

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

学士学位论文

THESIS OF BACHELOR



论文题目：Functional study on an uncharacterized gene involved in caste differentiation process in the pharaoh ant, *Monomorium pharaonis*

学生姓名：李儒岩 Li Ruyan

学生学号：516080910021

专 业：生物科学（致远荣誉计划）

Biological Sciences (Zhiyuan Honors Program)

指导教师：张国捷 Zhang Guojie 袁政 Yuan Zheng

学院(系)：生命科学技术学院 致远学院

School of Life Sciences and Biotechnology Zhiyuan College

上海交通大学

本科生毕业设计（论文）任务书

课题名称：Functional study on an uncharacterized gene involved in caste differentiation process in the pharaoh ant, *Monomorium pharaonis*

执行时间：2019 年 4 月 至 2020 年 4 月

教师姓名：张国捷 袁政 职称：教授 教授

学生姓名：李儒岩 学号：516080910021

专业名称：生物科学（致远荣誉计划）

学院(系)：生命科学技术学院 致远学院

毕业设计（论文）基本内容和要求：

真社会性（eusociality）是一种高等的社会组织形式。真社会性昆虫有蚂蚁、白蚁、胡蜂、蜜蜂等，其特点是同一巢穴内部的不同个体之间存在基于生殖能力的社会分工。对于蚁科来说，所有蚂蚁的雌蚁都拥有二元等级分化，即可以繁育后代的个体（生殖蚁）和不育或几乎不育的个体（工蚁）。早期对于蚂蚁的研究大多是从宏观生态学的层面出发，研究蚂蚁不同等级的个体与环境之间的相互作用；然而近年来随着基因组学技术和各类基因编辑手段的兴起，从遗传学层面研究蚂蚁不同社会等级之间存在的发育和行为上的差异成为了可能。

来源于同一蚁后的雌蚁蚂蚁后代个体之间具有相同的基因组和遗传背景，但是却可以在发育过程中分化形成生殖蚁和工蚁，所以某些关键基因在不同个体间的差异性表达可能是引起蚁后和工蚁在发育和行为上差异的遗传调控机制。本实验室以法老蚁，*Monomorium pharaonis* 为模式生物，已通过各个发育阶段的等级特异转录组比较筛选出了若干个在不同等级之间具有明显差异表达的候选基因。本课题重点研究其中一个在生殖蚁和工蚁的各发育阶段均有显著差异表达的，同时没有任何已知功能的基因（geneID: 13359），尝试探索其在蚂蚁个体发育及等级分化中的调控功能。

研究内容：

1. 研究该基因的起源过程，在多个不同蚂蚁物种的基因组中寻找这个基因的同源序列，结合该基因的上下游基因共线性，分析这个基因的起源时间

需要与相关研究人员合作，利用已有蚂蚁物种的基因组注释数据来分析此基因的情况。

2. 研究该基因在生殖蚁和工蚁不同发育时期的表达情况

通过微液滴数字 PCR（droplet digital PCR, ddPCR），验证该基因在蚁后和工蚁不同发育时期的表达情况与 RNA-seq 所呈现的表达趋势是否吻合。

3. 研究该基因在幼虫或蛹期的空间表达情况

首先通过 ddPCR 初步验证该基因是否在某一个特定的组织中表达，之后再通过原位杂交（in situ hybridization, ISH）进一步研究该基因的空间表达情况。

4. RNAi 研究该基因的生物学功能

通过在蚁后和工蚁的不同发育时期注射一定浓度人工合成的 siRNA 或 long dsRNA 使该基因的 mRNA 水平显著降低，使用 ddPCR 进行验证，并观察注射个体在发育或行为上是否出现差异。

要求：

1. 具有一定的分子生物学研究经验；
2. 熟悉法老蚁中不同等级个体的发育过程及其相关行为；
3. 熟悉显微注射操作；
4. 具备一定的探索精神和抗挫能力。
5. 分析数据并撰写研究报告

毕业设计（论文）进度安排：			
序号	毕业设计（论文）各阶段内容	时间安排	备注
1	掌握法老蚁的各项背景知识，探索其胚胎及幼虫时期的显微注射条件	2019.4-2019.5	探索制备显微注射所需针的程序、注射部位、注射液体积等条件
2	验证该基因在不同发育时期的表达情况，初步探索其空间表达情况	2019.6	包括设计引物、学习 RNA 提取、ddPCR 等技能
3	设计 siRNA，并体外合成 long dsRNA	2019.7	需要在大肠杆菌中完成一些实验
4	RNAi，大量注射个体，并观察其发育及行为上的变化	2019.8-2020.1	根据实际情况，上一步可能需要反复进行
5	探索原位杂交的实验条件，尝试研究该基因的空间表达情况	2020.2-2020.4	以上研究均在哥本哈根大学（UCPH）开展，此阶段研究在中科院昆明动物所（KIZ）开展

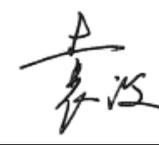
课题信息：

课题性质：设计 论文 ✓

课题来源*：国家级 ✓ 省部级 校级 横向 预研

项目编号 31671260

其他 _____

指导教师签名： _____  

2019 年 10 月 30 日

学院（系）意见：

院长（系主任）签名：_____

年 月 日

学生签名： 李儒岩

2019 年 10 月 30 日

上海交通大学

毕业设计（论文）学术诚信声明

本人郑重声明：所呈交的毕业设计（论文），是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名：李儒岩

日期：2020年5月26日

上海交通大学

毕业设计（论文）版权使用授权书

本毕业设计（论文）作者同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海交通大学可以将本毕业设计（论文）的全部或部分内 容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本毕业设计（论文）。

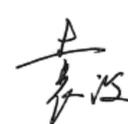
保密 ，在 5 年解密后适用本授权书。

本论文属于

不保密 。

（请在以上方框内打“√”）

作者签名：李儒岩

指导教师签名： 

日期：2020 年 5 月 26 日

日期：2020 年 5 月 26 日

法老蚁 *Monomorium pharaonis* 等级分化过程中一个功能未知基因的研究

摘要

物种从独居生活到社会性群居生活是演化上的一次重大转变。真社会性的膜翅目昆虫（蚁科、蜂科的一些种和胡蜂科的一些种）具有能够生殖的蚁后等级和不育的工蚁等级，这种生殖上的分工是这一重大转变的最重要的特征，然而其背后的遗传学原理还鲜为人知。本课题关注在法老蚁 *Monomorium pharaonis* 中一个功能未知但是可能调控蚁后和工蚁不同发育进程的基因（ID: 13359），并且从以下三个方面进行研究：通过构建系统树和正选择分析研究此基因的演化历史；通过收集不同发育时期和身体部位的样品研究此基因的时间和空间表达模式；通过 RNAi 研究此基因在个体发育中的生物学功能。研究发现 gene13359 的演化过程与蚁科中的物种分化过程基本相符；它在 poneroid ants 和 formicoid ants（蚁科中的两个进化分支）中受到不同的选择压力（ ω ratios, $\omega = d_N/d_S$ ）；尽管该基因的编码序列在蚁科中呈现高度不同，但是蛋白质都被预测在 N 端有一段信号肽并且定位在细胞外。研究发现在幼虫期至蛹期的发育过程中，gene13359 在蚁后中的表达量远高于在雄蚁和工蚁中的表达量；在蚁后的蛹期中该基因在全身均有表达，但是在腹部的表达量最高。RNAi 的结果显示 gene13359 既是在生殖蚁（蚁后和雄蚁）的化蛹过程中不可或缺的基因，同时也是控制蚁后和工蚁某些器官的可塑性发育的雌蚁特异性基因。我们认为在法老蚁中，gene13359 很可能是控制蚁后和工蚁等级分化的关键通路中的一部分。

关键词： 蚂蚁，等级分化，基因功能，法老蚁 *Monomorium pharaonis*

**Functional study on an uncharacterized gene involved in
caste differentiation process in the pharaoh ant,
*Monomorium pharaonis***

ABSTRACT

The transition from solitary organisms to social life is one of the major transitions in evolution. Eusocial hymenopterans (ants, some bees, and some wasps) best exemplify the transition for their division of labor between highly reproductive castes (queens) and functionally sterile castes (workers), the nature of which is currently unknown. We focus on an uncharacterized gene (ID: 13359) that might regulate differential development of queen and worker castes in the pharaoh ant *Monomorium pharaonis* and study its evolutionary history by constructing phylogenetic tree and performing positive selection analysis, temporal and spatial expression patterns by collecting samples from different developmental stages and body parts, and biological functions during ontogeny based on RNAi. We found that the phylogeny of gene13359 almost matched the speciation process across the Formicidae; it had differential ω ratios ($\omega = d_N/d_S$) between poneroid ants and formicoid ants; although the coding sequence exhibited highly divergence across the Formicidae, the encoded protein was always predicted to have a signal peptide at N-terminus and be localized extracellularly. We found that gene13359 demonstrated much higher expression in gyne than in male and worker from larval to pupal stages; it was expressed across the whole body in gyne pupa with the highest expression level in abdomen. The RNAi results revealed that gene13359 was both an essential gene functioning during pupation in sexual individuals (gynes and males), and a female-specific gene controlling plastic development of certain organs between gynes and workers. We conclude that gene13359 is a promising candidate involved in the key pathways controlling caste differentiation in *Monomorium pharaonis*.



Key words: ants, caste differentiation, gene functions, *Monomorium pharaonis*

Contents

General Introduction	1
Insect eusociality: a question need to be further studied	1
Ant colony as a superorganism	1
Study of social insects polyphenism: from Tinbergen's perspective	2
Summary of the recent functional genetics study in ants	3
Biology of <i>Monomorium pharaonis</i>	4
Advantages of using <i>Monomorium pharaonis</i> as a model organism	4
Colony-level regulation of <i>Monomorium pharaonis</i> caste allocation	6
Caste system and developmental scheme of <i>Monomorium pharaonis</i>	8
Aim of this study	10
Chapter 1	12
The evolutionary history of gene13359 across the Formicidae	
Chapter 2	27
The temporal and spatial expression patterns of gene13359	
Chapter 3	45
Study of the biological function of gene13359 through RNAi	
Conclusion	71
References	73
Supplementary Information	81
Acknowledgements	86

General Introduction

Insect eusociality: a question need to be further studied

Ant colony as a superorganism

The transition from solitary organisms to social life is one of the evolutionary major transitions. ([1], Smith & Szathmary, 1997). Eusocial hymenopterans, which are ants, some bees, and some wasps, best exemplify the evolutionary transition to eusociality from a solitary ancestor. The defining feature that differentiates these eusocial hymenopterans from other solitary insects is reproductive division of labor between highly reproductive castes (queens) and functionally sterile castes that perform all the non-reproductive tasks necessary to maintain the colony (workers) ([2], Hölldobler & Wilson, 1990). Wheeler recognized more than a century ago that in ant colonies, caste differentiation into reproductive and non-reproductive castes was analogous to multicellular organism cell differentiation into a sequestered germline and diverse somatic cell lines, therefore, the ant colony was considered as a superorganism ([3], Wheeler, 1911). Since eusocial hymenopterans are derived from subsocial wasp-like ancestors that alternate between reproductive and brood care phases ([4], Hunt, 2012), the evolution of eusociality involves a decoupling of these phases into reproductive queen caste and non-reproductive worker caste, which is responsible for brood care, the process of which is analogous to the differentiation of totipotent stem cells in a multicellular organism. The mechanisms under natural selection during the evolution of such decoupling process are indispensable for us to understand the evolutionary transition to eusociality.

Study of social insects polyphenism: from Tinbergen's perspective

It is notable that although the queen caste and the worker caste demonstrate dramatic differences on their morphologies, behaviors, physiologies and social roles in the colony ([5], Feldmeyer et al., 2014), they are alternative phenotypes developed from the same genotype as they are the decedents of the same parents. In other words, different environmental cues cause the organism to develop along separate pathways, resulting in two distinct phenotypes (castes). Thus, insect caste is a polyphenic trait.

Tinbergen (1963) proposed that a trait should be described from four aspects: mechanistic explanations (what the molecular mechanisms of the trait are and how they are regulated by environmental cues), developmental explanations (what effects of the trait are on the ontogeny of an individual across different life stages), phylogenetic explanations (how the trait evolves over the evolutionary history of a species, i.e., how the precursor trait is sequentially shaped in a species across time by natural selection, the process by which should be reconstructed from the fossil precursors and DNA evidence), evolutionary explanations (how the trait offers a selective advantage based on its function, i.e., why individuals with this trait on average have a better adaptation to the environment and have more offspring than others) ([6], Tinbergen, 1963; [7], Nesse, 2000). The first two are proximate questions, explaining how the trait works; the other two are ultimate questions, explaining how the trait came to exist ([8], Mayr, 1982). In terms of the polyphenism in the caste system of social insects, although the phylogenetic and evolutionary significance has received thorough attention, the developmental genetic basis is still rarely known ([9], Moczek et al., 2011). That's because many social insects have widespread global distribution and ecological abundance, which make them excellent subjects for field-based studies and comparative work ([10], Kapheim, 2019), while many species require mating flights and generation time is typically years or even decades, which make them impossible to be bred in the lab and readily kept across generations.

Summary of the recent functional genetics study in ants

In recent years, more and more studies have focused on unraveling the developmental genetic basis of queen-worker caste differentiation in the study of social insects. The comparative genomic and transcriptomic studies for an increasing number of social insect species facilitate the functional genetic research ([11], Morandin et al., 2016; [12], Warner et al., 2017; [13], Libbrecht et al., 2018; [14], Walsh et al., 2018; [15], Warner et al., 2019; [16], Warner et al., 2019). A few ant species were exploited as laboratory-reared organisms and the protocols of genetic manipulation techniques like RNAi-based gene knockdown and CRISPR/Cas9-based gene knockout have been developed to study the gene functions in these non-model animals. Some key genes have been identified and considered to participate in caste differentiation in some ant species. For instance, ILP2, an insulin-like peptide which is consistently upregulated in reproducing castes of many ant species, is instrumental in promoting and suppressing reproduction ([17], Chandra et al., 2018). The neuropeptide corazonin controls behavioral transitions between ant workers and gamergate, the individuals with reproductive ability, in *Harpegnathos* ([18], Gospic et al., 2017). CRISPR-based *orco* mutant ants display asocial behavior as well as defective reproduction ([19], Yan et al. 2017; [20], Tribble et al., 2017). *vestigial*, a gene coordinates the growth of rudimentary wing discs, is shown to control worker castes determination to maintain the ratio of minor workers to soldiers in *Pheidole* ([21], Rajakumar et al., 2018). Other interesting works regarding the differentiation within worker castes have also been conducted to explore the potential epigenetic mechanisms. For instance, histone acetylation can enhance foraging and scouting behaviors in worker castes in the carpenter ant *Camponotus floridanus* ([22], Simola et al., 2016). However, core gene regulatory networks mediating ontogeny of queen and worker castes in ants remain to be uncovered.

Biology of *Monomorium pharaonis*

Advantages of using *Monomorium pharaonis* as a model organism

Monomorium pharaonis (Linnaeus, 1758), also known as pharaoh ant, is considered to be the most widespread and oldest invasive ant: it has been spread worldwide for at least 200 years from its native range in tropical Asia. It exhibits several tramp ant characteristics such as generalist diet, colony reproduction by budding (the process of colony division of a group of queens and worker ants from the original colony to establish a new colony at a new location ([23], Vail & Williams, 1994)), extensive polydomy (a nesting strategy that ants live within a network of multiple related nests near the main colony, usually within a confined area ([24], Tsutsui & Suarez, 2003)), unicoloniality (a social structure of ants that the workers can move freely between neighboring colonies without hostility ([25], Schmidt et al., 2010)), extreme polygyny (the presence of multiple queens in the same colony ([26], Keller, 1995)), and close association with humans ([27], Passera 1994).

M. pharaonis is well studied and has the following unique set of characteristics of being an emerging model organism for the sociogenomics study and functional genetics study of development, collective behavior and so on:

As a global pest species, *M. pharaonis* has extremely high worldwide genetic differentiation among its global populations ([25], Schmidt et al., 2010), while gene flow between colonies may be rare because the species exclusively disperses by budding under certain circumstances, such as chemical or physical disturbance, sudden depletion of the food supply, or overcrowding ([28], Buczkowski et al., 2005). The species can also be inadvertently transported via human-mediated jump dispersal ([29], Edwards & Baker, 1981). These genetically highly structured populations can be collected from all over the world and maintained in laboratory as genetically distinct stock lineages. Since pharaoh ants generally display only low levels of intraspecific

aggression, colonies of increased intra- and inter-individual genetic diversity can be created in the laboratory by crossing and mixing of different colony lineages without any significant agonistic interaction between unrelated workers or queens. ([30], Schmidt et al., 2011). This implies the flexibility in both queen number and worker number in the colony of pharaoh ant, providing unique opportunity for the studies under experimentally controlled conditions.

M. pharaonis is a highly polygynous species. A colony may contain up to several thousand workers and hundreds of fertile queens. The winged but flightless new queens and males become reproductively active shortly after eclosion. They do not do nuptial flight, instead, they have intranidal mating, i.e. mating within their natal nest without dispersal of queen or male ([31], Fowler et al., 1993). All queens are inseminated only once; each male inseminate two to four queens, and rarely one ([32], Petersen & Buschinger, 1971). Mated queens produce sexual broods at intervals of three to eight months without any seasonality in their reproductive cycles which may be due to their tropical origin ([33], Petersen-Braun, 1975). Similar to most polygynous budding ants, the production of gynes and males in *M. pharaonis* is inhibited in colonies containing fertile queens ([34], Boonen & Billen, 2017). When the existing queens senesce, die, or are experimentally removed from the colony, new sexual individuals are almost immediately reared from the existing eggs and 1st instar larvae, while the colony continues to produce workers concurrently ([35], Edwards, 1991). Therefore, inbred lineages can be easily maintained in the laboratory conditions across generations ([36], Peacock & Baxter, 1949)

M. pharaonis has an unambiguous separation of the worker and queen caste. Queens in pharaoh ant colony are morphological and physiological adapted specifically for reproduction and are clearly morphologically distinguishable from workers at an early age (2nd instar). All brood comes from queen-laid eggs as workers emerge with no ovary and are thus completely sterile, which means there will be no worker-produced male to

complicate breeding programs ([37], Berndt & Eichler, 1987). *M. pharaonis* contrasts with many other ant species of which workers are still able to lay unfertilized eggs ([38], Ichinose & Lenoir, 2009) or lay both fertilized and unfertilized eggs like the workers of gamergates species such as *Dinoponera quadriceps* ([39], Monnin et al., 2002).

The generation time of *M. pharaonis* is only six weeks, which is relatively shorter than most ant species (usually years).

M. pharaonis is one of the few social insect species with an available genome.

Colony-level regulation of *Monomorium pharaonis* caste allocation

The key to the success of social insect colonies, particularly in the case of highly polygynous societies, is the flexible allocation of resources to alternate queen and worker castes, which is analogous to the allocation of energy to somatic and reproductive tissues in multicellular organisms. In general, queen and worker caste fates in most eusocial hymenopteran species are determined at the larval stage in response to socially regulated environmental signals, such as worker-regulated quality and quantity of nutrition ([40], Linksvayer et al., 2011; [41], Tribble & Kronauer, 2017; [42], Lillico-Ouachour & Abouheif, 2017), queen pheromone ([43], Vargo & Passera, 1991; [44], Boulay et al., 2007), maternal effects and temperature ([45], Schwander et al., 2008; [46], Libbrecht et al., 2013).

In *M. pharaonis*, since queen and worker castes are likely to be determined blastogenetically (i.e., female caste is determined during embryonic stage via compounds endowed by the queen), cannibalism of male- and gyne-destined embryos or first instar larvae is probably the mechanism of caste allocation in the presence of fertile queens rather than maternal effect and later regulation of the caste fate of developing larvae based on feeding different nutritional quality or quantity ([47], Warner et al., 2018). The culling process executed by workers might be driven by the

presence of freshly laid eggs via its emitted pheromone signal ([48], Edwards, 1987). The mechanism of how workers discriminate between sexual- and worker- destined embryos or first instar larvae in *M. pharaonis* is still not clear, but it is likely that workers identify them by pheromone present on the cuticle which reflects endocrine changes associated with development, as workers in *Harpegnathos saltator* do ([49], Penick & Liebig, 2017). Because sexual individuals are killed at extremely young developmental stages before substantial resources have been invested in them, and they are likely eaten ([50], Tschinkel, 1993) so nutrients are recycled within the colony, the cost of laying excess sexual eggs may be low, but with this strategy, *M. pharaonis* colonies can invest nearly all resources in colony growth while maintaining the ability to produce new queens whenever necessary without using extra resources.

Colonies of *M. pharaonis* produce reproductives as a result of queen removal, which is the ultimate driver of the caste allocation shift, but it is still regulated by other colony components, for example, fewer reproductives are reared proportionally in the colony with larger size ([51], Schmidt et al., 2011) and with the lower ratio of the number of workers to the number of eggs because a high number of eggs presumably signals the presence of many fertile queens. ([47], Warner et al., 2018). The presence of worker 3rd instar has been shown to have a stimulatory effect on a colony's ability to rear reproductives because they can enhance a colony's nutrient assimilation by processing solid insect protein, and the secretions will be fed to sexual larvae as their nourishment ([52], Warner et al., 2016). In addition, if new-founded colonies that were producing sexual broods were infected by endosymbiont *Wolbachia*, the colonies displayed strong female bias in sex ratio and a bit gyne bias in caste ratio ([53], Pontieri et al., 2016).

Although the mechanisms regulating colony-level resource allocation into alternate castes are well studied, the developmental genetic basis of individual-level caste fate determination is largely unknown. *M. pharaonis* and other *Monomorium* species likely possess blastogenic caste determination. In the congeneric *M. emersoni*, female caste

appears to be determined in the egg stage because in worker-destined embryos, the germ cells responsible for the formation of the adult gametes were observed to degenerate early in development ([54], Khila & Abouheif, 2010). This enlightens the following research focusing on the evolutionary genetic underpinnings of the developmental processes of caste determination and differentiation.

Caste system and developmental scheme of *Monomorium pharaonis*

The larval stages in *M. pharaonis* were determined by hair presence and morphology as well as body size and shape ([55], Berndt & Kremer, 1986). Sexualls and workers cannot be differentiated morphologically in embryos and 1st larval stage. The 1st instar is one to two times the size of embryo and the tummy panel is hairless. From 2nd larval stage, worker larvae are covered with hairs all over the body while sexual larvae are larger and almost hairless except sparse hairs on the head. Worker 2nd instar has straight hairs and is around four times the size of embryo, which is nearly the same size as adult worker head; sexual 2nd instar is about 1/2 size of worker pupa. Worker 3rd instar has branched hairs and was defined as “small” until it reached 3/4 size of worker pupa, “medium” until the size of worker pupa, and “large” thereafter; sexual 3rd instar was defined as “small” until it reached the size of worker pupa, “medium” until 1.5 size of worker pupa, and “large” thereafter. Demographic classes are defined as follows: young brood (eggs and 1st instar larvae), middle-aged larvae (2nd instar and small 3rd instar), late-instar larvae (medium 3rd instar and large 3rd instar), prepupae and pupae, and adults (**Figure 1**). Gyne and male cannot be differentiated morphologically until it develops into pupa (**Figure 2**). Compared to gyne pupa, male pupa has larger eyes, longer antennae and a bit shorter body length. Male usually develops faster than gyne and becomes into adult several days earlier, so that males are ready to mate as soon as gynes come out. The embryo stage lasts about 8 days ([56], Peacock & Baxter, 1950); each larval stage (1st, 2nd, small 3rd, medium 3rd and large 3rd) last about 3 days with prepupal and pupal stage lasting about 4 days and 12 days, respectively (personal

observation), thus the generation time of *M. pharaonis* from embryo to adult is around six weeks.

