INTEGRATED STRATEGIES FOR INVESTIGATING ENDOCRINE MECHANISMS IN *BIOMPHALARIA GLABRATA* AS A TEST ORGANISM FOR ANDROGENIC CHEMICAL TESTING

A thesis submitted for the degree of Doctor of Philosophy

By

Satwant Kaur

Institute for the Environment, Health and Societies

Brunel University

March 2015

DECLARATION

The work submitted in this thesis was conducted between 2010 and 2014 at Brunel University (Uxbridge, West London, UK). This work was carried out independently and has not been submitted for any other degree.

Abstract

Endocrine and metabolic disease or dysfunctions are of growing concern in modern societies across the globe, underlining the need for continued focus on the development of pharmaceuticals. Subsequent scientific research has revealed a trend in the increase of such abnormalities and expansion of chemical industries, highlighting concerns that these disorders may, in part, be caused by exposure to environmental pollutants. This has led to changes in legislation concerning chemicals safety testing involving an increasing number of vertebrate animal tests as a part of environmental risk assessment process, at significant financial and ethical costs. A solution that is appropriate and aligned with the three R's (reduction, refinement and replacement) in relation to animal research is to exploit the use of small invertebrate organisms as possible replacements for mammals. In line with the above approach/solution, this thesis is based on the null hypothesis that common genes, proteins and processes in gastropod molluscs and humans underlie the response of male reproductive organs to androgenic chemicals. Using a freshwater pulmonate snail, *Biomphalaria glabrata*, physiological effects of two steroid androgens on the development of mollusc secondary sexual organs were studied. Furthermore, an exhaustive investigation on the mollusc nuclear receptor repertoire and reproductive type neuropeptides was conducted. This also included the study of the evolutionary degree of conservation of these genes in non-model molluscs. The results obtained suggest that the snails did not respond to, and were not affected by exposure to the androgens. These results were supported by the absence of the members of subfamily 3C of nuclear receptors, which includes some of the "vertebrate" steroid hormone targets, suggesting that this mollusc may be an inappropriate model for steroid hormone mediated mammalian endocrine function. The nuclear receptor (NR) repertoire of *B. glabrata* comprised of 39 nuclear receptors representing all the known subfamilies of the NR superfamily. 21 reproductive type neuropeptide genes were identified encoding precursors that are predicted to release over 124 bioactive cleavage products. The consequence of these findings is significant in the context of the development of alternative model organisms for chemical testing as well as elucidating the taxonomic scope of nuclear receptor mediated endocrine disruption.

Acknowledgement

Table of contents

3   CHAPTER 3: EXPOSURE OF THE FRESHWATER SNAIL *BIOMPHALARIA GLABRATA* TO STEROID ANDROGENS DURING DEVELOPMENT DOES NOT AFFECT REPRODUCTIVE PHYSIOLOGY

4   CHAPTER 4: INVESTIGATION OF NUCLEAR RECEPTORS IN *BIOMPHALARIA GLABRATA* AND *LOTTIA GIGANTEA*

List of figures

List of Tables

1 Chapter 1: General Introduction

1.1 Endocrine disruption and animal testing

Endocrine and metabolic diseases are among the most contemporary human afflictions, the prevalence of which has been well defined in large population-based studies (Golden *et al.*, 2009; Ford *et al.*, 2008). To name a few these, neurodevelopmental disorders, diabetes, obesity and birth defects have been reported in pediatric population and cancer, obesity, diabetes in adult population (Bloom, Cohen and Freeman, 2010; Pleis JR, Ward BW, Lucas JW, 2010). A snapshot from different studies has been presented in the Table 1.1. The prevalence and incidence of certain disorder has been well defined in large population based studies (Golden *et al.*, 2009). The inability to effectively reduce the incidence of these diseases and disorders stems from the failure to completely understand the mechanisms underpinning these diseases. Scientific research on the relationship between the environment and health as well as a recently released World Health Organisation (WHO) 2012 report has revealed that a larger portion of these disorders in humans may be caused by exposure to environmental pollutants than previously thought possible (Prins, 2008; Grandjean *et al.*, 2008; Heindel, 2006). There is the suspicion that exposure to endocrine disrupting chemicals (EDCs) in the environment may be partially to blame (WHO report, 2012) (http://www.unep.org/).

Table 1.1: Snapshot of the effects of endocrine disruptors observed in wild life.

Contaminant	Gender	Observation	Reference
Diethylbesterol	Male	Decreased testosterone level	(Yamamoto *et al.*, 2005)
	Female	Advanced development of primary and secondary follicles in ovary	
Tribuyltin	Male	Reduced no of sertoli cells in fetal testis	(Kishta *et al.*, 2007)
	Female	Reduced number of germ cells in fetal ovary	
Vinclozolin	Male	Apoptosis in germ cells in testis	(Cowin *et al.*, 2010)
Lead	Female	Delay in puberty	(Wu, Buck and Mendola, 2003)
Bisphenol A	Male	Chronic hyperinsulinemia	(Alonso-Magdalena P, Vieira E, Soriano S, Menes L, Burks D, Quesada I, Nadal A, 2010)

According to WHO (2012) the definition of an endocrine disruptor is as follows:

'An exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism or its progeny or (sub) population'.

Typically, EDCs are natural or synthetic compounds present in food we eat, water we drink and air we breathe (Diamanti-Kandarakis *et al.*, 2009). Modern societies produce thousands of chemicals that have the potential to interfere with the endocrine system.

These chemicals come from different sources such as energy production, agriculture, manufacturing, mining, transportation and health care (http://deainfo.nci/nih.gov/). The widespread presence of EDCs in the environment has led to proposed changes in legislation concerning the safety testing of chemicals. As a consequence, more rigorous testing of potential chemicals and drugs is required which will inevitably lead to significantly more vertebrate animals being sacrificed. Available figures for usage of rodents in reproductive toxicity tests in the UK and Europe state that around 210,000 rodents/year are used in chemical and drug testing for reproductive toxicity and that this number is rising (http://ec.europa.eu/environment/). Thus it is timely to find more ethically acceptable alternative methods for testing endocrine disrupting chemicals and this is the focus of this thesis.

The use of animals in experiments has been a long standing concern within society dating back to 1865 when a government policy related to Prevention of Cruelty to Animals was framed. Nonetheless it was not until the second half of the 20th century, with increasing scientific knowledge, that it was established that at least all vertebrate animals, are sentient beings (capable of experiencing pain and distress). This led to a significant shift in community attitudes towards greater compassion for the welfare of animals. This led to the Prevention of Cruelty to Animals Act (1986) followed by the Prevention of Cruelty of Animals Regulations (1996) and the current Prevention of Cruelty of Animals Regulations (2008). These proposed changes are in response to people's increasing concern that animals should be treated humanely and the need to avoid causing unnecessary or unjustified suffering. The European Directive 2010/63/EU specifically mentions the need to use alternatives to living animals for testing chemicals and promotes the development of alternative methods. Moreover, the Organisation of Economic Co-operation and Development (OECD) has also committed itself to the implementation of the 3Rs-principle that stands for (*Replacement, Refinement and Reduction*). In 2005 all European Union member countries adopted a guidance document '*The Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*' that is central to animal welfare and emphasises a harmonised approach towards validation of new alternative test methods for regulatory purposes. In addition to the European and UK specific regulations, the US have a number of institutions who have guidelines on to how to work towards animal welfare and implementation of alternative to animal tests. These are namely the Institute for Laboratory Animal Research (ILAR) Guide for the Care and Use of Laboratory Animals 2010, and the ICLAS guiding principles (2010) are some of the many institutions that are working towards animal welfare and implementation of

alternatives to animal tests.

In order to design any test system for evaluating potential endocrine disruptors, it is important to understand how these chemicals alter the function of the endocrine system causing adverse health effects. A brief description of the endocrine system therefore follows:

1.2   The Vertebrate endocrine system

The endocrine system in vertebrates is a system of ductless glands and the chemical messengers (hormones) that they produce. This system functions through a series of cascade mechanisms, triggered by external or internal stimuli causing secretion of hormones into the blood that then travel to target organs binding to specific receptors leading to a coordinated physiological response (Hadley, 1992). These coordinated physiological responses are regulated by negative and positive feedback mechanisms to maintain homeostasis and prevent alterations in hormone secretions that could cause adverse effects.

The endocrine system cascade acts through a number of axes or sub-systems, e.g. the hypothalamus-pituitary-gonadal (HPG) axis consisting of the hypothalamus, pituitary and gonads (ovaries and testes), regulating reproductive and immune function; the hypothalamus-pituitary-thyroid (HPT) axis and hypothalamus-pituitary-adrenal (HPA) axis for maintaining energy homeostasis and energy maintenance respectively. The different glands within (and sometimes between) each axis, signal each other in sequence. During male sexual maturation, for example, (Figure 1.1) after genetic sex determination, genes located on the Y chromosome send a signal for sexual differentiation through neurotransmitters to the hypothalamus in the brain. In response the hypothalamus secretes the gonadotropin-releasing hormone (GnRH), which travels to the anterior pituitary (adenohypophysis), binds to the GnRH receptor and stimulates the secretion of gonadotropins such as follicle stimulating hormone (FSH) and Luteinizing hormone (LH). FSH and LH travel to the testis, where they stimulate the production and secretion of testosterone from the Leydig cells in the testis (Dorner, 1978). In the testis LH binds to the receptors on the Leydig cells and stimulates the synthesis of testosterone whilst FSH binds to the receptors on the Sertoli cells and release and metabolise factors necessary for spermatogenesis (Campbell and Reece, 2005). Testosterone is responsible for differentiation and development of the male urinogenital system. Testosterone then circulates back to the pituitary and hypothalamus and this leads to a reduction in the

secretion of GnRH (negative feedback). The testis then undergoes quiescence until puberty when they are once again activated by GnRH secretion from the pituitary. The hypothalamus is the neural control center for most of the endocrine systems of glands and hormones involved in physiological and behavioural responses that are necessary to survive, through the anterior pituitary gland. The hypothalamic neurosecretory cells produce hormones (some releasing and some inhibitory) that reach the anterior pituitary gland via hypothalamo-hypophyseal system to regulate its hormone (tropic hormone) production. These tropic hormones regulate the endocrine activities of the thyroid gland (secreting thyroid hormone), the adrenal cortex (secreting corticosteroids), gonads (secreting reproductive steroids) and the liver (secreting growth factors). This control system acting through hypothalamus does not directly control a number of endocrine organs like the adrenal medulla, the parathyroid glands, the pancreas, the thymus, the gastrointestinal tract, pineal gland and the kidney. The feedback mechanism in these glands is controlled by signals of the regulated variable they control for example the blood glucose levels control the secretion and inhibition of insulin and glucagon from the pancreas.
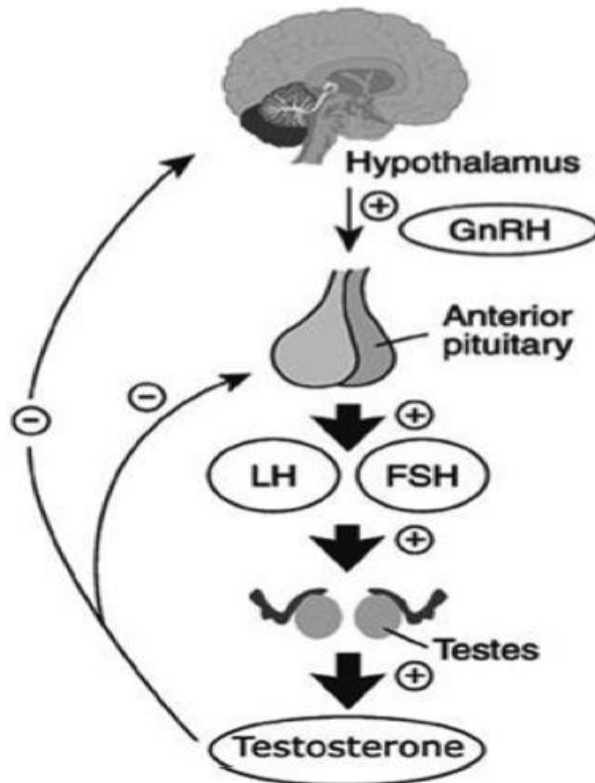
Figure 1.1: Schematic diagram of the vertebrate endocrine system showing feedback mechanism of hypothalamus-pituitary-gonadal axis.
The positive feedback mechanism stimulates the secretion of Gonadotropic-releasing hormone (GnRH) from hypothalamus to anterior lobe of pituitary gland that produces and secretes gonadotropins (LH and FSH). These gonadotropins travel to testes and enhance the production of testosterone. The negative feedback mechanism is inhibitory to reduce the secretion of GnRH and thus maintain homeostasis (Taken from http://pubs.niaaa.nih.gov/publications/arh25-4/282-287.htm*l)*

1.2.1 Endocrine glands

As mentioned above endocrine glands are ductless glands, they are highly vascularised and as secretory cells they are located close to the blood vessels to allow rapid transmission of their secretions. These glands have specific biological organisation to suit their function like the peptide secreting cells have well-developed rough endoplasmic reticulum rich in protein filled storage granules and the mitochondria has flat cristae while steroid secreting cells have more of smooth endoplasmic reticulum rich in lipid droplets containing cholesterol (precursor of steroids) and the mitochondria has tubular cristae (Norris and Carr, 2013). The following Table 1.2 shows the various vertebrate endocrine glands and their hormones.

Gland	Hormones	Target organs	Chemical class
Hypothalamus	Releasing and inhibiting hormones	Anterior Pituitary	Peptides
Pituitary Anterior	Trophic hormones	Endocrine glands	Peptides
Pituitary Posterior	Antidiuretic hormones	Kidney and blood vessels	Peptides
	Oxytocin	Uterus and mammary glands	
Gonads Testes Ovary	Androgen	Prostate, seminal vesicles and other organs	Steroid
	Estrogen	Female reproductive gland and mammary glands	Steroid
	Progesterone		Steroid
Thyroid	Thyroxine	Most organs	Steroid
	Triidothyronine		Amine
	Calcitonin		Peptide
Pancreas	Insulin	Most organs	Protein
	Glucagon	Liver and adipose tissue	Protein
Adrenal Cortex	Glucocorticoid Mineralocorticoid	Liver and muscles Kidneys	Steroid
Medulla	Epinephrine Norepinephrine	Heart, bronchioles	Amine
Stomach	Gastrin	Stomach	Peptide
Heart	Atrial natriuretic hormone	Kidney	Peptide
Pineal	Melatonin	Hypothalamus and Anterior Pituitary	Amine
Thymus	Thymopoietin	Lymph nodes	Peptide
Parathyroid glands	Parathyroid hormone	Bone, Small intestine and Kidney	Peptide
Liver	Somatomedins	Cartilage	Peptides
Small Intestine	Secretin and cholecystokinin	Stomach, Liver and pancreas	Peptides
Kidney	Eythropoietin	Bone marrow	Peptides

Table 1.2: Vertebrate endocrine glands, target glands, their hormone and chemical nature of the hormones.
(Adapted from (Adkins-Regan, 2007).

1.2.2 Hormones

A hormone is a biochemical molecule produced by an endocrine gland that travels via blood to distant cells and tissues where it produces irreversible effects (Melmed *et al.*, 2011). Hormones regulate a variety of physiological processes in the body from cell differentiation during embryonic development to the control of organ functions in adulthood and behavioural activities such as responding in appropriate ways to other individuals (social behaviour), that varies with age, between and with the sexes (Adkins-Regan, 2007). A single hormone can affect more than one function and several hormones

control each function. Vertebrate hormones fall into three chemical classes:

1. Amino acid derivatives are the derivatives of tyrosine such as catecholamines (epinephrine or norepinephrine) secreted by adrenal medulla and thyroxine secreted by the thyroid gland.

2. Peptide hormones are glycoproteins secreted from anterior pituitary gland such as thyroid stimulating hormone and luteinizing hormone. These also include gastrointestinal tract hormones and pancreatic hormones such as insulin.

3. Lipid derivatives are cholesterol derived steroid hormones like androgen, estrogen, and those secreted by adrenal cortex. These are synthesised on demand and have the same basic ring structure but their side chain and spatial orientation renders them specific. Another class of lipid derivatives includes the major classes of eicosanoids including prostaglandins, prostacyclins, thromboxanes and leukotrienes (Nussey and Whitehead, 2001).

1.2.3 Hormone synthesis

Amino acid derivatives like catecholamines are hormone regulators and are synthesised from amino acid phenylalanine or tyrosine by enzymatic reaction taking place in the cytosol. These are then stored in vesicles and are exocytosed as and when required. Being water soluble (hydrophilic) they do not require any transport proteins and can be easily transported through plasma.

The synthesis of peptide hormones starts with transcription of a single gene in the presence of RNA polymerase II and various transcription factors followed by modification of 3' and 5' ends of the mRNA (messenger RNA). This mRNA is used as a template that docks with the signal recognition particle on the surface of rough endoplasmic reticulum where protein synthesis takes place. The peptides in their inactive state (pro-hormone) are packed in secretory granules within the endoplasmic reticulum from where they pass into the Golgi apparatus and in the presence of peptidases undergo cleavage forming active hormones. Like catecholamines, these are also transported through plasma (Norris and Carr, 2013; Nussey and Whitehead, 2001).

Steroid hormones are synthesised through a biosynthetic pathway called steroidogenesis in the mitochondria and endoplasmic reticulum. The complete process of steroidogenesis involves a number of steps starting with signal transduction pathway, which causes the increased transport of cholesterol to mitochondria in Leydig cells. Cholesterol is the starting precursor molecule for all steroid hormones and all steroid hormones have similar basic structure. Enzymatic conversion of cholesterol into pregnenolone in the presence of

cytochrome P450 cholesterol side-chain cleavage enzyme (P450$_{SCC}$), initiates this biochemical pathway that culminates with the formation of hormones (Figure 1.2). For the purpose of illustration the pathway for the formation of androgens will be described. Pregnenolone is converted into progesterone by 3β-hydroxysteroid dehydrogenase. Progesterone is converted into 17α-hydroxyprogesterone (by 17α-hydroxylase), which is then converted into androstenedione (by 17, 20 lyase). 17β-hydroxysteroid dehydrogenase converts androstenedione into testosterone, which is then converted into dihydrotestosterone (DHT) by 5α-reductase (Hinson et al., 2007). DHT is significantly more potent androgen compared to testosterone and is an end-product hormone. The last enzyme of steroidogenic pathway is aromatase that converts testosterone into estradiol, an estrogen.
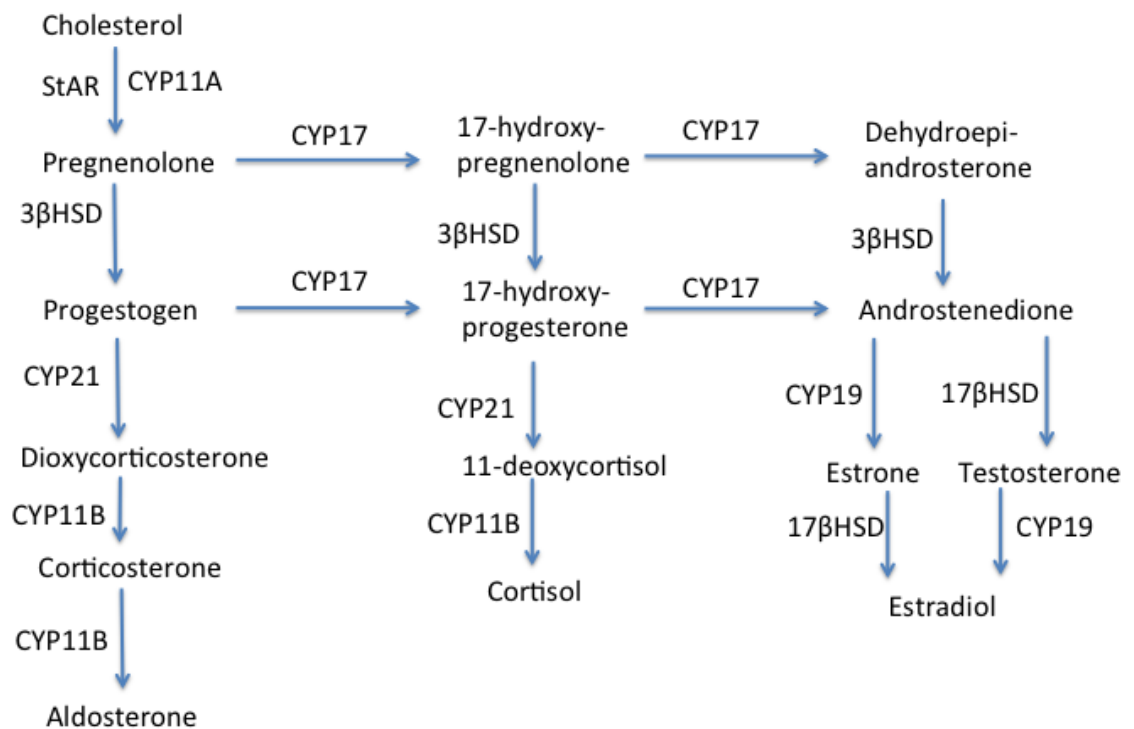
Figure 1.2: The steroid hormone synthesis pathway in vertebrates showing the enzymes and cytochrome P450 (CYP) genes participating in this process.
The enzymes are 3β-HSD: 3beta-Hydroxysteroid dehydrogenase; 17β-HSD: 17beta-Hydroxysteroid dehydrogenase. (Adapted from (Fretter, 1984).

1.2.4 Hormone action

Most hormones initiate a cellular response by combining with an intracellular or cell membrane associated receptor. Hormone receptors provide specificity to hormone actions, both in terms of the time and the place of hormone action. Hormones cannot act on all the cells but only those where the respective receptors are present and also many hormones do

not act throughout the life span of an organism but at specific time periods of life.

- Peptides are large water soluble molecules that do not readily cross the hydrophobic lipid cell membrane. For this reasons their receptors are membrane receptors lying extracellularly and they couple with the intracellular signal transducing molecules by transversing the cell membrane (Nussey and Whitehead, 2001). Most peptide receptors are members of G-protein coupled receptor family and these may either have a short extracellular amino terminal domain (For example, GnRH) or a long extracellular domain (For example thyroid stimulating hormone) (Darlison and Richter, 1999). Hormone-receptor interaction causes dissociation of associated G-protein, activation of membrane bound enzymes that stimulate the synthesis of cyclic AMP (cAMP) (second messenger), which triggers protein kinases (Cooper and Hausman, 2000). The transcription factor targets for kinases include c-jun and c-fos that form heterodimeric complex, which may be targeted by mitogen activated protein kinase (MAPK) or cAMP response element binding protein (CREB) that initiate transcription (Nussey and Whitehead, 2001). A GTP-binding membrane protein mediates between the peptide receptor protein and the enzyme for the synthesis of AMP, adenylate cyclase. However, some peptides like vasopressin have more than one receptor subtype (Goodson and Bass, 2001). There is evidence that steroids, acting through intracellular receptors stimulate the transcription of membrane receptor genes thus upregulating peptide action (Norris, Madsen and Freestone, 1996). In addition, peptides and their receptors are key players in the downstream pathway for steroid action on brain and behaviour (Ferguson, Young and Insel, 2002; Albers and Bamshad, 1999).

- Unlike peptides, steroid and thyroid hormones are lipophilic and can easily diffuse across the cell membrane. Their receptors are intracellular proteins acting as transcription factors regulating the expression of target genes in target tissues. These receptors belong to the superfamily of nuclear receptors (NRs) that are classified based on their ability to form dimers and the sequence of DNA they bind (Mangelsdorf *et al.*, 1995). Not all the NRs are activated by hormones (also called ligands), for some NRs regulatory ligands are not known or might not exist and these could be regulated by dietary components. All these receptors have a similar structure with the highly conserved DNA binding domain and less conserved the ligand binding domain regions (discussed in detail in chapter 4) (Germain *et al.*, 2006). The hormone binds the NR in the ligand binding domain causing structural changes in the receptor that allows the DNA binding domain to bind with the

hormone response elements, localized in the 5' flanking region of the target genes and bring about transcription. In addition to the nuclear receptors, there are also membrane receptors for some steroid hormones like estrogens and progestins. The membrane receptors for these receptors are coupled with fast acting pathways that result in rapid effects, unlike NRs, which take several hours to stimulate production of protein and exert action.

1.3 Endocrine disruptor action

Endocrine disrupting chemicals (EDCs) mimicking the endogenous hormones can act to disrupt physiology by interfering with the synthesis, secretion, transport, binding, action or elimination of hormones in the body so changing homeostasis, reproduction or development and/or behaviour (Vos *et al.*, 2000; Tyler, Jobling and Sumpter, 1998). These chemicals have been shown to affect the hormone delivery to the receptor, have effect on hormone receptors and alter their functions (Diamanti-Kandarakis *et al.*, 2009). Some EDCs may have mixed properties or they may metabolise to yield substance having different properties. For example estrogen agonist dichlorodiphenyltrichloroethane (DDT) metabolises into dichlorodiphenyldichloroethylene (DDE), which is androgen antagonist (Diamanti-Kandarakis *et al.*, 2009; Sohoni and Sumpter, 1998). These conversions can have significant biological effects as human reproduction (male and female) is an interplay of both androgens and estrogens. Along with direct effects on androgens and estrogens they can also interfere in the negative and positive feedback mechanisms of the different axes in vertebrates (HPG, HPT and HPA) (Kester *et al.*, 2002).

Among the first compounds to be identified as an EDC is diethylstilbestrol (DES). This was specifically designed drug to help prevent miscarriage and restore hormonal balance in pregnant women (Bamigboye and Morris, 2003; Smith, 1949; Smith and Smith, 1949). However in 1971, Herbst and colleagues found that *in utero* exposure to DES is associated with vaginal adenocarcinoma in female offspring (Herbst, Ulfelder and Poskanzer, 1971). This study was followed by reports on DES induced cryptorchidism in male offspring (Newbold, 2004; Lassurguere *et al.*, 2003; Skakkebaek, Rajpert-De Meyts and Main, 2001; Giusti, Iwamoto and Hatch, 1995). Transgenerational studies using DES suggest that in utero exposure to DES alters the normal programming of genes that play an important role in development and differentiation of reproductive tract, which increased risk of reproductive tract anomalies and dysfunction in their sons and daughters (Titus-Ernstoff *et al.*, 2010; Palmer *et al.*, 2005; Schrager and Potter, 2004; Kim *et al.*, 2002) and

have been associated with 80% increased risk of endometriosis (Missmer *et al.*, 2004). Subsequent research confirmed the transplacental mechanism of DES transmission (Kerlin, 2005; Maydl *et al.*, 1983) and classified DES as a mutagen (Stopper, Schmitt and Kobras, 2005), teratogen and carcinogen (Mittendorf, 1995). Animal studies using prenatanally DES exposed outbred mice showed numerous transgenerational effects including uterine fibroids (Hoover *et al.*, 2011); ovarian cancer (Blatt *et al.*, 2003); increased menstrual irregularities (Titus-Ernstoff *et al.*, 2006) and others. One study by Viglietti-Panzica et al. (2005) investigating sexual behaviour of Japanese quail concluded that embryonic exposure to DES caused sex reversal of behavioural phenotype and a significant decrease in vasotocin in the males (Viglietti-Panzica *et al.*, 2005). In another recent study using Balb/c pregnant mice it was demonstrated that 67μg/kg body weight of DES injected on gestation day 13 induced symptoms of cervical carcinogenesis in the F1 generation (Zulfahmi *et al.*, 2013). Even though this medicine was banned for use during pregnancy in the 1980s, animal studies are being performed to understand the mechanism of action. In a review Newbold (2012) suggested that DES exposed population should be continuously monitored as they age and the mechanism of DES induced toxicities would be useful in predicting effects of other environmental estrogens (Newbold, 2012).

Another example of cause and effect of EDCs is the case of the skewed sex ratio of fish in the rivers receiving effluent from paper mills effluents and sewage treatment plants which has been reported consistently across geographical regions and habitats and different species (*Lepomis megalotis, Catostomus catostomus, Coregonus clupeaformis, Perca fluviatilis, Gambusia affinis; Pimephales promelas*) (Hinck *et al.*, 2009; Blazer *et al.*, 2007; Fentress *et al.*, 2006; Orlando *et al.*, 2004; Jobling *et al.*, 1998; Munkittrick *et al.*, 1998; Gagnon *et al.*, 1994; Munkittrick *et al.*, 1991). Masculinisation of fish was reported among the embryos of viviparous eelpout *Zoarces viviparous* in the vicinity of a paper mill in Swedish Baltic coast (Larsson and Forlin, 2002). During the period of mill shutdown for 17 days, which coincided with the duration of embryonic gonad differentiation, sex ratios returned to normal, and then subsequently shifted towards males again when the mill was reactivated the following year, providing evidence in favour of chemicals in the effluent causing the shift in sex ratio. Compounds present in pulp/Kraft mill effluents have been shown to bind to androgen receptor in both humans and fish (Hewitt *et al.*, 2002; Parks *et al.*, 2001) and also bind to the retinol receptors, RXR and RAR (Alsop *et al.*, 2003).

In contrast, Lange et al. (2011) found sex ratios of roach deviating from 1:1 ratio in uncontaminated environment to 98% feminised phenotype, when exposed to wastewater

effluent from one month post hatch to 3.5 years of age (Geraudie *et al.*, 2011; Lange *et al.*, 2011). Long-term studies in a lake in northwestern Ontario, Canada, indicated that exposure to the potent estrogen, ethinylestradiol, at 4–6 ng/L over a period of 3 years resulted in the collapse of the population of the fathead minnow (*Pimephales promelas*) (Harris *et al.*, 2011; Kidd *et al.*, 2007). A strong correlation between predicted steroidal estrogens in the UK rivers and feminisation in the wild fish was demonstrated by Jobling et al. (2006) (Jobling *et al.*, 2006). Statistical modeling of different steroids in the UK wastewater also provided a strong argument of multicausal etiology of the widespread feminisation of fish (Jobling *et al.*, 2009).

There are many more instances of endocrine disruption having huge effects like the population-level consequences of Tributyltin (TBT) on molluscs (will be discussed in later chapters) and the story goes on but the important issue to be tackled is the increasing number of chemicals that are being introduced in the environment through varied sources and the increased number of pre-clinical tests to meet safety requirements.

1.4   Endocrine Disruptor Animal Testing Programmes

The increased awareness that EDCs might play a role in the increasing prevalence of reproductive endocrine diseases has led to stringent laws for chemical testing being devised both by the US Environmental Protection Agency (EPA) and by the OECD. Currently these test protocols focus mainly on interactions of chemicals with the estrogen, androgen and thyroid systems. The Endocrine Disruptor Screening and Testing Advisory Committee (EDSAT, 1998) stated in their final report, that as a part of the protocol, every chemical has to undergo a two-tiered screening process firstly consisting of *in vitro* and *in vivo* assays to identify and classify substances relative to their potential interaction with the endocrine hormone (androgen, estrogen and thyroid) system (Tier 1) Table 1.3. Tier 2 tests focused on developing a dose-response relationship in animal models (http://www.epa.gov/). The screening process is a battery of tests designed to ensure estrogen and androgen mediated effects through various modes of action (receptor binding and steroidogenesis pathway) and also feedback mechanisms of the HPG- and HPT-axis are detected. The battery is comprised of a combination of *in vitro* and *in vivo* assays that complement each other and help identify specific endocrine mechanisms of chemicals and relate them to physiological effects observed at higher levels of biological organisation. If a chemical is identified as a potential endocrine disruptor, it is further screened using Tier 2 battery of processes. The main purpose of Tier 2 is to provide a comprehensive profile

of the effects the chemical could induce via different modes/mechanisms of action.

Table 1.3: EDSAT Tier 1 Battery tests to detect hormone-mediated effects on rodents.
Pubertal male: male rats; pubertal female: female rats; HPG-axis: Hypothalamus-Pituitary-Gonadal axis; HPT-axis: Hypothalamus-Pituitary- Thyroidal axis. Adapted from Hecker and Hollert, 2011.

Screening Assays	Modes of action							
	Receptor Binding				Steroidogenesis		HPG-Axis	HPT-Axis
	Estrogen	Anti-estrogen	Androgen	Anti-androgen	Estrogen production	Androgen production		
In Vitro								
ER Binding	×	×						
ERα Transcriptional activation	×							
AR Binding			×	×				
Steroidogenesis H295R					×	×		
Aromatase Recombinant					×			
In Vivo								
Uterotrophic	×							
Hershberger			×	×		×		
Pubertal Male			×	×		×	×	×
Pubertal Female	×	×			×		×	×
Amphibian Metamorphosis								×
Fish short-term Reproduction (Male and Female)	×	×	×	×	×	×	×	

In short Tier1 should inform Tier2 in terms of dose range, life stage of exposure and most appropriate organism in which adverse effects are likely to occur. Tier 2 tests are long-term animal studies that could even be transgenerational with a broad range of taxa (Hecker and Hollert, 2011). In addition to the US EPA activities in the same year, the OECD also initiated an Endocrine Disrupter Testing and Assessment (EDTA) process. The EDTA has five levels of organisation that are not to be used as a tiered testing strategy but rather as a conceptual framework with which to provide information regarding the hazards of a chemical with regards to its potential endocrine disrupting properties Table 1.4. In 2007 the European community released a summary scientific report for prioritising chemicals for their endocrine disrupting effects. Of the 575 chemicals investigated, 320 substances were identified as potential EDCs (European Commission, 2007). The European Commission has also co-operated with the WHO (WHO, 2012). Besides these, REACH (Registration, Evaluation, Authorisation, and restriction of CHemical substances) is another European community regulation on chemicals and their safe use and came into force on 7th June 2007. REACH allows identification of intrinsic properties of chemicals such as carcinogenicity, mutagenicity, reproductive toxicity and their persistence and potential for bioaccumulation. The testing requirements for chemicals are according to their supply in tonnage. The tests are those included in OECD conceptual framework (Table 1.4). Taken together, the available information suggests that national and international governments are in the process of establishing testing programs and strategies to assess the risk associated with endocrine

disrupting chemicals.

Table 1.4: OECD (2012) conceptual framework for testing and assessment of endocrine disruptors.
Adapted from OECD (2012).

Level 1	Existing data/ literature
Level 2	*In vitro* assays providing data about selected endocrine mechanisms/pathways
Level 3	In vivo assays providing data about selected endocrine mechanisms/pathways
Level 4	*In vivo* assays providing data on adverse effects on endocrine relevant endpoints
Level 5	*In vivo* assays providing more comprehensive data on adverse effects on endocrine relevant endpoints encompassing different time-points in life-cycle of organisms.

One of the key issues remaining in context with the increasing number of testing programs across the globe to evaluate the likelihood of a chemical to disrupt endocrine functions is the reliance on a large number of *in vivo* tests using live animals. With the increasing testing demands the number of rodents, amphibians and fish that will be required will increase proportionally. It is difficult to estimate the number of animals being used globally for animal testing. In 1995, researchers at Tufts University, US Center for Animals and Public Policy had an estimate of 14–21 million animals being used in American laboratories in 1992, which was a reduction from 50 million used in 1970 (Andrew NR, 1995). According to the US Department of Agriculture (USDA) the total number of animals used in 2005 was almost 1.2 million (USDA report on the Enforcement of the Animal Welfare Act, 2005), excluding rodents, which make up about 90% of research animals. The British Union for the Abolition of Vivisection (BUAV) has estimated that 115 million vertebrates are experimented on around the world every year of which approximately 27 million of them are in European Union laboratories (http://www.buav.org). The US Department of Agriculture (USDA) has estimated over 950,000 animals were used in research in 2012 as compared to 1.1 million in 2010 (http://speakingofresearch.com/) of which 98% of the research was done using rodents, birds and fish. According to the UK Home Office report HC317 on animal testing (2011), just over 2.18 million animals were used in scientific experimentation in 2011. Sector-wise, the report shows an increase of 2% in the number of scientific studies for toxicological safety testing as compared to 2010; 78% of these procedures were for pharmaceutical procedures and 73% involved rodent species in testing. Of the total species used, mice contributed to 71% in these studies Figure 1.3 (https://www.gov.uk/).
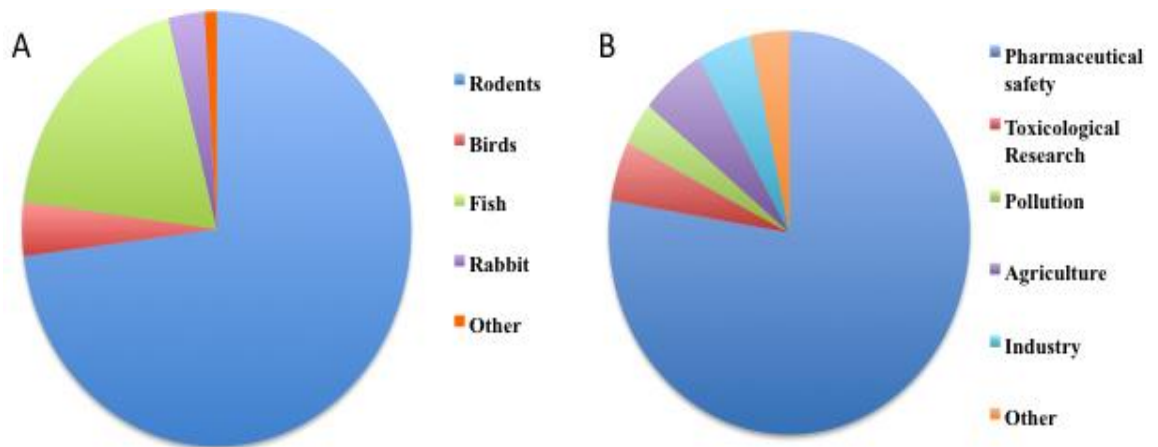
Figure 1.3: (A) Pie chart representing animal usage in different areas of research. (B) Pie chart showing percentage of animals used in toxicological research. Adapted from Home Office Report HC317(2011).

According to the Annual Statistics of Scientific Procedure on living animals, Great Britain (2012) this number went down to 2.13 million animals used in scientific experimentation in 2012 (excluding breeding of genetically modified and harmful mutants) (https://www.gov.uk/). The US EPA's EDSP (Endocrine Disruptor Screening Program) report highlights that subjecting a single chemical to the complete Tier 1 testing battery would require approximately 130 rats, 30 tadpoles/froglets and 60 fish (Hecker and Hollert, 2011). These estimates give us an overview of the large number of animals being used. None of these figures include animals bred in surplus for research studies that are then killed as surplus and or animals used for breeding purposes. Invertebrates like fruit flies, nematodes or molluscs are not part of these calculations, as they are not yet regulated. It is not only the huge number of animals used that should be the reason for concern but also the cost involved in these studies. Table 1.5 and Table 1.6 below shows the estimated total cost for conducting EDSAT Tier1 and Tier2 screening for a chemical which totals $922300. This does not include the feeding, housing, and other maintenance costs applied.

With the changes in the legislation concerning the safety testing of chemicals to incorporate endocrine disruption, increased numbers of vertebrate animal tests involving rats, mice, fish and frogs will grow proportionally (http://www.epa.gov/endo/pubs/edsat/) and almost 80,000 chemicals need testing (Krisberg, 2010).

Table 1.5: The estimated costs per chemical for Tier 1 screening, adapted from EDSAT final report.

The values listed are the mean of estimates for each assay obtained from a cost estimate survey conducted by Applied Pharmacology and Toxicology, Inc. in May 1998.

Name of Assay	Cost Required endpoints in mammalian assay
Rat estrogen receptor equilibrium exchange assay	$7,200.00
Rat androgen receptor equilibrium exchange assay	$7,300.00
Steroidogenesis assay in minced testes	$11,600.00
Uterotrophic Assay in Ovariectomized Rats	$29,600.00
The Hershberger Assay in Male Rats	$37,200.00
Pubertal Assay in Female Rat	$38,900.00
Fish Gonadal Recrudescence Assay	$37,800.00
Frog Metamorphosis Assay	$22,200.00
Subtotal (Assays only)	$191,800.00
Analytical costs	$30,000.00
Total for Tier 1 screening	$221,800.00

Table 1.6: The estimated costs per chemical for Tier 2 screening, adapted from EDSAT final report.

The values listed are the mean of estimates for each assay obtained from a cost estimate survey conducted by Applied Pharmacology and Toxicology, Inc. in May 1998.

Name of Test	Cost of test	Analytical cost
Avian Reproductive Toxicity Test	$172,100.00	$55,000.00
Fish Life Cycle Toxicity Test	$221,400.00	$80,000.00
Mysid Toxicity Test	$38,600.00	$23,000.00
Amphibian Reproductive / Developmental Toxicity Test	$75,400.00	$35,000.00
Subtotal	$507,500.00	$193,000.00
Total for Tier 2 screening	$700,500.00	

The use of alternative methods for testing endocrine disruptors could bring substantial economic savings in terms of time and resource. *In vitro* assays, high throughput screening assays and working with invertebrates with short life cycles is increasing the speed with which results can be obtained, and there is decreased cost of maintaining microorganisms/invertebrates compared with maintaining vertebrate animals in laboratories. A good example is the use of *in vitro* tests in high-throughput screening (HTS) for drug discovery in the pharmaceutical industry. The search for compounds with the ability to interact with particular molecular targets among millions of chemical

structures would be practically impossible without this technology. HTS technologies have found application in a wide range of important topics in biology and medicine, now allowing researchers to address biological questions using innovative approaches such as discovery and development of new chemical probes as research tools for use by the research community to enhance the understanding of biological functions and disease mechanisms (Soon, Hariharan and Snyder, 2013). One of the candidate drugs, FT1050 (16,16-dimethyl Prostaglandin E2), that influences adult stem cells was discovered through HTS on zebra fish. This drug has entered clinical trials for optimising transplantation of human hematopoietic stem cells from umbilical cord blood (North *et al.*, 2007). Newly developed gene sequencing methods have provided new perceptions into the global organisation of eukaryotic genomes that were previously unattainable. Analyses of individual regions in the genome have revealed the role of some regulatory factors in controlling the expression of genes such as transcription factors and non-coding RNAs (Eskiw *et al.*, 2010; Branco and Pombo, 2007; Fraser and Bickmore, 2007). This information has helped researchers to map genetic variants to gene regulatory regions and to analyse multiple pathways linked to disease (Boyle and Silver, 2012).

Both the governmental and non-governmental funding bodies do realise that animal research must continue for the well being of all communities (animal and plants) on the earth and that immediate abolition of animal testing is not possible. Nevertheless, more research is being focused on finding alternatives to vertebrate animal testing for endocrine disruptors in different areas of research such as biomedical science, endocrinology and drug development. My thesis describes one of the examples of many such projects being funded by the NC3Rs (National Centre for Replacement, Refinement and Reduction of animals in research).

1.5 The 3Rs (Replacement, Refinement and Reduction of animals in testing chemicals)

The 3Rs principle was first set out by William Russell and Rex Burch in 1959 as 'The Principles of Humane Experimental Technique'. Since then, the 3Rs have become established as an internationally accepted approach to research involving animals, and have been written into legislation in many countries (Guhad, 2005; Russell, Burch and Hume, 1959). Besides this the European Union Reference Laboratory for Alternatives to Animal Testing (ECVAM), Registration, Evaluation, Authorisation and restriction of CHemicals (REACH) and EU funded, European Phased Adaptive Approach (EPAA) are

also conducting research into alternative methods and their validation.

Reduction – includes methods that minimise the number of animals used per experiment or study including maximising the information obtained from a single animal. Examples include improved experimental design and statistical analysis, sharing data and resources within the scientific community and organisations by depositing experimental design and results in public repositories and the use of technologies, such as imaging to enable longitudinal studies in the same animal.

Refinement includes methods that minimise pain, suffering and distress or lasting harm experienced by the animals like surgical and other painful procedure. Pain, distress causes physiological changes in the animal and could lead to variability in results. Refinement reduces variability in experimental results and thus reduces the number of animal tests that need to be done. Experiments encompassing painful procedures or surgeries should not be carried out without proper sedation or anesthesia and if any procedure could lead to chronic pain then the animals should be humanely killed. Refinement also means using techniques and choice of animal that is best suited for the projects in terms of their behaviour, genetic constitution and general health.

Replacement includes the use of:

- Human volunteers, tissues and cells
- Non-testing approaches like mathematical and computer models
- Established animal cell lines, or cells and tissues taken from animals killed solely for this purpose (i.e. not having been subject to a regulated procedure)
- Immature forms of vertebrates, or invertebrates, such as *Drosophila* and nematode worms.
- Use of 'omics' technologies like transcriptomics, proteomics and metabolomics.

The research described in this thesis focuses on invertebrates as substitutes for vertebrate testing. One of the dysfunctions for which there seems at present to be no current invertebrate alternative to vertebrate animal testing is reproductive disorders and this is also a particular focus of this thesis.

1.6 Invertebrates as replacements for vertebrates in reproductive toxicity testing

Historically, invertebrates have made a major contribution to biomedical research (Wilber, 1964) because certain aspects of their physiology are very similar to that of mammals. There has always been a strong belief that organisms closest to humans in evolutionary

terms will provide the best 'model' for experiments on human biology and medicine. For example, a monkey is a better model for the human than is a rat, and a rat is better than a worm. However recent research has shown this is not necessarily the case as the very basic biological functions amongst all animals are quite conserved. For example, Pourquie (2003) reported that genetic machinery responsible for somitogenesis, a multistep patterning process forming mesoderm of the embryo, is conserved throughout the animal kingdom from rodents to protostomes through deuterostomes (Pourquié, 2003). With the growing wealth of sequenced genomes, computational comparative genomic tools and system biology approaches there is a potential to utilise the divergence of DNA sequences and physiology between species to find a suitable invertebrate model. Invertebrates raise less societal concern as compared to vertebrates and are not included in the list of protected animals under the UK Home Office Animals (Scientific procedures) Act 1986 (except *Octopus vulgaris*) (https://www.gov.uk). They generally have a short life span so developmental studies can yield faster results and they are easy and less expensive to culture and maintain in laboratories. Invertebrates have a simpler body plan as compared to vertebrates and this has been pivotal in understanding basic biological systems in vertebrates (Andrews, 2011). The ability of invertebrates to regenerate, as seen in echinoderms (Mashanov *et al.*, 2010; Carneveli, 2006) can be quite useful in optimising techniques such as tissue engineering. The primary advantage of designing assays using lower vertebrates/invertebrates is that the embryos develop rapidly and we can get the complete repertoire of their gene expression and molecular signaling from fertilization to organogenesis in short span of time. During this critical window of development, embryos can be monitored for a series of physiological or behavioural endpoints and there is a high probability of detecting an adverse interaction between the chemical and its molecular targets. In addition to the rapid developmental profile, invertebrate embryos are often transparent and develop externally (in contrast to mammalian *in utero* development), allowing for non-invasive microscopy techniques to observe them across different developmental stages. Moreover, using whole animals is quite useful when integrative endpoints are assessed, keeping in mind the complexity of an endocrine system functioning in a living organism, and especially that of a developing organism. For example in the area of reproductive toxicity, one of the problems is the lack of understanding of many integrated mechanisms of action leading to toxicity (EURL-ECVAM report, 2013) (Zuang *et al.*, 2013) and real time exposure studies in invertebrates could be an answer. Nevertheless, working with whole animals/embryos would also allow observing physiological processes without the knowledge of mechanisms, including

currently unknown mechanisms.

Drosophila melanogaster (fruit fly) is the best understood animal model for detecting mutagenicity (Brusick, 1980), teratogenicity (Bournias- Vardiabasis *et al.*, 1983) and female reproductive toxicity (Avanesian, Semnani and Jafari, 2009; Goss and Sabourin, 1985). The importance of *D. melanogaster* to research can also be estimated by the Nobel prizes awarded to Thomas H Morgan in 1933, for his discovery related to the role of chromosomes in heredity using *D. melanogaster* as model organism.

D. melanogaster is a poikilothermic, diploid outbreeding organism and adult flies have body parts divided into three regions (head, thorax with attached wings and abdomen) with distinct tissues and homologues of most human organs (Bier and McGinnis, 2004). However, *Drosophila*, like all insects, does not have a calcareous endoskeleton like mammals, but a chitinous exoskeleton. Further distinct features are the lack of an adaptive immune system, its circulatory system comparable to a mix of lymph and blood, the tracheal breathing system, and the structure of the nervous system, to name a few. *D. melanogaster* is bisexual with male and female flies and has a high rate of reproduction. It undergoes indirect development passing through distinct phases from egg to larva and pupa, reaching adulthood in a short time span of about 10 days under common laboratory conditions. The reproductive organs of female fly include a pair of ovaries, accessory glands, spermatheca, spermathecal gland, gonophore and vagina making it structurally similar in many regards to mammalian reproductive organs (King, Aggarwal and Aggarwal, 1968). The male reproductive anatomy comprises of testes, vas deferens, seminal vesicle, accessory gland and ejaculatory duct leading into the gonophore from where the sperm exits. There is a pair of testes comprised of testicular follicles and encased in epithelial tissue (nutritive layer) and all the follicles are encased in connective tissue. The accessory glands are responsible for secreting chemical infusion for stimulating spermatozoa. There is increasing evidence of conserved sexual development genes between *D. melanogaster* and mammals (Nanda *et al.*, 2009; Lalli and Sassone-Corsi, 2003; Carney and Bender, 2000; De Loof and Huybrechts, 1998) and there are studies suggesting *D. melanogaster* as an alternate animal model for quick screening of chemicals for their reproductive toxicity. Indeed, the *Drosophila* female reproductive system has been studied to get a better understanding of human ovarian cancer (Naora and Montell, 2005) and some studies have demonstrated ovarian impairment in female *Drosophila* when exposed to 0.12 and 0.24mg/ml of methotrexate (causes folic acid deficiency) (Avanesian, Semnani and Jafari, 2009; Affleck and Walker, 2007; Affleck *et al.*, 2006). Xenobiotics like dichlorvos and chlorpyrifos caused tissue damage in the

accessory gland of adult male flies in a dose dependent manner (Gupta *et al.*, 2007). However, the major limitation for using *Drosophila* as a model organism for male reproductive toxicity is the absence of mammalian molecular machinery, such as the estrogen receptor and androgen receptor as well as mammalian steroid hormones. A few studies have suggested ecdysone as a potential steroid hormone-mimicking mammalian steroid hormone (Rauschenbach *et al.*, 2000; De Loof and Huybrechts, 1998) but clear evidence is still lacking (Tiwari *et al.*, 2011).

Caenorhabditis elegans- The round worm (*Caenorhabditis elegans*) is a model organism popularised by Sydney Brenner in the field of developmental biology (Brenner, 1974). Since then *C. elegans* have been used to study neurodegeneration, signal transduction, apoptosis, ageing and RNAinterference (Teschendorf and Link, 2009; Leung *et al.*, 2008; Antoshechkin and Sternberg, 2007; Timmons *et al.*, 2003). Moreover Nobel Prizes have been awarded to three geneticists working with *C. elegans*. In 2002, the Nobel Prize in Medicine was awarded to Sydney Brenner, H. Robert Horvitz and John Sulston for their work on the genetics of organ development and programmed cell death in *C. elegans*. In 2006, Andrew Fire and Craig C. Mello were awarded the Nobel Prize for their discovery of RNA interference in *C. elegans*. The 2008 Nobel Prize in Chemistry was shared between Martin Chafie, Osamu Shimomura and Roger Y. Tsien for their work on green fluorescent protein in *C. elegans*. Furthermore, the intensively studied genome, complete cell lineage map, knockout mutant libraries and established genetic methodologies such as mutagenesis and RNAinterference to name a few, provide a variety of methods to manipulate and study *C. elegans* at the molecular level (Leung *et al.*, 2008).

Like *Drosophila, C. elegans* is also poikilothermic, can be easily cultured in the laboratory and has an even shorter life span of 3 weeks. Unlike *Drosophila* it is small in size (adult worm measuring 1mm) and has a transparent body. It is naturally a highly inbred organism of mostly self-fertilising hermaphrodites and occasional males with high reproductive output and a short time span of 3 days from egg to sexual maturation (Kaletta and Hengartner, 2006). The hermaphrodite germ line produces both oocytes and sperms. Sperms are produced by maturing germ cells only during L4 larval stage (approximately 37 hours post oviposition) while in the adult stage germ cells mature to oocytes (Zanetti *et al.*, 2012). Development is indirect passing through four distinct larval stages L1-L4 in three days at 20^0C. *C. elegans* as a model organism has its own advantages and disadvantages. The small size (adult is approximately 1mm long) of *C. elegans* is useful as *in vivo* assays can be done in 96-welled plate but is experimentally challenging for biochemical assay due to the amount of material in single worm. The transparent body of

the worm is useful for non-invasive microscopic examination of the internal anatomy and development but the thick outer cuticle forms a barrier for certain hydrophilic chemicals limiting its use to monitor chemical toxicity (Pandey and Nichols, 2011; Avanesian, Semnani and Jafari, 2009). A big disadvantage of *C. elegans* is its evolutionary distance from humans with less conserved genes, absence of steroid hormones as well as steroid hormone nuclear receptors and the lack of metabolising enzymes like cytochrome P450s. Furthermore, almost one third of the organism is comprised of neurons with less defined organs and tissues.

Other invertebrate models – As recently highlighted by Bolker (2012), very few invertebrate models have been investigated, and in order to find the best model organism for endocrine function we may need to look beyond the traditional non-vertebrate model organisms (flies, worms) (Bolker, 2012). Apart from *Drosophila* and *C. elegans* other invertebrates like the slime mould, *Dictyostelium* have been used in research on epilepsy and bipolar disorder (Williams *et al.*, 2006). In toxicology, effects such as mutation can be conveniently detected using several strains of the bacterium *Salmonella typhimurium* in the Ames test. This test is used to identify almost all types of chemicals for mutagenicity (Mortelmans and Zeiger, 2000). The sea urchin has long been the test organism for basic developmental biology (McClay, 2011; Semenova, Kiselyov and Semenov, 2006; Anderson, Hose and Knezovich, 1994). Apart from all the research done on these organisms, there are no satisfactory invertebrate models for male reproductive endocrinology, but as the basic concept of sexual reproduction is similar in most animal phyla, it could be hypothesised that regulation of reproduction must be highly conserved between invertebrates and vertebrates. Lower vertebrates like zebra fish were not used in this study even though it is well-established model organism to study drug-induced developmental toxicity. Like *C. elegans*, the embryos of this fish are transparent allowing non-invasive observation of developmental changes but the difference in reproductive morphology such as lack of prostate glands makes them unsuitable for preliminary testing for androgenic compounds.

1.7 Molluscs as proposed test systems for reproductive endocrine disrupter testing

Recent research has shown that the effects of androgenic and anti-androgenic chemicals on reproduction and development in crustaceans, gastropod molluscs and echinoderms seem to be comparable to those in mammals. Of the three phyla, molluscs are the most

promising group because the sex steroids in molluscs seem to be chemically similar to vertebrates (Matthiessen, 2000). Gastropod molluscs have long been used as models of neuronal
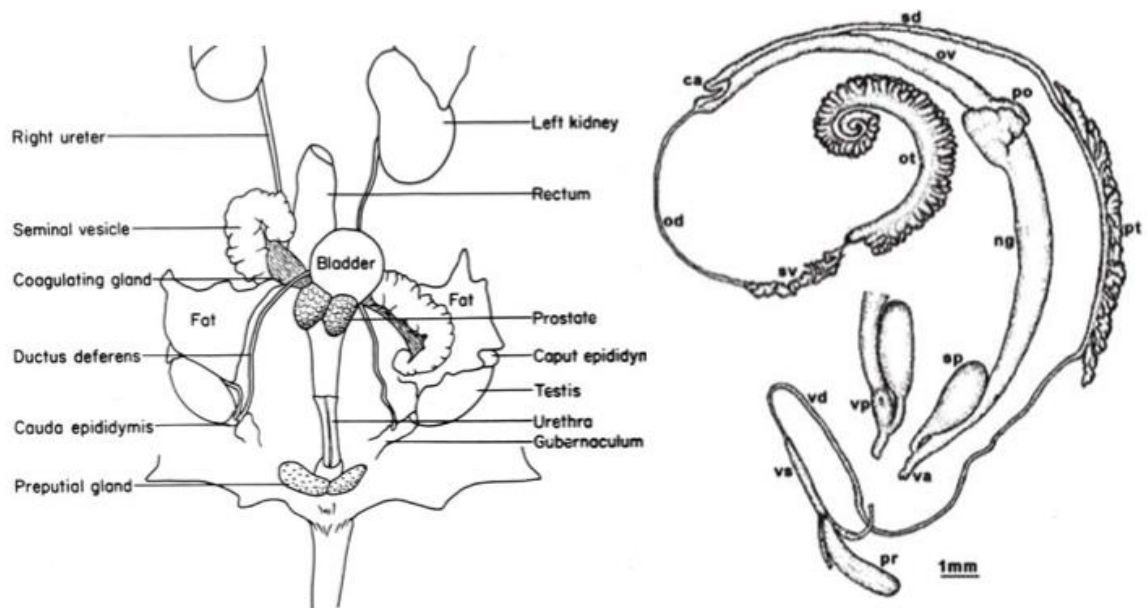
Figure 1.4: Comparison of anatomy of male mouse urogenital tract (A) with reproductive tract in gastropod snail (B).
ca: Carrefour; ng: Nidamental gland; od: Ovispermiduct; ot: Ovotestis; ov: Oviduct; po: Pouch of Oviduct; pr: Prepuce penis; pt: Prostate; sd: Spermiduct; sp: Spermatheca; sv: Seminal vesicle; va: Vagina; vd: Vasdeferens; vp: Vaginal pouch; vs: Vergic sac. Adapted from (www.informatics.jax.org) (A) with reproductive tract in gastropod snail (De Jong-Brink, 1969) (B)

mechanisms of learning and memory (Glanzman, 2006), including the work of Eric Kandel who won a Nobel Prize in 2000 for his studies using *Aplysia californica* (Californian sea slug). Molluscs have also been suggested as models of autism (Hatfield, 2005). Although gastropods have been used as models of neuronal mechanisms, its use as a surrogate for testing reproductive toxicity can be suggested owing to the similarities between mollusc and mammalian reproductive system. Gastropod molluscs, in particular have a reproductive anatomy that, at least superficially, is very similar to that of vertebrates. Like mammals, the snail male reproductive tract extends from the testis (or ovotestis in hermaphrodite snails), has seminal vesicles, and a prostate gland (shown to be involved in stimulating sexual drive) (Adiyodi and Adiyodi, 1988) Figure 1.4 and terminates in a penile glandular complex (around the preputium) containing sebaceous glands and a penis. The eversion of the penis is controlled by a retractor muscle (analogous to the levator ani and bulbocavernosus muscle in mammals) (discussed in more detail in Chapter 2).

There is evidence that the gonads of molluscs are able to synthesise estrogen and testosterone like molecules in the reproductive cycle (Lafont and Mathieu, 2007), known

to regulate vertebrate reproductive function (Bannister *et al.*, 2007). Preliminary studies have shown that a molluscan testosterone may have a role in male reproductive tract development, but there is no evidence of whether this activity is dependent on an androgen receptor (Sternberg, 2007). The presence of androgen receptors has been suggested but not yet reported   (Koehler *et al.*, 2007). Biosynthetic pathway of the vertebrate steroids (involved in reproduction) uses 3β-hydroxy-Δ$_5$-steroid dehydrogenase (3β-HSD) enzyme to convert pregnenolone to progesterone or dehydroepiandrsterone (DHEA) to androstenedione (Ad). 3β-HSD enzyme activity has been reported in different mollusc species (including, *Crassostrea gigas* (oyster), *Octopus vulgaris* (octopus) and *Pecten yessoensis* (scallop)) (Matsumoto *et al.*, 1997; D'aniello *et al.*, 1996; Krusch *et al.*, 1979; Mori, Muramatsu and Nakamura, 1969). The evidence for the existence of 17β-hydroxysteroid dehydrogenase that catalyse the interconversion of Ad and testosterone (T), estrone and 17β-estradiol in vertebrates is very strong in molluscs (Lyssimachou, Ramón and Porte, 2009; Lyssimachou, Bachmann and Porte, 2008; Janer *et al.*, 2005). Another enzyme important in metabolising T into 5α-dihydrotestosterone (DHT) in vertebrates has been identified in molluscs tissues (Lyssimachou, Bachmann and Porte, 2008; Janer *et al.*, 2005; Ronis and Mason, 1996). Developmental exposure of gastropods to the exogenous androgen, Testosterone (T) or to the environmental androgen tributyltin (TBT) caused hypertrophy of the prostate gland and inhibition of spermatogenesis in males (Wang and Croll, 2004). In the female these exposures led to development of a penis and transformation of the albumen gland into a prostate gland (Watermann *et al.*, 2008). This raises the possibility that the molluscan male reproductive tract is a functional analogue of the mammalian male reproductive tract. When exposed to mixture of androgens and anti-androgens the above mentioned effects seemed to be suppressed, indicating the effects to be hormonal dependent (Hagger *et al.*, 2009; Hagger *et al.*, 2006). Spooner *et al*. (1991) and Bettin *et al*. (1996) suggested that TBT could prevent the conversion of testosterone to E2 by inhibiting aromatase (Bettin, Oehlmann and Stroben, 1996; Spooner *et al.*, 1991). Later on Gooding *et al*. (2003) and Le Blanc *et al*. (2005) suggested that TBT prevented esterification of testosterone, thus leading to free testosterone and the above effects (Leblanc, Gooding and Sternberg, 2005; Gooding *et al.*, 2003).

Recently, retinoid X receptor has since been identified as being affected by TBT and leading to masculinisation in different species of mollusc (*Thais clavigera* (sea snail), *Nucella lapillus* (dog whelk) and *Nassarius reticulatus* (netted dog whelk)) (Pascoal *et al.*, 2012; Sousa *et al.*, 2010; Horiguchi *et al.*, 2007). Targeted deletion of RXR-alpha in

mice led to prostatic growth, but not invasive carcinoma and metastasis (Tanaka and De Luca, 2009; Huang *et al.*, 2002). There may be some similarity in the mode of action of development of prostate gland in molluscs and mammals. To understand the likelihood of the effects of chemicals on vertebrates and invertebrates, a basic understanding of their endocrine system is required.

1.8   The Gastropod neuro-endocrine system

The neuro-endocrine system in Gastropods is by far one of the best investigated and consists of neurohemal organs made up of neuro-secretory cells, nerve cells and endocrine glands (cerebral, pleural, pedal and abdominal) that produce neuro-peptides. These neuro-peptides are secreted to elicit appropriate response to the stimuli (external and physiological) and regulate processes such as development, growth and reproduction (Ketata *et al.*, 2008). The nervous system with cerebroid ganglia and a ventral nerve chain has a central role in coordination along with endocrine glands as in vertebrates. Many hormonally active regions have been identified as integral parts of their neuro-endocrine system. The dorsal body present in the dorsal part of the cerebral ganglia produces dorsal body hormone (DBH), which is a female gonadotropic hormone stimulating vitellogenesis, oocyte maturation and development of female accessory organs. The caudo-dorsal cell hormone (CDCH) produced by caudo-dorsal cells (CDC) in the cerebral ganglia are involved in control of ovulation egg mass formation and egg laying behaviour in freshwater pulmonates (Fig. 1.5) (Geraerts, 1976a; Geraerts and Algera, 1976; Geraerts, WPM and Joosse, J, 1975). Moreover, an egg laying hormone (ELH) has been reported in the sea hare *Aplysia californica* (Garden *et al.*, 1998; Mahon *et al.*, 1985; Arch and Smock, 1977; Arch, 1976; Arch, Earley and Smock, 1976). The pleural ganglia have the dark green cell system, which generates a hormone structurally analogous to vertebrate thyroid-stimulating releasing hormone (TRH) having a diuretic effect (Pinder *et al.*, 1999). The dark green cells are supported by yellow and yellow-green cells that produce peptides stimulating sodium influx in the body and thus control osmosis in the body (Pinder *et al.*, 1999). Male copulatory behaviour in pulmonates is controlled by Ala-Pro-Gly-Trp-amide-related peptide (APGWamide) synthesised in the right arterial lobe (AL) of the cerebral ganglion that is connected to the penial complex in both *Lymnaea* and *Aplysia* (Koene, 2010; Fan *et al.*, 1997; Croll and Van Minnen, 1992) (Figure 1.5). Although some peptide hormones are clearly specific to molluscs, there are few that seem to be functionally, or structurally, similar between vertebrates and molluscs. The lateral lobes in

26

molluscs generate gonadotropin-releasing hormone (GnRH) gene, a peptide hormone that plays a role in reproduction and on gamete differentiation in gonads in invertebrates (Gorbman and Sower, 2003; Rastogi *et al.*, 2002). In vertebrates GnRH (via other peptide hormones) acts on steroidogenic cells in the gonads, and it is the steroid hormones that affect gametogenesis and reproductive behaviour (Hartenstein, 2006). Using anti-peptide antibodies, several studies have substantiated the presence of GnRH-like peptides in a wide range of molluscs including gastropods (*Helisoma trivolvis* (freshwater snail), *Lymnaea stagnalis* (great pond snail)*, Loligo edulis* (European squid) and *A. californica*) (Onitsuka *et al.*, 2009; Zhang *et al.*, 2008; Young, Chang and Goldberg, 1999; Goldberg *et al.*, 1993), cephalopods (*O. vulgaris*) (Iwakoshi- Ukena *et al.*, 2004; Di Cristo *et al.*, 2002), bivalves (*Crassostrea gigas*, *Mytilus edulis* (blue mussel)*, Patinopecten yessoensis* (Yesso scallop)) (Treen *et al.*, 2012; Rodet *et al.*, 2005) and the squid (*Loligo bleekeri*) (Amano *et al.*, 2008).The light green cell system with the cerebral ganglia secretes molluscan insulin-like peptides (MIP). Five members (I, II, III, V and VII) of the MIP family have been characterised from the pulmonate snail *L. stagnalis* (Lardans *et al.*, 2001; Smit *et al.*, 1996), land snail *Otala lacteal* (Abdraba and Saleuddin, 2000) and in *C. gigas* (Hamano, Awaji and Usuki, 2005). Studies show that MIPs control body growth, development and metabolism (Gricourt *et al.*, 2003; Lebel *et al.*, 1996; Smit *et al.*, 1996) and are excellent candidates for insulin signaling pathways in schistosoma-host relationships (Dissous *et al.*, 2006). These peptides stimulate adenylcyclase and guanosine 5' triphosphate binding activity of G-proteins similar to insulin and insulin related factor1 in vertebrates (McClellan-Green, 2013). Another vertebrate peptide hormone analogue found in molluscs is Conopressin belonging to vasopressin/oxytocin family (neurohypophyseal hormones) (Van Kesteren *et al.*, 1995a). This family of peptides is of very ancient origin, found in organisms as diverse as worms and man (Beets *et al.*, 2013; Hoyle, 1998) and is believed to play a role in modulation of male sexual organs and contribute to the control of complex behaviour such as copulation (Garrison *et al.*, 2012; Van Soest and Kits, 2002).
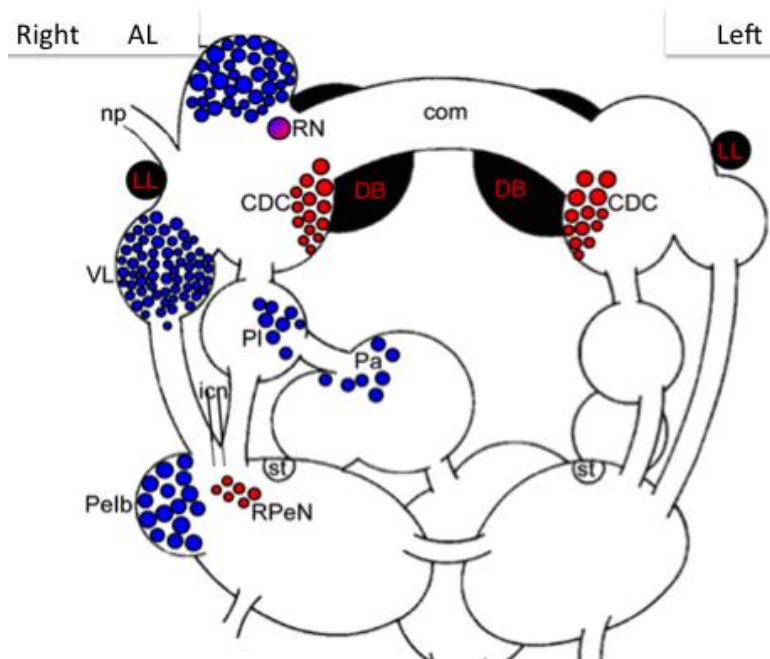
Figure 1.5: Neuroendocrine system of gastropod snails (Koene et al., 2010) showing involved ganglia of the central nervous system.
The neuronal clusters involved in male reproduction are indicated in blue and of female is indicated in red. AL: Anterior lobe; CDC: Caudo dorsal cells; com: Cerebral commissure; icn: Inferior cervical nerve; ni: Nervusintestinals; np: Nervous penis; RN: Ring neuron; RPeN: Right pedal neuron; st: Statocyst; VL: Ventral lobe.

In addition to the peptide hormones, recent research has been focused on the possibility that vertebrate type steroid hormones are functionally active and that they are endogenously produced in molluscs, although there is some controversy as to which steroids are naturally produced (Scott, 2012a). Vertebrate-like steroids (17β-estradiol, androgens and progesterone), have however, been measured in several species of molluscs and in numerous publications the potential for biosynthesis of vertebrate-type steroids has been demonstrated for molluscs (For example (Fernandes, Loi and Porte, 2011; Lafont and Mathieu, 2007; Janer and Porte, 2007; Osada, Tawarayama and Mori, 2004; Siah *et al.*, 2002; Le Curieux-Belfond *et al.*, 2001; Bose, Majumdar and Bhattacharya, 1997; D'aniello *et al.*, 1996; Jong-Brink *et al.*, 1981; De Longcamp, Lubet and Drosdowsky, 1974). Moreover, enzymes involved in steroidogenesis have also been found in gonad tissue (Osada, Tawarayama and Mori, 2004; Di Cosmo, Di Cristo and Paolucci, 2001; D'aniello *et al.*, 1996). Although the functional role of sex steroids in molluscs is questionable, there is evidence suggesting that steroids have marked effects when administered to molluscs. In 1980, aquaculture operators used immersion of oyster spats in either methyltestosterone (MT) or 17β-estradiol to produce unisex clutches of oysters (Koehler *et al.*, 2007). Research suggested that testosterone administration to castrated male slugs (*Euhadra prelionphala*) could stimulate the production of male secondary sex

characteristics (Koehler *et al.*, 2007). Similarly, testosterone exposure of female gastropods (*N. lapillus, N. reticulatus, Ilyanassa obsolete* (Eastern mud snail) and *Marisa cornuarietis* (Giant rams horn snail) caused pseudohermaphroditism or imposex measured by vas deferns sequence index (an index to measure the development of vas deferens and blockage of oviduct) (Oehlmann *et al.*, 1996) and penis length (Bettin, Oehlmann and Stroben, 1996b; Spooner *et al.*, 1991). Developmental exposure of snails to the potent steroidal androgen, MT or to the environmental androgen tributyltin (TBT) caused hypertrophy of the prostate gland and inhibition of spermatogenesis in males and transformation of the female albumen and capsule gland to prostate tissue and penis development in females (Watermann *et al.*, 2008; Oehlmann *et al.*, 1996). The growth induced by TBT were blocked by co-administration of the anti-androgenic drug CPA (Hagger *et al.*, 2009; Bettin, Oehlmann and Stroben, 1996) suggesting that the growths caused by TBT may be androgen-dependent as are many prostate tumours in humans and rodents. Overall there are many findings that suggest the existence of steroids similar to vertebrates but the primary issue is the precise mechanism in which these steroids exert endocrine control in molluscs, which is still not clear and require more research.

1.9   Conclusion

Modern societies produce thousands of chemicals some of which have the potential to interfere with the endocrine system of the organisms causing diseases. Declining health is of particular current concern, underlining the need for continued focus on the development of pharmaceuticals with which to treat these disorders and diseases. Moreover, scientific research on the relationship between the environment and health has highlighted concerns that these disorders may, in part, be caused by exposure to environmental pollutants. This has led to more stringent laws for screening of chemicals, devised by government bodies like EPA in the US and OECD in Europe. The screening process consists of a battery of tests designed to ensure steroid hormones such as estrogens and androgens do not mediate the mode of action of the chemicals. These tests included animal studies with a broad range of taxa such as fish, amphibians, rodents and a few other species of mammals. More tests meant more animals being used in research and all this led to finding alternatives to vertebrate animal testing for endocrine disruptors. These alternatives include *in vitro* assays, high throughput screening assays, use of invertebrate animals with short life cycle. My thesis discusses the potential of *Biomphalaria glabrata*, a mollusc as test organism for

androgenic chemical testing. Additionally, this new knowledge provides a broad evolutionary perspective on the molecular conservation of metabolic ad developmental pathways.

My research was carried out to test the following null hypothesis:
Androgenic chemicals affect the reproductive tract development in Biomphalaria glabrata and the process is mediated through homologues of sex steroid receptors as reported in vertebrates.

1.10 Aim of the thesis

The aim of my Ph.D. project was to test the hypothesis that common genes, proteins and physiological processes in both molluscs and humans underlie the response of the male reproductive system to androgens. In order to fulfill this aim the following objective are being addressed:

- Examine the physiological responses of androgenic chemicals (17α-Methyl testosterone and Dihydrotestosterone) on the development of reproductive tract in the snail (*Biomphalaria glabrata*).

- Investigate the fully sequenced genomes from *Biomphalaria glabrata* and *Lottia gigantea* to identify putative nuclear receptors from both the species of snails and cross-species comparison, looking at the similarities and differences that might exist.

- Identification and annotation of reproductive type neuropeptides in *Biomphalaria glabrata*.

2 Chapter 2: Introduction to the test species, *Biomphalaria glabrata* and preliminary work

2.1 Introduction

The phylum Mollusca is one of the most diverse group of triploblastic protostomes (Brusca *et al.*, 2005) ranking second to arthropoda with approximately 200000 extant marine, freshwater and land species (Ponder and Lindberg, 2008) that trace their origin to the earliest period of Cambrian at the commencement of the radiation of coelomate animals (animals with a true body cavity) (Brusca *et al.*, 2005). Based on morpho-anatomical evidence and molecular data, the phylum is divided into seven or eight classes, Aplacophora (Solenogastres, Caudofoveata), Monoplacophora, Polyplacophora, Scaphopoda, Cephalopoda, Bivalve and Gastropoda (Stöger *et al.*, 2013; Ponder and Lindberg, 2008). Of these, the focus of my thesis is on gastropods.

2.2   Gastropoda

The class Gastropoda emerged during the Ordovician period of important evolutionary differentiation within the molluscan phylum (Passamaneck, Schander and Halanych, 2004). If we consider the origin of life on the earth on a 24 hour clock with each second equivalent to 52,000 years then humans that diverged 2.6 million years ago, appear 54 seconds before the clock complete its 24 hours and the gastropods appear 500 million years ago, which is approximately 3 hours. Since their origin they have undergone dramatic radiations with over 100,000 extant species comprising about 80% of all mollusc species, forming one of the truly diverse classes of molluscs (Haszprunar and Wanninger, 2012). Based on the anatomical layout of the respiratory system of the gastropods, it has been classified into Opisthobranchia (gills behind the heart), Prosobranchia (gills in front of the heart) and Pulmonata (lungs instead of gills) (Brusca *et al.*, 2005). The main focus of this thesis is a freshwater pulmonate (*Biomphalaria glabrata*), although a marine prosobranch (*Lottia gigantea*) was studied for comparative cross-species computational analyses. What follows is an overview of the physiology and anatomy of these mollusc groups followed by detailed description of the test species.

2.2.1   Prosobranch

Prosobranch are the most primitive of all gastropods and are mainly gonochoristic with two different organisations, the diotocardian and the more advanced monotocardian. In diotocardian, fertilization is external and it discharges by the opening of the functional right kidney, whereas the monotocardian practice internal fertilization and they discharge through a separate glandular pallial genital duct (Fretter, 1984). In addition to the normal

sexual reproduction, some species of prosobranch exhibit parthenogenesis for example *Potamopyrgus antipodarum* (Collin, 1995; Wallace, 1992) and some species show protandric sex reversal (e.g. *Crepidula fornicata*, (Collin, 1995); *Busycon carica*, (Castagna and Kraeuter, 1994); *Lottia gigantea,* (Lindberg and Wright, 1985). A male phase precedes the female phase and is linked by a hermaphrodite phase (Geraerts, 1976b).

The male reproductive tract in the diotocardian is simple, having a single gonad and the tubules of the gonad join a short testicular duct that opens into the right kidney and discharge sperm directly into the water Figure 2.1. In the more advanced monotocardian, it is comprised of a long and highly coiled testicular duct that runs along the visceral mass leading into a ciliated vas deferens, which opens into the pallial duct and the prostate (Fretter, 1984). Usually sperm is produced in the tubules and passes along the vas deferens and the prostate where nutritive secretions are added. The sperm is then discharged through the penis, situated on the head behind the tentacle (Fretter, 1984) (Figure 2.1). The female reproductive tract in diotocardian is similar to males and the gonad forms the protective covering for the eggs, which are shed singly into the water bodies. In monotocardian, there is a more elaborate female reproductive tract starting with an ovary that runs into the ovarian duct and the fertilization pouch. The fertilized embryos move from the fertilization pouch along the oviduct into the proximal albumen gland, where a nutritive secretion is molded around the eggs, and then into the distal capsule gland or the jelly gland that produces a viscous fluid in which the egg capsules are embedded. This provides an extra protection to the egg capsules from osmotic pressure. The egg capsules are then harbored in the pallial duct (modified into a brood pouch) until hatched and then discharged through the vaginal opening (Fretter, 1984).
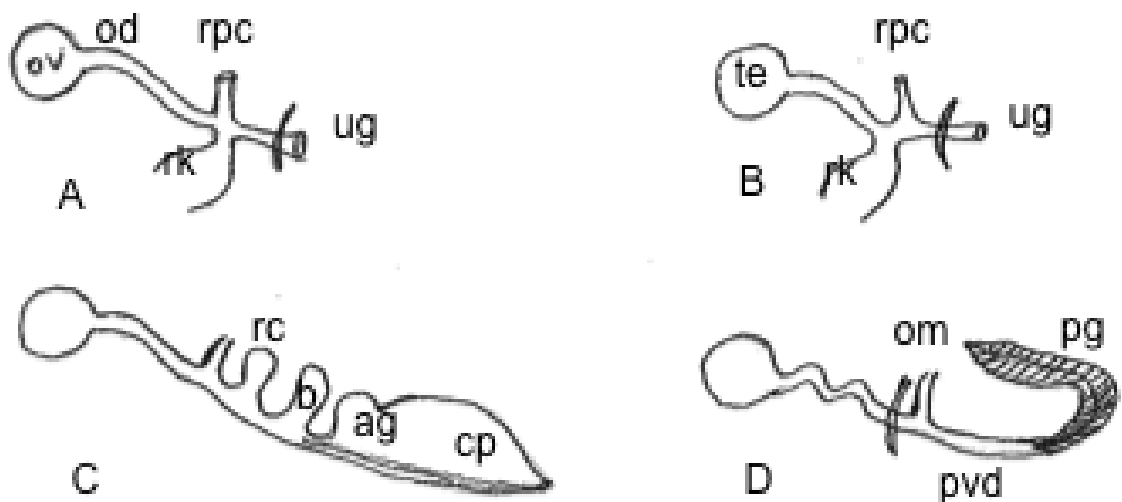
Figure 2.1: Reproductive tract of Diotocardian (A,B) female and male; Monotocardian (C,D) gemale and male.

Thickened lines indicate boundary of mantle cavity. ag: albumen gland; b: bursa copulatrix; cp: capsule gland; od: ovarian duct; om: opening to mantle cavity; ov: ovary; rc: receptaculum seminis; rk: right kidney; rpc: renopericardial canal; te: testis; ug: urinogenittal papilla (Adapted from (Fretter, 1984)).

2.2.1.1   Test species 1- Lottia gigantea (*for computational purposes*)

The prosobranch snail *L. gigantea* was chosen as the first lophotrochozoan for whole genome sequencing to provide a foundation for elucidating the evolutionary and ecological success of the species and its lineage for conservation purposes (Simakov *et al.*, 2012). The genome data was significant to define the core animal gene set, enabling the identification of many ancient animal genes that have apparently been lost in flies and nematodes but retained in molluscs (Chapman, 2007). The genome is small relative to other molluscs, consisting of about 500 mega base pairs (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html). The genome organisation, gene structures and functional content of these species are more similar to invertebrate deuterostomes (sea urchin and amphioxus) than to the sequenced protostomes (flies, nematodes and planarians). These similarities reveal features that might have been lost or diverged in the protostomes studied so far (Simakov *et al.*, 2012). *L. gigantea* has been used for computational analysis of the relevant genes in my thesis (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html).

2.2.2   Pulmonates

The pulmonates include the large majority of land snails and slugs; freshwater snails and a few families of estuarine and shallow water marine snails and slugs (Ponder and Lindberg, 2008) and are comprised of two groups; the Stylommatophora (one pair of tentacles with eyes on the end of the tentacle) and the Basommatophora (two pairs of tentacles and eyes at the base of the posterior tentacle). The Stylommatophora are entirely terrestrial (land slugs and snails) and the Basommatophora are mainly freshwater snails inhabiting ponds, ditches, rivers and lakes (Geraerts *et al.*, 1984). Basommatophora are hermaphrodite and show internal fertilization. Their reproductive system consists of a bisexual gonad (ovotestis), male and female ducts and accessory glands (albumen gland in female and prostate gland in male). Male and female gametes are produced in the ovotestis, which is located in the apical part of the visceral hump. When mature, both male and female gametes pass anteriorly into the hermaphrodite duct. With an exception of a few species, most of them have a short seminal vesicle along this duct (Geraerts *et al.*, 1984) and then male and female gametes are transported via separate paths. The junction of male and female ducts and hermaphrodite duct is called carrefour. Both albumen gland and the

fertilization pocket also diverge at this point. At copulation the sperm passes through the carrefour and are transported to the male duct. This region also receives sperm from cross-fertilization, which fertilize in the fertilization pouch and secretions from the albumen gland cover the oocytes preventing further sperm entrance (Egonmwan, 2007). The female duct is mainly glandular and secretes nutritive and packaging materials to support and protect the developing embryos (Heller, 1993). The female duct receives and often stores the sperms from a copulative partner to fertilize the eggs until the stores are exhausted (Geraerts, 1976b). The fertilized egg receives yolk from the albumen gland and passes into the uterus where tertiary membranes are added to the zygote before it moves to the oviduct and is extruded through the muscular vagina (Egonmwan, 2007). When the albumen gland stops secreting fluid, the ovulated oocytes that remain in the hermaphrodite duct and carrefour are transported to the bursa copulatrix where they are hydrolyzed and reabsorbed (Boyle and Yoshino, 2000).

The male duct is muscular and lined by ciliated epithelium helping to transport the sperm and to transfer it to its partner during mating. A prostate gland is usually present but there is a wide range of species-specific variation in structure. The prostate gland in higher Basommatophora is an elaborate and glandular with a number of diverticula (De Jong-Brink, 1969) associated with vas deferens, but in lower Basommatophora the prostate gland is just an elongated part of the male duct (Geraerts *et al.*, 1984). The male copulatory organ consists of penis transversed by the vas deferens and enclosed in the penis sheath, that has a terminal opening through which the penis can be everted. This is the general anatomical plan of the reproductive tract of all pulmonates, although a wide range of species-specific variation in these structures occurs (Geraerts *et al.*, 1984).

As evident from the above discussion the pulmonates differ in an important respect from prosobranch snails. Of the three extant subclasses, Prosobranchia diverged about 400 million years ago and underwent strong radiation, occupying most diverse habitats and different reproductive mechanisms (bisexual, hermaphrodite, parthenogenic) (Ram *et al.*, 1998) compared to pulmonates that are all hermaphrodite (Adiyodi and Adiyodi, 1988). Development in prosobranch is planktotrophic where they spend days or weeks in the plankton stage depending on the species. As they age, the larvae spend more time at greater depths and under favourable conditions undergo metamorphosis (Fretter and Shale, 1973). Studies have shown that chemotactic and thigmotactic responses initiate settlement, which may be induced by organic properties of the substratum beneficial to the adult (Struhsaker and Costlow, 1968). Conversely, pulmonates undergo direct development in eggs passing through morula, blastula, gastrula, trochophore and veliger stages (Little *et*

al., 1985) hatching out as miniature versions of adult snails. The rate of development is dependent on abiotic factors like temperature and salinity (Geraerts *et al.*, 1984). Unlike prosobranch, there is no evidence of pulmonates showing chemotactic and thermotactic responses but the Basommatophora are known to have a neuroendocrine center that responds to external and internal stimuli through various hormones and neuropeptides (Lever and Boer, 1983). Endocrine control of reproduction in prosobranch is achieved by first order neurosecretory mechanisms through cerebral ganglia, pedal ganglia, cerebropleural connective areas controlling non-endocrine target organs, but in pulmonates, both first and second order neurosecretory mechanisms occur; that is neurosecretory cells control endocrine and non-endocrine target organs (discussed in detail in Chapter 1).

2.2.2.1   *Test species 2-Biomphalaria glabrata*

The most extensively studied order of pulmonates is Basommatophora (basal eye-bearing stalk). More specifically, *Biomphalaria* in the genus Planorbidae has become a focus for research due to the integral role it plays in the transmission of human intestinal schistosomiasis. The tropical freshwater snail *Biomphalaria glabrata* is an intermediate host for several digenean trematode parasitic worms, including *Schistosoma mansoni*, the causative agent of schistosomiasis. This is the most widespread trematode infection affecting around 200 million people, leading to a chronic debilitating disease and up to 200,000 deaths per year, across 75 developing countries (WHO report, 2011). Schistosomiasis is a disease acquired by people when they come in contact with freshwater infested with larval form (cercariae) of *Schistosoma mansoni. B. glabrata* is the intermediate host of *S. mansoni* that is infected by free-swimming larvae (miracidium) of *S. mansoni* (Bayne, 2009a). For all the above reasons *B. glabrata* has been a focus of research in the field of taxonomy, morphology, cytology, physiology, ecology, genetics, and reproduction (Raghavan and Knight, 2006; Camey T and Verdok NH, 1970).

My test species, *Biomphalaria glabrata*, strain BB02, has been widely studied as a model intermediate host for *Schistosoma mansoni* since the mid-1950s when it was found that susceptibility in *B. glabrata* to *S. mansoni* is a heritable trait (Ittiprasert and Knight, 2012; Richards and Shade, 1987; Newton WL,). Laboratories across the world are actively studying the molecular biology of *B. glabrata* trying to identify genes and their products that interfere with parasite survival in the snail and may reduce transmission of schistosomiasis to humans e.g. (Ittiprasert and Knight, 2012; Bayne, 2009a; Raghavan *et al.*, 2003). From a biological perspective, *B. glabrata* possesses some unique properties. It

is a tropical, aquatic, hermaphrodite organism and has been used as a research organism in comparative immunology (Rollinson, 2011; Loker, 2010), evolutionary systematics, biogeography (DeJong *et al.*, 2001) and environmental toxicological studies (Ansaldo *et al.*, 2009; Salice, Miller and Roesijadi, 2009; Ansaldo *et al.*, 2006). It can be easily maintained under laboratory conditions and its natural diet consists of aquatic plants and the cyanobacterium, *nostoc* (found in soil and mud) (Toledo and Fried, 2011; Eveland and Haseeb, 2011) and in the laboratory can be fed dried lettuce or dried fish food.

2.2.2.1.1 General Anatomy

The shell of *B. glabrata* is composed of calcium carbonate embedded in small amount of organic matrix secreted by the mantle. The mantle is a line of skin, which transcends the shell and is divided into three chambers (pulmonary, water inflow and water outflow) that help in flotation, oxygen reservoir, excreta circulation and elimination (Toledo and Fried, 2011; Jurberg, Cunha and Rodrigues, 1997). Inner layers of the outer mantle fold secrete the outermost organic layer of the shell, the periostracum, during embryogenesis (Hasse *et al.*, 2000). The *Biomphalaria* genus (as well as all gastropods) undergoes a process called torsion and the shell undergoes counter clockwise spiral (known as sinistral coiling). As the snail ages, the shell slowly increases in size from the apex (specifically the prontoconch) towards the aperture (Marxen *et al.*, 2003). As a consequence, the mantle cavity (as well as the anus) is prostrated over the head-foot Figure 2.2. As the name suggests, the head-foot functions both, as a sensory organ (which has protruding tentacles, mouth and feeding structures) and provides locomotion. The heart is located in a pericardial chamber and contains a distinctive atria and ventricle.
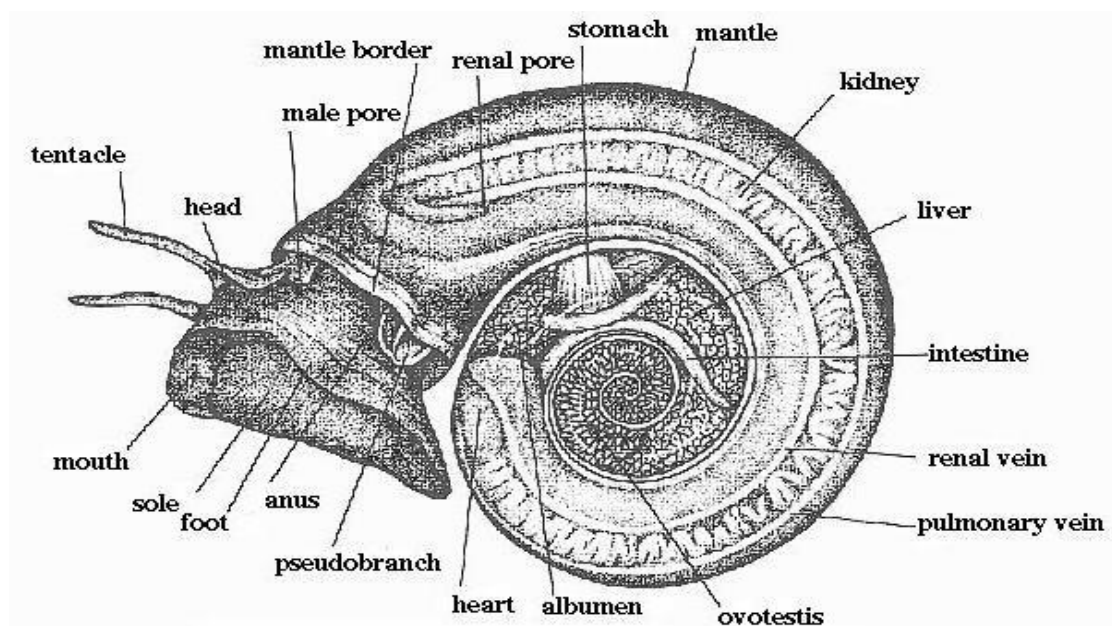
Figure 2.2: Figure showing anatomical parts of the Biomphalaria glabrata.

Although the anatomical features involved in food ingestion by *B. glabrata* are basically similar to other pulmonates, there are some unique features. The adult snail has a total of about 4425 radula teeth arranged in 75 rows well adapted to feed upon diatoms from the surface of freshwater macrophytes and the area grazed can be regulated by the change in direction and velocity of head (Thomas, Nwanko and Sterry, 1985).

2.2.2.1.2   Life cycle and reproduction

B. glabrata snails start their reproductive cycle as protandric males and then change to simultaneous hermaphrodites (Trigwell and Dussart, 1998; Vernon and Taylor, 1996) with the ability to both self-fertilize (when isolated) and cross fertilize, although they prefer to cross fertilize (Vianey-Liaud, Dominique and Dussart G, 1996; Jarne, Vianey-Liaud and Delay, 1993). They can exchange bundles of sperm to avoid self-fertilization.

The reproductive tract of *B. glabrata* is subdivided into three parts (hermaphrodite, male and female), each including several structures (accessory gland, ducts, etc. Figure 2.3). Both male and female gametes are produced within the hermaphrodite gonad called the ovotestis, situated in the inner whorls of the shell and dorso-laterally covered by the orangish digestive gland (hepatopancreas). The collecting canal of the ovotestis carries male and female gametes into a hermaphrodite duct, which leads into seminal vesicles. The spermatozoa are stored in seminal vesicles until maturation, whereas oocytes proceed anteriorly to the receptaculum seminis, where fertilization occurs. The carrefour is present at the junction of the hermaphrodite duct and the receptaculum seminis and its main role is to receive sperms from copulating partner and transfer it to fertilization pouch
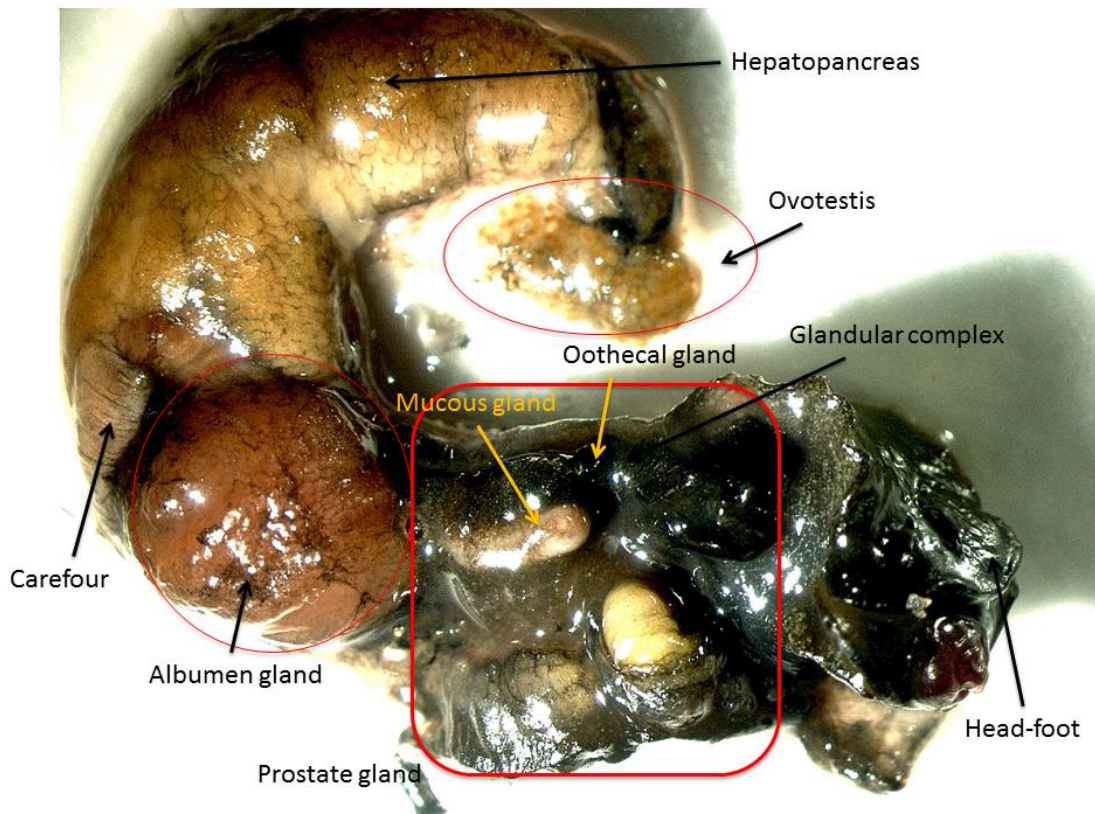
Figure 2.3: Figure showing the schematic of reproductive tract in B. glabrata.
Starting with ovotestis leading into the hermaphrodite duct to the carrefour where fertilization occurs and the male and female ducts get separated. The female accessory gland is shown that leads into the glandular complex comprising of both male (prostate diverticulum and spermduct) and female accessory glands (mucous gland and membrane gland) closely intertwined into each other.

(Egonmwan, 2007). Albumen gland is one of the female accessory reproductive glands, is a large tubulous gland that pours its secretions into the Carrefour. These secretions (galactogen – a polysaccharide) form a part of the perivitelline fluid that is used as a nutrient by the developing embryos in their eggs and after hatching (Mukai *et al.*, 2004; Ramasubraniam, 1979). The first part of the female duct, the oviduct, has a tube like appearance, continuing into a swollen convoluted part, the muciparous gland. This gland is not distinctly separate but is closely associated to the oothecal gland, the uterus, and the bursa copulatrix, which are connected with the vagina by a small duct. The external orifice of vagina opens a short distance behind the male genital outlet. The male part starts with sperm duct and runs parallel to the oviduct, curving along the muciparous gland. Between the transitions of the muciparous gland, oothecal gland and bursa copulatrix, the sperm duct has numerous branched and unbranched diverticula forming the prostate gland (De Jong-Brink, 1969). The foremost diverticulum is usually inserted between the oothecal gland and the muciparous gland, but in some instances it partially covers the latter part of the prostate diverticula. This runs into the vas deferens that continues into the penial

complex, consisting of penis sheath and prepuce. The penis sheath is uniformly cylindrical enclosing the penis consisting of well developed erectile tissue surrounding the penis canal and enveloped by a double muscular coat. The vas deferens and prostate help in conveying the semen and also aids in propelling the penis through the prepuce during copulation (De Jong-Brink, 1969). The main function of the male duct is to transport sperm to the copulative partner.

Reproductive behaviour in this species has been divided into four phases by (Trigwell, Dussart and Vianey-Liaud, 1997) and starts with precourtship during which an active searching of mate takes place and the most mobile snail acts as a male. This is slightly different from the pond snail, *Lymnaea stagnalis* in which the period of isolation is the determining factor for the male actor (Van Duivenboden and Maat, 1985). This phase is followed by courtship, copulation and post copulation stages. Copulation involves penetration of female genital pore by the penis of the partner followed by insemination, when the sperm is introduced into the female genital tract and fertilizes the egg internally producing embryos. This could be followed by a second copulation, which is not necessarily with the same individual except in isolated pairs (Vernon and Taylor, 1996), or copulation could end with both the snails moving apart. Ovulation and oviposition follows thereafter, laying capsules with an average of 40-60 eggs and this continues until the allosperms stored in the female tract are exhausted (Vianey-Liaud *et al.*, 1989). There is a typical 4-6 week egg-egg generation time. Since the generation time is approximately 5 weeks, the snail can undergo several generations in a year and under ideal conditions one snail can produce 14,000 eggs in its life span (Toledo and Fried, 2011).

2.2.2.1.3 Molecular tools to study the biology of *B. glabrata*

Biomphalaria is probably the most widely studied mollusc as a model intermediate host for the human parasite *Schistosoma mansoni* with more than 50 years of research invested in its genetics and physiology. There are several genomic tools available for *B. glabrata* that make this species a potential model for various purposes.

2.2.2.1.3.1 Bacterial artificial chromosome (BAC) Library

A BAC is an engineered DNA molecule used to clone large DNA segments (100,000-300,000bps) from the organism, say *B. glabrata*, by inserting these segments into the bacterial cells (for example, *E. coli*) (Shizuya *et al.*, 1992; O'Connor, Peifer and Bender, 1989). As the bacterial cells grow and divide, they amplify the BAC DNA, which can then be isolated and used in sequencing DNA. The sequenced parts are rearranged *in silico*, resulting in the genomic sequence of the organism. BAC libraries provide a platform for

the creation of physical maps and complete genome sequencing, elucidation of gene function and regulation, map-based cloning of quantitative trait loci used in modeling genetic diseases, and comparisons of specific regions between different species or ecotypes, including establishment of synteny and functional single nucleotide polymorphism (SNP) analyses in coding and regulatory sequences (Farrar and Donnison, 2007; Shizuya and Kouros-Mehr, 2001). Within molluscs, BAC libraries have been constructed from bivalves, Eastern oyster and Pacific oyster (Cunningham *et al.*, 2006); Scallop (*Chlamys farreri*) (Scheuring, 2008) and *Aplysia californica* as a prerequisite for sequencing. *B. glabrata* snails were collected from Brazil, South America and the genomic DNA from this strain (BB02 (*Biomphalaria* from Barreiro caught in 2002)) was used for the production of BAC library (Adema *et al.*, 2006a). A BAC library from BB02 provided collectively a total of 61824 clones with an average insert size of 136.3kb and provides 9.05X coverage of the 931Mb genome. Odoemelam and colleagues (2009) developed protocol to perform physical mapping of BAC onto chromosome by Fluorescence in-situ hybridization (FISH) for *B. glabrata* (Odoemelam *et al.*, 2009).

