

Unraveling Insect Biology: The Role of Multi-Omics Approaches in Advanced Insect Research

Achyutananda Samal, Anwesha Pradhan, Sajan Sahoo, Srikanta Jena, Luna Samanta, Soumya Ranjan Jena

The multi-omics approach, encompassing genomics, transcriptomics, proteomics, and metabolomics, has revolutionized insect research by enabling a comprehensive view of the molecular interactions that drive insect adaptation, evolution, and ecological roles. Insect species, as the most diverse group of organisms, offer significant potential for understanding evolutionary processes, and multi-omics allows researchers to dissect these processes at multiple molecular levels. Through the integration of bioinformatics tools, complex datasets from different omics platforms can be aligned and interpreted to reveal patterns and networks within insect biology. These tools are essential for managing large volumes of biological data and for visualizing interactions across genes, proteins, and metabolites. In agricultural and environmental contexts, multi-omics enables a detailed exploration of plant-insect interactions, revealing signaling pathways and molecular dialogues that impact pest control and crop resistance. In medical entomology, the study of insect vectors is enhanced, helping researchers understand pathogen transmission and develop vector control strategies. Moreover, multi-omics allows for genome-wide analyses, identification of insect resistance genes, and the development of insect-targeted interventions. This integrated, data-driven approach thus offers unprecedented insights into insect biology, promising advancements in pest management, evolutionary study, biodiversity conservation, and the control of insect-borne diseases.

Keywords: *Insect, Multi-omics, Data integration, Genomics, Proteomics, Transcriptomics, Metabolomics, Entomological database*

Achyutananda Samal¹, Anwesha Pradhan¹, Sajan Sahoo¹, Srikanta Jena¹, Luna Samanta¹, Soumya Ranjan Jena^{1,2*}

¹Redox Biology and Proteomics Laboratory, Department of Zoology, Ravenshaw University, Cuttack, Odisha, India.

²ICMR-PDRA, Department of Zoology, Ravenshaw University, Cuttack, Odisha 753003, India.

*Email: srjena@ravenshawuniversity.ac.in

©Access: CC BY-NC

Publisher: Cornous Publications LLP, Puducherry, India.

Applied Entomology and Zoology (Volume 1)

Editor: Dr. Anil Kumar

ISBN: 978-81-976294-7-1

DOI: <https://doi.org/10.37446/volbook082024/38-61>

Introduction

Insects are among the most significant and diverse organisms on Earth, representing over 80% of all described animal species. Their ecological, agricultural, and biomedical importance makes them indispensable to life. Insects pollinate approximately 85% of angiosperms, regulate pest populations, recycle nutrients, and serve as a vital part of food webs. In terms of evolutionary success, insects are unparalleled: their lineage length, species variety, adaptive diversity, biomass, and ecological influence are unrivalled. However, insects also significantly impact human health as vectors of diseases like malaria, dengue, and chagas disease. Despite their critical roles, insect populations are under increasing threat from habitat destruction, climate change, and pesticide use, highlighting an urgent need for extensive research to understand and conserve this invaluable group of organisms. Recent advancements in molecular biology and bioinformatics have revolutionized the study of insects, providing unprecedented insights into their biology, evolution, and interactions with the environment. Traditional methods of insect study, while foundational, are often labor-intensive and limited in scope. In contrast, modern techniques like genomics, transcriptomics, proteomics, and metabolomics collectively known as omics that enable researchers to analyze complex biological processes with remarkable precision and efficiency. For instance, genomic sequencing of insect vectors such as *Anopheles gambiae* has provided critical information about malaria transmission, facilitating the development of targeted vector control strategies (Touré et al., 2004). Similarly, transcriptomic studies in pollinators like honeybee (*Apis mellifera*) have revealed the genetic underpinnings of behaviours crucial to ecosystem health, such as foraging and communication (Ma et al., 2019). Additionally, innovative methods like environmental DNA (eDNA) and metagenomics are being harnessed to monitor insect biodiversity and track population changes in response to environmental challenges (Ruppert et al., 2019). Comparative genomics has shed light on evolutionary relationships between insect species, identifying conserved pathways that reveal shared mechanisms across taxa. Hox gene studies in *Drosophila melanogaster* and other insects have elucidated the genetic basis of body plan development, providing insights into evolutionary conservation and divergence (Hajirnis, N., & Mishra, R. K. 2021). Multi-omics, an interdisciplinary approach within system biology allows researchers to investigate and analyse these interactions at all scales, from genes to ecosystems. Bioinformatics tools have become integral to these advancements, unlocking the potential of multi-omics integration by allowing researchers to integrate and analyze vast datasets generated by molecular studies. By combining genomic, proteomic, and metabolomic data, researchers can achieve a holistic understanding of the various aspects of insect biology including resistance proteins, gene expression, sequence pattern of gene mutations, qualitative and quantitative analysis of proteins and their interactions at the genomic level. These multi-omics analyses provide a better knowledge of vector biology, pathogen transmission, and the development of vector control measures, such as genetically engineering insects to restrict their ability to spread the infectious diseases (Singh et al., 2011; Iquebal et al., 2015). Researchers can identify important genes or pathways that can be targeted to disrupt disease transmission by analysing the molecular interactions between pathogens and insect vectors. Moreover, a thorough comprehension of the structural and functional mechanisms underlying biological processes is approached by multi-omics data integration techniques in insect study. In addition to helping to manage the growing volume of biological data, it includes statistical and mathematical methods using computational algorithms, as well as the visualisation and interpretation of wet bench results. This chapter explores the transformative impact of multi-omics approaches in advancing insect research, emphasizing their ecological, agricultural, and evolutionary importance. It further highlights the role of insect-specific databases, and discusses the challenges and future directions in this rapidly evolving field.

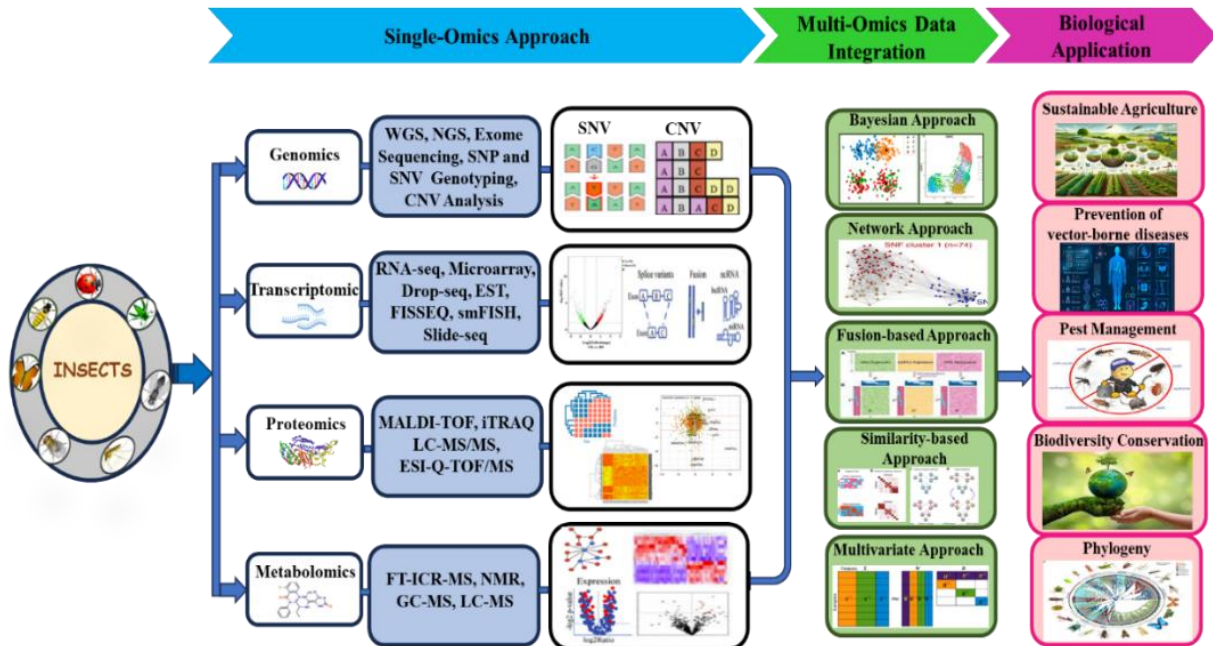


Figure 1. Overview of multi-omics approach in insects and its diverse application.

Omics platforms

Omics refers to a set of technologies used to explore the roles, relationships, and functions of various biological molecules in a comprehensive manner. It aims to understand the complexity of life at different levels, from genes to proteins and metabolites (Gutierrez Reyes et al., 2024). Omics platforms are significant in biology as they provide broad insight into cellular processes, allowing researchers to study organisms systematically rather than focusing on individual components. Key types of omics include genomics, which examines an organism's complete genetic material; transcriptomics, analyzes RNA transcripts to understand gene expression; proteomics, investigates the structure and function of proteins; and metabolomics, focuses on the metabolites involved in metabolic processes. These platforms, when integrated through multi-omics approaches, provide a comprehensive understanding of biological systems, supporting advancements in disease research, agriculture, and environmental biology (Figure 1).

