# Nuclear receptors in the Pacific oyster, *Crassostrea gigas*, as screening tool for determining response to environmental contaminants

**Susanne Vogeler**

University of Exeter & Centre for Environment, Fisheries and Aquaculture Science

Submitted by Susanne Vogeler to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Biological Sciences in July 2016.



I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other university.

Signature:.....

## Abstract

Marine environments are under constant pressure from anthropogenic pollution. Chemical pollutants are introduced into the aquatic environment through waste disposal, sewage, land runoff and environmental exploitation (harbours, fisheries, tourism) leading to disastrous effects on the marine wildlife. Developmental malformations, reproduction failure including sex changes and high death rates are commonly observed in aquatic animal populations around the world. Unfortunately, the underlying molecular mechanisms of these pollution effects, in particular for marine invertebrate species, are often unknown. One proposed mechanism through which environmental pollution affects wildlife, is the disruption of nuclear receptors (NRs), ligand-binding transcription factors in animals. Environmental pollutants can directly interact with nuclear receptors, inducing incorrect signals for gene expression and subsequently disrupt developmental and physiological processes. Elucidation of the exact mechanism in invertebrates, however, is sparse due to limited understanding of invertebrate endocrinology and molecular regulatory mechanisms. Here, I have investigated the presence, expression and function of NRs in the Pacific oyster, *Crassostrea gigas*, and explored their interrelation with known environmental pollutants. Using a suite of molecular techniques and bioinformatics tools I demonstrate that the Pacific oyster possesses a large variety of NR homologs (43 NRs), which display individual expression profiles during embryo/larval development and supposedly fulfil distinct functions in developmental and physiological processes. Functional studies on a small subset of oyster NRs provided evidence for their ability to regulate gene expression, including interactions with DNA, other NRs or small molecules (ligand-binding). Oyster receptors also show a high likeliness to be disrupted by environmental pollutants. Computational docking showed that the retinoid X receptor ortholog, CgRXR, is able to bind and be activated by 9-*cis* retinoic acid and by the well-known environmental contaminant tributyltin. A potential interaction between tributyltin and the peroxisome proliferator-activated receptor ortholog CgPPAR has also been found. In addition, exposure of oyster embryos to retinoic acids and tributyltin resulted in shell deformations and developmental failure. In contrast, computer modelling of another putative target for pollutants, the retinoic acid receptor ortholog CgRAR, did not indicate interactions with common retinoic acids, supporting a recently developed theory of loss of retinoid binding in molluscan RARs. Sequence analyses revealed six residues in the receptor

sequence, which prevent the successful interaction with retinoid ligands. In conclusion, this investigative work aids the understanding of fundamental processes in invertebrates, such as gene expression and endocrinology, as well as further understanding and prediction of effects of environmental pollutants on marine invertebrates.

## Table of contents

## List of Tables and Files



## List of Figures

## Abbreviations

13RA: 13-*cis* retinoic acid
9RA: 9-*cis* retinoic acid
acyl-CoA: acyl coenzyme A
AF: activation function
AR: androgen receptor
ASW: artificial seawater
ATRA: all-*trans* retinoic acid
bp: base pair
BPA: bisphenol A
BS: bootstrap
CAR: constitutive androstane receptor
CDS: coding DNA sequence
CMV: human cytomegalovirus
COUP: chicken ovalbumin upstream promoter
CYP: cytochrome P450
DBD: DNA binding domain
DDT: dichlorodiphenyltrichloroethane
DHA: docosahexaenoic acid
DHR: *Drosophila* hormone receptor
DMSO: dimethylsulfoxide
dpf: days post fertilisation
DR: direct repeat
DSF: dissatisfaction receptor
E: efficiency
EcR: ecdysone receptor
EDC: endocrine disrupting chemicals
EF-1: elongation factor-1 α
ER: estrogen receptor
ERR: estrogen related receptor
FTZ-F1: fushi tarazu factor 1
FXR: farnesoid X receptor
GCNF: germ cell nuclear factor
GR: glucocorticoid receptor
HAT: histone acetyltransferase
HDAC: histone deacetylase
HEK: human embryonic kidney cells
HNF4: hepatocyte nuclear factor 4
hpf: hours post fertilisation
HR: hormone receptor
HSV-TK: herpes simplex virus thymidine kinase
ICES: International Council for the Exploration of the Sea
INR: invertebrate specific NR
IP: everted repeat
IR: indirect repeat
LA: linoleic acid
LBD: ligand binding domain
LBP: ligand binding pocket
LRH1: liver receptor homolog-1
LXR: liver X receptor
MCMC: Markov chain Monte Carlo
ML: maximum likelihood
Mpf: month post fertilisation
MR: mineralocorticoid receptor
MZT: maternal-to-zygotic transition
NCBI: National Centre for Biotechnology Information
NCoR: NR co-repressor
NF: normalisation factor
NGFI-B: nerve growth factor I-B
NHR: nuclear hormone receptor
NJ: neighbor-joining
NOR1: neuron derived orphan receptor-1

NR: nuclear receptor
nt: nucleotides
NURR1: nuclear receptor related 1 protein
PA: Promoter activity
PAH: polycyclic aromatic hydrocarbons
PC: Principal component
PCA: Principal component analysis
PCR: polymerase chain reaction
PFOA: perfluorooctanoic acid
PNR: photoreceptor cell-specific nuclear receptor
POP: persistent organic pollutants
PP: posterior probabilities
PPAR: peroxisome proliferator-activated receptor
PR: progesterone receptor
PXR: pregnane X receptor
qPCR: quantitative RT-PCR
RA: retinoic acid
RAR: retinoic acid receptor
Re: relative expression
RE: response element
RL7: ribosomal protein L7
ROR: retinoid acid related receptor
Rosi: rosiglitazone
RS18: ribosomal protein S18
RXR: retinoid X receptor
RXRE: RXR response element
SF1: steroidogenic factor 1
SHP: small heterodimer partner
SMRT: silencing mediator for retinoid and thyroid hormone receptors
SVP: seven-up receptor
TBT: tributyltin
TBTO: Bis(tributyltin) oxide
THR: thyroid receptor
TLL: tailless receptor
TPT: triphenyltin
USP: ultraspiracle protein
VDR: vitamin D receptor
ZGA: zygotic genome activation

## Species list

*Amphimedon queenslandica*
*Aplysia californica*
*Biomphalaria glabrata*
*Caenorhabditis elegans*
*Crassostrea gigas*
*Daphnia pulex*
*Drosophila melanogaster*
*Homo sapiens*
*Lottia gigantea*
*Lymnaea stagnalis*
*Marisa cornuarietis*
*Mytilus edulis*
*Mytilus galloprovincialis*
*Nucella lapillus*
*Octopus vulgaris*
*Photinus pyralis*
*Renilla reniformis*
*Schistosoma mansoni*
*Scorbicularia plana*
*Thais clavigera*
*Xenopus laevis*

## Author’s declaration

My thesis is presented as five individual chapters including my three research publications (published, accepted or submitted) and two additional chapters, a literature review and a final chapter discussing dual-luciferase reporter assays conducted by me. I am the lead author of my three publications. All molecular and embryo development experiments including oyster handling, spawning and fertilisation were planned, organised and executed by myself in consultation of my supervisors. The contribution of my supervisors and co-authors are described at the beginning of each experimental chapter. The publications were reformatted to provide a unified editorial and referencing style throughout, with figures embedded within the text. Supplementary files and references are compiled into a separate section and a single bibliography, respectively, at the end of the thesis.

## Acknowledgements

First and foremost I would like to thank my supervisors Prof Tamara Galloway and Dr Tim Bean for their support and guidance over the time of my Ph.D. I am very grateful for all the time and ideas they contributed, their motivation and advices, which made my Ph.D. such a wonderful and inspiring experience. It was a pleasure working in the lab with Tim, who had always been a great teacher. I have learnt so much. Tamara was always a source of encouragement and enthusiasm, and helped me pick up the pieces in times when things did not go as planned. I would also like to thank my third supervisor Dr Brett Lyons, whose advice, especially at the beginning of my Ph.D., was much appreciated.

Special thanks go to members of the Environmental Biology group: Paul Holcraft, Darren Rowe, Anke Lange and Eduarda Santos for their advice and help in the lab, oyster fertilisation experiments and data analysis. Many thanks to Nick Taylor and John Bignell for helping with the statistics and imaging. I would also like to acknowledge Dr Michail Isupov for his collaboration during the computational docking experiments. Personal thanks to Andrew Watts and Mauricio Urbina for being such lovely neighbours in the office.

My sincerest gratitude to Prof Taisen Iguchi and Dr Shinichi Miyagawa from the National Institute for Basic Biology in Okazaki for the opportunity to work at your institute and instructing me in the dual-luciferase reporter assay technique. I very much appreciated your hospitality.

I would like to thank my parents and my brother. Without your support completing this Ph.D. would have been a lot more difficult. I am indebted to Kai Diederichsen for his love, unfailing support, infinite patience and wonderful editing skills.

I also want to sincerely thank my friends, who made my life in Exeter such a joyful and fantastic time: Philippa Holder, Rik Lindsay, Michael Sieber, William Kay, Tom Laver, Emma Wright and Mark and Charlotte Hewlett.

I gratefully acknowledge the funding source that made my Ph.D. work possible. I was funded by the University of Exeter and the Centre for Environment, Fisheries and Aquaculture Science.

## Introduction

Environmental pollution has become a global concern, with chemical contaminants finding their way into the environment via waste disposal, agriculture, atmospheric deposition, and anthropogenic use. Marine ecosystems are under particular pressure from direct anthropogenic impacts, such as coastal tourisms, harbours, fishery and aquaculture, but also through terrestrial pollution. Eighty percent of the marine pollution comes from land, entering the aquatic environment from various sources [UNESCO, 2016]. Land runoff from agriculture and urban areas, especially after heavy rain falls, enters rivers and ground and coastal waters. Additionally, different types of litter arrive in marine environments, with plastic debris making up a significant percentage of all marine litter [Hammer et al., 2012]. Sewage effluents and industrial waste also contain metals, pharmaceuticals, pesticides, industrial chemicals and additives of personal care products.

Chemical compounds in the environment can have many effects on marine species. The World Health Organisation, the United Nations Environment Programme and the Endocrine Society have recently updated their concerns about the potentially adverse health effects of chemicals on humans and wildlife [Diamanti-Kandarakis et al., 2009; Bergman et al., 2013; Gore et al., 2015]. Xenobiotics, exogenous compounds not naturally produced by an organism or present in an abnormal concentration, can be taken up by an organism or cell and disrupt their developmental and physiological processes. Specific focus is set on endocrine disrupting chemicals (EDCs), which disrupt the endocrine system, an organism's collection of glands, which secrete hormones to regulate metabolism, growth, reproduction and development amongst other processes.

The effects of xenobiotics on marine species vary depending on the level and timing of exposure, which makes predicting xenobiotic effects particularly challenging. This is complicated even further due to a lack of information on precisely how xenobiotics evoke their disruptive effects. One of their proposed mechanisms is a direct effect on gene expression, a fundamental process in all living species, at which the information of genes is transcribed and translated into functional gene products such as proteins and functional RNAs. Gene expression is regulated by a range of regulating elements, which directly or indirectly interact with the DNA. The nuclear receptor superfamily is a class of transcription factors, found exclusively in metazoan species. These nuclear receptors (NRs) directly bind to specific sequences in the promoter

region of genes and control their expression. One of this superfamily's main features is the ability of some receptors to interact with endogenous compounds (e.g. steroid or thyroid hormones) by directly binding them as ligands, which subsequently initiate or inhibit gene transcription. Nuclear receptors are also able to respond to environmental stimuli by interacting with exogenous compounds [Germain et al., 2006]. This characteristic makes NRs vulnerable to xenobiotic compounds. Small molecules such as metals or synthetic compounds (e.g. pesticides, pharmaceuticals, phenols), but also natural and synthetic environmental hormones, can disrupt the NR function by mimicking a ligand or blocking a receptors' ligand binding ability [McLachlan, 2001]. However, knowledge on which and how xenobiotics interact with NRs is sparse. Additionally, the functions of NRs in many species are only partially or not at all characterised, making it difficult to predict the respective xenobiotic effects. This is particularly apparent for invertebrate species, for which the understanding of endocrinology and underlying molecular mechanisms is very limited.

Invertebrates make up more than 95% of all described species. Marine invertebrates vary remarkably from very simple to highly complex organisms, each with a distinctive endocrine system. They are under constant pressure from marine pollution, which affects a vast range of physiological and developmental processes in invertebrates. The ability to understand and predict these processes is highly desirable for the protection of not only marine invertebrates, but humans and wildlife in general.

Here, I investigate NRs in the Pacific oyster, *Crassostrea gigas*, and their putative participation in xenobiotic effects. The Pacific oyster is a key ecological species, belonging to the habitat-forming species, and serves as an important food source for wildlife and humans. It was recently established as a model species. This thesis comprises a literature review, in which I provide general information on NR structure and function as well as examples of chemical pollutant effects linked to a potential NR disruption. Subsequently, I present four experimental chapters, in which I investigate the hypotheses that: (1) the Pacific oyster possesses a large variety of NR homologs; (2) NRs are differently expressed during embryo, larval and adult life stages; (3) a subset of oyster NRs have the potential to bind to small ligands and regulate gene expression; (4) exogenous compounds such as environmental pollutants have the potential to interact with NRs and disrupt the normal receptor's gene expression function, which subsequently results in disruptions of physiological

processes. The results are discussed separately for each experimental chapter, leading up to a combined discussion on all chapters.

**Research aims:**

The objectives of my Ph.D. research are to:

- Provide an in-depth literature review on nuclear receptors, including structure, mechanism and function, as well as an overview of the effects of marine pollution on invertebrates, which are linked to potential nuclear receptor disruptions.

- Investigate the presence of nuclear receptors in the Pacific oyster, *Crassostrea gigas*, including their phylogenetic relationship.

- Assess the nuclear receptor expression during oyster embryo and larval development.

- Characterise three oyster nuclear receptors (retinoid X receptor, retinoic acid receptor, peroxisome proliferator-activated receptor), which show putative key functions in oyster endocrinology, as well as a high potential to interact with xenobiotic compounds. Their respective ability to bind natural and synthetic ligands is assessed and their connections to observed xenobiotic effects are discussed.

Chapter one

# Nuclear receptors and the effect of marine environmental pollution on invertebrates - a review


## Abstract

Nuclear receptors are a superfamily of ligand-binding transcription factors, which can regulate gene expression in metazoan species. They attach to specific response elements in the promoter region of a target gene. Some nuclear receptors, such as hormone regulated receptors, have the ability to react to endogenous and exogenous signals and bind small hydrophobic ligands, which have either agonistic or antagonistic effects on gene transcription. Different forms of dimerization, as well as other transcriptional machinery such as co-regulators and post-transcriptional modification also impact the gene expression regulatory system. Chemical pollutants, including natural and synthetic compounds, are introduced into aquatic environments through waste disposal, sewage, land runoff and anthropogenic environmental exploitation. Some of these exogenous compounds have the ability to enter organisms and interfere with nuclear receptor functions by binding to receptors and inducing incorrect signals. This can lead to disruption to reproduction, development, the endocrine system and other physiological processes. Understanding of these effects on invertebrate systems is sparse due to a limited understanding of endocrinology and molecular regulatory mechanisms in invertebrates. The Pacific oyster, *Crassostrea gigas*, provides many benefits as a model organism in the research of invertebrate endocrinology, molecular mechanisms and disruptive effects through environmental pollution.

## 1.1. Introduction

Gene expression is an essential aspect of life on earth and a key regulatory stage in the synthesis of functional gene products such as proteins and functional RNAs. This macromolecular process involves a sequence of sub-processes including gene transcription, RNA splicing, protein translation and post-translational modifications. Regulation of gene expression controls timing, location and quantity of gene expression and ensures an unobstructed process. Transcription factors are an example of such gene regulating elements. They promote or block the recruitment of the required gene transcription machinery. In metazoan species a specific type of transcription factors has evolved: the nuclear receptor (NR) [Laudet & Gronemeyer, 2002; Owen & Zelent, 2000; Laudet, 1997]. Nuclear receptors have the ability to directly bind to the DNA [Germain et al., 2006; Aranda & Pascual, 2001]. Some NRs are able to interact with ligands, endogenous or exogenous organic compounds, such as steroid and thyroid hormones and retinoids, which function either as agonists or antagonists. Other receptors (orphan receptors), for which no ligand exists, operate in a more constitutively activated manner, without ligand binding required. Embryo development, reproduction, metabolism and homeostasis are some of the processes regulated by NRs.

Nuclear receptors are common in multi-cellular metazoan species and are present across all phyla. Phylogenetic analyses have revealed that NRs have emerged early in metazoan evolution. Two NRs have even been discovered in demosponges, simplistic animals which do not have organs or clearly separated tissues [Larroux et al., 2006; Bridgham et al., 2010], but they are absent in yeast [Goffeau et al., 1996], plants [Arabidopsis, Genome & Initiative, 2000; in Baker, 2005] and choanoflagellates [King et al., 2005], the closest known relatives to animals. Up to 2011 more than 900 NR genes had already been identified [Sladek, 2011], ranging from a few NRs in sponges and *Trichoplax* [Baker, 2008; Bridgham et al., 2010; Larroux et al, 2006; Srivastava et al., 2008;], to approximately 21 NRs in insects [Adams et al., 2000; King-Jones & Thummel, 2005] and 48 NR genes in humans [Robinson-Rechavi et al., 2001] (Figure 1.1). The divergence of NRs is most likely driven by gene duplication and gene loss. Evidence for periods of gene duplication was found in the NR subfamilies [Bertrand et al., 2004; Bridgham et al., 2010; Laudet et al., 1992; Laudet, 1997].

The discovery of two NRs in the demosponge, *Amphimedon queenslandica,* indicated that the ancestral metazoans contained only one single NR, which doubled in sponges by gene duplication [Bridgham et al., 2010]. The deep conservation of NR sequences between different animal phyla also suggested this presence of a single ancestral NR [Bridgham et al., 2010; Escriva et al., 1997, 2000; Laudet et al., 1992; Sladek, 2011]. Controversy of the ligand status of the ancestral NR still exists, but recent theories propose that this receptor operated as a sensor for nutritional compounds [Markov & Laudet, 2011; Sladek, 2011], such as linoleic acid, an



**Figure 1.1: Gene duplication and gene loss history in the nuclear receptor evolution.** Duplications (green bars) and losses (red bars) of genes are labelled with named NR common names or subfamily groups. The large green bar comprises numerous duplications that cannot be temporally ordered. * many nematode genes are expansions of a single NR gene (HNF4-like, NR2A). ** vertebrate specific duplication. INR: invertebrate specific NR. Schematic NR history tree adapted from [Bridgham et al., 2010]. Quantity NR in phyla/classes adapted from [Sladek, 2011].

essential omega-6 fatty acid, which most animals cannot synthesize [Sladek, 2011]. The binding of these ligands is supposed to be of relatively low affinity. Another theory suggests a ligand-dependent ancestral NR. This is supported by the findings of the two demosponge NRs, which indicate binding to fatty acids and the regulation of gene transcription [Bridgham et al, 2010]. A ligand-independent ancestral NR is also possible [Escrivia et al., 1997; Mangelsdorf et al., 1995], given that orphan receptors exist. However, structure analyses of the demosponge NRs suggested that more mutations would be required for a ligand-dependent receptor to evolve from a ligand-independent receptor than for the loss of the ligand-dependency [Bridgham et al, 2010]. Nevertheless, additional waves of gene duplication and gene loss led to a variety of representatives of NR subfamilies before and after the divergence of protostomes and deuterostomes took place [Bertrand et al., 2004; Bridgham et al., 2010]. The NR superfamily is divided into six subfamilies (NR1-NR6) based on phylogeny, through sequence alignment and tree construction of two functional NR domains, the DNA binding domain (DBD) and the ligand binding domain (LBD) [Nuclear Receptors Nomenclature Committee, 1999] (Figure 1.2). Abnormally structured NRs, which do not contain one of the two conserved domains [Nuclear Receptors Nomenclature Committee, 1999] or hold two DBDs [Wu et al., 2006, 2007], are grouped in a separate subfamily (denoted NR0 or NR2DBD, respectively) irrespective of their phylogenetic relationship. Phylogenetic trees often display several genes for single types of NR (e.g. retinoid X receptor RXRα, RXRβ, RXRγ). Such isotypes are products of related paralogous genes, and often relate to phenotypic variations in NRs. Isoforms, on the other hand, are NR products of the same gene, but are produced by an alternative translation initiation, promoter usage, splicing or post-translational modifications (e.g. RXRα1, RXRα2) [Germain et al., 2006]. Receptors are named based on subfamily classification and are further categorised into subfamily groups [Nuclear Receptors Nomenclature Committee, 1999]. This classification system allows for newly identified NRs among different phyla to be placed in an appropriate evolutionary and functional context. Common names for receptors are also regularly used and generally refer to the first identified ligand or the respective anatomical region.

Nuclear receptors play essential roles in a large variety of metazoan processes including signalling during reproduction, development, physiology and metabolism [reviewed in Cheek et al., 1998; Omiecinski et al., 2011]. Due to their ability to interact with ligands, NRs have the potential to be unintentionally disrupted. Chemical

**Figure 1.2: Phylogenetic tree of 65 nuclear receptor genes in vertebrates, arthropods and nematodes.** Subfamilies are indicated by large Arabic numerals, groups by capital letters and brackets, and individual genes by small Arabic numerals next to the abbreviation of the common name. The length of the bootstrap is proportional to the bootstrap value. The bootstrap values defining the subfamilies are boxed, except in the case of GCNF1 defined by only one member. [Nuclear Receptors Nomenclatures Committee, 1999].

pollutants and xenobiotics, chemicals that are of exogenous origin or are present in an abnormal quantity, induce incorrect signals by interfering with receptor binding abilities [McLachlan et al., 2001]. This can also lead to a disturbance of the endocrine

system of an organism. Various xenobiotics have been characterized as “endocrine disrupting chemicals” (EDCs) and reports published on EDCs interpreted the deleterious effects of chemicals on the health conditions of humans and wildlife [Bergman et al., 2013; Diamanti-Kandarakis et al., 2009]. Effects on the reproductive system, neurodevelopment and participation in cancer, metabolic, immune and cardiovascular disorders and diseases have been tentatively assigned to EDCs. Therefore, it is crucial to further investigate NR presence and function to help understand and prevent said negative effects caused by xenobiotics and EDCs in particular.

## 1.2. Structure and mechanism of action

### 1.2.1. *Structure*

The structure of NRs is well characterised and most receptors contain five common structural features (Figure 1.3) [Germain et al., 2006].

The A/B region is the N-terminal of the NR and is highly variable in length and sequence between receptors and receptor isotypes. It is also a commonly used region for alternative splicing (isoforms) and post-translational modifications. Receptor and cell-type specific responses are regulated via this variable site. The N-terminal contains an activation function AF-1, which can operate autonomously, and produces a more stable up-regulation of gene transcription. This function can interact with various co-regulators, but also with its own receptor domains and the activation function AF-2 positioned towards the C-terminal of the receptor.

The C-region, commonly named the DNA binding domain (DBD), is the most conserved domain of the receptor, and is responsible for binding to specific DNA sequences of target promotors. It is composed of two cysteine-rich $C_4$ zinc finger motifs, each placed in one of the two main alpha-helixes of the DBD. The first zinc finger contains the proximal or P-box between two cysteines, and determines the DNA binding specificity. The second zinc finger holds the distal or D-box, which allows for dimerisation (homo- or heterodimer) of the receptor. Additional C-terminal extensions, like the T-box and A-box, further regulate DNA binding activity, receptor dimerization and protein-protein contacts.

Adjacent to the DBD, the D-region or so called hinge domain connects the DBD with the ligand binding domain (LBD) in the E-region. The hinge domain is highly variable and allows for folding, 3D structure and conformational change in the protein structure. It also contains nuclear localisation signals and sites for post-translation modification, protein-protein interaction, DNA binding and intracellular mobility.

The LBD is moderately conserved in sequence, but structurally it shows high similarities between receptors. In a sandwich-like arrangement up to twelve α-helixes (H1-H12) are organised around a hydrophobic binding pocket, the ligand binding pocket (LBP). Residues within the LBP can interact with lipophilic molecules. For some receptors, which are true orphans, the LBP is filled with hydrophobic amino acid side chains, which prevents ligand binding. Other features of the LBD domain are binding sites for dimerization, co-regulatory binding surfaces, and an activation

function AF-2 near the C-terminal (F-region). Activation function AF-2, usually in the helix 12, is responsible for ligand-dependent conformational change and recruiting of essential co-regulators.

Nevertheless, not all receptors follow this general structure. Most of the anomalies belong to the miscellaneous receptor subfamily NR0. The human receptors DAX-1 and SHP (small heterodimer partner) lack DBDs completely and act as repressors for other receptors. *Drosophila melanogaster* contains types of NRs, which lack the LBD instead and operate in a constitutively manner. Nuclear receptors containing two DBD have been identified in few protostome species [Vogeler et al., 2014; Wu et al., 2006; 2007].

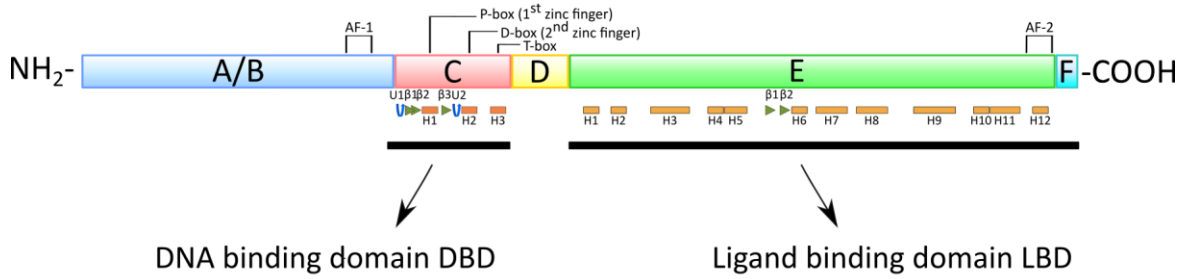





**Figure 1.3: Nuclear receptor structure and crystallographic structure of DNA (DBD) and ligand (LBD) binding domains.** Organisation of the five receptor domains including positioning of the main α-helixes, β-turns and u-turns of the DBD and LBD. The crystallographic structures of DBD [Vanden Heuvel, 2009] and LBD [Ng et al, 2014] display models derived from the estrogen receptor. Blue turns (U): u-turn; green triangle (β): β-turns; orange/yellow bars (H): α-helixes; AF: activation function; Zn : zinc.

### 1.2.2. *Dimerisation and DNA recognition*

Nuclear receptors bind to DNA either as monomers, homodimers or heterodimers [Mangelsdorf et al., 1995]. Some receptor types can work as monomers, including thyroid like receptors (THR, NR1A) [Forman et al., 1992; Lazar et al., 1991], the estrogen-related receptor (ERR; NR3B) [Johnston et al., 1997], and the nerve growth factor IB-like receptors (NR4A group) [Wilson et al., 1993], to name only a few.

Dimerisation between receptor partners occurs between both main functional domains, the DBD and LBD [Mangelsdorf et al., 1995; Germain et al., 2006; Aranda & Pascual, 2001; Helsen & Claessens, 2014]. In the DBD, specific dimer interfaces; often unique for different partners, and the D-box (receptor dependent) are responsible for DBD dimerisation. The LBD domains also connect via dimer interfaces, which comprise residues in several helixes, β-turns (additional receptor motifs) and loops. Several receptors work as homodimers (Table 1.1) such as steroid receptors (NR3 subfamily) and THR (NR1A), but also members of the NR2 subfamily and the NR4A group. Other receptors work as heterodimers, in combination with the RXR, which is the predominant heterodimer partner found in many species [Mangelsdorf & Evans, 1995; Dawson & Xia, 2012]. The RXR binds to receptors of various subfamily groups (Table 1.1), each with specific heterodimer interfaces in the DBD and LBD [reviewed in Helsen & Claessens, 2014]. In solution homodimers and heterodimers form more "open" dimers with the hinge region giving the opportunity

**Table 1.1: Selection of nuclear receptors (NRs) including dimerisation state and DNA binding sites.** AR: androgen receptor; ER: estrogen receptor; GR: glucocorticoid receptor; MR: mineralocorticoid receptor; PPAR: peroxisome proliferator –activated receptor; PR: progesterone receptor; RAR: retinoic acid receptor; RXR: retinoid X receptor; THR: thyroid receptor; VDR: vitamin D receptor. RE: response element; RA: retinoic acid.

| Nomenclature | Name | Ligand | P-box | Half-site | Dimerisation | RE |
|---|---|---|---|---|---|---|
| NR1A | THR | thyroid hormones | CEGCKG | 5'-AGGTCA-3' | monomer<br>homodimer<br>heterodimer | half-site<br>DR4, IP6, IR0<br>DR3, DR4 |
| NR1B | RAR | retinoic acids | CEGCKG | 5'-AGGTCA-3' | homodimer<br>heterodimer | IR0<br>DR1, DR2, DR5 |
| NR1C | PPAR | fatty acids | CEGCKG | 5'-AGGTCA-3' | heterodimer | DR1, DR2 |
| NR1I | VDR | vitamin D | CEGCKG | 5'-AGGTCA-3' | homodimer<br>heterodimer | DR3<br>DR3 |
| NR2B | RXR | 9-*cis* RA | CEGCKG | 5'-AGGTCA-3' | homodimer<br>heterodimer | DR1<br>DR1-DR5 |
| NR3A | ER | estrogens (17β estradiol) | CEGCKA | 5'-AGGTCA-3' | homodimer | IR3 |
| NR3C1 | GR | cortisol | CGSCKV | 5'-AGAACA-3' | homodimer | IR3 |
| NR3C2 | MR | aldosterone | CGSCKV | 5'-AGAACA-3' | homodimer | IR3 |
| NR3C3 | PR | progesterone | CGSCKV | 5'-AGAACA-3' | homodimer | IR3, DR3 |
| NR3C4 | AR | testosterone | CGSCKV | 5'-AGAACA-3' | homodimer | IR3, DR3 |

**Figure 1.4: Solution and crystal structures of nuclear receptor hetero- and homodimers.** The solution structures were obtained via SAXS, SANS and cryo-EM. The crystal structures were determined via crystallography. Triangles: DBD; circles: LBD; small cylinders: ligands; turquoise helix: DNA; DR: response elements in direct repeats, small number indicating number of residues between motifs. RXR: retinoid X receptor; PPAR: peroxisome proliferator-activated receptor; RAR: retinoic acid receptor; VDR: vitamin D receptor; HNF-4: hepatocyte nuclear factor 4. Adapted from [Helsen & Claessens, 2014].

for different binding options to the DNA. The DBDs and LBDs bind independently (Figure 1.4). Crystal structure analyses using X-ray crystallography, on the other hand, has found a more compact dimerisation of the same dimers, including binding between DBD and LBD. Both structures are suggested as snapshots of the receptor complex, which is in constant equilibrium between several energy-favourable conformations. This enables an important increase of structural variation.

Nuclear receptor complexes bind to response elements (RE) in the promotor sequence of target genes [Mangelsdorf & Evans, 1995; Laudet & Gronemeyer, 2002]. A RE contains two hexameric core half-site motifs, which form direct (DR), indirect (IR) or everted repeats (IP) and are separated by a short spacer of various residues with differences in length (residue length 0-10) (Figure 1.5). Each receptor binds to one half-site motif with its P-box (Figure 1.3). The P-boxes of all receptors show variations of 5'-CxxCKx-3' (Table 1.1). A receptor can also attach to REs with one half-site, but binds it as monomer. Based on their structure and mechanism NRs can be categorized in three classes. Class I receptors, steroid receptors, bind to the DNA as homodimer and attach to the indirect repeat with the half-site 5'-AGAACA-3' (NR3C group) or 5'-AGGTCA-3' (estrogen receptor ER, NR3A). The indirect repeats

are separated by three residues functioning as a spacer (IR3) [Mangelsdorf et al., 1995]. The androgen receptor (AR, NR3C4) and progesterone receptor (PR, NR3C3) (originally Class III by [Mangelsdorf et al., 1995]) also binds to direct repeats DR3 [Denayer et al., 2010; Kerkhofs et al., 2012]. The second class (Class II) includes most NR1 subfamily members (THR, NR1A; retinoic acid receptor RAR, NR1B; peroxisome proliferator-activated receptor PPAR, NR1C; liver X receptor LXR, farnesoid X receptor FXR, ecdysone receptor EcR, NR1H; vitamin D receptor VDR, constitutive androstane receptor CAR, pregnane X receptor PXR, NR1I), binding as a heterodimer with RXR to direct repeats of 5'-AGGTCA-3' with different spacer lengths [Mangelsdorf et al., 1995; Rastinejad et al., 1995; reviewed in Dawson & Xia, 2012]. Some of these receptors can also bind as homodimer or monomers (RXR, THR, and RAR) and attach to direct, indirect or everted repeats. The third class (Class III; originally Class IV) comprises all remaining receptors binding either as homodimer or monomer to variations of 5'-AGGTCA-3' [Mangelsdorf et al., 1995].

**a)**

This image has been removed by the author of this thesis/dissertation for copyright reasons.

**b)**

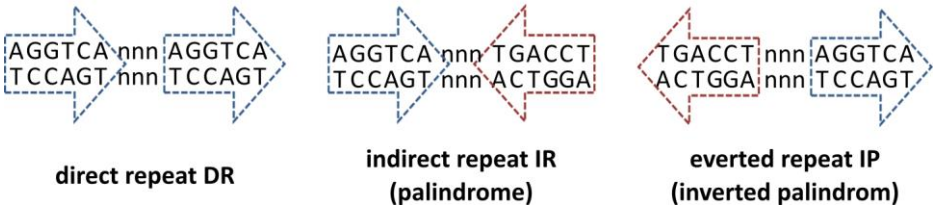



**Figure 1.5: Nuclear receptor DNA binding sites and response element (RE) types. a)** Organisation of the DBD α-helixes 1 and 2 including the two zinc-finger motifs (zn), P-box and D-box. The blue letters indicate the main residues interacting with the DNA [Vanden Heuvel, 2009]. **b)** Response element types including orientation of the two motifs forming the RE: direct repeat DR, indirect repeat (palindrome) IR, and everted repeat (inverted palindrome) IP.

### 1.2.3. *Basic mechanism of action*

The basic mechanism of action for NRs can be categorised in two main groups (Figure 1.6). The first group includes all Class I receptors, steroid receptors,

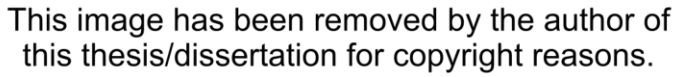



**Figure 1.6: Basic mechanism of nuclear receptors as homodimers or heterodimers.** In the cytoplasm homodimer-forming nuclear receptors are bound by various types of heat shock proteins. In the event of ligand binding to the receptor, the complex of receptor/heat shock proteins will be terminated due to conformational changes. The nuclear receptor forms a homodimer with the same type of receptor, migrates into the nucleus and attaches to the promoter sequences of a target gene. Gene expression can be initiated after recruitment of the required transcription factors and co-activators. Heterodimer complexes are already bound to the gene promoter region, but kept inactive by co-repressors. After successful ligand binding, the co-repressor complex detaches from the heterodimer and the co-activator complex can be recruited. The gene transcription is initiated. [Vanden Heuvel, 2009].

operating as homodimers [Mangelsorf et al., 1995]. As inactive receptor monomers they are mainly present in the cytoplasm, bound to several proteins such as heat shock proteins and form a large complex [Mangelsorf et al., 1995; Vanden Heuvel, 2009]. When a ligand binds to the monomer, the receptor undergoes a conformational change and is released from the protein complex. The monomer forms a homodimer with another free monomer and migrates into the nucleus where it binds to a designated RE in the promoter region. If the ligand is an agonist, the homodimer recruits the required co-activators and RNA polymerase to induce the gene transcription. In case of an antagonist the monomer does not detach from the complex or bind to co-activators and polymerase. Co-repressors are recruited instead, which inhibit the transcription process.

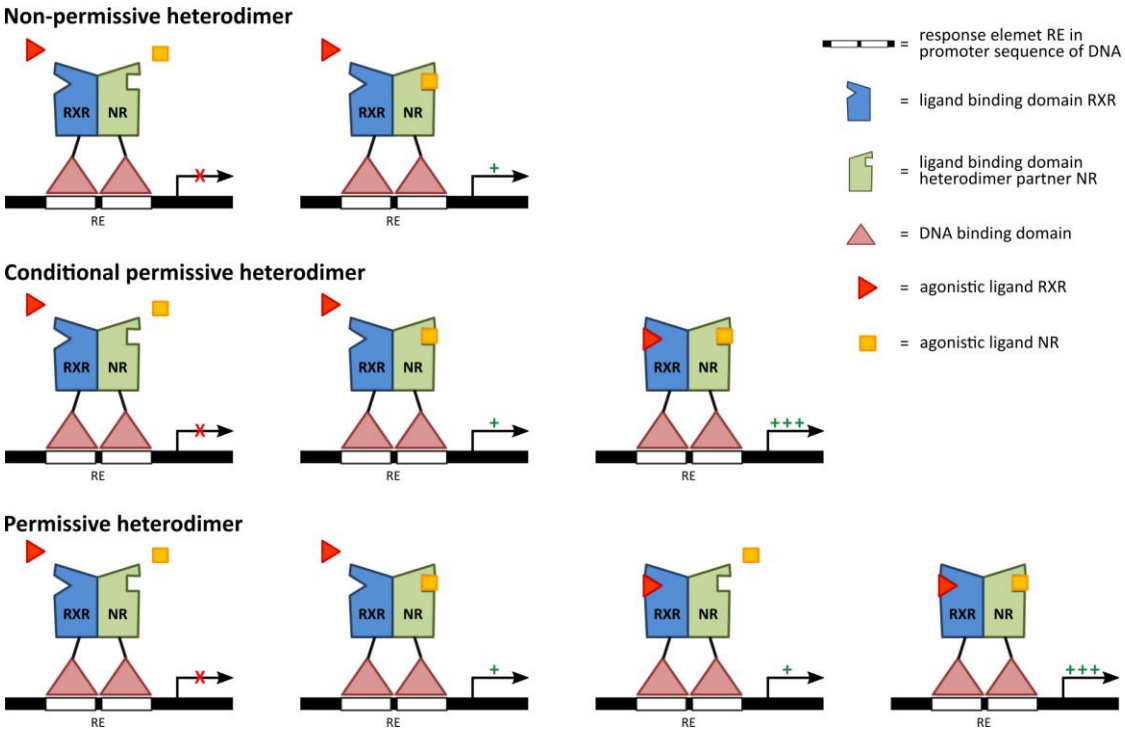



**Figure 1.7: Transcriptional activation by non-permissive, conditional permissive and permissive heterodimers of retinoid X receptor RXR and other nuclear receptors (NRs).** Red X: no transactivation; green +: activation; triple +: enhanced activation.

The second group combines all receptors working as heterodimer with RXRs (Class II receptors). They are attached to the REs as heterodimers and are inactivated through co-repressors. Agonistic ligands can bind to the heterodimer complex, which let the co-repressor detach and recruit co-activators and RNA polymerase. Three different types of ligand activation heterodimers exist (Figure 1.7) [reviewed in Xia & Dawson]. The non-permissive heterodimers (e.g. VDR, THR) can only be activated through the RXR partner. Permissive heterodimers (e.g. PPAR, FXR, PXR, LXR) are activated by both, the RXR and its partner. When both partners bind to a ligand the response will be increased. The conditional permissive heterodimer (RAR) is activated through the RXR partner, but the signal will be enhanced through ligand binding to RXR after the partner has been activated. Antagonistic ligands, on the other hand, will not induce the recruitment of transcriptional factors, but will block the ligand binding pocket.

Most Class III receptors belong to the first group of mechanisms, but can exhibit a slightly different behaviour. Hepatocyte nuclear factor 4 (HNF4) binds as homodimer similar to group one, but it is present in the nucleus instead of the cytoplasm [Jiang et al., 1995]. Other Class III receptors operate similar to group two, and the receptors

such as RAR-related orphan receptors ROR [Jetten et al., 2001] bind as monomer and reside in the nucleus and cytoplasm.

Nevertheless, not all receptors are activated by a ligand. Several receptors such as the ERR initiate upregulation of gene expression without any ligand binding required and operate as constitutively activated receptors [Horard & Vanacker, 2003]. These receptors have a constant activity, which is regulated through the receptor presence. Antagonistic ligands are an additional regulation factor by recruiting co-repressors and silencing the constitutive activity of the receptor [Germain et al., 2006; Greschik et al., 2004].

### 1.2.4. *Ligands and ligand binding*

Nuclear receptors belong to the ligand interacting transcription factor superfamily, which employs a communication between the receptor and a ligand to activate/deactivate gene expression [reviewed in Germain et al., 2006]. Orphan receptors, on the other hand, are receptors that do not have a ligand (true orphans), or those for which a ligand has not yet been identified. When a ligand later is discovered, the NR can be referred to as an “adopted orphan”. Most true orphans operate in a constitutive manner, but all other ligand interacting receptors require a ligand signal to function. Ligand binding receptors can be found among all subfamilies and it is also suggested in evolutionarily old phyla such as the Cnidaria [Reizel & Tarrant, 2009].

Nuclear receptor ligands are small, hydrophobic molecules, which vary greatly in their structure, mechanism and function [Germain et al., 2006; Sladek, 2011]. They are endogenous or exogenous compounds, which the latter being either from a natural or synthetic origin. Natural ligands range from different hormones such as steroids and thyroids, to retinoids (vitamin A derivatives), heme and fatty acids (Figure 1.8). Synthetic ligands for NRs are often pharmaceutical drugs made for the treatments of diseases, such as diabetes, cancer or hormone resistance syndromes [reviewed in Ottow & Weinmann, 2008]. However, synthetic ligands also include any man-made compounds such as pesticides, industrial chemicals and environmental pollutants. Some of these synthetic compounds were not originally generated as a ligand, but still interact with receptors. This interaction can even occasionally result in an orphan receptor receiving an interactive ligand [Greschik et al., 2004].

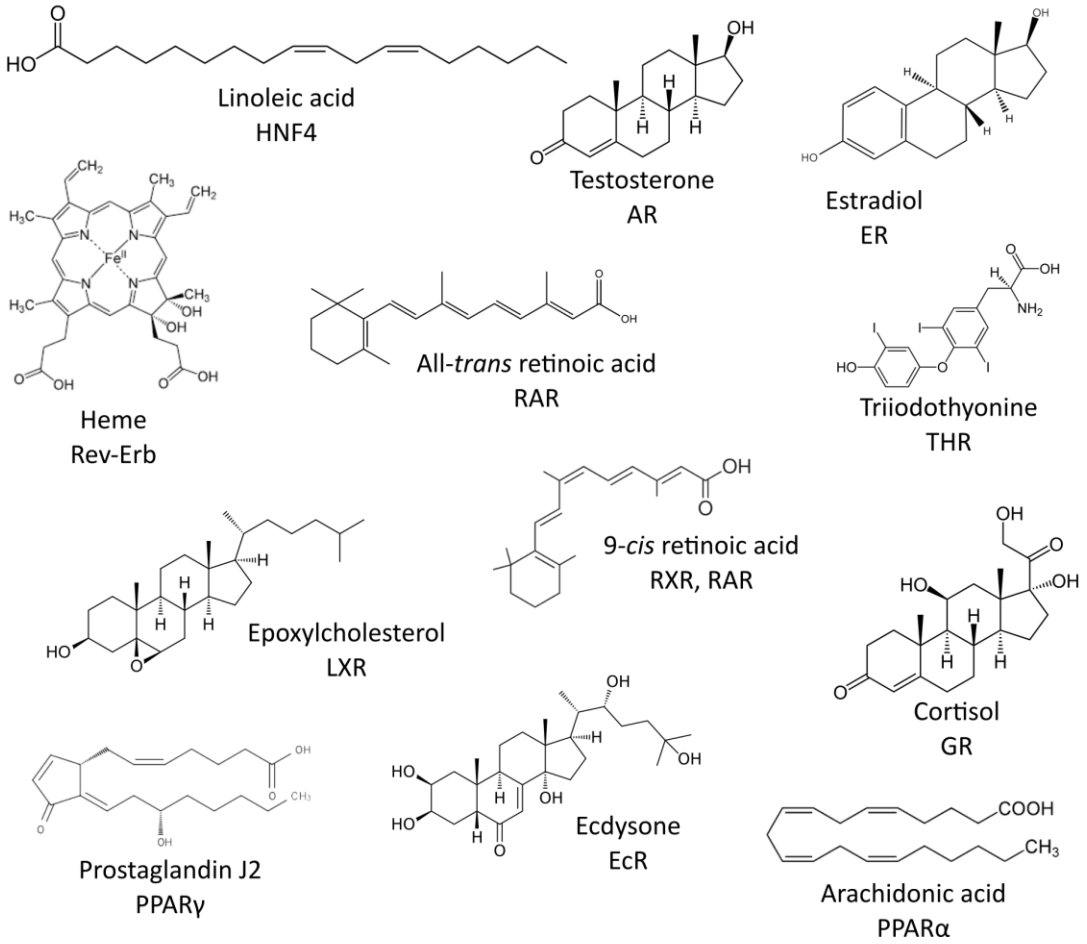



**Figure 1.8: Selection of natural agonistic ligands for nuclear receptors.** HNF4: Hepatocyte nuclear factor 4 (NR2A); AR: androgen receptor (NR3C); ER: estrogen receptor (NR3A); Rev-erb: Rev-erb receptor (NR1D); RAR: retinoic acid receptor (NR1B); RXR: retinoid X receptor (NR2B); THR: thyroid receptor (NR1A); LXR: liver X receptor (NR1H); GR: glucocorticoid receptor (NR3C); PPAR: peroxisome proliferator –activated receptor (NR1C); EcR: ecdysone receptor (NR1H).

Some ligands are unique to a single receptor (e.g. 20-hydroxyecdysone to the EcR in arthropods [Koelle et al., 1991]), whilst others bind to a wide range of receptors irrespective of the receptors phylogenetic relatedness. Retinoic acids (RA) activate the RXR (NR2B), but also the RAR (NR1B) [Heyman et al., 1997; Idres et al., 2002; Levin et al., 1992]. Many steroids in vertebrates bind to members of the NR3 subfamily [reviewed in Beato & Klug, 2000], but can also interact with NR1 members (e.g. NR1I xenobiotic-sensing receptors) [Banerjee et al., 2013]. Closely related receptors and receptor isotypes can have different sets of ligands, instead. The PPARγ is activated by a prostaglandin (active lipid) [Forman et al., 1995a], but PPARα by leukotrienes (eicosanoid inflammatory mediator) [Devchand et al., 1996]. The same ligand can also have an activating or repressing effect on closely related

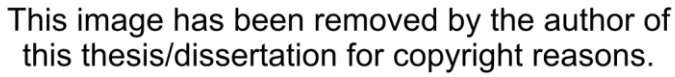



**Figure 1.9: Schematic representation of the 3D structure of LBD binding mechanism to ligands.** **(a)** Canonical apo-structure of RXR (unliganded) **(b)** holostructure of RAR (ligand bound), **(c)** antagonist bound to RAR LBD and **(d)** holo-structure bound by co-activator Grip 1/TIF2 with H12 occupying the NR box binding site (adapted from [Germain et al., 2003]). **(e-g)** simple schematic illustration of ligand binding to agonist, antagonist and partial agonist including involved helixes and co-activator (adapted from [Vanden Heuvel, 2009]).

receptors. The synthetic ligand BMS453 activates the RARβ isotype, but inactivates RARγ [Germain et al., 2004].

Ligands interact either as antagonist, agonist or partial agonist and bind to the hydrophobic LBP of the NRs [reviewed in Germain et al., 2003, 2006]. The induced conformation changes and the recruiting/enhancing of co-regulators can be very specific and unique for NR types. The general principle can still be summarised as follows (Figure 1.9) [Wurtz et al., 1996]. The LBD, with its up to 12 α-helixes (H1-H12) and a β-turn, is folded in a three layer sandwich form. The un-liganded state of the LBD is the so called apo-form (Figure 1.9a). When an agonistic ligand binds to the LBP the receptor changes its conformation to the holo-form. Agonists initiate the transcription by recruiting co-activators and other transcription factors to

the receptor and promoter sequence (Figure 1.9b,e). When the ligand binds, the H11 is repositioned and H12 swings back, flips over H6 and gets close to H4. The H12 works as a lid and seals the ligand inside the receptor cavity. The ligand is stuck (mousetrap principle). Thereby, the AF-2 gets displayed and becomes transcriptionally competent. Co-activators can bind with their LxxLL sequence (NR box [Heery et al., 1997]) to a receptor surface (formed by H3, H4 and H12), which emerged due to conformational changes (Figure 1.9d). These changes can also release any pre-bound co-repressors.

Antagonists, on the other hand, are ligands, which prevent transcription by not inducing the correct conformational changes to recruit the required co-activators [reviewed in Germain et al., 2003, 2006]. Many antagonists (but not all) have bulky side chains which cannot fit in the LBP similar to the agonists (Figure 1.9c, f). The C-terminus of H11 unwinds, resulting in a longer loop L11-12. The H12 flips back differently and covers the co-activator recognition surface (cleft). H12 is prevented from adopting the holo-position and co-activators cannot be bound. If a co-repressor is pre-bound, its detachment cannot be initiated. Some ligands operate as inverse agonists (an antagonist type). For un-liganded receptors, which do not interact (e.g. constitutively activated receptors) or only weakly interact with co-repressors, inverse agonists can induce conformational changes (H12 flips back to the N-terminus of H3), which leaves enough space to reinforce co-repressors (e.g. SMRT and NCoR, see below) and stabilize their interactions.

The final ligand type is the partial agonist which exhibits a mixture of agonistic and antagonistic features [reviewed in Germain et al., 2006]. When bound to such a ligand, H11 is partly unwound, but not in the same extent as for antagonists (Figure 1.9g). H12 flips back similarity to the agonists, but not identically (loss of interaction of H12, ligand and/or H3), slightly destabilising the agonistic conformation. It is assumed that the present concentration equilibrium of co-activators and co-repressors determines whether partial agonistic ligands inhibit or prevent transcription activation.

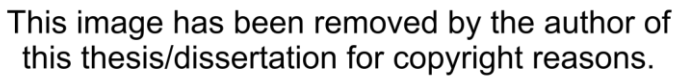



**Figure 1.10: selection of co-activators and co-regulators.** Co-activator complexes (green) include factors that contain ATP-dependent chromatin remodelling activity, histone arginine methyltransferases, histone acetyltransferases (HAT), as well as factors, which are involved in RNA processing and components of the so-called Mediator complex, which mediate the interaction with the RNA polymerase II (pol II) machinery. Conversely, co-repressors (red) include ATP-dependent chromatin remodelling complexes, basal co-repressors, such as NCoR and SMRT, which function as platforms for the recruitment of several subcomplexes that often contain histone deacetylase activity (HDAC), and specific co-repressors recruit general co-repressors on ligand induction. TATA: TATA–box; general transcription factors: IIA, B, E, F, H, J; associated TATA-binding protein factors: TBP, TAFs; Pol II: RNA polymerase II [Perissi & Rosenfeld, 2005].

### 1.2.5. *Co-regulators*

Co-regulators are macromolecules which help modulate transcription activity in association with NRs (for an extensive review on co-regulators [Rosenfeld et al., 2006]). They are divided into two groups: co-activators induce transcriptions and assure upregulation, while co-repressors inhibit the transcription (downregulation). Many co-regulators have been identified, most of which are not NR specific. They are grouped according to their functional properties (Figure 1.10). Co-activators usually operate when a ligand is bound to the NR and the conformational change of the

receptor allows for the recruiting co-activators. They are bound to hydrophobic grooves generated by the C-terminal part of H3, H4, H12 and loop L3-4 [Germain et al., 2006; Perissi & Rosenfeld, 2005]. Most co-activators work as chromatin-remodelling enzymes, such as the histone acetyltransferases (HATs), which acetylate lysine residues in histones [Vo & Goodman, 2001]. This acetylation neutralizes the positive charges of the histone, which weakens the histone tail with the nucleosome DNA. The transcriptional machinery can be recruited to the promoter sequence. Members of the p160 family, CREB-binding proteins (CBP) and p300 proteins are examples of HATs. Some of these co-activators have specific binding sequences in the LBD of the receptor. For example, p160 proteins are able to bind with their LxxLL α-helical motif (NR-box) to the LBD [Heery et al., 1997]. ATP-depending chromatin remodelling also disrupts the histone and DNA contact, which leads to a release of the nucleosome [Vignali et al., 2000]. RNA processing activators and activating mediators interact directly with the transcription machinery, recruit the required factors and process the mRNA [Perissi & Rosenfeld, 2005].

Co-repressors silence the activation of gene expression by binding to un-liganded receptors such as RXR heterodimers. The NR co-repressor (NCoR) and silencing mediator for retinoid and thyroid hormone receptors (SMRT) are the best characterised co-repressor complexes [Chen & Evans, 1995; Hörlein et al., 1995]. A co-repressor contains two NR-interacting domains, each including an LxxxIxxxI/L motif (CoRNR-box), with which it binds to the LBD in specific hydrophobic grooves in the surface of the LBD. NCoR and SMRT do not have intrinsic enzymatic activity and recruit enzymes complexes to help with the repressive action [Perissi et al. 1999]. Histone deacetylases (HDAC) deacetylate the lysine residues in the N-terminal of the histones, and thus increase the positive charges and promote a closed nucleosomal structure. The receptor interacting protein 140 and ligand dependent CoRs bind to liganded receptors and compete with co-activators [Fernandes & White, 2003]. In few cases receptors themselves can operate as co-repressor. Members of the NR0A subfamily DAX1 and SHP, do both not contain a DBD, and act as inhibitory partners for other NRs (transrepression) [Bavner et al., 2005; Germain et al., 2006]. They bind to the LxxLL motifs, competing with co-activators and actively repress the transcription.

### 1.2.6. *Post-translational modification*

Regulatory activity of NRs and co-regulator proteins is influenced by post-translational modification, another piece of the gene-transcription regulation process [Germain et al., 2006]. Post translational modifications can be divided into two categories: The first category holds all reversible modifications by addition or removal of functional groups to specific amino acids of the proteins through phosphorylation, acetylation and methylation; the second category groups modifications by attaching other proteins or polypeptides to the receptors or co-regulators. SUMO (SUMOlation) or ubiquitin (ubiquitination) proteins are two examples of such post-translational modification.

Phosphorylation is a common post-translational modification occurring through the addition of phosphoryl groups ($PO_3^{2-}$) to amino acids of proteins. Most of the phosphorylation in NRs appears in the A/B region but also in the DBD and LBD [Rochette-Egly et al., 1995; Delmotte et al., 1999] and can lead to DNA disconnection, ligand binding termination and a decrease in ligand binding affinity. However, it can also have a positive effect on transcription. Phosphorylation of a serine residue surrounded by prolines in the A/B region can enhance the activation [Germain et al., 2006]. Co-activator phosphorylation can increase the NR's affinity and HAT recruitment.

## 1.3. Marine environmental pollution and nuclear receptors

Eighty percent of marine environmental pollution comes from land, entering the ecosystem from point and non-point sources [Bergman et al., 2013; UNESCO, 2016]. Point sources are single, identifiable sources such as sewage effluents and industrial waste, containing metals, pesticides, industrial chemicals, pharmaceuticals and additives of personal care products. Non-point sources occur over a wider area and are often termed “diffuse” pollution. Land runoff from agriculture and urban areas contain a wide range of chemical contaminates, plastics and other litter, and arrive in rivers and coastal waters. Emissions of a diversity of chemicals from urban areas and especially from industrial areas, lead to air pollution and enter the aquatic environment through precipitation [Bergman et al., 2013]. Additionally, direct anthropogenic use of the ocean and coastal areas (e.g. harbours, fisheries, aquaculture and tourism) introduces further pollutants to the marine environment.

The effects of environmental pollutants on the health conditions of wildlife and humans are well documented [Bergman et al., 2013; Diamanti-Kandarakis et al., 2009] and range from effects on reproduction, development, growth and immune system, to the point of mortality. Many studies focus on xenobiotics, in particular on EDCs, and their effects on the physiological and developmental processes of vertebrates, including the endocrine system. Research on invertebrates, however, is scarce.

