The role of *Numb* in honeybee ovary activation

Shannon Taylor

A report submitted in partial fulfilment of the Degree of Bachelor of Biomedical Science with Honours.

Genetics, University of Otago, Dunedin, Aotearoa New Zealand

October 2018
Abstract

Environmental signals in utero are thought to induce plastic changes to metabolism and gene regulation, causing diseases such as Type 2 Diabetes. For example, fetal undernutrition is associated with obesity, hyperglycemia, and high circulating lipids. These multifactorial responses are challenging to study in humans, so we need a better model of phenotypic plasticity than humans to understand the fundamental mechanisms underpinning these diseases. Honeybees are an ideal model for phenotypic plasticity, as they have a manipulable response to a defined stimulus, queen mandibular pheromone (QMP): when QMP exposure is lost, worker honeybees become fertile. One molecular mechanism underpinning this response is Notch signalling: when Notch cell signalling is active, the worker ovary is repressed, while a loss of Notch signalling causes ovary activation. Notch signalling is active in the honeybee germarium, the region of the ovary where oocytes are specified and produced, but its function there is unknown. Changes in expression of Numb, a known inhibitor of Notch, coincides spatially and temporally with Notch activation. Therefore Numb is hypothesised to regulate Notch signalling.

To investigate the regulation of Numb in honeybee ovary activation, I predicted a cis-regulatory element within the Numb gene. This drove expression in the gut of Drosophila, and may be QMP-responsive, suggesting Notch signalling may mediate the Drosophila response to QMP in the gut. In addition, to investigate the role of Notch signalling in the honeybee ovary, I performed hybridization chain reaction (HCR) to visualize expression of Numb and the germ-cell marker vasa. HCR in this context was very successful, revealing a new expression domain of vasa, and when quantified, gave results similar to other studies. HCR is, therefore, a powerful
method to investigate and quantify gene expression in an unbiased manner. I found that *Numb* is expressed in a broad domain in the anterior germarium, instead of localizing to a specific cell type. Notch signalling may, therefore, prevent the cell clusters in the ovary from dividing and differentiating into more mature oocytes. This hypothesis is supported by data in the wasp *Nasonia vitripennis*, where more information is available about the structure of the germarium: Notch protein expression correlates spatially with the presence of undifferentiated, self-renewing cell clusters, though the Notch modifier *fringe* is expressed throughout the whole germarium. Therefore, the role of Notch signalling in the honeybee (and *Nasonia*) germarium may be to hold cell clusters in an undifferentiated state, preventing them from becoming mature oocytes, thus repressing worker bee reproduction.
Acknowledgements

Firstly, a colossal thank you to my supervisor Prof. Peter Dearden, for encouragement, advice, and guidance. Thank you for all the opportunities you’ve given me, and for reminding me that my Honours should be my top (science) priority.

Thank you also to Mackenzie Lovegrove, for teaching me wet-lab skills, and frequent illuminating conversations. And to Dr Tom Harrop, for teaching me how to do bioinformatics, and about reproducible research.

Thank you too, to Dr Elizabeth Duncan and Dr Meagan Leask, for generously sharing unpublished data, and to Prof. Marc Halfon, for assistance getting SCRMshaw running.

To the rest of the Laboratory for Evolution and Development- thanks for making this such an enjoyable year!

Thank you to my family, for your encouragement, and for proof-reading this.

This thesis was written using the template of Pollard et al. (2016).
# Table of Contents

Abstract

Acknowledgements

List of Figures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iii
List of Tables . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iii

Abbreviations

1 Introduction

1.1 Environmental influences on metabolic disease . . . . . . . . . . . . . 2
1.2 Developmental Origins of Metabolic diseases . . . . . . . . . . . . . . 3
1.3 Honeybees Ovary Activation . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.4 Connections between ovary activation and human disease . . . . . . 6
1.5 Notch signalling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
1.6 Diversity in insect ovary ultrastructure . . . . . . . . . . . . . . . . . . . 9
1.7 Potential roles of Notch in the honeybee . . . . . . . . . . . . . . . . . 11
1.8 Using Hybridization Chain Reaction (HCR) to investigate gene ex-

pression . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
1.9 Current work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15

2 Methods

2.1 Dissection of ovary and gut tissue . . . . . . . . . . . . . . . . . . . . . 16
2.2 Hybridization Chain Reaction (HCR) . . . . . . . . . . . . . . . . . . 16
2.3 in situ hybridization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
2.4 Immunohistochemistry . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
2.5 Nasonia RNAi . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
2.6 Cloning . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
  2.6.1 Restriction Digests . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
  2.6.2 Gel purification . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
  2.6.3 Ligation and Transformation . . . . . . . . . . . . . . . . . . . . 20
  2.6.4 Colony PCR . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
  2.6.5 Plasmid purification and sequencing . . . . . . . . . . . . . . . 21
  2.6.6 Transformation of Drosophila . . . . . . . . . . . . . . . . . . 22
  2.6.7 Drosophila embryo fixation . . . . . . . . . . . . . . . . . . . . 22
  2.6.8 Staining for LacZ . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
2.7 Bioinformatics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
  2.7.1 CLARE . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
  2.7.2 ClusterDraw and ClusterBuster . . . . . . . . . . . . . . . . . . 23
List of Figures

1.1 Schematic of the Notch signalling pathway ........................................... 8
1.2 Phylogeny and diversity of insect ovary structure ................................. 10
1.3 Diagram of HCR v3.0 technique .......................................................... 14
3.1 Validation of transcription factors predicted by CLARE ..................... 29
3.2 ClusterDraw prediction of motif binding densities ............................ 31
3.3 Predicted CREs mapped to Am-Numb ................................................. 32
3.4 Expression of the CRE in various Drosophila tissues ......................... 34
4.1 Positive and negative controls for the HCR experiment .................... 43
4.2 Numb and vasa expression in the Queen ovary ................................ 45
4.3 Diagram of each step of the image analysis pipeline ......................... 47
4.4 Quantification of HCR data ............................................................... 48
4.5 Immunohistochemistry against pH3 in various insect species ............ 51
4.6 Models and data for cell division in the Nasonia and Drosophila germarium 52
4.7 Role of components of the Notch signalling pathway in Nasonia ovaries . 57
6.1 pH-Pelican vector .............................................................................. 73
6.2 pUC57 vector .................................................................................... 73

List of Tables

2.1 PCR primers ....................................................................................... 21
3.1 Regulatory motifs predicted by CLARE .............................................. 27
3.2 Proportion of CRE-positive and negative flies .................................... 33
3.3 Expression of LacZ in the guts of QMP exposed and unexposed Drosophila . 34
4.1 Quantification of cell division in the Nasonia and Drosophila ovary ........ 53
4.2 Results of Pearson’s χ² statistical test on aggregated data ................ 54
6.1 HCR reagents list .............................................................................. 69
6.2 in situ hybridization reagents list ..................................................... 70
6.3 Immunohistochemistry reagents list ................................................ 70
6.4 Cloning reagents list ........................................................................ 71
6.5 LacZ staining reagents list .............................................................. 71
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Br-C</td>
<td>Broad Complex</td>
</tr>
<tr>
<td>CLARE</td>
<td>Cracking the Language of Regulatory Elements</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular Disease</td>
</tr>
<tr>
<td>DAPI</td>
<td>4 ,6-Diamidino-2-Phenylindole, dihydrochloride</td>
</tr>
<tr>
<td>DAPT</td>
<td>N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester</td>
</tr>
<tr>
<td>dHCR</td>
<td>digital Hybridization Chain Reaction</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DOHaD</td>
<td>Developmental Origins of Health and Disease</td>
</tr>
<tr>
<td>E(spl)-C</td>
<td>Enhancer of Split Complex</td>
</tr>
<tr>
<td>HCR</td>
<td>Hybridization Chain Reaction</td>
</tr>
<tr>
<td>ISH</td>
<td>in situ hybridization</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>NICD</td>
<td>Notch Intracellular Domain</td>
</tr>
<tr>
<td>p-value</td>
<td>probability value</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffer Saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>pH3</td>
<td>Phospho Histone H3</td>
</tr>
<tr>
<td>QMP</td>
<td>Queen Mandibular Pheromone</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNAseq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristic</td>
</tr>
<tr>
<td>Su(H)</td>
<td>Suppressor of Hairless</td>
</tr>
<tr>
<td>T2D</td>
<td>Type 2 Diabetes</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription Factor</td>
</tr>
</tbody>
</table>
We’re all capable of the most incredible change. We can evolve while still staying true to who we are.

*The Doctor*

One global health challenge in the 21st century is the rise of metabolic diseases such as Type 2 Diabetes (T2D), cardiovascular disease (CVD), and metabolic syndrome. In developed nations such as Aotearoa New Zealand, Australia, and the United States, obesity rates are already high and causing an endemic of metabolic diseases. In addition, Maori find the current biomedical framework unhelpful, so we need more inclusive ways of understanding these diseases (Bell et al. 2017). Therefore, a better understanding of the mechanisms, and social factors, underpinning these diseases is of critical importance.

Phenotypic plasticity is the ability of organisms to respond to the environment; plastic responses are thought to underpin some metabolic diseases. According to the Developmental Origins of Health and Disease (DOHaD) hypothesis, fetuses may ‘predict’ the environment they will be exposed to postnatally, based on prenatal stimuli (Gluckman 2004). This may explain observations linking birth weight and maternal under-nutrition to cardiovascular and metabolic disease later in life (Roseboom et al. 2011; Barker & Osmond 1986). However, studying such complex effects in humans is challenging. Instead, in this project, I investigated mechanisms of phenotypic plasticity in honeybee ovary activation: a well defined, easy to manipulate plastic response. This can give insight into the mechanisms organisms use to achieve phenotypic plasticity, that transfer to human disease.
1.1 Environmental influences on metabolic disease

Dramatic lifestyle changes, increasing affluence, and the growing availability of fatty, high energy food are thought to underly the increasing prevalence of T2D, CVD, and obesity (Drewnowski & Popkin 2009). Biomedically, overnutrition causes obesity and increasing insulin resistance, leading to T2D, as well as hypertrophy and hyperplasia of adipose tissue. This causes hypoxia, necrosis, and inflammation, contributing to CVD and cancer risk (O’Neill & O’Driscoll 2015). In New Zealand, one-third of Pākehā and nearly half of Māori are obese (Ministry of Health 2015a). Māori also have higher rates of CVD hospitalisation and mortality than non-Māori, and higher rates of diabetes and diabetes-related complications (Ministry of Health 2015b). Therefore, understanding the causes and consequences of the metabolic syndrome, and related diseases, is important to reduce heath inequality in New Zealand.

The rise of obesity and diabetes in the Pacific provides a case study for understanding the environmental causes of metabolic diseases. In Nauru, a rapid rise in the incidence of T2D coincided with a shift in diet towards importing fatty, nutrient-dense foods (Diamond 1992). Therefore, changes in diet likely underpin the rise of T2D in Nauru (McLennan & Ulijaszek 2015a). This raises the question of why Nauru was so dramatically impacted by these lifestyle changes, while Western countries were not. One explanation is that “thrifty genes” promoting energy storage (as fat) have been more intensely selected for by famine in Nauru than in Western countries (Diamond 1992; Dowse et al. 1991; Neel 1962). However, there is little empirical evidence of increased famine in Nauru, and complex food preservation techniques that exist in the island would buffer the impacts of famine (McLennan & Ulijaszek 2015a). Instead, changes due to colonialization may have caused the increased incidence of T2D in Nauru (McLennan & Ulijaszek 2015a). Colonizers brought more
intensive farming practices, resulting in a population expansion, and increased food importation. The social structure of Nauru also shifted from one reliant on connections between islanders, to a more individualistic structure. Such sociological change should be considered alongside biomedical knowledge when researching the aetiology of metabolic diseases in Nauru and other Pacific Islands (McLennan & Ulijaszek 2015a; McLennan & Ulijaszek 2015b; Gosling et al. 2015). This impact of colonialism and social change is echoed in research on Māori understandings of obesity (Bell et al. 2017). Māori see incorporating the Indigenous worldview, and social interconnectedness, as health-promoting, while colonialism and social trauma diminish health (Bell et al. 2017). Therefore, in New Zealand, to properly promote Māori health, we need to consider environmental effects alongside biomedical causes of metabolic disease. One theory describing the role of the environment in metabolic disease is the DOHaD hypothesis.

1.2 Developmental Origins of Metabolic diseases

The DOHaD hypothesis arose to explain the observation that low birthweight correlates with ischaemic heart disease, and predicts that the maternal environment can influence fetal development and cause disease (Barker & Osmond 1986; Wadhwa et al. 2009). There are two key interpretations of DOHaD. According to the “thrifty phenotype” hypothesis, adult responses to developmental events are responses to noxious stimuli. For example, poor maternal and early life nutrition may damage the pancreas and therefore glucose homeostasis (Hales & Barker 1992). Another theory is the “predictive adaptive response” model: organisms use maternal cues during development to predict the type of environment they will be exposed to as adults (Gluckman 2004). For example, a fetus may receive limited nutrition in utero and make adaptations such as slowing its metabolism and storing energy as fat, because it predicts that it will be exposed to limited nutrients after birth. These adaptations become maladaptive if the maternal environment does not match the
adult one (Bateson & Gluckman 2011). Even invertebrates such as nematode worms arrest development in response to both starvation and crowding; these conditions cause distinct changes in gene expression late in life that are thought to be adaptive to their early life conditions (Ow et al. 2018).

The DOHaD hypothesis is a compelling idea, with important implications. If the gestational environment of an individual is important to their future health, then the health of mothers becomes crucially important to the future population. Emphasis on the environmental impacts on health provides an opportunity to consider population-wide interventions that may drastically alter health outcomes. Māori also recognise the importance of holistic health, and the role of the environment in obesity (Bell et al. 2017). Combining these worldviews suggests a way forwards: focus on the health of communities, rather than merely the individual. This addresses Māori concerns about the way obesity is treated in New Zealand (Bell et al. 2017), and proactively addresses the predictions of DOHaD. Focussing on community health allows us to ensure that the next generation will have the best start to life, both biologically and socially (Ulijaszek et al. 2016).

The ability of organisms to respond to their environment is called phenotypic plasticity. This is very highly conserved in animals: even sea sponges are able to alter their morphology and metabolism in response to depth and season (Morley et al. 2016; López-Legentil et al. 2010). Plasticity is also thought to provide the variation necessary for evolution, so is of fundamental importance (West-Eberhard 1989). The foetal responses to the environment predicted by DOHaD are plastic responses. However, plasticity is difficult to study in humans, particularly in the context of T2D, CVD, and DOHaD. Though the phenotype of these diseases is clear, a variety of environmental and genetic factors contribute to their pathogenesis (Sloboda et al. 2009), and distinguishing developmental effects from these other causes is challenging. Model systems are useful for understanding complex biological systems: studies of Drosophila have had a huge impact on how we understand developmental
genetics. A good model system for understanding developmental plasticity would have consistent, easy to activate, and predictable plastic responses to the environment. Honeybees are such a model. Firstly, clearly definable alternative phenotypes exist, so it is easy to identify whether a given stimulus or mechanism is associated with a particular state (Cridge et al. 2017). Secondly, the stimuli controlling caste development and ovary activation, the two major plastic events in honeybee development, are known. These plastic responses can be manipulated in the lab, both by removing Queens from experimental hives and CO$_2$ necrosis of workers (Cridge et al. 2017). Though the genetic tools available for honeybees are not as powerful as for Drosophila, robust detection of RNA and protein expression patterns is possible (Cridge et al. 2017).

