon from pupation until adult
89	eclosion like <i>B. mori</i> , <i>S. litura</i> stays in the soil from the wandering stage until the moth
90	emerges. This exposure to soil requires CPs to form a more protective cuticle for
91	protection against abrasion and fungal and bacterial infection compared with <i>B. mori</i> . It
92	is also informative to compare S. litura with M. sexta, a close relative of B. mori that
93	stays underground during pupal and pharate adult stages.
94	In this study, we annotated <i>CP</i> genes based on the recently published complete
95	genome sequence of S. litura (Cheng et al., 2017). Combined with transcriptome
96	analysis, we then estimated what different kinds of CPs contribute to different types of
97	cuticles among various tissues and in different metamorphic stages. In addition,
98	comparative genomics and phylogenetic analysis among Lepidoptera provided
99	information on how CP genes evolved to adapt to the different ecological niches of
100	each of these three species.

ļ

101

102 **2. Materials and Methods**

103 **2.1 Annotation of cuticular protein genes**

104	To predict	putative CP	genes of S. litura.	reported sets of le	pidopteran CPs were
-			/ · · · · · · · · · · · · · · · · · · ·		

- 105 collected from KAIKObase (http://sgp.dna.affrc.go.jp/KAIKObase/) and the NCBI
- 106 Reference Sequence database (https://www.ncbi.nlm.nih.gov/genbank/). CP genes
- 107 were predicted for the *S. litura* genome assembly (Cheng *et al.*, 2017) using
- 108 TBLASTN (E-value $<10^{-5}$) and BLASTP in the non- redundant GenBank database.
- 109 Predicted *CP* genes were further examined by HMMER3 search (cutoff E-value =
- 110 0.001) using the Pfam database to confirm conserved domains and subsequently
- 111 classified into 9 families based on conserved motifs with the help of an online tool
- 112 (http://bioinformatics.biol.uoa.gr/CutProtFam-Pred/) (Ioannidou et al., 2014). The
- annotation sequences were deposited at GenBank under BioProject accession
- 114 PRJNA344815.

115 **2.2 Phylogenetic tree construction**

116 A total of 199 RR-1 and 361 RR-2 CDS sequences from S. litura, B. mori, and M. sexta

- 117 were aligned using ClustalW in MEGA6 (Tamura et al., 2013). The tree was
- 118 constructed using the Maximum Likelihood method based on the Jukes-Cantor model



- 139 RNA-seq data are deposited in the Sequence Read Archive database under SRA
- 140 accession: PRJNA498147 (https://www.ncbi.nlm.nih.gov/sra/PRJNA498147). The
- 141 Log₂ (FPKM+0.01) (fragments per kilobase per million fragments mapped) value was
- 142 used for making the heat maps.

143 **3. Results and Discussion**

144 **3.1 Annotation of cuticular proteins**

- 145 As shown in Table 1, 287 putative CP coding genes (Supplementary Table 1) were
- 146 predicted and classified into eight CP families: CPR subdivided into three groups,
- 147 (RR-1, RR-2 and RR-3), Tweedle, CPF, CPFL, CPLCA, CPCFC, CPAP1, CPAP3 and
- 148 Glycine-rich (CPG). In addition, a collection of unclassified CPs was named CPH
- 149 following Futahashi et al. (2008).
- 150 Among the 3 R&R consensus groups used to subdivide the CPR family in *S. litura*,
- 151 63 RR-1 protein genes and 129 RR-2 protein genes were identified with
- 152 CutProtFam-Pred, and one RR-3 protein gene was identified by manual annotation
- 153 based on its similarity to RR-3 proteins in *B. mori*.
- 154 CPAP1 and CPAP3 families, which contain one or three peritrophin A-type chitin-
- 155 binding domains, are expressed only in cuticle-forming tissues (Jasrapuria et al., 2010).
- 156 Thirteen *CPAP1s* and 9 *CPAP3s* were annotated in the *S. litura* genome (Table 1)
- 157 (Tetreau et al., 2015).

158	The CPLCA family is defined by the presence of the retinin domain (pfam04527)
159	and richness in alanine residues, varying from 13-26% (Cornman and Willis, 2009).
160	Four CPLCA genes (Sl_Aki-270, 271, 272 and 273) which formed a gene cluster in S.
161	litura chromosome 28 (Chr28) were annotated. Two CPs, named BmorCPH6 and
162	BmorCPH7 by Futahashi et al. (2008) in B. mori, were related to CPLCA by homology
163	search. This is the first time CPLCAs, originally described in Anopheles gambiae
164	(Cornman and Willis, 2009; Willis et al., 2012), were found in Lepidoptera.
165	CPF has a conserved region with 44 amino acids (Togawa et al., 2007). One putative
166	CP gene belonging to the CPF family in the S. litura genome (SWUSI0111680) was
167	identified which we named Sl_Aki-CPF. Seven CPFL genes, which have a conserved
168	C-terminal region similar to CPF (Togawa et al., 2007), were found. All of the CPFL
169	genes together with two other CPG genes (Sl_Aki-235 and 242) formed a cluster on
170	Chr25. This genomic structure was similar to <i>B. mori</i> , in which three CPFL genes
171	(BmorCPFL2, BmorCPFL3 and BmorCPFL4) and two CPG genes (BmorCPG23 and
172	<i>BmorCPG42</i>) form a gene cluster on Chr23. In addition, <i>Sl_Aki-235</i> and <i>Sl_Aki-242</i>
173	were orthologous to BmorCPG23 and BmorCPG42, respectively.
174	We identified five genes with a Tweedle motif, among which Sl_Aki-292, 293 and
175	294 formed a cluster on ChrZ; the other two CPT genes were not linked to other CPs.
176	The CPG family contains GG repeats but does not have any definite motif described
177	in other CP families (Futahashi. et al., 2008; Willis et al., 2012). However, although

178	some glycine-rich cuticular proteins also had an R&R Consensus or Tweedle motif
179	(i.e., <i>Sl_Aki-008</i> (RR-1), <i>Sl_Aki-185</i> (RR-2), and <i>Sl_Aki-34</i> (CPT)), we left them in
180	their well defined families. In S. litura, 28 putative CP genes were classified as CPGs.
181	Futahashi et al. (2008) classified a group of 34 proteins as CPH which stands for
182	cuticular protein, hypothetical, among which we assigned BmorCPH6 and 7 to CPLCA
183	and BmorCPH1 to CPCFC. They all have signal peptides and some sequence similarity
184	with known CP genes or with the AAP (A/V) motif often found in CPs (Futahashi. et
185	al., 2008; Magkrioti et al., 2004). Twenty-six putative CP genes which showed
186	sequence similarity with 31 already known CPH proteins in <i>B. mori</i> were classified as
187	CPH in S. litura. Most of them had signal peptides; additionally, the AAP (A/V) motif
188	was found in 20 CPH proteins. Only 6 of these had sequence similarity with already
189	known CPH proteins in <i>B. mori</i> .
190	3.2 Characterization of histidine–rich CPs

191 The amino acid composition of the 287 CPs revealed 112 His-rich CPs in which

192 histidine residues ranged from 6.00–30.14% (Fig. 2). Interestingly, most His-rich CPs

193 belonged to the RR-2 family, which is common to both *B. mori* and *S. litura*. *B. mori*

194 has one extremely high His-rich CP, BmorCPR152, of 45%, while the orthologous CP

- 195 with the highest His content in *S. litura* is Sl_Aki-166 with 30.14%. Figure 3 shows a
- 196 large expansion of His-rich CPs of the RR-2 family in the S. litura genome. As
- 197 Andersen and Roepstorff (2007) stated, histidine and several other amino acid residues

- 198 in CPs could be involved in cuticular adduct formation based on biochemical analysis
- 199 of the cuticular hydrolysates from different cuticular types of *M. sexta*, desert locust
- 200 Schistocerca gregaria, and yellow mealworm, Tenebrio molitor (Andersen, 2007;
- 201 Andersen and Roepstorff, 2007; James L.Kerwin, 1999; Kerwin, 1999; SO Andersen
- 202 1997). EM immunolocalization utilized by Vannini and Willis (2017) support the
- 203 hypothesis about the deployment of RR-1 and RR-2 localization first proposed by
- Andersen (1998). Namely, the RR-1 CPs are mostly found in soft cuticle like
- 205 inter-segmental membranes, whereas RR-2s are restricted to hard cuticles in A.
- 206 *gambiae* (Vannini and Willis, 2017). Histidino-β-dopamine is the dominating adduct in
- 207 hard cuticles like those of adult beetle cuticle, lepidopteran pupae and dipteran puparia
- 208 (Andersen, 2008). Consistent with these reports is that His-rich CPs were mostly found
- 209 in RR-2, but not RR-1 CPs of S. litura.

210 **3.3 Major clusters of** *CPR* **genes in** *S. litura*

- 211 Compared to *B. mori* which has a cluster of CPR genes on chromosome 22, major
- 212 expansions of *S. litura CPR* genes derived from this family were on *S. litura* Chr9.
- 213 Figure 3 shows that 34 RR-1 (Sl-Aki-48 80) genes and 82 RR-2 (Sl-Aki-86, 88 94,
- 214 96-169) genes were present as two large clusters located on different regions of Chr9.
- 215 Intriguingly, all of the RR-2 CP members belonged to the cluster encoding His-rich
- 216 CPs, whereas none of the 34 RR-1 CPs in the large cluster on Chr9 were His-rich.
- 217 Chr22 in *B. mori* also has the orthologs of these RR-1 genes, but none of them were
- found to be His-rich (Fig. 3). Although much smaller than S. litura, there was a

- 219 separate His-rich RR-2 CP cluster (BmorCPR79-129) on Chr22 in B. mori. Thirteen
- 220 RR-1 (Sl-Aki-001 013) genes were also localized on S. litura Chr1 as a cluster
- 221 (Supplementary Fig. 1). Their orthologs in *B. mori* also formed a cluster on Chr9.
- 222 3.4 Phylogenetic analysis of S. litura, B. mori and M. sexta CPs
- 223 To compare the S. litura CPs with B. mori and M. sexta CPs, phylogenetic trees of
- 224 RR-1 and RR-2 CPs were constructed separately (Fig. 4). A single species clade was
- very rare in the RR-1 tree. Although a small number of RR-1 CP genes were expanded
- in comparison with *B. mori*, more than half of them showed one-to-one correspondence
- among the three lepidopteran species. Sl_Aki-2, 3, 4, 5 and 6 and MsCPR13, 174, 175,
- 228 176 and 177 formed separate clades, which corresponded to *BmorCPR13* (black star in
- Fig. 4A). Another small species-specific RR-1 CP clade was observed (thick bar in Fig.4A).
- In sharp contrast to the situation with RR-1s, more than half of the RR-2 CPs formed
- 232 species-specific clades (Fig. 4B), indicating species-specific expansions by gene
- 233 duplication events. Intriguingly, all CP members of S. litura, M. sexta and B. mori
- belonging to large species specific clades (green branch in Fig. 4B) were His-rich.
- 235 Twenty-five RR-2 CPs (Sl_Aki-100-108, Sl_Aki-110-117, Sl_Aki-122, 124, 127, 129,
- 132, 134 136) formed the biggest S. litura specific clade (Blue star in Fig.4B), all of
- which belonged to the *RR-2 CP* gene cluster in chr9. BmorCPR152, MsCPR152,
- 238 MsCPR153 and Sl_Aki-166, which had the highest His-residue content in each species,

made a single clade (red star in Fig.4B), indicating that this highe	st His CP is
--	--------------

240 conserved and may play a common role in some specific structure among the three

Lepidoptera. Five of the CPFLs (Sl_Aki-236, 237, 238, 239, 240) formed a clade in S.

242 *litura* (Supplementary Fig. 2).

243 **3.5 Transcript abundance of CPs**

We conducted RNA-Seq analysis to study the transcript distribution of CP genes in 244 245 various developmental stages and tissues. In total, transcriptional evidence was 246 obtained for 283 of the 287 annotated CPs (see Supplementary Table 2 for numbers of genes from each CP group expressed per library). Transcripts were found for a 247 maximum of 233 CP genes in the 2nd instar larvae and a minimum of 94 CP genes in 248 249 the epidermis of the 6th day pupa (Fig. 5A). The pattern of transcript levels of different 250 CP groups (Fig. 5B) showed dynamic changes in epidermis from larval to pupal stages (Supplementary Table 3 for the total FPKM value for CP groups in each library). The 251 252 transcript pattern of RR-1 CPs contrasted extremely compared to RR-2 CPs (Fig. 5B). 253 Most CP transcripts in larval epidermis were derived from RR-1 CP genes whereas 254 RR-2 transcripts predominated in pharate adults. The difference in transcript pattern 255 between RR-1 and RR-2 genes was consistent with published reports that the RR-1 256 transcripts are much more abundant in soft and flexible cuticles typical of larvae than 257 in hard cuticles, whereas RR-2 transcripts are associated with hard structures typical of adults (Dittmer et al., 2015; Ettershank, 1964; Futahashi. et al., 2008; Vannini and
Willis, 2017).

260	The transcripts of CPG and CPH amount to a large percentage in pupal and pharate
261	adult stages, despite their small numbers of genes compared with the CPR family
262	(CPG 16-22; CPH 10-21; CPR 38-151) (Fig. 5A,B). In <i>B. mori</i> , transcripts from <i>CPG</i>
263	genes are also reported to be present in some hard cuticles such as pheromone gland,
264	compound eye and maxillary galea (Futahashi. et al., 2008). Most CPHs contain AAP
265	(A/V) repeats which might cause a protein to fold into a more or less regular helix,
266	leading to an elastin-like structure which is easily and reversibly deformed by external
267	forces (Andersen, 1995). Transcripts of CPH genes were abundant in pupal epidermis
268	and wing, especially in early pupal epidermis (Fig. 5B). Further study is needed to
269	determine the function of the CPGs and CPHs in cuticle formation.
270	Wolfgang and Riddiford (1986) reported changes of CP synthesis correlated with
271	changes of lamellar structure in <i>M. sexta</i> cuticle during the final larval instar when they
272	dig into soil in preparation for pupation. Our RNA-Seq analysis of S. litura suggested
273	that epidermal cuticular layers of 6 th larval instar would be mainly composed of RR-1s
274	as well as CPAPs, whereas, based on transcript abundance, CPH and CPGs together
275	with CPAPs would contribute to the cuticular layers of the wandering stage (Fig. 5A;
276	Fig. 6). Especially, CPHs encoded by Sl_Aki-260 (Supplementary Fig. 4) and Sl_Aki-
277	261 (Supplementary Fig. 5) were extremely highly expressed in the wandering stage at
278	night compared to the daytime. Nevertheless, the ratios of expressed gene numbers

l

among *CP* gene families did not change so much in epidermis between 6th larval instar
and wandering stage (Fig. 5B).