Genomics

Genomics is the multidisciplinary science that explores the structure, function and evolution of genes and their interactions by identification, molecular characterization and cloning of whole genomes for the ultimate goal of understanding phenomics of organisms. Overall genomics includes sequencing of genomes, determination of complete set of genes and understanding the interspecific and intraspecific difference through single nucleotide polymorphisms, loss of heterozygotic, insertion and deletion. Sequence data provides important biological insights into an organism by identifying the genes that make up vital biochemical pathways and the essential cellular components (Richard, 2015). Genome sequencing beginning with the enzymatic dideoxy DNA sequencing technique developed by Frederick Sanger and advancing to next-generation sequencing (NGS) methods (Kamps et al., 2017; Kim et al., 2020). The sequencing data allows researchers to understand how an organism's genetic composition influences its morphology and establish an evolutionary relationship between and within phyla to identify the molecular basis for traits that are unique (Nayduch et al., 2019). Insects represent the most diversified group of animals, with more than

one million species already named, though these represent less than 20% of the total estimated number of insect species (Stork, 2018). Insects play important roles in ecosystems, as well as have significant impacts on agricultural food production and human and animal health. Therefore, it makes more sense to interest our curiosity about the genetic, genomic, and evolutionary foundations of insects. Whole Genome Sequencing (WGS) has become an indispensable tool in evolutionary biology, enabling the comprehensive analysis of insect genomes. This technique involves the sequencing of the entire genome, providing detailed information on gene content, structure, and variation (Ellegren, 2014). The availability of advanced sequencing technologies, such as next-generation sequencing (NGS) platforms, has significantly reduced the cost and time required for WGS, making it feasible to sequence even non-model organisms (Goodwin et al., 2016). For insects, WGS has revolutionized the study of evolutionary processes by offering insights into the genetic basis of traits like development, behaviour, and adaptation to environmental changes. A notable example is the sequencing of the *Drosophila melanogaster* genome, which has been instrumental in understanding evolutionary genetics and molecular biology (Adams et al., 2000). Similarly, the WGS of the malaria mosquito *Anopheles gambiae* has provided key insights into the genetic basis of insecticide resistance, which is crucial for vector control efforts in public health (Holt et al., 2002). The whole-genomic sequencing of insect species now goes a long way in the identification of insect-specific proteins and orthologous groups, greatly enriching our understanding of their evolutionary and adaptive biology compared to other eukaryotes. By analyzing the genomes of various insect species, such as *Drosophila melanogaster*, *Anopheles gambiae*, and *Apis mellifera*, researchers can pinpoint unique proteins that have evolved within the Insecta clade, which may play crucial roles in their ecological success and diversification. *Drosophila melanogaster* is the first insect whose 120-megabase euchromatic genomic region has been sequenced by using a whole-genome shotgun sequencing strategy, enhanced by extensive clone-based sequencing and a high-quality physical map resolution of bacterial artificial chromosomes. The analysis revealed that the genome encodes approximately 13,600 genes (Adams et al., 2000). But till now, 1,345 insect genome size estimates have been published varying 240-fold from 69 Mb in chironomid midges (Cornette et al., 2015) to 16.5 Gb in the mountain grasshopper *Podisma pedestris* (Westerman et al., 1987). Comparative genomics, which involves comparing the genomes of related insect species, plays a crucial role in tracing evolutionary patterns across taxa. By identifying conserved and divergent genomic regions, scientists can infer the evolutionary relationships between species and understand how specific genetic changes have contributed to adaptation and diversification (Rokas & Abbot, 2009). For example, comparing the genomes of different ant species has revealed key insights into the evolution of social behaviour, with evidence showing that variations in gene expression, rather than gene content, are responsible for caste differentiation in social insects (Simola et al., 2013). Similarly, comparisons between the genomes of honeybees and other Hymenoptera have provided valuable information on the evolution of eusociality, a major evolutionary transition (Kapheim et al., 2015). These comparative studies demonstrate how genomic data can be used to identify genetic signatures of adaptation, such as genes involved in immunity, metabolism, and development, offering a deeper understanding of the genetic basis of insect diversity. Population genomics involves studying genetic variation within and between populations of a species to understand evolutionary processes such as adaptation, gene flow, and speciation (Luikart et al., 2019). By analyzing population-level genetic data, researchers can track evolutionary changes over time and across geographic regions. In insects, population genomics has been used to investigate the genetic basis of adaptation to diverse environmental conditions, such as climate and host plant specialization (Martin et al., 2012). For instance, studies on the *Heliconius* butterflies have shown how population genomic approaches can identify genes involved in mimicry and colour pattern evolution, which are key for survival and reproductive success (Martin et al., 2012). Additionally, population genomics has provided insights into speciation, revealing how genetic divergence between populations can lead to the formation of new species (Ellegren, 2014). This approach is particularly

valuable for studying insects with large, geographically dispersed populations, as it allows researchers to assess how genetic variation contributes to local adaptation and ecological diversification. Phylogenomics integrates genomic data to resolve evolutionary relationships between species, offering more precise phylogenetic reconstructions than traditional methods that rely on morphological or single-gene data (Misof et al., 2014). By using large-scale genomic datasets, phylogenomics can overcome limitations such as homoplasy, where similar traits evolve independently in unrelated species. This approach has been particularly useful in resolving complex insect phylogenies, such as those of the Holometabola, a group that includes beetles, flies, and butterflies (Zhan et al., 2011). A major advantage of this approach is the simultaneous analysis of thousands of genes, greatly increasing statistical power and accuracy in resolving evolutionary relationships (Misof et al., 2014). However, challenges remain, such as the computational complexity of analyzing large genomic datasets and the potential for incomplete lineage sorting, where gene trees may not fully reflect species trees (Kapli et al., 2020). Despite these challenges, phylogenomics has provided groundbreaking insights into the timing and pattern of insect diversification, revealing, for example, that major radiations of insect groups coincided with key environmental changes in Earth's history.

Transcriptomics

Transcriptomics, the study of RNA transcripts produced by the genome, provides insights into gene expression and regulation at various stages of development, in different tissues, and under varying environmental conditions. Insects, being the most diverse group of organisms on earth, serve as excellent models for studying gene expression because of their ecological importance, varied life cycles, and remarkable adaptability. Transcriptomic studies in insects have gained substantial traction with the advent of next-generation sequencing (NGS) technologies such as RNA sequencing (RNA-seq), providing high-throughput and detailed data on gene expression patterns. Messenger RNAs (mRNAs), microRNAs (miRNAs), and different kinds of long noncoding RNAs (lncRNAs) are all included in transcriptomics (Milward et al., 2016). In the context of insect-plant interactions, the host plant's transcriptome acts as an arsenal, enabling it to deploy defense responses against herbivores. In the past, transcriptome research was carried out by creating expressed sequence tags (ESTs) using complementary deoxyribonucleic acid (cDNA) clones. These ESTs were then examined using automated Sanger sequencing methods or serial study of gene expression (SAGE). However, transcriptomic investigations presently employ RNA sequencing (direct sequencing) and DNA microarray (indirect sequencing) (Milward et al., 2016). Transcriptomics can also be used to study certain cell activities. The position and form of the cells can also reflect their function, among the many other elements that affect how the cells work. Different microenvironments may exist at different sites of the same tissues in complex biological systems, like mammals. Recently, single-cell transcriptomic and spatial transcriptomic technologies have been created to gain a deeper understanding of the activities of cells. The scientific community has long used fluorescence-activated cell sorting (FACS) and laser microdissection (LMD) to examine the transcriptomes of individual cells. Nonetheless, single-cell transcriptomics has advanced with the advent of droplet-based sequencing tools like Drop-seq. Using a microfluidics-based technique that wraps individual cells in droplets with barcoded beads, the Drop-seq approach measures the mRNA level of transcripts in thousands of single cells. This enables high-throughput screening of single-cell transcriptomes (Rich-Griffin et al., 2020). However, single-cell transcriptomics and spatial transcriptomics are utilised to study the activities of cells that are only found at certain sites. Spatial transcriptomics makes use of technologies such as fluorescent in situ RNA sequencing (FISSEQ) (Moor & Itzkovitz 2017), single-molecule fluorescence in situ hybridisation (smFISH) (Raj et al., 2008), and Slide-seq (Rodriques et al., 2019). Slide-seq is a revolutionary spatial transcriptomics technique that offers spatial transcription analysis capabilities, much as Drop-seq, which is utilised for spatial transcription analysis. To

enable sequencing to determine the locations of the RNA, tissue pieces are placed onto a surface coated in DNA-barcoded beads with predetermined positions in a process called Slide-seq. According to Rodriques et al., it makes possible to examine spatial gene expression at the level of individual cells (Rodriques et al., 2019). Transcriptomics has played a key role in understanding the developmental biology of insects. For instance, studies have used RNA-Seq to uncover stage-specific gene expression in insects like *Drosophila melanogaster* and various Lepidopteran species. For example, transcriptomic analysis of the silkworm (*Bombyx mori*) during metamorphosis identified key regulatory genes involved in moulting and pupation (Zhang et al., 2021). Such findings not only improve our understanding of insect physiology but also have practical applications in pest control, particularly by targeting developmentally essential genes. Transcriptomics has also been employed to investigate insect-plant interactions, particularly in agricultural pests. A transcriptomic analysis of the brown planthopper (*Nilaparvata lugens*), a major rice pest, elucidated the molecular mechanisms behind its adaptation to resistant rice varieties (Xu et al., 2022). Their work revealed differentially expressed genes related to digestion, detoxification, and immunity, providing insights that could aid in the development of resistant crop varieties. The study of insect immunity has greatly benefited from transcriptomics, revealing complex gene networks involved in defense responses against pathogens. An example is the research on the immune response of the honeybee (*Apis mellifera*) to viral infections. A similar transcriptome analysis identified several immune-related genes that are upregulated in response to Deformed Wing Virus (DWV) infections (Pizzorno et al., 2021). Such data is critical for understanding and potentially mitigating the decline in bee populations due to pathogens. Environmental stressors, including climate change, are known to affect insect populations, influencing their behavior, physiology, and survival. Transcriptomic studies have shown how insects respond to temperature fluctuations, drought, and pollution at the molecular level. For instance, transcriptomic analysis of the migratory locust (*Locusta migratoria*), revealed how it modulates gene expression in response to temperature changes (Zhao et al., 2023). These insights are crucial for predicting how insect populations might shift due to global climate change. Transcriptomics has become a key tool in studying the evolution of insects, particularly in understanding speciation and adaptation processes. For example, a study that compared comparing the transcriptomes of two closely related butterfly species revealed divergent expression of genes associated with flight and pigmentation, traits thought to be important for ecological adaptation and reproductive isolation (Liu et al., 2021).