1.3 Honeybees Ovary Activation

Notch signalling is essential for ovary activation in honeybees (Duncan et al. 2016). In the inactive ovary germarium, Notch signalling is active, and downstream E(spl)-C (enhancer of split complex) genes are upregulated. In active worker and Queen bee ovaries, on the other hand, Notch signalling is inactive and E(spl)-C genes are downregulated. Blocking Notch signalling causes workers to activate their ovaries, even in the presence of Queen Mandibular Pheromone (QMP), indicating that Notch signalling is dominant over QMP repression (Duncan et al. 2016). Numb, an inhibitor of Notch signalling, is more highly expressed in the active ovary than the inactive ovary, and may control ovary activation upstream of Notch. Unfortunately, limitations of the type of in situ hybridization performed in this study tell us little about which cell types Numb and other Notch pathway genes are expressed in, and so what the function of Notch signalling in the ovary may be.

Though Notch signalling is essential to honeybee ovary activation, other mechanisms are also involved. Apoptosis in the adult worker ovary represses reproduction by destroying oocytes after they are specified in the germarium, and is dependent on
the gene *Anarchy* (Hartfelder et al. 2018; Ronai, Barton, et al. 2015; Ronai, Oldroyd, et al. 2015). Notch can cause apoptosis in *Drosophila* (Lundell 2003), so it is possible apoptosis is occurring downstream of Notch in the worker ovary. QMP signalling to the ovary also involves dopamine signalling: when worker bees are removed from the Queen and fed dopamine, they activate their ovaries at a higher rate than those not fed dopamine (Dombroski et al. 2003). Differential methylation of *Krüppel* is thought to be involved in the switch between active and inactive worker bee ovaries (Kilaso et al. 2017). Epigenetic marks are also thought to mark the boundaries of genes plastically expressed in the honeybee ovary (Duncan et al. in press). Therefore, many complex molecular mechanisms control honeybee ovary activation; Notch is central to this process.

The aim of this project was to investigate the function and regulation of Notch signalling in the honeybee ovary, one important known mechanism of phenotypic plasticity.

### 1.4 Connections between ovary activation and human disease

Why study honeybees in the context of human disease? As argued, models for complex biological process can help us discover underlying mechanisms. Additionally, similar molecular mechanisms may underpin both honeybee ovary activation and metabolic diseases. Notch signalling, which is so crucial to honeybee ovary activation, also plays multiple roles in metabolic syndrome, including in the liver, adipose tissue, and immune system (Bi & Kuang 2015). For example, in the mouse liver, increased Notch signalling appears to increase liver size and TAG levels, but inhibits insulin signalling. A high protein or fat diet also increases Notch expression in the liver (Pajvani et al. 2013; Li et al. 2014). In the human liver, Notch target gene expression correlates with *glucose-6-phosphatase* and *protein kinase 1* expression in overweight and obese patients, suggesting Notch also has a role in the
human liver (Valenti et al. 2013). Therefore, Notch signalling plays an important role in human disease. Another example of a signalling pathway that is important in metabolic syndrome is insulin signalling (O’Neill & O’Driscoll 2015). This signalling pathway also underpins caste differentiation in hymenoptera, and is thought to be important in the evolution of eusociality (Azevedo & Hartfelder 2008; Chandra et al. 2018). In the non-social Drosophila, insulin signalling controls the size of the GSC niche via Notch, providing an intriguing link between reproduction, nutrition, and Notch signalling (Hsu & Drummond-Barbosa 2009; Hsu & Drummond-Barbosa 2011). However it is unclear whether these mechanisms are hallmarks of plasticity, or whether this is an example of convergent evolution. Changes in the copy number of Notch orthologues in humans are thought to contribute to the growth of our brains during evolution (Suzuki et al. 2018); it does not follow that is related to metabolic syndrome or honeybee ovary activation. In addition, Notch signalling has an extraordinary number of roles in multiple tissue types, including seven roles in the Drosophila ovary alone (Xu & Gridley 2012). Therefore, as Notch is involved in so many diverse biological processes, its involvement in metabolic syndrome is probably unrelated to its role in ovary activation.

Of more interest than individual signalling pathways are the principles underpinning how these signalling pathways interact to control complex phenotypes, and more fundamental, as-yet undiscovered cellular mechanisms for achieving plasticity. These principles are likely to be involved in both the honeybee and multiple human diseases, not just metabolic syndrome and T2D. As an example of the types of fundamental mechanisms that may underpin plasticity, histone marks in the honeybee ovary appear to define clusters of genes that are differentially regulated during ovary activation (Duncan et al, in press). Such clusters of differentially regulated genes have been reported before: in C. elegans, clusters of ‘see-saw’ genes are differentially regulated in adulthood in response to starvation-induced lifepaths. Thus, epigenetic control of clusters of genes that respond to the environment may be a
conserved mechanism of phenotypic plasticity, involved in not just the honeybee but also human disease.

1.5 Notch signalling

The Notch signalling pathway is an incredibly well-conserved signalling pathway with multiple roles in development. The biochemical mechanism of Notch signalling is likewise highly conserved, and works in the following manner (see also Figure 1.1). The Notch receptor is a transmembrane protein expressed on the receiving cell surface. Binding of a ligand such as Delta or Serrate (transmembrane proteins expressed on neighboring cells) causes cleavage of Notch by γ-secretase, and migration of the Notch intracellular domain (NICD) into the cell nucleus, where it can
activate transcription of relevant genes (Bray 2016). There are multiple possible points of regulation of Notch signalling; in honeybee ovary activation, endocytosis and degradation of the Notch receptor by Numb is most likely (Duncan et al. 2016).

1.6 Diversity in insect ovary ultrastructure

Insects are a diverse clade, and contain several different ovary types. All insect ovaries contain two main regions, the germarium and vitellarium. The germarium is where the oocytes are specified and produced; in the vitellarium, the oocyte develops further, preparing to be laid (Lynch & Roth 2011). Drosophila and hymenoptera such as the honeybee and Nasonia have ovaries of the panostic merositic type, meaning the developing oocyte has ‘nurse cells’ that provide nutrients to the developing oocyte (Lynch & Roth 2011). In Drosophila, the oocyte and nurse cells are encapsulated together within a layer of follicle cells; in Nasonia and the honeybee, the follicle cells only surround the oocyte, and nurse cells are connected via a nutritive pore (see Figure 1.2 for images of different ovary structures) (King & Richards 1969; Lynch & Roth 2011). Beetles such as Tribolium castaneum and the Argentine stem weevil (Listronotus bonariensis), use a very different strategy: nurse cells exist in the germarium and supply nutrients to the oocyte via nutritive cords (see Figure 1.2). Other insect species lack nurse cells, instead all oocyte components are synthesised within the egg (Lynch & Roth 2011). This is an extraordinary variation in structure which all produce a similar output, the mature insect oocyte.

In the Drosophila germarium, the signalling pathways governing both the maintenance of the germ stem cell niche, and further differentiation in the germarium, are extremely well understood (Ting 2013). Cap cells control GSC self-renewal via BMP signalling. GSCs can differentiate into cystoblasts, which undergo four synchronous cell divisions to produce a 16-cell cystocyte, connected via fusomes then ring canals. One of these cystoblasts will become the oocyte, while the others become nurse cells (Ting 2013). In other insects, little is known about the signalling pathways
Figure 1.2: Phylogeny and diversity of insect ovary structure. Blue represents nuclei, green is F-actin. In all images, mature oocytes are to the left and terminal filament (anterior) are to the right. Bars over figures represent known roles (or expression) of Notch. Phylogeny and confocal slices not to scale. 

a: Schematic and representative confocal images of the honeybee ovary. 

b: Representative confocal images of the Nasonia ovary and germarium. Note the blue staining, representing nurse cell dumping RNA, and the apoptotic nurse cells posterior to this. Note also the different stages of the Nasonia germarium (which lacks a terminal filament). A white dotted line surrounds the cap cell. 

c: Representative confocal images of Drosophila ovary and germarium. 

d: Confocal images of Argentine stem weevil ovary (representative of other beetle species). All images are either unpublished data from the Dearden lab (used with permission of the author) or were taken during this project. Cartoon of the honeybee germarium adapted from one kindly provided by P. Dearden. oc: oocyte. fc: follicle cells. fs: fusome. nc: nurse cell.
operating in the germarium. Like *Drosophila*, the anterior honeybee germarium contains a series of cystocyte clusters, connected by fusomes (see top of Figure 1.2 for a schematic). The oocyte is then specified, and nurse cells are connected via ring canals (Tanaka & Hartfelder 2004). The oocyte further develops, associated with nurse cells, and is eventually enclosed with follicle cells, making a mature follicle (Tanaka & Hartfelder 2004).

The existence of discrete germ stem cells in the adult insect ovary does not appear to be well conserved. In *Nasonia*, there is evidence that four-cell clusters self-renew and differentiate, rather than the single germ stem cell (Griebel & Rübsam 2014). A study in *Tribolium* also found no evidence for the existence of GSC or cystoblasts (Trauner & Büning 2007). In honeybees, the location of the germ stem cells is unknown. Cell division is detected in the honeybee terminal filament, and at the base of the germarium, where the germ stem cells exist in *Drosophila* (Tanaka & Hartfelder 2004). Vasa, a germ stem cell marker, has been detected in the honeybee terminal filament (Dearden 2006; Tanaka & Hartfelder 2009). Therefore, if germ stem cells exist in honeybees, they most likely exist in one of these locations.

### 1.7 Potential roles of Notch in the honeybee

Notch plays multiple roles during development, many of which are conserved. Therefore, describing the role of Notch signalling in the ovary of other species generates predictions about how this pathway controls honeybee ovary activation. In the *Drosophila* germarium, germ stem cells self-renew and differentiate to produce cystoblasts that further mature into oocytes. This is controlled in part by Notch signalling: overexpression of Notch causes an increase in both cap cell and GSC number (Song et al. 2007). Figure 1.2 depicts the known roles of Notch signalling in the insect ovary. Note that little is known about the signalling pathways active in ovarian development in insects other than *Drosophila*, so a lack of a described role for Notch does not mean the pathway has no role. Notch appears to have a con-
served role in controlling follicle cells, as it has multiple roles regulating follicle cells in *Drosophila* (Xu & Gridley 2012), and in *Tribolium*, Notch signalling is necessary for follicle cell encapsulation (Bäumer et al. 2012). Notch protein is also expressed in a curious patched pattern in the honeybee follicle cells (Wilson et al. 2011), so presumably also regulates follicle cells in this species. Notch could regulate ovary activation via follicle cells, as inactive honeybee workers do not develop oocytes containing follicle cells. However, this is thought to be because differentiated oocytes apoptose (Ronai et al. 2016). In addition, ovaries of insects where *Notch* has been knocked down by RNAi still produce mature oocytes, which is not the case in the honeybee ovary (Bäumer et al. 2012; Duncan et al. 2016). Another plausible role for Notch signalling in the honeybee ovary is in maintenance of the germ stem cell niche. Components of the Notch signalling pathway are differentially expressed in this region of the ovary in the honeybee (Duncan et al. 2016), and Notch has a well-understood role in maintenance of the germ stem cell niche in *Drosophila*. Therefore a similar role in another species is plausible. In *Nasonia*, a species more closely related to the honeybee, disruption of Notch signalling via RNAi against Delta and Notch causes sterility and massive disruption to ovary structure, but the etiology of this phenotype has not been further examined (Rolleston 2016). Therefore, Notch could control honeybee ovary activation by controlling germ stem cells, or by an as-yet undiscovered mechanism.

1.8 Using Hybridization Chain Reaction (HCR) to investigate gene expression

In invertebrates, *in situ* hybridization (ISH) is typically performed using dioxygenin-labelled RNA probes. Though this method is central to analysis of gene expression, it is subject to several limitations. RNA probes have to be synthesized using run-off transcription from a plasmid template, which is time-consuming and technically challenging, as RNA readily degrades. Detection of the probe relies on enzymatic de-
position of a coloured pigment, limiting the resolution of the technique, and meaning
that the experimenter must decide when to stop the reaction, introducing bias (Jin
& Lloyd 1997). This can also cause comparatively low RNA expression domains to
be missed, and means the technique is not quantitative. In addition, ISH typically
detects only one gene per sample, meaning colocalization experiments are challeng-
ing or impossible (Jin & Lloyd 1997). HCR, in contrast, uses fluorescent amplifier
probes to allow detection of up to four genes in one sample. Under sufficiently high
resolution, each RNA molecule appears as one ‘diffraction-limited dot’ (Choi et al.
2016), meaning that expression domains can be unambiguously detected. These
dots can also be used to quantify RNA expression (Choi et al. 2018; Shah et al.
2016; Daniel Green et al. 2018).

The version of HCR used in this study (version 3.0, Choi et al. (2018)) works in the
following way. An initiator set specific to the gene of interest is purchased. This
initiator set consists of 5-10 pairs of “split initiator” probes: each half of the probe
pair binds to adjacent sections of the genome, and contains half of the initiator
sequence needed to trigger hairpin dimer formation. This use of a pair of probes
reduces non-specific RNA detection and background (Choi et al. 2018). The detect-
ing hairpins are pairs of self-assembling DNA oligonucleotides, containing a sequence
complementary to the initiator sequence, and the other hairpin. The hairpins are
also conjugated to a fluorescent dye (Choi et al. 2018; Choi et al. 2016). When
the hairpins are added to sample with initiator probes bound to an RNA of interest,
the hairpins bind to both the exposed initiator sequence, and each other. They
therefore self-assemble into long, fluorescent chains of DNA, which are detectable
using fluorescence microscopy (see Figure 1.3) (Choi et al. 2018).

There are limitations to HCR. Though the initiator probes are considerably cheaper
than genes clones, the hairpins are expensive (initiator: $360USD/100 reactions,
hairpins: $450USD/20 reactions, Molecular Instruments, Oct 2018). Therefore de-
pending on the experimental design, sample number, and tissue type, HCR exper-
Figure 1.3: Diagram of HCR v3.0 technique (Choi et al. 2018). Note that usually, 5-10 initiator pairs are used, instead of the two shown here, and that assembled hairpin chains are much longer.

Experiments may be considerably more expensive than ISH. Though four-channel HCR is possible, the spectral overlap between the Alexa Fluor 594 and 564 fluorophores makes distinguishing these channels challenging (E. Clark, pers. comm). Additionally, samples must be imaged using a confocal (or other fluorescent) microscope, which is time consuming and requires access to expensive equipment. However, the ability to quantitatively image up to four genes at once, at the single-cell resolution, has the potential to greatly aid both biomedical and fundamental biological research. HCR has been used to validate single-cell RNAseq data (Daniel Green et al. 2018), and to investigate genetic regulatory networks during development (Clark 2018; Nandagopal et al. 2018). A detailed comparison of HCR and in situ hybridization is in the discussion of this work.
1.9 Current work

In this project, I investigated two key questions about the role of Notch signalling in honeybee ovary activation. How is \textit{Numb} regulated in the honeybee ovary? And what is the cellular role of Notch signalling? Answering both of these will reveal mechanisms of phenotypic plasticity operating in the honeybee, that may therefore function in the context of human disease.

