Abstract: Abnormal expressions of microRNA (miRNA) can result in human diseases such as cancer and neurodegenerative diseases. MiRNA mainly exert their biological functions via repressing the expression of their target genes. Drosophila melanogaster (D. melanogaster) is an ideal model for studying the molecular mechanisms behind biological phenotypes, including human diseases. In this study, we collected human and D. melanogaster miRNA as well as known human disease-related genes. In total, we identified 136 human disease-related miRNA that are orthologous to 83 D. melanogaster miRNA by mapping “seed sequence”, and 677 human disease-related genes that are orthologous to 734 D. melanogaster genes using the DRSC Integrative Ortholog Prediction Tool. Furthermore, we revealed the target relationship between genes and miRNA using miRTarBase database and target prediction software, including miRanda and TargetScan. In addition, we visualized interaction networks and signalling pathways for these filtered miRNA and target genes. Finally, we compiled all the above data and information to generate a database designated DHDD. This is the first comprehensive collection of human disease-related miRNA and their targeting genes conserved in a D. melanogaster database. The DHDD provides a resource for easily searching human disease-related miRNA and their disease-related target genes as well as their orthologs in D. melanogaster, and conveniently identifying the regulatory relationships among them in the form of a visual network.

Keywords: database; miRNA; disease; Drosophila melanogaster

1. Introduction

Over 3700 human genes with phenotype-causing mutations are identified and stored in the Online Mendelian Inheritance in Man database (OMIM database). Studies of human disease gene orthologs in model organisms have contributed to improved understanding of the molecular mechanisms of human disease [1]. Genome-wide analysis shows that approximately 70% of all known human disease-related genes have orthologous genes in D. melanogaster [2]. The major signal pathways involved in human disease are also evolutionarily conserved between human and D. melanogaster, for example, EGFR/RTK-Ras, PI3K, Notch, Wnt, Jak-STAT, Hedgehog and TGF-b pathways, which were first discovered from genetic studies of D. melanogaster [3–6]. In addition, body structures of adult
D. melanogaster can perform equivalent functions to mammalian organs, such as the heart, kidney, lung, and reproductive system [7]. D. melanogaster is one of the most powerful models for studying the genetics and biology of human diseases [8–11], and is widely used as a model organism to study the pathogenesis of human diseases, such as cancer, neurodegenerative diseases, metabolic diseases [12–15], etc. The employment of D. melanogaster as a model system has successfully revealed the genetic and molecular mechanisms behind the process of neuronal degeneration, angiogenesis, innate immune response, stem cell selection and maintenance, cell and tissue polarity, signal transduction, growth control, behavioural neural control and organogenesis [11,14,16–18].

Identifying human gene orthologs in D. melanogaster is necessary to facilitate the use of D. melanogaster as the model system to study the function and mechanisms of human genes. Several databases, such as the Homophila database, the orthodisease database and the DIOPT-DIST database have been established [19–21], to provide valuable information for the further use of D. melanogaster as a model for the exploration of the molecular mechanisms of human diseases via cross-species genomic data analysis between human and D. melanogaster.

MiRNA play important roles in various human diseases, including cancers, cardiovascular diseases, neurodegenerative disorders, etc. [22–26]. In humans, miRNA are associated with diseases via regulation of the expression of disease-related genes. Many human disease-related miRNA databases are available, for example, miR2Disease, The Human microRNA Disease Database [27–30]. Several databases, for example, miRGeneDB and miROrtho, provide information about orthologous miRNA between humans and D. melanogaster. However, no database provides information on the orthologs of human disease-related miRNA and their disease-related target genes in D. melanogaster to date.