Despite the numerous records of the effects of EDCs on animals, their distinct mode of action is unknown for most of them. Understanding and studying the mechanisms of EDCs and other xenobiotics is particularly difficult as they vary widely in their chemical properties and environmental fate and range from natural hormones (17β-estradiol, estrone, testosterone), metals (e.g. mercury, silver), pesticides (e.g. DDT, atrazine, vinclozolin, tributyltin) and pharmaceuticals (e.g. levonorgestrel, fluoxetine), to persistent and bioaccumulative chemicals (e.g. persistent organic pollutants (POPs)), as well as less persistent and bioaccumulative chemicals such as phenols and plasticisers [Bergman et al., 2013]. One observed feature of xenobiotics is their ability to directly interact with NRs. They can mimic a natural ligand (both their agonistic or antagonistic function) or block the LBD, leading to disruption of the normal receptor function [McLachlan, 2001]. Predicting the effects of xenobiotics is challenging as the effects vary depending on the level and timing of exposure. Additionally, many xenobiotics can target a variety of different NRs. Bisphenol A, a

synthetic monomer and an additive in many plastics, for example, is able to bind to different members of the NR superfamily: ER [Lutz and Kloas, 1999; Suzuki et al., 2004], ERR [Toheme et al., 2014], AR [Ekman et al., 2012], PPAR [Rui et al., 2014] and THR [reviewed in Zoeller et al., 2005]. Many molecular mechanisms such as gene expression regulation and the endocrine system are conserved among vertebrates. Nuclear receptors, including their ability to bind to specific natural or synthetic ligands, also show high conservation across phyla [Mangelsdor & Evans, 1995]. Research on xenobiotic effects in vertebrates, including those caused by EDCs, is abundant, particularly due to the major interest for human health. Knowledge on endocrinology and gene expression regulation, including underlying molecular mechanisms, in invertebrates is limited to a few model species belonging mainly to the Insecta and Crustacea genera.

### 1.3.1. *Xenobiotic effects on marine invertebrates*

The marine environment provides the best documented examples of xenobiotic disruption in protostome invertebrate species caused by environmental contaminants. The organotin tributyltin (TBT), and in a lesser extent triphenyltin (TPT), was introduced as a biocide in antifouling paints for boats in the 1960/70s. Although these paints were very effective, TBT and TPT slowly leached into the surrounding waters and caused adverse effects on aquatic environments world-wide at low concentrations. Female masculinisation in species of the class Gastropoda (Phylum: Mollusca) has been widely reported as an effect of organotin exposure and ~20 gastropods species are known to be affected by TBT or TPT [Title-O'Neal et al., 2011]. Three types of female masculinisation are distinguished: imposex, a superimposition of male organs (penis and/or vas deferens) on females (Figure 1.11) [Matthiessen et al., 1999]; intersex, the shift of the pallial organs towards male morphological structures [Matthiessen et al., 1999]; and ovo-testis, supressed oogenesis and/or occurrence of spermatogenesis or seminiferous tube-like structure in females [Gibbs, et al., 1988 in Titley-O'Neal et al., 2011]. Gastropods are not the only class affected by TBT. In bivalve species (Phylum: Mollusca) shell thickening, growth reduction, developmental failure, reproduction impairment and a high rate of mortality were observed [Heral et al., 1989; Higuera-Ruiz & Elorza, 2011; Inoue et al., 2004; Park et al., 2015; Ruiz et al., 1995a, 1995b; Thain, 1986; Tsunemata &

Okamura, 2011; Waldock & Thain, 1983]. The effects were also noticeable in the shellfish industry. Arcachon Bay in France had an annual oyster production of 10,000-15,000 tons until the mid-1970s [Evans, 2000]. Around that time the number of vessels and boats in the bay increased, introducing TBT to the water in high concentrations (up to 8 kg TBT per day) and, consequentially, the annual oyster production decreased dramatically to only 3,000 tons by 1981 [Ruiz et al., 1996]. After TBT was linked to the observed effects in gastropods and shellfish, step-by-step legal actions were undertaken, which finally resulted in a global ban of TBT and other organotins in antifouling systems in 2008 [CD, 2002, IMO, 2001]. Long-term recovery of the macrofaunal communities has been reported along various coast lines since [Langston et al., 2015; Morton, 2009; Nicolaus & Barry, 2015]. However, although TBT has a short half-life in water, degrading within a few days, it has a high affinity to particles and it is persistent in soil for days, years and even decades depending on the environmental conditions. Therefore, the issue of TBT is still present in many areas today.



**Figure 1.11: Drawings of generalised female whelk without and with superimposition of male organs (imposex) due to contaminant exposure.** Unaffected female whelks pass their fertilised eggs through capsule glands and enclosed them in capsules, which will be released via the genital pore and attached to the sea bottom. Affected females develop a non-functional penis and form a vas deferens, which overgrows and blocks the genital pore. Encapsulated eggs cannot be released and are reabsorbed by the capsule gland. The female whelk becomes functionally sterile. [Carefoot, n.d.; Bright & Ellis, 1990]

The underlying molecular mechanism of TBT disruption in molluscan species is still not resolved, but research strongly suggests a participation of specific NRs. TBT and TPT were proposed as mimics for the natural ligands of RXR ortholog in gastropods [Horiguchi et al., 2007, 2010a, 2010b; Nishikawa et al., 2004]. RXR orthologs were

characterised in many molluscan species [Abidli et al., 2013; Bouton et al., 2005; Carter et al., 2010; Castro et al., 2007; Lv et al., 2013; Nishikawa et al., 2004; Sternberg et al., 2008] and TBT has been reported as a ligand for RXR orthologs in a gastropod species (*Thais clavigera*) and in vertebrates [Grün et al., 2006; le Maire et al., 2009; Urushitani et al., 2011]. 9-*cis* retinoic acid, a vitamin A derivative, is a proposed natural ligand for RXR and it has been shown to bind to gastropods RXRs (*T. clavigera* [Urushitani et al., 2011], *Nucella lapillus* [Gutierrez- Mazariegos, et al., 2014] and *Biomphalaria glabrata* [Bouton et al., 2005]). Injections of 9-*cis* retinoic acid led to imposex development in the females of two gastropod species (*N. lapillus* & *Nassarius reticulatus*), which further supports the NR theory [Castro et al., 2007; Lima et al., 2011; Sousa et al., 2010]. Additionally, the vertebrate RXR/PPAR permissive heterodimer has been proposed as a TBT target. In vertebrates TBT binds to the RXR receptor and intensifies its effect by also binding to PPAR [Grün et al., 2006; Harada et al., 2015; Kanayama et al., 2005; le Maire et al., 2009]. Injection of rosiglitazone, a synthetic ligand for vertebrate PPARγ, in the gastropod *N. lapillus* induced imposex, which lead to the assumption of a RXR/PPAR involvement of TBT-induced disruption in gastropods [Pascoal et al., 2013]. This theory has been supported by identified PPAR homologs in other molluscan species (*B. glabrata* & *Lottia gigantea* [Kauer et al., 2015], *C. gigas* [Vogeler et al., 2014 (see Chapter 2)]).

Endocrine disruption through environmental pollutants has also been reported for crustacean species (reviewed in Rodriguez et al., 2007; Weis, 2015), and in a lesser extent for annelids [reviewed in Krajniak, 2005] and nematodes [reviewed in Höss & Weltje, 2007]. Heavy metals, xenoestrogens, pesticides and other natural and synthetic compounds have shown to affect moulting, metamorphosis, reproduction and growth processes in invertebrates. Special concern are proposed regarding insecticides affecting crustaceans as non-target organism due to their close relationship to insects, as they share endocrine features regulating physiological processes [Rodriguez et al., 2007; Weis, 2015]. The molecular mechanisms of endocrine disruption are mainly unknown, but suggested mechanisms range from disruptions of NRs and other receptors, hormone levels, transduction pathways, enzymatic activities and protein production.

### 1.3.2. *The Pacific oyster as model organism*

Invertebrates make up more than 95 % of all living animal species on earth and are essential elements of many ecosystems. But yet, invertebrates are fairly underrepresented as model organisms, even though they provide several advantages over vertebrate species. They are, for example, easier to culture, have shorter generation times and also fewer legal/regulatory requirements. The Pacific oyster, *Crassostrea gigas*, is a sessile marine bivalve belonging to the phylum Mollusca. Bivalve species are commonly used for biomonitoring programmes and are sentinel organism tools for marine ecotoxicology tests [Matthiessen, 2008; Davies & Vethaak, 2012]. The Pacific oyster fulfils many criteria for a suitable model organism. It is a keystone species living in the intertidal and subtidal zones along coasts and estuaries world-wide and belongs to the habitat-forming species [Coen et al., 2007]. Oyster reefs are inhabited by various macrofauna such as other molluscs, crustaceans, cnidarians and fish, and provide a nursery habitat for juveniles [Zimmerman et al., 1989]. Additionally, oysters are a food source for many aquatic species and as filter feeders they eliminate organic matter from the water column. Oyster reefs also prevent sediment deposition and stabilise shorelines.

As a research species the Pacific oyster is relatively inexpensive, easy to handle in the laboratory and can tolerant wide ranges of salinity (20-35 ‰) and temperature (-1.8-35 °C) [Helm, 2005]. Additionally, oyster specimens can easily be acquired from oyster farms or simply collected from the shores. Oysters have separated sexes, but can switch sex from male to female (protandric) after a few annual cycles, which makes them suitable for studies on sexual dimorphism. Embryo and larval development can easily be monitored and sampled using the well-established oyster embryo-larval bioassay protocol [Leverett & Thain, 2013]. It usually takes 24-30 h from fertilisation to a free swimming D-shelled larva, including all stages of embryonic development (see Figure 1.3.1., Chapter 3). Additionally, the Pacific oyster has a high rate of reproductive cycles and can be conditioned in the laboratory throughout the year. As filter feeding species oysters take up particles and chemicals from the water column and are known to accumulate pollutants [Salazar; 2002]. Compared to many other bivalve species *C. gigas* has one decisive advantage. The genome of the Pacific oyster has been sequenced [Zhang et al., 2012], providing information on genetic and molecular data. Research on oysters is highly demanded, particularly due to its high economic interest as a popular food source. Further understanding of

oyster endocrinology, development and health effects of pollutants on oyster stocks, including underlying molecular mechanisms, is therefore of particular interest for aquaculture farms, since it could allow for an effective manipulation of various attributes, such as size, sex or reproduction cycles [Camara & Symonds, 2014]. This is primarily of relevance for the food industry in terms of productivity and profit, but also from a human health perspective.

Chapter two

# The nuclear receptor gene family in the Pacific oyster, *Crassostrea gigas*, contains a novel subfamily group


## Abstract

Nuclear receptors are a superfamily of transcription factors important in key biological, developmental and reproductive processes. Several of these receptors are ligand- activated and through their ability to bind endogenous and exogenous ligands, are potentially vulnerable to xenobiotics. Molluscs are key ecological species in defining aquatic and terrestrial habitats and are sensitive to xenobiotic compounds in the environment. However, the understanding of nuclear receptor presence, function and xenobiotic disruption in the phylum Mollusca is limited. Here, forty-three nuclear receptor sequences were mined from the genome of the Pacific oyster, *Crassostrea gigas.* They include members of NR0-NR5 subfamilies, notably lacking any NR6 members. Phylogenetic analyses of the oyster nuclear receptors have been conducted showing the presence of a large novel subfamily group not previously reported, which is named NR1P. Homologs to all previous identified nuclear receptors in other mollusc species have also been determined including the putative heterodimer partner retinoid X receptor, estrogen receptor and estrogen related receptor. *C. gigas* contains a highly diverse set of nuclear receptors, including a novel NR1 group, which provides important information on presence and evolution of this transcription factor superfamily in invertebrates. The Pacific oyster possesses two members of NR3, the sex steroid hormone receptor analogues, of which there are nine in humans. This provides increasing evidence that steroid ligand specific expansion of this family is deuterostome specific. This new knowledge on divergence and emergence of nuclear receptors in *C. gigas* provides essential information for studying regulation of molluscan gene expression and the potential effects of xenobiotics.



This chapter is a revised and reformatted copy of the publication: **Vogeler S, Galloway TS, Lyons BP, Bean TP. The nuclear receptor gene family in the Pacific oyster, *Crassostrea gigas*, contains a novel subfamily group. BMC Genomics. 2014; 15:369. Small changes to the original publication may occur.** SV, TG, BL and TB contributed to study concept, design, interpretation of data and manuscript preparation. SV and TB performed the BLAST searches and SV conducted the phylogenetic analyses and laboratory studies to verify expression.

## 2.1. Introduction

### 2.1.1. *Nuclear receptors*

Nuclear receptors (NRs) are transcription factors, which regulate the expression of specific genes involved in embryonic development, homeostasis and physiologically regulated processes. They are of particular interest as this regulation often requires interaction with endogenous or exogenous ligands. Nuclear receptors bind to a response element in target gene promoters and activate gene transcription in cooperation with bound co-factors [Germain et al., 2006]. Although activation of NRs often requires interaction with ligands, there are many proteins which can be 'constitutively activated' and perform their biological response without a ligand. Nuclear receptors are usually found in protein complexes as monomers, homodimers, or heterodimers [Mangelsdorf et al., 1995], with one member of the NR2 family, the retinoid X receptor (RXR) operating as the predominant heterodimer partner in vertebrates [Mangelsdorf & Evans, 1995]. Structures of NRs are well characterized and usually contain six common structural features (Figure 2.1). The A/B region and the final F-region are highly variable and account for most of the difference observed between genes. The A/B region contains the activation function AF-1, which is able to synergize with AF-2 in region E to produce a more stable up-regulation of gene expression. Region C, the central specific DNA binding domain (DBD), is a highly conserved domain including two $C_4$ zinc-fingers (alpha helices): (1) a five amino acid sequence (P-box) determining the specificity of DNA binding, and (2) the D-box, which mediates receptor dimerization. Region D is a flexible "hinge" domain and connects the DBD with region E, the ligand binding domain (LBD). The LBD is highly conserved in structure, and moderately conserved in sequence. It is often able to bind specific hormonal (e.g. thyroids, steroids) or non-hormonal (morphogens, dietary components) hydrophobic ligands and it can induce or inhibit the expression of a gene by a conformational change of the receptor [Germain et al., 2006].

Nuclear receptors are exclusive to multicellular metazoans. Their numbers in animals range from a few receptors in sponges and *Trichoplax* [Baker, 2008; Bridgham et al., 2010; Larroux et al., 2006; Srivastava et al., 2008], to approximately 21 NRs in *Drosophila melanogaster* [King-Jones & Thummel, 2005] and 48 NRs in humans. The nematode *Caenorhabditis elegans* possesses the highest number of NRs

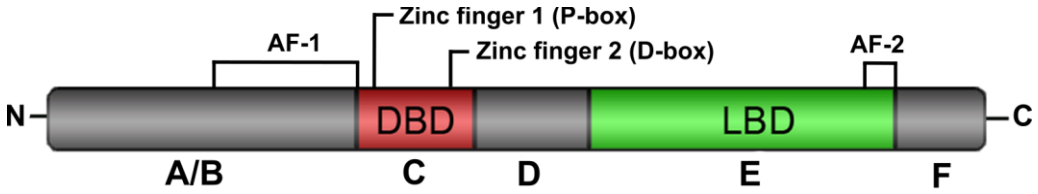



**Figure 2.1: Nuclear receptor gene structure.** The six regions (A-F) of nuclear receptors: the A/B region contains the AF-1 activation function. The highly conserved central DNA-binding-domain DBD (C region) comprises two zinc fingers, including the P-box and D-box. The ligand-binding-domain LBD (E region) contains the AF-2 activation function helix. Situated between the DBD and the LBD is the variable "hinge" region (D- region). The C-terminal region F is located at the end of the NR and varies in length among different nuclear receptors.

identified in a species with over 270 NRs [Maglich et al., 2001]. Two NRs have been found in the demosponge *Amphimedon queenslandica*, suggesting that NRs originated from a single NR in the ancestral metazoans [Bridgham et al., 2010]. This theory is supported by the deep conservation of the DBD and the LBD sequences between different animal phyla and suggests the divergence of NRs is most likely driven by gene duplication and gene loss [Bertrand et al., 2004; Bridgham et al., 2010; Laudet, 1997; Laudet et al., 1992].

Nuclear receptors are divided into six subfamilies based on phylogenetic reconstructions of the DBD and LBD [Nuclear Receptors Nomenclature Committee, 1999]. Abnormally structured NRs, which do not contain one of the two conserved regions (DBD or LBD), are grouped in a separate subfamily (NR0) irrespectively of their phylogenetic relationship. A novel group of NRs has been identified in Platyhelminthes *Schistosoma mansoni,* containing tandem DBDs and a single LBD, which do not belong to the miscellaneous NR0 subfamily and are categorised as 2DBDNR group [Wu et al., 2006, 2007].

### 2.1.2. *Nuclear receptors in the Mollusca*

Mollusca (gastropods, bivalves, cephalopods and relatives) diverged rapidly during the Cambrian period resulting in a large range of morphology and life histories, becoming the second most species-rich phylum among the invertebrates [Parkhaev, 2008]. Molluscs sit within the Lophotrochozoa, one of the two major groups among the protostomes. Marine mollusc species are common inhabitants of rocky, intertidal and estuary flats world-wide. They occupy important ecological niches as filter feeders and decomposers, and serve as a protein source for animals, including

humans, linking them directly with human health. Molluscs are recommended as ideal sentinel species in a number of marine monitoring programmes including those supported by international bodies such as ICES and OSPAR [Davies & Vethaak, 2012]. Their large differences in anatomy and life cycle, their wide global distribution, bioaccumulation of chemicals by filtration and relatively straightforward capacity to be cultured and handled in the laboratory makes molluscs an ideal species for studying biological processes. They are also often considered as surrogates for vertebrate models in laboratory based chemical risk assessment studies [Matthiessen, 2008]. However, the information on similarities and differences between vertebrate and mollusc endocrine system and gene regulation is limited and a deeper insight as to how molluscs are affected by chemicals will therefore directly aid the development of ecological and chemical risk assessment and enhance protection of the marine environment.

Several NR sequences, including the conserved domains, have previously been isolated and characterized in molluscs. The estrogen receptor ER (NR3A) in the gastropod *Aplysia californica* [Thornton et al., 2003] was the first NR identified in a mollusc species. Since then single ER homologs have been identified in more than eleven species among three main classes of the phylum Mollusca: gastropods (6), bivalve (4) and cephalopods (1) (Supplementary Table 2.1). A second member of the NR3 subfamily, the estrogen related receptor, ERR, has been cloned in the gastropod snail *Marisa cornuarietis* [Bannister et al., 2007] and a single RXR (NR2B) representative has been identified in at least six species among the molluscs (Supplementary Table 2.1). Additionally, a retinoic acid receptor RAR (NR1B) has been cloned from the central nervous system of the pond snail *Lymnaea stagnalis* [Carter & Spencer, 2009]. Finally, three molluscan receptors have been reported in the bivalve *Mytilus galloprovincialis,* possessing one homolog to the NR1D group, one NR that is closely related to the NR1D, NR1E and NR1F groups, and one receptor that was related to the nematode and trematode receptors SEX-1 (NR1G) [Raingeard et al., 2013].

### 2.1.3. *Nuclear receptors and xenobiotics*

Due to their ligand binding abilities, some NRs are inherently vulnerable to xenobiotics, which can modulate normal gene expression by mimicking a ligand or

blocking the LBD of NRs [Omiecinski et al., 2011]. This can lead to abnormal gene expression and hence, to disruption of development and/or endocrinology of an organism. Various xenobiotics, which have a mode of action mediated through NRs, have thus been characterized as endocrine disrupting chemicals (EDCs). Published reports have interpreted EDCs as having caused serious effects of chemicals to the health conditions of human and wildlife [Bergman et al., 2013; Diamanti-Kandarakis et al., 2009]. Mass mortalities and population declines in approximately 20 gastropod species worldwide [Bergman et al., 2013; Fioroni et al., 1991; Horiguchi et al., 1997, Shi et al., 2005, Sternberg et al., 2010] have been associated with exposure to tributyltin (TBT). This biocide was employed in antifouling paint on ships and fishing nets from the early 60s until 2005, when its use was legally restricted. In gastropods, exposure to TBT causes irreversible superimposition of male genitals on females, a condition termed imposex, whilst in bivalve species, exposure to TBT causes growth reduction [Salazar & Salazar, 1996; Thain, 1986; Widdows & Page, 1993], and shell thickening [Chagot et al., 1990; Higuera-Ruiz & Elorza, 2009, 2011; Waldock & Thain, 1983]. The mechanism through which TBT affects mollusc species remains unclear, although hypotheses have been raised related to binding to and disruption of a putative molluscan RXR or RXR/ peroxisome proliferator-activated receptor (PPAR) heterodimer [Castro et al., 2007; Horiguchi et al., 2007, 2008, 2010a, 2010b; Lima et al., 2011; Nishikawa, 2006; Nishikawa et al., 2004; Pascoal et al., 2013; Sternberg et al., 2008, 2010; Urushitani et al., 2011].

In this study, we took advantage of the recently released complete genome of the Pacific oyster *Crassostrea gigas* [Zhang et al., 2012]. The published oyster genome has previously been automatically annotated and NR genes were identified in the genome. However, automatic annotations often contain errors and potentially denoted homologs might not be actual homologs to a known gene [Koonin & Galperin, 2003; Schnoes et al., 2009]. Therefore, we used a combination of bioinformatics and phylogenetics to analyse the NR gene family in the Pacific oyster. Here we report the phylogenetic relationship of 43 NRs, confirm expression and discuss their homology to *Homo sapiens*, *D. melanogaster*, *C. elegans* and previously cloned molluscan NRs. The data are assessed from the perspective of putative function, evolution and the potential for the binding of xenobiotics, based on previous functional studies on NR homologs in other species.

## 2.2. Materials and Methods

### 2.2.1. *Identification of nuclear receptors in* C. gigas *genome*

Putative *C. gigas* NR sequences were identified through a local combination of tBLASTn and BLASTp searches of genome, CDS and protein databases using the published Pacific oyster genome [Zhang et al., 2012]. The protein sequences (full length sequences, isolated DBD regions and isolated LBD regions) of the 48 *H. sapiens* and 21 *D. melanogaster* NRs were downloaded from GenBank and were used as templates for interrogating the *C. gigas* databases. The DBD (zf-C4) and LBD (hormone_rec) of identified putative oyster NRs were verified by using Pfam (Pfam 26.0) [Finn et al., 2010] and in addition annotated by using the Conserved Domain Database at NCBI [Marchler-Bauer et al., 2011]. A BLASTp search using the conserved domains and the full length sequences against the non-redundant Metazoan protein database at NCBI was used for a first characterization of the putative NRs.

Nomenclature of the putative *C. gigas* NRs was based on phylogenetic analyses using conserved domains and sequence similarities of full length sequences to the NRs from *H. sapiens* and *D. melanogaster*. Genes were classified to NR subfamily groups based on the nomenclature guidelines [Nuclear Receptors Nomenclature Committee, 1999], if a single representative was identified. For groups, which include several representatives, the nomenclature name of the closest orthologs were given or listed with the Greek suffix α-δ. Nuclear receptors, which showed similarities to two or more NR subfamily groups, were named after all group names. Nuclear receptors, which could not be assigned to a NR subfamily group or for which sequence could not be resolved, were labelled as CgNHRs.

### 2.2.2. *Verification* C. gigas *nuclear receptor expression*

Six *C. gigas* individuals were sampled from the coastline close to Starcross, Devon, UK (50.6167 °N, 3.4500 °W). Five shucked whole animals were frozen in liquid nitrogen and ground to a medium fine powder by hand under constant supply of liquid nitrogen to prevent RNA degradation. Total RNA was extracted from each individual using TRI Reagent RNA Isolation Reagent (Sigma-Aldrich) following the

manufacturer's protocol [Chomczynski & Mackey, 1995]. For the RNA extraction 1 ml TRI Reagent has been mixed with 100 mg powder and homogenised using a handheld homogeniser. DNA was removed with RQ1 RNase-Free DNase (Promega). RNA was cleaned using the RNeasy Mini Kit (QIAGEN) and pooled. Cleaned total RNA was converted to cDNA with the ThermoScript RT/PCR System (Invitrogen) using oligo (dT) primers. Forward and reverse primers for 42 putative NRs were designed with Primer-Blast at National Centre for Biotechnology Information (NCBI) [Ye et al., 2012] to amplify either parts of the hinge domain plus LBD, or parts of LBD or whole DBDs with predicted amplicons of 177-984 bp (Supplementary Table 2.2). Primers for CgNR1D were also designed to obtain the NR sequence. One microlitre of undiluted cDNA was used for PCR amplification with the GoTaq system (Promega) under the following conditions: 95 °C 5 min, thirty-five cycles at 95 °C for 15 s, 57 °C for 30 s, 72 °C for 1.5 min, and a final extension at 72 °C for 5 min. Amplified PCR products were visualized on a 1.5 % agarose gel and amplicons purified with the QIAquick PCR Purification Kit, or with QIAquick Gel Extraction Kit (Qiagen, UK) and sequence verification conducted by Eurofins MWG Operon (Ebersberg, Germany).

### 2.2.3. *Phylogenetic analysis*

Thirty-eight classically and five abnormally structured putative NRs from *C. gigas* were compared to 48 *H. sapiens*, 21 *D. melanogaster*, 12 *C. elegans*, 3 *S. mansoni* and 16 previously cloned NR amino acid sequences from different mollusc species (*A. californica, Biomphalaria glabrata, C. gigas, L. stagnalis, M. cornuarietis, Mytilus edulis, M. galloprovincialis, Nucella lapillus, Octopus vulgaris, Thais clavigera*). The lineage-specific expansion of *C. elegans* in the NR2A subfamily was disregarded as preliminary data of *C. gigas* NRs did not suggest a similar NR2A expansion. For a better readability of the phylogenetic trees only one representative of the *C. elegans* NR2A subfamily have been used. The amino acid sequence GenBank accession numbers of all NR used in the phylogenetic analysis is available in Supplementary Table 2.1. The DBD and LBD amino acid sequences were aligned using default parameters in MUSCLE v3.8.31 [Edgar, 2004] and edited manually in case of errors. LBD domains were trimmed to allow efficient alignment of conserved sequences. Three separate maximum likelihood (ML) analyses were conducted, the first using

only the DBD, the second with a portion of the LBD, and the third with the DBD plus a portion of the LBD. Trees were constructed using PhyML v3.0 [Guindon et al., 2010] with a LG +I +G matrix (model determined by AIC criteria with ProtTest v2.4) [Abascal et al., 2005]. Nodes were supported by ML analyses assessed with 1,000 bootstraps. The same three data sets were also tested by Bayesian Inference, carried out under a proportion of invariable sites and gamma-distributed rate heterogeneity among sites with a mixed amino acid replacement model using MrBAYES v3.2.2 [Huelsenbeck & Ronquist, 2001]. The trees started randomly with four simultaneous Markov chains running for 5 million generations with chains sampled every 100 generations and with a burnin of 5000 trees. The JTT model [Jones et al., 1992] was selected as the best fitting substitution model. The Bayesian posterior probabilities (PPs) were calculated using a Markov chain Monte Carlo (MCMC) sampling approach implemented in MrBAYES v3.2.2. Additional phylogenetic support was conducted by a distance neighbour-joining (NJ) analysis of the DBD plus the portion of LBD using Seaview v4.0 [Gouy et al., 2010]. Default characteristics were used and the branch support was measured by bootstrap analysis with 1,000 replicates. Phylogenetic trees were visualized and illustrated with FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

## 2.3. Results

### 2.3.1. *Nuclear receptor genes*

Forty-three putative NR genes were identified in the *C. gigas* genome. Transcription of all NR genes was successfully confirmed by sequencing. Putative NR affiliation was verified based upon the conserved domains, DBD and LBD, using a Pfam analysis and a conserved domain search resulting in 38 NRs showing the classical structures of the NR superfamily. One of the NR identified putatively appeared to suggest a sequencing error in the genome project, with a single LBD and a lack of DBD. Therefore, the full gene was re-sequenced as the NR homolog, CgNR1D, showing a single DBD and a single LBD. Five putative oyster NRs have abnormal structures including two NRs containing two DBDs and a single LBD, one NR lacking the DBD but containing a single LBD, and two NRs with only a single DBDs and lack of a LBD. A full list of annotated protein sequences of *C. gigas* NRs including accession numbers is provided in the Supplementary Table 2.1.

### 2.3.2. *Phylogenetic analysis*

Phylogenetic analyses were performed using the amino acid sequences of the 43 *C. gigas* NRs. Several trees were constructed using different phylogenetic analyses: the DBD tree, based on just DNA binding domains (Maximum Likelihood (ML) and Bayesian Inference analyses, 38 classically and 4 abnormally structured *C. gigas* NRs); LBD tree, based on a portion of the LBD (ML and Bayesian Inference analyses, 38 classically and 3 abnormally structured *C. gigas* NRs); and DBD plus LBD trees, based on a composition of DBD and a portion of LBD (ML, Bayesian Inference and neighbour-joining (NJ) analyses, 38 classically structured *C. gigas* NRs). The ML and Bayesian Inference phylogenetic analyses of the DBD plus LBD alignment showed similar pattern and both segregated in a monophyletic group NR1 and a second major clade containing the subfamilies NR2-NR6 (Figure 2.2). The second major group further subdivided in five sub-clades NR2, NR4, NR3, NR5 and NR6. Nodes for NR1, NR3-NR6 were supported by high ML bootstrap scores (BS= 81-100) and high posterior probabilities (PP= 0.99-1). The NR2 clade was more moderately supported (BS= 76), but highly supported by posterior probabilities

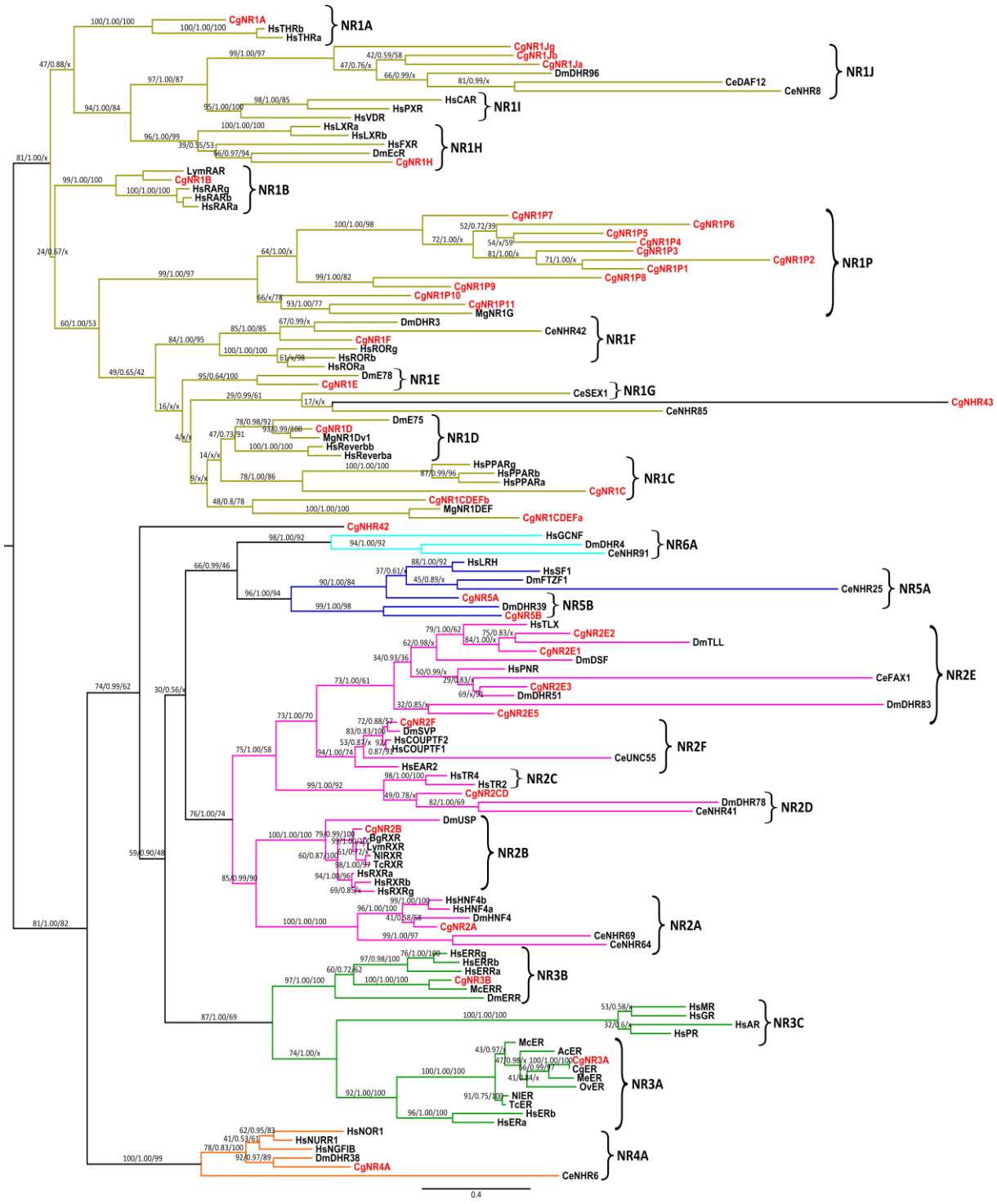



Figure 2.2 legend: see next page

(PP= 1). The NJ analysis of the DBD plus LBD segregated in three major clades: two including NR1 subfamily members and the third subdivided in the other NR2-NR6 subfamilies, which displayed different positioning of the NR3, NR5 and NR6 subfamilies compared to the ML and Bayesian Inference analyses. In general, the NJ analysis provided the least resolution of the DBD plus LBD trees and therefore, was only used as additional support for individual receptor's placements. The individual ML and Bayesian Inference analyses of the separate DBD and LBD sequences resulted in less supported nodes for the six receptor subfamilies and were also not able to assign some of the receptor subfamilies to the existing monophyletic subfamilies (Supplementary Figure 2.1 & 2.2). Therefore, the phylogenetic relationships of the putative *C. gigas* NRs were deduced from the DBD plus LBD ML analysis supported by the Bayesian Inference analysis and NJ bootstraps values (Figure 2.2).

*C. gigas* possesses NRs belonging to six of the seven NR subfamilies. Twenty-three out of 38 classically structured oyster NRs are members of the NR1 subfamily (Figure 2.2). A novel NR1 group, NR1P, has been formed including eleven *C. gigas* receptors and the *M. galloprovincialis* NR MgNR1G. Subfamily NR2 is represented by eight oyster NRs. There are also two NR3 members, one NR4 member and two NR5 members. No homologs to the NR6 receptor subfamily were identified in the *C. gigas* genome. The abnormally structured *C. gigas* receptor CgNR0B showed homologies to the miscellaneous subfamily NR0B (Supplementary Figure 2.2). CgNHR40 and CgNHR41, the two single DBD sequences, assigned to the CgNR1Ja receptor and outside the NR4 group, respectively (Supplementary Figure 2.1). The

**Figure 2.2: Phylogenetic relationship of nuclear receptors in *Crassostrea gigas*, *Homo sapiens*, *Drosophila melanogaster*, *Caenorhabditis elegans* and mollusc species.** The alignment was constructed using the DBD plus portion of LBD and phylogenetic relationship was conducted by a Maximum likelihood (ML), Bayesian Inference and neighbour-joining (NJ) analyses. ML bootstrap support values (percentage of 1000 BS), Bayesian posterior probabilities (PPs) and NJ bootstrap support values (percentage of 1000 BS) are provided above the nodes separated by slash. Star indicates the node obtained from the Bayesian Inference and NJ analyses, which were different from that obtained by ML method. Highlighted clades display the six NR subfamilies, olive: NR1, pink: NR2, green: NR3, orange: NR4, dark blue: NR5, light blue: NR6. *C. gigas* NRs highlighted in red. Ac: *Aplysia californica,* Bg: *Biomphalaria glabrata,* Ce: *C. elegans*, Cg: *C. gigas,* Dm: *D. melanogaster*, Hs: *H. sapiens*, Lym: *Lymnaea stagnalis*, Mc: *Marisa cornuarietis*, Me: *Mytilus edulis*, Mg: *Mytilus galloprovincialis*, Nl: *Nucella lapillus,* Ov: *Octopus vulgaris,* Tc: *Thais clavigera.*

two NRs containing double DBDs and a single LBD display phylogenetic relationship to the 2DBD NR group previously identified in *S. mansoni* (Supplementary Figure 2.1 & 2.2). Cg2DBDγ showed structural and amino acid identity to the Sm2DBDγ, with two DBDs and the LBD sequence identities of 55%, 60% and 25% respectively. Cg2DBDδ, however, does not display such a high relationship and its second DBD clusters outside the 2DBD receptor group close to CgNR1CDEFα and MgNR1DEF (ML analysis DBD). In addition, its LBD was weakly supported in the ML analysis and not supported by the Bayesian Inference analysis.

The large majority of *C. gigas* NRs assignments to subfamily groups were supported by high ML bootstrap scores (BS= 89-100) and Bayesian posterior probabilities (PP= 1). Exceptions include CgNR1C and the entire NR2E group containing four *C. gigas* NRs, which were highly supported by the Bayesian Inference analysis (both PP=1) but only moderately supported by ML (BS= 78 & 73, respectively). Further positions within the groups among phylogenetic trees were fixed apart from the following exceptions. Two classically structured NRs (CgNHR42 and CgNHR43) could not be assigned to one of the receptor subfamilies. CgNHR42 was located as outgroup to NR2/3/5/6 clade in the DBD plus LBD analyses and the LBD analyses, while the DBD analyses nested it inside the NR1 subfamily. BLASTp search against the non-redundant Metazoan database search of conserved domains and the full length sequence showed homologies to RXRs, RARs and ERRs of various species, further suggesting this is an outlier. CgNHR43 DBD grouped with the NR6 subfamily for the ML analysis, but with relatively weak support, and with NR1 for the Bayesian analysis. This was the only *C. gigas* NR that displayed any homology to the NR6 subfamily for either the separate or combined conserved domains sequence. However, the LBD and the DBD plus LBD analyses nested CgNHR43 deeper in the NR1 subfamily. Although CgNH42 and CgNH43 have a classical NR structure and are expressed in oyster tissue, there is potential that they are unitary pseudogenes. However, as it is difficult to resolve the nearest common ancestor to these genes, this theory has not been tested and will require future functional assessment.

CgNR1CDEFα and CgNR1CDEFβ associated with the MgNR1DEF and nested with the NR1C, NR1D, NR1E and NR1F groups. The individual conserved domain analyses assigned CgNR1CDEFβ either to the NR1F group (DBD analyses) or NR1E group (ML LBD analysis) than to CgNR1CDEFα or MgNR1DEF. Results of BLASTp search against metazoan protein database found homologies to MgNR1DEF and

NR1C-NR1F members of various species for the DBD, LBD and full length sequences.

There were few differences within receptor groups among the phylogenetic trees. The novel receptor group NR1P was highly supported in the DBD plus LBD and the individual LBD phylogenetic analyses (BS= 97-100, PP= 1)) and weakly supported in the DBD analysis (BS= 41, PP= 0.81), but the MgNR1G always associated with this group. However, the arrangement of the *C. gigas* receptors within the NR1P group varied marginally between trees. Similar small differences were identified for the four *C. gigas* homologs of the NR2E group. The relationship between all members of this group changed depending on which receptor unit and analysis were used. The DBD analyses showed a few dissimilarities. CgNR1H was placed closer to the human farnesoid X receptor (FXR) compared to the other trees, and revealed higher homologies to the *D. melanogaster* ecdysone receptor EcR; the Hepatocyte Nuclear Factor 4 (HNF4) homolog CgNR2A is more closely related to the human members of NR2A than to the *Drosophila* homolog; and the molluscan orthologs of NR2B assigned closer to the *Drosophila* ultraspiracle protein USP than to the human RXR homologs.

## 2.4. Discussion

In this study, 43 NRs were identified in the bivalve *C. gigas,* representing six of the seven common NR subfamilies. This large set of *C. gigas* NRs provides an overview of NR presence in the class Bivalvia and it can enhance the understanding of NR evolution in invertebrates and the biological processes in which NRs are involved. Furthermore, it presents information on possible xenobiotic targets in mollusc species, which occupy an important position in terms of the evolution of the protostomes and as key ecological species in aquatic habitats.

### 2.4.1. *Novel group NR1P*

The phylogenetic analyses of *C. gigas* NRs revealed a novel monophyletic group among the NR1 subfamily including eleven oyster NRs. In addition this group includes one previously characterised *M. galloprovincialis* receptor, MgNR1G. The novel group was supported by high bootstrap values and therefore named as NR1P. The results indicate that the MgNR1G receptor, which was previously assigned to the *C. elegans* Sex-1 NR [Raingeard et al., 2013], is not a real homolog to the NR1G group, but a member of the novel NR1P group. Putative functions of these NRs cannot be deduced based on the phylogeny as no close homolog could be identified. However, many members of the NR1 subfamily are involved in growth and development in humans or in moulting and metamorphosis processes in *Drosophila*. A BLASTp search of the conserved domains of NR1P1-NR1P9 against the non-redundant metazoan database showed weak homologies to NR1C-NR1F invertebrate and vertebrate members, but also relationships to other NR1 groups. CgNR1P10 and CgNRP11 domains displayed homologies to the NR2E group. These differences in homology are reflected in the different sup-grouping of the NR1P group. However, disagreement in CgNR1P8 and CgNR1P9 positioning among the phylogenetic trees does not allow a separation of NR1P in two separate groups.

The phylogenetic analysis suggests that NR1P segregated from a common ancestor of the NR1C, NR1D, NR1E and NR1F groups. The molluscan phylum separated early among the Protostomia [Winnepenninckx et al., 1998] and could have evolved a unique group of NRs. However, it is not clear if this novel group is mollusc specific or also present in other lophotrochozoans. Interestingly, C*. gigas* possesses homologs of NR1C, NR1D, NR1E and NR1F groups, which are also present in

Ecdysozoa, the sister clade of the Lophotrochozoa. Therefore, it is possible that some ecdysozoans contain NR1P homologs. However, no NR1P homologs have been identified for *D. melanogaster* (Arthropoda) or *C. elegans* (Nematoda). The NR set identified in *Daphnia pulex* (Crustacea) revealed a novel group among the NR1 subfamily, but this group showed high sequence similarities to the invertebrate NR1J group [Thomson et al., 2009]. It is also possible that the ecdysozoans have lost this particular receptor group as it has been reported for other NRs [Bertrand et al., 2004; Bridgham et al., 2010].

The existence of another novel receptor group in *C. gigas* is implied by the presence of CgNR1CDEFα and CgNR1CDEFβ, which associated with the MgNR1DEF as outgroup to NR1C-NR1F. However, CgNR1CDEFα and CgNR1CDEFβ were not consistent in their positions for all phylogenetic analyses and therefore, unambiguous assignments are difficult. This inconsistency could be a consequence of rapid evolutionary divergence [Raingeard et al., 2013]. It is presumed that the members of the NR1C-NR1F groups originate from a common ancestor, but separated very early in invertebrate evolution [Bertrand et al., 2004; Thornton et al., 2003]. Alternatively, the changing position of CgNR1CDEFβ in the phylogenetic analyses could be a result of sequence similarities. This would be supported by vertebrate RAR-related orphan receptors RORα-γ (NR1F1-3) and REV-ERB receptors, which compete for the same response elements with their DBD [Forman et al., 1991]. Additional NR sets of more closely related protostome species are required to confirm the final relationship of these two NRs.

### 2.4.2. C. gigas *receptors with functionally characterised homologs*

Analysis of the *C. gigas* genome identified a homolog, CgNR1A, to the human thyroid receptors, THRα and THRβ. Vertebrate THRs bind thyroid hormones and play important roles in growth, development and metabolism, and bind either as monomers, homodimers or form heterodimers with RXRs [Yen, 2001]. Though the function of THRs in invertebrates is still unknown, the previously identified flatworm (*S. mansoni*) THRs are able to bind to DNA either as monomers or homodimers and can function as a heterodimer with SmRXR [Wu et al., 2006, 2007].

CgNR1B is a homolog to the human RARα-γ paralogs and showed high homology to the cloned freshwater snail *L. stagnalis* RAR [Carter & Spencer, 2009], with the

DBDs and LBDs of these having an amino acid identity of 90 % and 73 %, respectively. Exposures to all-*trans* retinoid acids (RA) and 9-*cis* RA, known agonistic ligands to vertebrate RAR [Allenby et al., 1994], and to a human RARβ-selective antagonist caused significant deformations to the eyes and shell in *L. stagnalis* embryos [Carter & Spencer, 2009; Creton et al., 1993]. In vertebrates, RARs regulate the expression of genes involved in morphogenesis and especially embryonic development [Marlétaz et al., 2006].

CgNR1C grouped with the human paralogs of NR1C (PPARα-γ), which bind endogenous ligands, including eicosanoids, fatty acids and fatty acid derivatives. PPARα controls the uptake of fatty acid and their esterification into triglycerides. PPARγ is the main regulator of adipogenesis, fat storage and glucose homeostasis. PPARγ, together with its heterodimer partner RXR, is a potent inducer of adipogenesis in vertebrates when exposed to organotin compounds [Grün et al., 2006]. PPARβ/δ is involved in fatty acid oxidation, as well as energy consumption and thermogenesis. PPARs are also a target of the fibrate and thiazolidinedione drugs. These are classified as PPARα and PPARγ activators and are used in hyperlipidemia and hyperglycemia treatments [Singh et al., 2011].

CgNR1D and MgNR1Dv1 group together with the NR1D group human and *Drosophila* homologs. The insect E75 receptor (NR1D3) is induced by ecdysteroids and is involved in moulting and metamorphosis [Dubrovskaya et al., 2004; King-Jones & Thummel, 2005]. The human counterparts, REV-ERBα (NR1D1) and REV-ERBβ (NR1D2), display some similar functions to PPARs, playing important roles in lipid and glucose metabolism, gas-response, inflammation and circadian rhythm [Ramakrishnan & Muscat, 2006].

The oyster genome also possesses a homolog, CgNR1E, to the *Drosophila* E78 receptor. E78 is directly related to ecdysone signalling and it is another important receptor during metamorphosis [Stone & Thummel, 1993].

CgNR1F is an ortholog to the *D. melanogaster* DHR3 and *C. elegans* NHR23 receptors. DHR3 is inhibited by E75 and regulates metamorphosis by repressing genes [Lam et al., 1999]. Expression of the *E75* gene is regulated by another NR, ftz-transcription-factor-1, FTZ-F1 (NR5A3) [Dubrovsky et al., 2011]. The human members of the NR1F group, RORα-γ, play a role in circadian rhythm, immune response and other important physiological processes [Jetten, 2009].

The oyster genome contains a NR (CgNR1H), which is a homolog to the *D. melanogaster* EcR of NR1H group. EcR has also been found in crustaceans [Kato et

al., 2007; Thomson et al., 2009], in the genome of the mollusc *Lottia gigantea*, leeches and Polychaeta [Laguerre & Veenstra, 2010]. EcR is involved in moulting, developmental and reproductive processes in insects [King-Jones & Thummel, 2005] and crustaceans [Kato et al., 2007]. In addition, EcR agonists and antagonists are commonly used as insecticides [Retnakaran et al., 2003]. The vertebrate homologs are liver X receptors (LXRs) and FXR, which regulate lipid and cholesterol metabolism, bile salt synthesis and control expression of certain cytochrome P450s (CYP) [Kalaany & Mangelsdorf, 2006].

*C. gigas* possesses three NRs of the NR1J group, of which CgNR1Jα and CgNR1Jβ grouped together and CgNR1Jδ assigned on the fringe of the NR1J group. A fourth homolog might be present in the genome as well, indicated by the putative incomplete CgNHR40 NR (Supplementary Figure 2.1), the sequence of which could not be fully resolved. This group appears to be unique to invertebrates, including ecdysozoans [King-Jones & Thummel, 2005, Maglich et al., 2001], crustaceans [Thomson et al., 2009] and Platyhelminthes [Wu et al., 2006]. All *C. gigas* NR1J representatives contained the group-unique residues ESCKAFFR in their DBD sequence [Antebi et al., 2000]. Characterised NR1J receptors include DHR96, which is believed to play a role in xenobiotic stress response [King-Jones et al., 2006] and is able to bind cholesterol to regulate cholesterol homeostasis [Horner et al., 2009]. Xenobiotic defence in *C. elegans* is thought to be managed by the NHR-8 [Lindblom et al., 2001]. DAF-12, also a *C. elegans* NR1J, is involved in dauer formation, in which larval development is diverted under adverse environmental conditions to a form of stasis termed the dauer stage [Antebi et al., 2000]. The NR1I subgroup is the vertebrate group homologous to NR1J and shares common receptor ancestors prior to the divergence of deuterostomes and protostomes [Bertrand et al., 2004]. Its three representatives, pregnane X receptor (PXR), constitutive androstane receptor (CAR) and vitamin D receptor (VDR), have all been implicated in the vertebrate response to xenobiotic stress [Reschly et al., 2006]

A single member, CgNR2A, of the NR2A (HNF4) group has been identified in the *C. gigas* genome. The NR2A group contains the most ancient NRs found in animals and have been discovered in simple metazoans [Baker, 2008; Bridgham et al., 2010; Grasso et al., 2001; Larroux et al., 2006]. Only a single ortholog is encoded in *D. melanogaster*. It is involved in the development of the digestive tract [Zhong et al., 1993], lipid metabolism and mobilization [Palanker et al., 2009]. In humans, HNF4

receptors play a significant role in diseases like diabetes [Love-Gregory & Permutt, 2007] and colon cancer [Chellappa et al., 2012].

The *C. gigas* RXR ortholog (CgNR2B) clustered together with other identified molluscan RXRs and the conserved regions had identities of over 93% to the retinoid X receptors' DBDs and LBDs of *B. glabrata* and *L. stagnalis*. The *B. glabrata* RXR is able to act as a heterodimer partner to human NRs and is also able to form homodimers [Bouton et al., 2005]. Retinoid acid (9-*cis* RA) and docosahexaenoic acid (DHA), natural ligands of vertebrate RXRs [Dawson & Xia, 2012], have been identified as putative gastropod agonistic ligands [Bouton et al., 2005], while vertebrate RXR pan-antagonists successfully inhibited growth cone turning in adult gastropod CNS and produced eye and shell deformation in embryos during the gastrulation stage [Carter et al., 2010].

The *C. gigas* receptor CgNR2CD could not be unambiguously assigned to either the NR2C or NR2D group. In humans the NR2C proteins, TR2 and TR4, act as transcriptional repressors in cooperation with co-factors [Cul et al., 2011]. The *Drosophila* NR2D ortholog DHR78 uses a similar repression mechanism (binding-site competition) and inhibits ecdysone signalling [Zelhof et al., 1995]. The *C. elegans* homolog NHR41 is also involved in moulting processes and morphogenesis [Gissendanner et al., 2004].

CgNR2E1, CgNR2E2, CgN2E3 and CgNR2E5 represent four putative NRs from the NR2E group. CgNR2E1 and CgNR2E2 have most identity to the *Drosophila* homologs tailless (DmTLL), dissatisfaction (DmDSF) and human TLX receptors. They are all involved in anterior-posterior axis formation and have important roles in vision and forebrain development [King-Jones & Thummel, 2005], as well as in emotional behaviour [Abrahams et al., 2005]. Furthermore, TLX regulates adult vertebrate neural stem cell proliferation [Sun et al., 2007]. CgNR2E3 shows homology to vertebrate photoreceptor cell-specific nuclear receptor (PNR), *C. elegans* FAX-1 and *Drosophila* DHR51. PNR is required for controlling neural differentiation and retina development [Halder et al., 2000; Milam et al., 2002] and neuron identity in *C. elegans* is regulated by FAX-1 [Much et al., 2000]. CgNR2E5 shows homologies to DHR83, a *Drosophila* function-unknown receptor [King-Jones & Thummel, 2005].

The *C. gigas* genome possesses one homolog (CgNR2F) to the NR2F group, which has a close phylogenetic relationship to the *D. melanogaster* seven-up receptor, DmSVP. Similar to other members of NR2E the DmSVP and HsCOUP-TF1/2

receptors are responsible for neural development and photoreceptor cells [Broadus et al., 1995; Mlodzik et al., 1990; Qiu et al., 1997; Zhou et al., 2001].

Two NR3 homologs were identified in the *C. gigas* genome. CgNR3A, previously identified as CgER, grouped well with other molluscan ERs and has been shown to be unresponsive to estrogen [Matsumoto et al., 2007]. The second, CgNR3B is a member of the NR3B group, representing the 'constitutively activated' ERRs, and shows a high degree of similarity to the previously identified *M. cornuarietis* molluscan ERR (DBD= 94.6% and LBD= 65.6%, respectively). McERR has been tested for modulation by vertebrate estrogens and other putative ligands *in vitro* and *in vivo*. However, no significant evidence for modulation could be identified and it is assumed that McERR is 'constitutively activated' [Bannister et al., 2007, 2013]. CgNR3B can be assumed to work in a similar way of action regarding ligand binding and activation, particularly as ligand-activated-requirement for a NR3B representative has not been identified either in invertebrates or vertebrates. No further NR3 subfamily receptor were identified by the genome analysis of *C. gigas*, which is consistent with the theory of the NR3 subfamily evolution [Bertrand et al., 2004; Eick et al., 2011; Markov & Laudet, 2011; Saez et al., 2010; Thomson et al., 2009] that the expansion of steroid receptors including the deuterostome specific NR3C group occurred after the divergence of protostomes and deuterostomes. However, we cannot rule out that additional steroid receptors exist in protostomes, or even in other mollusc species, and have then been lost during gene deletion events [Bertrand et al., 2011].

The CgNR4A receptor is the sole *C. gigas* member of the NR4 subfamily. There might be another homolog, CgNHR41, but this NR contains only a single DBD; the LBD could not be identified (Supplementary Figure 2.1). The human NR4 subfamily includes nerve growth factor I-B (NGFI-B), NR related 1 protein (NURR1), neuron derived orphan receptor 1 (NOR1), and is involved in a broad array of cellular metabolic processes; vascular remodelling and cancer [Mohan et al., 2012; Zhao et al., 2010]. In *Drosophila* species the NR4 gene DHR38 mediates an ecdysteroid signalling pathway [Baker et al., 2003] and the *C. elegans* homolog NHR-6 is involved in ovulation processes [Gissendanner et al., 2004]. The LBDs of NR4 genes found in humans, *D. melanogaster* and *S. mansoni* contain phenylalanines, which fill the entire volume of the ligand binding pockets and NR4 subfamily members are therefore suggested to be "true orphans" requiring no ligand [Baker et al., 2003; Codina et al., 2004, Flaig et al., 2005; Wu et al., 2008]. CgNR4A contains

phenylalanines at the same positions in its LBD, suggesting it too is a “true orphan” receptor.

The NR5 group is represented by two NR homologs in *C. gigas*, designated CgNR5A and CgNR5B. The CgNR5A DBD contains a highly conserved sequence (FTZ-F1 box), which is characteristic of the NR5A group. This sequence is located at the boundary between the DBD and the hinge region and is essential for the high-affinity interaction with the DNA [Ueda et al., 1992]. The *D. melanogaster* FTZ-F1 receptor is part of the ecdysteroid regulated NR group including members of different NRs subfamilies. Expression profiles of various fly NRs showed the close relationship between EcR, E75, E78, DHR3, DHR4, FTZ-F1 and DHR39 [King-Jones & Thummel, 2005; Sullivan & Thummel, 2003] and as a part of this regulatory cascade FTZ-F1 has a crucial role during embryonic development and metamorphosis [King-Jones & Thummel., 2005; Sullivan & Thummel., 2003]. The *C. elegans* NHR25 receptor is also associated with reproduction, embryogenesis and moulting processes in Nematodes [Antebi et al., 2000]. The human NR5A genes include liver receptor homolog-1 (LRH1), which regulates bile acid and cholesterol metabolism [Fayard et al., 2004], and the steroidogenic factor 1 (SF1), which is involved in reproductive development and endocrine function [Hoivik et al., 2010]. CgNR5B is a member of the NR5B group represented by the *D. melanogaster* DHR39 receptor. Besides its role in embryonic development and metamorphism DHR39 is also involved in female reproductive tract development [Allen & Spradling, 2008]. Homologs of this group are also identified in crustaceans [Thomson et al., 2009] and the invertebrate flatworm *S. mansoni* [De Mendonca et al., 2002]. *D. melanogaster* DHR39 and FTZ-F1 receptors also provide a good example that NRs are able to compete for the same DNA binding site as a putative target-gene regulated mechanism [King-Jones & Thummel, 2005; Ohno et al., 1994].

