# Preliminary characterisation of aggregation behaviour in *Drosophila melanogaster*

**A Thesis Submitted**

**In Partial Fulfillment of the Requirements**

**for the Degree of Master of Science**

**by**

**Rutvij Kulkarni**

**Jawaharlal Nehru Centre for Advanced Scientific Research,**

**Jakkur, Bengaluru 560 064, India**

**May, 2019**

*In loving memory of Fidget*

*who tried helping me with analysis once*

## TABLE OF CONTENTS

## DECLARATION

I hereby declare that the work embodied in this thesis, entitled "**Preliminary characterisation of aggregation behaviour in *Drosophila melanogaster***", has been carried out by me under the supervision of Dr. Sheeba Vasu, Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru, and that it has not been submitted for any degree or diploma to any other institution.

Following prevalent scientific practice, acknowledgements have been accorded wherever due. Any omissions, due to oversight or error of judgement, are deeply regretted.

Place: Bengaluru (Rutvij Kulkarni)

May, 2019

## CERTIFICATE

This is to certify that the work presented in this thesis, entitled "**Preliminary characterisation of aggregation behaviour in *Drosophila melanogaster***", has been carried out by Mr. Rutvij Kulkarni under my supervision at the Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru, and that the results presented in this thesis have not previously formed the basis for the award of any diploma, degree or fellowship.

(Dr. Sheeba Vasu)

## ACKNOWLEDGEMENTS

This thesis is an account of my work, a labour of love and other emotions, that has overshadowed everything in my life for the past few years. Over the course of these years, I have faced ups and many downs, learnt some and forgotten more, and emerged all the better for it. Yet, as is customary with academic writing, this thesis tells the story of my scientific efforts from these years shorn of the human experiences behind them. However, as I discuss later in this thesis, an organism's phenotype is affected significantly by its social environment. Hence, I would be remiss if I forgot the irreplaceable cast of people who have been an integral part of my social environment and without whom this thesis would be decidedly different.

My greatest debt of gratitude is owed to the Evolutionary and Organismal Biology Unit (EOBU) at JNCASR, my first home away from home, which is hugely responsible for shaping me as a scientist and a person. These people have taught me much and more about science, and about life in general, that I will cherish forever. I am eternally grateful to Dr. Sheeba Vasu without whom this project would have never begun. It was because of the creative space that I found in her lab during my first rotation that the idea for this project first came to me. And it is because of her unwavering faith in me that I was eventually able to give form to it. I am deeply thankful to the entire EOBU faculty, Professors Amitabh Joshi and Vijay Kumar Sharma, Dr. TNC Vidya and Dr. Sheeba Vasu, for the incredible learning experiences that I have had during coursework and lab rotations. I have grown tremendously, both as a person and a scientist, thanks to my interactions with them. I owe special thanks to Dr. Keerthipriya P., Abhilash Lakshman and Manan Gupta for being unbelievably helpful whenever I felt lost and confused with work. They have been a huge source of knowledge and inspiration for me and this thesis would have certainly been a poorer work of science without their guidance. I also thank all members of the Behavioural Neurogenetics and Chronobiology labs for being an eager and honest audience for my work. Despite it being unfamiliar and unrelated to

their work, they have invested time and effort to understand my work and have often provided me with newer perspectives. I owe a huge debt of thanks to Rajanna and Muniraju for all their help during fly cooks and maintenance. I will never forget how once, when I ran out of vials in the middle of a painful experiment, Muniraju calmed me down and made sure that I had all the vials I needed within a matter of few hours. I truly could not have completed my experiments successfully without their help. I am also thankful to Vybhav for helping me find the right camera for my experiments and to Mr. Raju from the JNC workshop who fabricated the assay arena for me. Finally, I am thankful to the hostel wardens for allowing my stay in the hostel despite the delay in the submission of this thesis.

Science is an unforgiving muse. I learnt this the hard way and the only reason I managed to persist was because of the comradeship I have had with my fellow students. Alex and Aditi have been my closest companions for much of my time at JNC and have seen me through the best and the worst of times. They have always been there for me, regardless of whether I wanted them or not, and have been a constant source of reassurance when I was at my worst. I truly cannot thank them enough for their friendship and love. I am also lucky to have found a bunch of bright and honest friends in Aishwariya, Srikant, Viveka, Sheetal, Keerthi, Manan, Anuj, Abhilash, Aishwarya, Chitrang, Mayank and Ranjitha, who have been a constant source of support and camaraderie over the past few years. From lunch time banter to intense analyses of TV shows, I have shared the simplest of joys with them. I have also learnt immensely, about anything and everything under the sun, from them. I would also like to thank the all the critters - Mess cat, Constance, Fidget, Boopy, Dusty, Crookshanks, Bushy and Frisky- who found their way into my life and left it significantly happier.

I would not be the person I am today without the support and guidance of my friends and family, who deserve much of the credit for ensuring the completion of this thesis. My parents and my brother have been perennial sources of support and encouragement ever since I chose to pursue

science. Without their faith in me, I would have never found the confidence to ignore societal misgivings about my career choices. I have also been exceptionally lucky to have befriended an amazing set of people- Chinmay, Akash, Sanika, Aishwarya, Trisha, Naomi, Shawn, Alex, Aditi and Keerthi- without whom my life would be markedly duller. They have shaped the way I see the world, and myself, and I honestly can't imagine the person I would be without them. Finally, I want to acknowledge Fidget who found me in some of my darkest moments and introduced me to a side of myself that I didn't know existed.

## SYNOPSIS

Aggregation is the tendency of organisms to gather in space and in time. This behaviour is seen in a wide variety of species and is typically brought about by predation threats, presence of common resources or social factors. The fruit fly, *Drosophila melanogaster*, is a non-group living organism which is commonly seen to aggregate on food sources. Interestingly, *Drosophila* aggregate even in the absence of food (and predators) suggesting that social factors may play some role during aggregation. Although such aggregation has been studied before, these studies have failed to describe many aspects of the aggregation pattern such as aggregate size, stability of aggregates and the strength of association among flies. Hence, my goal was to comprehensively describe fly aggregation patterns and to understand the individual behaviours that underlie these patterns. I also wanted to understand the effects of social factors such as mating status and prior social experience on these behaviours.

I obtained fly spatial patterns by video recording groups of ~30 females in a circular arena for two hours. Social environment of these groups was varied by varying their mating status and exposure to other female flies. I analysed these patterns using the Ripley's K method and identified clear instances of non-random proximity between flies. Using empirical estimates of such proximity I could identify individual aggregates formed by the flies. I measured the features of these aggregates and also quantified the tendencies of individual flies to join and stay in these aggregates. Finally, to understand why flies may aggregate, I made preliminary behavioural observations of social interactions between flies.

I found that variation in properties of the aggregate pattern such as aggregate size and aggregate stability could be explained by variation in the tendencies of individual flies to join or stay in aggregates. However, these tendencies were not fully explained by social interactions among flies,

suggesting that other behaviours may be involved. Interestingly, flies showed non-random associations with each other, which suggests that flies may form some kinds of social relationships. The biological relevance of such relationships, however, is unclear. Finally, mating and prior social experience had distinct effects on aggregation behaviour which suggests that aggregation may serve distinct roles in different social contexts.

# INTRODUCTION

When a cat detects prey, it stalks the prey before attacking it. Prey detection, thus, causes the cat to perform a series of non-random actions to conceal itself and capture the prey. These non-random actions are repeated every time the cat encounters any sort of prey. However, these actions may not be seen when the cat detects other non-food objects or if the prey is unfamiliar to the cat. Sometimes, the cat may fail to show these actions even when it encounters familiar prey simply because it may have fed recently. Such non-random actions are called behaviours and, as the above observations illustrate, they occur in response to specific environmental stimuli in a context dependent manner. This context specificity often depends on whether the organism stands to receive a net benefit by performing the behaviour. Thus, behaviours serve as responses to changes in an organism's environment that may allow the organism to adapt to these changes.

Given the importance of responding to constantly changing environments, behaviours are ubiquitous across taxa. This ubiquity is accompanied by great diversity in form and function of behaviours. Consequently, a variety of questions may be posed to understand how and why such behavioural diversity exists in nature. Tinbergen (1963) broadly organized these questions into four complementary categories - causation, ontogeny, function and phylogeny (or evolution) - and thereby provided a framework to address different aspects of any given behaviour. These questions may be further grouped into proximate (or *how?*) questions (causation and ontogeny), which explore the biological mechanisms underlying behaviours, and ultimate (or *why?*) questions (function and phylogeny) which probe the evolutionary processes that underlie behaviours. Between them, these questions provide a comprehensive approach to studying any behaviour. Hence, I have used this framework to organize and discuss the literature available for aggregation behaviour in the fruit fly, *Drosophila melanogaster*.

### 1.1. What is aggregation behaviour?

Social behaviours are a subset of behaviours that occur in response to other conspecifics in an individual's environment. Although conspecifics may include individuals of both sexes, social behaviours typically involve behavioural interactions between individuals of the same sex. Inter-sexual interactions are largely studied in the context of mating and reproduction but may be considered social if they are non-sexual in nature (such as parental care). Social behaviours may agonistic i.e. involving conflict, or non-agonistic i.e. involving tolerance or active co-operation. In the former case, inter-individual interactions are often asymmetric such that one individual benefits at the expense of the other. For example, individuals engage in aggressive encounters with each other to determine access to resources such as food or mates (Alcock, 2001, p. 267-268, 327-329). The latter category includes a range of non-hostile interactions from simple, non-specific associations between individuals such as aggregations, to more complex, co-operative behaviours such as social grooming (Alcock, 2001, p. 431). While agonistic interactions may be easily explained in terms of conflict over resources, non-agonistic behaviours may depend on a variety of environmental factors. Interactions among these environmental factors contribute to a great diversity in the types of non-agonistic behaviours that can be observed across species and across different populations of a given species.

Aggregation is one such non-agonistic, social behaviour which is commonly observed across diverse taxa. It refers to the tendency of individuals to gather in space and time. Such gathering is shaped by ecological factors such as predation, resource distributions and social environment. Organisms aggregate in response to predation, as the per capita risk of predation is lowered in groups due to dilution effects (Hamilton, 1971) or increased vigilance (Alcock, 2001, p. 202-203). At the same time, grouping can potentially be costly as it increases competition for resources among individuals by increasing local density of individuals (Alcock, 2001, p. 423). If these costs are low (or cost of engaging in conflict is high, such that net cost is low), then aggregates may form

near resource patches simply due to shared preferences for resources. Aggregations may also form as a result of social relationships between individuals, whereby individuals gain social benefits such as offspring care, coalitionary support or opportunities for social learning by being with other conspecifics (Alcock, 2001, p. 451-453, 427, 227-229).

Proximity brought about by aggregate formation may have several secondary consequences for the aggregating individuals. Individuals within the aggregate may engage in social interactions, which would allow for the development of social relationships among them (reviewed in Aureli et al., 2002). These relationships may, in turn, affect cohesion of the aggregate as well as other aspects of an individual's biology, such as the transfer of pathogens (Patterson and Ruckstuhl, 2013; Sah et al., 2018) and information among individuals (Fernandez-Juricic and Kacelnik., 2004; Lind and Lindenfors, 2010). In some cases, this proximity may also perform novel, emergent functions such as social thermoregulation (reviewed in Terrien et al., 2011).

Aggregations may be described using different group-level properties such as size (i.e. number of individuals in the aggregate), temporal stability and composition (i.e. identities of individuals in the aggregate and relationships among them). As the ecological forces shaping aggregation change, these properties may change to accommodate them (reviewed in Aureli et al., 2008). For example, large aggregates may form when benefits, in terms of predator protection or social value, are high or when costs of grouping are low or non-existent. However, if resource competition or social conflict were to increase, then individuals may leave these aggregates resulting in a reduction of size. These properties can thus provide insight into the ecological forces that shape aggregation. Epiphenomena of aggregation are also often linked to the properties of the aggregate such as its size (disease spread- reviewed in Patterson and Ruckstuhl, 2013; thermoregulation- Willis and Brigham, 2007) and the network of relationships within the aggregate (social relationships- Chase et al., 1982; disease spread- VanderWaal and Ezenwa, 2016; Sah et al., 2018). For example, rates of exchange of information and/or pathogens would be higher in groups with highly interconnected

individuals. Thus, these properties are also useful for understanding other phenotypes associated with aggregation.

### 1.2. *Drosophila* as a model to study behaviour

The fruit fly, *Drosophila melanogaster,* is known to exhibit a wide variety of behaviours which can be easily observed and quantified. The proximate bases of many such behaviours have been studied extensively thanks to the abundance of molecular and genetic tools available in *Drosophila* (circadian rhythms- reviewed in Sheeba, 2008; aggression- reviewed in Hoopfer, 2016; courtship- reviewed in Yamamoto and Koganezawa, 2013; thermal preference- reviewed in Dillon et al., 2009; sleep- Griffith, 2013). These tools can be used to identify neuronal circuits and molecular players underlying behaviours by precisely manipulating properties of different neuronal populations and studying their effects. Owing to its ease of handling and maintenance, *Drosophila* is also ideal for studying ultimate questions under laboratory conditions. Populations comprising several hundred flies can be easily maintained under laboratory conditions. It is also possible to measure behavioural and fitness related traits for large numbers of flies. These benefits have allowed for the study of relationships between trait variation and fitness at the level of populations for a variety of traits (gregarious oviposition- Ruiz-Dubreuil et al., 1994; circadian behaviours- Sheeba et al., 1999; aggression- Hoffmann and Cacoyianni, 1989, 1990). Fruit flies are thus ideal systems for conducting comprehensive experimental studies of behaviour.

### 1.3. Aggregation behaviour in *Drosophila melanogaster*

As *Drosophila* are not known to be typical group living organisms, their repertoire of social behaviours is believed to be limited to aggression and aggregation. Aggression commonly occurs

between male flies in the presence of patchy food and females (Jacobs, 1960) and also, to a lesser extent, between female flies in the presence of high quality food (Ueda and Kidokoro, 2002). Proximate and ultimate bases of such behaviour have been widely studied previously. In contrast, aggregation behaviour has been studied less extensively and also less systematically. Aggregation in flies occurs over large distances, whereby flies locate and gather at food sources such as rotting fruits in their environment. Flies are attracted to these sources in response to food odours and any pheromonal cues that may be left at these sources by other conspecifics (Bartelt et al., 1985; Wertheim et al., 2006). Aggregation also occurs over much smaller distances, as flies form non-random clusters on such food sources (Saltz and Foley, 2011; Soto-Yeber et al., 2018). This short-range aggregation can occur even in the absence of food (Navarro and del Solar, 1975; Simon et al., 2012) suggesting that conspecifics alone may be mediating such aggregation. While some studies have attempted to understand how these aggregates may be formed, we know very little about the behavioural bases for such aggregation as well as the ultimate reasons that may promote it.

In the following section, I provide a brief overview of the experimental approaches that have been used previously to study aggregation. I also discuss available literature for *Drosophila* aggregation and some other studies that may be useful to understand how and why such aggregation may occur.

*1.3.1. Studying aggregation in flies*

Unlike other behaviours in *Drosophila*, there are no standardized protocols to study aggregation behaviour. Consequently, assay conditions as well as the measures used for quantifying aggregation vary widely across studies. One of the simpler approaches to study aggregation involves comparison of observed spatial patterns of fly positions to those expected to form purely by random chance (Navarro and del Solar, 1975; Lefranc et al. 2001). Such an approach allows for identification of statistically non-random patterns, but is not useful to compare multiple, non-

random spatial patterns with each other. Other approaches involve the use of nearest neighbour distances as a proxy for the individual tendency to aggregate (Simon et al., 2012; Anderson et al., 2017). Although these distances are easy to measure and useful for comparing spatial patterns, they fail to comprehensively describe spatial patterns. Since they focus only on the nearest neighbour, nearest neighbour distances ignore other individuals that may be aggregating with the focal fly. As a result, they fail to account for differences in the number of aggregating individuals. The number of aggregating individuals, or aggregate size, may also be used as a measure of aggregation. Since it often reflects the underlying factors mediating aggregation, it represents a more biologically meaningful alternative to nearest neighbour distance for measuring aggregation. Unfortunately, as aggregates have been rarely defined in case of *Drosophila*, sizes of aggregates formed by flies have not been quantified previously. The few studies that have attempted to define aggregates in flies have defined aggregates with respect to some environmental patchiness such as food sources (Saltz and Foley, 2011; Saltz, 2011) or low temperature refuges (Philippe et al. 2016). However, this approach also has some limitations, as it is unclear if the flies present on such patches truly constitute a single aggregate. Additionally, this approach is not useful for identifying aggregates in the absence of resource heterogeneities. To my knowledge, no study has fully accounted for these limitations while describing aggregation patterns. Therefore, newer methods need to be developed that can accurately identify and quantify aggregates formed by flies.

It is important to note that limitations of existing methods do not invalidate results from previous studies. Instead, they highlight the need for caution while comparing results across studies.

### *1.3.2. Causal factors underlying aggregation in flies*

#### 1.3.2.1. Behavioural bases of aggregation:

The spatial pattern seen during aggregation is an outcome of spatial choices made by individual flies and the interactions among them. To study these individual level choices, Philippe et al.

(2016) recorded the behaviour of female flies kept in a high temperature arena containing two low temperature refuges. They found that the tendency of a fly to join and stay in a refuge was higher for the more crowded refuge, suggesting that flies make aggregation choices based on the size of aggregates. A similar tendency to preferentially join larger aggregates has also been seen in males (Saltz, 2011). In addition to size, inter-fly interactions can also influence individual choices to join or leave groups. For example, aggressive males drive away other males from food patches resulting in an over-dispersed spatial pattern for males but not for females (Saltz and Foley, 2011; Foley et al. 2015). Other behaviours such as physical interactions, which have been found to be negatively correlated with nearest neighbour distance across mutant fly strains, may also influence spatial patterns (Anderson et al., 2017).

Individual tendencies to aggregate are thus influenced by features of existing aggregates as well as by the behaviours shown by aggregating flies. It is important to note that fly behaviour often varies across assay environments and thus the relevance of these factors is likely to differ across assays.

1.3.2.2. Genetic and neuronal bases of aggregation:

Since aggregation is an outcome of different individual behaviours, neuronal mechanisms underlying aggregation are likely to be quite complex. Different aspects of aggregation, such as inter-individual distance, preferred number of associates and stability of associations, may be regulated by different neural circuits, which may function with varying degrees of independence. Hence, it is important to note that most studies discussed here have quantified aggregation using nearest neighbour distance only, due to which, their results may apply only to the regulation of inter-individual distance.

Mutant based studies have long been used to understand the genetic and neuronal underpinnings of behaviour in *Drosophila*. Such studies have revealed the existence of considerable genetic variation for nearest neighbour distance across different *Drosophila* strains (McNeil et al., 2015) and

mutants (Anderson et al., 2017). Such studies have also been useful for identifying the sensory modalities that are required during aggregation. Although olfaction is known to be important for long-range aggregation via pheromonal cues such as cis-vaccenyl acetate (cVA) (Bartelt et al., 1985; Xu et al., 2005), it is not essential for aggregation at smaller spatial scales (Simon et al., 2012). Short-range aggregation instead seems to depend mainly on vision, as mutants showing low visual acuity show larger nearest neighbour distances (Simon et al., 2012, Burg et al., 2013). Non-sensory mutations, such as those at the *foraging* locus, have also been found to affect aggregation behaviour. In a rare study where individual-level tendencies to join and stay in aggregates were measured, Philippe et al. (2016) found that $for^{s}$ flies showed a preference to join and stay in larger aggregates while $for^{R}$ mutants did not. Since these mutations are known to affect several social phenotypes in flies (Kohn et al., 2013; Foucaud et al., 2013), the *for* locus may be involved in establishing an internal state for sociability, which may consequently affect social tendencies such as aggregation.

Although the higher order processing circuits involved in aggregation behaviour have not yet been identified, neurogenetic studies have identified some neurotransmitters and neuronal subtypes that are important for aggregation. Dopamine secreting neurons seem to be involved in aggregation, as modification of dopamine levels is known to change nearest neighbour distances in a sex-specific manner (Fernandez et al., 2017). These sex-specific responses are surprising because nearest neighbour distances have been reported to not vary across sexes (Simon et al., 2012). These results may be explained if dopamine levels vary across sexes in a environment-specific manner, such that sexual dimorphism in aggregation is visible only under some environmental conditions. Other neuronal types such as cholinergic and glutamatergic neurons have been reported to be essential for mediating the effects of halothane, an anaesthetic which is known to increase nearest neighbour distances in flies. Mushroom body neurons have also been found to be essential for mediating such effects of halothane (Burg et al., 2013). Unfortunately, we do not yet know the specific neurons

among these different subsets that may regulate aggregation. We also lack an understanding of the processing performed by these subsets. Further studies are thus required to identify specific subsets of neurons that may be important for processing different sensory inputs and for regulating tendencies to aggregate.

Since aggregation is an easy behaviour to assay, it has been widely used to study social behaviour in models of neurophysiological disorders in *Drosophila.* Such studies have demonstrated the effects of mutations in *rugose* (Wise et al., 2015) and FoxP genes (Castells-Nobau et al., 2019) as well as early exposure to bisphenol A (Kaur et al., 2015) on nearest neighbour distances in flies. Although the mechanisms underlying these effects are unclear, they are likely to involve errors in the development of neurons underlying aggregation behaviour. These models may thus help us how the circuitry underlying aggregation may develop.

*1.3.3. Ontogeny of aggregation in flies*

Mechanisms of learning allow an organism to vary the intensity or nature of responses to different stimuli depending on its past experiences with these stimuli. As a consequence, behaviours often show variation over the course of an organism's lifetime. Such experience-dependent changes have been documented for several fly behaviours (Ueda and Kidokoro, 2002; Ganguly-Fitzgerald et al., 2006; Kacsoh et al., 2018) including aggregation. Simon et al. (2012) tested the effect of early social exposure on aggregation by regulating access to male and female conspecifics for focal flies. They found that mated flies and group-housed flies showed smaller nearest neighbour distances compared to virgin flies and singly-housed flies respectively. In addition to these, age also affects nearest neighbour distances between flies as younger (<30 days old) flies tend to have shorter values than older flies (Brenman-Suttner et al., 2018). While these changes have been shown to be linked to ageing, it is unclear if they result from degeneration of the neural machinery underlying aggregation, or from age-related changes in behaviour. Interestingly, effects of age are trans-

generational as parental age can influence nearest neighbour distances in the progeny (Brenman-Suttner et al., 2018). These parental effects are likely to be developmental in nature as they are not transmitted beyond the first generation. Curiously, these effects of age have been identified in several lab adapted strains but not in wild caught ones (Brenman-Suttner et al., 2018). Effects of ageing could thus be an unintended consequence of lab adaptation, but further study is required to verify this. Taken together, these results show that mating, social experience and age contribute to ontogenetic changes in nearest neighbour distances. The biological relevance of these changes, however, remains unclear.

*1.3.4. Functional bases of aggregation*

As was discussed in Section 1.1., aggregation is typically brought about by ecological factors such as predators, resource distributions and social environments. Although the role of these factors has not been systematically evaluated in *Drosophila*, evidence from different studies may be pieced together to speculate on the function of aggregation in flies.

Since most studies in flies have tested and observed aggregation in the absence of predators, predation alone seems insufficient to explain aggregation behaviour in flies. While it is possible that aggregation may serve as a predator avoidance strategy under more natural settings, anecdotal evidence from Soto-Yeber et al. (2018) do not lend support to this hypothesis. These authors observed considerable variation in aggregation patterns across different geographically proximate fruit orchards. However, this variation did not match the variation in the abundance of predatory ants across these orchards. Since their observations were only preliminary, further study is required to verify test the role of aggregation as a strategy for predator avoidance.

Aggregation is not dictated by patterns of resource distribution either, as aggregation occurs even in the absence of resources and other environmental heterogeneities (Navarro and del Solar, 1975; Simon et al., 2012). While this suggests that resources patterns may not be essential for

aggregation, they may still shape patterns of aggregation by affecting levels of competition among flies. However, such effects of resource patterns have not yet been tested.

Unlike predation and resource distributions, the presence of conspecifics is sufficient for flies to aggregate, which suggests that aggregation may serve some social function. However, the nature of these social benefits remains unclear. These benefits may likely result from social behaviours among flies that ultimately contribute to an individual's fitness. Since males and females experience distinct social ecologies, social behaviours, and the benefits associated with them, may also vary across sexes. Hence, the possible roles of aggregation in males and in females have been discussed separately below.

It is important to note that while predation and resource distributions may not be necessary for aggregation, they may still influence aggregation patterns directly or indirectly by modifying the social benefits of aggregation. Thus, a systematic study of all these factors remains essential.

1.3.4.1. Aggregation in females:

Gregarious oviposition, i.e. the tendency to lay eggs near one another, is a well studied group behaviour in *Drosophila* (del Solar and Palomino, 1966). Such clustered egg laying is thought to be adaptive as it increases the density of larvae on a food patch. Such increases in density can improve larval survival by reducing the per capita risk of parasitism (seen in *D. pseudoobscura* by Rohlfs and Hoffmeister, 2004) and/or by increasing the efficiency of feeding (Dombrovski et al., 2017). Since grouped egg laying may require flies to gather in space, aggregation has been commonly thought to mediate this behaviour. However, previous studies fail to provide clear evidence for such a relationship. Anecdotal reports suggest that gregarious oviposition may not require aggregation, as flies do not lay eggs simultaneously (del Solar and Palomino, 1966). Instead, they are known to use pheromonal cues, such as cVA, to identify and visit patches that have been used by other flies (Wertheim et al., 2006; Sarin and Dukas, 2009). In contrast, Ruiz-Dubreuil et al.

(1994) observed correlated evolution of female aggregation patterns in response to selection against gregarious oviposition. These seemingly conflicting results suggest that aggregation and gregarious oviposition may not be causally linked, but may instead show genetic correlations i.e. may share common genetic variation. However, it is important to note that the results discussed here are either anecdotal (del Solar and Palomino, 1966) or suffer from limitations such as lack of population-level replication (Ruiz-Dubreuil and del Solar, 1986). Thus these inferences are only preliminary and further systematic studies are required to verify the relationship between aggregation and gregarious oviposition.

Besides gregarious oviposition, other forms of collective behaviour have also been reported to occur in *Drosophila*. Flies in large groups have been shown to find food patches faster (Lihoreau et al., 2016), choose between palatable and unpalatable food more easily (Tinette et al., 2004) and also avoid noxious stimuli better (Ramdya et al., 2014) as compared to flies kept in smaller groups. It is important to note that these behaviours are not social in nature as they can also be performed by single individuals. However, the presence of other individuals seems to make these behaviours more efficient. Flies may thus choose to stay near each other, i.e. may co-ordinate their movement, to facilitate such collective behaviour. Some evidence for co-ordinated movement has been previously reported in *Drosophila*. Soto-Yeber et al. (2018) found that female flies arrived on grapes in groups of 2-3 individuals. Of these, one fly was often seen to scan the surface of the grape using its abdomen while the others stayed motionless. This behaviour seems to be important for selecting oviposition sites, as 15% of the visited grapes contained eggs or larvae and these pre-imaginal stages were almost always (~99%) homospecific (i.e. of the same species) despite the presence of several *Drosophila* species at the field site. Co-ordinated movement may also be involved during dispersal, as patterns of dispersal have been found to be consistent with movement of fly groups instead of single individuals (Lefranc et al., 2001). Overall, these studies suggest that several *Drosophila* behaviours may occur collectively, and may thus require aggregation.

However, as the prevalence of such collective behaviours is unknown, the adaptive value of such behaviours remains unclear.

Social learning, i.e. the ability of individuals to gain and employ information from each other, is another form of social behaviour that is known to occur in *Drosophila*. Naive flies are known to suppress oviposition after being exposed to females who have been exposed to parasitoid wasps (Kacsoh et al., 2018). Similarly, oviposition site preferences are also known to be transmitted socially (Sarin and Dukas, 2009). Exchange of such information also occurs in the absence of oviposition substrates (Battesti et al., 2012), which suggests that flies may actively communicate their preferences and may not simply mimic each other. Social learning has also been demonstrated for behaviours such as mate choice (Mery et al., 2009) whereby females preferentially mate with males that look similar to males chosen by other females. Such 'mate-copying' can ultimately give rise to 'mate-choice traditions' in fly populations via cultural transmission of information over several generations (Danchin et al., 2018). These results, thus, suggest that flies may use social information more commonly than previously expected. While the mechanisms underlying suchsocial learning are largely unknown, recent studies have shown that it depends on the ecology of information transfer. For example, flies are seen to conform to majority opinion while choosing mates (Danchin et al., 2018). Similarly, the efficiency of transmission for oviposition preferences also depends on the homogeneity in preferences of a fly's associates (Battesti et al., 2015; Pasquaretta et al., 2016). Thus, flies may need to access several individuals to gain social information for accurately modifying their behaviour. Consequently, aggregation may be essential for mediating the spread of social information among flies.

1.3.4.2. Aggregation in males:

As in case of females, it is helpful to look at male aggregation in the context of male social behaviours. Social interactions between males are largely aggressive as males exhibit territoriality

in the presence of food and females (Jacobs et al., 1960). They defend individual patches of food by fighting intruder males and driving them away. Most males thus tend to be alone, yielding overdispersed spatial patterns for males in the presence of food (Foley et al., 2015). Such defensive behaviour is adaptive as it confers mating benefits on territorial males (Hoffmann and Cacoyianni, 1990; Saltz and Foley, 2011). Incidence of aggression depends strongly on the density of flies in a given environment (Hoffmann and Cacoyianni, 1990; Saltz and Foley, 2011). As the number of flies per patch increases, a single male may not be able to defend large resource patches or may be forced to defend smaller patches. In both cases, the mating advantage gained by defending resources is likely to reduce, consequently reducing the frequency of such encounters (Hoffmann and Cacoyianni, 1990). Interestingly, increase in the density of flies also increases the mean and, more importantly, the variance of male group sizes. These results suggest that males actively start aggregating, instead of moving randomly, following a reduction in aggressive behaviour. Aggregation may thus serve as an alternative strategy to aggression under conditions where resource defense does not yield mating benefits.

Several taxa are known to form such male aggregations, called leks, where males engage in aggressive interactions with each other and/or exhibit courtship displays towards females (Alcock, 2001, p. 385). Such lekking behaviour is known to occur in several species of *Drosophila* (Shelly, 1987; Hodosh et al., 1979; Aspi and Hoffmann, 1998) but it remains poorly studied in *Drosophila melanogaster*. Under field conditions, male-biased aggregations of *D. melanogaster* are seen to form near, but not on, food sources (Taylor and Kekic, 1988; Soto-Yeber et al., 2018). Male flies in these aggregations do not show aggression but do court females that approach the aggregates (Taylor and Kekic, 1988). Consequently, most matings are also seen to occur near such aggregates. Although these observations indicate correlations between mating and aggregation, they do not explain how aggregation may influence male mating success. Thus, to verify if aggregation indeed serves as a mating-related strategy, it is necessary to measure the mating benefit accrued to males

as a consequence of aggregation. These benefits may be understood in the context of the following hypotheses that have been proposed to understand why leks may form (Alcock, 2001, p. 386-389) :

- Female preference hypothesis : Females prefer to mate with aggregated males.

- Hotspot hypothesis : Males maximise access to females by gathering near locations frequented by females such as food sources.

- Hotshot hypothesis : Less successful males gain access to females by being near 'hotshot' males i.e. the most attractive or successful males.

Tests of these hypotheses in other *Drosophilids* suggest that lekking seems to benefit males in species specific ways (Shelly, 1987; Droney et al., 1994; Aspi and Hoffmann, 1998). Although these hypotheses have been not been tested in case of *D. melanogaster*, available data suggest that the female preference and hotspot hypotheses may be insufficient to describe male aggregation. The female preference hypothesis is not supported as females do not show any preference for males from larger groups (Taylor and Kekic, 1988). The hotspot hypothesis also fails to explain aggregation as male aggregates are seen to form even in the absence of resources that may attract females (Simon et al., 2012). Although the hotshot hypothesis is not refuted by the available data, it still lacks supporting evidence. Since females are capable of identifying and choosing specific males (Mery et al., 2009), other males may also be able to identify the preferred individuals amongst themselves. While it is unclear if less preferred males actually stay near preferred ones, they may be able to improve their access to females by interrupting courtship between females and the preferred males (Taylor and Kekic, 1988). The efficiency of such a strategy, however, remains unclear.

Aggregation may also occur as a result of behaviours such as 'co-operative male courtship'. This behaviour was described under field conditions by Soto-Yeber et al. (2018) who observed that grouped males often followed a courting male while he chased a female. These males tended to

disperse if the courting male ended the chase. However, if courtship was successful, then these males were seen to stay and vibrate their wings near the mating flies until copulation was completed. Although the authors called this behaviour co-operative, it is unclear if males actually aid each other in their mating efforts. Further study is required to understand the relevance of such behaviour and to understand its relationship with aggregation.

Aggregation may also provide non-sexual benefits to males by facilitating collective behaviours such as those discussed for females. Behaviours such as collective foraging (Lihoreau et al., 2016; Tinette et al., 2004) and dispersal (Lefranc et al., 2001) are similar in case of both males and females. Similarly, males also show social learning of behaviours such as odour avoidance (Kohn et al., 2013). Since these traits are not sex-specific, any fitness benefits associated with them may thus promote aggregation in both sexes.

*1.3.5. Phylogeny of aggregation*

To my knowledge, only one study, Schultzaberger et al. (2019), has compared spatial patterns of aggregation across *Drosophilids*. The authors found that kinetics of aggregation on food, as well as its sensitivity to fly density (i.e. the extent of collective behaviour *sensu* Tinette et al., 2004) varied widely across ten *Drosophilid* species. This variation seems to correlate with foraging ecologies of these species. For example, specialist species (*D. erecta*) that encounter temporally patchy resource distributions showed rapid rates of arrival on food that did not depend on fly density, while specialists (*D. arizonae*, *D. sechellia*) that encounter abundant resources showed slow rates of arrival (effect of density was not reported). In contrast, generalists (*D. melanogaster*, *D. simulans*, *D. willistoni*, *D. pseudoobscura*) showed intermediate rates of arrival on food which were sensitive to density. These results highlight that resource distributions may indeed exert strong selective pressures on aggregation patterns. Additionally, the authors found that mean nearest neighbour distances and aggression levels varied across species, although they did not report any overall

relationship between these traits. They also found considerable inter-species variation in the effects of social isolation on these traits. Since this study used only single strains to represent a species, these results are sensitive to inter-strain variation. Nevertheless, these results suggest that aggregation behaviour, as well as its correlations with other social behaviours, may evolve differently across species.

### *1.4. Scope of the present study*

Aggregation in *D. melanogaster* has been studied under a variety of contexts. While these studies have yielded a trove of information, variation in their approaches to quantify aggregation crucially limits our ability to interpret these results. Hence, the aim of my study was to develop an approach that could be used to comprehensively study how and why flies may aggregate. Further, I wanted to understand the role of social environment in shaping aggregation behaviour.

To achieve the first objective, it was necessary to define and identify individual aggregates formed by flies. For this, I used the Ripley's K method for spatial point analysis to empirically estimate proximity among flies. and used these data to cluster proximate individuals into aggregates. After defining such aggregates, I quantified their properties such as size, density, duration and composition to obtain a more detailed description of the aggregation pattern. In an attempt to understand the processes that underlie aggregation, I quantified the behaviour of individual flies in terms of their tendencies to join and stay in aggregates. Finally, I made preliminary observations of fly behaviour within these aggregates to try and understand any functional aspects of such aggregation. To understand the effect of social environment on aggregation, I varied the mating status and prior social experience of flies before assaying their aggregation behaviour. These factors are known to influence nearest neighbour distances in flies (Simon et al., 2012) and were thus expected to affect different aggregate-level properties and individual-level behaviours.

My results suggest that variation in aggregation patterns may be explained by the tendencies of flies to join and stay in aggregates. However, these tendencies may be only partially explained by social interactions between flies. My results also suggest that flies show non-random patterns of association with each other, which may be indicative of some form of social relationships among them. Finally, I found that both mating and prior social experience had distinct and largely independent effects on different aspects of aggregation behaviour.

# MATERIALS AND METHODS

## *2.1. Fly maintenance*

Canton-S flies were used for all experiments. These flies have been maintained under laboratory conditions for several generations in vial cultures that follow a fourteen day generation cycle. Prior to the experiment, 300-500 adult flies were transferred to a small Plexiglas cage (20.4 cm x 16.4 cm x 13.6 cm) and a 30 mm Petri plate, containing standard cornmeal medium with added charcoal, was provided to these flies for feeding and egg laying. These flies formed the parental population from which eggs were collected to derive the experimental flies. Since, the fecundity of the parental flies was found to reduce after about nine days after transfer into the cages, one last batch of eggs was collected from them to establish a second parental population. Experimental flies for subsequent replicates were derived from this second population. Both parental cages were maintained at a mean temperature of 25°C under a strict 12:12 light-dark (LD) regime.

## *2.2. Exposure of flies to different social environments*

30 mm Petri plates containing charcoal food were placed inside parental cages for eight to nine hours to obtain eggs for a single cohort of experimental flies. These eggs were allowed to complete development in vials containing approximately 6 ml of standard cornmeal medium (without charcoal) at a density of approximately sixty eggs per vial. Around the $9^{th}$ day post egg collection, all darkened pupae were transferred to Petri plates using a paint brush moistened with water. Since these flies had completed pigmentation, the sex combs present on the forelimbs of males could be seen through the pupal case. These sex combs were used to identify and separate male pupae. The female pupae were then transferred to new vials where they were allowed to emerge. Density of

females in these vials was controlled to provide different levels of social exposure to the emerging flies. Social enrichment was provided by housing approximately thirty female flies per vial while deprivation was provided by housing female flies in isolation. Two high density vials and sixty low density vials were used to ensure similar number of flies for each treatment. These vials were maintained under constant light conditions at 25°C until the $11^{th}$ day post egg collection.