2.2.2.1.3.2 Microarrays

DNA microarray technology is used to examine the expression of thousands of genes at once (Schena *et al.*, 1995) in a hybridisation-based assay. In microarray experiments, mRNA from one target is made into cDNA labelled with red fluorescent dye, and the other target is labelled with green fluorescent dye. The two cDNA samples are combined into a single solution and hybridized to the microarray slide, which is a collection of single stranded DNA spots, each representing a gene sequence. The cDNA binds to the DNA spots and the colour of the fluorescence indicates the gene expression. When both red and green fluorescence exist on the same dot it appears to be yellow (Hoopes, 2008; Heller, 2002). The application of cDNA microarray technology has brought about a significant change in the breadth of hypotheses that can be explored. For example, transcriptional response patterns can generate a better understanding of mechanisms of disease, identifying disease sub phenotypes or predicting disease progression in different organisms. The first application of microarray technology in *B. glabrata* identified 98 differentially expressed genes between snail strains resistant or susceptible to *S. mansoni*, 94 resistance-associated and 4 susceptible-associated (Lockyer *et al.*, 2008). A 70-mer oligonucleotide-based microarray approach has also been used in *B. glabrata* to investigate defense response of the snails depending on the nature of stimulus (wounding, exposure to gram +ve bacteria (*Micrococcus luteus*), gram −ve bacteria (*Escherichia coli*)

or infection with trematode parasites (*Schistosoma mansoni* and *Echinostoma paraensei*)). Wounding yielded a modest response in the differential expression and it was different from the other infections. Distinct differential expression was obtained of stress/defense related features by bacterial infection but significantly altered expression of immune related genes was obtained from the trematodes and the immunosuppressive interference of *E. paraensei* is stronger than *S. mansoni* (Adema and Loker, 2009). The application of these types of approaches are a step forward in understanding the molecular interplay between parasite and host and the availability of these resources make *B glabrata* a potential model organism for other purposes.

2.2.2.1.3.3 RNAinterference

RNAinterference (RNAi) is a process of introducing an exogenous single-stranded (siRNA) or double-stranded (dsRNA) RNA corresponding to a mRNA sequence and through specific binding, the RNA sequence coding for the target gene is depleted resulting in significant reduction in the levels of target mRNA (Campbell and Choy, 2005; Scherer and Rossi, 2003; Sharp, 2001; Fire *et al.*, 1998). The effects of this reduction can be assessed by its phenotype change, which could indicate the biological role of the gene. Since RNAi may not totally abolish expression of the gene, this technique is sometimes referred as a "knockdown", to distinguish it from "knockout" procedures where the gene is completely erased from the genome through homologous recombination (exchange of nucleotide sequences between non-sister chromatids). The first reported RNAi knockdown for a pulmonate snail was in *L. stagnalis,* which was injected with dsRNA corresponding to the transcript for nitric oxide synthase that led to the reduction in expression of this gene, affecting the feeding behaviour of treated snails (Korneev *et al.*, 2002). In *B glabrata* RNAi mediated knockdown of the fibrinogen-related protein (FREP2) gene involved in molluscan internal defense was successful (Jiang, Loker and Zhang, 2006). Recently a non-invasive technique of post-transcription gene silencing using cationic polymer polyethyleneimine (PEI) mediated delivery of long double-stranded and small interfering RNA (soaking the snails in dsRNA/PEI or siRNA/PEI) was tried in juvenile snails. The technique was successful and led to knockdown in expression for the gene peroxiredoxin at nucleotide and protein level (Knight *et al.*, 2011).

2.2.2.1.3.4 Biomphalaria glabrata genome

The *B. glabrata* genome initiative is an effort by an international consortium to identify comparable sequence information from the snail host to that has already been achieved for *S. mansoni* (parasite)   (Berriman *et al.*, 2009) and human (definitive host) genomes

(Lander *et al.*, 2001; Venter *et al.*, 2001), which may provide predictors for epidemiological studies on parasite transmission. With the advent of modern genome biology, genomic information plays a significant role in the development of treatment and control measures for such widespread diseases, as emphasised by the recent publication of the complete genome sequences of the malaria parasite (*Plasmodium falciparum*) and its mosquito vector (*Anopheles gambiae*) (Gardner *et al.*, 2002; Hoffman *et al.*, 2002; Holt *et al.*, 2002). The genome sequencing project for *B. glabrata* was initiated in 2005 at the Washington University Genome Sequencing Centre (WUGS) (Raghavan and Knight, 2006). The first large scale set of genomic sequence data employing whole genome shotgun sequencing was contributed by WUGS and the expressed sequence tags (ESTs) derived either from whole snail tissue or different tissues. The genome of *B. glabrata*, using BB02 (susceptible) strain from Brazil, is composed of 2n = 36 small chromosomes (Toledo and Fried, 2011; Burch, 1960) with a genome size of ~ 950Mb with high heterozygosity and is AT rich (74.6%) (Ittiprasert and Knight, 2012).

All the background information and the availability of the draft genome make *B. glabrata* an appropriate test species for my project. The fact that it can be easily bred in the laboratory with a generation time of approximately two months is helpful in terms of undertaking developmental studies in a short time span that is one of the main objectives of the project. *L. gigantea* genome will be used for cross-species comparison of the relevant genes of interest.

2.2.2.2 *Preliminary work done on test species (B. glabrata) and optimization of techniques*

Before commencing the main experimental study, a number of preliminary experiments and procedures were conducted with *B. glabrata* to inform the main experimental design. These were primarily divided into the following subsections,

- Determination of the likely critical window for androgen/anti-androgen exposure in immature/developing snails (the time period between hatching and onset of puberty). Identification and description of the reproductive tract/accessory sex organs and optimization of dissection and histopathological techniques.

- Optimisation of the protocol for gonadectomy, including the success rate; survival following gonadectomy and the utility of the gonadectomised snail model as a part of this project.

- Pilot study exposing gonadectomised snails to one of the androgenic chemical, 5α-dihydrotestosterone to be used in the exposure study.

2.2.2.2.1 Sexual development in *B. glabrata* from day of hatch to sexual maturity

In *B. glabrata* there is literature related to the embryogeny of development (Vasta, LeSage and Fried, 2011; De Jong-Brink, 1969; Boyce, Tieze-Dagevos and Larman, 1967) but not much information is available on time-specific development, which would be helpful to identify the time of gonadal differentiation and the onset of puberty. This information was needed before the androgen exposure study could be initiated.

Materials and methods

Biomphalaria glabrata acquisition

An initial stock of *B. glabrata* (BB02 strain) was kindly supplied by the Natural History Museum, London, UK in 2010. Thereafter a captive-bred stock has been maintained at our facility at Brunel University, London (UK). For the preliminary developmental study, one group of 10 adult snails (diameter>12mm) from same breeding stock was selected.

General laboratory culture conditions

During non-experimental periods, adult and juvenile snails were housed in a recirculation system in glass aquaria supplied by de-chlorinated tap water. Snails were fed Tetramin fish food flakes (roughly ~10% of their body mass) *ad libitum* three times a week. Any uneaten flakes or rotting vegetation and dead snails were siphoned out prior to feeding. Ammonia (0mg/l), nitrate (0mg/l), nitrite (<40mg/l) (API aquarium test strips) and dissolved oxygen levels (7 ± 0.5 mg/L) using Haan Lange probe were monitored in the water system 3 times a week, as was pH (7 ± 0.21) and general water hardness (120mg/l). Low pH and carbonate levels were adjusted by the addition of commercial aquaculture re-mineralizer (Remin) to sustain snail growth and prevent shell thinning (White *et al.*, 2007).

Temperature and photoperiod

The snail room temperature was maintained at 27^0C by the use of an internal air-conditioning unit and similar temperature was maintained in the snail tanks. The photoperiod for all work conducted in the snail room was16:8 Light:Dark incorporating 20 min dawn/dusk transition periods and this was controlled externally by a digital timer switch.

Routine collection of egg masses

Egg masses were easily observed sticking to the glass or pipe-work in the aquaria. Egg

masses laid on flat surfaces were removed using a sterile scalpel pressed against the edge of the egg mass taking care not to damage the eggs inside the egg mass. If the egg mass was laid on pipe-work or in a corner, it could be easily taken off using thumbnail in similar manner. Once removed, egg masses were placed, 2-3 egg masses per basket, in metal mesh baskets, which were suspended using glass rods in the tanks to keep the egg masses at optimal temperature and provide sufficient mobility space for the juveniles.

Biomphalaria glabrata rearing protocol

Under these optimum conditions *B. glabrata* eggs start hatching 5 days post oviposition and hatched snails were fed finely crushed Tetramin fish food flakes (approximately 0.05gms). Most of the snails hatched 6-7 days post-oviposition and hatching completed by day 11 post-oviposition. The hatchlings were left in the baskets, fed *ad libitum* three times a week until big enough not to escape the tanks. Once released into the tanks at the density of 100-150 hatchlings per tank with working volume of 3.7L, they were fed @ 0.5 gm/tank/alternate-day and this was increased to 3.0 gm/tank/alternate-day as the juvenile snail's age. Under these conditions they attain reproductive age by 30 days post oviposition and they lay, on average, 1-2 egg masses a day (containing 20-50 eggs per mass) until six months of age, after which their reproductive fitness steadily declines.

For experimental purposes, egg masses were taken from adult snails of 2-3 months of age to maintain ideal hatching and developmental conditions. The snails were monitored for two weeks to get an estimate of egg masses laid per day. The snail tanks were cleaned one day before the beginning of this study and any egg masses present or uneaten flakes or rotting vegetation and dead snails were siphoned out.

Husbandry of hatchlings and juvenile snails

Four egg masses laid on the same day were collected and one egg mass was placed in each of the four wells of a six-well plate (nunc, sterile non-treated), filled with tank water, and floated in the snail flow through recirculation system maintained at optimum conditions mentioned above. A 50% water change was conducted every 48 hours and the wells were refilled with the tank water. They were closely monitored for any hatching. Hatching started on day 5 post-oviposition with maximum hatching in the 6-7 days post-oviposition. When hatching started egg masses containing unhatched embryos were moved to a new well and the egg mass were washed off using a small jet of water from a pipette to remove any newly hatched snails not yet left the surface of the egg mass. The well containing the newly hatched snails was re-labeled specifying their date of hatch. This process was repeated until all the snails had hatched. Hatchlings were fed on finely crushed Tetramin

fish food flakes twice a week. At 10 days post hatch, the hatchlings were taken out of the six well plates using a small 1ml plastic pipette with the nozzle diagonally cut and transferred into a basket, suspended in the tanks using glass rods. The hatchlings were kept at a density of ~30-40 snails per basket but due to natural mortality or continual sacrificing for histology, the density was not constant. Cleaning and feeding continued every 48 hours with the feed increasing @ 0.5gm/tank/day, after 15 days post-hatch (dph). These juvenile snails were kept in the baskets in the tanks till 60 days post hatch.

Sampling of the hatchlings and juvenile snails

Hatchlings and juvenile snails were sacrificed and fixed for histology at 3dph and then on 5, 7, 9, 11, 13, 15, 19, 21, 23, 27, 30, 34, 42, 49, 56 and 60dph. Snails were narcotized in 5% w/v Magnesium Chloride hexahydrate ($MgCl_2 6H_2O$) (≤99%, Sigma-Aldrich) solution (using purite water) for 30 minutes, or until the snails showed no mobility, which ever was the longer. Once the snails were fully narcotized, the hatchlings from 3dph − 15dph were fixed whole (with shells) in Kahel's fixative (Frase, 1986) and kept at room temperature in 7ml glass vials until further histological processing. The bottles were labeled with the species name, age in dph and the sampling date.

Histopathology processing and analyses

Prior to processing, the hatchlings (3dph -9dph) were placed in biopsy cassettes (Histosette II) and to prevent any loss of sample, the tissue was sandwiched between two Whatman filter papers dipped in 70% IMS. Each specimen was designated its own biopsy cassette and labelled with the individual snail reference number. The fixed tissues were serially dehydrated in an automatic tissue processor (TP 1020, Leica Inc.) passing the samples through a series of increasing alcohol concentrations and cleared with Histoclear II (National diagnostics) before being vacuum impregnated with wax and subsequently embedded in wax blocks taking care that the tissue was not too much folded, especially the ovotestis Table 2.1. Tissue samples were then sectioned on a rotary microtome (Lieca RM2235) using disposable microtome blades (Shanndon MB35 premier 35^0/ 80 mm), at a thickness of 5µm. All sections were mounted on glass polysine treated microscopic slides (VWR) and each slide was labelled with the individual snail reference number and slide number on the opaque end. They were dried overnight (minimum 12 hours) on warming racks (@45^0C photoxdishwarmer) before staining with Mayer's haematoxylin (VWR) and 1% aqueous eosin (VWR) according to the protocol in Table 2.2. After staining the slides were left to dry for a few minutes in the fume cupboard to allow Histoclear II (RA Lamb) to evaporate. A droplet of histomount was placed on the sections of tissue and a coverslip

(22 x50 mm, No. 1 thickness, RA Lamb) was mounted onto the slide. Care was taken to avoid any bubbles formation on or around the tissue sections and any excess of histomount on the slide was removed using tissue paper dipped in histoclear. The slides were dried in the fume cupboard for at least 24 hours and then were examined under an Olympus (BX51) light microscope using x2, x10 and x40 objectives (total magnification x20, x100 and x400 respectively). For each tissue, the cell types and structure and any abnormalities were assessed. Photomicrographs were taken using a digital camera (Q Imaging Micropublisher 5.0RTV) linked to a personal computer. Q Capture Pro 5.1 software was used to capture and view images from the digital camera.

Table 2.1: Summary of the tissue processing method of the Leica Tissue Processor, indicating the various solutions used and time taken for each step.

CYCLE NUMBER	TREATMENT	HOURS
1	70% IMS	3
2	90% IMS	2.5
3	95% IMS	1.5
4	100% IMS	1.5
5	100% IMS	1.5
6	100% IMS	1.5
7	100% IMS	1.5
8	HISTOCLEAR	1.5
9	HISTOCLEAR	1.5
10	HISTOCLEAR	1.5
11	WAX	1.25
12	WAX	1.25
		20 HOURS TOTAL

Table 2.2: The staining protocol for histological slides of various tissues from B. glabrata.

Step	Bath	Chemical	In bath time (secs)	In bath agitation	Out bath time (secs)	Out bath agitation
0	0	Histoclear	900	N	8	Y
1	1	100% IMS	120	N	8	Y
2	2	90% IMS	120	N	8	Y
3	4	70% IMS	120	N	8	Y
4	3	Water	120	N	8	Y
5	5	Haematoxylin	600	N	8	Y
6	3	Water	600	N	8	Y
7	6	Acid alcohol	20	N	8	Y
8	3	Water	20	Y	8	Y
9	7	Sat. Li_2CO_3	20	N	8	Y
10	3	Water	20	Y	8	Y
11	8	Eosin	40	N	8	Y
12	3	Water	300	N	8	Y
13	4	70% IMS	120	N	8	Y
14	9	90% IMS	120	N	8	Y
15	10	100% IMS	300	N	8	Y
16	11	Histoclear	300	N	0	N
17	11	Histoclear	0	N	0	N

Results and Discussion

B. glabrata undergoes direct development and when hatched the snails look like

miniatures of an adult snail with a transparent shell that is separated from the head region by the mantle fold, which is depressed to form the mantle cavity. The relative position of the digestive gland and gut were established at 3dph. The first evidence of the hepatopancreas could be seen at 6dph while the kidney appeared on the 9th day after hatch. The glandular complex comprising of prostate, membrane gland and mucous gland could be seen in primary stages of development at 9dph. No gonad was observed in *B. glabrata* hatchlings from the day of hatch until 12dph Figure 2.4(A).   There is evidence that juvenile snails (less than 8mm in diameter) invest first in the development of ovotestis and male organs while development of the albumen gland and female organs starts later in the sub adult stage (when they are greater than 8mm in diameter) (Koene and Ter Maat, 2007; Trigwell and Dussart, 1998). The genital rudiment had extended its full length posteriorly showing early stages of gonadal acini formation from 14dph to 16dph Figure 2.4(B). Mature gonads with large numbers of sertoli cells were seen at 20dph Figure 2.4(C) and full spermatogenesis with spermatozoa and oocytes in the lumen was seen at 28dph Figure 2.4(D). The albumen gland and glandular complex were fully developed on the 22dph.
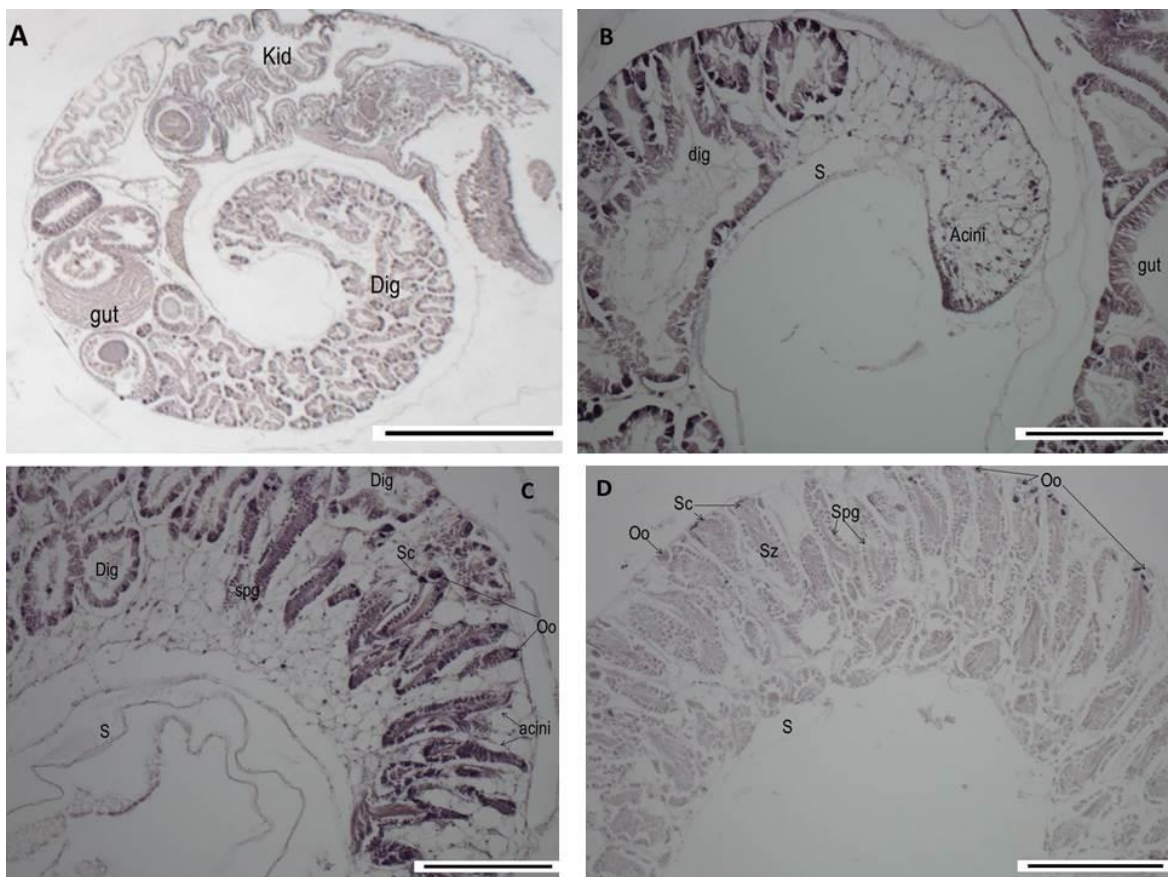
Figure 2.4: Pictomicrograph of B. glabrata hatchlings showing the gonad at different stages of development.
(A) 9dph with no gonad visible (x100); (B) 14dph: developing acini first become apparent (x100); (C) 20 dph showing a number of acini with sertoli cells, spermatogonium, spermatocytes and early stage oocytes (x100); (D) 24dph acini have full spermatogenesis and oogenesis to stage 3 (x100). 5μm sections stained with H & E. digestive tissue – dig; foot – f; gut – g; kidney –k; oocyte – O; shell – S; Sertoli cells – ser; spermatocytes – spc; spermatids –spd; spermatogonium – spg; spermatozoa – spz; tentacle – t.

Overall the preliminary experimental results indicate that the time period between hatching and onset of puberty is very short, between 12dph to 24 dph, which only provides a short opportunity to evaluate effects of the androgens on the developing reproductive system before natural puberty begins. In order to avoid any confounding factors that might affect the response or missing this critical period, it was decided to conduct androgen exposures from the egg mass stage through to sexual maturity (approximately 30 days post hatch (dph)).

2.2.2.2.2 Optimization of gonadectomy technique

In the rodent Hershberger assay, castration (Moon *et al.*, 2009) is usually practiced in the adult version of the assay as it removes any endogenous sources of androgen and improves the response rate. Procedures for castration of adult snails have been described in the older scientific literature (Vianey-Liaud and Dussart, 2002; Miksys and Saleuddin, 1987) where castration of hermaphroditic pulmonate snails has been reported to cause acceleration in the growth of female accessory sex organs (albumen gland) (Miksys and Saleuddin, 1987; Geraerts and Mohamed, 1981; De Jong-Brink *et al.*, 1979) and the size of the male accessory sex organ (prostate organ) remained unaffected (Harry, 1965). In contrast, Harry (1965) reported that castration of juvenile *B. glabrata* prevented the development of accessory sex organs and concluded that the growth of the accessory sex organs is dependent on the presence of the gonad and thus on hormones (Harry, 1965). Geraerts and Josse (1975) also reported that two blood-borne hormones were responsible for the development of prostate gland and female glands respectively (Geraerts, WPM and Joosse, J, 1975).

The hypothesis was that the gonadecotomised adult snail could serve as a useful model with which to explore whether steroidal substances produced by the gonads are necessary for development and maintenance of the snail prostate gland as is the case in mammals. If so, removing the source of endogenous hormones could produce a snail equivalent of the castrated rodent, which could be used to look at the response of the snail to androgens.

Material and Method

Procedure for gonadectomy

The snails used in this study were sexually mature snails from the same breeding stock maintained at our facility at Brunel University, London (UK). Snails having shell diameter

within the range of between 10-12mm (measured across the widest portion of the shell) were selected for the procedure (Dussart G, personal communication). Snails were removed from their tanks, blotted dry, weighed (g) using an electronic balance and their shell diameter measured to the nearest 0.1mm using hand-held digital calipers. Castration was carried out on ten snails at one time, by gently breaking open the apex of the central whorl of the shell on the right side of the posterior body wall with fine forceps, thereby exposing the hepatopancreas and the ovotestis Figure 2.5. The ovotestis can be observed as a narrowing pink or orange spotty structure at the very posterior of the animal; it is often curled (much like the shell) while the digestive tissue is much larger and generally brown in colour, dorso-laterally covering the ovotestis. The membrane overlying the ovotestis was cut using a pair of dissecting scissors with fine point open shanks curved (Fishers Scientific, UK) and with fine forceps along the length of the curvature and stripped away on either side laying bare the ovotestis. The hermaphrodite duct was ligatured using synthetic absorbable sterile surgical sutures (Vicryl), and the ovotestis was gently aspirated using plastic pipette not exerting too much force, taking care not to damage the hepatopancreas. The shell opening was then sealed with autoclaved shell pieces of dead snails with sterile veterinary tissue adhesive (Vetloc). After surgery, animals were allowed to recover in standing water containing an antibiotic gentamycin (Sigma-Aldrich) at a final concentration of 40µg/ml and 0.25µg/ml Amphotericin B (fungizone) (Sigma-Aldrich) for 1-2 hours. The gonadectomised snails were transferred in newly assigned tanks in the flow through system. Ten sham-operated controls were treated the similar manner as the gonadectomised snails, except that no gonadal tissue was removed. The process was repeated until >30% survival rate could be obtained (in line with others performing a similar procedure, Dussart G, personal communication). A batch of 50 snails of similar age was then gonadectomised and 50 snails were sham-operated. The gonadectomised and sham-operated snails were kept in separate tanks, and monitored for mortality and growth for six weeks.
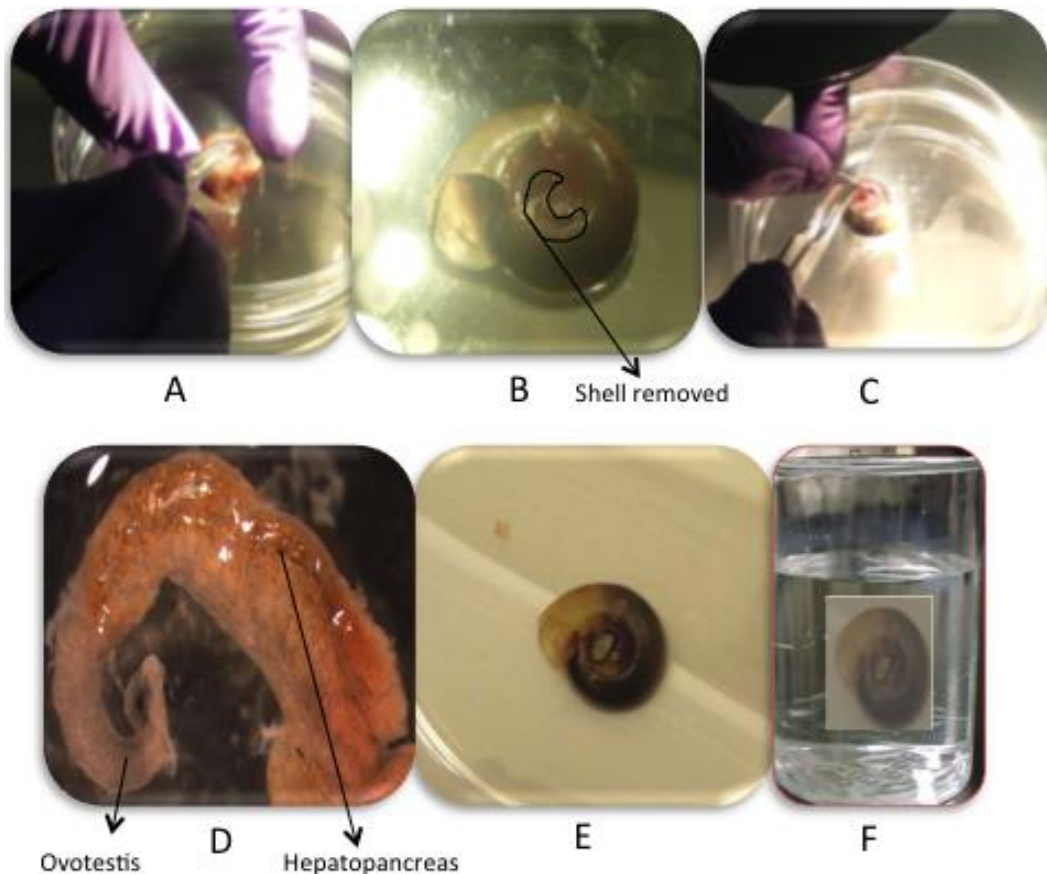
Figure 2.5: Technique of gonadectomy
A: Making an incision in the innermost whorl of the shell using a fine forceps. B: showing the incision exposing the ovotestis and hepatopancreas, C: the hermaphrodite duct being ligatured to prevent excessive blood loss, D: the light orangish region towards the tip is the ovotestis and the dark orangish part is the hepatopancreas, E: The incision sealed using autoclaved shell pieces from dead snail, F: the recovering snail.

The snails were weighed and their diameter was measured every week and any mortality noted. At the end of six weeks, the snails were euthanized by quick decapitation with a scalpel. Once euthanized, the shell was carefully crushed and removed and the wet weight of each snail was recorded. Dissections took place under a stereomicroscope (Leica); the visceral cavity was opened up and the albumen gland was removed, weighed and fixed in a minimum of three times the tissue volume of Bouin's solution (Sigma-Aldrich) for 1 hour, rinsed twice in Industrial Methylated Spirits (IMS, Charles Tenant) (70% IMS: 30 % distilled water) and then stored in 70% IMS at room temperature before processing. The histopathological processing and analysis was done on the albumen gland following the procedure described earlier in this chapter.

Statistical Analyses

All statistical analyses were carried out using Sigma-stat (v3.5). Shell diameter and total weight pre- and post-experiment were each compared by analysis of variance (ANOVA) followed by Dunnett's test to compare the treatment means with the respective controls.

Data sets found to lack variance homogeneity were alternatively subjected to non-parametric analyses (Kruskal-Wallis), followed by Dunn's post hoc test (Zar, 1999). The normalized albumen gland weight of the sham operated snails and the castrated snails at the time of termination of the experiment was compared by Student's T-test or Mann-Whitney U test depending on whether the data were parametric or non-parametric.

Results and Discussion

Maximum mortality was observed in the first week after gonadectomy. This could have been due to excessive hemolymph loss and fungal growth and prior to death the snails were much less mobile and did not consume food. The survivors behaved normally in terms of mobility and food consumption. After one week there were 30 survivors from the gonadectomised snails and 45 survivors from the 50 sham-operated snails. To keep the numbers similar, 30 from each gonadectomised and sham-operated snails were monitored and the remaining sham-operated snails were killed. In the gonadectomised snail tank, there were 5 mortalities in week 3 followed by 2 in week 4 and 1 in week 5 Figure 2.6.
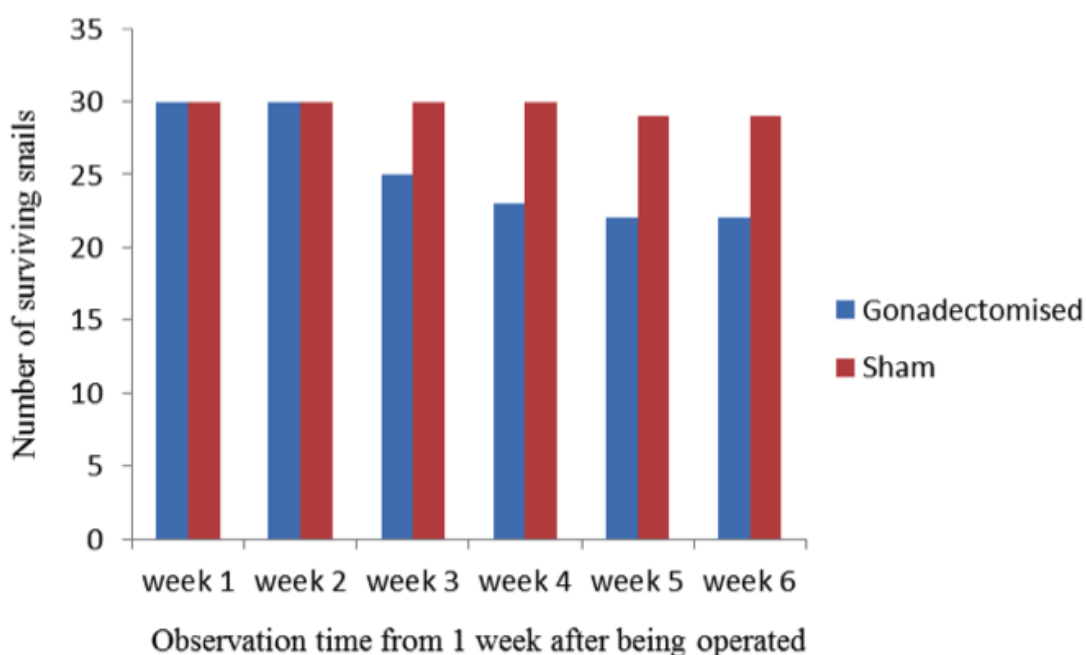
Figure 2.6: figure showing mortalities in gonadectomised and sham operated snails over a period of six weeks after a recovery period of one week post-gonadectomization.

In the sham-operated snail tank, only 1 snail died in week 5. At the end of the six week period there were 22 gonadectomised snails as compared to 29 sham-operated snails. In the gonadectomised snails the increase in mean weight was from 0.363g ±0.05 in the first week to 0.844g±0.19 in the sixth week while the sham-operated snail had mean weight of

0.441±0.064 in the first week that increased to 0.753±0.24 in the sixth week Figure 2.7. In the first week the gonadectomised snails had a mean diameter of 11.39mm± 0.08 that increased to a mean value of 15.79mm±1.84 in week six Figure 2.8. The sham-operated group had mean diameter of 11.38mm±0.19 in the first week that increased to a mean value of 15.62mm±1.78 in the sixth week.
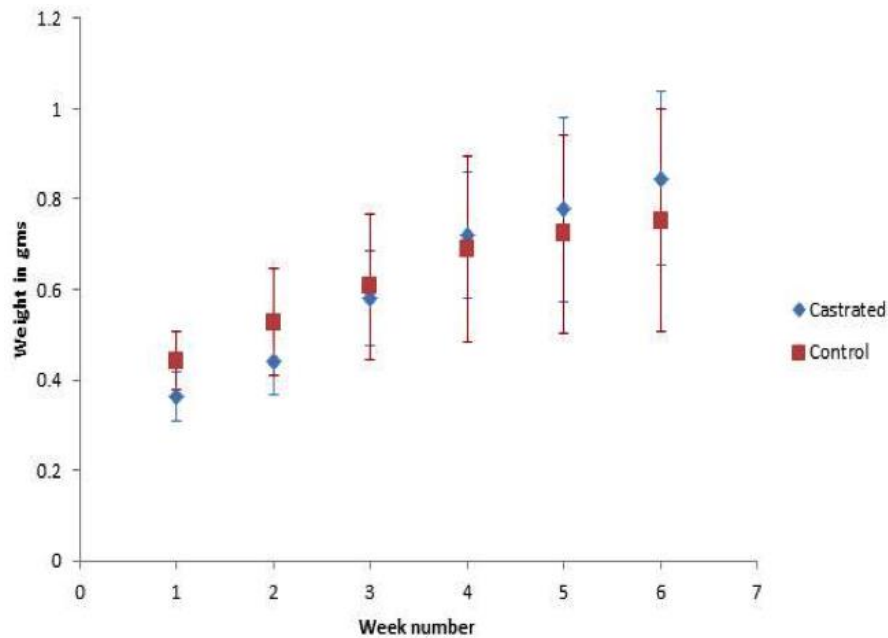
Figure 2.7: Change in weight measured in grams (mean ±SD) of gonadectomised and sham-operated (controls) snails.

During the first two weeks the gonadectomised snails weighed less than sham-operated snails Figure 2.7 that was likely due to loss of the gonadal tissue. A maximum increase in weight in the gonadectomised snails was observed between weeks three and four (24% increase) and then from week 4 onwards, their weight has risen over the sham-operated animals and this difference persisted until the end of the study period. Overall the gonadectomised snails showed an average increase of 132% in weight when compared to sham-operated snails with only a 71% increase. The overall average increase in diameter for the gonadectomised snails was 39% and a similar increase of 37% in diameter was measured in sham-operated snails. The increase in shell diameter and total wet weight of the gonadectomised snails was significant compared to the sham-operated snails ($P=<0.001$). These results are similar to earlier studies looking at the effects of castration on somatic growth in *Bulinus truncatus* (De Jong-Brink *et al.*, 1979), *Heloisoma duryi* (Miksys and Saleuddin, 1987) and *B. glabrata* (Vianey-Liaud and Dussart, 2002).
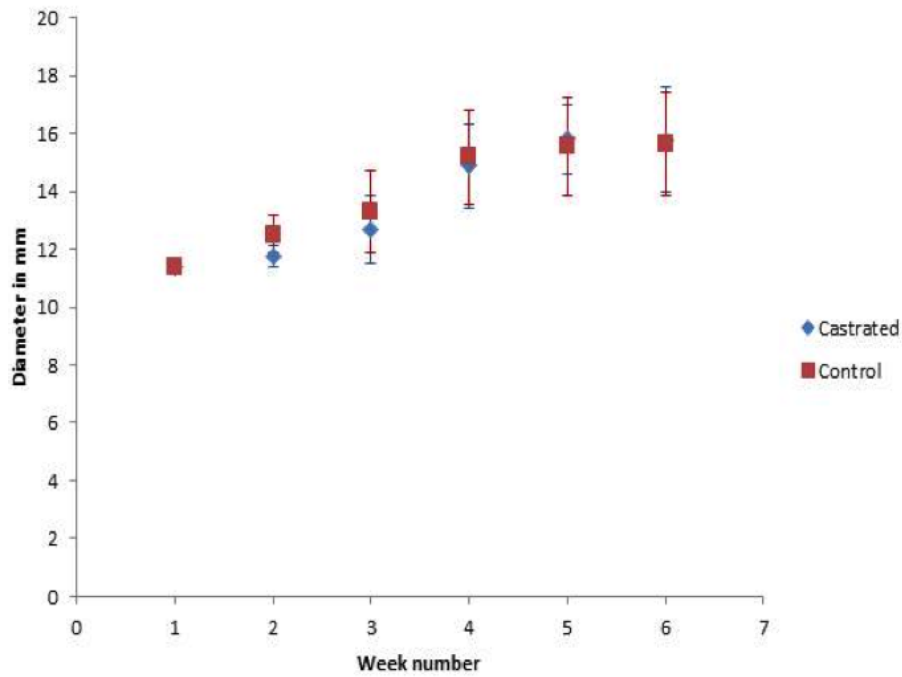
Figure 2.8: Change in diameter measured in grams (mean ±SD) of gonadectomised and sham-operated (controls) snails.

Castration in hermaphrodites seems to favor somatic growth due to reduced reproductive activity (Miksys and Saleuddin, 1987).

The weight of the albumen gland was normalized against the total body wet weight ($\frac{weight\ of\ albumen\ gland\ (g)}{Total\ wet\ weight\ of\ the\ snail\ body\ (g)}$) in both gonadectomised and sham-operated snails. The normalized albumen gland weight (NAW) for the sham-operated snails had a mean of 0.0419 but the gonadectomised snails had a larger albumen gland and the mean normalized albumen gland weight was 0.0739 Figure 2.9. There was a significant difference in the normalized albumen gland weight between the gonadectomised snails and the sham operated snails (controls) (P=<0.001). The increased body weight cannot be the only reason for the increased albumen gland weight in the gonadectomised snails. One of the reasons could be accumulation of secretory products (perivitelline fluid) due to reduced reproductive activity (Miksys and Saleuddin, 1987). A similar increase in the size of albumen gland post-gonadectomy has been reported earlier in *H. duryi* and *B. truncatus* (Miksys and Saleuddin, 1987; De Jong-Brink *et al.*, 1979). The smaller overall size as well as the smaller albumen glands is due to egg laying in the control snails compared to gonadectomised snails where all energy gained through food was invested in growth (Koene and Ter Maat, 2007).
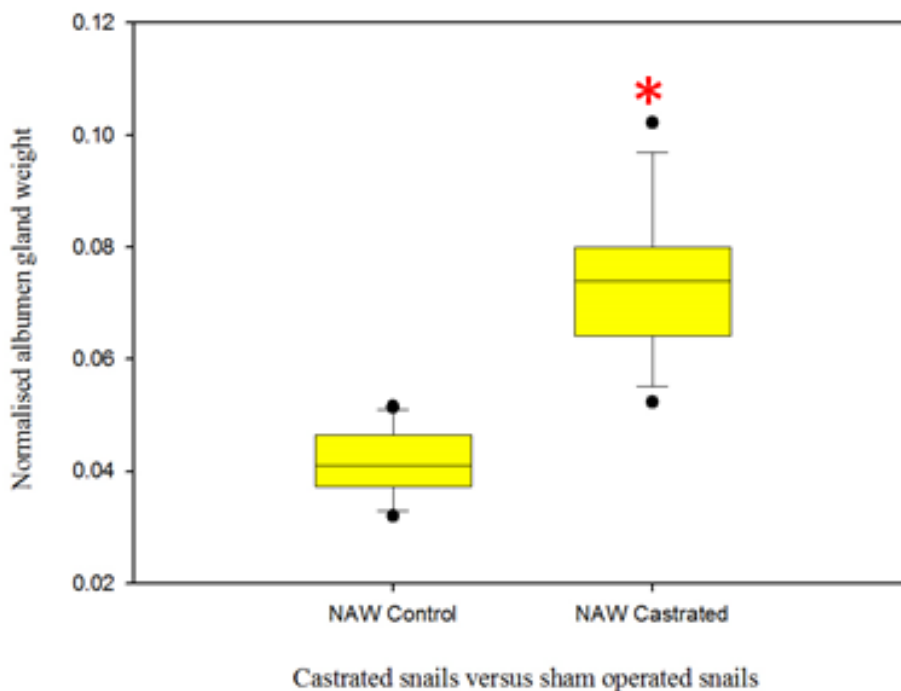
Figure 2.9: Boxplot showing the normalized albumen gland weight (NAW) of the gonadectomised snails compared to the sham operated snails after six weeks of observation. Boxes represent median with 25th and 75th percentiles. The bars extend to the 10th and 90th percentiles with outliers represented as dots. Sample size is indicated above each treatment label on the x-axis. * p< 0.05 compared to the Control group.

Thus it could be concluded that the technique for gonadectomy with a survival rate of ~51% was successful. Post-gonadectomy increase in body size (both body weight and shell diameter) and the increased albumen gland weight agree with previous studies on other species of snail (Miksys and Saleuddin, 1987; De Jong-Brink *et al.*, 1979).

2.2.2.2.3 Exposure of gonadectomised snails to 5α-dihydrotestosterone

The next step was to conduct a preliminary study exposing gonadectomised snails to androgen, DHT with the aim of investigating effects of DHT on the overall condition (mortality and growth) of the gonadectomised snails.

Materials and Methods

Test compounds

All chemicals were purchased in pure form (≥97%). The androgen 5α-dihydrotestosterone (DHT) (CAS No. 521-18-6) was purchased from Sigma Aldrich Co. Ltd (UK) and HPLC grade N, N-dimethylformamide (DMF) >99% (CAS No. 68-12-2) was purchased from Fisher Scientific, UK.

Stock solution preparation

Chemical super stock solution was prepared by dissolving 0.025g (weighed on an

electronic balance to four decimal places) DHT in 50ml of DMF making a concentration of 0.5 g/L and were stored in glass screw top bottles at 4°C until required to make dosing stock solutions. Dosing stocks were then prepared in 1litre amber glass bottles by spiking double distilled water with the desired volume of concentrated stock in DMF (Table 2.3). 1 litre of each stock solution was prepared with DMF as the carrier solvent in all stock solutions. Total DMF solvent concentration in all dosing stock bottles was maintained at 6%. Fresh stock solutions were prepared twice a week. Stock bottles (1 litre) were washed and were double solvent rinsed (using methanol and ethanol) prior to new stock storage.

Table 2.3: Volume of DHT added to prepare dosing stocks made up to 1L and nominal concentration in the tanks.

Dosing stock (DHT/MT)	0.5g/l DHT in DMF (ml)	100% DMF (ml)	ddH$_2$O (ml)	Total (ml)	Nominal tank concentration
500 μg/l	1	59	940	1000	500 ng/l (DHT/MT)
250 μg/l	0.5	59.5	940	1000	250 ng/l (DHT/MT)
125 μg/l	0.25	59.75	940	1000	125 ng/l (DHT/MT)
62.5 μg/l	0.125	59.875	940	1000	62.5 ng/l (DHT/MT)
0	0	60	940	1000	Solvent control

The continuous flow through dosing system

Dechlorinated carbon filtered (5 and 10 μm) tap water (27°C ±1) was delivered via flow control-meters into individual glass mixing chambers via medical grade silicon tubing (VWR, UK) at a rate of 200 mL/min (Figure 2.10). From each mixing chamber, water was evenly distributed into two replicate tanks at the rate of 100mL/min. Test compound dosing was controlled via a peristaltic pump (Watson Marlow, UK). Dosing stock solutions were pumped from 1 L amber glass bottles through pumpsil platinum-cured silicone tubing (Watson Marlow), to the mixing chambers at a rate of 0.2 mL/min, in order to achieve 1000 fold dilution and nominal concentrations in the exposure tanks of 62.5, 125, 250 and 500 ng/L (DHT). The solvent control (SC) tank received DMF at the same rate; such that the water in all tanks (except DWC) contained DMF at 0.006% v/v (below the suggested maximum solvent concentration of 0.01% v/v OECD limit for aquatic vertebrates OECD (2000) (http://www.epa.gov.uk/endo/pubs/ref-2_oecd) and gastropods OECD (2010) (http://www.oecd.org/). The glass exposure tanks (two per treatment) each had a working volume of 3.5 L (16x 15 x 23.5cm) and received approximately 41 tank volume renewals per day. Tank water was gently aerated via 2 pumps and 12 air stones to maintain optimum levels of dissolved oxygen. Dilution water and dosing flow rates were

monitored and if necessary corrected daily and dosing stock solutions were freshly prepared and replaced every 3 days to avoid any potential chemical degradation.
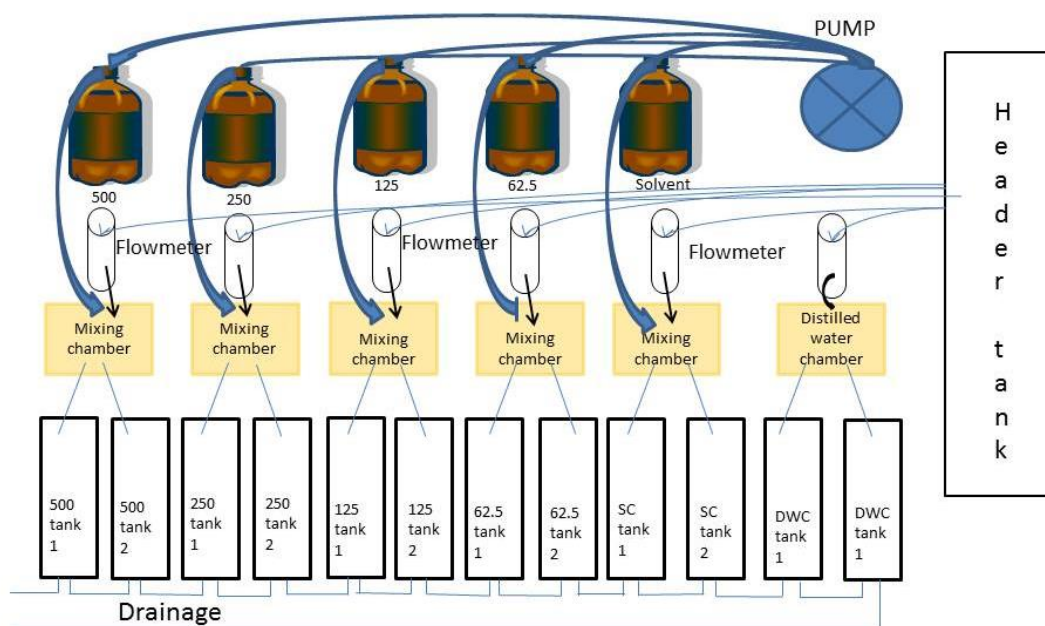
Figure 2.10: Diagram of the layout of the experiment in the exposure room.
The diagram shows the main water supply from the header tank flowed through 6 flow meters into 6 mixing chambers at the rate of 200ml/min. The chemical from the dosing bottle flowed into the mixing chamber at the rate of 0.2ml/min and then from the mixing chamber the diluted chemical flows into the tanks @100ml.min.

Experimental procedure

This preliminary study was designed to chemically expose 5 adult gonadectomised snails in each replicate tank (working volume of 3.5L) for 10 days following the "rodent castrated adult male assay" (Moon *et al.*, 2009), to observe the effects of the chemical on accessory reproductive glands. The study consisted of six treatments (tested in duplicate): Dilution water control (DWC), solvent control 0.006% DMF (SC) and DHT at nominal concentrations of 62.5ng/l, 125ng/l, 250ng/l and 500ng/l. Test concentrations were chosen to be in the range of androgen concentrations previously reported to affect gastropod molluscs and fish (Margiotta-Casaluci and Sumpter, 2011; Lyssimachou, Ramón and Porte, 2009; Watermann *et al.*, 2008; Schulte-Oehlmann *et al.*, 2006; Pawlowski *et al.*, 2004; Schulte-Oehlmann *et al.*, 2004; Tillmann *et al.*, 2001). The study was designed to chemically expose 5 adult gonadectomised snails in each replicate tank (working volume of 3.5L) for 10 days. Based on the previous gonadectomisation study survival rate, 140 snails were gonadectomised over a period of 2 days, to allow for expected mortalities (i.e. aprrox. 50% mortality expected) following the same gonadectomy protocol mentioned above in section 2.2.2.2.2. In total 60 gonadectomised snails were used in the DHT study

(6 treatments X 2 replicate tanks X 5 gonadectomised snails in each of the 12 tanks). At the end of 10 days of DHT exposure, the snails were removed from their exposure tanks, blotted dry, weighed (g) using an electronic balance and shell diameter measured to the nearest 0.1mm using hand-held digital callipers. Snails were then euthanized by quick decapitation with a scalpel. Once euthanized the shell was crushed, shell pieces carefully removed and each unshelled snail weighed.

Statistical Analysis

All statistical analyses were carried out using Sigma-Stat (v3.5). Shell diameter and total weight pre- and post-experiment were each compared by analysis of variance (ANOVA) followed by the Dunnett's test to compare the treatment means with the respective controls. Data sets found to lack variance homogeneity were alternatively subjected to non-parametric analyses (Kruskal–Wallis), followed by Dunn's post hoc test (Zar, 1999). A Spearman's Rank Order correlation between shell diameter and total weight of the snails was performed to determine overall condition of the snails.

Results and discussion

The weight of the gonadectomised snails in the different teeatments at the beginning of the experiment varied between 0.304-0.336±0.018 (g) and their diameter varied between 10.51-10.93±0.22 (mm). The average weight post-exposure varied from 0.283±0.008 (g) in solvent control to 0.351±0.045 (g) in DWC. In the DWC, 62.5ng/l DHT and 250ng/l DHT treatments the average weight increased but a decrease in average weight was observed in the SC, 125ng/l DHT and 500ng/l DHT Figure 2.11(A). A similar trend was observed for diameter except in SC, where even though the average weight post exposure decreased but diameter did show a slight increase (1.62%) Figure 2.11(B). Both the weight and diameter of the snails post-exposure was significantly positively correlated (R^2=0.69,P<0.001). Overall there were no significant differences observed in the average weight and average diameter of the gonadectomised snails in the different treatments post-exposure (ANOVA; P<0.05).
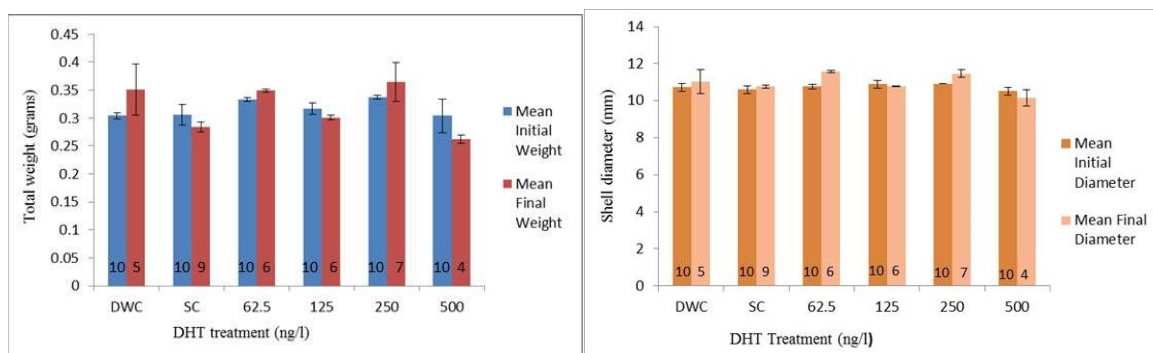
Figure 2.11: (A). Mean weight±SD and (B). mean diameter±SD of the snails pre- and post DHT exposure preliminary study.
A group of 60 sexually mature B. glabrata was exposed to 62.5, 125, 250 and 500ng/l nominal concentration of DHT.
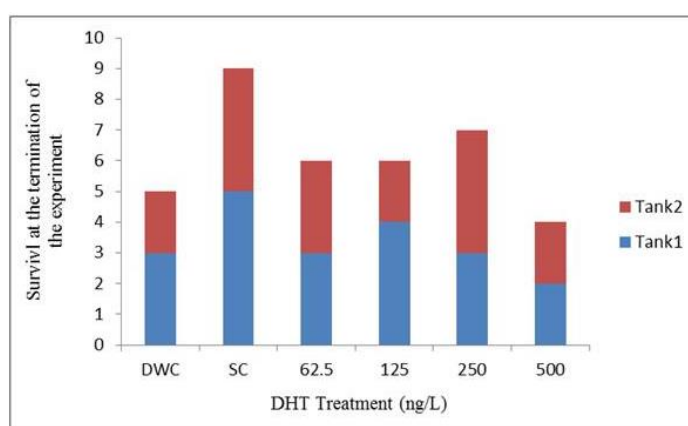
Figure 2.12: Survival of snails at the termination of DHT exposure preliminary study.
A group of 60 sexually mature gonadectomised B. glabrata exposed to 62.5, 125, 250 and 500ng/L of DHT for 10 days.

High mortality was recorded over the duration of the exposure, with only 5 survivors (2 in tank1+3 in tank 2) in DWC, 6 survivors in 62.5ng/l DHT (3 in tank1+3 in tank 2) and 125ng/l DHT (4 in tank1+2 in tank 2), 4 survivors in 500ng/l (2 in tank1+2 in tank 2) (Figure 2.12). The surviving snails were not very healthy, which was evident by the fact that they were not consuming food and were lethargic (not moving actively in the tanks). These surviving snails had pale hemolymph, as observed at the time of dissection.

The initial aim of this study was to quantify any changes in the accessory reproductive glands post-exposure. Due to low survival there was not enough data to provide statistically meaningful results. The survivors were sick and so could not be truly representative. The results obtained during this preliminary study indicated the likely unreliability of using the gonadectomised snail model. Another factor was that the number of snails needed to be gonadectomised for the main study would be approximately 500, which would be very labour intensive taking approximately seven days (approximately 70 snails per day). Taking together the work involved and the mortality rate during the

exposure study it was not considered feasible to have a gonadectomised snail version of the mammalian Hershberger test.

2.2.3 Conclusion

All the information gained about the molluscs and the test species through the available literature was vital for my PhD project as it would set the ground for the main exposure studies and the bioinformatics investiogation in *B. glabrata* and *L. gigantea*, that will be discussed in the later chapters. The preliminary experiments conducted in *B. glabrata* provided baseline information to compare the reproductive tract development of the experimental snails. The protocol gor gonadectomising snails was well optimized but the pilot study provided vital evidence that led to the decision of excluding gonadectomy as a part of my exposure study protocol.

3 Chapter 3: Exposure of the freshwater snail *Biomphalaria glabrata* to steroid androgens during development does not affect reproductive physiology

3.1 Introduction

Androgens are steroid hormones that regulate a wide range of physiological processes in male vertebrates, often mediated through the androgen receptor (AR). Various synthetic androgens and androgen antagonists can therefore act via the androgen receptor and are used in pharmaceutical medicines for treatment of some common diseases, such as androgen deficiency and prostate cancer, respectively (Tabb and Blumberg, 2006). In pre-clinical and safety testing of these pharmaceutical drugs, the Hershberger assay, originally developed to identify the drivers of male sexual development, is the gold standard test (Hershberger, Shipley and Meyer, 1953). This assay employs either immature or castrated peripubertal male rats, which are dosed with the compound to be tested over a number of days (3-10 days depending on assay). At the end of the exposure the reproductive organs and tissues that are known targets for androgens are assessed (e.g. weight, morphology, histopathology etc.). Chemicals, which increase organ weights are considered stimulatory, i.e. a compound, which increases the weight of the prostate in the Hershberger assay, would be considered androgenic. Similarly a reduction in target organ weight would suggest anti-androgenic or inhibitory properties of the chemical. These types of mammalian assays, including pubertal assays (to detect developmental effects) have been validated by the Organisation for Economic Co-operation and Development (OECD) and the United States Environment Protection Agency (US EPA) for screening suspected endocrine disrupting chemicals for androgenic and anti-androgenic activity (Moon *et al.*, 2009; Owens *et al.*, 2007; Shin *et al.*, 2007; Owens *et al.*, 2006).

New legislation concerning chemical safety testing will undoubtedly involve increased numbers of vertebrate animal tests and will bring with it significant financial and ethical costs. As a consequence fish tests are also on the increase, as they are part of the new environmental risk assessment process involved in the safety testing of pharmaceuticals and industrial chemicals both for toxicity and endocrine disrupting properties. As discussed earlier in chapter 1, the Annual Statistics of Scientific Procedure on living animals, Great Britain (2012) stated that over 2.13 million animals were used in scientific experimentation in 2012 and the estimated total cost for conducting EDSTAC Tier 1 and Tier 2 screening for a chemical sums up to approximately $900000 (approximately over £540,000). Considering the vast amount of chemicals ultimately to be tested will involve a huge expenditure. Now more than ever, there is increased pressure to implement the three R's (reduction, refinement and replacement) in relation to the testing of chemicals and medicines and in general biomedical research. By exploiting the use of small invertebrates

as surrogate mammalian and fish test organisms the number of vertebrates can be reduced or in some cases replaced. This has been successfully achieved with *C. elegans* and *D. melanogaster* in, for example, the fields of ageing, development and physiology (Harrington *et al.*, 2010; Ghosh- Roy and Chisholm, 2010; Beckingham *et al.*, 2007; Roberts, 2006). Neither of these organisms, however, have male reproductive anatomy remotely similar to that of the mammals. Even in the non-mammalian vertebrate classes (e.g. fish), suitable models for mammals have not been found, with respect to the existence of a prostate and other accessory sexual glands used as endpoints in the Hershberger assay. Notwithstanding this, very few invertebrates have been investigated as potential surrogates of male or female reproductive function, for either biomedical or chemical safety testing.

At the time this research was initiated, homologues of androgens and steroidogenic androgen and estrogen producing enzymes had, been reported in the gonads and digestive tube of several invertebrate groups (Markov *et al.*, 2009; Lafont and Mathieu, 2007; Lavado, Janer and Porte, 2006a; Lehoux and Sandor, 1970), including molluscs and annelids. In addition molluscs have a reproductive anatomy that is, at least superficially, similar to that of vertebrates. Like mammals, the snail's male reproductive tract extends from the testis (or ovotestis in hermaphrodite snails), has seminal vesicles, and a prostate gland and terminates in a penile glandular complex (around the preputium) containing sebaceous glands and a penis (Adiyodi and Adiyodi, 1988) (discussed in detail in Chapter 2).

There are also various examples in the literature that suggest developmental exposure of gastropod snails (*Marisa cornuarietis*) to the synthetic androgen, 17α-methyltestosterone (MT) or to the environmental pollutant "androgen" tributyltin (TBT) causes hyperplasia in the epithelial tissue of reproductive glands and inhibition of spermatogenesis in males (Schulte-Oehlmann *et al.*, 2006; Schulte-Oehlmann *et al.*, 2004). Indeed TBT and triphenyltin (TPT) have been documented to alter female reproductive tissue (the albumen gland) to histologically resemble male reproductive tissue (the prostate gland) in some gastropod snails (Watermann *et al.*, 2008; Albanis *et al.*, 2006). Female adult *M. cornuarietis* (a freshwater prosobranch gastropod mollusc) developed imposex (imposition of penis growth in females) after 150 days of MT exposure (Janer *et al.*, 2006a; Janer *et al.*, 2005b) and adult pond snails, *Lymnaea stagnalis* (a pulmonate gastropod), demonstrated weak histological damage to the albumen and prostate glands after 8 weeks of exposure to 100ng/L of MT (Czech, Weber and Dietrich, 2001). Reduction in fertility and embryo production was observed in *Potamopyrgus*

antipodarum, *M. cornuarietis* and *Nassarius reticulates* when exposed to environmentally relevant concentrations of MT (Duft *et al.*, 2007). Similar "androgenic" effects were reported to be significantly reduced by co-exposure to the anti-androgen cyproterone acetate and to an aromatase inhibitor (1-methy-1, 4-androstadiene-3, 17-dione) (Ketata *et al.*, 2008; Oberdörster and McClellan-Green, 2002; Bettin, Oehlmann and Stroben, 1996), suggesting the existence of vertebrate-like mechanisms of male reproductive function in molluscs (Hagger *et al.*, 2009). Moreover, exposure of males from a number of gastropod molluscan species (*Marias cornuarietis, Nucella lapillus and Hinia reticulata*) to 17α-MT caused a reduction in the length of the penis and penis sheath (Watermann *et al.*, 2008; Oehlmann *et al.*, 2007; Tillmann *et al.*, 2001), adding further strength to the suggestion that mammalian-like mechanisms through which androgens and anti-androgens could act are present in molluscs. These documented impacts (mentioned in detail in Chapter 1) raise the possibility that the effects of androgenic and anti-androgenic chemicals on reproduction and development in molluscs might be comparable to those in mammals (Oehlmann and Schulte-Oehlmann, 2003) and that snails may hold promise as a model for male reproductive diseases and for testing medicines and chemicals for interference with male reproductive system.

As the basis of the mammalian Hershberger assay is the absolute requirement of the male reproductive tract for testosterone (T) and/or dihydrotestosterone (DHT), it seemed sensible to test the sensitivity of *B. glabrata* reproductive tract to androgens. In rodents, the Hershberger assay has been tested at three different life stages, the prepubertal intact male (Ashby and Lefevre, 2000), the castrated adult male and the peripubertal castrated male (Gray, Furr and Ostby, 2005). The choice of time points is based on the fact that accessory sex glands and tissues (seminal vesicles and prostate, coagulating glands and Cowper's glands, preputial glands and glans penis) are dependent upon androgen stimulation to gain and maintain weight during and after puberty (Owens *et al.*, 2006).

3.2 Aim of the study

The aim of this study was to investigate if an invertebrate "Hershberger-type" test could be developed. These tests included intact unhatched snails exposed until sexual maturity and adult castrated (gonadectomised) snails to replicate the "rodent prepubertal intact assay" and the "rodent castrated adult male assay" respectively. These tests employed 5α-Dihydrotestosterone (DHT) and 17α-methyltestosterone (MT) as model androgenic compounds, to see if similar stimulatory effects could be produced in developing and adult snails as would be expected in rodents.

5α-Dihydrotestosterone (DHT) is one of the most physiologically important and active androgens in many male vertebrates (George, 1997), and is formed by the 5α−reduction of testosterone by the enzyme 5α-reductase. Both 5α-reductase enzyme activity and endogenous levels of testoterone have been previously reported in molluscs (Fernandes, Loi and Porte, 2011; Lavado, Janer and Porte, 2006a; Janer, LeBlanc and Porte, 2005a), even though their functions have not been clearly demonstrated, as the knowledge of steroid metabolism in invertebrates is still limited and fragmentary. The drug, Finasteride has been shown to inhibit 5-alpha reductase activity in the ovaries and testis of the sea urchin *Lytechinus variegatus* with an IC50 of 2.7 μM (Wasson *et al.*, 2000; Wasson and Watts, 1998b). Moreover, 5-alpha reduced metabolites were found in the gonads and digestive tube of another sea urchin, *Paracentrotus lividus* when exposed to 10 μM Fenarimol imidazole-like fungicide (Lavado *et al.*, 2006b; Janer, LeBlanc and Porte, 2005a). Subcellular fractions isolated from the visceral coil of the gastropod snail *M. cornuarietis* reported 5-alpha reduced metabolites such as 5α-androstenedione and 5α DHT (Janer, LeBlanc and Porte, 2005a) indicating that steroid 5α reductase activity appears to be involved in androgen metabolism in these invertebrate species.

17α-methyltestosterone (MT) was tested in addition to DHT, as it is known to have similar binding affinity as T and the addition of the 17 α-alkyl group has been shown to slow metabolism at the 17-position and prolong its *in vivo* half-life (Gao, Bohl and Dalton, 2005). This synthetic androgen is widely applied in aquaculture to control sex determination and sex-reversal of genetic female fish to phenotypic males (Vick and Hayton, 2001; Kitano *et al.*, 2000; Papoulias, Noltie and Tillitt, 2000; Hunter and Donaldson, 1983) and is therefore considered to be an endocrine disruptor in fish (Rivero-Wendt *et al.*, 2013; Kang *et al.*, 2008; Pawlowski *et al.*, 2004; Hornung *et al.*, 2004). The specific mechanism through which MT acts in invertebrates remains to be resolved, but dose, timing and duration of treatment can influence the effects, such as inducing masculinization in prosobranch snails and crustaceans (Fujioka, 2002; Czech, Weber and Dietrich, 2001). In another study conducted by Schulte-Oehlmann and colleagues (2004), 0.1μg/L MT caused virilization in female *M. cornuarietis* and also affected the formation of germ cells in male and female gonads (Schulte-Oehlmann *et al.*, 2004).

The aim of this study, therefore, was to test the hypothesis that common physiological processes might underlie the response of the male reproductive system to androgens in both molluscs and mammals.

3.3 Experimental Approach

The overall approach was to expose *B. glabrata* to two potent androgens; the natural vertebrate steroid, 5α-dihydrotestosterone (DHT) and the anabolic steroid 17α-Methyltestosterone (MT), after which reproductive tissues and organs were assessed for stimulatory (e.g. male tissue reproductive growth or advanced spermatogenesis) or possible disruptive (e.g. to female reproductive tissues or oocyte development) effects. Due to the unreliability of the gonadectomised model (high rate of mortality and unhealthy snails), it was not considered feasible to have a gonadectomised snail version of the mammalian Hershberger test. In vertebrates, early life and pubertal stages are more susceptible to endocrine interference by natural or exogenous steroid androgens (Mensink et al., 1996) than in adulthood. Therefore, to investigate possible androgenic effects in *B. glabrata*, the snails were developmentally exposed with the expectation that a similar 'sensitive window' during sexual development would also present in our test species.

3.3.1 Test compounds

All chemicals were purchased in pure form (≥97%). The androgen 5α-dihydrotestosterone (DHT) (CAS No. 521-18-6) was purchased from Sigma Aldrich Co. Ltd (UK) and HPLC grade N, N-dimethylformamide (DMF) >99% (CAS No. 68-12-2) was purchased from Fisher Scientific, UK.

3.3.2 Stock solution preparation

Chemical super stock solution was prepared by dissolving 0.025g (weighed on an electronic balance to four decimal places) either MT or DHT in 50ml of DMF making a concentration of 0.5 g/L and were stored in glass screw top bottles at 4°C until required to make dosing stock solutions. Dosing stocks were then prepared in 1litre amber glass bottles by spiking double distilled water with the desired volume of concentrated stock in DMF (Table 3.1). 1 litre of each stock solution was prepared with DMF as the carrier solvent in all stock solutions. Total DMF solvent concentration in all dosing stock bottles was maintained at 6%. Fresh stock solutions were prepared twice a week (Figure 3.1). Stock bottles (1 litre) were washed and were double solvent rinsed (using methanol and ethanol) prior to new stock storage. Personal protective equipment was worn while preparing and handling stock solutions.

Table 3.1: Volume of Chemical (DHT/MT) added to prepare dosing stocks made up to 1L and nominal concentration in the tanks.

Dosing stock (DHT/MT)	0.5g/l DHT in DMF (ml)	100% DMF (ml)	ddH$_2$O (ml)	Total (ml)	Nominal tank concentration
500 µg/l	1	59	940	1000	500 ng/l (DHT/MT)
250 µg/l	0.5	59.5	940	1000	250 ng/l (DHT/MT)
125 µg/l	0.25	59.75	940	1000	125 ng/l (DHT/MT)
62.5 µg/l	0.125	59.875	940	1000	62.5 ng/l (DHT/MT)
0	0	60	940	1000	Solvent control

3.3.3 Experimental Design

Two consecutive 30 days exposure studies (one for DHT and one for MT) were conducted in April-May, 2011. Each experiment was designed to have 20 hatchlings in each tank, in six treatments, tested in duplicate (Figure 2.10) up until sexual maturity (240 animals in total per chemical). The range of doses for exposure was informed by the doses used in the preliminary adult castrated snail study which itself had been informed by literature on fish developmental studies (Margiotta-Casaluci and Sumpter, 2011) and other mollusc research (Lyssimachou, Ramón and Porte, 2009; Watermann *et al.*, 2008; Schulte-Oehlmann *et al.*, 2006; Schulte-Oehlmann *et al.*, 2004; Pawlowski *et al.*, 2004; Tillmann *et al.*, 2001). Based on this, the six treatments for the experiment (tested in duplicate) were dilution water control (DWC), solvent control 0.006% DMF (SC), and MT or DHT at nominal concentrations of 62.5 ng/L, 125 ng/L, 250 ng/L, and 500 ng/L. Tanks were pre-dosed for 3 days before introducing egg masses to allow for the dosing to equilibrate and water samples were collected for analysis prior to and during the snail exposure (see Figure 3.1).

Developmental exposure study Gantt chart
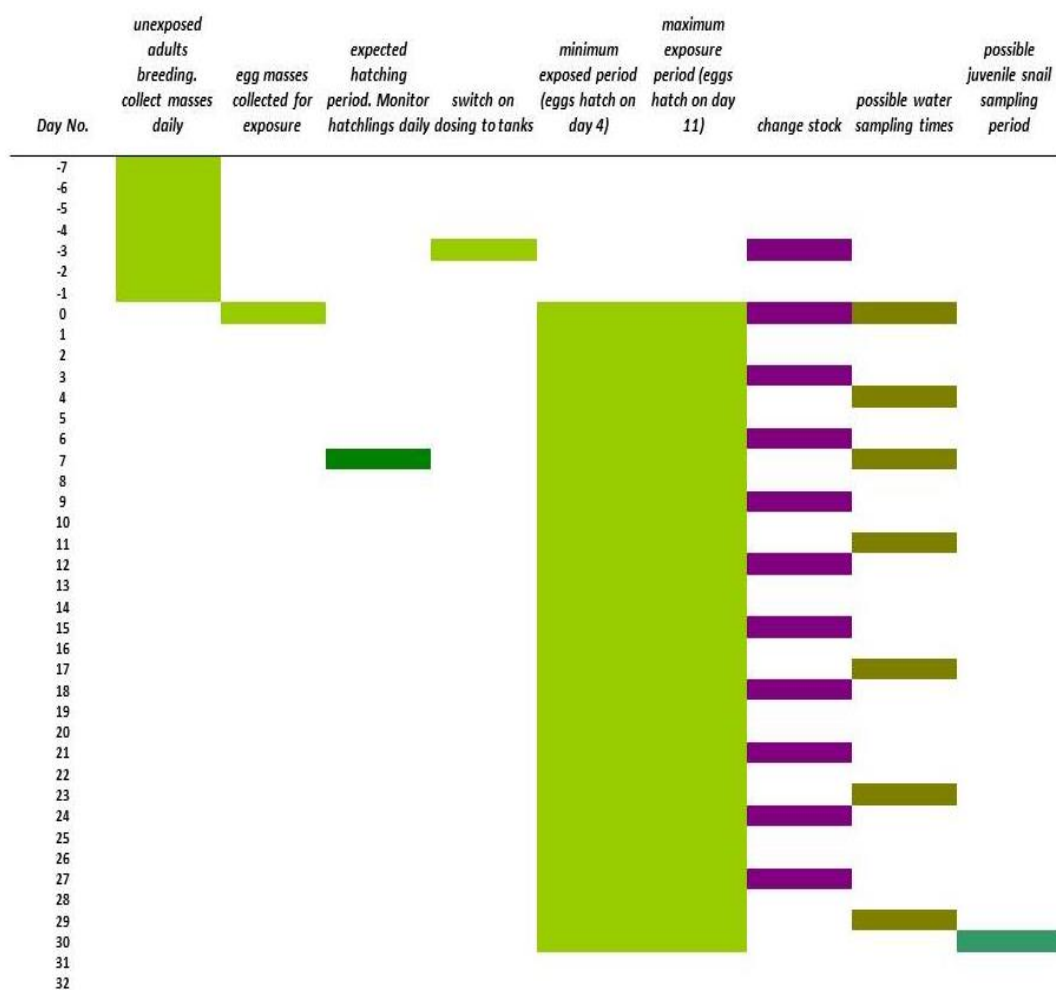
Figure 3.1: Gantt chart showing the experimental design for the 30 days exposure study using B. glabrata.

3.3.4 Egg collection and initial static-renewal exposure of embryos

Snails were exposed to each of the test chemicals, MT or DHT, throughout embryonic and early development until sexual maturation (age at which egg laying begins) total of 30 days. 24 egg masses, laid on the same day, were collected from a breeding group of 30 adult *B. glabrata* (BB02). Egg masses were carefully removed from the glass walls of the parental tank using a scalpel. Each egg mass was inspected under dissecting microscope for any damage prior to being included in the study. *B. glabrata* naturally hatch out of the egg mass over a period of days (i.e. within a single egg mass some hatchlings will emerge before others), however for this experiment design, to remove variations in age and development at the point of sampling, hatchlings were required to be monitored (and selected) over the hatching period. Therefore, egg masses were initially exposed to the test solutions individually within wells of six-well plates. The multi-well plates (nunc, sterile non-treated) were floated on top of the flow through exposure tanks to maintain

temperature and light regime consistency. Four egg masses were exposed per treatment. Each egg mass was inspected daily for development and hatching and 50% of the dosed water (3ml total volume) was renewed every 24 hours. All newly hatched snails were carefully separated from egg masses (and continued to be dosed in a new well) and the date that they hatched was recorded. The majority of snails hatched on days 6-7, at which point 20 hatchlings were collected and placed into the appropriate tank (40 snails per treatment), enabling all hatchlings to be the same age at the point of sampling and for the density of snails to be fixed at the start of the experiment. After hatchling snails were placed into the tanks (day 7 of the exposure) and were exposed under continuous dosing (flow-through) conditions for 23 days until sexual maturation.

3.3.5 Tank water analyses

Water samples (250 mL) were collected, in amber glass bottles, from each tank on day 0 (day before egg-masses were exposed), and subsequently every 7 days until the end of the experiment (Figure 3.1). The target chemical was extracted from the water sample by solid phase extraction (SPE) using C18 SPE cartridges (Sep-pak Plus, Waters Ltd., Watford, UK) and a Supelco SPE manifold (Sigma Aldrich) as per the manufacturers instructions. Additional "blank" and "spiked" samples were also taken to check SPE efficiency, these were double distilled water 'blanks', MT/DHT 'spiked' using10µl of 0.5g/l chemical in 250ml double distilled water. These were extracted at the same time and using the same SPE method as the 250 ml tank water samples.

3.3.5.1 *Extraction and concentration of androgenic chemicals from water samples*

Extraction and concentration of androgenic chemicals from water samples was accomplished using a 'visiprep' vacuum manifold (Supelco) and reverse-phase C18 cartridges (Sep-Pak C18, Waters Associates, inc). A silicone liner was inserted into each port of the manifold and C18 cartridge was positioned on the top of the liner. A tube adapter and plastic reservoir were placed above the cartridge. The vacuum unit was then attached to the manifold and turned on. 5ml of double distilled water was added to each sample well, followed by 5ml of methanol and then a further 5ml of double distilled water to condition the C18 cartridge (these were sucked thought the cartridge by the vacuum). While conditioning, it was important that the C18 cartridge did not dry out so another 5ml of double distilled water was filled in the reservoir while the well bung was being applied to the reservoir. This bung had an attached plastic straw, which was suspended into the sample bottles, and weighed down by a stainless steel weight at the other end. Once the vacuum was started and taps opened, water was sucked through the plastic straw (at the

rate of ≤5ml/min), into the reservoir, and through the C18 cartridge (where the test chemicals were bound). The water sample dripped through the disposable liner (attached to the C18 cartridge) and into the base of the manifold. The water collected at the bottom of the manifold was removed through a rubber tube to a collecting vessel for disposal. Once the entire water sample had passed through the C18 cartridge, the bung and attached straw were removed. The vacuum was left on for at least 30 minutes to remove any water from C18 cartridge. 7ml glass vial double solvent rinsed were placed inside the manifold beneath the C18 cartridge and disposable liners taking care that the cartridge/liner/vials were properly aligned. The C18 cartridges were eluted using 5ml of analytical grade methanol (Fisher scientific, UK) under a vacuum @ approximately 2ml/min. Then the methanol-elutant was dried (evaporated) under nitrogen gas using the supelco 'visidry' attachment for the extraction manifold. The extract was then re-suspended in 1 ml of analytical grade ethanol (Hayman, UK). The concentration factor calculated as $(\frac{Volume\ of\ water\ extracted}{Volume\ of\ ethanol\ in\ which\ the\ sample\ was\ resuspended})$ was recorded. Vials containing extract in ethanol were stored with tight fitting lids at 4ºC to prevent evaporation until further analysis using a recombinant Yeast androgen screen (YAS) (Sohoni and Sumpter, 1998).

3.3.5.2 *Yeast Androgen Screen (YAS) Assay*

The yeast assay is a receptor/reporter assay using recombinant yeast plasmid for detecting estrogenic or androgenic activity of the chemical compounds. The activity can be monitored by eye and measured by spectrophotometer. The assay was carried out in a type II laminar flow cabinet (to minimize aerosol formation) following the standard protocol (Sohoni and Sumpter, 1998).

A day before the YAS assay was conducted, yeast growing medium was prepared (Appendix 3.1). 125 µl of 10X concentrated yeast stock, from a cryogenic vial, was added to prepared sterile yeast growth medium and then incubated at 28^0C for approximately 24 hours on an orbital shaker (250 r.p.m). Chemicals in the extract (prepared from water extraction in the earlier section) were serially diluted in absolute ethanol in a 96-well microtitre plate. To start with, 100µl aliquots of absolute ethanol were pipetted into each row of wells, except for the first well. In the first and the second well 100µl of sample was pipetted. The 200µl mixture of extract and ethanol were mixed using the pipette and then 100µl of the 50:50 mixtures of extract and ethanol was removed from the second well and transferred to the third well. This was mixed carefully using a pipette and then 100µl was taken from this well and transferred into the fourth well and this process was repeated until all 12 wells had been completed (the final well contained 100µl of absolute ethanol

and 100µl of the mixture). The dilution series for the Dihydrotestosterone (DHT) and 17α-methyltestosterone (MT) standard curves were prepared in similar manner (DHT starting concentration 1×10^{-6} M and MT 6.6×10^{-7} M) (Figure 3.2). MT (6.6×10^{-7} M) showed similar activity in the YAS screen with DHT standard solution and so this concentration was selected as the starting concentration for MT. 10µl aliquot of each concentration was then transferred from the dilution plate to a new 96-well microtitre assay plate, starting from the most dilute concentration from the right side of the dilution plate. Each sample dilution was replicated so two identical rows were produced and to avoid cross contamination, between each double row of extract one row was left completely empty (blank). DHT and MT were used as positive controls in their respective screens and a negative control (absolute ethanol) was used in both the experiments. The purpose of the negative control was to account for any kind of activity due to possible chemical contamination in the ethanol. Once all assay plate dilutions were complete the ethanol was evaporated to dryness.
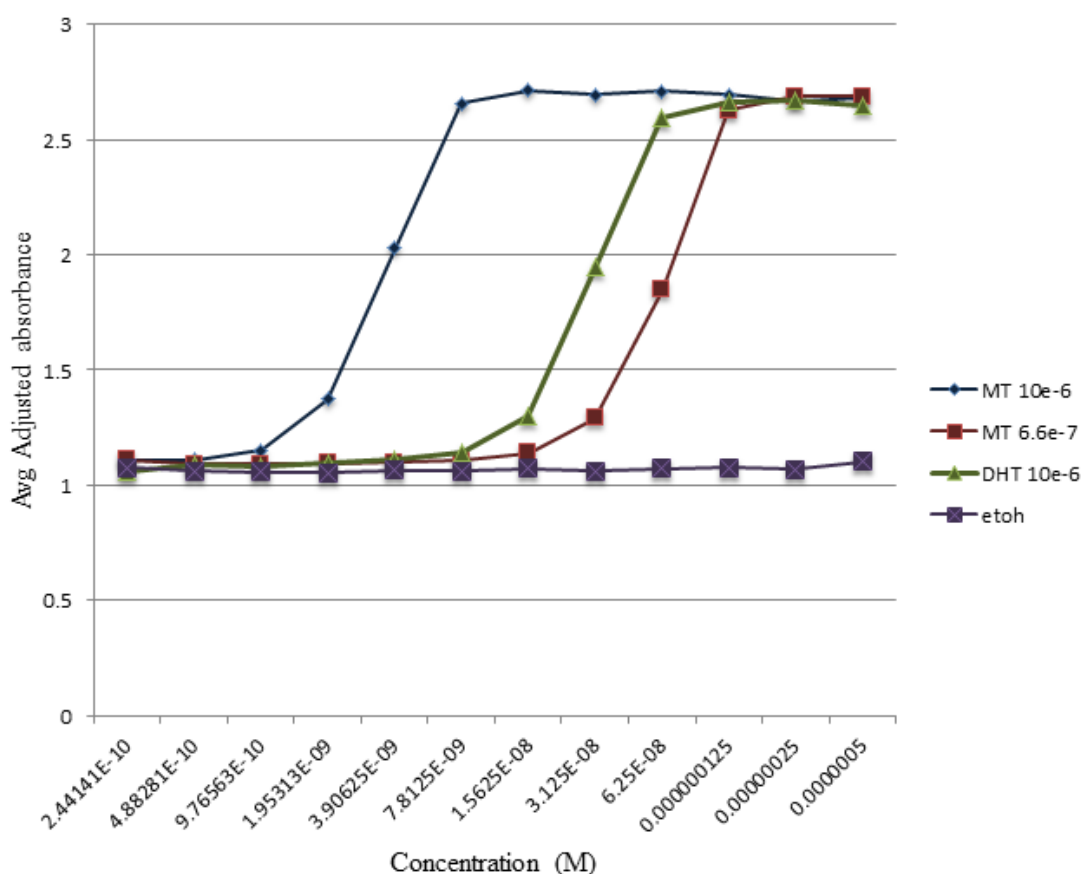
Figure 3.2: Standard curves showing the starting molar concentrations of DHT and MT with respect to ethanol.

While the assay plates dried the final assay medium was prepared as per the primary

culture (appendix 3.1) with the addition of 0.5ml of Chlorophenol red-β-D-galactopyranoside (CPRG) and 4×10^7 yeast cells from the 24-h culture prepared the previous day, per 50ml of growth medium. The volume containing 4×10^7 yeast cells was calculated by taking two 100μl aliquots of the 24-hour culture, in wells of a 96-well microtitre plate and the absorbance at 640nm were measured using a spectrophotometer plate reader. The mean of the two readings obtained was used to extrapolate the number of cells per ml using the absorbance graph. The volume required to give 4 x 107 cells per 50ml of growth medium was calculated using the following equation.

$$\frac{4 \, X \, 10^7}{Number \, of \, cells \, calculated \, from \, the \, absorbance \, graph} = Vol \, of \, 24 \, hour \, culture$$

200μl of the yeast inoculated final assay medium was then added to each well of the assay plate using a multichannel pipette. To prevent evaporation of the medium (which might change the concentration of CPRG) and therefore the colour of the background, each plate was sealed round its edges using autoclave tape. The plates were then agitated briskly for 2 minutes on a titre plate shaker before being incubated at 32°C in a naturally ventilated heating cabinet. The plates were removed after 24 hours and incubated overnight at 28^0C to avoid a significant increase in background expression of the enzyme β-galactosidase. Colour development of the medium was checked at regular intervals using a Titertek Multiscan MCC/340 spectrophotometer plate reader. The plates were read at an absorbance of 540nm (for colour) and 620nm (for turbidity). Following incubation, a deep red colour indicates positive activity and due to slight background production of β-galactosidase, the negative control wells appear light orange in colour. Growth of the yeast is accompanied by a turbid appearance. Therefore the corrected value = chemical absorbance at 540nm-(chemical absorbance at 620nm-blank absorbance at 620nm) to correct for the turbidity.

3.3.5.3 *Calculating androgenic equivalents*

The androgenic equivalents were calculated by first plotting the corrected absorbance values, the DHT standard curve and ethanol blank (to check for contamination), against the relative concentration of each extract. When plotted the DHT standard and extract dilutions produced sigmoidal curves. Only the linear part of the graph was used for calculations. The midpoint of this linear part was noted and then two corrected absorbance values on the either side of the selected midpoint were obtained. Based on the well numbers of these two values, the corrected absorbance was extrapolated using the standard

curve and the concentration of the first well was calculated (concentration before dilution). The concentration obtained were then multiplied by the molecular weight of DHT (290.44)/ MT (288.424), to provide a concentration in weight per volume, rather than a Molar concentration. The concentration (weight per volume) was then multiplied by 20 to take into account the volume of assay medium. To obtain the concentration of the chemical in initial water sample in g/l, these values were divided by the original concentration factor. An average value of the two concentrations was used as the final androgenic equivalent concentration of the chemical in the sample.

3.3.6 Biological Sampling and analyses

3.3.6.1 *Terminal sampling*

At the culmination of the experiment (after 30 days of exposure), snails were removed from their exposure tanks, blotted dry, weighed (g) using electronic balance and their shell diameter measured to the nearest 0.1mm using hand-held digital callipers. Snails were then euthanized by quick decapitation with a sterile scalpel. Once euthanized the shell was carefully crushed and removed. Once the snail was removed from the shell it was washed in a petri dish containing distilled water to remove any remains of shell pieces. When handling the de-shelled snail with forceps, mantle was used to hold the tissue, to avoid damaging any reproductive organs. The soft body weight of the whole snail was recorded and were quickly fixed in RNAlater and stored at 4°C in 2ml cryovial tubes (Nalgene).

3.3.6.2 *B. glabrata dissection and preservation*

Each snail fixed in RNAlater was carefully dissected under stereomicroscope (Leica), the visceral cavity was opened up and the albumen gland, glandular complex and the ovotestis were removed Figure 3.3. The ovotestis is situated in the innermost whorls of the shell and is dorso-laterally covered by the digestive gland (hepatopancreas). Using forceps and a sterile scalpel the ovotestis was carefully teased away from the digestive tissue, weighed on an electronic balance up to 4 decimal places, weight recorded on the data sheet and cut transversally into two equal sized pieces, placing one half into the new RNAlater cryovial and the other into Bouin's fixative. The albumen gland or the female accessory reproductive gland is a large tubulous gland, that can be distinctly seen above the hepatopancreas as a ball like or folded structure, was removed using forceps (Figure 3.3), weighed and divided into two equal sized pieces for molecular analysis and for histopathological analysis. The glandular complex (GC) is the term used to describe the prostate diverticulum, membrane gland (oothecal gland) and mucous glands. This term has

been previously used in the literature to describe the albumen gland and the mucous gland (Pal, 2007). Due to its complex nature (i.e. 3 gland types intertwining into each other) GC were only dissected and preserved for histopathology and not split between molecular and histopathology analysis. Total tissue weights or sizes (gland complex, surface area cm^2, see section below) were normalized by dividing them by the individuals total body weight to produce, normalized glandular complex, normalized ovotestis or normalized albumen gland;

e.g. $(\frac{Albumen\ gland\ weight\ (g)}{Total\ body\ weight\ (g)})$.
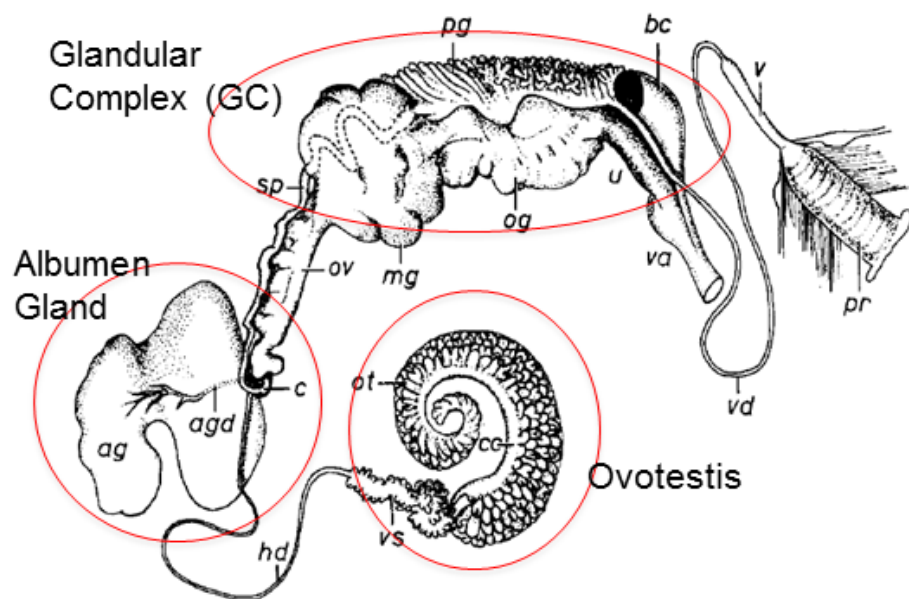
Figure 3.3: Figure showing the schematic of reproductive tract in B. glabrata.
aga: albumen gland, agd: secretory duct of the albumen gland, bc: bursa copulatrix, c: carrefour, cc: collecting canal, hd: hermaphroditic duct, mg: muciparous gland, og: othecal gland, ot: ovotestis, ov: oviduct, pg: prostate gland, pr: preputium, sp: spermduct, u: uterus, v: verge, va: vagina, vd: vas deferens, vs: vesiculae seminales (De Jong-Brink et al., 1979).

To prevent cross treatment contamination, dissection equipment (scalpel, scissors, and forceps) were washed in 100% ethanol between different snails and a new plastic petri dish and weigh boat used for each snail. Tissues for histology were fixed in a minimum of three times the tissue volume of Bouin's solution (Sigma-Aldrich) for 1 hour, rinsed twice in Industrial Methylated Spirits (IMS, Charles Tenant) (70% IMS: 30 % distilled water) and then stored in 70% IMS at room temperature before processing.

3.3.6.3 *Morphometric analysis of glandular complex*

Prior to fixation for histopathological analysis, the whole GC were dissected out and placed on a grid (cm^2) to assess their cross-sectional area through image analysis (Figure

3.4). To determine the area of the GC, photos were taken using a Leica MZ FLIII binocular dissecting microscope and captured with a Leica DC300F digital camera. The images were captured using 8bit/channel colour depth. The resolution of the objective was varied between 0.8x to 1.25x. The images were then objectively analysed using a computer-assisted image analysis program, ImageJ (http://rsb.info.nih.gov). The program allows the user to analyze, edit, enhance and print the images. Image intensity was measured using grey values with an arbitrary scale of 0 (dark, black) to 255 (clear, white) units. The scale was calibrated depending on the resolution of the image. Images were cropped and using freehand selection tool, an outline was drawn for every image before it could be converted to RGB stacks consisting of three primary colour channels. All the images in the red channel of the RGB stack were selected and were threshold at a constant value and their cross-sectional area measured.

Figure 3.4: Photomicrographs of part of glandular complex tissue from DWC of 17α-MT exposed snails.
Tissue section dissected out from the visceral cavity and placed on a cm² grid to measure the cross sectional area through image analysis. Glandular complex comprises of Mg: mucous gland, Memgl: membrane gland and Pd: prostate diverticulum. Scale bar 100μm is used.

3.3.6.4 Statistical Analysis

All statistical analyses were carried out using Sigma-Stat (v3.5). Possible intra replicate differences in shell diameter, total weight, CG surface area, albumen gland weight and ovotestis weight were analysed with either Student's T-test or Mann-Whitney U test depending on whether the data were

75

parametric or non-parametric. A Spearman's Rank Order correlation between shell diameter and total weight of the snails was performed to determine overall condition of the snails. Correlations between snail density, in terms either of total number of snails (n) present at the end of the exposure were conducted using either Pearson's correlation coefficient or Spearman's rank order to assess the effects of different survival/density profiles on overall condition. When no significant intra replicate differences were found using T-test/Mann Whitney tests then shell diameter, total weight, GC area, ovotestis and albumen gland weights were compared by analysis of variance (ANOVA) followed by the Dunnett's test to compare the treatment means with the respective controls. Data sets found to lack variance homogeneity were alternatively subjected to non-parametric analyses (Kruskal–Wallis), followed by Dunn's post hoc test (Zar, 1999). Statistical significance was considered at $\alpha \leq 0.05$ for all comparisons.

3.4 Results

3.4.1 Measured water concentration of MT and DHT

Analysis of biological activity in tank water extracts dosed with MT and DHT using the YAS assay showed that snails were exposed to the test substances in the expected range of concentrations. No androgenic activity was detected in either the dilution water control (DWC), solvent control (SC) or extraction 'blanks' at any time point measured. Solid phase extraction efficiencies, calculated from spiked samples, were $93\pm37.7\%$ for DHT and $71.13\pm21.1\%$ for MT. Overall the actual tank concentrations (as measured by the YAS assay) were between 80 and 110% of nominal concentrations for MT and 65% to 70% of nominal concentration for DHT and increased in a dose-dependent manner (Figure 3.5). The lower level of DHT relative to the nominal is likely due to lower stability of DHT in water as compared to MT (Katsiadaki *et al.*, 2006). No significant intra-replicate differences were found in the concentrations of either chemical dosed (P<0.05).
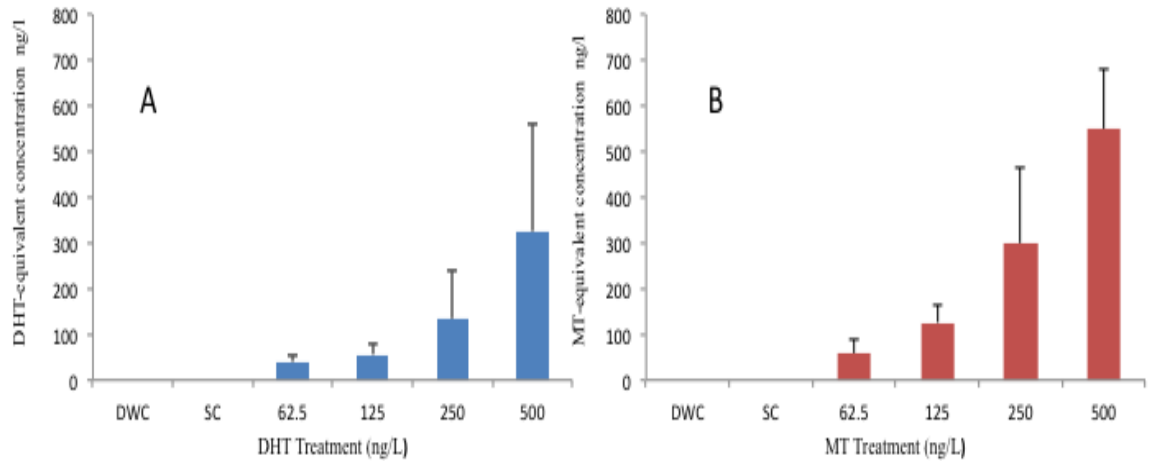
Figure 3.5: Mean tank water concentration of (A) DHT and (B) MT during exposure period.
Water samples analyzed using in vitro recombinant yeast androgen screen (YAS)
(Sohini and Sumpter, 1998). Nominal tank concentration for MT and DHT were Dilution water control,
solvent control, 62.5, 125, 250 and 500ng/l.

3.4.2 Hatchling survival and growth (shell diameter and weight)

Hatchling survival and growth of *B. glabrata* exposed to either Methyltestosterone (MT) or Dihydrotestosterone (DHT) was monitored over the exposure period to assess the effects of these chemicals on snails. For both DHT and MT exposures, 20 hatchling snails were placed into each replicate tank at the start of the experiment. At the end of the 30 day exposure hatchling survival varied between treatments, and in some cases between replicate tanks within a treatment (Figure 3.6A). Overall survival in the DHT experiment was lower (61.7%) than in the MT experiment (80.4%). In the DHT treatment most snails survived in the SC (39 total) and 62.5ng/l DHT (38 total) nominal treatments, fewer snails survived in the 125ng/l DHT (31 total) and DWC (25 total), with much lower survival found in the 500ng/l (13 total) and 250ng/l DHT (2 total) treatments.
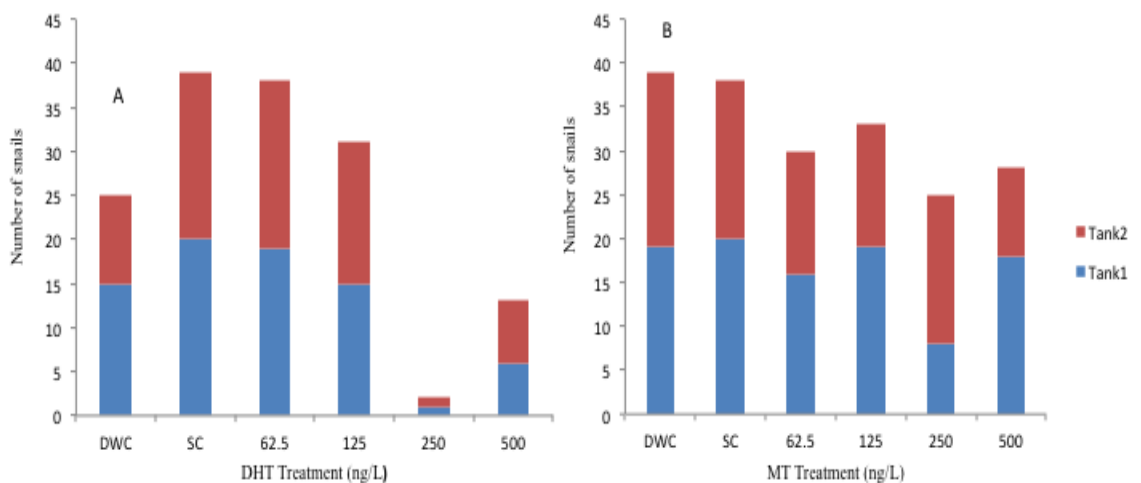
Figure 3.6: Numbers of B. glabrata snails surviving until experiment termination (30 days exposure) in each replicate tank, exposed to A) Dihydrotestosterone (DHT) and B) Methyltestosterone (MT).

In the MT treatment maximum survival was seen in the DWC (39 total) and SC (38 total) treatments Figure 3.6(B). Survival was slightly lower in the 125ng/l nominal MT (33 total) and 62.5ng/l MT (30 total); as with the DHT exposure the lowest survival in MT treatment was seen in the 500ng/l (28 total) and 250ng/l MT (25 total). Although overall survival was higher in the MT exposure compared to that of DHT, more variation between replicate tanks was observed with MT, with more than double the number of snails surviving in one 250ng/l MT replicate compared to the other (Figure 3.6B).

Shell diameters in the DHT exposure ranged from 6.78 to 15.76mm with an average of 11.52mm over the whole experiment, wet weight ranged from 0.108g to 0.701g, with an average of 0.377g (Figure 3.7 A, C). The mean diameter ranged from dilution water control (DWC) (12.04 ± 1.52mm) to 62.5ng/l (10.39 ± 1.32mm), 125ng/l DHT had the largest mean diameter (12.83 ± 1.49mm) and 500ng/l DHT (11.93 ± 1.66mm) had a smaller mean diameter. The mean wet weight also followed similar pattern decreasing from dilution water control (DWC) (0.383 ± 0.117 g) to 62.5ng/l (0.300 ± 0.072 g), 125ng/l had the largest mean weight (0.473 ± 0.124 g) and then 500ng/l DHT (0.428 ± 0.113 g). No intra replicate differences were found in the shell diameter and weight of the DHT treated snails (Figure 3.8A). Although significant differences ($P < 0.05$) were found between the shell diameter and weight of the snails from the DWC and DHT dosed snails (62.5, 125 and 500ng/l). However when compared to the solvent control (SC) treatment, no dose dependent trend was seen for mean shell diameter or wet weight, with the largest snails occurring in the 125ng/l DHT treatment and the smallest from the 62.5ng/l DHT treatment (Figure 3.7A, C). There was also significant variability in snail size within the 62.5ng/l DHT treatment (Figure 3.8A).

