Genome-wide identification and developmental expression profiling of long noncoding RNAs during Drosophila metamorphosis

Bing Chen1,*, Yi Zhang1,2,*, Xia Zhang1,4, Shili Jia2, Shuang Chen3 & Le Kang1,4

An increasing number of long noncoding RNAs (lncRNAs) have been discovered with the recent advances in RNA-sequencing technologies. lncRNAs play key roles across diverse biological processes, and are involved in developmental regulation. However, knowledge about how the genome-wide expression of lncRNAs is developmentally regulated is still limited. We here performed a whole-genome identification of lncRNAs followed by a global expression profiling of these lncRNAs during development in Drosophila melanogaster. We combined bioinformatic prediction of lncRNAs with stringent filtering of protein-coding transcripts and experimental validation to define a high-confidence set of Drosophila lncRNAs. We identified 1,077 lncRNAs in the given transcriptomes that contain 43,967 transcripts; among these, 646 lncRNAs are novel. In vivo expression profiling of these lncRNAs in 27 developmental processes revealed that the expression of lncRNAs is highly temporally restricted relative to that of protein-coding genes. Remarkably, 21% and 42% lncRNAs were significantly upregulated at late embryonic and larval stage, the critical time for developmental transition. The results highlight the developmental specificity of lncRNA expression, and reflect the regulatory significance of a large subclass of lncRNAs for the onset of metamorphosis. The systematic annotation and expression analysis of lncRNAs during Drosophila development form the foundation for future functional exploration.

The rapid accumulation of data from tilling arrays and high-throughput sequencing indicates that most of a eukaryotic genome is actively transcribed. Many of these transcribed RNAs are untranslatable into proteins and are called noncoding RNAs (ncRNAs). More than half of the mammalian transcriptome is comprised of ncRNAs1. These transcripts consist of small ncRNAs (including microRNAs and piRNAs) and long ncRNAs (lncRNA, >200 nucleotides). There is increasing in vitro and in vivo evidence that lncRNAs play key roles across diverse biological processes2–5, although functions of lncRNAs are largely unknown6,7. For example, lncRNAs are involved in dosage compensation, genomic imprinting, and regulation of epigenetic marks and gene expression8–10. Many lncRNAs are associated with human disease11–13. Therefore, the discovery and annotation of lncRNAs have attracted increasing attentions in recent years.

Intriguing studies on individual lncRNAs have revealed distinct functions for different lncRNA families. One of the major functions mediated by lncRNAs is related to developmental regulation for pluripotency, differentiation, specialization and homeostasis14,15. Several findings that support the biological significance of lncRNAs demonstrate that many lncRNAs exhibit tissue- or cell type-specific expression16–18. Some lncRNAs exhibit expression at specific developmental time point19, and thus can be only discovered during narrow developmental time windows. These findings suggest that the complex task involved in developmental regulation requires the fine timing, space, and rate of expression of specific lncRNAs2. However, knowledge about how the genome-wide

1 State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. 2 Department of Mathematics, Hebei University of Science and Technology/Hebei Laboratory of Pharmaceutical Molecular Chemistry, Shijiazhuang 050018, China. 3 Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China. 4 Beijing Institutes of Life Sciences, Chinese Academy of Sciences, Beijing 100101, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to L.K. (email: lkang@ioz.ac.cn)
expression of lncRNAs is developmentally regulated is still limited. A systematic profiling of the temporal or spatial expression of lncRNAs during development undoubtedly will facilitate such understanding.

*D. melanogaster* is a cheaper and more amenable genetic model organism for the functional study of lncRNAs compared with mammal models. *Drosophila* also has many benefits for the evolutionary and experimental investigation of IncRNA loci. Many *Drosophila* IncRNAs have been predicted and verified the predicted IncRNAs with bioinformatic and experimental approaches. The expression profiles of IncRNAs in 27 developmental processes in *D. melanogaster* were analyzed. The results showed high developmental specificity of lncRNA expression and extensive involvement of a large subclass of IncRNAs during *Drosophila* metamorphosis. The systematic annotation and expression profiling of *Drosophila* IncRNAs highlight the fundamental importance of lncRNA enrichment for developmental regulation and lay the foundation for future functional and evolutionary studies.

