Transcriptome Analysis of the Oriental Fruit Fly *Bactrocera dorsalis* Early Embryos

Wei Peng [1,2], Shuning Yu [1], Alfred M. Handler [3] and Hongyu Zhang [1,*]

[1]   Key Laboratory of Horticultural Plant Biology (MOE), State Key Laboratory of Agricultural Microbiology,
      China-Australia Joint Research Centre for Horticultural and Urban Pests,
      College of Plant Science and Technology, Huazhong Agricultural University,
      Wuhan 430070, China; pengwei@cjlu.edu.cn (W.P.); shuning_yu@yeah.net (S.Y.)
[2]   College of Life Sciences, China Jiliang University, Hangzhou 310018, China
[3]   USDA/ARS, Center for Medical, Agricultural and Veterinary Entomology, 1700 SW 23rd Drive,
      Gainesville, FL 32608, USA; al.handler@usda.gov
[*]   Correspondence: hongyu.zhang@mail.hzau.edu.cn

check for updates

Abstract: The oriental fruit fly, *Bactrocera dorsalis* (Hendel), is one of the most devastating and highly invasive agricultural pests world-wide, resulting in severe economic loss. Thus, it is of great interest to understand the transcriptional changes that occur during the activation of its zygotic genome at the early stages of embryonic development, especially the expression of genes involved in sex determination and the cellularization processes. In this study, we applied Illumina sequencing to identify *B. dorsalis* sex determination genes and early zygotic genes by analyzing transcripts from three early embryonic stages at 0–1, 2–4, and 5–8 h post-oviposition, which include the initiation of sex determination and cellularization. These tests generated 13,489 unigenes with an average length of 2185 bp. In total, 1683, 3201 and 3134 unigenes had significant changes in expression levels at times after oviposition including at 2–4 h versus 0–1 h, 5–8 h versus 0–1 h, and 5–8 h versus 2–4 h, respectively. Clusters of gene orthology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations were performed throughout embryonic development to better understand the functions of differentially expressed unigenes. We observed that the RNA binding and spliceosome pathways were highly enriched and overrepresented during the early stage of embryogenesis. Additionally, transcripts for 21 sex-determination and three cellularization genes were identified, and expression pattern analysis revealed that the majority of these genes were highly expressed during embryogenesis. This study is the first assembly performed for *B. dorsalis* based on Illumina next-generation sequencing technology during embryogenesis. Our data should contribute significantly to the fundamental understanding of sex determination and early embryogenesis in tephritid fruit flies, and provide gene promoter and effector gene candidates for transgenic pest-management strategies for these economically important species.

Keywords: *Bactrocera dorsalis*; embryogenesis; sex-determination; embryonic cellularization; gene expression

1. Introduction

The oriental fruit fly, *Bactrocera dorsalis* (Hendel), is one of the most devastating and highly invasive agricultural pests in the world that causes severe economic loss due to damage to over 250 types of fruits and vegetables throughout Southeast Asia and several Pacific Islands [1,2]. As *B. dorsalis* has the characteristics of polyphagy, high reproductive capacity, and adaptability, it is a dreaded invasive species with great potential to invade new geographic areas and host plants [1]. Chemical insecticides

are currently considered to be the most effective tool to control fruit flies; however, resistance to commonly applied chemical insecticides in *B. dorsalis* has been increasing [3,4], and it is, thus, urgent to develop new environmentally acceptable methods to control this pest and related tephritid pests.

The primary biological method for tephritid fruit fly control has been the sterile insect techniques (SIT); however, sterile males generated by irradiation or chemosterilants in traditional methods have decreased mating performance, which ultimately reduces population control efficiency [5]. However, genetically-enhanced SIT, that include transgenic genetic male sterility and sexing strains, have shown significant potential for fruit fly management [6–12]. Based on the identification of key sex-determining genes in insects, it has been possible to obtain male-only progeny by knocking down the female determining genes *transformer* (*tra*) and *transformer-2* (*tra-2*) for male-only populations [13–19]. In addition, biotechnology-based control strategies have been tested for the survival of male-only populations for sterile release. For example, female-specific dominant lethal constructs use the alternative sex-specific splicing of sex-determining genes for the female-specific expression of a lethal effector gene [20–23].

The early stages of insect embryonic development include a complex interaction between maternal and zygotic genetic information, in which maternal transcripts degrade while the zygotic genome is activated [24]. In *Drosophila melanogaster*, the degradation of maternal mRNA is regulated by the activity of the Pumilio, Smaug proteins and miRNAs in the developing embryo [25–27], while the initial activation of zygotic gene transcription is controlled by the zinc-finger protein Zelda (*zelda*, *zld*) which binds to TAG team sites in the promoter regions of downstream genes, activating their expression [28,29]. Such TAG team promoter sites have been found in sex-determination genes such as *sisterless A* (*sisA*) and cellular blastoderm formation genes such as *serendipity a* (*sry-a*), *nullo*, *bottleneck* (*bnk*) and *slow as molasses* (*slam*) [30–34].

The sex-determination cascade genes are widely conserved in dipteran insects—most prominently, the terminal downstream sex-determining gene, *doublesex* (*dsx*) and its upstream regulatory genes *tra* and *tra-2* [35–38]. In *D. melanogaster*, a combination of X-chromosome linked signal elements (XSE), namely *sisA*, *scute* (*sc*), *outstretched* (*os*) and *runt* as the primary signal, initiate the female-specific embryonic splicing of *Sxl* that results in the continuing production of functional SXL protein in females [39–41]. The sex-specific splicing of *tra* in females is also regulated by SXL, resulting in fully functional TRA protein, while the absence of SXL in males results in truncated non-functional TRA proteins [42,43]. In contrast to in *Drosophila*, *Sxl* transcripts have not been found to be sex-specifically spliced in other dipteran insects in the Tephritidae, Muscidae and Calliphoridae and thus have no apparent role in determining somatic and germline sexual differentiation [13,44–46]. In those insects, a Y chromosome-linked male determining factor (M-factor) is the primary determinant of sex differentiation, in which the paternally inherited M factor prevents the maternal activation of the zygotic *tra* positive self-regulatory loop, resulting in the successive male-specific mode of *tra* and *dsx* splicing [47–51].

In the current study, we used Illumina sequencing technology to analyze the early embryonic transcriptome of *B. dorsalis* (0–1, 2–4 h, and 5–8 h post-oviposition) during the key stages for the initiation of sex determination during embryogenesis. Our results, for the first time, provide a broader insight into the expression profiles of sex-determination genes in the *B. dorsalis* early embryo transcriptome. We also present a preliminary scenario of early zygotic gene expression during the maternal-to-zygotic transition (MTZ) in this non-model organism. The early embryonic transcriptome and genome data should assist in the identification of the primary Y-linked M-factor signal and in further elucidating the genetic regulation of sex determination in *B. dorsalis*.

2. Materials and Methods

2.1. Insect Rearing and Embryo Collection

B. dorsalis were reared at the Institute of Horticultural and Urban Entomology at Huazhong Agricultural University (Wuhan, China). Larvae were maintained on a banana diet while adult flies were fed on an artificial diet consisting of yeast extract and sugar. All life stages were cultured at 28 °C under a 12 h light/12 h dark photoperiod in cages [52]. Embryos were collected from females by inducing egg laying through a perforated paper-cup coated with orange juice for a period of 15 min (first egg batch discarded) and then placed on moist filter paper and maintained at 25 °C and 70% humidity for set time intervals.

2.2. RNA Extraction and Library Preparation for Transcriptome Sequencing

Embryos were collected at 0–1, 2–4, and 5–8 h after egg laying (AEL) and stored in RNAlater® Solution (Ambion, Austin, TX, USA). Total RNA was extracted using RNAiso Plus reagent (TaKaRa, Dalian, China) according to the manufacturer's protocol, with the quantity and purity examined using a Bioanalyzer 2100 (Agilent, CA, USA). A total amount of 1 µg RNA from 0–1, 2–4, and 5–8 h embryos was used as the input material to prepare RNA samples. Sequencing libraries were produced using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, USA) according to the manufacturer's instructions. The index codes were added to sequences to distinguish each sample.

In brief, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. The divalent cations were used for fragmentation under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X), and the random hexamer primer and M-MuLV Reverse Transcriptase (RNase H−) were used for first strand cDNA synthesis. Subsequently, DNA Polymerase I and RNase H were used for second strand cDNA synthesis, after which the exonuclease and polymerase were used to convert the remaining overhangs into blunt ends. Hybridizations were prepared by ligating the NEBNext Adaptor with a hairpin loop structure after the adenylation of the DNA fragment 3′ ends. The library fragments were purified with the AMPure XP system (Beckman Coulter, Beverly, MA, USA) to preferentially select 250–300 bp cDNA fragments. Then, 3 µL of USER Enzyme (NEB, Ipswich, MA, USA) was incubated with the size-selected and adaptor-ligated cDNA at 37 °C for 15 min, followed by 5 min at 95 °C before PCR. The Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer were used to conduct PCR. PCR products were then purified (AMPure XP system), with library quality being assessed using the Agilent Bioanalyzer 2100 system.