Snails from the MT exposure were generally smaller than those in the DHT experiment, ranging from 2.97 to 15.03mm shell diameter, with an average of 9.42mm (Figure 3.7 B, D). However, the mean total weight of the snails from the dilution water control (DWC) was the largest (10.86 ±1.26) and smallest in 250ng/l MT (7.57 ±1.9) and 500ng/l MT (8.93 ±2.65)) compared to other treatments (Figure 3.7D). Individual wet weight of snails from the MT experiment ranged from 0.023g to 0.685g with an average of 0.260g. The mean wet weight was the largest in DWC snails (0.332 ± 0.071) and followed a decreasing trend in the exposed snails and 250ng/l was the smallest (0.153 ± 0.096) (Figure 3.7B). Significant intra replicate differences both, in the shell diameter and total wet weight were found within a number of treatments (DWC ($P = 0.013$), SC ($P < 0.001$), 125ng/l MT ($P < 0.001$) and 500ng/l MT ($P = 0.006$)) (Figure 3.8B). Due to lack of uniformity between

replicates further ANOVA analysis was not conducted for diameter and wet weight in the MT study.
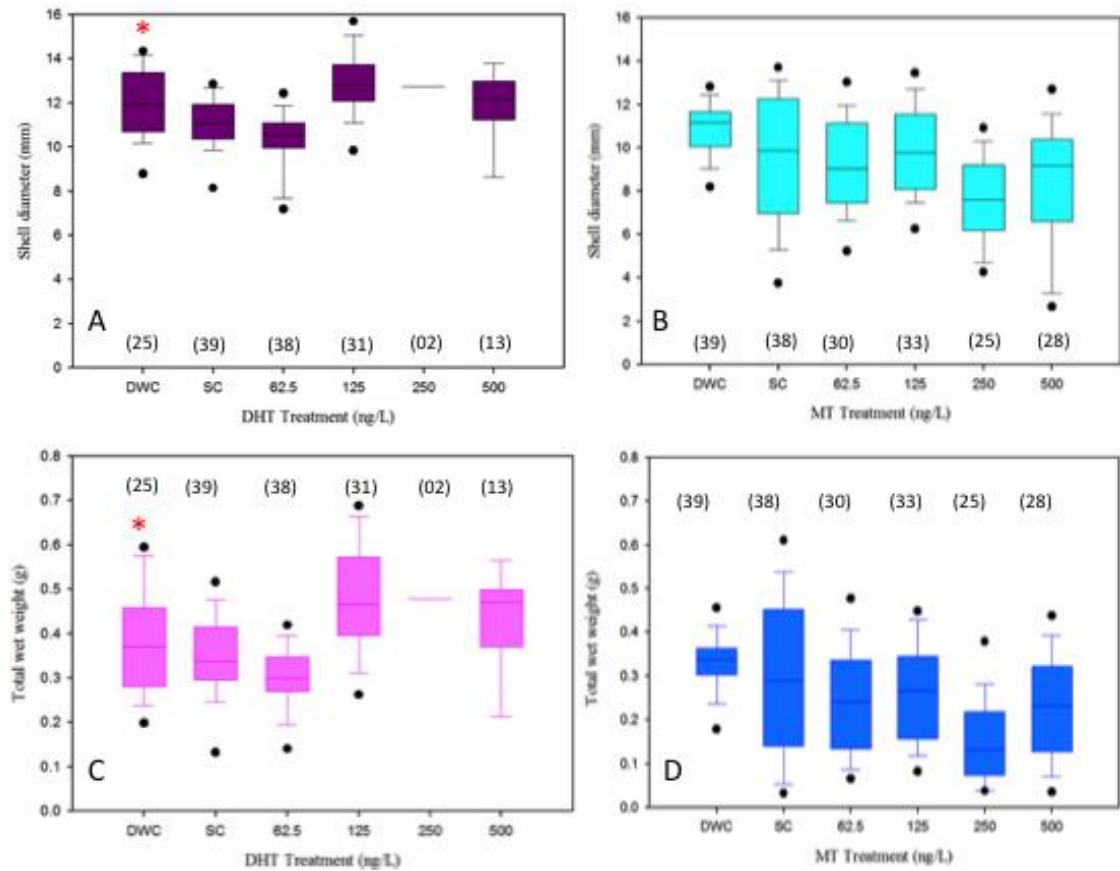
Figure 3.7: Boxplots of shell diameter (A, B) and wet weight (C, D) of B. glabrata developmentally exposed to Dihydrotestosterone (A, C) and Methyltestosterone (B, D).
Boxes represent median (full line) with 25th and 75th percentiles. The bars extend to the 10th and 90th percentiles, with outliers represented as dots. * $p < 0.05$ compared to the Solvent Control group. Sample size is indicated above each treatment label on the x-axis.
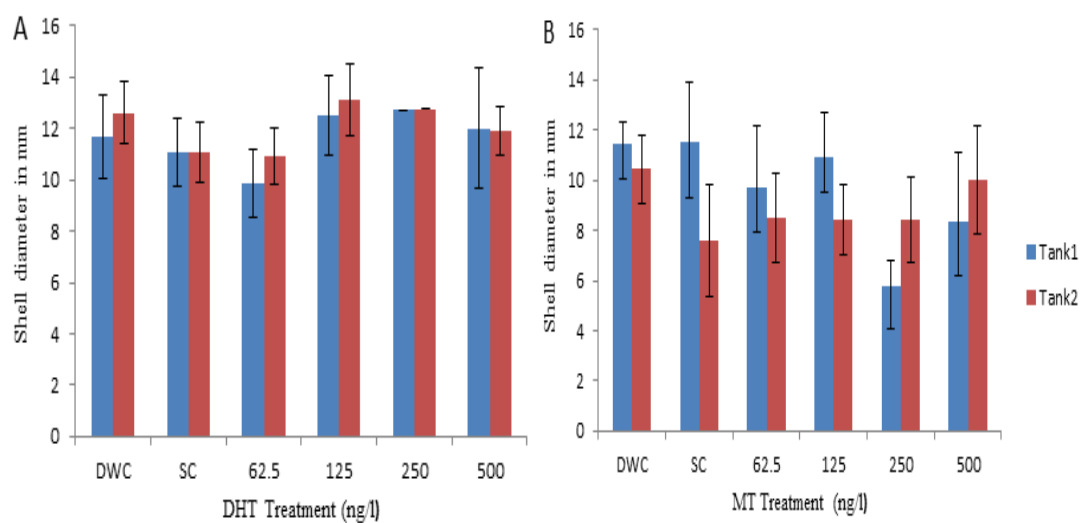
Figure 3.8: Mean shell diameter of snails from replicate tanks exposed to (A) Dihydrotestosterone (DHT) and (B) Methyltestosterone (MT).
Figure showing intra replicate differences in shell diameter of snails exposed to dilution water control (DWC), solvent control (SC), 62.5ng/l, 125ng/l, 250ng/l and 500ng/l of DHT (A) and MT (B) for 23 days post hatch.

To assess possible impacts of androgens (MT/DHT) on somatic development of the snails linear regression of snails shell diameter and total body weight (square root transformation of weight) was conducted. Significant linear correlation was observed for both the DHT ($R^2 = 0.884$, P = 0.003) and MT ($R^2 = 0.976$, P = 0.001) studies (Figure 3.9A and B). For both studies no obvious deviation from the expected pattern of growth was found for any
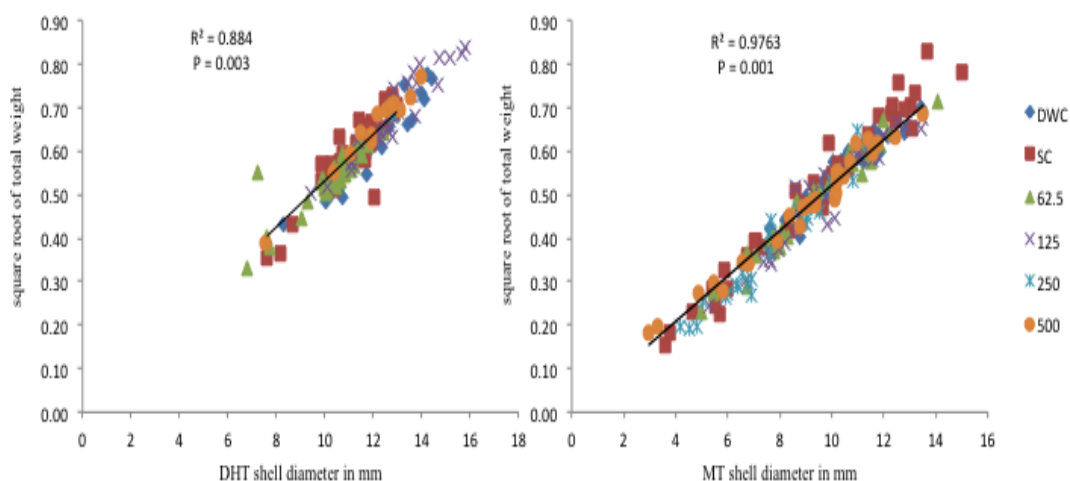
Figure 3.9: Square root transformation of the total body weight as a function of shell diameter for the snails exposed to (A) Dihydrotesterone and (B) Methyltestosterone.
The figure shows coefficient of linear regression (R^2) and statistical significance at (P<0.05) of
B. glabrata exposed to dilution water control (DWC), solvent control (SC), 62.5ng/l, 125ng/l, 250ng/l
and 500ng/l.

of the treatments i.e. all treatments followed the similar growth pattern as the controls. Significant relationships were found between density (number of surviving snails per litre) and diameter as well as density and wet weight in both the MT and DHT exposures (P<0.05) (ANCOVA) (Figure 3.10 A, B). However, the two studies showed opposite trends, in the DHT study, a significant negative trend was observed between mean shell diameter and snail density ($R^2 = -0.402$, P = 0.026) (Figure 3.10 A) and also between mean wet weight and snail density ($R^2 = -0.502$, P = 0.009) meaning tanks with lower survival had larger snail size (both weight and shell diameter). Whereas in the MT study, both the mean diameter ($R^2 = .384$, P = 0.031) (Figure 3.10 B) and the mean wet weight ($R^2 = .424$, P = 0.021) of the snails was significantly positively correlated to snail density, therefore, in this case tanks with higher survival (and therefore higher density) was associated with larger snail size.
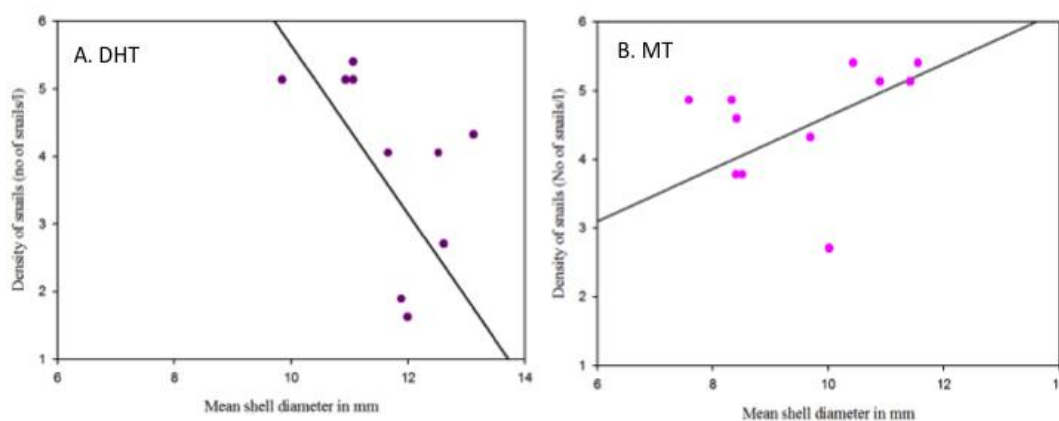
***Figure 3.10: Linear regressionog B. glabrata mean shell diameter (mm) against density (no. of
snails per litre) from the (A). DHT exposure and (B). MT exposure.***
*Figure showing weak but significant correlation in both experiments. Each treatment tank is
represented by a dot (DHT: violet and MT: pink). Linear trend line, Regression coefficient (R^2)
and significance (P) are highlighted on each graph.*

3.4.3 Quantitative assessment of the size of reproductive glands

To assess possible androgenic effects on growth and development of reproductive glands
in developmentally exposed *B. glabrata,* the size or weight of reproductive tissues were
compared. In the DHT study, no significant intra-replicate variation was found for
normalized glandular complex (NGC) size, although no comparisons could be made for
the 250ng/l dose due to very small sample size (n = 2). No significant treatment related
effects on NGC size was observed (P>0.05). Overall the mean NGC size was the largest in
62.5ng/l (0.236±0.052), similar size in DWC, 500ng/l DHT and SC (0.219±0.077,
0.214±0.042 and 0.20±0.067) and smallest in 125ng/l DHT (0.191±0.065) (Figure 3.11 B
and Table 3.2).

***Table 3.2: Mean values (±SD) of normalized glandular complex surface area (mm²), normalized
albumen gland weight (g) and normalized ovotestis weight (g) for B. glabrata developmentally
exposed to DHT.***
Asterisks represent a significant difference compared to the solvent control (SC).

Endpoint	Treatment					
	DWC	SC	62.5	125	250	500
Normalized Glandular complex(mm²)	0.219±0.077	0.20±0.067	0.236±0.052	0.191±0.065	xxxxxxx	0.214±0.042
Normalized Albumen gland weight(g)	0.019±0.007	0.032±0.009	0.029±0.009	0.023±0.008	xxxxxxx	0.023±0.006
Normalized Ovotestis weight(g)	0.023±0.008	0.021±0.009	0.21±0.006	0.019±0.005	xxxxxxx	0.018±0.003

***Table 3.3: Mean values (±SD) of normalized glandular complex surface area (mm2), normalized
albumen gland weight (g) and normalized ovotestis weight (g) for B. glabrata developmentally
exposed to MT.***
Asterisks represent a significant difference compared to the solvent control (SC).

Endpoint	Treatment					
	DWC	SC	62.5	125	250	500
Normalized Glandular complex(mm²)	0.243±0.132	0.21±0.069	0.287±0.098	0.209±0.047	0.253±0.099	0.23±0.136

Normalized Albumen gland weight(g)	0.024±0.007	0.027±0.012	0.037±0.016	0.03±0.016	0.026±0.02	0.027±0.012
Normalized Ovotestis weight(g)	0.027±0.027	0.017±0.007	0.023±0.012	0.022±0.007	0.018±0.008	0.022±0.01

In the MT study no dose dependent trend was observed in the size of the normalized glandular complex (NGC). The gland size decreased from DWC to solvent control. The gland size was largest in 62.5ng/l MT and smallest in 125ng/l MT (Table 3.3). However significant intra replicate differences were found for NGC size in 62.5ng/l (P = 0.027) nominal treatment (Figure 3.11A) in this experiment.

In DHT exposed snails no intra replicate differences were found for normalized albumen gland weight (NAW). Snails from the dilution water control (DWC) had smallest normalized albumen gland weight (0.019±0.007 g) compared to all the other treatments (Figure 3.11D and Table 3.2). No comparisons could be made for the 250ng/l DHT dose due to very small sample size. The mean normalized albumen gland weight (NAW) in the MT treatment was the largest in 62.5ng/l (0.036±0.018 g) compared to the controls and different treatments (Figure 3.11C and Table 3.3) but significant intra replicate differences were found for NAW size in 62.5ng/l (P = 0.034) and 125ng/l (P=0.042) nominal treatment.

The mean normalized ovotestis weight (NOW) seemed to follow a decreasing trend in the DHT treatment with the largest NOW in dilution water control (DWC) and smallest in the 500ng/l nominal treatment (Table 3.2 and Figure 3.11F). However no significant treatment related effects were found for NOW in the DHT study. No comparisons could be made for the 250ng/l DHT dose due to very small sample size. In MT exposure, the mean normalized ovotestis weight (NOW) was largest in DWC (0.027±0.027 g) and smallest in SC (0.017±0.007 g) (Table 3.3 and Figure 3.11E). However, there was significant intra-replicate variation in ovotestis weight in the solvent control (SC) treatment (P = 0.014) in the MT treatment.
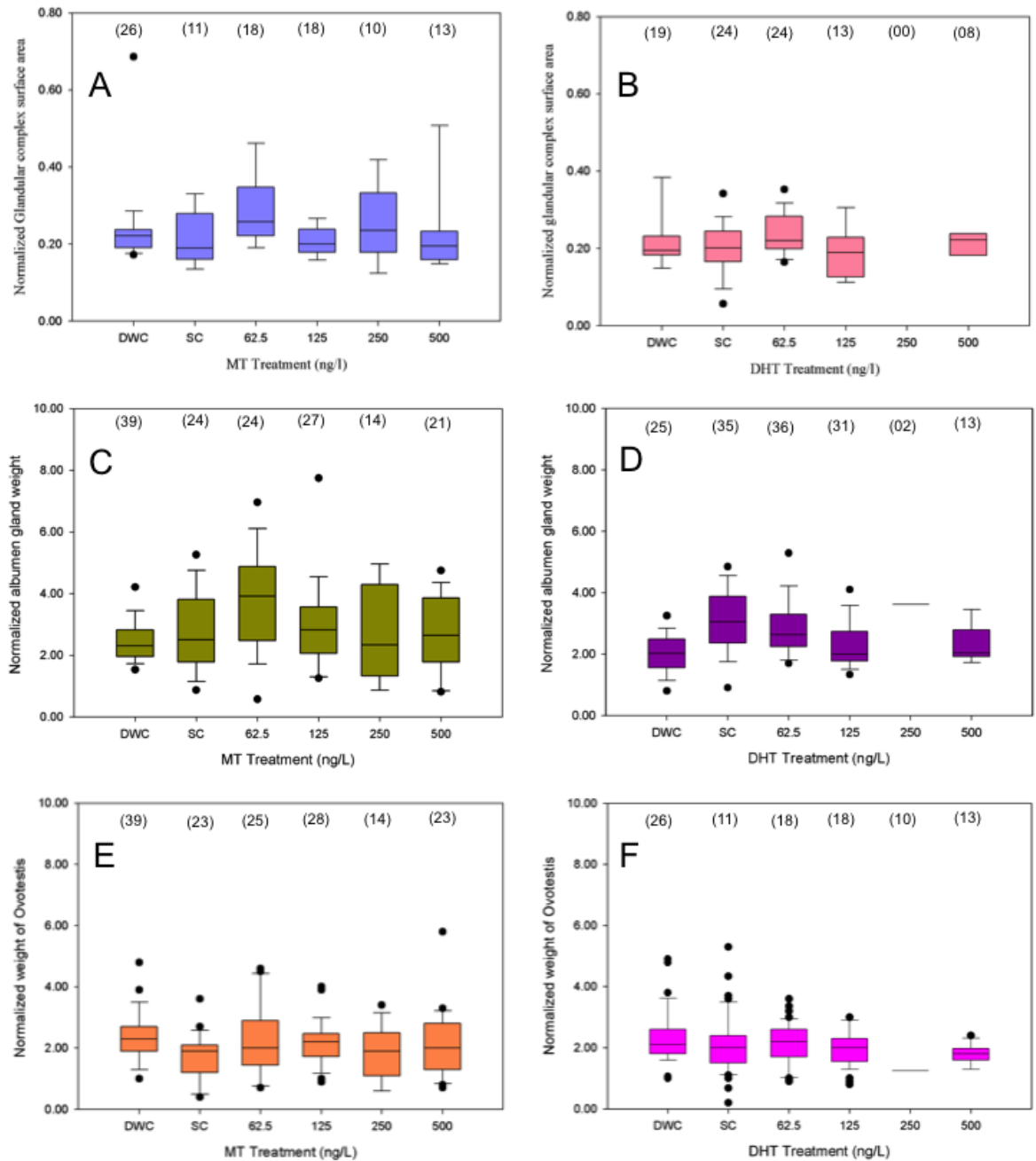
Figure 3.11: Boxplots of Normalized surface are of glandular complex (A, B), Normalized albumen gland weight (C, D) and Normalized ovotestis weight (E, F) of B. glabrata developmentally exposed to dihydrotestosterone (A, C, E) and Methyltestosterone (B, D, F).
*Boxes represent median (full line) with 25th and 75th percentiles. The bars extend to the 10th and 90th percentiles, with outliers represented as dots. * p< 0.05 compared to the Solvent Control group. Sample size is indicated. Due to high mortality observed in the 250ng/lDHT treatment, the two snails were not included in analysis.*

3.4.4   Histopathological analyses of reproductive tissues

To assess possible androgenic effects on the reproductive tissue structure or cell type, histopathology was conducted on the reproductive tissues (albumen gland, glandular complex and ovotestis) from developmentally exposed *B. glabrata* from both MT and DHT studies. For these experiments RNAlater was used to preserve whole de-shelled

83

snails, enabling quick preservation followed by later dissection of tissues for both molecular and histopathological analysis. RNAlater did sufficiently preserve the tissue and aid dissection (firmed up snail tissue) and had previously been reported to have similar histopathology fixation results to formalin (Florell *et al.*, 2001). However, compared to Bouin's solution (more often used in our laboratory for histopathology of reproductive tissues) RNAlater was not as fine a histology fixative with some tissue integrity and fine structural detail lost (Figure 3.12(A-F)) but it was good enough for my purpose to determine that chemicals did not affect these tissues. For example fine details of oocyte development stage could not be determined due to disruption to the shape of the oocytes, supporting cells (follicles) and follicle cleft. Much of the shape, size and staining of the acini, including the structure of the acini walls were distorted. However, some spermatogenic stages, spermatids and spermatogonia as well various stages of oogenesis as described by De Jong-Brink et al. (1976, 1977) could be identified (De Jong-Brink *et al.*, 1977; De Jong-Brink *et al.*, 1976). In RNAlater fixed GC staining was muted and differential staining between different tissues types (membrane gland, mucous gland and prostate gland) was less pronounced. Sectioning artifacts also increased in the tissues due to slight hardening up of the tissue. Therefore RNAlater is not recommended as a fixative, especially for gonadal or ovotestis tissues. Histopathology was conducted on the ovotestis; albumen gland and glandular complex of all exposed and control snails from both the MT and DHT exposures.
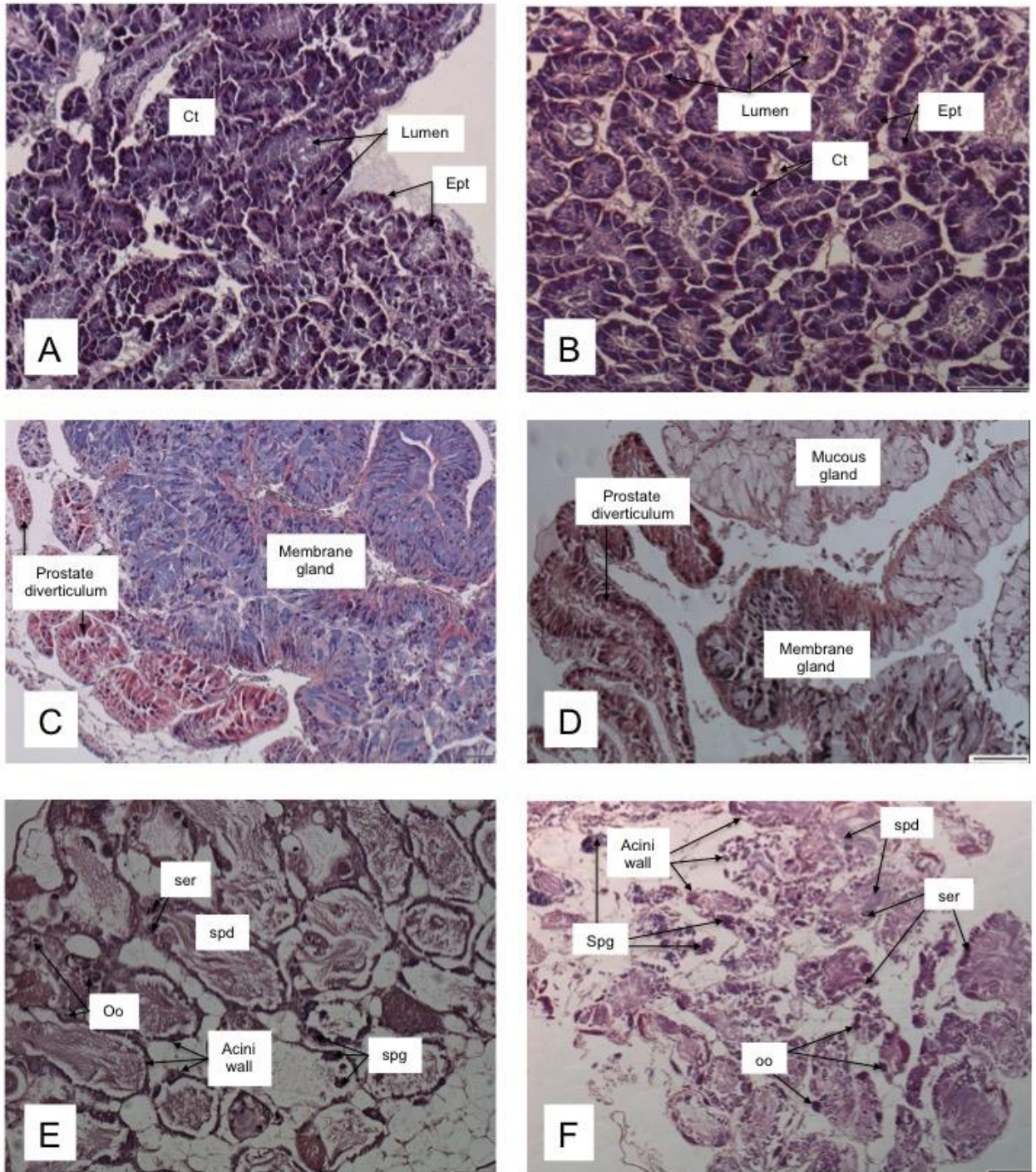
Figure 3.12: Photomicrographs of reproductive tissues (Albumen gland (A-B), Glandular complex (C-D), and Ovotestis (E-F)) preserved in Bovin's (A, C, E) and in RNAlater followed by Bouin's (B, D, F).

Tissue section stained with H & E. Scale bar 100µm in each case. Staining has been affected in albumen gland and prostate diverticulum and some wear and tear can be seen, but ovotestis structure looks degraded. C: connective tissue, E: epithelium, L: lumen of a gland tubule, MG: mucous gland, Memgl: Membrane gland and Pd: prostate diverticulum, Ooc: Oocyte, Oog: Oogonia, Sc: Sertoli cells,
Spd: Spermatid, Spg: Spermatogonia.
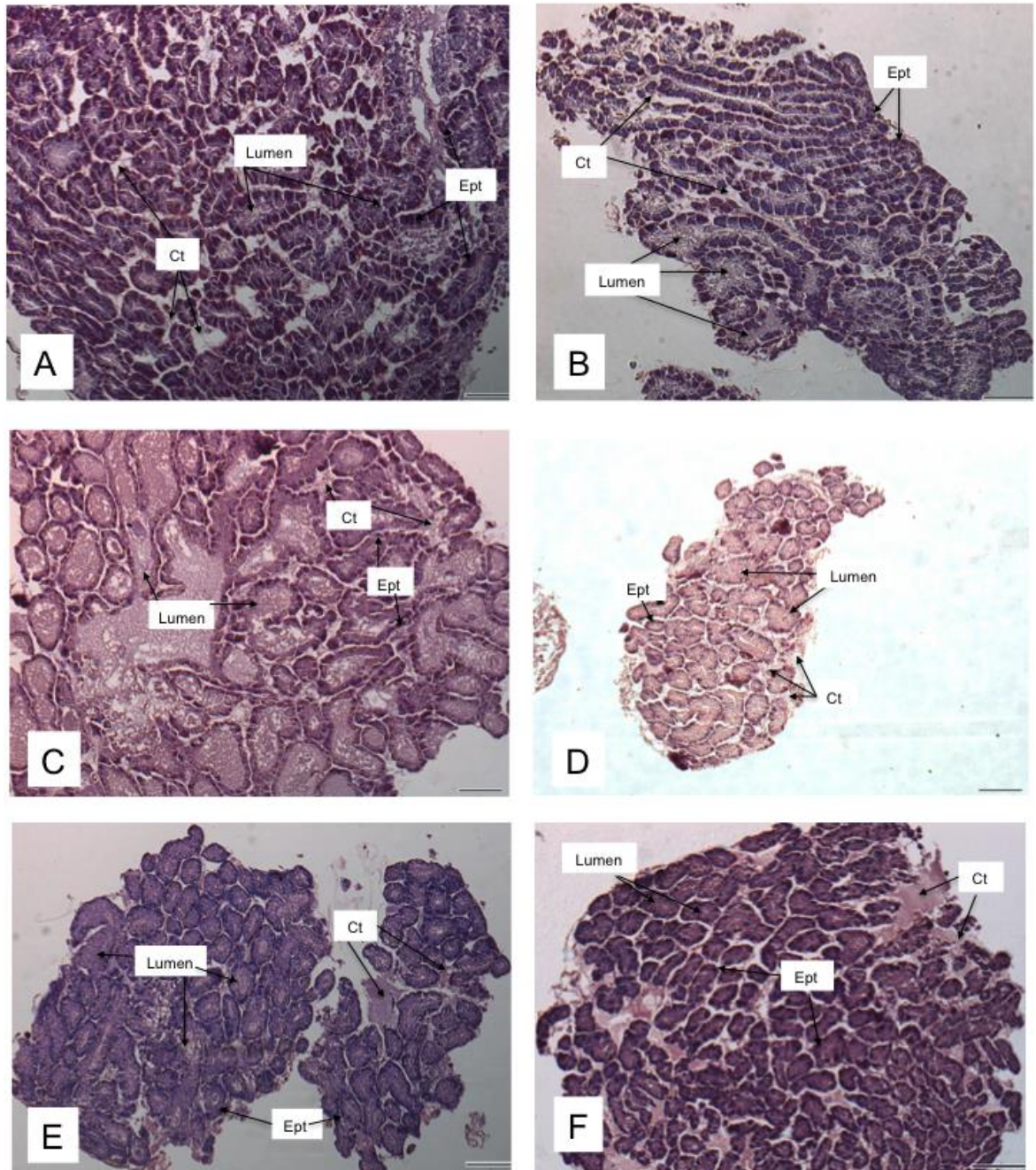
Figure 3.13: (A-F) Pictomicrographs of part of albumen gland tissue rom DWC, SC, 62.5ng/l, 125ng/l. 250ng/l and 500ng/l of 17α-MT exposed snails.
Tissue sections stained with H&E. Scale bar 100μm in each case. C: connective tissue, E: epithelium, L: lumen of a gland tubule.

The histopathological results in both the MT (Figure 3.13(A-F)) and the DHT (Figure 3.14(A-E)) studies revealed no significant compound/dose related morphological changes in the albumen gland. Structurally, the albumen gland in *B. glabrata* consists of densely packed tubules consisting of columnar glandular epithelial cells with large and conspicuous nuclei towards their base (that stained pink) and apically located cuboidal secretory vesicles (which stained dark pink to violet) (Klussmann-Kolb, 2001). Some of the glandular epithelial cells were heterogeneously stained but this was observed in all

treatments (including the DWC) irrespective of the dose. This could be due to variation in the granules and cytoplasm present in the secretory cells, depending on the egg laying process (Klussmann-Kolb, 2001; De Jong-Brink, 1973). The tissues become slightly grainier when RNAlater was used and this has been also reported by Guo and Catchpoole (2003) but they reported no adverse effects on tissue integrity (Guo D, 2003). There was no evidence of the albumen gland changing in structure or becoming more like a prostate gland, as has been reported in other gastropod species exposed to environmental 'androgens' (Watermann *et al.*, 2008) (COMPRENDO, Final Report 3/2006), (http://ec.europa.eu/).

The GC comprised of membrane gland, mucous gland and prostate gland (Figure 3.15(A-F) and Figure 3.16(A-E)). The histological sections of GC from both the MT study and the DHT study did not show any chemical-related or dose-related morphological changes. The membrane gland (Klussmann-Kolb, 2001) or the oothecal gland (De Jong-Brink *et al.*, 1979) consists of glandular cells (stained light pinkish) filled with secretory fluid in the form of small mucous droplets that stained violet. The membrane gland showed a continuous transition to the mucous gland that was the largest part of the GC. The mucous gland consisted of a glandular mass made up of glandular cells that were columnar in shape, widely coiled and stained light pinkish. Interspersed were cells containing small densely packed secretory droplets that stain dark pinkish. The prostate gland consists of the prostate duct and its diverticula contains long glandular cells with distinct vesicles of homogeneous or amorphous texture with conspicuous nuclei that normally stain blue (Klussmann-Kolb, 2001) but due to the RNAlater stained dark pink in these experimental samples. The supporting cells bear short cilia and stained pink to violet.

All ovotestis from the MT and DHT exposures were initially assessed for chemical dose related effect on oogenesis and spermatogenesis. As far as could be determined there were no treatment related effects, all stages of oogenesis and spermatogenesis, as described by De Jong-Brink et al. (1976, 1977), were present (De Jong-Brink *et al.*, 1977; De Jong-Brink *et al.*, 1976). However, due to preservation in RNAlater detailed analysis (e.g. scoring) of possible alterations to spermatogenesis or oogenesis was not possible.
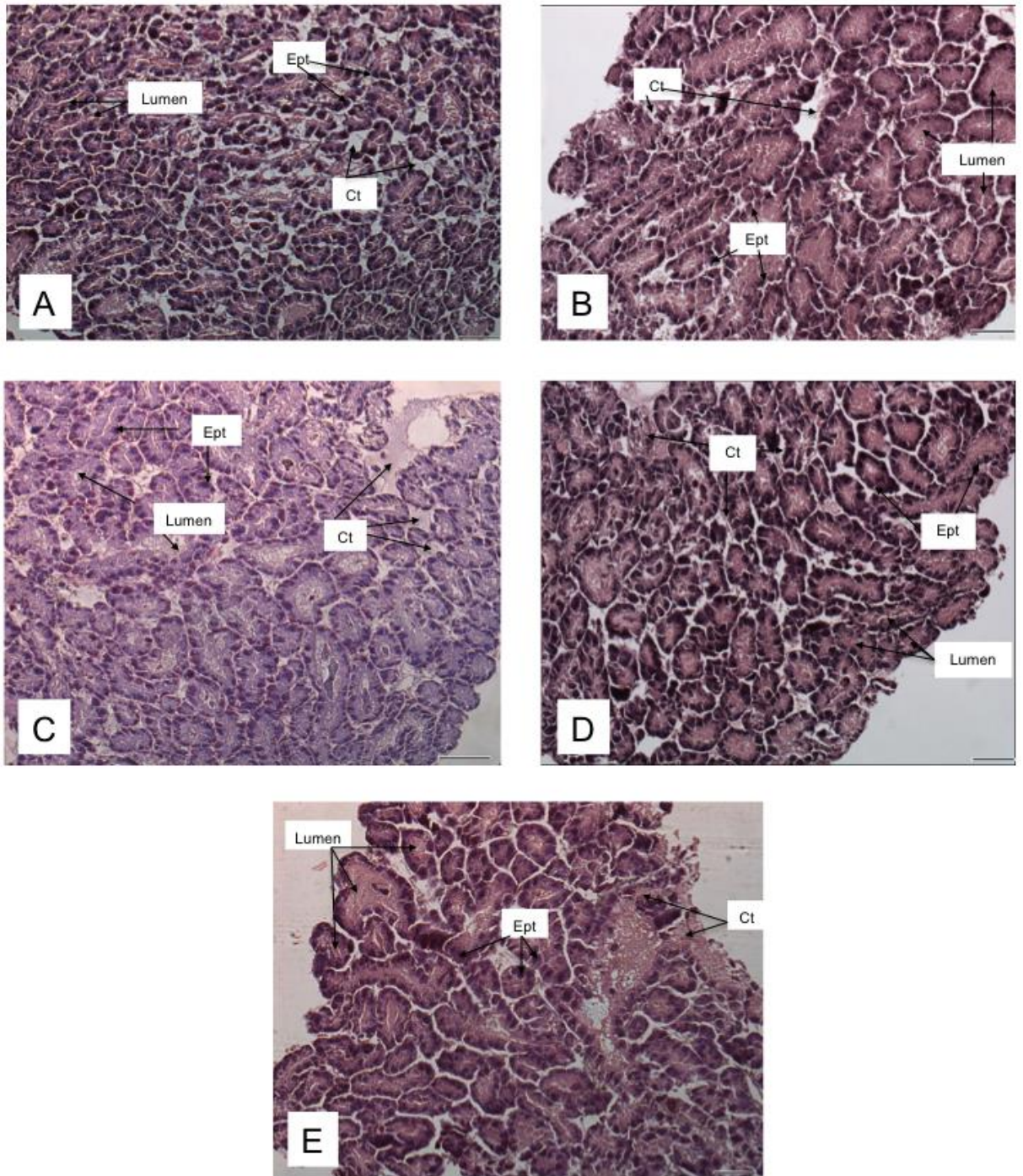
Figure 3.14: (A-F): Photomicrographs of part of albumen gland tissue from DWC, SC, 62.5ng/l, 125ng/l and 500ng/l of DHT exposed snails.

Tissue section stained with H & E. Scale bar 100μm in each case. Due to very high mortality observed in 250ng/L, the tissue from two survivor snails has not been included. Histopathological analysis did not show any significantly obvious effects on the albumen gland of the exposed groups as compared to controls (DWC and SC). C: connective tissue, E: epithelium and L: lumen of a gland tubule.
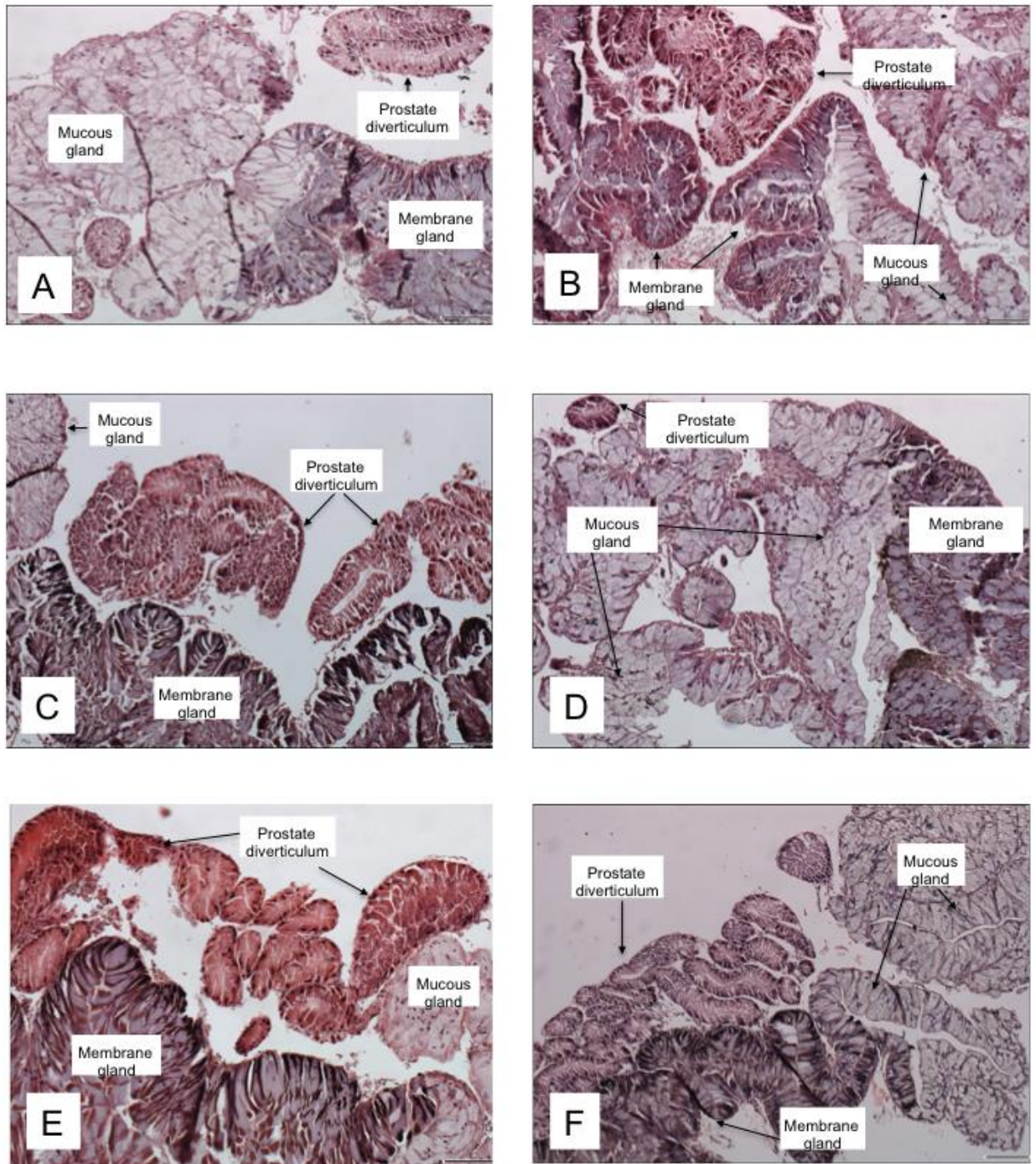
Figure 3.15: (A-F): Photomicrographs of part of glandular complex tissue from DWC, SC, 62.5ng/l, 125ng/l, 250ng/l and 500ng/l of 17α-MT exposed snails.
Tissue section stained with H & E. Scale bar 100µm in each case. Glandular complex comprises of mucous gland, membrane gland and prostate diverticula. Histopathological analysis did not show any significantly obvious effects on the albumen gland of the exposed groups as compared to controls (DWC and SC). MG: mucous gland, Memgl: Membrane gland and Pd: prostate diverticulum.
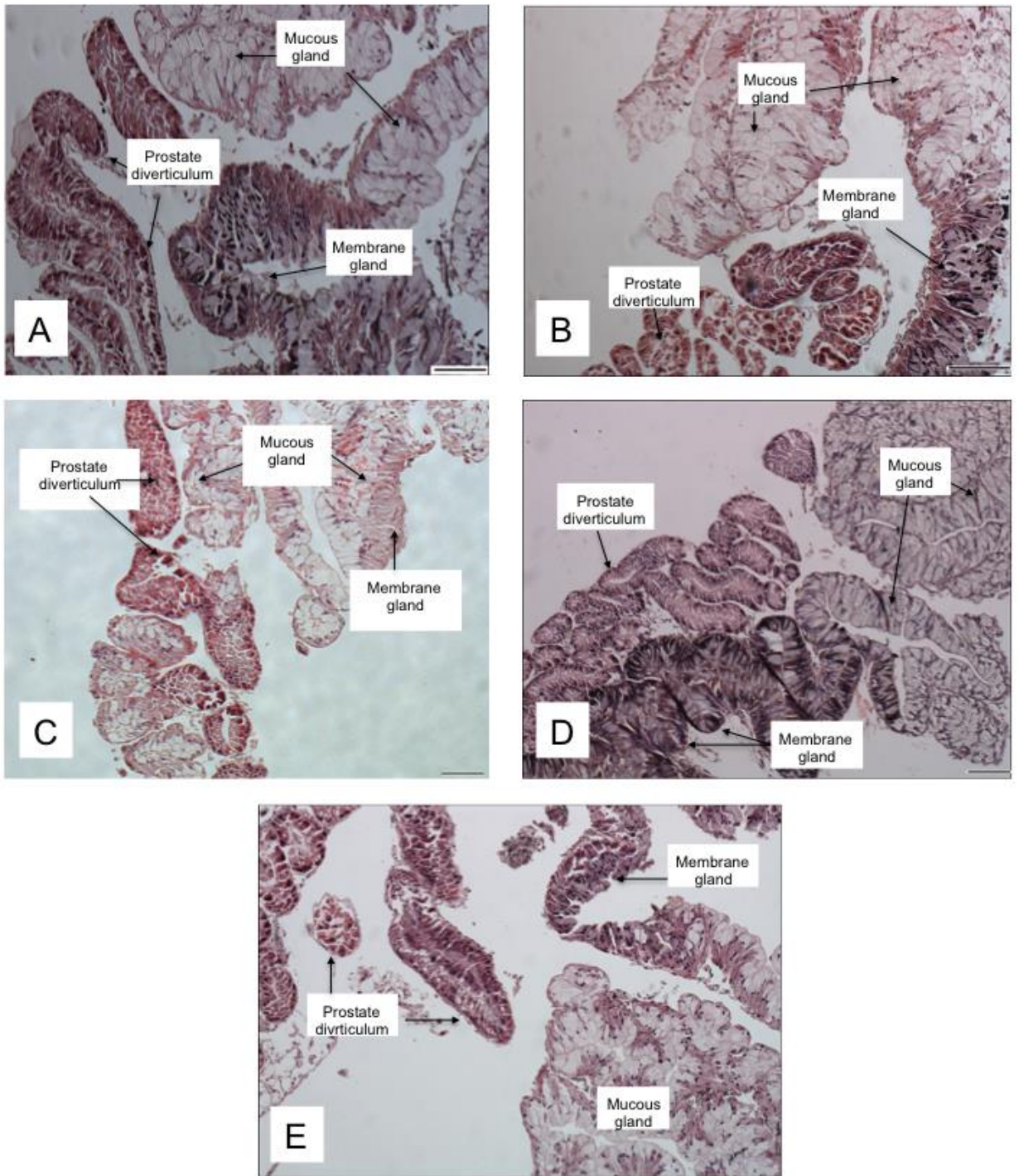
Figure 3.16: (A-E): Photomicrographs of glandular complex tissue from DWC, SC, 62.5ng/l, 125ng/l and 500ng/l of DHT exposed snails.

Tissue section stained with H & E. Scale bar 100μm in each case. Due to very high mortality observed in 250ng/l, the tissue from two survivor snails has not been included. Glandular complex comprises of mucous gland, membrane gland and prostate diverticula. Histopathological analysis did not show any significantly obvious effects on the albumen gland of the exposed groups as compared to controls (DWC and SC). MG: mucous gland, Memgl: Membrane gland and Pd: prostate diverticulum.

3.5 Discussion

The first part of my hypothesis was tested by *in vivo* developmental exposure of the freshwater snail *B. glabrata* to 5α-Dihydrotestosterone and 17α-Methyltestosterone, compounds known to be potent androgens in vertebrates. The primary aim of this study was to investigate whether developmental exposure to these androgens would affect the differentiation, development and growth of the reproductive organs and glands of the freshwater snail *B. glabrata* thus providing a non-vertebrate alternative model for male reproductive toxicity testing. The main result of the two chemical exposure studies is that no chemical related dose-dependent effects on sexual development were found, in terms of reproductive organ development, using any of the morphological and histopathology endpoints.

3.5.1 Possible effects on survival and growth

The general pattern in hatchling survival in both the DHT and MT studies was that mortality was higher in the 250ng/l and 500ng/l (MT or DHT) than the dilution water control (DWC) or solvent control (SC) treatments. This suggests there may have been some level of toxicity to the developing snails at these concentrations (although mortality in the MT study was generally lower than in the DHT study). However, toxicity has not been previously reported in other gastropod species exposed to similar nominal concentrations of androgens (e.g. (Rivero-Wendt *et al.*, 2014; Giusti *et al.*, 2014; Duft *et al.*, 2007; Czech, Weber and Dietrich, 2001; Tillmann *et al.*, 2001)) and the dosing series for these experiments were designed to mimic concentrations documented to have endocrine effects in molluscs (e.g. (Duft *et al.*, 2007; Czech, Weber and Dietrich, 2001; Tillmann *et al.*, 2001) and aquatic vertebrates (e.g. (Margiotta-Casaluci and Sumpter, 2011; Seki *et al.*, 2004) but not acute toxic effects. The experimental design of the other gastropod studies differs from this study, in that many were static-renewal rather than continuous dosing using flow-through water system. In addition those studies which measured exposure concentrations in the water often found much lower (sometimes <20%) than nominal exposure concentrations (Giusti *et al.*, 2014), whereas in this study we measured exposure concentrations of 80-110% nominal MT and 65-70% nominal DHT. Therefore snails in these experiments may have been exposed to higher concentrations than similar species in other studies.

Growth is an important factor to consider when assessing the reproductive organ size

endpoints. In vertebrates, exposure to androgens has been reported to increase somatic growth, for example in developmentally exposed fathead minnows (*Pimephales promelas*), DHT significantly increased weight and length in males (20ng/l DHT) and females (200ng/l DHT) (Margiotta-Casaluci and Sumpter, 2011). Not much is known about effects of androgens on somatic growth in molluscs but somatic growth was enhanced in *Radix balthica* (a pulmonate snail) when exposed to 5 µg/l 17α-Ethylestradiol (Hallgren *et al.*, 2012). Mimicking the rodent prepubertal intact male Hershberger assay protocol, *B. glabrata* was exposed to chemicals from egg mass to sexual maturity. Although some significant differences in size were found between the solvent control (SC) and DHT exposed *B. glabrata,* no dose dependent trend could be identified, with significantly larger snails in the 125ng/l DHT treatment and significantly smaller snails in the 62.5ng/l DHT treatment, indeed significant differences were also found between the SC and the dilution water control (DWC). Solvents are used as carriers in aquatic toxicology testing because at high concentrations (such as stock solutions) some test chemicals have low water solubility. The possibility of 'solvent effects' have arisen in a number of publications, and the use of solvents is a topic of concern in endocrine toxicology testing (Turner *et al.*, 2012; Hutchinson *et al.*, 2006; Zhang and Baer, 2000). The OECD guidelines suggested maximum solvent concentration of 0.01% v/v (100µl/l) for aquatic vertebrates (OECD 2000) (http://www.epa.gov/endo/pubs/ref-2_oecd) and gastropods (OECD 2010) (http://www.oecd.org/). In this study the nominal concentration of dimethylformamide (DMF) in the SC, MT and DHT tanks was 0.006% v/v (60ul/l), and therefore below the OECD recommendations. However, in a review of the use of solvents in aquatic toxicology testing, Hutchinson *et al*. (2006) suggest using a maximum concentration of 0.002% solvent, and more recently, Turner *et al.* (2012) and Lecomte *et al.* (2013) have reported molecular and reproductive effects at solvent concentrations of 0.0025 and 0.002% v/v respectively (Lecomte *et al.*, 2013; Turner *et al.*, 2012; Hutchinson *et al.*, 2006). Notwithstanding this, studies in our laboratory have used similar concentrations of DMF (0.0067% v/v) to those used in the *B. glabrata* tests and found no significant solvents effects on growth or condition index in fish (e.g. (Margiotta-Casaluci and Sumpter, 2011). All these studies highlight the importance of testing solvent controls alongside dilution water controls and comparing chemically treated subject to the appropriate controls when analyzing outcomes. Interestingly, in the MT study a slightly different pattern emerged, in this study the largest snails were found in the DWC treatment, whereas MT exposed snails were on average smaller than the SC or DWC snails, suggesting that there may have been a negative effect on growth of MT in *B.*

glabrata in the higher concentrations, however due to significant intra dose variation (i.e. between replicate tanks) no statistically significant effects could be seen. In general the variation in shell size between the replicate tanks, in the DHT exposure was much lower than in the MT exposure and much of this variability seems to be caused not by the higher levels of mortality in the DHT exposure but by the higher survival in the MT exposure consisting of smaller, possibly not well grown/developed snails at the end of the 30 exposure. In the DHT study only one snail was less than 7mm in diameter at the termination of exposure compared to 36 snails in the MT experiment, in the MT study these small snails were found in all treatments bar the dilution water control (DWC).

3.5.2 Possible effects of androgens on reproductive tissue size and histopathology

One of the primary markers of androgen exposure in vertebrate models, such as the Hershberger test, is increase size/weight of male specific organs and a reduction in those that are female specific. For example, in a pubertal rat assay (enhanced OECD test guideline 407) MT exposure resulted in increased prostate weight in male rats treated with 200 mg/kg/day (Wason *et al.*, 2003). In contrast no chemical related effect on normalized glandular complex (NGC) (including prostate gland) size was found for *B. glabrata* snails exposed to any of the DHT concentrations tested. Although the normalized glandular complex (NGC) of the 62.5ng/l MT exposed snails were much larger than those from the SC treated snails, there were also significant differences in NGC between the two 62.5ng/l MT replicate tanks, suggesting this was not an effect of MT exposure.

There were no consistent significant dose-related effects on albumen gland weight in the MT or DHT exposure studies. In the MT study due to intra replicate differences in the normalized albumen gland size (NAW) no statistically significant effects could be assessed. In the DHT study, the snails from the dilution water control (DWC) had significantly smaller albumen glands compared to all the other treatments, suggesting the addition of the carrier solvent (DMF), rather than the test compound, had positive effects on the snails albumen gland, perhaps by providing an additional source of nutrition for the production and accumulation of secretory substances. In general, gastropod albumen gland size is thought to be related to accumulation of secretory products in the tissue, due to different stages of egg-laying cycle (Egonmwan, 2007; Miksys and Saleuddin, 1987; De Jong-Brink *et al.*, 1979) but this variability factor will hold true for all the snails irrespective of the chemical exposure.

In vertebrate models, histopathological changes to gland structure or development are also reported in androgen exposed individuals, for example, in the pubertal male rats histology

of the seminal vesicle and prostate gland show enhanced glandular development (1 mg/kg/day Testosterone Propionate by subcutaneous injection) (Tinwell *et al.*, 2007). In comparison, no histological impact of MT or DHT exposure was observed in the albumen gland or the Glandular Complex (GC) of androgen-exposed *B. glabrata*. Differences were observed between individual snails in the staining of the epithelium of the tubules in the albumen gland, but these variations were found across all treatments, and were due to the presence of more vacuolated cytoplasm and less secretory granules in the epithelium (De Jong-Brink, 1973), possibly due to variation in the egg-laying cycle of the snails at the time of sampling. Varying structural appearances of the albumen gland in conjunction with different physiological states due to metabolic processes and sexual activity has also been reported by Duncan (Duncan, 1960). The lack of histological effects of MT or DHT exposure on secondary sexual glands in *B. glabrata* is in contrast to previously published result in another freshwater pulmonate species, *L. stagnalis*, which showed weak histological damage to the albumen after 8 weeks exposure to 100ng/L MT (Czech, Weber and Dietrich, 2001). The timing of exposure (adult versus juvenile) and duration of exposure (8 weeks versus 30 days) could be an important factor deciding the range of effects seen in these organisms.

Although a detailed histopathology comparison of androgen exposure on the ovotestis could not be conducted due to sub-optimal tissue preservation in RNAlater, there were no indications of effects at 10X resolution and no gross disruption to oogenesis or spermatogenesis were observed in any of the MT or DHT exposed snails. No significant effect of androgen exposure (MT or DHT) was found on the ovotestis weight (although some differences were found between the ovotestis size between the dilution water control (DWC) and solvent control (SC) in the MT study). The lack of gross histopathological changes to the ovotestis in *B. glabrata* exposed to MT is in contrast to the recent work of Rivero-Wendt *et al*. (2014), in which adult *B. glabrata* were exposed semi-statically to MT (0.01mg/l to 1mg/l MT). In that study, Rivero-Wendt *et al.* (2014) reported that nominal MT concentrations of 0.1 mg/l and 1.0 mg/l significantly increased the area of ovotestis acini occupied by mature spermatozoa. However, their statistical comparisons were between MT dosed and water control snails, rather than between MT dosed and solvent control (0.001% ethanol) exposed snails. Indeed, in Rivero-Wendt *et al.* (2014) the solvent control snails showed the largest average area occupied by sperm (almost double that of the water control) (Rivero-Wendt *et al.*, 2014). This suggests MT may not have been the overriding factor in increasing sperm production in these snails and again highlights the importance of assessing, and factoring in, possible solvent mediated affects

in toxicology studies. A recent study from Bandow and Weltje (2012), reported enhanced embryonic development in *Lymnaea stagnalis* when exposed to 89.3μl/l of triethylene glycol. More recently accelerated embryonic development has also been reported in *P. antipodarum* at 20μl/l of acetone and inhibition of juvenile growth at similar concentration of ethanol (Lecomte *et al.*, 2013).

No changes in gene expression were found in ovotestis from MT snails compared to the controls were seen using suppressive subtractive hybridization methods is conducted in our laboratory by Dr Anne Lockyer. Overall this suggests it is unlikely that there were androgenic effects on spermatogenesis or oogenesis in these animals as has been documented in other molluscs. Similar experimental procedures to those conducted with the snails i.e. developmental exposure to MT via the water, have been conducted on a number of fish species producing androgen-induced effects. For example, Japanese Medaka (*Oryzias latipes*), developmentally exposed to androgens via the water in a similar way to *B. glabrata* in this study, exhibited male skewed sex ratios and testis-ova at average concentrations of 27.75ng/l MT (Seki *et al.*, 2004). Another fish species, the Fathead Minnow (*Pimephales promelas*), when exposed to MT as adults, displayed atretic follicles in their ovary ($\geq$ 100ng/l MT) and inhibited ovarian growth ($\geq$ 50μg/l MT) (Pawlowski *et al.*, 2004).

The absence of significant dose-dependent physiological or molecular responses to MT or DHT on *B. glabrata* was surprising considering the number of publications that suggest gastropod molluscs have physiological levels of vertebrate-type androgens naturally occurring in their tissues; use vertebrate-type androgens in reproduction or are reproductively disrupted by exogenous androgen or anti-androgen exposure (Giusti *et al.*, 2013; Rivero-Wendt *et al.*, 2013; Goto *et al.*, 2012; Schulte-Oehlmann *et al.*, 2006; Janer *et al.*, 2005b; Tillmann *et al.*, 2001). For example, in the freshwater prosobranch gastropod, *Marisa cornuarietis,* MT showed imposex promoting effects from 100ng/l nominal to complete virilization of females at 1μg/l (Schulte-Oehlmann *et al.*, 2006)and extreme alternation of albumen gland histopathology have been reported in gonochoristic prosobranch gastropods environmentally or experimentally exposed to TBT, TPT or MT (Watermann *et al.*, 2008; Schulte-Oehlmann *et al.*, 2006; Schulte-Oehlmann *et al.*, 2004). However, it is worth highlighting that the majority of publications documenting effects of androgens are in prosobranch (generally gonochoristic) gastropods rather than pulmonate (hermaphrodite) gastropods as tested here. Considering the evolutionary radiation of gastropods it might not be unexpected that they might show different responses.

Therefore, in terms of basic molluscan endocrinology research, the question that arises

from this work is that if the classic vertebrate steroids, such as androgens, are not controlling sexual development, spermatogenesis and oogenesis then what are the relevant genes and molecules involved? There is the well-known example of masculinizing (imposex) effects of the marine pollutant tributyltin (TBT) on prosobranch gastropods (Matthiessen and Gibbs, 1998) and study of this chemical in molluscs came up with several hypothesis. One of the groups suggested that TBT might be acting through androgen-mediated pathway (Oehlmann *et al.*, 2007; Matthiessen and Gibbs, 1998) and another group in Japan suggested that TBTs virilizing effects were acting via the retinoid X receptor (RXR) (Horiguchi *et al.*, 2007; Nishikawa *et al.*, 2004). More recently Pascoal et al. (2012) demonstrated the involvement of both retinoids and androgen signaling in inducing imposex in molluscs (Pascoal *et al.*, 2012).

3.6 Conclusion

The data presented provides convincing evidence that growth and development of the reproductive tract in *B. glabrata* are unaffected by developmental exposure (embryo through to sexual maturity) to MT and DHT, both potent androgens in vertebrates. The consequence of these findings are significant as it suggests that *B. glabrata* and possibly other molluscs species would not be appropriate models for testing androgenic or anti-androgenic compounds. The reported effects of androgens seen in other mollusc species could be due to alternative NRs interacting with androgens or alternative, perhaps non-genomic, pathways for androgen signalling. The next most feasible step was to investigate and identify nuclear receptors in the publically available genome for *B. glabrata*.

Nevertheless, it must be borne in mind that molluscs consist of a group with considerable diversity that might be reflected in its endocrine system. This disparity among different species, even within the gastropods, makes it difficult to extrapolate steroid exposure effects seen in one species to another and warrants further research to examine and understand mechanisms for gonadal development and other important phenomena in molluscs, which are second only to arthropods in the number of species occupying the widest variety of habitats in the world (Gruner H, 1993). All this should not limit the investigation and use of new invertebrate models in the field of toxicology (thus supporting important 3Rs strategies) but it does highlight the importance of having supporting molecular programs and mechanistically relevant endpoints for the proposed test species.

4 Chapter 4: Investigation of nuclear receptors in *Biomphalaria glabrata* and *Lottia gigantea*

4.1 Background

One of the main stimuli for this thesis was that the effects of androgenic and anti-androgenic chemicals on reproduction and development in crustaceans, molluscs and echinoderms seem to be comparable to those in mammals (For example (Oehlmann and Schulte-Oehlmann, 2003), the consequent hypothesis that androgens in snails must act through signal transduction pathways similar to those found in vertebrates and that androgen receptors must be present in molluscs was inferred (Omran, 2012; Koehler *et al.*, 2007). The results of Chapter 3 however show clearly that vertebrate androgens do not support this hypothesis in my model organism, *B. glabrata*. In this chapter I therefore sought to investigate the nuclear receptor repertoire in Biomphalaria glabrata and compare these to known vertebrate nuclear receptors (NRs), with a view to assessing the snail's potential as a new invertebrate model organism, both as a prospective new test organism and to elucidate fundamental genetic and mechanistic causes of endocrine diseases. For comparative purposes, the genome of second gastropod, the owl limpet, *Lottia gigantea* was also investigated for nuclear receptors.

4.2 Introduction to Hormone signalling

In vertebrates, the classical genomic model of hormone action is directly related to the ability of the steroid hormone to regulate the expression of genes containing DNA response elements in their promoter region. When the steroid hormone (ligand) binds to a specific intracellular steroid receptor protein (also called nuclear receptor), the bound ligand-activated steroid receptor complex acts as a transcription factor and binds as a homodimer/heterodimer to DNA response elements in the target gene promoter causing activation or repression of transcription and thus protein synthesis (Nikolenko and Krasnov, 2007; Shang, Myers and Brown, 2002). In addition, steroids may use a short-cut, referred to as the 'rapid' or non-genomic' pathway in the absence of the nuclear receptor and operating through a membrane receptor that is possibly associated with the plasma membrane (Heinlein and Chang, 2002). This involves initiation of second messenger signal transduction cascades which leads to an increase in free intracellular calcium, activation of protein kinase A (PKA) and protein kinase C (PKC), resulting in activation of MAP kinase cascade and phosphorylation of cyclic AMP (cAMP) responsive element binding (CREB) protein transcription factor (Walker, 2010a). Steroid hormones can act through either of these mechanisms in order to regulate biological processes.

4.2.1 General structure of nuclear receptors

NRs are members of a large superfamily of evolutionary related multi- domain transcription factors that are believed to have emerged early in animal evolution, as reflected by their biologically essential roles (Bridgham *et al.*, 2010; Bertrand *et al.*, 2004; Owen and Zelent, 2000). Distinct NRs regulate and coordinate multiple physiological processes such as embryonic development, sex determination (Ganter *et al.*, 2011; Finley *et al.*, 1998; Shen *et al.*, 1994), fatty acid metabolism and toxin metabolism (Chiang, 2002), detoxification of foreign substances by integrating internal and external signals, thereby maintaining homeostasis (Markov and Laudet, 2011a; Sladek, 2011). Outside of normal physiology, NRs have been identified to play a role in many endocrine related pathological processes, such as cancer (breast and prostate) (Peehl and Feldman, 2004), hormone-resistance syndromes, diabetes, and rheumatoid arthritis (Ottow and Weinmann, 2008; Moller, 2001).
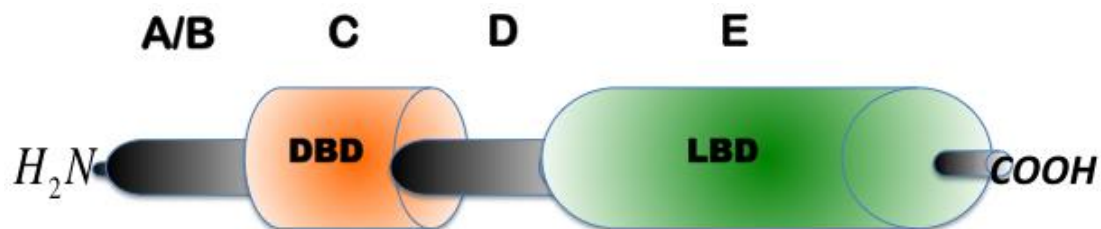
Figure 4.0.1: Schematic representation of an nuclear receptor (NR).
The variable N-terminal region (A/B) contains the ligand independent AF-1 transactivation domain. The C region contains conserved DNA-binding domain (DBD). A hinge region D connects the DBD to the less conserved E region that contains the ligand binding domain (LBD) as well as the dimerization surface and the ligand-dependent AF-2 core activation domain.

Despite their diversity in function, NRs exhibit strong similarities due to their conserved molecular structure, which consists of single polypeptide chains that can be divided into five to six modular domains, from the N-terminal to C-terminal end (Mangelsdorf and Evans, 1995; Giguère *et al.*, 1986) Figure 4.0.1. Although there is undoubtedly some interaction between the domains, each domain is responsible for encoding a motif that exerts a specified molecular action. The activation function 1 (AF-1), which is located on the N-terminal A/B domain, is important for the ligand-independent constitutive transcriptional activity of the NR (Wärnmark *et al.*, 2003). The length and sequence of the A/B domain is hyper variable between receptors and receptor subtypes (RXRα and β) revealing a weak evolutionary conservation. This region contains a variety of kinase recognition sequences with specific attributed functions and is the most frequent site of alternative splicing. For these reasons, it is thought that the variable N-terminal sequences may be responsible for the species-, receptor-, promoter-context- and cell-specific

properties of NR transactivation (Germain *et al.*, 2006; Renaud and Moras, 2000). The C domain or the DNA binding domain (DBD) is based on two highly conserved cysteine-rich zinc finger motifs and a carboxy-terminal extension (CTE). This is the most conserved region of the nuclear receptors. The first zinc finger contains the proximal-or P-box region, an alpha helix that is responsible for the sequence specificity of the receptor and DNA binding response element (Laudet *et al.*, 2002). Located within the second zinc finger is the distal or D-box, which lies vertical to P-box helix, and is a site responsible for receptor dimerization (Aranda and Pascual, 2001). The CTE provides the interface for protein-DNA or protein-protein monomerization (Khorasanizadeh and Rastinejad, 2001). The hinge region is a mobile linker between the DBD and the C-terminal LBD facilitating the DBD and the LBD to adopt different conformations. Moreover, the hinge domain contains the C-terminal end of the DBD, which may be involved in identifying the 5' end of the hormone response element (HRE). This region may also contain nuclear localization signals and serve as binding site for corepressor proteins (Aranda and Pascual, 2001). The LBD (E domain) region of NRs is a multifunctional domain that mediates homo- and heterodimerisation, ligand binding, interaction with coregulators, nuclear localization and transcriptional activity (Aranda and Pascual, 2001). The sequence of the LBD varies substantially between NRs although the overall architecture of the LBD is comprised of 11-13 anti-parallel α-helices. Helices 1-11 form a structure comprising a ligand-binding pocket that diverges sufficiently among the NRs to guarantee selective ligand recognition. The pocket is guarded by flexible helix 12 (H12), which is commonly referred to as the ligand-induced activation function (AF- 2) and contains residues crucially involved in transcriptional coregulator interactions (Gronemeyer, Gustafsson and Laudet, 2004). A further structurally distinct region within the LBD is a dimerization surface, which mediates the interaction with the partner LBDs (Bourguet, Germain and Gronemeyer, 2000; Gampe Jr *et al.*, 2000). The C-terminal F domain, which is contiguous with the E domain, is not present in all nuclear receptors and its function is poorly understood (Germain *et al.*, 2006).

There are several NRs that do not follow this basic structure function paradigm. For example, the constitutive androstane receptor (CAR) undergoes constitutive transcription (in the absence of the ligand) of the target gene (Dussault *et al.*, 2002). A subfamily (NR0) of nuclear receptor lacks either the DBD or the LBD. The dosage-sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, gene 1 (DAX-1) and small heterodimer partner (SHP) are the two NRs in humans that lack the DBDs completely and *knirps* and embryonic gonad (EGON) in *Drosophila* lack the LBD completely (Germain

et al., 2006). A novel subfamily of NRs with two DBDs have been identified in flatworm species (*Schmidtea mediterranea* and *Dugesia japonica*), from molluscs (*L. gigantea*) and from arthropods (*Daphnia pulex*) (Wu *et al.*, 2007) but not yet reported in mammals.

4.2.2 Mechanism of action

Extensive studies over the last three decades have provided a detailed understanding of the molecular mechanisms through which nuclear receptors convert a hormone signal into a transcriptional response. Nuclear receptors bind to DNA as a heterodimer, homodimer, or monomer, depending on the class of NR. The steroid hormone receptors glucocorticoid receptor (GR), progesterone receptor (PR), estrogen receptor (ER), androgen receptor (AR) and Mineralocorticoid receptor (MR) function as homodimer, binding to DNA inverted repeat with a gap of 3 neucleotide (AGAACAnnnTGTTCT). The thyroid hormone receptor (THR) and retinoid acid receptor (RAR) as well as many of the other orphan receptors, heterodimerize with retinoid X receptor (RXR) (Lefebvre, Benomar and Staels, 2010) and activate direct repeats of (GGTCA) and the gap between the repeat motifs depends on the heterodimer. For example RXR-RAR heterodimer has a gap of 5 nucleotides and RXR-THR has a gap of 4 nucleotides (Laudet, 1997). The ER activates a similar repeat to the THR but requires a 3bp gap like other steroid receptors (Forman *et al.*, 1992). However, NR family 5 NRs (steroidogenic factor-1) bind as monomers to an extended core site (Mangelsdorf *et al.*, 1995). Based on sequence homology, NRs have been further classified into subfamilies and owing to their mechanism of action NRs have been categorized into classes.

Class I NRs (includes members of NR subfamily 3) are located in the cytoplasm in the absence of ligands. Ligand (hormone) binding to Class I NRs lead to conformational change resulting in dissociation of heat shock proteins, see Figure 4.0.2(A). Translocation to the nucleus is followed by steroid receptor homodimerisation and binding to specific hormone response elements (HREs). The NR-DNA complex in turn recruits other proteins that are responsible for the transcription of downstream DNA into mRNA. Subsequent translation of this mRNA into protein results in changes in cellular functions.

Class II NRs (NR subfamily 1) are located in the nucleus regardless of their ligand binding status and form obligate heterodimers with RXR to form a functional transcription factor, see Figure 4.0.2(B). In the absence of the ligand, the receptor is bound to corepressor proteins, which is dissociated in the presence of the ligand followed by association with coactivator proteins. These coactivator proteins recruit additional proteins such as RNA polymerase II forming a complex that bring about the translation of the downstream DNA

into mRNA and eventually protein synthesis takes place (Rosenfeld, Lunyak and Glass, 2006; Glass and Rosenfeld, 2000).
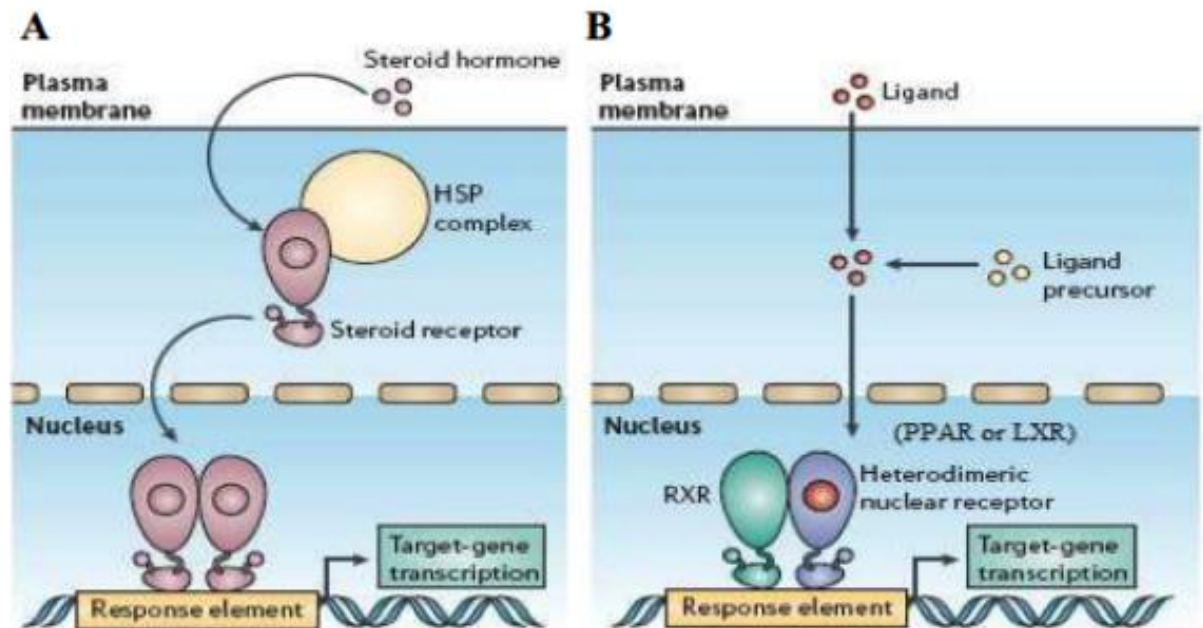
Figure 4.0.2: Molecular mechanism of class I and class II NRs:
A). Hormone-binding dissociates HSP from steroid receptor and enables DNA-binding and activation of the associated transcription machinery. B). Class II NRs like PPAR or LXR form a heterodimer RXR and bind DNA target gene in presence/absence of ligand. Adapted from Glass et al., 2006.

Both Class III NRs (NR family 2) and Class IV NRs (NR family 5) mostly comprise of orphan receptors (Mangelsdorf *et al.*, 1995). Class III NRs bind to direct repeats of the HRE as homodimers while Class IV NRs bind as monomers or dimers but only a single half site of the inverted repeat of the HRE, that is preceded by a 5'- flanking A/T rich sequence, is bound to the DNA binding domain of the receptor (Aranda and Pascual, 2001). Monomeric NRs use C-terminal extension (CTE) to recognize their HRE (Glass, 1994).

4.2.3 Species distribution of nuclear receptors and their nomenclature

The NR superfamily genes are specific to metazoans and are not found in algae, fungi (*Saccharomyces cerevisae*), yeast, choanoflagellates and plants (*Arabidopsis thaliana*) (King *et al.*, 2008; Escriva, Bertrand and Laudet, 2004; Launhardt and Munder, 2000). However, there are some suggestions that fungi can use ligand regulated transcription factors that might share structural and functional features with metazoan NRs (Naar and Thakur, 2009; Phelps *et al.*, 2006). With recent progress in genome sequencing projects from the early-branching animal lineages, 2 NRs have been reported from the sponge *Amphimedon queenslandica* (Bridgham *et al.*, 2010; Srivastava *et al.*, 2010), 2 from the

ctenophore *Mnemiopsis leidyi* (Reitzel *et al.*, 2011), 4 from the placozoa *Trichoplax adhaerens* (Bridgham *et al.*, 2010; Srivastava *et al.*, 2008) and 17 from the cnidarian *Nematostella vectensis* (Reitzel and Tarrant, 2009).Over 284 NRs have been reported in *Caenorhabditis elegans* (Sluder and Maina, 2001). Among arthropods, 21 NR genes are found in *Drosophila melanogaster* (Adams *et al.*, 2000), 22 in *Apis mellifera* (Velarde, Robinson and Fahrbach, 2006), and 21 in *Tribolium castaneum* (Bonneton, Chaumot and Laudet, 2008). 21 NRs have also been identified in the trematode *Schistosoma mansoni* (Wu *et al.*, 2006) and recently 43 NRs were identified from the Pacific oyster, *C. gigas*, a marine bivalve, (Vogeler *et al.*, 2014a). Moving on to vertebrates, 48 NRs have been identified in the human genome (Robinson-Rechavi, Garcia and Laudet, 2003), 49 in mice and 47 in rats (Zhang *et al.*, 2004). The unusually large number of NRs in *C. elegans* is due to extensive proliferation of one gene (HNF4) within the nematode phylum (Antebi *et al.*, 2000).

Nuclear receptors are strong phylogenetic markers showing diversification of metazoans from the simplest sponges to vertebrate-specific paralogues at different evolutionary distances (García, Laudet and Robinson-Rechavi, 2003) through the structural conservation of their various domains. Comparative studies examining the evolution of various protein families and their nomenclature in major groups of organisms have been done using various approaches to phylogeny. Based on sequence homology using phylogenetic approaches the Nuclear Receptors Nomenclature committee (1999) organized NRs into six distinct subfamilies (Escriva, Delaunay and Laudet, 2000; Thornton and DeSalle, 2000). Each of these subfamilies is a robust monophyletic group, in which all receptors clustered in a subfamily originate from a single ancestor (Markov and Laudet, 2011b) (Figure 4.0.3). The committee defined a gene nomenclature containing the nuclear receptor abbreviation NR followed by subfamily, group and receptor number. Subfamily 1 includes thyroid hormone receptors (THR)-like receptors; Subfamily 2 has RXR-like receptors; Subfamily 3, estrogen receptors (ER)-like receptors; Subfamily 4, nerve growth factor IB (NGFIB)- like receptors; Subfamily 5, steroidogenic factor like receptors; Subfamily 6, germ cell nuclear factor (GCNF)-like receptors and the last Subfamily 0 includes all the nuclear receptors lacking either of the DBD or LBD regions (Zhang *et al.*, 2004; Maglich *et al.*, 2001) (Figure 4.0.3). In Table 4.0.1 all the members of NR superfamily are listed with their subfamily and groups, nomenclature, gene name, accession numbers and their respective ligands (Chawla *et al.*, 2001).

Table 4.0.1: Nomenclature of nuclear receptors and their associated ligands.
(Germain et al.*, 2006; King-Jones and Thummel, 2005; Smirnov, 2002)*

Subfamilies and groups	Name	Nomenclature	Accession number	Ligand
1A	THRa	NR1A1	AAH08851	Thyroid hormone
	THRb	NR1A2	AAI06931	Thyroid hormone
1B	RARa	NR1B1	AAH08727	Retinoic acid
	RARb	NR1B2	AAH60794	Retinoic acid
	RARg	NR1B3	AAA63254	Retinoic acid
	RAR	NR1B4	AF378827	Retinoic acid
1C	PPARa	NR1C1	AAB32649	Fatty acids fibrates
	PPARb	NR1C2	AAA36469	Fatty acids
	PPARg	NR1C3	AAH06811	Fatty acids, prostoglandins
1D	Rev-erb-a	NR1D1	AAH56148	Orphan
	Rev-erb-b	NR1D2	AAH45613	Orphan
	E75	NR1D3	AAN11687	Orphan
	HR-85	NR1D4	AAO39185	Orphan
1E	E78	NR1E1	AAF51692	Orphan
1F	RORa	NR1F1	AAH08831	Cholesterol
	RORb	NR1F2	AAH93774	Retinoic acid
	RORg	NR1F3	AAA64751	orphan
	HR3	NR1F4	AAA28461	orphan
	HR-23	NR1F5	P41828	orphan
1G	HR14	NR1G1	AAA96982	orphan
1H	ECR	NR1H1	NP724456	20-Hydroxyecdysone
	LXRb	NR1H2	NP_009052	Oxysterols
	LXRa	NR1H3	AAV38218	Oxysterols
	FXR	NR1H4	AAI30574	Bile acids, fexaramine
	FXRB	NR1H5	AY094586	Lanosterol
1I	VDR	NR1I1	AAB95155	Vitamin D
	PXR	NR1I2	AAD05436	Xenobiotics steroids
	CAR	NR1I3	AAY56401	Xenobiotics Phenobarbital
1J	HR96	NR1J1	AAC46928	Orphan
	Daf-12	NR1J2	AAD34462	Orphan
	HR-8	NR1J3	AAP31437	Orphan
	HR-48	NR1J4	CAD36502	Orphan
1K	HR-1	NR1K1	AAC48174	Orphan
2A	HNF4	NR2A1	NP_849180	Orphan
	HNF4g	NR2A2	NP_004124	Orphan
	HNF4B	NR2A3	Z49827	Orphan
	DHNF4	NR2A4	AAF52702	Orphan
	HR-49	NR2A5	CAD57702	Orphan
2B	RXRa	NR2B1	AAI10999	9-cis retinoic acid
	RXRb	NR2B2	AAC18599	9-cis retinoic acid
	RXRg	NR2B3	AAA80681	9-cis retinoic acid
	USP	NR2B4	AAF45707	20-Hydroxyecdysone
2C	TR2	NR2C1	AAA36761	Orphan
	TR4	NR2C2	NP_003289	Orphan
	TR2-4	NR2C3	AF378828	Orphan
2D	HR78	NR2D1	Q24142	Orphan
	HR-41	NR2D2	NP_500073	orphan
2E	TLL/TLX	NR2E1	AAL05871	Orphan
	PNR	NR2E2	AAD28301	Orphan
	DSF	NR2E3	AAF52303	Orphan
	FAX-1	NR2E4	AAF54133	Orphan
	HR-67	NR2E5	CAA97428	orphan
2F	COUP-TFa	NR2F1	AAH04154	Orphan
	COUP-TFb	NR2F2	AAH42897	Orphan
	SVP	NR2F3	M28863	Orphan
	EAR2	NR2F4	AAH02669	Orphan
	UNC55	NR2F5	CAP16277	orphan
2H	AmNR4	NR2H1	AF323683	orphan
3A	ERa	NR3A1	AAI28575	Estradiol-17β
	ERb	NR3A2	AAV31779	Estradiol-17β
3B	ERRa	NR3B1	AAH63795	Orphan
	ERRb	NR3B2	AAC99409	Diethylstilbestrol
	ERRg	NR3B3	EAW93335	Diethylstilbestrol
	ERR	NR3B4	AAL37554	Diethylstilbestrol
3C	GR	NR3C1	AAH15610	Cortisol
	MR	NR3C2	AAA59571	Aldosterone
	PR	NR3C3	AAA60081	Progesterone
	AR	NR3C4	EAX05380	Testosterone
4A	NGFI-B	NR4A1	CAG32985	Orphan

	NURR1	NR4A2	AAH09288	Orphan
	NOR1	NR4A3	CAI95138	Orphan
	HR38	NR4A4	AAF53914	Orphan
	HR-6	NR4A5	CAA85271	orphan
5A	SF1	NR5A1	AAH32501	Orphan
	LRH1	NR5A2	AAI18572	Orphan
	FTZ-F1	NR5A3	AAN11667	Orphan
	HR-25	NR5A4	CAA91028	orphan
5B	HR39	NR5B1	AAN11107	Orphan
6A	GCNF	NR6A1	AAB96828	Orphan
	HR4	NR6A2	AAX73355	Orphan
	HR-91	NR5B2	CAB60329	orphan
0A	Knirps	NR0A1	CAA31709	Orphan
	Knirps-like	NR0A2	AAF51627	Orphan
	EGON	NR0A3	CAA34626	Orphan
	ODR7	NR0A4	AAC46497	Orphan
	Trithorax	NR0A5	M31617	Orphan
0B	DAX1	NR0B1	AAH11564	Orphan
	SHP	NR0B2	AAH30207	Orphan
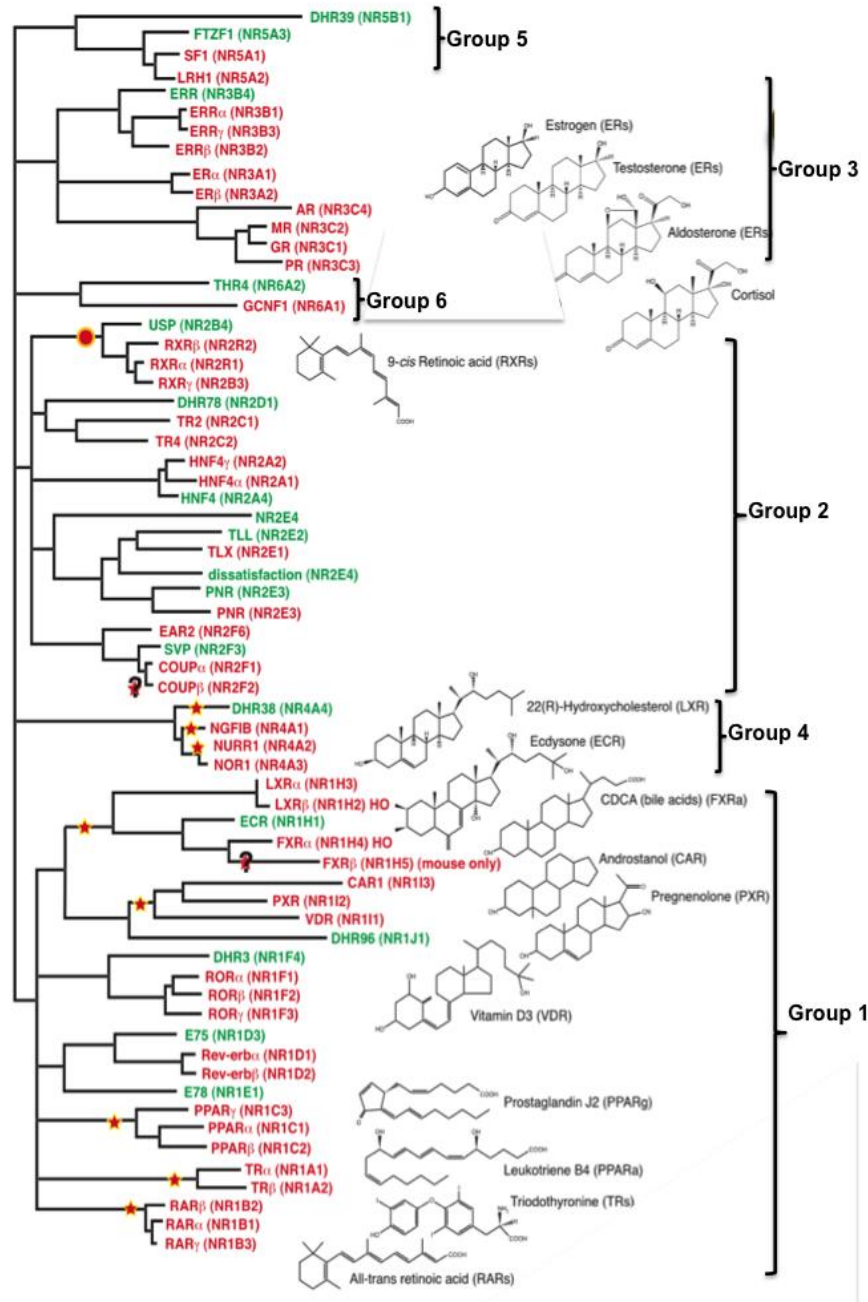
Figure 4.0.3:Phylogenetic tree showing nuclear receptors identified in humans and fly and the structure of identified ligand in the nuclear receptors as mentioned in table 4.0.2.
The nuclear receptors in green belong to D. melanogaster and those in red are found in humans (Robinson-Rechavi, Garcia and Laudet, 2003).

4.2.4 Nuclear receptor repertoire in molluscs: evolutionary and endocrine disruption insights

Nuclear receptors became a focus of research in molluscs after the discovery that an anti-fouling compound, Tributyltin (TBT), caused penis growth in female molluscs living near harbours and estuaries (Oehlmann *et al.*, 2007; Matthiessen and Gibbs, 1998). This is one of the most cited examples for the effect of endocrine disrupting compounds in the environment on masculinization of female snails and has been reported in over 150

species of gastropods worldwide (Janer *et al.*, 2006a; Oehlmann and Schulte-Oehlmann, 2003; Matthiessen and Gibbs, 1998). This led to several studies in molluscs trying to differentiate the mode of action of sex steroids, specifically estrogen receptor (Stange *et al.*, 2012; Benstead *et al.*, 2011a; Bannister *et al.*, 2007; Castro *et al.*, 2007; Kajiwara *et al.*, 2006; Keay, Bridgham and Thornton, 2006). Several *in vivo* studies in molluscs indicated a possible role for "vertebrate-like" endocrine (e.g. androgen and estrogen) hormones in the reproductive cycle (Lafont and Mathieu, 2007; Bannister *et al.*, 2007; Thornton, Need and Crews, 2003). Sex steroids in molluscs may be chemically vertebrate-like (Matthiessen, 2008), although vertebrate antibodies can cross-react with other steroids or steroid metabolites, including non-vertebrate steroids, leading to false positive results. Experiments focussing on blocking of sex steroid receptors with chemical antagonists suggested receptor-mediated sex steroid signalling in the scallop *Placopecten magellanicus* and in the gastropod *N. lapillus* (Croll and Wang, 2007; Santos *et al.*, 2006; Bettin, Oehlmann and Stroben, 1996) but (Koehler *et al.*, 2007) pointed out the fact that the presence of binding sites does not necessarily indicate the presence of a receptor.

Recently, Thornton et al. (2003) found that molluscan neural and reproductive tissues express a nuclear receptor gene that is unambiguously orthologous to the vertebrate-like estrogen receptor (ER) (Thornton, Need and Crews, 2003). This was followed by vertebrate-like ER being reported in other molluscan species, including the bivalves, *Crassostrea giga*s (Matsumoto *et al.*, 2007), *Mytilus edulis* (Puinean *et al.*, 2006) and *Chlamys farreri* (Zhang, Pan and Zhang, 2012), and the gastropod, *Aplysia californica* (Thornton, Need and Crews, 2003), *Lymnaea ollula*, *Marisa cornuarietis* (Castro *et al.*, 2007), *Thais clavigera* (Kajiwara *et al.*, 2006), *Nucella lapillus* (Bannister *et al.*, 2007), and in the cephalopod, *Octopus vulgaris* (Keay, Bridgham and Thornton, 2006). The vertebrate-like ER in molluscs activates transcription constitutively in the absence of ligand, different from vertebrate ER (Bannister *et al.*, 2013; Eick and Thornton, 2011; Matsumoto *et al.*, 2007; Castro *et al.*, 2007; Keay, Bridgham and Thornton, 2006; Kajiwara *et al.*, 2006). The presence of an vertebrate-like ER in molluscs and its similarity with vertebrate ER is also confounded by the absence of aromatase (CYP19) (Markov *et al.*, 2009), the enzyme required for the conversion of testosterone to estradiol, but it is possible that another enzyme catalyzes the aromatase reaction (Thornton, Need and Crews, 2003). Another NR that was likely to be involved in the induction of imposex was RXR, found in the mollusc *T. clavigera,* which could bind 9-cis retinoic acid and organotins (Horiguchi *et al.*, 2007). Recent work of Pascoal et al. (2012) using *N. lapillus* demonstrated that retinoid signalling is involved in inducing imposex and the group also

indicated the involvement of androgen signalling (Pascoal *et al.*, 2012). Moreover, at the time my experiments were conducted, no AR gene had yet been reported in any mollusc species and this remains an area of controversy (Scott, 2012b).

During the course of my studies, the draft genome for *B. glabrata* became available. This presented the ideal opportunity both to search for AR-like receptors in this species and to characterize the full complement of NRs in this species. In tandem with this, NRs were further investigated in the mollusc *Lottia gigantea,* for which 26 nuclear receptors had already been identified (Bridgham *et al.*, 2010; Wu, Niles and LoVerde, 2007; Wu *et al.*, 2007), to compare with the snail NR repertoire. These two species are from different families and, thus represent distant lineages within the class gastropoda. I have also assessed the evolutionary conservation, divergence, and uniqueness of molluscan NRs in comparison with the receptors of a vertebrate (*Homo sapiens)*, insect (*Drosophila melanogaster)*, and nematode worm (*Caenorhabditis elegans).*

4.3   Materials and Methods

The following methodological steps were performed to identify the NRs in *B. glabrata* and *L. gigantea*:

1) Screening of *B. glabrata* and *L. gigantea* sequence databases using the Basic Local Alignment Search Tool (BLAST) looking for expressed sequence tags (ESTs), contigs, transcripts and proteins showing low, medium and high degrees of similarity to known NRs from human, fly and worm;

2) 12 'local' databases were made from the 12 tissue transcript data FASTA files using Seqtools (http://www.seqtools.dk/) to reaffirm the prediction of NRs.

3) *B. glabrata*12 tissue transcriptome database in Seqtools were screened (BlastN) using GenBank ESTs (E-value: 1e-98; Descriptions: 10; Alignments: 2) and the same was done against the genome (Vectorbase E-value: 1e-80).

4) Identification of 40 putative NRs from *B. glabrata* and 33 NRs from *L. gigantea.* For *L. gigantea*, the available expressed sequence tags (EST) were screened against assembled *L. gigantea* transcripts from the filtered gene models (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html); experimental assessment of NR gene expression in *B. glabrata* by RT-PCR to generate products for sequencing;

5) Editing of *B. glabrata* gene sequence chromatograms and computational analyses to identify domains and motifs followed by reverse blast to reaffirm the sequence characteristics.

6) Multiple sequence alignment followed by comparative cross-species phylogenetic analysis of mollusc NRs with the human, fly and worm.

4.3.1 Identification of putative NR genes from *B. glabrata* and *L. gigantea*

Human AR (GenBank: EAX05380) and ER (GenBank: AAI28575) gene sequences were used to search for ESTs in *B. glabrata* and *L. gigantea* on National Centre for Biotechnology Information (NCBI) using TBlastX with default parameters. ESTs are short fragment of complementary DNA (cDNA) that represents the expressed portion of the genome (Parkinson and Blaxter, 2009). These ESTs were then used to search for homologues at the preliminary *B. glabrata* genome interface (version 4.01) (1e-04) (http://129.24.144.93/blast_bg/2index.html) and assembled *L. gigantea* transcripts from the filtered gene models in the JGI Genome Portal (http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html) using BlastN (1e-99) with scoring matrix as BLOSUM62. This was followed by a wider search using nuclear receptors spanning all the groups in humans and *D. melanogaster,* Only 15 NRs from *C. elegans* have been included in this study for comparative analysis as they are orthologues of NRs that are broadly conserved among animal phyla (Robinson-Rechavi, Garcia and Laudet, 2003; Sluder and Maina, 2001). A secondary search using the identified *Lottia* and *Biomphalaria* NR genes against both the genomes using a lower threshold value came up with additional family members, some of which were distinct from human and fly NRs. The query sequences were filtered for low complexity regions. All of the identified contigs from *B. glabrata* were downloaded and processed through GENSCAN (1e-10) (http://genes.mit.edu/GENSCAN.html) to identify predicted coding regions, intron-exon boundaries, and peptide sequences. The putative peptide sequences were then assessed using the BLASTP (1e-10) interface against the non-redundant GenBank database to check for the presence of the NR domains (LBD/DBD). All the nucleotide sequences from both molluscs containing NR domains were aligned separately using Clustal X user interface in Seqtools (8.4ver) and analysed for redundancy using clustering analysis with sequence overlap cut-off 0.5 and segment coverage cut-off of 0.25.

A systematic nuclear receptor nomenclature (functional annotation) for the identified genes was based on the sequence similarity and phylogenetic analysis of the full-length sequence to that of previously characterised human and *Drosophila* NRs. Where there was no vertebrate or invertebrate homologue, nuclear receptor abbreviation NR followed by subfamily, a capital letter for the group and receptor number for the individual gene was used (Nuclear Receptors Nomenclature committee, 1999) (Germain *et al.*, 2006).

4.3.2 Comprehensive search for NRs in RNA-Seq 12 tissue transcript data and Vectorbase

Towards the end of my PhD, I got an opportunity to be part of an international effort on *Biomphalaria* genome annotation and had access to the final *B. glabrata* genome assembly and all raw data discussed in detail in Chapter 5. This gave me another chance to do an exhaustive search and look for any NRs that did not come up in the previous search on the draft genome. Local libraries were created in Seqtools (8.4ver) corresponding to 12 tissue RNA-Seq transcript data provided by the consortium. ESTs for NRs submitted to GenBank were used to search for identical transcripts in the 12 local libraries in Seqtools using BlastN (E-value: 1e-98; Descriptions: 10; Alignments: 2). These ESTs were also used to search for predicted transcripts from the genome by performing BlastN searches in Vectorbase (https://www.vectorbase.org/) using default parameters. The best hit from transcripts and scaffold was then scanned (Genome Browse button on the top right end of the Blast result page in Vectorbase) to get the entire protein summary and all the domains and conserved features in the identified transcript. The nucleotide sequence of the identified transcripts were downloaded and aligned using MUSCLE in Seqtools and Clustal Ω (http://www.clustal.org/) and analysed for splice variants and if not then got a consensus sequence. An exhaustive search was performed in the genome (Vectorbase) using PFAM domain analysis (PF00104 and PF00105) (Bateman *et al.*, 2004) and PANTHER, a hidden Markov model-based method (PTHR24082) (Thomas *et al.*, 2003) and INTERPRO (IPR001628; IPR000536 and IPR001723) (http://www.ebi.ac.uk/Tools/services/web/toolresult.ebi?jobId=iprscan-I20131216-231529-0380-30756832-oy) to identify all the predicted transcripts with NR domains. TBlastX with default parameters was done using these sequences on National Centre for Biotechnology Information (NCBI) to perform a cross species check for the domains. A local alignment was performed between the RNA-Seq transcript and the *B. glabrata* scaffold to identify splice sites to signal exon/intron boundaries. The BLAST results for the scaffold with the lowest E-value for the query were selected and downloaded and its corresponding GFF3 (Generic feature format v3, described in details in Chapter 6) file containing the genomic features such as exon-intron boundaries was also downloaded.[1]

---

[1] The detailed methods section describing the annotation process can be found in Chapter 6.*

The rationale for using the above mentioned bioinformatics methods was to perform an exhaustive search to identify homologues for androgen receptor and other NRs in the *B. glabrata* draft genome that was released at the start of the project and in *L. gigantea* genome. The steps were repeated when the RNA sequencing data was made available to reaffirm the results with the final assembly released towards the end of my PhD.

4.3.3 PCR amplification of *B. glabrata* NRs

4.3.3.1 Snail material

Snail material was obtained from different stages of the life cycle (Adult, 96 hours and 120 hours embryo, 14 dph juvenile). Two adult snails were swiftly killed by chopping off the head/foot region through the shell and their shell removed starting from the outermost whorl to inner whorls. Once the snail was removed from its shell, the wet weight of the body was recorded and the whole snail was transferred into individual clean micro-centrifuge tubes containing 4ml RNAlater^{TM} (Ambion) to prevent RNA (ribonucleic acid) degradation. Six egg masses (2 X 3) each from embryos 96 hours post hatch and 120 hours post hatch embryos were collected and transferred into clean micro-centrifuge tubes containing 3ml of RNAlater^{TM} (Ambion). All micro-centrifuge tubes were subsequently stored in the fridge (4^0C).

4.3.3.2 RNA extraction

Total RNA was isolated from whole adult *B. glabrata* snails using TRI-reagent (Sigma-Aldrich). The whole snails were homogenized using motor and pestle adding 1ml of Tri-reagent per 100mg of snail tissue until no lumps could be seen. The tissue homogenate was centrifuged (12,000 x g, for 10 minutes at 4^0C) and the supernatant was mixed with 0.1ml of 1-bromo-3-chloropropane per ml of TRI reagent used. The sample was shaken vigorously for 15 seconds and then allowed to stand at room temperature for 15 minutes. The mixture was then centrifuged (13,000 x g for 15 minutes at 4^0C). The upper aqueous phase containing RNA was collected in a fresh tube and 0.5ml of isopropanol per 1ml of TRI reagent added. The mixture was allowed to stand for 10minutes at room temperature before centrifugation (13,000 x g for 15 minutes at 4^0C), which led to precipitation of RNA as a pellet. The precipitated RNA pellet was washed with 1ml of 75% ethanol per 1ml of TRI reagent, vortexed followed by centrifugation (13,000 x g for 15 minutes at 4^0C) (Based on manufacturer's instructions). The extracted RNA was solubilized in 30μL of RNase-free H$_2$O. Each RNA sample was quantified using spectrophotometry (Nanodrop, Fisher Scientific Ltd.), and was visually analysed by agarose gel

electrophoresis. All samples were then stored at – 80 °C.

For the egg masses, the RNeasy Fibrous tissue mini kit (Qiagen kit) was used. The egg masses were homogenized using motor and pestle adding 300µl of lysis buffer per 30mg of tissue for 15-20 seconds. The lysate was mixed with 590µl of RNase free water and 10µl of proteinase K per 30mg of tissues followed by incubation at 55^0C for 10 minutes to aid in tissue digestion. The homogenate was centrifuged, and the supernatant was collected and mixed with 96-100% ethanol to provide appropriate binding conditions. The samples were subsequently applied to RNeasy Mini Spin Columns, washed and eluted following manufacturer's instructions. The kit has a DNase step to remove contaminating genomic DNA. The extracted RNA was solubilized in 30µL of RNase-free H_2O. Each RNA sample was quantified using spectrophotometry (Nanodrop, Fisher Scientific Ltd.), and was visually analysed by agarose gel electrophoresis. All samples were then stored at – 80 °C. (See Table 4.0.2 for the yield).

Table 4.0.2: RNA extraction report using Nanodrop showing the yield from adult, 96 hours and 20 hours embryos (3 egg masses (Ems) in each time point X two replicates).

Time point	Egg mass no.	ng/µl	Yield (µg)	260/280	260/230	Total weight (g)	Wet weight (g)	Diameter (cms)
Adult	1a	2321.53	116.07	1.71	1.09	15.24	0.6239	0.1549
Adult	1b	2271.84	113.592	1.88	1.01	15.54	0.644	0.1808
120hrs	2a	191.97	5.7591	2.13	0.75	0.1273	0.1273	--------
120hrs	2b	212.12	6.3636	1.88	1.31	0.1708	0.1708	---------
96hrs	3a	72.33	2.1699	2.12	2.15	0.1462	0.1462	---------
96hrs	3b	64.05	1.9215	2.07	1.70	0.1128	0.1128	---------

4.3.3.3 Generation of complementary DNA using reverse transcription

Superscript III First-Strand synthesis system for reverse transcription-PCR kit (Invitrogen) was used to synthesise complementary DNA. For each adult sample, 4µg of total RNA and for embryos 1.5µg of total RNA (Table 4.0.2) was reverse transcribed in a 20µl reaction containing 1µl of 100µM oligo(dTAP) primer (TGACTCGAGTCGACATCGAT$_{20}$) and 1.25mM dNTPs (Bioline). RNA and primer mix was briefly centrifuged, incubated at 65 °C for 5 minutes and then placed on ice for 1 minute. This step allowed the annealing of the oligo (dT) primer to the mRNA poly-A tail. Successively, 2µL of 10X RT buffer, 4µL of 25mM MgCl2, 2 µL of 0.1M DTT, 1 µL of RNase (40 U/µL) and 1 µL of Superscript III RT (200 U/µL) were added, mixed and briefly centrifuged. The mixture was then incubated at 50 °C for 50 minutes to allow the synthesis of the cDNA. The reaction was terminated by incubation at 85 °C for 5 minutes. Any residual RNA was removed by adding 1 µL of RNase H (2 U/µL) to the reaction and

incubating it at 37 °C for 20 minutes. cDNA samples were stored at -20 °C until further analysis. Reverse transcriptase-PCR (RT-PCR) with 18S primers (18S-F: CGCCCGTCGCTACTATCG and 18S-R: ACGCCAGACCGAGACCAA) (Hertel, Adema and Loker, 2005) verified successful cDNA synthesis.

