

Analysis of Expressed Sequence Tags of the Spider, *Araneus ventricosus*

Eun Hwa Chung, Kwang Sik Lee, Ji Hee Han, Hung Dae Sohn and Byung Rae Jin*

College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Korea.

(Received 30 August 2001; Accepted 2 October 2001)

We have constructed a cDNA library from the whole body of the spider, *Araneus ventricosus*. Sequence analysis of randomly selected cDNA clones was performed to obtain genetic information on the spider *A. ventricosus*, of which genetic information is currently not available. We have partially sequenced randomly selected 385 clones of the cDNA library constructed from *A. ventricosus*. This expressed sequence tags (EST) analysis revealed 383 genes had high homologies to known genes in GenBank. In this report, 241 independent genes were analyzed in detail.

Key words : Spider, *Araneus ventricosus*, Expressed sequence tags

Introduction

Spiders can generate several protein silks from specialized glands designed for diverse functions, unlike insects which produce only one type of silk. These silks are related by a predominance of alanine, serine, and glycine in their amino acid compositions. They are synthesized in specialized cells at the tail of their respective glands, secreted into a glandular lumen, and finally extruded on demand through a duct and valve system ending in a spinneret. Most molecular and structural investigations of spider have focused on spider silk with a tensile strength and extensibility (Hinman *et al.*, 2000; Vollrath, 2000). Silk strength is widely attributed to crystalline β -sheet structures. Such protein domains are found in both lepidopteran silks (e.g. *Bombyx mori*) (Mita *et al.*, 1994) and spider silks (Xu and Lewis, 1990; Hinman and Lewis,

1992). Furthermore, insecticidal toxins from arthropod venoms have also focused because these neurotoxic peptides have proved a high insecticidal potency and great selectivity for insects. The peptides have been identified from the venoms of several spiders (Adams *et al.*, 1989; Entwistle *et al.*, 1982; Krapcho *et al.*, 1995; Ornberg *et al.*, 1976; Quistad *et al.*, 1991; Skinner *et al.*, 1992). Most of these peptide toxins from the spiders appear to act by altering the function of presynaptic sodium or calcium channels, and are generally quite potent in insects (Adams *et al.*, 1989). The peptide toxin genes have been inserted into the baculovirus genome to produce enhanced insect viral pesticides (Hughes *et al.*, 1997; Krapcho *et al.*, 1995).

The spider, *Araneus ventricosus* is the abundant species in Korea (Kim *et al.*, 1997b). The body size of *A. ventricosus* is relatively larger than that of the other species. The mechanical properties of the web of the Korean native spiders, *A. ventricosus* and *Nephila clavata*, have been reported (Kim *et al.*, 1997a, b). However, the genetic information on the spider *A. ventricosus* is not currently available.

The recent rapid progress of several genome projects has brought the advancement of a new research area, structural genomics. It is required to establish tools for the utilization of the information hidden in the genome. An efficient way for gathering information on the genome of an organism is to generate and analyze expressed sequence tags (ESTs). EST analysis is one of the powerful strategies for such approaches. Partial sequencing of cDNAs is a relatively rapid and efficient way to establish genetic information on structure and expression-profile of the genes expressed in different tissues and developmental stages of various organs (Adams *et al.*, 1992; Lee *et al.*, 2000a, b; Lim *et al.*, 2000; Yun *et al.*, 2000).

To elucidate the genetic information on the spider *A. ventricosus*, in this study, we constructed and analyzed ESTs from *A. ventricosus* cDNA library. Here, we report the identification of partial nucleotide sequences of 385

*To whom correspondence should be addressed.

College of Natural Resources and Life Science, Dong-A University, Pusan 604-714, Korea. Tel & Fax: +82-51-200-7594; E-mail: brjin@mail.donga.ac.kr

Table 1. Summary of the expressed sequence tags isolated from the cDNA library of the spider *A. ventricosus*

