Research in Biotechnology 2024, 15: 15-25
doi: 10.25081/rib.2024.v15.9214
https://updatepublishing.com/journal/index.php/rib

Research Article

ISSN: 2229-791X

Biophysical and molecular evolutionary analysis reveals evidence of micro-evolution in the seminal fluid protein-Diazepam-binding inhibitor (DBI) of a Heliothine insect *Helicoverpa armigera*

Kiran Kumar Halagur Bhogegowda[1]*, D. J. Kiran Kumar[2], Rama Thyloor[3], G. M. Manohar[4]

[1]Former Post-doc NCBS, Bangalore, Wilson Garden, Bangalore-560030, affiliated to Nrupathunga University, Bengaluru-560001, Karnataka, India, [2]Department of Biotechnology, Government Science College, (Nrupathunga University), Bengaluru, 560001, Karnataka, India, [3]Department of Biotechnology, Government Science College, (Nrupathunga University), Bengaluru, 560001, Karnataka, India, [4]HalliLabs, Ragihalli, Bangalore-560083, Karnataka, India

ABSTRACT

Several seminal fluid peptides (SFP) are critical for gametogenesis and reproduction related physiological processes in insects. Male seminal fluid induced mating results in post mating physiological responses (PMR) in the female, further impacting the reproductive success of the mating pair. We have previously reported Diazepam-binding inhibitor (DBI) protein in the lepidopteran *Helicoverpa armigera* with species specific PMR response. In the present study, we study the biophysical properties of the DBI protein with bioinformatics methods, further; we map its origin and diversification in the Heliothine clade using molecular evolutionary methods. Our analysis suggests unique biophysical properties of the protein such as four α helices, high exposed and disordered regions. Further, the Proteins B-Value and ProNA values are indicative of its roles in lipid metabolism. High aliphatic amino-acid composition and conservation of protein domain at unique residues along with the hydrophobicity and transmembrane index are indicative of the relative solubility of amino acid residues conferring adaptability. Evolutionary analysis indicated the gene has undergone selection. Further, several unique evolutionary-constrained domain residues/regions (ECRs) in the protein are suggestive of their roles in reproduction related physiological mechanisms. Our data implicate that the protein DBI has undergone evolutionary variation through the micro-evolution process enabled through adaptive mutation in the proteins conferring flexibility and adaptability, thus ensuring genus and species specific specificity.

Received: September 09, 2024
Revised: November 05, 2024
Accepted: November 06, 2024
Published: November 27, 2024

***Corresponding Author:**
Kiran Kumar Halagur Bhogegowda
kirankumarhb@gmail.com

KEYWORDS: Seminal fluid (SF), Post-mating response (PMR), Diazepam binding inhibitor (DBI), Divergence and *de novo* origin, Adaptive mutations

INTRODUCTION

Mating is an important event for all insect orders since it facilitates genes to be transferred and is essential to the reproduction and evolution of the species (Nanfack-Minkeu & Sirot, 2022). Over the last decade, research has focused on the behavioral patterns and physiological alterations related to reproduction (Anholt *et al.*, 2020). In addition to transferring their spermatozoa, male insects during mating also impart non-sperm components in their ejaculate, causing physiological changes in females that facilitate reproduction referred to as PMR. Seminal fluid proteins (SFPs) have a range of relative molecular weights, from 36 amino to 200 amino acids. Proteases and their inhibitors, lectins, prohormone precursors, peptides, and antioxidants are among the diverse protein classes that are found in the seminal fluid (SF) (Ravi Ram & Wolfner, 2007; Avila *et al.*, 2011; Mcgraw *et al.*, 2015). Several insects, including the model organism *Drosophila melanogaster* belonging to all orders, have had SFPs extensively defined (Gioti, 2012; Meuti

& Short, 2019). This is suggestive conservation of the proteins and their significance in gametogenesis and reproduction related physiology. Few of the changes associated with PMR include increased egg production (Ameku & Niwa, 2016), ovulation (Rubinstein & Wolfner, 2013), oviposition and feeding (Carvalho *et al.*, 2006). Additionally, there is a decline in female receptivity (Sirot *et al.*, 2011). SFPs indirectly increase the male's chances of reproductive success through competition. Males induce both short- and long-term memories through SFPs and neural networks. PMRs in insects include a switch from protein to neuronal signaling resulting in a substantial impact on the female behavior and organ functions. Consequently, this affects the resolution of successful reproduction and impacts sexual conflict. Several unique structural features of the SFPs may account for their biochemical and physiological roles in insects.

Sexual reproduction involves both cooperation and antagonism between male and female insects. Seminal proteins delivered by the male often show evidence of rapid evolution under positive selection (Haerty *et al.*, 2007). Significant selection pressure is applied at the molecular level during sexual conflict, necessitating the quick evolution of SFPs to survive competition for resources (Ramm *et al.*, 2009; Sirot *et al.*, 2014). Additionally, sexual conflicts between males (sperm rivalry) and the "interests" and strategies of females and males have an impact on these characteristics (Swanson & Vacquier, 2002; O'Grady & Markow, 2012). These wooing and mating rituals lead to mechanisms for species isolation, which suggests the rapid evolution of reproductive traits. This highly plastic nature can quickly establish barriers between species.

Seminal Fluid Proteins in *Helicoverpa armigera*

Understanding insect reproductive molecules and their functions requires an examination of the nature and function of SFPs even in diverse order and in closely related species. Since, SFPs are essential for insect reproduction and their functions across insect order suggesting conservation (Walters & Harrison, 2010; Sirot *et al.*, 2011; Goenaga *et al.*, 2015). However, identification is challenging especially in non-model insects. Molecular and proteomics methods including EST with proteomic analysis and computational methods have been developed to address these shortcomings and are under use widely. The genus *Helicoverpa* belongs to the family Heliothine which includes *H. armigera*, a devastating agricultural pest in several areas of the world (CABI, 2018). Sex pheromone mediated chemosensory mechanism mediated mating is common in noctuid moths (Zhang *et al.*, 2015). These mechanisms promote reproductive and evolutionary uniqueness and enable isolation both within and across species. Researchers have identified several peptides in the genus *Helicoverpa* such as the 6.03 kDa in *H. armigera* (Kiran *et al.*, 2021), 4.9 kDa (Rama *et al.*, 2024a). A list of diverse classes of SFPs identified through proteomics is listed in Table 1. As evident from the table a wide range of molecules including viz., chaperones, endopeptidase, enzymes and hormones are present in the SF. These peptides modulate a range of functions including protein folding, protein binding, inhibitor activity and catalysis. Further, each of the molecules has unique domains enabling various PMR responses. In our previous studies we have shown several PMR in *H. armigera* such as reduced receptivity, oviposition rates and longevity (Thyloor *et al.*, 2016, 2021). Our proteomic study identified a DBI protein unique to *H. armigera*. The initial structure-function relationship investigation revealed unique biophysical features which augur its PMR functions.

In the present study, we extend the investigation to several protein biophysical and molecular evolutionary methods which would shed light on the proteins species-specific adaptation.

MATERIALS AND METHODS

Protein Analysis

The fasta sequence of the protein was downloaded from (https://www.ncbi.nlm.nih.gov/protein/) and saved for all subsequent analysis. Several URLs were used for the analysis as listed below.