To answer the first question, I used various bioinformatic techniques to predict a cis regulatory element (CRE) regulating \textit{Numb}, cloned this into the expression vector pH-Pelican, had this transformed into \textit{Drosophila}, and investigated its ability to drive expression in a variety of tissues. I also predicted a series of transcription factors that may regulate honeybee ovary activation, and \textit{Numb} (See Chapter 3).

To answer the second question, I first investigated the expression of \textit{Numb} in active and inactive honeybee ovaries. I also investigated the role of Notch signalling in the non-social hymenoptera \textit{Nasonia}. In addition, I investigated the existence and location the honeybee germ stem cell niche, by staining for dividing cells in the germarium (see Chapter 4).
Methods

All materials are listed in Appendix 1.

2.1 Dissection of ovary and gut tissue

Honeybees were dissected in PBS under a Leica dissection microscope, after Dearden et al. (2009). Ovaries were fixed for five (inactive workers) or ten (active workers) minutes in a 1:1 mix of heptane and 4% formaldehyde in PBS. The lower heptane layer was replaced with 100% ice cold methanol, followed by the formaldehyde. Ovaries were washed 3 times with ice cold methanol, and stored in the freezer. Nasonia, Drosophila, and aphid ovaries, and Drosophila guts, were dissected in a similar manner, except they were dissected into the 1:1 heptane:formaldehyde mix for a maximum of 25 minutes, and fixed for 25 (Nasonia) and 10 minutes (others), respectively.

2.2 Hybridization Chain Reaction (HCR)

Tissue was prepared for hybridization as for conventional honeybee in situ hybridization (Dearden et al. 2009). Hybridization was performed as per generic samples (Choi et al. 2016). Reagents are given in Table 6.1. In-depth protocols are as follows. Stored ovaries were rehydrated in successive five minute 75%, 50%, and 25% methanol washes. Ovaries were then washed 3x five minutes in PTw. Ovarioles were separated under a dissecting microscope in PTw, then were digested in 1 mL PTw and 1 μL 20 mg/g Proteinase K for 20 minutes, washed once with PTw, and refixed in 4% formaldehyde for 15 minutes. Fixative was removed by washing 6x with PTw. Ovaries were then pre-hybridized in 500 μL of probe hybridization buffer for 30
minutes at 37°C. The probe solution, consisting of 1 μL of each pair of probes in 500 μL of probe wash buffer, prepared at 37°C, was added to the ovaries. The probe solution and ovaries were incubated overnight at 37°C. Excess probes were removed by washing four times, for 15 minutes each, in probe wash buffer. Samples were then washed three times, for 5 minutes each, with 5X SSCT, then left at room temperature until the end of the day. Samples were incubated in 500 μL of amplification buffer for 30 minutes at room temperature. The hairpin mixture was prepared by incubating 10 μL of each 3M hairpin stock at 95°C for 30 seconds, then left at room temperature for 30 minutes. To prevent bleaching of the hairpins, all subsequent steps were performed in the dark. 500 μL of amplification buffer was then added to the hairpins. The amplification buffer in the ovaries was replaced with the hairpin mixture and incubated overnight at room temperature. The hairpin solution was then removed and stored in the freezer, and the ovaries washed with 5X SSCT, in the following sequence: 2x 5 minutes, 2x 30 minutes, 1x 5 minutes. Ovaries were then counterstained with 1 μL DAPI in 1 mL 5x SSCT for 30 minutes, before being washed once more and stored in 70% glycerol. Ovaries were mounted on slides and imaged under the Leica FV1000 confocal, using the UPLSAPO30X 30X silicon oil lens.

2.3 in situ hybridization

Traditional in situ hybridization in both honeybee and Nasonia ovaries was performed as per honeybee protocols (Dearden et al. 2009). Reagents are given in Table 6.2. Preparation of ovaries was performed as described in Section 2.2, except Nasonia ovaries were digested with Proteinase K for 12 minutes, not 15. Instead of pre-hybridizing in probe hybridization buffer, ovaries were incubated in in situ hybridization buffer for 2 hours at 52°C. Hybridization solution was then replaced, 2 μL of (undigested) probe added, and the solution incubated at 52°C overnight. The ovaries were then washed with wash buffer three times at room temperature, before
being washed in successive 10, 15, 30, and 60 minute washes. Ovaries were then incubated overnight at 52°C in wash buffer. Wash buffer was removed and ovaries washed at room temperature 3x with PTw, before being blocked for 30 minutes in PBTw. PBTw was replaced with a 1:1000 solution of anti-DIG AP in PBTw, and incubated for another 90 minutes. Tissue was then rinsed 3x with PTw, and washed in successive 10, 15, 30, and 60 minute washes. AP buffer was used to detect staining. Ovaries were incubated twice in AP buffer, for 5 minutes each. Ovaries were then placed into staining solution and immediately transferred to a staining dish, and staining was monitored. When a clear expression pattern was visible, ovaries were destained in 100% methanol, before being washed 6x in PTw. Ovaries were then mounted in 70% glycerol and imaged on the Olympus BX61 light microscope.

2.4 Immunohistochemistry

*Drosophila*, *Nasonia*, and aphid ovaries were stained using *Drosophila* protocols (König & Shcherbata 2013). Ovaries were washed four times for 15 minutes each in PBT (see Table 6.3 for reagents), blocked in PBTB for one hour at room temperature, then left overnight in primary antibody in the fridge. Ovaries were then washed four times for 15 minute each in PBT, before being blocked for an hour in PBTB at room temperature. After adding pre-diluted secondary antibody ovaries were incubated in the dark overnight, before being washed for 15 minutes in PBT, stained with 1 μL/mL DAPI and/or 1 μL/mL Phalloidin-488, washed twice more in PBT for 15 minutes, and mounted in 70% glycerol. For honeybee ovaries, a similar protocol was used, except PTx and PBTx were used instead of PBT and PBTB, and ovaries were permeabilised in PTx for at least two hours before adding primary and secondary antibody. Imaging was performed under the FV1000 confocal microscope. Because the nuclear staining in *Drosophila* and *Nasonia* germaria was unambiguous, it was not necessary to control for microscope settings for these tissue types. Cell clusters were manually counted.
2.5 *Nasonia* RNAi

*Nasonia* RNAi was performed by pupal microinjection, following protocols from Lynch & Desplan (2006). Pupae (at the yellow-eye stage) were glued to coverslips using a thin layer of Crayola school glue. The coverslips were stuck to slides using a drop of water. Microinjection was performed using a Warner Instruments borosilicate needle pulled on the Narishige PC-10 at 60°C and briefly sanded to break the needle\(^1\). Wasps were microinjected at 60-100kPa for 60-90ms using the PLI-100 (Harvard Apparatus), depending on needle quality, then placed in vials and left to develop at 25°C for one week. dsRNA had previously been synthesized (Rolleston 2016), concentrations of 100-200ng/μL were used\(^2\).

2.6 Cloning

2.6.1 Restriction Digests

Restriction digests with the appropriate enzymes were set up as follows: 10 μL vector (with DNA > 200 ng/μL), 3 μL 10x digestion buffer, 1 μL of each enzyme, 15 μL dH₂O (see Table 6.4). The CRE insert was cloned from the pUC57 vector to the pH-Pelican vector with restriction enzymes Xba1 and BamH1 (plasmid maps in Figures 6.1 and 6.2).

2.6.2 Gel purification

Gel purification of plasmids was performed using the following protocol, adapted from Sun et al. (2012). The restriction digest reaction was run in a 1% agarose gel, and the DNA fragment was excised using a sterile blade. To purify the DNA, a 500 μL Eppendorf tube with holes in the bottom and a thin piece of cotton wool was used. The gel fragment was placed on the cotton wool, the small Eppendorf placed in a 1.7 mL Eppendorf, and the device centrifuged at 5000rpm for 5-10 minutes. To

\(^1\)Needle quality has a substantial impact on survival in microinjection experiments, hence the level of detail

\(^2\)Such low concentrations have given good results in the past
remove salts for ligation, the eluate was ethanol precipitated. The volume of eluate was made up to 100 μL using dH$_2$O, and 10 μL of 3M sodium acetate added. After a brief vortex, 250 μL of ice cold ethanol was added, and the mixture incubated in the freezer for two hours or overnight. The sample was then centrifuged for 15 minutes at 17,000rpm, the supernatant removed, and 250 μL of 75% ethanol added to wash the DNA pellet. After another 5 minute centrifuge, the ethanol was removed, and remaining ethanol left to evaporate. The DNA pellet was resuspended in dH$_2$O. Initial validation of this method was performed by running the resuspended DNA on a gel, and by HS-DNA Qubit before ligation.

### 2.6.3 Ligation and Transformation

Cloning was performed after Sambrook & Russell (2006). The ligation reaction was set up as follows: 25ng vector DNA, 75ng insert DNA, ligase buffer at 1X concentration, 1 μL T4 DNA ligase, H$_2$O to 10 μL (or the volume of the vector and insert ligation reaction). A control, vector only, ligation reaction also performed for each ligation. The mixture was incubated overnight at 16°C.

Competent XLI-Blue *Escherichia coli* stored at -80°C were used for transformation. These bacteria were thawed on ice, before 0.1 – 5 μL of ligation reaction was added, and the bacteria incubated on ice for 15 minutes. Bacteria were heatshocked at 42°C for one minute, before 1 mL LB was added, and incubation for 30 minutes at 37°C. 100 μL of culture was spread on LB + Ampicillin plates. The transformed bacteria were then spun down at 4000g, the supernatant removed, and the remaining culture spread onto plates. Plates were incubated at 37°C overnight.

### 2.6.4 Colony PCR

To identify colonies containing the insert, large colonies from plates incubated overnight were picked with a pipette tip and suspended in 6 μL LB. 3 μL of this mixture was added to 8 μL 0.5% Tween-20, and heated at 100°C for 30sec in a thermal cycler. 1 μL of this mixture was used as template for polymerase chain reaction.
(PCR). The PCR mix was set up with 2 μL buffer, 1 μL dNTPs, 2 μL each forward and reverse primer (see Table 2.1 for sequences), 0.2 μL Taq polymerase, and 11.8 μL nuclease free water, and 1 μL of template. The PCR reaction was denatured at 94°C for 3 minutes. For 35 cycles, the reaction was denatured at 94°C for 30 seconds, annealed at 55°C for 30 seconds, and extended at 72°C for one minute. It was then incubated at 72°C for five minutes, before being held at 4°C for further analysis. The PCR reaction was run on a 1% agarose gel to identify colonies with the insert. The remaining 3 μL LB/colony mix was used to set up overnight cultures in 5 mL of LB and 5 μL of 50 mg/mL ampicillin.

**2.6.5 Plasmid purification and sequencing**

Plasmids were purified using the Invitrogen midiprep kit, per the manufacturers instructions. To obtain high-concentration, salt-free DNA for embryo microinjection, DNA was ethanol precipitated and cleaned as described earlier, resuspended in 20 μL of ultrapure water, nanodropped, then diluted to 1 μg/μL DNA. Sequencing was performed using the Otago University Genetics Analysis Services, on the ABI 3730xl DNA Analyser. Sequencing reactions were set up with 200 ng of plasmid DNA, 3.2 pmol of primer (see Table 2.1 for sequences), in 5 μL H2O. The identity of the original insert and vector plasmids were also validated in this way. Sequences were processed, checked for quality, and aligned using the Emboss suite command line tools `abiview`, `revseq`, and `merger` (Rice et al. 2000). BLAST (Madden 2013) was used to validate the aligned sequences.

<table>
<thead>
<tr>
<th>Table 2.1: PCR primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
</tr>
<tr>
<td>pH-Pelican forward</td>
</tr>
<tr>
<td>pH-Pelican reverse</td>
</tr>
</tbody>
</table>
2.6.6 Transformation of Drosophila

Microinjection and transformation of vectors into Drosophila was performed using the Genetivision Drosophila microinjection service (http://www.genetivision.com/).

2.6.7 Drosophila embryo fixation

After heat-shocking for 2 hours, Drosophila embryos were methanol fixed as per standard protocols (Sullivan et al. 2000, pp.143–151). Flies were placed into laying cages and provided with an apple juice plate and yeast. After 24 hours, embryos washed into a cell strainer and rinsed with water. Embryos were rinsed in 7% NaCl before being dechorionated in 50% bleach for ~5 minutes, before the bleach was rinsed off with tap water. The embryos were then transferred to a 1:1 mix of heptane:4% formaldehyde in PBS, before being vortexed for 20 minutes. The bottom layer was replaced with ice-cold methanol, shaken again, and rinsed 3x in ice-cold methanol. Embryos were then rehydrated before being stained.

2.6.8 Staining for LacZ

Adult Drosophila were exposed to 26 queen equivalents (10 $\mu$L) of QMP for 48 hours before being heat-shocked for 2 hours. LacZ detection relied on enzymatic cleavage of X-gal to galactose and 5-bromo-4-chloro-3-hydroxyindole, which oxidizes to a blue product (Sullivan et al. 2000). Embryos, ovaries, and gut tissue were rehydrated (if necessary), washed 3x in PBS, before being incubated in staining solution + 8 $\mu$L/ml of 8% X-Gal overnight. The staining solution was then removed, and the solution washed 3x in PBS, before being mounted in 70% glycerol and imaged.

2.7 Bioinformatics

All scripts are available at my github: github.com/Shannon-E-Taylor. Script URLs are given in footnotes.
2.7.1 CLARE

CLARE uses a machine learning technique to predict transcription factor binding sites that are overexpressed in a set of co-regulated sequences (Taher et al. 2012). A series of overexpressed motifs, and transcription factors from the JASPAR database, are given. To predict regulatory motifs controlling worker ovary activation, a custom R script was used³ to identify the sequences 10kb upstream of differentially expressed genes. To ensure genes identified were those involved in the switch between active and inactive ovaries, rather than general reproductive status, only genes differentially expressed (false discovery rate <= 0.05) between both active and inactive ovaries, and active and Queen ovaries, were used. Regulatory motifs obtained from CLARE were run through both ClusterDraw and ClusterBuster.

2.7.2 ClusterDraw and ClusterBuster

ClusterDraw predicts CREs by taking into account the frequency of transcription factor binding sites in a particular “cluster”/region of the genome (Papatsenko 2007). High cluster scores correspond to likely binding site clusters. The program was run using the online web server⁴, using default parameters and the *Apis mellifera* GC content.

Like ClusterDraw, ClusterBuster uses motif-binding sites to predict regulatory regions of DNA (Frith et al. 2003). This tool was run using the online web server⁵, using default parameters.

2.7.3 Alignments

BLAST was used to identify the Numb orthologues in *Apis cerana*, *Bombus terrestris*, *Bombus impatiens*, and *Nasonia vitripennis*. Alignments between the honeybee and other hymenopteran Numb were performed using the “Align two sequences”

⁴http://line.bioinfolab.net/webgate/submit.cgi; this has since been taken down
⁵http://cagt.bu.edu/page/ClusterBuster_submit
function.

