



ISSN: 2321-9122

www.biosciencejournals.com

EJBB 2015; 3 (2): 45-55

Received: 13-02-2015

Accepted: 28-02-2015

**Monika Miglani**

Centre for Medical  
Biotechnology, Maharshi  
Dayanand University,  
Rohtak, Haryana 124001,  
India.

**S. K. Gakhar**

Centre for Medical  
Biotechnology, Maharshi  
Dayanand University,  
Rohtak, Haryana 124001,  
India.

**Correspondence:**

**Monika Miglani**

Centre for Medical  
Biotechnology, Maharshi  
Dayanand University,  
Rohtak, Haryana 124001,  
India.

Email:

monika3185@gmail.com

## In silico analysis of insect vitellogenin genes

**Monika Miglani, S. K. Gakhar**

**Abstract**

Vitellogenins and yolk proteins are synthesized in oviparous animals to provide nutrients to the developing embryo. The complete vitellogenin gene sequences belonging to six different orders of Class Insecta have been retrieved from the GenBank to prepare a database. A comparative sequence analysis was carried out to gain better understanding of the evolution of vitellogenin genes. The conservation of residues was considerable and widespread throughout the coding region. However, the high level of divergence was observed in the size and positions of introns. This could be ascribed to the presence of repetitive DNA elements. Great sequence and structural diversity suggests that vitellogenin genes are highly evolving with selective functional constraints.

**Keywords:** Insects, *In silico* analysis, Mosquito, Vitellogenin, Yolk proteins

**1. Introduction**

Developing oocytes of all oviparous species, including insects, incorporate massive amounts of yolk to ensure ample supply of nucleic acids, proteins, lipids, phosphate, carbohydrates, ion, and vitamins necessary for autonomous development of the future embryo. Although several types of yolk protein precursors (YPPs) are accumulated by insect oocytes, vitellogenin (Vg), with few exceptions, is by far the most abundant in all insect species. Vitellogenin (Vg), the major yolk protein precursor is synthesised in the fat body (analogous to the vertebrate liver) and highly expressed in a sex-, tissue- and stage- specific manner after the female ingests blood meal<sup>[1, 2]</sup>. It is secreted in to the haemolymph and subsequently sequestered in to the developing oocytes through receptor mediated endocytosis<sup>[3, 4, 5]</sup>.

Vitellogenins are known to be rapidly evolving protein sequences, probably due to the fact that the main function of these proteins is to provide amino acids during embryonic development. Thus, in addition to synonymous substitutions, most nonsynonymous amino acid substitutions in the coding region may not have a high degree of detrimental effects on protein function. Vitellogenin gene sequences may, therefore, serve as an ideal molecular marker for inferring phylogenetic relationships among insects and even more specifically among mosquitoes. Several lines of evidence discussed earlier suggest that insect vitellogenins are ideal model genes and gene products to investigate their evolution. Therefore, in the present study, the comparative nucleotide sequence analysis of insect vitellogenin genes has been carried out to explore its evolution and functional conservation. In addition, nucleotide sequence alignment, intron analysis and nucleotide composition was carried out.

**2. Methodology**

The database of complete vitellogenin gene sequences of Class Insecta has been prepared by retrieving from the GenBank. This included 27 genomic sequences belonging to four orders (Table 1), and 37 cDNA sequences belonging to six orders (Table 2). The order wise multiple sequence alignments were performed using Clustal W to detect the conserved sequences<sup>[6]</sup>. Intron sequences of Vg genes were analysed for the presence of conserved sequence motif. The intron – exon splice sites were also analysed. In addition, the secondary structure of the introns of dipteran Vgs and all the introns of Vgs belonging to other orders of Class Insecta was also predicted using Mfold web server<sup>[7]</sup>. The A+T & G+C content was detected by using Genomics % GC content calculator. The comparative analysis of A+T and G+C contents were performed to depict the biasness of the gene.

### 3. Results and Discussion

The entire open-reading frame of 64 vitellogenin gene sequences belonging to 42 insect species and 34 genera have been retrieved from various data bases at NCBI, GenBank etc (Table 1 & 2) and has been analysed by using various bioinformatic tools.

#### Nucleotide Analysis

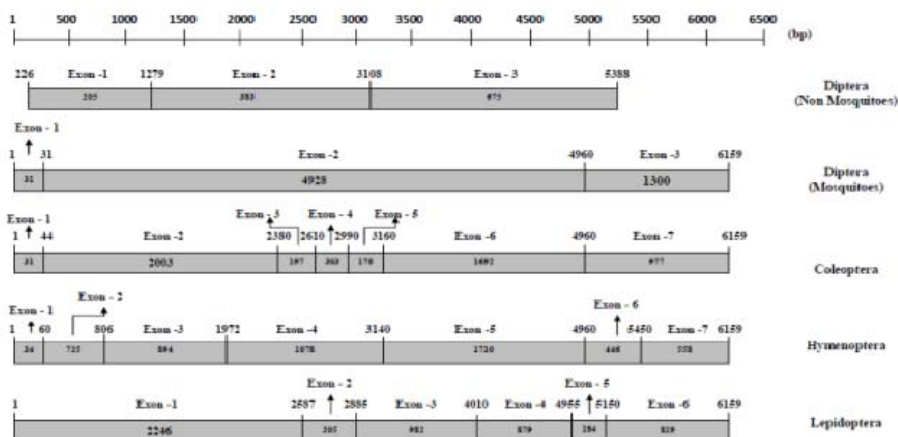
Insect Vgs have diverged considerably from one another, however the complete sequence can be aligned along with their entire lengths. The nucleotide coding sequence was found to be somewhat similar amongst all insects irrespective of the different number and position of exons. Moreover, vitellogenin genes in mosquitoes reveal strong sequence similarities to each other (Table 3).

#### 3.1 Exon Analysis

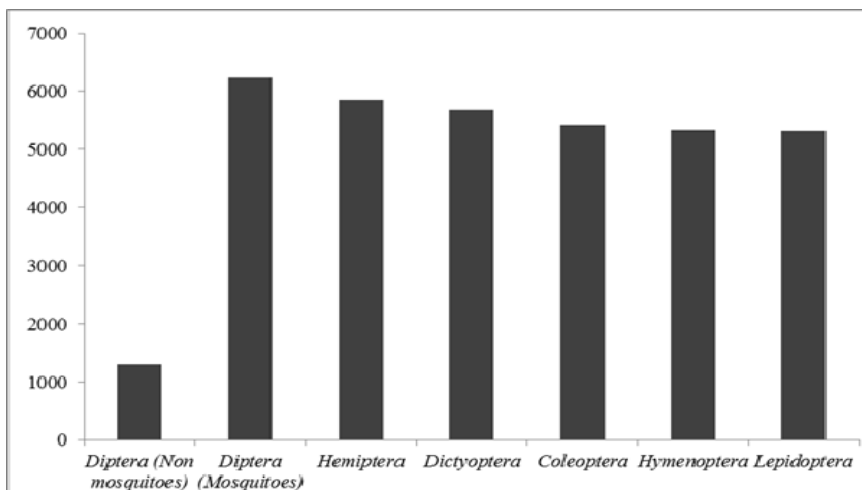
The exons of Vg genes belonging to various orders were aligned respectively (Fig. 1). The comparative analysis of Vg sequences revealed 3 exons in Diptera, 6 in Lepidoptera and 7 in Coleoptera and Hymenoptera. It was also observed that the relative position of exons was also different in different orders. However, the coding sequence was found to be almost similar in all insects. The majority of insect Vgs possess the same ORF size (6-7kb) except few Vgs

belonging to the order Coleoptera, Diptera and Hymenoptera. The Vgs belonging to non mosquito group of order Diptera revealed an ORF ranging from 1.2 kb to 1.35 kb (Supplementary file 1). However, the size of mRNA in mosquitoes ranged from 6-7 kb as also present in all other insects (Fig. 2) [8]. Nucleotide sequence alignments of insect Vgs has been shown to have conservation at few limited short regions [9,10]. However, the present study indicated that conservation of residues was considerable and widespread throughout the coding region (Fig. 1).

