國立臺灣大學生物資源暨農學院昆蟲學系

碩士論文

Department of Entomology
College of Bioresources and Agriculture
National Taiwan University

Master Thesis

熱帶火蟻之性別決定機制探討

What type of sex determination does *Solenopsis geminata* use?

A focus on the homologous *SDL* region

陳沐恩

Mu-En Chen

指導教授：王忠信 博士

黃榮南 博士

Advisor: John Wang, Ph.D.

Rong-Nan Huang, Ph.D.

中華民國 106 年 1 月

January 2017

doi:10.6342/NTU201703365

國立臺灣大學（碩）博士學位論文
口試委員會審定書

熱帶火蟻之性別決定機制探討

What type of sex determination does *Solenopsis geminata* use?

本論文係陳沐恩君（R03632006）在國立臺灣大學昆蟲學系、所完成之碩士學位論文，於民國 106 年 1 月 13 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

王忠信（指導教授）　　王淑芳（指導教授）

王照林　　林宗岐

王弘毅

系主任　張俊彥

doi:10.6342/NTU201703365

致謝

事實上，這篇論文的完成讓我百感交集。從小我對於大自然的各樣事物就充滿了熱情，希望可以知道更多一點。但不幸的是，”興趣大於能力” 也同時一直是我身上的標籤。這樣的情況持續到我碩士班第一年，甚至一度在實驗室待不下去。但感謝上帝，在這裡相遇相處的人們都或多或少的幫助著我，無論是我最廢最無貢獻的時候還是我近期比較有 ”產出” 的時候我都可以感受到大家的溫暖與支持。謝謝楊景程學長與我分享關於螞蟻獨特的一切，而且更重要的是若沒有約六年前您的邀約，我現在完全不可能在這個實驗室裡，進行這些研究。感謝李志琦學長與黃裕清學長，謝謝您們與我一起發想討論，無論是當時有用的沒用的都是很棒的啟發，尤其您們分享的很多關於入侵紅火蟻的資料讓我可以一一比對並尋找適合我的研究的方法，這幾乎是這篇論文的根基。謝謝張家寧(Tiffany)學姐、高嘉宜學姊、蕭丞凱學長、鄧小美(Viet Dai Dang)學姐、丘祐坤學長，謝謝您們與我分享很多技術相關的事務，完成這篇論文您們真的幫助很大。謝謝張倪禎、羅韻華、許蓉禎、楊昉蓉、方森雅(Silvia Fontana)，您們的鼓勵、支持與與日常的垃圾話(其實想一想這些真的很有幫助)給了我很多繼續下去的動力。謝謝我的好同學張少濬還有林羿岑，你們給了我很多很實際的幫助，讓我不至於因為實驗室不在學校而與世隔絕。當然，真的非常感謝黃榮南老師願意收留我以及 John 的耐心指導，沒有您們我現在絕對不是如此。最後，感謝我的家人接納我壓力很大的壞脾氣以及持續的支持我。當然，感謝我的上帝一路的帶領，這篇論文為您而發。謝謝大家。

doi:10.6342/NTU201703365

中文摘要

對於有性生殖之生物而言，性別決定是非常重要的。多數昆蟲性別決定機制之下游部分是相當保守，但上游基因卻不然。入侵紅火蟻(*Solenopsis invicta*) 是使用單基因座互補性別決定機制（single-locus complementary sex determination, sl-CSD）的真社會性膜翅目昆蟲。 因此，異型合子（heterozygosity）個體將發育成雌性，但半型合子（hemizygosity）或同型合子 (homozygosity) 個體則會發育為雄性。 我們實驗室過去的研究顯示入侵紅火蟻決定性別基因有別於已被深入研究之另一真社會昆蟲蜜蜂 (*Apis mellifera*)，我們將其命名為 *sex determination locus* (*SDL)*。 即便目前尚未得到此基因之序列，此基因已被證實位於一高度變異區域（hypervariable region）。本研究欲確認另一同屬的螞蟻：熱帶火蟻 （*Solenopsis geminata)*是否亦是使用此新發現之性別決定基因，但由於目前為止尚未有雙套體雄蟲之紀錄，因此本研究將測試雌性個體是否全為異型合子。本研究記錄了與入侵紅火蟻 hypervariable region 同源區域附近的 10 個微衛星（microsatellite）基因座之基因型，並檢定其是否如同 CSD 機制一樣偏離哈溫平衡 （Hardy-Weinberg equilibrium），結果顯示並未有顯著不同，後續的模擬資料顯示這是由於樣本數不足而造成。因此，這些微衛星資料被組成單倍群（haplotype)， 並確認是否有任一區域是如同 CSD 機制預期之雌性個體均為異型合子，合乎預期之位置在一橫跨連續五個基因座的區域內被找到，且與隨機得到相同結果的機率有顯著不同。另外，我們亦比較了 hypervariable region 序列在此兩種火蟻間的區別。系統發生樹

doi:10.6342/NTU201703365

（phylogenic tree）顯示在此區域有跨種多型性（trans-species polymorphism）之現象，而在平衡選汰（balancing selection）下此為一常見之模式。 CSD 機制必然與平衡選汰相關，因此綜合上述證據，我們可以間接了解此兩種全球入侵性螞蟻，即便分別已久，SDL 同源區域仍在其性別決定機制中扮演著重要的角色。

關鍵字：性別決定機制、入侵紅火蟻、熱帶火蟻、平衡選汰、跨種多型性

doi:10.6342/NTU201703365

Abstract

Sex determination is absolutely critical for sexual organisms. For most insect species the downstream components of the sex determination pathway are conserved, but the upstream genes and mechanisms are very diverse. The red imported fire ant (*Solenopsis invicta*, RIFA) is a eusocial Hymenoptera which uses the single-locus complementary sex determination (sl-CSD) mechanism whereby heterozygosity at a single sex locus results in females and hemizygosity (or homozygosity) at this locus yields males. Our lab has shown that the master trigger gene for RIFA sex determination, *sex determination locus* (*SDL*), is different from that of the honey bee (*Apis mellifera*). Although SDL has not been cloned in RIFA, it is characterized by a hypervariable region. I tested whether the tropical fire ant (*Solenopsis geminata*, TFA) also uses SDL, which I hypothesize is likely because these two species are congeneric. Conservation of function between these two species predicts that TFA SDL will always be heterozygous in females and homozygous in males. Diploid males were not available so this study focused on females.

I genotyped 10 microsatellite loci in TFA that are putatively orthologous to RIFA loci near the hypervariable region. Absolute female heterozygosity would violate Hardy-Weinberg equilibrium, but analysis of these microsatellite data could not rejected the Hardy-Weinberg model, likely because of insufficient power given the large number of sex alleles. Nevertheless, by constructing 5-loci haplotypes, I never found any homozygous

doi:10.6342/NTU201703365

haplotypes in females, consistent with the CSD model. In addition, the phylogenic tree of the hypervariable region and the nearby gene *EGF-like* revealed trans-species polymorphisms between TFA and RIFA indicating that the alleles may be older than these two species and suggesting that *SDL* is likely evolving under balancing selection. Together my results indicate that the homologous SDL locus could be important for sex determination in *S. geminata*.

Keywords: sex determination, fire ant, *Solenopsis invicta*, *Solenopsis geminata*, balancing selection, trans-species polymorphism

doi:10.6342/NTU201703365

Table of Contents

doi:10.6342/NTU201703365

doi:10.6342/NTU201703365

doi:10.6342/NTU201703365

List of Figures

doi:10.6342/NTU201703365

List of tables

doi:10.6342/NTU201703365

Abbreviation Table

Abbreviation	Full name
Blastn	Basic Local Alignment Search Tool
CSD	Complementary Sex Determination
INDEL	Insertion and Deletion
MAFFT	Multiple Alignment using Fast Fourier Transform progress
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RIFA	Red Imported Fire Ant
TFA	Tropical Fire Ant
TSP	Trans-Species Polymorphism

doi:10.6342/NTU201703365

1. Introduction

Life on Earth can be divided into asexual and sexual organisms. Compared to the simpler self-replication of asexuality, sexual reproduction is more complex and energetically expensive. Despite these costs, there are some distinct advantages for sexual reproduction. First, sexual reproduction involves the combining and mixing of genetic traits through their gametes which permits faster evolution and adaptation. Second, sexual reproduction can purge the genome of deleterious mutations through meiotic recombination (Kondrashov, 1988; Muller, 1964) and thereby also increasing the fitness of a population.

In animals, males and females are often sexually dimorphic. Examples include the larger body size of female spiders or the horns of male goats. An individual's sex is also associated with sex-specific behaviors. These sex-specific morphologies and behaviors often intersect through sexual selection. Thus, proper sex determination is absolutely critical for the fitness of sexual organisms.

1.1 Sex determination systems

Although some species employ environment-dependent sex determination (e.g., temperature-dependent sex determination), the majority of species use genetic sex determination mechanisms with the most familiar being the XY sex determination system. In this system, females are homozygous for one of the sex chromosomes, XX, and males are heterozygous, XY. This system is referred to as male heterogamety and includes humans and most vertebrates. Another kind of sex determination is female heterogamety, or the ZW system, where females are heterozygous for the sex chromosomes, ZW, and males are homozygous, ZZ. This second system is very common in avian and lepidopteran (butterflies and

doi:10.6342/NTU201703365

moths) species. A third system, which is the focus of this study, is haplodiploid sex determination where females are diploid and males are haploid (Heimpel and de Boer, 2008). Wasps, honeybees, and ants employ this. Outside of the honeybee, the genetic and molecular mechanisms in hymenoptera are unclear.

1.2 Core sex determination genes are conserved

Even though there are many mechanisms to initiate the sex determination pathway, they all seem to converge on two components. The first is the splicing factor, *transformer* (*tra*), which causes sex-specific splicing of the second component, the transcription factor, *doublesex* (*dsx*). In contrast to these two genes, the master trigger genes and mechanisms are very diverse (Gempe and Beye, 2011). The fruit fly (*Drosophila melonogaster*) uses *sex-lethal* (*sxl*) to determine their sex (Bell *et al.*, 1988). The housefly (*Musca domestica*) uses *Musca domestica male determiner* (*Mdmd*) to determine whether an individual will develop into a male (Sharma *et al.*, 2017). The red flour beetle (*Tribolium castaneum*) and a wasp (*Nasonia vitripennis*) both have maternal control of their sex determination, although the maternally regulated genes are different (Beukeboom and Van De Zande, 2010; Shukla and Palli, 2012). Finally, honeybees (*Apis mellifera*) use the complementary sex determination (CSD) system via the eponymously named gene, *csd* (Beye *et al.*, 2003).

The master trigger genes evolve relatively fast (Bachtrog *et al.*, 2014). The housefly is a good example. In most housefly populations, the *male-determining factor* (*M factor*) is on the Y chromosome and inhibits the feminizing factor which is on the X chromosome. However, in other populations the M factor may be located on a former autosome or even on the prior X chromosome (Inoue *et*

doi:10.6342/NTU201703365

al., 1983).

1.3 Honeybees and the red imported fire ants (RIFA)

Honeybees are haplodiploid organisms, and also are the most well understood hymenopteran system. The discovery of viable diploid males permitted the genetic dissection of the honeybee sex determination pathway. They use the complementary sex determination system. In this mechanism, there are one or several loci that determine sex. The simplest case is single locus complementary sex determination (sl-CSD) where the individuals who are heterozygous at this locus develop into females. Individuals that are homozygous or hemizygous at this locus, develop into males (Ross, 1985). It is worth noting that diploid males are often sterile.

At the molecular level, when a honeybee individual is heterozygous at the master trigger gene, *csd*, the protein product of this gene causes female-specific splicing of the downstream gene *feminizer* (*fem*) (Gempe *et al.*, 2009). Then the female-specific isoform of fem causes female-specific splicing of *doublesex* (*dsx*). Note that *fem* is a homolog of the gene *tra* and *csd* is a paralog of the gene *fem* (Biewer *et al.*, 2015). On the other hand, for individuals that are homozygous or hemizygous for *csd*, the *csd* protein product does not promote female-specific *fem* splicing; *fem* is spliced to the default male form. Consequently, *dsx* is also spliced as the male form. Thus like in other insects, the sex-specific isoforms of *dsx*, ultimately regulate proper sexual fate (Cho *et al.*, 2007).

Although the honeybee and the red imported fire ant (*Solenopsis invicta* Buren, RIFA) are both eusocial Hymenoptera and use the single-locus CSD mechanism, quantitative trait locus (QTL) analysis and genetic mapping

doi:10.6342/NTU201703365

conducted in our lab has revealed that the RIFA sex locus of RIFA is not at the *tra* homologous region. Instead, the sex locus, which we call *sex determination locus* (*SDL*), is on a different chromosome (Huang *et al*., unpublished).

Hymenopteran species are no exception with respect to quickly evolving upstream sex determination genes. In addition to the fire ant case, allelic variability of the *tra* homolog or paralog is absent from the genome of bumblebee (*Bombus terrestris* and *Bombus impatiens*) that indicates they determine their sex by something else (Sadd *et al.*, 2015). To understand the evolution of the master trigger gene for sex in the genus *Solenopsis*, the related tropical fire ant (*Solenopsis geminata*, TFA) is examined to determine whether it also uses *SDL* like RIFA or some other gene(s).

1.4 Tropical fire ant (*Solenopsis geminata* Fabricius, TFA)

TFA is an invasive species. Their native habitat is in Central America and northern South America. The current invasion model is that they spread throughout the world from the southwest of Mexico because of global commerce in the 16th century to Manila. After TFA moved from Manila to Taiwan, and then from Taiwan to many Old World tropical regions (Gotzek *et al.*, 2015). In Taiwan, they are located in the southern, central (Lai *et al.*, 2009), and eastern (Retrieved July 24, 2017, from http://taibif.tw/zh/namecode/341929) regions. They are a very common ant species currently and have even become agricultural pests.

Phylogenetically, both RIFA and TFA belong to the genus Solenopsis, but TFA is classified as a relatively basal species in the “fire ant” group based on morphology (Pitts *et al.*, 2005) and the mitochondrial genome (Gotzek *et al.*, 2010). Furthermore, the genome of RIFA is already available (Wurm *et al.*, 2011),

doi:10.6342/NTU201703365

and TFA was chosen to compare to the RIFA sex determination region because they are common in Taiwan and their genome is also available (C.C. Lee and J. Wang, personal communication).

1.5 Key questions and the hypothesis

In this study, I used the two invasive ants, TFA and RIFA, to investigate the evolution of the master sex determination genes after speciation. Conservation or diversification of function at the orthologous sex locus between these two species will be informative of the rate of sex locus change in ants, at least for the *Solenopsis* clade. I hypothesized that TFA uses the orthologous SDL locus as RIFA for sex determination because of their phylogenetic relatedness. Other mechanisms are possible, and examples occur in Hymenoptera. For instance, sex determination may be multiple-locus CSD like some parasitic wasps (Snell, 1935) and the ant *Vollenhovia emeryi* (Miyakawa and Mikheyev, 2015). In another case, the master trigger gene may be epigenetically regulated such as in the jewel wasp *Nasonia vitripennis* (Verhulst *et al.*, 2010).

In my hypothesis, we would expect to observe that all females are heterozygous at the SDL locus and diploid males are homozygous. In principle, the strongest evidence would come from diploid males. However, there are no reports of diploid males in TFA. This may be partly due to the fact that all reports of TFA in Taiwan are monogyne colonies (with only one queen) (Lai *et al.*, 2009), which are unlikely to produce diploid males. In RIFA, newly "match-mated" queens can survive in polygyne colonies but will fail to found independent monogyne colonies because the diploid males take too much of the resources from the first batch of workers (Ross and Fletcher, 1986). TFA is likely to be

doi:10.6342/NTU201703365

similar to RIFA in this regard. Thus, obtaining diploid males is extremely unlikely at this time.

An alternative tactic is to test if the sex locus is always heterozygous in females. The risk of this method is that if there are many alleles at this locus, as predicted from theory for complementary sex determination, then the expected heterozygosity may already be high from simple random mating. Therefore, very large sample sizes may be necessary. Nevertheless, this is the method I have chosen because it is the only obtainable evidence.

1.6 General research strategy

I conducted three genetic analyses to test if the homologous *SDL* locus may function as the sex locus in TFA.

First, since the CSD mechanism implies that all heterozygous individuals develop into females, the prediction is that female data should not fit the Hardy-Weinberg equilibrium model. This test assumes an ideal population genetic model with infinite size and random mating as well as no mutation, migration, or selection. Since the genotypes of workers are identical to that of virgin queens (next generation queens), the analysis of worker genotypes should reflect the allele frequency of colonies in next generation. Therefore, I obtained microsatellite genotypes from many workers and conducted a Hardy-Weinberg equilibrium test of this basic population parameter.

Microsatellites are short tandem repeats in a genome that often have length polymorphism among individuals. Multiple loci can be composed into haplotypes. We can quickly determine if a diploid individual is heterozygous or homozygous through this method. In addition, I collected additional information of unlinked

doi:10.6342/NTU201703365

microsatellite loci as a control group for comparison. Analysis of single microsatellites could be problematic if the mutation rate of the microsatellite markers is much faster than for the *SDL*, resulting in some true homozygotes at *SDL* appearing falsely heterozygous. The use of multilocus haplotype markers can circumvent this issue. Thus, for the Hardy-Weinberg analyses, I constructed and examined microsatellite the haplotypes of the microsatellite loci that were tightly linked to *SDL*.

Second, Hardy-Weinberg equilibrium just tests for a reduction in female homozygosity, but actually the sex locus should never be homozygous. Direct inspection of the microsatellite haplotypes may be informative. Specifically, homozygous regions can be used to exclude sex locus regions. Conversely, invariably heterozygous regions would suggest candidate sex locus regions.

Third, the CSD sex locus should be under balancing selection because the fitness of heterozygotes is greater than homozygotes (heterozygote advantage) and lower frequency alleles have a selective advantage (negative frequency-dependent selection). Thus, selection would maintain or even increase multiple alleles over long evolutionary time (Garrigan *et al.*, 2003); some alleles can become trans-species polymorphisms (TSP). TSP means that the allele sequences are more similar among species than within species (Klein, 1980). If three alternative possibilities, convergence (Andersson *et al.*, 1991), introgression (Wegner and Eizaguirre, 2012), and new speciation (Nagl *et al.*, 1998) could be excluded, TSP would be the strongest evidence for balancing selection (Takahata, 1990).

I examined the phylogenetic trees of the homologous regions which are tightly link to *SDL*. For controls, I also examined some neutral or independent

doi:10.6342/NTU201703365

region. Finding evidence for balancing selection at the target region would indirectly support the hypothesis that TFA uses the same or similar sex determination mechanism with RIFA.

doi:10.6342/NTU201703365

2. Materials and methods

2.1 Sample collection

Tropical fire ant colonies were collected in central and southern Taiwan between 2014 November to 2015 October. The locations were the eastern district of Tainan, Yunlin Dounan, Taichung Wurih, and Taichung Wuqi. Colonies and the surrounding dirt were dug out with a shovel and placed into buckets. Upon were returned to the lab, water was dripped into the buckets to slowly force ants to move to the surface. Unlike *S. invicta*, in which workers and queens efficiently move into paper cups placed on top of the dirt, the queens of S. geminata are difficult to collect. Thus, we modified the standard fire ant protocol. Instead, a bridge was placed from the bucket to a plastic box (39 * 28 * 11 cm) containing an artificial nest. This allowed us to verify that the brood and queen had successfully moved into the artificial nest.

I also obtained some individual samples collected by friends. In total, 35 colonies were included in this study. Sample information is in appendix 1.