2.7.4 SCRMshaw

SCRMshaw (Kazemian et al. 2014) was run using the following parameters\(^6\), and is now available as a singularity container\(^7\):

- `thig 300 --imm --hexmcd --pac`
- `genome` Amel4.5 assembly (Elsik et al. 2014)
- `gene` and `exon`: gene and exon positions\(^8\)
- `traindirlst`: the *Drosophila* regulatory regions provided with SCRMshaw

2.7.5 Image analysis

Image analysis was performed in Fiji. All HCR images in this thesis have been background-subtracted in Fiji to improve contrast\(^9\). For dHCR, the `vasa` channel was used as a mask to define the region of `Numb` expression to be measured. The DAPI, `Numb` and `vasa` channels were then background-subtracted, auto thresholded, and the `Analyze Particles` command was used to count particles\(^10\). Justification of this workflow is provided in the Results section.

\(^6\)https://github.com/Shannon-E-Taylor/apis-numb/blob/master/Snakefile  
\(^7\)shub://TomHarrop/singularity-containers:scrmshaw_20180523  
Chapter 3

Regulation of *Numb*

The next two chapters present and discuss my results, first in relation to the regulation of *Numb*, and the second in relation to the role of Notch signalling in ovary activation. To investigate the regulation of *Numb*, I predicted a *Numb* CRE, and expressed it in *Drosophila* to investigate its regulation under different environmental conditions.

3.1 Results

Two general methods are used to predict CREs. Motif-based methods use the frequency of transcription factor binding motifs in a particular genome location to predict CRE location. Motif-independent methods include machine-learning methods, genome alignments, and chromatin structure information to predict CREs without needing to understand the transcription factors (TFs) involved. I used both approaches. To predict potential TFs binding Numb for use with motif-based CRE prediction tools, TFs binding *Drosophila Numb* were obtained from the RedFly database (Gallo et al. 2011). Suppressor of Hairless (Su(H)) was the only TF predicted from this technique (Rebeiz et al. 2011). Additionally, I used CLARE to predict TFs that may be involved in honeybee ovary activation. These TFs, and motifs, were used in the ClusterBuster and ClusterDraw tools. These tools predict TF binding sites based on both strength of TF binding to a particular site, and density of binding sites within a given region of the genome.

In addition, I used motif-independent methods to predict CREs within *Numb*: SCRMshaw, a machine-learning pipeline to predict CREs, and alignments of
Am-Numb against other insect species. I investigated the entire NCBI annotation of Numb, but no sequence upstream, as Numb lies directly adjacent to another gene at the 5’ end, and alignment against Drosophila Numb revealed that the NCBI annotation was most complete.

3.1.1 Predicting regulatory motifs: Cracking the LAnguage of Regulatory Elements (CLARE)

To predict transcription factors that may be involved in honeybee ovary activation, I used the CLARE tool, which takes a set of co-regulated sequences and identifies overexpressed motifs (Taher et al. 2012). These co-regulated sequences were 10kb upstream of genes differentially expressed between both Queens and active workers, and active workers and inactive worker honeybees1.

CLARE predicted a total of 27 potential regulatory motifs: 17 from the JASPAR database and ten de novo predicted (see Table 3.1). Motifs with a high positive weight are likely to be activators, negative motifs are likely to be repressors (or activators not currently expressed). Of these, Optix and run::Bgb occurred with the highest frequency in the training set. Run::Bgb was also frequently present in the background data (see Figure 3.1). One way to test the accuracy of a machine learning pipeline is to plot the true positive rate against the false positive rate, producing a receiver operating characteristic (ROC). A ROC of 1.0 represents perfect specificity: no false negatives or false positives (Taher et al. 2012). This is the ROC CLARE produced for my input data. Therefore, according to the algorithm defining true and false positives in the CLARE program, all of the transcription factors predicted were truly involved in honeybee ovary activation. I also ran CLARE on regulatory sequences 1kb and 5kb upstream from differentially expressed genes; these runs gave no JASPAR transcription factors, and smaller ROC values (ROC = 0.7 and 0.98,

---

respectively), so I did not pursue them further.

Of the predicted TFs, 12 were expressed in the honeybee ovary (Duncan & Dearden, unpublished data), and five were differentially expressed between active and inactive ovaries (see Table 3.1). These non-expressed TFs could due to misannotation of the honeybee genome, or poor power in the data set. For differentially expressed TFs, all positively weighted genes are unregulated in active ovaries, while the negatively weighted TF is downregulated. Non-differentially expressed genes could be regulated at the protein level. The upregulation of positively weighted TFs does validate the predictions of CLARE. The motifs for Awh and CG15696 closely resemble each other; it is likely that only one of these genes is involved in ovary activation.

Table 3.1: Regulatory motifs predicted by CLARE. Named motifs are from the JASPAR database, priority motifs were de novo predicted.

<table>
<thead>
<tr>
<th>TF name</th>
<th>Sequence</th>
<th>Weight</th>
<th>Expression in active vs inactive ovaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eip74EF</td>
<td>GGAA</td>
<td>405.15</td>
<td>Up</td>
</tr>
<tr>
<td>brk</td>
<td>TGGC</td>
<td>281.87</td>
<td>Not expressed</td>
</tr>
<tr>
<td>hb</td>
<td>AAAAA</td>
<td>261.16</td>
<td>Not significant</td>
</tr>
<tr>
<td>CG42234</td>
<td></td>
<td>251.93</td>
<td>?</td>
</tr>
<tr>
<td>Deaf1</td>
<td></td>
<td>151.86</td>
<td>Up</td>
</tr>
<tr>
<td>Awh</td>
<td>TAATTA</td>
<td>97.05</td>
<td>Up / Not significant</td>
</tr>
<tr>
<td>4_PRIORITY CGATMTWCGA</td>
<td>84.14</td>
<td>Not Expressed</td>
<td></td>
</tr>
<tr>
<td>CG15696</td>
<td></td>
<td>64.13</td>
<td>Not Expressed</td>
</tr>
<tr>
<td>3_PRIORITY CGATCGTTCG</td>
<td>52.36</td>
<td>Not Expressed</td>
<td></td>
</tr>
<tr>
<td>1_PRIORITY TTCGATCGAT</td>
<td>38.93</td>
<td>Not Expressed</td>
<td></td>
</tr>
<tr>
<td>6_PRIORITY ATCRTCGTCG</td>
<td>35.78</td>
<td>Not Expressed</td>
<td></td>
</tr>
<tr>
<td>TF name</td>
<td>Sequence</td>
<td>Weight</td>
<td>Expression in active vs inactive ovaries</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>--------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>h</td>
<td><img src="image" alt="h sequence" /></td>
<td>27.2</td>
<td>Not significant</td>
</tr>
<tr>
<td>10_PRIORITY</td>
<td>TCGAWCGAAT</td>
<td>19.06</td>
<td>Not Expressed</td>
</tr>
<tr>
<td>Optix</td>
<td><img src="image" alt="Optix sequence" /></td>
<td>4.18</td>
<td>Up</td>
</tr>
<tr>
<td>run::Bgb</td>
<td><img src="image" alt="run::Bgb sequence" /></td>
<td>-2.7</td>
<td>Not significant</td>
</tr>
<tr>
<td>fkh</td>
<td><img src="image" alt="fkh sequence" /></td>
<td>-29.61</td>
<td>Not significant</td>
</tr>
<tr>
<td>till</td>
<td><img src="image" alt="till sequence" /></td>
<td>-48.53</td>
<td>Not significant</td>
</tr>
<tr>
<td>sd</td>
<td><img src="image" alt="sd sequence" /></td>
<td>-59.02</td>
<td>Not significant</td>
</tr>
<tr>
<td>br_Z3</td>
<td><img src="image" alt="br_Z3 sequence" /></td>
<td>-96.97</td>
<td>Not significant</td>
</tr>
<tr>
<td>lbe</td>
<td><img src="image" alt="lbe sequence" /></td>
<td>-99.66</td>
<td>Not Expressed</td>
</tr>
<tr>
<td>slp1</td>
<td><img src="image" alt="slp1 sequence" /></td>
<td>-103.24</td>
<td>Not Expressed</td>
</tr>
<tr>
<td>mirr</td>
<td><img src="image" alt="mirr sequence" /></td>
<td>-131.93</td>
<td>Down</td>
</tr>
</tbody>
</table>

### 3.1.2 An attempt to validate the CLARE predicted TFs

If these predicted TFs were involved in honeybee ovary activation, some would be differentially expressed in the honeybee gerarium. I therefore attempted to investigate the expression of the CLARE predictions *hairy* (*h*) and *hunchback* (*hb*) via *in situ* hybridization in both inactive worker ovaries, and Queen ovaries. After incubation with AP buffer, staining was only visible in the intima (a membrane surrounding the honeybee ovariole) for all probes (data not shown). The ovaries were then left overnight in AP buffer, but staining became saturated, so no expression pattern was visible (see Figure 3.1). This failed *in situ* was not surprising, as the
probes were approximately 10 years old and likely degraded. I did not attempt validation of any other genes, due to the lack of accessible probes or antibodies.

### 3.1.3 Predicting motif-binding sites: ClusterDraw and ClusterBuster

To predict CREs that might bind transcription factors of interest within the *Numb* gene, I used the tools ClusterBuster and ClusterDraw. These tools predict genome regions that bind TFs, based on the density of TF binding sites. ClusterDraw is more sensitive, but only six TF motifs can be used, meaning interactions between
different TFs may be missed by the analysis. Therefore, ClusterBuster, which allows more TF motifs, was also used.

There are three predicted Su(H) binding sites in *Am-Numb* (see Figure 3.2). The motifs predicted by CLARE had fewer binding sites: of the JASPAR motifs, only hairy and brk had regulatory sites predicted by ClusterDraw. Both of these TFs were positively weighted by CLARE (see Table 3.1), so may activate *Numb* expression. The *de novo* predicted motifs were more successful: all were present in at least one cluster (see Figure 3.2). ClusterBuster predicted four CREs, and detected hairy and Su(H) binding; all of the *de novo* predicted motifs had at least one binding site. Therefore, undiscovered TFs, or miRNAs, may play an important role in *Numb* regulation.

### 3.1.4 Motif independent techniques

Motif-dependent CRE prediction methods rely on knowing the transcription factors that may be functioning within a particular tissue. This knowledge is lacking with regard to the honeybee. Though I attempted to predict transcription factors in honeybee ovary activation using CLARE and the RedFly database, also using motif-independent methods would have increased the chance of successfully predicting a CRE, if the same gene region was predicted by distinct methods. Therefore, I also used motif-independent methods: genome alignments and the machine-learning tool SCRMshaw. Nucleotide alignments to *Drosophila melanogaster* and *Nasonia vitripennis* revealed very little sequence conservation outside of the exons, likely because these species are too far diverged from the honeybee (Kazemian et al. 2014). Alignment to *Drosophila* also revealed that the version of *Numb* deposited in Beebase is truncated, lacking the 5’ end, while the NCBI version of *Numb* is complete. Therefore, the NCBI *Numb* annotation was used in downstream analysis. *Apis cerana* was too closely related to the honeybee to be useful for CRE prediction, as virtually the entire gene aligned. The *Bombus* species were more useful, as some
distinct ~1000bp regions within introns aligned with honeybee Numb (see Figure 3.3). These regions are possible CREs. SCRMshaw also predicted four regulatory regions, of varying length (see Figure 3.3). SCRMshaw always predicts 500bp CREs, the three CREs that were smaller than this contained some unassigned nucleotides.
3.1.5 Cloning the CRE into pH-Pelican

A graphic summary of all predicted CREs are presented in Figure 3.3. I chose to use the SuH binding site at ~25kb after the start of Numb (see box in Figure 3.3), because of its proximity to a large SCRMshaw-predicted CRE, and therefore prediction by two distinct methods. This was cloned into the pH-Pelican vector, at the Xba1 and BamH1 restriction sites, to allow easy detection in Drosophila (Barolo et al. 2000). After ligation and transformation of the cloned vector, a ~700bp sequence (the same size as the CRE insert) was visible after double restriction enzyme digestion (see Figure 3.3). There was a very faint band in the uncut vector just below the CRE insert, which was not visible in the initial gel photo. This may represent contamination or fragmentation of the vector. Sequencing confirmed the CRE insert was successfully cloned into pH-Pelican. The vector was then transformed into Drosophila using the Genetivision service.

3.1.6 The CRE is expressed in the Drosophila gut

---

2This is also not always visible in printed copies of the gel; see the pdf
Table 3.2: Proportion of CRE-positive (white) and negative (red/yellow) flies. Results from two independently transformed fly lines (Populations 3 and 4) are presented separately.

<table>
<thead>
<tr>
<th>Population</th>
<th>White eyes</th>
<th>Red/yellow eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population 3</td>
<td>34</td>
<td>38</td>
</tr>
<tr>
<td>Population 4</td>
<td>33</td>
<td>31</td>
</tr>
</tbody>
</table>

I investigated expression of the predicted CRE in three places: first, the embryo, as Su(H) drives *Numb* expression in the *Drosophila* embryo, second, the ovary, as differential expression of *Numb* in the honeybee ovary controls ovary activation. Thirdly, I also investigated the gut, because Notch signalling has an important role in the *Drosophila* gut (Sallé et al. 2017; Obniski et al. 2018). Due to time limitations, I did not use virgin female flies for QMP exposure experiments, as is the normal procedure in our lab. This meant fly age (and sex) may have confounded the results. All experiments were done on flies from two independent transformations of the pH-Pelican vector. As expected, ~50% of each population was white-eyed (see Table 3.2), meaning the CRE was present in these flies. Only white-eyed (CRE positive) flies (females and males) were exposed to QMP. Embryos were collected from mixed populations, which should have resulted in ~25% of embryos expressing LacZ.

LacZ expression was detectable in the anterior and posterior *Drosophila* gut, but not in the embryo or ovary (see Figure 3.4). Though non-specific detection of LacZ is common in the gut (P. Dearden, pers. comm.), negative controls (wild-type OregonR flies) did not express LacZ. Therefore, the predicted CRE drove LacZ expression in the *Drosophila* gut. As the CRE is a Su(H) binding site, it is possible that it was Su(H), not factors intrinsic to *Numb*, that were driving this expression. This expression was responsive to QMP in population three ($\chi^2=3.97$, p-value=0.046), but not population four ($\chi^2=0.041$, p-value=0.83) (see Table 3.3). This difference in regulation by QMP could be due to the reporter inserting into different locations,
with varying response to QMP, or bias because of fly age and sex.

Table 3.3: Expression of LacZ in the guts of QMP exposed and unexposed Drosophila. Results from two independently transformed fly lines (Populations 3 and 4) are presented separately.

<table>
<thead>
<tr>
<th>Population</th>
<th>LacZ expressed in gut</th>
<th>Not expressed in gut</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 +QMP</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>3 -QMP</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Population</td>
<td>LacZ expressed in gut</td>
<td>Not expressed in gut</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>4 +QMP</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4 -QMP</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

### 3.2 Discussion

To investigate how *Numb* is regulated in the honeybee ovary, I predicted a series of CREs, and chose one to clone and express in *Drosophila*. This CRE was chosen because the same region of *Numb* was predicted by two methods: ClusterDraw and SCRMshaw. This was expressed in the gut, rather than the ovary or embryo. Biologically, this is plausible, as *Numb* regulates differentiation of the enteroendocrine cells of the *Drosophila* midgut, and Su(H) (the transcription factor binding the CRE) has a related role (Sallé et al. 2017; Bardin et al. 2010). Notch signalling alters differentiation of enteroendocrine cells in the *Drosophila* gut in response to altered cholesterol intake via altering Delta stability; similar changes in Notch signalling occur in the ovary (Obniski et al. 2018). Changes in the *Drosophila* gut also mediate the the relationship between fertility and diet (Reiff et al. 2015). Therefore, the expression of this *Numb* CRE in the *Drosophila* gut, in a potentially QMP responsive manner, suggests that Notch signalling in the gut may be important in the *Drosophila* response to QMP. However these data are confounded by age and sex, so the experiment needs to be repeated in a more controlled manner before definitive conclusions can be drawn.

