Intersexuality and Endocrine Disruption in the Amphipod Echinogammarus marinus – From Genes to Physiology

Yang, Gongda

DOCTOR OF PHILOSOPHY (AWARDED BY OU/ABERDEEN)

Award date:
2013

Awarding institution:
The University of Edinburgh

Link URL to thesis in UHI Research Database

General rights and usage policy
Copyright, IP and moral rights for the publications made accessible in the UHI Research Database are retained by the author, users must recognise and abide by the legal requirements associated with these rights. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement, or without prior permission from the author.

Users may download and print one copy of any thesis from the UHI Research Database for the not-for-profit purpose of private study or research on the condition that:

1) The full text is not changed in any way
2) If citing, a bibliographic link is made to the metadata record on the the UHI Research Database
3) You may not further distribute the material or use it for any profit-making activity or commercial gain
4) You may freely distribute the URL identifying the publication in the UHI Research Database

Take down policy
If you believe that any data within this document represents a breach of copyright, confidence or data protection please contact us at RO@uhi.ac.uk providing details; we will remove access to the work immediately and investigate your claim.

Download date: 24. Aug. 2020
Intersexuality and Endocrine Disruption in the Amphipod *Echinogammarus marinus* – From Genes to Physiology

A THESIS PRESENTED FOR THE DEGREE OF PH.D AT THE UNIVERSITY OF ABERDEEN

by

Gongda Yang

December 2011
# ABSTRACT


# CHAPTER 1 GENERAL INTRODUCTION

## 1.1 BACKGROUND INTRODUCTION

## 1.2 BIOLOGY OF CRUSTACEANS

### 1.2.1 Sex determination in crustaceans

### 1.2.2 Endocrinology of crustaceans

### 1.2.3 Male reproductive system in amphipods

### 1.2.4 Biology of Echinogammarus marinus

## 1.3 ENDOCRINE DISRUPTION

### 1.3.1 Endocrine disruption in crustaceans

### 1.3.2 Intersexuality in crustaceans

## 1.4 MICROSPORIDIAN PARASITE

## 1.5 MOLECULAR TECHNIQUES EXPLORING TRANSCRIPTOME PROFILING

### 1.5.1 DNA Microarray

#### 1.5.1.1 General Introduction of Microarray

#### 1.5.1.2 DNA microarray studies in crustaceans

### 1.5.2 DNA Sequencing

#### 1.5.2.1 A Brief Review of DNA Sequencing Techniques

#### 1.5.2.2 Roche GS FLX (454) Sequencing

## 1.6 PROJECT OBJECTIVES

# CHAPTER 2 INTERSEXUALITY AND MALE FERTILITY IN ECHINOGRAMMARUS MARINUS

## 2.1 INTRODUCTION

## 2.2 MATERIALS AND METHODS

### 2.2.1 Field Collection of Amphipod Specimens

#### 2.2.1.1 Sampling of E. marinus in Three Scottish Sites

#### 2.2.1.2 Sampling of E. marinus from Two Portsmouth Sites

#### 2.2.1.3 Sampling of E. marinus in Inverkeithing for Moul Cycle Monitoring

### 2.2.2 Moul Cycle Monitoring in Normal and Intersex E. marinus
2.2.3 Sperm Count....................................................................................................................63

2.3 RESULTS........................................................................................................................................65

2.3.1 Sex Ratio and Male Intersexuality in Five E. marinus Populations in the UK........65
  2.3.1.1 Sex Ratio and Male Intersexuality in Three Scottish Sites.................................65
  2.3.1.2 Prevalence of male intersexuality in Langstone Harbour and Portsmouth Harbour

E. marinus populations.................................................................................................................66

2.3.2 Body Weight Comparison Between Normal and Intersex Males.................................67

2.3.3 Molt Frequency in Normal and Intersex E. marinus......................................................71

2.3.4 Sperm Count Comparison..................................................................................................74
  2.3.4.1 Sperm Counts in Normal and Intersex Males from Three Scottish E. marinus

Populations....................................................................................................................................74
  2.3.4.2 Sperm Counts of two E. marinus Populations in Portsmouth...............................78

2.4 DISCUSSION............................................................................................................................81

CHAPTER 3 MICROSPORIDIAN INFECTION IN E. MARINUS..................................................89

3.1 INTRODUCTION.....................................................................................................................90

3.2 MATERIALS AND METHODS..............................................................................................95

3.2.1 Specimen Collection...........................................................................................................95
  3.2.1.1 E. marinus Samples for Microsporidian Identification by Histology.......................95
  3.2.1.2 Sampling for Parasite Screening in Inverkeithing E. marinus Population.............95
  3.2.1.3 Sampling for Parasite Screening in Portsmouth E. marinus Populations...............96
  3.2.1.4 E. marinus for Microsporidian Infection Quantification by QPCR......................96

3.2.2 Histology..........................................................................................................................97

3.2.3 RNA Extraction..................................................................................................................97
  3.2.3.1 RNA Extraction Using Tri-reagent............................................................................98
  3.2.3.2 RNA Extraction Combing Tri-reagent and RNeasy Plus Micro Kit (Qiagen)........99

3.2.4 Single-Stranded cDNA Synthesis......................................................................................99

3.2.5 Genomic DNA Extraction.................................................................................................100

3.2.6 Nested PCR Based Parasite Screening Method.............................................................101

3.2.7 PCR Products Cloning........................................................................................................104
3.2.7.1 PCR Products Purification ......................................................... 104
3.2.7.2 Ligation of PCR Products to pGEMT Easy Vector ..................... 104
3.2.7.3 Plasmid Transformation .............................................................. 105
3.2.8 DNA Sequencing .......................................................................... 105
3.2.9 Microsporidian Infection Intensity Quantification by QPCR .............. 106

3.3 RESULTS .......................................................................................... 108
3.3.1 Microsporidian infection identified by histology ................................. 108
3.3.2 Microsporidian infection and parasite species identification in Inverkeithing E. marinus population ................................................................. 111
   3.3.2.1 Microsporidian Infection Ratio in Four Sexual Phenotypes .......... 111
   3.3.2.2 Stochasticism of Nested PCR .................................................... 112
   3.3.2.3 PCR Products Sequencing ......................................................... 114
   3.3.2.4 PCR Products Cloning and Sequencing .................................... 117
3.3.3 Microsporidian Infection in an E. marinus Population in Portsmouth Harbour ....... 120
   3.3.3.1 Sex-Phenotype Ratio of E. marinus in Portsmouth ..................... 120
   3.3.3.2 Microsporidian Infection in Normal and Intersex E. marinus .......... 121
   3.3.3.3 Microsporidian Species Identification ........................................ 124
   3.3.3.4 Differential Distribution of the Two Microsporidian Parasites .......... 126
3.3.4 Microsporidian Bi-infection in Inverkeithing and Portsmouth E. marinus ........ 130
3.3.5 Microsporidian Infection Intensity Quantification by QPCR ................. 136

3.4 DISCUSSION ...................................................................................... 138

CHAPTER 4 TRANSCRIPTOMIC COMPARISON OF NORMAL AND INTERSEX MALE E. MARINUS USING A CROSS-SPECIES DNA MICROARRAY ........................................ 148

4.1 INTRODUCTION .................................................................................. 149
4.2 MATERIALS AND METHODS .............................................................. 153
   4.2.1 Sampling ..................................................................................... 153
   4.2.2 RNA Purification ........................................................................ 153
   4.2.3 DNA Microarray .......................................................................... 154
   4.2.3.1 First-stranded cDNA Synthesis and Labeling ............................ 154
4.2.3.2 cDNA Purification.................................................................155
4.2.3.3 CyDye and cDNA Coupling......................................................156
4.2.3.4 CyDye Labeled cDNA Purification and Quality Assessment ..........157
4.2.3.5 Preparing Probes for Microarray Hybridisation and Array Pre-Wash ....160
4.2.3.6 Microarray Hybridisation........................................................161
4.2.3.7 Microarray Post-Wash............................................................162
4.2.3.8 Scanning and Fluorescent Quantification......................................163
4.2.3.9 Microarray Data Analysis.........................................................164
4.2.4 Microarray Validation by Quantitative-PCR........................................165

4.3 RESULTS ....................................................................................166
4.3.1 Total RNA and Labeled Probes Purification.........................................166
4.3.2 Microarray Calibration and Data Analysis............................................168
4.3.3 Genes Showing Differential Expression Amongst Three Male Phenotypes......171
   4.3.3.1 ANOVA, Volcano plot and PCA analysis........................................171
   4.3.3.2 Annotation of Genes Showing Differential Expressions Amongst Three male Phenotypes.........................................................178

4.2 DISCUSSION.............................................................................186

CHAPTER 5 TRANSCRIPTOME SEQUENCING IN NORMAL AND INTERSEX E. MARINUS BY GS FLX (454).................................................................190

5.1 INTRODUCTION........................................................................191
5.2 MATERIALS AND METHODS.........................................................194
   5.2.1 Sampling...............................................................................194
   5.2.2 RNA Purification.................................................................194
   5.2.3 Microsporidian Screening and RNA Samples Preparation for GS FLX Sequencing ....195
   5.2.4 RNA Amplification...............................................................195
   5.2.4.1 First-Strand cDNA Synthesis..................................................195
   5.2.4.2 Second-Strand cDNA Synthesis.............................................196
   5.2.4.3 cDNA Purification.............................................................196
   5.2.4.4 In Vitro RNA Amplification................................................197
Abstract

The intertidal amphipod (*Echinogammarus marinus* LEACH) exhibits several forms of intersexuality (intersex males and intersex females) and has been considered an ideal model to study reproductive endocrine disruption in Crustacea. This study aimed to investigate both the physiology and gene expression in intersex *E. marinus* with the objective of better understanding their reproductive biology and aid the development of biomarkers of de-masculinisation in Crustacea. *E. marinus* populations from three Scottish (Thurso, Inverkeithing and Loch Fleet) and two English sites (Portsmouth Harbour and Langstone Harbour) were assessed for intersexuality prevalence, sperm counts or microsporidian infection rate. Increased incidence of intersexes and reduced sperm counts were revealed in amphipods from industrial impacted sites. The microsporidian species infecting *E. marinus* were identified for the first time, and *Dictyocoela duebenum* and *Dictyocoela berillonum* were found to be dominant in Inverkeithing and Portsmouth Harbour *E. marinus* populations, respectively. Microsporidian has been reported as a potential factor inducing intersexes in amphipods, and this study also revealed a strong association between microsporidian infection and intersexes in *E. marinus*. A cross-species cDNA microarray was used to characterise the gene expressions of three male phenotypes (normal males, internal and external intersex males), and PCA analysis clearly differentiates the three groups into three separate patches. The *de novo* transcriptome sequencing was carried out on *E. marinus* gonadal tissue, by
employing Roche 454 pyrosequencing producing one of the largest cDNA libraries for a crustacean species. A total of 12,645 gonadal contigs were assembled from 213,212 sequencing reads, and 1206, 1745 and 782 contigs were found to be male-, female- and intersex-specific genes, respectively. This study revealed that the majority of internal intersex males were without microsporidian infection, indicating multiple factors causing intersexuality in *E. marinus*. Molecular studies identified differential gene expression patterns amongst sexual phenotypes and suggest the various forms of intersexuality in *E. marinus* are different at both the biological and genetic level. A large number of strongly male and female sex biased gene sequences have been identified which will provide a powerful resource for future studies into the reproductive biology of crustaceans.
Chapter 1

General Introduction
1.1 Background Introduction

In recent decades, there have been increasing concerns on the impacts of endocrine disrupting chemicals (EDCs), which have been reported to have the ability to alter the function of endocrine system of wildlife (Colborn et al, 1993; Guillette et al, 1995; Tyler et al, 1998; LeBlanc, 2007; Rodríguez et al, 2007). The endocrine system, also known as the hormone system which can be found in the majority of animal groups, plays a critical role on regulating various important bodily functions, such as sex differentiation, development and reproduction. Great concerns have been arisen on EDCs, because of their adverse effects of reducing the fertility and fecundity, and could lead the population to extinction, and subsequently trigger a cascade of ecological issues (Cooper and Kavlock 1997; Jobling and Tyler, 2006).

