



## Database tool

# InSexBase: an annotated genomic resource of sex chromosomes and sex-biased genes in insects

Xi Chen<sup>†</sup>, Yang Mei<sup>†</sup>, Mengyao Chen<sup>†</sup>, Dong Jing, Yumin He, Feiling Liu, Kang He and Fei Li<sup>ID\*</sup>

Ministry of Agriculture and Rural Affairs Key Laboratory of Molecular Biology of Crop Pathogens and Insects & Key Laboratory of Biology of Crop Pathogens and Insects of Zhejiang Province, Institute of Insect Sciences, Zhejiang University, Yuhangtang Rd 866, Xihu District, Hanzhou, 310058, China

\*Corresponding author: Tel: 86-0571-88982679; Fax: 86-0571-88982679; Email: [lfei18@zju.edu.cn](mailto:lfei18@zju.edu.cn)

Citation details: Chen, X., Mei, Y., Chen, M. *et al.* InSexBase: an annotated genomic resource of sex chromosomes and sex-biased genes in insects. *Database* (2020) Vol. 2020: article ID baab001; doi:10.1093/database/baab001

<sup>†</sup>These authors contributed equally to this work.

Received 2 September 2020; Revised 9 November 2020; Accepted 6 January 2021

## Abstract

Sex determination and the regulation of sexual dimorphism are among the most fascinating topics in modern biology. As the most species-rich group of sexually reproducing organisms on Earth, insects have multiple sex determination systems. Though sex chromosomes and sex-biased genes are well-studied in dozens of insects, their gene sequences are scattered in various databases. Moreover, a shortage of annotation hinders the deep mining of these data. Here, we collected the chromosome-level sex chromosome data of 49 insect species, including 34 X chromosomes, 15 Z chromosomes, 5 W chromosomes and 2 Y chromosomes. We also obtained Y-linked contigs of four insect species—*Anopheles gambiae*, *Drosophila innubila*, *Drosophila yakuba* and *Tribolium castaneum*. The unannotated chromosome-level sex chromosomes were annotated using a standard pipeline, yielding a total of 123 030 protein-coding genes, 2 159 427 repeat sequences, 894 miRNAs, 1574 rRNAs, 5105 tRNAs, 395 snoRNAs (small nucleolar RNA), 54 snRNAs (small nuclear RNA) and 5959 other ncRNAs (non-coding RNA). In addition, 36 781 sex-biased genes were identified by analyzing 62 RNA-seq (RNA sequencing) datasets. Together with 5707 sex-biased genes from the *Drosophila* genus collected from the Sex-Associated Gene Database, we obtained a total of 42 488 sex-biased genes from 13 insect species. All these data were deposited into InSexBase, a new user-friendly database of insect sex chromosomes and sex-biased genes.

**Database URL:** <http://www.insect-genome.com/Sexdb/>.

## Introduction

As the most species-rich metazoans on Earth, insects contain the most variable mechanisms of sex determination in animals, including male heterogamety and female heterogamety (1, 2). Thus, insects form a good model system on which to study the biology of sex, the mechanisms of sex determination and how they evolve (3). In the past decades, investigations of fruit flies, mosquitos and silkworms have yielded copious knowledge in the field of sex determination (4–6). Comparative genomic analyses between insect species have increased our understanding of the evolution of sex chromosomes (7–10). Moreover, understanding insect sex biology has contributed to improvements in both pest control and resource-insect production. For example, by destroying spermatid X chromosomes (11), the sterile insect technique can be used to manipulate sex ratios in vector mosquitoes, thereby dramatically reducing or even crashing their populations (11). In silkworms, a sequence containing an egg color gene is translocated to the W chromosome by radiation treatment. This enables males and females to be easily distinguished before hatching, hence increasing the efficiency of the silk production industry (12). Furthermore, insects display a wide array of sexual dimorphism, including body size, ornamentation and coloration. So far, these have been found to be controlled by different sex-biased genes (13). Understanding how these phenotypic differences arose from sex-biased genes can reveal novel insights into species evolution (14).

Thanks to the rapid development of sequencing technologies, an increasing number of insect sex chromosomes and sex-biased genes have been reported (15–21). Sex-biased genes can be obtained by comparing RNA-seq data from females and males. Studies on sex-biased genes and their functions and uneven chromosomal distribution have drawn more attention recently (22–32). Unfortunately, these data are scattered throughout various genome databases, which often lack annotation information. The Sex-Associated Gene Database (SAGD) catalogs the sex-biased genes of many species, but only a single insect genus, *Drosophila*, is available therein (33). This provides very limited support for in-depth studies of sex biology and sex determination in insects.