Interestingly, the Vgs belonging to the order Diptera possess 3 exons but the size and position of exons varied in this order particularly in mosquitoes and in non- mosquitoes (Fig. 1). Also, the mosquito Vgs revealed certain conserved sites such as; translation start site, ATG along with conserved exon-intron splice sites and similar termination codon site. The first exon was also found to be constant in size (31bp) in mosquitoes (Table 4). The role of this particular exon in functional conservation or any translation initiation needs to be further investigated. However, the sequence comparisons of the vitellogenin genes from taxonomically divergent organism showed that there are many nonsynonymous substitutions, indicating that the vitellogenin genes are rapidly evolving genes [11, 12, 13].



**Fig 1:** Showing respective position of exons amongst different orders of insect with respect to nucleotide sequence alignment on one scale. The organisms representing different groups are *Drosophila melanogaster* (Non mosquitoes), *Anopheles culicifacies* (Mosquitoes), *Anthonomous grandis* (Coleoptera), *Encarsia formosa* (Hymenoptera), *Actias selene* (Lepidoptera). Digits inside and outside the box represents size and position (bp) respectively.



**Fig 2:** Average ORF size of vitellogenin gene of different orders of Class Insecta

**Table 1:** List of genomic sequences of vitellogenin genes of various insects.

Sr. No.	Order	Family	Organism	Accession No.	Length (bp)	ORF (bp)	No. of Exons	No. of Introns	AA	Mol. Wt. (kDa)	References	
1	<b>Coleoptera</b>	Curculionidae	<i>Anthonomus grandis</i>	M72980	10017	5373	7	6	1790	205.8	Trewitt <i>et al.</i> , 1992 <sup>[11]</sup>	
2		<b>Diptera</b>	Culicidae	<i>Aedes aegypti</i> Vg A1	L41842	8780	6457	3	2	2148	250.2	Romans <i>et al.</i> , 1995 <sup>[10]</sup>
3				<i>Aedes aegypti</i> Vg B	AY380797	12957	6429	3	2	2142	249.4	Isoe, J and Hagedorn, HH, 2007 <sup>[14]</sup>
4				<i>Aedes aegypti</i> Vg C	AY373377	9101	6264	3	2	2087	242.0	Isoe, J and Hagedorn, HH, 2007 <sup>[14]</sup>
5				<i>Anopheles albimanus</i> Vg C	AY691327	7454	6089	3	2	2029	237.4	Isoe, J and Hagedorn, HH, 2007 <sup>[14]</sup>
6				<i>Anopheles culicifacies</i>	JN113091	8911	6159	3	2	2052	238.0	Miglani and Gakhar, 2013 <sup>[13]</sup>
7				<i>Anopheles gambiae</i>	AF281078	12312	6156	3	2	2051	239.3	Romans, 2000 (unpub)
8				<i>Anopheles minimus</i>	KJ000491	6569	6150	3	2	2049	239.4	Miglani and Gakhar, 2014 (unpub)
9				<i>Anopheles splendidus</i>	KF577805	6177	6177	3	2	2058	240.2	Miglani and Gakhar, 2013 (unpub)
10				<i>Anopheles stephensi</i>	DQ442990	9342	6177	3	2	2058	240.0	Nirmala <i>et al.</i> , 2006 <sup>[15]</sup>
11				<i>Anopheles subpictus</i>	KJ000061	6523	6156	3	2	2052	239.2	Miglani and Gakhar, 2014 (unpub)
12				<i>Culex pipiens quinquefasciatus</i>	AY691324	9886	6336	3	2	2111	241.5	Isoe, J and Hagedorn, HH, 2007 <sup>[14]</sup>
13				<i>Culex tarsalis</i> Vg 1a	GU017909	9676	6354	3	2	2117	242.4	Chen <i>et al.</i> , 2010 <sup>[16]</sup>
14				<i>Culex tarsalis</i> Vg 1b	GU017910	8499	6360	3	2	2119	242.3	Chen <i>et al.</i> , 2010 <sup>[16]</sup>
15				<i>Culex tarsalis</i> Vg 2a	GU017911	7123	6243	3	2	2080	233.7	Chen <i>et al.</i> , 2010 <sup>[16]</sup>
16				<i>Culex tarsalis</i> Vg 2b	GU017912	7704	6258	3	2	2085	234.5	Chen <i>et al.</i> , 2010 <sup>[16]</sup>
17				<i>Toxorhynchites amboinensis</i>	AY691326	8565	6099	3	2	2032	233.4	Isoe, J and Hagedorn, HH, 2007 <sup>[14]</sup>
18				<b>Drosophilidae</b>	<i>Drosophila melanogaster</i> YP 1	V00248	1723	1320	2	1	439	48.7
19		<i>Drosophila melanogaster</i> YP 2	M15898		4280	1263	3	2	420	46.1	Garabedian <i>et al.</i> , 1987 <sup>[18]</sup>	
20		<b>Tephritidae</b>	<i>Ceratitis capitata</i> Vg 1	X54661	2364	1314	3	2	437	48.1	Rina and Savakis, 1991 <sup>[19]</sup>	
21			<i>Ceratitis capitata</i> Vg 2	X54662	1702	1269	2	1	422	45.5	Rina and Savakis, 1991 <sup>[19]</sup>	
22	<b>Hymenoptera</b>	Aphelinidae	<i>Encarsia formosa</i>	AY553878	6654	5444	7	6	1814	204.5	Donell, D.M. 2004 <sup>[20]</sup>	
23	<b>Lepidoptera</b>	Saturniidae	<i>Actias selene</i>	GU361974	7329	5325	6	5	1774	201.3	Qian <i>et al.</i> , 2011 <sup>[21]</sup>	
24			<i>Antheraea pernyi</i>	EF683091	8142	5337	6	5	1778	201.6	Liu <i>et al.</i> , 2011 <sup>[22]</sup>	
25			<i>Antheraea yamami</i>	AB247378	13260	5337	7	6	1778	201.4	Meng and Liu 2006a <sup>[23]</sup>	
26		Bombycidae	<i>Bombyx mori</i>	D30733	9664	5349	7	6	1782	203.0	Yano <i>et al.</i> , 1994b <sup>[9]</sup>	
27		Erebidae	<i>Lymantria dispar</i>	U90756	16133	5244	7	6	1747	198.1	Hiremath, S. and Lehtoma, K. 1997b <sup>[24]</sup>	

Unpub- Unpublished;