### Results

**Drosophila** IncRNA identification and validation. We identified putative IncRNAs from *D. melanogaster* transcriptomes deposited in modENCODE. The database contains 43,967 transcripts from both poly(A) and total RNAs that are expressed at 27 distinct developmental stages from embryo to adult. We first filtered the transcriptomes of transcripts that are shorter than 200 bp in total length or predicted to harbor maximum open reading frames (ORFs) encoding longer than 100 amino acids (Fig. 1a). Protein coding potential of transcripts was evaluated by TransDecoder (http://transdecoder.github.io). We then used the software CPAT to predict IncRNA, which is expected to yield more than 96% sensitivity and specificity. As a result, we identified 1,077 putative IncRNAs (>200 base pair, or bp) in *D. melanogaster*. We last evaluated the false positive rate of prediction based on protein annotation in Flybase (http://flybase.org/). Blast against the *Drosophila* protein-coding transcript sequences (CDS) showed that 9% of the putative IncRNAs, i.e., 101 transcripts, were annotated as CDS in Flybase, even though the coding proteins of most (62%) of the CDS have not been reported. Therefore, only the remaining 976 IncRNAs were accepted and used for the following examination (Supplementary data 1). The predicted IncRNAs are distinct from CDS in coding probability (Supplementary Fig. S1). Among these IncRNAs, 330 transcripts have been annotated as IncRNAs in modENCODE, Flybase, or previous publications consisting of 34% of all IncRNAs predicted. The 646 *Drosophila* IncRNAs are therefore putatively novel.

The predicted IncRNAs had an average length of 621 bp. Most (87%) of the IncRNAs were less than 1,000 bp long (Fig. 1b and Supplementary data 1). The averaged size of CDS is 11,254 bp. The size of IncRNAs is much shorter than that of CDS. The averaged size of the 330 IncRNAs is 3,490 bp, and that of the 646 novel IncRNAs is 1,099 bp (Supplementary data 1). The 330 IncRNAs are longer than the 646 IncRNAs mainly because a large proportion (47%) of the novel IncRNAs range between 200–500 bp; in contrast, 39% of the annotated IncRNAs range between 500–1,000 bp, and 19% are longer than 3,000 bp (Supplementary Fig. S2a). The CDS have 6.1 exons in average. The 646 IncRNAs have 1.4 exons while the 330 IncRNAs have 2.2 exons (Supplementary data 1). The averaged size of introns of CDS is 8,270 bp. The size of introns of the 330 IncRNAs is 2,910 bp, and that of the 646 IncRNAs is 501 bp (Supplementary Fig. S2b). The results indicate that we have identified many previously annotated IncRNAs as well as a subset of novel IncRNAs that are distinct in gene structure in *D. melanogaster*.

To verify that these IncRNAs were transcribed RNA but not artificial products or amplified from genomic DNA or contaminants, 20 IncRNAs, including 12 intergenic IncRNAs, were randomly selected from the novel IncRNAs and experimentally validated by RT-PCR. Among the 20 IncRNAs tested, 19 yielded PCR products of the right size and correct sequence from their cDNA templates (Fig. 1c). The 12 intergenic IncRNAs were randomly selected from the 19 IncRNAs and further tested by RNA dot hybridization. Eleven IncRNAs exhibited clear RNA blot with antisense probe but not with sense probe, i.e., the negative control (Fig. 1d). The genomic structures of the 12 IncRNAs were annotated based on Flybase (Supplementary Fig. S3). Detailed analysis on one intergenic IncRNA, i.e., chr2R_135227_135517, revealed its prominent expression in many developmental stages, even though the expression level was low compared with that of the neighboring coding genes (Fig. 1e).