Clone	Putative identification	Species	Overlength (length)	Homology (%)
323	Ribosomal protein S11	<i>Plasmodium falciparum</i>	267	47
832	Cytochrome C oxidase subunit VIc	<i>Homo sapiens</i>	196	70
132	Histone H2B	<i>Psammechinus miliaris</i>	237	93
1386	NADH dehydrogenase subunit 4	<i>Hypsiglena torquata</i>	247	38
1417	Astacin like metalloendopeptidase	<i>Coturnix japonica</i>	204	57
1482	Tissue factor pathway inhibitor 2	<i>Homo sapiens</i>	110	62
1395	Saxphilin precursor	<i>Rana catesbeiana</i>	237	44
93	Transforming growth factor beta 1 precursor	<i>Xenopus laevis</i>	132	58
79	Trypsin inhibitor	<i>Indian cobra</i>	153	46
90	Tissue factor pathway inhibitor 2	<i>Mus musculus</i>	63	67
102	Proteinase inhibitor	<i>Sarcophaga bullata</i>	162	66
98	6-pyruvoyl tetrahydropterin synthase	<i>Rattus norvegicus</i>	288	75
113	Cytochrome b	<i>Cryptotermes brevis</i>	284	80
115	Apoprotein A2	<i>Porphyra purpurea</i>	144	49
119	40S ribosomal protein S23	<i>Brugia malayi</i>	245	88
125	Ribosomal protein S4	<i>Plasmodium falciparum</i>	297	44
147	ADP/ATP translocase	<i>Drosophila melanogaster</i>	208	78
148	Catechol 1,2-dioxygenase	<i>Acinetobacter calcoaceticus</i>	138	38
163	Insulin-like growth factor-binding protein 1 precursor	<i>Rattus norvegicus</i>	96	57
245	Muscle LIM protein	<i>Epiblema scudderiana</i>	247	87
243	Phosphoglycerate dehydrogenase	<i>Oryza sativa</i>	108	60
271	Defender against cell death1	<i>Xenopus laevis</i>	180	59
287	Alpha-N-acetylglucosaminidase	<i>Mus musculus</i>	171	66
197	Gamma-interferon inducible lysosomal thiol reductase	<i>Homo sapiens</i>	180	51
201	Xeroderma pigmentosum gene	<i>Mus musculus</i>	204	46
205	Cytochrome p450 2P1	<i>Fundulus heteroclitus</i>	310	47
214	UDP-glucuronyltransferase-like protein	<i>Arabidopsis thaliana</i>	84	70
218	P-type cation translocating ATPase	<i>Plasmodium falciparum</i>	267	50
221	Pectinesterase	<i>Citrus sinensis</i>	300	44
225	Nucleoside diphosphate kinase A	<i>Cavia porcellus</i>	321	60
439	Astacin precursor	<i>Astacus astacus</i>	282	42
431	Merpin 1 beta	<i>Rattus norvegicus</i>	159	59
435	Progesteron-binding protein-like	<i>Arabidopsis thaliana</i>	135	55
426	Actin isoform A	<i>Mimosa pudica</i>	198	76
462	Protein kinase	<i>Mus musculus</i>	108	49
467	Transposon protein	<i>Arabidopsis thaliana</i>	117	51
498	Kynurenine 3-manoxygenase	<i>Homo sapiens</i>	113	52
483	Protease	<i>Oryzias latipes</i>	87	43
489	Proteasome subunit C10-11	<i>Oncorhynchus mykiss</i>	213	86
1418	Heat-labile hemolysin	<i>Virrio mimicus</i>	108	54
1465	High choriolytic enzyme 2 precursor	<i>Oryzias latipes</i>	186	56
529	Choriolysin H	<i>Oryzias latipes</i>	197	68
516	CAAX BOX 1	<i>Homo Sapiens</i>	1346	57
1137	Cell cycle protein p55CDC	<i>Rattus norvegicus</i>	389	76
1715	Cuticle protein	<i>Bombyx mori</i>	246	48

