EGYPTIAN ACADEMIC JOURNAL OF
BIOLOGICAL SCIENCES
ENTOMOLOGY

A

ISSN
1687-8809

WWW.EAJBS.EG.NET

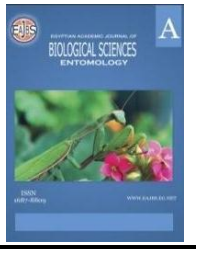
Vol. 18
No. 3 (2025)

Citation: *Egypt. Acad. J. Biolog. Sci. (A. Entomology) Vol. 18(3) pp.61-86 (2025)*
DOI: 10.21608/EAJBSA.2025.463258

Egypt. Acad. J. Biolog. Sci., 18(3):61-86 (2025)

Egyptian Academic Journal of Biological Sciences
A. Entomology
ISSN 1687- 8809

http://eajbsa.journals.ekb.eg/



From Flies to Findings: The Function of Insect Models in Predicting Protein Function and Neurodegenerative Disease

Eshraka A. Al-Ayari[1,2] *, Mona G. Shaalan[2] and Magdi G. Shehata[2]

[1]Bioinformatics Department, Egypt Center for Research and Regenerative Medicine, Cairo, Egypt.

[2]Entomology Department, Faculty of Science, Ain Shams University, Cairo, Egypt.

*E-mail: eshrakaali_p@sci.asu.edu.eg

REVIEW INFO

Article History
Received: 17/9/2025
Accepted: 1/11/2025
Available: 5/11/2025

Keywords:
Neurodegenerative disease, Insect model, Bioinformatics, computational tools, homology modelling, insect databases, Modeling disease.

ABSTRACT

The increasing prevalence of aging diseases, particularly neurodegenerative diseases, has profoundly affected the lives of diseased individuals and their families. Researchers have employed model organisms to mimic the disease and look into possible cures in order to better understand these conditions and create efficient treatments. While mammalian models exhibit similarities to humans, insects are frequently favored due to their abbreviated lifespans and reduced ethical considerations. The application of sequence analysis and comparative genomics to identify human disease genes within insect model organisms has the potential to expedite research and pharmaceutical development. The ongoing enhancement of specialized protein databases, in conjunction with advanced computational tools, has augmented the precision and dependability of protein function prediction. By reviewing various studies and articles on insect modelling, this review discusses the use of insects as model organisms for studying neurodegenerative diseases with their potential benefits, current developments in protein function prediction, and their implications for accelerating research and drug development in human diseases.

INTRODUCTION

Current projections indicate that the average human lifespan is approximately 80 years in most countries, with this figure on an upward trajectory (Vincent, 2023). By 2050, it is anticipated that the elderly population will constitute 21% of the global demographic (Ali, A. M., & Kunugi, H. 2020 b). This significant increase in the aging population correlates with a heightened prevalence of neurodegenerative diseases. Recognized neurodegenerative disorders include Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD) (Checkoway *et al.,* 2011). These conditions result from a progressive decline in brain function due to the gradual neuronal death, adversely impacting the quality of life for patients and their families both socially and economically (Batista, P., & Pereira, A. 2016). A comprehensive understanding of the molecular and cellular mechanisms underlying neurodegeneration is essential for identifying novel therapeutic targets and developing effective strategies to combat these diseases. One of the most promising outcomes of genome sequencing projects is the enhanced understanding of human

diseases and the development of new therapies. The genomes of model organisms have been sequenced alongside human genomes, providing valuable tools for investigating human gene function. Many genes in insect model organisms share common ancestry and function with human genes, particularly orthologous genes. The whole genome sequencing of *Drosophila melanogaster* (fruit fly) marked a significant milestone (Adams *et al.,* 2000). Currently, hundreds of insect genomes from various orders have been identified, with ongoing advancements in genome assemblies and annotations (Li *et al.,* 2019). Progress in bioinformatics has facilitated the management of vast data sets, playing a crucial role in genome assembly and annotation. The availability of multiple insect genomes offers exceptional potential for comparative genomics among insects and between insects and humans, leading to a wealth of information on organismal biology (Roy, S. 2013).

Relevant studies were identified through systematic searches conducted between January 2010 and August 2025 using PubMed and Google Scholar. The search combined terms related to neurodegenerative diseases ("neurodegenerative disease" OR "Alzheimer's disease" OR "Parkinson's disease" OR "Huntington's disease") with insect model keywords ("*Drosophila melanogaster*" OR "fruit fly" OR "*Bombyx mori*" OR "silkworm" OR "*Apis mellifera*" OR "honeybee" OR "*Tribolium castaneum*" OR "red flour beetle"). Boolean operators (AND/OR) were used to combine terms. Studies were included if they (1) involved experimental or computational work using insect models to investigate molecular, genetic, or phenotypic aspects of neurodegenerative diseases, or (2) evaluated relevant analytical tools or genetic manipulations. Exclusion criteria were review papers, non-English articles, conference abstracts, and studies not directly addressing ND mechanisms.

Screening was performed in two stages: title/abstract screening followed by full-text review. Reference lists of included articles were also manually searched to capture additional relevant studies. Data extraction emphasized study design, target genes, experimental assays, and quantitative outcomes when available.

Neurodegenerative Diseases:

Neurodegenerative diseases (NDs) are caused by a progressive decline in brain function as a result of gradual neuronal death resulting from protein misfolding as tau protein, which has a neurotoxic effect on Alzheimer's disease. When these toxic proteins exceed the cell capacity to dispose of them or escape cell quality control, they cause diseases (Pandey, U. B., & Nichols, C. D. 2011). They are incurable, cause problems with movement and mental functioning, affect a person's ability to move, speak, and breathe, and may lead to death. Current treatments for neurodegenerative diseases are largely symptomatic and provide only limited relief, failing to address the underlying disease mechanisms or halt the progressive neuronal loss (Tello, 2022). NDs severely affect the social and economic quality of life of patients and their families (Chekani, F., *et al.,* 2016).

Some NDs are caused by inherited gene mutations from one or both parents, such as HD and AD, but most result from a combination of genetic and environmental factors, such as PD. The research approach of NDs is divided into two main areas: With no treatment, one group of scientists is dedicated to discovering a cure or at least developing or repurposing drugs that can slow down symptoms or halt the progression of the disease, and research is being conducted to better understand the prevention and treatment of these diseases. Meanwhile, another group is focused on creating tools to identify the earliest and most subtle indicators of the disease as different mutations on genetic or proteomic levels (Al-Ayari, E. A., *et al.,* 2023).

Alzheimer's Disease:

Alzheimer's disease became the seventh leading cause of death worldwide after COVID-19 entered the ranks of the top ten causes of death, as mentioned by the World Health Organization (WHO), due to pneumonia or other immobility complications (Better,

M. A., 2023). Therefore, it is considered the most frightful and debilitating neurological disorder. Its main key features are the abnormal accumulation of beta-amyloid protein forming plaques in the spaces surrounding synapses and microtubule-associated protein (tau) in the cell bodies of neurons (Ali, A. M., & Kunugi, H. 2020 b; Alvarez-Mora *et al.,* 2022; el Tallawy *et al.,* 2019). Reductions occur in the indicators of several neurotransmitters, including acetylcholine and glutamate, which allow cells to communicate with one another. Clinical symptoms result from damage to these brain networks, which are essential for learning, attention, memory, and higher cognitive capacities (Ali, A. M., & Kunugi, H. 2020 b). Acetylcholine aids in learning and short-term memory. Because plaques raise acetylcholinesterase activity, which breaks down ACh, they lower ACh levels in AD patients. Glutamate aids with memory and learning. Glutamate is released in excess by dying brain cells in AD patients, which overstimulates healthy brain cells and damages or kills them.

According to Alvarez-Mora *et al.* (2022), there are two types of the disease: early onset familial AD (EOAD), which affects individuals between the ages of 30 and 65 and is linked to three autosomal dominant inheritance genes: amyloid precursor protein (APP), presenilin 1, and two genes (PSEN1 and PSEN2). Late-onset AD (LOAD), which develops after the age of 60 and is linked to a single gene, the Apolipoprotein E (APOE) gene, is the second most prevalent sporadic form of AD. Interestingly, APOE ɐ3 has a more neutral effect on risk than the ɐ4 variant. This gene includes apolipoprotein E, which transports cholesterol in the central nervous system (CNS) and influences amyloid β aggregation and clearance in the brain (Alvarez-Mora *et al.,* 2022; El-Metwally *et al.,* 2019; Kloske *et al.,* 2021). Furthermore, a cell surface receptor on microglia (immune cells in the brain) is encoded by the triggering receptor expressed on the myeloid cell 2 gene (TREM2), which also plays a role in the microglial response to amyloid plaques that characterize AD. Microglia respond less to plaques when TREM2 function is lost, and they then seem to be in a more hazardous state. The distribution of neurofibrillary tangles and plaque density is observed to influence the occurrence or susceptibility to AD in neuropathological diagnosis. While tangles originating from mutations in the microtubule-associated protein tau (MAPT) gene have not been genetically linked to AD, Aβ is the main component of plaques, and the component most closely associated with AD is Aβ 42, which is derived from the APP gene. The following link illustrates a continuous update of AD genes, mutations, and risk factors: https://www.alzforum.org/ (Kloske *et al.,* 2021).

Parkinson's Disease:

Nearly 1.5 million Americans and over 6.3 million people worldwide suffer from Parkinson's disease (Ali, A. M., & Kunugi, H. 2020 a). It accounts for approximately 2% of the world's population. In addition to impairments in motor movements, such as resting tremors, rigidity, bradykinesia, and postural instability. Parkinson's disease can also lead to changes in non-motor brain functions, including mood disorders, rapid eye movement sleep behavior disorder (RBD), hyposmia, and constipation (Guadagnolo *et al.,* 2021). This is due to the death of dopamine-producing cells in the midbrain, a part of the brain known as the substantia nigra pars compacta. A considerable number (~40%) of neurons must be lost before symptoms occur. The Hoehn and Yahr scale is widely used to assess PD patients according to five disease stages, which range from only unilateral involvement in stage I to wheelchair or bed-bound involvement in stage V. Evidence indicates that the loss of dopaminergic neurons in the substantia nigra pars compacta in patients with PD occurs primarily in stage III (Ali, A. M., & Kunugi, H. 2020 a; Carrarini *et al.,* 2019).

