Characterisation of the M-locus and functional analysis of the male-determining gene in the housefly

Wu, Yanli

English summary

The housefly, *Musca domestica* is particularly suited to investigate the evolution of sex determination and sex chromosomes because it has a polymorphic sex determination system. The male-determining *M*-locus, typically located on the Y-chromosome, can also be present on any of the five autosomes or even the X-chromosome. Recently, based upon differential expression analysis, a male-determining gene was identified and termed *Mdmd* for *Musca domestica male determiner*. *Mdmd* appears to have arisen as a duplication of the splicing regulatory gene *CWC22*, called *nucampholin* (*ncm*) in insects. To further characterise the *M*-loci in terms of genomic organisation and function, I addressed several questions about *Mdmd* structure and function: What is the genomic organisation of *M*-loci on different chromosomes? What is the coding sequence of *Mdmd*? To what extent are the different *M*-loci conserved? What is the evolutionary relationship between *Mdmd* and its paralog *CWC22/nucampholin*? When and where is *Mdmd* expressed? Is *Mdmd* sufficient for male function?

Although *Mdmd* was identified as the male-determining gene in the housefly, the complete sequence of *Mdmd* and its embedding in the *M*-locus remained unknown. In **Chapter 2**, I investigated the complex nature of *M*-loci in two autosomal *M* strains, M^{III} (*M*-locus on autosome III) and M^{V} (*M*-locus on autosome V). I found that the *M*-loci contain multiple copies of different sequences of *Mdmd* sequences, with various levels of homology to each other. Interestingly, the M^{III}-locus and the M^{V}-locus share common sequences. On the basis of these common sequences, I identified an open reading frame (ORF) that is part of the *Mdmd* gene (**Chapter 3**). Sequences with high similarity to the *Mdmd* ORF were also detected in M^{II} (*M*-locus on autosome II) and M^{Y} (*M*-locus on Y-chromosome) strains, but not in the M^{I} (*M*-locus on autosome I) strain, which probably has a different male-determining gene (s). This ORF is assumed to be the coding sequence of *Mdmd*, the functional male-determining gene.

The liability and turnover of sex chromosomes is a remarkable aspect of sex determination evolution. Sex chromosomes are believed to evolve from ordinary autosomes that lost recombination after having acquired a sex-determining role. What drives the evolution of new sex chromosomes is not yet well understood. My results in *M. domestica* provide support for the birth-decay-rebirth model of sex chromosome evolution. The high sequence similarity of $Mdmd^{II}$, $Mdmd^{III}$, $Mdmd^{V}$ and $Mdmd^{Y}$ suggests that all *Mdmd* genes originated from a common ancestral sequence. A comparison of Mdmd protein sequences and its paralog CWC22/NCM in **Chapter 3** suggests a scenario of *M*-locus evolution, whereby the male-determining gene *Mdmd* evolved after a single duplication event of *Md-ncm* generating a proto-Y chromosome. Whether this happened on the ancestral Y or on an autosomal pair that was not yet involved in sex determination cannot be

answered at this moment.

The next stage of Y-chromosome evolution would be the reduction of recombination in the surrounding *Mdmd* region, followed by accumulation of repetitive sequences and transposons due to the lack of recombination on the proto-sex chromosome. Consistent with this model, I found that *M*-loci in the M^{III} and M^{V} strain contain transposable elements and repetitive sequences (**Chapter 2**). Subsequent amplification of *Mdmd* appears to have led to the complex structure of the *M*-locus, as multiple tandemly repeated copies of *Mdmd* are found in M^{III} and M^{V} males in **Chapter 2**. After amplification, the *M*-locus may have translocated multiple times as a cluster from the Y to an autosome and/or subsequently between autosomes, generating novel Y-chromosomes. In addition, the data presented in **Chapter 2** revealed that to some extent different sequences exist in different autosomes, indicating that after translocation, the *M*-locus underwent further independent genomic changes on each autosome. The existence of multiple different autosomal *M* variants in the housefly provides a unique opportunity for further study of early stages of sex chromosome evolution.

As *Mdmd* is a crucial gene for male development, localising *Mdmd* mRNA in different embryonic developmental stages is needed to understand its regulation in the sex determination pathway. In **Chapter 4**, I demonstrate the ubiquitous expression of *Mdmd* mRNA throughout embryonic development. This suggests that *Mdmd* acts at a very early embryonic stage and that it needs to be continuously active in embryos to sustain male development. Sharma et al. (2017) showed that targeted disruption of *Mdmd* turns genotypic males into females. Although this indicated that *Mdmd* plays a crucial role in male development, it did not proof that *Mdmd* is sufficient for male determination. To test whether *Mdmd* is solely sufficient to perform the male-determining function, in **Chapter 4**, I introduced $Mdmd^{V}$ mRNA into early blastoderm stage embryos from the M^{III} strain and tested for sex-reversal. Transient expression of $Mdmd^{V}$ mRNA in female embryos did not yield any masculinised flies, although an insignificant bias towards more males was observed in injected offspring. These results either indicate that expression of $Mdmd^{V}$ alone is not sufficient to turn genotypic females into males, or alternatively, it is caused by an experimental shortcoming, i.e. insufficient translation of $Mdmd^{V}$ mRNA. An alternative approach to determine whether expression of *Mdmd* is sufficient to turn genotypic females into males, would be to use *piggyBac* germline transformation to repeatedly express $Mdmd^{V}$ during development. In **Box 4.1**, I describe how I constructed a pBac[3×P3-EGFP, hsp70-$Mdmd^{V}$] transgene. This transgene will be used in future experiments to assess the masculinising activity of $Mdmd^{V}$.

My work has shed light on the complex structure of the *M*-loci in the housefly and on the evolution of sex chromosomes in the housefly and in insects in general.