281	The heat map of transcripts for each CP gene in S. litura showed several
282	characteristic patterns (Fig. 6). The transcripts from RR-1 and CPG genes were
283	continuously and highly abundant in epidermis of larval stages (Fig. 6A). However,
284	transcripts from some RR-1 genes (Fig. 6B) that were abundant in larval epidermis
285	were also found in the early pupal stage (P-2) or pharate adult stage (P-12). This is
286	similar to An. gambiae, where Willis (2010) reported that both RR-1 and RR-2
287	transcripts are present in pharate adults and post-eclosion, but many fewer are RR-1
288	compared with RR-2.
289	The secondary structure predicted by online software Phyre ² (Kelley et al., 2015)
290	(<u>http://www.sbg.bio.ic.ac.uk/</u>) for RR-1 CPs of pattern B (Fig. 6B) suggested that most
291	of them shared a common structure homologous to the Polo-Box domain (Park et al.,
292	2010), which comprises a six-stranded antiparallel β -sheet shielded by one α -helix.
293	However, we could not find such a common domain in RR-1 CPs of pattern A (Fig.
294	6A). As Vannini and Willis (2017) reported, the location of RR-1s and RR-2s depends
295	on properties of individual proteins in An. gambiae. It will be interesting to learn how
296	RR-1s are involved in specific protein structures of adult cuticle.
297	The CPH genes (Sl_Aki-262, 265, 266, 267and 268) which formed a cluster on

298 Chr28 and a small cluster of *RR-2* genes (*Sl_Aki-186, 187, 188, 189* and *190*) in Chr15

1

299 (Supplementary table. 1) had transcripts mainly in late larval and early pupal stages 300 (Fig. 6C).

301	RR-2 genes, which are the main members of "pattern D" of the heat-map (Fig. 6D),
302	were highly expressed in the pharate adult wing and epidermis (P-9 and P-12). These
303	genes were also well conserved among the three moth species with one-to-one
304	correspondence in the phylogenetic trees (Fig. 4B). Sl_Aki-166, encoding the highest
305	His-rich CP, had abundant transcripts in the pharate adult stage. This expression
306	pattern was similar to its B. mori ortholog, BmorCPR152, which encodes the highest
307	His-rich CP (Suetsugu et al., 2013). These CP genes, which were mainly expressed in
308	the pharate adult stage, may contribute to scales or other specific structures in the adult.
309	Not only RR-2 transcripts, but also transcripts from other CP families such as RR-1,
310	CPG, CPFL, CPAP and CPH, were also observed in high abundance in pupal wing and
311	epidermis (Fig. 6C, D). The finding that two RR-1 genes (Sl_Aki-182 and 210) were
312	expressed highly only in the pupal stage, but not in the larval stage (Fig. 6D), is
313	interesting since most RR-1 genes had high transcript levels in larval epidermis. The
314	contrasting expression patterns of the genes encoding these two RR-1 proteins suggests
315	that their function merits further study.
316	Transcripts from a few CPAP, RR-1 and CPH genes were observed in abundance in
317	epidermis throughout the development and wing of pharate adult (Fig. 6E). The
318	transcripts of three CPAP genes (Sl_Aki-20, 22 and 23) showed a steady high
319	abundance in epidermis in all stages of pupal wing. These three genes are orthologs of
	16

320 CPAP3-A1, CPAP3-B and CPA3-C in both *M. sexta* and *T. castaneum*. Their

- 321 biological importance is indicated by reports that RNAi knockdown of TcCPAP3-A1
- 322 causes adult lethality, down-regulation of the *TcCPAP3-B* causes a walking defect and
- 323 RNAi treatment for *TcCPAP3-C* leads to molting arrest at the pharate adult stage
- 324 (Jasrapuria et al., 2012; Petkau et al., 2012).
- 325 His-rich *RR-2 genes* (*Sl_Aki-100-108, Sl_Aki-110-117, Sl_Aki-122, 124, 127, 129,*
- 326 132, 134 136), which form the large species-specific clade in the phylogenetic tree
- 327 (Blue star in Fig.4B), unexpectedly showed few transcripts in epidermis or other
- 328 samples (Supplementary Fig. 3). However, since the transcripts were only examined in
- 329 limited tissues in this study, we estimate these species-specific *His-rich CP* genes may
- be expressed in some hard cuticle structures such as the cornea of the compound eye,
- 331 maxilla, or antenna.
- Another unexpected finding was that transcripts from CPG, CPAP, CPT, CPH,
- 333 *RR-1* and *RR-2* were present in relatively high abundance at the 2^{nd} instar

334 (Supplementary Fig. 4), which was quite different from those of other larval stages.

335 Other CPs that were abundant at later stages also had high levels of transcripts in the

336 2nd instar (Supplementary Fig. 5). Among several possibilities to explain this peculiar

- 337 expression pattern across all the samples is the large/major morphological change that
- 338 occurs in larvae during the transition from 1^{st} to 2^{nd} instar. Further study of the
- 339 ultrastructure and physical properties of larval cuticle at these stages may help to
- 340 explain this observation. Another likely possibility is related to the timing of sample

341 collection. If the sampling was performed at the pharate third instar stage, CP

342 expression would be higher than in mid-instar larvae. It is also of interest to check the

343 transcripts of *CP* genes in the pharate stage. Levels of other *CP* transcripts are shown

in Supplementary Figure 5.

345 **4.** Conclusion

S. litura, among the most economically important global agricultural pests, is 346 347 characterized by a short life cycle, direct contact with soil during both late larval and 348 pupal/pharate adult stages, long distance migration and quick adaptation to diverse ecological niches. RNA-Seq analysis suggests that differential expression of various 349 350 CP groups to produce cuticle layers with different physical properties at different 351 stages. This study aimed to illustrate how the CP genes have been adapted to the 352 characteristic life style of this pest by analyses of their genome organization, phylogenetics and transcriptomics. Similar to M. sexta (Dittmer et al., 2015), the RR-2 353 354 group has expanded more than other CP groups, largely by gene duplication events 355 after speciation. Additionally, we found that 91 of 129 RR-2 CPs are His-rich, 356 amounting to twice the number and a greater fraction than in *B. mori* (from 58% in *B.* 357 mori to 71% in S. litura). These His-rich proteins are likely involved in cuticular 358 sclerotization since a His residue, which contains a nucleophilic imidazole group, can 359 react with ortho-quinone or dehydrobenzodioxine (Andersen, 2010, 2012). Although 360 more evidence is needed, we speculate that the high content of His-residues and the

361	large expansion of <i>His-rich RR-2 CP</i> genes that occur in <i>S. litura</i> could contribute to
362	construct tougher cuticles that could protect against fungal and bacterial attacks and
363	abrasion throughout their larval and pupal stages. Comparison with M. sexta CPs also
364	strengthened this idea, since both species which stay under soil during pupal/pharate
365	adult stages without cocoons share a large expansion of His-rich RR-2 CP genes
366	compared with B. mori.

367	The present work clearly showed how this pest established its unique life cycle
368	through the expansion of <i>His-rich</i> CP genes, which provides insights into the possible

regulation of CPs for pest control and their properties as new biomaterials. 369

370 Acknowledgements

371	We are very grateful to Dr	. Judith H.	Willis for many	helpful discussions and	1

372 comments for this MS. We also thank a reviewer and Prof. H. Kishino (University

373 Tokyo) for their valuable comments and discussion on phylogenetic analysis. This

- 374 research was supported by a grant from the One Thousand Foreign Experts
- 375 Recruitment Program of the Chinese Government (No. WO20125500074).

376 References

377 Andersen, S.O., Roepstorff, P., 1997. Sequence studies of proteins from larval and pupal cuticle of the yellow meal worm, Tenebrio molitor. Insect Biochem. Mol. Biol. 378 379 27, 121-131.

380	Andersen,	S.O.,	1998.	Amino	acid	sequence	studies	on end	locut	icular	proteins	form	the
-----	-----------	-------	-------	-------	------	----------	---------	--------	-------	--------	----------	------	-----

- desert locust *Schistocerca gregaria*. Insect Biochem. Mol. Biol. 28, 421.
- 382 Andersen, S.O., Hojrup, P., Roepstorff P., 1995. Insect Cuticular Proteins. Insect
- 383 Biochem. Mol. Biol. 25, 153-176.
- 384 Andersen, S.O., 2000. Studies on proteins in post-ecdysial nymphal cuticle of locust,
- 385 Locusta migratoria, and cockroach, Blaberus craniifer. Insect Biochem. Mol. Biol. 30,
- 386 569-577.
- 387 Andersen, S.O., 2007. Involvement of tyrosine residues, N-terminal amino acids, and
- 388 beta-alanine in insect cuticular sclerotization. Insect Biochem. Mol. Biol 37, 969-974.
- 389 Andersen, S.O., 2008. Quantitative determination of catecholic degradation products
- 390 from insect sclerotized cuticles. Insect Biochem. Mol. Biol. 38, 877-882.
- 391 Andersen, S.O., 2010. Insect cuticular sclerotization: a review. Insect Biochem. Mol.
- 392 Biol. 40, 166-178.
- 393 Andersen, S.O., 2012. Cuticular Sclerotization and Tanning. In: Gilbert, L.I., Iatrou, K.,
- 394 Gill, S. (Eds.), Comprehensive Molecular Insect Science, vol. 4. Elsevier Press, Oxford,
- 395 UK, pp. 145-170.
- 396 Andersen, S.O., Roepstorff, P., 2007. Aspects of cuticular sclerotization in the locust,
- 397 Scistocerca gregaria, and the beetle, Tenebrio molitor. Insect Biochem. Mol. Biol. 37,
- 398 223-234.
- Baton, L.A., Robertson, A., Warr, E., Strand, M.R., Dimopoulos, G., 2009.
- 400 Genome-wide transcriptomic profiling of *Anopheles gambiae* hemocytes reveals

l

- 401 pathogen-specific signatures upon bacterial challenge and Plasmodium berghei
- 402 infection. BMC genomics. 10, 257.
- 403 Cheng, T., Wu, J., Wu, Y., Chilukuri, R.V., Huang, L., Yamamoto, K., Feng, L., Li,
- 404 W., Chen, Z., Guo, H., Liu, J., Li, S., Wang, X., Peng, L., Liu, D., Guo, Y., Fu, B., Li,
- 405 Z., Liu, C., Chen, Y., Tomar, A., Hilliou, F., Montagne, N., Jacquin-Joly, E.,
- 406 d'Alencon, E., Seth, R.K., Bhatnagar, R.K., Jouraku, A., Shiotsuki, T., Kadono-Okuda,
- 407 K., Promboon, A., Smagghe, G., Arunkumar, K.P., Kishino, H., Goldsmith, M.R.,
- 408 Feng, Q., Xia, Q., Mita, K., 2017. Genomic adaptation to polyphagy and insecticides in
- 409 a major East Asian noctuid pest. Nat Eco Evol. 1, 1747-1756.
- 410 Cornman, R.S., Willis, J.H., 2009. Annotation and analysis of low-complexity protein
- 411 families of *Anopheles gambiae* that are associated with cuticle. Insect Mol Biol. 18,
- 412 607-622.
- 413 Dittmer, N.T., Tetreau, G., Cao, X., Jiang, H., Wang, P., Kanost, M.R., 2015.
- 414 Annotation and expression analysis of cuticular proteins from the tobacco hornworm,
- 415 Manduca sexta. Insect Biochem. Mol. Biol. 62, 100-113.
- 416 Dong, Z., Zhang, W., Zhang, Y., Zhang, X., Zhao, P., Xia, Q., 2016. Identification and
- 417 Characterization of Novel Chitin-Binding Proteins from the Larval Cuticle of
- 418 Silkworm, *Bombyx mori*. J. Proteome Res. 15, 1435-1445.
- 419 Felsenstein J, 1985. Confidence limits on phylogenies: An approach using the
- 420 bootstrap. Evolution 39, 783-791.

- 421 Futahashi., R., Okamoto., S., Kawasaki., H., Zhong, Y.-S., Iwanaga., M., Mita, K.,
- 422 Fujiwara, H., 2008. Genome-wide identification of cuticular protein genes in the
- 423 silkworm, *Bombyx mori*. Insect Biochem. Mol. Biol. 38, 1138-1146.
- 424 Gu, S., Willis, J.H., 2003. Distribution of cuticular protein mRNAs in silk moth
- 425 integument and imaginal discs. Insect Biochem. Mol. Biol. 33, 1177-1188.
- 426 He, N., Botelho, J.M., McNall, R.J., Belozerov, V., Dunn, W.A., Mize, T., Orlando, R.,
- 427 Willis, J.H., 2007. Proteomic analysis of cast cuticles from Anopheles gambiae by
- 428 tandem mass spectrometry. Insect Biochem. Mol. Biol. 37, 135-146.
- 429 Ioannidou, Z.S., Theodoropoulou, M.C., Papandreou, N.C., Willis, J.H., Hamodrakas,
- 430 S.J., 2014. CutProtFam-Pred: detection and classification of putative structural
- 431 cuticular proteins from sequence alone, based on profile hidden Markov models. Insect
- 432 Biochem. Mol. Biol. 52, 51-59.
- 433 Kerwin, J.L., Turecek, F., Xu, R., Kramer, K.J., Hopkins, T.L., Gatlin, C.L., John
- 434 R.Yates, J.R., III, 1999. Mass spectrometric analysis of catechol-histidine adducts from
- 435 insect cuticle. Anal Biochem. 268, 229-37.
- 436 Jasrapuria, S., Arakane, Y., Osman, G., Kramer, K.J., Beeman, R.W., Muthukrishnan,
- 437 S., 2010. Genes encoding proteins with peritrophin A-type chitin-binding domains in
- 438 *Tribolium castaneum* are grouped into three distinct families based on phylogeny,
- 439 expression and function. Insect Biochem. Mol. Biol. 40, 214-227.

- 440 Jasrapuria, S., Specht, C.A., Kramer, K.J., Beeman, R.W., Muthukrishnan, S., 2012.
- 441 Gene families of cuticular proteins analogous to peritrophins (CPAPs) in Tribolium
- 442 *castaneum* have diverse functions. PloS one. 7, e49844.
- 443 Kerwin, J.L., Turecek, F., Xu, R., Kramer, K. J., Hopkins, T. L., Gatlin, C. L., Yates,
- 444 J.R., 1999. Mass spectrometric analysis of catechol-histidine adducts from insect
- 445 cuticle. Anal Biochem. 268, 229-237.
- 446 Locke, M., 2001. The Wigglesworth Lecture: Insects for studying fundamental
- 447 problems in biology. J Insect Physiol 47, 495–507.
- 448 Magkrioti, C.K., Spyropoulos, I.C., Iconomidou, V.A., Willis, J.H., Hamodrakas, S.J.,
- 449 2004. cuticleDB: a relational database of Arthropod cuticular proteins. BMC
- 450 bioinformatics. 5, 138.
- 451 Nei, N.S.M., 1987. The neighbor-joining method- a new method for reconstructing
- 452 phylogenetic trees. Mol Biol Evol 4:406-25.
- 453 Pan, P.L., Ye, Y.X., Lou, Y.H., Lu, J.B., Cheng, C., Shen, Y., Moussian, B., Zhang,
- 454 C.X., 2018. A comprehensive omics analysis and functional survey of cuticular
- 455 proteins in the brown planthopper. Pro. Natl. Acad. Sci. USA 115, 5175-5180.
- 456 Park, J.E., Soung, N.K., Johmura, Y., Kang, Y.H., Liao, C., Lee, K.H., Park, C.H.,
- 457 Nicklaus, M.C., Lee, K.S., 2010. Polo-box domain: a versatile mediator of polo-like
- 458 kinase function. Cell Mol Life Sci. 67, 1957-1970.