The cBot Cluster Generation System was employed to conduct the clustering of index-coded samples using the TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. After cluster generation, an Illumina HiSeq platform was used to sequence the library preparations and to generate 125 bp/150 bp paired-end reads.

2.3. Transcript Sequence Analysis

The in-house perl scripts were used first to process raw data (raw reads) of fastq format. After removing reads containing adaptor, reads containing poly-N, and low quality reads from raw data, the clean data (clean reads) were generated. The Q20, Q30 and GC-content of the clean data were calculated simultaneously, and high quality clean data were used for all of the downstream analyses. The paired-end clean reads were then aligned to the whole genome sequence (WGS) of *B. dorsalis* (NCBI Assembly #ASM78921v2) using HISAT2 v2.0.4. Since HISAT2 can generate a database of splice junctions based on the gene model annotation file, it was the preferred mapping tool versus non-splice mapping tools. The read numbers mapped to each gene were counted using HTSeq v0.9.1, and the expression level of each gene was further calculated by FPKM (Fragment Per Kilobase of exon model per million mapped reads) based on the length of the gene and the read counts mapped to this gene. FPKM is presently the most commonly utilized method for assessing gene expression levels as an effect of sequencing depth and gene length for the concurrent read counts [53,54].

2.4. Comparison of Differentially Expressed Genes

Prior to differential gene expression analysis, the read counts were adjusted by the edgeR program package through one scaling normalized factor in each sequenced library. Differential expression analysis of two conditions was performed using the DEGSeq R package (1.20.0). The Benjamini and Hochberg's method was used to adjust the P values. Genes with a corrected P-value less than 0.005 and log$_2$(Fold change) of 1 identified by DESeq were assigned as differentially expressed [55]. The GOseq was applied for the Gene Ontology (GO) enrichment analysis of differentially expressed genes, and the gene length bias was corrected. Significantly enriched differentially expressed genes were determined by GO terms with a corrected P-value less than 0.05. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource to check the statistical enrichment of differentially expressed genes using the molecular-level information from biological systems and large-scale molecular datasets from genome sequencing and other high-throughput experimental technologies (http://www.genome.jp/kegg/) [56]. The statistical enrichment of differentially expressed unigenes (DEGs) in KEGG pathways was analyzed with the KOBAS software [57]. Both known and novel transcripts were constructed and identified using the Cufflinks v2.1.1 Reference Annotation Based Transcript (RABT) assembly method from TopHat alignment results [58]. The software rMATS v3.2.5 was used to classify the alternative splicing (AS) events as five basic types. The number of AS events was estimated separately for each compared group.

2.5. Quantitative Real-Time PCR

The developmental transcript expression profiles of sex determination and early zygotic genes were investigated using quantitative Real-Time PCR (qPCR). Total RNA was extracted using RNAiso Plus reagent (TaKaRa, Dalian, China) from 100 embryos per replicate, with 200 ng for each sample subjected to reverse transcription for mRNA using the PrimeScript^{TM} RT Master Mix (TaKaRa, Dalian, China). The reverse transcription products were used for qPCR using the primers listed in Table S1. qPCR was performed using the SYBR Green qPCR mix following the manufacturer's instructions in a real-time thermal cycler (Bio-Rad, Hercules, CA, USA) using the cycling conditions: 95 °C for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 30 s. Three biological and three technical replicates were performed, with expression data being analyzed by the $2^{-\Delta\Delta ct}$ method [59]. Dissociation curves were determined for each mRNA to confirm unique amplification. The expression of ribosomal protein 49 (*rp49*) was used as an internal control to normalize the expression of mRNA.

2.6. Statistical Analysis

Conceptual amino acid sequences encoded by the *tra*, *tra-2* and *dsx* genes in indicated species were used to construct a phylogenetic tree using MEGA5.0 (Molecular Evolutionary Genetics Analysis, Version 5.0, Sudhir Kumar, AZ, USA) with the pair-wise deletion option under the JTT empirical amino acid substitution model. Branch support was assessed by bootstrap analysis with 1000 replicates. All experiments were repeated at least three times and analyzed using GraphPad Prism 5.0 (GraphPad Software, San Diego, CA, USA) or Microsoft Excel (Microsoft, Redmond, WA, USA), with results being expressed as the mean ± SEM. Data were compared with either a two-way ANOVA—with subsequent *t* tests using Bonferroni post-tests for multiple comparisons—or with Student's *t* test. For all tests, differences were considered significant when $p < 0.05$.

3. Results

3.1. B. Dorsalis Early Embryo Sequencing and Assembly

Transcriptome libraries (accession number: GSE118472) for the *B. dorsalis* early embryonic developmental stages (0–1, 2–4, and 5–8 h embryos AEL) were constructed by paired-end (PE) Illumina sequencing, which generated a total of 272,036,418 125 bp/150 bp-long PE reads with high sequence quality. The filtered sequence reads from all samples were assembled and produced 13,489 unigenes. The average size of these unigenes was 2185 bp with lengths that ranged from 61

to 57,167 bp (Table 1). There were 9809, 5227, 2873 and 1066 unigenes that were larger than 1000, 2000, 3000 and 5000 bp, respectively, with all having a complete open reading frame (ORF) (Figure S1). After removing low-quality reads, 0–1, 2–4, and 5–8 h libraries generated 83, 85 and 98 million clean reads, respectively. Among these clean reads, 72–84 million (85.12%–86.97%), were mapped to unigenes in the whole genome sequence (WGS) of *B. dorsalis* (Table 2). The percentage of clean reads ranged from 98.08% to 98.61%, and the percentage of reads mapped to the genome exon regions ranged from 91.2% to 97.1%, reflecting high quality sequencing (Figures S2 and S3).

Table 1. Summary of the *B. dorsalis* early embryo transcriptome.

Total Reads	272,036,418
Total number of unigenes	13,489
Average transcript length (bp)	2185
Minimum (bp)	61
Maximum (bp)	57,167
Number of transcripts >1 Kb	9809
Number of transcripts >2 Kb	5227
Number of transcripts >3 Kb	2873
Number of transcripts >5 Kb	1066
Number of transcripts >10 Kb	128

Table 2. Alignment statistics of the *B. dorsalis* early embryo RNA-Seq analysis.

Sample Name	0–1 h	2–4 h	5–8 h
Raw reads	84,969,396	87,332,604	99,734,418
Clean reads	83,787,930	85,655,698	98,332,254
Clean bases	12.57 G	12.85 G	14.75 G
Total mapped	73,099,585 (87.24%)	73,131,953 (85.38%)	84,393,170 (85.82%)
Multiple mapped	228,708 (0.27%)	223,940 (0.26%)	336,793 (0.34%)
Uniquely mapped	72,870,877 (86.97%)	72,908,013 (85.12%)	84,056,377 (85.48%)
Error rate (%)	0.02	0.02	0.02
Q 20 (%)	97.48	96.93	97.38
Q 30 (%)	92.72	91.65	92.58
GC content (%)	40.94	40.59	40.71