To make it easier for researchers to obtain and analyze sex-related genes, we constructed the Insect Sex Chromosome and Sex-Biased Genes Database (InSexBase) by integrating data from 49 sequenced chromosome-level genomes and 62 RNA-seq datasets. InSexBase is intended to provide a comprehensive platform of chromosome-level sex-related genome and gene resources for insects. To the best of our knowledge, InSexBase collects nearly all available chromosome-level insect sex chromosome sequences

and sex-biased genes. It offers some widely used web services such as BLAST and JBrowse. Users can browse data by species, sex chromosomes, sex chromosome genes or sex-biased genes. Data grouped by species or chromosome can be downloaded easily, furnishing a valuable resource for researchers in the fields of sex evolution, pest control and public health.

## Materials and methods

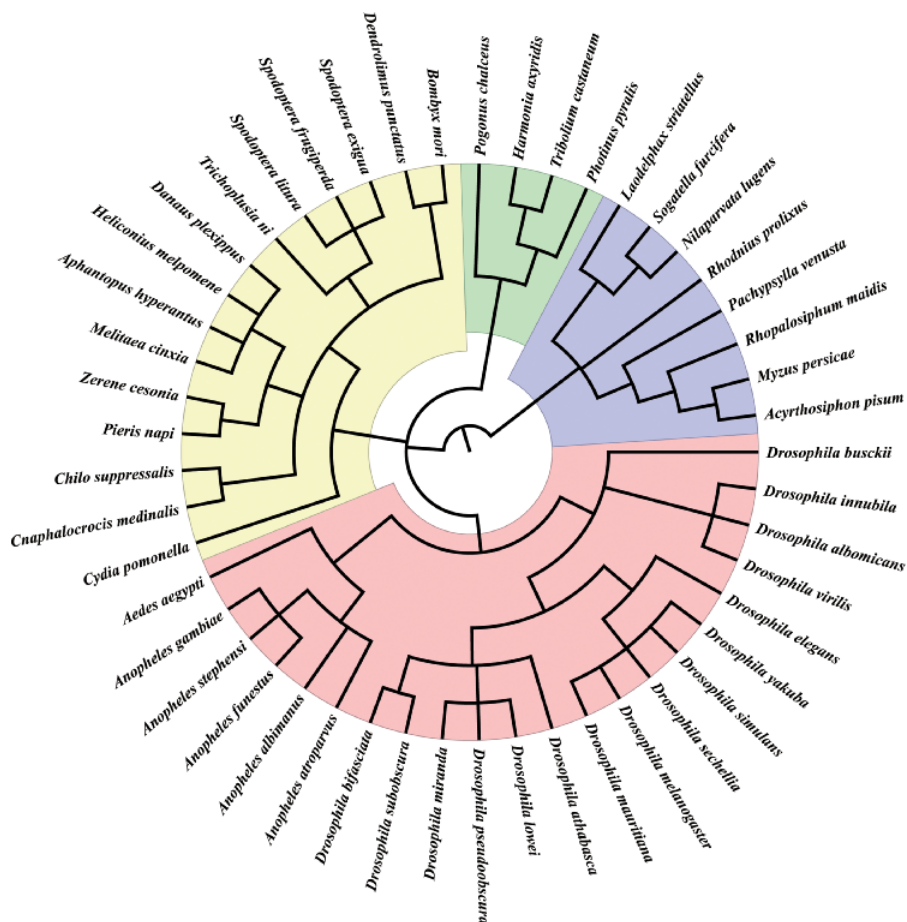
### Data collection

We collected sex chromosome sequences from 49 insect species, including 34 X chromosomes, 15 Z chromosomes, 5 W chromosomes and 2 Y chromosomes from various public databases (Supplemental Table S1). We also obtained Y-linked contigs of four insect species—*Anopheles gambiae*, *Drosophila innubila*, *Drosophila yakuba* and *Tribolium castaneum*. Further annotations of Y chromosomes did not include these contigs because they are very short (Supplemental Table S1). These species represent 28 genera in four orders (Figure 1). We also collected 62 RNA-seq datasets from NCBI and EMBL-EBL (Supplemental Table S2) and 5707 sex-biased genes from the *Drosophila* genus from SAGD (<http://bioinfo.life.hust.edu.cn/SAGD#!/>).

### Gene annotation of sex chromosomes

We obtained the gene annotation information of the genomes with available annotations (Supplemental Table S1). For those genomes without gene annotations, we carried out gene annotation with the following methods.

