Identifying Novel Transcriptional Effectors of the Juvenile Hormone Pathway in *Aedes aegypti*

Megan L. Richardson

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science in Life Sciences
In
Biochemistry

Jinsong Zhu, Committee Chair
Chloe Lahondére
Igor Sharakhov
Jake Tu

April 21, 2020
Blacksburg, VA

Keywords: (Aedes aegypti, juvenile hormone, fat body, transcription factor, RNA sequencing)

Identifying Novel Transcriptional Effectors of the Juvenile Hormone Pathway in *Aedes aegypti*

Megan L. Richardson

GENERAL ABSTRACT

The mosquito, *Aedes aegypti*, is responsible for the spread of a myriad of viruses such as dengue, zika, and chikungunya. Currently, these infections have no vaccine or treatment available and transmission rates continue to steeply rise in response to the spread of breeding grounds. Popular insecticides carry detriments such as off-species toxicity and continuous application to treatment areas. Our lab proposes an alternative to these chemical insecticides by manipulating a developmental pathway in the mosquito. The Juvenile Hormone pathway is conserved in arthropods, responsible for growth and reproduction, and the hormone is nontoxic to mammals. Through the combination of bioinformatics and genomics studies, we have identified two JH-responsive gene candidates that are potential regulators of this pathway.

Identifying Novel Transcriptional Effectors of the Juvenile Hormone Pathway in *Aedes aegypti*

Megan L. Richardson

ABSTRACT

Aedes aegypti is the primary vector for dengue, zika, chikungunya, and yellow fever viruses. Disease transmission through this mosquito places over 40% of the world's population at risk of contracting one or more of these pathogens. Current control strategies such as insecticide application have failed or carry additional burdens, such as off-target toxicity to mammals and birds. Our lab proposes utilizing a conserved arthropod hormone pathway, juvenile hormone (JH), related to growth and reproduction to curb these vector populations and reduce disease transmission. Additionally, JH is nontoxic to birds and mammals; it requires incredibly high doses to have lethal effects. We hypothesize that JH-responsive genes expressed early in the adult are responsible for her reproductive capacity and by manipulating the signaling downstream of the receptor, we will be able to decrease the female's fecundity and limit vector populations.

Via bioinformatics screening of RNA-sequencing data using the New Tuxedo pipeline, we identified 47 potential transcription factor candidates. With the use of *in vitro* culturing of the mosquito's reproductive tissues in the presence of a translation inhibitor, we identified two early JH responsive gene candidates, FoxA and zinc finger 519, p-value <0.05. The functional characterization of these two remains to be seen, however, in *Drosophila melanogaster*, they both have roles in chromatin remodeling and require protein partners to carry out long range interactions.

Acknowledgements

I would like to thank the Virginia Tech Department of Biochemistry, specifically Dr. Jinsong Zhu for the opportunity to work with his team. These last three years have been eye-opening to my understanding of science as a whole. His and Dr. Pengcheng Liu's guidance through my time here holds deep value and appreciation. Without their patient discussions, this work would not be the same and I would not be a better scientist. I would also like to thank my lab mate Tahmina for all of her support and friendship while we have gone down this road together.

Finally, I would like to thank my family for the support needed to pull through this degree program. By family, I mean my cohort, my support system here in Blacksburg, and my immediate family members. They are the reason I am able to do good work and the reason I continue to be passionate about this field of research.

Table of Contents

List of Figures

List of Tables

Chapter One: Introduction

I. Global Impact of the Vector, *Aedes aegypti*

Aedes aegypti has been studied for the last 50 years due to its role as the main vector of dengue, yellow fever, Zika, and chikungunya viruses[3]. *Aedes* females consume blood meals to make yolk proteins necessary for egg production and, ultimately, propagation of the species. These feeding behaviors passively facilitate disease transmission through the exchange of the vector and hosts' bodily fluids, i.e. saliva for blood, respectively (Figure 1). For example, dengue virus (serotypes I-V) is passed to mosquitoes via infected blood

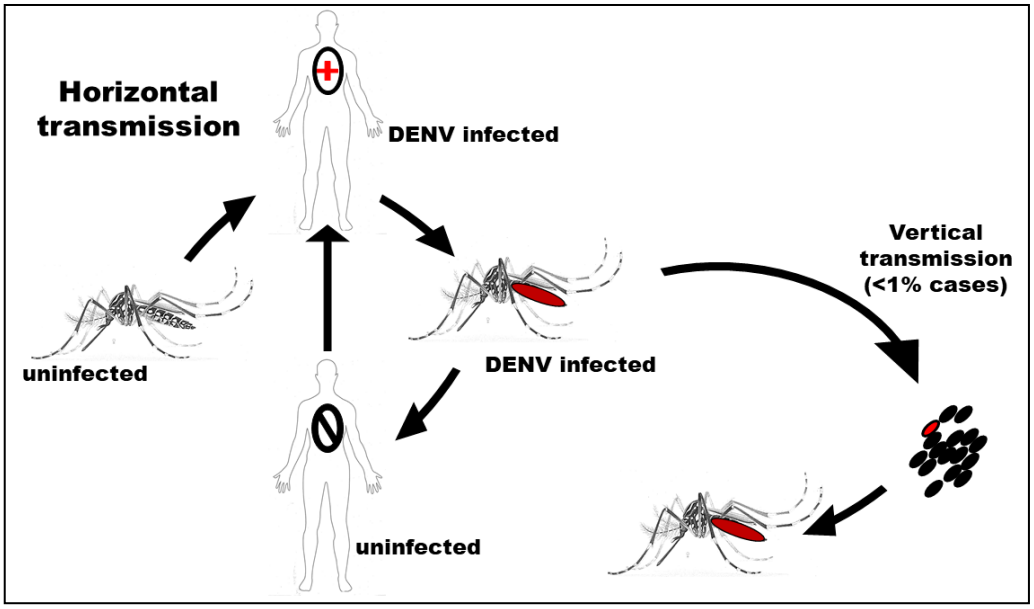


Figure 1: Model of dengue virus transmission.

meals (horizontal transmission) and establishes itself in the midgut of the mosquito. More rarely, although significantly, vertical transmission through the mosquito germline is also possible for the spread of dengue and is an additional issue in control programs success [4-6]. Knockdown studies of the midgut proteasome catalytic subunits demonstrated that dengue virus spreads to the various organs inside the mosquito via the ubiquitin

proteasome pathway (UPP) and finally establishes itself in the salivary glands, where all subsequent feedings now carry the possibility of spreading dengue infection to humans [7]. This was verified with intra-thoracic injection of DENV-II (separating initial viral establishment from viral spread) and subsequent RNA sequencing showing upregulation of key components of the UPP [7]. A single blood meal can generate ~120 eggs per female and a wild mosquito can sustain approximately three to four blood meals in her lifetime [8]. Thus, the probability of contact with virally infected hosts and subsequent disease spread raises concern for worldwide transmission. *Aedes* mosquitoes have utilized sea trade routes for hundreds of years as a means of spreading to regions now encompassing most of the Americas, Southern Africa, and Southeast Asia [9]. Dengue transmission alone has risen dramatically with over 100 countries now experiencing severe dengue outbreaks [10]. Current models suggest that the incidence and risk of infection have increased more than 30-fold in recent decades [11]. There are 390 million new dengue infections per year, of which 96 million develop into severe dengue, involving major hemorrhagic fever, shock, and death [11].

Viewing this global issue as a tripartite association between virus, vector, and infected human populations provides alternative solutions to this ever-growing problem [3]. Current efforts to curb transmission include insecticide-treated bed netting, inoculative releases of sterile males (SIT), and treatment of aquatic breeding grounds [4]. Despite various control efforts, *Ae. aegypti* has demonstrated an affinity to urban landscapes and an ability to thrive even in countries with strict control programs [12]. Insecticide resistance is quickly on the rise and many ecologically important, off-target species are being affected by

current chemical methods [10, 13, 14, 15]. Sterile males help mitigate off-species detriments, but require lasting cooperativity of local governments where the releases take place. Additionally, the resources required for the continued rearing and release of sterile males are not feasible when looking for a lasting solution to the problem. Instead, we should be targeting the female mosquito's genetics for our efforts into population control. Juvenile hormones are unique to arthropods; at least one juvenile hormone has been identified in over 100 insect species, crossing ten insect orders[16]. The main structure for JH is a 12-carbon chain containing a methyl ester at one end and an epoxide at the opposite (Figure 2). There is a keto-group on the first carbon and R groups at carbon positions 3, 7, and 11. In different insect species, the R groups differ in methyl or ethyl chains on the carbon positions as well as potentially containing additional epoxide rings. The family of juvenile hormones contains seven known members thus far: JH 0-III, 4-methyl JH I, JH III bisepoxide, and JH III skipped bisepoxide. Of these, JH-III is the most commonly found in dipteran species. This illustrates that our approach to manipulating the JH pathway would not be unique to the dengue vector, but capable of managing many medical and agricultural pests with no harmful effects to mammals. JH analogs such as methoprene, pyriproxyfen, and fenoxycarb act as agonists of JH and mimic its antimetamorphic roles during development and affect reproductive stages [17]. These JH mimics are currently used in mosquito control. To understand the current hormonal mimics and provide a higher level of specificity and control for insecticide development, the field must have a deeper understanding of this pathway attributed to metamorphosis and reproduction.

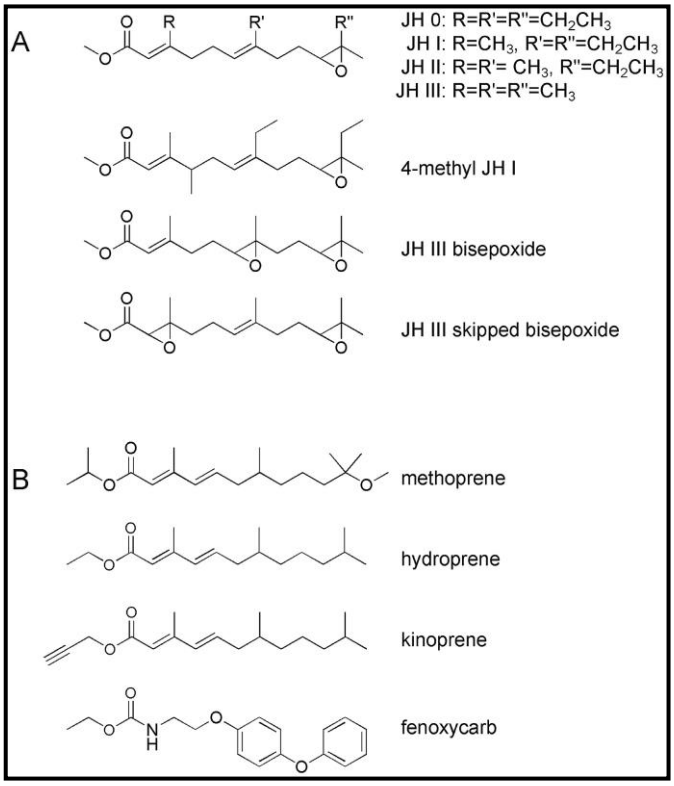


Figure 2: (A). JH III is the principal form of JH that is found in dipteran insects. JHA insecticides (B) are structural and/or biological mimics of JH. Reproduced from © 2011 Kamita et al. Juvenile Hormone (JH) Esterase of the Mosquito Culex quinquefasciatus Is Not a Target of the JH Analog Insecticide Methoprene. PLoS ONE 6(12): e28392. https://doi.org/10.1371/journal.pone.0028392

II. Juvenile Hormone Synthesis and Signaling

A. Overview

Juvenile Hormone (henceforth referred to as JH) is a sesquiterpenoid hormone synthesized in the corpora allata (CA) of the majority of insects and has been studied for over half a century [18, 19]. JH works in concert with 20-hydroxyecdysone (20-E) for correct developmental timing related to growth and reproduction. JH functions as a metamorphic repressor, preventing ascendance to the next stage of the invertebrate's life cycle until a status quo has been reached. *Ae. aegypti* pupae eclosing into new adults have low to no JH present [20]. These titers quickly rise and peak in the 48 hours following

emergence, from there they begin to decline due to the female's sexual readiness. 20-E remains at basal levels after adult emergence and rises following a blood meal to activate vitellogenesis [19, 21, 22, 23]. 20-E titers reach a peak at 18 hours post-blood meal and then sharply drop at 30 hours; at this point JH levels rise again to arrest the second round of

