CHARACTERIZATION OF THE ONTHOPHAGUS NIGRIVENTRIS INSULIN RECEPTOR

By

LAURA LISA HAHN

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN ENTOMOLOGY

WASHINGTON STATE UNIVERSITY
Department of Entomology

MAY 2010
To the Faculty of Washington State University:
The members of the Committee appointed to examine the thesis of LAURA L. HAHN find it satisfactory and recommend that it be accepted.

____________________________
Laura S. Lavine, Ph.D., Chair

____________________________
Stephen F. Garczynski, Ph.D.

____________________________
Buel D. Rodgers, Ph.D.
Acknowledgement

I would like to first of all thank my advisor, Dr. Laura Lavine, for providing the opportunity and assistance needed to obtain my degree. I would like to thank Dr. Stephen Garczynski for allowing me to work in his lab and for the time spent helping me troubleshoot my project. Last, I would like to thank Dr. Dan Rodgers for advice given about my project, difficult decisions, and the future.
CHARACTERIZATION OF THE ONTHOPHAGUS NIGRIVENTRIS INSULIN RECEPTOR

Abstract

by Laura Lisa Hahn, M.S.
Washington State University
May 2010

Chair: Laura S. Lavine

Coordinating body size and trait size is an important aspect of animal development. The body size of an individual is determined by genetic and environmental factors. Nutrition is the most important environmental cue. Organisms need a way to communicate their nutritional condition to each cell in their body in order to attain the correct body size, and thus trait size. The insulin signaling pathway is one mechanism for accomplishing this. Beetle horn size is a trait that is highly dependent on larval access to nutrition. There is a link between overall body size and amount of horn growth. The insulin signaling pathway has been suggested as a mechanism for coordinating horn size with body size in scarab beetles. As a first step to determine the extent to which the insulin signaling pathway plays a role in beetle horn growth, the insulin receptor transcript was cloned from a horned scarab beetle, Onthophagus nigriventris. In addition, the expression of this transcript during the late prepupal period was determined in different imaginal tissues and in horn imaginal tissues specifically. The full length O. nigriventris insulin receptor transcript is highly conserved in sequence when compared to the nucleotide and amino acid sequences of insulin receptors of other species. It was found in all imaginal tissues tested
including horn, wing, leg, and genital discs of males and horn, wing and genital discs of females. Relative expression levels did not differ between the horn tissues of small male beetles and large female beetles. These results are an important contribution and first step in understanding of the role the insulin receptor, and the insulin signaling pathway, in allometric growth and horn development in scarab beetles and in the evolution of beetle horns in general.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Horned beetles</td>
<td>3</td>
</tr>
<tr>
<td>Insect and beetle horn development</td>
<td>4</td>
</tr>
<tr>
<td>The insulin signaling pathway</td>
<td>7</td>
</tr>
<tr>
<td>History</td>
<td>7</td>
</tr>
<tr>
<td>Insulin, insulin-like growth factors and receptors in mammals</td>
<td>9</td>
</tr>
<tr>
<td>Insulin-related peptides in insects</td>
<td>11</td>
</tr>
<tr>
<td>The insulin receptor in invertebrate animals</td>
<td>14</td>
</tr>
<tr>
<td>Components of the insulin/IGF signaling pathway</td>
<td>15</td>
</tr>
<tr>
<td>Growth</td>
<td>19</td>
</tr>
<tr>
<td>Tissues and hormones involved in growth control</td>
<td>24</td>
</tr>
<tr>
<td>Lifespan</td>
<td>28</td>
</tr>
<tr>
<td>Reproduction</td>
<td>30</td>
</tr>
<tr>
<td>Metabolism</td>
<td>32</td>
</tr>
<tr>
<td>2. MATERIALS AND METHODS AND RESULTS</td>
<td>34</td>
</tr>
</tbody>
</table>
Beetle rearing........................................................................................................34
RNA extraction and cDNA synthesis for cloning and sequencing.......................35
Cloning of the *O. nigriventris* insulin receptor transcript................................35
Annotating the sequence.....................................................................................37
Phylogenetic analysis of OnInR..........................................................................38
RNA extraction and cDNA synthesis for RT-PCR and quantitative PCR ..........39
RT-PCR of imaginal disc tissues........................................................................40
Quantitative PCR..............................................................................................40
Statistical analysis..............................................................................................41
*O. nigriventris* insulin receptor transcript..........................................................42
N-terminal region..............................................................................................42
The processing site and fibronectin type III domains.........................................43
Transmembrane and juxtamembrane domains..................................................47
Tyrosine kinase domain......................................................................................48
Evolutionary relationship of OnInR to other species..........................................48
Distribution of OnInR transcript in imaginal discs..............................................54
Relative expression of OnInR mRNA in horn tissues........................................54
3. DISCUSSION....................................................................................................58
Structure of the *O. nigriventris* insulin receptor..............................................58
Expression of OnInR..........................................................................................62
Summary............................................................................................................64
APPENDIX

A. Insulin receptor protein alignment of the L1 and L2 domains and the cysteine-rich region............ 83
B. Insulin receptor protein alignment of the Fibronectin type III domains ..................................... 86
C. Insulin receptor protein alignment of the transmembrane, juxtamembrane, and tyrosine kinase domains ........................................................................................................................................... 89
D. Expression of the OnInR mRNA relative to 28S in the horn tissues of small males and large females ........................................................................................................................................... 92
LIST OF TABLES

1. Degenerate and gene specific primers used to clone the OnInR .......................... 37
2. Insulin receptors used in protein alignment .................................................................. 38
3. Insulin receptors used in phylogenetic analysis ............................................................. 39
4. RT-PCR primers ........................................................................................................... 40
5. Gene specific primers for qPCR .................................................................................... 41
LIST OF FIGURES

1. Diagram of the insulin receptor homodimer .............................................................. 10
2. The components of the insulin signaling pathway ...................................................... 18
3. OnInR α subunit ............................................................................................................. 44
4. Fibronectin type III domains and processing site ...................................................... 46
5. OnInR β subunit ............................................................................................................. 50
6. Phylogenetic tree of the known vertebrate and invertebrate IRs ................................ 52
7. Comparative IR phylogenetic tree .............................................................................. 53
8. Distribution of OnInR in late prepupal stage imaginal discs ...................................... 54
9. Single primer controls for RT-PCR of imaginal discs .............................................. 56
10. Expression of OnInR mRNA in the horn tissue ........................................................ 57
DEDICATION

This thesis is dedicated to my mother and father who provided both emotional and financial support and to my husband who was there through the best and the worst of it.
CHAPTER 1
INTRODUCTION

How growth, body size, and trait size are determined and coordinated are recurring themes in studies of animal development. Great strides have been made toward elucidating the answers to these questions, particularly in insect model systems such as *Drosophila melanogaster*. Genetic and environmental factors determine the body size of an individual. Animals live in changing environments and have generated plastic responses that allow for modifying the size of an individual according to environmental cues. Among the most important of these environmental cues is nutrition which influences two parameters of final adult size in many metazoans: (1) growth rate, and (2) duration of the growth period. Although the physiology of growth control in insects is different from that of mammals, the signaling pathways involved are extraordinarily conserved. Insulin/insulin-like growth factor signaling controls factors that affect growth rate (cell growth, nutrient use, cell size and body size) in insects and mammals while steroid hormones regulate transitions between life stages.

Allometry is defined as either the relationship between the size of one trait versus the size of another trait or trait size versus body size (Shingleton *et al*., 2008). Some traits vary in direct proportion to body size. This is called static allometry. However, the expression of some morphological traits is regulated by thresholds of sensitivity to the environment (Emlen and Nijhout, 2001). Threshold traits—polyphenisms—occur as one of several different forms expressed in response to predictable environmental conditions. Allometric growth is controlled by physiological signaling mechanisms that directly respond to nutrition during development from the whole body to the cell (Shingleton *et al*., 2008). Threshold traits provide unique
opportunities for studying the variation in physiological processes that regulate the development of static and plastic morphological traits (Emlen and Nijhout, 2001).

Beetle horns provide an ideal system for studying threshold traits. Beetle horn growth is highly dependent on access to nutrition during larval development. Both horn size and body size are susceptible to changes in nutrition, and this sensitivity results in a link between overall body size and amount of horn growth (Emlen et al., 2006). The dung beetle, *Onthophagus nigriventris*, is an especially relevant model to study the genetic control of polyphenism because males exhibit dimorphism in horn growth dependent upon body size (male dimorphism) and females never produce horns regardless of their body size (sexual dimorphism) (Emlen et al., 2006). Male *O. nigriventris* only produce horns if they reach a critical threshold weight during larval development, but once horns are produced their growth is correlated with body size (Emlen et al., 2006). Both phylogenetic and developmental evidence to date suggests that horn dimorphism involves one or several secondary mechanisms that arose after the evolution of horns (Emlen et al., 2006). The hypothesis is that once horns evolved, physiological mechanisms to shut off horn growth evolved secondarily (Emlen et al., 2006). Arguably the most important nutrition signaling pathway in animals is the insulin signaling pathway (Edgar 2006; Emlen & Allen 2004). Thus, the insulin signaling pathway is a candidate physiological signaling network for the regulation of phenotypic plasticity in beetle horn dimorphisms in both males and females. My central hypothesis is that the insulin signaling pathway, via the insulin receptor, functions in horn development in *O. nigriventris* male dimorphism and sexual dimorphism. My goal for this study was to characterize the *O. nigriventris* insulin receptor through cloning and sequencing the mRNA encoding the receptor, examine its expression pattern in horn, wing, leg, and genital
imaginal disc tissues from male and female prepupae, and to quantify relative levels of transcript in these tissues. The data from my study is a first step in providing new information to be used to test the extent to which the insulin receptor and the insulin signaling pathway is involved in horn development in beetles.

**Horned Beetles**

Beetle horns are rigid outgrowths of the exoskeleton typically used by males in competition with other males for access to females and resources used by females (Emlen 1997; Hunt and Simmons 1997; Iguchi 1998; Moczek & Emlen, 2000; Hongo 2003). There are about 120 different families of beetles, and many of these have isolated representatives that bear some form of cuticular outgrowth (Emlen et al., 2005). However, most horned species are located within the family Scarabaeidae (Emlen et al., 2006). Recent phylogenetic analysis of 48 species from the scarab genus *Onthophagus* showed 25 changes in horn location and extensive variation in horn size and shape (Emlen et al., 2005). While this study is impressive in the total number of species that were compared, it is important to point out that it represents approximately 1.6% of the species in this genus (Emlen et al., 2006). Nevertheless, this variation suggests that there have been four principle axes of horn evolution: (1) physical location of the horn, (2) horn shape, (3) horn allometry—a reflection of the relative size of the horn and of the coupling of horn growth with individual variation in body size, and finally (4) the presence or absence of horn expression and the nature of dimorphism (Emlen et al., 2006). Evidence from other insect models and preliminary evidence from *O. nigriventris* suggests that the insulin signaling pathway is a likely candidate for the mechanisms of allometry and dimorphism in this beetle (Emlen et al., 2006).
Insect and beetle horn development

All holometabolous insects undergo complete metamorphosis during development in which they pass through a genetically specified number of larval instars until they reach a critical size during the final larval instar. At this point, the larvae begin preparations to enter a pupal stage during which they undergo metamorphosis to their adult form. Scarab beetles are holometabolous insects and after three larval instars and a pupal stage, they molt into an adult after which time their morphology is fixed.

Insect appendages form from clusters of cells called imaginal discs, which behave as autonomous units. In many insect species, these cells remain dormant for most of the larval period, and then undergo rapid proliferation after the larvae have ceased feeding during what is known as the gut purge and prepupal periods. Proliferation of imaginal discs determines the final sizes of wings, legs, antennae, mouthparts, compound eyes and genitalia in adults (Emlen and Allen, 2004). Development of insect appendages is remarkably similar to limb development in vertebrates. Both develop from regions of cells which have a self-contained identity. Genetic information directs growth of these regions of cells to specified approximate final sizes (Stern and Emlen, 1999).

Beetle horns arose as novel morphological structures. They have no obvious homolog in other arthropod structures; they exist alongside other insect appendages but are not modified from them (Emlen et al., 2005; Moczek, 2009). Their development shares many similarities with the development of traditional imaginal discs, but horns appear to have arisen as new regions of epidermal tissue that at some point in scarab history began to develop like imaginal discs (Emlen et al., 2006). In species where it has been studied, horns form during the larval period (Emlen et
al., 2006). For most metamorphic insects, the last larval instar is when the final sizes of adult body parts are determined and is likely the period when scaling of body parts to body size occurs (Emlen and Allen, 2004). Horns in *O. nigriventris* delay growth until the prepupal period. During this period larvae begin to purge their guts in preparation for metamorphosis, and these clusters of epidermal cells undergo rapid cell division forming discs of densely folded tissue that will unfurl to their full length in the pupa to form the horns (Emlen and Nijhout, 1999; Moczek and Nagy, 2005; Emlen et al., 2005).

Several factors have been suggested to control the amount of growth in developing horns. First, horn development is highly dependent on larval nutrition (Moczek and Emlen, 1999; Hunt and Simmons, 2000; Karino et al., 2004). Horn size and body size are sensitive to variation in nutrition resulting in a coupling of the amount of horn growth with body size, and this is reflected in tight scaling relationships between horn length and body size (Emlen et al., 2006). Emlen (1994) provided evidence that horn growth was a function of nutrition and not genetics when he performed an experiment on a related species possessing differential horn expression, *Onthophagus acuminatus*, in which he tested environmental factors versus paternal inheritance. By experimentally manipulating larval food quantity he demonstrated that horn length variation in *O. acuminatus* was influenced primarily by environmental factors. In both experimental and control populations he showed that horn lengths of male progeny were a function of individual differences in body size and not of paternal horn lengths.

Second, horn growth in many scarabs is regulated by developmental ‘switch’ mechanisms that prevent or alter patterns of horn growth in subsets of individuals (West-Eberhard, 2003). For example, most females do not produce horns regardless of their nutritional
environment or body size (sexual dimorphism). Also, males that do not attain a threshold body
size often do not produce horns or produce small horns that scale very differently with body size
than those of large males (male dimorphism). Mechanisms of dimorphism appear only to affect
the growth of horns in horned beetle species (Emlen et al., 2006).