Repeat sequences were identified by both homology-based and *de novo* prediction methods. For *de novo* predictions, RepeatModeler v1.0.11 was used with default parameters to construct a *de novo* repeat library. For homology-based predictions, RepeatMasker v4.0.7 (34) was used with the Repbase library (35). We used GeneMark-ES (version 4.32) to annotate expected protein-coding genes identified via *de novo* prediction (36). Sequences were BLASTed against the Swiss-Prot database (37) with an *e*-value of  $1 \times 10^{-5}$ . Then, the Retrieve/ID mapping function (<https://www.uniprot.org/uploadlists/>) was used to obtain the OrthoDB ID and PDB ID. Gene Ontology analysis was carried out using Blast2GO v5.2 (38). We further mapped these genes to data from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using the BlastKOALA v2.2 (39) online service (KOALA: KEGG Orthology and Links Annotation). Noncoding RNA genes were predicted by the INFERNAL software against Rfam database, release 14.2 (40). All algorithms were run with default parameters.



**Figure 1.** Insect species in InSexBase. InSexBase contains 49 insect species from 28 genera belonging to four orders.

### Identifying sex-biased genes

RNA-seq reads were first filtered using Trimmomatic v0.38 (41). The clean reads were then mapped to the corresponding assembled genome using RSEM v1.3.0 (42) to calculate Fragments Per transcripted Kilobase per Million mapped reads (FPKM) and a read count table. Only genes with both male and female FPKM >1 were considered for further analysis. The read count tables for male and female samples were used as input for DESeq2 v1.28.1 (43). Genes were accepted as differentially expressed when DESeq2 returned a  $q$ -value <0.05 and  $|\log_2(\text{Fold Change})| > 1$ . Genes satisfying these conditions comprised the final set of sex-biased genes.

### Database implementation

The database was built with Django 3.0.5 web framework (<https://www.djangoproject.com/>), and all data were stored in a MySQL 5.7.17 database (<https://dev.mysql.com/>) on a CentOS 7.4.1708 web server. Web page templates used Bootstrap framework (<https://getbootstrap.com/>), DataTables (<http://datatables.net>), jQuery (<http://jquery.com>)

and Bootstrap Table (<https://bootstrap-table.com/>) libraries to establish a user-friendly front-end interface. On the tools page, the JBrowse Genome Browser (<http://www.jbrowse.org/>) was used to build a fast and scalable genome browser (Figure 2).

