



國立臺灣大學與中央研究院

生命科學院基因體與系統生物學學位學程

碩士論文

Genome and Systems Biology Degree Program

College of Life Science

National Taiwan University & Academia Sinica

Master thesis

擬黃果蠅性別比減數分裂驅動之轉錄體學分析

Transcriptomic Analyses of *Sex-ratio* Meiotic Drive in

*Drosophila simulans*

林頌雅

Sung-Ya Lin

指導教授：丁照棟

Advisor: Chau-Ti Ting

中華民國 107 年 8 月

August, 2018

## 謝辭



兩年前，如願拿到一個有趣研究主題的我，還正興奮地構思著各種可能的研究方向，野心勃勃地定了三個不小的研究目標。一年後，面對沒有太大進展的研究，初始的雄心壯志轉化成了現實的求生意志，我下定決心專注於一個目標，才在很多人的幫助下，完成了這篇碩士論文。過程中有了小發現的喜悅和遇到問題的挫折，都在曲折了兩年後化為一份努力過後的踏實。

能夠走到今天，我很感謝丁照棟老師耐心和無私的指導。猶記得當初和您討論研究主題時，您那神采奕奕的樣子，而那次的對話也開啟了一扇我以往不曾踏入的門，讓我開始嘗試用演化的角度想事情。每當有好的演講和課程，您總鼓勵我並給予我很大的時間和空間去學習，讓我在研究路上一直保持自我精進。不管實驗結果是好是壞，您也永遠是笑著鼓勵我，時而支持我的決定，必要時指出我的盲點。很感謝有您的指導，這兩年我收穫了很多。

陪伴我完成這趟旅程的還有方淑老師、曹順成老師和口委老師們。方淑老師是個極端仔細和認真的人，很感謝您總是指出我的錯誤並給予建議。曹順成老師是個認真又風趣的人，很感謝您在實驗上的協助和建議。另外還要特別感謝口試委員呂俊毅老師、蔡怡陞老師和莊樹諄老師，在口試過程中，您們對於我研究提出的問題和建議，讓我有重新思考組織的機會，也使這篇論文變得更加完整。



促使我繼續堅持的另一份力量來自實驗室的夥伴們，感謝你們不僅提供我研究上的幫助，更豐富了我的研究生生活。家宣教了我很多實驗技術和觀念，也總用玩笑的方式鼓勵我。志冠在分析上提供我很多有用的建議，讓我免去很多獨自摸索的時間。宇謙總能精準地指出我的問題，並給予我在研究和論文中很實用的建議。李娟在分析上給我很大的幫助，一針見血地指出我的問題，並熱心地教導大家生物資訊的工具。光庭在實驗上給我的協助，讓我能更順利地進行實驗。陳曦在生活上一一直很照顧我，讓我很快就融入了這個環境。泓瑋、勝瑜、文宇和 Blaise 是和我一起學習的好夥伴，泓瑋細心，勝瑜執著，文宇自信，Blaise認真，很開心和你們一起進步。文喬、孟萱和美靜讓實驗室能順利地運作，文喬對人的關心、孟萱說的有趣語錄和美靜散播出的快樂，也讓我的研究生生活多了份溫度。Duyen和Romain是來自德國和法國的學生，Duyen的開朗外向和Romain的積極認真，為我的生活畫上特別的一筆。很感謝實驗室的大家，有你們真好。

為我打氣並默默支持我的還有我的 GSB 戰友、家人和朋友們。很感謝有 GSB 的大家和我一起學習和討論研究上的問題，有了大家的陪伴，我的研究路上不再孤單。而我最要感謝的是我的家人和朋友們，你們無條件地支持和包容我，理解我的忙碌，也為我的成果喝采，讓我有有了有力的後盾，由衷地感謝你們！

不管未來的路怎麼走，我會帶著大家的祝福，繼續踏實地走下去。

## 摘要



性別比 (*sex-ratio*, *SR*)減數分裂驅動會讓 X 染色體比 Y 染色體更容易被傳到下一代，因此造成偏向產生雌性的子代。*SR* 減數分裂驅動曾經在不同的物種中被發現過，但我們對於造成這個現象的分子機制還是不甚瞭解。在擬黃果蠅 (*Drosophila simulans*)的巴黎 *SR* 系統中，有很多參與作用的基因，但至今只有 *HP1* 基因家族裡的 *HP1D2* 是被確認的。*HP1D2* 在 *SR* 品系中有一段 *chromo shadow domain* 的缺失，並且表現量較野生型標準品系(*ST*)低。為了調查 *HP1D2* 的基因型和 *SR* 現象的關聯性，我做了基因型鑑定，並發現在 *ST* 品系內和 *SR* 品系內都有基因型差異。這項結果代表只有 *HP1D2* 基因型並無法預測 *SR* 的現象。為了更系統性地發現其他與 *SR* 相關的基因，我比較了 *ST* 與 *SR* 品系果蠅在精巢基因表現的轉錄體差異。在這些基因之中，有很高比例的多細胞生物生殖與免疫反應相關的基因。我接著做 RT-qPCR 來確認 *SR* 候選基因中的 34 個基因。雖然有五個基因在三個 *SR* 品系中都有較高表現量，包括 *CG16772*、*CG15209*、*Ser7*、*CG34265* 和 *CG43348*，但其他有表現差異的基因在不同 *SR* 品系中都不同。*SR* 品系的機制和遺傳基礎可能是不一樣的。依照現有的結果，要區分主要機制是殺手-目標驅動 (*killer-target drive*)或毒藥-解毒劑驅動 (*poison-antidote drive*)還很困難，但殺手-目標驅動是跟目前結果比較吻合的。如果可以找到 Y 染色體

上的目標，就有可能闡明巴黎 *SR* 系統的機制。




關鍵字：性別比、減數分裂驅動、轉錄體、擬黃果蠅、*HP1D2*

## Abstract



*Sex-ratio (SR)* meiotic drives, favoring the transmission of X over Y chromosome, lead to strong female-biased progeny of affected males. *SR* meiotic drives have been reported in several independent lineages. Yet, the molecular mechanism remains largely unclear. In *Drosophila simulans*, it has been known that many genes are involved in the Paris *SR* system, but only *HP1D2*, a member of the Heterochromatin Protein 1 (HP1) gene family, was identified. *HP1D2* possesses a deletion of chromo shadow domain in *SR* strains and is expressed at lower level in *SR* relative to wild-type standard (ST) strains. To examine the correlation between the *HP1D2* genotype and the *SR* phenotype, genotyping was performed. The observation of variation of *HP1D2* in both ST and *SR* strains indicates that the genotypes alone cannot predict *SR* phenotypes. To systematically identify other *SR*-related genes, the transcriptomic differences of testicular expression between three ST and three *SR* strains were compared. Among these genes, they are highly enriched in genes associated with multicellular organism reproduction and immune response. The RT-qPCR analysis was then performed to validate 34 genes from the *SR* candidate genes. Although five genes, namely *CG16772*, *CG15209*, *Ser7*, *CG34265*, and *CG43348*, were consistently up-regulated in three *SR*



strains, other differentially expressed genes differed among these strains. The underlying mechanisms and genetic bases of *SR* strains may be different. Based on the current results, although it is still difficult to distinguish major mechanisms, the killer-target drive or the poison-antidote drive, the killer-target drive is more consistent with the current results. If the target on Y chromosome can be identified, it is possible to elucidate the underlying mechanism of the Paris *SR* system.

**Keywords:** sex ratio, meiotic drive, transcriptome, *Drosophila simulans*, *HP1D2*

## Contents



Acknowledgements (Chinese).....	i
Abstract (Chinese).....	iii
Abstract.....	v
List of Figures.....	viii
List of Tables.....	x
Introduction.....	1
Materials and Methods.....	11
Results.....	22
Discussion.....	46
References.....	56
Appendix.....	69

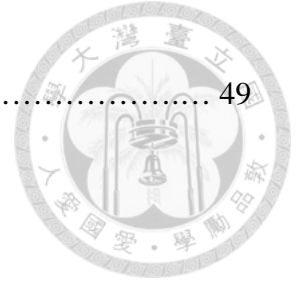


## List of Figures



Fig. 1 The two major drive loci of the Paris <i>SR</i> system.....	9
Fig. 2 Element 1 of the <i>SR</i> drive locus.....	9
Fig. 3 The crossing scheme of X chromosome isolation.....	13
Fig. 4 The maintenance of ST and <i>SR</i> strains.....	14
Fig. 5 The RNA-Seq data analyses pipeline.....	16
Fig. 6 Variation existing in the <i>HPID2</i> genotype of both ST and <i>SR</i> strains.....	25
Fig. 7 Testicular expression of genes related to heterochromatin organization and chromatin assembly.....	27
Fig. 8 ST and <i>SR</i> strains are in two different clusters.....	30
Fig. 9 Neighbor-joining tree of the X chromosome transcriptome sequences of three ST and three <i>SR</i> strains.....	32
Fig. 10 Up-regulated genes were more than down-regulated genes in <i>SR</i> strains.....	33
Fig. 11 Summary of the experimental design and <i>SR</i> candidate genes of transcriptomic analyses.....	34
Fig. 12 Gene ontology enrichment results of <i>SR</i> candidate genes.....	35
Fig. 13 Testicular expression of <i>SR</i> candidate genes.....	43

Fig. 14 Possible mechanisms of the X chromosome drive..... 49



## List of Tables



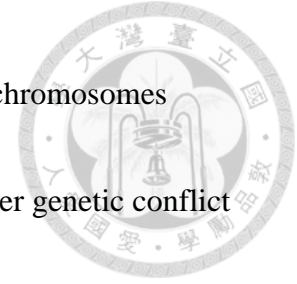
Table 1 <i>Drosophila simulans</i> strains used in this study.....	12
Table 2 Sex ratio phenotype assays of ST and <i>SR</i> strains.....	23
Table 3 The <i>SR</i> phenotype and <i>HP1D2</i> genotype of ST and <i>SR</i> strains.....	25
Table 4 The left reads after trimming and mapping rate of RNA-Seq reads.....	28
Table 5 The proportion of expressed genes among all genes.....	29
Table 6 The proportion of <i>SR</i> candidate genes on each chromosome.....	33
Table 7 Testicular expression of <i>SR</i> candidate genes.....	40
Table 8 The spermatogenesis stage at which the <i>SR</i> candidate genes are mainly expressed in the testes.....	51
(Supplementary Tables)	
Table B1 List of primers used in the RT-qCPR experiments.....	70

## Introduction




Sex ratio is the ratio of different sexes in a population of a species which produces offspring through sexual reproduction. Among these species, the sex ratio of most species is 1:1 ratio between males and females (Fisher 1930). However, the 1:1 ratio can sometimes be violated, resulting in male-biased or female-biased offspring, called sex ratio distortion. The factors affecting the 1:1 ratio can be environmental, pathogenic, or genetic factors.

For genetic factors, for example, in haplodiploid systems, which females are diploid and males are haploid, PSR (paternal sex ratio) chromosomes cause male-biased sex ratio in the parasitic wasp *Nasonia vitripennis* and *Trichogramma kaykai*. The PSR chromosomes act by destructing the paternal set of chromosomes in the sperm, leading to PSR-carrying males with only maternal set of chromosomes (Werren and Stouthamer 2003; Camacho et al. 2011). Another example is the sex chromosome meiotic drive in diploid systems, such as female-biased *Rodentia* and *Drosophila* (reviewed in Helleu et al. 2015). The preferentially transmission of X chromosome leads to female-biased offspring.



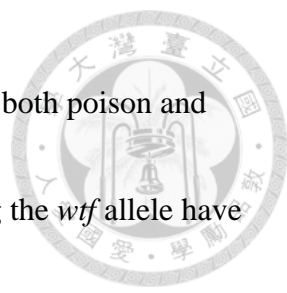
Meiotic drive is the non-Mendelian transmission of alleles or chromosomes during gametogenesis. It is a selfish genetic element which can trigger genetic conflict because it enhances its own transmission but is often detrimental to the rest of the genome and the organism (reviewed in Werren et al. 1988; Werren 2011). It can be fixed, lost, or become a stable polymorphism. Meiotic drivers are widespread in nature, but the cases which are reported are probably largely underestimated because meiotic drive is hard to detect. The difficulty of detection results from that the meiotic drives can be observed only in heterozygotes rather than homozygotes. When geneticists choose inbred lines to do experiments or when the meiotic driver is fixed in the population, it is impossible to find meiotic drivers because there are only homozygotes in both cases. Another reason why the meiotic drive is hard to detect lies in the presence of suppressors which act on the driver (reviewed in Bravo Nunez et al. 2018).

Despite the difficulty to discover meiotic drives, there are a few identified genes in different systems. There are two general classes of meiotic drivers, including true meiotic drivers and killer meiotic drivers. True meiotic drivers act during meiotic divisions (Bravo Nunez et al. 2018). For example, stronger centromeres with increased kinetochore protein levels and altered interaction with spindle microtubules in mouse



oocytes are preferentially segregated to the eggs (Chmatal et al. 2014). Also, spindle asymmetry in mouse oocytes can retain the meiotic drivers in the eggs (Akeru et al. 2017). Another example is the loci called “knobs” in maize, in which the *Kinesin driver* (*Kindr*) interacts specifically with neocentromeres and promotes meiotic drive (Dawe et al. 2018).

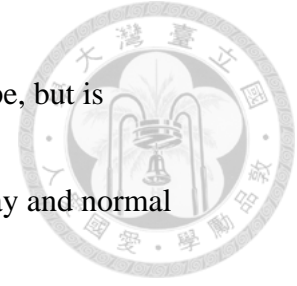
Killer meiotic drivers are the destruction of gametes without the drivers, thus enabling the drivers to preferentially transmit to the functional gametes (Lindholm et al. 2016). For example, the *t*-haplotype in mice transmits itself at the expense of its wild-type homologous chromosome in heterozygous *+/t* males (Silver 1993). The *t*-haplotype is located on the chromosome 17 and consists of four inversions relative to the wild-type, with four *t*-complex distorters (*Tcds*) and a *t*-complex responder (*Tcr*). Three of the four *Tcds*, including *Tagap1*, *Fgd2*, and *Nme3*, act in *trans* to disrupt flagellar function by overactivating the sperm motility kinase SMOK1. *Tcr* rescues the defect of sperm motility of only *Tcr*-carrying sperms, resulting in an advantage for sperms bearing *t*-haplotype (Herrmann et al. 1999; Schimenti 2000; Bauer et al. 2005; Bauer et al. 2007; Bauer et al. 2012). Another example is the *wtf* genes in *Schizosaccharomyces pombe*. The *wtf* genes are autonomous spore-killing meiotic drive genes, resulting in



destroying of spores not inheriting the driver. The *wtf* genes produce both poison and antidote using dual, overlapping transcripts. Only the spores carrying the *wtf* allele have a better chance to be viable, biasing transmission of *wtf* into >70% of the viable spores (Hu et al. 2017; Nuckolls et al. 2017).

In *Drosophila*, the most known meiotic drive is the *Segregation Distorter (SD)* in *D. melanogaster*, which is a killer meiotic drive. *SD* preferentially produces *SD*-bearing progeny of *SD/SD*<sup>+</sup> males by inducing dysfunction of *SD*<sup>+</sup> spermatids (Larracuent and Presgraves 2012). *SD* is an autosomal driver which is mainly composed of two loci, the *Segregation distorter (Sd)* (Sandler and Carpenter 1972) and the target of drive, *Responder (Rsp)* (Hartl 1974). In addition, there is a modifier of distortion, called *Enhancer of Segregation Distorter (E(SD))* (Ganetzky 1977).

The *Sd* locus on the chromosome arm 2L encodes a partial duplication of the gene *RanGAP (Ran GTPase activating protein)*, called *Sd-RanGAP* (Merrill et al. 1999), while *Rsp* on the chromosome arm 2R contains variable numbers of a block of satellite DNA correlated with sensitivity to *SD* (Wu et al. 1988; Lyttle 1989; Pimpinelli and Dimitri 1989; Lyttle 1991). The *Sd-RanGAP* protein is a truncated form of enzyme with wild-type cytoplasm function which stimulates hydrolysis of Ran-GTP to Ran-GDP and



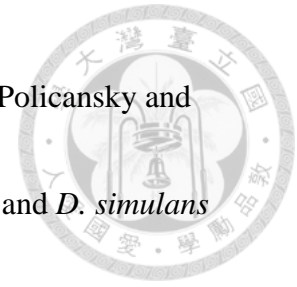
maintains a steep Ran-GTP concentration across the nuclear envelope, but is mislocalized to the nucleus thus disrupting the Ran signaling pathway and normal nuclear transport. (Kusano et al. 2001, 2002). It promotes meiotic drive by disrupting the spermiogenesis only in drive-sensitive spermatids ( $Rsp^s$ ) with many  $Rsp$  repeats, but not in drive-insensitive spermatids ( $Rsp^i$ ) with only a few or no  $Rsp$  repeats. However, how Sd-RanGAP causes the defect of  $Rsp^s$ -carrying spermatids remains unknown. Moreover, the established  $SD$  haplotype also promotes the evolution of enhancers of distortion and suppressors of recombination (Larracuenta and Presgraves 2012).