Last, the limb-patterning pathway, a network of local gene interactions, acts within the
developing disc to determine the shape and final size of the horn. Studies of *Onthophagus
taurus*, focusing on dimorphic switches between horned and hornless patterns of development,
identified two critical periods during the third larval stage when hormones appeared to influence
the growth of horns (Emlen and Nijhout, 1999; Emlen and Nijhout, 2001; Moczek and Nijhout,
2002). These studies suggest that interactions between ecdysone and juvenile hormone (JH)
control whether the growth of horn cells is permitted or suppressed. The details of these events
have been described in other papers (Emlen and Nijhout, 2001; Moczek and Nijhout 2002;
Emlen et al., 2005). This study focuses on the second critical period of hormone sensitivity. This
is the period of horn growth. Larvae have stopped feeding and all appendages undergo rapid cell
division. It is during this period that trait growth is likely to be modified by response to insulin
and other circulating signals (for example, in *Drosophila*: Mirth et al., 2005; Shingleton et al.,
2005). This is also the period when shape and size of the horn is determined by the limb-
patterning pathway (Moczek and Nagy, 2005). During this time, these pathways may be altered
in the horn discs of small males and females to repress horn growth (Emlen et al., 2006). A
detailed description of the limb-patterning pathway can be found elsewhere (Emlen et al., 2006;
Moczek and Nagy; 2005). There is evidence that insulin signaling affects some elements of the
limb-patterning pathway (Chen et al., 1996). Other studies of insulin receptor mutants suggest an interaction between insulin, JH, and ecdysone (Tu et al., 2002; Tu and Tatar, 2003).

As I have briefly described above, the insulin signaling pathway is an important pathway in development and a good candidate for regulation of dimorphism in horned beetles. In the following section, I provide an extensive review on what is known about this pathway in insects to this point, beginning with its role in growth and body size control, important signaling tissues, and interactions with other hormones. I also discuss its effects on lifespan, reproduction, and metabolism.

**The Insulin Signaling Pathway**

Nutrition is a critical factor in an organism’s existence. It is necessary for an organism to be able to assess its nutritional condition and relay this information to every cell in the body. The insulin signaling pathway evolved early in the history of life as a mechanism for accomplishing this task (Skorokhod et al., 1999). First discovered in an attempt to treat diabetes, the roles of insulin and its pathway now include growth, development, metabolic homeostasis, reproduction, and lifespan.

**History**

Type 2 diabetes mellitus—the most common metabolic disorder worldwide—results from the failure of insulin to regulate carbohydrate metabolism. For this reason, insulin is one of the most well studied peptide hormones. At the beginning of the 20th century, investigators attempted to use pancreas extracts to treat diabetes mellitus in test animals which often resulted in infection and abscess formation due to impurities. In 1920, Frederick Banting, a general surgeon and teacher of anatomy and physiology at what is now known as the University of
Western Ontario, approached Dr. John J. R. Macleod at the University of Toronto with the idea of ligating the pancreatic duct to destroy the enzyme-secreting acinar cells before attempting to extract the secretions of the islets of Langerhans (Rosenfeld, 2002; Rendel, 2008). In May, 1921 Banting began working in MacLeod’s laboratory with student assistant Charles H. Best to obtain more pure secretions from the canine pancreas in order to treat diabetes symptoms in dogs whose pancreas had been removed. In December after much difficult work, Banting and Best developed a technique for extracting secretions from bovine pancreas using 95% alcohol and successfully used the extract to lower the blood sugar of a diabetic dog (Rosenfeld, 2002). At that point Canadian biochemist James B. Collip joined the team to develop a method for purification of the extract. The first clinical trial was conducted on January 11, 1922 on a patient suffering from Type 1 diabetes. The treatment was relatively unsuccessful and the patient suffered a severe reaction caused by the impurities in the extract (Rosenfeld, 2002; Rendel, 2008). Collip further purified the extract and subsequent clinical trials were very successful. Until this point, patients suffering from juvenile diabetes had little hope of a long life (Rosenfeld, 2002; Rendel, 2008). In 1923, Banting and MacLeod were awarded the Nobel Prize for the discovery of insulin. John J. Abel of Johns Hopkins University prepared the first crystalline insulin in 1926 and Fredrick Sanger determined the molecular structure of insulin in the mid-1950s and later determined the base sequences of nucleic acids—discoveries for which he was respectively rewarded the Nobel Prize in 1958 and 1980 (Rosenfeld, 2002). In 1982, the first recombinant human insulins were synthesized by Novo and Eli Lilly and Company (Rendel, 2008).
**Insulin, Insulin-like Growth Factors and Receptors in Mammals**

In mammals, the insulin signaling pathway functions to regulate metabolism while growth regulation and differentiation is under the control of the IGF-1 regulatory axis. Some evidence suggests that these pathways may also function in the regulation of lifespan (Chistyakova, 2008). Together the insulin/IGF-1 pathways consist of multiple ligands and receptors. The insulin family in mammals consists of insulin, insulin-like growth factors (IGF) I and II and seven members of the relaxin-like peptide family which includes gene 1 (H1) relaxin, gene 2 (H2) relaxin, gene 3 (H3) relaxin, insulin-like peptides Insl 3, 4, 5, and 6 (Adham et al., 1993; Bathgate et al., 2002; Chassin et al., 1995; Conklin et al., 1999; Hudson et al., 1983, 1984; Lok et al., 2000). In humans, the active form of insulin that circulates throughout the body is a monomer composed of two chains—an A chain of 21 amino acids and a B chain of 30 amino acids that are linked by two disulfide bridges located between A7-B7 and A20-B19. There is also an intra-chain disulfide bridge between A7 and A11 (De Meyts, 2004). Receptors that can bind insulin and the IGFs are the insulin receptor (IR), IGF-1 receptor (IGF-1R), and IGF-2 receptor (IGF2R). IR and IGF1R are synthesized as single chain preproreceptors and then directed by the signal peptide to the endoplasmic reticulum for further processing (De Meyts, 2004). They are ligand activated receptor tyrosine kinases which form homodimers of two identical α/β monomers (Figure 1) or as heterodimers with one IR monomer and one IGF-1R monomer. The IR is alternatively spliced into A and B isoforms that differ in their affinity for insulin. Both forms can bind the IGFs but with a lower affinity than insulin (Taguchi and White, 2008). The B isoform is the dominant form in insulin-responsive tissues such as the adult liver, muscle, and adipose tissue (Taguchi and White, 2008). The A isoform is dominant form in fetal tissues, adult
central nervous tissues, and hematopoietic cells (Taguchi and White, 2008). The IGF-1R binds IGF-2 and insulin but with a lower affinity than IGF-1. No function is known for the hybrid receptors (DeMeyts, 2004). It is clear from this brief summary that mammals have a complex and functionally diverse insulin/IGF signaling network.

**Figure 1. Diagram of the human insulin receptor homodimer.** The leucine-rich repeat regions are denoted and L1 and L2, the cysteine-rich regions as CR, the three fibronectin type III domains as FnIII-1, FnIII-2a/FnIII-2b and FnIII-3 and the insert region of FnIII-2 as ID. The transmembrane, juxtamembrane and tyrosine kinase (TKD) domains are indicated. The α-α and α-β disulfide bonds are indicated. Modified from Mulhern et al., 1998.
Insulin-related peptides in insects

In 1984, the first insect insulin-like peptide (ILP), bombyxin from the brain of the silkmoth *Bombyx mori*, was discovered. At that time the peptide did not display activity in *B. mori*, but it did stimulate the prothoracic glands (PG) of *Samia cynthia* larvae to release ecdysone which regulates molting and metamorphosis in insects. Thirty-eight bombyxins were subsequently identified in *B. mori* (Kondo *et al*., 1996; Yoshida *et al*., 1997, 1998), and ILPs were also identified in *Samia cynthia ricini* (the saturniid moth; Nagata *et al*., 1999) and *Agrius convolvuli* (the hornworm; Iwami *et al*., 1996). All lepidopteran ILP genes except for five pseudogenes in *B. mori* have A and B peptides connected by a C peptide (Kondo *et al*., 1996; Yoshida *et al*., 1997, 1998). The genes coding for bombyxins lack introns unlike those of vertebrate insulins (Kondo *et al*., 1996; Yoshida *et al*., 1997, 1998). In Lepidopterans, the Medial Neurosecretory Cells (MNCs) are the site of synthesis for ILPs which are then stored in and released from either the corpora cardiaca (CC) and the corpora allata (CA) depending on the species. The CC are structurally and functionally analogous to the posterior pituitary of vertebrates (Wu and Brown, 2006).

The second insect ILP to be discovered was isolated from the CC of *Locusta migratoria* (the migratory locust) (Hetru *et al*., 1995; Lagueux *et al*., 1990). In contrast to Lepidopteran ILP genes, the locust insulin-related peptide (LIRP) gene contains introns. This gene is expressed as two transcripts differing in their 5’ untranslated region (Kromer-Metzger and Lagueux, 1994). One is expressed in the MNCs of the *pars intercerebralis* while the other is expressed in nearly all tissues (Kromer-Metzger and Lagueux, 1994). Other orthopteran species have exhibited
insulin immunoreactivity although no further orthopteran ILPs have been isolated (reviewed in Wu and Brown, 2006).

Seven ILPs were indentified in *D. melanogaster* (DILP 1-7) through the genome database of the fly (Brogiolo *et al.*, 2001; Vanden Broeck, 2001). The first 5 of these genes are located on chromosome 3 while DILP6 & 7 are located on the X chromosome (Brogiolo *et al.*, 2001; Vanden Broeck, 2001). Like LIRP, these genes all have introns (Brogiolo *et al.*, 2001; Vanden Broeck, 2001). Transcripts for the DILPs are located in the brain MNCs (DILP1, 2, 3, 5), the imaginal discs and salivary glands (DILP2), the midgut (DILP 4, 5, and 6), and the ventral nerve cord (DILP7) of larva (Brogiolo *et al.*, 2001; Broughton *et al.*, 2005). Transcripts of DILP5 were detected in the follicle cells of the ovaries of adult female flies (Ikeya *et al.*, 2002). The MNCs are the primary source of ILPs in larvae and females (Rulifson *et al.*, 2002).

Seven ILP genes, *AgamILP1*-7, were identified from the genome of the African malaria mosquito, *Anopheles gambiae*. The A and B peptides predicted from the sequences were remarkably similar to the DILPs (Krieger *et al.*, 2004; Riehle *et al.*, 2002). The distribution of transcripts for these genes suggested growth factor and neurohormonal functions (Krieger *et al.*, 2004). Eight genes encoding ILPs in the yellow fever mosquito, *Aedes aegypti*, have been discovered (Riehle *et al.*, 2006). All eight ILPs share the conserved features of the insulin superfamily prepropeptides. Transcripts for five of the ILPs were expressed mainly in the heads of larval, pupal, and adult mosquitoes (Riehle *et al.*, 2006). Transcripts of two other genes were present in the head, thorax and abdomens of all stages (Riehle *et al.*, 2006). The final ILP was predominantly expressed in abdomen (Riehle *et al.*, 2006).
There has been a small amount of evidence presented for the presence of ILPs in species from Hemiptera, Coleoptera, and Hymenoptera (Wu and Brown, 2006). Two insulin receptors and two ILPs were predicted using the honey bee genome and their expression during caste determination has been described (Wheeler et al., 2006; de Azevedo and Hartfelder, 2008). A *B. mori* ILP was used to stimulate the prothoracic glands of the last instar of *Rhodnius prolixus*, a hemipteran, to secrete ecdysteroids in vitro (Steel and Vafopoulou, 1997). ILPs were detected in extracts of larval heads and midguts of the beetle, *Tenebrio molitor*, and in the brain, CA, and subesophageal ganglion of other life stages (Sevala et al., 1993). An ILP that altered carbohydrate metabolism in the larval fat body in vitro was purified from the midguts of this beetle (Teller et al., 1983). Four genes encoding ILPs were identified in the *Tribolium* genome (Li et al., 2008).

The presence of ILP genes has also been reported in other invertebrates. There are 38 ILPs in the *C. elegans* genome and some of these have been shows to affect insulin signaling in the worm (Panowski and Dillin, 2009). Insulin-like gene 1 (*INS-1*) is the *C. elegans* gene most closely related to human insulin in sequence and structure (Panowski and Dillin, 2009). However, *INS-1*, unlike human insulin, seems to be an antagonist of DAF-2, the *C. elegans* homolog of the human insulin receptor (Panowski and Dillin, 2009). Other ILPs appear to function as DAF-2 agonists (Panowski and Dillin, 2009). ILPs have also been identified in molluscs (Smit et al., 1998; Floyd et al., 1999; Hamano et al., 2005; Moroz et al., 2006), and the sea anemone (Putnam et al., 2007).
The insulin receptor in invertebrate animals

The *Drosophila* insulin receptor (DIR) is similar to mammalian insulin receptors, except that both the NH2 terminus and the COOH terminus include approximately 300 additional amino acid residues (Fernandez *et al*., 1995). The presence of the additional C-terminus region leads to the formation of a β-subunit with an extension which can undergo cell specific proteolytic cleavage to form a smaller β-subunit (Fernandez *et al*., 1995). Mammalian insulin binds the DIR and stimulates tyrosine phosphorylation of both types of β-subunits, although with slightly decreased affinity (Fernandez *et al*., 1995). However the DIR signaling capacity is similar to that of the human insulin receptor which suggests that the basic functions of the receptor are conserved from insects to mammals (Fernandez *et al*., 1995; Yamaguchi *et al*., 1995). The DIR tail contains three tyrosine phosphorylation sites (YXXM) which align with tyrosine phosphorylation sites on mammalian IRS-1, which allows DIR to bind directly to PI3K and activate it without associating with the *Drosophila* insulin receptor substrate (CHICO) (Yenush *et al*., 1995). When the C-terminal extension is cleaved, DIR must associate with CHICO in order to activate the phosphatidylinositol 3-kinase (PI3K) pathway (Fernandez *et al*., 1995). This association is required to produce a full insulin response (Yenush *et al*., 1995).

The mosquito insulin receptor (MIR) was isolated from ovarian mRNA of *A. aegypti* (Graf *et al*., 1997). Like its vertebrate counterpart in consists of an α-subunit with a ligand binding domain, a β-subunit with a tyrosine kinase domain, and a putative cleavage signal for post-translational processing (Graf *et al*., 1997). The mature MIR is a 400-kDa tetrameric protein localized in the membranes of follicle cells surrounding the oocyte and nurse cells (Riehle and Brown, 2002). Levels of transcript and protein increased in ovaries during the first 24 hours after
a blood meal but then disappeared (Riehle and Brown, 2002). The MIR is not present in eggs, larvae, pupae or pharate adults (Riehle and Brown, 2002).