**Developmental specificity of IncRNA expression in *D. melanogaster***. We examined the possible functions of the 976 IncRNAs in developmental regulation by analyzing their developmental profiling and specificity of expression at different stages. We investigated the pattern of gene expression changes with embryonic development at 0 h to 24 h and then compared the expression profiles between IncRNAs and protein-coding loci (Fig. 2a). Some IncRNAs were highly expressed in one of the 12 time points of embryonic development. For example, 133 IncRNAs consisting of 20.7% of the total expressed IncRNAs were specifically highly expressed at the 22h to 24 h embryonic stage, the critical time point close to the metamorphosis from embryo to larva. This percentage of IncRNAs was significantly higher than that (5.2%) of protein genes highly expressed at this specific stage (Fisher’s exact test, $P < 3.6 \times 10^{-13}$). The majority (69%) of expressed IncRNAs with 454 transcripts demonstrated a more developmental stage-specific expression, relative to 39% of the protein-coding RNAs (Fisher’s exact test, $P < 3.4 \times 10^{-14}$) (Fig. 2a). The exclusion of 941 housekeeping gene transcripts (Supplementary data 2) from CDS didn’t alter the proportion (i.e., 38.6%) of CDS exhibiting developmental-specific high expression (Supplementary Fig. S4).
To further test whether or not lncRNAs have a more temporally restricted expression than coding genes, we introduced a measure of expression level divergence over developmental time, i.e., a Shannon entropy-based specificity score per locus (JS)\textsuperscript{20}. The JS value ranged from 0 to 1, and the expression pattern with JS = 1 indicated perfect specificity. To rule out that the different specificities resulted from the low expression level of lncRNAs,
we compared the specificity values of lncRNAs at different expression levels. The developmental specificities of lncRNA expression at different levels were compared in two ranges of maximal JS specificity score. The JS
The difference between lowly and moderately expressed lncRNAs was significant at $J_S < 0.9$ (Mann–Whitney $U$ Test, $P < 2.3 \times 10^{-12}$). The difference between moderately and highly expressed lncRNAs was also significant ($P < 9.6 \times 10^{-15}$). The proportion of lncRNAs with extreme specificity, with $J_S$ in the range of 0.9 to 1.0, was also compared. Approximately 26.5% lowly expressed lncRNAs, 3.8% moderately expressed lncRNAs, and 1.1% highly expressed lncRNAs exhibited extreme temporal-restricted expression. A significant difference in the proportion was detected between lowly and moderately expressed lncRNAs ($P < 6.5 \times 10^{-12}$) but not between moderately and highly expressed lncRNAs ($P = 0.071$) (Fig. 2b). The results suggest that the lowly expressed lncRNAs significantly contribute to the extreme $J_S$ value. Therefore, to compare the expression specificities of lncRNAs and coding RNAs, we excluded the lowly expressed lncRNA, and only compared the coding RNAs and 331 highly expressed lncRNAs in embryos. A significant difference in developmental specificity was observed between the coding genes and highly expressed lncRNAs (Mann–Whitney $U$ Test, $P = 0.039$) (Fig. 2c).

The developmental specificity of lncRNA expression was also examined at the larval stage (Fig. 3). The majority (79%) of the lncRNAs had high stage-specific expression. By contrast, half (50%) coding genes were specifically highly expressed ($P < 3.9 \times 10^{-14}$). The percentage of highly expressed genes (15,585 out of 31,212) remained unchanged after the housekeeping genes were excluded for the analysis (Supplementary Fig. S4). Approximately 42% lncRNAs were specifically highly expressed at the last larval stage, the critical time for the transformation from larval to prepupal stage (Fig. 3a). The maximal $J_S$ specificity of the coding genes was obtained at 0.36, whereas that of the lncRNAs was obtained at 0.5. Up to 11% of the lncRNAs reached an extreme specificity, which was significantly higher than that (4%) of the protein-coding genes ($P < 4.1 \times 10^{-11}$) (Fig. 3b). The developmental specificity of the lncRNA expression observed at the pupal stage was not as prominent as that at the embryonic and larval stages (Supplementary Fig. S5). No significant difference in the maximum $J_S$ specificity was detected between the lncRNAs and coding RNAs (Supplementary Fig. S5).
Discussion
An increasing number of ncRNAs have been discovered with the deeper-sequencing of transcriptomes and the accumulation of transcript data. Thousands of ncRNA loci (both small and long RNAs) have been suggested to exist in the *D. melanogaster* genome. In the present study, we combined high-accuracy prediction of IncRNAs with a stringent filtering of putative protein-coding transcripts to define a high-confidence set of IncRNAs. As IncRNAs typically present a cell type-specific expression as well as a lower expression than protein-coding genes, we searched IncRNAs in a rich source of transcriptomes that are originated from 27 developmental stages. We identified 1,077 IncRNAs from 43,967 transcripts, among which 646 IncRNAs have not been previously reported. These IncRNAs were missed from previous annotation mainly because different methods for ncRNA annotation had been used. For example, we used CPAT program for ncRNA prediction, which relies on a logistic regression model built with sequence features. Many novel IncRNAs are short and located in intergenic region. However, previous methods identified ncRNA mainly based on the annotation of gene structure/gene model in flybase. Experimental validations further confirm the reliability of the identified IncRNAs. These IncRNAs provide a high-quality resource of IncRNAs for future functional and evolutionary studies. Nevertheless, more IncRNAs in *D. melanogaster* are expected to be discovered with deeper-sequencing in different tissues and in individuals exposed to different environmental conditions.