M. pharaonis workers are monomorphic but go through age-based division of labor (age polyethism) ([57], Mikheyev & Linksvayer, 2015). Young workers with light yellow color (called callows) perform tasks inside the nest, such as brood tending and colony cleaning, while the darker older workers perform outside tasks such as foraging. Older workers will also perform callows' tasks if there is no callow in the colony. Some young workers use their crop to store nourishment secreted by third-instar larvae, which will be mainly provided to reproductive queens to boost egg production; such workers form a temporary caste called replete workers ([58], Børgesen, 2000). The lifespan of worker is around 10 weeks, but queens live much longer, usually about 6–8 months ([59], Pontieri & Linksvayer, 2019).

Aim of this study

Compared to the cellular differentiation process, which is regulated by complex gene regulation networks controlled by some key transcription factors, we predict that differential expression of some key genes during ontogeny probably also account for the differential developmental processes of castes. My host group has identified several candidate genes with significant queen-worker caste-differentiated expression among nearly all developmental stages through analysis of the transcriptome of sexual individuals and worker individuals in *M. pharaonis*. This project focuses on one of the candidate genes (gene ID: 13359) which has never been studied before but shows significantly expression difference since the castes can be morphological recognized at 2nd instar in *Monomorium pharaonis*. The main objective of this study is to uncover the evolutionary history of this gene and its biological functions in ontogeny and caste determination in *Monomorium pharaonis*. I explored this mainly by knocking down the gene expression at larval stages through RNAi and then observing the consequent morphological changes during development.

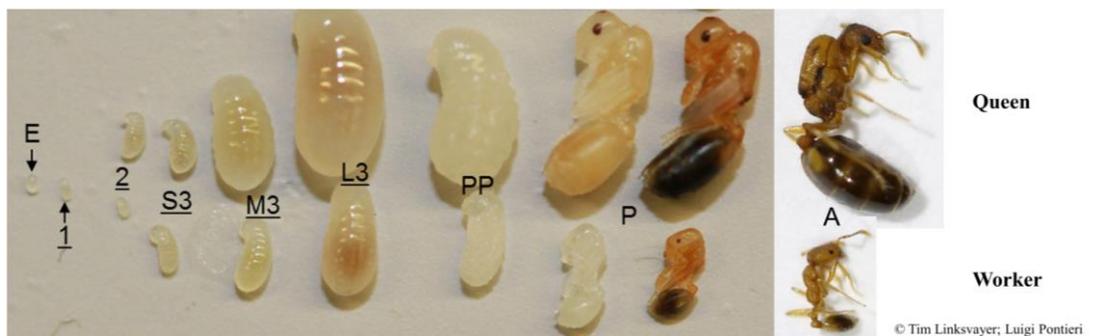


Figure 1. Morphological characters for gyne (unmated queen) and worker at different developmental stages in *M. pharaonis*

E, 1, 2, S3, M3, L3, PP, P and A represent embryo, 1st, 2nd, small 3rd, medium 3rd, large 3rd instar, prepupae, pupae and adults, respectively. The body length of gyne pupa is around 3 mm. © Tim Linksvayer; Luigi Pontieri



Figure 2. Development of gyne and male pupa in *M. pharaonis*

The pupal stage lasts about 12 days. With the maturation of pupa, the body color of gyne gradually turns yellow and male turns black. The eyes and ocelli develop further, and the color turns darker. After pupal stage the individual ecloses into adult with the help from workers in the colony. © Luigi Pontieri

Chapter 1

The evolutionary history of gene13359 across the Formicidae

Results

To study the evolutionary history of gene13359 across the Formicidae, the phylogenetic analysis and positive selection analysis were performed. Two wasps (outgroup) and ninety ants (ingroup) were included in this study (**Supplementary Figure 1.1 and Supplementary Table 1.1**). The homolog of gene13359 in each of these species was identified based on sequence homology and microsynteny, and the coding sequences were used in the analysis.

Construction of the phylogenetic tree of gene13359

The phylogenetic analysis revealed that the evolution of gene13359 reflected the speciation process across the Formicidae. The phylogeny constructed from the coding sequences of gene13359 homologs (**Figure 1.1**) was consistent with taxonomic relationships among species (**Figure 1.2**) on the level of subfamily except Ectatomminae. The gene phylogeny showed that the Ectatomminae formed an outgroup to the Formicinae, which was contradictory with the taxonomic relationship between these two subfamilies. However, the study only included one Ectatomminae species, *Gnamptogenys bicolor*, which might bring uncertainty to the construction of gene phylogenetic tree. Three Myrmicinae ants, *Temnothorax ambiguous*, *Strongylognathus testaceus* and *Myrmecina graminicola*, had different phylogenetic positions within the gene tree from the species taxonomy. According to the taxonomic relationship, *Temnothorax ambiguous* should be gathered with *Temnothorax* species which were belonged to the tribe of Crematogastrini, not with *Messor* species which was belonged

to the tribe of Stenammini as shown in the gene tree. *Strongylognathus testaceus* was a Crematogastrini species, which shouldn't be gathered with Solenopsidini species such as *Monomorium pharaonis* and *Megalomyrmex milenae* shown in the gene tree. *Myrmecina graminicola* was a Crematogastrini species, which shouldn't be gathered with Attini species like *Acromyrmex* and *Pheidole* shown in the gene tree. The sample and genome data of these species need to be further checked to verify that it is not because of the errors during the identification of samples or the contamination of DNA that accounts for such results. Except these three species, other species in the gene tree were gathered in the same way as species taxonomy.

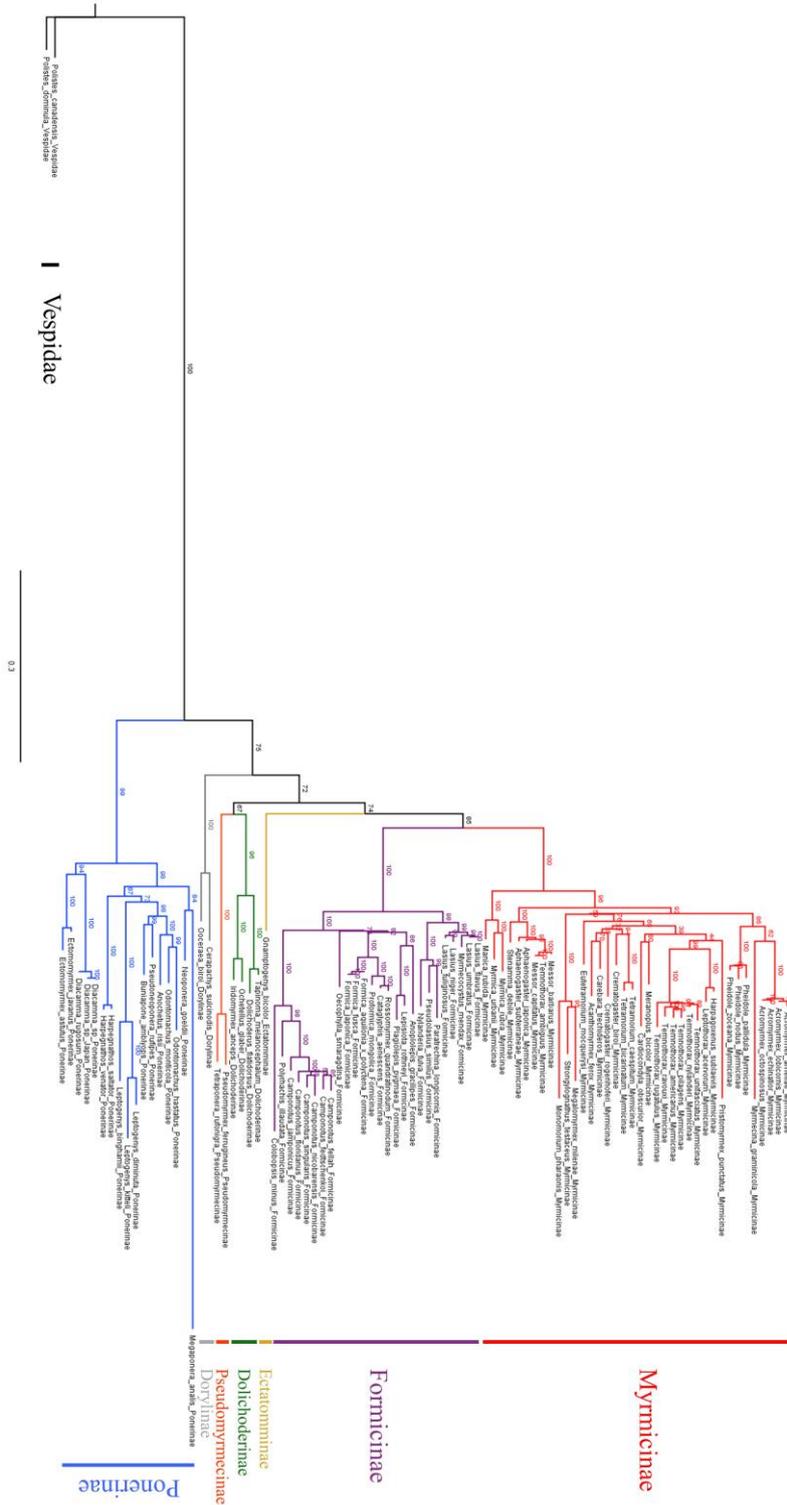


Figure 1.1. Phylogenetic tree of gene13359 across the Formicidae constructed from the coding sequences of ninety-two species (ninety ants and two wasps) with wasps set as roots

Homologs of gene13359 in species other than *M. pharaonis* were identified based on sequence homology and synteny (see Materials and methods). The tree was created by IQ-TREE2.0 program with MGK+ F1X4+ R3 model.

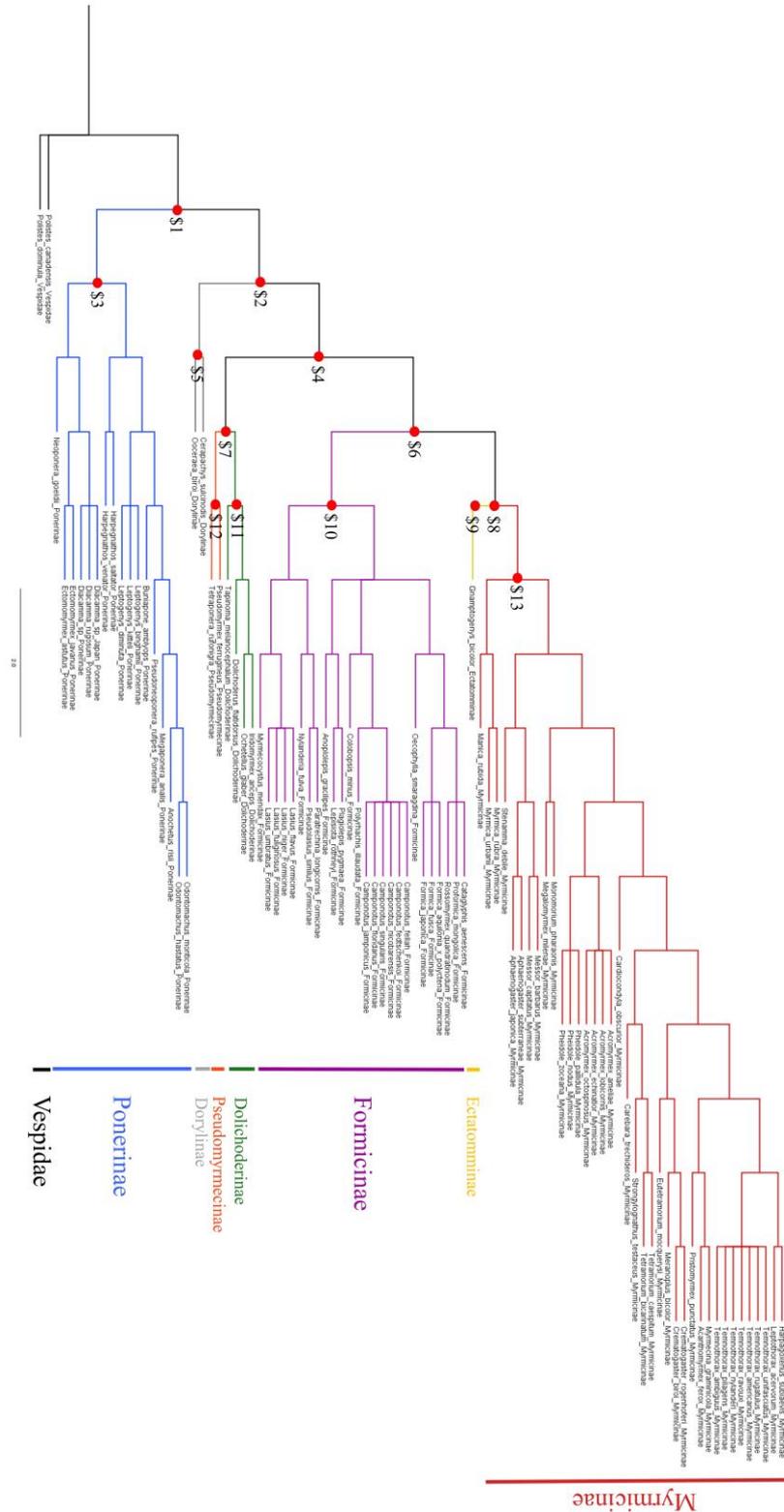


Figure 1.2. Phylogenetic tree showing the taxonomic relationships among ninety-two species (ninety ants and two wasps) that were included in the selection analysis

The relationships were mainly based on the AntWiki website while some of the species were based on the papers (see Materials and methods).

Prediction of properties of gene13359 protein

In addition to the construction of gene phylogenetic tree, the coding sequences were used in prediction of domains and subcellular localization of gene13359 protein. Each protein of these 92 species was predicted. The results showed that there was no difference in the composition of domains of gene13359 protein among wasp and ant species. According to the Phobius program on InterPro website, each protein was predicted to have a signal peptide which was about 25-amino-acid-long at the N-terminus, while the remaining part of the protein was predicted to be a non-cytoplasmic domain (**Figure 1.3**). No transmembrane structure was detected. DeepLoc website also showed that each protein was soluble with 80% probability on average to be extracellular. Uniprot website was utilized to make blast for gene13359 protein but no known protein family was detected to rank top on the list.

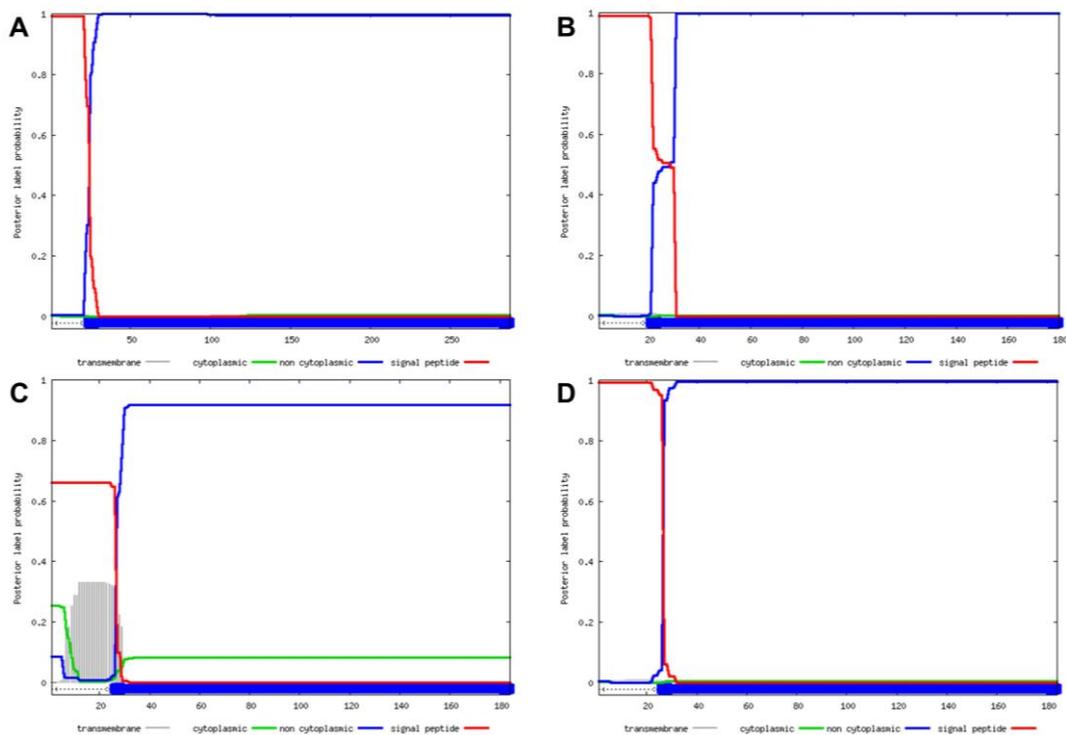


Figure 1.3. Prediction of domains in gene13359 protein according to the Phobius program

(A) *Polistes canadensis* (wasp in the Vespidae), (B) *Harpegnathos saltator* (Ponerinae ant), (C) *Camponotus floridanus* (Formicinae ant) and (D) *Monomorium pharaonis* (Myrmicinae ant).

Probability of the scheme varies among species, but in general it was larger than 0.5.

Positive selection analysis on the coding sequences of gene13359

The amino acid sequences of gene13359 homologs among different subfamilies exhibited highly divergence. For example, the average identity between *Acromyrmex echinator* and Formicinae ants was about 50%, and for Ponerinae ants it was about 40%. After removing unreliable alignment regions by Gblocks, 45 positions were selected, which was only 25% of the whole length of the original protein. Since the remove of unreliable alignment regions was not a necessary step for positive selection analysis, the original sequences without being treated by Gblocks were used for CodeML analysis. Ninety-two species including two wasps and ninety ants were included in the study and 1395 sites were analyzed for each species.

CodeML-based positive selection analysis was performed on the coding sequences for gene13359. In branch model, H0 was fitted by specifying model=0, which assumed only one ω ratio ($\omega = d_N/d_S$, the ratio of nonsynonymous/synonymous substitution rates) for all branches; H1 was specified by model=2, which allowed several ω ratios to fit the specified branches. In this study, each H1 model had two or three branch types by specifying one or two clade labels \$ in the tree file. Thirteen H1 models with two branch types (from \$1 to \$13 in **Figure 1.2**) and one H1 model with three branch types were tested to detect which particular lineages (called labeled foreground branches) had a different ω ratio from background branches. The results (**Table 1.1**) showed that gene13359 had different ω ratios between Ponerinae ants ($\omega=0.68149$) within the poneroid clade and ants within the formicoid clade ($\omega=0.50988$) according to the Comparison H0-H1 (\$3; \$2).

Table 1.1. Results of the positive selection analysis on gene13359 homologs in ninety-two species using branch model in CodeML program

Comparison	2 Δ lnL	d.f.	P-value	Significant after LRT	ω ratio (background; foreground1; foreground2)
H0	/	/	/	/	0.53339
H0-H1 (\$1)	4.716	1	0.030	Yes	0.31920; 0.53937
H0-H1 (\$2)	4.785	1	0.029	Yes	0.60880; 0.50972
H0-H1 (\$3)	13.839	1	0	Yes	0.49709; 0.68225
H0-H1 (\$4)	1.103	1	0.29	No	0.56472; 0.52044
H0-H1 (\$5)	7.955	1	0.005	Yes	0.54511; 0.31718
H0-H1 (\$6)	0.201	1	0.65	No	0.54390; 0.52652
H0-H1 (\$7)	0.553	1	0.46	No	0.53794; 0.49161
H0-H1 (\$8)	0.0007	1	0.98	No	0.53299; 0.53404
H0-H1 (\$9)	0.0234	1	0.88	No	0.53379; 0.51062
H0-H1 (\$10)	0.0705	1	0.79	No	0.53597; 0.52381
H0-H1 (\$11)	0.0558	1	0.81	No	0.53450; 0.51506
H0-H1 (\$12)	0.824	1	0.36	No	0.53678; 0.45108
H0-H1 (\$13)	0.004	1	0.95	No	0.53248; 0.53495
H0-H1(\$3; \$2)	23.063	2	0	Yes	0.27532;0.68149;0.50988

Because only two branch types were allowed in the branch-site model, three comparisons that were significant after LRT from branch models (\$1, \$2, \$3) were tested for branch-site model. H0 was specified by fix_omega=1 and omega=1, assuming that ω ratio for all sites was fixed as one; H1 was specified by fix_omega=0, which allowed ω ratio to vary among sites. Branch-site models were run to detect which sites were affected by positive selection within foreground branches. The results (**Table 1.2**) showed that comparisons H0-H1 (\$1, \$2) were significant and the sites under positive selection (ω ratio > 1) for foreground lineages were selected according to the results of BEB (Bayes Empirical Bayes) analysis. Since the alignment sequences used for CodeML analysis were not filtered by Gblocks, gaps surrounding the selected sites might bring uncertainty to the calculations of their ω ratios, thus the results were not put emphasis on in this study.

Table 1.2. Sites under positive selection for foreground lineages by branch-site model in CodeML program

Comparison	2 $\Delta\ln L$	d.f.	P-value	Significant after LRT	No. of sites selected (*Prob>0.95; **Prob>0.99)
H0-H1 (\$1)	32.581	1	0	Yes	3 with*; 3 with **
H0-H1 (\$2)	24.173	1	0	Yes	3 with*; 2 with **
H0-H1 (\$3)	0.337	1	0.56	No	/

Abbreviations: Prob, Probability

Discussion

The results of positive selection analysis showed that gene13359 had differential ω ratios between Ponerinae ants ($\omega=0.68149$) within the poneroid clade and ants within the formicoid clade ($\omega=0.50988$)

While ant eusociality originated in the last common ancestor of the Formicidae, the formicoid ants were considered to be advanced eusocial species; the poneroid ants, sister clade to the formicoid ants, were primitive eusocial species. The most important characteristic of formicoid ants was that life-long unmated workers arose, which signaled the evolutionary transition to an advanced eusociality, i.e., superorganismality, and led to a remarkable success of formicoid ants. Such life-long unmated workers lost mating ability because of degeneration of the spermatheca, an organ for long-term sperm storage ([60], Gobin et al., 2008); workers in some species such as *Monomorium pharaonis* even lost their ovaries. There was also a significant difference between the morphology of queens and workers. Worker caste in formicoid ants could be compared to permanent somatic cells in multicellular organisms ([61], Boomsma & Gawne, 2018). In poneroid ants, queens showed slight morphological difference from their workers except wings. Importantly, workers still retained functional spermathecas like the queen did, and in many species, workers could mate and reproduce to become gamergates

once the original queen in the colony died. Thus, the eusociality in poneroid ants was considered to stay at a primitive state.