Proteomics

Proteomics, the large-scale study of proteins, has become an essential tool in insect research, providing insights into physiology, development, behavior, and interactions with the environment (Palli et al., 2012). Unlike genomics, which identifies the potential of a genome, proteomics reveals the active molecular players, shedding light on how insects respond to various stimuli, pathogens, and environmental changes. With technological advancements in mass spectrometry (MS), tandem mass spectrometry (MS/MS), and bioinformatics, the proteomic analysis of insects has dramatically improved. Techniques such as isobaric tags for relative and absolute quantification (iTRAQ) and label-free quantification methods allow for accurate protein identification and quantification in different developmental stages, tissues, and species. Furthermore, the use of advanced bioinformatics tools helps in the annotation of proteins and understanding their functional implications, particularly in species with limited genomic data. Proteomics has been instrumental in unraveling the molecular pathways underlying insect development. For example, research on *Drosophila melanogaster* has utilized proteomic techniques to identify proteins involved in metamorphosis, reproduction, and immune responses (Casas-Vila et al., 2017). Similarly, studies on agricultural pests like *Helicoverpa armigera* have highlighted proteins involved in detoxification, providing targets for pest control (Dawkar et

al., 2016). Proteomics enables mapping of protein profiles across developmental stages. Furthermore, studies have explored the changes in the proteome during metamorphosis in insects like *Drosophila melanogaster* and *Bombyx mori* (Lepidoptera) (Si et al., 2020). These studies provide a detailed understanding of the physiological processes governing insect development and how they are regulated at the protein level. Proteomics has been used to study the immune response of insects to pathogens. This is particularly relevant for vector species like mosquitoes, which transmit diseases such as malaria and dengue fever. Proteomic analyses have identified key proteins involved in immune signaling pathways and responses to infections. Another study examined the proteomic changes in *Anopheles gambiae* following infection with *Plasmodium falciparum* (Marie et al., 2014). The authors identified novel proteins associated with immune defense and pathogen recognition, which may help develop strategies for controlling disease transmission by interfering with mosquito immunity. Proteomics has proven to be a valuable tool in understanding the interactions between insects and plants. For herbivorous pests, studying the proteome of both the insect and the host plant can reveal how plants defend themselves and how insects counteract these defenses. For instance, a study on the rice pest *Nilaparvata lugens* (brown) utilized proteomics to investigate the changes in the insect's proteome when feeding on resistant versus susceptible rice plants (Zha & You, 2020). The study highlighted proteins involved in digestion, detoxification, and stress response, providing insights into mechanisms of resistance and susceptibility. Proteomics has also contributed to our understanding of insect evolution through comparative analysis of conserved and species-specific proteins. Comparative proteomic studies help identify conserved and species-specific proteins, contributing to our understanding of insect evolution. A comparative proteomic analysis between alate (winged) and apterous (wingless) morphs of the pea aphid (*Acyrtosiphon pisum*) provided novel insights into wing development and highlighted molecular differences associated with morphological diversification, contributing to our understanding of evolutionary adaptations in aphids (Song et al., 2018). Understanding the tritrophic relationships between plants, insects, and their endosymbionts is another benefit of proteomic technology. Francis discovered that the endosymbionts, which contribute to aphid adaptation to host plant resistance, were the source of the many proteins that were variably regulated throughout these interactions (Francis et al., 2010). Matrix-Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) and LC-MS were used to examine the proteome alterations in plants caused by feeding them the specialised herbivore *Manduca sexta*. The genes of the phytophagous beetles, namely those belonging to the Chrysomeloidea and Curculionoidea families, were believed to encode several enzymes that break down plant cell walls. Proteomics was used to analyse the midgut contents of mustard leaf beetle *Phaedon cochleariae* larvae in order to demonstrate this. Thirteen proteins that break down plant cell wall polysaccharides were found to be members of the xylanase, polygalacturonase, and cellulase families (Kirsch et al., 2012). Aphids' (*Macrosiphum euphorbiae*) reactions to plant stressors, such as water stress and beetle (*L. decemlineata*)-induced defoliation, showed that symbiosis was given priority under stressful circumstances (Nguyen et al., 2007).

Metabolomics

The metabolic profiles of living beings often reflect the responses they have to genetic or environmental alterations. Metabolomics is the high-throughput characterisation of all small-molecule metabolites, including those generated as a result of metabolic pathways on a global scale (Idle & Gonzalez 2007; Palli et al., 2012; Liu & Locasale 2017). Metabolomics has become a pivotal tool in entomology, offering profound insights into insect physiology, behavior, and evolutionary adaptations. By examining the complete set of small molecules in insects, researchers can elucidate metabolic pathways, identify biomarkers, and understand how insects respond to environmental changes. It employs three methods: metabolic profiling, which gathers and examines data from crude extracts to classify them according to all metabolites rather than

breaking them down into individual metabolites; targeted metabolomics, which involves extremely precise quantification of metabolites; and untargeted metabolomics, which screens metabolites of known and unknown identities (qualification). In metabolomics, three different methods are employed: nuclear magnetic resonance (NMR), Fourier transform-ion cyclotron resonance associated with mass spectrometry (FT-ICR-MS), and gas or liquid chromatography paired to mass spectrometry (GC- or LC-MS). While NMR techniques are utilised to detect unknown metabolites, gas chromatography and LC-MS are used to detect and measure known metabolites (Palli et al., 2012). An experiment was conducted on the metabolomic analysis of the larvae during desiccation and rehydration cycles in the anhydrobiotic insect *Polypedilum vanderplanki* revealed the species is capable of surviving almost complete dehydration (Ryabova et al., 2020). This study found that trehalose stabilizes cellular structures during dehydration, provides energy, and aids antioxidant recovery. This study also identified key genes involved in metabolism and stress response providing insights into desiccation tolerance and suggesting potential strategies for developing dry-storage technologies. The study investigates the roles of metabolites in plant-insect interactions in *Chlosyne lacinia* larval development and diapause. The main metabolites identified were sesquiterpenes, fatty acids, long chain alkanes and steroids in *Aldama robusta* leaves, while *C. lacinia* biosynthesize its most of the metabolites were plant-derived. Sesquiterpenes and triterpenes were also detected across larval development and the findings highlight the importance of metabolites in ecosystem evolution (Gallon et al., 2024). The study explores the application of metabolomics in the emerging field of "entometabolomics" to address safety, quality, and traceability challenges in edible insect-based food systems. It highlights the potential of these analytical techniques to improve the sustainability and acceptance of insect-derived food and feed while addressing technical and regulatory hurdles in their industrial application (Poma et al., 2022). The study investigates the metabolic mechanisms underlying paternal defects caused by *Wolbachia* infection in *Drosophila melanogaster*. Using untargeted and targeted metabolomics, the researchers identified significant changes in carbohydrate, lipid, and cofactor metabolism, which are linked to reproductive defects. *Wolbachia* infection enhances energy metabolism but induces oxidative stress, mitochondrial dysfunction, and impaired spermatogenesis. These findings provide insights into how *Wolbachia* manipulates host metabolism to affect reproduction and have implications for controlling pest populations and disease vectors (Zhang et al., 2021).

Integrated omics approach

The integration of omics datasets, known as multi-omics integration, is crucial for uncovering the complex interactions between genes, proteins, and metabolites in insects. It enables researchers to correlate genomic variations with phenotypic outcomes, protein expression with metabolic changes, and gene expression with physiological traits. Multi-omics approaches have been particularly useful in understanding insecticide resistance, pathogen-insect interactions, and adaptations to environmental stressors like temperature and humidity. Data integration in multi-omics involves combining heterogeneous datasets generated from different omics platforms. This process requires advanced computational tools and statistical methods to ensure accurate data alignment, normalization, and interpretation. Several bioinformatics tools and databases have been developed for multi-omics data integration in insect research (Roychowdhury et al., 2023). One key challenge in multi-omics data integration is the heterogeneity of the datasets, as they vary in data structure, resolution, and dynamic range. Various computational approaches, such as machine learning algorithms and network-based methods, have been employed to integrate these datasets. These methods help identify patterns, interactions, and regulatory networks, providing a comprehensive view of the molecular mechanisms at play in insects.

