NEUROGENOMIC DYNAMICS FOLLOWING SOCIAL INTERACTIONS IN MALE THREESPINED STICKLEBACKS

BY

SYED ABBAS BUKHARI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Informatics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Professor Alison M Bell, Chair
Professor Lisa Stubbs
Professor Saurabh Sinha
Professor Dave Zhao

ABSTRACT

Social interactions provoke changes in brain and behavior, however molecular changes associated with social interactions remain obscure. This thesis explored the neurogenomic responses to aggressive and affiliative social interactions in male threespined sticklebacks (*Gasterosteus aculeatus*), a small fish famous for their rich behavioral repertoire. In chapter one I provided the detailed overview of the research included in this dissertation. The second chapter tested the hypothesis that there are conserved transcriptional responses to social interactions in sticklebacks and fruit fly. There was stronger evidence for this hypothesis for one type of social behavior – a territorial challenge – than for a social interaction at the opposite end of the continuum: a courtship opportunity. In chapter three and chapter four I tracked the temporal dynamics of neurogenomic plasticity in male sticklebacks. I focused on two brains regions (diencephalon and telencephalon), which contain several nuclei of the social decision-making network. The third chapter focused on the transcriptomic and epigenomic responses to a territorial challenge. Results showed that the genome dynamically responds to a territorial challenge, with waves of transcription associated with different functions, e.g. hormone activity and immune response. The fourth chapter focused on males' transition to fatherhood, and compared and contrasted the neurogenomics of paternal care with the neurogenomics of the response to a territorial challenge. Males experienced dramatic neurogenomic shifts while they were providing paternal care. Genes related to hormones that change in mammalian mothers during pregnancy and maternal care, were differentially expressed in stickleback fathers. Gene regulatory analysis suggested that shared regulators were responsive to both a territorial challenge and paternal care and these were regulated differently along with their targets. This analysis offers a glimpse into how genes differentially acting within the social decision-making

network in the brain can generate responses to opposing social stimuli. Altogether, this thesis adds to the growing repertoire of studies examining social behavior at the molecular level and draws attention to the neurogenomic dynamics associated with behavioral plasticity.

ACKNOWLEDGEMENTS

This thesis wouldn't have been possible without the support, encouragement, and guidance of my adviser, Alison Bell. She helped me forge my abilities to bridge between computational biology and ecology. I am grateful for both research and collaborative opportunities she provided me along the way.

I am indebted to my committee members, Lisa Stubbs, Dave Zhao and Saurabh Sinha, who helped me develop scientific intellect and pushed me to the best of my abilities to find the most suitable solutions. I would like to thank Michael Saul, who helped me with statistical computing in R, and useful weekly discussions and feedback. I am grateful for the members of the Bell lab who critically evaluated my work and provided me with feedback.

I am thankful to my parents for their unconditional love and for being the constant source of encouragement and inspiration. I thank my siblings: Afshan, Hassan, and Saddam who always supported me and were there for me whenever I needed them. I thank my wife Bushra and my son Wali for being a loyal companion and source of entertainment. I also thank the friends I have made in Urbana Champaign with whom I have enjoyed both cricket and squash. Finally, I am grateful to Allah who always showed me the right path and guided me in my hour of need.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION.. 1

CHAPTER 2: COMMON NEURO-TRANSCRIPTOMIC RESPONSES TO SOCIAL STIMULI IN A FISH AND AN INSECT 11

CHAPTER 3: TEMPORAL DYNAMICS OF NEUROGENOMIC PLASTICITY IN RESPONSE TO SOCIAL INTERACTIONS IN MALE THREESPINED STICKLEBACKS... 39

CHAPTER 4: NEUROGENOMIC INSIGHTS INTO PATERNAL CARE AND ITS CONNECTION TO TERRITORIAL AGGRESSION 80

APPENDIX A: SUPPLEMENTARY INFORMATION FOR CHAPTER 2...................... 109

APPENDIX B: SUPPLEMENTARY INFORMATION FOR CHAPTER 3...................... 112

APPENDIX C: DIFFERENTIALLY EXPRESSED GENES FOR CHAPTER 4.............. 115

APPENDIX D: SUPPLEMENTARY INFORMATION FOR CHAPTER 4...................... 116

CHAPTER 1: INTRODUCTION

PROBLEM STATEMENT

Social behavior is complex – it reflects genetic and environmental sources of variation, it varies within and among species and it involves dynamic interactions among conspecifics. Social behaviors like mating, parental care and territorial aggression are crucial for survival and reproductive success. Social behavior varies along a continuum ranging from strongly affiliative, pro-social behaviors such as mating and parental care to strongly aggressive behaviors such as territorial defense. At the brain level, these social interactions are resolved by the social decision-making network, where different gene pathways underlying the neural circuits integrate prior experiences with currents physiological state to generate a behavioral response (O'Connell & Hofmann 2011a).

Until recently, the molecular basis of social behavior has been difficult to dissect. However, genomic tools have offered new opportunities to tackle the complexities of social behavior. For example, studies have attempted to find genetic variants contributing to variation in social behavior via quantitative trait locus mapping (Flint 2003; Greenwood et al. 2015; Southey et al. 2016; Giray et al. 2000). Other studies have approached the molecular basis of social behavior by studying the dynamic, environmentally sensitive side of the genome by quantifying gene expression on a genome-wide scale following social interactions (e.g., Robinson 2002, 2004; Whitfield et al. 2002; Robinson et al. 2005, 2008; Cash et al. 2005; Smith et al. 2008; Liang et al. 2012; Greenwood et al.; Sanogo et al. 2011) or between different behavioral types of individuals (e.g., Bell et al. 2016).

This thesis explores the neurogenomic responses to aggressive and affiliative social interactions in male threespined sticklebacks (*Gasterosteus aculeatus*), a small fish famous for

their rich behavioral repertoire (Pelkwijk & Tinbergen 1937). Chapter 2 explores the extent to which transcriptomic responses to social stimuli are conserved in two species from highly diverged lineages: threespined sticklebacks and fruit fry (*Drosophila melanogaster*). Chapters 3 and 4 reveal the temporal dynamics of neurogenomic plasticity in response to a territorial challenge and paternal care, respectively in male sticklebacks.

MODEL ORGANISM

Threespined sticklebacks are a model organism for studies of evolution and behavior. Renowned for their phenotypic variation within (Huntingford & Turner 1987) and among (Bell & Foster 1994) populations, sticklebacks have been dubbed a "supermodel" for studying the genetic basis of adaptation (Gibson 2005). Sticklebacks have also been a favorite subject for studying behavior dating back to the founders of the scientific study of animal behavior (ethology), including Niko Tinbergen (Pelkwijk & Tinbergen 1937). Part of what makes sticklebacks so popular is their readily-quantified and conspicuous social behaviors such as territorial aggression, courtship and paternal care, all of which can be studied both in the field and in the lab.

For male sticklebacks, the territory is hub of family life. The territory is where a father constructs his nest, attracts females with courtship behaviors and where he rears his offspring. Males must constantly defend their nest from territorial intruders, some of which are conspecific males seeking to usurp the territory or cannibalize the eggs, while others are heterospecific predators. Therefore, successful territory defense is critical for males' reproductive success.

In addition to managing the boundaries of their territory, male sticklebacks spend much of the breeding season engaged in paternal care, which is necessary for the survival of their

offspring. Males undergo a series of stages as they become fathers, where each stage depends on successful completion of the previous stage. After establishing a territory, males construct a nest; nest completion is marked by an overt “creeping through” behavior where the male burrows a tunnel through the nest. Only after completing the nest will males begin to court females and attract them to lay eggs in the nest. After a female swim through the nest and deposits her eggs, the male fertilizes the eggs, marking the end of the “nest” phase and the beginning of the “egg care” stage. Males provide care for the developing embryos in the form of direct paternal behaviors including fanning and tending the embryos as well as nest defense. The eggs hatch over the course of the fifth day after they are fertilized, and males continue to provide care for their newly-hatched fry for approximately one week.

THESIS ORGANIZATION

This thesis has been organized into three chapters. In chapter 2, I compared the neurotranscriptomic responses to a territorial challenge and a courtship opportunity between stickleback and fruit fly, to test the hypothesis that there are conserved roots of socially and ecologically comparable social behaviors in a vertebrate and in an insect (the “toolkit” hypothesis (Toth & Robinson 2007; Rittschof et al. 2014). There was stronger evidence for the toolkit hypothesis for one type of social behavior – a territorial challenge – than for a social interaction at the opposite end of the continuum: a courtship opportunity. Findings from this chapter shed new light on toolkit genes for social behavior and suggest that the importance of toolkit genes can vary between social contexts.

In chapter 3, I tracked the neurogenomic changes over time due to territorial challenge in male sticklebacks using RNA-Seq. Both transcriptomics and epigenomic changes were

quantified following a territorial challenge in two brain areas (diencephalon and telencephalon). Gene expression differences were measured 30, 60 and 120 minutes following a 5-minute intrusion. Differences in chromatin accessibility were measured using H3K27Ac as a marker 30 minutes and 120 minutes following an intrusion in diencephalon only. By integrating the time course gene expression data with a transcriptional regulatory network and changes in chromatin accessibility, I identified transcription factors that are predicted to coordinate waves of transcription associated with different components of behavioral plasticity. This study revealed rapid and dramatic epigenomic plasticity in response to a brief, highly consequential social interaction and has been published (Bukhari et al. 2017).

In chapter 4, I tracked the neurogenomic landscape of male sticklebacks as they transitioned to fatherhood. In addition to characterizing the neurogenomic landscape of fathers, an additional goal of this study was to compare and contrast the neurogenomics of paternal care with the neurogenomic response to a territorial challenge (Bukhari et al. 2017). I compared the brain gene expression profiles of males before, during and after they became parents, relative to the appropriate control. Gene expression was measured in two brain areas (diencephalon and telencephalon) using RNA-Seq. Males experienced huge neurogenomic shifts as they became fathers. Genes related to hormones that change in mammalian mothers during pregnancy and maternal care were differentially expressed in stickleback fathers. Overall, gene expression continuity and specificity was observed across stages, some of which might be analogous to changes associated with female pregnancy, parturition and postpartum periods. Gene regulatory analysis suggested that shared regulators are responsive to a territorial challenge and paternal care and these are being regulated differently along with their targets. This analysis offers a

glimpse into how genes acting within neural circuits of the social decision-making network in the brain can generate responses to opposing social stimuli.

CONCLUSIONS

Altogether this work highlights the insights that can be gained by examining behavioral plasticity at the molecular level. Chapter 2 shows that comparing the molecular mechanisms responsive to social stimuli in different organisms can offer insights into the ways in which behavior evolves, i.e. the extent to which similar genes are independently recruited to solve similar problems in highly diverged different organisms. Chapter 3 reveals heretofore underappreciated neurogenomic dynamism following social interactions. Whereas the literature to date has measured gene expression at one time point following a behavioral interaction, my results show that static experiments that measure gene expression at a single time point are likely to only catch a glimpse of what is a very dynamic and coordinated process. My study provides support for the hypothesis that there are waves of transcription associated with perceiving social information, responding to social information, maintaining a behavioral response, recovering from the social interaction and modifying future behavior. Chapter 4 illustrates the ways in which studying paternal care at the molecular level offers insights into the evolution of parental care. I show that the molecular building blocks of maternal care in mammals are operational in a fathering fish. These results suggest that (1) maternal care in mammals is a derivation of an ancient and highly conserved process; (2) there are deep commonalities between maternal and paternal care. Finally, I show that studying behavior at the molecular level offers a way to gain new insights into the modularity of behavior: two opposing social behaviors provoked opposite responses at the gene regulatory level, which suggests that the two behaviors are mechanistically

linked but that there are gene regulatory mechanisms operating within neural networks in the brain that facilitate their flexibility.

LITERATURE CITED

Bell, A.M., Bukhari, S.A. & Sanogo, Y.O., 2016. Natural variation in brain gene expression profiles of aggressive and nonaggressive individual sticklebacks. *Behaviour*, 153(13–14), pp.1723–1743. Available at: http://booksandjournals.brillonline.com/content/journals/10.1163/1568539x-00003393 [Accessed November 11, 2016].

Bell, M.A. & Foster, S.A., 1994. *The evolutionary biology of the threespine stickleback*, Oxford University Press. Available at: https://global.oup.com/academic/product/the-evolutionary-biology-of-the-threespine-stickleback-9780198577287?cc=us&lang=en& [Accessed May 25, 2018].

Bukhari, S.A. et al., 2017. Temporal dynamics of neurogenomic plasticity in response to social interactions in male threespined sticklebacks D. Clayton, ed. *PLOS Genetics*, 13(7), p.e1006840. Available at: http://dx.plos.org/10.1371/journal.pgen.1006840 [Accessed May 20, 2018].

Cash, A.C. et al., 2005. Behavior and the limits of genomic plasticity: power and replicability in microarray analysis of honeybee brains. *Genes, Brain and Behavior*, 4(4), pp.267–271. Available at: http://doi.wiley.com/10.1111/j.1601-183X.2005.00131.x [Accessed November 11, 2016].

Flint, J., 2003. Analysis of quantitative trait loci that influence animal behavior. *Journal of Neurobiology*, 54(1), pp.46–77. Available at: http://doi.wiley.com/10.1002/neu.10161 [Accessed November 11, 2016].

Gibson, G., 2005. Evolution. The synthesis and evolution of a supermodel. *Science (New York, N.Y.)*, 307(5717), pp.1890–1. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15790836

[Accessed May 25, 2018].

Giray, T. et al., 2000. Genetic variation in worker temporal polyethism and colony defensiveness in the honey bee, Apis mellifera. *Behavioral Ecology*, 11(1), pp.44–55. Available at: http://www.beheco.oupjournals.org/cgi/doi/10.1093/beheco/11.1.44 [Accessed November 11, 2016].

Greenwood, A.K. et al., 2015. Genetic Mapping of Natural Variation in Schooling Tendency in the Threespine Stickleback. *G3; Genes|Genomes|Genetics*, 5(5), pp.761–769. Available at: http://g3journal.org/cgi/doi/10.1534/g3.114.016519 [Accessed November 11, 2016].

Huntingford, F.A. & Turner, A.K., 1987. Conflict in the animal world. In *Animal Conflict*. Dordrecht: Springer Netherlands, pp. 3–12. Available at: http://www.springerlink.com/index/10.1007/978-94-009-3145-9_1 [Accessed October 11, 2016].

Liang, Z.S. et al., 2012. Molecular Determinants of Scouting Behavior in Honey Bees. *Science*, 335(6073).

O'Connell, L.A. & Hofmann, H.A., 2011. The Vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *Journal of Comparative Neurology*, 519(18), pp.3599–3639.

Pelkwijk, J.J. & Tinbergen, N., 1937. Eine reizbiologische Analyse einiger Verhaltensweisen von Gasterosteus aculeatus L. *Zeitschrift für Tierpsychologie*, 1(3), pp.193–200. Available at: http://doi.wiley.com/10.1111/j.1439-0310.1937.tb01422.x [Accessed October 11, 2016].

Rittschof, C.C. et al., 2014. Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proceedings of the National Academy of Sciences of the United States of America*, 111(50), pp.17929–34. Available at:

http://www.pnas.org/content/111/50/17929.full.

Robinson, G.E., 2002. Sociogenomics Takes Flight. *Science*, 297(5579).

Robinson, G.E., Fernald, R.D. & Clayton, D.F., 2008. Genes and Social Behavior. *Science*, 322(5903), pp.896–900. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1159277 [Accessed March 27, 2016].

Robinson, G.E., Grozinger, C.M. & Whitfield, C.W., 2005. Sociogenomics: social life in molecular terms. *Nature Reviews Genetics*, 6(4), pp.257–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15761469.

Robinson, G.F., 2004. Beyond Nature and Nurture. *Science*, 304, pp.397–398.

Sanogo, Y.O. et al., 2011. Brain transcriptomic response of threespine sticklebacks to cues of a predator. *Brain, behavior and evolution*, 77(4), pp.270–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21677424 [Accessed October 10, 2016].

Smith, C.R. et al., 2008. Genetic and genomic analyses of the division of labour in insect societies. *Nature Reviews Genetics*, 9(10), pp.735–748. Available at: http://www.nature.com/doifinder/10.1038/nrg2429 [Accessed November 11, 2016].

Southey, B.R. et al., 2016. Characterization of Genomic Variants Associated with Scout and Recruit Behavioral Castes in Honey Bees Using Whole-Genome Sequencing. PloS one, 11(1), p.e0146430

Toth, A.L. & Robinson, G.E., 2007. Evo-devo and the evolution of social behavior. *Trends in Genetics*, 23(7), pp.334–341. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0168952507001552 [Accessed March 27, 2016].

Whitfield, C.W. et al., 2002. Annotated expressed sequence tags and cDNA microarrays for

studies of brain and behavior in the honey bee. *Genome research*, 12(4), pp.555–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11932240 [Accessed November 11, 2016].

CHAPTER 2: COMMON NEURO-TRANSCRIPTOMIC RESPONSES TO SOCIAL STIMULI IN A FISH AND AN INSECT

ABSTRACT

It is possible that complex behaviors might evolve using the same molecular 'toolkit' in different species. Here, we test this hypothesis by comparing the neurogenomic responses to social stimuli between a fish (sticklebacks, *Gasterosteus aculeatus*) and an insect (fruit fly, *Drosophila melanogaster*). Specifically, we use published data on male brain gene expression profiles in response to a territorial challenge by a rival male and a courtship opportunity with a potential mate in both species. Using the PANTHER orthology map and Monte Carlo random resamplings, we found that there were gene families and subfamilies which were responsive to a territorial challenge in both species. Functional enrichment of common family and subfamily genes suggests that metabolic processes such as carboxylase and oxidoreductase activities are core and conserved processes involved in social behavior. We also report overrepresentation of racemase and epimerase activity genes in response to a territorial challenge in both species, which suggests a common mechanism of D-amino acid modulation related to aggression. We found less evidence for conserved responses to a courtship opportunity: there were gene families, but not gene subfamilies, which were responsive to a courtship opportunity in both species. Altogether, these results provide partial support for the hypothesis that there is a molecular 'toolkit' for social behavior that is conserved between invertebrates and vertebrates.

INTRODUCTION

Despite differences among species in brain structure, endocrine systems, behavior and ecology, it has recently been proposed that there might be a common set of molecular mechanisms involved in regulating social behaviors across species, or a molecular toolkit for social behavior (Toth & Robinson 2007; Rittschof & Robinson 2014). According to this hypothesis, certain genes are used and reused during social evolution, much in the same way that morphological traits have convergently evolved via the same molecular mechanisms along different lineages (Toth & Robinson 2007; Stern 2013; Rittschof & Robinson 2014). Support for the toolkit hypothesis has the potential to help unravel the molecular underpinnings of social behaviors and can help understand the evolution of complex phenotypes (Toth et al. 2010).

The hypothesis that there is a molecular 'toolkit' for social behavior was originally proposed and tested in eusocial insects (Toth & Robinson 2007; Toth et al. 2014; Toth et al. 2010; Woodard et al. 2014; Woodard et al. 2011; Mikheyev & Linksvayer 2015; Berens et al. 2015) has since been applied over much greater evolutionary distances (Rittschof et al. 2014). Rittschof et al. (2014) mined for shared neuromolecular mechanisms underlying the behavioral response to a territorial challenge across distantly related species (honey bee, mouse and sticklebacks). They found support for the toolkit hypothesis when they identified a set of conserved metabolic genes and homologous transcription factors in response to a territorial challenge.

However, we know little about whether the toolkit hypothesis applies equally to different types of social behavior or if the molecular basis for some social behaviors is more conserved than for others. For example, perhaps there is a molecular toolkit for responding to a social threat by an intruder (aggression), but not for responding to a mating opportunity (courtship). Recent

studies in several species have shown that both a territorial challenge by a conspecific and a mating opportunity provoke the differential expression of hundreds of genes in the brain (Sanogo et al. 2012; Sanogo & Bell 2016; Zayed & Robinson 2012; Ellis & Carney 2011; Cummings et al. 2008a). Therefore we now have an opportunity to compare support for the toolkit hypothesis across two social contexts. Specifically, two independent but similar experiments quantified neurogenomic changes in response to a territorial challenge and a courtship opportunity in fruit flies (*Drosophila melanogaster*) and three-spined stickleback fish (*Gasterosteus aculeatus*). Experiments in both species involved presenting a focal male with either a male or female conspecific for 5-10 min and then quantifying changes in the focal male's brain transcriptome within 30-45 min relative to an appropriate control. The use of genome wide microarrays provides us with an opportunity to test the 'toolkit genes' hypothesis for aggression and courtship in these two highly diverged species (~570 Mya) (Peterson et al. 2004).

In both stickleback and fruit fly, males defend territories and exhibit territorial aggression toward intruders. In contrast to fruit flies, where males defend mating territories but not not provide parental care (Zwarts et al. 2012; MØller & Thornhill 1998) male sticklebacks defend nesting territories and provide sole paternal care that is necessary for offspring survival (Wootton 1984). Males in both species exhibit courtship behavior toward potential mates, including wing shaking and body wagging (fruit fly) and a conspicuous courtship dance (zig zag, sticklebacks) (Wootton 1984).

Here using the PANTHER orthology map and Monte Carlo random resamplings. we ask if there are common genomic and neuromolecular mechanisms involved in the response to a territorial challenge and a courtship opportunity across both species. Specifically, within each social context, we mine for common gene families and gene subfamilies in both species.

METHODS

Previous studies have measured transcriptomic responses to social challenges and opportunities in male fruit flies (Ellis & Carney 2011) and sticklebacks (Sanogo et al. 2012). Here, we mine these published data to look for commonalities between the two species. As the experimental methods are previously published, they are briefly summarized here. To measure the response to a territorial challenge, males were confronted by another male. To measure the response to a courtship opportunity, males were presented with a receptive female. In all experiments, the social stimulus was briefly presented (15-20 mins, control: no stimulus) and males were sacrificed within an hour for brain gene expression analysis using microarrays. An overview of methodology is sketched in Figure 2.1.

Microarray statistical analysis

We used the LIMMA package in R Bioconductor to analyze both the stickleback and fruit fly microarray datasets (Smyth 2005). For stickleback, we used separate channel analysis in LIMMA, which is suitable for two-dye Agilent microarrays. We used loess normalization to correct gene expression intensities with a background correction offset of 50. For the between array normalization, we used the Aquantile normalization. A linear model was fit to the data that took into consideration the effect of treatment, dye, fish, and brain region. For fruit fly, we used the LIMMA protocol for Affymetrix microarrays. A similar linear model was fit which allowed the following contrasts: territorial males vs control males and courting males vs control males.

We focus our analyses on the sets of genes that were differentially expressed at raw p-values <0.05. Although this threshold is looser than the FDR <0.05 criteria, we assume that most false positives will be filtered by the cross-species comparison (similar to Rittschof et al. 2014).

We took the average of multiple probes for same transcript. Each transcript was assigned a gene ID.

The fruit fly microarray experiments were conducted on whole head tissue while the stickleback experiments were conducted on four brain regions separately. In order to compare across species, we took the union of the differentially expressed genes across the four brain regions in sticklebacks.

Altogether we generated lists of differentially expressed genes in four different conditions, which are hereafter referred to as "*stickleback territorial challenge*", "*fruit fly territorial challenge*, "*stickleback courtship opportunity*" and "*fruit fly courtship opportunity.*"

PANTHER annotation assignments and functional enrichment test

The PANTHER HMM scoring tool was used with PANTHER HMM library v 9.0 to assign a PANTHER gene family or subfamily ID to each gene in the stickleback and fruit fly genomes. This resulted in the creation of PANTHER Generic Mapping Files with two columns "gene id" and "best PANTHER HMM hit". Therefore, genes in both genomes have the same annotations of PANTHER gene families or subfamilies.