Similar to AD, PD is a proteinopathy in which the toxic aggregation of characteristic proteins in specific areas of the brain is the main diagnostic hallmark. α-Synuclein is a major contributor to neuronal death in PD, which can be found in enteric

nerves and enteroendocrine cells of PD patients (Ali, A. M., & Kunugi, H. 2020 a). In general, there is a complex combination of multiple genetic risks and protective factors, including smoking, nicotine, caffeine, nonsteroidal anti-inflammatory drugs, omega-3 fatty acids, and hormone replacement therapy, with environmental factors developing the disease at a given point in life. Apart from the PD genes, PARK2 (E3 ubiquitin ligase complex) and PINK1 (phosphatase and tensin homolog-induced kinase 1) are specific to it, whereas all other genes, LRRK2 (leucine-rich repeat kinase 2), PARK7 (Parkinsonism Associated Deglycase), and SNCA (alpha-synuclein), are associated with more than one clinical diagnosis of AD, where PD has an infrequent monogenic component (Cruts *et al.,* 2012). Knowledge of monogenic mutations that lead to neurodegenerative diseases is important. Clinical genetic counseling will help to develop efficient diagnostic screening protocols to support or specify a clinical diagnosis.

Huntington's Disease:

Huntington's disease is among the most prevalent inherited brain disorders in the world. It affects approximately 30,000 Americans and places 200,000 people at risk (Brown, D. G., & Wobst, H. J. 2023). It exhibits important characteristics of aberrant protein aggregation in the brain cortex, which is the hub for cognition, perception, and memory, as well as the basal ganglia, which regulate coordination. Death frequently results from heart failure, pneumonia, or other complications. Cognitive decline and movement disorders, such as chorea and loss of coordination, are hallmarks of Huntington's disease, an autosomal dominant disorder. The motor defects are caused by an expanded triplet repeat CAG in the HTT gene that codes for an abnormal version of the huntingtin protein with a toxic function (El-Jaafary *et al.,* 2020; McColgan, P., & Tabrizi, S. J. 2018). Huntington's disease symptoms can start as early as age 2 or as late as age 80, but they typically manifest between the ages of 30 and 50. Psychiatric, cognitive, and motor abnormalities are symptoms of Huntington's disease. Unintentional weight loss, sleep and circadian rhythm abnormalities, and autonomic nervous system dysfunction are some of the other less well-known but common and frequently incapacitating symptoms of HD (McColgan, P., & Tabrizi, S. J. 2018). It is frequently hard to diagnose, which makes it hard to manage, particularly in nations where a large proportion of marriages are between relatives (El-Jaafary *et al.,* 2020).

Insects as Model Organisms:

By increasing the human lifespan, aging-related research has also increased, using different model organisms to mimic human diseases. Despite the inability to mimic all the disease symptoms under study, they are phylogenetically distant from humans. In contrast to vertebrate models, which have numerous drawbacks, including long generation times that range from 8 to 10 years for the grey mouse lemur, a commonly used non-human primate model of aging (Lacreuse, A., & Herndon, J. G., 2008), low fecundity and few offspring, high housing costs, and ethical permissions, insects are favored due to their global distribution, meaning that the number of species of insects is greater than any other group. In the world, some 900 thousand different kinds of living insects are known. This represents approximately 80 percent of the world's species (May RM, 1988), comparatively low rearing costs, and high fecundity that allows high-throughput screening. For insects, especially fruit flies and other social insects that are fully identified and have available mutant stocks; their small size means small breeding sites, short generation time with high fecundity, short lifespan investigations, homology between insects and humans, Table 1 and the presence of thousands of mutants for human disease-causing genes, as in fruit flies, make them good models for neurodegenerative disease research (Haleem H. F. & Grace B. D., 2019).

Table 1: Homologous neurodegenerative disease genes between insects and humans (Adamski *et al.,* 2019; McGurk *et al.,* 2015; Song *et al.,* 2017; Tabunoki *et al.,* 2013; Tabunoki *et al.,* 2016), most of which are found in *Drosophila melanogaster*, and a little in *Tribolium castaneum* and *Bombyx mori*. Therefore, further studies on other insect models are needed to investigate other pathways to find innovative cures.

Disease/Gene	Insect Model	References
Alzheimer's disease		McGurk *et al.* (2015)
APP	*Drosophila melanogaster*	
A-beta peptide	*Drosophila melanogaster*	
PSEN 1 and 2	*Drosophila melanogaster*	
MAPT (Tau)	*Drosophila melanogaster*	
PSEN 1 and 2	*Tribolium castaneum*	Al-Ayari, E. A., *et al.* (2023)
PSEN 1 and 2	*Bombyx mori*	
APP	*Apis mellifera*	
Huntington's disease		McGurk *et al.* (2015)
PolyQ domains	*Drosophila melanogaster*	
SCA1	*Drosophila melanogaster*	
SCA2/(ALS)	*Drosophila melanogaster*	
SCA3/MJD	*Drosophila melanogaster*	
SCA7	*Drosophila melanogaster*	
SCA8	*Drosophila melanogaster*	
SCA17	*Drosophila melanogaster*	
VCPl	*Tribolium castaneum*	Al-Ayari, E. A., *et al.* (2023)
HTT	*Apis mellifera*	
UBQLN2	*Apis mellifera*	
DMBK	*Apis mellifera*	
VPS35	*Bombyx mori*	
UCHL1	*Bombyx mori*	
Parkinson's disease		McGurk et *al*. (2015)
LRRK2	*Drosophila melanogaster*	
Parkin (loss of function)	*Drosophila melanogaster*	
Pink1 (loss of function)	*Drosophila melanogaster*	
SNCA (α-synuclein)	*Drosophila melanogaster*	
AUX	*Drosophila melanogaster*	Song *et al.* (2017)
TcPINK1	*Tribolium castaneum*	Adamski *et al.* (2019)
PARK7/DJ-1	*Bombyx mori*	Tabunoki *et al.* (2013)
GAK	*Tribolium castaneum*	Al-Ayari, E. A., *et al.* (2023)
HTRA2	*Tribolium castaneum*	
LRRK2	*Tribolium castaneum*	
EIF4G1	*Tribolium castaneum*	

Over the years, the number of sequenced insect genomes has greatly increased, starting from *Drosophila melanogaster, Apis mellifera, Tribolium castaneum, Galleria mellonella, Bombyx mori, Periplaneta americana, Anopheles gambiae,* and *Locusta migratoria* Table 2. Also, the provided tools and databases that contain different types of insect data increase to cope with the current situation of continuous work on these insect models as shown in Table 3. All of that enables us to further speed progress toward finding a cure for NDs such as AD, HD, and PD (Adamski *et al.,* 2019).

Table 2: Updated total number of genes and proteins in four different insect orders and their insect sequence lengths in May 2025.

Order	Insect		Chromosome no.	No. of genes	No. of protein-coding	GC content	Genome size (Mb)	Release date	Link
	Common name	Scientific name							
Diptera	Fruit fly	*Drosophila Melanogaster*	7	17872	13962	42	143.7	8/1/2014	https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_000001215.4/
	Mosquito	*Anopheles gambiae*	3	15165	12519	44.5	264.5	9/9/2022	https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_943734735.2/
Lepidoptera	Silkworm	*Bombyx Mori*	29	18210	13459	38.5	461.7	6/7/2023	https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_030269925.1/
Coleoptera	Flour beetle	*Tribolium castaneum*	11	15518	12172	31.5	241.8	14/9/2023	https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_031307605.1/
Hymenoptera	Honeybee	*Apis Mellifera*	16	12398	9935	32.5	225.2	9/13/2018	https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_003254395.2/

Based on nearly five decades of basic research in a few model systems, most animals are quite similar, especially at the cellular and molecular level (Batista, P., & Pereira, A. 2016). According to Ugur *et al.* (2016), decision-making centers in the brains of mammals and insects evolved separately but have many similarities. These findings could aid in the understanding of many neurodegenerative diseases. The central complex in arthropods and the basal ganglia in vertebrates are very similar; the central complex mediates the selection and maintenance of behavioral actions. Thus, severe mental health issues ranging from autism to neurodegenerative diseases, dementia, and attention deficits can result from basal ganglia dysfunction. Also, arthropod mushroom bodies and vertebrate hippocampuses serve comparable purposes. Insects' malfunctioning brains provide insights into human brain disorders (Strausfeld, N. J., & Hirth, F. 2013). Refer to Figure 3.

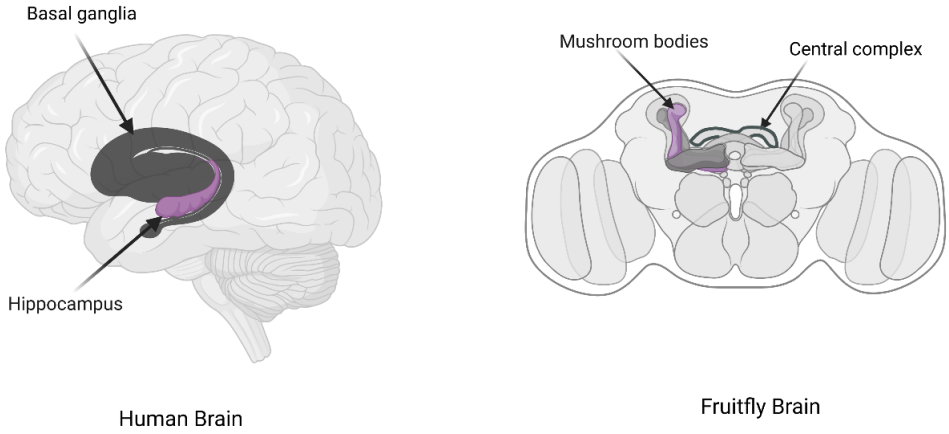


Fig. 3: Structural and functional similarities between human and fruit fly brains. The central complex in fruit flies parallels the basal ganglia in humans in dark grey, and the mushroom bodies in fruit flies parallel the hippocampus in humans in pink color. Created with BioRender.com

Table 3: Insects databases and tools accessed in May 2025.