1. Secondary structure-Alphafold prediction was done using the https://alphafold.ebi.ac.uk
   Structure annotation was done using the https://predictprotein.org/, the variables analyzed were
   a. Solvent accessibility
   b. Disordered region

Table 1: List of SFPs and their key features identified in *H. armigera* through proteomics methods

S. No.	Description	Function (GO-ontology)	Cellular location	Unique feature of protein domain
1	Odorant-binding protein, partial	Odorant-binding	Secreted	Chain
2	Heat shock protein 70	ATP-dependent protein folding chaperone	endoplasmic reticulum lumen	Coiled coil
3	Diazepam-binding inhibitor	Fatty-acyl-CoA binding	Mitochondria	Ligand- Acyl-CoA-binding
4	Elongation factor 1-alpha, partial	GTP binding	cytoplasm	tr-type G
5	Thioredoxin	Thioredoxin-disulfide reductase (NADP) activity	Mitochondria	FAD/NAD(P)-binding Pyridine nucleotide-disulphide oxidoreductase dimerisation
6	Serpin	Serine-type endopeptidase inhibitor activity	extracellular space	Serpin domain-containing protein
7	Bombyxin B-12	Hormone	Secreted	insulin family. C peptide like
8	PBAN-type neuropeptides	Hormone-Neuropeptide	Secreted	Disordered
9	Antitrypsin	Serine protease inhibitor	Secreted	Disordered
10	Sorbitol dehydrogenase	Oxidoreductase	Enzyme	Enoyl reductase (ER)

c. Relative b-value and
d. Protein binding
e. Amino acid composition

2. https://predictprotein.org/was used to analyze the follow properties
   a. Transmembrane helices DeepTMHMM
   b. Hydrophobicity Plot
   c. Relative mutability
   d. Evolutionary analysis
3. The fasta sequence of the nucleotide was downloaded from (https://www.ncbi.nlm.nih.gov/nucleotide/) and saved for all subsequent analysis. The URLs used for the analysis are listed below.
   a. Multiple sequence alignment and Phylogenetic tree were constructed using http://www.phylogeny.fr/muscle and https://www.ebi.ac.uk/goldman-srv/webprank
   b. Interpro domain analysis (https://www.ebi.ac.uk/interpro/)
   c. Ka/ks analysis (KaKs_Calculator 3.0)
   d. Adaptive selection pressure analysis (https://www.datamonkey.org/)
   e. Protein constraint analysis (http://www.aminode.org/)

RESULTS

Our hypothesis was that the protein's biophysical characterization would reveal its special characteristics and shed light on how evolutionary selection has shaped it over a timescale. The prevailing premise for globular proteins like DBI that are inherently disordered, is that the incomplete conservation of physicochemical characteristics and domain architecture may exert larger effects on the sequence adaptability (Dunker *et al.*, 2008). The following biophysical properties were carried out to infer the protein secondary structure, hydrophobicity plot, transmembrane tendency, mutability and the evolutionary analysis.

Acyl-CoA-binding protein (ACBP), also known as diazepam binding inhibitor (DBI) or endozepine (EP) because of its ability to displace diazepam from the benzodiazepine (BZD) recognition site. By acting along with (GABA) type A receptor it is involved in several physiological processes (Uniprot). ACBP is a 10 Kd membrane protein that binds acyl-CoA esters and function as an intracellular carrier. The intact ACB domains of the protein are its defining feature. These proteins have been found in several eukaryotic species (Islinger *et al.*, 2020). Secondary structure analysis of the protein suggested that the acyl-CoA-binding site of the ACB domain composed of four bowl-shaped alpha helices. The protein's four α helices (67% of amino acids deduced from human acyl-CoA binding domain 7) were proposed by 3D modeling, which used the crystal structure of the insect *Maduca sexta* acyl-CoA-binding protein as a template (Figure 1). The acyl chain is positioned between the hydrophobic surfaces of CoA and the protein, the ligand is bound by conserved positive interactions with residues on the protein with phosphate group on the adenosine-3'phosphate moiety. Figures 2a and b depict the conservation residues, secondary structure and hydrophobicity 3D map of the protein. As evident from the Figure 2 buried residues were distributed throughout the protein with high residues in the leader peptide. Signal cleavage is at amino acid 26 with score of 0.5429 (SignalP-5.0). The hydrophobic residues were high in the core region of the protein. Predict protein suggests the secondary structure to contain higher percentage of helix, followed by strand and others. An important property of proteins that determines folding and stability is solvent accessibility quantified by the ratio of accessible or buried residues in a protein's three-dimensional

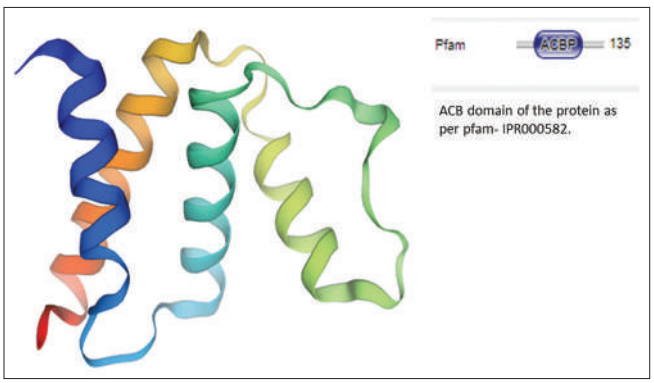

Figure 1: 3D model (Ribbon representation) of *H. armigera* DBI modeled using Maduca sexta acyl-CoA-binding protein depicting α helices and architecture of ACB domain

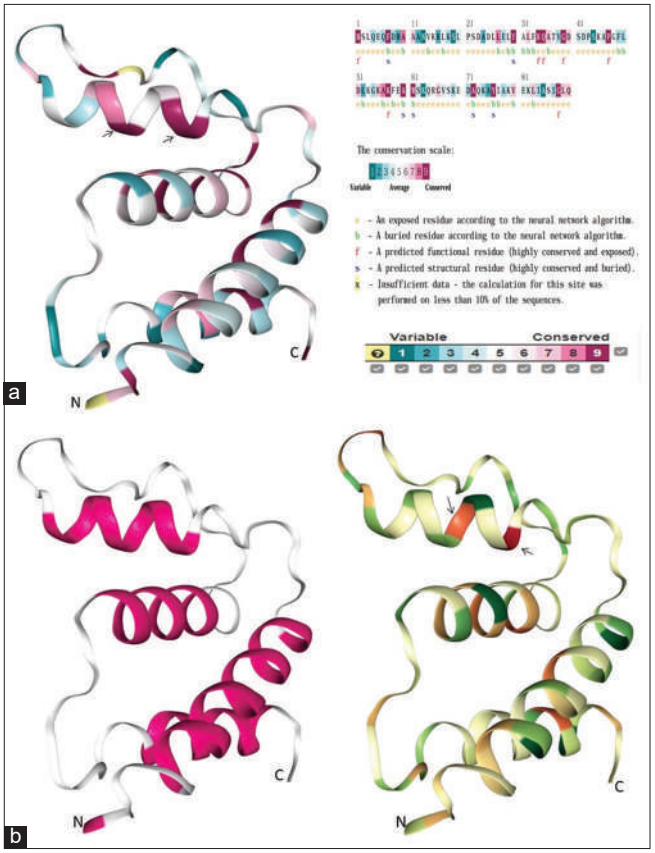

Figure 2: a) ALPHAFOLD MONOMER V2.0 based identification of functionally conversed regions in Diazepam-binding inhibitor protein. Arrow head indicated highly conserved residues, and b) ALPHAFOLD MONOMER V2.0 based identification of conserved secondary structure and hydrophobicity regions in Diazepam-binding inhibitor protein. Arrow head indicates hydrophobic residues

structure (Raghunathan, 2024). The exposed portions were higher than the buried remnants. Hydrophobic residues that are buried and form the core allow the integrity of proteins to be maintained enabling interaction with other proteins or ligands through the exposed regions. Since the protein lacked b-strands, we reasoned if disordered regions were more common. The disordered areas were observed at amino-acid sequences (1-49), (52), and (54-135) respectively. This highlights its importance in molecular recognition with ligands; additionally, the disorder-to-order transitions enable interactions with various partners and confer flexibility (Craveur *et al.*, 2015). The ProteinB and ProNA analysis is depicted in Figures 3a and b. The proteins residues, function and mobility on the surface are tightly related to rigid or flexibility (Almeida *et al.*, 2021). ProteinB with intermediate values were distributed throughout the protein, followed by high and low values. Protein Binding (ProNA) values are indicators of a protein's binding properties with other proteins, DNA, and RNA, and are essential for understanding its role in biological processes (Littmann *et al.*, 2021). Several blocks of high (RI: 00-33) were observed and few blocks of low (RI: 34-66) across the protein (reliability score: 21; GO: 0005515). The indices of aromaticity and aliphaticity were 7.1% and 89.07%, respectively. Further as evident from the Figure 4, the protein's amino acid composition indicates that it has high residues of the following amino acids: Lysine (K), leucine (Leu/L), alanine (Ala/A), serine (Ser/S), valine (Val/V), Phenylalanine (Phe/F) also other amino acids in decreasing order of percentages. As evident from the Figures 5a, b and c the relative solubility of amino acid residues (hydrophobic exterior, hydrophilic interior) is indicated by hydrophobicity index. The transmembrane(TM) tendency and surface-exposed areas are implicated by the Kyte-Doolittle scale of 2.278, highlighting the protein's molecular interaction property (Roterman *et al.*, 2022). Analysis of the amino acids with propensity for TM for the protein is suggestive of the appreciable effects of hydrophobic segments on transmembrane helix formation. Hydrophobicity scale and the transmembrane tendency scale are correlated. Grand average of hydropathicity (GRAVY) was -0.332 indicative of its globular and transmembrane properties. The likelihood that a specific amino acid can change to another resulting in a sequence of biophysical changes during evolution is indicated by the relative mutability (Rm) (Hecht *et al.*, 2013). Three peak values above 90 were observed viz., peak1-aa residue between