\*Only complete sequences have been included

**Table 2:** List of cDNA sequences of vitellogenin genes of various insects.

Sr. No.	Order	Family	Organism	Accession No.	Location	ORF	Signal peptide	Mat peptide	Poly A signal	AA	Mol. Wt. (kDa)	Reference
1	Dictyoptera	Ectobiidae	<i>Blatella germanica</i>	AJ005115	25-5613	5589	25-75	76-5610	NR	1862	213.6	Comas <i>et al.</i> , 2000 [25]
2		Blattidae	<i>Periplaneta americana</i> Vg1	AB034804	13-5703	5691	13-60	61-5700	5816-5821	1896	212.7	Tufail <i>et al.</i> , 2000 [26]
3			<i>Periplaneta americana</i> Vg2	AB047401	16-5646	5631	16-63	64-5643	5783-5788	1876	215.9	Tufail <i>et al.</i> , 2001 [27]
4		Blaberidae	<i>Rhyparobia maderae</i> Vg1	AB052640	13-5754	5742	13-57	58-5751	5868-5873	1913	219.5	Tufail & Takeda, 2002 [28]
5			<i>Rhyparobia maderae</i> Vg2	AB194976	16-5751	5736	16-63	64-5748	5865-5870	1911	219.6	Tufail <i>et al.</i> , 2007 [29]
6	Hemiptera	Cicadidae	<i>Graptopsaltria nigrofuscata</i>	AB026848	12-5975	5964	12-59	60-5972	6137-6142	1987	221.9	Lee <i>et al.</i> , 2000b [30]
7		Cicadellidae	<i>Homalodisca coagulata</i>	DQ118408	1-5673	5673	NR	NR	NR	1890	211.5	Hunter & Hunnicutt, 2006 [31]
8		Delphacidae	<i>Laodelphax striatella</i> Vg1	KC469580	59-6196	6138	59-112	113-6193	NR	2045	227.4	Huo <i>et al.</i> , 2014 [32]
9			<i>Laodelphax striatella</i> Vg2	KC469581	19-6192	6174	19-78	79-6189	NR	2057	228.8	Huo <i>et al.</i> , 2014 [32]
10			<i>Nilaparvata lugens</i>	AB353856	13-6204	6192	13-66	67-6201	6264-6269	2063	229.6	Tufail <i>et al.</i> , 2010 [33]
11		Belostomatidae	<i>Lethocerus deyrollei</i>	AB425334	42-5729	5688	42-98	99-5726	NR	1895	212.9	Nagaba <i>et al.</i> , 2011 [34]
12		Pentatomidae	<i>Plautia stali</i> Vg1	AB033498	20-5743	5712	20-70	71-5740	5782-5787	1907	216.2	Lee <i>et al.</i> , 2000a [35]
13			<i>Plautia stali</i> Vg2	AB033499	1-5571	5571	1-51	52-5568	5592-5597	1856	209.9	Lee <i>et al.</i> , 2000a [35]
14			<i>Plautia stali</i> Vg3	AB033500	12-5723	5724	12-65	66-5720	5809-5814	1903	211.1	Lee <i>et al.</i> , 2000a [35]
15		Alydidae	<i>Riptortus clavatus</i>	U97277	26-5656	5631	NR	NR	NR	1876	210.6	Hirai <i>et al.</i> , 1998 [36]
16	Coleoptera	Tenebrionidae	<i>Tenebrio molitor</i>	AY714212	31-5496	5466	31-78	79-5493	5570-5575	1821	206.0	Warr <i>et al.</i> , 2006 [37]
17	Diptera	Culicidae	<i>Aedes aegypti</i>	U02548	10-6429	6420	NR	NR	NR	2139	249.3	Chen <i>et al.</i> , 1994 [38]
18		Drosophilidae	<i>Drosophila grimshawi</i> YP1	XM_001991852	1-1302	1302	NR	NR	NR	433	47.7	Zimin <i>et al.</i> , 2008 [39]
19			<i>Drosophila grimshawi</i> YP2	XM_001991853	1-323	1323	NR	NR	NR	NR	49.1	Zimin <i>et al.</i> , 2008 [39]
20	Hymenoptera	Apidae	<i>Apis mellifera</i>	AJ517411	25-5337	5312	25-72	73-5334	5374-5379	1770	201.0	Piulachs <i>et al.</i> , 2003 [40]
22			<i>Bombus hypocrita</i>	GQ340749	32-5350	5319	NR	NR	NR	1772	201.9	Li <i>et al.</i> , 2010 [41]
23			<i>Bombus ignitus</i>	FJ913883	30-5348	5319	NR	NR	NR	1772	202.0	Li <i>et al.</i> , 2010 [41]
24		Ichneumonidae	<i>Pimpla nipponica</i>	AF026789	6-5429	5424	NR	NR	NR	1807	203.4	Nose <i>et al.</i> , 1997 [42]
25		Formicidae	<i>Solenopsis invicta</i> Vg 1	AF512520	34-4959	4925	NR	NR	NR	1641	186.2	Lewis <i>et al.</i> , 2001 [43]
26			<i>Solenopsis invicta</i> Vg 2	AY941795	36-5459	5424	NR	NR	NR	1807	204.9	Tian <i>et al.</i> , 2004 [44]
27			<i>Solenopsis invicta</i> Vg 3	AY941796	37-5322	5286	NR	NR	NR	1761	201.0	Tian <i>et al.</i> , 2004 [44]
28		Pteromalidae	<i>Pteromalus puparum</i>	EF468683	35-5446	5412	NR	NR	NR	1803	202.9	Ye <i>et al.</i> , 2008 [45]
29		Tenthredinidae	<i>Athalia rosae</i>	AB007850	16-5634	5619	16-63	64-5631	NR	1872	206.3	Nose <i>et al.</i> , 1997 [42]
30		Vespidae	<i>Vespula vulgaris</i>	JN794080	1-5271	5271	NR	NR	NR	1756	199.7	Blank <i>et al.</i> , 2013 [46]

31	Lepidoptera	Bombycidae	<i>Bombyx mandarina</i>	AB055845	43-5385	5343	NR	NR	NR	1780	202.9	Meng <i>et al.</i> , 2006b <sup>[47]</sup>
32			<i>Bombyx mori</i>	D13160	43-5391	5349	43-87	88-5388	5715-5720	1782	203.0	Yano <i>et al.</i> , 1994 <sup>[48]</sup>
33.		Crambidae	<i>Cnaphalocrocis medinalis</i>	JN408698	80-5449	5370	NR	NR	NR	1789	204.0	Gu <i>et al.</i> , 2011 (unpub)
34.		Noctuidae	<i>Helicoverpa armigera</i>	JX504706	1-5271	5271	NR	NR	NR	1756	200.8	Zhang and Xiao, 2012 (unpub)
35.		Saturniidae	<i>Samia cynthia pryeri</i>	AB190810	66-5405	5340	NR	NR	NR	1779	201.6	Meng <i>et al.</i> , 2008 <sup>[49]</sup>
36.			<i>Samia ricini</i>	AB055844	64-5403	5340	NR	NR	NR	1779	201.5	Meng <i>et al.</i> , 2008 <sup>[49]</sup>
37.			<i>Saturnia japonica</i>	AB190809	57-5387	5331	NR	NR	NR	1776	203.2	Meng <i>et al.</i> , 2008 <sup>[49]</sup>