Meiotic drives can be located on both the autosomes and the sex chromosomes.

Sex chromosome meiotic drive is the sex chromosome-linked meiotic drive, leading to sex ratio distortion. In *Drosophila* species, sex chromosome meiotic drive favors the transmission of X relative to Y chromosome, leading to strong female-biased progeny of affected males. The observation of X-linked meiotic drive has been reported in 13 *Drosophila* species (Helleu et al. 2015). In 1925, sex ratio distortion was first observed in *D. affinis* (Morgan et al. 1925). Later in 1928, female-biased offspring from some *D. obscura* males were also observed, resulting from an X-linked genetic element sex-ratio ( $SR$ ) (Gershenson 1928). Since then, the sex ratio distortion phenotype has been



observed in other *Drosophila* species, including *D. pseudoobscura* (Policansky and Ellison 1970), *D. subobscura* (Hauschteckjungen and Maurer 1976) and *D. simulans* (Helleu et al. 2015).



Among these cases, *D. simulans* has been described in three X-linked *sex-ratio* (*SR*) systems, including the Durham (Tao et al. 2001), Winters (Tao et al. 2007a,b), and Paris (Mercot et al. 1995a,b) *SR* systems.

#### Durham *SR* system

The sex ratio distortion of the *D. simulans* from the Durham *SR* system was discovered by introgressing the third chromosome from *D. mauritiana* genome segments into *D. simulans* genome. A dominant autosomal suppressor of sex ratio distortion, called *Too much yin* (*Tmy*), was found in the *D. simulans* by replacing it with a nonsuppressing allele *tmy* from *D. mauritiana*. In addition to controlling sex ratio distortion, *tmy* also controls hybrid male sterility. There is a tight association between a suppressor of sex ratio distortion and hybrid male sterility (Tao et al. 2001).

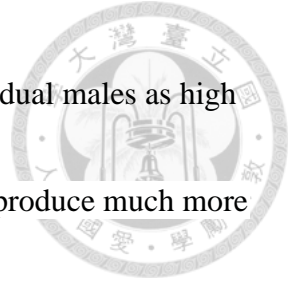


### Winters *SR* system

The sex ratio distortion of the *D. simulans* from the Winters *SR* system was first discovered when introgressing the *D. sechellia* genome into the *D. simulans* genome (Dermitzakis et al. 2000). An X-linked distorter of sex ratio and an autosomal suppressor were found in this system. The X-linked distorter *Dox* (*Distorter on the X*) is a new gene which arises from another new gene *MDox* (*Mother of Dox*). Both *Dox* and *MDox* are either non-coding RNAs or as mRNA with very limited coding potential. The dominant suppressing allele *Not much yang* (*Nmy*) was identified on the third chromosome of *D. simulans*. *Nmy* originated from *Dox* through a retrotransposition event, which is likely to suppress the distorter through RNA interference mechanism. In addition, the *SR* phenotype is resulted from the failure of the Y-bearing sperm maturation when lacking the suppressor *Nmy* (Tao et al. 2007a,b).

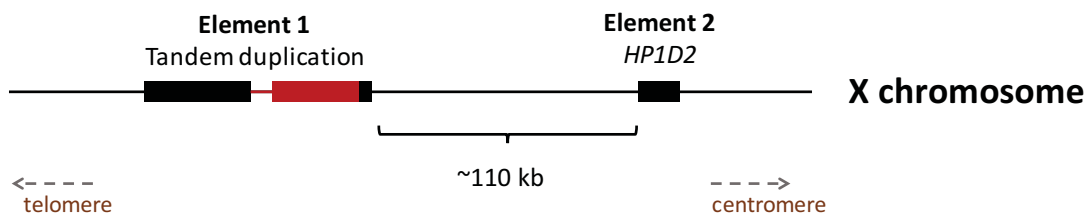
### Paris *SR* system

The Paris *SR* system is used in my study. The *D. simulans* flies used in the Paris *SR* system were initially collected from Seychelles (Mercot et al. 1995a,b), with the

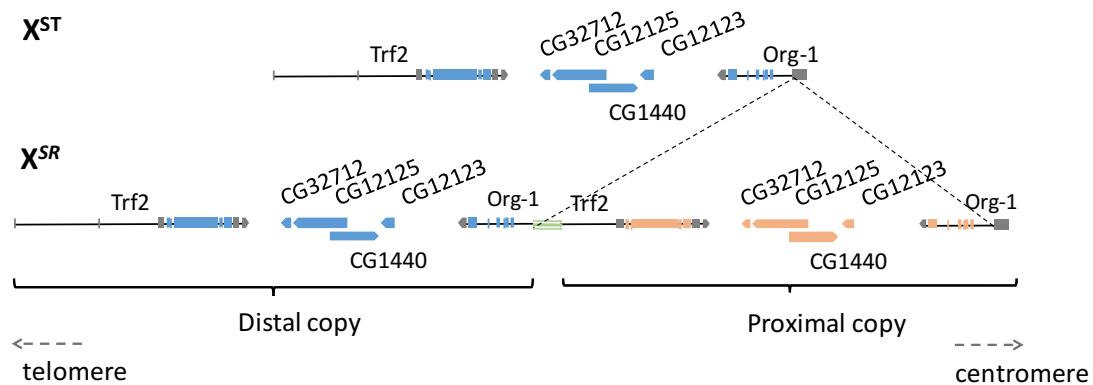


average sex ratio (the proportion of females) of the progeny of individual males as high as 91% (Atlan et al. 1997). A previous study showed that *SR* strains produce much more X spermatids than Y spermatids during spermatogenesis, while the wild-type standard (ST) strains produce equal numbers of X and Y spermatids during spermatogenesis. Moreover, there is nondisjunction of Y chromatids in meiosis II in *SR* strains, leading to abnormal Y chromosome segregation and less Y spermatids production (Cazemajor et al. 2000). These results suggest that the Paris *SR* system bears an *SR* meiotic drive on the X chromosome.

From previous recombinant mapping experiment of the major *SR* drive genetic loci, there are two major loci (Fig. 1). The primary locus was found to contain at least two required interacting elements on the X chromosome. The first element in the *SR* strains contains a tandem duplication of six genes (*Trf2*, *CG32712*, *CG12125*, *CG1440*, *CG12123*, and *org-1*) (Fig. 2) (Montchamp-Moreau et al. 2006). The second element in the *SR* strains is a truncated allele of *HP1D2*, a member of the Heterochromatin Protein 1 (HP1) gene family (Helleu et al. 2016). It is known that many genes are involved in the Paris *SR* system, but only *HP1D2* was identified. *Trf2* in the first element is a transcription factor gene (Hochheimer et al. 2002; Andersen et al. 2017), which is



**Figure 1. The two major drive loci of the Paris *SR* system.** Element 1 is a tandem duplication of six genes, while element 2 is *HP1D2*. The two elements are 110 kb apart.



**Figure 2. Element 1 of the *SR* drive locus.** There is only one copy for *Trf2*, *CG32712*, *CG12125*, *CG1440*, *CG12123*, and *org-1* in ST strains, while there are two copies for the six genes in SR strains.

partially duplicated in the *SR* strains. This raises a possibility that the gene expression level of some genes will be altered, leading to the *SR* phenotype.

In this study, to systematically identify other *SR*-related genes, the transcriptomic differences of testicular expression between the ST strains (XMa23.1, XDz2.2, and XSe3) and *SR* strains (XTa4, Rf50, and XSR7) of the Paris *SR* system were compared.

After identifying these *SR*-related genes, we may be able to understand the possible mechanisms underlying the sex chromosome meiotic drive.



## Materials and Methods



### *D. simulans* stocks

#### Standard ST8 stock

*Standard* (ST) is a standard reference stock without *sex-ratio* distorters, drive suppressors, and any cytoplasmic parasite. The stock was collected in Nasr'allah, Tunisia in 1983 (Mercot et al. 1995a). The sex ratio of this stock is about 52.4% (Cazemajor et al. 1997). ST8 is a highly inbred standard stock, derived from the ST stock with 20 generations of sib-pair mating.

#### Attached-X ST8 stock

Attached-X ST8 represents ST8/C(1)RM *y w*, a stock in which the females carry the attached-X chromosome from the *l<sub>z</sub>[sp]/C(1)RM y w* stock (Bloomington Stock Center, Indiana University) in a standard ST8 background (Montchamp-Moreau et al. 2001).



### Tested strains

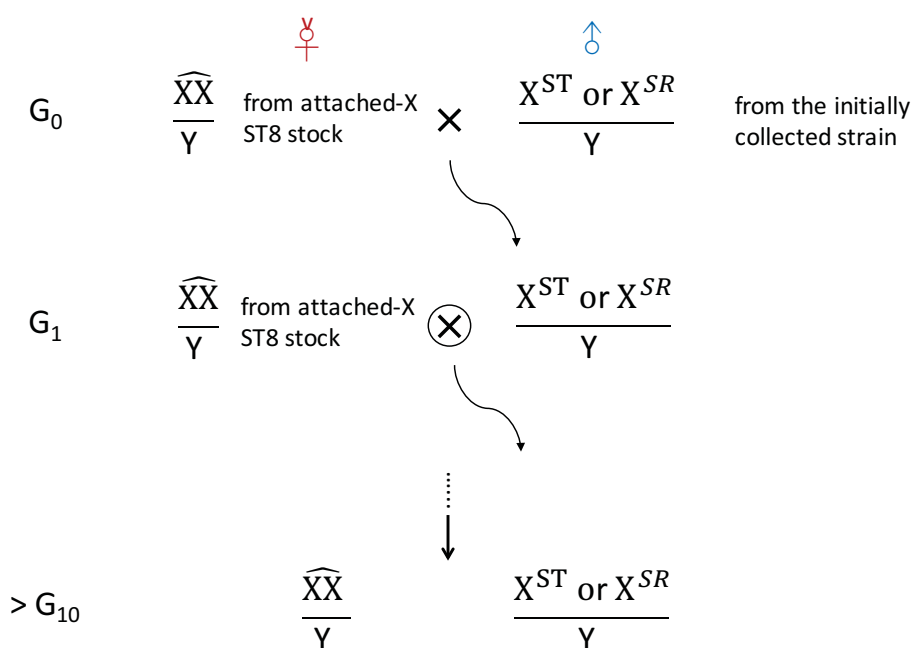
All tested strains, including wild-type standard (ST) and *sex-ratio* (SR), were initially collected from different locations or used in the genetic mapping of *sex-ratio* distorters (Table 1) (Helleu et al. 2016). Then, the X chromosomes of the tested strains were isolated by repeated backcrosses of the males with ST8/C(1)RM *y w*, the attached-X chromosome-carrying females for more than 10 generations to get the X

**Table 1. *Drosophila simulans* strains used in this study**

Phenotype	Name	Origin
ST	XDz2.2	Mayotte
	XMa23.1	Madagascar
	XSe3	Seychelles
	Xsn+5	genetic mapping
SR	Rf50	Mayotte
	XDi6	Comores
	Xsn+13	genetic mapping
	XSR6	Seychelles
	XSR7	Seychelles
	XTa4	Madagascar
	XVou8	Comores

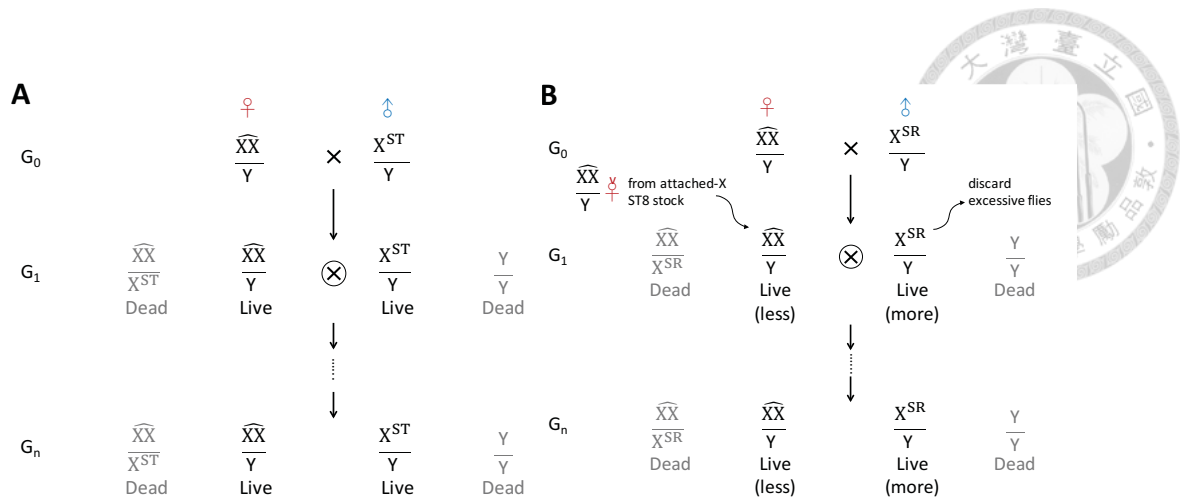


chromosome-isolated strains under attached-X ST8 background (Fig. 3). The maintenance of the strains is shown in Fig. 4. Progeny with three X chromosomes or that with two Y chromosomes die, so females in these strains all carry the ST8/C(1)RM,  $y, w$  attached-X chromosomes and males all bear an X chromosome isolated from the initially collected strains. In *Drosophila* males, because there is no chromosome recombination, the X chromosomes can remain intact in males.



**Figure 3. The crossing scheme of X chromosome isolation.** The male from the initially collected strain was crossed with the virgin females from the attached-X ST8 stock. The male progeny of the G<sub>0</sub> were then repeatedly backcrossed with the virgin females from the attached-X ST8 stock for more than 10 generations.



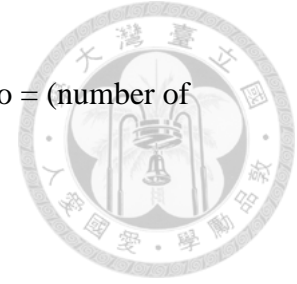


**Figure 4. The maintenance of ST and SR strains.** (A) ST strains were kept with females carrying the attached-X and males bearing the  $X^{ST}$ . (B) SR strains were kept by adding virgin females from attached-X ST8 stock and discarding excessive males for each generation, keeping the  $X^{SR}$  in males and attached-X in females.

### *Sex ratio phenotype assays*

The sex ratio of each tested strain was measured by crossing one 3- to 5-day-old male with three 3- to 5-day-old standard ST8 virgin females with three replicate experiments. Also, the sex ratio of each tested strain under attached-X ST8 background was measured by crossing one 3- to 5-day-old male with three 3- to 5-day-old attached-X ST8 virgin females with three replicate experiments. For each cross, the pair was transferred into a new vial twice a week for two weeks. The progenies were all sexed and counted until no more flies emerged. Only crosses producing more than 50 flies were considered.

The average sex ratio was calculated as below: average sex ratio = (number of females / total number of flies)  $\times$  100%  $\pm$  standard deviation



### ***HP1D2 genotyping***

Genomic DNA of a single male from the tested strains was extracted from the whole fly. PCR was performed using the primer pair (HP1D2-234F: 5'-CACTATACGATGAAAGCGAGCAC-3' and HP1D2-1661R: 5'-TAACCGAAAGCCTATGGACACAC-3') to distinguish between the *HP1D2<sup>ST</sup>* allele (PCR product: 1428 bp) and the *HP1D2<sup>SR</sup>* allele (PCR product: 1057 bp). PCR conditions were as followed: 35 cycles at 95°C for 30 s, 60°C for 30 s, and 72°C for 2 min 9 sec.

### ***Transcriptomic analyses***

#### **RNA sequencing experimental design**

RNA sequencing of the testes of wild-type (ST) and *sex-ratio* (SR) strains were performed to identify candidate genes associated with the SR phenotype. ST strains include XMa23.1, XDz2.2, and XSe3, while SR strains include XTa4, Rf50, and XSR7.

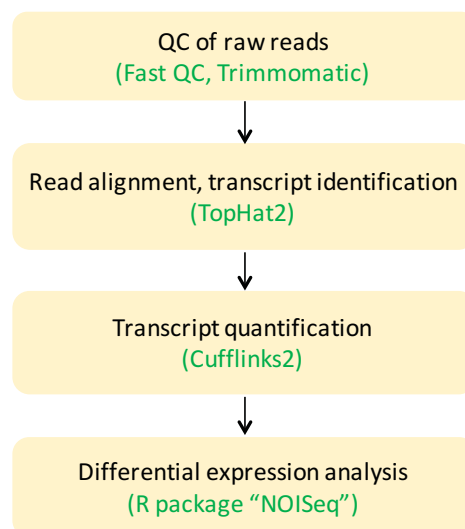


For the details of RNA sequencing sample preparation, please see Appendix A.