Analysis of the developmental in vivo expression profile of *Drosophila* IncRNAs highlighted the high developmental specificity of IncRNA expression. The expression profiling of the 976 IncRNAs showed that IncRNAs were expressed in a much narrower time window than protein-coding genes. Thus, the former was more highly temporally restricted than the latter. The developmental specificity of expression was demonstrated at the embryonic, larval, and pupal stages examined. Recent results prove that IncRNAs have important functions in cell fate decision, differentiation, migration and signaling. Our findings suggest that IncRNAs could be involved in regulation of the timing of developmental transition.

An interesting finding revealed in the study is that a large subclass of IncRNAs are prominently upregulated at the developmental stage of metamorphosis. Twenty-one percent of IncRNAs were highly upregulated at 22–24 h of embryos (Fig. 2) and 42% of IncRNAs were greatly upregulated at the puff stage 7–9 (i.e., L6) of larvae (Fig. 3). The massive upregulation of IncRNAs coincides with the onset of metamorphosis from embryos to larvae and from larvae to prepupae. A similar phenomenon was reported in the developmental expression of IncRNAs in zebrafish. The results indicate that IncRNA enrichment in development could be important for the transformation and organogenesis in animals, even though further experiments are needed to demonstrate it.

Numerous studies have provided evidence that IncRNAs are crucial players in cell differentiation and development. For example, IncRNA *Fendrr* presents lateral mesoderm-specific expression which is critical for proper wall development. IncRNAs also exhibit conserved functions in animal embryonic development. Some IncRNAs are involved in processes directing embryonic stem cell pluripotency and differentiation. Interestingly, expression of many IncRNAs is regulated in response to hormone. The coordinated high expression of numerous IncRNAs in late embryos and larvae are highly in synchrony with the high titer pulses of the steroid hormone ecdysone in *Drosophila*. Temporal boundaries in the *Drosophila* life cycle is defined by ecdysone, whose titrate pulses trigger the major postembryonic developmental transitions, including molting and metamorphosis. Ecdysone exerts its effects through a regulatory cascade of ecdysone - ecdysone receptors (e.g., EcR and USP) - target genes. Some IncRNAs act as coactivator or repressor of hormone receptors and respond to hormone titration. The upregulation of IncRNAs could be responding to a temporal signal that occurs in association with or in parallel with the ecdysone - EcR signaling pathway. Therefore, IncRNAs may mediate the timing of embryo-to-pupae cell fate decisions by interacting with ecdysone and temporal protein regulators of differentiation.

Methods

*D. melanogaster* transcriptomes and transcript expression. The transcriptome data of *D. melanogaster* were obtained from modENCODE (http://www.modencode.org). The transcriptomes have an annotation of 43,967 transcripts that are pooled from RNA expressed at 27 distinct developmental stages from embryo to adult. These stages include 12 embryonic stages divided at 2 h intervals for 24 h, 6 larval stages, 6 pupal stages, and 3 male and 3 female adult stages. The transcriptomes of each stage were generated from poly(A) RNAs, but embryonic transcriptomes were generated from both poly(A) RNAs and total RNAs. A variety of methods and sequencing data, such as complementary tiling microarray and RNA-Seq, long read sequencing, and genome mapping, were used for the prediction of novel splice junctions and the mapping and annotation of transcripts and genes. RNAs from three biological replicates of each sample were prepared for tiling arrays. Here 'transcript' refers to a RNA molecule that transcribes from a genomic loci whereas 'gene' refers to one or more transcripts that share exons. The expression data from poly(A) RNAs were used in this study.