Table 1. Continued

Clone	Putative identification	Species	Overlength (length)	Homology (%)
1713	IgG Fc binding protein	<i>Homo sapiens</i>	131	57
1707	Selectin, Platelet	<i>Rattus norvegicus</i>	147	51
1692	Voltage-gated sodium channel	<i>Takifuge pardalis</i>	128	57
1571	Heat shock factor 3	<i>Arabidopsis thaliana</i>	234	50
1642	3- hydroxyacyl-coA dehydrogenase precussor	<i>Sus scrofa</i>	363	63
1854	Glycoprotein hormone receptor II	<i>Drosophila melanogaster</i>	137	58
1855	Histone H2A like protein	<i>Bombyx mori</i>	181	75
1773	Endoprotease	Bovine adenovirus 2	93	57
1843	Calcium binding protein	<i>Meleagris gallopavo</i>	176	66
1636	Cytochrome C oxidase polypeptide vic-2	<i>Rattus norvegicus</i>	202	73
1482	Vitellogenin I precursor	<i>Gallus gallus</i>	174	57
1879	Cytochrome C oxidase subunit IV	<i>Urechis caupo</i>	150	65
1882	Sligene product	<i>Drosophila melanogaster</i>	285	51
1888	Exportin 1	<i>Homo sapiens</i>	189	78
2313	Polycystin	<i>Homo sapiens</i>	169	67
2096	Contains similarity to cell wallplasma membrane linker protein	<i>Arabidopsis thaliana</i>	309	33
2087	Hypothetical zinc metalloproteinase	<i>Caenorhadbitis elegans</i>	207	78
2057	Cyclic nucleotide-gated channel beta subunit	<i>Rattus norvegicus</i>	240	61
2070	Dj 733D15.1(zinc-finger protein)	<i>Homo sapiens</i>	222	51
2146	Syntaxin-related protein At-SVR1	<i>Arabidopsis thaliana</i>	240	52
2244	Ribosomal protein L18a	<i>Podocoryne carnea</i>	204	81
2353	Trypsin inhibitor	<i>Heteractis</i> sp.	8	64
1993	Ferritin	<i>Echinococcus granulosus</i>	98	84
2315	Beta-hexosaminidase, beta chain	<i>Anopheles gambiae</i>	188	61
2340	6-Phosphofructokinase	<i>Mesorhizobium loti</i>	98	60
2391	3-oxoacyl-acyl carrier protein synthase	<i>Deinococcus radiodurans</i>	110	53
2393	Protein kinase	<i>Arabidopsis thaliana</i>	144	45
2421	Nuclear RNA binding protein SLM-1	<i>Rattus norvegicus</i>	179	61
2248	Elastin like protein	<i>Drosophila melanogaster</i>	41	100
2473	Saxiphilin precursor(SAX)	<i>Rana catesbeiana</i>	132	72
2607	Cytochrome b	<i>Lateolabrax latius</i>	252	78
2475	Elongation factor-1 alpha	<i>Idiogaryops pallidus</i>	354	89
2628	Putative gypsy type retrotransposon	<i>Oryza sativa</i>	48	75
2609	Chymotrypsin inhibitor (AMC1)	<i>Oryctoagus cuniculus</i>	117	53
2527	Choline dehydrogenase	<i>Rattus rattus</i>	186	71
2627	alveolin	<i>Oryzias latipes</i>	179	55
2641	Tyr protein kinase	<i>Gallus gallus</i>	69	82
2648	Adhesin protein	<i>Mycoplasma pneumoniae</i>	150	38
2588	Defensin B	<i>Ornithodoros moubata</i>	153	43
2671	Capsular polysaccharide repeating unit polymerase	<i>Streptococcus agalactiae</i>	132	58
2688	Exportin 1	<i>Homo sapiens</i>	288	89
2660	40S Ribosomal protein S25	<i>Drosophila melanogaster</i>	135	92
2766	Mitochondrial carrier homolog	<i>Drosophila melanogaster</i>	267	66
2913	Ribosomal protein S11	<i>Plasmodium falciparum</i>	246	47

Table 1. Continued

Clone	Putative identification	Species	Overlength (length)	Homology (%)
2898	Ovomucin alpha-subunit	<i>Gallus gallus</i>	150	53
2924	Fibronectin leucine rich transmembrane protein 3	<i>Homo sapiens</i>	225	47
2917	Metalloproteinase	<i>Hydra vulgaris</i>	312	61
2920	Excinuclease ABC subunit B	<i>Mycobacterium leprae</i>	132	49
2932	Putative cell division protein	<i>Streptococcus pyogenes</i>	183	48
2975	Myosin heavy chain	<i>Rattus sp.</i>	110	70
2957	Cytochrome B	<i>Blackburnia kauwa</i>	315	63
2967	Heat shock protein hsp 20.8A	<i>Bombyx mori</i>	299	54
3012	Astacin family metalloendopeptidase FARM-1	<i>Hydra vulgaris</i>	216	72
3068	Blatula protease 10 precursor	<i>Paracentrotus lividus</i>	123	58
3034	Collagen alpha 1 (chain precursor	<i>Gallus gallus</i>	171	44
3048	Intestinal toxin 1	<i>Dendroaspis polylepis</i>	144	43
3071	High choriolytic hatching proteinase	<i>Oryzias latipes</i>	267	48
3074	Heme oxygenase-2non-reducing isoform	<i>Rattus norvegicus</i>	186	49
3075	Vitellogenin B	<i>Melanogrammus aeglefinus</i>	233	46
3097	Heme oxygenase 2	<i>Mus musculus</i>	207	48
3107	Cytochrome oxydase	<i>Dryadula phaetusa</i>	348	54
3108	Succinate:cytochrome c oxidoreductase subunit	<i>Cyanidioschyzon merolae</i>	207	52
3154	Muscle-specific protein 20	<i>Drosophila melanogaster</i>	264	73
3173	Nucleotide binding protein 1	<i>Mus musculus</i>	348	82
3178	Mitochondrial carrier homolog	<i>Bombyx mori</i>	183	76
3161	Ribosomal protein S26	<i>Rattus norvegicus</i>	291	91
3209	Mitochondrial precursor receptor	<i>Rattus sp.</i>	98	69
3224	AKR / J normal thyroglobulin	<i>Mus susculus</i>	171	46
3256	Putative cell division protein	<i>Streptococcus pyogenes</i> <i>MI GAS</i>	183	48
3294	Chitinase	<i>Anopheles gambiae</i>	273	72
3275	Beta-actin	<i>Ovis aries</i>	315	86
3290	Elastin like protein	<i>Drosophila melanogaster</i>	42	100
3315	Phosphatidylinositol 4-kinase	<i>Solanum tuberosum</i>	44	93
3261	Cytochrome C oxidase subunit Va preprotein	<i>Drosophila melanogaster</i>	276	80
3266	Cu / Zn Superoxide dismutase	<i>Onchocerca volvulus</i>	179	78
3310	Integrin homolog-yeast	<i>Saccharomyces cerevisiae</i>	248	52
3373	Putative C2H2-type zinc finger protein	<i>Arabidopsis thaliana</i>	225	45
3319	Zinc finger protein mfg1 mRNA	<i>Mus musculus</i>	187	73
3419	Dipeptidyl-peptidase	<i>Rattus norvegicus</i>	369	72
3392	Alpha-A crystallin	<i>Clarias fuscus</i>	240	59
3394	Serum/glucocorticoid regulated kinase	<i>Homo sapiens</i>	401	67
3401	Multiprotein bridging factor	<i>Bombyx mori</i>	272	85
3527	NADH Dehydrogenase subunit 4	<i>Chrysomya chloropyga</i>	336	72
3532	Protease	<i>Human adenovirus type 17</i>	183	47
3429	Eye development protein canoe	<i>Drosophila melanogaster</i>	225	44
3638	Chordin	<i>Danio rerio</i>	246	42
4089	Chitinase	<i>Clostridium thermocellum</i>	129	47
4197	Ribosome associated membrane protein RAMP4	<i>Rattus norvegicus</i>	147	80