Using PANTHER web services, we performed statistical over-representation tests to compare the over-representation of functional categories in a given gene list against its corresponding reference list (whole genome PANTHER Generic Mapping File); a p-value of enrichment was computed via the binomial distribution. For each functional term, the PANTHER software employs a binomial test to identify statistically significant over-representation (or under-representation) of the genes in an input list relative to the genes in a reference. No assumptions are made about the processes used to generate either input or reference lists. The null hypothesis is that both input and reference lists are drawn from the same

population, such that each functional category is equally well represented in the two lists. Here we only focus on overrepresented functional terms.

Monte Carlo random samplings to detect significant overlaps

Within each social context, we computed the number of shared and non-shared PANTHER gene families and subfamilies between the two species. We used Monte Carlo repeated random sampling to determine if an observed PANTHER gene family or subfamily overlap between species for differentially expressed genes in either context was statistically significant (Ernst 2004). For example, suppose t^* is the observed overlap between gene families in the stickleback territorial challenge and fruit fly territorial challenge experiments. n_1 and n_2 are the sizes of these two lists respectively. We repeatedly and randomly drew samples of size n_1 from the stickleback genome and samples of n_2 from the fruit fly genome for M times (M = 10^5) with replacement and detected an overlap t_i for each iteration of M and computed an estimated p-value using the following equation.

$$estimate\ p = \frac{1 + \sum_{i=1}^{M} I(t_i \geq t^*)}{1 + M} \qquad (1)$$

Where *I(.)* is an indicator function.

Here we used a more conservative test for overlap by fixing the larger set n_1 as the real observed set and randomly sampled n_2, then estimated p using equation (1). This test statistic is more conservative because it randomly samples one rather than both sets (Figure A.1).

RESULTS

Gene families and subfamilies that were socially responsive in both sticklebacks and fruit fly

There were 654 PANTHER gene families and 91 PANTHER subfamilies that were responsive to a territorial challenge in both sticklebacks and fruit fly (Figure 2.2). Both overlaps are greater than expected (p-value <0.001).

There were 155 PANTHER families and 11 PANTHER subfamilies that were responsive to a courtship opportunity in both sticklebacks and fruit fly (Figure 2.2). The overlap at the family level is greater than expected (p-value < 0.001); the overlap at the subfamily level is not greater than expected.

Functional enrichment

To identify core biological processes conserved between sticklebacks and fruit fly in response to social stimuli, we focus our analysis on differentially expressed genes present in gene families or subfamilies that were common to both species, which are hereafter referred to as common family and common subfamily genes, respectively. There were 1665 stickleback and 1258 fruit fly common family genes in response to a territorial challenge, and 97 sticklebacks and 100 fruit fly common subfamily genes in response to a territorial challenge. There were 517 stickleback genes and 207 fruit fly common family genes in response to a courtship opportunity. Within each species, the common family genes were functionally enriched using the PANTHER database with species-appropriate reference lists (see METHODS). We did not test for functional enrichment for the courtship opportunity common subfamily genes because the overlap between species at the subfamily level was not statistically significant. Functional enrichment for the territorial challenge common subfamily genes are reported in Figure A.2.

This analysis revealed molecular functions that were elicited in response to a territorial challenge in both species (Figure 2.3), including RNA and protein binding activities, necessary for cellular transcription and with various enzyme activities e.g. peptidase, lyase, oxidoreductase, isomerase and racemase etc. Molecular functions elicited in response to a courtship opportunity in both species included chromatin, DNA, and RNA binding activities, transcription and translation cofactor activities, receptor activities, ion channel activities and various enzyme activities, which seem to be involved in cell signaling and metabolic processes (Figure 2.3).

Biological processes elicited in response to a territorial challenge in both species included cell adhesion, vesicle-mediated transport, immune system processes, RNA splicing and various metabolic processes involving lipid, protein and cyclic nucleotide metabolic processes (Figure A.3). Biological processes elicited in response to a courtship opportunity in both species included cell-cell signaling, cytokines, ion transport, RNA splicing and polyadenylation, neuronal action potential propagation and various metabolic processes involving rRNA, polysaccharide, cyclic nucleotide, carbohydrate, lipid and steroid metabolic processes (Figure A.3).

Shared and non-shared functional categories

Overall, metabolic processes are common to both social contexts in both species (Figure 2.3). Functional processes related to cell-signaling and stereochemical inversion of biological molecules are more specific to the territorial challenge common family genes in both species. Functional processes related to ion-channel activities and neuronal action potential are more specific to courtship opportunity common family genes in sticklebacks. Other potentially interesting categories related to synaptic plasticity and generation of activity mediated neural response includes GABA and acetylcholine receptor activities. These two functional pathways

are shared between territorial challenge in fruit fly and courtship opportunity in stickleback and are important in excitation or inhibition of postsynaptic potentials (Richmond & Jorgensen 1999; Schuske et al. 2004; Rand 2007).

Selected genes found in enriched functional pathways are described in Table 2.1.

Genes that were responsive to both a social challenge and opportunity in both species

Functional enrichment analysis suggests an important role for metabolic processes in both species in response to social challenges and opportunities. For instance, oxidoreductate and carboxylase activities contain genes e.g. *ddc* involved in the synthesis of serotonin and dopamine (Alekseyenko et al. 2010; Hull et al. 2004), and *CSAD* and *ADC* involved in taurine metabolism and biosynthesis (Tang et al. 1996; Liu et al. 2013; Wu & Prentice 2010) (Table 2.1). Studies have shown that inhibiting the transmission of serotonin and dopamine decreases the level of aggression (Alekseyenko et al. 2010), while dopamine facilitates sexual motivation and sexual behavior (Hull et al. 2004).

Genes that were responsive to a territorial challenge in both species

Functional categories in response to a territorial challenge in both species suggest an important role of cell-signaling genes related to aggression. For example, Abelson interaction proteins were responsive to a territorial challenge in both species. Studies have shown (Ch'ng & Martin 2011; Karpova et al. 2012) that Abi-1 is a synaptically localized protein that is known to translocate to the nucleus via the actin and microtubule cytoskeleton following specific stimuli. Nuclear Abi-1 regulates the gene expression via c-Myc/Max transcription factor complex (Proepper et al. 2007). Also genes such as son of sevenless (SOS) and rab3-interacting molecules (RIM) are involved in activity mediated gene expression and presynaptic active zone formation,

respectively (Yang & Bashaw 2006; Liu et al. 2011). Other signal transduction related genes (e.g. ArfGaps) were also enriched in both species (see Table 2.1).

Another functional category that is overrepresented in territorial challenge common family genes is racemase and epimerase activity, which mediate stereochemical inversion of amino acids and sugars. D-amino acids are unusual amino acids, which recently have attracted significant research interests because of their signaling properties in the nervous system. Serine dehydratase (*sds*) is enriched in stickleback territorial challenge and is implicated with the degradation of D-serine in pyruvate and ammonia by elimination of water molecules (Tanaka et al. 2011). Interestingly, *sds* was down regulated in brain stem in response to a territorial intrusion in stickleback, which suggests an increase of D-serine concentration in response to a territorial challenge. Lastly, glutathione S-transferase (GST), which acts as cellular defense mechanisms against toxicities of electrophiles and reactive oxygen/nitrogen species (Lushchak 2012), is common to both species, suggesting a common detoxification mechanism to cater with aggression-induced oxidative stress (Costantini et al. 2008; Coccaro et al. 2016).

Genes that were responsive to a courtship opportunity in stickleback

Ion channel genes were overrepresented in the courtship opportunity experiment, but not in the territorial challenge experiment. This suggests that electrochemical mediated synapse-to-nucleus signaling, which allows for extremely rapid communication in the cell body and nucleus than soluble proteins is important for responding to a courtship opportunity. Most of the genes from this category are involved in different phases of synaptic plasticity. For example, transient receptor potential are localized in the plasma membrane, where they control the transport of divalent cations, which are essential for several physiological processes e.g. temperature sensing, taste transduction and pheromone signaling (Clapham 2003; Gees et al. 2010; Nilius & Owsianik

2011). Another important gene found in this functional category is reported to regulate male mating behavior. For example, disruption of inositol 1,4,5-triphosphate receptor in male *Caenorhabditis elegans* results in dramatic loss of male fertility (Gower 2005).

DISCUSSION

Here, we report the results of a comparative whole genome, brain transcriptomic analysis of two social contexts across two highly diverged species. To test the toolkit hypothesis, we first computed homologies at the gene family and subfamily levels using PANTHER orthology map, which makes use of both sequence similarities and gene trees (Mi et al. 2005), and then we used unbiased and conservative permutation schemes to test the extent of gene overlaps. We found that despite millions of years of divergence, dramatically different neuroanatomy and technical differences between experiments in the two species (e.g., different array platforms), the response to a social challenge invokes a common set of molecular mechanisms governed by homologous genes in distantly related species (consistent with Rittschof et al 2014): there were gene families and subfamilies that were responsive to a territorial challenge in both fruit flies and sticklebacks. A highly significant p-value was observed at the gene family level, even when one set was fixed and the other was chosen randomly (Figure A.1). We found less support for the toolkit hypothesis in the social opportunity (courtship) context: there were commonalities between the two species at the gene family but not the gene subfamily level.

One explanation for the failure to find strong support for the toolkit hypothesis in the courtship context has to do with differences between the two species with respect to their reproductive behavior, particularly their choosiness and parental investment. While male *Drosophila* do not have strong mating preferences and do not provide care for offspring (MØller

& Thornhill 1998), male sticklebacks are choosy and provide paternal care that is necessary for offspring survival (Smith & Wootton 1999). Moreover, female sticklebacks can represent a social threat as well as a social opportunity because females often cannibalize the contents of males' nests (Belles-Isles et al. 1990) Therefore, a potential mate might represent a very different social stimulus to a male fruit fly compared to a male stickleback.

By comparing the brain transcriptomic reaction to social stimuli in two distantly-related species, we gained insights into the extent to which molecular mechanisms associated with different social behaviors have been conserved throughout evolution. For example, consistent with other studies (Chandrasekaran et al. 2015; Li-Byarlay et al. 2014), our results suggest an important role for metabolism in conserved aspects of social behavior. We also note a trend of cell signaling-related activity genes in response to a territorial challenge, and ion-channel related activity genes in response to a courtship opportunity.

For instance, we noted overrepresentation of small GTPase regulators, and racemase and epimerases pathways in response to a territorial challenge in both species (Figure 2.3). Genes such as ABI-related proteins present in the small GTPases GO term suggest common synapse-to-nucleus signal transduction, and genes such as SDS present in the racemase and epimerase activity GO term suggest a role of D-amino acids in both species due to aggression. Moreover, overrepresentation of ion channel and ligand-gated ion channel activity genes in response to a courtship opportunity in sticklebacks possibly suggests a combination of electrochemical and calcium influx mediated synapse-to-nucleus signal transduction.

CONCLUSIONS

Here we catalogued and discussed the neuromolecular mechanisms that were responsive to social stimuli in sticklebacks and fruit fly. Our findings can be summarized as follows:

1. There may be a toolkit of genes conserved at flexible levels of homologies (e.g. at gene family and subfamily), responsive to territorial challenge and courtship opportunity in both species. We found more support for toolkit genes in territorial challenge social context.
2. Metabolic genes were common to both social contexts and to both species.
3. Cell signaling genes were responsive to a territorial challenge in both species.
4. We speculate that D-amino acids are involved in mediating the response to a territorial challenge in both species.
5. Over representation of ion channel activity genes in response to a courtship opportunity in sticklebacks suggests potentially different routes of initial synapse to nucleus signal processing in response to a courtship opportunity vs. a territorial challenge.

Further refinements of the toolkit hypothesis are expected to emerge from more detailed studies examining gene expression in response to social stimuli over time and in different parts of the brain. It may be, for example, that there are toolkit genes, but they operate at different points in time and space in different species. In conclusion, we report transcriptomic commonalities to social stimuli at the gene family and subfamily levels in a fish and an insect, which may help improve our understanding of the evolution of complex traits like social behavior.

ACKNOWLEDGEMENTS

We thank Charles Blatti for helpful discussion on computing overlaps among gene lists using permutations. We also thank members of the Bell lab for useful discussion and comments on the manuscript.

FIGURES AND TABLE

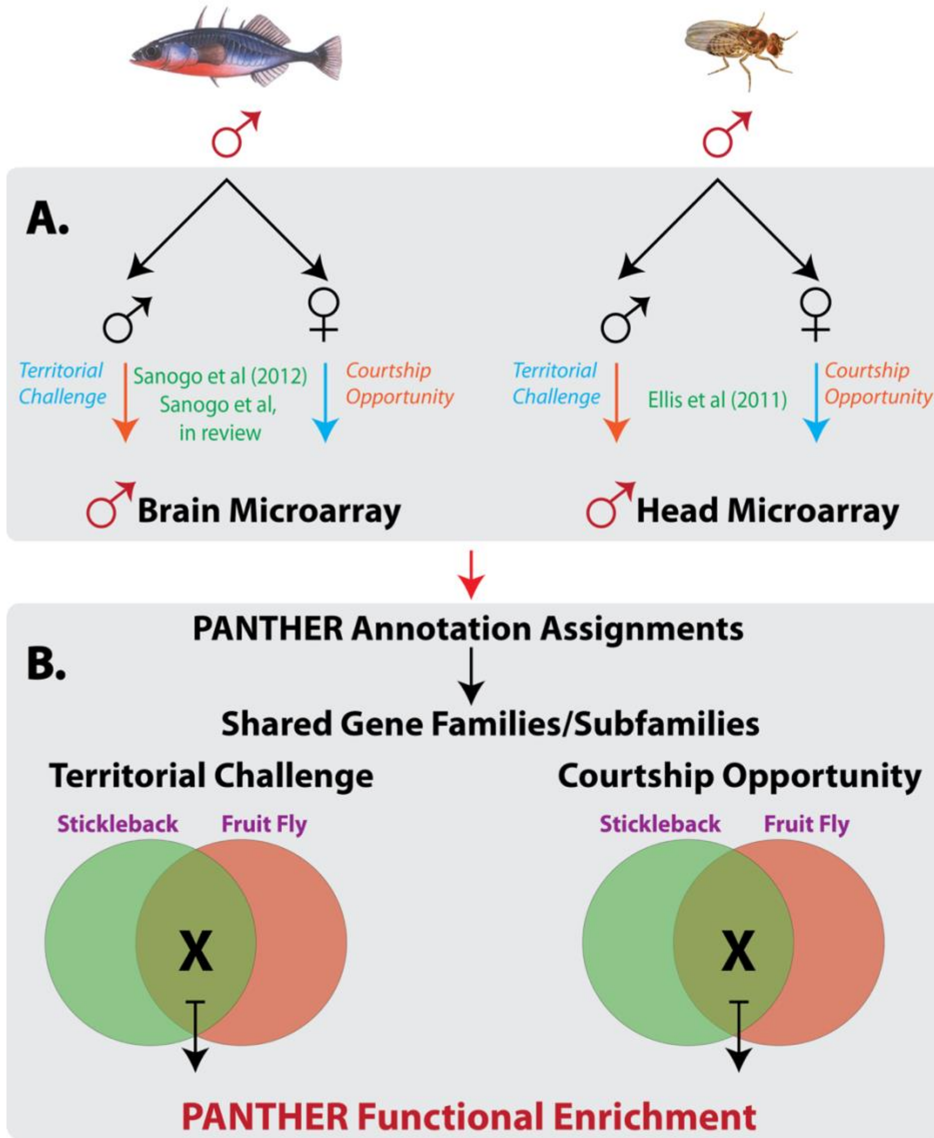


Figure 2.1 – Study workflow. (A) Territorial challenge and courtship opportunity experiments were performed in sticklebacks and fruit fly, and gene expression was measured using microarrays. (B) A comparative transcriptomics pipeline was adopted to mine for common gene families or subfamilies and neuromolecular mechanisms.

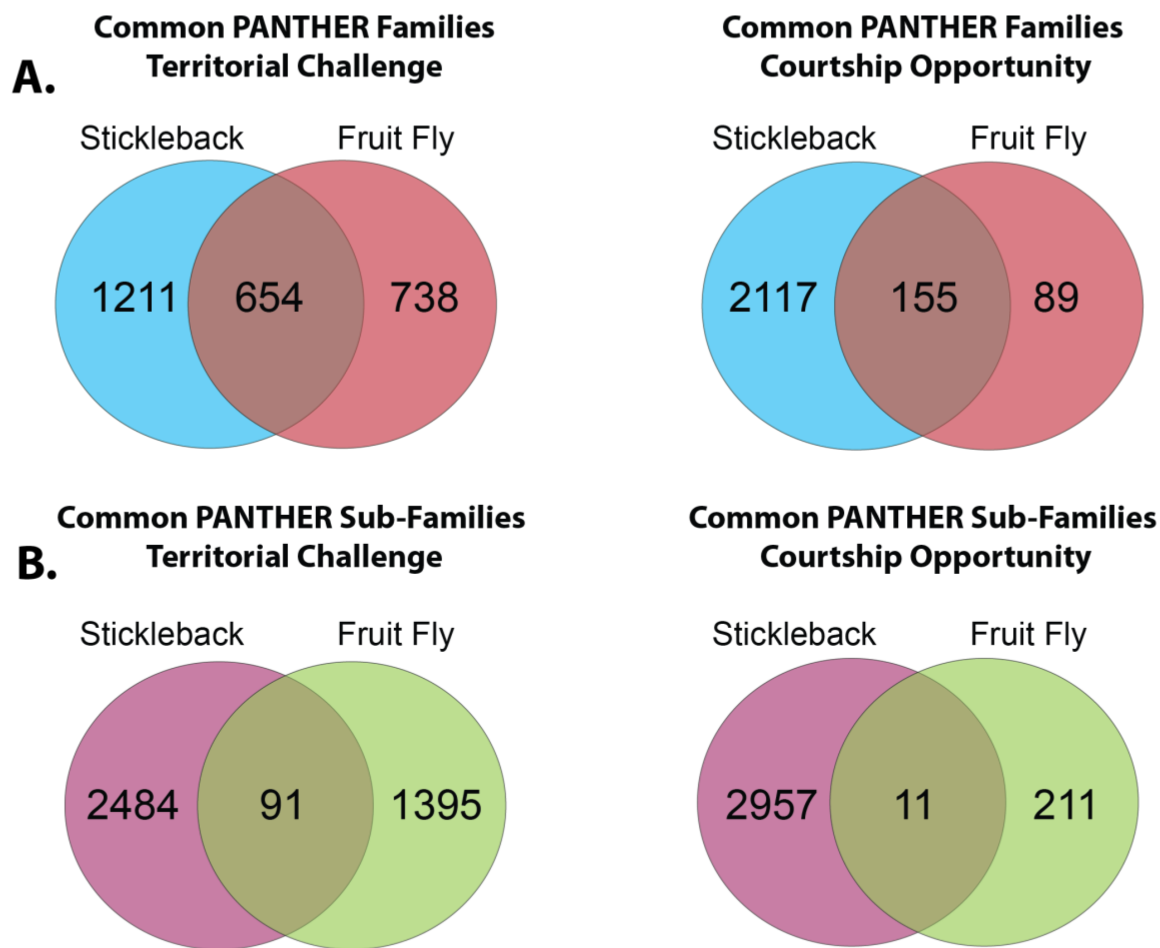


Figure 2.2 – Gene families (A) and subfamilies (B) that were common to both species, separated by social context (left: territorial challenge; right: courtship opportunity).

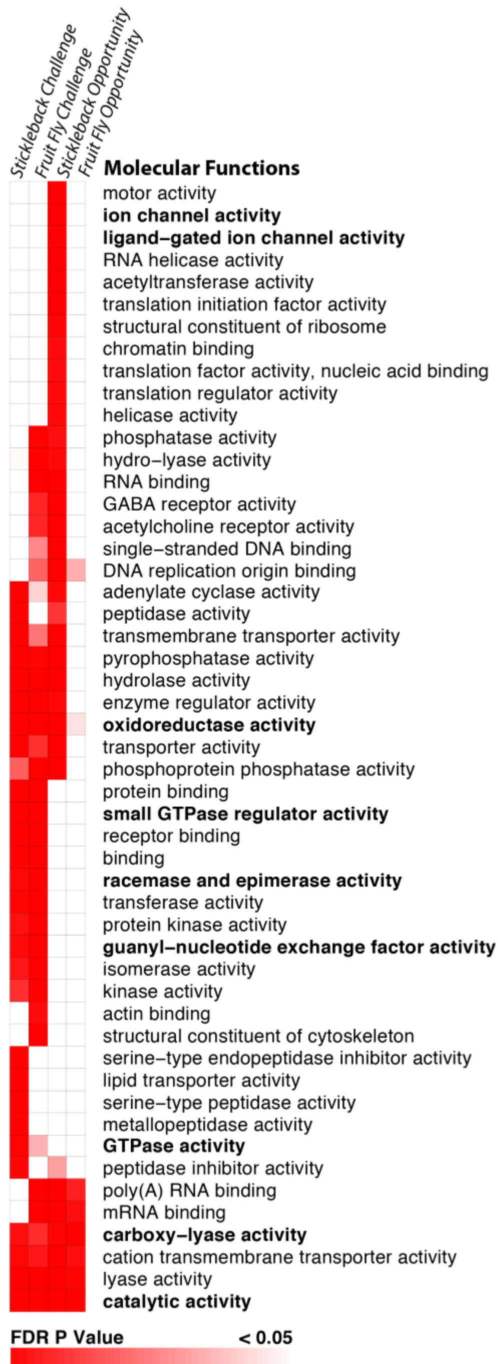


Figure 2.3 – Functional enrichment of stickleback and fruit fly common gene families. Only functional terms significant at FDR < 0.05 were used to create the heatmap. Terms whose font is bold are discussed more in the text.

Table 2.1 – Selected genes that were responsive to a social challenge and/or opportunity in either one or both species.

Social Context	Functional Category	Stickleback Gene	Fruit Fly Gene	Comments	Reference
territorial challenge, courtship opportunity	metabolic pathways	Dopa decarboxylase: ddc	*ddc*	synthesis of dopamine and serotonin. inhibition of these decreases aggression while dopamine facilitates sexual motivation and behavior.	(Alekseyenko et al. 2010; Hull et al. 2004)
territorial challenge, courtship opportunity	metabolic pathways	cysteine sulfonic acid decarboxylase *CSAD*	aspartate decarboxylase *ADC*	Both play important role in taurine metabolism and biosynthesis. Neuroprotector against glutamate excitocity.	(Liu et al. 2013; Wu & Prentice 2010; X. W. X. W. Tang et al. 1996)
territorial challenge	signal transduction	Abelson interaction protein: *Abi-1a, ABI3 (2 of 2)*	*Abi-1*	Activity mediated signal transduction and gene expression	(Ch'ng & Martin 2011; Karpova et al. 2012)

Table 2.1 (continued)

territorial challenge	signal transduction	Son of sevenless: *sos2*	*Sos*	Activity mediated signal transduction and gene expression	(Yang & Bashaw 2006; Liu et al. 2011)
territorial challenge	signal transduction	rab3-interacting molecules: *RIMS4, RIMS2a, RIMS2b*	*RIM*	Essential component of presynaptic active zone and plays an important part in neurotransmitter release.	(Yang & Bashaw 2006; Liu et al. 2011)
territorial challenge	signal transduction	Arf GTPase-activating proteins (ArfGaps): *ARFGAP1, asap1b, ADAP1 (1 of 2), acap3b, acap3a, agfg1b*	*Asap1*	Membrane trafficking and actin cytoskeleton remodelling	(Inoue & Randazzo 2007)
territorial challenge	metabolic pathways	Serine dehydrogenase: *SDS*	unknown	Degradation of D-Serine	(Tanaka et al. 2011; Wang et al. 2012)

Table 2.1 (continued)

territorial challenge	metabolic pathways	glutathione S-transferase: *gstt1a*	*gste2*, *gste3*, *gste5*, *gste7*, *gste9* and *gstd9*	common detoxification mechanism to cater with aggression-induced oxidative stress	(Costantini et al. 2008; Coccaro et al. 2016)
courtship opportunity (stickleback)	Ion channel activity	transient receptor potential cation channels: *trpm3, trpm4, trpm5 and trpm6*		important for temperature sensing, taste transduction and pheromone signaling.	(Clapham 2003; Gees et al. 2010; Nilius & Owsianik 2011)
courtship opportunity (stickleback)	Ion-channel activity	inositol 1,4,5-triphosphate receptor: *itpr2*		disruption results in dramatic loss of male fertility	(Gower 2005)

LITERATURE CITED

Alekseyenko, O. V., Lee, C. & Kravitz, E.A., 2010. Targeted Manipulation of Serotonergic Neurotransmission Affects the Escalation of Aggression in Adult Male Drosophila melanogaster B. D. McCabe, ed. *PLoS ONE*, 5(5), p.e10806. Available at: http://dx.plos.org/10.1371/journal.pone.0010806 [Accessed March 27, 2016].