To further facilitate the use of D. melanogaster as human disease model, in this study we constructed a Dme-Hsa disease database (DHDD) (http://bioinf.njnu.edu.cn/dhdd/home.php), an open web service, by identifying and analysing these orthologs of human disease-related miRNA and genes in D. melanogaster. In the DHDD, we provided information on human disease-related miRNA and genes and their orthologs in D. melanogaster. This database includes 83 miRNA and 734 genes of D. melanogaster that are homologous to 136 human miRNA and 677 human disease-related genes, respectively, as well as those related genes involved in 65 D. melanogaster pathways and 212 human pathways.

2. Database Description

2.1. Data Sources

Data were retrieved from miRBase (v22), Ensembl (release 89) and OMIM. In brief, mature sequences of the miRNA of Homo sapiens and D. melanogaster were downloaded from the miRBase (v22) database. Nucleotide sequences of H. sapiens were downloaded from Ensembl (release 89), and nucleotide sequences of D. melanogaster were downloaded from FlyBase (r6.08) [31–33], respectively. The list of disorders, disease genes, and associations among them was obtained from the Online Mendelian Inheritance in Man (OMIM) [34]. We classified these disease genes into 22 disorder classes based on their effects on different physiological systems [35].

2.2. Database Construction

The workflow of the database construction is shown in Figure 1. In Step 1, we downloaded human disease-related genes from OMIM, and predicted their Drosophila orthologs using the DRSC Integrative Ortholog Prediction Tool (v7.1) (http://www.flyrnai.org.diopt). In Step 2, we downloaded human and Drosophila miRNA from miRBase (v22), then identified 156 human miRNAs with 94 orthologs of D. melanogaster miRNA by matching the seed sequences of miRNA (position 2–7 of 5′ end in mature miRNA) between human and D. melanogaster. In Step 3, we identified the orthologous pairs of miRNA to genes (OPMG). The process for identification of the OPMG was briefly as follows. Firstly, the 3′UTR
of human and fly genes were downloaded from Ensembl database (release 89). Secondly, Targetscan (targetscan_70.pl) and miRanda (version 3.3a) were used to predict the regulation relationships between miRNA and genes in humans and Drosophila, respectively. The intersections of the predicted results of the two softwares were accepted, and the weak site type (6 mer) was filtered out. The experimentally validated relationships of miRNA regulating target genes were obtained from the miRTarBase, and then the intersections of the prediction and the experimental results were gained [36–39]. In Step 4, we used the KEGG database to annotate human and fly genes [40,41]. In Step 5, we used the Cytoscape program to visualize the OPMG pairs, pathways, and networks [42].

![Diagram](image_url)

**Figure 1.** The analysis workflow of human disease-related miRNA and genes with orthologs in the fruit fly.

The DHDD database has been developed as a web page with ‘Home’, ‘Search’, ‘Browser’, ‘Download’ and ‘Help’ pages using an Apache HTTP web server and MySQL database server. The web page was constructed using the PHP language, and data was stored by MySQL. All data were processed using Perl language script. The networks were presented as pictures based on the analysis results. Queries from the ‘Search’ and ‘Browse’ pages retrieve the results from MySQL. All data in the DHDD database can be downloaded from the ‘Download’ page. Through the ‘Help’ page, users can get easily information on how to use our database and how to interpret the results of the search.

2.3. Database Content

2.3.1. Identification of Orthologs of Human Disease-Related Genes in *D. melanogaster*

In total, 3747 human disease-related genes were retrieved from the OMIM morbid map and classified into 22 groups based on their involvement in physiological systems. The results showed that 3160 (84.33%) human disease-related genes had 4252 orthologs in *D. melanogaster*. The number
of human disease-related genes and their *D. melanogaster* orthologs in different disease groups were calculated in Figure 2.

**Figure 2.** Statistical number of human disease-related genes and miRNA orthologs with the fly in different disease groups.

2.3.2. Identification of Orthologs of Human Disease-Related miRNA in *D. melanogaster*

We identified 200 orthologous miRNA pairs containing 156 human miRNA and 94 *D. melanogaster* miRNA, which contribute to 13.52% and 8% of total miRNA number in human and *D. melanogaster*, respectively. The number of human miRNA target to disease-related genes and their orthologs in *D. melanogaster* were also calculated in Figure 2.