4.3.3.4 Polymerase Chain Reaction

Each of the identified NR sequences from the B. glabrata draft genome were used to design specific PCR primers (

Table 4.0.3) using PRIMER3 web software (www.biotools.umassmed. edu/bioapps/primer3_www.cgi), for amplifying fragments from DBD and/or LBD to confirm transcription and sequence. The primers were designed keeping in mind the following parameters: guanine and cytosine (G and C) content of 40-60%, G and/or C nucleotide at the 3" end, and melting point between 55-63 °C. The primer pairs had similar melting point ($\Delta T < 4$ °C). NetPrimer and the java tool OligoAnalyzer 3.1 (www.eu.idtdna.com/analyzer/Applications/ OligoAnalyzer) were used to assess the likelihood to form secondary structures such as hetero- dimers, self-dimers and hairpin. All primers were between 18 and 23 nucleotides long and were purchased from Sigma Aldrich. Stock primer solutions were made in molecular grade water (100μM).

For conducting PCRs, a 25μl mix was prepared containing 2μl cDNA (diluted 1 in 20 for adults and 1 in 10 for embryos), 1 X PCR Buffer (Bioline), 2.5mM $MgCl_2$, 0.2mM dNTPs, 0.5μM forward and reverse primers and 1.25U *Taq* DNA polymerase (Bioline). The optimal annealing temperature for each primer pair was determined by performing RT- PCR reactions with a gradient of temperatures ($\Delta = 0.8$ °C). After gel electrophoresis of the PCR products, the temperature at which the amplification of the target fragment had an apparent better yield and specificity was chosen for the subsequent reactions.

Generally, PCR amplifications were performed as follows: -

- Hot Start: 95 °C for 2 minutes ⎫ 1 cycle

- Denaturation: 95 °C for 30 seconds

- Annealing: 55-63 °C for 30 seconds ⎫
- Extension: 72 °C for 1.30 minute ⎭ 34 cycles

- Final elongation: 72 °C for 5 minutes ⎫ 1 cycle

- Final hold: 4 °C

Table 4.0.3: Oligonucleotide primers used in RT-PCR analyses of nuclear receptor gene products with their product size (as submitted to GenBank) and their corresponding GenBank accession numbers.

Putative NRs	Forward primer	Reverse primer	Genbank Acc.	Fragment size (bps)
COUPTFA	GTCCGAGAGCAACAACAACA	CTGACGAGAACAGAGCGATG	JZ390916	773
NRU1	TGAGTTTTTCAGGCGGAGTT	GCAGCAGACAAGACACCATC	JZ390900	867
HR38	CACCGCACAATAGGTCTCC	TTGTAGGTGCAGTGGTCTCG	JZ390932	816
HR3	AGCAGCCTTCAGTTCTCACC	CAGGGCGATTGAGATTGG	JZ390936	913
PPAR2DBD	GGTGACAAAGCGTCTGGATT	CTTTTTGGCAAGGGTGGTAA	JZ390908	492
PPAR2LBD	AAGAGTGCCAGCCATTTAGC	CATTGCCAGTTCTGTTTCCA	JZ390909	469
HNF4	CAAGTTGGCGACTCTGTGAA	GGCTGTTGAGGACTGGGTTA	JZ390914	914
ER	CACCACCCCTGAATCTGTCT	ACTTGCCGCATCTCTTGAAT	JZ390913	1097
ERR	GATCCTTCCCACGACGACTA	TCAGCATCCAACATTTCCAA	JZ390912	1033
ROR	CAGCCCTACCCTCAACACTC	CTATGGCAACAGGTCAAAG	JZ390933	879
EcR	AGCAGTGCTTACCACCCATT	AGTCCAGGACGCTCTGAAAA	JZ390899	999
FAX1	GCAAGCACTACGGGGTCTAC	TCAACACGCTTTGCCTCTAA	JZ390935	731
HR4DBD	CGTCCTCTTCTGGGGTGATA	CTGGGATTGTCTCTGGCATT	JZ390920	779
HR4I	GGATGGTTCTTGCAGCTG	GTCGATGCGTTCAACTTC	JZ390921	833
HR4LBD	CCCTCCAACTCATCAAGCAT	CTGAGAAGGTGGGCCTGTAG	JZ390922	372
NR1D1	TGTCAACCTAACCCGTGGAT	GGAAGTGGGCTGAGAACATC	JZ390906	938
NR1D2	GCTGTAAGGCTTGTCGGTTC	GACTCGGTTGGCTTCTGTTC	JZ390907	1070
NR1D3	AACCCCACTGAACGAGACAC	GGGCTCCATCCTTTTTGAA	JZ390902	814
NR1J3DBD	GAGACTCGGAGCACAACACA	GATCGGCCAGCTTCTCTATG	JZ390897	855
NR1J3LBD	CAGTGGAGAGACAGCAACGA	CAGCAACATCTTGGAGTGGA	JZ390898	661
TRDBD	GCCTCTGTGGAAACCAAAAA	CCCCGACAAGCATAACCTAA	JZ390926	376
TRI	GAACGTCGTCCAAGTGAC	GTCTGGAGGTCAGCATTG	JZ390927	701
TRLBD	CGTTCAGTCACAGCCTTTCA	ACCAGTAATCCCAGCCACAG	JZ390928	608
NR4A	GACAGCATCACCCCAGAGTT	ACGGCAGTCGCAAAAGTATC	JZ390905	928
NRU2DBD	AATGGGCCGAAAAAGAAAAC	AGAAGGGGTAAGCGGTGAAT	JZ390903	805
NRU2LBD	ATTCACCGCTTACCCCTTCT	GCTAGGCTGGACTCAACTGG	JZ390904	894
2DBDNR	CCTCCAATCATCCATTCCAC	CTCCAGTCGGGGCTTTTATC	JZ390896	737
TLX	CGGAATATGGCGGAGTTTTA	CACGGAGGGATGGTAGAAGA	JZ390910	699
RXR	ATCTGTGCAATATGCGGTG	AAATGTGTCAATGGGCTGGT	JZ390931	1059
RAR	TTCCTGTGAGGGCTGTAAG	GCTGGACAATGATGGTGATG	JZ390929	1030
E78C	CTTTCGACGCAGCATACAGA	CGGAGTTCAGGCAGTTTCAT	JZ390894	905
HR39	CCCTGTGTGTGGGGATAAAG	CTGGAAGCCTCAAGAGCAAG	JZ390901	772
NRU3DBD	GGCAACAAAACCAAACGACT	TTTCACGAGGACATTCACCA	JZ390923	379
NRU3I	CTGCCCGATTCTATTGAG	CTTCGGCATTGGCTTTG	JZ390924	633
NRU3LBD	TGCACCAAGACACCAGAGAG	TGTTGGAGTAGGAAGCACGA	JZ390925	908
PPAR1	CGTCCGACAGTGGCAAACA	TGGCATGACATGGCTGAAAT	JZ390930	1074
NR1J2	ACAACTTTGACGCCATCTCC	CAACTCAATGCTCCCACCTT	JZ390911	582
NR1J4	GAAAACAGAGCAAAACGTGAGC	CATTAGGGGTTCCACATCCA	JZ390915	862
NR1J1	CAATCCAAGCGTGAACAAAG	GGGGAGGCAGTGTCAAAGTA	JZ390895	509
FTZ-F1DBD	CTTACCGACGGAAACTCTGC	CCCATTCCACCATCAAAAAC	JZ390917	573
FTZ-F1I	GCCACCTGGCTCTCCTG	CAGACACTGTGGACATTC	JZ390918	907
FTZ-F1LBD	GTTTTTGATGGTGGAATGGG	ACTTGTCCGTCATCTCTGGG	JZ390919	415
DSF	CCTCAAGTCACCGGAAAGA	AATGAAGATGGATGGCTGC	JZ390934	518
E75	GTCTTATGCAGGGTGTGTG	CCTTGTGGTTGGTGTTGATG	JZ390939	945
Rev-erb DBD	GTGACGAGTCCTCCGGTTTT	TGAAGGCTGTGTGGTTGTTG	JZ390937	367
Rev-erb LBD	CCTCTTCAACCACACAGCCT	GACTAAAGCGAGCCACCATC	JZ390938	560

Some of the NRs had separate primers for DBD, intermediate (I) and LBD regions.

4.3.3.5 Gel electrophoresis

The DNA fragments amplified by PCR were visualized by gel electrophoresis using 1-1.2% agarose gel with 5 µL of Gel red (Sigma) added and mixed (0.005%). The gel was decanted into a plastic mould with a comb inserted, in order to create empty wells. 7 µL of gel-loading buffer (Sigma Aldrich) were added to 3µL of each sample and loaded onto the gel. The gel-loading buffer increased the density of the sample, preventing the settlement of the PCR product in the gel's well. The dye (bromophenol blue) is negatively charged in neutral buffers and migrates in the gel in the same direction as the DNA during

electrophoresis. This allows the monitoring of the progress of the electrophoresis. The wells on both the extreme ends were loaded with 4 µL of 1 Kb DNA ladder (Bioline) to detect the size of the PCR fragment. The gels were typically run at 85 Volts for 40-60 minutes, and successively examined under ultraviolet (UV) illumination. Each sample was analyzed for the presence of amplified fragments having the expected size as calculated during primer design, see (Figure 4.0.4). All the PCR products with expected fragment size and a clear single band were sent for sequencing to the Sequencing facility, Wolfson Welcome Biomedical Laboratory, Natural History Museum.
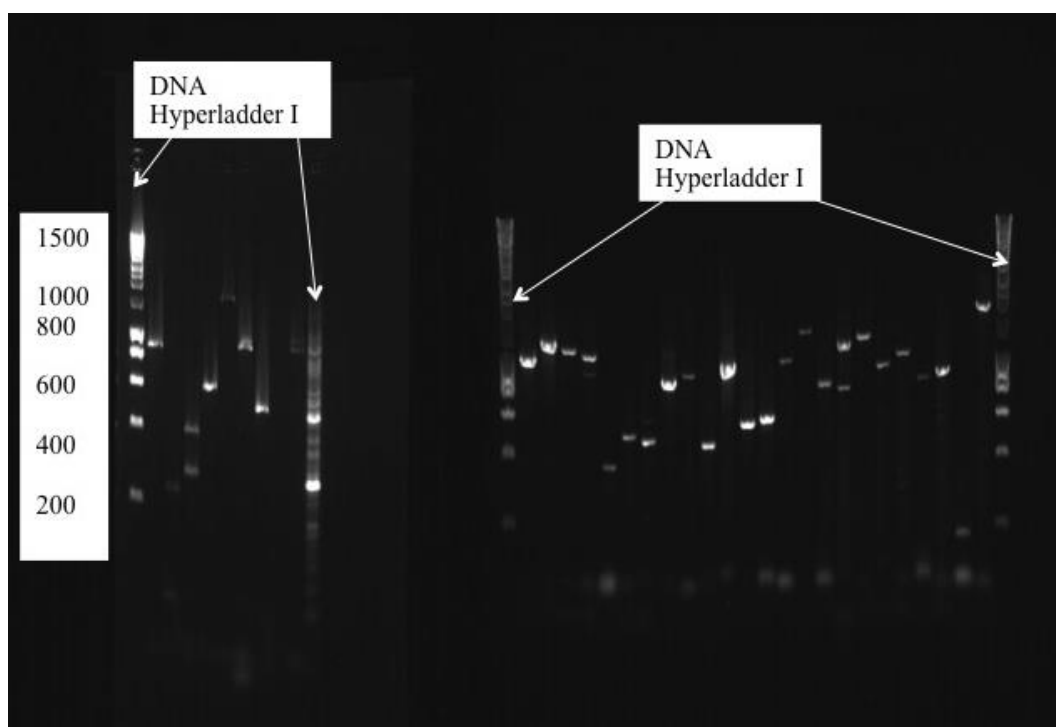
Figure 4.0.4: : The UV image of the PCR products representing NRs in B. glabrata.
The product bands obtained are clean, clear and at approximately the correct size in base pairs (bps).

4.3.3.5 *Extraction of PCR product from the gel*

For PCR products that showed double banding when run on an agarose gel, or which suggested the presence of more than one product when the sequencing electropherogram was examined, cloning was employed to obtain usable sequences for further analysis. The 25µL PCR product was run on the agarose gel and a fragment of the correct size was excised from the agarose gel under UV light using a clean, sharp scalpel (Swan-Morton, Sheffield). DNA extraction was done from the gel using the MinElute gel extraction kit (Qiagen), following the protocol indicated by the manufacturer (not described here). The eluted DNA was collected, quantified by spectrophotometry, and stored at -20 °C until further analysis.

4.3.3.6 *Cloning DNA fragments*

DNA cloning is a method to assemble recombinant DNA molecules for producing large amount of identical DNA fragments. The fragment of DNA to be amplified is inserted in a vector usually a highly modified phage or plasmid that can replicate in a host cell. Below is the brief description of the various steps involved.

STEP 1: *Ligation step to create recombinant DNA:* The PCR reaction mix was prepared using 1μl of 10X buffer, 1μl of 2mM dNTPs, 0.5μl of 50mM Mgcl$_2$, 0.1μl of Taq and made to 10μl by adding clean extracted DNA. The mix was incubated for 10 minutes at 72^{0}to replace A overhangs. 3μl of this PCR product was added to 5μl of Rapid Ligation buffer (manufacturer/concentration), 1μl pGEM$^®$ easy vector (50ng) and 1μl of T4 DNA ligase in 5μl DNase free micro-centrifuge tube. The contents were spun and left overnight at 4^{0}C for maximum number of transformants.

STEP 2: *Preparation for transformation step*: Lactobacillus (LB)/ampicillin/isopropyl thiogalactoside (IPTG)/5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-Gal) plates were prepared to grow the colonies. LB medium was prepared following laboratory protocol. The medium was divided into two parts and 3.75g of agar added to one part before autoclaving. The autoclaved mixture (LB+agar) was then allowed to cool to 50^{0}C before adding 175μl of 100mg/ml ampicillin, as excessive heat would degrade the antibiotic. 25ml of LB+agar was poured into each petri dish (measurement) and stored at 4^{0}C. Before use, these plates were supplemented with 20μl of 500mM IPTG and 20ul of 50mg/ml X-Gal spread over the surface of an LB-ampicillin plate and allowed to absorb at 37^{0}C.

STEP 3: *Transformation*: The tubes containing ligation reaction were centrifuged and 2μl of each ligation reaction was added to a sterile 1.5ml micro-centrifuge tube on ice. JM109 high efficiency competent cells, (which are cloning strain of *Escherichia coli* used to generate high quality plasmid DNA) were thawed on ice and then mixed by gently flicking it to avoid disrupting the cells. 50μl of cells were carefully transferred to the prepared ligation reaction tubes and incubated on ice for 20 minutes. The cells were heat-shocked for 45-50 seconds in a water bath at exactly 42^{0}C and without shaking returned to the ice bath for 2 minutes. This was followed by adding 950μl of SOC medium (Bioline) at room temperature to the ligation reaction transformations, incubating it for 1.5 hours at 37^{0}C on shaking incubator. 50-100μl of each transformation reaction was spread onto duplicate LB/ampicillin/IPTG/X-Gal plates and incubated overnight at 37^{0}C to grow colonies. A plasmid that does not have DNA insert in the α-peptide coding region of β-galactosidase will produce blue colour and if the α-peptide is inactivated by DNA insert, it will have

white/beige colour. Single white colonies were picked using a sterile pipette tip and grown up overnight in 10ml of LB and 5μl of ampicillin (100mg/ml) in a 50ml of centrifuge tube shaking at 37^0C. The next day, the mix is quite cloudy and ready to be used as a template to run PCR.

STEP 4: *PCR using cloned cDNA*: The culture was diluted 1:10 and the cells denatured at 95^0C for 10 minutes. It was then placed on ice, spun down to collect all the content at the bottom and then 5μl of this template used to run a 50μl PCR reaction as described above.

Amplified products were analysed by gel electrophoresis using 1% agarose gel and 5μl of Gel Red for the presence of amplified fragments and the insert size calculated using the formula:

M13 reverse sequence primer binding site………………………...197 base pairs (bp)

Total vector size………………………………………………3015 bp

T7 RNA polymerase promotor…………………………………...2993bp

M13 forward sequence binding site………………………………...22bp

Total estimated product size =197+22+estimated insert size. This is the distance from the primer on the gene to the end of the gene and the distance from the primer on the vector to the end of the vector (assuming that one primer sits on the cloned gene and the other sits on the vector).

4.3.3.7 *Sequencing*

The PCR products and 10mM of the forward and reverse primers were prepared and were sent to the Sequencing facility, Wolfson Wellcome Biomedical Laboratory, Natural History Museum for Sanger sequencing using Applied Biosystems 3730xl DNA analyser. The results from the sequencing were delivered as chromas plot showing the nucleotide sequence of the DNA product, as shown in Figure 4.0.5. The quality of the sequencing at the beginning of the process and towards the end is usually low, as indicated by the height and definition of the peaks (indicated in grey Figure 4.0.5). When the nucleotide peaks become well defined and higher, the base calls are considered reliable. Most of the sequences obtained were less than 1000bp. The sequences obtained were analyzed by computational methods, including reverse Blast, ClustalX alignment and Seqtools with the predicted sequences to get the consensus sequence. Sequences were submitted to GenBank accession: JZ390894-JZ390939.
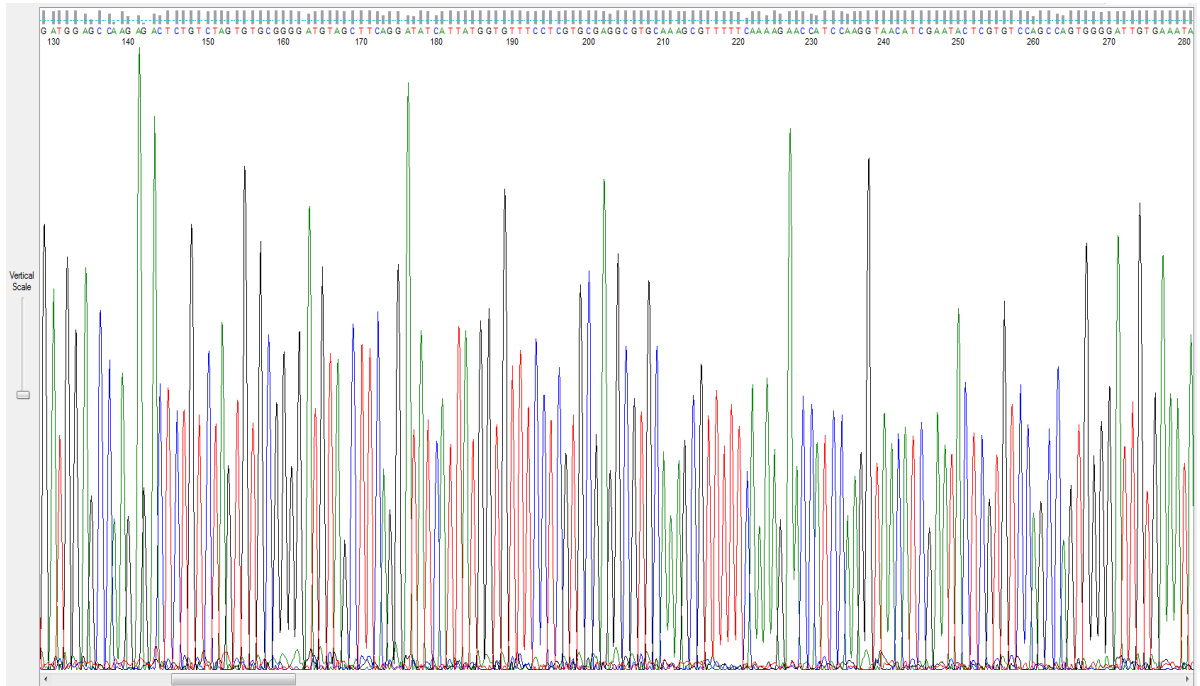
Figure 4.0.5: Example of a chromatogram file indicating the nucleotide sequence of a fragment of a NR expressed B. glabrata analysed after sequencing.

Note the increasing quality of the peaks, indicated by the quality bar in grey at the top of the graph going above the green line threshold. Typically the first 20-40 sequenced bases show low quality peaks and are not shown in the figure above.

4.3.4 Computational analysis

NR nucleotide sequences were translated using proteomics tools available on the EXPASY website (http://ca.expasy.org/sprot/). The translation process normally generated six possible reading frames, three in forward frame and three in reverse direction. Translation for frame starts from the first nucleotide, frame 2 from the second and frame 3 from the third. Frames 4 to 6 correspond to 1 to 3 but in the reverse direction. The frame that produced the longest reading frame with no stop codons was selected. The translated peptide sequences from *B. glabrata* and *L. gigantea* were analysed using different signatures on InterProScan (Thomas *et al.*, 2003) Figure 4.0.6. A further analysis was done with the Simple Modular Architecture Research Tool (SMART) for the identification of DNA-binding domain (DBD) and ligand binding domain (LBD) regions (Letunic *et al.*, 2006). Multiple sequence alignment of the concatenated domain regions (DBD and LBD) from mollusc NR sequences with the human (*H. sapiens)*, fruit fly (*D. melanogaster)* and nematode (*C. elegans*) NRs domains (obtained using conserved domain database (CDD)) (Marchler-Bauer *et al.*, 2007) were performed. Altogether 158 sequences from all the five species (*H. sapiens*, *D. melanogaster, C. elegans*, *B. glabrata* and *L. gigantea*) were compiled into a data matrix and aligned using ClustalX software (Thompson *et al.*, 1997). The aligned file (Appendix 4.1) was then converted into Nexus format, for phylogenetic analysis using Mesquite (ver2.75).
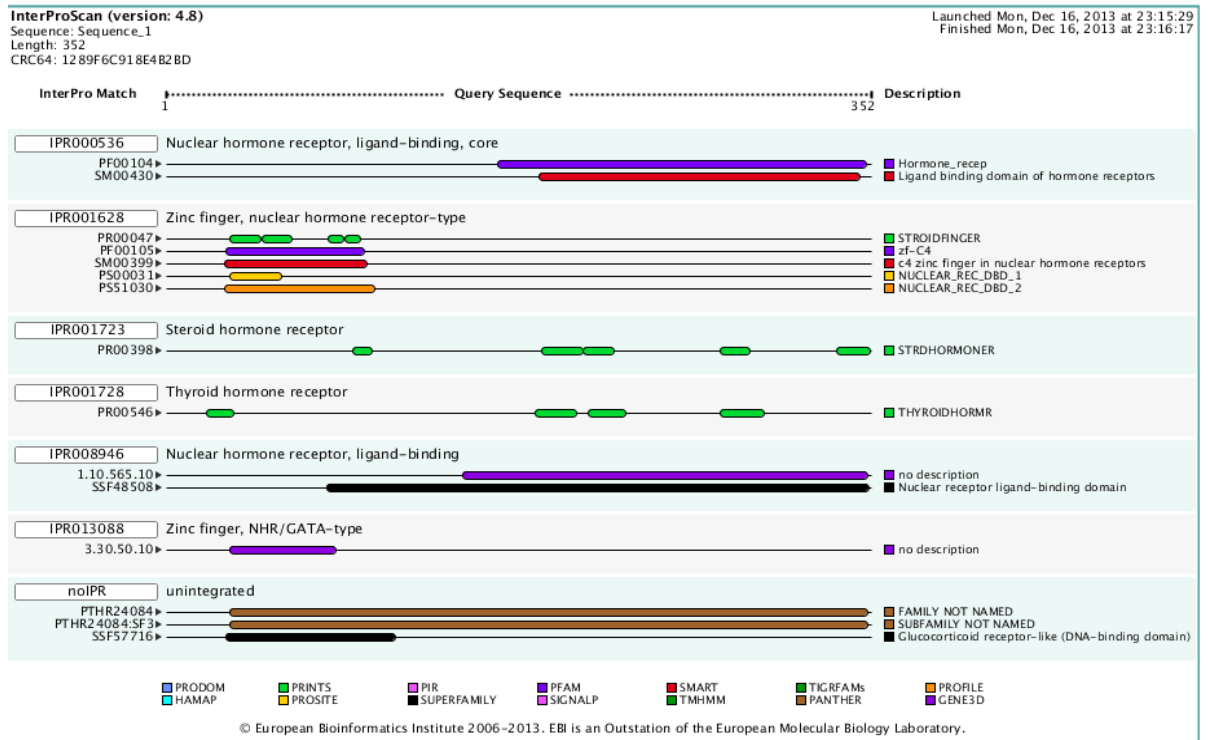
Figure 4.0.6: Example of the domain analysis using InterProScan4.
An online bioinformatics tool able to identify protein "signatures", such as functional domains,
repeats and important sites in the amino acid sequence of the nuclear receptor entered as input.

4.3.5 Phylogenetic analysis

Phylogenetics is an approach of reconstructing evolutionary history and understanding the biological classification of genes, organisms and species (Thain, M. & Hickman, M., 1995). Molecular phylogenetic techniques use sequenced fragments of nucleotides/amino acids to build phylogenetic trees and the analysis is carried out using various methods, including maximum parsimony (MP), Neighbour joining (NJ), maximum likelihood (ML) and Bayesian inference (BI).

BI is a probabilistic model-based method using the software MrBayes version 3.1.2 (Ronquist 1 and Huelsenbeck2, 2003), freely available computing cluster Bioportal (http://www.bioportal.uio.no) run by the University of Oslo. This method of phylogenetic analyses is similar to maximum likelihood. BI methods use a set of prior assumptions about the data matrix to infer the probability that a hypothesis may be true. The posterior probability for a hypothesis is proportional to the likelihood multiplied by the prior probability of the hypothesis. BI analyses allow complex models of molecular evolution to be included (Lewis, Holder and Holsinger, 2005; Holder and Lewis, 2003). The best model for a particular set of alignment can be calculated using ModelTest

(http://www.molecularevolution.org/software/phylogenetics/modeltest) for nucleotide sequences and Prottest (Abascal, Zardoya and Posada, 2005) for protein sequences. BI uses the Markov chain Monte Carlo (MCMC) algorithm that is used to sample probability distribution of the sample data. A Markov chain is a sequence of random variable and the distribution of each variable depends on the value of previous random variable. The algorithm works by taking a series of steps, each step forming a new link in the chain that moves through different trees and models of evolution. BI was conducted for the identified NRs with MrBayes version 3.1.2 using the WAG model that was selected as the best fitting substitution model with both Akaike information criterion and Bayesian information criterion frameworks (determined using default parameters by Prottest (v1.4)). There were two independent runs of the MCMC analysis, with four chains (one cold and three hot chains) run for five million generations with temperature 0.10 Priors having default values and sampled every 1000 generations.

The first 1.25 million generations were discarded (burn-in fraction) to avoid bias in posterior probability calculations. For this the Log likelihood values were plotted against the generations and it was found to be asymptotic well before the burn-in fraction. Convergence between the independent MCMC runs was examined by the average deviation of the split frequencies (0.006) and the potential scale-reduction factor (PSRF), which was 1.00. Clades with posterior probabilities > 95% were considered well supported. An additional BI tree was constructed using subfamily 1 of the NR superfamily to explore the position of newly identified BgTHR homologue, with respect to other selected species. There were two independent runs of the MCMC analysis, with four chains (one cold and three hot chains) run for one million generations, priors having default values and sampled every 500 generations. The first 600 samples (300000 generations) were discarded as burn-in because the Log likelihood values were plotted and found to be asymptotic well before the burn-in fraction. Convergence between the independent MCMC runs was examined by the average deviation of the split frequencies (0.01) and the potential scale-reduction factor (PSRF), which was 1.00. The tree was midpoint rooted in dendroscope (ver 3.2.1).

The maximum Likelihood method is similar to BI and uses probabilities of change from one character to another to calculate the likelihood that a given phylogenetic tree would lead to the dataset observed (Felsenstein, 1981). This is computationally very intensive and consequently demands a lot of computer time when large data matrices are involved. ML was carried out using PHYML v2.4.4 (Guindon and Gascuel, 2003), under Jones-Taylor-Thorton (JTT) substitution model (Jones, Taylor and Thornton, 1992) with a

gamma distribution of rates between sites (eight categories, parameter alpha, estimated by the program).

Maximum parsimony is a character-based method and looks for trees requiring the minimum number of changes to explain the data or in other words by minimizing the total tree length. This method uses PAUP version 4.0b10 (Swofford, 2003) to run the simulations. This process is only valid if evolution itself always proceeds along the most parsimonious route, a situation that seems very unlikely to be the case (Felsenstein, 1985). Furthermore, parsimony also suffers from the problem of long branch attraction. This means that the sequences at the end of long evolutionary branches are similar and this is not because of the close evolutionary relationship but due to convergent evolution (Holder and Lewis, 2003; Felsenstein, 1978). For MP, the settings used were:

Heuristic searches

Branch swapping set to tree-bisection-reconnection

Topological constraints not enforced and

Initial maximum tree setting at 100.

4.3.5.1   *Statistical support in phylogenetic analyses*

In MP analyses, statistical support for clades was calculated using bootstrapping, in which the original data matrix was randomly re-sampled and replaced with duplications of other characters (amino acids) chosen randomly from the same matrix (Felsenstein, 1985). The tree-building algorithm was then run using the modified dataset. This procedure was repeated 1000 times, providing a percentage value (bootstrap value) for each clade that represents the proportion of retrieved trees when the given clade is seen at that position. For example, if the procedures is repeated for 1000 times and a clade occurs in the same position 750 times, the bootstrap would be 75%. The higher the bootstrap value, the more robust is the position of that clade in the tree (Swofford, 2003). Bootstrap was also used as statistical method for ML trees and branch support was evaluated by 1000 bootstrap replicates.

Statistical support for BI analyses is given as posterior probability for each clade (Holder and Lewis, 2003), which is dependent on prior knowledge (prior probability) and new data, weighed by their relative precision (Likelihood function) (http://www.fil.ion.ucl.ac.uk/spm/course/slides10-vancouver/08_Bayes.pdf). As the run proceeds, the posterior probability of the new tree location is calculated and if higher than the previous location, the new value is accepted as the next link in the chain. If the new location has a lower posterior probability, then it can be only accepted as the new position

if this location appears more number of times in proportion to the present location. Effectively, small moves downward in probability are usually accepted, whereas large falls in probability are rejected. If the new position is rejected, the present position is used again as the new position (Holder and Lewis, 2003; Ronquist 1 and Huelsenbeck2, 2003). The simulation is repeated millions of times and a chain of tree locations is created preferring areas of high posterior probability. The number of generations that a chain reiterates in a particular position is an estimate of the posterior probability of the region. At the end of the simulation the output is an estimated value of posterior probabilities of the given tree being accurate (Holder and Lewis, 2003).

In my analyses of NRs an appropriate out-group to root the sequences was difficult for such a diverse and ancient nuclear receptor family and so the phylogenies were mid-point rooted using PAUP.

4.4 Results

4.4.1 Nuclear receptor genes

40 putative NR sequences were identified from the genome of *B. glabrata* and 33 from *L. gigantea* including the previously identified 26 NRs (Bridgham *et al.*, 2010; Wu, Niles and LoVerde, 2007; Wu *et al.*, 2007). This compares to the 48 NRs reported in the human genome (Robinson-Rechavi, Garcia and Laudet, 2003). Of the 40 NRs identified in *B. glabrata*, 15 were human NR homologues (For example DAX1, THR, TR etc.) and 19 were homologues of fly NRs. In Table 4.0.5, a cross-species comparison of nuclear receptors is shown with NRs from all the selected five species (humans, fly, worm, *B. glabrata* and *L. gigantea*). Table 4.0.4 show the tissue distribution of the NR genes in the RNA-Seq 12 tissue data and the pink shading in the table indicated the presence of a particular NR (identified by contig number) in the corresponding tissue. As evident from Table 4.0.4, *B. glabrata* NRs are most prominently expressed within the central nervous system (CNS) (34/40), Head foot region (Foot) (32/40) and amoebocyte producing organ (APO) (32/40). Twelve NRs show expression in all the 12 tissues. Most prominent of these are BgER, BgERR, BgFTZ-F1, Bg2DBDNRs, BgRAR, BgNR1D members and the unique NR (BgNRU4). An exhaustive search with mammalian and *Drosophila* NRs found no further sequence homologs. It should be noted that representative genes for each of the major nuclear receptor groups have been detected, suggesting the assembled genome provides a good first approximation of total gene content. Investigation of the genes in the final assembly with physical coverage of the genome being approximately 27.5X (GenBank: APKA00000000.1) identified

thyroid hormone receptor (BgTHR), which had not been picked up in the previous version (ver 4.01). Exhaustive search through the RNA-Seq transcript data and Vectorbase did not identify any of the subfamily 3C NRs thus supporting the previous search results.

Table 4.0.4: Distribution of putative neuropeptide precursors in the 12 tissues of adult BB02 snail.
The 12 tissuesare AG -albumen gland; BUC - buccal mass; CNS - central nervous system; DG_HP - digestive gland; Foot - headfoot; HAPO_APO – heart/amebocyte-producing organ; KID - kidney; MAN - mantle edge; OVO - ovotestis; SAL - salivary glands; STO - stomach and TRG – terminal genitalia.

Contig	AG	BUC	CNS	DG-HP	APO	FOOT	OVO	SAL	STO	TRG	KIGD	MAN
534	-	-	-	-	-	-	-	-	-	-	-	-
48	-	-	-	-	-	-	-	-	-	-	-	-
304	-	-	-	-	-	-	-	-	-	-	-	-
2148	-	-	-	-	-	-	-	-	-	-	-	-
1481	-	-	-	-	-	-	-	-	-	-	-	-
221	-	-	-	-	-	-	-	-	-	-	-	-
3667	-	-	-	-	-	-	-	-	-	-	-	-
201	-	-	-	-	-	-	-	-	-	-	-	-
293	-	-	-	-	-	-	-	-	-	-	-	-
2842	-	-	-	-	-	-	-	-	-	-	-	-
1939	-	-	-	-	-	-	-	-	-	-	-	-
1958	-	-	-	-	-	-	-	-	-	-	-	-
1275	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-
954	-	-	-	-	-	-	-	-	-	-	-	-
751	-	-	-	-	-	-	-	-	-	-	-	-
469	-	-	-	-	-	-	-	-	-	-	-	-
2702	-	-	-	-	-	-	-	-	-	-	-	-
3233	-	-	-	-	-	-	-	-	-	-	-	-
5703	-	-	-	-	-	-	-	-	-	-	-	-
1318	-	-	-	-	-	-	-	-	-	-	-	-
1866	-	-	-	-	-	-	-	-	-	-	-	-
180	-	-	-	-	-	-	-	-	-	-	-	-
2718	-	-	-	-	-	-	-	-	-	-	-	-
398	-	-	-	-	-	-	-	-	-	-	-	-
2052	-	-	-	-	-	-	-	-	-	-	-	-
3695	-	-	-	-	-	-	-	-	-	-	-	-
27	-	-	-	-	-	-	-	-	-	-	-	-
74	-	-	-	-	-	-	-	-	-	-	-	-
42	-	-	-	-	-	-	-	-	-	-	-	-
148	-	-	-	-	-	-	-	-	-	-	-	-
73	-	-	-	-	-	-	-	-	-	-	-	-
160	-	-	-	-	-	-	-	-	-	-	-	-
29	-	-	-	-	-	-	-	-	-	-	-	-
22311	-	-	-	-	-	-	-	-	-	-	-	-
3827	-	-	-	-	-	-	-	-	-	-	-	-
250	-	-	-	-	-	-	-	-	-	-	-	-
1296	-	-	-	-	-	-	-	-	-	-	-	-
182	-	-	-	-	-	-	-	-	-	-	-	-
3.151												

Group	*C. elegans*	Accession Number	*D. melanogaster*	Accession Number	*B. glabrata*	Contig Number	*L. gigantea*	Protein Id	*H. sapiens*	Accession Number
0A			KNI KNRL EGON	CAA31709 AAF51627 CAA34626						
0B					DAX1	182‡	DAX1	153776	DAX1 SHP	AAH11564 AAH30207
1A					THR**	3.151	THR	207623	THRa THRb	AAH08851 AAI06931
1B	CeODR-7	AAC46497			RAR	398	RAR	207867	RARa RARb RARg	AAH08727 AAH60794 AAA63254
1C					PPAR1 PPAR2	2052 1275	PPAR1 PPAR2	174409 140450	PPARa PPARb PPARg	AAB32649 AAA36469 AAH06811
1D	NHR-85	AAO39185	E75	AAN11687	E75 Rev-erb NR1D1+ NR1D2+	29 160 1939* 1958*	E75 Rev-erb	170913 168854	Rev-erb-a Rev-erb-b	AAH56148 AAH45613
1E			E78	AAF51692	E78C	534	E78C	163301		
1F	NHR-23	P41828	DHR3	AAA28461	HR3 ROR	73 74	HR3 ROR	167096 155536	RORa RORb RORg	AAH08831 AAH93774 AAA64751
1G	CNR14	AAA96982								
1H			EcR	NP724456	EcR	1481	EcR	170342	LXRa LXRb FXR	AAV38218 NP_009052 AAI30574
1I									VDR PXR CAR	AAB95155 AAD05436 AAY56401
1J	DAF-12 NHR-8 NHR-48	AAD34462 AAP31437 CAD36502	DHR96	AAC46928	NR1J1 NR1J2 NR1J3 NR1J4	48* 954 2148 3233	NR1J1 NR1J2 LgNR1J3	163618 63892 163956		
1K	NHR-1	AAC48174								
2A	NHR-49	CAD57702	HNF4	AAF52702	HNF4	2702	HNF4	73247	HNF4 HNF4g	NP_849180 NP_004124
2B			USP	AAF45707	RXR	3695	RXR	206562	RXRa RXRb RXRg	AAI10999 AAC18599 AAA80681
2C/D	NHR-41	NP_500073	DHR78	Q24142	TR	2718	TR	231979	TR2 TR4	AAA36761 NP_003289
2E	FAX-1 NHR-67	AAD55066 CAA97428	TLL PNR DSF FAX-1	AAB71371 AAF58145 AAF52303 AAF54133	TLX DSF FAX1 PNR	3* 42 148 3827*‡	TLX DSF FAX1 NR2E PNR	167709 135472 168960 232946 171552	TLX PNR	AAL05871 AAD28301
2F	UNC55	CAP16277	SVP	NP_524325	COUP-TFa	5703	COUP-TFa	126514	COUP-TFa COUP-TFb EAR2	AAH04154 AAH42897 AAH02669
3A					ER	469	ER	132166	ERa ERb	AAI28575 AAV31779
3B			ERR	AAL37554	ERR	751	ERR	168715	ERRa ERRb ERRg	AAH63795 AAC99409 EAW93335
3C									GR MR PR AR	AAH15610 AAA59571 AAA60081 EAX05380
4A	NHR-6	CAA85271	DHR38	AAF53914	HR38 NR4a	27 2842	HR38 NR4a	94341 182258	NGFIB NURR1 NOR1	CAG32985 AAH09288 CAI95138
5A	NHR25	CAA91028	FTZ-F1	AAN11667	FTZ-F1 NR5+	1318 201*	FTZ-F1	196916	SF1 LRH1	AAH32501 AAI18572
5B			DHR39	AAN11107	HR39	3667	HR39	94029		
6A	NHR91	CAB60329	DHR4	AAX73355	HR4	1866	HR4a HR4b	128948 162354	GCNF	AAB96828
2DBD NR					2DBDNR1 2DBDNR2	304* 1296	2DBDNR1 2DBDNR2	168696* 173632*		

NR-Unclassified		NRU2	221
		NRU3	293
		NRU4	180
			250*‡

*NRs from fly, human and worm (NCBI accession numbers), compared to identified L. gigantea NRs (protein identification numbers: JGI genome portal version 1.1) and B. glabrata NRs (Contig numbers: Preliminary Bg Genomic Data (version 4.01). Of 40 putative NRs predicted in B. glabrata, 1 was short fragment, which could not be classified and were therefore not included in the table. *NRs not used in phylogenetic tree; + NRs are short fragments that could be multiple partial genes; ‡ NRs not found expressed in B. glabrata **NR found in the final genome assembly.*

4.4.2 Expression of nuclear receptor mRNAs

For *B. glabrata,* the presence of nuclear receptors was evaluated using PCR with receptor-specific primers, to demonstrate that the predicted genes were expressed as transcripts and to confirm their sequence. For some of the genes in *B. glabrata,* the DBD and LDB regions were amplified and sequenced separately to obtain fragments suitably sized for direct sequencing. Fragments were amplified for 31 of the 40 putative receptors from cDNA derived from whole adult snail, while 3 further nuclear receptors, BgTLX (contig 3), BgDSF (contig 42) and BgFAX1 (contig 148), were not expressed in adult snails but were isolated from 96hrs/120hrs embryos. These three NRs were also not expressed in any of the RNA Seq 12 tissues further confirming the results. These receptors belong to the NR2E sub family and may be involved in embryonic development as reported in *D. pulex* (Thomson *et al.*, 2009). All amplicons corresponded to the predicted sequence length and sequencing further confirmed the identity of the gene products. Thus mRNA expression was confirmed for 34 out of 40 identified receptor genes [GenBank accession: JZ390894-JZ390939]. Due to time constraints no PCR was conducted for BgTHR, which was identified from the final assembly. For *L. gigantea*, the available expressed sequence tags (EST) were screened using BlastN and identified transcripts confirmed the expression of the predicted genes and mRNA expression was confirmed for 22 out of 33 identified genes, (see Table 4.0.6). E-value of zero indicates a 100% similarity between the ESTs in the hit accession column and the corresponding NR.

Table 4.0.6: Summary of all the top blast expressed sequence tag (EST) hits for L. gigantea NRs, including the EST description, Genbank accession number and E-value.

Sequence name	Sequence description	E-Value	Hit ACC
LgDAX1	caxs lottia gigantea from radula and heart lottia gigantea cdna clone caxs4873 5 mrna	0	FC610206
LgTHR	caxu lottia gigantea from female gonad lottia gigantea cdna clone caxu8659 5 mrna	0	FC646817
LgRAR	cbgc lottia gigantea 15h 18h embryos lottia gigantea cdna clone cbgc7034 5 mrna	0	FC805285
LgPPAR1	NA	--	---
LgPPAR2	cbbn lottia gigantea larvae lottia gigantea cdna clone cbbn9244 5 mrna	0	FC775242
LgE75	NA	--	---
LgRev_erb	cbbn lottia gigantea larvae lottia gigantea cdna clone cbbn8629 5 mrna	0	FC774887
LgE78C	caxw lottia gigantea from female gonad lottia gigantea cdna clone caxw9550 5 mrna	0	FC673655
LgHR3	cbbn lottia gigantea larvae lottia gigantea cdna clone cbbn11308 5 mrna	0	FC759257
LgROR	caxx lottia gigantea from male gonad lottia gigantea cdna clone caxx3392 5 mrna	0	FC692720
LgEcR	caxx lottia gigantea from male gonad lottia gigantea cdna clone caxx9325 5 mrna	0	FC701413
LgNR1J1	NA	--	---
LgNR1J2	caxp lottia gigantea from radula and heart lottia gigantea cdna clone caxp4376 5 mrna	0	FC587866
LgNR1J3	NA	--	---
LgHNF4	caxp lottia gigantea from radula and heart lottia gigantea cdna clone caxp2428 5 mrna	0	FC584194
LgRXR	cbbn lottia gigantea larvae lottia gigantea cdna clone cbbn11953 5 mrna	0	FC760240
LgTR	caxp lottia gigantea from radula and heart lottia gigantea cdna clone caxp1708 5 mrna	0	FC582901
LgTLX	NA	--	---
LgDSF	NA	--	---
LgFAX1	NA	--	---
LgNR2E	caxp lottia gigantea from radula and heart lottia gigantea cdna clone caxp7152 5 mrna	0	FC592666
LgPNR	NA	--	---
LgCOUP-TFa	cbbn lottia gigantea larvae lottia gigantea cdna clone cbbn11647 5 mrna	0	FC759752
LgER	caxp lottia gigantea from radula and heart lottia gigantea cdna clone caxp6663 5 mrna	0	FC591742
LgERR	cbgc lottia gigantea 15h 18h embryos lottia gigantea cdna clone cbgc18627 5 mrna	0	FC788262
LgHR38	NA	--	---
LgNR4a	caxx lottia gigantea from male gonad lottia gigantea cdna clone caxx755 5 mrna	0	FC699271
LgFTZ-F1	cDNA clone CBBN9480 5-, mRNA sequence 3,4,5,6.5d Larvae (M) Lottia gigantea	0	FC775593
LgHR4a	caxp lottia gigantea from radula and heart lottia gigantea cdna clone caxp4376 5 mrna	0	FC587866
LgHR4b	NA	--	---
LgHR39	NA	--	---
Lg2DBDNR1	caxx lottia gigantea from male gonad lottia gigantea cdna clone caxx20789 5 mrna	0	FC689646
Lg2DBDNR2	NA	--	---

4.4.3 Functional Annotation

Blast2go, InterProScan and GO annotation confirmed the identities of all the NRs, predicting the presence of specific conserved features. Table 4.0.7 shows the top protein Blast hits, their corresponding weight in terms of E-value, different protein signatures identified by InterProScan and GO annotation for the NRs. Overall InterProScan identified protein signatures specific to NRs like DNA binding Zinc finger (IPR001628) and ligand binding (IPR000536) and steroid hormone receptor signature (PR00398) in all the NRs and GO annotation identified biological activities like steroid hormone receptor activity; steroid hormone mediated signalling pathway in all the NRs. The biological process identified for COUP TFII was radial pattern formation/skeletal muscle tissue

development/neuron migration/anterior-posterior pattern specification, which is characteristic of COUP TFII NR (http://www.uniprot.org/uniprot/O09018). The structural annotation defined the gene boundaries and coding regions of the corresponding transcripts in GFF3 files that was manually edited using Artemis/Apollo.

4.4.4 Phylogenetic analyses

Phylogenetic analyses were performed using the conserved DBD and/or LBD regions of expressed NR sequences from *B. glabrata* and predicted sequences from *L. gigantea* aligned to *D. melanogaster*, *C. elegans* and *H. sapiens* NR sequences. Phylogenetic trees were constructed using 4 different approaches: Bayesian Inference (BI), Maximum Parsimony (MP) and Maximum Likelihood (ML). All the models agreed at family and subfamily level (Figure 4.0.7) (Appendix 4.2, Appendix 4.3 and Appendix 4.4) but BI showed greatest resolution at the base of the phylogram. (Figure 4.0.7) shows the BI tree with nodal support generated from all the tree construction methods. Phylogenetic analysis of the nuclear receptor genes from *B. glabrata* and *L. gigantea* compared with *D. melanogaster*, *C. elegans* and *H. sapiens* (Figure 4.0.7) reveals two major clades (NR1/0 and NR2/3/4/5/6) each further dividing and sub-dividing into several sub-clades. All the trees predicted the position of BgNRU1/2/3 close to NR0A subfamily within the NR1 group; however the NR0 subfamily was derived to accommodate nuclear receptors that lack either the DBD or the LBD both of which are present in BgNRU1/2/3. These are different genes with overlapping regions but the exact relationship of each of these receptors to each other is also not resolved. Unidentified NRs have been previously reported in two cnidarian species (*Nematostella vectensis* and *Hydra magnipapillata*) (Reitzel *et al.*, 2011; Reitzel and Tarrant, 2009) but our phylogenetic analysis using these NRs (assigned names same as in (Reitzel and Tarrant, 2009); NVNR1, NVNR2, NVNR3, (Data not shown) revealed that these do not cluster with BgNRU1/2/3 leaving open the question as to whether these NRs were a result of species-specific divergence. The NR1D1/2/3 receptors are partial fragments and, due to the fragmentary nature of the preliminary genome used for this study, we cannot currently determine if they are multiple partial genes or parts of the same transcript. Phylogenetic analyses (Bayesian Inference) of BgTHR were done using NR1 subfamily genes. The BgTHR clustered with human and Lottia THR and the branch is well supported with BMCMC posterior probability of 0.98 (Figure 4.0.8).

In subfamily 2, the position of members of Groups NR2C/D is also not resolved (Figure 4.0.7). Both mollusc genomes each contain one member of this clade that we have

designated LgTR/BgTR based on their sequence similarity with HsTRs, but they cluster with Ce_HR41 and Dm_HR78. The Bayesian tree predicts LgTR to be more closely related to Ce_HR41 with DM_HR78 as the most distal member but the ML tree shows both the mollusc sequences to be closely related. However the ML tree placed Ce_HR41 in NR2D, although this NR is usually classified in NR2C. Clearly, the NR2C/D groups seem to be related and not well defined.

Within the NR3 subfamily, only the position of the GR receptor showed non-concordance between the BI tree and the trees from the alternative methods. All of the trees were in agreement for the NR4/NR5/NR6 subfamilies, with one difference in the position of BgNR4a/LgNR4a and NR5B in the MP tree. The identified genes have been named in agreement with the unified nomenclature system for NRs based on their sequence homology and phylogenetic position.

Overall, the identified mollusc NRs encompass almost all the known NR subfamilies, with representatives divided between 21 groups (Table 4.0.5). Similarities and differences in the NR complement between the species used in the study were also seen. Eight of these sub-family groups 1D, 1F, 2A, 2E, 2F, 4A, 5A and 6A were represented in both invertebrates and vertebrates. The presence of these groups among protostomes and deuterostomes supports the suggestion that these receptors originate in a common ancestor of the bilateria before the divergence occurred (Bridgham *et al.*, 2010; Bertrand *et al.*, 2004). The mollusc genomes also contain NR homologues in 1E, 1J and 5B that are present in insects and nematodes but not in vertebrates. Representation of these groups in *L. gigantea* and *B. glabrata* shows that these nuclear receptors pre-date the ecdysozoa/lophotrochozoan split. Four groups (0B, 1A, 1C and 3A) are present in both molluscs and *H. sapiens* but not in *D. melanogaster* and *C. elegans*. The presence of these receptors in molluscs indicates the possibility of commonality in shared signalling pathways in molluscs and vertebrates, elucidating the evolutionary development of the endocrine system. The presence of additional receptors is most likely the result of gene or lineage-specific duplication events.
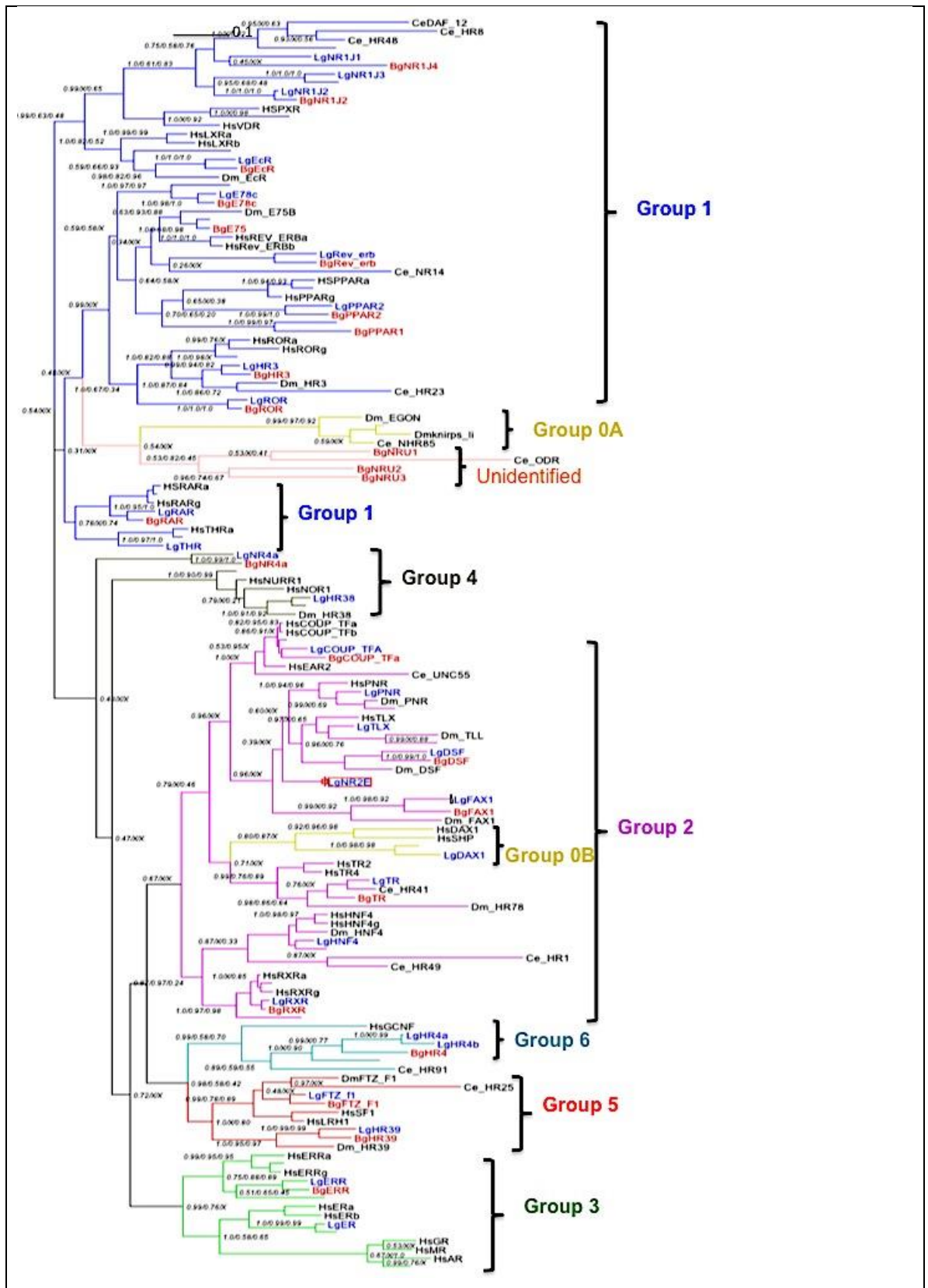
Figure 4.0.7: Phylogenetic relationships of NRs in molluscs, humans, fly and worm.
NRs from five species were subjected to phylogenetic comparisons using Bayesian Inference, Maximum Parsimony and Maximum Likelihood methods. The Bayesian tree is shown with posterior probability values from Bayesian tree and bootstrap values from Maximum Parsimony and Maximum Likelihood trees. The value of 1 on each node represents 100% posterior probability/bootstrap support; an X indicates an area of disagreement from the Bayesian tree (Appendix 4.2: expandable pdf format of this tree with all node support values). Notations Bg, Lg, Hs, Dm and Ce in association with receptor names denote sequences from B. glabrata (highlighted in red), L. gigantean (highlighted in blue), H. sapiens, D. melanogaster and C. elegans, respectively. Scale bar, 0.1 expected changes/site.
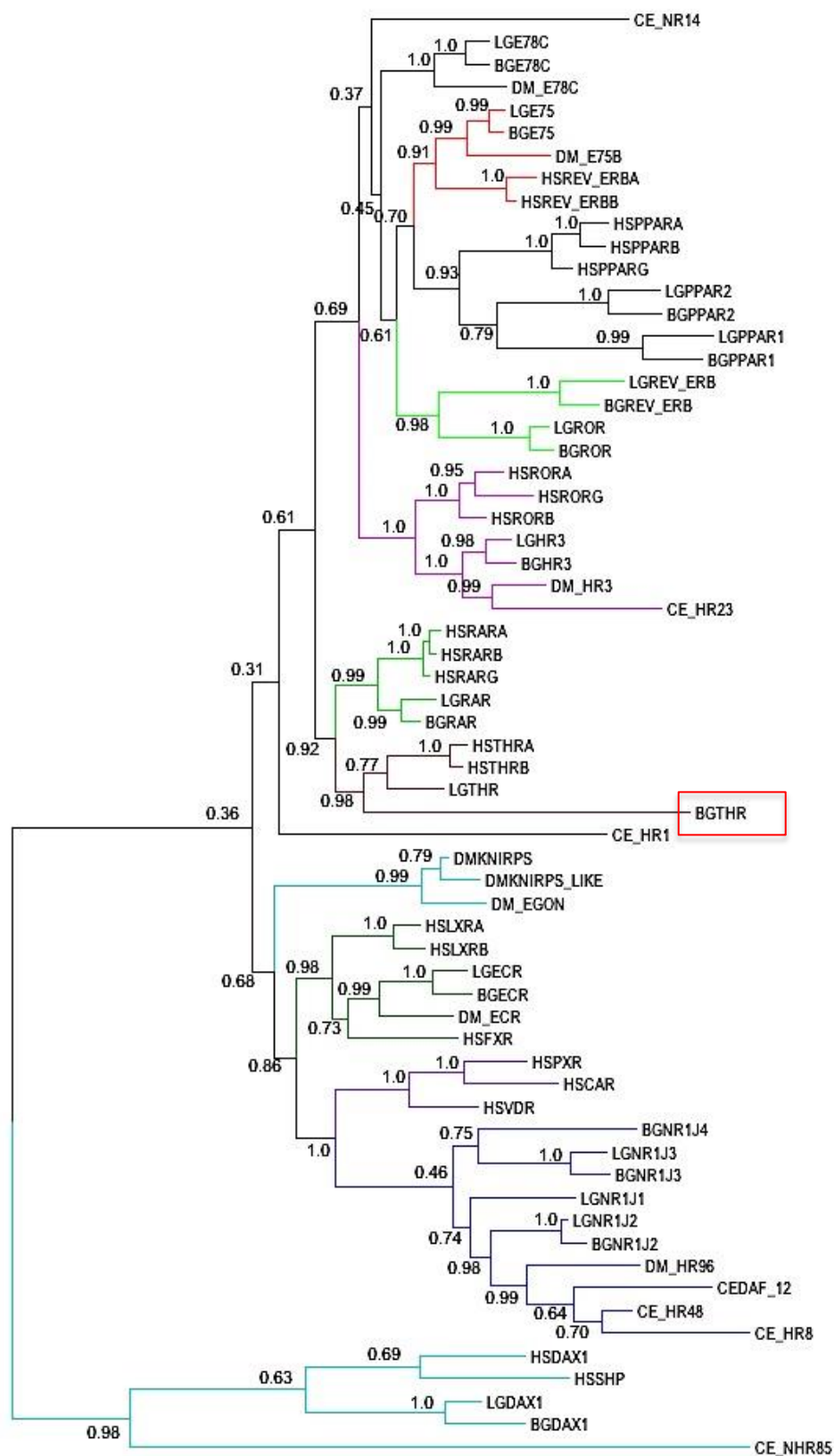
Figure 4.0.8: *Phylogenetic relationships of THRs in molluscs, humans, fly and worm using Bayesian inference showing posterior probability values.*

The value of 1 on each node represents 100% posterior probability. Notations Bg, Lg, Hs, Dm and Ce in association with receptor names denote sequences from B. glabrata, L. gigantea, H. sapiens, D. melanogaster and C. elegans, respectively. Scale bar, 0.1 expected changes/site.

4.5 Discussion

4.5.1 Steroid receptors (Group 3)

Steroid receptors were originally thought to be a vertebrate-specific gene family but the identification of the genes with a clear homology to human ERR in Trichoplax (Baker, 2008), human ER in molluscs and SR in amphioxus (Baker *et al.*, 2003) suggest that these receptors arose early in metazoan evolution and subsequently proliferated in vertebrates through a series of gene duplication events. We have identified vertebrate-like ERs in *B. glabrata* and *L. gigantea*, which have been previously reported in other molluscan species, including bivalves, *Crassostrea gigas* (Matsumoto *et al.*, 2007), *Mytilus edulis* (Puinean *et al.*, 2006) and *Chlamys farreri* (Zhang, Pan and Zhang, 2012), and gastropods, *Aplysia californica* (Thornton, Need and Crews, 2003), *Lymnaea ollula* (Kumkate *et al.*, 2009), *Marisa cornuarietis* (Bannister *et al.*, 2007), *Thais clavigera* (Kajiwara *et al.*, 2006), *Nucella lapillus* (Castro *et al.*, 2007), and cephalopod, *Octopus vulgaris* (Keay, Bridgham and Thornton, 2006). The phylogenetic position of the mollusc ER with the vertebrate ER is well supported with a BMCMC posterior probability of 99% and a bootstrap values of 0.95 (Figure 4.0.7), supporting the suggestion that vertebrate and invertebrate ER diverged from a common ancestor, before the evolution of the deuterostomes (Eick and Thornton, 2011; Baker, 2008; Thornton, Need and Crews, 2003). Structural similarities between ER DBDs of all molluscs, annelids, cephalochordates and vertebrates have been documented showing them to bind to and regulate transcription through estrogen response elements (Katsu *et al.*, 2010; Keay and Thornton, 2009; Bridgham *et al.*, 2008; Keay, Bridgham and Thornton, 2006). The ER LBD in non-vertebrates has unique functionalities, with the molluscs activating transcription constitutively in the absence of ligand (Bannister *et al.*, 2013; Eick and Thornton, 2011; Matsumoto *et al.*, 2007; Castro *et al.*, 2007; Kajiwara *et al.*, 2006; Keay, Bridgham and Thornton, 2006). In annelids the ER has been shown to activate transcription in presence of estrogens (Keay and Thornton, 2009) and in cephalochordates ER LBD transcriptional activity in response to estrogens is lost (Bridgham *et al.*, 2008). No ERs have been identified in ecdysozoa (Maglich *et al.*, 2001), echinoderms (Howard-Ashby *et al.*, 2006)and urochordates (Dehal *et al.*, 2002). One possible explanation could be that independent gain/loss of ligand binding ability has occurred during the evolution of these taxa; alternatively there could be a different origin of estrogen signalling mechanisms in molluscs and vertebrates (Eick and Thornton, 2011; Baker *et al.*, 2003; Escriva, Delaunay and Laudet, 2000). We also identified a single estrogen-related receptor (ERR) in both *B. glabrata* and *L. gigantea*, which clusters with

Drosophila and human ERRs. In molluscs, ERRs (group 3B) have been reported previously in *M. cornuarietis* (Bannister *et al.*, 2007), *C. gigas* [GenBank accession: EKC20050] and *M. edulis* (Puinean *et al.*, 2006). ERRs have no known endogenous ligand, although they are thought to bind to estrogen response elements and may play a role in estrogen signalling and energy metabolism (Bardet, Laudet and Vanacker, 2006). The presence of another mammalian steroid hormone receptor, the androgen receptor (AR), in molluscs has been hotly contested, its existence having been inferred from the effects of androgens in several species of mollusc (Sternberg, Hotchkiss and LeBlanc, 2008; Koehler *et al.*, 2007), rather than by identification of a homologous sequence. We identified no convincing orthologues in either of the two molluscs for the AR or for other members of group 3C (Figure 4.0.7), which also contains the glucocorticoid receptor (GR), mineralocorticoid receptor (MR) and progesterone receptor (PR). The absence of an AR sequence homologue, both in our systematic genomic searches, as well as specific lab-based searches in other molluscs (Santos *et al.*, 2005) (an approach which has successfully identified other receptors) strongly suggests no vertebrate AR homologue exists in these two molluscs.

4.5.2 Group 1 and 2 nuclear receptors

The absence of a molluscan AR suggests that alternative pathways must exist for spermatogenesis/oogenesis in molluscs. I identified homologues of Group 2 C/D vertebrate testicular receptor 4 (TR4) in *B. glabrata* and *L. gigantea* designated as BgTR and LgTR, that in mice have been shown to control spermatogenesis (Mu *et al.*, 2004) and folliculogenesis (Chen *et al.*, 2008). TR4 can cross talk with other NRs like peroxisome proliferator-activated receptor (PPAR), retinoic acid receptor (RAR), thyroid hormone receptor (THR) and estrogen receptor (ER) to control its target genes (Collins *et al.*, 2011). This suggests that there is some degeneracy in the alternative pathways for spermatogenesis and other processes, which exist in vertebrates and which may be important in invertebrates (Greenspan, 2005; Greenspan, 2001).

The identification of PPAR and retinoid X receptor (RXR) has particular significance in molluscs as a recent study of *N. lapillus* suggests PPAR pathways may induce imposex (Pascoal *et al.*, 2012) in response to tributyltin (TBT), which binds to the RXR (le Maire *et al.*, 2009) in this species. I identified two PPAR homologues in both *L. gigantea* and *B. glabrata,* both containing identical sequences to the P-box of human PPAR (CEGCKGFFRRTI). Convergence of 9-cis retinoic acid, a natural ligand for RXR, and PPAR signalling pathways through PPAR-RXRα heterodimerisation has been

demonstrated (Kliewer *et al.*, 1992). *B. glabrata* RXR has been previously shown to strongly heterodimerize with vertebrate PPARα (Bouton *et al.*, 2005), which might suggest the possibility of conservation of several important signalling pathways in molluscs.

A THR homologue has been identified in *B. glabrata* (Figure 4.0.8) that has been previously reported in *L. gigantea* and other invertebrates (Wu, Niles and LoVerde, 2007). This hormone receptor has been reported to be involved in the control of metamorphosis in vertebrates and in prenatal development of central nervous system (Paris *et al.*, 2010; Bernal, 2007a). Besides THR, both the molluscs seem to have ecdysone receptor (EcR), considered play an important role in in *D. melanogaster* metamorphosis/molting (Li and Bender, 2000). Metamorphosis in most of the animals occurs in post-embryonic developmental stages and has been well studied in insects like drosophila for which ecdysteroids and juvenile hormones are the main regulators and in vertebrates where THR seems to be actively involved. On an evolutionary scale molluscs lie somewhere between the insects and vertebrates, so it would be quite interesting to explore the possible mechanisms for these hormone receptors homologues in them. Ecdysone signalling in insects is transduced by EcR and ultraspiracle (USP), which is an insect orthologue of the vertebrate RXR (King-Jones and Thummel, 2005), also present in molluscs. The presence of ecdysone has been reported in pulmonate snails, in *B. glabrata* and *Lymnaea stagnalis* (Whitehead and Sellheyer, 1982) but its origin (endogenous or exogenous) is still contested (Garcia, Griffond and Lafont, 1995). The interaction between the ligand binding pocket of *L. gigantea* EcR homolog and an ecdysone-related steroid and its strong expression in the testis (Laguerre and Veenstra, 2010), may indicate an important role in molluscan developmental processes, potentially by heterodimerizing with mollusc RXR.

Comparison of the NR complements in human, fly, worm and snails show that *L. gigantea* has an additional gene in the NR2E group. In vertebrates, the two members of this NR group, tailless (TLX) and the photoreceptor-specific nuclear receptor (PNR) have important roles in brain and retina development (Benoit *et al.*, 2006). In *L. gigantea*, in addition to PNR, dissatisfaction (DSF), TLX and FAX1 (also present in *B. glabrata*), there is an additional NR2E gene. Since these genes function during development of the visual system in both *Drosophila* and humans (Miyawaki *et al.*, 2004; Yu *et al.*, 1994), it is tempting to speculate that an additional NR2E receptor in *L. gigantea* might assist in its relatively complex visual system as an adaption to the intertidal habitat. Functional analyses of the *L. gigantea* NR2E receptors will provide very interesting insights into the evolution of novel and shared roles for these genes.

The absence of NR1I homologues, which include vitamin D receptor (VDR) and pregnane X receptors (PXR), in *B. glabrata* and *L. gigantea*, is significant as these are considered as natural sensors and are involved in vertebrate response to xenobiotic stress (Benoit *et al.*, 2006). NR1J group in protostomes shares similarity with vertebrate NR1I group (Snow and Larsen, 2000)and there is evidence that both NR1I/NR1J group share a common ancestor (Bertrand *et al.*, 2004). Moreover, recent studies have indicated that NR1J members might regulate xenobiotic responses in *D. melanogaster* and *C. elegans* (King-Jones *et al.*, 2006; Lindblom and Dodd, 2006). DHR96 in *D. melanogaster* plays a role in xenobiotic stress response (King-Jones *et al.*, 2006) and also binds to cholesterol, regulating cholesterol metabolism (Horner *et al.*, 2009). In *C. elegans*, NR1J receptor like NHR-8 plays a role in xenobiotic defence (Lindblom and Dodd, 2006). DAF-12 is involved in dauer formation, in which larval development id diverted as a response to adverse environmental conditions (Antebi *et al.*, 2000). Our results show that *L. gigantea* has three representatives and *B. glabrata* three full-length sequences and one fragment (not shown in phylogenetic tree) (Table 4.0.5) representing NR1J, which segregated from the other invertebrate representatives to form a separate clade (Figure 4.0.7). The mollusc NR1J group members possess well-conserved base contact residues (ESCKAFFR) within the DBD, characteristic of the NR1J sub family (Thomson *et al.*, 2009). In molluscs the NR1J group receptors may be able to perform similar functions to the closely related NR1I of vertebrates.

4.5.3   Group 5 and 6 nuclear receptors

Germ cell nuclear factor (GCNF) homologues (NR6) have previously been reported in bilateria (Enmark and Gustafsson, 2001) and we have identified them in both molluscs. BgHR4/LgHR4a/b cluster with the fly hormone receptor-like (Dm_HR4). During embryonic stages in vertebrates, GCNF can interfere with retinoic acid signalling affecting the expression of *cyp*26*A*1, which is essential for normal hindbrain patterning and early developmental stages (Chung and Cooney, 2003). In Group 5 I identified BgFTZ-F1/LgFTZ-F1 and BgHR39/LgHR39, homologues of *Drosophila* NR5 subfamily members, with a conserved stretch of 23 amino acids adjoining the C-terminal end of the zinc finger motif (AVR**S**DRMRGGRNKFG**P**MYKRDRA). This sequence is placed immediately after the DBD and plays an important role in high affinity interactions of the receptor with DNA (Ueda *et al.*, 1992). All the receptors in NR5 and NR6 have an evolutionary origin preceding that of vertebrate receptors and are functionally active during embryonic development suggesting that developmental NRs are conserved in

bilateria (King-Jones and Thummel, 2005).

4.5.4 Group 0 receptors

The *L. gigantea* and *B. glabrata* genomes each contain one nuclear receptor homologue to the Human DAX1 receptor. Both mollusc receptors lack DBDs characteristic of the NR0B group. The predicted amino acid sequence is part of the conserved domain, LBD and shares 24 and 30% homology to human DAX1 and 64 and 68% similarity to the bivalve, *C. gigas* (EKC26839). DAX1 is a determinant for early mammalian embryonic development and development of embryonic germ cell lineages (Wang *et al.*, 2006) and regulates key steps for androgen production by repressing the transcriptional activation of SF-1 on the promoters of steroidogenic acute regulatory protein (*StAR)* (Lalli *et al.*, 1998). I was unable to generate an amplicon for *B. glabrata* DAX1 in either adult or embryonic stages (96hrs/120hrs). I cannot exclude the possibility that this gene is expressed during specific stages of development or under specific circumstances, not examined in this study. No NR0A members were identified in either of the mollusc genomes based upon the similarity to the invertebrate nuclear receptors of this group.

4.5.5 DBD nuclear receptors

I identified one 2DBD-NR in *B. glabrata* compared to two 2DBD-NRs reported from *L. gigantea*, three from *S. mansoni* [GenBank: XP_002582074] and one from *D. pulex* [JGI Genome Portal Protein Id: 442743]. These NRs have two tandem DNA binding domains, a well-conserved P-box sequence (CEACKK) in the first DBD and a P-box (CEGCK) in the second DBD. The P-box of the second DBD is similar to DBD of NR1 family but the P-box of the first DBD is unique to this NR and may determine novel targets specific to DNA (Wu *et al.*, 2007). The discovery of homologous receptors in both host (*B. glabrata*) and parasite (*S. mansoni*) may open up avenues to further characterize host-parasite interaction and for possible treatments to revert the parasite adaptation capabilities (Maule and Marks, 2006).

4.5.6 Novel Nuclear receptors

Differences in the NRs of *B. glabrata* and *L. gigantea*, indicate interspecies variation within the molluscs, which, given their evolutionary divergence is perhaps unsurprising. A unique subset of four NRs has been identified in *B. glabrata*, which did not associate with any specific NR group. These have been designated as BgNRU1, BgNRU2, BgNRU3, and BgNRU4 respectively and placed in a separate group (Table 4.0.5). These receptors may have originated from a specific duplication event in a *Biomphalaria* precursor, or

alternatively, they could be remnants of ancient NR subfamilies, whose representatives have been secondarily lost in the other represented species.

4.5.7 Molluscs as model organisms

Molluscan endocrinology is poorly understood, so further study to determine the functionality of these NRs promises new insights, especially concerning the many unanswered questions regarding the effects of steroids and other EDCs on molluscs. The identification of the NR genes in molluscs is a first step to understanding the evolutionary development of these important transcription factors. Our results suggest that the vertebrate-like NRs make molluscs a suitable model candidate to investigate functional relationships of individual receptors with their mammalian counterparts. The ultimate choice of model organisms lies not only with the biology of the organism, but also on its tractability for study and manipulation. Aside from the genome, the full range of genomic tools available for *B. glabrata*, such as RNAi (Jiang, Loker and Zhang, 2006), BAC library (Adema *et al.*, 2006b), microarrays (Bayne, 2009b), as well as traditional approaches of enhancer and suppressor genetics and transgenics to explore regulatory networks, make *B. glabrata* a possible candidate. However, the advent of next generation sequencing makes transcriptome study from non-model organisms a real possibility and therefore mollusc species may ultimately make better model systems. In particular, this approach could be used to identify inputs from signal transduction pathways, potential hormone metabolism genes, co-activators, corepressors and other unknown factors that may impinge upon receptor activity. We have shown that there are NR groups with a single molluscan representative for which more than one homologue NRs may be present in humans. For these NR groups, *B. glabrata* could provide a simple model system, not only to study the development and diversification of endocrine system evolution, but also to study potential therapeutic connections. In the NR1 subfamily, both molluscs have single homologues of several significant NRs, such as the retinoic acid receptor (RAR), PPAR, and ROR. RARs have also been identified in the bivalve, *C. gigas* [GenBank: EKC30866] and the echinoderm, *S. purpuratus* [GenBank: XP_779976]. In vertebrates RAR binds retinoic acid, the biologically active form of vitamin A, which mediates cellular signalling in embryogenic anteroposterior patterning of the central nervous system (Chung and Cooney, 2003). Similar functional hypotheses have been made for the RAR of *L. gigantea*, and the annelid, *Capitella teleta* [P.id: 168520] (Albalat, 2009). The presence of THR in molluscs indicates that further investigation of molluscs as models of thyroid disease/dysfunction is warranted. Based on the absence of Group3C NRs, it can be

136

concluded that these two molluscs are inappropriate models for mammalian steroid hormone function mediated via NRs (with the possible exception of ERs), as the genes for several major steroid hormone receptors are not present in either species. However since steroid hormones have also been shown to act via non-genomic mechanisms in vertebrates, using membrane bound receptors (mPR and GPR30) (Levin, 2008)there is still potential to examine alternative pathways in these organisms.

4.6 Conclusions

We identified 40 nuclear receptor genes in the genome of *B. glabrata* and 33 in the *L. gigantea* genome, representing all seven principal vertebrate nuclear receptor groups. An interesting feature emerges when we examine the list of nuclear receptors that are unique to vertebrates. The absence of NRs such as VDR, CAR and PXR, as well as steroid hormone NRs, AR/PR/GR/MR for which we found no orthologues in the two mollusc species examined in this study, indicates that several significant signalling pathways are absent in molluscs. Nevertheless, we have identified an array of NRs common to both vertebrates/molluscs and molluscs/flies. These results could add weight to Thummel's speculation about the convergent regulation of NRs in vertebrates and invertebrates (Thummel, 1996). Elucidation of NR targets in molluscs may unlock their potential as new model organisms allowing the discovery of new pathways leading to similar phenotypes as in vertebrates or, indeed, similar pathways that produce a different phenotype. The range of phenotypes, and their underlying genetic mechanisms, available for study in different species may enable the identification of alternative pathways mediated by NRs that might also be exploited (Greenspan, 2005). The potential of many invertebrate species, for endocrine study, is yet to be explored, but the underlying fundamental similarities and differences between molluscs and man may be the solution to determining not only the evolution of endocrine systems but also the full intricacies of our own.

Table 4.0.7: Results from Blast2go, InterProScan and GO annotations showing the best hits, their corresponding E-value and protein signatures identified using different InterProScan database.

Contig number	Transcript (Vectorbase)	Sequence desc.	Hit description	E-Value	InterProScan	Go annotation
534	BGLTMP002 793+ BGLTMP010 904	ecdysone-induced protein 78c-like	gi|524873840|ref|XP_005093662.1|PREDICTED: ecdysone-induced protein 78C-like [Aplysia californica]	4.29E-85	Zinc finger, nuclear hormone receptor-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway/Thyroid hormone receptor activity
2052	BGLTMP003 233	peroxisome proliferator-activated receptor gamma	gi|524915752|ref|XP_005112667.1|PREDICTED: nuclear hormone receptor E75-like [Aplysia californica]	2.41E-81	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
304	BGLTMP000 531	nuclear hormone receptor	gi|405971703|gb|EKC36526.1|Nuclear receptor ROR-beta [Crassostrea gigas]	5.49E-73	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway
1275	BGLTMP004 271	ecdysone inducible protein 75	gi|524897443|ref|XP_005105182.1|PREDICTED: flocculation protein FLO11-like [Aplysia californica]	1.45E-108	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
1481	BGLTMP013 350	ecdysone receptor	gi|405951903|gb|EKC19773.1|Ecdysone receptor [Crassostrea gigas]	2.01E-77	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
1939	BGLTMP007 582	nuclear receptor subfamily 1 group d member 2	gi|524890231|ref|XP_005101663.1|PREDICTED: nuclear receptor ROR-beta-like [Aplysia californica]	3.10E-70	nuclear hormone receptor-type/ligand binding/Steroid hormone receptor/Thyroid hormone receptor	steroid hormone receptor activity/ DNA binding /nucleus/steroid hormone mediated signalling pathway/Thyroid hormone receptor activity
1958	BGLTMP006 728	nuclear receptor subfamily 1 group d member 2	gi|443721770|gb|ELU10950.1|hypothetical protein CAPTEDRAFT_222512 [Capitella teleta]	3.56E-49	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
201	BGLTMP001 061	nuclear receptor ror-alpha	gi|524898328|ref|XP_005105614.1|PREDICTED: nuclear hormone receptor FTZ-F1-like, partial [Aplysia californica]	2.25E-64	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
2148	BGLTMP004 695	nuclear hormone receptor hr96	gi|524916664|ref|XP_005113105.1|PREDICTED: vitamin D3 receptor B-like [Aplysia californica]	6.74E-64	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
221	BGLTMP012 725	nuclear hormone receptor e75	gi|524915597|ref|XP_005112592.1|PREDICTED: retinoic acid receptor beta-like [Aplysia californica]	9.21E-149	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway
2702	BGLTMP013 098	hepatocyte nuclear factor 4-beta-like	gi|405972081|gb|EKC36868.1|Hepatocyte nuclear factor 4-alpha [Crassostrea gigas]	7.03E-160	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor/Retinoid X receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway
2842	BGLTMP005 503	retinoic acid receptor rxr-alpha	gi|524887228|ref|XP_005100202.1|PREDICTED: steroid hormone receptor ERR2-like [Aplysia californica]	5.20E-157	nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor/Thyroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway/Thyroid hormone receptor activity

293	BGLTMP013426	nuclear hormone receptor e75	gi|524900298|ref|XP_005106570.1|PREDICTED: putative GPI-anchored protein PB15E9.01c-like [Aplysia californica]	9.19E-55	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
3233	BGLTMP004658+ BGLTMP004659	vitamin d3 receptor b	gi|524864409|ref|XP_005089041.1|PREDICTED: nuclear hormone receptor HR96-like [Aplysia californica]	0	ligand binding/Steroid hormone receptor	steroid hormone mediated signalling pathway/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
3667	BGLTMP013844	nuclear hormone receptor ftz-f1 beta	gi|524911108|ref|XP_005110423.1|PREDICTED: nuclear hormone receptor FTZ-F1 beta-like isoform X2 [Aplysia californica]	0	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway
48	BGLTMP010830	protein daf- isoform b	gi|524864713|ref|XP_005089191.1|PREDICTED: nuclear hormone receptor HR96-like isoform X2 [Aplysia californica]	5.37E-63	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
5703	BGLTMP006781+ BGLTMP006780	coup transcription factor 2	gi|405977938|gb|EKC42362.1|Nuclear receptor subfamily 2 group F member 1-A [Crassostrea gigas]	0	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/COUP transcription factor	radial pattern formation/skeletal muscle tissue development/neuron migration/anterior-posterior pattern specification
954	BGLTMP003373+ BGLTMP003372	vitamin d3 receptor	gi|524900058|ref|XP_005106454.1|PREDICTED: nuclear hormone receptor HR96-like isoform X2 [Aplysia californica]	0	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor/Thyroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway/Thyroid hormone receptor activity
469	BGLTMP011477	estrogen receptor	gi|325296813|ref|NP_001191648.1|estrogen receptor [Aplysia californica] >gi|37620222|gb|AAQ95045.1| estrogen receptor [Aplysia californica]	5.21E-174	nuclear hormone receptor-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/DNA binding/nucleus/steroid hormone mediated signalling pathway
751	BGLTMP006863+ BGLTMP006865	estrogen-related receptor gamma-like	gi|115394376|gb|ABI97120.1|estrogen receptor-related receptor [Marisa cornuarietis]	3.23E-161	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway
1318	BGLTMP002005	nuclear receptor subfamily 5 group a member 2 isoform x3	gi|524871094|ref|XP_005092323.1|PREDICTED: nuclear hormone receptor FTZ-F1-like [Aplysia californica]	3.04E-120	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
1866	BGLTMP004549	isoform h	gi|321456767|gb|EFX67867.1|hypothetical protein DAPPUDRAFT_203579 [Daphnia pulex]	4.05E-39	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
180	XXXXXXXX	flocculation protein flo11-like	gi|524872016|ref|XP_005092775.1|PREDICTED: uncharacterised protein LOC101852903 [Aplysia californica]	3.62E-58	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/steroid hormone mediated signalling pathway
2718	BGLTMP006933+ BGLTMP006934	nuclear receptor subfamily 2 group c member 1	gi|524864962|ref|XP_005089314.1|PREDICTED: orphan steroid hormone receptor 2-like isoform X3 [Aplysia californica]	1.24E-118	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus

3695	BGLTMP009887	retinoid x receptor	gi|75017953|sp|Q8T5C6.1|RXR_BIOGLRecName: Full=Retinoic acid receptor RXR; AltName: Full=RXR-like protein; AltName: Full=Retinoid X receptor; Short=BgRXR >gi|19386469|gb|AAL86461.1| RXR-like protein [Biomphalaria glabrata]	0	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone mediated signalling pathway/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus/retinal binding
182	BGLTMP006937	nuclear receptor subfamily 0 group b member 1-like	gi|524911207|ref|XP_005110471.1|PREDICTED: uncharacterised protein DDB_G0271670-like isoform X1 [Aplysia californica]	3.74E-114	nuclear hormone receptor-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/DNA binding/nucleus/steroid hormone mediated signalling pathway
27	BGLTMP008415+ BGLTMP006500	retinoid x	gi|443700853|gb|ELT99619.1|hypothetical protein CAPTEDRAFT_167161 [Capitella teleta]	2.72E-144	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor/Orphan nuclear receptor	ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity/steroid hormone receptor activity/nucleus/zinc ion binding/steroid hormone signalling pathway/squence-specific DNA-binding
74	BGLTMP006044	nuclear receptor subfamily 1 group d member 2	gi|524895845|ref|XP_005104404.1|PREDICTED: flocculation protein FLO11-like [Aplysia californica]	2.66E-132	nuclear hormone receptor-type/ligand binding/Steroid hormone receptor/Thyroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/DNA binding/nucleus/Thyroid hormone receptor activity
398	BGLTMP008526	retinoic acid receptor beta isoform 1	gi|294854043|gb|ADF43963.1|retinoic acid receptor [Lymnaea stagnalis]	0	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor/Thyroid hormone receptor	steroid hormone receptor activity/ DNA binding /nucleus/steroid hormone mediated signalling pathway/Thyroid hormone receptor activity/squence-specific binding
3827	BGLTMP009437	nuclear receptor subfamily 2 group f member 6-like	gi|443688970|gb|ELT91492.1|hypothetical protein CAPTEDRAFT_53417, partial [Capitella teleta]]	3.84E-35	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
250	BGLTMP011358	nuclear receptor ror-alpha	gi|524900302|ref|XP_005106572.1|PREDICTED: putative GPI-anchored protein PB15E9.01c-like [Aplysia californica]	9.43E-25	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
1296	BGLTMP002838	nuclear hormone receptor	gi|405967531|gb|EKC32680.1|Retinoic acid receptor RXR-alpha [Crassostrea gigas]	1.04E-74	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
3.151+3.152	BGLTMP000581	thyroid hormone receptor alpha isoform 1	gi|6094470|sp|Q91241.1|THAA_PAROLRecName: Full=Thyroid hormone receptor alpha-A; Short=THR-alpha-A; AltName: Full=Nuclear receptor subfamily 1 group A member 1-A	1.35E-33	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
29	BGLTMP009749	e75 nuclear receptor	gi|524892572|ref|XP_005102813.1|PREDICTED: mucin-17-like isoform X3 [Aplysia californica]	0	Nuclear hormone receptor, ligand-binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/nucleus/steroid hormone mediated signalling pathway
160	BGLTMP006451	ecdysone-induced protein 78c	gi|405954484|gb|EKC21910.1|Ecdysone-induced protein 75B, isoforms C/D [Crassostrea gigas]	2.20E-76	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
73	BGLTMP011388	nuclear receptor ror-beta	gi|405950647|gb|EKC18621.1|Putative nuclear hormone receptor HR3 [Crassostrea gigas]	1.94E-174	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone	steroid hormone receptor activity/ zinc ion binding /nucleus/steroid hormone mediated signalling pathway/Thyroid hormone receptor activity

					receptor/Thyroid hormone receptor	
3	BGLTMP002 448	nuclear receptor subfamily 2 group e member 1	gi|443724903|gb|ELU12704.1|hypothetical protein CAPTEDRAFT_226190 [Capitella teleta]	2.75E-12	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding/Steroid hormone receptor	steroid hormone mediated signalling pathway/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
22141	BGLTMP008 289	nuclear receptor	gi|524872628|ref|XP_005093075.1|PREDICTED: vitamin D3 receptor-like [Aplysia californica]	1.96E-13	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus
42	BGLTMP000 234	nuclear receptor	gi|524866526|ref| XP_005090084.1|PREDICTED: nuclear receptor subfamily 2 group E member 1-like [Aplysia californica]	1E-83	Zinc finger, nuclear hormone receptor-type/GATA-type/ligand binding	steroid hormone receptor activity/sequence-specific DNA binding transcription activity/Zinc ion binding/nucleus

5 Chapter 5: Reproductive type neuropeptides in *Biomphalaria glabrata* - identification from the *B. glabrata* genome

5.1 Background

Gastropod molluscs are used extensively as model species for neuro-endocrine processes that regulate reproductive behaviour (Benjamin and Kemenes, 2013; Koene, 2010). There is strong evidence of neuro-endocrine control of reproduction in molluscs (York *et al.*, 2012; Koene, 2010). Investigations, primarily in egg-laying species suggest that regulation is achieved not by any single peptide, but rather by a complex combination of neurohormones (Henry, Zatylny and Boucaud-Camou, 1999; Nässel, 1996; Van Golen *et al.*, 1996; Geraerts, 1992). Some of the genes involved in reproduction such as conopressin, egg laying hormone, Ala-Pro-Gly-Trp-amide-related peptide (APGWa) and myomodulins have been identified in several molluscs (Veenstra, 2010). APGWamide has been found to aid peristalsis in the vas deferens of the great pond snail *Lymnaea stagnalis* (Koene, 2010; Van Golen *et al.*, 1996) and act as a spawning stimulant in the male tropical abalone *Haliotis asinine* (Chansela *et al.*, 2008). It antagonizes the effects of conopressin that causes the vas deferens to contract and ejaculate semen (Koene, 2010). The albumen gland produces epidermal growth factor and trypsin inhibitor, both of which are involved in embryonic development (Nagle *et al.*, 2001). In *L. stagnalis*, the caudo-dorsal cell hormone (CDCH) contributes to oocyte release, egg packaging and oviposition (Geraerts, 1992; Dogterom, Bohlken and Geraerts, 1983) and the dorsal body hormone (DBH) regulates vitellogenesis and female cell maturation (Geraerts, WPM and Joosse, J, 1975) while the schistosomin peptide hormone acts antagonistically, preventing reproduction by inhibiting CDCH and DBH (Geraerts, 1992). There is evidence that cerebral ganglia containing dorsal body cells and caudo-dorsal cells may be the site for synthesis of steroid hormones and are not of neuronal origin (Van Minnen, Smit and Joosse, 1989; Boer, Slot and Van Andel, 1968). Moreover, gonadotropin-releasing hormone (GnRH) has been found to be important in controlling reproduction in both vertebrates and invertebrates (Gorbman and Sower, 2003). An endogenous homologue of vasopressin, a mammalian neuropeptide, has also been reported in several mollusc species (*A. californica; L. stagnalis; O. vulgaris; Conus straitus*) (Gruber, 2014; Cruz *et al.*, 1987).

5.2 An introduction to neuropeptides

Neuropeptides are small proteinaceous signalling molecules produced and released by neurons through a regulated secretory pathway. These can act as transmitters, modulators of the central and peripheral nervous system in both vertebrates and invertebrates (Strand *et al.*, 1991). They are involved in most, if not all, physiological processes in metazoans

such as embryonic development (Ugriumov, 2009), organogenesis (Pickart *et al.*, 2006), disease-host response (Boldajipour *et al.*, 2008; Merritt and Sood, 2007) and intercellular signalling (Caneparo *et al.*, 2007). There is evidence that many peptides can have dual role, as neurohormones and neurotransmitters within the central nervous system (CNS). For example, caudo-dorsal cell hormone (CDCH) acts as a hormone in stimulating ovulation in *L. stagnalis* and also acts like a transmitter, organizing movements underlying egg-laying behaviour. About 70 genes encoding classical neuropeptides have been recognized, that have been grouped into 18 families based on their structural organisation or function (Burbach, 2011). Molecular mapping data from *L. stagnalis* studies on the central ganglia of the CNS show that approximately 2500 neurons of the total 20,000 neurons are peptidergic (secrete peptide hormones) (Benjamin and Kemenes, 2013). A recent peptidomic analysis of neuropeptides involved in copulatory behaviour indicates that the penis complex contains 44 different peptides arising from 10 different genes (El Filali *et al.*, 2006).

Despite a range of functional roles, all classical neuropeptides are targeted to and processed by the regulated secretory pathway through conserved protein motifs (Burbach, 2011). Classically, neuropeptides are short sequence amino acids that are part of larger, inactive molecule precursor proteins (Southey, Sweedler and Rodriguez-Zas, 2008) and contain a signal peptide at the N-terminal that guides them through the ribosome into the lumen of rough endoplasmic reticulum (ER). The N-terminal signal peptide is a hydrophobic sequence of 17-30 amino acids, which is cleaved off upon entry into the rough ER by the action of the signal endopeptidase (Figure 5.1). Only those proteins lacking a transmembrane domain will pass through the ER and eventually be secreted outside of the outer cell membrane. The protein molecule with signal peptide removed is a biologically inactive pro-hormone composed of one peptide or a chain of several peptide (identical or dissimilar) (Hökfelt *et al.*, 2000) which must be proteolytically processed into smaller active peptides. The pro-hormone is transferred to the Golgi apparatus where posttranslational processing starts.

In the Golgi apparatus, packaging of the pro-hormones into large dense core vesicles with processing enzymes (also called convertases) takes place and they undergo extensive cleavage prior to producing the actual bioactive peptides. Proprotein convertase cleaves the pro-hormone at specific cleavage sites that could be flanked by monobasic, dibasic or tetrabasic amino acid combination of lysine-arginine (KR); arginine-arginine (RR) or lysine-lysine (KK) from the C-terminal region, although some separations may occur at a single basic amino acid (Veenstra, 2000). After cleavage from the pro-hormone, the
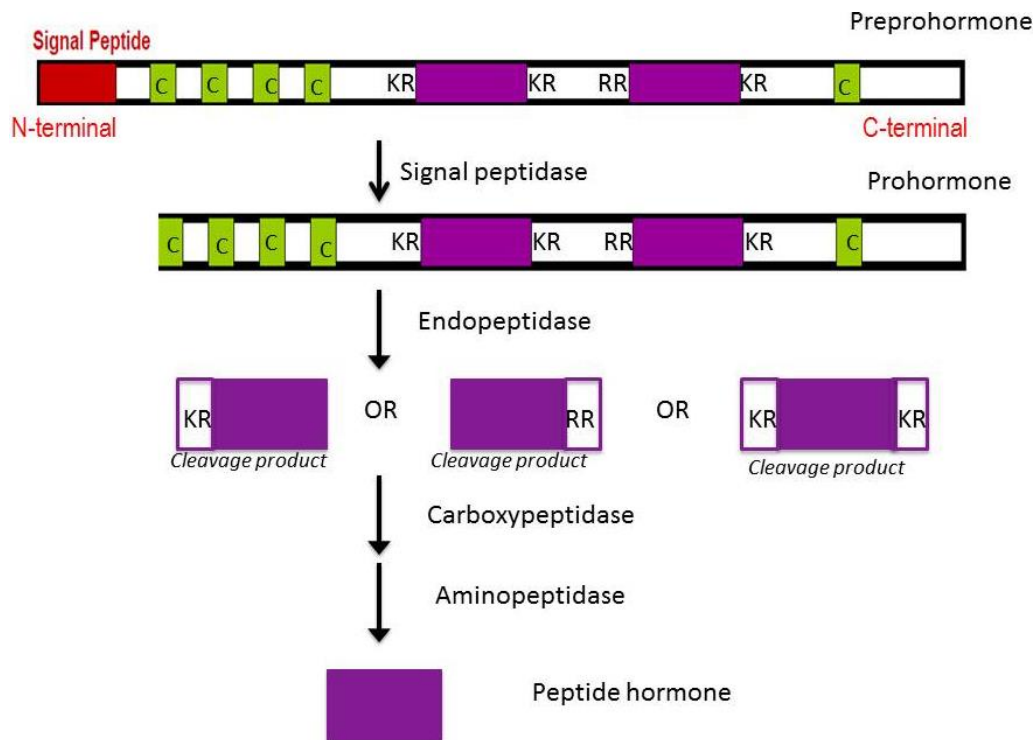
Figure 5.1: Schematic of the posttranslational processing of the prepro-hormone.
The process starts with removal of the signal peptide by signal peptidase forming a pro-hormone. This undergoes cleavage at endoproteolytic cleavage sites forming cleavage products that might undergo amidation, sulphation, and/or acetylation forming peptides. The green boxes indicate cysteine bases that form disulphide bonds during secondary coiling of the protein. The cleavage sites can be dibasic Lysine-Arginine (KR) or Arginine-Arginine (RR).

resulting intermediate peptide contains carboxyl terminal basic amino acids, which are removed by a specific carboxyl peptidase. The carboxyl terminal glycine residue, if present is transformed into an amide (Figure 5.1). This amidation protects peptides from degradative enzymes and is essential for peptide bioactivity (Spijker *et al.*, 1999). Some pre-hormones produce structurally related peptides possessing comparable activities, whereas others have been shown to yield peptides with different potencies (occasionally opposite) (Clynen et al., 2006). For example ghrelin precursor molecule produced two peptides with antagonistic roles in weight regulation (Zhang *et al.*, 2005). *L. stagnalis,* on the other hand offers an example of encoding of structurally related amino acid motifs on one gene, 13 copies of F(X)RI amide encoded on the same gene (Benjamin and Kemenes, 2013). Therefore it is important to identify all peptides originating from a single precursor since each of these peptides might have a different physiological relevance.

Based on the structural and functional characteristics discussed above, a biologically active neuropeptide as defined by Veenstra (2000) is a peptide which has structural similarity to a known neuropeptide, or a peptide being produced in multiple copies which are identical and surrounded by typical neuropeptide convertase sites within the precursor that contains a signal peptide at the N-terminal of the peptide. This definition does not

include spacer peptides, which have often a large number of acidic amino acid residues and in general do not have well conserved copies on the same precursor, or proteins, which have repeated sequences but are unlikely to be neuropeptide precursors (Veenstra, 2000).

Although there is not much information regarding the impact of environmental chemicals on the expression of reproductive type neuropeptides in snails, TBT, a widespread contaminant that has been linked to the induction of imposex in molluscs, was found in higher concentration in nerves and ganglia of *N. lapillus* (Bryan *et al.*, 1993). This led to the hypothesis that accumulation of TBT in ganglia may disturb reproductive-hormonal regulators acting through neuropeptides in molluscs (Oetken *et al.*, 2004). In support of this, when *Ilyanassa obsoleta* were injected with 10^{-12} – 10^{-10} M APGWamide/animal/per day, it caused imposex after seven days of exposure (Oberdorster, Romano and McClellan-Green, 2005; Oberdörster and McClellan-Green, 2000). Conversely, another study with *Bolinus brandaris* (a neogastropod) reported no imposex when exposed to 100 fold higher molar concentration of APGWamide (Santos *et al.*, 2006). Castro et al. (2007) also found no effect after injecting 10^{-9} M/g body weight APGWamide in *Nucella Lapillus* (Castro *et al.*, 2007). The contradiction in these results could be due to the differences in the reproductive behaviour of the snails as discussed earlier (Chapter 2). Other contaminants, which may be found in urban effluents, include neuro-active chemicals like antidepressants. Fluoxetine (a drug used to treat depression in humans and pet animals) induces spawning and parturition in estuarine and freshwater bivalves (Fong and Ford, 2014; Metcalfe *et al.*, 2010) and studies have demonstrated the potential of serotonin uptake inhibitors to disrupt reproductive activity in *Spisula subtruncata* clam populations in the vicinity of sewage effluent outfalls (de la Ossa Carretero, JA *et al.*, 2008).

My fascination for bioinformatics and the availability of complete *B. glabrata* genome was an opportunity for me to gather knowledge regarding the reproductive type neuropeptide precursors in *B. glabrata* that would help to increase our knowledge of the basic reproductive biology and the possible underlying neuro-endocrine processes that might be regulating the physiological responses (in terms of reproductive output) of the snail when exposed to chemicals. The aim of this work was to identify putative reproductive type neuropeptides in *B. glabrata* and assess their structural organisation with known neuropeptides from other species.

5.3 Annotation of reproductive type neuropeptides

Annotation is a process of identifying 'regions of interest/genes' and attempting to assign them a name and/or function. It is a powerful approach to unravel and gather valuable biological information of genes and to assert the accuracy and reliability of the identified genes. As more genomes are sequenced and more annotations become available the evidence based approach to annotate genes becomes more powerful as the chances of homology increases.

The complete process of annotation describes manually validated gene names, and gene names transferred from other annotated genomes due to strong protein sequence homology. It includes:

- Structural annotation: This gives information about the physical location of the gene in the genome (chromosome, scaffold and contig) and its structural organisation (Intron-exon structure);

- Functional annotation: This predicts the different protein domains, links it to three functional signatures characteristics of animal physiology- reproduction, development and metabolism and suggests a gene name based on the presence of homologous/similar functional domains.

The main aim of annotation is to identify genes of interest for the annotators and make a record of this for the community. The work with neuropeptides has been performed in collaboration with Dr Scott Cummins (University of the Sunshine Coast, Queensland, Australia) and Dr Joris Koene (VU University, Amsterdam).

5.4   Methods and materials

5.4.1   Data provided by the consortium

The data used for *B. glabrata* gene annotation was

- 12 tissues transcript data files from adult *B. glabrata* (BB02) having transcripts longer than 500 base pairs (bps). The 12 tissues were AG -albumen gland; BUC – buccal mass; CNS - central nervous system; DG_HP - digestive gland; Foot – head foot; HAPO_APO – heart/amoebocyte-producing organ; KID - kidney; MAN - mantle edge; OVO - ovotestis; SAL - salivary glands; STO - stomach and TRG – terminal genitalia.

- Data used from Vectorbase (https://www.vectorbase.org)

- BB02 Contigs (BglabB1.fa)
  BB02 Scaffold (BglaB1.fa)

BB02 Transcripts (BglaB1.fa)

BB02 Base features (BglaB1.o.GFF3)

BB02 Contig2Scaffold (BglaB1.agp)

BB02 Repeats.lib

- Expressed sequence tags (ESTs) from SnailDB.org

Clustered

Unclustered

The following methodological steps were performed to identify genes in *B. glabrata*

1) 12 'local' databases were made from the 12 tissue transcript data FASTA files using Seqtools (http://www.seqtools.dk/).

2) *B. glabrata*12 tissue transcriptome database in Seqtools were screened (BlastN) using GenBank ESTs (E-value: 1e-98; Descriptions: 10; Alignments: 2) and the same was done against the genome (Vectorbase E-value: 1e-80).

3) Identified sequences from different tissues were aligned using Muscle (Seqtools) and Clustal omega (http://www.clustal.org/) to generate a consensus sequence.

4). Exon-intron boundaries, domains, signal peptides, peptides and repeats were identified using a variety of web-based software.

5). Cross-species phylogenetic analysis of the newly identified genes was performed using Maximum Likelihood trees using MEGA5 (http://www.megasoftware.net). MEGA was used for makinh Maximum Likelihood due to problems with the older version of PAUP and non-availability of the newer version.

6). Identified gene information was aligned with the genomic sequence using Artemis and the edited GFF3 files uploaded in the Community Annotation Portal (CAP) of Vectorbase.