In the nematode *C. elegans*, DAF-2 is the single identified ortholog of the mammalian IR and IGF-1 (Kimura *et al.*, 1997). DAF-2 also has a C-terminal tail with several putative tyrosine phosphorylation sites that can directly bind to PI3K (Morris *et al.*, 1996, Wolkow *et al.*, 2002). An insulin receptor substrate ortholog called IST1 has been identified which is somewhat similar to CHICO and the mammalian insulin receptor substrate (IRS) proteins although it has only one tyrosine phosphorylation site (Wolkow *et al.*, 2002). There is no evidence that IST1 activates PI3K or is required for insulin signaling in *C. elegans* (Wolkow *et al.*, 2002).

**Components of the Insulin/IGF Signaling Pathway**

Organisms need a way to communicate nutritional information to the entire body in order to orchestrate the growth of tissues and organs. This is done by humoral signaling through the insulin/IGF pathway, and one of the key mechanisms through which this pathway affects tissue growth by influencing translation initiation and protein synthesis is via activation of the PI3K pathway (Figure 2).

Insulin binds to the α-subunits of the insulin receptor (IR) and causes a conformational change which induces autophosphorylation of specific tyrosine residues on the β-subunits. This phosphorylation produces a conformational change that activates the intrinsic protein tyrosine kinase activity of the receptor (reviewed in Taguchi and White, 2008). The receptor then phosphorylates tyrosine residues of many target proteins. Although each of these target proteins serves an important function, evidence suggests that insulin responses involved in somatic growth and nutrient homeostasis are mediated through IRS 1 and 2 (reviewed in Taguchi and
In mammals, the insulin signaling cascade interacts with many other pathways. Activation of the IR leads to induction of the GRB2 (growth-factor-receptor-bound protein 2)-SOS (Son of sevenless) complex through IRS1. In this way, Ras is activated and subsequently ERK (extracellular-signal-related kinase) (reviewed in Taguchi and White, 2008). ERK promotes transcriptional activity by directly phosphorylating ELK1 (Ets LiKe gene1) and activating p90<rsk> (p90 ribosomal S6 kinase) which phosphorylates FOS (v-los FBJ murine osteosarcoma viral oncogene homolog) (reviewed in Taguchi and White, 2008).

Another pathway, and one of the best studied, is the PI3K pathway (reviewed in Taguchi and White, 2008). Phosphorylated IRS proteins bind the Src homology 2 (SH2) domain of the p85 regulatory subunit of PI3K, which then catalyzes the formation of phosphatidylinositol-3,4,5-trisphosphate (PIP<sub>3</sub>) from phosphatidylinositol-4,5-diphosphate (PIP<sub>2</sub>) (reviewed in Taguchi and White, 2008). The tumor suppressor phosphatase and tensin homolog (PTEN) is a negative regulator of the insulin signaling pathway. It is a lipid phosphatase that converts PIP<sub>3</sub> back to PIP<sub>2</sub> by removing the 3’ phosphate (reviewed in Taguchi and White, 2008). The PTEN enzyme acts as part of a chemical pathway that signals cells to stop dividing and triggers cells to self-destruct (undergo apoptosis) when necessary. These functions prevent uncontrolled cell growth that can lead to the formation of tumors (reviewed in Taguchi and White, 2008). PIP3 recruits phosphoinositide-dependent kinase (PDK-1, Akt kinase) and serine/threonine kinase protein kinase B (PKB, Akt) to the plasma membrane where PDK-1 phosphorylates and activates Akt. Akt phosphorylates Forkhead Box O 1 (FOXO1) transcription factor. FOXO1, which is involved in the regulation of gene expression, is inactivated by phosphorylation and is sequestered in the cytosol (reviewed in Taguchi and White, 2008). Akt also phosphorylates Tsc 2 of the tuberous
sclerosis complex (TSC) which inhibits Tsc 2 and leads to removal of its inhibition on a small G protein Ras Homolog Enriched in Brain (RHEB). RHEB activates the Target of Rapamycin (TOR) kinase. PDK1 phosphorylates p70S6K, priming it for activation. TOR ultimately phosphorylates and activates p70S6K, which alters gene transcription and translation. TOR also inactivates the 4E-binding protein (4E-BP) that binds and inhibits the eukaryotic initiation factor 4E (eIF4E) which allows eIF4E to participate in ribosome complex formation and translation (reviewed in Taguchi and White, 2008).
Figure 2. The components of the insulin signaling pathway. Functional relationships between components are indicated. Arrows indicate activation and bar-ended lines indicate inhibition. Details of interactions are described in the text. Modified from Taguchi and White (2008).
Growth

Studies of the insulin signaling pathway in *Drosophila* led to the discovery of the pathway as an important component of growth regulation. Ablation of the IPCs or deletion of *DILP*$_1$-5 results in developmental delays and reduced growth, with cells in the wings of these flies exhibiting reductions in size and number (Rulifson *et al.*, 2002; Zhang *et al.*, 2009). Overexpression of DILPs results in bigger flies owing to a DIR-dependent increase in cell size and cell number of individual organs (Brogiolo *et al.*, 2001). Loss of function mutations in the *inr* gene lead to recessive phenotypes that are embryonic or early larval lethal (Fernandez *et al.*, 1995; Chen *et al.*, 1996). Some mutant alleles exhibit heteroallelic complementation and produce viable adults that are developmentally delayed and growth deficient due to a reduction in cell number (Chen *et al.*, 1996). The DIR appears to regulate cell proliferation during development and may be required in the embryonic epidermis and nervous system because formation of the cuticle and the peripheral and central nervous systems are affected by IR mutation (Fernandez *et al.*, 1995; Chen *et al.*, 1996).

An important target of the *Drosophila* insulin signaling pathway is chico, the homologue of the mammalian insulin receptor substrate proteins (IRS). Flies lacking both copies of this gene develop to only half the size of normal flies (Bohni *et al.*, 1999). This lack of body size results from a decrease in both cell size and cell number (Bohni *et al.*, 1999). Recently, evidence has been provided that another adaptor protein may act in parallel to chico (Werz *et al.*, 2009). Lnk is the single SH2B adaptor protein encoded in the *Drosophila* genome and shares a common domain structure with the mammalian SH2B family of adaptor proteins (Werz *et al.*, 2009). Mutations in the gene encoding this protein cause phenotypes similar to those caused by
mutations in chico, including developmental delay and female sterility (Werz et al., 2009). The growth deficiency phenotypes of chico and lnk mutants resembles that of DIR mutants.

Although, the IR can activate the Ras/MAPK and PI3K signaling pathways—two major branches that effect growth, proliferation and organismal size during development—through interaction with different IRS proteins, evidence from early studies in Drosophila strongly suggests that it is the PI3K pathway that controls growth and proliferation in this organism (Oldham et al., 2002). Drosophila possess one class I_A PI3K, dp110, bound to a phosphotyrosine-binding Src homology 2 (SH2) domain containing subunit, p60 (Weinkove et al., 1997; Leevers 2001). Overexpression of wild-type or dominant-negative dp110 in either developing wing or eye imaginal discs results in larger or smaller organs respectively due to changes in cell size and cell number. However, while adult wings are patterned normally, ommatidial pattern is slightly irregular (Weinkove et al., 1999).

The Drosophila ortholog of mammalian PTEN is dPTEN. Mutations of dPTEN in the Drosophila eye result in ommatidia that are larger and more abundant than wild-type ommatidia (Huang et al., 1999; Gao et al., 2000) Overexpression of dPTEN inhibits cell cycle progression early in mitosis and triggers apoptosis during eye development which results in eyes that are smaller than wild-type and also rough (Huang et al., 1999). Overexpressing dPTEN in the wing and the whole organism results in an overall reduction in size for either the wing or the organism (Gao et al., 2000). Overexpressing a dominant negative form of dp110 partially reduces this effect (Gao et al., 2000).

Loss of DIR function can be rescued by increasing PIP3 levels (Oldham et al., 2002). As indicated above, PIP3 is the effector of multiple downstream targets of the PI3K pathway. In
Drosophila these targets are the homolog of phosphoinositide-dependent kinase 1 (dPDK1) and Drosophila protein kinase B or dAkt—a growth factor regulated serine/threonine kinase.

Drosophila Akt activity is abolished in flies that are null mutants of dp110 (Radimerski et al., 2002). Consistent with the role of these two kinases in this growth promoting pathway, Drosophila embryos deficient in the dPDK-1 gene exhibit a dramatic increase in apoptotic activity and die (Cho et al., 2001). In addition, co-overexpression of dAKT and dPDK1 in the Drosophila eye led to a size increase in the eye that is cell-autonomous (Rintelen et al., 2001).

Inside the cell there are critical transcriptional regulatory factors that respond to molecular cues induced by ligand/receptor interactions, which are translated into changes in gene expression and mRNA production. One such transcription factor is Drosophila forkhead transcription factor (dFOXO), which is inhibited by insulin signaling typically resulting from high nutrient availability (Junger et al., 2003; Puig et al., 2003; Kramer et al., 2003; Kramer et al., 2008). When nutrients are abundant insulin is secreted and activates the IR which in turn leads to the phosphorylation of dFOXO through dAkt. Phosphorylated dFOXO remains in the cytoplasm of the cell and is unable to activate its nuclear target genes. When nutrients decrease, insulin levels decrease and dFOXO is dephosphorylated and moves into the nucleus where its target genes—genes implicated in growth inhibition—become activated. One such gene is the Drosophila homolog of 4E-BP (d4E-BP) (Puig and Tjian, 2006). Insulin receptor is also a target gene for dFOXO regulation, revealing a novel feedback mechanism in the insulin signaling pathway (Puig et al., 2003).

There is an intricate connection between the insulin signaling pathway and the Target of Rapamycin pathway (TOR) (Wu and Brown, 2006). Drosophila possess one TOR homolog.
dTOR. *Drosophila* mutant for TOR have severe growth defects, reductions in larval body size and imaginal discs, and die during the second instar (Oldham et al., 2000; Zhang et al., 2000). Imaginal disc cells mutant for TOR are smaller than wild-type cells and proliferate more slowly, and the phenotype of partial TOR mutation is the same as that caused by withdrawal of amino acids (Zhang et al., 2000). This indicates that TOR regulates cell growth and proliferation autonomously by sensing nutrient levels.

Mutation of TSC1 and TSC2 in *Drosophila* causes increased growth and cell size while co-overexpression causes a decrease in cell size, cell number, and organ size (Potter et al., 2001; Tapon et al., 2001). It has recently been shown that mutations in Tsc1 result in increased FOXO levels and elevated expression of FOXO-regulated genes, some of which antagonize growth-promoting pathways (Harvey et al., 2008). DAkt regulates growth by directly phosphorylating TSC2 and disrupting the subcellular colocalization of TSC1 and TSC2 (Potter et al., 2002). As in mammals, TSC1 and TSC2 in *Drosophila* form an active complex that suppresses the functioning of dRheb. *Drosophila* Rheb—a highly conserved member of the Ras superfamily of G-proteins that promotes cell growth—functions downstream of TSC 1/2, and its overexpression causes activation of this pathway, leading to increased cell proliferation (Garami et al., 2003; Zhang et al., 2003). Mutation of dRheb results in growth defects similar to phenotypes caused by mutation of chico (Saucedo et al., 2003; Stocker et al., 2003 Zhang et al., 2003).

DTOR may influence translation through *Drosophila* S6K (DS6K) and D4E-BP. TOR is required for growth factor-dependent phosphorylation of S6K (Schmelzle and Hall, 2000). The activity and phosphorylation of DS6K is reduced in larvae that have been deprived of amino acids, are mutant for dTOR, or that have been fed rapamycin (Oldham et al., 2000; Radimerski et
al., 2002). Overexpression of DS6K can partially rescue reduced larval growth resulting from DTOR hypomorphic mutations (Zhang et al., 2000). Though the size of cells in dS6K mutant flies is reduced, the number of cells in these flies is normal. Through response to growth factors and nutrient sensing, S6K provides a mechanism by which individual cells can coordinate their response to growth factors with nutrient availability (Zhang et al., 2003).

*Drosophila* eIF4E (deIF4E) is the homolog of mammalian eIF4E. In *Drosophila* cells, association of deIF4E with d4E-BP is inhibited by insulin stimulated phosphorylation of d4E-BP (Miron et al., 2001). Insulin inhibition is removed by treatment with rapamycin (Miron et al., 2001). Strong loss-of-function mutations in deIF4E reduced larval growth and result in larval death (Lachance et al., 2002). However loss of d4E-BP does not appear to be sufficient to increase growth (Miron et al., 2001).

As the preceding review has shown, loss-of-function mutations, or over-expression, of various components of the insulin signaling pathway in flies give rise to altered organ size due to changes in cell size, cell number, or both (Gao et al., 2000; Junger et al., 2003; Miron et al., 2001; Puig et al., 2003; Verdu et al., 1999). Cell proliferation requires high levels of protein synthesis, and this process is regulated by the insulin pathway in both vertebrates and insects (Weinkove and Leevers, 2000; Ikeya et al., 2002; Johnston and Gallant, 2002; Nijhout and Grunert, 2003). In insects, insulin-like peptides, likely in cooperation with growth factors, bind to the insulin receptor and activate a cascade that controls the activity of protein translation machinery (Weinkove and Leevers, 2000; Claeyss et al., 2002). Organisms adjust their rate of growth depending on nutrient availability (Puig and Tjian, 2006). Based on this evidence it was concluded that the insulin pathway is an ideal candidate for the developmental mechanism of
horn development in phenotypically plastic scarab beetles because levels of both insulin and growth factor signals are sensitive to larval nutrition (Kawamura et al., 1999; Britton et al., 2002; Ikeya et al., 2002; Nijhout and Grunert, 2003; Emlen and Allen 2004).

**Tissues and hormones involved in growth control**

How insects respond to variation in their nutritional environment and the hormones and tissues involved have been recently reviewed by Edgar, 2006 and Hietakangas and Cohen, 2009 so I will briefly summarize these interactions with the addition of important findings since these reviews were published.