On the $11^{th}$ day, males were anaesthetised using carbon dioxide and introduced into half of all vials for each density. Thus, one vial of the socially enriched treatment received thirty males while thirty vials of the socially deprived treatment received one male each. The flies were allowed to mate for 24 hours after which the males were separated using ice anaesthesia. To control for the effects of anaesthesia, ice anaesthesia was also provided to virgin flies. After separation, the flies were left undisturbed for a day after which their aggregation behaviour was assayed. A single cohort of eggs thus yielded four sets of flies, each of which was exposed to a different social environment (Refer Figure 2.1. for summary of the experimental scheme).

| Treatment | Social environment |
|---|---|
| Mated and Socially Enriched (ME) | Exposed to males for one day ($11^{th}$ day) and housed with other females for three days ($10^{th}$-$12^{th}$ day) |
| Mated and Socially Deprived (MD) | Exposed to males for one day ($11^{th}$ day) and housed alone for three days |
| Virgin and Socially Enriched (VE) | Not exposed to males and housed with other females for three days ($10^{th}$-$12^{th}$ day) |
| Virgin and Socially Deprived (VD) | Not exposed to males and housed alone for three days |

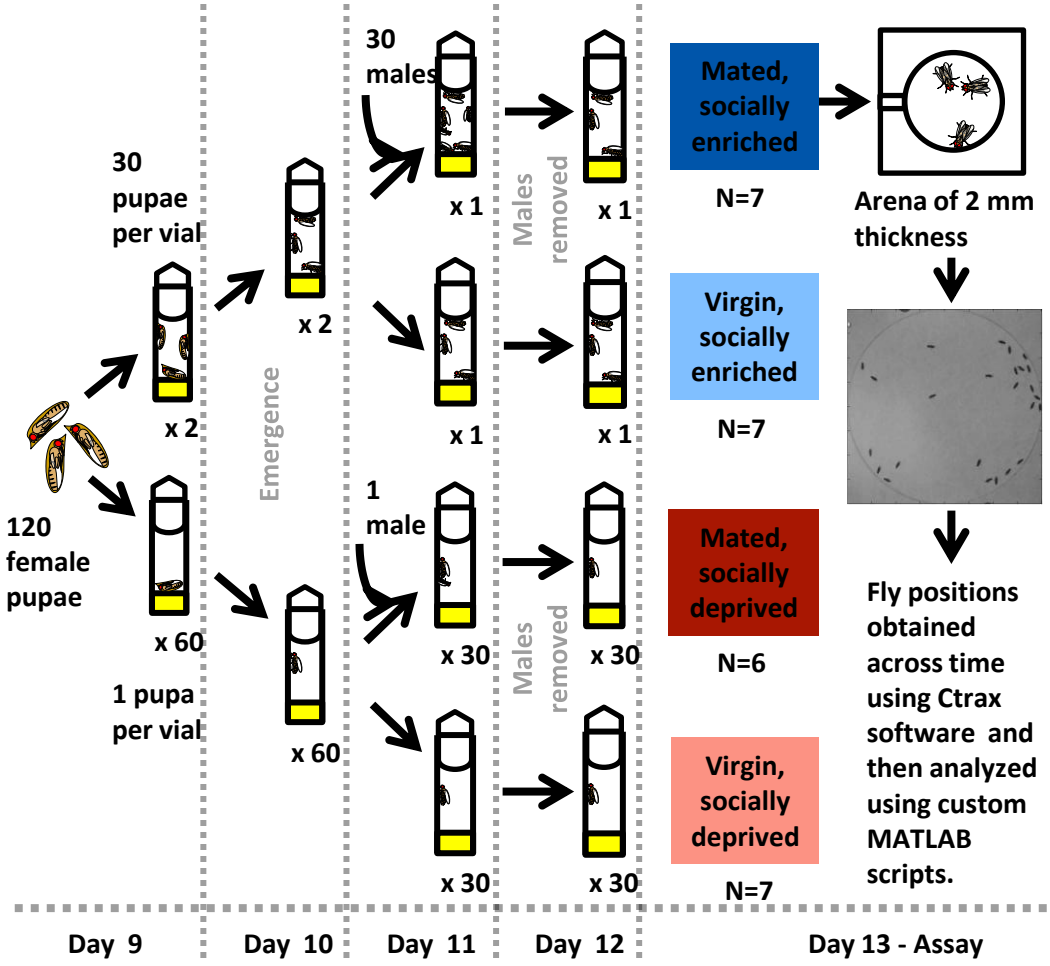



**Figure 2.1. Scheme of a single replicate aggregation experiment.**

The effect of social environment on female aggregation behaviour was tested using a fully factorial design with mating status and social condition as fixed factors. Mating status was determined by a female's exposure to males post emergence. Mated flies were housed with males for a single day while virgin flies were kept isolated from males. Social condition was determined by a female's exposure to other females post emergence. Enriched flies were exposed to other females while deprived flies were isolated from other females until the assay. A combination of these two factors yielded four distinct social environments- Mated and socially enriched (ME), mated and socially deprived (MD), virgin and socially enriched (VE) and virgin and socially deprived (VD). A single cohort of 120 female flies was divided equally and exposed to each of these environments before being assayed for their aggregation behaviour.

All sets of flies were assayed on the 13$^{th}$ day post egg collection. However, as it was not possible to set up multiple assays at the same time, each set had to be assayed at different times of the day. To ensure that all flies were assayed at the same phase of their daily activity rhythm, vials containing darkened pupae were first maintained under constant light conditions (LL) until the 11$^{th}$ day post egg collection. Then, on the 11$^{th}$ day, each set of flies was transferred to separate light boxes. Each box maintained a 12:12 light-dark cycle but timings of the light-dark transitions varied across boxes. As flies adjust their activity behaviour with reference to these transitions, variation in the timing of these transitions allowed different sets of flies to reach the same phase of their activity at different local times. Thus, by setting the time of lights on (called Zeitgeber Time 0 or ZT 0) to different local times for different boxes, different sets of flies could be assayed at different local times but at the same phase of their activity rhythms. To avoid any confounding effects associated with local time, the order in which each social environment was assayed through the day was randomised.

On the day of the experiment (13$^{th}$ day post egg collection), flies were habituated to the assay room for two hours prior to the assay. This room was maintained at 25°C and was illuminated by a single source of white light. Aggregation behaviour of flies was assayed from ZT 5 to ZT 7 which is when flies show low levels of locomotor activity.

### *2.3. Aggregation assay*

The assay arena consisted of three 2 mm thick sheets of Plexiglas held together using binder clips. A circle of radius 4 cm was cut from the centre of the middle sheet to create an empty circular space. Flies could be aspirated into this space via a canal cut from the edge of the middle sheet to the circle. This canal was plugged using the same cut piece of Plexiglas after introduction of the flies (Fig. 2.2).

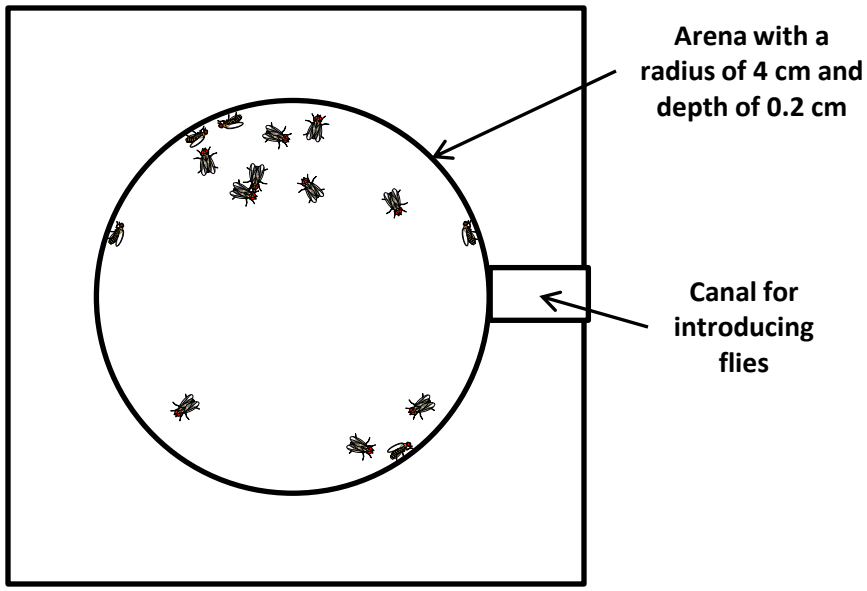



**Figure 2.2. View of the arena during video recording.**

The arena used to assay the aggregation behaviour of flies is depicted above. It was comprised of a circular space with a radius of 4 cm and a depth of 0.2 cm that was enclosed between two transparent pieces of Plexiglas. Flies were aspirated into the arena through a canal cut from the circular space to the edge of the arena. After introduction of the flies, this canal could be plugged using the same cut piece of Plexiglas. The arena was placed horizontally and provided with uniformly illuminated from above. A camera was placed below the arena and movement of the flies within the arena was recorded for two hours.

The arena was suspended horizontally above a Nikon D5500 camera with a macro lens (AF-S VR Micro-Nikkor 105mm f/2.8G IF-ED). A white paper was placed above the arena to evenly disperse light from the overhead source of white light. The camera was connected to a computer using a USB 2.0 cable and recordings were made directly to the computer using SparkoCam 2.4.1 software. Fly behaviour within the arena was recorded for 130 minutes after which the flies were discarded. The arena was cleaned thoroughly with soap and finally with 95% alcohol before reuse.

All videos were recorded at a resolution of 1920 x 1080 sq. pixels and frame rate of approximately thirty frames per second. However, as videos were saved directly to the computer, frame rates were affected by the rates of transmission and were thus seen to vary slightly. As this deviation was not expected to affect any measurements significantly, a frame rate of thirty frames per second was assumed for all analyses. All videos were saved in the .wmv format.

### *2.4. Video processing*

The MATLAB based Caltech multiple walking fly tracker or Ctrax (Branson et al. 2009) was used for tracking fly movement in the recorded videos. As Ctrax requires video files to be in the micro fly movie format (.ufmf) all videos were first uncompressed to .avi format using Virtual Dub 1.10.4. The first and last five minutes of the video were discarded to ignore any disturbance due to handling. As the size of the uncompressed file was prohibitively large, each video was divided into half hour long .avi files. To further reduce size, the audio component of the video was removed and the video was converted to greyscale. Finally, the video was cropped to exclude everything outside the arena and then converted to .ufmf using the any2ufmf video converter provided by the Ctrax developers. Each video was thus saved as four separate half hour long .ufmf files.

Ctrax tracked fly movement in these .ufmf files by identifying each fly and tracing its position across frames. For this, the software identified all contiguous dark (relative to the background)

pixels as potential flies and fit ellipses to them. Ellipses that fit certain size and shape parameters were classified as flies while those that didn't were discarded. Each fly's position was traced across time using a velocity model, which matched the position of each ellipse in a frame to the most probable position in the next frame. Thus, each fly could be uniquely traced across time. Position of a fly was measured as the co-ordinates of the centre of its ellipse while its size was measured using the lengths of the ellipse's major and minor axes. A fly's orientation was inferred from the direction of the its velocity.

Occasionally, identities were seen to switch between flies when they were close to one another. To identify and fix these errors, the FixErrors MATLAB GUI provided with the Ctrax software could be used. Unfortunately, FixErrors was found to ignore several mismatch errors in the videos. To overcome this problem, custom MATLAB scripts were used to identify suspicious changes within the data, such as when flies showed abrupt orientation changes near each other. To ground truth these scripts, all errors were scored manually for the first half hour of a randomly chosen replicate video for each treatment. Unlike the FixErrors GUI, these scripts were able to identify almost all observed errors. Thereafter, these custom scripts were used to identify errors in each video and FixErrors was used only to correct the identified errors. After correction, the software yielded per-frame data for the positions, orientations and physical dimensions of each fly as its output.

### *2.5. Behavioural Analysis*

For performing behavioural observations, the BehavioralMicroarray MATLAB Toolbox provided with the Ctrax software was used. The showtrx function was used to visualise all flies in a video along with the identities assigned to them by Ctrax. The sampling procedure for these analyses will be discussed later in Section 3.4 of Results and Inferences.

### *2.6. Statistical analyses*

Although the assay was performed several times for each treatment, only six to seven usable videos could be obtained per treatment. This attrition was due to technical failures during the experiment or due to compatibility errors during data processing. All analysis has thus been performed only using the error-free replicates.

Position data obtained from these replicates were used to quantify different variables (discussed under Results and Inferences). These variables were analysed using the analysis of variance (ANOVA). To test the effects of mating and social experience on aggregation, these were included as fixed factors in the analysis. To understand the evolution of aggregation behaviour across time, a video was divided into five minute long intervals and each variable was measured in each of these intervals. Since these measurements were performed on the same set of individuals at different time points, time interval was included as a repeated measure during analysis. Simple two-factor ANOVAs were used for analyses where time was not a factor. Any variable that was found to violate the assumptions of normality or homoscedasticity was transformed appropriately before analysis. In cases where transformations failed, non-parametric tests were used for analysis. However, most graphs were plotted using untransformed data.

Most analyses were performed using Statistica v.7 software. For tests that could not be performed using Statistica, resources provided by McDonald (2014) were used. The Benjamini-Hochberg procedure for controlling false discovery rates was carried out in Microsoft Excel 2007, as described in McDonald (2014). Welch's ANOVA and Games-Howell tests were also performed using excel resources freely provided by McDonald (2014). Fisher's exact tests for 2 x 4 and 2 x 2 contingency tables were performed using online resources provided by http://www.physics.csbsju.edu/stats/exact_NROW_NCOLUMN_form.html following references by McDonald (2014).

# RESULTS AND INFERENCES

## *3.1. Quantifying spatial patterns of aggregation in Drosophila*

### *3.1.1. Rationale*

Previous studies that have characterized short-range aggregation in *Drosophila* have used different methods to quantify aggregation. As discussed previously, a popularly used measure for quantifying aggregation is the mean distance to the nearest neighbour. Since the nearest neighbour distance may be thought of as the distance at which attractive and repulsive forces between individuals balance out (Mogilner et al., 2003), it serves as a useful estimate of the preferred degree of proximity for an individual. Although nearest neighbour distance is an individual level measure, previous studies have used it to quantify the overall pattern of aggregation by simply averaging it across individuals that make up the aggregation pattern. While such an approach is not necessarily problematic, this approach fails to account for emergent properties of the group behaviour such as the number of neighbours for an individual i.e. the size of an aggregate. This may be illustrated by the simple example in Fig. 3.1.1. Patterns A and B have the same distance between nearest neighbours. Yet they represent different aggregation patterns due to the difference in the distribution of points around a given point. Such a difference may result from differences in the underlying biological context. For example, Pattern A may be obtained if flies approached each other to engage in purely pair-wise interactions such as aggression. In contrast, B may result from multi-fly interactions such as collective behaviours. Such differences may be ignored as a consequence of relying purely on nearest neighbour distances.

To overcome these limitations, and thereby better describe *Drosophila* aggregation patterns, I used the Ripley's K method for spatial point analysis. This method measures the average numbers of flies at different distances from an individual and tests if they tend to be different from those

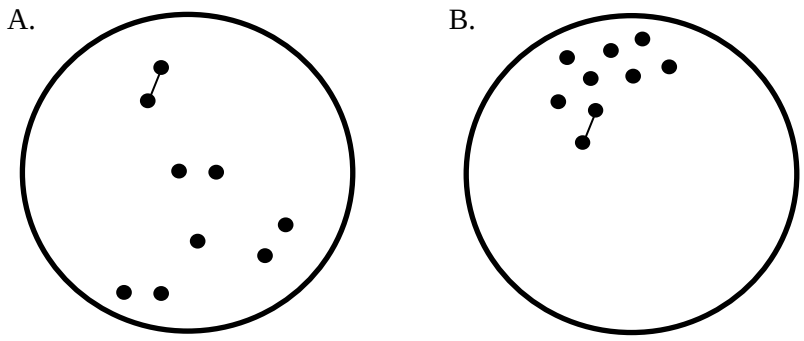



**Figure 3.1.1. Comparison of spatial patterns showing differences in aggregation without differences in nearest neighbour distance.**

Points in **A)** are clustered largely in pairs that are dispersed throughout the available space while points in **B)** are all clustered together in one portion of the available area. Nearest neighbour distance is the same in both cases (indicated by the line joining neighbouring points).

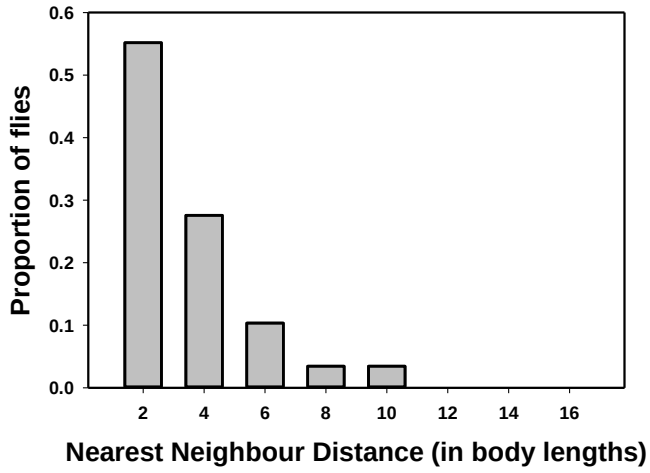



**Figure 3.1.2. Distribution of nearest neighbour distances in a given frame.**

Distribution of nearest neighbour distances typically showed positive skew. Social space index was quantified as the difference between the first two bins of the above histogram.

expected by chance alone. Using this method, I could quantify the distances over which non-random aggregation occurred and also measure the number of flies at these distances. I used these parameters to describe aggregation patterns of flies exposed to different social environments. I also quantified aggregation for these flies using nearest neighbour methods to test for concordance between these methods.

*3.1.2. Methods*

Pair-wise distances between flies were calculated in each frame of each replicate assay using the position values obtained after tracking. These distances were scaled by the average body length observed in a given replicate before performing any analysis. Since fly positions, and thereby distance values, are correlated across consecutive frames, frames were sampled for analysis at intervals of 300 frames (or 10 seconds). This interval was chosen as the expected change in position of a fly over this interval was similar to the change in position between any two randomly chosen points. Different variables were measured (discussed below) using data from these frames and analysis was performed only on these variables. The entire video was divided into five minute intervals and all sampled frames occurring within a given five minute interval were considered to be replicate observations for that interval. By averaging a given variable across such replicate frames, mean values for each five minute interval could be obtained. All analysis was ultimately performed on the time series of these values for each variable.

3.1.2.1. Nearest neighbour based methods:

Using the list of pair-wise distances for each sampled frame, the nearest neighbour could be identified for each fly. Since distributions of nearest neighbour distances were skewed (Fig. 3.1.2), median values were used to represent central tendency. Median nearest neighbour distance was measured for each sampled frame and averaged across frames within a five minute interval to obtain the average median nearest neighbour distance for that interval.

Using the distribution of nearest neighbour values, a social space index (SSI) was also calculated for each spatial pattern. The SSI was developed by Simon et al. (2012) as a simple measure to quantify social space. SSI values greater than zero were found to be indicative of non-random social space (Simon et al., 2012).

$$SSI = (fraction\ of\ flies\ showing\ nnd\ values > two\ body\ lengths\ and < four\ body\ lengths) - (fraction\ of\ flies\ showing\ nnd\ values < two\ body\ lengths)$$

The SSI was calculated for each sampled frame (refer Fig. 3.1.2.) and a time series of SSI was obtained, as above, which was used for analysis.

3.1.2.2. Ripley's K method:

The Ripley's K method, in its simplest form, assumes that positions of points in space are a result of complete spatial randomness (CSR) or a two-dimensional Poisson process (Young and Young, 1998). Given this assumption, the expected number of points found in the vicinity of a randomly chosen point can be predicted by simply multiplying the sampled area with the overall density of points for the total area. If, however, the point process is not random and instead is aggregative, then points will tend to be closer to each other, such that the observed number of points around a randomly chosen point will be higher than expected under CSR. Consequently, the area occupied by the observed number of points should be much larger under CSR. The Ripley's K is an estimate of this expected value that may then be compared with the sampled area (illustrated using Fig. 3.1.3a,b). More formally, Ripley's K or *K(r)* represents the area expected to be occupied by the number of points observed within a circle of radius *r* given a density of $\lambda$ for the entire sampling area.

$$K(r) = \frac{E\left(number\ of\ points\ (flies)\ within\ distance\ 'r'\ of\ a\ randomly\ chosen\ point\ (fly)\right)}{Density\ of\ points\ (flies) in\ the\ arena\ (\lambda)}$$

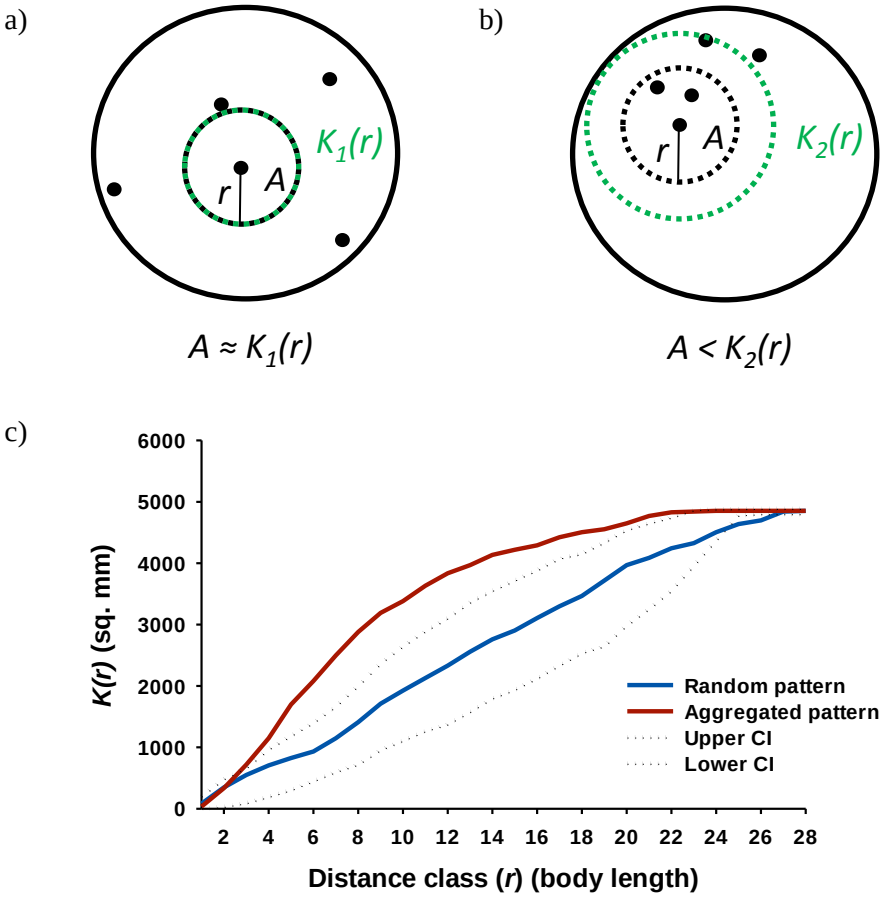



**Figure 3.1.3. Comparison of random and aggregated spatial patterns using the Ripley's K method.**

**a)** For a random pattern, the value of *K(r)* is approximately equal to the area *A* described by distance *r*. **b)** For an aggregated pattern with the same overall density, *K(r)* is greater than *A*, as more points are found within *A* than expected by chance alone. **c)** Dotted lines represent the 5$^{th}$ and 95$^{th}$ percentile values of *K(r)* for different values of *r* obtained using the 2D position histogram for single flies. The blue plot represents a spatial pattern than can be fully explained by spatial preferences of the observed flies. The red plot represents a pattern showing aggregation at distances of 3-22 body lengths owing to the greater than 95$^{th}$ percentile values of *K(r)* at those distances.

[*K(r)* is calculated for the overall pattern and not for a single point. Here *K(r)* is depicted around a point only to illustrate the relationship between *K(r)* and *A*]

A plot of *K(r)* against *r* is called the Ripley's K function which can be calculated for the observed spatial pattern. This observed function can then be compared with the K function expected under CSR. If significant aggregation exists at a distance *r* in the observed data, then the value of *K(r)* would be significantly greater than the random expectation. Similarly, a regularly ordered point pattern would yield *K(r)* values that are significantly smaller than the random expectation.

Since flies are known to have a preference for edges of the arena (Valente et al., 2007), flies may stay near the edge more frequently than towards the interior of the arena. Thus, unlike in case of CSR, probability of a fly being near the edge is higher than the probability of it being near the interior of the arena. Consequently, the distribution of a fly in the arena is fundamentally heterogeneous. To account for this, a null model incorporating such heterogeneity was used for testing significance of the observed K functions. To generate such a null model, movement of 32 single ME flies was recorded in the arena for two hours. Fly positions across these recordings were combined to obtain a two-dimensional position histogram for the arena which reflected the spatial preferences of a single fly in the absence of other flies. Any heterogeneities in the distribution of flies in the arena due to edge preferences of a single fly were thus accounted for. Using this position histogram as a reference, a thousand artificial datasets, containing thirty virtual fly positions each, were generated. Since these datasets represent the spatial patterns that would be expected under purely non-social conditions, K functions for these datasets were used to obtain a 95% confidence band for the K function expected to occur purely as a result of fly preferences for different parts of the arena. An observed pattern was considered to be significantly aggregated only if its K function lay above this confidence band for at least one value of *r* (Fig. 3.1.3c).

K functions were calculated for each sampled frame and tested for significant aggregation. Probability of aggregation over the course of the assay was calculated as the proportion of all sampled frames that yielded non-random K functions. To obtain a time series of aggregation over the course of the assay, the probability of being aggregated in each five minute interval was

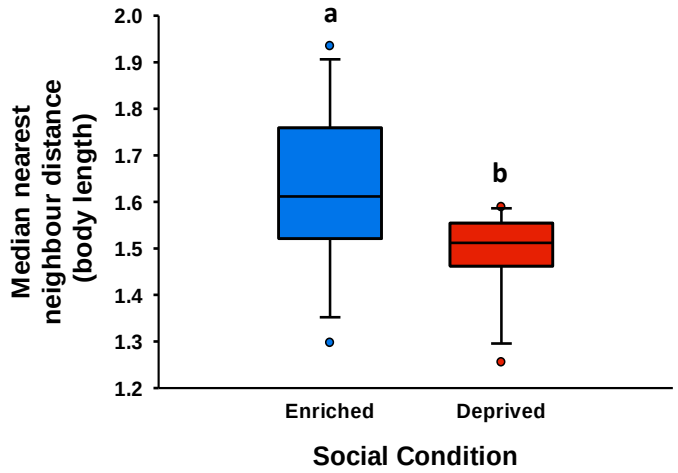
calculated by measuring the proportion of sampled frames showing non-random K functions within that interval. If a K function was found to be non-random, then the smallest and largest distance classes showing significant aggregation ( i.e. *r* values with non-random K values) were identified. The means of these values were also calculated for each five minute interval by averaging across all sampled frames showing significant aggregation within that interval. These values were used to quantify the minimum and maximum scales of aggregation within the observed patterns. The number of flies in each significant distance class was also recorded for each frame showing significant aggregation. Mean number of flies in a distance class was calculated by averaging the number of flies in that distance class across all frames showing significant aggregation for that distance class. Increase in number of flies was calculated for each distance class as the difference between mean numbers of flies in consecutive distance classes.

### *3.1.3. Results*

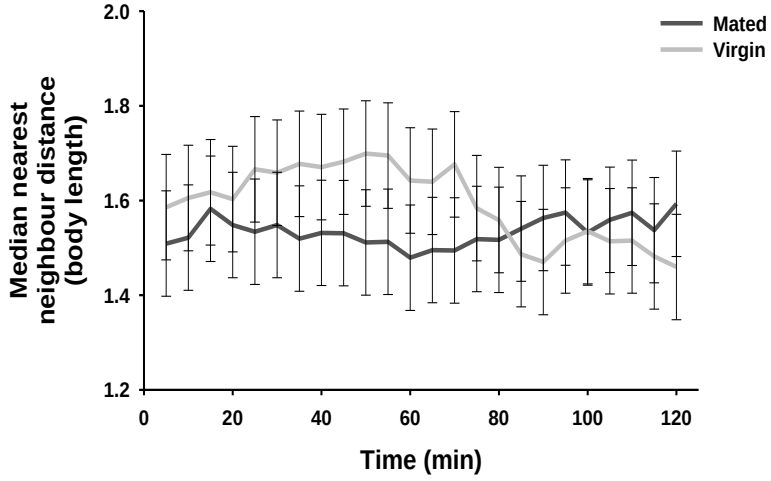
#### 3.1.3.1. Effect of social environment on the distribution of nearest neighbour distances:

Socially enriched flies were observed to show a significantly greater median nearest neighbour distance compared to socially deprived flies ($F=6.257$, $df=1$, $p=0.02$) (Fig. 3.1.4a). A significant interaction effect was also observed between time and mating status ($F=3.067$, $df=23$, $p<0.0001$)(Fig. 3.1.4b). To better understand effects of social environment across time, linear slope was calculated for each replicate time series and analysed using a two-factor ANOVA. No differences in slope were found, suggesting that there was no difference in the trend across time. The interaction between time and mating status was, thus, likely due to the higher values of median nearest neighbour distance seen for virgin flies in the middle third (approximately between 25-75 minutes) of the assay. Visual examination of the time-series for individual treatments suggested that this effect may be primarily due to the behaviour of VE females as VD females were seen to show nearest neighbour distances similar to the other two treatments.

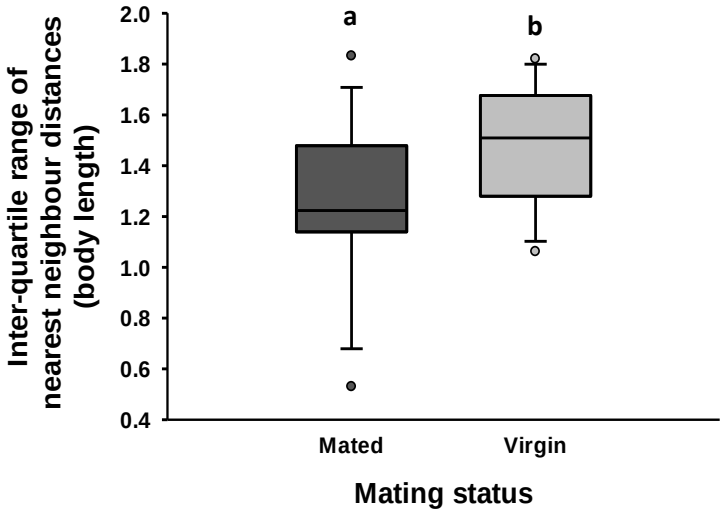
a)


Median nearest neighbour distance (body length)
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
a
b
Enriched
Deprived
Social Condition

b)


Median nearest neighbour distance (body length)
2.0
1.8
1.6
1.4
1.2
Mated
Virgin
0
20
40
60
80
100
120
Time (min)

c)


Inter-quartile range of nearest neighbour distances (body length)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
a
b
Mated
Virgin
Mating status

d)

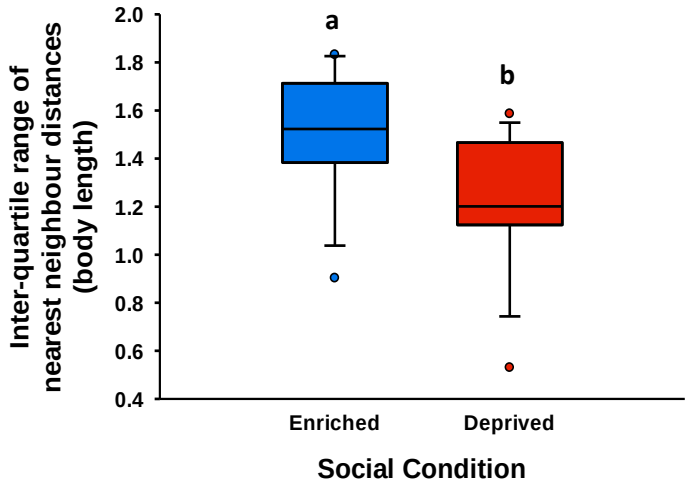



**Figure 3.1.4. Effects of social environment on nearest neighbour distances.**

**a)** Median nearest neighbour distances were smaller for flies maintained in social isolation (Main effect of social condition ($p=0.02$)). **b)** Time profiles of nearest neighbour distance were different between mated and virgin flies (Interaction between time and mating status ($p<0.0001$)) but no significant differences in slope were observed. Variation in nearest neighbour distances was higher for **c)** virgin flies relative to mated ones and for **d)** socially enriched flies relative to deprived flies (Main effect of mating status($p=0.0213$) and social condition ($p=0.0053$)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote 95% CI].

The inter-quartile range (IQR) (difference between 75th and 25th percentiles) of the distribution of nearest neighbour distances was significantly larger for virgin flies ($F=6.1027$, $df=1$, $p=0.0213$) (Fig. 3.1.4c) and for socially enriched flies ($F=9.4615$, $df=1$, $p=0.0053$) (Fig. 3.1.4d).

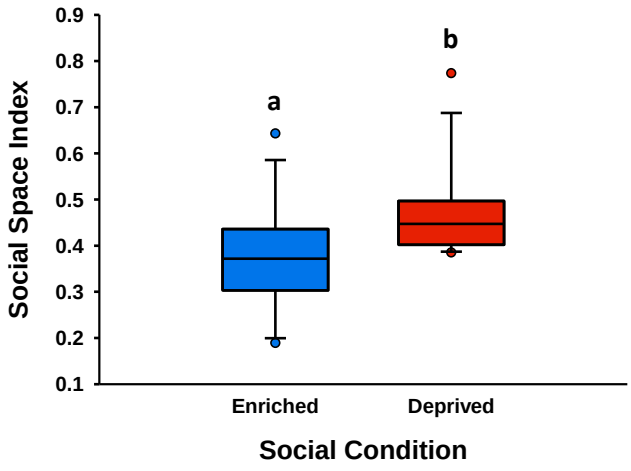
3.1.3.2. Effect of social environment on the social space index:

SSI was significantly larger for socially deprived flies compared to enriched flies ($F=4.9692$, $df=1$, $p=0.036$) (Fig. 3.1.5a). A significant interaction between mating status and time was observed ($F=2.9052$, $df=23$, $p<0.0001$) (Fig. 3.1.5b) but linear slopes were not affected by social environment. Thus, the effect of mating x time was likely due to virgin flies showing smaller SSI than mated flies in the first two thirds (~80 min) of the video.

3.1.3.3. Effect of social environment on probability of significant aggregation:

As probability data were skewed, they were transformed before analysis ($y=\sin^{-1}(\sqrt{x})$). Despite transformation, variances were seen to be heterogeneous. Hence, results of the ANOVA were verified using a Welch's ANOVA and were found to be consistent. The overall probability of aggregation was significantly higher for socially deprived flies compared to socially enriched flies ($F=6.3623$, $df=1$, $p=0.019$)(Fig. 3.1.6a). A similar trend was seen for mated flies compared to virgins but it failed to reach significance ($F=3.6057$, $df=1$, $p=0.07$)(Fig. 3.1.6b). The effect of time on the probability of aggregation could not be analysed as there was considerable variation across replicates in the shape of the time series. As a result, poor fits were obtained when sigmoid curves were used to model these time series. Analysis was thus limited to a visual inspection of these trends. The time profile for the probability of aggregation was visually very distinct for VE flies relative to the other treatments (Fig. 3.1.6c). VE flies showed low probability of aggregation for around one hour of the video after which the probability increased before stabilizing. In contrast, the other profiles showed a weak monotonic increase with time.

a)




b)

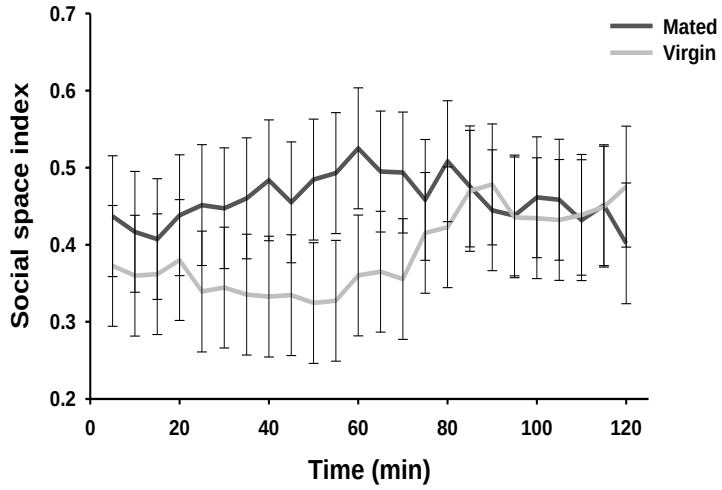
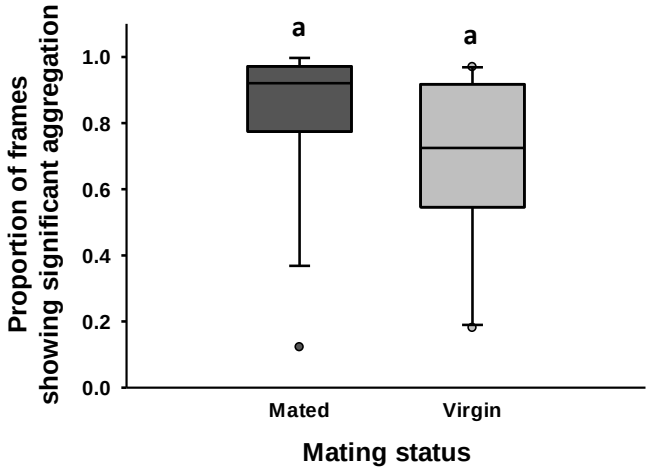



**Figure 3.1.5. Effect of social environment on the social space index (SSI).**

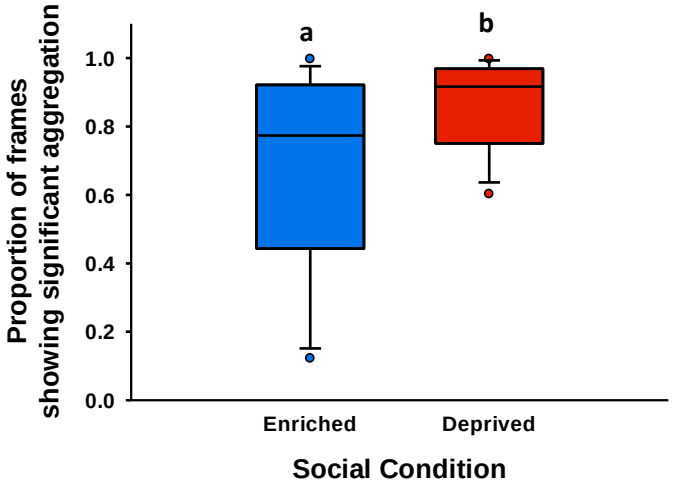
**a)** SSI was significantly higher for socially deprived flies compared to socially enriched flies (Main effect of social condition ($p=0.035$)). **b)** Time profiles of SSI were significantly different between mated and virgin flies (Interaction between mating status and time ($p<0.0001$)) but no significant differences in slope were observed.