Belles-Isles, J.-C., Cloutier, D. & FitzGerald, G., 1990. Female cannibalism and male courtship tactics in threespine sticklebacks. *Behavioral Ecology and Sociobiology*, 26(5), pp.363–368. Available at: http://link.springer.com/10.1007/BF00171103 [Accessed December 23, 2016].

Berens, A.J., Hunt, J.H. & Toth, A.L., 2015. Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. *Molecular biology and evolution*, 32(3), pp.690–703. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25492498 [Accessed March 27, 2016].

Ch'ng, T.H. & Martin, K.C., 2011. Synapse-to-nucleus signaling. *Current opinion in neurobiology*, 21(2), pp.345–52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21349698 [Accessed March 27, 2016].

Chandrasekaran, S. et al., 2015. Aggression is associated with aerobic glycolysis in the honey bee brain(1). *Genes, brain, and behavior*, 14(2), pp.158–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25640316 [Accessed March 27, 2016].

Clapham, D.E., 2003. TRP channels as cellular sensors. *Nature*, 426(6966), pp.517–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14654832 [Accessed March 27, 2016].

Coccaro, E.F., Lee, R. & Gozal, D., 2016. Elevated Plasma Oxidative Stress Markers in Individuals With Intermittent Explosive Disorder and Correlation With Aggression in Humans. *Biological psychiatry*, 79(2), pp.127–35. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/24582164 [Accessed March 27, 2016].

Costantini, D. et al., 2008. Aggressive and non-aggressive personalities differ in oxidative status in selected lines of mice (Mus musculus). *Biology Letters*, 4(1), pp.119–122. Available at: http://rsbl.royalsocietypublishing.org/cgi/doi/10.1098/rsbl.2007.0513 [Accessed March 27, 2016].

Cummings, M.E. et al., 2008. Sexual and social stimuli elicit rapid and contrasting genomic responses. *Proceedings. Biological sciences / The Royal Society*, 275(1633), pp.393–402. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18055387 [Accessed October 10, 2016].

Ellis, L.L. & Carney, G.E., 2011. Socially-Responsive Gene Expression in Male Drosophila melanogaster Is Influenced by the Sex of the Interacting Partner. *Genetics*, 187(1), pp.157–169. Available at: http://www.genetics.org/cgi/doi/10.1534/genetics.110.122754 [Accessed March 27, 2016].

Ernst, M.D., 2004. Permutation Methods: A Basis for Exact Inference. *Statistical Science*, 19(4), pp.676–685. Available at: http://projecteuclid.org/Dienst/getRecord?id=euclid.ss/1113832732/ [Accessed March 27, 2016].

Gees, M., Colsoul, B. & Nilius, B., 2010. The role of transient receptor potential cation channels in Ca2+ signaling. *Cold Spring Harbor perspectives in biology*, 2(10), p.a003962. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20861159 [Accessed March 27, 2016].

Gower, N.J.D., 2005. Inositol 1,4,5-Trisphosphate Signaling Regulates Mating Behavior in Caenorhabditis elegans Males. *Molecular Biology of the Cell*, 16(9), pp.3978–3986. Available at: http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E05-02-0096 [Accessed March 27, 2016].

Hull, E.M., Muschamp, J.W. & Sato, S., 2004. Dopamine and serotonin: influences on male sexual behavior. *Physiology & behavior*, 83(2), pp.291–307. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15488546 [Accessed March 27, 2016].

Inoue, H. & Randazzo, P.A., 2007. Arf GAPs and their interacting proteins. *Traffic (Copenhagen, Denmark)*, 8(11), pp.1465–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17666108 [Accessed March 27, 2016].

Karpova, A., Bär, J. & Kreutz, M.R., 2012. Long-distance signaling from synapse to nucleus via protein messengers. *Advances in experimental medicine and biology*, 970, pp.355–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22351064 [Accessed March 27, 2016].

Li-Byarlay, H. et al., 2014. Socially responsive effects of brain oxidative metabolism on aggression. *Proceedings of the National Academy of Sciences*, 111(34), pp.12533–12537. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1412306111 [Accessed March 27, 2016].

Liu, K.S.Y. et al., 2011. RIM-binding protein, a central part of the active zone, is essential for neurotransmitter release. *Science (New York, N.Y.)*, 334(6062), pp.1565–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22174254 [Accessed March 27, 2016].

Liu, P. et al., 2013. Mechanism of cysteine-dependent inactivation of aspartate/glutamate/cysteine sulfinic acid α-decarboxylases. *Amino acids*, 44(2), pp.391–404. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22718265 [Accessed March 27, 2016].

Lushchak, V.I., 2012. Glutathione Homeostasis and Functions: Potential Targets for Medical Interventions. *Journal of Amino Acids*, 2012, pp.1–26. Available at: http://www.hindawi.com/journals/jaa/2012/736837/ [Accessed March 27, 2016].

Mi, H. et al., 2005. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic acids research*, 33(Database issue), pp.D284-8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15608197 [Accessed March 27, 2016].

Mikheyev, A.S. & Linksvayer, T.A., 2015. Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. *eLife*, 4. Available at: http://elifesciences.org/lookup/doi/10.7554/eLife.04775 [Accessed March 27, 2016].

MØller & Thornhill, 1998. Male parental care, differential parental investment by females and sexual selection. *Animal behaviour*, 55(6), pp.1507–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9641996 [Accessed April 1, 2016].

Nilius, B. & Owsianik, G., 2011. The transient receptor potential family of ion channels. *Genome biology*, 12(3), p.218. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21401968 [Accessed March 27, 2016].

Peterson, K.J. et al., 2004. Estimating metazoan divergence times with a molecular clock. *Proceedings of the National Academy of Sciences of the United States of America*, 101(17), pp.6536–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15084738 [Accessed December 19, 2016].

Proepper, C. et al., 2007. Abelson interacting protein 1 (Abi-1) is essential for dendrite morphogenesis and synapse formation. *The EMBO journal*, 26(5), pp.1397–409. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17304222 [Accessed March 27, 2016].

Rand, J., 2007. Acetylcholine. *WormBook*. Available at: http://www.wormbook.org/chapters/www_acetylcholine/acetylcholine.html [Accessed March 27, 2016].

Richmond, J.E. & Jorgensen, E.M., 1999. One GABA and two acetylcholine receptors function

at the C. elegans neuromuscular junction. *Nature neuroscience*, 2(9), pp.791–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10461217 [Accessed March 27, 2016].

Rittschof, C.C. et al., 2014. Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proceedings of the National Academy of Sciences of the United States of America*, 111(50), pp.17929–34.

Rittschof, C.C. & Robinson, G.E., 2014. Genomics: moving behavioural ecology beyond the phenotypic gambit. *Animal behaviour*, 92, pp.263–270. Available at: http://www.sciencedirect.com/science/article/pii/S0003347214001134 [Accessed March 27, 2016].

Sanogo, Y.O. et al., 2012. Transcriptional regulation of brain gene expression in response to a territorial intrusion. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), pp.4929–4938. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.2087 [Accessed November 11, 2016].

Sanogo, Y.O. et al., 2012. Transcriptional regulation of brain gene expression in response to a territorial intrusion. *Proceedings. Biological sciences / The Royal Society*, 279(1749), pp.4929–38. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23097509 [Accessed March 27, 2016].

Sanogo, Y.O. & Bell, A.M., 2016. Molecular mechanisms and the conflict between courtship and aggression in three-spined sticklebacks. *Molecular Ecology*, 25(17), pp.4368–4376. Available at: http://doi.wiley.com/10.1111/mec.13766 [Accessed November 11, 2016].

Schuske, K., Beg, A.A. & Jorgensen, E.M., 2004. The GABA nervous system in C. elegans. *Trends in neurosciences*, 27(7), pp.407–14. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/15219740 [Accessed March 27, 2016].

Smith, C. & Wootton, R.J., 1999. Parental energy expenditure of the male three-spined stickleback. *Journal of Fish Biology*, 54(5), pp.1132–1136. Available at: http://doi.wiley.com/10.1111/j.1095-8649.1999.tb00866.x [Accessed April 1, 2016].

Smyth, G.K., 2005. limma: Linear Models for Microarray Data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York: Springer-Verlag, pp. 397–420. Available at: http://link.springer.com/10.1007/0-387-29362-0_23 [Accessed December 23, 2016].

Stern, D.L., 2013. The genetic causes of convergent evolution. *Nature Reviews Genetics*, 14(11), pp.751–764. Available at: http://www.nature.com/doifinder/10.1038/nrg3483 [Accessed March 27, 2016].

Tanaka, H. et al., 2011. Crystal structure of a zinc-dependent D-serine dehydratase from chicken kidney. *The Journal of biological chemistry*, 286(31), pp.27548–58. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21676877 [Accessed March 27, 2016].

Tang, X.W. et al., 1996. Multiplicity of Brain Cysteine Sulfinic Acid Decarboxylase - Purification, Characterization and Subunit Structure. *Journal of biomedical science*, 3(6), pp.442–453. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11725125 [Accessed December 22, 2016].

Toth, A.L. et al., 2010. Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1691).

Toth, A.L. et al., 2014. Shared genes related to aggression, rather than chemical communication, are associated with reproductive dominance in paper wasps (Polistes metricus). *BMC*

Genomics, 15(1), p.75. Available at: http://www.biomedcentral.com/1471-2164/15/75 [Accessed March 27, 2016].

Toth, A.L. & Robinson, G.E., 2007. Evo-devo and the evolution of social behavior. *Trends in Genetics*, 23(7), pp.334–341. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0168952507001552 [Accessed March 27, 2016].

Wang, C.-Y. et al., 2012. Modulating the function of human serine racemase and human serine dehydratase by protein engineering. *Protein engineering, design & selection : PEDS*, 25(11), pp.741–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23112234 [Accessed March 27, 2016].

Woodard, S.H. et al., 2011. Genes involved in convergent evolution of eusociality in bees. *Proceedings of the National Academy of Sciences*, 108(18), pp.7472–7477. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1103457108 [Accessed March 27, 2016].

Woodard, S.H. et al., 2014. Molecular heterochrony and the evolution of sociality in bumblebees (Bombus terrestris). *Proceedings of the Royal Society B: Biological Sciences*, 281(1780), pp.20132419–20132419. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2013.2419 [Accessed March 27, 2016].

Wootton, R.J., 1984. *A Functional Biology of Sticklebacks*, Boston, MA: Springer US. Available at: http://link.springer.com/10.1007/978-1-4615-8513-8 [Accessed March 27, 2016].

Wu, J.-Y. & Prentice, H., 2010. Role of taurine in the central nervous system. *Journal of biomedical science*, 17 Suppl 1, p.S1. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20804583 [Accessed March 27, 2016].

Yang, L. & Bashaw, G.J., 2006. Son of sevenless directly links the Robo receptor to rac activation to control axon repulsion at the midline. *Neuron*, 52(4), pp.595–607. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17114045 [Accessed March 27, 2016].

Zayed, A. & Robinson, G.E., 2012. Understanding the Relationship Between Brain Gene Expression and Social Behavior: Lessons from the Honey Bee. *Annual Review of Genetics*, 46(1), pp.591–615.

Zwarts, L., Versteven, M. & Callaerts, P., 2012. Genetics and neurobiology of aggression in Drosophila. *Fly*, 6(1), pp.35–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22513455 [Accessed April 1, 2016].

CHAPTER 3: TEMPORAL DYNAMICS OF NEUROGENOMIC PLASTICITY IN RESPONSE TO SOCIAL INTERACTIONS IN MALE THREESPINED STICKLEBACKS

ABSTRACT

Animals exhibit dramatic immediate behavioral plasticity in response to social interactions, and brief social interactions can shape the future social landscape. However, the molecular mechanisms contributing to behavioral plasticity are unclear. Here, we show that the genome dynamically responds to social interactions with multiple waves of transcription associated with distinct molecular functions in the brain of male threespined sticklebacks, a species famous for its behavioral repertoire and evolution. Some biological functions (e.g., hormone activity) peaked soon after a brief territorial challenge and then declined, while others (e.g., immune response) peaked hours afterwards. We identify transcription factors that are predicted to coordinate waves of transcription associated with different components of behavioral plasticity. Next, using H3K27Ac as a marker of chromatin accessibility, we show that a brief territorial intrusion was sufficient to cause rapid and dramatic changes in the epigenome. Finally, we integrate the time course brain gene expression data with a transcriptional regulatory network, and link gene expression to changes in chromatin accessibility. This study reveals rapid and dramatic epigenomic plasticity in response to a brief, highly consequential social interaction.

Chapter 3 was published in PLOS Genetics in 2017: Bukhari, S.A., Saul, M.C., Seward, C.H., Zhang, H., Bensky, M., James, N., Zhao, S.D., Chandrasekaran, S., Stubbs, L. and Bell, A.M., 2017. Temporal dynamics of neurogenomic plasticity in response to social interactions in male threespined sticklebacks. PLOS genetics, 13(7), p.e1006840.

INTRODUCTION

Animals exhibit remarkable behavioral plasticity. Social interactions in particular can provoke moment-to-moment changes in behavior. These changes are coordinated at the neural level, but social interactions also elicit transcriptional changes within the brains of behaving animals (Burmeister et al. 2005). For example, genome-wide transcription studies show that roughly ~10% of the genome responds to a mating opportunity (McGraw et al. 2008; Mack et al. 2006; Lawniczak & Begun 2004; Cummings et al. 2008b; Carney et al. 2007; Fraser et al. 2014), predation risk (Sanogo et al. 2011; Lavergne et al. 2014; Jansen et al. 2013), or a territorial challenge (Alaux et al. 2009; Rittschof & Robinson 2013; Sanogo et al. 2012).

However, we know little about the temporal and spatial dynamics of neurogenomic plasticity in response to social interactions. It is likely that there are waves of transcription associated with perceiving social information, responding to social information, maintaining a behavioral response, recovering from the social interaction and modifying future behavior (Aubin-Horth & Renn 2009; Bell & Aubin-Horth 2010) . Static experiments that measure gene expression at a single time point can only catch a glimpse of what is probably a very dynamic and coordinated process.

Studies in development have linked changes in chromatin accessibility with the time course of changes in gene expression and the activity of transcription factors operating within gene regulatory networks (M. N. Arbeitman et al. 2002; Bar-Joseph et al. 2012). This tactic has also proven to be successful for examining acute, short-term responses of other types, for example, in response to pathogens (Amit et al. 2009; Huang et al. 2011). However, whether the same principles apply to behavioral stimuli, and social interactions in particular, is unknown.

Here, we test the hypothesis that a brief social interaction, albeit one with strong implications for fitness, is sufficient to induce transcriptomic and epigenomic responses that change over time. We test this hypothesis in threespined sticklebacks (*Gasterosteus aculeatus*), a species for which successful territorial defense is critical for Darwinian fitness. Sticklebacks are small fish whose behavioral repertoire has attracted attention since the early ethologists (Pelkwijk & Tinbergen 1937). Freshwater sticklebacks must quickly establish territories because they have a short window of opportunity to breed in the spring, and die at the end of the breeding season. Male sticklebacks typically occur in neighborhoods and function in a dynamic social environment where they vigorously defend individual nesting territories against intrusions by rival males and predators. The territory is the hub of family life, where the father constructs a nest, attracts females to mate and where he rears the offspring without any help from the mother. If a male fails to defend a territory, he will not obtain a mate and he will not produce offspring therefore effective defense of that territory is necessary for reproductive success. Like other territorial animals, male sticklebacks exhibit experience-dependent changes in behavior following a territorial intrusion, as they learn the boundaries of their territory and how to detect and repel intruders (Peeke 1969; Peeke & Veno 1973).

We provide evidence that the genome and the epigenome are highly responsive to social interactions during territory defense. We characterize transcriptomic and epigenomic plasticity in response to social interactions by measuring changes in gene expression at three points in time following a brief territorial intrusion using RNASeq. We compare expression in two parts of the brain containing nodes in the social decision-making network (O'Connell & Hofmann 2012): diencephalon and telencephalon. The diencephalon includes the hypothalamus – a key integrator of social information with the neuroendocrine system – while the telencephalon is a part of the

forebrain, and includes the teleost homolog of the hippocampus. Using these data, we construct a transcriptional regulatory network that links temporal changes in gene expression in different parts of the brain to the activity of transcription factors operating within a gene regulatory network. Finally, we measure changes in chromatin accessibility in response to a social interaction on a genome-wide scale, using acetylated lysine 27 on histone H3 (H3K27Ac) as a marker of accessible chromatin, and link changes in chromatin accessibility to changes in gene expression. We show that many of the same principles that characterize transcriptomic and epigenomic changes unfolding over development (Michelle N Arbeitman et al. 2002; Bar-Joseph et al. 2012) also apply to the brain's response to brief, but potent, social behavior.

RESULTS

Spatiotemporal dynamics of the transcriptomic response to territorial challenge

Within both brain regions, we identified genes whose expression was influenced by a territorial challenge at three time points: 30, 60 and 120 minutes. The greatest transcriptional response to a territorial challenge occurred 60 minutes after the challenge (Figure 3.1a). Generally, gene expression was down-regulated 30 and 60 minutes after the social challenge but was up-regulated at the 120 minute time point in diencephalon (Figure 3.1a).

The transcriptomic response to a territorial challenge changed rapidly over time in both brain regions (Figure 3.1b, c); in fact, there was little overlap between the differentially expressed genes detected at each time point. Functional enrichment analysis revealed that early responding genes (30 minutes) were related to hormones and post-translational modifications (PTMs), whereas a strong signature of genes related to metabolism dominated at the 60 minute

time point (consistent with Rittschof et al. 2014). By 120 minutes, differentially expressed functions shifted toward transcription, immune response and homeostasis (Figure 3.1g).

Not surprisingly, we detected strong differences in gene expression between brain regions. However, there were some genes that were differentially expressed in both brain regions, and these genes showed a remarkably concordant quantitative pattern of expression across brain regions at 30 and 60 minutes (Figure 3.1d, e). Specifically, genes that were strongly upregulated in diencephalon in response to a territorial challenge were also strongly upregulated in telencephalon (correlation > 0.9). The pattern at 120 minutes was different, with a subset of genes (n=18, hereafter referred to as 'discordant genes') showing the opposite pattern of regulation in the two brain regions. These 18 genes, which were upregulated in diencephalon and downregulated in telencephalon after the territorial challenge (Figure 3.1f), are primarily related to visual perception and include retinal genes (e.g. *rom1b, rom 1a,* opsins), circadian genes (e.g. *crx*, opsins) and phosphodiesterases, which have been repeatedly duplicated in the stickleback genome and acquired new functions (Sato et al. 2009).

Waves of transcription in response to a territorial challenge

In order to find genes that changed in a coordinated fashion in response to a territorial challenge, we first analyzed the gene expression data by testing for main effects and interactions between them. We built separate generalized linear models for each brain region and were particularly interested in genes whose time course of expression was influenced by the territorial intrusion (time x treatment interaction term). There were 758 and 739 such genes in diencephalon and telencephalon, respectively, hereafter referred to as DEG_x (FDR < 0.1).

We next used hierarchical clustering of the DEG_x to determine whether there were clusters of genes that changed in concert together. We identified 12 and 13 clusters in

diencephalon and telencephalon, respectively (Figure 3.2a, b). Each cluster had a particular expression profile over time in response to a territorial challenge. For example, cluster D1 comprised a set of genes that were downregulated at 30 mins, upregulated at 60 mins and downregulated again at 120 mins. On the other hand, cluster D2 comprised a set of genes that were upregulated at 30 mins, downregulated at 60 mins and then strongly upregulated at 120 mins.

Functional analysis revealed that genes with similar time-course profiles also tended to have similar functions. For instance, cluster D1 comprised hormonal genes such as *tshb*, *prl*, *cga*, *lhb* and *gh1*, and a nuclear receptor transcription factor *nr5a1b,* which binds to a *prl* (encoding prolactin) promoter (Hu et al. 1997). In contrast, cluster T2 included transcription factors such as *pax7* (both *a* and *b* paralogs*), irx (2a, 3a and 5a)*, *tfap2b*, *shox*, *sp5l*, *DMBX1* and *pou4f2* and two hormonal genes known to be very important to social behavior (*avp* and *oxt*). The clustering of hormonal genes with transcription factors suggests a complex interplay between hormones with transcription factors in response to a territorial challenge. Clusters D2 and T3 included the genes that exhibit the discordant pattern of expression across brain regions at 120 minutes (Figure 3.1f).

Functional enrichment analysis confirmed that each cluster of genes is associated with its own unique set of functions (Figure 3.2c, d). GO terms enriched in cluster D10, for example, were not shared with any other cluster, while cluster D4 also associated with its own unique set of GO terms. These results are consistent with the hypothesis that there are waves of transcription associated with different biological functions following a social interaction. Some biological functions peak early then subside (e.g., cluster T3), while others peak at 60 minutes

(e.g., D1, D8, T5, T8); still others peak hours (120 mins, e.g., D2, D4, T1, T4) following a social interaction.

Transcription factors within a gene regulatory network coordinate waves of transcription

Next, to identify genes that regulate transcriptional changes in the stickleback brain in response to a territorial challenge, we reconstructed a transcriptional regulatory network (TRN) model using the ASTRIX approach (Chandrasekaran et al. 2011). We used the gene expression data to identify regulatory interactions between transcription factors and their predicted target genes (see methods). ASTRIX infers a genome-scale TRN model capable of making quantitative predictions about the expression levels of genes given the expression values of the transcription factors. The full TRN is in Figure B.1.

We then integrated the DEG_x from the hierarchical clustering analysis with the TRN in order to find transcription factors that are predicted to regulate the clusters. This integration proved to be insightful because it connected dynamic gene expression to interacting transcription factors within a gene regulatory network. For example, the transcription factors *dlx4a*, *grhl3* and *si:ch211-157c3.4* were predicted to regulate cluster D9 (enriched for energy metabolism and immune response) and were connected to each other in the network. This analysis therefore allows us to identify transcription factors within a gene regulatory network that we hypothesize are regulating clusters of genes that change in a coordinated fashion in response to a territorial challenge (Figure 3.3).

Indeed, closer examination of the dynamics of expression of transcription factors and their targets revealed that many of the transcription factors in the TRN showed expression patterns consistent with the cluster they were predicted to regulate. For example, the expression pattern over time of *irf8* and *cebpb* was very similar to the expression pattern of their predicted

targets (D4, D5, D6, D9: no change, down, up).

The TRN offered a number of insights into the spatiotemporal dynamics of gene expression in response to a territorial challenge. For example, the TRN can help explain striking patterns in the gene expression results, such as the discordant genes that were upregulated in diencephalon and downregulated in telencephalon 120 minutes after a social challenge (Figure 3.1f). The discordant genes were in clusters D2 and T3, which were predicted to be regulated by the set of connected transcription factors *otx5*, *vsx1* and *CRX*. *CRX,* implicated with circadian rhythm in addition to visual functions (Furukawa et al. 1999), is noteworthy because its expression profile was consistent with the expression pattern of the discordant genes: *CRX* was upregulated in diencephalon and downregulated in telencephalon at 120 minutes.

Linking changes in gene expression to changes in chromatin accessibility

While chromatin is suspected to change relatively slowly compared to mRNA in adult tissues changes in chromatin accessibility can be an important driver of changes in gene expression (Hon et al. 2009). However, little is known about the impact of short-term behavioral interactions on the chromatin landscape.