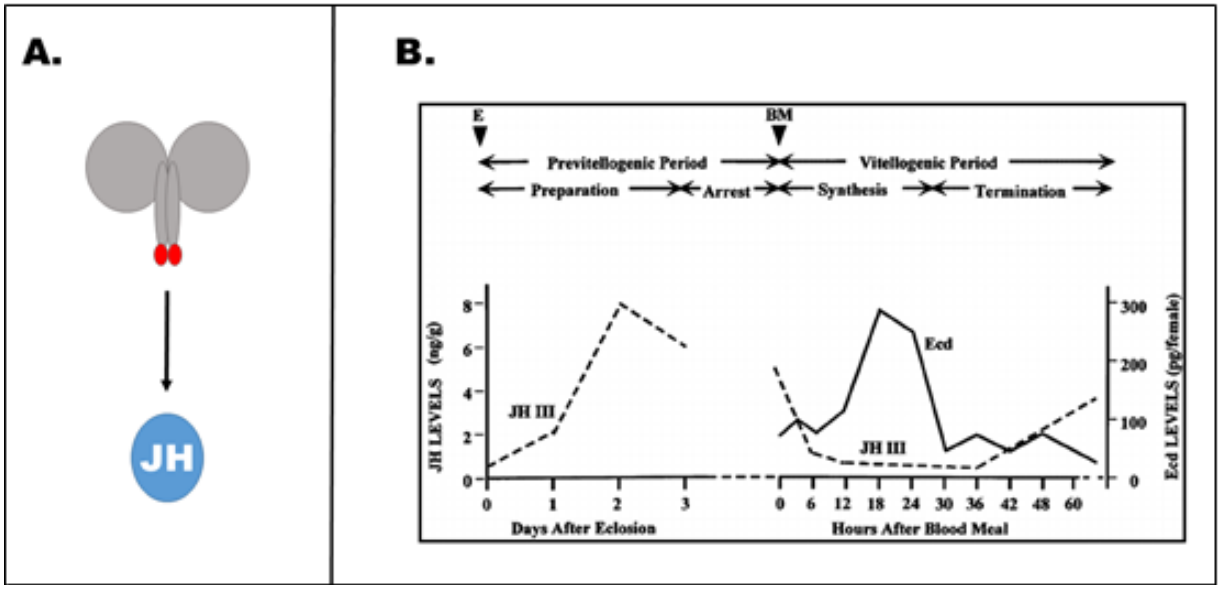


Figure 3: (A) JH is synthesized by the corpora allata (red). (B) Hormonal titers during the first vitellogenic cycle of the anautogenous mosquito. A. aegypti. BM, blood meal; E, eclosion; JH III, juvenile hormone titers (modified from Shapiro et al, 1986); Ecd, ecdysteroid titers (Modified from Hagedorn et al. 1975). Adapted with permission from EMBO-J Creative Commons License 4.0. Zhu, J., Miura, K., Chen, L., & Raikhel, A. S. (2000). AHR38, a homolog of NGFI-B, inhibits formation of the functional ecdysteroid receptor in the mosquito Aedes aegypti. *The EMBO journal*, *19*(2), 253–262. https://doi.org/10.1093/emboj/19.2.253

eggs from being laid until a nutritional status quo has been reached. These alternating waves ebb and flow for all subsequent gonotrophic cycles during the adult's life (Figure 3).

Allatectomy, removal of the CA structure, results in loss of JH synthesis as viewed by an *in vitro* radiochemical assay [24]. Interestingly, both surgical implantation of CA from a

different mosquito and topical application of JH permitted normal development to proceed [25, 26]. In other species, JH's dependence on the CA for synthesis was confirmed. Decapitations of immature silkworm moths, *Bombyx mori*, resulted in the immediate spinning of a cocoon and a small adult emerging [27]. Conversely, if the CA of a young silkworm is implanted into the brain of a fully mature larvae this results in no metamorphic transitions [27]. Instead, it undergoes an additional molt and produces an extra-large caterpillar [27]. With these studies, it has become common knowledge in the field that JH is the master regulator of metamorphoses; later studies demonstrated it is also crucial to reproductive development in adults.

B. 20-E signaling in *Aedes aegypti*

The molecular mechanism related to 20-E signaling has been well described and fits a classical model for steroid hormone action. 20-E exerts its effects through a protein heterodimer composed of the ecdysone receptor (EcR) and Ultraspiracle [28, 29]. The receptor complex binds to an ecdysone response element within the promoters of target genes, altering their expression and thereby leading to changes in the genetic programs associated with molting and metamorphosis [30]. Within these tissues exists a signaling hierarchy with 20-E at the precipice. In *Drosophila*, 20-E bound to the receptor directly activates a small group of early genes: *Broad-Complex* (*BR-C*), *E74* and *E75*, each of which encodes several isoforms of DNA-binding transcriptional regulators [31, 32]. The protein products of early genes activate a much bigger group of late genes that directly or indirectly perform distinct metamorphic processes such as cell death of obsolete larval

tissues, cell proliferation/ differentiation of imaginal tissues, and pupal cuticle formation [33].

In adult reproduction, blood meal-stimulated neural tissues trigger the neuropeptide Ovarian Ecdysteroidogenic Hormone (OEH) to be secreted into the hemolymph and OEH stimulates the ovaries to begin production of ecdysone [34]. OEH is a prominent regulator of ecdysone synthesis, but is by no means the only factor involved. 20-E synthesis occurs from the conversion of ecdysone to its bioactive form, 20-hydroxyecdysone, within the fat body of the females and activates vitellogenesis by upregulating yolk precursor proteins (YPPs) such as Vitellogenin (Vg), Cathepsin b, and Vitellogenic carboxypeptidase [35]. This, in conjunction with the nutritional deposits from the blood meal, results in the mass production of yolk proteins for developing eggs [36]. Yolk proteins provide the essential nutrients required for embryonic development and are transferred from the fat body to the developing oocytes as early as 4 hours PBM [37]. The primary oocytes repress the secondary oocytes until egg laying is complete and the first gonadotrophic cycle is finished [37].

Interestingly, 20-E is also found in the male accessory glands of the malaria mosquito, *Anopheles gambiae,* and large amounts are transferred to females during mating [37]. This suggests that 20-E plays an additional role in modulating post-mating effects in females [37]. However, a role for male derived 20-E in YPP gene regulation has not yet been demonstrated in this mosquito [38].

C. JH signaling during the *Aedes* life cycle

During *Aedes* embryonic development, JH patterning is transiently appearing and subsequently rises in the remaining life cycle stages [20]. It maintains high levels of expression to repress ecdysone titers and prevent early instar shifts and pupal commitment [20]. During the mosquito's larval stage, the fat body, equivalent to the human liver, accumulates nutrients for use in the adult stage. The nutritional state of the larval fat body is crucial to the longevity and reproductive success of adult females [39]. Following emergence, JH activates many transcription factors responsible for previtellogenic (PV) oocyte growth, proliferation of ribosomes, and ploidy of fat bodies [40]. Fat body tissue is the primary organ for metabolism in insects, functioning as a storage for carbohydrates, lipids, and proteins. In response to a blood meal, the fat body is capable of synthesizing large amounts of proteins, ribosomes, and oil droplets for maturing eggs, also making it the primary organ of reproduction. JH levels rise dramatically in adult mosquitoes and peak during the first 2-3 days post-emergence (PE), remaining relatively high until a blood meal is consumed [26, 41]. About 25% of topically applied, radio-labeled juvenile hormone appears in the hemolymph as the acid diol byproduct, and 50% of this is excreted in the urine immediately following the blood meal [26]. Topical application of an inhibitor for the JH precursor enzyme, JH esterase, resulted in an absence of JH acid diol, confirmed with gas chromatography coupled with mass spectroscopy (GC/MS), as well as poor follicle development and a failure of the ovaries to reach the resting stage [26]. These results confirm that JH is responsible for previtellogenic growth of the oocytes and that its decline may be necessary to allow the release of 20-E in response to a blood meal.

It is assumed that the endocrine and nervous systems monitor teneral reserves and regulate physiological, developmental, and behavioral processes in response [42, 43]. In one study, larval diets were manipulated to be nutritionally rich or poor, and *in vitro* production of JH was measured via the incorporation of ^{3}H-labeled methionine. The data showed that adults that were fed a nutrient-rich larval diet had a 3-fold increase in JH biosynthesis compared to the low-nutrient females [39]. In addition, there was also a significant increase in the dry mass of the nutrient-rich female as well as their lipid production [39]. The primary follicles that make up the ovaries (those first undergoing vitellogenesis) are undifferentiated during emergence [44]. Over the next 60 hours, the PV follicles will double in size (100 μm long) and enter a state-of-arrest prior to a blood meal [45]. During vitellogenesis, yolk proteins and lipids are secreted by the fat body, transported to the ovaries, and deposited into the developing eggs [46]. Body size (assessed by wing length) and fecundity were measured as well; the number of mature primary follicles in 72h PBM females determined potential fecundity. *Ae. aegypti* females emerging from a high-nutrient larval environment had both larger body size (~30%) and matured a greater number of eggs, 118 vs. 64, $P<0.001$ [39]. Development of the follicles to the resting stage was determined to be independent of the larval diet, although low-reserve females required 3% sucrose to develop to the resting stage [39]. Additionally, topical application of JH restored follicle development in low-nutrient females and allowed the resting stage to be reached without additional nutrition from sugar feeding [47]. When ecdysone was added with JH, the follicles failed to grow [47].

From these data, a model was developed to explain how a female's nutritional condition and hormones interact to regulate oocyte development. In this model, high levels of glycogen and protein surpass a threshold set in the nervous system that activates ovarian ecdysteroid production and inhibits JH biosynthesis in the CA, allowing vitellogenesis and egg maturation. When glycogen and protein levels are below this threshold, the CA secretes high JH levels, ovarian ecdysteroid levels are low, and egg maturation is delayed or arrested [39, 48]. Lipid assays showed that following blood-feeding, between 50-80% of the fat body's lipid storages are mobilized for oocyte development [49]. The fate of the follicles post-blood meal (PBM) is dependent on PV nutrition and JH [42, 49]. PV females treated topically with JH during the resting stage were offered a blood meal and subsequently had their ovaries dissected [46]. RNA sequencing of PV and blood-fed females coupled with GC/MS showed that poor diet females are lacking lipids, glycogen, and protein in their fat body reserves [48-50]. Low-nutrient PV females have a larger proportion of follicle resorption following a blood meal and require two or more blood meals to have ovaries developed to the resting stage [46, 49, 51].

Additionally, when provided a blood meal, larger females (high larval nutrition) consumed more than twice that of the small females and increased fecundity roughly 4-fold as determined by fecal haematin analysis [42]. An average blood meal triggers apoptosis and resorption of approximately 27% of the follicles for uniform egg maturation or to divert nutritional reserves to multiple processes [46]. The efficiency of the female to progress through vitellogenesis was also measured and the nutrient-rich females were able to complete their first gonotrophic cycle in a significantly shorter time

with a larger proportion of mature eggs (18 vs. 116 eggs) [42]. However, researchers observed that the low-nutrition, smaller females had an increased life span by 40%, having one speculate that the increased longevity is an offset of delayed development directly correlated to the nutritional status of the female, both larval and adult [52]. These observational data are crucial when examining disease transmission mechanics and the epidemiology of viruses such as dengue.

III. *Methoprene-tolerant* and gene regulation

In the fat body tissue, JH becomes bound to an accessory binding protein (JHBP) for receptor-mediated endocytosis and transport to the cytoplasm of the cell [23]. JH then binds to its receptor *Methoprene-tolerant* (*Met*) for nuclear translocation. Upon binding of JH, *Met* recruits the coactivator *Taiman* (*Tai*, *SRC,* or *FISC*) which facilitates DNA binding of *Met* to the JH response elements of the genes that are under the control of JH. There are four known splice variants of *Tai* (A-D), each of which are involved in hormone-specific and time-specific expression. Our lab demonstrated that JH controls the production of the different *Tai* isoforms [53]. In the fat body of newly emerged mosquitoes, *Tai* C and D were the predominant isoforms while the transcripts of isoforms A and B were not detectable. When the fat body was cultured *in vitro* with JH-III, the mRNA levels of A/B increased considerably while the abundance of the C/D mRNA decreased accordingly [53]. The A/B isoforms coincide with the gained competence of a 20-E response in the PV fat body [53]. The isoform composition remained unchanged from 48 h PE to 30 h post blood meal (PBM). A decline of *Tai* A and B transcripts started at 30 h PBM, accompanied with a rise of the levels of *Tai* C and D. This timing coincides with

the drop of 20-E titers back to a basal level as the yolk protein synthesis ends and the first gonotrophic cycle comes to a close [22].

A screen for *Drosophila* mutants resistant to the insecticide methoprene identified Methoprene-tolerant as the receptor of JH and kinetic studies determined a high binding affinity for JH (K_d = 4.4± 1.9 nM) [26, 54, 55]. RNAi experiments of *Met* and *Tai* showed a necessity for the heterodimer to exert its influence; either alone could not facilitate the binding to gene regulatory elements [56-58]. *Met* is a transcription factor of the basic helix-loop-helix (bHLH)/Per-Arnt-Sim (PAS) family. This family of proteins is involved in vast cellular processes across all kingdoms, with over 125 different proteins identified in humans and 145 identified in *Arabidopsis*, including sex determination and nervous system development [59]. Many bHLH proteins are comprised of two alpha-helices separated by a short loop. At the amino-terminal end of the bHLH domain is the first helix, the part that binds to DNA at a consensus hexanucleotide sequence (E-box) [59]. The HLH domain mediates interactions with other proteins. Further beyond the HLH domain is the PAS domain, a region approximately 260-310 amino acids long; the domain is made up of two repeats of about 50 amino acids (PAS-A and PAS-B) separated by poorly conserved residues [59]. The PAS-B domain facilitates a dimerization interface of polar and hydrophobic residues allowing a myriad of proteins to interact and facilitate the sweeping gene signaling we know of this protein superfamily [60, 61]. Our lab previously performed a yeast two-hybrid screening using the bHLH-PAS domain of *Met* as bait to isolate partners that binds to *Met* in response to JH [56]. We were able to isolate a single clone encoding the bHLH-PAS domain of *Tai* [56]. In *Tribolium castaneum*, a potential JH-

binding pocket in the PAS-B domain was verified by site-directed mutagenesis of select residues to alter its binding efficacy [62]. Predicted key binding pocket residues were altered to amino acids with bulkier side chains, causing a bulge and steric hindrance in the pocket [62]. This did not prevent *Met-Met* dimer formation, but did halt the dissociation of the complex and the production of the *Met-Tai* heterodimer. RNA interference (RNAi) mediated depletion of *Met* inhibits the growth of follicles in newly emerged adults[63]. In *Met*-depleted female adults, we see severe retardation in growth of primary follicles, a phenotype similar to that of JH-depleted and low-nutrient PV females [63].