Insulin producing cells (IPCs) are the prominent DILP expressing neuroendocrine cells of the brain (Brogiolo et al., 2001; Rulifson et al., 2002). These cells, along with the adipokinetic hormone (AKH)-producing cells (APCs) in the corpora cardiaca, are considered analogous to β and α cells respectively of the pancreas—the site of vertebrate insulin production (Géminard et al., 2009). Changes in nutrient levels are sensed by the insect fat body, an organ that is functionally related to the vertebrate liver and adipose tissue (Colombani et al., 2003). Recent evidence has revealed that amino acid shortage blocks signaling of the TOR pathway in the fat body, which prevents ILP secretion (Géminard et al., 2009). Ex vivo tissue co-culture demonstrated that a humoral link exists between the fat body and the brain which controls ILP secretion in *Drosophila* (Géminard et al., 2009).

Insulin signaling in *Drosophila* is controlled, at least partially, by a nucleostemin family GTPase, NS3, which acts in serotonergic neurons that are in close proximity to the IPCs (Kaplan et al., 2008). Flies mutant for NS3 display delayed development and reduced body size. Delayed growth is associated with reduced peripheral insulin signaling and can be restored by promoting
PI3K/ Akt signaling in target tissues (Kaplan et al., 2008). It is not yet known whether the IPCs or these serotonergic neurons are the target of the factor secreted by the fat body (Géminard et al., 2009). Also, *Drosophila* short neuropeptide F 1 (sNPF1) and sNPF2, homologues of mammalian neuropeptide Y (NPY), regulate expression of DILP transcripts through activation of extracellular signal-related kinases (ERK) in IPCs of larvae and adults (Lee et al., 2008; Lee et al., 2009). Increased signaling through the sNPFR in the MNCs promoted expression of DILPs 1 and 2 and increased body size. Reduction of signaling through this receptor leads to reduced production of ILPs and reduced peripheral AKT activity (Lee et al., 2008). This mechanism appears to be conserved since insulin expression in rat insulinoma cells is under control of NPY and increased signaling through this peptide activates the MAPK/ERK pathway which is needed for insulin expression mediated by this peptide.

In another study it was found that altering the activity of genes of the protein kinase A pathway within the IPCs altered insulin signaling, a situation very much like the transcription activation of the mammalian insulin gene and the IRS by PKA and is downstream transcription factor Creb (Walkiewicz and Stern, 2009). When DILPs are released, *Drosophila* acid-labile subunit (dALS), the fly ortholog of the vertebrate insulin-like growth factor-binding protein (IGFBP) ALS, forms a complex with circulating DILPs. The complex consists of one molecule dALS, one molecule DILP, and one molecule of Imaginal morphogenesis protein-Late 2 (Imp-L2), an immunoglobulin G-family molecule evolutionarily related to IGFBPs (Arquier et al., 2008). These complexes serve to antagonize the growth function of DILPs and also their carbohydrate and fat metabolism function by limiting the availability of DILPs (Arquier et al., 2008). *Drosophila* ALS could be a candidate for the humoral signal released by the fat body.
since it is expressed mainly in the fat body and regulated by amino acid availability (Colombani et al., 2003).

Nutrition has an intense effect on developmental timing in *Drosophila* (Layalle et al., 2008) and it seems that the signaling mechanisms involved in nutrient sensing must interact with the mechanisms that control developmental timing. A key regulator of developmental timing is ecdysone, a steroid hormone secreted by the PG (Nijhout, 2008). Body size is assessed from the relative growth of the PG. A peak of 20-hydroxyecdysone at the end of the last larval instar causes larvae to stop feeding and growing and begin metamorphosis (Nijhout, 2008). The mechanism that controls the timing of this pulse controls developmental timing and final adult body size (Nijhout, 2008) and to date, four mechanisms underlying this are known (Nijhout, 2008). In *Rhodnius*, *Dipetalogaster*, and *Oncopeltus*, feeding causes abdominal stretch reception which prompts ecdysone secretion and therefore growth stops at a constant body size (Nijhout, 2008). In *Manduca*, the terminal growth period between critical size and ecdysone secretion is constant (Nijhout et al., 2006). How much the animal grows during this period is dependent upon nutrition. In *Drosophila*, the length of the terminal growth period is dependent on nutrition which is sensed by the PTG through TOR (Layalle et al., 2008). Low nutrition reduces ecdysone production (Layalle et al., 2008). TOR only affects development during the TGP. In *Onthophagus*, ecdysone is secreted 48 hours after the larva’s food supply is exhausted (Shafiei et al., 2001).

Imaginal discs are also a regulator of critical size in *Drosophila* (Stieper et al., 2008). Damage to, or slow growth of, the imaginal discs impedes metamorphosis by increasing critical
size and extending the TGP. This does not change final adult body size, however (Stieper et al., 2008).

Critical weight is defined as the size at which the PGs produce enough ecdysone to maintain growth and patterning in imaginal tissues in the absence of nutrition (Mirth and Riddiford, 2007). Critical weight depends on insulin dependent growth of the PTG (Colombani et al., 2005; Mirth et al., 2005). Enhancing insulin signaling in the IPCs increases the expression of an ecdysone response gene, \textit{E74B} (Walkiewicz and Stern, 2009). Recent evidence indicates that ecdysone mediates the switch of imaginal discs to a post-critical weight, nutrition-independent mode of development through the derepression of genes repressed by unliganded ecdysone receptor (Mirth et al., 2009). In \textit{Manduca}, the imaginal discs and primordia become pupally committed during the first day of the feeding stage of the final larval instar and require nutrient input (Koyama et al., 2009). Insulin signaling regulates this change in commitment by overriding the effects of JH which prevents changes earlier instars (Koyama et al., 2009). Bombyxin stimulates cell proliferation in the wing discs of lepidopteran larvae (Nijhout et al., 2007).

Levels of JH and ecdysone are sensitive to larval nutrition (Tu and Tatar, 2003). JH is produced by the CA in growing larvae and suppresses 20-hydroxyecdysone levels by inhibiting PTTH release (Tu and Tatar, 2003). When larvae achieve critical weight, JH esterase is produced by the fat body and other organs and degrades JH. PTTH is released and signals the PG to produce ecdysone production (which suppresses growth, ends feeding and causes metamorphosis) (Tu and Tatar, 2003).
The interaction between tissues and hormones in the regulation of growth is extremely complex. Knowledge of how nutrition is coupled with the growth of a novel trait could provide valuable clues as to how organisms develop and evolve phenotypically plastic traits in response to environmental variation as in the case of beetle horn dimorphisms.

**Life Span**

The insulin signaling pathway is the most thoroughly studied pathway that affects aging in *C. elegans*, and it shares many homologues with the vertebrate pathway (Wolff and Dillin, 2006). DAF-2 signals through AGE-1—a conserved homologue of the catalytic subunit p110 of PI3K—, which activates PDK-1 and Akt-1/Akt-2/SGK-1 through conversion of PIP2 to PIP3 (Wolff and Dillin, 2006). Mutations in *daf-2* and *age-1* significantly extend the life span of *C. elegans* (Panowski and Dillin, 2009). Akt1, 2 and SGK-1 phosphorylate the *C. elegans* ortholog of mammalian FOXO transcription factors, DAF-16 (Wolff and Dillin, 2006). This pathway is inactivated through the *C. elegans* ortholog of PTEN, DAF-18 (Wolff and Dillin, 2006). Null mutations of *daf-16* reversed the extended life span phenotypes caused by *daf-2* and *age-1* mutations (Christyakova, 2008). Constitutive nuclear localization of DAF-16 is not sufficient to extend life span and other regulators are also required for full IIS-regulated longevity. However nearly all *daf-2* mutant phenotypes are dependent on DAF-16. Many genes regulated in *daf-2* mutant worms in a *daf-16*-dependent manner play a part in stress resistance. Many of these target genes affect life span, and knockdown of these target genes by RNAi shortens the life span of *daf-2* mutants (Panowski and Dillin, 2009).

As mentioned above, *C. elegans* ILPs function as both agonists and antagonists of DAF-2. Overexpression of INS-1 results in reduced signaling through DAF-2 and an increased life
span in wild-type worms (Pierce et al., 2001). ILPs known to function as agonists of DAF-2 are DAF-28 and INS-7 (Panowski and Dillin, 2009). Loss of function of either of these genes results in extended life span and nuclear localization of DAF-16 (Panowski and Dillin, 2009).

Evidence suggests that there are particular tissues that regulate longevity in relation to insulin signaling (Panowski and Dillin, 2009). The effects of age-1 mutation are reversed by expression of age-1 in the neurons or intestine (Iser et al., 2007; Wolkow et al., 2000). Loss of daf-2 function in a subset of neurons resulted in extended life span (Panowski and Dillin, 2009). Expression of daf-16 in the intestine of daf-2;daf-16 double mutants and in the intestine and neurons of age-1;daf-16 double mutants restored life span extension (Iser et al., 2007; Libina et al., 2003). These results indicate that the intestine is a key tissue in extending life span. This tissue serves as the adipocyte tissue of C. elegans and is the primary place of fat accumulation. It is interesting to note that fat-specific insulin receptor knock-out (FIRKO) mice are also long-lived (Kimura et al., 1997).

Diet has an effect on life span (Taguchi and White, 2008) and recent evidence indicates that a low-sugar diet may be the key to extending life span in higher organisms (Lee et al., 2009). Feeding glucose to wild-type C. elegans reduced life span in these animals by inhibiting DAF-16 and a cooperating heat shock factor, HSF-1. The gene responsible for this effect is the aquaporin gene aqp-1 which encodes a glycerol channel (Lee et al., 2009). Insulin downregulates gene expression of similar aquaporin glycerol channels in mammals (Lee et al., 2009).

Knockdown of the dir gene is fatal indicating that insulin signaling is essential in Drosophila (Fernandez et al., 1995; Chen et al., 1996). Life span in flies mutant for DIR or chico is significantly extended (Clancy et al., 2001; Tatar et al., 2001). Treatment with methoprene, a
juvenile hormone analog, of DIR mutant flies restored normal life span (Tatar et al., 2001). Over-expression of dFOXO in the pericerebral fat bodies of adult flies and in the fat body reduced the expression of ILPs in neurons and extended the life span of female flies (Hwangbo et al., 2004; Giannakou et al., 2004).

There is evidence that insulin action is responsible for deterioration of heart function in older flies (Wessells et al., 2004). Flies in which insulin signaling had been reduced exhibited nominal heart degeneration over the course of their lives and had a prolonged life span (Wessells et al., 2004). Suppression of insulin signaling in the heart itself eliminated deterioration of heart function but life span was normal (Wessells et al., 2004).

Age-related locomotor impairment (ARLI) is of particular concern in aging humans (Jones et al., 2009). Recent evidence from Drosophila indicates that disruption of PDK1, dp110 and Akt leads to delayed development of ARLI and points to insulin signaling as a key regulator of age-related decline in Drosophila (Jones et al., 2009). Eliminating germ cells (GCs) late in development or in the adult fly leads to increased life span and affects insulin signaling (Flatt et al., 2008). Flies exhibit increased production of DILPs and hypoglycemia; however, dFOXO target genes such as d4E-BP, l(2)efl, and IMP-L2 are upregulated (Flatt et al., 2008).

**Reproduction**

Mosquitoes are effective vectors of pathogens because of repeated cycles of blood ingestion and egg development by females (Roy et al., 2007). In some female mosquitoes the reproductive cycle is triggered by the ingestion of a blood meal from vertebrate host (Roy et al., 2007). Feeding stimulates release of neuropeptides from the brain MNCs of female A. aegypti, which induce the ovaries to produce ecystyroid hormones (Brown et al., 1998). Both porcine
and bovine insulin activate ecdysteroid production by the ovaries in vitro (Graf et al., 1997, Brown et al., 1998). Inhibitors of the IR and PI3K blocked this stimulation (Riehle and Brown, 1999). One of the ILPs (ILP3) was synthesized and shown to stimulate yolk uptake by oocytes and ecdysteroid production by the ovaries at significantly lower concentrations than bovine insulin (Brown et al., 2008).

In addition to being located in the ovaries (Graf et al., 1997; Riehle and Brown, 2003), the IR and PKB/Akt have also been detected in the fat body of adult female yellow fever mosquitoes, A. aegypti (Roy et al., 2007). Evidence suggests that insulin in conjunction with 20-hydroxyecdysone from the ovaries of the female mosquito and amino acid signaling through TOR and S6K stimulate the fat body to produce the yolk protein precursor, vitellogenin (Vg) (Roy et al., 2007). Recently in A. aegypti, six splice variants of AaegPTEN were identified and shown to be differentially expressed during development and in adult tissues (Riehle and Brown, 2007). The splice variant AaegPTEN6 particularly was expressed in the key reproductive tissues—the fat body and ovary (Riehle and Brown, 2007). Knockdown of AaegPTEN—or the variant AaegPTEN6—lead to an increase in egg production during the first reproductive cycle without affecting egg viability (Arik et al., 2009).

Ablating the MNCs in D. melanogaster produces females that are smaller than wild type females and have reduced fecundity (Broughton et al., 2005). Deletion of DILPs1-5 generated homozygous adults that were poorly viable and displayed reduced fertility (Zhang et al., 2009). In females mutant for DIR the ovaries are arrested at the previtellogenic stage (Tatar et al., 2001). Treatment with methoprene, a juvenile hormone (JH) analog, restored vitellogenesis
which suggests a link between these two pathways. Mutations in chico results in females that are sterile (Bohni et al., 1999).

**Metabolism**

Insulin functions in mammals to regulate blood glucose (Wu and Brown, 2006). It induces cells to take up glucose and convert it to glycogen for storage (Wu and Brown, 2006). Insulin inhibits glycogen break down and gluconeogenesis and shifts cell metabolism from catabolic to anabolic (Wu and Brown, 2006). Many early studies of the conserved action of insulin focused on injecting insects with vertebrate insulins and studying the effects (Wu and Brown, 2006). However due to the discovery of insect ILPs and the ability to synthesize or genetically manipulate them to study the regulation of carbohydrate and lipid metabolism in insects, scientists can now determine if effects of vertebrate insulins in earlier studies are real (Wu and Brown, 2006; Brown et al., 2008). Bombyxin II lowered hemolymph trehalose levels—the major hemolymph sugar in insects—, raised trehalase activity, and lowered glycogen content and elevated glycogen phosphorylase activity in the fat body when injected into neck-ligated *B. mori* larvae. The effect was not the same in adult *B. mori* (reviewed in Wu and Brown, 2006). Contrary to insulin function in mammals, the function of bombyxin II was to advance the breakdown and utilization of carbohydrate reserves (reviewed in Wu and Brown, 2006). Yet results from adult *B. mori* indicate that regulation of metabolism throughout the life of an insect may be complex (reviewed in Wu and Brown, 2006). In support of this view, *A. aegypti* synthetic ILP3 reduced circulating sugar levels and elevated carbohydrate and lipid storage in decapitated adult female mosquitoes (Brown et al., 2008). In addition, deletion of *Drosophila ILPs1-5* produced larvae that displayed metabolic abnormalities analogous to those caused by
loss of insulin action in mammals such as elevated circulating sugar levels and starvation responses in fat tissues (Zhang et al., 2009). These symptoms are similar to those for human Type I diabetes. Expression of DILP2 in these larvae partially rescued the starvation response and lowered circulating sugar levels to or below wild-type levels (Zhang et al., 2009). Ablation of IPCs in larvae elevated circulating trehalose and glucose levels (Broughton et al., 2005; Rulifson et al., 2002). Although the effects of deletion of ILP genes and ablation of IPCs are similar, IPC ablated larvae had higher hemolymph sugar levels, indicating that ablation of IPCs may affect additional signals (Zhang et al., 2009). Flies with mutations in DIR and CHICO displayed up to a fivefold increase in lipids (Bohni et al., 1999).