The obvious next experiment (aside from improving the execution of the previous one) is to stain for Notch and Delta in the QMP exposed and unexposed *Drosophila* gut, as a more thorough test of the involvement of Notch in the QMP response. In addition, as the predicted CRE binds Su(H), it was unclear whether its expression was because *Numb* is involved in regulating Notch in this region, or whether Su(H)
is activating the CRE independent of its role in Numb. Staining for Numb protein in the Drosophila gut would resolve this. Once the role (if any) of Notch in the Drosophila response to QMP has been better defined, these tests could be extended to the honeybee gut of both active and inactive workers. In addition, if better controlled experiments reveal that the different fly lines do have distinct responses to QMP, identifying the location of the CRE insert would help identify genome regions that may be QMP responsive in Drosophila.

3.2.1 How reliable are the predicted CRES?

I predicted several CREs that may be involved in regulating Numb, and therefore honeybee ovary activation, and generated a fly line to test the expression of one CRE. This was able to drive LacZ expression in the Drosophila gut, meaning the CRE prediction was successful. As the true positive rate for CRE prediction range from 20-80%, with lower true positive rates being much more common, without experimental validation it is impossible to confirm whether the other predicted CREs were real (Kazemian et al. 2014). There was limited agreement between CREs predicted based on motif-binding sites using motif-based methods. The largest ClusterBuster-predicted motif binding site was close to two Su(H) binding sites predicted by ClusterDraw, and two CLARE-predicted motif binding sites. Moreover, this location in Numb, at the 3’ end of the second intron, is where a Su(H) enhancer exists in Drosophila (Rebeiz et al. 2011). The strongest Su(H) binding site predicted by ClusterDraw also lay adjacent to a SCRMshaw-predicted CRE. This prediction of the same region of the genome by two distinct methods suggests it was a true CRE, and indeed, it was able to drive LacZ expression when transformed into Drosophila.

ClusterBuster predicted CREs within other genes: one inside a long non-coding RNA in the first intron of Numb, another within the second Numb exon. These are unlikely to be real CREs, and this result suggests that the ClusterBuster predictions are not always reliable. Ultimately, it is impossible to draw meaningful conclusions
from these predictions without experimental validation, especially considering that the location of CREs in the genome tells us little about their function.

3.2.1.1 Challenges running SCRMshaw

As a command-line program, SCRMshaw was more challenging to run than the other web interfaces used in this study. It did not run with the unprocessed data we provided, and substantial data processing and troubleshooting was required to get useful information out of the program\(^3\). A description of this process follows. The genome annotation file available at Beebase was incompatible with SCRMshaw, so a genes and exons input file was generated. Initially, SCRMshaw only returned hits to the unscaffolded regions of the honeybee genome, again because the input genome was not formatted as expected. This was initially solved by renaming the scaffolds to a format SCRMshaw could handle\(^4\), this has been resolved in the publicly available code. Once SCRMshaw was running, it was important to run the code on the entire honeybee genome: as SCRMshaw ranks predicted CREs to determine hits, rather than using an absolute probability that the hit is real, using just one chromosome led to the entire Numb gene being predicted as a CRE\(^5\). I accepted CREs with a 'local rank' $\geq 3$ as possible CREs. SCRMshaw is now packaged into Singularity\(^6\) so that it will run on any computer. Thus, this troubleshooting process has resulted in a more reproducible system.

3.2.2 Interesting predicted transcription factors

The transcription factors predicted by CLARE may give insight into genes controlling honeybee ovary activation. CLARE uses a set of co-expressed genes to predict overexpressed motifs, either predicted de novo or from the JASPAR database (Taher

\(^3\)Tom Harrop was of enormous help in getting the program to run, and wrote all early versions of scripts used for this part of the project (https://github.com/Shannon-E-Taylor/apis-numb)

\(^4\)Initial efforts to achieve this had amusing results: the entire honeybee genome was overwritten with the letter 'p', because I did not know how to use 'RegEx' properly.

\(^5\)Another amusing result, once I understood why.

\(^6\)By Tom Harrop
et al. 2012; Khan et al. 2018). CLARE gave a very high ROC value, so according to this algorithm the predicted transcription factors are all correct. Bioinformatics tools are only as reliable as the input data, however, and there were limitations with the data I gave to CLARE. Firstly, many genes were differentially expressed between active and inactive worker honeybees. As CLARE can only take a 2Mb (~2Mbp) file of sequence, this was too many genes to use. Therefore I filtered data to only include genes differentially expressed between active and inactive workers, to identify genes important in ovary activation. I further filtered this set to only include genes differentially expressed between active workers and queens, to remove genes required for reproduction. This gave a substantially smaller list of genes (~50 instead of ~5000). It is unclear how well this smaller list represents genes involved in ovary activation. Secondly, the RNAseq data used was from stage three active ovaries. These are already able to lay eggs, so the gene set may not represent genes involved in early ovary activation. Thirdly, I had to decide which region of these genes to use as a regulatory sequence. To do this, I took 10kb of sequence upstream of the start of a gene. This almost certainly missed some regulatory DNA, as CREs can lie a substantial distance from the transcription start side of a gene (Bentovim et al. 2017), and would have included some portions of genes. Therefore, though these limitations may bias the CLARE data, a more precise algorithm was beyond the scope of this project.

Limitations of the data aside, CLARE predicted some interesting transcription factors, including two ecdysteriod-response genes, *E74* and *Broad Complex (BR-C)*. Ecdysteroids are known to be involved in both ovary activation, and larval fate choice between Queen and Worker (Hartfelder et al. 2015). *E74* is expressed in the Queen follicle cells, posterior nurse cells, and oocyte, and may regulate uptake of egg chamber components such as vitellogenin (Paul et al. 2005). *BR-C* is also expressed in the Queen follicle cells (Paul et al. 2006). The expression of these genes in worker ovaries is unknown. Interestingly, binding sites for both these transcription factors
were identified within the Essential for Life cluster of genes, which are coordina-
tively regulated in the terminal filament of honeybee ovaries (Lovegrove 2013). The
prediction of two transcription factors that could reasonably be expected to have a
role in ovary activation, due to their function and expression, suggest that CLARE
is making accurate predictions.

One class of genes that was not present in the CLARE predictions were the Enhancer
of Split complex genes, a set of transcription factors that respond to Notch signaling
(Dearden 2015). A Notch-responsive gene, \textit{hairy}, was predicted by CLARE, but not
investigated by Duncan et al. (2016), and is not differentially expressed in the ovary
activation dataset. As \textit{hairy} binding sites are present in honeybee \textit{Numb}, this gene
may be an upstream regulator of \textit{Numb} and Notch signalling. Another missing
honeybee ovary activation gene was \textit{gemini}: RNAi against this gene causes ovary
activation (Jarosch et al. 2011). Though Su(H) is known control \textit{Numb} expression
in \textit{Drosophila}, it was not predicted by this analysis in honeybees. These missing
genes may be false negatives by CLARE, or the genes may not be involved in ovary
activation either generally or in this dataset.

CLARE predicted some TFs, where speculation as to their role in ovary activation is
possible. One such TF was \textit{brinker} (\textit{brk}), a transcriptional repressor that modulates
BMP signalling. In the \textit{Drosophila} ovary, \textit{brk} has a role in eggshell patterning, and
loss of \textit{brk} causes downregulation of \textit{Br-C} expression. Though \textit{brk} controls BMP
signaling in the \textit{Drosophila} ovary, it is unlikely to be involved in BMP signaling in
the \textit{Drosophila} GSC niche (Chen & McKearin 2003); a plausible mechanism of action
for \textit{brk} signaling in honeybee ovary activation is via modulation of Br-C. According
to my ClusterDraw analysis, \textit{brk} also binds \textit{Numb}, so may regulate this gene. The
gene \textit{hunchback}, also predicted by CLARE, is involved in early embryonic patterning
in honeybees (Wilson & Dearden 2011), and is expressed in the honeybee ovary, but
there is no known regulatory role for the gene in either the \textit{Drosophila} or honeybee
ovary. It is worth noting that many insect developmental genes (including \textit{hb}) are provided maternally and are therefore expressed in the ovary (Lynch & Roth 2011). This does not mean these genes play a role in ovary development or function. A possible link to ovary function is that in early \textit{Nasonia} (and \textit{Drosophila}) patterning, the germ cell marker \textit{nanos} represses \textit{hunchback} (Lynch et al. 2010). This suggests a possible link between GSC proliferation and \textit{hb}. \textit{Scalloped} (\textit{sd}) is activated by the Hippo signalling pathway (Staley & Irvine 2012). Hippo is involved in pole cell specification and controls organ size and stem cell proliferation in multiple organs, including the \textit{Drosophila} ovary (Sarikaya & Extavour 2015). \textit{Mirror} integrates Notch signaling to specify the \textit{Drosophila} pole cells (Jordan et al. 2000), so may play a role in ovary activation downstream of Notch.

Other genes with a less clear potential role in germ cell development were also predicted by CLARE. \textit{Deaf1} is involved in aggression in honeybees, and is necessary for segmentation in \textit{Drosophila} (Alaux et al. 2009; Veraksa et al. 2002). \textit{Awh} has a role in specifying the \textit{Drosophila} imaginal disks, structures that give rise to adult organs (Curtiss & Heilig 1995). \textit{Tailless} is known to play a role in terminal patterning of both the \textit{Drosophila} and honeybee embryo, but has no known role in ovary development (Wilson & Dearden 2009). \textit{Optix} is a Hox gene required for \textit{Drosophila} eye and brain development (Gold & Brand 2014), while \textit{runt} and \textit{slp} are involved in segmentation (Clark 2017). \textit{Forkhead} is specific to secretory cells (Kerman et al. 2006), while \textit{lbe} is a hox gene with a role in mesoderm patterning (Jagla et al. 2001). Speculation on the role of these genes, if any, in ovary activation is impossible with the data available.

In this section, to investigate the regulation of \textit{Numb}, I predicted a CRE that was able to drive LacZ expression in the \textit{Drosophila} gut. If this is truly QMP-responsive, it would suggest that Notch-mediated changes in the \textit{Drosophila} gut control QMP response. The transcription factors predicted by CLARE are also worth investigat-
ing further, as potential drivers and effectors of honeybee ovary activation. *brinker* and *hairy* binding sites are present in the *Numb* gene, so these transcription factors may regulate Notch upstream of *Numb*. The ecdysteroid response genes and *mir-
or* also have intriguing biological roles in the insect ovary that are worth further exploring in the context of honeybee ovary activation. This provides steps towards better understanding how plastic responses in the honeybee ovary are regulated.
Chapter 4

Notch signalling in the honeybee

Notch signalling is essential in honeybee ovary activation (Duncan et al. 2016). However, its cellular function in this species is unknown. Therefore, I investigated plausible roles of Notch signalling in the honeybee germarium in three ways. First, I investigated the cell-level expression of the Notch regulator *Numb* using HCR, as the cell types this gene localizes to, will give insight regarding its function. Secondly, to define the region of the germarium where *Numb* is required for ovary activation, and to further establish quantitative HCR, I quantified *Numb* and *vasa* expression using the HCR data generated. Thirdly, I investigated whether the honeybee contains a germ stem cell niche (as in *Drosophila*), or contains self-renewing cell clusters (as in *Nasonia*), as this will inform interpretation of the HCR data. Finally, I investigated the role of Notch signalling in the related species *Nasonia*, under the assumption that the role of Notch signalling in the honeybee germarium may be conserved in *Nasonia*.

4.1 Results

Notch may maintain the germ stem cell niche in honeybee ovary activation (Duncan et al. 2016). Describing a cellular resolution expression pattern of *Numb* would test this theory. Notch is involved in the germ stem cell niche, *Numb* localization to particular cell types should be detectable. HCR is an emerging technique to perform high-resolution imaging of multiple genes simultaneously (Choi et al. 2018). Therefore I used HCR to investigate expression of *Numb* and the germ cell marker *vasa*. 
4.1.1 Numb expression in the honeybee ovary

HCR has not yet been validated in the honeybee. To validate the sensitivity and specificity of HCR in the honeybee ovary, I investigated whether the technique could accurately reproduce existing expression patterns. As a positive control, I performed HCR against orthodenticle1 (otd1), a gene with well-characterised ovarian expression (Wilson & Dearden 2011). In developed oocytes, otd1 was expressed throughout the nurse cells and oocyte, as also reported by Wilson & Dearden (2011). Otd1 was also expressed in the gerarium, a pattern not previously reported (see Figure 4.1b-c). Both Numb and vasa expression recapitulated published gene expression domains in the vitellarium (see Figure 4.1d) (Dearden 2006; Duncan et al. 2016). Negative controls consisted of Queen ovarioles, treated as per usual HCR protocols, but without any initiator probes added. Autofluorescence of the ovary was detectable in the negative controls, as is common in fluorescent imaging (Lee & Kitaoka 2018), but none of the bright spots representing RNA transcripts were present (see Figure 4.1a). Therefore, HCR can robustly identify gene expression patterns in the honeybee.

![Figure 4.1: Positive and negative controls for the HCR experiment. a: negative control, with hairpin H1 (Alexa488, usually detecting Numb) and H2 (Alexa 546, detecting vasa) added without initiators. b: otd1 expression in the vitellarium. c: otd1 expression in the gerarium. d: vasa (magenta) and Numb (green) expression in the vitellarium. Scale bars represent 100um. In all images the terminal filament (anterior) is to the right, maturing oocytes (posterior) to the left. Blue represents DAPI staining, green represents numb, magenta vasa and red otd1.](image)

While vasa and Numb expression have been previously described in the honeybee,
their relative expression at the cellular level is unknown. Having demonstrated the efficacy of HCR in the honeybee, I used HCR to investigate the cell-level expression of Numb and vasa. In both the Queen and worker gerarium, vasa was expressed in three cell types: the cystocyte clusters of the anterior gerarium, the newly-specified oocyte, and the anterior nurse cells (Figure 4.2). Though vasa has previously been reported as only expressed in the worker gerarium (Dearden 2006), HCR detected clear vasa expression patterns in the Queen gerarium. This difference in expression patterns may have been due to increased sensitivity of the HCR probes; the expression patterns’ consistency and localization to particular cells implies it was not an artifact. Nuclear Numb and vasa expression was detectable in the terminal filament, meaning these genes were actively transcribed in these cells (see Figure 4.2). Therefore vasa appears to have a role in oocyte specification, and in the cystocyte clusters.

Numb was also similarly expressed in Queen and worker ovaries. Expression was high in the presumptive oocytes, and present in the cystocyte clusters (see Figure 4.2). Presumably, the honeybee germ stem cells exist in the gerarium anterior to oocyte differentiation. No localization of Numb to cell subtypes was seen in this region of the gerarium, meaning Numb probably does not control the honeybee germ stem cell niche. The magnitude of Numb expression appears to differ between Queens and inactive workers in the anterior gerarium, within the cystocyte clusters. Numb, therefore, appears to specify the honeybee oocyte, and may have its function in ovary activation in the anterior ovary.