To test the hypothesis that a brief social interaction has consequences for the epigenome, we used chromatin immunoprecipitation on histone H3 subunits with acetylated lysine 27 (H3k27Ac ChIPseq), a marker of accessible chromatin, to assess changes in genome-wide chromatin accessibility at two time points (30 and 120 minutes) following a territorial challenge in diencephalon. These experiments revealed tens of thousands of H3K27Ac peaks in each sample tested with robust *p* values and enrichment (see Methods for details). We distinguish between areas of the genome that were accessible in controls (‘baseline accessible peaks’) from

areas of the genome whose accessibility changed in response to a territorial challenge (i.e. differed between control and experimental males, 'differentially accessible peaks', DAPs).

Most of the genes were accessible at baseline (Figure B.2). There were 23656 and 18797 baseline accessible peaks ($\geq$4-fold change in peak difference between sample and input, and $p < 10^{-4}$) associated with 12630 and 11723 genes (within 20kb) at 30 minutes and 120 minutes, respectively.

However, there were a large number of genes whose accessibility was affected by a territorial challenge, particularly 120 minutes following the challenge (Figure 3.4a, b). There were 2868 differentially accessible peaks (DAPs) that were associated with 1975 genes (within 20kb, DAPs; 2-fold and $p < 10^{-4}$). Representative DAPs are shown in Figure 3.4d, e, f. Many (n=97) of the peaks that were differentially accessible at 120 minutes were near genes whose expression profile changed over time in response to a territorial challenge ($DEG_{x,}$ Figure 3.4c). The DEG_x associated with nearby DAPs (hereafter referred to as $DAPDEG_x$) were not a random set, but also enriched in specific functional categories: functions related to stimulus response, cell signaling and development were highly enriched in this gene set.

The territorial challenge had dramatic consequences for chromatin in terms of peak size, with fold enrichment or depletion in specific peaks after challenge as high as 30-40x (Figure 3.4a). There was a general trend toward more accessibility at the 120 min following a territorial challenge, consistent with the general pattern of up-regulation of DEGs in diencephalon at the 120 min time point (Figure 3.1a).

Computational analysis suggests a small set of peaks (associated with 24 genes) that were inaccessible in the control group but which became accessible in response to a territorial challenge at 120 minutes, and are possible pioneer factors. A representative sample is in Figure

3.4f, which shows a differential peak within 5kb upstream of *C4B*. The peaks that were inaccessible at baseline but became accessible in response to a territorial challenge included several that are near genes associated with the immune response, e.g. *irg1*, *lcp1, ccr8.1, pstpip1b, PRF1, C4B, zc3h12a*.

Table 3.1 illustrates how changes in chromatin are linked to the activity of transcription factors in the TRN and the expression of their targets over time. All of the transcription factors in Table 3.1 are in the TRN and the genes encoding these TFs were all associated with DAPs that either became accessible, or became more accessible, in response to a territorial challenge at 120 minutes. All but one of these TFs (*NFATC3*) regulate clusters of genes that are upregulated at 120 minutes (e.g. clusters D9, D12, D6, D12, D5). Several transcription factors (*NFATC3*, *irf8*, *pparg* and *cebpb*) were themselves differentially expressed over time, and their expression tracks the expression of their targets. The overall pattern of chromatin becoming more accessible at 120 minutes suggests that there are transcriptomic consequences of a brief territorial challenge that persist for more than two hours afterwards.

DISCUSSION

By integrating different types of transcriptomic and epigenomic data (RNASeq, H3K27Ac ChIPSeq) with rigorous computational analyses, we show heretofore underappreciated consequences of social interactions for the brain transcriptome and epigenome. We detected dramatic changes in gene expression over time in response to a brief territorial challenge: clusters of genes enriched for particular biological functions changed in a coordinated fashion, over a period extending for hours afterwards. Using a TRN and generalized linear model, we linked changes in gene expression to the activity of transcription factors, which we propose to be

factors that regulate them. Moreover, we demonstrate that a brief social interaction was sufficient to cause changes in the accessibility of chromatin elements throughout the genome, including many linked to DEGs. While conventional wisdom is that chromatin changes relatively slowly in adult tissues, there is some precedent for our findings of rapid response in adult brain; for example, epigenetic responses to strong stimuli such as cocaine can happen quickly, e.g. within an hour (Maze et al. 2011). Indeed, there is growing evidence from the learning and memory literature that chromatin can be very dynamic (Yang et al. 2016; Sweatt 2013; Halder et al. 2016; Hirano et al. 2016), and changes in chromatin accessibility in response to a social challenge have been reported in other species (Saul et al. 2017). The magnitude of epigenomic plasticity we observed in response to a territorial challenge is also noteworthy. Hundreds of genes were closely linked to differentially accessible chromatin peaks, and for many of these we found dramatic differences in the degree of accessibility, measured by enrichment for H3K27Ac, following a social interaction (Figure 3.4b). Indeed, a territorial challenge was sufficient to cause some genes that were not clearly associated with accessible chromatin prior to a territorial challenge to become accessible afterwards (Table B.1). We hypothesize that a territorial intrusion provoked dramatic responses at the transcriptomic and epigenomic level in male sticklebacks because successful territory defense is so consequential in this species, with strong implications for fitness.

Changes in gene expression over time were similarly dramatic, consistent with the hypothesis that there are waves of transcription associated with different temporal aspects of behavioral plasticity – some genes are involved in the initial reaction to a conspecific, others with responding to social information and still others involved in recovery and preparing for the future. The early hormonal response parallels time course of the neuroendocrine response to

aggression, which involves both the hypothalamic-pituitary-adrenal (interrenal in fishes) axis and the HPG axis in vertebrates (Wingfield et al. 1990), including in sticklebacks (Bell et al. 2007). Interestingly, prolactin – a hormone associated with maternal care – was also recruited in response to a territorial challenge. This result is consistent with the hypothesis that aggression and parental care share common neuroendocrine and neurogenomic substrates (Wingfield et al. 1990). The relatively large number of upregulated DEG and increased chromatin accessibility at the 120 minute time point implies that much of the neurogenomic response to a brief territorial intrusion is related to recovery and preparing for the future, rather than producing the immediate behavioral response. The increased accessibility and expression of genes related to immunity and learning at 120 minutes is consistent with this idea. For example, GO terms related to learning (calmodulin binding and calcium ion binding, involved in the activation of CamK) were enriched in clusters of genes that peak at 120 mins (D2 and D3). Similarly, the expression of *CAMKK1* (important for long term memory (Blaeser et al. 2006), for example, changed over time in response to a territorial challenge and was upregulated at 120 minutes. Finally, actin binding, important for late long term potentiation and long term memory (Chen et al. 2016), was implicated in the differentially accessible genes (Figure 3.4c). Indeed, there is growing appreciation for the relationship between immunity and learning, especially spatial learning (Filiano et al. 2016; Stamps & Krishnan 2001). Increased chromatin accessibility at 120 minutes is also consistent with the idea that the transcriptomic response to social interactions might be even faster in the future, i.e. priming.

The involvement of learning and memory-related genes makes sense in light of the biology of territorial animals (Hollis 1999). During an intrusion, territory holders gain information about the spatial boundaries of their territory, the competitive ability of their

neighbors and their own resource holding potential, and use that information to guide future behavior. Indeed, territorial animals improve their ability to detect and repel intruders with experience (Hollis 1999; Bronstein 1994; Jenkins & Rowland 1996; Losey & Sevenster 1995) and the behavioral literature is rife with examples of experience-dependent changes following a territorial challenge such as the prior residency advantage (Huntingford & Turner 1987), the winner effect (Hsu et al. 2006) and the dear enemy phenomenon (Temeles 1994). Social interactions during territory defense are especially likely to influence spatial learning. For example, fishes actively patrol sites where they've had previous encounters with intruders (Bronstein 1986; Bronstein 1988). We speculate that a brief territorial challenge triggers the expression of learning-related genes and that changes in chromatin are associated with the formation of memories of where the social interaction occurred.

A growing number of studies are implicating metabolic genes with aggression (Rittschof et al. 2014; Li-Byarlay et al. 2014; Rittschof et al. 2015; Chandrasekaran et al. 2015). Consistent with this, the expression and accessibility of peroxisome proliferator activated receptors gamma (*pparg*), which participates in the regulation of lipid metabolism and glucose homeostasis, changed over time in response to a territorial challenge. *Cebpb* is another transcription factor enriched in the $DAPDEG_x$ which directly binds at the *pparg* promoter and can regulate its expression (Kawai & Rosen 2010). Other studies have shown that *pparg* is expressed in the hypothalamus and is important for CNS energy balance (Ryan et al. 2011; Sarruf et al. 2009). For instance, *pparg* agonists, which are insulin-sensitization drugs such as thiazolidinedione (TZD), are widely prescribed to diabetes mellitus 2 patients (Ryan et al. 2011). *Pparg* and its targets are downregulated at 60 min and then up-regulated at 120 minutes, possibly reflecting changes in energy balance and homeostasis following an aggressive interaction.

From an ethological perspective it is staggering to consider these results in light of the richness of social animals' lives. Animals that live in social groups are constantly engaged in social interactions. Indeed, rates of territorial intrusions in natural populations of sticklebacks have been reported to be as high as 76 intrusions per hour (Bakker & Goldschmidt 1989). Moreover, territory holders interact not only with competitors but also with predators, potential mates and offspring. How animals in natural populations behave during these interactions influences their current and future social environment as well as their fitness. Our results prompt the intriguing hypothesis that meaningful social interactions (even brief ones) can provoke waves of transcription and changes to the epigenome which lead to changes in neural functioning, and those changes are a mechanism by which animals update their assessment of their social world.

METHODS

Animals

Adult males were collected from Putah Creek, a freshwater population, in spring 2013 and maintained in the lab on a 16:8 (L:D) photoperiod and at 18° C in separate 9-liter tanks. Males were provided with nesting material including algae, sand and gravel and were visually isolated from neighbors. All males were in the 'territorial' phase of the nesting cycle, i.e. defending a territory. Sneaking is rare in this population.

Territorial challenge

We employed a method to simulate a territorial challenge initially developed by van Iersel (Iersel 1958) and used in previous studies (Sanogo et al. 2012; Rittschof et al. 2014). Males were randomly assigned to either the experimental or control group. Males in the experimental group were presented with a smaller, unrelated male intruder confined to a flask.

Males in the control groups were presented with an empty flask. At the same time as a confined intruder was introduced to an experimental male's tank, an empty flask was introduced into a paired control male's tank. After 5 min the flask was removed, and after a predetermined period (see below) males were quickly netted and sacrificed by decapitation within seconds following an IACUC approved protocol (#15077) of the University of Illinois at Urbana-Champaign.

RNA Sequencing

Tissue Preparation - Males for RNA Sequencing were collected 30, 60 or 120 minutes after the flask was introduced, with n=10 males per time point. Heads were flash frozen in liquid nitrogen and the telencephalon and diencephalon were carefully dissected and placed individually in Eppendorf tubes containing 500 μL of TRIzol Reagent (Life Technologies). Total RNA was isolated immediately using TRIzol Reagent according to the manufacturer's recommendation and subsequently purified on columns with the RNeasy kit (QIAGEN). RNA was eluted in a total volume of 30 μL in RNase-free water. Samples were treated with DNase (QIAGEN) to remove genomic DNA during the extraction procedure. RNA quantity was assessed using a Nanodrop spectrophotometer (Thermo Scientific), and RNA quality was assessed using the Agilent Bioanalyzer 2100 (RIN 7.5-10). RNA was immediately stored at −80 °C until used in sequencing library preparation.

Library Preparation - Poly-A RNA was enriched from 1–2 μg of total RNA by using Dynabeads Oligo(dT)25 (Life Technologies), following the manufacturer's protocol. Two rounds of poly(A) enrichment were performed with a final elution in 14μL of water. The poly-A–enriched RNA was used to prepare RNA-seq libraries, using the NEXTflex Directional RNA-seq Kit (dUTP based) with Illumina compatible adaptors (Bio Scientific). Manufacturer's instructions were followed and 13–15 cycles of PCR amplification were performed depending on

the starting input of total RNA. Libraries were quantified on a Qubit fluorometer, using the dsDNA High Sensitivity Assay Kit (Life Technologies), and library size was assessed on a Bioanalyzer High Sensitivity DNA chip (Agilent). Libraries were pooled and diluted to a final concentration of 10 nM. Final library pools were quantified using real-time PCR, using the Illumina compatible kit and standards (KAPA) by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois). Single-end sequencing was performed on an Illumina HiSeq 2500 instrument by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois). The samples were sequenced on 20 lanes.

ChIP Sequencing

Tissue Preparation – Diencephalons from a new set of males were collected for ChIP-seq at 30 or 120 minutes after the flask was introduced. Prior to nuclei isolation, brain tissue was pooled into groups of 5 and kept at 0° C in PBS with Protease Inhibitor Cocktail (PIC, Roche 04693132001). Tissue was homogenized by motor pestle and then fixed in PBS+PIC with 1% formaldehyde for 10 minutes. The fixing reaction was stopped with addition of Glycine to a final concentration of 0.125M. Fixed cells were washed 2x with PBS+PIC to remove formaldehyde. Washed cells were lysed to nuclei with L1 lysis solution – 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, 0.1% v/v NP-40, 10% v/v glycerol, and protease inhibitor cocktail (PIC) – for 30 minutes on ice. Cell debris was washed away with PBS and PIC. Nuclei were then pelleted and frozen on dry ice. Prior to pelleting, a small (2 µL) sample of nuclei was taken, stained with Trypan, and checked for quality and quantification via hemacytometer. Nuclei were sonicated at high power for 7 x 7 minute cycles (30 s on, 30 s off) in a Diagenode Biorupter Sonicator

(Diagenode). Remaining cellular debris was pelleted by centrifugation for 10 minutes at 13,000 x g.

Fragmented chromatin was processed in duplicate for histone H3K27Ac ChIP with Diagenode iDeal ChIP kits, according to manufacturer's specifications with minor adjustments. Six million nuclei were used for each IP. 25 µl of each IP was reserved for input samples. Technical replicate inputs were pooled to 50 µl. 2 g of H3K27Ac antibody (Abcam ab4729) was used for each IP. An additional wash in TE buffer was performed after the initial four IP washes.

Library Preparation - After ChIP, IP DNA was quantified by Qubit with a dsDNA High Sensitivity quantification kit (Invitrogen). Libraries were prepared using KAPA LTP library kits, with protocol as written, using Bioo index adapters. Libraries were size selected using AmpureXP beads, with protocol as written, selecting for DNA between 200-500bp in size. Library quality was checked by a Qubit flourometer and Bioanalyzer. Samples were sequenced with an Illumina HiSeq 2500 sequencer.

RNA Seq Informatics

FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess the quality of the reads. Adaptor sequences and low quality bases were clipped from 100 bp single-end sequences using Trimmomatic. RNA-seq produced an average of ~59 million reads per sample. We aligned reads to the *Gasterosteus aculeatus* reference genome (the repeat masked reference genome, Ensembl release 75), using TopHat (2.0.8) and Bowtie (2.1.0) (Langmead & Salzberg 2012). Reads were assigned to features according to the Ensembl release 75 gene annotation file (http://ftp.ensembl.org/pub/release-75/gtf/gasterosteus_aculeatus/).

ChIP Seq Informatics

Libraries from each technical replicate and the input control were sequenced with average depth of 7.6 M reads with average quality score > 35. Technical replicates were pooled with average sequence depth of 16M reads. Sequence data were mapped with Bowtie2 (Langmead & Salzberg 2012) to the *Gasterosteus aculeatus* reference genome (the repeat masked reference genome, Ensembl release 75), using default settings, yielding 3.87-7.96 uniquely mapped reads (averaging approximately 1X whole genome coverage of the stickleback genome for each replicate). Mapped sequence data were analyzed for peaks using HOMER (Hypergeometric Optimization of Motif EnRichment) v4.7 (Heinz et al. 2010). Samples were converted into tag directories, and QC was performed using read mapping and GC bias statistics. Histone peaks were then called from the Tag Directories with default factor settings, except local filtering was disabled (-L 0) and input filtering was set at three-fold over background (-F 3), to increase the sensitivity of the peak calling and identify individual subunits of multi-histone peaks, identifying tens of thousands peaks for each sample with average tag counts ranging from 42.7-58.1. Replicates were assessed for correlation, displaying >80% correlation in these filtered peaks across the two samples, which were then pooled for final peak identification. Peaks were highly associated with annotated gene promoters with average distance to transcription start sites (TSS) ranging from 75-328.4 bp, as expected for H3K27Ac (Hon et al. 2009); these data confirmed the robustness of the ChIP-seq data. After peak calling, peak files were annotated to the stickleback genome using HOMER's annotation script to assign peaks to genes, and associate peaks with differential expression data. BigWiggle pileup files were generated using HOMER's makeBigWig.pl script with default settings.

Defining differentially expressed genes (DEGs)

HTSeq read counts were generated for genes using stickleback genome annotation. Any reads that fell in multiple genes were excluded from the analysis. We included genes with at least one counts per million (cpm) in at least two samples. Count data were TMM (trimmed mean of M-values) normalized in R using edgeR. To assess differential expression a nested interaction model (~time+treatment:time) was fitted separately for diencephalon and telencephalon in edgeR (see edgeR manual section 3.3.2). A tagwise dispersion estimate was used after computing common and trended dispersions. Finally, to call differential expression between treatment groups, a 'glm' approach was used. We FDR-adjusted the p-values from all contrasts at once. A FDR cutoff < 0.1 was used to call for differentially expressed genes.

Hierarchical Clustering Analysis

An agglomerative clustering was done separately on DEG_X from each brain region. A hierarchical dendogram was generated using hclust function in R (R version 3.2.2), whereas "ward.D" objective criterion was used to merge the pair of cluster at each step. Trees were cut at height 25 to obtain clusters. Each cluster's fold change values at each time point were plotted as profile plots using ggplot2 in R.

Defining differentially accessible peaks (DAPs)

H3K27Ac peaks and their differences between experimental and control groups were calculated at 30 minutes and 120 minutes in each brain region using HOMER's getDifferentialPeaks functionality. For each time point and brain region, two sets of results were calculated: one treating the experimental group as background and the other treating the control group as background. An H3K27ac peak was termed to be "differentially accessible" if it had a fold change of larger than 2 in either set of results, and if it had a p-value less than 10^{-4}.

Differential peak sets were then annotated using a custom R script to search for all transcripts with transcript start or end sites within 20 kb on all Ensembl-annotated splice variants built using biomaRt.

Associating DAPs with genes

A chromatin domain was defined for each gene in the Ensembl build (v.1.75) of the stickleback genome. First, for each transcript corresponding to a gene, a window was defined that began 20 kb upstream of the transcription start site and ended 20 kb downstream of the transcription end site. A 20 kb window was chosen based on the estimated intergene interval in the stickleback genome. Next, this window was truncated so that it did not intersect with any transcript of any other gene. The union of these windows for all transcripts of a gene constituted that gene's domain. All peaks that had any overlap with the domain of a gene were considered as potential regulators of that gene's expression.

Transcriptional regulatory network (TRN) analysis

ASTRIX uses gene expression data to identify regulatory interactions between transcription factors and their target genes. A previous study validated ASTRIX-generated TF-target associations using data from ModENCODE, REDfly and DROID databases. The predicted targets of TFs were defined as those genes that share very high mutual information ($P < 10^{-6}$) with a TF, and can be predicted quantitatively with high accuracy (Root Mean Square Deviation (RMSD) < 0.33 i.e prediction error less than 1/3rd of each gene expression profile's standard deviation. The list of putative TFs in the stickleback genome was obtained from the Animal Transcription Factor Database. Given TFs and targets sets ASTRIX infers a genome-scale TRN model capable of making quantitative predictions about the expression levels of genes given the expression values of the transcription factors. The ASTRIX algorithm was previously used to

infer a TRN model for the honeybee brain that showed remarkably high accuracy in predicting behavior-specific gene expression changes. ASTRIX identified transcription factors that are central actors in regulating aggression, maturation and foraging behaviors in the honey bee brain. Transcription factors that are predicted to regulate a cluster (from the hierarchical clustering analysis) were determined according to whether they had a significant number of targets in a cluster as assessed by a Bonferroni FDR-corrected hypergeometric test. TFs with at least 3 targets were used and a FDR cutoff of < 0.05 was used to call for significant associations.

Functional analysis

We derived GO assignments, using protein family annotations from the database PANTHER. Stickleback protein sequences were blasted against all genomes in the database (PANTHER 9.0 ~85 genomes). This procedure assigns proteins to PANTHER families based on structural information as well as phylogenetic information. Genes were then annotated using GO information derived from the ~82 sequenced genomes in the PANTHER database.

GO analysis were performed in R using TopGo v.2.16.0 and Fisher's exact test. A p-value cut off <0.01 was used to select for significantly enriched functional terms wherever possible. For visualization we found dissimilarity among GO terms using zebrafish as closest organism and "Wang" algorithm in GOSemSim package (Yu et al. 2010). We then plotted their similarity using the non-metric isoMDS function in MASS. We used the individual terms and the genes inside each term to manually annotate names for clusters appearing in MDS plots. This study has been submitted to GEO. The RNASeq data and ChIPSeq data are accessible with this GEO ID: GSE96673.

ACKNOWLEDGEMENTS

We thank members of the Bell lab, Saurabh Sinha and Gene Robinson for advice, comments and suggestions which have significantly improved this manuscript.

FIGURES AND TABLE

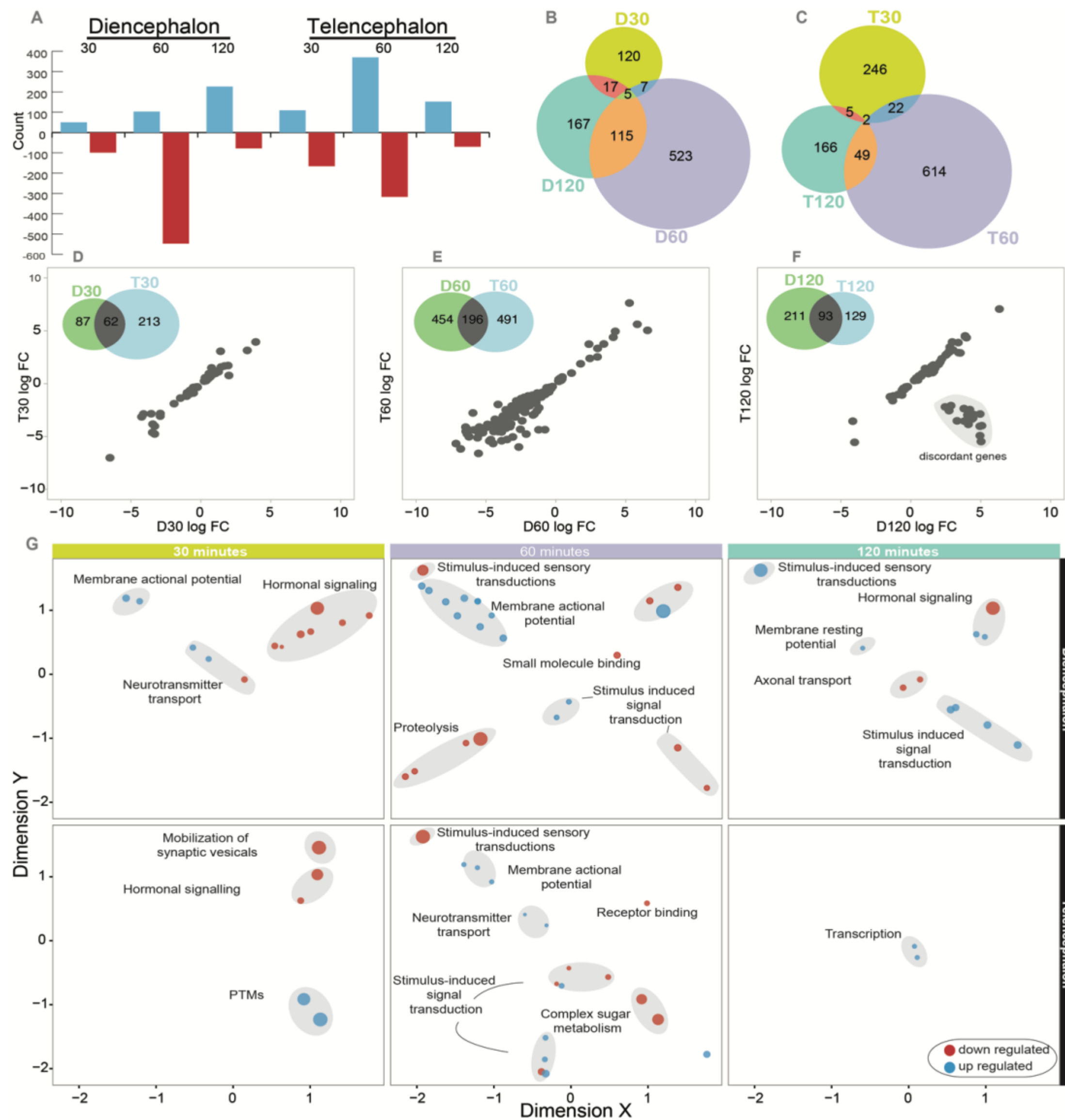


Figure 3.1 – Brain region-specific changes in gene expression in response to a territorial challenge over time. (A) Numbers of up- (blue) and down (red)-regulated genes at 30, 60 and 120 minutes after a territorial challenge in diencephalon and telencephalon. Overlap between differentially expressed genes across time points in diencephalon (B) and telencephalon (C). Correlation between expression in diencephalon (X axis) and telencephalon (Y axis) at 30 min (D), 60 min (E) and 120 min (F) after a territorial challenge. The numbers in the Venn diagram

Figure 3.1 (continued)

indicate the number of differentially expressed genes in each brain region and the overlap between them at a given time. Scatterplots show the expression pattern of the genes that were shared between brain regions at a time point. Note the cluster of genes in the lower right corner of 1f, hereafter referred to as 'discordant genes', which were differentially expressed in both brain regions at 120 minutes but in opposite directions: they were upregulated in diecenphalon and downregulated in telencephalon. (G) Functional enrichment of DEGs by time point (columns) and by brain region (rows), shown as revigo-like MDS graphs. Blue indicates enrichment of up-regulated genes, red indicates enrichment of down-regulated genes. Groups of terms with similar functions are highlighted.