To date, the majority of researches on endocrine disruption have centered on vertebrates, such as fish, amphibian, alligators and mammals (Guillitte et al, 1994; Lye et al, 1998; Choi and Jeung, 2002; Hayes et al, 2002). Due to the sparse knowledge on endocrinology, studies on endocrine disruption in invertebrates have been restricted in a limited number of animal groups such as insects, molluscs and crustaceans (Au et al, 2000; Dunn et al, 2006; Price et al, 2008). The studies upon molluscs documented one of the best known EDCs in invertebrates -- Tributyltin (TBT), which is found to be responsive for the incidence of imposex in molluscs. Penis-structure has been found in female dogwhelk (Nucella lapillus) since 1970s.
(Blaber, 1970; Smith, 1971). Smith (1981) pointed out that tributyltin, the chemical contained in the antifouling paints, is responsive for the masculinization of female molluscs, and the association between imposex individuals and TBT has been reported in various molluscs populations (Gibbs and Bryan, 1986; Horiguchi et al, 1995; Mattheissen and Gibbs, 1998). Intersexuality, an atypical sex-phenotype having both male and female secondary sex characteristics in one gonochoristic individual, has been widely reported in crustaceans, such as lobsters, crabs and amphipods (Bulnheim, 1965; IEH, 1999; Ginsburger-Vogel, 1991; Sangalang et al, 1997; Zou and Fingerman, 2000), and has been considered as a great model to study endocrine disruption in crustaceans (Ford et al, 2004a). Although the mechanism of intersexuality in Crustacea is not fully understood in all species, it provides an opportunity to study dysfunctional endocrine system in invertebrates.

This study employs an intertidal amphipod, *Echinogammarus marinus*, as the model species, with the aim to develop biomarkers for intersexuality or endocrine disruption in crustaceans. In this study, the cost of intersexuality on male fecundity was investigated, by comparing the sperm counts between normal and intersex male amphipods. The transcriptome profilings of normal and intersex *E. marinus* were compared by employing both DNA microarray and *de novo* transcriptome sequencing (Roche 454 GS-FLX sequencing platform). Potential molecular biomarkers for de-masculinisation in amphipods can be developed based on genes showing differential expression levels between normal and intersexes.
1.2 Biology of crustaceans

Crustaceans make up a subphylum under the largest phylum of Arthropoda. There are over 66,000 species of crustaceans described and documented (LeBlanc, 2007). Crustacea has been further classified into five recognized classes, and three classes were commonly met, namely Branchiopoda, Maxillopoda and Malacostraca (LeBlanc, 2007). Crustaceans widely distribute in various environments (Rodríguez et al, 2007). Apart from the limited number of terrestrial species, most crustaceans live in aquatic environments, such as rivers, lakes, estuaries and the sea. They make up one of the most important groups in aquatic ecology. The majority of crustaceans comprise the base of the ecosystem, and they are the preys of many high level consumers, such as fish, birds and mammals (Sommer et al, 2002).

The majority of crustaceans have segmented body, and most of them can be divided into three main parts: head, thorax and abdomen. Usually, there are two pairs of antennae, one pair of eye and three pairs of mouthparts on the head. Thorax and abdomen bear several pairs of appendages, including gills (in aquatic species) and brood plates (in female animals). Almost all of the crustaceans have a hard exoskeleton, which provides them a protection against dangers in the environment, for example predators. Due to the importance of moult behaviour, a lot of bodily functions have been reported to synchronise to the moult stage in crustaceans, especially the reproductive activities (Patel and Crisp, 1961; Nelson, 1991; McCurdy...
et al, 2000; Okumura and Hara, 2004; Rodríguez et al, 2007). Some of essential biological characters in crustaceans are reviewed in the following sections.

1.2.1 Sex determination in crustaceans

The gender of a variety of animals is determined genetically, by the number of sex chromosomes (Mittwoch, 1967). There are mainly two heterogametic systems in nature, XY/XX and ZZ/ZW, depending on the two identical sex-chromosomes determine female or male (Ezaz et al, 2006). Although both mammals and insects apply the XY/XX system, the mechanisms of sex determination are slightly different. In mammals, the SRY gene on the Y chromosome is essentially involved the process of sex determination (McElreavey et al, 1993). While in insects, the sex is determined by the balance of autosomes and sex chromosomes (Bridge, 1925). There are female-determination factors in the sex chromosome X, and male-determination factors in the autosomes. Therefore, an individual with the sex chromosome of XO will be a female in mammals, but a male in insects. Apart from the sex determination by the sex chromosome, environmental factors, such as temperature and photoperiod, have been discovered to determine the gender in some animal groups, such as fish, reptile and crustaceans (Bulnheim, 1978; Adams et al, 1987; Janzen and Paukstis, 1991; Baroiller and D'Cotta, 2001).

The mechanism of sex determination in crustaceans is not fully understood. Genetic, cytoplasmic and environmental factors are all presumed to contribute to sex
determination in crustaceans (Bulnheim, 1978). The gender of isopod has been reported to under control of the sex chromosome, and two heterogametic systems (XY/XX & ZZ/ZW) have been reported in isopods (Rigaud et al, 1997). Feminising parasites, such as the bacteria Wolbachia, have been found to influence the sex determination in isopods (Juchault et al, 1992). Palmer (1926) revealed two heterochromosomes out of 26 chromosomes found in spermatogonia from male amphipod *Gammarus chevreuxi*. The author suggested these two heterochromosomes might be sex chromosomes X and Y, but experiments supporting the hypothesis were missing. To date, sex chromosomes have not been reported in amphipods, and the knowledge on sex determination mechanism in amphipod is very limited.

The androgenic gland, which was firstly reported in *Orchestia gammarella* (Charniaux-Cotton, 1954), has been known as an essential organ maintaining male characters in amphipods as well as other crustaceans (Charniaux-Cotton, 1962). In amphipods, undifferentiated juveniles have both oviduct and vas deferens anlagen (Lockwood, 1968). When a male is determined, it develops androgenic glands which trigger a cascade of male secondary sex characters by reproducing the androgenic gland hormone (AGH). The oviducts degenerate in two or three mouls after a male is determined, whilst in intersex males the degeneration of oviducts is slower than normal males (Ginsburger-Vogel 1972a; 1972b). Females do not develop androgenic glands, and the gonads turn into ovaries. Two types of hormones have been reported
in female ovaries, permanent ovary hormone (POH) and temporary ovary hormone (TOH), which controls permanent (oostegites/brood plates) and temporary female characters (setae on brood plates), respectively (Charniaux-Cotton, 1952; 1962). Development of brood plates was induced in androgenic gland ablated male \textit{O. gammarella} after implanting ovaries from females (Charniaux-Cotton, 1962). Setae on brood plates of females were found to degenerate at the next moult after removing ovaries, and it grows back when the ovaries were re-implanted (Charniaux-Cotton, 1952).

Environmental factors, such as photoperiod and temperature have been revealed to influence the sex ratio of amphipod (Bulnheim, 1978; Dunn et al, 2001). Low temperature and long day length were reported to bias the sex ratio to male, whilst their impacts were found to be inconsistent in different populations (Naylor et al, 1988). Microspora is a group of intracellular parasites, which have been found to harbour in various amphipod species (Terry et al, 2004). Some microsporidian species have been reported to have feminising effects, and lead to the population sex ratio biased to females (Terry et al, 1998, 2004). The activity of microsporidian parasites could be impaired when exposed to low temperature and high salinity, therefore both the salinity and temperature are able to affect the sex determination via the effect of microsporidians (Bulnheim, 1969; Dunn and Hatcher, 1997).
1.2.2 Endocrinology of crustaceans

Endocrine system, which can be found in majority of animals, is a chemical communication network coordinating the function of different systems. There are three main components in endocrine system, namely glands, hormones and receptor cells (Lockwood, 1968). Glands are a clump of specialized cells which are able to produce and release hormones or their precursors. Unlike the secretion process of digestive system, in which enzymes are excreted though a tube or duct, the hormones in endocrine system are directly released into the inter-cell fluid, and transported though the body by circulatory system. The effect of endocrine system is very slow compared to the rapid response of nerve system (Lockwood, 1968).

The study on endocrine system in crustaceans dates back to the beginning of the 20th century, with the discovery of regulation functions of eyestalks (Rodríguez et al, 2007). The crustacean eyestalk harbours the X-organ-sinus complex which is comprised of two glands, X-organ and sinus gland (Carlisle, 1954). They are directly connected and controlled by the neuron system therefore they are termed as neuroendocrine organs, or neuro-excretory glands. In some crustaceans like amphipods, which lack the eyestalk structure, the X-organ-sinus complex locates in the central neuron system in the head (Rodríguez et al, 2007). Sinus gland does not actually synthesize hormones. Instead, it stores and releases the hormones produced by the X-organ. The hormones produced and secreted by the X-organ-sinus complex can be classified into three categories: hyperglycemic regulation hormones (CHH),...
negative regulators (molting inhibiting hormone -- MIH, mandibular organ inhibiting hormone -- MOIH, gonadal inhibiting hormone -- GIH) and stimulation-respons hormones responsive for colour change and neurodepressing (NDH).

Y-organs, which are paired organs in the head of higher Crustacea such as Malacostraca, are responsive for initiating the occurrence of moulting through the effect of ecdysteroid it excretes (Lachaise et al, 1993). Apart from the effect on molt, ecdysteroids are also involved in embryonic development. The embryo needs the ecdysteroids of maternal origin in the egg for normal development before it is able to synthesize ecdysteroids by itself (Mu and LeBlanc 2004). The levels of ecdysteroids and vitellogenin are found to increase simultaneously during the stage of ovarian maturation. But it is not clear whether increased ecdysteroids are responsible for stimulating the synthesis of vitellogenin (Tseng et al, 2002).

Methyl farnesoate (MF), a terpenoid hormone secreted by mandibular organ, involve in many bodily functions in crustaceans, such as molting and reproduction. MF is the unexpoxidated form of juvenile hormone of insects. Laufer et al (1993) reported high level of MF in the blood of reproductively active male and female crustaceans, and low level of MF in immature females and males in diapause. MF has been found to have the ability to stimulate the synthesis of ecdysteroids in Y-organ, as well as the growth of both male and female gonad (Laufer et al, 1993; Kalavathy et al, 1999). Olmstead and LeBlanc (2002) reported exposing eggs to 400 nM methyl farnesoate
resulted in all-male broods of offspring in *Daphnia magna*, which indicate that MF is able to act as a sex determination factor in some crustacean species.

The androgenic gland, lying at the terminal region of vas deferens, is a cluster of cells producing androgenic gland hormone (AGH) which stimulates and maintains the male sex characteristics (Charniaux-Cotton, 1954). Implantation of androgenic gland into juvenile female will cause masculinisation including preventing the development of female secondary characteristics and negatively regulating vitellogenin synthesis. (Taketomi and Nishikawa, 1996; Suzuki and Yamasaki, 1997). Female is the default sex in crustaceans, and the ablation of androgenic glands will lead to reduction in spermatogenesis and feminisation (Nagamine et al, 1980; Sagi et al, 1990; Khalaila et al, 1999). AGH has been identified only in malacostracan crustaceans, such as isopods, decapods and amphipods (LeBlanc, 2007). The AGH has been investigated in three isopods species – *Armadillidium vulgare, Porcellio scaber, Porcellio dilatatus*, and three chains are found to form the structure (Ohira et al, 2003). A chain (29 amino acid residues) and B chain (44 amino acid residues) are linked by two disulfide bonds and an additional peptide C chain. The A and B chains are found to be highly conserved among the three isopods. AGH influences the sex differentiation and drives the individual to develop as a male, for both primary and secondary sex characteristics. AGH is presumed to exert its effects during or after the sex differentiation stage, since AGH has been reported to be able to stimulate a female crustacean to develop male sex characteristics, but has no effect on
undifferentiated females (Suzuki 1999). Therefore AGH acts as a sex differentiation factor in crustaceans instead of a sex determination factor. Charniaux-Cotton (1952) reported the loss of ovigerous setae on brood plates after removing ovaries in female *Orchestia gammarella*, and the ovigerous setae reappeared when the ovaries were transferred back. Charniaux-Cotton and Payen (1985) suggested the ovary could produce and release two kinds of hormones, permanent ovarian hormone (POH) and temporary ovarian hormone (TOH). But none of these hormones have been identified yet (Zou, 2003).