- 459 Petkau, G., Wingen, C., Jussen, L.C.A., Radtke, T., Behr, M., 2012. Obstructor-A Is
- 460 Required for Epithelial Extracellular Matrix Dynamics, Exoskeleton Function, and
- 461 Tubulogenesis. J Biol Chem. 287: 21396–21405.
- 462 Rebers, J.E, Willis, J.H., 2001. A conserved domain in arthropod cuticular proteins
- 463 binds chitin. Insect Biochem. Mol. Biol. 31, 1083-1093.
- 464 Andersen, S.O., Rafn, K., Roepstorff, P., 1997. Sequence studies of proteins from
- 465 larval and pupal cuticle of the yellow meal worm, *Tenebrio molitor*. Insect Biochem.
- 466 Mol. Biol. 27, 121-131.
- 467 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and
- 468 post-analysis of large phylogenies. Bioinformatics. 30, 1312-1313.
- 469 Suetsugu, Y., Futahashi, R., Kanamori, H., Kadono-Okuda, K., Sasanuma, S.,
- 470 Narukawa, J., Ajimura, M., Jouraku, A., Namiki, N., Shimomura, M., Sezutsu, H.,
- 471 Osanai-Futahashi, M., Suzuki, M.G., Daimon, T., Shinoda, T., Taniai, K., Asaoka, K.,
- 472 Niwa, R., Kawaoka, S., Katsuma, S., Tamura, T., Noda, H., Kasahara, M., Sugano, S.,
- 473 Suzuki, Y., Fujiwara, H., Kataoka, H., Arunkumar, K.P., Tomar, A., Nagaraju, J.,
- 474 Goldsmith, M.R., Feng, Q., Xia, Q., Yamamoto, K., Shimada, T., Mita, K., 2013.
- 475 Large scale full-length cDNA sequencing reveals a unique genomic landscape in a
- 476 lepidopteran model insect, *Bombyx mori*. G3. 3, 1481-1492.
- 477 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6:
- 478 Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 30, 2725-2729.

- 479 Tang, L., Liang, J., Zhan, Z., Xiang, Z., He, N., 2010. Identification of the
- 480 chitin-binding proteins from the larval proteins of silkworm, *Bombyx mori*. Insect
- 481 Biochem. Mol. Biol. 40, 228-234.
- 482 Tetreau, G., Dittmer, N.T., Cao, X., Agrawal, S., Chen, Y.R., Muthukrishnan, S.,
- 483 Haobo, J., Blissard, G.W., Kanost, M.R., Wang, P., 2015. Analysis of chitin-binding
- 484 proteins from *Manduca sexta* provides new insights into evolution of peritrophin
- 485 A-type chitin-binding domains in insects. Insect Biochem. Mol. Biol. 62, 127-141.
- 486 Togawa, T., Augustine Dunn, W., Emmons, A.C., Willis, J.H., 2007. CPF and CPFL,
- 487 two related gene families encoding cuticular proteins of Anopheles gambiae and other
- 488 insects. Insect Biochem. Mol. Biol. 37, 675-688.
- 489 Togawa, T., Nakato, H., Izumi, S., 2004. Analysis of the chitin recognition mechanism
- 490 of cuticle proteins from the soft cuticle of the silkworm, *Bombyx mori*. Insect Biochem.
- 491 Mol. Biol. 34, 1059-1067.
- 492 Vannini, L., Willis, J.H., 2017. Localization of RR-1 and RR-2 cuticular proteins
- 493 within the cuticle of *Anopheles gambiae*. Arthropod Struct Devl. 46, 13-29.
- 494 Willis, J.H., Papandreou, N.C., Iconomidou, V.A., Hamodrakas, S.J., 2012. Cuticular
- 495 proteins. In: Gilbert, L.I. (Ed.), Insect Molecular Biology and Biochemistry. Ac-
- 496 ademic Press, London, Waltham & San Diego, pp. 134e166.
- 497 Wolfgang, W.J., Riddiford, L.M., 1986. Larval cuticular morphogenesis in the tobacco
- 498 hornworm: *Manduca sexta*, and its hormonal regulation. Devl. Biol. 113, 305-316.

499	Yang, Z., 1994 . Maximum likelihood phylogenetic estimation from DNA sequences
500	with variable rates over sites: approximate methods. J Mol Evol. 39, 306-314.
501	
502	Figure Captions
503	Figure 1. Differences in behavior and life cycle between S. litura and B. mori. L1-L6
504	refer to larval instars 1-6; W is the wandering stage; numbers in parentheses represent
505	days at a particular stage.
506	Figure 2. The number of His-rich CPs in each CP group. His-rich RR-2 CPs in S. litura,
507	91; B. mori, 54; His-rich RR-1 CPs in S. litura, 0; B. mori: 2; His-rich CPGs in S. litura,
508	11; B. mori 9; His-rich CPLCA in S. litura, 3; B. mori, 1; His-rich CPAP in S. litura, 3;
509	B. mori, 1; His-rich CPH in S. litura, 4; B. mori, 2. Blue, B. mori CPs; red, S. litura
510	CPs.
511	Figure 3. Chromosomal location of the largest cluster of CPR genes in S. litura
512	compared with <i>B. mori.</i> RR-1 genes (red) and RR-2 genes (blue) are located primarily
513	in separate clusters on chromosome 9. Arrows indicate direction of transcription.
514	Figure 4. Phylogenetic trees of annotated RR-1 and RR-2 proteins among S. litura, B.
515	mori and M. sexta. (A) RR-1 protein genes. Black star, BmorCPR13; thick bar,
516	species-specific clade. (B) RR-2 protein genes. blue, M. sexta CPs; yellow, B. mori
517	CPs; magenta, S. litura CPs; blue star, 25 RR-2 CPs forming the biggest S. litura
518	specific clade; red star, highest His-content RR-2 CP among the three species, forming

519	a clade. Green branch, His-rich CP genes. Bootstrap values of 70 or higher are shown
520	in the branches.

- 521 Figure 5. Transcript distribution of CP genes in RNA-Seq libraries. (A) Total FPKM
- 522 values of transcripts for CP genes in each family. (B) Numbers of CP genes with
- 523 transcripts in each family. EP, Epidermis; D, daytime; N, nighttime; Wg, wing.
- 524 Figure 6. Heatmap of transcripts of *S. litura CP* genes grouped into five distinct
- 525 patterns. (A) CP transcripts mainly found in the larval epidermis. (B) CP transcripts
- 526 expressed highly in larval and pharate adult stages. (C) CP transcripts expressed highly
- 527 in epidermis of late larval to early pupal stages. (D) *CP* transcripts found mainly in the
- 528 pharate adult stage. (E) Transcripts of *CP* genes highly abundant through the larval to
- 529 pupal stage.











Numbers of CP genes with transcripts in each family





Table1. Size	e of each	cuticular protein family in Spodoptera litura and Bombyx mori
Motif	S.litura	B.mori
RR-1	63	56
RR-2	129	93(4) ^a
RR-3	1	3
Tweedle	5	4
CPF	1	1
CPFL	7	4
CPLCA	4	$2(\text{BmorCPH6},7)^{\text{b}}$
CPCFC	1	1(BmorCPH1) ^c
CPAP1	13	14
CPAP3	9	9
Glycine-Rich	28	29
СРН	26	31
Total	287	247
^{a, b, c} items in pa	rentheses	refer to CPs reported in Futahashi et al. (2008)
^a 4 more cuticul	ar RR-2 g	enes were identified in S. litura than B. mori

Supplementary Table 1. Classification of CP coding genes.