It is worth noting that *C. gigas* does not contain a NR6 subfamily homolog. Homologs have been identified both in protostomes (ecdysozoans e.g. *D. melanogaster* DHR4 [Sullivan & Thummel, 2003], *C. elegans* NHR91 [Maglich et al., 2001] and crustacean [Thomson et al., 2009]) and deuterostomes (germ cell nuclear factor (GCNF) in *H. sapiens*). Thus, it is likely that the NR6 subfamily homolog in the Pacific oyster could have been lost due to a gene loss event either during the separation of ecdysozoans and lophotrochozoans or during one of the evolutionary differentiations to the Pacific oyster.

CgNR0B is a predicted member of the miscellaneous NR-subfamily NR0B, lacking the DNA binding domain. The first human member of NR0B, DAX-1, plays major roles in steroidogenesis and reproductive development [Ehrlund & Teuter, 2012] and acts as a dominant-negative regulator of other NR transcription, e.g. SF1 and ER [Iyer & McCabe, 2004]. SHP, the second human NR0B representative, is involved in maintaining cholesterol and glucose homeostasis [Ehrlund & Teuter, 2012].

Two NRs, Cg2DBDγ and Cg2DBDδ, associate with the 2DBDNR group found in *S. mansoni* [Wu et al., 2006]. Cg2DBDγ and the *S. mansoni* receptors each contain the same unique P-box sequences, 'CEACKK', in the first DBD sequence [Wu et al., 2006, 2007].The two DBDs and the LBD sequence show amino acid identities of 52% and 25% respectively, to the Sm2DBDγ. Cg2DBDδ, which also possesses two DBDs, does not assign as closely to Sm2DBD receptors as Cg2DBDγ for its second DBD and LBD. In addition, it does not contain the unique P-box sequence in its first DBD. However, it contains another unique P-box sequence, 'CLPCKS', which has not been identified in any only NR's DBD. It seems that this 2DBD receptor may be unique to molluscs. Further phylogenetic analyses of NRs in other species and functional studies of Cg2DBDδ will reveal if this is a functional molluscan specific 2DBD receptor.

### 2.4.3. *Molluscan nuclear receptors as xenobiotic targets*

The variety of NRs in *C. gigas* and the known propensity for NRs to bind ligands prolifically provides the opportunity for xenobiotic disruption and chemically induced biological effects. Exposure to TBT can cause multiple developmental problems in *C. gigas*, including shell thickening [Chagot et al., 1990; Higuera-Ruiz & Elorza, 2009, 2011; Waldock & Thain, 1983]. TBT, a ligand for vertebrate RXR/PPAR heterodimers [Kanayama et al., 2005, le Maire et al., 2009], and thought to interact with molluscan RXRs [Castro et al., 2007; Horiguchi et al., 2007, 2008, 2010a, 2010b; Lima et al., 2011; Nishikawa, 2006; Nishikawa et al., 2004; Pascoal et al., 2013; Sternberg et al., 2008, 2010; Urushitani et al., 2011] may also target the *C. gigas* RXR homolog. Consequently, disruption of RXR function, the putative exclusive heterodimer partner of NRs, could cause alteration of the gene regulation and the reported malformations. Additionally, CgNR1C, the homolog to the human PPARs, could be affected by TBT. Gastropods exposed to rosiglitazone, a known vertebrate PPAR ligand, exhibited

similar effects (imposex) to TBT exposed animals [Pascoal et al., 2013]. A PPAR homolog in gastropods has not yet been identified, but it is likely that gastropod species also contain a PPAR homolog due to the high conservation of similar NR complements between related species [Bertrand et al., 2004].

Furthermore, the *C. gigas* PPAR homolog could be affected by environmental xenobiotics in addition to TBT and peroxisome proliferation in bivalve species is actually used as a biomarker for monitoring the health of aquatic environments [Zorita et al., 2007]. Organic xenobiotics such as polycyclic aromatic hydrocarbons (PAHs), phthalates and bisphenol-A, increase the number and volume of peroxisomes and induce peroxisomal β-oxidation enzyme acyl coenzyme A (acyl-CoA) oxidase in *Mytilus edulis* and *M. galloprovincialis* [Cajaraville & Zarragoitia, 2006; Cancio et al., 1998; Orbea & Cajaraville, 2006]. Vertebrate peroxisome proliferation and acyl-CoA are regulated by PPARs [Latruffe et al., 2003; Rakhshandehroo et al., 2010] and disturbance of PPAR regulated genes has been observed after exposure to the aforementioned xenobiotics [Desvergne et al., 2009; Kim et al., 2005; Kwintkiewicz et al., 2010].

The effect of xenobiotic vertebrate sex steroids is also a widely debated and has been investigated in a large range of mollusc species. Several studies have reported effects on reproductive output and morphology in different molluscan classes when exposed to vertebrate estrogen E2, synthetic estrogens and estrogen mimics [Scott, 2013], but their response remains ambiguous and characterised largely through hypothesis and homology. It was assumed that estrogen and other sex steroids (androgens, progestins, and corticoids) are used as reproductive hormones, operating through steroid receptors of the NR3 subfamily [Benstedead et al., 2011; De Lisa et al., 2012; Köhler et al., 2007; Matsumoto et al., 1997] and possessing a vertebrate-like sex steroid biosynthetic pathway [Janer & Porte, 2007; Lafont & Mathieu, 2007]. However, functional studies have shown that neither molluscan ERs nor ERR bind to estrogens or other sex steroids [Baker, 2007; Bannister et al., 2007, 2013; Kajiwara et al., 2006; Keay et al., 2006; Matsumoto et al., 2007; Thornton et al., 2003] and ER and ERR gene transcription was not affected by exogenous estrogens [Bannister et al., 2013; Puinean et al., 2006]. This present study identified two NR3 members, including an ER homolog, that does not bind estrogen [Matsumoto et al., 2007] and an ERR homolog, which is unlikely to bind estrogen. Additional NR3 members, which could interact with vertebrate sex steroids, could not be identified. This supports the hypothesis that any mechanism of sex steroid actions

in molluscs does not operate in a similar way to those in vertebrates and does not mediate via the NR3 group of NRs [Scott, 2012, 2013]. However, our findings do not provide any new information for the debate on the ligand state of the putative ancestral steroid receptor before the deuterostomes and protostomes have separated, which hypothesizes either a ligand-regulation, a sensory function or a constitutive action [Eick & Thornton, 2011; Markov & Laudet, 2011].

Nevertheless, estrogens and sex steroids still might have xenobiotic effects on mollusc species and possibly mediate via alternative NRs. For example, *C. gigas* possesses three NR1J homologs, which are known to respond to estrogens in *D. pulex* (DHR96). [Karimullina et al., 2012]. Similar results exist for vertebrate NR1I members, PXR and CAR [Kawamoto et al., 2000; Kretschmer et al., 2005].

## 2.5. Conclusion

This study verified the presence of 43 NRs in the Pacific oyster, *C. gigas.* Phylogenetic analyses demonstrate that the majority of *C. gigas* NRs are homologs to *D. melanogaster* and human NRs supporting the theory that these receptor groups emerged prior to the divergence of the Bilateria [Bertrand et al., 2004; Bridgham et al., 2010]. A novel group, NR1P, was discovered in *C. gigas*, which could not be identified in ecdysozoans or humans. Further studies of NRs in closely related mollusc species and in non-molluscan lophotrochozoans will discover if this novel group is mollusc specific or also present in other lophotrochozoan phyla. The *C. gigas* NR family does not contain any additional homolog to NR3 groups beside the ER and ERR and therefore, supports the theory that steroid ligand expansion of sex steroid NR3 subfamily is deuterostome specific.

*C. gigas* is a key ecological species and an important food source for humans, but due to its filter feeding lifestyle, it is susceptible to environmental pollution. This set of NRs provides important information on putative xenobiotic targets and the discovery of PPAR and RXR homologs in *C. gigas* encourages the theory of an RXR/PPAR heterodimer involvement in effect caused by TBT contamination. Additionally, we found further evidence that exogenous estrogens do not operate through a NR3 subfamily member, simply by the absence of an adequate NR3 candidate for estrogen binding.

The *C. gigas* NRs provide an important illustration of the presence and importance of this superfamily of ligand-regulated transcription factors, leaving the way open to future studies to analyse their functional significance.

Chapter three

# Dynamics of nuclear receptor gene expression during Pacific oyster development


## Abstract

Nuclear receptors are a highly conserved set of ligand binding transcription factors, with essential roles regulating aspects of vertebrate and invertebrate biology alike. Current understanding of nuclear receptor regulated gene expression in invertebrates remains sparse, limiting our ability to elucidate gene function and the conservation of developmental processes across phyla. Here, we studied nuclear receptor expression in the early life stages of the Pacific oyster, *Crassostrea gigas*, to identify at which specific key stages nuclear receptors are expressed. We used quantitative RT-PCR to determine the expression profiles of 34 nuclear receptors, revealing three developmental key stages, during which nuclear receptor expression is dynamically regulated: embryogenesis, mid development from gastrulation to trochophore larva, and late larval development prior to metamorphosis. Clustering of nuclear receptor expression patterns demonstrated that transcriptional regulation was not directly related to gene phylogeny, suggesting closely related genes may have distinct functions. Expression of gene homologs of vertebrate retinoid receptors suggests participation in organogenesis and shell-formation, as they are highly expressed at the gastrulation and trochophore larval initial shell formation stages. The ecdysone receptor homolog showed high expression just before larval settlement, suggesting a potential role in metamorphosis. Throughout early oyster development nuclear receptors exhibited highly dynamic expression profiles, which were not confined by gene phylogeny. These results provide fundamental information on the presence of nuclear receptors during key developmental stages, which aids elucidation of their function in the developmental process. This understanding is essential as ligand sensing nuclear receptors can be disrupted by xenobiotics, a mode of action through which anthropogenic environmental pollutants have been found to mediate effects.



This chapter is a reformatted copy of my publication: **Vogeler S, Bean TP, Lyons BP, Galloway TS. Dynamics of nuclear receptor gene expression during Pacific oyster development.' BMC Developmental Biology (accepted).** SV, TB, BL and TG contributed to study concept and design. Interpretation of data and manuscript preparation were conducted by all authors. The laboratory studies for quantitative PCR for gene expression analysis and the statistically analyses were carried out by SV.

## 3.1. Introduction

Nuclear receptors (NRs) are one of the largest classes of transcription factors in metazoan species and regulate many cellular functions through manipulation of gene expression. Although NRs are present even in the simplest animals in low numbers (demosponge: 2 NRs [Bridgham et al., 2010]), extensive diversification of NR families has occurred in Bilateria through gene duplications, gene loss and diversification [Bertrand et al., 2004; Bridgham et al., 2010 Laudet, 1997; Laudet et al., 1992]. In vertebrate species, NRs are essential for regulating gene expression during complex processes, in particular during development, which is one of the most dynamic periods of NR activity [Bruce & Campbell, 2010; Taneja, 2006]. For invertebrates, however, information on NR regulated gene expression is limited. Nevertheless, many developmental processes in metazoans are highly conserved and similar features are shared among diverse phyla. Understanding the presence and participation of NRs in phases of tightly controlled gene expression, particularly during developmental stages, is highly desirable. Not only is this a requirement for understanding the intrinsic biology, but also because alteration of NR function is one of the key routes through which normal biology can be disrupted by external factors, often resulting in abnormal phenotypes.

Typical NRs consist of five different domains, which include the highly conserved C domain, also denoted as the DNA binding domain (DBD), and the moderately conserved ligand binding (E-) domain (LBD). The high sequence conservation of the DBD and LBD in particular, allows for phylogenetic classification of NR subfamilies (NR0-NR6) and their constituent subgroups [Nuclear Receptor Nomenclature Committee, 1999; Wu et al., 2007]. NRs regulate gene expression by attaching the DBD to specific response elements in the promoter of target genes, with the whole protein structure functioning as monomer, homodimer or heterodimer [Germain et al., 2006; Mangelsdorf et al., 1995].

One of the most interesting characteristics of NRs is their capability to interact with endogenous or exogenous compounds through ligand binding, a feature which has been described for a subset of NRs, such as for hormonal (steroid and thyroid hormones) and retinoid regulated receptors. In addition, a few NRs do not require any ligand binding and function in a constitutively activated manner [Germain et al., 2006]. Ligand-binding NRs are able to bind to exogenous compounds. When present in high doses or as mimics of natural ligands, these xenobiotics can lead to disruption

of normal NR function [McLachlan, 2001]. Xenobiotics in the environment are commonly introduced by anthropogenic pollution and can affect various developmental and physiological processes in humans and wildlife [Bergman et al., 2013; Gore et al., 2015]. In many fish species exposure to xenoestrogens such as bisphenol A (BPA), 17α-ethinylestradiol and dichlorodiphenyltrichloroethane, have caused developmental malformations and had negative effects on reproduction [Bergman et al., 2013; Canesi & Fabbri, 2015; Segner, 2011]. Many of the effects of environmental estrogens are known to be mediated through interaction with NRs [Canesi & Fabbri, 2015; Delfosse et al., 2015; Lange et al., 2012; Miyagawa et al., 2014]. Tributyltin (TBT), an organotin and an additive in antifouling paints for boats, has been shown to cause imposex, (superimposition of male organs on females) in > 20 gastropod species [Titley-O'Neal et al., 2011] as well as developmental failure and reproduction impairment in bivalves [Inoue et al., 2004; Park et al., 2015; Ruiz et al., 1995a, 1995b]. TBT has been identified as a xenobiotic ligand for vertebrate and gastropod NRs [Grün et al., 2006; le Maire et al., 2009; Urushitani et al., 2011] and a link between NRs and TBT-mediated disruption has been proposed [Grün et al., 2006; Iguchi & Katsu, 2008; Nishikawa et al., 2004]. As the expression of NRs varies between different life stages, the response of an animal to a xenobiotic can vary according to the life stage.

The function and presence of NRs in development, reproduction and homeostasis in vertebrate species is well studied, but knowledge of receptor participation in invertebrate systems has been less well investigated. Previously, we [Vogeler et al., 2014] reported the presence of 43 NRs in the Pacific oyster, *Crassostrea gigas,* and described their phylogenetic relationship to other known NR homologs in human and *Drosophila*. The Pacific oyster is a bivalve species (Clade: Lophotrochozoa, Phylum: Mollusca) and as a sessile filter feeder, it is a commonly used organism for biological monitoring [McCellan-Green, 2013; Suarez-Ulloa et al., 2013; Zhou et al., 2008]. Oysters live along coasts and estuaries worldwide and are under constant anthropogenic pressure including industrial, agricultural and sewage pollution. Although oyster development has been well studied due to the high economic interests in aquaculture for food, the underlying molecular mechanisms of gene regulation during development remain mostly unknown. Here, we provided an overview of the presence of NRs in Pacific oyster life stages, including the early embryo and larval stages. We studied 34 of the 43 NRs genes, for which expression

could be verified by quantitative RT-PCR (qPCR), and assessed the expression patterns for phylogenetically related groups of genes across early life stages, also including a comparison with adults of both sexes. We discuss the results in the context of the three identified developmental stages (early, mid and late development) and the potential participation of NR genes in embryonic development and sexual differentiation. In particular we focused on those receptors which have previously shown a potential for xenobiotic disruption in other species; the retinoid X receptor (RXR), retinoic acid receptor (RAR), thyroid receptor (THR), estrogen receptor (ER), estrogen-related receptor (ERR), peroxisome proliferator-activated receptor (PPAR), ecdysone receptor (EcR), and xenobiotic-sensing receptor subfamily group NR1J.

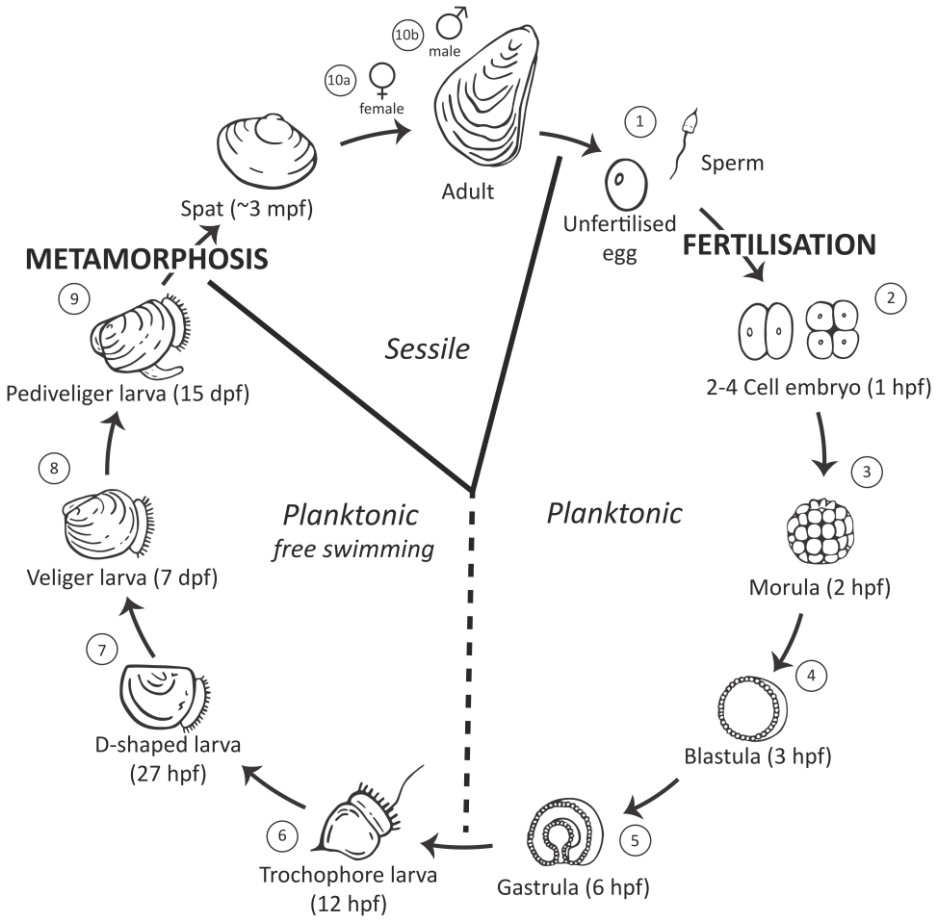



**Figure 3.1: Life cycle of the Pacific oyster, *Crassostrea gigas*.** Numbers 1-10 represent sampling points for nuclear receptor expression analysis. hpf: hours post fertilisation; dpf: days post fertilisation; mpf: month post fertilisation.

## 3.2. Materials and Methods

### 3.2.1. *Oyster husbandry*

Development studies were adapted from the oyster embryo-larval bioassay protocol [Leverett & Thain, 2013]. Three independent fertilisations were performed using four female and five male conditioned adult individuals (Guernsey Sea Farm, Guernsey, UK). Artificial seawater (ASW) was prepared several days prior at 34 psu and pH 8.25. Approximately 3,000–4,000 eggs/ml in a total volume of 1 L were fertilized per biological replicate three hours post fertilization (hpf) each replicate was diluted 1:20 (to ~150-200 embryos/ml) in ASW to prevent oxygen depletion. Life stages were microscopically assessed and samples were taken of unfertilized eggs (~100,000 eggs/sample) and at 1 hpf, 2 hpf, 3 hpf (~100,000 embryos/sample), 6 hpf, 12 hpf and 27 hpf (3000-6000 embryos/sample) (Figure 3.1: 1–7) from the three independent developmental experiments. The experiments were validated by assessing the percentage of dead/abnormally developed D-shelled larvae (ranging from 0-15 % abnormal development), which did not exceed the critical rejecting value of 30 %.

Veliger (7 days post fertilisation (dpf)) and pediveliger (15 dpf) larvae, as well as adult individuals were also obtained from Guernsey Sea Farm. The larvae were re-suspended in 12 °C ASW for one hour and sampled (Figure 3.1: 8–9). Three samples of veliger (1200-2000 larvae/sample) and pediveliger (500-1000 larvae/samples), respectively, were taken for further analysis. Three male and three female adult individuals at the beginning of the gametogenesis were shucked, tissues homogenised and pooled separately by sex (Figure 3.1: 10a–10b). Three homogenized pooled tissue samples (~100 mg) for each sex were taken for further analysis. The sex of the individuals was determined through examination of the gonads to verify whether sperm or eggs were present. Only individuals in the early stages of gametogenesis were used.

### 3.2.2. *RNA extraction and reverse transcription*

The RNA was extracted from three biological replicates at each time point (developmental stages: unfertilized eggs, 1 hpf – 27 hpf; veliger larvae (7 dpf); pediveliger larvae (15 dpf); pooled male and female individuals). Total RNA was

extracted using TRI Reagent RNA Isolation Reagent (Sigma-Aldrich), following the manufacturer's protocol, and genomic DNA was removed with RQ1 RNase-Free DNase (Promega). RNA purity and quantity were determined by ND-1000 spectrophotometer (Nanodrop). For each sample, 900 ng of total RNA, divided into two independent 20 µl reactions (each 450 ng RNA), was converted to cDNA with the ThermoScript RT/PCR System (Invitrogen), using oligo(dT) primers, and then pooled together and further diluted (1:1) with nuclease-free water.

### 3.2.3. *Primer design and optimisation*

Forward and reverse oligonucleotide primers were designed with Primer-Blast at the National Centre for Biotechnology Information (NCBI) [Ye et al., 2012] to amplify each of the 34 *C. gigas* NRs. Primers were 18–23 nt, with a GC content of 40–60 % and produced predicted amplicons of 100–200 bp (Supplementary Table 3.1). The primer pairs were optimized by changing final primer concentration, temperature and/or final $MgCl_2$ concentration to reach a primer pair efficiency between 90-115 %. The efficiency was tested by a dilution series resulting in a standard curve with a slope between 3.0–3.55. The efficiency was calculated as follows [Pfaffl, 2001]:

$$\text{Efficiency (E)} = 10^{(-1/\text{slope})}$$

Each primer pair amplification product was verified by sequencing, using a common polymerase chain reaction (PCR) with the GoTaq system (Promega) for amplification and the products were purified with the QIAquick PCR Purification Kit (Qiagen, UK). Sequencing was conducted by Eurofins MWG Operon (Ebersberg, Germany).

### 3.2.4. *Quantitative RT-PCR*

Quantitative RT PCR was performed using the SsoFast EvaGreen Supermix (Bio-Rad), and the reactions were run on a CFX96 Real-Time PCR Detection System (Bio-Rad). For each gene, each of the biological replicates per time point was run in duplicate (technical replicates) on a single plate and 0.5 µl of cDNA dilution was used per 10 µl reaction. The $MgCl_2$ concentration of the SsoFast EvaGreen supermix and the primer concentration were adjusted for primer optimisation (Supplementary Table 3.1). qPCR conditions were as follows: 95 °C for 2 min, 45 cycles of 95 °C for 15 s,

60–63.1 °C for 30 s (primer dependent) and 72 °C for 1 min. A melt curve was run after each PCR (65–95 °C at a temperature transition rate of 0.05 °C/s). For each reaction the melt curves were analysed to verify the specificity of the amplified product, and to confirm that a single PCR product had been amplified. A non-template control was analysed in parallel for each gene as well as a positive control.

### 3.2.5. *Data analysis*

The calculation of the relative expression (Re) for gene transcripts (mRNA copies) of each NR was based on the modified comparative Ct method [Filby & Tyler, 2005; Pfaffl, 2001], using the average Ct ($avCt_{target}$) of each biological replicate per time point, corrected for efficiency (E) and compared to the normalisation factor of combined reference genes ($NF_{refs}$). Three housekeeping genes (elongation factor-1 α (EF-1), ribosomal protein S18 (RS18), ribosomal protein L7 (RL7)) were chosen as reference genes (Supplementary Figure 3.1). The normalisation factor of the combined reference gene ($NF_{refs}$) has been determined by the programme geNorm v3 [Vandesompele et al., 2002], using the Ct values of the reference genes corrected for their efficiency:

$$\text{Normalisation Factor } (NF_{refn}) = \sqrt[n]{E_{ref1}^{avCt_{ref1}} * E_{ref2}^{avCt_{ref2}} * \ldots * E_{refn}^{avCt_{refn}}}.$$

Thus, the relative expression was calculated as follows [Filby & Tyler, 2005]:

$$\text{Relative expression } (Re) = \frac{NF_{refn}}{E_{target}^{avCt_{target}}}.$$

Relative expressions of all NRs were statistically analysed using RStudio v0.98.1091 (RStudio, Inc.). The data were transformed (log or sqrt) to normal distributions, tested using a Shapiro-Wilk test, and the expression patterns were analysed with a one-way ANOVA follows by multiple pairwise comparisons with Tukey's Honestly Significant Difference Test. Cluster analysis was performed to demarcate the expression patterns during all stages using Cluster 3.0 v1.52 [De Hoon et al., 2004]. Hierarchical and *k*-means clusters were obtained by logarithmically transformed centred data, by gene, using the Euclidean similarity metric. The hierarchical cluster was visualized using Java TreeView [Saldanha, 2004]. In addition, to examining gene expression patterns along NR subfamilies and developmental stages excluding and including

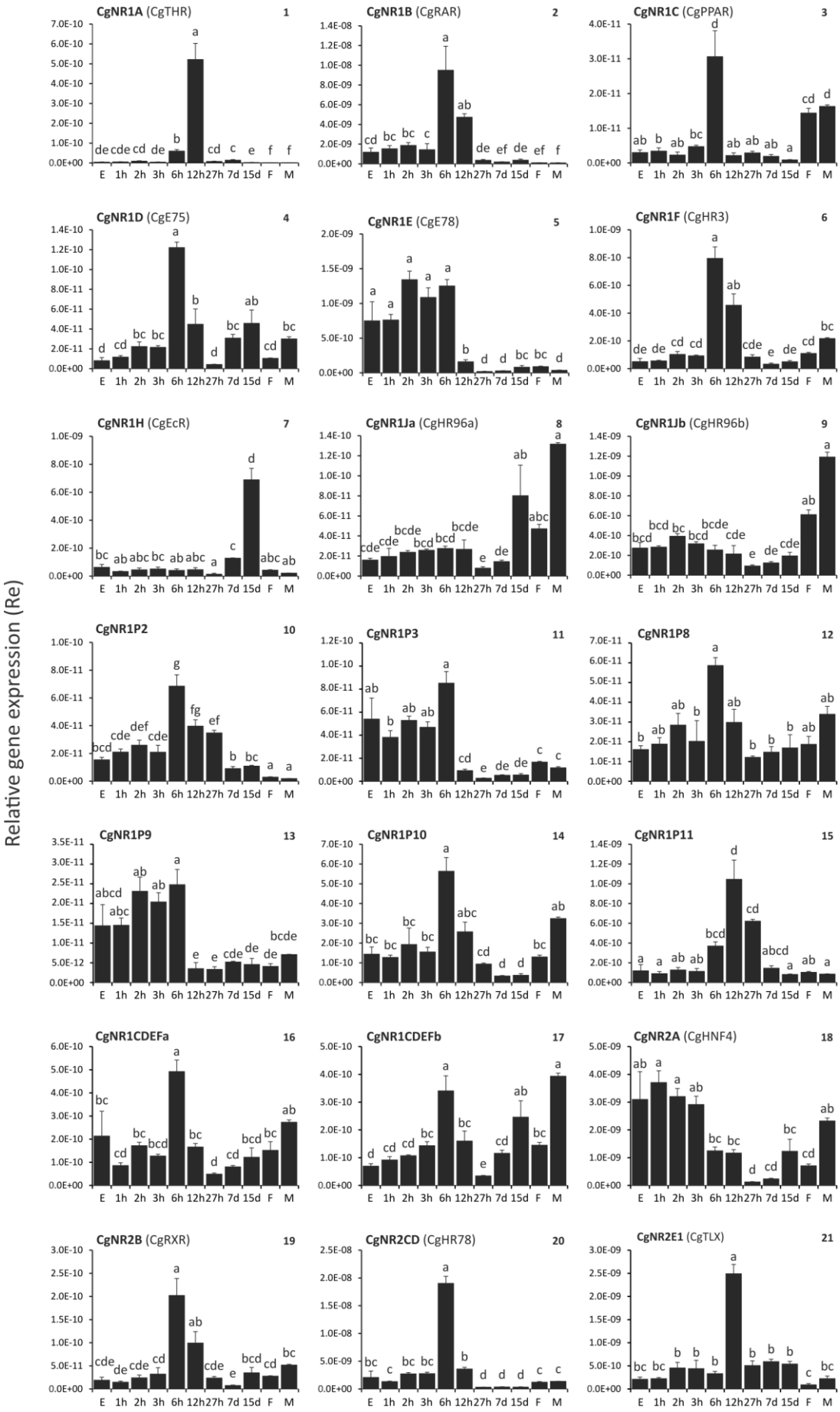
adult stages, principal component analyses (PCA) were conducted in R version 3.2.4 [R Development Core Team, 2008], using prcomp(), a built-in function in the R stats package, and the packages FactorMineR [Vienna et al., 2008] and factoextra [Factoextra, n.d.].

## 3.3. Results

Quantitative RT-PCR was used to measure the expression of 34 NRs in ten different life stages of the Pacific oyster, including nine developmental stages, and male and female adult individuals (Figure 3.1). In general, NR expression was measured during all life points and showed variation between developmental stages (Figure 3.2). A few NR transcripts (CgNR2E2, CgNR2E3 and CgNR2F), were below the limit of detection (N/A) at some life stages. Expression of each receptor was calculated relative to a normalisation factor derived from three reference genes. Elongation factor-1 α (EF-1), ribosomal protein S18 (RS18) and ribosomal protein L7 (RL7) were verified as the most suitable reference genes among other commonly used housekeeping genes (glyceraldehyde-3-phosphate dehydrogenase, glutathione S-transferase, α-tubulin) by the programme geNorm v3 (Supplementary Figure 3.1).

Principal component analysis (PCA) for developmental stages (excluding adult stages) (Figure 3.3) was conducted for 31 of 34 NRs, excluding those receptors which showed an expression below detection limit. Principal component 1 (PC1) accounted for 42 % of the overall variance in gene expression among developmental stages and principal component 2 (PC2) accounted for 25 % of the overall variance. Based on the observed variance in the expression levels of the NRs, four distinct life events are clearly distinguishable (Figure 3.3A): early embryo development (unfertilized egg – 3 hpf), mid development, divided into gastrulation (6 hpf) and trochophore larval stage (12 hpf), and late larval development (27 hfp – 15 dpf).

Similar clustering was detected by separating the NRs based on the observed variance in different life stages (Figure 3.3B). Three clusters of NRs emerged and could be categorised in three of the four previously detected key stages: early, mid and late development. Within the mid development group, the gastrulation (6 hpf) and trochophore larval (12 hpf) stages were not as clearly distinguished as they had been during the previous observation of the variance in the expression levels of NRs. Most of the NRs in the detected early and mid development stages contribute to the observed principle components (Figure 3.3C): early development: CgNR1P3, CgNR1E, CgNHR42, CgNR5B, Cg2DBDγ; CgNR1P9; mid-development: CgNR2CD, CgNHR43, CgNR1B, CgNR1F, CgNR2B, CgNR2E5, CgNR1CDEFa, CgNR1P10, CgNR1P8. Only two NRs (Cg2DBDδ, CgNR0B) display a contribution higher than the expected average contribution for the late development stage. CgNR1H and CgNR1Ja could not be clearly assigned to one of the three key developmental stages


CgNR1A (CgTHR) 1
CgNR1B (CgRAR) 2
CgNR1C (CgPPAR) 3
CgNR1D (CgE75) 4
CgNR1E (CgE78) 5
CgNR1F (CgHR3) 6
CgNR1H (CgEcR) 7
CgNR1Ja (CgHR96a) 8
CgNR1Jb (CgHR96b) 9
CgNR1P2 10
CgNR1P3 11
CgNR1P8 12
CgNR1P9 13
CgNR1P10 14
CgNR1P11 15
CgNR1CDEFa 16
CgNR1CDEFb 17
CgNR2A (CgHNF4) 18
CgNR2B (CgRXR) 19
CgNR2CD (CgHR78) 20
CgNR2E1 (CgTLX) 21
E 1h 2h 3h 6h 12h 27h 7d 15d F M
Relative gene expression (Re)
Life stage

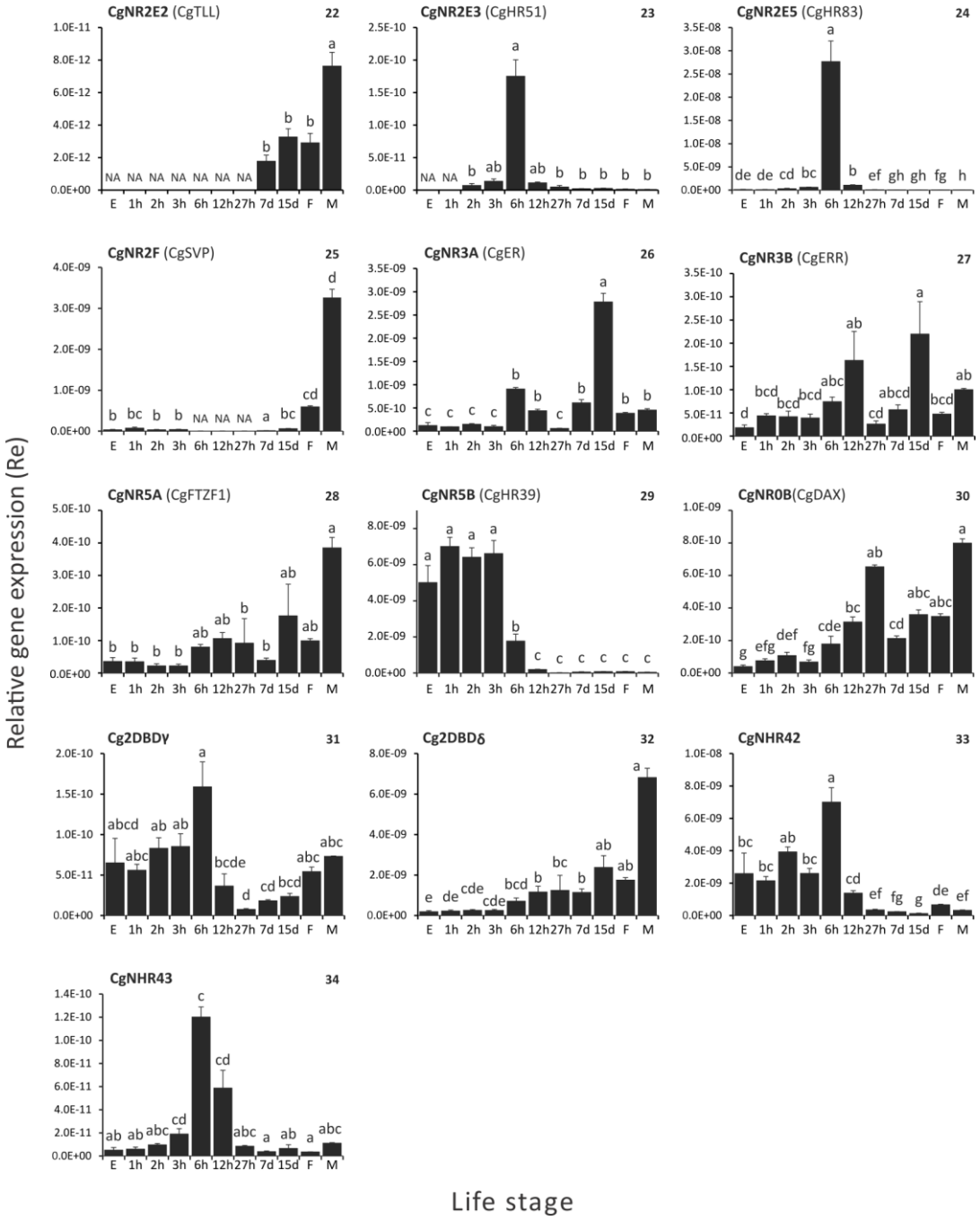



**Figure 3.2: Relative gene expression of 34 *Crassostrea gigas* nuclear receptors in different life stages.** Gene expression was measured using quantitative RT-PCR. Relative gene expression was calculated using a normalisation factor computed with the three reference genes and statistically analysed as described in the methods section. Alternative names for oyster nuclear receptors based on their closest identified homologs in *Homo sapiens* or *Drosophila melanogaster* [Vogeler et al., 2014], are provided in brackets. Bars indicate the mean ± standard error of three independent measurements per time point. Different letters above each bar represent groups that were significantly different (p=0.05); same letters: no significant difference. N/A: expression below detection limit; h: hour post fertilisation; d: days post fertilisation; E: unfertilised eggs; F: female; M: male.

by the PCA (Figure 3.3) and did not contribute towards the observed variance in different developmental life stages (Figure 3.3C). The expression profile of CgNR1H (Figure 3.2) showed an increase at 15 dpf compared to all other developmental and adult stages. CgNR1Ja, showed a comparable expression pattern to its closest oyster receptor paralog CgNR1Jb, which assigned with the NR group expressed during early development. The moderate expressions in early and mid development of both receptors are replaced by low expression during later development and high expression in adult life stages. In addition, CgNR1Ja showed a high expression peak at 15 dpf. Previously we identified a novel subfamily group, NR1P, in the Pacific oyster comprising eleven NRs [Vogeler et al., 2014]. The expressions of six of these receptors (NR1P2, NR1P3, CgNR1P8, CgNR1P9, CgNR1P10, CgNR1P11) were analysed and they all display differences in their expression profiles among different life stages (Figure 3.3B): two receptors (NR1P3 (11), NR1P9 (13)) are mainly expressed at early development; three at mid-development (NR1P2 (10), NR1P8 (12), NR1P10 (14)); and one at late development (NR1P11 (15)).

The sex of adults was determined by visually observing the presence of developed oocytes (eggs) or spermatozoa in the gonads. Gene expression data (Figure 3.2) shows a difference in expression patterns between unfertilized eggs and adult females. Therefore, we assume that the expression measured in female adults, which were at the beginning of the gametogenesis, was not entirely due to the presence of eggs. The same is assumed for males as RNA concentration in sperm is very low [Goodrich et al., 2013].

Individual analyses of adults (Figure 3.2) show equal levels of expression for most males and females, with only five NRs showing significant differences between sexes. An additional PCA was conducted (Supplementary Figure 3.2) to identify general expression patterns between males and females among developmental and adult stages. Males and females vary in their NR expression and each sex also separates from most of the developmental stages (likenesses to pediveliger stage (15 dpf)). The cladogram of a heat map analysis of all NRs (Figure 3.4) shows a similar separation of adults and developmental stages, in particular to early and mid developmental stages. The clustering of the gene expression analysis does not correspond with the phylogenetic clustering of the NR subfamily groups, which also shows divergent temporal expression patterns during development. This becomes particularly apparent in the heat map analysis of all NR expressions (Figure 3.4), in

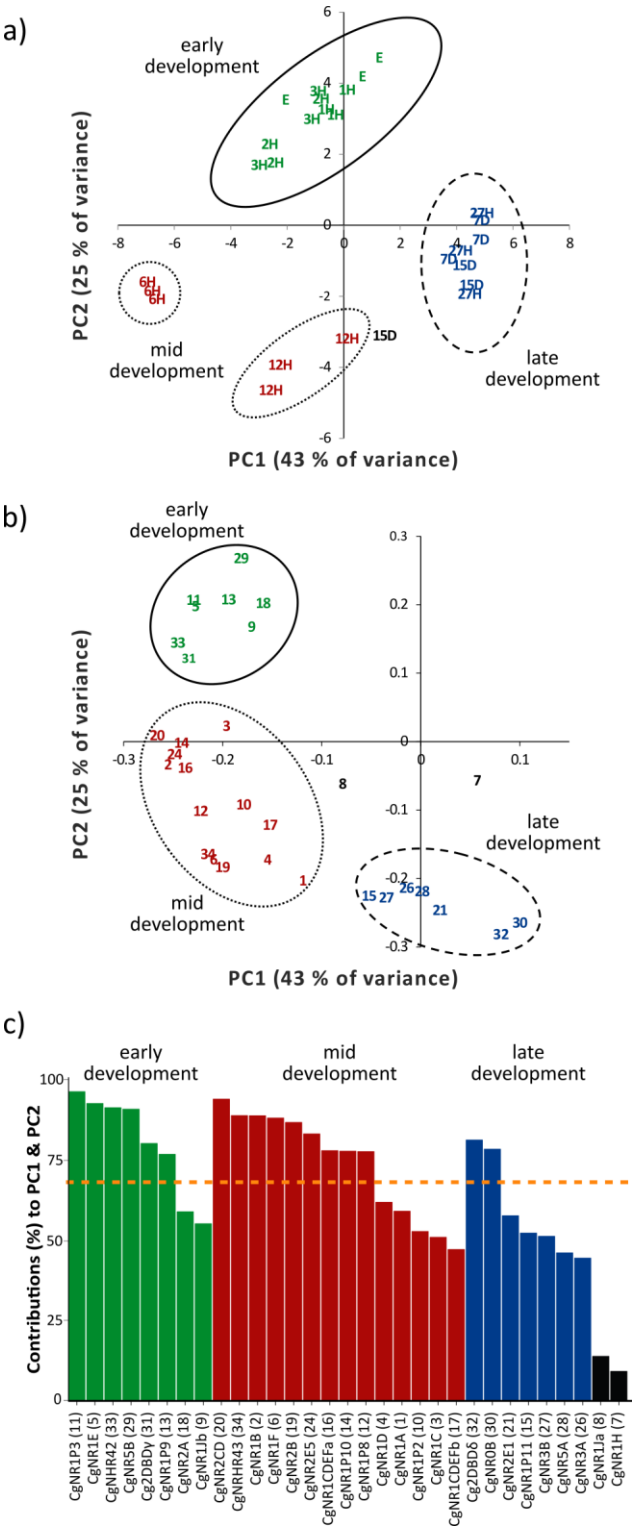



**Figure 3.3: Principal component analysis (PCA) of 31 *Crassostrea gigas* nuclear receptor (NR) gene expression data.** **(a)** Scatterplot of the first two PCA components of developmental stages indicating separation based upon the variance observed in the expression levels of 31 of the 34 NR genes. Principal component 1 (PC1) and 2 (PC2) explain 43 % and 25 % of variance, respectively. **(b)** Scatterplot of PC1 and PC2 scores indicating the separation of 31 NRs based on expression across life stages. Circles around measurements represent distinct clustering for developmental stages or NRs: early development (solid line), mid development (dotted lines), late development (dashed line). Number codes for NRs can be found in (c). **(c)** Bar chart showing the contribution (in percentage) of each NR towards the variability of PC1 and PC2. The orange dashed line: expected average contribution. Letters/numbers/bars: early development (green), mid development (red), late development (blue), receptors not assigned (black). h: hour post fertilisation; d: days post fertilisation; E: unfertilised eggs.

which all developmental stages and both adult life stages have been included. The heat map shows comparable results to the PCA; three main developmental stages, early, mid and late development, plus a further adult life stage were identified with hierarchical clustering of the life stages. Clustering of NRs based on their expression profile does not concur with the phylogenetic relationship of these NRs. The dendogram of the NR hierarchical clustering shows no coherence between

phylogenetically related NRs. For example, during early development (unfertilised eggs – 3hpf), receptors belonging to subfamily NR1 (e.g. 11, 13, 5, 10, 2) cluster together with receptors member of subfamily NR2 (e.g. 18, 23) or non-assigned receptors (e.g. 33) (Figure 3.4).

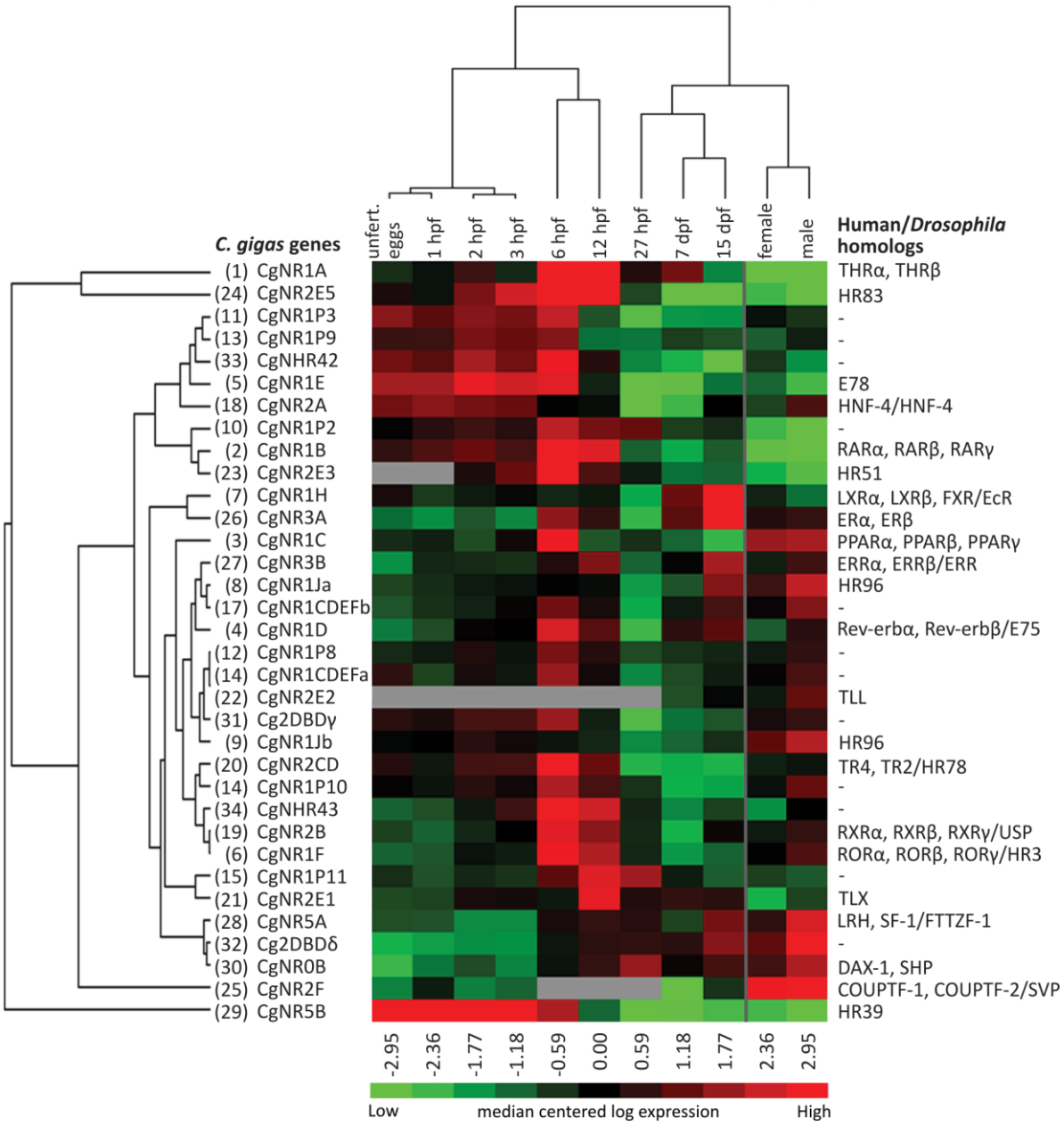



**Figure 3.4: Heat map of all 34 *Crassostrea gigas* nuclear receptor (NR) gene expression data among different life stages**. The mean of the relative expression of the three biological replicates were log transformed and centred. Cladogram of NRs (left) and of life stages (above) indicate groups with similar expression pattern. *C. gigas* NR homologs in human and/or *Drosophila melanogaster* are provided in a table next to it. Numbers in brackets: Number codes for NRs accordant to Figure 3.2. Grey boxes: expression below detection limit. hpf: hours post fertilisation; dpf: days post fertilisation.

## 3.4. Discussion

We show here that the Pacific oyster, like many other complex multicellular metazoans, differentially expresses a variety of NRs during its lifetime. The expression pattern analysis of 34 NRs throughout nine developmental time points clusters into three distinct life stages, showing dynamic changes in receptor expression: an early, a mid and a late developmental stage. Observationally, the mid developmental stage can be further divided into two separate developmental processes: gastrulation and trochophore larval stages, which include organ differentiation and shell development. Male and female adult life stages show an overall separation from the other developmental stages as well as from each other. However, the individual analysis of male and female expression patterns demonstrates only a few cases for which NRs are differentially expressed. The observed clustering of NR expression does not correlate with each receptor's phylogenetic relationship, and belonging to the same subfamily does not result in similar patterns of expression, indicating differential or redundant functions. The members of the novel subfamily group NR1P also show differences in their expression among each life stage, which suggests that these receptors fulfil different functions, irrespective of their close phylogenetic relationship.

### 3.4.1. *Early development: embryogenesis*

Embryogenesis, defines the first few steps in the life of a freshly fertilised egg in metazoans. During this event, RNA gene transcripts are thought to be maternal or zygotic in origin: maternally synthesized RNAs and proteins are stocked in oocytes during female gametogenesis [Davidson, 1986]. A Pacific oyster zygote undergoes its first 2-4 cell division within an hour of fertilisation and is shortly after followed by the morula (2 hpf), blastula (3 hpf) and gastrula (6 hpf) stages (Figure 3.1). The expression of NRs at these early life stages did not significantly differ from the expression in unfertilised eggs for any tested NR. This suggests that the first NRs expressed during early development in oysters are provided by the mother. Nevertheless, maternal RNA is not universally stable and will degrade over time. It is replaced or diluted by zygotic gene expression, a process that is called maternal-to-zygotic transition (MZT) [Davidson, 1986; Tandros & Lipshitz, 2009]. This RNA destabilisation varies spatially and temporally, depending on gene, degradation

mechanism and species, but the event of maternal RNA decay appears to be evolutionarily conserved among metazoans [Bashirullah et al., 1999, 2001]. Embryonic transcriptional activation through zygotic genome activation (ZGA) for a specific gene usually results in an increase in expression during early embryogenesis [Tandros & Lipshitz, 2009]. In this study, the PCA of gene expression for 34 oyster NRs did not distinguish a subgrouping within the early developmental group (Figure 3.3A), which would indicate a collective ZGA onset for the majority of the NRs. The data suggests that the MTZ time schedule is not concordant for most NRs. It is unclear if and which NRs are actually regulated by ZGA in invertebrates. Although research on *Drosophila* suggests that most of the transcription factors are strictly zygotic [De Renzis et al., 2007], other data indicates that some transcription factors are of maternal origin and are required for ZGA onset and maternal RNA destabilisation [Benoit et al., 2009; De Renzis et al., 2007; Liang et al., 2008]. Nevertheless, the resolution of this study for distinguishing between maternal/zygotic RNA is low and additional research including a higher temporal resolution is required to comprehensively identify the origin of early NR expression or proteins.

Compared to their expression level during other life stages, a few NRs showed a high expression level during early embryogenesis: CgNR2A, CgNR5B, CgNR1P9 mostly decrease in expression after 3 hpf, and CgNR1E, CgNR1P3, Cg2DBDɣ, CgNHR42 mostly decrease after 6 hpf. Assumptions of their putative function can only be made based on their closest homologs present in other species (this is a possibility for CgNR2A, CgNR1E and CgNR5B). Early development is shaped by mitotic division, germ cell layer formation and initiation of organogenesis. CgNR2A is a homolog to the hepatocyte nuclear factor 4 (HNF4) [Vogeler et al., 2014], a maternally transferred NR [Holewa et al., 1996, 1997; Zhong et al., 1993]. CgNR2A expression also suggests a maternal origin for this transcript and its high expression indicates an important role during early development. In the frog species *Xenopus laevis* HNF4 is thought to contribute to zygotic activation of a transcription factor regulating tissue specification [Holewa et al., 1998; Nastos et al., 1998; Ryffwl & Lingott, 2000]. Mouse HNF4 participates in regulation processes for primary endoderm development [Duncan et al., 1994] and gastrulation [Chen et al., 1994], and in *Drosophila* it plays a role in gut formation [Zhong et al., 1993]. Knockout experiments in insect species indicated that E78, a CgNR1E ortholog, and HR39, a CgNR5B ortholog, are required

for successful early embryogenesis, beside their main functions in female reproduction [Ables et al., 2015; Xu et al., 2010].

### 3.4.2. *Mid development: gastrulation and shell development*

The mid development stage combines two of the most decisive developmental processes in the bivalve life cycle. Gastrulation (~6 hpf) is defined by three germ layers, basic body structure development and organogenesis. At the trochophore larval stage (~12 hpf), shell development is initiated and larvae become free living, using their circular ciliary bands for locomotion [Gosling, 2004; Kakoi et al., 2008]. Differentiation and development of organs and shell begins, and, in conjunction with axial patterning, requires intensive gene regulation. Many of the oyster NRs are highly expressed during gastrulation and the trochophore larval stage, indicating a participation during these pivotal developmental stages.

CgNR2B, an ortholog to vertebrate and invertebrate retinoid X receptors (RXR), is highly expressed at 6 hpf and moderately expressed at 12 hpf. Research on different molluscan species suggests participation of RXR in organogenesis and shell development. In embryos of fresh-water snails of three different gastropod species (sister-class to bivalves), exposure to 9-*cis* retinoic acid and all-*trans* retinoic acid, natural metabolites of vitamin A, caused significant eye and shell deformation. Some embryos were even developmentally arrested at the trochophore larval stage [Carter et al., 2010, 2015; Creton et al., 1993]. 9-*cis* retinoic acid was identified as a ligand of RXR in a fresh-water and a marine gastropod snail species [Bouton et al., 2005; Urushitani et al., 2011], which in combination with reported exposure effects in other gastropods, indicates a biological function of RXR during optic development and shell formation in snails. Additionally, exposure of developing molluscs, including oyster species, to the organotin TBT caused serious disruption of development including shell deformities, delayed growth and larval development to the point of high death rates even at low TBT concentrations [Salazar & Salazar, 1996; Thain, 1986; Thain & Waldock, 1986]. TBT has been identified as a xenobiotic ligand for gastropod [Urushitani et al., 2011] and deuterostome RXRs [Grün et al., 2006; le Maire et al., 2009].

Heterodimerisation is a common feature of NRs and RXR is the preferred heterodimer partner for many species studied so far [Mangelsdorf & Evans, 1995;

Szantos et al., 2004]. CgNR1B, an ortholog to the retinoic acid receptor RAR, is mainly expressed between 6 hpf and 12 hpf and shows a similar expression pattern as CgNR2B. This provides room for speculation about a possible interaction between these receptors, a hypothesis, which is supported by the findings in two gastropod species, for which gastropod RAR orthologs were observed to heterodimerise with their RXR homologs and are able to regulate gene expression *in vitro* [Gutierrez-Mezariegos et al., 2014; Urushitani et al., 2013]. To what extent the oyster RAR is sensitive to xenobiotic disruption is currently unknown. In contrast to vertebrate RARs, which are highly responsive to natural and synthetic retinoic acids, molluscan RARs seem to have no ability to bind to ligands [Gutierrez-Mezariegos et al., 2014; Urushitani et al., 2013]. More research on oyster RXR and RAR homologs could reveal links between their prominent mid developmental expression, chemical exposure and putative binding, which has been reported for their molluscan relatives. Nevertheless, retinoid metabolism in invertebrates is proposed to be partially conserved among bilaterians [André et al., 2014]. In vertebrates the RAR/RXR heterodimers are involved in the regulation of a diverse variety of genes contributing to organogenesis, axial patterning and neuronal differentiation [Niederrether & Dolle, 2008].

Two additional oyster NRs, which are highly expressed during mid development, should also be mentioned: The first is the peroxisome proliferator-activated receptor (PPAR) homolog CgNR1C. Protostome homologs to PPARs have only been identified in one bivalve species, *C. gigas* [Vogeler et al., 2014], and two gastropod species, *Lottia gigantea* and *Biomphalaria glabrata,* across all of the currently studied protostomes [Kaur et al., 2015], but information on their putative function or mode of action is sparse. The high expression at 6 hpf could potentially point towards a participation of CgNR1C in gastrulation supported by research on vertebrate species. In vertebrates PPARs fulfil different functions during embryo/larval development, especially during gastrulation, ranging from fat metabolism and adipocyte development [Michalik et al., 2002] to cell, neural and muscular differentiation [Michalik et al., 2002; Rotman et al., 2013]. TBT has also been proposed as a xenobiotic modulator of the RXR/PPARγ heterodimer in vertebrates [Grün et al., 2006; Harada et al., 2015; Kanayama et al., 2005; le Maire et al., 2009 68-69].