Our results showed that gene13359 was under stronger selection pressure in Ponerinae ants than in formicoid ants. In our knowledge, gene13359 was supposed to be responsible for the caste differentiation in formicoid ants or at least in *Monomorium pharaonis*, so it was hypothesized to be positively selected within the formicoid clade which had advanced eusociality. So far, we could not give a clear answer to why this gene evolved faster in Ponerinae ants, but we know that in Ponerinae, there was a large diversity on social organizations, mating systems, morphology, etc., between genera ([78], Schmidt, 2013), while in formicoid ants the difference between genera was not so big, so it seems that Ponerinae ants were experiencing adaptive radiation, during the process gene13359 was favored by natural selection. Because in formicoid ants there was a clear reproductive division of labor, it might be that in Ponerinae ants, some functions of gene13359 were related to the specified female physiology of queens and their sister workers that still maintained plasticity in reproduction, and such functions were selected during the evolution.

It was shown that the phylogeny of gene13359 almost matched the taxonomic relationships among species across the Formicidae. The amino acid sequences of gene13359 homologs were divergent among ninety ant species in seven subfamilies in the Formicidae and two Polistinae wasps in the Vespidae, while there was no difference in the composition of domains of gene13359 protein. Each protein was predicted to be soluble and localized extracellularly, with a signal peptide at the N-terminus and a non-cytoplasmic domain.

Materials and methods

Identification of gene13359 homologs and phylogenetic reconstruction

Two wasps and ninety ant species were included in this study (**Supplementary Figure 1.1 and Supplementary Table 1.1**). Genome data of ant species were from The Global Ant Genomics Alliance, GAGA (<http://www.antgenomics.dk/>) and were stored on the BGI server. We noticed that the sequences of gene13359 homologs exhibited large divergence among different subfamilies, so the amino acid sequences of gene13359 homologs of *Harpegnathos saltator* (taxid on NCBI:610380), *Camponotus floridanus* (taxid on NCBI:104421) and *Acromyrmex echinator* (taxid on NCBI:103372) were used as inputs to find homologs in Ponerinae ants, Camponotini ants in Formicinae and other ant species, respectively, through tblastn research where the E value threshold was set as 1e-5.

In addition to the sequence homology, the synteny analysis was also conducted to help identify gene13359 homologs by using the genomes of *Drosophila melanogaster* and *Atta cephalotes* as references. Among those ninety ant species, three Ponerinae ants (*Megaponera analis*, *Neoponera goeldii*, *Pseudoneoponera rufipes*), one Dorylinae ant (*Ooceraea biroi*), one Dolichoderinae ant (*Dolichoderus flatidorsus*), one Ectatomminae ant (*Gnamptogenys bicolor*), seven Formicinae ants (*Lepisiota rothneyi*, *Lasius flavus*, *Nylanderia fulva*, *Paratrechina longicornis*, *Camponotus jamponicus*, *Camponotus nicobarensis*, *Camponotus singularis*) and three Myrmicinae ants (*Harpagoxenus sublaevis*, *Meranoplus bicolor*, *Manica rubida*) were selected for the synteny analysis. The results showed that the microsynteny of gene13359 was conserved in all those species except *Ooceraea biroi*, which experienced chromosomal translocations. It should be noted that *Ooceraea biroi* had unique biology because it was a queenless species where workers reproduced through thelytokous parthenogenesis, which meant females were hatched from unfertilized eggs ([62], Rabeling & Kronauer, 2013), but so far there was no evidence showing relationship

between these two events. We thus concluded that the microsynteny of gene13359 was conserved across the Formicidae. The microsynteny of gene13359 in two wasp species *Polistes canadensis* and *Polistes dominula* was the same as in the Formicidae.

After identifying the homolog from genome data, the coding sequence (CDS) was extracted by using the scripts (genewise.pl, gw2gffWithShift.pl, getGene.pl) from BGI. The coding sequences from ninety ant species and two wasp species were used to reconstruct the phylogenetic tree of gene13359 across the Formicidae by IQ-TREE2.0 program with MGK+ F1X4+ R3 model. Wasp was set as root. The coding sequence was converted into amino acid sequence by cds2aa.pl script, which was then used for the prediction of protein domains based on InterPro website (<https://www.ebi.ac.uk/interpro/>) and protein subcellular localization based on DeepLoc website (<http://www.cbs.dtu.dk/services/DeepLoc/>).

Multiple sequence alignment and positive selection analysis

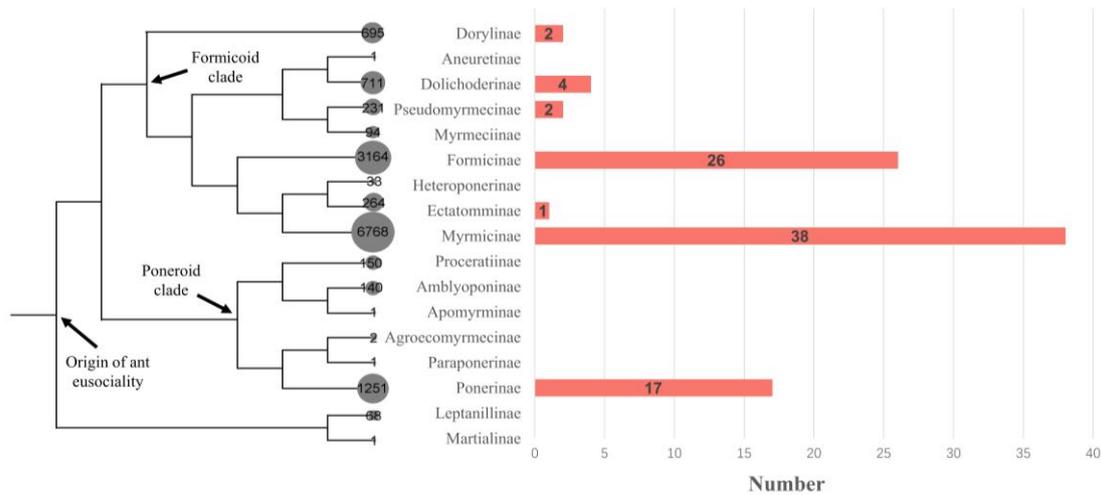
PRANK v.170427 program was used to align the amino acid sequences of gene13359 homologs of the ninety-two species (ninety ants and two wasps), which was then run by the script pepMfa_to_cdsMfa.pl to convert into coding sequences used for positive selection analysis.

After alignment of the amino acid sequences by PRANK, the unreliable alignment regions that might account for the uncertainty of multiple parameters within the selection analysis could be removed by Gblocks 0.91b program as an optional step. In this study, since the sequences were highly divergent, the parameter of b4 (minimum length of a block) was set as 2 and other parameters were set as default. The selected regions were converted into coding sequences by the scripts (read_gbhtm.pl and ch2phy.pl) and were then used for the following positive selection analysis.

The phylogeny tree used in positive selection analysis was created based on the taxonomic relationships among species shown on the AntWiki website (<https://www.antwiki.org/>) (**Figure 1.2**). Since not all genus were shown on the website, papers were searched to determine the phylogenetic relationships of such species. *Strongylognathus testaceus* was together with *Tetramorium* species ([63], Sanetra & Buschinger, 2000); *Rossomyrmex quadratinodum* was together with *Proformica* and *Cataglyphis* ([64], Hasegawa et al., 2002); *Colobopsis minus* was together with *Camponotus* and *Polyrhachis* ([65], Ward et al., 2016); *Lepisiota rothneyi* was together with *Plagiolepis pygmaea* and *Anoplolepis gracilipes* ([66], Blaimer et al., 2015).

CodeML program in PAML4.9h package was used for the positive selection analysis based on the coding sequences of homologs and the species tree. Likelihood ratio tests (LRTs) (*G*-tests) were performed to test whether the difference between null hypothesis (H₀) and alternative hypothesis (H₁) was significant (p-value < 0.05). Branch model and branch-site model were performed in the analysis. The parameters in each model were set as follows (**Supplementary Table 1.2**). The analysis was performed on the BGI server.

Supplementary materials



Supplementary Figure 1.1. Cladogram of the ninety ant species used in PAML analysis in the context of the entire ant phylogeny with all subfamilies labeled

Genome data of these ant species come from The Global Ant Genomics Alliance, GAGA (<http://www.antgenomics.dk/>). The phylogeny is based on [67], Borowiec et al., 2019. Species numbers come from AntCat website (<https://www.antcat.org/>).

Supplementary Table 1.1. List of the two wasp species and ninety ant species used in PAML analysis

Formica aquilonia x *polycтена* was the species given by the hybridization between *Formica aquilonia* and *Formica polycтена*. *Diacamma* sp. and *Diacamma* sp. nov. from Japan were two species belonging to the genus of *Diacamma* but they were not identified at the species level.

Family: Vespidae (2 species)		
<i>Polistes canadensis</i>	<i>Polistes dominula</i>	
Family: Formicidae (90 species)		
Subfamily: Myrmicinae (38 species)		
<i>Myrmica rubra</i>	<i>Myrmica urbanii</i>	<i>Manica rubida</i>
<i>Messor barbarus</i>	<i>Messor capitatus</i>	<i>Aphaenogaster subterranea</i>
<i>Stenamma debile</i>	<i>Monomorium pharaonis</i>	<i>Megalomyrmex milenae</i>
<i>Acromyrmex ameliae</i>	<i>Acromyrmex lobicornis</i>	<i>Acromyrmex echinator</i>
<i>Acromyrmex octospinosus</i>	<i>Pheidole pallidula</i>	<i>Pheidole nodus</i>

Supplementary Table 1.1

<i>Meranoplus bicolor</i>	<i>Pheidole zoceana</i>	<i>Aphaenogaster japonica</i>
<i>Temnothorax unifasciatus</i>	<i>Temnothorax rugatulus</i>	<i>Myrmecina graminicola</i>
<i>Temnothorax americanus</i>	<i>Temnothorax ravouxi</i>	<i>Temnothorax nylanderi</i>
<i>Temnothorax pilagens</i>	<i>Temnothorax ambiguus</i>	<i>Tetramorium caespitum</i>
<i>Tetramorium bicarinatum</i>	<i>Pristomyrmex punctatus</i>	<i>Crematogaster biroi</i>
<i>Crematogaster rogenhoferi</i>	<i>Acanthomyrmex ferox</i>	<i>Cardiocondyla obscurior</i>
<i>Carebara trechideros</i>	<i>Eutetramorium mocquerysi</i>	<i>Harpagoxenus sublaevis</i>
<i>Strongylognathus testaceus</i>	<i>Leptothorax acervorum</i>	
Subfamily: Formicinae (26 species)		
<i>Anoplolepis gracilipes</i>	<i>Lepisiota rothneyi</i>	<i>Plagiolepis pygmaea</i>
<i>Lasius flavus</i>	<i>Lasius niger</i>	<i>Lasius fuliginosus</i>
<i>Lasius umbratus</i>	<i>Myrmecocystus mendax</i>	<i>Nylanderia fulva</i>
<i>Paratrechina longicornis</i>	<i>Pseudolasius similis</i>	<i>Camponotus japonicus</i>
<i>Camponotus fellah</i>	<i>Camponotus fedtschenkoi</i>	<i>Camponotus nicobarensis</i>
<i>Camponotus singularis</i>	<i>Colobopsis minus</i>	<i>Polyrhachis illaudata</i>
<i>Formica fusca</i>	<i>Formica japonica</i>	<i>Formica aquilonia</i> x <i>polycтена</i>
<i>Cataglyphis aenescens</i>	<i>Proformica mongolica</i>	<i>Rossomyrmex quadratinodus</i>
<i>Oecophylla smaragdina</i>	<i>Camponotus floridanus</i>	
Subfamily: Ponerinae (17 species)		
<i>Diacamma rugosum</i>	<i>Diacamma sp.</i>	<i>Diacamma sp. nov. from Japan</i>
<i>Leptogenys diminuta</i>	<i>Leptogenys kitteli</i>	<i>Leptogenys binghamii</i>
<i>Ectomomyrmex javanus</i>	<i>Ectomomyrmex astutus</i>	<i>Odontomachus hastatus</i>
<i>Odontomachus monticola</i>	<i>Anochetus risii</i>	<i>Buniapone amblyops</i>
<i>Harpegnathos venator</i>	<i>Harpegnathos saltator</i>	<i>Megaponera analis</i>
<i>Neoponera goeldii</i>	<i>Pseudoneoponera rufipes</i>	
Subfamily: Dolichoderinae (4 species)		
<i>Iridomyrmex anceps</i>	<i>Ochetellus glaber</i>	<i>Dolichoderus flatidorsus</i>
<i>Tapinoma melanocephalum</i>		
Subfamily: Ectatomminae (1 species)		
<i>Gnamptogenys bicolor</i>		
Subfamily: Pseudomyrmecinae (2 species)		
<i>Pseudomyrmex ferrugineus</i>	<i>Tetraoponera rufonigra</i>	
Subfamily: Dorylinae (2 species)		
<i>Cerapachys sulcinodis</i>	<i>Ooceraea biroi</i>	

Supplementary Table 1.2. List of the parameter settings used in branch model and branch-site model in CodeML selection analysis

	Branch model H0	Branch model H1	Branch-site model H0	Branch-site model H1
noisy	3	3	3	3
verbose	1	1	1	1
runmode	0	0	0	0
seqtype	1	1	1	1
CodonFreq	2	2	2	2
clock	0	0	0	0
model	0	2	2	2
NSsites	0	0	2	2
icode	0	0	0	0
fix_kappa	0	0	0	0
kappa	2.5	2.5	2	2
fix_omega	0	0	1	0
omega	0.2	0.2	1	1.5

Chapter 2

The temporal and spatial expression patterns of gene13359

Results

Temporal expression patterns of gene13359 based on transcriptome analysis

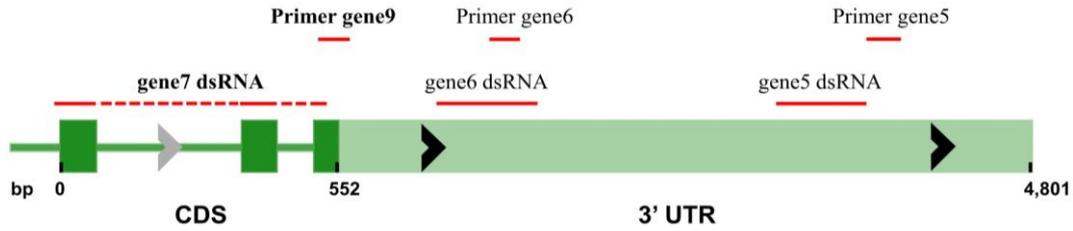
The previous RNA-seq results from Bitao Qiu's work showed that in *Monomorium pharaonis*, gene13359 had a significant expression difference since the castes could be morphological recognized at 2nd instar (**Supplementary Figure 2.1A**). The patterns showed that the expression level in worker caste was very low from 2nd instar to adult while in gyne the level was much higher. With the development of gyne pupa, the expression level became lower and lower and finally reached nearly the same level as worker in adult. For male, in the 2nd and 3rd larval stages, the level was generally lower than gyne but still higher than worker; from prepupal stage to adult, the level was close to worker. The homolog of gene13359 in the ant species *Acromyrmex echinator* was identified according to its sequence homology and microsynteny, and its expression patterns were similar to gene13359 in *Monomorium pharaonis* (**Supplementary Figure 2.1B**).

Selection of the efficient ddPCR primers for gene13359

To validate these expression patterns, ddPCR (droplet digital PCR) was performed to quantify the transcriptional level of gene13359 across different castes and developmental stages. Since gene13359 had never been studied before and the functional structure of this gene was almost unknown, three pairs of qualified ddPCR primers targeting at different sites on the transcript were used to test which pair of primer was the most efficient one to detect the transcriptional level (**Figure 2.1A**).

Primer gene5 targeted at the far end of 3' UTR of the transcript; primer gene6 targeted at nearer 3'UTR and primer gene9 targeted at the junction part of coding sequence (CDS) and 3' UTR. RNA from each of the 15 sexual individuals from different developmental stages was extracted and was then reverse transcribed into cDNA which was then used as template in ddPCR. Three pairs of primers were tested by using the same cDNA template and thus 15 replicates were performed. ddPCR results showed that the region targeted by primer gene9 had a much higher concentration value in the unit of copies/ μ l, which represented a higher level of transcript copies, than the regions targeted by primer6 and primer5. If the concentration values resulted from primer gene6 were standardized to one, gene9 was around 17-fold higher than gene6, and gene5 was around 3-fold lower than gene6, demonstrating that the CDS of gene13359 had a higher transcriptional level than the far end of 3' UTR (**Figure 2.1B**) and primer gene9 could reflect the most accurate expression level of the functional part of gene13359. This pattern also suggested that the functional part of the 3' UTR on the transcript of gene13359 was much shorter than expected.

A



B

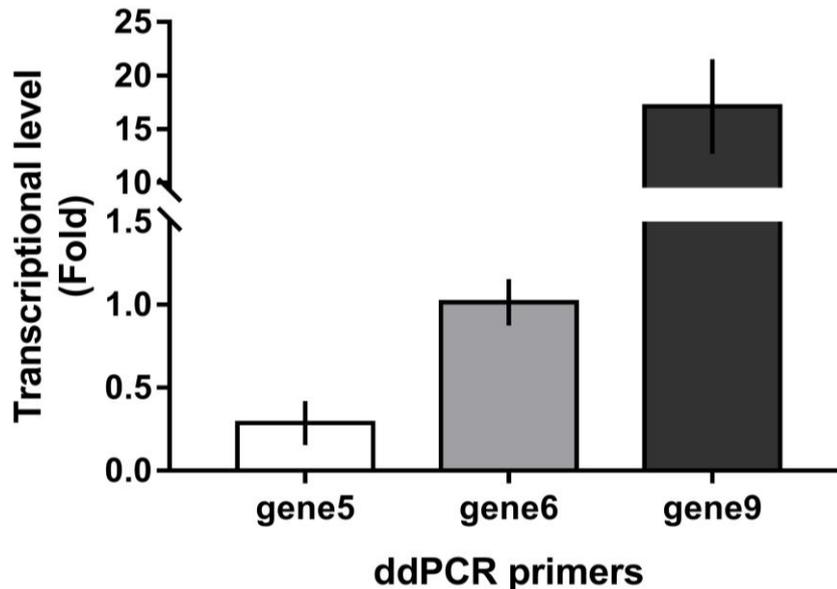


Figure 2.1. Efficiency of ddPCR primers to detect the transcriptional level of gene13359

(A) Locations of three ddPCR primers and three dsRNAs (double-stranded RNAs, see Chapter 3) on the coding sequence (CDS) and 3' UTR of gene13359 transcript. CDS was 552 bp long and 3' UTR was 4,249 bp long according to the NCBI database. The dark green blocks represented CDS and the light green block represented 3' UTR. The arrows represented the direction of transcription. 5' UTR was not shown in the figure because there were four different isoforms. (B) Transcriptional levels of three different regions of gene13359 transcript targeted by ddPCR primers. Primer gene5 targeted at the far end of 3' UTR; primer gene6 targeted at nearer 3'UTR and primer gene9 targeted at the junction part of the CDS and 3' UTR.

Temporal expression patterns of gene13359 based on ddPCR

After selecting primer gene9 to measure the transcriptional level of gene13359 by ddPCR, *NADH* (*NADH-ubiquinone oxidoreductase subunit 8*) was used as a reference gene as it had previously been validated as a reference gene for quantitative PCR in ants ([68], Livramento et al., 2018). *RPL18* (*Ribosomal Protein L18*) was used as another reference gene ([69], Jia et al., 2018). Multiple individuals of worker caste and sexual caste with different developmental stages (3rd instar, prepupa, young pupa, medium pupa and old pupa) were collected from queenless sub-colonies and RNA from each individual was extracted. For worker pupae, four individuals were pooled together to get enough RNA for performing ddPCR. During the RNA extraction process of sexual individuals, DNA was also extracted to be further used in microsatellite genotyping in order to differentiate male (haploid) from gyne (diploid). The extracted RNA was reverse transcribed into cDNA which was used as template in ddPCR. The ddPCR results using *NADH* as a reference gene (**Figure 2.2**) showed the same expression patterns as the results from transcriptomic data (**Supplementary Figure 2.1A**). *RPL18* was also used as a reference gene and showed the similar patterns, except that fewer replicates were conducted (**Supplementary Figure 2.2**).

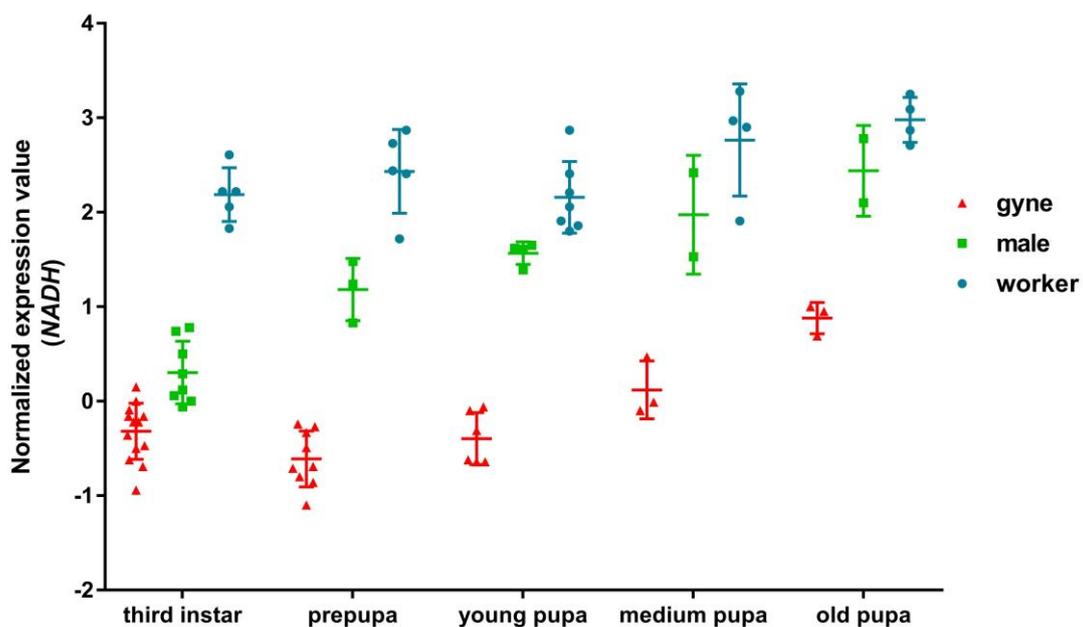


Figure 2.2. Transcriptional level of gene13359 across different castes and developmental stages using *NADH* as a reference gene

Each dot represented one individual except that for worker pupae, each dot represented four individuals. The normalized expression value of gene13359 in the sample was calculated by taking the logarithmic value of the ratio between the reference gene transcript concentration of the same sample and gene13359 transcript concentration, which meant that the lower normalized expression value represented the higher transcriptional level, and the values differed by one represented the levels differed by ten folds.