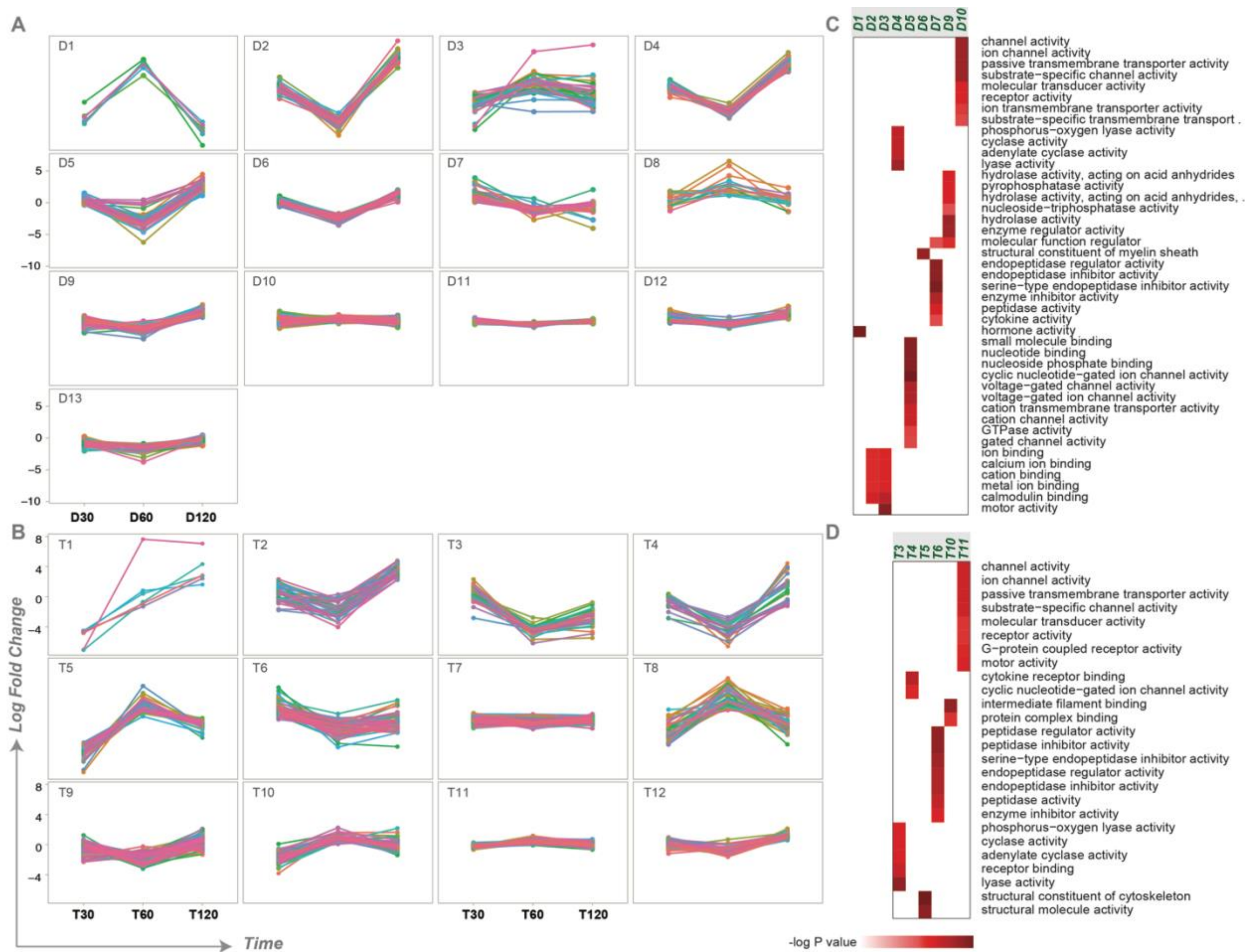


Figure 3.2 – Hierarchical clustering of genes whose expression profiles changed over time in response to a territorial challenge (DEG_x) and their functional enrichments. Hierarchical clustering grouped together genes with similar expression profiles over time. 13 clusters were identified in diencephalon (D1-D13, A). 12 clusters were identified in telencephalon (T1-T12, B). Each line represents the expression pattern of a different gene, where positive fold change indicates upregulation and negative fold change indicates downregulation in response to a territorial challenge. Clusters of genes with similar expression profiles (columns) had different GO molecular functions associated with them (rows); C) diencephalon; D) telencephalon. Some clusters did not have significant functional enrichment.

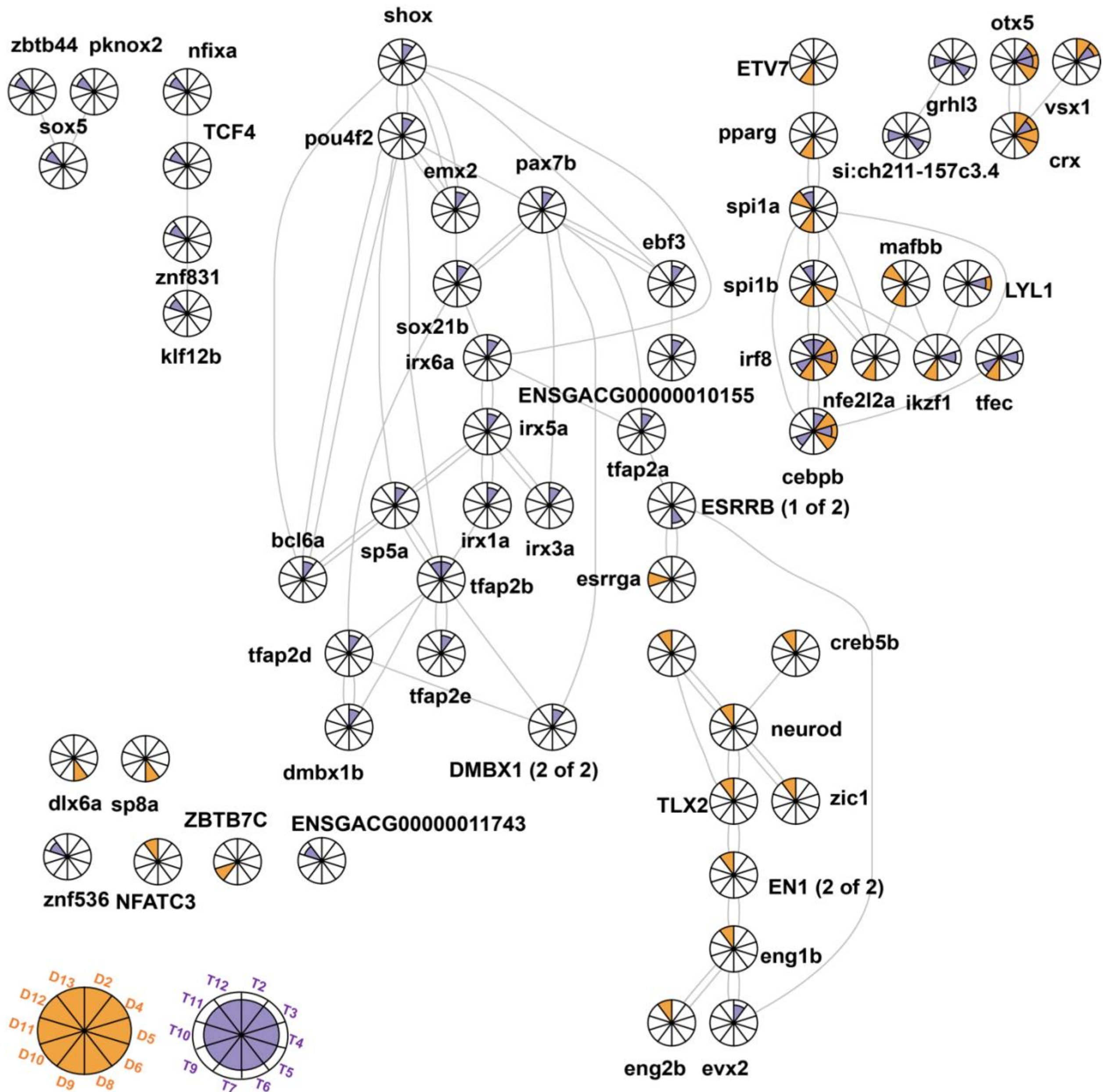


Figure 3.3 – Network of interacting transcription factors (TFs) in the transcriptional regulatory network highlighting enrichments of TFs in clusters of DEG_x. Each node represents a TF. Slices of pie correspond to different clusters in diencephalon or telencephalon; the key to the clusters is in the lower left corner. A full orange slice represents a diencephalon cluster. A purple half slice

Figure 3.3 (continued)

represents a telencephalon cluster, a purple and orange slice represents clusters in both brain regions. For example, *cebpb* is predicted to regulate D4, D5, D6, T3, T5, T9.

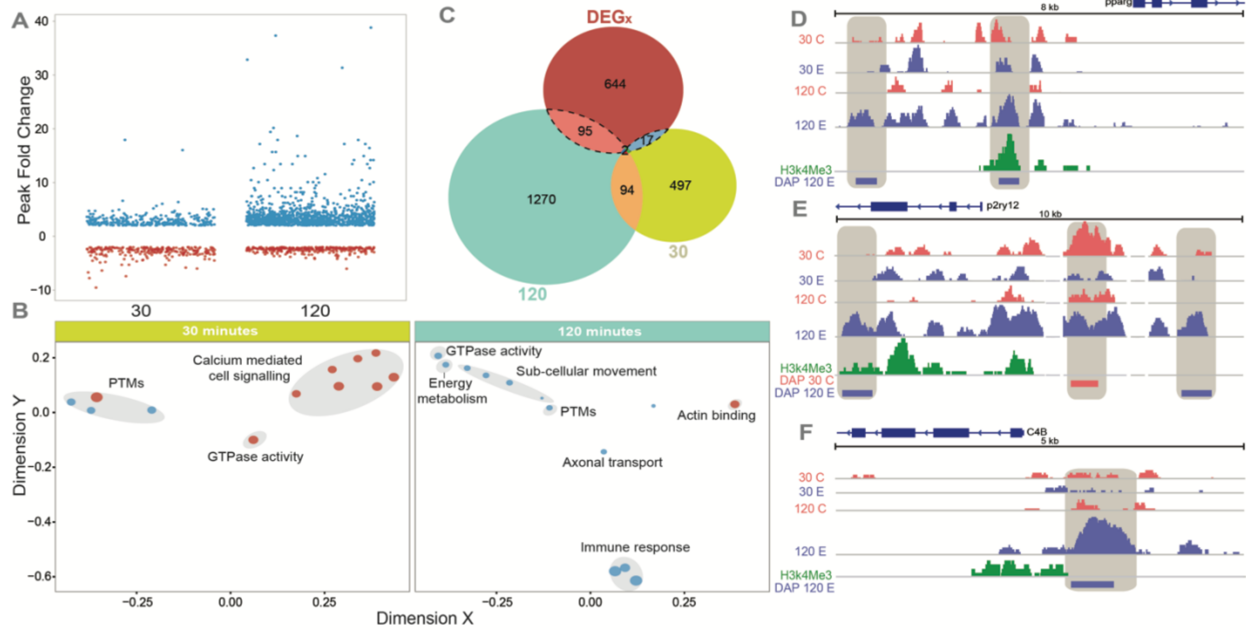


Figure 3.4 – Connecting gene expression and chromatin accessibility in diencephalon. (A) Fold change of differentially accessible peaks at 30 minutes and 120 minutes; blue indicates up in challenged, red indicates down in challenged. (B) Functional enrichment (molecular function) of genes associated with differentially accessible peaks at 30 minutes and 120 minutes. Blue indicates up in challenged, red indicates down in challenged. (C) Overlap of genes whose expression profile changed over time in response to a social interaction (DEG_x) with genes associated with differentially accessible peaks at 30 minutes and 120 minutes. The overlap between DEG_x and accessibility at 120 minutes is statistically significant ($P<0.0001$). (D-F) Examples of differentially accessible peaks around DEG_x. Separate tracks are shown for H3K27Ac peaks in control 30 min, experimental 30 min, control 120 min, experimental 120 min, and H3Kme3, which marks the location of the promoter. (D) *Pparg* (a TF in D9 and also present TRN) was more accessible at 120 minutes and was also up-regulated at 120 minutes. (E) *P2ry12* (cluster D9) is purinergic receptor involved in synaptic plasticity [64] that was more accessible in controls at 30 minutes then become more accessible in experimental animals at 120 minutes.

Figure 3.4 (continued)

P2ry12 is known to stimulate microglia migration toward neuronal damage [65]. (F) *C4B* (cluster D9) was not accessible at baseline but became accessible at 120 mins.

Table 3.1 – Integrating TFs with DEG_x and chromatin accessibility. These TFs are in the TRN and are enriched in the $DAPDEG_x$ with accessibility indicated. Some of the TFs (in bold) were differentially expressed and in a cluster. The general expression pattern of their cluster is indicated.

TF	Description	Cluster	Expression pattern	Accessibility
pparg	peroxisome proliferator activated receptor gamma	D9	Up, down, up	More accessible at 120E
ikzf1	IKAROS family zinc finger 1 (Ikaros)			Became accessible at 120E
ETV7	ets variant 7			More accessible at 120E
mafbb	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B, duplicate b			Became accessible at 120E
cebpb	CCAAT/enhancer binding protein (C/EBP), beta	D9	No change, down, up	More accessible at 120E
spi1b	spleen focus forming virus (SFFV) proviral integration oncogene spi1b			Became accessible at 120E
irf8	interferon regulatory factor 8	D9	No change, down, up	Became accessible at 120E

Table 3.1 (continued)

NFATC3	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3	D11	Down, down, no change	More accessible at 120E

LITERATURE CITED

Alaux, C. et al., 2009. Honey bee aggression supports a link between gene regulation and behavioral evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 106(36), pp.15400–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19706434 [Accessed October 10, 2016].

Amit, I. et al., 2009. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science (New York, N.Y.)*, 326(5950), pp.257–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19729616 [Accessed October 11, 2016].

Arbeitman, M.N. et al., 2002. Gene Expression During the Life Cycle of Drosophila melanogaster. *Science*, 297(5590), pp.2270–2275. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12351791 [Accessed May 24, 2018].

Arbeitman, M.N. et al., 2002. Gene expression during the life cycle of Drosophila melanogaster. *Science (New York, N.Y.)*, 297(5590), pp.2270–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12351791 [Accessed October 10, 2016].

Aubin-Horth, N. & Renn, S.C.P., 2009. Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Molecular Ecology*, 18(18), pp.3763–3780. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19732339 [Accessed May 24, 2018].

Bakker, T.C.M. & Goldschmidt, T., 1989. Determinants of Reproductive Success of Male Sticklebacks in the Field and in the Laboratory. *Netherlands Journal of Zoology*, 40(4), pp.664–687.

Bar-Joseph, Z., Gitter, A. & Simon, I., 2012. Studying and modelling dynamic biological processes using time-series gene expression data. *Nature Reviews Genetics*, 13(8), pp.552–

564. Available at: http://www.nature.com/doifinder/10.1038/nrg3244 [Accessed October 11, 2016].

Bell, A.M. et al., 2007. Variable neuroendocrine responses to ecologically-relevant challenges in sticklebacks. *Physiology & behavior*, 91(1), pp.15–25. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17321556 [Accessed October 11, 2016].

Bell, A.M. & Aubin-Horth, N., 2010. What can whole genome expression data tell us about the ecology and evolution of personality? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1560), pp.4001–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21078652 [Accessed May 24, 2018].

Blaeser, F. et al., 2006. Long-Term Memory Deficits in Pavlovian Fear Conditioning in Ca2+/Calmodulin Kinase Kinase -Deficient Mice. *Molecular and Cellular Biology*, 26(23), pp.9105–9115. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17015467 [Accessed May 24, 2018].

Bronstein, J.L., 1994. Conditional outcomes in mutualistic interactions. *Trends in ecology & evolution*, 9(6), pp.214–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21236825 [Accessed October 11, 2016].

Bronstein, P.M., 1988. Aggression, by John Klama. London: Longman Scientific, 1988, 169 pp. *Aggressive Behavior*, 14(6), pp.463–464.

Bronstein, P.M., 1986. Socially mediated learning in male Betta splendens. *Journal of comparative psychology (Washington, D.C. : 1983)*, 100(3), pp.279–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3769447 [Accessed October 10, 2016].

Burmeister, S.S., Jarvis, E.D. & Fernald, R.D., 2005. Rapid Behavioral and Genomic Responses to Social Opportunity J. Godwin, ed. *PLoS Biology*, 3(11), p.e363. Available at:

http://dx.plos.org/10.1371/journal.pbio.0030363 [Accessed October 10, 2016].

Carney, G.E. et al., 2007. A rapid genome-wide response to Drosophila melanogaster social interactions. *BMC Genomics*, 8(1), p.288. Available at: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-8-288 [Accessed October 10, 2016].

Chandrasekaran, S. et al., 2015. Aggression is associated with aerobic glycolysis in the honey bee brain. *Genes, brain, and behavior*, 14(2), pp.158–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25640316 [Accessed March 27, 2016].

Chandrasekaran, S. et al., 2011. Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. *Proceedings of the National Academy of Sciences of the United States of America*, 108(44), pp.18020–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21960440 [Accessed October 11, 2016].

Chen, Z., Stockwell, J. & Cayabyab, F.S., 2016. Adenosine A1 Receptor-Mediated Endocytosis of AMPA Receptors Contributes to Impairments in Long-Term Potentiation (LTP) in the Middle-Aged Rat Hippocampus. *Neurochemical Research*, 41(5), pp.1085–1097. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26700433 [Accessed May 24, 2018].

Cummings, M.E. et al., 2008. Sexual and social stimuli elicit rapid and contrasting genomic responses. *Proceedings. Biological sciences*, 275(1633), pp.393–402. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18055387 [Accessed December 22, 2016].

Filiano, A.J. et al., 2016. Unexpected role of interferon-γ in regulating neuronal connectivity and social behaviour. *Nature*, 535(7612), pp.425–429. Available at: http://www.nature.com/articles/nature18626 [Accessed May 24, 2018].

Fraser, B.A. et al., 2014. Phenotypic and genomic plasticity of alternative male reproductive

tactics in sailfin mollies. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1781).

Furukawa, T. et al., 1999. Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nature Genetics*, 23(4), pp.466–470. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10581037 [Accessed May 24, 2018].

Halder, R. et al., 2016. DNA methylation changes in plasticity genes accompany the formation and maintenance of memory. *Nature Neuroscience*, 19(1), pp.102–110. Available at: http://www.nature.com/articles/nn.4194 [Accessed May 24, 2018].

Heinz, S. et al., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell*, 38(4), pp.576–89. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20513432 [Accessed October 11, 2016].

Hirano, Y. et al., 2016. Shifting transcriptional machinery is required for long-term memory maintenance and modification in Drosophila mushroom bodies. *Nature Communications*, 7, p.13471. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27841260 [Accessed May 24, 2018].

Hollis, K.L., 1999. The Role of Learning in the Aggressive and Reproductive Behavior of Blue Gouramis, Trichogaster trichopterus. *Environmental Biology of Fishes*, 54(4), pp.355–369. Available at: http://link.springer.com/10.1023/A:1007529628117 [Accessed October 10, 2016].

Hon, G., Wang, W. & Ren, B., 2009. Discovery and Annotation of Functional Chromatin Signatures in the Human Genome E. Segal, ed. *PLoS Computational Biology*, 5(11), p.e1000566. Available at: http://dx.plos.org/10.1371/journal.pcbi.1000566 [Accessed May

24, 2018].

Hsu, Y., Earley, R.L. & Wolf, L.L., 2006. Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biological reviews of the Cambridge Philosophical Society*, 81(1), pp.33–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16460581 [Accessed October 10, 2016].

Hu, Z. et al., 1997. Steroidogenic factor-1 is an essential transcriptional activator for gonad-specific expression of promoter I of the rat prolactin receptor gene. *The Journal of biological chemistry*, 272(22), pp.14263–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9162060 [Accessed October 11, 2016].

Huang, Y. et al., 2011. Temporal Dynamics of Host Molecular Responses Differentiate Symptomatic and Asymptomatic Influenza A Infection N. J. Schork, ed. *PLoS Genetics*, 7(8), p.e1002234. Available at: http://dx.plos.org/10.1371/journal.pgen.1002234 [Accessed October 10, 2016].

Huntingford, F.A. & Turner, A.K., 1987. Conflict in the animal world. In *Animal Conflict*. Dordrecht: Springer Netherlands, pp. 3–12. Available at: http://www.springerlink.com/index/10.1007/978-94-009-3145-9_1 [Accessed October 11, 2016].

Iersel, Jja., 1958. Some aspects of territorial behaviour of the male three-spined stickleback. *Archives Néerlandaises de Zoologie*, (13), pp.381–400.

Jansen, M. et al., 2013. Gene expression profiling of three different stressors in the water flea Daphnia magna. *Ecotoxicology*, 22(5), pp.900–914. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23564370 [Accessed May 24, 2018].

Jenkins, J.R. & Rowland, W.J., 1996. Pavlovian conditioning of agonistic behavior in male

threespine stickleback (Gasterosteus aculeatus). *Journal of comparative psychology (Washington, D.C. : 1983)*, 110(4), pp.396–401. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8956509 [Accessed October 10, 2016].

Kawai, M. & Rosen, C.J., 2010. PPARγ: a circadian transcription factor in adipogenesis and osteogenesis. *Nature reviews. Endocrinology*, 6(11), pp.629–36. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20820194 [Accessed October 11, 2016].

Langmead, B. & Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), pp.357–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22388286 [Accessed October 11, 2016].

Lavergne, S.G. et al., 2014. Impact of high predation risk on genome-wide hippocampal gene expression in snowshoe hares. *Oecologia*, 176(3), pp.613–624. Available at: http://link.springer.com/10.1007/s00442-014-3053-0 [Accessed October 10, 2016].