### RNA sequencing data analyses

For the RNA sequencing (RNA-Seq) data analyses, it was performed under four steps: Quality control (QC) of raw reads, transcript identification, transcript quantification, and differential expression analysis (Fig. 5) (Conesa et al. 2016).

For QC of raw reads, the quality of reads was evaluated using FastQC (version 0.11.5) (Andrews et al. 2010), while the adapter trimming and low-quality bases and reads removal were done using Trimmomatic (version 0.36) (Bolger et al. 2014). The quality of a base is calculated as below: Phred quality score ( $Q$ ) =  $-10 \log_{10}P$ , in

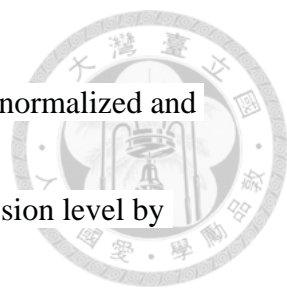


**Figure 5. The RNA-Seq data analyses pipeline.** The main steps of RNA-Seq data analyses include quality control (QC) of raw reads, read alignment and transcript identification, transcript quantification, and differential expression analysis.



which  $P$  represents the base-calling error probability. The trimming procedure are as follows: (1) Remove adaptors; (2) Remove leading 10 bases; (3) Remove leading low-quality bases with quality below 10; (4) Remove trailing low-quality bases with quality below 10; (5) Perform a 4-base sliding window scanning, cutting when the average quality within the window drops below 15; (6) Remove reads with length below 50 bases.

For transcript identification, the *D. simulans* genome (flybase release version FB2016\_05) and its gene annotations were downloaded from flybase FTP. The read alignment and transcript identification were done by TopHat2 (version 2.1.1) (Kim et al. 2013) and its embedded Bowtie2 (version 2.2.6) (Langmead et al. 2009). The parameters used for read alignment are as follows: (1) Final read alignments with more than two mismatches were discarded; (2) Final read alignments with more than two edit distance were discarded. After read alignment, the mapped reads of two housekeeper genes *GD17524* (an ortholog of *eIF2B- $\alpha$*  in *D. melanogaster*) and *GD18948* (and ortholog of *rpII140* in *D. melanogaster*) were also visualized using Integrative Genomics Viewer (IGV) (Robinson et al. 2011; Thorvaldsdottir et al. 2013) to examine whether the number of mapped reads were similar.



For transcript quantification, the number of mapped reads were normalized and quantified as RPKM (Mortazavi et al. 2008) to represent gene expression level by Cufflinks (version 2.2.1) (Trapnell et al. 2010; Roberts et al. 2011). The parameters used were mostly in default except that the maximum number of fragments a locus may have is adjusted to  $10^{12}$ .

For differential expression analysis, three ST strains XMa23.1, XDz2.2, and XSe3 were used as ST biological replicates, while three SR strains XTa4, Rf50, and XSR7 were used as SR biological replicates. Gene expressions between ST and SR strains were compared. Only genes with RPKM values higher than 1 in at least one sample were classified as expressed genes and preserved for the following analyses. Batch effect correction and differential expression analysis were computed by the R package, “NOISeq” (Tarazona et al. 2011; Tarazona et al. 2015). The probability of differential expression is equivalent to  $1 - FDR$  (Benjamini-Hochberg adjusted  $p$ -value). Only the genes with  $|\log_2(\text{Fold change})|$  higher than 0.6 were defined as differential expressed genes (DEGs) and performed with following functional annotation analyses.



### Clustering analysis

After batch effect correction, all expressed genes were first z-score transformed.

Then, the dissimilarity between samples was calculated as below:

Dissimilarity between samples = 1 - correlation


The samples were then hierarchically clustered by complete linkage clustering in the R package, “stats” (R Core Team 2017). The cut-off for the clustering of samples was set to be 1.2. The heatmap was plotted by the R package, “gplots”.

### Phylogenetic tree construction

After transcript identification, the consensus sequences of the X chromosome transcriptome of all six strains were obtained by using the samtools pipeline according to the samtools-1.9 manual (Li et al. 2009). The six consensus sequences were then performed with the neighbor-joining phylogenetic tree construction by muscle3.8.31 (Edgar 2004).

### Functional annotation analysis


To search for the functions in which the DEGs are enriched, functional annotation



analysis of the DEG list was done by a web-based tool DAVID (Huang et al. 2009). The dataset of *D. melanogaster* is much more complete than *D. simulans*, so the flybase gene ID of the DEGs were converted to the Flybase gene ID of their *D. melanogaster* homologs and input into DAVID. The gene ontology (GO) terms, including biological process, cellular component, and molecular function of the gene list were analyzed. The *p*-value of the gene annotation analysis was corrected by Benjamini-Hochberg false discovery rate. Only the GO terms with the adjusted *p*-value smaller than 0.05 were listed.

### ***Measuring testicular gene expression***

To validate the *SR* candidate genes, the testicular gene expression level of the three ST strains XMa23.1, XDz2.2, and XSe3, and the three *SR* strains XTa4, Rf50, and XSR7 were measured by RT-qPCR (Reverse transcription-quantitative PCR). For each strain, at least 60 pairs of testes were dissected in PBS from males less than 5-day old. RNA extractions were conducted using the TRIzol™ Reagent (Thermo Fisher Scientific, Waltham, MA, USA) according to their protocol. The extracted RNA was then treated with Deoxyribonuclease I, Amplification Grade (Thermo Fisher Scientific)



to digest genomic DNA if present in the sample. Reverse transcription was performed using SuperScript<sup>®</sup> III First-Strand Synthesis System for RT-PCR (Thermo Fisher Scientific). Complementary DNA quantification was done using iQ<sup>™</sup> SYBR<sup>®</sup> Green Supermix (Bio-Rad, Hercules, CA, USA) in a CFX96 Touch<sup>™</sup> Real-Time PCR Detection System or a CFX Connect<sup>™</sup> Real-Time PCR Detection System (Bio-Rad). For each sample, gene expression was measured with three technical replicates using one or two reference genes *rpIII140*, *eIF2B-α*, *light*, and *Act5C* as internal controls. Differences in the expression level between XMa23.1 and XTa4, XDz2.2 and Rf50, and XSe3 and XSR7 strains were tested with the unpaired Student's t-test (CFX Manager<sup>™</sup> software, Bio-Rad).



## Results



### *Sex ratio phenotype data of ST and SR strains*

To examine the sex ratio phenotype of all wild-type standard (ST) and *sex-ratio* (SR) strains, I crossed single males of each tested strains to ST8 virgin females and counted their progeny. XMa23.1, XDz2.2, XSe3, and Xsn+5 showed around 50% average sex ratio. This indicated that these four strains were the ST phenotype and that these strains were either free of meiotic drivers or lacking some components causing sex chromosome meiotic drive. XTa4, Rf50, XSR7, XSR6, XDi6, XVou8, and Xsn+13 showed an average sex ratio higher than 87%. The highest sex ratio was observed in XDi6, with 96% average sex ratio. This result confirmed that these strains were SR phenotype, with meiotic drivers located on their X chromosomes (Table 2).

As all these strains were maintained under attached-X ST8 background, I also examined the sex ratio of all strains by crossing single males of each tested strains to attached-X ST8 virgin females and counted their progeny. If the sex ratio under attached-X ST8 background is consistent with that under ST8 background, the average sex ratio of a particular strain when crossed to attached-X ST8 virgin females should be  $100 - (\text{average sex ratio when crossed to ST8 virgin females})$ . ST strains XMa23.1,



XDz2.2, XSe3, and Xsn+5 showed an average sex ratio close to 50%, consistent with the previous phenotype assays by crossing to ST8 virgins. *SR* strains XTa4, Rf50, XSR7, XSR6, XDi6, XVou8, and Xsn+13 also showed consistent average sex ratio with the previous phenotype assays by crossing to ST8 virgins (Table 2).

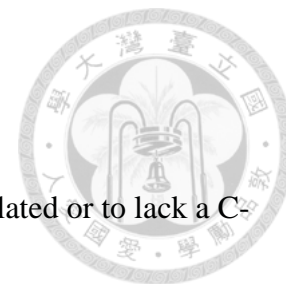
**Table 2. Sex ratio phenotype assays of ST and *SR* strains**

G0 male	G1 progeny average sex ratio (%) <sup>a</sup>	
	Females from ST8 stock:	Females from attached-X ST8 stock:
ST		
XMa23.1	52.91 ± 3.63	45.34 ± 10.41
XDz2.2	55.84 ± 1.28	48.50 ± 4.33
XSe3	52.24 ± 2.15	45.75 ± 4.20
Xsn+5	52.30 ± 5.69	49.18 ± 7.68
<i>SR</i>		
XTa4	95.02 ± 2.26	5.34 ± 2.83
Rf50	91.01 ± 2.78	7.82 ± 1.23
XSR7	89.19 ± 7.18	11.48 ± 6.36
XSR6	87.03 ± 3.06	20.36 ± 6.75
XDi6	96.03 ± 1.15	3.71 ± 1.70
XVou8	94.38 ± 1.91	3.90 ± 1.83
Xsn+13	93.60 ± 2.28	6.15 ± 1.27

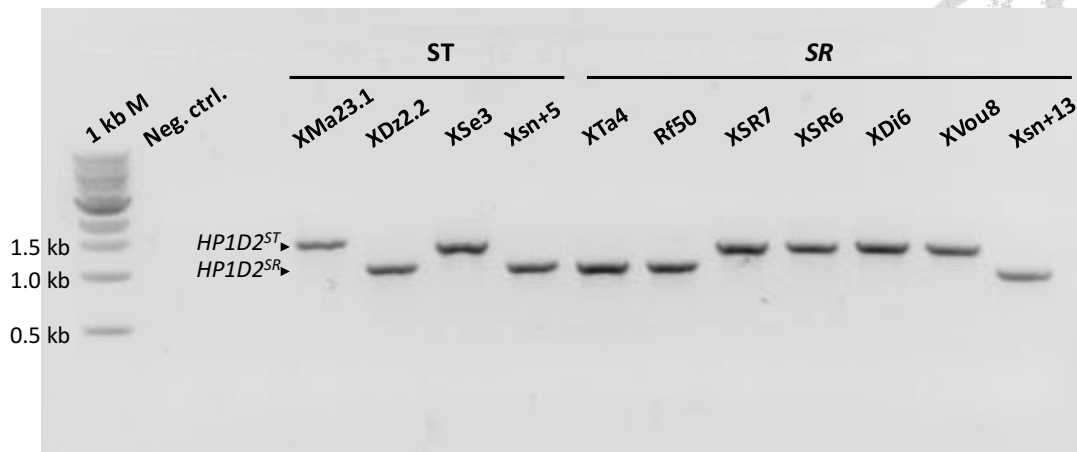
<sup>a</sup>average sex ratio = (number of females / total number of flies) × 100% ± standard deviation

Note: A single 3- to 5-day-old male was mated with three 3- to 5-day-old virgin females from the standard ST8 stock or attached-X ST8 stock. The progenies were sexed and counted. ST, wild-type standard strain; *SR*, *sex-ratio* strain.

### ***HP1D2* genotyping of ST and SR strains**



A previous study on SR have revealed *HP1D2* to be down-regulated or to lack a C-terminal chromo shadow domain (CSD), leading to the SR phenotype (Helleu et al. 2016). To examine the genotypes of *HP1D2* of all tested strains, I performed genotyping for *HP1D2* of ST and SR strains. ST strains XMa23.1, XDz2.2, XSe3, and Xsn+5 showed variation in the *HP1D2* genotype. XMa23.1 and XSe3 carried the *HP1D2<sup>ST</sup>* allele, while XDz2.2 and Xsn+5 carried the *HP1D2<sup>SR</sup>* allele. SR strains XTa4, Rf50, XSR7, XSR6, XDi6, XVou8, and Xsn+13 also showed variation in the *HP1D2* genotype. XSR7, XSR6, XDi6, and XVou8 carried *HP1D2<sup>ST</sup>* allele, while XTa4, Rf50, and Xsn+13 carried *HP1D2<sup>SR</sup>* allele (Fig. 6 and Table 3). The results indicated that there exists variation in the *HP1D2* genotype of both ST and SR strains. The *HP1D2<sup>SR</sup>* allele did not necessarily lead to the SR phenotype.



**Figure 6.** Variation existing in the *HP1D2* genotype of both ST and SR strains. The arrowheads indicate the full-length *HP1D2*<sup>ST</sup> allele (1428 bp) or the *HP1D2*<sup>SR</sup> allele lacking chromo shadow domain (CSD) (1057 bp).

**Table 3.** The SR phenotype and *HP1D2* genotype of ST and SR strains

Phenotype	Strains	<i>HP1D2</i> genotyping result
ST	XMa23.1	<i>HP1D2</i> <sup>ST</sup>
	XDz2.2	<i>HP1D2</i> <sup>SR</sup>
	XSe3	<i>HP1D2</i> <sup>ST</sup>
	Xsn+5	<i>HP1D2</i> <sup>SR</sup>
SR	XTa4	<i>HP1D2</i> <sup>SR</sup>
	Rf50	<i>HP1D2</i> <sup>SR</sup>
	XSR7	<i>HP1D2</i> <sup>ST</sup>
	XSR6	<i>HP1D2</i> <sup>ST</sup>
	XDi6	<i>HP1D2</i> <sup>ST</sup>
	XVou8	<i>HP1D2</i> <sup>ST</sup>
	Xsn+13	<i>HP1D2</i> <sup>SR</sup>

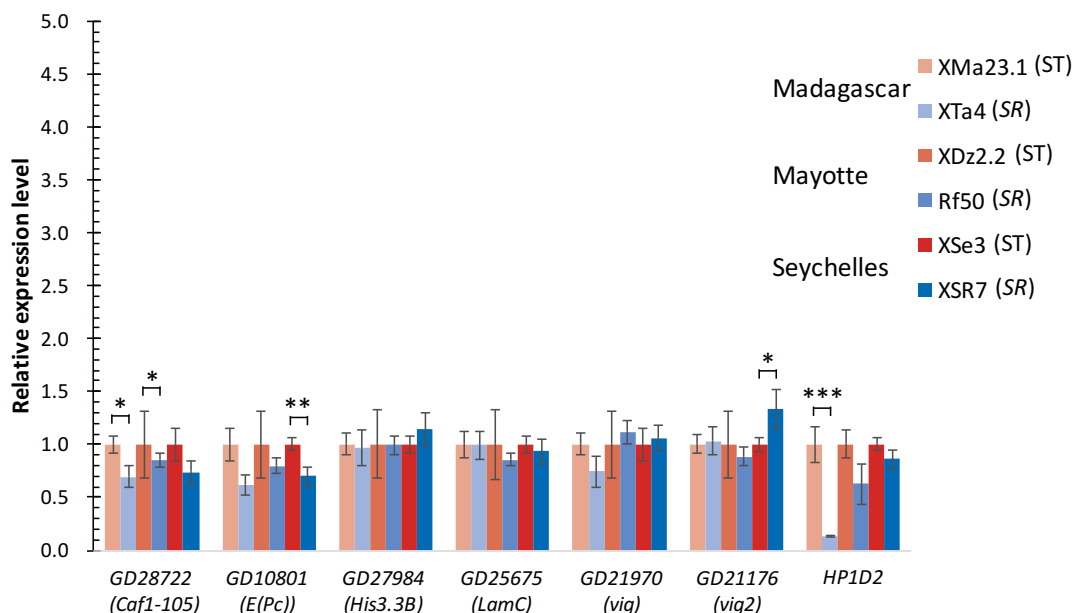
***Testicular expression of genes related to heterochromatin organization and chromatin assembly***



Previous studies on *SR* have shown that there was nondisjunction of Y chromatids in meiosis II in *SR* strains, causing abnormal Y chromosome segregation and less Y spermatids production (Cazemajor et al. 2000). In addition, *HP1D2* has previously been shown to be one of the *SR* meiotic drive genes. *HP1D2* is a member of the Heterochromatin Protein 1 (HP1) gene family. Dysfunction of *HP1D2*, either its down-regulation or lack of a C-terminal chromo shadow domain (CSD), leads to the *SR* phenotype (Helleu et al. 2016). It interferes with the Y chromosome segregation (Helleu et al. 2016) and may also be involved in heterochromatin organization (Levine et al. 2012). This raises a possibility that genes related to heterochromatin organization or chromatin assembly are involved in the *SR* phenotype. To examine whether these genes are related to the *SR* phenotype, the testicular expression of *Caf1-105*, *E(Pc)*, *His3.3B*, *LamC*, *vig*, *vig2*, and also *HP1D2* were compared between ST and *SR* strains. *Caf1-105* is down-regulated in *SR* strains collected from two of the three locations. *E(Pc)* is down-regulated in one of the *SR* strains. *vig2* is up-regulated in one of the *SR* strains. However, *His3.3B*, *LamC*, and *vig* showed no differences in gene expression between



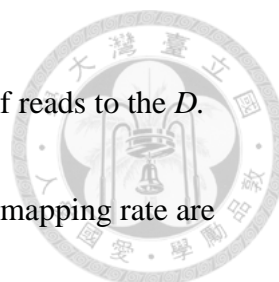
ST and *SR* strains. In addition, *HP1D2*, which was expected to be down-regulated in *SR* strains, only showed down-regulation in one of the *SR* strains (Fig. 7).