3.2. Comparison of Gene Expression Profiles in Three Sequential Early Embryo Stages

To evaluate the relative expression level of unigenes in the *B. dorsalis* early embryo transcriptome, unigene paired-end read counts were normalized by transforming them into Fragments Per Kilobase of transcript per Million mapped reads (FPKM). We obtained a wide range of expression levels from less than 1 FPKM to approximately 6595 FPKM (Table S2). We observed that 47.76%, 47.26%, and·42.38% of the unigenes had a low expression level (FPKM 0–1); 37.96%, 40.40%, and 45.52% had a mid-expression level (FPKM 1–60); and 14.28%, 12.34%, and 12.10% exhibited a high expression level (FPKM > 60) in the 0–1, 2–4 and 5–8 h transcriptome libraries, respectively (Figure 1). The three early embryo stages were evaluated in three pairwise comparisons: 2–4 h vs. 0–1 h, 5–8 h vs. 0–1 h, and 5–8 h vs. 2–4 h, and unigenes found to have significant differences in expression were identified in each comparison (Figure 2). A comparison of the results showed the expression of 1683 unigenes was significantly different between 2–4 and 0–1 h. Of these unigenes, 369 were up-regulated and 1314 were down-regulated (Figure 2a and Table S3). In the comparison of 5–8 and 0–1 h, the expression profiles of 3201 unigenes had altered, of which 1451 were up-regulated and 1750 were down-regulated (Figure 2b and Table S3). When comparing 5–8 and 2–4 h, the expression of 3134 unigenes was significantly different as 1675 unigenes were up-regulated and 1459 unigenes were down-regulated (Figure 2c and Table S3). The expression of 368 unigenes was significantly different in three pairwise comparisons (Figure 3).
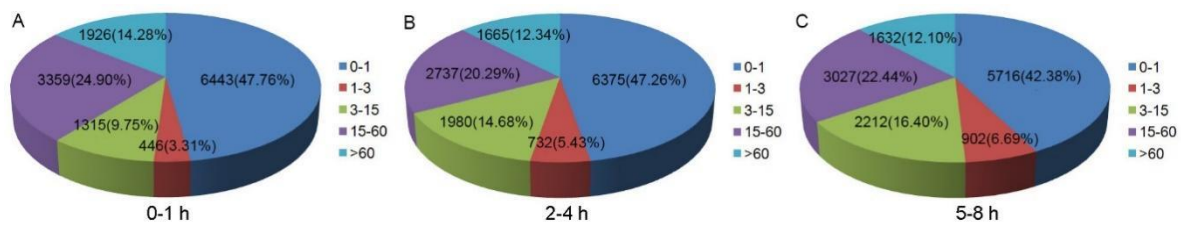
Figure 1. Fragments Per Kilobase of transcript per Million mapped reads (FPKM) distribution of the *B. dorsalis* early embryo transcriptomes.
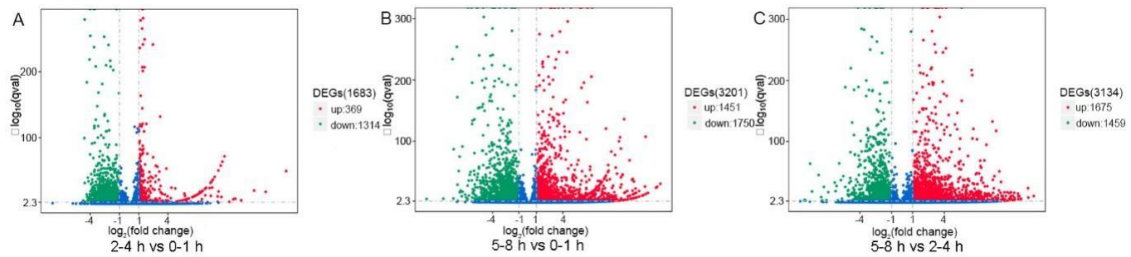

Figure 2. Volcano plot showing differentially expressed unigenes (DEGs, FDR < 0.01 and |log2 ratio| ≥ 1) in 2–4 h vs. 0–1 h, 5–8 h vs. 0–1 h, and 5–8 h vs. 2–4 h libraries of the *B. dorsalis* early embryos' transcriptomes. The up-regulated unigenes are represented by a red dot and the down-regulated unigenes, by a green dot.
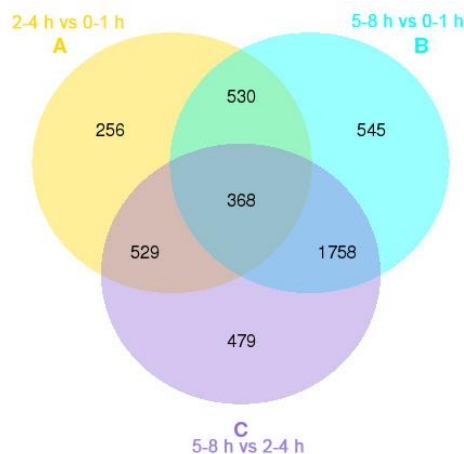

Figure 3. Venn diagram representing the distribution of differentially expressed unigenes in the 0–1, 2–4 and 5–8 h libraries of *B. dorsalis* early embryos. A, B and C represent the number of differentially expressed unigenes between 2–4 and 0–1 h, between 5–8 and 0–1 h, and between 5–8 and 2–4 h libraries, respectively.

3.3. Unigene Function Annotation and Analysis

Unigene sequences were annotated by searching the nonredundant NCBI protein database using BLASTX. A total of 11,693 distinct sequences (86.69% of the unigenes) matched known genes (Table S2). Based on the gene ontology (GO) classification, differently expressed unigenes were characterized within three groups: biological processes, cellular components, and molecular function (Table S4). The results of each comparison showed high concordance with unigenes related to biological processes (Figure 4). In the comparison of 2–4 and 0–1 h, and 5–8 and 0–1 h, 127 and 421 biological process unigenes were involved in the cellular macromolecule metabolic process (Figure 4a,b). In the comparison of 5–8 and 2–4 h, 607 biological process unigenes were involved in the cellular metabolic process (Figure 4c). In the molecular function category, enrichment comparisons showed that 93 unigenes involved in

nucleic acid binding were enriched at 2–4 h vs. 0–1 h, 804 unigenes involved in binding were enriched at 5–8 h vs. 0–1 h, and 496 unigenes involved in organic cyclic and heterocyclic compound binding were enriched t 5–8 h vs. 2–4 h (Figure 4). In addition, the UniGenes metabolic pathway analysis was conducted using the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation system in each comparison (Table S5). We found that spliceosome pathway was enriched in the comparison of stages 2–4 and 0–1 h, and 5–8 and 0–1 h (Figure 5). The observation that the RNA binding and spliceosome pathway were highly enriched and overrepresented in the GO term and KEGG analysis during *B. dorsalis* early stage embryogenesis may be an indication that RNA binding and spliceosome pathway enrichment is consistent with sex-specific alternative splicing.
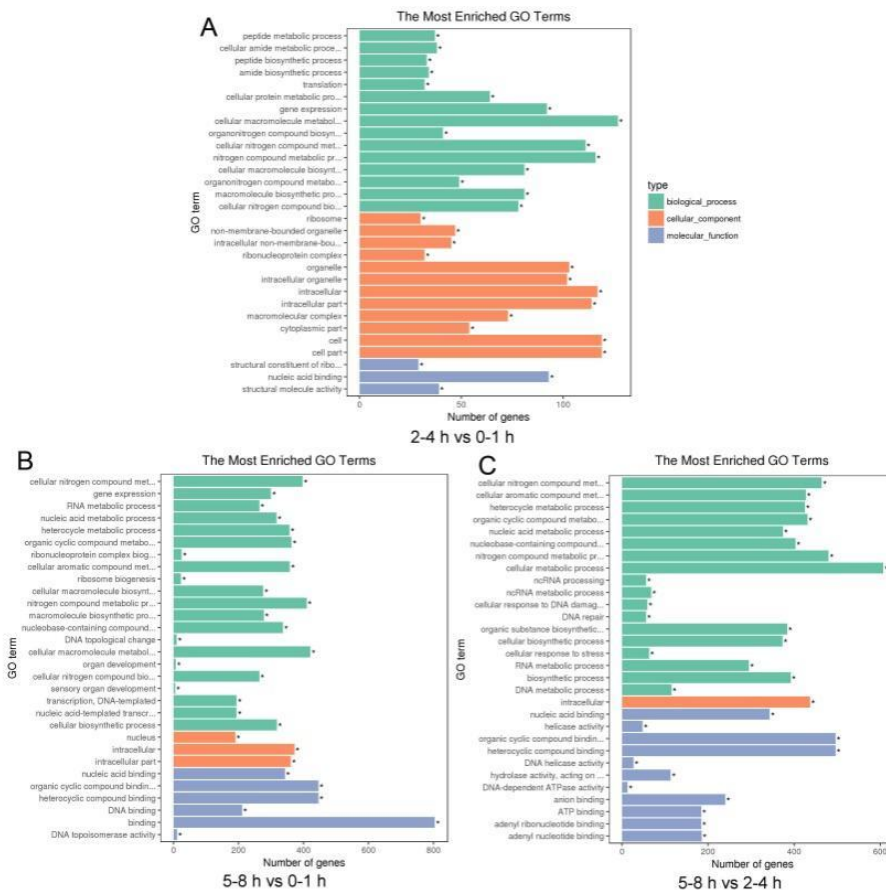