Lawniczak, M.K.N. & Begun, D.J., 2004. A genome-wide analysis of courting and mating responses in Drosophila melanogaster females. *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada*, 47(5), pp.900–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15499404 [Accessed October 10, 2016].

Li-Byarlay, H. et al., 2014. Socially responsive effects of brain oxidative metabolism on aggression. *Proceedings of the National Academy of Sciences*, 111(34), pp.12533–12537. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1412306111 [Accessed March 27, 2016].

Losey, G.S. & Sevenster, P., 1995. Can three-spined sticklebacks learn when to display? Rewarded displays. *Animal Behaviour*, 49(1), pp.137–150.

Mack, P.D. et al., 2006. Mating-responsive genes in reproductive tissues of female Drosophila

melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 103(27), pp.10358–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16798875 [Accessed October 10, 2016].

Maze, I. et al., 2011. Cocaine dynamically regulates heterochromatin and repetitive element unsilencing in nucleus accumbens. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7), pp.3035–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21300862 [Accessed October 10, 2016].

McGraw, L.A., Clark, A.G. & Wolfner, M.F., 2008. Post-mating Gene Expression Profiles of Female Drosophila melanogaster in Response to Time and to Four Male Accessory Gland Proteins. *Genetics*, 179(3).

O'Connell, L.A. & Hofmann, H.A., 2012. Evolution of a Vertebrate Social Decision-Making Network. *Science*, 336(6085), pp.1154–1157.

Peeke, H.V.S., 1969. Habituation of Conspecific Aggression in the Three-Spined Stickleback (Gasterosteus Aculeatus L.). *Behaviour*, 35(1), pp.137–156. Available at: http://booksandjournals.brillonline.com/content/journals/10.1163/156853970x00178 [Accessed October 11, 2016].

Peeke, H.V.S. & Veno, A., 1973. *Stimulus specificity of habituated aggression in the stickleback (Gasterosteus aculeatus)*. Behavioral biology, 8(3), pp.427-432.

Pelkwijk, J.J. & Tinbergen, N., 1937. Eine reizbiologische Analyse einiger Verhaltensweisen von Gasterosteus aculeatus L. *Zeitschrift für Tierpsychologie*, 1(3), pp.193–200. Available at: http://doi.wiley.com/10.1111/j.1439-0310.1937.tb01422.x [Accessed October 11, 2016].

Rittschof, C.C. et al., 2014. Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proceedings of the National Academy of*

Sciences, 111(50), pp.17929–17934. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25453090 [Accessed May 20, 2018].

Rittschof, C.C., Grozinger, C.M. & Robinson, G.E., 2015. The energetic basis of behavior: Bridging behavioral ecology and neuroscience. *Current Opinion in Behavioral Sciences*, 6, pp.19–27.

Rittschof, C.C. & Robinson, G.E., 2013. Manipulation of colony environment modulates honey bee aggression and brain gene expression. *Genes, Brain and Behavior*, 12(8), pp.802–811. Available at: http://doi.wiley.com/10.1111/gbb.12087 [Accessed October 10, 2016].

Ryan, K.K. et al., 2011. A role for central nervous system PPAR-γ in the regulation of energy balance. *Nature medicine*, 17(5), pp.623–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21532595 [Accessed October 10, 2016].

Sanogo, Y.O. et al., 2011. Brain transcriptomic response of threespine sticklebacks to cues of a predator. *Brain, behavior and evolution*, 77(4), pp.270–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21677424 [Accessed October 10, 2016].

Sanogo, Y.O. et al., 2012. Transcriptional regulation of brain gene expression in response to a territorial intrusion. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), pp.4929–4938. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.2087 [Accessed November 11, 2016].

Sanogo, Y.O. et al., 2012. Transcriptional regulation of brain gene expression in response to a territorial intrusion. *Proceedings. Biological sciences / The Royal Society*, 279(1749), pp.4929–38. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23097509 [Accessed October 10, 2016].

Sarruf, D.A. et al., 2009. Expression of peroxisome proliferator-activated receptor-gamma in key neuronal subsets regulating glucose metabolism and energy homeostasis. *Endocrinology*, 150(2), pp.707–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18845632 [Accessed October 10, 2016].

Sato, Y., Hashiguchi, Y. & Nishida, M., 2009. Evolution of multiple phosphodiesterase isoforms in stickleback involved in cAMP signal transduction pathway. *BMC Systems Biology*, 3(1), p.23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19232106 [Accessed May 24, 2018].

Saul, M.C. et al., 2017. Transcriptional regulatory dynamics drive coordinated metabolic and neural response to social challenge in mice. *Genome research*, 27(6), pp.959–972. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28356321 [Accessed May 20, 2018].

Stamps, J.A. & Krishnan, V. V., 2001. How Territorial Animals Compete for Divisible Space: A Learning-Based Model with Unequal Competitors. *The American Naturalist*, 157(2), pp.154–169. Available at: http://www.journals.uchicago.edu/doi/10.1086/318634 [Accessed October 10, 2016].

Sweatt, J.D., 2013. The emerging field of neuroepigenetics. *Neuron*, 80(3), pp.624–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24183015 [Accessed October 10, 2016].

Temeles, E.J., 1994. The role of neighbours in territorial systems: when are they “dear enemies”? *Animal Behaviour*, 47(2), pp.339–350.

Wingfield, J.C. et al., 1990. The Challenge Hypothesis: Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. *The American Naturalist*, 136(6), pp.829–846. Available at: http://www.journals.uchicago.edu/doi/10.1086/285134 [Accessed October 11, 2016].

Yang, Y. et al., 2016. Chromatin remodeling inactivates activity genes and regulates neural coding. *Science*, 353(6296).

Yu, G. et al., 2010. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics (Oxford, England)*, 26(7), pp.976–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20179076 [Accessed October 11, 2016].

CHAPTER 4: NEUROGENOMIC INSIGHTS INTO PATERNAL CARE AND ITS CONNECTION TO TERRITORIAL AGGRESSION

ABSTRACT

Decades of research have shown that motherhood is a period of intense behavioral and neural activation. However, we know less about the molecular mechanisms that accompany the transition to fatherhood. Here we evaluate changes in neurogenomic state during paternal care in male threespined sticklebacks, a species in which males are the sole providers of parental care that is necessary for offspring survival. In addition to characterizing the neurogenomic landscape of paternal care, we also compare and contrast the neurogenomics of paternal care with the neurogenomic response to a territorial challenge. We compared the brain gene expression profiles of males before, during and after they became parents, relative to the appropriate controls. Gene expression was measured using RNA-Seq in two brain areas (telencephalon and diencephalon) containing nodes within the social behavior network. Males experienced dramatic changes in neurogenomic state as they became fathers. Genes related to hormones that change in mammalian mothers during pregnancy and maternal care were differentially expressed in stickleback fathers. Continuity and specificity of brain gene expression was observed across stages, some of which might be analogous to the changes associated with female pregnancy, parturition and postpartum periods in mammals. Finally, gene regulatory analysis nominated regulators which are responsive to both paternal care and a territorial challenge. Transcription factors that are predicted to regulate genes that were differentially expressed as a function of paternal care and a territorial challenge were regulated in opposite directions in different brain regions in the two experiments. This analysis sheds light onto the neurogenomic dynamics of

paternal care and offers a glimpse into how differential modulation of the social decision-making network in the brain can generate responses to opposing social stimuli.

INTRODUCTION

In many species, parents provide care for their offspring, which can improve offspring survival. There is fascinating diversity in the ways in which parents care for their offspring, from the piggyback behavior of poison dart frogs and spiders to parental provisioning of offspring in burying beetles and birds (Clutton-Brock 1991). The burden of parental care does not always land exclusively on females, indeed in some species both parents provide care and in other males are solely responsible for care.

Our understanding of the molecular and neuroendocrine basis of parental care has been largely influenced by studies in mammals, where maternal care is the norm. In mammals, females experience a series of changes including cycles of estrus, pregnancy, child birth and lactation as they become mothers, all of which are coordinated by pulses and the sustained release of hormones. Males also experience dramatic changes in physiology and behavior as they become fathers, even though they do not experience cycles of estrus, pregnancy, child birth and lactation. For example, men experience increased oxytocin (Gordon et al. 2010) and a drop in testosterone (Storey et al. 2000) following the birth of a child.

Fishes are particularly good subjects for understanding the molecular orchestrators of paternal care. Unlike mammals, paternal care is relatively common in fishes (30-80% of care-giving fish species are paternal) (Gross & Sargent 1985). Moreover, the basic building blocks of parental care are ancient and deeply conserved in vertebrates. For example, the hormone prolactin was named for its essential role in lactation in mammals, but had functions related to parental care in fishes long before mammals evolved (Whittington & Wilson 2013). Similarly,

growing evidence for deep homology of brain circuits related to social behavior (Newman 1999; Goodson 2005; O'Connell & Hofmann 2011a) suggests that the evolution of conserved genes operating within conserved neural circuits has led the diversity of parental care among vertebrates today.

In addition to providing care, parents must be vigilant to defend their vulnerable dependents from potential predators or other threats. Tradeoffs between parental care and territory defense have been particularly well studied in the ecological literature, e.g. (Ketterson et al. 1992), and represent the extremes on a continuum of social behavior – from strongly affiliative to strongly aggressive. The challenge hypothesis originally posited that patterns of testosterone secretion in birds reflects tradeoffs between territory defense and parental care, assuming that testosterone is incompatible with parental care in males (Wingfield et al. 1990). Subsequent studies have shown that testosterone is not always inhibitory of parental care (Hirschenhauser & Oliveira 2006), and that a territorial challenge activates gene regulatory pathways that do not depend on the action of testosterone (Rosvall & Peterson 2014). Regardless of the specific neuromodulators or hormones, a mechanistic link between territory defense and parental care is likely to operate through the social behavior network in the brain because most nodes of this network express receptors for neuromodulators and hormones that are involved with both aggression and parental care (Cardoso et al. 2015).

In this study, we tracked the neurogenomic shifts that accompany the transition to fatherhood in male stickleback fish. In this species, fathers are solely responsible for the care of the developing offspring, and paternal care is necessary for offspring survival. Male sticklebacks go through a predictable series of changes as they become fathers, from territory establishment and nest building to mating, caring for eggs, hatching and caring for fry (Wootton 1984). Each

stage is characterized by a particular set of behaviors and events, and the transition to the next stage depends on the successful completion of the preceding stage. We sampled males across these stages and measured brain gene expression via RNA-Seq in two brain regions that contain nodes within the social behavior network, diencephalon and telencephalon. Paternal care in sticklebacks is costly both in terms of time and energy (Smith & Wootton 1999), infanticide and cannibalism are common (Wootton 1984), and males must be highly vigilant to challenges from predators and rival males throughout the process.

The time-series sampling approach used in this experiment allowed us to capture the temporal dynamics of stability and change at the molecular level during the process of becoming a father. Conceptually, the temporal dynamics of the neurogenomic response across stages can be explained by considering four simple models (Figure 4.1). The first model posits that fatherhood has a particular neurogenomic state that is activated as soon as males become reproductively active, and which remains activated throughout the process. This model predicts that there is no effect of stage on brain gene expression, i.e. similar neurogenomic states across stages. The second model posits that each stage has a unique neurogenomic state associated with it, such that the neurogenomic state of males with nests is entirely different from the neurogenomic state of males with eggs, or males carrying for fry. According to this model, there is a strong effect of stage on brain gene expression and very little to no overlap or carryover between differentially expressed genes across stages. The third model posits that there are genes associated with having a nest, a set of genes associated with caring for eggs, a set of genes associated with caring for fry, etc., and genes for the next stage are added to the previous set as males go through the nesting cycle. According to this model, when males are caring for eggs, the “egg” genes are added to the previously activated “nest” genes, and so on in an additive fashion.

This simplistic hypothesis predicts that there is a strong carryover signal across stages. Finally, a fourth model is intermediate between models 1-3, allowing for carryovers, additivity and unique genes to each stage. According to this model, there are genes that are activated early in the process and remain "on", and other genes that are only recruited during particular stages. Genes whose expression persists over time could be involved in maintaining the previous neurogenomic state, while genes exhibiting a transient expression pattern could be involved in facilitating the next stage, priming and/or responding to a particular event or stimulus during that stage, e.g. the arrival of offspring.

In addition to characterizing the neurogenomic landscape of paternal care, an additional goal of this study was to compare and contrast the neurogenomics of paternal care with the neurogenomic response to a territorial challenge (Bukhari et al. 2017). In a previous study, we employed a time-series sampling approach to measures the transcriptomic response to a territorial challenge. Brain gene expression was measured 30, 60 or 120 minutes after a five-minute territorial challenge, which reflects the temporal dynamics of the important biology of male sticklebacks during territorial interactions with neighbors and intruders. As both territorial aggression and paternal care are social behaviors and utilize the similar circuitary of the social behavior network, we might expect to observe similarities between a territorial challenge and paternal care at the molecular level. However given their position at opposite ends of the continuum of social behavior, along with neuroendocrine tradeoffs as predicted by the challenge hypothesis (Wingfield, Hegner et al. 1990), we were also interested in whether theire is evidence for their opposition at the gene regulatory level.

RESULTS AND DISCUSSION

Neurogenomic dynamics of paternal care

Males experienced dramatic neurogenomic shifts while they were providing paternal care. A large number of genes – almost 10% of the transcriptome – were differentially expressed over the course of the breeding cycle (Figure 4.2A; APPENDIX C). A comparable number of genes were up- and down-regulated at each stage. There were significant gene expression differences between the control and experimental groups within both brain regions; relatively more genes were differentially expressed in diencephalon.

Genes related to hormones that change in mammalian mothers during pregnancy and maternal care were differentially expressed in stickleback fathers. For example, in mammals, levels of progesterone, estrogen and their receptors increase during pregnancy and then subside after childbirth. A similar pattern was observed in the diencephalon of male sticklebacks: both *Esr* and *Pgr* were upregulated during early hatching and then subsided (Figure 4.2B). Oxytocin (and its teleost homolog isotocin) is another important hormone that has been well-studied for its role in social affiliation and parental care in mammals (Gordon et al. 2010; Galbally et al. 2011) and fish (Kleszczyńska et al. 2007; Kleszczyńska et al. 2012; O'Connell et al. 2012; Kulczykowska & Kleszczyńska 2014; Lema et al. 2015). The gene for isotocin was upregulated in diencephalon when male sticklebacks were caring for eggs in their nests, and upregulated in telencephalon mid-way through the hatching process. Altogether these patterns suggest that paternal care involves significant neurogenomic shifts in stickleback males. Moreover, commonalities with what is known about maternal care in mammals suggest that the neurogenomic shifts during paternal care in a fish are deeply conserved and are not sex-specific.

Functional enrichment analysis of the differentially expressed genes (DEGs) suggests that paternal care required energy metabolism in the brain along with modifications of immune system and transcription. Genes associated with the immune response were down regulated in both brain regions and during most stages. Genes associated with energy metabolism and the adaptive component of the immune response were upregulated in telencephalon. Genes associated with the stress response were downregulated in both brain regions around the day of hatching. And, genes associated with energy metabolism were downregulated as fry emerged (Figure 4.2C).

Two genes implicated with infanticide and parental care in mammals (galanin receptor and progesterone receptor) were differentially expressed in fathering sticklebacks. Galanin and galanin-like peptide neurons regulate both feeding and parental behavior in mice, and the inhibition of infanticidal behavior in mice is facilitated by galanin neurons in the preoptic area of the hypothalamus (Wu et al. 2014). In this study, levels of expression of the galanin gene in diencephalon (which includes the preoptic area) was relatively high during the nest, eggs and early hatching stages. In contrast, the galanin receptor gene was downregulated during the middle to late hatching stages in both brain regions (Figure 4.2B). In male mice, progesterone and progesterone receptor mediate aggressive behavior toward pups and the suppression of parental behavior (Schneider et al. 2003). In this study, the expression of the *Pgr* gene gradually dropped as hatching progressed in both regions (Figure 4.2B). These patterns could reflect parents' need to manage tradeoffs between feeding and parental care (Fischer & O'Connell 2017), which is a particularly acute problem for fishes, where egg cannibalism is common. Specifically, both up-regulation of galanin during the egg stage and down-regulation of

progesterone receptor during the hatching stage could reflect how male sticklebacks inhibit their cannibalistic behavior while providing care.

Gene expression carryover and uniqueness across stages

Within each brain region, some DEGs were shared across stages of paternal care while other DEGs were unique to a stage. In general, stages closer in time had more overlapping DEGs than stages further apart in time (Figure 4.3A).

To explore whether a previously acquired neurogenomic state persisted into subsequent stages, we looked at the expression profiles of genes that were shared between successive stages. In particular, we wanted to know how many of the genes that were differentially expressed in one stage remained differently expressed in the subsequent stages. To do this, we generated lists of genes that were differentially expressed between the control and experimental group at each stage within each brain region. Then, we examined the overlapping DEGs between stages, and examined the expression profile of these "carryover genes" in a heatmap.

This analysis showed that the genes that were shared across stages tend to be similarly regulated over time (Figure 4.3B, D). For example, shared DEGs that were upregulated in males that had nests were also upregulated at subsequent stages, especially during stages close in time to the nesting stage. Similarly, there was a transcriptional signal of "eggs" which persisted after the "egg" stage. These patterns suggest that the events and behaviors that characterize a particular stage of paternal care (e.g. finishing a nest, the arrival of eggs, hatching) trigger a neurogenomic state that persists, perhaps for as long as those events and behaviors continue. The similar regulation across stages suggests that a previously acquired neurogenomic state is maintained into subsequent stages.

There were also genes that were unique to each stage. Those "unique" genes tended to exhibit an expression pattern that was stage-specific (Figure 4.3C, D). Genes exhibiting a transient stage-specific expression pattern might be involved in facilitating the next stage, priming and/or responding to a particular event or stimulus during that stage, e.g. the arrival of offspring. A recent study in mice compared brain gene expression between pregnant, post partum and virgin females and found a large number of overlapping genes between the pregnant and postpartum stages, which suggests that the signal of pregnancy carries over during maternal care (Ray et al. 2016). Overall our results show that the neurogenomic dynamics of paternal caregiving shows elements of both continuity and change, and are consistent with changes in brain gene expression as a function of pregnancy, childbirth and the postpartum period in mammalian females.

Tradeoffs between paternal care and a territorial challenge at the molecular level

To better understand how different social demands are resolved in the brain at the level of gene regulation, we compared these data to a previous study on the neurogenomic response to a territorial challenge in male sticklebacks (Bukhari et al. 2017). The two experiments are at the opposite ends of a continuum of social behavior: a territorial challenge provokes aggressive behavior while paternal care provokes affiliative behavior. Commonality at the molecular level could reflect shared mechanisms associated with these opposing social behaviors, while genes that are specific to an experiment could reflect the unique biology of territorial aggression versus paternal care. Given the different time courses in the two experiments (30, 60, 120 minutes after a territorial challenge versus five stages of paternal care over approximately one week), our comparison of the two experiments is likely to be conservative.

To look at commonalities at the gene level, we pooled DEGs (FDR < 0.01) across time points, stages and brain regions within each experiment, which resulted in two sets of genes associated with either a territorial challenge or paternal care (Figure 4.4A). There were 177 genes that were shared between the two experiments (Figure 4.4B); this overlap is highly statistically significant (hypergeometric test, fdr < 1e-10, APPENDIX D).

Genes that were specific to either a territorial challenge or paternal care were enriched with non-overlapping functional categories (APPENDIX D). For example, some of the genes that were specific to a territorial challenge were related to sensory and perception, whereas some of the genes that were unique to paternal care were related to energy metabolism and biosynthesis, which might reflect the high metabolic needs of males as they are providing care (Smith & Wootton 1999).

The large number of genes that were differentially expressed both in response to a territorial challenge and during paternal care prompted us to test for evidence of their common regulation at the gene regulatory level. Therefore, we used the data from both experiments to build a transcriptional regulatory network and asked if there are transcription factors whose targets are significantly associated with the DEG sets from the territorial challenge experiment, the paternal care experiment or both experiments (Figure 4.4F). There were 10 transcription factors that were significantly enriched in both experiments. 8 out of 10 transcription factors were regulated in opposite directions in the two experiments (Figure 4.5). Transcription factors like *NR3C1* and *klf7b* has been implicated with context to social behavior. For instance, NR3C1 codes for glucocorticoid receptor which is involved in several aspects of HPA axis modulation and has been implicated in psychosocial stress during pregnancy (Palma-Gudiel et al. 2015). Whereas krupple like factors has been linked with neurodevelopment disorders such intellectual

disability and austim spectrum disorder (Powis et al. 2018). These patterns suggest that different salient experiences – providing paternal care and territorial aggression –trigger opposite gene regulatory responses.

Interestingly, the transcription factors showing the opposite expression pattern were differentially expressed in different brain regions in the two experiments. Specifically, shared transcription factors and their predicted targets were *up*-regulated in telencephalon in response to a territorial challenge and *down*-regulated in diencephalon during parental care. A similar pattern was observed at the transcriptomic (rather than gene regulatory) level when neurogenomic states were compared between territorial aggression and courtship in male threespined sticklebacks (Sanogo & Bell 2016). These patterns suggest that paternal care and territory defense differentially modulate the social behavior network (Newman 1999; Goodson 2005; O'Connell & Hofmann 2011b) in the brain. While previous studies have explored circuit-level changes in the social behavior network in response to different social stimuli (Newman 1999), our results point to the molecular basis of differential modulation of the social behavior network.

CONCLUSION

Altogether this study shows that paternal care involves significant neurogenomic changes in stickleback males. Commonalities with what is known about maternal care in mammals suggests that some of the neurogenomic shifts during paternal care in a fish are deeply conserved and are not sex-specific. Finally, we show that both a territorial challenge and paternal care share the same underlying gene regulatory machinery, most of which is regulated in opposite directions, possibly to account for tradeoffs between these opposing social behaviors.

METHODS

Animals

Adult males were collected from Putah Creek, a freshwater population, in spring 2013 and maintained in the lab on a 16:8 (L:D) photoperiod and at 18° C in separate 9-liter tanks. Males were provided with nesting material including algae, sand and gravel and were visually isolated from neighbors.

To track transcriptional dynamics associated with becoming a father, in this experiment we sampled males for brain gene expression profiling at five different points during the reproductive cycle (n=5 males per time point): nest, eggs, early hatching, middle hatching and late hatching (control: reproductively adult males with no nests). Males in the "nest" condition had a nest but had not yet mated. Males in the "eggs" condition were sampled 4 days after their eggs were fertilized. Because males in the "eggs" condition were sampled 4 days after mating, the transcriptomic effects of mating are likely to have attenuated by the time males were sampled at this stage. Hatching takes place over the course of the fifth day after fertilization, and a previous study found that brain activation as assessed by *Egr-1* expression was highest while male sticklebacks were caring for fry as compared to males with nests or eggs (Kent & Bell 2018). To capture males' response to the new social stimulus of their fry, we focused on three time points on the day of hatching which capture the start of the hatching process (9am), when approximately half of the clutch is hatched (1pm) and when all of the eggs have hatched (5pm). Males in the nest, eggs and early hatching conditions were sampled at 9am, males in the mid-hatching condition were sampled at 1pm and males in the late hatching condition were sampled at 5pm.

Males in these conditions were compared to reproductively mature circadian-matched control males that did not have a nest (n=5 males per control group). Wild-caught females from the same population were used as mothers. Males were quickly netted and sacrificed by decapitation within seconds following an IACUC approved protocol (#15077) of the University of Illinois at Urbana-Champaign.

RNA Sequencing

Tissue Preparation - Heads were flash frozen in liquid nitrogen and the telencephalon and diencephalon were carefully dissected and placed individually in Eppendorf tubes containing 500 µL of TRIzol Reagent (Life Technologies). Total RNA was isolated immediately using TRIzol Reagent according to the manufacturer's recommendation and subsequently purified on columns with the RNeasy kit (QIAGEN). RNA was eluted in a total volume of 30 µL in RNase-free water. Samples were treated with DNase (QIAGEN) to remove genomic DNA during the extraction procedure. RNA quantity was assessed using a Nanodrop spectrophotometer (Thermo Scientific), and RNA quality was assessed using the Agilent Bioanalyzer 2100 (RIN 7.5-10). RNA was immediately stored at −80 °C until used in sequencing library preparation.