*Drosophila* IncRNA identification. *Drosophila* IncRNAs were predicted from the 43,967 transcripts. Open reading frames (ORFs) in the transcripts were evaluated using the program TransDecoder (http://transdecoder.github.io). Transcripts possessing a maximum ORF encoding >100 amino acids were filtered. A state-of-the-art and popular software CPAT was used to predict ncRNAs in the transcripts. CPAT is an alignment-free method and uses a logistic regression model built with four sequence features including ORF size, ORF coverage, Fickett TESTCODE statistic and hexamer usage bias. CPAT has been optimized for prediction of *Drosophila* IncRNA with coding probability cutoff at 0.39 (i.e., coding probability <0.39 indicates ncRNA, also see: http://rna-cpat.sourceforge.net). Prediction with CPAT yields a high sensitivity (0.96) and high specificity (0.97) in theory. The predicted transcripts were finally blasted against CDS annotated in Flybase (built 5.54) to evaluate false positive rate of the prediction. The ncRNAs that were filtered of all CDS and longer than 200 bp genes were labeled as putative *Drosophila* IncRNAs for further experimental validation and expression analysis.
Validation of *Drosophila* lncRNAs by RT-PCR and dot-blot hybridization. Twenty intergenic lncRNAs that were randomly selected from the predicted lncRNAs were validated by RT-PCR and RNA dot hybridization (Supplementary Table S1). Total RNA was isolated from whole body tissues of 5-d-old female adults of *D. melanogaster* (strain OreR) and was treated with Dnase I. Reverse transcription was then performed using Oligo-d(T) following the Omniscript RT kit instruction (QUIAGEN, Cat. no. 205110). We first used RT-PCR to validate the candidate lncRNAs. The amplification conditions were as follows: 94 °C for 1 min and 40 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s. Thirty cycles of amplification were used for the amplification. The RT-PCR products were purified and were sent for direct sequencing to verify their sequences. Positive and negative controls were also subjected to RT-PCR. The positive control contained genomic DNA as the template for PCR to test the validity of the designed primers and the PCR system used. The negative control had no reverse transcriptase added for reverse transcription. Expression of 12 lncRNAs were further validated by RNA hybridization. These lncRNAs except one, i.e. chr2L_21798044_21799854, are single exoned. High specific RNA probes were first synthesized using Biotin RNA labeling mix (Roche, Cat. no. 11685597910) and T7 RNA polymerase (Promega, Cat. no. p2075). Total 10 μg RNA were hybridized with the probes on BrightStar-Plus nylon membrane (Ambion, Cat. no. AM10102). After blocking and wash with Chemiluminescent Nucleic Acid detection Module (Thermal Scientific, Cat. no. 89880), the biotin-labeled RNA was detected using CCD imaging station system (Carestream Health, USA). Primers for hybridization probes were the same as those used for the RT-PCR. Details on the selected ncRNAs and PCR primers can be found in Supplementary Table S1. Note that all these primers were designed to amplify exonic region.

Developmental specificity of lncRNA expression. FPKM expression were normalized to obtain the relative expression levels over the course of development and then to visualize developmental expression profiles via heatmaps. The lncRNAs and protein-coding loci were separately clustered using Cluster 3.0 software using k-means with a city-block distance matrix. The transcripts were clustered into 10 groups at the embryonic profiles via heatmaps. The lncRNAs and protein-coding loci were separately clustered using Cluster 3.0 software.

Developmental specificity of lncRNA expression. FPKM expression were normalized to obtain the relative expression levels over the course of development and then to visualize developmental expression profiles via heatmaps. The lncRNAs and protein-coding loci were separately clustered using Cluster 3.0 software using k-means with a city-block distance matrix. The transcripts were clustered into 10 groups at the embryonic stage, 8 groups at the larval stage, and 9 groups at the pupal stage to achieve distinct clusters. The 976 lncRNAs identified in this study and the 33,074 protein-coding transcripts were analyzed for the heatmap and developmental specificity of expression. The protein-coding transcripts were further filtered to exclude housekeeping genes for the analysis of developmental specificity. By comparing with human housekeeping genes, a total of 941 orthologous transcripts from 172 genes were taken as *Drosophila* housekeeping genes (Supplemental data 2).