Tools for data integration

Multi-omics methods yield data in a number of forms, the majority of which are machine-readable. The genes, transcripts, proteins, and metabolites that are often represented in generated omics data from different platforms may be linked in network pathways. A deeper comprehension of the biology of insect system is made possible by the integration and interception of omics data, which is made possible by bioinformatics. Bioinformatics is the application of mathematical methods and algorithms to biological data using tools such as databases and data mining software. Data processing, molecular identification, statistical data analysis, pathway and network analysis, and system modelling are all included in omics data analysis. Additionally, it makes possible to combine heterogeneous high-throughput data sets produced by research with pre-existing data sets by using knowledge management, annotation and text mining tools, pathway discovery, and network inference and analysis (Picard et al., 2021). These endeavours seek to clarify the molecular mechanisms underlying physiology in order to characterise a system that combines physiological and environmental measures to enhance monitoring and detection of a phenomenon, like insect damage in plant protection research, and to make treatment and management easier. Understanding various aspects of insect like behaviour, physiology etc requires the application of tools and techniques for combining diverse omics data sets. The following standards will be used to choose the tools. The process must start with an integrative phase that examines several data sets simultaneously. Second, at least two omics data sets from samples that are somewhat comparable must be included in the methodology. Lastly, as a tool or collection of tools that can be used to any data set, the approach or methodology should be simple to use. According to their ability to handle different biological case studies, the tools and techniques are categorised in the parts that follow (Subramanian et al., 2020).

Bayesian approach: The mathematical process of applying probability to statistical issues is known as Bayesian statistics. It gives people the means to revise their views in light of fresh information. A method of data analysis that yields both a prior probability distribution and a posterior probability distribution for a parameter (such as the treatment effect) based on observed data. The posterior distribution serves as the foundation for statistical reasoning. These tools include Bayesian Consensus Clustering (BCC), iCluster, iClusterPlus, Multiple Dataset Integration (MDI), LRAcluster, Bayesian Random Effects Mixture model for joint clustering Single Cell multi-omics data (BREM-SC) and Pathway Recognition Algorithm using Data Integration on Genomic Models (PARADIGM). By employing Bayesian techniques to integrate network route analysis in them, it would be simple to study the development of pierid butterflies and their hosts as they vie for supremacy in a major arms race (Braga et al., 2021).

Network approach: Metabolomics and other omics data sets are increasingly being integrated and interpreted using network-based techniques, such as Similarity Network Fusion (SNF). Kantsa used a network-based technique to study how pollinators and plant volatiles interact (Kantsa et al., 2019).

Fusion-based approaches: The process of merging data from several sources to provide information that is more reliable, accurate, and practical than any one source alone is known as data fusion. Depending on the processing step at which fusion takes place, data fusion procedures are commonly categorised as low, middle, or high. Low-level data fusion creates new raw data by combining raw data from many sources. Fusional data, such as Pattern Fusion Analysis (PFA), is thought to be more synthetic and insightful than the original inputs.

Similarity-based approaches: Similarity-based techniques group data based on sample distances or similarities. Using these techniques, the similarities between samples in each omics are calculated independently before being integrated in various ways. The integration stage uses just similarity values. The benefit of using similarity-based approaches is that they can handle both categorical and ordinal data, as well as a variety of omics types. All that is needed for each omics, such as PINSPPlus and Neighbourhood-based multi-omics clustering (NEMO), is the formulation of a similarity measure. Muto-Fujita integrated the omics data to investigate the host plant-insect interaction using a similarity-based method (Muto-Fujita et al., 2017). They stated that specific butterfly families choose plants that are taxonomically connected to the host plant.

Other multivariate approaches: Integrative non-negative matrix factorisation (iNMF), mixOmics, moCluster, joint and individual variation explained (JIVE), multiple co-inertia analysis (MCIA), multiple factor analysis (MFA), and rMKL-LPP. Hepialid moths' responses to alien plants were examined using comparative metabolomics. MixOmics was used to analyse the metabolomic data (Atijegbe et al., 2020). Additionally, tools such as Cytoscape (DOI: 10.1093/bioinformatics/btq675) and MetaboAnalyst (DOI: 10.1093/nar/gkv380) are also widely used for the visualization and analysis of multi-omics data. Cytoscape allows researchers to construct and visualize complex molecular interaction networks, while MetaboAnalyst enables the analysis of metabolomic data through advanced statistical and machine learning methods. These tools are instrumental in bridging the gap between different omics layers, helping researchers understand how molecular changes at one level affect other levels within the biological system. Some of the studies related to multi-omics approaches in insects were listed in Table 1.

Table 1. Multi-omics approaches in insects.

SI No.	Insects	Methods	Multimomics approaches	Tools	References
1	<i>Tribolium castaneum</i>	EST, ESI-Q-TOF/MS, MALDI-TOF/MS	Genomics, Transcriptomics and Peptidomics	NCBI Database and trace archives, BLAST, ClustalW, GENBOREE	Li et al., (2008)
2	<i>Acyrtosiphon pisum</i>	EST, MALDI-TOF/MS, LC-MS/MS	Transcriptomics and Proteomics	NCBI EST library, EGAssembler, Blast2GO, PhylomeDB, InterProScan, ClustalX	Carolan et al., (2011)
3	<i>Phaedon cochleariae</i>	2-DE, RNA-seq, LC-MS/MS	Transcriptomics and proteomics	NCBI database, InterProScan, CLC Genomics Workbench, Blast2GO, KEGG, Mascot	Kirsch et al., (2012)
4	<i>Bemisia tabaci</i>	RNA-seq, qRT-PCR, nLC-MS/MS	Transcriptomics and proteomics	NCBI database, Blast2GO, KEGG, SOAPdenovo, Mascot	Yang et al., (2013)
5	<i>Myzus persicae</i> <i>Myzus cerasi</i>	RNA-seq LC-MS/MS	Transcriptomics and Proteomics	NCBI database, BLASTP, Blast2GO, MASCOT,	Thorpe et al., (2016)

	<i>Rhopalosiphum padi</i>			MUSCLE, MEGA6, Pfam, AphidBase database	
6	<i>Nilaparvata lugens</i> , <i>Laodelphax Striatellus</i> , <i>Sogatella furcifera</i>	FT-ICR-MS, qRT-PCR	Transcriptomics and Proteomics	GenBank databases, TGI clustering tool, SOAP, InterProScan software, Blast2go, COG database, Estscan, blastx, BGI WEGO, NCBI database	Huang et al., (2018)
7	<i>Aphis gossypii</i>	LC-ESI-MS/MS, RNA-seq, RACE-seq	Transcriptomics and Proteomics	UniProt/ Swiss-Prot Homo Database, Trinotate, TransDecoder, Gene Ontology analysis, and Cluster of Orthologous Group annotation	Chen et al., (2019)
8	<i>Leguminivora glycinivorella</i>	LC-MS/MS, RNA-seq	Transcriptomics and Proteomics	NCBIInr protein database, UniProt database, Protein family database, KEGG, COG, KO, Swiss-Prot, GO, Blast2GO	Yang et al., (2020)
9	<i>Spodoptera frugiperda</i>	iTRAQ analysis, LC-MS/MS	Transcriptomics and Proteomics	NCBIInr, BLAST+, Swiss-Prot protein database, COG, KEGG, Blast2GO	Cui et al., (2020)
10	<i>Pseudoregma bambucicola</i>	qRT-PCR, LC-MS/MS	Transcriptomics and Proteomics	BLAST, Swiss-Prot, non-redundant protein database, KEGG, KOG, Pfam and GO databases.	Zhang et al., (2023)
11	<i>Acyrtosiphon pisum</i>	DIGE, 2-DE, CID-MS/MS, NMR spectroscopy	Proteomics and metabolomics	NCBIInr, EST database, AphidBase, Blast2GO, InterProScan, AcypicCyc database	Wang et al., (2010)
12	<i>Spodoptera frugiperda</i>	RNA-seq, UHPLC-MS/MS	Transcriptomics and Metabolomics	GO, KEGG, COG, NR, Swiss-Prot, and Pfam databases	Yan et al., (2024)

Entomological databases

Entomological databases are essential for multi-omics analysis as they enable a wide range of researchers to access and examine data from across the globe. These databases are designed to store and arrange biological data incorporating information from scientific experiments, published literature and computer studies (Wren & Bateman, 2008; Attwood et al., 2013). Various databases, such as genome databases, protein structure databases, protein sequence databases, taxonomy databases, and proteomics, metabolomics, microarray gene expression, phylogenetics are kept and preserved (Altman, 2013). These computerised web databases aid in extracting different omics data performed by different research groups, data integration and multi-omics analyses in order to address environmental and disease-related concerns. Insect species may still be identified and their connections with similar species can be compared using these databases. Table 2 highlights some of the widely used entomological databases.

Table 2: List of entomological databases.