The second NR is CgNR1A, an ortholog to human thyroid receptors (THR) [Vogeler et al., 2014]. Our data demonstrates high levels of CgNR1A expression during the

trochophore larval stage and lower, but significant expression during gastrulation. The vertebrate THR reacts to stimuli of thyroid hormones and is a common ligand-activated NR. The thyroid, and THRs in particular, play important roles in vertebrates during organogenesis and neural development [Darras et al., 2011] and it is suggested that the thyroid signalling pathway is conserved in invertebrate protostomes [Huang et al., 2015; Wu et al., 2007]. Exposure to BPA causes disruption of developmental, reproduction and physical processes in vertebrates and invertebrates [Canesi & Fibbri, 2015]. The xenoestrogen BPA is also a known antagonist to THR in rats *in vitro* [Moriyama et al., 2002; Zoeller et al., 2005].

### 3.4.3. *Late development: pre-metamorphosis*

The later stages of development in bivalves and gastropods are mainly defined by growth and shell expansion as free swimming planktonic larvae. Some of the fundamental organs and the central nervous system are further defined [Ellis & Kempf, 2011], and new features appear, so called larval organs (velum for swimming and feeding, a foot for crawling, an eyespot as a light sensing organ). The duration of the free swimming stage varies depending on species and environmental cues and terminates in metamorphosis at which larvae transform to their juvenile form (spat). After a substantial re-organisation of body parts and disappearance of the larval organs, the individual attaches to a substrate and becomes sessile [Helm et al., 2004; Gosling, 2004]. Some NRs are expressed just before this life changing event. CgNR1H, an ortholog to the ecdysone receptor (EcR) in ecdysozoans [Vogeler et al., 2014], the sister clade of lophotrochozoans, is highly expressed at 15 dpf. In conjunction with an RXR homolog, EcR initiates and regulates life changing events such as moulting and metamorphosis in ecdysozoans, binding to ecdysteroids [Riddiford et al., 2010]. After receiving an ecdysone signal the EcR/RXR homolog heterodimer initiates a transcriptional cascade of NRs, which are responsible for further gene regulation [Thummel, 2001]. EcR is a common target for pesticides, which results in a disturbance of this cascade, leading to disruption in insect moulting [Fahrbach et al., 2012]. The presence of an EcR ortholog in oyster pediveliger larvae raises the question whether a similar cascade in oyster individuals is activated to initiate metamorphosis and settlement. *In silico* modelling of EcR orthologs in other lophotrochozoans already suggested the possibility for the binding of EcR to

ecdysone or ecdysone-like compounds outside the ecdysozoan lineage [Laguerre & Veenestra, 2010]. Further research on oyster EcR will be needed to investigate its putative participation in metamorphosis and their potential for xenobiotic disruption through pesticides.

The estrogen receptor (ER) and estrogen-related receptor (ERR) homologs CgNR3A and CgNR3B [Matsumoto et al., 2007; Vogeler et al., 2014], both highly expressed at the end of the larvae stages (15 hpf), are also worth mentioning. The ERR, also present in ecdysozoans, is a proposed precursor gene of metamorphosis in insects along with many other functions in animal development [Bardet et al., 2006; Palanker et al., 2006; Tennessen et al., 2011]. The ER is not present in insects, but has been widely studied in fish, regulating brain development and sexual differentiation in larvae [Fröhlicher et al., 2009; Strobl-Mazzulla et al., 2008]. The relatively high expressions of ER and ERR homologs during the mid development stage in oysters, although unresponsive to any estrogen or estrogen-like ligands [Bannister et al., 2007, 2013; Matsumoto et al., 2007], suggest these genes might be fulfilling similar functions, in a constitutive, rather than ligand dependant, manner.

### 3.4.4. *Adults: males and females*

The adult life style of the Pacific oyster differs distinctly from its free-living planktonic developmental stages and requires the regulation of different genes. Being sessile, adult individuals depend completely on their surrounding environment and do not undergo any further fundamental reconstruction of their body plan. In general, our data suggests a separation of NR expression between most of the developmental stages and adult male/female individuals, indicating that different sets of NRs are switched on during the high dynamic developmental and the stationary adult stages.

Our data also indicates a difference in NR expression for male and female adults. The Pacific oyster has separate sexes and has an annual reproductive cycle, which includes the development of oocytes and spermatozoa. As a protandric species, most Pacific oyster individuals first develop as males and may change sex to female after a few annual cycles [Gosling, 2004; Guo et al., 1998]. In the nematode roundworm *Caenorhabditis elegans* [Carmi et al., 1998], as in vertebrates [Norris & Lopez, 2013], sex change and reproduction is directly or indirectly regulated by NRs. The oysters used in this study were at an early stage of gametogenesis for sex

identification. However, the data for individual NR comparison showed only five receptors with significant differences (CgNR1E, CgNR1CDEFb, CgNR2A, CgNR2E2, CgNR2E5), of which most have not been linked to sex dimorphism in other species. Only homologs to CgNR1E in insects have been connected to female reproduction and oogenesis [Ables et al., 2015]. CgNR1E, although expressed highly in early development and less in late and adult life stages, shows a higher expression in females than in males. Expression of the ER and ERR homologs (CgNR3A and CgNR3B respectively) was at similar levels. Previous research in gastropod species showed differences in expression for male and female reproductive tissues and no differences for other tissues [Bannister et al., 2007, 2013; Kajiwara et al., 2006]. Nevertheless, gene expression was measured in whole individuals (not separated by tissue) and sex-dependent expression could still be neutralized by pooling of tissues (e.g. high expression in gonads, but low expression in mantle and vice versa).

CgNR1C, the PPAR ortholog, is highly expressed during gastrulation, but also moderately expressed in adult life stages. TBT, a known gastropod and vertebrate RXR ligand [Grün et al., 2006; Urushitani et al., 2011] causes shell thickening in adult bivalve [Higuera-Ruiz & Elorza, 2011; Waldock & Thain, 1983], as well as the development of male sex organs in females (imposex) in many gastropod species [Titley-O'Neal et al., 2011]. There have been previous hypothesise that this is triggered through disruption of the RXR/PPAR heterodimer, since exposure to rosiglitazone, a strong PPARγ agonist for vertebrates [Nakanishi, 2007], also causes imposex in gastropods [Pascoal et al., 2013]. Additionally, the vertebrate RXR/PPARγ heterodimer has been identified as a target for TBT, which binds to either RXR alone or both receptors [Grün et al., 2006; Harada et al., 2015; Kanayama et al., 2005, le Maire et al., 2009]. Expression of a PPAR ortholog in combination with CgNR2B in adult individuals in the Pacific oyster supports the conceivable theories of interaction between RXR and PPAR, and further adds to the theory that this may be the primary mechanism of TBT-based endocrine disruption in molluscs.

CgNR1Jb, originally assembled with the early development group, displayed its highest expression in the adult life stages (Figure 3.2 & 3.4), but also showed measurable expression during most of the other life stages. CgNR1Jb and CgNR1Ja, which showed a similar expression profile to CgNR1Jb, are members of the protostomes subfamily group NR1J, a homologous group to the deuterostome

subfamily group NR1I. Representatives of the NR1I and NR1J subfamilies have been linked to xenobiotic-sensing [Cruzeiro et al., 2016; Karimullina et al., 2012; King-Jones & Thummel, 2005; Richter & Fiddler, 2014; Timsit et al., 2007], a mechanism of defence against natural and anthropogenic environmental stressors through which expression of genes involved in detoxification is induced. [Prakash et al., 2015; Richter & Fiddler, 2014]. Research in the bivalve *Scorbicularia plana* demonstrated that the bivalve homolog NR1Jβ is also able to interact with such compounds, suggesting a conserved xenobiotic-sensing mechanism in bivalves [Cruzeiro et al., 2016].

## 3.5. Conclusion

This study provides a detailed overview of the NR expression dynamics in the Pacific oyster. We have demonstrated that a large variety of NRs are expressed at different respective stages throughout oyster lifetime, ranging from fertilisation, through embryo and larval development, to the point of adulthood. Different NRs cluster together into groups in a non-phylogenetic manner, representing different life events such as early, mid and late development. Differences between sexes were also recognized.

NRs are known to interact with ligands, which makes them vulnerable to exogenous xenobiotic compounds. Therefore, detecting the expression dynamics in different life stages is important in predicting putative functions of NRs and helping to uncover at which life points the Pacific oyster is vulnerable to xenobiotics. Our study on NRs in a molluscan species is therefore an important step towards understanding invertebrate development and for the study of anthropogenic impacts on the environment.

Chapter four

# Characterisation of Pacific oyster nuclear receptors, retinoid X receptor CgRXR, retinoic acid receptor CgRAR and peroxisome proliferator-activated receptor CgPPAR


## Abstract

Endocrine disruption has been reported globally in many aquatic invertebrate species. Understanding of the mechanism of effect, however, is sparse due to limited knowledge of invertebrate endocrinology and molecular regulatory mechanisms. One proposed mechanism of pollution effects is through the disruption of nuclear receptors, a transcription factor superfamily which regulates gene expression in animals. Environmental pollutants have the ability to interfere with nuclear receptor function by binding to the receptor and inducing incorrect signals. Here, we isolated and characterised four nuclear receptors of the Pacific oyster, *Crassostrea gigas*: two isoforms of the retinoid X receptor, CgRXR-1 and CgRXR-2, a retinoic acid receptor ortholog CgRAR, and a peroxisome proliferator-activated receptor ortholog CgPPAR. The amino acid sequences are highly homologous to those of other molluscan, invertebrate and vertebrate receptors and possess the features required for DNA binding and receptor dimerisation. Computer modelling of the receptors based on 3D crystal structures of human proteins was used to predict each receptor's ability to bind to different ligands. CgRXR showed high potential to bind and be activated by 9-*cis* retinoic acid (9RA) and the organotin tributyltin (TBT). Computer modelling of CgRAR supports the theory of loss of retinoid binding in molluscan RARs, suggesting no binding of 9RA, all-*trans* retinoic acids (ATRA) or synthetic retinoids (TTNPB). Examination of the protein sequence revealed six residues in the ligand binding domain, which prevent the successful interaction with retinoid ligands. Modelling of CgPPAR was less reliable due to high discrepancies in sequence to its human ortholog. Yet, there are suggestions of binding to TBT, but not to rosiglitazone. Early oyster embryos of 2-4 cell stage were exposed for 22-24 h to two different concentrations of ATRA, TBT oxide, rosiglitazone and perfluorooctanoic acid (PFOA), respectively. Embryo toxicity was assessed on percentages of normal and abnormal D-shaped larvae. TBT oxide (0.2 µg/l), ATRA (0.06mg/L) and PFOA

(20 mg/L) showed high effects on development (>74 % abnormal larvae) at their lowest tested concentration, while rosiglitazone (4-40 mg/L) showed no effect on larval development. These findings suggest a nuclear receptor disruption through contaminants, either in a direct effect (TBT) or an indirect effect potentially caused by an isomerisation of chemical compounds (ATRA to 9RA). Our data contributes to the understanding of receptor mechanisms in molluscs and the effects of environmental pollution on aquatic invertebrates.

This chapter is a copy of my research paper: **Vogeler S, Galloway TS, Isupov M, Bean TP. Characterisation of Pacific oyster nuclear receptors, retinoic X receptor, retinoic acid receptor and peroxisome proliferator-activated receptor. Aquatic Toxicology (submitted).** SV, TB and TG contributed to study concept and design as well as to the interpretation of data and manuscript preparation. Receptor cloning, protein analysis, phylogeny, qPCR and statistical analyses were performed by SV. TB and SV conducted the embryo toxicity test and SV and MI were responsible for the conduction and interpretation of the computer modelling data.

## 4.1. Introduction

One specific mechanism through which endocrine disrupting pollutants can impact wildlife is the disruption of gene expression regulation by interfering with the function of nuclear receptors. Nuclear receptors (NR) are ligand binding transcription factors in metazoan species, regulating the transcription of many fundamental genes involved in development, reproduction and homeostasis. These receptors bind to specific response elements in a gene promotor sequence [Germain et al., 2006] and function either as monomers, homodimers, or heterodimers [Mangelsdorf et al., 1995]. Some of these receptors are able to interact with ligands including endogenous or exogenous organic compounds, such as steroids, thyroid hormones and retinoids, which operate either as agonist or antagonist [Germain et al., 2006]. Environmental pollutants can have the same ability to interact with NRs and subsequently induce incorrect signalling. Due to their implicit roles in the endocrine system xenobiotic agonists of NRs are commonly cited as a key mode of action in cases of endocrine disruption [Bergman et al., 2013; Gore et al., 2015].

A classic example for aquatic pollution, which is hypothesised to be linked to NR disruption, is the effect of tributyltin (TBT) on coastal wildlife. This synthetic organotin was introduced as an effective active ingredient of antifouling paints in the 1960s and banned world-wide in 2008 [CD, 2002, IMO, 2001]. The leeching and accumulation of TBT into the environment resulted in severe effects on marine wildlife. Exposure to TBT causes imposex, a superimposition of male genitalia on female gastropods, an effect documented in >20 species [Titley-O'Neal et al., 2011]. In bivalves, such as the Pacific oyster, *Crassostrea gigas*, shell thickening, growth reduction, developmental disruption and a high rate of mortality were observed [Heral et al., 1989; Higuera-Ruiz & Elorza, 2011; Ruiz et al., 1995; Thain, 1986; Tsunemata & Okamura, 2011; Waldock & Thain, 1983]. It is currently assumed that TBT alters the normal function of a specific NR, the retinoid X receptor RXR [Horiguchi et al., 2007; Nishikawa et al., 2004], and reports on TBT as a ligand for RXR orthologs in gastropods and vertebrates [Grün et al., 2006; le Maire et al., 2009; Urushitani et al., 2011] support this theory.

Disruption of RXR receptors by pollutants like TBT could also affect other NRs. The RXR receptor is the predominant heterodimer partner for NRs in various species [Mangelsdorf & Evans, 1995]. Heterodimer constructs can be activated through a ligand by binding to both partners (permissive) or just to the RXR partner

(non-permissive) [reviewed in Dawson & Xia, 2012]. The peroxisome proliferator-activated receptor (PPARs) is a common permissive partner of RXRs in vertebrates [Kliewer et al., 1992]. The RXR/PPAR heterodimer has been proposed as the pathway, through which TBT has its disruptive effect by binding to the RXR receptor and then possibly intensifying the effect by also binding to PPAR [Grün et al., 2006; Harada et al., 2015; Kanayama et al., 2005; le Maire et al., 2009]. In molluscan species PPAR homologs have been found, although not characterised [Kauer et al., 2015; Vogeler et al., 2014].

The RXR receptors belong to the retinoid activated receptors and their natural ligand is 9-*cis* retinoic acid (9RA), a vitamin A derivate [Heyman et al., 1997; Levin et al., 1992]. In vertebrates retinoic acids (RAs) are morphogens involved in pattern formation, cell differentiation and proliferation as well as embryonic development and reproduction [Kam et al., 2012; Niedereiter & Dolle, 2008; Rhinn & Dolle, 2012]. 9RA and other RA derivatives such as all-*trans* RA (ATRA) and 13-*cis* RA are natural ligands for the second retinoic activated receptor type RAR (retinoic acid receptor) [Idres et al., 2002]. RAR is another RXR partner (although non-permissive) [Forman et al., 1995; Leid et al., 1992]. However, RAs are among the most potent known teratogens for animals [Soprano & Soprano, 1995]. When present at an inappropriate titre or time point, RAs affect normal development via binding to retinoid receptors (RXR and RAR) and initialising incorrect signals [Shmarakov, 2015]. Many xenobiotic pollutants like organochlorine pesticides, styrene dimers, monoalkylphenol and parabens have been identified as RAR agonists [Kamata et al., 2008]. Disruptive RAR agonistic activity by unidentified pollutants has also been detected in aquatic environments in North America and Asia [Inoue et al., 2010]. Research on several gastropod species has shown that exposure to natural and synthetic RAs causes eye defects, shell malformation, incorrect neuronal differentiation and abnormalities in sex organ development [Carter et al., 2010, 2015; Creton et al., 1993]. RAR orthologs have previously been characterised in only a few gastropod species [Carter et al., 2015; Gutierrez–Mezariegos et al., 2014; Urushitani et al., 2013], but all of them seem to be unresponsive to RAs [Gutierrez–Mezariegos et al., 2014; Urushitani et al., 2013]. However, research on the RA machinery including RARs in invertebrates is limited and includes only three gastropod species and no other molluscan or protostome representatives.

Molluscs are the second largest group of invertebrates and represent roughly a quarter of all characterised marine species. The presence of NRs has been established in many gastropods and bivalves [Kauer et al., 2015; Vogeler et al., 2014], but the exact functions of most receptors have yet to be discovered. Here we clone and characterise three NRs of the Pacific oyster, *C. gigas*: CgRXR, CgRAR and CgPPAR. All three receptors are phylogenetically assessed and their sequences are structurally analysed for typical features of NRs including the highly conserved DNA binding domain (DBD) and the ligand binding domain (LBD), wherewith some receptors bind ligands. Additional 3D modelling and computational docking are used to predict whether these receptors have the potential to bind to common natural and/or synthetic ligands of vertebrate and gastropod homologous receptors: 9RA, the natural ligand of RXR and RAR [Heyman et al., 1997; Idres et al., 2002; Levin et al., 1992]; ATRA, another natural ligand of RAR [Idres et al., 2002]; TTNPB, a synthetic RAR agonist [Pignatello et al., 1997]; bis(tributyltin) oxide (TBTO), a synthetic ligand of RXR and vertebrate PPAR [Harada et al., 2015; Kanayama, 2005; le Maire et al., 2009; Urushitani et al., 2011]; rosiglitazone, a diabetic drug binding to vertebrate PPARγ receptors [Lehmann et al., 1995]; and perfluooctanoic acid (PFOA), a highly persistent synthetic compound, which interacts with vertebrate PPARα [Vanden Heuvel et al., 2006]. Exposure experiments are conducted using the hypothetical ligands (except 9RA) to test for potential effects on early oyster development.

## 4.2. Material and Methods

### 4.2.1. *Animals and chemical reagents*

Male and female conditioned adult Pacific oyster (*C. gigas*) individuals were obtained from the Guernsey Sea Farm (Guernsey, UK). Individuals for RNA extraction were frozen in liquid nitrogen and stored at -80 °C. Individual broodstock were held in filtered (0.2 µm) and UV treated natural seawater (26.0 ‰; pH 7.8) at 10-15 °C.
Bis(tributyltin) oxide (TBTO), rosiglitazone, perfluorooctanoic acid (PFOA), and all-trans retinoic acid (ATRA) were purchased from Sigma-Aldrich. Stock solution and dilutions were freshly prepared with dimethylsulfoxide (DMSO).

### 4.2.2. *Cloning* C. gigas *nuclear receptors CgRXR, CgRAR and CgPPAR*

Oligonucleotide primers for sequencing the full length of the oyster NR sequences CgRXR, CgRAR and CgPPAR (Supplementary Table 4.1) were designed with Primer-Blast at NCBI [Ye et al., 2012] based on *C. gigas* genomic DNA data [Zhang et al., 2012] for each gene (GenBank: CGI_10004075 (CgRXR); CGI_10028545 (CgRAR); CGI_10011509 (CgPPAR)). Total RNA was extracted from frozen whole adults and mixed embryo oyster individuals (embryo toxicity test). Extraction, DNA digestion, reverse transcription, and amplicon visualization and purification were performed as described previously [Vogeler et al., 2014]. Amplicons were obtained by RT-PCR under the following conditions: 95 °C for 2 min, thirty cycles of 95 °C for 30 s, 60 °C for 30 s, 72 °C for 2 min, and a final extension at 72 °C for 5 min, and cloned into a pGEM-T Easy vector (Promega). Vectors were purified using the PureLink Plasmid miniprep kit (Invitrogen), and were subsequently sequenced by Eurofins MWG Operon (Cologne, Germany). Each identified receptor sequence was confirmed by three independent successful cloning attempts.
The obtained coding DNA sequences (CDS) for all receptors (GenBank accession numbers: CgRXR1: KX590999; CgRXR2: KX591000; CgRAR: KX591001; CgPPAR: KX591002) were aligned to their associated genomic DNA sequence to identify isoforms and their intron/exon structure.

### 4.2.3. *Protein analysis and phylogeny*

The deduced amino acid sequences of each receptor (GenBank accession numbers: CgRXR1: KX590999; CgRXR2: KX591000; CgRAR: KX591001; CgPPAR: KX591002) protein sequences: Supplementary File 4.1), including different isoforms, were annotated by using the Conserved Domain Database at NCBI [Marchler-Bauer et al., 2011]. The sequence identities of each receptor domain were assessed against homologs of other closely related molluscan species (RXR: *Thais clavigera, Nucella lapillus, Chlamys farreri, Biomphalaria glabrata, Lymnaea stagnalis;* RAR: *T. clavigera, N. lapillus, L. stagnalis;* PPAR: *B. glabrata, Lottia gigantea*). Identities to *Homo sapiens* receptors were also calculated.

Phylogeny of oyster receptors was inferred using the Maximum Likelihood and Bayesian Inference methods as previously described [Vogeler et al., 2014]. The DBD and selected parts of the LBD of each receptor were combined and then aligned with NR homologs of species across phyla (Supplementary File 4.2) using default parameters in MUSCLE v3.8.31 [Edgar, 2004]. The Maximum Likelihood phylogenetic tree was constructed using PhyML v3.0 [Guidon et al., 2010] with an LG matrix plus optimized invariable sites (+I) and gamma distributed rate heterogeneity among sites (+G) and 1000 bootstrap replicates. The Bayesian Inference tree was calculated using MrBayes v3.2.2 [Huelsenbeck & Ronquist, 2001] with the JTT+I+G model. Four randomly started simultaneous Markov chains were running for 5 million generations with chains sampled every 100 generations and a burin of 5000 trees. The phylogenetic trees were visualized and illustrated with FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

### 4.2.4. *Three-dimensional (3D) modelling of CgRXR, CgRAR and CgPPAR LBD*

Crystal structures of human RXR, RAR or PPAR LBD in complex with either 9RA (1FBY [Egea et al., 2000], 3LBD [Klaholz et al., 1998]), ATRA (2LBD [Klaholz et al., 1998]), TTNPB (1XAP [Germain et al., 2004]), TBT (3E94 [le Maire et al., 2009]) or rosiglitazone (4EMA [Liberato et al., 2012]) were obtained from the RCSB Protein Data Bank (PDB) [Berman et al., 2000]. Ligand dictionaries for docking were also obtained from the PDB site including the molecular structure of estradiol for negative control. *C. gigas, T. clavigera* and *N. lapillus* models for RXR and RAR and oyster models for PPAR were constructed by the modelling server SWISS-MODEL [Biasini

et at., 2014] using the crystal structures of the human NR LBDs with the investigated ligand bound as templates. Computational docking of ligands to human and generated mollusc NR models was conducted using AutoDock Vina [Trott & Olson, 2010] with AutoDockTools and Phython Molecular Viewer PMV graphical interface [Sanner, 1999]. The mean ligand binding energy was estimated from three independent computational docking calculations with the assumption that more negative values equates stronger ligand binding. The tin atom was replaced with zinc for the docking calculation of TBT, since atom parameters for tin are not included in the AutoDock Vina dictionary [Trott & Olson, 2010]. Zinc was chosen as the best tin mimic among other metals available in the dictionary due to its similar tetragonal coordination and related preference to sulfhydryl side chains of cysteine residues. The programmes Coot [Emsley et at., 2010] and CCP4 Molecular Graphics [McNicholas et at., 2011] were used for assessment and visualizing of ligand docking results.

#### 4.2.5. *Embryo toxicity test*

The embryo toxicity assays were executed following the standardised ICES oyster embryo bioassay (OEB) protocol [Leverett & Thain, 2013]. Fertilisation was carried out in natural seawater using eggs from two females and a sperm mix of three male conditioned adult oyster individuals. Three hours post fertilisation (hpf) approximately 200 embryos/ml in a total volume of 250 ml were exposed to chemicals with two final concentrations (low & high): TBT low: 0.2 µg/L; TBT high: 2 µg/L; Rosi low: 4 µg/L; Rosi high: 40 µg/L; ATRA low: 0.06 mg/L; ATRA high: 0.6 mg/L; PFOA low: 20 mg/L; PFOA high: 50 mg/L. Additionally, water and DMSO (0.4 %) controls were prepared and three replicates per treatment were performed. The three replicates are determined as pseudo-replicates, since the experiment took place on the same day and the same parental individuals, chemicals stocks and water source were used. As recommended by the protocol, a gradient of zinc concentration was used as the reference toxicant. The assay was terminated 24 hours post fertilisation (hpf). Approximately 40 ml from each treatment at trochophore (12 hpf) and D-shaped larval stage (24 hpf) by retaining them on a 20 µm filter, followed by snap-freezing in liquid nitrogen. D-shaped larvae samples were also preserved in a buffered formaldehyde solution (final concentration/sample 0.4 % formaldehyde) for later

observation. Per replicate, 100-140 oyster individuals were microscopically assessed based on their larval appearance and the numbers of perfect and abnormal D-shaped larvae counted. Differential interference contrast microscopic pictures of exposed larvae were taken using a Nikon Eclipse E800 microscope and the Nikon element BR image analysis software. Abnormal larvae were further categorized: extruding velum, protruding soft tissue, partly developed shell, arrested growth/ shell not developed. Mean percentages including the standard error (±SE) for each larval category of each pseudo-replicate set were calculated. Additional D-shaped larval samples from the water column of each replicate were taken, concentrated, and preserved. Mean numbers of larvae/ml including the standard error (±SE) were calculated for each water column sample of each pseudo-replicate set. Throughout the assay embryos of the water controls were checked for normal development and samples of key developmental stages were taken for RNA extractions.

#### 4.2.6. *Gene expression*

Gene expression of CgRXR, CgRAR and CgPPAR (not differentiated between isoforms) of control (water & DMSO) and treated oyster embryos at trochophore and D-shaped larval stage were assessed by using quantitative RT-PCR (qPCR) as previously described [Vogeler et al., accepted]. No RNA could be extracted for ATRA high and was therefore excluded from this analysis.

## 4.3. Results

### 4.3.1. *Isolation, phylogeny and ligand binding modelling of* C. gigas *nuclear receptors CgRXR, CgRAR and CgPPAR*

The genomic DNA data of each receptor was used as template for sequencing the full CDS (coding sequence) of three oyster NRs and their isoforms. Alignment of CDS to the genomic DNA sequence showed clear exon/intro structure for each NR (Figure 4.1). Each CDS encodes for a protein, which is includes the expected domain structure of NRs: a highly variable A/B domain (N-terminal), a DBD, a flexible "hinge" region in between, a LBD, and for most receptors a final highly variable F-domain (C-terminal). Furthermore, two conserved zinc finger motifs of the DBD and additional features, which show high homology to those used for dimerization, cofactor recognition and ligand binding, were also identified (Supplementary Figure 4.1). Phylogenetic analyses (maximum likelihood and Bayesian Inference) of DBD

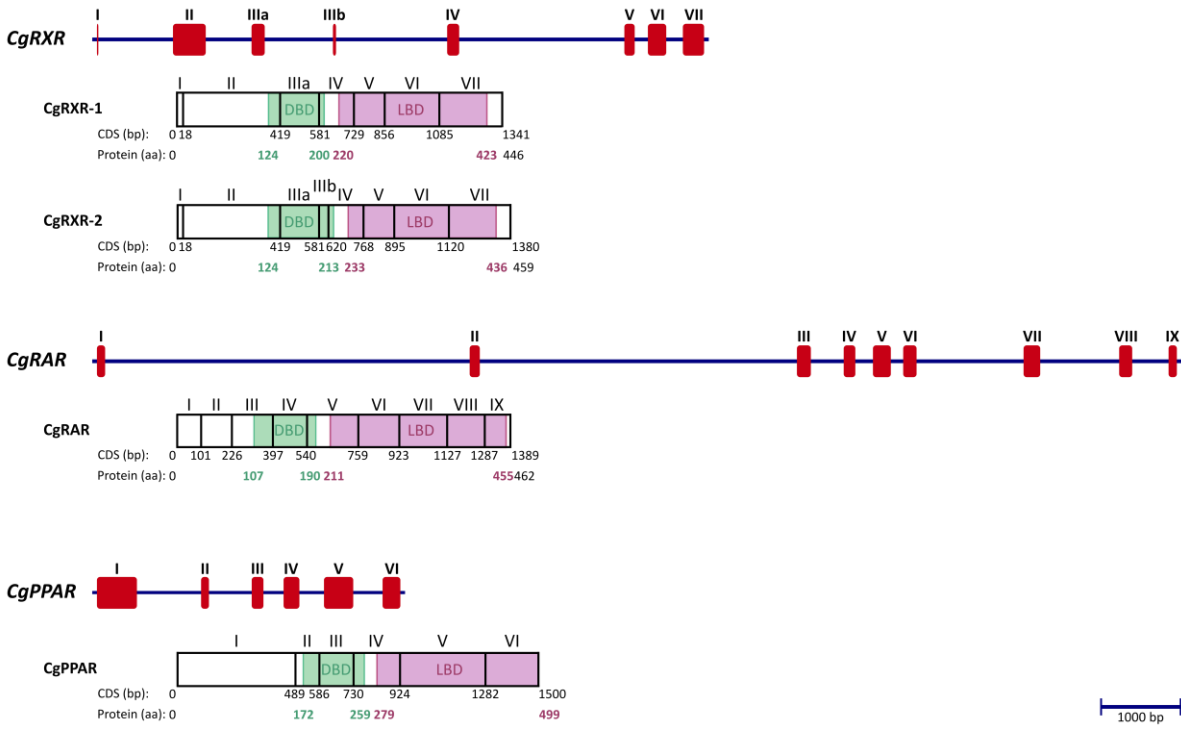



**Figure 4.1: Exon/intron structure, coding sequence and protein organization of CgRXR-1, CgRXR-2, CgRAR and CgPPAR.** Blue line: genomic sequence (bp); red rectangles: exons forming CDS (bp); roman nummerals: number of exon; Arabic numbers: position and length for either CDS (bp) or protein (aa). Green boxes/numbers: DBD position in protein; purple boxes/numbers: LBD position in protein.

plus LBD amino acid alignments were conducted to confirm homology for each NR (Figure 4.2). Human estrogen receptors alpha (HsERα) and beta (HsERβ) were used as outgroup of the combined tree of both analyses. The 3D models for CgRXR and CgRAR were successfully created based on crystal structures of human homolog NRs bound to specific ligands. Computational docking calculated the binding affinity for retinoids (9RA, ATRA, TTNPB) and TBT to CgRXR and CgRAR models (Table 4.1, Supplementary Table 4.2). The same ligands, docked to human source models, provide reference affinity values for successful binding, possible induction of

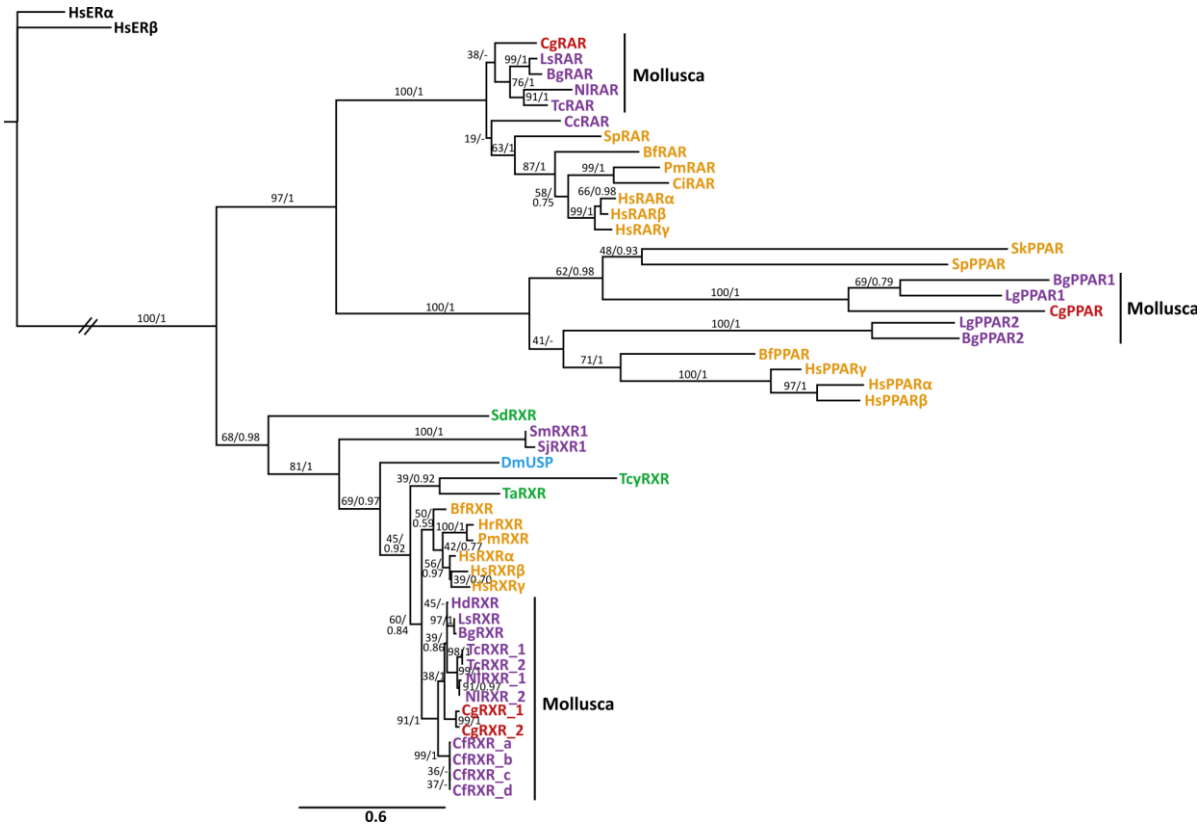



**Figure 4.2: Phylogenetic tree of nuclear receptors homologs RXR, RAR and PPAR among various phyla.** The alignment was constructed using the DBD plus portion of LBD and phylogenetic relationship was conducted by a Maximum Likelihood and Bayesian Inference. Maximum Likelihood bootstrap support values (percentage of 1000 BS) and Bayesian posterior probabilities are provided above the nodes separated by slash. Human estrogen receptors alpha (HsERa) and beta (HsERb) were used as outgroup. Bf: *Branchiostoma florida*; Bg: *Biomphalaria glabrata*; Cc: *Ciona intestinalis*; Cf: *Chlamys farreri*; Cg: *Crassostrea gigas;* Dm: *Drosophila melanogaster*; Hd: *Haliotis diversicolor*; Hr: *Halocynthia roretzi*; Hs: *Homo sapiens*; Lg: *Lottia gigantea*; Ls: *Lymnaea stagnalis*; Nl: *Nucella lapillus*; Pm: *Polyandrocarpa misakiensis*; Sd: *Suberites domuncula*; Sj: *Schistosoma japonicum*; Sk: *Saccoglossus kowalesvski*; Sm: *Schistosoma mansoni*; Sp: *Strongylocentrotus purpuratus*; Ta: *Trichoplax adhearensis*; Tc: *Thais clavigera*; Tcy: *Tripedalia cystophora.* Red: *C. gigas* receptors; Orange: Deuterostomia; Purple: Lophotrochozoa; Blue: Ecdysozoa; Green: Pre-Bilateria

conformational change and signal transmission. Models of the 3D structure of CgPPAR LBD were considered unreliable as its sequence identity to human homologs is low (<24%) (Supplementary Table 4.3), but docking of TBT and rosiglitazone to these models was still conducted for comprehensiveness. Binding probabilities of CgPPAR to PFOA could not been tested as no human HsPPARα template bound to PFOA were available. Binding affinity values between positive and negative controls show small differences. Docking of 9RA to the positive control model of HsRXRα (pdb ID: 1FBY) resulted in binding affinities of -10.6 kcal/mol (Supplementary Table 4.2). The negative controls estradiol and ATRA, which do not induce agonistic signals in HsRXRα, were assigned less negative binding affinities (-9.5 & -9.1 kcal/mol, respectively), but displayed ligand positioning unlikely to induce conformational changes when docked to the human model. These comparably small differences between the docking energies of the negative and positive control are a drawback of the used docking method. The docking of the expected natural ligand was usually accompanied by hydrogen bond/salt bridge formation and by a better fit of the hydrophobic part of the ligand into the receptor cavity. This results in lower energy values in comparison to the negative controls. The high binding energy values for the negative controls, on the other hand, are probably due to the binding of the bulky hydrophobic ligands into the corresponding LBD pockets, including a burial of the significant hydrophobic surface area.

#### 4.3.1.1. *Retinoid X receptor CgRXR*

Two isoforms for the oyster RXR homolog were identified and named CgRXR-1 and CgRXR-2, encoding a 446 amino acid (aa) and a 459 aa protein, respectively (GenBank accession numbers: CgRXR1: KX590999; CgRXR2: KX591000) (Figure 4.1). The difference between CgRXR-1 and CgRXR-2 is a 13 aa long insertion/deletion in the T-box of the DBD. Sequence alignment to known RXR homologs in other mollusc species shows that the T-box, a conserved section of the DBD required for dimerization, is likely to be common for molluscan RXR isoforms (Supplementary Figure 4.1). However, the CgRXR-2 sequence is unique and does not show any similarities to other molluscan isoforms, suggesting it is an insertion. Sequence identities for the conserved regions to other molluscan RXRs range from 90-97 % for the DBD and 90-93 % for the LBD (Supplementary Table 4.3). Identical

P-box sequences (first zinc finger; ‘CEGCKG’), a DNA recognition motif, were identified for human and molluscan RXRs. The D-box (second zinc finger; ‘RDDRN’), responsible for dimerization, shows 100 % identity to the bivalve *C. farreri* and one amino acid difference to gastropod RXR’s D-boxes (‘R’). Conserved domain searches identified similarities to features known for DNA binding, ligand binding and co-activator recognition sites. A homodimer interface, recognised as a requirement for RXR homodimers on response elements DR1 and DR2 (two half site motifs as direct repeats (DR) separated with a short spacer of one or two nucleotides), was also identified, as well as equivalent heterodimer interfaces for RAR, PPAR and ecdysone (ECR) heterodimer formations. The phylogenetic analysis places both CgRXR isoforms inside the group of other known molluscan RXRs (Figure 4.2). The high sequence identity of the LBD of CgRXR-1 to human RXRs allowed reliable CgRXR 3D models to be built. Computational docking showed similar binding affinities

(-10.6 kcal/mol) of 9RA to CgRXR, using HsRXRα model as a positive control, and to the two snail receptors NlRXR and TcRXR (Supplementary Table 4.2), which are both known to respond to 9RA [Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2011]. The oyster RXR ligand docking position was comparable to human RXRα (Supplementary Figure 4.2). Binding affinities to TBT were also consistent between the CgRXR model, the HsRXRα control model (pbd ID: 3E94) and TBT responding TcRXR [Urushitani et al., 2011] models (-5.4 - -5.6 kcal/mol) (Supplementary Table 4.2). According to these results, TBT does not bind tightly to the receptor. However, the modelling always positions the tin atom in the vicinity of the side chain of the conserved cysteine C415 in the H11 helix.

#### 4.3.1.2. *Retinoic acid receptor CgRAR*

Nine exons form the 1389 base pair (bp) long CDS of CgRAR, which encodes a 462 aa long protein (GenBank accession number: CgRAR: KX591001) (Figure 4.1). Additional RAR isoforms could not be confirmed by three independent cloning attempts, but sequencing of one, possibly rare isoform, suggests presence of a CgRAR isoform showing two supplementary amino acids in the T-box of the DBD (Supplementary Figure 4.1). Sequence identity to three gastropod RARs ranged from

**Table 4.1: Calculated binding affinity values (kcal/mol) by computational docking of retinoid ligands to the created 3D models of human, oyster and gastropod retinoic acid receptors (RARs).** (VGSMNL): single-letter code of residues changed in CgRAR ligand binding domain; pdb template: pdb ID providing the template for 3D modelling; RA: retinoic acid; ATRA: all-*trans* RA; Cg: *Crassostrea gigas*, Hs: *Homo sapiens*; Nl: *Nucella lapillus*; Tc: *Thais clavigera*

| Ligand | Receptors | | | | | | pdb template |
|---|---|---|---|---|---|---|---|
| | HsRARγ HsRARγ HsRARβ | CgRAR | TcRAR | NlRAR | CgRAR (VSGMNL) | NlRAR (mutated) | |
| 9-*cis* RA | -11.7 | -10.2 | -9.8 | -10.3 | -11.8 | -11.6 | 3LBD |
| ATRA | -11.8 | -10.5 | -9.7 | -9.8 | -11.4 | -11.0 | 2LBD |
| TTNPB | -14.9 | -12.6 | -11.6 | -12.2 | -14.7 | -14.7 | 1XAP |

90-95 % for the DBD and 58-60 % for the LBD, respectively (Supplementary Table 4.3). The P-box ('CEGCKG') is identical to its molluscan and human homologs. The D-box ('HKDKN') shows differences to human ('HRDKN'), *L. stagnalis* ('HKEKN') and *N. lapillus* ('HKDQT'), but it is identical to the D-box of *T. clavigera*. DNA, ligand and co-regulatory recognition sites as well as the heterodimer interface site could also be recognized. CgRAR groups together with related molluscan RARs (Figure 4.3). Computational docking suggests that CgRAR is unlikely to be activated by RAs. Neither 9RA, ATRA nor TTNPB occupy the correct position to cause required induced conformational changes in the receptor for signal transduction. Modelled CgRAR binding affinity energies are similar to TcRAR and NlRAR (Table 4.1), which were both shown not to respond to RAs *in vitro* [Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2013]. Human RARs, on the other hand, are able to respond to RAs and accordingly display better binding affinities to different types of RAs. When analysing the ligand binding of ATRA to CgRAR, six residues could be identified, which prohibit the binding required for the 'induced fit' conformational changes (Figure 4.3). Three residues (S271, M308, N326) sterically prohibit the correct ligand positioning, including its carboxyl group, which in turn results in a weakening of the salt bridge to arginine (A315). Three additional residues (V267, G274, L445) are less bulky than their equivalents in human RARs. They match the surface of the ligand unfavourably and cannot provide the required induced fit conformational change in the receptor domain. When these six residues in CgRAR were mutated into the equivalent HsRARs residues, the ligand binding affinity of ATRA, 9RA and TTNPB to CgRAR reverts to the human RA affinities (Table 4.1). Ligand position and induced fit

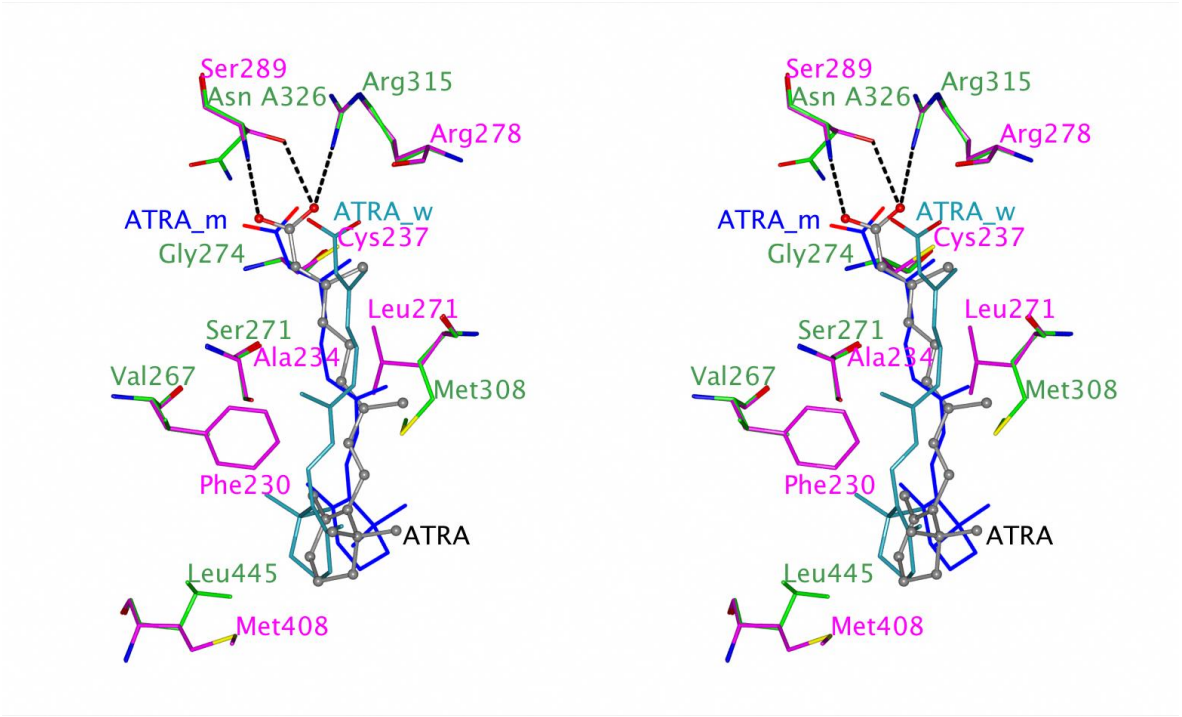



**Figure 4.3: Stereo view of ATRA bound to the ligand binding pocket of a CgRAR model.** Superimposition of model CgRAR (purple) on the crystal structure of HsRARγ LBD (green) bound to human RAR agonist ATRA. Original ATRA (grey) bound to HsRARγ LBD template (pdb ID: 2LBD); ATRA (light blue) to CgRAR wildtype, ATRA (dark blue) bound to CgRAR mutated. Divergent residues as well as arginines binding to the COOH group of ATRA including hydrogen bonds are indicated.

become more advantageous for an induction of conformational changes, suggesting a recovery of ligand response (Figure 4.3). Similar results have been reported for *N. lapillus* RAR. Seven residues in the ligand binding pocket of NlRAR, of which five are shared with CgRAR, have previously been shown to prevent the receptor from responding to retinoic acid [Gutierrez-Mazariegos et al., 2014]. All six mutations result in a ligand binding affinity increase similar to the HsRAR level (Table 4.1).

#### 4.3.1.3. *Peroxisome proliferator-activated receptor CgPPAR*

CgPPAR, the longest of the three investigated receptors, is a six exon, 1500 bp long sequence, encoding for a 499 aa long protein (GenBank accession number: CgPPAR: KX591002) (Figure 4.1). No evidence for any additional CgPPAR or transcript isoform has been found. Sequence identity shows medium identities (75-78 %) for the DBD to PPARs identified in other molluscan species, but a low sequence identity for the LBD (29-38 %) (Supplementary Table 4.3). The LBD is also

shorter than the human LBD (approx. 50 aa) and sequence alignment indicates an absence of the helixes H2, H2' and H12 (Supplementary Figure 4.1). Although the P-box ('CEGCK') is identical, the D-box ('ENPKG') does not show any similarities to other molluscan or human PPARs. CgPPAR groups together with homologous PPARs, and gastropods BgPPAR1 and LgPPAR1 are the closest identified relatives to CgPPAR (Figure 4.2). The reliability of 3D structure models of CgPPAR are limited by the low sequence homology to the only physically characterised PPAR receptor: HsPPARγ, and hence computational docking results for this receptor are less trustworthy. Rosiglitazone is predicted to dock to the CgPPAR model with similar binding affinity values (~-8.2 kcal/mol) to HsPPARγ (pdb ID: 4EMA) (Supplementary Table 4.2), but its positioning does not suggest an induction of a conformational change. TBT docks loosely to the hydrophobic pockets of CgPPAR and HsPPARγ (pdb ID: 3WJ4) (-5.1 kcal/mol) and positions its tin atom towards a cysteine in the H3 helix in both receptor models.

### 4.3.2. *Embryo toxicity tests and resulting gene expression*

The effects of TBTO, rosiglitazone, PFOA and ATRA on oyster development were tested by oyster embryo bioassay. Larval appearance and developmental status were microscopically assessed after completion of the assay (24 hpf) (Figure 4.4). Five categories of larval appearance could be identified. Development of a perfect D-shaped larva (Figure 4.4a,g) indicates a normal development of an oyster larva at around 24 hpf. The abnormal group includes four categories: (1) extruding velum, exposing structures of the velum such as the cilia (Figure 4.4b,h); (2) protruding soft tissue, showing enlarged soft tissue at one side of the normal sized D-shell (Figure 4.4c,i); (3) shell partly developed, development of a much smaller D-shaped shell with much the animal's body exposed (Figure 4.4d); (4) arrested/ shell not developed, including individuals trapped in the trochophore larvae stage (Figure 4.4fe,k) or atrophied larval animals without a shell (Figure 4.4f). Arrested and no-shell development could not be clearly distinguished under a standard light microscope, and were therefore grouped together. This last category is rated as the most severe as it does not allow progression to D-shaped larvae. Differential interference contrast microscopy showed that in the latter group of individuals had, on occasion, developed shell-endowment/disposition (Figure 4.4k, orange-red structure), but could

not continue to develop a full D-shaped shell. The reference treatment with zinc showed an increasing negative effect on the development perfectly D-shaped oyster embryos with increasing zinc concentration (Supplementary Figure 4.3).

The untreated water control shows a high percentage of perfectly developed D-shaped larvae (92 %) (Figure 4.4l). The solvent control (DMSO) did not have any visible effects on the oyster development, and neither did the chemical rosiglitazone (perfect D-shaped larvae: 94-96 %). TBTO, ATRA and PFOA, on the other hand, have severe effects on oyster development at both, low and high, concentrations. TBTO, the chemical with the lowest concentration (0.2 µg/L), displays effects on the oyster development resulting in only 26 % normal perfectly developed D-shaped larvae. A tenfold increase in TBTO concentration (2 µg/L) leads to more severe effects with almost 40 % of embryos not even reaching the D-shaped larval stage. This is comparable to previously reported data of oyster embryos exposed to TBT for 24 h with an $LC_{50}$ of 3.9 µg/L [Tsunemasa & Okamura, 2011] and an $EC_{50}$ of 1.7 µg/L [IPCS, 1990]. The effects of TBTO also intensified at lower embryo concentration (data not shown). At a ten times lower initial embryo concentration (20 embryos/ml) at low TBTO concentration 18 % normal development occurred and at high concentration all embryos are arrested or the shell could not been developed.

Low dose of ATRA (0.06 mg/L) appears to disturb organ and tissue development as well as shell formation. Most of the individuals exposed to a low dose of ATRA show normal sized shell development, but have extruding velum or protruding soft tissue. Exposure of the pond snail *L. stagnalis* to range of ATRA doses ($10^{-7}$-$10^{-5}$ M) displayed comparable levels of organ (eye) and shell deformations as well as arrested trochophores during early development [Carter et al., 2015; Creton et al., 1993]. At the higher concentration of ATRA (0.6 mg/L) none of the oyster individuals reached the D-shaped larval stage. Microscopic examination of the high ATRA samples showed no surviving individuals, not even as arrested trochophore individuals. PFOA has a similar impact on development, albeit at much higher concentrations. The lower concentration (20 mg/L) leads to extruding velum, protruding soft tissues and enlargments. The higher concentration (50 mg/L) prevents embryos from developing a full size D-shaped shell or reaching the D-shaped larval stage. This concurs with acute toxicity values in microalgae, marine invertebrates and fish ($EC_{50}$: 12-164 mg/L) [Mhadhbi et al., 2012].

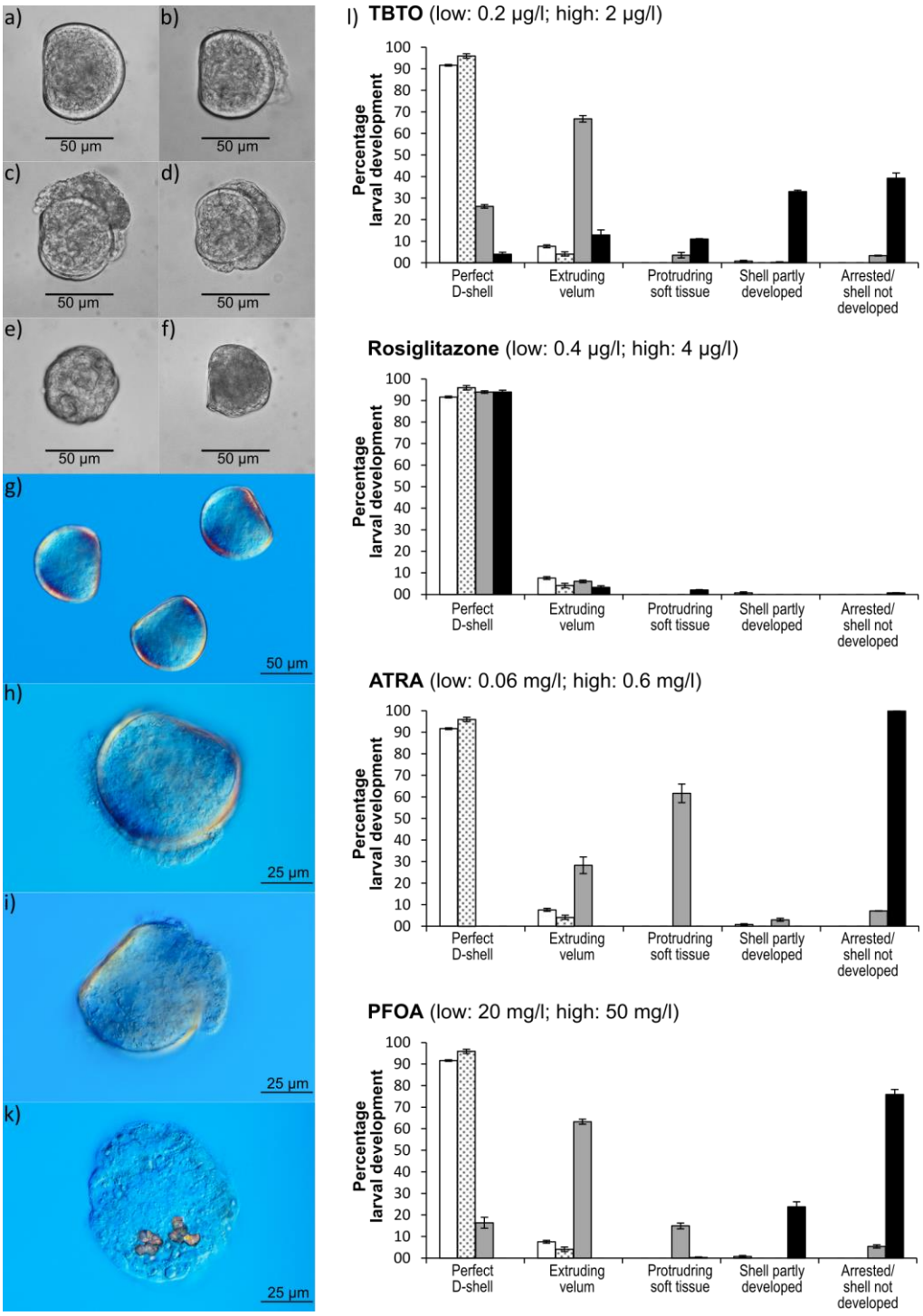



**Figure 4.4: Oyster embryo development after 21 h of exposure to TBTO, rosiglitazone (Rosi), all-*trans* retinoic acid (ATRA) and perfluorooctanoic acid (PFOA). a – k)** Example of oyster development under a light (grey) and differential interference contrast (blue) microscope belonging to the five categories: perfectly developed D-shaped larvae (**a, g**), extruding velum (**b, h**), protruding soft tissue, (**c, i**), shell partly developed (**d**), arrested shell/shell not developed (**e-f, k**). **l)** Percentage of perfectly developed D-shaped larvae and abnormal developed larvae described in four abnormal development categories. Error bars show standard error of three pseudo-replicates per treatment. White bars: water control; spotted bars: DMSO control (0.4%); grey bars: low concentration of treatment; black bars: high concentration of treatment.

The primary measures of effect from the oyster embryo bioassay are measured on the animals that remain in the water column (as free swimming or floating larvae) (Table 4.2), but in fact the effects of the chemicals can also been seen in the actual numbers of larvae that remain in the water column. While approximately half of the controls and rosiglitazone larvae were swimming in the water column, TBTO, ATRA and PFOA (only high conc.) showed a decrease in floating/swimming larvae.