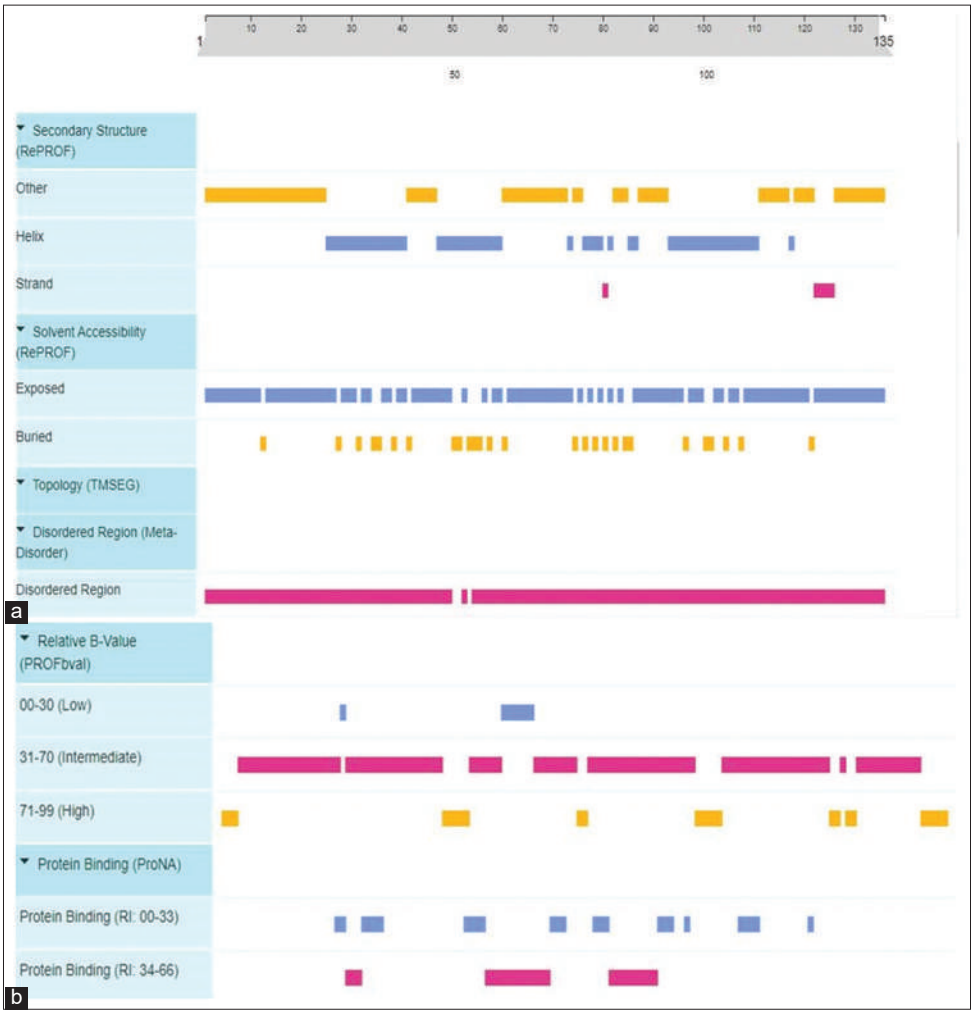

Figure 3: a) PredictProtein analysis of protein Topology – Secondary structure, Solvent accessibility, Disordered region, and b) PredictProtein Relative – value, Protein binding

(20-40), 2-(70), and 3-aa (110-120). In summary these unique biophysical properties of the protein recapitulates its functions as an adaptable globular and membrane protein.

Evolutionary Analysis

Multiple sequence alignment and phylogenetic tree generation

The Figures 6a and b depicts the sequence alignment and Phylogenetic tree of DBI protein generated using few insects' species belonging to all taxa. Overall conservation across the protein was low also several amino acids residues were unique to a particular species. Further, very few amino acid residues were conserved in all the taxa, followed by individual amino acids and blocks of conservation across taxa. The rooted phylogenetic analysis grouped the proteins in two groups viz., a-Diptera including *Drosophila melanogaster* and *Aedes aegypti* and b-Lepidoptera includes - *Papilio polytes* and *Manduca sexta*. The lepidopteran *Danus plexipus* separated from the main branch. A Maximum likelihood (ML) phylogenetic tree of Lepidoptera (500 boot strap values) constructed using the DBI protein sequence is depicted in Figure 6c.

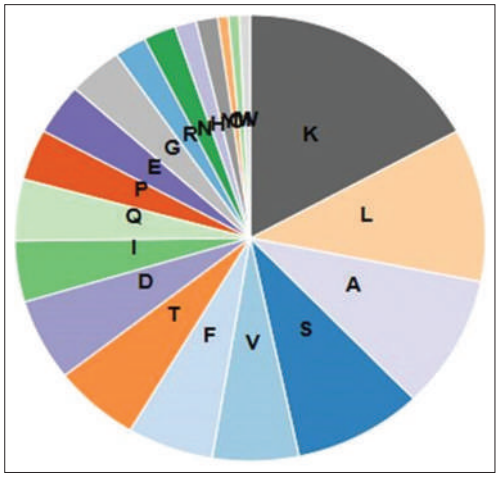

Figure 4: Amino acid composition

Interpro analysis of the ACBP domain shows conservation across eukaryotes (Eubacteria and Archaea) and in certain viruses (Figure 7). In eukaryotes, the domain was distributed in metazoans (mammals, aves, insects, nematodes and few plant species). Further, their presence in rotifers, fungi and algae is suggestive of their role in reproduction related physiological roles across the tree of life.

Ka/Ks Analysis

Measuring the rate of evolutionary changes in protein sequence is frequently determined by the Ka/Ks ratio (number of nonsynonymous substitutions (Ka) to the number of synonymous substitutions synonymous (Ks)). The Ka/Ks substitution, for the DBI was (0.87) indicating that the gene could be evolving as a result of positive selection.

Adaptive selection pressure analysis suggested 81 (probability-0.9) and 5 (p-value threshold 0.1) episodic selection sites of purifying selection through FUBAR and MEME methods respectively.