Database	Description	URL
Agripestbase	A community database that provides web-accessible databases of genomic information for pests	http://agripestbase.org/
AphidBase	A database for genomic resources of aphid	http://w3.rennes.inra.fr/AphidBase
BmTEdb	A database of transposable elements in the silkworm genome	http://gene.cqu.edu.cn/BmTEdb/
ButterflyNet	A database for butterfly genomes and transcriptomes, providing tools for evolutionary and ecological studies.	https://www.butterflynet.org/
ChiloDB	A genomic and transcriptomic database & it provides advanced search tools and downloadable datasets for comparative genomics and gene annotation	http://ento.njau.edu.cn/ChiloDB
DBM-DB	The genome database for diamondback moth	http://iae.fafu.edu.cn/DBM/
Dung beetle database	Comparison with other invertebrate transcriptomes	http://flylab.wits.ac.za/EI/est2uni/home.php
FlyBase	A database of genetic and genomic data on the model organism <i>Drosophila melanogaster</i> and the entire insect family <i>Drosophilidae</i>	http://flybase.org/
Hymenoptera Genome Database	A comprehensive genome informatics resource for insects of the order Hymenoptera	http://hymenopteragenome.org
i5k Initiative (i5k Workspace @ NAL)	To sequence and annotate the genomes of five thousand species of insects and allied arthropods.	i5K Consortium 2013

InsectBase	A database for insect gene resources	http://www.insect-genome.com
InsectBase 2.0	It provides web services like BLAST and Synteny Viewer for efficient searching and visualization of insect genomic data.	http://v2.insect-genome.com/
InsectBrainDatabase (IBdb)	A database for insect neuroanatomical and functional data, facilitating cross-species comparisons and data visualization	https://insectbraindb.org/
KAIKObase	The silkworm genome database	http://sgp.dna.affrc.go.jp/KAIKObase
KONAGAbase	A genomic and transcriptomic database for <i>Plutella xylostella</i> (Diamondback moth)	http://dbm.dna.affrc.go.jp/px/
MonarchBase	The monarch butterfly genome database	http://monarchbase.umassmed.edu/
STARdbi	A pipeline and database for insect monitoring based on automated image analysis	https://stardbi.cs.bgu.ac.il/home/welcome

i5k Initiative (i5k Workspace @ NAL)

The goal of the i5k Initiative is to sequence and annotate the genomes of five thousand species of insects and allied arthropods (i5K Consortium 2013). Users can access several computational tools for analysis, annotations, and insect genomes through the i5k Workspace at National Agricultural Library (NAL). Arthropod genome projects serve relatively small scientific communities, but providing further resources through a community genome portal to improves their worth for comparative genomics (Poelchau et al., 2018). The i5k Workspace@NAL has two major objectives. First, its goal is to assist i5k "data producers," specifically "orphaned" groups that lack the resources to host genomes, at the interface of sequence retrieval and analysis. The second objective of the i5k Workspace is to provide a unified framework for data consumers to retrieve genomic information from data providers, aids in disseminating and curating genome assemblies, and explores future directions for improving services for the i5k community (Poelchau et al., 2015). The most effective visualization method for comparing genomic data in the i5k Workspace is the JBrowse Genome Browser. JBrowse integrates with the BLAST+ server, allowing for visualization of sequence similarity searches and comparing genes across genomes. By providing tools for visualization, dissemination, and curation, the i5k Workspace@NAL plays a vital role in supporting genomic research across multiple disciplines. This is particularly important as the initiative aims to coordinate the sequencing of 5000 genomes, which will generate a substantial amount of data. The workspace has already accumulated genomic data for 35 arthropod species. These species are significant for various research areas, including agriculture, invasion biology, systematics, ecology, evolution, and developmental biology (Poelchau et al., 2015). An open-source pipeline for genome-scale functional annotation has been developed, allowing researchers to identify under-represented gene sets and facilitate focused community annotation efforts (Saha et al., 2021).

Flybase

Flybase is a vital resource for insect research, providing a comprehensive database of genetic information about *Drosophila*. It offers a range of data types, including mutant phenotypes, molecular characterization, cytological maps, and expression data, which aid in comparative genomics and understanding insect behavior (Grumblin G., & Strelets V., 2006). However, one of the major contributions flybase makes is to bridge *Drosophila* research with another insect research. For instance, it has enabled scientists to examine genetic relationships between *Drosophila* and other insects, including aphids, using BLAST-derived sequence data in identifying homologous genes across species. In 2018, 'flybase 2.0' was released with improved user interface and new tools, including interactive lists, reference lists, protein domain graphics, and experimental tools. FlyBase search tools often provides results in an interactive format called a HitList, which is a central hub for Flybase searches. It allows users to view, filter, export, and analyze results (Thurmond et al., 2019). Responsive tables are currently implemented for lists of related alleles, insertions, and transgenic constructs (on Gene reports, the frequently used GAL4 drivers' resource, and in results of the "GAL4 etc" QuickSearch tab) and for the members table on Pathway reports. This approach helps to accommodate larger screens and provide more relevant information for users (Gramates et al., 2022).

Hymenoptera genome database

The hymenoptera genome database (HGD; <http://hymenopteragenome.org>) is a comprehensive genome informatics resource for insects of the order Hymenoptera, including bees, ants, and wasps. It offers genome browsers, manual annotation tools (JBrowse/Apollo), BLAST, bulk data download, and a data mining warehouse called HymenopteraMine. The database also features built-in template queries, QueryBuilder tool, List Analysis and Genomic Regions search tools, and facilitates cross-species data mining based on orthology and supports meta-analyses by tracking identifiers across gene sets and genome assemblies (Elsik et al., 2018). The Hymenoptera Genome Database (HGD) has expanded its species representation to fifty-eight across various hymenopteran groups, including bees, ants, wasps, and sawflies. The study highlights a significant increase in Gene Ontology (GO) annotations, with more than a ten-fold increase in the number of GO-annotated genes. The introduction of the HGD-Ortho dataset provides a web-based resource for orthologue analysis, which is particularly beneficial for twelve of the species in HGD. This dataset allows for better comparative studies among hymenopteran taxa (Walsh et al., 2022).

Aphidbase

Aphidbase is a specialized database that analyzes and annotates aphid genetic resources. This database intends to save and make genomic data available for aphids, allowing researchers to study, compare, retrieve, and annotate sequences from multiple aphid species. The database contains a collection of unique transcripts from the pea aphid (*Acyrtosiphon pisum*), as well as gene transcripts from other insects including *Drosophila melanogaster* and *Anopheles gambiae*. This allows for comparative investigations across different species. Aphidbase includes functional annotation tools, which allow users to describe the potential functions of aphid sequences using linkages to known databases such as Flybase and Entrez (Gauthier et al., 2007). Data Architecture of Aphidbase is built on the Generic Model Organism Database (GMOD) framework and uses Chado for database architecture and GBrowse for genomic data visualization, allowing users to easily traverse and view genomic information. The database also uses text mining techniques on *Drosophila melanogaster* literature to generate a network of co-cited gene or protein names, which improves the functional annotation of aphid homolog sequences (Legeai et al., 2010). Aphidbase represents one of the first

genomic databases for a hemipteran insect. Additionally, studies have employed molecular techniques to explore the phylogenetic relationships among aphid species, revealing significant insights into their evolutionary history (Abdulla et al., 2023).

Insectbase

The availability of insect genomes and transcriptomes provides excellent resources for entomological research. Hundreds of insect genomes and transcriptomes have been sequenced, but insect community lacks an integrated, up-to-date dataset of insect genes. Insectbase, the first release of insect gene data, is available online at <http://www.insect-genome.com> (Yin et al., 2016). In the Linux operating system (Redhat 6.5), Apache HTTP server was used to construct insectbase, while PostgreSQL was used to host the database. PHP, HTML, Bootstrap, Cascading Style Sheets (CSS), JavaScript (JS) framework, and Layer JS were used in the creation of the webpages. The database was designed with a user-friendly interface through the usage of Perl scripts. PHP scripts are used by the Apache server to process queries from Web clients and conduct searches. The data is stored in Chado, a relational database structure created to manage intricate representations of biological knowledge (Guignon et al., 2012). Genome visualization was done using the Generic Model Organism Project's generic Genome Browser software (Stein, 2013). The database contains over 12 million sequences and offers search functionality, a BLAST server (Mount, 2007), GBrowse, insect pathway construction, a Facebook-like network for the insect community (iFacebook), and phylogenetic analysis of selected genes. It includes 138 insect genomes, 116 insect transcriptomes, 61 insect gene sets, 36 gene families, 7544 miRNAs, 96 925 piRNAs, 2439 lncRNAs, 22 536 pathways, 678 881 untranslated regions, and 160 905 coding sequences (Yin et al., 2016).

Challenges in multiomics approach in insects

Despite the significant potential of multi-omics in insect research, several challenges remain. One major hurdle is the lack of high-quality reference genomes for many insect species complicate data annotation and comparative analyses (Li et al., 2023). It can also be challenging to get enough material for thorough multi-omics analysis from insects because of their tiny body sizes, particularly when examining certain tissues or developmental stages. Insect-specific databases and resources are still limited for some species, hindering comprehensive studies. Another challenge is the biological variability within insect populations as well as the tissue heterogeneity with diverse cell types, which can lead to inconsistencies in results, making it difficult to draw general conclusions (Thorpe et al., 2024). Moreover, environmental factors, such as temperature or exposure to chemicals, can influence gene expression and metabolic pathways, adding another layer of complexity (Graw et al., 2021). Overcoming these challenges will require continued advancements in technology, improved data integration methods, and more extensive insect-specific genomic resources.