Table 1. Continued

Clone	Putative identification	Species	Overlength (length)	Homology (%)
4199	Slit protein	<i>Drosophila melanogaster</i>	225	47
4195	Translationaly controlled tumor protein	<i>Lumbricus rubellus</i>	66	81
4293	Alpha2-macroglobulin protein	<i>Limulus</i> sp.	269	53
4398	TEP2 protein	<i>Drosophila melanogaster</i>	213	62
2679	Putative secreted effector protein	<i>Escherichia coli</i>	162	53
4416	Fibroin-1	<i>Araneus diadematus</i>	182	56
4511	Alpha-2-macroglobulin	<i>Limulus</i> sp.	384	52
4649	Odorant receptor S50	<i>Mus musculus</i>	132	52
4665	Proteosome subunit, alpha type 4	<i>Rattus norvegicus</i>	225	94
4499	DJ54B20.4 (novel KRAB containing C2H2 type Zine finger protein)	<i>Homo sapiens</i>	309	65
4543	Homocyanin subunit d	<i>Eurypelma californicum</i>	210	84
4487	Major tail protein	<i>Escherichia coli</i>	329	84
4559	Tryptophanyl- tRNA synthetase	<i>Carsonella ruddii</i>	165	47
4568	Serum/glucocorticoid regulatedkinase	<i>Squalus acanthias</i>	396	67
4565	endosialin	<i>Mus musculus</i>	156	45
4567	chitinase	<i>Streptomyces coelicolor</i>	114	51
4636	Mucin glycoprotein MUC3	<i>Mus musculus</i>	171	44
4496	E-selectin precursor	<i>Oryctolagus cuniculus</i>	264	38
4659	60S Ribosomal protein L7A	<i>Gallus gallus</i>	303	87
4768	Putative auxin efflux carrier protein:AtPIN 1	<i>Arabidopsis thaliana</i>	119	57
4762	DNA-directed RNA polymerase beta chain	<i>Plasmodium falciparum</i>	213	49
4872	thyroglobulin	<i>Mus musculus</i>	243	45
4726	Tyrosine kinase C receptor	<i>Gallus gallus</i>	246	49
4867	Gilbulin-1	<i>Zea mays</i> subsp. <i>parviglumis</i>	198	42
4832	DNA polymerase alpha 1, 180KDa	<i>Mus musculus</i>	57	84
4831	Thiolase(AA 1-397)	<i>Rattus norvegicus</i>	357	76
4817	Glutathione S-transferase,GST	<i>Nematospiroides dubius</i>	255	48
4714	Predicted membrane protein	<i>Clostridium acetobutylicum</i>	144	53
4704	Polypeptide deformylase	<i>Thermotoga maritima</i>	129	57
4807	NADH dehydrogenase subunit 1	<i>Albinaria caerulea</i>	252	59
4825	Protein-tyrosine phosphatase	<i>Homo sapiense</i>	231	60
4812	alanine dehydrogenase	<i>Synechocystis</i> sp. PCC6803	111	66
4994	Actin	<i>Aplysia californica</i>	216	97
4900	Cystatin C precursor	<i>Mus musculus</i>	279	48
4880	Alternative oxidase precursor	<i>Sauromatum guttatum</i>	384	37
4964	Putative nucleic acid binding protein	<i>Oryza sativa</i>	181	48
4893	Homoglobin alpha-3 chain	<i>Silurana tropicalis</i>	255	43
4898	L-cystatin	<i>Tachypleus trimesurus</i>	288	41
4977	Fibrinogen-binding protein	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50	195	55
4959	Testican-3 protein	<i>Mus musculus</i>	174	46
4935	Peptide transport protein	<i>Candida albicans</i>	138	46
5013	Esterase/lipase VbF,putative	<i>Vibrio cholerae</i>	147	69
5086	Span protein precursor	<i>Strongylocentrotus purpuratus</i>	216	60