5.4.2 Identification of putative neuropeptides from *B. glabrata* transcript data files

The neuropeptide precursors datasets from *Aplysia californica*, *Lymnaea stagnalis*, *Crassostrea gigas* (Feng *et al.*, 2009; Fleury *et al.*, 2009; Moroz *et al.*, 2006), *Lottia gigantea* (Veenstra, 2010), *Helix aspersa* and *Theba pisana* (Scott et al., personal communication) were compared to three *Biomphalaria* datasets, the genome assembly at Vectorbase, RNA-Seq transcripts in the 12 local libraries (corresponding to 12 tissue transcript data that was uploaded in Seqtools(8.4ver), and a dataset of all available *B. glabrata* ESTs at that time (October 2013), which consisted of 15224 entries from GenBank and 54304 (www.ncbi.nlm.nih.gov) entries from ESTs. The aim was to identify the sequence homologs of mollusc neuropeptide precursors and any additional NRs from the *B. glabrata* datasets with homology based searches.

TBlastN (protein query against the 6-frame translation of a nucleotide database) was used to search with default values and BLOSSUM62 as a substitution matrix as search parameters. A secondary search was performed with the identified sequences with default parameters to identify if there were isoforms present. The identified transcripts were aligned using MUSCLE (Edgar, 2004) in Seqtools and ClustalΩ, analysed for splice variants and a consensus sequence obtained.

5.4.3   Computational analyses

The identified nucleotide sequences were translated using proteomics tools at EXPASY website (http://www.expasy.org). The translation process normally generated six possible reading frames. The frame that produced the longest reading frame with no stop codons was selected. The peptide sequences were analysed using SMART (http://smart.embl-heidelberg.de) confirming their conserved domains, repeats, transmembrane regions etc.

5.4.4   Gene Feature finding for neuropeptides

5.4.4.1   *Identification of signal peptides*

The translated reproductive type neuropeptide data set was analysed with TargetP1.1 (http://www.cbs.dtu.dk/services/TargetP/). All the peptide sequences (fasta format) were fed to TargetP using default parameters. Sequences were considered to be targeted to the secretory pathway if they had SP (secretory pathway) score >0.8, indicating they had a signal peptide. In order to ascertain the presence of signal peptides, SignalP3.0 (http://www.cbs.dtu.dk/), which implements the algorithm described by Bendtsen *et al*. (2004) was used (Dyrløv Bendtsen *et al.*, 2004). This method uses both Neural Network method and the Hidden Markov Model methods and the criteria for selection were based on scoring method for discriminating between secretory and non-secretory proteins, length and composition of the sequence and cleavage site prediction of the protein (Appendix 5.1). Both Neural Network method and the Hidden Markov Model methods and the 'short' and 'no graphics' options were selected. Proteins were considered to have a signal sequence if SignalP returned a "YES" for four out of five scores for the Neural Network method and the Hidden Markov Model method predicted that it contained a signal peptide. Transmembrane domains were then predicted using TMHMM 2.0c (http://www.cbs.dtu.dk) using the default settings. Subsequently all the neuropeptides were analysed with the web version of Phobius (www.phobius.sbc.su.se) using a 'short' option (Appendix 5.1) that predicts transmembrane topology and signal peptide position. Consistency in predictions from each method was assessed to determine if the predictions

were robust.

5.4.4.2 *Identification of the cleavage sites*

Once the pro-hormone sequence was determined, this was analysed for possible cleavage sites with Neuropred (http://neuroproteomics.scs.illinois.edu) that is based on logistic regression models trained on experimentally verified cleavage information (Southey *et al.*, 2006). Advanced search was done with the peptide sequences in fasta file format using following parameters:

- Model selection: Known motif and Mollusc;
- Output Selection tasks: Only predict cleavage site, Obtain mass of predicted sites and Model accuracy statistics;
- Post Translational Modifications: N- and O-glycosylation, Amidation, Acylation, phosphorylation and sulphonation.

 The pro-hormone is processed to yield peptides by cleavage at specific cleavage sites preceded by a basic amino acid residue at position -4, -6 or -8 (Veenstra, 2000). These positions refer to the number of amino acid before the sessile bond, e.g. in the Lysine (Lys)-Arginine (Arg) pair, the Arg occupies position -1, the Lys -2. Generally in vertebrates and insects cleavage after a single Lys residue is very rare (Cohen, Rholam and Boussetta, 1995; Devi, 1991) but this seems to be more common in molluscs (Veenstra, 2000). In *Aplysia* amide gene, convertase cleavage occurs systematically at a single Lys residue supported by the Arg in the -4 position (Schaefer *et al.*, 1985). In pro-hormone the amino acid residue in position +1, i.e. immediately after the cleavage site has a significant influence on cleavage (Cohen, Rholam and Boussetta, 1995; Nakayama *et al.*, 1992; Devi, 1991). Generally cleavage of neuropeptide precursors seems to occur at basic amino acid residues but from *Aplysia* neuropeptide precursors there is good evidence that processing can also occur at other sites (Hummon *et al.*, 2003).

5.4.4.3 *Intron-exon boundaries*

BlastN was performed in Vectorbase using default parameters using the RNA-Seq transcript identified from 12 tissues to identify the scaffold/contig and physical location of exons corresponding to the gene sequence of interest, in the genome. Then a local alignment was performed between the RNA-Seq transcript identified from 12 tissues, GenBank EST (if available), the predicted coding sequence from the Vectorbase and its corresponding scaffold using MUSCLE in Seqtools using default parameters. The splice sites were worked out using the transcript data. As a general rule, most introns start from

the sequence G (U/T) and end with the sequence AG (in the 5' to 3' direction) and identified as 5' and 3' splice sites (also referred to as splice donor and splice acceptor site) (Padgett *et al.*, 1986) (Figure 5.2). Studies have shown that apart from the donor 5' and acceptor 3' splice sites, introns contain a branch site towards the 3' end of the intron that becomes covalently bonded to the 5' end during the splicing process (Guthrie and Patterson, 1988) and do not always follow GT-branch-AG rule (Jackson, 1991; Kiss *et al.*, 1989). To overcome this, extra evidence based on homology-based approach was used to confirm the position of splice sites. The homology-based approach can often predict intron location by identifying gaps in the sequence alignment when the candidate gene is aligned to their homologs.
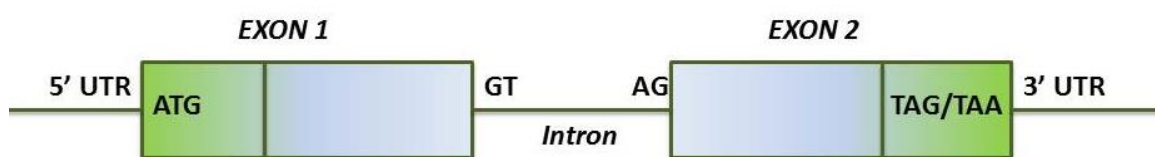
Figure 5.2: Figure showing Standard gene structure.
Starting with untranslated region (UTR) at 5' end, the gene has coding region that translates into a protein, starts with start codon ATG and stops at stop codon (TAA, TAG, TGA) followed by UTR at the 3'end. GT and AG are the standard splice sites present at the start and end of the introns.

5.4.5   Functional annotation of identified reproductive type neuropeptide

The predicted function of the identified genes was confirmed using Blast2go to search public databases, for gene homology as well as additional functional annotation including gene ontology (GO) classification and Kyoto Encyclopaedia of Genes and Genomes object identifiers (KEGG object IDs). FASTA files containing protein sequences for either neuropeptides were uploaded separately in Blast2go (http://www.blast2go.com) (BlastP E-value: 1e-06, word size (default): 3), BlastP searches were performed to identify homologs from different species with e-value less than or equal to 1e-06 and the results were used to map GO terms based on a look-up table matching UNIPROT results to GO terms. BLAST2GO also has the additional benefit of being able to incorporate InterProScan data, and adds this information during the GO annotation step. The InterProScan tool detects protein predictive models or signatures using the integrative data stored in the InterPro database (Hunter *et al.*, 2009) which aggregates diverse information from multiple databases, including Gene3D, PANTHER (Protein Analysis Through Evolutionary Relationships), Pfam (Protein family), PIR (Protein Information Resource), PRINTS, ProDom (Protein Domain), ProSITE, SMART (Simple Modular Architecture Research Tool),   HAMAP (High-quality Automated and Manual Annotation of microbial

Proteomes), SUPERFAMILY and TIGRFAM. In the 2014 release, close to 133026 different InterPro signatures were present in the database, and together with the over 25586 UniProtKB entries makes it a valuable tool for protein functional annotation.

The final set of proteins from *B. glabrata* was aligned with known neuropeptides from different species using ClustalΩ to identify the degree of conservation of peptide sequences. Phylogenetic analysis was performed using the alignment with the Maximum likelihood (ML) method. ML was carried out using MEGA5 (Hall, 2005), under Jones-Taylor-Thorton (JTT) substitution model (Jones, Taylor and Thornton, 1992) with a gamma distribution of rates between sites (eight categories, parameter alpha, estimated by the program). Due to lack of time, MrBayes was not used as it is more computer intensive and PAUP version 4.0b10, could not be used as there was some technical problem with the software CD. So I selected MEGA5 for doing phylogenetic tree analyses as it quite powerful and flexible interface and performs very fast.

5.4.6 Structural Annotation and final Gene prediction

Most of the genes predicted from the genome had their associated annotation files in GFF3 (Generic feature format v3) file containing the predicted genomic features such as exon-intron boundaries. A GFF3 file is a standard text file with nine tab delimited columns (Table 5.1) and was developed as part of the "Sequence Ontology Project" (SO) (Eilbeck *et al.*, 2005), replacing the GFF2 file format developed at the Sanger Institute (www.sanger.ac.uk). The main advantage of GFF3 over its predecessors is the mechanism for representing hierarchical information for the gene with a three-tier system (See top left window in the Artemis gene builder in (Figure 5.3). The gene attribute is the parent to mRNA, which is then parent to the CDS.

Table 5.1: Explanation of the Gff3 file format edited during annotation process.

Column	Name	Description
1	Sequence id	Scaffold or contig number
2	Source	Free text qualifier that generated this feature, Ensembl in this case
3	Type	Term used to define the sequence such as gene, exon, CDS, mRNA etc.
4	Start	Starting base of the feature
5	End	Ending base of the feature
6	Score	A floating point number representing E-value or P-value.
7	Strand	Strand of the feature with respect to the sequence object, - for minus strand and . If not stranded
8	Phase	Reference to reading frame. Can be 0, 1 or 2 depending on the number of bases that are removed to reach the first base of next codon and only applicable to CDS.
9	Attributes	Description of the feature such as name; Parent; target etc

The two genome editing programs considered suitable for this task were Artemis (Rutherford *et al.*, 2000) and Apollo (http://apollo.berkeleybop.org), both of which have been used by many genome projects and can work directly on GFF3 files. Apollo is a user-friendly editor that allows the annotator to update gene models based on the lines of evidence. The sequence aligner lets the annotator detect pseudogenes and highlight sequence discrepancies between expressed and genome sequences (Madupu *et al.*, 2010). Moreover, Apollo can link directly on a remote database and so annotation can theoretically be performed anywhere and fed directly back to the database, useful for groups working in collaboration on similar gene families. Artemis, on the other hand, is java based and displays annotation derived from EMBL feature keys like repeats, promoters, coding sequence etc. (Carver *et al.*, 2008). The Artemis Comparison Tool (ACT) allows comparison between complete and draft genome sequences and their associated annotation making it useful for comparison between different species. Both of the annotation methods were equally valid for this work but due to some technical issue with the web interface of Apollo, Artemis was chosen for annotating genes.

Structural annotation of the identified genes using Artemis enables the visualization of the predicted genes in the updated genome assembly and, using the available evidence (Blast results, ESTs, RNA-Seq data), editing of the exon-intron boundaries and identification of the complete coding sequence (CDS) with start and stop codons and untranslated regions

(UTRs), both 5-prime and 3-prime of the gene. UTRs are difficult to predict as they show high degree of sequence variation and sequence content, but evidence may be present in transcript data. GFF3 files were available for predicted genes, which were used directly in Artemis, in association with the appropriate scaffold sequence. The files were visualized, edited and saved with comments referring to the changes made and evidence used. Each gene annotation on Artemis is composed of 4 main elements (gene, transcript, CDS and exon) (Figure 5.3) that can be visualized in the gene builder in hierarchical order. Modifications were made including intron or exon elimination, exon creation or intron creation as necessary by adding coordinates of the exons in the Gene builder location window. For genes that were not predicted by the automated annotation pipeline, a new annotation element was created by inputting all co-ordinates for exons and introns in relevant reading frame, merging the exons to form the CDS and saving with a new identifier (For example, NEWGENESK01). A community portal for the *B. glabrata* genome project has been constructed (https://www.vectorbase.org/). The site is controlled with a username/- password login and could be accessed to upload the genes. Some of the putative neuropeptides were split over multiple scaffolds, which could be due to fragmented nature of the genome, making the process of gene finding more complex. For these genes editing was done in their respective GFF3 files (if predicted) and rather than uploading these gene directly on the community portal, they were emailed to the Vectorbase team informing them the order of the exons and the scaffolds in which they are located. Three of the RNA-Seq derived transcripts of pro-hormone neuropeptides did not map to the predicted genes in Vectorbase. For example putative FFamide was not found in Vectorbase or the 12 tissue data, but an EST was found in the GenBank entries. For such genes new GFF3 files were created and given an identifier as stated above.
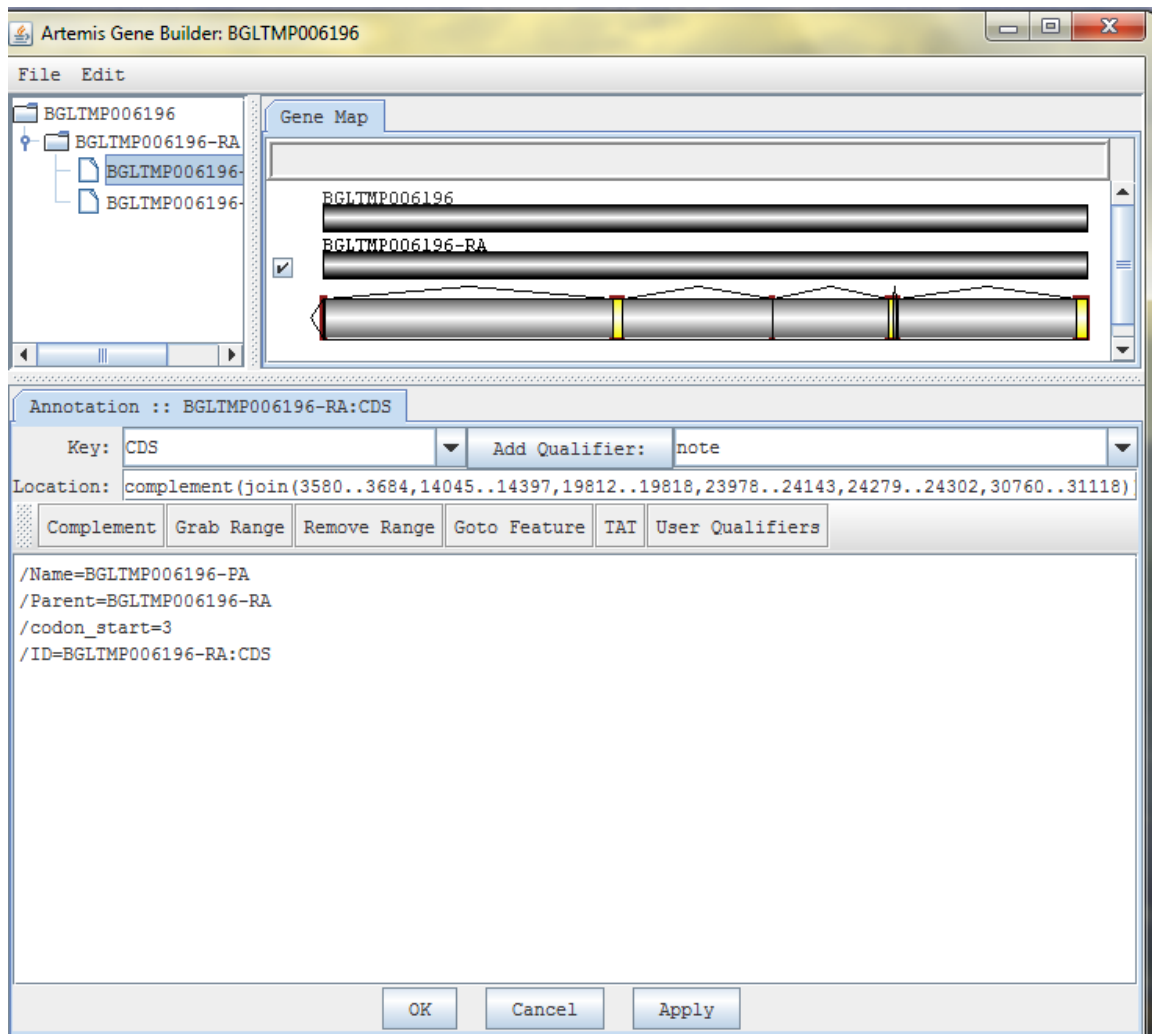
Figure 5.3: Screen shot of the Artemis Gene Builder for APGWamide showing the hierarchical order of gene, transcript, CDS and exon.
The location window shows the start and end base pair co-ordinates of all exons joined together.
The Gene map shows the complete gene model with exon and introns.

5.5   Results and Discussion

5.5.1   Reproductive type neuropeptide genes

Here I describe an evidence-based approach that allowed identification of both evolutionarily conserved neuronal genes - the putative pro-hormones and neuropeptides in the genome of *B. glabrata.* Using computational *in silico* data mining, 17 putative reproductive-type neuropeptide genes and 4 insulin-like neuropeptides were identified from the genome and transcriptome databases for *B. glabrata,* encoding precursors that are predicted to yield a repertoire of small neuropeptides with over 124 peptide sequences similar to known bioactive peptides. Table 5.2 shows the tissue distribution of the neuropeptide precursor genes in the RNA-Seq 12 tissue data. The shaded portions in the table indicate the presence of a particular neuropeptide in the corresponding tissue.

As evident from Table 5.2, reproductive type neuropeptides are most prominently expressed within the terminal genitalia (TRG) (13/21) and central nervous system (CNS) (12/21). ELH-1 and myomodulin-3 are exclusive to the TRG while GnRH and Insulin like peptide -2 and -3 are exclusive to the CNS. Albumen gland, ovotestis and salivary glands expressed fewer of the identified neuropeptides with albumen gland expressing only pedal

Table 5.2: Distribution of putative neuropeptide precursors in the 12 tissues of adult BB02 snail.

Neuropeptide Precursor gene	AG	BUC	CNS	DG-HP	APO	FOOT	OVO	SAL	STO	TRG	KID	MAN
APGWamide			X		X					X	X	
Schistosomin		X			X				X			X
Conopressin		X		X					X			
ELH1										X		
ELH2			X							X		
FFamide												
FMRFamide												X
GNRH			X									
Whitnin						X				X		
Pedal Peptide 1					X				X			X
Pedal Peptide 2	X		X			X				X		
Pedal Peptide 3			X	X								
LFRFamide		X								X		
NPY		X			X	X	X	X	X	X	X	X
Myomodulin 1		X										X
Myomodulin 2		X			X				X		X	
Myomodulin 3										X		
Insulin-like peptide 1			X			X			X			X
Insulin-like peptide 2			X									
Insulin-like peptide 3			X									
Insulin-like peptide 4		X		X					X	X		
TOTAL	1	7	12	4	8	5	2	2	8	13	3	8

The 12 tissuesare AG -albumen gland; BUC - buccal mass; CNS - central nervous system; DG_HP - digestive gland; Foot - headfoot; HAPO_APO – heart/amebocyte-producing organ; KID - kidney; MAN - mantle edge; OVO - ovotestis; SAL - salivary glands; STO - stomach and TRG – terminal genitalia.

peptide-2. As previously reported in other mollusc species, I found that the *B. glabrata* genome contains gene families encoding multiple genes such as; pedal peptides (pedal peptide 1, 2, 3); insulin-like peptides (insulin-like peptide 1, 2, 3, 4) and myomodulin

(myomodulin 1, 2, 3). All these genes are related but quite divergent from each other due to differences in cleavage products. An exhaustive search with known reproductive type neuropeptide precursors and peptides found no further sequence homologs.

5.5.2 Gene features of putative neuropeptides

The translated peptide sequences were analysed for gene features such as signal peptide, transmembrane regions, internal repeats and other conserved domains. All the neuropeptides had secretory pathway score >0.80, indicating the presence of signal peptide (TargetP1.1) (Emanuelsson, von Heijne and Schneider, 2001) and the position of signal peptide was identical between the different software used (see Appendix 5.1). No transmembrane domain was detected in any of the identified neuropeptides except for 3 (conopressin, GnRH and insulin-like peptide4), each of which had one transmembrane domain in the first 60 amino acids. Conopressin shows transmembrane domain from 1-25aa and the signal peptide identified is between 1-23aa. Similarly GnRH has a transmembrane domain from 1-23aa, which overlaps with the signal peptide region, and the insulin-like peptide4 transmembrane domain is from 1-29aa while the signal peptide is 1-25aa. In all the above cases the transmembrane region overlaps with (partially or fully) signal peptide. This shows that a single transmembrane domain identified in these precursor molecules may be a signal sequence falsely identified as a transmembrane domain.

Neuropred predicted a large number of peptides in *B. glabrata* pro-hormones. Of the predicted cleavage products, those that undergo α-amidation at the C-terminal were assumed to be the potential bioactive cleavage products (Spijker *et al.*, 1999). Table 5.3 shows the predicted amides (C-terminal amides arise from post-translational conversion of glycine) and other cleavage products from the *B. glabrata* pro-hormones which were identified by Neuropred and confirmed based on homology searches in other species by multiple alignments to identify conserved cleavage sites. A notable feature observed was cleavage occurring at monobasic cleavage site (Lysine or arginine residue) and supported by either Arg- or Lys- at -4 positions in *B. glabrata* neuropeptide genes, which has been previously reported in *Aplysia* (Schaefer *et al.*, 1985) Monobasic cleavage sites were found in FMRFamide, conopressin, ELH1 and myomodulin1. In FMRFamide, convertase cleavage occurs at a single Lys residue supported by Arg in -4 positions but in conopressin cleavage occurs at single arginine residue supported by lysine at -4 positions. In ELH1, there is one monobasic cleavage site at lysine supported by arginine at -8 positions. In myomodulin1 monobasic cleavage sites are common with single arginine residue

supported by lysine at -4 positions or single Lys residue supported by the Arg in the -4 position.

Table 5.3:Identified reproductive type neuropeptides from B. glabrata showing putative bioactive cleavage products and the number of peptides encoded in the prohormone.

Neuropeptide	No of peptides encoded	Putative Bioactive Cleavage Product
APGWamide (AKA: C-terminally located anterior lobe peptide (CALP))	8	APGW[Amide]
Conopressin	1	CFIRNCPKG[Amide]
ELH 1 (AKA: Caudodorsal cell hormone 1)	1	YPIGNDLELLANMIVGIKKQQNRQEVMDMLEEF[Amide]
ELH 2 (AKA: Caudodorsal cell hormone 2)	1	VPIGNDLLALADLLFIERQKQAYSALKAMMDEA[Amide]
		QPYGDDSEYamide
FFamide	1	RVDSKDMIQQPLLF[Amide]
	1	GVNPNLNSLFF[Amide]
FMRFamide	8	FMRFamide
	16	(X)nFLRFamide
	1	FIRFamide
	1	FIRFamide
	1	QFYRIamide
	1	SDKGPVYMRFamide
	1	EPSRISYPRYamide
GnRH	1	HMSQAQNYHFSNGWYA[Amide]
Insulin-like peptide 1	1	EYDHTCNVLSRPHRN
Insulin-like peptide 2	1	AADQPLSRamide
Insulin-like peptide 3	1	QSSFTVICECCFNQCTVSEMLDYCSLFGRamide
Insulin-like peptide 4	1	QDISACDVSKGTHamide
LFRFamide	1	NTLFRF[Amide]
	1	QGAWFRY[Amide]
	1	AGTLLRF[Amide]
	1	GTLLRF[Amide]
	1	GGSLLRF[Amide]
	1	TLFRF[Amide]
Myomodulin 1	7	ASSFVRI[Amide]
	1	PSSFVRI[Amide]
	1	ASNFVRI[Amide]
	1	FIRI[Amide]
	1	KSSFVRI[Amide]
	3	SPSSFVRI[Amide]
	1	APSSFVRI[Amide]
	1	VPSSFVRI[Amide]
	2	VSSFVRI[Amide]
	1	SSAFVRI[Amide]
	1	IPSSAFVRI[Amide]
	1	EGQSSFVRI[Amide]
Myomodulin 2	1	GLAMLRL[Amide]
	2	SMKMLRL[Amide]
	1	QLKMLRL[Amide]
	1	SLKMLRL[Amide]
	4	PMNMLRL[Amide]
Myomodulin 3	1	SMEMLSL[Amide]
	8	QLKMLKL[Amide]

NPY	1	MDNMLTPPERPAEFKNANELRK YLKALNEY]YAIVGRPRF[Amide]
Pedal peptide 1	6	PFDSISGSHGLSGFA
Pedal peptide 2	1	EVPDGDI
	2	SIDSIGSGFI
	10	PLDSIGSSFI
	1	GIDSIGSAFI
	2	PIDSIGSSFV
	1	NIDSIGSAFI
	9	GIDSIGSGFI
	3	PIDSIGSSFI
	1	SLNSIGSGFI
	1	ALDSIGSSFI
Pedal peptide 3	3	PFDRIGTSSFTSF [Amide]
	1	PFDRIGTSAFTSF [Amide]
	2	PFDRIGNSAFASF [Amide]
	2	PFDRIGNSAFTSF [Amide]
	3	PFDRIGNSAFTTF [Amide]
	1	PFDRIGTSVFTSF [Amide]
	2	PFDRIGSSAFTSF [Amide]
	1	PFDRIGTSAITSF [Amide]
	1	RFDSISDSSAFNHF [Amide]
	1	RFDRISKNSQFNPF [Amide]
	1	RLDRISSISGFGKF [Amide]
	1	KFDSIADGSRFSSF [Amide]
	1	RFDRISSGFSHF [Amide]
	1	RFDRIDRGSAFSRF [Amide]
	1	SSFDRIDKNSAFNAF [Amide]
Schistosomin	1	DNYRCPNPGDAFECFESDATARFCVS [Amide]
Whitnin /SPTR /PKYMDT /Proctolin	1	PKYMDT

5.5.3   Intron-exon boundaries

Compared to *B. glabrata* NR genes, where most of the genes have seven to eight exons, the number of intron/exon boundaries found in selected neurosecretory genes is relatively small, since there are few having more than two exons. Most of the exon-intron boundaries follow the consensus GT-AG rule except one (Pedal peptide 1) that has AC-AG between its two exons. The exons of eight neuropeptides span over two scaffolds and schistosomin has three exons that span over three scaffolds. In most instances the two adjacent exons overlapped by more than a few base pairs, of which only one is likely to be the real exon. Even the best gene predictors and genome annotation pipelines rarely exceed accuracies of 80% at the exon level, which means at least one mis-annotated exon per gene (Reese *et al.*, 1997). Automated annotation programs frequently have difficulty in choosing the correct ORF depending on the training data set or the settings and are most

likely to pick the longest ORF and it is here that manual curation becomes important.

5.5.4 Confirmation of peptide identity and function

Blast2go and InterProScan (http://www.ebi.ac.uk) confirmed the identities of all the reproductive type neuropeptides, predicting the presence of peptide specific conserved features. Table 5.4 shows the protein top Blast hits, their corresponding weight in terms of E-value and different protein signatures identified by InterProScan in the neuropeptides. 5 distinct protein families were detected and the sequences were associated with 9 Protein Information Resource (PIR) domain identifiers. The presence of transmembrane regions in conopressin, insulin-like peptide 4 and GnRH was reconfirmed by TMHMM database, which predicted transmembrane helices in the proteins. The GO annotation predicted myomodulins, APGWamide, FMRFamide, pedal peptides and NPY as part of neuropeptide signalling pathway (GO:0007218). Conopressin has neurophyseal domain, which is the characteristic of the family of neuropeptides showing neurophyseal hormone activity, For example, vasopressin in humans. Hormone activity was indicated for most of the neuropeptides except GnRH, Whitnin and LFRFamide (see Table 5.4)

The structural annotation that defined the gene boundaries and coding regions of the corresponding transcripts in GFF3 files was manually edited using Artemis. Three of the identified neuropeptides (FFamide, Insulin-like peptide1 and Pedal peptide1) did not have a predicted transcript in Vectorbase so a new gene model was created. All the proteins had a complete coding sequence with start and stop codons. UTRs were difficult to predict due to variation in the sequence but based on the 12 tissue sequence data. Only 5 precursor molecules were assigned a 5'UTR region.

Here I describe the structural characteristics of the *B. glabrata* reproductive-type neuropeptides. For some of these detailed cross-species comparison of the peptides using multiple sequence alignment and phylogenetic analyses was performed to elaborate on the global organisation of mature peptides.

Table 5.4: Results from Blast2go, InterProScan and GO annotations showing the best hits, their corresponding E-value and protein signatures identified using different InterProScan database.

Sequence name	Sequence description	Hit description	E-Value	InterProScan	Go annotation
APGWamide	Cerebral peptide 1 precursor	gi|228795|prf||1811269Aprepro-APGWamide	5.62E-64	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0007218 GO:0005179
Schistosomin	Schistosomin-like precursor	gi|159145016|gb|ABW90145.1|schistosomin [*Biomphalaria glabrata*]	6.79E-63	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0005576
Conopressin	Lys-conopressin-prepro-hormone	gi|159454|gb|AAA29289.1| preproconopressin [*Lymnaea stagnalis*]	9.56E-65	Neurohypophyseal hormone; IPR000981; IPR022423; PF00184 (PFAM); SM00003 (SMART); PTHR11681 (PANTHER); PS00264 (PROSITE_PATTERNS); PR00831 (PRINTS); PIRSF001815 (PIRSF); SSF49606 (SUPERFAMILY); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK); TM helix (TMHMM)	GO:0005185 GO:0005576
FFamide	Hypothetical protein CGI_10013719	gi|405950308|gb|EKC18304.1|hypothetical protein CGI_10013719 [*Crassostrea gigas*]	1.67E-08	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	No Go Id
ELH1	Hypothetical protein	gi|516516102|rf|WP017904540.1|hypothetical protein [*Pseudomonas fuscovaginae*]	2.1	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	No Go Id
ELH2	Egg-laying hormone	gi|129299|sp|P06308.2|Ovulation pro-hormone [*Lymnaea stagnalis*]	3E-13	Egg-laying hormone/atrial gland peptide; Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK) IPR003424; PF02323 (PFAM)	GO:0005576 GO:0005179
FMRFamide	FMRFamide partial	gi|1169711|sp|P19802.2| FMRFamide [*Lymnaea stagnalis*]	1.26E-68	FMRFamide-related peptide-like; IPR002544 PF01581 (PFAM)	GO:0007218
GnRH	Gonadotropin-releasing hormone-like protein	gi|325296899|ref|NP_001191482.1|preprogonadotropin-releasing hormone-like protein precursor [*Aplysia californica*]	8.47E-17	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK); TM helix (TMHMM)	No GO id
Whitnin	Whitnin precursor	gi|7108581|gb|AAF36485.1|AF129397_1SPTR pro-hormone [*Lymnaea stagnalis*]	2.95E-47	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	No GO id

Myomodulins1	Myomodulin neuropeptides 1-like	gi|524885775|ref|XP_005099491.1|PREDICTED: myomodulin neuropeptides 1-like [*Aplysia californica*]	1.48E-98	FMRFamide-related peptide-like; Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK) IPR002544; PF01581 (PFAM)	GO:0007218
Myomodulins2	Myomodulin neuropeptide precursor	gi|1359919|emb|CAA65635.1| myomodulin neuropeptide precursor [*Lymnaea stagnalis*]	5.96E-56	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0007218 GO:0005185
Myomodulins3	Myomodulin-like precursor	gi|156567667|gb|ABU82757.1|myomodulin-like precursor [*Tritonia diomedea*]	1.64E-26	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK) ; G3DSA: 3.80.10.10 (GENE3D)	GO:0007218 GO:0005179
LFRFamide	LFRFamide precursor	gi|54873678|gb|AAV41057.1|LFRFa precursor [*Lymnaea stagnalis*]	4.28E-58	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	No GO id
NPY	Neuropeptide y	gi|392872573|gb|AFM85446.1|neuropeptide Y prepropeptide [*Biomphalaria glabrata*]	6.79E-50	Pancreatic hormone-like; IPR001955; IPR020392; SM00309 (SMART); G3DSA: 1.20.5.570 (GENE3D); PF00159 (PFAM); PS00265 (PROSITE_PATTERNS); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0007218 GO:0005179 GO:0005576
Pedal_peptide1	Pedal peptide-1 precursor	gi|524884604|ref|XP_005098921.1|PREDICTED: dentin sialophosphoprotein-like [*Aplysia californica*]	7.48E-94	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0007218 GO:0005179 GO:0005576
Pedal_peptide2	Pedal peptide-1 precursor	gi|31745702|gb|AAP57098.1|pedal peptide prepro-hormone [*Lymnaea stagnalis*]	6.75E-65	Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0007218 GO:0005179 GO:0005576
Pedal_peptide3	Pedal peptide-1 precursor	gi|325296771|ref|NP_001191623.1|pedal peptide 2 precursor [*Aplysia californica*]	4.81E-176	Armadillo-type fold; IPR016024; SSF48371 (SUPERFAMILY); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0005488
Insulin-like peptide 1	Insulin precursor	gi|325296757|ref|NP_001191615.1|insulin precursor [*Aplysia californica*]	4.30981E-23	Insulin-like peptide 1; IPR016179; IPR022352; PR00276 (PRINTS); SM00078 (SMART); PF00049 (PFAM); G3DSA: 1.10.100.10 (GENE3D); PS00262(PROSITE_PATTERNS); SSF56994 (SUPERFAMILY); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0005179 GO:0005576
Insulin-like peptide 2	Insulin-like peptide 5 precursor	gi|524866407|ref|XP_005090025.1|PREDICTED: uncharacterised protein LOC101847921 [*Aplysia californica*]	2.90516E-28	Insulin-like peptide 2; IPR016179; G3DSA: 1.10.100.10 (GENE3D); SSF56994 (SUPERFAMILY); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-	GO:0005179 GO:0005576

				no TM (SIGNALP_EUK)	
Insulin-like peptide 3	Insulin-like peptide 3 precursor	gi|544545|gb|AAB28954.1| insulin-related peptide III [*Lymnaea stagnalis*]	9.55296E-15	Insulin-like peptide 3; IPR016179; IPR022353; G3DSA:1.10.100.10 (GENE3D); PS00262 (PROSITE_PATTERNS); SSF56994 (SUPERFAMILY); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK)	GO:0005179 GO:0005576
Insulin-like peptide 4	Insulin-like peptide 7 precursor	gi|1836031|gb|AAB46831.1| prepro-insulin-related peptide VII [*Lymnaea stagnalis*]	7.39198E-43	Insulin-like peptide 4; IPR016179; IPR022353; SM00078 (SMART); G3DSA: 1.10.100.10 (GENE3D); PF00049 (PFAM); PS00262 (PROSITE_PATTERNS); SSF56994 (SUPERFAMILY); Signal peptide (Phobius); non-cytoplasmic domain; SignalP-no TM (SIGNALP_EUK); TM helix (TMHMM)	GO:0005179 GO:0005576

The biological functions predicted by GO Ids are 0007218: neuropeptide signalling pathway; 0005179: hormone activity; 0005576: extracellular region; 0005185: neurohypophyseal hormone activity; 0005488: binding.

164

5.5.4.1 APGWamide

Bg-APGWamide encodes 8 APGWamide, one copy of SVDDYCQTLDQIMDEYIHKALEVK amide, 4 spacer peptides and one C-terminal peptide (Fig 5.4) identical to previously reported APGWamide in *L. stagnalis*, *A. californica*, *L. gigantea* and *H. asinine* (York *et al.*, 2012; Veenstra, 2000; Smit *et al.*, 1993; Painter *et al.*, 1989). Two APGWamide ESTs were found in the GenBank (EV823649, EV817552) and the 12 tissue RNA-Seq data shows that the gene is expressed in CNS, TRG, APO and KID (Figure 5.4). Multiple sequence alignment shows that the cleavage sites in Bg-APGWamide are conserved with other molluscan homologues and the Bg-APGWamide is most similar to *L. stagnalis* APGWamide (data not shown). Immunohistochemistry of APGWamide demonstrated the bioactivity of this tetrapeptide (Oberdörster and McClellan-Green, 2000; De Lange and Van Minnen, 1998; Smit *et al.*, 1993) and APGWamide seems to control part of the male mating behavior by causing penis eversion (Benjamin, 2008) and its presence in abdominal tissue of imposex female *I. obsoleta* indicates its role in sexual reproduction (Oberdörster and McClellan-Green, 2000). In bivalves such as *Sepia officinalis*, a dipeptidyl aminopeptidase located in the central nervous system has been reported to process APGWamide to dipeptide GWamide that inhibits the contraction of distal oviduct (Henry, Favrel and Boucaud-Camou, 1997). Several related GWamide dipeptides such as TPGWamide, KPGWamide and RPGWamide have been identified from *S. officinalis* and *M. edulis* (Henry, Zatylny and Favrel, 2000; Henry, Favrel and Boucaud-Camou, 1997). In addition to its wide distribution in molluscs, APGWamide shows some structural similarity to crustacean neuropeptide precursor RPCH. RPCH consists of pQLNFSPGWamide (Kornthong *et al.*, 2013; Martínez-Pérez *et al.*, 2002) and the last three amino acids at the C-terminal of the amide (PGW) are identical (Martínez-Pérez *et al.*, 2002b; De Lange and Van Minnen, 1998; Croll and Van Minnen, 1992). Although APGWamide and RPCH share some structural similarities, their functional diversity explains the independent evolution of their receptors (Martínez-Pérez *et al.*, 2002; Henry, Zatylny and Boucaud-Camou, 1999).

Figure 5.4: PreproAPGWamide in B. glabrata.
Figure showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Spacer peptides are indicated in grey; glysine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue.

5.5.4.2 Myomodulins

Myomodulins comprise of a family of peptides that control neuromuscular signaling in the specific muscle of the buccal mass controlling feeding system in *A. californica* (Miller *et al.*, 1993a). They also have well-documented role in modulating various reproductive processes (Greenberg *et al.*, 1997; Van Golen *et al.*, 1996). In *L. stagnalis* myomodulins were shown to modulate neuromusculature of penis (Kellett *et al.*, 1996) but its exact role in reproduction in *H. asinine* is not known (Cummins *et al.*, 2009).

Three structurally related myomodulin-like neuropeptides have been identified from B. glabrata genome similar to those previously found in *L. stagnalis* (Kellett *et al.*, 1996), *A. californica* (Miller *et al.*, 1993b; Miller *et al.*, 1993a), *H. asinine* (Cummins *et al.*, 2009), *L. gigantea* (ESO97371), *C. gigas* (EKC28655) and *Tritonia diomedea* (ABU82757). These neuropeptide precursors are characterised by multiple distinct but structurally related peptides cleaved during post-translational processing (Miller *et al.*, 1993b). All the three Bg-myomodulin genes are differently expressed in the 12 tissues RNA Seq data (Table 5.2). Both Bg-myomodulin-1 and -2 are expressed in 7 different tissues in the snail but Bg-myomodulin 3 is only expressed in the terminal genitalia. The Biomphalaria myomodulin-1precursor molecule is very similar to the FVRIamide neuropeptide precursor in the annelid, Platynereis dumerilii (Conzelmann *et al.*, 2011; Jékely *et al.*, 2008), and F(X)RIamide in L. stagnalis (El Filali *et al.*, 2006). Since FVRI does not appears to be the dominant naming convention used in the model gastropod L. stagnalis as well as in other molluscs so we (Scott et al, personal communication) have followed similar naming convention. Bg-myomodulin-1 is 412aa sequence encoding 21 copies of $X_nF(X)RI$ (where n can be any number between 3 and 31), one copy of TPMEGAGSYEDDSSEEEPIGDD and SGADEVEDE, along with 8 spacer peptides (Figure 5.5). All the peptides have glycine at the C-terminus suggesting that the peptides are amidated after cleavage from the propeptide. Each of the $X_nF(X)RI$ is flanked by monobasic Arg/Lys at C-terminal with another Arg at -4 position similar to FMRFamide

(Schaefer *et al.*, 1985). Bg-myomodulin-2 is the classical myomodulin precursor as described in other mollusc species (Kellett *et al.*, 1996; Miller *et al.*, 1993b; Miller *et al.*, 1993a). 8 Bg-myomodulin ESTs were found that are overlapping partial fragments (ES490100; EX000262; EV819989; EX003480; ES746942; ES751452; EV820260; EX004598) for Bg-myomodulin-2. Based on the 12 tissue transcript data Bg-myomodulin-2 is a 299aa coding sequence consisting of five different myomodulin-like peptides (GLAMLRL, SMKMLRL, QLKMLRL, SLKMLRL and PMNMLRL) and 5 spacer peptides. The Bg-myomodulin-2 gene encodes 4 copies of highly conserved PMNMLRL as compared to 7 copies is H. asinine (York *et al.*, 2012) and 9 to 10 copies in other gastropods (Veenstra, 2010; Kellett *et al.*, 1996; Miller *et al.*, 1993b) (Figure 5.5). These are flanked at the N-terminus by Lys-Arg basic amino acid pairs and have glycine followed by pair of basic amino acids except one of the SMKMLRL that has monobasic Arg- after glycine at the C-terminus. Bg-myomodulin-3 is the shortest transcript of all the three Bg-myomodulins with 164aa and 8 copies of QLKMLKL, one copy of SMEMLSL and one spacer peptide. As in Bg-myomodulin-2, Bg-myomodulin-3 is flanked at the N-terminus by Lys-Arg basic amino acid pairs and has glycine followed by pair of basic amino acids at the C-terminal except for SMEMLSL that has Arg- after glycine at the C-terminus (Figure 5.5). Peptides of this family exhibit considerable homology in their carboxy terminal sequence (MLRL) and unique N-terminus extension (Kellett *et al.*, 1996; Van Golen *et al.*, 1996; Hirata *et al.*, 1989). Bg-myomodulin-3 peptides exhibit an exception as these peptides share a common C-terminal sequence (MLKL), where Arg- is replaced by Lys-and one of the peptide, SMEMLSL, in which Lysine is substituted for Serine. A multiple sequence alignment (Figure 5.6) highlights the conserved peptides and cleavage sites in *B. glabrata, A. californica, L. stagnalis, L. gigantea, H. asinina* and *C. gigas*.

MYOMODULIN1

MFINIILISVTVLLHGVSGDIAEDNLEDDKRASSFVRIGRPSSFVRIGRGDNVEDLETDPNYVDL
EKKASNFVRIGRYPTMSRFIRIGRTPMEGAGSYEDDSSEEEPIGDDGKRASSFVRIGKRKSSF
VRIGKSLAEEDSDEDVDKRASSFVRIGKSPSSFVRIGKAPSSFVRIGKSPSSFVRIGKSPSSFV
RIGKVPSSFVRIGRSIENDLKSELDDIDEEKKASSFVRIGKSPNEELVDSEEEKKRASSFVRIG
KSGLNDQDLFKRVSSFVRIGKSQGEEDKRVSSFVRIGKSGADEVEDEGKRASSFVRIGKSD
TPMDKKASSFVRIGKSSTSPAETSSDSANSAISDEDPINIASRSSAFVRIGKIPSSAFVRIGKNT
NLLTAPSENWKLGFRRGSREGQSSFVRIGK

MYOMODULIN2

MKGTFFLSISVAVTCLLSLGQSEEAKVQSVHSDQTNDAADHPRLKRGLAMLRLGKRDSADID
NLAELLYQMRESEELGAGSEGEDPEEVEELEVPAQHPRVRRSAQSPTENPHHEVLENSSNVE
EFVEHEPRLVDDSYFYVFPDGSEELEPSEGGEDADNFDELEKRSMKMLRLGRSATDDDIFT
NDKRQLKMLRLGKRGDASFDEEFDDEMDKRSLKMLRLGKRPMNMLRLGKRPMNMLRL
GKRPMNMLRLGKRPMNMLRLGKRPADLTDAAEDEKRSMKMLRLGKRSAH*

MYOMODULIN3

MFLTNLVLCLCLSVILLVPGQSEAKVKSPKVLLKSLNVLPDGSDEIIKSREGGEENIDHLEKRS
MEMLSLGRSASDDGILTIEKRQLKMLKLGKRQLKMLKLGKRQLKMLKLGKRQLKMLKL
GKRQLKMLKLGKRQLKMLKLGKRQLKMLKLGKRQLKMLKLGK*

Figure 5.5: Prepromyomodulin in B. glabrata.
Figure showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Spacer peptide are indicated in grey; Glysine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue.
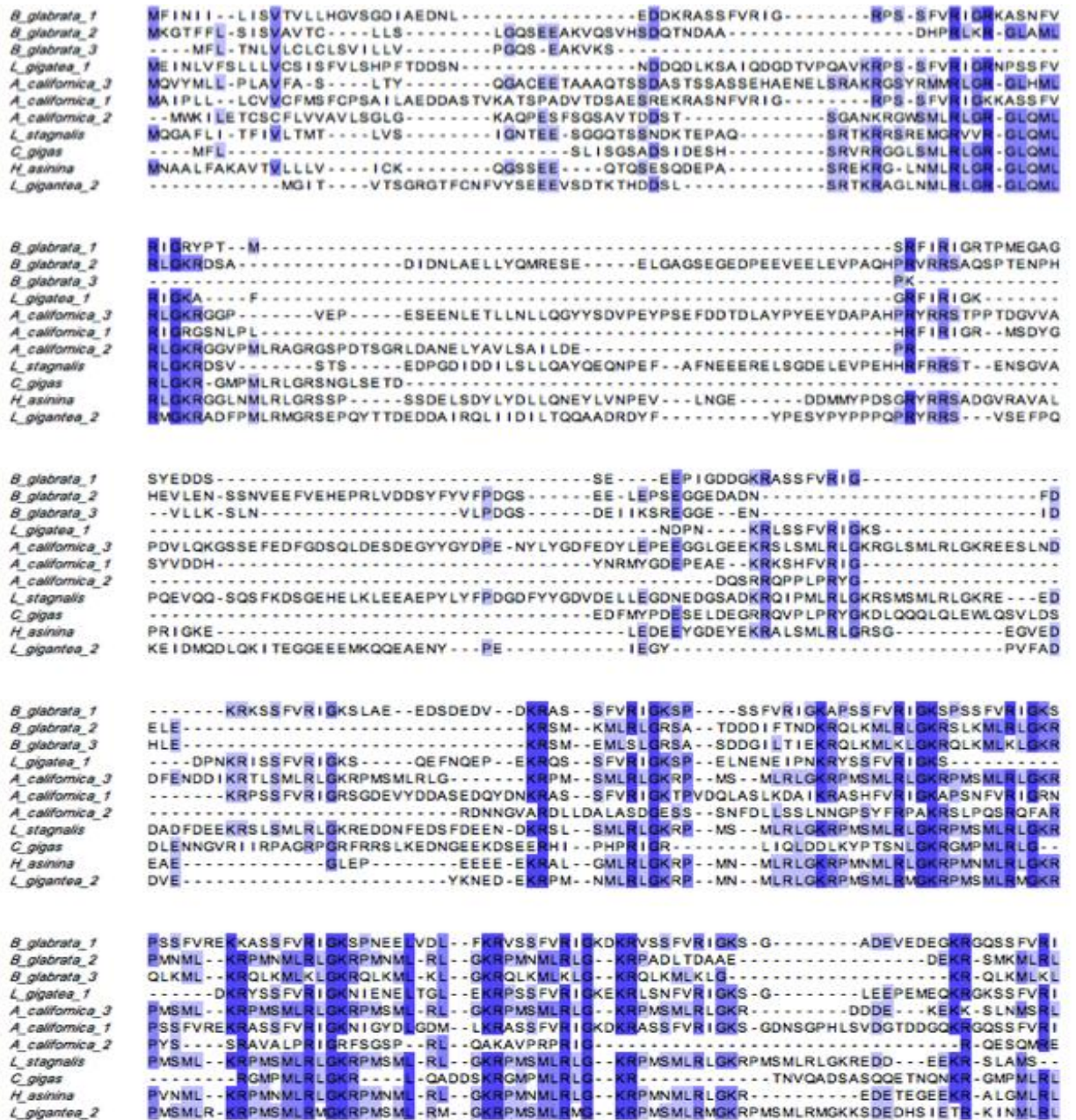
Figure 5.6: Multiple sequence alignment showing cleavage sites in prepromyomodulin in B. glabrata.

The accession numbers used in the alignment were: A. californica_1: XP_005099491; A. californica_2: NP_001191658; A. californica_3: NP_001191423; L. stagnalis : CAA65635;;C. gigas: EKC28655; L. gigantea I: XP_009051543;L. gigantea II:XP_009051970; H. asinine: AFN20272;.

5.5.4.3 Pedal peptides

"Pedal" peptide is a neuropeptide identified from the pedal ganglia from *A. californica* (Willows, Pavlova and Phillips, 1997; Lloyd and Connolly, 1989), which has four different pedal peptide precursors (NP_001191623, NP_001191625, NP_001191585, NP_001191585). One of the physiological roles of pedal peptide in *A. californica is* to control muscular contraction of the foot and thus mediate locomotory behavior (Hall and

Lloyd, 1990). In *Tritonia diomedea*, the pedal peptides mediate the activity of ciliated cells responsible for locomotion and maybe egg transport (Willows, Pavlova and Phillips, 1997). Their function in *L. stagnalis* is still not known (Koene, 2010).

Three similar genes were identified in *B. glabrata* and named pedal peptide 1,2 and 3 based on their sequence similarity with *Aplysia* pedal peptides (Figure 5.7). Bg-pedal peptide-1 pro-hormone contains 6 copies of PFDSISGSHGLSGFA as in *H. lucorum* (AAB51694) and *L. stagnalis* (AAP57098) and 1 copy of SVGTENEVDLD and SMDEESLYGAYKESDDDIY (Figure 5.8). Bg-pedal peptide-2 contains 12 different peptides of which 5 are present in multiple copies and the consensus sequence is SIGS(X)FI. Bg-pedal peptide-3 has 14 different peptides with glycine at their C-terminal followed by a pair of basic amino acids. Of the identified peptides, only two peptides are present in 3 copies (PFDRIGTSSFTSF and PFDRIGNSAFTTF). Overall, Bg-pedal peptide-1 seems to lack C-terminal amide which is similar to *Aplysia* and *Lottia* pedal peptides (Veenstra, 2010). Bg-pedal peptide-2 has one C-terminal amide (PDDDVEEDFENGQLVR) and Bg-pedal peptide-3 has 19 copies of C-terminal amides. This high variability in the peptides makes it difficult to find their homologs but multiple sequence alignment shows that the endoproteolytic cleavage sites are conserved in this gene family in mollusc species, all having dibasic amino acids except for Bg-pedal peptide-3 that has 5 tribasic cleavage sites (Figure 5.7). The multiple sequence alignment of the Bg-pedal peptides with other molluscs reveales conserved cleavage sites (Figure 5.8).

PEDAL PEPTIDE1

*MKMLSTVTIYSAIVLFAAVVSIQA*DEELSSNTNHNEI**KR**SVGTENEVDLD**KR**PFDSISGSHGLSG
FA**KR**PFDSISGSHGLSGFA**KR**PFDSISGSHGLSGFA**KR**PFDSISGSHGLSGFA**KR**PFDSISGSH
GLSGFA**KR**PFDSISGSHGLSGFA**KR**SMDEESLYGAYKESDDDIY**RR**EIEEEDDKLFNVAEEE
LQTVSENTV*

PEDAL PEPTIDE2

*MRTELFLLVILPYSLMSA*EVPDGDI**KKK**SQSQVFQYETRPGDPQGSEMN**KR**SIDSIGSGFI**KRP**
LDSIGSSFI**KK**GIDSIGSAFI**KR**PIDSIGSSFV**KK**NIDSIGSAFI**KR**PLDSIGSSFI**KRR**FHHYAFN
KRPLDSIGSSFI**KR**GIDSIGSGFI**KR**PLDSIGSSFI**KR**GIDSIGSGFI**KR**PLDSIGSSFI**KR**GIDSIGS
GFI**KR**PLDSIGSSFI**KR**GIDSIGSGFI**KR**PLDSIGSSFI**KR**GIDSIGSGFI**KR**PLDSIGSSFI**KR**GIDS
IGSGFI**KR**PIDSIGSSFI**KR**GIDSIGSGFI**KR**PLDSIGSSFI**KR**GIDSIGSGFI**KR**PIDSIGSSFI**KK**SI
DSIGSGFI**KR**PIDSIGSSFV**KK**GIDSIGSGFI**KR**PLDSIGSSFI**KK**SLNSIGSGFI**KR**PIDSIGSSFI**K**
RDYDEEMEDD**KR**ALDSIGSSFI**KR**PDDDVEEDFENGQLV**RGKR**SPTDVVVNEPSDSSYDA
MAAGLHRNDMVHAQLPLWFAEMEIPEYLLEEAQHNQQL**RR***

PEDAL PEPTIDE3

*MPCFIIFFGVCHYVLM*SLHNLSTDDYLDSKESSQENVD**KR**PFDRIGTSSFTSF**GKR**PFDRIGTS
SFTSF**GKR**PFDRIGTSAFTSF**GKR**PFDRIGNSAFASF**GKR**PFDRIGNSAFASF**GKR**PFDRIGNS
AFTSF**GKR**PFDRIGNSAFTSF**GKR**PFDRIGNSAFTTF**GKR**PFDRIGTSVFTSF**GKR**PFDRIGNS
AFTTF**GKR**PFDRIGSSAFTSF**GKR**PFDRIGTSAITSF**GKR**PFDRIGNSAFTTF**GKR**PFDRIGSSA
FTSF**GKR**PFDRIGTSSFTSF**GKR**EDEEGAENFEVTSVNHLEDAGNAKMKEASNKDNSSL**RK**
KRALESNSKTERKEAAIV**KR**EAE**KR**TAEGELTLSMVANNADLSGPIS**KR**FSGTFNRFDNEG
DDNSNEISSDKLDEFLN**KRR**FDSISDSSAFNHF**GKRR**FDRISKNSQFNPF**GKRR**LDRISSISGF
GKF**GKR**KFDSIADGSRFSSF**GKRR**FDRISSGFSHF**GKRR**FDRIDRGSAFSRF**GKR**NYPFNLG
KSSFDRIDKNSAFNAF**GKR**DWEPVMIDENMLEPTYSQDLDLVPENSRSFS**K**PFKVLIDSPWP
EQHINNAIEGLGYHDDEAEKGLTDMDESSDEEAKQLVKYLLNL*

Figure 5.7: Prepropedal peptides in B. glabrata.
Figure showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Sparse peptide are indicated in grey; Glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue.

Figure 5.8: Multiple sequence alignment showing cleavage sites in prepropedal peptides in B. glabrata.

The accession numbers used in the alignment were: A. californica_1: NP_001191585; A. californica_2: NP_001191623; A. californica_3: NP_001191625; L. stagnalis : AAP57098;T. diomedea: ABU82763; L. gigantea : XP_009053476; ; Helix lucorum: AAB51694.

5.5.4.4 FMRFamide

FMRFamide is a member of a large family of structurally related neuropeptides that was first isolated in the clam *Macrocallista nimbosa* and now has been isolated from species spanning all phyla of the animal kingdom (reviewed by (Greenberg and Price, 1992). Previous studies have shown that these peptides are often highly repeated within a protein

precursor and are found with a number of novel peptides (*A. californica* (Schaefer *et al.*, 1985); *D. melanogaster* (Nambu *et al.*, 1988; Schneider and Taghert, 1988); *L. stagnalis* (Saunders *et al.*, 1991); *Cepaea nemoralis* (Rőszer *et al.*, 2006); *C. elegans* (Rosoff *et al.*, 1993). The Bg-FMRFamide (Figure 5.9) consists of one copy of FIRFamide, one copy of a pentapeptide QFYRI, 8 copies of FMRF and 15 copies of (X_n) FLRF where n could be any number between n=0 to n=70, two copies of novel decapeptide (EPSRISYPRY and SDKGPVYMRF) and 18 spacer peptides. Each F(X)RP propeptide is flanked by lysine-arginine or by a single arginine at the amino terminus with arginine/lysine at the carboxy terminus. Each FaRP propeptide has a glycine residue at the carboxy terminus, which is converted to an amide (Loi and Tublitz, 1997; Eipper, Stoffers and Mains, 1992). The tetrapeptide identified in *B. glabrata* have been previously reported in bivalves, cephalopods and gastropods, pentapeptide, QFYRI and heptapeptide (GDPFLRF) are found only in gastropods. The first decapeptide related to FaRPs was reported in *Mytilus edulis* and later in *O. vulgaris* and *L. gigantean* (Zatylny-Gaudin and Favrel, 2014) but the decapeptide identified in *B. glabrata* is different from that previously reported in other mollusc classes (Lopez-Vera et al., 2008).

Figure 5.9: PreproFMRFamide in B. glabrata.
Figure showing signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Sparse peptide are indicated in grey; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue.

In molluscs and other invertebrates FaRPs regulate various physiological activities. Studies have reported that FMRFamide related peptides also act in coordination with other neuropeptides like APGWamide in regulating male copulation behavior (Zatylny-Gaudin and Favrel, 2014; van Golen *et al.*, 1995).

5.5.4.5 LFRFamide

The LFRFamide precursor has been previously identified in *A. californica* (Cropper *et al.*, 1994), *L. stagnalis* (Hoek et al., 2005), *L. gigantea* (Veenstra, 2010), *Sepia officinialis* (Zatylny-Gaudin et al., 2010) and is also present in *B. glabrata* genome. This is 187aa (Figure 5.10) prepro-hormone encoding one copy each of NTLFRF, QGAWFRY, AGTLLRF, GTLLRF, GGSLLRF and TLFRF and four spacer peptides (Figure 5.10). Of these, NTLFRF, GGSLLRF and TLFRF have been found in *L. stagnalis*. GSLLRF is the conserved peptide reported in all molluscs species investigated so far and TLFRF in *L. stagnalis* and *B. glabrata* seem to be the N-terminus truncated form of *Aplysia* hexapeptide STLFRF (Zatylny-Gaudin et al., 2010). Cross-species comparison shows that the cleavage sites are conserved with all the peptides having identical amino acids (RF) at their C-terminus except QGAWFRY where Arg- is substituted with Tyrosine (data not shown). Similar to FaRPs, LFRFamide peptides also regulate main physiological activities (Zatylny-Gaudin and Favrel, 2014). Functional characterization of *C. gigas* NPF like receptor identified LFRFamide as a specific ligand and was reported to be involved in reproduction and feeding (Bigot *et al.*, 2014). In *L. stagnalis*, schistosome parasites induce LFRF gene expression to suppress host metabolism and reproduction (Hoek *et al.*, 2005).


Figure 5.10: PreproLFRFamide in B. glabrata.
Figure showing signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Sparse peptides are indicated in grey; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue.

5.5.4.6 Insulin-like peptide

The *Biomphalaria* genome contains four insulin-like peptides (Figure 5.11) that are homologs of *Lymnaea* molluscan insulin-related peptides (MIP) (Smit *et al.*, 1993; Li *et al.*, 1992). The lymnaea MIPs encode a preprohormone with similar organization of preproinsulin. A typical insulin-related peptide is the member of insulin superfamily consisting of a signal peptide, a B chain, a C chain and an A chain (Smit et al., 1991) (Figure 5.11) and plays a vital role in regulating metabolism, growth, reproduction and

ageing processes (Claeys et al., 2002; Pertseva and Shpakov, 2002). Multiple sequence alignment with insulin-related peptides from mollusc species (*L. stagnalis, A. californica, Tritonia diomedea, Haliotis corrugate*), annelid *(Platynereis dumerilii)* and fly (*D. melanogaster*) exhibited an overall low sequence similarity with insulin of mammals. However, these peptides are principally similar by their structural organization with well-conserved residues essential for the basic insulin core structure (Figure 5.12). The spacing between the cysteine residues was conserved but the cysteine organisation in molluscs is unusual in the sense that they have four intramolecular disulphide bridges rather than the three typically found in the insulin family (Hamano et al., 2005). In *B. glabrata* three cysteine groups were present in the B chain and five cysteine groups were present in the A chain of the insulin-like peptides. Similar cysteine organisation has been reported in other mollusc species (Smit *et al.*, 1998). This is unlike a typical vertebrate insulin family where two cysteine groups are present in the B chain and four cysteine groups are present in the A chain. Interestingly, there was a difference in the position of these extra cysteines in the Bg-insulin-like peptide-2 and the *Lottia* insulin-like peptide-4, which are closest to *Drosophila* insulin-like peptide-7 (Veenstra, 2010). Like the insect peptide (Li *et al.*, 2008), both these genes (Bg-insulin-like peptide-2 and *Lottia* insulin-like peptide-4) have characteristic four amino acid residues between the second and the third Cys- residue of the A chain while classical insulin-like peptides typically have three amino acid residues at the similar position.

Another significant aspect of insulin-related peptides in invertebrates is the polymorphism of the genes encoding multiple isoforms, which is seldom seen in insulin of mammals (Rusakov *et al.*, 2003). Polymorphism of insulin-related peptide in C. gigas was associated with the overall growth of the snail (Cong, Li and Kong, 2013).

INSULIN-LIKE PEPTIDE1

Signal peptide　　　　**B Chain**　　　　**C Chain**

MAYVSRLLLLLLTLAHSTSVTVA EYDHTCNVLSRPHRNGRCGRLLVDTVNLLCNSFSGLLVKRD

TTERVNENLKHILNKKEALSYLTKRETRGSIVCECCYHTCISELLKYCSYDVYNSKHAAAY

A Chain

INSULIN-LIKE PEPTIDE 2

MFSTCVLAYLVLSWHIGTSSASVAEEQAFYTMARNRFATANRDDLLYIWHHDCHRRCFYQIEE

HFEISCAYDPYIARPRRSVPEIANTTTRLTANKQPKKEPSINGNELETRENSARFPFLSKRDAK

SFLRSKRAADQPLSRGKRGIMHECCFTKACSWEEYAEYCHRNPRTTTAMENNCSYS*

INSULIN-LIKE PEPTIDE 3

MLLTTAMLVLVLKVSGVHGVATKVCTSTDRSNSRGLCTGHLYSTIMLACDIFRPVNKRDVMT

DPVNTMDDIVLAKPNALSFLTKKQSSFTVICECCFNQCTVSEMLDYCSLFGRGKK

INSULIN-LIKE PEPTIDE 4

MNAGIDVCLSFTVVCVTLFVGLSLGQDISACDVSKGTHGRGICGARLARAHANLCFLLRNQFP

DVFQRKRSVENSIGDVFSIPLSTLADVDLARDDSNKKTNIPPLMAYTDTNALLSTEPRNFRDL

FLTLRYSLDNSRPKRQSDQQPRDIVCECCINNCTARQLATYC*

Figure 5.11: PreproInsulin-like peptides in B. glabrata.

Figure showing signal peptide highlighted in yellow; the different regions of the gene indicate as B Chain, C Chain and A Chain; likely endoproteolytic cleavage sites are indicated in red; Sparse peptide are indicated in grey; cysteine residues are made dark green; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue.


Figure 5.12: Multiple sequence alignment showing the conserved cysteine organisation in different

5.5.4.7   *Schistosomin*

Schistosomin has been reported in four gastropods (*A. californica, L. stagnalis, P. acuta and B. glabrata*). Previous studies show that schistosomin in *L. stagnalis* inhibits reproduction and had a role in parasitic castration in the snail (Hordijk *et al.*, 1992). But the *B. glabrata* schistosomin differed functionally from other molluscs as it was expressed in all life stages from embryonic development to sexually mature snails and seemed to help in development (Zhang *et al.*, 2009).

The previously identified schistosomin *B. glabrata* (Zhang *et al.*, 2009), is a 97 AA mature peptide with 18aa signal peptide, 8 cysteine residues that form 4 disulphide bonds (De Jong-Brink *et al.*, 1997; Hordijk *et al.*, 1991) and has a peptide (DNYRCPNPGDAFECFESDATARFCVS) with glycine at its C-terminal followed by Lys-Arg pair forming cleavage sites. 4 ESTs for schistosomin were found in GenBank (ES488206; ES485584; DW474143; EU126802) and two amino acid sequences that differ by only two amino acids, one in the signal peptide region and one in the mature peptide (Figure 5.13). Extensive BLAST searches did not reveal significant identity of molluscan schistosomin to other known protein domains, suggesting that schistosomin are unique to the gastropods (Zhang *et al.*, 2009) but its absence in *L. gigantea* narrows this neuropeptide to Euthyneura (Veenstra, 2010). One notable feature is that the tribasic cleavage site found in the schistosomin of *B. glabrata* is RRK, whereas in all other molluscan schistosomin it is RRR.

Figure 5.13: Preproschistosomin in B. glabrata.
Figure showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Sparse peptide are indicated in grey; cysteine residues are made dark green; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue. The two amino acid sequences just differ by two amino acids highlighted in pink.

5.5.4.8 *Gonadotropin-releasing hormone*

Gonadotropin-releasing hormone (GnRH) is present throughout the bilateral animals regulating the production of follicle stimulating hormone and luteinizing hormone thereby regulating gametogenesis and steroidogenesis in vertebrates (Sower *et al.*, 2012; Tsai, 2006). In gastropods, the first suggestion of a molluscan GnRH-like peptide was made in *Hellisoma trivolvis* using heterologous GnRH antibodies (Goldberg *et al.*, 1993) and then in *L. stagnalis* (Young, Chang and Goldberg, 1999), *A. californica* (Tsai, Maldonado and Lunden, 2003) and *L. gigantea* (Veenstra, 2011). A comparative cross species analysis of the identified *B. glabrata* GnRH with *A. californica, L. stagnalis, C. gigas, L. gigantea, Octopus vulgaris, Homo sapiens, Uroteuthis edulis, Mizuhopecten yessoensis, Oncorhynchus masou and Schmidtea mediterranea,* showed that GnRH in all these species have a decapeptide with glycine at its C-terminus and followed by tribasic cleavage site. Overall, the GnRH identified in *B. glabrata* seems to exhibit similar organisation as in other molluscs with the cleavage sites and cysteine organisation being conserved (Figure 5.14). Mollusc GnRH differed from vertebrate GnRH in cysteine organisation with human GnRH_2 and *O. masou* GnRH having none and human GnRH_1 having one cysteine compared to two cysteines in other species in the study. Phylogenetic analysis shows that the vertebrate GnRH clustered separately from the invertebrate GnRH. *B. glabrata* GnRH is most closely related to *A. californic*a with good branch support and is ancestral to other mollusc species.

A
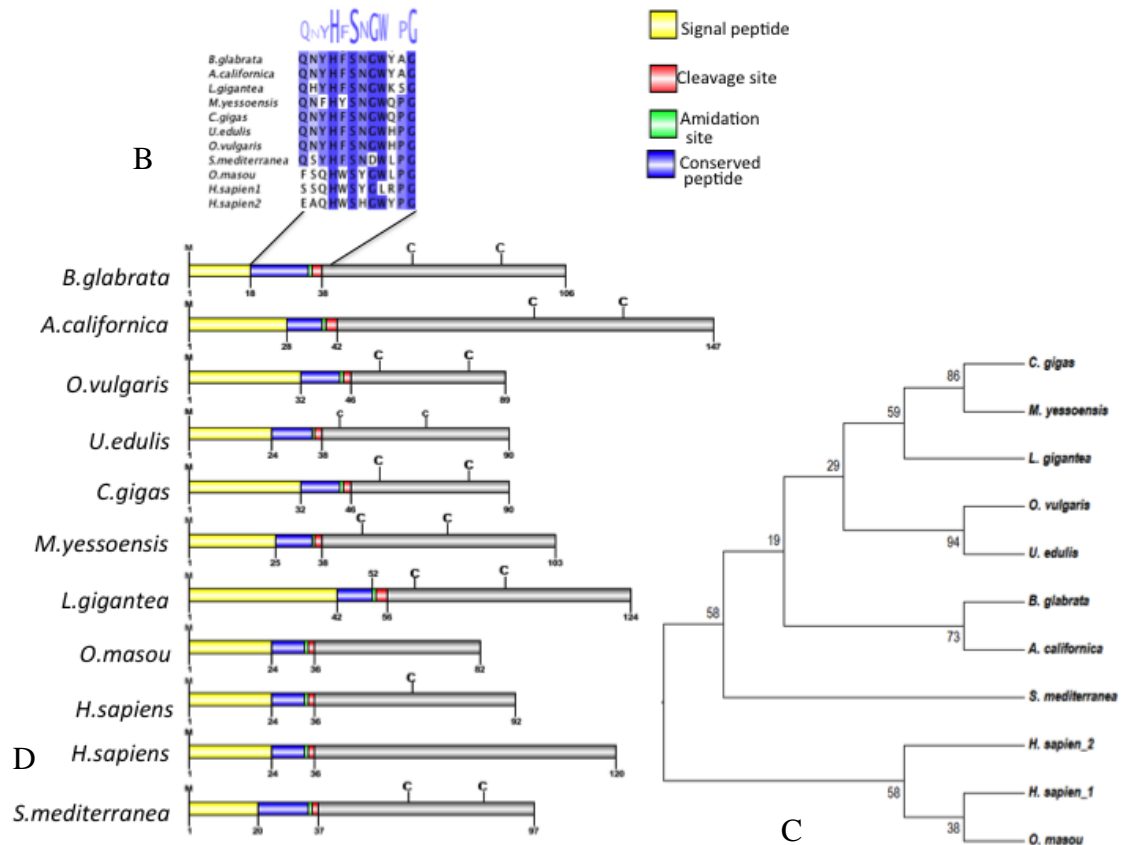
Figure 5.14: Characterization of B. glabrata GnRH.
(**A**) *Predicted GnRHprepro-hormone showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; cysteine residues are made dark green; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue. (**B**)Multiple sequence alignment of B. glabrata neuropeptide with different species where alignments are shaded according to ClustalX colouring scheme, as detailed in Clamp et al. (2004) (Clamp* et al.*, 2004). (**C**) Comparative cross species phylogenetic analyses using Maximum likelihood method. The scale bar indicates the number of amino acid substitutions per site. (**D**) Schematic representation of neuropeptides precursors in different species showing signal peptide, cleavage sites, cysteine organisation, amidation sites and peptides bearing sequence identity to known bioactive peptides (conserved peptides). A. californica: NP_001191482; O. vulgaris:BAB86782; U. edulis:BAH09303; C. gigas: ADZ17180; M. yessoensis:BAH47639; L. gigantea: FC805607; O. masou: P30973; H. sapiens: AAX36880 and AAC02981; S. mediterranea: ADC84439.*

5.5.4.9 Egg-laying hormone

Egg laying hormones have been previously reported in *Aplysia* (McAllister *et al.*, 1983; Chiu *et al.*, 1979), *L. stagnalis* (e.g. (Vreugdenhil *et al.*, 1988; Ebberink *et al.*, 1985) and *L. gigantea* (Veenstra, 2011) and a related hormone has also been identified in the leech *Theromyzon tessulatum* (Salzet *et al.*, 1997). This peptide induces ovulation, egg-mass formation and controls egg laying behavior (Li *et al.*, 1994; Nagle *et al.*, 1988; Vreugdenhil *et al.*, 1988). Two genes (Figure 5.15) coding ELH homologues were identified in *B. glabrata*. Both Bg-ELH precursors consist of a well-conserved region of 33 amino acids, near the carboxy terminus of the pro-hormone followed by glycine at the C-terminal of ELH causing amidation of ELH. This region is flanked on both sides by a

dibasic amino acid motif: Lys-Arg in Bg-ELH-2 but Bg-ELH-1is flanked by monobasic lysine near the carboxy terminal (Figure 5.15). Additionally, the Bg-ELH-2 precursor consists of one more C-terminal peptide and has multiple spacer peptides flanked by mono-, di- and tribasic cleavage sites (Figure 5.15).

The number of cleavage sites (8) is conserved between Bg-ELH-1 and BgELH-2 found in *A. californica* and *L. stagnalis*, but there was a difference in the number of basic amino acids in endoproteolytic cleavage sites. The ELH domain (Pfam02323) identified in Bg-ELH-2 was also found in *Aplysia* and *Lymnaea* but not found in Bg-ELH-1. No homologs for Bg-ELH-1 were found but the core basic organisation of the prepro-hormone is similar to other identified neuropeptides (Figure 5.15).
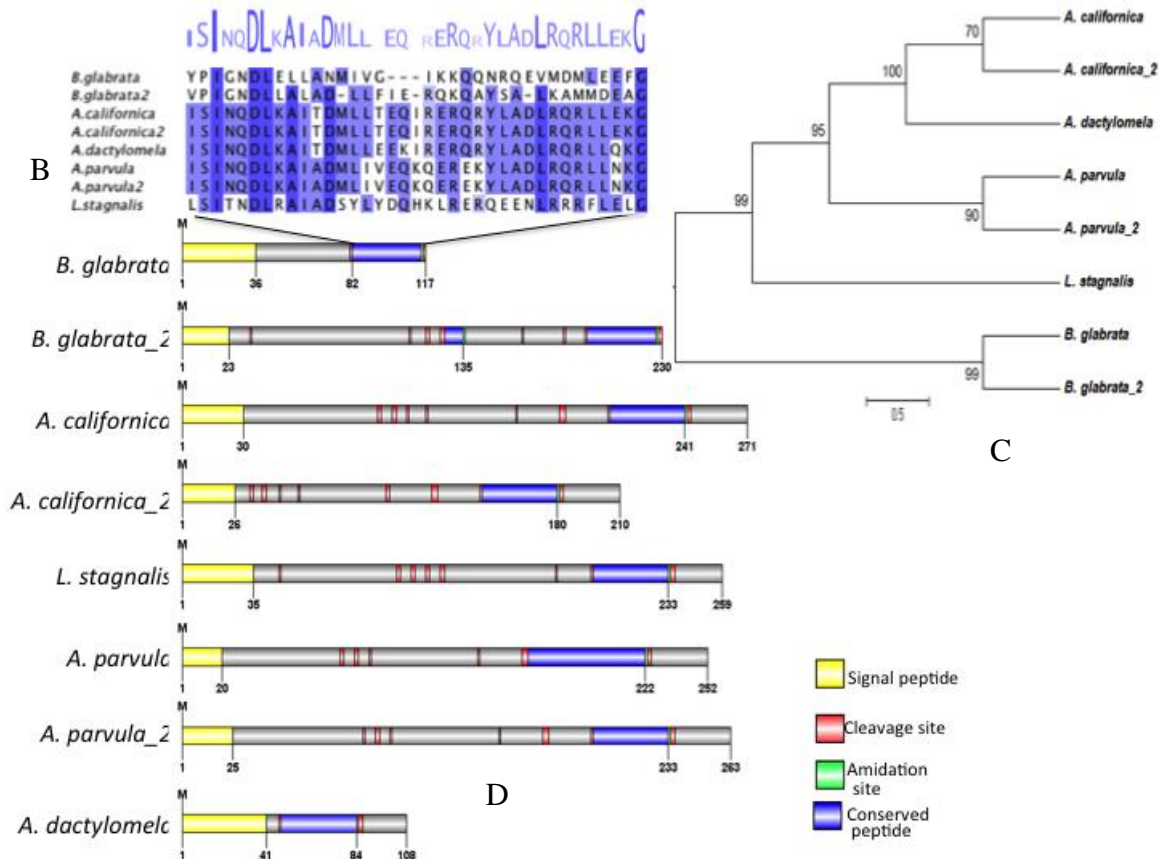
Figure 5.15: Characterization of B. glabrata ELH.
(A)Two PreproELH in B. glabrata showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Sparse peptide are indicated in grey; cysteine residues are made dark green; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue. (B) Multiple sequence alignment of B. glabrata neuropeptide with different species where

5.5.4.10 Conopressin

Conopressin, a member of the vasopressin/oxytocin superfamily (also referred as neurohypophyseal hormones) was first identified in the venom of the fish hunting marine snail (Cruz *et al.*, 1987). These have been previously identified in *L. stagnalis* (Van Kesteren *et al.*, 1995b), *Octopus vulgaris* (Kanda *et al.*, 2003) as well as *Aplysia kurodai* (BAB40371), *A. californica* (e.g. GD239415), *C. gigas* (AM854257and AM853403), *E. scolopes* (DW266122), *L. gigantea* (Veenstra, 2010) and the leech *Erpobdella octoculata* (Salzet *et al.*, 1993). Not only molluscs but also peptides of this family have been reported in annelids, insects and nematodes suggesting an ancient lineage of this family (Dutertre *et al.*, 2008). In *B. glabrata*, one gene encoded a prepro-hormone of 157aa (Figure 5.16) and a propeptide consisting of a nine amino acid peptide (CFIRNCPKG), a neurophysin domain (Pfam00184) and a C-terminal copeptin-homologous extension of the neurophysin domain similar to that previously reported in mollusc conopressin and pro-hormones of the vertebrate vasopressin/oxytocin superfamily (Veenstra, 2010; Salzet *et al.*, 1993). Comparative schematics show a well conserved basic structure and cysteine organisation between different mollusc species. All the identified species had 14 cysteines with similar location in the neuropeptide precursor. The cross-species phylogenetic analyses show that the *Biomphalaria* conopressin is closest to *Lymnaea* in terms of sequence similarity (Figure 5.16).

This is one of the well-characterized nonapeptide that seems to be involved in osmoregulation, neurotransmission and metabolism in molluscs and annelids (Gruber, 2014). In *C. elegans* this neuropeptide plays a role in reproduction. Similar function has also been reported in leeches molluscs and earthworms (Levoye *et al.*, 2005). In *L. stagnalis* conopressin was isolated from penis complex and vas deferens suggesting a possible role of conopressin in controlling male copulatory behavior (Van Kesteren *et al.*, 1995b).
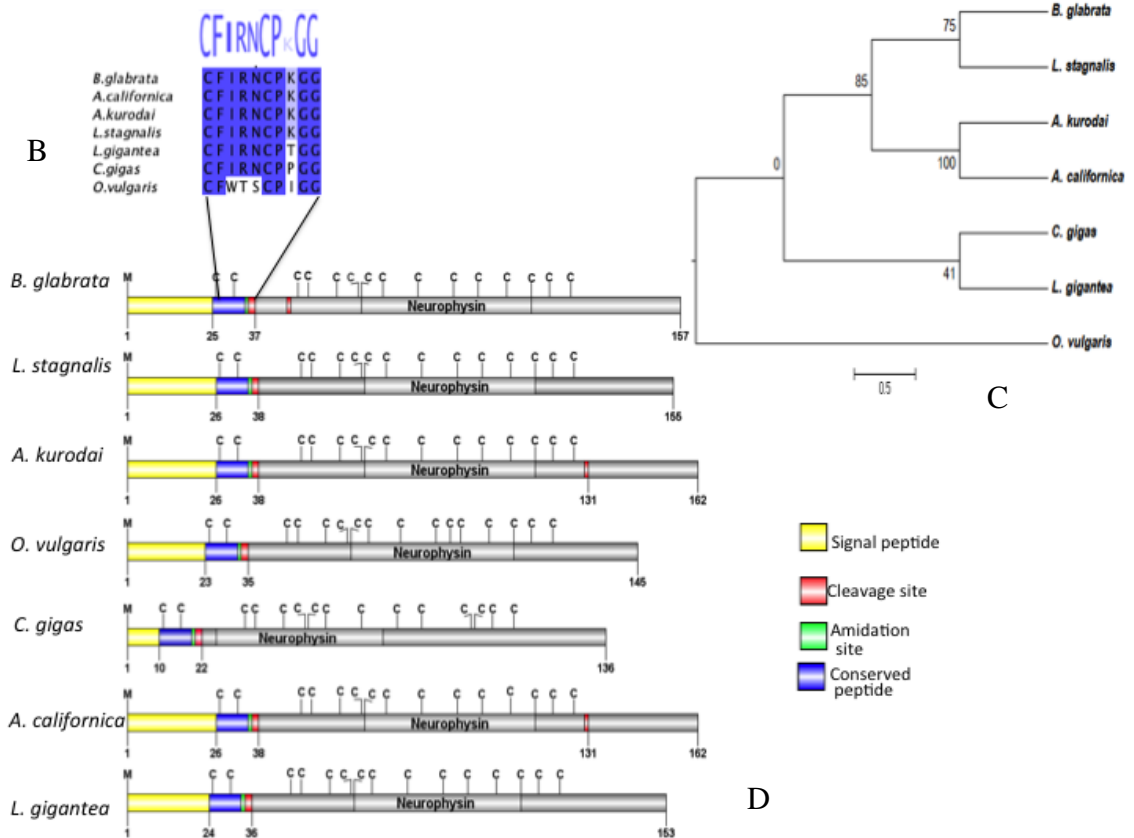
Figure 5.16: Characterization of Bg-Conopressin.

(A) Bg-Conopressin showing the signal peptide highlighted in yellow; likely endoproteolytic cleavage sites are indicated in red; Sparse peptide are indicated in grey; cysteine residues are made dark green; glycine residues predicted to be converted to amide in green and likely bioactive biological peptides in light blue. (B) Multiple sequence alignment of B. glabrata neuropeptide with different species where alignments are shaded according to ClustalX colouring scheme, as detailed in Clamp et al. (2004). (C) Comparative cross species phylogenetic analyses using Maximum likelihood method with MEGA5, under Jones-Taylor-Thorton(JTT) substitution with a gamma distribution of rates between sites (eight categories, parameter alpha, estimated by the program). (D) Schematics representation of neuropeptides precursors in different species showing signal peptide, cleavage sites, cysteine organisation, amidation sites and peptides bearing sequence identity to known bioactive peptides (conserved peptides). L. stagnalis: Q00945; A. kurodai: BAB40371; O. vulgaris: BAC82436; C. gigas: EKC41686; A. californica:NP_001191416; L. gigantea:ESO90380.

Overall, the structural organisation of *B. glabrata* reproductive type neuropeptides follows similar organisation to that found in other species. While homology searches provided a wealth of information when predicting neuropeptides known in other species, bioinformatics approaches and computational resources have proved most useful for identifying novel transcripts based on the structural organisation of these genes (presence of signal peptide; cysteine organisation; cleavage sites; post-translational modifications). Cross-species screening through multiple sequence alignment and phylogenetic tree construction revealed a low prepro-hormone sequence similarity but a high degree of conservation in the basic core structure of the neuropeptides in different species and this may represent evolutionary homology in these conserved peptides. Intron-exon boundaries were also analysed to determine if the conservation/variation was due to genomic structure but preliminary analyses did not reveal any

significant differences between the genomic organisation of each type of neuropeptide. One of the limitations of this whole process was that some of the neuropeptide precursors were split over multiple scaffolds and that could lead to error due to overlapping splice sites. A more complete genome would vastly improve the quality of the gene predictions, as identifying genes would be less error prone if they were lying on single scaffolds. Secondly, putative neuropetides were predicted using *in silico* methods and then confirmed based on homology of the predicted peptides with the previously reported ones but it should be borne in mind that the proposed processing of the predicted neuropeptide are subject to the limitations of the splice-site predicting software.

5.6 Conclusion

Neuropeptides are a diverse class of chemical messengers instrumental in orchestrating complex physiological events which control the basic life processes of metabolism to growth to reproduction and immunity. This initiative has enhanced our knowledge of the neuropeptides present in the snail *B. glabrata* and the possible underlying neuro-endocrine processes that might be regulating reproductive activity. The overall analysis demonstrates conserved molluscan neuropeptide domains and precursor organisation as well as elucidating many previously unrecognized gastropod idiosyncrasies. This data of identified and conserved neuropeptides makes a significant contribution to understanding the evolution of hermaphroditic mode of reproduction in molluscs. This work is a useful addition to previous research as it has the potential to greatly aid the study of molluscan basic reproductive biology. The availability of the complete *B. glabrata* genome opens avenues for further exploring the biochemical pathways through which these neuropeptides exert their function. In this way *B. glabrata* could provide a simple model system to initially investigate the biodynamic interplay between neuropeptides and the nuclear receptors, leading eventually to identification of the mechanisms that might be involved in their interaction with the environment.

6 Chapter 6: General Discussion

6.1   Overview of the thesis

This thesis investigates the similarities between the molluscs and the mammals in their physiological responses to androgenic chemicals and examines if common genes and/or proteins in both molluscs and humans underlie similar mechanisms for male reproductive development. To answer these scientific questions an integrated strategy was developed involving two approaches. 1. *In vivo* exposure of the snails to vertebrate androgens. 2. *In silico* investigation in the snail genome for the presence of an androgen receptor as well as other genes present in the vertebrate androgen signalling pathway. The degree of conservation in snail responses and signalling molecules between snails and humans was expected to give insight into the conservation of functionality in the experimental model chosen.

The overall results showed that the biochemical pathways controlling androgen-dependent reproduction in this mollusc species are not similar to vertebrate androgen signalling pathways. A developmental exposure study using two androgenic chemicals did not reveal any comprehensive pattern of response in either snail growth or reproductive organ development. No significant changes in the weight of gonads (ovotestis) or sex accessory glands (albumen gland and glandular complex) as compared to controls/solvent controls were found in either of the chemical treatments. Therefore, these results were not similar to Hershberger's prepubertal rodent assay where the accessory sex organs weight increase and the weight of testes decreases at the termination of the prepubertal exposure of rodents to androgenic chemicals (Moon *et al.*, 2009) (Chapter 3). *In silico* investigation into two snail genomes (*B. glabrata* and *L. gigantea*) for the presence/absence of androgen receptor identified a repertoire of 40 nuclear receptors in *B. glabrata* and 33 in *L. gigantea*, representing all major subfamilies of the nuclear receptor superfamily also found in vertebrates (Chapter 4). Of the mammalian sex steroid receptor homologs, the whole clade of sex steroid receptor subfamily (NR3C) was missing, although estrogen receptor-like and estrogen related receptor-like homologs were identified. Cross-species phylogenetic analyses with model organisms (fly, worm and human) showed that the test species shared NR homologs with both invertebrates and vertebrates. On an evolutionary scale, molluscs lie somewhere between the insects and vertebrates, exemplified by the presence of both a vertebrate-like THR and an insect-like EcR in molluscs. THR is reported to control metamorphosis and prenatal development of central nervous system in vertebrates (Paris *et al.*, 2010; Bernal, 2007b; Brown, 2005) and EcR plays an important role in *D. melanogaster* metamorphosis/moulting (Li and Bender, 2000). Both these receptors are

functionally similar (analogous) and are present in molluscs. It would be interesting to look whether either of these molecules control development in snails.

Neuropeptide genes controlling reproduction were also identified in our test species and exhaustive "data mining" came up with 17 reproductive type neuropeptide precursors and 4 insulin-like peptides that had 124 putative bioactive peptides in *B. glabrata.* The snail appears to have the molluscan homologues of vertebrate reproductive type neuropeptides (Chapter 5) revealing a similar structural organisation to that found in other species. Moreover, some of the reproductive type neuropeptides like GnRH, Conopressin, NPY and Insulin-like peptides show sequence similarities to their vertebrate counterparts. These findings set forth an avenue of research for the potential use of mollusc in understanding the neuro-endocrine evolution of vertebrate neuropeptides.

Taken together, this knowledge is significant given the intricacies of molluscan endocrine systems due to complexity of reproductive systems and life histories as it gives an insight into the evolution of endocrine control systems and their similarities/differences with the vertebrate endocrine mechanisms. Looking back at the available literature there are still many unanswered questions as previous studies have shown effects of synthetic androgens on molluscs and indicated probable or possible conservation in androgen signalling mechanisms. The following questions are foremost:

- ➢ Is the lack of chemical effect seen in *B. glabrata* my studies a species-specific response (pulmonates/prosobranch snails)?
- ➢ Is it to do with the dose of exposure? Was the dose used lower than that required to effect a response?
- ➢ What is controlling spermatogenesis and reproductive behaviour in *B. glabrata*?

Taking the first of these questions, differences in snail responses to chemicals were reported in the past. Reproductive output was inhibited in *P. antipodarum* at 160ng/l of MT (Duft *et al.*, 2007) whereas *L. stagnalis* when exposed to the same chemical at 100ng/l for 8 weeks showed weak aberration in the albumen gland and no dose dependent effect was seen on the ovotestis (Czech, Weber and Dietrich, 2001). In *M. Cornuarietis,* MT induced imposex and displayed inhibition of spermatogenesis in males at 100ng/l (Albanis *et al.*, 2006; Schulte-Oehlmann *et al.*, 2004). Interestingly it appears that 3 gastropod snails with different reproductive strategies (parthenogenic, hermaphrodite, gonochoric) exhibited different sensitivities towards chemicals (in this example androgen). Similarly, prosobranchs like *N. lapillus* and *Nassarius obsoletus* were affected at very low concentrations of TBT (3 ng/l) (Peña *et al.*, 1988) compared to the pulmonate *L.*

stagnalis, which was only affected at relatively high (230-3200 ng/L) concentrations, causing vacuolization and enlargement of the prostate gland in juvenile snails (Segner *et al.*, 2003). In another study a LOEC of 1 μg/L for embryonic development in *L. stagnalis* was found during TBT exposure (Leung *et al.*, 2007), whereas adult *P. antipodarum* was shown to have a LOEC of 50ng/L for 8-week exposure to TBT (Duft *et al.*, 2007). These results suggest that prosobranchs might be more sensitive to androgenic chemicals than pulmonates.

Not only androgens but also estrogenic chemicals, such as ethinylestradiol (EE2) have been shown to have a greater effect at lower concentrations in prosobranchs (*P. antipodarum*) than in pulmonates. Embryo production was enhanced at 1ng/l in *P. antipodarum* (Jobling *et al.*, 2003) while similar effects were seen in *L. stagnalis* at 500 ng/L (Segner *et al.*, 2003; Anonymous). A one-day exposure of adult *P. antipodarum* to 100ng/L EE2 caused a significant increase in ER-mRNA expression (Stange *et al.*, 2012) as compared to controls while no significant change was reported in adult *Radix balthica* exposed to similar concentration (Börjesson, 2012). Two species *R. balthica* and *Bithynia tentaculata* when exposed to identical concentrations of EE2 responded differently with *B. tentaculata* (a prosobranch) showing a significant inhibition of somatic growth and *R. balthica* (a pulmonate) showing enhanced somatic growth (Hallgren *et al.*, 2012). It seems likely that prosobranchs are more sensitive to EDCs as compared to pulmonates.

Differences in the results of chemical exposures have been also reported between studies of the same species of snail. In contrast to my study, *B. glabrata* snails have previously been reported to be affected by semi-static MT exposure (concentration ranging between 0.01mg/l to 1.0mg/l) in a 4week exposure (Rivero-Wendt *et al.*, 2013). However in this study the statistical comparisons were made between MT dosed and water control snails, rather than between MT dosed and solvent control (0.001% ethanol) exposed snails. The solvent control snails showed that largest average area occupied by sperm. Indeed the authors themselves suggested that the stimulatory effects on sperm production were due to the solvent (ethanol) used, which suggests that MT might not be the predominant factor (Rivero-Wendt *et al.*, 2013). Moreover, the effective exposure concentration reported was much higher than what the snails in our laboratory were exposed to. It is possible that the effects seen in pulmonates at higher reported concentrations than in my experiments were due to an inflammatory effect (Czech, Weber and Dietrich, 2001) rather than due to endocrine disruption; endocrine disruption is generally considered as one of the modes of toxicological action that has potential to impact organisms at very low concentrations in the range of ng/l or lower μg/l concentrations (Oehlmann *et al.*, 2007) depending on their

sensitivities and presence/absence of endogenous hormones. In another study using *L. stagnalis*, Segner et al. (2003) showed that adult survival was significantly reduced by TBT exposure, with an LC_{50} at 21 days of 290 ng Sn/L (nominal concentration), which is in contrast to a 21-d exposure of juvenile *L. stagnalis* to TBT (in the range of 19ng Sn/L to 473ng/L), which did not result in significant mortality (Giusti *et al.*, 2013; Segner *et al.*, 2003). No significant mortality was observed after 49 days and 84 days of exposure of adult *L. stagnalis* to 100 ng Sn/L (Czech, Weber and Dietrich, 2001). Female *N. lapillus* showed a significant effect on penis growth in response to testosterone (Spooner *et al.*, 1991). Conversely, a more recent study found that testosterone had no effect at all on penis growth of *N. lapillus* (Castro *et al.*, 2007)(Castro *et al.*, 2007). One of the possibilities for the contradictory results obtained in the above studies devoted to the same species and the same compound could be that only a single dose (1µg/g body weight) was tested in the latter study even though it was within the range of doses used in the previous study (0.01.0.1,10µg/g body weight). Single dose effects can be interpreted in multiple ways, as it is difficult to explain random variation from 'real' biological effects. Moreover, statistical data analysis is more prone to type I errors or 'false discovery' (Fitts, 2011) as all animals in the same tank are not replicates to each other; infection or disease in one animal could skew the results and the absence of any supporting evidence to characterize these uncertainties in the data affects the reliability of the scientific interpretation. A review comparing sensitivity of three prosobranch snails with different habitats (fresh-, brackish- and marine waters) and different reproductive strategies (parthenogenesis, hermaphrodism and gonochorism) to TBT, reported their sensitivities to be species and sex-dependent (Ketata *et al.*, 2008). Freshwater gastropods (and freshwater species in general) are less sensitive than their marine relatives due to differences in chemical kinetics of TBT in freshwater as compared to marine water (Leung *et al.*, 2007; Leung *et al.*, 2004).

Given their evolutionary divergence in gastropods, interspecies variation within the phylum mollusca is perhaps unsurprising. Morphologically the phylum includes species as different as meiofaunal "worms", now classed as molluscs, and the giant squid (Kocot, 2013). The extreme disparity in morphology among the major lineages within molluscs has led to several conflicting phylogenetic hypotheses (Haszprunar and Wanninger, 2012; Ponder and Lindberg, 2008). Even within gastropods, according to Smith et al. (2011), the different phylogenetic analyses studies completed to date have not been able to cover the diversity of 100,000 living species of this group (Smith *et al.*, 2011).

There are some species of molluscs in which chemicals seem to be having an effect not

only at the individual level, for example, TBT exposure causing imposex affecting molluscs not only at the individual level but also at the population level (Janer *et al.*, 2006a; Oehlmann and Schulte-Oehlmann, 2003; Matthiessen and Gibbs, 1998) and so one might consider all the possible mechanisms/endpoints in which endocrine disrupting chemicals can affect the body's signalling system (Matthiessen, 2008) in these classes of molluscs. Several hypotheses have been tested to investigate the mechanism through which TBT might affect the reproductive physiology in molluscs either through androgen receptor-mediated pathways (Oehlmann *et al.*, 2007; Matthiessen and Gibbs, 1998a) or through retinoids (RXR) (Ciocan *et al.*, 2012; Horiguchi *et al.*, 2007; Nishikawa *et al.*, 2004). Recently Pascoal et al. (2012) suggested the involvement of both retinoids and androgen signalling in inducing imposex in *N. lapillus* (Pascoal *et al.*, 2012). There have been several studies trying to differentiate the mode of action of sex steroids in molluscs (Bannister *et al.*, 2013; Stange *et al.*, 2012; Benstead *et al.*, 2011b; Bannister *et al.*, 2007) but the absence of the whole clade of sex steroid receptors (androgen receptor, progesterone receptor, mineralocorticoids and glucocorticoids) identified in my studies in *B. glabrata* and *L. gigantea* reduces the possibility of finding an androgen receptor in gastropods and possibly in this phylum. If the AR is not involved in androgen signaling in molluscs, the alternative hypothesis, that retinoids are involved, may be valid. Retinoids have been suggested to be involved in inducing imposex in molluscs due to the ability of TBT to form dual ligand heterodimers RXRα:PPARγ (le Maire *et al.*, 2009). The presence of these nuclear receptors in molluscs increases the chances of the retinoid signalling pathway being an active mechanism controlling reproductive development. It has also been shown that nanomolar concentrations of TBT can activate the RXRα:PPARγ heterodimer promoting adipocyte differentiation in amphibian, mouse and humans (le Maire *et al.*, 2009; Iguchi *et al.*, 2007; Grun and Blumberg, 2006; Kanayama *et al.*, 2005) and cause deregulation of the aromatase gene expression (Kotake, 2012; Nakanishi and Nishikawa, 2009; Nakanishi, Nishikawa and Tanaka, 2006). TBT is one of the examples of endocrine disrupting chemicals causing adverse effects in both gastropods and mammals (Iguchi *et al.*, 2008; Iguchi *et al.*, 2007).

Alternatively pathways using membrane bound receptors could be an active mechanism in molluscs as these hormones/synthetic hormones can also trigger biological effects by activating and coordinating a number of transcription factors like cyclic AMP (cAMP) responsive element binding protein (CREB) via MAPK (Mitogen Activated protein Kinase) signaling pathway (Walker, 2010b; Porte *et al.*, 2006; Canesi *et al.*, 2004). Biological and genetic evidence points out that cAMP-dependent pathway is an essential

factor controlling myriad of biological systems like gonadal functions (Khotimchenko YS, 1989), development and homeostasis (Scheide and Dietz, 1986) and metabolism (Fabbri and Capuzzo, 2010). Scarce information on cAMP-dependent signaling was available for invertebrates when it was first reported in *M. edulis* (Stefano *et al.*, 1989) and later in freshwater snail *L. stagnalis* and the marine mollusc *A. californica* (Bailey CH, 2008). Various synthetic estrogenic chemicals (17β-estradiol, DES) were reported to affect haemocyte lysosomal function in *Mytilus* spp by rapid activation of p38 MAPK and STATs (Canesi *et al.*, 2006).

Understanding these pathways and identifying their components genes in molluscs through computational bioinformatics might bring forth interesting evolutionary aspects, guiding further investigation.

6.2 Agreement and disagreement with the previous mollusc exposure studies

The first report of the existence of vertebrate-like steroids in molluscs appeared in the 1950s (Rohlack, 1959; Hagerman, Wellington and Villee, 1957). Since then there have been many accounts in different species and different classes of molluscs showing the presence of these hormones in pg/g wet body weight, tissue distribution, and variation by sex and season. This is a short discussion of scientific studies conducted on different species of snails for endocrine disruptors looking at the experimental details and the interpretation of results. There is no disagreement with the fact that chemicals do have effects on some snail species (reviewed in (Pascoal *et al.*, 2012; Matthiessen, 2008). A closer examination of the experimental details, however, reveals that a few of these studies, looking into the effect of vertebrate-like steroids on growth, fecundity and development of secondary sexual characteristics in molluscs, have not performed any statistical analyses, eg (Sakr, Osman and Abo-Shafey, 1992) (*Theba pisana*); (Bettin, Oehlmann and Stroben, 1996) (*N. lapillus*); (Mori, Muramatsu and Nakamura, 1969) (*C. gigas*); (Tosti *et al.*, 2001) (*O. vulgaris*); (Takeda, 1979) (*Deroceras reticulatum*); (Moss, 1989) (*Mulinia lateralis*)) and so these cannot be taken into account when evaluating the weight of evidence. In some molluscan exposure studies, only one replicate has been used (Oehlmann *et al.*, 2006; Gagne and Blaise, 2005; Janer *et al.*, 2004; Osada *et al.*, 2003; Gagné *et al.*, 2002; Schulte-Oehlmann *et al.*, 2001; Czech, Weber and Dietrich, 2001; Gagné *et al.*, 2001). Such results are open to more than one interpretation because of the absence of any variability arising from examining more than one replicate. Moreover, it is

impossible to determine whether any difference between control and treatment represents a true effect of the treatment or of uncontrolled variables (Scott, 2012b).Any results observed could be due to the treatment, or conversely, could only be specific to the tank that holds the animals. If all animals in any one treatment are held together in the same container, it is impossible to determine intra- and inter- snail differences that might be contributing towards test endpoint variability (Forbes *et al.*, 2007b; Forbes *et al.*, 2007a). Tank effect could also be due to the position of the tank in the exposure room or could be due to an individual in the tank carrying some sort of infection. Replication of *in vivo* animal studies is a vital part of experimental design and should be done keeping in mind endpoint variability (Forbes *et al.*, 2007b; Forbes *et al.*, 2007a).