Name	**Description URL address**	**Notes**	References
ANOBASE: Anopheles gambiae	http://www.anobase.org/	A database contains genomic and biological data about anopheline mosquitoes	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
Aphidbase	http://www.aphidbase.com/aphidbase/	AphidBase is a reference information system providing genomic resources for the study of aphids. And a bioinformatics platform for agrosystem arthropods.	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
BDGP: Berkeley Drosophila Genome Project	http://www.fruitfly.org/	This project aims to present up-to-date information about the annotated sequence to the research community and develop informatics tools that support the experimental process.	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
Butterflies and Moths of North America	http://www.butterfliesandmoths.org	It shares species information and occurrence data of butterflies, moths, and caterpillars.	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
CAPS	http://pest.ceris.purdue.edu/	It curates many details on invasive insects and diseases. Last update: 2021	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
Caterpillars: Australian region	http://lepidoptera.butterflyhouse.com.au/	It describes biology, behaviour, distribution, life histories, and images of Australian Lepidoptera species	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
Drosophila Immune Response	http://www.fruitfly.org/expression/immunity/	Preview only	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
Drosophila Interaction Database (CuraGen)	http://www.droidb.org/Index.jsp	Assembled gene and protein interaction data from a variety of sources. last updated 29 August 2018	(Pandey, U. B., & Nichols, C. D. 2011)
Drosophila Species Genomes	http://insects.eugenes.org/DroSpeGe/	A comparative database of *Drosophila* species. 2019-April: Now archived at Jetstream-cloud.org	(Pandey, U. B., & Nichols, C. D. 2011)

East Asian Distribution Center	https://shigen.nig.ac.jp/fly/nigfly/	It provides antisera to segmentation gene proteins	(Pandey, U. B., & Nichols, C. D. 2011)
FlyExpress	http://flyexpress.net/	An online platform for biological discovery. It contains a digital library of standardised images from BDGP, Fly-FISH, and peer-reviewed publications capturing the spatial expression of thousands of genes at different developmental stages in *Drosophila melanogaster*	(Chilana, P., Sharma, A., & Rai, A. 2012)
FlyMine	http://www.flymine.org	An integrated database for *Drosophila melanogaster* and *Anopheles spp.* Genomics	(Pandey, U. B., & Nichols, C. D. 2011)
Gene Disruption Project P-Screen Database	http://flypush.imgen.bcm.tmc.edu/pscreen	A searchable database of gene disruption strains	(Pandey, U. B., & Nichols, C. D. 2011)
Hymenoptera Genome Database	http://hymenopteragenome.org/	A genome informatics resource that supports the research of insects of the order Hymenoptera and provides tools for data mining (HymenopteraMine), sequence searching (BLAST), genome browsing (JBrowse), genome annotation (Apollo), and data download.	(Chilana, P., Sharma, A., & Rai, A. 2012)
i5k Workspace@NAL	https://i5k.nal.usda.gov/	Workspace strives to facilitate public data access, visualization, and community curation across arthropod species	(Poelchau MF, et al., 2018)
iBeetle	http://ibeetle-base.uni-goettingen.de/	A database dedicated to sequences, phenotype information, RNAi, and homologs of *Tribolium castaneum*	(Hakeemi et al., 2022; Schmitt-Engel et al., 2015)
Insect protein Domain Analysis	http://bioinf.ibun.unal.edu.co/insecta/	Preview only	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
KSU Tribolium Genetics Program	http://www.ksu.edu/tribolium/	Preview only	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
modENCODE	http://www.modencode.org	The tool identifies all of the sequence-based functional elements in the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
Pherobase - database	http://www.pherobase.com/	Database of pheromones and semiochemicals	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)
REDfly	http://redfly.ccr.buffalo.edu/	A curated collection of known insect transcriptional cis-regulatory modules (CRMs), cis-regulatory module segments (CRMsegs), predicted cis-regulatory modules (pCRMs), and transcription factor binding sites (TFBSs).	(Chilana, P., Sharma, A., & Rai, A. 2012)
The California beetle databases	http://www.sbcollections.org/cbp/cbp_db_informatics.aspx	A database contains nearly 150,000 records of beetle species in California, including collection and literature records	(Habeeb, S.K.M, & Raman Chandrasekear. 2014)

Thermofisher Arrays	https://www.thermofisher.com/eg/en/home/life-science/microarray-analysis.html	It provides innovative products, tools, and resources that help advance the work of researchers via microarray analysis	(Pandey, U. B., & Nichols, C. D. 2011)
VectorBase	https://vectorbase.org/vectorbase/app/	It contains genome information for three mosquito species: *Aedes aegypti, Anopheles gambiae, and Culex quinquefasciatus*	(Chilana, P., Sharma, A., & Rai, A. 2012)

Fruit Fly (*Drosophila melanogaster):*

Drosophila melanogaster, a dipteran insect, was first used as a genetic model in 1900 by Woodworth (Munkácsy, E., & Pickering, A. M. 2021). *Drosophila melanogaster* has a short life cycle, and it provides fast and simple gene manipulation because of its smaller genome (1.2×10^8 bp) compared to humans (3.3×10^9 bp), probably a million-fold fewer neurons in flies, a small number of genes, and a complex nervous system that shares similarities with vertebrate brains (McGurk *et al.,* 2015). According to Pandey and Nichols (2011), this fly has functional homologs for nearly 75% of human disease-causing genes. Additionally, the fly has highly conserved disease pathways that help us better understand neurological diseases. By using it at various stages as a model that exhibits a notable response to numerous medications that act on the mammalian central nervous system, it offers a new tool for drug discovery to overcome existing limitations. nervous system (CNS) instead of differences in the fly brain (Ugur *et al.,* 2016). Interestingly, as shown in Table 4, fruit flies can express or model different phenotypes of induced AD, PD, or HD (Munkácsy, E., & Pickering, A. M. 2021; Pandey, U. B., & Nichols, C. D. 2011; Ugur *et al.,* 2016). For Alzheimer's, it is one of the best models according to the average protein identity percentage of disease proteins (Al-Ayari, E. A., *et al.,* 2023). Drosophila melanogaster and other models can be used in three ways to investigate human diseases: First, reverse genetics is used to study the phenotypes of human genes in flies by causing mutations in these fly homologs. According to second-forward genetics, the animals are screened for a specific phenotype, and mutations are introduced at random. Mutations can be generated by radiation, chemicals [e.g., alcohol], transposons, knockdown techniques, or even inbreeding. The third diagnostic strategy includes initiatives designed to identify human disease-causing genes by sequencing the whole exome or genome of patients and their parents to determine the causative gene variant in which three or fewer individuals are assessed (Song *et al.,* 2017; Ugur *et al.,* 2016).

The reduced complexity and ease of studying the nervous system of *D. melanogaster* allows for in-depth evaluation of the functions of genes and neural networks. Therefore, *D. melanogaster* has been used more and more over the past 20 years to model neurological dysfunction, such as neurodegeneration, epilepsy, dementia, and other brain diseases, enabling researchers to create novel treatment approaches (Ugur *et al.,* 2016). It also allows the identification of genetic modifiers (suppressors or enhancers) that affect degeneration caused by polyQ; for example, (dHDJ, dtpr2, MLF1, QBP1) "homologs to human" suppress the pathogenesis of polyQ, including AD. Geldanamycin, an "inhibitor of Hsp90," suppresses PD. The SAHA "HDAC inhibitor," cystamine, and Congo red "aggregation inhibitor" reduced the pathology of HD (Batista, P., & Pereira, A. 2016). Also, Expression of human α-synuclein (A53T mutant) in *Drosophila* dopaminergic neurons leads to a ~60% reduction in climbing ability by day 20 post-eclosion compared to controls ($p < 0.001$) (Pesah Y, *et al.,* 2005).

Table 4: Phenotypic characteristics of *D. melanogaster* as a model for human AD, PD, HD (Pandey, U. B., & Nichols, C. D. 2011).

Diseases/Gene	Animal	Phenotypes
Alzheimer's Disease		
Amyloid protein	*D. melanogaster*	Eye degeneration, accumulation of amyloid plaques, reduced life span, locomotor defect, and vacuolation of the brain. (Luheshi LM, *et al.,* 2007)
Presenilin	*D. melanogaster*	Pupal lethality, dorsoscutellar bristle duplications, wing notching, and wing vein defects. (Seidner GA, *et al.,* 2006)
Tau	*D. melanogaster*	Eye degeneration, disruption of the microtubular network at presynaptic nerve terminals, axonal degeneration, neuromuscular junctions, and morphological defects. (Chen X, *et al.,* 2007)
Parkinson's Disease		
α-Synuclein	*D. melanogaster*	Age-dependent loss of dopaminergic neurons and progressive climbing defect. (Pesah Y, *et al.,* 2005)
Parkin and Pink	*D. melanogaster*	Dopaminergic neuron loss, age-dependent motor deficits, reduced lifespan, locomotor defects, male sterility, and mitochondrial pathology. (Xi Y, *et al.,* 2010)
Huntington's disease		
Triplet repeat expansion	*D. melanogaster*	Axonal transport defect, lethality, neurodegeneration, and behavioral and electrophysiological defects. (Romero E, *et al.,* 2008)

Red Flour Beetle *(Tribolium castaneum):*

Beetles are the largest insect order, with *Tribolium castaneum* being the first species to have a completely sequenced genome in 2008. They have survived major disasters and many mass extinctions over 300 million years of history and still represent 25% of all known animals. One-third of the genome consists of repetitive sequences. The unassembled regions in *T. castaneum* are thought to be heterochromatin with highly repetitive sequences. Although *T. castaneum* has more conserved genes than other insects, it also has more non-conserved genes. Because *T. castaneum* has more olfactory receptor ORs and detoxification genes than *D. melanogaster* and other insects, it may be better adapted to its environment (Li *et al.,* 2019). Instead of being economically important, it is a stored-grain beetle. *T. castaneum* exhibits some developmental mechanisms that are more representative and similar to those found in mammals. As a result, *T. castaneum* is among the best genetic models for post-genomic research, including functional genomics and proteomics, which refer to the ability of genetic crosses to effectively screen for and knock down particular gene products in any tissue. It is believed that early-onset hereditary Parkinson's disease is caused by missense mutations in the human protein kinase (hPINK1) gene. Despite several

structural variations from hPINK1, the *T. castaneum* beetle's TcPINK1 protein shows catalytic activity toward ubiquitin, Parkin, and generic substrates and contributes to additional research on Parkinson's disease in humans (Adamski *et al.,* 2019). Furthermore, it provides phenotypic characteristics that represent PD (Brandt, A. *et al.,* 2019). It has been shown that motor defects of *Drosophila* PINK1 null flies, similar to those that occur in Parkinson's disease, can be rescued in vivo by crossing lines that overexpress TcPINK1 (Woodroof *et al.,* 2011). According to Al-Ayari, E. A., *et al.* (2023), *T. castaneum* is one of the best models for Huntington's disease based on protein-level homology with humans, as it represents 13 SNPs of 37 suggested mutations in the VCPl protein.