Future directions

Future multi-omics research is expected to involve more interdisciplinary collaboration, integrating expertise from fields such as bioinformatics, systems biology, computational biology, and clinical research. Collaboration will be essential to addressing complex biological concerns and developing comprehensive disease models. Multidisciplinary research will also be necessary to integrate multi-omics discoveries into therapeutic therapy. For instance, data scientists and doctors may collaborate to help bridge the gap between complex multi-omics data analysis and practical healthcare applications, ensuring that the insights gained are both actionable and therapeutically effective. For multi-omics research to be reliable and reproducible,

standardised procedures and quality control methods are crucial. The requirement for strong statistical techniques to manage the high dimensionality and complexity of multi-omics data is another difficulty. Accurately evaluating these data and deriving significant conclusions will need the development of more complex statistical models and algorithms. Future paths in insect multiomics research are set to investigate unconventional omics categories as lipidomics, metallomics, and epitranscriptomics in order to obtain a more comprehensive knowledge of insect biology and interactions with their surroundings (Yang et al., 2016; Machado & Gambino, 2024). Integrating these omics methods with cutting-edge techniques like single-cell sequencing, spatial transcriptomics, and high-resolution mass spectrometry can offer unparalleled insights into cellular heterogeneity and functional dynamics in insect systems (Wu et al., 2020; Sun et al., 2023). Artificial intelligence (AI) and machine learning algorithms, in conjunction with increasingly advanced bioinformatics tools, have the potential to transform the study of intricate information through the identification of patterns, prediction of functions, and modelling of biological systems. By enhancing data interpretation and hypothesis development, AI-driven methodologies, such as deep learning for image-based omics and predictive modelling of gene regulatory networks, can greatly enhance the research. These developments might reveal new facets of insect physiology, adaptability, and evolution, opening the door to creative uses in ecology, agriculture, and the prevention of vector-borne diseases.

Conclusion

Multi-omics approaches have opened new frontiers in insect research, providing a comprehensive view of the molecular interactions governing insect biology. Through data integration and the use of advanced bioinformatics tools, machine learning and specialized databases, researchers can uncover the intricate mechanisms underlying insect adaptation, evolution, and responses to environmental challenges. This holistic understanding is critical for developing targeted strategies, such as improving pest management strategies in agriculture, developing biotechnological innovations like bioengineered insects for environmental benefits, and leveraging insect models to unravel human disease pathways. Looking ahead, future directions in multiomics research will focus on exploring dynamic omics interactions under environmental pressures, expanding datasets to cover diverse and underexplored insect taxa, and enhancing computational frameworks for data integration. Achieving these goals will require more extensive studies and multidisciplinary collaborations among entomologists, ecologists, computational biologists, and biotechnologists. Such efforts will pave the way for transformative discoveries and practical applications that address global challenges in food security, ecosystem conservation, and human health.

Conflict of interest statement

The authors declare no conflict of interest.

References

- Abdulla Agha, W. N., Muhammad, H. M., & Taha, K. M. (2023). Phylogenetic analysis among some species of aphids (Homoptera: Aphididae) using DNA sequencing molecular technique. *Tikrit Journal for Agricultural Sciences*, 23(3), 71–78.
- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., ... & Venter, J. C. (2000). The genome sequence of *Drosophila melanogaster*. *Science*, 287(5461), 2185–2195. <https://doi.org/10.1126/science.287.5461.2185>

- Altman, R. B. (2004). Building successful biological databases. *Briefings in Bioinformatics*, 5(1), 4–5. <https://doi.org/10.1093/bib/5.1.4>
- Atijegbe, S. R., Mansfield, S., Ferguson, C. M., Worner, S. P., & Rostás, M. (2020). Host range expansion of an endemic insect herbivore is associated with high nitrogen and low fibre content in exotic pasture plants. *Journal of Chemical Ecology*, 46(6), 544–556. <https://doi.org/10.1007/s10886-020-01186-3>
- Attwood, T. K., Gisel, A., Eriksson, N. E., & Bongcam-Rudloff, E. (2011). Concepts, historical milestones and the central place of bioinformatics in modern biology: A European perspective. In *Bioinformatics – Trends and Methodologies* (pp. 1–27). IntechOpen.
- Braga, M. P., Janz, N., Nylin, S., Ronquist, F., & Landis, M. J. (2021). Phylogenetic reconstruction of ancestral ecological networks through time for pierid butterflies and their host plants. *Ecology Letters*, 24(10), 2134–2145. <https://doi.org/10.1111/ele.13798>
- Carolan, J. C., Caragea, D., Reardon, K. T., Mutti, N. S., Dittmer, N., Pappan, K., ... Edwards, O. R. (2011). Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrtosiphon pisum*): A dual transcriptomic/proteomic approach. *Journal of Proteome Research*, 10(4), 1505–1518. <https://doi.org/10.1021/pr100881q>
- Casas-Vila, N., Bluhm, A., Sayols, S., Dinges, N., Dejung, M., Altenhein, T., ... Butter, F. (2017). The developmental proteome of *Drosophila melanogaster*. *Genome Research*, 27(7), 1273–1285. <https://doi.org/10.1101/gr.213694.116>
- Chen, X., Xia, J., Shang, Q., Song, D., & Gao, X. (2019). UDP-glucosyltransferases potentially contribute to imidacloprid resistance in *Aphis gossypii* Glover based on transcriptomic and proteomic analyses. *Pesticide Biochemistry and Physiology*, 159, 98–106. <https://doi.org/10.1016/j.pestbp.2019.06.012>
- Cornette, R., Gusev, O., Nakahara, Y., Shimura, S., Kikawada, T., & Okuda, T. (2015). Chironomid midges (Diptera: Chironomidae) show extremely small genome sizes. *Zoological Science*, 32(3), 248–254. <https://doi.org/10.2108/zs140249>
- Cui, G., Sun, R., Veeran, S., Shu, B., Yuan, H., & Zhong, G. (2020). Combined transcriptomic and proteomic analysis of harmine on *Spodoptera frugiperda* Sf9 cells to reveal the potential resistance mechanism. *Journal of Proteomics*, 211, 103573. <https://doi.org/10.1016/j.jprot.2019.103573>
- Dawkar, V. V., Chikate, Y. R., More, T. H., Gupta, V. S., & Giri, A. P. (2016). The expression of proteins involved in digestion and detoxification are regulated in *Helicoverpa armigera* to cope with chlorpyrifos insecticide. *Insect Science*, 23(1), 68–77. <https://doi.org/10.1111/1744-7917.12175>
- Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29(1), 51–63. <https://doi.org/10.1016/j.tree.2013.09.008>
- Elsik, C. G., Tayal, A., Unni, D. R., Burns, G. W., & Hagen, D. E. (2018). Hymenoptera genome database: Using HymenopteraMine to enhance genomic studies of hymenopteran insects. In *Eukaryotic Genomic Databases: Methods and Protocols* (pp. 513–556). Springer. https://doi.org/10.1007/978-1-4939-7737-6_20

- Francis, F., Guillonnet, F., Leprince, P., De Pauw, E., Haubruge, E., Jia, L., & Goggin, F. L. (2010). Tritrophic interactions among *Macrosiphum euphorbiae* aphids, their host plants and endosymbionts: Investigation by a proteomic approach. *Journal of Insect Physiology*, 56(6), 575–585. <https://doi.org/10.1016/j.jinsphys.2009.12.006>
- Gallon, M. E., Silva-Junior, E. A., & Gobbo-Neto, L. (2024). GC–MS-based metabolomics unravels metabolites across larval development and diapause of a specialist insect. *Chemistry & Biodiversity*, 21(3), e202301779. <https://doi.org/10.1002/cbdv.202301779>
- Gauthier, J. P., Legeai, F., Zasadzinski, A., Rispe, C., & Tagu, D. (2007). AphidBase: A database for aphid genomic resources. *Bioinformatics*, 23(6), 783–784. <https://doi.org/10.1093/bioinformatics/btl671>
- Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), 333–351. <https://doi.org/10.1038/nrg.2016.49>
- Gramates, L. S., Agapite, J., Attrill, H., Calvi, B. R., Crosby, M. A., Dos Santos, G., ... Strelets, V. B. (2022). FlyBase: A guided tour of highlighted features. *Genetics*, 220(4), iyac035. <https://doi.org/10.1093/genetics/iyac035>
- Graw, S., Chappell, K., Washam, C. L., Gies, A., Bird, J., Robeson, M. S. II, & Byrum, S. D. (2021). Multi-omics data integration considerations and study design for biological systems and disease. *Molecular Omics*, 17(2), 170–185. <https://doi.org/10.1039/D0MO00176G>
- Grumblin, G., & Strelets, V. (2006). FlyBase: Anatomical data, images and queries. *Nucleic Acids Research*, 34(suppl 1), D484–D488. <https://doi.org/10.1093/nar/gkj068>
- Guignon, V., Droc, G., Alaux, M., Baurens, F. C., Garsmeur, O., Poirion, C., ... Bocs, S. (2012). Chado controller: Advanced annotation management with a community annotation system. *Bioinformatics*, 28(7), 1054–1056. <https://doi.org/10.1093/bioinformatics/bts076>
- Gutierrez Reyes, C. D., Alejo-Jacuinde, G., Perez Sanchez, B., Chavez Reyes, J., Onigbinde, S., Mogut, D., ... & Mechref, Y. (2024). Multi omics applications in biological systems. *Current Issues in Molecular Biology*, 46(6), 5777–5793. <https://doi.org/10.3390/cimb46060287>
- Hajirnis, N., & Mishra, R. K. (2021). Homeotic genes: Clustering, modularity, and diversity. *Frontiers in Cell and Developmental Biology*, 9, 718308. <https://doi.org/10.3389/fcell.2021.718308>
- Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., ... & Hoffman, S. L. (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*, 298(5591), 129–149. <https://doi.org/10.1126/science.1076181>
- Huang, H. J., Lu, J. B., Li, Q., Bao, Y. Y., & Zhang, C. X. (2018). Combined transcriptomic/proteomic analysis of salivary gland and secreted saliva in three planthopper species. *Journal of Proteomics*, 172, 25–35. <https://doi.org/10.1016/j.jprot.2017.10.020>
- Idle, J. R., & Gonzalez, F. J. (2007). Metabolomics. *Cell Metabolism*, 6(5), 348–351. <https://doi.org/10.1016/j.cmet.2007.10.005>