2.3.3. Identification the Orthologous Pairs of miRNA to Gene (OPMG) in Human and *D. melanogaster*

We identified 4104 OPMG data entries, include 136 human miRNAs with orthologous to 83 *D. melanogaster* miRNAs, which respectively target 677 human disease-related genes and 734 orthologous genes in *D. melanogaster*. Remarkably, the gene orthologous pairs and miRNA orthologous pairs of *Drosophila* and human are not one-to-one, and the interaction in an OPMG may also be orthologous between human and fly.

2.3.4. Pathway Annotation of Ortholog miRNA Target Genes between Human and *D. melanogaster*

In order to annotate the biological functions of OPMG, we analysed the genes with KEGG (Kyoto Encyclopedia of Genes and Genomes) annotations and identified 212 pathways for human disease-related genes and 65 pathways for *D. melanogaster* genes. The common pathways between human and *D. melanogaster* include Wnt, mTOR, Hippo, MAPK, TGF-beta, FoxO, and the Jak-STAT signaling pathway, etc. It is worth noting that the large number of genes from both human and *D. melanogaster* were annotated to be involved in ‘metabolic pathways’.
2.3.5. Construction of Visual Networks

The network of each disease group, human pathway, *Drosophila* pathway, human miRNA and their target genes as well as the orthologs in *D. melanogaster* were constructed based on the OPMG analysis. Based on the disease group and pathway, we constructed 22 networks of disease groups, including 216 networks of human pathways and 60 networks of *D. melanogaster* pathways in total. There are four types of nodes (gene, miRNA, disease group and pathway) and three types of edges (miRNA-target gene, orthologous and gene pathway annotation). Figure 3 shows an example diagram of a network. All of these networks were presented in the DHDD database as a dynamic network. At present, using *D. melanogaster* as a model organism to study human disease-related genes is mainly concentrated in the two groups ‘cancer’ and ‘neurological’ disease [43–45]. Among these networks, the ‘neurological’ network has 25 human genes and 27 *D. melanogaster* genes, and the ‘cancer’ network has 15 human genes and 16 *D. melanogaster* genes. In addition, the number of genes/miRNA in each network are listed in Table 1.

![Example diagram of a network](image)

**Figure 3.** Example diagram of a network. Oval nodes represent genes, triangle nodes represent miRNA, squares nodes represent disease groups and pathways. Orthologous miRNA or genes are linked with a yellow line; miRNA are connected to genes by red lines with an arrow; genes are linked to pathways by gray lines.

2.3.6. Web Server Introduction and Use

On the search page of the DHDD database, the user can conveniently retrieve information (gene, miRNA and pathway) using keywords, such as the gene symbol and miRNA name (Figure 4a). For example, on the search web page, if the user searches the keyword “fly miRNA name: mir-11”, a return list includes *Drosophila* miRNA and genes, human miRNA and genes, the type of disease associated with mir-11, as well as the information on targeting relationships. Whilst the information on targeting relationships can be classified into two types, acquired from the miRTarbase (link to miRTarbase) or the results predicted by the software, and the database presents the miRTarbase ID (there could also be a link to the miRTarbase or the scores, energy and site type for target prediction). In addition, the human genes are also linked to the Ensembl database, the *Drosophila* genes are linked to the FlyBase database, and all miRNA are linked to the miRBase database. On the browse page, the user can easily retrieve information from the OPMG based on different categories (disease group, human pathway or fly pathway) (Figure 4b). The retrieval results include all OPMG information under this category and provides a visual network. The help page provides more details about the database. In addition, the download page can download all the data in the DHDD database.
Table 1. The different nodes in all disease sub-networks.