As an important component in the endocrine system, hormones serve as message chemicals to control and coordinate many bodily functions. Like the activities of other organs, the excretion behavior of one gland could be under the control of other hormones. For example, the excretory activity of Y-organ is negatively controlled by the X-organ-sinus complex (LeBlanc, 2007). MIH maintains ecdysteroids in a low level during the intermolt stage. When the level of MIH declines, the level of ecdysteroids increases and then triggers the progress of moult (LeBlanc, 2007). Figure 2 depicts the functions and inter-reactions of the glands of endocrine system in crustaceans.
The crustacean hyperglycemic hormone (CHH), mandibular organ inhibiting hormone (MOIH), molting-inhibiting hormone (MIH) and gonadal inhibiting hormone (GIH) are produced in X-organ, and stored and released by the sinus. Ecdysteroids, methyl farnosoate (MF) and androgenic gland hormone (AGH) are produced and released by Y-organ, mandibular organ and androgenic gland, respectively. MOIH and MIH are known to have the ability to negatively control the effect of MF and ecdysteroids, respectively. GIH suppresses the growth of gonadal tissue, either the ovary in females or the testis in males. MF positively controls the process of molt, by stimulates the synthesis of ecdysteroids in Y-organ (Laufer et al, 1993). Evidence showed that MF is able to stimulate the growth of gonadal tissue, whilst it is not known whether MF exerts the effect through increasing the synthesis of the hormones controlling the growth of the gonad, for example AGH.
1.2.3 Male reproductive system in amphipods

This study mainly focused on the impacts of endocrine disruption and intersexuality onto the male amphipods. The reproductive organs in Crustacean are paired slender tubular structures, namely testes in males and ovaries in females (Schmitz, 1992). In amphipods, the testes extend from the second segment of pereiomere to the seventh segment of pereiomere, or the second segment pleomere (Schmitz 1967). The testes expand to form a lobed structure called seminal vesicle, which is used to store the matured spermatozoa. The seminal vesicle constrains to form vas deferens, which terminates and connects to genital papillae. The anterior of testis contains numerous primary germs cells, and the vacuolated testicular epithelium is presumed to secret mucus which facilitates the smooth ejaculation of spermatozoa (Charniaux-Cotton, 1957; Schmitz, 1967). Okumura and Hara (2004) reported the spermatogenesis in the freshwater prawn *Macrobrachium rosenbergii* showed close correlation to the moult cycle, with increased number of spermatogonia, spermatocytes in the late premolt stage, postmolt stage, respectively, and spermatozoa were formed in intermolt and early premolt stages. Androgenic gland locates in the distal end of vas deferens, and it produces androgenic gland hormone which is essential for maintaining the male secondary sex characteristics in crustaceans (Charniaux-Cotton, 1954).
1.2.4 Biology of *Echinogammarus marinus*

*Echinogammarus marinus*, previously named as *Chaetogammarus marinus* or *Marinogammarus marinus*, is a ubiquitous intertidal amphipod, which is widely distributed in northwest Europe (Sexton and Spooner, 1940; Lincoln, 1979; Leineweber, 1985). *E. marinus* are relatively large amphipods, with the maximum life span of approximately 12-14 months (Skadsheim, 1982). Individuals feed on seaweed, whilst cannibalistic behaviour has been observed (Maranhão and Marques, 2003). By analysing the gut content, *E. marinus* were found to consume both algae and some macro-invertebrates, and active predatory of *E. marinus* have been observed in the laboratory (Dick et al, 2005).

As a gonochoristic species, the male and female genders are separate in *E. marinus*, however, intersexes specimens have been reported in a variety of *E. marinus* populations (Ford et al, 2003a; Yang et al, 2008). In general, matured males (up to 25mm) are bigger than females (up to 20 mm), and males have larger gnathopods and slimmer body shape compared to females (Leineweber, 1985; Ford et al, 2003a; 2004a). Precopulation, the behaviour that a male monopolises and guards a female before mating, is very common in aquatic isopods and amphipods (Ridley, 1983). The behaviour is presumed to evolve because the female receptivity window during moult cycle is limited to a very short period (Parker, 1974). During precopulation, male amphipods hold the females using their gnathopods, and the pair closely live together from several hours to weeks, depending on the species (Jormalainen, 1998).
Like other amphipods, male *E. marinus* guard females to form pre-copula pairs until the females moult. The fertilisation takes place in the brood chamber of females, when the male amphipod releases spermatozoa into the water between the moulting and ovulation of the female (Ford et al, 2003b; Maranhão and Marques, 2003). Female *E. marinus* reproduce approximately 20 eggs each time (Cheng 1942), and the embryonic development has been found to be affected by temperature, with approximately 33, 32 and 17 days at 10 °C, 15 °C and 20 °C, respectively (Maranhão and Marques, 2003).

![Fig. 1.2. A picture of a male *E. marinus*.](image)
1.3 Endocrine disruption

Over the past two decades, research into endocrine disrupting chemicals has drawn a lot of attention amongst the scientific community, with the concerns that developmental and reproductive abnormalities caused by endocrine disruption could lead to population and ecosystem level effects. Endocrine disruption occurs when the function of hormone system is altered. Since the hormone system coordinates and controls many essential bodily functions, such as development and reproduction, dys-functioned endocrine system could lead to many serious problems, for instance impaired fecundity and fertility, which are able to impact the whole population rather than only the individual itself.

The first evidence of endocrine disruption dates back to 1940s (Ankley and Giesy, 1998). One of the most famous cases is the dramatically declined population of bald eagle in United States through 1940s to 1960s, due to the exposure of their preys to the pesticide DDT. The bald eagle population recovered since the ban of DDT in 1970s (Giesy et al, 1994; Ankley and Giesy, 1998). To date, endocrine disruption has been reported in many species, such as alligators, amphibiaons, fish and molluscs (Blaber, 1970; Smith, 1971; Guillitte et al, 1994; Lye et al, 1998; Hayes et al, 2002).

The specific mechanisms of how EDCs affect wildlife are not clear yet. Several pathways are able to explain the potential effects of EDCs onto the hormone system. Some EDCs are known to have the ability to influence the hormone level, either by
stimulating/inhibiting the synthesis of certain hormones in the gland, or accelerate/decrease the rate of hormone release. Some EDCs have endocrine activities themselves and they interfere with the hormone system by simulating the function of certain hormone, and finally enhance/reduce its effects. Competing with hormone on the binding sites of the receptors is another way how EDCs impact the function of the endocrine system (Rodríguez et al, 2007).

1.3.1 Endocrine disruption in crustaceans

To date, most studies on endocrine disruption have focused on vertebrate groups, and relative sparse evidence has been found in invertebrates. The lack of knowledge on invertebrate hormone system limits the investigations on invertebrates, and the high variety of invertebrate endocrine system makes it even more difficult. So far, endocrine disruption studies on invertebrates have mainly centred onto three groups, insects, molluscs and crustaceans, due to their ecological and economic importance (Matthiessen and Gibbs, 1998; Tyler et al, 1998; Soin and Smagghe, 2004; LeBlanc, 2007; Rodríguez et al, 2007).

Crustaceans are a very diverse group and occupy a variety of ecological niches, either in freshwater, seawater or even terrestrial environment. Concern has been raised recently, regarding to the huge impacts on the environment by human activities. Due to their wide distribution, crustaceans have the potential to suffer from a wide range of pollutants in the environment, and present as a great model to
study the impact of environmental contamination on invertebrates. Meanwhile, their position in the food chain makes them an important group of animals from the ecological point of view. Many eco-toxics have cumulative effect. They will be passed through the food chain, and concentrate in the tissue of high-level predators. It is critical to monitor the eco-toxicological response on those organisms which form the base of the ecosystem.

A variety of endocrine disrupting chemicals (EDCs) have been reported to have the ability to interfere with the hormone system in crustaceans. In recent decades, various endocrine disruption cases have been reported in crustaceans. Some EDCs, such as cadmium and DDT, are able to induce hyperglycemia in crustaceans by altering the level of CHH in eyestalk (Fingerman et al, 1981; Reddy et al, 1994). Fenarimol, a fungicide, and testosterone have anti-ecdysteroidal activity (LeBlanc, 2007). They are responsive of embryo development abnormality in *Daphnia magna* by decreasing the ecdysteroid level (LeBlanc et al, 2000; Mu and LeBlanc 2002). Pyriproxyfen, a methyl farnosoate agonist, has been reported to induce all-male offspring in *D. magna*, at very low concentration (Rider et al, 2005). Wang et al (2005) reported that dieldrin is able to suppress the effect of endogenous methyl farnesoate, while it also has positive actions when the methyl farnesoate is absent in the body.
1.3.2 **Intersexuality in crustaceans**

Intersexuality has been widely reported in various Crustacea species, such as crayfish, lobsters, crabs and amphipods (Sagi et al, 1996; Sangalang et al, 1997; Zou and Fingerman, 2000; Ford et al, 2003). Intersexes are more often found in male crustaceans (LeBlanc, 2007). To date, the mechanism of intersexuality is not fully understood. Pollution and feminising parasites have been suggested as potential causes of intersexuality in crustaceans (Kelly et al, 2004; Rodgers-Gray et al, 2004; Ford et al, 2004a).

Moore and Stevenson (1991) reported elevated levels of intersexuality in copepods nearby a sewage discharge off Edinburgh, which suggested the impacts of industrials pollution on the occurrence of intersexes in crustaceans. However, no relationship has been established between sewage exposure and intersexes incidence (Moore and Stevenson, 1991). The occurrence of intersexes in crustaceans is not always found to be consistent with the incidence of environmental pollutions. Ladewig et al (2006) examined the population structure and the population dynamics of two *Gammarus fossarum* populations in Germany, but failed to find the relationship between intersex levels and sewage effluent exposure.

Infection with feminising parasites has been revealed to be associated with increased level of intersexes in several Crustacea groups, for example microsporidians in amphipods and *Wolbachia* bacteria in isopods (Ginsburger 1991; Rigaud 1997).
Microsporidian parasites have been reported in a variety of amphipod hosts, for example *Gammarus duebeni*, *Gammarus roeseli*, *Gammarus pulex* and *Orchestia gammarallus* (Bulnheim, 1965; Bulnheim and Vavra 1968; Haine et al, 2004; Terry et al, 2004). The relationship between microsporidian infection and incidence of intersexuality in amphipods has been revealed by a number of studies (Terry et al, 1998; 2004; Kelly et al, 2004). The fact that females infected by feminising microsporidian produced female biased broods indicated the feminisation effects by the parasites (Ironside et al 2003; Mautner et al, 2007). Kelly et al (2004) reported that intersex *G. duebeni* were only found in offspring produced by microsporidian infected females, which suggests the incomplete feminising effect of microsporidian is responsible for the occurrence of intersexuality in amphipods. Similar associations between feminising parasite and the incidence of intersexes have been reported in isopods (Rigaud, 1997). Various degrees of intersexuality were observed in the isopod *Armadillidium vulgare*, and high temperature was presumed to impair the feminisation activity of *Wolbachia* bacteria therefore led to the sterile intersex phenotypes in their hosts (Rigaud and Juchault, 1998).

Although the mechanism of intersexuality in crustaceans is not clear yet, several studies have successfully induced intersexes in some Crustacea species. Ginsburger-Vogel and Carre-Lecuyer (1976) reported the intersexuality could be induced by transplant organs, including androgenic glands, muscle and testes or ovaries, from intersex males or thelygenous female *Orchestia gammarallus* to
normal male *Orchestia cavimana* and *Talitrus saltator*. Depending on the different transplanted organs, 33% to 48% normal males were turned into intersexes. The same experiment was carried out on *Orchestia aestuarensis* and *Orchestia mediterranea*. Only the testes from intersex were able to induce intersex in *Orchestia gammarallus*, and the seminal vesicle or androgenic gland failed to do so (Ginsburger-Vogel, 1991). Jungmann et al (2004) reported various levels of intersexes in *Gammarus fossarum* populations from five streams in Germany. Significantly increased incidence of intersex was observed in gammarids from a location with a low level of intersexuality after being exposed to the water from a stream with a high level of intersexes. Neither chemicals from municipal sewage treatment plants nor microsporidian parasite was found to associate with the occurrence of intersexuality, and some environmental factors might be responsible for inducing intersexes in *Gammarus fossarum* (Jungmann et al, 2004).

Matured intersexes are usually found to be larger than normal specimens, which indicated developmental abnormalities, for example delayed maturation (Plaistow et al, 2003; Ford et al, 2004b; Pastorinho et al, 2009). Male and female individuals cannot be differentiated until the sixth moult in *Gammarus chevreuxi*, whilst intersex individuals can be distinguished as early stage as the third moult, by the relatively small body size (Sexton, 1924). The fact that intersex individuals are significantly smaller than their normal counterparts suggests the occurrence of delayed development. Androgenic gland has been found to govern and maintain male sexual
characteristics in crustaceans (Charniaux-Cotton, 1954). Reduced level of androgenic gland hormone (AGH) has been revealed in intersex amphipods, suggesting intersexuality in crustaceans might be caused by the impaired activity of androgenic glands (Rodgers-Gray et al 2004; Ford et al 2005).