Honeybee *(Apis mellifera):*

Sociobiological theory proposes that similarities between human and animal societies reflect similar evolutionary roots, that is, the systematic study of how natural selection shapes the biological basis of all social behaviors. Therefore, comparative genomics is used to test whether similarities in behavior between humans and honeybees reflect shared mechanisms (Shpigler *et al.,* 2017).

Apis mellifera, as a representative of the order Hymenoptera, is a key model of social behavior. Social behavior is represented by multiple generations living in the same nest at the same time, the cooperation of some members in caring for offspring that are not their own, and their specialization according to specialized tasks in the beehive (Honeybee Genome Sequencing Consortium, 2006; Shpigler *et al.,* 2017).

In terms of circadian rhythm, deoxyribonucleic acid (DNA) methylation, and ribonucleic acid (RNA) interference, honeybees are more similar to vertebrates. Despite having a brain with only one million neurons—five times fewer than the human brain and only four times larger than that of *D. melanogaster*, which has a much simpler behavioral repertoire it exhibits sophisticated cognitive abilities. Worker bees learn to associate a flower's color, shape, scent, or location with a food reward, which increases foraging efficiency. They communicate new food discoveries with 'dance language,' the only non-primate symbolic language. Recent studies have revealed that *A. mellifera* can learn concepts such as 'same' and 'different' (Honeybee Genome Sequencing Consortium, 2006).

This enhances the potential to serve as a model for Alzheimer's disease, in conjunction with the capacity to detect novel genome-derived single-nucleotide polymorphisms (SNPs). Interestingly, bees offer different life spans for workers, from 20 to 40 days in summer and about 140 days in winter, while the queen can live for two years, which refers to vitellogenin expression. Furthermore, it was found that reducing feeding on nectar and pollen for workers to an hour increases their lifespan, which in turn could explain the mechanism of life expansion (Haleem H. F. & Grace B. D., 2019). A Parkinson's disease (PD) model has also been established in honeybees *Apis mellifera* through exposure to rotenone, a mitochondrial complex I inhibitor commonly used to induce PD-like symptoms, as well as through colonization with gut microbiota derived from PD patients. In this model, bees exhibited behavioral impairments reminiscent of PD-related motor and cognitive dysfunctions. The study demonstrated a strong association between altered gut microbiota composition and neurological deficits, suggesting a conserved role of the gut–microbiota–brain axis in the pathogenesis of Parkinson's disease (Zeng J, *et al.,* 2024).

Silkworm (*Bombyx mori):*

Furthermore, being the most important industrial insect for over 5000 years, *Bombyx mori* has played a significant role in disease modeling and medical research. Silkworms are 4–5 cm in length with a flightless nature; easy lab rearing in large numbers with simple facilities makes them easy to handle, and they can undergo experiments similar to those conducted on mammals, such as oral administration and intravenous injection. A draft silkworm genome sequence was completed by Chinese and Japanese groups in 2008 (The

International Silkworm Genome Consortium, 2008). The silkworm genome contains 16,823 gene loci, based on sequence analysis of the cDNA dataset, and more than 400 visible mutant silkworms are available (Suetsugu *et al.,* 2013), and for now a total of 18,397 proteins were predicted using over 95 Gb of mRNA-seq derived from 10 different organs, covering 96.9% of the complete orthologs of the lepidopteran core genes (Waizumi *et al.,* 2023). To annotate silkworm genome information. Human homologs common to *B. mori* and *D. melanogaster* were identified using cDNA sequence sets from Ensembl Metazoa and a systematic BLAST search. *B. mori* contains 8,469 homologs, and *D. melanogaster* contains 8,815 human homologs. A recent study found that approximately 5006 genes of silkworm moths are orthologs of human disease-related genes, mainly skeletal, neurological, and growth diseases (Haleem H. F. & Grace B. D., 2019; Tabunoki *et al.,* 2016).

Thus, *B. mori* shares 58% homology with human genes. These genes are the most conserved in pathways involved in neurodegenerative diseases, oxidative stress, and protein degradation. *B. mori* is a good model for age-related and neurodegenerative diseases such as PD (Tabunoki *et al.,* 2013). While not a model for the full human disease, silkworms have been used to evaluate the potential neuroprotective effects of compounds that could help prevent or treat AD (Baek *et al.,* 2020). Preventive Nutrition and Food Science, 25(4), 389-399. 10.3746/pnf.2020.25.4.389 Translucent silkworm larvae exhibit a significant downregulation of the PARK7 (DJ-1) gene, a key component of the antioxidant defense system. This reduction in DJ-1 expression leads to elevated oxidative stress and neuronal dysfunction, thereby recapitulating a major pathological hallmark of human PD (Zhang H, *et al.,* 2024). It may accumulate uric acid as urate granules in their integument, which has a protective role against oxidative stress, as in humans against PD, which is unique to the *B. mori* model (Tabunoki *et al.,* 2016; Zhu *et al.,* 2022). Experiments have shown that chronic feeding of silkworms with 6-hydroxydopamine (6-OHDA) results in similar PD phenotypes, including a decrease in dopamine levels, degeneration of dopaminergic neurons, and motor dysfunction, despite the fact that silkworms are phylogenetically distant from humans (Zhu et al., 2022). Consequently, the p-translucent silkworm mutant has been proposed as a promising invertebrate model for investigating the molecular mechanisms underlying Parkinson's disease. According to Al-Ayari, E. A., *et al.* (2023), for HDs, *B. mori* has higher identical VPS35 and UCHL1 to *Homo sapiens* than *D. melanogaster.*

Computational Analysis and Databases:

History:

Huntington's disease was linked to the huntingtin (HTT) gene on chromosome 4 in 1983, according to a study done on the DNA of families affected by the condition. The first hereditary illness to be mapped using polymorphism data (G8 DNA probe/genetic marker) was Huntington's disease. Remarkably, it took an additional decade to identify the HTT sequence and pinpoint the specific type of mutation linked to Huntington's disease (Bromberg, Y. 2013).

Over the past 20 years, comparative biology and genetics have evolved into comparative genomics. Comparative genomics is a relatively new field of biological research that compares genome sequences from different species and helps determine physiological mechanisms related to humans by identifying and studying common processes that occur both in humans and model organisms. Nevertheless, comparative genomics began nearly 200 years ago when animal models were first proposed to mimic humans (Moreno *et al.,* 2008).

Scientists have found that it is better to combine the data collected by Online Mendelian Inheritance in Man (OMIM), LocusLink, and FlyBase in Homophila (human disease gene cognates in *Drosophila*), which enables merging genes of diseases and their related genes to determine the function of related genes in model organisms (Chien *et al.,*

2002). The genomes of model organisms show common ancestry, and their function with human genes is an effective tool for investigating human gene function and modeling human genetic diseases. This investigation, in turn, requires the correct assignment of orthology, which is provided through Orthodisease, which is considered an advanced step from the Homophila database (O'Brien *et al.,* 2004). With the availability of a complete sequence of many model organisms, comparative genomics and sequence analysis will enable the identification of the counterparts of many human disease genes and find more applications in the field of drug targeting (Dornburg *et al.,* 2022). Experimentally demonstrating a causal relationship between a gene and a disease is costly and time-consuming. Consequently, the associated expenditures are significantly decreased by thoroughly prioritizing potential genes prior to experimental testing. In Bombyx mori, spontaneous mutants and human disease-homolog network analyses have been proposed as models of protein aggregation diseases, although a full human APP-expressing silkworm line with measured amyloid aggregation remains to be reported (Tabunoki H, *et al.,* 2016).

Genome Annotation:

Genome annotation refers to the characterization of functional elements in the genome, where single-copy genes usually play critical roles in the main organism activities (Li *et al.,* 2019). Genome annotation is classified into two steps: structural annotation, which identifies the region responsible for specific features, and functional annotation, which implies the function and identity of genes and other elements, depending on sequence similarities. 1) Repeat sequences can be found using two different methods: homology searching, which uses BLAST, for example, to find homologous repeat sequences, and the ab initio prediction method, which uses structural features of the repetitive sequence, such as Augusts, to find novel repeat sequences. In the majority of insect genomes, it offers significant benefits for anticipating repeating sequences with unique structural characteristics, like long terminal repeats. A complete dataset of repeat sequences is created using both techniques (Liu *et al.,* 2014). 2) Identification of noncoding RNA: Noncoding RNA is a type of RNA gene that does not translate into protein products but plays critical regulatory roles in different biological processes, such as microRNAs (miRNAs), small interfering RNAs (siRNAs), piwi-interacting RNAs (piRNAs), transfer RNA (tRNA), and ribosomal RNA (Huntzinger, E., & Izaurralde, E. 2011). 3) Prediction of protein-coding genes: Prediction is the most important part and is performed using one of three methods: (a) determining homologues of known protein-coding genes based on sequence similarity (Pearson, W. R. 2013), (b) de novo protein-coding gene prediction using software developed through machine learning of protein-coding gene structures, and (c) determination of exons by direct transcriptome sequencing and alignment to the assembled scaffolds.

Prediction of Mutations of Diseased Genes and Gene Prioritization:

Growing genomic data resulting from annotation of genetic variation data for humans and other model organisms and prediction of mutations that have a molecular effect as single nucleotide polymorphism data improve our understanding of their functional elements.

Different methods are used to prioritize functional SNPs and mutations through the prediction of genes associated with the disease under study using OMIM or by identifying functional features near variability (genetic variation data) at different levels (Bromberg, Y. 2013; Mooney *et al.,* 2010). At the protein level, by determining the protein structure annotation, SNP location, annotation of known function sites through prediction, or by prediction of whether an amino acid substitution affects protein function (experimental amino acid substitution) for a conserved region. At the transcript level, mutations occur within the splicing factor-binding site or intron, exon splice sites, enhancers, and silencers. At the genome level, identifying SNPs that may affect gene expression or genetic features are near a candidate SNP, where genetic variation influences gene expression and phenotype

(Mooney *et al.,* 2010).