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote 95% CI].

a)


Proportion of frames showing significant aggregation
1.0
0.8
0.6
0.4
0.2
0.0
a
a
Mated
Virgin
Mating status

b)


Proportion of frames showing significant aggregation
1.0
0.8
0.6
0.4
0.2
0.0
a
b
Enriched
Deprived
Social Condition

c)

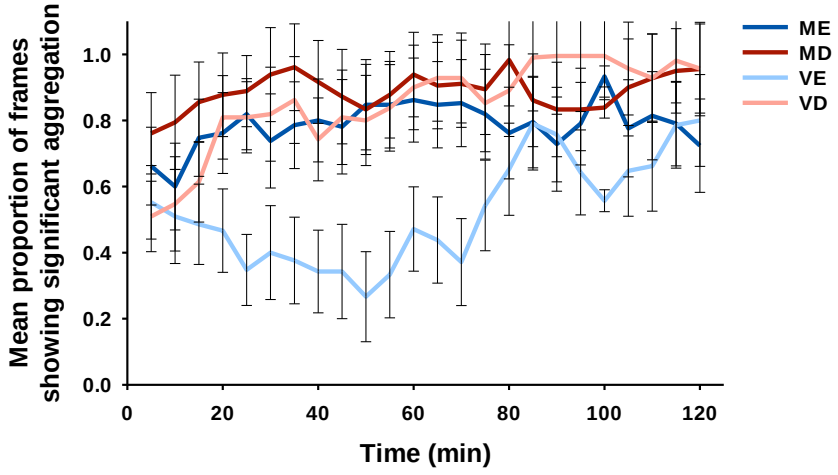

Mean proportion of frames showing significant aggregation
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
Time (min)
ME
MD
VE
VD

**Figure 3.1.6. Effect of social environment on the tendency to aggregate.**

Tendency to aggregate was measured as the proportion of frames in a five minute interval in which significant aggregation was detected. Proportion of frames showing significant aggregation was **a)** greater for mated flies compared to virgin and **b)** smaller for enriched flies compared to deprived flies although only the latter was significant (Main effect of social condition ($p=0.019$)). **c)** Virgin and socially enriched flies showed a visibly lower tendency to aggregate in the first hour of the assay compared to the other three treatments. However, these results could not be tested statistically.

[For box-whisker plots, horizontal line denotes median, box edges denote $25^{th}$ and $75^{th}$ percentiles and whiskers denote $10^{th}$ and $90^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM].

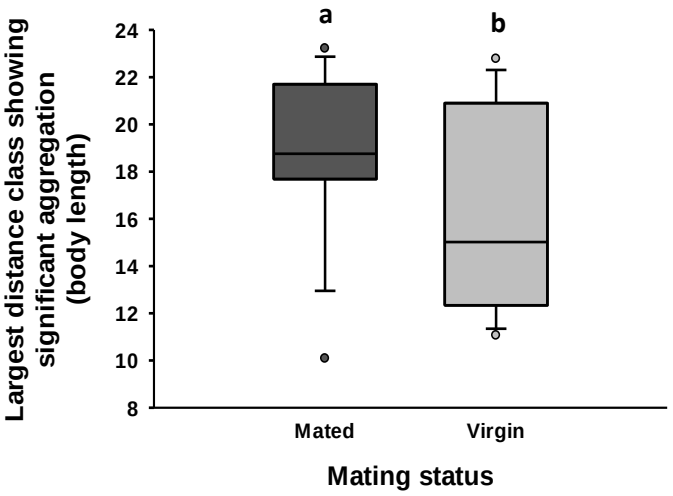
3.1.3.4. Effect of social environment on the scales of aggregation:

As the distributions of the smallest and largest distance class showing aggregation were skewed, the data were transformed ($y=1/x$ and $y=x^2$ respectively) before analysis. The smallest distance class showing significant aggregation was not different across treatments. However, the largest distance class showing aggregation was significantly affected by social environment. Mated flies showed aggregation at a significantly larger distances compared to virgins ($F=5.929$, $df=1$, $p=0.025$) (Fig. 3.1.7a). Socially deprived flies also showed aggregation at larger distances compared to enriched flies, but this difference failed to reach significance ($F=4.064$, $df=1$, $p=0.058$) (Fig. 3.1.7b). A significant interaction effect of mating status and social condition was observed ($F=4.58$, $df=1$, $p=0.045$) as the maximum distance class showing aggregation was significantly smaller for VE flies relative to the other treatments (Fig. 3.1.7c) (Tukey's HSD tests- ME *vs*. VE- $p=0.023$; MD *vs*. VE- $p=0.037$; VE *vs* VD- $p=0.035$). The value of the largest distance class showing aggregation was seen to increase over time for all treatments ($F=3.2367$, $df=23$, $p<0.0001$).

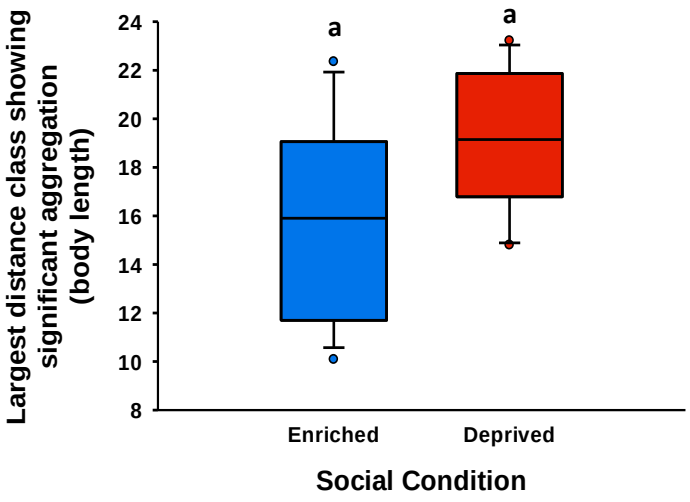
3.1.3.5. Effect of social environment number of flies in each distance class:

Since mean numbers of flies in consecutive distance classes (Fig. 3.1.8a) are cumulative in nature, these values are not independent of each other and hence may not be used for analysis. This lack of independence was corrected by calculating the increase in mean number of flies seen for each distance class (Fig. 3.1.8b). A significant interaction effect of mating status and distance class was seen for the increase in number of flies ($F=4.164$, $df=20$, $p<0.0001$)(Fig. 3.1.8c). Visually, mated flies showed a steeper decline in the number of flies added with each distance class compared to virgin flies. To quantify this, the slope of this function was calculated (the first distance class was excluded as it showed a slope distinct from later distance classes in each replicate video) for different treatments. Comparison of these slopes suggested that decline in the number of flies

a)


Largest distance class showing significant aggregation (body length)
24
22
20
18
16
14
12
10
8
a
b
Mated
Virgin
Mating status

b)


Largest distance class showing significant aggregation (body length)
24
22
20
18
16
14
12
10
8
a
a
Enriched
Deprived
Social Condition

c)

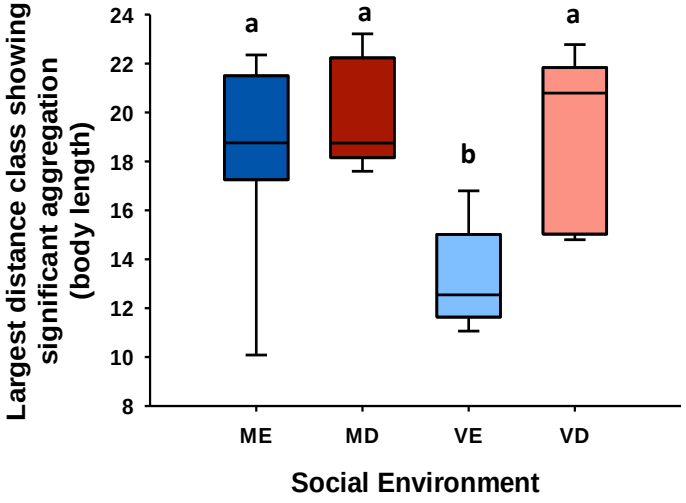
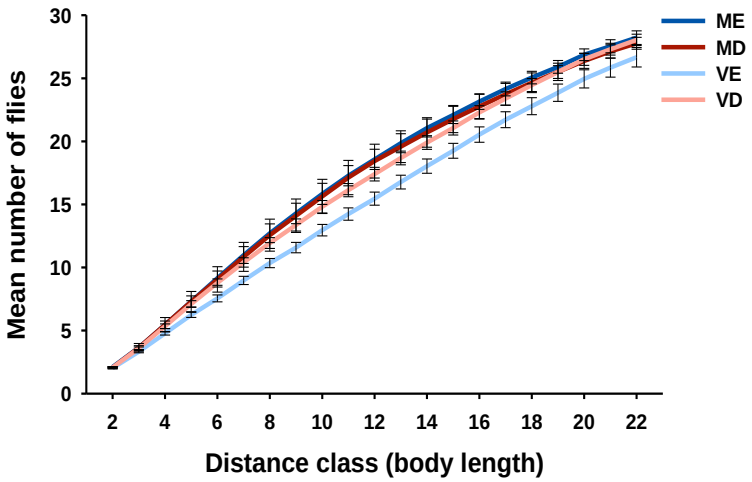

Largest distance class showing significant aggregation (body length)
24
22
20
18
16
14
12
10
8
a
a
b
a
ME
MD
VE
VD
Social Environment

**Figure 3.1.7. Effect of social environment on the scale of aggregation.**

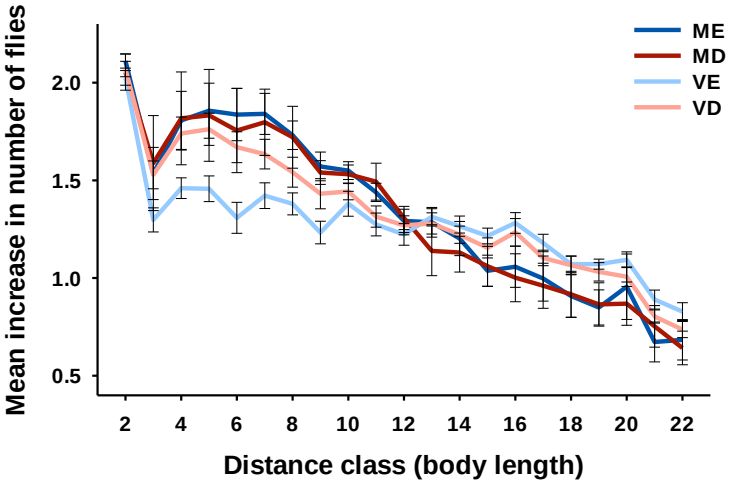
Scale of aggregation was quantified using the largest distance class showing significant aggregation. **a)** Significant aggregation was seen to occur at significantly larger distance classes for mated flies compared to virgins (Main effect of mating status on transformed data ($y=x^2$)). **b)** Significant aggregation occurred at larger distance classes for deprived flies compared to enriched ones but this difference was marginally non-significant. **c)** Values of the largest distance class showing aggregation were significantly smaller for virgin and socially enriched (VE) flies compared to the other treatments (Interaction between mating status and social condition (p=0.045) on transformed data ($y=x^2$). Tukey's HSD tests, ME *vs*. VE- p=0.023; MD *vs*. VE- p=0.037; VE *vs*. VD- p=0.035).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other].

a)


Mean number of flies
Distance class (body length)
ME
MD
VE
VD

b)


Mean increase in number of flies
Distance class (body length)
ME
MD
VE
VD

c)

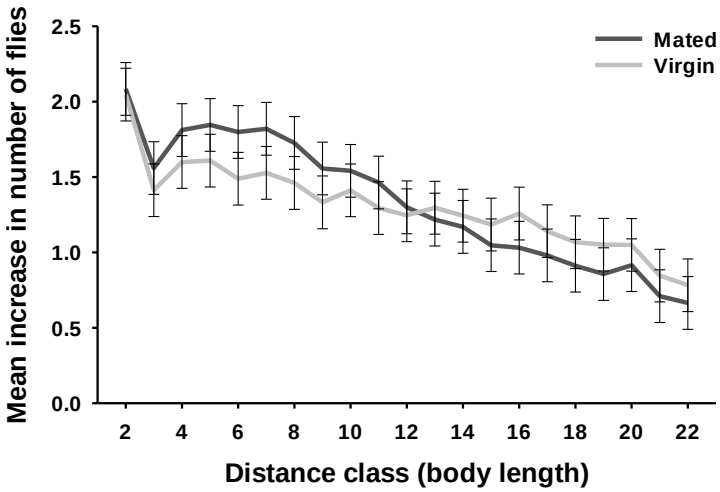

Mean increase in number of flies
Distance class (body length)
Mated
Virgin

**Figure 3.1.8. Distribution of flies across distance classes.**

**a)** Plots of the number of flies in each distance class showed a large amount of overlap across treatments. **b)** Increase in the number of flies with increase in distance was calculated as the difference between neighbouring distance class values in a). **c)** Mated flies showed a significantly larger number of flies in smaller distance classes than virgin flies suggesting that flies mated flies tended to cluster together more than virgins (Interaction between distance class and mating status ($p=0.042$)).

[Error bars are SEM for a) and b) and 95% CIs for c)]

added with each new distance class was indeed significantly greater for mated flies compared to virgins ($F=4.6121$, $df=1$, $p=0.042$).

### *3.1.4. Inferences*

Although the Ripley's K method has been employed mainly to analyze patterns that are stationary in time, it may be applicable to dynamic patterns also. Movement of the points has little effect on the value of K as it only affects the variance in the number of points in a distance class and not the mean, which is used for the computation of the K value. Additionally, this increase in the variance may be reduced by sampling the same set of points across time and averaging the results. However, I am unaware of any previous study that has employed such an approach to theoretically or empirically test the Ripley's K method for moving points. Nevertheless, I employed this approach for this study as it yielded largely the same results as the nearest neighbour methods while also providing additional information. A systematic study, however, is essential to verify the suitability of this approach.

#### 3.1.4.1. Scale and size of fly aggregations:

Ripley's K analysis indicated that significant aggregation occurred at several consecutive distance classes. As each distance class is inclusive of the previous shorter distances, K values are not independent of each other. Thus, presence of aggregation at consecutive distance classes implies that fly aggregates spread across several distance classes as opposed to aggregation occurring independently at different distances. This result, along with the observation that several flies were present in these distance classes, shows that non-random aggregation in flies probably occurs due to the presence of multi-fly gatherings within the arena.

3.1.4.2. Effect of social environment on aggregation patterns:

All three methods showed consistent results for the effect of social environment on aggregation. Contrary to results reported by Simon et al. (2012), socially deprived flies showed more aggregation compared to enriched ones. I speculate that this discrepancy may be due to the longer duration (7 days *vs*. 3 days) of social isolation (or enrichment) provided in their study. I also saw consistent trends of greater aggregation for mated flies compared to virgins across all three analyses but these were not found to be significant (p-values range from 0.07 to 0.28). However, this trend was similar to the results reported by Simon et al. (2012). The lack of significance in this study may be due to the differences in duration of male presence (3-4 days *vs*. 1 day). Further studies with social environments similar to those used by Simon et al. can help clarify the effect of the duration of social experience on aggregation.

Patterns of aggregation across time were similar regardless of the type of analysis used. Virgin flies showed lesser aggregation compared to mated flies but only for the initial seventy or so minutes of the assay. This seems to largely result from behaviour of virgin and socially enriched flies (VE). It is possible that these flies show different behavioural patterns during the initial part of the assay compared to the other treatments and hence show less aggregation.

The Ripley's K method allowed for the measurement of different properties of the spatial pattern such as scale of aggregation and the number of flies in each distance class. I saw that the smallest distance class showing significant aggregation was not different across treatments despite differences in the nearest neighbour distance. This is not really surprising as the resolution of the spatial scale was larger than the range of differences between nearest neighbour distances, making it incapable of reflecting these differences. Although no effects of mating status were detected using nearest neighbour methods, Ripley's K analysis showed that mated flies aggregated at larger spatial scales than virgins and also contained more flies within them. These results suggest that

aggregation is different, both in terms of spatial scale as well as fly numbers, between mated and virgin flies even if these differences are not reflected in nearest neighbour distances. Like mated flies, deprived flies also showed larger scales of aggregation compared to deprived flies. In this case, however, the number of flies in each distance class was not different between enriched and deprived flies. Thus, mating status and social condition seem to affect aggregation patterns in distinct ways.

### *3.2. Features of Drosophila aggregates*

#### *3.2.1. Rationale*

In the previous section, I observed that flies tended to form multi-individual gatherings or aggregates. I also identified differences in the aggregation patterns of flies exposed to different social environments. These differences are likely to be a consequence of differences in the features of aggregates, such as number, size, spread and duration, formed by different flies. Hence, to understand bases of the observed differences in spatial patterns I quantified features of the aggregates that make up the spatial pattern.

Fly groups or aggregates have been previously defined using environmental boundaries such as those of a food patch (Saltz and Foley, 2011) or thermal refuges (Philippe et al., 2016). Since I was interested in studying aggregation brought about purely due to social factors, my experiments did not include any environmental heterogeneities such as food patches which could serve as sites of aggregation. As a result, aggregates could form anywhere within the arena and at any point in time during the course of the assay. To identify such aggregates, I relied on the fact that all aggregates involve non-random proximity between aggregating flies. Since I had estimates of such non-random proximity from the previous section, I was able to group individuals showing such proximity into clusters. These clusters were considered to be aggregates if they were stable across time, i.e. if associations were maintained across time.

After identifying aggregates, I tracked them across time and made note of the identity and number of flies that joined and left these aggregates over the course of the assay. With this information I was able to quantify the abundance, size, duration and density of the observed aggregates. I also quantified changes in these features across time to understand if and how they evolved across time. As earlier, I also tested the effect of social environment on these variables.

*3.2.2. Methods:*

3.2.2.1. Defining aggregates:

Fly clusters were identified using the DBSCAN (Density Based Spatial Clustering of Applications with Noise) algorithm (Yarpiz, 2015). This algorithm uses two parameters that need to be defined by the user- the expected proximity between individuals of a cluster, or clustering distance (ε), and the minimum size of a cluster. The algorithm starts by grouping any points that lie within ε units of each other. Any points that lie within ε units of the already grouped points are also added to this group. This process continues iteratively until every neighbour in proximity (i.e. ε distance) of each grouped point has been included. A group is considered to be a cluster if the number of points in the group equals or exceeds the minimum size. Groups smaller than the minimum size are dismissed as noise. Parameter values are typically chosen based on previous knowledge or using subjective estimates. These values and the methods used to identify them are described below.

- Minimum size of a cluster:

As there was no a priori expectation for the size of the aggregate, every group was considered to be a valid cluster. Hence, the minimum size was set to two flies.

- Clustering distance (ε):

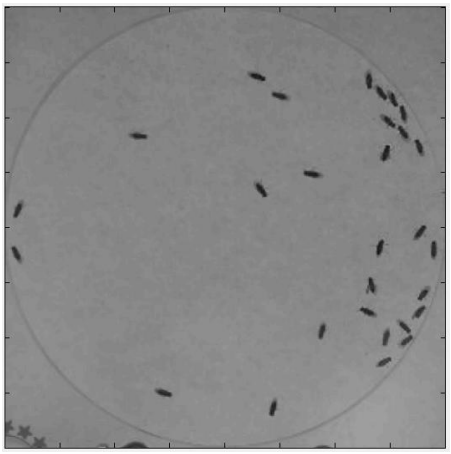
To determine the value of ε, all possible estimates of proximity were identified from previous analyses. These included the median nearest neighbour distance as well as the smallest distance classes at which aggregation was detected using the Ripley's K method. To identify the most appropriate ε from among these values, fly aggregates were assumed to be clearly identifiable by human observers. Conversely, any cluster of flies that could be identified by humans was assumed to correspond to a biologically meaningful aggregate. Given these assumptions, ability of a given ε to detect aggregates could be evaluated by measuring the overlap between aggregates identified

using that value and those identified by a human observer. To measure such overlap, DBSCAN was used to identify clusters in a total of thirty six frames (frames at the $5^{th}$, $55^{th}$ and $115^{th}$ minutes), chosen from three randomly sampled videos from each treatment. For each frame, clusters were identified using all prospective values of ε. These frames were provided to seven human volunteers who were asked to visually identify clusters within them (Fig. 3.2.1a). They were primed to identify a cluster as a group of flies that were close to each other but were not instructed on how this proximity was defined. Each volunteer was free to use her subjective estimates of proximity. After the volunteers had identified subjective clusters for a given frame, they were provided with a visual representation of the clusters identified by DBSCAN using one of the ε values (without disclosing the value) (Fig. 3.2.1b) and were asked to score the mismatch between their subjective classification of clusters in the image and the result provided. Table 3.2.1. was provided as a reference for scoring each image. This process was repeated for each frame and each value of ε. In this fashion, each volunteer provided thirty six scores for each value of ε. These were analysed using a repeated-measures ANOVA with Epsilon as a fixed factor and Volunteer as the repeated-measure. Least mismatch was seen for ε values of 2.5 and 3 body lengths (Fig. 3.2.1c). All results reported here were obtained using 3 body lengths as the value of ε. In the future, the analysis will be repeated using 2.5 body lengths as ε, to verify the consistency of these results.

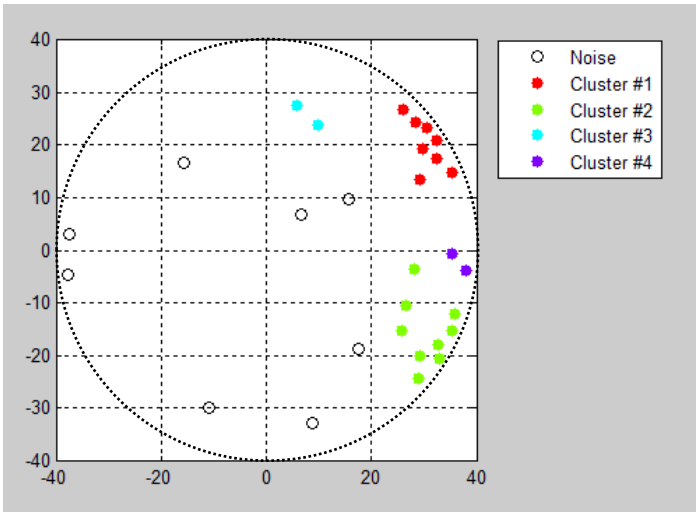
- Minimum duration of an aggregate:

The DBSCAN algorithm identifies every instance of proximity between flies. Thus, it also classifies flies engaging in social interactions as a cluster. As will be discussed in Section 3.4, social interactions involve distinct forms of physical contact between a pair of flies, which typically last for a few frames. Since aggregates may include several individuals and are also stable across time, interactions and aggregation are likely to constitute distinct behaviours. It would thus be inaccurate in include social interactions as a form of aggregation. To exclude such interactions from being classified as aggregates, interactions were distinguished from aggregates on the basis of

a)



b)


Noise
Cluster #1
Cluster #2
Cluster #3
Cluster #4
40
30
20
10
0
-10
-20
-30
-40
-40
-20
0
20
40

c)

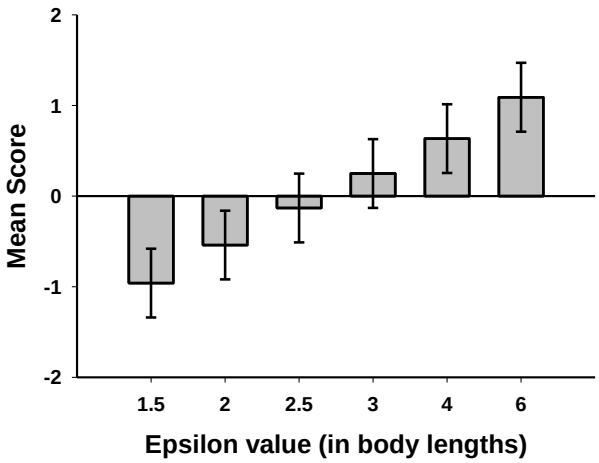

Mean Score
2
1
0
-1
-2
1.5
2
2.5
3
4
6
Epsilon value (in body lengths)

**Figure 3.2.1. Scheme to identify appropriate value of ε**

**a)** Sample image showing fly aggregation. **b)** Clusters were identified within a) using 2.5 body lengths as ε. Volunteers were asked to compare a) and b) and score mismatch between them using Table 3.2.1. as a reference. Clusters were identified using several values of ε for a given image and mismatch was scored for each value of ε. **c)** Mean score was calculated for each ε across multiple sample images. These means were compared using repeated-measures ANOVA with volunteer as a repeated-measure.

[Error bars are 95 % CIs].

| Extent of match | Description | Score |
|---|---|---|
| Bad fit | Many flies missing from aggregates (or misclassified) | -2 |
| Poor fit | Few flies missing from aggregates (or misclassified) | -1 |
| Good fit | Displayed result matches subjective classification | 0 |
| Over fit | Few extra flies included in aggregates | 1 |
| Extremely over fit | Many extra flies included in aggregates | 2 |

**Table 3.2.1. Scoring system used for identifying ideal value of ε.**

Human volunteers were instructed to quantify mismatch between their subjective classification of aggregates and the aggregates identified by DBSCAN using the above guidelines. An ideal score of zero was given for strong agreement between DBSCAN results and the volunteer's subjective classification. Positive and negative scores were indicative of mismatch with respect to the subjective classification due to inclusion or exclusion of flies respectively.

their duration. As most interactions were seen to last less than 120 frames (or 4 seconds), any cluster that persisted for longer than that could be inferred to be an aggregate. Hence, 120 frames was used as the minimum duration for an aggregate and all clusters with shorter durations were ignored.

3.2.2.2. Identifying aggregates:

Custom MATLAB scripts were used for identifying and tracking aggregates across time. These scripts executed the DBSCAN algorithm in each frame using the above parameters and identified each cluster and the flies that were present in it. To track a cluster across time, the composition of a cluster (i.e. identities of flies in the cluster) was compared to compositions of all clusters observed in the previous frame. A cluster was considered to be novel if it was absent in the previous frame (i.e. if none of the flies from that cluster were clustered in the previous frame). Each novel cluster was assigned a unique numerical identity which was logged into a list along with the frame at which it was formed and the identities of the flies that comprised it. A cluster was considered to be a continuation of a previous cluster if the identities of flies in these clusters overlapped maximally (i.e. had most flies in common). This cluster was assigned the identity of the previous cluster and the cluster list was updated to include any new flies that joined the cluster in the new frame. If a large cluster fissioned to form smaller ones, then the largest among them was assigned the identity of the older cluster and newer identities were created for the others. A cluster was considered to end when all flies within it had left it. The frame at which a cluster ended was also recorded in the cluster list. Using the start and end frames for a cluster, the duration of each cluster could be calculated. This duration was used to verify the duration requirement (i.e. duration >120 frames) of an aggregate. All clusters that satisfied this requirement were considered to be aggregates while the others were ignored.

3.2.2.3. Properties of an aggregate:

After all aggregates formed over the course of the assay had been identified, properties of the observed aggregates, such as abundance, size, spread, density and duration, were quantified. Abundance was measured by simply counting the number of unique aggregates observed in the video. Size of the aggregate was measured as the number of flies present in an aggregate. Spread of an aggregate was the space occupied by the aggregate within the arena. It was measured as the radial area of the aggregate i.e. area of the circle with a radius equal to the distance between the centroid of the aggregate and the fly furthest from it. Density of flies within the aggregate was calculated as the ratio of the size of an aggregate to its radial area. Since values of size, spread and density could vary across frames, only the mean values of these variables were used to describe a given aggregate. Finally, duration of an aggregate was measured as the number of frames for which the aggregate was seen to exist.

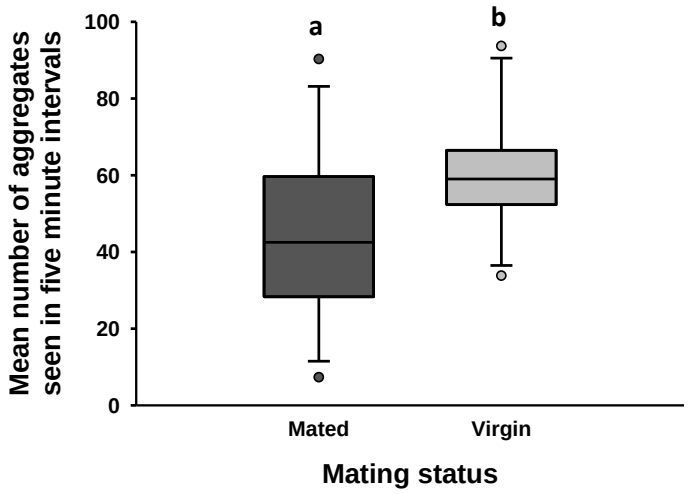
Having defined the properties of individual aggregates, the overall spatial pattern could be described using the mean values of these properties. Mean values of size, spread and density were calculated by averaging these variables across aggregates. Duration was measured using the median instead of the mean as the distribution of aggregate durations was skewed. To test the effect of time on the spatial pattern, these mean (and median) values were quantified for each five minute interval. For a given five minute interval, only those aggregates that were present during the interval were considered for calculating mean or median.

*3.2.3. Results*

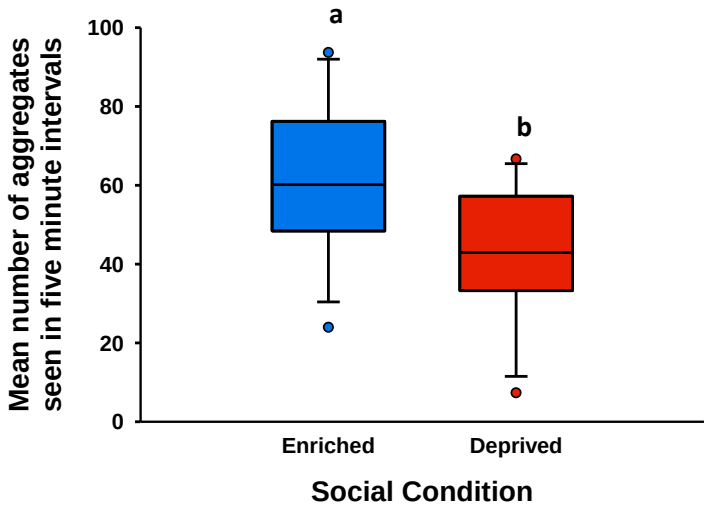
3.2.3.1. Effect of social environment on the total number of aggregates observed:

Mated flies showed significantly fewer aggregates compared to virgins ($F=6.5106$, $df=1$, $p=0.018$)(Fig. 3.2.2a) while socially enriched flies formed a significantly greater number of aggregates compared to deprived flies ($F=8.022$, $df=1$, $p=0.009$) (Fig. 3.2.2b). In general, the

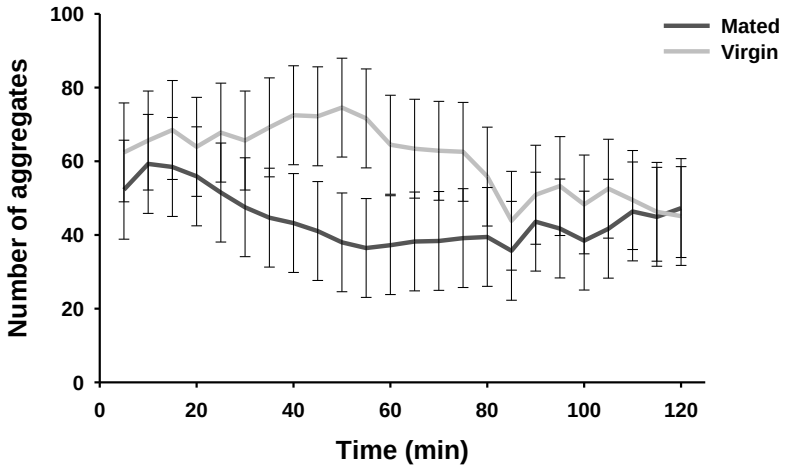
a)


Mean number of aggregates seen in five minute intervals
100
80
60
40
20
0
a
b
Mated
Virgin
Mating status

b)


Mean number of aggregates seen in five minute intervals
100
80
60
40
20
0
a
b
Enriched
Deprived
Social Condition

c)


Number of aggregates
100
80
60
40
20
0
0
20
40
60
80
100
120
Time (min)
Mated
Virgin

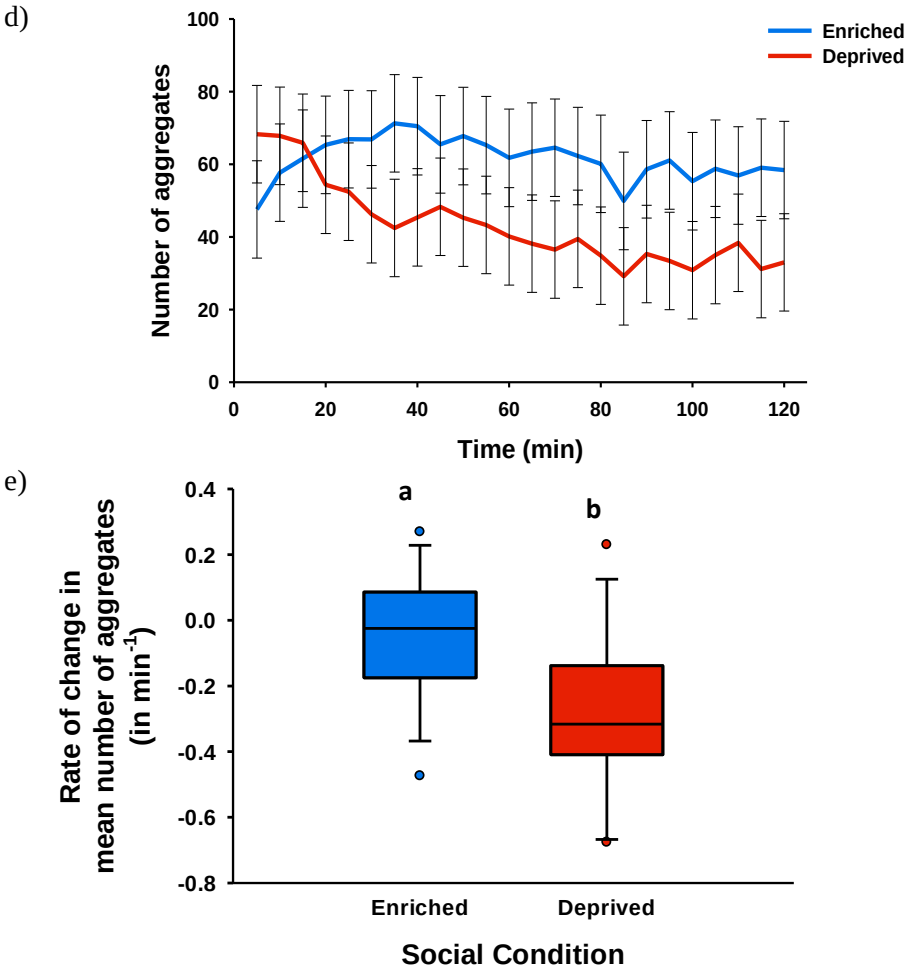



**Figure 3.2.2. Effect of social environment on the number of aggregates seen in a five min interval.**

Mean number of aggregates in a five minute interval was **a)** significantly lower for mated flies compared to virgins and **b)** significantly higher for enriched flies compared to deprived ones (Main effects of mating status ($p= 0.018$) and social condition ($p=0.009$) ). Time profiles of aggregate number differed between **c)** mated and virgin flies and between **d)** enriched flies and deprived flies (Interactions between time and mating status ($p<0.0001$) and between time and social condition ($p<0.0001$)). **e)** Analysis of slopes suggested that deprived flies showed steeper declines in aggregate number with time compared to enriched flies which showed little or no difference (Main effect of social condition ($p=0.012$)). No such differences were seen due to mating status.