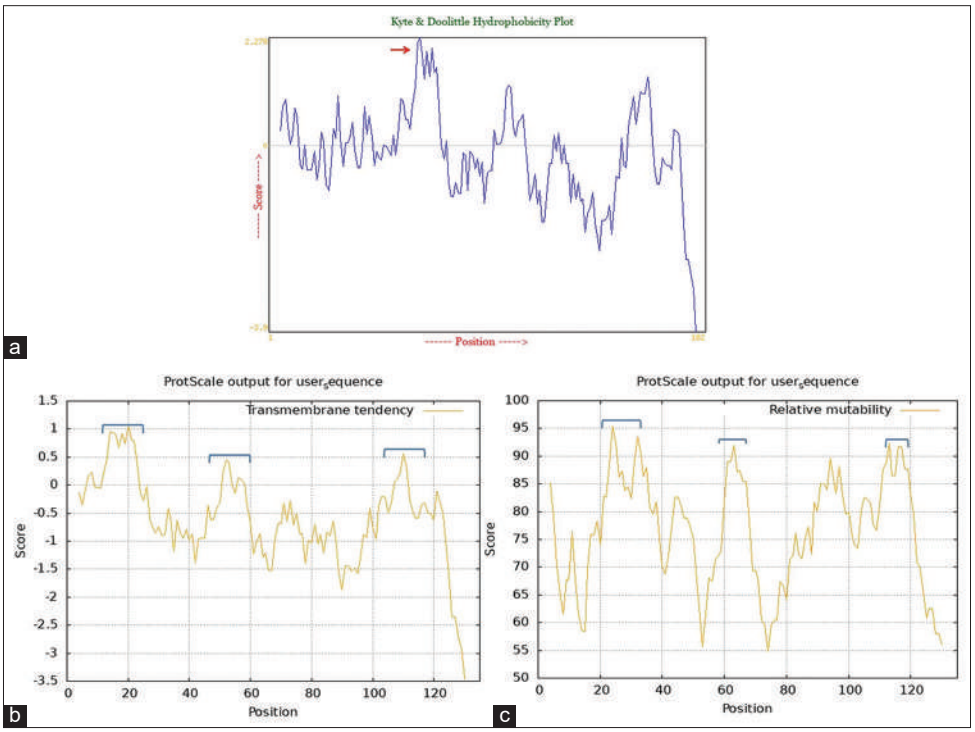

Figure 5: ProtScale analysis. a) Transmembrane tendency (Brackets indicate high score), b) Hydrophobicity Plot (Arrow indicates peak score) and c) Relative mutability (Brackets indicates score above 90)

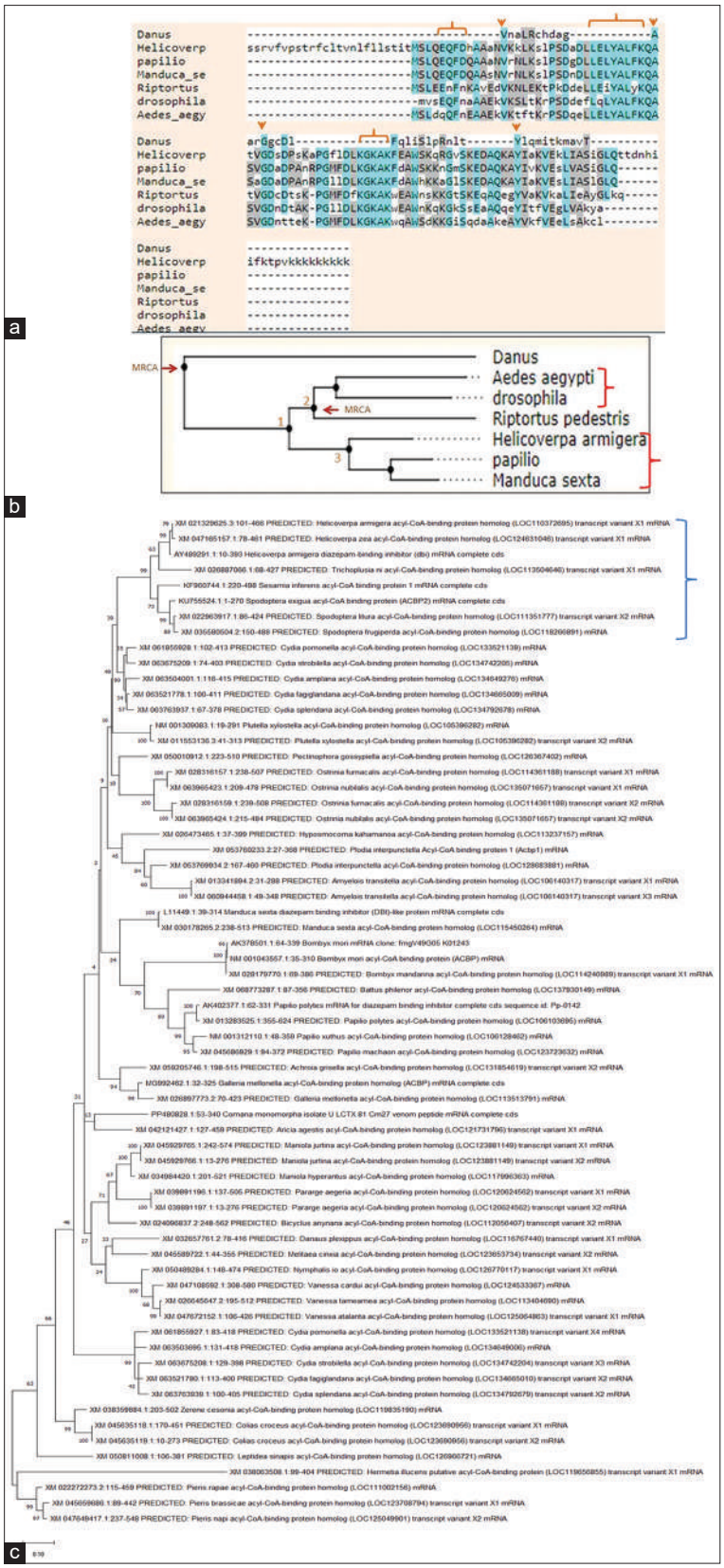

Figure 6: a) Multiple sequence alignment of DBI protein in selected insects belonging to all taxa. (Arrow and brackets highlight conserved amino acids. b) Phylogenetic tree of DBI protein constructed using insects sequence from all taxa (Arrow indicates MRCA and nodes) .c) An ML phylogenetic tree of Lepidoptera (500 boot strap values) constructed used the DBI protein sequence. (Bracket indicate Noctuidae). Details of alignment-Aligned in Mafft and refined suing CD HIT for removal of 100 percent identical sequences. The remaining 65 sequences were exported in MEGA X aligned using Muscle. The alignment was trimmed only to retain the CDS and other default parameters.

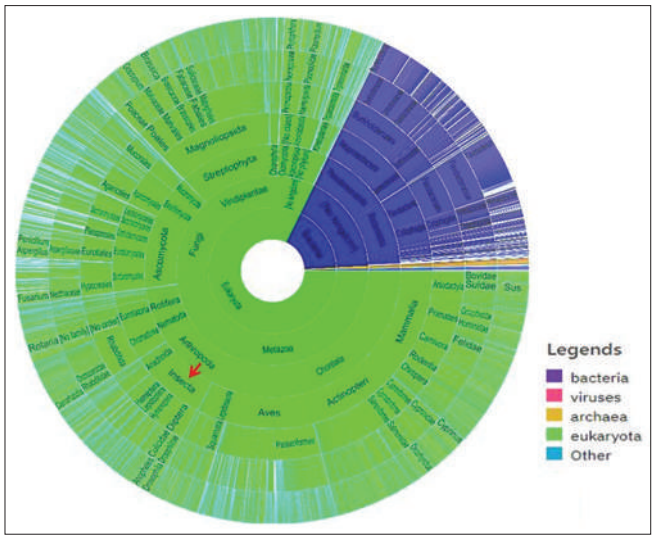

Figure 7: Interpro analysis of conservation of ACBP domain across Phyla (arrow indicates insects)

Evolutionary-constrained regions (ECRs) are unique protein motifs that have evolved through an evolutionary time scale (Wang *et al.*, 2019). Analysis suggested four ECR regions at a residues (10-22, 34-46, 55-64 and 79-88) in Figure 7. The protein is conserved with few residues (1-10) showing lower conservation. This is indicative of the unique cellular and biophysical properties of the protein and regions that are under functional constraint.