Gene prioritization is the process of determining the probability that a gene will be involved in the development of a disease phenotype. A number of pertinent characteristics, including gene expression and function, pathway involvement, and mutation effects are taken into account when assigning it. As a rule, disease genes have a tendency to interact with other disease genes, contain mutations that are functionally deleterious, code for proteins that are specific to the biological compartment in question (tissue, cellular space, or pathway), have unique sequence characteristics like greater length and more exons, and have more orthologs than paralogs (Bromberg, Y. 2013).

In general, functional evidence is used to infer gene-disease associations. This includes direct molecular interactions, transcriptional co-function (regulation, expression, and localization), genetic linkage, sequence/structure similarity, and paralogy (in-species homology resulting from a gene duplication event). Cross-species Evidence: Similar phenotypes in other organisms are attributed to homologs of the putative gene. Same-compartment Evidence: The suspected gene is active in tissues (like the liver), cellular compartments (like the cell membrane), and disease-associated pathways (like ion channels). Mutation Evidence: In the genomes of sick people, functionally harmful mutations impact suspected genes. Text Evidence: Terms related to genes and diseases frequently occur together in scientific texts (Bromberg, Y. 2013).

Identifying disease genes necessitates defining molecular pathways that may alter due to environmental disruptions (e.g., elevated temperatures from inflammation or decreased ligand concentrations from malnutrition), changes in the structure of the gene product (e.g., conformational change, binding site blockage, loss of ligand affinity), changes in gene expression (i.e., quantity and concentration of product), and the introduction of new pathway members (e.g., activation of previously silent genes). The study of single-nucleotide variations (SNV) and the prioritization of them for experimental characterization depend on computational methods for the effects of mutations on protein function. Table 5. Like the user-friendly web interface PredictSNP, http://loschmidt.chemi.muni.cz/predictsnp. This is a combination of the best prediction tools (MAPP, nsSNP Analyzer, PANTHER, PhD-SNP, PolyPhen-1, PolyPhen-2, SIFT, and SNAP) (Bendl *et al.,* 2014).

In coding regions, single nucleotide polymorphisms (SNPs) play a significant role in many human diseases. Synonymous SNPs (sSNPs) have no effect on translated proteins (Bromberg, Y. 2013). However, they can also affect mRNA stability and translation rates. Nonsynonymous SNPs (nsSNPs), which cause amino acid substitutions, have a direct impact on protein structure or function. SNPs in non-coding regions may affect gene splicing and other biological processes, such as RNA degradation and the transcription process (Tey, H. J., & Ng, C. H. 2019).

One of the specialized tools for deep studies and gene prioritization is an online database named Gene4PD (http://genemed.tech/gene4pd), which integrates published genetic data in PD to prioritize risk variants in PD. It provides researchers with comprehensive genetic knowledge, such as rare variants, associated SNPs, CNVs, differential expression, differential DNA methylation, and an analytic PD platform to improve their understanding of PD (Li *et al.,* 2021). Also, using RNA-seq–based prioritization, upregulation of DJ-1 homologs in Drosophila was shown to protect against oxidative stress, paralleling findings in Parkinson's disease patients (Lee Y, *et al.,* 2022).

Table 5: Different computational tools for similarity search, annotations, and predicting SNPs.

Database/ Tool	URL	Comment	Field	References
Alliance	https://www.alliancegenome.org/	Search Across Species to Explore model organism and human comparative genomics		(Agapite et al., 2022)
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi	Finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance.	sequence similarity search	(Johnson et al., 2008)
ConSurf	https://consurf.tau.ac.il/consurf_index.php	Evolutionary conservation analysis tools based on amino acid positions among the homologous protein sequences		(Ashkenazy et al., 2010)
DDIG	http://sparks-lab.org/server/ddig/	Detecting Disease-causing Genetic Variations due to Insertion/Deletion and Nonsense mutations	Gene Variation Analysis	(Livingstone et al., 2017)
EASE-MM	http://sparks-lab.org/server/ease-mm/	Sequence-based prediction of mutation-induced stability changes with feature-based multiple models	Gene Variation Analysis	(Folkman et al., 2016)
I-Mutant 2.0	https://folding.biofold.org/i-mutant/i-mutant2.0.html	Prediction of protein stability changes upon single-point mutation from protein sequence or structure; Reported Performance: ~80% prediction accuracy on a large dataset.	for nsSNPs	(Capriotti et al., 2005)
Interpro	https://www.ebi.ac.uk/interpro/	InterPro provides functional analysis of proteins by classifying them into families and predicting domains and important sites.	functional annotations	(Paysan-Lafosse et al., 2023)
MaxEntScan	http://www.ensembl.org/Tools/VEP	Predict splice site variants; Reported Performance: Considered a gold standard; high correlation with experimental data.		(McLaren et al., 2016)
Mutation Assessor	http://mutationassessor.org/r3/	Predict the effects of the nsSNPs	for nsSNPs	(Reva et al., 2011)
OMIM	https://www.omim.org/	An Online Catalog of Human Genes and Genetic Disorders	disease genes	(Hamosh et al., 2000)
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/index.shtml	Predicts possible impact of an amino acid substitution on the structure and function of a human protein; Reported Performance: AUC of ~0.85-0.90 on human disease variants	nsSNPs prediction	(Adzhubei et al., 2010)
Predict SNP2	https://loschmidt.chemi.muni.cz/predictsnp2/	A platform to evaluate the SNP accurately by using different characteristics of variants in distinct genomic regions; Reported Performance: Ranked best in benchmark with 87% accuracy and AUC 0.93.	nsSNPs prediction	(Bendl et al., 2016)
PROSITE	http://www.expasy.ch/prosite/	Database of biologically significant sites, patterns, and profiles in proteins that helps in function prediction.	functional annotations	(Ison et al., 2013)

RNAsnap2	http://sparks-lab.org/server/rnasnap2/	Single-sequence and Profile-based Prediction of RNA Solvent Accessibility Using a Dilated Convolutional Neural Network.	RNA prediction	(Hanumanthappa et al., 2021)
SIFT	https://sift.bii.a-star.edu.sg/	Predicts whether an amino acid substitution affects protein function; Reported Performance: Accuracy of ~80-85% in distinguishing neutral from deleterious variants.	nsSNPs prediction	(Ng, P.C., & Henikoff, S. , 2003).
SignalP - 6.0	https://services.healthtech.dtu.dk/services/SignalP-6.0/	The SignalP 6.0 server predicts the presence of signal peptides and the location of their cleavage sites in proteins	functional annotations	(Felix Teufel et al., 2021)
SMART	http://smart.embl-heidelberg.de/	Searchable collection of protein domains for detecting and analyzing domain architecture.	functional annotations	(Letunic et al., 2021)
SNPs&GO	https://snps.biofold.org/snps-and-go/snps-and-go.html	Predicting disease-associated variations using GO terms	for nsSNPs	(Thusberg et al., 2011)
SNPs&GO 3rd	https://snps.biofold.org/snps-and-go/snps-and-go-3d.html	Predicting disease-associated variations using GO terms	for nsSNPs	(Thusberg et al., 2011)
SNPs3D	http://www.snps3d.org/	Predicting functional nonsynonymous single nucleotide polymorphisms	nsSNPs prediction	(Yue et al., 2006)
SPARKS-X	http://sparks-lab.org/server/sparks-x/	Fold recognition and template-based structure prediction	Protein fold recognition	(Yang et al., 2011)
SPOT-RNA	http://sparks-lab.org/server/spot-rna/	RNA Secondary Structure Prediction using an Ensemble of Two-dimensional Deep Neural Networks and Transfer Learning; Reported Performance: Achieved high accuracy (>90%) on independent test sets.	RNA prediction	(Singh et al., 2019)
SPOT-RNA-1D	http://sparks-lab.org/server/spot-rna-1d/	RNA Backbone Torsion and Pseudotorsion Angle Prediction using Dilated Convolutional Neural Networks	RNA prediction	(Hanumanthappa et al., 2021)
SPOT-RNA2	http://sparks-lab.org/server/spot-rna2/	Improved RNA Secondary Structure and Tertiary Base-pairing Prediction using Evolutionary Profile, Mutational Coupling, and Two-dimensional Transfer Learning.	RNA prediction	(Singh et al., 2021)
SPRINT-Gly	http://sparks-lab.org/server/sprint-gly/	Predicting N- and O-linked glycosylation sites of human and mouse proteins by using sequence and predicted structural properties	Protein annotation	(Taherzadeh et al., 2019)
STRING	https://string-db.org/	Search Tool for the Retrieval of Interacting Genes/Proteins		(Szklarczyk et al., 2021)
SWISS-PDB Viewer and Swiss Model	http://www.expasy.ch/spdbv/	Powerful structure visualization and homology modeling system for personal computers.	Protein	(Johansson et al., 2012)
UTR Scan	https://web.archive.org/web/20190930171524/http://utrdb.ba.itb.cnr.it/	Predict UTR variants	Last update: 2015	(Tey, H. J., & Ng, C. H. 2019)

VAST	https://www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml	Structure similarity search tool based on the definition of the threshold of statistically significant structural similarity.	sequence similarity search	(Gibrat et al., 1996)

Benefit to Patients Suffering from Neurodegenerative Diseases:

The principal advantage of insect-based models lies in their accelerated Timeline for treatments. Species such as Drosophila melanogaster and Bombyx mori enable high-throughput screening of thousands of compounds within weeks rather than years, substantially shortening the preclinical stage of drug discovery. This acceleration increases the likelihood of translating promising leads into human trials within a much shorter timeframe.

Insect models offer a cost-effective alternative to mammalian systems due to their minimal maintenance and rapid generation cycles. This efficiency translates directly into lower research expenditures, potentially reducing the overall cost of therapy development. Such economic feasibility may attract greater investment in ND research, which is often constrained by the high attrition rates of traditional clinical trials.

While most current ND treatments remain symptomatic, insect models facilitate the exploration of fundamental disease mechanisms, such as protein aggregation (amyloid-β, tau, and α-synuclein) and mitochondrial dysfunction. By targeting these molecular processes, researchers can identify compounds that modify disease progression rather than strictly reducing symptoms for instance, molecules that enhance autophagic clearance of tau tangles in *Drosophila* neurons.