| Disease                  | Human gene | Human miRNA | Fly gene | Fly miRNA |
|--------------------------|------------|-------------|----------|-----------|
| Bone                     | 10         | 35          | 11       | 13        |
| Cancer                   | 15         | 54          | 16       | 31        |
| Cardiovascular disorder  | 6          | 10          | 5        | 7         |
| Dermatological           | 3          | 10          | 3        | 10        |
| Developmental            | 2          | 3           | 3        | 3         |
| Ear Nose Throat          | 6          | 25          | 7        | 11        |
| Endocrine                | 5          | 14          | 7        | 7         |
| Gastrointestinal         | 2          | 4           | 4        | 4         |
| Hematological            | 2          | 3           | 2        | 2         |
| Immunological            | 9          | 31          | 9        | 12        |
| Metabolic                | 8          | 33          | 7        | 11        |
| Muscular                 | 11         | 37          | 14       | 17        |
| Neurological             | 25         | 55          | 27       | 36        |
| Nutritional              | 1          | 1           | 1        | 1         |
| Ophthalmological         | 9          | 25          | 11       | 21        |
| Psychiatric              | 8          | 19          | 8        | 11        |
| Renal                    | 6          | 7           | 6        | 6         |
| Skeletal                 | 1          | 5           | 1        | 2         |
| Unclassified             | 4          | 19          | 4        | 15        |
| multiple                 | 28         | 53          | 30       | 31        |

Figure 4. Web user interface, (a) on the search page, after entering a keyword, the result will be displayed and the list of the results will be displayed below; (b) On the browse page, select one group to display, the result is similar as on the search page. Both on the search and browse page, each row of the list is composed of five columns, human miRNA, fly miRNA, human gene, fly gene, gene disease group, where the human miRNA is orthologous to the fly miRNA, the human gene is orthologous to the fly gene, and the miRNA target to gene in human and fly, respectively.
3. Discussion and Conclusions

In the present study, our results show that many human disease-related genes have corresponding orthologous genes in *D. melanogaster*, and human disease-related genes can be regulated by orthologous miRNA between human and *D. melanogaster* (Figure 2). We identified the relationship of miRNA-regulating target genes that are conserved between human and *D. melanogaster*. In addition, we found that these miRNAs regulating target genes are classified into 22 disease groups, and 212 human pathways, as well as 65 fly pathways.

*Drosophila* is widely used as a feasible model organism in the study of human diseases [2,37,46]. Moreover, human miRNA was found to be involved in cancers, cardiovascular and neurological disease and so on [23,47,48]. Recent works have revealed that *D. melanogaster* is an ideal model organism to study human miRNA-related diseases, such as miR-124, which is involved in Alzheimer’s disease, and miR-219, which results in neurofibrillary degeneration [49–51]. Therefore, the importance of using *D. melanogaster* as a model to study the miRNA roles in human disease process is conspicuous. At present, the study of human disease using *D. melanogaster* as a model organism mainly is focused on the function of the disease-related genes. Although some studies have considered the regulatory role of miRNA during disease progression, large-scale analysis of the regulation relationship between miRNA and disease genes across different species is still limited [26,51]. The DHDD database provides very important resources for the investigation of the target relationship of human miRNA regulating human disease-related genes and their *D. melanogaster* orthologs. For example, hsa-miR-216a has been reported to be involved in pancreatic cancer progression, and a significant downregulation of hsa-miR-216a in human pancreatic cancer tissue has also been reported [52,53], which is consistent with our result in the DHDD database, indicating that the results from the DHDD database are reliable. In particular, in currently existing database resources, either only cross-species orthologous gene data or only human miRNA related to disease can be acquired. Therefore, our present work has remedied the limitations of these existing databases by providing valuable resources for the study of the role of miRNA in human diseases using *D. melanogaster* as a model.

In conclusion, the DHDD database can present human disease-related miRNA and orthologs in *D. melanogaster*, and a visual network can also conveniently show the regulation relationship between the miRNA and target genes, as well as providing more information resources in comparison to previous database resources.