Figure 4. Gene Ontology (GO) terms classification for differentially expressed unigenes (DEGs) during the early embryo transcriptomes of *B. dorsalis*. (**A**) GO significant enrichment analysis for the DEGs between 2–4 and 0–1 h. (**B**) GO significant enrichment analysis for the DEGs between 5–8 and 0–1 h. (**C**) GO significant enrichment analysis for the DEGs between 5–8 and 2–4 h.
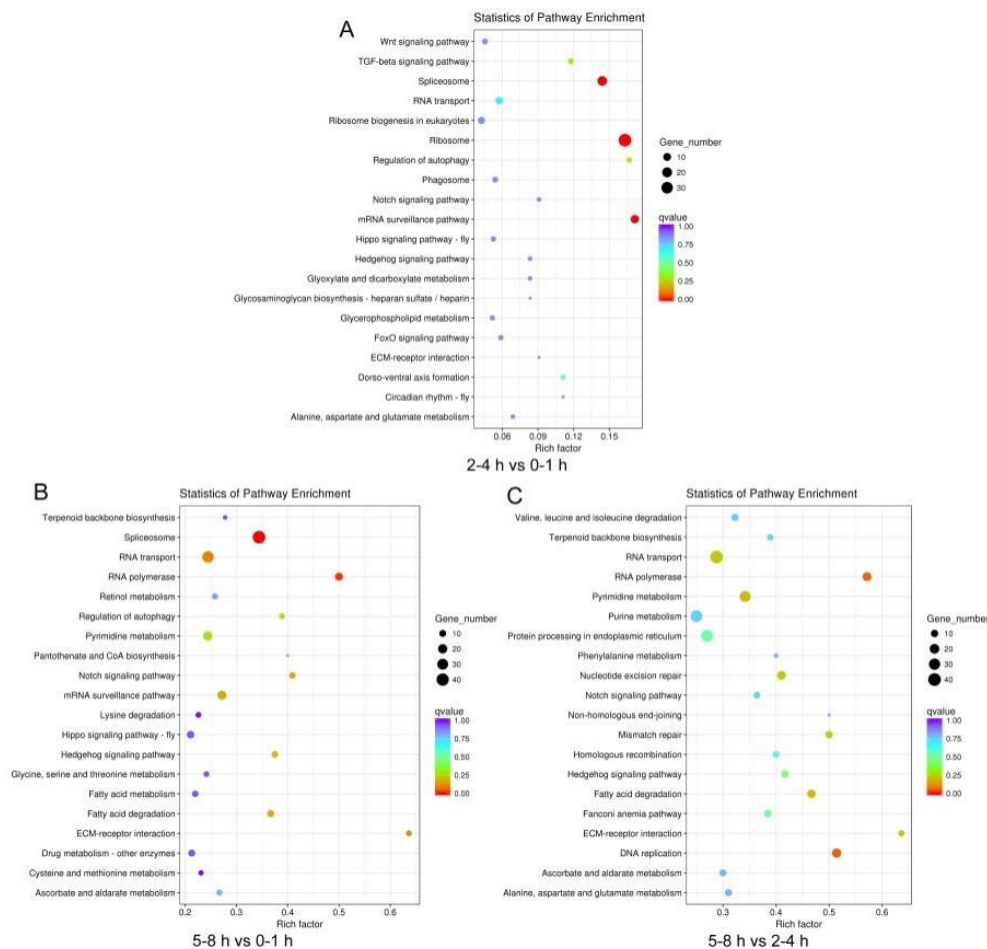
Figure 5. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway classification for differentially expressed unigenes (DEGs) during the early embryo transcriptomes of *B. dorsalis*. (**A**) KEGG significant enrichment analysis for the DEGs between 2–4 and 0–1 h. (**B**) KEGG significant enrichment analysis for the DEGs between 5–8 and 0–1 h. (**C**) KEGG significant enrichment analysis for the DEGs between 5–8 and 2–4 h.

3.4. B. dorsalis Sex Determination and Early Zygotic Genes

The alternative intron splicing events that occur during early embryogenesis in *B. dorsalis* include the skipped exons (SE), alternative 5′ splice sites (A5SS), alternative 3′ splice sites (A3SS), mutually exclusive exons (MXE), and retained intron (RI), of which exon skipping exhibited the highest distribution in the three developmental time periods compared (Figure S4). In addition to the *Bdtra*, *Bdtra-2* and *Bddsx* genes, we identified the expression of another 18 additional sex determination gene transcripts in early embryos (Table 3). We found that all sex determination genes contained full length ORFs and that most of the identified sex determination ortholog genes were highly expressed, except for *Bddsx*, where low levels (normalized FPKM < 1) did not change significantly among 0–1, 2–4 and 5–8 h embryos (Table 3). The phylogenetic analysis of the TRA, TRA-2 and DSX amino acid sequences showed that the putative orthologs of *Bdtra*, *Bdtra-2* and *Bddsx* were conserved in tephritid species, especially among the *Bactrocera* (Figure 6). The qPCR results showed that the expression of *Bdtra*, *Bdfru*, *BdSxl* and *Bddpn* increased significantly from 0–1 to 5–8 h, while the expression of *Bddsx*, *Bdtra-2*, *Bdda* and *Bdfl(2)d* had no substantial changes across the three early embryo stages, and the expression of *Bdotu* decreased significantly from 0–1 to 5–8 h (Figure 7). All these results were consistent with our deep sequencing data, which indicated that the current analysis is accurate. The maternal-to-zygotic transition is the period in *B. dorsalis* embryogenesis when maternal transcripts degrade and the zygotic genome expression is activated. In *B. dorsalis*, there were two waves of

zygotic transcription at 25 °C, a minor wave occurring early where *BdsisA* and *Bddpn* genes were activated before 4 h after oviposition, and a second major wave that occurred at 5 h when *Bdtra*, *BdSxl* and *Bdfl(2)d* gene transcription was initiated based on the qRT-PCR and mRNA-sequencing results (Figure 8). The waves of zygotic transcription during *B. dorsalis* embryogenesis are similar to those in *C. capitata* [60]; however, in *D. melanogaster*, the minor wave initiates before 2 h after oviposition and the second wave occurs at 3 h, at the time when zygotic *Dmtra* initiates transcription (Figure 8). In the *B. dorsalis* early embryonic transcriptome, we identified a *BdZelda (BdZld)* sequence that encodes a zinc-finger transcription factor protein (Table 3), but we failed to identify *Bdslam* encoding transcripts in the transcriptome. *BdZld* displayed a significant increase in expression from 0–1 to 5–8 h, which was involved in the maternal-to-zygotic transition (Table 3). Other two cellularization genes, *Bdsrya* and *Bdnullo*, were found through BLAST annotation and motif searches of the transcriptome (Table 3). *Bdsrya* showed greater amino acid similarity to the *D. melanogaster srya*-like gene, which is involved in cellular blastoderm formation.

Table 3. Expression profiles of the sex-determination and cellularization transcripts identified in the *B. dorsalis* early embryo transcriptomes. Expression in each sample is reported as the normalized FPKM value.