Molluscs are one of the most specious phyla and different scientific groups have been involved with different species. There are only a handful of papers that are based on studies carried out on any one species, and so there is no standardized protocol for mollusc chemical testing. Experiments are generally designed keeping in mind the life cycle, habitat and behaviour of the test species, even though efforts are being made to develop a standard protocol for mollusc life-cylce tests and partial life cycle tests (Matthiessen, 2008). Another major source of variability for all measured endpoints in *in vivo* mollusc studies is therefore due to variability within and among the test species in their response to the chemical. This complicates interpretation of the impacts of chemicals on molluscs as defined by the current literature. One of the many reasons for differences in the response seen could be difference in competition among animals in a group for food and space between different studies (Forbes *et al.*, 2007b; Forbes *et al.*, 2007a). For example, for TBT, there have been several similar studies (Stange *et al.*, 2012; Pascoal *et al.*, 2012; Lima *et al.*, 2011; Castro *et al.*, 2007; Santos *et al.*, 2006; Bettin, Oehlmann and Stroben, 1996) carried out on one species (*N. lapillus*). There was no replication of the TBT treatment in many of these published experiments with this species (Stange *et al.*, 2012; Castro *et al.*, 2007; Bettin, Oehlmann and Stroben, 1996) and the snails were maintained at different densities in different experiments. Studies have shown that density affects growth and development of snails (Yoshida *et al.*, 2013; Zachar and Neiman, 2013; Mangal TD, Paterson S, Fenton A, 2010). Other species of snails (*Ocenebra erinacea, Ostrea edulis, Urosalpinx cinerea* and *Ocinebrina aciculate*) have been tested with TBT. Female masculinization varied according to species and is a graded mechanism depended on the dose of exposure (Matthiessen, 2008).

6.3 Significance of the results

The results presented in this thesis have been obtained with careful scientific rigour, in order to overcome some of the problems associated with those studies mentioned above. This study showed clearly that reproductive physiology of *B. glabrata* is unaffected by androgenic chemicals at low doses. Mechanistically, these chemicals cannot act through androgen signalling pathways as claimed by many scientific studies previously reported, even though in different species (Pascoal *et al.*, 2012). No androgen receptor has been found in the fully sequenced genome of *B. glabrata, L. gigantea* and *C. gigas* (Vogeler *et al.*, 2014b). Taken together, these studies suggest that snails could be inappropriate models for mammalian steroid hormone function mediated via nuclear receptors. The whole clade of the NR3C subfamily is missing from both gastropod species investigated here and had also been shown to be absent in bivalve *C. gigas* (Vogeler *et al.*, 2014b) and it seems likely that this clade is vertebrate specific. This investigation has, however, raised the possibility of molecular conservation of some of the other metabolic and developmental pathways between snails and humans. Of the NRS identified in *B. glabrata*, 15

Figure 6.1: Venn diagram showing the nuclear receptor homologues in B. glabrata, D. melanogaster and H. sapiens.
B. glabrata has 40NRs of which 15 were homologues to human NRs and 19 were homologues to Drosophila NRs and 5 NRs were identified in all the three species.

NRs are found in both *B. glabrata* and humans while 5 NRs are found in all the three, *Drosophila*, *B. glabrata* and humans Figure 6.1. The mammalian homologs include retinoid receptors (RAR, ROR, RXR and PPAR); steroid hormone receptors like the

estrogen receptor and the estrogen related receptor; the thyroid hormone receptor; the testicular receptor (TR4) and DAX1.

Traditionally, only a few invertebrate animal models, primarily worms and flies, have been relied upon for investigation of many NR-controlled functions but this study reveals the snail's potential as possibly a new model organism for research and treatment of some human diseases and disorders. The identification of a valid invertebrate model might, at least partially, replace rodent use reducing the use of animals in scientific research and chemical testing.

The absence of a molluscan AR and the constitutive expression of the ER suggest that alternative pathways must exist for spermatogenesis/oogenesis in molluscs and other nuclear receptors have been proposed as initiating these pathways for reproductive processes, some of which exist in vertebrates and which may be particularly important in invertebrates. Testicular receptors identified in both the mollusc species have been shown to control spermatogenesis in male mice (Mu *et al.*, 2004) and folliculogenesis (Chen *et al.*, 2008) in female mice. TR knockout mice show degenerated primary spermatocytes and some necrotic tubules (Liu, 2010). Moreover, the significance of RXR signalling in conjunction with PPAR signalling on reproductive physiology in molluscs has also been demonstrated. Although vertebrates may have subsequently evolved more complex processes involving steroid hormones, the presence of orthologous NRs involved in other pathways in invertebrates offers the opportunity to study these underlying/fundamental conserved networks in a simplified system. Another notable finding was the thyroid hormone receptor that plays key roles in growth development and metabolism in vertebrates and makes *B. glabrata* a potential model for thyroid hormone processes. These findings are significant for addressing fundamental biological questions as to the evolution and origin of nuclear receptors as these are the critical components of the hormonal system.

Not only nuclear receptors but also neuropeptides are the main players participating in the snail response to some chemicals. Recently a study reported proteomic analysis of the reproductive organs of *L. stagnalis* exposed to 19.6μg/l of chlordecone (weak binder to the estrogen receptor in vertebrates) causing an increased expression of ovopostatin (inhibits oviposition in *L. stagnalis*) and PIWI gene in reproductive organs (Giusti *et al.*, 2014). PIWI is a key protein in germline stem cell differentiation (Cox, Chao and Lin, 2000) and has been associated with spermatogenesis in vertebrates as well as in invertebrates like hermaphrodites *C. elegans* (Juliano, Wang and Lin, 2011). This work has been culminated in collaboration with Dr Scott Cummins, classifying the secretory peptides employed by

B. glabrata for higher-order neural control of its functioning. This would be an additional knowledge to what is known about neuropeptides in molluscs facilitating the study of invertebrate neuropeptides with the hope of stimulating advances in model organism and parasite-host research.

6.4 Implications for unanswered questions in mollusc endocrinology

This thesis highlights the fact that molluscan endocrinology is not well understood. Even though we have moved a step forward in identifying the components of the neuro-endocrine system in gastropods there are still unanswered questions that need attention. Additional research will be necessary to understand the fundamental mechanisms and pathways that determine key developmental and physiological processes. In turn, this would help us to understand the morphological endpoints and improve tractability of the results in these studies. The fact that molluscs are the second most specious phyla inhabiting different habitats and have undergone different morphological adaptations is the most challenging aspect of research in this field. The evidence that endocrine disrupting chemicals affect molluscs is not very well understood but the phenotypic effects of some of these chemicals is still significant. The "take home message" is that we need more rigorous studies and there is a need for a widely agreed testing protocol for evaluating the strength of evidence associating exposure to chemicals and adverse effects seen in molluscs.

At the time when I started my PhD, only one mollusc genome was available (*L. gigantea*). Since then, however, there has been considerable progress in next generation sequencing, making sequencing of genomes cheaper and quicker than before. Concomitantly, there has been an increase in the number of snails with a fully sequenced genome and a comparative genomic analysis between snails with differing genetic diversities and this would deepen our knowledge related to their fundamental biological differences and similarities. This would facilitate identification of model systems in invertebrates, both to reveal the mechanistic causes of reproductive diseases and to develop test systems. The elucidation of the *B. glabrata* genome creates an invaluable resource to study the functional role of specific genes in these organisms. The conserved nature of the nuclear receptors across the metazoans helps to identify them, but the results should be treated with caution as molecular studies indicate that species-specific differences should be expected in ligand-receptor interaction (Daston et al., 2003). Not only receptor-mediated mechanisms but other mechanisms such as hormone synthesis, transport and metabolism, activation of nuclear receptors and gene methylation are

also equally important in endocrine disruption (Tabb and Blumberg, 2006).

6.5 Further Work

There exist a multitude of potential future avenues of investigation based on the work that I have described here. Nuclear receptors integrate between the central and peripheral nervous systems and communicate not only through established endocrine-hormone signalling but also various other physiological regulatory networks. This comprehensive analysis of the NR gene complement in this non-model species demonstartes the huge diversity of nuclear receptors found in these Lophotrochozoans, and indicates the scope of possible NR-mediated negative impacts of chemicals. Most importantly, the retinoid receptors, PPAR, ER, ERR, Thyroid receptor, COUP-TFII and Testicular orphan receptors identified in the snails represent the so called "ring of physiology" described in the mouse on the basis of gene expression of diverse anatomical systems (http://www.nursa.org/files/NURSA_Phase%20I_II_Final_Report.pdf). Studying the expression profile of NRs in snail tissues at different life stages would offer a simple but powerful way to obtain highly relational information about their ability to coordinate the transcriptional programs and their link with physiological pathways like reproduction, development and metabolism as an individual NR or as a superfamily as done in some recent rodent studies (Xie *et al.*, 2009; Bookout *et al.*, 2006; Yang *et al.*, 2006). Statistical analyses of the hierarchical clustering of NR expression profiling correlating receptors based on their tissue distribution and biological functions, will bring together NR pairs with highest correlation. In particular, this approach could be used to identify hormone metabolic genes, co-activators, corepressors and other unknown factors affecting hormone activity from the signal transduction pathways. Furthermore, this would elucidate common sets of transcription regulators that govern the expression of specific NRs and the physiological processes directed by those networks. There have been some studies using this approach in different vertebrate species (Martinović-Weigelt *et al.*, 2011; Holbeck *et al.*, 2010; Szakács *et al.*, 2006; Fu *et al.*, 2005). This is a large-scale hypothesis free approach so the knowledge of the exact mechanism is not essential. The incorporation of a further expression profiling approach to study chemical-induced changes in these molecules would lead to characterization of basic biological networks that collectively define an organ's response in an organism. Depending on the degree of conservation of the molecules in the network and on the conservation of their pathways, in snails and humans this could lead to the development of new model organisms for research into both host-parasite interaction and cross-species

toxicogenomics studies to understand off-target and mechanism based toxicities for chemical/drug testing purposes (Ulrich, 2003).

Given *B. glabrata's* significance as an intermediate host to *S. mansoni,* and the discovery of homologous receptors in host (*B. glabrata*) and parasite (*S. mansoni*), this thesis adds more avenues to further characterize host-parasite interaction and for possible treatments to break the parasite life-cylce. Recent experimental evidence suggests that parasites can not only evade immune responses actively, but also exploit the hormonal microenvironment within the host to favour their growth, development and reproduction (Escobedo *et al.*, 2005). Extensive analysis of the *S. mansoni* genome for the identification of potential novel drug targets emerged with a family of 21 nuclear receptors that included THR, RXR, COUP TFII and FTZ-F1, that could bind to the host hormones to direct downstream transcription and aid diverse developmental processes (Wu *et al.*, 2007; de Mendonça *et al.*, 2000). Immunization of BALB/c mice (albino, laboratory bred strain of house mice) with recombinant-THR-LBD extract from *S. japonicum* induced partial protection against schistosoma infection (27.52% of worm reduction and 29.5% of reduction in eggs produced in the liver by the parasite) and induced a higher level of T helper cell response in the mice (Qiu *et al.*, 2012). Studies have shown that Thyroid hormone, Insulin and Interlukin 7, all contribute to a host environment favourable to schistosome development in the mouse (de Mendonça *et al.*, 2000). One recent study in the fox tapeworm, *Echinococcus multilocularis* identified 17 NRs which broadly overlapped with the set of NRs expressed in *S. mansoni*. One of the unique receptors identified in the cestode belonged to NR1J subfamily of NRs that demonstrated a role in hormonal host-parasite cross-communication between NR signalling and TGF-β signalling during infection in the intermediate host (Förster *et al.*, 2011). DAF-12 is an evolutionarily conserved nuclear hormone receptor in several parasitic nematodes including *Strongyloides stercoralis, Ancylostoma spp.* and *Necator americanus*, involved in signaling pathway which governs development of the stage 3 infective larvae (iL3) (Wang *et al.*, 2009). Administration of dafachronic acid (a steroid like molecule and a ligand to DAF-12) has been shown to markedly reduce the pathogenic iL3 population in *S. stercoralis*, indicating the potential use of DAF-12 ligands to treat disseminated strongyloidiasis (Rana and Misra-Bhattacharya, 2013; Wang *et al.*, 2009). The *B. glabrata* genome data as well as the parasite data can be further exploited to identify differences between host-parasite metabolic pathways (believed to be regulated by some vertebrate orphan receptors) that make them dependent on the host. Reaction points in those metabolic pathways that produce or consume metabolites (chokepoints) could be

potential drug targets (Webster *et al.*, 2010).

Moreover, aside from the genome, the different genomic tools available for *B. glabrata*, such as RNAi (Jiang, Loker and Zhang, 2006) and the traditional approaches of enhancer and suppressor genetics and transgenics make *B. glabrata* a candidate to demonstrate gene functions that are not based only on homology.

6.6 Conclusion

This thesis has shown that androgen-signalling pathways in the pulmonate gastropod molluscs, *B. glabrata*, and *L. gigantea* are not similar to vertebrates. *In vivo* exposuresto androgens showed no effect on the snail reproductive physiology and the absence of an androgen receptor confirms the non-existence of the androgen signalling pathway in *L. gigantea* and *B. glabrata*. As a result, there is limited commonality with vertebrates with regards to the impacts of androgenic chemicals causing reproductive disorders in vertebrates, although there remains the possibility that other mollusc species might contain an AR considering the number of species this phylum holds. Wildlife is considered sentinel for human health, but care should be taken when extrapolating the impacts of chemicals between different species of organisms. There are a lot of variables that should be considered such as chemical kinetic, chemical dynamics, ADME (adsorption, distribution, metabolism and excretion) of the chemical in different species and genetic machinery to name a few. The elucidation of the *B. glabrata* genome and knowledge of nuclear receptors creates an invaluable resource to study the functional role of these transcription factors stimulating advances in cause-effect 'interactome' research.

Appendix 3.1

Recipe for minimal medium and other medium components for yeast androgenic screen obtained from Brunel University Laboratory.

Ingredients required for yeast screen and the method of their preparation:

Minimal Medium (pH 7.1)

Add 13.61 g KH_2PO_4, 1.98 g $(NH_4)SO_4$, 4.2 g KOH pellets, 0.2 g $MgSO_4$,1 ml $Fe_2(SO_4)_3$ solution, 50 mg L-leucine, 50 mg L-histidine, 50 mg Adenine, 20 mg L-arginine-HCl, 20 mg L-methionine, 30 mg L-tyrosine, 30 mg L-isoleucine, 30 mg L-lysine-HCl, 25 mg L-phenylalanine, 100 mg L-glutamic acid, 150 mg L-valine, and 375 mg L-serine to 1 L double-distilled water. All these ingredients are placed on heated stirrer to dissolve and then it is dispensed in 45 ml aliquots into glass bottles. The glass bottles with the contents are sterilized at 121°C for 10 min, and store at room temperature.

D-(+)-Glucose

Prepare a 20% w/v solution (20 parts by weight of powder to 1000 parts by volume of liquid). Then sterilize in 20 ml aliquots at 121°C for 10 min and stored at room temperature.

L-Aspartic Acid

Make a stock solution of 4 mg/ml; sterilize in 20 ml aliquots at 121°C for 10 min and store at room temperature.

Vitamin Solution

Add 8 mg thiamine, 8 mg pyridoxine, 8 mg pantothenic acid, 40 mg inositol, and 20 ml biotin solution (2 mg/100 ml H_2O) to 180 ml double-distilled water. This is then filtered through a 0.2-µm pore size disposable filter, in a laminar air flow cabinet into sterile glass bottles in 10 ml aliquots and stored at 4°C.

L-Threonine

Prepare a solution of 24 mg/ml; sterilize in 10 ml aliquots at 121°C for 10 min and stored at 4°C.

Copper (II) Sulfate

Prepare a 20mM solution. This is then filtered through a 0.2-µm-pore size filter, in a laminar flow cabinet into sterile glass bottles in 5 ml aliquots and stored at room temperature.

Chlorophenol red-ß-D-galactopyranoside (CPRG)

Make a 10mg/ml stock solution filtered through a 0.2-μm-pore size filter into sterile glass bottles, in a laminar flow cabinet and stored at room temperature.

Method of preparation:

We start by preparing the growth medium by adding 5ml glucose solution, 1.25ml L-aspartic acid solution, 0.5ml vitamin solution, 0.4ml L-threonine solution, and 125μl copper (II) sulfate solution to 45ml minimal medium. The mixture is then transferred to a sterile conical flask bringing the final volume to 50ml. To this mixture 125μl of 10X concentrated yeast stock from cryogenic vial stored at -20°C is added. The mixture is incubated for approximately 24hours at 28°C on an orbital shaker until it turns turbid.

Appendix 4.1

Alignment of all the sequences from human, fly, worm and owl limpet, used in the investigation of nuclear receptors in *B. glabrata*.

```
#NEXUS
[Written Thu Feb 14 10:07:51 GMT 2013 by Mesquite  version 2.75
(build 566) at X-UXIFE125/134.83.70.53]

BEGIN TAXA;
      TITLE Taxa;
      DIMENSIONS NTAX=146;
      TAXLABELS
            HsAR   HsCOUP-TFa   HsCOUP-TFb   HsEAR2   CeDAF_12   Dm_E78c
Dm_E75B  Dm_EGON  HsERa  HsERb  HsERRa  HsERRb  HsERRg  HsGR  HsMR  HsPR
HsGCNF  HsHNF4  HsHNF4g  DmFTZ_F1  HsLXRa  HsLXRb  HsFXR  HSNGFIB  HsNURR1
HsNOR1  HsPNR  HsDAX1  HsSHP  HsPPARa  HsPPARb  HsPPARg  HSPXR  HsCAR  HSRARa
HsRARb  HsRARg  HsREV_ERBa  HsRev_ERBb  HsRORa  HsRORg  HsRORb  HsRXRa
HsRXRb  HsRXRg  HsSF1  HsLRH1  HsTHRa  HsTHRb  HsTLX  HsTR2  HsTR4  HsVDR
LgHNF4  LgNR1J1  Dm_SVP  LgERR  LgFTZ_f1  LgHR3  LgE78c  LgDAX1  LgRev_erb
LgHR4a  LgHR4b  LgTR  LgEcR  LgFAX1a  LgHR38  LgPPAR2  LgRXR  LgNR2E1  LgTHR
LgNR1J3  LgCOUP_TFA  LgHR39  LgRAR  LgROR  LgER  LgTLX  LgFAX1b  LgPNR  LgDSF
LgNR4a  LgE75  LgPPAR1  LgNR2E2  LgNR1J2  BgCOUP_TFa  BgNR4A  BgERR  BgHR39
BgDSF  BgFAX1  BgDAX1  BgROR  BgHR3  BgER  BgRAR  BgHNF4  BgNRU1  BgE75
BgRev_erb  BgNR1J2  BgNR1J4  BgEcR  BgHR4  BgFTZ_F1  BgPPAR2  BgPPAR1  BgNRU2
BgE78c  BgTR  BgNR1J3  BgRXR  BgNRU3  BgHR38  Dmknirps  Dmknirps_li  Dm_ERR
Dm_HR3  Dm_USP  Dm_HR78  Dm_TLL  Dm_PNR  Dm_HR38  Dm_HR39  Dm_HR4  Ce_ODR
Ce_HR23  Ce_NR14  Ce_HR8  Ce_HR1  Ce_HR49  Ce_UNC55  Ce_HR25  Ce_HR91
Dm_HNF4  Dm_HR96  Dm_EcR  Dm_DSF  Dm_FAX1  Ce_HR48  Ce_HR41  Ce_FAX1  Ce_HR6
Ce_NHR85
      ;

END;


BEGIN CHARACTERS;
      TITLE  Character Matrix;
      DIMENSIONS NCHAR=308;
      FORMAT DATATYPE = Protein GAP = - MISSING = ?;
      MATRIX
      HsAR               QKTCLICGDEASGCHYGALTCGSCKVFFKRAAEGKQ--KYL-CAS--
RNDCTIDKFRRKNCPSCRLRKCYEAGMTLGELGERQLVHVVKWAKALPGFRNLHVDDQMAVIQYSWMGL
MVF-AMGWRSFTNVN--------------------SRMLYFAPDLV----------FNEY-R---
-MHK-----SRM-YSQCVRMRHLSQ-EFG-WLQITPQEFLCMKALLLFSI-----IPV----------
DGLKNQKFFDELRMNYIKELDRIIACKRK-NPTSCS-RRFYQLTKLLDSVQPIA
      HsCOUP_TFa         HIECVVCGDKSSGKHYGQFTCEGCKSFFKRSVRRNL--TYT-CRA--
NRNCPIDQHHRNQCQYCRLKKCLKVGMRREELAARLLFSAVEWARNIPFFPDLQITDQVSLLRLTWSEL
FVL-NAAQCSMPLHVA-PLLAAAG--------------------LH-A---------SPMSAD-----
-RVV-------AFMDHIRIFQEQVE-KLK-ALHVDSAEYSCLKAIVLFTS----------------
DACGLSDAAHIESLQEKSQCALEEYVRSQYP----NQP-SRFGKLLLRLPSLRTVS
      HsCOUP_TFb         HIECVVCGDKSSGKHYGQFTCEGCKSFFKRSVRRNL--SYT-CRA--
NRNCPIDQHHRNQCQYCRLKKCLKVGMRREELAARMLFSAVEWARNIPFFPDLQITDQVALLRLTWSEL
FVL-NAAQCSMPLHVA-PLLAAAG--------------------LH-A---------SPMSAD-----
-RVV-------AFMDHIRIFQEQVE-KLK-ALHVDSAEYSCLKAIVLFTS----------------
DACGLSDVAHVESLQEKSQCALEEYVRSQYP----NQP-TRFGKLLLRLPSLRTVS
```

```
      HsEAR2           QVDCVVCGDKSSGKHYGVFTCEGCKSFFKRSIRRNL--SYT-CRS--
NRDCQIDQHHRNQCQYCRLKKCFRVGMRKEELAARLLFSTVEWARHAPFFPELPVADQVALLRLSWSEL
FVL-NAAQAALPLHTA-PLLAAAG--------------------LH-A---------APMAAE-----
-RAV-------AFMDQVRAFQEQVD-KLG-RLQVDSAEYGCLKAIALFTP----------------
DACGLSDPAHVESLQEKAQVALTEYVRAQYP----SQP-QRFGRLLLRLPALRAVP
      CeDAF_12         QKTCRVCGDHATGYNFNVITCESCKAFFRRNALRPK--EFK-CPY--
SEDCEINSVSRRFCQKCRLRKCFTVGMKKENIMDVTMRRFVKVAKGVPAFREVSQEGKFSLLKGGMIEM
LTVRGVTRYDASTNS--------------------------FKTPTIKGQN-----------VSVN-
VDDMFAKLNAN----A-QAQKAKCLEFFG-FFDEEIKKNELAVYLVMLAVFSV-----RSDPP-----
----MNENDVRIVTERHNHFMSLLNRYLESLFG-EQ-A-R-RIFERIPKALGLLNEIA
      Dm_E78c          ---CKVCGDKASGYHYGVTSCEGCKGFFRRSIQKQI--EYR-CLR--
DGKCEIIRLNRNRCQYCRFKKCLAAGMSKDARVTPGVQRIVEFAKRVPGFCDFTQDDQLILIKLGFFEV
WLTHVARLIN--EAT---------------------------LTLDDGAY-----------LTRQ-Q---
LEI------LY-DSDFVNALLNFAN-TLN-AYGLSDTEIGLFSAMVLLAS-----DRAGLS--------
----EPKVIGRARELVAEALRVQILRSRA-GSPQAL-QLMPALEAKIPELRSLG
      Dm_E75B          TVLCRVCGDKASGFHYGVHSCEGCKGFFRRSIQQKI--QYRPCTK--
NQQCSILRINRNRCQYCRLKKCIAVGMSRDQRFAHVIRGVIDFAGMIPGFQLLTQDDKFTLLKAGLFDA
LFVRLICMFDSSINS--------------------------IICLNGQV-----------MRRD-A---
IQNG-----AN-ARFLVDSTFNFAE-RMN-SMNLTDAEIGLFCAIVLITP-----DRPGLR--------
----NLELIEKMYSRLKGCLQYIVAQNRP-D---QP-EFLAKLLETMPDLRTLS
      Dm_EGON         -QTCQVCGEPAAGFHHGAYVCEACKKFFMRSSNAISFTISE-CPN--
GGKCIIDKKNRTKCQACRYRKCLNVGMSLG----------------------------------------
----------------------------------------------------------------------
----------------------------------------------------------------------
----------------------------------------------------------
      HsERa           TRYCAVCNDYASGYHYGVWSCEGCKAFFKRSIQGHN--DYM-CPA--
TNQCTIDKNRRKSCQACRLRKCYEVGMMKGNLADRELVHMINWAKRVPGFVDLTLHDQVHLLECAWLEI
LMI-GLVWRSMEHP-----------------------GK-LLFAPNLL-----------LDRN-Q---
-GKCV----EGM-VEIFDMLLATSS-RFR-MMNLQGEEFVCLKSIILLNS-----GVYTFLS----
STLKSLEEKDHIHRVLDKITDTLIHLM-AKAGLTLQQQH-QRLAQLLLILSHIRHMS
      HsERb           AHFCAVCSDYASGYHYGVWSCEGCKAFFKRSIQGHN--DYI-CPA--
TNQCTIDKNRRKSCQACRLRKCYEVGMVKCKLADKELVHMISWAKKIPGFVELSLFDQVRLLESCWMEV
LMM-GLMWRSIDHP-----------------------GK-LIFAPDLV-----------LDRD-E---
-GKCV----EGI-LEIFDMLLATTS-RFR-ELKLQHKEYLCVKAMILLNS-----SMYPLVT----
ATQDA-DSSRKLAHLLNAVTDALVWVI-AKSGISSQQQS-MRLANLLMLLSHVRHAS
      HsERRa          KRLCLVCGDVASGYHYGVASCEACKAFFKRTIQGSI--EYS-CPA--
SNECEITKRRRKACQACRFTKCLRVGMLKEDLFDREIVVTISWAKSIPGFSSLSLSDQMSVLQSVWMEV
LVL-GVAQRSLPLQ-----------------------DE-LAFAEDLV-----------LDEE-G---
-ARAA----GL--GELGAALLQLVR-RLQ-ALRLEREEYVLLKALALANS-----DSVHIE--------
----DAEAVEQLREALHEALLEYEAGRAGPGGGAER-RRAGRLLLTLPLLRQTA
      HsERRb          KRLCLVCGDIASGYHYGVASCEACKAFFKRTIQGNI--EYS-CPA--
TNECEITKRRRKSCQACRFMKCLKVGMLKEDLADRELVVIIGWAKHIPGFSSLSLGDQMSLLQSAWMEI
LIL-GIVYRSLPYD-----------------------DK-LVYAEDYI-----------MDEE-H---
-SRLA----GL--LELYRAILQLVR-RYK-KLKVEKEEFVTLKALALANS-----DSMYIE--------
----DLEAVQKLQDLLHEALQDYELSQRH----EEP-WRTGKLLLTLPLLRQTA
      HsERRg          KRLCLVCGDIASGYHYGVASCEACKAFFKRTIQGNI--EYS-CPA--
TNECEITKRRRKSCQACRFMKCLKVGMLKEDLADRELVVIIGWAKHIPGFSTLSLADQMSLLQSAWMEI
LIL-GVVYRSLSFE-----------------------DE-LVYADDYI-----------MDED-Q---
-SKLA----GL--LDLNNAILQLVK-KYK-SMKLEKEEFVTLKAIALANS-----DSMHIE--------
----DVEAVQKLQDVLHEALQDYEAGQHM----EDP-RRAGKMLMTLPLLRQTS
      HsGR            PKLCLVCSDEASGCHYGVLTCGSCKVFFKRAVEGQH--NYL-CAG--
RNDCIIDKIRRKNCPACRYRKCLQAGMNLEMLGGRQVIAAVKWAKAIPGFRNLHLDDQMTLLQYSWMFL
MAF-ALGWRSYRQSS-----------------------ANLLCFAPDLI-----------INEQ-R---
-MTL-----PCM-YDQCKHMLYVSS-ELH-RLQVSYEEYLCMKTLLLLSS-----VPK----------
DGLKSQELFDEIRMTYIKELGKAIVKREG-NSSQNW-QRFYQLTKLLDSMHEVV
      HsMR            SKICLVCGDEASGCHYGVVTCGSCKVFFKRAVEGQH--NYL-CAG--
RNDCIIDKIRRKNCPACRLQKCLQAGMNLGARLAGKQMIQVVKWAKVLPGFKNLPLEDQITLIQYSWMCL
SSF-ALSWRSYKHTN-----------------------SQFLYFAPDLV-----------FNEE-K---
-MHQ-----SAM-YELCQGMHQISL-QFV-RLQLTFEEYTIMKVLLLLST-----IPK----------
DGLKSQAAFEEMRTNYIKELRKMVTKCPN-NSGQSW-QRFYQLTKLLDSMHDLV
      HsPR            QKICLICGDEASGCHYGVLTCGSCKVFFKRAMEGQH--NYL-CAG--
```

```
RNDCIVDKIRRKNCPACRLRKCCQAGMVLGQLGERQLLSVVKWSKSLPGFRNLHIDDQITLIQYSWMSL
MVF-GLGWRSYKHVS---------------------GQMLYFAPDLI-----------LNEQ-R---
-MKE-----SSF-YSLCLTMWQIPQ-EFV-KLQVSQEEFLCMKVLLLLNT-----IPL----------
EGLRSQTQFEEMRSSYIRELIKAIGLRQK-GVVSSS-QRFYQLTKLLDNLHDLV
        HsGCNF          QRTCLICGDRATGLHYGIISCEGCKGFFKRSICNKR--VYR-CSR--
DKNCVMSRKQRNRCQYCRLLKCLQMGMNRKRLADELLFRQIAWIKKLPFFCELSIKDYTCLLSSTWQEL
ILL-SSLTVYSKQ---------IFGELADVT--------AKYSPSDEE----------LHRF-S---
-DEG---------MEVIERLIYLYH-KFH-QLKVSNEEYACMKAINFLNQ-----DIRGLT-------
----SASQLEQLNKRYWYICQDFTEYKY---T-HQP-NRFPDLMMCLPEIRYIA
        HsHNF4          SALCAICGDRATGKHYGASSCDGCKGFFRRSVRKNH--MYS-CRF--
SRQCVVDKDKRNQCRYCRLKKCFRAGMKKEESMKEQLLVLVEWAKYIPAFCELPLDDQVALLRAHAGEH
LLL-GATKRSMV-FKD-VLLLGND--------------------Y-IVPRHCP---EL---------
--AE-------MSRVSIRILDELVL-PFQ-ELQIDDNEYAYLKAIIFFDP----------------
DAKGLSDPGKIKRLRSQVQVSLEDYINDRQY----DSR-GRFGELLLLLPTLQSIT
        HsHNF4g         NCLCAICGDRATGKHYGASSCDGCKGFFRRSIRKSH--VYS-CRF--
SRQCVVDKDKRNQCRYCRLKKCFRAGMKKEESMKQQLLVLVEWAKYIPAFCELPLDDQVALLRAHAGEH
LLL-GATKRSMM-YKD-ILLLGNN--------------------Y-VIHRNSC---E----------
--VE-------ISRVANRVLDELVR-PFQ-EIQIDDNEYACLKAIVFFDP----------------
DAKGLSDPVKIKNMRFQVQIGLEDYINDRQY----DSR-GRFGELLLLLPTLQSIT
        DmFTZ_F1        EELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQNKK--VYT-CVA--
ERSCHIDKTQRKRCPYCRFQKCLEVGMKLEKVLDQNLFSQVDWARNTVFFKDLKVDDQMKLLQHSWSDM
LVL-DHLHHRIHNGL-------------------PDETQLNNGQVFN-----------LMSL-G---
-LLG-------V-PQLGDYFNELQN-KLQ-DLKFDMGDYVCMKFLILLNP-----SVRGIVN------
--RK---TVSEGHDNVQAALLDYTLTCY---P-SVN-DKFRGLVNILPEIHAMA
        HsLXRa          NELCSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGA--HYI-CHS--
GGHCPMDTYMRRKCQECRLRKCRQAGMREEELAIVSVQEIVDFAKQLPGFLQLSREDQIALLKTSAIEV
MLLETSRRYNPGSES------------------------ITFLKDFS----------YNRE-D---
FAK------AGLQVEFINPIFEFSR-AMN-ELQLNDAEFALLIAISIFSA-----DRPNVQ-------
----DQLQVERLQHTYVEALHAYVSIHHP-HDR----LMFPRMLMKLVSLRTLS
        HsLXRb          HELCRVCGDKASGFHYNVLSCEGCKGFFRRSVVRGGARRYA-CRG--
GGTCQMDAFMRRKCQQCRLRKCKEAGMREQELAIISVQEIVDFAKQVPGFLQLGREDQIALLKASTIEI
MLLETARRYNHETEC------------------------ITFLKDFT----------YSKD-D---
FHR------AGLQVEFINPIFEFSR-AMR-RLGLDDAEYALLIAINIFSA-----DRPNVQ-------
----EPGRVEALQQPYVEALLSYTRIKRP-QDQ----LRFPRMLMKLVSLRTLS
        HsFXR           DELCVVCGDRASGYHYNALTCEGCKGFFRRSITKNA--VYK-CKN--
GGNCVMDMYMRRKCQECRLRKCKEMGMLAEEMATNHVQVLVEFTKKLPGFQTLDHEDQIALLKGSAVEA
MFLRSAEIFNKKLPS------------------------G----HSD----------LLEE-R---
IRN------SGISDEYITPMFSFYK-SIG-ELKMTQEEYALLTAIVILSP-----DRQYIK-------
----DREAVEKLQEPLLDVLQKLCKIHQP-ENP----QHFACLLGRLTELRTFN
        HSNGFIB         EGRCAVGDNASCQHYGVRTCEGCKGFFKRTVQKNA--KYI-CLA--
NKDCPVDKRRRNRCQFCRFQKCLAVGMVKEDLLSGSLEVIRKWAEKIPGFAELSPADQDLLLESAFLEL
FIL-RLAYRSKPGEG----------------------KLIFCSG-LV----------LHRL-Q---
------C-ARGF-GDWIDSILAFSR-SLH-SLLVDVPAFACLSALVLITD-----R------------
HGLQEPRRVEELQNRIASCLKEHVAAVAG-EP-QPA-SCLSRLLGKLPELRTLC
        HsNURR1         EGLCAVGDNAACQHYGVRTCEGCKGFFKRTVQKNA--KYV-CLA--
NKNCPVDKRRRNRCQYCRFQKCLAVGMVKEDLLTGSMEIIRGWAEKIPGFADLPKADQDLLFESAFLEL
FVL-RLAYRSNPVEG----------------------KLIFCNG-VV----------LHRL-Q---
------C-VRGF-GEWIDSIVEFSS-NLQ-NMNIDISAFSCIAALAMVTE-----R------------
HGLKEPKRVEELQNKIVNCLKDHVTFNNG-GL-NRP-NYLSKLLGKLPELRTLC
        HsNOR1          EGTCAVGDNAACQHYGVRTCEGCKGFFKRTVQKNA--KYV-CLA--
NKNCPVDKRRRNRCQYCRFQKCLSVGMVKENLLTASIDVSRSWAEKIPGFTDLPKEDQTLLIESAFLEL
FVL-RLSIRSNTAED----------------------KFVFCNG-LV----------LHRL-Q---
------C-LRGF-GEWLDSIKDFSL-NLQ-SLNLDIQALACLSALSMITE-----R------------
HGLKEPKRVEELCNKITSSLKDHQSKGQ-----ALE-PTESKVLGALVELRKIC
        HsPNR            SLQCRVCGDSSSGKHYGIYACNGCSGFFKRSVRRRL--IYR-CQV-
GAGMCPVDKAHRNQCQACRLKKCLQAGMNQDETSARLLFMAVKWAKNLPVFSSLPFRDQVILLEEAWSE
LFLL-GAIQWSLPLDS-CPLLAPPE-------------------AS-AAG--GA---QGR-------
---LT-------LASMETRVLQETIS-RFR-ALAVDPTEFACMKALVLFKP----------------
ETRGLKDPEHVEALQDQSQVMLSQHSKAHHP----SQP-VRFGKLLLLLPSLRFIT
        HsDAX1          ------------------------------------------------------
-------------------------EAASAGLLKTLRFVKYLC-
```

202

```
                         FQVLPLDQQLVLVRNCWASLLML-ELAQDRLQFET--VEVSEPS---MLQKILTTR----
RRETGGNEP-----------LPVP-T----LRKV------PSASQVQAIKCFLS-KCW-
SLNISTKEYAYLKGTVLFNP-----DVPGLQ-----------CVKYIQGLQWGTQQILSEHTRMT---
----HQ-GPHDRFIELNSTLFLLR
        HsSHP            --------------------------------------------------
--------------------------REALDVLAKTVAFLRNLS-
FWQLPPQDQRRLLQGCWGPLFLL-GLAQDAVTFEV--AEAPVPS---ILKKILLEE----
PSSSGGSGQ-----------LPD--------RPQ------PSLAAVQWLQCCLE-SFW-
SLELSPKEYACLKGTILFNP-----DVPGLQ-----------AASHIGHLQQEAHWVLCEVLEPWC--
-P-AAQ-GRLTRVLLTASTLKSIP
        HSPPARa         NIECRICGDKASGYHYGVHACEGCKGFFRRTIRLKL--VYDKCDR----
SCKIQKKNRNKCQYCRFHKCLSVGMSHNCTSVETVTELTEFAKAIPGFANLDLNDQVTLLKYGVYEAIF
AMLSSVMNKDGML-----------------------VAYGNGFI-----------TRE-F---
LK----SLRKP-FCDIMEPKFDFAM-KFN-ALELDDSDISLFVAAIICCG-----DRPGLL--------
----NVGHIEKMQEGIVHVLRLHLQSNHP-D---DI-FLFPKLLQKMADLRQLV
        HsPPARb         NMECRVCGDKASGFHYGVHACEGCKGFFRRTIRMKL--EYEKCER----
SCKIQKKNRNKCQYCRFQKCLALGMSHNCTTVETVRELTEFAKSIPSFSSLFLNDQVTLLKYGVHEAIF
AMLASIVNKDGLL-----------------------VANGSGFV-----------TRE-F---
LR----SLRKP-FSDIIEPKFEFAV-KFN-ALELDDSDLALFIAAIILCG-----DRPGLM--------
----NVPRVEAIQDTILRALEFHLQANHP-D---AQ-YLFPKLLQKMADLRQLV
        HsPPARg         AIECRVCGDKASGFHYGVHACEGCKGFFRRTIRLKL--IYDRCDL----
NCRIHKKSRNKCQYCRFQKCLAVGMSHNFRSVEAVQEITEYAKSIPGFVNLDLNDQVTLLKYGVHEIIY
TMLASLMNKDGVL-----------------------ISEGQGFM-----------TRE-F---
LK----SLRKP-FGDFMEPKFEFAV-KFN-ALELDDSDLAIFIAVIILSG-----DRPGLL--------
----NVKPIEDIQDNLLQALELQLKLNHP-E---SS-QLFAKLLQKMTDLRQIV
        HSPXR           PQICRVCGDKATGYHFNVMTCEGCKGFFRRAMKRNA--R-LRCPF--
RGACEITRKTRRQCQACRLRKCLESGMKKEDMSTYMFKGIISFAKVISYFRDLPIEDQISLLKGAAFEL
CQLRFNTVFNAETGT-----------------------WECGRL----------SYCLEDTA----
---G-----G-FQQLLLEPMLKFHY-MLK-KLQLHEEEYVLMQAISLFSP-----DRPGVL--------
----QHRVVDQLQEQFAITLKSYIECNRP-QPA-HR-FLFLKIMAMLTELRSIN
        HsCAR           LRNCVVCGDQATGYHFNALTCEGCKGFFRRTVSKSI--G-PTCPF--
AGSCEVSKTQRRHCPACRLQKCLDAGMRKDDINTFMVLQVIKFTKDLPVFRSLPIEDQISLLKGAAVEI
CHIVLNTTFCLQTQN-----------------------FLCGPL----------RYTIEDGAV---
SPTV-----G-FQVEFLELLFHFHG-TLR-KLQLQEPEYVLLAAMALFSPAPYLTDRPGVT--------
----QRDEIDQLQEEMALTLQSYIKGQQR-RPR-DR-FLYAKLLGLLAELRSIN
        HSRARa          YKPCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNM--VYT-CHR--
DKNCIINKVTRNRCQYCRLQKCFEVGMSKEELSTKCIIKTVEFAKQLPGFTTLTIADQITLLKAACLDI
LILRICTRYTPEQDT-----------------------MTFSDGLT----------LNRT-Q---
MHN------AG-FGPLTDLVFAFAN-QLL-PLEMDDAETGLLSAICLICG-----DRQDLE--------
----QPDRVDMLQEPLLEALKVYVRKRRP-S---RP-HMFPKMLMKITDLRSIS
        HsRARb          YKPCFVCQDKSSGYHYGVSACEGCKGFFRRSIQKNM--IYT-CHR--
DKNCVINKVTRNRCQYCRLQKCFEVGMSKEELATKCIIKIVEFAKRLPGFTGLTIADQITLLKAACLDI
LILRICTRYTPEQDT-----------------------MTFSDGLT----------LNRT-Q---
MHN------AG-FGPLTDLVFTFAN-QLL-PLEMDDTETGLLSAICLICG-----DRQDLE--------
----EPTKVDKLQEPLLEALKIYIRKRRP-S---KP-HMFPKILMKITDLRSIS
        HsRARg          YKPCFVCNDKSSGYHYGVSSCEGCKGFFRRSIQKNM--VYT-CHR--
DKNCIINKVTRNRCQYCRLQKCFEVGMSKEELATKCIIKIVEFAKRLPGFTGLSIADQITLLKAACLDI
LMLRICTRYTPEQDT-----------------------MTFSDGLT----------LNRT-Q---
MHN------AG-FGPLTDLVFAFAG-QLL-PLEMDDTETGLLSAICLICG-----DRMDLE--------
----EPEKVDKLQEPLLEALRLYARRRRP-S---QP-YMFPRMLMKITDLRGIS
        HsREV_ERBa      VLLCKVCGDVASGFHYGVHACEGCKGFFRRSIQQNI--QYKRCLK--
NENCSIVRINRNRCQQCRFKKCLSVGMSRDMSFTPAVREVVEFAKHIPGFRDLSQHDQVTLLKAGTFEV
LMVRFASLFNVKDQT-----------------------VMFLSRTT----------YSLQ-E---
LG------AMG-MGDLLSAMFDFSE-KLN-SLALTEEELGLFTAVVLVSA-----DRSGME--------
----NSASVEQLQETLLRALRALVLKNRP-L---ET-SRFTKLLLKLPDLRTLN
        HsRev_ERBb      VLLCKVCGDVASGFHYGVHACEGCKGFFRRSIQQNI--QYKKCLK--
NENCSIMRMNRNRCQQCRFKKCLSVGMSRDMSFTPAVKEVVEFAKRIPGFRDLSQHDQVNLLKAGTFEV
LMVRFASLFDAKERT-----------------------VTFLSGKK----------YSVD-D---
LH------SMG-AGDLLNSMFEFSE-KLN-ALQLSDEEMSLFTAVVLVSA-----DRSGIE--------
----NVNSVEALQETLIRALRTLIMKNHP-N---EA-SIFTKLLLKLPDLRSLN
        HsRORa          IIPCKICGDKSSGIHYGVITCEGCKGFFRRSQQ--SNATYS-CPR--
```

203

```
         QKNCLIDRTSRNRCQHCRLQKCLAVGMSRDIKITEAIQYVVEFAKRIDGFMELCQNDQIVLLKAGSLEV
         VFIRMCRAFDSQNNT------------------------VYFDGKY-----------ASP-D-V---
         FKSL-----GCED--FISFVFEFGK-SLC-SMHLTEDEIALFSAFVLMSA-----DRSWLQ-------
         ----EKVKIEKLQQKIQLALQHVLQKNHR-E----D-GILTKLICKVSTLRALC
         HsRORg          VIPCKICGDKSSGIHYGVITCEGCKGFFRRSQR--CNAAYS-CTR--
         QQNCPIDRTSRNRCQHCRLQKCLALGMSRDHHLTEAIQYVVEFAKRLSGFMELCQNDQIVLLKAGAMEV
         VLVRMCRAYNADNRT------------------------VFFEGKY-----------GGM-E-L---
         FRAL-----GCSE--LISSIFDFSH-SLS-ALHFSEDEIALYTALVLINA-----HRPGLQ-------
         ----EKRKVEQLQYNLELAFHHHLCKTHR-Q----S-I-LAKLP-PKGKLRSLC
         HsRORb          VIPCKICGDKSSGIHYGVITCEGCKGFFRRSQQ--NNASYS-CPR--
         QRNCLIDRTNRNRCQHCRLQKCLALGMSRDIQITHAIQYVVEFAKRITGFMELCQNDQILLLKSGCLEV
         VLVRMCRAFNPLNNT------------------------VLFEGKY-----------GGM-Q-M---
         FKAL-----GSDD--LVNEAFDFAK-NLC-SLQLTEEEIALFSSAVLISP-----DRAWLI-------
         ----EPRKVQKLQEKIYFALQHVIQKNHL-D----D-ETLAKLIAKIPTITAVC
         HsRXRa          KHICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDL--TYT-CRD--
         NKDCLIDKRQRNRCQYCRYQKCLAMGMKREQAADKQLFTLVEWAKRIPHFSELPLDDQVILLRAGWNEL
         LIA-SFSHRSIAVK--DGILLATG-------------------LH--VHRNS--------AH-----
         -SAG-------VGAIFDRVLTELVS-KMR-DMQMDKTELGCLRAIVLFNP----------------
         DSKGLSNPAEVEALREKVYASLEAYCKHKYP----EQP-GRFAKLLLRLPALRSIG
         HsRXRb          KRLCAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDL--TYS-CRD--
         NKDCTVDKRQRNRCQYCRYQKCLATGMKREQAADKQLFTLVEWAKRIPHFSSLPLDDQVILLRAGWNEL
         LIA-SFSHRSIDVR--DGILLATG-------------------LH--VHRNS--------AH-----
         -SAG-------VGAIFDRVLTELVS-KMR-DMRMDKTELGCLRAIILFNP----------------
         DAKGLSNPSEVEVLREKVYASLETYCKQKYP----EQQ-GRFAKLLLRLPALRSIG
         HsRXRg          KHICAICGDRSSGKHYGVYSCEGCKGFFKRTIRKDL--IYT-CRD--
         NKDCLIDKRQRNRCQYCRYQKCLVMGMKREHAADKQLFTLVEWAKRIPHFSDLTLEDQVILLRAGWNEL
         LIA-SFSHRSVSVQ--DGILLATG-------------------LH--VHRSS--------AH-----
         -SAG-------VGSIFDRVLTELVS-KMK-DMQMDKSELGCLRAIVLFNP----------------
         DAKGLSNPSEVETLREKVYATLEAYTKQKYP----EQP-GRFAKLLLRLPALRSIG
         HsSF1           DELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQNNK--HYT-CTE--
         SQSCKIDKTQRKRCPFCRFQKCLTVGMRLERMADQTFISIVDWARRCMVFKELEVADQMTLLQNCWSEL
         LVF-DHIYRQVQHGK-------------------EGSILLVTGQEVE----------LTTV-A---
         -TQAG----SLL-HSLVLRAQELVL-QLL-ALQLDRQEFVCLKFIILFSL-----DLKFLNN-------
         --HI---LVKDAQEKANAALLDYTLCHY---P-HCG-DKFQQLLLCLVEVRALS
         HsLRH1          EELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQNNK--RYT-CIE--
         NQNCQIDKTQRKRCPYCRFQKCLSVGMKLEKMADQTLFSIVEWARSSIFFRELKVDDQMKLLQNCWSEL
         LIL-DHIYRQVVHGK-------------------EGSIFLVTGQQVD----------YSII-A---
         -SQAG----ATL-NNLMSHAQELVA-KLR-SLQFDQREFVCLKFLVLFSL-----DVKNLEN-------
         --FQ---LVEGVQEQVNAALLDYTMCNY---P-QQT-EKFGQLLLRLPEIRAIS
         HsTHRa          DEQCVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPTYS-CKY--
         DSCCVIDKITRNQCQLCRFKKCIAVGMAMDKIITPAITRVVDFAKKLPMFSELPCEDQIILLKGCCMEI
         MSLRAAVRYDPESDT------------------------LTLSGEMA----------VKRE-Q---
         LKN------GG-LGVVSDAIFELGK-SLS-AFNLDDTEVALLQAVLLMST-----DRSGLL-------
         ----CVDKIEKSQEAYLLAFEHYVNHRKH-N---IP-HFWPKLLMKVTDLRMIG
         HsTHRb          DELCVVCGDKATGYHYRCITCEGCKGFFRRTIQKNLHPSYS-CKY--
         EGKCVIDKVTRNQCQECRFKKCIYVGMATDKIITPAITRVVDFAKKLPMFCELPCEDQIILLKGCCMEI
         MSLRAAVRYDPESET------------------------LTLNGEMA----------VTRG-Q---
         LKN------GG-LGVVSDAIFDLGM-SLS-SFNLDDTEVALLQAVLLMSS-----DRPGLA-------
         ----CVERIEKYQDSFLLAFEHYINYRKH-H---VT-HFWPKLLMKVTDLRMIG
         HsTLX           DIPCKVCGDRSSGKHYGVYACDGCSGFFKRSIRRNR--TYV-CKN--
         QGGCPVDKTHRNQCRACRLKKCLEVNMNKDESAARLLFMSIKWAKSVPAFSTLSLQDQLMLLEDAWREL
         FVL-GIAQWAIPVDAN-TLLAVS--------------------G-MNGDNTD---S----Q-----
         -KLN-------KIISEIQALQEVVA-RFR-QLRLDATEFACLKCIVTFKAVPTHS---------
         GSELRSFRNAAAIAALQDEAQLTLNSYIHTRYP----TQP-CRFGKLLLLLPALRSIS
         HsTR2           FDLCVVCGDKASGRHYGAVTCEGCKGFFKRSIRKNL--VYS-CRG--
         SKDCIINKHHRNRCQYCRLQRCIAFGMKQDESASRLLFLSMHWALSIPSFQALGQENSISLVKAYWNEL
         FTL-GLAQCWQVMNVA-TILATFV-------------------NCLHNSLQQD---KM-STE-----
         -RRK-------LLMEHIFKLQEFCN-SMV-KLCIDGYEYAYLKAIVLFSP----------------
         DHPSLENMELIEKFQEKAYVEFQDYITKTYP----DDT-YRLSRLLLRLPALRLMN
         HsTR4           VEYCVVCGDKASGRHYGAVSCEGCKGFFKRSVRKNL--TYS-CRS--
         NQDCIINKHHRNRCQFCRLKKCLEMGMKMEESASRLLFLSMHWARSIPAFQALGQDCNTSLVRACWNEL
```

```
FTL-GLAQCAQVMSLS-TILAAIV-------------------NHLQNSIQED---KL-SGD-----
-RIK-------QVMEHIWKLQEFCN-SMA-KLDIDGYEYAYLKAIVLFSP---------------
DHPGLTSTSQIEKFQEKAQMELQDYVQKTYS----EDT-YRLARILVRLPALRLMS
        HsVDR          PRICGVCGDRATGFHFNAMTCEGCKGFFRRSMKRKA--L-FTCPF--
NGDCRITKDNRRHCQACRLKRCVDIGMMKEDLVSYSIQKVIGFAKMIPGFRDLTSEDQIVLLKSSAIEV
IMLRSNESFTMDDMS----------------------WTCGNDY---------KYRVSDV-----
-TKA-----G-HSLELIEPLIKFQV-GLK-KLNLHEEEHVLLMAICIVSP-----DRPGVQ-------
----DAALIEAIQDRLSNTLQTYIRCRHP-PPG-SH-LLYAKMIQKLADLRSLN
      LgHNF4          SQYCAICGDRATGKHYGASSCDGCKGFFRRSVRKSH--VYT-CRF--
NRNCIVDKDKRNQCRYCRLRKCFRAGMKKEESMKQQLLILVEWAKYIPCFCELPLDDQVALLRAHAGEH
LVL-GVAKRSLEVNND-VLLLGND-------------------A-VIPRTAS---D----------
---G-------IGIIAGRVLDELVQ-PLR-EIQIDDTEFACIKAIVFFDP---------------
YARGLNDKQRIKMFRYQVHVNLEDYINDRQY----DSR-GRFGEILLLLPALQSIA
        LgNR1J1          ILKCGVCGDKALGYNFDAITCESCKAFFRRNASKNK--G-YNCSF--
EGNCKLDPHTRKFCSGCRLKKCFSIGMKKENLTDAFIRRLIKFGKVLPEFKELSQEDQIHILKGGIMEI
FILRSAMSFDKNSEK------------------------WKYSDTE--------GKKQMDPR-V---
IATA-------LSPDMYSSHMIFVTSLLE-LIKSDRYILLLLFVIELFSP-----DRQLVK--------
----NKESVTRSQEKYSNYLKVYLESIYP-VRK-AQ-TLYPKILMKLIEVRSLG
        Dm_SVP          NIECVVCGDKSSGKHYGQFTCEGCKSFFKRSVRRNL--TYS-CRG--
SRNCPIDQHHRNQCQYCRLKKCLKMGMRREELAARLLFSAVEWAKNIPFFPELQVTDQVALLRLVWSEL
FVL-NASQCSMPLHVA-PLLAAAG-------------------LH-A---------SPMAAD-----
-RVV-------AFMDHIRIFQEQVE-KLK-ALHVDSAEYSCLKAIVLFTT---------------
DACGLSDVTHIESLQEKSQCALEEYCRTQYP----NQP-TRFGKLLLRLPSLRTVS
      LgERR          KRVCLVCGDIASGYHYGVSSCEACKAFFKRTIQGNI--EYS-CPA--
SGECEITKRRRKACQACRFQKCLRVGMLREDLADRELVITISWAKQVPGFTSLNLSDQMNLLQHSWLEV
LCL-NLVYRSCPYS----------------------GY-LCFAEDLK----------VGSD-L---
-VKTY----NVP-PEFDNLTRKLCK-KFT-LLGVSKEEYVLLKAITLCNI-----DVVA-E--------
----VTDTVRCLQDKLQDSLMETIRQRHG----GDM-KRLGQLFLLLPPITHIK
      LgFTZ_f1          EELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQNKK--VYS-CVD--
NRSCHIDKSQRKRCPYCRFQKCLSVGMKLEKLVDQLLFLMVEWARSSIFFKDMRVEDQMKLLQNSWSEI
LFL-DFVYRQVHDCW-------------------GSEIVLENGQKIN-----------LDIL-D---
-KLG-------L-GDLKTRLLELIK-KTK-EVNLDIHEYACLKFIILLNP-----DVSGLEN-------
--RQ---YVEESQERVNASLLEYCLSFY---P-DVK-DKFGQILLRLPEIRLIS
        LgHR3          VIPCKVCGDKSSGVHYGVITCEGCKGFFRRSQS--GPVNYQ-CPR--
NKACVIDRVNRNRCQYCRLQKCLALGMSRDDKITLAVQQIIEFAKMIPGFMDLSQDDQIMLLKAGSFEL
ALLQLCRAYDVASNS-----------------------VIFGNSY-----------IPL-D-A---
FHAL-----DAEEKMVMEKIFEFVK-TVQ-IFKLTEIEMALFSALILISS-----DRPGLK--------
----DTGDIQKLHEKIQAALKLEVGNNHS-D---DS-ETLAKLIVQIPIVRNLS
        LgE78c          VIPCKVCGDRASGYHYGVTSCEGCKGFFRRSIQKQI--EYR-CLR--
DGKCTVIRLNRNRCQYCRFKKCLAVGMSRDKLITPSVHSVVEFAKRVPSFSELSQDDQLILIKAGFFEV
WLTRMSRMFNRNENL------------------------VTFEDGSM-----------IHKD-E---
LSV------VF-SPAFVSSLFDVST-SFN-HLNLNDTEVGLFTGIVLATA-----DRPGLS--------
----DSKGVEQIQDKLIEALKLQVSRNHS-T---EE-NLFASIIMKFPELRTLG
      LgDAX1          -------------------------------------------------
-------------------------RVSDWLVKVIQFAKSVE-
FCNLSHNDKVTLILNSWSKILLL-FMAENNFQFAV--TPLPSSP--------------------
VPT-----------EPVP-A----SPDE------PSMKSVESIQTFIR-KCQ-
HLNIDSEEYSFLRLLVLFNS-----GYIGLD-----------RADIIDQLNSTVQQLLQEHVRLTR--
-P-NDV-MHYSHLLLCLPTLYGTN
        LgRev_erb          SPHCKVCGDESSGFHYGVDSCEGCKGFFRRCITQGM--THK-CAN--
EEKCEITPFTRNSCQFCRLRKCFEVGMSRESAMAPEIENVVKFCKKIPGFKELDQEDQIKLLKRGSFDV
LLARFTLLVNPQTQE-----------------------MMDPTLSL----------RCPRH-V---
VEN------MP-MGMIIKEFYGISE-VLK-PLDLTDGEIAIFSAILIISP-----DRVSPK--------
----NRMTVTKVKGLLHHMFYELLKTNHP-D---PD-NLFLETLSTIPLFQIIN
      LgHR4a          MMICAICQDRATGLHYGIITCEGCKGFFKRTVQNR--IYS-CVA--
SGDCEINRAQRNRCQFCRFKKCLEKGMVLARIGDDIVVKLVHWIRNLPFYSEIPLKLQSQMLTKKWHEL
LLL-IMTAFSACP---------LFTDLFRQNISKIQCYLNKTFGNSLS----------TNQL-N---
-KEI---------GPVIEKVTHIML-KFL-SLQLSRQEFVCLQVILLMSQ-----ND--NIN-------
--EMSGQSIINIQERYMKVLQLYTKTQF---S-SQP-YRFGELLLLLPQIKDAS
      LgHR4b          DKYCGVCGDRALGYNFDAISCESCKAFFRRNAPKLD--YFK-CPY--
EEKCKMDVSNRRFCKRCRLRKCFEIGMRKERIGDDIVVKLVHWIRNLPFYSEIPLKLQSQMLTKKWHEL
```

```
LLL-IMTAFSACP----------LFTDLFRQNISKIQCYLNKTFGNSLS-----------TNQL-N---
-KEI---------GPVIEKVTHIML-KFL-SLQLSRQEFVCLQVILLMSQ-----ND--NIN------
--EMSGQSIINIQERYMKVLQLYTKTQF---S-SQP-YRFGELLLLLPQIKDAS
        LgTR             ---------------------------------------------
--------------------------
ESASRLLFLSMHWTRSINAFQMINPDLQTTLVRACWSELFTL-GLAQCSNVMCL----------
PTLLAAILNHLQ---------------------SNSS-Q-----DKNGDRVKVT-
IEHILRLQDYIT-AMQ-NLQPSTAEYAYLKALVLFQP-----DNPNIIN-----------
KKQVEKLQEKVEKEFVQYSSSVN---T-SAA-ERICQLLLRLQPLKAFN
        LgEcR            EELCRICGDRASGYHYNALSCEGCKGFFRRSITRGA--TYS-CRY--
GGNCEMDWMRRKCQACRLRRCKECGMKEEEMTVLITNLIVEFAKSLPGFTSLSKADQIVLLKAASSEV
MAIRAARCYDLETRC-----------------------IVFANGVP----------CTLE-N---
MKA------TGMG-EYAELTYQFCH-NMA-VMETDNAEYALLTAISIFS------ERPGLS--------
----EKDRVEEVQYRYVSALESYEHIKRH-KGA----CSFGRILSLLPDLRSIS
        LgFAX1a          PVSCEVCGDKSYGKHYGVFCCDGCSCFFKRSIRKNI--NYT-CI--
GKGTCIIDKARRNWCPYCRLKKCLSINMNRNTICQQILMSSLRRLHYNRIFRSLYPADQMLLVQNTWSE
LFLL-SAAYW--PVDI-----------------------CQIIHMA-KS-----------TNDQ-S--
-------CK-----------IKKILA-GCQ-SLKVDNTELPYLETILVLRN----D-----------
NSSKLIENSKVETLLDQAQLALAQYESSL---YP-DSP-TRFGKMLLTLPDLKTLS
        LgHR38           SLLCAVGDNAACQHYGVRTCEGCKGFFKRTVQKNA--KYV-CLA--
DKNCPVDKRRRNRCQFCRFQKCMAVGMVKEDILIQSIDIIKSWAEKVPGFTDLCKEDQELLFQSATLEL
FVL-RAAYRIQANDD----------------------NIIFDNG-LV-----------LHRL-Q---
------C-LGIF-GDWVNSIVDFGL-SLH-RMTLDLSSHCCMAALSMVTL-----R------------
HGLKEPHKMEELQMKIIDCLRDHCTYNSE-AQ-KKP-NFFSKILGKIAELRSLS
        LgPPAR2          DVLCRICGDRASGFHYGVHSCEGCKGFFRRTLKKSL--VYKPCQM--
GSQCKIDMGSRNKCQYCRYQRCLNAGMSQDELILPVMEASVKFAKKLPGFTNLSMADQICLMKNNGFMV
VQLALHILIDK----------------------------------DYIHFVSRRNAFRLNRF-S---
PY-----LCKE-VQNLLERTICVGD-RLH-NLQLTGGEIALFCGILL-TS-----ECSGLQ--------
----NPHQAERVQTDLIDAIRLELKHNHP-K---DR-LMMPKLLLLMPDLCQIV
        LgRXR            KHICAICGDRASGKHYGVYSCEGCKGFFKRTVRKEL--TYA-CRD--
DKNCYIDKRQRNRCQYCRYMKCLAMGMKREHAADKQLFTLVDWAKRIPHFTELPLEDQVIQLRAGWNEL
LIA-GFSHRSIAVK--DGILLSTG-------------------LH--VHRSS--------AR-----
-DAG-------VGTIFDRVLTELVA-KMR-EMKMDKTELGCLRAIVLFNP----------------
DAKGLTAVQEVEQLREKVYASLEEYCKTRYP----DEP-GRFAKLLLRLPALRSIG
        LgNR2E1          DALCLVCGDKASGKHYGVQSCDGCRGFFKRSIRRSL--DYV-CKE--
NGNCTVDVARRNQCQACRFRKCLDVKMNKDETAARLLFMSVKWARNIPSFLQLPFRDQAILLEESWSEL
FIL-SASQWSLPLDAA-TLLMAS--------------------G-VTSENSA---T----E-----
-KSV-------YISSQIRSFQDIIN-RFN-SLRVDATEYACLKALILFKS----------------
ETRGLRDPSQTDTLQDQAQLMMHEYIQSHHP----GNK-ARFGKLLLLLLPSLRAVS
        LgTHR            PEPCVVCGDAATGYHYRCMTCEGCKGFFRRTIQKEL--KYH-CKW--
NQNCEIDKTTRNQCQECRFQKCVRVGMATDEAMTPSIVKVVEFAKGVPGFTQLNVDDQILLLKSCCMEI
MCLRAARRYEPDEQI-----------------------LVLSNGMT----------IHKS-Q---
LNH------VS-VGALVEPIFEFAI-GLA-KLSLDETEVSLLAAVLLMQS-----DRTGLK--------
----DSESVEKLQDAILGAFKRYITCKRP-H---QP-VHWAKVLMKVTDLRTIS
        LgNR1J3          DKICLVCGDRALGYNFNAISCESCKAFFRRNAHKTI--R---GRC--
EGKCDITIESRSFCKRCRQAKCFSVGMRKDNIAEGFVRRIIKVTKHIVNFKKLCKDDQICLLKGAVVDV
MMLRSAVNYDPSTES-----------------------WSLSTMSYD-------SNRISAD-I---
LKCG-----GRETMEMFKSYSKFIKSLMS-TIRGDLLMLKILIVMAIFST-----DRGDLV--------
----DSGFVQKIQEEYAEILQKYIQSRFP-EE---K-TLFAQVVMKLTDLRNIN
        LgCOUP_TFA       HIECVVCGDKSSGKHYGQYTCEGCKSFFKRSVRRNL--NYT-CRG--
NKNCPIDQHHRNQCQYCRLRKCLKAGMRREELAARLLFSAVEWARNVHFFPELHVNDQVELLRISWSEL
FVL-NAAQSSMPCHLS-PLLAAAG-------------------LH-A---------SPMPAD-----
-HVV-------AFMENIRTFQDHVE-KLK-NLHIDTAEYSCLKAIALFSS---------------
DSRSLSDINQIESLQERSQCALEEYVRSQYP----NQP-TRFGKLLLRLPSLRAIN
        LgHR39           GNPCPICGDKISGYHYGIYTCESCKGFFKRTVQNNK--QFV-CHR--
NNDCEVNIFNRKKCPACRFMRCLNMGMRLEQLTELRLYKLVRWARNLPQFGAISTDDQILLLQSCWSDL
LAL-GVCWRTME--E-------------------PSSLRISATKTIS----------MEFA-T---
-SVG-------F-QEVVTRLLSLSE-HFH-KLKMDQVEYVALKVLLLITP-----DVKGVKE-------
--PA---KVSEHQEKLMEALEIYTCSHF---P-RSP-SKFGELLLRLPELARVS
        LgRAR            YRPCVVCLDKSSGYHYGVSSCEGCKGFFRRSVQNM--QYT-CHK--
DKNCPINKVTRNRCQYCRLQKCFSAGMSRELLSSGGIVKIVDFAKKINGFSSLCTSDQITLLKAACLEI
```

```
MILRLSFRYDPDLDS------------------------MVFSNKVC----------INRE-Q---
LEE------GG-FGVLAATIFNFAA-SLK-SMSTDETEFAVLSAVCLVSG-----DRSGLK-------
----ESVKIEQLQEPLLEALKHYVRSRRE-E---QP-HVFAKMLMKLTDLRSIS
        LgROR             VLKCLVCSDKSSGVHYGVLACEGCKGFFRRALQNVGPARKK-CFY--
NKNCEINLQTRNRCQYCRLQKCLALGMSRSLRMHKTIQDVVIFAKKIPGFTGLDQDDQISLIKGGCFEV
ACVVCAPFVDNDSNT------------------------IYLLGNGT----------LITRE-E---
MKIG-----FPLGEHFVELLFNLCN-RFN-AFHLLDSEKALFSALVMISP-----DRPGLK--------
----NRDKVCKLQEMLIQALQFEISDCHP-D---EV-GLFPRLLMSISSLRELG
        LgER              SKVCQVCNDASGFHYGVWSCEGCKAFFKRSIQGPV--DYV-CPA--
TNNCTIDKHRRKSCQACRLRKCYEVGMNKGKLADRELVYLINWAKHVPGYTDLSLSDQVHLIECCWMEL
LLL-NCAFRSMDHE----------------------GKRLVFAPDFH----------LDRQ-L---
-W-NL----TGM-TEILEQVAAVSE-LMV-QYCISKEELLLLQATVLVNA-----EVRRLA--------
----SFSRIQDMRQVILDALVETS-GRF--HA-DNS-RHVPSILLLLTHIRQAG
        LgTLX             DIPCKVCQDHSSGKHYGIYACDGCAGFFKRSIRKSR--HYI-CKG--
QGTCPVDKTHRNQCRACRLKKCEEAGMNKDEAAARLLFMSVKWAKTVPAFLSLPFRDQVILVEEAWREL
FVL-GAAQFQIPLEAG-PLLTSA----------------------G-ISTEQSP---P----E-----
-KVV-------SIMTEIRALQEVIT-KFK-AIQVDPTEYACLKGIVLFKTVFASS----------
TTEMKTLREFHTVATLQDQAQLTLNHYIRTAYP----SQP-FRFGKLLLLLPSLRSIS
        LgFAX1b                 PVSCEVCGDKSYGKHYGVFCCDGCSCFFKRSIRKNI--NYT-
CIEIGKGTCIIDKARRNWCPYCRLKKCLSINMNRNTICQQILMSSLRRLHYNRIFRSLYPADQMLLVQN
TWSELFLL-SAAYW--PVDI----------------------CQIIHMA-KS-----------
TNDQ-S---------CK-----------IKKILA-GCQ-SLKVDNTELPYLETILVLRN-----GKK--
------NNSSKLIENSKVETLLDQAQLALAQYESSL---YP-DSP-TRFGKMLLTLPDLKTLS
        LgPNR               GLVCLVCGDTSSGKHYGILACNGCSGFFKRSVRRKL--IYR-CQA-
GTGMCNVDKAHRNQCQACRLKKCITLGMNKDECSARLLFMAVKWSKNLPSFSNLPFRDQVILLEEAWSE
LFLL-CAIQWSMPMEA-CPLFLLPD-------------------HG-HSTQS---------------
--GKV-------SPYSDMRVLQEVTN-RFK-AVCVDPAEFACLKAIVLFKP----------------
ETRGLKDPSQVENLQDQAQVMLGQHIRAQHP----TRP-FRFGRLLLTLPALRYIP
        LgDSF              DIPCKVCGDRSSGKHYGIYSCDGCSGFFKRSIHKNR--AYT-CKA-
MKGMCPIDKTHRNQCRSCRLKKCFDADMNKDEVMARLLFTIIGWVKQIPAFMILTEADQVI-
LTTSWKDLFLL-GIVQWGLPVDV----------------------NGIHRSI-NF-----------
LG-N-----------VGEMNP-TTELDRLTETIA-RLR-EMALDPTELTCLKAICLFKS-----DG---
---------TNLVDVKTVENLQEQAHLMLTEHIQLQ---YP-RQM-VRFGRLLMLVSKLQTLK
        LgNR4a            GTKCLVCEDRASGFHYSVFSCEGCKGFFKRTVQKAL--TYV-CRE--
QGNCVINKFTRNSCQHCRYNKCVAQGMRKEQYGYLELKFIIEWAKKVPGFQEMSMDDQMCLLKSAFMEL
NVLRLAYRSMNLDRC------------------------IKFAEGVV----------IPTD-V---
AQG------MGWGRELVNATIEFTS-RLR-ELKIDRTEFAILNALVLTFP-----DAIGIQ--------
----DKLQVLGLQTKILDALRRYTSRNYP-T---ET-RRYGKILLRLPALRTV-
        LgE75             TVLCRVCGDKASGFHYGVHACEGCKGFFRRSIQQKI--QYRPCLK--
NQQCNIMRVNRNRCQYCRLKKCISVGMSRDESFTPAIKSVVDFAKGIPGFTLLNQDDQVTLLKAGTFEV
LLVRLSCLFDPDTNT------------------------LMFTGGRL----------FRRQ-P---
SNNTN-NISCG-AGFLLDSMFDFAE-RFN-KLNLGDEELALFSAIVLLSP-----DRPGLR--------
----NLEQVEKIQNALTESLHNIININYK-E---DN-TMFAKLLMKTTDLRTLN
        LgPPAR1            HILCRVCGDKASGFHYGVFSCEGCKGFFRRTIRQNI--TYKPCGN--
MKGCLIMRISRNRCQFCRLNKCVSVGMSHERYLPCIVKCITTFAKEIPQFLDLSLDEQRLLVKGCLLEV
AAIHDSTHILHSEEK------------------------W--EDEKL----------KFRL-D---
TKH-W-NEYGL-LGEVFHHMYDVVG-KLR-KYELTDVELSLLCAMVLFCP-----DREGLK--------
----TVKMLENLEMDLSMALKCQLLLNHG-N---GS-MLFVRCIETVVALRTIT
        LgNR2E2           DALCLVCGDKASGKHYGVQSCDGCRGFFKRSIRRSL--DYV-CKE--
NGNCTVDVARRNQCQACRFRKCLDVKMNKDETAARLLFMSVKWARNIPSFLQLPFRDQAILLEESWSEL
FIL-SASQWSLPLDAA-TLLMAS--------------------G-VTSENSA---T----E-----
-KSV-------YISSQIRSFQDIIN-RFN-SLRVDATEYACLKALILFKSVIFSK-----I----
FSETRGLRDPSQTDTLQDQAQLMMHEYIQSHHP----GNK-ARFGKLLLLLPSLRAVS
        LgNR1J2            -KYCGVCGDRALGYNFDAISCESCKAFFRRNAPKGL--DYFKCPY--
EEKCKMDVSNRRFCKRCRLRKCFEIGMRKENIAEISVRRVIDMSKKIKSFKALSQTDQISLLKGGSIEL
LIIRSVITFDREKQH------------------------FLDPYDN--------GKYAMTAE-Q---
LKMG-------EGSGLFDEHMKFVKSLAM-DLKADETVLILLLVISLFSP-----DRPSIS-------
----DKTYIAKEQDKYALLLQHYLESKYP-VSY-VR-TIYPKLLMKLTDIRNLN
        BgCOUP_TFa         QIECVVCGDKSSGKHYGQFTCEGCKCFFKRSVRRNL--TYT-CRG--
SRNCPIDQHHRNQCQYCRLKKCLKSGMRREELAARLLFSAVEWARNIPFFPDLQVTDQVALLRLGWSEL
FVL-NAAQCSMPLHVA-PLLAAAG-------------------LH-A---------SPMAAE-----
```

207

```
        -RVV-------AFMDHIRIFQEQVE-KLK-ALHVDSAEYSCLKAIALFSS----------------
AFDRDFKNP-MEQPMRNCPAYQKQNVRNLYH----DVT-ALYTYSVLRVTTFDLCE
        BgNR4A          IGACLICGDRASGYHYSVYSCEGCKGFFKRTVQKSL--QYI-CKD--
QSMCFINKFSRNSCQFCRFQKCLNMGMRRDHYGYLELRYIIEWAKKVPGFAELDFSDQIALLKSSFMEL
NVLRLSFRSMDFENK-----------------------IKFCEDLI-----------LPIN-Y---
CES------MGWGKELVDGTIEFAQ-RLK-DINLDLTEFCILNGIVLMYP-----DAHGIE--------
----DKIHIAELQSKILDCLRRH----------------------------
        BgERR          KRLCLVCGDVASGYHYGVSSCEACKAFFKRTIQGNI--EYS-CPA--
SGDCEITKRRRKACQACRFHKCLGVGMLREDLADRELVMTISWAKQVPGFSNLSLVDQMMLLQHSWLEI
LML-SLVFRSCPYK----------------------GN-ICFAEDLQ-----------ISEE-M---
-VETL----SFP-TELDLLLRKLCK-KFT-HLGVTKEEFVLLKAITLCNI-----DVIA-E--------
----LGECVSHLQDQLQDSLSDCVKSVFG----PNT-RRLGQLFLLLPSITHIK
        BgHR39         -------SFSCSGYHYGIFTCESCKGFFKRTVQNKK--NFV-CPK--
ASDCTVNSFNRKKCPACRFKRCLLMGMKLEQLAELRLYKLVRWARNLPQFGAISTDDQILLLQNCWSEL
IAL-SLCWRSID--T--------------------PNQISISSTHHIT-----------AEWA-Q---
-HLG-------F-EDVVIRLLNITD-NFR-KLFMDQFEYVALKVLLLITP-----DVKGLKE-------
--PK---KIREFQDKLTEALLTYTSTHY---P-HLT-DKFGALLLRLPELSRIS
        BgDSF          DIPCKVCGDRSSGKHYGIYSCDGCSGFFKRSIHKNR--AYT-CKA-
LKGRCPVDKTHRNQCRACRLNKCFQADMNKEEVMVRLMFTVVTWVRHIPAFISLCSADQVQLIASAWKE
LFLL-GLVQWGLPVEH----------------------LISREEL-AT-----------S-------
----------SVD-LDDCVTVCQMIA-RIR-ELSLDPTEVTCVKAIALFKP-----DV-----------
-PGLADKSQIQGIQDQAQLMLNRHIQLT---YP-RQM-IRFGRLLMLLSKMQTIH
        BgFAX1         PVSCEVCGDKSYGKHYGVYCCDGCSCFFKRSIRKNI--TYS-CI--
GKGSCLIDKARRNWCPHCRLQKCFAVNMNRNSTLQQILLATFRRAQYNQLFRLLHPHDQIVLLELRWAE
LFIL-TASYW--PVDV----------------------AKLINKL-RD-----------KSGS-K--
-------------------KQGLA-DSE----------------------------------------
----------------------------------------------------
        BgDAX1         ----------------------------------------------
-------------------------RVSDWLVKIVQFAKSIE-
FTALCHNDKVTLILNSWSRLMLL-YMAESNFQFAV--TPIPTVE--------------------
SSN-----------CDSN-P----SQDE------PIMKSVEGVQTFIK-KCQ-
QMSVDSKEYEFLRMLVLFNA-----GYVGLS------------SPSTIDQLNSLVQQLLQQHVRIHR--
-P-KEH-LYTQTLIHALLNIRSIN
        BgROR          NLKCLVCGDKSSGVHYGVLACEGCKGFFRRALQNVGPARKK-CFY--
TKNCDINMQTRNRCQYCRLQKCLQLGMSRAMRLNRTIEDVVCFAKKIPGFADLDQDDQISLIKGGCFEV
ACVVCAPFIDAETNS------------------------IFLLGNSS----------IAMRE-E---
MKLG-----FPLGEHFVELLFNLSI-RLN-AFSLQDTEKALFSALVLISP-----DRPGLK--------
----NREKVSKLQELLIQALQSEISVSHE-D---EG-GLFPRLLMSISSLRELG
        BgHR3          VIPCKVCGDKSSGVHYGVITCEGCKGFFRRSQS--GPVNYQ-CPR--
NKNCVIDRVNRNRCQYCRLQKCLALGMSRDDKITLAVQQIIEFAKMIPGFMELSQDDQIMLLKAGSFEL
ALMQLCRVFDAATNR------------------------VVFGNMF-----------LPL-E-A---
FANL----------------------------------------------------NRPGVK-------
----EVSEIQQLHDRILTALKLEMGVNHA-G---DE-DILKRFVEQAGVLRSLS
        BgER           AKLCQVCNDASGFHYGVWSCEGCKAFFKRSIQGPV--DYI-CPA--
TNTCTIDKHRRKSCQACRLRRCYEVGMNKGKLSDRELVYLINWAKHVPGYTCLTLSDQVHLIECCWMEL
LLL-NCAFRSMEHE----------------------GRTLVFAPDFH-----------LERH-Q---
-W-AL----TGM-GDVLEQVAAVSE-QMV-LYNLNKEELLLLQATVLVNA-----EVRRLD--------
----SFVKIQEMRQVILDVFMEVA-GRH--QGFSSW-RHAPSILLLLTHIRQAG
        BgRAR          YKPCVVCNDKSSGYHYGVSSCEGCKGFFRRSVQKNM--LYT-CHK--
EKTCVINKVTRNRCQYCRLQKCIARGMSKEELSSAGVVLIVDFAKKIPGFLSLSTSDQITLLKAACLEI
MILRISIRYDLETDT----------------------MQFSNGLS-----------LTRE-Q---
LKK------GG-FGPLTSTIFSFAA-SLK-SMSCDETEYALLSSICLISG-----DRSGLK-------
----DTEKIEQMQEPLLEALKHYIRSRRP-D---QK-HSFAKMLMKLTDLRSIS
        BgHNF4         SQECAICRDKSTGKHYGAFSCDGCKGFFRRSVRRNH--VYA-CRF--
SRNCVVDKDKRNQCRYCRLRKCFRAGMKKEESMKQQLLGLVEWAKLIPCFSELPLDDQVALLRAHAGEH
LIL-GVAKRSIP-LSE-VLLLSNE--------------------S-VITRQNI---D----------
--PD-------VSHIASRILDEIVS-PMR-DIQIDDSEFSCLKTLVFFDP---------------
IAPGLTNIAKVKQFRYQVQVNLEDYINDRQY----ESR-GRFGEILLLLPPLQSIC
        BgNRU1         LPPCRVCGNKASGLHFGVNTCEACNEFFRRSLKRGA--SYY-CTK--
NRECQVYGK-
KNACSYCRYRRCVEMGMSRDEMTEKWLRNYINYAKNVPGFKELALSDQASLVRGTWFEFWFL-GAYR-
```

```
GYNSDL-----------------------RVVYYPNRTF----------HEEE-I----IKVF----
--G-KEYTDFSFSLAD-RMA-SLRVTPEEMVLIKTVCLTFP-----DRTPLQD------------
RDAVESMHWQMVS--L-HTLEKNR----PGDSTVFPMVVNKLVELR---
        BgE75            TVLCRVCGDRASGFHYGVHACEGCKGFFRRSIQQKI--QYRPCLR--
NQQCNIQRVNRNRCQYCRLKKCISVGMSRDESFTPAIKSVVDFAKGIPGFVLLNQDDQVTLLKAGTFEV
LLVRLACLFDPVSNT------------------------MMFTGGKL-----------FRRQ-T---
STVT-----TN-VGFLLDSMFDFAE-RFN-RMCLNDDEIAVFSAVVLMSP-----DRPGLR--------
----NVEQVERIQNKLTEALQNIINTNHK-E---DS-TLFAKLLMKTTDLRTLN
        BgRev_erb        SPQCKVCGDESSGFHYGVDSCEGCKGFFRRCITQGM--THK-CAN--
DEKCEITPFTRNSCQYCRLKKCFEVGMSRESSMIPEITKVVKFCKKLPGFSEVDQEDQIKLIKQGSFEV
MVARFSLLVDEINEE------------------------MLDPQLKM---------KSSRK-V---
VQQ------MP-MGPFLDEFFEVAH-RFN-PLLLSDGEIGLFTSILLICP-----ERQGLS--------
----NTKAVSAIQMLYLQALYLLLQQHHK-D---HE-HLFLELISLIPLFRRIN
        BgNR1J2          DKYCGVCGDRALGYNFDAISCESCKAFFRRNAPKGL--EYFKCPY--
EEKCKMDVSNRRFCKRCRLRKCFEIGMRKENIAEISVRRVIDMSKKIKSFKALSQADQISLLKGGSIEL
LIIRSVITFDVDKMQ------------------------FLDPYDK--------EKYAMTTD-Q---
LKMG-------QGSGLFDEHMRFVKSLAI-DLQADETVLILLLMIALYSP-----DRPQIT--------
----DKHYVAEEQERYALLLLRYLESKHP-PHV-IR-WMYPKLLMKLVDIRNLN
        BgNR1J4          -KLCGVCGDKALGYNFNAITCESCKAFFRRNALKDK--TMV-CLF--
ENKCVVDIRTRRFCPACRIKKCFHIGMNRDNHSEICVRKMILFVKYLADFKLLSQDDQIAALKASVMKT
VLLRSALFYIIEKDA------------------------WITPKGEIPS-----------SILK-
KSTGFVSV-----------HNNHVYYCR-RVKSMAQDDLHIYSLLQALILFDP-----AGLNL-----
----TSRE---YVSTIQDKYLILLKHYMESKFS-YD-YSR-EYYVSALDRLSDLRVLS
        BgEcR            EELCRICGDRASGYHYNALSCEGCKGFFRRSITKTA--SYH-CKY--
GGNCEMDMWMRRKCQFCRLRRCREVGMKEESTFVIITQLIVEYAKCLPGFLSLSKADQITLLKSSTTEV
MTIRAARCYDLESRC------------------------IVFGNGQP-----------CSVD-N---
MKA------AGLG-HYAELLFEFCH-NMA-SVKTDNAEYTLFTAICIFS------ERPGLE--------
----AKQKVEKIQGDYVELLQAYENAKRG-KGG----NALARFLSRLSDLRTIS
        BgHR4            PLVCAICNDKATGLHYGIITCEGCKGFFKRTVQNKR--VYT-CVA--
DSNCEINKAQRNRCQYCRFKKCLKMGMVLAQVGGKIVMRFIQWMKQLPFYRDIPKELQTKILMSKWHEL
LLL-IMVS-DYRP---------TFSELYHKNMLRMQDYLDKSFNKFFT-----------LDQL-Q---
-SDI---------GEIMEKVTAIMT-YFW-ELGITKKELMCLKVILLLNH-----ES--IKK------
--ET---KLQHISASYKQALQQYILEKF---P-SEA-NRLGDVLSQLPYLQEAS
        BgFTZ_F1         EELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQNKK--TYQ-CVD--
SRNCQIDKSQRKRCPYCRFQKCLSVGMKLERIVDQLLFLMVEWARNSSFFKEIKVEDQMKLLHSSWSEI
LIL-DFMYRNVHSVW-------------------GKDIMKPNGTGIG-----------FDIL-D---
-KIG-------L-GDVKAKVFELIK-KGK-ELKLDSNEYMCLKFLILLNP-----DVSGLEN-------
--RN---FIEQSQEKVNAALFEYCLLFY---P-EMT-DKFSQLLMRLPEIRLIS
        BgPPAR2          EVLCRICGDRASGFHYGVHSCEGCKGFFRRTLKKQL--VYKPCQA--
GSRCKIDTGTRNKCQYCRYQRCLFAGMSQDELALPVMEGSVRFAKKIPGFTSLAMRDQIILMKRNGFMV
VHLALSELLTLLQEM------------------------QHCLRTVLATKNSRDSGLRIRRN-S---
YY-----ICDQ-MARLLGHAMNLID-KVH-ALGLTTGELALFSAVLL-TQ-----DCQGLS-------
----ASKQVEELQANLIEALRLDLKHNHP-K---EK-VLLAKLLMLVPDLCQIV
        BgPPAR1          HILCKICGDKASGFHYGVFSCEGCKGFFRRTIRHQL--TYK-PCD-
TPSQCLIMRISRNRCQYCRLQKCISLGMSHETQYVPSIMFITTLAKKIPQFLDVSQEDQRALIKGSILE
IAFLYDATHVKLIQDW------------------------WTNTKLK-FS-----------ISRS-H--
------LEQMSFI-GKIIEKFWSLME-KVH-KMALTDVEMSLICALLIFSP-----DR----------
-EGLTTIRYLENLETQLAMALKCQLILN---HS-DLP-YIFVHLIDIIIELRGVS
        BgNRU2           LPPCRVCDAPGAGFHYGANTCEACKGFFHRSLKLHE--QYR-CDG--
EGHCTIEHGL-
KLCQLCRYQKCLSVGMGKDIWLENMIRGMVKFAKAIPGFGKVITPDKVNLIKYSRQEFCVF-SIYT-
TLNCEL----------------------GVMRGGDEWK-----------CEYD-I----AKAG--
SFTAF-KEFIEKNFNFCK-TLQ-NLKLTEQEQVVIKAILIMAP-----DRFPPVE-----------
SSLAHEIYWYLIK--L-HLLSLND----PKPNLKFAKIISRLTEARTMT
        BgE78c           FVPCKVCGDKASGYHYGVTSCEGCKGFFRRSIQKQI--EYR-CLR--
EGKCMVIRLNRNRCQYCRFKKCLAVGMSRDALVTPSIHSVVEFSKRVPNFPDLSQDDQLILIKTGFFEI
WLARMSRMILKDEDI------------------------IVFEDGSC-----------IARQ-E---
LAV------VY-SPDFVTLLFEVAS-GFN-LLNLNDTEIGLFTGIVLATS-----DRQGLS-------
----DSQSVERIQDRLVEALKLQVSRNHS-T---EE-NLFPSIIMKLPELRTLG
        BgTR             YEPCVVCGDRASGRHYGAISCEGCKGFFKRSIRKQL--GYA-CRG--
ARDCPVTKLHRNRCQYCRLQKCLAVGMRSEESASRLLFLSMHWTRSVTAFQELNADLQTALVRSCWSEL
```

```
FTL-GLAQCSSTMCLP-TMLAAIL--------------------NHLQASLQRG---DHGNQE-----
-KVK-------NVIEHIIRLQDYVS-HTQ-SLNISPAEYAYLKSLVLFAP----------------
DNPNVRDKRLITRLQEKGRRELLLACSSP------GSS-ERVSQLLLRLIPLKTFV
        BgNR1J3         DKLCLVCGDKALGYNFNAVSCESCKAFFRRNAHKNI--R---GRC--
EGKCEVTVESRSFCKRCRLAKCFTVGMRKDNMAELFVRKVIKVAKNIASFRQLNKEDQMSLLKGSVVDI
MMLRSAVNYDPFTES------------------------WSLNTISQGSSQA--HGERISAD-I---
LKMG-----TPETKQLFMTYSKFIKSLMS-TIHGDLVVLKVLILMSLFSA-----DRMPMV--------
----DHHRVQEIQETYGRVLEEYLNHKFN-NE---A-TLFARIIMKLTDLRNIN
        BgRXR           KHICAICGDRASGKHYGVYSCEGCKGFFKRTVRKDL--TYA-CRD--
DKNCMIDKRQRNRCQYCRYMKCLSMGMKREQAADKQLFTLVEWAKRIPHFTELPLEDQVILLRAGWNEL
LIA-GFSHRSIMAK--DGILLATG-------------------LH--VHRSS--------AH-----
-QAG-------VGTIFDRVLTELVA-KMR-DMKMDKTELGCLRAVVLFNP---------------
DAKGLTAVQEVEQLREKVYASLEEYTKSRYP----EEP-GRFAKLLLRLPALRSIG
        BgNRU3          LPPCKICGEPASGFHYGLNSCGACKGFFLRSLLRPD--PYK-CST--
DGKCQVGPEIRKICQKCRFDKCLRLAIKIGLGIHNEIATVIKFFKQIPGFKHLPTSDQIKLVKGCSHEY
HIL-GLYR-GWDNTT----------------------RCMTFATETL----------TEQD-L---
-RKVF------P-DEALSAHIQMVM-TLQ-KLNLSSEQIVLLKAIVATAP-----DRDILET-------
-----VEAVTHIHWKLVN--L-YLLCKQD----LCPVIKFGQIMSVITEMRSYS
        BgHR38          SLLCAVGDNAACQHYGVRTCEGCKGFFKRTVQKNA--KYV-CLA--
DKNCPVDKRRRNRCQFCRFQKCLAVGMVKENILLQSMDVLRAWADKIPGFTDLHKDDQELLFQSASLEL
FVL-KAAYRVQVNDE----------------------KIIFENG-QV----------YHRL-Q---
------C-MKTF-GEWINSIVEFGL-SLH-RMGLDISSLACMSALAMVTL-----R------------
HGLREPEKMEELQMKIIDCLRDHCTYNSE-AQ-RKP-HFFSRILSKIAELRTLS
        Dmknirps        NQTCKVCGEPAAGFHFGAFTCEGCKSFFGRSYNNIS-TISE-CKN--
EGKCIIDKKNRTTCKACRLRKCYNVGMSKG-------------------------------------
------------------------------------------------------------------
------------------------------------------------------------------
-------------------------------------------------------
        Dmknirps_li     NQTCKVCGEPAAGFHFGAFTCEGCKSFFGRSYNNLS-SISD-CKN--
NGECIINKKNRTACKACRLKKCLMVGMSKS------------------------------------
------------------------------------------------------------------
------------------------------------------------------------------
-------------------------------------------------------
        Dm_ERR          RRLCLVCGDVASGFHYGVASCEACKAFFKRTIQGNI--EYT-CPA--
NNECEINKRRRKACQACRFQKCLLMGMLKEDIYDKELVSVIGWAKQIPGFIDLPLNDQMKLLQVSWAEI
LTL-QLTFRSLPFN----------------------GK-LCFATDVW----------MDEH-L---
-AKEC----GY--TEFYYHCVQIAQ-RME-RISPRREEYYLLKALLLANC-----DILL-D--------
----DQSSLRAFRDTILNSLNDVVYLLRHS---SAV-SHQQQLLLLLLPSLRQAD
        Dm_HR3          IIPCKVCGDKSSGVHYGVITCEGCKGFFRRSQS--SVVNYQ-CPR--
NKQCVVDRVNRNRCQYCRLQKCLKLGMSRDEKLTQMIQNIIEFAKLIPGFMRLSQDDQILLLKTGSFEL
AIVRMSRLLDLSQNA----------------------VLYGDVM----------LPQ-E-A---
FYTS-----DSEEMRLVSRIFQTAK-SIA-ELKLTETELALYQSLVLLWP-----ERNGVR-------
----GNTEIQRLFNLSMNAIRQELETNHA-PLKGDV-TVLDTLLNNIPNFRDIS
        Dm_USP          KHLCSICGDRASGKHYGVYSCEGCKGFFKRTVRKDL--TYA-CRE--
NRNCIIDKRQRNRCQYCRYQKCLTCGMKREQVVNKQLFQMVEYARMMPHFAQVPLDDQVILLKAAWIEL
LIA-NVAWCSIVSLDDGQLFLNQS-------------------FS--YHRNS--------AI-----
-KAG-------VSAIFDRILSELSV-KMK-RLNLDRRELSCLKAIILYNP---------------
DIRGIKSRAEIEMCREKVYACLDEHCRLEHP----GDD-GRFAQLLLRLPALRSIS
        Dm_HR78         VELCLVCGDRASGRHYGAISCEGCKGFFKRSIRKQL--GYQ-CRG--
AMNCEVTKHHRNRCQFCRLQKCLASGMRSDETGSRIIFLTIHTLRKVPVFEQLEAHTQVKLLRGVWPAL
MAI-ALAQCQGQLSV----------------------PTIIGQF-IQ----------STRQ-LAD-
IDKIEPLKISKM-ANLTRTLHDFVQ-ELQ-SLDVTDMEFGLLRLILLFNP-----TLLQQRKERS----
R-----GYVRRVQLYALSSLRRQGG-----IG-GGE-ERFNVLVARLLPLSSLD
        Dm_TLL          HVPCKVCRDHSSGKHYGIYACDGCAGFFKRSIRRSR--QYV-CK--
KQGLCVVDKTHRNQCRACRLRKCFEVGMNKDETAAEHLFKNVNWIKSVRAFTELPMPDQLLLLEESWKE
FFIL-AMAQYLMPMNF----------------------AQLLFVY-ES----------ENAN-R--
-------EIMGMV-TREVHAFQEVLN-QLC-HLNIDSTEYECLRAISLFRK-----
SPPSASSTEDSAESRGLLESGKVAAMHNDARSALHNYIQRT---HP-SQP-MRFQTLLGVVQLMHKVS
        Dm_PNR          GLICVVCGDTSSGKHYGILACNGCSGFFKRSVRRKL--IYR-CQA-
GTGRCVVDKAHRNQCQACRLKKCLQMGMNKDETSARLLFMAVKWAKNLPSFARLSFRDQVILLEESWSE
LFLL-NAIQWCIPLDPTCALFSVAE------------------HC-NNLENNA---NGDTCI----
```

```
--TKE-------ELAADVRTLHEIFC-KYK-AVLVDPAEFACLKAIVLFRP----------------
ETRGLKDPAQIENLQDQAHH-----TKTQFT----AQI-ARFGRLLLMLPLLRMIS
        Dm_HR38        SQLCAVGDTAACQHYGVRTCEGCKGFFKRTVQKGS--KYV-CLA--
DKNCPVDKRRRNRCQFCRFQKCLVVGMVKEQLLTSSVDVIKQFAEKIPGYFDLLPEDQELLFQSASLEL
FVL-RLAYRARIDDT---------------------KLIFCNG-TV-----------LHRT-Q---
------C-LRSF-GEWLNDIMEFSR-SLH-NLEIDISAFACLCALTLITE-----R------------
HGLREPKKVEQLQMKIIGSLRDHVTYNAE-AQ-KKQ-HYFSRLLGKLPELRSLS
        Dm_HR39        NSPCPICGDKISGFHYGIFSCESCKGFFKRTVQNRK--NYV-CVR--
GGPCQVSISTRKKCPACRFEKCLQKGMKLENVADHRLYKIVKWCKSLPLFKNISIDDQICLLINSWCEL
LLF-SCCFRSID--T---------------------PGEIKMSQGRKIT-----------LSQA-K---
-SNG-------L-QTCIERMLNLTD-HLR-RLRVDRYEYVAMKVIVLLQS-----DTTELQE-------
--AV---KVRECQEKALQSLQAYTLAHY---P-DTP-SKFGELLLRIPDLQRTC
        Dm_HR4         PLVCMICEDKATGLHYGIITCEGCKGFFKRTVQNRR--VYT-CVA--
DGTCEITKAQRNRCQYCRFKKCIEQGMVLQNIGDSIVHKLVSWTKKLPFYLEIPVEIHTKLLTDKWHEI
LIL-TTAAYQALHGKRGTPIPSPAQPLHKDSHLSTLQTCLTTLMGQPIA-----------MEQL-K---
-LDV---------GHMVDKMTQITI-MFR-RIKLKMEEYVCLKVYILLNK-----EV------------
-------ELESIQERYVQVLRSYLQNSS---P-QNPQARLSELLSHIPEIQAAA
        Ce_ODR         LHDCQVCSTHANGLHFGARTCAACAAFFRRTISDDK--R-YVCKR--
NQRCNNASRDRKICRSCRMKRCLEIGMLPE-------------------------------------
------------------------------------------------------------------
------------------------------------------------------------------
------------------------------------------------------
        Ce_HR23        VIPCKVCGDKSSGVHYGVITCEGCKGFFRRSQSSIV--NYQ-CPR--
QKNCVVDRVNRNRCQYCRLKKCIELGMSRDRELNPLIQAIIEFAKSIDGFMNLPQETQIQLLKGSVFEL
SLVFAAMYYNVD-AQ---------------------AVCGERY-SV----------P--F-A---
------CLIAED-DAEMQLINNTLQ-EIV-HLQPHQSELALLAAGLILEQ-----VSSS--------
HGIGILDTATIATAET-----LK-----NAL-YQ-SVM-PRIGCMEDTIHRIQDVE
        Ce_NR14        ISFCKVCGDKASGYHYGVTSCEGCKGFFRRSIQRKI--DYR-CLK--
QQVCEIKRESRNRCQYCRFKKCLDSGMSKDHEIDVDIKEVVNFVKEIPKFNFINGNDKAVLLRKNAFPL
YLL-RIVRGMSN--R---------------------GLMLRDG-RL-----------IDFK-S---
----LQLLYGSL-ADEM---LAFAN-HII-TIGCTDGDIALFIVLILCQP-----LTTEQQ------
FSTNFKSQLQLLEMFDFYKKVLFQKMT---C-RI-DGC-DTYKQLMKCIHELNRLN
        Ce_HR8         GRICTVCSDRANGYNFGVLTCESCKAFFRRNASKHK--EIK-CPF--
SDSCQITSASRKFCQACRLNKCFAVGMNSERIIEEALRRIVNMACQLSLFRELHVDDRKNLLKSGFGEL
LIVRGLMAYDKSDNS---------------------WNHSFGVRGK-----------MEVK-VEV-
--LKNPK----L-EEHYKAHMNLLS-TFGEDVRNNEHLMLIFNAAVIFHP-----HVSNL---------
R---DSKRVHSTQAKYFQMLLKLLTFEYG-KS-RAD-IAYSNLLNQVVELHRVN
        Ce_HR1         GELCAVCSDLATGYHYGVASCNGCKTFFRRTIVSEQ--TFI-CQY--
NGNCDVNKN-
RCACRHCRFNKCLLVGMDAKMWNIREMRICIEWAKTFDVYQRLNLFDQFALVRNFAFAFNLL-
NRVFYSPDHGP-----------------------DKIVFQGAFI-----------MRQP-Q----
QQVQ--LSGCR-PIYTRQMDEIMI-PFR-KLQLSVAEFATFKAALFFNP-----DALDLSP-------
--QAKQEVFEERNKYLGGLFTCITQKIG-IP-TGV-QKYGSLLMMTASIQNIL
        Ce_HR49        AEPCAVCGDKSTGTHYGVISCNGCKGFFRRTVLRDQ--KFT-CRF--
NKRCVIDKNFRCACRYCRFQKCVQVGMKREEISRTTLLLMVEWAKTINPFMDLSMEDKIILLKNYAPQH
LIL-MPAFRSPDTTR---------------------VCLFNNTYM----------TRDN-N---
-TDLNGAAFKTS-NITPRVLDEIVW-PMR-QLQMREQEFVCLKALAFLHP-----EAKGLSN-------
---SSQIMIRDARNRVLKALYAFILDQMP-D--DAP-TRYGNILLLAPALKALT
        Ce_UNC55       ATDCVVCGDKSSGKHYGQFSCEGCKSFFKRSIRRSL--SYT-CRA--
TKNCAIDVQHRNQCQYCRLTKCIRMGMRKEEFAAQTIFATVNWARTS--
MSNLVKGDQLILLRHSWTPIFIF-ALAQSNFALNL---------------------STHLTAV-
TA-----------TAAS-TED-EEKSEEKPERVFDEPQFQGFQAKID-KIR-
DFHLDVVESSSLRAVLLFSC-----DE-----------EALEEKGKIEEIVEKLKSAVDEYCKMN---
---KRS-ERYHQICECLQLLKSTR
        Ce_HR25        GEMCPVCGDRVSGYHYGLLTCESCKGFFKRTVQNKK--QYQ-CSA--
EANCHVDRTCRKRCPSCRFQKCLTMGMKMEQVAEENLKDIVIWAKNDQLFSKLSLDDQMILLQTSWTTV
HIV-DITNAMVHGNL--------------------LSQYKMSNGDEVP----------VGLV-A---
-LLGN--------QTFVSSWNDVVI-RLR-NMGFTNFDYCAFRFLALFDQ-----SMDSFPA-------
--VS------TARSRVLQAFLRYLWNLH---S-NCP-TNWEQFFLVLEMIRTTV
        Ce_HR91        SKLCSVCGDKSTGLHYGAATCEGCKGFFKRSVQNKK--VYH-CSQ--
DNCCEIDKQNRNRCQSCRFRKCISKGMLTERIGDEIVEQLVEWTKTLPFFDELPVEAHTHLLTQRWAEL
```

211

```
VLL-SAGYYDTDE----------ISFTNPQVNLRLLQNRLSLVLGKEIP-----------LEHV-A---
-KEA---------GPLVTRFTTLLH-SFS-NLKVSPEAYVCIKAITLLHL-----SADSTLD-------
--RSIIDKVNTLQDHFVKTLQIHLHQPG---E-DAQTTSLAQILDWLPDLRNAS
        Dm_HNF4        PTVCAICGDRATGKHYGASSCDGCKGFFRRSVRKNH--QYT-CRF--
ARNCVVDKDKRNQCRYCRLRKCFKAGMKKEESMKQQLLTLVEWAKQIPAFNELQLDDQVALLRAHAGEH
LLL-GLSRRSMH-LKD-VLLLSNN--------------------C-VITRHCP---DPLVSP-----
-NLD-------ISRIGARIIDELVT-VMK-DVGIDDTEFACIKALVFFDP----------------
NAKGLNEPHRIKSLRHQILNNLEDYISDRQY----ESR-GRFGEILLILPVLQSIT
        Dm_HR96        PKNCAVCGDKALGYNFNAVTCESCKAFFRRNALAKK--QFT-CPF--
NQNCDITVVTRRFCQKCRLRKCLDIGMKSENLTAVAIKRLIKMAKKITAFRDMCQEDQVALLKGGCTEM
MIMRSVMIYDDDRAA---------------------WKVPHTKENM-----------GNIR-TDL-
LKFAEGN----I----YEEHQKFIT-TFDEKWRMDENIILIMCAIVLFTS-----ARSRV--------
IHKD---VIRLEQNSYYYLLRRYLESVYS-GC-EAR-NAFIKLIQKISDVERLN
        Dm_EcR        EELCLVCGDRASGYHYNALTCEGCKGFFRRSVTKSA--VYC-CKF--
GRACEMDMYMRRKCQECRLKKCLAVGMRPEEITILTVQLIVEFAKGLPAFTKIPQEDQITLLKACSSEV
MMLRMARRYDHSSDS------------------------IFFANNRS-----------YTRD-S---
YKM------AGMA-DNIEDLLHFCR-QMF-SMKVDNVEYALLTAIVIFS------DRPGLE--------
----KAQLVEAIQSYYIDTLRIYILNRHC-GDSMSL-VFYAKLLSILTELRTLG
        Dm_DSF        DIPCKVCGDRSSGKHYGIYSCDGCSGFFKRSIHRNR--IYT-CKA-
LKGRCPVDKTHRNQCRACRLAKCFQSAMNKDETTARLLFMAVRWVKCLMPFQTLSKNDQHLLLQESWKE
LFLL-NLAQWTIPLDL----------------------TPILESP-LI-----------RE-R-V--
-------LQDEAT-QTEMKTIQEILC-RFR-QITPDGSEVGCMKAIALFAP-----ET----------
-AGLCDVQPVEMLQDQAQCILSDHVRLR---YP-RQA-TRFGRLLLLLPSLRTIR
        Dm_FAX1        FSACAVCGDQSSGKHYGVSCCDGCSCFFKRSVRRGS--SYA-CIA-
LVGNCVVDKARRNWCPSCRFQRCLAVGMNAA-------------------------------------
--------------------------------------------------------------------
--------------------------------------------------------------------
-----------------------------------------------------
        Ce_HR48        NKVCRVCGDKAFSYNFNVITCESCKAFFRRNANKEK--EIR-CPF--
NEQCEINMVSRRFCQRCRLTKCFAVGMKKE--------------------------------------
--------------------------------------------------------------------
--------------------------------------------------------------------
-----------------------------------------------------
        Ce_HR41        GELCVVGDKASGRHYGAVSCEGCKGFFKRSIRKRI--GYV-CRS--
QKDCPVTKFHRNRCQYCRLRKCLSMGMRSE--------------------------------------
--------------------------------------------------------------------
--------------------------------------------------------------------
-----------------------------------------------------
        Ce_FAX1        TLCCAVCGDVSSGKHYGILACNGCSGFFKRSVRRRL--IYR-CQA-
GTGNCVVDKAHRNQCQACRLKKCLNKGMNKD-------------------------------------
--------------------------------------------------------------------
--------------------------------------------------------------------
----------------------------------------------------
        Ce_HR6        DVDCRVCEDHSSGKHYSIFSCDGCAGFFKRSIRRHR--QYV-
CKNKSEGQCKVDKTHRNQCRACRLRKCLEIGMNKD----------------------------------
--------------------------------------------------------------------
--------------------------------------------------------------------
---------------------------------------------------
        Ce_NHR85        --------------------------------------------------
-------------------------
NFIVIPLKAAIDFANSIPAFLSITQTQRVHLLQNSVFDVMLL-ASASASTSQHFPG------------
-----------------LT-----------YDQS-S------AN--------PIIPQAIQSISA-RIR-
QL---PPQTVILTAIAVCQA-----DLLPESQ----------QPMLLAERLWCVLGKLG-GIQSLA--
---TAP--------SLLADVRTLR

;


END;
```
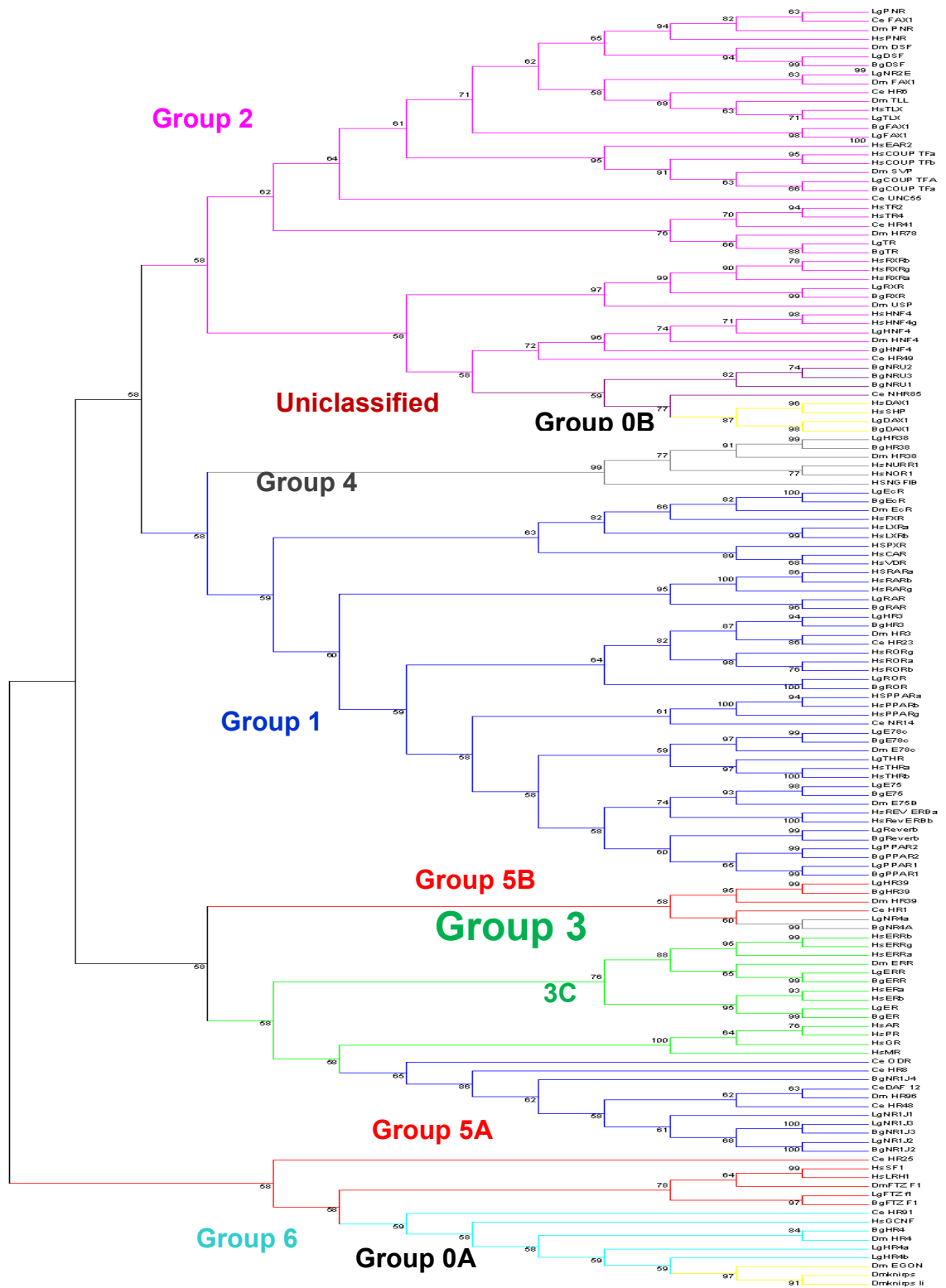
Appendix 4.2: Maximum Parsimony tree showing nuclear receptors in molluscs, human, fly and worm.