NR- Not Reported ;

\* Only complete sequences have been included

**Table 3:** Percent identity of mosquito vitellogenin genes

Sr. No.	Organism	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	<i>Ae. aegypti A1</i>	100	86	77	80	76	78	83	73	66	67	67	67	66	66	67	68	69	65	65	68	66	78	69	69
2	<i>Ae. aegypti B</i>	86	100	70	70	77	71	84	72	66	66	66	67	66	66	66	68	69	67	67	69	69	77	69	69
3	<i>Ae. aegypti C</i>	77	70	100	—	81	—	72	84	69	69	69	70	69	69	70	72	72	67	67	72	66	69	79	68
4	<i>Ae. albopictus A1</i>	80	70	—	100	—	85	71	—	—	—	—	—	—	—	—	—	—	—	—	—	—	68	—	—
5	<i>Ae. albopictus C</i>	76	77	81	—	100	—	76	89	70	70	69	—	70	70	—	71	71	—	—	72	—	71	75	72
6	<i>Ae. polynesiensis A1</i>	78	71	—	85	—	100	74	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	<i>Ae. polynesiensis B</i>	83	84	72	71	76	74	100	72	66	69	68	69	68	69	69	70	70	65	65	70	66	74	71	68
8	<i>Ae. Polynesiensis C</i>	73	72	84	—	89	—	72	100	67	70	69	69	69	69	69	71	70	66	65	71	66	70	77	69
9	<i>An. albimimus C</i>	66	66	69	—	70	—	66	67	100	85	86	85	86	86	84	74	74	71	72	76	71	64	69	67
10	<i>An. culicifacies</i>	67	66	69	—	70	—	69	70	85	100	90	95	93	93	91	70	70	72	71	70	70	65	69	69
11	<i>An. gambiae</i>	67	66	69	—	69	—	68	69	86	90	100	91	91	90	94	72	72	69	69	71	72	65	68	68
12	<i>An. minimus</i>	67	67	70	—	—	—	69	69	85	95	91	100	93	93	92	70	70	71	71	70	70	65	69	68
13	<i>An. splendidus</i>	66	66	69	—	70	—	68	69	86	93	91	93	100	96	92	71	71	70	72	70	72	65	69	69
14	<i>An. stephensi</i>	66	66	69	—	70	—	69	69	86	93	90	93	96	100	92	71	70	69	68	70	70	65	68	69
15	<i>An. subpictus</i>	67	66	70	—	—	—	69	69	84	91	94	92	92	92	100	71	71	69	69	71	69	65	69	68
16	<i>Cx. tarsalis 1a</i>	68	68	72	—	71	—	70	71	74	70	72	70	71	71	71	100	98	68	68	91	69	67	70	69
17	<i>Cx. tarsalis 1b</i>	69	69	72	—	71	—	70	70	74	70	72	70	71	70	71	98	100	68	68	90	70	67	70	69
18	<i>Cx. tarsalis 2a</i>	65	67	67	—	—	—	65	66	71	72	69	71	70	69	69	68	68	100	97	69	89	66	68	68
19	<i>Cx. tarsalis 2b</i>	65	67	67	—	—	—	65	65	72	71	69	71	72	68	69	68	68	97	100	69	90	67	68	68
20	<i>Cx. quinquefasciatus C1</i>	68	69	72	—	72	—	70	71	76	70	71	70	70	70	71	91	90	69	69	100	70	67	70	69
21	<i>Cx. quinquefasciatus C2</i>	66	69	66	—	—	—	66	66	71	70	72	70	72	70	69	69	70	89	90	70	100	66	68	69
22	<i>Oc. atropalpus B</i>	78	77	69	68	71	—	74	70	64	65	65	65	65	65	65	67	67	66	67	67	66	100	69	66
23	<i>Oc. atropalpus C</i>	69	69	79	—	75	—	71	77	69	69	68	69	69	68	69	70	70	68	68	70	68	69	100	72
24	<i>Tx. amboinensis C</i>	69	69	68	—	72	—	68	69	67	69	68	68	69	69	68	69	69	68	68	69	69	66	72	100

### 3.2 Intron Analysis

Among Class Insecta, Diptera was observed to have 2 introns, Lepidoptera 5 and Coleoptera & Hymenoptera had 6 introns (Tables 4-8). No significant sequence conservation was observed in introns of vitellogenins. However, it was observed that the two introns of mosquito Vgs had similar position as that in boll weevil (*Anthonomous grandis*- Coleoptera) and silkworm (*Bombyx mori* – Lepidoptera) Vgs where six introns were present (Data not shown).

The size of introns also varied in different orders of Class Insecta (Tables 5-8). In mosquitoes (Diptera), the first intron ranged from 62 bp (*Cx. tarsalis*) to 119 bp (*An. stephensi*) and the second intron ranged from 57 (*Ae. aegypti*) to 84 bp (*An. stephensi*) (Table 4). It was also interesting to note that the first intron in mosquitoes, i.e. in *Anopheles*, *Aedes* and *Culex* was bigger in size as compared to second intron. However, in the second duplicated copy of Vg in *Cx. tarsalis* (2a and 2b genes) <sup>[16]</sup>, the second intron was bigger in size than the first one (Fig. 3). On the contrary, the size of second intron was found to be almost similar in all the orders of Insecta (Fig. 4). However, the number and size of introns

varied in order Lepidoptera (Fig. 5) (Table 8).

The intron-exon splice sites showed a significant conservation pattern, particularly in Diptera (Tables 4 & 5). Similarly the splice sites in other orders i.e. Coleoptera, Hymenoptera and Lepidoptera, also had a particular pattern (Tables 6-8). These conserved splice sites might indicate some selective functional constraints during evolution for maintaining the overall function of the vitellogenins. The presence of repetitive DNA elements may be the reason for the variation in the size and position of introns as also observed among vertebrates <sup>[50]</sup>.

The secondary structure of all the Vg introns of Class Insecta was also predicted using Mfold web server <sup>[7]</sup>. The stem loop structures were showing significant resemblance to the standard clover leaf structure of tRNA genes. The structures predicted were not having a common representation at all because of the absence of any conservation pattern among introns (Supplementary File 2). This was in spite of the fact that the second intron of vitellogenins was having almost the same size in all the orders as discussed above, even it does not reveal any resemblance in its structure among insects.