DISCUSSION

During mating, the female of many insect species experiences physiological changes that increase reproductive fitness. This PMR is triggered by interactions between the female reproductive system and proteins found in SFP. Certain SFP proteins and genes are missing from insect lineages, raising questions over their genesis and evolutionary paths. Structural analysis of SFP proteins suggests three distinct structural and functional domains (Degner *et al.*, 2019): first is the N-terminal domain, which serves as an anchorage to mediate the peptide's attachment to the sperm tail, enabling SP sequestration in the female reproductive organs. Second, a C-terminal domain with an intramolecular di-sulphide domain activates the neuronal G protein-coupled receptor responsible for initiating many of the PMR responses. Additionally, it stimulates the biosynthesis of juvenile hormones. The third central domain, rich in hydroxyprolines, is responsible for eliciting an early gene response to release several peptides in post-mated females (Peng *et al.*, 2024). Thus, the female hemolymph experiences several biophysical changes that are determined by physiological factors like pH and metabolic rates as well as biological factors including insect taxonomy, diet, microbial load and cellular stress. These modifications are maintained even at the intra- and interspecies level. According to the biophysical investigations of the protein, the solvent accessibility and disorder of the protein revealed that the ratio of exposed to buried residues was higher, and the disordered residues further contributed to the flexibility. This observation is strengthened by the high indices of amino acid aliphaticity. The ProteinB and ProNA analysis suggest mobility of residues on the surface of a protein thereby contributing to its function in lipid metabolism. Since, the N and C terminals of the protein are involved in important functions such as anchorage and signaling functions it would be expected that evolution will disfavor flexibility and adaptation in these regions leading to perturbation of key signaling pathways and fitness loss.

The globular and transmembrane characteristics of the protein are suggested by the hydrophobicity plot, transmembrane helix propensity and relative mutability. Together, these physiological and biological alterations can be accommodated by the SFP owing to its biophysical flexibility. Therefore, it is possible to propose that the DBI protein undergoes several biophysical adaptations at the protein level through mutational mechanisms by changes to pre-existing protein over a timescale leading to divergence in genesis and evolution.

The two main theories that underpin the genesis and evolutionary routes of SFP are the divergence and *de novo* origin (Andersson *et al.*, 2015; Singh *et al.*, 2020). The process by which new genes arise from non-genic parental DNA sequences is referred to as "de novo gene birth". The mechanism involved is gain-of-function which facilitates adjustments of the regulatory regions and formation of new genes. This theory has enabled to explain fitness in several insect species and is a source of evolutionary innovation (Grandchamp *et al.*, 2023). The process by which an already-existing gene undergoes such significant alterations over time that the original sequence can no longer be preserved is known as divergence. Several insect genomes have evolved through this mechanism of evolution (Brand *et al.*, 2024). The function of natural selection in the evolution of proteins has long been a focus of molecular evolution research (Jayaraman *et al.*, 2022). Strong purifying (negative) selection pressure is applied on the proteins ensuring a reduction in amino acid alterations.

The DBI conservation across the protein was low with very few amino-acid residues conserved in all the taxa, followed by individual amino acids and blocks of conservation across taxa. This observation is suggestive of unique evolutionary trajectories in each species. The ACB domain with motif sequence LELYALFKQA was conserved with few variations in *Drosophila melanogaster* and *Riptortus pedestris*. This variability is supported by the differences in biophysical properties in proteins belonging to other species (data not shown).Also, through adaptive selection pressure of purifying selection. It could be therefore summarized that the protein has evolved with distinct residues to support its biological functions in a particular species contingent upon the variables. The rooted phylogenetic analysis supports this notion where the proteins are grouped to Diptera including *D. Melanogaster* and *Aedes aegypti* and Lepidoptera including *Papilio polytes* and *Manduca sexta* which show a higher degree of sequence conservation. The lepidopteran

Danus plexipus separated from the main MRCA (most recent common ancestor). Further, the *R. pedestris* branched from node 2. In summary the evolutionary analysis suggests that the protein has evolved under selection pressure through species specific domains which confer species specific advantages.

Purifying selection eliminates deleterious non-synonymous mutations, leading to the conservation of amino acid sequences leading to adaptation (Charlesworth *et al.*, 2015).

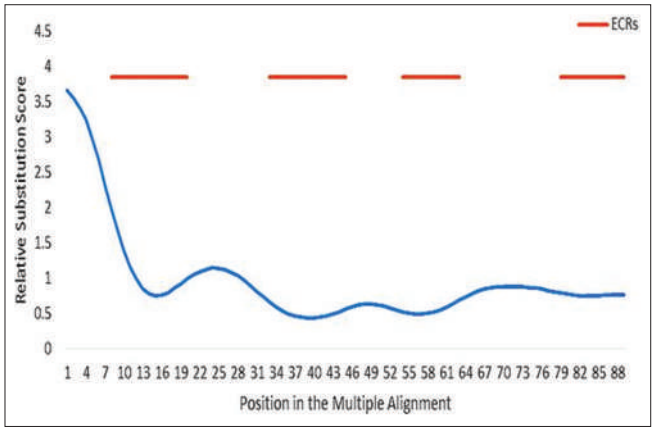

Figure 8: Evolutionary conserved constrain analysis depicted conserved domains in the protein

The ratio of non-synonymous to synonymous substitutions is an evolutionary method to identify genomic sequences that have undergone significant mutational changes and to quantify the degree of selection (Del Amparo *et al.*, 2021). This model captures the tendency to conserve the physico chemical properties of amino acids when they undergo mutations, like polar amino acids that are mostly substituted by other polar amino acids but less frequently by hydrophobic ones. The Ka/Ks values enable estimation of molecular evolutionary forces that have shaped the protein domain/s enabling stringency or flexibility to biochemical activity (binding, catalysis, etc.) and biophysical properties (folding, stability, etc.).

Few regions or Domains in a protein are naturally disordered or unstructured. These regions have unstable structures and are extremely flexible (van der Lee *et al.*, 2014). In order to study a protein's function during evolutionary time, biologists usually concentrate on these conserved sections of the protein. Its significance is seen in the four regions of constraint (Figure 8), in the DBI protein that the ECR analysis proposes for a broad range of taxa (data not shown). The sluggish development trend of such a short DBI peptide is further supported by this observation. Furthermore, it appears that the protein adopts various interconverting structural states based on our results and several others (Forman-Kay & Mittag 2013; Dishman & Volkman, 2018). Disordered areas facilitate fidelity in biological functions including reproduction (in

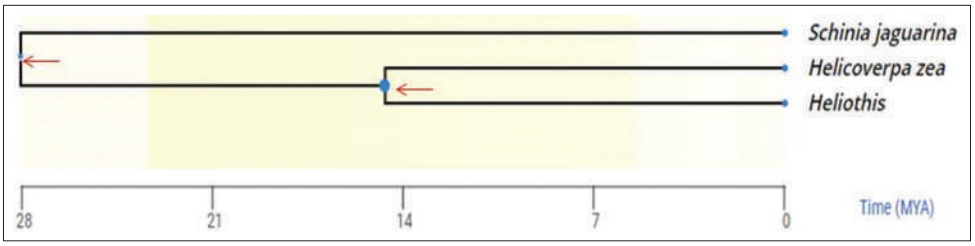

Figure 9: Evolutionary timescale of divergence of Heliothines

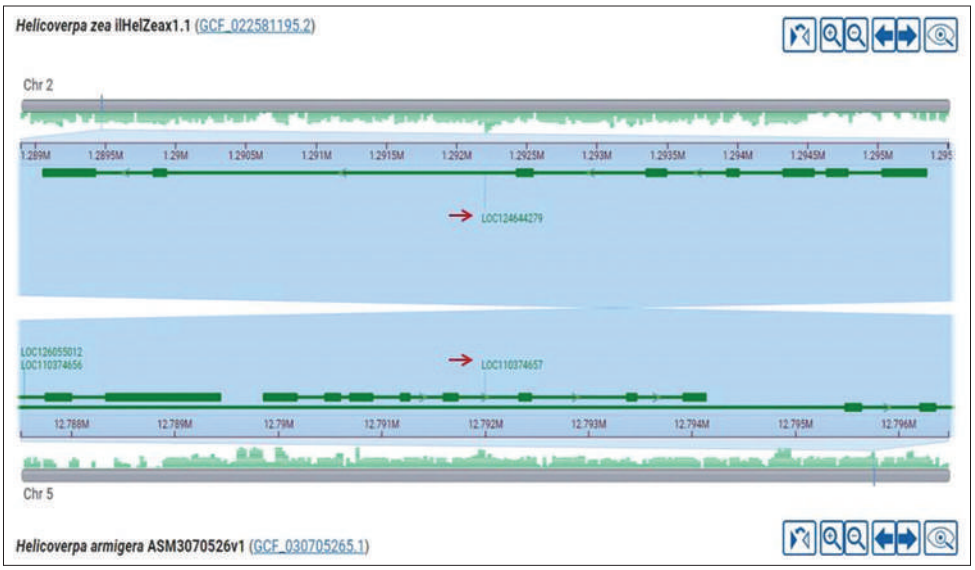

Figure 10: Gene re-arrangemnet in the helicoverpa spices (Arrow indicate DBI gene)

this case, lipid metabolism). Given that the DBI protein is relatively short with 90 amino acids, and assuming little to no selective pressure in maintaining its sequence, the protein has diverged significantly from its ancestral state to its present form in *H. armigeria*.