Elucidating JH signaling has proven to be more challenging than its ecdysone counterpart; some elements fit the hierarchical classical model, but there exists a confusing crosstalk of genes involved directly downstream of *Met*. Evidence shows that these genes can activate a subset of downstream genes alone and a different subset when working in combination with one another [64, 65]. These factors do not inhibit one another, leading to the speculation of a cooperativity between these rather than a hierarchy. Additionally, our lab demonstrated that one of the downstream effectors, *Kr-h1*, displays both an activating and a repressive role, depending on the time of the signaling event [66]. However, we still do not have a complete understanding of what accounts for all signaling events downstream of the receptor; we are currently only able to attribute approximately 20% of *Met*-repressed genes to *Kr-h1* and *Hairy* [64]. To date, no transcription factors that function solely as activators on this pathway have been identified, instead we see dualistic and synergistic roles to activate different subsets of genes.

IV. **Early JH-Responsive Genes**

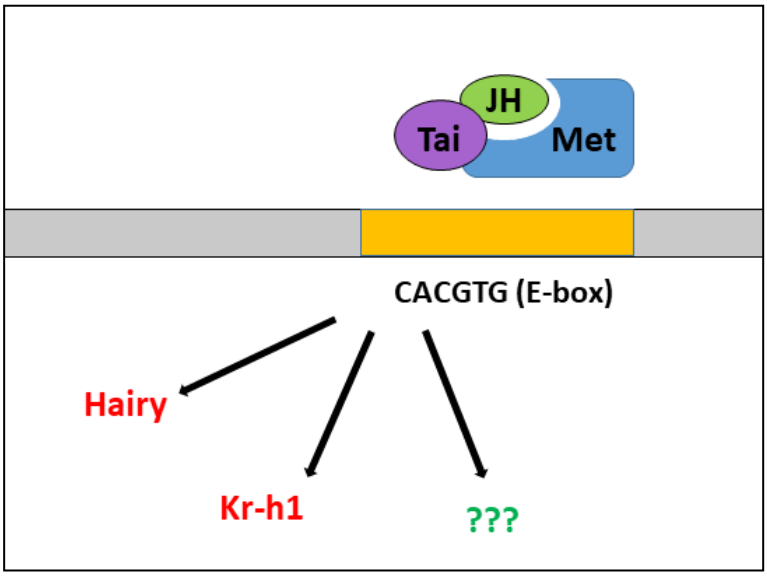


Figure 4: Following heterodimer formation, the *Met-Tai* complex binds to E-box motifs within a gene's promoter region.

The *Met-Tai* complex directly binds to a JH response element (JHRE), located in the promoters of the target genes of JH (Figure 4). Bioinformatics analysis and electrophoretic mobility shift assays (EMSA) revealed a well-conserved 6-mer E-box sequence (CACGTG/CACGCG), which was present in the promoters of a number of *Met* up-regulated genes [58]. The binding action can rapidly activate genes such as *Hairy* and *Krüppel homolog 1* (*Kr-h1)* [57, 67], which encode transcriptional regulators.

After microarray analysis of the transcriptomes in the fat body of previtellogenic female mosquitoes, 6,146 JH-regulated genes were split into early post-emergence (EPE), mid post-emergence (MPE), or late post-emergence (LPE) subsets, based on their peak expression at 6, 24, and 66 h PE, respectively [58]. The EPE (1,843 genes) cluster shows enrichment for carbohydrate and lipid metabolism, congruent with previous studies indicating nutrition to be vital for PV oocyte growth in female mosquitoes [46, 58]. Of those,

a subset of 1,385 genes displayed elevated expression in response to knockdown of the JH receptor (iMet), illustrating which genes *Met* normally represses [58]. These trends were validated using RT-PCR for seven early genes and nine late genes. LPE genes show enrichment for protein synthesis, signaling the fat body to transfer yolk proteins and other nutrients to the developing eggs. Poor diet females are lacking lipids, glycogen, and protein in their fat body reserves, leading one to speculate that the EPE genes repressed by *Met* are involved with the competency of ovaries to reach the resting stage [48]. The study also identified a trend in the differential expression of the gene sets where maximal down-regulated genes were displayed in early PE (0-6 hours) followed by a rise in up-regulation during late PE (60-66 hours) [58]. Gene ontology demonstrated that throughout PE development the majority of genes up-regulated switch from metabolism to transcription and translation. Overall, they determined the early gene subset are maximally expressed in a low JH environment whereas an opposite trend is observed for the cohort of late genes, requiring a high concentration of JH to maximally express.

Well annotated genes, like *Hairy* and *Kr-h1* fall into the trends of the LPE gene cohort for their maximal expression. *Hairy*, like *Met*, contains a bHLH domain in addition to an Orange domain that facilitates dimerization. Hairy uses a WRPW C-terminal motif to recruit its coactivator, *Groucho*. These domains are common in transcriptional repressors related to the regulation of cell differentiation. *Hairy* is a pair-wise gene involved in correct segmentation and neurogenesis of developing *Drosophila* embryos [68]. In adults, null mutation of these genes results in extra bristled antennae and retarded ovarian development, hence the name Hairy [68-70]. *Kr-h1* is often considered a stage-specification

gene for larval states, mediating the gene expression of JH for pupal commitment in holometabolan insects. It also prepares the adult fat body to transfer nutrients to developing follicles [71]. *Kr-h1* consists of eight C2-H2 type zinc-fingers that bind to DNA sequences with great specificity. Not all eight fingers are required for contact and signaling, creating a visual of fingers playing a piano; each combination holds unique notes of gene repression and activation [66]. This binding is not well understood and studies to determine the different interactions are currently underway in our lab (Ahmed, unpublished).

In adult mosquitoes, double knockdowns of *Hairy* and *Kr-h1* illustrated a synergistic regulation of follicle development [65]. RNAi knockdowns of either gene show an observed decrease in follicle length, but simultaneous knockdown had stronger effects and *Met* knockdown showed the most severe phenotype [65]. Sequencing analysis showed that within the promoters of about two-thirds of the dually repressed genes, 70% of the time the two binding sites were less than 300 bp away from each other [65]. The proximity of the two sites had no bearing in the promoters of the single knockdown gene sets, further strengthening the hypothesis of synergistic actions between transcription factors. Approximately 20% of *Met*-repressed genes are regulated by Kr-h1 and Hairy [65], leaving 80% with unknown regulators (Figure 5).

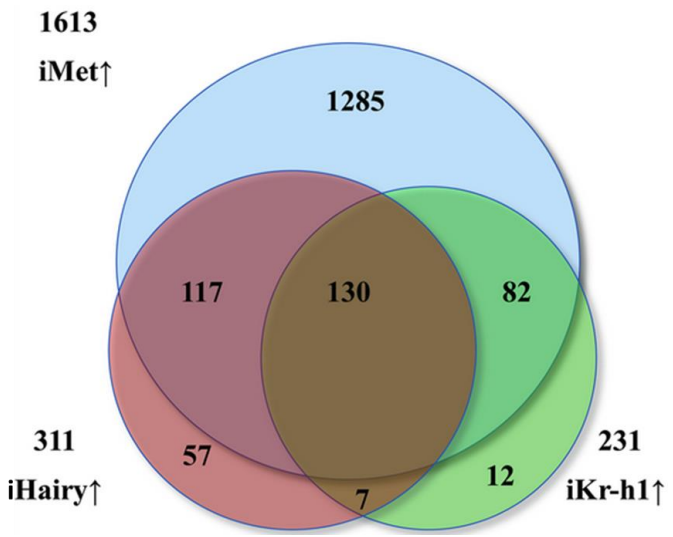


Figure 5: EPE ChIP results copied with permission from Saha et. al. 2019. Regions of overlap highlight multiple genes capable of binding to subsets.

During preparation for vitellogenesis, there are many signaling events occurring within the female. Of the 1,154 genes that were down-regulated by a fold change of ≥1.75 after RNAi-mediated knockdown of *Met*, only 68 genes carried the putative *Met/Tai* binding motif [58]. For 1,334 of the 1,385 genes that were upregulated by a fold change of ≥1.75 after knockdown of *Met*, the canonical *Met/Tai*-binding motif was not found within 2 kb from their transcriptional start sites [58]. These results indicate that the majority of *Met*-dependent genes are controlled indirectly by *Met* and require additional protein partners for effects observed [58]. We hypothesize that additional transcription activators and repressors act downstream of *Met* in the signaling cascades and regulate post-emergence development in adult female mosquitoes. Few effectors immediately downstream of *Met* have been identified to date, and only repressors of the pathway have been annotated thus far [72, 73]. There is a knowledge gap related to the identity and functions of intermediate transcriptional regulators in the JH signaling pathway [21]. Understanding JH's role has its

own set of challenges including its lipophilic nature, low concentration in tissues, tendency to bind non-specifically, and susceptibility to degradation [74].

We hypothesize that these transcriptional effectors work alone or act cooperatively to exact their regulatory nature; there is evidence of transcription factors acting in a collaborative effort for reproductive development [65]. Exacting our effects on the development of the ovaries provides avenues for genetic manipulation resulting in fewer offspring and decreased longevity of females, both of which will help lower the incidence of disease transmission. Through females, we can establish lasting impacts to the germline of this species by targeting genes crucial to their growth and reproductive capabilities. By utilizing RNA sequencing analysis and *in vitro* fat body tissue culture, we plan to identify and characterize new effectors in the JH pathway that function as early JH response genes to prepare newly emerged female mosquitoes for impending vitellogenesis. We will validate *Met* regulation of these effectors by detecting chromatin binding of *Met* to the promoters of these effector genes.

Chapter Two: Bioinformatics Approach for Identifying Candidates

I. Introduction

The intertwining of JH and 20-E signaling are nuanced and obfuscatory. To separate the events from each other, we require females with low to no levels of either hormone. Taking into account the adult profiles of JH and 20-E titers, newly emerged adults (0-2 h PE) are ideal for our purposes.

In our lab, X. Fu performed RNA sequencing of JH-stimulated fat bodies that were cultured *in vitro* to identify alternative splice isoforms of *Taiman*. Fat bodies were dissected from newly emerged female mosquitoes within two hours PE and cultured *in vitro* with 1 µM JH-III or an equivalent volume of the solvent (ethanol) as a negative control for three hours. The RNA libraries were sequenced on an Illumina high-output HiSeq platform, producing 100-bp paired-end reads. The RNA-Seq data have been deposited in the National Center for Biotechnology Information Sequence Read Archive database under accession number SRP136548. Reads were mapped to the genome (AaegL3) using STAR. An expanded transcript annotation was generated using the PASA pipeline with AaegL3.4 as a reference annotation. Alternative splicing events were identified and quantified using rMATS through unpaired differential testing. Differential gene expression analysis was conducted using DESeq2. Functional categories of the differentially spliced genes were determined using the dataset for insects: Nonsupervised Orthologous Groups (inNOG) and their proteins from the eggNOG database (v3.0). Approximately 30 candidates were obtained from this analysis (X.Fu, unpublished data).

As time passed after the original mappings, sequencing technologies improved, and programs that were once standard practice for differential gene expression analysis became out of date. The original pipeline utilized an older version of the *Ae. aegypti* genome, AaegL3.3, as well as Bowtie, TopHat, and Cufflinks to map the reads onto this genome. I performed the analysis of the sequencing data again with the “new Tuxedo” package to identify additional candidate transcription factors with its more sensitive sampling methods (Figure 6).

The *Ae. aegypti* genome is large (approximately 1.25 Gb) and highly repetitive; AaegL3 [75] was unable to produce a contiguous genome fully anchored to a physical chromosome map [76]. The next iteration of assembly, AaegL4 [77], produced chromosome-length scaffolds that made it possible to detect larger-scale syntenic genomic regions in other species but suffered from short contigs, roughly half of the assembly is found on contigs >84 kb and a correspondingly large number of gaps (31,018). The most recent genome (AaegL5.1) utilized sequencing of chromatin conformations (Hi-C) to produce a new reference genome that is highly contiguous and anchored end-to-end to the three chromosomes [78]. This method is based on Chromosome Conformation Capture, in which chromatin is cross-linked with formaldehyde, then digested, and re-ligated in such a way that only DNA fragments that are linked together form products [79]. The ligation products contain the information of not only where they originated from in the genomic sequence but also where they reside, physically, in the 3D organization of the genome [79]. This reference genome was also annotated with a combination of RNA-seq and an assay for transposase-accessible chromatin (ATAC-seq) [78]. This method probes DNA accessibility with a hyperactive Tn5 transposase, which inserts sequencing adapters into accessible regions of chromatin [80]. Sequencing reads are then used to infer regions of increased accessibility, as well as to map regions of transcription factor binding and nucleosome position [80].