**Table 4:** Exon lengths and intron boundaries of complete mosquito Vitellogenin genes.

Organism	Gene	Exon1 (bp)	Intron 1		Exon 2 (bp)	Intron 2		Exon 3 (bp)
			3' Splice Sites	5' Splice Sites		3' Splice Sites	5' Splice Sites	
<i>Anopheles culicifacies</i>	Vg	31	G/GTAAGT ... (107)* ...	TTACAG/T	4929	C/GTAAGT ... (76)* ...	TTTCAG/C	1199
<i>Anopheles splendidus</i>	Vg	31	G/GTAAGT ... (115)* ...	TTACAG/T	4947	C/GTAAGT ... (77)* ...	TTTCAG/C	1199
<i>Anopheles minimus</i>	Vg	31	G/GTAAGT ... (117)* ...	TTACAG/T	4929	C/GTAAGT ... (72)* ...	TTTCAG/C	1190
<i>Anopheles subpictus</i>	Vg	31	G/GTAAGT ... (115)* ...	TTACAG/T	4935	C/GTAAGT ... (81)* ...	TTTCAG/C	1189
<i>Anopheles gambiae</i>	Vg 1	31	G/GTAAGT ... (99)* ...	TTACAG/T	4941	C/GTAAGT ... (74)* ...	CTACAG/C	1184
<i>Anopheles stephensi</i>	Vg 1	31	G/GTAAGT ... (119)* ...	TTGCAG/T	4947	C/GTAAGT ... (84)* ...	CTCCAG/C	1199
<i>Anopheles albimanus</i>	Vg C	31	G/GTAAGT ... (101)* ...	CTGCAG/T	4881	C/GTAAGT ... (83)* ...	CATCAG/C	1178
<i>Aedes aegypti</i>	Vg A1	31	G/GTAAGT ... (70)* ...	CCACAG/C	5214	C/GTAAGT ... (57)* ...	TTTCAG/C	1202
<i>Aedes aegypti</i>	Vg B	31	G/GTAAGT ... (82)* ...	CCACAG/T	5199	C/GTAAGT ... (59)* ...	TTGCAG/C	1199
<i>Aedes aegypti</i>	Vg C	31	G/GTAAGT ... (67)* ...	CAATAG/T	4995	C/GTAAGT ... (68)* ...	TCGCAG/C	1238
<i>Ochlerotatus atropalpus</i>	Vg B	31	G/GTAAGT ... (74)* ...	TCGCAG/T	5190	C/GTAAGT ... (61)* ...	TTCTAG/C	1103
<i>Ochlerotatus atropalpus</i>	Vg C	31	G/GTCAGT ... (80)* ...	TTTCAG/T	4992	C/GTAAGT ... (68)* ...	TCTTAG/C	>453
<i>Culex tarsalis</i>	Vg 1a	31	G/GTAAGT ... (63)* ...	ACGCAG/T	5061	T/GTAAGA ... (61)* ...	CTACAG/C	1262
<i>Culex tarsalis</i>	Vg 1b	31	G/GTAAGT ... (64)* ...	ACGCAG/T	5067	T/GTAAGA ... (59)* ...	CTACAG/C	1262
<i>Culex tarsalis</i>	Vg 2a	31	G/GTAAGC ... (62)* ...	TTCCAG/T	5013	C/GTAAGT ... (75)* ...	TTCCAG/C	1199
<i>Culex tarsalis</i>	Vg 2b	31	G/GTAAGC ... (62)* ...	TTCCAG/T	5028	C/GTAAGT ... (75)* ...	TTCCAG/C	1199

\*Size of the intron has been mentioned in parenthesis

**Table 5:** Exon lengths and intron boundaries of complete non- mosquito Vitellogenin genes.

Organism	Gene	Exon1 (bp)	Intron 1		Exon 2 (bp)	Intron 2		Exon 3 (bp)
			3' Splice Sites	5' Splice Sites		3' Splice Sites	5' Splice Sites	
<i>Drosophila melanogaster</i>	YP 1	220	CT/GTGAGT ... (75)*... TTGTAG/AT		1100	-		-
<i>Drosophila melanogaster</i>	YP 2	205	CT/GTAAGT ... (68)*... CCGCAG/AC		383	-		-
<i>Ceratitis capitata</i>	Vg 2	211	CT/GTAAGT ... (89)* ... CTCTAG/AT		1058	-		-
<i>Ceratitis capitata</i>	Vg 1	223	CT/GTAAGT ... (67)* ... CTCTAG/AT		389	TC/GTAAGT... (68)*... CTCAG/AT		702
<i>Drosophila melanogaster</i>	YP 3	205	CT/GTGAGT ... (62)* ... ACCTAG/AT		383	TC/GTAAGT... (72)*... TGCAG/AT		675

**Table 6:** Exon lengths and intron boundaries of *Anthonomus grandis* (Coleoptera) vitellogenin gene.

Organism	<i>Anthonomus grandis</i>
Exon 1 (bp)	31
Intron 1 Splice sites (3' – (bp) – 5')	G/GTAAGT (2058)*TTACAG/T
Exon 2 (bp)	2003
Intron 2 Splice sites (3' – (bp) – 5')	G/GTAAGT (87)* TTTCAG/A
Exon 3 (bp)	197
Intron 3 Splice sites (3' – (bp) – 5')	G/GTAAGT(57)*TACTAG/A
Exon 4 (bp)	303
Intron 4 Splice sites (3' – (bp) – 5')	G/GTAAGT (103)* TCACAG/A
Exon 5 (bp)	170
Intron 5 Splice sites (3' – (bp) – 5')	G/GTAAGT(59)*TTGTAG/T
Exon 6 (bp)	1692
Intron 6 Splice sites (3' – (bp) – 5')	G/GTAAGT(87)*TTTTAG/C
Exon 7 (bp)	977

\*Size of the intron has been mentioned in parenthesis

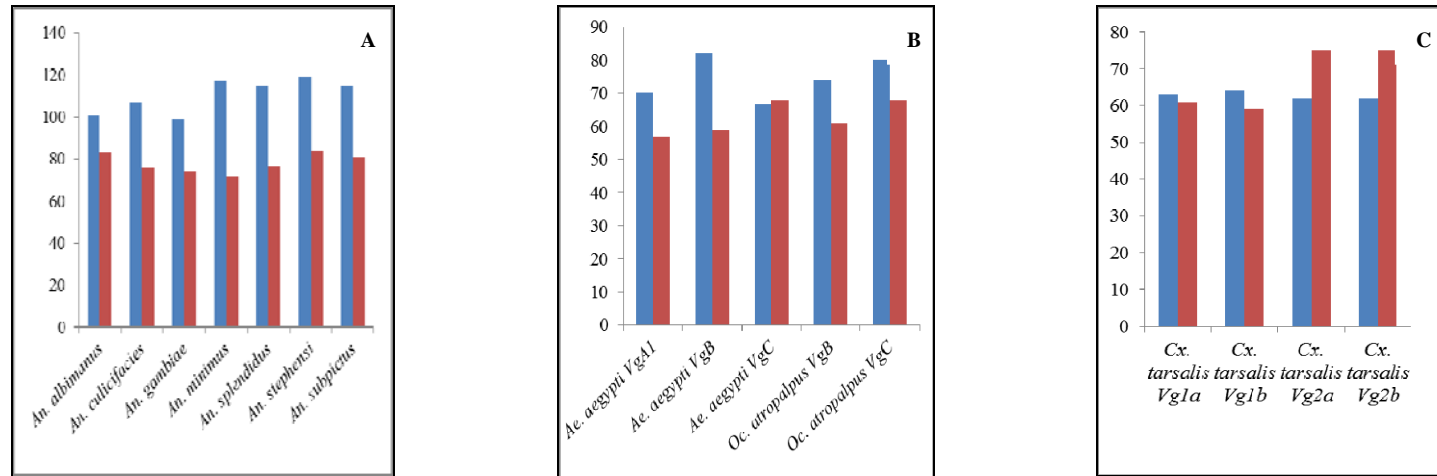
**Table 7:** Exon lengths and intron boundaries of *Encarsia formosa* (Hymenoptera) vitellogenin genes.