CONCLUSION

The fast evolution of fertilization proteins, which contribute to reproductive isolation between divergent taxa, is a distinct feature of insects (Carlisle & Swanson, 2021). Speciation requires an awareness of the contributors to reproductive isolation (Kulmuni *et al.*, 2020) as demonstrated in animals and plants (Moyle *et al.*, 2021; Hopkins & Perry 2022). An investigation of the DBI through the protein evolution lens suggests several unique evolutionary features such as the correlation between the ratios of exposed and buried solvent exposure sites in the folded protein as indicator of adaptive selection. Next in globular proteins, surface residues are mostly polar and charged, while core residues have a higher tendency to be hydrophobic. Mutations that conserve these properties are less likely to result in a large change in stability. In addition, the folding speeds of proteins with alpha-helices have increased throughout evolution compared to beta-sheets (Sikosek & Chan, 2014). Also, there is evolutionary pressure towards faster folding and stability. Finally, adaptation towards increased conformational flexibility could have acted as a check against proteins evolving to become extremely stable.

Evolutionary investigations on protein motifs or domains can reveal insights into protein/s involved in fertilization mechanics. Insect taxa use non-homologous proteins in PMR. For example, the moth *Helicoverpa zea* uses pheromonostatic peptide (Kingan *et al.*, 1995), the mosquito A. *aegypti* uses Head Protein 1 (Naccarati *et al.*, 2012), and *D. melanogaster* uses SP (Sex Peptide) (Yang *et al.*, 2024), *H. armigeria* uses DBI (Rama *et al.*, 2024b). This pattern suggests that high degree of evolutionary turnover of regulators with new regulators evolving and old regulators being lost from populations. Study of Heliothines genomes by Pearce *et al.* (2017) suggests extensive amplification and neofunctionalization of genes. Also, the nonsynonymous/synonymous sites have rapidly diverged between species and paralogs of other insects orders (Bartolomé *et al.*, 2005).

Evolution generates and optimizes new traits through "adaptation" facilitated by mutations that enable malleability to changing milieu. Compensatory mutations allow potential adaption of deleterious mutations to persist over a timescale (Ruelens *et al.*, 2023). In insects, this is demonstrated by the evolution of insect esterase's against insecticide resistance (Menozzi *et al.*, 2004). The micro-evolutionary events model of protein evolution proposes a confined 'random walk' through the fitness space model (Lynch *et al.*, 2005). The compensatory mutations buffer the biophysical strain created by adaptive mutations (Brown *et al.*, 2010).

Protein evolution occurs across length scales, stability, and time scales (Tiana *et al.*, 2004). Such timescales unfold episodes of specialization, reductive evolutionary tendencies of architectural repertoires in proteomes. In specific the evolutionary timescale of the Heliothis is estimated to be 28MYA and approximately 14MYA for the *Helicoverpa zea* (timetree.org). Thus it could be inferred that the DBI protein diverged along with *Schinia jaguarina* around 14MYA (Figure 9).

A recent study by Chen *et al.* (2023) implicates chromosomal re-arrangement, and movement of individual genes in Lepidoptera contributing to evolution (Stewart & Rogers, 2019). Comparative Genome Viewer suggests chromosomal gene movement event of the DBI in *H. zea* and *H. armigeria* (Figure 10) (https://www.ncbi.nlm.nih.gov/cgv/). The rates by which reproductive proteins evolve compared to others proteins vary within species (Garlovsky & Ahmed-Braimah, 2003; Turner & Hoekstra, 2008). This is facilitated by various biophysical adaptations in the protein backbone. In summary afore evolutionary results summarize that the DBI globular transmembrane proteins with variable residues in different species evolve over long time intervals with species specific endpoints.

The study has few limitations we have not included several biophysical methods in the analysis such as kinetics, folding, role of co-evolution of domains and networks. Also, the alternative genetic mechanisms such as drift, and epistasis were also not tested. Future studies directed towards including these variables and also detailed characterization of proteins (N, C terminals) and the characterization of cognate receptors will help delineate the role of this protein in species-specific PMR in *H. armigeria*.

ACKNOWLEDGEMENT

We thank the Honourable Vice-chancellor, Nrupathunga University for encouragement and logistics. Dedications - To the sacred feet of Bhaghavan and Gurus of Sringeri peetams. Sowmyashree for technical help in preparation of manuscript.

AUTHORS' CONTRIBUTION

Design, literature collation, data analysis and manuscript drafting Kiran Kumar H. B., literature collation, drafting Kiran Kumar D. J., literature review, collation of data and drafting Rama Thyloor, Data analysis, critical inputs and drafting Manohar G. M.

REFERENCES

Almeida, F. C. L., Sanches, K., Pinheiro-Aguiar, R., Almeida, V. S., & Caruso, I. P. (2021). Protein Surface Interactions-Theoretical and Experimental Studies. *Frontiers in Molecular Biosciences, 8*, 706002. https://doi.org/10.3389/fmolb.2021.706002

Ameku, T., & Niwa, R. (2016). Mating-Induced Increase in Germline Stem Cells via the Neuroendocrine System in Female *Drosophila*. *PLoS Genetics, 12*, e1006123. https://doi.org/10.1371/journal.pgen.1006123

Andersson, D. I., Jerlström-Hultqvist, J., & Näsvall, J. (2015). Evolution of new functions de novo and from preexisting genes. *Cold Spring Harbor Perspectives in Biology, 7*(6), a017996. https://doi.org/10.1101/cshperspect.a017996

Anholt, R. R. H., O'Grady, P., Wolfner, M. F., & Harbison, S. T. (2020). Evolution of reproductive behavior. *Genetics, 214*, 49-73. https://doi.org/10.1534/genetics.119.302263

Avila, F. W., Sirot, L. K., LaFlamme, B. A., Rubinstein, C. D., & Wolfner, M. F. (2011). Insect seminal fluid proteins: identification and function. *Annual Review of Entomology, 56*, 21-40. https://doi.org/10.1146/annurev-ento-120709-144823

Bartolomé, C., Maside, X., Yi, S., Grant, A. L., & Charlesworth, B. (2005). Patterns of selection on synonymous and nonsynonymous variants in Drosophila miranda. *Genetics, 169*(3), 1495-1507. https://doi.org/10.1534/genetics.104.033068

Brand, C. L., Oliver, G. T., Farkas, I. Z., Buszczak, M., & Levine, M. T. (2024). Recurrent Duplication and Diversification of a Vital DNA Repair Gene Family Across Drosophila. *Molecular Biology and Evolution, 41*(6), msae113. https://doi.org/10.1093/molbev/msae113

Brown, K. M., Costanzo, M. S., Xu, W., Roy, S., Lozovsky, E. R., & Hartl, D. L. (2010). Compensatory mutations restore fitness during the evolution of dihydrofolate reductase. *Molecular Biology and Evolution, 27*(12), 2682-2690. https://doi.org/10.1093/molbev/msq160

CABI. (2018). *Helicoverpa armigera*. Wallingford, UK: Center for Agriculture and Bioscience International.