Table 1. Continued

Clone	Putative identification	Species	Overlength (length)	Homology (%)
5117	Decidual prolactin-related protein	<i>Rattus norvegicus</i>	171	54
5068	Cystain precursor	<i>Tachypleus trimesurus</i>	267	50
5146	Phosphoribosylformylglycinamide cyclo-ligase	<i>Pyrococcus abyssi</i>	165	49
5165	Preprosomatostatin 28	<i>Carassius auratus</i>	90	69
5168	Collagen protein	<i>Suberites domuncula</i>	141	51
5185	DNA binding protein Gt-2	<i>Oryza sativa</i>	126	47
5186	Cytolethal distending toxin subunit A precursor	<i>Actinobacillus actinomycetemcomitans</i>	156	59
5203	Development protein tolkin	<i>Drosophila melanogaster</i>	126	57
5320	DNA- directed RNA polymerase II largest chain	<i>Norsema locustae</i>	159	52
5334	MHC class II associated invariant chain	<i>Gallus gallus</i>	174	59
5372	30S ribosomal protein S12	<i>Aeropyrum pernix</i>	126	80
5340	Filament-interacting protein	<i>Dictyostelium discoideum</i>	201	45
5342	collagen	<i>Ascaris suum</i>	117	56
5348	Olfactory receptor	<i>Rana esculenta</i>	276	44
5400	Ribosomal protein S23	<i>Spodoptera frugiperda</i>	174	96
1560	DNA binding protein	<i>Metanotermobacter thermautotrophicus</i>	111	61
1914	Heme oxygenase-2 non-reducing isoform	<i>Rattus norvegicus</i>	171	47
1906	DNA- dependent RNA polymerase II	<i>Trichophyton rubrum</i>	230	41
2184	Putative delta-7-sterol reductase	<i>Castanea sativa</i>	132	54
2066	DNA polymerase III subunit gamma	<i>Streptomyces coelicolor A3(2)</i>	159	38
2308	Gamete-specific hydroxyproline-rich glycoprotein a2	<i>Chlamydomonas reinhardtii</i>	213	32
2725	Zinc carboxypeptidase A	<i>Anopheles gambiae</i>	169	73
2809	Utrophin(homologous to dystrophin)	<i>Rattus norvegicus</i>	126	58
2897	Immunoglobulin A1 protease precursor	<i>Haemophilus influenzae</i>	81	59
2800	Cell division protein	<i>Ureaplasma urealyticum</i>	204	46
3035	Voltage- gated sodium channel homologue BdNa 1	<i>Bdelloura candida</i>	165	41
3049	acyltransferase	<i>Mesorhizobium loti</i>	231	39
3076	FliV-like protein	<i>Aeromonas punctata</i>	144	58
3062	Putative maturase	<i>Asterella tenella</i>	117	58
3289	tenascin	<i>Homo sapiens</i>	87	58
3262	Crystalline, alpha binding protein 1	<i>Mus musculus</i>	108	55
95	Kaliciudone 1	<i>Anemonia sulcata</i>	167	68
129	Potassium channel blocking toxin	<i>Centruroides limpidus</i>	85	56
3626	Ribosomal protein L36	<i>Rattus norvegicus</i>	294	88
3634	Vitellogenin	<i>Athalia rosae</i>	177	59
3643	Adenine nucleotide translocator	<i>Rattus norvegicus</i>	338	80
3645	Glutamate carboxypeptidase	<i>Homo sapiens</i>	251	92
3341	Proline dehydrogenase , P5C dehydrogenase	<i>Escherichia coli K12</i>	303	85
3778	Hemolectin	<i>Drosophila melanogaster</i>	147	50
3679	60S acidic ribosomal protein PO	<i>Sarcophaga crassipalpis</i>	204	87
3688	Zinc proteinase	<i>Astacus astacus</i>	309	62
3756	Nucleotide binding protein1	<i>Mus musculus</i>	252	85
3818	Heat shock protein HSP 90-alpha	<i>Danio rerio</i>	174	49