From the discovery of the first insect ILP, bombyxin, our knowledge of the insulin signaling pathway in insects has expanded to include multiple ILPs from different species and an intricate but evolutionarily conserved signaling pathway. Its functions are based on the studies of only a few model organisms. Future investigations that include non-model organisms such as scarab beetles may expand our knowledge of the functions of this pathway not only in growth and development but also in metabolism, longevity, and reproduction.
CHAPTER TWO
MATERIALS AND METHODS AND RESULTS

Materials and Methods

**Beetle rearing**

Laboratory beetle colonies were reared from *Onthophagus nigriventris* adults collected from cow and horse manure at approximately 3000 feet from the Kahua Ranch, Kamuela, Hawaii 96743 in late May 2007, 2008, and 2009. Adults were kept and bred in cylindrical plastic containers 21 cm tall and a diameter of 7.5 cm with approximately 600 ml of a sterilized soil/sand mixture and 200 ml of cow dung in an environmental chamber with 80% humidity, at 26 °C, under a 16:8 light:dark cycle. Typically 2-4 females were housed in each tube with 1-2 males, however, we found that under our breeding conditions, *O. nigriventris* performed better in single breeding pairs. Every fourteen days the beetles were removed from the plastic containers, and brood balls were collected. Brood balls were kept under the same environmental conditions described for adults in 16 oz plastic, covered, deli dishes on vermiculite and the dishes were dated. Newly emerged beetles require a period to feed and become sexually mature (approximately one week) and were kept separate from older beetles. Beetle larvae were sexed during the third instar according to the method of Moczek and Nijhout (2002). Beetle larvae were observed and weighed until the prepupal period when feeding stops and physiological changes such as the gut purge and imaginal disc proliferation begin. At the early prepupal period, larvae are characterized by a complete gut purge and translucent cuticle while late prepupal period (approximately 12 hours later at 26°C) larvae are characterized by a flattening of the typical C-shape of the larva and are extremely opaque.
RNA Extraction and cDNA synthesis for cloning and sequencing

Total RNA was extracted from whole tissues of late prepupal stage *O. nigriventris* using the Qiagen RNAlater reagent and RNeasy® Plus Mini kit and protocol (Qiagen, Valencia, CA). In order to convert RNA to cDNA, the SMART™ cDNA amplification kit (Clontech, Mountain View, CA, USA) and the SuperScript III First Strand Synthesis SuperMix kit and protocol (Invitrogen, USA) were used. At this point, 1 µl of the 5’ template was used in RACE reactions. For PCR and 3’ RACE reactions the template was further amplified by Long Distance PCR (LDPCR) according to the SMART™ cDNA amplification kit instructions (Clontech, Mountain View, CA, USA) using Advantage® 2 Taq polymerase mix (Clontech, Mountain View, CA, USA) and purified using the High Pure PCR Product Purification Kit and protocol from Roche Diagnostics (Indianapolis, IN, USA). The template was then normalized to insure that all transcripts were present in equal amounts in a sample and that more abundant transcripts did not overshadow rare targets. This step was accomplished using Duplex Specific Nuclease (DSN) enzyme, 10X DSN buffer, and protocol from Evrogen (Moscow, Russia). Using 2 µl of the normalized cDNA, LDPCR reactions were performed according to the protocol used for previous amplification of cDNA. The resulting template was diluted 1:10 before being used in PCR reactions. Incubation steps were performed in a Mastercycler gradient thermocycler (Eppendorf, Westbury, NY).

Cloning of the *O. nigriventris* insulin receptor transcript

In order to obtain the initial fragment of the *O. nigriventris* insulin receptor, degenerate primers were used (Table 1; Roovers *et al.*, 1995) in PCR. The cycling conditions were as follows: 94°C for 3 minutes, then 30 cycles of 94°C for 30 seconds, 50°C for 30 seconds, and
72°C for 3 minutes, with a final extension of 72°C for 5 minutes. This fragment, which was 450 bp in length, was from the conserved tyrosine kinase domain (TKD). From this fragment, gene specific primers were designed using Primer 3 (http://primer3.sourceforge.net/) (Rozen and Skaletsky, 2000) for 5’ and 3’ RACE and used along with the 3’-RACE CDS Primer A and the Universal Primer Mix (UPM) from the SMART™ RACE cDNA Amplification kit according to the manufacturer’s instructions (Clontech, Mountain View, CA, USA). These primers are listed in Table 1 along with their annealing temperatures (T_a). Gel electrophoresis was performed using a 1.2% agarose gel. Bands appearing at the predicted position were excised and placed in a GenElute™ Minus EtBr Spin column and the DNA was extracted using the manufacturer’s protocol (Sigma, Saint Louis, MO, USA).

Ligation and cloning were performed using the TOPO TA Cloning® kit from Invitrogen (Carlsbad, CA) following the manufacturer’s protocol. The pCR®II-TOPO® 4.0 kb vector and chemically competent One Shot® TOP10 E. coli cells were used. Fifty microliters of each reaction was spread on prepared LB plus 100 μg/ml ampicillin agar plates and incubated overnight at 37 °C. Transformed colonies chosen for purification were incubated overnight in a 37 °C air incubator, with shaking at 225 rpm, in sterile, covered test tubes containing 3 ml of LB broth plus 100 μg/ml of ampicillin. To isolate the plasmid DNA from E. coli, the Qiagen miniprep kit and protocol were used (QIAGEN Sciences, Maryland, USA). Fragment sizes were examined by digestion of the miniprep DNA using the restriction enzyme EcoR1 and protocol (Promega, Madison, WI, USA). Miniprep DNA was sent to MCLAB (South San Fransico, CA) for sequencing using the M13 forward and M13 reverse primers. Sequences were blasted using tblastx (NCBI) to compare them to sequences from other organisms to determine if any were
similar to the InR transcript. Seqman (Lasergene software, DNASTAR®, Inc.) was used to
generate putative OnInR fragments and produce a contiguous sequence from all sequencing
projects. Each nucleotide was sequenced at least 3 times but for many regions up to 10 times
coverage was obtained.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Direction</th>
<th>Sequence</th>
<th>Tₘ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>InR</td>
<td>Forward</td>
<td>CTTYGGNATG TAYGAR GG</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>C GTC ATN CCR AAR TCN CCR ATY TT</td>
<td>55</td>
</tr>
<tr>
<td>OnInR F18</td>
<td>Forward</td>
<td>AT TTA TCG GCG AAG AAG TTT GTC CAT CG</td>
<td>65</td>
</tr>
<tr>
<td>OnInR R22B</td>
<td>Reverse</td>
<td>AAG GAG GGC GGT TGC CCA ATA ATA CCA T</td>
<td>65</td>
</tr>
<tr>
<td>InR 5’UTR2</td>
<td>Reverse</td>
<td>GTT TCT ACA GAC GGA GCA TTG TTT G</td>
<td>65</td>
</tr>
<tr>
<td>InR 3’UTR2</td>
<td>Forward</td>
<td>GGA AGC TAG ATC AAC CAG AAA CAA TC</td>
<td>65</td>
</tr>
</tbody>
</table>

**Table 1. Degenerate and gene specific primers used to clone the OnInR.** Degenerate primers
were taken from Roovers *et al.*, 1995. Gene specific primers were designed using Primer 3. All
primers are shown in 5’ to 3’ orientation.

**Annotating the sequence**

The complete open reading frame (ORF) and translational start site of the OnInR was
predicted using the ORF Finder on the NCBI website (http://www.ncbi.nlm.nih.gov/gorf/gorf.htm). The signal peptide was predicted using Signal 3.0 (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen *et al.*, 2004; Nielsen *et al.*, 1997). Other conserved domains were predicted using the Conserved Domain Database (CDD) on the NCBI website and comparison with
*Drosophila* insulin receptor (DIR) (Fernandez *et al.*, 1995; Ruan *et al.*, 1995), molluscan insulin-
related peptide receptor (MIPR) from *Lymnaea stagnalis* (Roovers *et al.*, 1995) and mosquito
insulin receptor from *Aedes aegypti* (AaeInR) (Graf *et al.*, 1997). A protein alignment was
created by the Clustal W method using MegAlign (Lasergene software, DNASTAR®, Inc.) in
order to determine similarity between the OnInR and other IRs. Insulin receptors used in this
alignment are listed in Table 2 below along with their accession numbers and the number of amino acids that were used in the alignment.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession Number</th>
<th>Number of amino acids used</th>
</tr>
</thead>
<tbody>
<tr>
<td>O. nigriventris</td>
<td>N/A</td>
<td>1370</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>ACL83551</td>
<td>1615</td>
</tr>
<tr>
<td>H. sapiens</td>
<td>AAA59452</td>
<td>1382</td>
</tr>
<tr>
<td>A. aegypti</td>
<td>AAB17094</td>
<td>1371</td>
</tr>
<tr>
<td>B. mori</td>
<td>NP_001037011</td>
<td>1472</td>
</tr>
</tbody>
</table>

Table 2. **Insulin receptors used in protein alignment.** Accession numbers and number of amino acids used in the alignment are indicated.

**Phylogenetic analysis of OnInR**

A phylogenetic analysis of OnInR with IRs from different arthropods and mammals was performed using ClustalX 2.0.12 (Larkin et al., 2007). The entire prepropeptide of each IR was entered. For an outgroup, representatives were chosen from the molluscs: *Crassostrea gigas* (the pacific oyster), *Lymnaea stagnalis* (great pond snail), and *Biomphalaria glabrata* (a species of air-breathing freshwater snail). These and other species used in the analysis are listed in Table 3 along with accession numbers for their insulin receptors and the number of amino acids used from each sequence. Phylogenetic analysis was performed using neighbor joining analysis. A bootstrap analysis with 1000 replicates was performed.
Table 3. Insulin receptors used in phylogenetic analysis. Accession numbers and number of amino acids used in the analysis.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession Number</th>
<th>Number of amino acids used</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>H. sapiens</em></td>
<td>AAA59452</td>
<td>1382</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td>AAA39318</td>
<td>1372</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>ACL83551</td>
<td>2144</td>
</tr>
<tr>
<td><em>A. aegypti</em></td>
<td>AAB17094</td>
<td>1371</td>
</tr>
<tr>
<td><em>Anopheles gambiae</em></td>
<td>XP_320130</td>
<td>1316</td>
</tr>
<tr>
<td><em>M. sexta</em></td>
<td>ACI02334</td>
<td>1064</td>
</tr>
<tr>
<td><em>B. mori</em></td>
<td>NP_001037011</td>
<td>1472</td>
</tr>
<tr>
<td><em>O. nigriventris</em></td>
<td>N/A</td>
<td>1370</td>
</tr>
<tr>
<td><em>Tribolium castaneum</em></td>
<td>EFA11583</td>
<td>1363</td>
</tr>
<tr>
<td><em>C. elegans</em></td>
<td>ACC47715</td>
<td>1846</td>
</tr>
<tr>
<td><em>Schistosoma mansoni</em></td>
<td>AAV65745</td>
<td>1736</td>
</tr>
<tr>
<td><em>Echinococcus multilocularis</em></td>
<td>CAD30260</td>
<td>1736</td>
</tr>
<tr>
<td><em>Crassostrea gigas</em></td>
<td>CAD59674</td>
<td>804</td>
</tr>
<tr>
<td><em>Lymnaea stagnalis</em></td>
<td>CAA59353</td>
<td>1607</td>
</tr>
<tr>
<td><em>Biomphalaria glabrata</em></td>
<td>AAF31166</td>
<td>1671</td>
</tr>
</tbody>
</table>

**RNA extraction and cDNA synthesis for RT-PCR and quantitative PCR**

Larvae from the late prepupal period were weighed and anaesthetized on ice and horn, wing, leg, and genital (males) imaginal discs were dissected from large and small males and females of all sizes for RNA isolation. Size categories of large and small males were determined by the weight of the prepupa at the time of dissection: large males are greater than or equal to 0.25g and small males are less than or equal to 0.20g to ensure differences in size.

Total RNA was extracted from imaginal discs listed above using the TRIzol® reagent and protocol (Invitrogen, USA). In order to ensure no contamination from genomic DNA, the RNA was treated with DNaseI according to the manufacturer’s protocol (Ambion, Austin, TX, USA). First-strand cDNA was made using 1 µg of total RNA and the SMART™ cDNA amplification
kit and MMLV reverse transcriptase (Clontech, Mountain View, CA, USA). Ten microliters of
the template was further amplified by LDPCR according to the SMART™ cDNA amplification
kit instructions using Advantage® 2 Taq polymerase mix (Clontech, Mountain View, CA, USA).
The cycling conditions were as follows: 95 ºC for 1 minute and 25 cycles of 95 ºC for 15
seconds, 65 ºC for 30 seconds, and 68 ºC for 6 minutes, hold at 4 ºC.