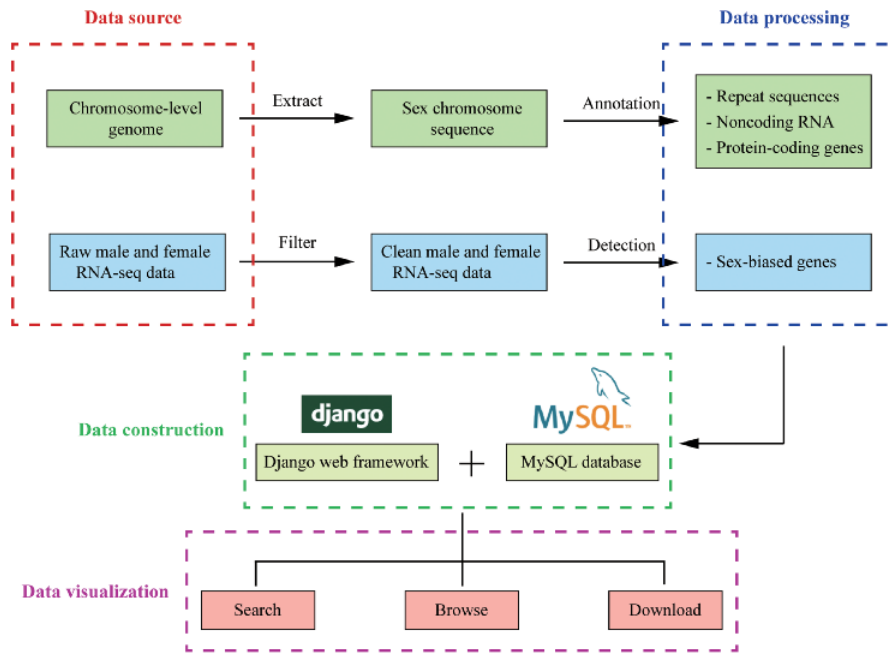
## Results and discussion

### Sex chromosomes in insects

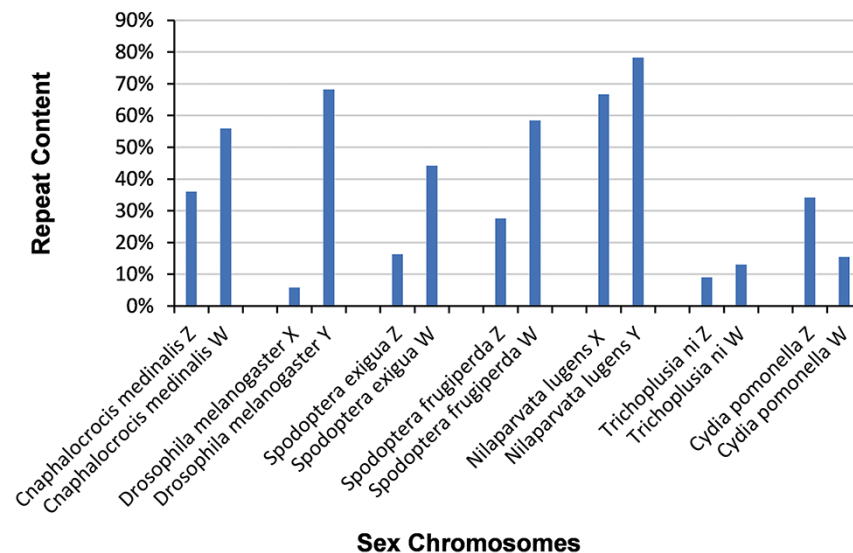
In total, we collected the sex chromosome sequences of 49 insect species from four orders, including 4 Coleoptera species, 8 Hemiptera species, 15 Lepidoptera species and 22 Diptera species (Figure 1).

Lepidopterans are female heterogametic and have both ZW and ZO systems. InSexBase contains 15 Z chromosomes and 5 W chromosomes of 15 species. Because the W chromosome is packed with repeats (44–48), it remains a difficult problem to obtain a complete W chromosome. The W chromosome sequences of these five species are useful references.

The majority of dipterans have XY sex chromosomes (49). The sex chromosomes differ between various families,



**Figure 2.** The overview of InSexBase. Genomes and sex-related RNA-seq data were collected from public databases. Sex chromosome sequences were extracted from those genome data. Repeat sequences, non-coding RNAs and protein-coding genes were annotated from sex chromosome. Sex-biased genes were detected from male and female RNA-seq clean data. Finally, these data were deposited in InSexBase built with Django 3.0.5 web framework and all data were stored in a MySQL 5.7.17 database. Users can use browse, search and download functions to explore the database.



**Figure 3.** Distribution of repeat sequences in 14 sex chromosomes. Sex chromosomes generally contain a high percentage of repeat sequences.

however, and there are numerous transitions of sex chromosomes in Diptera (50, 51). Unlike the malaria mosquito *A. gambiae* and *Drosophila* species, *Aedes aegypti* does not have heteromorphic (XY) sex chromosomes (52). The male phenotype determination is by a non-recombining M locus on one copy of autosome 1 (53). InSexBase contains 22 X

chromosomes of 22 species in Diptera and 1 Y chromosome of *D. melanogaster*. InSexBase also contains Y-linked contigs of three Diptera species, *A. gambiae*, *D. innubila* and *D. yakuba*. *Aedes aegypti* has a homomorphic sex chromosome system, and chromosome 1 is regarded as the X chromosome in InSexBase.

The majority of sexually reproducing beetles are male heterogametic (54). Unlike dipterans and lepidopterans, the base genetic mechanism of sex determination has not been identified for beetles (49). InSexBase contains four X chromosomes from four species in Coleoptera and Y-linked contigs of *T. castaneum*. Of these, *Photinus pyralis* has an XO system, and the other three have XY systems (54).

Most hemipteran insects employ the XY or XO sex determination system, while most aphids reproduce through cyclic parthenogenesis (49). InSexBase contains eight X chromosomes from eight species in Hemiptera (five aphids and three plant hoppers) and one Y chromosome from the plant hopper *Nilaparvata lugens*.