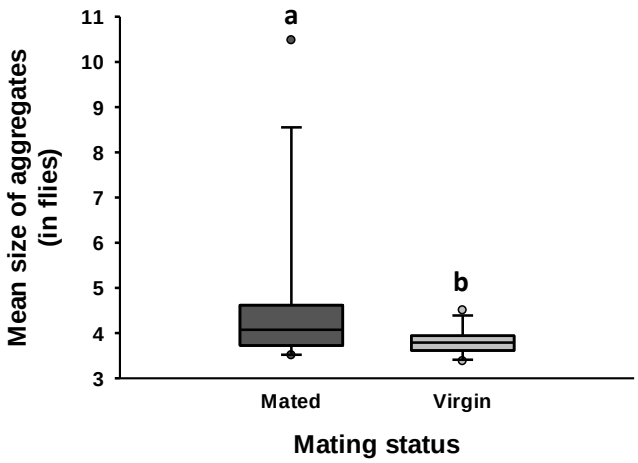
[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote 95% CIs].

number of aggregates was seen to reduce over time ($F=3.7025$, $df=23$, $p<0.0001$). Significant interaction effects were found between mating status and time ($F=2.6461$, $df=23$, $p<0.0001$) (Fig. 3.2.2c) and between social condition and time ($F=3.5171$, $df=23$, $p<0.0001$) (Fig. 3.2.2d). To better understand these interaction effects, slopes of the time series were calculated and analysed using a two-way ANOVA. Mean slope was found to be significantly different between socially enriched and deprived flies ($F=7.3167$, $df=1$, $p=0.012$) (Fig. 3.2.2e) as enriched flies showed a near zero slope while deprived flies showed a clear decline in aggregate number with time. Mated and virgin flies did not show any differences in the values of mean slope. Visual examination of these profiles suggested that mated and virgin flies formed similar numbers of aggregates for most of the video except between the 35th and 90th minutes (approximately) which is when virgin flies were seen to form more aggregates.

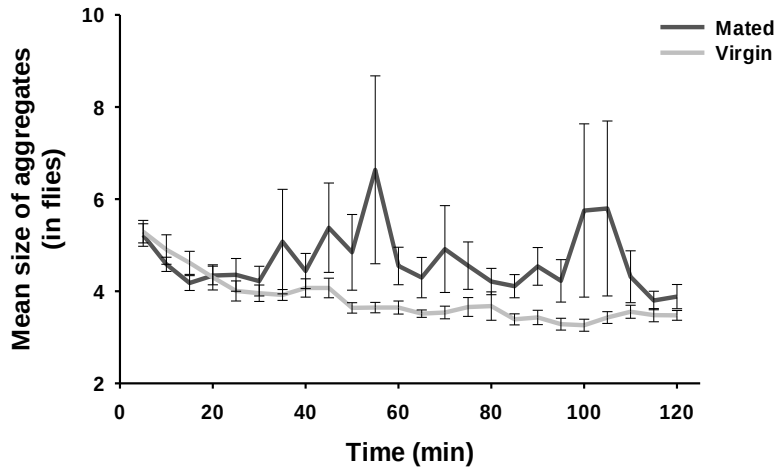
3.2.3.2. Effect of social environment on the mean size of aggregates:

Since distributions of size for different treatments violated the assumptions of normality and homoscedasticity, the data were transformed ($y=1/x^2$) prior to analysis. Mated flies were seen to form significantly larger sized groups than virgins ($F=4.7136$, $df=1$, $p=0.04$) (Fig. 3.2.3a). This result was biased due to the presence of a single replicate (mated and socially deprived) that showed an extremely large mean aggregate size. Exclusion of this replicate yielded a similar but marginally non-significant result ($F=3.1256$, $df=1$, $p=0.09$). Regardless of the treatment, size of the aggregate was seen to reduce over time ($F=7.7575$, $df=23$, $p<0.0001$). A significant interaction effect of social condition and time was also observed ($F=2.414$, $df=23$, $p<0.0001$) (Fig. 3.2.3c). To study the effects across time, slopes of the untransformed time series were compared across treatments. Significant differences in slope were detected between socially enriched and deprived flies ($F=10.2119$, $df=1$, $p=0.004$) (Fig. 3.2.3e) and between mated and virgin flies ($F=5.4531$, $df=1$, $p=0.028$) (Fig. 3.2.3b,d). Slopes were more negative for socially enriched flies and virgin flies compared to deprived flies and mated flies respectively.

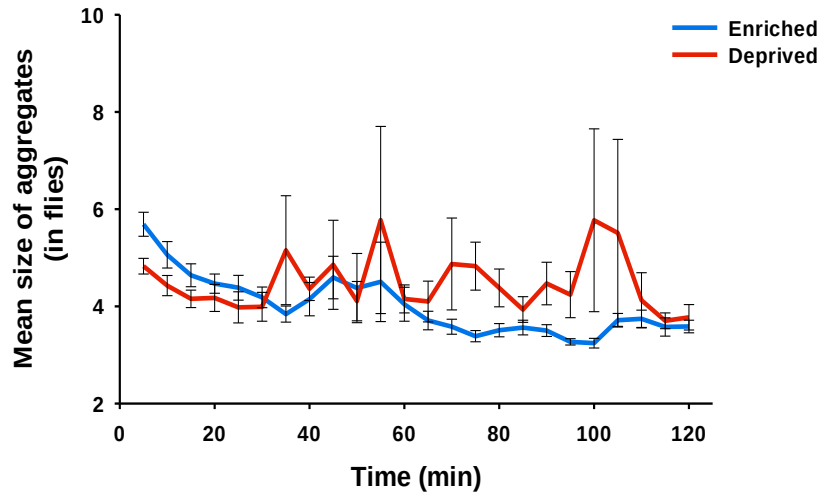
a)


Mean size of aggregates (in flies)
11
10
9
8
7
6
5
4
3
a
b
Mated
Virgin
Mating status

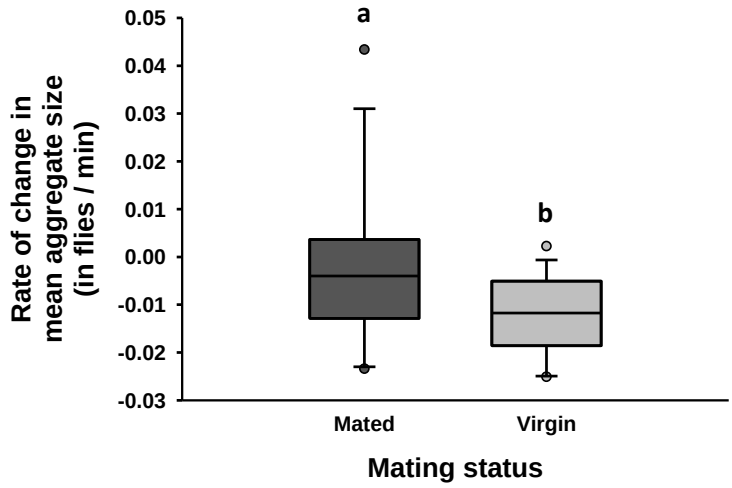
b)


Mean size of aggregates (in flies)
10
8
6
4
2
Mated
Virgin
0
20
40
60
80
100
120
Time (min)

c)


Mean size of aggregates (in flies)
10
8
6
4
2
Enriched
Deprived
0
20
40
60
80
100
120
Time (min)

d)




e)

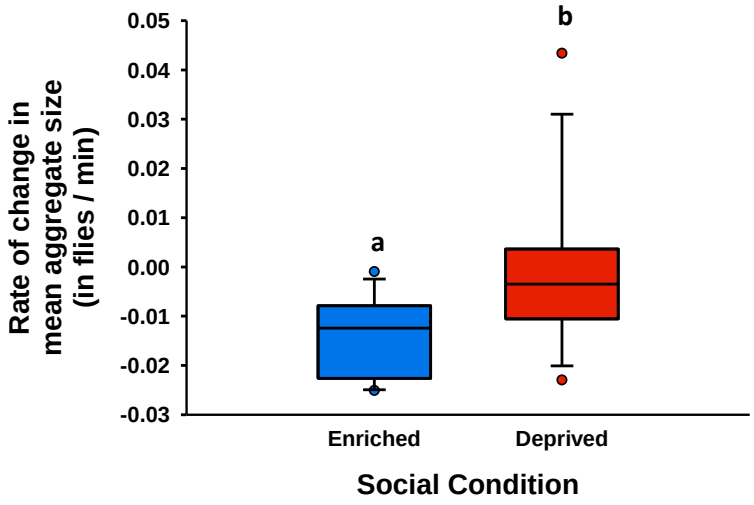



**Figure 3.2.3. Effect of social environment on the mean size of aggregates seen in a five minute interval.**

**a)** Mean number of flies per aggregate was significantly higher for mated flies compared to virgin flies. (Main effect of mating status ($p=0.04$) on transformed data ($y=1/x^2$)). Time profiles of number of flies per aggregate were **b)** not different for mated and virgin flies but **c)** significantly different between enriched and deprived flies (Interaction between time and social condition ($p<0.0001$) on transformed data ($y=1/x^2$)). Analysis of slopes showed that decline in aggregate size was **d)** significantly lower (i.e. less negative) for mated flies compared to virgins and **e)** significantly higher (i.e. more negative) for enriched flies compared to deprived flies (Main effects of mating status ($p=0.028$) and social condition ($p=0.004$) on slopes calculated using untransformed data).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM].

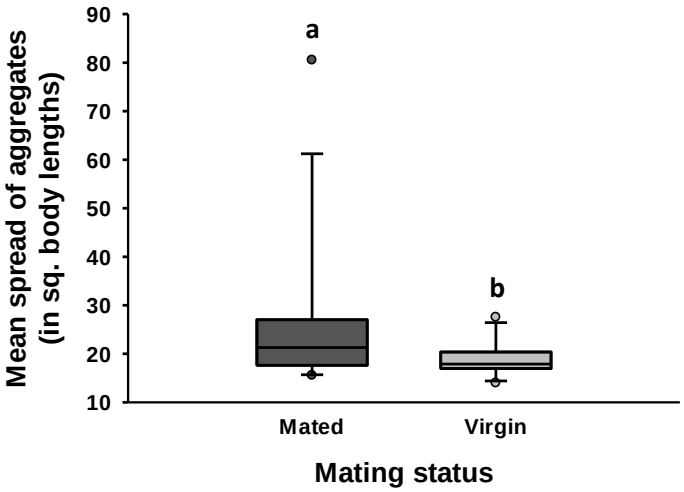
3.2.3.3. Effect of social environment on the mean radial area of aggregates:

Since these data violated assumptions of the ANOVA, they were transformed (y=1/x) before analysis. Aggregates formed by mated flies showed significantly larger spatial spread compared to virgin flies (F=4.9379, df=1, p=0.036) (Fig. 3.2.4a). As above, this result was biased by the presence of the same extreme replicate. Exclusion of this replicate yielded a non-significant effect of mating status although the pattern of differences was maintained (F=3.307, df=1, p=0.083). The area of the aggregate was seen to reduce significantly with time (F=8.47, df=23, p<0.0001). Significant interaction effects between social condition and time and between mating status and time were also detected but these were found to be non-significant after correcting for lack of sphericity (Greenhouse-Geisser correction for Social condition x Time - F=1.8327, $df_1$=7.9, $df_2$=181.7, p=0.074; Greenhouse-Geisser correction for Mating status x Time - F=1.6033, $df_1$=7.9, $df_2$=181.7, p=0.127). When slopes of the untransformed time series were analysed, significant differences could be detected between mated and virgin flies (F=7.1209, df=1, p=0.013) (Fig. 3.2.4b,d) and between socially enriched and deprived flies (F=9.4514, df=1, p=0.005) (Fig. 3.2.4c,e). Both socially enriched flies and virgin flies showed significantly steeper declines in aggregate area with time compared to socially deprived flies and mated flies respectively.

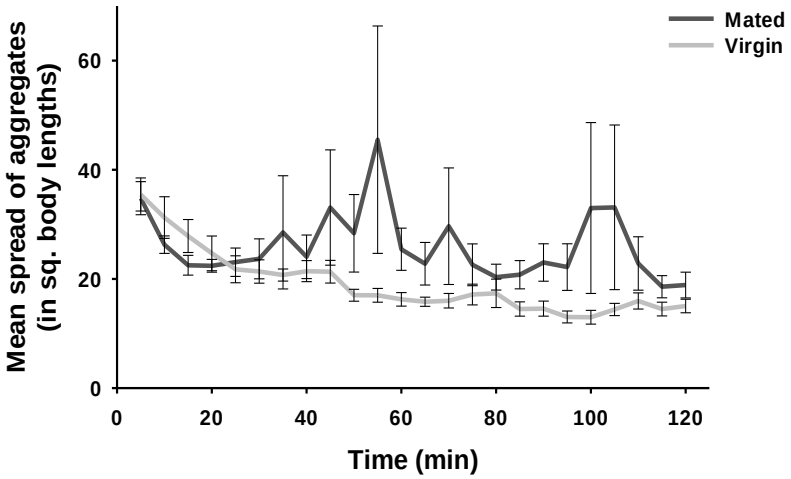
3.2.3.4. Effect of social environment on the mean density of aggregates:

Aggregates formed by mated flies were seen to be less dense compared to those formed by virgin flies but this effect was not found to be statistically significant (F=3.752, df=1, p=0.065) (Fig. 3.2.5a). However, a significant effect of the interaction between mating status and social condition (F=5.63, df=1, p=0.026) could be observed. Post hoc tests indicated that this was due to differences in density between MD flies and VD flies (Tukey's HSD test, p=0.031) (Fig. 3.2.5b). Overall, mean density was seen to increase over time (F=3.398, df=23, p<0.0001). All interactions between time and the fixed factors were seen to have significant effects on aggregate density but these effects

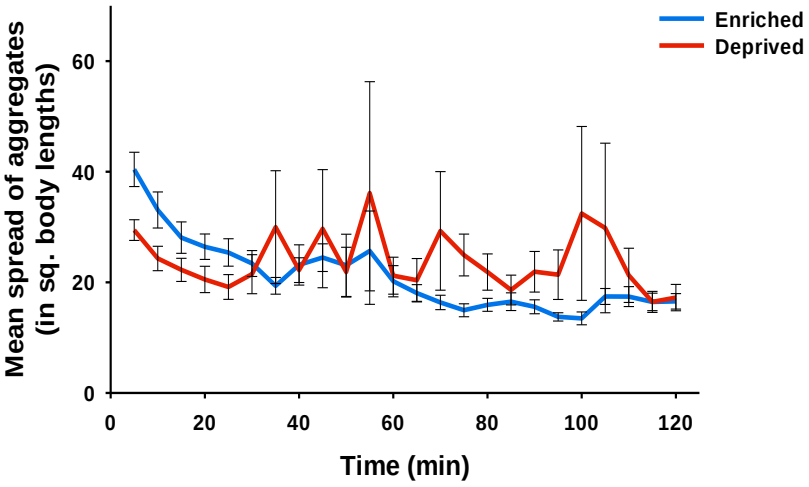
a)


Mean spread of aggregates (in sq. body lengths)
90
80
70
60
50
40
30
20
10
a
b
Mated
Virgin
Mating status

b)


Mean spread of aggregates (in sq. body lengths)
60
40
20
0
0
20
40
60
80
100
120
Time (min)
Mated
Virgin

c)


Mean spread of aggregates (in sq. body lengths)
60
40
20
0
0
20
40
60
80
100
120
Time (min)
Enriched
Deprived

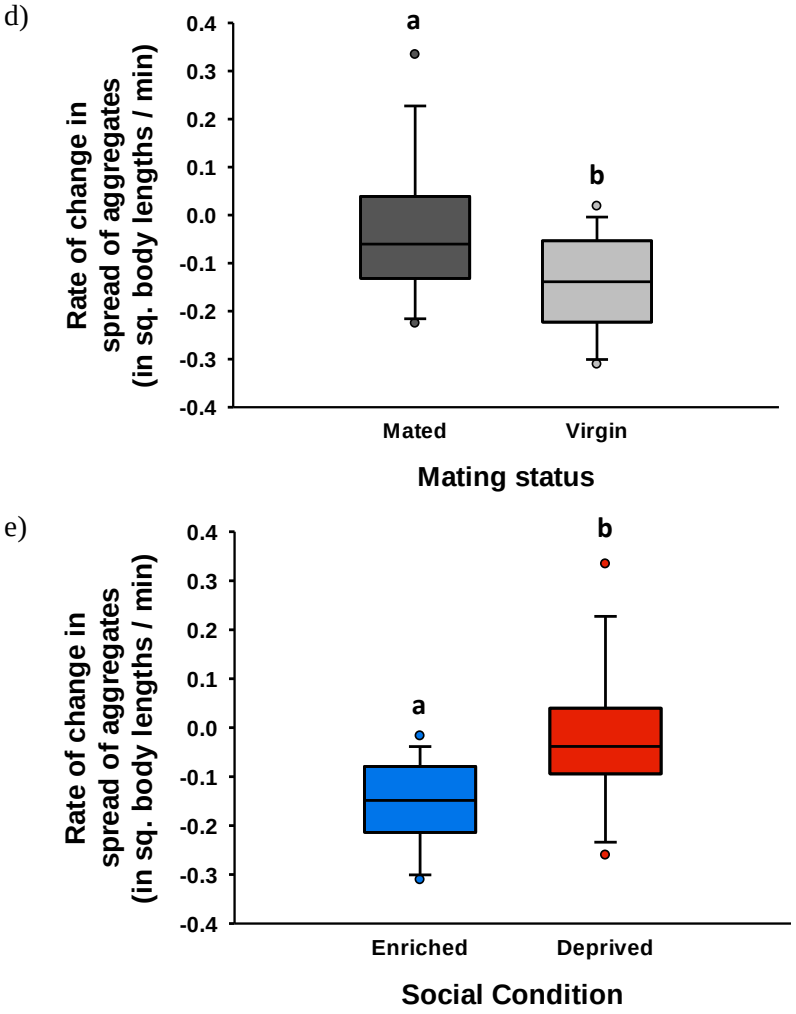
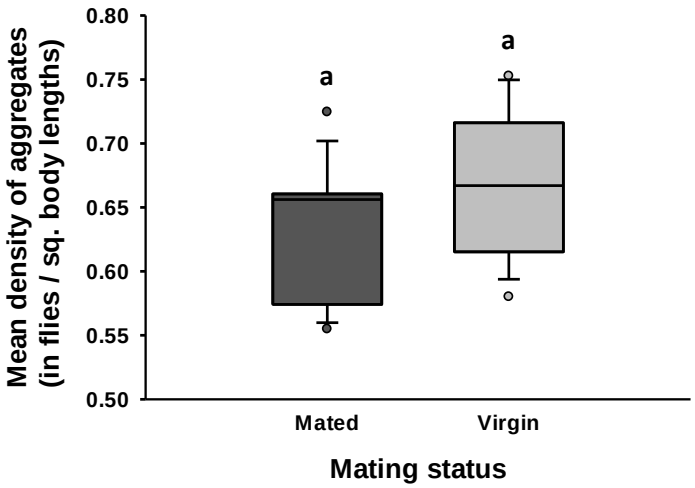



**Figure 3.2.4. The effect of social environment on the radial area of aggregates seen in a five minute interval.**

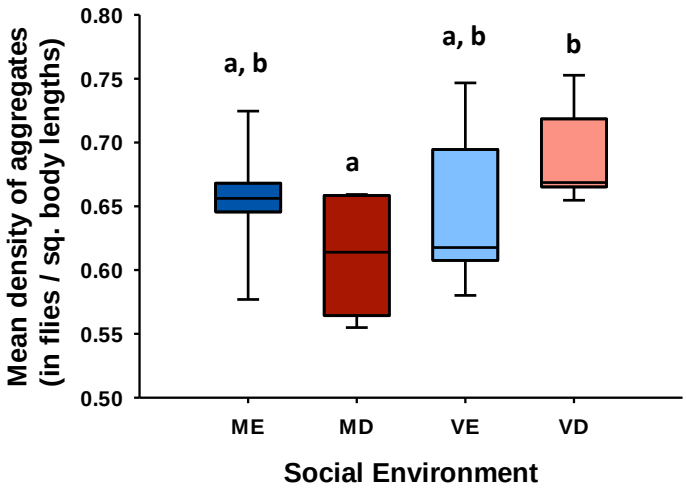
**a)** Mean radial area (area of circle traced using distance from the centroid of the aggregate to the furthest fly as radius) of an aggregate was significantly greater for mated flies compared to virgins (Main effect of mating status ($p=0.036$) on transformed data ($y=1/x$)). Time profiles of aggregate area were not significantly different **b)** between mated flies and virgins or **c)** between enriched flies and deprived flies. However, decline in aggregate spread over time was **d)** significantly smaller (i.e. less negative) for mated flies compared to virgins and **e)** greater (i.e. more negative) for enriched flies compared to deprived flies (Main effect of mating status ($p=0.013$) and social condition ($p=0.005$) on slopes calculated using untransformed data).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM].

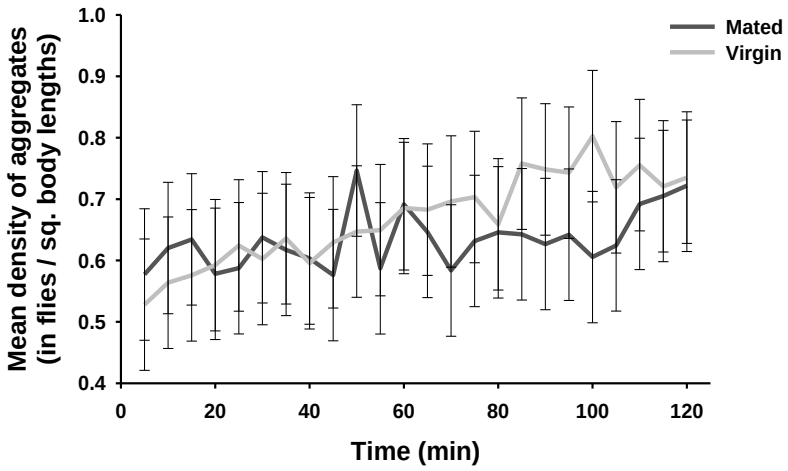
a)


Mean density of aggregates (in flies / sq. body lengths)
0.80
0.75
0.70
0.65
0.60
0.55
0.50
a
a
Mated
Virgin
Mating status

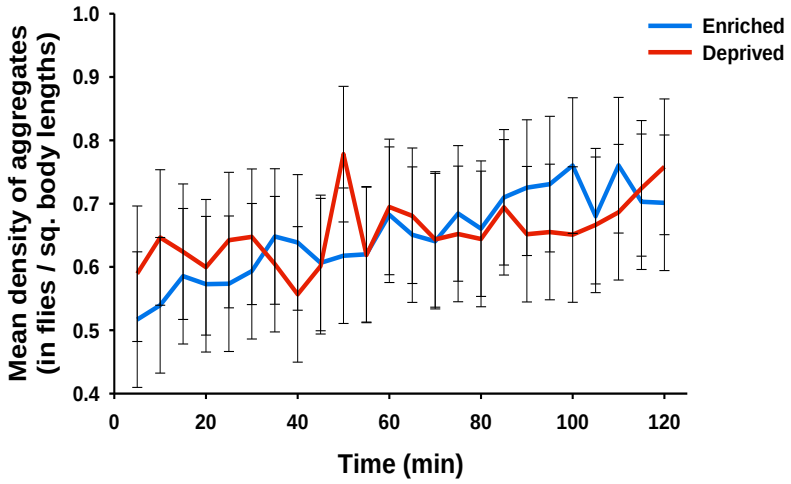
b)


Mean density of aggregates (in flies / sq. body lengths)
0.80
0.75
0.70
0.65
0.60
0.55
0.50
a, b
a
a, b
b
ME
MD
VE
VD
Social Environment

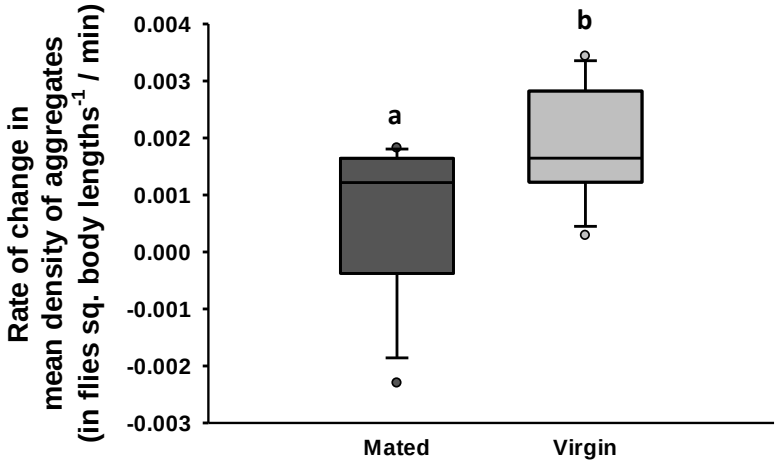
c)


Mean density of aggregates (in flies / sq. body lengths)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
20
40
60
80
100
120
Time (min)
Mated
Virgin

d)


Enriched
Deprived
Mean density of aggregates (in flies / sq. body lengths)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
20
40
60
80
100
120
Time (min)

e)


Rate of change in mean density of aggregates (in flies sq. body lengths$^{-1}$ / min)
0.004
0.003
0.002
0.001
0.000
-0.001
-0.002
-0.003
a
b
Mated
Virgin
Mating status

f)

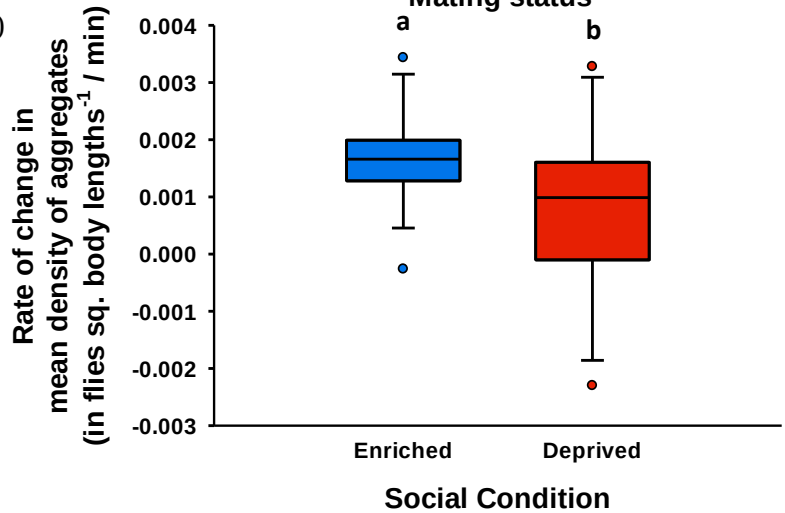

Rate of change in mean density of aggregates (in flies sq. body lengths$^{-1}$ / min)
0.004
0.003
0.002
0.001
0.000
-0.001
-0.002
-0.003
a
b
Enriched
Deprived
Social Condition

**Figure 3.2.5. Effect of social environment on the mean density of aggregates seen in a five minute interval.**

**a)** Aggregate density was lower for mated flies compared to virgins but this difference was marginally non-significant (Main effect of mating status ($p=0.065$)). **b)** Mated and socially deprived (MD) flies showed significantly lower density compared to virgin and socially deprived flies (VD) (Interaction between mating status and social condition ($p=0.026$); Tukey's HSD test -MD *vs.* VD, $p=0.031$). Time profiles of aggregate density did not differ between **c)** mated and virgin flies or **d)** between enriched or deprived flies. However, rates of increase in density across time were **e)** significantly lower for mated flies compared to virgins and **f)** significantly higher for enriched flies compared to deprived flies (Main effect of mating status ($p=0.007$) and social condition ($p=0.016$) on slope).

[For box-whisker plots, horizontal line denotes median, box edges denote $25^{th}$ and $75^{th}$ percentiles and whiskers denote $10^{th}$ and $90^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote 95 % CIs].

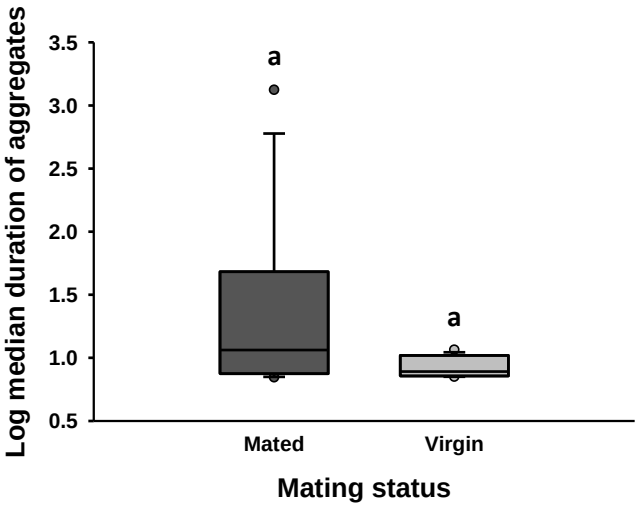
were not significant after correcting for lack of sphericity. However, when slopes of density across time were compared across treatments, significant effects of both mating status (F=8.6611, df=1, p=0.007) (Fig. 3.2.5c,e) and social condition (F=6.6706, df=1, p=0.016) were detected (Fig. 3.2.5d,f). Both socially enriched flies and virgin flies showed significantly greater positive slope compared to socially deprived and mated flies respectively.

3.2.3.5. Effect of social environment on the duration of aggregates:

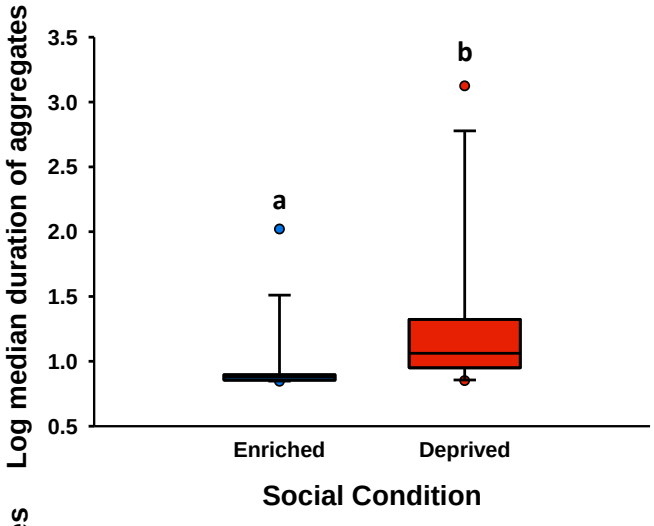
As the data violated the assumptions of normality and homoscedasticity, they were transformed ($y=1/x^{3.5}$) before analysis. Socially enriched flies were seen to form significantly short lived aggregates compared to deprived flies (F=25.9266, df=1, p<0.0001) (Fig. 3.2.6b). Marginally non-significant effects of mating status (F=4.0458, df=1, p=0.056) (Fig. 3.2.6a) and the interaction between mating status and social condition (F=2.9822, df=1, p=0.097) (Fig. 3.2.6c) were also observed. In general, aggregates were seen to become longer lived over time (F=5.4128, df=23, p<0.0001). A significant three-way interaction between mating status, social condition and time was also observed (F=2.345, df=23, p=0.0004) (Fig. 3.2.6d). To analyze these results, slope of each time series (untransformed data) was measured and analysed using a two-way ANOVA. As these data violated normality and homoscedasticity, they were transformed ($y=x^{0.2}$) before analysis. Socially enriched flies were seen to show significantly lower positive slope than deprived flies (F=11.689, df=1, p=0.002) (Fig. 3.2.6e). Significant effects of the interaction between mating status and social condition were also observed (F=10.09, df=1, p=0.004) (Fig. 3.2.6f). This effect was due to the significantly larger slope seen for MD flies (Tukey's HSD test, ME *vs.* MD- p=0.0008, VE *vs.* MD- p=0.045, VD *vs*. MD- p=0.065).

To better understand the above result for MD flies across time, the effect of social environment was tested on parameters of the distribution of aggregate durations, such as inter-quartile range (IQR), skewness and kurtosis. Only the results for IQR are discussed here as the three-way interaction had

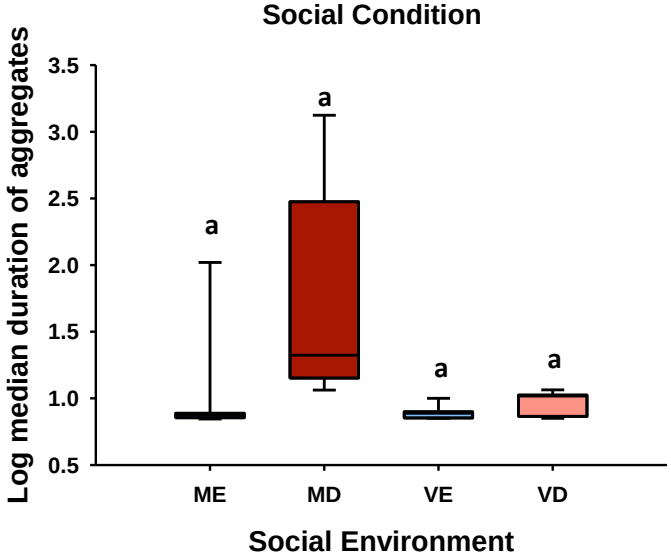
a)


Log median duration of aggregates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
a
a
Mated
Virgin
Mating status

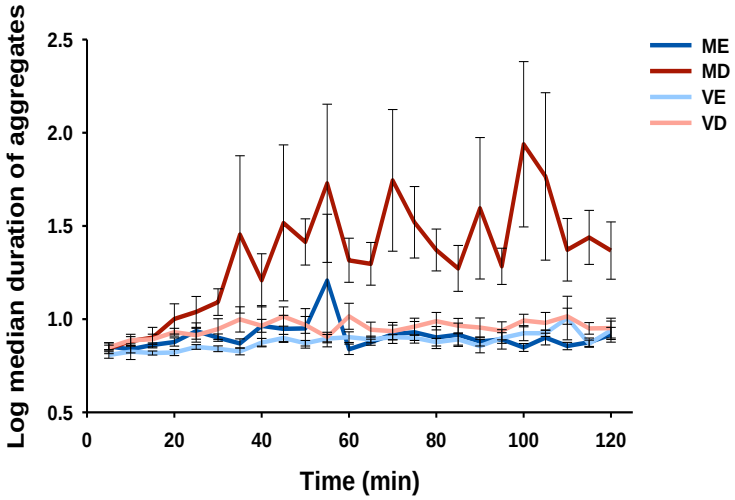
b)


Log median duration of aggregates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
a
b
Enriched
Deprived
Social Condition

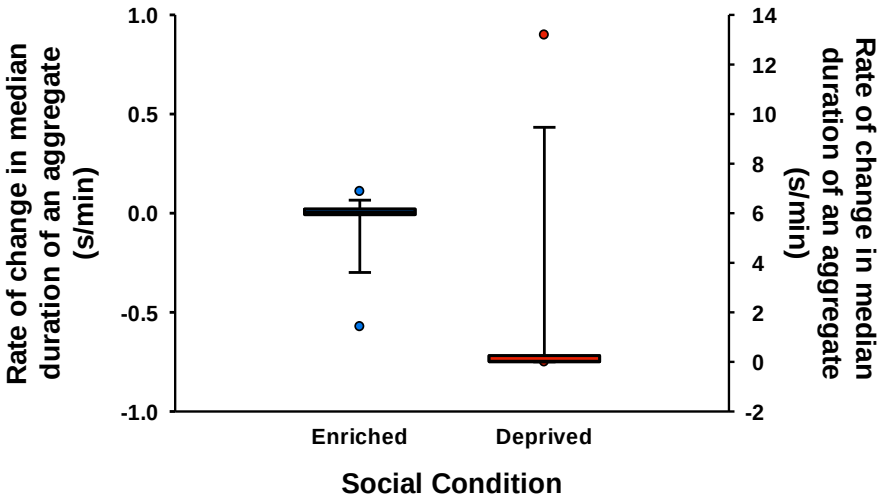
c)


Log median duration of aggregates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
a
a
a
a
ME
MD
VE
VD
Social Environment

d)


Log median duration of aggregates
2.5
2.0
1.5
1.0
0.5
0
20
40
60
80
100
120
Time (min)
ME
MD
VE
VD

e)


Rate of change in median duration of an aggregate (s/min)
1.0
0.5
0.0
-0.5
-1.0
14
12
10
8
6
4
2
0
-2
Rate of change in median duration of an aggregate (s/min)
Enriched
Deprived
Social Condition

f)

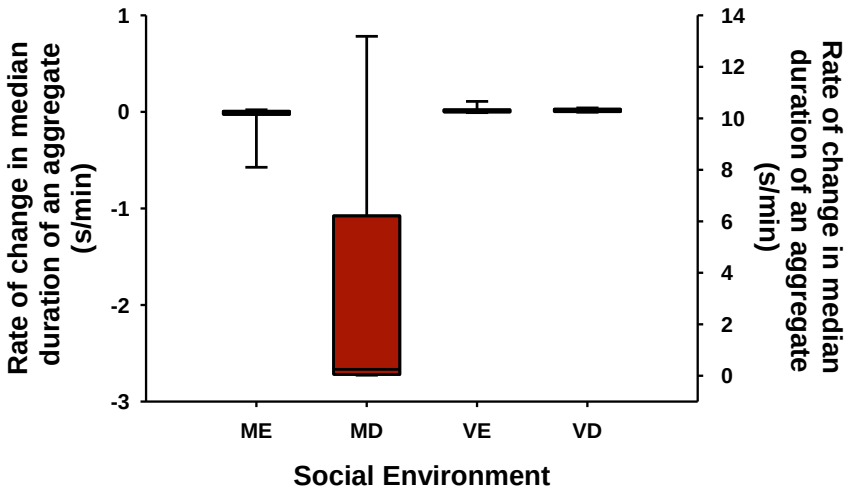

Rate of change in median duration of an aggregate (s/min)
1
0
-1
-2
-3
14
12
10
8
6
4
2
0
Rate of change in median duration of an aggregate (s/min)
ME
MD
VE
VD
Social Environment

**Figure 3.2.6. Effect of social environment on the median duration of aggregates seen in a five minute interval.**

**a)** Median duration of aggregates was higher for mated flies compared to virgins, although this difference was marginally non-significant (Main effect of mating status ($p=0.056$) on transformed data ($y=1/x^{3.5}$)). **b)** Median duration was significantly smaller for enriched flies compared to deprived flies (Main effect of social condition ($p<0.0001$) on transformed data ($y=1/x^{3.5}$)). **c)** Mated and deprived (MD) flies show a trend for higher duration values than the other treatment but this was non-significant (Interaction effect between mating status and social condition ($p=0.097$) on transformed data ($y=1/x^{3.5}$)). **d)** Time profiles of aggregate duration were not similar across treatments (Interaction between mating status, social condition and time ($p=0.0004$) on transformed data ($y=1/x^{3.5}$)). **e)** Rate of change in aggregate duration was significantly lower for enriched flies compared to deprived flies (Main effect of social condition on transformed values of slope ($y=x^{0.2}$) calculated using untransformed data ($y=1/x^{3.5}$)). **f)** Mated and socially deprived (MD) flies showed significantly higher increases in aggregate duration with time relative to the other treatments (Interaction between mating status and social condition on transformed values of slope ($y=x^{0.2}$) calculated using untransformed data ($y=1/x^{3.5}$). Tukey's HSD test, ME *vs.* MD- $p=0.0008$, VE *vs.* MD- $p=0.045$, VD *vs.* MD- $p=0.065$).