Table 1. Continued

Clone	Putative identification	Species	Overlength (length)	Homology (%)
3844	Cytoplasmic superoxide dismutase	<i>Haemonchus contortus</i>	138	72
3853	Cytochrome B	<i>Ensatina Eschscholtzii eschscholtzii</i>	324	74
3973	r-RNA-dimethyltransferase related protein	<i>Thermoplasma acidophilum</i>	168	58
3927	Xotch protein	<i>Xenopus laevis</i>	129	61
5216	Flagelliform silk protein	<i>Nephila madascariensis</i>	174	36
5161	nuclelin	<i>Xenopus laevis</i>	276	56
5158	Reticulocyte-binding protein 2	<i>Plasmodium vivax</i>	99	60
5218	Tu translation elongation factor,mitochondrial	<i>Homo sapiens</i>	456	62
5219	Probable short-chain dehydrogenase	<i>Pseudomonas aeruginosa</i>	93	73
5228	Putative histone	<i>Agaricus bisporus</i>	159	59
5229	Amylo-1,6-glucosidase	<i>Oryctolagus cuniculus</i>	147	81
5233	40S ribosomal protein S20	<i>Ictalurus punctatus</i>	233	88
5236	Blastula protease-10	<i>Paracentrotus lividus</i>	96	77
5242	Putative RNA binding protein 1	<i>Rattus norvegicus</i>	158	59
5271	ATP synthase A subunit 6	<i>Pemphigus spyrothecae</i>	317	73
5283	Major merozoite surface antigen-malaria parasite	<i>Plasmodium falciparum</i>	111	63
5286	Ribosomal protein S2	<i>Plasmodium falciparum</i>	210	52
5289	N-type calcium channel alphasubunit	<i>Mus musculus</i>	144	52
5297	Cytochrome P450 monooxygenase	<i>Oncorhynchus mykiss</i>	117	71
3929	Similar to heat shock factor binding protein	<i>Mus musculus</i>	156	81
3934	Acid-CoA ligase, putative	<i>Caulobacter crescentus</i>	174	48
3987	DNA Polymerase	<i>Porphyra purpurea</i>	174	49
3993	Immunoglobulin A1 protease precursor	<i>Haemophilus influenza</i>	81	59

cDNA clones from *A. ventricosus*. The information of the ESTs could be used in molecular genetic research of *A. ventricosus*.

Materials and Methods

Preparation of spider

The spider *Araneus ventricosus* was collected at Kimhae, Kyungnam Province in Korea. The collected spider was directly used in this study.

Purification of poly(A)+ mRNA

Total RNAs were isolated from the whole body of *A. ventricosus* following the procedure of Total RNA extraction kit (Promega). Poly (A)+ mRNA was purified using oligo (dT) columns of Quick mRNA isolation kit (Stratagene).

Construction of cDNA library, nucleotide sequencing and EST analysis

A cDNA library was constructed from poly(A)+ mRNA

isolated from the whole body of *A. ventricosus* by Uni-ZAP XR vector and Gigapack III Gold Packing Extract (Stratagene). The cDNA was ligated into *EcoRI-XhoI* sites of Uni-ZAP XR vector. Ligated library was transformed into *E. coli* XL1-Blue MRF strain. *E. coli* XL1-Blue MRF strain was infected by the Uni-ZAP XR library harboring *A. ventricosus* cDNA and cultured on the NZY agar medium. Each plaque was suspended in SM buffer [5.8 g/l NaCl, 2 g/l MgSO₄·7H₂O, 0.05 M TrisCl (pH 7.5) and 0.01% gelatin solution] containing 0.02% (v/v) chloroform and stored at 4°C for 1 day. The plaques were eluted into SM buffer. The pBluescript phagemids were *in vivo* excised from the Uni-ZAP XR vector using an ExAssist helper phage. *E. coli* strain, SOLR cell (Stratagene), was infected by the excised phagemids and plated on LB-Amp medium (50 µg/ml ampicillin). Plasmid DNA from the overnight culture was isolated. The size of inserted cDNA was estimated with a 1% agarose gel electrophoresis after treatment of restriction enzymes (*EcoRI* and *XhoI*). For DNA sequencing, plasmid DNA was extracted by Wizard mini-preparation kit (Promega). Sequence of the 5' end of

each cDNA clone was determined using ABI PRISM Big-Dye Terminator Cycle Sequencing Kit (Applied Biosystems). The sequences were translated into 6 reading frames and compared using the DNASIS and BLAST programs provided by the NCBI. GenBank, EMBL and SwissProt databases were searched for sequence homology using a BLAST algorithm program.