Spatial expression patterns of gene13359 in gyne pupa based on ddPCR

To preliminarily study the spatial expression pattern of gene13359, different body parts (head, thorax and abdomen) of young gyne pupae, which had been shown to exhibit the highest expression level, were collected. Six heads, thoraxes and abdomens were pooled together respectively in order to get enough RNA for performing ddPCR. The ddPCR results using *NADH* and *RPL18* as reference genes showed that in the young pupal stage of gyne, gene13359 exhibited the highest expression level in abdomen that was close to the level in whole body, which was about two-magnitude higher than worker pupa. The lowest expression level was exhibited in head, but it was still more than ten-fold higher than worker pupa (**Figure 2.3**).

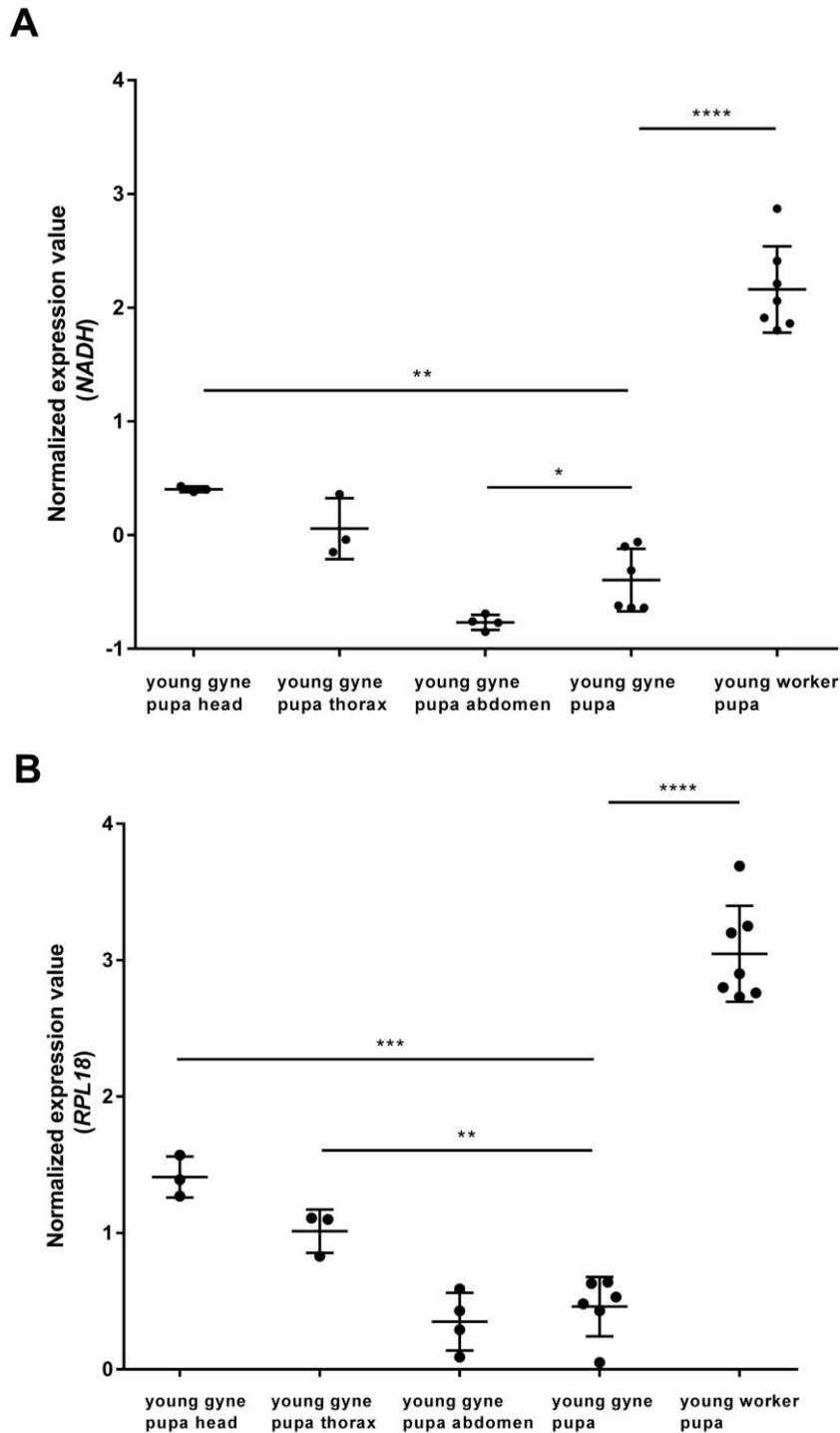


Figure 2.3. Transcriptional level of gene13359 in different body parts of young gyne pupa

Each dot in young gyne pupa head, thorax and abdomen represented six samples. Each dot in young gyne pupa and young worker pupa represented one individual and four individuals respectively. Unpaired t-test was used on the normalized values to examine if any statistical significance was present between two groups of samples. * P value<0.05; ** P value<0.01; *** P value<0.001; **** P value<0.0001. (A) *NADH* and (B) *RPL18* were used as reference genes.

Discussion

The expression of gene13359 was both gyne-male sex-biased and gyne-worker caste-biased during development from larval to young pupal stages in *Monomorium pharaonis* and *Acromyrmex echinator*

Previous studies hypothesized that genes with existed expression bias among tissues or between sexes, for example, in the nonsocial ancestor of eusocial insects like ants and honeybees, were more likely to be recruited in the gene regulatory networks (GRNs) underlying the caste differentiation in eusocial insects ([15], Warner et al., 2019; [70], Hunt et al., 2011). In other words, the gyne-worker biased expression patterns of certain genes might be derived from their pre-existing expression plasticity in the ancestor. Sex-biased genes had also been shown to evolve rapidly in amino acid sequences, possibly by experiencing stronger positive selection driving higher d_N/d_S ratio ([71], Ellegren & Parsch, 2007). We hypothesized that gene13359 was a differentially expressed gene during the development of both sexes in the nonsocial ancestor of ants. During the transition to eusociality it was under selection pressure so that it accumulated many nonsynonymous substitutions in coding sequence, which made changes in protein functions and its surrounding GRNs that was beneficial to the evolution of worker caste. As a result, it acquired caste-biased expression in eusocial insects from its pre-existing sex-biased expression in nonsocial ancestors. So far only the transcriptomic data from two Myrmicinae ants, *Monomorium pharaonis* and *Acromyrmex echinator*, was available to show the expression of gene13359 across development in gyne, male and worker. To decipher whether gene13359 has sex-specific expression in ancestors and sisters of formicoid ants like Myrmicinae, collecting data from wasps and poneroid ants like Ponerinae is worthy to be done.

The higher expression level of gene13359 in gyne compared to worker in *Monomorium pharaonis* was not only attributed to the abdomen, but also the head and the thorax

Our results showed that during young pupal stage, the expression levels of gene13359 in head, thorax and abdomen in gyne were significantly higher compared to worker. Among different body parts, the expression level was highest in abdomen and lowest in head. One big difference between gyne and worker in ants was that worker lacked ovary or its function was inhibited, while gyne maintained fully functional ovary, thus many ovary-related genes responsible for egg production such as *vitellogenin* and *insulin-like peptide (ILP)* were identified to differentially expressed between gyne and worker or reproductive and non-reproductive individuals in some queenless ants ([72], Araki et al., 2020). In spite of this, caste differentiation should not simply be driven by the presence of ovaries, instead, it was more likely to be driven by the overall female reproductive physiology that is almost conserved across insects, although abdominal tissues had more caste-biased genes than head and thorax ([15], Warner et al., 2019). Our results suggested that gene13359 had general female-related physiological functions instead of only ovary-related reproductive functions because it had significantly higher expression level in gyne than in worker across the whole body instead of being restricted in abdomen. In situ hybridization (ISH) needed to be performed to determine the localization of gene13359 at the tissue-specific level.

Materials and methods

Colony setup and maintenance

The ant species *Monomorium pharaonis* was used in this study. All stock colonies with known pedigree were kept at Centre for Social Evolution at University of Copenhagen in a climate chamber at 26 ± 2 °C and 50 % humidity. The main colonies were kept in 21x17x15 cm fluon-coated plastic boxes since 2004 and are thus mature colonies with high genetic relatedness. Colonies were fed twice a week with frozen crickets (*Acheta domesticus*) and sucrose agar. Water was provided in 50 ml falcon tubes sealed with a plug of cotton and was exchanged once a week. The cotton plug should not be too small because a leak in the plug could flood the nest box. Two *M. pharaonis* main colonies were used in this study (colony ID: 4139c and H59).

Queenless (Q⁻) sub-colonies were created from the main colonies to produce sexual (gyne and male) broods used for microinjection since they did not naturally exist in normal queen-existed main colonies. Around 5 ml random collection of workers, broods of different stages and queens were transferred from the main colony into 15 cm fluon-coated petri dishes. All queens were immediately removed from the sub-colonies to enable rearing of sexuals from the present broods (Tay et al., 2014). Sub-colonies were maintained in the same way as main colonies except that water was provided in a smaller cylindrical tube fastened by a paper clip. Normally 2nd sexual larvae appeared around the ninth day after queens had been removed.

Sucrose agar diet recipe (12.5% sucrose and 1.25% agar g/ml):

1200 g of sucrose (normal cooking sucrose) was added to 3000 ml of water (in a pot larger than 5400 ml). Heated it on a hotplate while stirring until the sucrose was dissolved in the water.

Took a separate pot and added 120 g of agar (Sigma) to 2400 ml of water. Thoroughly mixed the two ingredients. The agar solution was added to the hot sucrose solution and

was heated on hotplate set at 265°C under constant stir until the agar was dissolved and the solution had a clear surface.

Took a separate pot (larger than 9600 ml) and mixed 4200 ml of water with 25.92 g of vitamin mixture (Vanderzant Vitamin Mixture for insects). Stirred it until the vitamin was dissolved into the water.

The hot sucrose/agar solution was added slowly to the cold vitamin solution while stirring constantly. Stirred for a while without heating.

Water was added to reach the final volume of 9600 ml. The solution was then poured into the plates (ca 150 plates of 9 cm).

The plates were stored at -20°C after the solution was cooled down to room temperature.

Sample collection

Different developmental stages of workers, gynes and males were collected. Stages of larvae were estimated based on morphological characters and body length (**Figure 1 in General Introduction**), while stages of pupae were determined through the cuticle color using a color scale bar obtained from a test set of age-controlled pupae (**Figure 2 in General Introduction**). Since pupae enclosed into adults at the twelfth day after pupation, I defined early pupae as being 0-2 days old and old pupae as being 10-12 days old.

While workers and sexuals (males and gynes) could be differentiated by morphological traits from the 2nd larval stage, males and gynes could not be distinguished morphologically until pupal stage. To separate males and gynes in 2nd larval, 3rd larval and prepupal stage, DNA was extracted from each individual sample during RNA extraction, and then microsatellite genotyping was performed to identify whether a sexual individual was haploid (male) or diploid (female).

RNA and DNA extractions for each sample

The RNA from each individual sample was extracted using RNeasy Plus Micro Kit

(catalogue number: 74034, Qiagen, adapted protocol). During RNA extractions, gDNA Eliminator spin columns were kept for DNA extraction. The RNA concentration was determined through Qubit RNA HS Assay, and 1 μ l of extracted DNA was loaded onto the Nanodrop 1000 spectrophotometer (Thermo Scientific) to determine the concentration. The average RNA concentration of each sexual prepupa and pupa (around 40 ng/ μ l) was higher than each larva (around 10 ng/ μ l), while the average DNA concentration of each sexual sample is around 20 ng/ μ l. The average RNA concentration of each worker sample was lower than half of each sexual sample. DNA was not extracted from worker samples because worker caste could be morphologically distinguished from sexual caste since second larval stage.

Microsatellite genotyping

Five highly polymorphic nuclear microsatellite loci: Mp4, Mp8, Mph2, Mph9 and Mph23 ([73], Schmidt, 2010) were used for microsatellite genotyping to identify whether an individual was haploid (male) or diploid (female). Primer sequences for amplifying the nuclear microsatellite loci were listed in **Supplementary Table 2.1**.

1 μ l each of the forward and reverse primers (10 μ M) were run in separate PCR reactions in 10 μ l reaction volumes with 1 μ l extracted DNA as template, using 5 μ l VWR Red Taq DNA Polymerase Master Mix. The samples were run with a T100 Thermal Cycler (Bio-Rad) with the PCR reaction conditions consisted of an initial denaturing step of 95°C for 10 min, followed by 30 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec, and finally an extension step at 72°C for 10 min.

To verify and visualize the PCR products, samples were loaded on a 2% agarose gel. The 2% agarose was prepared in TAE buffer (0.04 M tris, 0.5 mM EDTA) and heated in a microwave until the solution was clear. 2 μ l of the PhiX174-HaeIII marker (catalogue number: N3026L, NEB) was used as ladder. 2 μ l of PCR product was mixed with 1 μ l GelRed -contained loading buffer (0.3 μ l 10,000X GelRed (catalogue number:

41003, Biotium) diluted in 1 ml loading buffer) and run in MultiSUB Choice wells (CS Cleaver Scientific Ltd) at 90 V for 25 min in TAE buffer. PCR products were visualized by GENi Bio Imaging Gel Documentation System (Syngene). In this section, PCR products were visualized to confirm that the microsatellite loci were amplified correctly and were qualified to be used in the next step (**Supplementary Figure 2.3**).

After gel visualization, 1 μ l of each PCR products was mixed with 6.75 μ l HiDi-Formamide and 0.25 μ l GeneScan LIZ500 in 10 μ l reaction volumes. The samples were run on 5% polyacrylamide gels using an ABI PRISM® 377 DNA automated sequencer. The loci were multiplexed on the ABI. GeneMapper™ Software 5 were used to score the alleles with the help from Rasmus Stenbak Larsen. If all the five loci turned out to be haploid, then the individual was considered as male; otherwise if at least one locus were diploid, it was a female.

Reverse Transcription and Droplet Digital PCR (ddPCR)

Reverse Transcription

iScript cDNA Synthesis Kit (catalogue number: 1708890, BIO-RAD) was used to reverse transcribed RNA into cDNA. The 20 μ l reaction system was made up of 4 μ l iScript Reaction Mix, extracted RNA (around 100 ng) and 1 μ l iScript Reverse Transcriptase. The samples were run with a T100 Thermal Cycler (Bio-Rad) with the PCR reaction conditions consisted of an initial priming step of 25°C for 5 min, followed by a reverse transcription step of 46°C for 20 min, and finally an inactivation step of 95°C for 1min. cDNA was stored at -20°C.

ddPCR primer design

ddPCR primers were designed with Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) and NCBI Primer designing tool (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>). The PCR product size ranged from 70 bp to 200 bp. The primer melting temperatures (T_m) ranged from

57.0°C to 63.0°C. Optimal T_m was set as 60.0°C and the maximum T_m difference was set as 3°C. Organism was set as *Monomorium pharaonis* (NCBI taxonomy id: 307658).

ddPCR primer validation

Primers were synthesized by TAG Copenhagen A/S Manufacturer in Denmark. The 20 µl PCR reaction system was made up of 10 µl VWR Red Taq DNA Polymerase Master Mix, 1 µl cDNA as template, 1 µl each of the forward and reverse primers (10 µM) and 3 µl ddH₂O. The PCR program consisted of an initial denaturing step of 95°C for 5 min, followed by 35 cycles of 95°C for 10 sec, 55°C for 20 sec, and 72°C for 30 sec, and finally an extension step at 72°C for 7 min. Products were loaded on a 2% agarose gel and were verified and visualized in the same way as the Section for microsatellite genotyping. The primer was qualified to be used in ddPCR if the band on the gel was bright, the size was the same as expected and no other band was visualized, suggesting the high efficiency and specificity of the qualified primers (**Supplementary Figure 2.4**). Sequences of the qualified ddPCR primers were listed in **Supplementary Table 2.2**.

ddPCR

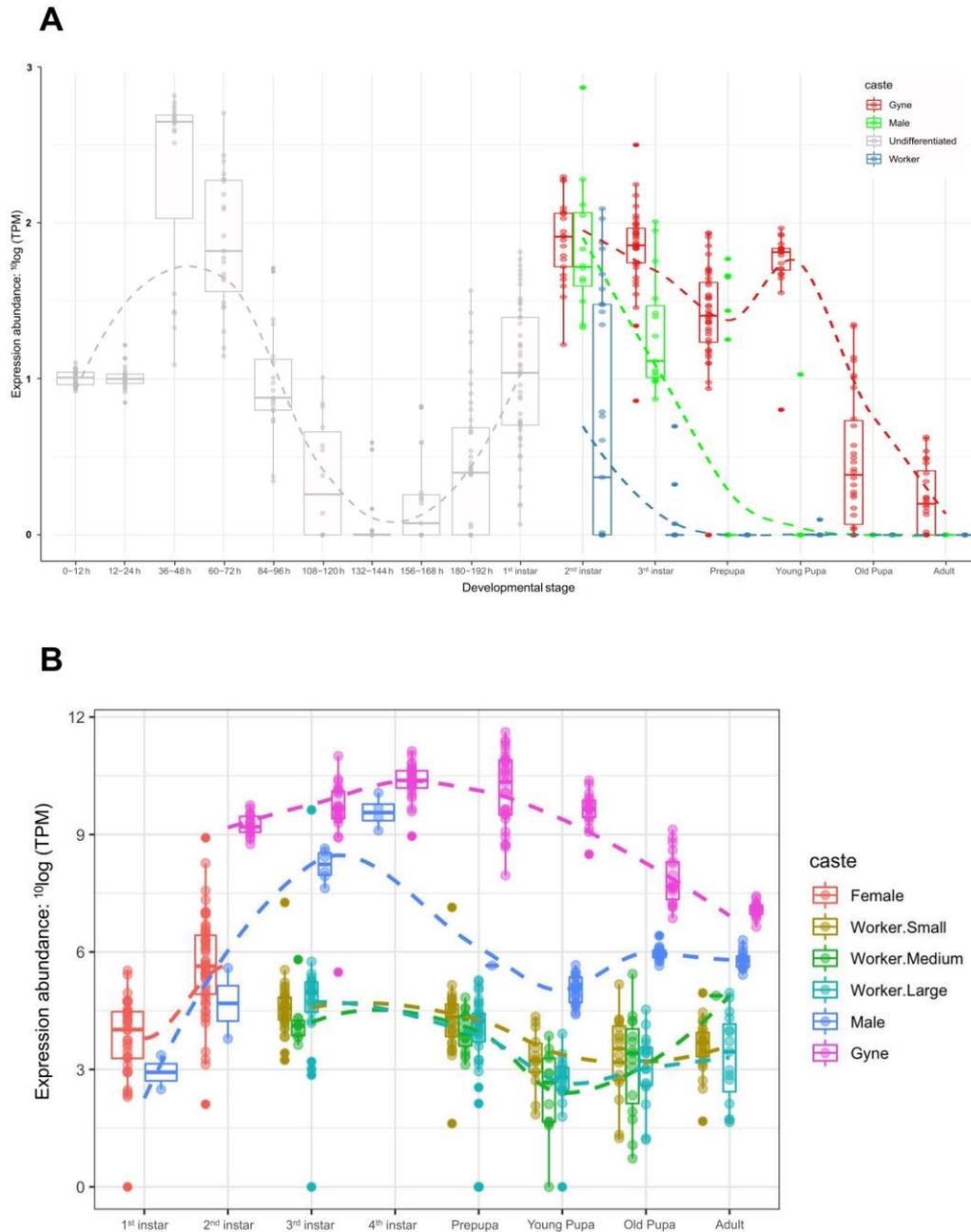
Gene expression levels were determined with the QX200 ddPCR system (Bio-Rad). For each cDNA sample, 11 µl EvaGreen supermix, 1 µl cDNA, 0.22 µl forward primer, 0.22 µl reverse primer and 9.56 µl ddH₂O were mixed for a total volume of 22 µl. Then 20 µl of each reaction mix was added to a DG8 Cartridge (Bio-Rad) sample well, followed by 70 µl of QX200 Droplet Generation Oil for EvaGreen into the oil wells, covered with the Droplet generator DG8 gasket (Bio-Rad) and run in the QX200 Droplet Generator (Bio-Rad). A non-template control, using ddH₂O, was made for each primer pair. The samples, now in droplet form with approximately 40 µl of each sample, was transferred to a ddPCR 96-well PCR plate (Bio-Rad), covered with a pierceable foil heat seal (Bio-Rad) and sealed at 180 °C for 0.5 sec with PX1™ PCR Plate Sealer (Bio-Rad). Then PCR was performed with a T100 Thermal Cycler (Bio-Rad) at 95°C

for 5 min (ramp rate at 2°C/sec.), followed by 40 cycles of 95°C for 30 sec. (ramp rate at 2°C/sec.) and 60°C for 1 min. Then 4°C for 5 min (ramp rate at 2°C/sec.), 90°C for 5 min (ramp rate at 2°C/sec.) and then kept at 4°C (ramp rate at 2°C/sec). Following the PCR, the samples were transferred to the ddPCR Droplet Reader (Bio-Rad) for gene expression quantification.

Droplet fluorescence data was initially analyzed with QuantaSoft Analysis software (Bio-Rad), and a manually adjusted threshold was set to get the concentration values in the unit of copies/μl for the droplets. The expression level of the candidate gene in the sample was normalized by taking the logarithmic value of the ratio between the reference gene transcript concentration of the same sample and the candidate gene transcript concentration.