[Duration values were log transformed before plotting for easier visualisation. Given the large variation in slope values, separate axes were used to plot different treatments. In e), left axis used for enriched flies. In f), left axis used for all treatments except MD. ]

[For box-whisker plots, horizontal line denotes median, box edges denote $25^{th}$ and $75^{th}$ percentiles and whiskers denote $10^{th}$ and $90^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM].

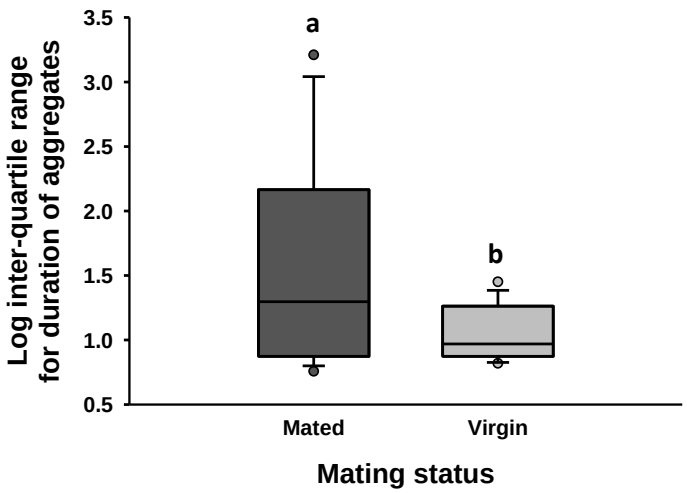
no significant effects on the other parameters. These values were transformed before analysis ($y=1/\sqrt{x}$) to account for lack of normality and homoscedasticity. Mated flies and deprived flies showed significantly larger IQR values compared to virgins and enriched flies respectively (Mated *vs*. Virgin, F=5.251, df=1, p=0.032; Enriched *vs*. Deprived, F=28.122, df=1, p<0.0001) (Fig. 3.2.7a,b). A significant effect of the interaction between mating status and social condition was also observed (F=7.894, df=1, p=0.01) (Fig. 3.2.7c). Post hoc tests revealed that this was due to larger values seen for MD flies compared to the other treatments. (Tukey's HSD test, ME *vs*. MD- p=0.0003, VE *vs*. MD- p=0.0004, VD *vs*. MD- p=0.011). Significant effects of the interaction between time and social condition (F=1.97, df=23, p=0.0048) (Fig. 3.2.7d) and among time, social condition and mating status (F=2.5414, df=23, p=0.0001) (Fig. 3.2.7e) were also detected and were seen to persist after correcting for lack of sphericity. To analyse the profile of IQR across time, a two-way ANOVA was performed on slopes of the time profiles. The data were transformed before analysis ($y=x^{1/3}$) to account for lack of homoscedasticity and normality. There were no significant effects of social environment on these data (Fig. 3.2.7f, g) but a trend of larger slope values for MD flies was observable. Analysis using a larger dataset may clarify this result further.

### *3.2.4. Inferences*

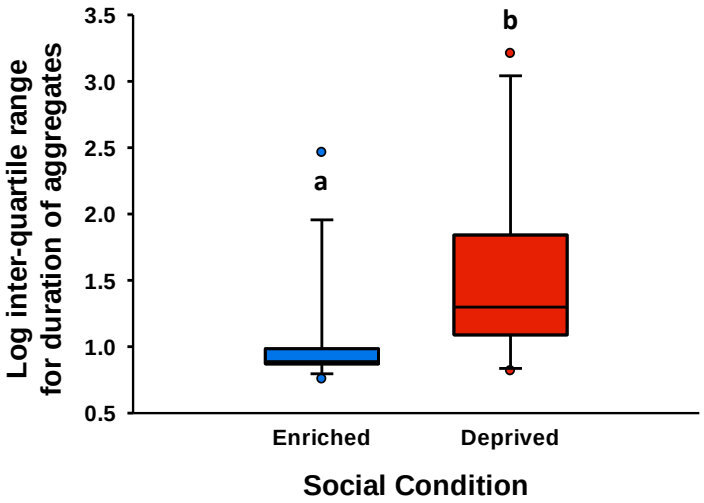
#### 3.2.4.1. Defining and understanding an aggregate:

In this section, I used a clustering based approach to identify aggregates that make up fly aggregation patterns. Since this method is susceptible to subjectivity in the choice of ε, I chose empirical estimates of proximity to avoid subjectivity. However, I had to rely on subjective opinions of human volunteers to choose from among several possible estimates of proximity. Thus, it is unclear if the chosen value of ε accurately represents the distance at which flies may perceive proximity with each other inside aggregates. To further refine the criteria for defining aggregates,

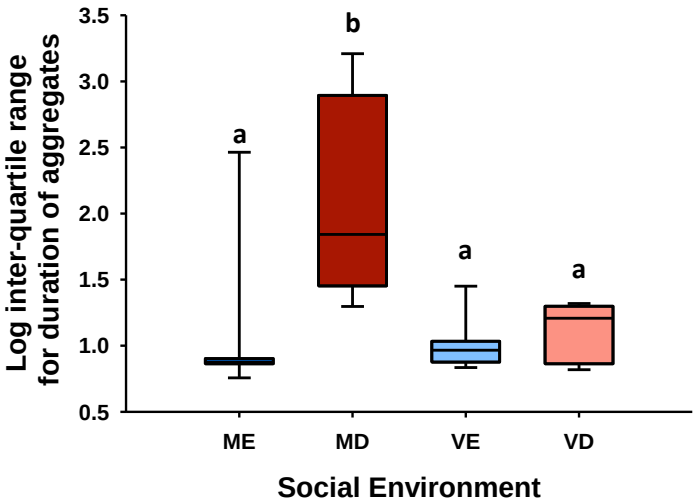
a)


Log inter-quartile range for duration of aggregates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
a
b
Mated
Virgin
Mating status

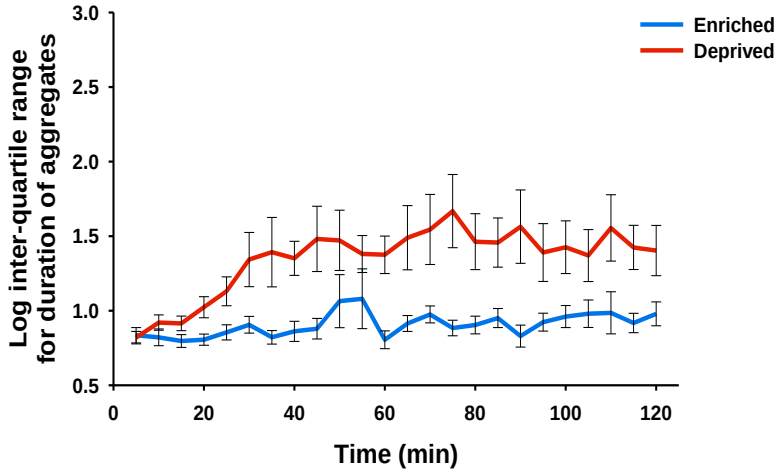
b)


Log inter-quartile range for duration of aggregates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
a
b
Enriched
Deprived
Social Condition

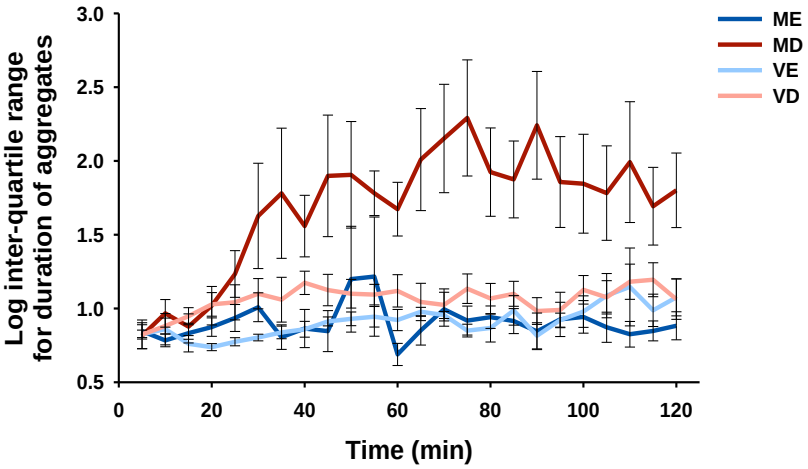
c)


Log inter-quartile range for duration of aggregates
3.5
3.0
2.5
2.0
1.5
1.0
0.5
a
b
a
a
ME
MD
VE
VD
Social Environment

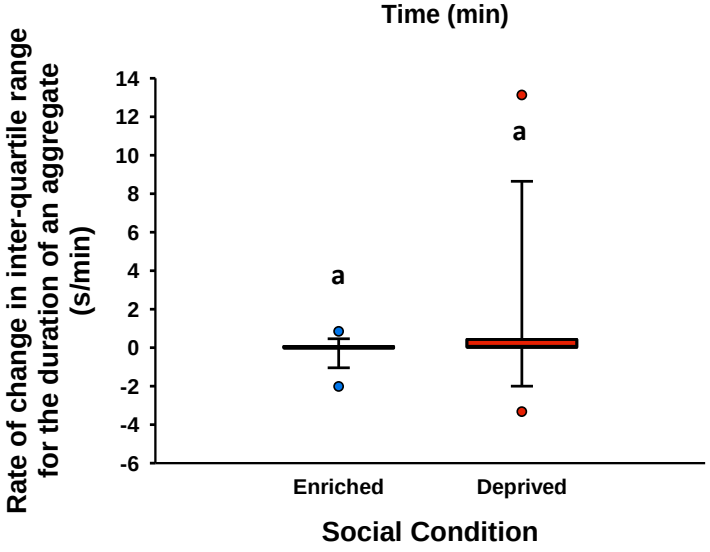
d)


Enriched
Deprived
Log inter-quartile range for duration of aggregates
3.0
2.5
2.0
1.5
1.0
0.5
0
20
40
60
80
100
120
Time (min)

e)


ME
MD
VE
VD
Log inter-quartile range for duration of aggregates
3.0
2.5
2.0
1.5
1.0
0.5
0
20
40
60
80
100
120
Time (min)

f)


Rate of change in inter-quartile range for the duration of an aggregate (s/min)
14
12
10
8
6
4
2
0
-2
-4
-6
a
a
Enriched
Deprived
Social Condition

g)

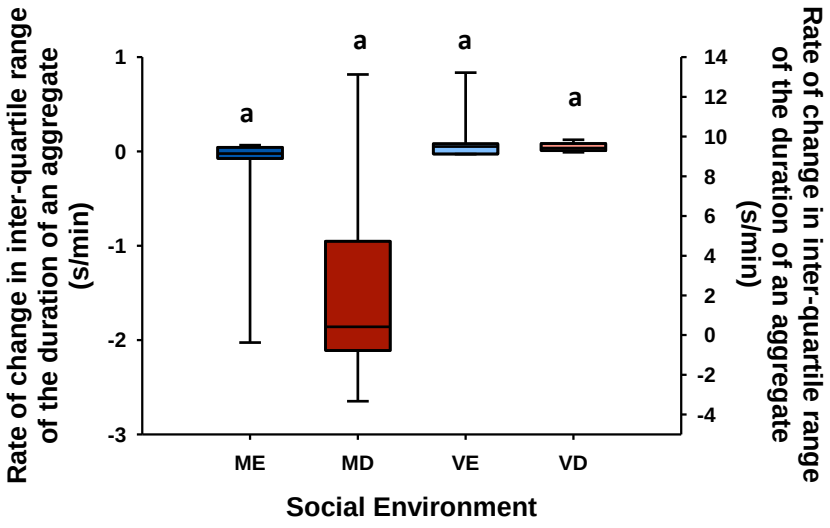



**Figure 3.2.7. Effect of social environment on variation in duration of aggregates seen in five minute intervals.**

Inter-quartile range (IQR) was **a)** significantly larger for mated flies compared to virgins and **b)** significantly smaller for enriched flies compared to deprived flies (Main effect of mating status ($p=0.032$) and social condition ($p<0.0001$) on transformed data ($y=1/\sqrt{x}$)). **c)** Mated and socially deprived (MD) flies showed significantly greater IQR than the other treatments (Interaction between mating status and social condition ($p=0.01$) on transformed data ($y=1/\sqrt{x}$). Tukey's HSD test, ME *vs.* MD- $p=0.0003$, VE *vs.* MD- $p=0.0004$, VD *vs.* MD- $p=0.011$). **d)** Time profile for IQR of aggregate duration was different for enriched and deprived flies (Interaction between time and social condition ($p=0.0048$) on transformed data ($y=1/\sqrt{x}$)). **e)** Time profiles for IQR of aggregate duration were dissimilar across the four treatments (Interaction between time, mating status and social condition ($p=0.0001$) on transformed data ($y=1/\sqrt{x}$)). **f)** and **g)** Rates of change in IQR were not different across treatments although mated and socially deprived (MD) flies tended to show larger values compared to the other treatments.

[Duration values were log transformed before plotting for easier visualisation. Given the large variation in slope values, separate axes were used to plot different treatments. In g), left axis used for all treatments except MD.]

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM].

behavioural markers of aggregation may be used. Behavioural analysis of flies within the aggregates described here may be useful for developing these criteria.

Ripley's K analysis from the previous section suggested that aggregation occurred at several large distance classes. I interpreted this result to reflect the spread of a single aggregate. However, the values of spread obtained for individual aggregates in this section were smaller (mean ≈ 4 body lengths) than those seen in the previous section (mean ≈ 17 body lengths). This discrepancy may arise from clustering of aggregates themselves, which was often observed during the assay. Thus, it is possible that flies show another level of non-random spatial organization whereby aggregates are themselves aggregated. If this is true, then social environment could affect the organization of aggregates in addition to its effects on the properties of an individual aggregate. Further study is required to verify the existence of such patterns and to understand their biological relevance.

3.2.4.2. Effect of social environment on features of aggregates:

My results show that both mating status and social condition exert detectable differences on features of the observed aggregates. Mated flies formed fewer aggregates than virgins but these aggregates were longer lived. These aggregates also tended to have more individuals and subsequently occupied more space. Mated flies thus seem to show both a tendency to aggregate in larger numbers as well as a tendency to stay together for longer. Socially enriched flies formed more numerous but short-lived aggregates compared to socially deprived flies. However, social condition did not affect the size of the aggregate. Social condition thus seems to affect only the tendency of flies to stay together. These results reiterate that these factors may influence aggregation in distinct ways, possibly by affecting different behaviours within the aggregate.

The absence of interaction effects between mating and social condition on various features of the aggregate suggest that effects of mating and social condition are largely additive. This again supports the idea that these factors may affect independent behaviours. However, interaction

effects were seen for the density of aggregates, as the effect of mating was evident only in deprived flies (MD *vs.* VD) but not in enriched flies (ME *vs.* VE). Some caution needs to be exercised while interpreting this result as the values of mean density, as measured here, are biased. Since radial area increases non-linearly with increase in size, the change in density associated with unit change in size is different at different aggregate sizes. As aggregate size was not accounted for while averaging densities across sizes, differences in density may arise due to differences in aggregate sizes across treatments. That being said, the results for density do not mirror those for aggregate size, which suggests that density differences may not be mere reflections of size differences. Density values need to recalculated to verify these results before interpretation.

3.2.4.3. Effect of social environment on the evolution of aggregates across time:

Features of the aggregates changed over the course of the assay in a somewhat linear fashion. Hence, I quantified these changes using slopes of their profiles across time. Mating status and social condition had independent effects on these slope values, with mated flies and deprived flies often showing similar trends with respect to virgin flies and enriched flies respectively. Aggregate sizes and, consequently, their spread reduced over time. These declines were smaller for mated flies and for deprived flies which imply that these flies may show greater tendencies to stay near each other. In contrast, density increased over time, with slopes being larger for virgin flies and for enriched flies. This result can be explained by non-linear relationship between aggregate size and spread discussed above. Since, aggregate spread reduces disproportionately with size, the larger declines in size seen for virgin flies and enriched flies result in disproportionately larger declines in spread. Consequently, their ratio, i.e. the density is seen to increase. The number of aggregates decreased across time for deprived flies but it showed little change for enriched flies. This may be explained by the steeper increase in aggregate duration for deprived flies. As aggregates formed by deprived flies became more stable over time, their turnover reduced and thus fewer aggregates could be observed overall. Interestingly, this increase in duration over time was markedly steeper

for MD flies compared to the other treatments. Given this result, total number of aggregates in an interval would be expected to be much lower for MD flies compared to the other treatments. However, I did not see such differences in turnover rate of aggregates for MD flies. These results may be explained by the steeper increases in inter-quartile range (IQR) of aggregate durations for MD flies (although the differences were non-significant). Large values of IQR suggest that these flies may form several short-lived aggregates along with long lived ones. Thus, presence of some short-lived aggregates may allow turnover rate to remain unaffected even as the median duration increases due to the presence of long lived aggregates. However, it remains unclear why MD flies may show increased variation in the duration of aggregates.

### *3.3. Individual tendencies underlying aggregation*

#### *3.3.1. Rationale*

My results, so far, have demonstrated that social environment affects the aggregation pattern in *Drosophila* and that these changes are observable as changes in the features of aggregates. While the functional relevance of these changes is unclear, we can be certain that these are proximately linked to changes in the aggregation behaviour of the flies tested within the assay. An individual's tendency to aggregate is likely to be an outcome of several individual-level behaviours. These include a fly's preference to join aggregates as well as its preference to stay in aggregates. Hence, to understand the aggregation behaviour of individual flies, I quantified different individual tendencies for a fly, such as the overall time spent aggregating, time spent per aggregate and the tendency to join any aggregate. I also quantified the overall activity levels for each fly.

Tendencies to join or stay in aggregates may not be the same for each aggregate as flies may treat each aggregate differently. They may evaluate different aspects of an aggregate, such as its size or composition, while choosing to join or stay within it. While the motivations underlying these choices remain poorly understood, there is evidence to suggest that flies may indeed evaluate the size of an aggregate while aggregating (Saltz and Foley, 2011; Philippe et al., 2016). Hence, I tested for size dependence of the individual behaviours mentioned above by measuring the tendency to join an aggregate and the time spent per aggregate for differently sized aggregates. Finally, I measured these tendencies across social treatments to test if they differed with social environment.

An important aspect of aggregation that has been underexplored in flies is the composition of the aggregate. Flies may choose to preferentially aggregate with or specifically avoid each other depending on their motivations to aggregate. For example, if aggregation is purely a form of predator avoidance, then flies would be expected to join any aggregate, regardless of the identity of

the flies within it. However, if aggregation allows flies to participate in agonistic (eg. aggression) or non-agonistic (interactions mediating social learning) behaviours, then flies may actively choose their interaction partners for these behaviours. These scenarios would be expected to give rise to different patterns of association between individuals. In the former case, we would expect associations between flies to be random as flies would be associating with whichever individuals they happen to encounter. In the latter case, however, we would expect non-random associations between flies that could involve preferential aggregation or avoidance. To understand how flies may associate with each other, I calculated association indices for each pair of flies to quantify the relationship between individuals. I used randomization tests to determine if these indices could be explained by chance alone and found that flies did, in fact, show non-random spatial associations with each other. To understand how these relationships changed across time, I tested if associations between fly pairs were correlated across the first and second hours of the assay. Finally, I studied the effects of social environment on the pattern of spatial relationships viz. the proportion of non-random associations, the proportion of flies participating in non-random associations and the strength of these associations. I also tested the effect of social environment on the time correlations seen for these associations.

### *3.3.2. Methods*

#### 3.3.2.1. Quantifying individual behaviour:

To study how individual behaviours varied across time, each variable was measured over five minute intervals. Each of these variables was averaged across flies to obtain mean values of these behaviours for each replicate assay.

The overall tendency to aggregate for a fly was measured as the total number of frames within an interval in which a fly was found to be aggregating. This tendency itself depends on a fly's tendency to join new aggregates and its tendency to stay within an aggregate. Tendency to join

aggregates was measured as the proportion of unique aggregates visited within a five minute interval. Tendency to stay in an aggregate was measured as the mean number of frames between a fly's entry into an aggregate and its exit from it. In addition to aggregation related behaviours, the overall activity levels were also measured for each individual. Activity was measured as the mean velocity of the fly within a five minute interval. Velocity was calculated as the distance covered by a fly in a single frame. These values were calculated for each frame and then averaged across all frames within a five minute interval to estimate the mean velocity.

To test the effect of aggregate size on the aggregation behaviour of a fly, the tendencies to join and stay in aggregates were measured for aggregates of different sizes. If flies prefer joining aggregates of specific sizes, then aggregates of the preferred size are more likely to increase in size than others. Thus, tendencies to join aggregates of different sizes may be indirectly compared by comparing proportions of aggregates that increase in size for different aggregate sizes. For this calculation, every frame in which a fly joined an aggregate was identified and the aggregate's size in the previous frame was noted. In this fashion, instances of joining were counted for each aggregate size and for each fly. Counts for each aggregate size were averaged across flies to obtain the mean numbers of aggregates that increased in size. Since mean counts for each aggregate size depend on the total number of aggregates for that aggregate size, these counts were scaled by their corresponding abundances to obtain proportions of aggregates that increased in size for different aggregate sizes. To study the effect of size on the tendency to stay in an aggregate, duration of a fly's stay in an aggregate was noted along with the median size of the aggregate during the stay. Using these data, mean durations of stay were calculated for each fly and for each aggregate size. Finally, these values were averaged across flies to obtain mean stay durations for each aggregate size for a given replicate assay. The effect of size on these data was tested using a repeated-measures ANOVA with mating status and social condition as fixed factors and aggregate size as the repeated measures variable.

3.3.2.2. Associations between flies:

- Measuring association:

The strength of association between a pair of flies, A and B, was measured by calculating an association index (AI) as follows-

$$AI_{AB} = \frac{\#\ frames\ where\ A\ and\ B\ are\ in\ the\ same\ aggregate}{\#frames\ where\ either\ are\ in\ any\ aggregate}$$

$$= \frac{\#\ frames\ where\ A\ and\ B\ are\ in\ the\ same\ aggregate}{\begin{array}{c}\#\ frames\ where\ A\ is\ in\ an\ aggregate \\ +\ \#\ frames\ where\ B\ is\ in\ an\ aggregate \\ -\ \#frames\ where\ A\ and\ B\ are\ in\ the\ same\ aggregate\end{array}}$$

$$= \frac{N(A \cap B)}{N(A)+N(B)-N(A \cap B)}$$

Any association between flies depends on the flies' tendency to seek or avoid each other, as well as on random encounters within aggregates which would depend on their general tendency to be in aggregates. In the absence of non-random relationships (i.e. preference or aversion), associations between flies would depend only on the latter. In such a scenario, AI values would be predictable simply from the tendency of flies to be in aggregates. Conversely, if AI values fail to match such predictions, then we may infer existence of differentiated relationships between individuals. Thus, to test if non-random relationships are present among flies, observed association indices may be compared with indices obtained purely as a result of tendencies to aggregate for different flies.

- Randomisation tests:

Randomisation or permutation tests are commonly used tests of significance to test associations between paired variables such as the occurrences of individuals A and B in the same aggregate. Under the null hypothesis of no association, values of these variables are paired randomly, but do not conform to any specific distribution. Instead, they are obtained by shuffling pairings between different values of these variables from the observed dataset. Such shuffling is expected to break

associations between the variables such that values obtained after repeated shuffling represent the data expected when the variables show only random associations with each other. Variation across several such randomized datasets can be used to test significance of the observed dataset.

Randomisation tests were performed using custom MATLAB scripts where the composition of each aggregate was changed by assigning it randomly chosen flies. The size of an aggregate was maintained constant across randomisations to ensure that differences in associations were only due to changes in aggregate composition. The code first identified all instances of entry into and exit from an aggregate. Next, each frame from the video was recreated simply as the set of aggregates observed during that frame. Each aggregate consisted of empty positions to which flies could be assigned. The number of empty positions was determined by the size of that aggregate in that frame. Flies were sampled without replacement and assigned to these empty positions. The probability of a given fly being chosen varied across time and was determined by its empirical tendency to aggregate in the given time interval. Once all flies had been assigned to an aggregate, the composition did not vary across time until a new fly joined or left the aggregate. For frames showing fly entry, new empty positions were created in the relevant aggregates and new flies were assigned to them as above. In frames where exits occurred, the number of available positions in an aggregate was reduced by removing flies. Probability of removal also depended on the empirical tendency of a fly to be in aggregates such that flies which aggregated more were less likely to be removed. All removed flies were returned to the pool of unassigned flies that were available for reassignment to other aggregates. This exercise was performed for each frame until the end of the video, after which the AI was calculated for each pair of flies.

For each replicate video, 400 randomisations were performed, yielding 400 AI values for each pair of flies. The mean and standard deviation of these randomised AI values were used to calculate the Z-score and p-value associated with the observed AI value for each pair of flies. To accurately measure the total number of fly pairs with non-random AI values, the $\alpha$-level for each pair needed

to be corrected to prevent inflation of the family-wise error rate. This correction was performed using the Benjamini-Hochberg procedure with a false discovery rate of 5%. This procedure evaluates p-values for all observed comparisons using a rank based method and identifies significant values among them. Using this approach, all non-random AI values for a replicate could be identified.

- Patterns of spatial associations between flies:

All non-random relationships were classified as being positive or negative depending on whether the observed AI was significantly larger or smaller than the random expectation. For each type of association, frequency of occurrence, magnitude (or strength) of association and the participation of flies in such relationships were measured. These variables represent some aspects of the overall network of associations. Frequency of occurrence was measured as the proportion of fly pairs that showed significant relationships. Magnitude or strength of the association was measured as the mean AI value across all relationships of a given type. Participation of flies was measured as the proportion of flies that showed at least one non-random relationship of a given type.

To understand the effect of social environment on these patterns, these variables were compared across treatments using two-factor ANOVA with mating status and social condition as fixed factors. Any data that were seen to violate assumptions of normality were transformed before analysis.

- Correlation of association indices across time:

To understand how non-random relationships changed across time, separate sets of association indices were calculated for each hour of the assay and the correlation between them was tested. A significant positive correlation between these sets would suggest that spatial relationships are sustained over the course of the assay. Negative correlation, on the other hand, would suggest that flies choose to associate with flies that they haven't previously associated with, at the expense of

prior associations. A lack of correlation would indicate absence of any pattern suggesting that relationships vary randomly across time.

Correlations were tested separately for three datasets, one using all relationships regardless of significance and one for either type (i.e. positive or negative) of non-random relationship. As the significance of hourly AI values was not tested using randomisation, a relationship was classified as positive or negative based on the results of randomisation for the two hour data. All correlations were performed using Mantel tests with Spearman's rho as the correlation coefficient. Custom MATLAB scripts were used to perform these tests.

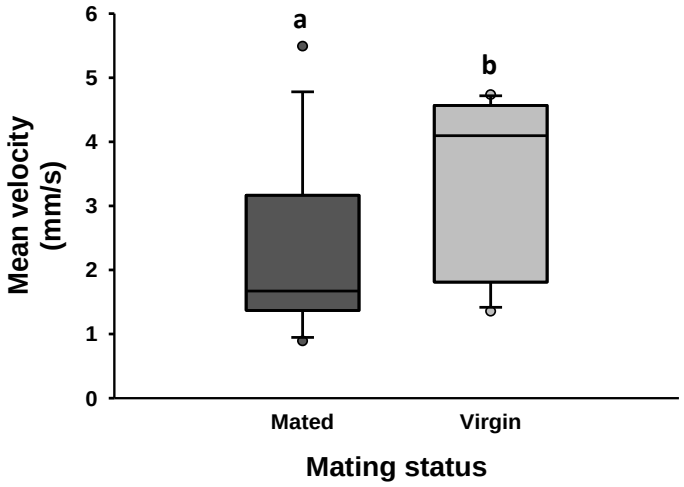
To understand the effect of social environment on changes in associations across time, correlation coefficients, as well as the proportion of replicates showing significant correlations, were compared across treatments. Correlation coefficients were compared using Kruskal-Wallis test while the proportion of replicates showing correlations were compared using Fisher's exact test.

*3.3.3. Results*

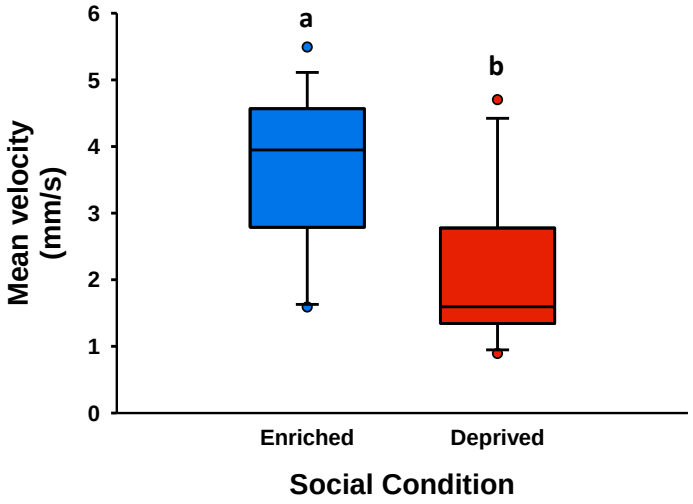
3.3.3.1. Effect of social environment on activity levels:

As the distribution of velocity values was skewed, the data were transformed ($y=x^{0.1}$) before analysis. Mated flies showed significantly lower activity compared to virgins ($F=11.75$, $df=1$, $p=0.0022$)(Fig. 3.3.1 while socially enriched flies showed higher activity levels compared to deprived flies ($F=24.81$, $df=1$, $p<0.0001$)(Fig. 3.3.1b). Velocity reduced with time ($F=61.71$, $df=23$, $p<0.0001$) in an exponential fashion. Although significant effects of all interactions between time and the other factors were detected, only the three-way interaction among social condition, mating status and time was found to be significant after correcting for lack of sphericity ($F=3.646$, $df_1=2.225$, $df_2=51.19$, $p=0.028$) (Fig.3.3.1c). To investigate the effects of social environment on the time profile of activity, exponential functions were fitted to the untransformed data ($y=a+be^{cx}$) and the parameters estimated for each replicate video. Two-way ANOVA on the estimated decay

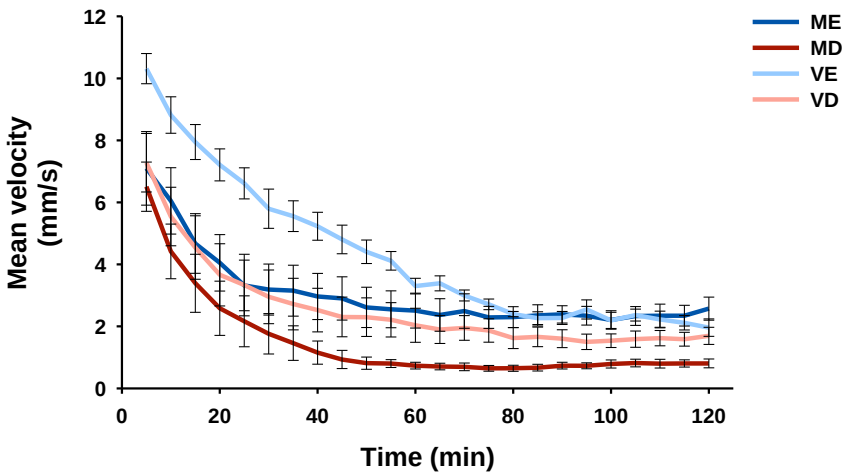
a)


Mean velocity (mm/s)
0
1
2
3
4
5
6
a
b
Mated
Virgin
Mating status

b)


Mean velocity (mm/s)
0
1
2
3
4
5
6
a
b
Enriched
Deprived
Social Condition

c)


Mean velocity (mm/s)
0
2
4
6
8
10
12
0
20
40
60
80
100
120
Time (min)
ME
MD
VE
VD

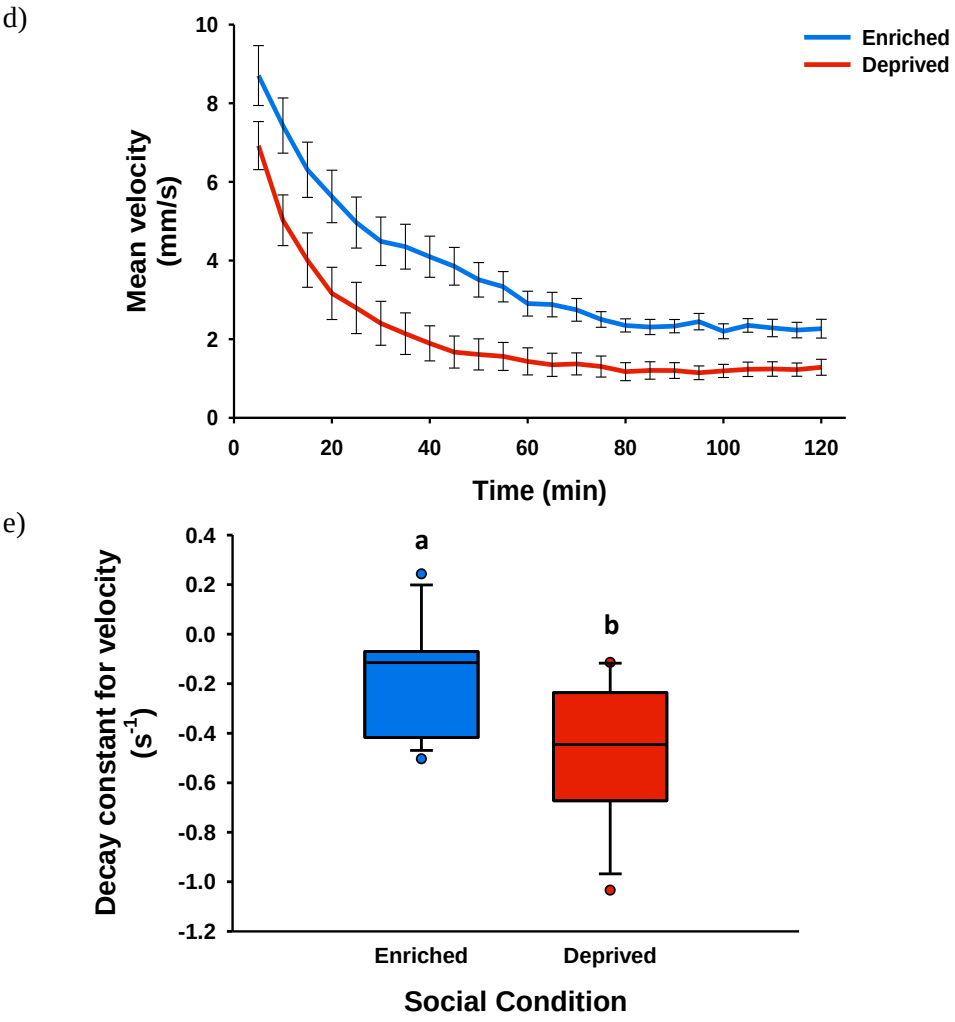



**Figure 3.3.1. Effect of social environment on activity levels in a five minute interval.**

Mean activity, quantified as the mean velocity, was **a)** significantly lower for mated flies compared to virgins and **b)** significantly higher for enriched flies compared to deprived flies (Main effect of mating status (p=0.0022) and social condition (p<0.0001) on transformed data ($y=x^{0.1}$)). **c)** Time profile for mean velocity was dissimilar across the four treatments (Interaction among mating status, social condition and time (p=0.028) on transformed data ($y=x^{0.1}$)). **d)** Differences in time profiles of mean velocity for enriched and deprived flies were not significant but **e)** decay constants ($c$) of exponential fits ($y=a+be^{cx}$) to these profiles were different between enriched and deprived flies. Reduction in velocity with time was significantly lower (i.e. value was less negative) for enriched flies compared to deprived flies (Main effect of social condition (p=0.003) on decay constants ($c$)).