New SI_Ak+001	Hs-Content (signal prptide removed)	RR-1	of signal peptide	corresponding SWUSI #(old) SWUSI10006476	corresponding SWUSI #(New) SWUSI0006070	Chr 1 mRNi	Start (chr) 1908343	End (chr) 1909789	dir +	scattoid scattoid227	Start (scaf) 377929	End (scaf) 376483	de C	DS (bp) a	a length	exon # 2	protein feature Gy54(1628),Ala46(13.77%),Ser39(11.7%),Pro33(0.9%)						
S_Ak+002 S_Ak+003 S_Ak+004 S_Ak+005		88-1 88-1 88-1 88-1	15 15 15	SWUS10006475 SWUS10006473 SWUS10006473	SWUSD006080 SWUSD006090 SWUSD006100 SWUSD006110	1 mP00 1 mP00 1 mP00 1 mP00	1926011 1930513 1935412 1942925	1927618 1931607 1936440 1943888	* * *	scattold227 scattold227 scattold227 scattold227	355759 350860 342384	354665 349832 343347	. s	46 2 93 2 90 2	281 230 229	4 4 4	LIB31 (8-W8,MED9(1-3-W8,J&HV6) (10.8%),AB52(1-1.5%) Ab44(1-5.7%),Sar30(10.7%),Gh24(8.5%);Gb(23(8.19%);Gh(Q)24(8.54%) Ab45(17.7%),Sar324(10.4%);Gb(232(19.57%) Ab46(17.5%),Sar21(9.1%);Gb(232(19.57%))				-		
2_AL-005 2_AL-007 2_AL-005		18-1 18-1 18-1	15 0 15	SWUST0008471 ISWUST0008470 SWUST0008469	SWUSI0006120 SWUSI0006130 SWUSI0006140	1 millio 1 millio 1 millio	1347817 1354832 2026662	1948628 1957664 2027410	+ +	scaffold227 scaffold227 scaffold227	337644 329440 259610	338435 328638 258862	- 5 - 6 - 7	24 1 00 1 02 2	87 199 232	3	Ab29(15.5%),Gy(G)18(9.63%); Ser14(8.6%),Gin15(8.0%),Gu13(7.0%) Ser21(10.6%),Pro(P17(8.54%); Gy(G)18(9.05%)); Aun14(7.0%),Ap)13(8.5%),Gu GyG2(26.7%),Ab38(16.4%)	2(6.0%)					
SLAki-010 SLAki-011 SLAki-011		88-1 88-1 88-1 88-1	17 17 17	SWUS1000E465 SWUS1000E464 SWUS1000E463 SWUS1000E462	SWUSIDOG180 SWUSIDOG6190 SWUSIDOG6463 SWUSIDOG6210	1 mRNo 1 mRNo 1 mRNo 1 mRNo	20070274 2073274 2100870 2114423	2051041 2102356 2118483	* - -	scattold227 scattold227 scattold227 scattold227	205231 183916 167789	212938 185402 171849	· · ·	22 1 56 1 31 1	173 185 176	4 3 4	Log25(17:3%),Gh16(12:3%),Gh17(2:5%); Ad(A)16(3:1%); Gy18(10:4%),Ah29(16:3%),Gh17(2:3%),Gh13(7:5%) Ah225(13:5%),Gh19(10:3%); Gy1(G16(8:63%); Sw(S)16(8:65%) Ah220(15:9%),Tv2D(11:4%)				-	-	
S_AkH013 S_AkH014 S_AkH015	Hzt5.34%	18-1 18-2 18-2	17	SWUST0008461 SWUST0008742 none	SWUSI0006220 SWUSI0006900	1 millio 1 millio 1 millio	2183105 3888642 3892472	2188858 3889806 3893190	* * *	scattoid227 scattoid145 scattoid145	103167 329196 333026	37414 330360 333744	- 5 - 5 - 5	22 1 014 3 67 1	73 537 188	4 3 2	Gly20(11.6%),Ab21(12.1%),Ser17(9.8%),Tyr16(9.3%) Gly58(17.2%),Ab44(13.1%),Lex38(11.3%),Pro42(12.3%),Tyr23(6.8%);His(H)13(6. ValV(20(10.64%);Pro(P)20(10.64%);Lex4(L17(9.04%);Gy(6)16(6.51%)	54%)					
SL_Aki-017 SL_Aki-018 SL_Aki-019		CPAP1-0 CPG CPG	16 18 18	10008760 SWUSI10002513 SWUSI10002514	SWUSI0007070 SWUSI0007210 SWUSI0007220	1 mRNA 1 mRNA 1 mRNA	451364	4401774 4652844 4659053	* * -	scattold148 scattold148 scattold475 scattold475	814723 26962 33171	25482 32005	- 2	178 2 67 2 57 2	725 288 518	5 5 5	Hab (12) (12) (12) (12) (12) (12) (12) (12)	e.r.en)					
SI_Aki-020 SI_Aki-021 SI_Aki-022		CPAP3-A1 CPAP3-A2 CPAP3-B	16/7 20 4	10014292 10014293 10014295	SWUSIO011690 SWUSIO011700 SWUSIO011720	2 118% 2 118% 2 118%	2527417 2554035 2514637	2538168 2573487 2626168	* * -	scattoid19 scattoid19 scattoid19	563624 590302 650904	\$74375 \$09694 \$62375	* 7 * 7 * 8	14 2 17 2 97 2	237 238 298		CBM_14_20-72_20-136_174-219: Ov802_15-68_09-135_174-213 CBM_14_22-74_22-140_175-221_Ov802_20-68_01-112_176-215 CBM_14_57-107_124-174_201-246_C04802_53-102_121-169_185-240 CBM_14_14_57-107_124-174_201-246_C04802_53-102_121-169_185-240				-	=	
SI_Ak+024 SI_Ak+025 SI_Ak+025		CPAP3-01 CPAP3-02 CPAP3-02 CPAP3-E1	19 0 17	10014294 10015026 10004597	SWUSIO011710 SWUSIO012000 SWUSIO012480	2 m/00 2 m/00 2 m/00 2 m/00	A 2584663 A 3508533 A 4756976	2612712 3590600 4767956	+ + -	scattold19 scattold19 scattold19	620870 1544740 4557	648919 1626807 15537	+ 6 + 5	70 2 75 2	128 189 190	5 7 6	CBM_14 [74-12][104-146][101-225] CHED2 [27-67][100-145][160-210] CBM_14 [74-122][143-183][195-241]; CHED2 [73-102][142-183][193 CBM_14 [65-111][65-180][207-252]; CHED2 [64-103][131-174][207-244]						
SL_Aki-027 SL_Aki-028 SL_Aki-029 SL_Aki-029	Hs:7.2% Hs:12.08%	DPAP3-E2 DG 88-2 89-7	20	10004596 SWUS110007758 SWUS110004457	SWUSIO012470 SWUSIO023630 SWUSIO024960 SWUSIO024700	2 mRNo 5 mRNo 5 mRNo 6 mRNo	4754218 6253148 11008737	4756301 6254637 11010425 8630287	-	scattoid317 scattoid179 scattoid333 scattoid147	1799 213819 123549 136204	3882 212330 125237 136561	- 6	90 2 272 4 10 1 27 1	129 123 169	6 2 4 2	CBM_14 [105:154]; ChtB02 [28-71][98-148] Cby47(11.1%),Var/5(17.7%),pro54(12.5%),Lys54(12.5%) Cby(C)10(9.4%);Site(H118(10.65%);Aap(D)16(3.47%); Cb/(C)12.5%);Cbw(21.0%)				=	=	_
S_Ak+031 S_Ak+032 S_Ak+033		82-1 82-1 82-1	12 18 19	SWUS10006212 SWUS10006211 SWUS10006210	SWUSI0028900 SWUSI0028910 SWUSI0028920	6 11890 6 11890 6 11890	A 2206793 A 2222242 A 2242940	9208785 9226193 9245784		scattold240 scattold240 scattold240	163876 146468 126877	165868 150419 129721	+ 4 + 4 + 3	26 1 29 1 95 1	141 142 131	4 4 4	Ala21(14.9%),Val16(11.4%)(Pto(P)13(9.22%); Ala18(12.7%),Sar15(10.6%);Val(V)14(9.86%)Pto(P)14(9.86%) Ala18(13.7%),Sar11(8.4%);Pto(P)11(8.46%);Yr10(7.6%);Val(V)11(8.46%)						
SL_Aki-034 SL_Aki-035 SL_Aki-036 SL_Aki-037		021 18-2 18-2 074	0 2 0 0	0008 SWUSI10012478 SWUSI10005425 SWUSI10005424	SWUSI0031190 SWUSI0034970 SWUSI0034980	6 π/0% 7 π/0% 7 π/0% 7 π/0%	4070112 4070112 14722300	10462919 4071480 14723922 14734131	- - -	scattold37 scattold63 scattold275 scattold275	597328 317084 271283 261428	203476 318432 269661 259452	- 4	21 2 52 1 236 4 554 6	196 62 111 227	2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Gb(G)113(38,18%);5er(5)30(10,14%);de(1)26(8,78%); ProP191(11,73%);Aa(A)18(11,11%);G5(C)14(8,64%);5er(5)13(8,02%) Gb(Q)55(13,38%);Pro(P)52(12,65%);Aa(A) 41(3,38%) ProG03(8,0%);Gbr(O)49(7,31%);			_	=	=	_
SL_Aki+038 SL_Aki+039 SL_Aki+040		89-1 89-2 CPAP1-J	16 16 0	SWUSI10005423 SWUSI10005421 10005116	SWUSI0034990 SWUSI0035010 SWUSI0035530	7 mRNG 7 mRNG 8 mRNG	A 14740589 A 14779444 A 465865	14746822 14782211 479760		scattold275 scattold275 scattold290	246761 211372 7953	252894 214139 21845	+ 9 + 8 + 3	62 3 52 2 939 1	122 183 1312	6 3 5	Pro54(16.8%),G841(12.7%)(AB4(A)29(9.01%); Pro55(12.4%),G848(17.0%),AB4(A)25(8.13%) G8M_14 [114-163]; Cht802 [112-160]						
SI_Aki-041 SI_Aki-042 SI_Aki-043 SI_Aki-044	Hat 5.79%	0PAP1-K 88-2 88-2 88-2	23	10008046 SWUSI10000941 SWUSI10000941 SWUSI10000943	SWUSI0036010 SWUSI0037080 SWUSI0037080 SWUSI0037100	3 11200 3 11200 3 11200 3 11200	2308511 6197309 6203090 6203259	2376767 6198596 6203803 6209910	- + +	scattold168 scattold719 scattold719 scattold719	750940 44772 50553 56722	819196 46059 51266 57373	- 2 * 4 * 4	232 D 80 1 80 1 80 1	43 59 62 79	8 3 3 2	CBM_14_1377-456]; ChtBD2_[375-429] Aba20(15.2%), V421(15.2%), Pro17(10.7%), Tyr12(7.6%), Gby(G)12(7.55%) Val19(11.7%), Pro17(10.5%), Tyr10(6.2%) Ab45(25, 1%), Val33(118.4%), Min 11(6.2%), Tyr13(7.3%), Pro[P15(6.38%)				=	=	_
S_Ak+045 S_Ak+046 S_Ak+047	Ha:6.82%	CPAP1-L R83 CPG	16	10016513 SWUSI10005702 SWUSI10006973	SWUSI0037740 SWUSI0039150 SWUSI0039470	5 mRN0 3 mRN0 3 mRN0	4 8306915 A 35018 A 879767	8317130 38125 881529	+ + -	scattold3a scattold270 scattold205	2092833 35018 629386	2103048 38125 631148	+ 9 + 9 + 1	62 3 87 3 30 1	122 128 109	2 7 2	CBM_14 (189-232); ChtED2 (189-229) Abs(A)35(10.06%);Ge(Q)50(9.15%);Ga(E)27(8.23%) Gb(G)21(19.27%);Tyr(Y)15(13.76%);Abs(A)12(11.01%);Lea(L)8(7.34%);						
SL_Aki-048 SL_Aki-049 SL_Aki-050 SL_Aki-051		88-1 88-1 88-1 88-1	17 17 18 17	SWUS10010733 SWUS10010734 SWUS10010734	SWUSIO041300 SWUSIO041320 SWUSIO041320	2 11976 2 11976 3 11976 3 11976	63326579 6332264 63336677 6343990	6332994 6337362 6344643	* * *	scattold22 scattold22 scattold22 scattold22	625033 630698 635111 642424	631428 635796 643077	* 3 * 1 * 3	42 1 32 1 32 1 42 1	12 12 13	3	Aartz(10.00%),Var(H(12.4%),Gar(H(12.4%)(SayGa),11(9.73%); Var(7(15.2%),Gar(12.4%),Sar(5)10(8.93%); Ta(172(15.71%),Aart(10.8%),Var(11(9.8%),Gar(119.8%),Aapd(7.1%),Sar(0(8.9%)) Var(4(12.4%),Gar(1(11.5%),Aapd(7.1%),Ya5(8.0%);Aar(A) 10(8.85%)				=	=	_
SI_Aki-052 SI_Aki-053 SI_Aki-054		段-1 段-1 段-1	17	SWUS10010736 SWUS10010737 SWUS10010738	SWUSI0041340 SWUSI0041350 SWUSI0041360	2 mRNo 2 mRNo 2 mRNo	A 6343952 A 6357430 A 6362254	6350485 6359574 6365344	+ + +	scattold93 scattold93 scattold93	648386 655864 660688	645919 658008 663778	• 3 • 8 • 4	42 1 13 2 71 1	113 270 156	3 4 4	App5(8.0%);Glub(8.0%);Lys5(7.1%) Pro(P)41(15.19%);Tyr(Y)24(8.89%); Pro17(10.9%);Glub(26(16.7%)						
S_A6H055 S_A8H056 S_A8H057 S_A8H057		88-1 18-1 18-1 18-1	15 15 0	SWUST0010740 SWUST0010741 SWUST0010742	SWUSIO041370 SWUSIO041380 SWUSIO041390 SWUSIO041400	2 πΡου 2 πΡου 2 πΡου 2 πΡου	6375964 6375964 6384086 6398507	6376538 6385488 6389233	* * *	acattold22 acattold22 acattold22 acattold22	652520 674398 682520 696941	465739 474972 483922 497667	* * * 1 * 5 * 4	64 1 64 1 62 1 80 1	47 27 93 59	2	ABI/AE2(17.01%E)P10(714(2.52%); Ab17(13.4%)Pc014(11.0%)(Gb22(0;5%)(Gb10(7.9%) Ab1A()(47)(24.55%)(7%)(V21(10.8%)(Pro(P)17)(8.81%) Gb22(20.1%)(Ser24(15.1%)						
S_Ak+059 S_Ak+050 S_Ak+051 S_Ak+051		076 88-1 88-1	16	SWUS10010743 SWUS10010744 SWUS10010745	SWUSI0041410 SWUSI0041420 SWUSI0041430	2 m/00 2 m/00 2 m/00 2 m/00	4 6404931 4 6414609 4 6417948	6407713 6415382 6418940	+ + +	acattold99 acattold99 acattold99 acattold99	703415 713043 716382	7D6147 713816 717374	+ 8 + 3 + 4	25 2 72 1 47 1	274 123 145	3	Gly51(18.0%)Pr024(8.0%),Tyr30(11.0%) Ala14(11.4%),Tyr11(8.0%)(Val(V)12(.9.76%); Ala22(15.5%),Pr018(10.1%) District 20%)District 20%(12-12/12/18/04/04/04/18/0.86%).						
52_Aki-053 52_Aki-054 52_Aki-055		88-1 88-1 88-1	0010	SWUSI10010747 SWUSI10010748 SWUSI10010750	SWUSI0041450 SWUSI0041460 SWUSI0041460	2 118% 2 118% 2 118%	6456133 6461526 6462122	6458055 6463766 6483444	+ - +	scattold99 scattold99 scattold99	754567 762200 780556	756489 759960 781878	* 3 - 7 * 5	05 1 62 2 85 1	101 153 194	2	Ala13(12:0%),Pro14(13:0%)(Cdy(G)8(8:01%) Thr23(0:1%),Glu22(8:7%),Lyn21(8:3%),Arg21(8:3%) Pro16(8:3%),Glu19(9:8%)				_		
S_Ak+055 S_Ak+057 S_Ak+058 S_Ak+058		89-1 89-1 89-1 89-1	0 17 25	SWUS10010751 SWUS10010752 SWUS10010753 SWUS10010754	SWUSI0041490 SWUSI0041500 SWUSI0041510 SWUSI0041520	2 mRNo 2 mRNo 2 mRNo 2 mRNo 2 mRNo	4 6497412 4 6516199 4 6525848 6525848	6501292 6520405 6526709 6532078	* * *	scattold99 scattold99 scattold99 scattold99	795846 814633 824282 827989	799726 818839 825143 830512	• 3 • 4 • 3	30 1 65 1 60 1	109 154 119	3	Gy11 (10.1%),Ala12 (11.0%),Pro13(11.0%),Ga10(8.2%) Ala17(11.0%),Gh30(19.3%);Gy(G)13(8.4%) Yal13(10.3%),Aqs12(10.1%),Am9(7.6%),Ala(A)11(9.24%) Alu12 98% Am119 58% Am9(7.2%),GA12(2.8%),Am9(111),2.48%),Gu(G)10(8.62%)				-	-	
SI_Aki-070 SI_Aki-071 SI_Aki-072		10-1 10-1 10-1	2002	SWUSI10010755 SWUSI10010756 SWUSI10010757	SWUSI0041530 SWUSI0041540 SWUSI0041550	2 118%	6534162 6553707 6565425	6534832 6554162 6565935	+ + +	scattoid?? scattoid?? scattoid??	832596 852141 863859	833266 852596 864369	+ 3 + 3 + 3	51 1 42 1 51 1	15	3	Val12(10.3%), Aan12(10.3%), Aap7(6.0%), Gu7(6.0%), GloB(6.3%); App13(11.5%), Aap7(6.2%), GOV(G)11(2.73%); Gy12(10.3%), Ala15(12.3%), Val12(10.3%), Gln13(11.2%)						
SI_Aki-075 SI_Aki-075 SI_Aki-075		80-1 80-1 80-1 80-1	15 15 15	SWUS10010759 SWUS10010760 SWUS10010760	SWUSIO41580 SWUSIO41580 SWUSIO41580 SWUSIO41580	2 11970 2 11970 2 11970 2 11970	4 6578924 4 6587290	6575744 6579679 6588020	* * *	acattold22 acattold22 acattold22	873410 877358 885724	876113 886454	+ 3 + 3 + 3	63 1 57 1 45 1	120 120 118	3	Lag13(10.00%),AB13(12.20%),VB14(11.7%),AB20(7.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(5.20%),AB7(6.20\%),AB7(6.20\%),AB7(~				-	
SI_Aki-077 SI_Aki-078 SI_Aki-079		10-1 10-1 10-1	17	SWUS10010763 SWUS10010764 SWUS10010765	SWUSI0041610 SWUSI0041620 SWUSI0041630	2 118%	6598391 6610771 6614940	6608913 6611988 6616458	+ + -	scattoid?? scattoid?? scattoid??	896825 909205 914892	907347 910422 913374	+ 1 + 1 - 1	21 1 12 1 60 1	105 103	2	Ala(A)11(10.38%): Gy(G)9(8.49%); Val(V19(8.74%) Ala23(19.3%),App(7.6%),Giz9(7.6%),Giz9(7.6%),Val(V)11(9.24%); acidic						
S_Ak+051 S_Ak+052 S_Ak+053		18-2 18-2 18-2	27 17 25	SWUSI10010768 SWUSI10010769 SWUSI10008518	SWUSI0041650 SWUSI0041660 SWUSI0041940	2 mRNJ 2 mRNJ 2 mRNJ	6660539 6678768 7119339	6663448 6709257 7121766	* * *	scattold99 scattold99 scattold99	958973 977202 596714	961852 1007691 599141	· · ·	605 5 062 3 374 4	534 553 157	4 4 4	Tarr (10539)(2016)(20170), 2019(10(2016)) Gly63(11396)(20160(11286)/Pro46(8.6%)(2045(8.4%)) Prof/154(8.2536)(201(8.7856)(201(130(8.50%);Tyr(Y)28(7.93%);Arg(R)26(7.3 Prof/9(66(14.44%);	7%);Asp[0	25(7.08%));GY(G)25(7	08%);Gh(0)25(7.08	1%);
SL Aki-054 SL Aki-056 SL Aki-055 SL Aki-055 SL Aki-055		DPAP1-M 88-2 88-2 88-2	0 12 0 12 12	10008194 SWUS10008112 SWUS10008109	SWUSI0042310 SWUSI0042390 SWUSI0043020 SWUSI0043020	2 millio 2 millio 2 millio 2 millio	4 8290351 9435634 9473801 9473801	8294236 9436149 9474790 9487.0FF	*	acaffold163 acaffold167 acaffold167 acaffold167	421184 569866 530899 520994	425069 568351 529710 518014	4	241 2 62 1 55 2	146 53 54 63	3	CBM_14 [39-117]; CHED2 [57-109] Abs/A114(9,159);Val(V)15(9,809);CgV(G13(8,599); Val(V)21(0,718);Cap(G12(9,528);CgV(G13(8,599); Val(V)21(0,718);Cap(G12(9,528);CgV(G12(8,599); Val(V)21(0,718);Cap(G12(9,728);Cap(G12(1,78);Cap(G12(1,738))))))))))))))))))))))))))))))))))))					1	