2.2 Ant husbandry

Colonies were housed in boxes with the inside walls coated with fluon to prevent ant escape. Occasionally, if the fluon was degraded and a new fluon-coated box was unavailable, baby powder would be applied to the rim of the box to temporarily prevent escape. Colonies were provided with 15 cm diameter petri dishes whose bottom were covered with plaster and to which water was added as needed to maintain humidity for shelter. Test tubes with water and plugged with cotton were also provided. All colonies were fed mealworms, crickets, and seeds 3 times per week. Temperature was maintained at 27^0C to 30^0C and the humidity was about 60%.

doi:10.6342/NTU201703365

2.3 Microsatellite data analysis

2.3.1 The reference genome

The TFA reference genome was sequenced and assembled by our lab. There are three versions. The first version was assembled de novo from Illumina Solexa data (Bentley *et al.*, 2008) derived from DNA extracted from a single TFA haploid male. The second one was assembled *de novo* from DNA extracted from a worker pool and sequenced using the PacBio Single Molecule, Real-Time (SMRT) Sequencing platform (Eid *et al.*, 2009). This worker pool was from another colony and may include three different alleles. The last one was another assembly from the raw-data of the second version but with the potential alternate allele contigs included (C.C. Lee and J. Wang, personal communication). The third assembly is also the sequence which was analyzed, see below.

2.3.2 Primer design

In a RIFA study, 15 microsatellite primers were used to help map the sex determination locus (Huang et al., unpublished). Of these, 11 were near the target site and 4 others were on different chromosomes. To check if these primers could be used in this study, their sequences were compared to the TFA draft genome using the nucleotide basic local alignment search tool (blastn) (Altschul *et al.*, 1990). Thirteen of the primer pairs were identical to at least 15 nucleotides at the 3' end, suggesting that they could be used in TFA. PCR tests for these 13 primer pairs all produced amplification products. For the fourteenth primer pair (Sol20), the fifth nucleotide from the 3' terminal in TFA is divergent from that in RIFA but PCR tests still yielded a product. The primers for the fifteenth primer pair, SDL6, differed at the last nucleotide at 3' end, so it was excluded in this study (Appendix 2).

For these 14 useable primer pairs, a fluorescent dye was added to the 5’ end of each

doi:10.6342/NTU201703365

forward primer while a “PIG-tail” sequence (5'-GCTTCT) was added to the 5' end of the reverse primers to facilitate accurate genotyping (Brownstein *et al.*, 1996). The primers were divided into six groups according to their fluorescence marker, size, and PCR difficulty (e.g., GC content; see also Appendix 3).

2.3.3 DNA extraction

Since there should be 2 worker genotypes in each presumably singly-mated queen colony, five random individuals were initially chosen from each colony to reduce the probability of obtaining only one genotype. Additional individuals were sampled if initial genotyping results were unclear. In total, 4 to 16 individuals were successfully genotyped per colony. Each sample was placed in a tube with 300 µl of a solution containing LabTurbo buffer LTL (195 µl), PBS (75 µl) and proteinase K (30 µl of 600 mAU/ml). Next, samples were snap frozen in liquid nitrogen prior to homogenization. Then, samples were homogenized by adding ceramic beads and beating with a bead shaker or by grinding with a plastic pestle. Subsequently the samples were heated overnight at 56^0C and then DNA was extracted using the LabTurbo® Genomic DNA mini Kit. The DNA was eluted in a final volume of 40 µl and stored at -20^0C.

2.3.4 Multiplex PCR

The general polymerase chain reaction (PCR) mix was 10 µl, consisting of 1 µl of 0.29 to 159 ng/µl of DNA, 1 µl of 10x Super-Thermo Gold Buffer, 0.8 µl of dNTPs (2.5 mM each dNTP), 0.1 µl of Super-Thermo Gold Taq DNA Polymerase (5 U/µl), 0.2 µl each of each set of forward and reverse primers (10 mM/µl), and water for the remaining volume. For the GC-rich microsatellite set, the PCR mix was modified with 2 µl of 5x Q-solution (Qiagen) replacing an equal volume of water. PCR reactions consisted of an initial denaturation temperature of 94^0C for 10 min, followed by 10 “touchdown” cycles,

doi:10.6342/NTU201703365

25 normal cycles, and a final extension at 72^{0}C for 30 min. The touchdown cycles consisted of denaturation at 94^{0}C 30 sec; annealing at 60^{0}C for 45 sec and decreasing 0.5^{0}C per cycle; and then extension at 72^{0}C for 1 min. The normal cycles were: denaturation at 94^{0}C for 30 sec, annealing at 55^{0}C for 45 sec and extension at 72^{0}C for 1 min. PCR reactions were carried out in an ABI 9700 thermal cycler (Applied Biosystems). All PCR reactions were checked by gel electrophoresis using 1% TBE agarose gels. The products of successful PCRs were then analyzed (below).

2.3.5 Capillary electrophoresis and microsatellite allele identification

Capillary electrophoresis of the multiplex microsatellite PCR products was carried out by Genomics BioSci & Tech. Because the primers used for PCR sometimes produced multiple peaks, microsatellite allele identities were called based on the following rules, ordered by importance. First, the signal was variable among individuals. Second, male samples exhibited only one peak and female samples had at most two peaks. Third, when there were "shadow" peaks, the strongest signal was chosen. While collecting diploid individual data, the strongest 2 different peaks with the same shadow pattern were chosen or it would be identified as 2 copies of the same allele. Fourth, if there was more than one peak with the strongest signal, then the peak whose size was the largest was chosen. Microsatellite allele calls are in Appendix 4.

2.3.6 Hardy-Weinberg test

In a singly mated, single-queen ant colony, there are two potential female genotypes, which differ at the maternal allele inherited. Using all the genotype data from a colony can create two biases. First, since both worker genotypes will inherit the paternal allele from the haploid father, analysis of all worker data will overweigh the paternal allele data by two. Second, multiple individuals from one colony will cause pseudo-replication. To

doi:10.6342/NTU201703365

avoid these cases, data was subsampled as follows. One random sample was chosen from each of the 35 colonies to construct a new dataset. This dataset was tested for departure from Hardy-Weinburg (HW) equilibrium using the "hw.test" function from the pegas package in R (alpha threshold was 0.05) (Paradis, 2010). Subsampling and HW testing was repeated 10,000 times. The final P-value was defined as 1 – [(counts of rejecting H0)/replicate number]. All tests were conducted in R version 3.3.1 (R Development Core Team, 2016).

2.3.7 Simulation data for the Hardy-Weinburg test

To check the statistical power of the Hardy-Weinburg test, a simulation was created in R. Three parameters were considered in this simulation: allele number, sample size, and the exact p-value. The last one was calculated as the extreme probability that the queens of all samples were heterozygous and not match-mated with allele frequencies weighted by total allele number. In the simulation, suppose there are n alleles in the population, then $\frac{1}{n}$ was used as the frequency for each allele because the probability for heterozygotes is highest. Additionally, the simulation chooses one diploid individual per colony as the sample. The calculation of the exact p-value is described as follows:

1. Suppose the mother's genotype is heterozygous and father's allele is a third (different) allele, then the frequency of this mated female is:

$$F_{ij} = f_i * f_j * (1 - f_i - f_j)$$

Here, fi and fj are the respective allele frequencies for alleles i and j and equal to $\frac{1}{n}$. In this situation, the sum of these probabilities is:

$$\Sigma i\,(\Sigma j\,(F_{ij})) = n^2 * \left(\frac{1}{n}\right)^2 * \left(1 - \frac{2}{n}\right)$$

doi:10.6342/NTU201703365

2. When the parents' genotypes only include 2 different alleles (e.g, i and j), there are 6 possible genotype combinations. In 2 combinations, the mother is a homozygote and the father has a different allele, resulting in all heterozygous progeny. In the remaining 4 combinations, the mother is a heterozygote and the father's allele is the same as one of the mother's alleles; all 4 of these will yield half homozygous diploid males. These homozygous probabilities are excluded from the final sum.

$$Prob_{Ho} = \left(\frac{4}{6} * \frac{1}{2} \right) * \Sigma k \left[\left(k^2 \right) * (1-2k) \right] = \left(\frac{n}{3} \right) * \left(\frac{1}{n} \right)^2 * \left(1 - \frac{2}{n} \right)$$

Where k is the total number of alleles and each allele's frequency is equal to $\frac{1}{n}$.

3. Finally, the formula of the exact p-value is:

$$\Sigma i \left(\Sigma j \left(F_{ij} \right) \right) - Prob_{Ho} = \left(\frac{n-2}{n} \right) * \left(1 - \frac{1}{3n} \right)$$

2.3.8 The determination of phased data and heterozygosity power analysis

Due to the haplodiploid system, phased data could be identified easily by the allele frequency from the same family. Generally, the allele frequency is 1 for alleles from the father and 0.5 for those from the mother.

The microsatellite genotypes from female samples were used to construct haplotypes. I considered haplotypes composed of 3, 4, and 5 adjacent loci. Since the CSD mechanism should only produce heterozygous females (and homozygous/hemizygous males), these haplotypes were examined for consistency with this model. The haplotypes were built using only phased data in the region to avoid noise such as recombination between haplotypes, at the cost of reduced sample sizes. Each haplotype was compared with the

doi:10.6342/NTU201703365

other haplotype from the same individual directly. If there was a locus which was never homozygous, I would calculate the exact p-value to express the relationship of this candidate region and the null hypothesis (Hardy-Weinberg equilibrium). Similarly, in the simulation, the exact p-value was calculate as the extreme probability that the queens of all samples were heterozygous but the precise allele number in all samples were used. Haplotype building and comparisons were performed in R.

2.4 SDL homologous sequence analysis

2.4.1 The homologous sequences of RIFA and outgroup species

In this study, the sequences of the Taiwan population and the population of the original habitat, South America, were analyzed. There are 25 sequences from the South American population, all from northern Argentina (A. Cohanim and E. Privman, personal communication). Since the coverage of these sequences was different, the analysis of different regions would not include all of them but the sequences which cover the analyzed homologous region. There were 10 sequences from the Taiwan population representing 10 different CSD allele sequences and all of them were include in every analysis (Huang et al., unpublished).

Monomorium pharaonis was used as the out-group species in this study. This is the species which is categorized in the same tribe, *Solenopsidini*, which includes TFA and RIFA. Additionally, the published genome of this species is the closest available one to *Solenopsis* genus (Mikheyev and Linksvayer, 2015) (GenBank assembly accession: GCA_000980195.3). The choice of this out-group species was to minimize possible long branch attraction which may result in random rooting (Kinene *et al.*, 2016). For genomic regions lacking data (sub-fragments, CoP, HVR and UFO) unrooted trees were constructed.

doi:10.6342/NTU201703365

2.4.2 DNA extraction

In order to get better sequence information, higher quality DNA was required. For this reason I modified the Puregene® DNA purification Kit (Gentra systems). In this part, all samples, include alates and workers, were disrupted and lysed as above, then the DNA was extracted with the kit. The DNA pellet was dissolved in 100 μl of DNA hydration solution and then stored at 4^0C.

2.4.3 Primer design

The master trigger gene for sex determination may be located in the hypervariable region. For convenience, we have divided it into three regions (CoP, HVR and UFO, see Fig 3) based on 2 conserved sequences found separating these fragments in RIFA (Huang et al., unpublished). Based on an alignment of 18 RIFA and 2 TFA hypervariable region sequences currently available, multiple primer pairs were designed to these conserved sequences to amplify each of the CoP, HVR, and UFO sub-regions. *EGF-like* is a gene adjacent to the hypervariable region. RIFA data showed the DNA sequence of the gene is slightly polymorphic associated with the different haplotype of the hypervariable region. Despite *EGF-like* being a conserved gene, since it is close to the hypervariable region, primer pairs for this gene were also designed.

2.4.4 Selection of candidate neutral loci

First, 7 contigs from the TFA PacBio genome assembly were selected randomly. Then I checked whether any >5kb region within each contig lacked any predicted open reading frame (ORF) >300 bp using the orffinder tool (https://www.ncbi.nlm.nih.gov/orffinder/). The candidate loci were also examined for any potential long coding RNAs by mapping to RIFA RNAseq data (Huang et al., unpublished) from each larval instar and sex. I chose a 1kb region satisfying these criteria on five of

doi:10.6342/NTU201703365

these contigs as the neutral control fragments for this study.

2.4.5. Cloning and sequencing

2.4.5.1 Specific fragment PCR

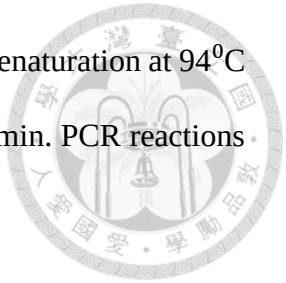
The EGF-like gene was amplified using the Kapa HiFi HotStart PCR Kit (Kapa Biosystems). The mix was 30 μl, consisting of 1 μl 0.19~159 ng/μl DNA, 6 μl 5x κ HiFi Buffer (Fidelity), 0.9 μl dNTPs (10 mM each dNTP), 0.5 μl κ HiFi HotStart DNA Polymerase (2.5 U/μl) and 0.9μl forward and reverse primers (10 mM/μl), and water for the remaining volume. PCR reactions consisted of an initial denaturation temperature of 95^0C for 3 min, followed by 35 normal cycles and a final extension at 72^0C for 30 min. The normal cycles were: denaturation at 98^0C for 20 sec, annealing at 58^0C for 15 sec and extension at 72^0C for 1.5 min. PCR reactions were carried out in an ABI 9700 thermal cycler (Applied Biosystems).

The conditions for the neutral control PCRs was the same as above, except from a premix containing both forward and reverse primers (10 mM /μl) was used. The PCR reactions consisted of an initial denaturation temperature of 95^0C for 4 min, followed by 35 normal cycles and a final extension at 72^0C for 10 min. The normal cycles were: denaturation at 98^0C for 30 sec, annealing at 60^0C for 30 sec and extension at 72^0C for 5min. PCR reactions were carried out in a SuperCycler Trinity (Kyratec.).

The target DNA was amplified using the Kapa Long Range HotStart PCR Kit (Kapa Biosystems). The mix was 30 μl, consisting of 1 μl 0.19~159 ng/μl DNA, 6 μl 5x κ LongRang buffer (without Mg^{2+}), 0.9 μl dNTPs (10mM each dNTP), 2.1 μl MgCL2 (25 mM), 0.5 μl κ LongRang HotStart DNA Polymerase (2.5 U/μl) and 0.5 μl forward and reverse primers (10 mM /μl), and water for the remaining volume. PCR reactions consisted of an initial denaturation temperature of 94^0C for 4 min, followed by 35 normal

doi:10.6342/NTU201703365

cycles and a final extension at 72^{0}C 7 min. The normal cycles were: denaturation at 94^{0}C for 20 sec, annealing at 50^{0}C for 15 sec and extension at 72^{0}C for 7 min. PCR reactions were carried out as above PCR machine.



2.4.5.2 Transformation, culturing, and sequencing

The EGF-like gene and neutral control fragment PCR products were separated by 1% TAE agarose gel electrophoresis at 60 V, 220 mA for 2 hr. Then, the gels with DNA were cut and DNA was purified using the Viogene® Gel / PCR DNA isolation systems. The DNA fragment was inserted into the pCRTM TOPO II blunt vector (Zero Blunt® TOPO® PCR Cloning Kit) and then plasmid was transformed into the E. coli competent cell DH5α (EGF-like gene was transformed into Fast-TransTM competent cells and neutral control fragments were transformed into RBC HIT competent cells). After 37^{0}C plate culture overnight the transformed colonies were confirmed by colony PCR. Positive colonies were used to inoculate a liquid culture overnight. Plasmid DNA was extracted using the QIAprep mini kit (Qiagen system) and then sent to the company Genomics BioSci&Tech for Sanger sequencing.

Cloning of the target (CoP, HVR, and UFO) DNA PCR products was similar as above with differences as follows. First, all transformations were with RBC HIT *E. coli* competent cell DH5α. Second overnight plating was at 30^{0}C, and then culture plates were incubated in 37^{0}C 4~6 hr prior to inoculating colonies into a liquid culture overnight. Third, all plasmids were checked for successful inserts by EcoRI restriction enzyme digestion for 2 hr in 37^{0}C followed by gel electrophoresis. Only clones with inserts were sent to Genomics BioSci&Tech for Sanger sequencing.

2.4.6 Analysis of nucleotide diversity and molecular evolution

All TFA sequences (this study) and RIFA sequences were combined and aligned by

doi:10.6342/NTU201703365

Multiple Alignment using Fast Fourier Transform program (MAFFT) (Katoh and Standley, 2013). Then I used the software Jmodeltest (Guindon and Gascuel, 2003) to determine the best substitution model under the Akaike information criterion. After, all the sequences were analyzed to make the phylogenic tree with 10,000 bootstrap replications in best model through the software MEGA 7 (Kumar *et al.*, 2016). For the nucleotide substitution models of the hypervariable region, gamma distributed General Time Reversible model (CoP, HVR) and gamma distributed with invariant site General Time Reversible model (UFO, *EGF-like*) were used. For the neutral regions, gamma distributed General Time Reversible model (LG1N), Hasegawa-Kishino-Yano model (LG3N, LG5N1 and LG10N) and uniformed General Time Reversible model (LG5N4) were used. The partial deletion option and the BioNJ initial tree were used on all trees. All trees were created with the online tool Interactive Tree Of Life (iTOL version 3.5.4) (Letunic and Bork, 2016).

Finally, the nucleotide diversity (π) was calculated by averaging the nucleotide difference per site between two randomly picked sequences extracted respectively from these two species or from one species. The Tajima's D statistic was calculated and show the p-value according to the coalescent model. These statistics were calculated using the DnaSP (Rozas *et al.*, 2003) software.

doi:10.6342/NTU201703365

3. Results

3.1 Microsatellite data and the Hardy-Weinburg test potentially suggest that the homologous hypervariable region could function in TFA sex determination.

3.1.1 Sequence data show that the SDL homologous region of TFA is similar to RIFA's.

In the aligned sequence (Fig.1) the genes and microsatellites appear to be in the same relative positions and orientation with the homologous sequence from RIFA. The only exception is microsatellite SDL11 whose nearby sequence did not overlap with the alignment. Importantly, the focal SDL candidate region was completely assembled so the subsequent analysis could be conducted. Therefore, I conclude that the region of TFA is similar to RIFA's.

3.1.2 Hardy-Weinberg equilibrium model could not be rejected based on single microsatellite locus analysis.

The CSD model presumes all heterozygous individuals develop into females and loci near CSD would violate HW equilibrium (null model), which assumes random sampling from a multinomial distribution. Thus, in principle, we can identify the candidate locus based on deviation from Hardy-Weinberg equilibrium.

The microsatellite data from 244 females from 35 colonies were tested for violation of Hardy-Weinberg equilibrium (Table 1). I assessed significance using 1000 adjusted-bootstrap replicates. This analysis revealed that no linked locus rejected the null hypothesis. Control unlinked loci (Sol20, SdagC487, SiMS2A-65, and Tramsa2) did not deviate from Hardy-Weinberg equilibrium, as expected. The analysis data are in Table 2.

Although the focal loci did not reject the Hardy-Weinberg model, the result does not

doi:10.6342/NTU201703365

exclude this SDL homologous region as the candidate. Analysis of only the female data for RIFA showed a similar result. The main evidence implicating the hypervariable region for RIFA is based on genotyping of diploid males, which was not possible for TFA. Given that there are probably many alleles at the sex locus in TFA, analysis of singlet microsatellite loci may not have enough power since some loci have only a few alleles.