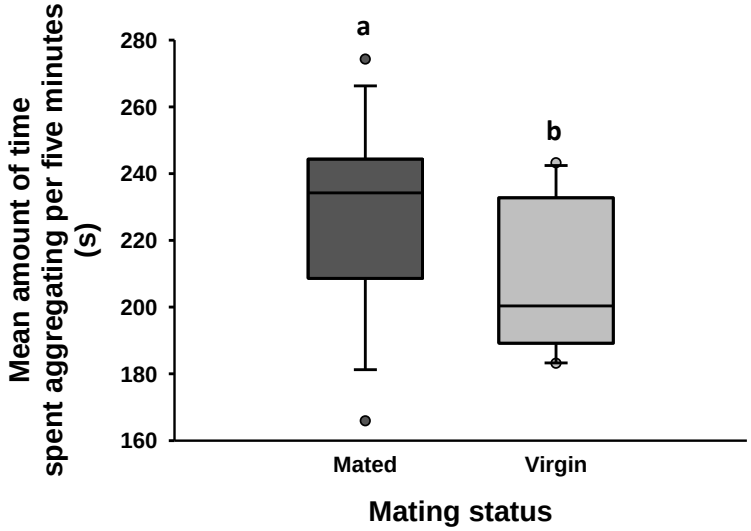
[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM].

parameters revealed an effect of social condition (F=11.09, df=1, p=0.003) (Fig. 3.3.1d,e) which could be attributed to the shallower decay seen for socially enriched flies compared to deprived flies. While it is unclear what contributes to the three-way interaction observed above, visual inspection of the time profiles suggests it may be due to higher activity seen for VE flies in the first 80 minutes of the video.

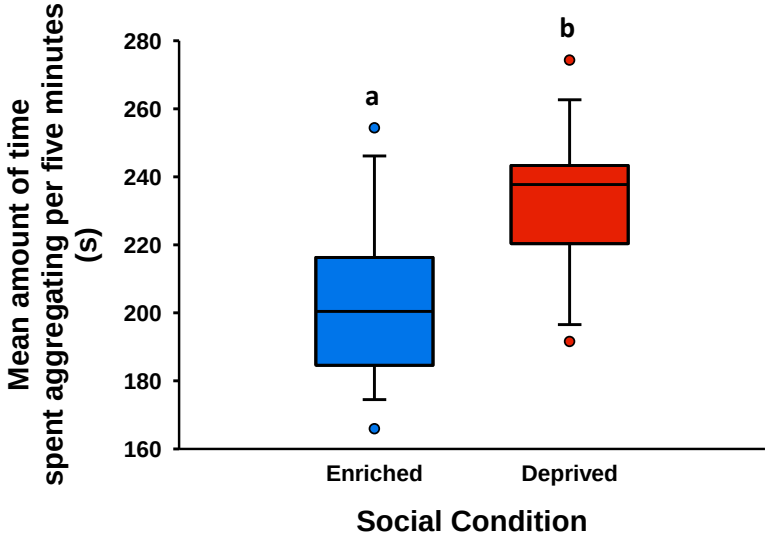
3.3.3.2. Effect of social environment on the time spent in aggregates:

Mated flies spent significantly greater time in aggregates compared to virgins (F=7.95, df=1, p=0.009) (Fig. 3.3.2a). Socially enriched flies spent significantly lesser time in aggregates than socially deprived flies (F=15.172, df=1, p=0.0007) (Fig. 3.3.2b). In general, the time spent in aggregates by a fly increased with time (F=23.76, df=23, p<0.0001). Significant effects of the interaction between time and mating status (Fig. 3.3.2c) as well as the three way interaction between time, mating status and social condition were detected. However, only the two way interaction was significant after correcting for lack of sphericity (F=2.673, $df_1$=3.33, $df_2$=76.6, p=0.047). To analyze the effect of social environment on profiles of time spent in aggregates, sigmoid curves were fitted to the data ($y=a+b/(1+e^{-cx})$) and a two-way ANOVA was performed on the estimated values of the steepness parameter using mating status and social condition as factors. As these data showed heteroscedasticity, they were transformed ($y=x^{1/3}$) before analysis. A significant effect of social condition on the steepness of the curve could be detected (F=5.59, df=1, p=0.026), with socially enriched flies showing a nearly linear increase with time (Fig. 3.3.2d,e). Socially deprived flies, on the other hand, showed a rapid increase in the time spent aggregating followed by a plateau. Owing to lack of differences in any parameter due to mating status, the interaction between time and mating status could not be explained. However, visual inspection of time profiles suggests that differences between mated and virgin flies reduced over time such that they spent similar amounts of time aggregating by the end of the video.

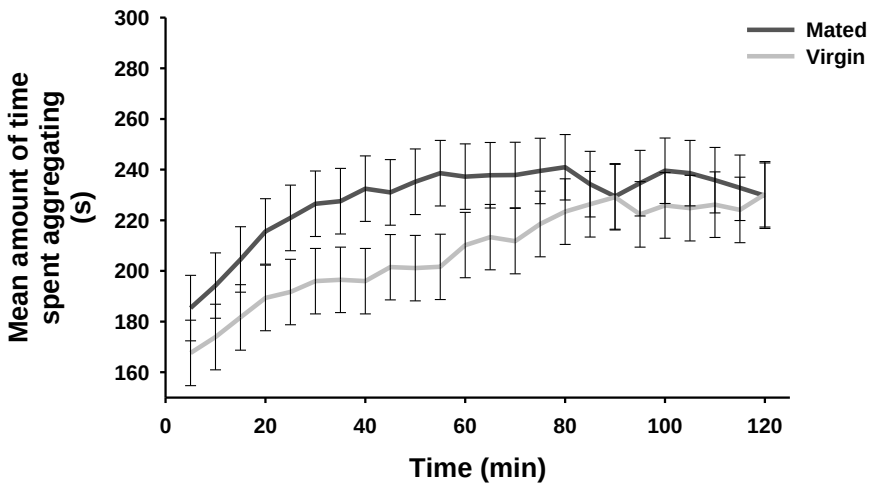
a)


Mean amount of time spent aggregating per five minutes (s)
280
260
240
220
200
180
160
a
b
Mated
Virgin
Mating status

b)


Mean amount of time spent aggregating per five minutes (s)
280
260
240
220
200
180
160
a
b
Enriched
Deprived
Social Condition

c)


Mean amount of time spent aggregating (s)
300
280
260
240
220
200
180
160
0
20
40
60
80
100
120
Time (min)
Mated
Virgin

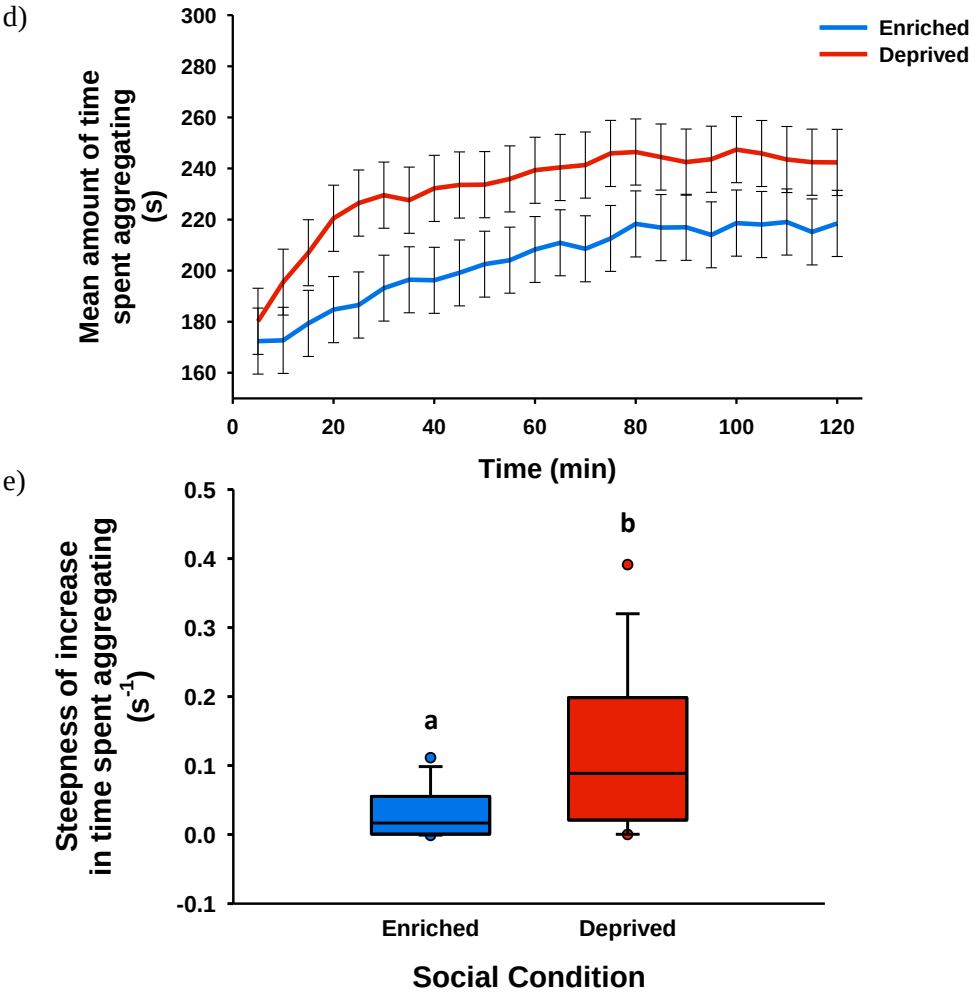



**Figure 3.3.2. Effect of social environment on a fly's tendency to aggregate in a five minute interval.**

Tendency of a fly to aggregate was quantified as the amount of time that a fly spent in aggregates in a five minute interval. Time spent in aggregates was **a)** significantly higher for mated flies compared to virgins and **b)** significantly lower for enriched flies compared to deprived flies (Main effect of mating status (p=0.009) and social condition (p=0.0007)). Time profiles of time spent in aggregates were **c)** significantly different between mated and virgin flies but **d)** not between enriched and deprived flies (Interaction between time and mating status (p=0.047)). When steepness parameters of sigmoid fits to these profiles ($y=a+b/(1+e^{-cx})$) were compared, no differences were observed between mated and virgin flies but **e)** significantly shallower increases in time spent aggregating were seen for enriched flies compared to deprived flies (Main effect of social condition (p=0.026) on steepness of curves ($c$)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote 95% CIs].

3.3.3.3. Effect of social environment on the tendency to join aggregates:

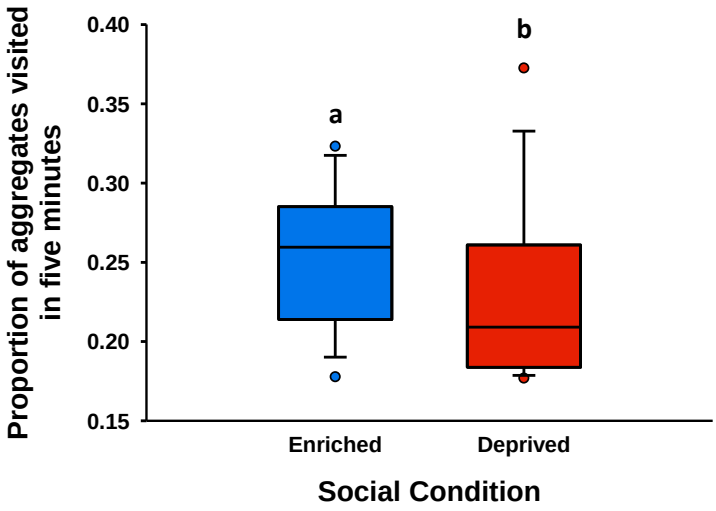
- Effect of time:

The data were transformed (y=1/x) to correct for skewness. Socially enriched flies were seen to join a significantly greater proportion of the available aggregates within a time interval (F=5.048, df=1, p=0.035) (Fig. 3.3.3a). A significant interaction between mating status and social condition was also seen (F=4.5, df=1, p=0.045). Tukey's post hoc analysis revealed that VE flies joined a significantly larger fraction of the available aggregates than VD flies (p=0.02) (Fig. 3.3.3b). However, none of the other comparisons were significant. In general, the fraction of aggregates that were visited by a fly reduced over time (F=18.1147, df=23, p<0.0001). Significant effects of the interactions between time and mating status and between time and social condition were also detected. However, only the interaction between time and mating status was found to be significant after correcting for lack of sphericity (F=2.367, $df_1$=6.92, $df_2$=159.2, p=0.025) (Fig. 3.3.3c). To further analyse these results, exponential decay functions ($y=a+b*e^{cx}$) were fitted to the time series. However, no effects of social environment could be detected for any of the fitted parameters. Visual inspection of the time series also did not reveal any clear differences between the time profiles for mated and virgin flies.

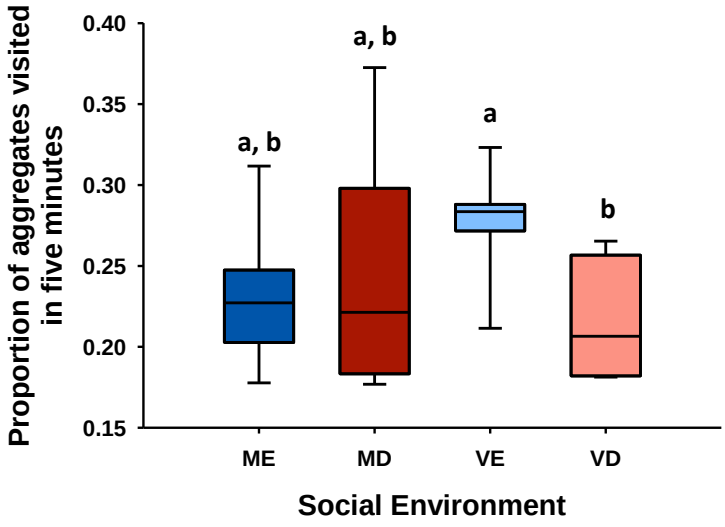
- Effect of aggregate size:

Mated flies showed a greater tendency to join aggregates compared to virgins but this difference was marginally non-significant (F=3.963, df=1, p=0.058). Tendency to join aggregates was affected significantly by the size of the aggregate (F=40,2765, df=23, p<0.0001) (Fig. 3.3.3d) as flies joined a greater proportion of mid-sized (9-12 flies) aggregates compared to small or large aggregates. A significant interaction between aggregate size and mating status could be detected but this effect was not significant (F=2.088, $df_1$=2.4, $df_2$=55.4, p=0.12) after correcting for the lack of sphericity.

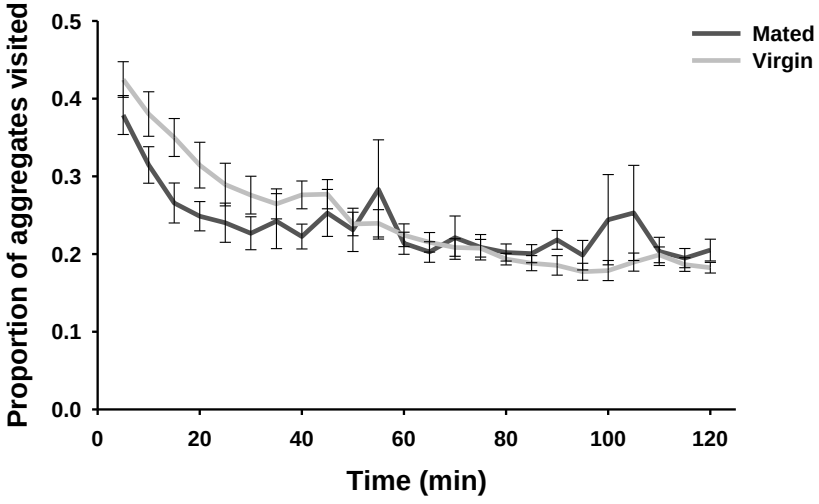
a)


Proportion of aggregates visited in five minutes
0.40
0.35
0.30
0.25
0.20
0.15
a
b
Enriched
Deprived
Social Condition

b)


Proportion of aggregates visited in five minutes
0.40
0.35
0.30
0.25
0.20
0.15
a, b
a, b
a
b
ME
MD
VE
VD
Social Environment

c)


Proportion of aggregates visited
0.5
0.4
0.3
0.2
0.1
0.0
0
20
40
60
80
100
120
Time (min)
Mated
Virgin

d)

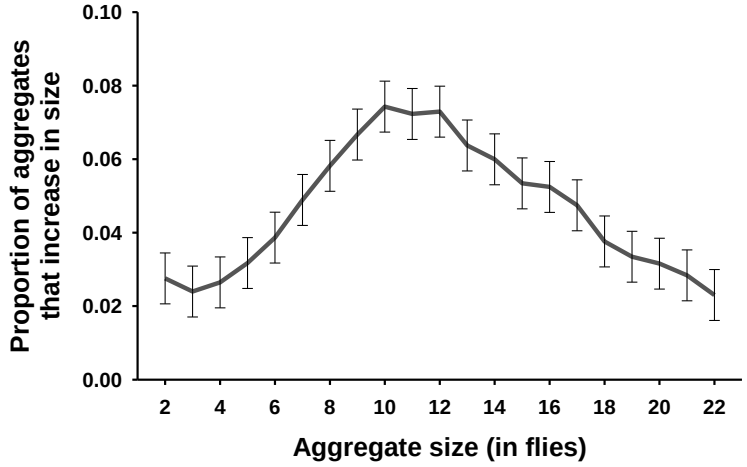



**Figure 3.3.3. Effect of social environment on the tendency to join aggregates.**

Tendency to join aggregates was quantified as the number of aggregates visited by a fly in a five minute interval. **a)** Enriched flies were seen to visit a greater proportion of aggregates compared to socially deprived flies (Main effect of social condition ($p=0.035$) on transformed data ($y=1/x$)). **b)** Virgin and socially enriched (VE) flies were seen to visit significantly larger proportion of aggregates compared to virgin and socially deprived (VD) flies (Interaction between mating status and social condition ($p=0.045$) on transformed data ($y=1/x$)). Tukey's HSD test, VE *vs.* VD- $p=0.02$). **c)** Time profiles of proportion of aggregates visited were found to be different for mated and for virgin flies (Interaction between time and mating status ($p=0.025$) on transformed data ($y=1/x$)) but values of fitted ($y=a+be^{cx}$) parameters were not different across treatments.
The effect of aggregate size on the tendency to join aggregates was tested by measuring the proportion of aggregates that grew in size for each aggregate size. **d)** Proportion of aggregates that grew in size was high for mid-sized aggregates (8-14) compared to smaller or larger aggregates (Main effect of size ($p<0.0001$)). This pattern was not different across treatments.

[For box-whisker plots, horizontal line denotes median, box edges denote $25^{th}$ and $75^{th}$ percentiles and whiskers denote $10^{th}$ and $90^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other]. [For time profiles, error bars denote SEM]. [For profiles across sizes, error bars denote 95% CIs].

3.3.3.4. Effect of social environment on time spent in an aggregate:

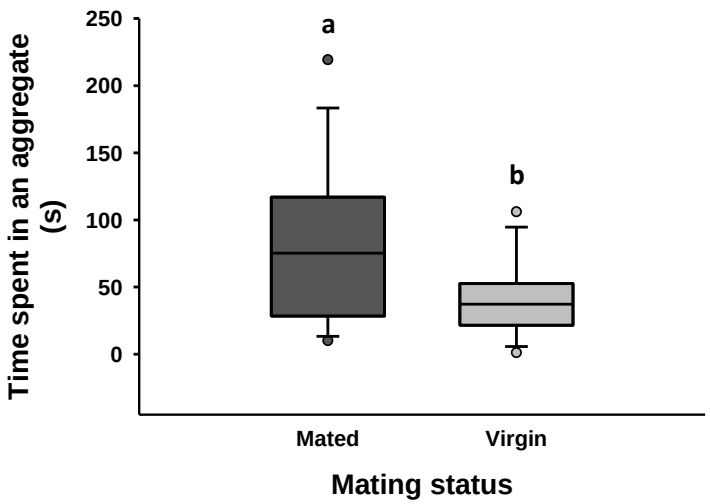
- Effect of time:

Since the data were skewed and showed heteroscedasticity, they were transformed ($y=x^{1/4}$) prior to analysis. Mated flies were seen to spend significantly more time per aggregate than virgin flies ($F=10.11$, $df=1$, $p=0.004$) (Fig. 3.3.4a). Socially enriched flies spent significantly less time per aggregate than socially deprived flies ($F=15.828$, $df=1$, $p=0.0006$) (Fig. 3.3.4b). In general, the time spent per aggregate increased with time ($F=26.096$, $df=23$, $p<0.0001$). A significant effect of the interaction between mating status and time was also seen ($F=2.821$, $df=23$, $p<0.0001$) (Fig. 3.3.4c). Owing to considerable variation in the shape of time profiles across replicates, a single non-linear function could not be found to describe them accurately. Hence, linear slopes were calculated to quantify the general trend across time. A two-way ANOVA on these slope values revealed a significant effect of social condition ($F=4.369$, $df=1$, $p=0.047$) (Fig. 3.3.4d,e). Socially enriched flies showed a gradual increase in the time spent per aggregate over time while the increase was seen to be steeper for socially deprived flies. Mating status had no significant effect on the slope of the time profile. Visual examination of the profiles for mated and virgin flies suggested that differences between mated and virgin flies were most pronounced in the middle of the video (30-90 min) but were largely similar towards the end.

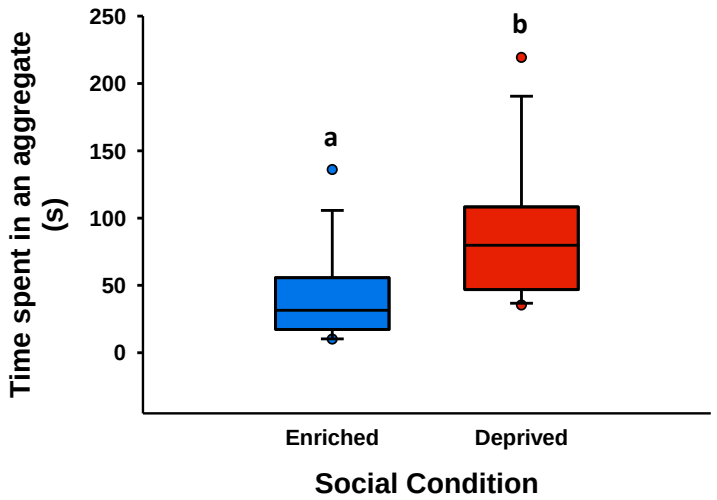
- Effect of aggregate size:

The data were transformed ($y=\log(x)$) before analysis to correct for the lack of normality and homoscedasticity. As above, a main effect of social condition ($F=13.68$, $df=1$, $p=0.001$) was detected, but the effect of mating status failed to show significance ($F=3.373$, $df=1$, $p=0.079$). Significant effects of aggregate size, an interaction between size and social condition and the three-way interaction between mating status, social condition and size were also detected. However, only the three-way interaction was found to be significant after correcting for the lack of sphericity

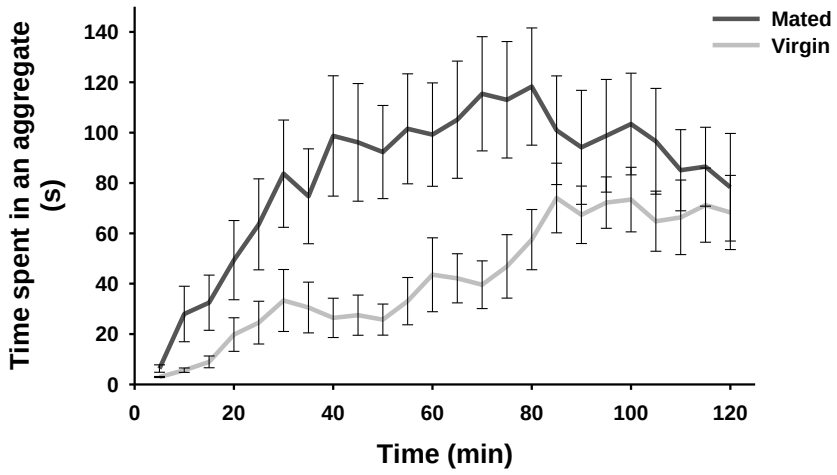
a)


Time spent in an aggregate (s)
250
200
150
100
50
0
a
b
Mated
Virgin
Mating status

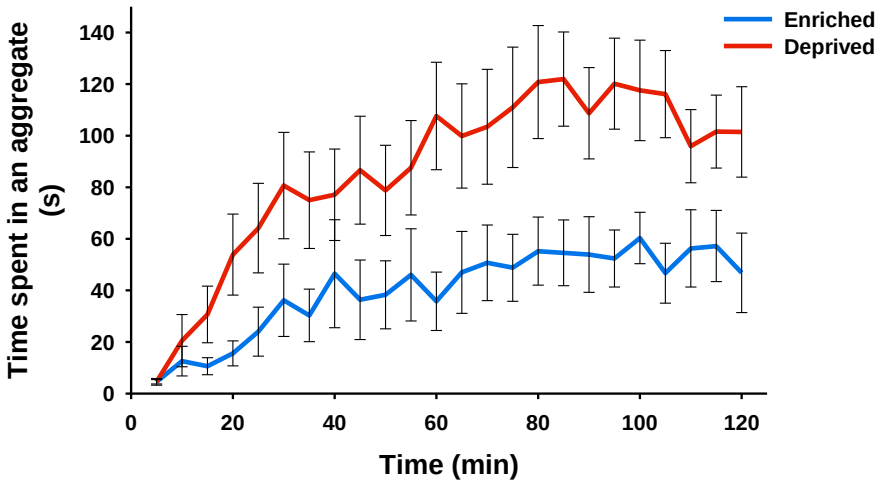
b)


Time spent in an aggregate (s)
250
200
150
100
50
0
a
b
Enriched
Deprived
Social Condition

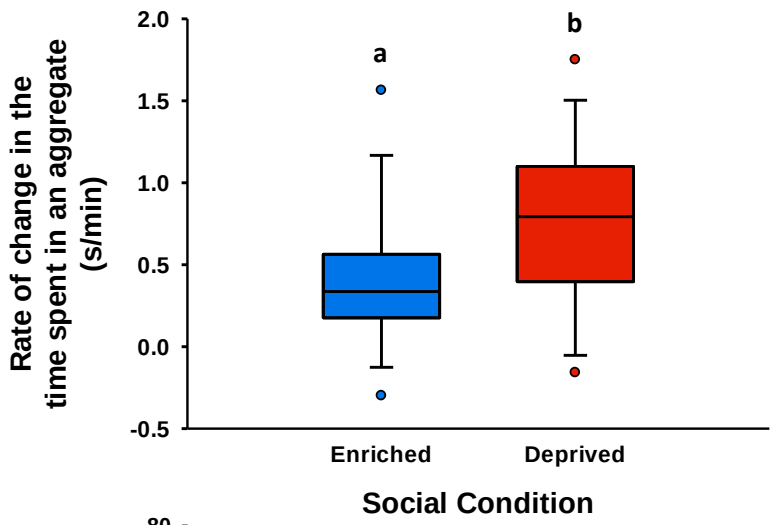
c)


Time spent in an aggregate (s)
140
120
100
80
60
40
20
0
0
20
40
60
80
100
120
Time (min)
Mated
Virgin

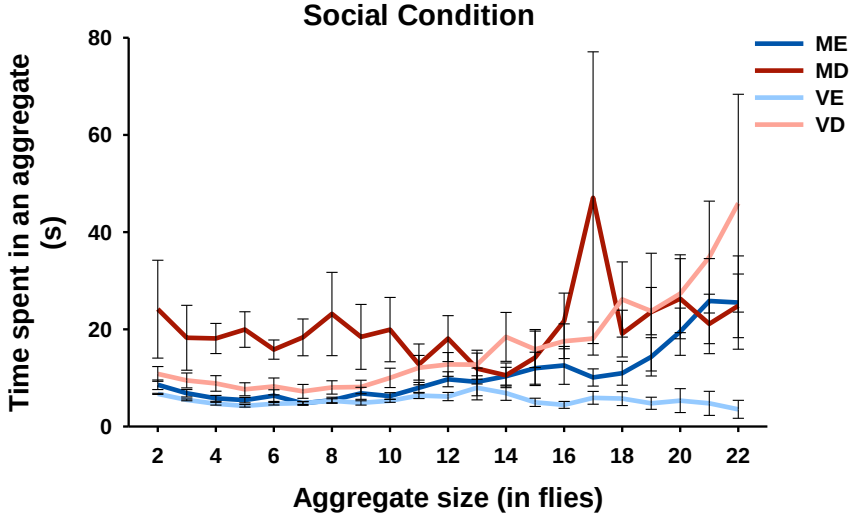
d)


Enriched
Deprived
Time spent in an aggregate (s)
0
20
40
60
80
100
120
140
Time (min)

e)


Rate of change in the time spent in an aggregate (s/min)
2.0
1.5
1.0
0.5
0.0
-0.5
a
b
Enriched
Deprived
Social Condition

f)


ME
MD
VE
VD
Time spent in an aggregate (s)
0
20
40
60
80
2
4
6
8
10
12
14
16
18
20
22
Aggregate size (in flies)

g)

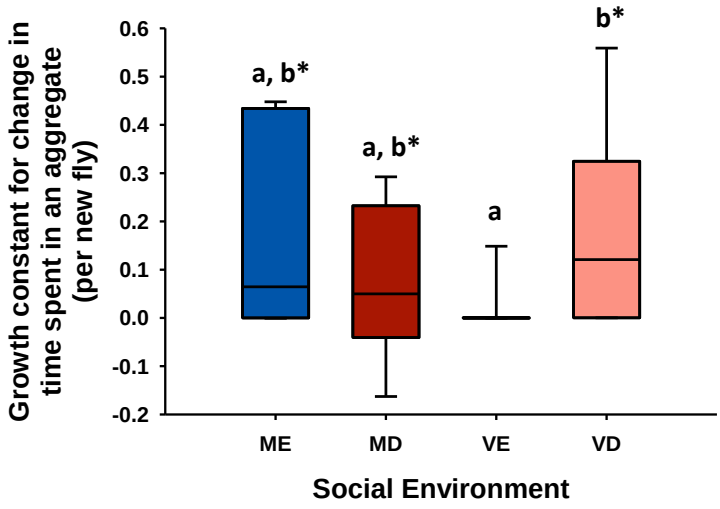



**Figure 3.3.4. Effect of social environment on tendency to stay in an aggregate.**

Tendency to stay in an aggregate was quantified as the mean amount of time spent in an aggregate in a five minute interval. Mean time spent in an aggregate was **a)** significantly higher for mated flies compared to virgins and **b)** significantly lower for enriched flies compared to deprived flies (Main effect of mating status ($p=0.004$) and social condition ($p=0.0006$) on transformed data ($y=x^{1/4}$)). Time profiles of time spent per aggregate were **c)** significantly different between mated and virgin flies but **d)** not significantly different between enriched and deprived flies (Interaction between time and mating status ($p<0.0001$)). Analysis of linear slopes showed that the rates of change in time spent per aggregate were not different for mated and virgin flies but were **e)** significantly lower for enriched flies compared to deprived flies (Main effect of social condition ($p=0.047$) on linear slope calculated using untransformed data).

Effect of aggregate size on tendency to stay in an aggregate was tested by measuring time spent in an aggregate as a function of the aggregate's median size. **f)** Profiles of time spent per aggregate across sizes were found to be dissimilar across treatments (Interaction among mating status, social condition and aggregate size on transformed data ($y=log(x)$). **g)** Growth constants ($c$) of exponential fits ($y=a+be^{cx}$) to these profiles were dissimilar across treatments (Interaction between mating status and social condition on transformed values($y=x^{1/3}$) of growth constants ($c$)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other. b* denotes that significance is seen using Fisher's LSD but not Tukey's HSD]. [For time and size profiles, error bars denote SEM].

($F=4.4137$, $df_1=3.684$, $df_2=84.74$, $p=0.003$) (Fig. 3.3.4f). To investigate these effects, time spent in aggregates was expressed as an exponential function of aggregate size ($y=a+be^{cx}$) and the growth constant was estimated. Values of growth constant were transformed ($y=x^{1/3}$) to reduce heteroscedasticity and analysed using a two-way ANOVA. This analysis revealed an effect of the interaction between mating status and social condition on the value of the growth constant. Although Tukey's HSD failed to identify differences between treatments, Fisher's LSD detected a significantly higher value of the growth constant for VD flies compared to VE flies ($Error_{MS}=0.1142$, $df=23$, $p=0.0426$) (Fig. 3.3.4g).

3.3.3.5. Effect of social environment on relationships between flies:

Significantly non-random relationships between flies were observed in every replicate. Both significantly larger (positive) and significantly smaller (negative) AI values could be observed. Since several proportion values were close to one, a factorial ANOVA could not be used to test the effect of social environment on these data. Hence, these data were analyzed using the Kruskal-Wallis test. For this analysis, each social treatment was considered to be a separate level of the same fixed factor.

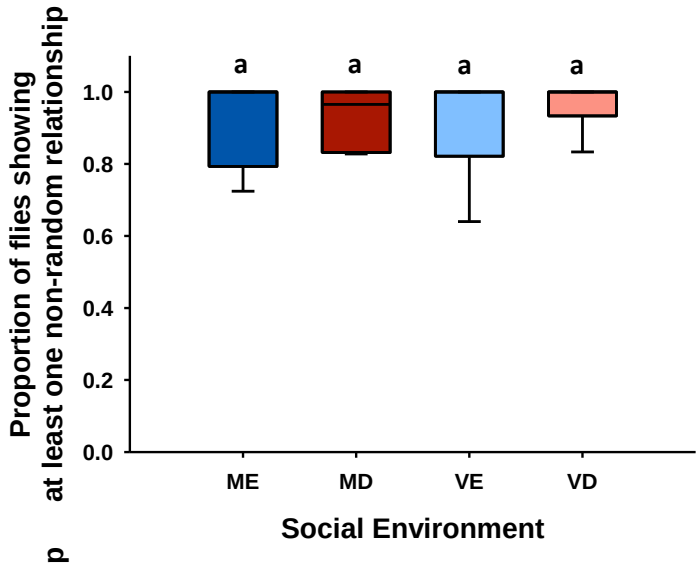
- Differences in the proportion of significant relationships-

The proportion of significant relationships was not affected significantly by social environment (Fig. 3.3.5d). No differences were seen when the proportions of positive (Fig. 3.3.5e) and negative relationships (Fig. 3.3.5f) were analyzed separately. (Values listed in Table 3.3.1.)

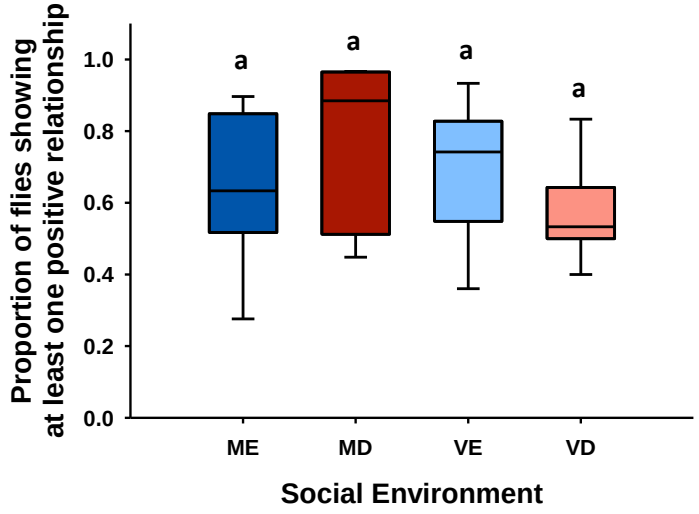
- Differences in the strength of association-

Mean AI value of positive relationships was significantly higher for mated flies compared to virgins ($F=4.4978$, $df=1$, $p=0.045$) (Fig. 3.3.6a).