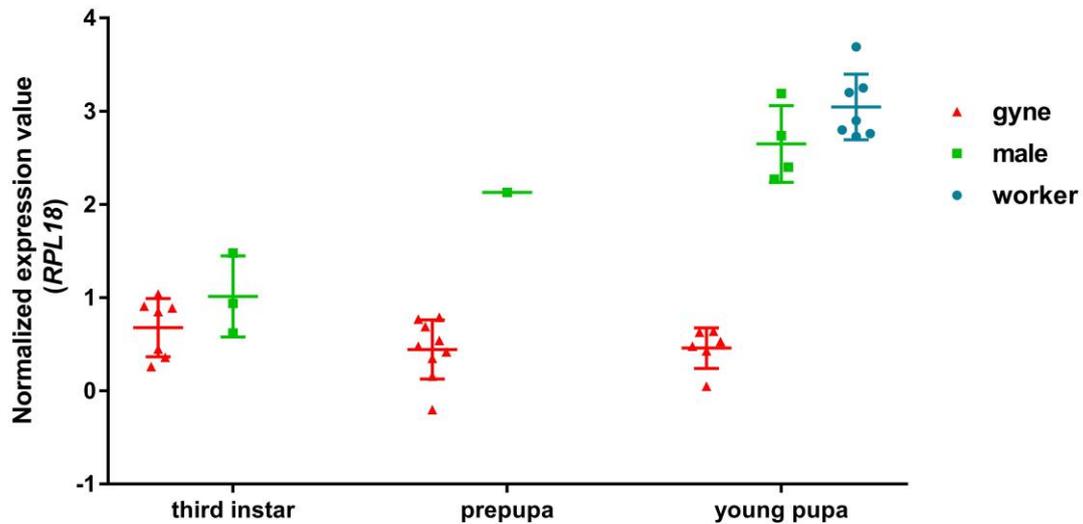
$$\text{Normalized expression value} = \log_{10} \left(\frac{\text{Reference gene transcript concentration}}{\text{Candidate gene transcript concentration}} \right)$$

Supplementary materials



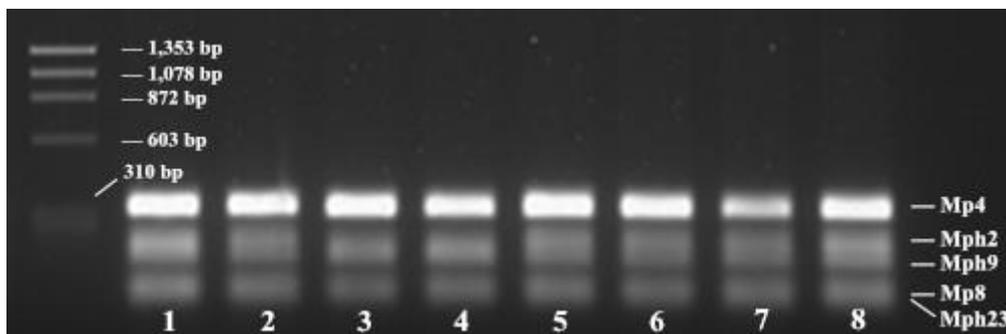
Supplementary Figure 2.1. Expression patterns of gene13359 in gyne, male and worker across developmental stages based on transcriptomic data

(A) Expression patterns of gene13359 from embryo to adult in *Monomorium pharaonis*. From embryonic stage to 1st larval stage, individuals from different castes couldn't be morphologically differentiated. (B) Expression patterns of the homolog of gene13359 from 1st larval stage to adult in *Acromyrmex echinator*. The profiles were expressed in the scale of $^{10}\log$ Transcripts Per Kilobase Million and were normalized by transcriptome differences between castes. © Bitao Qiu



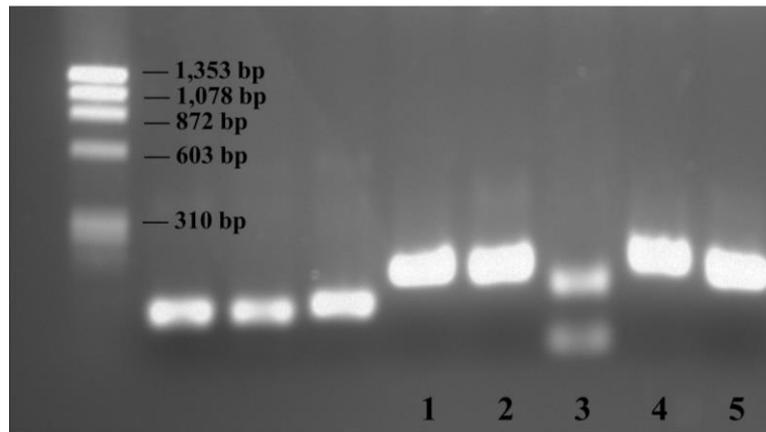
Supplementary Figure 2.2. Transcriptional level of gene13359 across different castes and developmental stages using *RPL18* as a reference gene

Each dot represented one individual except that for worker pupae, each dot represented four individuals. The normalized expression value of gene13359 was calculated in the same way as in **Figure 2.2**



Supplementary Figure 2.3. PCR products for microsatellite genotyping visualized by 2% agarose gel electrophoresis

DNA extracted from sexual individual sample was used as template to amplify the five highly polymorphic nuclear microsatellite loci: Mp4, Mp8, Mph2, Mph9 and Mph23. Lane 1-8: eight different sexual individuals to be identified for their genotype. The leftmost lane was PhiX174-HaeIII marker (size range: 72 bp to 1,353 bp).



Supplementary Figure 2.4. PCR products amplified by ddPCR primers and visualized by 2% agarose gel electrophoresis

Total RNA extracted from young gyne pupa was reverse transcribed into cDNA which was then used as template for PCR. The primer was qualified to be used in ddPCR if the band on the gel was bright, the size was the same as expected and no other band was visualized, suggesting the high efficiency and specificity of the qualified primers. Primers used from lane 1 to lane 5: NADH-1 (199 bp), RPL18-2 (189 bp), gene4 (123 bp), gene5 (193 bp) and gene6 (172 bp). Primer gene4 was not qualified to be used because it had a weak amplification of the target product (faint band) and a non-specific band below was visualized.

Supplementary Table 2.1. The microsatellite loci analyzed in the study

Locus	Primer sequences (5' - 3')	Repeat motif	No. of alleles	Product length (bp)
Mp4	F: CGGCAAATGCACAAACATTA	(GT) ₁₇	7	279-303
	R: CGTGAGGTCAAAAGTTCCGTA			
Mp8	F: TTGTATGGGAAAGGCGAAAG	(GA) ₆ GG(GA) ₁₂	10	97-119
	R: TAAGCCTCCTTGCACAATCC			
Mph2	F: CGCACTAAAGCGCGAAAG	(TC) ₁₃	22	190-236
	R: ATCGTCGGTGTCTCATTTC			
Mph9	F: TTCCTGCATCAAAAGTGCTG	(CT) ₃ CC(CT) ₁₇	9	161-203
	R: ACAGTTGGCCGCATAAATTG			
Mph23	F: CGCGAGGAGACGAACTACC	(GA) ₂₇	18	75-115
	R: TGCTCGTTTCTCGACGTATG			

Supplementary Table 2.2. The qualified ddPCR primers in the study

Target gene	Primer	Sequence (5'→3')	Product length (bp)
gene13359	gene5-F	GTTTCGTCTCCCTTCCAGCA	193
	gene5-R	GCAAGCCCAACCAAACGAAA	
	gene6-F	ATTCGCGGTGATCTGTGGAG	172
	gene6-R	TCGTGCTCGAATCAGGGATG	
	gene9-F	TATTCTACCTTGGTGCACCG	179
	gene9-R	CCCACAATAGTGTACGTCAC	
NADH	NADH-1-F	AGGTATTTTCGAGAACCAGCG	199
	NADH-1-R	GTCCTTCGAGATCCATCTGC	
RPL18	RPL18-2-F	CTATCGATTCCTGGCGAGAC	189
	RPL18-2-R	TCAAAGATCCTGGCGTCATC	

Chapter 3

Study of the biological function of gene13359 through RNAi

Results

Selection of M3rd and L3rd sexual larvae for microinjection

M. pharaonis was considered to have blastogenic caste determination system ([74], Bourke & Franks, 1995; [75], Petersen-Braun, 1977), in other words, the caste fate in *M. pharaonis* might be determined in early embryo stage. If a gyne-biased expression gene involved in caste differentiation process could be knocked down at an earlier developmental stage of gyne, the more likely the individual would be transferred to the worker developmental pathway and finally exhibited some characteristics of worker. Otherwise, if the gene was knocked down at a late developmental stage, then the developmental pathway was more canalized, and the individual was more difficult to be changed into another caste. Since gene13359 showed significantly expression difference from 2nd instar between gyne and worker, it would be better to knock it down in gyne individual before 2nd larval stage so that the gene regulatory network (GRN) would be transformed into worker's to a large extent and some obvious worker-like phenotypic traits might be observed. After trying hundreds of microinjections with ddH₂O into embryos, 1st instar and 2nd sexual instar, it was given up because the death rate caused by injection was as high as 90% no matter how different needle puller programs and injection volumes were tried. Finally, proper conditions including the needle puller program, injection site, time and pressure were explored for M3rd and L3rd sexual larvae, and the death rate caused by injection was proven to be lower than 50%, which was acceptable for the project. Thus, the M3rd and L3rd larval stages were selected for microinjection, although they were located at the downstream of developmental pathway which might be too late for the caste fate to change and

probably more difficult for the phenotype to be identified when the gene was knocked down.

siRNA complexed with transfection reagent was no longer used in RNAi assays

The synthetic small-interfering RNA (siRNA) targeting at the coding sequence was first used to functionally knock down gene13359. Because short RNA fragments were inefficiently taken in by insect cells ([76], Miller et al., 2012), the transfection reagent *in vivo*-jetPEI was used to help deliver siRNA into cells. siRNA oligo gene-1 targeted at the CDS of gene13359, while siRNA oligo yfp-3 targeted at the *yfp* (*yellow fluorescence protein*) gene, which was used as a control. At first, the final concentration of 2.5 μ M siRNA complexed with *in vivo*-jetPEI was injected into M3rd and L3rd sexual larvae. The concentration was selected in reference to another study conducted in *Harpegnathos saltator* ([18], Gospocic et al., 2017), but ddPCR analysis using primer gene9 showed that gene13359 could not be significantly knocked down in *M. pharaonis* at the third or seventh day after injection compared to the control group. Then the concentration was increased to 25 μ M and around 500 individuals were injected including the control group. The transcriptional level of gene13359 in around 50% of the survived individuals could be knocked down to 0.25 compared to control group at the third day after injection. One weird pupa with severe developmental defect was observed in treatment group (**Figure 3.1 B2**). This kind of phenotype had never been observed in control group or in the normal colony. In spite of this, siRNA complexed with the transfection reagent *in vivo*-jetPEI was no longer used in this project due to the high lethality in larvae. The survival rates in both the treatment group and the control group were about one-ninth, which was much lower than injection with ddH₂O. It should be noted that in the study of Gospocic et al., the siRNA- *in vivo*-jetPEI complex was injected into the head of adult ants. According to the manufacturer's instructions, the essence of the complex was 5% glucose solution, which might be fine for the use in adult individuals, but for larvae it was hypothesized to cause a high lethality.

Synthesis and qualification of different dsRNAs to functionally knock down gene13359

Double-stranded RNAs (dsRNAs) with the length between 500 bp to 600 bp were synthesized through *in vitro* transcription to knock down gene13359 in M3rd and L3rd larval stage. The synthesized dsRNA was dissolved in 1X injection buffer for microinjection. Two pieces of dsRNAs, gene5 dsRNA targeting at the far end of 3' UTR of the transcript and gene6 dsRNA targeting at nearer 3'UTR (**Figure 2.1A**) were first synthesized. *gfp* dsRNA targeting at the *gfp* (*green fluorescence protein*) gene was synthesized as control. The concentrations of 1 $\mu\text{g}/\mu\text{l}$ and 5 $\mu\text{g}/\mu\text{l}$ were selected to use. According to ddPCR analysis using primer gene9, only gene6 dsRNA in the concentration of 5 $\mu\text{g}/\mu\text{l}$ was detected to knock down gene13359 to a slightly lower level compared to control group in only a few individuals at the fifth day after injection (data not shown), but no phenotype had been observed because all survived individuals in treatment group developed normally as in control group. Some abnormal phenotypic traits, like intensive fat bodies throughout larvae's internal body cavities, different colors of meconium in the gut of larvae, faster developmental progress to prepupal stage, smaller body length of prepupae and pupae, deformed wings in young adults after eclosion and so on, used to be considered as phenotypes resulted from the knockdown of the target gene, but later on similar phenotypes were also observed in control group, suggesting that these phenotypes were just the variation among individuals in a normal colony or common results caused by the physical damage from injection. One good thing was that after injecting more than 600 M3rd and L3rd sexual larvae, it was concluded that the survival rate of injection with dsRNA dissolved in 1X injection buffer was nearly the same of the injection with ddH₂O, which was acceptable, and there was no obvious difference between treatment group and control group.

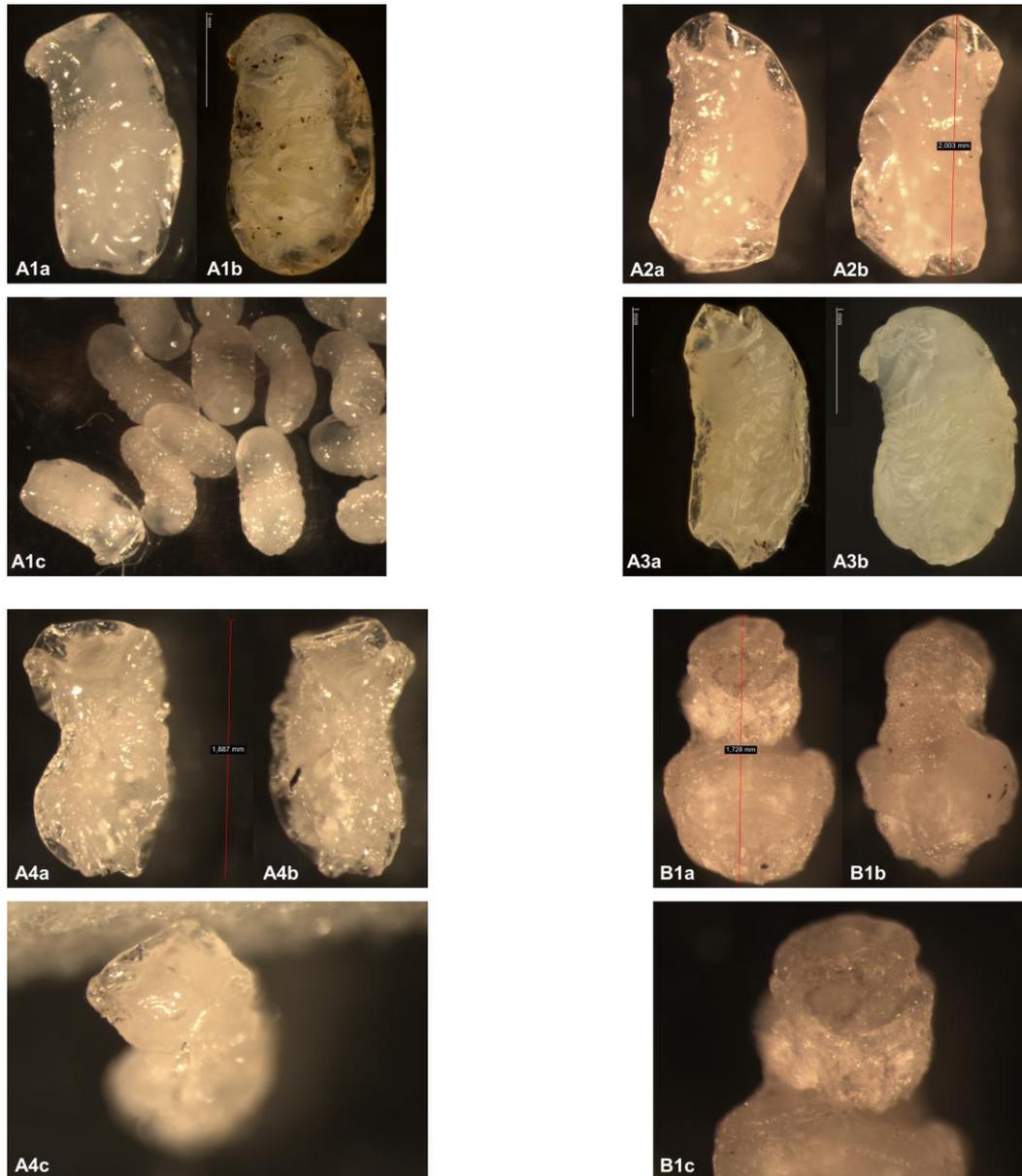
Because the regions targeted by gene5 dsRNA and gene6 dsRNA were located on the far end and near end of 3' UTR, which had much fewer transcript copies than CDS (**Figure 2.1B**), the copies of CDS might still remain at a relatively high level so that the

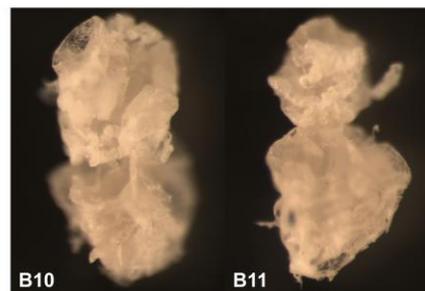
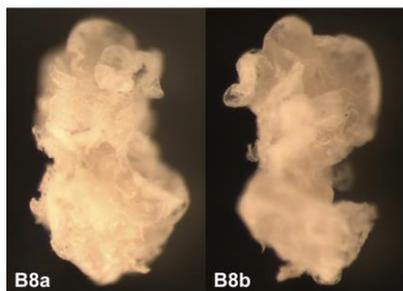
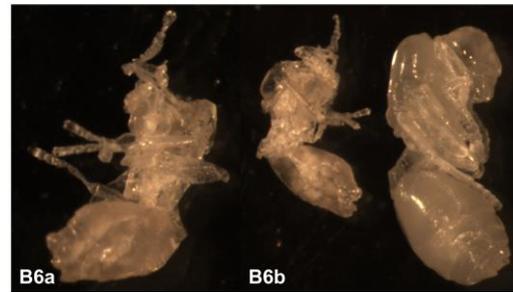
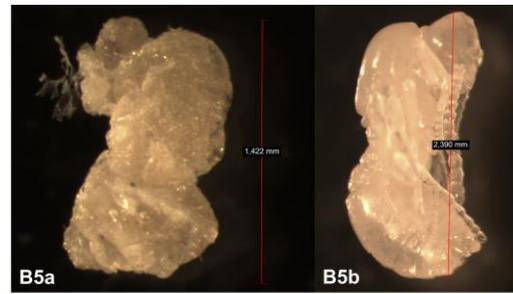
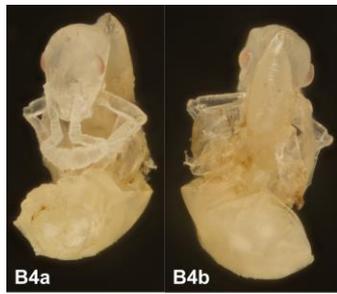
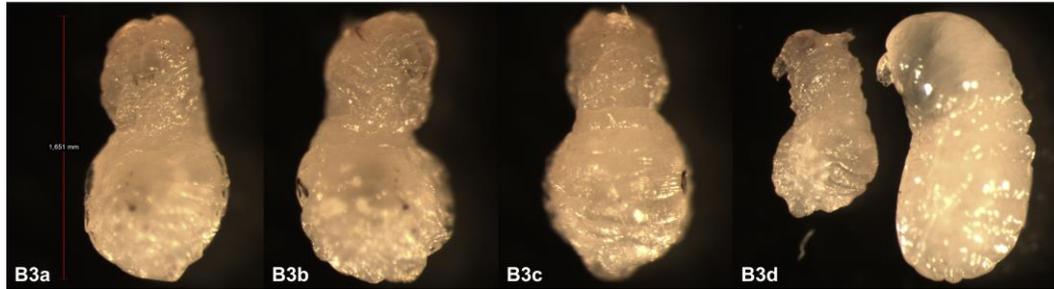
functional part of gene13359 transcript failed to be knocked down significantly in most of the individuals in the treatment group. Therefore, another piece of dsRNA targeting at the whole region of CDS, named gene7 dsRNA (**Figure 2.1A**), was synthesized to functionally knock down gene13359. It should be noted that the ddPCR product amplified by primer gene9 was not completely overlapping with the region targeted by gene7 dsRNA, so it was impossible for primer gene9 to amplify the region on the synthetic gene7 dsRNA which had been injected *in vivo*, thus primer gene9 could still detect the actual level of gene13359 transcript.

Knockdown of gene13359 by gene7 dsRNA and observation of phenotypes

gene7 dsRNA dissolved in 1X injection buffer was injected into M3rd and L3rd sexual larvae in the concentration of 5 µg/µl. Same concentration of *gfp* dsRNA was injected as control. In treatment group, several individuals for both genders (gynes and males) were observed with severe developmental defects in prepupal (**Figure 3.1A**) and pupal (**Figure 3.1B**) stages in multiple independent assays, which had never been observed in control group injected with *gfp* dsRNA, where all individuals had a normal developmental progress as in the wildtype colony. Prepupa exhibited wrinkled skin and the peripheral body cavity seemed to be hollowed and filled with liquid, thus it was softer and more likely to change the shape of surface when touched. The pupa-like outline could be observed inside the body cavity, but the individual couldn't develop further into pupa. The prepupa in control group had a full body cavity and was resistant to change the shape of surface. All pupae suffered from severe developmental defects and the size was much smaller than normal pupae. These individuals would die soon, and the bodies would be dehydrated and become a bit yellow. From B1 to B3, the three individuals were at a status that was between prepupal stage and pupal stage. The outline of body resembled prepupa because it didn't exhibit the segmentations among head, thorax and abdomen; some organs like antennae, legs and eyes appeared but were not fully developed as normal pupa did. This kind of status might exist in a short time during pupation, but it was very unlikely that the individual had such small size and

died soon. From B4 to B11, both gyne and male pupae were observed with smaller size and deformed head, thorax and abdomen. Various organs like antennae, legs and wings were also deformed. Lower concentration of 3.5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA had been tried and similar lethal phenotypes occurred in prepupae and pupae (**Figure 3.1C**).





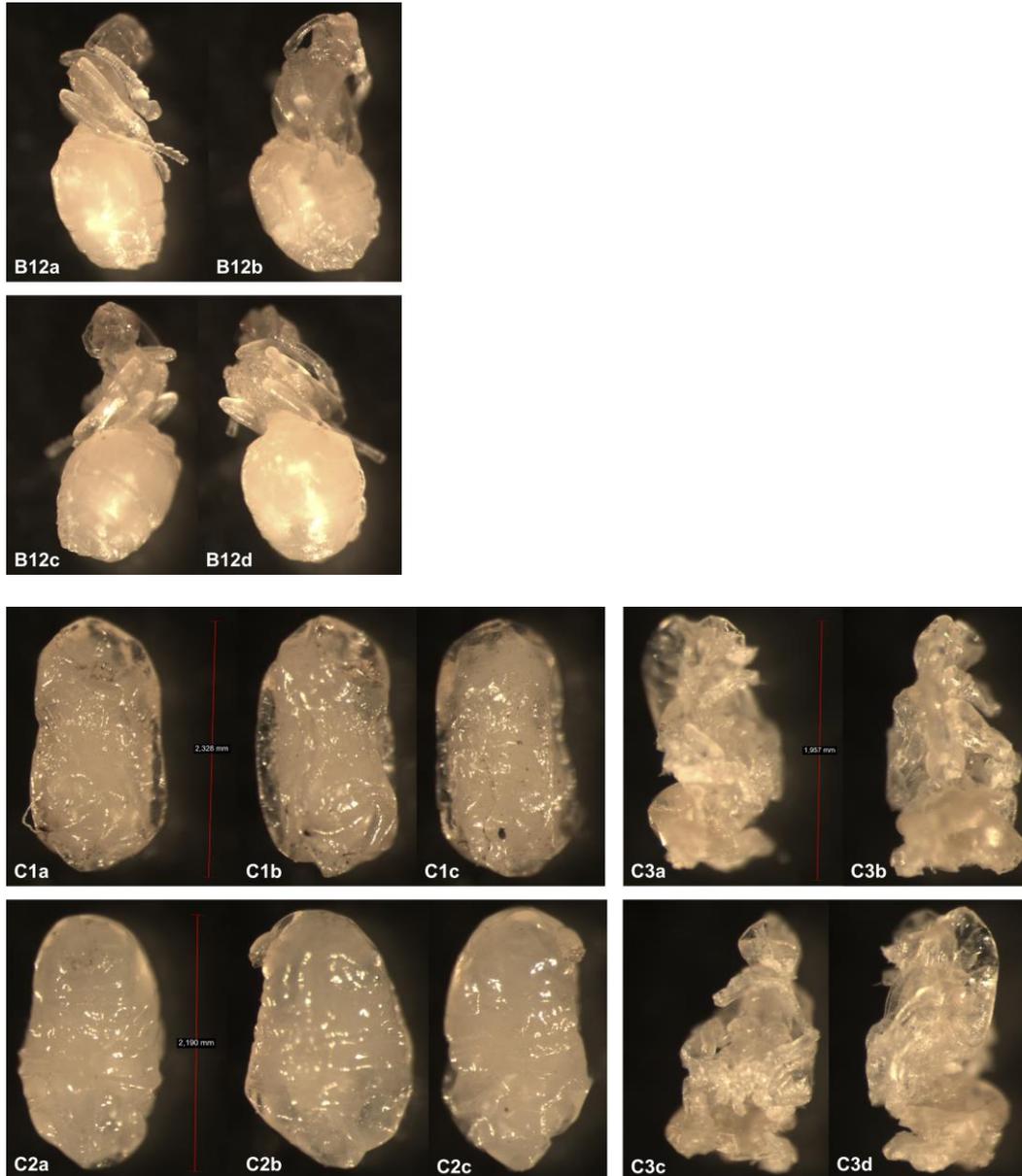


Figure 3.1. Lethal phenotypes observed in multiple individuals in treatment groups

(A) Prepupae injected with 5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA during larval stages. **1a-1c**: a gyne prepupa with the body length of 2.52 mm. **1a**: 12 days after injection; **1b**: 15 days after injection; **1c**: 12 days after injection, being together with other prepupae in the same treatment group without any abnormal phenotype. **2a-2b**: a male prepupa at 14 days after injection with the body length of 2.00 mm. **3a**: a male prepupa at 15 days after injection with the body length of 2.44 mm. A red eye can be observed from lateral side. **3b**: a sexual prepupa from the control group with the body length of 2.42 mm. **4a-4c**: a male prepupa at 5 days after injection with the body length of 1.89 mm. A pair of red eyes could be observed in **4c** which was recorded from the head of individual.