3.1.3 The exact simulation p-value indicates that a sample size of 1.3 times the allele number is necessary to make the analysis significant.

We next considered if the failure to reject Hardy-Weinberg equilibrium could have been due to the lack of statistical power. To test this possibility, I conducted a simulation with different allele numbers (1 to 100) and sample sizes (1 to 150) and then calculated the exact p-value (Fig. 2A). For this analysis, the exact p-value is the extreme probability that all extracted samples are heterozygous when in Hardy-Weinberg equilibrium. This simulation shows that the slope of detected allele number and the minimum needed sample size is about 1.3 if the type I error rate is fixed at 0.05 (Fig. 2B). I conclude that I would have enough power to find the candidate region if it is heterozygous in all samples and the sample size (colonies number) is about 1.3 times the actual allele number.

3.1.4 Analysis of heterozygosity of haplotype blocks potentially suggest that the region between THUMP to SDL6-4 as the candidate SDL locus.

In the CSD mechanism, homozygosity at the region should never be observed in females. Building haplotype blocks may allow enumerating all the alleles. Therefore I built up haplotype blocks to do this. A haplotype is a group of genetic markers (SNP, microsatellite, etc), which is transmitted to the next generation together if there is no recombination. So it could be regarded as a genetic unit. Haplotype blocks consisting of 3, 4, or 5 neighboring microsatellite loci located near the SDL locus were examined in all

doi:10.6342/NTU201703365

female data. Attempts at using >5 loci were unsuccessful because enough recombination occurred among individuals within colonies, thus precluding analysis. In this analysis only phased data was used and to reduce noise, the alleles (haplotypes) with missing data were excluded in the analysis despite reducing the sample size. The results (Table 3) show that only the region from THUMP to SDL6-4 was never homozygous. This never homozygous region only included 19 colonies and the random chance of obtaining no homozygotes for this haplotype was 0.119. This result only weakly suggests a candidate region.

3.2 Balancing selection is supported in the homologous sequence in the hypervariable region.

3.2.1 The description of the hypervariable region and the nearby region and genes.

After aligning and comparing the sequences of TFA and RIFA together, the structure of this region which lacks homozygotes could be described (Fig. 3). The nucleotide diversity (π) and Tajima's D were also calculated (Table 4). Basically, like RIFA, the homologous hypervariable region (about 52.5 Kb) is between the same flanking genes *THUMP* (about 2 Kb) and *EGF-like* (about 1.5 Kb) and there are no other obvious genes in this range. In addition, based on two blocks of conserved sequences the hypervariable region was been divided into three sub-fragments: UFO, HVR and CoP. From the gene orientations, the first gene is *EGF-like*. This is a coding gene with epidermal growth factor-like domains. Then, next is one sub-fragment of the hypervariable region, UFO, which is about 5 Kb. There are several alleles here, each one being very different from the others. All sequences from TFA belong to the same allele family. A 0.8 Kb conserved non-coding region, N4R, follows. The next sub-fragment is HVR, including two parts.

doi:10.6342/NTU201703365

The part closest to N4R is about 3 Kb, is an extremely diverse region, contains some large indels (INsertion/DELetion, more than 100 bp), and belongs to a few discrete clades spread among all samples. It is worth mentioning that one of the indels is shared by one RIFA sample and one TFA sample (Fig. 4). The other part is an about 2 Kb region and converges into 2 conserved alleles. Again, all TFA sequences are categorized into one of these two alleles. The sequences continuously converge into an about 2 Kb conserved region with relatively few SNPs and few indels. After this region is the CoP sub-fragment which has an about 3 Kb diverse region that is just less polymorphic than HVR. This is the region that is covered by allele sequence that the alignment software suggests (see method 3.1). The neighboring region is an about 37.5 Kb intervening region between the hypervariable region and the gene *THUMP*. These final two regions are relatively conserved and were not further investigated here.

3.2.2 Gene tree analyses support trans-species polymorphism (TSP) in the hypervariable region.

The homologous TFA hypervariable region sequence was compared to the RIFA sequence using blast. In the TFA draft genome, the blast results showed that HVR was similar to haplotype 3 of RIFA, but UFO and CoP which flank HVR was most similar to haplotype 1 of RIFA.

The phylogenetic tree could describe the relationship among the TFA and RIFA sequences and provide hints for exploring the sex locus under the CSD mechanism. If the hypervariable region is truly the only sex locus in TFA, then it should be under balancing selection. Strong balancing selection is often associated with TSP because old alleles were kept that can even exceed the species age. In contrast, if this region is unimportant (i.e., TFA evolved another locus for sex determination), most alleles in TFA should be lost by

doi:10.6342/NTU201703365

drift or possible selective sweeps due to hitchhiking. In this case, TFA sequences should be an out-group to the RIFA alleles or derived from one of the RIFA alleles. Thus by examining phylogenetic trees of the candidate region, we can distinguish which hypothesis is more likely. The TSP phenomenon also could be caused by convergence, introgression and new speciation. Thus, in addition to the hypervariable region, I cloned 5 neutral sequences and compared them to exclude these three possibilities.

Thus far, 5 CoP sequences, 4 HVR sequence, 3 UFO sequences, 4 *EGF-like* sequences and 6 to 9 samples for each neutral sequence in TFA were cloned. Additionally, the existing TFA genome assembly sequence was added as another sequence set. These sequences were analyzed and phylogenic trees constructed for each phylogenic tree. In the phylogenic tree for *EGF-like*, there were 5 clades and 4 of 5 TFA sequences belonged to one of these. The other one looked different than others and was classified in another clade. There were 5 clades in the UFO tree and the distances between each other were relatively far. All TFA sequences obviously belonged to the same clade. The first part of HVR was most polymorphic and TFA sequences were spread among 3 clades in the tree. The other part was separated into two clades and all TFA sequences were grouped into one of these. Although the CoP sequences also could be divided to two parts, the results were similar in that one of 6 TFA sequences was discrete from the others (Fig. 5C to Fig. 5H). In sum, the phylogenic trees of CoP, HVR and the gene *EGF-like* all exhibited the TSP phenomenon, even though an out-group was only available for *EGF-like*. This is consistent with our prediction. In contrast, all neutral sequences showed that the TFA sequences were in one cluster with no evidence for TSP (Fig. 6A to Fig. 6E). Together, these data indicate that in TFA, the hypervariable region shares some alleles with RIFA. In contrast, the other regions cluster as a group from RIFA, even for a control region (LG3N) that was nearby, only <100 kb away.

doi:10.6342/NTU201703365

In fact, part of intervening region (about 3.5 Kb with out-group sequence available) and the gene *THUMP* were analyzed with only one TFA sample, many RIFA samples and the out-group (Fig. 5A, 5B). The intervening region tree showed a similar result: the TFA sample is mixed within the RIFA clade despite only one sample, Although this cannot be considered as a solid evidence, this region is obviously different from the phylogenetic tree for the gene *THUMP* (which has with the same sample size) where the TFA sample is independent of the RIFA clade. This result indicates that the TSP phenomenon is in the hypervariable region and is attenuated to both sides, at least to the gene *THUMP* and the neutral region LG3N.

doi:10.6342/NTU201703365

4. Discussion

In this study, I tested the homologous TFA SDL region by conducting Hardy-Weinberg tests, heterozygosity analysis, and examining the phylogenic trees of the target and neutral regions. Although my data are not definitive, the results from heterozygosity power analysis and phylogenic trees are consistent with the possibility that TFA also uses the homologous SDL locus for sex determination as RIFA does.

4.1 Structural similarity between TFA and RIFA at the SDL homology candidate region

I used 3 versions of the TFA genome sequence to assemble the SDL homology candidate region of TFA and found that it is similar to the structure of RIFA's. All microsatellite loci and the genes in this region are in the same relative positions and direction. In the initial draft genome of TFA, the relative position of the hypervariable region could not be identified until the second version (PacBio version) was available. This was because there were likely 3 different alleles at this locus preventing the software from assembling this region. This polymorphism phenomenon is consistent with the RIFA model.

The characteristics of each sub-region were also similar except the sub-fragment, UFO. In my TFA data, I only found one allele whereas many alleles were detected in RIFA. The lack of finding other alleles could be due to sampling bias, such as PCR bias. For example, failed PCR and cloning would result in less sequenced alleles.

4.2 Information from single microsatellite markers are insufficient to confirm or refute the sex locus.

The Hardy-Weinberg test is a simple way to examine if the region is operating under the CSD mechanism because the expectation is that there should never be homozygous

doi:10.6342/NTU201703365

females at the sex locus and reduced homozygosity at nearby linked loci. Analysis of the female data for each microsatellite locus revealed none with significant deviation from HW equilibrium. This result is not surprising because analysis of comparable female-only data from RIFA also showed a similar result. Thus, multilocus haplotype blocks were examined. I found that a haplotype block consisting of 5 microsatellites in the region from THUMP to SDL6-4 (which encompasses the hypervariable region in the SDL candidate region) was always heterozygous ($p \cong 0.12$) using high quality, but low sample size data.

4.3 The highly diverse haplotypes at the SDL region in TFA.

If TFA uses a sl-CSD mechanism and the sex locus is orthologous to the SDL region in RIFA, then a prediction is that the SDL region in TFA would contain many haplotypes. My results support this prediction. Based on microsatellite data of all samples, I found 115 all-loci haplotypes (with 10 homologous microsatellite makers) from 39 colonies. Additionally, the observed haplotypes seem different in each colony. This indicates that the amount of shared haplotypes among colonies is very low despite most of the data coming from one population, Taichung Wuqi.

In RIFA, there are 10 sex alleles in Taiwan and each allele corresponds to 1 haplotype constructed from microsatellites in the SDL region. My results could imply that TFA has 47 sex alleles depends on the heterozygous power analysis (see table 3). Alternatively, new haplotypes may have been created by recombination since TFA's arrival to Taiwan resulting in multiple haplotypes mapping to each sex allele. Recombinant haplotypes are more probable for TFA since they arrived more than 4 hundred years ago (Gotzek *et al.*, 2015), compared to RIFA which arrived in Taiwan more recently (Ascunce *et al.*, 2011). Another possibility is that some microsatellites have mutated by expansion or contraction, and thus creating new haplotypes. Resolution of these possibilities will require cloning of

doi:10.6342/NTU201703365

the locus and additional detailed sequence analysis of the region.

4.4 Balancing selection could be supported by the TSP phenomenon.

The sex locus should be under balancing selection which predicts that alleles could persist as trans-species polymorphisms (TSP). At the honeybee sex locus, TSP is also present in some related species (Cho *et al.*, 2006). Cho *et al* showed that the type 2 csd alleles of *A. mellifera* are more similar to *A. dorsata*, suggesting balancing selection works at the honeybee sex determination locus. In addition, the major histocompatibility complex of vertebrates (Takahata and Nei, 1990) and the self-incompatibility locus of plants (Vekemans and Slatkin, 1994) also show the TSP phenomenon under balancing selection.

I tested whether the TFA alleles might exhibit TSP by sequencing three different sub-fragments of the hypervariable region, the gene *EGF-like*, and neutral regions LG1N, LG3N, LG5N1, LG5N4 and LG10N. Indeed, this analysis revealed that the alleles for hypervariable region and the gene EGF-like for TFA are dispersed in the phylogenetic tree with RIFA except the sub-fragment, UFO. Yet, the homologous sequence of the out-group could not be found, and consequently the trees of sub-fragments were unroot trees. In contrast, none of the neutral regions exhibited the TSP phenomenon; all were clustered in one group which could be split from RIFA sequences.

Although consistent with balancing selection, TSP may also arise by recent speciation, introgression, and convergence. Recent speciation can be excluded because TFA is a basal lineage in the fire ants and split from the lineage leading to RIFA a long time ago (Gotzek *et al.*, 2010; Pitts and McHugh Ross, 2005). Recent introgression is also unlikely because the two species are not known to hybridize where they overlap.

Convergence means the independent evolutional lineages have with similar features.

doi:10.6342/NTU201703365

An argument against convergence is that evidence from another ant suggests that the sex locus may be ancient. In the ant *V. emeryi*, which uses 2 locus CSD, the location of one of the sex determination genes is close to the homologous hypervariable region and may suggest *SDL* is conserved from ancestral species. Thus, balancing selection is the most likely explanation.

The fact that *V. emeryi* uses 2-loci CSD raises an interesting possibility regarding sex determination evolution in ants. The second QTL locus maps to the homologous tra gene, which is homologous to the honeybee *csd* gene. One intriguing possibility is that fire ants in the native range actually use 2-locus sex determination, like *V. emeryi*, but because of the bottlenecks associated with fire ant invasions into the USA and subsequently elsewhere, there was loss of alleles at the *tra* locus. Consequently, the invasive range now only has single-locus CSD. This scenario, termed sex determination meltdown event, has been demonstrated for a parasitoid wasp, *Cotesia rubecula* (de Boer *et al.*, 2012). If this is true, then the evolution of 2-locus sex determination followed by meltdown of one allele may suggest a mechanism of how CSD loci appear to "move" during evolution.

doi:10.6342/NTU201703365

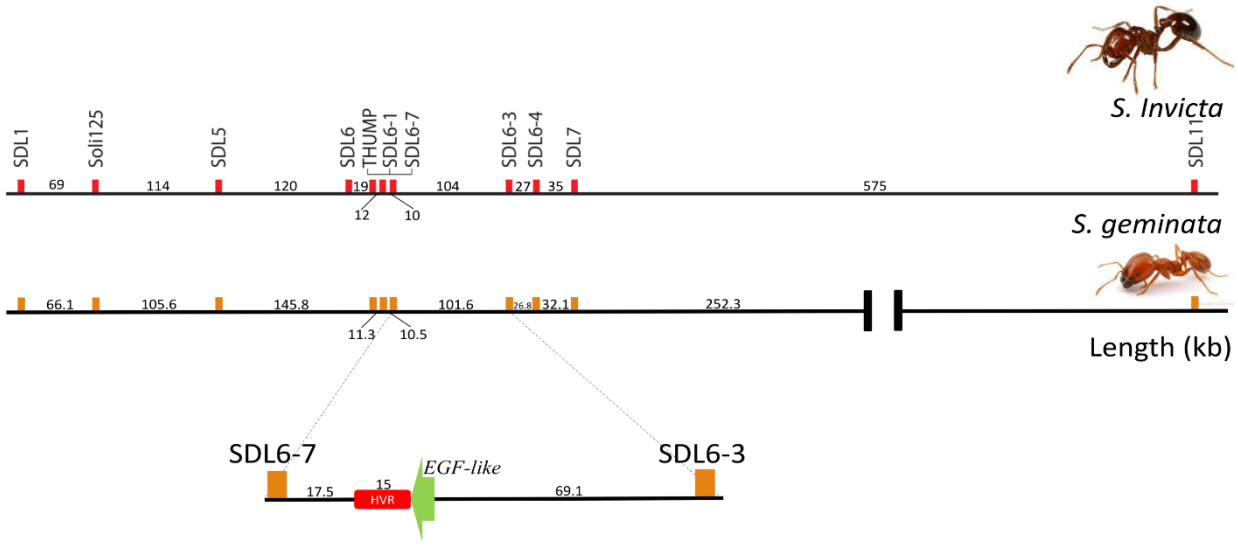
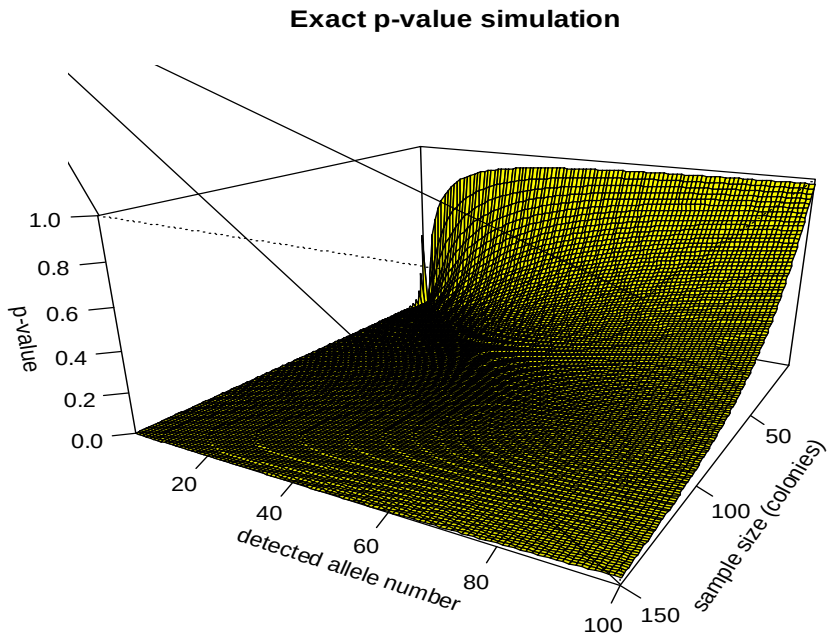


Figure 1. Comparison of the candidate sex locus within the SDL region of TFA and RIFA.

The TFA draft genome contains 3 overlapping fragments that align to the RIFA candidate sex locus. Using microsatellite loci which are linked to the RIFA SDL region as markers labels the homologous TFA SDL region. In addition to the hypervariable region (HVR) and the genes *THUMP* and *EGF-like*, the relative positions of all markers are similar between these two species. Only the SDL11 marker could not be confirmed because it is not assembled onto the same TFA contig as the other markers. Units are in kilobasepairs (kb).

doi:10.6342/NTU201703365

A.



B.