Library Preparation - Poly-A RNA was enriched from 1–2 µg of total RNA by using Dynabeads Oligo(dT)25 (Life Technologies), following the manufacturer's protocol. Two rounds of poly(A) enrichment were performed with a final elution in 14µL of water. The poly-A–enriched RNA was used to prepare RNAseq libraries, using the NEXTflex Directional RNA-seq Kit (dUTP based) with Illumina compatible adaptors (Bioo Scientific). Manufacturer's instructions were followed and 13–15 cycles of PCR amplification were performed depending on the starting input of total RNA. Libraries were quantified on a Qubit fluorometer, using the dsDNA High Sensitivity Assay Kit (Life Technologies), and library size was assessed on a

Bioanalyzer High Sensitivity DNA chip (Agilent). Libraries were pooled and diluted to a final concentration of 10 nM. Final library pools were quantified using real-time PCR, using the Illumina compatible kit and standards (KAPA) by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois). Single-end sequencing was performed on an Illumina HiSeq 2500 instrument by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois). The samples were sequenced on 27 lanes.

RNA Seq Informatics

FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess the quality of the reads. Adaptor sequences and low quality bases were clipped from 100 bp single-end sequences using Trimmomatic. RNA-seq produced an average of 60 million reads per sample. We aligned reads to the *Gasterosteus aculeatus* reference genome (the repeat masked reference genome, Ensembl release 75), using TopHat (2.0.8) (Kim et al. 2013) and Bowtie (2.1.0) (Langmead & Salzberg 2012). Reads were assigned to features according to the Ensembl release 75 gene annotation file (http://ftp.ensembl.org/pub/release-75/gtf/gasterosteus_aculeatus/).

Defining differentially expressed genes (DEGs)

HTSeq read counts were generated for genes using stickleback genome annotation. Any reads that fell in multiple genes were excluded from the analysis. We included genes with at least one counts per million (cpm) in at least two samples. Count data were TMM (trimmed mean of M-values) normalized in R using edgeR. To assess differential expression, pairwise comparisons between experimental and control conditions were made at each stage using appropriate circadian controls. Diencephalon and telencephalon were analyzed separately in edgeR. A

tagwise dispersion estimate was used after computing common and trended dispersions. To call differential expression between treatment groups, a 'glm' approach was used. We adjusted actual p-values via empirical FDR, where a null distribution of p-values was determined by permuting sample labels for 500 times for each tested contrast and a false discovery rate was estimated (Storey & Tibshirani 2003).

Unique DEGs

One of the goals of this study was to identify genes that uniquely characterize a particular condition, e.g. to a particular stage of paternal care, or to either the territorial challenge or the paternal care experiment. To address the possibility that putative "unique" genes barely passed the cutoff for differential expression in another condition (false negatives), we adopted an empirical approach. We kept the cutoff for DEGs at the focal condition at eFDR < 0.01 and relaxed the FDR threshold on the other conditions to eFDR < 0.20. This procedure was repeated for each condition and in each brain region separately.

Transcriptional regulatory network (TRN) analysis

ASTRIX uses gene expression data to identify regulatory interactions between transcription factors and their target genes. A previous study validated ASTRIX-generated TF-target associations using data from ModENCODE, REDfly and DROID databases (Chandrasekaran et al. 2011). The predicted targets of TFs were defined as those genes that share very high mutual information ($P < 10^{-6}$) with a TF, and can be predicted quantitatively with high accuracy (Root Mean Square Deviation (RMSD) < 0.33 i.e prediction error less than 1/3rd of each gene expression profile's standard deviation. The list of putative TFs in the stickleback genome was obtained from the Animal Transcription Factor Database. Given TFs and targets sets ASTRIX infers a genome-scale TRN model capable of making quantitative predictions

about the expression levels of genes given the expression values of the transcription factors. The ASTRIX algorithm was previously used to infer a TRN models for honeybee, mouse and sticklebacks (Chandrasekaran et al. 2011; Saul et al. 2017; Shpigler et al. 2017; Bukhari et al. 2017). ASTRIX identified transcription factors that are central actors in regulating aggression, maturation and foraging behaviors in the honeybee brain (Chandrasekaran et al. 2011).

Here we have used ASTRIX to infer a joint gene regulatory network by combining gene expression profiles from a previous study on the transcriptomic response to a territorial challenge in male sticklebacks (Bukhari et al. 2017) with the data from this experiment. Combining the two datasets should increase statistical power to help identify modules that are shared and unique to the two experiments. Transcription factors that are predicted to regulate differentially expressed genes in either experiment were determined according to whether they had a significant number of targets as assessed by a Bonferroni FDR-corrected hypergeometric test.

Functional analysis

We derived GO assignments, using protein family annotations from the database PANTHER (Mi et al. 2016). Stickleback protein sequences were blasted against all genomes in the database (PANTHER 9.0 85 genomes). This procedure assigns proteins to PANTHER families on the basis of structural information as well as phylogenetic information. Genes were then annotated using GO information derived from the 85 sequenced genomes in the PANTHER database.

GO analysis were performed in R using TopGo v.2.16.0 and Fisher's exact test. A p-value cut off of <0.01 was used to select for significantly enriched functional terms wherever possible.

The datasets generated during and/or analysed during the current study will be available in GEO upon publication; they are available from the corresponding author on request.

FIGURES

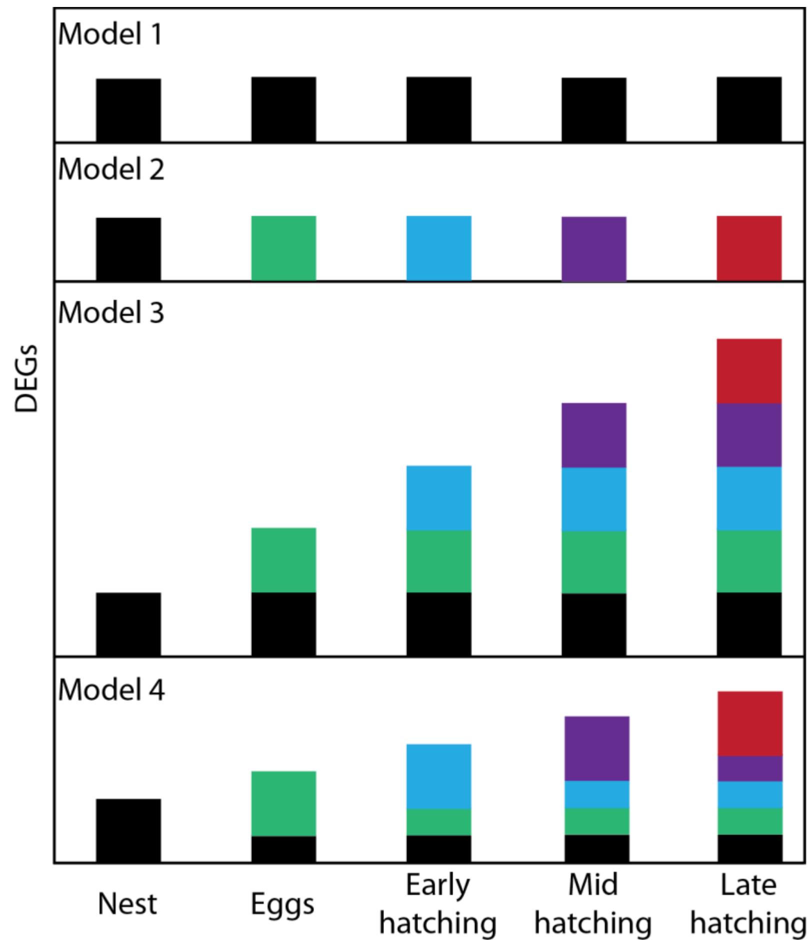
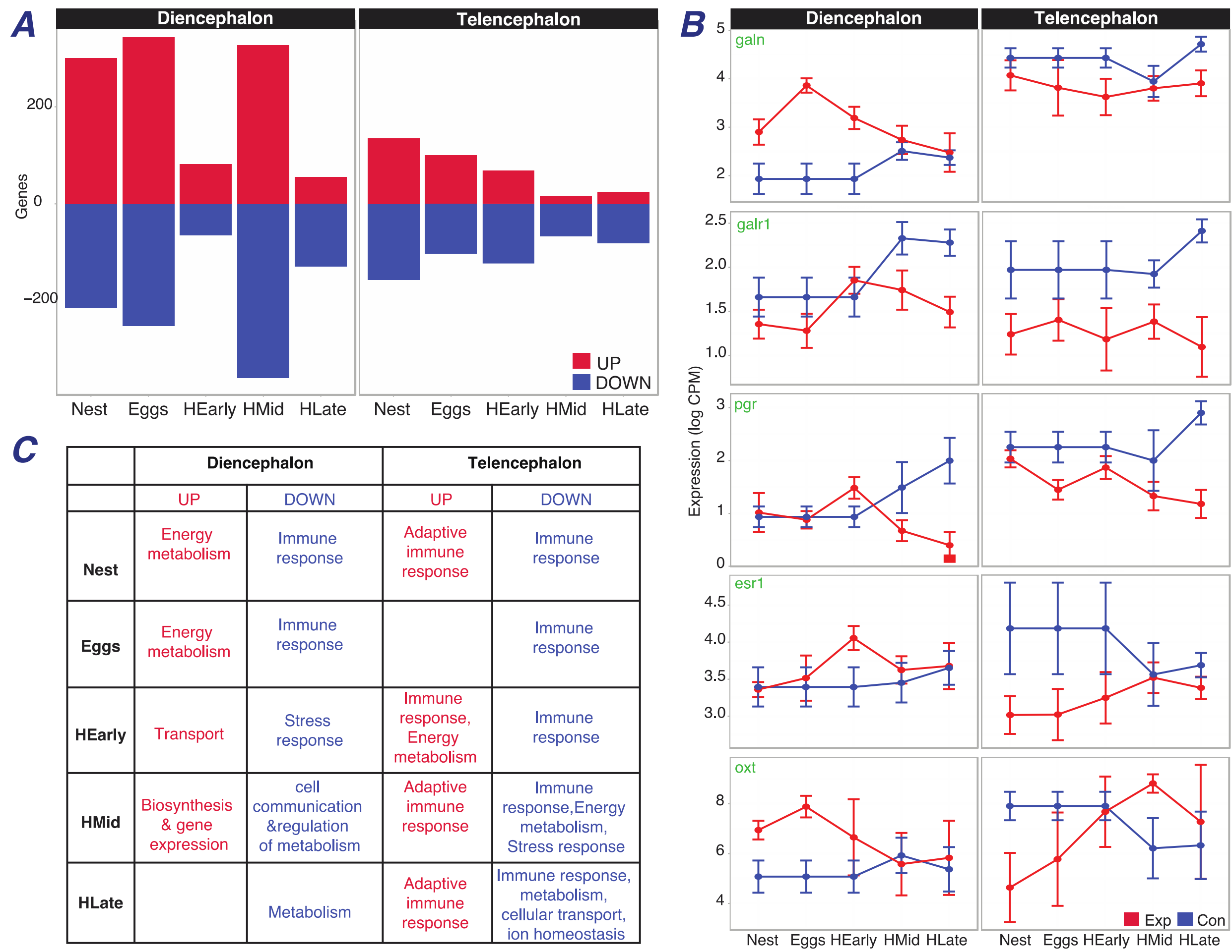


Figure 4.1 – Conceptual models to explain the neurogenomic dynamics across different stages of parenting. Model 1 suggests no effect of stage on neurogenomic state. Model 2 suggests strong effect of stage on neurogenomic state and there is no overlap across stages. Model 3 suggests additive carryover e.g., genes for the next stage are added to the previous set of as males go through the nesting cycle. Model 4 is between 1-3 which allow for carryover, additivity and unique genes to each stage.



	Diencephalon		Telencephalon	
	UP	DOWN	UP	DOWN
Nest	Energy metabolism	Immune response	Adaptive immune response	Immune response
Eggs	Energy metabolism	Immune response		Immune response
HEarly	Transport	Stress response	Immune response, Energy metabolism	Immune response
HMid	Biosynthesis & gene expression	cell communication ®ulation of metabolism	Adaptive immune response	Immune response,Energy metabolism, Stress response
HLate		Metabolism	Adaptive immune response	Immune response, metabolism, cellular transport, ion homeostasis

Figure 4.2 – Neuro-transcriptomic changes during paternal care. (A) The number of up- and down-regulated DE genes at each stage of paternal care in the two brain regions. (B) The expression profile of candidate genes related to maternal care across stages, with expression in the two brain regions plotted separately. Note the nest, eggs and early stages were sampled at 9am and were compared to a common circadian control, plotted thrice for visualization. (C) Summary of GO-terms that were enriched in up- and down-regulated genes at each stage in the two brain regions.

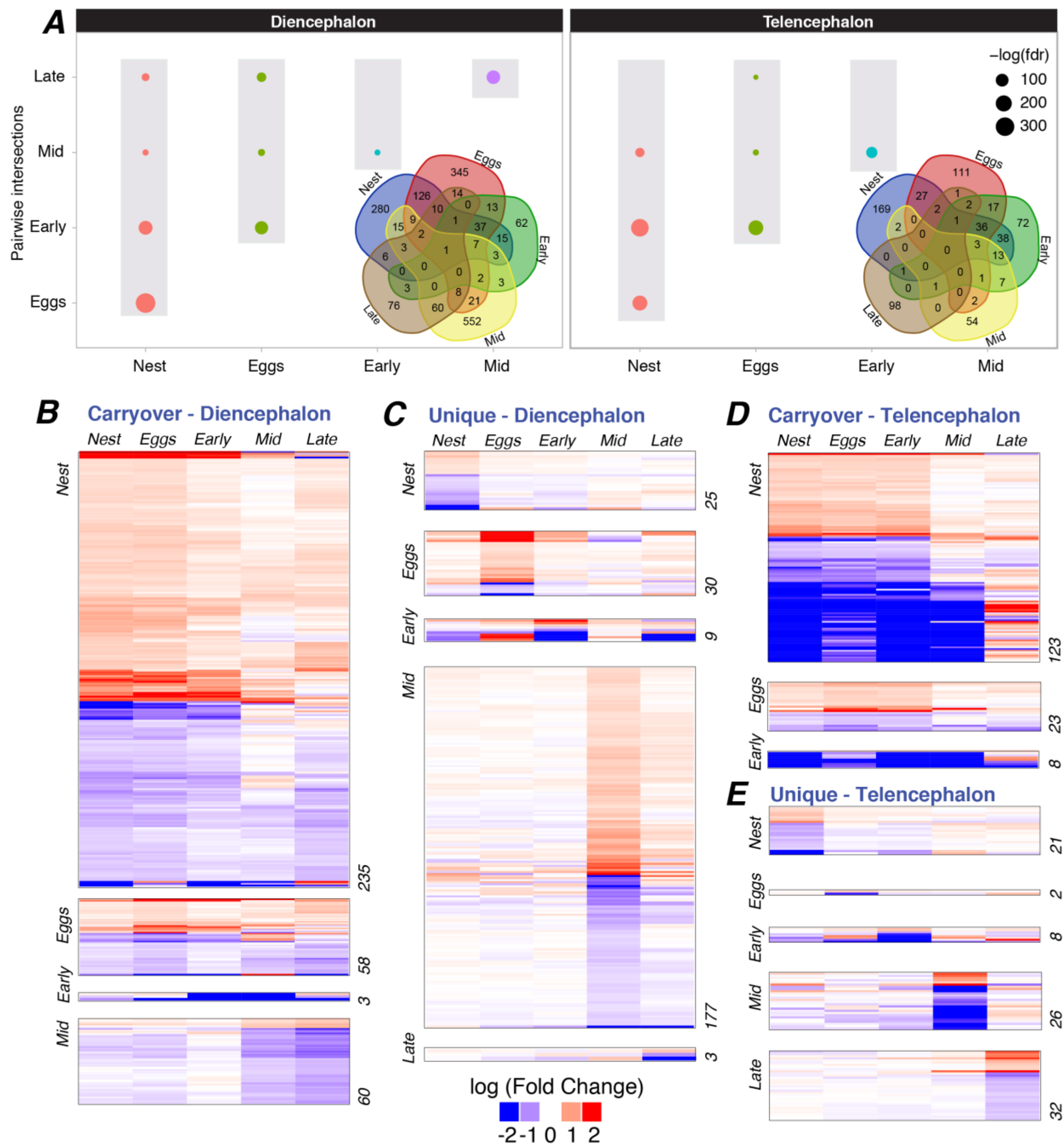


Figure 4.3 – Stability and change across stages. (A) The statistical significance of the pair-wise overlap between stages within each brain region ("carryover genes"). The size of the circle is proportional to the significance of p-value (fdr). A table version of this figure along with p values is available in supplementary table 2. The inset Venn diagram shows the number of

Figure 4.3 (continued)

overlapping and non-overlapping DE genes across stages. (B) The expression pattern of nest, eggs, hatch-early and hatch-middle carryover genes in diencephalon. (C) The expression pattern of unique genes in nest, eggs, hatch-early, hatch-middle and hatch-late stages in diencephalon. (D) The expression pattern of the nest, eggs, and hatch-early carryover genes in telencephalon. (E) The expression pattern of unique genes in the nest, eggs, hatch-early, hatch-middle and hatch-late stages in telencephalon. Numbers show the number of genes plotted in each heat map. Note that the number of genes in the heat maps is less than the number of unique genes in the venn diagrams due to stringent statistical filtering.

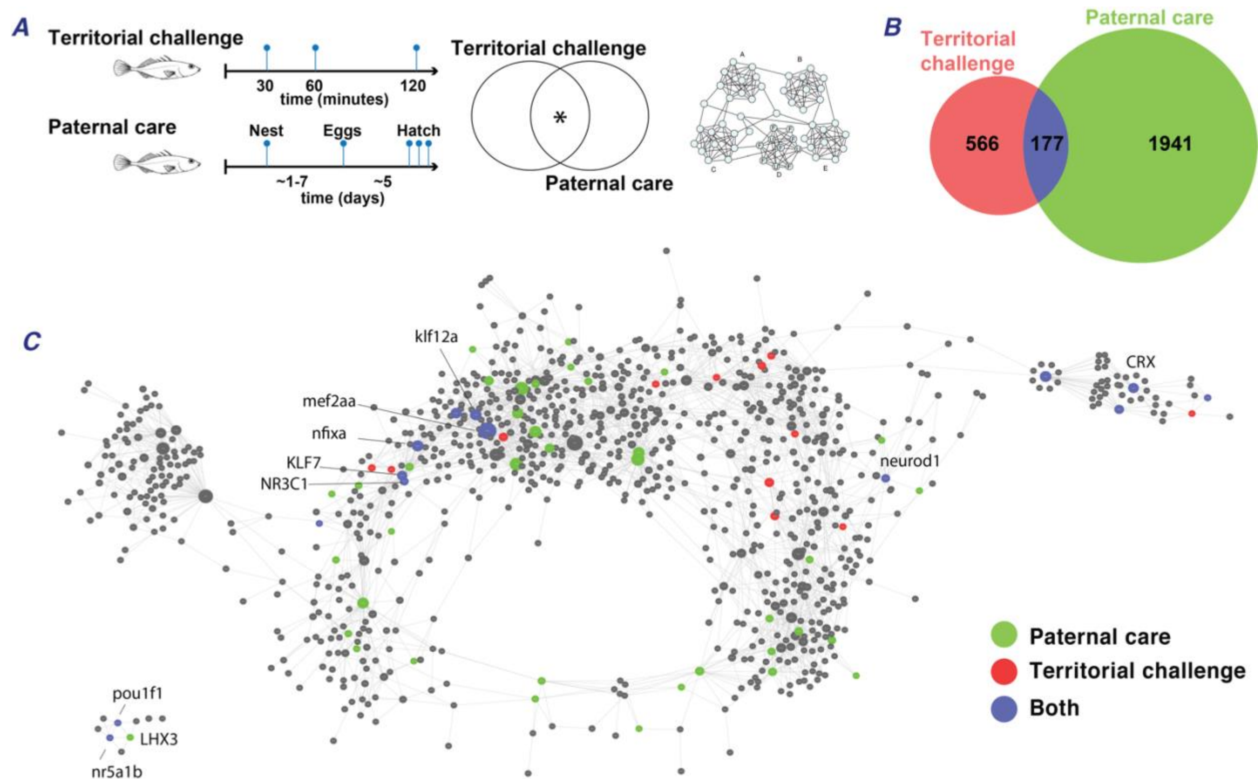


Figure 4.4 – Comparing the transcriptional regulatory dynamics associated with a territorial challenge and paternal care. (A) Experimental time course sampling design in the two experiments. (B) Overlap between territorial aggression and paternal care DEGs. DEGs were pooled across time points and brain regions. (C) ASTRIX-generated transcriptional regulatory network. Each node represents a TF or a predicted TF target gene. Oversized nodes are TFs where the size of the node is proportional to the number of targets. TFs whose targets are significantly enriched in either or both experiments are highlighted with different colors.

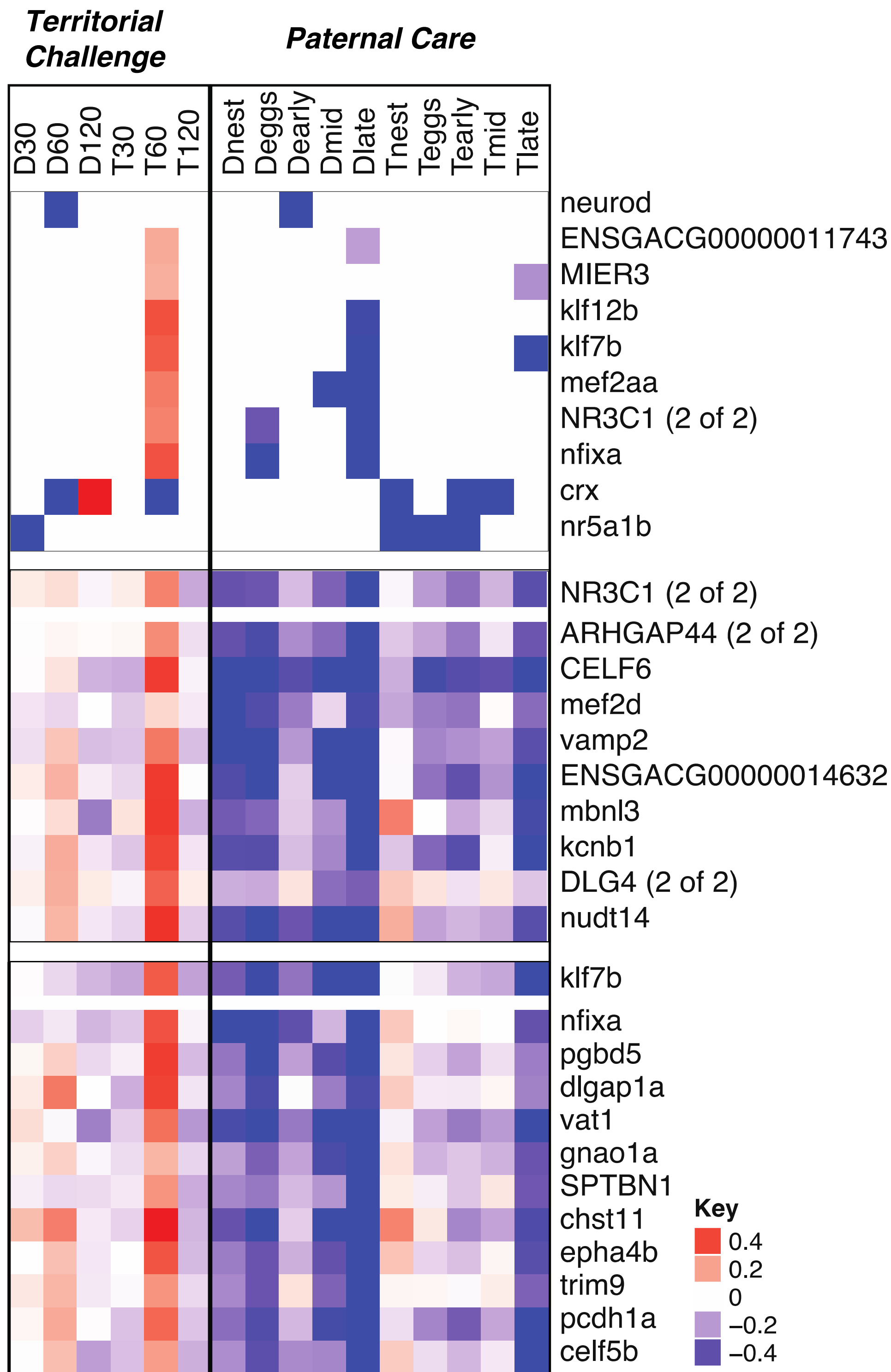


Figure 4.5 – Shared regulators of a territorial challenge and paternal care. The first panel shows the expression pattern of the 10 TFs that were enriched in both experiments (Figure 3). Columns

Figure 4.5 (continued)

are conditions within the two experiments (30, 60 or 120 minutes after a territorial challenge in diencephalon (D) or telencephalon (T); the five stages of paternal care). Note that 8 of the shared TFs were regulated in opposite directions and in different brain regions in the two experiments. The bottom two panels show the expression pattern of two shared, differentially regulated TFs (*NR3C1* and *Klf7b*) and their targets.