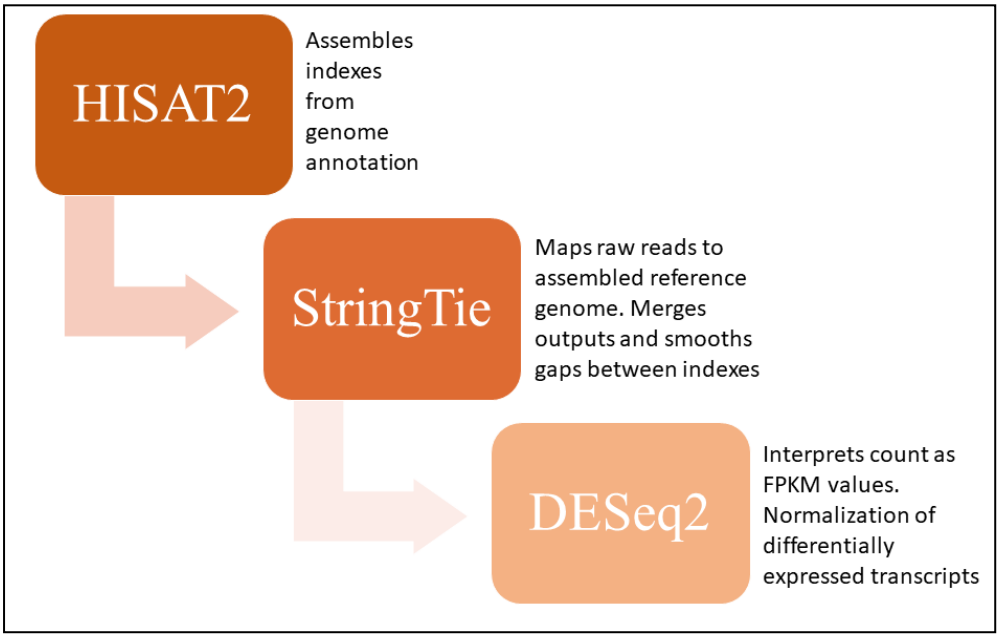


Figure 6: New 'Tuxedo' Pipeline for processing reads from RNA sequenced fat bodies.

II. Methods

A. HISAT2

In regards to the software used to map experimental data to the reference genome, the developers of the 'Tuxedo' package (Bowtie, TopHat, Cufflinks) software package now discourage users from performing their mapping with this outdated software. In 2017, the developers stated that one of their latest updates to the platform involved fixing major issues in Bowtie like corrupted output files and files being skipped during the build when multiple inputs are used; this gives one a lack of confidence in the performance of the software to accurately map, process, and normalize the data. TopHat2 is consistently among the worst performers on both human and malaria T2 and T3 libraries, which contain higher levels of polymorphisms [81].

HISAT2 was developed in 2015 to take the place of Tophat [82]. HISAT stands for hierarchical indexing for spliced alignment of transcripts. This new pipeline has the same fundamental workflow, but the software is more sophisticated, uses less memory, and is capable of analyzing novel splice variants. Any analysis pipeline (Figure 5) can be conceptually divided into four main tasks: (i) alignment of the reads to the genome/transcriptome; (ii) assembly of the alignments into full-length transcripts; (iii) quantification of the expression levels of each gene and transcript; and (iv) calculation of the differences in expression for all genes among the different experimental conditions.

The first two steps of the 'new Tuxedo' pipeline (Figure 5) involve building the index for the reference genome and extracting known splice sites from the reference file (GTF) associated with AaegL5. Because of the large file size, this was run with 8 threads/ processors in parallel with one another. Once the index was made, and with the known splice sites from the GTF as a reference, our experimental RNA-seq data was read. We selected for paired-end reads; one RNAseq fastq file held the forward strand data and another held the reverse for the contigs. Paired-end reads are also able to detect DNA rearrangements such as insertions, deletions, gene fusions, inversions and they improve the ability of the assembler to link together exons belonging to the same transcript [83]. Multiple outputs were generated from this, but for our purposes, the .sam file was necessary for further steps. A SAM file is a tab-delimited text file containing the sequence alignment data. The file contains 11 tabs with the sequence data for each transcript such as the reference name, the position of the first base, the mapping quality

of the read, and the neighboring reads [84]. This needed to be compressed into its binary form (.bam) for more efficient downstream processing.

B. StringTie

In a head to head comparison between Cufflinks (older) and StringTie (newer), StringTie consistently produced more complete and accurate reconstructions of genes and better estimates of expression levels. For example, on 90 million reads from human blood, StringTie correctly assembled 10,990 transcripts, whereas Cufflinks assembled 7,187 transcripts, a 53% increase in transcripts assembled [85]. On a simulated data set, StringTie correctly assembled 7,559 transcripts, which is 20% more than the 6,310 assembled by Cufflinks [85]. As well as producing a more complete transcriptome assembly, StringTie runs faster on all data sets tested to date compared with other assembly software, including Cufflinks.

In my new pipeline, StringTie used the AaegL5 GTF as a reference to assemble the mapped transcripts and provide genomic annotation information with the .bam data of the number of instances per gene they were expressed [86]. StringTie assembles the genes for each replicate separately, estimating the expression levels of each gene and each isoform as it assembles them. After assembling each sample, the full set of assemblies is merged to prevent transcripts in some of the samples partially covered by reads, and consequently, partially assembled [82]. This merging step also adds consistency between the assemblies, allowing real inferences to be made across samples.

An important mapping quality parameter is the percentage of mapped reads, which is a global indicator of the overall sequencing accuracy and of the presence of contaminating DNA [87]. For example, we expect between 70 and 90 % of regular RNA-seq reads to map to a portion of the genome [87]. The average percentage of mapped reads was 76.5% for the total replicates. There were multiple .bam files generated for each experimental data set that then was merged into one file containing all replicates. With the merged GTF file, StringTie then generated text files with the list of transcript abundances corresponding to its genomic identifier. These text files were then imported into R-Studio for further analysis with the DESeq2 package.

C. DESeq2

Count quantification is measured by the number of pairs of reads ("fragments" where one or both ends of a fragment are sequenced), which are normalized based on the total number of fragments sequenced (measured in millions) and by the length of the transcript (measured in kilobases), giving an estimate measured as fragments per kilobase of transcript per million fragments (FPKM) [85]. FPKM is preferable to RPKM (reads) in the case of paired-end RNA-seq experiments, where in some cases one of the two reads belonging to a fragment might be unmapped, possibly leading to underestimates of expression [85].

Transcript abundances were counted and normalized with the DESeq2 package in R-Studio [82 88]. An alternative to DESeq2's htseq-count function is the Cufflinks cuffdiff. The cuffdiff route provides an analysis of two-group experiments, whereas a counts table

provides the raw material for richer experimental designs [89]. It is for this reason that ddsHTSeq was used instead of cuffdiff.

This pipeline's stronger mapping of splice isoforms and normalization for gene count data yielded approximately 34,000 total genes in the cultured fat body after incubation with JH. These reads were filtered using DESeq2 in R Studio with a baseMean value >0, a >2-fold expression (1.0 in log2 scale), and a p-adjusted value of <0.01 which yielded 2,532 differentially expressed transcripts.

III. Results

These genes were hand-screened and filtered using NCBI's Entrez search function to identify corresponding gene names to their gene ID determined from ddsHTSeq; 61 transcripts showed promise as possible transcription factors. It is important to note that most of these transcripts were labeled as 'uncharacterized' as they had only been identified through bioinformatics analysis at this point and had no functional data to add to the annotation. Additionally, the two pipelines used to map the reads to their respective reference genomes obtained lists with similar candidates, such as AAEL002499 (Transcription factor Pipsqueak). Positive control genes such as *Kr-h1* were also identified through this filtering process, adding confidence to the results generated from the former method. These 61 transcripts were then analyzed for DNA quality and nucleotide length. Transcripts less than 300 nucleotides were discarded, the reasoning being that if they are too small or nonspecific, it would be very difficult to design primers and confirm their presence experimentally.

Next, the Simple Modular Architecture Research Tool (S.M.A.R.T) database was used to identify whether there was a presence of DNA-binding motifs in the protein FASTA sequences of the upregulated transcripts. S.M.A.R.T utilizes HMMER to identify known binding motifs and involved gene regulatory networks [90]. HMMER uses profile hidden Markov models (profile HHMs) to create probabilistic models for sequence homologs [91]. Sequences that score significantly better to the profile-HMM compared to a null model are considered to be homologous to the sequences that were used to construct the profile-HMM [91]. DNA binding motifs such as ankyrin repeats, zinc fingers, and homeobox domains are just some examples of what was sought after. My new pipeline yielded 17 additional candidates to undergo real-time PCR along with our previously identified 30 candidates (Table 1).

Table 1: Vectorbase Genes and Known Domains

ID	Name	Motifs
AAEL000287	zinc finger protein 519	ZnF
AAEL000402	cAMP REBP A	bZIP-CREB3
AAEL000849	ZF-43	ZnF
AAEL000894	huncback	ZnF
AAEL001373	LIM/HOX	Zn homeodomain
AAEL001569	dnaJ homolog	hsp70 chaperone
AAEL002456	b-cell lymphoma/leukemia	ZnF
AAEL002499	BTB-domain containing protein	BTB/POZ domain, AT hook
AAEL002853	CCAAT/enhancer-binding protein	bZIP-CEBP
AAEL003014	ZF-704	ZnF
AAEL003163	forkhead	"winged helix"
AAEL003327	ush	ZnF
AAEL004707	hhex	HOX

AAEL005480	protein hairy	Orange
AAEL006314	upstream stimulatory factor 1	HLH
AAEL006533	DNA-binding protein D-ETS-6	SAM-PNT & ETS
AAEL006690	protein bric-a-brac 1	BTB & HTH-psq
AAEL007040	does not exist in Vectorbase	
AAEL007202	nuclear hormone receptor HR96	NR-LBD-F1/CAR
AAEL007624	NF-kappa-B p110	IPT(Ig-like, plexins,trx factors) and ankyrin repeats
AAEL007696	embryonic polarity protein dorsal	IPT and a coiled-coil region
AAEL008005	hamlet	ZnF
AAEL008524	grauzone	zf-AD and ZnF
AAEL008528	tyrosine-protein phosphatase corkscrew	SH2 and PTPc
AAEL008551	grauzone	zf-AD and ZnF
AAEL008953	kayak	BRLZ (leucine zipper)
AAEL009092	ceramide synthase 6	homeodomain
AAEL009171	protein CREBRF homolog	bZIP-Aureo like
AAEL009212	lola (ABDL)	BTB and ZnF
AAEL009325	uncharacterized LOC5571806	transmembrane regions
AAEL009442	ATF3	BRLZ (leucine zipper)
AAEL009645	uncharacterized	bZIP
AAEL010222	GATA binding factor-4	ZnF-GATA
AAEL010513	cwo	Orange
AAEL010555	sterol regulatory element-binding protein 1	HLH and HMMR-C
AAEL011316	ZF-704	zf-AD and ZnF
AAEL011323	HNF-4gamma	ZnF and HOLI
AAEL011359	NF-activated Tcells 5	IPT
AAEL011371	uncharacterized LOC5574731	BRLZ and coiled coil
AAEL012224	oocyte zinc finger protein XlCOF6	Zf-AD, FOG, SFP1, C2H2
AAEL013752	Rfx5	"winged helix" RFX-like
AAEL014754	cytochrome b binding site	cytochrome b5 site

Chapter Three: Experimental Validation of JH Early Response Genes

I. Fat Body Culture (JH vs. Ethanol)

After preliminary bioinformatics analysis and primer design, 39 candidates underwent *in vitro* fat body culture to validate their presence and differential expression after JH treatment (Figure 7). Some of the genes identified through the old Tuxedo pipeline did not have corresponding Vectorbase FASTA sequences to design primers from. This is due to the data cooresponding for probes relating to other genes, however those specific genes did not show elevated expression in the presence of JH.

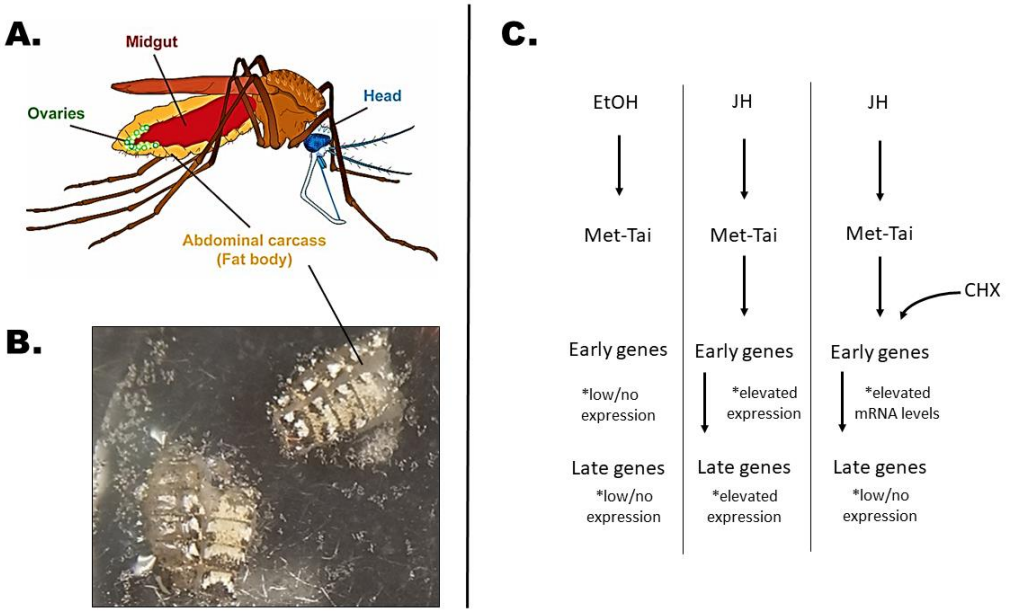


Figure 7: (A) Overview of mosquito biology and location of fat body tissues (yellow). (B) Dissected fat bodies in complex culture media. (C) Experimental design for *in vitro* fat body tissue culture.