| Gene Name	Gene ID	Gene Length (bp)	ORF Length (aa)	0–1 h	2–4 h	5–8 h
Bddaughterless (Bdda)	105221792	3380	710	6.19	4.73	7.30
Bddeadpan (Bddpn)	105228067	2409	576	0.01	0.59	44.00
Bddoublesex (Bddsx)	105226634	3858	389	0.25	0.60	0.17
Bdfemale-lethal(2)d (Bdfl(2)d)	105224548	3326	641	31.72	49.32	51.90
Bdfruitless (Bdfru)	105224692	3346	839	0	0.03	29.84
Bdhopscotch (Bdhop)	105225518	3921	1175	62.75	51.02	54.04
Bdscute (Bdsc)	105227369	1122	296	0	0.10	57.87
Bdsisterless A (BdsisA)	105231921	1075	259	72.23	164.06	524.97
BdSex-lethal (BdSxl)	105232012	1938	354	57.67	119.48	169.84
Bdtransformer (Bdtra)	105232904	2577	422	18.31	19.48	42.76
Bdtransformer-2 (Bdtra-2)	105222300	1676	252	97.08	121.79	91.93
Bdvirilizer (Bdvir)	105229158	5842	1873	33.11	24.47	47.36
Bdgroucho (Bdgro)	105227097	2981	727	53.30	99.88	197.99
BdMes-4	105222825	4961	1556	77.50	43.23	39.14
Bdmdg4	105224476	1808	509	27.18	12.89	33.83
Bdlongitudinals (Bdlo)	105223842	1964	128	208.09	323.25	216.68
BdRho1	105230670	1828	193	207.05	267.43	258.73
Bdspaghetti-squash (Bdsqh)	105227302	1793	498	43.11	5.09	13.02
Bdextra-macrochaetae (Bdemc)	105224429	1646	236	0.12	35.34	187.76
Bdovarian tumor (Bdotu)	105222775	5302	1260	133.30	57.03	17.55
Bdrunt	105226934	3210	496	0	69.26	45.79
BdZelda (BdZld)	105232172	7823	1598	63.76	147.98	120.75
Bdnullo	Novel00216	782	211	0	68.06	2.15
Bdserendipity a (Bdsrya)	105222678	2520	663	32.58	47.88	10.82
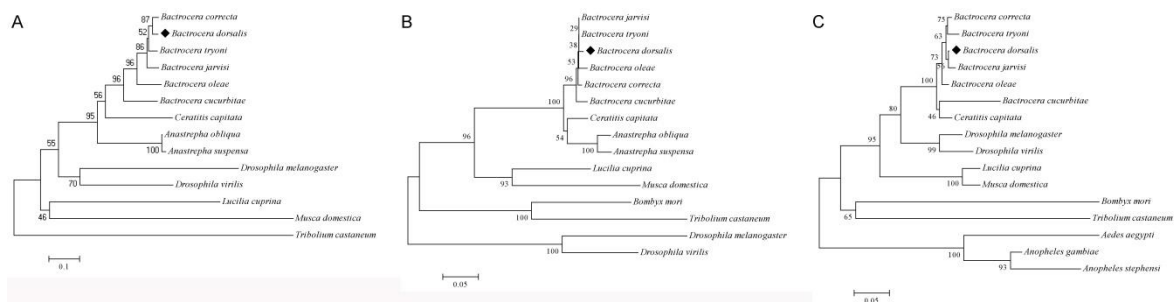
Figure 6. Neighbor-joining tree of insect transformer (TRA), transformer-2 (TRA-2) and doublesex (DSX) amino acid sequences. The scale represents the mean character distance.
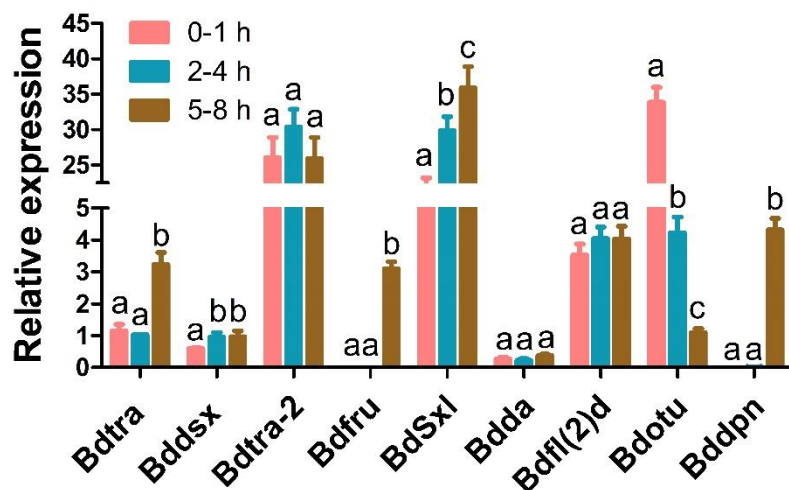
Figure 7. qPCR expression profiles of the sex determination genes during embryogenesis in *B. dorsalis*. Error bars indicate the SEM of three independent biological replicates and different letters above them indicate statistically significant differences (*p* < 0.05) among the three stages based on Student's *t*-test.
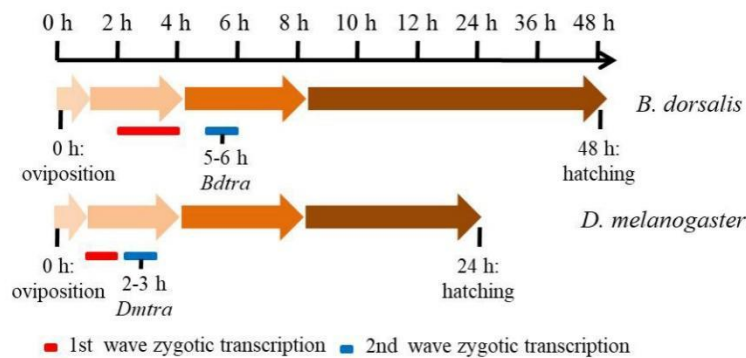

Figure 8. Timing of gene expression during embryogenesis in *B. dorsalis* and *D. melanogaster*. A schematic representation of the *B. dorsalis* and *D. melanogaster* embryogenesis stages, from oviposition (0 h) to hatching (48 and 24 h in *B. dorsalis* and *D. melanogaster*, respectively). The zygotic expression onset of *Bdtra* and the two zygotic expression waves are represented in both species.

4. Discussion

In previous studies, the developmental gene expression profiles of *B. dorsalis* have been investigated by constructing RNA-seq libraries for different developmental stages [61–64]. However, early embryogenesis transcripts in this dipteran species have been largely unexplored. To better define the broad gene expression profiles in early embryogenesis and the initiation of zygotic control of sex determination, we sequenced transcriptomes from three developmental time periods in *B. dorsalis* early embryos by next-generation sequencing (NGS) technology. Twenty-four sex determination and cellularization genes from a total of 13,489 unigenes were identified, and differential expression profiles of transcripts were validated from the early embryo libraries. Importantly, this study is the first transcriptome exploration of embryogenesis in *B. dorsalis*, which provides an essential gene expression resource necessary to fully understand the genetic basis of sex determination in this tephritid fruit fly.

In *D. melanogaster*, the combination of X-chromosome linked signal elements (XSEs), namely *sisA*, *sc*, *os* and *runt* as the primary signal, initiates the sex determination cascade [41,43,65,66]. In our study, we identified *B. dorsalis* homologs of *sisA*, *sc* and *runt*, but not *os*, expressed in the early embryonic stages. *BdsisA*, *Bdsc* and *Bdrunt* increased 7-, 57- and 45-fold, respectively, over this time period. Compared to the *sisA*, *sc*, *os* and *runt* genes that are all transcribed from the X chromosome in *D. melanogaster* [66,67], *BdsisA*, *Bdsc* and *Bdrunt* are autosomal genes in *B. dorsalis*. The transcripts

of *BdsisA* and *Bddpn* have the same expression profiles in *C. capitata*, which exhibit a significant increase in the zygote [60]. Although the primary sex determination signal is postulated to be a Y chromosome-linked male determining factor in tephritids [47–51], the identification of XSEs in the early embryonic developmental transcriptome may expand the function of these genes in the sex determination cascade of non-drosophilid insects [60,68–73]. In contrast to in *Drosophila*, *Sxl* is not sex-specifically spliced and is not involved in tephritid sex determination [46,74,75]. In our study, we identified the orthologs of *Drosophila Sxl*, *tra*, *tra-2*, and *dsx* genes in *B. dorsalis*, of which *BdSxl* and *Bdtra-2* have the same transcripts in males and females. *Bdtra* and *Bddsx* play an important role in *B. dorsalis* sex determination and reproduction, which are regulated by sex-specific alternative splicing [17,76–78]. Interestingly, we observed that RNA binding, the spliceosome pathway and alternative splicing events were highly enriched during the early stage of embryogenesis in *B. dorsalis*, which is also consistent with the role of sex-specific alternative splicing in *B. dorsalis* sex determination, in addition to that in other insects species [75,79,80]. The expression of *Bdtra* increased approximately two-fold during early embryogenesis, while *Bdtra-2* had consistently higher expression levels during the same periods. Given that TRA-2 performs a novel function by interacting with TRA to form a splicing factor that results in the female-specific splicing and autoregulation of tra^F and dsx^F [13–18,38,81,82], their significant expression levels in early embryos is not unexpected. However, the relative differences in the levels of *Bdtra-2* transcription had no substantial changes across the three early embryo stages.

Previous research shows that several genes expressed during the minor wave and major wave of embryonic transcription are involved in cellularization and sex determination [24,28,60,83]. Based on FPKM analysis, almost all the genes identified from the *B. dorsalis* early embryos have maternally derived transcripts (Table 3), except for *Bddpn*, *Bddsx*, *Bdfru*, *Bdsc*, *Bdemc*, *Bdrunt*, and *Bdnullo* having FPKM values <1 at 0–1 h. A *Bdslam* transcript was not detected previous to cellularization but has been reported as a maternal transcript in *C. capitata* and *Bactrocera jarvisi* [60,69]. *Bdsrya* was expressed during early embryogenesis at 0–1 h, and cognates of these genes have provided promoters that drive a transgenic tetracycline-suppressible embryonic lethality system, first tested in *D. melanogaster* [10] and then in the pest species *C. capitata*, *Anastrepha suspensa*, *Anastrepha ludens* and *Lucilia cuprina* [9,11,12,84–86]. Notably, for two of these species, *A. ludens* and *L. cuprina*, the use of the *srya* promoter resulted in embryonic lethality but also in maternal female sterility, presumably due to the pre-zygotic expression of the lethal effector in their ovaries that could be suppressed by a short-term tetracycline diet [84,85]. This might also occur in *B. dorsalis*, given the appearance of the pre-zygotic maternal *Bdsrya* transcript at 0–1 h. For *Lucilia*, however, the use of the *nullo* promoter avoided pre-zygotic cell lethality [86], and our results would suggest that the use of the *Bdnullo* promoter, which initiates expression zygotically at 2–4 h, would also be preferable for an embryo-specific lethality system in *B. dorsalis*. Other uses of embryonic and sex-determination genes for population control or the manipulation of various insects include the engineering of early zygotic gene (EZG) promoters in *Drosophila* for the zygotic-specific expression of antidote genes for the maternal-effect dominant embryonic arrest (*Medea*) gene-drive system [87]. For *Anopheles gambiae* mosquitoes, the CRISPR-Cas9 gene drive targeting of the terminal *dsx* sex determination gene has been used to induce female sterility, resulting in incomplete population suppression [88], and in the silkworm, *Bombyx mori*, a female-specific embryonic lethal system has been developed by the CRISPR-Cas9 targeting of the *tra-2* gene [89]. In our study, we have identified several sex determination and early zygotic genes that should also contribute to the development of novel genetic control strategies for *B. dorsalis* and related pest species.