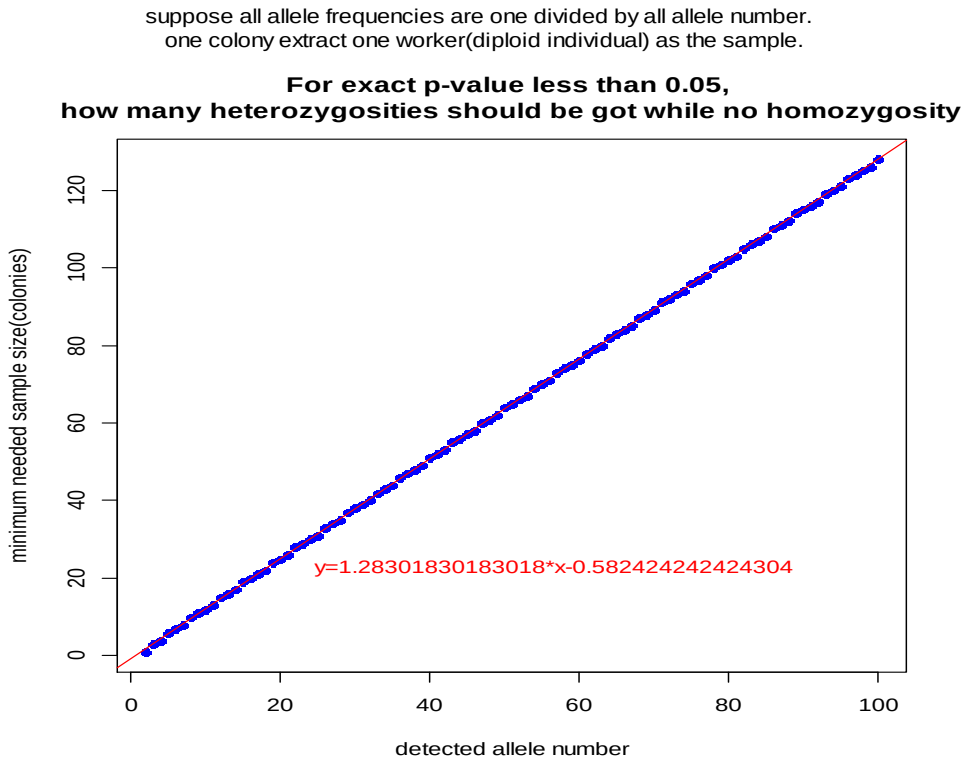


Figure 2. Hardy-Weinberg simulation analysis

(A)The simulation examined three parameters: allele number, sample size and the exact p-value. (B) When the type I error is fixed at 0.05, the largest exact p-value is similar to a line described by Y = 1.28 * X - 0.58.

doi:10.6342/NTU201703365

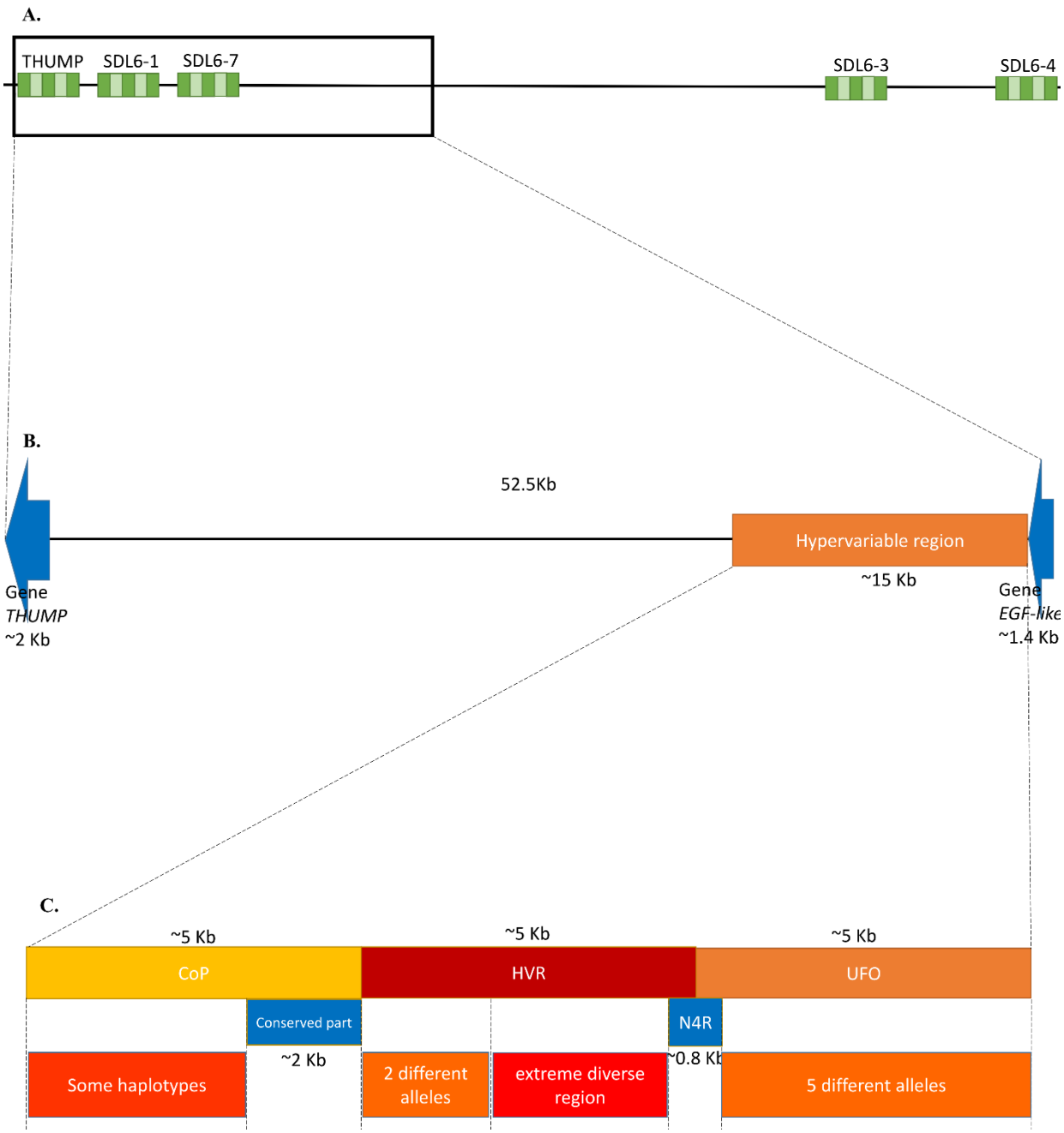


Figure 3. The hypervariable and surrounding region.

(A) Schematic of genomic region where all TFA females are never homozygous. This region encompasses 5 microsatellite markers, THUMP, SDL6-1, SDL6-7, SDL6-3 and SDL6-4. (B) Zoom of the region between the two genes, *THUMP* and *EGF-like*. (C) Detailed view of the region including CoP, HVR and UFO. Red, orange, and yellow are a rough indication of the diversity within each sub-region.

doi:10.6342/NTU201703365

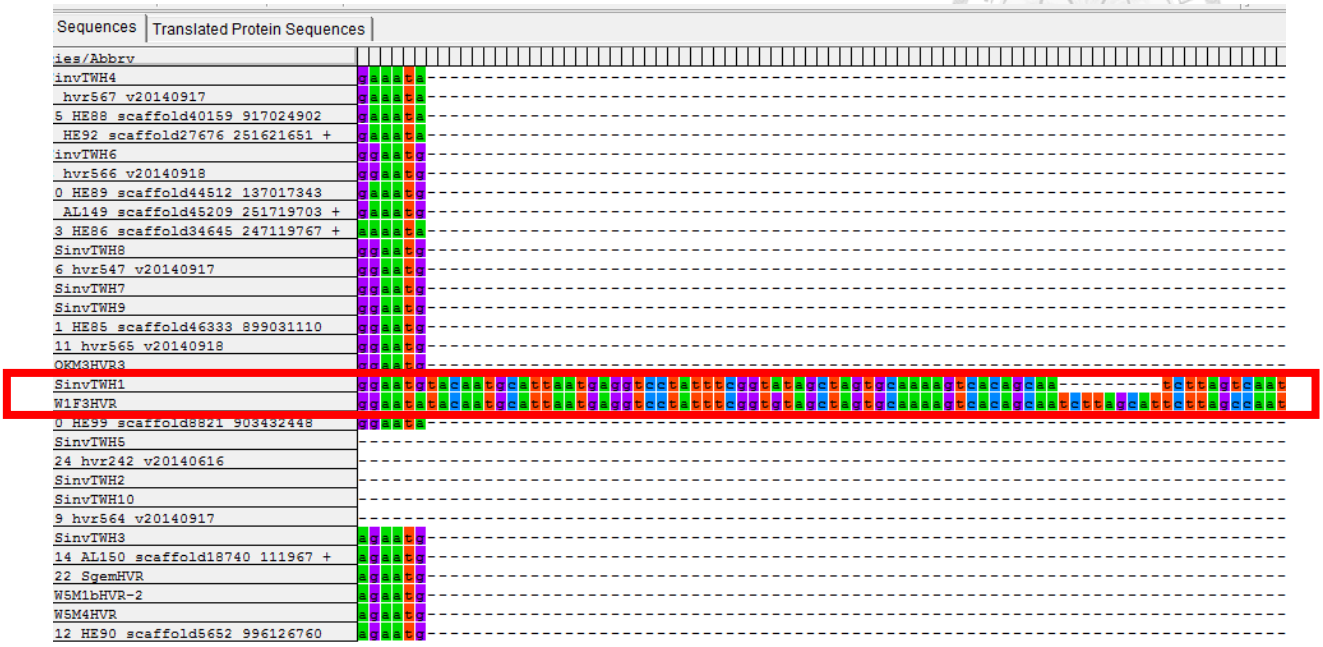


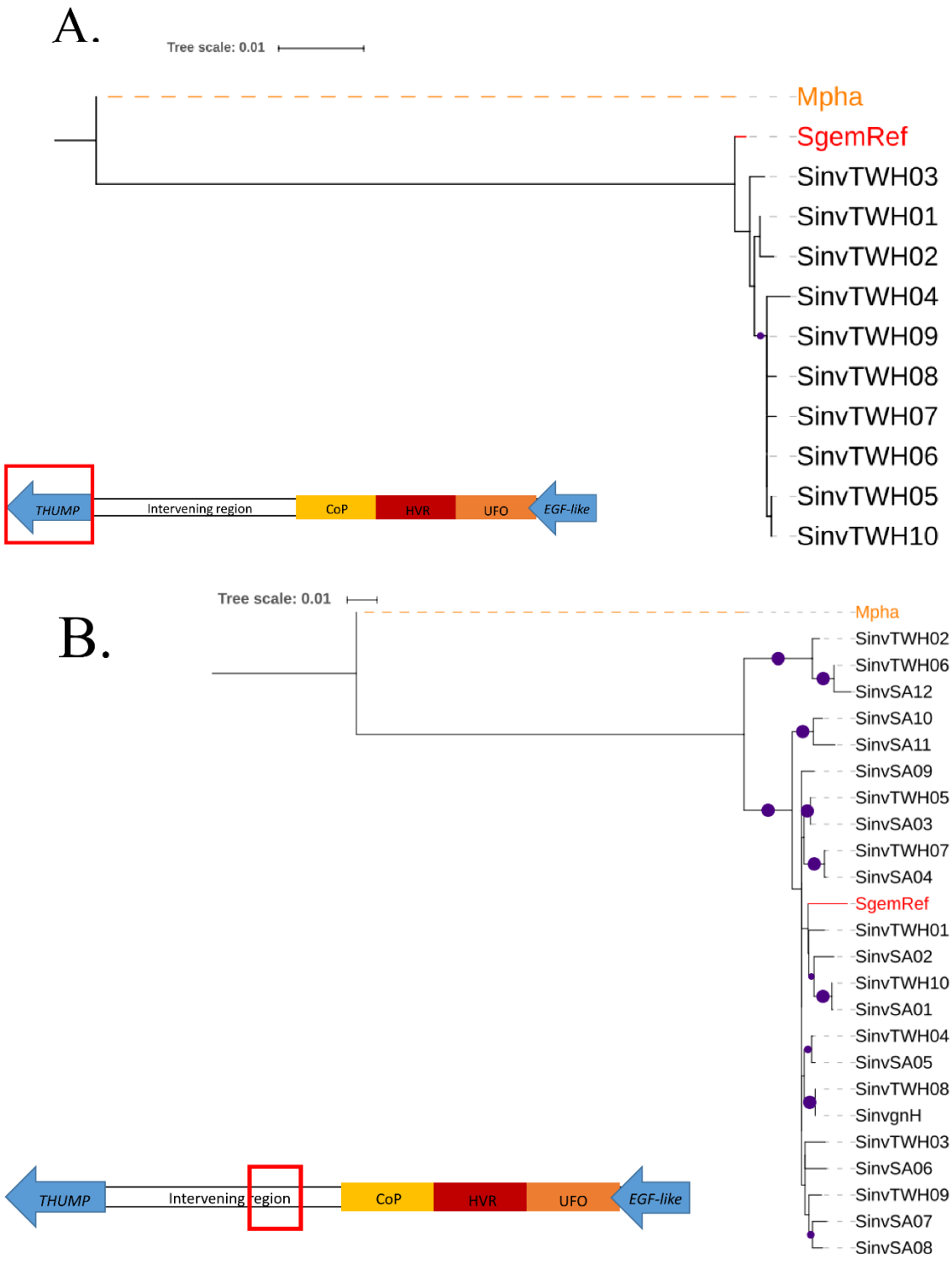
Figure 4. The shared "insertion" between TFA and RIFA.

A shared "insertion" between a TFA (W1F3HVR) and a RIFA (SinvTWH1) sample (red box). No other samples share this insertion. The length of this insertion is 167 bp.

doi:10.6342/NTU201703365

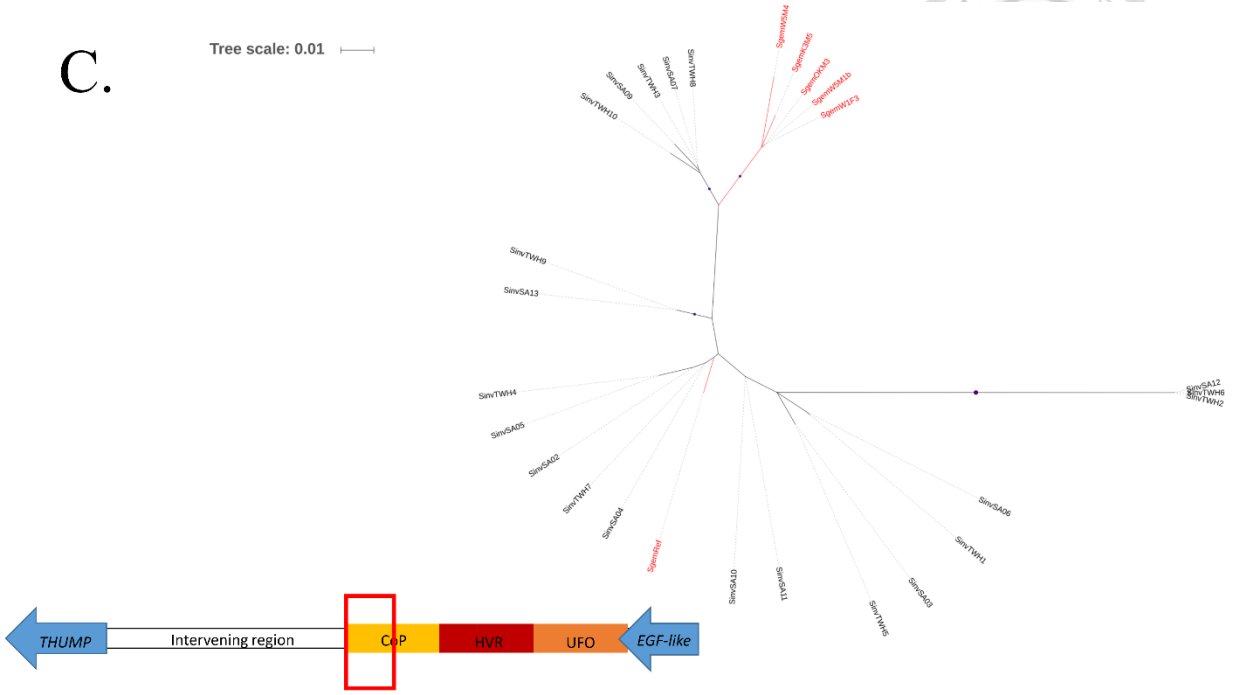
Figure 5. Phylogenic trees of the hypervariable and surrounding regions.

TFA samples, red; RIFA samples, black; out-group, orange. Purple dots indicate node support >70% by bootstrap replication. (A) The gene *THUMP*, (B) The intervening region, (C) CoP1, (D) CoP2, (E) HVR1, (F) HVR2, (G) UFO, (H) The gene *EGF-like*.

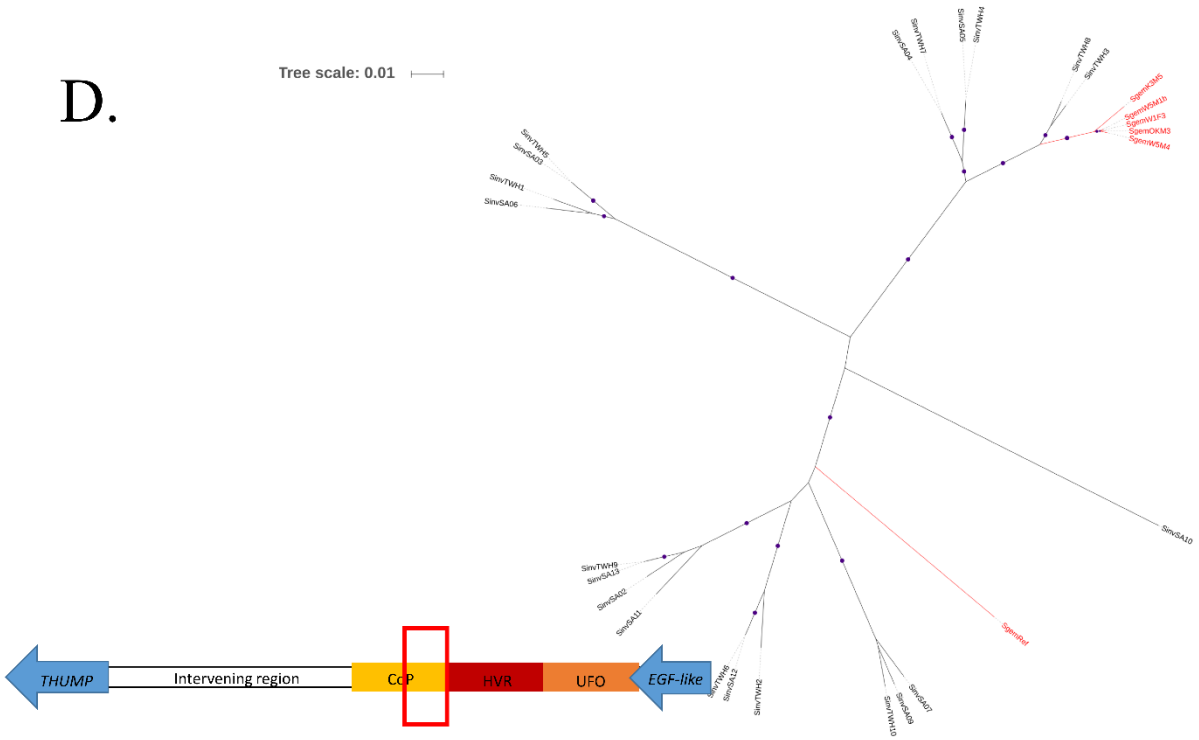


doi:10.6342/NTU201703365

C.

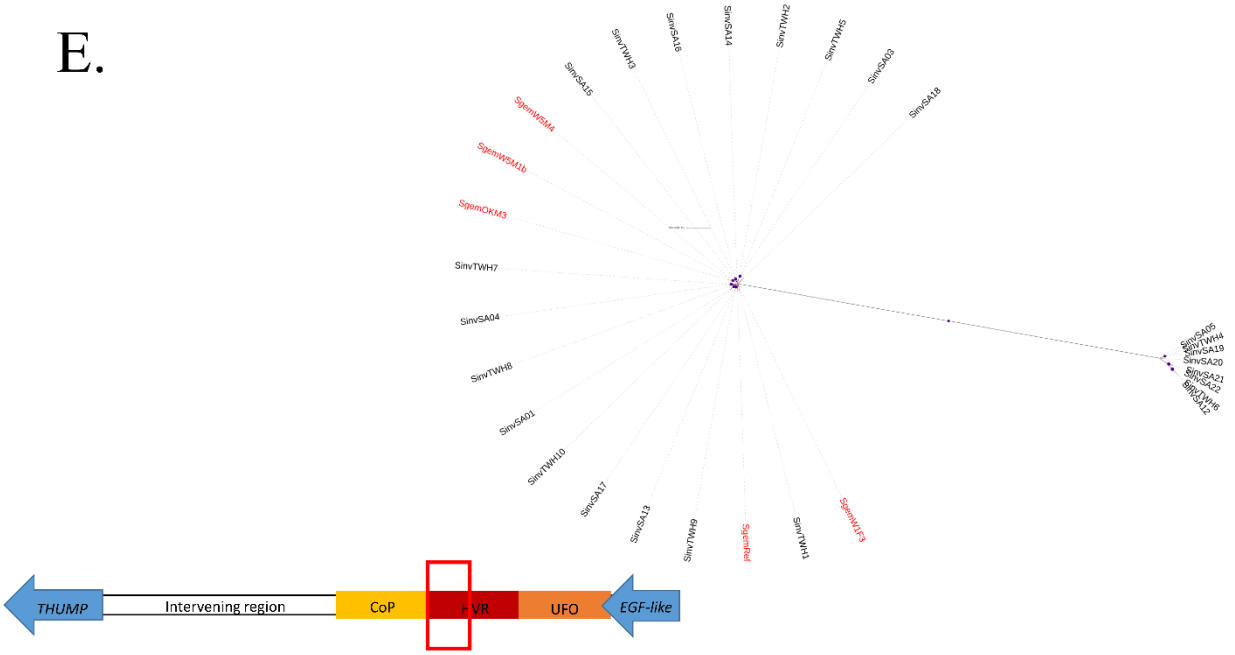


D.