### Repeat sequences in sex chromosomes

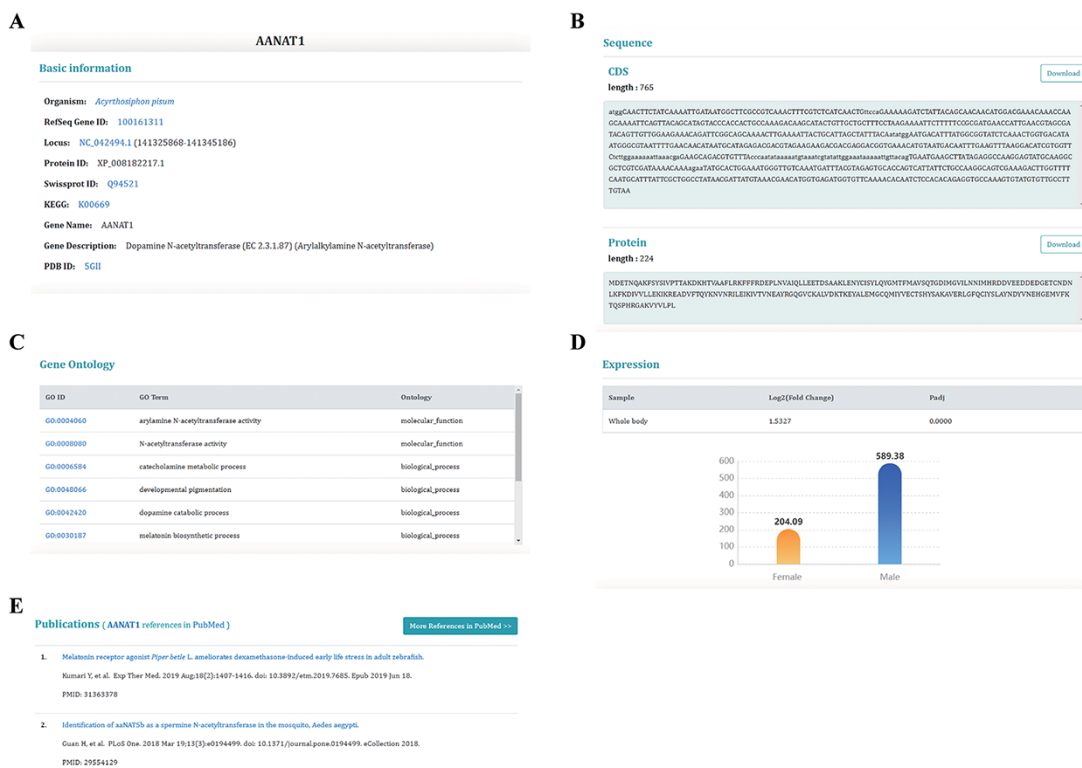
We identified repeat sequences in 55 sex chromosomes with RepeatModeler v1.0.11 and RepeatMasker v4.0.7 (34), yielding 2 159 427 repeat sequences. The evolution of sex-limited chromosomes is often accompanied by the acquisition of repetitive DNA (44–48). The occurrence of repeat sequences in the W and Y chromosomes of six species (*Cnaphalocrocis medinalis*, *D. melanogaster*, *Spodoptera exigua*, *Spodoptera frugiperda*, *N. lugens* and *Trichoplusia ni*) is much higher than that in their Z and X chromosomes (Figure 3). However, repeat sequences in the

W chromosome of *Cydia pomonella* are not particularly frequent (Figure 3).

### Protein-coding genes and non-coding RNA genes in sex chromosomes

To collect protein-coding genes and non-coding RNA genes in sex chromosomes, we obtained the gene annotation information of the genomes with available annotations. For those genomes without gene annotations, we carried out gene annotation as described in the ‘Materials and methods’ section. Consequently, a total of 123 030 protein-coding genes, 894 miRNAs, 1574 rRNAs, 5105 tRNAs, 395 snoRNAs, 54 snRNAs and 5959 other ncRNAs were predicted on 56 sex chromosomes in 49 species. The 34 X chromosomes were predicted to contain 101 060 protein-coding genes, 593 miRNAs, 1527 rRNAs, 4071 tRNAs, 389 snoRNAs, 54 snRNAs and 5773 other RNAs. The 15 Z chromosomes contain 16 799 protein-coding genes, 286 miRNAs, 26 rRNAs, 757 tRNAs, 6 snoRNAs and 59 other RNAs; 2 Y chromosomes contain 354 protein-coding genes, 11 miRNAs, 12 rRNAs and 236 tRNAs. The five W chromosomes contain 4817 protein-coding genes, 4 miRNAs, 9 rRNAs, 95 tRNAs and 14 other RNAs (more information provided in Supplemental Table S1).

In *D. melanogaster*, four X-related proteins encoded by the *scute*, *sisA*, *runt* and *unpaired* genes serve as the primary



**Figure 4.** Information on sex-biased genes. (A) Basic genetic information. (B) CDS and protein sequences. (C) The Gene Ontology information of this gene. (D) The expression of sex-biased genes. (E) Publications relating to this gene.



determinants of X dose (55–59), which plays an important role in sex determination in fruit flies. In the silkworm, a W-linked piRNA (PIWI-interacting RNA) is the primary determiner of sex (4). The protein-coding genes and non-coding RNA genes on the sex chromosomes may also be important for sex determination in non-model insects.