Conclusion

Bioinformatics and other multi-omics are used as the foundation of current and future biotechnology to find new drugs or better alternatives for human disease models and, moreover, to develop transgenic insects in applied insect science. In addition to the insects mentioned above, monarch butterflies, grasshoppers, houseflies, crickets, ants, American cockroaches, wax moths, and mosquitoes have been used in aging research, including neurodegenerative diseases, to investigate neurodegenerative diseases such as Huntington's disease, Alzheimer's disease, and Parkinson's disease. With the advent of new technologies for data sequencing and analysis, insects have a promising future in the field of biotechnology, as they can provide valuable insights into rare diseases. This review aims to provide comprehensive information on the available insect genomic resources for biotechnologists, molecular biologists, entomologists, and physiologists to develop new methods for modelling diseases and health management and to find answers to complex biological questions. Insect models can be used as tools for disease modelling and further investigation. For future research, it is recommended to identify specific insect species with potential importance for modelling different neurodegenerative diseases and discuss potential ethical considerations in the use of transgenic insects.

List of Abbreviations

6-OHDA	6-hydroxydopamine
A- beta peptide	succinate-CoA ligase ADP-forming subunit beta
a.a	amino acid
AD	Alzheimer's disease
APP	amyloid precursor protein
ASRT	Academy of Scientific Research and Technology
AUX	auxilin
bp	base pair

cDNA	Complementary deoxyribonucleic acid
CNS	Central nervous system
DNA	Deoxyribonucleic Acid
DRPLA	DentatoRubroPallidoLuysian Atrophy
EOAD	Early-onset familial AD
GO	Gene ontology
GRN	granulin precursor
HD	Huntington's disease
HDAC	histone deacetylase
HDL	Huntington Disease-Like
Homophilia	Human disease gene cognates in *Drosophila*
HTT	huntingtin
LDLR	low-density lipoprotein receptor
LOAD	Late-onset AD
LRRK2	Leucine-rich repeat kinase 2
MAPT	Microtubule-Associated Protein Tau
MLF1	Myeloid Leukemia Factor 1
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
NBIA	Neurodegeneration with Brain Iron Accumulation
NCBI	National Center for Biotechnology Information
ND	Neurodegenerative disease
Nep 2	Neprilysin 2
nsANP	Nonsynonymous SNP
OMIM	Online Mendelian Inheritance in Man
PARK2	E3 ubiquitin ligase complex
PARK7	Parkinsonism-Associated Deglycase
PD	Parkinson's disease
PINK1	(Phosphatase and tensin homologue)-induced kinase 1
PKAN	Pantothenate-Kinase-Associated-Neurodegeneration
PSEN1	presenilin 1
PSEN2	presenilin 2
PSN-1	presenilin protease
RBD	Rapid eye movement sleep behavior disorder
RNA	Ribonucleic acid
RNAi	RNA interference
SCA	Spinocerebellar ataxia
SNc	substantia nigra pars compacta
SNCA	alpha-synuclein
SNP	Single nucleotide polymorphisms
SORL1	Sortilin-Related Receptor 1
sSNP	Synonymous SNP
TOM	Transcriptomics of OMIM
TREM2	Triggering Receptor Expressed On Myeloid Cells 2
UTR	Untranslated regions
WHO	World Health Organization

Declarations:

Ethics Approval: Not applicable.

Authors' Contributions: All authors contributed to the study plan, preparation, interpretation of results, and writing, reviewing, and editing the manuscript.

Competing Interests: The author declares no conflict of interest of any kind.

Availability of Data and Materials: Authors confirm that all relevant data included in the article is in its supplementary information files.

Funding: Academy of Scientific Research and Technology (ASRT) for funding the Scholarship of Scientist of Next Generation (SNG) cycle 6, Grant no. ASRT/SNG/BGM/2018-8.

Acknowledgements: I would like to express appreciation to Dr. Mohamad Al-Hadidi, Dr. Enas Ghalab, Dr. Mohamad Awad, and Dr. Mohamad H. Moussa for their academic advice and assistance during the planning and development of this research work.

REFERENCES

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., … Venter, J. C. (2000). *The Genome Sequence of Drosophila melanogaster. Science,* 287(5461), 2185–2195. https://doi.org/10.1126/science.287.5461.2185

Adamski, Z., Bufo, S. A., Chowański, S., Falabella, P., Lubawy, J., Marciniak, P., Pacholska-Bogalska, J., Salvia, R., Scrano, L., Słocińska, M., Spochacz, M., Szymczak, M., Urbański, A., Walkowiak-Nowicka, K., & Rosiński, G. (2019). Beetles as model organisms in physiological, biomedical and environmental studies - A review. *In Frontiers in Physiology,* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fphys.2019.00319

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature methods,* 7(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Agapite, J., Albou, L.-P., Aleksander, S. A., Alexander, M., Anagnostopoulos, A. v, Antonazzo, G., Argasinska, J., Arnaboldi, V., Attrill, H., Becerra, A., Bello, S. M., Blake, J. A., Blodgett, O., Bradford, Y. M., Bult, C. J., Cain, S., Calvi, B. R., Carbon, S., Chan, J., … Zytkovicz, M. (2022). Harmonizing model organism data in the Alliance of Genome Resources. *Genetics, Volume 220, Issue 4, April 2022, iyac022,* https://doi.org/10.1093/genetics/iyac022

Al-Ayari, E.A., Shehata, M.G., EL-Hadidi, M. et al. In silico SNP prediction of selected protein orthologues in insect models for Alzheimer's, Parkinson's, and Huntington's diseases. Sci Rep 13, 18986 (2023). https://doi.org/10.1038/s41598-023-46250-5

Ali, A. M., & Kunugi, H. (2020 a). Apitherapy for Parkinson's Disease: A Focus on the Effects of Propolis and Royal Jelly. *In Oxidative Medicine and Cellular Longevity,* (Vol. 2020). Hindawi Limited. https://doi.org/10.1155/2020/1727142

Ali, A. M., & Kunugi, H. (2020 b). Royal jelly as an intelligent anti-aging agent— a focus on cognitive aging and Alzheimer's disease: A review. *In Antioxidants* (Vol. 9, Issue 10, pp. 1–46). MDPI. https://doi.org/10.3390/antiox9100937

Alvarez-Mora, M. I., Blanco-Palmero, V. A., Quesada-Espinosa, J. F., Arteche-Lopez, A. R., Llamas-Velasco, S., Palma Milla, C., Lezana Rosales, J. M., Gomez-Manjon, I., Hernandez-Lain, A., Jimenez Almonacid, J., Gil-Fournier, B., Ramiro-León, S.,

González-Sánchez, M., Herrero-San Martín, A. O., Pérez-Martínez, D. A., Gómez-Tortosa, E., Carro, E., Bartolomé, F., Gomez-Rodriguez, M. J., … Moreno-Garcia, M. (2022). Heterozygous and Homozygous Variants in SORL1 Gene in Alzheimer's Disease Patients: Clinical, Neuroimaging and Neuropathological Findings. International *Journal of Molecular Sciences,* 23(8), 4230. https://doi.org/10. 3390/ ijms23084230

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., & Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic acids research,* 38(Web Server issue), W529–W533. https://doi.org/10.1093/nar/gkq399

Baek, S. Y., Li, F. Y., Kim, J. H., Ahn, C., Kim, H. J., & Kim, M. R. (2020). Protein Hydrolysate of Silkworm Pupa Prevents Memory Impairment Induced by Oxidative Stress in Scopolamine-Induced Mice via Modulating the Cholinergic Nervous System and Antioxidant Defense System. Preventive nutrition and food science, 25(4), 389–399. https://doi.org/10.3746/pnf.2020.25.4.389

Batista, P., & Pereira, A. (2016). Quality of life in patients with neurodegenerative diseases. dimensions, 1(3).

Bendl, J., Musil, M., Štourač, J., Zendulka, J., Damborský, J., & Brezovský, J. (2016). PredictSNP2: A Unified Platform for Accurately Evaluating SNP Effects by Exploiting the Different Characteristics of Variants in Distinct Genomic Regions. *PLOS Computational Biology,* 12(5), e1004962. https://doi.org/10.1371/journal. pcbi.1004962

Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E. D., Zendulka, J., Brezovsky, J., & Damborsky, J. (2014). PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Computational Biology,* 10(1). https://doi.org/10.1371/journal.pcbi.1003440

Better, M. A. (2023). Alzheimer's disease facts and figures. *Alzheimers Dement,* 19(4), 1598-1695.

Brandt, A., Joop, G., & Vilcinskas, A. (2019). Tribolium castaneum as a whole-animal screening system for the detection and characterization of neuroprotective substances. *Archives of Insect Biochemistry and Physiology,* 100(3). e21532 https://doi.org/10.1002/arch.21532.

Bromberg, Y. (2013). Chapter 15: Disease Gene Prioritization. *PLoS Computational Biology,* 9(4). https://doi.org/10.1371/journal.pcbi.1002902

Brown, D. G., & Wobst, H. J. (2023). Analysis of orphan designation status for FDA approved drugs, and case studies in oncology, neuroscience and metabolic diseases. *Bioorganic & Medicinal Chemistry,* 117170.

Capriotti, E., Fariselli, P., & Casadio, R. (2005). I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic acids research,* 33(Web Server issue), W306–W310. https://doi.org/10.1093/nar/gki375

Carrarini, C., Russo, M., Dono, F., di Pietro, M., Rispoli, M. G., di Stefano, V., Ferri, L., Barbone, F., Vitale, M., Thomas, A., Sensi, S. L., Onofrj, M., & Bonanni, L. (2019). A Stage-Based Approach to Therapy in Parkinson's Disease. *Biomolecules,* 9(8), 388. https://doi.org/10.3390/biom9080388

Checkoway, H., Lundin, J. I., & Kelada, S. N. (2011). Neurodegenerative diseases. *IARC Scientific Publications,* 163, 407–419

Chekani, F., Bali, V., & Aparasu, R. R. (2016). Quality of life of patients with Parkinson's disease and neurodegenerative dementia: A nationally representative study. *Research in Social and Administrative Pharmacy,* 12(4), 604–613. https://doi.org/ 10.1016/j.sapharm.2015.09.007

Chen X, Li Y, Huang J, Cao D, Yang G, Liu W, Lu H, and Guo A (2007) Study of tauopathies by comparing Drosophila and human tau in Drosophila. C*ell Tissue Research,* 329:169 –178.

Chien, S., Reiter, L. T., Bier, E., & Gribskov, M. (2002). Homophila: human disease gene cognates in Drosophila. *In Nucleic Acids Research,* (Vol. 30, Issue 1). http://homophila.sdsc.edu.