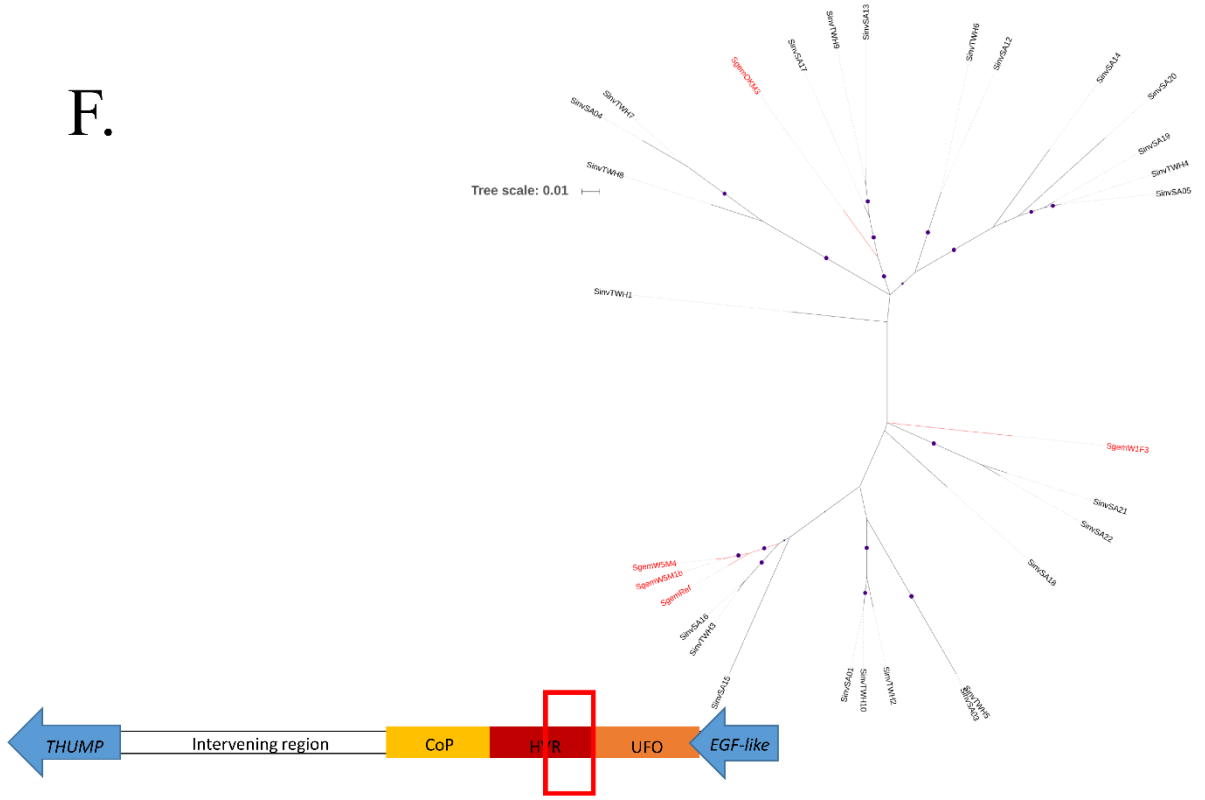


doi:10.6342/NTU201703365

E.

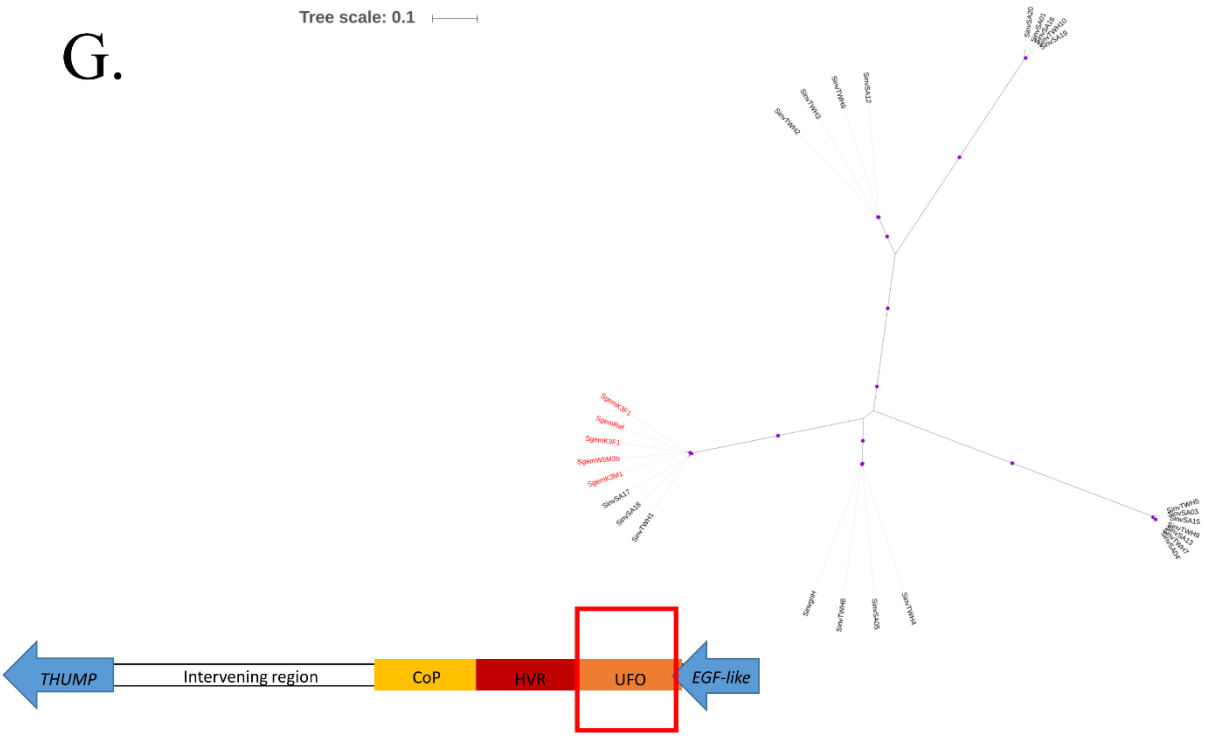


F.

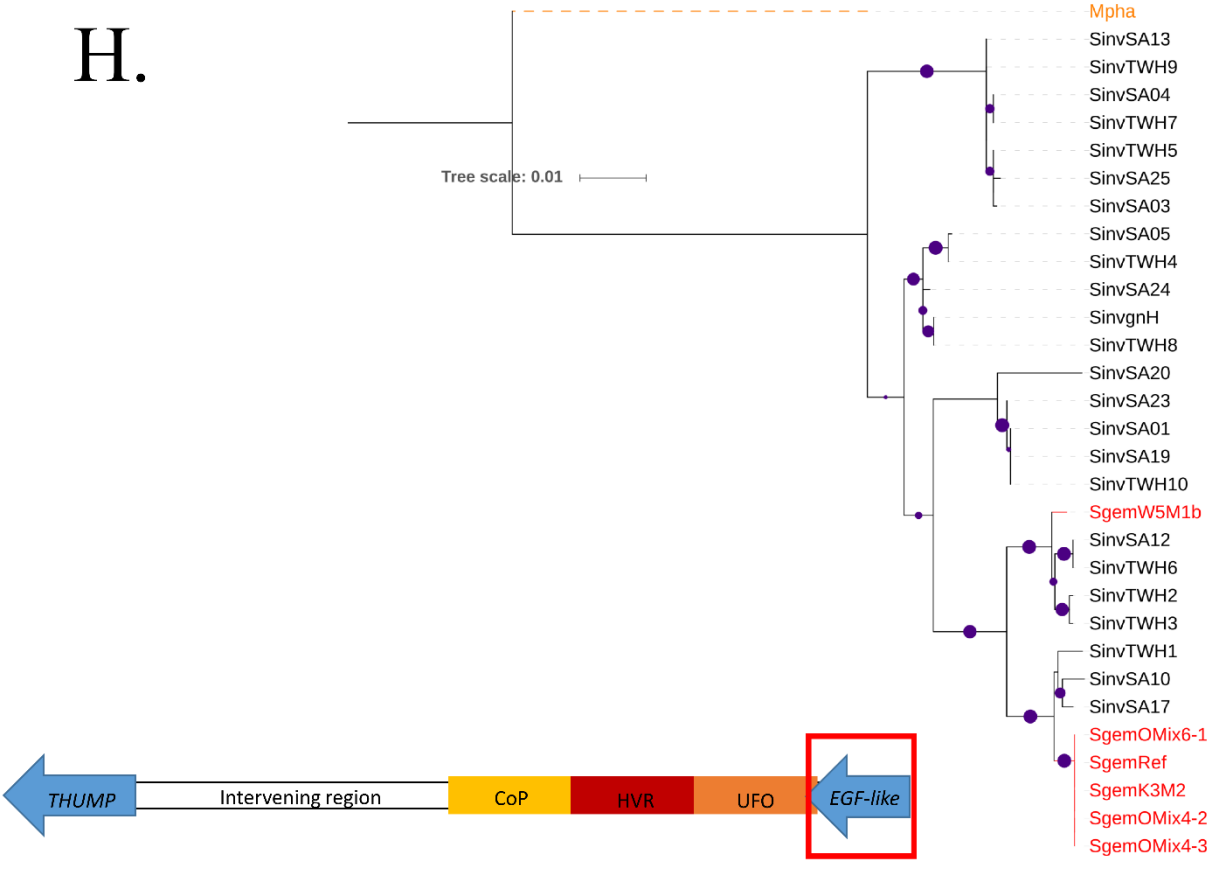


doi:10.6342/NTU201703365

G.



H.

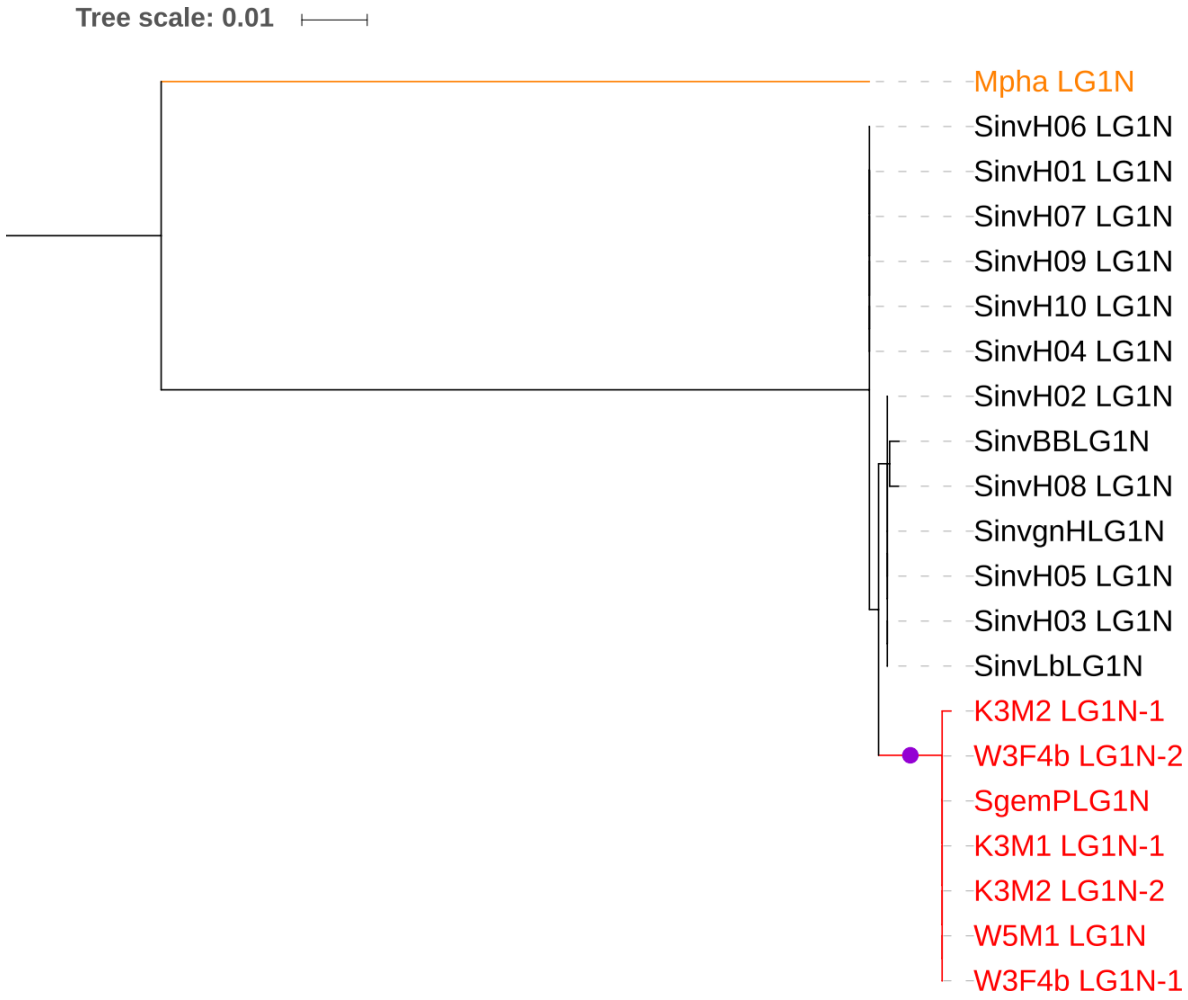


doi:10.6342/NTU201703365

Figure 6. Phylogenic trees of the neutral regions.

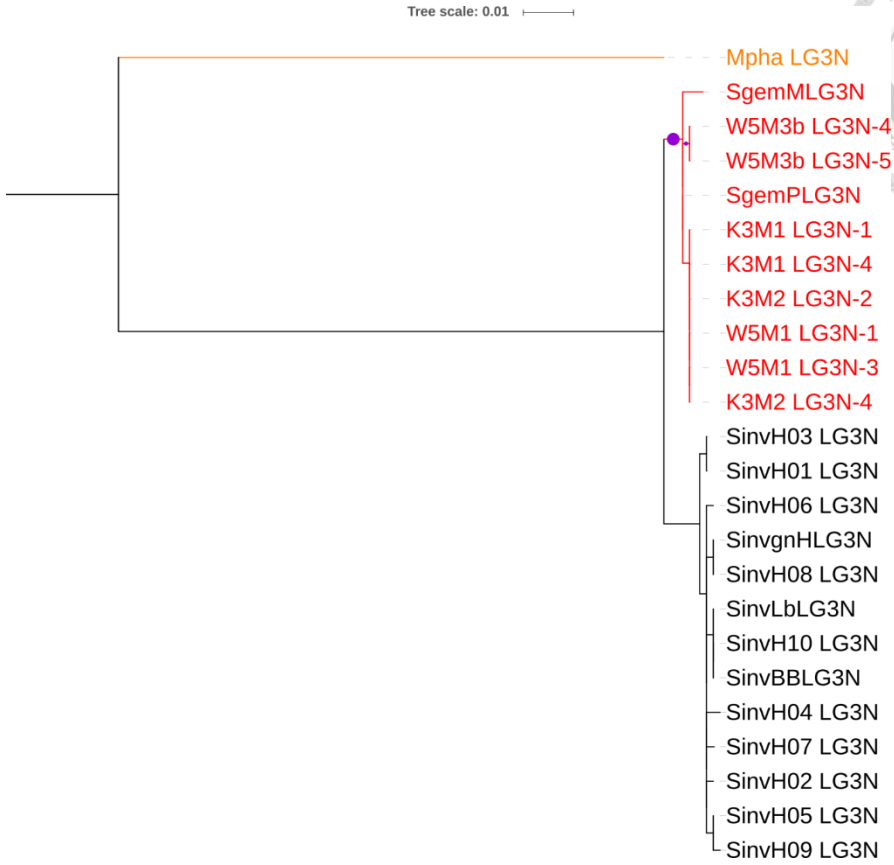
TFA samples, red; RIFA samples, black; out-group, orange. Purple dots indicate node support >70% by bootstrap replication. (A) LG1N, (B) LG3N, (C) LG5N1, (D) LG5N4, (E) LG10N.

A.

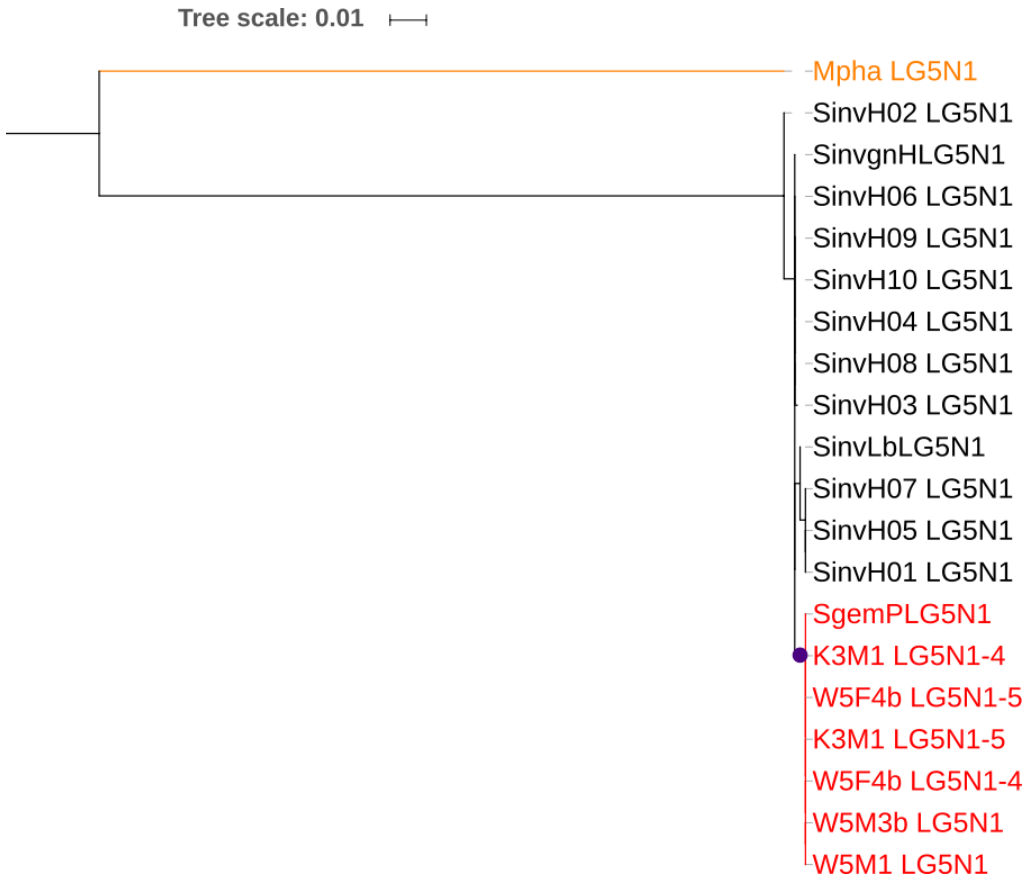


doi:10.6342/NTU201703365

B.