5. Conclusions

In summary, RNA-seq analysis was performed to elucidate or confirm the sex determination and cellularization processes at the molecular level during early embryogenesis in *B. dorsalis*, a major invasive agricultural tephritid pest. Numerous genes displayed significant changes in expression during embryogenesis, especially genes involved in sex determination and the maternal-to-zygotic transcriptional transition. The gene expression patterns of *Bdtra*, *Bdtra-2* and *BdZld*, in particular,

are an indication of their important roles in early embryogenesis. Overall, the identification of sex-determination and cellularization genes has demonstrated the value of *B. dorsalis* as a non-typical model for early embryonic development and sexual differentiation. This study should also result in the advancement of the identification and use of gene promoters and sex-specific splicing mechanisms from these genes for transgenic-based improvements in pest management strategies for *B. dorsalis* and related fruit fly species.

Supplementary Materials: The following are available online at http://www.mdpi.com/2075-4450/11/5/323/s1, Figure S1: Distribution of unigene lengths in *B. dorsalis* early embryo transcriptomes. The sizes of all unigenes were calculated. Figure S2: Evaluation of sequence quality for the three early embryo developmental stages of *B. dorsalis*. Figure S3: Percent of reads mapped to the genome regions (exon, intergenic and intron) for the three early embryo developmental stages of *B. dorsalis*. Figure S4: The alternative splicing events during the early embryogenesis of *B. dorsalis*. SE, A5SS, A3SS, MXE and RI represent the skipped exon, alternative 5' splice site, alternative 3' splice site, mutually exclusive exons and retained intron (RI) respectively. Table S1: Primers used in our study. Table S2: All genes description and FPKM value during the embryogenesis in *B. dorsalis*. Table S3: Differentially expressed genes in each pairwise comparison during the embryogenesis in *B. dorsalis*. Table S4: GO classification of the differentially expressed genes in each pairwise comparison during the embryogenesis in *B. dorsalis*. Table S5: KEGG pathway enrichment analysis for differentially expressed genes in each pairwise comparison during the embryogenesis in *B. dorsalis*.

Author Contributions: H.Z. and W.P. conceived and designed the study; W.P. and S.Y. performed the experiments; W.P. and A.M.H. analyzed the data; W.P. wrote the manuscript, which all authors have read and approved the final version of the manuscript.

Conflicts of Interest: The authors have declared that no competing interests exist. The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Clarke, A.R.; Armstrong, K.; Carmichael, A.E.; Milne, J.R.; Raghu, S.; Roderick, G.K.; Yeates, D.K. Invasive Phytophagous Pests Arising through a Recent Tropical Evolutionary Radiation: The Bactrocera Dorsalis Complex of Fruit Flies. *Annu. Rev. Èntomol.* **2005**, *50*, 293–319. [CrossRef] [PubMed]

2. Zhang, H.; Li, H. *Photographic Guide to Key Control Techniques for Citrus Disease and Insect Pests*; Chinese Agricultural Press: Beijing, China, 2012.

3. Hsu, J.-C.; Feng, H.-T.; Wu, W.-J. Resistance and synergistic effects of insecticides in Bactrocera dorsalis (Diptera: Tephritidae) in Taiwan. *J. Econ. Èntomol.* **2004**, *97*, 1682–1688. [CrossRef] [PubMed]

4. Vontas, J.; Hernández-Crespo, P.; Margaritopoulos, J.T.; Ortego, F.; Feng, H.-T.; Mathiopoulos, K.D.; Hsu, J.-C. Insecticide resistance in Tephritid flies. *Pestic. Biochem. Physiol.* **2011**, *100*, 199–205. [CrossRef]

5. Ant, T.; Koukidou, M.; Rempoulakis, P.; Gong, H.-F.; Economopoulos, A.; Vontas, J.; Alphey, L.S. Control of the olive fruit fly using genetics-enhanced sterile insect technique. *BMC Biol.* **2012**, *10*, 51. [CrossRef]

6. Robinson, A.S. Genetic sexing strains in medfly, Ceratitis capitata, sterile insect technique programmes. *Genetica* **2002**, *116*, 5–13. [CrossRef]

7. Dafa'Alla, T.; Fu, G.; Alphey, L.S. Use of a regulatory mechanism of sex determination in pest insect control. *J. Genet.* **2010**, *89*, 301–305. [CrossRef]

8. Koukidou, M.; Alphey, L.S. Practical Applications of Insects' Sexual Development for Pest Control. *Sex. Dev.* **2014**, *8*, 127–136. [CrossRef]

9. Schetelig, M.F.; Cáceres, C.; Zacharopoulou, A.; Franz, G.; Wimmer, E. Conditional embryonic lethality to improve the sterile insect technique in Ceratitis capitata (Diptera: Tephritidae). *BMC Biol.* **2009**, *7*, 4. [CrossRef]

10. Horn, C.; Wimmer, E. A transgene-based, embryo-specific lethality system for insect pest management. *Nat. Biotechnol.* **2002**, *21*, 64–70. [CrossRef]

11. Schetelig, M.F.; Handler, A.M. A transgenic embryonic sexing system for Anastrepha suspensa (Diptera: Tephritidae). *Insect Biochem. Mol. Biol.* **2012**, *42*, 790–795. [CrossRef]

12. Schetelig, M.F.; Handler, A.M. Strategy for enhanced transgenic strain development for embryonic conditional lethality in Anastrepha suspensa. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9348–9353. [CrossRef] [PubMed]

13. Concha, C.; Scott, M.J. Sexual Development in Lucilia cuprina (Diptera, Calliphoridae) Is Controlled by the Transformer Gene. *Genetics* **2009**, *182*, 785–798. [CrossRef] [PubMed]

14. Hediger, M.; Henggeler, C.; Meier, N.; Perez, R.; Saccone, G.; Bopp, D. Molecular Characterization of the Key Switch F Provides a Basis for Understanding the Rapid Divergence of the Sex-Determining Pathway in the Housefly. *Genetics* **2009**, *184*, 155–170. [CrossRef] [PubMed]

15. Lagos, D.; Koukidou, M.; Savakis, C.; Komitopoulou, K. The transformer gene in Bactrocera oleae: The genetic switch that determines its sex fate. *Insect Mol. Biol.* **2007**, *16*, 221–230. [CrossRef]

16. Pane, A.; Salvemini, M.; Bovi, P.D.; Polito, C.; Saccone, G. The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development* **2002**, *129*, 3715–3725.

17. Peng, W.; Zheng, W.; Handler, A.M.; Zhang, H. The role of the transformer gene in sex determination and reproduction in the tephritid fruit fly, Bactrocera dorsalis (Hendel). *Genetica* **2015**, *143*, 717–727. [CrossRef]

18. Salvemini, M.; Robertson, M.; Aronson, B.; Atkinson, P.; Polito, L.C.; Saccone, G. Ceratitis capitata transformer-2 gene is required to establish and maintain the autoregulation of Cctra, the master gene for female sex determination. *Int. J. Dev. Biol.* **2009**, *53*, 109–120. [CrossRef]

19. Schetelig, M.F.; Milano, A.; Saccone, G.; Handler, A.M. Male only progeny in Anastrepha suspensa by RNAi-induced sex reversal of chromosomal females. *Insect Biochem. Mol. Biol.* **2012**, *42*, 51–57. [CrossRef]

20. Fu, G.; Condon, K.C.; Epton, M.J.; Gong, P.; Jin, L.; Condon, G.C.; Morrison, N.I.; Dafa'Alla, T.H.; Alphey, L.S. Female-specific insect lethality engineered using alternative splicing. *Nat. Biotechnol.* **2007**, *25*, 353–357. [CrossRef]

21. Tan, A.; Fu, G.; Jin, L.; Guo, Q.; Li, Z.; Niu, B.; Meng, Z.; Morrison, N.I.; Alphey, L.S.; Huang, Y. Transgene-based, female-specific lethality system for genetic sexing of the silkworm, Bombyx mori. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6766–6770. [CrossRef]

22. Jin, L.; Walker, A.S.; Fu, G.; Harvey-Samuel, T.; Dafa'Alla, T.; Miles, A.; Marubbi, T.; Granville, D.; Humphrey-Jones, N.; O'Connell, S.; et al. Engineered Female-Specific Lethality for Control of Pest Lepidoptera. *ACS Synth. Biol.* **2013**, *2*, 160–166. [CrossRef] [PubMed]

23. Fu, G.; Lees, R.; Nimmo, D.; Aw, D.; Jin, L.; Gray, P.; Berendonk, T.U.; White-Cooper, H.; Scaife, S.; Phuc, H.K.; et al. Female-specific flightless phenotype for mosquito control. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4550–4554. [CrossRef] [PubMed]