- Iquebal, M. A., Jaiswal, S., Mukhopadhyay, C. S., Sarkar, C., Rai, A., & Kumar, D. (2015). Applications of bioinformatics in plant and agriculture. In *PlantOmics: The omics of plant science* (pp. 755–789). Springer.
- Kamps, R., Brandão, R. D., van den Bosch, B. J., Paulussen, A. D., Xanthoulea, S., Blok, M. J., & Romano, A. (2017). Next-generation sequencing in oncology: Genetic diagnosis, risk prediction and cancer classification. *International Journal of Molecular Sciences*, 18(2), 308. <https://doi.org/10.3390/ijms18020308>
- Kantsa, A., Raguso, R. A., Lekkas, T., Kalantzi, O. I., & Petanidou, T. (2019). Floral volatiles and visitors: A meta-network of associations in a natural community. *Journal of Ecology*, 107(6), 2574–2586. <https://doi.org/10.1111/1365-2745.13205>
- Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., ... & Zhang, G. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science*, 348(6239), 1139–1143. <https://doi.org/10.1126/science.aaa4788>
- Kapli, P., Yang, Z., & Telford, M. J. (2020). Phylogenetic tree building in the genomic age. *Nature Reviews Genetics*, 21(7), 428–444. <https://doi.org/10.1038/s41576-020-0233-0>
- Kim, S., Lee, J. W., & Park, Y. S. (2020). The application of next-generation sequencing to define factors related to oral cancer and discover novel biomarkers. *Life*, 10(10), 228.
- Kirsch, R., Wielsch, N., Vogel, H., Svatoš, A., Heckel, D. G., & Pauchet, Y. (2012). Combining proteomics and transcriptome sequencing to identify active plant-cell-wall-degrading enzymes in a leaf beetle. *BMC Genomics*, 13, 1–15. <https://doi.org/10.1186/1471-2164-13-587>
- Legeai, F., Shigenobu, S., Gauthier, J. P., Colbourne, J., Rispe, C., Collin, O., ... Tagu, D. (2010). AphidBase: A centralized bioinformatic resource for annotation of the pea aphid genome. *Insect Molecular Biology*, 19(suppl 2), 5–12. <https://doi.org/10.1111/j.1365-2583.2009.00930.x>
- Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Cazzamali, G., ... Park, Y. (2008). Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. *Genome Research*, 18(1), 113–122. <https://doi.org/10.1101/gr.6714008>
- Li, H., Tahir ul Qamar, M., Yang, L., Liang, J., You, J., & Wang, L. (2023). Current progress, applications and challenges of multi-omics approaches in sesame genetic improvement. *International Journal of Molecular Sciences*, 24(4), 3105. <https://doi.org/10.3390/ijms24043105>
- Liu, G., Liu, W., Zhao, R., He, J., Dong, Z., Chen, L., ... Li, X. (2021). Genome-wide identification and gene-editing of pigment transporter genes in the swallowtail butterfly *Papilio xuthus*. *BMC Genomics*, 22(1), 1–18. <https://doi.org/10.1186/s12864-021-07462-y>
- Liu, X., & Locasale, J. W. (2017). Metabolomics: A primer. *Trends in Biochemical Sciences*, 42(4), 274–284. <https://doi.org/10.1016/j.tibs.2017.01.004>
- Luikart, G., Kardos, M., Hand, B. K., Rajora, O. P., Aitken, S. N., & Hohenlohe, P. A. (2019). Population genomics: Advancing understanding of nature. In O. P. Rajora (Ed.), *Population genomics: Concepts, approaches and applications* (pp. 3–79). Springer. https://doi.org/10.1007/13836_2018_60

- Ma, R., Rangel, J., & Grozinger, C. M. (2019). Honey bee (*Apis mellifera*) larval pheromones may regulate gene expression related to foraging task specialization. *BMC Genomics*, 20(1), 592. <https://doi.org/10.1186/s12864-019-5942-2>
- Machado, I., & Gambino, D. (2024). Metallomics: An essential tool for the study of potential antiparasitic metallodrugs. *ACS Omega*, 9(14), 15744–15752. <https://doi.org/10.1021/acsomega.3c10119>
- Marie, A., Holzmüller, P., Tchioffo, M. T., Rossignol, M., Demettre, E., Seveno, M., ... Cornélie, S. (2014). *Anopheles gambiae* salivary protein expression modulated by wild *Plasmodium falciparum* infection: Highlighting of new antigenic peptides as candidates of *An. gambiae* bites. *Parasites & Vectors*, 7, 1–13. <https://doi.org/10.1186/1756-3305-7-424>
- Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., Halder, G., ... & Reed, R. D. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences*, 109(31), 12632–12637. <https://doi.org/10.1073/pnas.1204800109>
- Milward, E. A., Shahandeh, A., Heidari, M., Johnstone, D. M., Daneshi, N., & Hondermarck, H. (2016). Transcriptomics. In *Encyclopedia of cell biology* (Vol. 4, pp. 160–165). Academic Press. <https://doi.org/10.1016/B978-0-12-394447-4.40029-5>
- Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., ... & Zhou, X. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), 763–767. <https://doi.org/10.1126/science.1257570>
- Moor, A. E., & Itzkovitz, S. (2017). Spatial transcriptomics: Paving the way for tissue-level systems biology. *Current Opinion in Biotechnology*, 46, 126–133. <https://doi.org/10.1016/j.copbio.2017.02.004>
- Mount, D. W. (2007). Using the basic local alignment search tool (BLAST). *Cold Spring Harbor Protocols*, 2007(7), pdb.top17. <https://doi.org/10.1101/pdb.top17>
- Muto-Fujita, A., Takemoto, K., Kanaya, S., Nakazato, T., Tokimatsu, T., Matsumoto, N., ... Kotera, M. (2017). Data integration aids understanding of butterfly–host plant networks. *Scientific Reports*, 7, 43368. <https://doi.org/10.1038/srep43368>
- Nayduch, D., Fryxell, R. T., & Olafson, P. U. (2019). Molecular tools used in medical and veterinary entomology. In R. Wall & D. M. Voshall (Eds.), *Medical and veterinary entomology* (2nd ed., pp. 673–694). Academic Press. <https://doi.org/10.1016/B978-0-12-814043-7.00030-2>
- Nguyen, T. T. A., Michaud, D., & Cloutier, C. (2007). Proteomic profiling of aphid *Macrosiphum euphorbiae* responses to host-plant-mediated stress induced by defoliation and water deficit. *Journal of Insect Physiology*, 53(6), 601–611. <https://doi.org/10.1016/j.jinsphys.2007.02.014>
- Palli, S. R., Bai, H., & Wigginton, J. (2012). Insect genomics. In L. I. Gilbert (Ed.), *Insect molecular biology and biochemistry* (pp. 1–29). Academic Press.

- Picard, M., Scott-Boyer, M. P., Bodein, A., Périn, O., & Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis. *Computational and Structural Biotechnology Journal*, 19, 3735–3746. <https://doi.org/10.1016/j.csbj.2021.06.039>
- Pizzorno, M. C., Field, K., Kobokovich, A. L., Martin, P. L., Gupta, R. A., Mammone, R., Rovnyak, D., & Capaldi, E. A. (2021). Transcriptomic responses of the honey bee brain to infection with deformed wing virus. *Viruses*, 13(2), 287. <https://doi.org/10.3390/v13020287>
- Poelchau, M., Childers, C., Moore, G., Tsavatapalli, V., Evans, J., Lee, C. Y., ... Hackett, K. (2015). The i5k Workspace@NAL—Enabling genomic data access, visualization and curation of arthropod genomes. *Nucleic Acids Research*, 43(D1), D714–D719. <https://doi.org/10.1093/nar/gku983>
- Poma, G., Cuykx, M., Da Silva, K. M., Iturrospe, E., van Nuijs, A. L. N., van Huis, A., & Covaci, A. (2022). Edible insects in the metabolomics era: First steps towards the implementation of entometabolomics in food systems. *Trends in Food Science & Technology*, 119, 371–377. <https://doi.org/10.1016/j.tifs.2021.12.034>
- Raj, A., Van Den Bogaard, P., Rifkin, S. A., Van Oudenaarden, A., & Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods*, 5(10), 877–879. <https://doi.org/10.1038/nmeth.1253>
- Richards, S. (2015). It's more than stamp collecting: How genome sequencing can unify biological research. *Trends in Genetics*, 31(7), 411–421. <https://doi.org/10.1016/j.tig.2015.04.007>
- Rich-Griffin, C., Stechemesser, A., Finch, J., Lucas, E., Ott, S., & Schäfer, P. (2020). Single-cell transcriptomics: A high-resolution avenue for plant functional genomics. *Trends in Plant Science*, 25(2), 186–197. <https://doi.org/10.1016/j.tplants.2019.10.008>
- Rodrigues, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., ... & Macosko, E. Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 363(6434), 1463–1467. <https://doi.org/10.1126/science.aaw1219>
- Rokas, A., & Abbot, P. (2009). Harnessing genomics for evolutionary insights. *Trends in Ecology & Evolution*, 24(4), 192–200. <https://doi.org/10.1016/j.tree.2008.11.004>
- Roychowdhury, R., Das, S. P., Gupta, A., Parihar, P., Chandrasekhar, K., Sarker, U., Kumar, A., Ramrao, D. P., & Sudhakar, C. (2023). Multi-omics pipeline and omics-integration approach to decipher plant abiotic stress tolerance responses. *Genes*, 14(6), 1281. <https://doi.org/10.3390/genes14061281>
- Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, 17, e00547. <https://doi.org/10.1016/j.gecco.2019.e00547>
- Ryabova, A., Cornette, R., Cherkasov, A., Watanabe, M., Okuda, T., Shagimardanova, E., ... Gusev, O. (2020). Combined metabolome and transcriptome analysis reveals key components of complete desiccation tolerance in an anhydrobiotic insect. *Proceedings of the National Academy of Sciences*, 117(32), 19209–19220. <https://doi.org/10.1073/pnas.2000223117>