Intersexuality is believed to be the result of endocrine dysfunction. Therefore it provides a great model to study endocrine disruption in crustaceans (Ford et al, 2004). Ford et al (2003a) reported two types of intersex in *Echinogammarus marinus*, intersex males and intersex females. Subsequent studies revealed both intersex males and females suffer costs on development and reproduction, such as delayed maturation and impaired fertility (Ford et al, 2003a; Ford et al, 2004; Yang et al, 2008). This study aimed to develop biomarkers for endocrine disruption in crustaceans, by investigating the intersex *E. marinus* on physiological and genetic level.
1.4 Microsporidian parasite

Microspora is a phylum of primitive single-cell organisms, which are known as one of the smallest and simplest eukaryotic microorganisms (Biderre et al, 1995). They have some unique features which are different from other eukaryotes, for example they are amitochondrial and contain 70s ribosomal RNA instead of the typical 80s rRNA in eukaryotes (Curgy et al, 1980). Because of these prokaryotic-like characteristics, Microsporidians are often referred as “ancient” or “primitive” eukaryotes (Curgy et al, 1980; Vossbrinck et al, 1987; Kamaishi et al, 2005). Recent studies found that a selection of genes in microsporidian have close orthologues in fungi, such as the genes of α- and β- tubulins, mitochondrial 70HSP, TATA Box Binding Protein (TBP) and the largest subunit of RNA polymerase II (RPB1), which suggested that microsporidians and fungi are closely related (Keeling P.J. and Doolittle, 1996; Edlind et al, 1996; Hirt et al, 1997 and 1999; Fast et al, 1999). Keeling and Mcfadden (1998) proposed that microsporidians were neither ancient nor primitive eukaryotes. The lack of Golgi membranes and mitochondria could be the result of secondary organelle degenerating during the evolution of their highly specialised parasitism lifestyle (Germot et al, 1997; Van de Peer et al, 2000).

In 19th century, the first microsporidium was recognized, *Nosema bombycis*, which infects and leads to destructive pébrine disease in the silkworm (Pasteur, 1870). Since this initial identification, a variety of animal groups ranging from protists to
humans, have been reported to be infected by microsporidians (Sprague and Vavra, 1977; Canning, 1990; Goodgame, 1996). Microsporidians are obligate intracellular parasites, and they consume their host’s organelles to obtain energy (Canning, 1998). There are mainly two stages during their life cycle, proliferative stage (meronts) and transmission stage (spores) (Canning and Lom, 1986). The spore structure has been employed as a taxonomic indicator since it varies dramatically across the phylum (Vavra and Larsson, 1999). However, this phylogenetic approach is not always promising since several microsporidian species have been found to be able to produce multiple types of spores (Iwano and Ishihara, 1991). This has lead to the phylogenetic relationship determined by molecular techniques, based on the sequence of small sub-unite ribosomal RNA (SSU rRNA) (Dunn and Smith, 2001). Vossbrinck et al (1987) reported the first microsporidian SSU rRNA sequence, from the species Vairimorpha necatrix. The SSU rRNA has been used as an effective molecular biological tool to study the phylogenetic relationship amongst microsporidian species, because the sequence is conservative between species (Van de Peer et al, 2000). It is worth noting that some microsporidian species have been reported to exhibit polymorphism of SSU sequence, for example O’Mahony et al (2007) reported that there were two or three rRNA sequence variants in a single Nosema bombi spore. Therefore the polymorphism of SSU sequence should be taken into account when the SSU sequence is used to investigate the relationship of microsporidian species.
In obligate intracellular parasites, the transmission success amongst hosts is critical to the survival of microsporidians. Cross species host compatibility has been reported in a selection of microsporidians, and hosts harbouring multiple species of microsporidian parasites have also been revealed (Dunn et al, 2000; Haine et al, 2004). There are mainly two transmission strategies in microsporidians, vertical and horizontal transmission. Vertical transmission happens while microsporidians are passed from the parental generation of hosts (mainly transovarial) to their offspring. Horizontal transmission could occur between hosts of the same or different species and generations (Dunn and Smith, 2001). Horizontal transmission was reported as a common mode in microsporidians, and can be achieved by several pathways -- oral intake, epidermal invasion and venereal (Dunn and Smith, 2001). Terry et al (2004) reported vertically transmitted microsporidian parasites widely spread in amphipods, and the fact that closely related microsporidian parasites infect similar host species indicated the co-evolution between microsporidians and their amphipod hosts. Some microsporidium species, for example *Nosema bombycis*, is known to apply both strategies, however some microsporidians are exclusively either horizontal or vertical transmitted (Dunn et al, 2000; Dunn and Smith, 2001).

In most occasions, parasites employing horizontal transmission have stronger virulence and lead to higher burden in the host than the vertically transmitted ones (Kellen et al, 1965; Ebert and Herre, 1996). Unlike the strategy of high copy and strong virulence for horizontally transmitted parasites, vertically transmitted
parasites increase their survival chance by distorting the host’s sex ratio. Most of the vertical transmissions follow maternal line, with some very rare exceptions, for example the sigma virus which could be vertically transmitted by both male and female *Drosophila melanogaster* hosts (Bangham et al, 2007). In order to increase the chance of being past to the next generation by female hosts, some vertically transmitted microsporidians are capable of distorting the host’s sex ratio in favour of female, through the strategies of male killing or feminisation (Kellen et al, 1965; Dunn et al, 1996). Some microsporidian species have been reported to distort their host’s sex ratio by male-killing in mosquito (Andreadis and Hall, 1979; Bandi et al, 2001), and some exert feminising effects in amphipods (Dunn et al, 1993; Terry et al, 1998; Rodgers-Gray et al, 2004).
1.5 Molecular techniques exploring transcriptome profiling

Due to the lack of genetics knowledge, the molecular biology studies in crustaceans are relatively limited. Complete mitochondrial genome sequence has been reported in a variety of crustacean groups, such as daphnia, copepod, prawn, shrimp, crayfish and crab (Valverde et al, 1994; Crease, 1999; Wilson et al, 2000; Machida et al, 2002; Yamauchi et al, 2003; Miller et al, 2004). However, whole-genome sequencing has been limited in some crustaceans with relatively small genome size, such as *Daphnia pulex* and *Daphnia Magana*. Colbourne et al (2011) reported the draft genome of *D. pulex*, which became the first Crustacea species whose whole genome sequence is available, and the genome sequencing project for *D. Magana* is currently ongoing (Shaw et al, 2008). This study applied microarray and next generation sequencing technology to investigate the genetic difference between normal and intersex *E. marinus*, with the aim to identify the mechanism of intersexuality in crustaceans on the genetic level. Brief introductions on the two techniques employed by this study were reviewed in the following sections.
1.5.1 DNA Microarray

1.5.1.1 General introduction of microarray

DNA microarray is a powerful molecular technique of characterising the expression pattern of thousands of genes simultaneously (Fodor et al, 1991; 1993; Schena et al, 1995; Taniguchi et al, 2001; Lettieri, 2006). The mechanism for microarray is based on the exclusive specificity and affinity of complementary base-pairing between targets and probes (Brown and Botstein, 1999). The nucleotide sequences immobilised on the arrays are usually called “targets”, each individual spot as an “element” and DNA or RNA strands hybridised to targets on the microarrays as “probes” (Watson et al, 1998). However, some studies referred the immobilised sequence on the array as “probe” and the specimen DNA or RNA sequences as “targets” (Levicky and Horgan, 2005), and this study adopted definitions of “target”, “probe” and “element” defined by Watson et al (1998). Microarray technique evolved from Southern’s insight that due to the complementary double-stranded DNA structure, the fluorescence-tagged oligo nucleotide could be used to interrogate nucleic acid molecules attached to a solid support (Southern, 1975).

There have been two breakthrough innovations stimulating the wide application of microarrays -- non-porous solid supports and high-density spatial synthesis of oligonucleotides on solid supports (Haslett and Kunkel, 2002; Lettieri, 2006). The use of non-porous solid support, facilitated the miniaturisation of arrays and the detection of microscopic fluorescence in the target-probe hybridisation (Khrapko et
al, 1989; Lockhart et al, 1996; Schena et al, 1995; 1996). Glass is widely adopted as the suitable material to make microarrays, because of its non-porous solid surface, durable to high temperature and washes of high ionic buffers, as well as low auto-fluorescence (Cheung et al, 1999). The development of methods for high-density spatial synthesis of oligonucleotides on solid supports allowed the analysis of hundreds of thousands of genes at the same time (Fodor et al, 1991; 1993; Lettieri, 2006).

There are primarily three types of nucleic acid sequence utilised for printing microarrays, namely short oligonucleotides (15-25 nt), long oligonucleotides (50-120 nt) and cDNA sequences from PCR amplification (100-3000 bp) (Lettieri, 2006). Microarrays using short oligonucleotide are primarily employed for detecting single-nucleotide polymorphisms (SNPs), because the high sensitivity is required to discriminate one base pair miss-match for SNP arrays, and short sequence maximizes the destabilisation induced by single base pair miss-match (Lockhart et al, 1996). The cDNA microarrays are commonly used to analyse the gene expression patterns in organisms whose genomic information is limited, and the targets for printing the cDNA arrays can be readily obtained from the cDNA libraries by PCR amplification (Lettieri, 2006). Arrays using long oligonucleotides are able to discriminate genes sharing high sequence similarity, and are considered as a sensitive alternative to the cDNA microarray (Hughes et al, 2001; Bozdech et al, 2003).
Most of early studies applying microarrays to compare the gene expression patterns amongst organisms of different conditions were carried out on bacterial, yeast or clonal mammalian cells, which could be manipulated in vitro (Haslett and Kunkel, 2002). DNA microarray was used to explore the genes involving metabolic shift from fermentation to respiration in *Saccharomyces cerevisiae*, and the same arrays were also utilised to identify genes affected by the transcriptional co-repressor *TUP1* or the transcriptional activator *YAP1* (DeRisi et al, 1997). This work demonstrated the utility of the DNA microarray as a feasible approach to analyse gene expression patterns in a genome-wide scale (DeRisi et al, 1997). Nowadays, microarrays have been employed to investigate the transcriptome profiling in countless living organisms, including microorganisms, plants, animals (Wilson et al, 2001; Haslett and Kunkel, 2002; Maguire et al, 2002; Seki et al, 2002; Zhou, 2003; Wu et al, 2004; van der Vena et al, 2005).

For a typical workflow for applying cDNA microarrays, RNA is purified from tissues of examined organisms, and converted into fluorescence labeled cDNA, which is then hybridised with targets on the microarray (Lettieri, 2006). Cyanine (Fig. 1.3, Cy3 and Cy5) is a common fluorescent dye staining cDNA targets. After the “probe” and “target” hybridisation, the expression level of a gene could be visualised by detecting and quantifying the fluorescence intensity of the corresponding element by a fluorometer (Schena et al 1995; DeRisi et al 1996; Lashkari et al, 1997). The principle procedures of using cDNA microarray is
interpreted in Fig. 1.4.

Fig. 1.3. The molecular structure of Cy3 and Cy5.
Fig. 1.4. The workflow of cDNA microarray.

The RNA samples are purified from the interested tissues, and are then converted into fluorescent dye incorporated cDNA probes. The labeled cDNA is hybridised with immobilised targets on arrays, and the gene expression level of a gene could be visualised by scanning and quantifying the fluorescence intensity of the corresponding element on the microarray.
1.5.1.2 DNA microarray studies in crustaceans

Due to the lack of genetic information, the deployment of transcriptomic studies on crustaceans has been relatively limited. Microarray has been used as a putative molecular biology technique to explore changes of gene expression under different conditions, and provided valuable information to understand the mechanism on the genetic level. Several microarray platforms were constructed for *D. magna* and *D. pulex* prior to the complete of *D. pulex* genome in 2011 (Colbourne et al, 2011), making Daphnia the most frequently used model crustaceans utilised for ecotoxicogenomics (Soetaert et al, 2006; Soetaert et al, 2007a; 2007b; Poynton et al, 2007). The following table elucidates some ecotoxicogenomics studies on Daphnia by employing the microarray technique (Table 1.1).
<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Species</th>
<th>Main Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>Soetaert et al.</td>
<td><em>D. magna</em></td>
<td>Constructed a first version of cDNA microarray on <em>D. magna</em>. Vitellogenin gene was found to be down-regulated, which indicated the impaired reproduction by propiconazole.</td>
</tr>
<tr>
<td>2007</td>
<td>Watanabe et al.</td>
<td><em>D. magna</em></td>
<td>Developed a microarray on <em>D. magna</em>. Chemical-specific gene expression patterns were revealed by microarray analysis after characterising the transcriptomic profiles for daphnids exposed to copper sulfate, hydrogen peroxide, pentachlorophenol, or β-naphthoflavone.</td>
</tr>
<tr>
<td>2007</td>
<td>Soetaert et al.</td>
<td><em>D. magna</em></td>
<td>Cadmium affected molecular pathways were revealed by microarray analysis.</td>
</tr>
<tr>
<td>2007</td>
<td>Shaw et al.</td>
<td><em>D. pulex</em></td>
<td>Revealed up-regulated metallothionein in the experiment group by comparing control and cadmium exposed <em>D. pulex</em> groups.</td>
</tr>
<tr>
<td>2008</td>
<td>Watanabe et al.</td>
<td><em>D. magna</em></td>
<td>Developed a new microarray for <em>D. magna</em> by increasing the EST numbers, and characterised the gene expression pattern for beta-naphthoflavone (bNF) exposed daphnids.</td>
</tr>
<tr>
<td>2009</td>
<td>Vandenbrouck et al.</td>
<td><em>D. magna</em></td>
<td>A large proportion of affected genes were found to involve in metabolic processes when the organisms were exposed to nicke, and the results from mixed metals exposure indicated that additional pathways were affected by binary exposure rather than the simple sum of the influenced genes when the two heavy metals were separately administrated.</td>
</tr>
<tr>
<td>2010</td>
<td>Vandegehuchte et al.</td>
<td><em>D. magna</em></td>
<td>Transcriptomic profiles of three generations of <em>D. magna</em> exposed to Zn were studied by DNA microarray. The majority of genes differentially transcribed in F1 Zn exposed daphnids were not differentially transcribed in the F0 Zn exposed organisms. Acclimation was indicated by a significantly lower number of differentially transcribed genes in Zn exposed daphnids in F2, compared to the F0 and F1 generations</td>
</tr>
</tbody>
</table>

Table 1.1. A list of recent microarray studies on ecotoxicogenomics in *Daphnia*.  