4.1.2 Quantification of Numb expression

From these data, Numb appeared to be differentially expressed (between Queens and inactive workers) in the anterior, but not posterior, honeybee ovary. However visual inspection is neither a robust nor unbiased method of quantifying gene expression. As each dot of HCR fluorescence represents one RNA molecule, these
Figure 4.2: Numb and vasa expression in the honeybee Queen (top) and worker ovary. Scale bars represent 100um. In all images the terminal filament (anterior) is to the right, maturing oocytes (posterior) to the left. a-d: Queen ovaries. a: vitellarium. b: posterior germarium. c: middle germarium. d: anterior germarium. Images were background subtracted in ImageJ to improve clarity. Queen images are representative of >15 ovarioles from two individuals, over three biological replicates. Worker images are representative of >15 individuals over three replicates. oc: oocyte. nc: nurse cell.
dots can be counted to quantify gene expression (Shah et al. 2016; Daniel Green et al. 2018). This digital HCR (dHCR) method of quantifying gene expression differs from other types of quantitative imaging, where the image fluorescence is used as a proxy for gene/protein expression, and is a more direct measure. Therefore, to test whether Numb was truly differentially expressed in the anterior honeybee ovary, I wrote an ImageJ macro to detect and count RNA molecules in HCR images. Only inactive worker and Queen ovaries were quantified, because active worker ovaries have a disordered morphology and I was unable to obtain sufficient clear images for quantification. I performed the quantification on ovaries that had been incubated in hairpins overnight, rather than for the 1-2 hours previously used (Shah et al. 2016), as this improved the detectability of Numb expression.

I wrote an ImageJ macro\(^1\) to count RNA foci and quantify Numb and vasa expression (see Figure 4.3 for an illustration of the pipeline). This macro blurred and thresholded the vasa channel to generate a ‘mask’ that can be used to select the region of the image to quantify. To produce an image suitable for quantification, the image was blurred to remove noise using a Gaussian blur, the **background subtract** command used to remove large and small particles, and the image thresholded and segmented to produce a series of separated dots. Dots were counted using the **Analyze particles** function. In the posterior germarium, the region of the ovary to be quantified (between the beginning of the vasa expression domain and the beginning of the vitellarium) was manually defined. The pipeline did a reasonable, if imperfect job, of separating each dot (see Figure 4.3). The vasa channel in particular has less-well segmented dots.

In the anterior germarium, Numb, but not vasa, was differentially expressed between inactive workers and Queens. Neither gene was differentially expressed in the pos-

---

\(^1\)see https://github.com/Shannon-E-Taylor/dHCR/blob/master/scripts/dHCR.ijm; https://github.com/Shannon-E-Taylor/dHCR/blob/master/scripts/processing-functions.R is also useful for data processing
terior germarium (see Figure 4.4a-b). Presumably, then, *Numb* expression in the anterior, not posterior, ovary, was involved in honeybee ovary activation. Note that the DAPI (nuclei) channel was meaningless, as the nuclei are removed in the background subtraction step (see Figure 4.3 after thresholding), and was included as an example of nonsense data. There was a small fold change in *Numb* expression in the anterior: *Numb* was expressed 1.3 times higher in Queens as opposed to workers. This was surprising given that this change was detected in whole tissue using qPCR (Duncan et al. 2016). One possible explanation for this is the smaller detected *Numb*
Figure 4.4: Quantification of HCR data. All workers used for quantification were inactive, unpaired t-test. Error bars represent one standard deviation of the mean. a: Number of dots per micron squared in the posterior (more differentiated) ovary. Queens: n = 6 replicates (ovarioles from one individual). Workers: n = 8 replicates, mix of same and different individuals. b: Number of dots per micron squared in the anterior (less differentiated) ovary. Queens: n = 6 replicates (ovarioles from one individual). Workers: n = 5 replicates, mix of same and different individuals. c, d: average size of particles in the anterior and posterior ovary. e: Changes in number of particles counted with depth. Each panel represents a different ovariole. f: Percentage of the image covered by segmented dots in the anterior ovary.

particles in inactive worker ovaries (see Figure 4.4c). It is possible that small background particles are being detected by the pipeline, causing a smaller fold change to be detected than would be seen otherwise. As Numb also marks the oocyte, it
is also possible that increased oocyte numbers in activating workers numbers were detected up by qPCR.

Do other measures generated by the pipeline match expected values? The *vasa* particles were also larger than the *Numb* ones. This may be due to the longer wavelength used for imaging, which caused a larger diffraction limit, and resulted in larger dots detected. The poorer segmentation seen in the *vasa* channel may also have been contributing. Both the *vasa* and *Numb* particles detected were larger in the posterior germarium (see Figure 4.4c-d). This may also be due to poor segmentation by the pipeline, or may be due to different confocal settings causing larger fluorescent foci. (Though I was careful to image the anterior germarium using consistent confocal settings, I did not do this for the posterior germarium.) The theoretical diffraction limit, and therefore expected size of these particles, can be calculated using Abbe’s diffraction formula \( d = \frac{\lambda}{2 \times NA} \), where \( \lambda \) is the wavelength of light used for imaging, and \( NA \) is the refractive index of the microscope lens used (\( NA = 1.05; \) UPLSAPO30X 30X silicon oil lens) (Franklin et al. 2010). For the 488 (*Numb*) imaging channel, then, the theoretical RNA foci size was 0.042\( \mu m^2 \), and 0.057\( \mu m^2 \) for the 546 (*vasa*) channel. This is a ~10fold difference from the particle sizes seen, likely because the samples were not imaged under ideal conditions.

Fluorescence brightness decreases with tissue depth (Shah et al. 2016), which could bias dHCR quantification. Therefore, I investigated the change in number of foci detected in image stacks. In general, there was no change in foci detected with depth. However, in one image there was a dramatic difference, due to the pipeline identifying multiple small (one pixel) dots (see Figure 4.4d), presumably because of increased background. This demonstrated that while care must be taken to select a representative slice with low background when quantifying HCR, quantification will not be biased by tissue depth.

HCR data was inconsistent with the currently understood potential role of Notch in
honeybee ovary activation. *Numb* does not localize to any particular cell type in the anterior germarium, meaning *Numb* is presumably not expressed in the honeybee germ stem cells. Therefore *Numb* (and by extension Notch signalling) does not maintain the honeybee germ stem cell niche. In contrast, the broad expression of *Numb* throughout the cystocyte clusters of the germarium suggests that the gene may regulate this cell type. Clearer information about the location of the honeybee germ stem cell niche would help confirm this. I explore this next.

### 4.1.3 Cell division in the *Drosophila*, *Nasonia*, and honeybee ovary

The working hypothesis for this project was that Notch maintains the honeybee germ stem cell niche, as in *Drosophila* (Ting 2013). However, the location of this structure, and whether it exists in the honeybee, is unknown. In addition, it appears that neither *Nasonia* nor *Tribolium* contain a germ stem cell niche (Trauner & Büning 2007; Griebel & Rübsam 2014). As cystocyte clusters within the germarium divide at the same time, the number of dividing cells within each cluster reveals when self-renewal of undifferentiated clusters occurs (Griebel & Rübsam 2014). For example, an overrepresentation of one-cell clusters would imply that single cells are renewed to produce cystocytes. In addition, the location of single dividing cells could reveal the position of germ stem cell niche. Therefore, I stained for dividing cells in the honeybee, *Nasonia*, and *Drosophila* ovary using an antibody against phospho Histone H3 (pH3). To validate the conservation of the antibody, I also stained asexual aphid (*Acyrthosiphon pisum*) ovaries.

Immunohistochemistry against pH3 gave clear, unambiguous nuclear staining in *Nasonia*, *Drosophila*, and the aphid (see Figure 4.5a-c). Staining was detectable in both the germarium and follicle cells, but not the nurse cells, which do not divide, validating the antibody specificity. As the aphid is distantly related to the honeybee, immunogenicity to the antibody was well conserved, which is expected given how
Figure 4.5: Immunohistochemistry against pH3 in various insect species. Scale bars indicate 100um.

a-b: Single confocal sections of *Drosophila* and *Nasonia* ovaries. Arrow indicates dividing cell clusters.

c: single confocal section of an aphid embryo developing inside the ovary. Head is up, legs to the right.

d: Maximum intensity z-projections or epifluorescence of the honeybee germarium. Arrow indicates the start of the terminal filament. Autofluorescent foci are circled.

well conserved Histone H3 is (Hans & Dimitrov 2001). Cell division in the honeybee germarium was detected at two places: in cystocyte clusters in the germarium, and in the terminal filament (see Figure 4.5d). While the terminal filament staining was consistent within one technical replicate (consisting of several ovarioles), I was unable to reproduce this result; however, cystocyte staining was reproducible. Obtaining data for quantification of honeybee cell cluster division was challenging, as the immunohistochemistry worked inconsistently, and the size of honeybee germarium made collecting sufficient image data difficult. Therefore, I was unable to determine whether cell division in the honeybee ovary followed a more *Nasonia* or *Drosophila*-like mode, and thus whether there is any evidence for the existence of
germ stem cells in this species. I still analysed the cell cluster numbers in Naso-
nia and Drosophila, as the initial Nasonia study had not compared their data to
Drosophila. This was therefore a useful extension of the literature.

![Figure 4.6: Models and data for cell division in the Nasonia and Drosophila germarium. Top panel gives the proportion of cells in each cluster, for both Nasonia and Drosophila. Bottom panel depicts predicted cell division events in the Nasonia and Drosophila germarium. In the Drosophila germarium, a single germ stem cell self-renews, and can differentiate to produce a cystoblast. In Nasonia, there is hypothesised to be no germ stem cell, meaning single cells are dividing follicle cells, and four-cell clusterse self-renew to produce new oocytes."

The Nasonia germarium is much larger than that of Drosophila, and the transition
from germarium to vitellarium is not well defined. Many instances of single cells
dividing were detected in the posterior germarium. Due to their location, these were
likely to be follicle cells rather than germ cells, with the potential to bias my data.
Therefore, when quantifying dividing cell clusters, I excluded any single cells that
lay more posterior than a cell cluster, on the assumption that these cells were follicle
cells, rather than germ stem cells. Despite this exclusion, it is possible that follicle
cells were still counted, giving the relatively high number of single cells in Table 4.1.

To interpret the quantitative data, I compared the cell count data to the expected
Drosophila and Nasonia distributions. The expected Drosophila distribution was a 2:1:1:1 ratio of one:two:four:eight cell clusters (see Figure 4.6 for diagram). This distribution makes two assumptions: that the cell clusters take the same amount of time to divide, irrespective of size, and that there was one dividing germ stem cell for each dividing cystoblast (giving a 2:1 ratio of one:two cell clusters). I did not detect a single 16-cell cluster in Drosophila or Nasonia, so ignored 16-cell clusters. The Nasonia distribution was 1:0.1:1:1, assuming a 1:1 ratio of dividing follicle cell clusters detected to germ cell clusters, and a small proportion of 2-cell clusters. These values were chosen because they fit the data reasonably well; I did not experiment with these values when performing the statistical test.


<table>
<thead>
<tr>
<th>Number of cells in cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila</td>
<td>27</td>
<td>15</td>
<td>5</td>
<td>12</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nasonia</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Interpretation of the quantitative data was challenging, as some cell clusters had intermediate values outside the predicted 1, 2, 4, 8, 16-cell sequence. This was probably because the clusters were still dividing. It was unclear how these intermediate clusters should be grouped- whether three-cell clusters should be treated as two or four-cell clusters. As the pH3 antibody stained cells both before and after cell division, either scenario is biologically relevant (Hans & Dimitrov 2001). Therefore, I performed statistical tests on the data with intermediate values grouped both to the highest and lowest neighboring cluster values (see Table 4.2). Neither grouping resulted in a significant difference being detected between the raw Drosophila and Nasonia data (Pearson’s χ² test). However, when intermediate cell values were aggregated up, the data supported the models. The Nasonia model did not fit the Drosophila data, and vice versa, and there was insufficient evidence to reject the
models from the same species’ data. This trend did not hold when data was aggregated downwards, but as this analysis did not support the *Drosophila* model, this grouping is invalid. The *Drosophila* model, supported by vast experimental evidence, is unlikely to be wrong (Ting 2013). Another key piece of evidence in favour of the *Nasonia* model is the lack of 2-cell clusters detected in this species. As there are no 2-cell clusters in the *Nasonia* ovary, no 2-cell cystocyte clusters are being produced from the “germ stem cells”. Therefore, this analysis supports the model of Griebel & Rübsam (2014), and suggests that the *Nasonia* germarium does contain self-renewing 4-cell clusters instead of single germ stem cells. As the honeybee is much more closely related to *Nasonia* than *Drosophila*, it is more likely that the honeybee does not contain a germ stem cell niche, but this remains to be proven, using this method.

<table>
<thead>
<tr>
<th>Table 4.2: Results of Pearson’s $\chi^2$ statistical test on aggregated data. Data are presented to two significant figures. df: degrees of freedom. <em>Dm</em>: <em>Drosophila melanogaster</em>. <em>Nv</em>: <em>Nasonia vitripennis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aggregated up</strong></td>
</tr>
<tr>
<td>$Dm$ data vs $Dm$ model</td>
</tr>
<tr>
<td>$Nv$ data vs $Nv$ model</td>
</tr>
<tr>
<td>$Dm$ data vs $Nv$ model</td>
</tr>
<tr>
<td>$Nv$ data vs $Dm$ model</td>
</tr>
<tr>
<td>$Dm$ data vs $Nv$ data</td>
</tr>
</tbody>
</table>

### 4.1.4 Notch has multiple roles in the *Nasonia* ovary

The HCR data suggested that Notch signalling may prevent the differentiation of cystocyte clusters, thus repressing worker reproduction. As *Nasonia* is closely related to the honeybee, I reasoned that Notch signalling may play a similar role in *Nasonia*. If this is the case, Notch would be expressed in the region of the *Nasonia* germarium where self-renewing germ cell clusters exist. Indeed, I detected Notch intercellular
domain (NICD) expression in region 1a/b of the germarium, where cystocyte renewal occurs. In the first part of region 1c, where cystocytes differentiate, NICD expression was lost (see purple bar in Figure 4.7b). NICD expression returns later in the ovary. Thus, these data supported the hypothesis. NICD is also expressed in the pole cells, which are follicle cells at the termini of the oocyte (see Figure 4.7a). Therefore, Notch probably specifies the Nasonia pole cells, as in Drosophila (Xu & Gridley 2012). NICD also localized to the membrane between the follicle cells and oocytes, so likely maintains one of these cell populations. RNA interference (RNAi) to knock down Notch expression in Nasonia caused a thickening of the follicle cell layer (see Figure 4.7e), so Notch prevents division of the Nasonia follicle cells. This does not suggest a role for Notch in maintaining follicle cells in honeybee ovary activation, as this function occurred after the germarium. Notch knockdown ovaries also have mis-positioned oocytes, and nurse cells exhibit premature apoptosis (see Figure 4.7d). This premature apoptosis could be a response to remove malformed oocytes, or a direct effect of Notch. Fringe, a modulator of Notch signalling, was maternally provided to the oocyte, and was expressed throughout the germarium, suggesting it modulates Notch signalling in the germarium and early embryo (see Figure 4.7f-h).