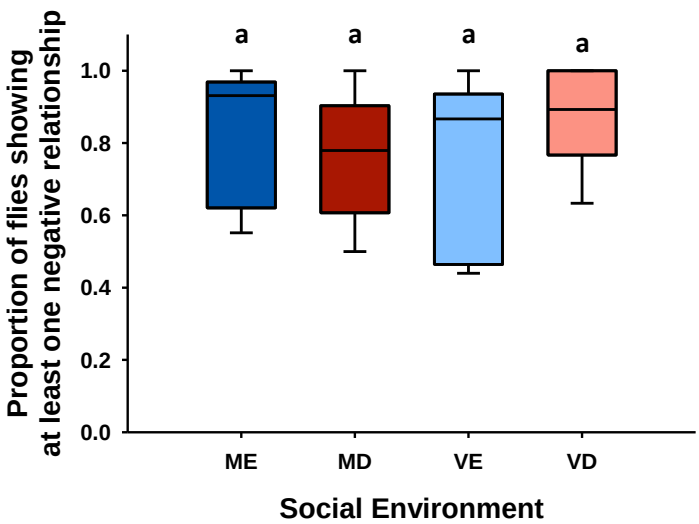
a)


Proportion of flies showing at least one non-random relationship
0.0
0.2
0.4
0.6
0.8
1.0
a
a
a
a
ME
MD
VE
VD
Social Environment

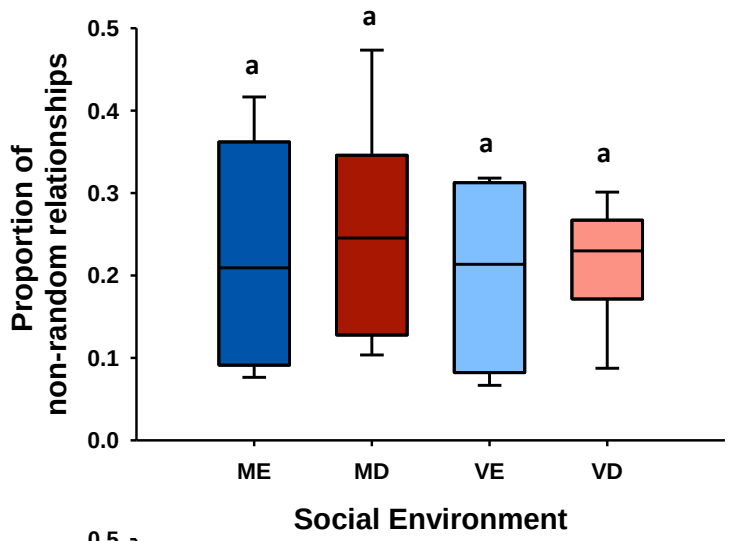
b)


Proportion of flies showing at least one positive relationship
0.0
0.2
0.4
0.6
0.8
1.0
a
a
a
a
ME
MD
VE
VD
Social Environment

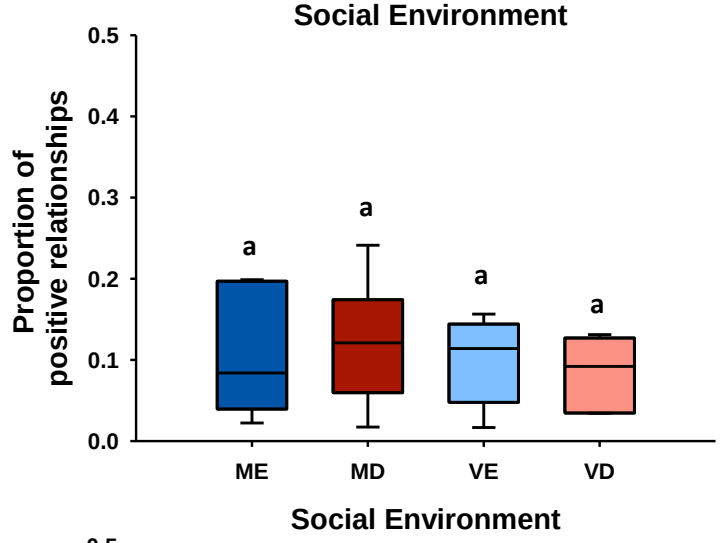
c)


Proportion of flies showing at least one negative relationship
0.0
0.2
0.4
0.6
0.8
1.0
a
a
a
a
ME
MD
VE
VD
Social Environment

d)


Proportion of non-random relationships
0.5
0.4
0.3
0.2
0.1
0.0
a
a
a
a
ME
MD
VE
VD
Social Environment

e)


Proportion of positive relationships
0.5
0.4
0.3
0.2
0.1
0.0
a
a
a
a
ME
MD
VE
VD
Social Environment

f)

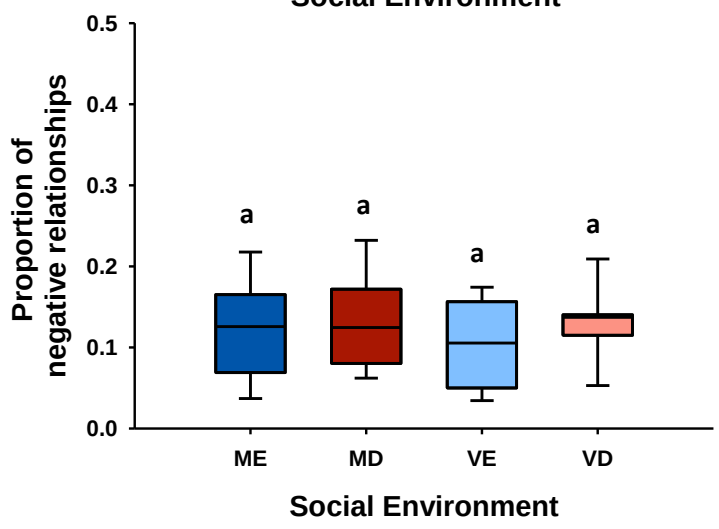

Proportion of negative relationships
0.5
0.4
0.3
0.2
0.1
0.0
a
a
a
a
ME
MD
VE
VD
Social Environment

**Figure 3.3.5. Effect of social environment on the structure of spatial relationships.**

Spatial relationships between flies were considered to be non-random if the observed values of association index (AI) were significantly different from randomized values. Relationships with AI values greater than the randomized expectation were considered to be positive while those less than the randomized expectation were considered to be negative. Social environment did not affect the proportions of flies showing **a)** at least one non-random relationship, **b)** at least one positive relationship or **c)** at least one negative relationship. Social environment did not affect the proportion of fly-pairs showing **d)** any non-random relationship, **e)** positive relationships or **f)** negative relationships.

[For box-whisker plots, horizontal line denotes median, box edges denote $25^{th}$ and $75^{th}$ percentiles and whiskers denote $10^{th}$ and $90^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other].

| | | ME | MD | VE | VD |
|---|---|---|---|---|---|
| Proportion of flies showing at least one non-random relationship | All relationships | 0.917 ± 0.116 | 0.932 ± 0.08 | 0.923 ± 0.142 | 0.958 ± 0.063 |
| | Positive relationships | 0.636 ± 0.207 | 0.78 ± 0.23 | 0.685 ± 0.193 | 0.573 ± 0.136 |
| | Negative relationships | 0.839 ± 0.181 | 0.762 ± 0.175 | 0.779 ± 0.231 | 0.866 ± 0.133 |
| Proportion of fly pairs showing a non-random relationship | All relationships | 0.227 ± 0.129 | 0.251 ± 0.132 | 0.207 ± 0.102 | 0.214 ± 0.074 |
| | Positive relationships | 0.103 ± 0.07 | 0.121 ± 0.076 | 0.100 ± 0.053 | 0.081 ± 0.04 |
| | Negative relationships | 0.124 ± 0.061 | 0.130 ± 0.061 | 0.107 ± 0.052 | 0.133 ± 0.046 |

**Table 3.3.1. Data for pattern of spatial relationships.**

Participation of flies in spatial relationships was measured as the proportion of flies showing at least one non-random relationship. No differences were seen in values of fly participation across treatments. Density of relationships was measured as the proportion of fly-pairs (i.e. possible relationships) that showed non-randomness. No differences were seen in values of density across treatments.

[Values shown as mean ± standard deviation].

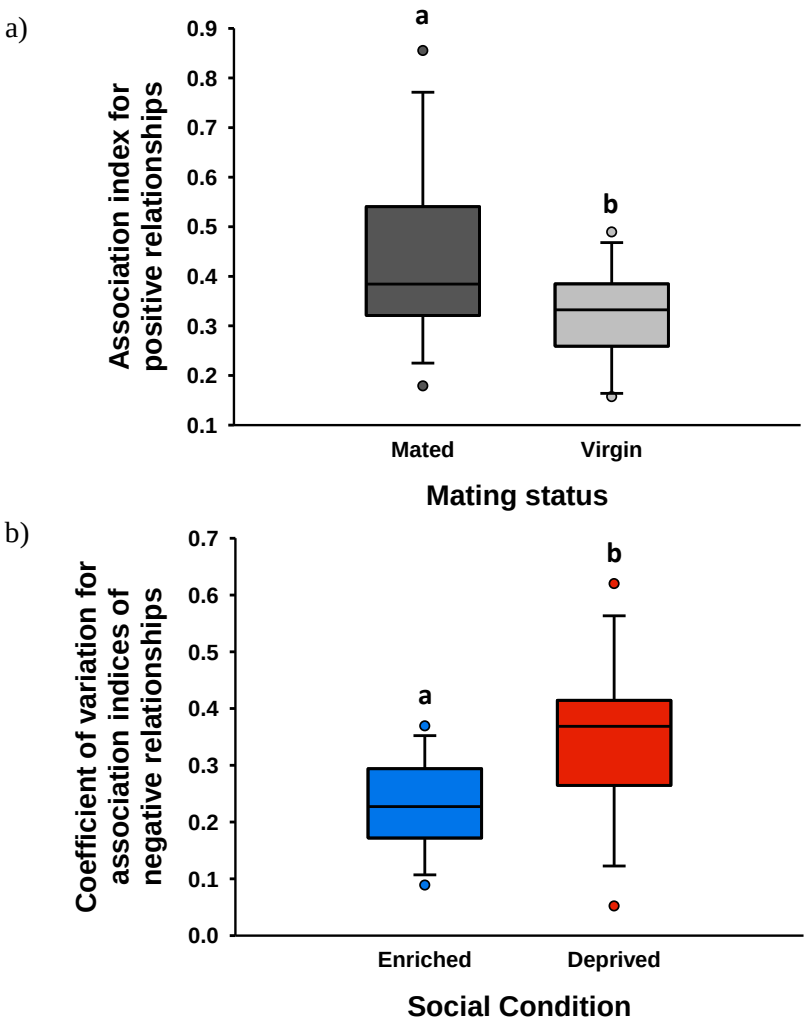



**Figure 3.3.6. Effect of social environment on strength of associations.**

**a)** Values of association index (AI) for positive relationships were significantly larger for mated flies compared to virgins (Main effect of mating status (p=0.045)). **b)** Variation in AI values for negative relationships was significantly lesser for enriched flies compared to deprived ones (Main effect of social condition (p=0.008)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other].

Data for AI values for negative relationships were transformed ($y=\sqrt{x}$) before analysis. Social environment had no effect on the mean AI for negative relationships. When the effect of social environment was tested on coefficients of variation, socially enriched flies were seen to show significantly lower values than socially deprived flies ($F=8.37$, $df=1$, $p=0.008$) (Fig. 3.3.6b).

- Differences in the proportion of flies showing significant relationships-

The overall proportion of flies that showed at least one significant relationship was not significantly different across treatments (Fig. 3.3.5a). Significant effects of social environment were not observed even when the proportions of flies showing positive (Fig. 3.3.5b) and negative relationships (Fig. 3.3.5c) were analyzed separately. (Values listed in Table 3.3.1.) Overall, a greater proportion of flies were seen to show negative relationships compared to positive relationships.

3.3.3.6. Effect of social environment on correlations between associations across time:

Several replicate videos showed significant correlations between association indices for the first and second hours. Fisher's exact test revealed that the proportion of videos showing correlations across time was significantly different across social treatments ($p=0.005$). To identify which treatments showed significant differences, pair-wise Fisher's exact tests were performed. Although VE flies were found to show correlations less frequently compared to MD ($p=0.005$) and VD ($p=0.029$) flies, only the former difference was significant after applying a Bonferroni correction. Correlation coefficients were not significantly different across treatments (see Table 3.3.2.).

When correlations were tested using only positive relationships, the proportion of replicates showing significant correlations was not significantly different across treatments. Correlation coefficients were affected by social environment but these effects failed to reach statistical significance ($p=0.055$) (see Table 3.3.2.).

| | | ME | MD | VE | VD |
|---|---|---|---|---|---|
| Proportion of replicates showing correlations across time | All relationships | 0.714 (5 out of 7) | 1 (6 out of 6) | 0.142 (1 out of 6) | 0.857 (6 out of 7) |
| | Positive relationships | 0.429 (3 out of 7) | 0.333 (2 out of 6) | 0.571 (4 out of 7) | 0.429 (3 out of 7) |
| | Negative relationships | 0 (0 out of 7) | 0.5 (3 out of 6) | 0.571 (4 out of 7) | 0.714 (5 out of 7) |
| Range of correlation coefficients | All relationships | 0.273 to 0.654 | 0.222 to 0.612 | 0.423 | 0.329 to 0.503 |
| | Positive relationships | 0.191 to 0.8 | 0.332 to 0.714 | -0.486 to -0.367 | -0.21 to 0.634 |
| | Negative relationships | 0 | -0.404 to 0.484 | -0.575 to -0.257 | -0.522 to 0.58 |

**Table 3.3.2. Data for correlations between association indices across time.**

Stability of spatial relationships was quantified by checking for correlations between association indices calculated separately for each hour of the assay. More MD replicates showed correlations across time for non-random relationships compared to VE replicates (Fisher's exact test, p=0.005). Unlike other treatments, ME flies did not show correlations across time for negative relationships but this difference was not significant. Large variation was seen in values of correlation coefficients across treatments and differences across treatments were not significant. Unlike other treatments, VE flies showed negative correlations across time for both positive and negative relationships.

[Highlighted cells are significantly different from each other]

A significant effect of social environment (p=0.034) was seen on the proportion of replicates showing significant correlations for negative relationships. This difference was due to lack of correlations in all replicate videos of ME flies (ME *vs.* MD, p=0.07 ; ME *vs*. VE, p=0.07; ME *vs*. VD, p=0.021) although these comparisons were not significant after applying a Bonferroni correction. Correlation coefficients did not vary across treatments. (see Table 3.3.2.)

*3.3.4. Inferences*

3.3.4.1. Activity levels and the tendency to aggregate:

Activity levels and the time spent aggregating were seen to show an inverse relationship with each other. Socially enriched flies and virgin flies showed greater activity and correspondingly spent lesser time in aggregates compared to deprived flies and mated flies respectively. This inverse relationship was also reflected in the time profiles of velocity and time spent aggregating seen for socially enriched and deprived flies. This relationship between activity and time spent aggregating suggests that flies may aggregate when they are less active.

Curiously, the results for activity did not match results reported previously in literature. Mated flies have been seen to be more active relative to virgins at the sampled time points (Isaac et al. 2010) while socially enriched flies are known to show lower activity per waking minute when compared to isolated flies (Ganguly-Fitzgerald et al. 2006). This reversal of differences may be explained by the fact that, unlike in my experiments, these studies quantified activity of flies that were kept in separate tubes and were thus moving in the absence of other flies. Thus, effects of prior social environment on activity levels may depend on the social environment of flies during the assay.

3.3.4.2. Tendency to join aggregates:

By and large, variation across treatments in the tendency to join aggregates was limited. Although tendency to join aggregates was higher for socially enriched flies, these differences seem to be due

to the large values seen for VE flies seen only in the first hour of the assay. This behaviour can be explained by the tendency of VE flies to spend less time per aggregate during the first hour (Fig.1c). Higher levels of activity shown by these flies during the first hour would also allow them to visit more aggregates.

To understand how a fly might be choosing aggregates to join, I looked at the proportions of aggregates that increased in size as a function of their sizes. My results show a curious non-linear pattern where the proportion of aggregates that attracted a new fly increased with aggregate size until a size of nine flies was reached and declined thereafter. While this suggests that a size of nine flies may be some preferred size for aggregation, the relevance of this number is unclear. This pattern is different from results reported previously (Saltz, 2011, Philippe et al., 2016) where flies were seen to preferentially join larger aggregates regardless of the absolute number of flies in the available aggregates. Since these studies defined an aggregate using environmental heterogeneities, either as a food patch (using males, Saltz, 2011) or a thermal refuge (using females, Philippe et al., 2016), these differences may be attributed to differences in the definition of an aggregate used in this study. However, it is unclear how or why this may impact the tendency to join aggregates. Alternatively, this decline may result from depletion of available flies with increases in size of the aggregate. This may be tested by assaying smaller or larger groups of flies within the assay. If depletion of flies underlies the observed patterns, then the preferred aggregate size may change with the total number of flies in the arena.

3.3.4.3. Tendency to stay in aggregates:

On average, socially enriched flies and virgin flies spent less time in an aggregate compared to deprived flies and virgin flies respectively. However, examination of stay durations across aggregate size suggests that these results depend on the size of the aggregate. Visually, ME flies and VD flies spent greater time in larger aggregates while MD flies and VE flies did not show

noticeable differences across aggregate sizes. Of these, only the difference between VD and VE was significant. It is possible that ME and VD show different behaviours inside an aggregate depending on its size, resulting in different stay durations for different aggregate sizes. It is unclear what these behaviours may be or why they may show such size dependence.

Tendency to stay in an aggregate was seen to change over the course of the assay. It increased steeply with time for deprived flies compared to enriched ones but no such differences were observed between mated and virgin flies. These results match results from previous sections where differences in time profiles due to mating status were different from those seen due to social condition. As discussed in the previous section, this suggests that mating and social condition likely affect different behaviours within the aggregate. The behaviours affected by mating do not seem to change across time while those affected by social condition seem to change over time such that aggregates become longer lived. These results highlight, once again, that mating status and social condition affect aggregation behaviour in distinct ways even though their effects appear to be similar.

3.3.4.4. Non-random spatial relationships between flies:

I found that flies showed selective association with some flies and selective avoidance of others over the course of the assay. Interestingly, I did not observe any effects of social environment on proportions of non-random relationships or on proportions of flies showing such relationships. Social environment only affected the strength of associations, as mated flies had significantly higher AI values compared to virgins. These results indicate that the broad structure of relationships may be retained across treatments with only the investment of time in these relationships being affected by social environment of the flies.

Since relationships were evaluated for the entire duration of the assay, temporal variation in these associations could be overlooked. To account for this, I performed correlation analyses of AI

values across time. However, these results are preliminary at best as few replicates were available for these analyses. All AI values, regardless of significance, were positively correlated across time, suggesting that flies maintained their relationships across time. However, I found fewer significant correlations when temporal correlations were tested using subsets of relationships (i.e. only positive or only negative). These results suggest that flies may prefer or disfavour the same flies across time but strengths of association with these individuals may vary somewhat randomly.

I found that correlation coefficients were often positive when only positive associations were compared across time, and negative when only negative associations were compared. Flies thus seem to reevaluate negative associations across time but continue to maintain positive associations. This pattern was true for all treatments except for VE flies who failed to show any correlations across time. When subsets of associations were analyzed, VE flies either lacked correlations or showed negative correlations. These results suggest that VE flies do not maintain their associations and seem to reevaluate both negative and positive associations. Thus, although largely inconclusive, these analyses hint at the existence of complex spatial relationships between flies.

### *3.4. Social behaviours within aggregates*

#### *3.4.1. Rationale*

Results from the previous section demonstrated that aggregation behaviour of individual flies was affected by social environment. Social environment was seen to primarily affect the tendency of flies to stay in aggregates. This tendency was also seen to be dependent on the size of an aggregate for some but not all social treatments. I also discovered that patterns of association during aggregation were non-random, with some flies aggregating preferentially with each other and some flies aggregating far less frequently. Flies thus seem to make choices about which aggregates to join and which flies to associate with. To better understand *why* flies might be showing such choices, I wanted to understand what events followed these choices i.e. the behaviours that were shown inside an aggregate.

To identify the behaviours that flies exhibited in aggregates, I first characterized all behaviours shown by a fly within the arena. Given the absence of food as well as mates, I mainly observed behaviours such as locomotion and self grooming. I also observed instances of close physical proximity between flies that were stable for short periods of time. These interactions typically involved stereotypical forms of physical contact between the participating flies. As these interactions involve proximity between flies, I expected that they may be relevant during aggregation. Hence, I quantified the number and duration of such interactions across social environments and used these data to understand how interactions were related to aggregation behaviour. Unfortunately, given the logistical constraints imposed by manual quantification of such behaviours, I could sample only a single video from each treatment. My results are thus preliminary and may be improved upon by using automated behavior classifiers.

*3.4.2. Methods*

3.4.2.1. Sampling procedure:

One video was randomly chosen from each treatment and divided into 900 frame (30 seconds) long sampling windows which occurred at intervals of 10 minutes. For each sampling window, five focal flies were chosen randomly and observed for the duration of sampling. All interactions shown by a focal fly during this period were identified and recorded. Start and end frames were noted for each interaction even if the interaction extended beyond the sampling window.

3.4.2.2. Defining an interaction:

Each interaction was defined as a set of stereotyped, non-random movements directed by one fly towards another. Movements were considered non-random if they occurred mainly in the proximity of another fly and were initiated towards the other fly. Instances of grooming (which is stereotypical but occurs in both presence and absence of neighbouring flies) were not quantified. Each interacting fly was defined either as the 'approacher' or as the 'approachee' depending on which fly initiated the interaction. Interactions were broadly classified as those involving association and those involving avoidance based on the extent of physical contact observed. They were further classified based on the relative orientations of the approacher and the approachee (Table 3.4.1.).

Although interactions could be sub-divided into different types, the biological meaning associated with these types is unclear. Hence, these categories were pooled into a single category for the analysis presented here. Avoidance interactions were excluded from the analysis as they were not expected to be relevant for understanding behaviour within the aggregate. Interaction behaviour were quantified in terms of the total amount of time spent interacting by a fly in a sampling window. The average number and duration of these interactions were also measured.

| Behavioural category | Sub-category | Description |
|---|---|---|
| **Avoidance** | Avoidance occurred when at least one individual in an interacting pair moved away shortly before or immediately after physical contact | |
| | ab | **A**voidance shown by <u>b</u>oth approacher and approachee |
| | ae | **A**voidance shown by approache**e** |
| | ar | **A**voidance shown by approache**r** |
| | ac | **A**voidance shown by either approacher or approachee after brief physical **c**ontact |
| **Association or proximity** | Association occurred when interacting flies maintained proximity with any part of their bodies (typically limbs) in contact | |
| | asf | **As**sociation where flies **f**ace each other |
| | asr | **As**sociation with contact between the front of one fly (typically approacher) and the **r**ear of the other |
| | ass | **As**sociation with contact between the front of one fly (typically approacher) and the **s**ide of the other |
| | asp | **As**sociation where the interacting flies lie roughly **p**arallel to each other.<br>(This includes instances where the flies are not exactly parallel but still maintain contact whilst facing away) |
| | ff | Close **f**ront to **f**ront physical contact |
| | fr | Close **f**ront to **r**ear physical contact |
| | fs | Close **f**ront to **s**ide physical contact |
| **Walk past** | Walk pasts were instances where the approacher established physical contact with the approachee and continued to move along the body of the approachee while maintaining contact | |
| | w | **W**alk past with physical contact between limbs of the interacting flies |
| | tw | **W**alk past with the approacher **t**ouching (close contact) the approachee |
| | e | Walk past where the approacher partially **e**ncircled the approachee |

**Table 3.4.1. Different types of physical interactions seen between flies.**

Interactions between flies were broadly classified as avoidance, association and walk pasts. These were further divided based on differences in relative orientations of flies with respect to each other.

3.4.2.3. Relationship between social behaviours and aggregation behaviour:

In Section 3.2., aggregates were defined as long-lived gatherings of two or more flies that are distinct from short-lived, pair-wise interactions. This distinction was made by presuming that the differences in duration and number of flies between these behaviours were indicative of underlying differences in the biological role of these behaviours. However, this distinction may not be valid as aggregates may form simply as a result of several flies interacting in groups. If this were the case, then flies would be expected to 1) aggregate whenever they interacted with each other and 2) spend most of their time interacting when in aggregates. To test these predictions, co-occurrence of aggregation and interaction behaviour was tested. The first prediction was tested by calculating the ratio of time spent interacting in aggregates to the time spent interacting overall. If most interactions occurred in aggregates, then this ratio would be close to one. The latter prediction was tested by calculating the proportion of time spent interacting by a fly when it was part of an aggregate. If aggregates were formed purely as a consequence of interactions between flies, then this value would be expected to be close to one.

Although interactions were the most obvious behaviour seen when flies were in proximity, instances of proximity without any physical contact between flies were also observed. These instances occurred when a fly approached another fly (or a group of flies) and stayed near it without making contact. This proximity was typically seen to be maintained for several frames. Although most actions shown by flies on such occasions were not clearly visible, some of these instances were clearly seen to correspond to preening or self-grooming behaviour. To test if such periods of immobility also contributed to aggregation, the total time spent being immobile while aggregating was quantified for each focal fly. A fly was considered to be immobile in a given frame if it was not interacting and had a velocity less than 1.5 mm/s.

3.4.2.4. Effect of social environment on social behaviours:

To test if associations between social behaviours and aggregation varied across flies exposed to different social environments, proportions of time spent in interactions and in periods of immobility while aggregating were compared across social treatments. The effect of social environment on interaction behaviour, in general, was also tested by comparing the total amount of time spent interacting in a sampling window, as well as the number and duration of these interactions, across treatments.

3.4.2.5. Statistical Analyses:

Since a single video was sampled per treatment, individual flies were used as replicates. Mean values of the above variables were calculated for each focal fly by averaging data collected for that fly across sampling windows. These mean values were analysed using ANOVAs with mating status and social condition as fixed factors. For variables that violated the assumptions of ANOVA, the data were transformed before analysis. Alternative tests were used for data that could not be transformed to conform to the assumptions of ANOVA.

*3.4.3. Results*

3.4.3.1. Effect of social environment on the total time spent in interactions:

Data were transformed to correct for the lack of normality ($y=\sqrt{x}$). As transformation failed to reduce heteroscedasticity in the data, a Welch's test was used to analyse the transformed data. The Games-Howell test was used for post-hoc comparisons. Social environment was seen to have a significant effect on the amount of time spent in interactions by flies ($F=4.299$, $df_1= 3$, $df_2=56.5$, $p=0.0008$). VD flies were seen to spend significantly greater amounts of time interacting than ME and VE flies ($p<0.05$) (Fig. 3.4.1).

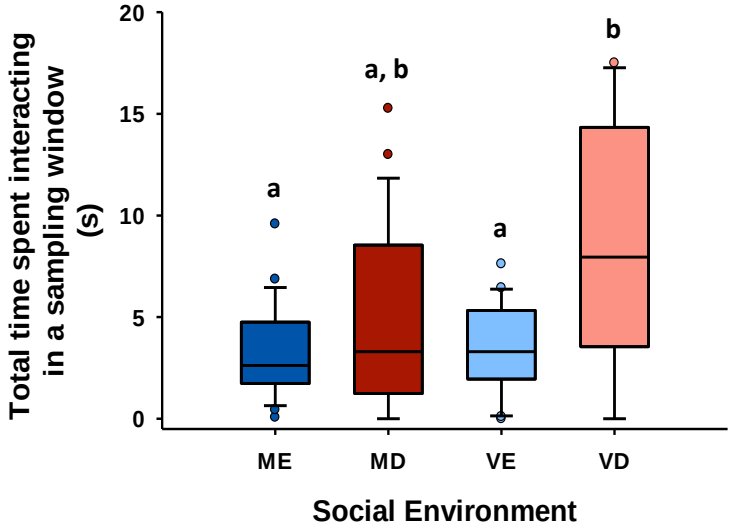



**Figure 3.4.1. Effect of social environment on total time spent interacting.**

Time spent interacting was significantly different across treatments (Welch's ANOVA on transformed data ($y=\sqrt{x}$), p=0.0008). Virgin and socially deprived (VD) flies spent significantly more time interacting compared to both types of enriched flies (Games-Howell tests, $p<0.05$).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other.]

3.4.3.2. Effect of social environment on the total number of interactions:

Socially enriched flies showed significantly more interactions than socially deprived flies ($F=17.973$, $df=1$, $p<0.0001$) (Fig. 3.4.2).

3.4.3.3. Effect of social environment on the duration of an interaction:

As the data did not exhibit normality or homoscedasticity, they were transformed ($y=x^{1/16}$) to reduce heterogeneity of variances and then analysed using a Kruskal-Wallis test. Duration of an interaction was affected significantly by social environment ($H=44.844$, $df=3$, $p<0.0001$). Multiple comparisons revealed that the two treatments involving social deprivation showed significantly longer interactions than the treatments with enriched flies (All comparisons were significant after Bonferroni correction ($p<0.0083$). ME *vs*. MD- $z'=4.655$; ME *vs*. VD, $z'=5.345$; MD *vs*. VE, $z'=4.028$; VE *vs*. VD, $z'=4.67$) (Fig. 3.4.3).

3.4.3.4. Effect of social environment on the proportion of interaction time that occurred in aggregates:

As the data were seen to be non-normal and heteroscedastic, they were transformed ($y=\text{asin}(\sqrt{x})^2$) to reduce heteroscedasticity and then analysed using a Kruskal-Wallis test. Social environment had a significant effect on the proportion of interaction time that occurred in aggregates ($H=34.33$, $df=3$, $p<0.0001$). Multiple comparison tests suggested that this proportion was significantly lower for ME flies compared to MD and VD flies (All comparisons significant after Bonferroni correction ($p<0.0083$). ME vs. MD- $z'=4.507$; ME vs. VD, $z'=4.861$; VE vs. VD, $z'=2.663$) (Fig. 3.4.4). Thus, ME flies seem to interact outside aggregates more frequently.

3.4.3.5. Effect of social environment on the proportion of time spent interacting in an aggregate:

As the mean proportion of time spent interacting in an aggregate was not normally distributed, these values were transformed ($y=\sqrt{x}$) before analysis. A Welch's test was performed as

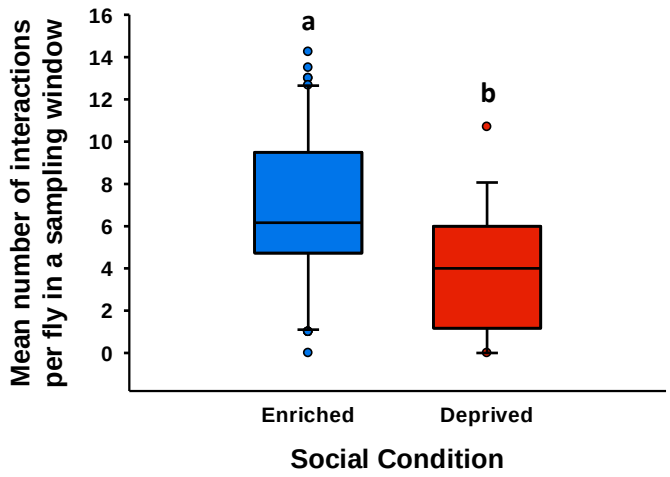



**Figure 3.4.2. Effect of social environment on number of interactions between flies in a sampling window.**

Mean number of interactions shown by a fly in an interval were significantly higher for enriched flies compared to deprived flies (Main effect of social condition ($p<0.0001$)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other.]

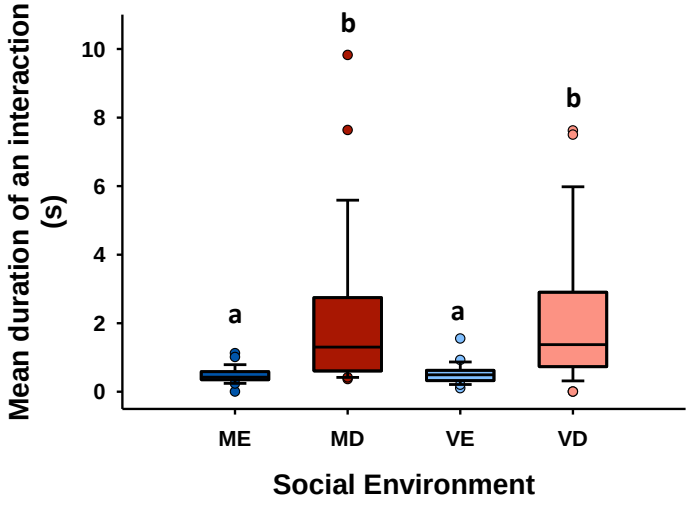



**Figure 3.4.3. Effect of social environment on the duration of an interaction.**

Duration of interactions was significantly different across treatments (Kruskal-Wallis test on transformed data ($y=x^{1/16}$), ($p<0.0001$). Duration was greater for deprived flies compared to enriched flies (z-tests using Bonferroni correction ($p<0.0083$)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other.]

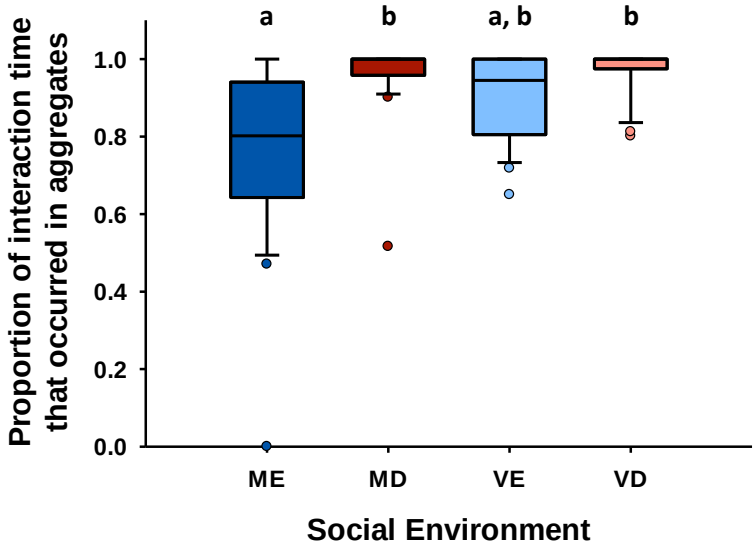



**Figure 3.4.4. Effect of social environment on proportion of interaction time that occurred in aggregates.**

Proportion of time spent interacting that occurred in aggregates was significantly different across treatments (Kruskal-Wallis test on transformed data ($y=asin(\sqrt{x})^2$). Proportion of time spent interacting that occurred in aggregates was significantly lower for mated and socially enriched flies compared to the two types of deprived flies (z-tests using Bonferroni correction ($p<0.0083$)).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other.]

heteroscedasticity could not be reduced using transformation. Social environment did not have a significant effect on the mean proportion of time spent interacting in an aggregate ($F=2.245$, $df_1=3$, $df_2=55.24$, $p=0.09$) (Fig. 3.4.5a).

The effect of social environment on the coefficient of variance for the proportion of time spent interacting in an aggregate was also tested. These data were transformed ($y=\sqrt{x}$) to reduce heteroscedasticity and then analysed using a Kruskal-Wallis test. The effect of social environment on the coefficient of variance did not reach statistical significance ($H=7.496$, $df_1=3$, $df_2=99$, $p=0.057$) (Fig. 3.4.5b).

3.4.3.6. Effect of social environment on the proportion of time spent being immobile in an aggregate:

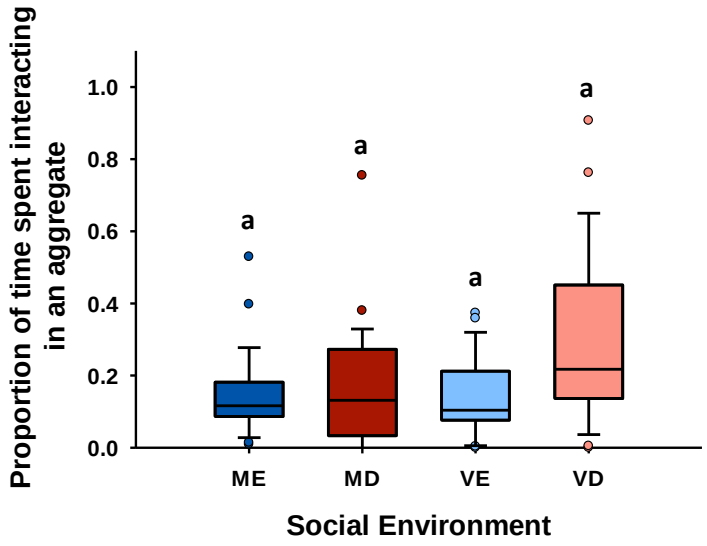
A significant effect of the interaction between mating status and social condition was seen on the proportion of time spent being immobile in aggregates ($F=6.856$, $df=1$, $p=0.01$). Tukey's post-hoc tests showed that this effect was due to the significantly larger values seen for MD flies compared to ME flies ($p=0.004$) (Fig. 3.4.6a).

When the effect of social environment was tested on the coefficient of variance for the proportion of time spent being immobile while aggregating, marginally non-significant effects of social condition ($F=3.937$, $df=1$, $p=0.05$) could be detected (Fig. 3.4.6b).