Results and Discussion

To generate the ESTs of the spider *A. ventricosus* by the single-run partial sequencing of the cDNA clones, we constructed a cDNA library from mRNA of *A. ventricosus* using oligo d(T) as a primer for the synthesis of the first-stranded cDNA. The size of inserted cDNA was estimated on an agarose gel (data not shown). The average size of the cDNA inserts in this cDNA library was approximately 1 kb. The partial sequencing of randomly selected clones harboring cDNA inserts was performed to generate the *A. ventricosus* ESTs. We mainly sequenced clones containing cDNA inserts larger than 500 bp. The nucleotide sequences from 385 randomly selected clones were obtained by single sequencing runs from the putative 5' end of the insert DNAs. Each EST sequence was compared with the protein sequences in the GenBank and EMBL databases using BLAST mail server (Altschul *et al.*, 1990).

Among 385 ESTs generated, 383 sequence tags showed homology with registered genes, and at least 241 (62.6%) of the clones had significantly high homology to the deduced amino acid sequence of the gene products that have been previously characterized. A summary of the results of these *A. ventricosus* ESTs is presented in Table 1. The targeted species with the highest similarity score were widely distributed from bacteria to human. Of these 385 clones, furthermore, 68 sequence tags are related to known insect genes such as those of *Drosophila melanogaster* and *Bombyx mori*, but only two clones were matched to previously identified spider genes. Among the 68 ESTs, related to the known insect genes, approximately 82.4% matched with *D. melanogaster* genes.

Among the ESTs listed in Table 1, many of them showed a homology to the enzymes of ubiquitous metabolic pathways, components of the transcriptional or translational apparatus, stress- or defense-related proteins, and developmental proteins. ESTs related to the spider silk proteins and toxins were also found. It is noteworthy that one cDNA clone encodes a defender against cell death 1 (DAD1). DAD is one of the inhibitor of programmed cell death, playing a role in prevention of apoptosis in certain cell types (Hong *et al.*, 1999). We

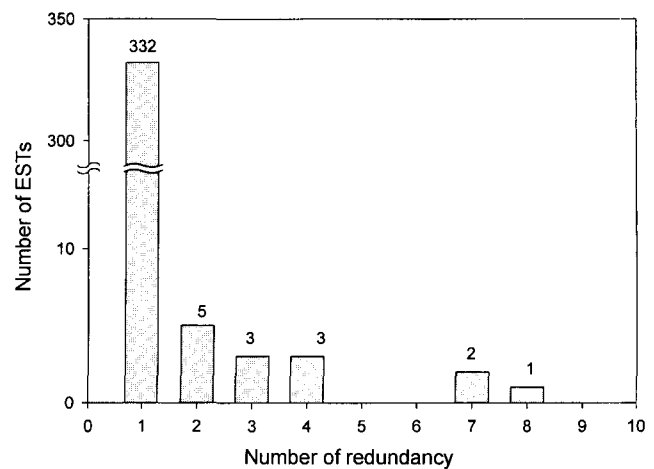


Fig. 1. Frequency distribution of ESTs showing redundancy in 385 clones. The number of the EST clones is given above each bar.

especially note that many of the genes identified in this work were related to the translational processes. Many ESTs were identified as various ribosomal proteins.

The limited extent of redundancy of EST sequencing is shown in Fig. 1. Of the 385 clones examined, 53 cDNA clones (13.8%) showed homology to other clones.

The accession numbers of EST were registered in GenBank. Results obtained from the *A. ventricosus* EST study can be applied for understanding of genetic information in spider.

Acknowledgments

We wish to acknowledge Researcher Young Bo Lee at the Department of Sericulture and Entomology, National Institute of Agricultural Science and Technology for the identification of the spider used in this study.