**Figure 7. Testicular expression of genes related to heterochromatin organization and chromatin assembly.** RT-qPCR analysis of testicular expression of genes in three ST and *SR* strains collected from three locations. XMa23.1 and XTa4 were collected from Madagascar; XDz2.2 and Rf50 were collected from Mayotte; XSe3 and XSR7 were collected from Seychelles. Gene names in the parentheses are *Drosophila melanogaster* homologs. Error bars represent SEMs. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (unpaired Student's t-test).

### *Transcriptomic analyses of ST and SR strains*

To systematically identify *SR*-related genes, I compared the transcriptomic differences of testicular gene expression between three ST and three *SR* strains. ST strains included XMa23.1, XDz2.2, and XSe3, while *SR* strains included XTa4, Rf50,



and XSR7. For the trimming of raw sequencing reads and mapping of reads to the *D. simulans* reference genome, the reads percentage after trimming and mapping rate are shown in Table 4. The results that the reads after trimming in all samples were above 92.7% and that the mapping rate of all samples were above 86.3% indicated that the quality of the reads was good. After transcript quantification, among the 15,476

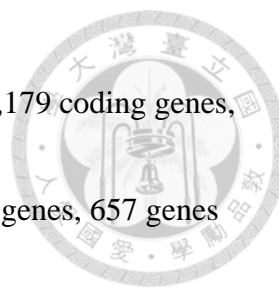
**Table 4. The left reads after trimming and mapping rate of RNA-Seq reads**

Sample	Number of raw reads	Number of trimmed reads	Number of mapped reads	Left reads after trimming (%) <sup>a</sup>	Mapping rate (%) <sup>b</sup>
XMa23.1	14,106,265	13,117,565	11,560,334	93.0	88.1
XDz2.2	13,134,515	12,225,070	10,726,776	93.1	87.7
XSe3	15,634,724	14,597,124	12,812,341	93.4	87.8
XTa4	16,030,685	14,899,847	12,860,067	92.9	86.3
Rf50	12,850,322	11,908,363	10,418,178	92.7	87.5
XSR7	15,772,634	14,731,250	12,887,315	93.4	87.5

<sup>a</sup>Left reads after trimming (%) = (number of trimmed reads / number of raw reads) × 100%

<sup>b</sup>Mapping rate (%) = (number of mapped reads / number of trimmed reads) × 100%

Note: The RNA-Seq raw reads of three ST strains, XMa23.1, XDz2.2, and XSe3, and three SR strains, XTa4, Rf50, and XSR7, were trimmed and mapped to the *Drosophila simulans* reference genome.



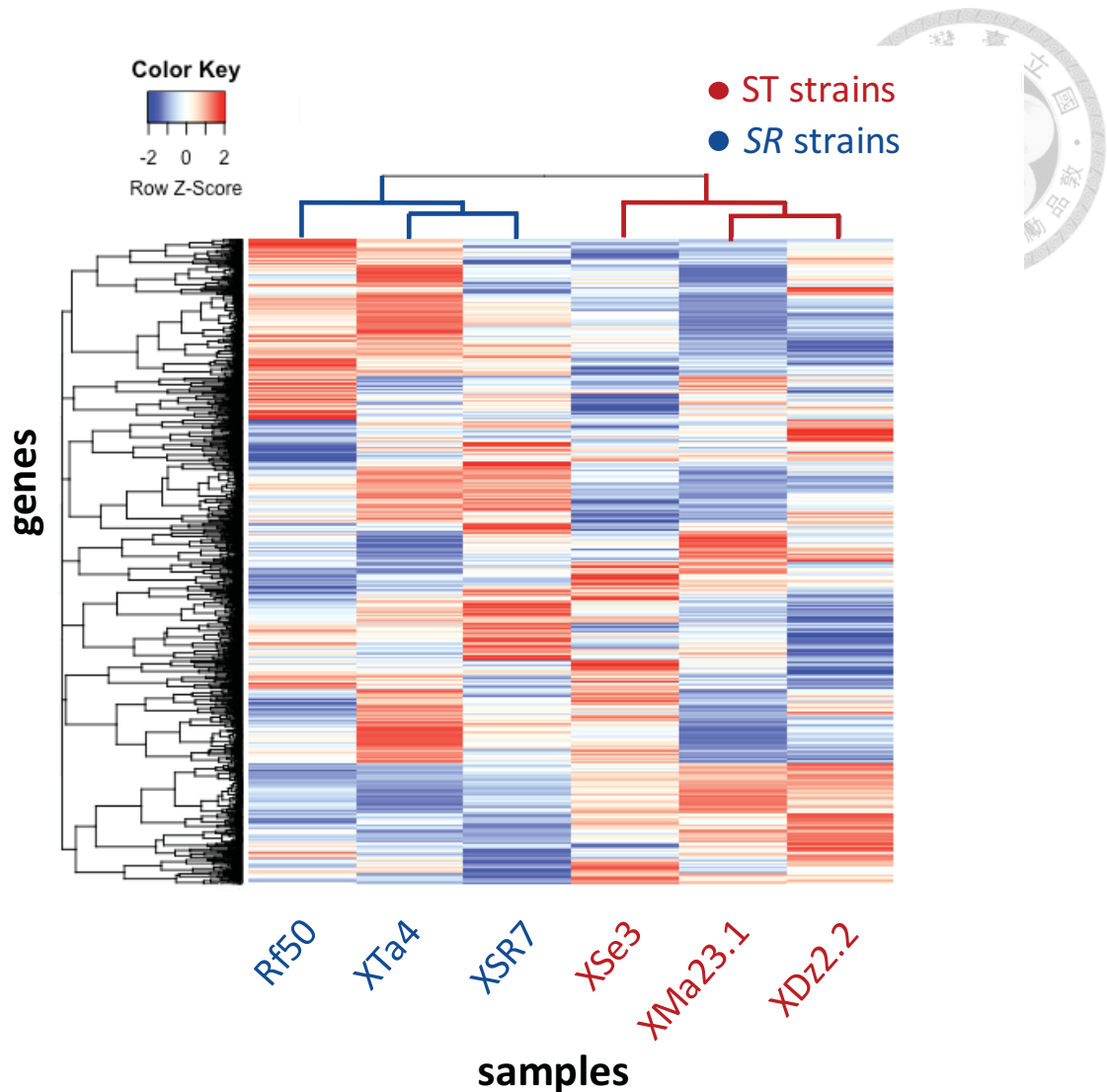
annotated genes, 11,298 genes (73%) were expressed. Among the 14,179 coding genes, 10,773 genes (76.0%) were expressed. Among the 1,185 non-coding genes, 657 genes (55.4%) were expressed (Table 5). After low-count filtering, I performed batch effect correction on the samples because there were two batches in the RNA sequencing process. XMa23.1 and XTa4 were in the same batch, while XDz2.2, XSe3, Rf50, and XSR7 were in the other batch.

**Table 5. The proportion of expressed genes among all genes**

	<b>Coding genes</b>	<b>Non-coding genes</b>	<b>All genes</b>
Expressed genes (RPKM > 1 in at least one sample)	10,773	657	11,298
Total	14,179	1,185	15,476
Proportion of expressed genes	76.0%	55.4%	73.0%


To examine whether between-group differences were larger than within-group differences, I performed hierarchical clustering of all six samples. The results showed that all three ST strains were clustered together and that all three SR strains were clustered together (Fig. 8). This indicated that the phenotype can be revealed by the





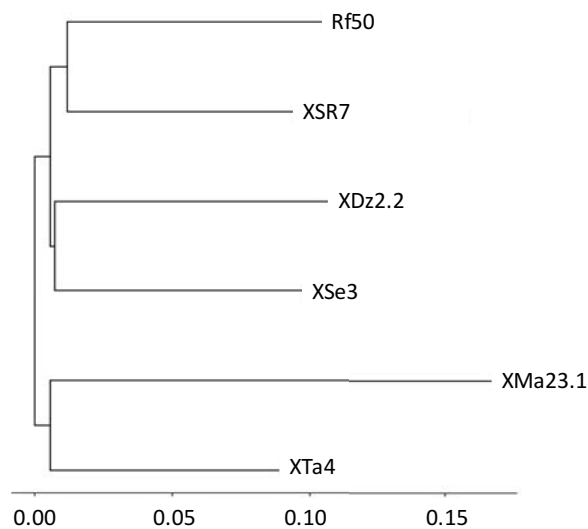
**Figure 8. ST and SR strains are in two different clusters.** Heatmap of gene expression profiles of ST and SR samples was plotted. Up-regulated genes are indicated as red, while down-regulated genes are indicated as blue. Both genes and samples were hierarchically clustered.

gene expression pattern. I also plotted a heatmap of all gene expression level of six samples (Fig. 8), with hierarchical clustering by both samples and genes. All six samples showed different clusters of up-regulated (shown in red) and down-regulated (shown in blue) genes. When examining the gene expression pattern similarity within ST strains or within SR strains, there were no apparent pattern from the transcriptome of



the same phenotype. However, if we compared the transcriptomes of the *ST* and *SR* samples pairwise according to their geographical locations, we can observe some patterns. For example, XMa23.1 and XTa4 were collected from the same geographical location. From this pair of transcriptome, we can see a completely opposite gene expression pattern. Those genes which were up-regulated in XMa23.1 were often down-regulated in XTa4, while genes which were down-regulated in XMa23.1 were often up-regulated in XTa4. This pattern also appeared between XDz2.2 and Rf50 as well as XSe3 and XSR7. The absence of apparent gene expression pattern within the same phenotype and pairwise opposite gene expression pattern between *ST* and *SR* strain of the same geographical location indicated that there may be different genes which were involved in *SR* among different geographical locations.

To examine whether the phylogenetic relationship of the three *ST* and three *SR* strains is grouped by geographical locations, I constructed a phylogenetic tree by the X chromosome transcriptome sequences of the six strains. Strains collected from Madagascar (XMa23.1 and XTa4) are grouped together and are more distant to those collected from Mayotte and Seychelles. For strains from Mayotte and Seychelles, *ST* and *SR* are in two different clusters (Fig. 9).



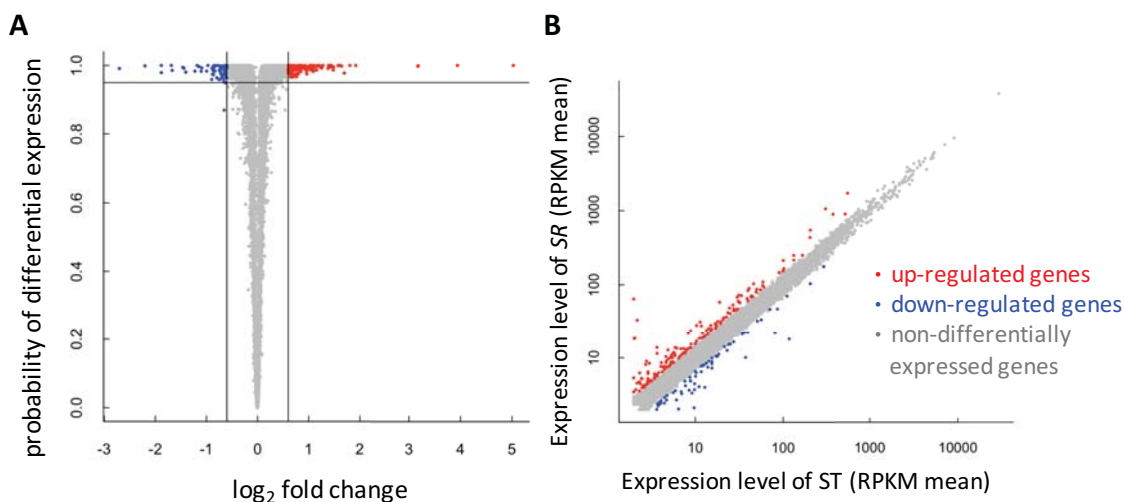
**Figure 9. Neighbor-joining tree of the X chromosome transcriptome sequences of three ST and three SR strains.** The X chromosome transcriptome sequences of strains collected from Madagascar (XMa23.1 (ST) and XTa4 (SR)), Mayotte (XDz2.2 (ST) and Rf50 (SR)), and Seychelles (XSe3 (ST) and XSR7 (SR)) were obtained and performed with the phylogenetic tree (neighbor-joining tree) construction.

Although there may be different groups of genes regulating *SR* phenotype, due to the limited number of samples, I treated three ST strains, XMa23.1, XDz2.2, and XSe3, as biological replicates, and three *SR* strains, XTa4, Rf50, and XSR7, as biological replicates as well. To identify *SR* candidate genes, I performed differential expression analysis. I compared gene expressions between ST and *SR* strains and picked out differentially expressed genes (DEGs) with  $|\log_2(\text{Fold change})|$  higher than 0.6 as *SR* candidate genes (Fig. 10). There were 221 *SR* candidates in total, with 154 up-regulated and 67 down-regulated. Among the 211 *SR* candidate genes, proportion of *SR* candidate genes on the X chromosome reaches 5.7%, higher than that on any other chromosomes



(Table 6). These 211 *SR* candidate genes are then performed with the Gene Ontology

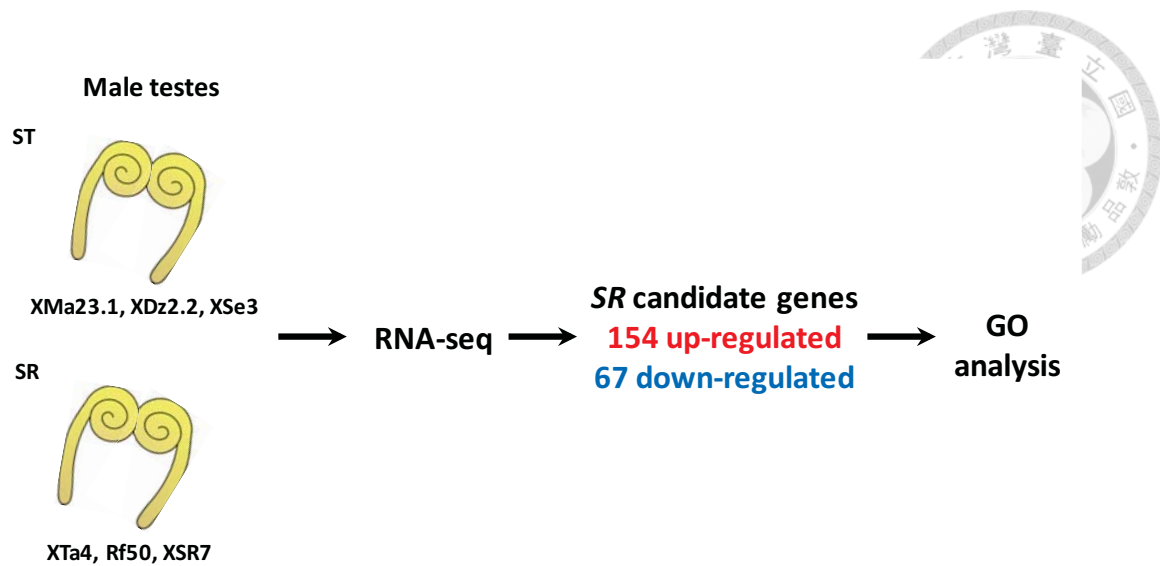
analysis (GO analysis) (Fig. 11).



**Figure 10. Up-regulated genes were more than down-regulated genes in *SR* strains.** The genes up-regulated in *SR* strains are indicated as red, while those down-regulated are indicated as blue. (A) The probability of differential expression (equivalent to  $1 - FDR$ ) was plotted against  $\log_2(\text{Fold change})$ . (B) The average expression level of *SR* strains for each gene was plotted against that of *ST* strains (shown in gray).