Chilana, P., Sharma, A., & Rai, A. (2012). Insect genomic resources: status, availability and future. *Current science,* 571-580.

Cruts, M., Theuns, J., & van Broeckhoven, C. (2012). Locus-specific mutation databases for neurodegenerative brain diseases. *Human Mutation,* 33(9), 1340–1344. https://doi.org/10.1002/humu.22117

Dornburg, A., Mallik, R., Wang, Z., Bernal, M. A., Thompson, B., Bruford, E. A., ... & Townsend, J. P. (2022). Placing human gene families into their evolutionary context. *Human Genomics,* 16(1), 56.

El-Jaafary S., Sabbah A., & Amer H.. (2020). Huntington's disease in Egypt; A big challenge for a mysterious disease. *MDS Virtual Congress.* 35, S96. https://www.mdsabstracts.org/ abstract /huntingtons- disease- in-egypt- a- big- challenge -for -a -mysterious-disease/

el Tallawy, H. N. A., Farghaly, W. M. A., el Hamed, M. A., Badry, R., Usama, K., Shehata, G. A., Tohamy, A. M., Abdulghani, K. O., Ghanem, M. K., Sayed, M. A., Yousef, A. H., Hashem, H. S., & Rageh, T. A. (2019). Prevalence of Alzheimer dementia in upper Egypt (Desert areas). *Egyptian Journal of Neurology, Psychiatry and Neurosurgery,* 55(1). https://doi.org/10.1186/s41983-019-0074-y

El-Metwally, A., Toivola, P., Al-Rashidi, M., Nooruddin, S., Jawed, M., Alkanhal, R., Razzak, H. A., & Albawardi, N. (2019). Epidemiology of Alzheimer's Disease and Dementia in Arab Countries: A Systematic Review. *Behavioural Neurology,* 2019. https://doi.org/10.1155/2019/3935943

Felix Teufel, José Juan Almagro Armenteros, Alexander Rosenberg Johansen, Magnús Halldór Gíslason, Silas Irby Pihl, Konstantinos D Tsirigos, Ole Winther, Søren Brunak, Gunnar Von Heijne and Henrik Nielsen. *Nature Biotechnology,* (2021), https://doi:10.1038/s41587-021-01156-3

Folkman, L., Stantic, B., Sattar, A., & Zhou, Y. (2016). EASE-MM: Sequence-Based Prediction of Mutation-Induced Stability Changes with Feature-Based Multiple Models. *Journal of molecular biology,* 428(6), 1394–1405. https://doi.org/10.1016/j.jmb.2016.01.012

GhazalehT, Abdollah D., Maryam G., Yaoqi Z., & Matthew P. C. (2019). SPRINT-Gly: Predicting N- and O-linked glycosylation sites of human and mouse proteins by using sequence and predicted structural properties. Bioinformatics. 15;35(20):4140-4146. doi: 10.1093/bioinformatics/btz215.

Gibrat, J. F., Madej, T., & Bryant, S. H. (1996). Surprising similarities in structure comparison. *Current opinion in structural biology,* 6(3), 377–385. https://doi.org/10.1016/s0959-440x(96)80058-3

Guadagnolo, D., Piane, M., Torrisi, M. R., Pizzuti, A., & Petrucci, S. (2021). Genotype-Phenotype Correlations in Monogenic Parkinson Disease: A Review on Clinical and Molecular Findings. In *Frontiers in Neurology,* (Vol. 12). Frontiers Media S.A. https://doi.org/10.3389/fneur.2021.648588

Habeeb, S.K.M, & Raman Chandrasekear. (2014). Short Views on Insect Biochemistry and Molecular Biology. Edition: First Edition. Chapter: 28. Publisher: International Book Mission, Academic Publisher, Manhattan, KS, USA.

Hakeemi, M. S., Ansari, S., Teuscher, M., Weißkopf, M., Großmann, D., Kessel, T., Dönitz,

J., Siemanowski, J., Wan, X., Schultheis, D., Frasch, M., Roth, S., Schoppmeier, M., Klingler, M., & Bucher, G. (2022). Screens in fly and beetle reveal vastly divergent gene sets required for developmental processes. *BMC Biology,* 20(1), 38. https://doi.org/10.1186/s12915-022-01231-4

Haleem Hany Farid, & Grace Bahaa Demian. (2019). Egypt Leads World in Huntington's Prevalence, But Lacks Clinical Trials on the Disease, Report Says. Huntington's Disease News. https://huntingtonsdiseasenews.com/news/egypt-leads-world-in-huntingtons-prevalence-but-lacks-clinical-trials-on-the-disease-report-says/?cn-reloaded=1

Hamosh, A., Scott, A. F., Amberger, J., Valle, D., & McKusick, V. A. (2000). Online Mendelian inheritance in man (OMIM). *Human mutation,* 15(1), 57-61.

Hanumanthappa, A. K., Singh, J., Paliwal, K., Singh, J., & Zhou, Y. (2021). Single-sequence and profile-based prediction of RNA solvent accessibility using dilated convolutional neural network. *Bioinformatics,* 36(21), 5169–5176. https://doi.org/10.1093/bioinformatics/btaa652

Honeybee Genome Sequencing Consortium. (2006). Insights into social insects from the genome of the honeybee Apis mellifera. *Nature,* 443(7114), 931.

Huntzinger, E., & Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature Reviews Genetics,* 12(2), 99–110. https://doi.org/10.1038/nrg2936

Ison, J., Kalas, M., Jonassen, I., Bolser, D., Uludag, M., McWilliam, H., Malone, J., Lopez, R., Pettifer, S., & Rice, P. (2013). EDAM: an ontology of bioinformatics operations, types of data and identifiers, topics and formats. *Bioinformatics (Oxford, England), 29*(10), 1325–1332. https://doi.org/10.1093/bioinformatics/btt113

Johansson, M.U., Zoete V., Michielin O. & Guex N. (2012) Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. *BMC Bioinformatics,* 13:173.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic acids research,* 36(Web Server issue), W5–W9. https://doi.org/10.1093/nar/gkn201

Kloske, C. M., Dugan, A. J., Weekman, E. M., Winder, Z., Patel, E., Nelson, P. T., Fardo, D. W., & Wilcock, D. M. (2021). Inflammatory Pathways Are Impaired in Alzheimer Disease and Differentially Associated With Apolipoprotein E Status. *Journal of neuropathology and experimental neurology,* 80(10), 922–932. https://doi.org/10.1093/jnen/nlab085

Lacreuse, A., & Herndon, J. G. (2008). Nonhuman primate models of cognitive aging. In Animal models of human cognitive aging (pp. 1-30). Totowa, NJ: Humana press.

Lee Y, Kim J, Kim H, Han JE, Kim S, Kang KH, Kim D, Kim JM, Koh H. (2022). Pyruvate Dehydrogenase Kinase Protects Dopaminergic Neurons from Oxidative Stress in Drosophila DJ-1 Null Mutants. ***Molecules and Cells,*** Jul 31;45(7):454-464. Epub 2022 Apr 21. PMID: 35444068; PMCID: PMC9260132. https://doi.org/10.14348/molcells.2022.5002

Letunic, I., Khedkar, S., & Bork, P. (2021). SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research,* 49(D1), D458–D460. https://doi.org/10.1093/nar/gkaa937

Li, B., Zhao, G., Zhou, Q., Xie, Y., Wang, Z., Fang, Z., Lu, B., Qin, L., Zhao, Y., Zhang, R., Jiang, L., Pan, H., He, Y., Wang, X., Luo, T., Zhang, Y., Wang, Y., Chen, Q., Liu, Z., … Li, J. (2021). Gene4PD: A Comprehensive Genetic Database of Parkinson's Disease. *Frontiers in Neuroscience,* 15.679568. https://doi.org/10.3389/fnins.2021.679568

Li, F., Zhao, X., Li, M., He, K., Huang, C., Zhou, Y., Li, Z., & Walters, J. R. (2019). Insect

genomes: progress and challenges. *Insect Molecular Biology,* 28(6), 739–758. https://doi.org/10.1111/imb.12599

Liu, J., Xiao, H., Huang, S., & Li, F. (2014). OMIGA: Optimized Maker-Based Insect Genome Annotation. *Molecular Genetics and Genomics,* 289(4), 567–573. https://doi.org/10.1007/s00438-014-0831-7

Livingstone, M., Folkman, L., Yang, Y., Zhang, P., Mort, M., Cooper, D. N., Liu, Y., Stantic, B., & Zhou, Y. (2017). Investigating DNA-, RNA-, and protein-based features as a means to discriminate pathogenic synonymous variants. *Human mutation,* 38(10), 1336–1347. https://doi.org/10.1002/humu.23283

Luheshi LM, Tartaglia GG, Brorsson AC, Pawar AP, Watson IE, Chiti F, Vendruscolo M, Lomas DA, Dobson CM, and Crowther DC (2007) Systematic in vivo analysis of the intrinsic determinants of amyloid beta pathogenicity. *PLoS Biology,* 5:e290.

May RM. How many species are there on Earth? Science. 1988 Sep 16;241(4872):1441-9. PMID: 17790039. https://doi.org/10.1126/science.241.4872.1441.

McColgan, P., & Tabrizi, S. J. (2018). Huntington's disease: a clinical review. *European journal of neurology, 25*(1), 24–34. https://doi.org/10.1111/ene.13413

McGurk, L., Berson, A., & Bonini, N. M. (2015). Drosophila as an in vivo model for human neurodegenerative disease. *Genetics,* 201(2), 377–402. https://doi.org/10.1534/genetics.115.179457

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome biology,* 17(1), 122. https://doi.org/10.1186/s13059-016-0974-4

Mooney, S. D., Krishnan, V. G., & Evani, U. S. (2010). Bioinformatic tools for identifying disease gene and SNP candidates. In Methods in Molecular Biology (Vol. 628, pp. 307–319). Humana Press Inc. https://doi.org/10.1007/978-1-60327-367-1_17

Moreno, C., Lazar, J., Jacob, H. J., & Kwitek, A. E. (2008). Comparative Genomics for Detecting Human Disease Genes (pp. 655–697). https://doi.org/10.1016/S0065-2660(07)00423-3

Munkácsy, E., & Pickering, A. M. (2021). Model organisms (invertebrates). In Handbook of the Biology of Aging (pp. 199-217). Academic Press.

Ng, P.C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research,* 31(13), 3812-3814.