Organism	<i>Encarsia formosa</i>
Exon 1 (bp)	34
Intron 1 Splice sites (3' – (bp) – 5')	G/GTAAGG (91)*TTTCAG/T
Exon 2 (bp)	725
Intron 2 Splice sites (3' – (bp) – 5')	T/GTAAGA (85)* TTTTAG/C
Exon 3 (bp)	894
Intron 3 Splice sites (3' – (bp) – 5')	T/GTAAGT(84)*CATTAG/G
Exon 4 (bp)	1078
Intron 4 Splice sites (3' – (bp) – 5')	G/GTAAAT (69)* TTCTAG/A
Exon 5 (bp)	1710
Intron 5 Splice sites (3' – (bp) – 5')	A/GTAAGT(74)*TTATAG/G
Exon 6 (bp)	446
Intron 6 Splice sites (3' – (bp) – 5')	G/GTAAGA (77)*TTACAG/G
Exon 7 (bp)	558

**Table 8:** Exon lengths and intron boundaries of Lepidoptera vitellogenin genes

Organism	<i>Actias selene</i>	<i>Antheraea pernyi</i>	<i>Antheraea yamami</i>	<i>Bombyx mori</i>	<i>Lymantria dispar</i>
Exon 1 (bp)	2246	2258	31	31	31
Intron 1 Splice sites (3' – (bp) – 5')	T/GTAAGT (107)*ATTTAG/A	T/GTGAGT(84)*TTTCAG/T	G/GTGGTT(4590)*TTTCAG/C	G/GTAAGA(1588)*TTACAG/C	G/GTAAGA(1667)*TTTCAG/C
Exon 2 (bp)	205	205	2227	2242	2143
Intron 2 Splice sites (3' – (bp) – 5')	G/GTGAAA (84)* TTCCAG/A	G/GTGAGA (78)* TTACAG/G	T/GTGAGT(77)*TTTCAG/T	T/GTGAGT(91)* TTTCAG/A	T/GTAAGT(254)*TTTTAG/G
Exon 3 (bp)	982	982	205	208	214
Intron 3 Splice sites (3' – (bp) – 5')	G/GTAAGT(626)*TTTCAG/A	G/GTAAGC(410)*TTTCAG/A	G/GTGAGC(77)*ATTCAG/G	A/GTAGGA(81)*TTACAG/A	G/GTTCGT(781)*TTTCAG/G
Exon 4 (bp)	879	879	982	985	970
Intron 4 Splice sites (3' – (bp) – 5')	C/GTGAGT (706)*TTACAG/C	C/GTGAGT(1055)* TTACAG/C	G/GTAAGC(423)*TTTCAG/A	G/GTAAGT(88)* TTTCAG/A	G/GTACGT(374)*TTACAG/A
Exon 5 (bp)	184	184	879	879	870
Intron 5 Splice sites (3' – (bp) – 5')	A/GTGAGT(230)*TTGCAG/A	A/GTGAGT(795)*TTTTAG/A	C/GTGAGT(896)*TTACAG/C	C/GTAAGT(740)*TTACAG/C	C/GTAAGT(3016)*CAACAG/C
Exon 6 (bp)	829	829	184	184	202
Intron 6 Splice sites (3' – (bp) – 5')	–	–	A/GTGAGT(806)*ATTTAG/A	A/GTTAGT(286)*CAACAG/G	A/GTAAGT(297)*TTTTAG/G
Exon 7 (bp)	–	–	829	820	814

\*Size of the intron has been mentioned in parenthesis



**Fig 3:** Showing relative size of two introns of vitellogenin gene. A-Anopheles, B- Aedes, C- Culex; ■ Intron 1 ■ Intron 2



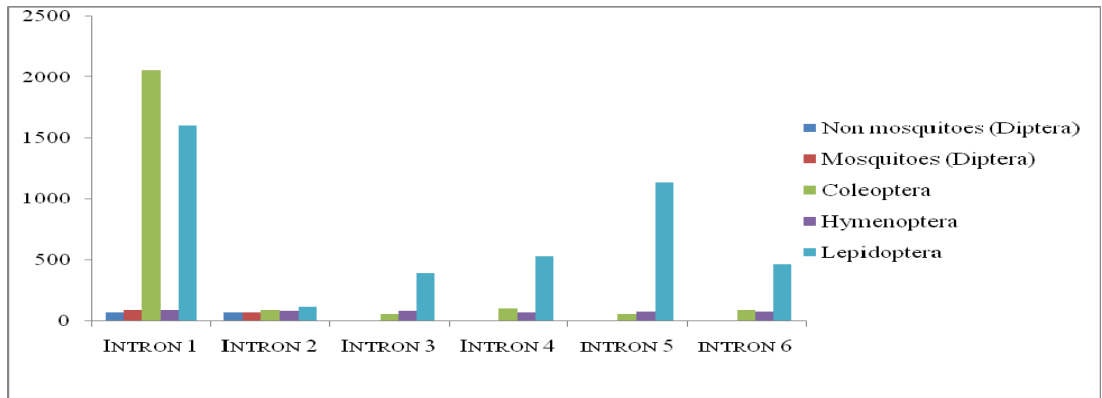


Fig 4: Bar chart showing the size of introns of vitellogenin gene in different orders of Class Insecta.

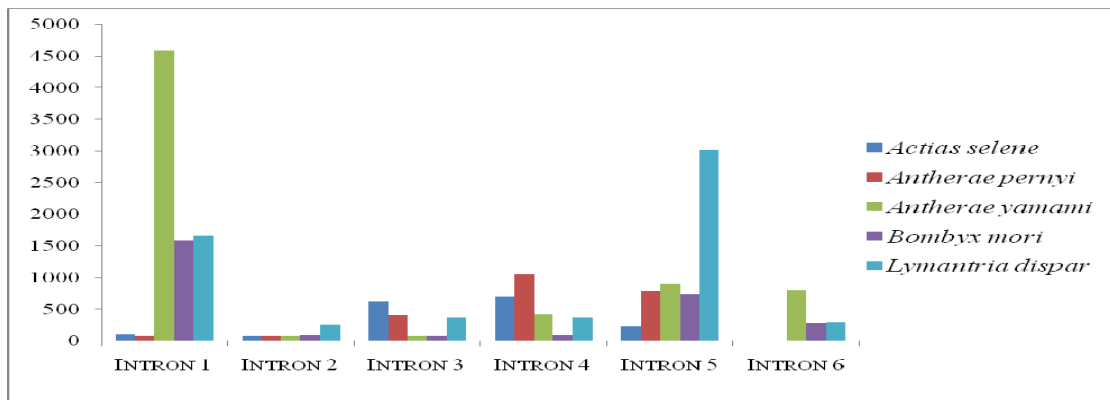


Fig 5: Bar chart showing the size of various introns of vitellogenin gene in Lepidoptera.

3.3 Nucleotide Composition

The nucleotide composition was analysed for coding region of insect vitellogenins (Supplementary File 3). Amongst all the orders, Coleoptera was found to have the highest average A + T content (57.95%), while Diptera was found to have the lowest average A + T content (46.50%) (Table 9). The biasness was observed in nucleotide composition as the A content was found to be slightly higher than the T content at 1<sup>st</sup> and 2<sup>nd</sup> codon position except for the third codon position in all the orders of Insecta except Diptera.