### Sex-biased genes in insects

We analyzed 62 RNA-seq datasets to deduce sex-biased genes with RSEM v1.3.0 (42) and DESeq2 (43), yielding 36 781 sex-biased genes. We also downloaded the 5707 identified sex-biased genes of three *Drosophila* species from SAGD. Finally, we obtained a total of 42 488 sex-biased genes from 13 insect species from seven genera in four orders. The study of sex-biased genes is important not only for understanding gene regulation and evolution but also for their application to insect reproduction and pest control (14, 60–63).

### Web interface and usage

InSexBase provides a user-friendly interface for searching, browsing and downloading data about insect sex chromosome and sex-biased genes. In the top navigation bar, there are eight units including ‘Home’, ‘Species’, ‘Dataset’, ‘Tools’, ‘Download’, ‘Help’, ‘Links’ and ‘About us’. The ‘Home’ page offers an introduction to InSexBase, a quick gene search function and some interface entrances. By clicking the left-side logo of the Webpage, users can find some insect pictures (Supplemental Table S3).

There is an unrooted phylogenetic tree in the ‘Species’ unit for users to browse species information and download data. The ‘Dataset’ unit currently includes five pages for users to browse data about sex chromosomes, sex chromosomal genes, ncRNAs, repeat information and sex-biased genes. The page dedicated to genes shows the most detailed information about genes (Figure 4). The database provides two bioinformatics tools: BLAST and JBrowse. Users can use ‘BLAST’ to search target genes and use ‘JBrowse’ to browse genome information. We provide convenient downloading options in the ‘Download’ unit to download anything from the database. The ‘Help’ unit offers detailed instructions and the ‘Links’ unit provides links to related databases.

### Future directions

1. Synteny analysis is helpful in studying the evolution of sex chromosomes. We plan to add online synteny analysis to InSexBase.
2. Due to the rapid development of sequencing technology in recent years, the genomes of an increasing number of insects will be sequenced and their

sex-related gene data will be identified. We plan to update InSexBase periodically to keep the database up-to-date.

3. We will add more expression evidence by integrating RNA-seq data into InSexBase.
4. 3D genomics, which explores the crucial role the three-dimensional configuration of the genome has in gene regulation, will be added in the future to study the genomic structure of sex chromosomes.

### Conclusion

To the best of our knowledge, InSexBase is the first database of sex-related chromosomes, genomes and gene resources in insects. By unifying and expanding insect sex information, InSexBase will be extremely useful for evolution analysis and comparative genomics in sex evolution, pest control, mosquito biology, resource-insect management and public health.

### Supplementary data

Supplementary data are available at Database Online.

### Funding

This work was funded by grants from the National Natural Science Foundation of China (31772238), the Natural Science Foundation of Zhejiang Province (LZ18C060001) and the Fundamental Research Funds for the Central Universities (2020QNA6024).

*Conflict of interest.* The authors declare no competing interests.

### References

1. Feener, D.H. Jr (1984) Evolution of sex determining mechanisms: a review of fact and theory. *Evolution*, **38**, 1163–1164.
2. Charlesworth, D. and Mank, J.E. (2010) The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. *Genetics*, **186**, 9–31.
3. Sanchez, L. (2008) Sex-determining mechanisms in insects. *Int. J. Dev. Biol.*, **52**, 837–856.
4. Kiuchi, T., Koga, H., Kawamoto, M. *et al.* (2014) A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature*, **509**, 633–636.
5. Cline, T.W., Dorsett, M., Sun, S. *et al.* (2010) Evolution of the *Drosophila* feminizing switch gene Sex-lethal. *Genetics*, **186**, 1321–1336.
6. Hall, A.B., Papathanos, P.-A., Sharma, A. *et al.* (2016) Radical remodeling of the Y chromosome in a recent radiation of malaria mosquitoes. *Proc. Nat. Acad. Sci.*, **113**, E2114–E2123.
7. Bernardo Carvalho, A., Koerich, L.B. and Clark, A.G. (2009) Origin and evolution of Y chromosomes: *Drosophila* tales. *Trends Genet.*, **25**, 270–277.
8. Pease, J.B. and Hahn, M.W. (2012) Sex chromosomes evolved from independent ancestral linkage groups in winged insects. *Mol. Biol. Evol.*, **29**, 1645–1653.