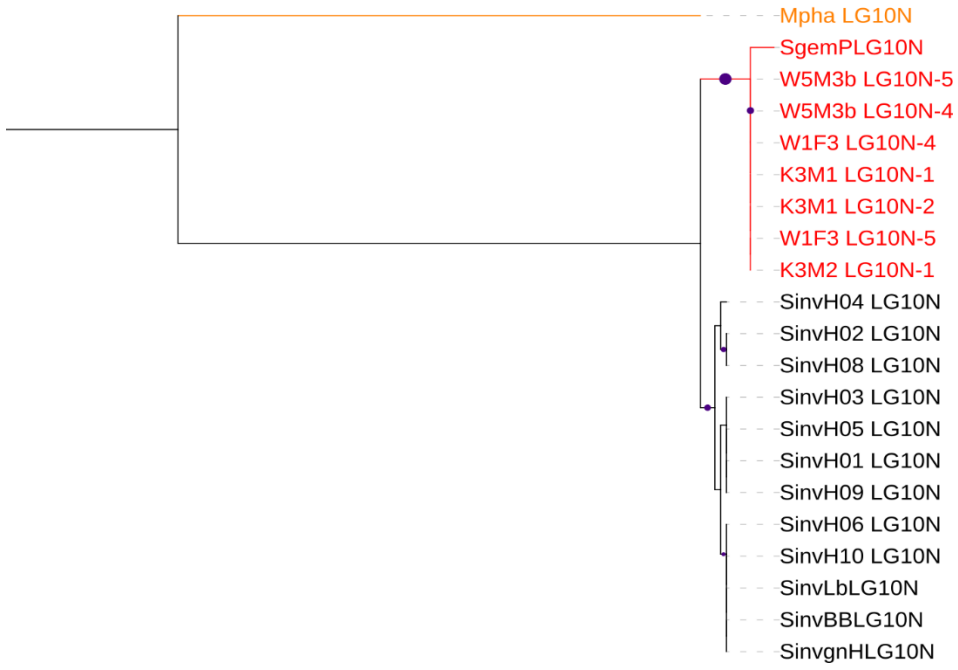
C.



doi:10.6342/NTU201703365

D.

E.



doi:10.6342/NTU201703365

Table 1. Descriptive information for the microsatellites used in this study.

Location	Locus	Allele number	Size	Core unit
LG3	SDL1	5	183~271	(CT)n
LG3	Soli125	2	196~202	(CGA)n
LG3	SDL5	3	247~253	(AGG)n
LG3	THUMP	16	310~347	(AT)n
LG3	SDL6-1	7	117~145	(GT)n
LG3	SDL6-7	2	359~361	(AT)n
LG3	SDL6-3	3	109~133	(GCG)n
LG3	SDL6-4	4	248~273	(GAG)n
LG3	SDL7	3	232~236	(CA)n
LG3	SDL11	10	376~402	(AG)n
LG10	Sol20	6	117~135	(TC)n
LG10	SdagC487	3	304~315	(CT)n
LG5	SiMS2A-65	5	137~153	(AG)n
LG1	Tramsa2	3	201~211	(AG)n

This study examined 256 diploid females and 60 haploid males (total 316 samples) from 39 colonies. The linkage group (LG) refers to the RIFA linkage map (Huang *et al.*, unpublished).

doi:10.6342/NTU201703365

Table 2. Summary of the Hardy-Weinberg analysis of diploid female microsatellite data.

	Allele number	Replicates	Proportion reject H_0, $\alpha = 0.05$	P-value
SDL1	5	1000	0	1
Soli125	2	1000	0	1
SDL5	3	1000	0	1
THUMP	16	1000	0.326	0.674
SDL6-1	7	1000	0.011	0.989
SDL6-7	2	1000	0.201	0.799
SDL6-3	3	1000	0.478	0.522
SDL6-4	4	1000	0	1
SDL7	3	1000	0	1
SDL11	10	1000	0.646	0.354
Sol20	6	1000	0.002	0.998
SdagC487	3	1000	0	1
SiMS2A-65	5	1000	0.243	0.757
Tramsa2	3	1000	0.427	0.573

256 diploid female data from 35 families were used. The analysis was tested with 1000 adjusted-bootstrap replicates, with the final P-value defined as 1 – [(number of tests rejecting H_0)/replicate number]. SDL1 to SDL11 are loci near the hypervariable region, and the four others are independent loci on different linkage groups.

doi:10.6342/NTU201703365

Table 3. Haplotype patterns for all 3, 4, or 5 adjacent loci combinations in the focal sex locus region

A

	SDL1 to SDL5	Soli125 to THUMP	SDL5 to SDL6-1	THUMP to SDL6-7	SDL6-1 to SDL6-3	SDL6-7 to SDL6-4	SDL6-3 to SDL7	SDL6-4 to SDL11
allele number	17	25	30	34	16	11	14	25
NA number	122	144	133	64	58	41	40	89
Homo. Number	38	15	13	16	37	70	58	22
Hetero. Number	74	75	88	154	139	123	136	123

B

	SDL1 to THUMP	Soli125 to SDL6-1	SDL5 to SDL6-7	THUMP to SDL6-3	SDL6-1 to SDL6-4	SDL6-7 to SDL7	SDL6-3 to SDL11
allele number	33	31	30	44	25	15	37
NA number	111	109	98	54	44	20	72
Homo. Number	12	12	13	5	22	49	12
Hetero. Number	76	78	88	140	133	130	115

C

	SDL1 to SDL6-1	Soli125 to SDL6-7	SDL5 to SDL6-3	THUMP to SDL6-4	SDL6-1 to SDL7	SDL6-7 to SDL11
allele number	36	31	40	47	30	39
NA number	96	94	83	50	34	61
Homo. Number	12	12	5	0	22	12
Hetero. Number	76	78	96	134	128	111

Haplotypes consisting of 3, 4, or 5 adjacent loci were constructed and revealed only one region, THUMP to SDL6-4, with no homozygous female individuals. NA, excluded data (due to the unphased data); allele number means total haplotype number; homo. number means total homozygote count; hetero. number means total heterozygote count. Bold boxed regions include the hypervariable region. (**A)** 3-loci haplotypes. (**B)** 4-loci haplotypes. (**C)** 5-loci haplotypes. The microsatellite marker order is: SDL1, Soli125, SDL5, THUMP, SDL6-1, SDL6-7, SDL6-4, SDL6-3, SDL7, and SDL11. The yellow block is the only no homozygote case (p-value $\cong$ 0.12).

doi:10.6342/NTU201703365

Table 4. The average nucleotide divergence (per site) within and between RIFA and TFA

RIFA						TFA				
fragment	N	site(bp)	π	Tajima's D	Dxy	fragment	N	site(bp)	π	Tajima's D
CoP1	22	3170	0.02894	0.40128	0.02986	CoP1	6	3170	0.01289	-1.52801
CoP2	22	4898	0.06501	0.30162	0.06761	CoP2	6	4898	0.03001	-1.48022
HVR1	25	2331	0.15871	1.94675	0.1239	HVR1	5	2331	0.01404	0.01218
HVR1a1	8	2331	0.03422	0.69474	0.32416					
HVR1a2	17	2331	0.03563	-0.71006	0.0402					
HVR2	25	4452	0.07428	0.19774	0.07624	HVR2	5	4452	0.04977	-0.11025
UFO	23	5802	0.25879	2.71455	0.27822	UFO	5	5802	0.00169	-1.23588
UFOa1	3	5802	0.00751	NA	0.01273					
UFOa2	4	5802	0.00394	2.29133	0.21076					
UFOa3	7	5802	0.00656	1.05954	0.38628					
UFOa4	5	5802	0.00660	-1.20208	0.38794					
UFOa5	4	5802	0.00592	1.48292	0.27757					
EGF-like	24	1490	0.03568	0.90387	0.03781	EGF-like	6	1490	0.00811	-1.51597
EGF-likeCDS	24	1140	0.0328	0.9214	0.03448	EGF-likeCDS	6	1140	0.00677	-1.49539
LG1N	13	738	0.00154	0.2084	0.0082	LG1N	7	737	0.00064	-1.00623
LG3N	13	897	0.00251	-1.29575	0.01529	LG3N	10	897	0.00229	-0.19251
LG5N1	12	960	0.00162	-0.57864	0.00494	LG5N1	7	966	0.00000	-0.30187
LG5N4	13	937	0.00000	0.00000	0.00647	LG5N4	11	939	0.00152	-0.64843
LG10N	12	911	0.00195	0.54369	0.00506	LG10N	8	910	0.00000	-1.53470

The within/between-species nucleotide divergence (π and Dxy) for each locus examined are indicated. Within species divergence (π), is the average divergence of two randomly picked sequences from one species. Between species nucleotide diversity (Dxy) is the average divergence of two randomly picked sequences selected respectively from these two species. Tajima's D is a summary statistic for the state of the evolution force. For example, positive value could indicate balancing selection or population contraction and negative value could indicate purify selection, balancing selection, bottleneck or others. Values in red are significant ($p < 0.05$). Fragment, the locus name; N, the number of sequences; site (bp), the length of locus after alignment.

doi:10.6342/NTU201703365

5. References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Andersson L, Sigurdardóttir S, Borsch C, Gustafsson K. 1991. Evolution of MHC polymorphism: Extensive sharing of polymorphic sequence motifs between human and bovine DRB alleles. Immunogenetics 33:188-193.

Ascunce MS, Yang C-C, Oakey J, Calcaterra L, Wu W-J, Shih C-J, Goudet J, Ross KG, Shoemaker D. 2011. Global Invasion History of the Fire Ant Solenopsis invicta. Science 331:1066-1068.

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R, Perrin N, Ross L, Valenzuela N, Vamosi JC, The Tree of Sex C. 2014. Sex Determination: Why So Many Ways of Doing It? PLoS Biol. 12:e1001899.

Bell LR, Maine EM, Schedl P, Cline TW. 1988. Sex-lethal, a Drosophila sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. Cell 55:1037-1046.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Cheetham RK, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IMJ, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DMD, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Catenazzi MCE, Chang S, Cooley RN, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fajardo KVF, Furey WS, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Jones TAH, Kang G-D, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ng BL, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos

doi:10.6342/NTU201703365

LL, Pickering L, Pike AC, Pike AC, Pinkard DC, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Rodriguez AC, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Sohna JES, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, vandeVondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ. 2008. Accurate Whole Human Genome Sequencing using Reversible Terminator Chemistry. Nature 456:53-59.

Beukeboom L, Van De Zande L. 2010. Genetics of sex determination in the haplodiploid wasp Nasonia vitripennis (Hymenoptera: Chalcidoidea). J Genet 89:333-339.

Beye M, Hasselmann M, Fondrk MK, Page Jr RE, Omholt SW. 2003. The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. Cell 114:419-429.

Biewer M, Schlesinger F, Hasselmann M. 2015. The evolutionary dynamics of major regulators for sexual development among Hymenoptera species. Front. Genet. 6:124.

Brownstein MJ, Carpten JD, Smith JR. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20:1004-1006, 1008-1010.

Cho S, Huang ZY, Green DR, Smith DR, Zhang J. 2006. Evolution of the complementary sex-determination gene of honey bees: Balancing selection and trans-species polymorphisms. Genome Res. 16:1366-1375.

Cho S, Huang ZY, Zhang J. 2007. Sex-Specific Splicing of the Honeybee doublesex Gene Reveals 300 Million Years of Evolution at the Bottom of the Insect Sex-Determination Pathway. Genetics 177:1733-1741.

de Boer JG, Kuijper B, Heimpel GE, Beukeboom LW. 2012. Sex determination meltdown upon biological control introduction of the parasitoid Cotesia rubecula? Evol. Appl. 5:444-454.

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-Time DNA Sequencing from Single Polymerase Molecules. Science 323:133-138.

doi:10.6342/NTU201703365

Garrigan D, Hedrick PW, Mitton J. 2003. Perspective: Detecting Adaptive Molecular Polymorphism: Lessons from the MHC. Evolution 57:1707-1722.

Gempe T, Beye M. 2011. Function and evolution of sex determination mechanisms, genes and pathways in insects. Bioessays 33:52-60.

Gempe T, Hasselmann M, Schiøtt M, Hause G, Otte M, Beye M. 2009. Sex determination in honeybees: two separate mechanisms induce and maintain the female pathway. PLoS Biol. 7:e1000222.

Gotzek D, Axen HJ, Suarez AV, Helms Cahan S, Shoemaker D. 2015. Global invasion history of the tropical fire ant: a stowaway on the first global trade routes. Mol. Ecol. 24:374-388.

Gotzek D, Clarke J, Shoemaker D. 2010. Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). BMC Evol. Biol. 10:300.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52.

Heimpel GE, de Boer JG. 2008. Sex Determination in the Hymenoptera. Annu. Rev. Entomol. 53:209-230.

Inoue H, Fukumori Y, Hiroyoshi T. 1983. Mapping of autosomal male-determining factors of the housefly, Musca domestica L., by means of sex-reversal. The Jpn. j. of genet. 58:451-461.

Katoh K, Standley DM. 2013. MAFFT MultipleSequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30:772-780.

Kinene T, Wainaina J, Maina S, Boykin L. 2016. Rooting Trees, Method for. Encycl. Evol. Biol. 3:489-493.

Klein J. 1980. Generation of diversity at MHC loci: implications for T-cell receptor repertoires. Immunology 80:239-253.

Kondrashov AS. 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336:435-440.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol.

Lai LC, Hua KH, Yang CC, Huang RN, Wu WJ. 2009. Secretion Profiles of Venom Alkaloids in Solenopsis geminata (Hymenoptera: Formicidae) in Taiwan. Environ. Entomol. 38:879-884.

Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44:W242-W245.

Mikheyev AS, Linksvayer TA. 2015. Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. eLife 4:e04775.

Miyakawa MO, Mikheyev AS. 2015. QTL Mapping of Sex Determination Loci Supports

doi:10.6342/NTU201703365

an Ancient Pathway in Ants and Honey Bees. PLoS Genet. 11:e1005656.

Muller HJ. 1964. The relation of recombination to mutational advance. Mutat. Res-Fund. Mol. M. 1:2-9.

Nagl S, Tichy H, Mayer WE, Takahata N, Klein J. 1998. Persistence of neutral polymorphisms in Lake Victoria cichlid fish. Proc. Natl. Acad. Sci. 95:14238-14243.

Paradis E. 2010. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26:419-420.

Pitts J, McHugh Ross KG. 2005. Cladistic analysis of the fire ants of the Solenopsis saevissima species-group (Hymenoptera: Formicidae). Zool Scr. 34.

Pitts JP, McHugh JV, Ross KG. 2005. Cladistic analysis of the fire ants of the Solenopsis saevissima species-group (Hymenoptera: Formicidae). Zool Scr. 34:493-505.

R Development Core Team. 2016. Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Ross K, Fletcher DC. 1986. Diploid male production — a significant colony mortality factor in the fire ant Solenopsis invicta (Hymenoptera: Formicidae). Behav. Ecol. Sociobiol. 19:283-291.

Ross KG. 1985. Genetic origin of male diploidy in the fire ant, Solenopsis invicta (Hymenoptera : Formicidae), and its evolutionary significance. Evolution 39:888-903.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19.

Sadd BM, Barribeau SM, Bloch G, de Graaf DC, Dearden P, Elsik CG, Gadau J, Grimmelikhuijzen CJ, Hasselmann M, Lozier JD, Robertson HM, Smagghe G, Stolle E, Van Vaerenbergh M, Waterhouse RM, Bornberg-Bauer E, Klasberg S, Bennett AK, Camara F, Guigo R, Hoff K, Mariotti M, Munoz-Torres M, Murphy T, Santesmasses D, Amdam GV, Beckers M, Beye M, Biewer M, Bitondi MM, Blaxter ML, Bourke AF, Brown MJ, Buechel SD, Cameron R, Cappelle K, Carolan JC, Christiaens O, Ciborowski KL, Clarke DF, Colgan TJ, Collins DH, Cridge AG, Dalmay T, Dreier S, du Plessis L, Duncan E, Erler S, Evans J, Falcon T, Flores K, Freitas FC, Fuchikawa T, Gempe T, Hartfelder K, Hauser F, Helbing S, Humann FC, Irvine F, Jermiin LS, Johnson CE, Johnson RM, Jones AK, Kadowaki T, Kidner JH, Koch V, Kohler A, Kraus FB, Lattorff HM, Leask M, Lockett GA, Mallon EB, Antonio DS, Marxer M, Meeus I, Moritz RF, Nair A, Napflin K, Nissen I, Niu J, Nunes FM, Oakeshott JG, Osborne A, Otte M, Pinheiro DG, Rossie N, Rueppell O, Santos CG, Schmid-Hempel R, Schmitt BD, Schulte C, Simoes ZL, Soares MP, Swevers L, Winnebeck EC, Wolschin F, Yu N,

doi:10.6342/NTU201703365

Zdobnov EM, Aqrawi PK, Blankenburg KP, Coyle M, Francisco L, Hernandez AG, Holder M, Hudson ME, Jackson L, Jayaseelan J, Joshi V, Kovar C, Lee SL, Mata R, Mathew T, Newsham IF, Ngo R, Okwuonu G, Pham C, Pu LL, Saada N, Santibanez J, Simmons D, Thornton R, Venkat A, Walden KK, Wu YQ, Debyser G, Devreese B, Asher C, Blommaert J, Chipman AD, Chittka L, Fouks B, Liu J, O'Neill MP, Sumner S, Puiu D, Qu J, Salzberg SL, Scherer SE, Muzny DM, Richards S, Robinson GE, Gibbs RA, Schmid-Hempel P, Worley KC. 2015. The genomes of two key bumblebee species with primitive eusocial organization. Genome biol. 16:76.

Sharma A, Heinze SD, Wu Y, Kohlbrenner T, Morilla I, Brunner C, Wimmer EA, van de Zande L, Robinson MD, Beukeboom LW, Bopp D. 2017. Male sex in houseflies is determined by *Mdmd*, a paralog of the generic splice factor gene *CWC22*. Science 356:642-645.

Shukla JN, Palli SR. 2012. Sex determination in beetles: Production of all male progeny by Parental RNAi knockdown of transformer. Sci. Rep. 2.

Snell GD. 1935. The Determination of Sex in Habrobracon. Proc. Natl. Acad. Sci. U. S. A. 21:446-453.

Takahata N. 1990. A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc. Natl. Acad. Sci. USA 87:2419.

Takahata N, Nei M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124:967-978.

Vekemans X, Slatkin M. 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics 137:1157-1165.

Verhulst EC, Beukeboom LW, van de Zande L. 2010. Maternal Control of Haplodiploid Sex Determination in the Wasp Nasonia. Science 328:620-623.

Wegner KM, Eizaguirre C. 2012. New(t)s and views from hybridizing MHC genes: introgression rather than trans-species polymorphism may shape allelic repertoires. Mol. Ecol. 21:779-781.