In all dipterans, vitellogenin genes were found to be GC rich and that too particularly having high G+C content at the third codon position as compared to at the 1<sup>st</sup> and 2<sup>nd</sup> codon position. High GC content clearly indicates biasness of the Dipteran Vg genes. The data observed could not be compared because no such analysis seems to have been performed earlier for insect vitellogenins.

Table 9: Comparative nucleotide composition of insect vitellogenin genes.

Order	A+T %			G+C%			AT %	GC %
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>		
Coleoptera	57.25	56.95	59.2	42.70	43.00	40.50	57.95	42.10
Dictyoptera	51.84	57.84	57.28	48.18	36.58	42.88	57.44	42.56
Diptera	46.71	62.77	29.61	53.13	37.08	70.26	46.50	53.47
Hemiptera	54.51	55.93	52.42	45.39	44.00	47.51	54.37	45.61
Hymenoptera	52.59	55.15	53.55	47.3	44.74	46.55	53.80	46.21
Lepidoptera	50.29	61.02	55.39	49.7	39.10	44.67	55.52	44.45

4. Conclusions

Since vitellogenin proteins have been known to play an important biological role in providing nutrients as amino acid building blocks for embryonic development. The sequential and structural differences among introns of the gene validates the fact that the Vgs are evolving at a higher rate although keeping the overall function conserved. The sequence comparisons of Insects vitellogenin genes showed that, except in conserved regions, there are many nonsynonymous substitutions, suggesting that the vitellogenin

genes are highly evolving genes along with selective functional constraints. Moreover, the variations in size and position of exons and introns indicate the divergence of Vg genes among insects. This also indicated that dipteran vitellogenin genes are highly biased.

5. Acknowledgements

The author Monika Miglani highly acknowledges the financial assistance provided by DBT-IPLS, Govt. of India.

## 6. References

- Ahmed AM, Maingon R, Romans P, Hurd H. Effects of malaria infection on vitellogenesis in *Anopheles gambiae* during two gonotrophic cycles. *Insect Molecular Biology* 2001; 10:347–356.
- Kokoza V, Ahmed A, Wimmer EA, Raikhel AS. Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the piggyback transposable element vector pBac [3xP3-EGFP afm]. *Insect Biochemistry and Molecular Biology* 2001; 31:1137-43.
- Raikhel AS, Dhadialla TS. Accumulation of yolk proteins in insect oocytes. *Annual Review of Entomology* 1992; 37:217–251.
- Sappington TW, Raikhel AS. Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect Biochemistry and Molecular Biology* 1998; 28:277–300.
- Van Antwerpen R, Pham DQD, Ziegler R. Accumulation of lipids in insects oocytes. In AS Raikhel & TW Sappington (Eds.) *Progress in Vitellogenesis: Reproductive Biology of Invertebrates*, Enfield, NH: Science Publishers Inc 2005; XII(B):265-288.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 1994; 11; 22(22):4673-80.
- Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 2003; 31(13):3406-15.
- Tufail M, Takeda M. Molecular characteristics of insect vitellogenins. *Journal of Insect Physiology* 2008; 54:1447–1458.
- Yano K, Sakurai M T, Izumi S, Tomino S. Vitellogenin gene of the silkworm, *Bombyx mori*: structure and sex-dependent expression. *FEBS Letters* 1994b; 356:207–211.
- Romans P, Tu Z, Ke Z, Hagedorn HH. Analysis of a vitellogenin gene of the mosquito, *Aedes aegypti* and comparisons to vitellogenins from other organisms. *Insect Biochemistry and Molecular Biology* 1995; 25:939–958.
- Trewitt PM, Heilmann LJ, Degnigillier SS, Kumaran AK. The boll weevil vitellogenin gene: nucleotide sequence, structure, and evolutionary relationship to nematode and vertebrate vitellogenin genes. *Journal of Molecular Evolution* 1992; 34:478-492.
- Hagedorn HH, Maddison DR, Tu Z. The evolution of vitellogenins, cyclorrhaphan yolk proteins and related molecules. *Advances in Insect Physiology* 1998; 27:335–384.
- Migliani M, Gakhar SK. Cloning and bioinformatical analysis of vitellogenin gene of the Indian malaria vector *Anopheles culicifacies* (Diptera: Culicidae). *Acta Entomologica Sinica* 2013; 56(9):1063-1074.
- Isoe J, Hagedorn HH. Mosquito vitellogenin genes: Comparative sequence analysis, gene duplication, and the role of rare synonymous codon usage in regulating expression. *Journal of Insect Science*, 2007; 7: 1-49.
- Nirmala X, Marinotti O, Sandoval JM, Phin S, Gakhar S, Jasinskiene N, James AA. "Functional Characterization of the Promoter of the Vitellogenin Gene, AsVg1, of the Malaria Vector, *Anopheles stephensi*." *Insect Biochemistry and Molecular Biology*. 2006; 36(9): 694–700.
- Chen S, Armistead JS, Javier KN, Sakamoto JM, Rasgon JL. Duplication, concerted evolution and purifying selection drive the evolution of mosquito vitellogenin genes. *BMC Evolutionary Biology*, 2010; 10: 142.
- Hovemann B, Galler R, Walldorf U, Kupper H, Bautz EK. Vitellogenin in *Drosophila melanogaster*: sequence of the yolk protein I gene and its flanking regions. *Nucleic Acids Research*, 1981; 9 (18): 4721-4734.
- Garabedian MJ, Shirras AD, Bownes M, Wensink PC. The nucleotide sequence of the gene coding for *Drosophila melanogaster* yolk protein 3. *Gene*, 1987; 55 (1): 1-8.
- Rina M, Savakis C. A cluster of vitellogenin genes in the Mediterranean fruit fly, *Ceratitis capitata*: sequence and structural conservation in dipteran yolk proteins and their genes. *Genetics*, 1991; 127: 769–780.
- Donnell DM. Vitellogenin of the parasitoid wasp, *Encarsia formosa* (Hymenoptera: Aphelinidae): gene organization and differential use by members of the genus. *Insect Biochemistry and Molecular Biology*, 2004; 34 (9): 951-961.
- Qian C, Liu CL, Zhu BJ, Meng Y, Wei GQ. Identification and expression analysis of vitellogenin from silk-producing insect, *Actias selene* Hubner. *African Journal of Biotechnology*, 2011; 10: 999-1010.
- Liu QN, Zhu BJ, Liu CL, Wei GQ, Wang ZG. Characterization of vitellogenin receptor (VgR) from the Chinese oak silkworm, *Antheraea pernyi*. *Bulletin of Insectology*, 2011; 64: 167-174.
- Meng Y, Liu CL. Vitellogenin gene organization of *Antheraea yamamai* and promoter activity analysis. *Wild. Silkmoth. Silk*. 2006a; 11: 29-40.
- Hiremath E, Lehtoma K. Complete nucleotide sequence of the vitellogenin mRNA from the gypsy moth: novel arrangement of the subunit encoding regions. *Insect Biochemistry and Molecular Biology*, 1997; 27: 27 -35.
- Comas D, Piulachs MD, Belle's X. Vitellogenin of *Blattella germanica* (L.) (Diptoptera, Blattellidae): nucleotide sequence of the cDNA and analysis of the protein primary structure. *Archives of Insect Biochemistry and Physiology*, 2000; 45: 1–11.
- Tufail M, Lee JM, Hatakeyama M, Oishi K, Takeda M. Cloning of vitellogenin cDNA of the American cockroach, *Periplaneta americana* (Diptoptera), and its structural and expression analyses. *Archives of Insect Biochemistry and Physiology*, 2000; 45: 37–46.
- Tufail M, Hatakeyama M, Takeda M. Molecular evidence for two vitellogenin genes and processing of vitellogenins in the American cockroach, *Periplaneta americana*. *Archives of Insect Biochemistry and Physiology*, 2001; 48: 72–80.
- Tufail M, Takeda M. Vitellogenin of the cockroach, *Leucophaea maderae*: nucleotide sequence, structure and analysis of processing in the fat body and oocytes. *Insect Biochemistry and Molecular Biology*, 2002; 32: 117–134.
- Tufail M, Bembenek J, Elgendy AM, Takeda M. Evidence for two vitellogenin-related genes in *Leucophaea maderae*: the protein primary structure and its processing. *Archives of Insect Biochemistry and Physiology*, 2007; 66: 190–203.
- Lee J M, Nishimori Y, Hatakeyama M, Bae TW, Oishi K. Vitellogenin of the cicada *Graptopsaltria nigrofuscata* (Homoptera): analysis of its primary structure. *Insect Biochemistry and Molecular Biology*, 2000b; 30: 1–7.
- Hunter WB, Hunnicutt LE. Sequencing and molecular characterization of a vitellogenin cDNA from the glassy-winged sharpshooter, *Homalodisca coagulata* (Hemiptera: cicadellidae). *Proceedings of Annual International Plant & Animal Genome Conference*. 2006; P. 509. p. 228.
- Huo Y, Liu W, Zhang F, Chen X, Li L. Transovarial Transmission of a Plant Virus Is Mediated by Vitellogenin of Its Insect Vector. *PLoS Pathogens*, 2014; 10(3): e1003949. doi:10.1371/journal.ppat.1003949
- Tufail M, Naemullah M, Elmogy M, Sharma PN, Takeda M, Nakamura, C. Molecular cloning, transcriptional regulation, and differential expression profiling of vitellogenin in two wing-morphs of the brown planthopper, *Nilaparvata lugens* Stal (Hemiptera: Delphacidae) *Insect Molecular Biology*, 2010; 19(6): 787-798.
- Nagaba Y, Tufail M, Inui H, Takeda M. Hormonal regulation and effects of four environmental pollutants on vitellogenin gene transcription in the giant water bug, *Lethocerus deyrollei* (Hemiptera: Belostomatidae). *Journal of Insect Conservation*, 2011; 15: 421-431.
- Lee JM, Hatakeyama M, Oishi K. A simple and rapid method for cloning insect vitellogenin cDNAs. *Insect Biochemistry and Molecular Biology*, 2000a; 30: 189–194.
- Hirai M, Watanabe D, Kiyota A, Chinzei Y. Nucleotide sequence of vitellogenin mRNA in the bean bug *Riptortus clavatus*: analysis of processing in the fat body and ovary. *Insect Biochemistry and Molecular Biology*, 1998; 28: 537– 547.
- Warr E, Meredith JM, Nimmo DD, Basu S, Hurd H, Eggleston P. A tapeworm molecule manipulates vitellogenin expression in the beetle *Tenebrio molitor*. *Insect Molecular Biology*, 2006; 15(4): 497-505.
- Chen J-S, Cho W-L, Raikhel AS. Analysis of mosquito vitellogenin cDNA similarity with vertebrate phosphatases and arthropod serum proteins. *Journal of Molecular Biology*, 1994; 237: 641–647.
- Zimin AV, Smith DR, Sutton G, Yorke JA. Assembly reconciliation. *Bioinformatics*, 2008; 24 (1): 42-45.
- Piulachs MD, Guidugli KR, Barchuk AR, Cruz J, Simoes ZL, Belles X. The vitellogenin of the honey bee, *Apis mellifera*: structural analysis of the cDNA and expression studies. *Insect Biochemistry and Molecular Biology*, 2003; 33(4): 459-465.