Carlisle, J. A., & Swanson, W. J. (2021). Molecular mechanisms and evolution of fertilization proteins. *Journal of Experimental Zoology - B Molecular and Developmental Evolution, 336*(8), 652-665. https://doi.org/10.1002/jez.b.23004

Carvalho, G. B., Kapahi, P., Anderson, D. J., & Benzer, S. (2006). Allocrine modulation of feeding behavior by the Sex Peptide of *Drosophila*. *Current Biology, 16*(7), 692-696. https://doi.org/10.1016/j.cub.2006.02.064

Charlesworth, B. (2015). Causes of natural variation in fitness: evidence from studies of Drosophila populations. *Proceedings of the National Academy of Sciences of the United States of America, 112*(6), 1662-1669. https://doi.org/10.1073/pnas.1423275112

Chen, X., Wang, Z., Zhang, C., Hu, J., Lu, Y., Zhou, H., Mei, Y., Cong, Y., Guo, F., Wang, Y., He, K., Liu, Y., & Li, F. (2023). Unraveling the complex evolutionary history of lepidopteran chromosomes through ancestral chromosome reconstruction and novel chromosome nomenclature. *BMC Biology,* 21(1), 265. https://doi.org/10.1186/s12915-023-01762-4

Craveur, P., Joseph, A. P., Esque, J., Narwani, T. J., Noël, F., Shinada, N., Goguet, M., Leonard, S., Poulain, P., Bertrand, O., Faure, G., Rebehmed, J., Ghozlane, A., Swapna, L. S., Bhaskara, R. M., Barnoud, J., Téletchéa, S., Jallu, V., Cerny, J.,… de Brevern, A. G. (2015). Protein flexibility in the light of structural alphabets. *Frontiers in Molecular Biosciences, 2*, 20. https://doi.org/10.3389/fmolb.2015.00020

Degner, E. C., Ahmed-Braimah, Y. H., Borziak, K., Wolfner, M. F., Harrington, L. C., & Dorus, S. (2019). Proteins, Transcripts, and Genetic Architecture of Seminal Fluid and Sperm in the Mosquito *Aedes aegypti*. *Molecular & Cellular Proteomics: MCP, 18*(Suppl 1), S6-S22. https://doi.org/10.1074/mcp.RA118.001067

Del Amparo, R., Branco, C., Arenas, J., Vicens, A., & Arenas, M. (2021). Analysis of selection in protein-coding sequences accounting for common biases. *Briefings in Bioinformatics, 22*(5), bbaa431. https://doi.org/10.1093/bib/bbaa431

Diana E. Wheeler 2009. *Encyclopedia of Insects* (Second Edition) 2009.

Dishman, A. F., & Volkman, B. F. (2018). Unfolding the Mysteries of Protein Metamorphosis. *ACS Chemical Biology, 13*(6), 1438-1446. https://doi.org/10.1021/acschembio.8b00276

Dunker, A. K., Silman, I., Uversky, V. N., & Sussman, J. L. (2008). Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology, 18*(6), 756-764. https://doi.org/10.1016/j.sbi.2008.10.002

Forman-Kay, J. D., & Mittag, T. (2013). From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure, 21*(9), 1492-1499. https://doi.org/10.1016/j.str.2013.08.001

Garlovsky, M. D., & Ahmed-Braimah, Y. H. (2023). Evolutionary Quantitative Proteomics of Reproductive Protein Divergence in *Drosophila*. *Molecular & Cellular Proteomics, 22*(8), 100610. https://doi.org/10.1016/j.mcpro.2023.100610

Gioti, A., Wigby, S., Wertheim, B., Schuster, E., Martinez, P., Pennington, C. J., Partridge, L., & Chapman, T. (2012). Sex peptide of Drosophila melanogaster males is a global regulator of reproductive processes in females. *Proceedings of the Royal Society B Biological Sciences, 279*(1746), 4423-4432. https://doi.org/10.1098/rspb.2012.1634

Goenaga, J., Yamane, T., Rönn, J., & Arnqvist, G. (2015). Within-species divergence in the seminal fluid proteome and its effect on male and female reproduction in a beetle. *BMC Evolutionary Biology, 15*, 266. https://doi.org/10.1186/s12862-015-0547-2

Grandchamp, A., Kühl, L., Lebherz, M., Brüggemann, K., Parsch, J., & Bornberg-Bauer, E. (2023). Population genomics reveals mechanisms and dynamics of de novo expressed open reading frame emergence in *Drosophila melanogaster*. *Genome Research, 33*(6), 872-890. https://doi.org/10.1101/gr.277482.122

Haerty, W., Jagadeeshan, S., Kulathinal, R. J., Wong, A., Ram, K. R., Sirot, L. K., Levesque, L., Artieri, C. G., Wolfner, M. F., Civetta, A., & Singh, R. S. (2007). Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila. *Genetics, 177*(3), 1321-1335. https://doi.org/10.1534/genetics.107.078865

Hecht, M., Bromberg, Y., & Rost, B. (2013). News from the protein mutability landscape. *Journal of Molecular Biology, 425*(21), 3937-3948. https://doi.org/10.1016/j.jmb.2013.07.028

Hopkins, B. R., & Perry, J. C. (2022). The evolution of sex peptide: sexual conflict, cooperation, and coevolution. *Biological Reviews of the Cambridge Philosophical Society, 97*(4), 1426-1448. https://doi.org/10.1111/brv.12849

Islinger, M., Costello, J. L., Kors, S., Soupene, E., Levine, T. P., Kuypers, F. A., & Schrader, M. (2020). The diversity of ACBD proteins - From lipid binding to protein modulators and organelle tethers. *Biochimica et biophysica acta. Molecular Cell Research, 1867*(5), 118675. https://doi.org/10.1016/j.bbamcr.2020.118675

Jayaraman, V., Toledo-Patiño, S., Noda-García, L., & Laurino, P. (2022). Mechanisms of protein evolution. *Protein Science, 31*(7), e4362. https://doi.org/10.1002/pro.4362

Kingan, T. G., Bodnar, W. M., Raina, A. K., Shabanowitz, J., & Hunt, D. F. (1995). The loss of female sex pheromone after mating in the corn earworm moth *Helicoverpa zea*: identification of a male pheromonostatic peptide. *Proceedings of the National Academy of Sciences of the United States of America, 92*(11), 5082-5086. https://doi.org/10.1073/pnas.92.11.5082

Kiran, T., Mangala, J. N., Anjana, K., & Manjulakumari, D. (2021). Heterologous expression of a substance which inhibits receptivity and calling in *Helicoverpa armigera* (Hübner). *Insect Molecular Biology, 30*(5), 472-479. https://doi.org/10.1111/imb.12718

Kulmuni, J., Butlin, R. K., Lucek, K., Savolainen, V., & Westram, A. M. (2020). Towards the completion of speciation: the evolution of reproductive isolation beyond the first barriers. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 375*(1806), 20190528. https://doi.org/10.1098/rstb.2019.0528

Littmann, M., Heinzinger, M., Dallago, C., Weissenow, K., & Rost, B. (2021). Protein embeddings and deep learning predict binding residues for various ligand classes. *Scientific Reports, 11*(1), 23916. https://doi.org/10.1038/s41598-021-03431-4

Lynch, M. (2005). Simple evolutionary pathways to complex proteins. *Protein Science, 14*(9), 2217–2227. https://doi.org/10.1110/ps.041171805

McGraw, L. A., Suarez, S. S., & Wolfner, M. F. (2015). On a matter of seminal importance. *BioEssays, 37*(2), 142-147. https://doi.org/10.1002/bies.201400117

Menozzi, P., Shi, M. A., Lougarre, A., Tang, Z. H., & Fournier, D. (2004). Mutations of acetylcholinesterase which confer insecticide resistance in *Drosophila melanogaster* populations. *BMC Evolutionary Biology, 4*, 4. https://doi.org/10.1186/1471-2148-4-4

Meuti, M. E., & Short, S. M. (2019). Physiological and Environmental Factors Affecting the Composition of the Ejaculate in Mosquitoes and Other Insects. *Insects, 10*(3), 74. https://doi.org/10.3390/insects10030074

Moyle, L. C., Wu, M., & Gibson, M. J. S. (2021). Reproductive Proteins Evolve Faster Than Non-reproductive Proteins among *Solanum* Species. *Frontiers in Plant Science, 12*, 635990. https://doi.org/10.3389/fpls.2021.635990

Naccarati, C., Audsley, N., Keen, J. N., Kim, J. H., Howell, G. J., Kim, Y. J., & Isaac, R. E. (2012). The host-seeking inhibitory peptide, Aea-HP-1, is made in the male accessory gland and transferred to the female during copulation. *Peptides, 34*(1), 150-157. https://doi.org/10.1016/j.peptides.2011.10.027