Wurm Y, Wang J, Riba-Grognuz O, Corona M, Nygaard S, Hunt BG. 2011. The genome of the fire ant Solenopsis invicta. Proc. Natl. Acad. Sci. U. S. A. 108.

doi:10.6342/NTU201703365

Appendix 1. TFA colony information

Colony ID	Location	Note
台東	台東	
梧棲	台中梧棲漁港	
O	台中烏日	with queen
G	台中烏日	
CK1	台南成功大學	
CK2	台南成功大學	
CK3	台南成功大學	
CK4	台南成功大學	
WL	台南林默娘公園	workers only
花	花蓮	
頂	屏東崁頂	males only
中 1	台中烏日	workers only
中 2	台中烏日	workers only
中 3	台中烏日	workers only
中 4	台中烏日	workers only
中 5	台中烏日	workers only
中 6	台中烏日	workers only
南 1	雲林斗南田徑場	workers only
南 2	雲林斗南田徑場	workers only
南 3	雲林斗南田徑場	workers only
南 4	雲林斗南田徑場	workers only
梧 11	台中梧棲漁港	workers only
梧 12	台中梧棲漁港	workers only
梧 13	台中梧棲漁港	workers only
U1	台中梧棲漁港	
U2	台中梧棲漁港	
U3	台中梧棲漁港	
W1	台中梧棲漁港	
W2	台中梧棲漁港	
W3	台中梧棲漁港	
W4	台中梧棲漁港	
W5	台中梧棲漁港	
K1	台中梧棲漁港旁公園	with queen
K2	台中梧棲漁港旁公園	
K3	台中梧棲港區公園	
P	台中梧棲台 61 線西濱道路旁	

doi:10.6342/NTU201703365

Appendix 2. Primer information

The PIG-tail sequence is underlined at the 5' end of the primer. The PET, NED, VIC and FAM are the fluorescent markers used in microsatellite analysis.

ID	Sequence	The part of this study	Note
SDL6_F	PET-CGCCGATTTGCTAATCGCGGG	Microsatellite analysis	
SDL6_R	GCTTCTTTCGAACGATCGAGACATTTC	Microsatellite analysis	mutation at 3' terminal
SdagC147_F	GAATTTCGTCGCAGCAGGG	Microsatellite analysis	
SdagC147_R	ACGACCGTCCTCTCGCTTTG	Microsatellite analysis	mutation at 3' terminal
Sinv18_F	VIC-CCAATATCGACGCCTATGGAG	Microsatellite analysis	
Sinv18_R	GTGTGTAATGCCCTGCCCTCT	Microsatellite analysis	No blast hit in TFA genome draft
SdagC485_F	NED-ATAGCGGGAATTGCAGGTCA	Microsatellite analysis	No blast hit in TFA genome draft
SdagC485_R	TCGGAAGGACTGAAGGAGTG	Microsatellite analysis	No blast hit in TFA genome draft
SDL1_F	PET-CGCGTTATATCGCAAACAAGCG	Microsatellite analysis	in LG3, SDL region
SDL1_R	GCTTCTTTCTCTCCCGTGAGATATCTTC	Microsatellite analysis	in LG3, SDL region
Soli125_F	FAM-GAAGCAGGCGTGAAGACAGAC	Microsatellite analysis	in LG3, SDL region
Soli125_R	GCTTCTTGTCTCAACATTGTCACATGTCCT	Microsatellite analysis	in LG3, SDL region
SDL5_F	VIC-AGTGAAAAGTGGACGAGGCTG	Microsatellite analysis	in LG3, SDL region
SDL5_R	CTAAGAAAGCCAGCGGAAAGC	Microsatellite analysis	in LG3, SDL region
THUMP_F	PET-CTAACAACTCACTTACGTGTGACGC	Microsatellite analysis	in LG3, SDL region
THUMP_R	GCTTCTATGAGTCTGGATAAGCTAGTGTTG	Microsatellite analysis	in LG3, SDL region
SDL6-1_F	NED-TGCAACGGGTTGGCGACGATCTC	Microsatellite analysis	in LG3, SDL region
SDL6-1_R	GCTTCTTGATTTATTTTATTAATTAGGCG	Microsatellite analysis	in LG3, SDL region
SDL6-7_F	NED-GCTTCCCTGTTTCGCTAACATTGTTCCGCG	Microsatellite analysis	in LG3, SDL region
SDL6-7_R	AACACCTTTGAAAATATTTGCACGG	Microsatellite analysis	in LG3, SDL region
SDL6-3_F	FAM-GACGCAGGGCGTTGAACCGTCC	Microsatellite analysis	in LG3, SDL region
SDL6-3_R	GCTTCTAAAATTCATCTACGAGCGAGAG	Microsatellite analysis	in LG3, SDL region
SDL6-4_F	FAM-CATCCGACCCGCTTCGGGTCTC	Microsatellite analysis	in LG3, SDL region
SDL6-4_R	TTATAAGATATTATGTTCGCCG	Microsatellite analysis	in LG3, SDL region
SDL7_F	NED-GTAGGTTTATCCAACGCGTTC	Microsatellite analysis	in LG3, SDL region

doi:10.6342/NTU201703365

SDL7_R	GCTTCTCTACGATTGAATTCAACGTTG	Microsatellite analysis	in LG3, SDL region
SDL11_F	VIC-GTCGATCGAGTCCAAGTACAC	Microsatellite analysis	in LG3, SDL region
SDL11_R	GCTTCTTTGCCGTTAACAAAAGTCCAC	Microsatellite analysis	in LG3, SDL region
Tramsa2_F	GCTTCTATTATCTGTTGATTGTATGAGTG	Microsatellite analysis	in LG1, *Tra* gene region
Tramea2_R	VIC-ATACTCAACTTATCGTGAGATAC	Microsatellite analysis	in LG1, *Tra* gene region
Sol20_F	FAM-TCGAAACGCTCCCTCTGT	Microsatellite analysis	in LG10, as contral
Sol20_R	AGCATGAAAAATCGGGAGC	Microsatellite analysis	in LG10, as contral
SdagC487_F	FAM-CCACGAGATGGAAGATGAGTTTACG	Microsatellite analysis	in LG10, as contral
SdagC487_R	GCTTCTCCTTTTGTCTGGCGAGGGAC	Microsatellite analysis	in LG10, as contral
SiMS2A-65_F	NED-ATTTTATCGAACGGGAGGAAAAAG	Microsatellite analysis	in LG5, as contral
SiMS2A-65_R	TGCTTTCAAATTAACCTTGCGAAT	Microsatellite analysis	in LG5, as contral
EGFlike_FN	ACTGTATGTTAAATAAGAGGTTATGTCTG	Cloning sequence	
EGFlike_R	TCCACGTTTGCCGTCGTATCTTG	Cloning sequence	
EGFlike_FNE	TTCCACAAGAATAGAACAAATACTTCTTTC	Cloning sequence	
N4R_F	GAACAATTCCTCCCTTTAAAACTTGAG	HVR positive contral	
N4R_RN	TGCGAAAGAGCGCTGTGATACGCGTTCACC	HVR positive contral	
preCoP_F	GTTATTATTATTTAAAACAACCCTTCGAGC	Cloning sequence	
CoP_R1	GTATATTTCGTCACATTAGAAATACCTGAAG	Cloning sequence	
CoP_R2	TTCTTTGCTTTGTCACTAGATCGTCAAACGCG	Cloning sequence	
HVRp_F1	TCTTCAGGTATTTCTAATGTGACAAAATATAC	Cloning sequence	
HVRp_F2	TTTACGCGTTTGACGATCTAGTGACAAAGC	Cloning sequence	
HVRN_R	TCGTGTCCACGTCGCGACCTTTCGAATCGGC	Cloning sequence	
UFOE_R	GCAAACGTGGATATTTTAGGACCAG	Cloning sequence	
UFON_F	AGCCGATTCGAAAGGTCGCGACGTGGACAC	Cloning sequence	

doi:10.6342/NTU201703365

Appendix 3. The six microsatellite PCR groups

Group	Locus	Size	Core unit	Labeled-compound
T	THUMP	310~347	(AT)n	PET
I	Tramsa2	201~211	(AG)n	VIC
	SDL6-7	359~361	(AT)n	NED
II	SDL6-1	117~145	(GT)n	NED
	SDL1	183~271	(CT)n	PET
	SDL6-4	248~273	(GAG)n	FAM
III	Soli125	196~202	(CGA)n	FAM
	SDL5	247~253	(AGG)n	VIC
	SDL11	376~402	(AG)n	VIC
IV	SDL7	232~236	(CA)n	NED
	SDL6-3	109~133	(GCG)n	FAM
V	Sol-20	117~135	(TC)n	FAM
	SdagC487	304~315	(CT)n	FAM
	SiMS2A-65	137~153	(AG)n	NED

doi:10.6342/NTU201703365

Appendix 4. All microsatellite information

1. Haploid male microsatellite genotypes

Locus	SDL1	Soli125	SDL5	THUMP	SDL6-1	SDL6-7	SDL6-3	SDL6-4	SDL7	SDL11
ID										
K3M1	271	202	247	320	145	359	112	265	234	380
K3M2	271	202	247	320	145	359	112	265	234	380
K3M3	271	196	247	320	145	359	109	265	234	380
K3M4	271	202		320	145	359	112	265	234	380
K3M5	271	202	247	320	145	359	112	265	234	380
W5M1	271	196	247	320	145	359		268	234	380
W5M2	271	196	247	324	119	359		273	234	380
W5M3	271	202	247	320	145	359		268	234	380
W5M4	271	202	247	320	145	359		268	234	380
W5M5	271	202	247	320	145	359		268	234	380
W5M6		196	247	324	119	359	136	273	234	380
W5M1b	271	202	247	320	145	359	136	268	234	380
W5M2b	271	202	247	320	145	359		268	234	380
W5M3b		196	247	324	119	359		273	234	380
W5M4b	271	202	247	320	145	359		268	234	380
CKM1-1	244	202	247	328	130	359	112	265	234	398
CKM1-2	244	202	247	328	130	359	112	265	234	398
CKM1-3	252	202	247	314	143	359	112	273	234	376
CKM1-4	252	202	247	314	143	359	112	273	234	376
CKM1-5	244	202	247	328	130	359	112	265	234	398
CKML41-6	253	196	247	314	143	359	112	268	234	376
CKMprep1-7	253	196	247	314	143	359	112	268	234	376
CKM3-1	253	196	247	314	143	359	112	265	234	398
CKM3-2	253	196	247	314	143	359	112	265	234	398
CKM3-3	252	202	247		119	359	112	265	232	392
CKM3-4	252	202	247	314	119	359	112	265	232	392
CKM3-5	252	202	247	314	119	359	112	265	232	392
烏日M1-1	244	202	247		119		112	265	234	400
烏日M1-2	252	202	247		119		112	265	234	400

doi:10.6342/NTU201703365

烏日 M2-1	271	196	250	310	143	359	109	265	236	398
烏日 M2-2	252	202	247	345	119	359	109	265	236	398
烏日 M2-3	252	202	247	345	119	359	109	265	236	398
烏日 M3-1	271	196	247	343	143	359	109	265	234	380
烏日 M3-2	252	202	247	320	117	359	112	265	234	380
烏日 M3-3	252	202	247	320	117	359	112	265	234	380
頂 1	244	202	247	310	143	359	112	268	234	398
頂 2	244	202	247	310	143	359	112	268	234	398
頂 3	244	202	247	310	143	359	112	268	234	398
頂 4	244	202	247	322	138	359	109	268	236	390
頂 5	244	202	247	310	143	359	112	268	234	398
頂 6	244	202	247	322	138	359	109	268	236	390
頂 7	244	202	247	322	138	359	109	268	236	390
花 m-1	244	196	247	347	123	361	112	265	234	376
花 m-2	244	202	247	320	117	359	112	268	234	398
梧棲 M 01	244	202	247	341	119	359	112	265	234	400
梧棲 M02	244	202	247	341	119	359	112	265	234	400
梧棲 M03	253	202	247	347	119	361	112	265	234	400
梧棲 M04	244	202	247	341	119	359	112	265	234	400
梧棲 M05	244	202	247	341	119	359	112	265	234	400
梧棲 M06	244	202	247	341	119	359	112	265	234	400
梧棲 M07	253	202	247	347	119	361	112	265	234	400
梧棲 M08	244	202	247	341	119	359	112	265	234	400

doi:10.6342/NTU201703365

梧棲 M09	244	202	247	341	119	359	112	265	234	400
梧棲 M10	244	202	247	341	119	359	112	265	234	400
梧棲 M11	253	202	247	347	119	361	112	265	234	400
梧棲 M12	253	202	247	347	119	361	112	265	234	400
梧棲 M13	244	202	247	341	119	359	112	265	234	400
梧棲 M14	253	202	247	347	119	361	112	265	234	400
梧棲 M15	244	202	247	341	119	359	112	265	234	400
梧棲 M16	244	202	247	341	119	359	112	265	234	400

doi:10.6342/NTU201703365

2. Diploid female microsatellite genotypes

Strange data (e.g. >3 alleles) are labeled in red. Manually modified values are in blue and could contain artifacts. Phased datas are denoted with an "|". Unphased datas are denoted with an "/".

Locus	SDL1	Soli125	SDL5	THUMP	SDL6-1	SDL6-7	SDL6-3	SDL6-4	SDL7	SDL11
ID										
O1	244|271	196|196	247|247	339|320	117|117	359|359	112|112	265/268	234|234	380|380
O2	244|271	196|196	247|247	339|339	117|143	359|359	112|109	265/268	234|234	380|380
O3	244|252	196|202	247|247	339|320	117|117	359|359	112|112	265/268	234|234	380|380
O4	244|252	196|202	247|247	339|320	117|117	359|359	112|112	265/268	234|234	380|380
O5	244|271	196|196	247|247	339|339	117|143	359|359	112|109	265/268	234|234	380|380
O6	244|271	196|196	247|253	339|339	117|143	359|359	112|109	265/268	234|234	380|380
O7	244|271	196|196	247|247	339|339	117|143	359|359	112|109	265/268	234|234	380|380
O8	244|271	196|196	247|253	339|339	117|143	359|359	112|109	265/268	234|234	380|380
O9	244|252	196|202	247|253	339|320	117|117	359|359	112|112	265/268	234|234	380|380
G1	252|252	202|202	247|247	310|310	143|143	359|359	109|112	265|265	234|234	380/392
G2	252|244	202|202	247|247	310|322	143|138	359|359	109|109	265|265	234|234	380/392
G3	252|244	202|202	247|247	310|322	143|138	359|359	109|109	265|265	234|234	380/392
G4	252|244	202|202	247|247	310|322	143|138	359|359	109|109	265|265	234|234	380/392
G5	252|252	202|202	247|247	310|310	143|143	359|359	109|112	265|265	234|234	380/392
WL1	252|252	202|202	247|247	318|310	117|143	359|359	109|109	265|265	234/236	392|400
WL2	252|244	202|196	247|247	318|318	117|117	359|359	109|109	265|268	234/236	392|398
WL3	252|252	202|202	247|247	318|310	117|143	359|359	109|109	265|265	234/236	392|400
WL4	252|244	202|202	247|247	318|310	117|143	359|359	109|109	265|265	234/236	392|400
WL5	252|252	202|202	247|247	318|310	117|143	359|359	109|109	265|265	234/236	392|400
南 1-1	244|252	202|202	247|247	320|310	117|143	359|359	112|109	265|265	234|236	390|398
南 1-2	244|244	202|202	247|247	320|343	117|123	359|361	112|112	265|265	234|236	390|398
南 1-3	244|244	202|202	247|247	320|343	117|123	359|361	112|112	265|265	234|234	390|398
南 1-4	244|244	202|202	247|247	320|343	117|123	359|361	112|112	265|265	234|234	390|398
南 1-5	244|252	202|202	247|247	320|310	117|143	359|359	112|109	265|265	234|236	390|398
南 2-1	253|244	202|196	247|247	318|324	145|145	359|359	109|109	265|265	234|234	396|396
南 2-2	253|253	202|202	247|247	318|318	145|145	359|359	109|112	265|265	234|234	396|396
南 2-3	253|253	202|202	247|247	318|318	145|145	359|359	109|112	265|265	234|234	396|396
南 2-4	253|244	202|196	247|247	318|324	145|145	359|359	109|109	265|265	234|234	396|396
南 2-5	253|244	202|196	247|247	318|324	145|145	359|359	109|109	265|265	234|234	396|396
南 2-6	253|244	202|196	247|247	318|324	145|145	359|359	109|109	265|265	234|234	396|396

doi:10.6342/NTU201703365

南 2-7	253|253	202|202	247|247	318|318	145|145	359|359	109|112	265|265	234|234	396|396
南 2-8	253|253	202|202	247|247	318|318	145|145	359|359	109|112	265|265	234|234	396|396
南 2-9	253|244	202|196	247|247	318|324	145|145	359|359	109|109	265|265	234|234	396|396
南 3-1	244|271	202|196	247|250	324|341	145|130	359|359	109|109	265|265	234|234	380/392
南 3-2	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 3-3	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 3-4	244|271	202|196	247|250	324|341	145|130	359|359	109|112	265|265	234|234	380/392
南 3-5	244|244	202|202	247|247	324|316	145|117	359|359	109|109	265|265	234|234	380/392
南 3-6	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 3-7	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 3-8	244|271	202|196	247|250	324|341	145|130	359|359	109|109	265|265	234|234	380/392
南 3-9	244|271	202|196	247|250	324|341	145|130	359|359	109|109	265|265	234|234	380/392
南 3-10	244|271	202|196	247|250	324|341	145|130	359|359	109|109	265|265	234|234	380/392
南 3-11	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 3-12	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 3-13	244|271	202|196	247|250	324|341	145|130	359|359	109|109	265|265	234|234	380/392
南 3-14	244|244	202|202	247|247	324|316	145|117	359|359	109|112	265|265	234|234	380/392
南 4-1	271|244	196/202	247/250	341|324	130|145	359|359	109|109	265|265	234|234	390|398
南 4-2	271|253	196/202	247/250	341|310	130|143	359|359	109|109	265|265	234|236	390|398
南 4-3	271|244	196/202	247/250	341|324	130|145	359|359	109|109	265|265	234|234	390|398
南 4-4	271|253	196/202	247/250	341|310	130|143	359|359	109|109	265|265	234|236	390|398
南 4-5	271|244	196/202	247/250	341|324	130|145	359|359	109|109	265|265	234|234	390|398
中 1-1	252|253	202|196	247|247	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-2	252|253	202|196	247|247	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-3	252|253	202|196	247|247	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-4	252|252	202|202	247|247	345|318	143|117	359|359	109|112	265|265	234|234	392/400
中 1-5	252|253	202|196	247|247	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-6	252|253	202|196	247|253	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-7	252|252	202|202	247|253	345|318	143|143	359|359	109|109	265|265	234|234	392/400
中 1-8	252|253	202|196	247|253	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-9	252|253	202|196	247|253	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 1-10	252|253	202|196	247|253	345|312	143|143	359|359	109|109	265|265	234|234	392/400
中 2-1	253|271	196|196	247|247	312|322	143|143	359|359	109|109	265|265	234|234	396|392
中 2-2	253|271	196|196	247|247	312|322	143|143	359|359	109|109	265|265	234|234	396|392
中 2-3	253|253	196|196	247|247	312|312	143|143	359|359	109|109	265|273	234|232	396|390
中 2-4	253|271	196|196	247|247	312|322	143|143	359|359	109|109	265|265	234|234	396|392