S_Ak+020 S_Ak+021 S_Ak+022	Hx:8.33% Hx:8.2%	18-2 18-2 18-2	000	SWUS10008107 SWUS10008107 SWUS10008105	SWUSIO04 3040 SWUSIO04 3040 SWUSIO04 3040	9 11950 9 11950 9 11950 9 11950	2502924 2502924 2509153 2655274	9504139 9512544 9656294	*	scaffold167 scaffold167 scaffold167	501576 495347 349226	500361 491956 347506	4	. 1 27 1 32 1 41 1	43	3 3 4	Val(V)2(8.33%)2(ii)(H)2(8.33%) Pro1(11.20%),Val(V)13(9.09%) Thr14(9.8%),Hin12(8.2%)					=	
SI_Aki-023 SI_Aki-025 SI_Aki-025 SI_Aki-025 SI_Aki-027	Hs:20.27% Hs:20.0%	88-2 88-2 88-2 88-2	0 17 16 0	SWUS10008104 SWUS10008103 SWUS10008102 none	SWUSI0043070 SWUSI0043080 SWUSI0043090	2 116%	A 2664925 A 2683323 A 2685974 A 2625974	9666591 9684597 9687179 9693858	*	acaffold167 acaffold167 acaffold167 acaffold167	339575 321107 318526 310642	337909 319903 317321 311854		137 3 60 1 17 2 58	578 119 138 185	3	App(D)44(11.64%)(2yn)(K)38(10.05%)(5ro(P)37(3.79%) Abs(A)2016.81%)(5p(G)19(15.97%)(2yn)(Y)10(8.40%)(1ays(L)9(7.56%)) Hs(H)45(18.91%)(2a)(A)22(9.66%)(5u)(22)(8.40%)(2a)(5)(6)(19(7.56%)) Hs(H)45(18.91%)(2a)(A)22(15.37%)(2a)(2a)(2a)(2a)(2a)(2a)(2a)(2a)(2a)(2a	UNXVIII V	18(7.56%	1.Tyr(0)17(Z.	14%)		_
SI_Aki-098 SI_Aki-099 SI_Aki-100	Hs:15.43% Hs:17.76% Hs:26.47%	85-2 85-2 85-2	17 17 16	SWUSI10008100 none none	SWUSI0043110	2 1500 2 1500 2 1500	2697929 2701627 2707439	9629559 9702778 9808350	+ +	scattoid167 scattoid167 scattoid167	304941 302803 297061	306571 301722 196150	+ 6 - 3	18 2 75 1 63 2	105 124 120	3	HILI(128)[4134](34)(32(12.68)](34)(23(2.76)](34)(4)15(2.270); Tyr(1)5 Gly22(17.76)]HI15(15.76) Gly22(17.76)]HI15(15.76) Gly24(10.96)]HI154(24.69)	7.32%)(GA	(E)15(7.32	56)			
si_Aki-101 Si_Aki-102 Si_Aki-103 Si_Aki-104	HE24.32% HE25.13% HE25.37% HE14.44%	202 202 202 202 202 202 202 202 202 202	10 17 16	none none none		2 11910 2 11910 2 11910 2 11910	2712175 2715794 2721706 2726117	9717180 9722600 9728683	*	scattoid167 scattoid167 scattoid167 scattoid167	288706 282794 277837	287320 281900 278383			204 217 106	3	mmuy routh/H845(22.4%) Hig(H47(23.04%);Cal(221(10.29%);Cal(2)17(8.33%); Gb24(11.1%);His1(23.5%) M12(11.3%);Syn1(10.4%);His13(12.3%)			Ħ	-	f	
SLAki-105 SLAki-105 SLAki-107 SLAki-107	Hs:24.71% Hs:27.44% Hs:25.53% Hs:15.54%	88-2 88-2 88-2 88-2	16 17 0	none none SWUS10008098	SWU20043130	2 11/20 2 11/20 2 11/20 2	2727026 2733252 2737350	9727806 9734211 9738084	*	acaffold167 acaffold167 acaffold167 acaffold167	277474 271248 266416 26494*	276694 270289 267150 264337		73 1 46 1 26 1 21	190 181 141	3	Gb21(11;78),His43(22:3%) His(H)46(25:41%);Gbu(E)20(11.05%);Gby(G)13(7:18%);Asp(D)13(7:18%); His(H)46(25:51%);Gbu(E)19(13:44%);Gby(G)13(7:22%);Asp(D)11(7:80%);Gby(G)10(0011100;44%):011100;44%(016):01110;01110;01110;01110;010;Gby(G)10(7.09%);			=	=	_
SI_Aki-108 SI_Aki-109 SI_Aki-110 SI_Aki-111	Hs:15.50% Hs:19.27% Hs:13.51% Hs:14.50%	10-2 10-2 10-2 10-2	16 18 17 17	none SWUS10008097 none none	SWUSI0043140	2 mR00 2 mR00 2 mR00 2 mR00	4 2732557 4 2741232 4 2743860 4 2745733	9740163 9742572 9744830 9746282	* * *	scattold167 scattold167 scattold167 scattold167	264943 263268 259670 258767	264337 261928 260640 258218	· 3 • 3	21 1 84 1 24 1 63 1	27 28 20	2	Val11(10.4%),Jyn11(10.4%),Jiin14(13.2%) Highi)21(16.54%),JynX)18(14.17%),Val(V)12(9.45%),Gla(E)9(7.09%); Tyr(T)9(7.0 Ala13(10.2%),Val13(10.2%),Jiin15(11.7%) Val15(10.0%),Jyn15(10.2%),Jiin15(12.5%)	2%)			-	=	
S_Ak+112 S_Ak+113 S_Ak+114	Hs:14.77% Hs:15.31% Hs:14.77%	18-2 18-2 18-2	17 18 17	none SWUSI1000.8096 none	SWUSI0043160	2 1100 2 1100 2 1100	2743060 2750731 2755348	2742544 2751323 2755229	•	scattoid167 scattoid167 scattoid167	254956 253769 248571	255440 253107 249152	• 3 • 3	18 1 51 1 18 1	105 116 105	3	Val11(10.5%),Lyn10(10.5%),Jiln13(12.4%) Ha(H)15(12.93%);Val(V)14(12.07%),Lyn(K)11(9.48%);Ala(A)10(8.62%); Val12(11.4%),Lyn11(10.5%),Jiln13(12.4%) Val12(11.4%),Lyn11(10.5%),Jiln13(12.4%)						
S_AK+115 S_Ak+116 S_Ak+117 S_Ak+118	HE17.00% HE12.71% HE15.66% HE26.97%	10-2 10-2 10-2 10-2	17 16 23	none none none		2 m/200 2 m/200 2 m/200 2 m/200	2761025 2762461 2764501	9763124 9765421	• • •	acattoid167 acattoid167 acattoid167	242715 242715 242039 239999	2495561 243405 241376 239079	· 4 · 3	03 1 00 5 32 7	135 22 112	2	Val11(10.0%),2y810(3.0%),0813(14.7%) Val10(10.1%),2y811(11.1%),J8813(13.1%) Ly811(0.2%),4825(23.2%)				-		
SLAk+119 SLAk+120 SLAk+121 SLAk+121 SLAk+121	Hs:20.79% Hs:29.89% Hs:22.12% Hs:11.2%	19-2 19-2 19-2 19-7	18 25 17	none sone SWUSI10008092	SWUSI0043190	2 m/8% 2 m/8% 2 m/8%	2775422 2775422 2776313 278084	9767191 9776234 9779286 9781703	•	scattoid167 scattoid167 scattoid167 scattoid167	237892 229078 226187 223616	237309 228266 225214 227797	- 3	20 1 32 1 21 1 74 1	19 112 119	3	Ha(H)21(17.55%)(Va(V)13(10.32%);Gy(G)10(8.40%);Gu(E)10(8.40%);Sys(K) 9(Va112(11.0%);Hs27(24.1%) Ha(H)23(19.01%);Hs27(24.1%) Ha(H)23(19.01%);Hs21(20.32%);Gy(G)10(8.26%);Sys(K) 10(8.26%); Sys(12(0.3%);Hs11(10.3%);Hs22(11.2%)	7.56%);Ty	(Y) 9(7.56	86)	-	=	
S_Ak+123 S_Ak+124 S_Ak+125	Hs:17.05% Hs:17.05% Hs:25.0%	秋-2 秋-2 秋-2	17 12 18	none none none		2 11600 2 11600 2 11600	2783275 2786680 2788440	9784013 9787265 9785909	* * *	scattold167 scattold167 scattold167	221225 217820 216060	220487 217235 21559	- 3	24 1 24 1 18 1	107 107 105	3	Val10(2.4%)_Up(10(162.3%)/48.25(2.4%) Val10(2.4%)_Up(10(2.4%),08.25(2.4%) Val11(10.5%)_Up(10(16.5%)) Val11(10.5%)_Up(10(16.5%))/48.13(12.4%)				-		
S_Ak+126 S_Ak+127 S_Ak+128 S_Ak+128	Hs:23.81% Hs:22.97% Hs:21.35% Hs:20.90%	18-2 18-2 18-2 18-7	18 17 18	none none none		3 mRN 3 mRN 3 mRN 3 mRN 3 mRN	4 9792226 4 9794161 4 9796621 4 9796621	9792751 9794857 9797011 9800205	• •	scattoid167 scattoid167 scattoid167 scattoid167	212274 210339 207879 205094	211749 209643 207489 204295	4	02 1 98 1 24 1 05 1	102 165 107	3	Gy12(11.8%),Vall 2(11.8%),Aap10(9.8%),Ha20(19.6%) His-ich(20.6%),Lya(K)18(10.91%),Val(V)12(7.27%) Gy11(10.3%),Vall 5(14.6%),Aap12(11.2%),Ha15(17.8%) His-HU39270 2004;Gu1115(11.1984), Val(18/11.9%) His-HU39270 2004;Gu1115(11.1984), Val(18/11.9%)				=	=	_
SL_Aki-130 SL_Aki-131 SL_Aki-132	Ha:25.0% Ha:25.0% Ha:11.5%	88-2 88-2 88-2	17	none none none		2 118%	4 2800823 4 2805333 4 2805788	9801344 9805804 9807299	+ + +	scattold167 scattold167 scattold167	203607 199167 197712	203156 198696 197201	- 3	05 1 02 1 62 1	101 102 122	3	Gy11(10.9%),Vall3(12.9%),Aep11(10.9%),Jili21(20.8%) Gy11(10.8%),Vall2,Aep11(10.8%),Jili21(20.6%) Hit-sch38a14(11.5%)				_		
SLAk+133 SLAk+134 SLAk+135 SLAk+135	Hs:22.626 Hs:22.626 Hs:22.626 Hs:22.626	18-2 18-2 18-2 18-7	18 17 18	none none none		3 mRN 3 mRN 3 mRN 3 mRN	4 9809525 4 9811252 4 9813549 9815800	9810073 9812131 9814451 9816562	* * *	scattoid167 scattoid167 scattoid167 scattoid167	194975 193248 190951 187918	194427 192369 190049 188700		02 1 65 1 65 1	02 55 102 54	2	Gyl1 (10.8%), Aug11(10.8%), Jilis21(20.6%) Lyn17(11.0%), Jilis31 (20.9%) Val12(11.8%), Aug11(10.8%), Jilis19(18.6%) Benl/121113 (Addis) varXi 1912 1346-CarlE12711 (2007: ValVi11X/8.4496) Ser(5)13	(8.44%)			=	=	_
SL_Aki-137 SL_Aki-138 SL_Aki-139	Hsc13.1% Hsc12.42% Hsc13.33%	88-2 88-2 88-2	000	none SWUS10008089 SWUS10008088	SWUSI0043200 SWUSI0043150	2 11/200 2 11/200 2 11/200	4 2818757 4 2820240 4 2823079	9819124 9820978 9823953	+ + -	scattold167 scattold167 scattold167	185743 184260 180547	185376 183522 181421	- 3 - 4 + 4	62 1 62 1	122 153 165	3	Ala13(10,7%),3815(13,1%) Ala(A)33(21,57%),Va(V)21(13,73%);580(0)(19(12,42%);587(5))2(7,84%);970(7)10 Ala(A)34(20,61%);180(022(13,33%);Va(V)21(12,73%);587(5)18(10,91%);	6.54%)					
S_A6+140 S_A6+141 S_A6+142 S_A6+143	Hs:12.05% Hs:10.74% Hs:5.59% Hs:10.34%	10-2 10-2 10-2 10-2	17	SWUSITOCOBOB/ SWUSITOCOBOBS SWUSITOCOBOBS none	SWUSD043210 SWUSD043230 SWUSD043230 SWUSD043230	2 π/00 2 π/00 2 π/00 2 π/00	4 2837038 4 2840489 4 2840489	2637924 2641287 2644243	•	scattoid167 scattoid167 scattoid167 scattoid167	167462 164011 160893	177409 166576 163213 160257	* 3 - 5 - 3 - 4	01 1 43 1 65 1	166 180 154	3	ABI(AF3)(23.509)(241)(241)3.119(310)(120(0.0339)(2413)16(0.748)(270))12 ABI(A352)(21.039)(241)(25(12.039)(5413)17(10.248)(240)(16(0.548)) ABI(A452)(2000)(241)(25(11.119)(340)(15(0.539)(571))17(1)47,759)(270)(11) ABI(A)(27(10.628)(241)(10(12419)(340)(15(10.349)(545))17(1)41)(2410)(270)(11)	7.22%)(Set 7.59%)(11	(5)13(7.22 (7.59%)	%i);	=	=	_
S_Ak+144 S_Ak+145 S_Ak+145 S_Ak+146	Hs:10.00% Hs:10.33% Hs:0.29%	数-2 数-2 数-2	0000	SWUS10008085 SWUS10008085 SWUS10008084	SWUSI0043230 SWUSI0043230 SWUSI0043240	2 mRN 2 mRN 2 mRN	A 2846292 A 2851576 A 2855178	9847177 9852466 9855983	+ + +	scattold167 scattold167 scattold167	158208 152924 149322	157323 152034 148517	4 0 0 0	83 1 55 1 52 1	160 184 183	3	Alar(A)32(20.00%);Ser(S)20(12:50%);Val(V)19(11.88%);Ha(H)16(10.00%);Tyr(Y)1 Alar(A)42(22:E3%);Val(V)22(11.36%);Ser(S)19(10.33%);Ha(H)19(10.33%);Pho(P)1- Alar(A)42(23:00%);Ser(S)22(12:02%);Val(V)20(10.93%);Ha(H)17(P)29%);Pho(P)13 Alar(A)42(23:00%);Ser(S)22(12:02%);Val(V)20(10.93%);Ha(H)17(P)29%);Pho(P)13	2(7.50%) k(7.61%)(1 (7.10%);	yr(Y)13(7.0	07%)	-		
S_Ak+145 S_Ak+145 S_Ak+149 S_Ak+150	HE10.34% HE7.85% HE9.38% HE15.27%	10-2 10-2 10-2 10-2	9 17 17 17	SWUST0003082 SWUST0008081 SWUST0008080	SWUSIO043260 SWUSIO043270 SWUSIO043280	2 πρου 2 πρου 2 πρου 2 πρου	2862334 2862439 2862750	9861111 9863632 9870613	• • •	scattoid167 scattoid167 scattoid167	144106 142061 134750	140868 140868 133887	· 6	20 27 2 42 1 47 1	45 103 148	2 2 3 3	AllicA32222.079E341(321741122000) AllicA327(27.40%);Val(V2811334%);Pro(P)16(7.69%);Sul(U115(7.21%)(7)Y1Y115(2)y1X1018.83%);Aap(D)07.30%(318))(017.90%);TyrY19(7.50%);Gy(C617.01%); AllicA323(15.54%);Sul(U20(13.51%);Sur(S) 11(7.43%);Pro(P)11(7.43%);	.21%) Se(()8(7.0	9%);Gu(E)	8(7.08%);	=	=	
SL_Aki-151 SL_Aki-152 SL_Aki-153 SL_Aki-153 SL_Aki-154	Hs:10.26% Hs:0.83% Hs:6.7%	19-2 19-2 19-2 19-7	17 0 0 21	SWUS10008079 none none SWUS10008077	SWUSI0043290	2 m/00/ 2 m/00/ 2 m/00/ 2 m/00/	A 28577945 A 2850360 A 2857291 A 2857291	9681280 9681280 9688615 9688615	• • •	scattold167 scattold167 scattold167 scattold167	126555 124140 117209 107829	124307 123220 115885 107171		22 1 27 1 15 1 01 1	73 158 104	3	Val(V)19(10.98%);Gin(Q)18(10.40%);Ha(H)(17(9.83%);Aa(A)16(9.25%); Ala22(21.3%);Val19(17.6%);Peo13(12.0%);Ala13(12.0%) Aap12(11.5%);J4la7(6.7%) Nov/V2122.5%);Sel7516(9.648);Pev/P015(9.048);Ter/P115(9.048);Fal(A)13(2	8 1867 Chris	512(7.23)	4)	-	-	_
S_Ak+155 S_Ak+156 S_Ak+157	Hs:15.80% Hs:14.48% Hs:14.52%	18.2 18.2 18.2	17 17 17	SWUSI10008076 none SWUSI10008075	SWUSI0043320 SWUSI0043320 SWUSI0043330	0 1100 0 1100 0 1100	2905769 2910364 2913755	9908596 9911098 9914371	• • •	scattoid167 scattoid167 scattoid167	98731 94136 90745	97904 93402 90129	- 4	52 1 52 1 26 1	62 62 41	4 4 3		02%).9ro 7.80%);	7)12(7.41	8)			
SI_Aki-158 SI_Aki-159 SI_Aki-160 SI_Aki-161	Hs:13.40% Hs:13.40% Hs:11.72% Hs:9.7%	10-2 10-2 10-2 10-2	17 17 17 17	none none swusi10008074	SWUSI0043330	2 mRN 2 mRN 2 mRN 2 mRN	A 2917125 A 2921102 A 2924175 A 2928167	9917835 9921890 9924810 9928948	+ + +	scattold167 scattold167 scattold167 scattold167	87375 83398 80325 76333	86665 82610 79690 75552		11 1 22 1 38 1 56 1	36 173 145	4 3 3	HB(H)25(16.31%)/AB(A)16(11.70%)/VB(V)12(8.82%)/AB(D)12(8.82%)/AB(D)12(8.82%)/AB(D)12(8.82%)/AB(D)12(8.87%)/AB(A)14(8 AB(A)3(10.03%)/AB(H)(121(2.14%)/VB(V)15(8.67%)/YF(Y)15(8.67%)/AB(1)14(8 AB(A)28(118.54%)/AB(1)(121(2.8%)/AB(1)(12.0%)/AB(1)(12.0%)/AB(1)(14(8.27%)/AB(1)(11)(12.8%)/AB(1)(11)(12.8%)/AB(1)(11)(12.8%)/AB(1)(11)(12.8%)/AB(1)(11)(12.8%)/AB(1)(11)(11)(11)(11)(11)(11)(11)(11)(11)	1.09%) 1.09%) 5(8.61%):5	a(v)11(7,3	20%).Pro(P)1	1(7,20%)	-	
S_Ak+162 S_Ak+163 S_Ak+164	Hs:13.33% Hs:16.48% Hs:26.4%	18.2 18.2 18.2	17	none SWUSI10008073 none	SWUSI0043350	2 π/80 2 π/80 2 π/80	2932158 2938646 2950409	9932832 9939585 9951839	• • •	scattoid167 scattoid167 scattoid167	72342 65854 54091	71668 64915 52661	- 5	54 1 52 1 26 1	67 93 141	3 4 3	VEX/V2.0(1.EX806346(0)20(1.5386);555(5)14(8.5386;557(6)12(7.1386;555(7)12); AB(A)42(21,7696346(0)20(1.5386);738(7)26(13.478()556(7)18(9.3386) Pto15(10.666);48:34(24.18)	.19%)					
SL_Aki-165 SL_Aki-165 SL_Aki-167 SL_Aki-169	Na:30.14% Na:5.06%	10-2 10-2 10-2 10-2	28 22 16	none none none		2 11970 2 11970 2 11970 2 11970	2065183 2067785 2083764	99553334 9955319 9959149 9957488	* *	acattoid167 acattoid167 acattoid167 acattoid167	49605 39317 35351 20736	46566 38181 36715 17012	· · ·	63 1 25 1 66 1 401 4	120 174 121 166	3	Har-RC204110(15.0%) Har-RC20445(25.5%) Abs25(18.2%),Val12(2.5%),Hat8(6.6%) Abs25(18.2%),Val12(2.5%),Hat8(6.6%)						_
S_Aki-170 S_Aki-171 S_Aki-172 S_Aki-172		BB-1 DPAP1-H DPAP1-F DPAP1-F	1800	SWUS10014762 10012959 10012958 10012955	SWUSIO043920 SWUSIO045030 SWUSIO045040 SWUSIO045040	2 mRN 10 mRN 10 mRN 10 mRN	4 11822991 4 508715 4 590758 4 590758	11829529 824379 898159 991745	*	scattold24 scattold54 scattold54 scattold54	870213 641474 567694 474108	863675 657138 575035 478785	· · · ·	206 4 240 1 14 1 16 7	101 1079 137	7 7 4 4 4	GMS(1128)Pro45(1128)Ser37(928)Aup(038(948%) CBM_14[118-170]CrHD2[116-162] CBM_14[35-105]CrHD2[33-103] CBM_14[65(270)CrHD2[65-243]				-	=	