**Author Contributions:** Conceptualization, G.W. and F.M.; Methodology, G.W.; Software, L.S.; Validation, S.Q., R.L. and P.J.; Formal Analysis, G.W. and L.S.; Investigation, L.S.; Resources, G.W.; Data Curation, S.Q.; Writing-Original Draft Preparation, L.S.; Writing-Review & Editing, G.W.; Visualization, G.W.; Supervision, L.C. and F.M.; Project Administration, F.M.; Funding Acquisition, F.M., P.J. and G.W.

**Funding:** This work was supported by the National Natural Science Youth Foundation of China (No. 31501863); National Natural Science Foundation of China (No. 31572324) and the Jiangsu Province Ordinary University Innovative Research Project (No. KLYX16_1280).

**Acknowledgments:** Thanks to Xu Xiaofeng for his suggestions on this work.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**

1. Read, R.D.; Goodfellow, P.J.; Mardis, E.R.; Novak, N.; Armstrong, J.R.; Cagan, R.L. A Drosophila model of multiple endocrine neoplasia type 2. *Genetics* 2005, 171, 1057–1081. [CrossRef] [PubMed]
2. Reiter, L.T.; Potocki, L.; Chien, S.; Gribskov, M.; Bier, E. A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 2001, 11, 1114–1125. [CrossRef] [PubMed]
3. Simon, M.A. Receptor tyrosine kinases: Specific outcomes from general signals. *Cell* 2000, 103, 13–15. [CrossRef]
4. Lee, L.A.; Orr-Weaver, T.L. Regulation of cell cycles in Drosophila development: Intrinsic and extrinsic cues. *Annu. Rev. Genet.* 2003, 37, 545–578. [CrossRef] [PubMed]
5. Karim, F.D.; Chang, H.C.; Therrien, M.; Wassarman, D.A.; Laverty, T.; Rubin, G.M. A screen for genes that function downstream of Ras1 during Drosophila eye development. Genetics 1996, 143, 315–329. [PubMed]

6. Gao, X.; Neufeld, T.P.; Pan, D. Drosophila PTEN regulates cell growth and proliferation through PI3K-dependent and-independent pathways. Dev. Biol. 2000, 221, 404–418. [CrossRef] [PubMed]

7. 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.; et al. The genome sequence of Drosophila melanogaster. Science 2000, 287, 2185–2195. [CrossRef] [PubMed]

8. Schneider, D. Using Drosophila as a model insect. Nat. Rev. Genet. 2000, 1, 218–226. [CrossRef] [PubMed]

9. Read, R.D. Drosophila melanogaster as a model system for human brain cancers. Glia 2011, 59, 1364–1376. [CrossRef] [PubMed]

10. Hirth, F. Drosophila melanogaster in the study of human neurodegeneration. CNS Neurol. Disord. Drug Targets 2010, 9, 504–523. [CrossRef] [PubMed]

11. Wolf, M.J.; Amrein, H.; Izatt, J.A.; Choma, M.A.; Reedy, M.C.; Rockman, H.A. Drosophila as a model for the identification of genes causing adult human heart disease. Proc. Natl. Acad. Sci. USA 2006, 103, 1394–1399. [CrossRef] [PubMed]

12. Tipping, M.; Perrimon, N. Drosophila as a model for context-dependent tumorigenesis. J. Cell. Physiol. 2014, 229, 27–33. [PubMed]

13. Herz, H.M.; Bergmann, A. Genetic analysis of ESCRT function in Drosophila: A tumour model for human Tsg101. Biochem. Soc. Trans. 2009, 37, 204–207. [CrossRef] [PubMed]

14. Lu, B.; Vogel, H. Drosophila Models of Neurodegenerative Diseases. Annu. Rev. Pathol. 2009, 4, 315–342. [CrossRef] [PubMed]

15. Alfa, R.W.; Kim, S.K. Using Drosophila to discover mechanisms underlying type 2 diabetes. Dis. Models Mech. 2016, 9, 365–376. [CrossRef] [PubMed]