The fat body is the insect equivalent of mammalian liver and adipose tissues and is regularly used in the field for interpreting genetic signaling related to reproduction and metabolism. By separating the reproductive tissues from the CA, where JH is synthesized, and then stimulating the JH pathway *in vitro*, we are able to have unencumbered gene signaling in the newly emerged adult females (Figure 7). Fat bodies were collected from female mosquitoes within two hours of emergence and cultured in complete fat body medium. These were incubated in a 24-well plate each containing

500µL complete fat body medium plus 5µM JH-III for three hours. All tissue sample groups (~8 females/replicate) were collected in 200µL TRIzol® and immediately ground with a motorized pestle to provide total RNA complete immersion in the chemical fixative. An additional 800µL were added to a total volume of 1mL/ sample. Samples

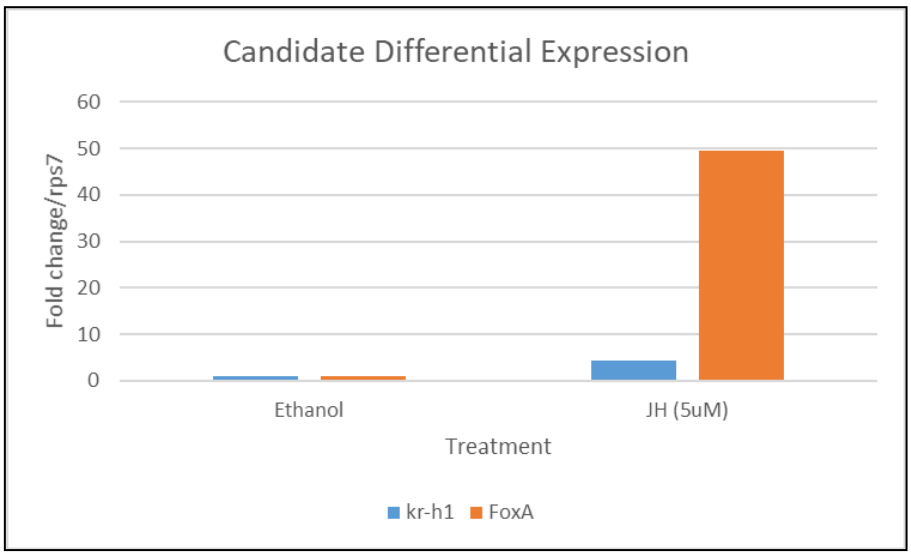


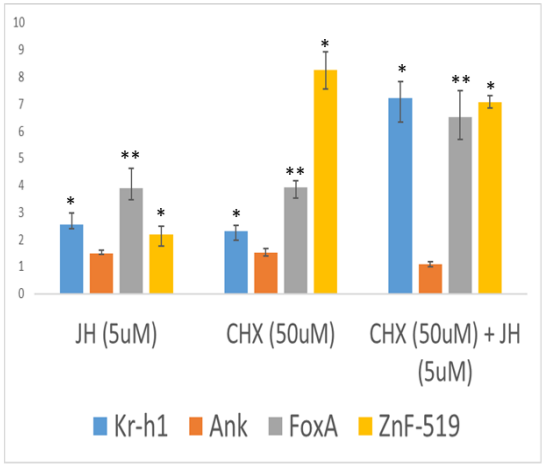
Figure 8: RT-PCR results for 3h cultured fat bodies, three replicates each. N=8 fat bodies. Results normalized against ribosomal protein 7.

were stored at -20°C until RNA extraction. Following total RNA extraction and cDNA synthesis, RT-PCR with primers specific for each candidate gene was performed. Relative expression of the candidates was calculated using ribosomal protein 7 (rps7) for the negative control and transcript abundance was quantified with the $2^{(-\Delta\Delta C_t)}$ method to determine the overall fold change of the gene (Figure 8). Kr-h1 was used as a positive control for this analysis, as it has been determined to always display an expression >2 fold (1.0 in log2 scale) and is an early gene product.

II. Fat Body Culture (Cycloheximide Treatment)

Further candidate screening involved tissue culture in the presence of an mRNA translation inhibitor, cycloheximide (CHX) (Figure 9). Fat bodies were collected at the same timepoints as stated in the previous section and were either treated with 50µM CHX for three hours or had a one hour pretreatment of the inhibitor followed by a washing of the fat bodies in complete medium and a three hour treatment in the presence of 50µM CHX and 5µM JH-III. The tissues were preserved and stored in the same way as described in the previous section. The understanding of CHX is that if the candidates are a true early gene of the JH pathway and directly regulated by *Met* [92], then analysis with real-time qPCR would show that this inhibitor is unable to abolish the JH-induced gene expression. The inability to diminish expression would demonstrate that this candidate is a primary product of *Met* and does not require additional protein synthesis for expression (Figure 7).

A.



B.

Gene ID	JH-III	Cycloheximide	CHX + JH
Kr-h1	2.32	2.06	6.91
FoxA	2.66	4.46	3.04
ZnF-519	2.19	8.25	7.06
Ank	1.48	1.52	1.08

Figure 9: RT-PCR results for 3h cultured fat bodies with or without mRNA inhibitor present, three replicates each. N=8 fat bodies. Results normalized against rps7 gene.

From this data, two genes showed promise as early gene products: FoxA and Zinc finger-519. It is important to note that this work is ongoing, FoxA has had 2 experiments completed and ZnF-519 has had one experiment. A pairwise T-test with a Bonferonni correction was performed for the candidates due to the small sample size per treatment (n=3). There was a significant effect for the two candidates at each of the treatments ($p<0.05$), however for FoxA the significance had a p-value<0.01 for each of the treatments. Additionally, the T-test for the negative control, an ankyrin repeat, was insignificant in regards to the effects of the treatments on gene expression, adding confidence to the data and to the abolishment of expression in the presence of cycloheximide. When examining Figure 9, one can see an elevated expression for the genes in the presence of the inhibitor; this goes against our understanding of inhibition. However, when one takes into account the disruption of negative feedback loops in the presence of cycloheximide, overexpression of these genes makes sense. It is known that Kr-h1 binds to itself to provide control over its expression, when the protein product of Kr-h1 is absent, there is no regulatory control over the expression. It remains to be seen whether FoxA and ZnF-519 regulate themselves or require additional protein partners for feedback.

A. Current knowledge of candidate transcription factor- FoxA

The Forkhead box (Fox) family of transcription factors are abundant, with over 2000 identified members in 108 species of animals and fungi [93]. The first Fox protein (FOXA) was identified as a homeotic gene in *Drosophila* in 1989 [94]. Its binding domain was delineated in the year following [95]. As of 2011, 19 subfamilies (A-S) have been

catalogued [96]. Insect Fox genes form well-resolved subfamilies corresponding to the human subfamilies A-Q [97]. The Forkhead domain (FHD) contains three N-terminal α-helices, three β-strands, and two loops towards its C-terminal region that is reminiscent of butterfly wings fluttering, hence it has also been referred to as a winged-helix domain (Figure 10) [98]. FOXA3/HNF3γ was the first member of this family to be resolved through crystallography [98]. These hepatocyte nuclear factors are transcriptional activators for liver-specific transcripts such as albumin and transthyretin and provide chromatin remodeling. Fox proteins tend to bind DNA as monomers but have demonstrated the

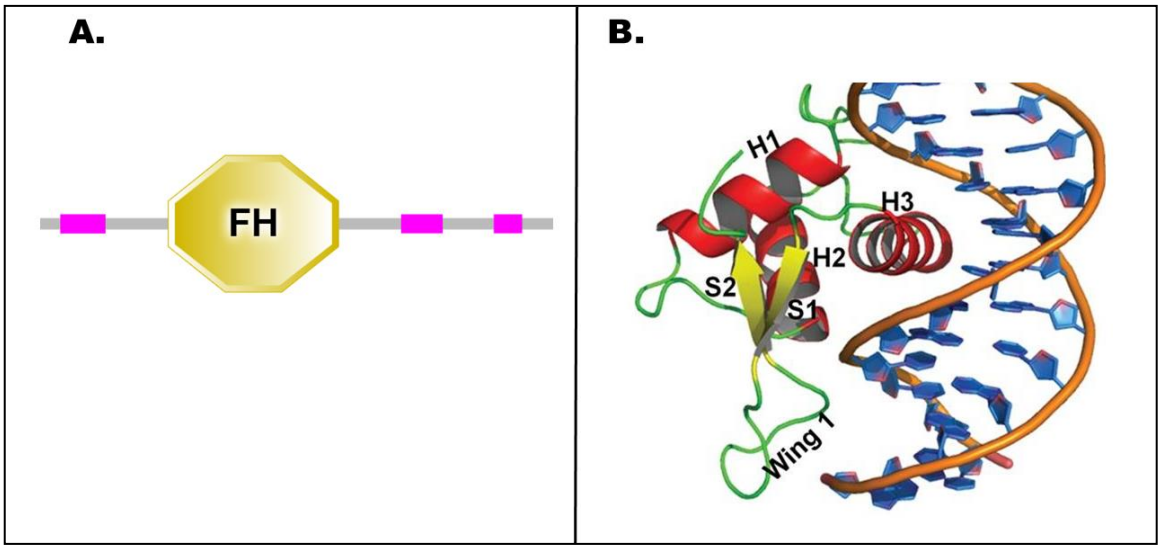


Figure 10: (A) SMART analysis of candidate FoxA protein FASTA. Regions in pink are low complexity. One Forkhead domain. (B) Pymol rendered crystal structure of Forkhead family member FoxP while bound to DNA. "Winged helix" characteristic in this family of transcription factors. Secondary structure elements are labeled. Adapted with permission from Li J. et al. 2017. Biochemistry. 56, 29, 3745-3753 Copyright 2020 American Chemical Society.

ability to create homo- and heterodimers of the various subfamilies for higher levels of organization [98-104]; they also can interact with co-activators, co-repressors, and other proteins [105, 106]. Fox-containing proteins have been implicated in processes related to development, metabolism, aging, immunoregulation, and tumorigenesis [107-111]. Fox

factors act as terminal effectors of many signal transduction pathways and hypothetically constitute nodes in cellular networks to allow cross talk and mediation of signaling in response to external fluctuations in the environment [93].

Many Fox factors control morphogenesis and cell differentiation during embryogenesis. In adults, they control carbohydrate and lipid metabolism, stress response, and energy homeostasis [96, 111]. For example, FOXA2 controls development of the liver and pancreatic tissue during embryogenesis and mediates insulin signaling and gluconeogenesis in the adult [112]. The highest expressional levels of hFOXA are found in the prostate and bladder (NCBI, Gene ID: 3169).

The best characterized invertebrate Fox gene is FOXO, a target of the insulin signaling pathway. dFOXO regulates insulin signaling in the brain and fat body as well as controlling fertility and lifespan through transcriptional feedback control of the insulin receptor [113, 114]. Fox factors have been suggested to be involved in mosquito reproduction; *in vitro* expressed dFOXA can bind to the promoter of the vitellogenin gene (Vg) and may work in coordination with other transcription factors to direct the tissue-specific expression of yolk protein precursor (YPP) genes in mosquito fat bodies [115]. The winged-helix fold is similar in structure to linker histones; hFOXA1 is known to direct nucleosome positioning at the albumin enhancer by binding nucleosomes with similar affinity of histones 1 and 5 [107]. This binding results in the opening of chromatin, transcriptional reactivation of the locus, and context-specific responses to ligand binding [116]. This suggests that the Forkhead family regulates gene expression in an

unconventional way by regulating the local state of chromatin at target loci, independent of their abilities as transcriptional enhancers/repressors [93].

RNA was isolated from previtellogenic females 72h PE and 24h PBM to determine the expressional profiles of each Fox gene [97]. Three Foxes shared a broad expressional profile through the head, midgut, malpighian tubules, thorax, fat body, and ovaries: FOXK1, FOXK2, and FOXO. FOXA was expressed in the thorax, midgut, and malpighian tubules of PV females and in the midgut and malpighian tubules of blood-fed females. One point to note is that the Vg gene contains a FOXA binding site, yet because of its lack of expression in the fat body, it was not considered for RNAi studies. RNAi of several FOX factors expressed in the ovaries followed a rise and fall of transcript abundance relative to the blood meal and resulted in a significant reduction of eggs deposited, suggesting that most Fox family members are expressed dynamically in response to the female's gonotrophic cycle. It would be a benefit for this study to be performed so as to see the indirect relationships to mosquito reproduction, knowing FOXA's roles in metabolism and development in humans as well as the Fox family's ability to mediate chromatin remodeling against conventional methods.

B. Current knowledge of candidate transcription factor- ZnF-519

Previous data collected on this gene was performed though bioinformatics analysis and domain prediction software. The most recent annotation of the *Aedes aegypti* genome allowed the identification of this gene. NCBI predicts the 371 amino acid protein to contain a zinc-finger associated domain (zf-AD) at the C-terminus and six C2H2 zinc

finger domains at the N-terminus (Figure 11). The only functional data for this novel gene comes from a whole transcriptome RNA-seq approach of 34 distinct time points in *Aedes* development [117].