- Si, W., Wang, Q., Li, Y., & Dong, D. (2020). Label-free quantitative proteomic analysis of insect larval and metamorphic molts. *BMC Developmental Biology*, 20(1), 1–12. <https://doi.org/10.1186/s12861-020-00215-4>
- Simola, D. F., Wissler, L., Donahue, G., Waterhouse, R. M., Helmkamp, M., Roux, J., ... & Gadau, J. (2013). Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Research*, 23(8), 1235–1247. <https://doi.org/10.1101/gr.155408.113>
- Singh, V. K., Singh, A. K., Chand, R., & Kushwaha, C. (2011). Role of bioinformatics in agriculture and sustainable development. *International Journal of Bioinformatics Research*, 3(2), 221–226.
- Song, L., Gao, Y., Li, J., & Ban, L. (2018). iTRAQ-based comparative proteomic analysis reveals molecular mechanisms underlying wing dimorphism of the pea aphid, *Acyrtosiphon pisum*. *Frontiers in Physiology*, 9, 1016. <https://doi.org/10.3389/fphys.2018.01016>
- Stein, L. D. (2013). Using GBrowse 2.0 to visualize and share next-generation sequence data. *Briefings in Bioinformatics*, 14(2), 162–171. <https://doi.org/10.1093/bib/bbs001>
- Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on Earth? *Annual Review of Entomology*, 63, 31–45. <https://doi.org/10.1146/annurev-ento-020117-043348>
- Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics data integration, interpretation, and its application. *Bioinformatics and Biology Insights*, 14, 1–24. <https://doi.org/10.1177/1177932219899051>
- Sun, C., Shao, Y., & Iqbal, J. (2023). Insect insights at the single-cell level: Technologies and applications. *Cells*, 13(1), 91. <https://doi.org/10.3390/cells13010091>
- Thorpe, P., Altmann, S., Lopez-Cobollo, R., et al. (2024). Multi-omics approaches define novel aphid effector candidates associated with virulence and avirulence phenotypes. *BMC Genomics*, 25, 1065. <https://doi.org/10.1186/s12864-024-09961-2>
- Thorpe, P., Altmann, S., Lopez-Cobollo, R., et al. (2024). Multi-omics approaches define novel aphid effector candidates associated with virulence and avirulence phenotypes. *BMC Genomics*, 25, 1065. <https://doi.org/10.1186/s12864-024-09961-2>
- Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., ... Calvi, B. R. (2019). FlyBase 2.0: The next generation. *Nucleic Acids Research*, 47(D1), D759–D765. <https://doi.org/10.1093/nar/gky1003>
- Touré, Y. T., Oduola, A. M., & Morel, C. M. (2004). The *Anopheles gambiae* genome: Next steps for malaria vector control. *Trends in Parasitology*, 20(3), 142–149.
- Walsh, A. T., Triant, D. A., Le Tourneau, J. J., Shamimuzzaman, M., & Elisk, C. G. (2022). Hymenoptera Genome Database: New genomes and annotation datasets for improved GO enrichment and orthologue analyses. *Nucleic Acids Research*, 50(D1), D1032–D1039. <https://doi.org/10.1093/nar/gkab1024>

- Wang, Y., Carolan, J. C., Hao, F., Nicholson, J. K., Wilkinson, T. L., & Douglas, A. E. (2010). Integrated metabonomic–proteomic analysis of an insect–bacterial symbiotic system. *Journal of Proteome Research*, 9(3), 1257–1267. <https://doi.org/10.1021/pr900901e>
- Westerman, M., Barton, N. H., & Hewitt, G. M. (1987). Differences in DNA content between two chromosomal races of the grasshopper *Podisma pedestris*. *Heredity*, 58(2), 221–228. <https://doi.org/10.1038/hdy.1987.34>
- Wren, J. D., & Bateman, A. (2008). Databases, data tombs and dust in the wind. *Bioinformatics*, 24(19), 2127–2128. <https://doi.org/10.1093/bioinformatics/btn392>
- Wu, J., *et al.* (2020). Spatial transcriptomics in insect development. *Nature Biotechnology*. <https://doi.org/10.1038/s41587-020-XXXX-X> (Incomplete reference—needs volume, issue, and page numbers. Please provide if available.)
- Xu, X., Li, X., Wen, D., Zhao, C., Fan, L., Wu, C., ... & Yao, Y. (2022). Sublethal and transgenerational effects of a potential plant-derived insecticide, β -asarone, on population fitness of brown planthopper, *Nilaparvata lugens*. *Entomologia Experimentalis et Applicata*, 170(7), 555–564. <https://doi.org/10.1111/eea.13201>
- Yan, X., Zhao, Z., Feng, S., Zhang, Y., Wang, Z., & Li, Z. (2024). Multi-omics analysis reveals the fall armyworm *Spodoptera frugiperda* tolerates high temperature by mediating chitin-related genes. *Insect Biochemistry and Molecular Biology*, 174, 104192. <https://doi.org/10.1016/j.ibmb.2024.104192>
- Yang, K., & Han, X. (2016). Lipidomics: Techniques, applications, and outcomes related to biomedical sciences. *Trends in Biochemical Sciences*, 41(11), 954–969. <https://doi.org/10.1016/j.tibs.2016.08.010>
- Yang, M., Wang, Z., Wang, R., Zhang, X., Li, M., Xin, J., ... Meng, F. (2020). Transcriptomic and proteomic analyses of the mechanisms of overwintering diapause in soybean pod borer (*Leguminivora glycinivorella*). *Pest Management Science*, 76(12), 4248–4257. <https://doi.org/10.1002/ps.5975>
- Yang, N., Xie, W., Yang, X., Wang, S., Wu, Q., Li, R., ... Zhang, Y. (2013). Transcriptomic and proteomic responses of sweetpotato whitefly, *Bemisia tabaci*, to thiamethoxam. *PLOS ONE*, 8(5), e61820. <https://doi.org/10.1371/journal.pone.0061820>
- Yin, C., Shen, G., Guo, D., Wang, S., Ma, X., Xiao, H., ... Li, F. (2016). InsectBase: A resource for insect genomes and transcriptomes. *Nucleic Acids Research*, 44(D1), D801–D807. <https://doi.org/10.1093/nar/gkv1204>
- Zha, W., & You, A. (2020). Comparative iTRAQ proteomic profiling of proteins associated with the adaptation of brown planthopper to moderately resistant vs. susceptible rice varieties. *PLOS ONE*, 15(9), e0238549. <https://doi.org/10.1371/journal.pone.0238549>
- Zhan, S., Merlin, C., Boore, J. L., & Reppert, S. M. (2011). The monarch butterfly genome yields insights into long-distance migration. *Cell*, 147(5), 1171–1185. <https://doi.org/10.1016/j.cell.2011.09.052>

Zhang, H. B., Cao, Z., Qiao, J. X., Zhong, Z. Q., Pan, C. C., Liu, C., ... Wang, Y. F. (2021). Metabolomics provides new insights into mechanisms of *Wolbachia*-induced paternal defects in *Drosophila melanogaster*. *PLOS Pathogens*, 17(8), e1009859. <https://doi.org/10.1371/journal.ppat.1009859>

Zhang, H., Lin, R., Liu, Q., Lu, J., Qiao, G., & Huang, X. (2023). Transcriptomic and proteomic analyses provide insights into host adaptation of a bamboo-feeding aphid. *Frontiers in Plant Science*, 13, 1098751. <https://doi.org/10.3389/fpls.2022.1098751>

Zhang, Z. J., Liu, X. J., Yu, Y., Yang, F. Y., & Li, K. (2021). The receptor tyrosine kinase torso regulates ecdysone homeostasis to control developmental timing in *Bombyx mori*. *Insect Science*, 28(6), 1582–1590. <https://doi.org/10.1111/1744-7917.12886>

Zhao, Y., Liu, W., Zhao, X., Yu, Z., Guo, H., Yang, Y., ... & Zhang, J. (2023). Lipophorin receptor is required for the accumulations of cuticular hydrocarbons and ovarian neutral lipids in *Locusta migratoria*. *International Journal of Biological Macromolecules*, 236, 123746. <https://doi.org/10.1016/j.ijbiomac.2023.123746>