16. Pandey, U.B.; Nichols, C.D. Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. Pharmacol. Rev. 2011, 63, 411–436. [CrossRef] [PubMed]

17. Wang, L.; Kounatidis, I.; Ligoxygakis, P. Drosophila as a model to study the role of blood cells in inflammation, innate immunity and cancer. Front. Cell. Infect. Microbiol. 2014, 3, 113. [CrossRef] [PubMed]

18. Baker, K.D.; Thummel, C.S. Diabetic larvae and obese flies-emerging studies of metabolism in Drosophila. Cell Metab. 2007, 6, 257–266. [CrossRef] [PubMed]

19. Chien, S.; Reiter, L.T.; Bier, E.; Gribskov, M. Homophila: Human disease gene cognates in Drosophila. Nucleic Acids Res. 2002, 30, 149–151. [CrossRef] [PubMed]

20. O’Brien, K.P.; Westerlund, I.; Sonnhammer, E.L. OrthoDisease: A database of human disease orthologs. Hum. Mutat. 2004, 24, 112–119. [CrossRef] [PubMed]

21. Hu, Y.; Flockhart, I.; Vinayagam, A.; Bergwitz, C.; Berger, B.; Perrimon, N.; Mohr, S.E. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinform. 2011, 12, 357. [CrossRef] [PubMed]

22. Ardekani, A.M.; Naeini, M.M. The Role of MicroRNAs in Human Diseases. Avicenna J. Med. Biotechnol. 2010, 2, 161–179. [PubMed]

23. Naeini, M.M.; Ardekani, A.M. Noncoding RNAs and Cancer. Avicenna J. Med. Biotechnol. 2009, 1, 55–70. [PubMed]

24. Zhao, W.G.; Yu, S.N.; Lu, Z.H.; Ma, Y.H.; Gu, Y.M.; Chen, J. The miR-217 microRNA functions as a potential tumor suppressor in pancreatic ductal adenocarcinoma by targeting KRAS. Carcinogenesis 2010, 31, 1726–1733. [CrossRef] [PubMed]

25. Zhao, W.; Zhao, S.P.; Zhao, Y.H. MicroRNA-143/-145 in Cardiovascular Diseases. BioMed Res. Int. 2015, 2015, 531740. [CrossRef] [PubMed]

26. Liu, N.; Landreh, M.; Cao, K.; Abe, M.; Hendriks, G.J.; Kennerdell, J.R.; Zhu, Y.; Wang, L.S.; Bonini, N.M. The microRNA miR-34 modulates ageing and neurodegeneration in Drosophila. Nature 2012, 482, 519–523. [CrossRef] [PubMed]

27. Wang, Y.; Cai, Y. A survey on database resources for microRNA-disease relationships. Brief. Funct. Genom. 2017, 16, 146–151. [CrossRef] [PubMed]

28. Jiang, Q.; Wang, Y.; Hao, Y.; Juan, L.; Teng, M.; Zhang, X.; Li, M.; Wang, G.; Liu, Y. miR2Disease: A manually curated database for microRNA deregulation in human disease. Nucleic Acids Res. 2009, 37, D98–D104. [CrossRef] [PubMed]
52. Yu, J.; Li, A.; Hong, S.M.; Hruban, R.H.; Goggins, M. MicroRNA alterations of pancreatic intraepithelial neoplasias. *Clin. Cancer Res.* **2012**, *18*, 981–992. [CrossRef] [PubMed]

53. Rachagani, S.; Macha, M.A.; Menning, M.S.; Dey, P.; Pai, P.; Smith, L.M.; Mo, Y.Y.; Batra, S.K. Changes in microRNA (miRNA) expression during pancreatic cancer development and progression in a genetically engineered KrasG12D; Pdx1-Cre mouse (KC) model. *Oncotarget* **2015**, *6*, 40295–40309. [CrossRef] [PubMed]