43
Apart from their exploitation in ecotoxicology, global transcriptomic approaches have also been applied to characterising the gene profiling of crustaceans in different physiological stages, for example molt. Kuballa and Elizur (2007) compared gene expression pattern among five moult stages in a fresh water crab, *Portunus pelagicus*. Up-regulated metallothionein was observed in the pre-moult and post-moult stages, and FaMeT, the enzyme involves in the process of converting farnesoic acid into methyl farnesoate (MF) was found be up-regulated in the inter-moult stage. Shechter et al (2007) applied the microarray to explore ecdysteroid-responsive genes in the red-claw crayfish *Cherax quadricarinatus*. The moult cycle was induced in the crayfish by either hydroxyecdysone treatment or the removal of X-organ–sinus gland, and a total of 102 genes showed similar response in the both manipulations and considered as ecdysteroid-responsive genes. Of these 102 genes, the majority of them were unknown genes, and the fact that 92 out of the 102 genes were down-regulated suggested the energetic shift before ecdysis. There have been very limited studies comparing normal and intersex crustaceans on the whole transcriptome scale. Ford et al (2008) employed cross-species microarrays to compare the gene expression profiling amongst four sexual phenotypes in *E. marinus*, and revealed 582 sex specific dimorphic genes in this amphipod species.
1.5.2 DNA sequencing

1.5.2.1 A brief review of DNA sequencing techniques

DNA sequencing technologies provide information on nucleotide alignment of nucleic acid sequences, such as genomic or complimentary DNA, and largely facilitate the biological studies by allowing researchers to decode the genome of living organisms (Pop and Salzberg, 2006). In 1970s, several sequencing techniques were reported, for example the specific chemical degradation based DNA sequencing approach described by Maxam and Gilbert (1977). Since the chain-termination sequencing method was firstly introduced by Sanger and his colleagues in 1977, this sequencing technique has been widely applied (Sanger et al, 1977). The chain-termination sequencing method, also named as the Sanger sequencing method, or the capillary sequencing when the fluorescence dye and capillary electrophoresis were introduced, employs 2',3'-dideoxynucleotides (ddNTPs) to terminate the extension of DNA synthesis therefore results in a selection of sequences with various lengths, which are then used to determine the ordering of nucleotide on the DNA sequence by electrophoresis (Sanger et al, 1977; Smith et al, 1986; Metzker, 2005).

In recent years, there emerged a variety of high throughput sequencing technologies, so called next generation sequencing, such as GS FLX (454) sequencing (Roche), Solexa (Illumina) and SOLiD technology (Applied Biosystems) (Margulies et al, 2005; Pop and Salzberg, 2006; Johnson et al, 2007; Mardis, 2008; Ondov et al, 2008). These revolutionary sequencing technologies are able to process millions of
sequencing reactions in parallel therefore dramatically increase the speed of DNA sequencing (Mardis, 2008). The Roche GS FLX Titanium sequences 100 mega bases per run, with the average 400 bp length of each read, and the three next generation sequencing technologies were compared in the Table 1.2. Although there have been skepticisms on these high throughput sequencing technologies, such as short sequence read and low fidelity compared to conventional Sanger method sequencing, with the unprecedented speed and relatively low cost in large scale sequencing, these next-generation sequencing technologies are leading the biological researches into the “post-genomics” era (Schuster, 2008).

<table>
<thead>
<tr>
<th>Sequencing Platform</th>
<th>Mechanism</th>
<th>Amplification Method</th>
<th>Read Length</th>
<th>Throughput per Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS FLX Titanium (Roche)</td>
<td>SBS</td>
<td>Emulsion PCR</td>
<td>~400 bp</td>
<td>Up to 400-600 Mp</td>
</tr>
<tr>
<td>Solexa (Illumina)</td>
<td>SBS</td>
<td>Bridge Amplification</td>
<td>30-150 bp</td>
<td>Up to 600 Gb</td>
</tr>
<tr>
<td>SOLiD (Applied Biosystems)</td>
<td>SBL</td>
<td>Emulsion PCR</td>
<td>35 bp</td>
<td>Up to 1-3 Gb</td>
</tr>
</tbody>
</table>

Table 1.2. Features of the three next-generation DNA sequencing technologies. SBS: Sequencing by Synthesis; SBL: Sequencing by Ligation.
1.5.2.2 Roche GS FLX (454) sequencing

The Roche GS FLX (454) sequencing employing the high-throughput DNA pyrosequencing principle is firstly introduced into commercial usage in 2004 (Margulies et al, 2005; Mardis, 2008). Unlike the Sanger sequencing methods which applies the concept of chain termination using ddNTPs, pyrosequencing is based on the “sequencing by synthesis” principle which relies on detecting the light signals emitted by the pyrophosphate (PPi) released from nucleotide incorporation (Ronaghi et al, 1998; Ronaghi, 2001; Ahmadian et al, 2000). During the process of pyrosequencing, in addition to the substrates for an ordinary PCR reaction, other reagents such as ATP sulfurylase, luciferase and apyrase, adenosine 5’ phosphosulfate (APS) and luciferin are added into the sequencing matrix in order to indicate the occurrence of chain elongation by detecting the emission of light signals (Ronaghi et al, 1998; Wang and Marincola, 2003). The solution flow with one of the four nucleotides A, T, C or G is added and removed to the synthesis system sequentially. When the added nucleotide is complementary to the template, the nucleotide will be incorporated and the pyrophosphate will be released. ATP sulfurylase then converts APS into ATP, using the released pyrophosphate. The generated ATP is utilised as the energy by luciferase to oxidise luciferin, through which visible light will be emitted and subsequently detected by the CCD camera (Ronaghi, 2001). In liquid-phase pyrosequencing, apyrase is utilised to degrade the non-used nucleotide and ATP, and the reaction will be triggered again by adding a new nucleotide flow (Ronaghi et al, 1998; Ronaghi, 2001).
To carry out the GS FLX sequencing, DNA sequences are fragmented and ligated with adaptors, and subsequently immobilized on onto DNA-capture beads. DNA strands attached to the bead are amplified by emulsion PCR, which simultaneously amplifies each individual sequence in an isolated bead in oil phase (Dressman et al, 2003; Williams et al, 2006). The hundreds of thousands of beads, each with millions of identical copies amplified from one single DNA fragment, are subsequently transferred a PicoTiterPlate device for pyrosequencing. The GS FLX sequencer is able to generate 100 Mb sequences per run, with the average read length at 250 bp, whilst the Genome Sequencer FLX Instrument (Fig. 1.5) using GS FLX Titanium series reagents can provide 1 million high quality 400 bp length reads per run. The workflow of GS FLX sequencing is interpreted in Figure 1.6.

Fig. 1.5. The Genome Sequencer FLX Instrument.

Image adapted from the manufacturer’s website.
Fig. 1.6. Flowchart of 454 sequencing technique.
1. Prepare DNA templates with the length of ~200 bp. 2. Ligate adaptors. 3. Immobilize the template on the DNA-capture bead. 4. Amplification in water-in-oil mixture. 5. Pyrosequencing. 6. Data analysis. Images adapted from the manufacturer’s website.
Due to its fast speed and relatively low cost, the 454 sequencing has dramatically facilitated the studies used to be limited by lack of DNA alignment information. The 454 sequencing has been utilised for a wide range of purposes, such as whole genome and transcripts sequencing (Cheung et al, 2006; Wicker et al, 2006; Ellegren, 2008). Margulies et al (2005) sequenced the whole genome of *Mycoplasma genitalium*, a bacterium with reference genome sequence available, with 96% coverage and at the accuracy of 99%. It was the first bacterial genome sequenced by the method other than Sanger sequencing. In 2008, Wheeler et al (2008) reported the complete genome of a single individual, James D. Watson, in two month by 454 sequencing technology. Vera et al (2008) sequenced the transcripts from larvae, pupae and adult of the non-model species, Glanville fritillary butterfly (*Melitaea cinxia*), using 454 sequencing technology. A total of 608 053 expressed sequence tags were yielded and assembled into 48 354 contigs.
1.6 Project objectives:

The aim of this study was to investigate in detail the physiology and transcriptomes of normal and intersex *Echinogammarus marinus* with the view to generating biomarkers of endocrine disruption and reproductive dysfunction in Crustacea. Due to the body of work already conducted on the female intersex phenotype this study primarily focuses on the normal and intersex male phenotypes. This was achieved through the following objectives:

1. Assess the prevalence of intersexes in various *E. marinus* populations
2. Compare the male fertility between amphipods from polluted and reference sites as well as between normal and intersex amphipods
3. Identify the distribution of microsporidian parasites in different sexual phenotypes of *E. marinus*, and determine the parasite species
4. Compare the transcriptome profilings amongst three male sexual phenotypes of *E. marinus* using a cross-species DNA microarray
5. Employ the second generation sequencing technique (454 GS FLX platform) to sequence the cDNA libraries generated from the gonads of different sexual phenotypes of *E. marinus*
6. Develop physiological and molecular biomarkers for de-masculinisation in crustaceans
Chapter 2

Intersexuality and Male Fertility in
_Echinogammarus marinus_
2.1 Introduction

In recent years, an increasing number of studies focused on the endocrine disruption in wildlife, with the concerns that a good number of man-made chemicals are able to affect the reproductive functions of wild animals and therefore lead to serious ecological issues (Tyler et al, 1998). Endocrine disruption has been reported to take place through several pathways: A) the synthesis of a certain hormone is stimulated or inhibited; B) an endocrine disrupting chemical interfere with hormone-receptor interaction by acting as an agonist or antagonist; C) the metabolism of circulating hormones is altered (Rodríguez et al, 2007). A lot of the researches centred on the oestrogen or anti-oestrogen effects of EDCs, and relatively few studies have investigated the cost of endocrine disruption in males (Depledge and Billinghurst, 1999).

Feminised males, with abnormal testicular development and reduced fecundity have been reported in a wide range of vertebrate groups, such as alligators (Guillette et al, 1994), fish (Lye et al, 1998), felines (Facemire et al, 1995), amphibians (Hayes et al, 2002). Guillette et al (1994) reported that the alligator population at Lake Apopka was dramatically declined, due to the contamination from nearby pesticide manufacturing. Further histological studies discovered that some male individuals had testicular abnormalities including poorly organised seminiferous tubules
Similar cases of intersexuality have been revealed in fish. For example, intersex Roach *Rutilus rutilus* found in contaminated rivers in the UK had milt reduced by more than 60%, suggesting that intersex fish could suffer reduced fertility (Jobling et al., 2002). Inhibited spermatogenesis and abnormal testes in wild flounder *Platichthys flesus* influenced by sewage effluent have also been reported (Lye et al., 1998).