**Table 6. The proportion of *SR* candidate genes on each chromosome**

chromosome	X	2L	2R	3L	3R	4	others	total
<i>SR</i> candidate	82	33	40	24	35	0	7	221
total	1441	2239	2413	2392	3058	53	3880	15476
Proportion of <i>SR</i> candidate (%)	5.7	1.5	1.7	1.0	1.1	0	0.2	1.4

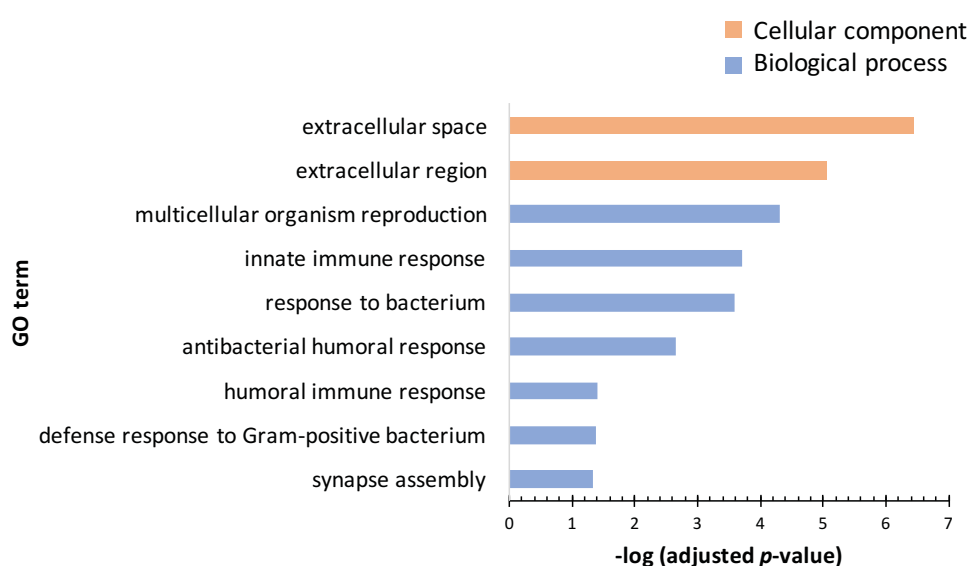


**Figure 11. Summary of the experimental design and *SR* candidate genes of transcriptomic analyses.** The male testes of three ST strains, XMa23.1, XDz2.2, and XSe3, and three *SR* strains, XTa4, Rf50, and XSR7, were dissected and performed with RNA-Seq. In total, there were 221 *SR* candidate genes (154 genes up-regulated and 67 genes down-regulated) in *SR* strains. These *SR* candidates were further performed with Gene Ontology (GO) analysis.

To further understand the functions of these *SR* candidate genes, I did a functional annotation analysis, GO analysis. As the database of *D. melanogaster* was more complete than that of *D. simulans*, I converted the flybase gene ID of the *SR* candidate genes to the Flybase gene ID of their *D. melanogaster* homologs. I performed GO analysis using *D. melanogaster* homologs of *SR* candidate genes. The results of GO analysis, including biological process, cellular component, and molecular function, were shown in Fig. 12. Only GO terms with Benjamini-Hochberg (BH)-adjusted *p*-value lower than 0.05 were listed. No significantly enriched terms was found in the category




of molecular function. For the category of biological process, there was an enrichment of genes related to multicellular organism reproduction, immune response, and synapse assembly. For the category of cellular component, genes located in extracellular space and region were also enriched.



**Figure 12. Gene ontology enrichment results of *SR* candidate genes.** The gene ontology terms, including biological process, cellular component, and molecular function, were inferred from the *SR* candidate gene list. The  $p$ -value of the gene annotation analysis was corrected by Benjamini-Hochberg false discovery rate. Only the GO terms with the adjusted  $p$ -value smaller than 0.05 were listed.

### *qPCR validation of SR candidate genes*

To validate the *SR* candidate genes selected from the RNA-Seq analyses, I set two criteria for choosing genes to perform with qPCR validation. First, among the 11298 expressed genes, I chose genes with average fold change larger than 2 both between and



after batch effect correction. There were 46 genes reaching this criteria, with 39 up-regulated genes and 7 down-regulated genes. I then chose genes with direction of up- or down-regulation consistent in three *SR* strains among the 46 genes. There are 36 genes fulfilling the criteria, with 32 up-regulated genes and 4 down-regulated genes. Among the 36 genes, there were 34 genes being successfully quantified by qPCR in total, with 31 up-regulated genes and 3 down-regulated genes (Table 7, Fig. 13). The six samples were collected from three different locations, with XMa23.1 (ST) and XTa4 (*SR*) from Madagascar, XDz2.2 (ST) and Rf50 (*SR*) from Mayotte, and XSe3 (ST) and XSR7 (*SR*) from Seychelles. Hence, I compared the testicular gene expression of ST and *SR* pairwise according to their location of collection.

There were five genes showing up-regulation in *SR* strains from all three locations, including *CG16772*, *CG15209*, *Ser7*, *CG34265*, and *CG43348* (Fig. 13A). Among these up-regulated genes consistent in all three locations, all five genes have *D. melanogaster* orthologs. *CG16772* is a mating-responsive, immune response gene which is regulated by the sex-determination hierarchy (Ellis and Carney 2010). *CG15209*, *CG34265*, and *CG43348* has no known function. *Ser7* is a serine protease (Ross et al. 2003). Besides these genes, other genes were only validated in one or two

*SR* strains (Fig. 13A-13D), indicating that different *SR* strains may use different sets of genes to cause the *SR* phenotype.




*Anp*, *IM1*, and *IM2* function in immune response or antibacterial response

(Samakovlis et al. 1991; Levy et al. 2004) and showed up-regulation in one of the *SR* strains (Fig. 13A). *Anp* was up-regulated in the *SR* strain from Seychelles, while *IM1* and *IM2* were up-regulated in the *SR* strain from Mayotte. *Tsp42Er* is a tetraspanin gene which serves in the cell surface receptor signaling pathway (Fradkin et al. 2002) and was up-regulated in the *SR* strains collected from Madagascar and Seychelles (Fig. 13A). *CG5402* is a seminal fluid protein transferred at mating in *D. melanogaster*, but is not transferred at mating in *D. simulans* (Findlay et al. 2008). *CG5402* was up-regulated in the *SR* strain from Mayotte (Fig. 13A).

*CG1640* is localized to cytosol and mitochondrion (Lye et al. 2014) with unknown function and was up-regulated in the *SR* strain from Mayotte (Fig. 13B). *CG34454* contains a Kazal domain, which usually serves as a serine protease inhibitor domain and is predicted to be localized to extracellular region and mitochondrion. *CG34454* was up-regulated in the *SR* strains from Madagascar and Mayotte (Fig. 13B). *CG12123* is within the tandem duplication which has been mapped to be one of the *SR* drive genetic





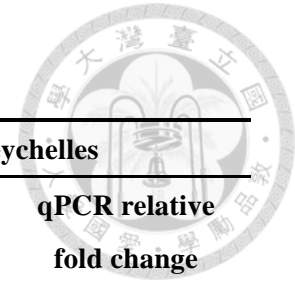
loci (Montchamp-Moreau et al. 2006), but its function is unknown. *CG12123* was up-regulated in the *SR* strain from Mayotte (Fig. 13B). *drm* is a zinc finger transcription factor functioning in developmental patterning and cell fate specification (Iwaki et al. 2001; Hatini et al. 2005) and was up-regulated in the *SR* strains from Madagascar and Seychelles (Fig. 13B). *CG43235* is predicted to function in metallocarboxypeptidase activity and proteolysis and was up-regulated in the *SR* strain from Mayotte (Fig. 13B). *Obp51a* is an odorant binding protein which may also be a seminal fluid protein (Findlay et al. 2008) and was up-regulated in the *SR* strain from Mayotte (Fig. 13B). Among the validated *SR* up-regulated genes, there were a group of genes without *D. melanogaster* orthologs with unknown functions, including *GD28318*, *GD11366*, *GD13193*, *GD17496*, *GD23211*, *GD27485*, *GD28356*, *GD28414*, and *GD28725* (Fig. 13B and 13C). These genes were all validated in one or two of the *SR* strains.

Besides up-regulated genes, there were also down-regulated genes in the *SR* strains. *ana* is a secreted glycoprotein that is expressed in the glial cells which inhibits premature neuroblast proliferation (Ebens et al. 1993). *ana* was down-regulated in the *SR* strains from Madagascar and Seychelles (Fig. 13D). *Ipod* is an interaction partner of

*Mt2* (*CG10692*) (Kunert et al. 2005) and was down-regulated in the *SR* strains from Madagascar and Seychelles (Fig. 13D).



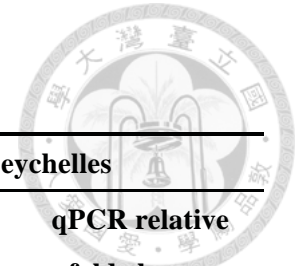
However, there were still some *SR* candidate genes which were not validated by the qPCR (Fig. 13E, Table 7). These genes had either no differential expression in the qPCR results or discrepancy in gene expression from the RNA-Seq results and the qPCR results. For example, *GD27972*, *GD29322*, and *Phk-3* were differentially expressed in the RNA-Seq result but showed no difference in expression in the qPCR result. *GD12342*, *Dup99B*, *GD24501*, and *GD27281* showed discrepancy in RNA-Seq and qPCR results. This indicated that these genes may not be *SR*-related genes because their gene expression was not consistent.



**Table 7. Testicular expression of SR candidate genes (to be continued)**

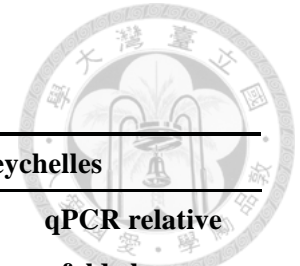
Gene	Chromosome	<i>Drosophila melanogaster</i> ortholog	Madagascar		Mayotte		Seychelles	
			RNA-Seq fold change	qPCR relative fold change	RNA-Seq fold change	qPCR relative fold change	RNA-Seq fold change	qPCR relative fold change
			SR/ST	(SR/ST) ± SEM	SR/ST	(SR/ST) ± SEM	SR/ST	(SR/ST) ± SEM
<i>Anp</i>	3R	<i>Anp</i>	2.44	1.13 ± 0.30	2.8	0.84 ± 0.04*	2.94	2.76 ± 0.25***
<i>GD10455</i>	2R	<i>Tsp42Er</i>	11.68	2.52 ± 0.21**	3.97	0.67 ± 0.09*	4.16	4.91 ± 1.25*
<i>GD10624</i>	2R	<i>ana</i>	0.32	0.33 ± 0.04***	0.34	0.91 ± 0.09	0.1	0.28 ± 0.02***
<i>GD11363</i>	2R	<i>IM1</i>	5.14	1.16 ± 0.24	1.92	2.08 ± 0.13***	4.82	0.90 ± 0.08
<i>GD11365</i>	2R	<i>IM2</i>	2.95	1.13 ± 0.24	1.42	1.32 ± 0.07*	6.75	0.89 ± 0.06*
<i>GD11366</i>	2R	-	5.45	0.95 ± 0.50	1.45	3.90 ± 0.49***	2	1.78 ± 0.30
<i>GD11952</i>	2L	<i>CG1640</i>	10.60/0	0.96 ± 0.13	0.82/0	1.36 ± 0.16*	35.66/0	1.03 ± 0.09
<i>GD12342</i>	3L	<i>Met75Ca,</i> <i>Met75Cb</i>	4.04	0.84 ± 0.09	3.22	0.11 ± 0.01***	3.48	1.44 ± 0.64
<i>GD13193</i>	3L	-	4.48	2.33 ± 0.64	4.32	1.19 ± 0.16	1.14	1.46 ± 0.12*
<i>GD16118</i>	X	<i>CG12123</i>	2.55	1.28 ± 0.17	1.98	1.10 ± 0.11*	2.48	1.31 ± 0.13
<i>GD16989</i>	X	<i>CG15209</i>	3.69	2.75 ± 0.25***	1.27	13.42 ± 1.45***	2.6	2.30 ± 0.18***

\*Note: The values on both sides of the slash in “RNA-Seq fold change SR/ST” indicate the original RPKM (Reads Per Kilobase of transcript per Million mapped reads) value of both ST and SR samples. Asterisks indicate a significant difference between the average relative normalized expression levels of the ST male testes and SR male testes. \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001 (unpaired Student’s t-test). SEM = Standard error of the mean.



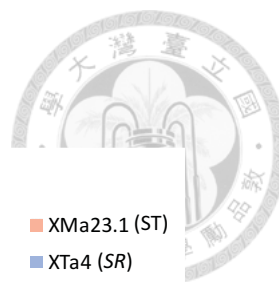
**Table 7. Testicular expression of SR candidate genes (continued)**

Gene	Chromosome	<i>Drosophila melanogaster</i> ortholog	Madagascar		Mayotte		Seychelles	
			RNA-Seq fold change	qPCR relative fold change	RNA-Seq fold change	qPCR relative fold change	RNA-Seq fold change	qPCR relative fold change
			SR/ST	(SR/ST) ± SEM	SR/ST	(SR/ST) ± SEM	SR/ST	(SR/ST) ± SEM
<i>GD17496</i>	X	-	21.58	0.50 ± 0.04*	3.28	1.71 ± 0.21*	1.06	16.57 ± 0.95***
<i>GD17687</i>	3R	<i>Dup99B</i>	2.7	0.68 ± 0.07*	2.65	0.16 ± 0.02*	5.3	1.13 ± 0.10
<i>GD18077</i>	3R	<i>CG5402</i>	2.26	0.93 ± 0.19	4.18	2.19 ± 0.14***	3.85	1.19 ± 0.07
<i>GD19402</i>	3L	<i>CG34265</i>	3.43/0	1.75 ± 0.17**	2.68/0	2.71 ± 0.40**	3.38	3.66 ± 0.23***
<i>GD23211</i>	2L	-	2.33	2.55 ± 0.17***	2.13	1.21 ± 0.13*	3.47	0.18 ± 0.01***
<i>GD23242</i>	2L	<i>drm</i>	16.01	1.33 ± 0.11**	11.91	1.13 ± 0.11	15.48	2.22 ± 0.34*
<i>GD24237</i>	2L	<i>CG16772</i>	26.43	10.34 ± 1.29***	5.67	5.36 ± 0.37***	5.2	4.70 ± 0.80**
<i>GD24501</i>	NODE_53243	<i>CG32588</i> , <i>CG33252</i> , <i>CG43075</i>	1.12	1.33 ± 0.29	33.87/0	0.00 ± 0.00***	1.55	0.44 ± 0.03***
<i>GD24651</i>	X	<i>Ipod</i>	0/10.73	0.00 ± 0.00***	0/0.08	1.25 ± 0.51	0.01	0.00 ± 0.00***
<i>GD24902</i>	2R	<i>Phk-3</i>	1.81	1.21 ± 0.17	2.38	1.47 ± 0.22	2.39	1.55 ± 0.43
<i>GD27028</i>	2L	<i>CG43348</i>	10.31	3.73 ± 0.23***	7.29	3.03 ± 0.25***	5.67	1.50 ± 0.20*