The nuclear receptors from five different species were subjected to phylogenetic comparison using Maximum Parsimony method with PAUP version 4.0b10. Notations Bg, Lg, Hs, Dm and Ce in association with the receptor name denote sequences from B. glabrata, L. gigantea, H. sapiens, D. melanogaster and C. elegans, respectively.

Appendix 4.3
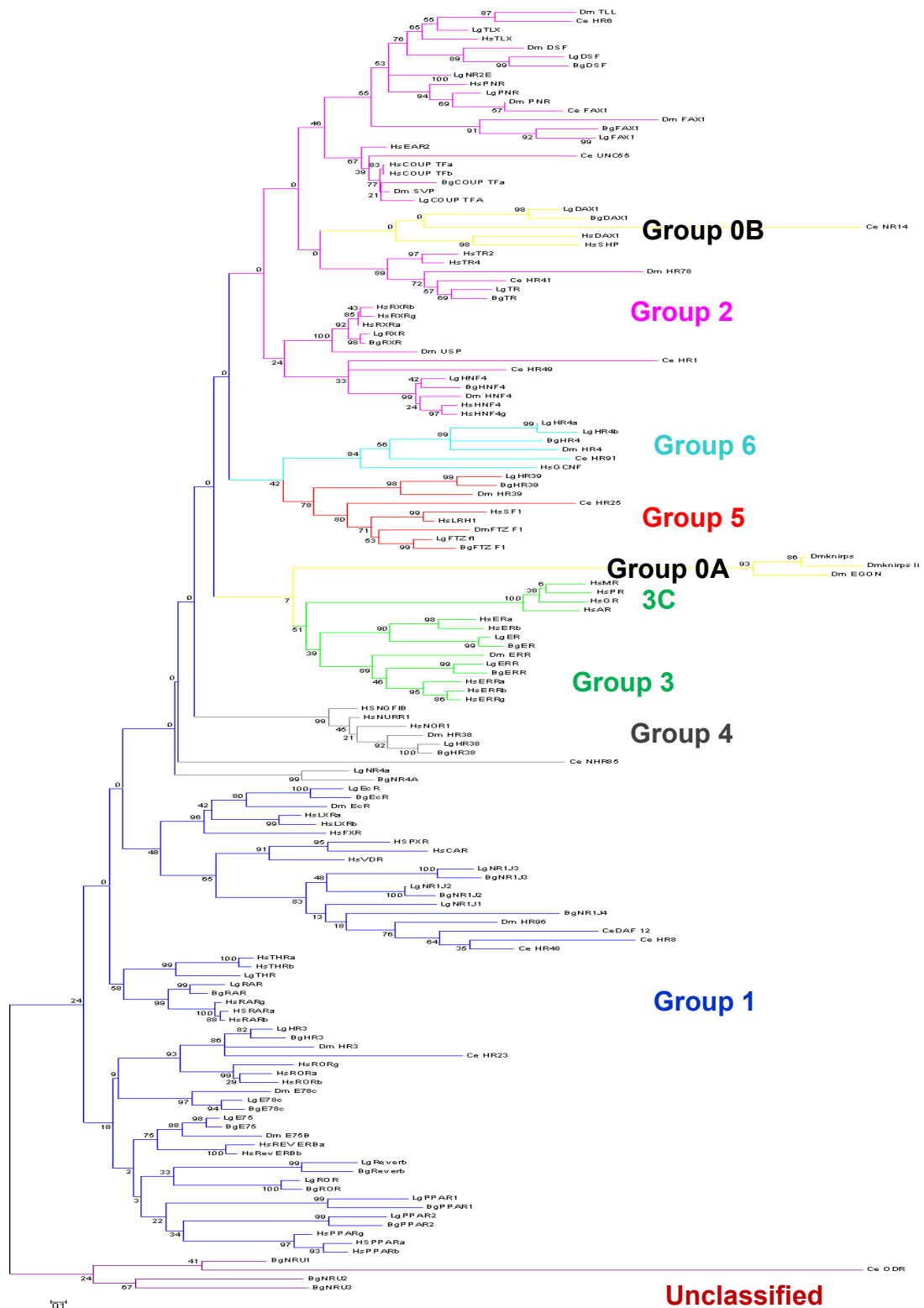
Appendix 4.3: Maximum Likelihood tree showing nuclear receptors in molluscs, human, fly and worm.

The nuclear receptors from five different species were subjected to phylogenetic comparison using Maximum Likelihood method using Jones-Taylor-Thorton (JTT) substitution model with PHYML v2.4.4 . Notations Bg, Lg, Hs, Dm and Ce in association with the receptor name denote sequences from B. glabrata, L. gigantea, H. sapiens, D. melanogaster and C. elegans, respectively.

Appendix 4.4

Appendix 4.5.pdf

Appendix 4.4: Bayesian Inference tree showing nuclear receptors in molluscs, human, fly and worm.
This nuclear receptor phylogenetic tree is the expanded for of Figure 4.0.7. Bg, Lg, Hs, Dm and Ce in association with the receptor name denote sequences from B. glabrata, L. gigantea, H. sapiens, D. melanogaster and C. elegans, respectively.

Appendix 5.1

TargetP v1.1 prediction results

```
Number of query sequences:  21
Cleavage site predictions included.
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC	TPlen
APGWamide	213	0.029	0.961	0.029	S	1	30
Schistosomin	96	0.022	0.964	0.093	S	1	17
Conopressin	157	0.030	0.946	0.056	S	1	25
ELH1	117	0.192	0.528	0.069	S	4	35
ELH2	230	0.338	0.622	0.085	S	4	22
FFamide	91	0.012	0.978	0.075	S	1	22
FMRFamide	292	0.028	0.963	0.035	S	1	23
GnRH	106	0.068	0.916	0.032	S	1	23
Whitnin	114	0.020	0.961	0.054	S	1	21
Myomodulin1	412	0.024	0.928	0.140	S	2	19
Myomodulin2	291	0.026	0.946	0.075	S	1	22
Myomodulin3	164	0.032	0.919	0.055	S	1	24
LFRFamide	187	0.021	0.879	0.139	S	2	22
NPY	90	0.014	0.975	0.057	S	1	21
Pedal_peptide1	610	0.035	0.918	0.154	S	2	23
Pedal_peptide2	200	0.024	0.920	0.072	S	1	24
Pedal_peptide3	508	0.026	0.918	0.144	S	2	18
Insulin_like_peptide	133	0.077	0.947	0.022	S	1	22
Insulin_like_peptide	183	0.035	0.926	0.046	S	1	21
Insulin_like_peptide	118	0.065	0.916	0.029	S	1	19
Insulin_like_peptide	168	0.008	0.978	0.095	S	1	25

```
Cut-off                                   0.780   0.000   0.730
```

SignalP-NN euk predictions

SignalP-HMM euk predictions

# name	Cmax	pos	?	Ymax	pos	?	Smax	pos	?	Smean	?	D	?	# name	!	Cmax	pos	?	Sprob	?
APGWamide	0.695	31	Y	0.766	31	Y	0.989	13	Y	0.872	Y	0.819	Y	APGWamide	S	0.744	31	Y	0.991	Y
Schistosomin	0.846	18	Y	0.853	18	Y	0.993	2	Y	0.953	Y	0.903	Y	Schistosomin	S	0.970	18	Y	1.000	Y
Conopressin	0.670	26	Y	0.695	26	Y	0.995	13	Y	0.872	Y	0.783	Y	Conopressin	S	0.487	26	N	0.994	Y
ELH1	0.825	36	Y	0.747	36	Y	0.963	23	Y	0.622	Y	0.684	Y	ELH1	S	0.960	36	Y	0.974	Y
ELH2	0.522	23	Y	0.457	23	Y	0.935	1	Y	0.575	Y	0.516	Y	ELH2	S	0.523	23	Y	0.990	Y
FFamide	0.493	23	Y	0.623	23	Y	0.985	13	Y	0.895	Y	0.759	Y	FFamide	S	0.769	23	Y	0.999	Y
FMRFamide	0.713	24	Y	0.742	24	Y	0.987	17	Y	0.940	Y	0.841	Y	FMRFamide	S	0.743	24	Y	0.997	Y
GnRH	0.975	24	Y	0.925	24	Y	0.996	14	Y	0.940	Y	0.933	Y	GnRH	S	0.997	24	Y	1.000	Y
Whitnin	0.690	22	Y	0.735	22	Y	0.959	11	Y	0.896	Y	0.815	Y	Whitnin	S	0.557	24	Y	0.977	Y
Myomodulin1	0.805	20	Y	0.749	20	Y	0.960	1	Y	0.781	Y	0.765	Y	Myomodulin1	S	0.983	20	Y	0.997	Y
Myomodulin2	0.577	23	Y	0.622	23	Y	0.954	6	Y	0.780	Y	0.701	Y	Myomodulin2	S	0.609	23	Y	0.997	Y
Myomodulin3	0.813	25	Y	0.795	25	Y	0.986	13	Y	0.877	Y	0.836	Y	Myomodulin3	S	0.639	25	Y	0.999	Y
LFRFamide	0.833	23	Y	0.744	23	Y	0.891	13	Y	0.704	Y	0.724	Y	LFRFamide	S	0.824	23	Y	0.987	Y
NPY	0.612	22	Y	0.683	22	Y	0.997	13	Y	0.924	Y	0.804	Y	NPY	S	0.690	22	Y	0.994	Y
Pedal_peptide1	0.218	24	N	0.335	24	Y	0.950	9	Y	0.669	Y	0.502	Y	Pedal_peptide1	Q	0.258	24	N	0.457	N
Pedal_peptide2	0.451	25	Y	0.594	25	Y	0.977	16	Y	0.841	Y	0.717	Y	Pedal_peptide2	S	0.737	25	Y	0.933	Y
Pedal_peptide3	0.637	19	Y	0.710	19	Y	0.934	9	Y	0.814	Y	0.762	Y	Pedal_peptide3	S	0.716	19	Y	0.987	Y
Insulin_like_peptide	0.914	23	Y	0.865	23	Y	0.992	1	Y	0.898	Y	0.882	Y	Insulin_like_peptide1	S	0.919	23	Y	0.999	Y
Insulin_like_peptide	0.908	22	Y	0.855	22	Y	0.982	13	Y	0.892	Y	0.874	Y	Insulin_like_peptide2	S	0.898	22	Y	0.999	Y
Insulin_like_peptide	0.898	20	Y	0.855	20	Y	0.995	13	Y	0.940	Y	0.897	Y	Insulin_like_peptide3	S	0.634	20	Y	0.995	Y
Insulin_like_peptide	0.746	26	Y	0.739	26	Y	0.973	12	Y	0.805	Y	0.772	Y	Insulin_like_peptide4	S	0.893	26	Y	0.921	Y

TMHMM 2.0c predictions

```
# APGWamide Length: 213
# APGWamide Number of predicted TMHs:  0
# APGWamide Exp number of AAs in TMHs: 12.75856
# APGWamide Exp number, first 60 AAs:  12.75856
# APGWamide Total prob of N-in:        0.51992
# APGWamide POSSIBLE N-term signal sequence
APGWamide        TMHMM2.0        outside     1    213
```

```
# Schistosomin Length: 96
# Schistosomin Number of predicted TMHs:  0
# Schistosomin Exp number of AAs in TMHs: 0.64892
# Schistosomin Exp number, first 60 AAs:  0.64892
```

```
# Schistosomin Total prob of N-in:        0.08835
Schistosomin    TMHMM2.0       outside     1    96
```

```
# Conopressin Length: 157
# Conopressin Number of predicted TMHs:  1
# Conopressin Exp number of AAs in TMHs: 16.65291
# Conopressin Exp number, first 60 AAs:  16.65269
# Conopressin Total prob of N-in:        0.08581
# Conopressin POSSIBLE N-term signal sequence
Conopressin     TMHMM2.0       outside     1     4
Conopressin     TMHMM2.0       TM helix    5    27
Conopressin     TMHMM2.0       inside      28   157
```

```
# ELH1 Length: 117
# ELH1 Number of predicted TMHs:  0
# ELH1 Exp number of AAs in TMHs: 8.89296
# ELH1 Exp number, first 60 AAs:  8.89113
# ELH1 Total prob of N-in:        0.37967
ELH1    TMHMM2.0       outside           1   117
```

```
# ELH2 Length: 230
# ELH2 Number of predicted TMHs:  0
# ELH2 Exp number of AAs in TMHs: 0.11823
# ELH2 Exp number, first 60 AAs:  0.11085
# ELH2 Total prob of N-in:        0.02623
ELH2    TMHMM2.0       outside           1   230
```

```
# FFamide Length: 91
# FFamide Number of predicted TMHs:  0
# FFamide Exp number of AAs in TMHs: 5.81686
# FFamide Exp number, first 60 AAs:  5.81669
# FFamide Total prob of N-in:        0.15353
FFamide TMHMM2.0       outside     1    91
```

```
# FMRFamide Length: 292
# FMRFamide Number of predicted TMHs:  0
# FMRFamide Exp number of AAs in TMHs: 3.21426
# FMRFamide Exp number, first 60 AAs:  3.21426
# FMRFamide Total prob of N-in:        0.18827
FMRFamide       TMHMM2.0       outside     1   292
```

```
# GnRH Length: 106
```

```
# GnRH Number of predicted TMHs:  1
# GnRH Exp number of AAs in TMHs: 16.72643
# GnRH Exp number, first 60 AAs:  16.72384
# GnRH Total prob of N-in:        0.68137
# GnRH POSSIBLE N-term signal sequence
GnRH    TMHMM2.0        inside              1     4
GnRH    TMHMM2.0        TMhelix             5    23
GnRH    TMHMM2.0        outside            24   106
```

```
# Whitnin Length: 114
# Whitnin Number of predicted TMHs:  0
# Whitnin Exp number of AAs in TMHs: 6.74482
# Whitnin Exp number, first 60 AAs:  6.74482
# Whitnin Total prob of N-in:        0.02809
Whitnin TMHMM2.0        outside             1   114
```

```
# Myomodulin1 Length: 412
# Myomodulin1 Number of predicted TMHs:  0
# Myomodulin1 Exp number of AAs in TMHs: 0.23007
# Myomodulin1 Exp number, first 60 AAs:  0.22997
# Myomodulin1 Total prob of N-in:        0.01416
Myomodulin1     TMHMM2.0        outside     1   412
```

```
# Myomodulin2 Length: 291
# Myomodulin2 Number of predicted TMHs:  0
# Myomodulin2 Exp number of AAs in TMHs: 0.13538
# Myomodulin2 Exp number, first 60 AAs:  0.13538
# Myomodulin2 Total prob of N-in:        0.01472
Myomodulin2     TMHMM2.0        outside     1   291
```

```
# Myomodulin3 Length: 164
# Myomodulin3 Number of predicted TMHs:  0
# Myomodulin3 Exp number of AAs in TMHs: 3.00215
# Myomodulin3 Exp number, first 60 AAs:  3.00215
# Myomodulin3 Total prob of N-in:        0.16742
Myomodulin3     TMHMM2.0        outside     1   164
```

```
# LFRFamide Length: 187
# LFRFamide Number of predicted TMHs:  0
# LFRFamide Exp number of AAs in TMHs: 4.54339
# LFRFamide Exp number, first 60 AAs:  4.54301
# LFRFamide Total prob of N-in:        0.05176
```

```
LFRFamide        TMHMM2.0        outside        1    187
```

```
# NPY Length: 90
# NPY Number of predicted TMHs:  0
# NPY Exp number of AAs in TMHs: 14.12354
# NPY Exp number, first 60 AAs:  14.12351
# NPY Total prob of N-in:        0.53621
# NPY POSSIBLE N-term signal sequence
NPY     TMHMM2.0        outside            1    90
```

```
# Pedal_peptide1 Length: 610
# Pedal_peptide1 Number of predicted TMHs:  0
# Pedal_peptide1 Exp number of AAs in TMHs: 0.08516
# Pedal_peptide1 Exp number, first 60 AAs:  0.08516
# Pedal_peptide1 Total prob of N-in:        0.00445
Pedal_peptide1 TMHMM2.0        outside        1    610
```

```
# Pedal_peptide2 Length: 200
# Pedal_peptide2 Number of predicted TMHs:  0
# Pedal_peptide2 Exp number of AAs in TMHs: 9.26740999999999999
# Pedal_peptide2 Exp number, first 60 AAs:  9.26723
# Pedal_peptide2 Total prob of N-in:        0.50095
Pedal_peptide2 TMHMM2.0        outside        1    200
```

```
# Pedal_peptide3 Length: 508
# Pedal_peptide3 Number of predicted TMHs:  0
# Pedal_peptide3 Exp number of AAs in TMHs: 0.03318
# Pedal_peptide3 Exp number, first 60 AAs:  0.03298
# Pedal_peptide3 Total prob of N-in:        0.00244
Pedal_peptide3 TMHMM2.0        outside        1    508
```

```
# Insulin-like_peptide1 Length: 133
# Insulin-like_peptide1 Number of predicted TMHs:  0
# Insulin-like_peptide1 Exp number of AAs in TMHs: 0.18066
# Insulin-like_peptide1 Exp number, first 60 AAs:  0.18019
# Insulin-like_peptide1 Total prob of N-in:        0.05416
Insulin-like_peptide1  TMHMM2.0        outside        1    133
```

```
# Insulin-like_peptide2 Length: 183
# Insulin-like_peptide2 Number of predicted TMHs:  0
# Insulin-like_peptide2 Exp number of AAs in TMHs: 1.0999
# Insulin-like_peptide2 Exp number, first 60 AAs:  1.0999
```

```
# Insulin-like_peptide2 Total prob of N-in:          0.08493
Insulin-like_peptide2  TMHMM2.0        outside     1    183


# Insulin-like_peptide3 Length: 118
# Insulin-like_peptide3 Number of predicted TMHs:  0
# Insulin-like_peptide3 Exp number of AAs in TMHs: 0.40773
# Insulin-like_peptide3 Exp number, first 60 AAs:  0.4066
# Insulin-like_peptide3 Total prob of N-in:          0.06364
Insulin-like_peptide3  TMHMM2.0        outside     1    118

# Insulin-like_peptide4 Length: 168
# Insulin-like_peptide4 Number of predicted TMHs:  1
# Insulin-like_peptide4 Exp number of AAs in TMHs: 17.61363
# Insulin-like_peptide4 Exp number, first 60 AAs:  17.60696
# Insulin-like_peptide4 Total prob of N-in:          0.81821
# Insulin-like_peptide4 POSSIBLE N-term signal sequence
Insulin-like_peptide4  TMHMM2.0        inside      1      6
Insulin-like_peptide4  TMHMM2.0        TM helix    7     29
Insulin-like_peptide4  TMHMM2.0        outside     30   168
```

Phobius predictions

```
ID    APGWamide
FT    SIGNAL          1     30
FT    REGION          1      9       N-REGION.
FT    REGION         10     25       H-REGION.
FT    REGION         26     30       C-REGION.
FT    TOPO_DOM       31    213       NON CYTOPLASMIC.
//
ID    Schistosomin
FT    SIGNAL          1     17
FT    REGION          1      2       N-REGION.
FT    REGION          3     13       H-REGION.
FT    REGION         14     17       C-REGION.
FT    TOPO_DOM       18     96       NON CYTOPLASMIC.
//
ID    Conopressin
FT    SIGNAL          1     23
FT    REGION          1      3       N-REGION.
```

221

```
FT      REGION          4      18         H-REGION.
FT      REGION         19      23         C-REGION.
FT      TOPO_DOM       24     157         NON CYTOPLASMIC.


//
ID      ELH1
FT      SIGNAL          1      35
FT      REGION          1      14         N-REGION.
FT      REGION         15      26         H-REGION.
FT      REGION         27      35         C-REGION.
FT      TOPO_DOM       36     117         NON CYTOPLASMIC.
//
ID      ELH2
FT      SIGNAL          1      27
FT      REGION          1       3         N-REGION.
FT      REGION          4      15         H-REGION.
FT      REGION         16      27         C-REGION.
FT      TOPO_DOM       28     230         NON CYTOPLASMIC.
//
ID      FFamide
FT      SIGNAL          1      22
FT      REGION          1       4         N-REGION.
FT      REGION          5      17         H-REGION.
FT      REGION         18      22         C-REGION.
FT      TOPO_DOM       23      91         NON CYTOPLASMIC.
//
ID      FMRFamide
FT      SIGNAL          1      23
FT      REGION          1       6         N-REGION.
FT      REGION          7      18         H-REGION.
FT      REGION         19      23         C-REGION.
FT      TOPO_DOM       24     292         NON CYTOPLASMIC.
//
ID      GnRH
FT      SIGNAL          1      23
FT      REGION          1       5         N-REGION.
FT      REGION          6      18         H-REGION.
FT      REGION         19      23         C-REGION.
FT      TOPO_DOM       24     106         NON CYTOPLASMIC.
//
ID      Whitnin
FT      SIGNAL          1      21
```

```
FT   REGION          1    4       N-REGION.
FT   REGION          5   16       H-REGION.
FT   REGION         17   21       C-REGION.
FT   TOPO_DOM       22  114       NON CYTOPLASMIC.
//
ID   Myomodulin1
FT   SIGNAL          1   19
FT   REGION          1    4       N-REGION.
FT   REGION          5   14       H-REGION.
FT   REGION         15   19       C-REGION.
FT   TOPO_DOM       20  412       NON CYTOPLASMIC.
//
ID   Myomodulin2
FT   SIGNAL          1   22
FT   REGION          1    5       N-REGION.
FT   REGION          6   17       H-REGION.
FT   REGION         18   22       C-REGION.
FT   TOPO_DOM       23  291       NON CYTOPLASMIC.
//
ID   Myomodulin3
FT   SIGNAL          1   24
FT   REGION          1    5       N-REGION.
FT   REGION          6   17       H-REGION.
FT   REGION         18   24       C-REGION.
FT   TOPO_DOM       25  164       NON CYTOPLASMIC.
//
ID   LFRFamide
FT   SIGNAL          1   22
FT   REGION          1    5       N-REGION.
FT   REGION          6   17       H-REGION.
FT   REGION         18   22       C-REGION.
FT   TOPO_DOM       23  187       NON CYTOPLASMIC.
//
ID   NPY
FT   SIGNAL          1   17
FT   REGION          1    3       N-REGION.
FT   REGION          4   12       H-REGION.
FT   REGION         13   17       C-REGION.
FT   TOPO_DOM       18   90       NON CYTOPLASMIC.
//
ID   Pedal_peptide1
FT   SIGNAL          1   16
```

```
FT   REGION         1      2          N-REGION.
FT   REGION         3     11          H-REGION.
FT   REGION        12     16          C-REGION.
FT   TOPO_DOM      17    610          NON CYTOPLASMIC.
//
ID   Pedal_peptide2
FT   SIGNAL         1     24
FT   REGION         1      3          N-REGION.
FT   REGION         4     16          H-REGION.
FT   REGION        17     24          C-REGION.
FT   TOPO_DOM      25    200          NON CYTOPLASMIC.
//
ID   Pedal_peptide3
FT   SIGNAL         1     17
FT   REGION         1      4          N-REGION.
FT   REGION         5     12          H-REGION.
FT   REGION        13     17          C-REGION.
FT   TOPO_DOM      18    508          NON CYTOPLASMIC.
//
ID   Insulin-like_peptide1
FT   SIGNAL         1     22
FT   REGION         1      6          N-REGION.
FT   REGION         7     14          H-REGION.
FT   REGION        15     22          C-REGION.
FT   TOPO_DOM      23    133          NON CYTOPLASMIC.
//
ID   Insulin-like_peptide2
FT   SIGNAL         1     21
FT   REGION         1      3          N-REGION.
FT   REGION         4     14          H-REGION.
FT   REGION        15     21          C-REGION.
FT   TOPO_DOM      22    183          NON CYTOPLASMIC.
//
ID   Insulin-like_peptide3
FT   SIGNAL         1     19
FT   REGION         1      1          N-REGION.
FT   REGION         2     12          H-REGION.
FT   REGION        13     19          C-REGION.
```