doi:10.6342/NTU201703365

中 3-1	253|244	196|196	247|247	312|339	143|117	359|359	112|112	265|268	234|234	400|400
中 3-2	253|244	196|196	247|247	312|339	143|117	359|359	112|112	265|268	234|234	400|400
中 3-3	253|244	196|196	247|247	312|339	143|117	359|359	112|112	265|268	234|234	400|400
中 3-4	253|271	196|196	247|247	312|343	143|143	359|359	112|109	265|265	234|234	400|400
中 3-5	253|244	196|196	247|247	312|339	143|117	359|359	112|112	265|268	234|234	400|400
中 4-1	253|271	196|196	247|247	312/343	143|143	359|359	NA|109	265|265	234|234	400|400
中 4-2	253|244	196|196	247|247	312/339	143|117	359|359	NA|112	265|268	234|234	400|400
中 4-3	253|244	196|196	247|247	312/339	143|117	359|359	NA|112	265|268	234|234	400|400
中 4-4	253|244	196|196	247|247	312/339	143|117	359|359	NA|112	265|268	234|234	400|400
中 4-5	253|271	196|196	247|247	312/343	143|143	359|359	NA|109	265|265	234|234	400|400
中 5-1	252/271	196/202	247/250	310|345	119|130	359|359	109|109	265|265	234|234	398|398
中 5-2	252/271	196/202	247/250	310|322	119|117	359|359	109|112	265|265	234|234	398|392
中 5-3	252/271	196/202	247/250	310|345	119|130	359|359	109|109	265|265	234|234	398|398
中 5-4	252/271	196/202	247/250	310|322	119|117	359|359	109|112	265|265	234|234	398|392
中 5-5	252/271	196/202	247/250	310|322	119|117	359|359	109|112	265|265	234|234	398|392
中 5-6	252/271	196/202	247/250	310|345	119|130	359|359	109|109	265|265	234|234	398|398
中 5-7	252/271	196/202	247/250	310|322	119|117	359|359	109|112	265|265	234|234	398|392
中 5-8	252/271	196/202	247/250	310|322	119|117	359|359	109|112	265|265	234|234	398|392
中 5-9	252/271	196/202	247/250	310|322	119|117	359|359	109|112	265|265	234|234	398|392
中 5-10	252/271	196/202	247/250	310|345	119|130	359|359	109|109	265|265	234|234	398|398
中 6-1	253|253	196|196	247|247	312|312	143|143	359|359	109|109	265|273	234|232	398|392
中 6-2	253|253	196|196	247|247	312|312	143|143	359|359	109|109	265|273	234|232	398|392
中 6-3	253|271	196|196	247|247	312|322	143|145	359|359	109|109	265|265	234|234	398|392
中 6-4	253|271	196|196	247|247	312|322	143|145	359|359	109|109	265|265	234|234	398|394
中 6-5	253|253	196|196	247|247	312|312	143|143	359|359	109|109	265|273	234|232	398|392
CKW1-1	252/253	196|202	247|253	314/328	143|130	359|359	112|112	268|273	234|234	376/398
CKW1-2	252/253	196|202	247|253	314/328	143|130	359|359	112|112	268|273	234|234	376/398
CKW1-3	252/253	196|202	247|253	314/328	143|130	359|359	112|112	268|273	234|234	376/398
CKW1-4	252/253	196|202	247|253	314/328	143|130	359|359	112|112	268|273	234|234	376/398
CKW1-5	244/252	196|202	247|253	314/328	143|130	359|359	112|112	268|265	234|234	376/398
CKW1-6	244/271	196|196	247|247	339/NA	143|117	359|359	112|109	268|265	234|234	380/380
CKW1-7	244/253	196|202	247|253	314/328	143|130	359|359	112|112	268|265	234|234	376/398
CKW1-8	244/271	196|196	247|247	339/NA	143|117	359|359	112|109	268|265	234|234	380/380
CKW1-9	244/271	196|196	247|247	339/NA	143|117	359|359	112|109	268|265	234|234	380/380
CKW1-10	252/253	196|202	247|253	314/328	143|130	359|359	112|112	268|265	234|234	376/398
CKW1-11	244/271	196|196	247|247	339/NA	143|117	359|359	112|109	268|265	234|234	380/380

doi:10.6342/NTU201703365

CKW2-1	252|253	196/202	247|247	320|310	143|143	359|359	112|112	265|265	234|234	400|400
CKW2-2	252|253	196/202	247|247	320|310	143|143	359|359	112|112	265|265	234|234	400|400
CKW2-3	252|253	196/202	247|247	320|310	143|143	359|359	112|112	265|265	234|234	400|400
CKW2-4	252|253	196/202	247|247	320|339	143|119	359|359	112|112	265|265	234|232	400|398
CKW2-5	252|253	196/202	247|247	320|339	143|119	359|359	112|112	265|265	234|232	400|398
CKW3-1	252|252	202|202	247|247	314|324	143|143	359|359	112|112	265|265	234|232	398/392
CKW3-2	252|252	202|202	247|247	314|324	143|143	359|359	112|112	265|265	234|232	398/392
CKW3-3	252|253	202|196	247|253	314|324	119|119	359|359	112|109	265|265	234|234	398/398
CKW3-4	252|253	202|196	247|253	314|324	119|143	359|359	112|109	265|265	234|234	398/398
CKW3-5	252|253	202|196	247|253	314|324	119|143	359|359	112|109	265|265	234|234	398/398
CKW3-6	252|244	202|196	247|253	320|339	119|143	359|359	112|109	265|265	234|234	380/380
CKW3-7	252|252	202|202	247|247	314|324	119|143	359|359	112|112	265|265	234|232	398/398
CKW3-8	252|252	202|202	247|247	314|324	119|119	359|359	112|112	265|265	234|232	398/398
CKW3-9	252|253	202|196	247|253	314|324	119|143	359|359	112|109	265|265	234|234	398/398
CKW3-10	252|252	202|202	247|247	314|324	119|119	359|359	112|112	265|265	234|232	398/392
CKW4-1	252|252	202|202	247|247	314|324	143|119	359|359	112|112	265|265	234|232	398|392
CKW4-2	252|253	202|196	247|247	314|324	143|143	359|359	112|112	265|265	234|234	398|392
CKW4-3	252|253	202|196	247|247	314|324	143|143	359|359	112|112	265|265	234|234	398|392
CKW4-4	252|252	202|202	247|247	314|324	143|119	359|359	112|112	265|265	234|232	398|392
CKW4-5	252|252	202|202	247|247	314|324	143|119	359|359	112|112	265|265	234|232	398|392
台東 1	252|252	202|202	247|247	322|328	143|130	359|359	112|109	265|265	234|236	400|400
台東 2	252|252	202|202	247|247	322|322	143|143	359|359	112|109	265|265	234|236	400|400
台東 3	NA|NA	202|202	247|247	322|328	143|130	359|359	112|112	265|265	234|236	400|400
台東 4	252|252	202|202	247|247	322|328	143|130	359|359	112|112	265|265	234|236	400|400
台東 5	252|244	202|202	247|247	322|328	143|143	359|359	112|112	265|268	234|234	400|392
台東 6	252|252	202|202	247|247	322|328	143|130	359|359	112|109	265|265	234|236	400|400
台東 7	252|244	202|202	247|247	322|328	143|130	359|359	NA|109	265|265	234|236	400|400
台東 8	252|252	202|202	247|247	322|322	143|143	359|359	112|109	265|268	234|236	400|400
梧棲 W01	252|244	196|202	247|247	320|341	145|119	359|359	112|112	268|265	234|234	380|400
梧棲 W02	252|244	196|202	247|247	320|341	145|119	359|359	112|112	268|265	234|234	380|400
梧棲 W03	252|244	196|202	247|247	320|341	145|119	359|359	112|112	268|265	234|234	380|400
梧棲 W04	252|253	196|202	247|247	320|347	145|119	359|361	112|112	268|265	234|234	380|400
梧棲 W05	252|253	196|202	247|247	320|347	145|119	359|361	112|112	268|265	234|234	380|400
梧棲 W06	252|253	196|202	247|247	320|347	145|119	359|361	112|112	268|265	234|234	380|400
梧棲 W07	252|244	196|202	247|247	320|347	145|119	359|361	112|112	268|265	234|234	380|400
梧棲 W10	252|244	196|202	247|247	320|341	145|119	359|359	112|112	268|265	234|234	380|400

doi:10.6342/NTU201703365

梧 11-1	253|252	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
梧 11-2	253|252	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
梧 11-3	253|253	196|196	247/253	310|324	119/143	359|359	112|136	273|268	234|234	380|402
梧 11-4	253|252	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
梧 11-5	253|253	196|196	247/253	310|324	119/143	359|359	112|136	273|268	234|234	380|402
梧 12-1	252|244	202|202	247/253	326|326	130/145	359|359	109|112	265|265	234|234	398|378
梧 12-2	252|253	202|196	247/253	326|320	130/145	359|359	109|109	265|265	234|234	398|402
梧 12-3	252|253	202|196	247/253	326|320	130/145	359|359	109|109	265|265	234|234	398|402
梧 12-4	252|244	202|202	247/253	326|326	130/145	359|359	109|112	265|265	234|234	398|378
梧 12-5	252|244	202|202	247/253	326|326	130/145	359|359	109|112	265|265	234|234	398|378
梧 13-1	252/253	196/202	247/253	324|310	123|143	359/361	NA|112	265|273	234|234	400|380
梧 13-2	252/253	196/202	247/253	324|310	123|143	359/361	NA|112	265|273	234|234	400|380
梧 13-3	252/253	196/202	247/253	324|320	123|145	359/361	NA|109	265|265	234|234	400|402
梧 13-4	252/253	196/202	247/253	324|310	123|143	359/361	NA|112	265|273	234|234	400|380
梧 13-5	252/253	196/202	247/253	324|320	123|145	359/361	NA|109	265|265	234|234	400|402
花 w-1	244|244	NA|196	247|247	347/347	117/117	NA/361	112|112	265|265	234|234	376/376
花 w-2	244|253	NA|202	247|253	347/320	117/123	359/361	112|112	265|265	234|234	376/398
花 w-3	244|253	NA|202	247|253	320/320	117/117	NA/359	112|112	265|268	234|234	NA
花 w-4	244|244	NA|196	247|250	347/339	123/143	359/361	112|112	265|265	234|234	NA
花 w-5	244|253	NA/196	247|250	320/320	117/117	NA/359	112|112	265|268	234|234	398/398
U1-1	244/252	202|202	247/NA	316|320	117|145	359|359	112|112	265|268	234|234	398|378
U1-2	244/252	202|202	247/253	316|320	117|145	359|359	112|112	265|268	234|234	398|378
U1-3	244/252	202|202	247/253	316|345	117|143	359|359	112|112	265|265	234|234	398|392
U1-4	244/252	202|202	247/253	316|345	117|143	359|359	112|112	265|265	234|234	398|392
U1-5	244/252	202|202	247/253	316|320	117|145	359|359	112|112	265|268	234|234	398|378
U1-6	244/252	202|202	247/253	316|345	117|143	359|359	112|112	265|265	234|234	398|378
U1-7	244/252	202|202	247/253	316|345	117|143	359|359	112|112	265|265	234|234	398|392
U2-1	271|244	196/202	247/250	318|341	145|143	359|359	112|112	265|265	234|236	400|398
U2-2	271|244	196/202	247/250	318|341	145|143	359|359	112|112	265|265	234|236	400|398
U2-3	271|252	196/202	247/250	318|318	145|117	359|359	112|112	265|273	234|234	400|376
U2-4	271|252	196/202	247/250	318|318	145|117	359|359	112|112	265|273	234|234	400|376
U2-5	271|252	196/202	247/250	318|318	145|117	359|359	112|112	265|273	234|234	400|376
U2-6	271|244	196/202	247/250	318|341	145|143	359|359	112|112	265|265	234|236	400|398
U2-7	271|252	196/202	247/250	318|318	145|117	359|359	112|112	265|273	234|234	400|376
U3-1	252|252	196/202	247|247	318|324	145|119	359|359	112|112	265|273	234|234	380/392
U3-2	252|252	196/202	247|247	318|324	145|119	359|359	112|112	265|273	234|234	380/392

doi:10.6342/NTU201703365

U3-3	252|252	196/202	247|247	318|324	145|119	359|359	112|112	265|273	234|234	380/392
U3-4	252|271	196/202	247|247	318|318	145|117	359|359	112|112	265|265	234|234	380/392
U3-5	252|271	196/202	247|247	318|318	145|117	359|359	112|112	265|265	234|234	380/392
U3-6	252|271	196/202	247|247	318|318	145|117	359|359	112|112	265|265	234|234	380/392
W1w1	253|244	196|202	247/253	320|326	145|145	359|359	NA|109	265|265	234|234	392|378
W1w2	253|244	196|202	247/253	320|326	145|145	359|359	NA|109	265|265	234|234	392|378
W1w3	253|253	196|196	247/253	320|320	145|119	359|359	NA|112	NA|NA	NA|NA	392|380
W1w4	253|244	196|202	247/253	320|326	145|145	359|359	NA|109	265|265	234|234	392|378
W1w5	253|253	196|196	247/253	320|320	145|119	359|359	NA|112	NA|NA	234|234	392|380
W1w6	253|244	196|202	247/253	320|326	145|145	359|359	NA|109	265|265	234|234	392|378
W1F1	NA|NA	196|196	247/253	320|326	145|119	359|359	NA|112	265|273	234|234	392|380
W1F2	NA|NA	196|202	247/253	320|320	145|145	359|359	NA|109	265|265	234|234	392|378
W1F3	NA|NA	196|196	247/253	320|326	145|119	359|359	NA|112	265|273	234|234	392|380
W2w1	265|252	196|202	247/253	330|322	117/130	359|359	112|112	247/265	234/236	376/380
W2w2	265|252	196|202	247/253	330|322	117/130	359|359	112|112	247/265	234/236	376/380
W2w3	265|252	196|202	247/253	330|322	117/130	359|359	112|112	247/265	234/236	376/380
W2w4	265|252	196|202	247/253	330|322	117/130	359|359	112|112	247/265	234/236	376/380
W2w5	265|253	196|196	247/253	330|310	117/130	359|359	112|112	247/265	234/236	376/380
W2w6	265|253	196|196	247/253	330|310	117/130	359|359	112|112	247/265	234/236	376/380
W3w1	253|253	196|196	247/253	310|324	119/143	359|359	112|112	NA|NA	234|234	380|380
W3w2	253|253	196|196	247/253	310|324	NA/NA	359|359	NA|NA	NA|NA	234|234	380|380
W3w3	253|252	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
W3w4	253|253	196|196	247/253	310|324	119/143	359|359	112|112	273|268	234|234	380|380
W3w5	253|252	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
W3w6	253|252	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
W-3F1b	NA|NA	196|202	247/253	310|320	119/143	359|359	112|109	273|265	234|234	380|402
W-3F2b	NA|NA	196|202	247/NA	310|320	119/NA	359|359	112|109	273|265	234|234	380|402
W-3F3b	NA|NA	196|202	247/NA	310|320	119/143	359|359	112|109	273|265	234|234	380|402
W-3F4b	NA|NA	196|196	247/253	310|324	119/143	359|359	112|112	273|268	234|234	380|380
W-3F5b	NA|NA	196|196	247/NA	310|324	119/143	359|359	112|112	273|268	234|234	380|380
W4w1	252/253	196/202	247/253	318/318	117|117	359|359	112|112	NA	234|234	376|392
W4w2	252/253	196/202	247/253	318/318	117|145	359|359	112|112	NA	234|234	376|392
W4w3	NA	NA	NA	NA	NA	NA	112|112	NA	234|234	376|392
W4w4	NA	NA	NA	NA	NA	NA	112|112	NA	234|236	376|392
W4w5	252/253	196/202	247/253	318/318	117|145	359|359	112|112	NA	234|234	376|390
W4w6	252/253	196/202	247/253	318/318	117|145	359|359	112|112	NA	234|234	376|392

doi:10.6342/NTU201703365

W5w1	244|244	196|202	NA|NA	341|320	119|145	359|359	NA|NA	265|268	234|234	376|380
W5w2	244|253	196|196	NA|NA	341|324	119|119	359|359	NA|NA	265|273	234|234	376|380
W5w3	244|253	196|196	NA|NA	341|324	119|119	359|359	NA|NA	265|273	234|234	376|380
W5w4	244|253	196|196	NA|NA	341|324	119|119	359|359	NA|NA	265|273	234|234	376|380
W5w5	244|244	196|202	NA|NA	341|320	119|145	359|359	NA|NA	265|268	234|234	376|380
W5w6	244|244	196|202	NA|NA	341|320	119|145	359|359	NA|NA	265|268	234|234	376|380
K1w1	253|252	196|202	NA|NA	324|310	119|143	359|359	NA|NA	268|265	234|234	376|376
K1w2	253|252	196|202	NA|NA	324|310	119|143	359|359	NA|NA	268|265	234|234	376|402
K1w3	253|253	196|196	NA|NA	324|320	119|145	359|359	NA|NA	268|268	234|234	376|376
K1w4	253|252	196|202	NA|NA	324|310	119|145	359|359	NA|NA	268|268	234|234	376|376
K1w5	253|253	196|196	NA|NA	324|320	119|145	359|359	NA|NA	268|268	234|234	376|376
K1w6	253|252	196|202	NA|NA	324|310	119|143	359|359	NA|NA	268|265	234|234	376|402
K2w1	253|244	196|202	NA|NA	320|326	145|145	359|359	NA/NA	265|265	234|234	376|378
K2w2	253|271	196|196	NA|NA	320|320	145|145	359|359	109/NA	265|265	234|234	376|402
K2w3	253|244	196|202	NA|NA	320|326	145|145	359|359	109/112	265|265	234|234	376|378
K2w4	253|244	196|202	NA|NA	320|326	145|145	359|359	109/NA	265|265	234|234	376|378
K2w5	253|244	196|202	NA|NA	320|326	145|145	359|359	109/NA	265|265	234|234	376|378
K2w6	253|271	196|196	NA|NA	320|320	145|145	359|359	NA/NA	265|265	NA|NA	NA|NA
K3w1	253|NA	196|196	NA|NA	347|320	119|145	361|359	NA|NA	265|265	NA|NA	376|380
K3w2	253|244	196|202	NA|NA	347|320	119|145	361|359	112|112	265|265	234|234	376|380
K3w3	253|244	196|202	NA|NA	347|320	119|145	361|359	112|112	265|265	234|234	376|380
K3w4	253|NA	196|196	NA|NA	347|320	119|145	361|359	112|109	265|265	234|234	376|380
K3w5	253|244	196|202	NA|NA	347|320	119|145	361|359	112|112	265|265	234|234	376|380
K3w6	253|244	196|202	NA|NA	347|320	119|145	361|359	112|112	265|265	234|234	376|380
K3F1	NA|NA	196|196	NA|NA	347|320	119|145	361|359	112|109	265|265	234|234	376|380
K3F2	NA|NA	196|196	NA|247	347|320	119|145	361|359	112|109	265|265	234|234	376|380
K3F3	NA|271	196|196	253|247	347|320	119|145	361|359	112|109	265|265	234|234	376|380
K3F4	NA|271	196|196	NA|247	347|320	119|145	361|359	112|109	265|265	234|234	376|380
Pw1	252|271	202|196	NA|NA	322|343	143|143	359|359	112|112	265/273	232/234	392|398
Pw2	252|271	202|196	NA|NA	322|343	143|143	359|359	112|112	265/273	232/234	392|398
Pw3	252|252	202|202	NA|NA	322|345	143|119	359|359	112|112	265/273	232/234	392|380
Pw4	252|271	202|196	NA|NA	322|343	143|143	359|359	112|112	265/273	232/234	392|398
Pw5	252|271	202|196	NA|NA	322|343	143|143	359|359	112|112	265/273	232/234	392|398
Pw6	252|252	202|202	NA|NA	322|345	143|119	359|359	112|112	265/273	232/234	392|380

doi:10.6342/NTU201703365