O'Brien, K. P., Westerlund, I., & Sonnhammer, E. L. L. (2004). OrthoDisease: A database of human disease orthologs. *Human Mutation,* 24(2), 112–119. https://doi. org/10. 1002/humu.20068

Pandey, U. B., & Nichols, C. D. (2011). Human disease models in drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacological Reviews,* 63(2), 411–436. https://doi.org/10.1124/pr.110.003293

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., Sigrist, C. J. A., … Bateman, A. (2023). InterPro in 2022. *Nucleic acids research,* 51(D1), D418–D427. https://doi.org/10.1093/nar/gkac993

Pearson, W. R. (2013). An Introduction to Sequence Similarity ("Homology") Searching. *Current Protocols in Bioinformatics,* 42(1). https://doi.org/10.1002/0471250953.bi0301s42

Pesah Y, Burgess H, Middlebrooks B, Ronningen K, Prosser J, Tirunagaru V, Zysk J, and Mardon G (2005) Whole-mount analysis reveals normal numbers of dopaminergic neurons following misexpression of alpha-Synuclein in Drosophila. *Genesis,* 41:154–159.

Poelchau, M.F., Chen, MJ.M., Lin, YY., Childers, C.P. (2018). Navigating the i5k Workspace@NAL: A Resource for Arthropod Genomes. In: Kollmar, M. (eds) Eukaryotic Genomic Databases. Methods in Molecular Biology, vol 1757. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7737-6_18

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research,* 39(17), e118. https://doi.org/10.1093/nar/gkr407

Romero E, Cha GH, Verstreken P, Ly CV, Hughes RE, Bellen HJ, and Botas J (2008) Suppression of neurodegeneration and increased neurotransmission caused by expanded full-length huntingtin accumulating in the cytoplasm. Neuron 57:27–40.

Roy, S. (2013). Genomics and Bioinformatics in Entomology. Entomology, Ornithology & Herpetology: *Current Research,* 02(03). https://doi.org/10.4172/2161-0983. 1000e107

Schmitt-Engel, C., Schultheis, D., Schwirz, J., Ströhlein, N., Troelenberg, N., Majumdar, U., Dao, V. A., Grossmann, D., Richter, T., Tech, M., Dönitz, J., Gerischer, L., Theis, M., Schild, I., Trauner, J., Koniszewski, N. D. B., Küster, E., Kittelmann, S., Hu, Y., … Bucher, G. (2015). The iBeetle large-scale RNAi screen reveals gene functions for insect development and physiology. *Nature Communications,* 6(1), 7822. https://doi.org/10.1038/ncomms8822

Seidner, G. A., Ye, Y., Faraday, M. M., Alvord, W. G., & Fortini, M. E. (2006). Modeling clinically heterogeneous presenilin mutations with transgenic Drosophila. *Current Biology,* 16(10), 1026-1033.

Shpigler, H. Y., Saul, M. C., Murdoch, E. E., Cash-Ahmed, A. C., Seward, C. H., Sloofman, L., Chandrasekaran, S., Sinha, S., Stubbs, L. J., & Robinson, G. E. (2017). Behavioral, transcriptomic and epigenetic responses to social challenge in honey bees. *Genes, Brain and Behavior,* 16(6), 579–591. https://doi.org/10.1111/ gbb.12379

Singh, J., Hanson, J., Paliwal, K., & Zhou, Y. (2019). RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nature Communications,* 10(1), 5407. https://doi.org/10.1038/s41467-019-13395-9

Singh, J., Paliwal, K., Zhang, T., Singh, J., Litfin, T., & Zhou, Y. (2021). Improved RNA secondary structure and tertiary base-pairing prediction using evolutionary profile, mutational coupling and two-dimensional transfer learning. *Bioinformatics,* 37(17), 2589–2600. https://doi.org/10.1093/bioinformatics/btab165

Song, L., He, Y., Ou, J., Zhao, Y., Li, R., Cheng, J., Lin, C. H., & Ho, M. S. (2017). Auxilin Underlies Progressive Locomotor Deficits and Dopaminergic Neuron Loss in a Drosophila Model of Parkinson's Disease. *Cell Reports,* 18(5), 1132–1143. https://doi.org/10.1016/j.celrep.2017.01.005

Strausfeld, N. J., & Hirth, F. (2013). Deep homology of arthropod central complex and vertebrate basal ganglia. *Science,* 340(6129), 157-161.

Suetsugu, Y., Futahashi, R., Kanamori, H., Kadono-Okuda, K., Sasanuma, S., Narukawa, J., Ajimura, M., Jouraku, A., Namiki, N., Shimomura, M., Sezutsu, H., Osanai-Futahashi, M., Suzuki, M. G., Daimon, T., Shinoda, T., Taniai, K., Asaoka, K., Niwa, R., Kawaoka, S., … Mita, K. (2013). Large Scale Full-Length cDNA Sequencing Reveals a Unique Genomic Landscape in a Lepidopteran Model Insect, Bombyx mori. *G3 Genes|Genomes|Genetics,* 3(9), 1481–1492. https://doi.org/10.1534/ g3.113.006239

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: customizable protein–protein networks, and functional

characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research,* 49(D1), D605–D612. https://doi.org/10.1093/nar/gkaa1074

Tabunoki, H., Bono, H., Ito, K., & Yokoyama, T. (2016). Can the silkworm (Bombyx mori) be used as a human disease model? *Drug Discoveries & Therapeutics,* 10(1), 3–8. https://doi.org/10.5582/ddt.2016.01011

Tabunoki, H., Ono, H., Ode, H., Ishikawa, K., Kawana, N., Banno, Y., Shimada, T., Nakamura, Y., Yamamoto, K., Satoh, J., & Bono, H. (2013). Identification of Key Uric Acid Synthesis Pathway in a Unique Mutant Silkworm Bombyx mori Model of Parkinson's Disease. *PLoS ONE,* 8(7), e69130. https://doi.org/10.1371/journal.pone.0069130

Taherzadeh, G., Dehzangi, A., Golchin, M., Zhou, Y., & Campbell, M. P. (2019). SPRINT-Gly: predicting N- and O-linked glycosylation sites of human and mouse proteins by using sequence and predicted structural properties. *Bioinformatics (Oxford, England),* 35(20), 4140–4146. https://doi.org/10.1093/bioinformatics/btz215

Tello, J. A., Williams, H., Eppler, R. M., Steinhilb, M. L., & Khanna, M. (2022). Animal models of neurodegenerative disease: recent advances in fly highlight innovative approaches to drug discovery. *Frontiers in Molecular Neuroscience.* https://doi.org/10.3389/fnmol.2022.883358

Tey, H. J., & Ng, C. H. (2019). Computational analysis of functional SNPs in Alzheimer's disease-associated endocytosis genes. *PeerJ,* 2019(9). https://doi.org/10. 7717/peerj.7667

The International Silkworm Genome Consortium. (2008). The genome of a lepidopteran model insect, the silkworm Bombyx mori. *Insect Biochemistry and Molecular Biology,* 38(12), 1036–1045. https://doi.org/10.1016/j.ibmb.2008.11.004

Thusberg, J., Olatubosun, A., & Vihinen, M. (2011). Performance of mutation pathogenicity prediction methods on missense variants. *Human mutation,* 32(4), 358–368. https://doi.org/10.1002/humu.21445

Ugur, B., Chen, K., & Bellen, H. J. (2016). Drosophila tools and assays for the study of human diseases. *In DMM Disease Models and Mechanisms* (Vol. 9, Issue 3, pp. 235–244). Company of Biologists Ltd. https://doi.org/10.1242/dmm.023762

Vincent, J. (2023). Life stage or Age? Reviewing perceptions of oldest digital technologies users. In Digital Ageism (pp. 36-52). Routledge.

Waizumi, R., Tsubota, T., Jouraku, A., Kuwazaki, S., Yokoi, K., Iizuka, T., ... & Sezutsu, H. (2023). Highly accurate genome assembly of an improved high-yielding silkworm strain, Nichi01. *G3: Genes, Genomes, Genetics,* 13(4), jkad044.

Woodroof, H. I., Pogson, J. H., Begley, M., Cantley, L. C., Deak, M., Campbell, D. G., et al. (2011). Discovery of catalytically active orthologues of the Parkinson's disease kinase PINK1: analysis of substrate specificity and impact of mutations. Open Biol. 1:110012. https://doi.org/10.1098/rsob.110012

Xi Y, Ryan J, Noble S, Yu M, Yilbas AE, and Ekker M (2010) Impaired dopaminergic neuron development and locomotor function in zebrafish with loss of pink1 function. *European Journal of Neuroscience,* 31:623– 633.

Yang, Y., Faraggi, E., Zhao, H., & Zhou, Y. (2011). Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics (Oxford, England),* 27(15), 2076–2082. https://doi.org/10.1093/bioinformatics/btr350

Yang, Y., Faraggi, E., Zhao, H., & Zhou, Y. (2011). Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native

properties of templates. *Bioinformatics,* 27(15), 2076-2082.

Yue, P., Melamud, E., & Moult, J. (2006). SNPs3D: candidate gene and SNP selection for association studies. *BMC bioinformatics,* 7, 166. https://doi.org/10.1186/1471-2105-7-166

Zeng J, Li Y, Yan J, Chang R, Xu M, Zhou G, Meng J, Liu D, Mao Z, Yang Y. Gut microbiota from patients with Parkinson's disease causes motor deficits in honeybees. *Front Microbiology,* 2024 Jul 12;15:1418857. PMID: 39070266; PMCID: PMC11272988. https://doi.org/10.3389/fmicb.2024.1418857.

Zhang H, Yang J, Guo Y, Lü P, Gong X, Chen K, Li X, Tang M. Rotenone-induced PINK1/Parkin-mediated mitophagy: establishing a silkworm model for Parkinson's disease potential. *Frontiers in Molecular Neuroscience*,2024 Apr 19;17:1359294. PMID: 38706874; PMCID: PMC11066238. https://doi.org/10.3389/fnmol.2024.1359294.

Zhu, F., Chen, H., Han, J., Zhou, W., Tang, Q., Yu, Q., Ma, S., Liu, X., Huo, S., & Chen, K. (2022). Proteomic and Targeted Metabolomic Studies on a Silkworm Model of Parkinson’s Disease. *Journal of Proteome Research,* 21(9), 2114–2123. https://doi.org/10.1021/acs.jproteome.2c00149