The HCR data suggested that Notch signalling may prevent the differentiation of cystocyte clusters, thus repressing worker reproduction. As Nasonia is closely related to the honeybee, I reasoned that Notch signalling may play a similar role in Nasonia. If this is the case, Notch would be expressed in the region of the Nasonia germarium where self-renewing germ cell clusters exist. Indeed, I detected Notch intercellular domain (NICD) expression in region 1a/b of the germarium, where cystocyte renewal occurs. In the first part of region 1c, where cystocytes differentiate, NICD expression was lost (see purple bar in Figure 4.7b). NICD expression returns later in the ovary. Thus, these data supported the hypothesis. NICD is also expressed in the pole cells, which are follicle cells at the termini of the oocyte (see Figure 4.7a). Therefore, Notch probably specifies the pole cells, as in Drosophila (Xu & Gridley 2012).
NICD also localized to the membrane between the follicle cells and oocytes, so likely maintains one of these cell populations. RNA interference (RNAi) to knock down Notch expression in Nasonia caused a thickening of the follicle cell layer (see Figure 4.7e), so Notch prevents division of the Nasonia follicle cells. This does not suggest a role for Notch in maintaining follicle cells in honeybee ovary activation, as this function occurred after the germarium. Notch knockdown ovaries also have mis-positioned oocytes, and nurse cells exhibit premature apoptosis (see Figure 4.7d). This premature apoptosis could be a response to remove malformed oocytes, or a direct effect of Notch. Fringe, a modulator of Notch signalling, was maternally provided to the oocyte, and was expressed throughout the germarium, suggesting it modulates Notch signalling in the germarium and early embryo (see Figure 4.7f-h).

Unfortunately, the ovaries of the five-day old wasps imaged had a morphologically normal early vitellarium and germarium, possibly because they had begun to recover from RNAi, as occurs at this time (Lynch & Desplan 2006). Therefore, I was unable to investigate the dynamics of cell division in the Nasonia germarium after reduction in Notch signalling, which would have allowed me to better test the hypothesis about the role of Notch cystocyte self-renewal. Overall, these data indicate Notch plays multiple roles in Nasonia oogenesis, including preventing cystocyte differentiation, inhibiting follicle cell division, and oocyte positioning.

4.2 Discussion

4.2.1 The role of Notch in the honeybee and Nasonia ovary

Based on its expression pattern, Numb appears to have two roles in the honeybee ovary: oocyte specification in the posterior, and ovary activation in the anterior. As Numb is differentially expressed between Queens and workers in the anterior, but not posterior, germarium, it is unlikely that the role in oocyte specification is related to ovary activation. Indeed, Numb appears to specify the oocyte in the non-social
Figure 4.7: Role of components of the Notch signalling pathway in *Nasonia* ovaries. Scale bars represent 100um unless indicated otherwise. a-c: single confocal sections of *Nasonia* ovaries after RNAi. a: Knockdown of *egfp*, a jellyfish gene, results in normal ovary morphology in *Nasonia*. b-c: Notch knockdown ovarian phenotypes. Green represents autofluorescence in the green laser line, and is included because it makes apoptotic cells clearer. d and e: Immunohistochemistry against Notch (magenta), counterstained with DAPI (blue) and Phalloidin (green). Z-projections of maturing oocytes and the germarium, respectively. f-h: in situ hybridization against *fringe*. f: DIC image of maturing oocytes. g: DIC image of germarium. h: fringe sense probe negative control. fc: follicle cell. fs: fusome. nc: nurse cell. oc: oocyte. pc: pole cell.

*Drosophila melanogaster* (Jambor et al. 2015). In the anterior germarium, *Numb* was expressed broadly, both in cystocytes expressing *vasa* and in the surrounding cells. There was no localization to a specific cell population, suggesting that *Numb* and Notch signaling did not have a role in maintaining the GSC niche, as hypothesized. One possibility is that Notch signalling functions to hold cystocytes in an undifferentiated state, preventing their maturing into oocytes and thus repressing worker reproduction. To test this hypothesis, I investigated the role of Notch signalling in *Nasonia*, a solitary wasp related to the honeybee. Expression of Notch in the *Nasonia* ovary supported this model: Notch was expressed in the region of the *Nasonia* germarium where cystocytes self-renew, rather than differentiate; Notch
expression was completely absent where oocytes differentiate. Notch expression re-
turned later in the germarium, where it presumably has some other role in the ovary. 
Therefore the role of Notch signalling in the honeybee may be as follows. In the 
inactive worker ovary, active Notch signalling may maintain cystocyte clusters in 
the undifferentiated state. As oocytes cannot form, the worker honeybees cannot 
reproduce. During ovary activation, increasing Numb transcription causes degrada-
tion of the Notch receptor, and subsequent differentiation of the cystocyte clusters 
to produce maturing oocytes and nurse cells.

Without knocking down Notch in honeybees and Nasonia, and investigating the 
effect this has on cystocyte production and differentiation, it is impossible to confirm 
this hypothesis. The presumed role of Numb in honeybee ovary activation is also 
based on correlative gene expression studies, not functional data (Duncan et al. 
2016), so it is possible Notch plays a role in the germarium independent of Numb. 
Staining for Notch pathway genes (such as the E(spl)-C genes) using HCR would 
rule this possibility out. In addition, it is unclear whether the honeybee germarium 
contains a germ stem cell niche, self-renewing cell clusters, or some other method 
for producing new oocytes. When I stained Queen ovaries for dividing cells, clusters 
of dividing cells were clearly detectable in the anterior germarium, and staining was 
inconsistently detectable in the terminal filament. Honeybee germ stem cells have 
been proposed to reside in either the terminal filament or base of the honeybee 
gerarium, based on detection of cell division and vasa staining in those areas 
(Tanaka & Hartfelder 2004; Tanaka & Hartfelder 2009; Dearden 2006). Either 
of these locations remain plausible given these data. Alternatively, the honeybee 
may not have germ stem cells, and instead produce self-renewing cell clusters like 
Nasonia. Completing the pH3 staining in the honeybee would test this. It would 
also be interesting to extend this analysis to other species, such as the aphid, to 
investigate whether the existence of the germ stem cell niche is the ancestral state, 
and whether the terminal filament may contain germ stem cells in other species.
This hypothesis, that Notch maintains cystocyte clusters in the undifferentiated state, makes some sense in light of the existing literature. The inactive worker ovary contains several fusomes, structures that connect cystocytes in several insect species (Duncan et al. 2016; Schmidt Capella & Hartfelder 2002). This suggests that the worker ovary is repressed after cystocyte clusters form, not in the germ stem cell niche. The hypothesis also suggests a mechanism for rapid ovary activation. Worker honeybees could rapidly differentiate all of the cystocyte clusters into maturing oocytes, depleting this region of the germarium entirely. However, it is unclear how Queen ovaries, which lack Notch signalling, would maintain the cystocyte clusters required for five years of continuous oocyte production (the lifespan of the Queen). New cystocyte clusters could be produced from the terminal filament, or Notch may be selectively activated in a subset of cells. Nuclear Notch protein was detected in the nuclei of cells in the Queen germarium by Wilson & Dearden (2011), but not Duncan et al. (2016). Evolutionarily, this mechanism of repression would be evolve comparatively easily, as it relies on existing regulators of reproduction.

4.2.2 Notch regulates *Nasonia* follicle cells

I also found that Notch signalling regulated the *Nasonia* follicle cells. NICD is expressed at the membrane between the oocyte and the follicle cells, and in the *Nasonia* pole cells, suggesting Notch maintains either the follicle cell or oocyte. RNAi against *Notch* causes thickening of the follicle epithelium, which is normally a single cell layer. Therefore, Notch in *Nasonia* appears to inhibit follicle cell proliferation. Notch has multiple roles in regulating the follicle cells of other species. In *Drosophila*, loss of Notch signalling causes defects in follicle cell migration and replication (Xu & Gridley 2012). In contrast, Notch signalling maintains follicle cells in the undifferentiated state in both *Tribolium* and *Blattella germanica* (Bäumer et al. 2012; Irles et al. 2016); loss of Notch prevents cell division. Interestingly, the honeybee lacks Notch expression at the oocyte membrane; instead, Notch-positive cells appear
to be migrating posteriorly in this species (Wilson et al. 2011). Notch may play a different role in the honeybee follicle cells, or have a similar role with a distinct expression pattern. Knockdown of Notch in the honeybee would test this. In addition, staining for, and knocking down, further components of the Notch pathway in *Nasonia* would further test the role of Notch in the *Nasonia* follicle cells.

### 4.2.3 HCR is a powerful method to study gene expression

HCR answered new questions about the role of Notch, and other genes, in the honeybee ovary. This study concluded that *Numb* does not localize to one cell type in the anterior ovary. This would have been impossible to prove without the ability of HCR to investigate cell-level gene expression. Quantitative HCR was also successful, and revealed that *Numb* was differentially expressed in the anterior, not posterior, honeybee germarium. In addition, the HCR data solved a puzzle in the honeybee literature. *Vasa* has been shown to be differentially expressed between the worker and Queen ovary (Dearden 2006). This made little sense: why would a germ cell marker be expressed in the inactive worker, but not Queen, ovary? HCR was able to detect *vasa* expression in the Queen germarium, and dHCR revealed that *vasa* was expressed at similar levels in Queen and worker ovaries. This is a more reasonable finding, and illustrates some advantages of HCR over traditional ISH experiments. In ISH experiments, the enzymatic reaction is left to develop until the experimenter decides the expression pattern is most visible. This decision involves a trade-off between missing regions of low expression, and overdeveloping the and ruining experiment. In HCR, however, individual RNA molecules are detected as fluorescent foci (Choi et al. 2018). The density, not intensity, of expression, reveals the magnitude of gene expression. Single RNA molecules are detectable: for example I was able to detect transcription of *vasa* and *Numb* in the honey bee terminal filament. Therefore, regions of low expression are much more likely to be detected with HCR. In the *vasa* expression pattern of Dearden (2006), *vasa* was highly expressed in the
most developed oocytes, suggesting that the reaction was terminated prematurely.

This comparison suggests that HCR is more sensitive than conventional *in situ* hybridization. Indeed, Patriarchi et al. (2018) used HCR as a more sensitive technique for detecting gene expression. HCR is also higher resolution, as it is able to image at the single-cell level. This ability is of enormous assistance to developmental biology. We are beginning to recognize the heterogeneity of cell types within tissues. Single-cell RNAseq is allowing us to investigate gene expression at the single-cell level, but validation of these data at the single-cell level is also required (Ståhlberg & Kubista 2018). HCR is such a method, and has been used to validate single cell RNAseq data (Daniel Green et al. 2018). Though single-cell qPCR is possible, it involves manually dissecting the individual cells of interest, flow cytometry, or macroaspiration, and is therefore technically challenging (Ståhlberg & Kubista 2018). HCR is an improvement on each of these points. The use of HCR in quantitative imaging is also important. The technique has been used to investigate the dynamics of Notch ligand-activation in cell culture, for example (Nandagopal et al. 2018). Therefore, HCR is a powerful developing method to investigate cell signaling and gene expression at the single-cell level.

The HCR pipeline developed as part of this project was able to detect differential *Numb* expression, as expected from the literature. This strengthens the argument that HCR is useful for quantitative imaging. However, the pipeline could be improved, to more accurately detect single RNA molecules. Analyzing the detected particle sizes suggests that the pipeline does not correctly segment the *vasa* channel, and detects small background particles. This could be improved, by excluding small particles from counting, and by better segmenting the image. Performing segmentation on greyscale, not binarized, images may also improve results. To generalize the technique, it needs to be tested in different tissues and organisms, and compared to more well-established techniques such as qPCR. It would be very interesting to per-
form dHCR as a replacement for *in situ* hybridization when validating an RNAseq experiment, and compare this data to the qPCR validation. The technique also needs to be extended to other fluorophores: other experiments were not able to detect RNA foci in the Alexa 594 and 633 channels (data not shown), so this method may only work with the Alexa 488 and 546 fluorophores. I am aware of three examples of dHCR in the literature. Shah et al. (2016) initially validated the technique. Choi et al. (2018) use single-molecule imaging of dual-labelled genes to validate v3.0 of HCR, while Daniel Green et al. (2018) used dHCR, in conjunction with other techniques, to validate single-cell RNAseq data and map this back to the cellular context. Single molecule HCR has been used to detect bacterial variants in sputum from cystic fibrosis patients, but this was imaging entire bacteria, not single RNA molecules (DePas et al. 2016). Therefore, developing a dHCR pipeline is a useful contribution to the literature.

### 4.2.4 Technical limitations

Because of time constraints, I was not able to fully verify all the antibodies used and knockdowns performed. I did not validate the anti-pH3 antibody, as I could not find a method to inhibit cell division with the tools available, and because the staining was unambiguous, wholly nuclear, and conserved in the more diverged aphid. I also did not validate the anti-Notch antibody, again because of time constraints. Both pH3 and Notch are extremely well conserved, and these antibodies had been validated in related insects (*Drosophila* and honeybees, respectively), so the results from these techniques are still reliable. For some gene expression studies, I did not have time to perform technical replicates, so these experiments need to be repeated. However, all experiments were performed with at least five biological replicates (ovaries or ovarioles). I did not confirm *Notch* knockdown in *Nasonia* using qPCR, so cannot confirm the RNAi was specific to this gene.

To investigate the role of Notch signalling in the honeybee ovary, I performed three
key experiments. I obtained cell-level expression patterns for Numb in Queen and worker ovaries. This revealed that instead of regulating follicle cells or maintaining the germ stem cell niche, Notch signalling may hold cystocytes in the undifferentiated state, thus repressing worker reproduction. This hypothesis explains the observation that fusomes, markers of cystocyte clusters, exist in the inactive honeybee ovary (Duncan et al. 2016; Hartfelder et al. 2006). I obtained expression patterns of Notch in the related Nasonia that support this hypothesis. Notch was expressed in the region of the ovary where cystocytes self-renew, and not expressed where the cystocytes differentiate. I also discovered that Notch regulated the Nasonia follicles, and appeared to specify the pole cells. I showed that Nasonia do contain self-renewing four-cell clusters (not discrete germ stem cells). This supported and extended the data of Griebel & Rübsam (2014). This provies a mechanism by which honeybees could repress their ovaries, and therefore a mechanism of phenotypic plasticity.

4.3 Future Directions

Directions for future research in follow three broad categories: technical advancements to make this project easier, experiments to build on the results of this work, and new approaches to answering outstanding questions about honeybee ovary activation. I will first outline future directions in reference to regulation of Numb, and then as per Notch signalling in the honeybee.

4.3.1 Technical advances

CRE prediction is a challenging task. Without experimental validation it is impossible to be sure whether the predicted CREs are real, and where they are expressed in the honeybee. Better methods of CRE prediction, preferably ones that do not require knowledge of transcription factors operating in a particular context, would be of enormous help in predicting CREs in non-model and model organisms. Alongside
this, methods for validation of CREs in non-model organisms would also be useful: Drosophila have a vastly different life-history to the honeybee, so it is unclear how well the expression of the CRE transfers to this species. Indeed, a system to validate enhancers in the beetle Tribolium has been developed (Lai et al. 2018). Given the regulatory difficulties in producing transgenic honeybees, however, such technologies are unlikely to be available for this species in New Zealand (Cridge et al. 2017).