SI_Aki-174 SI_Aki-175 SI_Aki-176	Ha:8.7% Ha:9.13%	CPOIC CPAP1-D CPH	14 17 15	SWUS10009335 10009369 SWUS10008269	SWUSI0045730 SWUSI0046060 SWUSI0046750	10 m/900 10 m/900 10 m/900	A 2398790 A 3059121 A 6368305	2329979 3063979 6369255	+ - +	scatfold132 scatfold132 scatfold161	33135 693466 41460	34324 698324 42410	+ 2 - 3 + 7	61 8 30 1 71 2	96 109 236	3 4 2	Ala15(17:4%),Pro18(20.9%) CBM_14 (49-100); ChtBD2 (47-96) Ala95(35.1%), Val24(9.4%),Jis22(8.6%)						
SL_Aki-177 SL_Aki-178 SL_Aki-179 SL_Aki-180	Hs:16.84% Hs:6.52% Hs:10.08%	014 054 88-2 88-1	15 0 18 17	SWUS10008270 SWUS10008271 none SWUS10008643	SWUSI0046760 SWUSI0046770 SWUSI0049030	10 mRN 10 mRN 11 mRN 11 mRN	6388849 6413900 5836636 5566328	6390276 6414472 5837922 5567025	* * *	scattoid161 scattoid161 scattoid150 scattoid150	62004 87035 578536 308228	63431 87627 579822 308925	+ 5 + 5 + 3	64 - 2 75 - 1 42 - 1 30 - 1	137 190 113 109	2 3 3	And5(22,7%), Val30(10,5%) An354(28,4%) Gy17(15,0%), Val13(11,5%), Jun16(14,2%), Tyr15(13,3%); An12(11,0%), Gin11(10,1%), Gu2(6,4%); Val(V)9(6,20%)				=	=	_
SL_Aki-181 SL_Aki-182 SL_Aki-183		0% 88-1 071	18 0 21 0	SWUS10012743 SWUS10006987 SWUS10002405	SWUSIO051830 SWUSIO057380 SWUSIO062550	11 mRNA 13 mRNA 14 mRNA	A 12161241 A 2126861 A 7901834	12164064 2131989 7905025	*	scattold56 scattold204 scattold488	815249 231198 187235	818072 236326 184044	* *	44 1 16 1 73 1	147 171 190	1 6 1	Val30(20.4%),Pro25(17.0%),Lys15(10.2%),Hs13(8.8%) Ghz4(14.0%),Ala5Leu16(9.4%),Pro16(9.4%) Pro42(22.1%),Ala(J20(10.53%),Val(V19(10.00%);Ser(5) 15(7.89%) Ser(115):1%(50007222-2000127211),2000127211,200012721,200012721,200012721,200012721,20000000000				-		
SL_Aki+185 SL_Aki+185 SL_Aki+185 SL_Aki+187	11-0.799	18-2 18-2 18-2	0 17 17 17 17	SWUSI10010664 SWUSI10012083 none	SWUSI0067230 SWUSI0070260	15 mRN	4870821	4871625 11246561 11311066	* - *	scattold102 scattold70 scattold70	1058120 284228 348184	1058224 283508 348733	* 5 - 4 * 5	07 1 05 1 75 1	168 134 24	2	Gby(G)34(20.24%)/Abs(A)2G(15.48%)/Leu(L)1G(9.52%)/Va(V) 15(8.93%) Abs28(20.9%)/Tur13(9.7%)/Sec(5)13(9.7%)/ Abs28(20.2%)/Tyr12(9.7%)/Va(V)10(8.09%)						
S_Ak+189 S_Ak+190 S_Ak+191	Hs:10.38% Hs:9.71% Hs:8.77%	18-2 18-2 18-2	17 17 17	none SWUSI10012085 SWUSI10012086	SWUSI0070290 5 SWUSI0070300	15 mRNJ 15 mRNJ 15 mRNJ	11317727	11318725 11323823 11326311	* *	scattold70 scattold70 scattold70	355324 360507 363978	356392 361490 362756	- 6 - 5	03 2 79 1 050 3	192 159	2 2 2	5024 (12.0%) Alad (26.7%) Val20 (12.0%) Ala (19.0%) Val20 (12.0%) Alad (12.0%) Val20 (12.0%) Alad (12.0%) Val20 (12.0\%) Val20 (1				-	-	
SL Aki-192 SL Aki-193 SL Aki-194 SL Aki-194 SL Aki-194		88-2 88-2 87-2 07AP1-G	19 17 17 19	SWUS10012087 SWUS10012088 SWUS10012089 10010858	SWUSI0070310 SWUSI0070320 SWUSI0070330 SWUSI0070210	15 m8%	11328621 11337706 11340851 12761707	11330078 11338443 11341695 12814695	*	acaffold70 acaffold70 acaffold70 acaffold70	366288 375373 379382 436018	367745 376110 378518 479661	-	28 2 25 1 52 1	275 174 83 515	2 2 2 6	GbyG948(17.45%);Ala(A)44(16.00%);Aan(N)(28(10.55%);Val(V)24(8.73%);Ser(5)24 Ala(A)49(28.16%);Val(V)28(16.02%);Pro(P)13(7.47%) Ala(A)45(29.51%);Val(V)28(16.02%);Pro(P)14(7.65%);Tyr(7)13(7.10%) CBM:14(157.470);ChtB2(115.20%)	(8.73%)					_
S_Ak+195 S_Ak+197 S_Ak+198		07G 07G 88-1	21 19 18	SWUS10003001 SWUS10003600 SWUS10013174	SWUSD071940 SWUSD072780 SWUSD073200	16 m/00 16 m/00 16 m/00	2215896 4701149 5333418	2217128 4703004 5336193	+ +	scattold140 scattold388 scattold50	632994 172820 471221	634226 174675 465446	+ 6 + 7	20 2 20 2 94 1	219 239 197	2 2 4	QGY repeat.Siy42(12.2%),Ty729(13.2%) Siy53(22.2%),Leu25(10.5%) Tyr20(11.2%),Lay15(0.6%)						
S_Aki-200 S_Aki-201 S_Aki-202		53 51 52 52	0 10 17	SWUSI10008912 SWUSI10002192 none	SMUSICO83940 SMUSICO83940	18 m8%	4888873 4888873 1179283 5079545	4890108 1179902 3080215	É	acaffold143 acaffold517 acaffold126	139508 112910 650506	140743 112291 649836	+ 1	194 2 37 1 85 1	197 178 194	2 2 2	mpon					#	
S_Aki-203 SL_Aki-204 SL_Aki-205 SL_Aki-205 SL_Aki-205	Har6.78% Har6.55% Har6.92% Har6.92%	18-2 18-2 18-2 18-2	17 17 17 17	none none none		20 11896 20 11896 20 11896 20 11896	3090897 3093616 3095965 3095965	3091546 3094237 3097603 3100564		acattoid126 acattoid126 acattoid126 acattoid126	661837 664528 667894 670855	661185 663907 667256 670214		34 1 31 1 31 1 58 1	176 176 185	2 2 2 2 2	ABA42.cc.286, Val31(16.896, Tyr17(8.296); Pto(P)17(9.1996) ABA57(21.0296,Val(19)29(16.4496,(5yr1(1)16(9.0996);Pto(P)15(8.5296) Aba(A33B(21.5996); Val(19)29(16.4496); Tyr(T)16(9.0996); Pto(P)15(8.5296) Aba(23.2396); Val29(15.796); Tyr17(9.286); Stefen(P)16(8.5296); Aba(23.2396); Val29(15.796); Tyr17(9.286); Stefen(P)16(8.5296);				T	Ŧ	
SL_Aki-207 SL_Aki-208 SL_Aki-209 SL_Aki-209	Hac6.55% Hac6.85%	89-2 89-2 88-1 89-1	17 17 16	SWUSI10009649 SWUSI10009650 SWUSI10010969	SWUSIO89540 SWUSIO089550 SWUSIO080520	20 m8%	3101400 3103335 6877778	3102013 3110438 6881160	*	acaffold126 scaffold126 scaffold23 scaffold23	671691 679686 1139422 1148243	472304 480729 1142804		92 1 13 2 67 1	163 270 38	2	Ala54(20, 9%), Val24(14,7%), Tyr14(8,6%) Ala75(27,8%), Val38(13,3%), Tyr24(8,9%): Pro(7)24(8,82%) Pro(7)17(0,24%), Tyr2(116(3,31%), Cala(A)), 5(7,33%), Thr(7)15(7,98%), Cala(3)15(7,33%), Cala(3)15(7,33\%), Cala(3)(7,33\%), Cala(3)(7,33\%), Cala(3)(7,33\%), Cala(3)(7,33\%), Cala(3)(7,33\%), Cala(3)(7,33\%),	9%)Lys(K)	15(7.98%				_
S_Aki-212 S_Aki-213 S_Aki-213	Hs:10.47% Hs:7.35% Hs:6.34%	88 88 88	19 19 18	SWUSI10015397 SWUSI10015395 SWUSI10015395	SNUSIO296450 SNUSIO296470 SNUSIO296480	21 mRN 21 mRN 21 mRN 21 mRN	1303529 1303529 13037502 13076032	13005223 13038906 13077578	Ē	acaffold15 acaffold15 acaffold15	1203668 1162985 1131313	1205362 1171389 1132859	+ 1 + 1 + 1	206 4 050 3 025 3	101 559 565	2 2 2	var77(19.2%),/hes/10.2%),/wi32(10.2%) var77(19.2%),/hes/10.2%),/wi32(10.2%) var72(21.2%),/hes/12(20.2%),/wi32(10.2%) var72(21.0%),/hes/1(19.5%)				f	₫	
SL Aki-214 SL Aki-215 SL Aki-216 SL Aki-216 SL Aki-216		88-1 88-1 88-1 (DAP1.8	16 16 16	SWUSI10015359 SWUSI10015358 10015690	SWUSI0296800 SWUSI0296810	21 m6%	13639352	13640534 13645094 13658432 455817*	* * *	acaffold15 acaffold15 acaffold15 acaffold15	568239 564379 551725 773261	568357 563797 550459 76819#		21 1 21 1 79	106 106 106	3	Ala17(16.0%),Val12(11.3%) Ala17(16.0%),Val12(11.3%),GAy(5)10(9.43%) Ala12(11.3%),Val12(15.1%) Mal12(11.3%),Val12(15.1%) Mal12(10.5%),Chill2(19.5.7%)					=	
S_Aki-218 S_Aki-219 S_Aki-220		CPAP1-82 88-2 CPG	21 0	10015691 SWUS10010676 SWUS10001933	SWUSD098950 SWUSD100940 SWUSD101330	22 mRN 22 mRN 23 mRN	4566167 12753202 11207	4590078 12765491 11838	+	scaffold11 scaffold101 scaffold550	796217 124784 11838	820128 137073 11207	+ 2	126 7 23 2 42 1	731 265 182	5 3 2	CBM_14 [SE1-613]; CHEBD2 [S39-609] Lyx30(11.3%) Gly60(33.0%),Val22(12.1%),Ser30(16.5%),Tyr17(9.3%)						
S_Aki-221 S_Aki-222 S_Aki-223 S_Aki-223 S_Aki-224	Hg15.11%	0% 0% 0%	16 16 16	5008 500510001935 500510001936 500510001937	SWUSID101340 SWUSID101350 SWUSID101360	23 mRN 23 mRN 23 mRN 23 mRN	15685 20737 25070 45858	16304 21356 25987 47295		acaffold550 scaffold550 scaffold550 scaffold550	16304 21356 25987 47295	1:685 20737 25070 45858		37 1 37 1 84 2 245 4	178 127 114	2 2 2 2	segse(2.c.498),Val22(12.498),Ser29(16.398),Tyr16(9.0798) Gy55(12.268),Val22(12.496),Ser29(16.398),Tyr16(9.0798) Gy54(15.098),Au27(11.398),Val27(11.398),Ser29(18.398),Tyr22(9.198) Au452(10.198),Lau54(12.496),Ptp-48(11.098),Ser111(26.398),Hat60(14.598)				Ŧ	Ŧ	_
S_Ak+225 S_Ak+226		676 676	0 18	SWUS10008448 SWUS10008453	SWUSID102380 SWUSID102410	23 mRNo 23 mRNo	A 2563635 A 2641871	2573347 2646309	•	acattoid155 acattoid155	431339 509395	441071 514033	• 1 • 6	245 4 84 2	114	2 2	Va)(V)78(18.94%);Gy(G)59(14.25%);Pro(P)31(12.32%);Ser(S)36(8.70%);Ala(A)31 G)(G)36(15.86%);Va(V)36(15.86%);Pro(P)33(14.54%);Se()(25(11.01%);Ser(S)18 (7.93%)	(7.49%)					
2 Aki-227 2 Aki-228 2 Aki-229 2 Aki-230	Hac6.63% Hac6.65% Hac14.92%	38.85	13 19 15	SWUS10015582 SWUS10015583 SWUS10015585 SWUS10006738	WUSID103030 SWUSID103040 SWUSID103060 SWUSID104050	23 m8% 23 m8% 23 m8% 23 m8%	4707065 4713522 4724706 7105557	4714673 4725829 7106143		scattoid12 scattoid12 scattoid12 scattoid214	1724223 1736530 115038	1725374 1735407 114452	-	10/ 2 053 2 91 1 37 2	126 278	2	ary					f	
S Aki-231 S Aki-233 S Aki-234 S Aki-234		DVL DVL	18 0 17	SWUST0011572 SWUST0013852 SWUST0011525	SWUSID109480 SWUSID110720 SWUSID112390	24 millio 25 millio 25 millio 25 millio	11506777 2760059 8271643	11509768 2785770 8272592	*	acaffold83 acaffold83 acaffold80 acaffold80	613494 1463141 929084 942794	610503 1488852 930033 945052	7	047 3 32 2 36 2	148 143 111	6 4 2 1	ChyG/46 (13.22%)(2ys)(X)35 (10.00%);Met(M)29 (8.33%) ChyG/40 (16.46%);Aha10 (12.35%);Pro/P(20(8.25%); Aha61 (36.4%);Ae20 (0.5%);Tyr26 (12.3%);Pro2 (10.0%) Chy1(21.2%);Aha12 (21.3%);ChyC26 (12.3%);Pro2 (10.0%)				-	=	_
SI_Aki-236 SI_Aki-237 SI_Aki-238		CPTL CPTL CPTL	17	SWUS10011528 SWUS10011529 SWUS10011529	SWUSID112410 SWUSID112420	25 m8%	8291897 8298557 8298557 8298557	8293040 8297773 8299981	*	scattold20 scattold20 scattold20 scattold20	949338 953998 936488	950481 955214 957422	+ 9 + 7 + 6	42 2 47 2 42 2	113 148 113	2	Aba21(23,7%),Va244(14,1%),Ser45(14,4%),Tyr22(12,2%) Aba21(23,7%),Va244(14,1%),Ser45(14,4%),Tyr22(12,2%) Aba21(23,7%),Va22(12,7%),Ser22(11,7%),Tyr22(11,7%) Aba23(23,7%),Va22(12,7%),Ser22(12,4%),Tyr22(12,4%) Aba24(12,7%),Va22(12,7%),Ser22(12,4%),Tyr22(12,4%) Aba24(12,7%),Va22(12,7%),Ser22(12,4%),Tyr22(12,4%) Aba24(12,7%),Va22(12,7%),Ser22(12,4%),Tyr22(12,4%) Aba24(12,7%),Ser22(12,4%),Tyr22(12,4%),Tyr22(12,4%) Aba24(12,7%),Ser22(12,4%),Tyr22(12,4\%),Tyr22(12,4						
S_A6-239 S_Aki-240 S_Aki-241 S_Aki-242			17 17 17 17	SWUS10011535 SWUS10011531 SWUS10011533 SWUS10012442	SMUSID112430 SMUSID112440 SMUSID112460 SMUSID113230	25 m8% 25 m8% 25 m8%	4 8304931 4 8317557 4 10832710	\$301751 \$306005 \$317604 10835676		acattold80 acattold80 acattold80 acattold82	963446 975045 905683	962372 974938 908649		2 2 13 2 13 2 245 4	270 270 114	2	ner_s(a.c.m, VBC2(1.1.2%),28832(1.7.2%),7924(11.5%) Alb2(20.64%),V430(1.1.1%),5844(10.5%),7931(11.5%) Alb2(21.5%),583(12.2%),5843(11.5%),7924(0.5%) Gb(GB2(21.15%),5642(13.7%),2483(11.5%),796(2.3%)	70%)			f		
5 Aki-243 5 Aki-245 5 Aki-246 5 Aki-246	1	88-2 CPH 88-2 88-2	0 18 19 0	SWUS10003648 SWUS10003718 SWUS10012781	SWUSID114190 SWUSID114520 SWUSID115480 SWUSID115480	26 m800 26 m800 26 m800 26 m800	GE4597 2447545 6259743 6259743	695005 2466941 6267836 628769	-	acaffold260 acaffold383 acaffold53 acaffold53	50866 211801 20194 35045	61274 231197 12101 31274		458 4 291 1 96 2 623	1096 231 240	5	Ann(N)43(8.87%);Ser(5)32(8.04%) Val122(12.0%),Pro110(10.0%) Ex(G3)1(13.42%);Ala(A)30(12.39%) Ex(G3)1(13.42%);Ala(A)30(12.39%)	_			=	=	_
AL_A01247		18-2	17	SWUSI10014164 SWUSI10014158 SWUSI10014158	SWUSID115940 SWUSID115940 SWUSID115980 SWUSID116010	26 m8% 26 m8% 26 m8%	. 02/8916 4 8487170 4 8796567 4 8865333	\$502996 \$502996 \$800969 \$867139	*	acelloid22 acelloid22 acelloid22	1313699 1020128 951362	1329525 1015726 949556	+ 4	14 1 12 2 72 1	34 903 123	3 6 3	nn (2), - (2), (2), (2), (2), (2), (2), (2), (2),	4) 27(8.91%)	Tyr(Y)23(7	59%)	f	≣	
S_Aki-249 S_Aki-252	Ha:14.17%	19.2 19.2	0	and a second second			_	10545555	-	acaffold33 acaffold377	2/1180 147573	2/9054 165443 907140		500 4 52 1 07 8	23 23 24 25	4 2 2	um(2)43(6.62%) CBM_14_(39-139); CHB02_(85-133] Tur(1)1725.00%) CBV(611217.65%) Sec(5)10(14.71%); Ala(A)7(10.29%) Leu(L)6(Ab3114.45%) VIS115.85%; Pro/P105.47%; Curr/Colo a 2%;	8.82%);		- T		_ T	_
S_Aki-240 S_Aki-252 S_Aki-253 S_Aki-254 S_Aki-255 S_Aki-255 S_Aki-255	Ha:14.17%	18-2 18-2 DAP1-A DH DH	0 23 25 17	SWUS10014139 10003769 SWUS10011623 SWUS10011625	SWUSI0116170 SWUSI0120830 SWUSI0124530 SWUSI0124500	26 π8% 27 π8% 28 π8% 28 +8*	4 2537631 12342242 4 7547384 762359	12363112 7550278 7625843	*	scattoid79 scattoid79	910034 832125	831575		P			the second s				-	-	_
2 Aki-240 2 Aki-252 2 Aki-253 2 Aki-254 2 Aki-255 2 Aki-255	Ne14.17%	88-2 88-2 DAP1-A DB1 DB1 DB1 DB1 DB1 DB1 DB1 DB1 DB1 DB1	0 23 25 17 17 16 17 18	SWUS10014139 10003769 SWUS10011623 SWUS10011678 SWUS10011678 SWUS10011677 SWUS10011677	SMUSIO116170 SMUSIO120830 SMUSIO124530 SMUSIO124450 SMUSIO124450 SMUSIO124460 SMUSIO124670 SMUSIO124680	26 1995 27 1995 28 1995 28 1995 28 1995 28 1995 28 1995 28 1995	A 2537631 A 12342242 A 7547384 A 7625253 A 7631214 A 7636447 A 7642810	12363112 7550278 7625843 7632834 7636946 7636946	+ + + + +	scattoid79 scattoid79 scattoid79 scattoid79 scattoid79	910034 832125 826204 820971 814608	831575 824584 820472 813725	- 1	536 5 17 1 96 2	511 138 231	2	Pro109(21.3%),GloS0(17.6%): Val(V)46(9.00%) Ala52(37.7%),Val20(14.5%),Pro23(16.7%) Ala62(05.0%),Val28(12.6%),Pro23(25.5%) Ala62(05.0%),Val28(12.6%),Pro53(25.5%)						