**RT-PCR of Imaginal disc tissues**

SMART cDNAs from imaginal disc tissues were investigated for the presence of the
OnInR transcript using gene specific forward and reverse primers listed in Table 4. Ribosomal
18S was used as a reference gene. These primers are also listed in Table 4. Large and small
males and large and small females were compared using horn, wing, leg, and genital (males)
tissues.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Direction</th>
<th>Sequence</th>
<th>Tₐ (ºC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OnInR F51</td>
<td>Forward</td>
<td>ccg aaa agt cca aat gga aa</td>
<td>65</td>
</tr>
<tr>
<td>OnInR R51</td>
<td>Reverse</td>
<td>tcg ccg tta gcc att aat tc</td>
<td>65</td>
</tr>
<tr>
<td>On18S 4</td>
<td>Forward</td>
<td>agg gaa gac acg ctg att cct tca</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>att ctt gga tgc tgc caa gac gga</td>
<td>65</td>
</tr>
</tbody>
</table>

**Table 4. RT-PCR primers.** Primers are shown in the 5’ to 3’ orientation.

**Quantitative PCR**

To quantify the relative abundance of *O. nigriventris* InR transcripts during late pre-pupal
development from small males, large males and female horn tissues, Quantitect SYBR green
PCR kit was used (QIAGEN, Valencia, CA, USA) on a GeneAmp 5700 sequence detection
system (Applied Biosystems, Foster City, CA) using gene specific primers listed in Table 5 for
the *O. nigriventris* InR and 28S rRNA genes. Primers were designed to be less than or equal to
150 bp with melting temperatures in the range of 56-58 °C with the computer program Primer3 (Rosen & Skaletsky 2000). Three replicates were performed for each sample, using 0.1 pg of cDNA for each replicate sample and 50 nM final primer concentration per reaction. Real-time qPCR cycling parameters were the same for both transcripts and were as follows: 95 °C for 5 minutes, then 40 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds and last a melting curve analysis of 80 repeats of 50 °C for 15 seconds with a temperature increase in increments of 0.5 °C and a 4 °C hold.

Statistical analysis

Relative expression of OnInR mRNA was determined using the comparative C_T method (ΔΔC_T). Standard errors (SE) and standard errors of the mean (SEM) are indicated for measures of relative expression. P values were calculated by using two-tailed Student’s t tests (Figure 10).

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Direction</th>
<th>Sequence</th>
<th>Fragment size (bp)</th>
<th>T_a (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>InR</td>
<td>Forward</td>
<td>aac acg caa caa acg cag aaa</td>
<td>119</td>
<td>56.5</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>taa tgt cgg ttg tcc ttg aga tac</td>
<td></td>
<td>54</td>
</tr>
<tr>
<td>28S</td>
<td>Forward</td>
<td>cgg atc etc cct aac acc aca ttt</td>
<td>103</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>aac aag gat tcc ctt agt agc ggc</td>
<td></td>
<td>58</td>
</tr>
</tbody>
</table>

Table 5. Gene specific primers for qPCR. All primers are listed in the 5’ to 3’ orientation.
Results

*O. nigriventris insulin receptor transcript*

The degenerate primers were used to amplify a 450 bp fragment from *O. nigriventris* cDNA. This fragment was analyzed using BLAST and exhibited high similarity to the tyrosine kinase domain of insulin receptors from other species. From these data we tentatively concluded that this cDNA fragment represented part of an OnInR homologue. Gene-specific primers were produced from this fragment using Primer 3 and used in RACE reactions to produce overlapping fragments of the full length transcript. Analysis of cloned fragments using SeqMan (Lasergene) revealed which fragments were identical in the areas of overlap. This indicated that these fragments originated from the same transcript. The complete contig was 4757 bp long including a complete 3’ UTR and a partial 5’ UTR and most likely encodes the putative OnInR. The longest open reading frame (ORF) begins at bp 55 and ends at 4162 and encodes a 1369-aa putative OnInR protein (Figure 3, 4, and 5).

*N-Terminal*

The predicted OnInR signal peptide extends from aa 1-37 (Figure 3). This sequence is extremely hydrophobic in character (22 of the 37 amino acids are hydrophobic). Like most proreceptor sequences it is preceded by an initiator methionine.

As in mammalian IR and other IR sequences known to date, the OnInR contains a cysteine-rich region (aa 202-356). This region of the OnInR contains 24 cysteines, 23 of which occupy conserved positions when compared with human, *Drosophila*, *A. aegypti* and *B. mori* IRs (Figure 3 and Appendix A). Overall similarity of this region when compared with the IRs listed above ranges from about 20% to 24%.
The two regions flanking the cystein-rich domain are known as the leucine-rich regions and are more conserved in sequence than the cysteine-rich region (Figure 3). It was once thought that these domains made up bilobal ligand binding sites and in the mammalian insulin receptor were important for binding insulin (DeMeyts, 2004). It is now known that while the first region is important for binding insulin, the second is not (Rentería et al., 2008). In the IGF-1 receptor the cys-rich region is known to contain the specificity determinants for IGF-1 binding (Hoyne et al., 2000). The N-terminal region (aa 74-186) displays 44 to 60% similarity with IRs listed above and the domain C-terminal to the cys-rich region (aa 365-486) displays about 37 to 53% similarity (Appendix A).
The processing site and fibronectin type III domains

Following the second ligand binding domain in HIR are 3 fibronectin type III domains (FnIII-1 to FnIII-3). The α-β cleavage site lies within FnIII-2 (Mckern et al., 2006). In the OnInR these domains correspond with aa positions 627-724 and 845-925 (Figure 4 and Appendix B). The region between these two corresponds with the FnIII-2 of the HIR and contains the conserved processing site but is not identified by CDD as a FnIII domain. The interdomain contacts are located at positions 626, 909, and 923. According to the HIR sequence, disulfide bonds between each α- and β- subunit are predicted to form between cysteines C678 and C824 and α-α disulfide bonds at C504 in the FnIII-1 domains (McKern et al., 2008). IRs are composed of two α-subunits and two β-subunits held together by disulfide bonds. Located at aa position 757-761 of the OnInR is a putative furin-like proteolytic processing site which is potentially the site where the subunits are liberated from each other (Figure 4). This stretch of aa follows the sequence rules for constitutive precursor cleavage (Watanabe et al., 1992).
The putative transmembrane domain of the OnInR consists of 20 aa (943-964), 19 of which are hydrophobic in nature (Figure 5). Though its hydrophobic character has been preserved, there is very little conservation of the sequence of this region when compared to human, Drosophila, A. aegypti and B. mori IRs (15-35%; Appendix C).

The putative juxtamembrane domain of the OnInR extends from aa 965-984 (Figure 5). There is limited conservation within this region when compared with species previously listed (25-40%; Appendix C) except for an NEPY motif that is implicated in coupling activated IRs to downstream signal transduction molecules (Sattar et al., 2007). It has previously shown that this

Transmembrane and Juxtamembrane domains

Figure 4 Fibronectin type III domains and processing site. Interdomain contacts (grey shade), FnIII-1 (italic) and 3 (underline) domains, cleavage site (fuchsia shade), conserved cysteines (yellow highlight, bold-disulfide bonds) and tyrosines (green highlight).
sequence is important but not absolutely necessary for coupling of HIR kinase to IRS 1 and p85 or for mediating insulin metabolic and mitogenic effects (Berhanu et al., 1997). But the role of the JM domain itself remains undetermined.

Tyrosine kinase domain

The prospective TK domain is the most conserved part of the putative OnInR (Figure 5). The percent identity when compared to the TKD of human, *Drosophila, A. aegypti* and *B. mori* IRs ranges from 68% to 77% (Appendix C). The OnInR TKD begins at aa 991 and ends at aa 1270. There is a potential ATP-binding site in a $^{1005}G-X-G-X-G^{1010}$ motif and a conserved Lysine$^{1031}$. Other conserved sequences include two motifs indicative of a protein kinase, $^{1154}D-F-G^{1156}$ and $^{1181}A-P-E^{1183}$, and specific tyrosine kinase consensus sequences, $^{1136}D-L-A-A-R-N^{1141}$ and $^{1176}P-V-R-W-M-A-P-E^{1183}$. When activated, the insulin receptor becomes autophosphorylated on specific tyrosine residues contained in a conserved YETDYY motif. In the OnInR these tyrosine residues are at positions 1162, 1166, and 1167. Additional conserved tyrosine residues are present at positions 1013, 1088, 1126, 1200, 1214, 1225, and 1243.
Figure 5. OnInR β subunit. Transmembrane domain (TMD; bold), juxtamembrane domain (JMD; italic), beginning and end of the tyrosine kinase domain (TKD; grey shade), ATP-binding site (fuchsia shade), protein kinase and protein tyrosine kinase motifs (blue and red shade).

Evolutionary relationship of OnInR to other species

A phylogenetic analysis of known arthropod and mammalian IRs was performed to determine the evolutionary relatedness of the OnInR to insect, non-insect arthropod, and mammalian IRs.

Molluscan IRs were used as an outgroup to anchor the neighbor-joining tree (Figure 6). The strongest link was between OnInR and T. castaneum IR. The next closest relationship appeared
to be between the Coleoptera and Diptera IRs. Bootstrap analysis did not support a relationship between the Coleoptera IRs and Lepidoptera IRs. There was, however, a strong association between the insect IRs and mammalian IRs. Nematode and flatworm IRs did not show any association with this group. The unrooted tree supports these groupings (Figure 7).
Figure 6. Phylogenetic tree of the known vertebrate and invertebrate IRs. Entire prepropeptide sequences were used. The tree was generated with ClustalW alignment and neighbor-joining tree with bootstrap analysis. Bootstrap values are shown at the nodes of the branches and represent the percentage of times that grouping was supported. Representatives from the molluscs were used as an outgroup.
Figure 7. Comparative IR phylogenetic tree. Entire prepropeptide sequences were
used. An unrooted phylogenetic tree was created using the Gonnet 250 matrix and
neighbor-joining algorithm. Support at each internal node was assessed using 1000
bootstrap samplings and the tree was visualized using Tree View. Branch lengths are
drawn to scale.
Distribution of OnInR transcript in imaginal discs

In order to determine the distribution of the OnInR transcript among different traits during the time of growth, gene specific primers were designed against the sequence and used in PCR to determine the presence or absence of the transcript in cDNA made from the RNA of imaginal discs. 18S ribosomal RNA was used as a reference. Horn, wing, and leg discs from one large female and one small female were investigated. In addition to these traits, genital discs were also investigated. All discs showed the presence of insulin receptor transcript (Figure 8). No template controls each containing one set of primers, either InR or 18S, showed no contamination. All single primer controls were clear (Figure 9).

Relative expression of OnInR mRNA in horn tissues

In order to determine the relative expression of the OnInR transcript in horn tissues, qPCR reactions were conducted using the *O. nigriventris* 28S ribosomal RNA (accession number EU162505) as a reference. Comparisons were conducted between large males, small males, large
females and small females to reveal any differences in OnInR expression levels between sex and/or body size. At this time no template controls and no reverse controls were not included in these experiments. Because these controls were not included, about one year after these tests were conducted further experiments were performed to determine if the RNA samples from the horn tissues were free of genomic DNA (gDNA) contamination. Due to gDNA contamination in many of horn tissue RNA samples despite treatment with DNase I, much of the qPCR data was not useable. Only RNA from small males and females was free of contamination and these comparisons are reported. Tests were conducted on five separate occasions. Four out of the five comparisons indicate that there is no difference in relative OnInR mRNA expression levels in horn tissues between small males and large females (see Appendix D). A representative set of data is shown below (Figure 10).
Figure 9. Single primer controls for RT-PCR of imaginal discs. Top row: OnInR primer controls. Bottom row: On18S primer controls. Molecular weight markers in nucleotides are indicated. Abbreviations: F (forward primer), R (reverse primer), LM (large male), SM (small male), LF (large female), SF (small female), H (horn), W (wing), L (leg), and G (genital) imaginal discs.
Figure 10. Expression of OnInR mRNA in the horn tissue. Expression levels of OnInR mRNA in small male and large female horn tissues is relatively the same (p=0.314989). Standard deviation (SD) and standard error of the mean (SEM) are indicated.
CHAPTER THREE

DISCUSSION

The insulin pathway represents an evolutionarily conserved mechanism for regulating growth and size in animals (Oldham et al., 2002). By 1975, insulin had been isolated and sequenced from all classes of vertebrates (Chan and Steiner, 2000). With advances in molecular techniques and genomics, genes for ILPs and other components of the insulin signaling pathway have been identified in species from Orthoptera, Diptera, Lepidoptera, Hymenoptera, and Coleoptera (Wu and Brown, 2006). Studies of Drosophila and C. elegans have shown that ILP signaling through a conserved insulin signaling pathway regulate development, longevity, metabolism, and female reproduction (Claeys et al., 2002; Garofalo, 2002; Goberdhan and Wilson, 2003; Wu and Brown, 2006; Edgar et al., 2006; Giannakou and Partridge, 2007; Chistyakova, 2008; Taguchi and White, 2008; Hietakangas and Cohen, 2009; Panowski and Dillin, 2009).

Structure of the O. nigriventris insulin receptor

The overall structure of the putative OnInR is extremely homologous to the mammalian insulin receptor and insulin receptors of other invertebrates. Like IRs from other species, it appears that the mature OnInR protein is a tetramer composed of two α and two β subunits. There are several important conserved features in the OnInR transcript which include potential insulin-binding domains, a cysteine-rich region, a tyrosine kinase domain, a furin-like site for proteolytic processing that potentially specifies the cleavage of the α from the β subunit, juxtamembrane and transmembrane domains. The potential α subunit contains the leucine-rich repeat domains (L1 and L2) and cysteine-rich domains. Conservation of cysteine residues in the
cysteine-rich domain may point to similar structural organization for the OnInR compared to other IRs. Intramolecular di-sulfide bonds between these cysteines link the α subunits together and the α subunits to the β and therefore are implicated in the IR tertiary (three-dimensional) structure.

Insulin binding sites have been characterized on the surfaces of the L1 domain and the cysteine rich region. Also, a region of conserved residues on the surface of the fibronectin type III domains has been identified as a strong candidate for an insulin binding site (Rentería et al., 2008). Early biochemical characterization of DIR showed preferential binding of the receptor to mammalian insulin over IGF-1 (Petruzelli et al., 1986; Fernandez-Almonacid and Rosen, 1987) but DIR is no more similar to HIR than IGFR when their corresponding leucine-rich and cysteine-rich regions are aligned. This is also true of MIPR. Future ligand binding experiments will need to be conducted in order to determine if the OnInR will bind any commercially available vertebrate or invertebrate insulins or IGF-1 or any Onthophagus ILPs that may be discovered.

Insulin receptor α and β subunits are produced from a common proreceptor by post-translational processing at a putative tetrabasic processing site, RKKR. This motif is present in HIR and other insect IRs and is thought to be recognized and cleaved by furin endoproteases (Watanabe et al., 1992). Since this tetrabasic motif is conserved in the OnInR, this suggests that the receptor will be processed in the same way as previously characterized IRs. However, further characterization will be needed to confirm this.