As an essential index to assess male fecundity, sperm counts have been investigated by many studies, on both animals and human beings (Sharpe and Skakkebaek, 1993; Vom Saal et al., 1998; Haubruge et al., 2000). Carlsen et al. (1992) reported a population-wide sperm count decline in humans. After investigating 61 studies on human sperm counts, the sperm counts found in male human’s semen have been found to have significantly decreased by nearly 50% in 50 years, from $113 \times 10^6$ ml$^{-1}$ in 1940 to $66 \times 10^6$ ml$^{-1}$ in 1990 (Carlsen et al., 1992). However, the conclusion that the sperm counts of human have been reduced in recent decades remains controversial (Bujan et al., 1992; Jouannet et al., 2001). It is possible that the observed decreased sperm counts might barely be caused by different measurement techniques (Handelsman, 1997) or region variation (Fisch et al., 1996).

Compared to studies of male infertility in vertebrate groups, a very limited number of investigations have been carried out in invertebrates. Those which have been completed are restricted to certain groups, such as insects, crustaceans, molluscs and
echinoderms (Byrne et al, 1998; Au et al, 2000; Dunn et al, 2006a; Price et al, 2008).

There has been evidence showing that dietary restriction constrains sperm production in Indian meal moths *Plodia interpunctella* (Gage and Cook, 1994). Nice (2005) reported that nonylphenol (NP) treated oysters produced decreased numbers of motile spermatozoa. This study demonstrated that the ability to produce motile sperm had been reduced by 70% in oysters exposed to 1g/µl NP whilst 100% of control specimens were able to produce motile sperm. Au et al (2000) reported that cadmium (Cd) and phenol had the ability to reduce the motility of sea urchin and mussel spermatozoa. In the both sea urchin and mussel, ultrastructure data showed thickened midpiece in spermatozoa exposed to Cd, which could explain the impaired motility.

To date, studies on sperm counts in crustaceans are very limited. Previous studies focusing on crustacean spermatozoa either centered upon their use in taxonomic classification (Jamieson, 1991) or for cryogenic studies within the growing aquaculture industry (Bray and Lawrence, 1998; Bart et al, 2006; Gwo, 2006). Dunn et al (2006a) recorded sperm counts in an amphipod demonstrating a link between lowered sperm counts and reduced fertilisation of eggs in the amphipod, *Gammarus duebeni*. In addition, Dunn et al (2006a) suggested that *G. duebeni* were able to alter their allocation of sperm with uninfected females receiving more sperm as opposed to females infected by microsporidian parasites. In lobster, the amount of spermatozoa transferred from males to females has been found to be dependent on
several factors, such as male size, female size and previous copulations (MacDiarmid and Butler, 1999). Although substantial sperm allocation in each reproductive event was observed in *Gammarus pulex*, sperm depletion was reported to have no impact on male amphipods’ reproductive availability, since the replenishment is fast and the sperm deposit in males is able to recover during the precopulation period (Lemaître et al, 2008).

Intersexuality has been reported in a variety of Cushtacea species (Bulnheim, 1965; Ginsburger-Vogel, 1991; Sangalang et al, 1997; Zou and Fingerman, 2000), and provides a great model to study endocrine disruption as well as male infertility in crustaceans (Ford et al, 2004a). In the model species used by this study, an intertidal amphipod *Echinogammarus marinus*, intersex specimens could be identified by the presence of internal or external secondary sex characteristics. Male *E. marinus* were identified by male’s secondary sex characteristics – two genital papillae – between their last pair of pereopods (Fig. 2.1 A). Females were recognized by the presence of brood plates developed besides the gills (Fig 2.1 B). Intersex females were identified by the presence of both brood plates and with one or two genital papillae formed between the seventh pair of pereopods. Some male specimens were found to have the female secondary sex characteristics – brood plates externally and oviducts internally (Ford et al, 2003; Yang et al, 2008). According to whether the appearance of female characteristics could be spotted externally or only internally, male intersexes could be sub-divided into two phenotypes, namely external or internal intersex male (Fig
Fig. 2.1. The secondary sex characters of five sexual phenotypes in *E. marinus*. A: normal male with genital papillae; B: normal female with brood plates; C: external intersex male with genital papillae and rudimental brood plates; D: internal intersex male with oviduct developed on testes.
To date, the cause of intersexes in crustaceans is not clear yet, whilst it has been considered as the consequence of dys-functioned endocrine system (Ford et al, 2004a). Although the mechanism of intersexuality in crustaceans is not fully understood, both industrial pollution and feminising parasites have been reported to be consistent with the incidence of intersexuality in crustaceans (Moore and Stevenson, 1991, 1994; Kelly et al, 2004; Rodgers-Gray et al, 2004; Ford et al, 2004a). Microsporidian parasites have been reported in a variety of amphipods and intersexuality has been presumed as a result of incomplete feminisation by microsporidians (Kelly et al, 2004; Terry et al, 2004). Despite the strong evidence on the association between microsporidian infection and the occurrence of intersexuality in amphipods, some studies reported the relationship between these two was not always promising (Buikema et al, 1980; Ladewig et al, 2002). Buikema et al (1980) found no microsporidian parasite in intersex Gammarus minus. Similarly, Ladewig et al (2002) collected intersex Gammarus fossarum from two streams in Germany, but no microsporidian infection was observed. The intersex individuals without microsporidian infection indicated that there were other factors inducing the occurrence of intersexuality in amphipods. The association of industrial pollution and high incidence of intersexuality in crustaceans has been reported by several studies (Moore and Stevenson, 1991; Takahashi et al, 2000; Ford et al, 2004a). Moore and Stevenson (1991) reported high frequencies of intersexuality in copepods from various sites impacted by either chemical pollution or sewage discharge. Significantly higher proportion of intersex Echinogammarus marinus was also
reported in industrial polluted sites comparing to reference sites (Ford et al, 2004a). Despite the increased intersexuality rate found in polluted sites, the mechanisms of industrial pollutants inducing intersexuality in crustaceans have not been established, and the indirect influence via parasites could not be ruled out (Ford et al, 2006).

This study aimed to investigate the reproductive abnormalities including testicular development and intersexuality in several *E. marinus* populations in Scotland and southern England. The prevalence of intersexuality was investigated in an industrially polluted site, Inverkeithing, and two reference sites, Loch Fleet and Thurso. The *E. marinus* population in Inverkeithing has been monitored on annual basis since 2002 (Ford et al, 2006), and the population structure has been well studied in order to assess the impact of intersexuality onto the population dynamics (Ford et al, 2007). Loch Fleet and Thurso were selected as reference sites according to the SEPA estuary classification scheme. Testicular abnormalities were quantified and the sperm counts were compared between normal and intersex males, as well as between the specimens from clean and polluted sites. A parallel study comparing sperm counts in males from clean and industrially impacted site was also carried out in two *E. marinus* populations in Portsmouth, southern England. To further elucidate whether physiological differences exist between normal and intersex *E. marinus*, the weight and moult frequency were compared between normal and intersex male phenotypes.
2.2 Materials and Methods

2.2.1 Field collection of amphipod specimens

*Echinogammarus marinus* were collected underneath the stones and seaweeds on the seashore of five sites in the UK, Thurso, Loch Fleet and Inverkeithing in Scotland, as well as Portsmouth Harbour and Langstone Harbour in England. Animals were brought back to the laboratory, and were anaesthetised in carbonated seawater and examined under stereo microscope for the sexual phenotype (Chapter 2, Section 2.1).

2.2.1.1 Sampling of *E. marinus* in three Scottish sites

In April and May, 2007, *E. marinus* were collected from three sites in Scotland – one polluted site (Inverkeithing, 56.025637, -3.385377) and two reference sites (Loch Fleet, 57.933809, -4.010696 and Thurso, 58.597759, -3.512685). Inverkeithing is located on the north of the Firth of Forth. It was once categorised by SEPA (Scottish Environment Protection Agency) as Class D (Seriously Polluted), and was upgraded to Class C (Unsatisfactory) in 2002. A shipbreaker’s yard and a decommissioned paper mill contributed to the main sources of pollution, leading to high levels of PCBs and heavy metals (SEPA, 2000; Table 2.1). Loch Fleet and Thurso are classified as Class A (excellent) under the coastal classification scheme. Thurso River, from the mouth of which amphipods were collected, was classified as Class B (good) under the SEPA river classification scheme.
Table 2.1. Contaminant levels in the common mussel, *Mytilus edulis* from Inverkeithing Bay between 1988 and 2005 (SEPA, unpublished data).

<table>
<thead>
<tr>
<th>Contaminant</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>National background</th>
<th>Substantially elevated</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cadmium</td>
<td>1.20</td>
<td>0.55</td>
<td>3.24</td>
<td>1</td>
<td>10</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Chromium</td>
<td>3.30</td>
<td>0.80</td>
<td>15.40</td>
<td>2</td>
<td>40</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Copper</td>
<td>9.10</td>
<td>5.91</td>
<td>14.40</td>
<td>6</td>
<td>20</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Lead</td>
<td>6.60</td>
<td>1.30</td>
<td>11.50</td>
<td>4</td>
<td>25</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mercury</td>
<td>0.30</td>
<td>0.19</td>
<td>0.75</td>
<td>0.15</td>
<td>1.5</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Nickle</td>
<td>3.30</td>
<td>0.70</td>
<td>15.80</td>
<td>1.5</td>
<td>15</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Zinc</td>
<td>161.30</td>
<td>107.00</td>
<td>204.00</td>
<td>90</td>
<td>400</td>
<td>mg/kg dry&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>HCB</td>
<td>0.20</td>
<td>0.15</td>
<td>0.19</td>
<td>1</td>
<td>10</td>
<td>μg/kg wet&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Dieldrin</td>
<td>0.50</td>
<td>0.19</td>
<td>0.76</td>
<td>2</td>
<td>20</td>
<td>μg/kg wet&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sum DDTs</td>
<td>1.80</td>
<td>0.61</td>
<td>2.82</td>
<td>20</td>
<td>100</td>
<td>μg/kg wet&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sum ICES 7</td>
<td>16.90</td>
<td>0.99</td>
<td>48.48</td>
<td>10</td>
<td>50</td>
<td>μg/kg wet&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> SEPA (unpublished data, 1988–2005).

<sup>b</sup> SEPA (unpublished data, 1994–2005).

<sup>c</sup> ICES guideline levels for contaminants.

2.2.1.2 Sampling of *E. marinus* from two Portsmouth sites

Amphipods were collected underneath seaweeds from two sites (Langstone Harbour, 50.789624, -1.042419, and Portsmouth Harbour, 50.827035, -1.095151) in Portsmouth, southern England. In order to examine the prevalence of intersex males in these two sites, *E. marinus* were collected from both sites in the summer of 2009, and the proportion of each sexual phenotype was recorded. For the sperm count comparison, two samplings were carried out in Langstone Harbour, in October and December 2009, respectively, and one sampling for Portsmouth Harbour in December, 2009. Elevated levels of TBT and the antifouling biocide, Irgarol 1051, have been reported in Portsmouth Harbour, due to its historical activities of
anti-fouling painting on boats (Gough et al, 1994; Zhou, 2008). The water quality of Langstone Harbour was categorised as “Excellent” by Havant Borough Council, for the period from 01/04/2008 to 30/09/2008. Animals were sexed in the lab, according to their secondary sex characteristics (Chapter 2, Section 2.1). From each site, approximately 30 normal males were examined for the number of spermatozoa in their seminal vesicle.

2.2.1.3 Sampling of E. marinus in Inverkeithing for moult cycle monitoring

*E. marinus* were collected underneath seaweeds from Inverkeithing, Scotland, in June 2008. Amphipods were sexed in the laboratory after being anaesthetised in the carbonated seawater, and the subsample of 120 animals were categorised into four sexual phenotypes, namely normal male, normal female, intersex male and intersex female, with 30 specimens for each of the four sexual phenotypes (Chapter 2, Section 2.1).

2.2.2 Moult cycle monitoring in normal and intersex E. marinus

Four sexual phenotypes of *E. marinus*, namely normal males, normal females, intersex males and intersex females (from the sampling in 2.2.1.3), were compared for their moult frequencies. For each sexual phenotype, 30 amphipod specimens were individually kept in labeled plastic pots with approximately 100 ml filtered seawater. All the 120 animals were kept in an incubator at 15 °C, with the photoperiod of 12:12 (L:D). Amphipods were fed by seaweeds, and the seawater was
changed every two days. The number of newly moulted amphipods as well as mortalities were counted and recorded every day, and exuviae were removed from the pot after examining. The moult frequency was monitored for a period of 33 days, and the moult frequency of each group was calculated and statistically compared by using Fisher’s exact test.