This approach increased the number of annotated genes more than 2-fold, identifying sex-specific and sex-differential expression profiles for ovaries, early embryos, testes, and adult male and female somatic tissues [117]. Of interest to this thesis, the gene

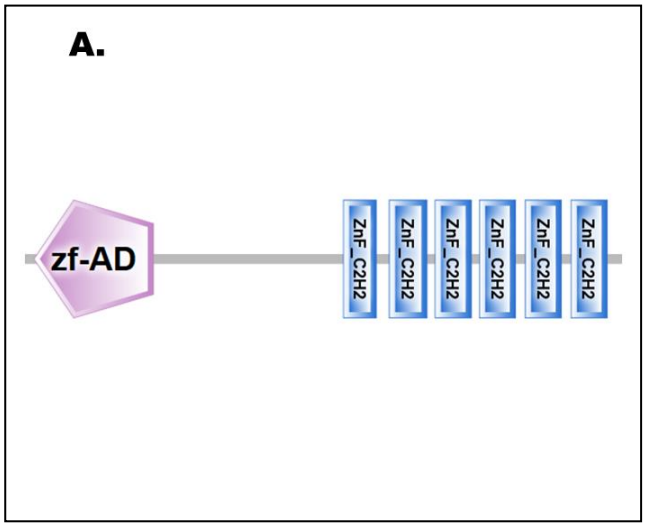


Figure 11: (A) SMART analysis of candidate ZnF-519 protein FASTA. Zinc-finger associated domain is colored purple, six C2H2-type zinc fingers colored in blue.

candidate displays an 11-fold expression increase in blood-fed ovaries compared to non-blood fed females as well as a pattern of maintained expression as the female completes her first gonotrophic cycle. This high expression (12-fold) is maintained in newly laid embryos and decreases to 1.69-fold as the embryo matures and hatches into its first larval instar [117].

The two gene candidates had their promoter regions analyzed for the canonical E-box sequence of *Met* and primers were designed to encompass these 5' regulatory regions for chromatin immunoprecipitation (ChIP). PCR primers were designed to encompass the *Met* E-box motif 'CACGTG' in the 5'UTR (Figure 12). FoxA had two E-box motifs located on the reverse strand of the 5'UTR, within 150 bp from the transcription start site (TSS). ZnF-519's was also found on the reverse strand, 2.5kb away from its' TSS. Motif enrichment analysis using CentriMo in the MEME Suite package revealed approximately 1,400 different motifs within the first 1000 bp of FoxA's promotoer region. These were examined using the *Drosophila* sequence databases, however, no preferred binding motif was identified. These data will allow us a wider vantage when desiging future experiments aimed at examining the effects of these transcription factors that initiate downstream gene signaling.

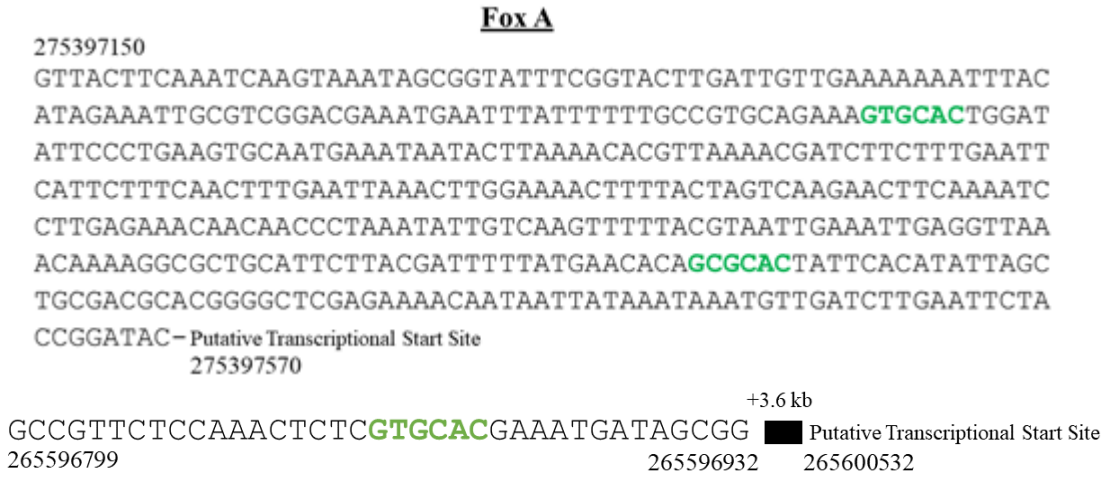


Figure 12: Promoter regions of candidate transcription factors featuring Met binding E-box motif (green text).

I. Materials and Methods

a. Mosquito rearing

Mosquitoes used in all experiments were *Aedes aegypti*, Liverpool strain. Mosquito larvae were raised in distilled H_2O (~200 larvae per 1.5 liter tray) and fed hand-ground TetraMin® fish food (Tetra Holding Inc.). Pupae were transferred to a large cage prior to adult emergence. Adults were fed 10% sucrose from soaked cotton balls. All life stages were raised at 28°C and 70% humidity in a 16 hr light / 8 hr dark cycle.

b. *In vitro* fat body tissue culture

Within two hours post-emergence, mosquitoes were collected and anesthetized in 4ºC cooler for 15 minutes. They were placed on ice for the duration of the dissections. Prepared fat bodies were treated with room temperature ethanol, 5μM JHIII, 50 μM cycloheximide, or with a combination of the hormones in a complex medium containing 137 mM NaCl, 1.2 mM Calcium chloride, 1 mM KCl, 0.6mM MgCl hexahydrate, and 1.8 mM sodium carbonate.

Total RNA was isolated from cultured mosquito fat bodies using a TRIzol reagent and a Direct-zol RNA MiniPrep Plus kit. Genomic DNA contamination was completely removed by incubating the total RNA with DNase I (Thermo Fisher Scientific) for 15 min at room temperature. Three micrograms of the purified RNA was used for reverse transcription to synthesize cDNA. cDNA synthesis was performed with Qiagen OmniScript RT Kit according to manufacturer's instructions. PCR product was diluted to a working concentration of 7.5 ng/μL for RT-PCR.

c. RT-PCR

Primers were designed using the NCBI tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Ideal primer parameters include a PCR product size between 120-200bp in length, a max Tm difference of two degrees; the default parameters for the program should otherwise remain the same. Of the suggested primer pairs, a 5'/3' self-complementarity max cutoff of 4 was established. Primers were prepared by Sigma Genosys Oligos (Sigma-Aldrich) to an OD of 3.0 mmol and a desalted purification. BioRad CX910 was used to perform all real-time PCR experiments.

Chapter Four: Conclusions and Discussion

The work presented in this thesis addresses the presence of novel transcription factors in the juvenile hormone pathway downstream of the JH receptor, Methoprene-tolerant. Utilizing bioinformatics analysis of RNA sequencing data, approximately 2,500 genes were significantly upregulated in the presence of JH-III. This sequencing data was originally run using an older pipeline ('old Tuxedo') and discovered 30 potential transcription factors. Using a more robust analysis method ('new Tuxedo') with similar cutoff parameters, 15 additional transcription factor candidates were found, creating a short list of 47 candidate transcription factors for experimental validation. These candidates were selected for significant differential expression in response to the addition of JH-III at 3 hours, 4 hours, and 6 hours following emergence of the female mosquito. Through *in vitro* fat body tissue culture, we were able to experimentally validate the results from RNA sequencing as well as determine whether these candiates function as early genes through inhibition studies utilizing cycloheximide. From preliminary qPCR

results, two candidates maintained levels of expression in the presence of the inhibitor, FoxA and ZnF-519.

Preliminary analysis of their expression profiles revealed that the Fox family is involved in processes with cell aging and chromatin remodeling. FoxA displays consistent expression through the different time points of the Aedes life cycle in Akbari, 2013. It did not display a specificity to either sex, and results were consistent from Hansen, 2007, where it displayed tissue specificity to the thorax, midgut, and malpighian tubules. This is in contrast with the results from our in vitro fat body tissue culture, where we determined expression in the fat bodies. There is not much information available for ZnF-519, but it is known to peak expression in developing ovaries and following blood feeding. Potential for a false positive gene candidate is possible with midgut cell contamination, these tissues must be fully removed first during the dissection process. However, this remains to be seen; three replicates of each gene were unable to be performed, FoxA had two replicates completed and ZnF-519 had one completed. Therefore, an immediate future direction of this work would be to complete the qPCR replicates of the fat body culture. From there, RNAi studies of these two genes should be performed to determine their function *in vivo* and whether there is a direct role of these factors related to reproduction. RNAi knockdowns coupled with RNA sequencing of the tissues will provide us identification of the subsets of genes affected by these two candidates and whether these candidate effectors share genes affected or work synergistically. As such, these two genes had their promoter regions analyzed for the presence of the canonical E-box motif for *Met* binding. Both genes contained this motif in the reverse strand of their promoter

regions. Due to time constraints and environmental pressures outside of our control, testing the direct binding of *Met* was unable to be experimentally validated.

Future work will compare the DNA binding sites of *Met* and the effectors and determine whether these share binding to the same subset of genes. Previous work from other labs demonstrated that there is often a Kruppel binding site or a Hairy binding site near the *Met* E-box in shared subsets of genes. For this reason, the promoter regions of these two genes were analyzed for the canonical KBS and HBS sequences and were determined to contain only Hairy binding sites (N-boxes). N-box motifs function similarly to the E-box motif but contains one less variable nucleotide (CACNAG vs CANNTG, respectively). Additionally, N-box sequences often come with a class C binding site nearby (CACG[A/C]G). FoxA contained four N-box motifs and one class C motif in its promoter region and ZnF-519 contained six N-box motifs and three class C motifs. A future direction of this work may be to additionally test the direct binding of Hairy and determine whether these candidate effectors exist in hierarchy with Hairy.

Elucidating additional effectors of the JH pathway will help provide a field of view whereby researchers may develop novel control strategies for biological chokeholds in this pathway. Additionally, by identifying other players in the signaling hierarchy, the field will be able to have a complete knowledge of the signaling events surrounding *Met*. The JH pathway is highly conserved in insects, allowing the possibility for expanding beyond *Ae. aegypti* control into a more broad pest control strategy.

Appendix

Terminal

Build index folders and identify splice regions within reference genome

1. hisat2-build -p 8 AaegL5.fa AaegL5_hisat2

2. hisat2_extract_splicesites.py -f AaegL5_genomic.gtf -o splice_sites.txt

Input RNA sequencing data and align to known splice sites

3. hisat2 -x AaegL5_hisat2 -known-splicesite-infile splicesites.txt -p 2 -1 pair_read_R1.fastq -2 pair_read_R2.fastq | samtools view -bS > new.file.bam

Sort transcripts, compress file, and merge calls to prevent fragment exclusion

4. samtools sort -@ 8 new.file.bam sampleID.sorted
5. stringtie -p 8 -G path/to/AaegL5_genomic.gtf -l path/to/bam/folder sampleID.sorted.bam
6. stringtie --merge -p 8 -G path/to/AaegL5_genomic.gtf -o merged.gtf mergelist.txt
7. stringtie -e -B -p 8 -G path/to/merged.gtf -o sampleID/ballgown sampleID.gtf path/to/sorted/bam

Turn merged transcripts into count data for FPKM analysis in R Studio

8. htseq-count accepted_hits.bam AaegL5.gtf -i Parent -f bam > E31_R1_exon.count (iterated for each replicate)

R-Studio

```
sampleTable = data.frame(
      sampleName =
c("J3R1_exon.count", "J3R2_exon.count", "J3R3_exon.count",
"E3R1_exon.count", "E3R2_exon.count", "E3R3_exon.count"),
      filename =
c("J3R1_exon.count", "J3R2_exon.count", "J3R3_exon.count",
"E3R1_exon.count", "E3R2_exon.count", "E3R3_exon.count"),
      condition = ("exp", "exp", "exp", "con", "con", "con"))

ddsHTSeq <- DESeqDataSetFromHTSeqCount(sampleTable = sampleTable,
directory = "~/path/to/exon.count", design = ~ condition)
dds <- DESeq(ddsHTSeq)
res <- results(dds)
res = as.data.frame(res)
normvalue = counts(dds, normalized=TRUE)
j3hr = apply(normvalue[,1:3],1,mean)
e3hr = apply(normvalue[,4:6],1,mean)
res = cbind(res,j3hr,e3hr)
filterRes = res[res$baseMean>0,]
filterRes = res[res$log2>1,]
filterRes = res[res$padj<0.01,]
View(filterRes)
```

Entrez

(https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?db=gen

e&id=

111111

222222…

&retmode=text&retmax=10000)

Table 2: RT-PCR Results of Fat Body Tissue Culture

Vectorbase ID	E	J	C	CJ
AAEL000287	1	2.19159	8.25211	7.063014
AAEL000849	1	1.184519	2.20098	0.928798
AAEL000894	1	0.379384	0.349431	0.256317
AAEL001373	1	0.974142	0.184378	0.550539
AAEL001569	1	1.092982	1.952674	0.917459
AAEL002353	1	0.690345	1.567224	1.059608
AAEL002456	1	1.24944	3.408322	3.528523
AAEL002499	1	0.391727	0.288867	0.184753
AAEL003014	1	1.534153	1.470088	1.525587
AAEL003075	1	0.816392	3.032542	3.045806
AAEL003163	1	2.658278	4.455719	3.039054
AAEL003678	1	0.850636	0.898837	1.065646
AAEL003726	1	0.909769	2.936109	3.629351
AAEL003737	1	0.84728	2.633005	2.318977
AAEL003740	1	1.481848	1.523781	1.083283
AAEL004707	1	3.030426	0.42065	0.22544
AAEL005152	1	1.345637	3.735759	4.173
AAEL007624	1	0.347795	0.074386	0.136417
AAEL007696	1	0.865384	0.433895	0.647655
AAEL008005	1	N/A	N/A	N/A
AAEL008524	1	1.006705	1.981886	2.30521
AAEL008528	1	0.998309	3.214394	4.048198
AAEL008551	1	0.874435	1.383345	1.083086
AAEL008953	1	0.45621	0.969473	1.218875
AAEL009171	1	0.76151	2.374924	2.549191
AAEL009212	1	0.694626	0.277647	0.479851

AAEL009325	1	1.368591	1.195621	1.605935
AAEL009442	1	0.660424	1.305832	1.854804
AAEL009626	1	0.586864	2.032569	1.574629
AAEL009645	1	0.426202	2.013678	2.886586
AAEL010222	1	0.651672	0.90127	1.108845
AAEL010513	1	1.158192	1.249327	2.763343
AAEL011056	1	0.697482	2.897289	2.054972
AAEL011316	1	1.013242	1.416431	1.450337
AAEL011323	1	1.08359	0.898868	0.863984
AAEL011338	1	0.487062	2.671772	1.512433
AAEL011359	1	1.021513	0.07404	0.568567
AAEL011371	1	0.859155	3.683568	5.9854
AAEL011943	1	0.613147	1.346473	1.819301
AAEL013752	1	1.009429	1.672972	2.550371
AAEL014695	1	0.948756	1.857582	2.152611
AAEL014743	1	1.588874	0.748689	1.45357
AAEL014754	1	0.825714	0.072928	0.246651
Kr-h1	1	2.31623	2.075931	6.90613
LOC110675204	1	0.996016	3.599033	2.501835

References

[1] Zhu, J., Miura, K., Chen, L., and Raikhel, A. (2000) AHR38, a homolog of NGFI-B, inhibits formation of the functional ecdysteroid receptor in the mosquito Aedes aegypti, *The EMBO journal 19*, 253-262.