The first domain of note in the β subunit region is the transmembrane domain (TMD). This domain links the β subunit with the α subunit and crosses the cell membrane. Observations
that single amino acid mutations in this domain lead to malfunction of human Receptor Tyrosine Kinases which causes certain types of cancer, indicates that interactions of single transmembrane helices are probably important for RTK activation and signaling (Finger et al., 2009). In a recent study, all human RTK TMDs were shown to be able to form oligomeric structures within a membrane. This evidence further indicates that TMD interaction is a property of RTKs and is important for activation and signaling of human RTKs (Finger et al., 2009). Given the evolutionary conservation of insulins and IRs across species, it is possible that this domain will serve the same function in insects.

The next domain of interest is the juxtamembrane domain (JMD). This domain is not well conserved in the OnInR except for an NPEY motif. This motif is a major site of phosphorylation and couples the activated HIR to downstream signal transduction molecules (Sattar et al., 2007). Phosphorylation of the tyrosine in this motif has also been implicated in ligand-induced receptor internalization (Fernandez et al., 1995). Endocytosis of the IR is stimulated by insulin and the receptor is internalized by pinching off of coated pits from the cell membrane that form coated vesicles (Backer et al., 1992). The coated pit-mediated endocytosis of many cell surface receptors relies on structural features in their cytoplasmic domains. Deletions and mutations in the cytoplasmic juxtamembrane region of the receptor prevent receptor internalization. These mutations include the residues NPEY (Backer et al., 1992). Through regulating the number of IRs expressed on the cell surface and by facilitating the intracellular decomposition of insulin, stimulation of the internalization of the insulin receptor plays a key role in the biological action of the hormone (Carpentier et al., 1992). Recent evidence indicates that the presence of intact HIR JMD is necessary for insulin-stimulated Shc—an Src
homology 2 (SH2)-domain containing protein—phosphorylation but not crucial for IRS-1 phosphorylation. Insulin signaling can also take place independent of the JMD of HIR and its NPEY motif. The mechanism is Shc independent and is dependent on insulin concentration. It was also shown that insulin internalization does not absolutely require the specific sequence of the HIR JM domain, which can be partially replaced by the JMD of a similar tyrosine kinase (Sattar et al., 2007). Although experiments in Drosophila indicate that the function of this domain could be conserved in insects (Yamaguchi et al., 1995), it is not yet know if O. nigriventris has IRS-1 or Shc-like molecules.

The predicted TK domain is the most conserved region of the OnInR. The first feature of note is the potential ATP-binding site (GXGXXG\textsuperscript{1010} and Lysine\textsuperscript{2012}; Hanks et al., 1988). Human insulin receptors mutated at this site are able to be processed into subunits and bind insulin but lack protein tyrosine kinase activity and fail to transduce the signal (Chou et al., 1987). Two other conserved motifs indicative of a protein kinase, 1154-D-F-G\textsuperscript{1156} and 1181-A-P-E\textsuperscript{1183}, are implicated in ATP-binding (the former) and increased catalytic activity due to phosphorylation (the latter) (Hanks et al., 1988). Specific protein tyrosine kinase consensus sequences, 1136-D-L-A-A-R-N\textsuperscript{1141} and 1176-P-V-R-W-M-A-P-E\textsuperscript{1183}, indicate substrate specificity (Hanks et al., 1988). The activation loop which extends from K\textsuperscript{1151} to L\textsuperscript{1174}, contains three conserved tyrosines at positions 1162, 1166, and 1167. These tyrosines become autophosphorylated when the HIR is activated and cause the receptor to undergo a conformational change which allows ATP and other substrates to bind to the active sites. Due to the conservation of these key features, it seems likely that the activation of the OnInR follows a similar mechanism as that of its invertebrate and
mammalian counterparts. The evolutionary relationships revealed by phylogenetic analysis suggest that growth and metabolic functions of the insulin receptor may also be conserved.

**Expression of OnInR**

The insulin signaling pathway has been suggested as a mechanism for modulating the amount of horn growth in scarab beetles in response to nutrition (Emlen et al., 2006, 2007). It has already been determined that beetle horn growth depends heavily on nutrition during the larval stages. In order to determine the likelihood of this pathway’s involvement in horn and trait growth, we investigated the distribution of OnInR mRNA in *O. nigriventris* imaginal discs during the late prepupal period, which is the period of maximum trait growth. OnInR mRNA was found in all imaginal discs from males and females both large and small. This result was not unexpected since during the larval stages of *Drosophila*, prominent expression of DIR mRNA is observed in the imaginal discs (Garofalo and Rosen, 1988). Later it was determined that DIR affects the size of body parts autonomously (Brogiolo et al., 2001). Over the years, genetic manipulation of insulin signaling pathway components within specific imaginal discs in *Drosophila* has been shown to affect the final sizes of those traits by affecting the rate of cell proliferation (Goberdhan and Wilson, 2002; Kramer et al., 2003; Puig et al., 2003; Shingleton et al., 2005).

Emlen et al. (2006) predicted that if the insulin pathway was involved in differential horn growth, then horn discs in large males would be sensitive to circulating insulin signals while horn discs in small males and females would not be. They predicted that in small males and females cell proliferation in the horn disc would be uncoupled from circulating insulin signals. Quantitative real-time PCR data (Emlen et al., 2006) showed that by the end of the period of
horn growth in *O. nigriventris*, relative OnInR expression was significantly lower in large males with horn growth than in small males and females that did not grow a horn. Also this data indicated that mRNA expression levels between small males and females were essentially the same (Emlen *et al.*, 2006). Although this may seem counterintuitive, it is not unlike the situation that occurs when a physiological mechanism arrests overall growth in animals under starvation conditions (Kramer *et al.*, 2003; Puig and Tijan, 2006). This mechanism is the feedback mechanism discussed in the introduction. High nutrient availability leads to insulin secretion which activates the insulin receptor and the eventual phosphorylation and inactivation of dFOXO. When nutrient levels decline, insulin levels and signaling will decrease and dFOXO becomes active again and moves into the nucleus where it activates transcription of growth inhibitory genes and other targets including the insulin receptor (Kramer *et al.*, 2003; Puig and Tijan, 2006). In a recent study of insulin pathway activity during caste development in *Apis mellifera*, a similar situation was described (de Azevedo and Hartfelder, 2008). During the fourth instar, queen larvae surpass worker larvae in growth rate and show a steeper sigmoid growth curve (de Azevedo and Hartfelder, 2008). This is also the time that worker larvae are shifted to a modified worker diet (de Azevedo and Hartfelder, 2008). At this time expression of both AmInR-1 and AmInR-2 mRNA declines greatly in queen larvae but not in worker larvae (de Azevedo and Hartfelder, 2008). Considering the differences in morphology of queen bees and worker bees and the fact that it is the queen that must reproduce, it appears that this feedback mechanism may also be active in this species.

In this study, relative quantification of OnInR mRNA levels in horn tissues of small males and large females at the end of imaginal disc growth revealed that there was essentially no
difference in expression between them (Figure 10). This confirms the data presented by Emlen et al. (2006) at least for small males and large females.

For each insect that has been shown to have an InR, corresponding ILPs have been cloned. For example, ILP3 from *Aedes aegypti* (Brown et al., 2008) has been shown to bind to the MIR. A recent study has provided evidence that insulin signaling plays a role in imaginal disc growth (Koyama et al., 2008). In *Manduca*, insulin overrides the growth inhibiting effect of JH on wing imaginal discs. However, if the insulin receptor is knocked down in imaginal discs, insulin is no longer able to affect the discs’ sensitivity to JH (Koyama et al., 2008). Distribution of the OnInR transcript in all imaginal discs during the period of disc growth suggests that insulin signaling will be involved in the mechanism of trait growth particularly that of differential horn growth. However, studies correlating receptor protein expression with transcript expression will be needed in addition to insulin injections or receptor knockdown studies in order to fully test this idea.

**Summary**

The results of this study are critical in beginning to understand the function of the insulin signaling pathway in horned scarab beetles and its role in the control of horn growth. Emlen et al., 2006 identified the insulin signaling pathway as one of the candidate pathways for the origin and diversification of scarab beetle horns. The insulin receptor of one of those beetles, *O. nigriventris*, has been identified and localized in the horn imaginal tissue during the period of horn growth in addition to other imaginal tissues. With the insulin receptor identified, molecular techniques can now be used to manipulate it and define its role during scarab development and horn growth. In order to study allometry, activity of the receptor can be measured between traits
in an individual, among individuals of a population, and populations that differ in allometry. These results contribute important new information to understanding how phenotypically plastic traits are regulated at the genetic level.
BIBLIOGRAPHY


Chou CK, Dull TJ, Russell DS, Gherzi R, Lebwohl D, Ullrich A, Rosen OM. 1987. Human insulin receptors mutated at the ATP-binding site lack protein tyrosine kinase activity and


Floyd PD, Li L, Rubakin SS, Sweedler JV, Horn CC, Kupfermann I, Alexeeva VY, Ellis TA, Dembrow NC, Weiss KR, Vilim FS. 1999. Insulin prohormone processing, distribution,


Nijhout HF. 2008. Size matters (but so does time), and it’s ok to be different. Developmental Cell 15: 491-492.


| Majority | XGXXVXRACXXLLXVAAXLXAXAXYAGX-X-X-X----------------GVCXSMDIRNXPX |
| OnInR protein.pro | MASNCSICLADGTGRCWLLVAVVWVFQFNCYLVPVTAGYT ---------------------EYNQFTYPKTSGYKSVDIRNGE 66 |
| DIR protein.pro | PISLSSLLFIILANTLIAQVVLPAHQQLHLNDANDGLKDTALSQGTQTRPRESNPMLSLQVPKCSMDIRNMS 349 |
| HIR protein.pro | MGMGGRAPAAAPLLVAVAALLLGAAGHLPG ------------------------EVCPSMDIRNNL 45 |
| AaeInR.pro | MASQQNMHNLG ---------------------------VCPSDVQRLNSPA 24 |
| Bom InR.pro | MEMLRGGMRACDRLGWPACCALLIAACARTGYTEV ------------------------VP ----------------ARGVCPSMDIRNLP 55 |
| Majority | HXRLXRCRQUEGLQLLIDVXDKAXPSXENXSFPKLXEXITYLVLRYVNLXSLXFLPNALVRGQX-LFXXYALIFE |
| OnInR protein.pro | MFNNRRCVVEGYIQIIFDHDENGSTFENIFSPPELEITGHLLLYRVNGLSTLGKFLPNALVRGQQLLLGIALIFE 146 |
| DIR protein.pro | HFNQLECNVIEGFLLLIDINDAS--LNSPFPKELVEYDIYIRYTGHLSSKLFPNSLVRGNK-LFDGYALVVS 425 |
| HIR protein.pro | RLHELENCVGEHLQIILLFMFKTPREDYRDLFPKLLMIDYLLLFRVYGLLLSGKLLDFNLTPVIRGSR-LFFNYALIFE 124 |
| AaeInR.pro | HLRLKDLCVVEGFHILDDYKIDSSSENYSFPLLTEITEYLLFRVYNGLKSLRLFNPALVIRGDA-LVGYDAMVIYE 103 |
| Bom InR.pro | QIRRLGCRVEGLQLISLIVMERATKNDMNSPQLREVQYLLIKTYKGLRLNGDLFPNLPVVRGMQ-LFKDFALVID 134 |
| Majority | XXHLEXCRRVEGLQLLIDIVXDKAXPSXENXSFPKLXEXITYLVLRYVNLXSLXFLPNALVRGQX-LFXXYALIFE |
| OnInR protein.pro | MPSLQELEGKSLTDIMGGIRKIDKQPCALCDTANRWIRHHKKNDKAILFISRSKNECPCMPMDN ------------------------G 215 |
| DIR protein.pro | NFDMLDGLHKLRSITRGVRIEKHNKLCYDRTIDWLEILADN-ESQLVLTVENGKEESCSKCPGEIRIEEHDNTAI 504 |
| HIR protein.pro | MHLLKELJNLYMNITRESSQKIERKNKCECYLTDWSSRSLDSV-EDNHVLNKKDNEECGGDCPQTAK -----------------------GKTNS 196 |
| AaeInR.pro | LMIHEEGKSLMDITRRGTVRIEKNPCLCYANTIDWSRXXX--EDEAIWXRNSXKENECPXCPXXX-X-X-X-XXXC |
| Bom InR.pro | NEHLAELGLNSLMDIRGQTIHRDLYCYNTIDWSRITR--------DAAEVRSVNYDTLRGLCPSNAQSRVLEE-DHPRSSL 209 |
| Majority | PAXPX----------------KXXXRELQWNSHKQXCPXSXS-GCDEXTCCNSX-CLGCGCSX--DPSSCSVCRNFSDX---X-XCX |
| OnInR protein.pro | -------------------K-PLCEYWSSKHCQCKCSSNCTSY-CNDFLGCNCT-CIGGCSN--DPKQSCVCRNITFYREFKNTVCL 280 |
| DIR protein.pro | EGELNASCIHLHNNRRLCNWLSKLCQXKPEKCRN-NCIDEHTECCSNQ-CLGCGVIDKNGNESCURSCRNVSNFNN----ICM 577 |
| HIR protein.pro | PATVII--NGQQFRVQWTSHKQCVPTICKSHQTAEGLCICHSE-CLGNCSQP-DPTKVCARCRNYFYLD---RVC 265 |
| AaeInR.pro | PAAAVGVRGKNDKHRLCWANHCQTPICPPEPK-ACSVKTVQCCDASLGCNGLIP--NTSSSCVHRSLIDPAG-KRQCV 255 |
| Bom InR.pro | PQCIPA----------------DTKGGKLLCWEDEKHQCQKICPSACGHGCSNGTCCNSA-CLGCGDGP--LASNCFVCKKFSFYEGS-EMTCM 281 |
### Insulin receptor protein alignment of the L1 and L2 domains and the cysteine-rich region.