LITERATURE CITED

Bukhari, S.A. et al., 2017. Temporal dynamics of neurogenomic plasticity in response to social interactions in male threespined sticklebacks D. Clayton, ed. *PLOS Genetics*, 13(7), p.e1006840. Available at: http://dx.plos.org/10.1371/journal.pgen.1006840 [Accessed October 29, 2017].

Cardoso, S.D., Teles, M.C. & Oliveira, R.F., 2015. Neurogenomic mechanisms of social plasticity. *Journal of Experimental Biology*, 218(1).

Chandrasekaran, S. et al., 2011. Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. *Proceedings of the National Academy of Sciences of the United States of America*, 108(44), pp.18020–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21960440 [Accessed October 11, 2016].

Clutton-Brock, T.H., 1991. *The evolution of parental care*, Princeton University Press. Available at: https://press.princeton.edu/titles/4850.html [Accessed May 21, 2018].

Fischer, E.K. & O'Connell, L.A., 2017. Modification of feeding circuits in the evolution of social behavior. *The Journal of experimental biology*, 220(Pt 1), pp.92–102. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28057832 [Accessed May 20, 2018].

Galbally, M. et al., 2011. The Role of Oxytocin in Mother-Infant Relations: A Systematic Review of Human Studies. *Harvard Review of Psychiatry*, 19(1), pp.1–14.

Goodson, J.L., 2005. The vertebrate social behavior network: evolutionary themes and variations. *Hormones and behavior*, 48(1), pp.11–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15885690 [Accessed May 20, 2018].

Gordon, I. et al., 2010. Oxytocin and the Development of Parenting in Humans. *Biological Psychiatry*, 68(4), pp.377–382. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/20359699 [Accessed May 20, 2018].

Gross, M.R. & Sargent, R.C., 1985. The Evolution of Male and Female Parental Care in Fishes. *American Zoologist*, 25, pp.807–822. Available at: https://www.jstor.org/stable/3883036 [Accessed May 20, 2018].

Hirschenhauser, K. & Oliveira, R.F., 2006. Social modulation of androgens in male vertebrates: meta-analyses of the challenge hypothesis. *Animal Behaviour*, 71(2), pp.265–277. Available at: https://www.sciencedirect.com/science/article/pii/S0003347205003623 [Accessed May 20, 2018].

Kent, M. & Bell, A.M., 2018. Changes in behavior and brain immediate early gene expression in male threespined sticklebacks as they become fathers. *Hormones and Behavior*, 97, pp.102–111. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29117505 [Accessed May 20, 2018].

Ketterson, E.D. et al., 1992. Testosterone and Avian Life Histories: Effects of Experimentally Elevated Testosterone on Behavior and Correlates of Fitness in the Dark-Eyed Junco (Junco hyemalis). *The American Naturalist*, 140(6), pp.980–999. Available at: https://www.journals.uchicago.edu/doi/10.1086/285451 [Accessed May 20, 2018].

Kim, D. et al., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, 14(4), p.R36. Available at: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-4-r36 [Accessed May 25, 2018].

Kleszczyńska, A. et al., 2007. Determination of the neuropeptides arginine vasotocin and isotocin in brains of three-spined sticklebacks (Gasterosteus aculeatus) by off-line solid phase extraction-liquid chromatography–electrospray tandem mass spectrometry. *Journal of*

Chromatography A, 1150(1–2), pp.290–294. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16996527 [Accessed May 25, 2018].

Kleszczyńska, A., Sokołowska, E. & Kulczykowska, E., 2012. Variation in brain arginine vasotocin (AVT) and isotocin (IT) levels with reproductive stage and social status in males of three-spined stickleback (Gasterosteus aculeatus). *General and Comparative Endocrinology*, 175(2), pp.290–296. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22137910 [Accessed May 25, 2018].

Kulczykowska, E. & Kleszczyńska, A., 2014. Brain arginine vasotocin and isotocin in breeding female three-spined sticklebacks (Gasterosteus aculeatus): The presence of male and egg deposition. *General and Comparative Endocrinology*, 204, pp.8–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24852350 [Accessed May 25, 2018].

Langmead, B. & Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), pp.357–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22388286 [Accessed October 11, 2016].

Lema, S.C., Sanders, K.E. & Walti, K.A., 2015. Arginine Vasotocin, Isotocin and Nonapeptide Receptor Gene Expression Link to Social Status and Aggression in Sex-Dependent Patterns. *Journal of Neuroendocrinology*, 27(2), pp.142–157. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25425529 [Accessed May 25, 2018].

Mi, H. et al., 2016. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic acids research*, 44(D1), pp.D336-42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26578592 [Accessed October 11, 2016].

Newman, S.W., 1999. The medial extended amygdala in male reproductive behavior. A node in the mammalian social behavior network. *Annals of the New York Academy of Sciences*, 877,

pp.242–57. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10415653 [Accessed May 20, 2018].

O'Connell, L.A. & Hofmann, H.A., 2011a. The Vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *Journal of Comparative Neurology*, 519(18), pp.3599–3639.

O'Connell, L.A. & Hofmann, H.A., 2011b. The Vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *The Journal of Comparative Neurology*, 519(18), pp.3599–3639. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21800319 [Accessed May 20, 2018].

O'Connell, L.A., Matthews, B.J. & Hofmann, H.A., 2012. Isotocin regulates paternal care in a monogamous cichlid fish. *Hormones and Behavior*, 61(5), pp.725–733. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22498693 [Accessed May 25, 2018].

Palma-Gudiel, H. et al., 2015. Maternal psychosocial stress during pregnancy alters the epigenetic signature of the glucocorticoid receptor gene promoter in their offspring: a meta-analysis. *Epigenetics*, 10(10), pp.893–902. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26327302 [Accessed May 25, 2018].

Powis, Z. et al., 2018. De novo variants in *KLF7* are a potential novel cause of developmental delay/intellectual disability, neuromuscular and psychiatric symptoms. *Clinical Genetics*, 93(5), pp.1030–1038. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29251763 [Accessed May 25, 2018].

Ray, S. et al., 2016. An Examination of Dynamic Gene Expression Changes in the Mouse Brain During Pregnancy and the Postpartum Period. *G3: Genes|Genomes|Genetics*, 6(1), pp.221–233. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26596646 [Accessed

May 20, 2018].

Rosvall, K.A. & Peterson, M.P., 2014. Behavioral effects of social challenges and genomic mechanisms of social priming: What's testosterone got to do with it? *Current zoology*, 60(6), pp.791–803. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27721823 [Accessed May 20, 2018].

Sanogo, Y.O. & Bell, A.M., 2016. Molecular mechanisms and the conflict between courtship and aggression in three-spined sticklebacks. *Molecular Ecology*, 25(17), pp.4368–4376. Available at: http://doi.wiley.com/10.1111/mec.13766 [Accessed November 11, 2016].

Saul, M.C. et al., 2017. Transcriptional regulatory dynamics drive coordinated metabolic and neural response to social challenge in mice. *Genome research*, 27(6), pp.959–972. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28356321 [Accessed May 20, 2018].

Schneider, J.S. et al., 2003. Progesterone receptors mediate male aggression toward infants. *Proceedings of the National Academy of Sciences*, 100(5), pp.2951–2956. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12601162 [Accessed May 20, 2018].

Shpigler, H.Y. et al., 2017. Behavioral, transcriptomic and epigenetic responses to social challenge in honey bees. *Genes, Brain and Behavior*, 16(6), pp.579–591. Available at: http://doi.wiley.com/10.1111/gbb.12379 [Accessed May 20, 2018].

Smith, C. & Wootton, R.J., 1999. Parental energy expenditure of the male three-spined stickleback. *Journal of Fish Biology*, 54(5), pp.1132–1136. Available at: http://doi.wiley.com/10.1111/j.1095-8649.1999.tb00866.x [Accessed April 1, 2016].

Storey, A.E. et al., 2000. Hormonal correlates of paternal responsiveness in new and expectant fathers. *Evolution and Human Behavior*, 21(2), pp.79–95. Available at: https://www.sciencedirect.com/science/article/pii/S1090513899000422 [Accessed May 20,

2018].

Storey, J.D. & Tibshirani, R., 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, 100(16), pp.9440–9445. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12883005 [Accessed May 20, 2018].

Whittington, C.M. & Wilson, A.B., 2013. The role of prolactin in fish reproduction. *General and Comparative Endocrinology*, 191, pp.123–136. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23791758 [Accessed May 20, 2018].

Wingfield, J.C. et al., 1990. The Challenge Hypothesis: Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. *The American Naturalist*, 136(6), pp.829–846. Available at: http://www.journals.uchicago.edu/doi/10.1086/285134 [Accessed October 11, 2016].

Wootton, R.J., 1984. *A Functional Biology of Sticklebacks*, Boston, MA: Springer US. Available at: http://link.springer.com/10.1007/978-1-4615-8513-8 [Accessed March 27, 2016].

Wu, Z. et al., 2014. Galanin neurons in the medial preoptic area govern parental behaviour. *Nature*, 509(7500), pp.325–30.

APPENDIX A: SUPPLEMENTARY INFORMATION FOR CHAPTER 2

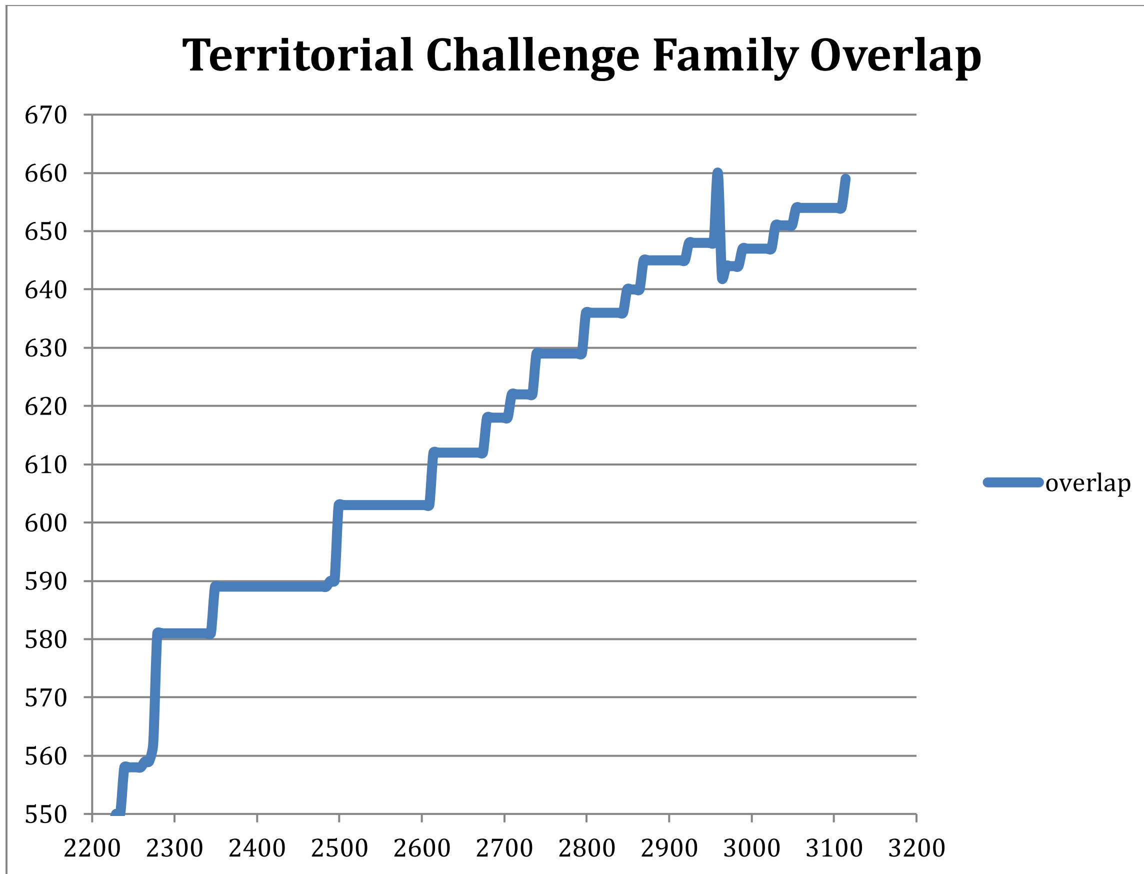


Figure A.1 - The observed overlap in gene families between species in the territorial challenge context was highly significant, therefore, we wanted to see when such an overlap is detected if we keep on increasing the size of our random set (n_2). It was found that at least such an overlap is expected when the n_2 size is of at least 2990, keeping n_1 as real observed set.

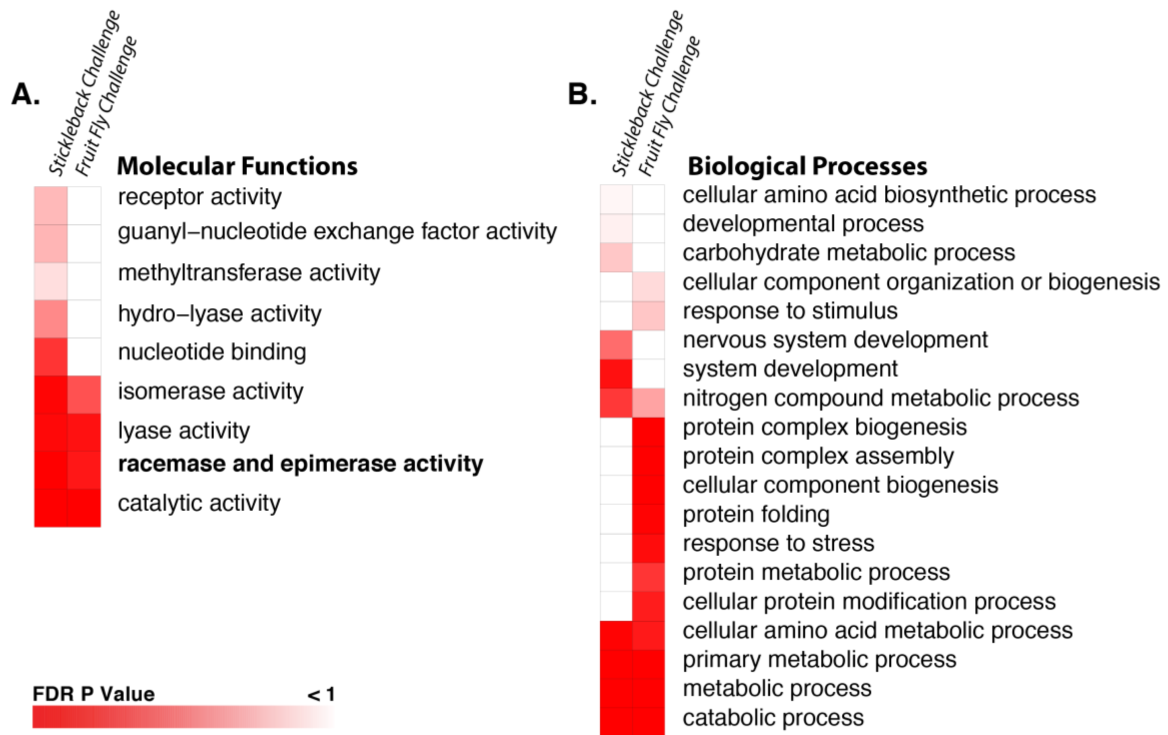


Figure A.2 - Functional enrichment of stickleback and fruit fly common gene subfamilies. (A) Molecular Function (B) Biological Processes.

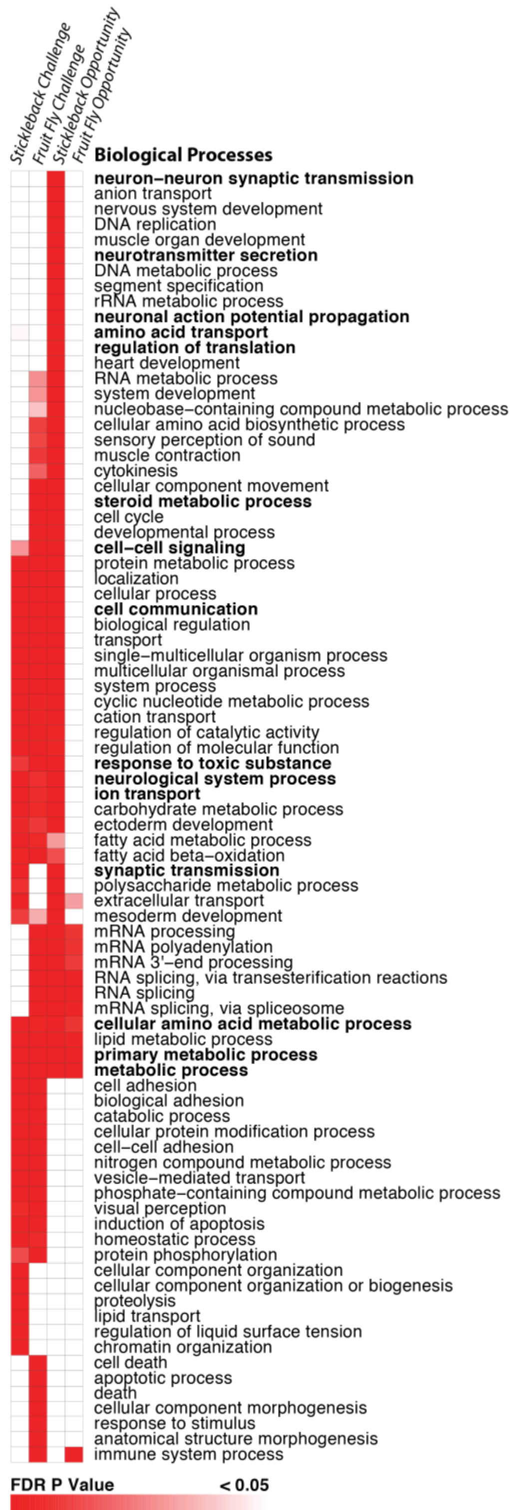


Figure A.3 – Functional enrichment of GO biological processes.

APPENDIX B: SUPPLEMENTARY INFORMATION FOR CHAPTER 3

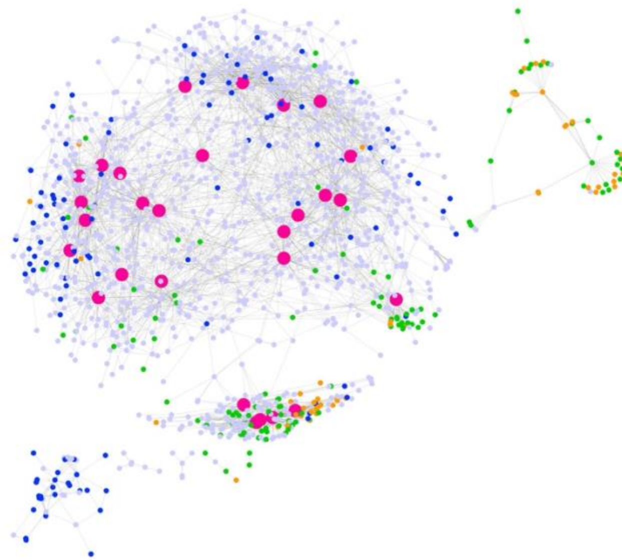


Figure B.1 – TRN highlighting top hubs in the network and region-specific DEGs. The TRN contains 352 TFs, which regulate 1155 genes through a total of 3683 interactions. The top 20 TFs ("hubs") with the highest number of targets (over 30 each) are highlighted in pink. Target genes that are differentially expressed in Diencephalon or Telencephalon (CFDR < 0.1) are shown as blue or green nodes respectively. Genes in the TRN that were differentially expressed in both the regions are shown in orange.

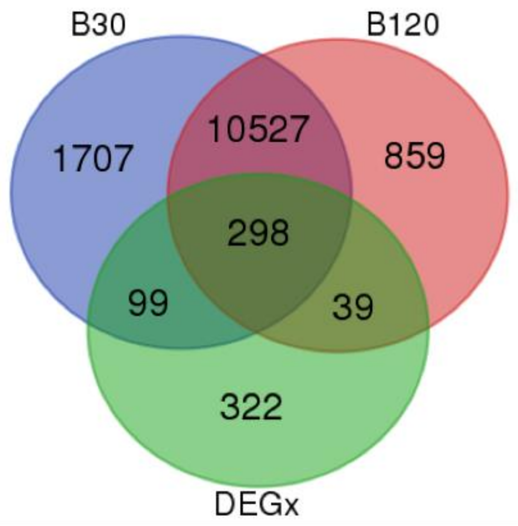


Figure B.2 – Baseline accessibility of diencephalon DEGx at 30 and 120 minutes.

Table B.1 – Enrichment of TFs in $DAPDEG_x$.

TFs	Description	Significance
pparg	peroxisome proliferator activated receptor gamma	P < 0.005
neurod	neurogenic differentiation	P < 0.0001
ikzf1	IKAROS family zinc finger 1 (Ikaros)	P < 0.0001
Irf4b	interferon regulatory factor 4b	P < 0.005
spi1a	spleen focus forming virus (SFFV) proviral integration oncogene spi1a	P < 0.0001
ETV7	ets variant 7	P < 0.001
pbx4	pre-B-cell leukemia transcription factor 4	P < 0.05
mafbb	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B, duplicate b	P < 0.01
cebpb	CCAAT/enhancer binding protein (C/EBP), beta	P < 0.0001
homez	homeodomain leucine zipper gene	P < 0.05
spi1b	spleen focus forming virus (SFFV) proviral integration oncogene spi1b	P < 0.0001
irf8	interferon regulatory factor 8	P < 0.0001
NFATC3	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3	P < 0.05
TLX2	T-cell leukemia homeobox 2	P < 0.05
tfec	transcription factor EC	P < 0.0001
LYL1	lymphoblastic leukemia associated hematopoiesis regulator 1	P< 0.05

APPENDIX C: DIFFERENTIALLY EXPRESSED GENES FOR CHAPTER 4

The supplementary file (APPENDIX C.xlsx) contains sixteen worksheets. First ten sheets are for paternal care DEGs at eFDR < 0.01, whereas last six are for territorial aggression DEGs at eFDR < 0.01

APPENDIX D: SUPPLEMENTARY INFORMATION FOR CHAPTER 4

The supplementary file (APPENDIX D.xlsx) contains two separate sheets D.1 and D.2.

Sheet D.1. Pairwise gene sets intersections along with their p values.

Sheet D.2. Functional enrichment of unique genes to paternal care and territorial aggression.