2.2.3 Sperm count

Subsamples of 262 normal male *E. marinus* were collected from the three Scottish sites (Thurso, Loch Fleet and Inverkeithing), and 58 male amphipods from two Portsmouth sites (Portsmouth Harbour and Langstone Harbour), and had their sperm counts compared. The method of quantifying spermatozoa in male *E. marinus* was adapted from the work of Dunn et al (2006a). The amphipod specimen was dapped by using a clean tissue to remove excessive water on the animal, and each male sample was then weighed before being sacrificed. To conduct sperm count, the testes were dissected out from males onto a cavity slide with 30 µl distilled water. Spermatozoa were squeezed out from the seminal vesicle, and subsequently transferred into a pre-weighted eppendorf tube by pipetting. Spermatozoa were diluted with 200 µl distilled water and pipetted several times with care to achieve homogeneity. The tube was weighted again to calculate the final volume of the sperm solution. Assumption was made that the density of the sperm solution was equal to the density of water (1.0 g/ml). A total of 30 µl sperm solution was applied on a haemocytometer. By counting the number of spermatozoa in the medium square of
the haemocytometer underneath microscope, the number of spermatozoa in every one micro-litre of sperm solution could be worked out. The total sperm count then could be estimated based on the density of spermatozoa and the total volume of the sperm solution. Sperm counts between normal males from clean and polluted sites, as well as between normal and intersex male specimens from the same site, were statistically compared by applying the ANOVA analysis (SPSS 16.0)
2.3 Results

2.3.1 Sex ratio and male intersexuality in five *E. marinus* populations in the UK

2.3.1.1 Sex ratio and male intersexuality in three Scottish sites

Approximately 400 *E. marinus* were sampled from each of the three Scottish sites, Thurso, Loch Fleet and Inverkeithing, in April and May 2007, and the prevalence of male intersexuality in each site was assessed. According to their internal and external sex characteristics, amphipod specimens were categorised into five sexual phenotypes, namely normal females, intersex female, normal males, internal intersex males and external intersex males. The male intersexuality ratio of three *E. marinus* populations was compared (Table 2.2).

<table>
<thead>
<tr>
<th>Site</th>
<th>Normal male</th>
<th>Intersex male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IIM</td>
</tr>
<tr>
<td>Thurso</td>
<td>94 (82.4%)</td>
<td>15 (13.2%)</td>
</tr>
<tr>
<td>Loch Fleet</td>
<td>76 (93.8%)</td>
<td>3 (3.7%)</td>
</tr>
<tr>
<td>Inverkeithing</td>
<td>47 (70.1%)</td>
<td>9 (13.4%)</td>
</tr>
</tbody>
</table>

Table 2.2. Sex ratio of sampled *E. marinus* from three sites.

Proportions of normal, internal and external intersex males in three *E. marinus* populations: Thurso, Loch Fleet and Inverkeithing. IIM: internal intersex male; EIM: external intersex male.
Significant differences were observed in the proportions of normal and intersex males amongst all three sites (G=19.1, d.f.=4, P<0.001). Inverkeithing *E. marinus* population was found to have the highest proportion (29.9%) of intersex males amongst the three sites, and Loch Fleet had the lowest (6.2%). In Inverkeithing and Loch Fleet, approximately equal numbers of internal and external intersex males were observed. Whilst in Thurso, a greater proportion of internal intersex males were revealed compared to external intersex males. However, the difference in the proportion of the two male intersex phenotypes between Thurso and Inverkeithing *E. marinus* population was not significant (Fisher’s exact test, p>0.05).

### 2.3.1.2 Prevalence of male intersexuality in Langstone Harbour and Portsmouth Harbour *E. marinus* populations

*E. marinus* specimens were collected from Langstone Harbour and Portsmouth Harbour respectively, and the number of each male phenotype was recorded in the Table 2.3. The proportion of male intersexes was found to be higher in Portsmouth Harbour (20.7%) than in Langstone Harbour (10.1%), whilst no significant difference was observed on the proportions of the three male phenotypes between the two *E. marinus* populations (Fisher’s exact test, p>0.05).
Table 2.3. Sex ratio of sampled *E. marinus* in two Portsmouth sites.
The number and proportion of five sex-phenotypes of *E. marinus* specimens collected from Portsmouth Harbour (England) in summer 2009. EIM: external intersex male; IIM: internal intersex males.

<table>
<thead>
<tr>
<th></th>
<th>Normal Males</th>
<th>Intersex males</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EIM</td>
<td>IIM</td>
</tr>
<tr>
<td>Langstone Harbour</td>
<td>89 (89.9%)</td>
<td>3 (3.0%)</td>
</tr>
<tr>
<td>Portsmouth Harbour</td>
<td>42 (79.2%)</td>
<td>5 (9.4%)</td>
</tr>
</tbody>
</table>

2.3.2 Body weight comparison between normal and intersex males

The body weight and the length of specimens were measured before dissection. In general, intersex males (both internal and external intersex males) were found to be heavier than normal specimens in all the three sites (Figure 2.2).
Significant difference has been observed in the body weight between normal and intersex males in both Thurso (t test, t= -2.646, p<0.05) and Inverkeithing (t test, t= -2.806, p<0.01). In Loch Fleet, the average body weight of intersex males was larger than normal males, but the difference was not significant (t test, t= -0.987, p>0.05). Normal males from Loch Fleet were found to have larger body weight than ones from Inverkeithing and Thurso. Significant difference was observed on the body weight amongst normal males from the three sites (ANOVA, F_{2, 149}=16.734,
p<0.001), and multiple comparison tests (Turkey) revealed that there were significant differences between any of the three normal male groups (p<0.05).

Further comparisons were made between weights of normal males, internal and external intersex males in Thurso and Inverkeithing (Figure 2.3). Since there were not enough intersex specimens collected from Loch Fleet, the body weight of normal males and two male intersex phenotypes was not statistically compared in this site. In both Thurso and Inverkeithing, external intersex males were found to be the heaviest, normal males were the lightest and internal intersex males formed an intermediate group. Significant difference of weight was observed between the three male phenotypes in Thurso (ANOVA, F\(_{2,75}\) = 3.544, P < 0.05; Fig. 2.3a). Multiple comparison tests (Turkey) revealed no significant difference between individual groups (P>0.05). In Inverkeithing, significant differences were observed between the weights of the three male phenotypes (ANOVA, F\(_{2,64}\) = 6.386, P < 0.01; Fig. 2.3b). Multiple comparison tests (Turkey) revealed significant differences between normal male and external intersexes (P < 0.05), whilst no significant difference was found between other groups (P > 0.05).
Fig. 2.3. **Body weight comparison amongst three male phenotypes in Thurso (A) and Inverkeithing (B).**

Mean body weight (±2S.E.) of normal, internal and external intersex male *E. marinus* were compared in (A) Thurso and (B) Inverkeithing (Scotland, UK), respectively. NM: normal males; IIM: internal intersex males; EIM: external intersex males.
2.3.3 Moult frequency in normal and intersex *E. marinus*

A total of 120 amphipods, including 30 specimens for each sexual phenotype (normal male, normal female, intersex male and intersex female), were individually cultured and daily examined for moult. During the 33-day observation period, a total of 43 (35.8%) samples moulted, and no amphipod was found to moult twice or more. Amongst the four groups, intersex males had the highest moult frequency (66.7%). Only two normal females moulted during the monitoring period, and the moult frequency for normal males and intersex females was 26.7% and 43.3%, respectively (Fig 2.4; Table 2.4).

**Fig. 2.4.** The number of moulted *E. marinus* in each sexual phenotype during the monitoring period of 33 days.

NM: normal male; NF: normal female; IM: intersex male; IF: intersex female.
Significant difference was observed in the proportion of moulted individuals amongst the four sexual phenotypes (Fisher’s exact test, p<0.001). Subsequently pairwise comparisons were made between each two of the four groups by using a Bonferroni corrected p value (p = 0.0083). The proportion of moulted individuals was found be significantly different between normal females and intersex females, normal females and intersex males, as well as normal males and intersex males (Fisher’s exact test, p<0.0083).

Three mortalities, one intersex male and two intersex females, were found out of the 120 specimens in the first 25 days (Fig 2.5; Table 2.5). Mortalities increased dramatically after 25 days in three groups (normal female, intersex male and intersex female), whilst the only mortality in the normal male group appeared on the last day of the monitoring period.

<table>
<thead>
<tr>
<th>Group</th>
<th>Moulted</th>
<th>Not moulted</th>
<th>Percentage of moulted samples of moulted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal male</td>
<td>8</td>
<td>22</td>
<td>26.7%</td>
</tr>
<tr>
<td>Normal female</td>
<td>2</td>
<td>28</td>
<td>6.7%</td>
</tr>
<tr>
<td>Intersex male</td>
<td>20</td>
<td>10</td>
<td>66.7%</td>
</tr>
<tr>
<td>Intersex female</td>
<td>13</td>
<td>17</td>
<td>43.3%</td>
</tr>
</tbody>
</table>

Table 2.4. The number of moulted *E. marinus* during 33 days culture in the lab.
Fig 2.5. Mortalities of four sexual phenotypes of *E. marinus*.
Accumulative mortality out of 30 *E. marinus* specimens in each sexual phenotype during the moult behaviour period of 33 days. NM: normal male; NF: normal female; IM: intersex male; IF: intersex female.

Normal males had the lowest mortality rate (3.3%) and intersex females had the highest (23.3%) amongst the four groups (Fig. 2.5). Two male groups were found to have less mortalities than two female groups, and the two normal groups had less mortalities than their intersex counterparts. However, there was no significant difference either amongst the four sexual phenotypes or between any two of them (Fisher’s exact test, p>0.05).
2.3.4 Sperm count comparison

2.3.4.1 Sperm counts in normal and intersex males from three Scottish *E. marinus* populations

Approximately 50 normal male *E. marinus* from each site (n = 58, 47, and 47 for Thurso, Loch Fleet and Inverkeithing Bay, respectively) and all collected intersex specimens were dissected for subsequent sperm count analyses. Figure 2.6 shows a typical spermatozoon from *E. marinus*. The spermatozoon consists of a main body and a straight non-motile pseudoflagellum.

![Fig. 2.6. A spermatozoon of *E. marinus*.](image-url)
Analysis of covariance (ANCOVA) was conducted to compare spermatozoa number in normal males from each site by using the body weight as the covariate (Figure 2.7). The sperm count data were normalised by $4^{th}$ root transformation. Significant differences were observed between sites and the covariate, weight (ANCOVA, $F_{2,151} = 15.419, P < 0.001; F_{1,151} = 77.931, P < 0.001$; respectively). The two reference sites had the highest mean number of spermatozoa (mean±S.E.M.: Loch Fleet 16.063±0.443; Thurso 15.550±0.492) whilst Inverkeithing (12.697±0.467) had the lowest (mean values taken at weight = 76.25 mg).

Fig. 2.7. Sperm counts comparison amongst normal males from three sites.
Numbers of spermatozoa plotted against weight in normal male *Echinogammarus marinus* from Inverkeithing (industrially impacted), Thurso (reference) and Loch Fleet (reference). $4^{th}$ root transformed.
Subsequent comparisons were made between the number of sperm produced in normal and all intersex (internal and external) males from Thurso and Inverkeithing (Fig. 2.8). Loch Fleet was excluded due to low samples sizes of intersex specimens. In Thurso, normal males produced on average (normalised to weight) approximately 15% more spermatozoa than intersex males. However, this result failed to reach 95% confidence limits (ANCOVA, $F_{1,77} = 2.97$, $P = 0.089$; $F_{1,77} = 21.52$, $P < 0.001$). No significant differences in sperm counts were observed between normal and intersex males in Inverkeithing (ANCOVA, $F_{1,66} = 2.135$, $P > 0.1$; $F_{1,66} = 30.896$, $P < 0.001$). In both Thurso and Inverkeithing, significant differences were observed in the co-variable weight (mg).
Fig. 2.8. Sperm counts comparison between normal and intersex males in Thurso (A) and Inverkeithing (B).

Numbers of spermatozoa plotted against weight in normal and intersex male *Echinogammarus marinus*, from (A) Thurso and (B) Inverkeithing (Scotland, UK). *4\textsuperscript{th} root transformed.*
2.3.4.2 Sperm counts of two *E. marinus* populations in Portsmouth

In order to assess the seasonal variation of spermatozoa in male *E. marinus*, sperm count was conducted on 30 and 28 male amphipods collected from Langstone Harbour in October and December 2009, respectively. The sperm counts of males sampled in October and December were compared by ANCOVA (SPSS 16) using the body weight as the covariate. The number of spermatozoa was normalised by quad-root transforming, and was plotted against its body weight (Figure 2.9). No significant difference was observed between the sperm counts of males collected in October and December, and the number of spermatozoa was found to be unrelated to the males’ body weight (ANCOVA, F_{2,57}=0.02, p>0.05; F_{1,57}=0.022, p>0.05).
Fig. 2.9. Sperm counts comparison for the seasonal variation.