2 Aki-249 2 Aki-232 3 Aki-233 2 Aki-233 2 Aki-235 2 Aki-255 2 Aki-255 2 Aki-255 2 Aki-255 2 Aki-255 2 Aki-255 2 Aki-259 2 Aki-259 2 Aki-260 2 Aki-261 2 Aki-263	Hx14.17%	88-2 88-2 07AP1-A 07H 07H 07H 07H 07H 07H 07H 07H 07H 07H		SNUS10014139 10003769 SNUS10011693 SNUS10011679 SNUS10011679 SNUS10011677 SNUS10011677 SNUS10011677 SNUS10011672	SMUSID116170 SMUSID120530 SAUSID124530 SMUSID124530 SMUSID124520 SMUSID124520 SMUSID124520 SMUSID124520 SMUSID124720 SAUSID124720	26 1890 27 1890 28 1800 28 1890 28 1800 28 1800000000000000000000000000000000000	A 2537631 A 12342242 A 7547384 A 7625233 A 76356447 A 7642810 A 7642810 A 7645313 A 7646878 A 7640315	12363112 7550278 7625843 7632834 7636946 7643693 7646111 7647931 7667591 767591	+ + + + + + + + + + + + + +	icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73 icaffoid73	910034 832125 826254 826254 814608 811367 810300 780540 787103	831575 824584 820472 813725 812105 809487 789827 789827	- 4	536 5 17 1 96 2 01 1 33 1 14 2 97	511 138 231 166 110 237 237	2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Not 05(21:35), (2560) (7:36), (Val(V)46(2:056)) MASC177, 7% (Val(V)45(2:056)) MASC172, 7% (Val(V)45(2:056)) MASC172, 7% (Val(V)45(2:056)) MASC122, 7% (96);Leu(L) 3998)	17(7.17%)				
9 A45-249 9 A45-252 9 A45-252 9 A45-253 9 A45-253 9 A45-253 9 A45-253 9 A45-255 9 A45-255 9 A45-255 9 A45-262 9 A45-262 9 A45-262 9 A45-263 9 A45-265 9 A5-265 9 A5-265 9	4x14.17%	88-2 88-2 DAPL-A DBH DBH DBH DBH DBH DBH DBH DBH DBH DBH		SMUST0014139 10003769 SMUST0011633 SMUST0011635 SMUST0011679 SMUST0011677 SMUST0011677 SMUST0011677 SMUST0011677 SMUST0011673 SMUST0011670 SMUST0011670 SMUST0011670 SMUST0011670	DAUSID116170 DAUSID120830 DAUSID124530 DAUSID124530 DAUSID124530 DAUSID124530 DAUSID124530 DAUSID124530 DAUSID124530 DAUSID124530 DAUSID124730 DAUSID124730 DAUSID124740 DAUSID124740 DAUSID124740 DAUSID124740	26 1000 27 1000 28 10000 28 10000 28 100000 28 10000 28 10000 28 10000 28 10	A 2537631 A 12342242 A 7247134 A 7625233 A 7625231 A 7625231 A 7642810 A 7642810 A 7642810 A 7642813 A 7642813 A 7642813 A 7642124 A 7672515 A 7622124 A 7622329 B 96675	12363112 7550278 7625843 7632844 7635946 7646033 7646111 7647331 76473	+ + + + + + + + + + + + + + + + + + +	acattology acattology acattology acattology acattology acattology acattology acattology acattology acattology acattology acattology acattology	21 0024 852125 82 6254 82 6254 81 606 81 1 307 81 1 307 81 1 307 78 1 0300 77 0540 78 7 05 78 4247 78 1 205 77 50 19 77 50 19 77 18 14	831575 824584 820472 813725 812105 802487 785507 785507 785294 781903 774555 774555		536 5 17 1 96 2 01 1 33 1 14 2 27 1 63 2 92 2 29 2 65 1	511 36 51 166 10 10 137 198 120 122 154 155	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Nex (592) 273 (Galacity 17 28), 1992 (Galacity 17 28), 1992 (Galacity 17 28), 1992 (Galacity 17 28), 1992 (Galacity 18), 1992	96)(cma(t)) 5996) 1091	17(7.17%)				
0 Alsi-249 0 Alsi-252 0 Alsi-252 0 Alsi-252 0 Alsi-253 0 Alsi-254 0 Alsi-254 0 Alsi-254 0 Alsi-255 0 Alsi-255 0 Alsi-253 0 Alsi-253 0 Alsi-253 0 Alsi-252 0 Alsi-252 0 Alsi-252 0 Alsi-262 0 Alsi-264	Ha:14.17% Ha:15.23% Ha:15.23%	85-2 85-2 85-2 85-2 85-2 85-2 85-2 85-2		SNUST 1001 4139 1000 3749 SNUST 1001 1623 SNUST 1001 1623 SNUST 1001 1629 SNUST 1001 1679 SNUST 1001 1679 SNUST 1001 1679 SNUST 1001 1679 SNUST 1001 1679 SNUST 1001 1679 SNUST 1001 1669 SNUST 1001 1669 SNUST 1001 1669 SNUST 1001 1669	WHISED 16 170 WHISED 20050 WHISED 24520 WHISED 24620 WHISED 24620 WHISED 24620 WHISED 24620 WHISED 24620 WHISED 24620 WHISED 24720 WHISED 24720 WHISED 24720 WHISED 24720 WHISED 24720 WHISED 24720	25 16%	A 2337631 A 12342424 A 1244242 A 7547264 A 7547264 A 762523 A 763524 A 763524 A 7642810 A 7645313 A 7645713 A 7646713 A 7666878 A 765515 A 765515 A 7655250 A 7685200 A 768500 A 771877	12363112 75550278 7625543 7625245 76326346 7645093 7645093 7647291 7667291 7667291 767591 767591 767591 7675213 7675213 7675213 7675213 7695280 7691265 7719748	* * * * * * * * * * * * * * * * * * *	icatio(7) icatio(7)	21 0034 82 12125 82 6254 82 60271 81 4606 81 1307 81 0300 78 7103 78 7103 78 7103 78 726 77 5019 77 1814 77 81 38 76 61 53 76 61 53 77 54 44 76 61 53	831375 824584 820472 813725 813725 802407 785507 785507 785507 785507 785507 785507 785503 774555 771347 768605 766748 766769		536 9 17 1 96 2 01 1 14 2 97 1 63 2 92 2 65 1 66 1 66 1 97 3 98 1 97 1 98 2 99 2 99 2 99 2 99 2 90 2	511 38 231 166 110 137 238 237 238 237 238 235 235 235 235 235 235 235 235		Mon 2021 TS (a)	9%)(Leu(L) 39%) 10%) 7.10%) 8(2.09%);	17(7.17%) Ha(H)15(7	53%)			
g Aib 240 g Aib 242 g Aib 252 g Aib 252 g Aib 252 g Aib 254 g Aib 255 g Aib 256 g Aib 262 g Aib 262 g Aib 265 g Aib 265 g Aib 267 g Aib 277 g Aib 277 g Aib 277 g Aib 277 g	4x:14.17% 4x:15.23% 4x:15.23% 4x:15.25% 4x:15.25% 4x:15.25% 4x:15.25%	822 822 827 827 827 827 827 827 827 827		Integration of 41.39 Integration of 41.39 Integration of 10.23 Integration of 10.23 Integration of 10.23 Integration of 10.23 Integration of 10.25 Integration of 1	PANJED 12, 01 20, 7 PANJED 12, 00 30 PANJED 12, 00 30 PANJED 12, 00 30 PANJED 12, 4450 PANJED 12, 4470 PANJED 12, 4470	26 1990 27 1990 28 1990 29 1990 29 1990 20 190	A 2557631 2557631 A 12242242 A 7252533 A 7252533 A 7252534 A 7252547 A 7262547 A 7264210 A 7264210 A 7264210 A 7264210 A 7264210 A 7264210 A 72672124 A 7265215 A 725219 A 725259 A 725259 A 725259 A 725259	12363112 7550278 7625343 7625343 7635343 7635246 7645033 7664111 76647331 76647331 7667331 7675231 7675231 7675231 7669230 7631265 7631265 7723937 7727259 7733917	* * * * * * * * * * * * * * * * * * *	1013/0279 1013/0279	21 0034 21 0034 21 22 5 22 62 5 22 62 5 22 62 5 22 62 5 22 62 5 23 7 24 7 27 105 27	8313775 234534 4303472 813725 813705 809487 7855267 7855264 781503 771535 771535 771535 771535 771535 7715405 773640 773769 773641 773199 774034 44265	- 2 - 4 - 4 - 5 - 5 - 7 - 5 - 7 - 7 - 7 - 7 - 7 - 7 - 7 - 7 - 7 - 7	536 5 17 1 16 2 17 1 16 2 17 1 16 2 17 1 16 2 17 1 16 2 17 2 16 2 17 2	511 338 331 346 10 337 337 337 337 337 337 337 33	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		95()5#25) 599() 996) 7.106) 7.109() 8(9.0996);	17(7.12%) (m(H)15(7	52%)			
2 Ab2-292 2 Ab2-292 2 Ab2-253 2 Ab2-257 2 Ab2-273 2 Ab2-273	4x234.12%	82 82 82 82 85 85 85 85 85 85 85 85 85 85 85 85 85		SNUST 1001 41 39 SNUST 1001 1430 SNUST 1001 1430 SNUST 1001 1430 SNUST 1001 1430 SNUST 1001 1430 SNUST 1001 1430 SNUST 1001 1473 SNUST 1001 1483 SNUST	WHEED 1 6 170 WHEED 2 0350 WHEED 2 4550 WHEED 2 44550 WHEED 2 4450 WHEED 2 44700 WHEED 2 4	26 mm 202 27 mm 202 28 mm 202 29 mm 202 20 mm 202	A 2537631 A 12342442 A 7247384 A 7242533 A 7252533 A 7252533 A 7252533 A 7242547 A 7242847 A 7242810 A 7242810 A 7242810 A 7242810 A 7242810 A 7242810 A 7242713 A 7242713 A 725515 A 7255515 A 7255515 A 7255515 A 7255555 A 72	12563112 7550278 7652545 7652545 7652545 7652546 7654503 7645131 7647331 7647331 7647331 7647331 7647331 7667331 767333 7769260 777330 7777330 7777330 7777330 7777330 7777330 7777330 7777330 7777330 7777330 77777330 77777777	* * * * * * * * * * * * * * * * * * *	actist?) catist?) catist??)	21 0034 837125 827634 877635 817607 817606 817507 817507 978247 978247 978247 978247 978247 978247 978247 97825 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 971824 97182 971824 97182 971824 97182 971824 97182 9718 9718 97717718 977718 977718 977718 9777718 977718 977718 977718 977718 977718	331375 234354 230472 13725 13725 13705 105244 105254 105254 105255 10525 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 105555 1055555 1055555 105555 105555 105	- 2 - 1 - 4 - 5 - 5 - 5 - 5 - 5 - 7 - 7 - 7 - 7 - 7 - 7 - 7 - 7 - 7 - 7	S316 2 177 1 177 1 177 1 177 1 170 2 111 1 1202 2 1333 1 114 2 277 1 633 2 265 2 663 1 707 1 707 1 703 2 703 2 703 2 703 2 703 2 703 2 704 1 11 1 11 1 71 7 721 7 733 2 734 1 71 7 721 7	511 511 518 521 521 520 520 520 520 520 525 525 525			96)1.eu(1) 5596) 996) 1066) 7.1096) 8(9.0996);	17(7.17%) 4a(H)15(7	5376)			
2 Ab3-252 2 Ab3-252 2 Ab3-253 2 Ab3-257 2 Ab3-275 2 Ab3-275	Re14378	822 852 852 852 852 852 852 852 852 852		SNUSH COT 4138 COCO 3759 ANUSH COT 625 SNUSH COT	9NUB0116170. 9NUB012677. 9NUB012650. 9NUB012650. 9NUB012650. 9NUB012650. 9NUB012660. 9NUB012660. 9NUB0126400. 9NUB0126400. 9NUB0126400. 9NUB0126400. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012640. 9NUB012630.	24. mm variable varia	13 25761 13 24247 13 24247 13 24247 14 747384 14 7427384 14 7427384 14 7427384 14 7424311 14 7424311 14 7424311 14 7424311 14 7424311 14 7424311 14 7424311 15 7425114 16 7425114 17425114 7456104 1725116 7725114 14 7725114 15 5101057 15 5101051 15 5101051 15 5101051 15 5200060 15 5202827 15 5202827	12363112 7550278 7623545 76252545 76252545 7625254 7625254 7625254 7645205 7645205 7645205 7645205 7645205 7645205 7645205 7625205 7725205 7725205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 77727205 7772727205 77727205 77727205 77727205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777272205 777727205 777727205 777727205 777727205 777727205 777727205 7777727205 7777727205 7777727205 7777777777	* *		21 0034 21 0034 21 22 22 21 22 22	3313775 234354 230472 313725 313725 313725 313725 313725 313725 313725 313725 313725 313725 313725 313725 313725 31109 313725 31109 313725 31109 313725 31109 313725 31109 313725 31109 313725 31109 313725 31109 313725 31109 313725 31109 31100 31000 31000 3100000000	- 2 - 1 - 4 - 4 - 5 - 5 - 5 - 5 - 5 - 5 - 5 - 5 - 5 - 5	5356 9 117 1 117 1 117 1 117 1 111 1 114 2 207 1 114 2 207 1 114 2 202 2 203 2 203 2	511 513 513 513 513 513 514 515 515 515 515 515 515 515			95(3.845) 5996) 996) 7.056) 7.1096) 8(9.0996);	17(7.17%) 6a(H)15(7	53%)			
2 Ab3-25 2 Ab3-	He113776 He113776 He115236 He115236 He11526 He11526 He11526 He11526 He11526 He11526 He11526	822 822 822 822 822 822 822 822 822 822	23 23 17 17 17 17 17 17 17 18 17 17 18 18 16 16 16 16 16 16 16 16 16 16	SAUST 1001 41 39 1003 7169 America 100 1420 America 1001 420 America 1001 420 A	2014/2011 (117) 2014/2011 (217) 2014/2012 (215) 2014/2012 (215) 2014/2014 (215	24	A 2537631 A 2537631 A 2447384 A 2447384 A 2447384 A 2447384 A 2447384 A 2453121 A 7642810 A 765312 A 765418 A 765114 A 77521145 A 1212020 A 1212020 A 1212020 A 1212020 A 1212314 A 1212314 A 1212314 A 1212314 A 1212314 A	12353112 12353112 12552218 176253545 176253545 176253545 176253546 17625354 17625351 17667351 17753557 1775557 1775557 1775557 1775557 1775557 1775557 1775757 177577 177577 177577 177577 177577 177577 177577 1775777 1775777 1775777 1775777 17757	0 0	sattakin sat	010034 21705 2	131375 23434 23434 23447 23434 23447 23452 23455 23452 23452 23455 234552 234552 234552 234552 234552 234552 2345552 234552 234552 234552 234552 234555 2345552 234555		5356 9 17 1 17 1 101 1 131 1 14 2 27 1 133 1 14 2 27 1 533 2 555 1 663 1 663 1 534 2 353 2 353 2 353 2 353 2 353 2 353 2 353 2 353 2 353 2 353 2 357 2 357 2 357 2 357 2 367 2 370 2	511 513 513 513 513 513 514 515 515 515 515 515 515 515			96();es(1) 598() 106() 7.1096() 7.1096() 7.1096()	8x(H)15(7	52%)			
2 A13-25 2 A13-	4:11176 	822 823 824 825 827 827 827 827 827 827 827 827	223 223 177 177 177 178 178 178 178 178	SALIST COLO 13 20 COLO 1749 MILES COLO 1420 MILES COLO 1420 MILES COLO 1420 MILES COLO 1420 MILES COLO 1420 MILES COLO 1420 MILES COLO 1427 MILES COLO 1427 MI	DNUSCI 10170 DNUSCI 10110000 DNUSCI 1011	246 - #999.02 277 - #959.02 288 - #650.02 288 -	N 9357631 1345242 1345242 1345242 1345244 1345242 1345344 141542 1453447 141542 1453447 141542 1453447 141542 1453447 141542 1454411 141542 1454411 141544 1454411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145411 141544 145544 141544 145454 1415454 145544 1415454 145544 1415454 145544 1415454 145544 1415454 145544 1415454 145544 1415454 145544 <	12353112 12353112 1250278 1	0 0	settistin catiliziti catiliziti catiliziti catiliziti settistin catiliziti ca	010034 010034 010034 010037 014066 014066 014067 014067 014067 014067 014067 014067 014067 014067 014067 01407 0	181 575 282 575 282 282 575 282 575	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	5356 9 17 17 17 18 201 1 1333 1 1 14 2 2 17 133 1 144 2 2 17 133 2 163 2 2 17 2 2 17 2 2 17 2 2 17 3 2 18 1 1 19 1 1 11 1 1 155 1 1 172 2 3 172 2 3 173 3 1 111 1 1 153 2 3 17 3 1 17 3 1 10 2 2 10 2 2 10	511 511 513 513 513 513 513 513			95(3.845) 950) 106() 1.06() 2.099()	840012(7	52%)			
2 Ab3-25 2 Ab3-	4:11178 4:11178 4:11	822 822 823 824 825 825 825 825 825 825 825 825	20 22 22 22 22 22 22 22 22 22 22 22 22 2	SALES (100 14) 39 (100) 1769 MICE (100) 1769 MICE (100) 1769 MICE (100) 1769 MICE (100) 1769 MICE (100) 1767 MICE (100) 1777 MICE (10		246 + #990.02 257 + #950.02 258 + #950.02 25	N 93.7611 13.45242 13.45242 13.45242 13.452447 14.75242 13.45447 14.75243 14.751447 14.752447 14.751447 14.75241 14.751447 14.75211 14.752114 14.752114 14.752114 14.752114 14.752114 14.752114 14.752114 14.752114 14.752114 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752215 14.65239 14.752315 14.652339 14.752315 14.652339 14.752315 14.652339 14.752316 14.552339 14.752317 14.552339 14.752317 14.552339 14.752317 14.552339 14.752	12353112 12353112 12353112 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 12553278 1255578 1255578 1255578 1255		settisty settis	0.0004 0.0004	181 1975 181 1975 182 542 182 545 182 545 182 545 182 545 182 545 182 545 1	- 2 2 - 4 4 + 9 +	336 9 17 17 17 17 17 101 17 133 101 17 14 2 297 1 297 1 297 1 297 2 251 2 263 2 292 2 263 1 203 2 204 1 11 1 165 1 171 1 165 1 171 1 165 1 171 1 165 1 177 2 213 2 2101 2 2101 2 2102 2 2103 2 2103 2 2104 2 2105 1 118	511 538 537 538 537 538 538 539 537 537 538 537 538 535 535 535 535 535 535 535			26()(2,402)) 2560) 2910) 2910) 20100) 2010 2010	98(H)1367	59%)			