*3.4.4. Inferences*

Aggregation is a consequence of the decisions made by flies to be near each other which are, in turn, dependent on the behaviours shown by flies towards each other. My results show that flies were in aggregates for most of the time that they spent physically interacting. This lends support to the idea that aggregates may be formed primarily due to interactions between flies. However, when the tendency of flies to interact in aggregates was measured, I found that the mean proportion of

a)




b)

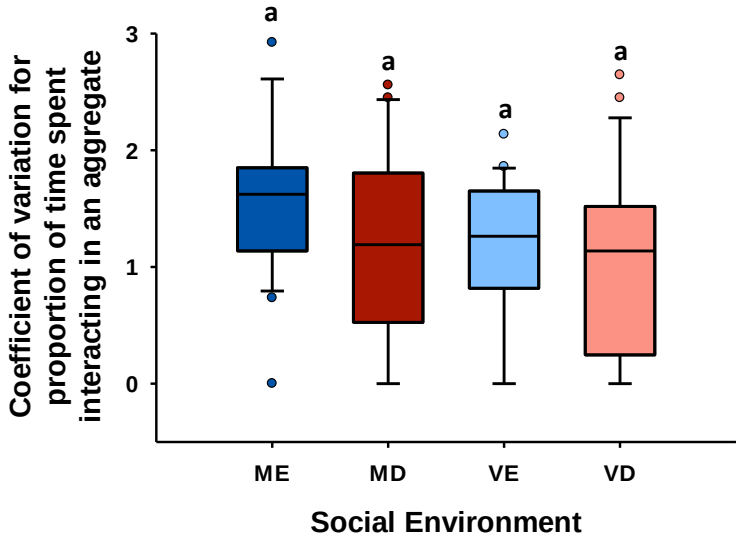
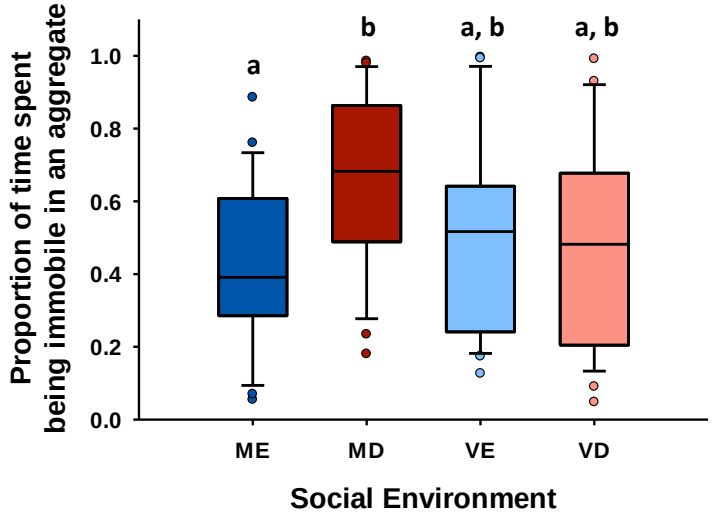



**Figure 3.4.5. Effect of social environment on proportion of time spent interacting when in an aggregate.**

Social environment did not have significant effects on the **a)** mean or **b)** coefficient of variation of the proportion of time spent interacting when in an aggregate (Welch's ANOVA and Kruskal-Walllis test respectively on transformed data ($y=\sqrt{x}$).

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other.]

a)




b)

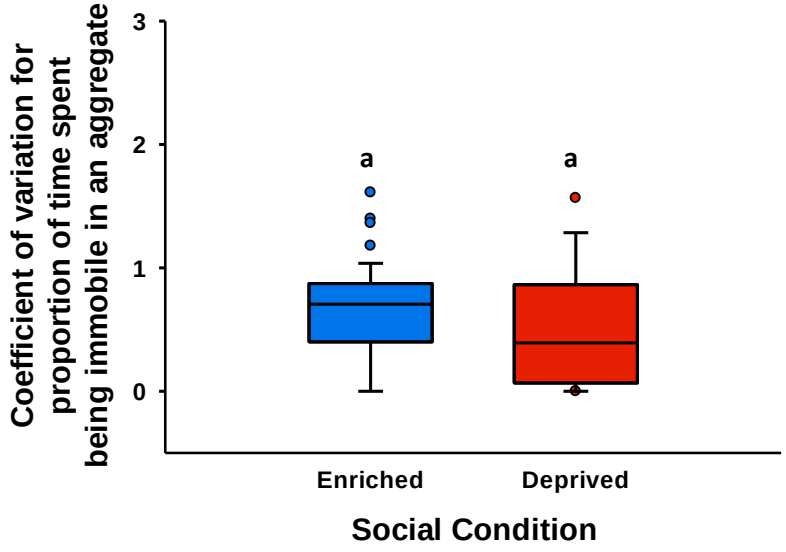



**Figure 3.4.6. Effect of social environment on proportion of time spent being immobile when in aggregates.**

**a)** Proportion of time spent being immobile in aggregates was significantly smaller for mated and socially enriched (ME) flies compared to mated and socially deprived flies (MD) (Interaction between mating status and social condition (p=0.01). Tukey's HSD tests, ME *vs.*MD- p=0.004). **b)** Coefficient of variation for the time spent being immobile in aggregates was greater for enriched flies compared to deprived flies but this difference was marginally non-significant.

[For box-whisker plots, horizontal line denotes median, box edges denote 25$^{th}$ and 75$^{th}$ percentiles and whiskers denote 10$^{th}$ and 90$^{th}$ percentiles. Outliers are shown as filled circles. Treatments denoted by different letters are significantly different from each other.]

time spent interacting inside an aggregate was actually quite small. In addition, coefficients of variation were large which suggests that flies interacted frequently in some aggregates but less so in others. Thus, aggregates are not formed simply as a result of several flies interacting in groups. Instead flies spend a large part of their time in aggregates being immobile. This immobility may result from different behaviours such as self-grooming, which has been previously suggested have some social role (Connolly 1968). The purpose of these behaviours, if any, remains a mystery.

Since I analyzed single videos for each treatment, I lack replication at the appropriate levels and hence, my results for the effects of social environment are, at best, preliminary. Deprived flies showed fewer interactions than enriched flies but these interactions tended to be of longer durations. Since deprived flies encounter conspecifics for the first time within the assay, this increase in the duration of interaction may result from increased socialisation between these flies. Curiously, deprived flies also seemed to show greater inter-individual variation in the duration of interactions compared to enriched flies. It is unclear what contributes to such differences in durations across individuals. Although the total time spent interacting was higher for deprived flies, there was no change in the proportion of time spent interacting within aggregates. This suggests that the behavioural program of flies within the aggregate, i.e. the relative amount of time allotted to each behaviour, remains the same across social environments. A more systematic analysis using more replicate assays may help verify these results.

# DISCUSSION

## *4.1. Describing aggregates in Drosophila*

Although aggregation refers to the behaviour of a group of individuals, it is the outcome of multiple, individual spatial decisions. Aggregation may thus be studied at the level of the individual as well as at the level of the aggregates. While individual-level studies can help us understand how aggregates may form, aggregate-level studies are useful to understand the biological consequences of aggregation, in terms of inter-individual competition or spread of information and/or pathogens. While several studies have studied fly aggregation at the level of individuals (Simon et al., 2012; Foley et al., 2015; Philippe et al., 2016; Anderson et al., 2017), aggregate-level approaches have been used far less frequently (Saltz and Foley, 2011). Additionally, aggregate-level properties such as density and composition, which can determine the extent of local competition as well as the scope for transfer of information or pathogens respectively, have been largely ignored. An important reason for this lack of comprehensive study of aggregate-level properties in fruit flies has been the inability to define an aggregate for *Drosophila.*

In this thesis, I attempted to address this problem by defining aggregates using empirical estimates of proximity between flies. Since this definition depends on the behaviour of flies and not on environmental heterogeneities, as has been used previously (Saltz and Foley, 2011; Philippe et al., 2016), it represents a more biologically meaningful approach to defining social aggregates. However, it is important to note that this approach is not entirely objective, as the choice of the $\varepsilon$ parameter used for clustering relies on human perceptions of proximity and may thus not reflect a fly's perspective. To estimate the perspective of a fly, we would need to use some behavioural marker of aggregation and then identify distances over which this marker is seen to occur.

Preliminary observations suggest that self-grooming behaviour may serve as such a marker. Although I did not quantify self-grooming here, we know from previous literature that flies groom more when assayed in a group compared to when assayed alone (Connolly, 1968). Given this result, we would expect flies to modify their grooming behaviour in response to the presence of other flies near them. Thus, the inter-individual distance at which flies show an increase in the frequency of grooming may be used as the distance at which flies experience proximity. A caveat to this approach is that flies may not necessarily groom when they experience proximity with other flies. To account for this, other potential markers of aggregation may be identified and tested similarly. Given the absence of human subjectivity, this approach may yield more accurate estimates of proximity in flies.

### *4.2. Individual behaviours underlying aggregation*

Observed patterns of aggregation are emergent outcomes of different individual behaviours and interactions between individuals (Foley et al., 2015; Philippe et al., 2016). In addition to identifying these behaviours it is also important to understand the biological context in which a given fly may show these behaviours. For example, Foley et al. (2015) showed that tendencies of male flies to join and stay in aggregates depended on the intensity of aggression among males. Hence, I tested the effect of mating and prior social experience on individual tendencies to aggregate and on the behaviours shown by flies while aggregating.

My results showed that mating primarily affected the size and the duration of aggregates. The differences in aggregate sizes likely result from the slightly higher preference to join larger aggregates shown by mated flies. Similarly, the longer durations can be explained by the tendency of mated flies to remain in aggregates for longer. However, I could not identify any clear differences in the behaviours shown by mated flies within the aggregate which may help explain

these tendencies. Although many of these results are somewhat inconclusive owing to large variation across replicates, the observed trends provide plausible explanations for how mated flies may form aggregates. As discussed in the Introduction, aggregation of mated flies may be linked to gregarious oviposition. It is unclear what these links may be, as flies have been reported to not lay eggs simultaneously (del Solar and Palomino, 1966. Given the effects of the male sex peptide (reviewed in Kubli, 2003) on several aspects of female behaviour, it is possible that mating modifies the tendency of individual flies to aggregate, independent of its effects on oviposition related behaviours. This hypothesis is supported by the observation that mated flies show an increased tendency to aggregate despite the absence of any oviposition sites within the arena. That being said, the overall size of the aggregate was only slightly larger for mated flies compared to virgins. It is thus possible that mated flies may form even larger aggregates if oviposition sites are also provided within the arena. Hence, further experiments in the presence of oviposition sites are essential to verify the relationship between oviposition and aggregation.

Unlike the effects of mating, social condition only affected the abundance and duration of aggregates. The tendency of socially deprived flies to form fewer but longer lived aggregates could be easily explained by their tendency to join fewer aggregates and to stay for longer within them. These individual tendencies also correlated well with the low frequency and longer durations of interactions observed for socially deprived flies. These results thus suggest that interactions may be driving the increases in aggregation seen after deprivation. Since physical interactions are known to mediate information transfer in flies (Battesti et al., 2015), these increases in interaction may involve information transfer between deprived flies, which may, in turn, aid in their socialisation. Curiously, however, flies interacted for only a small proportion of time within the aggregate. Instead, they spent much of their time being stationary within the aggregate. It is possible that these periods of immobility also involve some form of non-physical communication between flies (Kacsoh et al., 2018). To verify if aggregation mediates information transfer, naive flies may be

assayed along with flies trained to avoid specific odours. The ability of the naive flies to avoid these odours may then be correlated with features of the observed aggregates to test if aggregation affects efficiency of information transfer.

### *4.3. Social relationships in flies*

Aggregate composition refers to the identities of the aggregating individuals and, more importantly, the relationships among these individuals. Although such relationships have been largely ignored in case of *Drosophila*, recent studies on social networks in *Drosophila* (Schneider et al., 2012; Pasquaretta et al., 2016) show that patterns of interactions among flies are non-random. These non-random patterns may be reflective of underlying relationships between pairs of flies. My data show that such non-randomness also exists for patterns of association among flies. Flies aggregate preferentially with some individuals and avoid forming aggregates with others. This kind of discriminatory behaviour may indicate the presence of different kinds of relationships among flies.

Social environment had little or no effect on the overall patterns of relationships between individuals. These results concur with results reported by Schneider et al. (2012) who failed to detect any effect of social isolation on the properties of social interaction networks in flies. However, I was able to detect an effect of social environment on the strength and stability of spatial associations. Mated flies showed stronger spatial associations as they tended to be in aggregates with the same individuals more frequently than virgin flies. Such association between females may be important for co-ordinated movement of female *D. melanogaster* that has been seen to occur in the wild (Soto-Yeber et al., 2018). Spatial relationships were also maintained across time, as association values were correlated across the first and the second hours of the assay. Curiously, these correlations were largely absent in case of virgin and socially enriched (VE) flies suggesting

that mating and the need to socialise may play some role in driving these patterns of association. Broadly, these results suggest that flies show an innate, invariant pattern of spatial relationships but specific aspects of these relationships, such as the identity of associates and the extent of association with them, may be modulated by the social environment.

The presence of non-random relationships can influence aggregation behaviour as flies may choose to join or stay in aggregates more frequently if they contain preferred associates and less frequently if they contain disfavoured individuals. Conversely, flies may form relationships with individuals that they encounter most frequently in aggregates. Social relationships may also influence interaction behaviours among flies, as the types of interactions and their consequences (such as social learning) may depend on the relationship between these individuals. At the same time, the relationship may itself depend on the interaction history for a pair of flies. Social relationships are, thus, likely to introduce great complexity in patterns of association and interaction among flies. Further studies are required to build on the preliminary data presented here to fully understand the role of social relationships in the ecology of *Drosophila*.

### *4.4. Conclusions*

Fruit flies are not conventional social organisms as they do not show group living. As a result, inter-individual interactions in *Drosophila* have been understood typically in terms of inter-sexual encounters for mating or, more rarely, as intra-sexual conflict for resources. Yet recent research shows that sociality in flies includes complex networks of social interactions (Schneider et al., 2012, Pasquaretta et al., 2016) as well as different forms of social learning that involve both passive (Mery et al., 2009; Sarin and Dukas, 2009) and active information transfer (Battesti et al., 2012; Kacsoh et al., 2018). Results of this study add to this body of work by showing that fruit flies can evaluate and modify their social environment (i.e. the aggregate) and can also form social

relationships with each other. These observations do not, by any means, demand reclassification of *Drosophila* as a social animal but they do prompt a rethinking of how we study and interpret social behaviours for 'non-social' organisms. There is a tendency to consider complex social behaviour as a trait that is characteristic of more complex organisms who are believed to possess specialized cognitive abilities for sociality. However, we need to reconsider this assumption in the light of emerging evidence regarding the complexity in behaviour that may be achieved by brains of 'simpler' organisms.

## FUTURE STUDIES

Despite an increased interest in social behaviours of fruit flies much of the existing literature on aggregation in *Drosophila* remains patchy. Comparison across these studies is also complicated as they vary in terms of assay environments and the analytical approaches used to quantify aggregation. Most importantly, however, previous studies have tended to study aggregation in isolation from other social behaviours such as courtship and aggression. This approach is problematic for addressing both proximate and ultimate questions as these behaviours are likely to be correlated. Thus, to better understand aggregation behaviour, a systematic and comprehensive study is required to address a variety of proximate and ultimate questions that remain unexplored in case of aggregation behaviour. Some such questions that may be explored in the future are discussed below.

### *5.1. What are the mechanistic underpinnings of aggregation?*

Aggregation behaviour is a good example of a complex behaviour as it involves a sequence of distinct behavioural events. A fly must first choose an aggregate to join following which it may show different behaviours within the aggregate. Each of these constituent behaviours is likely to be influenced by several environmental factors such as resources as well as conspecifics. Consequently, a variety of sensory inputs may be involved in initiation of these behaviours. These inputs may be identified by assaying mutants for different sensory modalities and testing for differences in individual-level behaviours. To identify higher order processing centres that may be involved during aggregation, different transgenic constructs may be used to activate or silence different regions of the brain. Brain regions that give rise to changes in behaviour upon activation or silencing are likely to serve as processing centres for aggregation. As was discussed previously,

aggregation and aggression behaviours may be related. Thus, it would be interesting to test if brain regions known to be involved in aggression also modulate aggregation. Since aggregation is likely to be highly context dependent, different brain regions may be involved in mediating aggregation under different environmental conditions. The neuronal bases of such plasticity may be explored by assaying flies after varying environmental conditions such as the social environment or resource distributions.

### *5.2. Is aggregation behaviour adaptive?*

Resource distributions and social factors are known to be important factors that influence grouping behaviour. These factors often have fitness consequences for the grouping individuals, and it is these consequences that are thought to shape aggregation choices. In case of flies, resource distribution may affect fitness of individuals by determining their access to resources such as food and oviposition sites. Social factors on the other hand may affect fitness by influencing spread of information via social interactions or by affecting access to resources for different individuals via aggressive interactions. To study the effect of these factors on aggregation, different kinds of resource patterns and social environments may be incorporated into the assay setup described in this thesis. Resource patterns may be varied by changing food availability as well as by changing the spatial distribution of food. Social environment may be varied along several axes such as adult density, sex ratio, training status, disease status etc. Of these, adult density and sex ratio may be the most ecologically relevant and may thus be more useful for studying the adaptive value of aggregation. Since resource patterns and social environments are likely to influence each other, a fully factorial design involving these factors may be used.

For each combination of levels of these factors, aggregation behaviour may be recorded along with other fitness related traits such as feeding rates, oviposition rates for females and mating rates for

males. Other behaviours such as aggression, courtship and social interactions may also be quantified to understand how these traits are correlated with aggregation. If aggregation has some adaptive value, then we would expect fitness related traits to differ when individuals form aggregates compared to when they don't. Alternatively, they may vary with properties of the observed aggregates such as aggregate size. Even if direct correlations between fitness traits and aggregation are not observed, the relationship between aggregation and other behavioural traits may be verified. If such behavioural correlations exist, then aggregation may provide indirect fitness benefits to the aggregating individuals.

### *5.3. How may aggregation behaviour evolve?*

Aggregation behaviour is known to vary across commonly used laboratory fly strains (McNeil et al., 2015). This suggests that fly populations possess genetic variation for aggregation. If aggregation tendencies influence individual fitness, then this genetic variation may allow for the evolution of aggregation behaviour via natural selection. Flies exposed to different ecologies would experience distinct selection pressures which may result in distinct evolutionary trajectories for aggregation behaviour. Consequently, these flies may show distinct patterns of correlation between aggregation and other behavioural traits. To understand these patterns of correlation, we would need to identify the local ecologies for these flies as well as their ancestral trait values. Since these may not be accessible for wild populations, a long-term laboratory study involving adaptation of wild caught flies to laboratory conditions may be useful. As a preliminary study, wild caught flies could be reared under lab conditions for several generations and the evolution of aggregation behaviour may be charted. If aggregation does not evolve in response to lab adaptation, then we may infer that ecological factors associated with aggregation either do not differ between the wild and the lab, or are inert enough to not effect changes in behaviour. In either case, such a result

would suggest that aggregation behaviour under experimental conditions may, at least partially, reflect behaviour in the wild. However, if aggregation is seen to evolve under lab conditions then we may infer that aggregation is strongly shaped by ecology of the flies. To identify the factors shaping these patterns, these populations may be sub-divided and each sub-population may be reared under different environments. These environments may vary either in terms of resource patterns or social factors, such as adult density or sex ratio. Evolution of differences, if any, in aggregation behaviour of these populations after several generations may then be traced to these factors. Such long term experiments would allow for control over ancestral states and local ecologies of the evolving populations and thereby facilitate systematic study of the evolution of aggregation behaviour.

## REFERENCES


Alcock, J. (2001). *Animal behavior: An evolutionary approach* (7th ed.). Sunderland, Massachusetts: Sinauer Associates, Inc.

Anderson, B. B., Scott, A., & Dukas, R. (2017). Indirect genetic effects on the sociability of several group members. *Animal Behaviour*, *123*, 101–106. https://doi.org/10.1016/j.anbehav.2016.10.028

Aspi, J., & Hoffmann, A. A. (1998). Female encounter rates and fighting costs of males are associated with lek size in *Drosophila mycetophaga*. *Behavioral Ecology and Sociobiology*, *42*(3), 163–169. https://doi.org/10.1007/s002650050427

Aureli, F., Cords, M., & Van Schaik, C. P. (2002). Conflict resolution following aggression in gregarious animals: A predictive framework. *Animal Behaviour*, *64*(3), 325–343. https://doi.org/10.1006/anbe.2002.3071

Aureli, F., Schaffner, C. M., Boesch, C., Bearder, S. K., Call, J., Chapman, C. A., … van Schaik, C. P. (2008). Fission-Fusion Dynamics: New Research Frameworks. *Current Anthropology*, *49*(4), 627–654. https://doi.org/10.2307/20142694

Bartelt, R. J., Schaner, A. M., & Jackson, L. L. (1985). cis-Vaccenyl acetate as an aggregation pheromone in *Drosophila melanogaster*. *Journal of Chemical Ecology*, *11*(12), 1747–1756. https://doi.org/10.1007/BF01012124

Battesti, M., Moreno, C., Joly, D., & Mery, F. (2012). Spread of social information and dynamics of social transmission within *Drosophila* groups. *Current Biology*, *22*(4), 309–313. https://doi.org/10.1016/j.cub.2011.12.050

Battesti, M., Pasquaretta, C., Moreno, C., Teseo, S., Joly, D., Klensch, E., … Mery, F. (2015). Ecology of information: social transmission dynamics within groups of non-social insects. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801). https://doi.org/10.1098/rspb.2014.2480

Branson, K., Robie, A. A., Bender, J., Perona, P., & Dickinson, M. H. (2009). High-throughput ethomics in large groups of *Drosophila*. *Nature Methods*, *6*(6), 451–457. https://doi.org/10.1038/nmeth.1328

Brenman-Suttner, D. B., Long, S. Q., Kamesan, V., de Belle, J. N., Yost, R. T., Kanippayoor, R. L., & Simon, A. F. (2018). Progeny of old parents have increased social space in *Drosophila melanogaster*. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-21731-0

Burg, E. D., Langan, S. T., & Nash, H. A. (2013). *Drosophila* social clustering is disrupted by anesthetics and in narrow abdomen ion channel mutants. *Genes, Brain and Behavior*, *12*(3), 338–347. https://doi.org/10.1111/gbb.12025

Castells-Nobau, A., Eidhof, I., Fenckova, M., Brenman-Suttner, D. B., Scheffer-de Gooyert, J. M., Christine, S., … Schenck, A. (2019). Conserved regulation of neurodevelopmental processes and behavior by FoxP in *Drosophila*. In *Plos One* (Vol. 14). https://doi.org/10.1371/journal.pone.0211652

Chase, I. D. (1982). Dynamics of hierarchy formation: The sequential development of dominance relationships. *Behaviour*, *80*, 218–239. https://doi.org/10.1163/156853982X00364

Connolly, K. (1968). The social facilitation of preening behaviour in *Drosophila melanogaster*. *Animal Behaviour*, *16*(2–3), 385–391.

Danchin, E., Nöbel, S., Pocheville, A., Dagaeff, A.C., Demay, L., Alphand, M., … Isabel, G. (2018). Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions. *Science*, *362*(6418), 1025–1030. https://doi.org/10.1126/science.aat1590

del Solar, E., & Palomino, H. (1966). Choice of oviposition in *Drosophila melanogaster*. *The American Naturalist*, *100*(911), 127–133. https://doi.org/10.1086/282406

Dillon, M. E., Wang, G., Garrity, P. A., & Huey, R. B. (2009). Thermal preference in *Drosophila*. *Journal of Thermal Biology*, *34*(3), 109–119. https://doi.org/10.1016/j.jtherbio.2008.11.007

Dombrovski, M., Poussard, L., Moalem, K., Kmecova, L., Hogan, N., Schott, E., … Condron, B. (2017). Cooperative behaviour emerges among *Drosophila* larvae. *Current Biology*, *27*(18), 2821–2826. https://doi.org/.1037//0033-2909.I26.1.78

Droney, D. C. (1994). Tests of hypotheses for lek formation in a Hawaiian *Drosophila*. *Animal Behaviour*, Vol. 47, pp. 351–361. https://doi.org/10.1006/anbe.1994.1048

Fernández-Juricic, E., & Kacelnik, A. (2004). Information transfer and gain in flocks: The effects of quality and quantity of social information at different neighbour distances. *Behavioral Ecology and Sociobiology*, *55*(5), 502–511. https://doi.org/10.1007/s00265-003-0698-9

Fernandez, R. W., Akinleye, A. A., Nurilov, M., Feliciano, O., Lollar, M., Aijuri, R. R., … Simon, A. F. (2017). Modulation of social space by dopamine in *Drosophila melanogaster*, but no effect on the avoidance of the *Drosophila* stress odorant. *Biology Letters*, *13*(8), 20170369. https://doi.org/10.1098/rsbl.2017.0369

Foley, B. R., Saltz, J. B., Nuzhdin, S. V, & Marjoram, P. (2015). A novel, Bayesian approach to social structure uncovers cryptic regulation of group dynamics. *The American Naturalist*, *185*(6), 797–808. https://doi.org/10.1086/681084.A

Foucaud, J., Philippe, A.-S., Moreno, C., & Mery, F. (2013). A genetic polymorphism affecting reliance on personal versus public information in a spatial learning task in *Drosophila melanogaster*. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1760). https://doi.org/10.1098/rspb.2013.0588

Ganguly-Fitzgerald, I., Donlea, J., & Shaw, P. J. (2006). Waking experience affects sleep need in *Drosophila*. *Science*, *313*(5794), 1775–1781.

Griffith, L. C. (2013). Neuromodulatory control of sleep in *Drosophila melanogaster*: Integration of competing and complementary behaviors. *Current Opinion in Neurobiology*, *23*(5), 819–823. https://doi.org/10.1158/0008-5472.CAN-10-4002.BONE

Hamilton, W. D. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, *31*(2), 295–311. https://doi.org/10.1016/0022-5193(71)90189-5

Hodosh, R. J., Ringo, J. M., & McAndrew, F. T. (1979). Density and lek displays in *Drosophila grimshawi* (Diptera: *Drosophilidae*). *Zeitschrift Für Tierpsychologie*, *49*(2), 164–172.

Hoffmann, A. A., & Cacoyianni, Z. (1989). Selection for territoriality in *Drosophila melanogaster*: correlated responses in mating success and other fitness components. *Animal Behaviour*, *38*(1), 23–34. https://doi.org/10.1016/S0003-3472(89)80062-4

Hoffmann, A. A., & Cacoyianni, Z. (1990). Territoriality in *Drosophila melanogaster* as a conditional strategy. *Animal Behaviour*, *40*(3), 526–537. https://doi.org/10.1016/S0003-3472(05)80533-0

Hoopfer, E. D. (2016). Neural control of aggression in *Drosophila*. *Current Opinion in Neurobiology*, *38*, 109–118. https://doi.org/10.1016/j.conb.2016.04.007

Isaac, R. E., Li, C., Leedale, A. E., & Shirras, A. D. (2010). *Drosophila* male sex peptide inhibits siesta sleep and promotes locomotor activity in the post-mated female. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1678), 65–70. https://doi.org/10.1098/rspb.2009.1236

Jacobs, M. E. (1960). Influence of light on mating of *Drosophila melanogaster*. *Ecology*, *41*(1), 182–188.

Kacsoh, B. Z., Bozler, J., & Bosco, G. (2018). *Drosophila* species learn dialects through communal living. In *PLoS Genetics* (Vol. 14). https://doi.org/10.1371/journal.pgen.1007430

Kaur, K., Simon, A. F., Chauhan, V., & Chauhan, A. (2015). Effect of bisphenol A on *Drosophila melanogaster* behavior - A new model for the studies on neurodevelopmental disorders. *Behavioural Brain Research*, *284*, 77–84. https://doi.org/10.1016/j.bbr.2015.02.001

Keiser, C. N., Rudolf, V. H. W., Sartain, E., Every, E. R., & Saltz, J. B. (2018). Social context alters host behavior and infection risk. *Behavioral Ecology*, *29*(4), 869–875. https://doi.org/10.1093/beheco/ary060

Kohn, N. R., Reaume, C. J., Moreno, C., Burns, J. G., Sokolowski, M. B., & Mery, F. (2013). Social environment influences performance in a cognitive task in natural variants of the foraging gene. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0081272

Kubli, E. (2003). Sex-peptides : seminal peptides of the *Drosophila* male. *Cellular and Molecular Life Sciences CMLS*, *60*(8), 1689–1704. https://doi.org/10.1007/s00018-003-3052-5

Lefranc, A., Jeune, B., Thomas-Orillard, M., & Danchin, É. (2001). Non-independence of individuals in a population of *Drosophila melanogaster*: Effects on spatial distribution and dispersal. *Comptes Rendus de l'Academie Des Sciences-Serie III-Sciences de La Vie*, *324*(3), 219–227. https://doi.org/10.1016/S0764-4469(00)01297-X

Lihoreau, M., Clarke, I. M., Buhl, J., Sumpter, D. J. T., & Simpson, S. J. (2016). Collective selection of food patches in *Drosophila*. *The Journal of Experimental Biology*, *219*(5), 668–675. https://doi.org/10.1242/jeb.127431

Lind, J., & Lindenfors, P. (2010). The number of cultural traits is correlated with female group size but not with male group size in chimpanzee communities. *PLoS ONE*, *5*(3). https://doi.org/10.1371/journal.pone.0009241

McDonald, J.H. (2014). *Handbook of Biological Statistics* (3rd ed.). Baltimore, Maryland: Sparky House Publishing.

McNeil, A. R., Jolley, S. N., Akinleye, A. A., Nurilov, M., Rouzyi, Z., Milunovich, A. J., … Simon, A. F. (2015). Conditions Affecting Social Space in *Drosophila melanogaster*. *Journal of Visualized Experiments*, *e53242*(105), 1–10. https://doi.org/10.3791/53242

Mery, F., Varela, S. A. M., Danchin, É., Blanchet, S., Parejo, D., Coolen, I., & Wagner, R. H. (2009). Public versus personal information for mate copying in an invertebrate. *Current Biology*, *19*(9), 730–734. https://doi.org/10.1016/j.cub.2009.02.064

Mogilner, A., Edelstein-Keshet, L., Bent, L., & Spiros, A. (2003). Mutual interactions, potentials, and individual distance in a social aggregation. *Journal of Mathematical Biology*, *47*(4), 353–389.

Navarro, J., & del Solar, E. (1975). Pattern of spatial distribution in *Drosophila melanogaster*. *Behavior Genetics*, *5*(1), 9–16. https://doi.org/10.1007/BF01067577

Pasquaretta, C., Battesti, M., Klenschi, E., Bousquet, C. A. H., Sueur, C., & Mery, F. (2016). How social network structure affects decision-making in *Drosophila melanogaster*. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1826). https://doi.org/10.1098/rspb.2015.2954

Patterson, J. E. H., & Ruckstuhl, K. E. (2013). Parasite infection and host group size: A meta-analytical review. *Parasitology*, *140*(7), 803–813. https://doi.org/10.1017/S0031182012002259

Philippe, A.-S., Jeanson, R., Pasquaretta, C., Rebaudo, F., Sueur, C., & Mery, F. (2016). Genetic variation in aggregation behaviour and interacting phenotypes in *Drosophila*. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1827). https://doi.org/10.1098/rspb.2015.2967

Ramdya, P., Lichocki, P., Cruchet, S., Frisch, L., Tse, W., Floreano, D., & Benton, R. (2015). Mechanosensory interactions drive collective behaviour in *Drosophila*. *Nature*, *519*(7542), 233–236. https://doi.org/10.1038/nature14024

Rohlfs, M., & Hoffmeister, T. S. (2004). Spatial aggregation across ephemeral resource patches in insect communities: An adaptive response to natural enemies? *Oecologia*, *140*(4), 654–661. https://doi.org/10.1007/s00442-004-1629-9

Ruiz-Dubreuil, G., Burnet, B., & Connolly, K. (1994). Behavioural correlates of selection for oviposition by *Drosophila melanogaster* females in a patchy environment. *Heredity*, *73*(1), 103–110. https://doi.org/10.1038/hdy.1994.105

Ruiz-Dubreuil, G., & del Solar, E. (1986). Effect of selection on oviposition site preference in *Drosophila melanogaster*. *Australian Journal of Biological Sciences*, *39*(2), 155–160. Retrieved from http://www.publish.csiro.au/?paper=BI9860155

Sah, P., Mann, J., & Bansal, S. (2018). Disease implications of animal social network structure: A synthesis across social systems. *Journal of Animal Ecology*, *87*(3), 546–558. https://doi.org/10.1111/1365-2656.12786

Saltz, J. B. (2011). Natural genetic variation in social environment choice: Context-dependent gene-environment correlation in *Drosophila melanogaster*. *Evolution*, *65*(8), 2325–2334. https://doi.org/10.1111/j.1558-5646.2011.01295.x

Saltz, J. B., & Foley, B. R. (2011). Natural genetic variation in social niche construction: Social effects of aggression drive disruptive sexual selection in *Drosophila melanogaster*. *The American Naturalist*, *177*(5), 645–654. https://doi.org/10.1086/659631

Sarin, S., & Dukas, R. (2009). Social learning about egg-laying substrates in fruitflies. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1677), 4323–4328. https://doi.org/10.1098/rspb.2009.1294

Schneider, J., Dickinson, M. H., & Levine, J. D. (2012). Social structures depend on innate determinants and chemosensory processing in *Drosophila*. *Proceedings of the National Academy of Sciences*, *109*(Supplement 2), 17174–17179. https://doi.org/10.1073/pnas.1121252109

Sheeba, V. (2008). The *Drosophila melanogaster* circadian pacemaker circuit. *Journal of Genetics*, *87*(5), 485–493. https://doi.org/10.1007/s12041-008-0071-x

Sheeba, V., Sharma, V. K., Chandrashekaran, M. K., & Joshi, A. (1999). Persistence of eclosion rhythm in *Drosophila melanogaster* after 600 generations in an aperiodic environment. *Naturwissenschaften*, *86*(9), 448–449. https://doi.org/10.1007/s001140050651

Shelly, T. E. (1987). Lek behaviour of a Hawaiian *Drosophila*: male spacing, aggression and female visitation. *Animal Behaviour*, *35*(5), 1394–1404. https://doi.org/10.1016/S0003-3472(87)80012-X

Shultzaberger, R. K., Johnson, S. J., Wagner, J., Ha, K., Markow, T. A., & Greenspan, R. J. (2019). Conservation of the behavioral and transcriptional response to social experience among *Drosophilids*. *Genes, Brain and Behavior*, *18*(1). https://doi.org/10.1111/gbb.12487

Simon, A. F., Chou, M.-T., Salazar, E. D., Nicholson, T., Saini, N., Metchev, S., & Krantz, D. E. (2012). A simple assay to study social behavior in *Drosophila* : measurement of social space within a group. *Genes, Brain and Behavior*, *11*(2), 243–252. https://doi.org/10.1111/j.1601-183x.2011.00740.x

Soto-Yéber, L., Soto-Ortiz, J., Godoy, P., & Godoy-Herrera, R. (2018). The behavior of adult *Drosophila* in the wild. *PLoS ONE*, *13*(12). https://doi.org/10.1371/journal.pone.0209917

Taylor, C. E., & Kekic, V. (1988). Sexual selection in a natural population of *Drosophila melanogaster*. *Evolution*, *42*(1), 197–199.

Terrien, J., Perret, M., & Aujard, F. (2011). Behavioral thermoregulation in mammals: a review. *Frontiers in Bioscience*, *16*(4), 1428–1444. https://doi.org/10.2741/3797

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Fuer Tierpsychologie*, *20*(4), 410–433. https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

Tinette, S., Zhang, L., & Robichon, A. (2004). Cooperation between *Drosophila* flies in searching behavior. *Genes, Brain and Behavior*, *3*(1), 39–50. https://doi.org/10.1046/j.1601-183x.2003.0046.x

Ueda, A., & Kidokoro, Y. (2002). Aggressive behaviours of female *Drosophila melanogaster* are influenced by their social experience and food resources. *Physiological Entomology*, *27*(1), 21–28. https://doi.org/10.1046/j.1365-3032.2002.00262.x

Valente, D., Golani, I., & Mitra, P. P. (2007). Analysis of the trajectory of *Drosophila melanogaster* in a circular open field arena. *PLoS ONE*, *2*(10). https://doi.org/10.1371/journal.pone.0001083

VanderWaal, K. L., & Ezenwa, V. O. (2016). Heterogeneity in pathogen transmission: mechanisms and methodology. *Functional Ecology*, *30*(10), 1606–1622. https://doi.org/10.1111/1365-2435.12645

Wertheim, B., Allemand, R., Vet, L. E. M., & Dicke, M. (2006). Effects of aggregation pheromone on individual behaviour and food web interactions: a field study on *Drosophila*. *Ecological Entomology*, *31*(3), 216–226. https://doi.org/10.1111/j.1365-2311.2006.00757.x

Willis, C. K. R., & Brigham, R. M. (2007). Social thermoregulation exerts more influence than microclimate on forest roost preferences by a cavity-dwelling bat. *Behavioral Ecology and Sociobiology*, *62*(1), 97–108. https://doi.org/10.1007/s00265-007-0442-y

Wise, A., Tenezaca, L., Fernandez, R. W., Schatoff, E., Flores, J., Ueda, A., … Venkatesh, T. (2015). *Drosophila* mutants of the autism candidate gene neurobeachin (rugose) exhibit neuro-developmental disorders, aberrant synaptic properties, altered locomotion, and impaired adult social behavior and activity patterns. *Journal of Neurogenetics*, *29*(2–3), 135–143. https://doi.org/10.3109/01677063.2015.1064916.Drosophila

Xu, P., Atkinson, R., Jones, D. N. M., & Smith, D. P. (2005). *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron*, *45*(2), 193–200. https://doi.org/10.1016/j.neuron.2004.12.031

Yamamoto, D., & Koganezawa, M. (2013). Genes and circuits of courtship behaviour in *Drosophila* males. *Nature Reviews Neuroscience*, *14*(10), 681–692. https://doi.org/10.1038/nrn3567

Yarpiz. (2015). Implementation of DBSCAN Clustering in MATLAB. Retrieved from Yarpiz website: http://yarpiz.com/255/ypml110-dbscan-clustering

Young, L. J., & Young, J. H. (1998). *Statistical Ecology: a Population Perspective*. Boston, USA: Kluwer Academic Publishers.