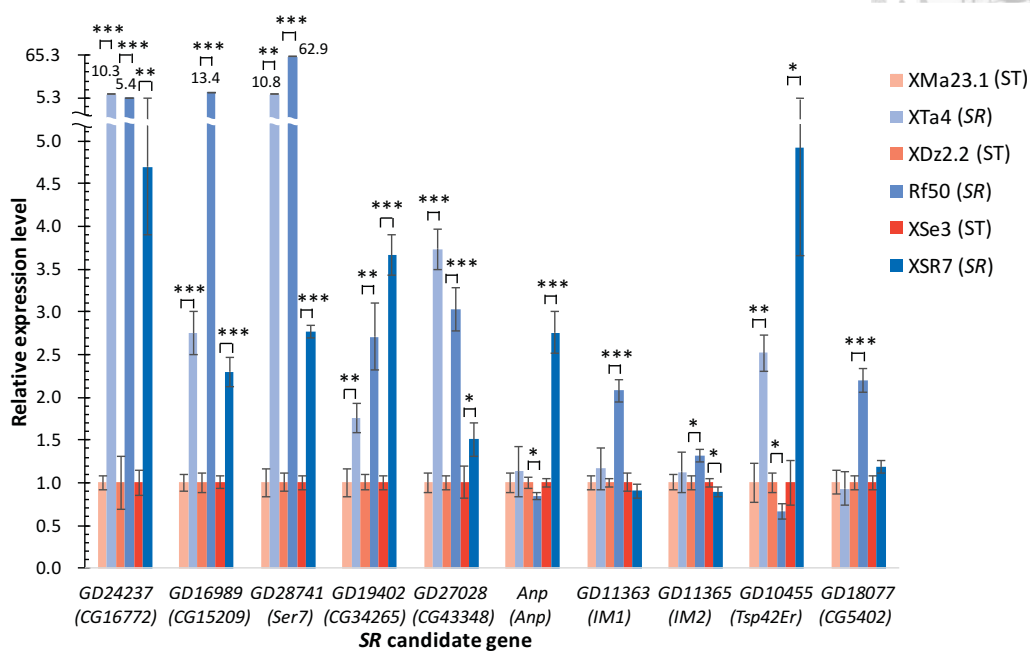


**Table 7. Testicular expression of *SR* candidate genes (continued)**

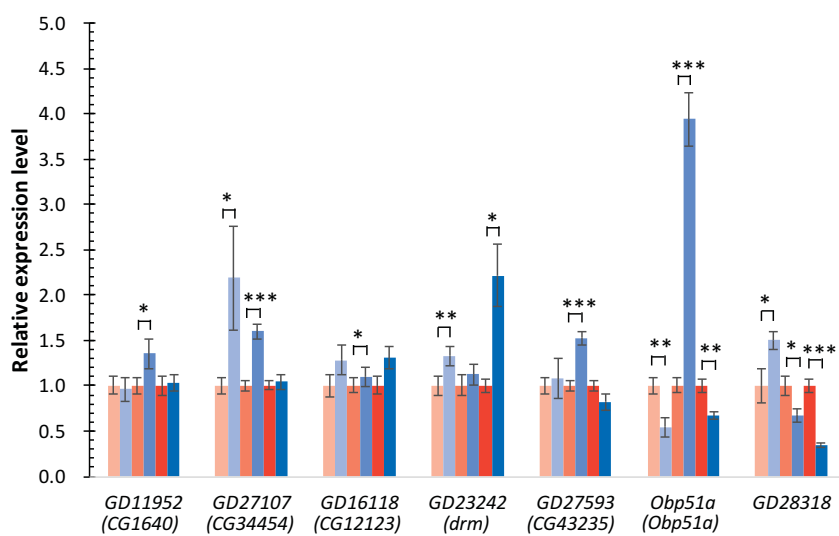
Gene	Chromosome	<i>Drosophila melanogaster</i> ortholog	Madagascar		Mayotte		Seychelles	
			RNA-Seq fold change	qPCR relative fold change	RNA-Seq fold change	qPCR relative fold change	RNA-Seq fold change	qPCR relative fold change
			SR/ST	(SR/ST) ± SEM	SR/ST	(SR/ST) ± SEM	SR/ST	(SR/ST) ± SEM
<i>GD27107</i>	3L	<i>CG34454</i>	3.61	2.19 ± 0.57*	2.23	1.60 ± 0.08***	2.3	1.04 ± 0.08
<i>GD27281</i>	2R	-	2.22	4.29 ± 1.79	1.83	1.14 ± 0.14	3.51	0.13 ± 0.01***
<i>GD27485</i>	X	-	45.93/0	1.36 ± 0.28*	53.03/0	0.94 ± 0.07	95.44/0	0.89 ± 0.09
<i>GD27593</i>	2L	<i>CG43235</i>	3.17	1.08 ± 0.22	3.77	1.53 ± 0.07***	2.74	0.82 ± 0.09
<i>GD27972</i>	X	-	1.89/0	1.29 ± 0.29	6.96/0	1.05 ± 0.09	1.72/0	1.18 ± 0.17
<i>GD28318</i>	2R	-	2.98	1.50 ± 0.10*	1.77	0.67 ± 0.07*	2.91	0.35 ± 0.02***
<i>GD28356</i>	2R	-	2.2	2.51 ± 0.19**	1.78	0.91 ± 0.10	3.6	0.12 ± 0.04***
<i>GD28414</i>	X	-	111230.51	1451.00 ± 165.21***	9946.21	0.68 ± 0.06**	2.96	434.52 ± 28.99
<i>GD28725</i>	X	-	5.79/0	14.63 ± 1.75**	8,75/0	24.76 ± 2.51***	1.46	0.92 ± 0.17
<i>GD28741</i>	X	<i>Ser7</i>	13.28	10.83 ± 1.46**	2.7	62.91 ± 9.20***	11.68	2.77 ± 0.07***
<i>GD29322</i>	3R	-	0.22	0.90 ± 0.16	0/6.87	0.82 ± 0.13	0.25	1.44 ± 0.24
<i>Obp51a</i>	2R	<i>Obp51a</i>	2.48	0.54 ± 0.11**	3.29	3.94 ± 0.29***	6.1	0.66 ± 0.04**



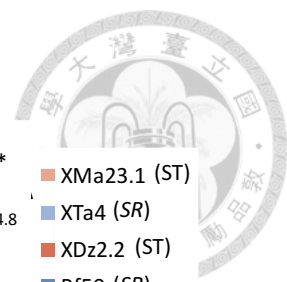
**A**



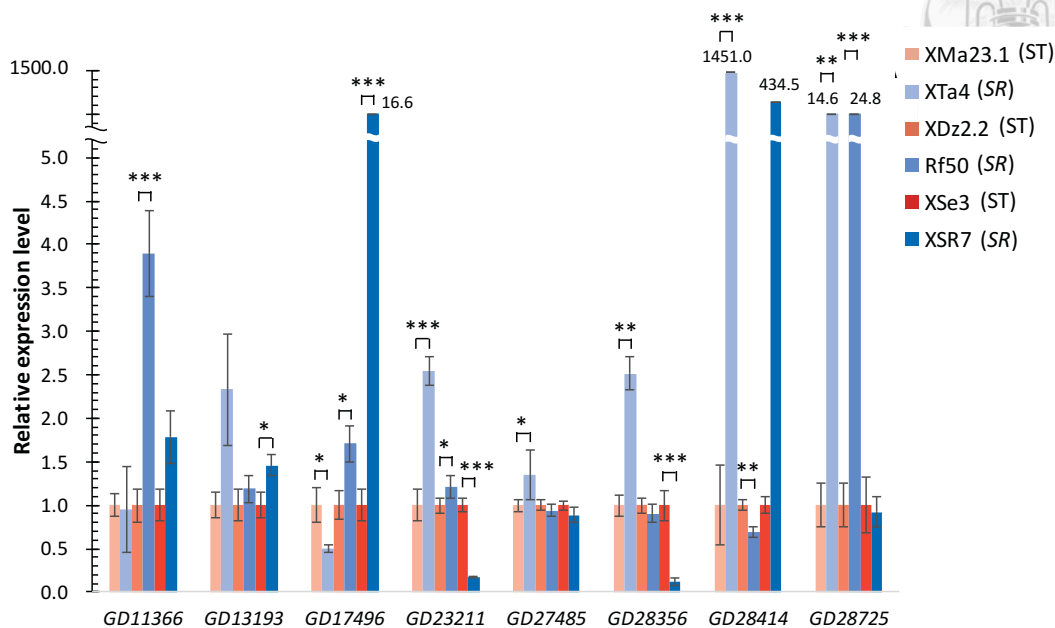
**B**



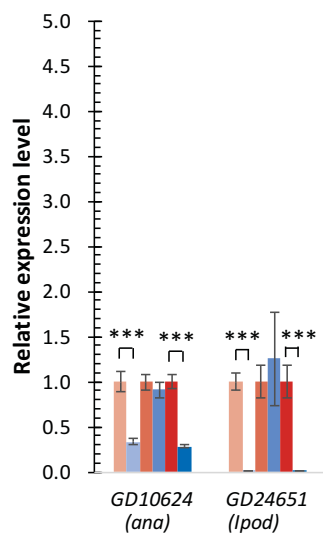
(Fig. 13A and 13B, figure legends are on p.44)



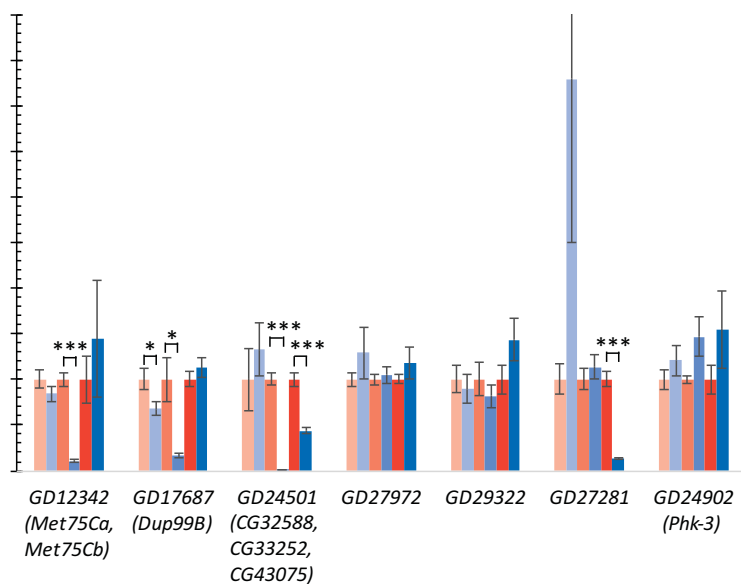
C



D



E



(Fig. 13C-13E, figure legends are on p.44)



**Figure 13. Testicular expression of *SR* candidate genes.** (A)-(C) Genes were validated to be up-regulated in the *SR* strain relative to the *ST* strain. (D) Genes were validated to be down-regulated in the *SR* strain relative to the *ST* strain. (E) Genes were not validated. (A)-(E) RT-qPCR analysis of testicular expression of genes in three *ST* and *SR* strains collected from three locations. XMa23.1 and XTa4 were collected from Madagascar; XDz2.2 and Rf50 were collected from Mayotte; XSe3 and XSR7 were collected from Seychelles. Gene names in the parentheses are *Drosophila melanogaster* homologs. Error bars represent SEMs. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (unpaired Student's t-test). Note that in Fig. 9C, because there was only one  $C_q$  value for GD28414 in XSe3, the unpaired Student's t-test was not performed.



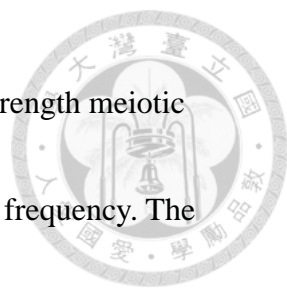
## Discussion



In this study, I performed *HP1D2* genotyping and found that there existed variation in both *ST* and *SR* strains, indicating that the *HP1D2* genotypes alone cannot predict *SR* phenotypes. Then, I compared the testicular expression of *ST* and *SR* strains and identified 221 differentially expressed genes. Among these genes, they were highly enriched in genes associated with multicellular organism reproduction, immune response, and genes localized to the extracellular region. The following qPCR validation of the *SR* candidate genes revealed that the *SR*-related genes in three locations were not identical and should be discussed separately. However, *CG16772*, *CG15209*, *Ser7*, *CG34265*, and *CG43348* were up-regulated in *SR* strains in all three locations, indicating that these genes played a similar role in *SR* in different *SR* strains.

### *Differences in the strength of meiotic drive*

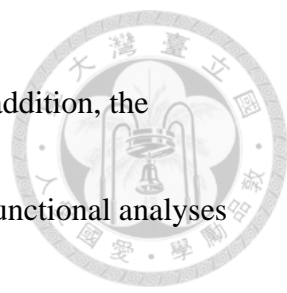
In the Paris *SR* system of *Drosophila simulans*, the meiotic drive on the sex chromosome is strong, leading to a strong female-biased progeny of drive-carrying males. The *SD* in *D. melanogaster* is also an example of a strong meiotic drive (Larracuente and Presgraves 2012). However, not all meiotic drives act with such strong



intensity. A previous study (Wei et al. 2017) has shown a moderate strength meiotic drive in *D. melanogaster*, with only ~8% increase in its transmission frequency. The variation in the strength of meiotic drives is an interesting issue in which few people have mentioned about it because of the difficulty to detect moderate ones. There are some possible explanations for the variation in the strength of meiotic drives: (1) Strong meiotic drive and moderate meiotic drive possess different mechanisms; (2) Both strong meiotic drive and moderate meiotic drive are in different stages of evolution; (3) There are more than one component in strong meiotic drives, and if only parts of the components exist, the strength of the meiotic drive will be weaker.

### ***Limitations of the transcriptomic analyses***

In the transcriptomic analyses of ST and SR samples, the SR strains collected from different geographical locations may be influenced by different mechanisms or different genes. If pooling all ST samples together and all SR samples together as biological replicates, it is possible to lose some genes specific in one or two geographical location(s). However, the common ones which play a role in all SR strains from different geographical locations should not be lost in my analyses. In this way, I can still

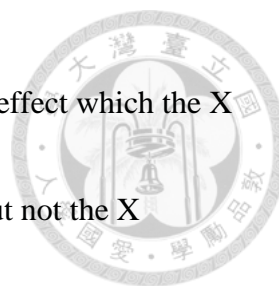


discover common, or probably more important *SR*-related genes. In addition, the differentially expressed genes can either be the cause or the effect. Functional analyses are necessary to confirm the causative relation between those genes and the *SR* phenotype.

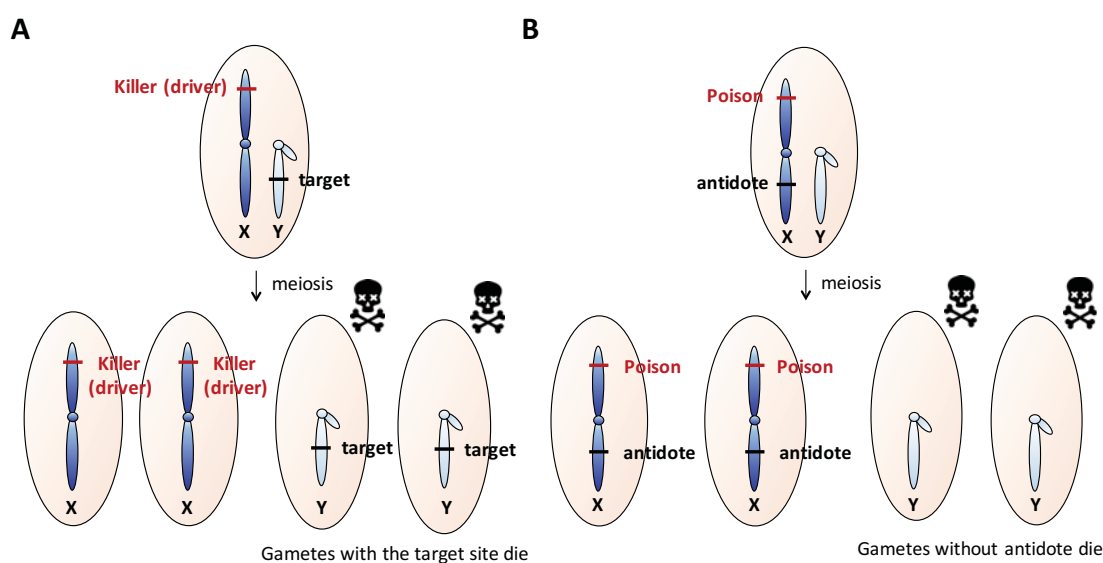
### ***Possible explanations of the underlying model of the sex-ratio phenotype***

From the GO analysis, there was an enrichment of genes related to multicellular organism reproduction, immune response, and synapse assembly. For the enrichment of immune response genes, it may be the effect of *SR*, not the cause. During spermatogenesis, the *SR* meiotic drive caused dying of non-driver-carrying spermatids. The death of spermatids may do harm to the fly. Therefore, the immune response genes were regulated to adapt to the defect which the *SR* driver caused.

Two possible mechanisms, killer-target meiotic drives and poison-antidote meiotic drives, have been reported in meiotic drives (Bravo Nunez et al. 2018). In the case of the Paris *SR* system in my study, it is possible that the drive acts specifically on the Y chromosome, but not on the X chromosome, that is, the killer-target drive (Fig. 14A). The second possibility is that the drive (poison) actually acts generally on both X and Y




chromosomes, but only X chromosome-carrying gametes rescue the effect which the X meiotic drive causes, resulting in the defect of the Y chromosome, but not the X chromosome, that is, the poison-antidote drive (Fig. 14B).



**Figure 14. Possible mechanisms of the X chromosome drive.** (A) Killer-target drivers kill only the gametes inherited with the target sites, the Y chromosome. (B) Poison-antidote drivers encode a *trans*-acting poison that acts generally on all gametes. Only the gametes which also inherit the antidote, the X chromosome-carrying gametes, can survive. (modified from Bravo Nunez et al. 2018)

Another question is that at which stage the drive acts to promote the preferential transmission of the X chromosome. Spermatogenesis in *Drosophila* can be divided into five stages: stem cell divisions, spermatogonial mitotic divisions, spermatocyte growth, meiotic divisions, and spermiogenesis. Germline stem cells (GSCs) produce



spermatogonial cells, which then divide four times to produce 16 primary spermatocytes from each spermatogonial cells. Primary spermatocytes mature and undergo meiotic divisions to produce 64 spermatids. These spermatids then remodel from round spermatids into needle-shaped mature sperms (Fuller 1993). The X meiotic drive may act during one of the stages.

If the Paris *SR* system is the case of the killer-target drive, in which the drive acts specifically on the Y chromosome, then it is more possible that the drive acts during meiosis II or spermiogenesis because there are nondisjunction of Y chromatids in spermatogenesis meiosis II and deficiencies of Y-carrying sperms (Cazemajor et al. 2000). However, if the X meiotic drive is the case of the poison-antidote drive, the drive may act before, during, or after meiosis II, but the Y chromosome defect shows during or after meiosis II. Also, the antidote must be on the X chromosome.