[2] Kamita, S. G., Samra Ai Fau - Liu, J.-Y., Liu Jy Fau - Cornel, A. J., Cornel Aj Fau - Hammock, B. D., and Hammock, B. D. (2011) Juvenile hormone (JH) esterase of the mosquito Culex quinquefasciatus is not a target of the JH analog insecticide methoprene, *PLOS ONE 6*, e28392.

[3] Organization, W. H. (2019) Dengue and Severe Dengue Fact Sheet.

[4] Murillo, D., Murillo, A., and Lee, S. (2019) The Role of Vertical Transmission in the Control of Dengue Fever, *Int J Environ Res Public Health 16*, 803.

[5] Joshi, V., Mourya Dt Fau - Sharma, R. C., and Sharma, R. C. Persistence of dengue-3 virus through transovarial transmission passage in successive generations of Aedes aegypti mosquitoes.

[6] Adams, B., and Boots, M. (2010) How important is vertical transmission in mosquitoes for the persistence of dengue? Insights from a mathematical model, *Epidemics 2*.

[7] Choy, M. M., Sessions, O. M., Gubler, D. J., and Ooi, E. E. (2015) Production of Infectious Dengue Virus in Aedes aegypti Is Dependent on the Ubiquitin Proteasome Pathway, *PLoS Negl Trop Dis 9*, e0004227-e0004227.

[8] Diseases, D. o. V.-B. Mosquito Life Cycle: *Aedes aegypti*, (Control, C. f. D., Ed.).

[9] Ellis, R. (2019) Where Disease-Carrying Mosquitoes Will Go In The Future, National Public Radio, Inc.

[10] Brady, O. J., et al. (2012) Refining the global spatial limits of dengue virus transmission by evidence-based consensus, *PLoS Negl Trop Dis 6*.

[11] Bhatt, S., et al. (2013) The global distribution and burden of dengue, *Nature 496*, 504-507.

[12] Gómez-Dantés, H., and Willoquet, J. R. (2009) Dengue in the Americas: challenges for prevention and control, *Cadernos de Saúde Pública 25*, S19-S31.

[13] Bass, C., Puinean, A. M., Zimmer, C. T., Denholm, I., Field, L. M., Foster, S. P., Gutbrod, O., Nauen, R., Slater, R., and Williamson, M. S. (2014) The evolution of insecticide resistance in the peach potato aphid, Myzus persicae, *Insect Biochemistry and Molecular Biology 51*, 41-51.

[14] Lu, C. (2014) Sub-lethal exposure to neonicotinoids impaired honey bees winterization before proceeding to colony collapse disorder, *Bulletin of Insectology 67*, 125-130.

[15] Alison Benjamin, A. H., and Ruth Spencer. (2013) Buzzfeeds: the effects of colony collapse disorder and other bee news, The Guardian.

[16] Goodman, W. G., and Cusson, M. (2012) 8 - The Juvenile Hormones, In *Insect Endocrinology* (Gilbert, L. I., Ed.), pp 310-365, Academic Press, San Diego.

[17] Dhadialla, T. S., Carlson, G. R., and Le, D. P. (1998) New Insecticides With Ecdysteroidal and Juvenile Hormone Activity, *Annual Review of Entomology 43*, 545-569.

[18] Bede, J. C., and Tobe, S. S. (2000) Insect Juvenile Hormones in Plants, In *Studies in Natural Products Chemistry* (Atta ur, R., Ed.), pp 369-418, Elsevier.

[19] Wigglesworth, V. B. (1958) Some methods for assaying extracts of the juvenile hormone in insects, *Journal of Insect Physiology 2*, 73-84.
[20] Truman, J. W. (2019) The Evolution of Insect Metamorphosis, *Current Biology 29*, R1252-R1268.
[21] Jindra, M., Palli, S. R., and Riddiford, L. M. (2013) The juvenile hormone signaling pathway in insect development, *Annu Rev Entomol 58*, 181-204.
[22] Raikhel, A. S., Brown, M. R., and Belles, X. (2005) 3.9 - Hormonal Control of Reproductive Processes, In *Comprehensive Molecular Insect Science* (Gilbert, L. I., Ed.), pp 433-491, Elsevier, Amsterdam.
[23] Wyatt, G. R., and Davey, K. G. (1996) Cellular and Molecular Actions of Juvenile Hormone. II. Roles of Juvenile Hormone in Adult Insects, In *Advances in Insect Physiology* (Evans, P. D., Ed.), pp 1-155, Academic Press.
[24] Li, Y., Hernandez-Martinez S Fau - Unnithan, G. C., Unnithan Gc Fau - Feyereisen, R., Feyereisen R Fau - Noriega, F. G., and Noriega, F. G. (2003) Activity of the corpora allata of adult female Aedes aegypti: effects of mating and feeding, *Insect Biochem Mol Biol 33*.
[25] Hernández-Martínez, S., Mayoral, J. G., Li, Y., and Noriega, F. G. (2007) Role of juvenile hormone and allatotropin on nutrient allocation, ovarian development and survivorship in mosquitoes, *Journal of insect physiology 53*, 230-234.
[26] Shapiro, A. B., Wheelock, G. D., Hagedorn, H. H., Baker, F. C., Tsai, L. W., and Schooley, D. A. (1986) Juvenile hormone and juvenile hormone esterase in adult females of the mosquito Aedes aegypti, *Journal of Insect Physiology 32*, 867-877.
[27] Kimball, J. W. (2008) Insect Hormones, Biology Pages, Internet Archive Wayback Machine.
[28] Thummel, C. S., and Chory, J. (2002) Steroid signaling in plants and insects—common themes, different pathways, *Genes & Development 16*, 3113-3129.
[29] Kozlova, T., and Thummel, C. S. (2000) Steroid Regulation of Postembryonic Development and Reproduction in Drosophila, *Trends in Endocrinology & Metabolism 11*, 276-280.
[30] Li, T.-R., and White, K. P. (2003) Tissue-Specific Gene Expression and Ecdysone-Regulated Genomic Networks in Drosophila, *Developmental Cell 5*, 59-72.
[31] Riddiford, L. M., Cherbas, P., and Truman, J. W. (2000) Ecdysone receptors and their biological actions, In *Vitamins & Hormones*, pp 1-73, Academic Press.
[32] Thummel, C. S. (2001) Molecular Mechanisms of Developmental Timing in C. elegans and Drosophila, *Developmental Cell 1*, 453-465.
[33] Thummel, C. S. (2002) Ecdysone-regulated puff genes 2000, *Insect Biochemistry and Molecular Biology 32*, 113-120.
[34] Brown, M. R., Graf, R., Swiderek, K. M., Fendley, D., Stracker, T. H., Champagne, D. E., and Lea, A. O. (1998) Identification of a Steroidogenic Neurohormone in Female Mosquitoes, *Journal of Biological Chemistry 273*, 3967-3971.
[35] Roy, S., Saha, T. T., Johnson, L., Zhao, B., Ha, J., White, K. P., Girke, T., Zou, Z., and Raikhel, A. S. (2015) Regulation of Gene Expression Patterns in Mosquito Reproduction, *PLOS Genetics 11*, e1005450.
[36] Roy, S. G., Hansen, I. A., and Raikhel, A. S. (2007) Effect of insulin and 20-hydroxyecdysone in the fat body of the yellow fever mosquito, Aedes aegypti, *Insect biochemistry and molecular biology 37*, 1317-1326.

[37] Hansen, I., Attardo, G., Rodriguez, S., and Drake, L. (2014) Four-way regulation of mosquito yolk protein precursor genes by juvenile hormone-, ecdysone-, nutrient-, and insulin-like peptide signaling pathways, *Frontiers in Physiology 5*.
[38] Pondeville, E., Maria, A., Jacques, J.-C., Bourgouin, C., and Dauphin-Villemant, C. (2008) Anopheles gambiae males produce and transfer the vitellogenic steroid hormone 20-hydroxyecdysone to females during mating, *PNAS 105*, 19631.
[39] Telang, A., Li, Y., Noriega, F. G., and Brown, M. R. (2006) Effects of larval nutrition on the endocrinology of mosquito egg development, *Journal of Experimental Biology 209*, 645.
[40] Raikhel, A. S., and Lea, A. O. (1990) Juvenile hormone controls previtellogenic proliferation of ribosomal RNA in the mosquito fat body, *General and Comparative Endocrinology 77*, 423-434.
[41] Bergot, B. J., Jamieson, G. C., Ratcliff, M. A., and Schooley, D. A. (1980) JH Zero: New Naturally Occurring Insect Juvenile Hormone from Developing Embryos of the Tobacco Hornworm, *Science 210*, 336.
[42] Briegel, H. (1990) Metabolic relationship between female body size, reserves, and fecundity of Aedes aegypti, *Journal of Insect Physiology 36*, 165-172.
[43] Briegel, H. (1990) Fecundity, metabolism, and body size in Anopheles (Diptera: Culicidae), vectors of malaria, *J Med Entomol 27*, 839-850.
[44] Noriega, F. G., Shah, D. K., and Wells, M. A. (2003) Juvenile hormone controls early trypsin gene transcription in the midgut of Aedes aegypti, *Insect Molecular Biology 6*, 63-66.
[45] Shapiro, J. P., and Hagedorn, H. H. (1982) Juvenile hormone and the development of ovarian responsiveness to a brain hormone in the mosquito, Aedes aegypti, *General and Comparative Endocrinology 46*, 176-183.
[46] Clifton, M. E., and Noriega, F. G. (2012) The fate of follicles after a blood meal is dependent on previtellogenic nutrition and juvenile hormone in Aedes aegypti, *J Insect Physiol 58*, 1879-1611 (Electronic).
[47] Hagedorn, H. H., Turner, S., Hagedorn, E. A., Pontecorvo, D., Greenbaum, P., Pfeiffer, D., Wheelock, G., and Flanagan, T. R. (1977) Postemergence growth of the ovarian follicles of Aedes aegypti, *Journal of Insect Physiology 23*, 203-206.
[48] Noriega, F. G. (2004) Nutritional regulation of JH synthesis: a mechanism to control reproductive maturation in mosquitoes?, *Insect Biochem Mol Biol 34*, 687-693.
[49] Naksathit, A. T., Edman Jd Fau - Scott, T. W., and Scott, T. W. Partitioning of glycogen, lipid, and sugar in ovaries and body remnants of female Aedes aegypti (Diptera: Culicidae) fed human blood.
[50] Price, D. P., Schilkey, F. D., Ulanov, A., and Hansen, I. A. (2015) Small mosquitoes, large implications: crowding and starvation affects gene expression and nutrient accumulation in Aedes aegypti, *Parasites & Vectors 8*, 252.
[51] Naksathit, A. T., Edman, J. D., and Scott, T. W. (1999) Amounts of Glycogen, Lipid, and Sugar in Adult Female Aedes aegypti (Diptera: Culicidae) Fed Sucrose, *Journal of Medical Entomology 36*, 8-12.
[52] Joy, T. K., Arik, A. J., Corby-Harris, V., Johnson, A. A., and Riehle, M. A. (2010) The impact of larval and adult dietary restriction on lifespan, reproduction and growth in the mosquito Aedes aegypti, *Experimental Gerontology 45*, 685-690.