4.3.2 Extending this work

The expression of the CRE in the Drosophila gut suggests a connection between QMP repression and Notch signalling in the gut. This connection may extend to the honeybee. This would be interesting to follow up on, as outlined in Section 3.2, but this experiment first needs to be repeated in a more controlled manner and with a larger sample size.

A clear extension of this project is to follow up on the transcription factors predicted by CLARE. Which, if any, are truly involved in ovary activation? Do they act upstream or downstream of Notch? Are they involved in the initial switch between active and inactive ovaries, or do they act later in ovarian development? The ecdysone response elements Br-C and E74 are particularly interesting in this regard, as they have a known role in honeybee ovary development, and ecdysone signaling is important in ovary activation. It would also be interesting to analyse the function of mirror, as this gene is involved in integrating Notch signalling at multiple stages in the Drosophila ovary (Jordan et al. 2000). Brinker and hairy may also regulate Numb, as ClusterDraw predicted binding sites for these transcription factors within Numb, so are worth investigating as upstream regulators of Notch. Obtaining gene and protein expression patterns, and knocking down these genes in the honeybee, would give insight into these questions regarding ovary activation.

This project could also be extended by validating my dHCR pipeline for other tissues and genes. In particular, comparing HCR with more established methods such as
qPCR, and more intensive validation of the system, would allow this HCR system to be generalized to other biological contexts.

More work needs to be done to test the hypothesis that Notch signalling controls cystocyte differentiation in both *Nasonia* and the honeybee, with a view to publication. One key experiment to test this would be to knock down Notch signalling components in *Nasonia*, and investigate the effects on cell division dynamics in the germarium (via pH3 staining). If the hypothesis is correct, we would expect to see changes in the location of dividing cell populations. It would also be worth investigating the dynamics of cell division between inactive worker and Queen ovaries: if the Notch model is correct, there should be an over-representation of dividing four and eight-cell clusters in the inactive worker germarium, relative to Queens and active workers, due to (active) Notch preventing differentiation to 16-cell clusters.

We would also expect to see this a switch to queenlike-dynamics in early activation stages (stage 0 and stage 1), and in ovaries treated with DAPT, a γ-secretase inhibitor that blocks Notch signalling. A further experiment that could support this model would be immunohistochemistry against both Notch and pH3 in both *Nasonia* and the honeybee. If the model is correct, we would expect to see active (nuclear) Notch in some cells undergoing cell division, as Notch prevents these cells from differentiating. A more technically challenging experiment would be culturing labelled ovaries with and without DAPT: in *Nasonia*, DAPT treatment should initially cause an increase in the number of oocytes produced, as all cystocyte clusters differentiate, followed by a total loss of cystocyte clusters as the population cannot renew. In honeybees it is difficult to predict the outcome of this experiment, as whether cystocyte formation resembles *Drosophila* or *Nasonia* in this species is unknown.

Work is also required to determine the interactions between different signalling pathways, particularly apoptosis and Notch signalling, as these pathways interact in
mammals (Lundell 2003). In particular, whether apoptosis is acting downstream of Notch could be investigated by staining DAPT-treated ovaries for markers of apoptosis.

4.3.3 Outstanding Questions

This project was hindered by a lack of knowledge about the honeybee germ stem cell niche. Two key questions are: does the honeybee germ stem cell niche resemble that of *Nasonia* or *Drosophila*, and are the signalling pathways from *Drosophila* conserved in the honeybee? Recent work on insect segmentation sheds light on these questions. There are two key morphological methods of insect development: long-germ patterning, where the entire embryo is segmented virtually simultaneously, and short-germ patterning, where segments are progressively added via a posterior “growth zone”. These are radically different morphological processes, but transitions between the two are evolutionarily common (Clark & Peel 2018). *Nasonia*, for example, utilises long-germ patterning in the anterior and short-germ in the posterior (Rosenberg et al. 2014). However, it appears that the cell signalling pathway governing both these processes is identical between long and short-germ insects (Clark & Peel 2018).

If this finding is correct, and generalizable, it may be that during evolution, gene-regulatory networks are highly conserved, while morphological processes are not. Curiously, these different modes of development are associated with different ovary morphologies (Lynch & Roth 2011). Applying this theory to the insect germ stem cell niche suggests that the gene regulatory network (ie interactions between Notch and BMP signalling) will be conserved, even in the face of radically different germarium dynamics, such as self-renewing 4-cell clusters versus a self-renewing germ stem cell.

This theory would need to be tested in two ways. Firstly, by showing that the signalling pathways active in the *Drosophila* germarium are conserved in the honeybee, using multi-channel HCR against genetic markers of different cell types and cellular
processes. Secondly, single-cell RNAsq could be used to show these are the main
signalling pathways involved. The RNAsq data could then be validated by HCR.
This approach could also be used to determine the cell-signalling pathways, and cell
types, involved in ovary activation: once genetic markers of particular cell types are
well defined, single-cell RNAsq could be performed on worker bee ovaries in varying
stages of activation. Changes in the number of cell types, and signalling pathways
active within them, would allow us to describe the genetic interactions underlying
honeybee ovary activation.
Conclusions

In this project, I examined two questions. How is Numb regulated? What is the role of Notch signalling in honeybee ovary activation? To address the first question, I predicted a CRE within the Numb gene, which drove LacZ expression in the anterior and posterior Drosophila gut, and may be QMP-responsive. This suggests the gut may play a role in the Drosophila response to QMP. I also identified several transcription factors that may be involved in honeybee ovary activation; two of these, hairy and brinker may regulate Numb, and therefore represent upstream regulators of Notch signalling.

To address the second question, I showed that instead of regulating the honeybee germ stem cell niche, Notch signalling in the honeybee ovary may prevent differentiation of the cystocyte clusters. This is supported by data in Nasonia. In addition, I validated HCR in the honeybee, and provided evidence that this is a powerful method to quantitatively measure gene expression for both fundamental and applied biological research. Overall, my work suggests a mechanism for achieving phenotypic plasticity in the honeybee ovary, and by extension mechanisms of phenotypic plasticity operating in human health and disease.
## Appendix 1: Reagents

### 6.1 Reagent Recipes and lists

**Table 6.1: HCR reagents list**

<table>
<thead>
<tr>
<th>Name</th>
<th>Recipe/supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PBS</td>
<td>80 g of NaCl, 2.0 g of KCl, 14.4 g of Na₂HPO₄, 2.4 g of KH₂PO₄, in 10L dH₂O</td>
</tr>
<tr>
<td>PTw</td>
<td>100mL 1X PBS, 100μL Tween 20 (Roche)</td>
</tr>
<tr>
<td>Proteinase K</td>
<td>20 mg/mL stock (Roche)</td>
</tr>
<tr>
<td>Probe hybridization buffer</td>
<td>30% formamide, 5X sodium chloride sodium citrate (SSC), 9 mM citric acid (pH 6.0), 0.1% Tween 20 (Roche), 50μg/mL heparin (Sigma Aldrich), 1X Denhardt’s Solution, 10% Dextran sulfate</td>
</tr>
<tr>
<td>100X Denhardt’s solution</td>
<td>2% Bovine serum albumin, 2% Ficoll 400, 2% Polyvinylpyrrolidone, in H₂O</td>
</tr>
<tr>
<td>Probe wash buffer</td>
<td>30% formamide, 5X SSC, 9 mM citric acid (pH = 6.0), 0.1% Tween 20 (Roche), 5 μg/mL heparin (Sigma Aldrich)</td>
</tr>
<tr>
<td>5X SSCT</td>
<td>5X SSC, 0.1% Tween 20 (Sigma Aldrich)</td>
</tr>
<tr>
<td>Amplification Buffer</td>
<td>5X SSC, 0.1% Tween 20 (Sigma Aldrich), 10% Dextran Sulfate</td>
</tr>
<tr>
<td>DAPI</td>
<td>4,6-Diamidino-2-Phenyindole, dihydrochloride (Invitrogen)</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Invitrogen UltraPure</td>
</tr>
<tr>
<td>Initiators and Hairpins</td>
<td>Molecular Technologies (now Molecular Instruments, <a href="https://www.molecularinstruments.com/">https://www.molecularinstruments.com/</a>)</td>
</tr>
</tbody>
</table>
Table 6.2: *in situ* hybridization reagents list (see also Table 6.1)

<table>
<thead>
<tr>
<th>Name</th>
<th>Recipe/supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybridization solution</td>
<td>50% formamide, 4X SSC, 1X Denhardt’s reagent, 250 μg/mL tRNA, 250 μg/mL ssDNA, 50 μg/mL heparin, 0.1% Tween 20 (Roche), 5% Dextran Sulfate</td>
</tr>
<tr>
<td>PBTw</td>
<td>100μL Tween 20 (Roche), 0.1g bovine serum albumin in 100mL PBS</td>
</tr>
<tr>
<td>AP buffer</td>
<td>50 mM MgCl₂, 100 mM NaCl, 100 mM Tris pH = 9.5, 0.1% Tween 20 (Roche)</td>
</tr>
<tr>
<td>anti-DIG AP</td>
<td>anti-Digoxigenein-AP Fab fragments, Roche</td>
</tr>
<tr>
<td>Staining solution</td>
<td>4.5 μL/mL 4-nitro blue tetrazolium chloride (Roche), 3.5 μL/mL BCIP in AP buffer</td>
</tr>
</tbody>
</table>

Table 6.3: Immunohistochemistry reagents list

<table>
<thead>
<tr>
<th>Name</th>
<th>Recipe/supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBT</td>
<td>1X PBS (see 6.1), 0.05% Tween 20</td>
</tr>
<tr>
<td>PBTB</td>
<td>0.1 g BSA, 2.5 ml Normal Goat Serum, 2.5 ml 20X Sodium Azide stock, 45 ml PBT</td>
</tr>
<tr>
<td>DAPI</td>
<td>4,6-Diamidino-2-Phenylindole, dihydrochloride (Invitrogen)</td>
</tr>
<tr>
<td>Phalloidin-488</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Invitrogen UltraPure</td>
</tr>
<tr>
<td>Name</td>
<td>Recipe/supplier</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>Primary antibodies (NICD: 1 in 50, Mouse anti-pH3: 1 in 1000 - 1 in 10 000)</td>
<td>Mouse-anti pH3 (abcam 14955), The Notch antibody (C17.9C6) developed by S. Artavanis-Tsakonas was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biology, Iowa City, IA 52242, USA</td>
</tr>
<tr>
<td>Secondary antibodies (1 in 400)</td>
<td>Donkey anti mouse 555, Thermo fisher</td>
</tr>
</tbody>
</table>

Table 6.4: Cloning reagents list

<table>
<thead>
<tr>
<th>Name</th>
<th>Recipe/supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xba1</td>
<td>BioLabs</td>
</tr>
<tr>
<td>BamH1</td>
<td>BioLabs</td>
</tr>
<tr>
<td>10x digestion buffer</td>
<td>BioLabs</td>
</tr>
<tr>
<td>Agarose</td>
<td>Applichem</td>
</tr>
<tr>
<td>Ethanol</td>
<td>Bulk, from Scharlab</td>
</tr>
<tr>
<td>3M Sodium Acetate</td>
<td>pH = 5.2</td>
</tr>
<tr>
<td>Ligase buffer</td>
<td>Roche</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>Roche</td>
</tr>
<tr>
<td>dNTPs</td>
<td>BioLabs</td>
</tr>
<tr>
<td>PCR buffer</td>
<td>BioLabs</td>
</tr>
<tr>
<td>E coli buffer</td>
<td>XKI-Blue</td>
</tr>
<tr>
<td>LB</td>
<td>10 g peptone, 5 g yeast, 10 g NaCl, in 1 L H2O</td>
</tr>
<tr>
<td>LB plates</td>
<td>LB, but with 15g agar/mL and 5mg/mL ampicillin</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>50 mg/mL (Sigma)</td>
</tr>
</tbody>
</table>

Table 6.5: LacZ staining reagents list

<table>
<thead>
<tr>
<th>Name</th>
<th>Recipe/supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple juice plates</td>
<td>18g agar, 20g sucrose, 20mL 20% nipagen, in 200mL apple juice</td>
</tr>
<tr>
<td>Yeast paste</td>
<td>Yeast mixed with water until thick</td>
</tr>
<tr>
<td>LacZ staining solution</td>
<td>50 mM ferricyanide, 50 mM ferrocyanide, in PBS</td>
</tr>
<tr>
<td>Name</td>
<td>Recipe/supplier</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>QMP</td>
<td>Intko Supply, Vancouver, Canada</td>
</tr>
</tbody>
</table>

### 6.2 Equipment

- FV1000 Olympus confocal microscope
- UPLSAPO30X 30X silicon oil lens
- BX61 Olympus light microscope
- Narishige PC-10 needle puller
- PLI-100 microinjection apparatus (Harvard Apparatus)

### 6.3 CRE sequence and vectors

```
TCTAGACTGAATGTATTTTCGAATGTAAATAAAAATTTTCA
TAAATGTTTTCAATCTGTCCGATGGATGGAGAATTGTATT
GGCAATTATTTCGATCAAATTGAACGCGAAGCCAGAAAAGC
GACCTCTTTGCATTAAACTCTTTCGTTAGATGTCCAAAAATGC
GGCTCGTTTTTTTTTTTCATTCAAAGCAGCATCGAAATTCAA
CGTTTGCGCCTTACGACTACCAACGCACGCGACGAGG
CGCAATGTAACCGTGACCAGGCTCTCCGAGTGTGTGTAAC
GTGTAACGTGTGCAGACACGTACAATTTCACCCCCGATAAA
ACGTTCCCATAGAACGGTGCGCGCCACCGGGGTGTCTCCGGG
AACGGTTTCACCCGATATATAAAGCATAACGTACGGCATTT
TTATGGGGCCGCTAGCAACTGCAAGTGTTGGGGATTT
CTTTGTAATGAACTCGACAGATTCTTCTCCGCGACTTTCCC
CCAAATTTTTCCCTCTTGTGCTTCTCCGGACCTTTCCC
TAACCCGGGTTGTTGTGTTGGACCCGAAAGTAGAAGTGAAGATTT
GATTATCAAGCTATTTGGCTATTAGGTGTTTGGGAAATTTT
TAGAAAAGAAATTTTGGAGAAATTTATTTATTTTTCAA
ACTTTTATTTACAAGTAGATAGAGATTGATTGTGTATTATGT
TGAAAAAAGTAGTGAAGAGTGGCCCTCCT
```
Insulated eGFP / lacZ Drosophila Transformation Vectors for Promoter/Enhancer Analysis
Annotated vector sequences: GenBank Accession Numbers AF242360-65.

Figure 6.1: pH-Pelican vector. Vector used in this project lies second from top. (Barolo et al. 2000)

Figure 6.2: pUC57 vector.


Daniel Green, C. et al., 2018. A comprehensive roadmap of murine spermatogenesis defined by


Franklin, K. et al., 2010. *Introduction to biological physics for the health and life sciences*, John Wiley & Sons.


Hsu, H.-J. & Drummond-Barbosa, D., 2009. Insulin levels control female germline stem cell main-


Lovegrove, M., 2013. Genome architecture and phenotypic plasticity: is the Lethal (2) Essential for Life cluster epigenetically regulated during ovary activation in the honeybee, Apis mellifera? Bachelor of Biomedical Science with Honours. University of Otago.


Veraksa, A., Kennison, J. & McGinnis, W., 2002. DEAF-1 function is essential for the early