(B) Pupae injected with 5 µg/µl gene7 dsRNA during larval stages. **1a-1c**: a female pupa at 11 days after injection with the body length of 1.73 mm. **1a**: ventral side; **1b**: dorsal side; **1c**: zoom in on the head and thorax of **1a**. **2a-2c**: a morphological gyne pupa at 14 days after injection with 25 µM siRNA gene-1. The body length was 1.54 mm. **2a**: lateral side; **2b**: ventral side; **2c**: lateral-dorsal side. **3a-3d**: a morphological male pupa at 7 days after injection with the body length of 1.65 mm. **3a**: ventral side; **3b**: lateral-ventral side; **3c**: dorsal side; **3d**: comparison with a sexual prepupa from the control group (right one, 2.43 mm in length). **4a-4c**: a gyne pupa at 18 days after injection with the body length of 2.31 mm. **4a**: ventral side; **4b**: dorsal side; **4c**: comparison with a gyne pupa from the control group (left one, 3.19 mm in length). **5a**: a morphological male pupa at 7 days after injection with the body length of 1.42 mm (lateral side). **5b**: a male pupa from the control group with the body length of 2.39 mm (lateral side). **6a-6b**: a sexual pupa at 9 days after injection with the body length of 1.90 mm. **6a**: lateral side; **6b**: comparison with a gyne pupa from the control group (right one, 2.93 mm in length). **7a-7b**: a male pupa at 7 days after injection with the body length of 1.88 mm. **7a**: dorsal side; **7b**: lateral side. **8a-8b**: a gyne pupa at 8 days after injection with the body length of 1.81 mm. **8a**: ventral side; **8b**: lateral side. **9a-9b**: a gyne pupa at 9 days after injection with the body length of 2.20 mm. **9a**: lateral side; **9b**: dorsal side. **10**: a gyne pupa at 8 days after injection with the body length of 1.89 mm (lateral-ventral side). **11**: a male pupa at 8 days after injection with the body length of 1.59 mm (ventral side). **12a-12d**: a morphological gyne pupa at 14 days after injection with the body length of 2.58 mm. **12a**: lateral side; **12b**: ventral side; **12c**: lateral side; **12d**: dorsal side.

(C) Individuals injected with 3.5 µg/µl gene7 dsRNA during larval stages. **1a-1c**: a gyne prepupa at 14 days after injection with the body length of 2.33 mm. **1a**: ventral side; **1b**: lateral side; **1c**: dorsal side. **2a-2c**: a male prepupa at 14 days after injection with the body length of 2.19 mm. **2a**: ventral side; **2b**: lateral side; **2c**: lateral side. **3a-3d**: a gyne pupa at 22 days after injection with the body length of 1.96 mm. **3a**: lateral-ventral side; **3b**: dorsal side; **3c**: ventral side; **3d**: lateral-dorsal side.

To detect whether the expression level of gene13359 in these individuals in **Figure 3.1** was indeed knocked down, some of these individuals were sacrificed for DNA and RNA extractions. DNA was used for microsatellite genotyping to determine gender and RNA was reverse transcribed into cDNA for ddPCR. Data of the control group was the same as data in **Figure 2.2 and Supplementary Figure 2.2**. Six prepupae with two females (A1, C1) and four males (A2, A3a, A4, C2), and eight pupae with six females (B1, B4, B8, B9, B10, C3) and two males (B7, B11) were collected. The ddPCR results showed that for gyne, the levels of gene13359 were significantly knocked down in both prepupae and pupae, but the knockdown effect in pupae was weaker than in prepupae. No significant knockdown effect was observed in both male prepupae and male pupae (**Figure 3.2**). One possible reason for the weaker knockdown effect in pupae might be that the expression level of gene13359 was rescued during pupation. Another possible reason was that strong knockdown effect would make individuals die in prepupal stage, like the phenotypes described in **Figure 3.1A**; weak knockdown effect would allow individuals to survive until pupal stage with severe developmental defect, as described in **Figure 3.1B**. The philosophy was that different levels of gene knockdown effect might result in different interactions in the related gene regulatory networks, which might produce different kinds of phenotypic consequences.

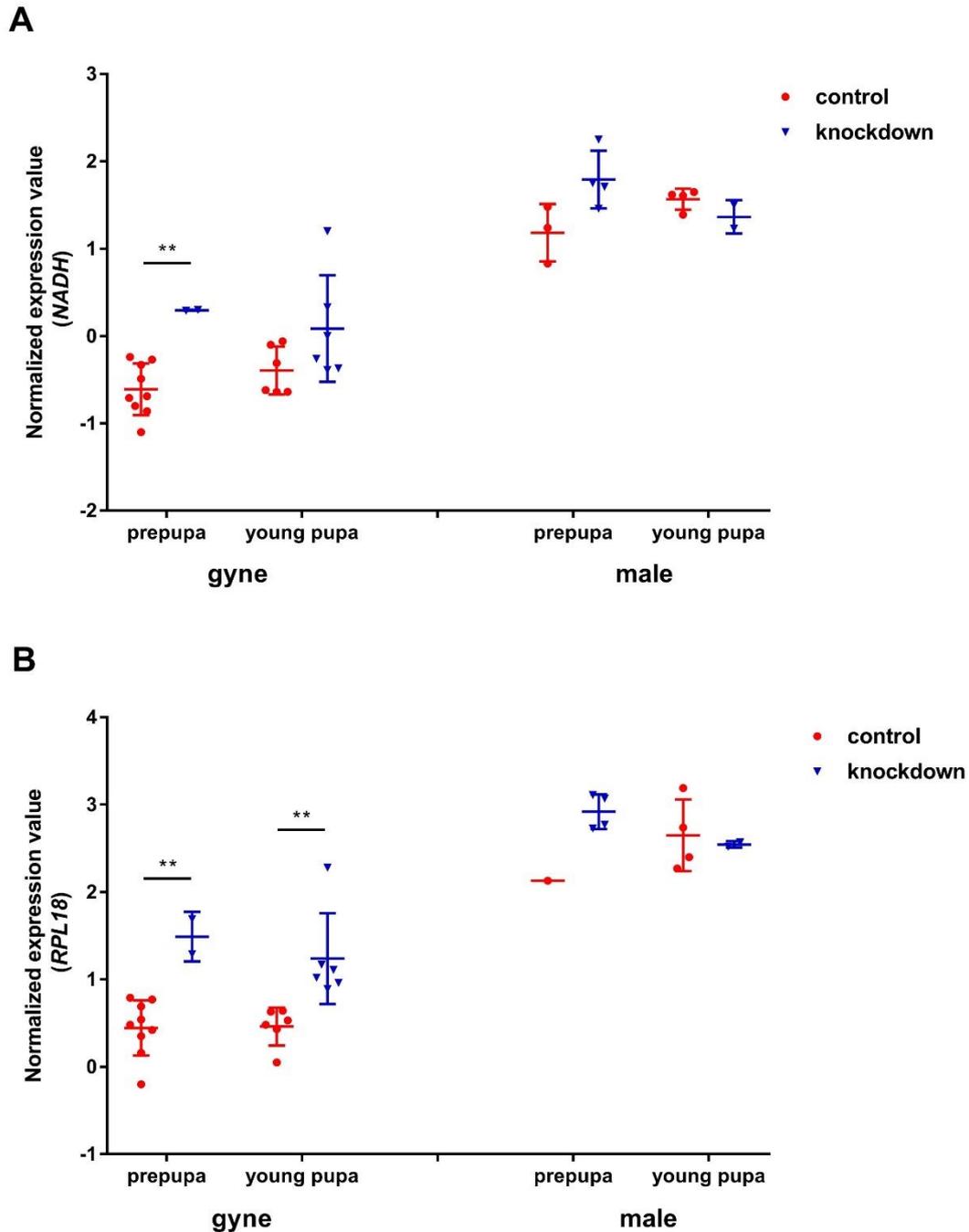


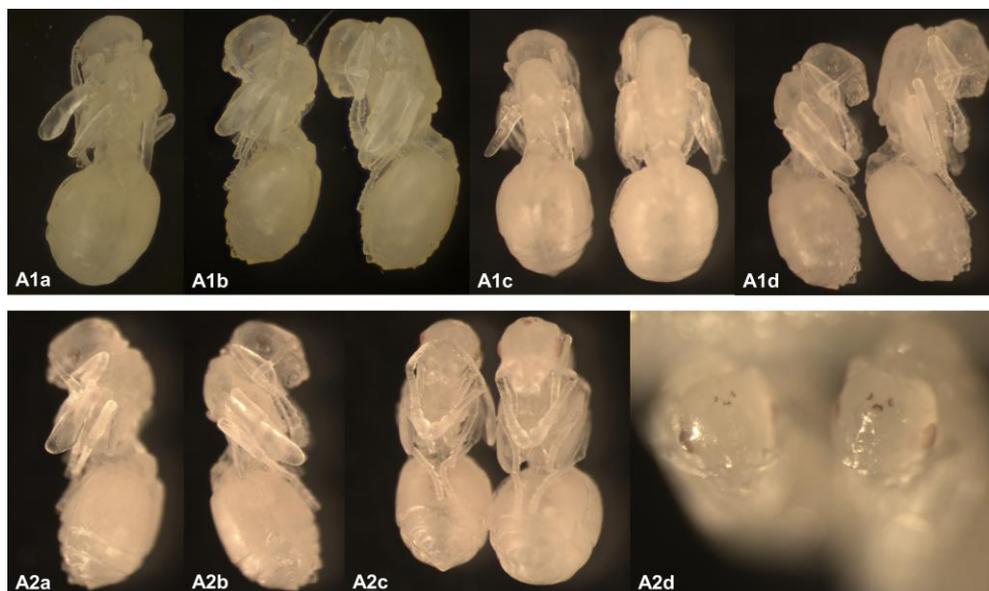
Figure 3.2. Transcriptional level of gene13359 in treatment individuals injected with gene7 dsRNA that exhibited certain phenotypes (blue) and control individuals from normal queenless sub-colonies (red)

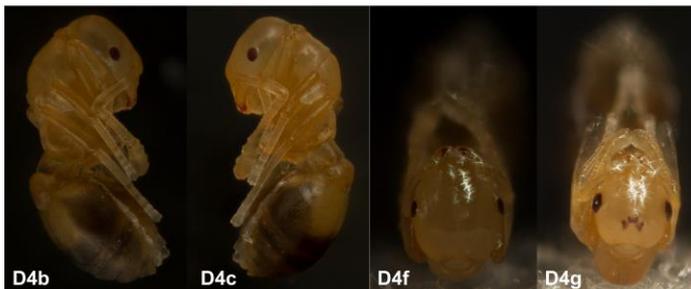
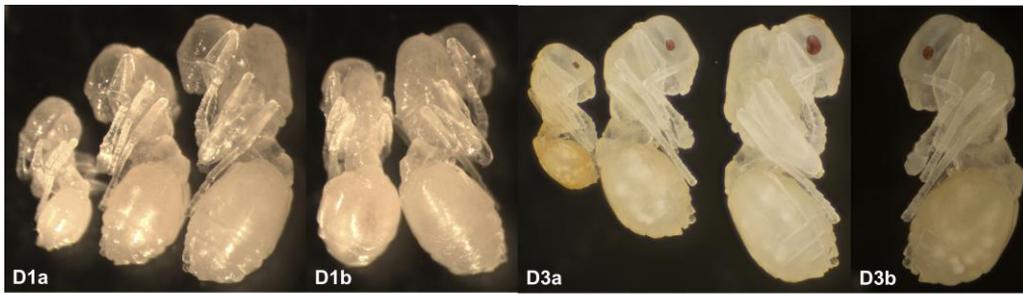
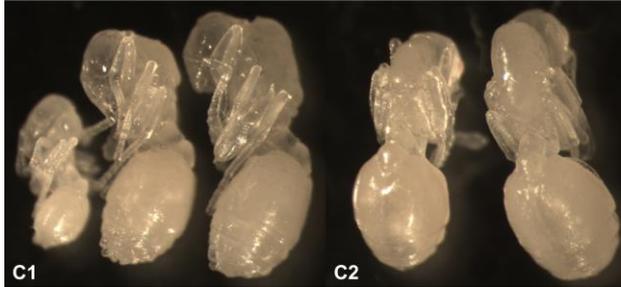
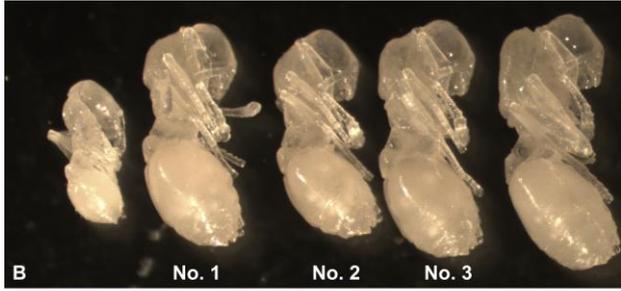
Each dot represented one individual. Unpaired t-test was used on the normalized values to examine if any statistical significance was present between two groups of samples. ** P value<0.01. (A) *NADH* and (B) *RPL18* were used as reference genes.

In addition to the lethal phenotypes described above, four gyne pupae with worker-like traits had been observed in treatment group (**Figure 3.3**). RNA was not extracted from these individuals thus gene13359 levels were not examined by ddPCR. Instead, they developed further into adults. In **panel A**, a gyne pupa appeared at 16 days after injection with 3.5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA. It had a smaller thorax compared to control group and the head was a little deformed. After two days' development, there was no change in the thorax and head, but the ocellus on head appeared to be smaller. The individual was killed by workers and thus disappeared in the colony at the fourth day after pupation (20 days after injection). It might be that workers detected its abnormal developmental traits so that it was no longer be tolerated in the colony. In treatment groups of 5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA, three gyne pupae with worker-like developmental traits were found in total (No.1, No.2 and No.3 in **panel B**). They have a worker-like thorax and smaller wings. The individual No.1 (**panel C**) was killed by workers two days later so its developmental progress could not be further recorded. The individuals No.2 (**panel D**) and No.3 (**panel E**) survived to adult. During the whole pupal stage, the individual No.2 maintained a worker-like thorax, shorter body length, smaller eyes, tiny wings, and no ocellus was found on the head. The individual No.3 also maintained a worker-like thorax, smaller eyes and wings, but ocelli developed. It should be noted that its ocelli were deformed and much smaller compared to normal gyne pupa, suggesting that the development of ocelli in individual No. 3 was suppressed but the effect was not as strong as that in individual No.2, which might be a signal of inter-caste between worker and gyne. After eclosion, the individual No.2 survived for two days and was finally attacked to death by workers from the same colony. During the first day workers were grooming it and trying to accept it as a colony member, but one day later workers became quite aggressive to it. It died soon after one of its antennae was bitten by workers. The individual No.3 survived for only one day because it didn't eclose well. The wings were folded and membrane from pupal stage was still remained on the surface of body. It might be that there were not enough workers in the colony because the process of eclosion demanded the help from workers. In short, the

behaviors of individuals No.2 and No.3 were not sufficiently recorded so that no conclusion could be drawn on whether their behaviors were more like gynes or workers. Because the expression levels of gene13359 in these individuals were not examined, so it was not one-hundred percent sure that such worker-like phenotypes in gyne pupae were caused by the knockdown of gene13359, in other words, it could not be ruled out that these phenotypes occasionally happened under normal conditions. More replicates of these worker-like gyne pupae in treatment groups and evidence from the molecular level were needed.

M3rd and L3rd worker larvae were also injected with 5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA even though gene13359 had a very low expression level in all the developmental stages of worker from 2nd instar. Same concentration of *gfp* dsRNA was injected as control. As expected, no difference had been observed in the developmental progress of the individuals between these two groups.





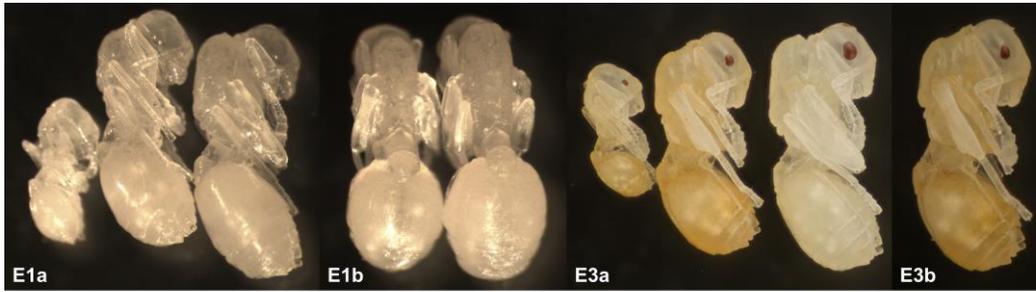




Figure 3.3. Gyne pupae with worker-like developmental traits in treatment groups

(A) A gyne pupa injected with 3.5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA during larval stage. The body length was 2.78 mm. **1a-1d**: 16 days after injection. **1a**: dorsal-lateral side; **1b**: lateral side with comparison (right one: a pupa from control group in the length of 2.94 mm); **1c**: dorsal side with comparison (right one: control); **1d**: lateral side with comparison (right one: control). **2a-2d**: 18 days after injection. **2a**: lateral side; **2b**: lateral side; **2c**: ventral side with comparison (right one: control); **2d**: head with comparison (right one: control). The individual was killed by workers in the colony at the twentieth day after injection.

(B) Three worker-like gyne pupae (No.1, No.2 and No.3) injected with 5 $\mu\text{g}/\mu\text{l}$ gene7 dsRNA during larval stages. No.1: 15 days after injection; No.2 and No.3: 11 days after injection. The leftmost was

a normal young worker pupa and the rightmost was a normal young gyne pupa.

(C) The individual No.1 at 15 days after injection with the body length of 2.54 mm. **C1**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost, 2.90 mm in length). **C2**: dorsal side comparison with a normal gyne pupa (right one).

(D) The developmental progress of individual No.2 with the body length of 2.40 mm. **D1**: the first day after pupation (11 days after injection). **1a**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost, 2.92 mm in length). **1b**: dorsal side comparison with a normal gyne pupa (right one). **D2**: the third day after pupation (13 days after injection). **2a**: lateral side; **2b**: lateral side; **2c**: dorsal side. **D3**: the seventh day after pupation (17 days after injection). **3a**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost); **3b**: lateral side; **3c**: lateral side; **3d**: dorsal side; **3e**: head. **D4**: the tenth day after pupation (20 days after injection). **4a**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost); **4b**: lateral side; **4c**: lateral side; **4d**: ventral side; **4e**: dorsal side; **4f**: head; **4g**: head of a normal gyne pupa. **D5**: two days after eclosion (15 days after pupation and 25 days after injection). **5a**: lateral side comparison with a normal worker adult (right one); **5b**: lateral side; **5c**: lateral side; **5d**: dorsal side; **5e**: head; **5f**: head of a normal queen.

(E) The developmental progress of individual No.3 with the body length of 2.66 mm. **E1**: the first day after pupation (11 days after injection). **1a**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost, 2.87 mm in length). **1b**: dorsal side comparison with a normal gyne pupa (right one). **E2**: the third day after pupation (13 days after injection). **2a**: lateral side; **2b**: dorsal side; **2c**: lateral side. **E3**: the seventh day after pupation (17 days after injection). **3a**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost); **3b**: lateral side; **3c**: lateral side; **3d**: head. **E4**: the tenth day after pupation (20 days after injection). **4a**: lateral side comparison with a normal worker pupa (leftmost) and a normal gyne pupa (rightmost); **4b**: lateral side; **4c**: lateral side; **4d**: ventral side; **4e**: dorsal side; **4f**: head; **4g**: head of a normal gyne pupa. **E5**: one day after eclosion (14 days after pupation and 24 days after injection). **5a**: lateral side; **5b**: dorsal side; **5c**: lateral side; **5d**: head; **5e**: head of a normal queen.

(F) Comparison between a normal worker adult and a normal gyne adult.

Discussion

Results of RNAi assays showed that gene13359 was an essential gene mainly functioning during the pupation of sexual individuals (gynes and males) in *Monomorium pharaonis*. The knockdown of gene13359 in gyne to a certain level during larval stages could result in worker-like developmental traits when it developed into pupa.

It should be noted that not only in gynes, the knockdown of gene13359 in males also caused lethal phenotypes in prepupae or in pupae, indicating that gene13359 functioned in a sexual-worker differentiated manner rather than a gyne-worker differentiated manner, which was reasonable because worker caste was a derived trait evolved from nonsocial ancestors where all individuals were sexual. We hypothesized that in sexual individuals, gene13359 was also an essential gene functioning during pupation in other ant species within formicoid clade, and probably even in wasps and ants within poneroid clade like Ponerinae, which are ancestors and sisters of formicoid ants.

We hypothesized that the regulatory mechanisms underlying the general female physiology in eusocial insects were different from the female physiology in their nonsocial ancestors for the following two experimental results in *Monomorium pharaonis*. First, the knockdown of gene13359 to a certain level in gynes caused some worker-biased developmental traits such as worker-like thorax, smaller eyes, tiny wings and deformed ocelli in pupae, while such phenotypes were never observed in males, indicating that some functions of gene13359 accounting for these phenotypes were not shared by males and were thus derived and gyne-specific. Second, such an essential gene in sexual individuals was inhibited from expression during all developmental stages in workers, which had the same genome as gynes, since the caste fate could be determined morphologically (2nd instar in *Monomorium pharaonis* and 3rd instar in *Acromyrmex echinator*), so other genes must take place of gene13359 to function during pupation in worker caste. It seemed that in *Monomorium pharaonis*, the

physiology of worker was different from gyne, and both are different from its nonsocial ancestors.