41. Li J, Huang J, Cai W, Zhao Z, Peng W, Wu J. The vitellogenin of the bumblebee, *Bombus hypocrite* : studies on structural analysis of the cDNA and expression of the mRNA. *Journal of Comparative Physiology B*, 2010; 180(2): 161-170.
42. Nose Y, Lee JM, Ueno T, Hatakeyama M, Oishi K. Cloning of cDNA for vitellogenin of the parasitoid wasp, *Pimpla nipponica* (Hymenoptera: Apocrita: Ichneumonidae): Vitellogenin primary structure and evolutionary considerations. *Insect Biochemistry and Molecular Biology*, 1997; 2: 1047- 1056.
43. Lewis DK, Campbell JQ, Sowa SM, Chen ME, Vinson SB, Keeley LL. Characterization of vitellogenin in the red imported fire ant, *Solenopsis invicta* (Hymenoptera : Apocrita : Formicidae). *Journal of Insect Physiology*, 2001; 47: 543-551.
44. Tian H, Vinson SB, Coates CJ. Differential gene expression between alate and dealate queens in the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). *Insect Biochemistry and Molecular Biology*, 2004; 34: 937-949.
45. Ye GY, Dong SZ, Song QS, Shi M, Chen XX, Hu C. Molecular cloning and developmental expression of the vitellogenin gene in the endoparasitoid, *Pteromalus puparum*. *Insect Molecular Biology*, 2008; 17(3): 227-233.
46. Blank S, Seismann H, McIntyre M, Ollert M, Wolf S, Bantleon FI, Spillner E. Vitellogenins Are New High Molecular Weight Components and Allergens (Api m 12 and Ves v 6) of *Apis mellifera* and *Vespula vulgaris* Venom. *PLoS ONE*, 2013; 8(4): E62009
47. Meng Y, Liu CL, Shiomi K, Nakagaki M, Banno Y, Kajiura Z. Genetic variations in the vitellogenin of Japanese populations of the wild silkworm *Bombyx mandarina*. *Journal of Insect Biotechnology and Sericology*, 2006b; 75: 127-134.
48. Yano K, Toriyama – sakurai M, Watabe S, Izumi S, Tomino S. Structure and expression of mRNA for vitellogenin in *Bombyx mori*. *Biochimica et Biophysica Acta*, 1994; 1218: 1 -10.
49. Meng Y, Liu CL, Shiomi K, Nakagaki M, Banno Y, Kajiura Z. Purification and cDNA cloning of vitellogenin of the wild silkworm, *Saturnia japonica* (Lepidoptera: Saturniidae). *Journal of Insect Biotechnology and Sericology*, 2008; 77: 35-44.
50. Huber SG, Nardelli D, Haefliger JA, Cooper DN, Givel F, Germond JE, Engel J, Green N M, Wahli W. Precursor-product relationship between vitellogenin and the yolk proteins as derived from the complete sequence of a *Xenopus* vitellogenin gene. *Nucleic Acids Research*, 1987; 15: 12.

## 7. Supplementary files:

1. Supplementary file 1:  
<http://www.biosciencejournals.com/Supplementary/1.pdf>
2. Supplementary file 2  
<http://www.biosciencejournals.com/Supplementary/2.pdf>
3. Supplementary file 3  
<http://www.biosciencejournals.com/supplementary/3.pdf>