**Table 4.2: Free swimming larval individuals (larvae/ml) in the water column after 21 h exposure to TBTO, rosiglitazone, ATRA, PFOA.** Negative controls: water or DMSO (0.4%) controls. Mean ±: standard error of three pseudo-replicates per treatment.

| Treatment | Individuals free swimming (larvae/ml) |
|---|---|
| Water Control | 94 ±13.8 |
| DMSO control | 92 ±8.7 |
| TBTO (0.2 µg/L) | 53 ±3.9 |
| TBTO (2 µg/l) | 25 ±3.9 |
| Rosiglitazone (0.4 µg/l) | 102 ±3.8 |
| Rosiglitazone (4 µg/l) | 89 ±9.2 |
| ATRA (0.06 mg/L) | 17 ±0.6 |
| ATRA (0.6 mg/L) | 0 |
| PFOA (20 mg/L) | 68 ±4.1 |
| PFOA (50 mg/L) | 62 ±5.7 |

The effect of chemicals on expression of three NR genes was tested with qPCR at trochophore larval and D-shaped larval stages (Supplementary Figure 4.4). TBTO, rosiglitazone, PFOA or ATRA do not seem to alter the expression of these receptor genes in any of these exposed oyster individuals

.

## 4.4. Discussion

### 4.4.1. *Cloning and phylogenetic analyses of* C. gigas *nuclear receptors*

We cloned three NRs, namely CgRXR, CgRAR and CgPPAR. CgRXR shows two different isoforms, CgRXR-1 and CgRXR-2, with CgRXR-2 having a 13 aa long insertion in the T-box, a locus typically seen in the isoformation of RXR homologues in molluscs. All receptors display the distinct NR domains including DBD and LBD for hypothetical/potential DNA and ligand binding, respectively. CgRXRs and CgRAR show high sequence identities for the conserved domains (>82 % DBD; 50-94 % LBD) to molluscan homologues and even to remotely related species such as the human RXRs. CgPPAR, on the other hand, displays much lower sequence identities to the two other identified molluscan and the three human PPARs (56-78 % DBD; 22-38 % LBD). Phylogeny of the receptors confirms this conservation pattern (Figure 4.2). Sequence analysis, identities and phylogeny are used to make assumptions for the *C. gigas* NR functions.

### 4.4.2. *DNA binding and dimerisation potential*

DNA binding, the most important mechanism for NR mediated regulation of gene transcription, is mainly achieved by the DBD binding to response elements in promoters of the genes to be transcribed [Germain et al., 2006]. In addition to general DNA binding sites, the P-box in the first zinc finger contains residues necessary for sequence discrimination of response elements. All three oyster receptors contain identical P-box sequences (‘CEGCKG’) to the NR orthologs in humans and gastropods, leading to the assumption of a similar binding behaviour to specific response elements in the oyster receptors. Human RXR, RAR and PPAR homologs recognise response elements with direct repeats (DR) of various length (5’-AGGTCA($n_x$)AGGTCA-3’) [reviewed in Helsen & Claessens, 2014]. For the gastropod rock shell *T. clavigera* TcRXRs, freshwater snail *B. glabrata* BgRXR and dog whelk *N. lapillus* NlRXR-1, binding to DR1, DR2 and/or DR5 response elements has been confirmed [Bouton et al., 2005; Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2011]. Similar findings were shown for *N. lapillus* NlRAR [Guterrez-Mazariegos et al., 2014] and *T. clavigera* TcRAR, for which a variant of the DR5

response element (5'-AGTTCA-3') was used [Urushitani et al., 2013]. Human PPARs bind to DR1 and DR2 response elements [Gervois et al., 1999; Kliewer et al., 1992]. Vertebrate RXRs are known to form homodimers, but also for being the predominant heterodimer partner for other receptors such as RAR and PPAR [Zhang et al., 1992; reviewed in Dawson & Xia, 2012]. Dimerisation is a complex process observed in many receptors. It involves several receptor domains such as the DBD and LBD including the D- and T-boxes, and the homo and heterodimer interfaces [Zechel et al., 1994a, 1994b; reviewed in Dawson & Xia, 2012]. CgRXR-1 and CgRAR both show identical residues to their molluscan homologues for the RXR/RXR homodimer and RXR/RAR heterodimer interfaces as well as for T- boxes. In gastropods BgRXR, TcRXRs and NlRXR-1 homodimer formation and binding to the response element DR1 as well as heterodimerisation of TcRXR-1/RAR (DR5) and NlRXR-1/RAR (DR1, DR2 & DR5) have been confirmed [Bouton et al., 2005; Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2011]. The high sequence identity and dimer interface identification suggests successful homodimerisation of CgRXR-1, and that heterodimerisation of CgRXR-1 and CgRAR is likely. The CgRXR-2, however, might not be effective in homodimer formation. The isoform insertion is localised in the T-box, which might inhibit the dimerisation as indicated for TcRXR-2 [Urushitani et al., 2011]. Few differences were detected between the RAR D-boxes of the different mollusc species, which could indicate that heterodimerisation is not possible. However, the D-boxes, normally involved in dimerisation of many heterodimers [Zechel et al., 1994b; reviewed in Dawson & Xia, 2012], seem not to be necessary for all heterodimer formations. Human RXR/RAR heterodimer binding to the DR2 response element exclude the D-box in their dimerization process [Zechel et al., 1994a]. Human PPARs are strictly heterodimers and only bind to RXRs, which dimerise via the T-box only, excluding the D-box from the process [IJpenberg et al., 1997]. Hence, the differences in length and sequence of the D-box for CgPPAR and human PPARs, as shown by our alignment results (Supplementary Figure 4.1), do not prevent heterodimerisation of the oyster PPAR. Sequence analysis reveals a putative dimer interface in CgPPAR, and CgRXRs show RXR/PPAR heterodimer interfaces.

### 4.4.3. *Ligand binding potential*

The ability to bind ligands is a common feature of RXRs, RARs and PPARs in many species [Dawson & Xia, 2012; Germain et al., 2006]. All three *C. gigas* receptors possess a LBD, including ligand binding sites. Sequence consensus between oyster, molluscan and human RXRs are very high for the LBD. The residues with which ligands bind are identical. Gastropod and human RXRs have been shown to bind to 9RA and TBT [Bouton et al., 2005; Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2011, 2013]. The 3D models of CgRXR and computational docking show a high likelihood that 9RA and TBT induce conformational changes in CgRXR, which would cause a ligand-dependent effect. A similar ligand binding positioning in human and oyster models is seen for 9RA including an induced fit in the hydrophobic pocket. This suggests successful binding of 9RA to CgRXR and a possible induction of an agonistic signal equivalent to human or snail RXRs. In HsRXRα the tin atom of the TBT molecule covalently binds to a cysteine thiol in the H11 helix [le Maire et al., 2009] and induces agonistic conformational changes for receptor dimer activation. In oysters this cysteine (C415) is conserved (Supplementary Figure 4.1) and with TBT preferred binding exposing its tin atom to the cysteine, we hypothesise that CgRXR is likely to respond to TBT.

In contrast to human RARs, which are able to bind a variety of natural and synthetic retinoic acids such as 9RA, 13RA, ATRA and synthetic agonists (e.g. TTPNB) [Idres et al., 2002; Pignatello et al., 1997], CgRAR may be unable to respond to such ligands. Based on the 3D models and the computational dockings, six residues in the LBD of the oyster RAR prevent binding of retinoids in the required conformation. Taking into account that neither of the two gastropod receptors TcRAR or NlRAR respond to RAs *in vitro* [Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2013], it is possible that molluscan species in general do not respond to RAs. This supports the theory of ligand binding loss for molluscan RARs [Gutierrez-Mazariegos et al., 2014]. It has been proposed that the urbilaterians, the last common ancestor of bilaterians before they split into deuterostomes and protostomes, possessed a proto-RAR, which was able to respond to RAs. Accordingly, loss of ligand binding could have emerged through just a few amino acid mutations in the LBD. In the dog whelk *N. lapillus,* ligand binding of NlRAR to 9RA and ATRA could be artificially restored *in vitro* through single or multiple mutations of up to seven amino acids to the equivalent human residues. CgRAR and NlRAR share five of the residues, known to prevent

successful binding. Both receptors, when these residues are mutated to the equivalent in human homologs, display binding energies similar to human retinoid binding RARs. CgRAR seems to have lost its ability to respond to RAs due to similar mutations in the LBD as seen for NlRAR.

CgPPAR is phylogenetically distinct from the human PPARs and its LBD is smaller than the human LBDs. The 3D models show that CgPPAR is missing a human PPAR typical helical region, H2, which results in a smaller ligand binding pocket (LBP). The LBP in the human protein is very large (~1300 $Å^3$), but ligands usually only occupy about 30-40 % of the cavity [Nolte et al., 1998]. Hence, the shorter LBD of CgPPAR does not automatically make the LBP too small for successful ligand binding. Computational docking shows that agonistic ligands such as rosiglitazone and TBT still fit in the LBP of the CgPPAR models. Rosiglitazone was chosen as putative CgPPAR ligand as it is an antidiabetic drug designed to interact with the human ortholog PPARγ [Lehmann et al., 1995]. However, rosiglitazone does not seem to position itself correctly to induce conformational changes. TBT, on the other hand, could stimulate a signal. TBT binds to a cysteine C285 in the H3 helix of HsPPARγ with an ionic bond [le Maire et al., 2009] and acts as a weak agonist [Harada et al., 2015; le Maire et al., 2009]. In the generated CgPPAR model, a cysteine C322 in the H3 helix would be in the position to bind to the tin compound. One must bear in mind that the generated 3D models for CgPPAR might be incorrect due to low sequence identity of LBD to human models. In contrast, since CgPPAR lacks the final H12 including an AF-2, which is required for ligand-dependent activation in most NRs, CgPPAR may not be activated by ligand binding or recruit different means of passing on the induction signal.

### 4.4.4. *Chemical effects on oyster embryos*

The oyster embryo bioassay showed that chemicals, such as the natural compound ATRA and synthetic compounds like TBTO and PFOA, can affect the oyster embryo development at different concentrations. These effects include visible impacts on shell development, as well as malformation of the soft tissue of the animal itself. The highest chemical concentrations lead to arrested development at the trochophore larvae stage and in few cases even to high mortality. TBTO affected embryo development at a low dose of 0.2 µg/L, which is comparable to previous research,

which also reported extruding velum in *C. gigas* embryos after 24 h exposure to the lowest tested TBTO concentration (1 µg/L) and a LC10 of 0.36 µg/L [Tsunemasa & Okamura, 2011]. We hypothesise that the observed effects of TBTO on oyster development are caused by disruption of the CgRXR function. Our results support an interaction of TBT and the oyster CgRXR receptor. Previous research also strongly suggests a correlation of observed effect in gastropods and RXR interaction with TBT [Nishikawa et al., 2004; Horiguchi et al., 2007]. In seawater TBTO ($Bu_3$-Sn-O-Sn-$Bu_3$) breaks down into two TBT units ($2Bu_3$-$Sn^+$) [Laughlin et al., 1986], which are available to interact with RXRs. The permissive RXR/PPAR heterodimer has been previously suggested as the specific target for TBT. Expression profiles of TBT-exposed dog whelk *N. lapillus* showed alteration of genes potentially involved in PPAR signalling pathways [Pascoal et al., 2013]. TBT can activate the vertebrate RXR/PPAR heterodimer via RXR alone or possible by both partners [Harada et al., 2015; Grün et al., 2006; Kanayama et al., 2005; le Maire et al., 2009]. The putative interaction of TBT and CgPPAR indicated by our sequence and docking analysis strengthens this theory. *C. gigas* embryos, on the other hand, were not visibly affected by rosiglitazone. This supports our computational docking result, which showed that rosiglitazone does not interact with CgPPAR. However, previous research on the gastropod *N. lapillus* indicated that there might be an RXR/PPAR heterodimer involvement in imposex formation after exposure to rosiglitazone, displaying similar effects to that of TBT-induced imposex [Pascoal et al., 2013]. This suggests that oysters and snail PPARs, even though closely related, react differently to rosiglitazone.

The effects of ATRA on oyster development, on the other hand, are possibly not caused via a traditional agonistic relationship of ligand and retinoid receptors similar to vertebrates. The computational docking results refute an agonistic vertebrate-like interaction between RAs and CgRAR. Studied gastropod RARs did also not show receptor activation by ATRA [Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2013]. Additionally, the expression of receptors in ATRA-exposed oyster larvae did not change compared to untreated oyster larvae. Human orthologs of RAR are self-regulated via an agonistic response to ATRA [Wan et al., 1998; Wu et al., 1992], which also affects the expression of RXR isoforms [Wan et al., 1994]. Oysters CgRAR and CgRXR do not vary in their expression suggesting a different regulation mechanism in *C. gigas* not related to ATRA. An indirect effect of ATRA on CgRXR is

a potential explanation, based on suggestions of an isomerization process of RAs; ATRA would be converted to 9RA and 13RA, with 9RA consequently interacting with CgRXR. Indeed, this mechanism is suggested to be present in several gastropod species [Gesto et al., 2013].

Our data of PFOA, a perfluoralkyl carboxylate used as synthetic salt, confirms the minor risk by a direct exposure from PFOA as previously reported for marine species [Mhadhbi et al., 2012] and it displays toxicity at what would be high levels of milligrams per litre. Although the environmental concentrations of PFOA are low (oceanic/coastal waters: 15 pg/L – 190 ng/L) [Yamashita et al., 2005], this chemical has been classified as a substance of a very high concern due to its high persistence (non-degradable) and ubiquitously existence in terrestrial and aquatic habitats, atmosphere, food, drinking water, plants, animals including humans [EFSA, 2011; Giesy et al., 2010; Mhadhbi et al., 2012; Vierke et al., 2012; Yamashita et al., 2005;]. The concern regarding its persistence and bioaccumulation abilities raises questions to its mode of action. PFOA is a known agonist of vertebrate PPARα and PPARβ [Vanden Heuvel et al., 2006] and successfully disrupts the PPAR pathways [Abbott et al., 2007; Cheng & Klaassen, 2008]. Due to the lack of a PPAR template bound to PFOA and the inability to generate a reliable CgPPAR 3D model the possible interaction of PFOA with CgPPAR was not be assessed. Nevertheless, the presence of a PPAR homolog in the Pacific oyster forms a starting point for further investigations of PFOA mode of action in protostome invertebrates.

## 4.5. Conclusion

In conclusion, we cloned and characterised three Pacific oyster nuclear receptors, CgRXR, CgRAR and CgPPAR, and examined their potential/inability to successfully induce gene expression and bind to agonistic ligands. Exposure to these ligands, natural (RAs) and synthetic (TBT, PFOA) chemicals, illustrated the sensitivity of oyster embryo development to disruption and the risk at which marine invertebrates are placed in polluted environment. The mode of action of many disruptive chemicals like TBT, ATRA and PFOA on invertebrates is far from being fully comprehended, and further understanding of chemical impact on marine species is urgently needed. This study helps to further understand the mechanism of environmental pollutants on marine invertebrates and aids in assessment and prediction of anthropogenic contaminants.

Chapter five

# Nuclear receptor transcriptional activation by ligand binding: dual-luciferase reporter assays


## Abstract

Gene expression in metazoan species is partially regulated by nuclear receptors, which react and bind to internal and external stimuli. Dual-luciferase reporter assays are used to test the ligand binding ability and gene expression activation of nuclear receptors *in vitro*. Here, the full sequences of three receptor homologs of the Pacific oyster, *Crassostrea gigas*, (retinoid X receptor CgRXR, retinoic acid receptor CgRAR & peroxisome proliferator-activated receptor CgPPAR) were co-transfected with reporter vectors containing DR1, DR2 or DR5 response elements and were exposed to proposed receptor-specific natural (all-*trans*, 9-*cis* or 13-*cis* retinoic acid) or synthetic (tributyltin oxide (TBTO), rosiglitazone, TTNPB, PFOA) ligands. None of the oyster receptors were activated by any of the chemical compounds tested. The human receptor, HsRXRα, and a gastropod *Thais clavigera* receptor, TcRXR, used as positive controls, showed ligand-dependent transactivation. However, the two additional human positive control receptors, HsRARβ and HsPPARγ, did not display transcriptional activation to their known ligands, leading to the assumption of a technical issue with the assay. Adjustment of observed differences in the spacer length of the expression vectors neither resulted in activation of the oyster receptors. The dual-luciferase reporter assay is a reinforced method, but differences in the implementation can be applied. The results of the performed assay suggest that a different dual-luciferase assay technique would be advisable to obtain answers about the oyster receptor ligand binding abilities.


This chapter is an unpublished experimental chapter. All experiments including the preparations were conducted by myself. Plasmid constructs were constructed in cooperation with my supervisor Dr Tim Bean. Practical teaching of the dual-luciferase reporter assay technique took place at the National Institute for Basic Biology in Okazaki under supervision of Dr Shinichi Miyagawa. The final experiments were conducted at the University of Exeter.

## 5.1 Introduction

Natural and synthetic compounds are introduced to the environment through numerous anthropogenic systems. Many of these compounds are harmless, but some can have severe effects on a wide range of species. Species like the Pacific oyster *Crassostrea gigas*, which are immobile, are under extreme pressure of environmental pollution as they cannot actively avoid the exposure. Exogenous compounds can activate gene expression by binding to a specific transcription factor group, called nuclear receptors (NRs) [McLachlan. 2001]. Several members of the NR superfamily are able to bind to small hydrophobic ligands [Germain et al., 2006]. Xenobiotics, foreign compounds, which are not naturally produced or are present in an abnormal concentration, can interfere with the NR gene expression function. They can interact directly with the receptor through binding and sending incorrect signals for gene regulation [McLachlan. 2001].

The presence of NRs in marine invertebrates has been established [Hwang et al., 2014; Kauer et al., 2015; Reizel & Tarrant, 2009; Vogeler et al., 2014; Wiens et al., 2003], but their function and interaction with putative ligands is poorly understood. In general, NRs regulate gene expression by attaching to a receptor specific response element (RE), two hexameric half-site motifs separated by short spacers, in the gene promoter region. They bind to the DNA either as monomer, homodimer or heterodimer with the retinoid X receptor (RXR) as the predominant heterodimer partner in many species [Mangelsdorf & Evans, 1995]. The Pacific oyster contains 43 NRs [Vogeler et al., 2014], but none of these are yet functionally resolved. Recently we cloned and characterised three oyster NRs: CgRXR (two isoforms), CgRAR and CgPPAR [Vogeler et al., submitted; see Chapter 4]. Sequence analyses and computational docking suggested putative DNA binding, receptor dimerisation and ligand interaction. CgRXR, a RXR ortholog, seems to be able to form homodimers and heterodimers and bind to REs with direct repeats (DR). The retinoic acid (RA) isomer 9-*cis* RA (9RA), a natural ligand of vertebrate RXRs [Heyman et al., 1997; Levin et al., 1992], and tributyltin (TBT), a highly toxic organotin agonist for vertebrate RXRs [Grün et al., 2006; le Maire et al., 2009], were identified as putative ligands. The retinoic acid receptor (RAR) ortholog in the Pacific oyster, CgRAR, on the other hand, appears to be unresponsive to RA and isomers, although vertebrate RARs interact with retinoids [Idres et al., 2002]. This is similar to findings in gastropod RARs (*Thais clavigera* and *Nucella lapillus*) [Gutierrez–Mezariegos, 2014;

Urushitani et al., 2013], a sister-class to bivalves. CgRAR seems to be able to form heterodimers with CgRXR, with a mechanism comparable to vertebrate RXR/RAR interactions. Such a heterodimer formation is also suggested for CgPPAR, a vertebrate peroxisome-proliferator-activated receptor (PPAR) ortholog, which could potentially dimerise with CgRXR. This is of particular interest as the vertebrate RXR/PPAR dimer belongs to the permissive heterodimers [reviewed in Dawson & Xia, 2012]. Many heterodimers can only be activated by the RXR heterodimer partner only (non-permissive) (e.g. RXR/RAR heterodimer). Other heterodimers, however, are triggered by both partners (permissive) increasing the risk of chemical disruption due to two possible xenobiotic targets. Predictions of putative CgPPAR ligands are sparse as CgPPAR showed high sequence dissimilarities to its human homolog. The ability to interact with ligands and the resulting potential disruption risk needs to be confirmed for these three Pacific oyster receptors.

The binding ability to putative ligands and subsequently activation of NRs can be tested with genetic reporter assays, a technology widely used to study gene expression as well as the associated cellular events. The dual-luciferase reporter assay, in particular, can be used to investigate ligand binding activity and the assay benefits from the bioluminescence genes of two different species. Cells from an immortalised cell line are co-transfected with different vector types (Figure 5.1). The expression vector contains a sub-cloned NR coding sequencing and a promoter sequence for efficient high level expression of the recombinant protein by the cell itself. The reporter vector expresses the firefly (beetle species) luciferase gene for which a promoter sequence including the NR specific response element motifs is sub-cloned ahead of the luciferase gene. The third vector is an internal control vector containing a promoter sequence and a *Renilla* (sea pansy species) luciferase gene. It is used to minimalize experimental variability such as cell viability, transfection and assay efficiency. After co-transfecting all required vectors the cells are exposed to the putative ligand. In case of successful ligand binding and activation of the receptor, firefly luciferase is transcribed and translated in the cell. If the ligand is incapable of activating gene transcription via receptor binding, only the internal control produces its *Renilla* luciferase. The luciferase production is quantified, using the respective firefly and *Renilla* luciferase bioluminescent reactions. The cells are lysed and luciferin and coelentrazine are each added successively. The firefly luciferase catalyses the oxidation of beetle luciferin to oxyluciferin. *Renilla* luciferase

catalyses the reaction of coelentrazine to coelenteramide. Both reactions produce detectable light, which is measured separately with a luminometer.

The dual-luciferase reporter assay is used to investigate the ligand binding ability of the Pacific oyster NRs CgRAR, CgPPAR and the two isoforms of CgRXR, CgRXR-1 and CgRXR-2. Known natural and synthetic ligands for vertebrate NR homologs are used as putative ligands. Human receptors HsRXRα, HsRARβ and HsPPARγ, as well as the sea snail *T. clavigera* receptor isoforms TcRXR-1 and TcRXR-2 operate as positive controls. Although the principle of dual-luciferase assays is relatively simple and comparable data is obtained within a few days, this assay requires individual preparation of each of the assay components. The numerous components of the assay also increase the chance for unpredictable complications.

Figure 5.1 on the next page

**Figure 5.1: Schematic representation of a dual-luciferase reporter assay.** Cells are co-transfected with expression (E), reporter (R) and control (C) vectors. The sub-cloned NR gene is expressed by the cell itself (green open circles) and the cells are exposed to putative ligands (yellow triangles, red circles). Successful ligand binding results in attachment of receptors to the promoter region of the reporter vector. The promoter region contains a specific receptor response element (RE motifs). Firefly luciferase (*luc* gene) is transcribed and translated. The cell produces low levels of *Renilla* luciferase (*Rluc* gene) of the control vector. In case of binding inability of ligand and receptor no firefly luciferase signal is produced. A luminometer separately measures the bioluminescent reaction of the luciferases after cell lysis and the addition of luciferine and coelenterazine, respectively. Firefly luciferase catalyses the oxidisation of luciferin to oxyluciferin. *Renilla* luciferase catalyses the oxidation of coelenterazine to coelenteramide.

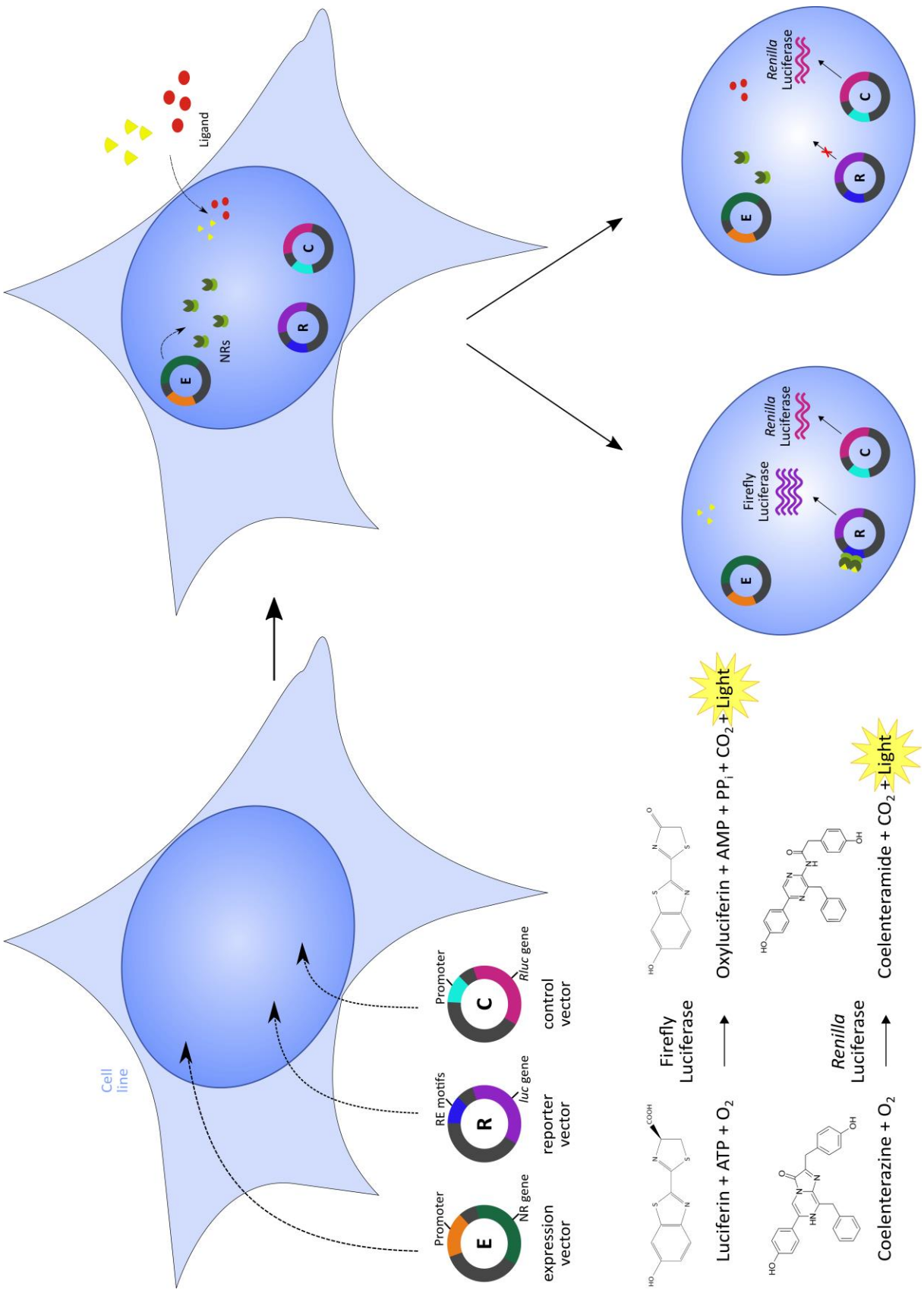



Figure 5.1 legend: see previous page

## 5.2. Materials and Methods

### 5.2.1. *Chemical reagents: ligands*

Natural all-*trans* retinoic acid (ATRA), 9-*cis* retinoic acid (9RA), 13-*cis* retinoic acid (13RA), the synthetic retinoic acid TTNPB, the synthetic organotin bis(tributyltin) oxide (TBTO), the antidiabetic drug rosiglitazone (Rosi), and the synthetic compound perfluorooctanoic acid (PFOA), were purchased from Sigma-Aldrich. Stock solution and dilution were freshly prepared with dimethylsulfoxide (DMSO). The concentration of DMSO in the culture medium did not exceed 0.1 %.

### 5.2.2. *Plasmid vector constructs*

The full coding region of the *C. gigas* NRs CgRXR-1, CgRXR-2, CgRAR and CgPPAR were amplified by RT-PCR with the GoTaq system (Promega), using forward and reverse primers including Kozak sequence (5'-GCCACC-3'), and start and stop codons (Table 5.1). The products were purified with the QIAquick PCR Purification Kit (Qiagen, UK) and cloned into the *KpnI-BamHI* site of the pcDNA3.1(+) plasmid (Invitrogen), which function as expression vectors. The expression vector HsRXRα, HsRARβ and HsPPARγ including the human receptors RXRα, RARβ and PPARγ, and TcRXR-1 and TcRXR-2 including the gastropod *T. clavigera* receptor isoforms RXR-1 and RXR-2 were kind gifts from the laboratory of Professor Taisen Iguchi at the National Institute of Natural Science in Okazaki, Japan, and Dr Toshihiro Horiguchi from the National Institute of Environmental Studies in Ibaraki, Japan. (HsRXRα, HsRARβ, TcRXR-1 and TcRXR-2 expression vectors have previously been published [Urushitani, et al., 2011, 2013]. The human cytomegalovirus (CMV) promoter of the pcDNA3.1(+) vector regulates the high-level expression of the introduced NR genes within mammalian cells. An empty vector was used as negative expression vector control (receptor absence). All expression vector inserts were verified by sequencing (T7_for: 5'-TAATACGACTCACTATAGG-3'; pCR3.1-BGH_rev: 5'-TAGAAGGCACAGTCGAG-3).

Three different reporter vectors were constructed using the pGL4.23 luciferase reporter vector (Promega), which contains a gene that encodes firefly (*Photinus pyralis*) luciferase. Double strained oligonucleotides were created by annealing two complementary single stranded oligonucleotides (Table 5.1), each including a

fourfold repeat of the half-site motif sequence 5'-AGGTCA-3' separated either by one, two or five base pairs (bp), representing direct repeat response elements DR1, DR2 and DR5 (Figure 5.2). The oligonucleotides were separately cloned into the *KpnI-HindIII* sites of the pGL4.23. To prevent the insertion of multiple copies of unwanted concatenated oligonucleotidesthe plasmid was digested with the restriction enzyme *KpnI* and re-annealed. All cloned DNA sequences were verified by sequencing. An empty pGL4.23 vector was used as negative reporter vector control (response element absence). The reporter vector RXRE (a gift from the Iguchi laboratory and Dr Horiguchi), a pGL4.23 vector containing a DR1 response element, was used as positive control for reporter vector testing.

The pRL-Tk vector (Promega) was used as internal control vector and encodes *Renilla* (*Renilla reniformis*) luciferase. It contains the herpes simplex virus thymidine kinase (HSV-TK) promoter for low to moderate levels of *Renilla* luciferase expression in co-transfected mammalian cells to provide a background luminescence for promoter activity quantification.

All vector constructs from a single transformed bacterial colony were preserved in 15% glycerol and stored for long term at – 80 °C.

**Table 5.1: Primer and oligonucleotide sequences for sub-cloning of the expression, reporter and adjusted expression vectors.** Cg: *Crassostrea gigas*; Hs: *Homo sapiens*, DR: direct repeat; underlined letters: response element motifs.

| | |
|---|---|
| ***Expression vectors: cloning into pcDNA3.1(+)*** | |
| CgRAR | For: 5'-CAAGGTACCGCCACCATGGACCCATCAGAAATGG-3'<br>Rev: 5'-AATGGATCCTATCAAGGAATACACACATTCT-3' |
| CgRXR | For: 5'-CAAGGTACCGCCACCATGAAGACGGACAGAATG -3'<br>Rev: 5'-AATGGATCCTATTAAGGTTTCCATTCCA -3' |
| CgPPAR | For: 5'-CAAGGTACCGCCACCATGGTGCACGGTTGACA-3'<br>Rev: 5'-AATGGATCCTAGCAGTTGGTGTCATCT -3' |
| ***Reporter vectors: cloning into pGL4.23*** | |
| DR1 | 5'-TAAGGTCACAGGTCACAGGTCACAGGTCAC-3'<br>5'-TCGAAGTGACCTGTGACCTGTGACCTGTGACCTTACCGTA-3' |
| DR2 | 5'-AAGGTCACAAGGTCACAAGGTCACAAGGTCAC-3'<br>5'-TCGAAGTGACCTTGTGACCTTGTGACCTTGTGACCTTCCATG-3' |
| DR5 | 5'-CCAAGGTCACCAGAGGTCACCAGGAGGTCACCAGGAGGTCACG-3'<br>5'-TCGAACGTGACCTCCTGGTGACCTCCTGGTGACCTCTGGTGACCTTGGCCATG-3' |
| ***Adjustment: promoter sequence insertion*** | |
| Cg_inserts | 5'-CTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGGTGTAC-3'<br>5'- ACCACTGGACTAGTGGATCCGAGCTCGGTACCAAGCTTAAGTTTAAACG-3' |
| Hs_inserts | 5'-CTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGG-3'<br>5'- CTAGCCACCACACTGGACTAGTGGATCCGAGCTCGGTACCAAGCTTAAGTTTAAACG-3' |

### 5.2.3. *Expression vector promoter sequence adjustment*

The promoter sequence of all expression vectors (Pacific oyster, human and *T. clavigera*) were sequenced, including the CMV and T7 promoter as well as the start region of the NRs to verify presence and integrity of the CMV promoter, T7 promoter, the Kozak sequence and start codon (CMV_for: 5'-TGTTGGAGGTCGCTGAGTAG-3'). Insertions of 34 bp or 57 bp – annealed double strained oligonucleotides - were sub-cloned into the *NheI-KpnI* sites of all oyster expression vectors (CgRXR-1, CgRXR-2, CgRAR and CgPPAR) and into the *NheI* site of the human HsRARβ and HsPPARγ expression vectors, respectively (Table 5.1). This created a promoter region of equal length (T7 promoter - start codon: 75bp) for all expression vectors.

### 5.2.4. *Dual-luciferase reporter assay*

Dual-luciferase reporter assays were performed on COS-7 cells (fibroblast-like cells from monkey kidney tissue), seeded into 24-well plates at 2 x $10^4$ cells/ml and cultured in phenol red-free DMEM (Sigma-Aldrich) supplemented with charcoal/dextran-treated fetal bovin serum (Fisher), L-glutamine (Sigma-Aldrich) and penicillin-streptomycin solution (Sigma-Aldrich). After 24 h cells were transfected with 0.2 µg expression vector (single or receptor combinations), 0.4 µg reporter vector and 0.1 µg pRL-Tk control vector by using 1.8 µl Fugene HD Transfection reagent (Promega). Chemicals (ligands) were added to the medium at various concentrations after 4-5 h of incubation. The total volume of chemical mixture required did not exceed a volume of 1 µl. Chemical combinations were all made up to a total volume of 1 µl. Treatment with DMSO (1 µl) for each expression/reporter vector combination was used as no-chemical control. The dual-luciferase reporter assay system was used to measure the luciferase activity of prior lysed cells after 40-42 h. A NanoQuant Infinite M200Pro (Tecan) luminometer was used to measure the firefly luciferase activity followed by a measurement of the *Renilla* luciferase. Vector and chemical combination was performed in triplets for each experiment. Selected experiments were independently performed three times.

Additional reporter assays using HEK-293 cells (human embryonic kidney cells) were performed following the described protocol with one exception: cells were seeded into 24-well plates at 5 x $10^4$ cells/ml.

5.2.5. *Data analysis*

The promoter activity was calculated as firefly luciferase activity per *Renilla* luciferase activity:

$$\text{Promoter activity (PA)} = \frac{\text{firefly luciferase}}{\textit{Renilla}\ \text{luciferase}}$$

The mean of the promoter activity of the triplet measurements were calculated. The fold activation for each expression/reporter vector combination through a chemical was quantified as followed:

$$\text{Fold activation} = \frac{PA_{chemical}}{PA_{non\text{-}chemical\ control}}$$

In case of three independent experiments the mean of each fold activation and associated standard error (±SE) were calculated. Statistical analysis was performed by a one-way ANOVA followed by multiple pairwise comparisons with Tukey's Honestly Significant Difference Test. Means were considered significantly different at $P<0.05$.

## 5.3. Results

### 5.3.1. *Vector construct and cell line assessing*

The successful cloning of all *C. gigas* expression vectors was verified by sequencing. The human and rock shell *T. clavigera* expression vector gifts HsRXRα, HsRARβ, HsPPARγ, TcRXR-1 and TcRXR-2 were also verified and contain the correct NR sequences. Additional sequencing also confirmed the presence and correct orientation of the four half-site motif sequences (5'-AGGTCA-3') including the associated spacers for the reporter vector DR1, DR2, and DR5 in a pGL4.23 vector (Figure 5.2a). The sequence of four half-site motifs allows for more binding opportunities for the NR complex. RXRE, a DR1 reporter vector used with the expression vector HsRXRα in previous research [Urushitani et al., 2011], was used as positive control for the reporter vector functioning of this study. A dual-luciferase reporter assay showed a fold activation of firefly luciferase for HsRXRα co-transfected with RXRE when exposed to 9RA ($10^{-5}$ M) indicating an HsRXRα receptor activation by 9RA (Figure 5.2b). 9RA is a natural ligand of human RXRα [Heyman et al., 1997; Levin et al., 1992]. HsRXRα co-transfected with the reporter vector DR1 also showed transcription activation, which indicates a successful functioning of the DR1 reporter vector. The stronger signal produced by HsRXRα/DR1 compared to RXRE suggests an enhanced functioning of DR1. The used RXRE holds a mutation of one base pair containing a thymine instead of a guanine in motif 4, which could explain the differences in fold activation (Figure 5.2a). Positive functioning for the reporter vector DR2 and DR5 are assumed due to similar assembly to DR1.

The successful functioning of the cell line COS-7 for molluscan species is verified by positive fold activation for the snail TcRXR-1 co-transfected with DR1 when exposed to 9RA (Figure 5.2b). TcRXR-2 displays a three times lower fold activation compared to TcRXR-1. Previous research showed similar results, but verified a receptor binding and activation for TcRXR-1 and TcRXR-2 by 9RA at $10^{-8}$ M and $10^{-7}$ M, respectively [Urushitani et al., 2011]. CgRXR-1 and CgRXR-2 showed similar levels of activation compared to TcRXR-2 indicating a receptor/ligand interaction *in vitro*.

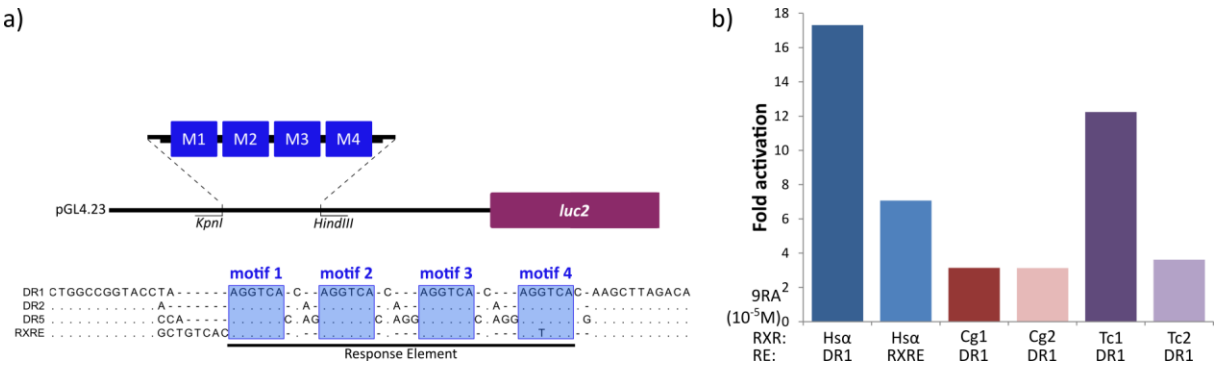



**Figure 5.2: Cloning reporter gene constructs and transcriptional activity of HsRXRα, CgRXRs and TcRXRs. a)** Schematic representation of response element (RE) insertion ahead of the firefly luciferase *luc2* gene of the reporter vector pGL4.23 including the restriction enzyme sites *KpnI* and *HindIII*. Three RE with four direct repeats of the motif 'AGGTCA' including different length of spacers between each motif (DR1, DR2 and DR5) were constructed. RXRE, a DR1 response element (gift from Iguchi laboratory in Okazaki), were aligned a long with the response element created in this study. Blue boxes M1-M4: motif 1-4; purple box *luc2*: Firefly luciferase gene. **b)** COS-7 cells were transfected with DR1 or RXRE reporter vectors in combination with human HsRXRα (Hsα), *Crassostrea gigas* CgRXR-1 (Cg1) and CgRXR-2 (Cg2), or *Thais clavigera* TcRXR-1 (Tc1) and TcRXR-2 (Tc2) fused expression vectors. Cells were incubated for 40-42 h with $10^{-5}$ M 9RA.

### 5.3.2. *Transcriptional activity of CgRXR, CgRAR and CgPPAR*

A single dual-luciferase reporter gene assay was performed to test CgRXR-1 and HsRXRα co-transfected with DR1 for a dose response to 9RA ($10^{-11}$ - $10^{-5}$ M). An empty pcDNA3.1(+) expression vector co-transfected with DR1 was used as base level for activation response to 9RA (negative control). CgRXR-1 showed an increase in response with increasing 9RA concentration (Figure 5.3a). The empty vector also displayed an increase in fold activation at a similar level to CgRXR-1. The response measured for CgRXR-1 was probably a response of the cell line to 9RA and CgRXR-1 is not directly activated by 9RA. However, the dose response of CgRXR-1 could also have been covered by the background noise of 9RA. Preliminary tests with no negative controls showed that the low level of CgRXR-1 response is consistent (Supplementary Figure 5.1). A similar performance was observed for CgRXR-2. HsRXRα co-transfected with DR1, on the other hand, showed an increase in activity with increasing concentration of 9RA (Figure 5.3a). HsRXRα seems activated at a 9RA concentration of $10^{-7}$ M ($P<0.05$) (Supplementary Figure 5.1), which matches the previously reported performance of this receptor [Urushitani et al., 2011].

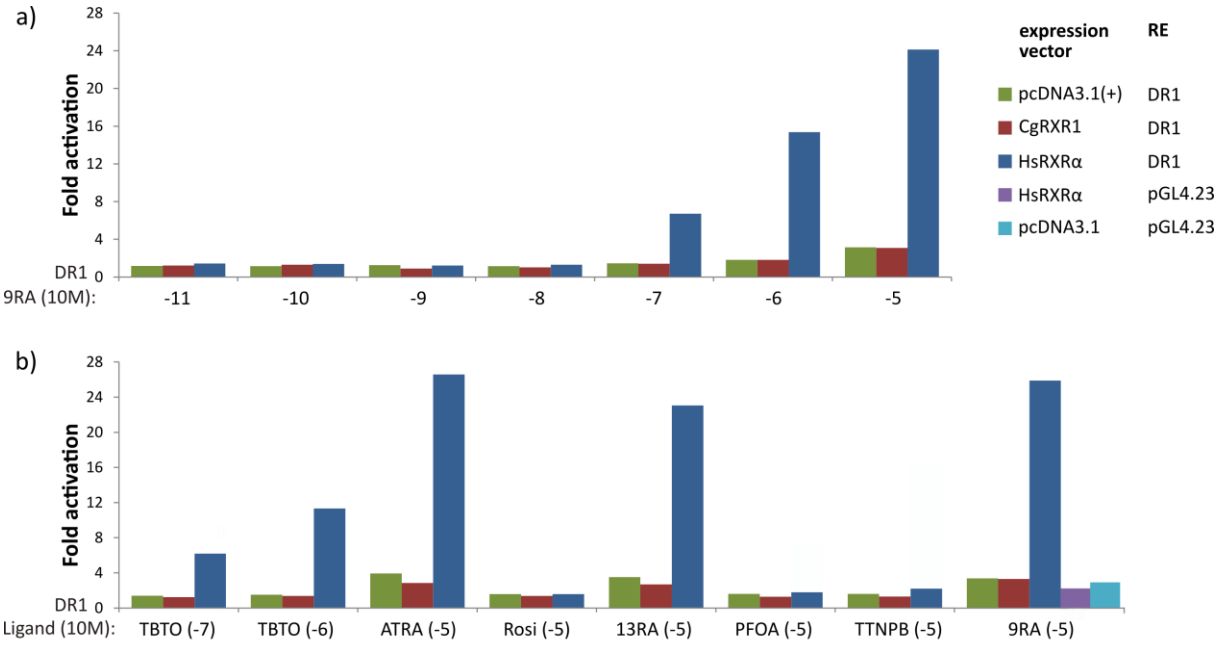



**Figure 5.3: Transcriptional activity of HsRXRα, CgRXR-1 and empty expression vector pcDNA3.1(+) dual-luciferase reporter assays.** COS-7 cells were transfected with either DR1 or empty reporter vector pGL4.23 in combination with either human HsRXRα, *Crassostrea gigas* CgRXR-1 or the empty expression vector. Cells were incubated for 40-42 h **a)** with 9RA in a dose response ($10^{-11}$-$10^{-5}$ M) and **b)** with different chemical at $10^{-5}$ M: ATRA, rosiglitazone (Rosi), 13RA, PFOA, TTNPB and 9RA. The TBTO exposures were performed with two different concentrations: $10^{-7}$ M and $10^{-6}$ M. RE: response element.

Different chemicals in different concentrations were tested for activation ability of CgRXR and HsRXRα (Figure 5.3b): TBTO ($10^{-7}$ M & $10^{-6}$ M), ATRA ($10^{-5}$ M), Rosi ($10^{-5}$ M), 13RA ($10^{-5}$ M), PFOA ($10^{-5}$ M), TTNPB ($10^{-5}$ M) and 9RA ($10^{-5}$ M). CgRXR-1 and the empty vector showed similar levels of low fold activation for all different chemicals indicating a non-activation of CgRXR-1 by all tested compounds. HsRXRα, on the other hand, responded to 9RA, ATRA and 13RA (significantly different, Supplementary Figure 5.1). The retinoids ATRA and 13RA are not natural ligands of vertebrate RXRs and the receptor transcription activity cannot be activated directly by these chemicals [Allenby et al., 1993; Heyman et al., 1992]. However, signal induction has been previously recorded for vertebrate RXRs exposed to ATRA and 13RA [Cai et al., 2010]. The measured response can, in fact, be an indirect response of HsRXRα to ATRA or 13RA. It has been reported that ATRA and 13RA could get, at least partially, isomerised to 9RA in cell lines [Shih et al., 1997; Urbach & Rando, 1994]. HsRXRα displayed a significant dose response to TBTO (Figure 5.3b, Supplementary Figure 5.1), a human RXRα agonist [Grün et al., 2006; le Maire et al., 2009; Urushitani et al., 2011]. Rosiglitazone, PFOA and TTNPB, on the other hand,

did not induce a response and none of these chemicals have been identified as human RXRα ligand. The responses of CgRXR-1 and HsRXRα fused with DR1 to the different chemicals were verified by three independent assays, but without empty vector controls conducted (Supplementary Figure 5.1).

Transcription activation of firefly luciferase was measured for RXR, RAR and in combination co-transfected with DR5, exposed to different types of natural RAs (ATRA, 9RA, 13RA) for all oyster and human receptors (Figure 5.4). Combinations of natural RAs (9RA+ATRA) and natural and synthetic RAs (9RA+TTNPB) were also tested. The response element DR5 was chosen for the CgRAR reporter assay. Preliminary tests showed the strongest fold activation of firefly luciferase for DR5 co-transfected with oyster RXR/RAR (Supplementary Figure 5.2). The RA isomers ATRA, 13RA and 9RA are natural ligands for vertebrate RAR [Idres et al., 2002] and successful signal induction has been reported for luciferase reporter assays using COS-7 and HEK-293 cell lines [Idrest et al., 2002; Moise et al., 2009]. Although transactivation activities were measured for HsRARβ, for CgRAR on its own as well as in combination with either HsRXRα or the putative heterodimer partner CgRXR, the results displayed several incongruities.

TTNPB is a highly potent synthetic vertebrate RAR agonist and induced firefly luciferase signals comparable to ATRA in previous studies [Moise et al., 2009]. In this study, however, the sole synthetic retinoid TTNPB induced luciferase gene expression at only very low levels for human and oyster RARs (Figure 5.4).

The levels of transactivation of co-transfected HsRXRα and HsRARβ are lower than single transfected HsRARβ and HsRXRα for the conducted luciferase assays (Figure 5.4a). Single transfected cells with HsRAR have been used to successfully verify binding and transactivation activities *in vitro* [Idres et al., 2002; Urushitani et al., 2013]. Additionally, weak binding to DR5 response elements and low transcription activity has been observed for HsRXRs *in vitro* [Cheskis & Freedman, 1997; Kurokawa et al., 1994; Parrado et al., 2001; Urushitani et al., 2013]. Human RXRs and RARs have a high affinity to form heterodimers with each other and induce gene transcription in the presence of an agonistic signal [reiewed in Mangelsdorf et al., 1994]. The resulting transcriptional activity signal is estimated to be stronger compared to the receptor partners alone due to an expected higher affinity to the natural DR5 response element of the heterodimer [Cheskis & Freedman, 1997; Mangelsdorf et al., 1994]. However, the results of the two conducted luciferase assay

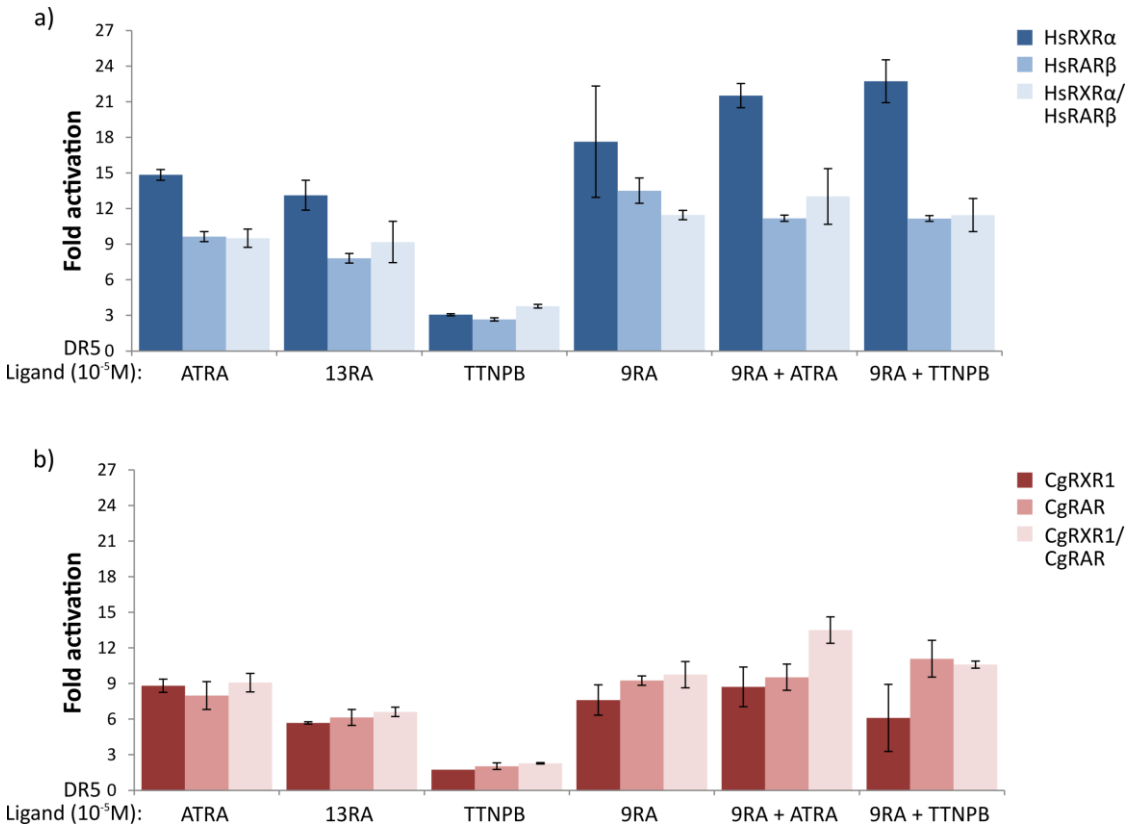



**Figure 5.4: Transcriptional activity of human and oyster RXR, RAR and RXR/RAR dual-luciferase reporter assays.** COS-7 cells were transfected with DR5 reporter vector and incubated for 40-42 h with natural (ATRA, 13RA, 9RA) and synthetic (TTNPB) RAs or a combination of natural (9RA+ATRA) and synthetic (9RA+TTNPB) RAs at a concentration of $10^{-5}$ M. The transfected expression vectors were either **a)** human receptors HsRXRα, HsRARβ, HsRXRα/HsRARβ or **b)** *Crassostrea gigas* receptors CgRXR-1, CgRAR, and CgRXR-1/CgRAR. The results of duplicates are shown as mean. Error bars display high and low fold activation of measured duplicates.

showed transcription activity approximately lower by a third lower for co-transfected human RAR and RXR receptors compared to HsRXRα single transfection when exposed to natural RAs, indicating either non activation of the receptors or unsuccessful heterodimerisation. Similar results were observed for the oyster receptors, which displayed equal levels of transcription activities for single exposures to natural RAs for all receptor combinations (Figure 5.4b). Only CgRXR/CgRAR exposed to 9RA+ATRA and 9RA+TTNPB as well as CgRAR to 9RA+TTNPB displayed stronger signal inductions compared to single-transfected CgRXR or CgRAR, respectively. This vaguely proposes a successful activation of CgRAR/CgRXR heterodimer by 9RA in combination with ATRA or TTNPB. The data could not bet tested for statistical significance as only two measurements per sample were available.

Putative chemical interaction with the transfected DR5 reporter vectors were, unfortunately, not tested, which could have explained the observed transactivation activity. As for DR1, chemicals could also produce misleading firefly luciferase signals without a transfected receptor. An additional assay set, however, showed only a weak activation increase for cells co-transfected with HsRARβ (4.8 fold activation) and DR5 compared to cells co-transfected with empty expression vectors (3.0 fold activation). This indicates a strong chemical reaction to the cell construct itself.

HsPPARγ co-transfected with DR1 or DR2, both known response elements of HsPPARγ [Gervois et al., 1999; Kliewer et al., 1992], could not be activated by rosiglitazone in a single dual-luciferase reporter assay (Figure 5.5). The measured fold activation did not exceed the base limit set by the empty vector. Rosiglitazone, an antidiabetic drug and a potent agonist for human PPARγ [Lehmann et al., 1995], activates vertebrate PPARγ co-transfected with a DR1 reporter vector in dual-luciferase reporter assays. PPARγ works as heterodimer and cannot act as homodimer [Kliewer et al., 1992]. Although previous reporter assays did not require a heterodimer partner for fold activation of PPARγ [Lin et al., 2012; Misra et al., 2008], HsRXRα has been provided as potential heterodimer partner. However, this also did not induce a detectable signal.

The HsRXR/PPAR dimer is a permissive heterodimer and can be activated via both

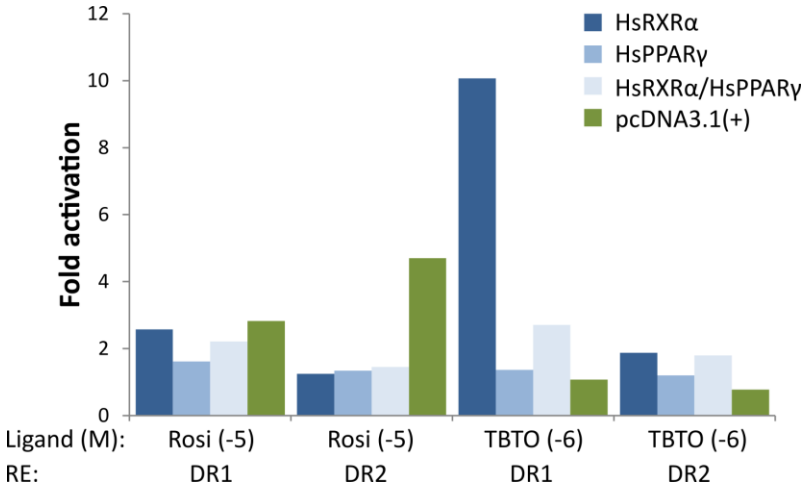



**Figure 5.5: Transcriptional activity of human HsRXRα, HsPPARγ, HsRXRα/HsPPARγ or empty expression vector pcDNA3.1(+) dual-luciferase reporter assays.** COS-7 cells were transfected with expression vectors fused with either HsRXRα, HsPPARγ, HsRXRα/HsPPARγ or an empty vector in combination with DR1 or DR2. The cells were exposed for 40-42 h with $10^{-5}$ M rosiglitazone (Rosi) or $10^{-6}$ M TBTO. RE: response element.