According to the results of RNAi assays in this study, we inferred that the lethal phenotypes during pupation reflected the ancestral functions of gene13359 which might be conserved in nonsocial ancestors, and the worker-like developmental traits in gyne pupa reflected the derived functions of gene13359 which were acquired later during the evolution of worker caste, the process of which gene13359 was under selection and rapidly accumulated nonsynonymous substitutions in the coding sequence. The mechanisms that how the ancestral functions and the derived functions of gene13359 interacted with each other remained to be further studied.

More work is needed to unify the observed diverse gene13359-knockdown phenotypes. The proper dsRNA concentration and sexual larval stage for injection need to be determined in order to increase the probability of producing worker-like gyne pupae while avoiding developmental lethal individuals as much as possible. Behavior assays including exploration, brood care, foraging and shelter seeking need to be performed and reproduction ability needs to be examined for the worker-like gynes. Besides RNAi-based knockdown in sexual individual, overexpression of gene13359 in worker during larval stages is another way to study the biological functions of this gene. In addition to *Monomorium pharaonis*, different species such as *Acromyrmex echinator* are also considered to conduct functional experiments. Sequencing the transcriptomes of individuals with gene13359-knockdown and -overexpression can help to uncover its downstream genes involved in the pathway.

Materials and methods

Microinjection setup

The M3rd and L3rd larvae were picked out from Q⁻ sub-colonies using a fine paint brush and were placed onto double-sided tape on a fluon-coated 15 cm petri dish with a 45-degree angle. Needles were made using a Sutter Instrument needle puller (Model P-97) with the program of P=500, HEAT=490, PULL=20, VEL=35, TIME=250. The type of capillaries used for pulling needles was Borosilicate Glass Capillaries 1B100F-4 produced by World Precision Instruments (WPI).

The liquid was loaded from top of the needle using GELoader Tips (catalogue number: 0030001222, Eppendorf) and then the needle was inserted into a Narishige IM-400 microinjection apparatus attached to an Olympus SZX9 dissection microscope fitted with a custom-made x-y movable platform. The tip of the needle was touched gently to the side of a glass capillary tube so that liquid could flow out of the needle.

Microinjection time and flow pressure were controlled by the microinjection apparatus. Normally the Time was set as 0.40 sec, the Pressure was set as 0.200 MPa and the Balance Pressure was turned off, but the parameters might be modified depending on the specific needle prepared. Each larva was injected through its dorsal-lateral side at the midline of the antero-posterior axis with around 0.17 μ l solution (6 μ l for about 36 larvae). An obvious enlargement of the larval body cavity should be observed during injection.

Setup of experimental colonies

After the injection, adult workers from the same Q⁻ sub-colony as injected larvae were added to the petri dish to take care of the injected larvae (in an approximately 3:1 adult: larva ratio). The nest consisted of a piece of thin cardboard (approximately 2x2 cm) with bent edges (**Supplementary Figure 3.1**). After a while, larvae would be gathered under the nest by workers and then the double-sided tape was removed by forceps. The experimental colonies were fed twice a week with frozen crickets (*Acheta domesticus*)

and sucrose agar. Water was provided in a cylindrical tube sealed with cotton and fastened by a paper clip. Most of the death caused by injection happened within three days after microinjection and the dead individuals would just disappear from the colony as they were eaten by workers.

siRNA design

The synthetic small-interfering RNA (siRNA) was used to functionally knock down gene13359. siRNA targeting *yfp* gene was used as control.

The siRNA targets on mRNA sequences were selected according to the following rules from ThermoFisher (<https://www.thermofisher.com/cn/zh/home/references/ambion-tech-support/rnai-sirna/general-articles/-sirna-design-guidelines.html>) and THE RNAi WEB (http://www.rnaiweb.com/RNAi/siRNA_Design/). The siRNA targets should be located 50-100 nt downstream of the start codon (ATG). Target sequences should have a G+C content between 35-60% without stretches of 4 or more nucleotide repeats. The target sequence should be 21 nt long starting with AA dinucleotide, i.e. with the motif of AA(N₁₉) where N is any nucleotide. The 19 nt sequences that shared a certain degree of homology with other genes according to NCBI Nucleotide BLAST (Organism: *Monomorium pharaonis*, taxonomy id: 307658; Program: blastn) should be avoided, i.e. complete match (19/19 matches), one mismatch (18/19 matches), two mismatches (17/19 matches) or three mismatches (16/19 matches).

After selecting target sequences, siRNAs were designed using siDirect version 2.0 (<http://sidirect2.rnai.jp/>). Sequences were listed in **Supplementary Table 3.1**.

The 21 nt siRNA oligos with 3' dTdT cantilevers were synthesized by TAG Copenhagen A/S Manufacturer in Denmark. The initial concentration of siRNA oligos synthesized by company was 100 μM.

siRNA preparation

Since RNA fragments with short length were inefficiently taken in by insect cells, the

transfection reagent *in vivo*-jetPEI (catalogue number: 201-10G, Polyplus Transfection) was used for *in vivo* siRNA delivery into cells with a N/P ratio of 8 following the manufacturer's instructions. First, 10 μ l 100 μ M siRNAs and 10 μ l 10% glucose solution were mixed to 20 μ l total. Next, 1.6 μ l *in vivo*-jetPEI and 8.4 μ l RNase-free water were added to 10 μ l 10% glucose solution. The two solutions were combined, resulting in 40 μ l 25 μ M siRNAs complexed with *in vivo*-jetPEI reagent (the final concentration of glucose was 5%). The complexes should be incubated for 15 min at room temperature before injections. From this time point, the complexes were stable for 4 hours at room temperature and for up to 7 days when stored at 4 °C.

dsRNA preparation

Double-stranded RNA (dsRNA) was synthesized through *in vitro* transcription to functionally knock down gene13359. dsRNA targeting *gfp* gene was used as control. The fragment was between 500 bp to 600 bp, which was of a size range that could efficiently knock down genes in insects ([76], Miller et al., 2012). The fragment was not overlapping with the region targeted by ddPCR primer.

Template preparation

Template for T7 RNA polymerase to synthesize the dsRNA was amplified by PCR using the qualified primers (using the same method as Section for ddPCR primer validation) designed against the target sequence with a T7 promoter sequence at their 5' ends (5'-TAATACGACTCACTATAGGG + target specific sequence-3'). Primers were synthesized by TAG Copenhagen A/S Manufacturer in Denmark. Sequences were listed in **Supplementary Table 3.2**.

The 20 μ l PCR reaction system was made up of 10 μ l VWR Red Taq DNA Polymerase Master Mix, 1 μ l cDNA as template, 1 μ l each of the forward and reverse primers (10 μ M) and 3 μ l ddH₂O. The PCR program consisted of an initial denaturing step of 95°C for 5 min, followed by 5 cycles of 95°C for 10 sec, 55°C for 20 sec, 72°C for 1 min,

and 30 cycles of 95°C for 10 sec, 60°C for 20 sec, 72°C for 1 min, and finally an extension step at 72°C for 15 min. PCR product was purified by the Smarter Nucleic Acid Sample Preparation kit (STRATEC Molecular) and was eluted with 15 µl ddH₂O. 1 µl of solution was used for Nanodrop concentration determination.

Cloning PCR templates with vectors and plasmid extraction

pGEM-T Easy Vector Systems (catalogue number: A1380, Promega) was used to clone PCR templates with vectors. White colonies were selected on the plate. Each colony was dissolved in 15 µl ddH₂O and 1 µl was used as template for colony PCR with the same reaction system and program as the Section for template preparation to make sure that the fragment was ligated successfully onto the vector. The rest of bacteria solution was added to 35 ml LB medium containing ampicillin placed in a 50 ml falcon tube and was incubated at 37°C overnight with 220 rpm shaking table. Plasmid was extracted from the medium using Plasmid Midi Kit (catalogue number: 12143, Qiagen) after taking 3 ml bacteria solution for preparation of *E.coli* culture glycerol stocks (820 µl bacteria solution and 180 µl 87% sterile glycerin sealed in a 2 ml screw-cap culture vial stored at -80°C).

DNA fragment preparation

2.5 µl plasmid (20 ng) was used as template in a 50 µl PCR reaction system containing 25 µl VWR Red Taq DNA Polymerase Master Mix and 2.5 µl each of the T7-forward and T7-reverse primers (10 µM) and ddH₂O. 100 µl PCR reaction system was prepared. The PCR program consisted of an initial denaturing step of 95°C for 5 min, followed by 35 cycles of 95°C for 10 sec, 60°C for 20 sec, and 72°C for 1 min, and finally an extension step at 72°C for 7 min. The product was visualized by running 2% agarose gel. Another PCR reaction was prepared using pUC/M13 sequencing primers and the purified product was used for sequencing using Mix2Seq Kit (Eurofins Genomics) to make sure that sequence of the fragment cloned onto the vector was the same as expected.

To eliminate the plasmid template from the PCR product, 11 μl 10X CutSmart Restriction Enzyme Buffer (catalogue number: B7203S, NEB) and 1 μl DpnI enzyme which could cleave the methylated recognition site (catalogue number: R0176S, NEB) were added. The reaction system was incubated at 37°C for 15 min for digestion of plasmid and then 80°C for 20 min for inactivation of DpnI enzyme. DNA fragment was purified by the purification kit and was eluted with 30 μl ddH₂O. 1 μl of solution was used for Nanodrop concentration determination.

dsRNA synthesis

The T7 RiboMAX Express RNAi System (catalogue number: P1700, Promega) was used to produce dsRNA through in vitro transcription. The 80 μl reaction system was set up in a sterile 1.5 ml tube at room temperature consisting of 40 μl RiboMAX Express T7 2X Buffer, 4 μg DNA fragment, Nuclease-Free Water and 8 μl T7 Express Enzyme Mix. The components were added in order, mixed gently and incubated at 37°C for 2 hours. Since the dsRNA size synthesized was smaller than 1000 bp, the step of annealing the RNA strands was not needed.

Diluted the supplied RNase Solution 1:200 by adding 1 μl RNase Solution to 199 μl Nuclease-Free Water. 4 μl freshly diluted RNase Solution and 4 μl RQ1 RNase-Free DNase were added to 80 μl reaction system and incubated at 37°C for 30 min to remove any remaining single-stranded RNA and the template DNA.

dsRNA purification

8.8 μl supplied 3M Sodium Acetate (pH 5.2) and 88 μl isopropanol were added to the reaction system and mixed. Placed on ice for 5 minutes. At this stage the reaction would appear cloudy. Centrifuged at full speed (14,000 rpm) for 10 minutes at 4°C. A white pellet should be visible after centrifugation. The supernatant was removed carefully.

The pellet was washed with 500 μl cold 70% ethanol. Centrifuged at full speed for 30 sec at 4°C and all ethanol was removed carefully. This washing step was repeated.

The pellet was dried in the air for 20-30 minutes at room temperature. It was very

important that no residual ethanol was remained, or it would cause a high mortality after injecting dsRNA solution containing ethanol into the organism.

The pellet was resuspended in 40 μ l 1X injection buffer. 0.5 μ l solution was diluted in 4.5 μ l RNase-free water and 1 μ l was used for Nanodrop concentration determination while another 4 μ l was used for running 2% agarose gel (**Supplementary Figure 3.2**).

The remaining dsRNA solution was stored at -20°C . It should be noted that normally 46.52 $\mu\text{g/ml}$ dsRNA was quantified as per A_{260} using UV spectrophotometry ([77], Nwokeoji et al., 2017). Since per A_{260} equals to 50 $\mu\text{g/mL}$ DNA, the concentration of dsRNA was about 93% of the concentration measured using DNA settings by Nanodrop. The concentration of dsRNA synthesized was around 10 $\mu\text{g}/\mu\text{l}$.

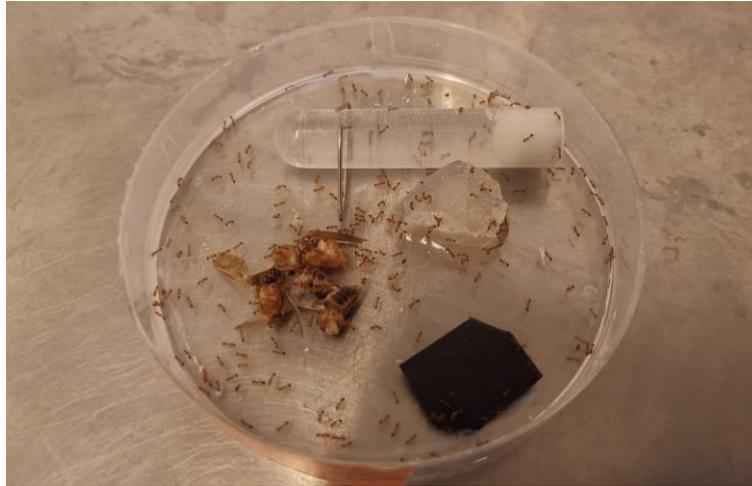
The solution was diluted to a proper concentration used for microinjection in 1X injection buffer.

Injection buffer preparation:

10X injection buffer was prepared by dissolving 4.09 mg NaCl, 0.497 mg Na_2HPO_4 , 0.204 mg KH_2PO_4 and 14.91 mg KCl into 5 ml RNase-Free H_2O followed by filter sterilization. The buffer was stored at room temperature.

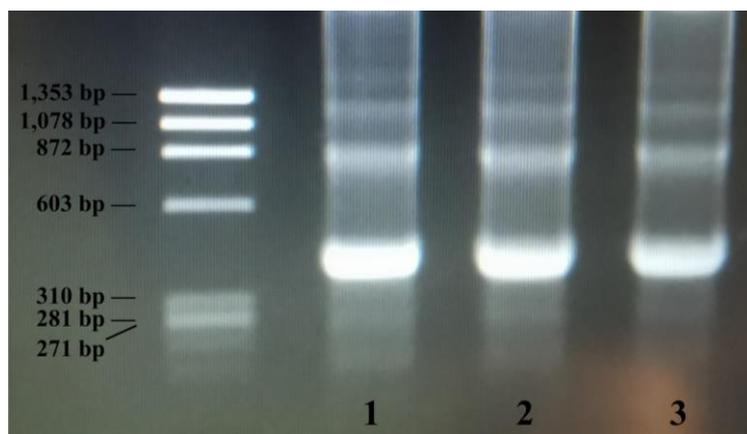
1X injection buffer was prepared when used by diluting 10X injection buffer stock into RNase-Free H_2O . The contents of 1X injection buffer were 1.4 mM NaCl, 0.07 mM Na_2HPO_4 , 0.03 mM KH_2PO_4 and 4 mM KCl.

Supplementary materials



Supplementary Figure 3.1. Setup of an experimental colony

The experimental colonies were set up with workers in order to rear injected larvae. The nest consisted of a piece of thin cardboard (approximately 2x2 cm) with bent edges. The colonies were kept in 15 cm flouon-coated petri dishes. They were fed with crickets and sucrose agar. Water was provided in a cylindrical tube sealed with cotton. The colonies were kept in a climate chamber at 26 ± 2 °C and 50 % humidity.



Supplementary Figure 3.2. dsRNA visualized by 2% agarose gel electrophoresis

PhiX174-HaeIII DNA ladder was used as marker. Since dsRNA exhibited a lower electrophoretic mobility than the same length of dsDNA in 2% agarose gel, dsRNA would run slower and the band would be higher than the corresponding DNA. It should be noted that the non-specific bands were not off-target sequences because Mix2Seq results showed the fragments cloned onto the vector were

correct. These bands might be caused by the annealing between two copies of product via double ended T7 promotor sequence during PCR and were thus the by-product that were two and three times the length of main product, which was not supposed to result in the off-target effect during RNAi. Lane 1-3: three replicates of gene7 dsRNA with the length of 520 bp.

Supplementary Table 3.1. siRNA oligo sequences

Target gene	gene13359
siRNA oligo name	gene- 1
Target sequence	AAGAATAAGGACTTCTAACAA
Oligo sequence (5'→3')	UUGUUAGAAGUCCUUAUUC dTdT
Target gene	<i>yfp</i>
siRNA oligo name	yfp- 3
Target sequence	AAGCAGAAGAACGGCATCAAG
Oligo sequence (5'→3')	CUUGAUGCCGUUCUUCUGC dTdT

Supplementary Table 3.2. Sequences for dsRNA primers and pUC/M13 sequencing primer

Target gene	Primer	Sequence (5' - T7 promoter sequence+ target specific sequence -3')	dsRNA length (bp)
gene13359	T7-gene5-F	TAATACGACTCACTATAGGG+ TGTTTGCGATCTTTGATTGG	524
	T7-gene5-R	TAATACGACTCACTATAGGG+ CTTTGCCGACGTGGTAGAAT	
	T7-gene6-F	TAATACGACTCACTATAGGG+ CCGAACGACTGTCTGTAGCA	573
	T7-gene6-R	TAATACGACTCACTATAGGG+ GCCAAAATTCCTGTCACGTT	
	T7-gene7-F	TAATACGACTCACTATAGGG+ TGTGCGGCTAAGAATAAGGACT	520
	T7-gene7-R	TAATACGACTCACTATAGGG+ TGTGCAATCGGGTATATCTGG	
<i>gfp</i>	T7-gfp-F	TAATACGACTCACTATAGGG+ AGTGCTTCAGCCGCTACCC	494
	T7-gfp-R	TAATACGACTCACTATAGGG+ CATGCCGAGAGTGATCCCCG	
pUC/M13 sequencing primer			
Forward (5'→3')		CAGGAAACAGCTATGAC	
Reverse (5'→3')		CGCCAGGGTTTTCCAGTCACGAC	

Conclusion

Our study focuses on an uncharacterized gene, gene13359, to uncover its evolutionary history and biological functions in differential developmental processes of gynes (unmated queens, reproductive caste) and workers (sterile caste) in *Monomorium pharaonis*.

Results from Chapter One showed that gene13359 was under stronger selection pressure in poneroid ants than in formicoid ants, which might be related to the female physiology of reproductive plasticity in poneroid ants; gene13359 protein was predicted to function extracellularly with a signal peptide at N-terminus. Results from Chapter Two showed that gene13359 exhibited both gyne-male sex-biased and gyne-worker caste-biased expression patterns during development, and we hypothesized that the caste-biased expression was derived from its sex-biased expression in nonsocial ancestors; gene13359 was expressed across the whole body with the highest expression level in abdomen, indicating that it had general female-related physiological functions instead of only ovary-related reproductive functions. Results from Chapter Three showed that gene13359 was an essential gene that mainly functions during the pupation of sexual individuals (gynes and males); knockdown of gene13359 also produced worker-like developmental traits in gyne pupa. We inferred that gene13359 had ancestral functions during pupation in nonsocial ancestors; during the transition to eusociality it acquired derived functions in differential development of gyne and worker castes.

More work is needed to make further study on the functions of gene13359 and other genes that shared the same pathway. As the core gene regulatory networks underlying differential development of queen and worker castes in ants still remain to be uncovered,

our study on an uncharacterized gene, gene13359, provides useful insights to make progress in this field. In conclusion, gene13359 is a promising gene that is likely to play an important role in the process of caste differentiation in ants.

References

- [1] Smith, J. M., & Szathmary, E. (1997). *The major transitions in evolution*. Oxford University Press.
- [2] Hölldobler, B., & Wilson, E. O. (1990). *The ants*. Harvard University Press.
- [3] Wheeler, W. M. (1911). The ant-colony as an organism. *Journal of Morphology*, 22(2), 307-325.
- [4] Hunt, J. H. (2012). A conceptual model for the origin of worker behaviour and adaptation of eusociality. *Journal of evolutionary biology*, 25(1), 1-19.
- [5] Feldmeyer, B., Elsner, D., & Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: worker-specific genes are more derived than queen-specific ones. *Molecular ecology*, 23(1), 151-161.
- [6] Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für tierpsychologie*, 20(4), 410-433.
- [7] Randolph M. Nesse, Tinbergen's Four Questions Organized, <http://nesse.us>, 2000.
- [8] Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*
- [9] Moczek, A. P., Sultan, S., Foster, S., Ledón-Rettig, C., Dworkin, I., Nijhout, H. F., ... & Pfennig, D. W. (2011). The role of developmental plasticity in evolutionary innovation. *Proceedings of the Royal Society B: Biological Sciences*, 278(1719), 2705-2713.
- [10] Kapheim, K. M. (2019). Synthesis of Tinbergen's four questions and the future of sociogenomics. *Behavioral ecology and sociobiology*, 73(1), 186.
- [11] Morandin, C., Tin, M. M., Abril, S., Gómez, C., Pontieri, L., Schiøtt, M., ... & Mikheyev, A. S. (2016). Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. *Genome biology*, 17(1), 43.

- [12] Warner, MR, Mikheyev, AS, & Linksvayer, TA (2017). Genomic signature of kin selection in an ant with mandatory sterile workers. *Molecular Biology and Evolution* , 34 (7), 1780-1787.
- [13] Libbrecht, R., Oxley, P. R., & Kronauer, D. J. (2018). Clonal raider ant brain transcriptomics identifies candidate molecular mechanisms for reproductive division of labor. *BMC biology*, 16(1), 89.
- [14] Walsh, J. T., Warner, M. R., Kase, A., Cushing, B. J., & Linksvayer, T. A. (2018). Ant nurse workers exhibited behavioral and transcriptomic signatures of specialization on larval stage. *Animal Behavior* , 141 , 161-169.
- [15] Warner, M. R., Qiu, L., Holmes, M. J., Mikheyev, A. S., & Linksvayer, T. A. (2019). Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. *Nature communications*, 10(1), 1-11.
- [16] Warner, MR, Mikheyev, AS, & Linksvayer, TA (2019). Transcriptomic basis and evolution of the ant nurse-larval social interactome. *PLoS genetics* , 15 (5), e1008156.
- [17] Chandra, V., Fetter-Pruneda, I., Oxley, P. R., Ritger, A. L., McKenzie, S. K., Libbrecht, R., & Kronauer, D. J. (2018). Social regulation of insulin signaling and the evolution of eusociality in ants. *Science*, 361(6400), 398-402.
- [18] Gospocic, J., Shields, E. J., Glastad, K. M., Lin, Y., Penick, C. A., Yan, H., ... & Liebig, J. (2017). The neuropeptide corazonin controls social behavior and caste identity in ants. *Cell*, 170(4), 748-759.
- [19] Yan, H., Opachaloemphan, C., Mancini, G., Yang, H., Gallitto, M., Mlejnek, J., ... & Perry, M. (2017). An engineered *orco* mutation produces aberrant social behavior and defective neural development in ants. *Cell*, 170(4), 736-747.
- [20] Tribble, W., Olivos-Cisneros, L., McKenzie, S. K., Saragosti, J., Chang, N. C., Matthews, B. J., ... & Kronauer, D. J. (2017). *Orco* mutagenesis causes loss of antennal lobe glomeruli and impaired social behavior in ants. *Cell*, 170(4), 727-735.
- [21] Rajakumar, R., Koch, S., Couture, M., Favé, M.J., Lillico-Ouachour, A., Chen, T., ... & Abouheif, E. (2018). Social regulation of a rudimentary organ generates complex worker-caste systems in ants. *Nature* , 562 (7728), 574-577.