Suppleme	Supplementary Table 3. FPKM values of CP groups expressed per library																
	L1D2_EP(whole)	L2D2_EP(whole)	L3D2_EP	L4D2_EP	L5D2_EP	L6D2_EP_D	L6D2_EP_N	W_EP_D	W_EP_N	P2_EP	P2_Wg	P6_EP	P6_Wg	P9_EP	P9_Wg	P12_EP	P12_Wg
RR-1	116705	278947	82667	355390	268002	455408	462561	26954	26775	31679	1660	1890	2669	42854	14808	58283	11567
RR-2	348	49969	569	2478	377	8517	5084	992	999	11732	1051	3025	5340	134920	184337	129479	162394
RR-3	642	303	1087	1332	949	1245	1443	635	207	2541	293	10	4	4	0	85	7
CPG	22935	116540	10206	4550	4322	3499	3163	3068	5817	2265	2335	120649	147493	93050	214203	118529	162159
CPH	149	25029	647	7192	1269	11518	8751	799	5365	315771	109135	3238	1526	327068	64099	102801	39777
CPAP1	203	882	1494	192	135	221	226	2348	5271	339	147	646	746	4272	4208	13776	25131
CPAP3	1271	21586	3937	3601	1572	8445	6543	2515	2740	2654	2432	27230	31503	24499	13611	81304	8913
CPFL	17	6231	29	181	18	1072	793	39	47	6635	653	35	143	4787	16926	37249	51676
CPLCA	129	3625	325	1900	713	2035	1356	0	18	1334	161	0	12	42	15	3628	233
CPT	8404	66386	10511	4451	3861	2222	934	65	1615	16	16	82	21	25446	2762	4573	1563
CPF	4	0	10	0	0	0	3	0	8	2	2	8	5	4	29	421	856
CPCFC	0	145	3	382	0	177	226	136	117	27017	327	144	40	56928	11963	1009	456
Total	150807	569643	111485	381649	281218	494359	491083	37551	48979	401985	118212	156957	189502	713874	526961	551137	464732



Supplementary Figure 1. Expansion of *RR-1 CP* gene cluster on Chr1 in *S. litura*. The orthologous RR-1CPs are clustered in Chr9 in *B. mori.*



Supplementary Figure 2. The phylogenetic tree of CPFL CPs of *S. litura*, *M. sexta* and *B. mori*. Five CPFL CPs formed a species specific clade in *S. litura*.



Supplementary Figure 3. *CPs* showing low expression in most RNA-seq libraries. Gene names colored red belong to the largest species-specific clade in the phylogenetic tree.



Supplementary Figure 4. *S. litura* CP transcripts heatmap. *S. litura* CPs were highly abundant in the 2nd instar larval epidermis exclusively.



Supplementary Figure 5. Heatmap of transcripts of all the other *CP* genes of *S. litura* not shown in Fig. 6, Supplementary Fig. S3 and S4.