References

- Adams, M. D., M. Dubnick, A. R. Kerlavage, R. Moreno, J. M. Kelly, T. R. Utterback, J. W. Nagle, C. Fields and C. Venter (1992) Sequence identification of 2375 human brain genes. *Nature* **355**, 632-634.
- Adams, M. D., E. E. Herold and V. J. Venema (1989) Two classes of channel-specific toxins from funnel web spider venom. *J. Comp. Physiol. A* **164**, 333-342.
- Altschul, S. F., W. Gish, M. E. W. Myers and D. J. Lipman (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
- Entwistle, I. D., R. A. W. Johnstone, D. Medzihradzky and T. E. May (1982) Isolation of a pure toxic polypeptide from the

- venom of the spider *Phoneutria nigriventer* and its neurophysiological activity on an insect femur preparation. *Toxicon* **20**, 1059-1067.
- Hinman, M. B., J. A. Jones and R. V. Lewis (2000) Synthetic spider silk: a modular fiber. *TIBTECH*. **18**, 374-379.
- Hinman, M. B. and R. V. Lewis (1992) Isolation of a clone encoding a second dragline silk fibroin. *J. Biol. Chem.* **267**, 19320-19324.
- Hong, N. A., N. H. Kabra, S. N. Hsieh, D. Cado and A. Winoto (1999) *In vivo* overexpression of Dad1, the defender against apoptotic death-1, enhances T cell proliferation but does not protect against apoptosis. *J. Immunol.* **163**, 1888-1893.
- Hughes, P. R., H. A. Wood, J. P. Breen, S. F. Simpson, A. J. Duggan and J. A. Dybas (1997) Enhanced bioactivity of recombinant baculoviruses expressing insect-specific spider toxins in lepidopteran crop pests. *J. Invertebr. Pathol.* **69**, 112-118.
- Kim, J. P., Y. B. Lee, S. J. Jang and M. A. Kim (1997a) The studies of properties of matter and chemical qualitative analysis on web of *Nephila clavata* L. KOCH, 1878. *Korean Arachnol.* **13**, 63-65.
- Kim, J. P., Y. B. Lee, S. J. Jang and M. A. Kim (1997b) The studies of properties of matter and chemical qualitative analysis on web of *Araneus ventricosus* L. KOCH, 1878. *Korean Arachnol.* **13**, 87-91.
- Krapcho, K. J., R. M. Kral Jr., B. C. Vanwagenen, K. G. Eppler and T. K. Morgan (1995) Characterization and cloning of insecticidal peptides from the primitive weaving spider *Diguetia canities*. *Insect Biochem. Molec. Biol.* **25**, 991-1000.
- Lee, E. K., S. B. Seo, T. H. Kim, S. K. Sung, G. An, C. H. Lee and Y. J. Kim (2000a) Analysis of expressed sequence tags of *Porphyra yezoensis*. *Mol. Cells* **10**, 338-342.
- Lee, K. S., J. S. Bae, T. W. Goo, S. E. Kim, J. G. Kim, H. D. Sohn and B. R. Jin (2000b) Analysis of expressed sequence tags of the firefly, *Pyrocoelia rufa*. *Int. J. Indust. Entomol.* **1**, 165-169.
- Lim, J. Y., C. S. Shin, E. J. Chung, J. S. Kim, H. U. Kim, S. J. Oh, W. B. Choi, C. S. Ryou, J. B. Kim, M. S. Kwon, T. Y. Chung, S. I. Song, J. K. Kim, B. H. Nahm, Y. S. Hwang, M. Y. Eun, J. S. Lee, J. J. Cheong and Y. D. Choi (2000) Analysis of expressed sequence tags from *Brassica rapa* L. ssp. *pekinensis*. *Mol. Cells* **10**, 399-404.
- Mita, K., S. Ichimura and T. James (1994) Highly repetitive structure and its organization of the silk fibroin gene. *J. Mol. Evol.* **38**, 583-592.
- Ornberg, R. L., T. Smyth Jr, and A. W. Benton (1976) Isolation of a neurotoxin with a presynaptic action from the venom of the black widow spider (*Latrodectus mactans* Fabr.). *Toxicon* **14**, 329-333.
- Quistad, G. B., C. C. Reuter, W. S. Skinner, P. A. Dennis, S. Suwanrumpha and E. W. Fu (1991) Paralytic and insecticidal toxins from the funnel web spider, *Hololena curta*. *Toxicon* **29**, 329-336.
- Skinner, W. S., P. A. Dennis, J. P. Li and G. B. Quistad (1992) Identification of insecticidal peptides from venom of the trap-door spider, *Aptostichus schlingeri* (Ctenizidae). *Toxicon* **30**, 1043-1050.
- Vollrath, F. (2000) Strength and structure of spiders silks. *Rev. Mol. Biotechnol.* **74**, 67-83.
- Xu, M. and R. Lewis (1990) Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci. USA* **87**, 7120-7124.
- Yun, E. Y., T. W. Goo, S. W. Kang, H. W. Lee, J. S. Hwang and H. R. Kim (2000) Analysis of expressed sequence tags generated from the posterior silk gland cDNA clones of *Antheraea yamamai*. *Korean J. Life Sci.* **10**, 188-195.