Abbreviations: *O. nigriventris* insulin receptor (OnInR), *Drosophila* insulin receptor (DIR), human insulin receptor (HIR), *A. aegypti* insulin receptor (AaeInR), *B. mori* insulin receptor (Bom InR).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>OnInR protein</td>
<td>S-RCCKVCPGIS-----VNIIETAESLEKETHITG---SIEIEVS---GKNVPAKLNNLMEIEEGYKLVKYSYPLSL</td>
<td>417</td>
</tr>
<tr>
<td>DIR protein</td>
<td>GGKCDKCECS-SG----LIDSLERAREFHGTITGTEPLTISIKRESGATAHVMDELKLYGALAAVHIQSSLMLVHLYTGLKSL</td>
<td>714</td>
</tr>
<tr>
<td>HIR protein</td>
<td>G-PCPKVCHLEEK-TIDSTSAQELRCGTVING--SLIIINIRG--GNLAEELANGLEIESGKLYKRISSWALVSL</td>
<td>406</td>
</tr>
<tr>
<td>AaeInR</td>
<td>G-TCPKRCEGSN----IDNIIAQDLLKGCEIIDG---SLEIQLRSRGGNIVKELENFLSSITEIKLYKLVRSYPLLSS</td>
<td>399</td>
</tr>
<tr>
<td>Bom InR</td>
<td>--PCPSRGVCHEQCLLGKSVDATAESFQGGCITKKEIKG--KMEFLTRRS--GGKTSSLLESALGEIRETHGCLQVTRSYPLVSL</td>
<td>431</td>
</tr>
<tr>
<td>Majority</td>
<td>XXCPGXYKXX---RCVTRXEICXXXXKPIXNSXXXXXXF---XXNGSCXXECPSGYTMXXX---RXCXPCX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>330                      340                      350                      360                      370                      380                      390                      400</td>
<td></td>
</tr>
<tr>
<td>OnInR protein</td>
<td>G-PCPKVCGXGSE-----IDSXEXAXELXGCTXITG--SGXIVXLEANLGXIEEIXGCLKVRYSYPLVSL</td>
<td></td>
</tr>
<tr>
<td>DIR protein</td>
<td>GGKCDKCECS-SG----LIDSLERAREFHGTITGTEPLTISIKRESGATAHVMDELKLYGALAAVHIQSSLMLVHLYTGLKSL</td>
<td>714</td>
</tr>
<tr>
<td>HIR protein</td>
<td>S-RCCKVCPGIS-----VNIIETAESLEKETHITG---SIEIEVS---GKNVPAKLNNLMEIEEGYKLVKYSYPLSL</td>
<td>417</td>
</tr>
<tr>
<td>AaeInR</td>
<td>G-PCPKVCHLEEK-TIDSTSAQELRCGTVING--SLIIINIRG--GNLAEELANGLEIESGKLYKRISSWALVSL</td>
<td>406</td>
</tr>
<tr>
<td>Bom InR</td>
<td>G-PCPKVCHLEEK-TIDSTSAQELRCGTVING--SLIIINIRG--GNLAEELANGLEIESGKLYKRISSWALVSL</td>
<td>406</td>
</tr>
<tr>
<td>Majority</td>
<td>XXCPGXYKXX---RCVTRXEICXXXXKPIXNSXXXXXXF---XXNGSCXXECPSGYTMXXX---RXCXPCX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>330                      340                      350                      360                      370                      380                      390                      400</td>
<td></td>
</tr>
</tbody>
</table>

Insulin receptor protein structure and function.
APPENDIX B
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority</td>
<td>AEDSGDXASCEIXVLKXSLSQXXSXXLXnxXPXXX-YQNVXFFGDRCAGGDWXVXD-</td>
<td>580</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>AENSGDKTACIEKDLIIIIEQSKSSTSVQVWVPFNT-DSRKLLSQLYHEAE-QVNSIFDSRDAGGDWSIQD-</td>
<td>570</td>
</tr>
<tr>
<td>DIR protein</td>
<td>GADSNGRNCGCTAVLSTVLSQGANSLNVTKEIGEPQPSATIVYKDPPAFIFGFYHMSPTQQLKAVTH-</td>
<td>872</td>
</tr>
<tr>
<td>HIR protein</td>
<td>ALKTNGDKACNELLKFSYRTSDFKILREWYWP-FDRDLGLMLFYKEAE-QVNTEDFGQDACGNSWTVQD-</td>
<td>560</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>LESNGNRAACSIETLTSKSIGEIQWAPFTELSARDMLGLYVYIYEAE-YANVFDDGRACNTEGWRLDD-</td>
<td>550</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>SQDSNYQASCLPDLVSXELQAVLTDIVYCPD-DMRKLGLSYLFMAT-EEKVSLYQRDAC-SDIKWILDP</td>
<td>585</td>
</tr>
<tr>
<td>Majority</td>
<td>--XXXISNXNSXEHAG-XTQKLKPYTQAYYVKTYYTXSSEX---GAQSDIKYFRNTPGQPS</td>
<td>650</td>
</tr>
<tr>
<td>OnInR protein</td>
<td>-VPVIPYDHVNSKIEAT-LKDLYATQYAVFVKTYYTIAQKEK---GGLSALKYFRTPQGQPS</td>
<td>627</td>
</tr>
<tr>
<td>DIR protein</td>
<td>--AMIAGKVSSEPKG-VMLSNLIPYNSYVRYRMAISLETT---NAESDVKNFRTPGQPS</td>
<td>930</td>
</tr>
<tr>
<td>HIR protein</td>
<td>IDPPRLSDPKSNQHPS-WLMRGLKPTQVAYVKTLVKTFSDDRTYGASKSDIVYQTDATNPS-</td>
<td>604</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>---ISDNMDKETTK-ILTLKYPYQTAYVKTYTGEGL-GGQSKIKYFTTAGTQPS</td>
<td>660</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>IEETRKVNNAKPSHARKTEVIFTNPMDPMDPLFPIQIQLPYRHYAYYVKTYTIDRQ---GAQSSIHYFRMLPQ</td>
<td>662</td>
</tr>
<tr>
<td>Majority</td>
<td>VPTDXXVXXXSXXSSXXVKKSPPXPPXPNGXTHXYXVVLNAQDRE-XXNDXYCXXPPPXXLAXSTXTPXX</td>
<td>730</td>
</tr>
<tr>
<td>OnInR protein</td>
<td>SPTFLNRANSSSIEVDWSPPPSPNGLTHYKVTIVLKTQDE---LEYSNCSVRSQSSPTTPDAPKTSPPS</td>
<td>701</td>
</tr>
<tr>
<td>DIR protein</td>
<td>--KVATIASDINVTWLSNDLMALRTYFKAKKLINRPTR---NNRDCYEPILKAMENDLPATTTPKISDPLA</td>
<td>1003</td>
</tr>
<tr>
<td>HIR protein</td>
<td>VPLDPISVNSSSSIIQKLWPPSPPNGNTHLYLVFWEQASEDE---LFELDYLCKLGLPSRTWSPPFEEDSQQ</td>
<td>700</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>VVNDVESPVL-KNMTLIVKLLPLKMNGLKEVLEVIELNADDNEQ-MLMRDYYC-DELRDIVPTEPSAPPKTSCT</td>
<td>680</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>PPTAVTVELTPTPSHPKSHPPMTMPGTVILYHEVQANSNRPQILAGABSGPTHNLAMMIKASLEAETPEEK</td>
<td>742</td>
</tr>
<tr>
<td>Majority</td>
<td>-----KKXGCXCECPKKXSXXLXTDXDEEEEXXINFENELQNXYXKKX---VNPXPPXXSTRXPS</td>
<td>810</td>
</tr>
<tr>
<td>OnInR protein</td>
<td>--KNDTCDTDDMRLRESANLKEATESRESFNEFELQNVRVIYIK---VAPPPQPSMTLSR</td>
<td>757</td>
</tr>
<tr>
<td>DIR protein</td>
<td>--GDCCKVEGSKTS-SQYEDRDKVQAGMENALQNFVFVPNIRSKSNSSDSDGAEAAALDNAIPNGATPS</td>
<td>1077</td>
</tr>
<tr>
<td>HIR protein</td>
<td>-EYEDSAEKCSCPKTDSSQILEKEESSFRKT-FEDYLVNIVALPV---TSSTGMTAEDPRPS</td>
<td>758</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>--ADQCRRNYCAPSKSSGSGTGTIVDTKENGITTEDQHNYVYIK---PLLRDKTTRKR</td>
<td>734</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>GVKMKNGVCACKEAKPSLRFSTRAEERDVSINFENELQNQVYRKS---VNPKINGHSIDTS</td>
<td>803</td>
</tr>
</tbody>
</table>
Majority

**O. nigriventer** insulin receptor (OnInR), **Drosophila** insulin receptor (DIR), human insulin receptor (HIR), **A. aegypti** insulin receptor (AaeInR), **B. mori** insulin receptor (Bom InR).

Insulin receptor protein alignment of the Fibronectin type III domains. Abbreviations: **O. nigriventer** insulin receptor (OnInR), **Drosophila** insulin receptor (DIR), human insulin receptor (HIR), **A. aegypti** insulin receptor (AaeInR), **B. mori** insulin receptor (Bom InR).
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>OnInR protein</td>
<td>WXIIILG--GLXXLFLXVGVXXYLRKKRPXXEXRMXXSVNPEYAS-----------------------------</td>
<td>1120</td>
</tr>
<tr>
<td>WXIIILG--IIILVLIIVGIVYLFFKRRQOEESPRMNPSINPEYM-----------------------------</td>
<td>1099</td>
<td></td>
</tr>
<tr>
<td>DIR protein</td>
<td>FFWLILG--IGLAFILSVLSFVGYCLHRRKVPNSD--LHMNETVNPFYAS------------------------</td>
<td>1379</td>
</tr>
<tr>
<td>HIR protein</td>
<td>AKIIGPLIVFLFSSVWGISLGFLRKRQPDGPLPGLYASNPYLELASSDVFPCSVYVPDEWESREKITTLLLRELQGSF</td>
<td>1034</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>TLIWL--TLIVIFLCSVGFVAFYWYKYRMKSRQIRMYPEVNPDYG-------------------------------</td>
<td>1014</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>WAWVG--AGGVLAVALLVGAWSYARRGLMSHAASKLFSASNPEYVPT----------------------------</td>
<td>1093</td>
</tr>
<tr>
<td>Majority</td>
<td>GMVYEGIXDIXKGAETRCAIKTVNEATDRERIXFLEASVMKAFDTFHVRRLLGVVSGQPTLTVMLMANGDLKSY</td>
<td>1088</td>
</tr>
<tr>
<td>OnInR protein</td>
<td>GMVYEGIVRDIK--GKATIRCAIKTVNEATNAERLNLKESVMKAFDTAVHIKLGVVSGQPTLTVMLMANGDLKSY</td>
<td>1459</td>
</tr>
<tr>
<td>HIR protein</td>
<td>GMVYEGNRDIKGAETRCAIKTVNEATDRENFLSASVMKAFDTYHVRRLLGVCSRGQPALVMELMKGDGLKSY</td>
<td>1114</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>GMVYKGILTQLRGEKCNQPCAIKTVNESATAREKDSFLASVMKQFNYHVRRLLGVVSGQPTLTVMLMANGDLKSY</td>
<td>1094</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>GMVYEGIAKEIVKGPETRCAIKTVNEATDRENFLSASVMKAFDTFHVRRLLGVVSGQPTLTVMLMANGDLKTY</td>
<td>1173</td>
</tr>
<tr>
<td>Majority</td>
<td>LRSHRPDAEG--------NXPGSPX--QPPTLXRQMAIEADGMAYLASKKFVHRDLAARNCMVADDLTIVQDFGMRTR</td>
<td>1159</td>
</tr>
<tr>
<td>OnInR protein</td>
<td>LRLHRPDSG--------YTHGIIG--QPPLSRLRILMAIEADGMAYLASKKFVHRDLAARNCMVADDLTIVQDFGMRTR</td>
<td>1182</td>
</tr>
<tr>
<td>HIR protein</td>
<td>LRAHRPEERDEAMTYLNHGTVNQPTYGRYQMAIEADGMAYLASKKFVHRDLAARNCMVADDLTIVQDFGMRTR</td>
<td>1164</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>LRRHRPDYPEN--------GEDPSP--QPPTLROIQMAIEADGMAYLASKKFVHRDLAARNCMVADDLTIVQDFGMRTR</td>
<td>1244</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>LRRMRPDAEG--------SLPSSPP--VPPTLQNILMAIEADGMAYLASKKFVHRDLAARNCMVADDLTIVQDFGMRTR</td>
<td>1239</td>
</tr>
<tr>
<td>Majority</td>
<td>DIYETDYYRKGTGKLLPVRRWAPELSKDGVSDDWVSGVVLWEMATLASKPYQLSNQVLYVIDGGMERPCP</td>
<td>1619</td>
</tr>
<tr>
<td>OnInR protein</td>
<td>DIYETDYYRKGTGKLLPVRRWAPELSKDGVSDDWVSGVVLWEMATLASKPYQLSNQVLYVIDGGMERPCP</td>
<td>1619</td>
</tr>
<tr>
<td>DIR protein</td>
<td>DIYETDYYRKGTGKLLPVRRWAPELSKDGVSDDWVSGVVLWEMATLASKPYQLSNQVLYVIDGGMERPCP</td>
<td>1619</td>
</tr>
<tr>
<td>HIR protein</td>
<td>DIYETDYYRKGTGKLLPVRRWAPELSKDGVSDDWVSGVVLWEMATLASKPYQLSNQVLYVIDGGMERPCP</td>
<td>1619</td>
</tr>
<tr>
<td>AaeInR.pro</td>
<td>DIYETDYYRKGTGKLLPVRRWAPELSKDGVSDDWVSGVVLWEMATLASKPYQLSNQVLYVIDGGMERPCP</td>
<td>1619</td>
</tr>
<tr>
<td>Bom InR.pro</td>
<td>DIYETDYYRKGTGKLLPVRRWAPELSKDGVSDDWVSGVVLWEMATLASKPYQLSNQVLYVIDGGMERPCP</td>
<td>1619</td>
</tr>
</tbody>
</table>
Insulin receptor protein alignment of the transmembrane, juxtamembrane, and tyrosine kinase domains. Abbreviations: *O. nirgiventer* insulin receptor (OnInR), *Drosophila* insulin receptor (DIR), human insulin receptor (HIR), *A. aegypti* insulin receptor (AaeInR), *B. mori* insulin receptor (Bom InR).
APPENDIX D
Expression of OnInR mRNA relative to 28S in the horn tissues of small males and large females. Standard errors, standard errors of the mean and p values are indicated for each trial. Levels of OnInR in only one comparison were significantly different (B).