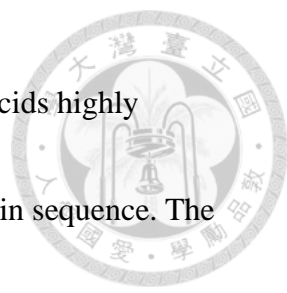
From the results of qPCR validation of *SR* candidate genes, there are five genes showing consistent up-regulation in all three *SR* strains relative to ST strains, including *CG16772*, *CG15209*, *Ser7*, *CG34265*, and *CG43348*. Among these genes, the immune response gene *CG16772*, uncharacterized gene *CG15209*, and serine protease *Ser7* show a higher testicular expression level in post-meiosis compared to that in meiosis



(Vibrantovski et al. 2009), indicating that these genes may play a role during sperm maturation. Since *CG15209* and *Ser7* are located on the X chromosome, it is possible that they are involved either in the killer-target X drive system by causing defect of Y-carrying sperms after meiosis or in the poison-antidote X drive system by rescuing X-carrying sperms from toxicity after X and Y separation (Table 8). *CG34265* and *CG43348* are uncharacterized genes, but there are interesting characteristics of their protein sequences. *CG34265* is predicted to encode a protein possessing a transmembrane helix. Its protein sequence is highly enriched in tyrosine, with almost a

**Table 8. The spermatogenesis stage at which the SR candidate genes are mainly expressed in the testes**

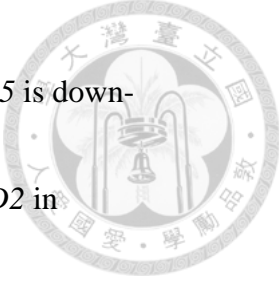
	Mitosis	Meiosis	Post-meiosis
			<i>CG16772</i> ,
Madagascar	-	<i>Tsp42Er, drm</i>	<u><i>CG15209</i></u> , <u><i>Ser7</i></u> , <i>Tsp42Er, ana</i>
			<i>CG16772</i> ,
Mayotte	<i>IM2, Obp51a</i>	<i>CG5402</i>	<u><i>CG15209</i></u> , <u><i>Ser7</i></u> , <u><i>CG1640</i></u> , <u><i>CG12123</i></u>
			<i>CG16772</i> ,
Seychelles	<i>Anp, E(Pc)</i>	<i>Tsp42Er, drm, E(Pc)</i>	<u><i>CG15209</i></u> , <u><i>Ser7</i></u> , <i>Tsp42Er, ana</i>



quarter of it. *CG43348* encodes a small protein with only 42 amino acids highly enriched in proline and histidine, constituting about 70% of the protein sequence. The characteristics of the protein sequences of both genes indicate that they may both serve in a signaling pathway which is related to the *SR* phenotype.

However, there are still discrepancies of the validated genes among the three *SR* strains. Moreover, *HP1D2*, the gene on the drive locus of Paris *SR* system, also shows variation in both genotypes and testicular gene expression level within ST and within *SR* strains. The variation indicates that there may be different mechanisms underlying the *SR* strains collected from different locations. Therefore, three ST and *SR* pairs are discussed separately.

In strains collected from Madagascar, *HP1D2*, a characterized *SR* drive genetic locus which is expressed in spermatogonia and specifically binds the Y chromosome (Helleu et al. 2016), is confirmed to be down-regulated in the *SR* strain as expected. *Caf1-105* is a subunit of chromatin assembly factor 1 (CAF1) complex, which functions in heterochromatin formation, chromatin assembly, and histone binding. CAF1 complex has also been observed to bind to HP1 protein and to be localized to heterochromatin (Murzina et al. 1999). In addition, CAF1-p75 encoded by processed *Caf1-105* mediates



assembly of protamine-based chromatin (Doyen et al. 2013). *Caf1-105* is down-regulated in the *SR* strain, which is likely to work together with *HP1D2* in heterochromatin formation. It may also cause the defect in histone-to-protamine transition during spermiogenesis, which may cause failure of sperm maturation. In addition, the transcription factor *drm*, the tetraspanin *Tsp42Er*, and the glycoprotein *ana* show a higher testicular expression level in meiosis compared to that in the mitosis stage (Vibrantovski et al. 2009). The immune response gene *CG16772*, uncharacterized gene *CG15209*, serine protease *Ser7*, and *ana* show a higher testicular expression level in post-meiosis compared to that in meiosis (Vibrantovski et al. 2009). Moreover, *Ipod*, an interaction partner of *Mt2* (*CG10692*) (Kunert et al. 2005) is down-regulated in the *SR* strain. Taken together, the mechanism of the X chromosome drive is more likely to be the killer-target drive, in which the drive acts specifically on the Y chromosome during meiosis II or spermiogenesis.

In strains collected from Mayotte, there are some antibacterial peptide genes and immune response genes which are up-regulated, such as *CG16772*, *IM1*, and *IM2*. *Caf1-105* is down-regulated in the *SR* strain, which may cause abnormality in heterochromatin formation. It may also cause the defect in histone-to-protamine



transition during spermiogenesis, which may cause failure of sperm maturation.

*CG1640* and *CG34454* are experimentally proved and predicted, respectively, to

localize to mitochondrion and are up-regulated in the *SR* strain. In addition, *CG1640*

shows a higher testicular expression level in post-meiosis compared to that in meiosis

(Vibrantovski et al. 2009). The giant mitochondria local remodeling is essential for

sperm elongation and maturation (Noguchi et al. 2011). Thus, it raises a possibility that

*CG1640* and *CG34454* localize to mitochondrion to promote local remodeling, in which

the abnormal overexpression of these genes may lead to defect in sperm maturation.

*CG12123* is one of the genes located in one of the *SR* meiotic drive loci, tandem

duplication of six genes on the X chromosome (Montchamp-Moreau et al. 2006), and is

up-regulated in the *SR* strain. Moreover, its testicular expression increases from mitotic

stage to meiotic stage, and further increases during post-meiotic stage (Vibrantovski et

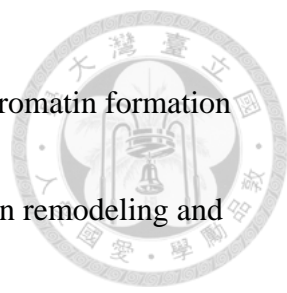
al. 2009). This raises a possibility that *CG12123* serves as a critical factor of the meiotic

drive and acts during or after meiosis. Taken together, the mechanism of the X

chromosome drive is more likely to be killer-target drive, in which the drive acts

specifically on the Y chromosome during meiosis II or spermiogenesis.

In strains collected from Seychelles, immune response and antibacterial genes



including *CG16772* and *Anp* are up-regulated. *vig2* impacts heterochromatin formation and H3K9me2 (Gracheva et al. 2009). *E(Pc)* plays a role in chromatin remodeling and histone exchange as a subunit of the Tip60 complex (Kusch et al. 2004). The up-regulation of *vig2* and down-regulation of *E(Pc)* may cause a defect in chromatin assembly before or during meiosis. The down-regulation of *Ipod* and up-regulation of *drm* may act in transcriptional regulation. *Ser7* may act in proteolysis. In addition, the transcription factor *drm*, the tetraspanin *Tsp42Er*, and the glycoprotein *ana* show a higher testicular expression level in meiosis compared to that in mitosis stage (Vibrantovski et al. 2009). Taken together, it is unlikely to deduce the possible mechanism of the X chromosome drive from this group of ST and SR strain.

According to the current results, it is still difficult to distinguish major mechanisms, the killer-target drive or the poison-antidote drive, but the killer-target drive is more consistent with the current results. In the future, if I could discover the target of the driver on the Y chromosome, the killer-target drive will be the most likely mechanism. In contrast, the poison-antidote drive or other unidentified drives would be the underlying mechanism.

## References



Akera T, Chmatal L, Trimm E, Yang K, Aonbangkhen C, Chenoweth DM, Janke C,

Schultz RM, Lampson MA. 2017. Spindle asymmetry drives non-Mendelian chromosome segregation. *Science* 358:668-672.

Andersen PR, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* 549:54-59.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data.

Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

Atlan A, Mercot H, Landre C, Montchamp-Moreau C. 1997. The sex-ratio trait in *Drosophila simulans*: geographical distribution of distortion and resistance. *Evolution* 51:1886-1895.

Bauer H, Schindler S, Charron Y, Willert J, Kusecek B, Herrmann BG. 2012. The nucleoside diphosphate kinase gene *Nme3* acts as quantitative trait locus promoting non-Mendelian inheritance. *PLoS Genet.* 8:e1002567.

Bauer H, Veron N, Willert J, Herrmann BG. 2007. The *t*-complex-encoded guanine nucleotide exchange factor Fgd2 reveals that two opposing signaling pathways promote transmission ratio distortion in the mouse. *Genes Dev.* 21:143-147.



Bauer H, Willert JR, Koschorz B, Herrmann BG. 2005. The *t* complex-encoded  
GTPase-activating protein Tagap1 acts as a transmission ratio distorter in mice.

Nat. Genet. 37:969-973.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina  
sequence data. Bioinformatics 30:2114-2120.

Bravo Nunez MA, Nuckolls NL, Zanders SE. 2018. Genetic villains: killer meiotic  
drivers. Trends Genet. 34:424-433.

Camacho JP, Schmid M, Cabrero J. 2011. B chromosomes and sex in animals. Sex Dev.  
5:155-166.

Cazemajor M, Joly D, Montchamp-Moreau C. 2000. Sex-ratio meiotic drive in  
*Drosophila simulans* is related to equational nondisjunction of the Y  
chromosome. Genetics 154:229-236.

Cazemajor M, Landre C, Montchamp-Moreau C. 1997. The sex-ratio trait in *Drosophila*  
*simulans*: genetic analysis of distortion and suppression. Genetics 147:635-642.

Chmatal L, Gabriel SI, Mitsainas GP, Martinez-Vargas J, Ventura J, Searle JB, Schultz  
RM, Lampson MA. 2014. Centromere strength provides the cell biological basis  
for meiotic drive and karyotype evolution in mice. Curr. Biol. 24:2295-2300.



- Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, et al. 2016. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17:13.
- Dawe RK, Lowry EG, Gent JI, Stitzer MC, Swentowsky KW, Higgins DM, Ross-Ibarra J, Wallace JG, Kanizay LB, Alabady M, et al. 2018. A kinesin-14 motor activates neocentromeres to promote meiotic drive in maize. *Cell* 173:1-12.
- Dermitzakis ET, Masly JP, Waldrip HM, Clark AG. 2000. Non-Mendelian segregation of sex chromosomes in heterospecific *Drosophila* males. *Genetics* 154:687-694.
- Doyen CM, Moshkin YM, Chalkley GE, Bezstarosti K, Demmers JA, Rathke C, Renkawitz-Pohl R, Verrijzer CP. 2013. Subunits of the histone chaperone CAF1 also mediate assembly of protamine-based chromatin. *Cell Rep* 4:59-65.
- Ebens AJ, Garren H, Cheyette BN, Zipursky SL. 1993. The *Drosophila* anachronism locus: a glycoprotein secreted by glia inhibits neuroblast proliferation. *Cell* 74:15-27.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-1797.
- Ellis LL, Carney GE. 2010. Mating alters gene expression patterns in *Drosophila*



*melanogaster* male heads. BMC Genomics 11:558.

Findlay GD, Yi XH, MacCoss MJ, Swanson WJ. 2008. Proteomics reveals novel

*Drosophila* seminal fluid proteins transferred at mating. PLoS Biol. 6:1417-

1426.

Fisher RA. 1930. The genetical theory of natural selection. Clarendon Press, Oxford

Fradkin LG, Kamphorst JT, DiAntonio A, Goodman CS, Noordermeer JN. 2002.

Genomewide analysis of the *Drosophila* tetraspanins reveals a subset with

similar function in the formation of the embryonic synapse. Proc. Natl. Acad.

Sci. USA 99:13663-13668.

Fuller, MT. 1993. Spermatogenesis. Cold Spring Harbor Press, Cold Spring Harbor,

NY.

Ganetzky B. 1977. On the components of segregation distortion in *Drosophila*

*melanogaster*. Genetics 86:321-355.

Gershenson S. 1928. A new sex-ratio abnormality in *Drosophila obscura*. Genetics

13:488-507.

Gracheva E, Dus M, Elgin SC. 2009. *Drosophila* RISC component VIG and its

homolog Vig2 impact heterochromatin formation. PLoS One 4:e6182.



Hartl DL. 1974. Genetic dissection of segregation distortion. I. Suicide combinations of

*SD* genes. *Genetics* 76:477-486.

Hatini V, Green RB, Lengyel JA, Bray SJ, DiNardo S. 2005. The drumstick/lines/bowl

regulatory pathway links antagonistic Hedgehog and Wingless signaling inputs

to epidermal cell differentiation. *Genes Dev.* 19:709-718.

Hauschteckjungen E, Maurer B. 1976. Sperm dysfunction in sex-ratio males of

*Drosophila subobscura*. *Genetica* 46:459-477.

Helleu Q, Gerard PR, Dubruille R, Ogereau D, Prud'homme B, Loppin B, Montchamp-

Moreau C. 2016. Rapid evolution of a Y-chromosome heterochromatin protein

underlies sex chromosome meiotic drive. *Proc. Natl. Acad. Sci. USA* 113:4110-

4115.

Helleu Q, Gerard PR, and Montchamp-Moreau C. 2015. Sex chromosome drive. *Cold*

*Spring Harb. Perspect. Biol.* 7:a017616.

Herrmann BG, Koschorz B, Wertz K, McLaughlin KJ, Kispert A. 1999. A protein

kinase encoded by the *t complex responder* gene causes non-mendelian

inheritance. *Nature* 402:141-146.

Hochheimer A, Zhou S, Zheng S, Holmes MC, Tjian R. 2002. TRF2 associates with



DREF and directs promoter-selective gene expression in *Drosophila*. *Nature*

420:439-445.

Hu W, Jiang ZD, Suo F, Zheng JX, He WZ, Du LL. 2017. A large gene family in fission yeast encodes spore killers that subvert Mendel's law. *Elife* 6:e26057.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4:44-57.

Iwaki DD, Johansen KA, Singer JB, Lengyel JA. 2001. drumstick, bowl, and lines are required for patterning and cell rearrangement in the *Drosophila* embryonic hindgut. *Dev. Biol.* 240:611-626.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36.

Kunert N, Marhold J, Kramer K, Lyko F. 2005. Identification and characterization of IPOD, a novel interaction partner of the Dnmt2 DNA methyltransferase in *Drosophila*. *A. Dros. Res. Conf.* 46:298A.

Kusano A, Staber C, Ganetzky B. 2001. Nuclear mislocalization of enzymatically active RanGAP causes segregation distortion in *Drosophila*. *Dev. Cell* 1:351-361.





Kusano A, Staber C, Ganetzky B. 2002. Segregation distortion induced by wild-type

RanGAP in *Drosophila*. Proc. Natl. Acad. Sci. USA 99:6866-6870.

Kusch T, Florens L, MacDonald WH, Swanson SK, Glaser RL, Yates JR, Abmayr SM,

Washburn MP, Workman JL. 2004. Acetylation by Tip60 is required for selective

histone variant exchange at DNA lesions. Science 306:2084-2087.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient

alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.

Larracunte AM, Presgraves DC. 2012. The selfish segregation distorter gene complex

of *Drosophila melanogaster*. Genetics 192:33-53.

Levine MT, McCoy C, Vermaak D, Lee YC, Hiatt MA, Matsen FA, Malik HS. 2012.

Phylogenomic analysis reveals dynamic evolutionary history of the *Drosophila*

heterochromatin protein 1 (HP1) gene family. PLoS Genet. 8:e1002729.

Levy F, Rabel D, Charlet M, Bulet P, Hoffmann JA, Ehret-Sabatier L. 2004. Peptidomic

and proteomic analyses of the systemic immune response of *Drosophila*.

Biochimie 86:607-616.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,

Durbin R, Genome Project Data Processing Subgroup. 2009. The sequence



alignment/map format and SAMtools. *Bioinformatics* 25:2078-2079.

Lindholm AK, Dyer KA, Firman RC, Fishman L, Forstmeier W, Holman L,

Johannesson H, Knief U, Kokko H, Larracuente AM, et al. 2016. The ecology and evolutionary dynamics of meiotic drive. *Trends Ecol. Evol.* 31:315-326.

Lye CM, Naylor HW, Sanson B. 2014. Subcellular localisations of the CPTI collection of YFP-tagged proteins in *Drosophila* embryos. *Development* 141:4006-4017.

Lyttle TW. 1989. The effect of novel chromosome position and variable dose on the genetic behavior of the *Responder* (*Rsp*) element of the *Segregation distorter* (*SD*) system of *Drosophila melanogaster*. *Genetics* 121:751-763.

Lyttle TW. 1991. Segregation distorters. *Annu. Rev. Genet.* 25:511-557.

Mercot H, Atlan A, Jacques M, and Montchamp-Moreau C. 1995a. Sex-ratio distortion in *Drosophila simulans*: co-occurrence of a meiotic drive and a suppressor of drive. *J. Evol. Biol.* 8:283-300.