```
FT    TOPO_DOM      20    118       NON CYTOPLASMIC.
//
ID    Insulin-like_peptide4
FT    SIGNAL        1     25
FT    REGION        1      6        N-REGION.
FT    REGION        7     20        H-REGION.
FT    REGION       21     25        C-REGION.
FT              TOPO_DOM                    26            168                    NON    CYTOPLASMIC
```

REFERENCES

Abascal, F., Zardoya, R. and Posada, D. (2005) 'ProtTest: selection of best-fit models of protein evolution', *Bioinformatics,* 21(9), pp. 2104-2105.

Abdraba, A. and Saleuddin, A.S. (2000) 'Localization and immunological characterization of insulin-like peptide (s) in the land snail *Otala lactea* (Mollusca: Pulmonata)', *Canadian journal of zoology,* 78(9), pp. 1515-1526.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A. and Galle, R.F. (2000) 'The genome sequence of *Drosophila melanogaster*', *Science,* 287(5461), pp. 2185-2195.

Adema, C. and Loker, E. (2009) '*Biomphalaria glabrata* (Gastropoda, Mollusca): Responses to immune challenges, a genomics perspective', *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology,* 154(1), pp. S1.

Adema, C.M., Luo, M., Hanelt, B., Hertel, L.A., Marshall, J.J., Zhang, S., DeJong, R.J., Kim, H., Kudrna, D. and Wing, R.A. (2006a) 'A bacterial artificial chromosome library for *Biomphalaria glabrata*, intermediate snail host of *Schistosoma mansoni*', *Memórias do Instituto Oswaldo Cruz,* 101, pp. 167-177.

Adema, C.M., Luo, M., Hanelt, B., Hertel, L.A., Marshall, J.J., Zhang, S., DeJong, R.J., Kim, H., Kudrna, D. and Wing, R.A. (2006b) 'A bacterial artificial chromosome library for *Biomphalaria glabrata*, intermediate snail host of *Schistosoma mansoni*', *Memórias do Instituto Oswaldo Cruz,* 101, pp. 167-177.

Adiyodi, K. and Adiyodi, R. (1988) *Reproductive Biology of Invertebrates, Accessory Sex Glands.* Wiley Interscience publication.

Adkins-Regan, E. (2007) 'Hormones and the development of sex differences in behavior', *Journal of Ornithology,* 148(1), pp. 17-26.

Affleck, J.G., Neumann, K., Wong, L. and Walker, V.K. (2006) 'The effects of methotrexate on *Drosophila* development, female fecundity, and gene expression', *Toxicological sciences : an official journal of the Society of Toxicology,* 89(2), pp. 495-503.

Affleck, J.G. and Walker, V.K. (2007) 'Transgenic rescue of methotrexate-induced teratogenicity in *Drosophila melanogaster*', *Toxicological sciences : an official journal of the Society of Toxicology,* 99(2), pp. 522-531.

Albalat, R. (2009) 'The retinoic acid machinery in invertebrates: ancestral elements and vertebrate innovations', *Molecular and cellular endocrinology,* 313(1), pp. 23-35.

Albanis, T., Allera, A., Bachmann, J., Barbaglio, A., Berntsson, P., Beresford, N., Dittmann, N., Carnivali, D., Ciceri, F. and Dagnac, T. (2006) *Comparative research on endocrine*

disrupters-Phylogenetic approach and common principles focussing on androgenic/antiandrogenic compounds-COMPRENDO: Final published report, .

Albers, H.E. and Bamshad, M. (1999) 'Role of vasopressin and oxytocin in the control of social behavior in Syrian hamsters (*Mesocricetus auratus)*', *Progress in brain research,* 119, pp. 395-408.

Alonso-Magdalena P, Vieira E, Soriano S, Menes L, Burks D, Quesada I, Nadal A (2010) '
Bisphenol A Exposure during Pregnancy Disrupts Glucose Homeostasis in Mothers and Adult Male Offspring. ', *Environmental Health Perspectives,* 118(9), pp. 1243-1250.

Alsop, D., Hewitt, M., Kohli, M., Brown, S. and Van Der Kraak, G. (2003) 'Constituents within pulp mill effluent deplete retinoid stores in white sucker and bind to rainbow trout retinoic acid receptors and retinoid× receptors', *Environmental toxicology and chemistry,* 22(12), pp. 2969-2976.

Amano, M., Oka, Y., Nagai, Y., Amiya, N. and Yamamori, K. (2008) 'Immunohistochemical localization of a GnRH-like peptide in the brain of the cephalopod spear-squid, *Loligo bleekeri*', *General and comparative endocrinology,* 156(2), pp. 277-284.

Anderson, S.L., Hose, J.E. and Knezovich, J.P. (1994) 'Genotoxic and developmental effects in sea urchins are sensitive indicators of effects of genotoxic chemicals', *Environmental toxicology and chemistry,* 13(7), pp. 1033-1041.

Andrew NR (1995) 'Alternatives and Replacement;', *The World Congress on Alternatives and Animal Use in the Life Sciences: Education, Research, Testing.* Mary Ann Liebert, 1.

Andrews, P.L. (2011) 'Laboratory Invertebrates: Only Spineless, or Spineless and Painless?', *ILAR J,* 52, pp. 121-125.

Ansaldo, M., Nahabedian, D.E., Di Fonzo, C. and Wider, E.A. (2009) 'Effect of cadmium, lead and arsenic on the oviposition, hatching and embryonic survival of *Biomphalaria glabrata*', *Science of the Total Environment,* 407(6), pp. 1923-1928.

Ansaldo, M., Nahabedian, D.E., Holmes-Brown, E., Agote, M., Ansay, C.V., Guerrero, N.R.V. and Wider, E.A. (2006) 'Potential use of glycogen level as biomarker of chemical stress in *Biomphalaria glabrata*', *Toxicology,* 224(1), pp. 119-127.

Antebi, A., Yeh, W., Tait, D., Hedgecock, E. and Riddle, D. (2000) 'daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in *C-elegans*', *Genes & development,* 14(12), pp. 1512-1527.

Antoshechkin, I. and Sternberg, P.W. (2007) 'The versatile worm: genetic and genomic resources for *Caenorhabditis elegans* research', *Nature Reviews Genetics,* 8(7), pp. 518-532.

Aranda, A. and Pascual, A. (2001) 'Nuclear hormone receptors and gene expression', *Physiological Reviews,* 81(3), pp. 1269-1304.

Arch, S. (1976) 'Neuroendocrine regulation of egg laying in *Aplysia californica*', *American Zoologist,* 16(2), pp. 167-175.

Arch, S. and Smock, T. (1977) 'Egg-laying behavior in *Aplysia californica*', *Behavioral biology,*

19(1), pp. 45-54.

Arch, S., Earley, P. and Smock, T. (1976) 'Biochemical isolation and physiological identification of the egg-laying hormone in *Aplysia californica*', *The Journal of general physiology,* 68(2), pp. 197-210.

Ashby, J. and Lefevre, P. (2000) 'The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti- androgens, oestrogens and metabolic modulators', *Journal of Applied Toxicology,* 20(1), pp. 35-47.

Avanesian, A., Semnani, S. and Jafari, M. (2009) 'Can *Drosophila melanogaster* represent a model system for the detection of reproductive adverse drug reactions?', *Drug discovery today,* 14(15), pp. 761-766.

Bailey CH, K.E. (2008) 'Synaptic remodelling, synaptic growth and the storage of long term memory in *Aplysia.*', *Prog Brain Res,* 169, pp. 179–198.

Baker, M.E. (2008) 'Trichoplax, the simplest known animal, contains an estrogen-related receptor but no estrogen receptor: Implications for estrogen receptor evolution', *Biochemical and biophysical research communications,* 375(4), pp. 623-627.

Baker, K.D., Shewchuk, L.M., Kozlova, T., Makishima, M., Hassell, A., Wisely, B., Caravella, J.A., Lambert, M.H., Reinking, J.L., Krause, H., Thummel, C.S., Willson, T.M. and Mangelsdorf, D.J. (2003) 'The *Drosophila* orphan nuclear receptor DHR38 mediates an atypical ecdysteroid signaling pathway', *Cell,* 113(6).

Bamigboye, A.A. and Morris, J. (2003) 'Oestrogen supplementation, mainly diethylstilbestrol, for preventing miscarriages and other adverse pregnancy outcomes', *Cochrane Database Syst Rev,* 3.

Bannister, R., Beresford, N., Granger, D.W., Pounds, N.A., Rand-Weaver, M., White, R., Jobling, S. and Routledge, E.J. (2013) 'No Substantial Changes in Estrogen Receptor and Estrogen-Related Receptor Orthologue Gene Transcription in *Marisa cornuarietis* Exposed to Estrogenic Chemicals', *Aquatic Toxicology, .*

Bannister, R., Beresford, N., May, D., Routledge, E.J., Jobling, S. and Rand-Weaver, M. (2007) 'Novel estrogen receptor-related transcripts in *Marisa cornuarietis*; a freshwater snail with reported sensitivity to estrogenic chemicals', *Environmental science & technology,* 41(7), pp. 2643-2650.

Bardet, P., Laudet, V. and Vanacker, J. (2006) 'Studying non-mammalian models? *Not a fool's ERRand*!', *Trends Endocrin Met,* 17.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths- Jones, S., Khanna, A., Marshall, M., Moxon, S. and Sonnhammer, E.L. (2004) 'The Pfam protein families database', *Nucleic acids research,* 32(suppl 1), pp. D138-D141.

Bayne, C.J. (2009a) 'Successful parasitism of vector snail *Biomphalaria glabrata* by the human blood fluke (trematode) *Schistosoma mansoni*: A 2009 assessment', *Molecular and biochemical parasitology,* 165(1), pp. 8-18.

Bayne, C.J. (2009b) 'Successful parasitism of vector snail *Biomphalaria glabrata* by the human blood fluke (trematode) *Schistosoma mansoni*: A 2009 assessment', *Molecular and biochemical parasitology,* 165(1), pp. 8-18.

Beckingham, K.M., Armstrong, J.D., Texada, M.J., Munjaal, R. and Baker, D.A. (2007) '*Drosophila melanogaster*-the model organism of choice for the complex biology of multi-cellular organisms', *Gravitational and Space Biology,* 18(2).

Beets, I., Temmerman, L., Janssen, T. and Schoofs, L. (2013) 'Ancient neuromodulation by vasopressin/oxytocin-related peptides', *Worm,* 2(2), pp. e24246.

Benjamin, P.R. (2008) 'Lymnaea', *Scholarpedia,* 3(1), pp. 4124.

Benjamin, P.R. and Kemenes, I. (2013) '*Lymnaea* neuropeptide genes', *Scholarpedia,* 8(7), pp. 11520.

Benoit, G., Cooney, A., Giguere, V., Ingraham, H., Lazar, M., Muscat, G., Perlmann, T., Renaud, J., Schwabe, J. and Sladek, F. (2006) 'International union of pharmacology. LXVI. Orphan nuclear receptors', *Pharmacological reviews,* 58(4), pp. 798-836.

Benstead, R.S., Baynes, A., Casey, D., Routledge, E.J. and Jobling, S. (2011a) '17β-Oestradiol may prolong reproduction in seasonally breeding freshwater gastropod molluscs', *Aquatic Toxicology,* 101(2), pp. 326-334.

Benstead, R.S., Baynes, A., Casey, D., Routledge, E.J. and Jobling, S. (2011b) '17β-Oestradiol may prolong reproduction in seasonally breeding freshwater gastropod molluscs', *Aquatic Toxicology,* 101(2), pp. 326-334.

Bernal, J. (2007a) 'Thyroid hormone receptors in brain development and function', *Nature Clinical Practice Endocrinology & Metabolism,* 3(3), pp. 249-259.

Bernal, J. (2007b) 'Thyroid hormone receptors in brain development and function', *Nature Clinical Practice Endocrinology & Metabolism,* 3(3), pp. 249-259.

Berriman, M., Haas, B.J., LoVerde, P.T., Wilson, R.A., Dillon, G.P., Cerqueira, G.C., Mashiyama, S.T., Al-Lazikani, B., Andrade, L.F. and Ashton, Peter D, et al. (2009) 'The genome of the blood fluke *Schistosoma mansoni*', *Nature,* 460(7253), pp. 352-358.

Bertrand, S., Brunet, F.G., Escriva, H., Parmentier, G., Laudet, V. and Robinson-Rechavi, M. (2004) 'Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems', *Molecular biology and evolution,* 21(10), pp. 1923-1937.

Bettin, C., Oehlmann, J. and Stroben, E. (1996c) 'TBT-induced imposex in marine neogastropods is mediated by an increasing androgen level', *Helgoländer Meeresuntersuchungen,* 50(3), pp. 299-317.

Bier, E. and McGinnis, W. (2004) 'Model organisms in the study of development and disease', *Oxford Monographs On Medical Genetics,* 49, pp. 25-45.

Bigot, L., Beets, I., Dubos, M.P., Boudry, P., Schoofs, L. and Favrel, P. (2014) 'Functional characterization of a short neuropeptide F-related receptor in a lophotrochozoan, the mollusk Crassostrea gigas', *The Journal of experimental biology,* 217(Pt 16), pp. 2974-2982.

Blatt, J., Van Le, L., Weiner, T. and Sailer, S. (2003) 'Ovarian carcinoma in an adolescent with transgenerational exposure to diethylstilbestrol', *Journal of pediatric hematology/oncology,* 25(8), pp. 635-636.

Blazer, V., Iwanowicz, L., Iwanowicz, D., Smith, D., Young, J., Hedrick, J., Foster, S. and Reeser, S. (2007) 'Intersex (testicular oocytes) in smallmouth bass from the Potomac River and selected nearby drainages', *Journal of Aquatic Animal Health,* 19(4), pp. 242-253.

Bloom, B., Cohen, R.A. and Freeman, G. (2010) 'Summary health statistics for U.S. children: National Health Interview Survey, 2009', *Vital and health statistics.Series 10, Data from the National Health Survey,* (247)(247), pp. 1-82.

Boer, H.H., Slot, J. and Van Andel, J. (1968) 'Electron microscopical and histochemical observations on the relation between medio-dorsal bodies and neurosecretory cells in the basommatophoran snails *Lymnaea stagnalis*, *Ancylus fluviatilis*, *Australorbis glabratus* and *Planorbarius corneus*', *Zeitschrift für Zellforschung und Mikroskopische Anatomie,* 87(3), pp. 435-450.

Boldajipour, B., Mahabaleshwar, H., Kardash, E., Reichman-Fried, M., Blaser, H., Minina, S., Wilson, D., Xu, Q. and Raz, E. (2008) 'Control of chemokine-guided cell migration by ligand sequestration', *Cell,* 132(3), pp. 463-473.

Bolker, J. (2012) 'There's more to life than rats and flies', *Nature,* 491(7422), pp. 31-33.

Bonneton, F., Chaumot, A. and Laudet, V. (2008) 'Annotation of *Tribolium* nuclear receptors reveals an increase in evolutionary rate of a network controlling the ecdysone cascade', *Insect biochemistry and molecular biology,* 38(4), pp. 416-429.

Bookout, A.L., Jeong, Y., Downes, M., Yu, R.T., Evans, R.M. and Mangelsdorf, D.J. (2006) 'Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network', *Cell,* 126(4), pp. 789-799.

Börjesson, L. (2012) 'The estrogen receptor in the freshwater pulmonate gastropod *Radix balthica*', .

Bose, R., Majumdar, C. and Bhattacharya, S. (1997) 'Steroids in *Achatina fulica* (Bowdich): Steroid Profile in Haemolymph and *In Vitro* Release of Steroids from Endogenous Precursors by Ovotestis and Albumen Gland', *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology,* 116(3), pp. 179-182.

Bourguet, W., Germain, P. and Gronemeyer, H. (2000) 'Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications', *Trends in pharmacological sciences,* 21(10), pp. 381-388.

Bournias- Vardiabasis, N., Teplitz, R.L., Chernoff, G.F. and Seecof, R.L. (1983) 'Detection of teratogens in the *Drosophila* embryonic cell culture test: assay of 100 chemicals', *Teratology,* 28(1), pp. 109-122.

Bouton, D., Escriva, H., de Mendonca, R.L., Glineur, C., Bertin, B., Noel, C., Robinson-Rechavi, M., de Groot, A., Cornette, J., Laudet, V. and Pierce, R.J. (2005) 'A conserved retinoid X

receptor (RXR) from the mollusk *Biomphalaria glabrata* transactivates transcription in the presence of retinoids', *Journal of Molecular Endocrinology,* 34(2).

Boyce, C.B., Tieze-Dagevos, J.W. and Larman, V.N. (1967) 'The susceptibility of Biomphalaria glabrata throughout its life-history to N-tritylmorpholine', *Bulletin of the World Health Organization,* 37(1), pp. 13-21.

Boyle, J. and Yoshino, T. (2000) 'The effect of water quality on oviposition in *Biomphalaria glabrata* (Say, 1818)(Planorbidae), and a description of the stages of the egg-laying process', *Journal of Molluscan Studies,* 66(1), pp. 83-94.

Boyle, P.M. and Silver, P.A. (2012) 'Parts plus pipes: synthetic biology approaches to metabolic engineering', *Metabolic engineering,* 14(3), pp. 223-232.

Branco, M.R. and Pombo, A. (2007) 'Chromosome organization: new facts, new models', *Trends in cell biology,* 17(3), pp. 127-134.

Brenner, S. (1974) 'The genetics of *Caenorhabditis elegans*', *Genetics,* 77(1), pp. 71-94.

Bridgham, J.T., Brown, J.E., Rodríguez-Marí, A., Catchen, J.M. and Thornton, J.W. (2008) 'Evolution of a new function by degenerative mutation in cephalochordate steroid receptors', *PLoS genetics,* 4(9), pp. e1000191.

Bridgham, J.T., Eick, G.N., Larroux, C., Deshpande, K., Harms, M.J., Gauthier, M.E., Ortlund, E.A., Degnan, B.M. and Thornton, J.W. (2010) 'Protein evolution by molecular tinkering: diversification of the nuclear receptor superfamily from a ligand-dependent ancestor', *PLoS biology,* 8(10), pp. e1000497.

Brown, D.D. (2005) 'The role of deiodinases in amphibian metamorphosis', *Thyroid,* 15(8), pp. 815-821.

Brusca, R., Findley, L., Hastings, P., Hendrickx, M., Cosio, J.T. and van der Heiden, A. (2005) 'Macrofaunal diversity in the Gulf of California', *Biodiversity, ecosystems, and conservation in Northern Mexico,* 179.

Brusick, D. (1980) 'Principles of genetic toxicology', 33.

Bryan, G., Bright, D., Hummerstone, L. and Burt, G. (1993) 'Uptake, tissue distribution and metabolism of 14C-labelled tributyltin (TBT) in the dog-whelk, *Nucella lapillus*', *Journal of the Marine Biological Association of the United Kingdom,* 73(04), pp. 889-912.

Burbach, J.P.H. (2011) 'What are neuropeptides?', in *Neuropeptides.* Springer, pp. 1-36.

Burch, J. (1960) 'Chromosome numbers of schistosome vector snails.', *Zeitschrift fur Tropenmedizin und Parasitologie,* 11(4), pp. 449-452.

Camey T and Verdok NH (1970) 'The early development of the snail *Biomphalaria glabrata*', *Biomphalaria glabrata,* , pp. 93-121.

Campbell, N. and Reece, J. (2005) 'Hormones and the endocrine system', *Biology, 7th edition, San Francisco, USA: Pearson, Benjamin Cummings.p,* , pp. 956-958.

Campbell, T.N. and Choy, F.Y. (2005) 'RNA interference: past, present and future', *Current Issues in Molecular Biology,* 7, pp. 1-6.

Caneparo, L., Huang, Y.L., Staudt, N., Tada, M., Ahrendt, R., Kazanskaya, O., Niehrs, C. and Houart, C. (2007) 'Dickkopf-1 regulates gastrulation movements by coordinated modulation of Wnt/beta catenin and Wnt/PCP activities, through interaction with the Dally-like homolog Knypek', *Genes & development,* 21(4), pp. 465-480.

Canesi, L., Betti, M., Ciacci, C., Lorusso, L., Pruzzo, C. and Gallo, G. (2006) 'Cell signalling in the immune response of mussel hemocytes', *Invertebrate Survival Journal,* 3, pp. 40-49.

Canesi, L., Lorusso, L.C., Ciacci, C., Betti, M., Zampini, M. and Gallo, G. (2004) 'Environmental estrogens can affect the function of mussel hemocytes through rapid modulation of kinase pathways', *General and comparative endocrinology,* 138(1), pp. 58-69.

Carneveli, M. (2006) 'Regeneration in Echinoderms: repair, regrowth, cloning', *Invertebrate Survival Journal,* 3, pp. 64-76.

Carney, G.E. and Bender, M. (2000) 'The *Drosophila* ecdysone receptor (EcR) gene is required maternally for normal oogenesis', *Genetics,* 154(3), pp. 1203-1211.

Carver, T., Berriman, M., Tivey, A., Patel, C., Bohme, U., Barrell, B.G., Parkhill, J. and Rajandream, M.A. (2008) 'Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database', *Bioinformatics (Oxford, England),* 24(23), pp. 2672-2676.

Castagna, M. and Kraeuter, J.N. (1994) 'Age, growth rate, sexual dimorphism and fecundity of knobbed whelk Busycon carica(Gmelin, 1791) in a western Mid-Atlantic lagoon system, Virginia', *Journal of Shellfish Research,* 13(2), pp. 581-585.

Castro, L.F.C., Lima, D., Machado, A., Melo, C., Hiromori, Y., Nishikawa, J., Nakanishi, T., Reis-Henriques, M. and Santos, M. (2007) 'Imposex induction is mediated through the Retinoid X Receptor signalling pathway in the neogastropod *Nucella lapillus*', *Aquatic Toxicology,* 85(1), pp. 57-66.

Chansela, P., Saitongdee, P., Stewart, P., Soonklang, N., Stewart, M., Suphamungmee, W., Poomtong, T. and Sobhon, P. (2008) 'Existence of APGWamide in the testis and its induction of spermiation in *Haliotis asinina* Linnaeus', *Aquaculture,* 279(1), pp. 142-149.

Chapman, M. (2007) 'Colonization of novel habitat: tests of generality of patterns in a diverse invertebrate assemblage', *Journal of experimental marine biology and ecology,* 348(1), pp. 97-110.

Chawla, A., Repa, J.J., Evans, R.M. and Mangelsdorf, D.J. (2001) 'Nuclear receptors and lipid physiology: opening the X-files', *Science (New York, N.Y.),* 294(5548), pp. 1866-1870.

Chen, L., Wang, R., Lee, Y., Liu, N., Chang, Y., Wu, C., Xie, S., Hung, Y. and Chang, C. (2008) 'Subfertility with defective folliculogenesis in female mice lacking testicular orphan nuclear receptor 4', *Molecular Endocrinology,* 22(4), pp. 858-867.

Chiang, J.Y. (2002) 'Bile acid regulation of gene expression: roles of nuclear hormone receptors', *Endocrine reviews,* 23(4), pp. 443-463.

Chiu, A.Y., Hunkapiller, M.W., Heller, E., Stuart, D.K., Hood, L.E. and Strumwasser, F. (1979) 'Purification and primary structure of the neuropeptide egg-laying hormone of *Aplysia*

californica', *Proceedings of the National Academy of Sciences of the United States of America,* 76(12), pp. 6656-6660.

Chung, A.C. and Cooney, A.J. (2003) 'The varied roles of nuclear receptors during vertebrate embryonic development', *Nuclear receptor signaling,* 1.

Ciocan, C.M., Cubero-Leon, E., Peck, M.R., Langston, W.J., Pope, N., Minier, C. and Rotchell, J.M. (2012) 'Intersex in *Scrobicularia plana*: Transcriptomic Analysis Reveals Novel Genes Involved in Endocrine Disruption', *Environmental science & technology,* 46(23), pp. 12936-12942.

Clamp, M., Cuff, J., Searle, S.M. and Barton, G.J. (2004) 'The Jalview Java alignment editor', *Bioinformatics (Oxford, England),* 20(3), pp. 426-427.

Cohen, P., Rholam, M. and Boussetta, H. (1995) '[9] Methods for identification of neuropeptide-processing pathways', *Methods in Neurosciences,* 23, pp. 155-194.

Collin, R. (1995) 'Sex, size, and position: a test of models predicting size at sex change in the protandrous gastropod Crepidula fornicata', *American Naturalist,* , pp. 815-831.

Collins, L.L., Lee, Y., Ting, H., Lin, W., Liu, N., Meshul, C.K., Uno, H., Bao, B., Chen, Y. and Chang, C. (2011) 'The roles of testicular nuclear receptor 4 (TR4) in male fertility-priapism and sexual behavior defects in TR4 knockout mice', *Reprod.Biol.Endocrinol,* 9(1), pp. 138.

Cong, R., Li, Q. and Kong, L. (2013) 'Polymorphism in the insulin-related peptide gene and its association with growth traits in the Pacific oyster Crassostrea gigas', *Biochemical systematics and ecology,* 46, pp. 36-43.

Conzelmann, M., Offenburger, S.L., Asadulina, A., Keller, T., Munch, T.A. and Jekely, G. (2011) 'Neuropeptides regulate swimming depth of Platynereis larvae', *Proceedings of the National Academy of Sciences of the United States of America,* 108(46), pp. E1174-83.

Cooper, G.M. and Hausman, R.E. (2000) *The cell.* Sinauer Associates Sunderland.

Cowin, P.A., Gold, E., Aleksova, J., O'Bryan, M.K., Foster, P.M., Scott, H.S. and Risbridger, G.P. (2010) 'Vinclozolin exposure in utero induces postpubertal prostatitis and reduces sperm production via a reversible hormone-regulated mechanism', *Endocrinology,* 151(2), pp. 783-792.

Cox, D.N., Chao, A. and Lin, H. (2000) 'Piwi Encodes a Nucleoplasmic Factor Whose Activity Modulates the Number and Division Rate of Germline Stem Cells', *Development (Cambridge, England),* 127(3), pp. 503-514.

Croll, R.P. and Van Minnen, J. (1992) 'Distribution of the peptide Ala- Pro- Gly- Trp- NH2 (APGWamide) in the nervous system and periphery of the snail *lymnaea stagnalis* as revealed by immunocytochemistry and in situ hybridization', *Journal of Comparative Neurology,* 324(4), pp. 567-574.

Croll, R.P. and Wang, C. (2007) 'Possible roles of sex steroids in the control of reproduction in bivalve molluscs', *Aquaculture,* 272(1), pp. 76-86.

Cropper, E.C., Brezina, V., Vilim, F.S., Harish, O., Price, D.A., Rosen, S., Kupfermann, I. and

Weiss, K.R. (1994) 'FRF peptides in the ARC neuromuscular system of *Aplysia*: purification and physiological actions', *Journal of neurophysiology,* 72(5), pp. 2181-2195.

Cruz, L.J., de Santos, V., Zafaralla, G.C., Ramilo, C.A., Zeikus, R., Gray, W.R. and Olivera, B.M. (1987a) 'Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from Conus geographus and Conus straitus venoms', *The Journal of biological chemistry,* 262(33), pp. 15821-15824.

Cummins, S.F., York, P.S., Hanna, P.J., Degnan, B.M. and Croll, R.P. (2009) 'Expression of prohormone convertase 2 and the generation of neuropeptides in the developing nervous system of the gastropod *Haliotis*', *Int J Dev Biol,* 53(7), pp. 1081-1088.

Cunningham, C., Hikima, J., Jenny, M.J., Chapman, R.W., Fang, G., Saski, C., Lundqvist, M.L., Wing, R.A., Cupit, P.M. and Gross, P.S. (2006) 'New resources for marine genomics: bacterial artificial chromosome libraries for the Eastern and Pacific oysters (*Crassostrea virginica* and *C. gigas*)', *Marine biotechnology,* 8(5), pp. 521-533.

Czech, P., Weber, K. and Dietrich, D.R. (2001) 'Effects of endocrine modulating substances on reproduction in the hermaphroditic snail *Lymnaea stagnalis*', *Aquatic Toxicology,* 53(2), pp. 103-114.

D'aniello, A., Di Cosmo, A., Di Cristo, C., Assisi, L., Botte, V. and Di Fiore, M.M. (1996) 'Occurrence of Sex Steroid Hormones and Their Binding Proteins in *Octopus vulgaris* Lam', *Biochemical and biophysical research communications,* 227(3), pp. 782-788.

Darlison, M.G. and Richter, D. (1999) 'Multiple genes for neuropeptides and their receptors: co-evolution and physiology', *Trends in neurosciences,* 22(2), pp. 81-88.

De Jong-Brink, M. (1973) 'The effects of desiccation and starvation upon the weight, histology and ultrastructure of the reproductive tract of *Biomphalaria glabrata*, intermediate host of *Schistosoma mansoni*', *Zeitschrift für Zellforschung und Mikroskopische Anatomie,* 136(2), pp. 229-262.

De Jong-Brink, M. (1969) 'Histochemical and electron microscope observations on the reproductive tract of *Biomphalaria glabrata* (Australorbis glabratus), intermediate host of *Schistosoma mansoni*', *Zeitschrift für Zellforschung und Mikroskopische Anatomie,* 102(4), pp. 507-542.

De Jong-Brink, M., Boer, H., Hommes, T. and Kodde, A. (1977) 'Spermatogenesis and the role of Sertoli cells in the freshwater snail *Biomphalaria glabrata*', *Cell and tissue research,* 181(1), pp. 37-58.

De Jong-Brink, M., de Wit, A., Kraal, G. and Boer, H.d. (1976) 'A light and electron microscope study on oogenesis in the freshwater pulmonate snail *Biomphalaria glabrata*', *Cell and tissue research,* 171(2), pp. 195-219.

De Jong-Brink, M., Hoek, R.M., Lageweg, W. and Smit, A.B. (1997) 'Schistosome Parasites Induce Physiological Changes in their Snail Hosts by Interfering with Two Regulatory Systems, the Internal Defense System and the Neuroendocrine System', in *Parasites and*

Pathogens. Springer, pp. 57-75.

De Jong-Brink, M., Ter Borg, J., Bergamin-Sassen, M.J. and Boer, H. (1979) 'Histology and histochemistry of the reproductive tract of the pulmonate snail, *Bulinus truncatus*, with observations on the effects of castration on its growth and histology', *International Journal of Invertebrate Reproduction,* 1(1), pp. 41-56.

de la Ossa Carretero, JA, del Pilar Ruso, Y., Casalduero, F.G. and Lizaso, J.S. (2008) 'Effect of sewage discharge in Spisula subtruncata (da Costa 1778) populations', *Archives of Environmental Contamination and Toxicology,* 54(2), pp. 226-235.

De Lange, R. and Van Minnen, J. (1998) 'Localization of the Neuropeptide APGWamide in Gastropod Molluscs by *in Situ* Hybridization and Immunocytochemistry', *General and comparative endocrinology,* 109(2), pp. 166-174.

De Longcamp, D., Lubet, P. and Drosdowsky, M. (1974) 'The *in vitro* biosynthesis of steroids by the gonad of the mussel (*Mytilus edulis*)', *General and comparative endocrinology,* 22(1), pp. 116-127.

De Loof, A. and Huybrechts, R. (1998) '"Insects do not have sex hormones": a myth?', *General and comparative endocrinology,* 111(3), pp. 245-260.

de Mendonça, R.L., Escrivá, H., Bouton, D., Laudet, V. and Pierce, R.J. (2000) 'Hormones and nuclear receptors in schistosome development', *Parasitology Today,* 16(6), pp. 233-240.

Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M. and Goodstein, D.M. (2002) 'The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins', *Science,* 298(5601), pp. 2157-2167.

DeJong, R.J., Morgan, J.A., Paraense, W.L., Pointier, J.P., Amarista, M., Ayeh-Kumi, P.F., Babiker, A., Barbosa, C.S., Bremond, P., Pedro Canese, A., de Souza, C.P., Dominguez, C., File, S., Gutierrez, A., Incani, R.N., Kawano, T., Kazibwe, F., Kpikpi, J., Lwambo, N.J., Mimpfoundi, R., Njiokou, F., Noel Poda, J., Sene, M., Velasquez, L.E., Yong, M., Adema, C.M., Hofkin, B.V., Mkoji, G.M. and Loker, E.S. (2001) 'Evolutionary relationships and biogeography of *Biomphalaria* (Gastropoda: Planorbidae) with implications regarding its role as host of the human bloodfluke, *Schistosoma mansoni*', *Molecular biology and evolution,* 18(12), pp. 2225-2239.

Devi, L. (1991) 'Consensus sequence for processing of peptide precursors at monobasic sites', *FEBS letters,* 280(2), pp. 189-194.

Di Cosmo, A., Di Cristo, C. and Paolucci, M. (2001) 'Sex steroid hormone fluctuations and morphological changes of the reproductive system of the female of *Octopus vulgaris* throughout the annual cycle', *Journal of Experimental Zoology,* 289(1), pp. 33-47.

Di Cristo, C., Paolucci, M., Iglesias, J., Sanchez, J. and Di Cosmo, A. (2002) 'Presence of two neuropeptides in the fusiform ganglion and reproductive ducts of *Octopus vulgaris*: FMRFamide and gonadotropin- releasing hormone (GnRH)', *Journal of Experimental Zoology,* 292(3), pp. 267-276.

Diamanti-Kandarakis, E., Bourguignon, J., Giudice, L.C., Hauser, R., Prins, G.S., Soto, A.M., Zoeller, R.T. and Gore, A.C. (2009) 'Endocrine-disrupting chemicals: an Endocrine Society scientific statement', *Endocrine reviews,* 30(4), pp. 293-342.

Dissous, C., Khayath, N., Vicogne, J. and Capron, M. (2006) 'Growth factor receptors in helminth parasites: signalling and host–parasite relationships', *FEBS letters,* 580(12), pp. 2968-2975.

Dogterom, G., Bohlken, S. and Geraerts, W. (1983) 'A rapid *in vivo* bioassay of the ovulation hormone of *Lymnaea Stagnalis*', *General and comparative endocrinology,* 50(3), pp. 476-482.

Dorner, G. (1978) 'Hormones, brain development and fundamental processes of life', *Hormones and Brain Development, Elsevier, Amsterdam,* , pp. 13-25.

Duft, M., Schmitt, C., Bachmann, J., Brandelik, C., Schulte-Oehlmann, U. and Oehlmann, J. (2007) 'Prosobranch snails as test organisms for the assessment of endocrine active chemicals—an overview and a guideline proposal for a reproduction test with the freshwater mudsnail *Potamopyrgus antipodarum*', *Ecotoxicology,* 16(1), pp. 169-182.

Duncan, C. (1960) 'The genital systems of the freshwater Basommatophora', *Proceedings of the Zoological Society of London.* Wiley Online Library, 339-356.

Dussault, I., Lin, M., Hollister, K., Fan, M., Termini, J., Sherman, M.A. and Forman, B.M. (2002) 'A structural model of the constitutive androstane receptor defines novel interactions that mediate ligand-independent activity', *Molecular and cellular biology,* 22(15), pp. 5270-5280.

Dutertre, S., Croker, D., Daly, N.L., Andersson, A., Muttenthaler, M., Lumsden, N.G., Craik, D.J., Alewood, P.F., Guillon, G. and Lewis, R.J. (2008) 'Conopressin-T from Conus tulipa reveals an antagonist switch in vasopressin-like peptides', *The Journal of biological chemistry,* 283(11), pp. 7100-7108.

Dyrløv Bendtsen, J., Nielsen, H., von Heijne, G. and Brunak, S. (2004) 'Improved prediction of signal peptides: SignalP 3.0', *Journal of Molecular Biology,* 340(4), pp. 783-795.

Ebberink, R.H., van Loenhout, H., Geraerts, W.P. and Joosse, J. (1985) 'Purification and amino acid sequence of the ovulation neurohormone of *Lymnaea stagnalis*', *Proceedings of the National Academy of Sciences of the United States of America,* 82(22), pp. 7767-7771.

Edgar, R.C. (2004) 'MUSCLE: multiple sequence alignment with high accuracy and high throughput', *Nucleic acids research,* 32(5), pp. 1792-1797.

Egonmwan, R.I. (2007) 'Gross Morphology and Ultrastructural Study of Albumen Gland of the Land Snail *Archachatina marginata* ovum (Pfeiffer)(Pulmonata: Achatinidae)', *Pakistan Journal of Biological Sciences,* 10(2), pp. 322-325.

Eick, G.N. and Thornton, J.W. (2011) 'Evolution of steroid receptors from an estrogen-sensitive ancestral receptor', *Molecular and cellular endocrinology,* 334(1), pp. 31-38.

Eilbeck, K., Lewis, S.E., Mungall, C.J., Yandell, M., Stein, L., Durbin, R. and Ashburner, M. (2005) 'The Sequence Ontology: a tool for the unification of genome annotations', *Genome biology,* 6(5), pp. R44.

Eipper, B.A., Stoffers, D.A. and Mains, R.E. (1992) 'The biosynthesis of neuropeptides: peptide alpha-amidation', *Annual Review of Neuroscience,* 15(1), pp. 57-85.

El Filali, Z., Minnen, J.v., Liu, W., Smit, A. and Li, K. (2006) 'Peptidomics Analysis of Neuropeptides Involved in Copulatory Behavior of the Mollusk *Lymnaea s tagnalis*', *Journal of proteome research,* 5(7), pp. 1611-1617.

Emanuelsson, O., von Heijne, G. and Schneider, G. (2001) 'Analysis and prediction of mitochondrial targeting peptides', *Methods in cell biology,* 65, pp. 175-187.

Enmark, E. and Gustafsson, J. (2001) 'Comparing nuclear receptors in worms, flies and humans', *Trends in pharmacological sciences,* 22(12), pp. 611-615.

Escobedo, G., Roberts, C.W., Carrero, J.C. and Morales-Montor, J. (2005) 'Parasite regulation by host hormones: an old mechanism of host exploitation?', *Trends in parasitology,* 21(12), pp. 588-593.

Escriva, H., Bertrand, S. and Laudet, V. (2004) 'The evolution of the nuclear receptor superfamily', *Essays Biochem,* 40, pp. 11-26.

Escriva, H., Delaunay, F. and Laudet, V. (2000) 'Ligand binding and nuclear hormone receptors evolution.', *Bioessays,* 22, pp. 717-727.

Eskiw, C.H., Cope, N.F., Clay, I., Schoenfelder, S., Nagano, T. and Fraser, P. (2010) 'Transcription factories and nuclear organization of the genome', *Cold Spring Harbor symposia on quantitative biology,* 75, pp. 501-506.

Eveland, L. and Haseeb, M. (2011) 'Laboratory rearing of *Biomphalaria glabrata* snails and maintenance of larval schistosomes in vivo and in vitro', in *Biomphalaria snails and larval trematodes.* Springer, pp. 33-55.

Fabbri, E. and Capuzzo, A. (2010) 'Cyclic AMP signaling in bivalve molluscs: an overview', *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology,* 313(4), pp. 179-200.

Fan, X., Croll, R.P., Wu, B., Fang, L., Shen, Q., Painter, S.D. and Nagle, G.T. (1997) 'Molecular cloning of a cDNA encoding the neuropeptides APGWamide and cerebral peptide 1: Localization of APGWamide- like immunoreactivity in the central nervous system and male reproductive organs of *Aplysia*', *Journal of Comparative Neurology,* 387(1), pp. 53-62.

Farrar, K. and Donnison, I.S. (2007) 'Construction and screening of BAC libraries made from Brachypodium genomic DNA', *Nature protocols,* 2(7), pp. 1661-1674.

Felsenstein, J. (1985) 'Confidence limits on phylogenies: an approach using the bootstrap', *Evolution,* , pp. 783-791.

Felsenstein, J. (1981) 'A likelihood approach to character weighting and what it tells us about parsimony and compatibility', *Biological Journal of the Linnean Society,* 16(3), pp. 183-196.

Felsenstein, J. (1978) 'Cases in which parsimony or compatibility methods will be positively misleading', *Systematic Biology,* 27(4), pp. 401-410.

Feng, Z.P., Zhang, Z., van Kesteren, R.E., Straub, V.A., van Nierop, P., Jin, K., Nejatbakhsh, N.,

Goldberg, J.I., Spencer, G.E., Yeoman, M.S., Wildering, W., Coorssen, J.R., Croll, R.P., Buck, L.T., Syed, N.I. and Smit, A.B. (2009) 'Transcriptome analysis of the central nervous system of the mollusc *Lymnaea stagnalis*', *BMC genomics,* 10, pp. 451-2164-10-451.

Fentress, J.A., Steele, S.L., Bart, H.L.,Jr and Cheek, A.O. (2006) 'Reproductive disruption in wild longear sunfish (Lepomis megalotis) exposed to kraft mill effluent', *Environmental health perspectives,* 114(1), pp. 40-45.

Ferguson, J.N., Young, L.J. and Insel, T.R. (2002) 'The neuroendocrine basis of social recognition', *Frontiers in neuroendocrinology,* 23(2), pp. 200-224.

Fernandes, D., Loi, B. and Porte, C. (2011a) 'Biosynthesis and metabolism of steroids in molluscs', *The Journal of steroid biochemistry and molecular biology,* 127(3), pp. 189-195.

Finley, K.D., Edeen, P.T., Foss, M., Gross, E., Ghbeish, N., Palmer, R.H., Taylor, B.J. and McKeown, M. (1998) '*Dissatisfaction* Encodes a Tailless-like Nuclear Receptor Expressed in a Subset of CNS Neurons Controlling *Drosophila* Sexual Behavior', *Neuron,* 21(6), pp. 1363-1374.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) 'Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*', *Nature,* 391(6669), pp. 806-811.

Fitts, D.A. (2011) 'Ethics and animal numbers: informal analyses, uncertain sample sizes, inefficient replications, and type I errors', *Journal of the American Association for Laboratory Animal Science : JAALAS,* 50(4), pp. 445-453.

Fleury, E., Huvet, A., Lelong, C., de Lorgeril, J., Boulo, V., Gueguen, Y., Bachere, E., Tanguy, A., Moraga, D., Fabioux, C., Lindeque, P., Shaw, J., Reinhardt, R., Prunet, P., Davey, G., Lapegue, S., Sauvage, C., Corporeau, C., Moal, J., Gavory, F., Wincker, P., Moreews, F., Klopp, C., Mathieu, M., Boudry, P. and Favrel, P. (2009) 'Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase', *BMC genomics,* 10, pp. 341-2164-10-341.

Florell, S.R., Coffin, C.M., Holden, J.A., Zimmermann, J.W., Gerwels, J.W., Summers, B.K., Jones, D.A. and Leachman, S.A. (2001) 'Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol', *Modern pathology,* 14(2), pp. 116-128.

Fong, P.P. and Ford, A.T. (2014) 'The biological effects of antidepressants on the molluscs and crustaceans: a review', *Aquatic Toxicology,* 151, pp. 4-13.

Forbes, V.E., Aufderheide, J., Warbritton, R., van der Hoeven, N. and Caspers, N. (2007a) 'Does bisphenol A induce superfeminization in *Marisa cornuarietis*? Part II: Toxicity test results and requirements for statistical power analyses', *Ecotoxicology and environmental safety,* 66(3), pp. 319-325.

Forbes, V.E., Selck, H., Palmqvist, A., Aufderheide, J., Warbritton, R., Pounds, N., Thompson, R., van der Hoeven, N. and Caspers, N. (2007b) 'Does bisphenol a induce superfeminization in

Marisa cornuarietis? Part I: Intra-and inter-laboratory variability in test endpoints', *Ecotoxicology and environmental safety,* 66(3), pp. 309-318.

Ford, E.S., Li, C., Zhao, G., Pearson, W.S. and Mokdad, A.H. (2008) 'Prevalence of the metabolic syndrome among US adolescents using the definition from the International Diabetes Federation', *Diabetes care,* 31(3), pp. 587-589.

Forman, B.M., Casanova, J., Raaka, B.M., Ghysdael, J. and Samuels, H.H. (1992) 'Half-site spacing and orientation determines whether thyroid hormone and retinoic acid receptors and related factors bind to DNA response elements as monomers, homodimers, or heterodimers', *Molecular endocrinology (Baltimore, Md.),* 6(3), pp. 429-442.

Förster, S., Günthel, D., Kiss, F. and Brehm, K. (2011) 'Molecular characterisation of a serum-responsive, DAF- 12- like nuclear hormone receptor of the fox- tapeworm Echinococcus multilocularis', *Journal of cellular biochemistry,* 112(6), pp. 1630-1642.

Fraser, P. and Bickmore, W. (2007) 'Nuclear organization of the genome and the potential for gene regulation', *Nature,* 447(7143), pp. 413-417.

Fretter, V. (1984) 'Prosobranchs', *The mollusca,* 7.

Fretter, V. and Shale, D. (1973) 'Seasonal changes in population density and vertical distribution of prosobranch veligers in offshore plankton at Plymouth', *Journal of the Marine Biological Association of the United Kingdom,* 53(03), pp. 471-492.

Fu, M., Sun, T., Bookout, A.L., Downes, M., Yu, R.T., Evans, R.M. and Mangelsdorf, D.J. (2005) 'A nuclear receptor atlas: 3T3-L1 adipogenesis', *Molecular endocrinology,* 19(10), pp. 2437-2450.

Fujioka, Y. (2002) 'Effects of hormone treatments and temperature on sex- reversal of Nigorobuna Carassius carassius grandoculis', *Fisheries Science,* 68(4), pp. 889-893.

Gagné, F., Blaise, C., Lachance, B., Sunahara, G. and Sabik, H. (2001) 'Evidence of coprostanol estrogenicity to the freshwater mussel *Elliptio complanata*', *Environmental Pollution,* 115(1), pp. 97-106.

Gagné, F., Blaise, C., Pellerin, J. and Gauthier-Clerc, S. (2002) 'Alteration of the biochemical properties of female gonads and vitellins in the clam *Mya arenaria* at contaminated sites in the Saguenay Fjord', *Marine environmental research,* 53(3), pp. 295-310.

Gagne, F. and Blaise, C. (2005) 'Review of biomarkers and new techniques for in situ aquatic studies with bivalves', *Environmental Toxicity Testing,* , pp. 408.

Gagnon, M.M., Dodson, J.J., Hodson, P.V., Kraak, G.V.D. and Carey, J.H. (1994) 'Seasonal effects of bleached kraft mill effluent on reproductive parameters of white sucker (Catostomus commersoni) populations of the St. Maurice River, Quebec, Canada', *Canadian Journal of Fisheries and Aquatic Sciences,* 51(2), pp. 337-347.

Gampe Jr, R.T., Montana, V.G., Lambert, M.H., Miller, A.B., Bledsoe, R.K., Milburn, M.V., Kliewer, S.A., Willson, T.M. and Xu, H.E. (2000) 'Asymmetry in the PPARγ/RXRα crystal structure reveals the molecular basis of heterodimerization among nuclear receptors',

Molecular cell, 5(3), pp. 545-555.

Ganter, G., Panaitiu, A., Desilets, J., Davis-Heim, J., Fisher, E., Tan, L., Heinrich, R., Buchanan, E., Brooks, K. and Kenney, M. (2011) '*Drosophila* male courtship behavior is modulated by ecdysteroids', *Journal of insect physiology,* 57(9), pp. 1179-1184.

Gao, W., Bohl, C.E. and Dalton, J.T. (2005) 'Chemistry and structural biology of androgen receptor', *Chemical reviews,* 105(9), pp. 3352-3370.

García, H.E., Laudet, V. and Robinson-Rechavi, M. (2003) 'Nuclear receptors are markers of animal genome evolution', in *Genome Evolution.* Springer, pp. 177-184.

Garcia, M., Griffond, B. and Lafont, R. (1995) 'What are the origins of ecdysteroids in gastropods?', *General and comparative endocrinology,* 97(1), pp. 76-85.

Garden, R.W., Shippy, S.A., Li, L., Moroz, T.P. and Sweedler, J.V. (1998) 'Proteolytic processing of the *Aplysia* egg-laying hormone prohormone', *Proceedings of the National Academy of Sciences of the United States of America,* 95(7), pp. 3972-3977.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E. and Bowman, S. (2002) 'Genome sequence of the human malaria parasite Plasmodium falciparum', *Nature,* 419(6906), pp. 498-511.

Garrison, J.L., Macosko, E.Z., Bernstein, S., Pokala, N., Albrecht, D.R. and Bargmann, C.I. (2012) 'Oxytocin/vasopressin-related peptides have an ancient role in reproductive behavior', *Science,* 338(6106), pp. 540-543.

George, F.W. (1997) 'Androgen Metabolism in the Prostate of the Finasteride-Treated, Adult Rat: A Possible Explanation for the Differential Action of Testosterone and 5α-Dihydrotestosterone during Development of the Male Urogenital Tract 1', *Endocrinology,* 138(3), pp. 871-877.

Geraerts, W.M. (1976a) 'The control of ovulation in the hermaphroditic freshwater snail *Lymnaea stagnalis* by the neurohormone of the caudodorsal cells', *General and comparative endocrinology,* 28(3), pp. 350-357.

Geraerts, W. (1992) 'Neurohormonal control of growth and carbohydrate metabolism by the light green cells in *Lymnaea stagnalis*', *General and comparative endocrinology,* 86(3), pp. 433-444.

Geraerts, W. (1976b) 'The role of the Lateral Lobes in the control of growth and reproduction in the hermaphrodite freshwater snail *Lymnaea stagnalis*', *General and comparative endocrinology,* 29(1), pp. 97-108.

Geraerts, WPM and Joosse, J (1975) 'Control of vitellogenesis and of growth of female accessory sex organs by the dorsal body hormone (DBH) in the hermaphroditic freshwater snail*Lymnaea stagnalis.*', *Gen Comp Endocrinol,* 27, pp. 450-464.

Geraerts, W. and Algera, L. (1976) 'The stimulating effect of the Dorsal-Body Hormone on cell differentiation in the female accessory sex organs of the hermaphrodite freshwater snail, *Lymnaea stagnalis*', *General and comparative endocrinology,* 29(1), pp. 109-118.

Geraerts, W., Joosse, J., Saleuddin, A. and Wilbur, K. (1984) 'Freshwater snails (Basommatophora)', *The mollusca,* 7.

Geraerts, W. and Mohamed, A. (1981) 'Studies on the role of the lateral lobes and the ovotestis of the pulmonate snail *Bulinus truncatus* in the control of body growth and reproduction', *International Journal of Invertebrate Reproduction,* 3(5), pp. 297-308.

Geraudie, P., Hinfray, N., Gerbron, M., Porcher, J., Brion, F. and Minier, C. (2011) 'Brain cytochrome P450 aromatase activity in roach (*Rutilus rutilus*): Seasonal variations and impact of environmental contaminants', *Aquatic Toxicology,* 105(3), pp. 378-384.

Germain, P., Staels, B., Dacquet, C., Spedding, M. and Laudet, V. (2006a) 'Overview of nomenclature of nuclear receptors', *Pharmacological reviews,* 58(4), pp. 685-704.

Ghosh- Roy, A. and Chisholm, A.D. (2010) '*Caenorhabditis elegans*: a new model organism for studies of axon regeneration', *Developmental Dynamics,* 239(5), pp. 1460-1464.

Giguère, V., Hollenberg, S.M., Rosenfeld, M.G. and Evans, R.M. (1986) 'Functional domains of the human glucocorticoid receptor', *Cell,* 46(5), pp. 645-652.

Giusti, A., Barsi, A., Dugué, M., Collinet, M., Thomé, J., Joaquim- Justo, C., Roig, B., Lagadic, L. and Ducrot, V. (2013) 'Reproductive impacts of tributyltin (TBT) and triphenyltin (TPT) in the hermaphroditic freshwater gastropod *Lymnaea stagnalis*', *Environmental Toxicology and Chemistry,* 32(7), pp. 1552-1560.

Giusti, A., Lagadic, L., Barsi, A., Thomé, J., Joaquim-Justo, C. and Ducrot, V. (2014) 'Investigating apical adverse effects of four endocrine active substances in the freshwater gastropod *Lymnaea stagnalis*', *Science of The Total Environment,* 493, pp. 147-155.

Giusti, R.M., Iwamoto, K. and Hatch, E.E. (1995) 'Diethylstilbestrol revisited: a review of the long-term health effects', *Annals of Internal Medicine,* 122(10), pp. 778-788.

Glanzman, D. (2006) 'The cellular mechanisms of learning in *Aplysia*: Of blind men and elephants', *Biological Bulletin,* 210(3), pp. 271-279.

Glass, C.K. (1994) 'Differential Recognition of Target Genes by Nuclear Receptor Monomers, Dimers, and Heterodimers*', *Endocrine reviews,* 15(3), pp. 391-407.

Glass, C.K. and Rosenfeld, M.G. (2000) 'The coregulator exchange in transcriptional functions of nuclear receptors', *Genes & development,* 14(2), pp. 121-141.

Goldberg, J.I., Garofalo, R., Price, C.J. and Chang, J.P. (1993) 'Presence and biological activity of a GnRH- like factor in the nervous system of *Helisoma trivoluis*', *Journal of Comparative Neurology,* 336(4), pp. 571-582.

Golden, S.H., Robinson, K.A., Saldanha, I., Anton, B. and Ladenson, P.W. (2009) 'Prevalence and incidence of endocrine and metabolic disorders in the United States: a comprehensive review', *Journal of Clinical Endocrinology & Metabolism,* 94(6), pp. 1853-1878.

Gooding, M.P., Wilson, V.S., Folmar, L.C., Marcovich, D.T. and LeBlanc, G.A. (2003) 'The biocide tributyltin reduces the accumulation of testosterone as fatty acid esters in the mud snail (*Ilyanassa obsoleta*)', *Environmental health perspectives,* 111(4), pp. 426-430.

Goodson, J.L. and Bass, A.H. (2001) 'Social behavior functions and related anatomical characteristics of vasotocin/vasopressin systems in vertebrates', *Brain Research Reviews,* 35(3), pp. 246-265.

Gorbman, A. and Sower, S.A. (2003a) 'Evolution of the role of GnRH in animal (Metazoan) biology', *General and comparative endocrinology,* 134(3), pp. 207-213.

Goss, L.B. and Sabourin, T.D. (1985) 'Utilization of alternative species for toxicity testing: an overview', *Journal of Applied Toxicology,* 5(4), pp. 193-219.

Goto, Y., Kajiwara, M., Yanagisawa, Y., Hirose, H., Yoshimi, T., Umemura, M., Nakano, H., Takahashi, S., Shida, Y. and Iguchi, T. (2012) 'Detection of vertebrate-type steroid hormones and their converting activities in the neogastropod *Thais clavigera* (Küster, 1858)', *Journal of Molluscan Studies,* 78(2), pp. 197-204.

Grandjean, P., Bellinger, D., Bergman, Å, Cordier, S., Davey- Smith, G., Eskenazi, B., Gee, D., Gray, K., Hanson, M. and Van Den Hazel, P. (2008) 'The Faroes statement: human health effects of developmental exposure to chemicals in our environment', *Basic & Clinical Pharmacology & Toxicology,* 102(2), pp. 73-75.

Gray, L.E., Furr, J. and Ostby, J.S. (2005) 'Hershberger Assay to Investigate the Effects of Endocrine- Disrupting Compounds with Androgenic or Antiandrogenic Activity in Castrate-Immature Male Rats', *Current Protocols in Toxicology,* , pp. 16.9. 1-16.9. 15.

Greenberg, M.J., Doble, K., Lesser, W., Lee, T., Pennell, N., Morgan, C. and Price, D. (1997) 'Characterization of myomodulin-related peptides from the pulmonate snail Helix aspersa', *Peptides,* 18(8), pp. 1099-1106.

Greenberg, M. and Price, D. (1992) 'Relationships among the FMRFamide-like peptides', *Progress in brain research,* 92, pp. 25-37.

Greenspan, R.J. (2005) 'No critter left behind: an invertebrate renaissance', *Current Biology,* 15(17), pp. R671-R672.

Greenspan, R.J. (2001) 'The flexible genome', *Nature Reviews Genetics,* 2(5), pp. 383-387.

Gricourt, L., Bonnec, G., Boujard, D., Mathieu, M. and Kellner, K. (2003) 'Insulin-like system and growth regulation in the Pacific oyster *Crassostrea gigas*: hrIGF-1 effect on protein synthesis of mantle edge cells and expression of an homologous insulin receptor-related receptor', *General and comparative endocrinology,* 134(1), pp. 44-56.

Gronemeyer, H., Gustafsson, J. and Laudet, V. (2004) 'Principles for modulation of the nuclear receptor superfamily', *Nature Reviews Drug Discovery,* 3(11), pp. 950-964.

Gruber, C.W. (2014a) 'Physiology of invertebrate oxytocin and vasopressin neuropeptides', *Experimental physiology,* 99(1), pp. 55-61.

Grun, F. and Blumberg, B. (2006) 'Environmental obesogens: organotins and endocrine disruption via nuclear receptor signaling', *Endocrinology,* 147(6), pp. s50-s55.

Gruner H (ed.)(1993) *Mollusca,
Sipunculida, Echgiurida, Annelida, Onychophora, Tardigrada, Pentastomida.* 5th edn. Gustav Fischer Verlag, Jena, Stuttgart, New York:

Lehrbuch der Speziellen Zoologie.

Guhad, F. (2005) 'Introduction to the 3Rs (refinement, reduction and replacement)', *Journal of the American Association for Laboratory Animal Science,* 44(2), pp. 58-59.

Guindon, S. and Gascuel, O. (2003) 'A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood', *Systematic Biology,* 52(5), pp. 696-704.

Guo D, C.D. (2003) 'Isolation of intact RNA following cryosection of archived frozen tissue.', *Biotechniques,* 34, pp. 48-50.

Gupta, S.C., Siddique, H.R., Mathur, N., Mishra, R.K., Saxena, D.K. and Chowdhuri, D.K. (2007) 'Adverse effect of organophosphate compounds, dichlorvos and chlorpyrifos in the reproductive tissues of transgenic *Drosophila melanogaster*: 70kDa heat shock protein as a marker of cellular damage', *Toxicology,* 238(1), pp. 1-14.

Guthrie, C. and Patterson, B. (1988) 'Spliceosomal snRNAs', *Annual Review of Genetics,* 22(1), pp. 387-419.

Hadley, M. (1992) *Endocrinology.* New Jersy, USA: Pentice Hall, Inc. Endocrinology.

Hagerman, D.D., Wellington, F.M. and Villee, C.A. (1957) 'Estrogens in marine invertebrates', *Biological Bulletin,* , pp. 180-183.

Hagger, J.A., Depledge, M.H., Oehlmann, J., Jobling, S. and Galloway, T.S. (2006) 'Is there a causal association between genotoxicity and the imposex effect?', *Environmental health perspectives,* 114(S-1), pp. 20.

Hagger, J.A., Galloway, T.S., Langston, W.J. and Jones, M.B. (2009) 'Application of biomarkers to assess the condition of European Marine Sites', *Environmental Pollution,* 157(7), pp. 2003-2010.

Hall, J.D. and Lloyd, P.E. (1990) 'Involvement of pedal peptide in locomotion in Aplysia: modulation of foot muscle contractions', *Journal of neurobiology,* 21(6), pp. 858-868.

Hallgren, P., Sorita, Z., Berglund, O. and Persson, A. (2012) 'Effects of 17α-ethinylestradiol on individual life-history parameters and estimated population growth rates of the freshwater gastropods *Radix balthica* and *Bithynia tentaculata*', *Ecotoxicology,* 21(3), pp. 803-810.

Hamano, K., Awaji, M. and Usuki, H. (2005) 'cDNA structure of an insulin-related peptide in the *Pacific oyster* and seasonal changes in the gene expression', *The Journal of endocrinology,* 187(1), pp. 55-67.

Harrington, A.J., Hamamichi, S., Caldwell, G.A. and Caldwell, K.A. (2010) '*C. elegans* as a model organism to investigate molecular pathways involved with Parkinson's disease', *Developmental Dynamics,* 239(5), pp. 1282-1295.

Harris, C.A., Hamilton, P.B., Runnalls, T.J., Vinciotti, V., Henshaw, A., Hodgson, D., Coe, T.S., Jobling, S., Tyler, C.R. and Sumpter, J.P. (2011) 'The consequences of feminization in breeding groups of wild fish', *Environmental health perspectives,* 119(3), pp. 306-311.

Harry, H. (1965) 'Evidence of a gonadal hormone controlling the development of the accessory reproductive organs in Taphius glabratus (Say)(Gastropoda, Basommatophora)', *Transactions*

of the American Microscopical Society, 84, pp. 157.

Hartenstein, V. (2006) 'The neuroendocrine system of invertebrates: a developmental and evolutionary perspective', *The Journal of endocrinology,* 190(3), pp. 555-570.

Hasse, B., Ehrenberg, H., Marxen, J.C., Becker, W. and Epple, M. (2000) 'Calcium carbonate modifications in the mineralized shell of the freshwater snail *Biomphalaria glabrata*', *Chemistry-A European Journal,* 6(20), pp. 3679-3685.

Haszprunar, G. and Wanninger, A. (2012a) 'Molluscs', *Current Biology,* 22(13), pp. R510-R514.

Hatfield, J.S. (2005) 'Gastropods as an Animal Model for Studying Autism and Other Behavioral Phenotypes', *Causes and risks for autism, ,* pp. 51.

Hecker, M. and Hollert, H. (2011) 'Endocrine disruptor screening: regulatory perspectives and needs', *Environmental Sciences Europe,* 23(1), pp. 1-14.

Heindel, J.J. (2006) 'Role of exposure to environmental chemicals in the developmental basis of reproductive disease and dysfunction', *Seminars in reproductive medicine.* Published in 2006 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA., 168-177.

Heinlein, C.A. and Chang, C. (2002) 'The roles of androgen receptors and androgen-binding proteins in nongenomic androgen actions', *Molecular Endocrinology,* 16(10), pp. 2181-2187.

Heller, J. (1993) 'Hermaphroditism in molluscs', *Biological Journal of the Linnean Society,* 48(1), pp. 19-42.

Heller, M.J. (2002) 'DNA microarray technology: devices, systems, and applications', *Annual Review of Biomedical Engineering,* 4(1), pp. 129-153.

Henry, J., Zatylny, C. and Boucaud-Camou, E. (1999) 'Peptidergic control of egg-laying in the cephalopod *Sepia officinalis*: involvement of FMRFamide and FMRFamide-related peptides', *Peptides,* 20(9), pp. 1061-1070.

Henry, J., Favrel, P. and Boucaud-Camou, E. (1997) 'Isolation and Identification of a Novel Ala-Pro-Gly-Trp-amide-Related Peptide Inhibiting the Motility of the Mature Oviduct in the Cuttlefish, *Sepia officinalis*', *Peptides,* 18(10), pp. 1469-1474.

Henry, J., Zatylny, C. and Favrel, P. (2000) 'HPLC and electrospray ionization mass spectrometry as tools for the identification of APGWamide-related peptides in gastropod and bivalve mollusks: comparative activities on *Mytilus* muscles', *Brain research,* 862(1), pp. 162-170.

Herbst, A.L., Ulfelder, H. and Poskanzer, D.C. (1971) 'Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumor appearance in young women', *New England journal of medicine,* 284(16), pp. 878-881.

Hershberger, L., Shipley, E.G. and Meyer, R.K. (1953) 'Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method.', *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY).* Royal Society of Medicine, 175-180.

Hertel, L.A., Adema, C.M. and Loker, E.S. (2005) 'Differential expression of *FREP* genes in two

strains of *Biomphalaria glabrata* following exposure to the digenetic trematodes *Schistosoma mansoni* and *Echinostoma paraensei*', *Dev. Comp. Immunol.,* 29(4), pp. 295-303.

Hewitt, L.M., Smyth, S.A.M., Dubé, M.G., Gilman, C.I. and MacLatchy, D.L. (2002) 'Isolation of compounds from bleached kraft mill recovery condensates associated with reduced levels of testosterone in mummichog (*Fundulus heteroclitus*)', *Environmental toxicology and chemistry,* 21(7), pp. 1359-1367.

Hinck, J.E., Blazer, V.S., Schmitt, C.J., Papoulias, D.M. and Tillitt, D.E. (2009) 'Widespread occurrence of intersex in black basses (*Micropterus* spp.) from US rivers, 1995–2004', *Aquatic Toxicology,* 95(1), pp. 60-70.

Hirata, T., Kubota, I., Imada, M., Muneoka, Y. and Kobayashi, M. (1989) 'Effects of the catch-relaxing peptide on molluscan muscles', *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology,* 92(2), pp. 283-288.

Hoek, R., Li, K., Van Minnen, J., Lodder, J., Jong- Brink, D., Smit, A. and Van Kesteren, R. (2005) 'LFRFamides: a novel family of parasitation- induced- RFamide neuropeptides that inhibit the activity of neuroendocrine cells in *Lymnaea stagnalis*', *Journal of neurochemistry,* 92(5), pp. 1073-1080.

Hoffman, S.L., Subramanian, G.M., Collins, F.H. and Venter, J.C. (2002) 'Plasmodium, human and Anopheles genomics and malaria', *Nature,* 415(6872), pp. 702-709.

Hökfelt, T., Broberger, C., Xu, Z.D., Sergeyev, V., Ubink, R. and Diez, M. (2000) 'Neuropeptides—an overview', *Neuropharmacology,* 39(8), pp. 1337-1356.

Holbeck, S., Chang, J., Best, A.M., Bookout, A.L., Mangelsdorf, D.J. and Martinez, E.D. (2010) 'Expression profiling of nuclear receptors in the NCI60 cancer cell panel reveals receptor-drug and receptor-gene interactions', *Molecular endocrinology,* 24(6), pp. 1287-1296.

Holder, M. and Lewis, P.O. (2003) 'Phylogeny estimation: traditional and Bayesian approaches', *Nature reviews genetics,* 4(4), pp. 275-284.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M., Wides, R., Salzberg, S.L., Loftus, B. and Yandell, M.e.a. (2002) 'The genome sequence of the malaria mosquito *Anopheles gambiae*', *Science (New York, N.Y.),* 298(5591), pp. 129-149.

Hoopes, L. (2008) 'Genetic diagnosis: DNA microarrays and cancer', *Nature Education,* 1(1), pp. 3.

Hoover, R.N., Hyer, M., Pfeiffer, R.M., Adam, E., Bond, B., Cheville, A.L., Colton, T., Hartge, P., Hatch, E.E. and Herbst, A.L. (2011) 'Adverse health outcomes in women exposed in utero to diethylstilbestrol', *New England Journal of Medicine,* 365(14), pp. 1304-1314.

Hordijk, P.L., de Jong-Brink, M., Ter Maat, A., Pieneman, A., Lodder, J. and Kits, K. (1992) 'The neuropeptide schistosomin and haemolymph from parasitized snails induce similar changes in excitability in neuroendocrine cells controlling reproduction and growth in a freshwater snail',

Neuroscience letters, 136(2), pp. 193-197.

Hordijk, P.L., van Loenhout, H., Ebberink, R.H., de Jong- Brink, M. and Joosse, J. (1991) 'Neuropeptide schistosomin inhibits hormonally- induced ovulation in the freshwater snail *Lymnaea stagnalis*', *Journal of Experimental Zoology,* 259(2), pp. 268-271.

Horiguchi, T., Nishikawa, T., Ohta, Y., Shiraishi, H. and Morita, M. (2007) 'Retinoid X receptor gene expression and protein content in tissues of the rock shell *Thais clavigera*', *Aquatic toxicology,* 84(3), pp. 379-388.

Horner, M.A., Pardee, K., Liu, S., King-Jones, K., Lajoie, G., Edwards, A., Krause, H.M. and Thummel, C.S. (2009) 'The Drosophila DHR96 nuclear receptor binds cholesterol and regulates cholesterol homeostasis', *Genes & development,* 23(23), pp. 2711-2716.

Hornung, M.W., Jensen, K.M., Korte, J.J., Kahl, M.D., Durhan, E.J., Denny, J.S., Henry, T.R. and Ankley, G.T. (2004) 'Mechanistic basis for estrogenic effects in fathead minnow (*Pimephales promelas*) following exposure to the androgen 17α-methyltestosterone: conversion of 17α-methyltestosterone to 17α-methylestradiol', *Aquatic toxicology,* 66(1), pp. 15-23.

Howard-Ashby, M., Materna, S.C., Brown, C.T., Chen, L., Cameron, R.A. and Davidson, E.H. (2006) 'Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*', *Developmental biology,* 300(1), pp. 90-107.

Hoyle, C.H. (1998) 'Neuropeptide families: evolutionary perspectives', *Regulatory peptides,* 73(1), pp. 1-33.

Huang, J., Powell, W.C., Khodavirdi, A.C., Wu, J., Makita, T., Cardiff, R.D., Cohen, M.B., Sucov, H.M. and Roy-Burman, P. (2002) 'Prostatic intraepithelial neoplasia in mice with conditional disruption of the retinoid X receptor alpha allele in the prostate epithelium', *Cancer research,* 62(16), pp. 4812-4819.

Hummon, A.B., Hummon, N.P., Corbin, R.W., Li, L., Vilim, F.S., Weiss, K.R. and Sweedler, J.V. (2003) 'From precursor to final peptides: a statistical sequence-based approach to predicting prohormone processing', *Journal of proteome research,* 2(6), pp. 650-656.

Hunter, G.A. and Donaldson, E.M. (1983) 'Hormonal sex control and its application to fish culture', *Fish physiology,* 9(Part B), pp. 223-303.

Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R.D., Gough, J., Haft, D. and Hulo, N., et al. (2009) 'InterPro: the integrative protein signature database', *Nucleic acids research,* 37(Database issue), pp. D211-5.

Hutchinson, T., Shillabeer, N., Winter, M. and Pickford, D. (2006) 'Acute and chronic effects of carrier solvents in aquatic organisms: a critical review', *Aquatic Toxicology,* 76(1), pp. 69-92.

Iguchi, T., Watanabe, H., Ohta, Y. and Blumberg, B. (2008) 'Developmental effects: oestrogen-induced vaginal changes and organotin- induced adipogenesis', *International journal of andrology,* 31(2), pp. 263-268.

Iguchi, T., Katsu, Y., Horiguchi, T., Watanabe, H., Blumberg, B. and Ohta, Y. (2007) 'Endocrine disrupting organotin compounds are potent inducers of imposex in gastropods and adipogenesis in vertebrates', *Mol Cell Toxicol,* 3, pp. 1-10.

Ittiprasert, W. and Knight, M. (2012) 'Reversing the resistance phenotype of the *Biomphalaria glabrata* snail host *Schistosoma mansoni* infection by temperature modulation', *PLoS pathogens,* 8(4), pp. e1002677.

Iwakoshi- Ukena, E., Ukena, K., Takuwa- Kuroda, K., Kanda, A., Tsutsui, K. and Minakata, H. (2004) 'Expression and distribution of octopus gonadotropin- releasing hormone in the central nervous system and peripheral organs of the octopus (*Octopus vulgaris*) by in situ hybridization and immunohistochemistry', *Journal of Comparative Neurology,* 477(3), pp. 310-323.

Jackson, I.J. (1991) 'A reappraisal of non-consensus mRNA splice sites', *Nucleic acids research,* 19(14), pp. 3795-3798.

Janer, G., Bachmann, J., Oehlmann, J., Schulte-Oehlmann, U. and Porte, C. (2006a) 'The effect of organotin compounds on gender specific androstenedione metabolism in the freshwater ramshorn snail *Marisa cornuarietis*', *The Journal of steroid biochemistry and molecular biology,* 99(2), pp. 147-156.

Janer, G., Lavado, R., Thibaut, R. and Porte, C. (2005) 'Effects of 17β-estradiol exposure in the mussel *Mytilus galloprovincialis*: A possible regulating role for steroid acyltransferases', *Aquatic toxicology,* 75(1), pp. 32-42.

Janer, G., LeBlanc, G. and Porte, C. (2005a) 'A comparative study on androgen metabolism in three invertebrate species', *General and comparative endocrinology,* 143(3), pp. 211-221.

Janer, G., Mesia-Vela, S., Porte, C. and Kauffman, F.C. (2004) 'Esterification of vertebrate-type steroids in the Eastern oyster (*Crassostrea virginica*)', *Steroids,* 69(2), pp. 129-136.

Janer, G. and Porte, C. (2007) 'Sex steroids and potential mechanisms of non-genomic endocrine disruption in invertebrates', *Ecotoxicology,* 16(1), pp. 145-160.

Janer, G., Sternberg, R., LeBlanc, G. and Porte, C. (2005b) 'Testosterone conjugating activities in invertebrates: are they targets for endocrine disruptors?', *Aquatic toxicology,* 71(3), pp. 273-282.

Jarne, P., Vianey-Liaud, M. and Delay, B. (1993) 'Selfing and outcrossing in hermaphrodite freshwater gastropods (Basommatophora): where, when and why', *Biological Journal of the Linnean Society,* 49(2), pp. 99-125.

Jékely, G., Colombelli, J., Hausen, H., Guy, K., Stelzer, E., Nédélec, F. and Arendt, D. (2008) 'Mechanism of phototaxis in marine zooplankton', *Nature,* 456(7220), pp. 395-399.

Jiang, Y., Loker, E.S. and Zhang, S. (2006a) 'In vivo and in vitro knockdown of *FREP2* gene expression in the snail *Biomphalaria glabrata* using RNA interference', *Developmental & Comparative Immunology,* 30(10), pp. 855-866.

Jobling, S., Casey, D., Rodgers-Gray, T., Oehlmann, J., Schulte-Oehlmann, U., Pawlowski, S.,

Baunbeck, T., Turner, A. and Tyler, C. (2003) 'Comparative responses of molluscs and fish to environmental estrogens and an estrogenic effluent', *Aquatic Toxicology,* 65(2), pp. 205-220.

Jobling, S., Burn, R.W., Thorpe, K., Williams, R. and Tyler, C. (2009) 'Statistical modeling suggests that antiandrogens in effluents from wastewater treatment works contribute to widespread sexual disruption in fish living in English rivers', *Environmental health perspectives,* 117(5), pp. 797-802.

Jobling, S., Nolan, M., Tyler, C.R., Brighty, G. and Sumpter, J.P. (1998) 'Widespread sexual disruption in wild fish', *Environmental science & technology,* 32(17), pp. 2498-2506.

Jobling, S., Williams, R., Johnson, A., Taylor, A., Gross-Sorokin, M., Nolan, M., Tyler, C.R., Aerle, R., Santos, E. and Brighty, G. (2006) 'Predicted exposures to steroid estrogens in UK rivers correlate with widespread sexual disruption in wild fish populations', *Environmental health perspectives,* 114, pp. 1.

Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992b) 'The Rapid Generation of Mutation Data Matrices from Protein Sequences', *Comput. Appl. Biosci.,* 8(3).

Jong-Brink, M.d., Schot, L., Schoenmakers, H. and Bergamin-Sassen, M.J. (1981) 'A biochemical and quantitative electron microscope study on steroidogenesis in ovotestis and digestive gland of the pulmonate snail *Lymnaea stagnalis*', *General and comparative endocrinology,* 45(1), pp. 30-38.

Juliano, C., Wang, J. and Lin, H. (2011) 'Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms', *Annual Review of Genetics,* 45, pp. 447-469.

Jurberg, P., Cunha, R.A. and Rodrigues, M.L. (1997) 'Behavior of *Biomphalaria glabrata* Say, 1818 (Gastropoda: Planorbidae): I. Morphophysiology of the Mantle Cavity', *Memórias do Instituto Oswaldo Cruz,* 92, pp. 287-295.

Kajiwara, M., Kuraku, S., Kurokawa, T., Kato, K., Toda, S., Hirose, H., Takahashi, S., Shibata, Y., Iguchi, T. and Matsumoto, T. (2006) 'Tissue preferential expression of estrogen receptor gene in the marine snail, *Thais clavigera*', *General and comparative endocrinology,* 148(3), pp. 315-326.

Kaletta, T. and Hengartner, M.O. (2006) 'Finding function in novel targets: C. elegans as a model organism', *Nature Reviews Drug Discovery,* 5(5), pp. 387-399.

Kanayama, T., Kobayashi, N., Mamiya, S., Nakanishi, T. and Nishikawa, J. (2005) 'Organotin compounds promote adipocyte differentiation as agonists of the peroxisome proliferator-activated receptor gamma/retinoid X receptor pathway', *Molecular pharmacology,* 67(3), pp. 766-774.

Kanda, A., Takuwa-Kuroda, K., Iwakoshi-Ukena, E., Furukawa, Y., Matsushima, O. and Minakata, H. (2003) 'Cloning of *Octopus* cephalotocin receptor, a member of the oxytocin/vasopressin superfamily', *The Journal of endocrinology,* 179(2), pp. 281-291.

Kang, I.J., Yokota, H., Oshima, Y., Tsuruda, Y., Shimasaki, Y. and Honjo, T. (2008) 'The effects

of methyltestosterone on the sexual development and reproduction of adult medaka (*Oryzias latipes*)', *Aquatic toxicology,* 87(1), pp. 37-46.

Katsiadaki, I., Morris, S., Squires, C., Hurst, M.R., James, J.D. and Scott, A.P. (2006) 'Use of the three-spined stickleback (Gasterosteus aculeatus) as a sensitive in vivo test for detection of environmental antiandrogens', *Environmental health perspectives,* 114, pp. 115.

Katsu, Y., Kubokawa, K., Urushitani, H. and Iguchi, T. (2010) 'Estrogen-dependent transactivation of amphioxus steroid hormone receptor via both estrogen and androgen response elements', *Endocrinology,* 151(2), pp. 639-648.

Keay, J., Bridgham, J.T. and Thornton, J.W. (2006) 'The *Octopus vulgaris* estrogen receptor is a constitutive transcriptional activator: evolutionary and functional implications', *Endocrinology,* 147(8), pp. 3861-3869.

Keay, J. and Thornton, J.W. (2009) 'Hormone-activated estrogen receptors in annelid invertebrates: implications for evolution and endocrine disruption', *Endocrinology,* 150(4), pp. 1731-1738.

Kellett, E., Perry, S.J., Santama, N., Worster, B.M., Benjamin, P.R. and Burke, J.F. (1996a) 'Myomodulin gene of Lymnaea: structure, expression, and analysis of neuropeptides', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 16(16), pp. 4949-4957.

Kerlin, S.P. (2005) 'Prenatal exposure to diethylstilbestrol (DES) in males and gender-related disorders: results from a 5-year study', *International Behavioral Development Symposium.*

Kester, M.H., Bulduk, S., van Toor, H., Tibboel, D., Meinl, W., Glatt, H., Falany, C.N., Coughtrie, M.W., Schuur, A.G. and Brouwer, A. (2002) 'Potent inhibition of estrogen sulfotransferase by hydroxylated metabolites of polyhalogenated aromatic hydrocarbons reveals alternative mechanism for estrogenic activity of endocrine disrupters', *The Journal of Clinical Endocrinology & Metabolism,* 87(3), pp. 1142-1150.

Ketata, I., Denier, X., Hamza-Chaffai, A. and Minier, C. (2008) 'Endocrine-related reproductive effects in molluscs', *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology,* 147(3), pp. 261-270.

Khorasanizadeh, S. and Rastinejad, F. (2001) 'Nuclear-receptor interactions on DNA-response elements', *Trends in biochemical sciences,* 26(6), pp. 384-390.

Khotimchenko YS, D.I. (1989) '
The effect of dopamine and galoperidol on cyclic AMP in the gonad of the bivalve mollusc *Mizuhopecten yessoensis* and the sea urchin *Strogylocentrotus intermedius*. ', *Comp Biochem Physiol,* 92C, pp. 23–26.

Kidd, K.A., Blanchfield, P.J., Mills, K.H., Palace, V.P., Evans, R.E., Lazorchak, J.M. and Flick, R.W. (2007) 'Collapse of a fish population after exposure to a synthetic estrogen', *Proceedings of the National Academy of Sciences of the United States of America,* 104(21), pp. 8897-8901.

Kim, H.S., Shin, J.H., Moon, H.J., Kim, T.S., Kang, I.H., Seok, J.H., Kim, I.Y., Park, K.L. and

Han, S.Y. (2002) 'Evaluation of the 20-day pubertal female assay in Sprague-Dawley rats treated with DES, tamoxifen, testosterone, and flutamide', *Toxicological sciences : an official journal of the Society of Toxicology,* 67(1), pp. 52-62.

King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y. and Letunic, I. (2008) 'The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans', *Nature,* 451(7180), pp. 783-788.

King, R.C., Aggarwal, S.K. and Aggarwal, U. (1968) 'The development of the female *Drosophila* reproductive system', *Journal of Morphology,* 124(2), pp. 143-165.

King-Jones, K., Horner, M.A., Lam, G. and Thummel, C.S. (2006) 'The DHR96 nuclear receptor regulates xenobiotic responses in *Drosophila*', *Cell metabolism,* 4(1), pp. 37-48.

King-Jones, K. and Thummel, C. (2005) 'Nuclear receptors - A perspective from *Drosophila*', *Nature Reviews Genetics,* 6(4), pp. 311-323.

Kishta, O., Adeeko, A., Li, D., Luu, T., Brawer, J.R., Morales, C., Hermo, L., Robaire, B., Hales, B.F. and Barthelemy, J. (2007) 'In utero exposure to tributyltin chloride differentially alters male and female fetal gonad morphology and gene expression profiles in the Sprague–Dawley rat', *Reproductive Toxicology,* 23(1), pp. 1-11.

Kiss, I., Deak, F., Holloway, R.G.,Jr, Delius, H., Mebust, K.A., Frimberger, E., Argraves, W.S., Tsonis, P.A., Winterbottom, N. and Goetinck, P.F. (1989) 'Structure of the gene for cartilage matrix protein, a modular protein of the extracellular matrix. Exon/intron organization, unusual splice sites, and relation to alpha chains of beta 2 integrins, von Willebrand factor, complement factors B and C2, and epidermal growth factor', *The Journal of biological chemistry,* 264(14), pp. 8126-8134.

Kitano, T., Takamune, K., Nagahama, Y. and Abe, S. (2000) 'Aromatase inhibitor and 17a-methyltestosterone cause sex-reversal from genetical females to phenotypic males and suppression of P450 aromatase gene expression in Japanese flounder (Paralichthys olivaceus)', *Mol Reprod Dev,* 56(1), pp. 1-5.

Kliewer, S., Umesono, K., Noonan, D., Heyman, R. and Evans, R. (1992) 'Convergence of 9-Cis Retinoic Acid and Peroxisome Proliferator Signaling Pathways through Heterodimer Formation of their Receptors', *Nature,* 358(6389), pp. 771-774.

Klussmann-Kolb, A. (2001) 'Comparative investigation of the genital systems in the Opisthobranchia (Mollusca, Gastropoda) with special emphasis on the nidamental glandular system', *Zoomorphology,* 120(4), pp. 215-235.

Knight, M., Miller, A., Liu, Y., Scaria, P., Woodle, M. and Ittiprasert, W. (2011) 'Polyethyleneimine (PEI) mediated siRNA gene silencing in the *Schistosoma mansoni* snail host, *Biomphalaria glabrata*', *PLoS neglected tropical diseases,* 5(7), pp. e1212.

Kocot, K.M. (2013) *A combined approach toward resolving the phylogeny of Mollusca,* .

Koehler, H., Kloas, W., Schirling, M., Lutz, I., Reye, A.L., Langen, J., Triebskorn, R., Nagel, R. and Schoenfelder, G. (2007) 'Sex steroid receptor evolution and signalling in aquatic

invertebrates', *Ecotoxicology,* 16(1), pp. 131-143.

Koene, J.M. (2010a) 'Neuro-endocrine control of reproduction in hermaphroditic freshwater snails: mechanisms and evolution', *Frontiers in behavioral neuroscience,* 4, pp. 167.

Koene, J.M. and Ter Maat, A. (2007) 'Coolidge effect in pond snails: male motivation in a simultaneous hermaphrodite', *BMC evolutionary biology,* 7, pp. 212.

Korneev, S.A., Kemenes, I., Straub, V., Staras, K., Korneeva, E.I., Kemenes, G., Benjamin, P.R. and O'Shea, M. (2002) 'Suppression of nitric oxide (NO)-dependent behavior by double-stranded RNA-mediated silencing of a neuronal NO synthase gene', *J Neurosci,* 22(11), pp. RC227.

Kornthong, N., Chotwiwatthanakun, C., Chansela, P., Tinikul, Y., Cummins, S.F., Hanna, P.J. and Sobhon, P. (2013) 'Characterization of red pigment concentrating hormone (RPCH) in the female mud crab (< i> Scylla olivacea) and the effect of 5-HT on its expression', *General and comparative endocrinology,* 185, pp. 28-36.

Kotake, Y. (2012) 'Molecular mechanisms of environmental organotin toxicity in mammals', *Biological and Pharmaceutical Bulletin,* 35(11), pp. 1876-1880.

Krisberg, K. (2010) 'Push is on to improve US chemical safety laws: Agencies lead drive', *The Nation's Health,* 39(10), pp. 1-20.

Krusch, B., Schoenmakers, H., Voogt, P. and Nolte, A. (1979) 'Steroid synthesizing capacity of the dorsal body of< i> Helix pomatia L.(Gastropoda)—an *in vitro* study', *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry,* 64(1), pp. 101-104.

Kumkate, S., Onmek, N., Boonburapong, B., Singhakaew, S. and Leardkamolkarn, V. (2009) 'Estrogen-Related Fecundity Reduction of *Lymnaea ollula* Following *Fasciola gigantica* Infection', *Journal of Parasitology,* 95(6), pp. 1391-1396.

Lafont, R. and Mathieu, M. (2007a) 'Steroids in aquatic invertebrates', *Ecotoxicology,* 16(1), pp. 109-130.

Laguerre, M. and Veenstra, J.A. (2010) 'Ecdysone receptor homologs from mollusks, leeches and a polychaete worm', *FEBS letters,* 584(21), pp. 4458-4462.

Lalli, E., Melner, M., Stocco, D. and Sassone-Corsi, P. (1998) 'DAX-1 Blocks Steroid Production at Multiple Levels', 39.

Lalli, E. and Sassone-Corsi, P. (2003) 'DAX-1, an unusual orphan receptor at the crossroads of steroidogenic function and sexual differentiation', *Molecular Endocrinology,* 17(8), pp. 1445-1453.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M. and FitzHugh, W. (2001) 'Initial sequencing and analysis of the human genome', *Nature,* 409(6822), pp. 860-921.

Lange, A., Paull, G.C., Hamilton, P.B., Iguchi, T. and Tyler, C.R. (2011) 'Implications of persistent exposure to treated wastewater effluent for breeding in wild roach (*Rutilus rutilus*) populations', *Environmental science & technology,* 45(4), pp. 1673-1679.

Lardans, V., Coppin, J., Vicogne, J., Aroca, E., Delcroix, M. and Dissous, C. (2001) 'Characterization of an insulin receptor-related receptor in *Biomphalaria glabrata* embryonic cells', *Biochimica et Biophysica Acta (BBA)-Biomembranes,* 1510(1), pp. 321-329.

Larsson, D.G. and Forlin, L. (2002) 'Male-biased sex ratios of fish embryos near a pulp mill: temporary recovery after a short-term shutdown', *Environmental health perspectives,* 110(8), pp. 739-742.

Lassurguere, J., Livera, G., Habert, R. and Jegou, B. (2003) 'Time- and dose-related effects of estradiol and diethylstilbestrol on the morphology and function of the fetal rat testis in culture', *Toxicological sciences : an official journal of the Society of Toxicology,* 73(1), pp. 160-169.

Laudet, V., Gronemeyer, H., Laudet, V. and Gronemeyer, H. (2002) *The Nuclear Receptor Factsbook.* Academic Press San Diego:.

Laudet, V. (1997) 'Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor', *Journal of Molecular Endocrinology,* 19(3), pp. 207-226.

Launhardt, H. and Munder, T. (2000) 'Post-translational regulation of Saccharomyces cerevisiae proteins tagged with the hormone-binding domains of mammalian nuclear receptors', *Molecular and General Genetics MGG,* 264(3), pp. 317-324.

Lavado, R., Janer, G. and Porte, C. (2006a) 'Steroid levels and steroid metabolism in the Mussel *Mytilus edulis*: The modulating effect of dispersed crude oil and alkylphenols', *Aquatic toxicology,* 78, pp. S65-S72.

Lavado, R., Sugni, M., Candia Carnevali, M.D. and Porte, C. (2006b) 'Triphenyltin alters androgen metabolism in the sea urchin *Paracentrotus lividus*', *Aquatic toxicology,* 79(3), pp. 247-256.

Le Curieux-Belfond, O., Moslemi, S., Mathieu, M. and Séralini, G. (2001) 'Androgen metabolism in oyster *Crassostrea gigas*: evidence for 17β-HSD activities and characterization of an aromatase-like activity inhibited by pharmacological compounds and a marine pollutant', *The Journal of steroid biochemistry and molecular biology,* 78(4), pp. 359-366.

le Maire, A., Grimaldi, M., Roecklin, D., Dagnino, S., Vivat- Hannah, V., Balaguer, P. and Bourguet, W. (2009a) 'Activation of RXR–PPAR heterodimers by organotin environmental endocrine disruptors', *EMBO reports,* 10(4), pp. 367-373.

Lebel, J., Giard, W., Favrel, P. and Boucaud- Camou, E. (1996) 'Effects of different vertebrate growth factors on primary cultures of hemocytes from the gastropod mollusc, *Haliotis tuberculata*', *Biology of the Cell,* 86(1), pp. 67-72.

Leblanc, G.A., Gooding, M.P. and Sternberg, R.M. (2005) 'Testosterone-Fatty Acid esterification: a unique target for the endocrine toxicity of tributyltin to gastropods', *Integrative and comparative biology,* 45(1), pp. 81-87.

Lecomte, V., Noury, P., Tutundjian, R., Buronfosse, T., Garric, J. and Gust, M. (2013) 'Organic solvents impair life-traits and biomarkers in the New Zealand mudsnail *Potamopyrgus antipodarum* (Gray) at concentrations below OECD recommendations', *Aquatic Toxicology,*

140, pp. 196-203.

Lefebvre, P., Benomar, Y. and Staels, B. (2010) 'Retinoid X receptors: common heterodimerization partners with distinct functions', *Trends in Endocrinology & Metabolism,* 21(11), pp. 676-683.

Lehoux, J.G. and Sandor, T. (1970) 'The occurrence of steroids and steroid metabolizing enzyme systems in invertebrates. A review', *Steroids,* 16(2), pp. 141-171.

Letunic, I., Copley, R.R., Pils, B., Pinkert, S., Schultz, J. and Bork, P. (2006) 'SMART 5: domains in the context of genomes and networks', *Nucleic acids research,* 34, pp. 242-244.

Leung, K.M., Grist, E.P., Morley, N.J., Morritt, D. and Crane, M. (2007) 'Chronic toxicity of tributyltin to development and reproduction of the European freshwater snail *Lymnaea stagnalis*', *Chemosphere,* 66(7), pp. 1358-1366.

Leung, K.M., Morley, N.J., Grist, E.P., Morritt, D. and Crane, M. (2004) 'Chronic toxicity of tributyltin on development and reproduction of the hermaphroditic snail *Physa fontinalis*: Influence of population density', *Marine environmental research,* 58(2), pp. 157-162.

Leung, M.C., Williams, P.L., Benedetto, A., Au, C., Helmcke, K.J., Aschner, M. and Meyer, J.N. (2008) '*Caenorhabditis elegans*: an emerging model in biomedical and environmental toxicology', *Toxicological sciences : an official journal of the Society of Toxicology,* 106(1), pp. 5-28.

Lever, J. and Boer, H. (1983) *Molluscan neuro-endocrinology: proceedings of the International Minisymposium on Molluscan Endocrinology, held in the Department of Biology, Free University, Amsterdam, the Netherlands, August 16-20, 1982.* North-Holland.

Levin, E.R. (2008) 'Rapid signaling by steroid receptors', *Am. J. Physiol. Regul. Integr. Comp. Physiol.,* 295(5), pp. R1425-R1430.

Levoye, A., Mouillac, B., Riviere, G., Vieau, D., Salzet, M. and Breton, C. (2005) 'Cloning, expression and pharmacological characterization of a vasopressin-related receptor in an annelid, the leech Theromyzon tessulatum', *The Journal of endocrinology,* 184(1), pp. 277-289.

Lewis, P.O., Holder, M.T. and Holsinger, K.E. (2005) 'Polytomies and Bayesian phylogenetic inference', *Systematic Biology,* 54(2), pp. 241-253.

Li, K.W., Geraerts, W.P., Ebberink, R.H. and Joosse, J. (1992) 'Purification and sequencing of molluscan insulin-related peptide I (MIP I) from the neuroendocrine light green cells of *Lymnaea stagnalis*', *Molecular and cellular endocrinology,* 85(1), pp. 141-150.

Li, T. and Bender, M. (2000) 'A conditional rescue system reveals essential functions for the ecdysone receptor (EcR) gene during molting and metamorphosis in *Drosophila*', *Development,* 127(13), pp. 2897-2905.

Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Cazzamali, G., Williamson, M., Arakane, Y., Verleyen, P., Schoofs, L., Schachtner, J., Grimmelikhuijzen, C.J. and Park, Y. (2008) 'Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the

red flour beetle Tribolium castaneum', *Genome research,* 18(1), pp. 113-122.

Li, K.W., Jimenez, C.R., Van Veelen, P.A. and Geraerts, W.P. (1994) 'Processing and targeting of a molluscan egg-laying peptide prohormone as revealed by mass spectrometric peptide fingerprinting and peptide sequencing', *Endocrinology,* 134(4), pp. 1812-1819.

Lima, D., Reis-Henriques, M., Silva, R., Santos, A., Filipe C Castro, L. and Santos, M. (2011) 'Tributyltin-induced imposex in marine gastropods involves tissue-specific modulation of the retinoid X receptor', *Aquatic toxicology,* 101(1), pp. 221-227.

Lindberg, D.R. and Wright, W. (1985) 'Patterns of sex change of the protandrie patellacean limpet *Lottia gigantea* (Mollusca: Gastropoda)', *The Veliger,* 27(3), pp. 261-265.

Lindblom, T.H. and Dodd, A.K. (2006) 'Xenobiotic detoxification in the nematode *Caenorhabditis elegans*', *Journal of Experimental Zoology Part A: Comparative Experimental Biology,* 305(9), pp. 720-730.

Little, C., Stirling, P., Pilkington, M. and Pilkington, J. (1985) 'Larval development and metamorphosis in the marine pulmonate Amphibola crenata (Mollusca: Pulmonata)', *Journal of zoology,* 205(4), pp. 489-510.

Liu, S. (2010) *Analysis of TR4 Nuclear Receptor Knockout Mouse: Roles of TR4 in Mitochondrial Function and Cellular Stress Defenses,* .

Lloyd, P.E. and Connolly, C.M. (1989) 'Sequence of pedal peptide: a novel neuropeptide from the central nervous system of *Aplysia*', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 9(1), pp. 312-317.

Lockyer, A.E., Spinks, J., Kane, R.A., Hoffmann, K.F., Fitzpatrick, J.M., Rollinson, D., Noble, L.R. and Jones, C.S. (2008) '*Biomphalaria glabrata* transcriptome: cDNA microarray profiling identifies resistant- and susceptible-specific gene expression in haemocytes from snail strains exposed to *Schistosoma mansoni*', *BMC genomics,* 9, pp. 634-2164-9-634.

Loi, P.K. and Tublitz, N. (1997) 'Molecular analysis of FMRFamide- and FMRFamide-related peptides (FaRPS) in the cuttlefish *Sepia officinalis*', *The Journal of experimental biology,* 200(Pt 10), pp. 1483-1489.

Loker, E.S. (2010) 'Gastropod immunobiology', in *Invertebrate immunity.* Springer, pp. 17-43.

Lyssimachou, A., Bachmann, J. and Porte, C. (2008) 'Short-term exposure to the organotin compound triphenyltin modulates esterified steroid levels in females of *Marisa cornuarietis*', *Aquatic toxicology,* 89(2), pp. 129-135.

Lyssimachou, A., Ramón, M. and Porte, C. (2009) 'Comparative study on the metabolism of the androgen precursor androstenedione in two gastropod species: *In vitro* alterations by TBT and TPT', *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology,* 149(3), pp. 409-413.

Madupu, R., Brinkac, L.M., Harrow, J., Wilming, L.G., Bohme, U., Lamesch, P. and Hannick, L.I. (2010) 'Meeting report: a workshop on Best Practices in Genome Annotation', *Database : the journal of biological databases and curation,* 2010, pp. baq001.

Maglich, J.M., Sluder, A., Guan, X., Shi, Y., Mckee, D.D., Carrick, K., Kamdar, K., Willson, T.M. and Moore, J.T. (2001) 'Comparison of complete nuclear receptor sets from the human, *Caenorhabditis elegans* and *Drosophila* genomes', *Genome biology,* 2(8), pp. 0029.1.

Mahon, A.C., Nambu, J.R., Taussig, R., Shyamala, M., Roach, A. and Scheller, R.H. (1985) 'Structure and expression of the egg-laying hormone gene family in *Aplysia*', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 5(7), pp. 1872-1880.

Mangal TD, Paterson S, Fenton A (2010) 'Effects of snail density on growth, reproduction and survival of *Biomphalaria alexandrina* exposed to *Schistosoma mansoni*.', *J Parasitol Res,* , pp. 186792.

Mangelsdorf, D.J. and Evans, R.M. (1995) 'The RXR heterodimers and orphan receptors.', *Cell,* 83(6), pp. 841.

Mangelsdorf, D., Thummel, C., BeatO, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. and Evans, R. (1995) 'The Nuclear Receptor Superfamily - the 2nd Decade', *Cell,* 83(6), pp. 835-839.

Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Krylov, D., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Thanki, N., Yamashita, R.A., Yin, J.J., Zhang, D. and Bryant, S.H. (2007) 'CDD: a conserved domain database for interactive domain family analysis', *Nucleic acids research,* 35.

Margiotta-Casaluci, L. and Sumpter, J.P. (2011) '5α-Dihydrotestosterone is a potent androgen in the fathead minnow (*Pimephales promelas*)', *General and comparative endocrinology,* 171(3), pp. 309-318.

Markov, G.V. and Laudet, V. (2011a) 'Origin and evolution of the ligand-binding ability of nuclear receptors', *Molecular and cellular endocrinology,* 334(1), pp. 21-30.

Markov, G.V. and Laudet, V. (2011b) 'Origin and evolution of the ligand-binding ability of nuclear receptors', *Molecular and cellular endocrinology,* 334(1), pp. 21-30.

Markov, G.V., Tavares, R., Dauphin-Villemant, C., Demeneix, B.A., Baker, M.E. and Laudet, V. (2009) 'Independent elaboration of steroid hormone signaling pathways in metazoans', *Proc. Natl. Acad. Sci.,* 106(29), pp. 11913-11918.

Martínez-Pérez, F., Becerra, A., Valdés, J., Zinker, S. and Aréchiga, H. (2002a) 'A possible molecular ancestor for mollusk APGWamide, insect adipokinetic hormone, and crustacean red pigment concentrating hormone', *Journal of Molecular Evolution,* 54(6), pp. 703-714.

Martinović-Weigelt, D., Wang, R., Villeneuve, D.L., Bencic, D.C., Lazorchak, J. and Ankley, G.T. (2011) 'Gene expression profiling of the androgen receptor antagonists flutamide and vinclozolin in zebrafish (*Danio rerio*) gonads', *Aquatic Toxicology,* 101(2), pp. 447-458.

Marxen, J.C., Becker, W., Finke, D., Hasse, B. and Epple, M. (2003) 'Early mineralization in *Biomphalaria glabrata*: microscopic and structural results', *Journal of Molluscan Studies,* 69(2), pp. 113-121.

Mashanov, V.S., Zueva, O.R., Rojas-Catagena, C. and Garcia-Arraras, J.E. (2010) 'Visceral regeneration in a sea cucumber involves extensive expression of survivin and mortalin homologs in the mesothelium', *BMC developmental biology,* 10, pp. 117-213X-10-117.

Matsumoto, T., Nakamura, A.M., Mori, K., Akiyama, I., Hirose, H. and Takahashi, Y. (2007) 'Oyster estrogen receptor: cDNA cloning and immunolocalization', *General and comparative endocrinology,* 151(2), pp. 195-201.

Matsumoto, T., Osada, M., Osawa, Y. and Mori, K. (1997) 'Gonadal estrogen profile and immunohistochemical localization of steroidogenic enzymes in the oyster and scallop during sexual maturation', *Comparative biochemistry and physiology Part B: Biochemistry and molecular biology,* 118(4), pp. 811-817.

Matthiessen, P. (2008) 'An assessment of endocrine disruption in mollusks and the potential for developing internationally standardized mollusk life cycle test guidelines', *Integrated Environ Assess Manag,* 4(3), pp. 274-284.

Matthiessen, P. (2000) 'Is endocrine disruption a significant ecological issue?', *Ecotoxicology,* 9(1-2), pp. 21-24.

Matthiessen, P. and Gibbs, P.E. (1998a) 'Critical appraisal of the evidence for tributyltin-mediated endocrine disruption in mollusks', *Environmental toxicology and chemistry,* 17(1), pp. 37-43.

Maule, A.G. and Marks, N.J. (2006) *Parasitic flatworms: molecular biology, biochemistry, immunology and physiology.* CABI.

Maydl, R., Newbold, R., Metzler, M. and McLachlan, J.A. (1983) 'Diethylstilbestrol metabolism by the fetal genital tract', *Endocrinology,* 113(1), pp. 146-151.

McAllister, L.B., Scheller, R.H., Kandel, E.R. and Axel, R. (1983) 'In situ hybridization to study the origin and fate of identified neurons', *Science (New York, N.Y.),* 222(4625), pp. 800-808.

McClay, D.R. (2011) 'Evolutionary crossroads in developmental biology: sea urchins', *Development (Cambridge, England),* 138(13), pp. 2639-2648.

McClellan-Green, P.D. (2013) 'Endocrine Disruption in Molluscs', *Endocrine Disrupters: Hazard Testing and Assessment Methods,* , pp. 143.

Melmed, S., Polonsky, K.S., Larsen, P.R. and Kronenberg, H.M. (2011) *Williams textbook of endocrinology: Expert consult.* Elsevier Health Sciences.

Merritt, W.M. and Sood, A.K. (2007) 'Markers of angiogenesis in ovarian cancer', *Disease markers,* 23(5), pp. 419-431.

Metcalfe, C.D., Chu, S., Judt, C., Li, H., Oakes, K.D., Servos, M.R. and Andrews, D.M. (2010) 'Antidepressants and their metabolites in municipal wastewater, and downstream exposure in an urban watershed', *Environmental Toxicology and Chemistry,* 29(1), pp. 79-89.

Miksys, S. and Saleuddin, A. (1987) 'Effects of castration on growth and reproduction of *Helisoma duryi* (Mollusca: Pulmonata)', *International journal of invertebrate reproduction and development,* 12(2), pp. 145-159.

Miller, M.W., Beushausen, S., Cropper, E.C., Eisinger, K., Stamm, S., Vilim, F.S., Vitek, A., Zajc, A., Kupfermann, I. and Brosius, J. (1993a) 'The buccalin-related neuropeptides: isolation and characterization of an *Aplysia* cDNA clone encoding a family of peptide cotransmitters', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 13(8), pp. 3346-3357.

Miller, M.W., Beushausen, S., Vitek, A., Stamm, S., Kupfermann, I., Brosius, J. and Weiss, K.R. (1993b) 'The myomodulin-related neuropeptides: characterization of a gene encoding a family of peptide cotransmitters in *Aplysia*', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 13(8), pp. 3358-3367.

Missmer, S.A., Hankinson, S.E., Spiegelman, D., Barbieri, R.L., Michels, K.B. and Hunter, D.J. (2004) 'In utero exposures and the incidence of endometriosis', *Fertility and sterility,* 82(6), pp. 1501-1508.

Mittendorf, R. (1995) 'Teratogen update: carcinogenesis and teratogenesis associated with exposure to diethylstilbestrol (DES) in utero', *Teratology,* 51(6), pp. 435-445.

Miyawaki, T., Uemura, A., Dezawa, M., Ruth, T.Y., Ide, C., Nishikawa, S., Honda, Y., Tanabe, Y. and Tanabe, T. (2004) 'Tlx, an orphan nuclear receptor, regulates cell numbers and astrocyte development in the developing retina', *The Journal of neuroscience,* 24(37), pp. 8124-8134.

Moller, D.E. (2001) 'New drug targets for type 2 diabetes and the metabolic syndrome', *Nature,* 414(6865), pp. 821-827.

Moon, H., Kang, T.S., Kim, T.S., Kang, I.H., Ki, H.Y., Kim, S.H. and Han, S.Y. (2009) 'OECD validation of phase 3 Hershberger assay in Korea using surgically castrated male rats with coded chemicals', *Journal of Applied Toxicology,* 29(4), pp. 350-355.

Mori, K., Muramatsu, T. and Nakamura, Y. (1969) 'Effect of steroid on oyster III. Sex reversal from male to female in *Crassostrea gigas* by estradiol-17β', *Bull.Jap.Soc.Sci.Fish,* 35, pp. 1072-1076.

Moroz, L.L., Edwards, J.R., Puthanveettil, S.V., Kohn, A.B., Ha, T., Heyland, A., Knudsen, B., Sahni, A., Yu, F. and Liu, L. (2006) 'Neuronal Transcriptome of *Aplysia*: Neuronal Compartments and Circuitry', *Cell,* 127(7), pp. 1453-1467.

Mortelmans, K. and Zeiger, E. (2000) 'The Ames *Salmonella* microsome mutagenicity assay', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 455(1), pp. 29-60.

Moss, S. (1989) 'Effects of exogenous androgens on growth, biochemical composition, and reproduction of the coot clam, Mulinia lateralis', *Pac.Sci,* 43, pp. 200.

Mu, X., Lee, Y., Liu, N., Chen, Y., Kim, E., Shyr, C. and Chang, C. (2004) 'Targeted inactivation of testicular nuclear orphan receptor 4 delays and disrupts late meiotic prophase and subsequent meiotic divisions of spermatogenesis', *Molecular and cellular biology,* 24(13), pp. 5887-5899.

Mukai, S.T., Hoque, T., Morishita, F. and Saleuddin, A.S.M. (2004) 'Cloning and characterization

of a candidate nutritive glycoprotein from the albumen gland of the freshwater snail, *Helisoma duryi* (Mollusca: Pulmonata)', *Invertebrate Biology,* 123(1), pp. 83-92.

Munkittrick, K., McMaster, M., McCarthy, L., Servos, M. and Van Der Kraak, G. (1998) 'An overview of recent studies on the potential of pulp- mill effluents to alter reproductive parameters in fish', *Journal of Toxicology and Environmental Health, Part B Critical Reviews,* 1(4), pp. 347-371.

Munkittrick, K., Portt, C., Kraak, G.V.D., Smith, I. and Rokosh, D. (1991) 'Impact of bleached kraft mill effluent on population characteristics, liver MFO activity, and serum steroid levels of a Lake Superior white sucker (Catostomus commersoni) population', *Canadian Journal of Fisheries and Aquatic Sciences,* 48(8), pp. 1371-1380.

Naar, A.M. and Thakur, J.K. (2009) 'Nuclear receptor-like transcription factors in fungi', *Genes & development,* 23(4), pp. 419-432.

Nagle, G.T., de Jong- Brink, M., Painter, S.D. and Li, K.W. (2001) 'Structure, localization and potential role of a novel molluscan trypsin inhibitor in *Lymnaea*', *European Journal of Biochemistry,* 268(5), pp. 1213-1221.

Nagle, G.T., Painter, S.D., Blankenship, J.E. and Kurosky, A. (1988) 'Proteolytic processing of egg-laying hormone-related precursors in Aplysia. Identification of peptide regions critical for biological activity', *The Journal of biological chemistry,* 263(19), pp. 9223-9237.

Nakanishi, T. and Nishikawa, J. (2009) 'Genetic impacts of organotin compounds', in *Ecotoxicology of Antifouling Biocides.* Springer, pp. 221-233.

Nakanishi, T., Nishikawa, J. and Tanaka, K. (2006) 'Molecular targets of organotin compounds in endocrine disruption: do organotin compounds function as aromatase inhibitors in mammals?', *Environmental sciences : an international journal of environmental physiology and toxicology,* 13(2), pp. 89-100.

Nakayama, K., Watanabe, T., Nakagawa, T., Kim, W.S., Nagahama, M., Hosaka, M., Hatsuzawa, K., Kondoh-Hashiba, K. and Murakami, K. (1992) 'Consensus sequence for precursor processing at mono-arginyl sites. Evidence for the involvement of a Kex2-like endoprotease in precursor cleavages at both dibasic and mono-arginyl sites', *The Journal of biological chemistry,* 267(23), pp. 16335-16340.

Nambu, J.R., Murphy-Erdosh, C., Andrews, P.C., Feistner, G.J. and Scheller, R.H. (1988) 'Isolation and characterization of a *Drosophila* neuropeptide gene', *Neuron,* 1(1), pp. 55-61.

Nanda, S., DeFalco, T.J., Loh, S.H., Phochanukul, N., Camara, N., Van Doren, M. and Russell, S. (2009) 'Sox100B, a *Drosophila* group E Sox-domain gene, is required for somatic testis differentiation', *Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation,* 3(1), pp. 26-37.

Naora, H. and Montell, D.J. (2005) 'Ovarian cancer metastasis: integrating insights from disparate model organisms', *Nature Reviews Cancer,* 5(5), pp. 355-366.

Nässel, D.R. (1996) 'Peptidergic neurohormonal control systems in invertebrates', *Current opinion*

in neurobiology, 6(6), pp. 842-850.

Newbold, R.R. (2004) 'Lessons learned from perinatal exposure to diethylstilbestrol', *Toxicology and applied pharmacology,* 199(2), pp. 142-150.

Newbold, R. (2012) 'Prenatal exposure to diethylstilbestrol and long-term impact on the breast and reproductive tract in humans and mice', *Journal of Developmental Origins of Health and Disease,* 3(02), pp. 73-82.

Newton WL 'The establishment of a strain of *Australorbis glabratus* which combines albinism and high susceptibility to infection with *Schistosoma mansoni*.', *J Parasitol,* 41, pp. 526-8.

Nikolenko, J. and Krasnov, A. (2007) 'Nuclear receptors: Structure and mechanisms of action', *Russian Journal of Genetics,* 43(3), pp. 234-240.

Nishikawa, J., Mamiya, S., Kanayama, T., Nishikawa, T., Shiraishi, F. and Horiguchi, T. (2004) 'Involvement of the retinoid X receptor in the development of imposex caused by organotins in gastropods', *Environmental science & technology,* 38(23), pp. 6271-6276.

Norris, D.O. and Carr, J.A. (2013) *Vertebrate endocrinology.* Academic Press.

Norris, V., Madsen, M.S. and Freestone, P. (1996) 'Elements of a unifying theory of biology', *Acta Biotheoretica,* 44(3-4), pp. 209-218.

North, T.E., Goessling, W., Walkley, C.R., Lengerke, C., Kopani, K.R., Lord, A.M., Weber, G.J., Bowman, T.V., Jang, I. and Grosser, T. (2007) 'Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis', *Nature,* 447(7147), pp. 1007-1011.

Nussey, S. and Whitehead, S. (2001) 'Principles of endocrinology', .

Oberdörster, E. and McClellan-Green, P. (2000) 'The neuropeptide APGWamide induces imposex in the mud snail, *Ilyanassa obsoleta*', *Peptides,* 21(9), pp. 1323-1330.

Oberdörster, E. and McClellan-Green, P. (2002) 'Mechanisms of imposex induction in the mud snail, *Ilyanassa obsoleta*: TBT as a neurotoxin and aromatase inhibitor', *Marine environmental research,* 54(3), pp. 715-718.

Oberdorster, E., Romano, J. and McClellan-Green, P. (2005) 'The Neuropeptide APGWamide as a Penis Morphogenic Factor (PMF) in Gastropod Mollusks', *Integrative and comparative biology,* 45(1), pp. 28-32.

O'Connor, M., Peifer, M. and Bender, W. (1989) 'Construction of large DNA segments in *Escherichia coli*', *Science (New York, N.Y.),* 244(4910), pp. 1307-1312.

Odoemelam, E., Raghavan, N., Miller, A., Bridger, J.M. and Knight, M. (2009) 'Revised karyotyping and gene mapping of the Biomphalaria glabrata embryonic (Bge) cell line', *International journal for parasitology,* 39(6), pp. 675-681.

Oehlmann, J., Fioroni, P., Stroben, E. and Markert, B. (1996) 'Tributyltin (TBT) effects on *Ocinebrina aciculata* (Gastropoda: Muricidae): imposex development, sterilization, sex change and population decline', *Science of the Total Environment,* 188(2), pp. 205-223.

Oehlmann, J. and Schulte-Oehlmann, U. (2003) 'Endocrine disruption in invertebrates', *Pure Appl. Chem.,* 75(11), pp. 2207-2218.

Oehlmann, J., Di Benedetto, P., Tillmann, M., Duft, M., Oetken, M. and Schulte-Oehlmann, U. (2007) 'Endocrine disruption in prosobranch molluscs: evidence and ecological relevance', *Ecotoxicology,* 16(1), pp. 29-43.

Oehlmann, J., Schulte-Oehlmann, U., Bachmann, J., Oetken, M., Lutz, I., Kloas, W. and Ternes, T.A. (2006) 'Bisphenol A induces superfeminization in the ramshorn snail *Marisa cornuarietis* (Gastropoda: Prosobranchia) at environmentally relevant concentrations', *Environmental health perspectives,* 114, pp. 127.

Oetken, M., Bachmann, J., Schulte-Oehlmann, U. and Oehlmann, J. (2004) 'Evidence for endocrine disruption in invertebrates', *International review of cytology,* 236, pp. 1-44.

Omran, N.E.E. (2012) 'Testosterone, gonadotropins and androgen receptor during spermatogenesis of *Biomphalaria alexandrina* snails (Pulmonata: Basommatophora)', *Reprod Biol,* 12(3), pp. 301-308.

Onitsuka, C., Yamaguchi, A., Kanamaru, H., Oikawa, S., Takeda, T. and Matsuyama, M. (2009) 'Molecular cloning and expression analysis of a GnRH-like dodecapeptide in the swordtip squid, *Loligo edulis*', *Zoological Science,* 26(3), pp. 203-208.

Orlando, E.F., Kolok, A.S., Binczik, G.A., Gates, J.L., Horton, M.K., Lambright, C.S., Gray, L.E.,Jr, Soto, A.M. and Guillette, L.J.,Jr (2004) 'Endocrine-disrupting effects of cattle feedlot effluent on an aquatic sentinel species, the *fathead minnow*', *Environmental health perspectives,* 112(3), pp. 353-358.

Osada, M., Takamura, T., Sato, H. and Mori, K. (2003) 'Vitellogenin synthesis in the ovary of scallop, *Patinopecten yessoensis*: Control by estradiol- 17β and the central nervous system', *Journal of Experimental Zoology Part A: Comparative Experimental Biology,* 299(2), pp. 172-179.

Osada, M., Tawarayama, H. and Mori, K. (2004) 'Estrogen synthesis in relation to gonadal development of Japanese scallop, *Patinopecten yessoensis*: gonadal profile and immunolocalization of P450 aromatase and estrogen', *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology,* 139(1), pp. 123-128.

Ottow, E. and Weinmann, H. (2008) *Nuclear receptors as drug targets: A historical perspective of modern drug discovery.* Wiley-VCH Verlag GmbH & Co KGaA: Weinheim, Germany.

Owen, G. and Zelent, A. (2000) 'Origins and evolutionary diversification of the nuclear receptor superfamily', *Cellular and Molecular Life Sciences CMLS,* 57(5), pp. 809-827.

Owens, W., Gray Jr, L.E., Zeiger, E., Walker, M., Yamasaki, K., Ashby, J. and Jacob, E. (2007) 'The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses: phase 2 dose–response studies', *Environmental health perspectives,* 115(5), pp. 671.

Owens, W., Zeiger, E., Walker, M., Ashby, J., Onyon, L. and Gray Jr, L.E. (2006) 'The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: use of a potent agonist and a potent antagonist to test

the standardized protocol', *Environmental health perspectives,* 114(8), pp. 1259.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) 'Splicing of messenger RNA precursors', *Annual Review of Biochemistry,* 55(1), pp. 1119-1150.

Painter, S.D., Gustavson, A.R., Kalman, V.K., Nagle, G.T. and Blankenship, J.E. (1989) 'Induction of copulatory behavior in *Aplysia*: Atrial gland factors mimic the excitatory effects of freshly deposited egg cordons', *Behavioral and neural biology,* 51(2), pp. 222-236.

Pal, P. (2007) 'Fine structure of reproductive glands in two primitive marine pulmonates (Basommatophora: Siphonariidae)', *Acta Zoologica,* 88(2), pp. 145-152.

Palmer, J.R., Wise, L.A., Robboy, S.J., Titus-Ernstoff, L., Noller, K.L., Herbst, A.L., Troisi, R. and Hoover, R.N. (2005) 'Hypospadias in sons of women exposed to diethylstilbestrol in utero', *Epidemiology,* 16(4), pp. 583-586.

Pandey, U.B. and Nichols, C.D. (2011) 'Human disease models in *Drosophila melanogaster* and the role of the fly in therapeutic drug discovery', *Pharmacological reviews,* 63(2), pp. 411-436.

Papoulias, D., Noltie, D. and Tillitt, D. (2000) 'Effects of methyl testosterone exposure on sexual differentiation in medaka, *Oryzias latipes*', *Marine environmental research,* 50(1), pp. 181-184.

Paris, M., Hillenweck, A., Bertrand, S., Delous, G., Escriva, H., Zalko, D., Cravedi, J. and Laudet, V. (2010) 'Active metabolism of thyroid hormone during metamorphosis of amphioxus', *Integr Comp Biol,* 50(1), pp. 63-74.

Parkinson, J. and Blaxter, M. (2009) *Expressed sequence tags: an overview.* Springer.

Parks, L.G., Lambright, C.S., Orlando, E.F., Guillette, L.J.,Jr, Ankley, G.T. and Gray, L.E.,Jr (2001) 'Masculinization of female mosquitofish in Kraft mill effluent-contaminated Fenholloway River water is associated with androgen receptor agonist activity', *Toxicological sciences : an official journal of the Society of Toxicology,* 62(2), pp. 257-267.

Pascoal, S., Carvalho, G., Vasieva, O., Hughes, R., Cossins, A., Fang, Y., Ashelford, K., Olohan, L., Barroso, C. and Mendo, S. (2012) 'Transcriptomics and *in vivo* tests reveal novel mechanisms underlying endocrine disruption in an ecological sentinel, *Nucella lapillus*', *Mol Ecol,* 22(6), pp. 1589-1608.

Passamaneck, Y.J., Schander, C. and Halanych, K.M. (2004) 'Investigation of molluscan phylogeny using large-subunit and small-subunit nuclear rRNA sequences', *Molecular phylogenetics and evolution,* 32(1), pp. 25-38.

Pawlowski, S., Sauer, A., Shears, J., Tyler, C. and Braunbeck, T. (2004) 'Androgenic and estrogenic effects of the synthetic androgen 17α-methyltestosterone on sexual development and reproductive performance in the fathead minnow (*Pimephales promelas*) determined using the gonadal recrudescence assay', *Aquatic toxicology,* 68(3), pp. 277-291.

Peehl, D.M. and Feldman, D. (2004) 'Interaction of nuclear receptor ligands with the Vitamin D signaling pathway in prostate cancer', *The Journal of steroid biochemistry and molecular*

biology, 92(4), pp. 307-315.

Peña, J., Guerra, M., Gaudencio, M.J. and Kendall, M. (1988) 'The occurrence of imposex in the gastropod Nucella lapillus at sites in Spain and Portugal', *Lurralde,* 11, pp. 445-451.

Phelps, C., Gburcik, V., Suslova, E., Dudek, P., Forafonov, F., Bot, N., MacLean, M., Fagan, R.J. and Picard, D. (2006) 'Fungi and animals may share a common ancestor to nuclear receptors', *Proceedings of the National Academy of Sciences of the United States of America,* 103(18), pp. 7077-7081.

Pickart, M.A., Klee, E.W., Nielsen, A.L., Sivasubbu, S., Mendenhall, E.M., Bill, B.R., Chen, E., Eckfeldt, C.E., Knowlton, M. and Robu, M.E. (2006) 'Genome-wide reverse genetics framework to identify novel functions of the vertebrate secretome', *PLoS One,* 1(1), pp. e104.

Pinder, L., Pottinger, T., Billinghurst, Z. and Depledge, M. (1999) *Endocrine function in aquatic invertebrates and evidence for disruption by environmental pollutants.* Environment Agency.

Pleis JR, Ward BW, Lucas JW (2010) '
Summary health statistics for U.S. adults: National Health Interview Survey, 2009. ', *Vital and Health Statistics. Series 10: Data from the National Health Survey,* 249, pp. 1-207.

Ponder, W.F. and Lindberg, D.R. (2008) *Phylogeny and Evolution of the Mollusca.* Univ of California Press.

Porte, C., Janer, G., Lorusso, L., Ortiz-Zarragoitia, M., Cajaraville, M., Fossi, M. and Canesi, L. (2006) 'Endocrine disruptors in marine organisms: approaches and perspectives', *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology,* 143(3), pp. 303-315.

Pourquié, O. (2003) 'Vertebrate somitogenesis: a novel paradigm for animal segmentation?', *International Journal of Developmental Biology,* 47(7/8), pp. 597-604.

Prins, G.S. (2008) 'Endocrine disruptors and prostate cancer risk', *Endocrine-related cancer,* 15(3), pp. 649-656.

Puinean, A., Labadie, P., Hill, E., Osada, M., Kishida, M., Nakao, R., Novillo, A., Callard, I. and Rotchell, J. (2006) 'Laboratory exposure to 17β-estradiol fails to induce vitellogenin and estrogen receptor gene expression in the marine invertebrate *Mytilus edulis*', *Aquatic toxicology,* 79(4), pp. 376-383.

Qiu, C., Liu, S., Hong, Y., Fu, Z., Wei, M., Ai, D. and Lin, J. (2012) 'Molecular characterization of thyroid hormone receptor beta from *Schistosoma japonicum* and assessment of its potential as a vaccine candidate antigen against schistosomiasis in BALB/c mice', *Parasit Vectors,* 5, pp. 172.

Raghavan, N. and Knight, M. (2006a) 'The snail (*Biomphalaria glabrata*) genome project', *Trends in parasitology,* 22(4), pp. 148-151.

Raghavan, N., Miller, A.N., Gardner, M., FitzGerald, P.C., Kerlavage, A.R., Johnston, D.A., Lewis, F.A. and Knight, M. (2003) 'Comparative gene analysis of *Biomphalaria glabrata* hemocytes pre-and post-exposure to miracidia of *Schistosoma mansoni*', *Molecular and*

biochemical parasitology, 126(2), pp. 181-191.

Ram, J.L., Gallardo, C.S., Ram, M.L. and Croll, R.P. (1998) 'Reproduction-associated immunoreactive peptides in the nervous systems of prosobranch gastropods', *The Biological bulletin,* 195(3), pp. 308-318.

Ramasubraniam, K. (1979) 'A histochemical study of the secretions of reproductive glands and of the egg envelopes of *Achatina fulica* (Pulmonata: Stylommatophora)', *International Journal of Invertebrate Reproduction,* 1(6), pp. 333-346.

Rana, A.K. and Misra-Bhattacharya, S. (2013) 'Current drug targets for helminthic diseases', *Parasitology research,* 112(5), pp. 1819-1831.

Rastogi, R.K., Di Fiore, M.M., D'Aniello, A., Iela, L. and Fiorentino, M. (2002) 'GnRH in the invertebrates: an overview', *Progress in brain research,* 141, pp. 19-29.

Rauschenbach, I.Y., Sukhanova, M.Z., Hirashima, A., Sutsugu, E. and Kuano, E. (2000) 'Role of the ecdysteroid system in the regulation of *Drosophila* reproduction under environmental stress', *Doklady Biological Sciences.* Springer, 641-643.

Reese, M.G., Eeckman, F.H., Kulp, D. and Haussler, D. (1997) 'Improved splice site detection in Genie', *Journal of computational biology,* 4(3), pp. 311-323.

Reitzel, A.M., Pang, K., Ryan, J.F., Mullikin, J.C., Martindale, M.Q., Baxevanis, A.D. and Tarrant, A.M. (2011) 'Nuclear receptors from the ctenophore Mnemiopsis leidyi lack a zinc-finger DNA-binding domain: lineage-specific loss or ancestral condition in the emergence of the nuclear receptor superfamily', *EvoDevo,* 2(1), pp. 3-3.

Reitzel, A.M. and Tarrant, A.M. (2009) 'Nuclear receptor complement of the cnidarian *Nematostella vectensis*: phylogenetic relationships and developmental expression patterns', *BMC evolutionary biology,* 9(1), pp. 230.

Renaud, J. and Moras, D. (2000) 'Structural studies on nuclear receptors', *Cellular and Molecular Life Sciences CMLS,* 57(12), pp. 1748-1769.

Richards, C.S. and Shade, P.C. (1987) 'The genetic variation of compatibility in *Biomphalaria glabrata* and *Schistosoma mansoni*', *The Journal of parasitology,* , pp. 1146-1151.

Rivero-Wendt, C.L., Borges, A.C., Oliveira-Filho, E.C., Miranda-Vilela, A.L., Ferreira, M.F. and Grisolia, C.K. (2014) 'Effects of 17alpha-methyltestosterone on the reproduction of the freshwater snail *Biomphalaria glabrata*', *Genetics and molecular research : GMR,* 13(1), pp. 605-615.

Rivero-Wendt, C.L., Miranda-Vilela, A.L., Ferreira, M.F., Amorim, F.S., da Silva, V.A., Louvandini, H. and Grisolia, C.K. (2013) 'Lack of genotoxicity in Astyanax bimaculatus and Oreochromis niloticus of 17alpha-methyltestosterone used in fish hatcheries to produce male monosex populations', *Genetics and molecular research : GMR,* 12(4), pp. 5013-5022.

Roberts, D.B. (2006) '*Drosophila melanogaster*: the model organism', *Entomologia Experimentalis et Applicata,* 121(2), pp. 93-103.

Robinson-Rechavi, M., Garcia, H.E. and Laudet, V. (2003b) 'The nuclear receptor superfamily',

Journal of cell science, 116(4), pp. 585-586.

Rodet, F., Lelong, C., Dubos, M., Costil, K. and Favrel, P. (2005) 'Molecular cloning of a molluscan gonadotropin-releasing hormone receptor orthologue specifically expressed in the gonad', *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression,* 1730(3), pp. 187-195.

Rohlack, S. (1959) 'Über das Vorkommen von Sexualhormonen bei der Meeresschnecke Littorina littorea L.', *Zeitschrift für vergleichende Physiologie,* 42(2), pp. 164-180.

Rollinson, D. (2011) '*Biomphalaria*: Natural history, ecology and schistosome transmission', in *Biomphalaria Snails and Larval Trematodes.* Springer, pp. 57-79.

Ronis, M. and Mason, A. (1996) 'The metabolism of testosterone by the periwinkle (*Littorina littorea*) *in vitro* and *in vivo*: Effects of tributyl tin', *Marine environmental research,* 42(1), pp. 161-166.

Ronquist 1, F. and Huelsenbeck2, J. (2003) 'MrBayes 3: Bayesian phylogenetic inference under mixed models', *Bioinformatics,* 19, pp. 1572-1574.

Rosenfeld, M.G., Lunyak, V.V. and Glass, C.K. (2006) 'Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response', *Genes & development,* 20(11), pp. 1405-1428.

Rosoff, M.L., Doble, K.E., Price, D.A. and Li, C. (1993) 'The *flp-1* propeptide is processed into multiple, highly similar FMRFamide-like peptides in *Caenorhabditis elegans*', *Peptides,* 14(2), pp. 331-338.

Rőszer, T., Kiss-Tóth, É, Petkó, M., Szentmiklósi, A.J. and Bánfalvi, G. (2006) 'Phe-met-arg-phe (FMRF)-amide is a substrate source of NO synthase in the gastropod nervous system', *Cell and tissue research,* 325(3), pp. 567-575.

Rusakov, Y.I., Kolychev, A., Shipilov, V. and Bondareva, V. (2003) 'Insulin-like peptides of the cerebropleural ganglion of the mollusc Anodonta cygnea: isolation, purification, and radioligand analysis', *Journal of Evolutionary Biochemistry and Physiology,* 39(4), pp. 425-432.

Russell, W.M.S., Burch, R.L. and Hume, C.W. (1959) 'The principles of humane experimental technique', .

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A. and Barrell, B. (2000) 'Artemis: sequence visualization and annotation', *Bioinformatics (Oxford, England),* 16(10), pp. 944-945.

Sakr, A.A., Osman, G.Y. and Abo-Shafey, A.E. (1992) 'Effect of testosterone on the ovotestis of the land snail Theba pisana', *Functional and developmental morphology,* 2(2), pp. 99-101.

Salice, C.J., Miller, T.J. and Roesijadi, G. (2009) 'Demographic responses to multigeneration cadmium exposure in two strains of the freshwater gastropod, *Biomphalaria glabrata*', *Archives of Environmental Contamination and Toxicology,* 56(4), pp. 785-795.

Salzet, M., Bulet, P., Dorsselaer, A. and Malecha, J. (1993) 'Isolation, structural characterization

and biological function of a lysine- conopressin in the central nervous system of the Pharyngobdellid leech *Erpobdella octoculata*', *European Journal of Biochemistry,* 217(3), pp. 897-903.

Salzet, M., Verger-Bocquet, M., Vandenbulcke, F. and Van Minnen, J. (1997) 'Leech egg-laying-like hormone: structure, neuronal distribution and phylogeny', *Molecular Brain Research,* 49(1), pp. 211-221.

Santos, M.M., Armanda Reis-Henriques, M., Natividade Vieira, M. and Solé, M. (2006) 'Triphenyltin and tributyltin, single and in combination, promote imposex in the gastropod *Bolinus brandaris*', *Ecotoxicology and environmental safety,* 64(2), pp. 155-162.

Santos, M., Castro, L.F.C., Vieira, M., Micael, J., Morabito, R., Massanisso, P. and Reis-Henriques, M. (2005) 'New insights into the mechanism of imposex induction in the dogwhelk *Nucella lapillus*', *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology,* 141(1), pp. 101-109.

Saunders, S.E., Bright, K., Kellett, E., Benjamin, P.R. and Burke, J.F. (1991) 'Neuropeptides Gly-Asp-Pro-Phe-Leu-Arg-Phe-amide (GDPFLRFamide) and Ser-Asp-Pro-Phe-Leu-Arg-Phe-amide (SDPFLRFamide) are encoded by an exon 3' to Phe-Met-Arg-Phe-NH2 (FMRFamide) in the snail *Lymnaea stagnalis*', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 11(3), pp. 740-745.

Schaefer, M., Picciotto, M.R., Kreiner, T., Kaldany, R., Taussig, R. and Scheller, R.H. (1985) '*Aplysia* neurons express a gene encoding multiple FMRFamide neuropeptides', *Cell,* 41(2), pp. 457-467.

Scheide, J.I. and Dietz, T.H. (1986) 'Serotonin regulation of gill cAMP production, Na, and water uptake in freshwater mussels', *Journal of Experimental Zoology,* 240(3), pp. 309-314.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) 'Quantitative monitoring of gene expression patterns with a complementary DNA microarray', *Science (New York, N.Y.),* 270(5235), pp. 467-470.

Scherer, L.J. and Rossi, J.J. (2003) 'Approaches for the sequence-specific knockdown of mRNA', *Nature biotechnology,* 21(12), pp. 1457-1465.

Scheuring, C.F. (2008) 'Construction of bacterial artificial chromosome libraries for Zhikong Scallop *Chlamys farreri*', *Chinese Journal of Oceanology and Limnology,* 26(2), pp. 215-218.

Schneider, L.E. and Taghert, P.H. (1988) 'Isolation and characterization of a *Drosophila* gene that encodes multiple neuropeptides related to Phe-Met-Arg-Phe-NH2 (FMRFamide)', *Proceedings of the National Academy of Sciences of the United States of America,* 85(6), pp. 1993-1997.

Schrager, S. and Potter, B.E. (2004) 'Diethylstilbestrol exposure', *American Family Physician,* 69(10), pp. 2395-2400.

Schulte-Oehlmann, U., Oetken, M., Bachmann, J. and Oehlmann, J. (2004) 'Effects of ethinyloestradiol and methyltestosterone in prosobranch snails', in *Pharmaceuticals in the*

Environment. Springer, pp. 233-247.

Schulte-Oehlmann, U., Albanis, T., Allera, A., Bachmann, J., Berntsson, P., Beresford, N., Carnevali, D.C., Ciceri, F., Dagnac, T. and Falandysz, J. (2006) 'COMPRENDO: Focus and approach', *Environmental health perspectives,* 114(S-1), pp. 98.

Schulte-Oehlmann, U., Tillmann, M., Casey, D., Duft, M., Markert, B. and Oehlmann, J. (2001) 'Östrogenartige wirkungen von bisphenol a auf vorderkiemenschnecken (Mollusca: Gastropoda: Prosobranchia)', *Umweltwissenschaften und Schadstoff-Forschung,* 13(6), pp. 319-333.

Scott, A.P. (2012b) 'Do mollusks use vertebrate sex steroids as reproductive hormones? II. Critical review of the evidence that steroids have biological effects', *Steroids, .*

Scott, A.P. (2012a) 'Do mollusks use vertebrate sex steroids as reproductive hormones? I. Critical appraisal of the evidence for the presence, biosynthesis and uptake of steroids', *Steroids, .*

Segner, H., Caroll, K., Fenske, M., Janssen, C., Maack, G., Pascoe, D., Schäfers, C., Vandenbergh, G., Watts, M. and Wenzel, A. (2003) 'Identification of endocrine-disrupting effects in aquatic vertebrates and invertebrates: report from the European IDEA project', *Ecotoxicology and environmental safety,* 54(3), pp. 302-314.

Seki, M., Yokota, H., Matsubara, H., Maeda, M., Tadokoro, H. and Kobayashi, K. (2004) 'Fish full life- cycle testing for androgen methyltestosterone on medaka (*Oryzias latipes*)', *Environmental Toxicology and Chemistry,* 23(3), pp. 774-781.

Semenova, M.N., Kiselyov, A. and Semenov, V.V. (2006) 'Sea urchin embryo as a model organism for the rapid functional screening of tubulin modulators', *BioTechniques,* 40(6), pp. 765.

Shang, Y., Myers, M. and Brown, M. (2002) 'Formation of the androgen receptor transcription complex', *Molecular cell,* 9(3), pp. 601-610.

Sharp, P.A. (2001) 'RNA interference--2001', *Genes & development,* 15(5), pp. 485-490.

Shen, W., Moore, C.C., Ikeda, Y., Parker, K.L. and Ingraham, H.A. (1994) 'Nuclear receptor steroidogenic factor 1 regulates the Müllerian inhibiting substance gene: a link to the sex determination cascade', *Cell,* 77(5), pp. 651-661.

Shin, J., Moon, H.J., Kang, I.H., Kim, T.S., Lee, S.J., Ahn, J.Y., Bae, H., Jeung, E.B. and Han, S.Y. (2007) 'OECD validation of the rodent Hershberger assay using three reference chemicals; 17α-methyltestosterone, procymidone, and p, p′-DDE', *Archives of Toxicology,* 81(5), pp. 309-318.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992) 'Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector', *Proceedings of the National Academy of Sciences of the United States of America,* 89(18), pp. 8794-8797.

Shizuya, H. and Kouros-Mehr, H. (2001) 'The development and applications of the bacterial artificial chromosome cloning system', *The Keio journal of medicine,* 50(1), pp. 26-30.

Siah, A., Pellerin, J., Benosman, A., Gagné, J. and Amiard, J. (2002) 'Seasonal gonad progesterone pattern in the soft-shell clam *Mya arenaria*', *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology,* 132(2), pp. 499-511.

Simakov, O., Marletaz, F., Cho, S., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D., Larsson, T., Lv, J. and Arendt, D. (2012) 'Insights into bilaterian evolution from three spiralian genomes', *Nature, .*

Skakkebaek, N.E., Rajpert-De Meyts, E. and Main, K.M. (2001) 'Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects', *Human reproduction (Oxford, England),* 16(5), pp. 972-978.

Sladek, F.M. (2011) 'What are nuclear receptor ligands?', *Molecular and cellular endocrinology,* 334(1), pp. 3-13.

Sluder, A.E. and Maina, C.V. (2001) 'Nuclear receptors in nematodes: themes and variations', *TRENDS in Genetics,* 17(4), pp. 206-213.

Smirnov, A. (2002) 'Nuclear receptors: nomenclature, ligands, mechanisms of their effects on gene expression', *Biochemistry (Moscow),* 67(9), pp. 957-977.

Smit, A., Spijker, S., Van Minnen, J., Burke, J., De Winter, F., Van Elk, R. and Geraerts, W. (1996) 'Expression and characterization of molluscan insulin-related peptide VII from the mollusc *Lymnaea stagnalis*', *Neuroscience,* 70(2), pp. 589-596.

Smit, A., Van Kesteren, R., Li, K., Van Minnen, J., Spijker, S., Van Heerikhuizen, H. and Geraerts, W. (1998) 'Towards understanding the role of insulin in the brain: lessons from insulin-related signaling systems in the invertebrate brain', *Progress in neurobiology,* 54(1), pp. 35-54.

Smit, A.B., van Marle, A., van Elk, R., Bogerd, J., van Heerikhuizen, H. and Geraerts, W.P. (1993) 'Evolutionary conservation of the insulin gene structure in invertebrates: cloning of the gene encoding molluscan insulin-related peptide III from *Lymnaea stagnalis*', *Journal of Molecular Endocrinology,* 11(1), pp. 103-113.

Smith, O.W. (1949) 'Diethylstilbestrol in the prevention and treatment of complications of pregnancy.', *Obstetrical & gynecological survey,* 4(2), pp. 190-191.

Smith, S.A., Wilson, N.G., Goetz, F.E., Feehery, C., Andrade, S.C., Rouse, G.W., Giribet, G. and Dunn, C.W. (2011) 'Resolving the evolutionary relationships of molluscs with phylogenomic tools', *Nature,* 480(7377), pp. 364-367.

Smith, O.W. and Smith, G.V. (1949) 'The influence of diethylstilbestrol on the progress and outcome of pregnancy as based on a comparison of treated with untreated primigravidas', *American Journal of Obstetrics and Gynecology,* 58(5), pp. 994-1009.

Snow, M.I. and Larsen, P.L. (2000) 'Structure and expression of *daf-12*: a nuclear hormone receptor with three isoforms that are involved in development and aging in *Caenorhabditis elegans*', *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression,* 1494(1), pp. 104-116.

Sohoni, P. and Sumpter, J.P. (1998b) 'Several environmental oestrogens are also anti-androgens', *The Journal of endocrinology,* 158(3), pp. 327-339.

Soon, W.W., Hariharan, M. and Snyder, M.P. (2013) 'High- throughput sequencing for biology and medicine', *Molecular systems biology,* 9(1).

Sousa, A., Barroso, C.M., Tanabe, S. and Horiguchi, T. (2010) 'Involvement of retinoid X receptor in imposex development in *Nucella lapillus* and *Nassarius reticulatus*–preliminary results', *Interdisciplinary studies on environmental chemistry–Biological responses to contaminants,* 3, pp. 189-196.

Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L. and Sweedler, J.V. (2006) 'NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides', *Nucleic acids research,* 34(Web Server issue), pp. W267-72.

Southey, B.R., Sweedler, J.V. and Rodriguez-Zas, S.L. (2008) 'Prediction of neuropeptide cleavage sites in insects', *Bioinformatics (Oxford, England),* 24(6), pp. 815-825.

Sower, S.A., Decatur, W.A., Joseph, N.T. and Freamat, M. (2012) 'Evolution of vertebrate GnRH receptors from the perspective of a basal vertebrate', *Frontiers in endocrinology,* 3.

Spijker, S., Smit, A.B., Eipper, B.A., Malik, A., Mains, R.E. and Geraerts, W.P. (1999) 'A molluscan peptide alpha-amidating enzyme precursor that generates five distinct enzymes', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology,* 13(6), pp. 735-748.

Spooner, N., Gibbs, P., Bryan, G. and Goad, L. (1991) 'The effect of tributyltin upon steroid titres in the female dogwhelk, *Nucella lapillus*, and the development of imposex', *Marine environmental research,* 32(1), pp. 37-49.

Srivastava, M., Begovic, E., Chapman, J., Putnam, N.H., Hellsten, U., Kawashima, T., Kuo, A., Mitros, T., Salamov, A. and Carpenter, M.L. (2008) 'The Trichoplax genome and the nature of placozoans', *Nature,* 454(7207), pp. 955-960.

Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M.E., Mitros, T., Richards, G.S., Conaco, C., Dacre, M. and Hellsten, U. (2010) 'The Amphimedon queenslandica genome and the evolution of animal complexity', *Nature,* 466(7307), pp. 720-726.

Stange, D., Sieratowicz, A., Horres, R. and Oehlmann, J. (2012) 'Freshwater mudsnail (*Potamopyrgus antipodarum*) estrogen receptor: Identification and expression analysis under exposure to (xeno-) hormones', *Ecotoxicology and environmental safety,* 75, pp. 94-101.

Stefano, G.B., Zhao, X., Bailey, D., Metlay, M. and Leung, M.K. (1989) 'High affinity dopamine binding to mouse thymocytes and *Mytilus edulis* (Bivalvia) hemocytes', *Journal of neuroimmunology,* 21(1), pp. 67-74.

Sternberg, R.M. (2007) *The Roles of Androgen, Estrogen, and Retinoid Signaling in Reproductive Recrudescence of the Eastern Mud Snail (Ilyanassa Obsoleta): Implications for the Mechanism of Tributyltin-induced Imposex.* ProQuest.

Sternberg, R.M., Hotchkiss, A.K. and LeBlanc, G.A. (2008) 'The contribution of steroidal

androgens and estrogens to reproductive maturation of the eastern mud snail *Ilyanassa obsoleta*', *General and comparative endocrinology,* 156(1), pp. 15-26.

Stöger, I., Sigwart, J., Kano, Y., Knebelsberger, T., Marshall, B., Schwabe, E. and Schrödl, M. (2013) 'The Continuing Debate on Deep Molluscan Phylogeny: Evidence for Serialia (Mollusca, Monoplacophora', *BioMed research international,* 2013.

Stopper, H., Schmitt, E. and Kobras, K. (2005) 'Genotoxicity of phytoestrogens', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 574(1), pp. 139-155.

Strand, F.L., Rose, K.J., Zuccarelli, L.A., Kume, J., Alves, S.E., Antonawich, F.J. and Garrett, L.Y. (1991) 'Neuropeptide hormones as neurotrophic factors', *Physiological Reviews,* 71(4), pp. 1017-1046.

Struhsaker, J.W. and Costlow, J.D. (1968) 'Larval development of *Littorina picta* (Prosobranchia, Mesogastropoda), reared in the laboratory', *Journal of Molluscan Studies,* 38(2), pp. 153-160.

Swofford, D.L. *PAUP*: Phylogenetic Analysis using Parsimony (*and Other Methods). Version 4.*(Version 4) [Computer program].

Szakács, G., Paterson, J.K., Ludwig, J.A., Booth-Genthe, C. and Gottesman, M.M. (2006) 'Targeting multidrug resistance in cancer', *Nature Reviews Drug Discovery,* 5(3), pp. 219-234.

Tabb, M.M. and Blumberg, B. (2006) 'New modes of action for endocrine-disrupting chemicals', *Molecular Endocrinology,* 20(3), pp. 475-482.

Takeda (1979) 'Induction of egg laying by steriod hormones in slugs.', *Comp. Biochem. Physlol.,* 62(A), pp. 273-278.

Tanaka, T. and De Luca, L.M. (2009) 'Therapeutic potential of "rexinoids" in cancer prevention and treatment', *Cancer research,* 69(12), pp. 4945-4947.

Teschendorf, D. and Link, C.D. (2009) 'What have worm models told us about the mechanisms of neuronal dysfunction in human neurodegenerative diseases', *Mol Neurodegener,* 4(1), pp. 1-13.

Thain, M. & Hickman, M. (1995) *The Penguin Dictionary of Biology.* 9th ed. edn. London, UK.: Penguin Books,.

Thomas, J., Nwanko, D. and Sterry, P. (1985) 'The feeding strategies of juvenile and adult *Biomphalaria glabrata* (Say) under simulated natural conditions and their relevance to ecological theory and snail control', *Proceedings of the Royal Society of London.Series B, Biological Sciences,* , pp. 177-209.

Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A. and Narechania, A. (2003) 'PANTHER: a library of protein families and subfamilies indexed by function', *Genome research,* 13(9), pp. 2129-2141.

Thompson, J., Gibson1, T., Plewniak, F., Jeanmougin, F. and Higgins, D. (1997) 'The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools', *Nucleic Acids Res.,* 25, pp. 4876-4882.

Thomson, S.A., Baldwin, W.S., Wang, Y.H., Kwon, G. and LeBlanc, G.A. (2009) 'Annotation, phylogenetics, and expression of the nuclear receptors in *Daphnia pulex*', *BMC genomics,* 10(1), pp. 500.

Thornton, J.W. and DeSalle, R. (2000) 'Gene family evolution and homology: genomics meets phylogenetics', *Annual review of genomics and human genetics,* 1(1), pp. 41-73.

Thornton, J.W., Need, E. and Crews, D. (2003) 'Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling', *Sci. Signal.,* 301(5640), pp. 1714.

Thummel, C. (1996) 'Flies on steroids - *Drosophila* metamorphosis and the mechanisms of steroid hormone action', *Trends in Genetics,* 12(8), pp. 306-310.

Tillmann, M., Schulte-Oehlmann, U., Duft, M., Markert, B. and Oehlmann, J. (2001) 'Effects of endocrine disruptors on prosobranch snails (Mollusca: Gastropoda) in the laboratory. Part III: Cyproterone acetate and vinclozolin as antiandrogens', *Ecotoxicology,* 10(6), pp. 373-388.

Timmons, L., Tabara, H., Mello, C.C. and Fire, A.Z. (2003) 'Inducible systemic RNA silencing in *Caenorhabditis elegans*', *Molecular biology of the cell,* 14(7), pp. 2972-2983.

Tinwell, H., Friry-Santini, C., Rouquie, D., Belluco, S., Elies, L., Pallen, C. and Bars, R. (2007) 'Evaluation of the antiandrogenic effects of flutamide, DDE, and linuron in the weanling rat assay using organ weight, histopathological, and proteomic approaches', *Toxicological sciences : an official journal of the Society of Toxicology,* 100(1), pp. 54-65.

Titus-Ernstoff, L., Troisi, R., Hatch, E.E., Palmer, J.R., Hyer, M., Kaufman, R., Adam, E., Noller, K. and Hoover, R. (2010) 'Birth defects in the sons and daughters of women who were exposed in utero to diethylstilbestrol (DES)', *International journal of andrology,* 33(2), pp. 377-384.

Titus-Ernstoff, L., Troisi, R., Hatch, E.E., Wise, L.A., Palmer, J., Hyer, M., Kaufman, R., Adam, E., Strohsnitter, W., Noller, K., Herbst, A.L., Gibson-Chambers, J., Hartge, P. and Hoover, R.N. (2006) 'Menstrual and reproductive characteristics of women whose mothers were exposed in utero to diethylstilbestrol (DES)', *International journal of epidemiology,* 35(4), pp. 862-868.

Tiwari, A., Pragya, P., Ravi Ram, K. and Chowdhuri, D.K. (2011) 'Environmental chemical mediated male reproductive toxicity: *Drosophila melanogaster* as an alternate animal model', *Theriogenology,* 76(2), pp. 197-216.

Toledo, R. and Fried, B. (2011) *Biomphalaria snails and larval trematodes.* Springer.

Tosti, E., Di Cosmo, A., Cuomo, A., Di Cristo, C. and Gragnaniello, G. (2001) 'Progesterone induces activation in *Octopus vulgaris* spermatozoa', *Molecular reproduction and development,* 59(1), pp. 97-105.

Treen, N., Itoh, N., Miura, H., Kikuchi, I., Ueda, T., Takahashi, K.G., Ubuka, T., Yamamoto, K., Sharp, P.J. and Tsutsui, K. (2012) 'Mollusc gonadotropin-releasing hormone directly regulates gonadal functions: a primitive endocrine system controlling reproduction', *General and comparative endocrinology,* 176(2), pp. 167-172.

Trigwell, J.A. and Dussart, G. (1998) 'Functional protandry in *Biomphalaria glabrata* (Gastropoda: Pulmonata), an intermediate host of *Schistosoma*', *Journal of Molluscan Studies,* 64(2), pp. 253-256.

Trigwell, J.A., Dussart, G. and Vianey-Liaud, M. (1997) 'Pre-copulatory behaviour of the freshwater hermaphrodite snail *Biomphalaria glabrata* (Say, 1818)(Gastropoda: Pulmonata)', *Journal of Molluscan Studies,* 63(1), pp. 116-120.

Tsai, P. (2006) 'Gonadotropin-releasing hormone in invertebrates: structure, function, and evolution', *General and comparative endocrinology,* 148(1), pp. 48-53.

Tsai, P., Maldonado, T.A. and Lunden, J.B. (2003) 'Localization of gonadotropin-releasing hormone in the central nervous system and a peripheral chemosensory organ of *Aplysia californica*', *General and comparative endocrinology,* 130(1), pp. 20-28.

Turner, C., Sawle, A., Fenske, M. and Cossins, A. (2012) 'Implications of the solvent vehicles dimethylformamide and dimethylsulfoxide for establishing transcriptomic endpoints in the zebrafish embryo toxicity test', *Environmental Toxicology and Chemistry,* 31(3), pp. 593-604.

Tyler, C., Jobling, S. and Sumpter, J. (1998) 'Endocrine disruption in wildlife: a critical review of the evidence', *CRC critical reviews in toxicology,* 28(4), pp. 319-361.

Ueda, H., Sun, G.C., Murata, T. and Hirose, S. (1992) 'A Novel Dna-Binding Motif Abuts the Zinc Finger Domain of Insect Nuclear Hormone Receptor Ftz-F1 and Mouse Embryonal Long Terminal Repeat-Binding Protein', *Molecular and cellular biology,* 12(12), pp. 5667-5672.

Ugriumov, M.V. (2009) 'Endocrine functions of the brain in adult and developing mammals', *Ontogenez,* 40(1), pp. 19-29.

Ulrich, R.G. (2003) 'The toxicogenomics of nuclear receptor agonists', *Current opinion in chemical biology,* 7(4), pp. 505-510.

Van Duivenboden, Y. and Maat, A.T. (1985) 'Masculinity and receptivity in the hermaphrodite pond snail, *Lymnaea stagnalis*', *Animal Behaviour,* 33(3), pp. 885-891.

Van Golen, F.A., Li, K.W., Chen, S., Jiménez, C.R. and Geraerts, W.P. (1996) 'Various isoforms of myomodulin identified from the male copulatory organ of *Lymnaea* show overlapping yet distinct modulatory effects on the penis muscle', *Journal of neurochemistry,* 66(1), pp. 321-329.

van Golen, F.A., Li, K.W., de Lange, R.P., Jespersen, S. and Geraerts, W.P. (1995) 'Mutually exclusive neuronal expression of peptides encoded by the FMRFa gene underlies a differential control of copulation in *Lymnaea*', *The Journal of biological chemistry,* 270(47), pp. 28487-28493.

Van Kesteren, R., Tensen, C., Smit, A., Van Minnen, J., Van Soest, P., Kits, K., Meyerhof, W., Richter, D., Van Heerikhuizen, H. and Vreugdenhil, E. (1995a) 'A novel G protein-coupled receptor mediating both vasopressin-and oxytocin-like functions of Lys-conopressin in *Lymnaea stagnalis*', *Neuron,* 15(4), pp. 897-908.

Van Kesteren, R.E., Smit, A.B., De Lange, R.P., Kits, K.S., Van Golen, F.A., Van Der Schors,

R.C., De With, N.D., Burke, J.F. and Geraerts, W.P. (1995b) 'Structural and functional evolution of the vasopressin/oxytocin superfamily: vasopressin-related conopressin is the only member present in *Lymnaea*, and is involved in the control of sexual behavior', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 15(9), pp. 5989-5998.

Van Minnen, J., Smit, A. and Joosse, J. (1989) 'Central and peripheral expression of genes coding for egg-laying inducing and insulin-related peptides in a snail', *Arch.Histol.Cytol,* 52, pp. 241-252.

Van Soest, P.E. and Kits, K.S. (2002) 'Role of Lys-Conopressin in the Control of Male Sexual Behavior in *Lymnaea stagnalis*', *Hormones, Brain and Behavior,* 3, pp. 317-330.

Vasta, J.D., LeSage, K.E. and Fried, B. (2011) 'A simple method for culturing neonatal Biomphalaria glabrata snails', *Journal of Parasitology,* 97(4), pp. 746-747.

Veenstra, J.A. (2011) 'Neuropeptide evolution: Neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*', *General and comparative endocrinology,* 171(2), pp. 160-175.

Veenstra, J.A. (2010) 'Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects', *General and comparative endocrinology,* 167(1), pp. 86-103.

Veenstra, J.A. (2000) 'Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors', *Archives of Insect Biochemistry and Physiology,* 43(2), pp. 49-63.

Velarde, R.A., Robinson, G.E. and Fahrbach, S.E. (2006) 'Nuclear receptors of the honey bee: annotation and expression in the adult brain', *Insect molecular biology,* 15(5), pp. 583-595.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C. and Broder, S.e.a. (2001) 'The sequence of the human genome', *Science (New York, N.Y.),* 291(5507), pp. 1304-1351.

Vernon, J. and Taylor, J.K. (1996) 'Patterns of sexual roles adopted by the schistosome-vector snail *Biomphalaria glabrata* (Planorbidae)', *Journal of Molluscan Studies,* 62(2), pp. 235-241.

Vianey-Liaud, M. and Dussart, G. (2002) 'Aspects of pairing and reproduction in the hermaphrodite freshwater snail *Biomphalaria glabrata* (Gastropoda: Pulmonata)', *Journal of Molluscan Studies,* 68(3), pp. 243-248.

Vianey-Liaud, M., Dominique, J. and Dussart G (1996) 'Sperm competition in the simultaneous hermaphrodite freshwater snail *Biomphalaria glabrata* (Gastropoda: Pulmonata)', *Journal of Molluscan Studies,* 62(4), pp. 451-457.

Vianey-Liaud, M., Nassi, H., Lancastre, F. and Dupouy, J. (1989) 'Duration of pairing and use of

allosperm in *Biomphalaria glabrata* (Gastropoda: Planorbidae)', *Memórias do Instituto Oswaldo Cruz,* 84(1), pp. 41-45.

Vick, A.M. and Hayton, W.L. (2001) 'Methyltestosterone pharmacolinetics and oral bioavailability in rainbow trout (*Oncorhynchus mykiss*)', *Aquatic toxicology,* 52(3), pp. 177-188.

Viglietti-Panzica, C., Montoncello, B., Mura, E., Pessatti, M. and Panzica, G. (2005) 'Organizational effects of diethylstilbestrol on brain vasotocin and sexual behavior in male quail', *Brain research bulletin,* 65(3), pp. 225-233.

Vogeler, S., Galloway, T.S., Lyons, B.P. and Bean, T.P. (2014a) 'The nuclear receptor gene family in the Pacific oyster, Crassostrea gigas, contains a novel subfamily group', *BMC genomics,* 15(1), pp. 369.

Vogeler, S., Galloway, T.S., Lyons, B.P. and Bean, T.P. (2014b) 'The nuclear receptor gene family in the Pacific oyster, *Crassostrea gigas*, contains a novel subfamily group', *BMC genomics,* 15(1), pp. 369.

Vos, J.G., Dybing, E., Greim, H.A., Ladefoged, O., Lambré, C., Tarazona, J.V., Brandt, I. and Vethaak, A.D. (2000) 'Health effects of endocrine-disrupting chemicals on wildlife, with special reference to the European situation', *CRC critical reviews in toxicology,* 30(1), pp. 71-133.

Vreugdenhil, E., Jackson, J.F., Bouwmeester, T., Smit, A.B., Van Minnen, J., Van Heerikhuizen, H., Klootwijk, J. and Joosse, J. (1988) 'Isolation, characterization, and evolutionary aspects of a cDNA clone encoding multiple neuropeptides involved in the stereotyped egg-laying behavior of the freshwater snail *Lymnaea stagnalis*', *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 8(11), pp. 4184-4191.

Walker, W.H. (2010a) 'Non-classical actions of testosterone and spermatogenesis', *Philosophical transactions of the Royal Society of London.Series B, Biological sciences,* 365(1546), pp. 1557-1569.

Walker, W.H. (2010b) 'Non-classical actions of testosterone and spermatogenesis', *Philosophical transactions of the Royal Society of London.Series B, Biological sciences,* 365(1546), pp. 1557-1569.

Wallace, C. (1992) 'Parthenogenesis, sex and chromosomes in Potamopyrgus', *Journal of Molluscan Studies,* 58(2), pp. 93-107.

Wang, C. and Croll, R.P. (2004) 'Effects of sex steroids on gonadal development and gender determination in the sea scallop, *Placopecten magellanicus*', *Aquaculture,* 238(1), pp. 483-498.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W. and Orkin, S.H. (2006) 'A protein interaction network for pluripotency of embryonic stem cells', *Nature,* 444(7117), pp. 364-368.

Wang, Z., Zhou, X.E., Motola, D.L., Gao, X., Suino-Powell, K., Conneely, A., Ogata, C., Sharma, K.K., Auchus, R.J., Lok, J.B., Hawdon, J.M., Kliewer, S.A., Xu, H.E. and Mangelsdorf, D.J.

(2009) 'Identification of the nuclear receptor DAF-12 as a therapeutic target in parasitic nematodes', *Proceedings of the National Academy of Sciences of the United States of America,* 106(23), pp. 9138-9143.

Wärnmark, A., Treuter, E., Wright, A.P. and Gustafsson, J. (2003) 'Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation', *Molecular endocrinology,* 17(10), pp. 1901-1909.

Wason, S., Pohlmeyer-Esch, G., Pallen, C., Palazzi, X., Espuña, G. and Bars, R. (2003) '17α-methyltestosterone: 28-day oral toxicity study in the rat based on the "enhanced OECD test guideline 407" to detect endocrine effects', *Toxicology,* 192(2), pp. 119-137.

Wasson, K.M., Gower, B.A., Hines, G.A. and Watts, S.A. (2000) 'Levels of progesterone, testosterone, and estradiol, and androstenedione metabolism in the gonads of *Lytechinus variegatus* (Echinodermata: Echinoidea)', *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology,* 126(2), pp. 153-165.

Wasson, K.M. and Watts, S.A. (1998b) 'Proscar (Finasteride) inhibits *5 α*-reductase activity in the ovaries and testes of *Lytechinus variegatus* Lamarck (Echinodermata: Echinoidea)', *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology,* 120(3), pp. 425-431.

Watermann, B., Thomsen, A., Kolodzey, H., Daehne, B., Meemken, M., Pijanowska, U. and Liebezeit, G. (2008) 'Histopathological lesions of molluscs in the harbour of Norderney, Lower Saxony, North Sea (Germany)', *Helgoland Marine Research,* 62(2), pp. 167-175.

Webster, J.P., Oliviera, G., Rollinson, D. and Gower, C.M. (2010) 'Schistosome genomes: a wealth of information', *Trends in parasitology,* 26(3), pp. 103-106.

White, M.M., Chejlava, M., Fried, B. and Sherma, J. (2007) 'The concentration of calcium carbonate in shells of freshwater snails', *American Malacological Bulletin,* 22(1), pp. 139-142.

Whitehead, D.L. and Sellheyer, K. (1982) 'The Identification of Ecdysterone (20-Hydroxyecdysone) in 3 Species of Mollusks (Gastropoda, Pulmonata)', *Experientia,* 38(10).

Wilber, C.G. (1964) 'Some Contributions of Exotic Animals to Biomedical Research', *Bioscience,* 14(10), pp. 21-24.

Williams, R.S., Boeckeler, K., Gräf, R., Müller-Taubenberger, A., Li, Z., Isberg, R.R., Wessels, D., Soll, D.R., Alexander, H. and Alexander, S. (2006) 'Towards a molecular understanding of human diseases using *Dictyostelium discoideum*', *Trends in molecular medicine,* 12(9), pp. 415-424.

Willows, A.O., Pavlova, G.A. and Phillips, N.E. (1997) 'Modulation of ciliary beat frequency by neuropeptides from identified molluscan neurons', *The Journal of experimental biology,* 200(Pt 10), pp. 1433-1439.

Wu, W., Niles, E.G. and LoVerde, P.T. (2007) 'Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*', *BMC Evol Biol,* 7(1), pp. 150.

Wu, W., Niles, E., Hirai, H. and LoVerde, P. (2007) 'Evolution of a novel subfamily of nuclear receptors with members that each contain two DNA binding domains', *BMC Evol Biol,* 7(1), pp. 27.

Wu, T., Buck, G.M. and Mendola, P. (2003) 'Blood lead levels and sexual maturation in U.S. girls: the Third National Health and Nutrition Examination Survey, 1988-1994', *Environmental health perspectives,* 111(5), pp. 737-741.

Wu, W., Niles, E., El-Sayed, N., Berriman, M. and LoVerde, P. (2006) '*Schistosoma mansoni* (Platyhelminthes, Trematoda) nuclear receptors: Sixteen new members and a novel subfamily', *Gene,* 366(2), pp. 303-315.

Xie, C., Jeong, Y., Fu, M., Bookout, A.L., Garcia-Barrio, M.T., Sun, T., Kim, B., Xie, Y., Root, S. and Zhang, J. (2009) 'Expression profiling of nuclear receptors in human and mouse embryonic stem cells', *Molecular endocrinology,* 23(5), pp. 724-733.

Yamamoto, M., Shirai, M., Tamura, A., KOBAYASHI, T., KOHARA, S., MURAKAMI, M. and ARISHIMA, K. (2005) 'Effects of maternal exposure to a low dose of diethylstilbestrol on sexual dimorphic nucleus volume and male reproductive system in rat offspring', *The Journal of toxicological sciences,* 30(1), pp. 7-18.

Yang, X., Downes, M., Yu, R.T., Bookout, A.L., He, W., Straume, M., Mangelsdorf, D.J. and Evans, R.M. (2006) 'Nuclear receptor expression links the circadian clock to metabolism', *Cell,* 126(4), pp. 801-810.

York, P.S., Cummins, S.F., Degnan, S.M., Woodcroft, B.J. and Degnan, B.M. (2012) 'Marked changes in neuropeptide expression accompany broadcast spawnings in the gastropod *Haliotis asinina*', *Frontiers in zoology,* 9(1), pp. 9-9994-9-9.

Yoshida, K., Hoshikawa, K., Wada, T. and Yusa, Y. (2013) 'Patterns of density dependence in growth, reproduction and survival in the invasive freshwater snail Pomacea canaliculata in Japanese rice fields', *Freshwater Biology,* 58(10), pp. 2065-2073.

Young, K.G., Chang, J.P. and Goldberg, J.I. (1999) 'Gonadotropin-releasing hormone neuronal system of the freshwater snails *Helisoma trivolvis* and *Lymnaea stagnalis*: possible involvement in reproduction', *Journal of comparative neurology,* 404(4), pp. 427-437.

Yu, R.T., Mckeown, M., Evans, R.M. and Umesono, K. (1994) 'Relationship between *Drosophila* Gap Gene Tailless and a Vertebrate Nuclear Receptor Tlx', *Nature,* 370(6488).

Zachar, N. and Neiman, M. (2013) 'Profound Effects of Population Density on Fitness-Related Traits in an Invasive Freshwater Snail', *PloS one,* 8(11), pp. e80067.

Zanetti, S., Grinschgl, S., Meola, M., Belfiore, M., Rey, S., Bianchi, P. and Puoti, A. (2012) 'The sperm-oocyte switch in the *C. elegans* hermaphrodite is controlled through steady-state levels of the fem-3 mRNA', *RNA (New York, N.Y.),* 18(7), pp. 1385-1394.

Zar, J. (1999) 'Biological statistics', .

Zatylny-Gaudin, C. and Favrel, P. (2014) 'Diversity of the RFamide peptide family in mollusks', *Frontiers in endocrinology,* 5.

Zhang, H., Pan, L. and Zhang, L. (2012) 'Molecular cloning and characterization of estrogen receptor gene in the Scallop *Chlamys farreri* : Expression profiles in response to endocrine disrupting chemicals', *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology,* 156(1), pp. 51-57.

Zhang, L. and Baer, K. (2000) 'The influence of feeding, photoperiod and selected solvents on the reproductive strategies of the water flea, *Daphnia magna*', *Environmental Pollution,* 110(3), pp. 425-430.

Zhang, L., Tello, J.A., Zhang, W. and Tsai, P. (2008) 'Molecular cloning, expression pattern, and immunocytochemical localization of a gonadotropin-releasing hormone-like molecule in the gastropod mollusk, *Aplysia californica*', *General and comparative endocrinology,* 156(2), pp. 201-209.

Zhang, S., Nian, H., Wang, B., Loker, E.S. and Adema, C.M. (2009) 'Schistosomin from the snail *Biomphalaria glabrata*: Expression studies suggest no involvement in trematode-mediated castration', *Molecular and biochemical parasitology,* 165(1), pp. 79-86.

Zhang, Z., Burch, P.E., Cooney, A.J., Lanz, R.B., Pereira, F.A., Wu, J., Gibbs, R.A., Weinstock, G. and Wheeler, D.A. (2004) 'Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome', *Genome research,* 14(4), pp. 580-590.

Zhang, J.V., Ren, P.G., Avsian-Kretchmer, O., Luo, C.W., Rauch, R., Klein, C. and Hsueh, A.J. (2005) 'Obestatin, a peptide encoded by the ghrelin gene, opposes ghrelin's effects on food intake', *Science (New York, N.Y.),* 310(5750), pp. 996-999.

Zuang, V., Schäffer, M., Tuomainen, A.M., Amcoff, P., Bernasconi, C., Bremer, S., Casati, S., Castello, P., Coecke, S. and Corvi, R. (2013) 'EURL ECVAM progress report on the development, validation and regulatory acceptance of alternative methods (2010-2013)', .

Zulfahmi, S., Yazan, L.S., Ithnin, H. and Armania, N. (2013) 'The improvement of *in vivo* model (Balb/c mice) for cervical carcinogenesis using diethylstilbestrol (DES)', *Experimental and Toxicologic Pathology,* 65(7), pp. 1083-1089.

Websites

Annual Statistics of Scientific Procedures on Living Animals (2012). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/212610/spanimals12/

Artemis. www.sanger.ac.uk.

ASPA Draft (2013). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/116843/aspa-draft-guidance.pdf.

Blast2GO. http://www.blast2go.com/b2ghome.

*BUAV*http://www.buav.org/international-work.

Clustal. http://www.clustal.org/.

The COMPRENDO Project report. http://ec.europa.eu/research/environment/pdf/env_health_projects/chemicals/c-comprendo.pdf.

EDSTAT Report (1998). http://www.epa.gov/endo/pubs/edstac/chap5v14.pdf.

EPA EDSTAC. http://www.epa.gov/endo/pubs/edspoverview/edstac.htm.

EU Directive (2010). http://ec.europa.eu/environment/chemicals/lab_animals/3r/alternative_en.htm.

ExPASy. http://www.expasy.org/.

ImageJ. http://rsb.info.nih.gov.

InterProScan5. http://www.ebi.ac.uk/Tools/pfa/iprscan5/.

Lottia gigantea genome. http://genome.jgi-psf.org/Lotgi1/Lotgi1.info.html.

MEGA. http://www.megasoftware.net.

ModelTest. http://www.molecularevolution.org/software/phylogenetics/modeltest.

NC3Rs. http://www.nc3rs.org.uk.

NCBI Blast. http://blast.ncbi.nlm.nih.gov/Blast.cgi.

NeuroPred Application. http://neuroproteomics.scs.illinois.edu/neuropred.html.

New GENSCAN Web Server at MIT. http://genes.mit.edu/GENSCAN.html.

NIC/NIH Annual Report (2010). (http://deainfo.nci.nih.gov/advisory/pcp/annualReports/pcp08-09rpt/PCP_Report_08-09_508.pdf.

NURSA report (2012). http://www.nursa.org/files/NURSA_Phase%20I_II_Final_Report.pdf.

OECD (2000). http://www.epa.gov/endo/pubs/ref-2_oecd_gd23_difficult_substances.pdf.

OECD (2009). www.oecd.org/document/62/0,2340,en_2649_34377_2348606_1_1_1_1,00.html.

OECD (2010). http://www.oecd.org/chemicalsafety/testing/46801244.pdf.

OligoAnalyzer 3.1. www.eu.idtdna.com/analyzer/Applications/ OligoAnalyzer.

Preliminary Bg Genome data. http://129.24.144.93/blast_bg/2index.html.

Primer3. http://bioinfo.ut.ee/primer3-0.4.0/.

Seqtools. http://www.seqtools.dk/.

SignalP-3.0. http://www.cbs.dtu.dk/services/SignalP-3.0/.

SMART. http://smart.embl-heidelberg.de.

SnailDB. http://snaildb.org/webfm.

SnailDB. http://snaildb.org/webfm.

TMHMM Server v.2.0. http://www.cbs.dtu.dk/services/TMHMM/.

UK Home Office report on animal testing (2011). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/328319/guide-animals-procedures.pdf.

USDA report (2012). http://speakingofresearch.com/2014/02/17/usda-statistics-for-animal-used-in-research-in-2012/.

VectorBase. https://www.vectorbase.org.
Web Apollo. http://apollo.berkeleybop.org.
WHO (2013). http://www.unep.org/pdf/WHO_HSE_PHE_IHE_2013.1_eng.pdf.