[53] Liu, P., Fu, X., and Zhu, J. A.-O. X. (2018) Juvenile hormone-regulated alternative splicing of the taiman gene primes the ecdysteroid response in adult mosquitoes, *PNAS 115*, E7738-E7747.
[54] Wilson, T. G. F., J. (1986) A Drosophila melanogaster mutant resistant to a chemical analog of juvenile hormone, *Dev Biol 118*, 190-201.
[55] Li, M., Liu, P., Wiley, J. D., Ojani, R., Bevan, D. R., Li, J., and Zhu, J. (2014) A steroid receptor coactivator acts as the DNA-binding partner of the methoprene-tolerant protein in regulating juvenile hormone response genes, *Mol Cell Endocrinol 394*, 47-58.
[56] Li, J., Mead, E., and Zhu, J. (2011) Heterodimer of two bHLH-PAS proteins mediates juvenile hormone-induced gene expression, *Proceedings of the National Academy of Sciences of the United States of America 108*, 638-643.
[57] Shin, S. W., Zou Z Fau - Saha, T. T., Saha Tt Fau - Raikhel, A. S., and Raikhel, A. S. (2012) bHLH-PAS heterodimer of methoprene-tolerant and Cycle mediates circadian expression of juvenile hormone-induced mosquito genes, *PNAS 109*.
[58] Zou, Z., Saha, T. T., Roy, S., Shin, S. W., Backman, T. W., Girke, T., White, K. P., and Raikhel, A. S. (2013) Juvenile hormone and its receptor, methoprene-tolerant, control the dynamics of mosquito gene expression, *Proc Natl Acad Sci U S A 110*, E2173-2181.
[59] Jones, S. (2004) An overview of the basic helix-loop-helix proteins, *Genome biology 5*, 226-226.
[60] Ashok, M., Turner, C., and Wilson, T. G. (1998) Insect juvenile hormone resistance gene homology with the bHLH-PAS family of transcriptional regulators, *Proceedings of the National Academy of Sciences 95*, 2761.
[61] Sailsbery, J. K., Atchley, W. R., and Dean, R. A. (2011) Phylogenetic Analysis and Classification of the Fungal bHLH Domain, *Molecular Biology and Evolution 29*, 1301-1318.
[62] Charles, J.-P., Iwema, T., Epa, V. C., Takaki, K., Rynes, J., and Jindra, M. (2011) Ligand-binding properties of a juvenile hormone receptor, Methoprene-tolerant, *Proceedings of the National Academy of Sciences 108*, 21128.
[63] Zhang, Z., Xu, J., Sheng, Z., Sui, Y., and Palli, S. R. (2011) Steroid receptor co-activator is required for juvenile hormone signal transduction through a bHLH-PAS transcription factor, methoprene tolerant, *J Biol Chem 286*, 8437-8447.
[64] Dubrovsky, E. B. (2005) Hormonal cross talk in insect development, *Trends in Endocrinology & Metabolism 16*, 6-11.
[65] Saha, T. T., Roy, S., Pei, G., Dou, W., Zou, Z., and Raikhel, A. S. (2019) Synergistic action of the transcription factors Krüppel homolog 1 and Hairy in juvenile hormone/Methoprene-tolerant-mediated gene-repression in the mosquito Aedes aegypti, *PLOS Genetics 15*, e1008443.
[66] Ojani, R., Fu, X., Ahmed, T., Liu, P., and Zhu, J. (2018) Kruppel homologue 1 acts as a repressor and an activator in the transcriptional response to juvenile hormone in adult mosquitoes, *Insect Mol Biol 27*, 268-278.
[67] Zhu, J., Busche, J. M., and Zhang, X. (2010) Identification of juvenile hormone target genes in the adult female mosquitoes, *Insect Biochem Mol Biol 40*, 23-29.
[68] Lardelli, M., and Ish-Horowicz, D. (1993) Drosophila hairy pair-rule gene regulates embryonic patterning outside its apparent stripe domains, *Development 118*, 255.

[69] Orenic, T. V., Held, L. I., Paddock, S. W., and Carroll, S. B. (1993) The spatial organization of epidermal structures: hairy establishes the geometrical pattern of Drosophila leg bristles by delimiting the domains of achaete expression, *Development 118*, 9.
[70] Rushlow, C. A., Hogan A Fau - Pinchin, S. M., Pinchin Sm Fau - Howe, K. M., Howe Km Fau - Lardelli, M., Lardelli M Fau - Ish-Horowicz, D., and Ish-Horowicz, D. (1989) The Drosophila hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc, *EMBO J 8*, 3095-3103.
[71] Belles, X., and Santos, C. G. (2014) The MEKRE93 (Methoprene tolerant-Kruppel homolog 1-E93) pathway in the regulation of insect metamorphosis, and the homology of the pupal stage, *Insect Biochem Mol Biol 52*.
[72] Cruz, J., Sieglaff, D. H., Arensburger, P., Atkinson, P. W., and Raikhel, A. S. (2009) Nuclear receptors in the mosquito Aedes aegypti, *The FEBS Journal 276*, 1233-1254.
[73] Coons, L. (1999) Effect of Methoprene and 20-Hydroxyecdysone on Vitellogenin Production in Cultured Fat Bodies and Backless Explants from Unfed Female Dermacentor Variabilis., *Journal of Insect Physiology 45*, 755-761.
[74] Rivera-Perez, C., Nouzova, M., and Noriega, F. G. (2012) A Quantitative Assay for the Juvenile Hormones and Their Precursors Using Fluorescent Tags, *PLOS ONE 7*, e43784.
[75] Nene, V. e. a. (2007) Genome sequence of Aedes aegypti, a major arbovirus vector, *Science 316*, 1718-1723.
[76] Timoshevskiy, V. A., Severson Dw Fau - Debruyn, B. S., Debruyn Bs Fau - Black, W. C., Black Wc Fau - Sharakhov, I. V., Sharakhov Iv Fau - Sharakhova, M. V., and Sharakhova, M. V. (2013) An integrated linkage, chromosome, and genome map for the yellow fever mosquito Aedes aegypti, *PLoS Negl Trop Dis 7*, e2052.
[77] Waterhouse, R. M., Seppey, M., Simao, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V., and Zdobnov, E. M. (2018) BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics, *Mol Biol Evol 35*, 543-548.
[78] Matthews, B. J. e. a. (2018) Improved reference genome of Aedes aegypti informs arbovirus vector control, *Nature 563*, 501-507.
[79] Belton, J. M., McCord Rp Fau - Gibcus, J. H., Gibcus Jh Fau - Naumova, N., Naumova N Fau - Zhan, Y., Zhan Y Fau - Dekker, J., and Dekker, J. (2012) Hi-C: a comprehensive technique to capture the conformation of genomes, *Methods 58*, 268-276.
[80] Buenrostro, J. D., Wu, B., Chang, H. Y., and Greenleaf, W. J. (2015) ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide, *Current protocols in molecular biology 109*, 21.29.21-21.29.29.
[81] Baruzzo, G., Hayer, K. E., Kim, E. J., Di Camillo, B., FitzGerald, G. A., and Grant, G. R. (2017) Simulation-based comprehensive benchmarking of RNA-seq aligners, *Nature Methods 14*, 135-139.
[82] Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown, *Nat Protoc 11*, 1650-1667.

[83] Wang, Z., Gerstein M Fau - Snyder, M., and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics, *Nat Rev Genet 10*, 57-63.
[84] Li, H. e. a. (2009) The Sequence Alignment/Map format and SAMtools, *Bioinformatics 25*, 2078-2079.
[85] Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads, *Nat Biotechnol 33*, 290-295.
[86] Merchant, N., Lyons, E., Goff, S., Vaughn, M., Ware, D., Micklos, D., and Antin, P. (2016) The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences, *PLOS Biology 14*, e1002342.
[87] Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2012) STAR: ultrafast universal RNA-seq aligner, *Bioinformatics 29*, 15-21.
[88] Biology, J. H. U. C. f. C. (2018) StringTie-Transcript assembly and quantification for RNA-Seq.
[89] Turner, S. (2012) Bioconductor Support, Bioinformatics Core, University of Virginia School of Medicine.
[90] Letunic, I., and Bork, P. (2017) 20 years of the SMART protein domain annotation resource, *Nucleic Acids Research 46*, D493-D496.
[91] Johnson, L. S., Eddy, S. R., and Portugaly, E. (2010) Hidden Markov model speed heuristic and iterative HMM search procedure, *BMC Bioinformatics 11*, 431.
[92] Chung, H.-N. e. a. (2017) Fat Body Organ Culture System in Aedes Aegypti, a Vector of Zika Virus, *JoVE*, e55508.
[93] Benayoun, B. A., Caburet, S., and Veitia, R. A. (2011) Forkhead transcription factors: key players in health and disease, *Trends in Genetics 27*, 224-232.
[94] Weigel, D., Jürgens, G., Küttner, F., Seifert, E., and Jäckle, H. (1989) The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo, *Cell 57*, 645-658.
[95] Weigel, D. (1990) The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? [letter], *Cell 63*, 455-456.
[96] Hannenhalli, S., and Kaestner, K. H. (2009) The evolution of Fox genes and their role in development and disease, *Nat Rev Genet 10*, 233-240.
[97] Hansen, I. A., Sieglaff, D. H., Munro, J. B., Shiao, S.-H., Cruz, J., Lee, I. W., Heraty, J. M., and Raikhel, A. S. (2007) Forkhead transcription factors regulate mosquito reproduction, *Insect biochemistry and molecular biology 37*, 985-997.
[98] Clark, K. L., Halay Ed Fau - Lai, E., Lai E Fau - Burley, S. K., and Burley, S. K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5, *Nature 364*, 412-420.
[99] Obsil, T., and Obsilova, V. (2008) Structure/function relationships underlying regulation of FOXO transcription factors, *Oncogene 27*, 2263-2275.
[100] Caburet, S., Vaiman, D., and Veitia, R. A. (2004) A Genomic Basis for the Evolution of Vertebrate Transcription Factors Containing Amino Acid Runs, *Genetics 167*, 1813.
[101] Seoane, J., Le, H.-V., Shen, L., Anderson, S. A., and Massagué, J. (2004) Integration of Smad and Forkhead Pathways in the Control of Neuroepithelial and Glioblastoma Cell Proliferation, *Cell 117*, 211-223.

[102] Madureira, P. A., Varshochi, R., Constantinidou, D., Francis, R. E., Coombes, R. C., Yao, K.-M., and Lam, E. W. F. (2006) The Forkhead Box M1 Protein Regulates the Transcription of the Estrogen Receptor α in Breast Cancer Cells, *Journal of Biological Chemistry 281*, 25167-25176.
[103] Blount, A. L., Schmidt, K., Justice, N. J., Vale, W. W., Fischer, W. H., and Bilezikjian, L. M. (2009) FoxL2 and Smad3 Coordinately Regulate Follistatin Gene Transcription, *Journal of Biological Chemistry 284*, 7631-7645.
[104] van der Vos, K. E., and Coffer, P. J. (2008) FOXO-binding partners: it takes two to tango, *Oncogene 27*, 2289-2299.
[105] Dansen, T. B., Smits, L. M. M., van Triest, M. H., de Keizer, P. L. J., van Leenen, D., Koerkamp, M. G., Szypowska, A., Meppelink, A., Brenkman, A. B., Yodoi, J., Holstege, F. C. P., and Burgering, B. M. T. (2009) Redox-sensitive cysteines bridge p300/CBP-mediated acetylation and FoxO4 activity, *Nature Chemical Biology 5*, 664-672.
[106] Kitamura, Y. I., Kitamura, T., Kruse, J.-P., Raum, J. C., Stein, R., Gu, W., and Accili, D. (2005) FoxO1 protects against pancreatic β cell failure through NeuroD and MafA induction, *Cell Metabolism 2*, 153-163.
[107] Carlsson, P., and Mahlapuu, M. (2002) Forkhead Transcription Factors: Key Players in Development and Metabolism, *Developmental Biology 250*, 1-23.
[108] Coffer, P. J., and Burgering, B. M. T. (2004) Forkhead-box transcription factors and their role in the immune system, *Nature Reviews Immunology 4*, 889-899.
[109] Jonsson, H., and Peng, S. L. (2005) Forkhead transcription factors in immunology, *Cellular and Molecular Life Sciences CMLS 62*, 397-409.
[110] Kaufmann, E., and Knöchel, W. (1996) Five years on the wings of fork head, *Mechanisms of Development 57*, 3-20.
[111] Lehmann, O. J., Sowden, J. C., Carlsson, P., Jordan, T., and Bhattacharya, S. S. (2003) Fox's in development and disease, *Trends in Genetics 19*, 339-344.
[112] Le lay, J., and Kaestner, K. H. (2010) The Fox Genes in the Liver: From Organogenesis to Functional Integration, *Physiological Reviews 90*, 1-22.
[113] Hwangbo, D. S., Gersham, B., Tu, M.-P., Palmer, M., and Tatar, M. (2004) Drosophila dFOXO controls lifespan and regulates insulin signalling in brain and fat body, *Nature 429*, 562-566.
[114] Puig, O., and Tjian, R. (2005) Transcriptional feedback control of insulin receptor by dFOXO/FOXO1, *Genes & Development 19*, 2435-2446.
[115] Kokoza, V. A., Martin, D., Mienaltowski, M. J., Ahmed, A., Morton, C. M., and Raikhel, A. S. (2001) Transcriptional regulation of the mosquito vitellogenin gene via a blood meal-triggered cascade, *Gene 274*, 47-65.
[116] Belikov, S., Åstrand, C., and Wrange, Ö. (2009) FoxA1 Binding Directs Chromatin Structure and the Functional Response of a Glucocorticoid Receptor-Regulated Promoter, *Molecular and Cellular Biology 29*, 5413.
[117] Akbari, O. S., Antoshechkin, I., Amrhein, H., Williams, B., Diloreto, R., Sandler, J., and Hay, B. A. (2013) The developmental transcriptome of the mosquito Aedes aegypti, an invasive species and major arbovirus vector, *G3 (Bethesda, Md.) 3*, 1493-1509.