24. Tadros, W.; Lipshitz, H.D. The maternal-to-zygotic transition: A play in two acts. *Development* **2009**, *136*, 3033–3042. [CrossRef] [PubMed]

25. Tadros, W.; Goldman, A.L.; Babak, T.; Menzies, F.; Vardy, L.; Orr-Weaver, T.; Hughes, T.R.; Westwood, J.T.; Smibert, C.A.; Lipshitz, H.D. SMAUG Is a Major Regulator of Maternal mRNA Destabilization in Drosophila and Its Translation Is Activated by the PAN GU Kinase. *Dev. Cell* **2007**, *12*, 143–155. [CrossRef]

26. Bushati, N.; Stark, A.; Brennecke, J.; Cohen, S.M. Temporal Reciprocity of miRNAs and Their Targets during the Maternal-to-Zygotic Transition in Drosophila. *Curr. Biol.* **2008**, *18*, 501–506. [CrossRef]

27. Benoit, B.; He, C.H.; Zhang, F.; Votruba, S.M.; Tadros, W.; Westwood, J.T.; Smibert, C.A.; Lipshitz, H.D.; Theurkauf, W.E. An essential role for the RNA-binding protein Smaug during the Drosophila maternal-to-zygotic transition. *Development* **2009**, *136*, 923–932. [CrossRef]

28. Bosch, J.R.T. The TAGteam DNA motif controls the timing of Drosophila pre-blastoderm transcription. *Development* **2006**, *133*, 1967–1977. [CrossRef]

29. Liang, H.-L.; Nien, C.-Y.; Liu, H.-Y.; Metzstein, M.M.; Kirov, N.; Rushlow, C.A. The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* **2008**, *456*, 400–403. [CrossRef]

30. Erickson, J.W.; Cline, T.W. A bZIP protein, sisterless-a, collaborates with bHLH transcription factors early in Drosophila development to determine sex. *Genome Res.* **1993**, *7*, 1688–1702. [CrossRef]

31. Lecuit, T.; Samanta, R.; Wieschaus, E. slam Encodes a Developmental Regulator of Polarized Membrane Growth during Cleavage of the Drosophila Embryo. *Dev. Cell* **2002**, *2*, 425–436. [CrossRef]

32. Stein, J.A.; Broihier, H.T.; Moore, L.A.; Lehmann, R. Slow as Molasses is required for polarized membrane growth and germ cell migration in *Drosophila*. *Development* **2002**, *129*, 3925–3934. [PubMed]

33. Schejter, E.D.; Wieschaus, E. Bottleneck acts as a regulator of the microfilament network governing cellularization of the Drosophila embryo. *Cell* **1993**, *75*, 373–385. [CrossRef]

34. Rose, L.S.; Wieschaus, E. The Drosophila cellularization gene nullo produces a blastoderm-specific transcript whose levels respond to the nucleocytoplasmic ratio. *Genes Dev.* **1992**, *6*, 1255–1268. [CrossRef] [PubMed]

35. Verhulst, E.C.; Van De Zande, L.; Beukeboom, L.W. Insect sex determination: It all evolves around transformer. *Curr. Opin. Genet. Dev.* **2010**, *20*, 376–383. [CrossRef] [PubMed]

36. Ruiz, M.F.; Eirin-Lopez, J.M.; Stefani, R.N.; Perondini, A.L.P.; Selivon, D.; Sánchez, L. The gene doublesex of Anastrepha fruit flies (Diptera, Tephritidae) and its evolution in insects. *Dev. Genes Evol.* **2007**, *217*, 725–731. [CrossRef]

37. Ruiz, M.F.; Milano, A.; Salvemini, M.; Eirin-Lopez, J.M.; Perondini, A.L.P.; Selivon, D.; Polito, C.; Saccone, G.; Sánchez, L. The Gene Transformer of Anastrepha Fruit Flies (Diptera, Tephritidae) and Its Evolution in Insects. *PLoS ONE* **2007**, *2*, e1239. [CrossRef]

38. Sarno, F.; Ruiz, M.F.; Eirin-Lopez, J.M.; Perondini, A.; Selivon, D.; Sánchez, L. The gene transformer-2 of Anastrepha fruit flies (Diptera, Tephritidae) and its evolution in insects. *BMC Evol. Biol.* **2010**, *10*, 140. [CrossRef]

39. Salz, H.; Erickson, J.W. Sex determination in Drosophila the view from the top. *Fly* **2010**, *4*, 60–70. [CrossRef]

40. Penalva, L.O.F.; Sánchez, L. RNA Binding Protein Sex-Lethal (Sxl) and Control of Drosophila Sex Determination and Dosage Compensation. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 343–359. [CrossRef]

41. Erickson, J.W.; Quintero, J.J. Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in Drosophila. *PLoS Biol.* **2007**, *5*, e332. [CrossRef]

42. Evans, D.S.; Cline, T.W. Drosophila switch gene Sex-lethal can bypass its switch-gene target transformer to regulate aspects of female behavior. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4474–E4481. [CrossRef] [PubMed]

43. Hashiyama, K.; Hayashi, Y.; Kobayashi, S. Drosophila Sex lethal Gene Initiates Female Development in Germline Progenitors. *Science* **2011**, *333*, 885–888. [CrossRef] [PubMed]

44. Lagos, D.; Ruiz, M.F.; Sánchez, L.; Komitopoulou, K. Isolation and characterization of the Bactrocera oleae genes orthologous to the sex determining Sex-lethal and doublesex genes of Drosophila melanogaster. *Gene* **2005**, *348*, 111–121. [CrossRef]

45. Saccone, G.; Peluso, I.; Artiaco, D.; Giordano, E.; Bopp, D.; Polito, L.C. The Ceratitis capitata homologue of the Drosophila sex-determining gene sex-lethal is structurally conserved, but not sex-specifically regulated. *Development* **1998**, *125*, 1495–1500. [PubMed]

46. Meise, M.; Hilfiker-Kleiner, D.; Dübendorfer, A.; Brunner, C.; Nöthiger, R.; Bopp, D. Sex-lethal, the master sex-determining gene in Drosophila, is not sex-specifically regulated in Musca domestica. *Development* **1998**, *125*, 1487–1494. [PubMed]

47. Hall, A.B.; Basu, S.; Jiang, X.; Qi, Y.; Timoshevskiy, V.A.; Biedler, J.K.; Sharakhova, M.V.; Elahi, R.; Anderson, M.A.E.; Chen, X.-G.; et al. A male-determining factor in the mosquito Aedes aegypti. *Science* **2015**, *348*, 1268–1270. [CrossRef]

48. Krzywinska, E.; Dennison, N.J.; Lycett, G.J.; Krzywinski, J. A maleness gene in the malaria mosquito Anopheles gambiae. *Science* **2016**, *353*, 67–69. [CrossRef]

49. Sharma, A.; Heinze, S.D.; Wu, Y.; Kohlbrenner, T.; Morilla, I.; Brunner, C.; Wimmer, E.; Van De Zande, L.; Robinson, M.D.; Beukeboom, L.W.; et al. Male sex in houseflies is determined byMdmd, a paralog of the generic splice factor geneCWC22. *Science* **2017**, *356*, 642–645. [CrossRef]

50. Criscione, F.; Qi, Y.; Tu, Z. GUY1 confers complete female lethality and is a strong candidate for a male-determining factor in Anopheles stephensi. *eLife* **2016**, *5*, 219. [CrossRef]

51. Meccariello, A.; Salvemini, M.; Primo, P.; Hall, A.B.; Koskinioti, P.; Dalíková, M.; Gravina, A.; Gucciardino, M.A.; Forlenza, F.; Gregoriou, M.-E.; et al. Maleness-on-the-Y (MoY) orchestrates male sex determination in major agricultural fruit fly pests. *Science* **2019**, *365*, 1457–1460. [CrossRef]

52. Peng, W.; Tariq, K.; Zheng, W.-W.; Yu, S.-N.; Zhang, H.-Y. MicroRNA Let-7 targets the ecdysone signaling pathway E75 gene to control larval–pupal development in *Bactrocera dorsalis*. *Insect Sci.* **2017**, *26*, 229–239. [CrossRef] [PubMed]

53. Trapnell, C.; Pachter, L.; Salzberg, S.L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, *25*, 1105–1111. [CrossRef] [PubMed]

54. Florea, L.; Song, L.; Salzberg, S.L. Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues. *F1000research* **2013**, *2*, 188. [CrossRef] [PubMed]

55. Wang, L.; Feng, Z.; Zhang, X. DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **2009**, *26*, 136–138. [CrossRef] [PubMed]

56. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2007**, *36*, D480–D484. [CrossRef] [PubMed]

57. Mao, X.; Cai, T.; Olyarchuk, J.G.; Wei, L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **2005**, *21*, 3787–3793. [CrossRef]

58. Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **2013**, *14*, R36. [CrossRef]

59. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* **2001**, *25*, 402–408. [CrossRef]

60. Gabrieli, P.; Falaguerra, A.; Siciliano, P.; Gomulski, L.M.; Scolari, F.; Zacharopoulou, A.; Franz, G.; Malacrida, A.; Gasperi, G. Sex and the single embryo: Early deveopment in the Mediterranean fruit fly, Ceratitis capitata. *BMC Dev. Biol.* **2010**, *10*, 12. [CrossRef]

61. Chen, E.-H.; Hou, Q.; Dou, W.; Wei, D.-D.; Yue, Y.; Yang, R.-L.; Yu, S.-F.; De Schutter, K.; Smagghe, G.; Wang, J.-J. RNA-seq analysis of gene expression changes during pupariation in Bactrocera dorsalis (Hendel) (Diptera: Tephritidae). *BMC Genom.* **2018**, *19*, 693. [CrossRef]

62. Zheng, W.; Luo, D.; Wu, F.; Wang, J.-L.; Zhang, H. RNA sequencing to characterize transcriptional changes of sexual maturation and mating in the female oriental fruit fly Bactrocera dorsalis. *BMC Genom.* **2016**, *17*, 194. [CrossRef] [PubMed]

63. Zheng, W.; Peng, T.; He, W.; Zhang, H. High-Throughput Sequencing to Reveal Genes Involved in Reproduction and Development in Bactrocera dorsalis (Diptera: Tephritidae). *PLoS ONE* **2012**, *7*, e36463. [CrossRef] [PubMed]

64. Shen, G.-M.; Dou, W.; Niu, J.-Z.; Jiang, H.-B.; Yang, W.-J.; Jia, F.-X.; Hu, F.; Cong, L.; Wang, J.-J. Transcriptome Analysis of the Oriental Fruit Fly (Bactrocera dorsalis). *PLoS ONE* **2011**, *6*, e29127. [CrossRef] [PubMed]

65. Vicoso, B.; Bachtrog, D. Reversal of an ancient sex chromosome to an autosome in Drosophila. *Nature* **2013**, *499*, 332–335. [CrossRef]

66. Lu, H.; Kozhina, E.; Mahadevaraju, S.; Yang, D.; Avila, F.W.; Erickson, J.W. Maternal Groucho and bHLH repressors amplify the dose-sensitive X chromosome signal in Drosophila sex determination. *Dev. Biol.* **2008**, *323*, 248–260. [CrossRef]

67. Harrison, D. Sex determination: Controlling the master. *Curr. Biol.* **2007**, *17*, R328–R330. [CrossRef]

68. Salvemini, M.; Arunkumar, K.P.; Nagaraju, J.; Sanges, R.; Petrella, V.; Tomar, A.; Zhang, H.; Zheng, W.; Saccone, G. De Novo Assembly and Transcriptome Analysis of the Mediterranean Fruit Fly Ceratitis capitata Early Embryos. *PLoS ONE* **2014**, *9*, e114191. [CrossRef]

69. Morrow, J.L.; Riegler, M.; Frommer, M.; Shearman, D. Expression patterns of sex-determination genes in single male and female embryos of two Bactrocera fruit fly species during early development. *Insect Mol. Biol.* **2014**, *23*, 754–767. [CrossRef]

70. Jiménez-Guri, E.; Huerta-Cepas, J.; Cozzuto, L.; Wotton, K.R.; Kang, H.; Himmelbauer, H.; Roma, G.; Gabaldon, T.; Jaeger, J. Comparative transcriptomics of early dipteran development. *BMC Genom.* **2013**, *14*, 123. [CrossRef]

71. Tomancak, P.; Beaton, A.; Weiszmann, R.; Kwan, E.; Shu, S.; Lewis, S.E.; Richards, S.; Ashburner, M.; Hartenstein, V.; Celniker, S.E.; et al. Systematic determination of patterns of gene expression during Drosophila embryogenesis. *Genome Biol.* **2002**, *3*, 1–14. [CrossRef]

72. Lemke, S.; Antonopoulos, D.A.; Meyer, F.; Domanus, M.H.; Schmidt-Ott, U. BMP signaling components in embryonic transcriptomes of the hover fly Episyrphus balteatus (Syrphidae). *BMC Genom.* **2011**, *12*, 278. [CrossRef] [PubMed]

73. Gomulski, L.M.; Dimopoulos, G.; Xi, Z.; Soares, M.B.; Bonaldo, M.D.F.; Malacrida, A.; Gasperi, G. Gene discovery in an invasive tephritid model pest species, the Mediterranean fruit fly, Ceratitis capitata. *BMC Genom.* **2008**, *9*, 243. [CrossRef] [PubMed]

74. Gempe, T.; Beye, M. Function and evolution of sex determination mechanisms, genes and pathways in insects. *BioEssays* **2010**, *33*, 52–60. [CrossRef] [PubMed]

75. Bopp, D.; Saccone, G.; Beye, M. Sex Determination in Insects: Variations on a Common Theme. *Sex. Dev.* **2014**, *8*, 20–28. [CrossRef]

76. Liu, G.; Wu, Q.; Li, J.-W.; Zhang, G.; Wan, F.-H. RNAi-Mediated Knock-Down of transformer and transformer 2 to Generate Male-Only Progeny in the Oriental Fruit Fly, Bactrocera dorsalis (Hendel). *PLoS ONE* **2015**, *10*, e0128892. [CrossRef]

77. Chen, S.-L.; Dai, S.-M.; Lu, K.-H.; Chang, C. Female-specific doublesex dsRNA interrupts yolk protein gene expression and reproductive ability in oriental fruit fly, Bactrocera dorsalis (Hendel). *Insect Biochem. Mol. Biol.* **2008**, *38*, 155–165. [CrossRef]

78. Laohakieat, K.; Aketarawong, N.; Isasawin, S.; Thitamadee, S.; Thanaphum, S. The study of the transformer gene from Bactrocera dorsalis and B. correcta with putative core promoter regions. *BMC Genet.* **2016**, *17*, 34. [CrossRef]

79. Sánchez, L. Sex-determining mechanisms in insects. *Int. J. Dev. Biol.* **2008**, *52*, 837–856. [CrossRef]

80. Salz, H. Sex determination in insects: A binary decision based on alternative splicing. *Curr. Opin. Genet. Dev.* **2011**, *21*, 395–400. [CrossRef]

81. Hoshijima, K.; Inoue, K.; Higuchi, I.; Sakamoto, H.; Shimura, Y. Control of doublesex alternative splicing by transformer and transformer-2 in Drosophila. *Science* **1991**, *252*, 833–836. [CrossRef]

82. Hediger, M.; Siegenthaler, C.; Moser, M.; Bopp, D. The transformer2 gene in Musca domestica is required for selecting and maintaining the female pathway of development. *Dev. Genes Evol.* **2005**, *215*, 165–176. [CrossRef]

83. Pritchard, D.K.; Schubiger, G. Activation of transcription in Drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Dev.* **1996**, *10*, 1131–1142. [CrossRef] [PubMed]

84. Schetelig, M.F.; Targovska, A.; Meza, J.S.; Bourtzis, K.; Handler, A.M. Tetracycline-suppressible female lethality and sterility in the Mexican fruit fly, Anastrepha ludens. *Insect Mol. Biol.* **2016**, *25*, 500–508. [CrossRef] [PubMed]

85. Yan, Y.; Linger, R.J.; Scott, M.J. Building early-larval sexing systems for genetic control of the Australian sheep blow fly Lucilia cuprina using two constitutive promoters. *Sci. Rep.* **2017**, *7*, 2538. [CrossRef]

86. Yan, Y.; Williamson, M.E.; Davis, R.J.; Andere, A.A.; Picard, C.J.; Scott, M.J. Improved transgenic sexing strains for genetic control of the Australian sheep blow fly Lucilia cuprina using embryo-specific gene promoters. *Mol. Genet. Genom.* **2019**, *295*, 287–298. [CrossRef]

87. Chen, C.-H.; Huang, H.; Ward, C.M.; Su, J.T.; Schaeffer, L.V.; Guo, M.; Hay, B.A. A synthetic maternal-effect selfish genetic element drives population replacement in Drosophila. *Science* **2007**, *316*, 597–600. [CrossRef]

88. Kyrou, K.; Hammond, A.; Galizi, R.; Kranjc, N.; Burt, A.; Beaghton, A.K.; Nolan, T.; Crisanti, A. A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nat. Biotechnol.* **2018**, *36*, 1062–1066. [CrossRef]

89. Zhang, Z.; Niu, B.; Ji, D.; Li, M.; Li, K.; James, A.A.; Tan, A.; Huang, Y. Silkworm genetic sexing through W chromosome-linked, targeted gene integration. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 8752–8756. [CrossRef]