- [22] Simola, D. F., Graham, R. J., Brady, C. M., Enzmann, B. L., Desplan, C., Ray, A., ... & Berger, S. L. (2016). Epigenetic (re) programming of caste-specific behavior in the ant *Camponotus floridanus*. *Science*, 351(6268), aac6633.
- [23] Vail, K. M., & Williams, D. F. (1994). Foraging of the Pharaoh ant, *Monomorium pharaonis*: an exotic in the urban environment. *Exotic ants: biology, impact, and control of introduced species*. Westview, Boulder, CO, 228-239.
- [24] Tsutsui, N. D., & Suarez, A. V. (2003). The colony structure and population biology of invasive ants. *Conservation biology*, 17(1), 48-58.
- [25] Schmidt, A. M., d'Ettorre, P., & Pedersen, J. S. (2010). Low levels of nestmate discrimination despite high genetic differentiation in the invasive pharaoh ant. *Frontiers in zoology*, 7(1), 20.
- [26] Keller, L. (1995). Social life: the paradox of multiple-queen colonies. *Trends in Ecology & Evolution*, 10(9), 355-360.
- [27] Passera, L. (1994). Characteristics of tramp species, *Exotic ants: biology, impact, and control of introduced species*. Westview, Boulder, CO, 23-43. In D. F. Williams (ed.).
- [28] Buczkowski, G., Scharf, M. E., Ratliff, C. R., & Bennett, G. W. (2005). Efficacy of simulated barrier treatments against laboratory colonies of pharaoh ant. *Journal of economic entomology*, 98(2), 485-492.
- [29] Edwards, J. P., & Baker, L. F. (1981). Distribution and importance of the Pharaoh's ant *Monomorium pharaonis* (L) in National Health Service Hospitals in England. *Journal of Hospital Infection*, 2, 249-250.
- [30] Schmidt, A. M., Linksvayer, T. A., Boomsma, J. J., & Pedersen, J. S. (2011). No benefit in diversity? The effect of genetic variation on survival and disease resistance in a polygynous social insect. *Ecological Entomology*, 36(6), 751-759.
- [31] FOWLER, H. G., ALVES, L. E., & BUENO, O. C. (1993). Reproductive strategies of the exotic Pharaoh's ant, *Monomorium pharaonis* (L.) (Hymenoptera: Formicidae) in Brazil. *Invertebrate reproduction & development*, 23(2-3), 235-238.
- [32] Petersen, M., & Buschinger, A. (1971). Das Begattungsverhalten der Pharaoameise,

Monomorium pharaonis (L.) 1. *Zeitschrift für Angewandte Entomologie*, 68(1 - 4), 168-175.

[33] Petersen-Braun, M. (1975). Untersuchungen zur Sozialen Organisation der Pharaoameise *Monomorium pharaonis* (L.) (Hymenoptera, Formicidae) I. Der Brutzyklus und seine Steuerung durch Populationseigene Faktoren. *Insectes Sociaux*, 22(3), 269-291.

[34] Boonen, S., & Billen, J. (2017). Caste regulation in the ant *Monomorium pharaonis* (L.) with emphasis on the role of queens. *Insectes Sociaux*, 64(1), 113-121.

[35] Edwards, J. P. (1991). Caste regulation in the pharaoh's ant *Monomorium pharaonis*: recognition and cannibalism of sexual brood by workers. *Physiological Entomology*, 16(3), 263-271.

[36] Peacock, AD, & Baxter, AT (1949). Studies in Pharaoh's ant, *Monomorium pharaonis* (L.). I. The rearing of artificial colonies. *Ent. Mon. Mag.*, 85, 256-260.

[37] Berndt, K. P., & Eichler, W. (1987). Die Pharaoameise, *Monomorium pharaonis* (L.) (Hym., Myrmicidae). *Mitteilungen aus dem Museum für Naturkunde in Berlin. Zoologisches Museum und Institut für Spezielle Zoologie (Berlin)*, 63(1), 3-186.

[38] Ichinose, K., & Lenoir, A. (2009). Reproductive conflict between laying workers in the ant *Aphaenogaster senilis*. *Journal of ethology*, 27(3), 475-481.

[39] Monnin, T., Ratnieks, F. L., Jones, G. R., & Beard, R. (2002). Pretender punishment induced by chemical signalling in a queenless ant. *Nature*, 419(6902), 61-65.

[40] Linksvayer, TA, Kaftanoglu, O., Akyol, E., Blatch, S., Amdam, GV, & Page Jr, RE (2011). Larval and nurse worker control of developmental plasticity and the evolution of honey bee queen – worker dimorphism. *Journal of evolutionary biology*, 24 (9), 1939-1948.

[41] Tribble, W., & Kronauer, D. J. (2017). Caste development and evolution in ants: it's all about size. *Journal of Experimental Biology*, 220(1), 53-62.

[42] Lillico-Ouachour, A., & Abouheif, E. (2017). Regulation, development, and evolution of caste ratios in the hyperdiverse ant genus *Pheidole*. *Current opinion in*

insect science, 19, 43-51.

[43] Vargo, E. L., & Passera, L. (1991). Pheromonal and behavioral queen control over the production of gynes in the Argentine ant *Iridomyrmex humilis* (Mayr). *Behavioral Ecology and Sociobiology*, 28(3), 161-169.

[44] Boulay, R., Hefetz, A., Cerdá, X., Devers, S., Francke, W., Twele, R., & Lenoir, A. (2007). Production of sexuals in a fission-performing ant: dual effects of queen pheromones and colony size. *Behavioral Ecology and Sociobiology*, 61(10), 1531-1541.

[45] Schwander, T., Humbert, J. Y., Brent, C. S., Cahan, S. H., Chapuis, L., Renai, E., & Keller, L. (2008). Maternal effect on female caste determination in a social insect. *Current Biology*, 18(4), 265-269.

[46] Libbrecht, R., Corona, M., Wende, F., Azevedo, D. O., Serrão, J. E., & Keller, L. (2013). Interplay between insulin signaling, juvenile hormone, and vitellogenin regulates maternal effects on polyphenism in ants. *Proceedings of the National Academy of Sciences*, 110(27), 11050-11055.

[47] Warner, MR, Lipponen, J., & Linksvayer, TA (2018). Pharaoh ant colonies dynamically regulate reproductive allocation based on colony demography. *Behavioral Ecology and Sociobiology*, 72 (3), 31.

[48] Edwards, J. P. (1987). Caste regulation in the pharaoh's ant *Monomorium pharaonis*: the influence of queens on the production of new sexual forms. *Physiological entomology*, 12(1), 31-39.

[49] Penick, C. A., & Liebig, J. (2017). A larval 'princess pheromone' identifies future ant queens based on their juvenile hormone content. *Animal Behaviour*, 128, 33-40.

[50] Tschinkel, W. R. (1993). Resource allocation, brood production and cannibalism during colony founding in the fire ant, *Solenopsis invicta*. *Behavioral Ecology and Sociobiology*, 33(4), 209-223.

[51] Schmidt, A. M., Linksvayer, T. A., Boomsma, J. J., & Pedersen, J. S. (2011). Queen-worker caste ratio depends on colony size in the pharaoh ant (*Monomorium pharaonis*). *Insectes sociaux*, 58(2), 139-144.

[52] Warner, M. R., Kovaka, K., & Linksvayer, T. A. (2016). Late-instar ant worker

larvae play a prominent role in colony-level caste regulation. *Insectes sociaux*, 63(4), 575-583.

[53] Pontieri, L., Schmidt, A. M., Singh, R., Pedersen, J. S., & Linksvayer, T. A. (2017). Artificial selection on ant female caste ratio uncovers a link between female-biased sex ratios and infection by *Wolbachia* endosymbionts. *Journal of evolutionary biology*, 30(2), 225-234.

[54] Khila, A., & Abouheif, E. (2010). Evaluating the role of reproductive constraints in ant social evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1540), 617-630.

[55] Berndt, K. P., & Kremer, G. (1986). Die Larvenmorphologie der Pharaoameise *Monomorium pharaonis* (L.) (Hymenoptera, Formicidae). *Zoologischer Anzeiger*, 216(5-6), 305-320.

[56] Peacock, A. D., & Baxter, A. T. (1950). Studies in Pharaoh's ant, *Monomorium pharaonis* (L.). 3. Life history. *Entomologist's Monthly Magazine*, 86, 171-178.

[57] Mikheyev, AS, & Linksvayer, TA (2015). Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. *Elife*, 4, e04775.

[58] Børgesen, L. W. (2000). Nutritional function of replete workers in the pharaoh's ant, *Monomorium pharaonis* (L.). *Insectes Sociaux*, 47(2), 141-146.

[59] Pontieri L., & Linksvayer T.A. (2019) *Monomorium*. In: Starr C. (eds) *Encyclopedia of Social Insects*. Springer, Cham.

[60] Gobin, B., Ito, F., Billen, J., & Peeters, C. (2008). Degeneration of sperm reservoir and the loss of mating ability in worker ants. *Naturwissenschaften*, 95(11), 1041-1048.

[61] Boomsma, J. J., & Gawne, R. (2018). Superorganismality and caste differentiation as points of no return: how the major evolutionary transitions were lost in translation. *Biological Reviews*, 93(1), 28-54.

[62] Rabeling, C., & Kronauer, D. J. (2013). Thelytokous parthenogenesis in eusocial Hymenoptera. *Annual review of entomology*, 58, 273-292.

[63] SANETRA, M., & BUSCHINGER, A. (2000). Phylogenetic relationships among social parasites and their hosts in the ant tribe Tetramoriini (Hymenoptera:

- Formicidae). *European Journal of Entomology*, 97(1), 95-118.
- [64] Hasegawa, E., Tinaut, A., & Ruano, F. (2002). Molecular phylogeny of two slave-making ants: *Rossomyrmex* and *Polyergus* (Hymenoptera: Formicidae). In *Annales Zoologistes Fennici* (pp. 267-271). Finnish Zoological and Botanical Publishing Board.
- [65] Ward, P. S., Blaimer, B. B., & Fisher, B. L. (2016). A revised phylogenetic classification of the ant subfamily Formicinae (Hymenoptera: Formicidae), with resurrection of the genera *Colobopsis* and *Dinomyrmex*. *Zootaxa*, 4072(3), 343-357.
- [66] Blaimer, B. B., Brady, S. G., Schultz, T. R., Lloyd, M. W., Fisher, B. L., & Ward, P. S. (2015). Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC evolutionary biology*, 15(1), 271.
- [67] Borowiec, M. L., Rabeling, C., Brady, S. G., Fisher, B. L., Schultz, T. R., & Ward, P. S. (2019). Compositional heterogeneity and outgroup choice influence the internal phylogeny of the ants. *Molecular phylogenetics and evolution*, 134, 111-121.
- [68] Livramento, K. G. D., Freitas, N. C., Máximo, W. P., Zanetti, R., & Paiva, L. V. (2018). Gene expression profile analysis is directly affected by the selected reference gene: the case of leaf-cutting *Atta sexdens*. *Insects*, 9(1), 18.
- [69] Jia, L. Y., Chen, L., Keller, L., Wang, J., Xiao, J. H., & Huang, D. W. (2018). *Doublesex* evolution is correlated with social complexity in ants. *Genome biology and evolution*, 10(12), 3230-3242.
- [70] Hunt, B. G., Ometto, L., Wurm, Y., Shoemaker, D., Soojin, V. Y., Keller, L., & Goodisman, M. A. (2011). Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proceedings of the National Academy of Sciences*, 108(38), 15936-15941.
- [71] Ellegren, H., & Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*, 8(9), 689-698.
- [72] Araki, M., Miyakawa, M. O., Suzuki, T., & Miyakawa, H. (2020). Two insulin-like peptides may regulate egg production in opposite directions via juvenile hormone signaling in the queenless ant *Pristomyrmex punctatus*. *Journal of Experimental*

Zoology Part B: Molecular and Developmental Evolution, 334(4), 225-234.

- [73] Schmidt, A. M. (2010). *The invasion biology and sociogenetics of pharaoh ants* (Doctoral dissertation, Museum Tusulanum).
- [74] Bourke, A. F., & Franks, N. R., & Franks, N. R. (1995). *Social evolution in ants* (Vol. 62). Princeton University Press.
- [75] Petersen-Braun, M. (1977). Untersuchungen zur sozialen Organisation der Pharaoameise *Monomorium pharaonis* L. (Hymenoptera, Formicidae) II. Die Kastendeterminierung. *Insectes Sociaux*, 24(4), 303-318.
- [76] Miller, S. C., Miyata, K., Brown, S. J., & Tomoyasu, Y. (2012). Dissecting systemic RNA interference in the red flour beetle *Tribolium castaneum*: parameters affecting the efficiency of RNAi. *PloS one*, 7(10).
- [77] Nwokeoji, AO, Kilby, PM, Portwood, DE, & Dickman, MJ (2017). Accurate quantification of nucleic acids using hypochromicity measurements in conjunction with UV spectrophotometry. *Analytical chemistry*, 89 (24), 13567-13574.
- [78] Schmidt, C. (2013). Molecular phylogenetics of ponerine ants (Hymenoptera: Formicidae: Ponerinae). *Zootaxa*, 3647(2), 201-250.

Supplementary Information

RNA extraction protocol

The protocol was modified from the manufactures' protocol, and performed as followed:

The sample was collected in a 2 ml Eppendorf tube.

A 5 mm sterile stainless-steel bead was added to the tube.

350 μ l Buffer RLT Plus, 3.5 μ l β -mercaptoethanol (β -ME) and 1.75 μ l Reagent DX (cat. no. 19088) was added to the tube. Buffer RLT Plus was used for lysing tissue; β -ME for protecting RNA from degradation; Reagent DX for reducing foaming.

Disrupted and homogenized the tissue using the TissueLyser instruments (Qiagen) for 3 min at 50 Hz and centrifuged the lysate for 3 min at full speed (14,000 rpm).

Around 320 μ l supernatant was transferred by pipetting to the gDNA Eliminator Spin Column placed in a 2 ml collection tube and was centrifuged for 30 sec at 11,000 rpm. gDNA Eliminator Spin Column was kept for further DNA extraction.

320 μ l DEPC-treated 70% ethanol was added to the flow-through and mixed well by pipetting.

Mixed flow-through was transferred to the RNeasy MinElute Spin Column placed in a 2 ml collection tube and centrifuged for 15 sec at 11,000 rpm. Flow-through was discarded.

700 μ l Buffer RW1 was added to column and centrifuged for 15 sec at 11,000 rpm. Flow-through was discarded.

500 μ l Buffer RPE was added to column and centrifuged for 15 sec at 11,000 rpm. Flow-through was discarded.

500 μ l DEPC-treated 80% ethanol was added to column and centrifuged for 2 min at 11,000 rpm. Flow-through was discarded.

The spin column was transferred to a new 2 ml collection tube and centrifuged at full speed for 5 min at 4°C with the lid open to dry the membrane.

The spin column was transferred to a new 1.5 ml collection tube. 14 μ l RNase-free water was added directly to the center of the spin column membrane and was centrifuged at full speed for 1 min at 4°C to elute the RNA.

2 μ l elution was transferred to a 0.2 ml tube for further concentration determination through Qubit RNA HS Assay. The extracted RNA was stored at -80°C.

DNA extraction protocol

500 μ l AW1 Buffer (cat. No. 19081) was added to the gDNA Eliminator Spin Column placed in a 2 ml collection tube. Centrifuged at 11,000 rpm for 2 min. Flow-through was discarded.

500 μ l AW2 Buffer (cat. No. 19072) was added to the spin column and centrifuged at full speed for 2 min. Flow-through was discarded.

The empty spin column was centrifuged at full speed for 2 min to dry the membrane.

50 μ l pre-warmed (~40°C) AE Buffer (cat. No. 19077) was added directly to the center of the spin column membrane and was incubated for approximately 5 min at room temperature. Centrifuged at 11,000 rpm for 1 min to elute the DNA.

1 μ l elution was used for Nanodrop to determine the concentration, and the rest was stored at -20°C.

Qubit RNA HS Assay protocol

The Qubit RNA HS (High Sensitivity) Assay Kit (catalogue number: Q32852, Thermo Scientific) was used for RNA quantitation with 0.5 ml Qubit assay tubes (catalogue number: Q32856, Thermo Scientific) and Qubit 3.0 Fluorometer (Thermo Scientific). The Qubit working solution was made by diluting the Qubit RNA HS Reagent 1:200 in Qubit RNA HS Buffer.

Added 190 μ l working solution to 10 μ l of each Qubit standard in standard tubes; added 198 μ l working solution to 2 μ l extracted RNA in individual assay tubes. Mixed tubes by vortexing 2–3 seconds.

All tubes were incubated at room temperature for 2 minutes.

RNA concentration was measured by Qubit 3.0 Fluorometer after reading standard tubes and assay tubes.

PCR product purification protocol

250 μ l Binding Buffer (catalogue number: 10202220) was added to the PCR sample and mixed well by vortexing. Transferred the sample completely onto a Spin Filter placed into a 2.0 ml Receiver Tube, incubated for 1 min at room temperature and centrifuged for 4 min at full speed (14,000 rpm).

Added 15 μ l ddH₂O directly onto the center of the Spin Filter placed into a new 1.5 ml Receiver Tube, incubated for 1 min at room temperature and centrifuged for 1 min at 11,000 rpm.

PCR template cloning protocol

The 10 μ l ligation reaction consists of 5 μ l Rapid Ligation Buffer, 1 μ l pGEM-T Easy Vector (50 ng), 30 ng PCR template (usually 0.5 μ l), 1 μ l T4 DNA Ligase and ddH₂O. The reaction was mixed by pipetting and incubated for one hour at room temperature. 2 μ l ligation reaction was added to 50 μ l JM109 High Efficiency Competent Cells placed in a sterile 1.5 ml tube on ice. After incubating on ice for 20 minutes, the tube was heat-shocked for 45 to 50 sec in a water bath at exactly 42°C and then was immediately returned on ice for 2 minutes.

950 μ l room temperature SOC medium was added to the ligation reaction transformations and was incubate for 1.5 hours at 37°C with shaking (~150rpm).

100 μ l of each transformation culture was plated onto LB/ampicillin/IPTG/ X-Gal plates. The plates were incubated overnight at 37°C and white colonies were selected.

Preparation of LB/ampicillin/IPTG/ X-Gal plates:

20 mg/ml X-Gal in DMF (covered with aluminum foil, stored at -20 °C for 12 months in maximum), 50 mg/ml ampicillin in dH₂O (stored at 4 °C) and 0.1 M IPTG in dH₂O (119.15 mg IPTG + 5 ml dH₂O) were prepared and filter-sterilized.

10 g tryptone, 5 g yeast extract, 5 g NaCl and 15 g agar were added to 1 L dH₂O. The solution was autoclaved and cooled to 50 °C. 1 ml 50 mg/ml ampicillin solution was added to the solution. The medium was allocated to 15 cm petri dishes with around 40 ml medium for each plate. Plates were put in laminar hood with lid open for 30 min to let the medium become solid.

120 µl 20 mg/ml X-Gal solution and 40 µl 0.1 M IPTG solution were plated onto each plate. Plates were sealed by biofilm and stored at 4°C.

Preparation of LB medium containing ampicillin (1L):

10 g tryptone, 5 g yeast extract and 5 g NaCl were added to 1 L dH₂O. The solution was autoclaved and cooled to room temperature. 1 ml 50 mg/ml ampicillin solution was added to the solution. The medium was stored at 4°C for six months in maximum.

Preparation of SOC medium (100 ml):

2.0 g tryptone, 0.5 g yeast extract, 1 ml 1 M NaCl and 0.25 ml 1 M KCl were added to 97 ml dH₂O. The solution was autoclaved and cooled to room temperature. 1 ml 2 M Mg²⁺ stock and 1 ml 2 M glucose were then added to the solution. The medium was stored at 4°C for six months in maximum.

(1) 1 M NaCl (58.5 g/L): 0.2925 g + 5 ml dH₂O, filter-sterilized

(2) 1 M KCl (74.5 g/L): 0.3725 g + 5 ml dH₂O, filter-sterilized

(3) 2 M Mg²⁺ stock: 1.0165 g MgCl₂•6H₂O + 0.602 g MgSO₄ + 5 ml dH₂O, filter-sterilized

(4) 2 M glucose: 1.9817 g glucose monohydrate + 5 ml dH₂O, filter-sterilized

Plasmid extraction protocol

The bacterial cells from about 30 ml solution were harvested by centrifugation at 6000 x g for 15 min at 4°C.

4 ml Buffer P1 was added to resuspend the bacterial pellet.

4 ml Buffer P2 was then added with thoroughly mixing and was incubated at room temperature for 5 min.

4 ml Buffer P3 was added with thoroughly mixing and was incubated on ice for 15 min. Centrifuged at full speed (20,817 x g) for 30 min at 4°C. QIAGEN-tip 100 was equilibrated by applying 4 ml Buffer QBT and was emptied by gravity flow.

The supernatant was applied to the equilibrated tip and entered the resin by gravity flow.

The tip was washed with 10 ml Buffer QC for two times.

Plasmid DNA was eluted with 5 ml Buffer QF and was precipitated by adding 3.5 ml room-temperature isopropanol. Mixed and centrifuged immediately at full speed for 30 min at 4°C. The supernatant was decanted carefully.

The pellet was washed with 2 ml room-temperature 70% ethanol, and was centrifuged at full speed for 10 min. The supernatant was decanted carefully. This step was repeated.

The pellet was dried in air for 5-10 min was dissolved in 50 µl ddH₂O. 1µl of solution was used for concentration determination by Nanodrop. Usually the plasmid concentration was around 1000 ng/µl. The solution was diluted to 20 ng/µl and was stored at -20°C.

Mix2Seq protocol

15 µl purified PCR product (the concentration was around 5 ng/µl) was transferred to one of the Mix2Seq tubes provided by the kit.

2 µl pUC/M13 forward primer or pUC/M13 reverse primer (10 µM) was added to the tube and mixed well by pipetting. The total volume of the mixed sample should be exactly at 17 µl.

The tube was sealed with a lid from the enclosed cap pad and was placed in a plastic bag which will be sent to Eurofins Genomics in Germany for sequencing.

Acknowledgements

I would like to thank the following people for this project:

I would like to give many thanks to my supervisor, Professor Guojie Zhang, for providing many guidance on the project and also the opportunity for me to do research at University of Copenhagen as part of my bachelor thesis. I would like to thank other members at CSE including Rasmus Stenbak Larsen for helping me with ant colonies setup and maintenance, guidance on all laboratory work and introduction to the biology of *Monomorium pharaonis*; Bitao Qiu for visualizing the transcriptome data; Sylvia Mathiasen for helping me order the reagents; Justinn Hamilton, who is my office mate, Wenjing Xu and Kristjan Germer for many useful discussions on my project; Antonin Jean Johan Crumiere, Luigi Pontieri, Manuel Nagel and Morten Schiøtt for their help during the beginning of my project. I would like to thank the members at Kunming Institute of Zoology including Guo Ding for guidance on double-strand RNA synthesis; Weiwei Liu and Qionghua Gao for guidance on microinjection setup. I would also like to thank the members at BGI in Shenzhen including Zijun Xiong and Fang Li for their patient guidance on gene blast and positive selection analysis that were operated on the BGI server.

I would like to thank Zhiyuan College at Shanghai Jiao Tong University for giving me permission to work abroad for ten months, which is uncommon for a bachelor student in China.

Finally, I give my sincere thanks to my family for their support and care on my daily life during this ten-month period in Denmark.