The developmental specificity of the expression of a transcript was evaluated following the measure used by Trapnell et al. and Cabili et al. The entropy-based specificity measure quantifies the similarity between the expression pattern of an ncRNA or coding transcript and another predefined transcript expressed at only one stage. The developmental specificity is defined as the maximal specificity score of the transcript expression pattern across all developmental stages, i.e., maximal JS specificity score. In the present study, we examined the specificity measures of transcripts expressed at the embryonic, larval, and pupal stages.

Statistical analyses. Fisher’s exact test was performed to compare the difference in the number of genes. All lncRNAs were sorted based on their averaged expression level in different periods of a developmental stage, and categorized according to their ranking of expression. The difference in JS values between gene sets was analyzed by Mann-Whitney U Test.

References

1. Carninci, P. & Hayashizaki, Y. Noncoding RNA transcription beyond annotated genes. *Curr Opin Genet Dev* **17**, 139–144, doi: 10.1016/j.gde.2007.02.008 (2007).
2. Pauli, A., Rinn, J. L. & Schier, A. F. Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet* **12**, 136–149, doi: 10.1038/nrg2904 (2011).
3. Batista, P. J. & Chang, H. Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298–1307, doi: 10.1016/j.cell.2013.02.012 (2013).
4. Pandey, R. R. & Kanduri, C. Transcriptonal and posttranscriptional programming by long noncoding RNAs. *Prog Mol Subcell Biol* **51**, 1–27, doi: 10.1007/978-3-642-16502-3_1 (2011).
5. Wang, K. & Chang, H. Molecular Mechanisms of long noncoding RNAs. *Mol cell* **43**, 904–914 (2011).
6. Fatica, A. & Bozzoni, I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet* **15**, 7–21, doi: 10.1038/nrg3606 (2014).
7. Mattick, J. S. & Rinn, J. L. Discovery and annotation of long noncoding RNAs. *Nat Struct Mol Biol* **22**, 5–7, doi: 10.1038/nsmb.2942 (2015).
8. Guttmann, M. & Rinn, J. L. Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339–346, doi: 10.1038/nature10887 (2012).
9. Heo, J. B. & Sung, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* **331**, 76–79, doi: 10.1126/science.1197349 (2011).
10. Lee, J. T. & Bartolomei, M. S. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* **152**, 1308–1323, doi: 10.1016/j.cell.2013.02.016 (2013).
11. Esteller, M. Non-coding RNAs in human disease. *Nat Rev Genet* **12**, 861–874, doi: 10.1038/nrg3074 (2011).
12. Ng, S. Y., Lin, L., Soh, B. S. & Stanton, L. W. Long noncoding RNAs in development and disease of the central nervous system. *Trends Genet* , doi: 10.1016/j.tig.2013.03.002 (2013).
13. Qureshi, I. A. & Mehler, M. F. Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat Rev Neurosci* **13**, 528–541 (2012).
14. Guttmann, M. et al. IncRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**, 295–300, doi: 10.1038/nature10398 (2011).
15. Dinger, M. E. et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* **18**, 1435–1445, doi: 10.1101/gr.073878.108 (2008).
16. Cabili, M. N. et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* **25**, 1915–1927, doi: 10.1101/gad.1746611 (2011).
17. Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F. & Mattick, J. S. Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci USA* **105**, 716–721, doi: 10.1073/pnas.0706729105 (2008).
users will need to obtain permission from the license holder to reproduce the material. To view a copy of this unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, and 11171088) and the Province Natural Science Foundation of Hebei (A2015208108).

Acknowledgements
We thank the modENCODE consortium, particularly Brenton Graveley et al. for sharing the developmental transcriptome data of Drosophila. We thank Ji-Long Liu for providing Drosophila IncRNA sequences previously published. The work was supported by the National Science Foundation of China (grant nos 31472048, 31430023, and 11171088) and the Province Natural Science Foundation of Hebei (A2015208108).

Author Contributions
B.C. and L.K. conceived the study. B.C., Y.Z., S.-L.J. and S.C. analyzed the data. B.C. and X.Z. performed the

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Chen, B. et al. Genome-wide identification and developmental expression profiling of long noncoding RNAs during Drosophila metamorphosis. Sci. Rep. 6, 23330; doi: 10.1038/srep23330 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/