NR partners. To test functional HsRXRα and HsPPARγ heterodimerisation, the expression vector combination, co-transfected with either DR1 or DR2, was exposed to the RXR agonist TBTO ($10^{-6}$ M). The TBTO induced a strong firefly luciferase signal in cells co-transfected with DR1 and HsRXRα, but a similar signal for HsRXRα/HsPPARγ could not be acquired, indicating unsuccessful heterodimerisation activation. The weak measured fold activation is probably due to activation of HsRXRα by TBTO.

### 5.3.3. *Dual-luciferase reporter assay optimisation attempts*

Sequencing of the promoter region of all expression vectors revealed that the human and *T. clavigera* receptors were inserted into different restriction sites of the pcDNA3.1(+) vector (Figure 5.6a). This resulted in different spacer lengths between the CMV promoter and the Kozak sequence (Figure 5.6b). Number of base pairs (bp) downstream of the T7 promoter: HsRXRα, TcRXR-1: 75 bp; CgRAR: 45 bp; CgRXR-1, CgRXR-2: 42 bp; CgPPAR: 41 bp; HsRARβ, HsPPARγ: 18 bp. The oyster receptors vary slightly due to additional base pairs ahead of the Kozak sequence. The CMV and T7 promoter sequence were identical for all expression vectors. To test if the spacer length influences the assay activation success, the spacer lengths were adjusted to the functioning HsRXRα length. The spacer sequence of CgRXR-1, CgRXR-2 and CgPPAR has been successfully extended (Figure 5.6c). Dual-luciferase assays using the adjusted CgRXR-1 and CgRXR-2, however, did not indicate a successful agonistic response to 9RA or TBTO based on a luciferase reporter assay (Figure 5.6d). The fold activation level did not exceed the response levels of the empty vector to both chemicals. The two positive controls, HsRXRα and TcRXR-1, strongly reacted to the chemical exposure. Additionally, the adjusted CgPPAR were tested alone and in combination with adjusted CgRXR-1 and exposed to rosiglitazone and PFOA. Clear affiliations of CgPPAR to one of the three isoforms of HsPPAR are not given [Vogeler et al., 2014, submitted]. Thus, agonistic ligands for HsPPARα (PFOA) [Vanden Heuvel et al., 2006] and HsPPARγ (rosiglitazone) [Lehmann et al., 1995] are tested for putative ligand binding to CgPPAR. Both chemicals, however, did not induce a luciferase signal (Supplementary Figure 5.4). The spacer length seems not to interfere with the activation success of the performed assay.

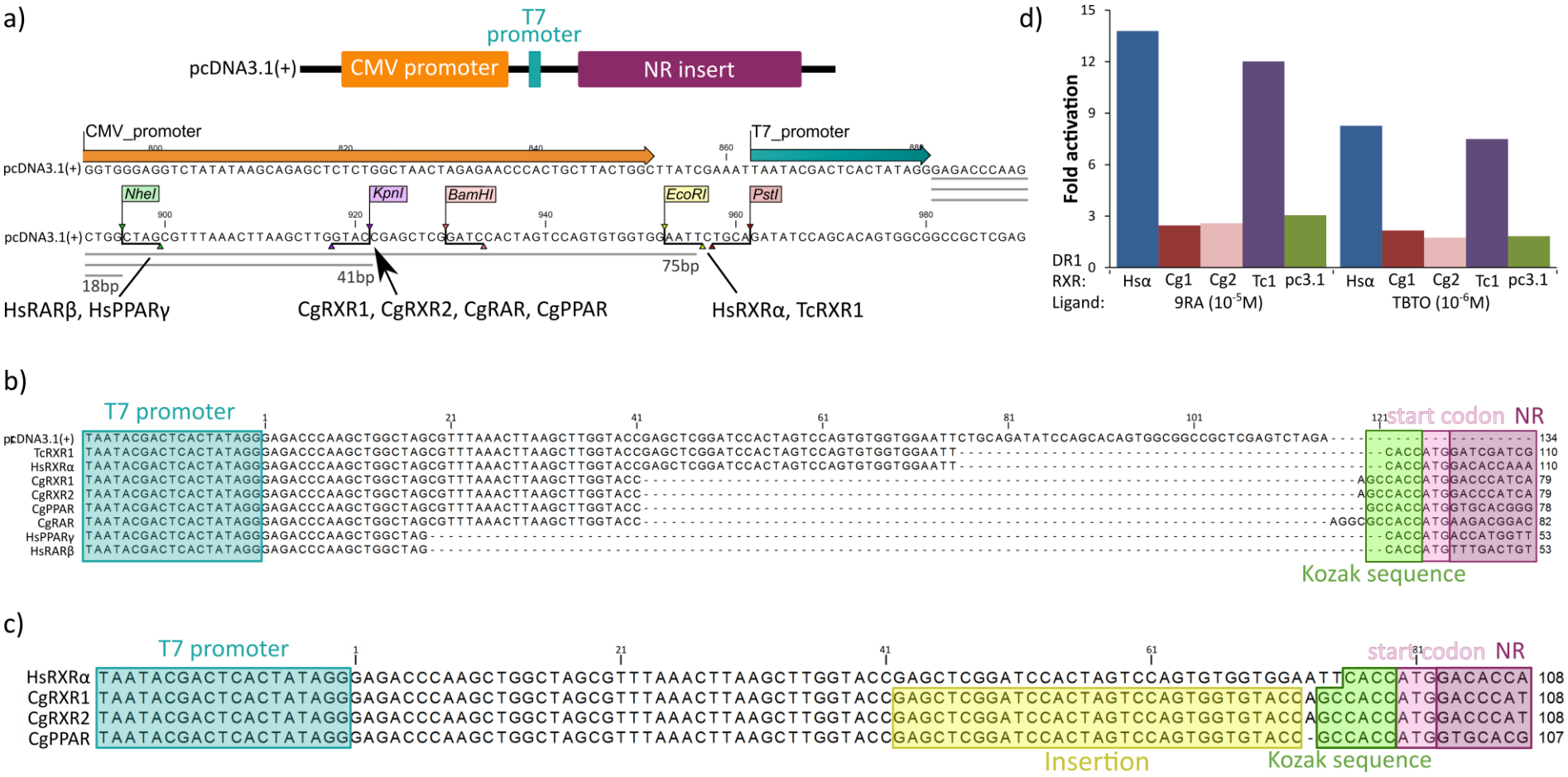



Figure 5.6 legend: see next page

**Figure 5.6: Space length and adjustment of human, *Crassostrea gigas* and *Thais clavigera* expression vectors and transcriptional activity of HsRXRα, CgRXR-1, TcRXR-1 and empty expression vector pcDNA3.1(+) dual-luciferase reporter assays. a)** Position of NR coding sequence insertion downstream of the T7 promoter of the expression vector pcDNA3.1(+), including specific restriction enzyme sites (*NheI; KpnI -BamHI; EcoRI – PstI*). Lines and numbers below sequences downstream of T7 promoter indicate spacer length (bp) for HsRXRα, HsRARβ, HsPPARγ, CgRXR-1, CgRXR-2, CgRAR, CgPPAR and TcRXR-1. **b)** Allignment of spacer sequence between T7 promoter and start codon of all expression vectors fused with a receptor and empty expression vector. **c)** Allignment of expression vector with successful addition of insertion lengthening the spacer region for CgRXR-1, CgRXR-2 and CgPPAR. Orange bar: CMV promoter; Blue box: T7 promoter; green box: Kozak sequence; light purple box: start codon; purple box: beginning of NR coding sequence; yellow box: insertion**. d)** COS-7 cells were transfected with DR1 reporter vector in combination with either human HsRXRα (Hsα), *C. gigas* CgRXR-1 (Cg1), CgRXR-2 (Cg2), *T. clavigera* TcRXR-1 (Tc1) or the empty expression vector (pc3.1). Cells were incubated for 40-42 h either with $10^{-5}$ M 9RA or with $10^{-6}$ M TBTO.

Different cell lines for receptor activation abilities were also tested. In contrast to COS cells (monkey kidney cells) the HEK-293 cells are human kidney cells and are widely used in reporter gene assays. However, using HEK-293 cells, luciferase gene expression were not activated by HsRXRα, adjusted CgRXR-1 and TcRXR-1 exposed to 9RA (Figure 5.7), identifying it as unsuitable cell line for the present vector/chemical composition. The positive control HsRXRα transfected into COS-7 cells produced a luciferase signal after exposure to 9RA. Additional assays with HsPPARγ, adjusted CgPPAR, HsRARβ and CgRAR did not lead to a chemically induced signal (Supplementary Figure 5.5).

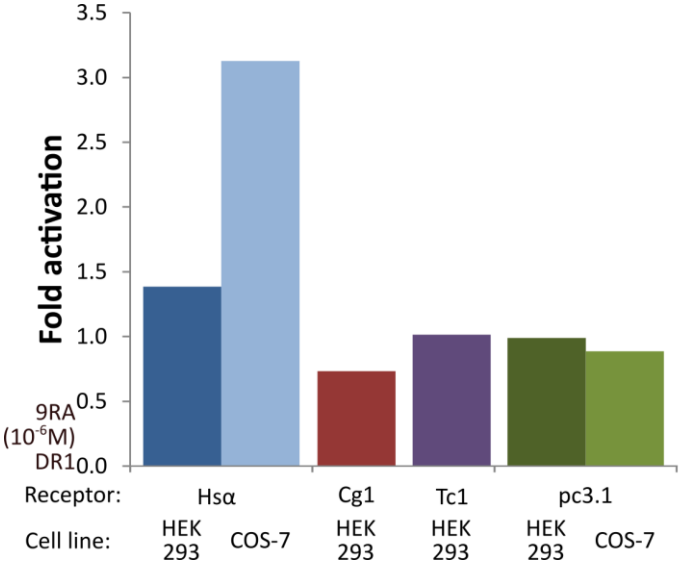



**Figure 5.7: Transcriptional activity of HsRXRα, CgRXR-1, TcRXR-1 and empty expression vector pcDNA3.1(+) dual-luciferase reporter assays.** HEK-293 or COS-7 cells were co-transfected with DR1 reporter vector in combination with either human HsRXRα (Hsα), *Crassostrea gigas* CgRXR-1 (Cg1), *T. clavigera* TcRXR-1 (Tc1) or the empty expression vector (pc3.1). Cells were incubated for 40-42 h with $10^{-6}$ M 9RA.

## 5.4. Discussion

The ability of NRs to regulate gene expression in the presence of a bound ligand opens up a putative pathway for xenobiotic disruption. Given that only some members of the NR superfamily are able to interact with ligands, it is important to understand which NR is able to bind to specific ligands. Here, three recently cloned NRs of the Pacific oyster *C. gigas* - CgRXR including the two isoforms, CgRAR and CgPPAR- were assessed for their ability to bind to specific ligands by using dual-luciferase reporter assays. The measured results raise questions about the accuracy of the assay and potential sources of error are discussed.

### 5.4.1. *Ligand binding ability of oyster receptors*

The results of the dual-luciferase reporter assay suggest that the CgRXR isoforms, CgRXR-1 and CgRXR-2, are not able to interact with the tested chemicals; neither in a dose response to 9RA nor to different chemicals including natural RA isomers and the synthetic compound TBTO. This is unexpected as previous sequence analyses and computational docking experiments of CgRXRs indicated an interaction with 9RA and TBT [Vogeler et al., submitted]. Additionally, gastropod RXRs (*T. clavigera* & *N. lapillus*) are activated by these chemicals [Gutierrez–Mezariegos, 2014; Urushitani et al., 2011]. This supported the theory of a successful activation of oyster RXRs given that the functional domains of NRs are highly conserved in sequence and structure due to their function in fundamental processes of gene expression. Unlike CgRXR, CgRAR showed high fold activation for all natural RAs. This contradicts previous CgRAR analyses and research on gastropod RARs, which did not detect nor suggest interaction with RAs [Gutierrez–Mezariegos, 2014; Urushitani et al., 2013, Vogeler et al., submitted]. However, the measured transactivation activity could also be a possible reaction of the chemical itself to the transfected DR5 reporter vector. Tests on CgPPAR did not indicate interaction with rosiglitazone or PFOA and preliminary evidence on ligand binding ability was sparse.

The performed dual-luciferase reporter assay displayed incongruities with previous research on human receptors, which questions the reliability of the conducted assay. The two positive controls for RXR, the human HsRXRα and the gastropod TcRXR-1, showed transactivation activity comparable to previously reported reporter assays [Urushitani et al., 2011], indicating a correctly operating dual-luciferase assay.

However, the HsRARβ luciferase assay, co-transfected with the DR5 reporter vector, did not fulfil the expectations. The highly potent human RAR synthetic agonist TTNPB did not activate the receptor HsRARβ. Additionally, the HsRXRα receptor shows unusually high activation for a receptor not established to strongly interact with DR5 as homodimer. Co-transfection of HsRXRα and HsRARβ, forming heterodimers highly attracted to DR5 response elements, resulted in a weaker signal compared to HsRXRα and HsRARβ alone. The HsPPARγ is a permissive heterodimer [Kliewer et al., 1992] and rosiglitazone is an antidiabetic drug developed as an HsPPARγ agonist [Lehmann et al., 1995]. However, HsPPARγ could not be activated; neither the receptor alone, nor co-transfected with its heterodimer partner HsRXRα. The permissive heterodimer was also not activated after exposure to TBTO, further indicating a problem with the heterodimer formation. In summary, HsRXRα and TcRXR-1 seem to be the only receptors successfully working in this assay. The other two positive controls, HsRARβ and HsPPARγ, yielded ambiguous, unreliable results, indicating technical problems with the conducted assay. Without correctly working positive controls, interpretation of the Pacific oyster receptor performance becomes very difficult.

### 5.4.2. *Dual-luciferase reporter assays: potential sources of error*

The dual-luciferase reporter assay, although not a complex assay, comprises several individually prepared assay components, which increases the chance for technical errors. In the following paragraph possible sources of technical problems will be described, in the hope to identify causes of the anomalous results of the conducted assay.

The most common mistake for all molecular work is contamination of reagents, chemicals or of other assay components. Contamination was prevented by using aliquots of working stocks for all substances. Furthermore, new working stocks for all vector constructs were obtained from glycerol stocks after half of the luciferase assays had been performed. The integrity of the vector constructs was confirmed by sequencing. No differences between old and new vector sets were detected.

For a successful dual-luciferase reporter assay the three individual vector components (expression vector, reporter vector and control vector) are co-transfected into cells of an immortal cell line. A transfection problem could explain

the absence of a transactivation signal. However, the transfection with FuGENE HD transfection reagent (Promega) appeared to be successful since all assays produced an internal control signal and clear firefly luciferase signals were detected for HsRXRα and TcRXRs. In case of no or low transfection efficiency, the FuGENE HD technical manual (Promega, TM328, revised Feb. 2013) suggests to use only high-quality purified vectors and adjust the FuGENE HD/vector ration. Suboptimal cell growth could also cause variability in transfection efficiencies.

The three vector constructs themselves are susceptible to technical errors. The expression vector contains the recombinant protein sequence for each NR, which will be expressed by the mammalian cell line using the CMV promoter. The integrity of promoter and NR sequences was confirmed. However, differences in spacer length between CMV promoter and the start codon of the receptor gene were detected. The working HsRXRα and TcRXR expression vectors contained a longer spacer sequence compared to all other human and oyster expression vectors. This could cause differences in the recombinant receptor expression efficiency. After adjustment of the three oyster receptors, the results suggested that the spacer length seemed not to interfere with the assay success. CgRXR-1 is highly likely to interact with 9RA due to previous sequence analyses [Vogeler et al., submitted] and receptor conservation to gastropods ligand binding RXRs [Bouton et al., 2005; Carter et al., 2010; Gutierrez–Mezariegos, 2014; Urushitani et al., 2011], but an activation could not be induced. However, it is possible that CgRXR does not interact with a ligand and, as such, the spacer length would produce a successful induction for HsRARβ and HsPPARγ, which both unfortunately could not be adjusted (due to a lack of convenient restriction enzyme sites). To completely exclude spacer length impact, all expression vectors should contain the identical vector composition with differences in the recombinant protein sequence only.

Failure or incorrect expression of the recombinant NR protein would result in an unsuccessful assay since no receptor would be available to expresses the firefly luciferase gene. Unfortunately, the expression of the NR sequences and production of the recombinant protein were not tested for the conducted assay. The positive firefly luciferase signal induction of HsRXRα and TcRXR-1 indicated an efficient NR protein expression. Gene expression of the recombinant protein can be tested with the quantitative RT-PCR (qPCR) method. For this, RNA from transfected cells is extracted and qPCR quantifies the presence of recombinant protein transcripts. This is a simple and quick method, especially when optimised primer pairs are available.

Another viable technique is the immunoblot analysis: the NR sequence is fused with a protein tag to which antibodies are available and sub-cloned into the expression vector (Proteins from transfected cells are extracted and the presence of recombinant proteins is detected with a western blot analysis). Another reason for incorrect expression could be the promoter sequence of the expression vector. The selected cell line might not be compatible with the chosen promoter sequences. The promoter of the expression vector could also cross-interact with the control promoter. The use of a different promoter could prevent this issue. It has been advised to use a strong promoter for the expression vector (e.g. CMV) and a weak for the internal control vector (e.g. TK) [Promega, 2015]. The technical manual of the pcDNA3.1(+) expression vector used (Invitrogen, 28-0104, Nov. 2010 ) also suggested linearizing the expression vector prior of transfection. This could increase the expression success. The expression and control vector used for the conducted assays contain two different promoter sequences (expression v.: CMV; control v.: HSV-TK), but the expression vectors were not linearized prior transfection. In the case of the recombinant protein still not being produced by the cell itself after adjustments, another expression vector backbone would be advisable.

The reporter vector DR1 used in this assay worked satisfactory for HsRXRα and TcRXR-1 and displayed better results than the previously used RXRE reporter vector. The reporter vector for DR1, DR2 and DR5 were all assembled with the same method and the integrity was verified by sequencing. Therefore, it can be suggested that all reporter vectors work correctly. However, a problem stated with some reporter vector backbones is the cryptic transcription activity to transcription factors [Thirunavukkarasu et al., 2000] or to the chemicals used [Giannakis et al., 2003]. The used pGL4 reporter vector backbone (Promega) is advertised to have reduced these cryptic anomalous expressions [Paguio et al., 2005]. Nevertheless, such cryptic transcription activity can explain the signal induction caused by reactions of chemical with cells transfected with empty expression vectors, as well as unexpected signal induction for NR expression vectors co-transfected with reporter vectors (e.g. HsRXRα/DR5). Cryptic signals are difficult to eliminate. Therefore, it is very important to conduct negative controls for empty vectors and non-chemical treatment along with each assay set to detect background signals. As seen in the conducted CgRXR dual-luciferase assays, firefly luciferase signals can look like actual transactivation signals, but sometimes background signals are induced by the tested chemicals. The same can be assumed for the luciferase assay using human and oyster RAR

expression vectors, but could not be proven due to the missing negative expression vector controls.

The *Renilla* luciferase control vector is an internal control, which adjusts for differences in cell viability and assay efficiency. Previous research has shown that the internal control function can be disturbed by different experimental factors, including chemicals, transcription factors (including NRs), other recombinant proteins and cell lines, which all result in variability of *Renilla* luciferase signals [Shifera & Hardin, 2010]. However, none of the reported experimental factor combinations match with conducted assay conditions. *Renilla* luciferase signals vary along the assay, but evidence for effects on the control vector by the co-transfected vectors or chemicals were not detected. Nevertheless, possible control vector disturbance could be tested for by comparing the luciferase results using a different control vector.

Finally, the choice of cell line can also influence the assay results. The cell line should be stable and express all critical components for the tested processes as well as being efficiently transfectable and not susceptible to toxicity of the used reagents. A transactivation signal was successfully induced in the COS-7 cell line co-transfected with the two positive controls HsRXRα and TcRXR-1, confirming cell line suitability. The COS-7 has been previously used for luciferase assays co-transfected with human receptors [Idres et al., 2002; Sun et al., 2014]. The positive induction of the gastropod receptor also suggested COS-7 as suitable cell line for molluscan species, though incompatibility with oyster receptors cannot be unambiguously excluded. Tests with HEK-293, however, did not result in positive transactivation results for any receptor, although HEK-293 has been successfully used as cell line for some of the tested human receptors/chemical combinations [Iwaki et al., 2003; Moise et al., 2009]. This is strengthening the assumption of an unknown issue with the assay.

## 5.5. Conclusion and remarks on further research

The dual-luciferase reporter assay technique was used to investigate the ligand binding ability of the Pacific oyster nuclear receptors CgRXR-1, CgRXR-2, CgRAR and CgPPAR. Unfortunately, the conducted assays did not produce reliable results. Serious concerns regarding the correct functioning of the conducted assays were raised due to incongruities in the performance of well-characterised human receptor homologs used as positive controls. Therefore, significant statements on the oyster receptor ligand binding abilities could, unfortunately, not be made.

Nevertheless, the dual-luciferase reporter assay has been established as a reliable method to test nuclear receptor activation abilities through putative ligands. For further research on oyster receptor binding abilities I suggest to perform a new set of dual-luciferase reporter assays, testing gene expression activation, ligand interaction and heterodimerisation in separate assays. Performing negative controls for expression and reporter vectors as well as non-chemical treatments along each assay is highly advised. Additionally, successful recombinant receptor protein expression should be verified by immunoblots.

Nuclear receptor binding ability to specific response elements and subsequent activation of gene expression can be tested by generating chimeric receptor constructs. The receptor DNA binding domain (DBD) is fused with a ligand binding domain (LBD) of a known receptor and co-transfected into a cell line with a reporter vector containing a specific response element. Transactivation ability can then be measured after exposure to a ligand of the known LBD. Ligand binding ability can be tested by a similar chimeric construct technique. The LDB of the receptor will be fused with a known DBD, most commonly a Gal4 DBD, and co-transfected with the corresponding reporter vector (e.g. possessing a Gal4 binding site). If firefly luciferase signal is measured after exposure to putative ligands, receptor:ligand interaction is verified. A mammalian two-hybrid assay can be used to test for heterodimerisation. For this specific form of a dual-luciferase assay, cells are co-transfected with a chimeric construct of a Gal4DBD fused with the LBD of one of the heterodimer partners and a second chimeric construct containing the second heterodimer LBD fused with different known DBD. Transcription activity is enhanced due to successful heterodimerisation and produces a much stronger luciferase signal in comparison to the negative controls.

Finally, an additional electrophoretic mobility shift assay is suggested to support DNA binding and heterodimer formation of the receptors [Garner & Revzin, 1981; Hellman & Fried, 2007] This assay uses the full receptor sequence. The receptor proteins are mixed with short double strained oligonucleotides containing the response element of interest. The oligonucleotides are labelled with either a radioactive, fluorescent or biotin label for detection. Binding of receptors to oligonucleotides is analysed by electrophoretic separation using polyacrylamide gels. Free oligonucleotides and receptors as well as putative receptors:oligonucleotide complexes are separated by size and charge, identifying whether a receptor binds to a specific response element. This can also be used to test for dimerisation as receptor partners vary in their sizes/charges, producing different sized/charged receptor:oligonucleotide complexes.

## Discussion

The effects of xenobiotics on marine wildlife can be observed around the globe [Bergman et al., 2013; Diamanti-Kandarakis et al., 2009], yet the understanding of these effects and the precise mode of action of many xenobiotics is still not resolved. Nuclear receptors (NRs) have been proposed as one possible target of xenobiotics, enabling environmental pollutants to evoke their harmful effects on the health of animals and humans. The ability of some xenobiotics to mimic ligands of a specific NR or to induce an incorrect gene expression signal for a set time point, can lead to disruption of reproduction, development and other physiological processes in animals [McLachlan, 2001]. However, knowledge on which xenobiotic compounds interact with NRs is sparse. Additionally, due to the highly complex nature of NRs, including numerous different modes of actions, this information alone may not even be enough to understand the disruptive effects. Some receptors function as monomer or as homo- or heterodimer, some bind a variety of ligands and others do not require ligand binding at all [Germain et al., 2006]. Certain receptors can be activated by only one of the dimer partners, others by both. They work in combination with a vast range of co-regulators and each receptor regulates the expression of a variety of genes. These processes are not even close to being fully understood in any species.
Research on xenobiotic effects and NRs in invertebrate species is very limited compared to research in vertebrates, particularly in humans, although invertebrates greatly outnumber vertebrates. Nevertheless, for all wildlife one of the best documented cases of xenobiotic effects can be found in marine invertebrates. The disruptive effects of tributyltin (TBT) on molluscan species have been extensively reported over the last forty years, but the exact mechanism has still not been fully resolved. Hypotheses regarding the involvement of NRs, specifically the retinoid X receptor (RXR) and the peroxisome proliferator-activated receptor (PPAR) [Grün et al., 2006; Harada et al., 2015; Horiguchi et al., 2007; Kanayama et al., 2005; le Maire et al., 2009; Nishikawa et al., 2004; Pascoal et al., 2013], have been established. However, the main obstacle for research on xenobiotic effects on invertebrates is the lack of fundamental knowledge of their endocrinology and especially on NRs (e.g. presence, expression, mode of action and which genes are regulated).

Over the course of my Ph.D. research, I investigated NRs in the Pacific oyster, *Crassostrea gigas,* which has been established as a model organism, principally due

to its genome sequencing [Zhang et al., 2012]. Forty-three oyster NR homologs to human (Bilateria Clade: Deuterostomia; vertebrate representative), and *Drosophila melanogaster* and *Caenorhabditis elegans* receptors (Bilateria Clade: Protostomia; Superphylum: *Ecdysozoa*; invertebrate representative) as well as to other molluscan receptors were identified [Vogeler et al., 2014]. It is noticeable that the Pacific oyster (Bilateria Clade: Protostomia; Superphylum: *Lophotrochozoa*) possesses NR homologs to human receptors (such as the PPAR, retinoic acid receptor RAR, estrogen receptor ER), which are not present in *Drosophila*, although *Drosophila* and oysters belong to sister classes in the protostomes clade. In contrast, the oyster also possesses receptors, which are only present in protostomes (xenobiotic-sensing receptor group NR1J) or thought to be unique for ecdysozoans (ecdysone receptor EcR). Hence, NR homology is not limited to one specific class of species. In addition, this knowledge can be used to further investigate the distinct function of oyster receptors and provides information on putative xenobiotic disruption. For example, insects, such as *Drosophila* species, possess sets of NRs, which are involved in the regulation of early development, moulting and metamorphosis. EcR regulates the expression of NRs genes after an ecdysteroid hormone signal, which further regulates genes important for the transition from early larval to late larval or to adult life stages [King-Jones et al., 2006; King-Jones & Thummel, 2005]. The Pacific oyster possesses NR homologs to most of these insect NRs (e.g. CgNR1H-EcR, CgNR1D-E75, CgNR1E-E78, CgNR1F-DHR3, CgNR5A-FTZ-F1, CgNR5B-DHR39; see Figure 2.2) and given that oysters also undertake metamorphosis to transform from a larval to an adult life stage, it can be speculated that these oyster NRs perform similar functions. In general, nuclear receptors are highly conserved in function due to their fundamental role in gene expression, and homologs share similarities in their mode of action and function between species even if they are not closely related. This also includes their potential for xenobiotic disruption.

The absence of specific NR groups can also be used to investigate and exclude xenobiotics effects via NR interactions. The Pacific oyster does not possess any members of the NR3C subfamily group (androgen receptor, glucocorticoid receptor, mineralocorticoid receptor, progesterone receptor), which supports the theory that the expansion of steroid receptors is deuterostome specific [Bertrand et al., 2004; Eick et al., 2011; Markov & Laudet, 2011; Saez et al., 2010; Thomson et al., 2009]. Vertebrate sex steroids such as testosterone and progesterone are able to interact with the receptors of the NR3C group and environmental pollution with sex steroids

or sex-steroid-like compounds has caused disastrous effects, for example, in many fish and other vertebrate species [reviewed in Bergman et al., 2013]. However, due to the absence of NR3C, exogenous vertebrate sex steroids might not be a hazard for oysters at all, at least not via the NR3C receptors. This is also the case for NRs, which have lost their ability to interact with ligands. Unlike vertebrate estrogen receptors (ERs), molluscan ERs have lost their ability to bind to estrogens and work as constitutive receptors [Keay et al., 2006; Matsumoto et al., 2007; Thornton et al., 2003]. Disruptive effects of estrogens via the ER are therefore unlikely. However, effects of vertebrate sex steroids on molluscan development and sexual differentiation have been extensively reported, although a critical evaluation of these studies questioned the reliability of these results [Kauer et al., 2016; Scott, 2012, 2013]. Thus, potential xenobiotic effects of vertebrate sex steroids might be caused through a different pathway, either another NRs or independent of NRs.

*C. gigas* also holds NR homologs, which may be unique to the Pacific oyster [Vogeler et al., 2014]. The NR1P group consists of eleven NRs, for which no direct homologs could be identified in human, *D. melanogaster* or *C. elegans*. Whether this novel group is unique to oysters or is also present in other molluscan species is currently unknown. An identified *Mytilus galloprovincialis* receptor MgNR1G clusters together with the novel NR1P group, suggesting at least a bivalve specific NR1 subgroup expansion. Apart from *C. gigas*, large sets of NRs have only been revealed for two other molluscan species so far [Kaur et al., 2015]. The freshwater snail *Biomphalaria glabrata* showed three unclassified receptors, located in the NR1 subfamily, which could assign to the novel subfamily group NR1P. Further research is needed to elaborate on this question. Proposing potential functions of the NR1P members based on homology is very difficult, since no direct characterised homologs to this group could be identified. Members of the NR subfamily groups in insects (NR1D-NR1F), which share the same origin as NR1P, are known to regulate genes for organ development, moulting and metamorphosis [King-Jones & Thummel, 2005]. NR1P receptors might fulfil similar functions. Expression profiles of six of the eleven NR1P receptors showed that most of them are expressed during life stages at which organs are developed (6-12 hpf; see Figure 3.2) [Vogeler et al., accepted]. It is also possible that this group is a molluscan specific adaptation, for example, to the loss of estrogen binding of the ER ortholog. Molluscs are a bilaterian class, which possess an ER ortholog that lost its ligand binding ability [Baker, 2011]. NR1P receptors could substitute for this loss by interacting with estrogens or other potential sex steroids.

Detailed studies on the NR1P subfamily group are required to determine its underlying mechanisms.

Nuclear receptors vary in their expression over time, adapting to developmental and environmental cues. During embryo and larval development oyster NRs appear to be expressed in waves rather than being expressed constantly, indicating distinct functions for each of them at specific life stages. Three key developmental stages were identified in the Pacific oyster: embryogenesis, mid development from gastrulation to trochophore larva, and late larval development prior to metamorphosis [Vogeler et al., accepted]. Early development life stage (up to 3 – 6 hpf) is mainly driven by synchronous cell division cycles. Mid development (6-12 hpf), including gastrulation, organogenesis and shell development (trochophore larval stage), on the other hand, is characterised by cell differentiation. The anterior-posterior axis is defined and organs develop. The late development stage (27 hpf- 15 d) is the life stage of a larva, in which it is free swimming, feeds and grows in size until in it settles onto the bottom and metamorphosis takes place. Each of these developmental life stages requires the expression of unique sets of genes, which might be regulated by specific NRs.

The information of when NRs are highly expressed can be used for assumptions regarding their potential function in oyster development. This can be further defined using the information on function of highly homologous NRs in other species. For example, receptors, which are expressed at 6 hpf or 12 hpf, might fulfil an important function in gastrulation (approx. at 6 hpf) or initial shell development (approx. at 12 hpf). The Pacific oyster highly expresses RAR and RXR homologs during these two stages. Vertebrate homologs of RAR and RXR are involved in a wide range of developmental processes. They regulate the anterior-posterior patterning during gastrulation, are responsible for nervous system and sensory organ development, as well as the development of limbs, heart, lung and digestive/urogenital tracts [Dolle, 2009; Escriva et al., 2002]. Hence, assumptions can be made that oyster RAR and RXR are involved in similar developmental processes. In relation to this, the human TLX receptor, an orphan receptor, is involved in eye development and regulates the expression of RAR [Kobayashi et al., 2000]. An oyster TLX homolog is also highly expressed during trochophore larval stage (12 hpf) and might fulfil similar regulation mechanism of oyster RAR during development.

Furthermore, life stages which are particularly sensible to xenobiotics can be identified by NR expression patterns. If a NR is known to be affected by specific xenobiotics, the life stages at which this receptor is expressed are more likely to be disrupted. This can aid the environmental conservation and the aquaculture industry to generate risk assessments. The EcR in *Arthropoda* species regulates life changing events such as metamorphosis and moulting by receiving ecdysone signals, and it is used as a target for many insecticides to prevent growth and successful metamorphosis/moulting of unwanted bugs in agriculture [Fahrbach et al., 2012; King-Jones & Thummel., 2005; Riddiford et al., 2010]. The oyster EcR homolog is highly expressed in the late larval stage (pediveliger; 15 d), which opens up speculations for a putative function in metamorphosis and settlement for this receptor [Vogeler et al., accepted]. This also raises the question of whether the oyster EcR can be disrupted by xenobiotics, such as pesticides, introduced to the water by land runoff or sewage. In addition, this information might be used in the aquaculture industry to identify specific compounds, which could prevent/induce the metamorphosis and settlement and benefit the oyster production.

Nevertheless, expression patterns among different life stages can only provide limited information about putative function of NRs and further research is needed. Molecular techniques such as In-situ hybridisation (ISH) or immunohistochemistry (IHC) could identify single NRs in specific tissues, providing further information on potential functions. If a receptor is present in a tissue which, for example, is developing into a specific organ, it can be assumed that this receptor is involved in the development of this particular organ. Tissue specific assessment, in general, is a very powerful tool to investigate NR functions. The observed differences between male and female NR expressions imply sex-specific NRs presence, possibly involved in the development of sex-separated reproductive organs such as gonads. However, distinct conclusions could not be made in this study. Whole animals were analysed and any sex-dependent tissue specific NR expression might have been lost. Here, tissue specific assessment using qPCR, ISH or IHC could uncover sex-dependent NR function.

NR expression analysis also showed differences between developmental and both adult groups, indicating differing NR demand between developmental and adult life stages. Adult individuals do not undergo massive transformation in body structure, except during reproduction. As sessile species their life is strongly influence by their surroundings. Interactions with exogenous compounds, including environmental

pollutants, cannot be avoided and adequate responses need to be executed. Indications for those responses can be found in the NR expression profile. The two adult groups displayed an increase in expression of members of the NR1J group. This NR subfamily group is involved in xenobiotic-sensing processes, defines mechanisms in many protostome species against natural and anthropogenic environmental stressors [King-jones & Tummel, 2005; Prakash et al., 2015; Richter & Fiddler, 2014]. Recent research suggests that this defence mechanism is also conserved in bivalves [Cruzeiro et al., 2016].

Knowledge on the presence of NRs at specific life stages and the impact of chemicals on oysters, as well as functional studies on receptor ligand interacting abilities can be combined to understand the observed effects of environmental pollution. My research on the mode of action of three NRs, CgRXR, CgRAR and CgPPAR, provides further information on this topic [Vogeler et al., submitted]. The sequence analysis of all three receptors suggests that these receptors are functionally active including the ability to interact with DNA, form dimers and regulate the expression of genes. Although the dual-luciferase reporter assay did not deliver reliable results (Chapter 5) on whether these receptors are able to bind compounds *in vitro*, computer based 3D modelling of these receptors highly indicates interaction with specific ligands. These results are further supported by findings in gastropod and vertebrate species. The CgRXR seems to be able to be activated by 9-*cis* retinoic acid (9RA), which has also been reported previously for gastropod species [Bouton et al., 2005; Carter et al., 2010; Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2011]. This makes 9RA a putatively harmful compound for oysters when present at a promotive time point. Additionally, partial isomerization mechanisms for retinoid acid (RA) isomerisation have been identified in several gastropod species [Gesto et al., 2013], suggesting a potential for the conversion of all-*trans* RA (ATRA) to 9RA and 13-*cis* RA. This opens up possible indirect disruptive pathway via RXRs for other RAs. This could explain the observed effects in oyster embryos after exposure to ATRA, which resulted in soft tissue malformations, shell deformation and high numbers of arrested trochophore larvae [Vogeler et al., submitted]. Similar malformations and effects have been reported previously in gastropod snail embryos [Carter et al., 2015; Creton et al., 1993]. The CgRXR is highly expressed at the trochophore and D-shelled larval stage [Vogeler et al., in review], but analyses did not provide evidence for a direct interaction with ATRA. The observed effects could

therefore be the result of an indirect effect of ATRA, isomerized to 9RA. Conclusively, the ligand-binding ability of molluscan RXRs and the observed effects of retinoids on molluscan embryos indicate that ATRA, 9RA and other retinoids function as teratogens for molluscan species. Retinoids are already well known teratogens for vertebrate species [Soprano & Soprano, 1995], but information on invertebrates is still limited.

CgRXR is highly likely to interact with other receptors as a heterodimer partner [Vogeler et al., submitted], which is a common feature for most characterised RXR orthologs in invertebrate and vertebrate species. Heterodimerisation between gastropod RXRs and RARs has been identified [Gutierrez-Mazariegos et al., 2014; Urushitani et al., 2011]. If the NR feature of permissive RXR heterodimers in vertebrates is also conserved in molluscan species, RXR could be a receptor through which many xenobiotics can procure their disruptive effects. This is presumed in the case of TBT and its disruptive effect on molluscan species. Exposure experiments of oyster embryos to TBT oxide (TBTO) have caused malformations of soft tissue, shell deformations and arrested larvae [Vogeler et al., submitted]. TBT can activate the vertebrate RXR/PPAR heterodimer via RXR alone or possibly through both partners [Harada et al., 2015; Grün et al., 2006; Kanayama et al., 2005; le Maire et al., 2009]. A similar pathway for TBT disruption is indicated by computational docking results for CgRXR and CgPPAR [Vogeler et al., submitted]. TBT is likely to activate CgRXR, a theory which is supported by the successful *in vitro* activation of the gastropod *Thais clavigera* RXR by TBT [Urushitani et al., 2011]. Implicative evidence for TBT binding has also been found for CgPPAR [Vogeler et al., submitted]. The theory of RXR/PPAR involvement in the TBT-induced disruption is further strengthened by the high expression of CgPPAR in adults [Vogeler et al., accepted].

Knowledge on the binding abilities of NRs also helps to exclude pathways. Similar to the gastropod *Nucella lapillus* RAR [Gutierrez-Mazariegos et al., 2014], CgRAR does not seem to be activated by RAs (ATRA, 9RA, 13RA, TTNPB). Hence, a direct effect of RAs on CgRAR is highly unlikely and other pathways for the observed disruptive effects should be considered; for example another NR, an indirect effect caused by isomerization of the compound or a pathway unrelated to NRs.

It must be noted, however, that using 3D modelling and computational docking experiments to test receptor ligand binding abilities comes with many limitations. First, template 3D structures of homologs for creating 3D models need to be available. Those 3D structures are initially obtained by crystallography and are

provided for download on protein data base websites such as Protein Data Bank (www.rcsb.org). However, these models must hold the 3D structure bound to the ligand of interest to provide a suitable template, which minimizes the amount of available receptor templates. As seen in Chapter 4 of this thesis, assumption of a putative binding of CgPPAR to PFOA could not be assessed due to a missing human template. The human 3D structure for PPARα is available, but not bound to the ligand PFOA. Secondly, a high sequence homology is required to secure the accuracy of the created 3D model. If sequences are too diverse, the modelled 3D structure might be incorrect and the performed computational docking loses its reliability. CgPPAR and human PPAR display only weak sequence homology of their ligand binding domains and the created 3D structure of CgPPAR is therefore not reliable. Although the docking experiment suggested an inability of rosiglitazone to activate CgPPAR transcription activity and exposure of oyster embryos to this chemical did not cause any visible effect, these results need to be interpreted with caution.

Nevertheless, the mode of action of many disruptive chemicals in invertebrates is far from being fully comprehended, and further understanding of chemical impacts on marine species is urgently needed. Although my Ph.D. research contributes to the understanding of fundamental processes of gene expression and endocrinology in the Pacific oyster and sheds light on how xenobiotics might interact with NRs, many unanswered questions still remain. Further experiments, for example adjusted dual-luciferase reporter assays, could solve the question whether oyster NRs do or do not bind specific chemicals, as suggested by the computational docking experiments. Additionally, different ligands could be tested, which show a high potential as EDCs, such as bisphenol A for which binding to various NRs in different species has been verified [Canesi & Fabbri, 2015]. If know-how and equipment is available, X-ray crystallography of a NR (with and without a bound ligand) would resolve the receptors' protein structure and how it binds to ligands as well as the subsequently induced conformational changes. Electrophoretic mobility shift assays could identify the response element sequence for each NR monomer or dimer. Furthermore, different molecular techniques and bioinformatics tools could be used to further investigate the function of NRs. Immunohistochemistry experiments can discover the presence of NRs in specific body parts, for example during embryo development, metamorphosis or in adult life stages. Enhanced NR expression could suggest

putative functions during the development of this body part. Using the published oyster genome [Zhang et al., 2012], bioinformatic tools could identify putative NR target genes based on the sequence information of NR response elements. Temporary gene knockdown of NRs or NR target genes, e.g. using the RNA interference technique, could further elucidate the search for NR functions.

Additional research on NRs could also aid the aquaculture industry and environmental conservation. Some NRs, such as the EcR, have high potential to be active during sensible time points (e.g. metamorphosis) and to interact with chemicals (e.g. pesticides, steroids). Information on these processes and interrelations can be used to generate risk assessments or increase the productivity of oyster farms. For example, chemicals could be identified, which synchronise metamorphosis and settlement in an oyster population; an approach highly desired in the aquaculture industry to control oyster spat production. Further exposure experiments could assess at which concentrations chemicals affect oyster populations and, in case of risk assessments, if these concentrations are environmentally relevant.

The Pacific oyster is increasingly used as a model organism for aquatic invertebrates mainly due to its many advantages as laboratory animal, increasing molecular and genomic information and its high economical interest. The identification of NR homologs further promotes its use as model organism. In general, NR functions are conserved across species and therefore oyster NR homologs to invertebrate- and vertebrate-like NRs are suitable for the investigation of functional relationships of individual receptors. Additionally, the potential of some oyster NRs to interact with xenobiotics can help to investigate pathways through which environmental pollutants trigger their disastrous effects in molluscs and other aquatic invertebrates. As shown in this thesis oyster embryos can also be used to assess the effects of xenobiotics on oyster development. The information on disruptive chemical effects, which can be gained through those investigations, can ultimately help to develop rules and laws to protect marine invertebrates and conserve our environment. Oysters might also be used as surrogates for vertebrate systems. However, it must be assured that the oyster system operates comparable to the respective vertebrates, which might not always be the case. For example, oysters are no suitable as model organism for vertebrate sex steroid pathways given that they are lacking important functional NRs such as members of the NR3C group or ligand-binding ER homologs. Nevertheless, understanding of oyster endocrine systems and xenobiotic disruption pathways is still

increasing and additional applications for oysters as model species for vertebrate systems might be identified in the future.

In conclusion, the Pacific oyster possesses a large variety of NR homologs, for which different expression patters during developmental/adult life stages and their high sequence homology to other species suggest distinct functions in developmental and physiological processes. For a small subset of NRs, evidence for receptor:DNA and receptor:receptor interaction, as well as the ability of receptors to bind to ligands, has been found. In addition, exposure experiments of oyster embryos to natural and synthetic compounds demonstrated sensitivity during embryo development and the risk at which marine invertebrates are placed in polluted environments. This new knowledge on NRs in the Pacific oyster helps to further understand the mechanisms of environmental pollutants on marine invertebrates and aids in the assessment and prediction of the effects of anthropogenic contaminants.

## Supplementary files

Note: The complete supplementary file data set is provided on the CD enclosed with this thesis. All supplementary files captions are presented in this section including selected figures and tables. Large datasets and files are only available as digital version on the enclosed CD and in few cases online.

### Supplementary Table 2.1

Available on CD: Supplementary_table_2_1.doc
or Online (Assessed May 2016):
http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-369

**Supplementary Table 2.1: Nuclear receptor amino acid sequence GenBank accession numbers for *Crassostrea gigas*, *Homo sapiens, Drosophila melanogaster, Caenorhabditis elegans* and mollusc species.**

### Supplementary Table 2.2

Available on CD: Supplementary_table_2_2.doc
or Online (Assessed May 2016):
http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-369

**Supplementary Table 2.2: Table of primer sequences including amplicon length (bp) and parts of nuclear receptor domains sequenced**. A/B= N-terminal domain, C= DNA binding domain, D= hinge domain, E= ligand binding domain, F= C-terminal domain.

### Supplementary Figure 2.1

Available on CD: Supplementary_figure_2_1.doc
or Online (Assessed May 2016):
http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-369

**Supplementary Figure 2.1: Phylogenetic tree using only DBD of NR alignment conducted by a Maximum Likelihood (ML) and Bayesian Inference analyses.** ML bootstrap support values (percentage of 1000 BS) and Bayesian posterior probabilities (PPs) are provided above the nodes separated by slash. Star indicates the node obtained from the Bayesian Inference analysis, which was different from that obtained by ML method. *Crassostrea gigas* NRs highlighted in red. Ac: *Aplysia californica,* Bg: *Biomphalaria glabrata,* Ce: *Caenorhabditis elegans*, Cg: *C. gigas,* Dm: *Drosophila melanogaster*, Hs: *Homo sapiens*, Lym: *Lymnaea stagnalis*, Mc: *Marisa cornuarietis*, Me: *Mytilus edulis*, Mg: *Mytilus galloprovincialis*, Nl: *Nucella lapillus,* Ov: *Octopus vulgaris,* Sm: *Schistosoma mansoni,* Tc: *Thais clavigera.*

## Supplementary Figure 2.2

Available on CD: Supplementary_figure_2_2.doc

or Online (Assessed May 2016):

http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-369

**Supplementary Figure 2.2: Phylogenetic tree using only a portion of LBD of NR alignment conducted by a Maximum Likelihood (ML) and Bayesian Inference analyses.** ML bootstrap support values (percentage of 1000 BS) and Bayesian posterior probabilities (PPs) are provided above the nodes separated by slash. Star indicates the node obtained from the Bayesian Inference analysis, which was different from that obtained by ML method. *Crassostrea gigas* NRs highlighted in red. Ac: *Aplysia californica,* Bg: *Biomphalaria glabrata,* Ce: *Caenorhabditis elegans*, Cg: *C. gigas,* Dm: *Drosophila melanogaster*, Hs: *Homo sapiens*, Lym: *Lymnaea stagnalis*, Mc: *Marisa cornuarietis*, Me: *Mytilus edulis*, Mg: *Mytilus galloprovincialis*, Nl: *Nucella lapillus,* Ov: *Octopus vulgaris,* Sm: *Schistosoma mansoni,* Tc: *Thais clavigera.*

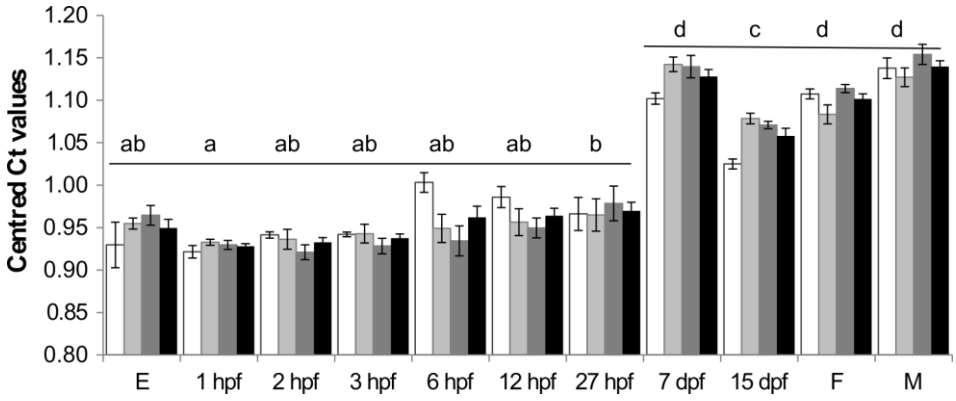
## Supplementary Table 3.1

Available on CD: Supplementary_table_3_1.doc

**Supplementary Table 3.1: Primer information used in qPCR analysis:** Table of forward and reverse primer sequences for 34 *Crassostrea gigas* nuclear receptors and three reference genes including amplicon length (bp), annealing temperature, final primer concentration and $MgCl_2$ concentration, primer efficiency, and accession number of receptor gene.

## Supplementary Figure 3.1

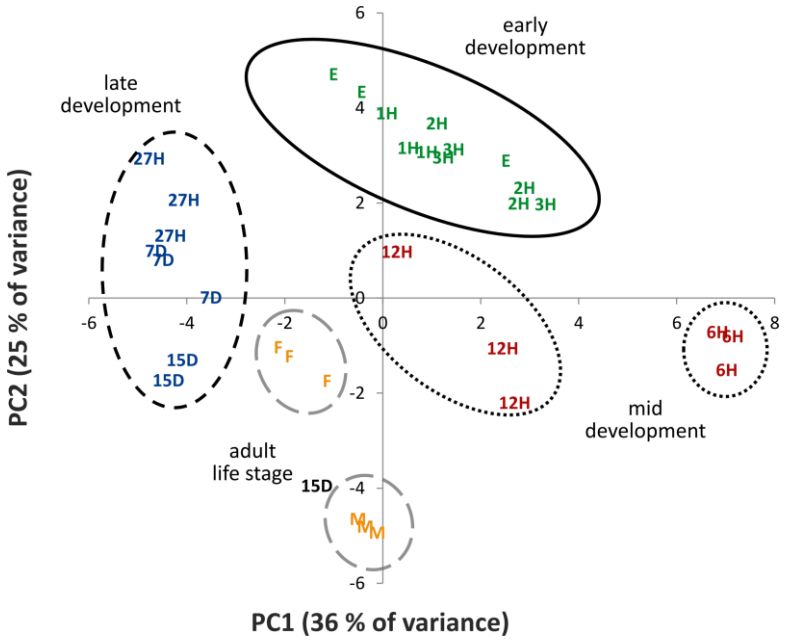
Available on CD: Supplementary_figure_3_1.doc




**Supplementary Figure 3.1: Gene expression of selected reference genes in *Crassostrea gigas* among different life stages.** Three housekeeping genes selected as reference genes (centred Ct values): elongation factor-1α: white bars; ribosomal protein S18: light grey bars; ribosomal protein L7: dark grey bars; mean of all reference genes including significantly different groups: black bars and letters. hpf: hour post fertilisation. dpf: days post fertilisation. E: unfertilised eggs. F: female. M: male.

## Supplementary Figure 3.2

Available on CD: Supplementary_figure_3_2.doc




**Supplementary Figure 3.2: Principle component analysis (PCA) of 31 *Crassostrea gigas* nuclear receptor gene expression data across developmental and adult life stages.** Scatterplot of the first two PCA components of developmental stages including adult life stages indicating separation of all life stages based upon the variance observed in the expression levels of 31 of the 34 nuclear receptor genes. Principal component 1 (PC1) and 2 (PC2) explain 36 % and 25 % of variance, respectively Circles around measurements and colours of measurements representing distinct clustering for all life stages or nuclear receptors: early development (green numbers+letters, black solid line), mid development (red numbers+letters, black dotted lines), late development (blue numbers+letters, black dashed line), adult life stages (yellow letters, grey dashed lines). h: hour post fertilisation; d: days post fertilisation; E: unfertilised eggs; F: female; M: male.

## Supplementary Table 4.1

Available on CD: Supplementary_table_4_1.doc

**Supplementary Table 4.1: Primers for sequencing full length RNA sequences of CgRXR, CgRAR and CgPPAR.**

## Supplementary File 4.1

Available on CD: Supplementary_file_4_1.fa

**Supplementary File 4.1: Full protein sequences of CgRXR-1, CgRXR-2, CgRAR and CgPPAR.**

## Supplementary File 4.2

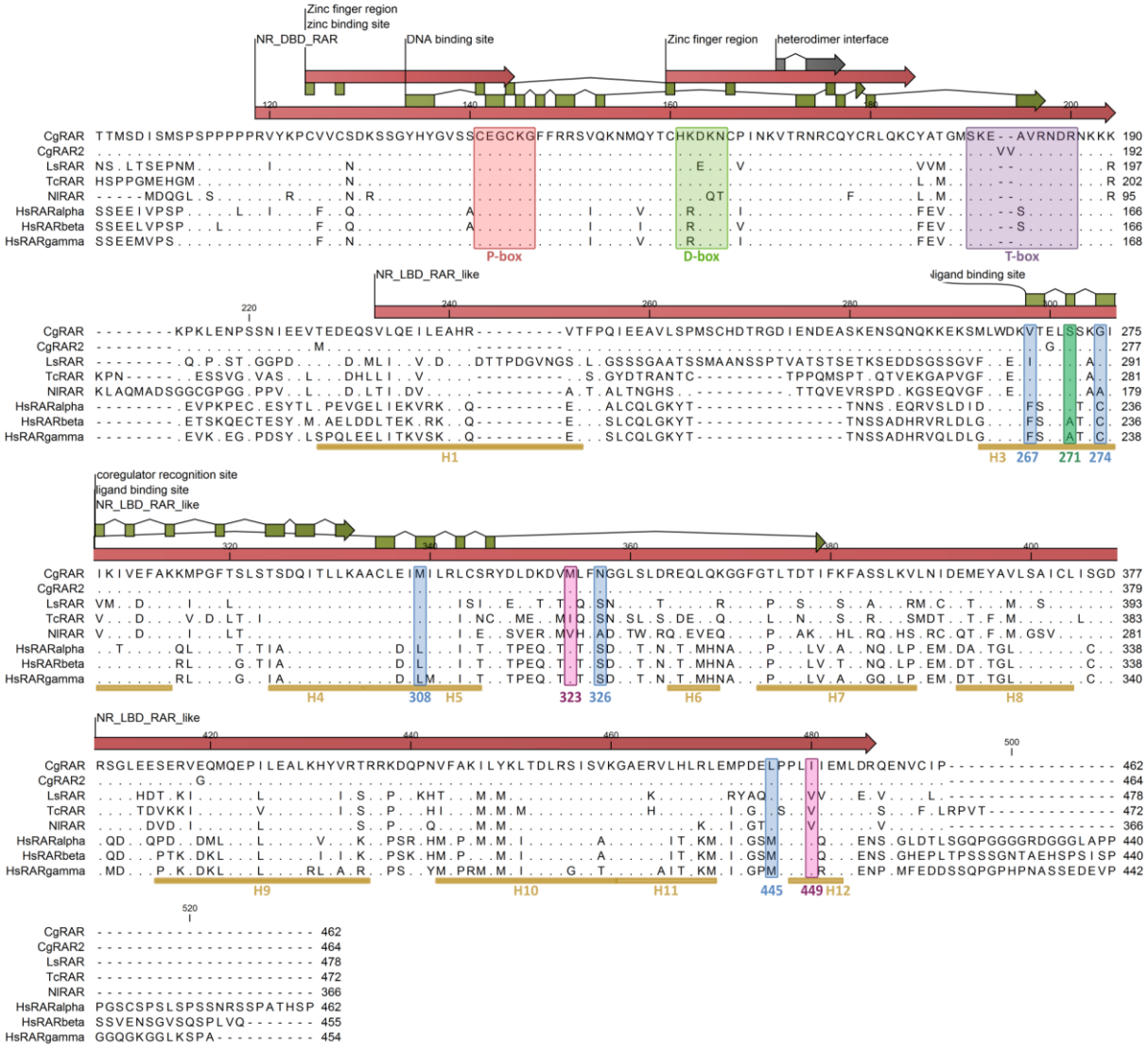
Available on CD: Supplementary_file_4_2.fst

**Supplementary File 4.2: Amino acid alignment of DBD and LBD of nuclear receptor homologs RXR, RAR and PPAR across various phyla**. Name code: species code + receptor homolog_genebank ID.