Nanfack-Minkeu, F., & Sirot, L. K. (2022). Effects of Mating on Gene Expression in Female Insects: Unifying the Field. *Insects, 13*(1), 69. https://doi.org/10.3390/insects13010069

O'Grady, P. M., & Markow, T. A. (2012). Rapid morphological, behavioral, and ecological evolution in *Drosophila*: Comparisons between the endemic Hawaiian *Drosophila* and the cactophilic repleta species group. In R. S. Singh, J. Xu, R. J. Kulathinal (Eds.), *Rapidly Evolving Genes and Genetic Systems* (pp. 176-186) Oxford, UK: Oxford University

Press. https://doi.org/10.1093/acprof: oso/9780199642274.003.0018
Pearce, S. L., Clarke, D. F., East, P. D., Elfekih, S., Gordon, K. H. J., Jermiin, L. S., McGaughran, A., Oakeshott, J. G., Papanikolaou, A., Perera, O. P., Rane, R. V., Richards, S., Tay, W. T., Walsh, T. K., Anderson, A., Anderson, C. J., Asgari, S., Board, P. G., Bretschneider, A.,… Wu, Y. D. (2017). Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive Helicoverpa pest species. *BMC Biology, 15*(1), 63. https://doi.org/10.1186/s12915-017-0402-6
Peng, J., Svetec, N., Molina, H., & Zhao, L. (2024). The Origin and Evolution of Sex Peptide and Sex Peptide Receptor Interactions. *Molecular Biology and Evolution, 41*(4), msae065. https://doi.org/10.1093/molbev/msae065
Raghunathan, S. (2024). Solvent accessible surface area-assessed molecular basis of osmolyte-induced protein stability. *RSC Advances, 14*(34), 25031-25041. https://doi.org/10.1039/d4ra02576h
Rama, T., Kiran Kumar, H. B., & Doddamane, M. (2024b). Characterization of seminal fluid peptides/proteins (SFPs) of male *Helicoverpa armigera* and their plausible role in post-copulatory modulation of female reproductive behavior. *Environmental and Experimental Biology,* Accepted Unpublished Article.
Rama, T., Kiran Kumar, H. B., Saraswathi, S., & Doddamane, M. (2024a). Role of Mating Receptivity Peptides in Helicoverpa armigera: Molecular and Evolutionary Perspectives. *Journal of Advances in Biology & Biotechnology, 27*(8), 1427-1447. https://doi.org/10.9734/jabb/2024/v27i81266
Ramm, S. A., McDonald, L., Hurst, J. L., Beynon, R. J., & Stockley, P. (2009). Comparative proteomics reveals evidence for evolutionary diversification of rodent seminal fluid and its functional significance in sperm competition. *Molecular Biology and Evolution, 26*(1), 189-198. https://doi.org/10.1093/molbev/msn237
Ravi Ram, K., & Wolfner, M. F. (2007). Seminal influences: Drosophila Acps and the molecular interplay between males and females during reproduction. *Integrative and Comparative Biology, 47*(3), 427-445. https://doi.org/10.1093/icb/icm046
Roterman, I., Stapor, K., & Konieczny, L. (2022). The Contribution of Hydrophobic Interactions to Conformational Changes of Inward/Outward Transmembrane Transport Proteins. *Membranes, 12*(12), 1212. https://doi.org/10.3390/membranes12121212
Rubinstein, C. D., & Wolfner, M. F. (2013). Drosophila seminal protein ovulin mediates ovulation through female octopamine neuronal signaling. *Proceedings of the National Academy of Sciences of the United States of America, 110*(43), 17420-17425. https://doi.org/10.1073/pnas.1220018110
Ruelens, P., Wynands, T., & de Visser, J. A. G. M. (2023). Interaction between mutation type and gene pleiotropy drives parallel evolution in the laboratory. *Philosophical transactions of the Royal Society of London Series B, Biological Sciences, 378*(1877), 20220051. https://doi.org/10.1098/rstb.2022.0051
Sikosek, T., & Chan, H. S. (2014). Biophysics of protein evolution and evolutionary protein biophysics. *Journal of the Royal Society, Interface, 11*(100), 20140419. https://doi.org/10.1098/rsif.2014.0419
Singh U, & Syrkin Wurtele E.How new genes are born. (2020) Causes and consequences of the evolution of reproductive proteins. *Elife.* Feb 19;9:e55136. doi: 10.7554/eLife.55136.
Sirot, L. K., Wolfner, M. F., & Wigby, S. (2011). Protein-specific manipulation of ejaculate composition in response to female mating status in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America, 108*(24), 9922-9926. https://doi.org/10.1073/pnas.1100905108
Sirot, L. K., Wong, A., Chapman, T., & Wolfner, M. F. (2014). Sexual conflict and seminal fluid proteins: a dynamic landscape of sexual interactions. *Cold Spring Harbor Perspectives in Biology, 7*(2), a017533. https://doi.org/10.1101/cshperspect.a017533
Stewart, N. B., & Rogers, R. L. (2019). Chromosomal rearrangements as a source of new gene formation in *Drosophila yakuba*. *PLoS Genetics, 15*(9), e1008314. https://doi.org/10.1371/journal.pgen.1008314
Swanson, W. J., & Vacquier, V. D. (2002). The rapid evolution of reproductive proteins. *Nature Reviews Genetics, 3*, 137-144. https://doi.org/10.1038/nrg733
Thyloor, R., Kiran, T., & Doddamane, M. (2016) Cost of mating: A study on physiological trade-offs between fecundity and longevity in laboratory reared *Helicoverpa armigera* (Hubner) female moths. *Journal of Entomology and Zoology Studies, 4*(6), 768-771.
Thyloor, R., Kiran, T., Manjulakumari, D. (2021). Unreceptivity–a phenomenon induced by male accessory gland secretions in female *Helicoverpa armigera*. *Environmental and Experimental Biology, 19*(2), 97-102. https://doi.org/10.22364/eeb.19.10
Tiana, G., Shakhnovich, B. E., Dokholyan, N. V., & Shakhnovich, E. I. (2004). Imprint of evolution on protein structures. *Proceedings of the National Academy of Sciences of the United States of America, 101*(9), 2846-2851. https://doi.org/10.1073/pnas.0306638101
Turner, L. M., & Hoekstra, H. E. (2008). Causes and consequences of the evolution of reproductive proteins. *The International Journal of Developmental Biology, 52*(5-6), 769-780. https://doi.org/10.1387/ijdb.082577lt
van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E., & Babu, M. M. (2014). Classification of intrinsically disordered regions and proteins. *Chemical Reviews, 114*(13), 6589-6631. https://doi.org/10.1021/cr400525m
Walters, J. R., & Harrison, R. G. (2010). Combined EST and proteomic analysis identifies rapidly evolving seminal fluid proteins in Heliconius butterflies. *Molecular Biology and Evolution, 27*(9), 2000-2013. https://doi.org/10.1093/molbev/msq092
Wang, F., Wang, K., Forknall, N., Patrick, C., Yang, T., Parekh, R., Bock, D., & Dickson, B. J. (2020). Neural circuitry linking mating and egg laying in *Drosophila* females. *Nature, 579*(7797), 101-105. https://doi.org/10.1038/s41586-020-2055-9
Wang, S.-W., Bitbol, A.-F., & Wingreen, N. S. (2019). Revealing evolutionary constraints on proteins through sequence analysis. *PLoS Computational Biology, 15*(4), e1007010. https://doi.org/10.1371/journal.pcbi.1007010
Yang, Y. T., Hu, S. W., Li, X., Sun, Y., He, P., Kohlmeier, K. A., & Zhu, Y. (2023). Sex peptide regulates female receptivity through serotoninergic neurons in *Drosophila*. *iScience, 26*(3), 106123. https://doi.org/10.1016/j.isci.2023.106123
Zhang, J., Walker, W. B., & Wang, G. (2015). Pheromone reception in moths: from molecules to behaviors. *Progress in Molecular Biology and Translational Science, 130*, 109-128. https://doi.org/10.1016/bs.pmbts.2014.11.005