9. Zhou,Q. and Bachtrog,D. (2012) Sex-specific adaptation drives early sex chromosome evolution in *Drosophila*. *Science*, 337, 341–345.
10. Zhou,Q., Ellison,C.E., Kaiser,V.B. *et al.* (2013) The epigenome of evolving *Drosophila* neo-sex chromosomes: dosage compensation and heterochromatin formation. *PLoS Biol.*, 11, e1001711.
11. Simoni,A., Hammond,A.M., Beaghton,A.K. *et al.* (2020) A male-biased sex-distorter gene drive for the human malaria vector *Anopheles gambiae*. *Nat. Biotechnol.*, 38, 1054–1060.
12. Hashimoto,H. (1948) An X-ray mutation in the silkworm (in Japanese with English summary). *J. Sericult. Sci. Jpn.*, 16, 62–64.
13. Bonduriansky,R. (2007) The evolution of condition-dependent sexual dimorphism. *Am. Nat.*, 169, 9–19.
14. Papa,F., Windbichler,N., Waterhouse,R.M. *et al.* (2017) Rapid evolution of female-biased genes among four species of *Anopheles* malaria mosquitoes. *Genome Res.*, 27, 1536–1548.
15. Cheng,T., Wu,J., Wu,Y. *et al.* (2017) Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest. *Nat. Ecol. Evol.*, 1, 1747–1756.
16. Chen,W., Yang,X., Tetreau,G. *et al.* (2019) A high-quality chromosome-level genome assembly of a generalist herbivore, *Trichoplusia ni*. *Mol. Ecol. Resour.*, 19, 485–496.
17. Kawamoto,M., Jouraku,A., Toyoda,A. *et al.* (2019) High-quality genome assembly of the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.*, 107, 53–62.
18. Wan,F.H., Yin,C.L., Tang,R. *et al.* (2019) A chromosome-level genome assembly of *Cydia pomonella* provides insights into chemical ecology and insecticide resistance. *Nat. Commun.*, 10, 4237.
19. Ma,W., Zhao,X., Yin,C. *et al.* (2020) A chromosome-level genome assembly reveals the genetic basis of cold tolerance in a notorious rice insect pest, *Chilo suppressalis*. *Mol. Ecol. Resour.*, 20, 268–282.
20. Xiao,H., Ye,X., Xu,H. *et al.* (2020) The genetic adaptations of fall armyworm *Spodoptera frugiperda* facilitated its rapid global dispersal and invasion. *Mol. Ecol. Resour.*, 20, 1050–1068.
21. Ye,X., Yan,Z., Yang,Y. *et al.* (2020) A chromosome-level genome assembly of the parasitoid wasp *Pteromalus puparum*. *Mol. Ecol. Resour.*, 20, 1384–1402.
22. Arunkumar,K.P., Mita,K. and Nagaraju,J. (2009) The silkworm Z chromosome is enriched in testis-specific genes. *Genetics*, 182, 493–501.
23. Baker,D.A., Nolan,T., Fischer,B. *et al.* (2011) A comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector, *Anopheles gambiae*. *BMC Genom.*, 12, 296.
24. Baker,R.H. and Wilkinson,G.S. (2010) Comparative genomic hybridization (CGH) reveals a neo-X chromosome and biased gene movement in stalk-eyed flies (genus *Teleopsis*). *PLoS Genet.*, 6, e1001121.
25. Bellott,DW., Skaletsky,H., Pyntikova,T. *et al.* (2010) Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. *Nature*, 466, 612–616.
26. Hahn,M.W. and Lanzaro,G.C. (2005) Female-biased gene expression in the malaria mosquito *Anopheles gambiae*. *Curr. Biol.*, 15, R192–R193.
27. Parisi,M. *et al.* (2003) Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science*, 299, 697–700.
28. Prince,E.G., Kirkland,D. and Demuth,J.P. (2010) Hyperexpression of the X chromosome in both sexes results in extensive female bias of X-linked genes in the flour beetle. *Genome Biol. Evol.*, 2, 336–346.
29. Rinn,J.L. and Snyder,M. (2005) Sexual dimorphism in mammalian gene expression. *Trends Genet.*, 21, 298–305.
30. Storchova,R. and Divina,P. (2006) Nonrandom representation of sex-biased genes on chicken Z chromosome. *J. Mol. Evol.*, 63, 676–681.
31. Sturgill,D., Zhang,Y., Parisi,M. *et al.* (2007) Demasculinization of X chromosomes in the *Drosophila* genus. *Nature*, 450, 238–241.
32. Li,Y., Zhang,B. and Moran,N.A. (2020) The aphid X chromosome is a dangerous place for functionally important genes: diverse evolution of hemipteran genomes based on chromosome-level assemblies. *Mol. Biol. Evol.*, 37, 2357–2368.
33. Shi,M.-W., Zhang,N.-A., Shi,C.-P. *et al.* (2019) SAGD: a comprehensive sex-associated gene database from transcriptomes. *Nucleic Acids Res.*, 47, D835–D840.
34. Tarailo-Graovac,M. and Chen,N. (2009) Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics*, Chapter 4:Unit 4.10.
35. Bao,W., Kojima,K.K. and Kohany,O. (2015) Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA*, 6, 11.
36. Lomsadze,A. *et al.* (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.*, 33, 6494–6506.
37. UniProt Consortium. (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.*, 47, D506–D515.
38. Conesa,A., Gotz,S., Garcia-Gomez,J.M. *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21, 3674–3676.
39. Kanehisa,M., Sato,Y. and Morishima,K. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.*, 428, 726–731.
40. Nawrocki,E.P., Kolbe,D.L. and Eddy,S.R. (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics*, 25, 1335–1337.
41. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.
42. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.*, 12, 323.

43. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
44. Dalikova, M., Zrzava, M., Hladova, I. *et al.* (2017) New insights into the evolution of the W chromosome in Lepidoptera. *J. Hered.*, **108**, 709–719.
45. Fukova, I., Traut, W., Vitkova, M. *et al.* (2007) Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. *Chromosoma*, **116**, 135–145.
46. Sahara, K., Marec, F., Eickhoff, U. *et al.* (2003) Moth sex chromatin probed by comparative genomic hybridization (CGH). *Genome*, **46**, 339–342.
47. Bachtrog, D. (2013) Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.*, **14**, 113–124.
48. Ellegren, H. (2011) Erratum: Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nat. Rev. Genet.*, **12**, 736.
49. Blackmon, H., Ross, L. and Bachtrog, D. (2017) Sex determination, sex chromosomes, and karyotype evolution in insects. *J. Hered.*, **108**, 78–93.
50. Vicoso, B. and Bachtrog, D. (2013) Reversal of an ancient sex chromosome to an autosome in *Drosophila*. *Nature*, **499**, 332–335.
51. Vicoso, B. and Bachtrog, D. (2015) Numerous transitions of sex chromosomes in Diptera. *PLoS Biol.*, **13**, e1002078.
52. Craig, G.B. Jr, Hickey, W.A. and Vandehey, R.C. (1960) An inherited male-producing factor in *Aedes aegypti*. *Science*, **132**, 1887–1889.
53. Toups, M.A. and Hahn, M.W. (2010) Retrogenes reveal the direction of sex-chromosome evolution in mosquitoes. *Genetics*, **186**, 763–766.
54. Blackmon, H. and Demuth, J.P. (2015) Coleoptera karyotype database. *Coleopt. Bull.*, **69**, 174–175.
55. Cline, T.W. (1988) Evidence that sisterless-a and sisterless-b are two of several discrete “numerator elements” of the X/A sex determination signal in *Drosophila* that switch Sxl between two alternative stable expression states. *Genetics*, **119**, 829–862.
56. Duffy, J.B. and Gergen, J.P. (1991) The *Drosophila* segmentation gene runt acts as a position-specific numerator element necessary for the uniform expression of the sex-determining gene Sex-lethal. *Genes Dev.*, **5**, 2176–2187.
57. Jinks, T.M., Polydorides, A.D., Calhoun, G. *et al.* (2000) The JAK/STAT signaling pathway is required for the initial choice of sexual identity in *Drosophila melanogaster*. *Mol. Cell*, **5**, 581–587.
58. Sánchez, L., Granadino, B. and Torres, M. (1994) Sex determination in *Drosophila melanogaster*: X-linked genes involved in the initial step of Sex-lethal activation. *Dev. Genet.*, **15**, 251–264.
59. Sefton, L., Timmer, J.R., Zhang, Y. *et al.* (2000) An extracellular activator of the *Drosophila* JAK/STAT pathway is a sex-determination signal element. *Nature*, **405**, 970–973.
60. Ellegren, H. and Parsch, J. (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.*, **8**, 689–698.
61. Grath, S. and Parsch, J. (2016) Sex-biased gene expression. *Annu. Rev. Genet.*, **50**, 29–44.
62. Hall, A.B., Basu, S., Jiang, X.F. *et al.* (2015) A male-determining factor in the mosquito *Aedes aegypti*. *Science*, **348**, 1268–1270.
63. Parsch, J. and Ellegren, H. (2013) The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.*, **14**, 83–87.