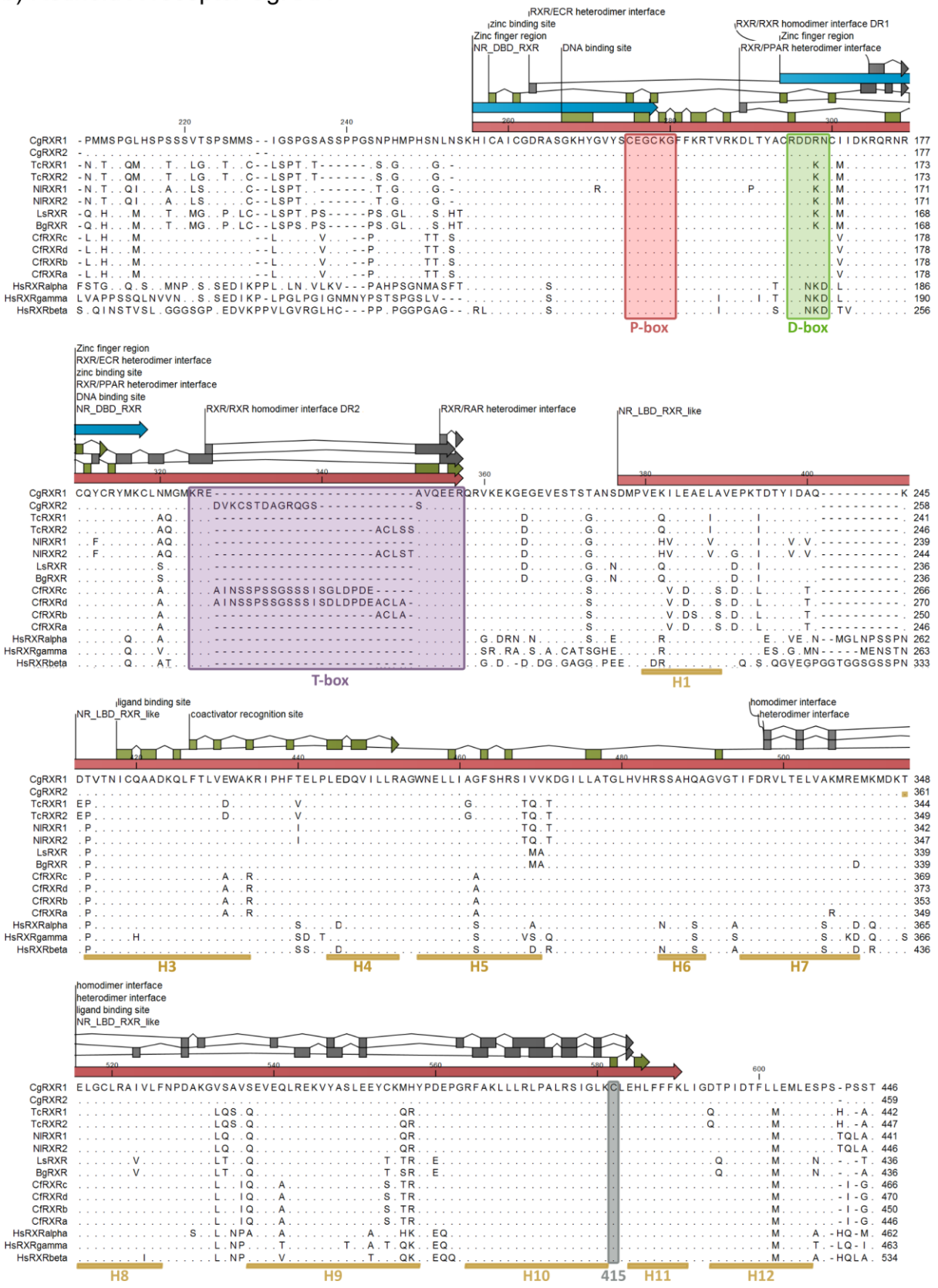
## Supplementary Figure 4.1

Available on CD: Supplementary_figure_4_1.doc

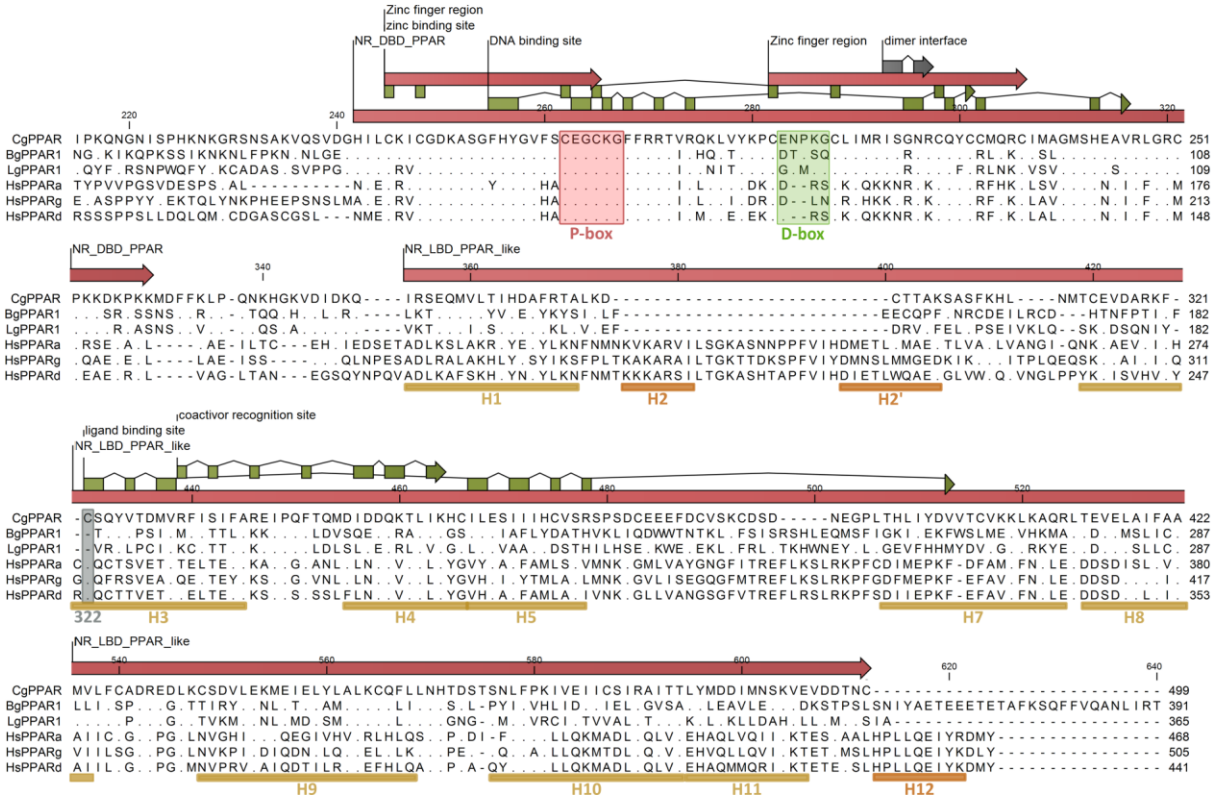
a) Retinoic acid receptor CgRAR



## b) Retinoid X receptor CgRXR



c) Peroxisome proliferator-activated receptor CgPPAR




**Supplementary Figure 4.1: Sequence alignment of CgRXR, CgRAR and CgPPAR and their isoforms to molluscan and human homologs including sequence features and motifs.** A/B domain partial; H1-12 (yellow): α-helixes of LBD; H2, H2', H12 (orange): non-CgPPAR specific helixes; red: P-box; green: D-box; purple: T-box; grey: TBT interacting cysteine; blue: CgRAR & NlRAR shared residues; dark green: CgRAR specific residues; violet: NlRAR specific residues.

## Supplementary Table 4.2

Available on CD: Supplementary_table_4_2.doc

**Supplementary Table 4.2: Calculated binding affinity values (kcal/mol) by computational docking of ligands to the created 3D models of human, oyster and gastropod retinoid X receptor (RXR) and peroxisome proliferator-activated receptor (PPAR).** pdb template: pdb ID providing the template for 3D modelling; RA: retinoic acid; ATRA: all-*trans* RA; TBT: tributyltin; Cg: *Crassostrea gigas*, Hs: *Homo sapiens*; Nl: *Nucella lapillus*; Tc: *Thais clavigera*

| Ligand | Receptors | | | | | | pdb template |
|---|---|---|---|---|---|---|---|
| | HsRXRα | CgRXR | TcRXR | NlRXR | HsPPARγ | CgPPAR | |
| 9-*cis* RA | -10.6 | -10.6 | -10.6 | -10.5 | | | 1FBY |
| TBT | -5.6 | -5.4 | -5.5 | -5.4 | | | 3E94 |
| ATRA | -9.2 | -9.1 | | | | | 1FBY |
| estradiol | -9.5 | -9.6 | | | | | 1FBY |
| rosiglitazone | | | | | -8.4 | -8.2 | 4EMA |
| TBT | | | | | -5.2 | -5.2 | 3WJA |

## Supplementary Table 4.3

Available on CD: Supplementary_table_4_3.doc

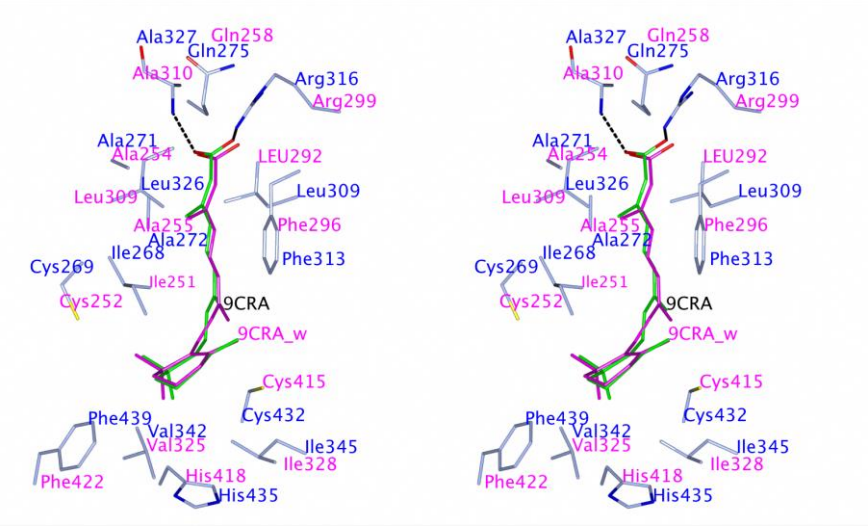
**Supplementary Table 4.3: Sequence identity (percent %) of amino acid sequences of *C. gigas* CgRXR-1, CgRXR-2, CgRAR and CgPPAR to molluscan and human receptor homologs.** Cg: *C. gigas*; Tc: *T. clavigera;* Nl: *N. lapillus*; Cf: *C. farreri*; Bg: *B. glabrata*; Lg: *L. gigantea*; Ls: *L. stagnalis*; Hs: *H. sapiens*. aa: amino acids

| NR | A/B | C* | T-box isoform | T-box length | D | E | F | Length (aa) |
|---|---|---|---|---|---|---|---|---|
| CgRXR-1 | | | | | | | | 446 |
| CgRXR-2 | 100 (1aa) | 99 (1aa) | 13 aa | | 100 | 98 (5aa) | 100 | 459 |
| TcRXR | 60 | 94-95 | -1:<br>-2: | -<br>5 aa | 89 | 90-91 | 79 % | 442<br>447 |
| NlRXR | 61 | 90-94 | -1:<br>-2: | -<br>5 aa | 89 | 90-92 | 75 % | 441<br>446 |
| CfRXR | 69 | 96-97 | -a:<br>-b:<br>-c:<br>-d: | -<br>4 aa<br>20 aa<br>24 aa | 95 | 90-92 | 75 % | 446<br>450<br>466<br>470 |
| BgRXR | 61-62 | 95-96 | - | | 84 | 91 | 71 | 436 |
| LsRXR | 60 | 95-96 | - | | 84 | 92-93 | 71 | 436 |
| HsRXRα | 46 | 88-90 | - | | 63 | 81-83 | 67 | 462 |
| HsRXRβ | 12 | 82-83 | - | | 37 | 77-78 | 67 | 533 |
| HsRXRγ | 41 | 86-87 | - | | 32 | 80-81 | 67 | 463 |
| CgRAR | | | | | | | | 462 |
| TcRAR | 14 | 95 | - | | 46 | 60 | 85 | 472 |
| LsRAR | 18 | 91 | - | | 64 | 58 | 93 | 478 |
| NlRAR | 14 | 90 | - | | 25 | 60 | 97 | 365 |
| HsRARα | 20 | 86 | - | | 39 | 51 | 2 | 462 |
| HsRARβ | 19 | 87 | - | | 39 | 51 | 17 | 455 |
| HsRARγ | 25 | 90 | - | | 43 | 49 | 24 | 454 |
| CgPPAR | | | | | | | | 499 |
| BgPPAR1 | 26 | 78 | | | 59 | 29 | | 391 |
| LgPPAR1 | 29 | 75 | | | 76 | 38 | | 365 |
| HsPPARα | 14 | 56 | - | | 14 | 22 | - | 468 |
| HsPPARβ | 9 | 56 | - | | 14 | 23 | - | 441 |
| HsPPARγ | 13 | 57 | - | | 0 | 24 | - | 505 |

* C domain (DBD) without isoform region in T-box

## Supplementary Figure 4.2

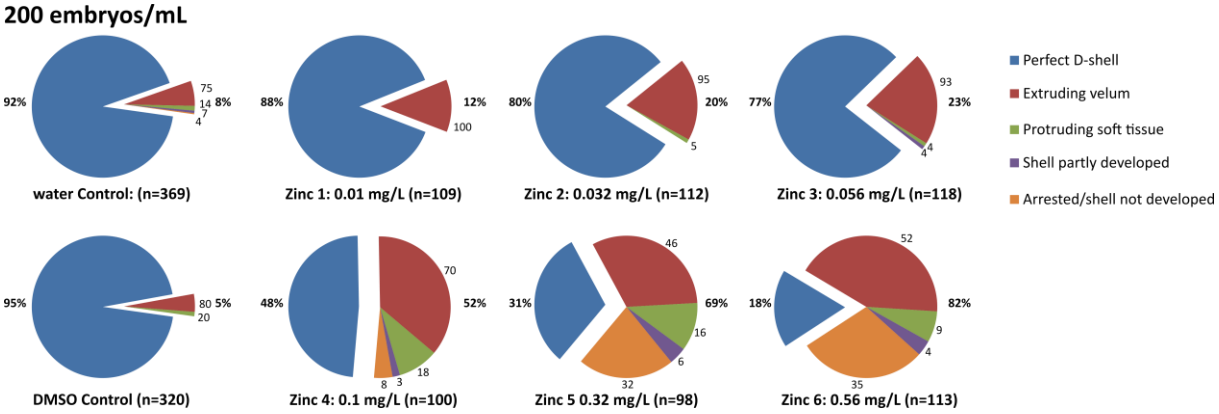
Available on CD: Supplementary_figure_4_2.doc




**Supplementary Figure 4.2: Stereo view of 9RA bound to the ligand binding pocket of a CgRXR model.** Superimposition of model CgRXR (purple) on the crystal structure of HsRXRα LBD (blue) bound to human RXR agonist 9RA. Original 9RA (green) bound to HsRXRα LBD template (pdb ID: 1FBY); 9RA (pink) to CgRXR. Divergent residues as well as arginines binding to the COOH group of 9RA are shown as stick models. Hydrogen bonds are indicated as dashed lines.

## Supplementary Figure 4.3

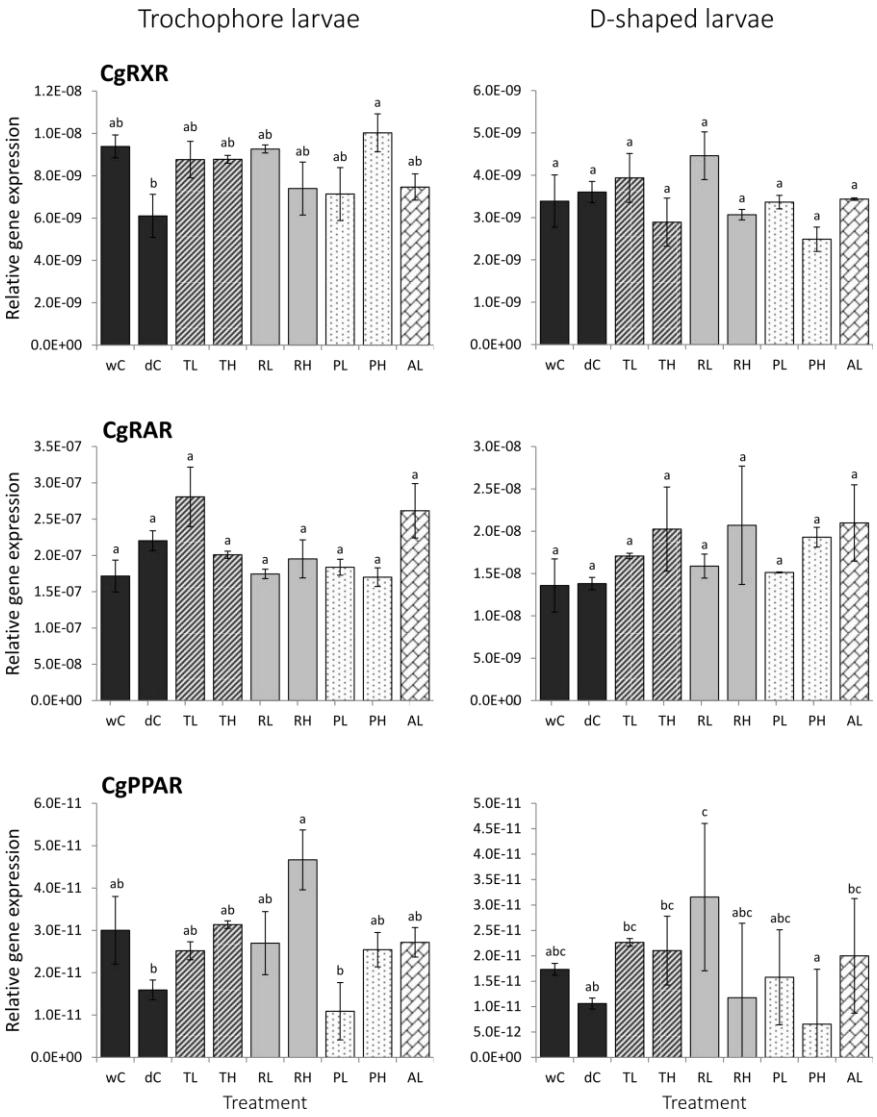
Available on CD: Supplementary_figure_4_3.doc




**Supplementary Figure 4.3: Oyster embryo development after exposure to zinc and TBTO.** Control exposure (200 embryos/ml) to increasing zinc concentration and water and DMSO Control, and TBTO low and TBTO high. Percentage of perfect developed D-shaped larvae (blue), and abnormal developed larvae grouped in four categories: extruding velum (red), protruding soft tissue (green), shell partly developed (purple), and arrested/shell not developed (orange). Bold numbers next to pie charts: percentage perfect D-shaped (left) and total abnormal D-shaped (right) larvae. Non-bold numbers: percentage of abnormal developed categories to total percentages of abnormal developed D-shell larvae. n: total count embryos/treatment.

## Supplementary Figure 4.4

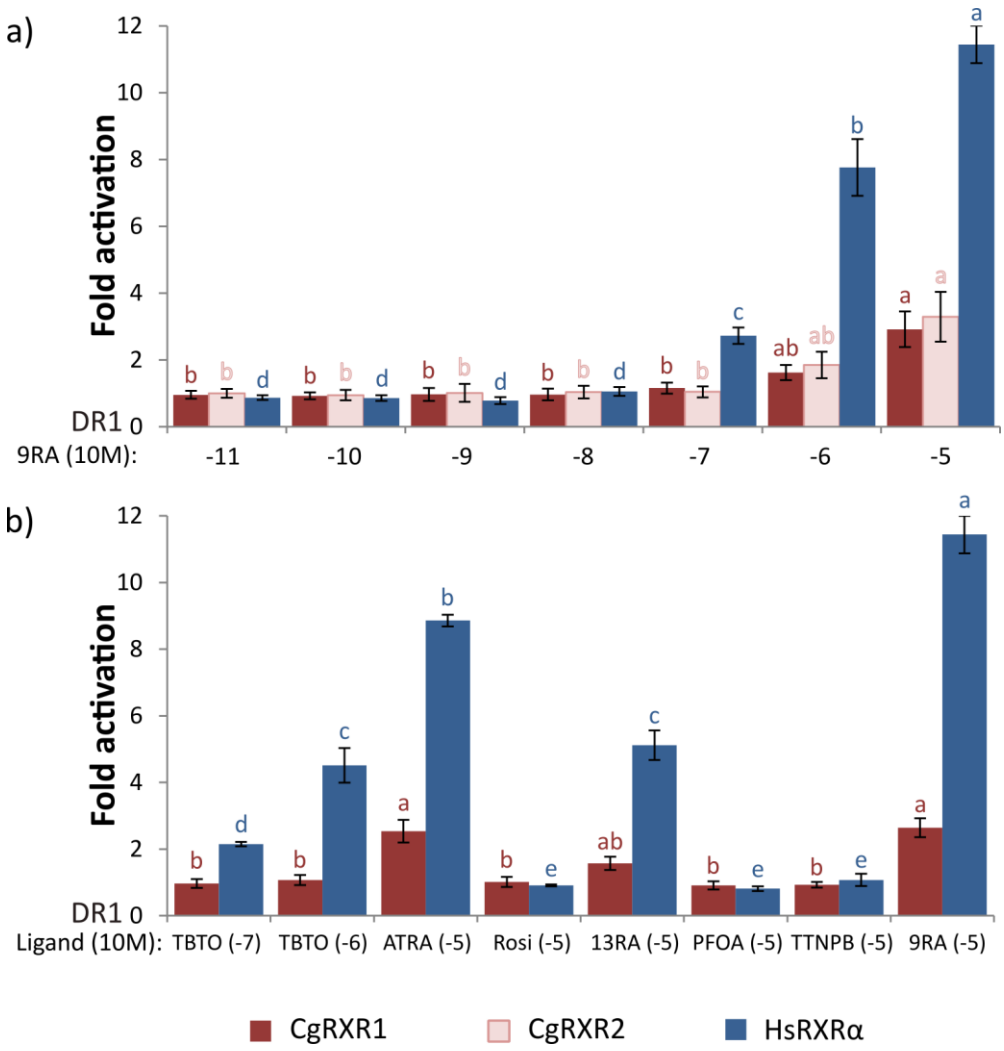
Available on CD: Supplementary_figure_4_4.doc




**Supplementary Figure 4.4: Relative gene expression of CgRXR, CgRAR and CgPPAR nuclear receptors of oyster trochophore and D-shaped larvae exposed to TBTO, Rosi, PFOA and ATRA.** Gene expression was measured with quantitative RT-PCR. Relative gene expression was calculated using a normalisation factor computed with the three reference genes and statistically analysed as described in the methods section. Bars indicate the mean ± standard error of three independent measurements per time point. Letters above each bar represent groups that were significantly different (p=0.05). wC: water Control; dC: DMSO Control; TL: TBTO low (0.2 µg/L); TH: TBTO high (2 µg/L); RL: rosiglitazone low (4 µg/L); RH: rosiglitazone high (40 µg/L); PL: PFOA low (20 mg/L); PH: PFOA high (50 mg/L); AL: ATRA low (0.06 mg/L).

**Supplementary Figure 5.1**

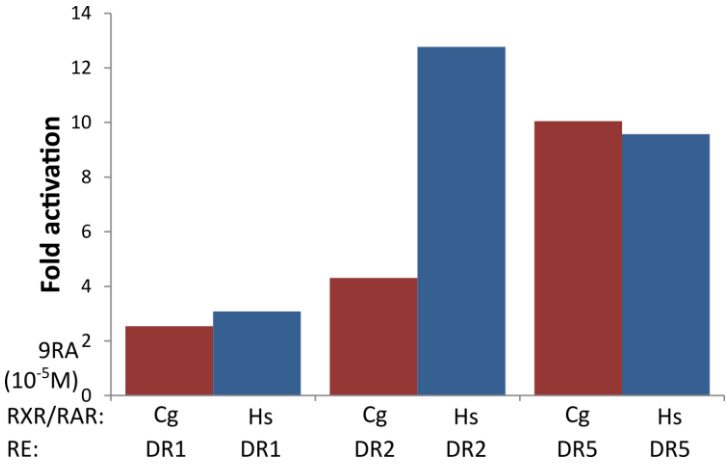
Available on CD: Supplementary_figure_5_1.doc




**Supplementary Figure 5.1: Transcriptional activity of HsRXRα, CgRXR-1 and CgRXR-2 dual-luciferase reporter assays.** COS-7 cells were transfected with DR1 reporter vector in combination with either human HsRXRα, *Crassostrea gigas* CgRXR-1 or CgRXR-2. Cells were incubated for 40-42 h **a)** with 9RA in a dose response ($10^{-11}$-$10^{-5}$ M) and **b)** with different chemical at $10^{-5}$ M: ATRA, rosiglitazone (Rosi), 13RA, PFOA, TTNPB and 9RA. The TBTO exposures were performed with two different concentrations: $10^{-7}$ M and $10^{-6}$ M. The results of triplicates are shown as mean ±SE. Lower case letters: significant differences between different chemical concentrations and chemicals types for each receptor ($P < 0.05$). No empty pcDNA3.1(+) vector control was conducted along these assay sets.

## Supplementary Figure 5.2

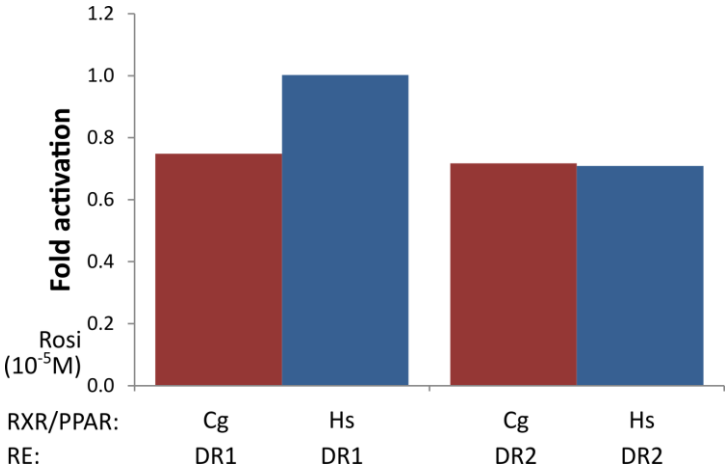
Available on CD: Supplementary_figure_5_2.doc




**Supplementary Figure 5.2: Transcriptional activity of human and oyster RXR/RAR dual-luciferase reporter assays.** COS-7 cells were co-transfected with expression vectors fused with either human HsRXRα/HsRARβ (Hs) or *Crassostrea gigas* CgRXR-1/CgRAR (Cg) in combination with DR1, DR2 or DR5 reporter vector. The cells were exposed for 40-42 h with $10^{-5}$ M 9RA. RE: response element.

## Supplementary Figure 5.3

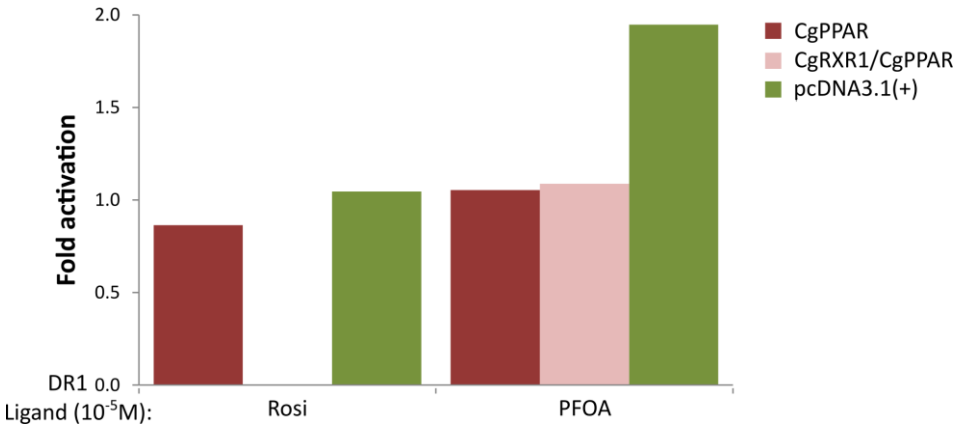
Available on CD: Supplementary_figure_5_3.doc




**Supplementary Figure 5.3: Transcriptional activity of human and oyster RXR/PPAR dual-luciferase reporter assays.** COS-7 cells were transfected with expression vectors fused with either human HsRXRα/HsPPARγ or *Crassostrea gigas* CgRXR-1/CgPPAR in combination with DR1 or DR2 reporter vector. The cells were exposed for 40-42 h with $10^{-5}$ M rosiglitazone (Rosi). RE: response elements.

## Supplementary Figure 5.4

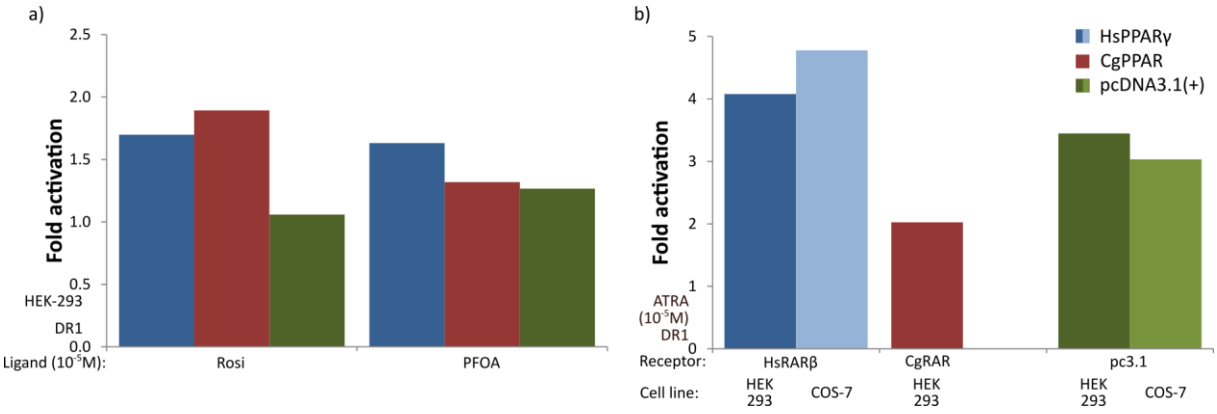
Available on CD: Supplementary_figure_5_4.doc




**Supplementary Figure 5.4: Transcriptional activity of Pacific oyster CgPPAR, CgRXR-1/CgPPAR and empty expression vector pcDNA3.1(+) dual-luciferase reporter assays.** COS-7 cells were transfected with DR1 reporter vector in combination with either *Crassostrea gigas* CgPPAR, CgRXR-1/CgPPAR or an empty expression vector. Cells were incubated for 40-42 h either with $10^{-5}$ M rosiglitazone (Rosi) or PFOA.

## Supplementary Figure 5.5

Available on CD: Supplementary_figure_5_4.doc




**Supplementary Figure 5.5: Transcriptional activity of HsPPARγ, HsRARβ, CgPPAR, CgRAR and empty expression vector pcDNA3.1(+) dual-luciferase reporter assays. a)** HEK-293 cells were transfected with DR1 reporter vector in combination with either human HsPPARγ, *Crassostrea gigas* CgPPAR, or the empty expression vector. Cells were incubated for 40-42 h either with $10^{-5}$ M rosiglitazone (Rosi) or PFOA. **b)** HEK-293 or COS-7 cells were transfected with DR1 reporter vector in combination with either HsRARβ, or CgRAR, or the empty expression vector (pc3.1). Cells were incubated for 40-42 h with $10^{-5}$ M ATRA.

# Bibliography


Abascal F, Zardoya R, Posada D. ProtTest: Selection of best-fit models of protein evolution. Bioinformatics. 2005; 21:2104-2105.

Abbott BD, Wolf CJ, Schmid JE, Das KP, Zehr RD, Helfant L, Nakayama S, Lindstrom AB, Strynar MJ, Lau C. Perfluorooctanoic acid induced developmental toxicity in the mouse is dependent on expression of peroxisome proliferator-activated receptor-alpha. Toxicol Sci. 2007; 98(2):571-581.

Abidli S, Castro LF, Lahbib Y, Reis-Henriques MA, Trigui El Menif N, Santos MM. Imposex development in *Hexaplex trunculus* (Gastropoda: Caenogastropoda) involves changes in the transcription levels of the retinoid X receptor (RXR). Chemosphere. 2013; 93(6):1161-1167.

Ables ET, Bois KE, Garcia CA, Drummond-Barbosa D. Ecdysone response gene E78 controls ovarian germline stem cell niche formation and follicle survival in *Drosophila*. Dev Biol. 2015; 400(1):33-42.

Abrahams BS, Kwok MC, Trinh E, Budaghzadeh S, Hossain SM, Simpson EM. Pathological aggression in “Fierce” mice corrected by human nuclear receptor 2E1. J Neorosci. 2005; 25:6263-6270.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li, PW, Hoskins RA, Galle RF, et al. The genome sequence of *Drosophila melanogaster*. Science. 2000; 287:2185-2195

Allen AK, Spradling AC. The Sf1-related nuclear hormone receptor Hr39 regulates *Drosophila* female reproductive tract development and function. Development. 2008; 135:311-321.

Allenby G, Janocha R, Kazmer S, Speck J, Grippo JF, Levin AA. Binding of 9-*cis*-retinoic acid and all-*trans*-retinoic acid to retinoic acid receptors alpha, beta, and gamma. Retinoic acid receptor gamma binds all-*trans*-retinoic acid preferentially over 9-*cis* retinoic acid. J Biol Chem. 1994; 269:16689-16695.

Allenby G, Bocquel MT, Saunders M, Kazmer S, Speck J, Rosenberg M, Lovey A, Kastner P, Grippo JF, Chambon P. Retinoic acid receptors and retinoid X receptors: interactions with endogenous retinoic acids. Proc Natl Acad Sci USA. 1993; 90(1):30-34.

André A, Ruivo R, Gesto M, Castro LF, Santos MM. Retinoid metabolism in invertebrates: when evolution meets endocrine disruption. Gen Comp Endocrinol. 2014 ;208:134-145.

Antebi A, Yeh WH, Tait D, Hedgecock EM, Riddle DL. Daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in *C. elegans*. Genes Dev. 2000; 14:1512-1527.

Aranda A, Pascual A. Nuclear hormone receptors and gene expression. Physiol Rev. 2001; 81(3):1269-1304.

Baker ME. Xenobiotics and the evolution of multicellular animals: emergence and diversification of ligand-activated transcription factors. Integr Comp Biol. 2005; 45(1):172-178.

Baker ME. *Trichoplax*, the simplest known animal, contains an estrogen-related receptor but no estrogen receptor: Implications for estrogen receptor evolution. Biochem Biophys Res Commun. 2008; 375:623–627.Baker ME.

Baker ME. Origin and divers:ification of steroids: co-evolution of enzymes and nuclear receptors. Mol Cell Endocrinol. 2011; 334:14-20.

Baker ME, Chandsawangbhuwana C. Analysis of 3D models of octopus estrogen receptor with estradiol: evidence for steric clashes that prevent estrogen binding. Biochem Biophys Res Commun. 2007; 361:782-788.

Baker KD, Shewchuk LM, Kozlova T, Makishima M, Hassell A, Wisely B, Caravella JA, Lambert MH, Reinking JL, Krause H, Thummel CS, Willson TM, Mangelsdorf DJ: The *Drosophila* orphan nuclear receptor DHR38 mediates an atypical ecdysteroid signaling pathway. Cell. 2003; 113:731–742.

Banerjee M, Robbins D, Chen T. Modulation of xenobiotic receptors by steroids. Molecules. 2013; 18:7389-7406.

Bannister R, Beresford N, May D, Routledge EJ, Jobling S, Rand-Weaver M. Novel estrogen receptor-related transcripts in *Marisa cornuarietis*; a freshwater snail with reported sensitivity to estrogenic chemicals. Environ Sci Technol. 2007; 41:2643-2650.

Bannister R, Beresford N, Granger DW, Pounds NA, Rand-Weaver M, White R, Jobling S, Routledge EJ. No substantial changes in estrogen receptor and estrogen-related receptor orthologues gene

transcription in *Marisa cornuarietis* exposed to estrogenic chemicals. Aquat Toxicol. 2013; 140-141:19-26.

Bardet PL, Laudet V, Vanacker JM. Studying non–mammalian models? Not a fool's ERRand! Trends Endocrinol Metab. 2006; 17(4):166–169.

Bashirullah A, Cooperstock RL, Lipshitz HD. Spatial and temporal control of RNA stability. Proc Natl Acad Sci USA. 2001; 98(13):7025–7028.

Bashirullah A, Halsell SR, Cooperstock RL, Kloc M, Karaiskakis A, Fisher WW, Fu W, Hamilton JK, Etkin LD, Lipshitz HD. Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in Drosophila melanogaster. EMBO J. 1999; 18(9):2610–2620.

Båvner A1, Sanyal S, Gustafsson JA, Treuter E. Transcriptional corepression by SHP: molecular mechanisms and physiological consequences. Trends Endocrinol Metab. 2005; 16(10):478-488.

Beato M, Klug J. Steroid hormone receptors: an update. Hum Reprod Update. 2000; 6(3):225-236.

Benoit B, He CH, Zhang F, Votruba SM, Tadros W, westwood JT, Smibert CA, Lipshitz HD, Theurkauf WE. An essential role for the RNA–binding protein Smaug during the *Drosophila* maternal–to–zygotic transition. Development. 2009; 136:923–932.

Benstead RS, Baynes A, Casey D, Routledge EJ, Jobling S. 17β-Oestradiol may prolong reproduction in seasonally breeding freshwater gastropod molluscs. Aquat Toxicol. 2011; 101:326-334.

Bergman Å, Heindel JJ, Jobling S, Kidd KA, Zoeller RT. The state-of-the-science of endocrine disrupting chemicals - 2012. Geneva: UNEP/WHO; 2013.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000; 28(1):235-242.

Bertrand S, Belgacem MR, Escriva H. Nuclear hormone receptors in chordates. Mol Cell Endocrinol. 2011; 334:67-75.

Bertrand S, Brunet FG, Escriva H, Parmentier G, Laudet V, Robinson-Rechavi M. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. Mol Biol Evol. 2004; 21:1923–1937.

Bouton D, Escriva H, de Mendoconça RL, Glineur C, Bertin B, Noël C, Robinson-Rechavi M, de Groot A, Cornette J, Laudet V, Pierce RJ. A conserved retinoid X receptor (RXR) from the mollusk *Biomphalaria glabrata* transactivates transcription in the presence of retinoids. J Mol Endocinol. 2005; 34:567-582.

Bridgham JT, Eick GN, Larroux C, Deshpande K, Harms MJ, Gauthier MEA, Ortlund EA, Degnan BM, Thornton JW. Protein evolution by molecular tinkering: diversification of the nuclear receptor superfamily from a ligand–dependent ancestor. PLoS Biol. 2010 ;8:e1000497.

Bright DA, Ellis DV. A comparative survey of imposex in northeast Pacific neogastropods (Prosobranchia) related to tributyltin contamination, and choice of a suitable bioindicator. Can J Zool. 1990; 68:1915-1924.

Broadus J, Doe CQ. Evolution of neuroblast identity: seven-up and prospero expression reveal homologous and divergent neuroblast fates in *Drosophila* and *Schistocerca*. Development. 1995; 121:3989–3996.

Bruce CM, Campbell MJ. Nuclear Receptors: current concepts and future challenges. Springer; 2010.

Cai SY, He H, Nguyen T, Mennone A, Boyer JL. Retinoic acid represses CYP7A1 expression in human hepatocytes and HepG2 cells by FXR/RXR-dependent and independent mechanisms. J Lipid Res. 2010; 51(8):2265-2274.

Cajaraville MP, Ortiz-Zarragoitia M. Specificity of the peroxisome proliferation response in mussel exposed to environmental pollutants. Aquat Toxicol. 2006; 78S:S117-S123.

Camara MD, Symonds JE. Genetic improvement of New Zealand aquaculture species: programmes, progress and prospects. New Zeal J Mar Fresh. 2014; 48:466-491.

Cancio I, Orbea A, Völkl A, Fahimi HD, Cajaraville MP. Induction of peroxisomal oxidases in mussels: comparison of effects of lubricant oil and benzo(a)pyrene with two typical peroxisome proliferators on peroxisome structure and function in *Mytilus galloprovincialis*. Toxicol Appl Pharmacol. 1998; 149:64–72.

Canesi L, Fabbri E. Environmental effects of BPA: focus on aquatic species. Dose Response. 2015; 13(3):1559325815598304.

Carefoot T. Reproduction & development. Imposex. n.D. http://www.asnailsodyssey.com/LEARNABOUT/WHELK/whelImpo.php (Accessed 26 May 2016)
Carmi I, Kopczynski JB, Meyer BJ. The nuclear hormone receptor SEX–1 is an X–chromosome signal that determines nematode sex. Nature. 1998; *396*:168–173.
Carter CJ, Spencer GE. Cloning of a retinoic acid receptor (RAR) from the CNS of a non-chordate, invertebrate protostome and its role in embryonic development [abstract]. In: Program 225.11, 2009 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience; 2009. p. 152-157.
Carter CJ, Farrar N, Carlone RI, Spencer GE. Developmental expression of a molluscan RXR and evidence for its novel, nongenomic role in growth cone guidance. Dev Biol. 2010; 343:124–137.
Carter CJ, Rand C, Mohammad I, Lepp A, Vesprini N, Wiebe O, Carlone R, Spencer GE. Expression of a retinoic acid receptor (RAR)–like protein in the embryonic and adult nervous system of a protostome species. J Exp Zool. 2015; 324B:51–67.
Castro LF, Lima D, Machado A, Melo C, Hiromori Y, Nishikawa J, Nakanishi T, Reis-Henriques MA, Santos MM. Imposex induction is mediated through the Retinoid X Receptor signalling pathway in the neogastropod *Nucella lapillus*. Aquat Toxicol. 2007; 85:57-66.
Castro LF, Melo C, Guillot R, Mendes I, Queiros S, Lima D, Reis-Henriques MA, Santos MM. The estrogen receptor of the gastropod *Nucella lapillus*: Modulation following exposure to an estrogenic effluent. Aquat Toxicol. 2007a; 84:465-468.
Chagot D, Alzieu C, Sanjuan J, Grizel H. Sublethal and histopathological effects of trace levels of tributyltin fluoride on adult oysters *Crassostrea gigas*. Aquat Living Resour. 1990; 3:121–130.
Cheek AO, P M Vonier, E Oberdörster, B C Burow, and J A McLachlan. Environmental signaling: a biological context for endocrine disruption. Environ Health Perspect. 1998; 106 Suppl 1:5–10.
Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CLS, Chan C, Dent OF, Clarke SJ, Robertson GR, Sladek FM. Src tyrosine kinase phosphorylation of nuclear receptor HNF4α correlates with isoform-specific loss of HNF4α in human colon cancer. Proc Natl Acad Sci USA. 2012; 109:2302-2307.
Chen JD, Evans RM. A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature. 1995 Oct 5; 377(6548):454-457.
Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, Darnell JE Jr.. Disruption of the HNF–4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. Genes Dev. 1994; 8:2466–2477.
Cheng X, Klaassen CD. Critical role of PPAR-alpha in perfluorooctanoic acid- and perfluorodecanoic acid-induced downregulation of Oatp uptake transporters in mouse livers. Toxicol Sci. 2008; 106(1):37-45.
Chomczynski P, Mackey K. Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich source. Biotechniques. 1995; 19:942-945.
Codina A, Benoit G, Gooch JT, Neuhaus D, Perlmann T, Schwabe JW. Identification of a novel co-regulator interaction surface on the ligand binding domain of Nurr1 using NMR footprinting. J Biol Chem. 2004; 279:53338–53345.
Coen LD, Brumbaugh RD, Bushek D, Grizzle R, Luckenbach MW, Posey MH, Powers SP, Tolley SG. Ecosystem services related to oyster restoration. Mar Ecol Prog Ser. 2007; 341:303-307.
Commission Directive (CD). Commission Directive 2002/62/EC of 9 July 2002. O J Eur Com. 2002; L183:58-59.
Creton R, Zwaan G, Dohmen R. Specific developmental defects in molluscs after treatment with retinoic acid during gastrulation. Dev Growth Differ. 1993; 35:357–364.
Cruzeiro C, Lopes-Marques M, Ruivo R, Rodrigues-Oliveira N, Santos MM, Rocha MJ, Rocha E, Castro LF. A mollusk VDR/PXR/CAR-like (NR1J) nuclear receptor provides insight into ancient detoxification mechanisms. Aquat Toxicol. 2016; 174:61-69.
Cui S, Kolodziej KE, Obara N, Amaral-Psarris A, Demmers J, Shi L, Engel JD, Grosveld F, Strouboulis J, Tanabe O. Nuclear receptors TR2 and TR4 recruits multiple epigenetic transcriptional

corepressors that associate specifically with the embryonic β-type globin promoters in differentiated adult erythroid cells. Mol Cell Biol. 2011; 31:3298-3311.
Darras VM, van Herck SLJ, Heijlen M, de Groef B. Thyroid hormone receptors in two model species for vertebrate embryonic development: chicken and zebrafish. J Thyroid Res. 2011; 402320.
Davidson E. Gene activity in early development. 3rd ed. Orlando: Academic Press, Inc.; 1986.
Davies IM, Vethaak D. Integrated marine environmental monitoring of chemicals and their effects. ICES Cooperative Research Report. 2012; 315:1-277
Dawson MI, Xia Z: The retinoid X receptors and their ligands. Biochim Biophys Acta. 2012; 1821:21-56.
De Hoon MJL, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics. 2004; 20(9):1453–1454.
De Lisa E, Paolucci M, Di Cosmo A. Conservative nature of oestradiol signalling pathways in the brain lobes of *Octopus vulgaris* involved in reproduction, learning and motor coordination. J Neuroendocrinol. 2012; 24:275-284.
De Mendonça RL, Bouton D, Bertin B, Escriva H, Noël C, Vanacker JM, Cornette J, Laudet V, Pierce RJ. A functionally conserved member of the FTZ-F1 nuclear receptor family from *Schistosoma mansoni*. Eur J Biochem. 2002; 269:5700–5711.
De Renzis S, Elemento O, Tavazoie S, Wieschaus F. Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. PLoS Biol. 2007; 5:e117.
Delfosse V, Maire AL, Balaguer P, Bourguet W. A structural perspective on nuclear receptors as targets of environmental compounds. Acta Pharmacol Sin. 2015; 36(1):88–101.
Delmotte MH, Tahayato A, Formstecher P, Lefebvre P. Serine 157, a retinoic acid receptor alpha residue phosphorylated by protein kinase C in vitro, is involved in RXR/RARalpha heterodimerization and transcriptional activity. J Biol Chem. 1999; 274(53):38225-38231.
Denayer S, Helsen C, Thorrez L, Haelens A, Claessens F. The rules of DNA recognition by the androgen receptor. Mol Endocrinol. 2010; 24(5):898-913.
Desvergne B, Feige JN, Casals-Casas C. PPAR-mediated activity of phthalates: A link to the obesity epidemic? Mol Cell Endocrinol. 2009; 304:43-48.
Devchand PR1, Keller H, Peters JM, Vazquez M, Gonzalez FJ, Wahli W. The PPARalpha-leukotriene B4 pathway to inflammation control. Nature. 1996; 384(6604):39-43.
Diamanti-Kandarakis E, Bourguignon JP, Giudice LC, Hauser R, Prins GS, Soto AM, Zoeller RT, Gore AC. Endocrine-disrupting chemicals: an Endocrine Society scientific statement. Endocr Rev. 2009; 30:293-342.
Dollé P. Developmental expression of retinoic acid receptors (RARs). Nucl Recept Signal. 2009; 7:e006.
Dubrovskaya VA, Berger EM, Dubrovsky EB. Juvenile hormone regulation of the E75 nuclear receptor is conserved in Diptera and Lepidoptera. Gene. 2004; 340:171-177.
Dubrovsky EB, Dubrovskaya VA, Bernardo T, Otte V, DiFilippo R, Bryan H. The *Drosophila* FTZ-F1 nuclear receptor mediates juvenile hormone activation of E75A gene expression through an intracellular pathway. J Biol Chem. 2011; 286:33689-33700.
Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, Darnell JE Jr.. Expression of transcription factor HNF–4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF–4 is a marker for primary endoderm in the implanting blastocyst. Proc Natl Acad Sci USA. 1994; 91:7598–7602.
Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32:1792-1797.
Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P, Moras D. Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-*cis* retinoic acid. EMBO J. 2000; 19:2592-2601.
Ehrlund A, Treuter E. Ligand-independent actions of the orphan receptors/corepressors DAX-1 and SHP in metabolism, reproduction and disease. J Steroid Biochem Mol Biol. 2012; 130:169-179.
Eick GN, Thornton JW. Evolution of steroid receptors from an estrogen-sensitive ancestral receptor. Mol Cell Endocrinol. 2011; 334:31-38.

Ellis I, Kempf SC. Characterization of the central nervous system and various peripheral innervations during larval development of the oyster *Crassostrea virgonica*. Invert Biol. 2011; 130(3):236–250.

Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2004; 66:486-501.

Escriva H*,* Delaunay F, Laudet V. Ligand binding and nuclear receptor evolution. BioEssays*.* 2000; 22(8):717-*727.*

Escriva H, Holland ND, Gronemeyer H, Laudet V, Holland LZ. The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. Development. 2002; 129(12):2905-16.

Escriva, H, Safi R, Hanni C, Langlois MC, Saumitou-Laprade P, Stehelin D, Capron A, Pierce R, Laudet V. Ligand binding was acquired during evolution of nuclear receptors. Proc Natl Acad Sci USA. 1997; 94:6803-6808.

European Food Safety Authority (EFSA). Perfluoroalkylated substances in food: occurrence and dietary exposure. EFSA J. 2012; 10(6):2743.

Evans SM. Marine antifoulants. In: Sheppard C, editor. Seas at the millennium: An environmental evaluation, Vol. III: Global issues and processes. Oxford: Elsevier Science Ltd; 2000. p. 247–256,

Factoextra. n.D. http://www.sthda.com/english/wiki/factoextra-r-package-easy-multivariate-data-analyses-and-elegant-visualization. (Accessed 3 May 2016).

Fahrbach SE, Smagghe G, Velarde RA. Insect nuclear receptors. Annu Rev Entomol. 2012; 57:83-106.

Fayard E, Auwerx J, Schoonjans K. LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis. Trends Cell Biol. 2004; 14:250-260.

Fernandes I, White JH. Agonist-bound nuclear receptors: not just targets of coactivators. J Mol Endocrinol. 2003; 31(1):1-7.

Filby AL, Tyler CR. Molecular characterization of estrogen receptors 1, 2a, and 2b and their tissue and ontogenetic expression profiles in fathead minnow (*Pimephales promelas*). Biol Reprod. 2005; 73:648–662.

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslunder K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A. The Pfam protein families database. Nucleic Acids Res. 2010; 38:D211-D222.

Fioroni P, Oehlmann J, Stroben E. The pseudohermaphroditism of prosobranchs: morphological aspects. Zool Anz. 1991; 226:1–26.

Flaig R, Greschik H, Peluso-Iltis C, Moras D. Structural basis for the cell-specific activities of the NGFIB and the Nurr1 ligand-binding domain. J Biol Chem. 2005; 280:19250–19258.

Forman BM, Umesono K, Chen J, Evans RM. Unique response pathways are established by allosteric interactions among nuclear hormone receptors. Cell. 1995b; 81:541–550.

Forman BM, Casanova J, Raaka BM, Ghysdael J, Samuels HH. 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. Mol Endocrinol. 1992; 6(3):429-442.

Forman BM, Tontonoz P, Chen J, Brun RP, Spiegelman BM, Evans RM. 15-Deoxy-delta 12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR gamma. Cell. 1995a; 83(5):803-12.

Forman BM, Chen J, Blumberg B, Kliewer SA, Henshaw R, Ong ES, Evans RM. Cross-Talk among RORα1 and the Rev-erb family of orphan nuclear receptors. Mol Endocrinol. 1994; 8:1253-1261.

Fröhlicher M, Liedtke A, Groh K, Lopez–Schier H, Neuhauss SCF, Segner H, Eggen RIL. Estrogen receptor subtype β2 is involved in neuromast development in zebrafish (*Danio rerio*) larvae. Dev Biol. 2009; 330:32–43.

Garner MM, Revzin A. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. Nucleic Acids Res. 1981; 9(13):3047-3060.

Germain P, Altucci L, Bourguet W, Rochette-Egly C, Gronemeyer H. Nuclear receptor superfamily: principles of signaling. Pure Appl Chem. 2003; 75:1619-1664.

Germain P, Staels B, Dacquet C, Spedding M, Laudet V. Overview of nomenclature of nuclear receptors. Pharmacol Rev. 2006; 58:685–704.

Germain P, Kammerer S, Perez E, Peluso-Iltis C, Tortolani S, Zusi FC, Starrett J, Lapointe P, Daris JP, Marinier A, De Lera AR, Rochel N, Gronemeyer H. Rational design of RAR-selective ligands revealed by RARbeta crystal structure. EMBO Rep. 2004; 5:877-882.

Gervois P, Chopin-Delannoy S, Fadel A, Dubois G, Kosykh V, Fruchart JC, Naj'b J, Laudet V, Staels B. Fibrates increase human REV-ERBalpha expression in liver via a novel peroxisome proliferator-activated receptor response element. Mol Endocrinol. 1999; 13:400–409

Gesto M, Castro LF, Santos MM. Differences in retinoid levels and metabolism among gastropod lineages: imposex-susceptible gastropods lack the ability to store retinoids in the form of retinyl esters. Aquat Toxicol. 2013; 142-143:96-103.

Giannakis G, Edmondson SR, Favaloro JM, Zajac JD, Greenland KJ. Aberrant cryptic responsiveness of the pCAT 3- and pGL3-promoter reporter vectors. Biotechniques. 2003; 35(2):332-339.

Gibbs PE, Pascoal PL, Burt GR. Sex change in the female dog-whelk, *Nucella lapillus*, induced by tributyltin from antifouling paints. J Mar Biol Assoc UK, 1988; 68:715-731.

Giesy JP, Naile JE, Khim JS, Jones PD, Newsted JL. Aquatic toxicology of perfluorinated chemicals. Rev Environ Contam Toxicol. 2010; 202:1-52.

Gissendanner CR, Crossgrove K, Kraus KA, Maina CV, Sluder AE. Expression and function of conserved nuclear receptor genes in *C. elegans*. Dev Biol. 2004; *266*:399–416.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG. Life with 6000 genes. Science. 1996; 274(5287):546, 563-567.

Goodrich RJ, Anton E, Krawetz SA. Isolation mRNA and small noncoding RNAs from human sperm. Methods Mol Biol. 2013; 927:385–396.

Gore AC, Chappell VA, Fenton SE, Flaws JA, Nadal A, Prins GS, Toppari J, Zoeller RT. Executive summary to EDC-2: the endocrine society's second scientific statement on endocrine-disrupting chemicals. Endocr Rev. 2015; 36(6):593-602.

Gosling E. Bivalve molluscs: biology, ecology and culture. Reprint. Bodim: Blackwell Science; 2004.

Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 2010; 27:221-224.

Grasso LC, Hayward DC, Trueman JWH, Hardie KM, Janssens PA, Ball EE. The evolution of nuclear receptors: evidence from the coral *Acropora*. Mol Phylogenet Evol. 2001; 21:93-102.

Greschik H, Flaig R, Renaud JP, Moras D. Structural basis for the deactivation of the estrogen-related receptor gamma by diethylstilbestrol or 4-hydroxytamoxifen and determinants of selectivity. J Biol Chem. 2004; 279(32):33639-33646.

Grün F, Watanabe H, Zamanian Z, Maeda L, Arima K, Cubacha R, Gardiner DM, Kanno J, Iguchi T, Blumberg B. Endocrine-disrupting organotin compounds are potent inducers of adipogenesis in vertebrates. Mol Endocrinol. 2006; 20:2151-2155.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010; 59:307-321.

Guo X, Hedgecock D, Hershberger WK, Cooper K, Allen Jr SK. Genetic determinations of podrandric sex in the Pacific oyster, *Crassostrea gigas* Thunberg. Evolution. 1998; 52(2):394–402.