Mercot H, Llorente B, Jacques M, Atlan A, Montchampmoreau C. 1995b. Variability within the Seychelles cytoplasmic incompatibility system in *Drosophila simulans*. *Genetics* 141:1015-1023.

Merrill C, Bayraktaroglu L, Kusano A, Ganetzky B. 1999. Truncated RanGAP encoded



by the *Segregation Distorter* locus of *Drosophila*. *Science* 283:1742-1745.

Montchamp-Moreau C, Ginhoux V, Atlan A. 2001. The Y chromosomes of *Drosophila*

*simulans* are highly polymorphic for their ability to suppress sex-ratio drive.

*Evolution* 55:728-737.

Montchamp-Moreau C, Ogereau D, Chaminade N, Colard A, Aulard S. 2006.

Organization of the sex-ratio meiotic drive region in *Drosophila simulans*.

*Genetics* 174:1365-1371.

Morgan TH, Bridges CB, Sturtevant AH. 1925. The genetics of *Drosophila*. *Bibliogr.*

*Genet.* 2:1-262.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and

quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5:621-628.

Murzina N, Verreault A, Laue E, Stillman B. 1999. Heterochromatin dynamics in mouse

cells: Interaction between chromatin assembly factor 1 and HP1 proteins. *Mol.*

*Cell* 4:529-540.

Noguchi T, Koizumi M, Hayashi S. 2011. Sustained elongation of sperm tail promoted

by local remodeling of giant mitochondria in *Drosophila*. *Curr. Biol.* 21:805-

814.



Nuckolls NL, Bravo Nunez MA, Eickbush MT, Young JM, Lange JJ, Yu JS, Smith GR,

Jaspersen SL, Malik HS, Zanders SE. 2017. *wtf* genes are prolific dual poison-antidote meiotic drivers. *Elife* 6:e26033.

Pimpinelli S, Dimitri P. 1989. Cytogenetic analysis of segregation distortion in

*Drosophila melanogaster*: the cytological organization of the *Responder (Rsp)* locus. *Genetics* 121:765-772.

Policansky D, Ellison J. 1970. "Sex-Ratio" in *Drosophila Pseudoobscura*: spermiogenic

failure. *Science* 169:888-889.

R Core Team. 2017. R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Roberts A, Pimentel H, Trapnell C, Pachter L. 2011. Identification of novel transcripts

in annotated genomes using RNA-Seq. *Bioinformatics* 27:2325-2329.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov

JP. 2011. Integrative genomics viewer. *Nat. Biotechnol.* 29:24-26.

Ross J, Jiang H, Kanost MR, Wang Y. 2003. Serine proteases and their homologs in the

*Drosophila melanogaster* genome: an initial analysis of sequence conservation



and phylogenetic relationships. *Gene* 304:117-131.

Samakovlis C, Kylsten P, Kimbrell DA, Engstrom A, Hultmark D. 1991. The andropin

gene and its product, a male-specific antibacterial peptide in *Drosophila*

*melanogaster*. *EMBO J.* 10:163-169.

Sandler L, Carpenter AT. 1972. A note on the chromosomal site of action of *SD* in

*Drosophila melanogaster*. *Edinburgh Symposium on the Genetics of the*

*Spermatozoon*, edited by R. A. Beatty and S. Gluecksohn-Waelsch. University

of Edinburgh, Edinburgh, Scotland. pp. 247-268.

Schimenti J. 2000. Segregation distortion of mouse *t* haplotypes - the molecular basis

emerges. *Trends Genet.* 16:240-243.

Silver LM. 1993. The peculiar journey of a selfish chromosome: mouse *t* haplotypes

and meiotic drive. *Trends Genet.* 9:250-254.

Tao Y, Araripe L, Kingan SB, Ke Y, Xiao H, Hartl DL. 2007a. A sex-ratio meiotic drive

system in *Drosophila simulans*. II: an X-linked distorter. *PLoS Biol.* 5:e293.

Tao Y, Hartl DL, Laurie CC. 2001. Sex-ratio segregation distortion associated with

reproductive isolation in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 98:13183-

13188.



Tao Y, Masly JP, Araripe L, Ke Y, Hartl DL. 2007b. A sex-ratio meiotic drive system in

*Drosophila simulans*. I: an autosomal suppressor. PLoS Biol. 5:e292.

Tarazona S, Furio-Tari P, Turra D, Pietro AD, Nueda MJ, Ferrer A, Conesa A. 2015.

Data quality aware analysis of differential expression in RNA-seq with NOISeq

R/Bioc package. Nucleic Acids Res. 43:e140.

Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, Conesa A. 2011. Differential

expression in RNA-seq: a matter of depth. Genome Res. 21:2213-2223.

Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV):

high-performance genomics data visualization and exploration. Brief.

Bioinform. 14:178-192.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL,

Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq

reveals unannotated transcripts and isoform switching during cell differentiation.

Nat. Biotechnol. 28:511-515.

Vibrantovski MD, Lopes HF, Karr TL, Long MY. 2009. Stage-specific expression

profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome

inactivation drives genomic relocation of testis-expressed genes. PLoS Genet.

5:e1000731.



Wei KH, Reddy HM, Rathnam C, Lee J, Lin D, Ji S, Mason JM, Clark AG, Barbash

DA. 2017. A pooled sequencing approach identifies a candidate meiotic driver in *Drosophila*. *Genetics* 206:451-465.

Werren JH. 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. USA* 108(Suppl 2):10863-10870.

Werren JH, Nur U, Wu CI. 1988. Selfish genetic elements. *Trends Ecol. Evol.* 3:297-302.

Werren JH, Stouthamer R. 2003. PSR (paternal sex ratio) chromosomes: the ultimate selfish genetic elements. *Genetica* 117:85-101.

Wu CI, Lyttle TW, Wu ML, Lin GF. 1988. Association between a satellite DNA sequence and the *Responder of Segregation Distorter* in *D. melanogaster*. *Cell* 54:179-189.

## Appendix A



### RNA sequencing sample preparation

For ST strains, XMa23.1, XDz2.2, and XSe3 and SR strains, XTa4, Rf50, and XSR7, at least 200 pairs of testes were dissected in PBS from males less than two-day old from each strain. RNA extraction was performed using the TRIzol™ reagent (Thermo Fisher Scientific, Waltham, MA, USA) according to their protocol. The samples were then sent for single-end RNA sequencing on Illumina platform with 101 bp read length. The RNA-Seq run are in two batches, with XMa23.1 and XTa4 in one batch, while XDz2.2, XSe3, Rf50, and XSR7 in the other batch.

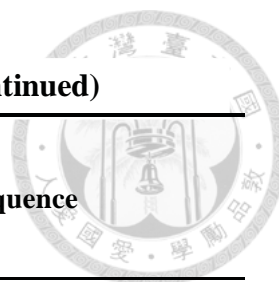


## Appendix B



**Table B1. List of primers used in the RT-qCPR experiments (to be continued)**

<i>D. melanogaster</i>		Primer name	Primer sequence
Gene	ortholog		
<i>Anp</i>	<i>Anp</i>	Anpq-F	TACTTTGTGtTCCTTGTCGTCCT
		Anpq-R	TAGCTgTGCCtATTCCCGCTTG
<i>GD10455</i>	<i>Tsp42Er</i>	GD10455q-F	CAAAATGATACGCaCTACAAGGAC
		GD10455q-R	TGAAGGCACCAATCgCAATGAC
<i>GD10624</i>	<i>ana</i>	GD10624q-F	GTCCTCAAATCCcCGTCACCGTCT
		GD10624q-R	CGCACTTCAtGTCGATTTCCACC
<i>GD10801</i>	<i>E(Pc)</i>	E(Pc)q-F	CTaAACCAGGACGACGAGACCA
		E(Pc)q-R	ACGAAAAGCCAAGgACGGA
<i>GD11363</i>	<i>IM1</i>	GD11363q-F	GCACTCAGTATCCAAAACCcGAGAA
		GD11363q-R	GCCGTTGATGAcCACATTGCC
<i>GD11365</i>	<i>IM2</i>	GD11365q-F	CTTCTCAGTCcTTACCGTCTTCGT
		GD11365q-R	TTTGCAGTCGCCGTTcATCACC
<i>GD11366</i>	-	GD11366q-F	ATTgCTATCAgTCGCCTTCGTT
		GD11366q-R	ATGAaATTGCCAGGAGTCAGT
<i>GD11952</i>	<i>CG1640</i>	GD11952q-F	ACACGaCTGCTGGATTCACCCGA
		GD11952q-R	ATGCCGCCATCCCTTTTCaCGATA

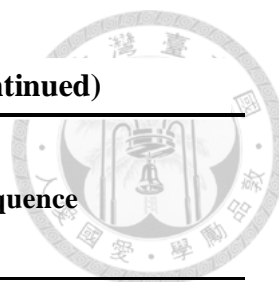


**Table B1. List of primers used in the RT-qCPR experiments (continued)**

Gene	<i>D. melanogaster</i>		Primer name	Primer sequence
		ortholog		
<i>GD12342</i>	<i>Met75Ca</i> ,		GD12342q-F	CCAGGAAACCACGGAGGACAA
	<i>Met75Cb</i>		GD12342q-R	CTAGTACGGCAGGCAAGTGTG
<i>GD13193</i>	-		GD13193q-F	CGCTCCCAGAAAATCAGGT
			GD13193q-R	AGCCAGTTCCAAAGAGATgTAGCAA
<i>GD16118</i>	<i>CG12123</i>		GD16118q-F	GCCTGCGACTCcGTGATCCCTT
			GD16118q-R	ATGCTTCcCCATCCGTAGTTGACC
<i>GD16989</i>	<i>CG15209</i>		GD16989q-F	CCGcTCATCTTCCTGGTGATCCT
			GD16989q-R	CGCAATTCTACCGACcTACTCCGACT
<i>GD17496</i>	-		GD17496q-F	CTGGAAGCGACGAGtTCTTGAC
			GD17496q-R	CGAGGAATAAATaCATGAAGCGAATA
<i>GD17687</i>	<i>Dup99B</i>		GD17687q-F	TCCGCTGTTTCTCCTCTTGgTC
			GD17687q-R	CACCACTTCTCACGCTCCATC
<i>GD18077</i>	<i>CG5402</i>		GD18077q-F	CTCAAATACATTGCCAGCGTCT
			GD18077q-R	GCCCATGAaGCCCAAAACTCCC
<i>GD19402</i>	<i>CG34265</i>		GD19402q-F	AACaATTACTACCAGACGcCGCCAT
			GD19402q-R	GCCATAAACATTCGCATAGTcGCTT
<i>GD21176</i>	<i>vig2</i>		vig2q-F	TTGTTCtTGGACGACGATGACTCCT
			vig2q-R	CGGCTTGTTCTgCTTCTCGGACTT

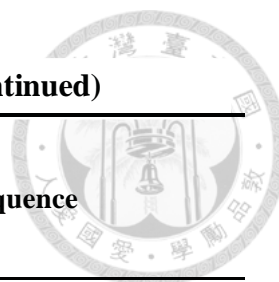
**Table B1. List of primers used in the RT-qCPR experiments (continued)**

Gene	<i>D. melanogaster</i> ortholog	Primer name	Primer sequence
GD21970	<i>vig</i>	vigq-F	ATGAATCTcGCACTGAACCACAAACC
		vigq-R	TTGCTGTcCCTTCCACTCGTCC
GD23211	-	GD23211q-F	TTCCGCAGAGAAAGAGaACCCA
		GD23211q-R	GCCGCCaTTGCCATAACCG
GD23242	<i>drm</i>	GD23242q-F	TCCGATgTCCGCCGCAAG
		GD23242q-R	CCGCACACCcCGCACGAA
GD24237	<i>CG16772</i>	CG16772q-F	CATCCCCATTCTCATTCGCATCCT
		CG16772q-R	CGGTGATTCCTTGTtCGCATCT
GD24501	<i>CG32588,</i> <i>CG33252,</i> <i>CG43075</i>	GD24501q-F	ATGAATgTCTTCGAGCAGATTAG
		GD24501q-R	CCCATGtAAAGGTTACGAAATC
GD24651	<i>Ipod</i>	GD24651q-F	CCAGtTGCTCCCATTGCCTATC
		GD24651q-R	AACCATcTCCCGAAGCATAACGAC
GD24902	<i>Phk-3</i>	GD24902q-F	TCTTCTtCCCGACGCCCTG
		GD24902q-R	TGCTTTGgTCTGTAGATGCCTT
GD25675	<i>LamC</i>	LamCq-F	ATCTCGCCCAGcACACCGTCA
		LamCq-R	GCCACcTTAGCCTCCTTCGTCTT
GD27028	<i>CG43348</i>	GD27028q-F	CACCTCgCCAGCtCCACCAAC
		GD27028q-R	GTGATGTCtATGATtGGGCTTCGGAT



**Table B1. List of primers used in the RT-qCPR experiments (continued)**

Gene	<i>D. melanogaster</i>		Primer name	Primer sequence
	ortholog			
<i>GD27107</i>	<i>CG34454</i>		GD27107q-F	ACCACATCGCAATACAATCCC
			GD27107q-R	AATGTCAGCTCCACAGAAAC
<i>GD27281</i>	-		GD27281q-F	ATGCTGAaCAACCGACATTCC
			GD27281q-R	GCCAAGTcCATTTTCCACCCT
<i>GD27485</i>	-		GD27485q-F	ACCGATCtGACCGTGGCATT
			GD27485q-R	TGGTGCGGAGaCTGCTGGAA
<i>GD27593</i>	<i>CG43235</i>		GD27593q-F	AGCCTATCAGATCCTATACTCC
			GD27593q-R	GACTACCTTTAAGCCACGATG
<i>GD27972</i>	-		GD27972q-F	TTCtAATATGACTATCGgCACCT
			GD27972q-R	CTTGGCCTTTGtGCTCGAC
<i>GD27984</i>	<i>His3.3B</i>		His3.3Bq-F	CGTGAGAgCCGTCGTTACC
			His3.3Bq-R	ATCCTGAGCGATTTCAaGAACC
<i>GD28318</i>	-		GD28318q-F	ATTCAaGATTAGgTTCCCAGTCAGA
			GD28318q-R	GTCACTTAcAAAATGAGATATGCGT
<i>GD28356</i>	-		GD28356q-F	ATCATCGGTCCACTATtCGTTT
			GD28356q-R	CGTTCTGCCTTTCAAcTATACACCT
<i>GD28414</i>	-		GD28414q-F	ATGCCGTCAcCTTTGCCtGCTC
			GD28414q-R	ATAACCGATTCCcGCTCCATAACTCC



**Table B1. List of primers used in the RT-qCPR experiments (continued)**

Gene	<i>D. melanogaster</i>		Primer name	Primer sequence
	ortholog			
GD28722	<i>Caf1-105</i>		Caf1-105q-F	CCCACTCCCATCGCCATC
			Caf1-105q-R	CGTTATTCCCCTTATCCTCCtGCTT
GD28725	-		GD28725q-F	TAGTTCGCCtCCCAAAaTGTCACC
			GD28725q-R	CAAATTATCCTGACTCCTCtAGCAA
GD28741	-		GD28741q-F	CTGCGACTTTTATGCCGTT
			GD28741q-R	TgATTACTGATCCTGAGTGCCAA
GD29322	-		GD29322q-F	TCACA <sub>g</sub> CACAGCATTAGGGTT
			GD29322q-R	TTTTATGG <sub>g</sub> GGACATC <sub>g</sub> CTCGT
HP1D2	-		HP1D2q-F	gGCATCGTAAAAGGTCGTCT
			HP1D2q-R	GCTTCCACT <sub>c</sub> GCTCCCAT <sub>c</sub> TGCTC
<i>Obp51a</i>	<i>Obp51a</i>		Obp51aq-F	ATTTTCCGCACAGCAGTC
			Obp51aq-R	ACCATACTTATCATAGCTCTCC
GD18948	<i>RpIII40</i>		rpIII40q1-F	ATGGTGGCTTGCgTTTCGGTG
			rpIII40q1-R	ATTgTTGCGCAGATTGGCGATGG
GD17524	<i>eIF2Ba</i>		eIF2B-qF1	CCGCAAATTACGGAAAATGGCCAGr
			eIF2B-qR1	GCaGCCACCGCTTCCCTCAT
GD26811	<i>lt</i>		Light410F	CCGATTCCAAAGCTCACATT
			Light535R	TtGACAAAACACTGCCTTCG

**Table B1. List of primers used in the RT-qCPR experiments (continued)**

Gene	<i>D. melanogaster</i> ortholog	Primer name	Primer sequence
<i>GD16764</i>	<i>Act5C</i>	Act5C_c418+	GGCACCACACCTTCTACAAT
		Act5C_c504-	TGGGTCATCTTCTCACGGTT
		Act5C_c984+	CACGAGACCACCTACA ACTCC
		Act5C_c1232-	GATCCACATCTGCTGGAAGG

\*Note: modified bases are in lowercase.