Gutierrez–Mezariegos J, Nadendla EK, Lima D, Pierzchalski K, Jones JW, Kane M, Nishikawa JI, Hiromori Y, Nakanishi T, Santos MM, Castro FC, Bourguet W, Schubert M, Laudet V. A mollusc retinoic acid receptor (RAR) ortholog sheds light on the evolution of ligand binding. Endocrinology. 2014; 155(11):4275–4286.

Haider NB, Jacobson SG, Cideciyan AV, Swiderski R, Streb LM, Searby C, Beck G, Hockey R, Hanna DB, Gorman S, Duhl D, Carmi R, Bennett J, Weleber RG, Fishman GA, Wright AF, Stone EM, Sheffield VC. Mutation of a nuclear receptor gene, NR2E3, causes enhanced S cone syndrome, a disorder of retinal cell fate. Nat Genet. 2000; 24:127–131.

Hammer J, Kraak MH, Parsons JR. Plastics in the marine environment: the dark side of a modern gift. Rev Environ Contam Toxicol. 2012; 220:1-44

Harada S, Hiromori Y, Nakamura S, Kawahara K, Fukakusa S, Maruno T, Noda M, Uchiyama S, Fukui K, Nishikawa J, Nagase H, Koboyashi Y, Yoshida T, Ohkubo T, Nakanishi T. Structural basis for PPARγ transactivation by endocrine-disrupting organotin compounds. Sci Rep. 2015; 5:8520.

Heery DM, Hoare S, Hussain S, Parker MG, Sheppard H. Core LXXLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. J Biol Chem. 2001; 276(9):6695-6702.

Hellman LM, Fried MG. Electrophoretic Mobility Shift Assay (EMSA) for detecting protein-nucleic acid interactions. Nat Protoc. 2007; 2(8):1849-1861.

Helm MM.Cultured aquatic species information programme. *Crassostrea gigas.* Cultured Aquatic Species Information Programme. FAO Fisheries and Aquaculture Department. 2005. http://www.fao.org/fishery/culturedspecies/Crassostrea_gigas/en (Accessed 26 May 2016).

Helm MM, Bourne N, Lovatelli A. Hatchery culture of bivalve. A practical manual. Rome: FAO Fisheries Technical Paper. 471; 2004.

Helsen C, Claessens F. Looking at nuclear receptors from a new angle. Mol Cell Endocr. 2014; 382:97-106.

Heral M, ALzieu C, Deslous-Paoli JM. Effects of organotin compounds (TBT) used in antifouling paints on cultured marine molluscs – a literature study. In: De Pauw N, Jaspers E, Ackefors H, Wilkins N, editors. Aquaculture – a biotechnology in progress. Bredene: European Aquaculture Society; 1989. p. 1081-1089.

Heyman RA, Mangelsdorf DJ, Dyck JA, Stein RB, Eichele G, Evans RM, Thaller C. 9-*cis* retinoic acid is a high affinity ligand for the retinoid X receptor. Cell. 1992; 68(2):397-406.

Higuera-Ruiz R, Elorza J. Biometric, microstructural, and high-resolution trace element studies in *Crassostrea gigas* of Cantabria (Bay of Biscay, Spain): Anthropogenic and seasonal influences. Est Coast Shelf Sci. 2009; 82:201–213.

Higuera-Ruiz R, Elorza J. Shell thickening and chambering in the oyster *Crassostrea gigas:* natural and anthropogenic influence of tributyltin contamination. Environ Technol. 2011; 32:583-591.

Hoivik EA, Lewis AE, Aumo L, Bakke M. Molecular aspects of steroidogenic factor 1 (SF-1). Mol Cell Endocrinol. 2010; 315:27-39.

Holewa B, Pogge van Strandmann E, Zapp D, Lorenz P, Ryffel GU. Transcriptional hierarchy in *Xenopus* embryogenesis: HNF4 a maternal factor involved in the developmental activation of the gene encoding the tissue specific transcription factor HNF1 alpha (LFB1). Mech Dev. 1996; 54:45–57.

Holewa B, Zapp D, Drewes T, Senkel S, Ryffel GU. HNF4beta, a new gene of the HNF4 family with distinct activation and expression profiles in oogenesis and embryogenesis of *Xenopus laevis*. Mol Cell Biol. 1997; 17:687–694.

Horard B, Vanacker JM. Estrogen receptor-related receptors: orphan receptors desperately seeking a ligand. J Mol Endocrinol. 2003; 31(3):349-357.

Horiguchi T, Shiraishi H, Shimizu M, Morita M. Imposex in sea snails, caused by organotin (tributyltin and triphenyltin) pollution in Japan: a survey. Appl Organomet Chem. 1997, 11:451–455.

Horiguchi T, Nishikawa T, Yasuhiko O, Shiraishi H, Morita M. Retinoid X receptor gene expression and protein content in tissues of the rock shell *Thais clavigera*. Aquat Toxicol. 2007; 84:379-388.

Horiguchi T, Nishikawa T, Ohta Y, Shiraishi H, Morita M. Time course of expression of the retinoid X receptor gene and induction of imposex in the rock shell, *Thais clavigera*, exposed to triphenyltin chloride. Anal Bioanal Chem. 2010a; 396:597-607.

Horiguchi T, Urushitani H, Ohta Y, Iguchi T, Shiraishi H. Establishment of a polyclonal antibody against the retinoid X receptor of the rock shell *Thais clavigera* and its application to rock shell tissues for imposex research. Ecotoxicology. 2010b; 19:571-576.

Horiguchi T, Ohta Y, Nishikawa T, Shiraishi F, Shiraishi H, Morita M. Exposure to 9-cis retinoic acid induces penis and vas deferens development in the female rock shell, *Thais clavigera*. Cell Biol Toxicol. 2008; 24:553-562.

Hörlein AJ, Näär AM, Heinzel T, Torchia J, Gloss B, Kurokawa R, Ryan A, Kamei Y, Söderström M, Glass CK, Rosenfeld MG. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature. 1995; 377(6548):397-404.

Horner MA, Pardee K, Liu S, King-Jones K, Lajoie G, Edwards A, Krause HM, Thummel CS. The *Drosophila* DHR96 nuclear receptor binds cholesterol and regulates cholesterol homeostasis. Genes Dev. 2009; 23:2711-2716.

Höss S, Weltje L. Endocrine disruption in nematodes: effects and mechanisms. Ecotoxicology. 2007; 16(1):15-28.

Huang W, Xu F, Qu T, Li L, Que H, Zhang G. Iodothyronine deiodinase gene analysis of the Pacific oyster *Crassostrea gigas* reveals possible conservation of thyroid hormone feedback regulation mechanisms in mollusks. Chinese J Ocean Limnol. 2015; 33:997–1006.

Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001; 17:754-755.

Hwang DS, Lee BY, Kim HS, Lee MC, Kyung DH, Om As, Rhee JS, Lee JS. Genome-wide identification of nuclear receptor (NR) superfamily genes in the copepod *Tigriopus japonicus*. BMC Genomics. 2014; 15(1):993.

Idres N, Marill J, Flexor MA, Chabot GG. Activation of retinoic acid receptor-dependent transcription by all-trans-retinoic acid metabolites and isomers. J. Biol. Chem. 2002; 277(35):31491-31498.

Iguchi T, Katsu Y. Commonality in signaling of endocrine disruption from snail to human. *BioScience.* 2008; 58(11):1061-1067.

IJpenberg A, Jeannin E, Wahli W, Desvergne B. Polarity and specific sequence requirements of peroxisome proliferator-activated receptor (PPAR)/retinoid X receptor heterodimer binding to DNA. A functional analysis of the malic enzyme gene PPAR response element. J Biol Chem. 1997; 272(32):20108-20117.

Inoue D, Sei K, Ike M. Disruption of retinoic acid receptor signalling by environmental pollutants. J Health Sci. 2010; 56:221-230.

Inoue S, Oshima Y, Nagai K, Yamamoto T, Go J, Kai N, Honjo T. Effect of maternal exposure to tributyltin on reproduction of the pearl oyster (*Pinctada fucata martensii*). Environ Toxicol Chem. 2004; 23(5):1276-81.

International Marine Organisation (IMO). International convention on the control of harmful antifouling systems on ships. http://www.imo.org/en/About/Conventions/ListOfConventions/Pages/International-Convention-on-the-Control-of-Harmful-Anti-fouling-Systems-on-Ships-%28AFS%29.aspx (Accessed 6 May 2016).

International Programme on Chemical Safety (IPCS).Tributyltin Compounds. Environmental Health Criteria 116. WHO. Geneva; 1990.

Iyer AK, McCabe ER. Molecular mechanisms of DAX1 action. Mol Genet Metab. 2004; 83:60–73.

Iwaki M, Matsuda M, Maeda N, Funahashi T, Matsuzawa Y, Makishima M, Shimomura I. Induction of adiponectin, a fat-derived antidiabetic and antiatherogenic factor, by nuclear receptors. Diabetes. 2003; 52(7):1655-1663.

Janer G, Porte C. Sex steroids and potential mechanisms of non-genomic endocrine disruption in invertebrates. Ecotoxicology. 2007; 16:145-160.

Jetten AM. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. Nucl Recept Signal. 2009; 7:e003.

Johnston SD1, Liu X, Zuo F, Eisenbraun TL, Wiley SR, Kraus RJ, Mertz JE. Nothing Estrogen-related receptor alpha 1 functionally binds as a monomer to extended half-site sequences including ones contained within estrogen-response elements. Mol Endocrinol. 1997; 11(3):342-52.

Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992; 8:275-282.

Kajiwara M, Kuraku S, Kurokawa T, Kato K, Toda S, Hirose H, Takahashi S, Shibata Y, Iguchi T, Matsumoto T, Miyata T, Miura T, Takahashi Y. Tissue preferential expression of estrogen receptor gene in the marine snail, *Thais clavigera*. Gen Comp Endocrinol. 2006; 148:315-326.

Kakoi S, Kin K, Miyazaki K, Wada H. Early development of Japanese spiny oyster (*Saccostrea kegaki*). Characterization of some genetic markers. Zoolog Science. 2008; 25:455–464.

Kalaany NY, Mangelsdorf DJ. LXRs and FXR: the Yin and Yang of cholesterol and fat metabolism. Annu Rev Physiol. 2006; 68:159-191.

Kam RKT, Deng Y, Chen Y, Zhao H. Retinoic acid synthesis and functions in early embryonic development. Cell Biosci. 2012; 2:11.

Kamata R, Shiraishi F, Nishikawa J, Yonemoto J, Shiraishi H. Screening and detection of the in vitro agonistic activity of xenobiotics on the retinoic acid receptor. Toxicol in Vitro. 2008; 22:1050-1061.

Kanayama T, Kobayashi N, Mamiya S, Nakanishi T, Nishikawa J. Organotin compounds promote adipocyte differentiation as agonists of the peroxisome proliferator-activated receptor gamma/retinoid X receptor pathway. Mol Pharmacol. 2005; 67(3):766-74.

Karimullina E, Li Y, Ginjupalli G, Baldwin WS. *Daphnia* HR96 is a promiscuous xenobiotic and endobiotic nuclear receptor. Aquat Toxicol. 2012; 116-117:69–78.

Kato Y, Koboyashi K, Oda S, Tatarazako N, Watanabe H, Iguchi T. Cloning and characterization of the ecdysone receptor and ultraspiracle protein from the water flea *Daphnia magna*. J Endocrinol. 2007; 193:183-194.

Kaur S, Jobling S, Jones CS, Noble LR, Routledge EJ, Lockyer AE. The nuclear receptor of *Biomphalaria glabrata* and *Lottia gigantea*: implications for developing new model organisms. PLoS One. 2015; 10:e0121259.

Kaur S, Baynes A, Lockyer AE, Routledge EJ, Jones CS, Noble LR, Jobling S. Steroid Androgen Exposure during Development Has No Effect on Reproductive Physiology of Biomphalaria glabrata. PLoS One. 2016;11(7):e0159852.

Kawamoto T, Kakizaki S, Yoshinari K, Negishi M. Estrogen activation of the nuclear orphan receptor CAR (constitutive active receptor) in induction of the mouse Cyp2b10 gene. Mol Endocrinol. 2000; 14:1897-1905.

Keay J, Bridgham JT, Thornton JW. The *Octopus vulgaris* estrogen receptor is a constitutive transcriptional activator: evolutionary and functional implications. Endocrinology. 2006; 147:3861-3869.

Kerkhofs S, Dubois V, De Gendt K, Helsen C, Clinckemalie L, Spans L, Schuit F, Boonen S, Vanderschueren D, Saunders PT, Verhoeven G, Claessens F. A role for selective androgen response elements in the development of the epididymis and the androgen control of the 5α reductase II gene. FASEB J. 2012; 26(10):4360-4372.

Kim JH, Yamaguchi K, Lee SH, Tithof PK, Sayler GS, Yoon JH, Baek SJ. Evaluation of polycyclic aromatic hydrocarbons in the activation of early growth response-1 and peroxisome proliferator activated receptors. Toxicol Sci. 2005; 85:585-593.

King-Jones K, Thummel CS. Nuclear receptors – a prospective from *Drosophila*. Genetics. 2005; 6:311-323.

King-Jones K, Horner MA, Lam G, Thummel CS. The DHR96 nuclear receptor regulates xenobiotic responses in *Drosophila*. Cell Metab. 2006; 4:37-48.

Klaholz BP, Renaud JP, Mitschler A, Zusi C, Chambon P, Gronemeyer H, Moras D. Conformational adaptation of agonists to the human nuclear receptor RAR gamma Nat Struct Biol. 1998; 5:199-202.

Kliewer SA, Umesono K, Noonam DJ, Heyman RA, Evans RM. Convergence of 9-*cis* retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. Nature. 1992; 358:771–774.

Kobayashi M, Yu RT, Yasuda K, Umesono K. Cell-Type-Specific Regulation of the Retinoic Acid Receptor Mediated by the Orphan Nuclear Receptor TLX. Mol Cell Biol. 2000; 20(23):8731-8739.

Koelle MR, Talbot WS, Segraves WA, Bender MT, Cherbas P, Hogness DS. The *Drosophila* EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. Cell. 1991; 67(1):59-77.

Köhler HR, Kloas W, Schirling M, Lutz I, Reye AL, Langen JS, Triebskorn R, Nagel R, Schönfelder G. Sex steroid receptor evolution and signalling in aquatic invertebrates. Ecotoxicology. 2007; 16:131-143.

Koonin EV, Galperin MY. Chapter 3: Information sources for genomics. In: Koonin EV, Galperin MY, editor. Sequence - Evolution - Function: Computational Approaches in Comparative Genomics. Boston: Kluwer Academic; 2003.

Krajniak KG. Annelid endocrine disruptors and a survey of invertebrate FMRFamide-related peptides. Integr Comp Biol. 2005; 45(1):88-96.

Kretschmer XC, Baldwin WS. CAR and PXR: xenosensors of endocrine disrupters? Chem Biol Interact. 2005; 155:111-128.

Kwintkiewicz J, Nishi Y, Yanase T, Giudice LC. Peroxisome proliferator-activated receptor-gamma mediates bisphenol A inhibition of FSH-stimulated IGF-1, aromatase, and estradiol in human granulosa cells. Environ Health Perspect. 2010; 118:400-406

Lafont R, Mathieu M. Steroids in aquatic invertebrates. Ecotoxicology. 2007; 16:109-130.

Laguerre M, Veenstra JA. Ecdysone receptor homologs from mollusks, leeches and a polychaete worm. FEBS Lett. 2010; 584:4458–4462.

Lam G, Hall BL, Bender M, Thummel CS. DHR3 is required for the prepupal–pupal transition and differentiation of adult structures during *Drosophila* metamorphosis. Dev Biol. 1999; 212:204–216.

Lange A, Katsu Y, Miyagawa S, Ogino Y, Urushitani H, Kobayashi T, Hirai T, Shears JA, Nagae M, Yamamoto J, Ohnishi Y, Oka T, Tatarazako N, Ohta Y, Tyler CR, Iguchi T. Comparative responsiveness to natural and synthetic estrogens of fish species commonly used in the laboratory and field monitoring. Aquat Toxicol. 2012; 109:250-258.

Langston WJ, Pope ND, Davey M, Langston KM, O' Hara SC, Gibbs PE, Pascoe PL. Recovery from TBT pollution in English Channel environments: A problem solved? Mar Pollut Bull. 2015; 95(2):551-564.

Larroux C, Fahey B, Liubicich D, Hinman VF, Gauthier M, Gongora M, Green K, Wörheide G, Leys SP, Degnan BM. Developmental expression of transcription factor genes in a demosponge: insight into the origin of metazoan multicellularity. Evol Dev. 2006; 8:150-173.

Latruffe N, Vamecq J, Cherkaoui Malki M. Genetic-dependency of peroxisomal cell functions – emerging aspects. J Cell Mol Med. 2003; 7:238-248.

Laudet V. Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. J Mol Endocrinol. 1997; 19:207-226.

Laudet V and Gronemeyer H. *The Nuclear receptor Facts Book*. San Diego: Academic Press; 2002.

Laudet V, Hänni C, Coll J, Catzeflis C, Stéhelin D. Evolution of the nuclear receptor gene family. EMBO J. 1992; 11:1003–1013.

Laughlin RB, Guard HE, Coleman WM. Tributyltin in seawater: speciation and octanol-water partition coefficient. Enviro. Sci Technol. 1986; 20(2):201-4

Lazar MA, Berrodin TJ, Harding HP. Differential DNA binding by monomeric, homodimeric, and potentially heteromeric forms of the thyroid hormone receptor. Mol Cell Biol. 1991; 11(10):5005-5015.

le Maire A, Grimaldi M, Roecklin D, Dagnio S, Vivat-Hannah V, Balaquer P, Bourguet W. Activation of RXR-PPAR heterodimers by organotin environmental endocrine disruptors. EMBO Rep. 2009; 10:367-373.

Lehmann JM, Moore LB, Smith-Oliver TA, Wilkison WO, Willson TM, Kliewer SA. An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). J Biol Chem. 1995; 270(22):12953-12956.

Leid M, Kastner P, Lyons R, Nakshatri H, Saunders M, Zacharewski T, Chen JY, Staub A, Garnier JM, Mader S, Chambon P. Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently. Cell. 1992; 68(2):377-395.

Leverett D, Thain J. Oyster embryo–larval bioassay (Revised). In: Keizer PD, editor. ICES Techniques in Marine Environmental Sciemce No. 54. Denmark; 2013. pp. 34.

Levin AA, Sturzenbecker LJ, Kazmer S, Bosakowski T, Huselton C, Allenby G, Speck J, Kratzeisen C, Rosenberger M, Lovey A, Grippo JF. 9-*cis* retinoic acid stereoisomer binds and activates the nuclear receptor RXR alpha. Nature. 1992; 355(6358):359-361.

Liang HL, Nien CY, Liu HY, Metzgenstein MM, Kirov N, Rushlow C. The zinc–finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. Nature Letters. 2008; 456:400–404.

Liberato MV, Nascimento AS, Ayers SD, Lin JZ, Cvoro A, Silveira RL, Martinez L, Souza PC, Saidemberg D, Deng T, Amato AA, Togashi M, Hsueh WA, Phillips K, Palma MS, Neves FA, Skaf MS, Webb P, Polikarpov I. Medium chain fatty acids are selective peroxisome proliferator activated receptor (PPAR) Gamma activators and pan-PPAR partial agonists. Plos One. 2012; 7:e36297

Lima D, Reis-Henriques MA, Silva R, Santos AI, Castro LF, Santos MM. Tributyltin-induced imposex in marine gastropods involves tissue-specific modulation of the retinoid X receptor. Aquat Toxicol. 2011; 101:221-227.

Lin S, Han Y, Shi Y, Rong H, Zheng S, Jin S, Lin SY, Lin SC, Li Y. Revealing a steroid receptor ligand as a unique PPARγ agonist. Cell Res. 2012; 22(4):746-756.

Lindblom TH, Pierce GJ, Sluder AE. A *C. elegans* orphan nuclear receptor contributes to xenobiotic resistance. Curr Biol. 2001; 5:864-868.

Love-Gregory L, Permutt MA. HNF4A genetic variants: role in diabetes. Curr Opin Clin Nutr Metab Care. 2007; 10:397–402.

Lutz I, Kloas W. Amphibians as a model to study endocrine disruptors: I. Environmental pollution and estrogen receptor binding. Sci Total Environ. 1999; 225:49–57.

Lv J, Feng L, Bao Z, Guo H, Zhang Y, Jiao W, Zhang L, Wang S, He Y, Hu X. Molecular characterization of RXR (Retinoid X Receptor) gene isoforms from the bivalve species *Chlamys farreri*. PLoS One. 2013; 8(9):e74290.

Maglich JM, Slunder A, Guan X, Shi Y, McKee DD, Carrick K, Kamdar K, Willson TM, Moore JT. Comparison of complete nuclear receptor sets from the human, *Caenorhabditis* elegans and *Drosophila* genomes. Genome Biol. 2001; 2:research0029.1.-research0029.7.

Mangelsdorf DJ, Evans RM. The RXR heterodimers and orphans receptors. Cell. 1995; 83:841–850.

Mangelsdorf D. J., Umesono K., Evans R. M. The retinoid receptors. In: Sporn MB, Roberts AB, Goodman DS, editors. Retinoids: Biology, Chemistry. New York: Raven Press; 1994. P. 319-349.

Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Kastner P, Mark M, Chambon P, Evans RM. The nuclear receptor superfamily: the second decade. Cell. 1995; 83:835–839.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 2011; 39:D225-9.

Markov GV, Laudet V. Origin and evolution of the ligand-binding ability of nuclear receptors. Mol Cell Endocrinol. 2011; 334:21-30.

Marlétaz F, Holland LZ, Laudet V, Schubert M. Retinoic acid signaling and the evolution of chordates. Int J Biol Sci. 2006; 2:38–47.

Matsumoto T, Osada M, Osawa Y, Mori K. Gonadal estrogen profile and immunohistochemical localization of steroidogenic enzymes in the oyster and scallop during sexual maturation. Comp Biochem Physiol B. 1997; 118:811–817.

Matsumoto T, Nakamura AM, Mori K, Akiyama I, Hirose H, Takahashi Y. Oyster estrogen receptor: cDNA cloning and immunolocalization. Gen Comp Endocrinol. 2007; 151:195-201.

Matthiessen P. An assessment of endocrine disruption in mollusks and the potential for developing internationally standardized mollusc life cycle test guidelines. Integr Environ Assess Manag. 2008; 4:274-284.

Matthiessen P, Reynoldson T, Billinghurst Z, Brassard DW, Cameron P, Chandler GT, Davies IM, Horiguchi T, Mount DR, Oehlmann J, Pottinger TG, Sibley PK, Thompson HM, Vethaak AD. Field assessment of endocrine disruption in invertebrates In: DeFur PL, Crane M, Ingersoll C, Tattersfield L, editors. Endocrine disruption in invertebrates: Endocrinology, testing and assessment. Pensacola: SETAC; 1999. p. 199–270.

McClellan–Green PD. Endocrine disruption in molluscs: processes and testing. In: Matthiessen P, editor. Endocrine disrupters: hazard testing and assessment methods. Hoboken: John Wiley & Sons, Inc; 2013. p. 143-184.

McLachlan JA. Environmental signaling: what embryos and evolution teach us about endocrine disrupting chemicals. Endocr Rev. 2001; 22(3):319-341.

McNicholas S, Potterton E, Wilson KS, Noble MEM. Presenting your structures: the CCP4mg molecular-graphics software. Acta Cryst. 2011; D67 386-394.

Mhadhbi L, Rial D, Pérez S, Beiras R. Ecological risk assessment of perfluorooctanoic acid (PFOA) and perfluorooctanesulfonic acid (PFOS) in marine environment using *Isochrysis galbana, Paracentrotus lividus, Siriella armata* and *Psetta maxima*. J Environ Monit. 2012; 14(5):1375-1382.

Michalik L, Desvergne B, Dreyer Ch, Gavillet M, Laurini RN, Wahli W. PPAR expression and function during vertebrate development. Int J Dev Biol. 2002; 46(1):105–114.

Milam AH, Rose L, Cideciyan AV, Barakat MR, Tang WX, Gupta N, Aleman TS, Wright AF, Stone EM, Sheffield VC, Jacobson SG. The nuclear receptor NR2E3 plays a role in human retinal photoreceptor differentiation and degeneration. Proc Natl Acad Sci. 2002; 99:473–478.

Misra P, Chakrabarti R, Vikramadithyan RK, Bolusu G, Juluri S, Hiriyan J, Gershome C, Rajjak A, Kashireddy P, Yu S, Surapureddi S, Qi C, Zhu YJ, Rao MS, Reddy JK, Ramanujam R. PAT5A: a partial agonist of peroxisome proliferator-activated receptor gamma is a potent antidiabetic thiazolidinedione yet weakly adipogenic. J Pharmacol Exp Ther. 2003; 306(2):763-771.

Miyagawa S, Lange A, Hirakawa I, Tohyama S, Ogino Y, Mizutani T, Kagami Y, Kusano T, Ihara M, Tanaka H, Tarazako N, Ohta Y, Katsu Y, Tyler CR, Iguchi T. Differing species responsiveness of estrogenic contaminants in fish is conferred by the ligand binding domain of the estrogen receptor. Environ Sci Technol. 2014; 48(9):5254-5263.

Mlodzik M, Hiromi Y, Weber U, Goodman CS, Rubin GM. The *Drosophila* seven-up gene, a member of the steroid receptor gene superfamily, controls photoreceptor cell fates. Cell. 1990; 60:211–224.

Mohan HM, Aherne CM, Rogers AC, Baird AW, Winter DC, Murphy EP. Molecular pathways: the role of NR4A orphan nuclear receptors in cancer. Clin Cancer Res. 2012; 18:3223-3228.

Moise AR, Alvarez S, Domínguez M, Alvarez R, Golczak M, Lobo GP, von Lintig J, de Lera AR, Palczewski K. Activation of retinoic acid receptors by dihydroretinoids. Mol Pharmacol. 2009; 76(6):1228-1237.

Moriyama K, Tagami T, Akamizu T, Usui T, Saijo M, Kanamoto N, Hataya Y, Shimatsu A, Kuzuya H, Nakao K. Thyroid hormone action is disrupted by bisphenol A as an antagonist. J Clin Endocrinol Metab. 2002; 87(11):5185-5190.

Morton B. Recovery from imposex by a population of the dogwhelk, *Nucella lapillus* (Gastropoda: Caenogastropoda), on the southeastern coast of England since May 2004: a 52-month study. Mar Pollut Bull. 2009; 58(10):1530-1538.

Much JW, Slade DJ, Klampert K, Garriga G, Wightman B. The fax-1 nuclear hormone receptor regulates axon pathfinding and neurotransmitter expression. Development. 2000; 127:703-712.

Nakanishi T. Potential toxicity of organotin compounds via nuclear receptor signaling in mammals. J Health Sci. 2007; 53:1–9.

Nastos A, Pogge van Strandmann E, Weber H, Ryffel GU. The embryonic expression of the tissue–specific transcription factor HNF1alpha in *Xenopus*: rapid activation by HNF4 and delayed induction by mesoderm inducers. Nucleic Acids Res. 1998; 26:5602–5608.

Ng HW, Zhang W, Shu M, Luo H, Ge W, Perkins R, Tong W, Hong H. Competitive molecular docking approach for predicting estrogen receptor subtype α agonists and antagonists. BMC Bioinformatics. 2014; 15 Suppl 11:S4.

Nicolaus EE, Barry J. Imposex in the dogwhelk (*Nucella lapillus*): 22-year monitoring around England and Wales. Environ Monit Assess. 2015; 187(12):736.

Niederreither K, Dolle P. Retinoic acid in development towards an integrated view. Nat Rev Genet. 2008; 9:541–553.

Nishikawa J. Imposex in marine gastropods may be caused by binding of organotins to retinoid X receptor. Mar Biol. 2006; 149:117-124.

Nishikawa J, Mamiya S, Kanayama T, Nishikawa T, Shiraishi F, Horiguchi T. Involvement of the retinoid X receptor in the development of imposex caused by organotins in gastropods. Environ Sci Technol. 2004; 38:6271–6276.

Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, Kurokawa R, Rosenfeld MG, Willson TM, Glass CK, Milburn MV. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. Nature. 1998; 395(6698):137-143.

Norris DO, Lopez KH. Hormones and reproduction of vertebrates. 1st ed. Elsevier: 2013.

Nuclear Receptors Nomenclature Committee. A unified nomenclature system for the nuclear receptor superfamily. Cell. 1999; 97:161–163.

Ohno CK, Ueda H, Petkovich M: The *Drosophila* nuclear receptors FTZ-F1 alpha and FTZ-F1 beta compete as monomers for binding to a site in the fushi tarazu gene. Mol Cell Biol. 1994; 14:3166-3175.

Omiecinski CJ, Vanden Heuvel JP, Perdew GH, Peters JM. Xenobiotic metabolism, disposition, and regulation by receptors: from biochemical phenomenon to predictors of major toxicities. Toxicol Sci. 2011; 120:49-75.

Orbea A, Cajaraville MP. Peroxisome proliferation and antioxidant enzymes in transplanted mussels of four basque estuaries with different levels of polycyclic aromatic hydrocarbon and polychlorinated biphenyl pollution. Environ Toxicol Chem. 2006; 25:1616-26.

Ottow E, Weinmann H. Nuclear receptors as drug targets: a historical perspective of modern drug discovery. In: Ottow E, Weinmann H, editor. Nuclear receptor as drug targets. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2008.

Owen GI and Zelent A. Origins and evolutionary diversification of the nuclear receptor superfamily. Cell Mol Life Sci. 2000; 57:809-827.

Paguio A, Almond B, Fan F, Stecha P, Garvin D, Wood M, Wood K. pGL4 Vectors: A new generation of luciferase reporter vectors. Promega Notes. 2005; 89:7–10.

Palanker L, Tennessen JM, Lam G, Thummel CS. *Drosophila* HNF4 regulates lipid mobilization and beta-oxidation. Cell Metab. 2009; 9:228-239.

Palanker L, Necakov AS, Sampson HM, Ni R, Hu Ch, Thummel CS, Krause HM. Dynamic regulation of *Drosophila* nuclear receptor activity in vivo. Development. 2006; 133:3549–3562.

Park JJ, Shin YK, Hung AAO, Romano N, Cheom YP, Kim JW. Reproductive impairment and intersexuality in *Gomphina veneriformis* (Bivalvia: Veneridae) by the tributyltin compound. Animal Cells System. 2015; 19:61-68.

Parkhaev PY: The Early Cambrian Radiation of Mollusca. In: Ponder WF, Lindberg DR, editors. Phylogeny and evolution of the Mollusca. University of California Press; 2008. p. 33-69.

Pascoal S, Carvalho G, Vasieva O, Hughes R, Cossins A, Fang Y, Ashelford K, Olohan L, Barroso C, Mendo S, Creer S. Transcriptomics and in vivo tests reveal novel mechanisms underlying endocrine disruption in an ecological sentinel, *Nucella lapillus*. Mol Ecol. 2013; 22:1589-1608.

Perissi V, Rosenfeld MG. Controlling nuclear receptors: the circular logic of cofactor cycles. Nat Rev Mol Cell Biol. 2005; 6(7):542-554.

Perissi V, Staszewski LM, McInerney EM, Kurokawa R, Krones A, Rose DW, Lambert MH, Milburn MV, Glass CK, Rosenfeld MG. Molecular determinants of nuclear receptor-corepressor interaction. Genes Dev. 1999; 13(24):3198-3208.

Pfaffl MW. A new mathematical model for relative quantification in real–time RT–PCR. Nucleic Acids Res. 2001; 29(9):e45.

Pignatello MA, Kauffman FC, Levin AA. Multiple factors contribute to the toxicity of the aromatic retinoid, TTNPB (Ro 13-7410): binding affinities and disposition. Toxicol Appl Pharmacol. 1997; 142(2):319-27.

Prakash C, Zuniga B, Song CS, Jiang S, Cropper J, Park S;Chatterjee B. Nuclear receptors in drug metabolism, drug response and drug interactions. Nucl Receptor Res. 2015; 2:101178.

Promega. Technical Manual TM259. pGL4 Luciferase reporter vectors. Revised March 2015.

Puinean AM, Labadie P, Hill EM, Osada M, Kishida M, Nakao R, Novillo A, Callard IP, Rotchell JM. Laboratory exposure to 17β-estradiol fails to induce vitellogenin and estrogen receptor gene expression in the marine invertebrate *Mytilus edulis*. Aquat Toxicol. 2006; 79:376-383.

Qiu Y, Pereira FA, DeMayo FJ, Lydon JP, Tsai SY, Tsai MJ. Null mutation of COUP-TFI results in defects in morphogenesis of the glossopharyngeal ganglion, axonal projection, and arborization. Genes Dev. 1997; 11:1925–1937.

R Development Core Team. R: A language and environment for statistical computing. Vienna: the R Foundation for Statistical Computing; 2008.

Raingeard D, Bilbao E, Cancio I, Cajaraville MP. Retinoid X receptor (RXR), estrogen receptor (ER) and other nuclear receptors in tissues of the mussel *Mytilus galloprovincialis*: Cloning and transcription pattern. Comp Biochem Physiol A Mol Integr Physiol. 2013; 165:178-190.

Rakhshandehroo M, Knoch B, Müller M, Kersten S. Peroxisome proliferator-activated receptor alpha target genes. PPAR Res. 2010; 2010:612089.

Ramakrishnan SN, Muscat GEO. The orphan Rev-erb nuclear receptors: a link between metabolism, circadian rhythm and inflammation? Nucl Recept Signal. 2006; 4:e009.

Rastinejad F, Perlmann T, Evans RM, Sigler PB. Structural determinants of nuclear receptor assembly on DNA direct repeats. Nature. 1995; 375(6528):203-211.

Reizel AM, Tarrant AM. Nuclear receptor complement of the cnidarian *Nematostella vectensis:* phylogenetic relationships and developmental expression patterns. BMC Evo Biol. 2009; 9:320.

Reschly EJ, Krasowski MD. Evolution and function of the NR1I nuclear hormone receptor subfamily (VDR, PXR, and CAR) with respect to metabolism of xenobiotics and endogenous compounds. Curr Drug Metab. 2006; 7:349-365.

Retnakaran A, Krell P, Feng Q, Arif B. Ecdysone agonists: mechanism and importance in controlling insect pests of agriculture and forestry. Arch Insect Biochem Physiol. 2003; 54:187-199.

Rhinn M, Dollé P. Retinoic acid signalling during development. Development. 2012; 139:843–858.

Richter I, fiddler AE. Marine invertebrate xenobiotic-activated nuclear receptors: their application as sensor elements in high-throughput bioassay for marine bioactive compounds. Mar Drugs. 2014; 12(11):5590-5618.

Riddiford LM, Truman JW, Mirth CK, Shen Y. A role for juvenile hormone in the prepupal development of *Drosophila melanogaster*. Development. 2010; 137(7):1117–1126.

Riu A, McCollum CW, Pinto CL, Grimaldi M, Hillenweck A, Perdu E, Zalko D, Bernard L, Laudet V, Balaguer P, Bondesson M, Gustafsson JA. Halogenated bisphenol-A analogs act as obesogens in zebrafish larvae (*Danio rerio*). Toxicol Sci. 2014; 139(1):48-58.

Robinson-Rechavi M, Carpentier AS, Duffraisse M, Laudet V. How many nuclear hormone receptors in the human genome? Trends Genet. 2001; 17:554-556.

Rochette-Egly C, Oulad-Abdelghani M, Staub A, Pfister V, Scheuer I, Chambon P, Gaub MP. Phosphorylation of the retinoic acid receptor-alpha by protein kinase A. Mol Endocrinol. 1995; 9(7):860-871.

Rodríguez EM, Medesani DA, Fingerman M. Endocrine disruption in crustaceans due to pollutants: a review. Comp Biochem Physiol A Mol Integr Physiol. 2007; 146(4):661-671.

Rosenfeld MG, Lunyak VV, Glass CK. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. Genes Dev. 2006; 20(11):1405-1428.

Rotman N, Guex N, Gouranton E, Wahli W. PPARβ Interprets a chromatin signature of pluripotency to promote embryonic differentiation at gastrulation. PLoS One. 2013; 18:e83300.

Roy AK, Lavrovsky Y, Song CS, Chen S, Jung MH, Velu NK, Bi BY, Chatterjee B. Regulation of androgen action. Vitam Horm. 1999; 55: 309–352.

Ruiz JM;Bryan GW, Gibbs PE. Acute and chronic toxicity of tributyltin (TBT) to pediveliger larvae of the bivalve *Scrobicularia plana*. Mar Biol. 1995a; 124:119-126.

Ruiz JM;Bryan GW, Wigham GD, Gibbs PE. Effects of tributyltin (TBT) exposure on the reproduction and embryonic development of the bivalve *Scrobicularia plana.* Mar Envir R. 1995b; 40(4):363-379.

Ruiz JM, Bachelet G, Caumette P, Donard OFX. Three decades of tributyltin in the coastal environment with emphasis on Arcachon Bay, France. Environ. Pol.1996; 93(2):195–203.

Ryffel GU, Lingott A. Distinct promoter elements mediate endodermal and mesodermal expression of the HNF1alpha promoter in transgenic *Xenopus*. Mech Dev. 2000; 90:65–75.

Sáez PJ, Lange S, Pérez-Acle T, Owen GI. Nuclear receptor genes: evolution. In: eLS. John Wiley & Sons Ltd, Chichester, 2010.

Salazar MH. Biological factors influencing bioaccumulation in bivalves. In: Newman MC, Unger MA, editors. Fundamentals of Ecotoxicology. 2nd ed. Florida: CRC Press LLC; 2002. p. 90-94.

Salazar MH, Salazar SM. Mussel as bioindicators: effects of TBT on survival, bioaccumulation, and growth under natural conditions. In: Champ MA, Seligman PF, editors. Organotin: environmental fate and effects. London: Chapman and Hall; 1996. p. 305-330.

Saldanha AJ. Java Treeview – extensible visualization of microarray data. Bioinformatics. 2004; 20(17):3246–3248.

Sanner MF. Python: A Programming Language for Software Integration and Development. J Mol Graphics Mod. 1999; 17:57-61.

Schnoes AM, Brown SD, Dodevski I, Babbitt PC. Annotation Error in Public Databases: Misannotation of Molecular Function in Enzyme Superfamilies. PLoS Comput Biol. 2009; 5(12):e1000605.

Scott AP. Do mollusks use vertebrate sex steroids as reproductive hormones? Part I: Critical appraisal of the evidence for the presence, biosynthesis and uptake of steroids. Steroids. 2012; 77:1450-1468.

Scott AP. Do mollusks use vertebrate sex steroids as reproductive hormones? II. Critical review of the evidence that steroids have biological effects. Steroids. 2013; 78:268-281.

Segner H. Reproductive and developmental toxicity in fishes. In: Gupta RC, editor. Reproductive and developmental toxicology. Amsterdam: Elsevier; 2011. p. 1145-1166.

Shi HH, Huang CJ, Zhu SX, Yu XJ, Xie WY. Generalized system of imposex and reproductive failure in female gastropods of coastal waters of mainland China. Mar Ecol Progr Ser. 2005; 304:179–189.

Shifera AS, Hardin JA. Factors modulating expression of *Renilla* luciferase from control plasmids used in luciferase reporter gene assays. Anal Biochem. 2010; 396(2):167-172.

Shih TW, Lin TH, Shealy YF, Hill DL. Non-enzymatic isomerization of 9-*cis*-retinoic acid catalyzed by sulfhydryl compounds. Drug Metab Dispos. 1997; 25: 27–32

Shmarakov IO. Retinoid-xenobiotic interactions: the Ying and the Yang. Hepatobiliary. Surg Nutr. 2015; 4(4):243–267.

Singh MP, Pathak D, Sharma GK, Sharma CS. Peroxisome proliferator-activated receptors (PPARs): A target with a broad therapeutic potential for human disease: an overview. Pharmacologyonline 2011; 2:58-89.

Sladek FM. What are nuclear receptor ligands? Mol Cell Endocrinol. 2011; 334(1-2):3-13.

Soprano DR, Soprano KJ. Retinoids as teratogens. Annu Rev Nutr. 1995; 15:111-132.

Sousa ACA, Barroso CM, Tanabe S, Horiguchi T. Involvement of retinoid X receptor in imposex development in *Nucella lapillus* and *Nassarius reticulatus* – preliminary results. In: Hamamura N, Suzuki S, Mendo S, Barroso CM, Iwata H, Tanabe S, editors. 189 Interdisciplinary Studies on Environmental Chemistry — Biological Responses to Contaminants. Terrapub; 2010. p.189-196.

Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, Kuo A, Mitros T, Salamov A, Carpenter ML, Signorovitch AY, Moreno MA, Kamm K, Grimwood J, Schnutz J, Shapiro H, Grigoriev IV, Buss LW, Schierwater B, Dellaporta SL, Rokhsar DS. The *Trichoplax* genome and the nature of placozoans. Nature. 2008; 454:955-960.

Stange D, Sieratowicz A, Horres R, Oehlmann J. Freshwater mudsnail (*Potamopyrgus antipodarum*) estrogen receptor: Identification and expression analysis under exposure to (xeno-)hormones. Ecotoxicol Environ Saf. 2012; 75:94-101.

Sternberg RM, Hotchkiss AK, Leblanc GA. Synchronized expression of Retinoid X Receptor mRNA with reproductive tract recrudescence in an imposex-susceptible mollusc. Environ Sci Technol. 2008; 42:1345-1351.

Sternberg RM, Hotchkiss AK, LeBlanc GA. The contribution of steroidal androgens and estrogens to reproductive maturation of the eastern mud snail *Ilyanassa obsoleta*. Gen Comp Endocrinol. 2008a; 156:15-26.

Sternberg RM, Gooding MP, Hotchkiss AK, LeBlanc GA. Environmental-endocrine control of reproductive maturation in gastropods: implications for the mechanism of tributyltin-induced imposex in prosobranchs. Ecotoxicology. 2010; 19:4-23.

Stone BL, Thummel CS. The *Drosophila* 78C early late puff contains E78 and ecdysone-inducible gene that encodes a novel member of the nuclear hormone receptor superfamily. Cell. 1993; 75:307-320.

Strobl–Mazzulla PH, Lethimonier Ch, Gueguen MM, Karube M, Fernandino JI, Yoshizaki G, Patino R, Strüssmann CA, Kah O, Somozoa GM. Brain aromatase (Cyp19A2) and estrogen receptors, in larvae and adult pejerrey fish Odontesthes bonariensis: neuroanatomical and functional relations. Gen Comp Endocrinol. 2008; 158:191–201.

Suárez–Ulloa V, Fernández–Tajes J, Manfrin C, Gerdol M, Venier P, Eirín–López JM. Bivalve Omics: state of the art and potential applications for the biomonitoring of harmful marine compounds. Mar Drugs. 2013; 11:4370–4389.

Sullivan AA, Thummel CS. Temporal profiles of nuclear receptor gene expression reveal coordinate transcriptional responses during *Drosophila* development. Mol Endocrinol. 2003; 17:2125-2137.

Sun G, Yu RT, Evans RM, Shi Y. Orphan nuclear receptor TLX recruits histone deacetylases to repress transcription and regulate neural stem cell proliferation. Proc Natl Acad Sci. 2007; 104:15282–15287.

Sun J, Narayanasamy S, Curley RW, Harrison EH. β-Apo-13-carotenone regulates retinoid X receptor transcriptional activity through tetramerization of the receptor. J Biol Chem. 2014; 289(48):33118-33124.

Suzuki T, Nakagawa Y, Takano I, Yaguchi K, Yasuda K. Environmental fate of bisphenol A and its biological metabolites in river water and their xeno-estrogenic activity. Environ Sci Technol. 2004; 38:2389–2396

Szantos A, Narkar V, Shen Q, Uray IP, Davis PJA, Nagy L. Retinoid X receptors: X–ploring their (patho)physiological functions. Cell Death Diff. 2004; 11:S126–S143.

Tandros W, Lipshitz HD. The maternal–to–zygotic transition: a play in two acts. Development. 2009; 136:3033–3042.

Taneja R. Nuclear receptors in development. San Diego: Elsevier; 2006.

Tennessen JM, Baker KD, Lam G, Evans J, Thummel CS. The *Drosophila* estrogen–related receptor directs a metabolic switch that supports developmental growth. Cell Metabol. 2011; 13:139–148.

Teuten EL1 Saquing JM, Knappe DR, Barlaz MA, Jonsson S, Björn A, Rowland SJ, Thompson RC, Galloway TS, Yamashita R, Ochi D, Watanuki Y, Moore C, Viet PH, Tana TS, Prudente M, Boonyatumanond R, Zakaria MP, Akkhavong K, Ogata Y, Hirai H, Iwasa S, Mizukawa K, Hagino Y, Imamura A, Saha M, Takada H. Transport and release of chemicals from plastics to the environment and to wildlife. Philos Trans R Soc Lond B Biol Sci. 2009; 364(1526):2027-2045.

Thain JE. Toxicity of TBT to Bivalves: effects on reproduction growth and survival. In: Proceedings of the Oceans ′86 Organotin Symposium, Vol. 4: 23-25 Sept. 1986, Washington D.C. New York: The Institute of Electrical and Electronics Engineers and Marine Technology Society; 1986. p. 1306-1313.

Thain JE, Waldock MJ. The impact of tributyl tin (TBT) antifouling paints on molluscan fisheries. Wat Sci Tech. 1986; 18:193-202.

The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature. 2000; 408:796-815.

Thirunavukkarasu K, Miles RR, Halladay DL, Onyia JE. Cryptic enhancer elements in luciferase reporter vectors respond to the osteoblast-specific transcription factor Osf2/Cbfa1. Biotechniques. 2000; 28(3):506-510.

Thomson SA, Baldwin WS, Wang YH, Kwon G, LeBlanc G. Annotation, phylogenetics, and expression of the nuclear receptors in *Daphnia pulex*. BMC Genomics. 2009; 10:500.

Thornton JW, Need E, Crews D. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science. 2003; 301:1714-1717.

Thummel CS. Molecular mechanisms of developmental timing in *C. elegans* and *Drosophila*. Dev Cell. 2001; 1(4):453-465.

Timsit YE, Negishi M. CAR and PXR: the xenobiotic-sensing receptors. Steroids. 2007; 72(3):231-246.

Titley-O'Neal CP, Munkittrick KR, Macdonald BA. The effects of organotin on female gastropods. J Environ Monit. 2011; 13(9):2360-2388.

Tohmé M, Prud'homme SM, Boulahtouf A, Samarut E, Brunet F, Bernard L, Bourguet W, Gibert Y, Balaguer P, Laudet V. Estrogen-related receptor γ is an *in vivo* receptor of bisphenol A. FASEB J. 2014; 28(7):3124-3133.

Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010; 31(2):455-461.

Tsunemasa N, Okamura H. Effects of organotin alternative antifoulants on oyster embryo. Arch Environ Contam Toxicol. 2011; 61:128-134.

Ueda H, Sun GC, Murata T, Hirose S. 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. Mol Cell Biol. 1992; 12:5667-5672.

UNESCO. Facts and figures on marine pollution. 2016 http://www.unesco.org/new/en/natural-sciences/ioc-oceans/priority-areas/rio-20-ocean/blueprint-for-the-future-we-want/marine-pollution/facts-and-figures-on-marine-pollution/ (Accessed on 26 May 2016).

Urbach J, Rando RR. Isomerization of all-*trans*-retinoic acid to 9-*cis*-retinoic acid. Biochem J. 1994; 299:459–465.

Urushitani H, Katsu Y, Ohta Y, Shiraishi H, Iguchi T, Horiguchi T. Cloning and characterization of retinoid X receptor (RXR) isoforms in the rock shell, *Thais clavigera.* Aquat Toxicol. 2011; 103:101-111.

Urushitani H, Katsu Y, Ohta Y, Shiraishi H, Iguchi T, Horiguchi T. Cloning and characterization of the retinoic acid receptor–like protein in the rock shell, *Thais clavigera*. Aquat Toxicol. 2013; 142–143:403–413.

Vanden Heuvel JP. Nuclear hormone receptors: a brief overview. Stage College PA: INDIGO Bioscience Inc.; 2009. http://nrresource.org/reprintspublications/research_article_nr_general.pdf (Accessed 26 May 2016)

Vanden Heuvel JP, Thompson JT, Frame SR, Gillies PJ. Differential activation of nuclear receptors by perfluorinated fatty acid analogs and natural fatty acids: a comparison of human, mouse, and rat peroxisome proliferator-activated receptor-alpha, -beta, and -gamma, liver X receptor-beta, and retinoid X receptor-alpha. Toxicol Sci. 2006; 92(2):476-89.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. Accurate normalisation of real–time quantitative RT–PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002; 3(7):research0034.1.

Lê S, Josse J, Husson F. FactoMineR: An R package for multivariate Analysis. J Stat Soft. 2008; 25(1):1-18.

Vierke L, Staude C, Biegel-Engler A, Drost W, Schulte C. Perfluorooctanoic acid (PFOA) — main concerns and regulatory developments in Europe from an environmental point of view. Environ Sci Eur. 2012; 24:16.

Vignali M, Hassan AH, Neely KE, Workman JL. ATP-Dependent Chromatin-Remodeling Complexes. Molecular and Cellular Biology. 2000; 20(6):1899-1910.

Vo N, Goodman RH. CREB-binding protein and p300 in transcriptional regulation. J Biol Chem. 2001; 276(17):13505-13508.

Vogeler S, Bean TP, Lyons BP, Galloway TS. Dynamics of nuclear receptor gene expression during Pacific oyster development.' BMC Developmental Biology (accepted).

Vogeler S, Galloway TS, Isupov M, Bean TP. Characterisation of Pacific oyster nuclear receptors, retinoic X receptor, retinoic acid receptor and peroxisome proliferator-activated receptor. Aquatic Toxicology (submitted).

Vogeler S, Galloway TS, Lyons BP, Bean TP. The nuclear receptor gene family in the Pacific oyster, *Crassostrea gigas*, contains a novel subfamily group. BMC Genomics. 2014; 15:369.

Waldock MJ, Thain JE. Shell thickening in *Crassostrea gigas*: organotin antifouling or sediment induced? Mar Pollut Bull. 1983; 14:411-415.

Wan YJ, Cai Y, Magee TR. Retinoic acid differentially regulates retinoic acid receptor-mediated pathways in the Hep3B cell line. Exp Cell Res. 1998; 238(1):241-7.

Wan YJ, Wang L, Wu TC. The expression of retinoid X receptor genes is regulated by all-*trans*- and 9-*cis*-retinoic acid in F9 teratocarcinoma cells. Exp Cell Res. 1994; 210(1):56-61.

Weis JS. Physiological, developmental and behavioural effects of marine pollution. Springer Netherlands; 1994.

Widdows J, Page DS. Effects of tributyltin and dibutyltin on the physiological energetics of the mussel, *Mytilus edulis*. Mar Environ Res. 1993; 35:233-249.

Wiens M, Batel R, Korzhev M, Müller WE. Retinoid X receptor and retinoic acid response in the marine sponge *Suberites domuncula.* J Exp Biol. 2003; 206:3261-3271.

Wilson TE, Fahrner TJ, Milbrandt J. The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. Mol Cell Biol. 1993; 13(9):5794-5804.

Winnepenninckx BM, Reid DG, Backeljau T. Performance of 18S rRNA in littorinid phylogeny (Gastropoda: Caenogastropoda). J Mol Evol. 1998; 47:586-596.

Wu W, LoVerde PT. *Schistosoma mansoni*: Identification of SmNR4A, a member of nuclear receptor subfamily 4. Exp Parasitol. 2008; 120:208-213.

Wu TC, Wang L, Wan YJ. Retinoic acid regulates gene expression of retinoic acid receptors alpha, beta and gamma in F9 mouse teratocarcinoma cells. Differentiation. 1992; 51(3):219-224.

Wu W, Niles EG, Hirai H, LoVerde PT. Evolution of a novel subfamily of nuclear receptors with members that each contain two DNA binding domains. BMC Evol Biol. 2007; 7:27.

Wu W, Niles EG, El-Sayed N, Berriman M, LoVerde PT. *Schistosoma mansoni* (Platyhelminthes, Trematoda) nuclear receptor: Sixteen new members and a novel subfamily. Gene. 2006; 366:303-315.

Wurtz JM, Bourguet W, Renaud JP, Vivat V, Chambon P, Moras D, Gronemeyer H. A canonical structure for the ligand-binding domain of nuclear receptors. Nat Struct Biol. 1996; 3(1):87-94.

Xu J, Tan A, Palli SR. The function of nuclear receptors in regulation of female reproduction and embryogenesis in the red flour beetle, *Tribolium castaneum*. J Insect Physiol. 2010; 56(10):1471-1480.

Yamashita N, Kannan K, Taniyasu S, Horii Y, Petrick G, Gamo T. A global survey of perfluorinated acids in oceans. Mar Pollut Bull. 2005; 51(8-12):658-668.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T. Primer–BLAST: a tool to design target–specific primers for polymerase chain reaction. BMC Bioinformatics. 2012; 13:134.

Yen PM. Physiological and molecular basis of thyroid hormone activation. Physiol Rev. 2001; 81:1097-1142.

Zechel C, Shen XQ, Chambon P, Gronemeyer H. Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements. EMBO J. 1994a; 13(6):1414–1424.

Zechel C, Shen XQ, Chen JY, Chen ZP, Chambon P, Gronemeyer H. The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. EMBO J. 1994b; 13:1425–1433.

Zelhof AC, Yao TP, Evans RM, McKeown M. Identification and characterization of a *Drosophila* nuclear receptor with the ability to inhibit the ecdysone response. Proc Natl Acad Sci. 1995; 92:10477–10481.

Zhang XK, Lehmann J, Hoffmann B, Dawson MI, Cameron J, Graupner G, Hermann T, Tran P, Pfahl M. Homodimer formation of retinoid X receptor induced by 9-*cis* retinoic acid. Nature. 1992; 358(6387):587-591.

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PW, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Lošo T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Wang J, Chen M, He Y, Tan F, Song X, Zheng Q, Huang R, Yang H, Du X, Chen L, Yang M, Gaffney PM, Wang S, Luo L, She Z, Ming Y, Huang W, Zhang S, Huang B, Zhang Y, Qu T, Ni P, Miao G, Wang J, Wang Q, Steinberg CE, Wang H, Li N, Qian L, Zhang G, Li Y, Yang H, Liu X, Wang J, Yin Y, Wang J: The oyster genome reveals stress adaptation and complexity of shell formation. Nature. 2012; 490:49-54.

Zhang H, Pan L, Zhang L. Molecular cloning and characterization of estrogen receptor gene in the Scallop *Chlamys farreri*: Expression profiles in response to endocrine disrupting chemicals. Comp Biochem Physiol C Toxicol Pharmacol. 2012a; 156:51-57.

Zhao Y, Bruemmer D. NR4A orphan nuclear receptors: transcriptional regulators of gene expression in metabolism and vascular Biology. Arterioscler Thromb Vasc Biol. 2010; 30:1535-1541.

Zhong W, Sladek FM, Darnell JE, Jr.. The expression pattern of a *Drosophila* homolog to the mouse transcription factor HNF–4 suggests a determinative role in gut formation. Embo J. 1993; 12:537–544.

Zhou C, Tsai SY, Tsai MJ. COUP-TFI: an intrinsic factor for early regionalization of the neocortex. Genes Dev. 2001; 15:2054–2059.

Zhou Q, Zhang J, Fu J, Shi J, Jiang G. Biomonitoring: an appealing tool for assessment of metal pollution in the aquatic ecosystem. Analytical Chimc Acta. 2008; 606:135–150.

Zimmerman RJ, Minello TJ, Baumer TJ, Castiglione MC. Oyster reef as habitat for estuarine macrofauna. NOAA Technical Memorandum, NMFS-SEFC-249,; 1989. pp. 16.

Zoeller RT, Bansal R, Parris C. Bisphenol-A, an environmental contaminant that acts as a thyroid hormone receptor antagonist in vitro, increases serum thyroxine, and alters RC3/neurogranin expression in the developing rat brain. Endocrinology. 2005; 146:607–612.

Zorita I, Apraiz I, Ortiz-Zarragoitia M, Orbea A, Cancio I, Soto M, Marigómez I, Cajaraville MP. Assessment of biological effects of environmental pollution along the NW Mediterranean Sea using mussels as sentinel organisms. Environ Pollut. 2007; 148:236-250.