Number of spermatozoa plotted against body weight in male *E. marinus* collected from Langstone Harbour in October and December, respectively. *4*\(^{th}\) root transformed.
The normalised data of sperm count was subsequently compared between male *E. marinus* from Portsmouth Harbour and Langstone Harbour, by ANCOVA with body weight as the covariate (Fig. 2.10). Significantly higher number of spermatozoa were found in males from Portsmouth Harbour than those from Langstone Harbour (ANCOVA, $F_{1,57} = 4.347, P < 0.05$; $F_{1,57} = 5.606, P < 0.05$). Sperm count in males from Langstone harbour was found not correlated to their body weight (Pearson correlation test, p>0.05).

![Fig. 2.10. Sperm counts comparison between males from two Portsmouth sites.](image)

Number of spermatozoa plotted against body weight in male *E. marinus* collected from Portsmouth Harbour and Langstone Harbour (England UK), respectively. * 4th root transformed.
2.4 Discussion

In this chapter, the prevalence of male intersexuality in five *E. marinus* populations in the UK was investigated, and the cost of intersex in male amphipods was assessed at the physiological level. Intersex specimens were found in all of the five sites examined by this study, with the ratio of male intersexuality ranging from 6.5% to 29.8%. Two phenotypes of male intersexes, namely internal intersex males and external intersex males were revealed in these sites. Although the internal character of intersexuality in male amphipod has already been reported by other studies (Ginsburger-Vogel 1972a; 1972b; Ford et al, 2006), this is the first study to quantify the prevalence of this intersex phenotype in amphipod populations. Internal intersex male provides an extra phenotype to study de-masculinisation in amphipods. However, it is not clear whether the internal intersex male is an intermediate phenotype between normal and external intersex male, or it is a completely different intersex phenotype.

Intersexes have been revealed in a variety of amphipod species, such as *Gammarus duebeni, Gammarus fossarum, Gammarus minus* and *Orchestia gammarella* (Bulnheim, 1965; Ginsburger-Vogel, 1972a; Buikema et al, 1980; Ladewig et al, 2002), and have been considered as a great model to study endocrine disruption in crustaceans. In this study, higher proportion of male intersexes was revealed in the industrially impacted site, Inverkeithing, compared to *E. marinus* populations from
the two reference sites, Loch Fleet and Thurso. Increased incidence of male intersexuality was also observed in a polluted site in England, Portsmouth Harbour (20.7%) compared to the reference site Langstone Harbour (10.1%). Increased levels of intersexuality in crustaceans have been reported in polluted sites by other studies (Moore and Stevenson, 1991, 1994; Ladewig et al, 2002; Ford et al, 2004a).

To date, the mechanism for intersexuality in crustaceans has not been fully understood, and it is not clear that whether industrial pollution exerts direct influence inducing intersexes in crustaceans, or the impact is indirect, in other words, via other factors such as feminising parasite. There has been strong evidence demonstrating the relationship between microsporidian parasite infection and the occurrence of intersexuality in amphipods (Rigaud and Juchault, 1998; Kelly et al, 2002, 2004; Terry et al, 2004). The correlation of microsporidian infection and intersexes in *E. marinus* is investigated in the Chapter 3 of this thesis.

Significant difference of the normal male body weight has been observed amongst three Scottish sites, with Loch Fleet population had the highest body weight and normal males from Thurso had the lowest. Intraspecific body weight difference amongst various populations has been revealed in other amphipods species (Strong, 1972). Environment factors have been reported to associate with the body weight of the population, for example populations from high latitude were observed to have higher body weight than the ones from low latitude (Poulin and Hamilton, 1995).
However, exceptional cases were also discovered (Sainte-Marie, 1991), and this study did not observe the association between population body weight and the latitude of the habitat either. A large number of amphipods from Loch Fleet were found to be infected by trematode parasite, which has been reported to induce gigatism in snails (Mouritsen, 1994; Gorbushin, 1997; Chapuis, 2009), and possibly contributed to the elevated body weight of amphipods from Loch Fleet.

The body weight of normal and intersex males amphipods were compared in this study. Intersex males were found to be heavier than their normal counterparts, and the weight decreased following the order below: external intersex males > internal intersex males > normal males. Similarly, Ford et al (2004b) reported that the weight of three female intersex phenotypes decreased in the order as: intersex females with two genital papillae > intersex females with one genital papilla > normal females. Ford et al (2004b) suggested that the increased weight observed in intersexes might be caused by endocrine dysfunction, resulting in a delay in the timing of maturation.

In this chapter, two intersex groups were found to have higher moult frequency than their normal counterparts. A variety of factors have been reported to be able to influence the moult frequency in crustaceans. High temperature is found to increase moult frequency therefore reduce the moult interval in amphipods (Morino, 1978). In some amphipod species, such as *Gammarus duebeni* and *Gammarus salinus*, the moult interval has been reported to be postponed in both males and females when the
other sex is missing (Skadsheim, 1990). The animals were individually kept when they were monitored for moult frequency. It is possible that the moult frequency observed by this study is not able to reflect the true moult frequency of this amphipod species in wild. Sexton reported that male and female individuals could not be differentiated until the sixth moult in *Gammarus chevreuxi*, whilst intersex individuals can be distinguished from their normal counterparts as early stage as the third moult, by the relatively small body size (Sexton, 1924). The facts that intersex individuals are significantly smaller than normal amphipods, and more mouls were needed to display sexual characters in intersexes than in normals (Sexton 1924), suggested the occurrence of delayed development in intersexes. Crustaceans have to shed their hard shell to increase their body size, and the inter-moult period in adults has been revealed to be gradually elongated along their age (Sexton, 1924; Chang, 1995; Caddy, 2003). The increased weight in intersexes might be related to the high moult frequency observed in intersexes.

Ford et al (2003a; 2004b) had previously observed costs associated with female intersexuality in *E. marinus*, such as reduced fecundity, decreased embryo survival rate and delayed maturation. In order to assess the cost of intersexuality, as well as the impact of environmental contamination, on male fecundity in amphipods, sperm counts between normal males and intersex males as well as between males from clean and polluted sites were compared in this chapter. The correlation between sperm counts and body weight was revealed in the three Scottish *E. marinus*
populations investigated by this study, which has also been reported in other Crustacea species (Lemaître et al., 2008; Sato et al., 2008).

A comparison of sperm counts from three Scottish sites has revealed lower sperm counts (~20%) in crustaceans inhabiting in an industrially impacted site compared to two reference sites. In Thurso, intersex male amphipods had reduced sperm numbers (~15%) compared to normal males, however this latter result was not significant. Previous studies across many Phyla have highlighted links between reduced sperm counts and fertility (Zenick and Clegg, 1989; Dunn et al., 2006a). Although some studies on rats have also pointed out that fertility is not affected by reduced sperm counts (Kirby et al., 1992; Kuriyama et al., 2005). Reduced sperm counts in intersex male *E. marinus* were observed at one of the reference sites in Scotland, Thurso, however this result was not replicated at the polluted site. Why this result was not replicated at Inverkeithing is unclear, however, greater variation observed in sperm count data may be a result of the contamination. McCurdy et al. (2004) found that when intersex male specimens of the amphipod *Corophium volutator* were paired with normal females the resultant broods were smaller than when compared to mating with normal males. Results from the current study suggest low sperm counts could well have resulted in fewer fertilised eggs therefore lead to the decline of the population.

A parallel study comparing sperm counts in male *E. marinus* from clean and polluted
sites was conducted on two *E. marinus* populations in Portsmouth, southern England. Due to its historical activity of anti-fouling painting on boat, Portsmouth Harbour has been reported to be impacted by the antifouling biocides, such as TBT and Irgarol 1051 (Gough et al, 1994; Zhou, 2008). Langstone Harbour was considered as the reference site in this study. No significant difference was observed in the sperm count between males from Langstone Harbour and Portsmouth Harbour. However, the co-relationship between sperm counts and the male’s body weight was observed to be different in the two sites. In the polluted site, Portsmouth Harbour, sperm counts were found to correlate to the male’s body weight, whilst the co-relationship was not revealed in the clean site, Langstone Harbour.

The independence of sperm deposit and body weight of males from Langstone Harbour is a potential sign of sperm depletion. Dunn et al (2006a) reported that in sperm limited male *Gammarus duebeni*, the number of spermatozoa were significantly affected by the times of mating, but not by the body weight. The *E. marinus* population in Langstone Harbour is found to be female biased (data not shown in this study), and the shortage of males will lead to sperm limitation. However, the Inverkeithing *E. marinus* population was also found to be female-biased, whilst significant co-relationship between sperm count and body weight was observed in this site. It is possible that apart from the sex ratio, there might be other factors affecting the mating frequency of the males, or the amount of spermatozoa released in each mating, and the specific mechanism needs to be further
In this chapter, three Scottish and two English *E. marinus* populations were investigated. In both areas, higher proportion of male intersexes was observed in polluted sites comparing to reference sites. However, more *E. marinus* populations need to be investigated in order to further investigate the relationship between industrial pollution and intersexuality in amphipods. Normal and intersex *E. marinus* were subsequently compared on several physiological factors, such as body weight, sperm counts and moult frequency. Reduced sperm counts in intersex males were observed in the reference site, Thurso, but not in the polluted site, Inverkeithing. Normal males from Inverkeithing were found to have significantly less spermatozoa than those from two reference sites, indicating the impaired male fertility caused by the direct or indirect impact from industrial pollution. Intersexes were observed to be heavier than their normal counterparts, and it was presumed to be caused by delayed maturation. Higher moult frequency was revealed in intersexes than in normal specimens, which suggested that the increased body weight observed in intersexes might be in the result of constant growth and moult, even in the adult stage. To date, the mechanisms of sex determination and sex differentiation in crustaceans are not clear, and cause of intersexuality also remains unknown. There has been evidence showing the strong relationship between feminising parasite infection and the occurrence of intersexes in crustaceans. Studies identifying the association between the two will be able to shed some light on the
mechanism of intersexuality in crustaceans.
Chapter 3

Microsporidian Infection in *E. marinus*
3.1 Introduction

Intersexuality has been reported in a variety of Crustacea groups, such as crabs, lobsters and amphipods (Bulnheim, 1965; Ginsburger-Vogel, 1991; Sangalang and Jones, 1997; Zou and Fingerman, 2000). In Chapter 2, the distributions of intersexes in various *E. marinus* populations were investigated, and normal and intersex individuals were compared on the physiological level, for example the sperm counts between normal and intersex males, and the moult frequency amongst different sexual phenotypes. To date, the mechanism of intersexuality is not clear yet, whilst genetic controls, industrial pollutions, environmental conditions and cytoplasmic parasites have been reported as the potential factors inducing intersexes in crustaceans (Bulnheim, 1978; Ginsburger-Vogel, 1991; Moore and Stevenson 1991; Dunn et al, 1993b; 1996; Rogers-Gray et al, 2004). Feminising parasites have been revealed in some crustacean groups, for example *Wolbachia* bacteria in isopods and microsporidians in amphipods (Bouchon et al, 1998; Rigaud and Juchault, 1998; Terry et al, 2004). Strong evidence on the relationship between the occurrence of intersexes and the incidence of feminising parasite infection indicates that intersexuality in crustaceans is a possible result of incomplete feminisation by cytoplasmic parasites (Rigaud and Juchault, 1998; Kelly et al, 2002, 2004). However, there have been some studies reporting no significant relationship between microsporidian infection and intersexes in amphipods (Buikema et al, 1980; Dunn et al, 1993b; Ladewig et al, 2002), which also suggested that other factors besides
microsporidian infection might contribute to the occurrence of intersexuality in amphipods.

Microsporidian infection has been revealed in a variety of amphipod species, such as *Gammarus duebeni*, *Corophium volutator* and *Orchestia mediterranea* (Ginsburger-Vogel, 1991; Dunn et al, 1993a; Mautner et al, 2007). Terry et al (2004) reported 11 species of microsporidians in 16 amphipod species (100% of investigated amphipod populations). Coexistence of different microsporidian species in the same amphipod population has been reported by several studies, but dual infection by two microsporidian species in a single amphipod host is reported to be very rare (Ironside et al, 2003a; Terry et al, 2004; Haine et al, 2004).

Some microsporidian species, such as *Nosema granulosis* and *Dictyocoela duebenum*, have been demonstrated to exhibit feminising effects in their amphipod hosts (Terry et al, 1999; Ironside et al, 2003a; Kelly et al, 2004; Dunn et al, 2006b). Kelly et al (2002) reported that female *Gammarus duebeni* infected by *Nosema granulosis* produced female biased broods, whilst uninfected females had male biased offspring. Female biased broods in microsporidian infected female amphipods were also reported by other studies (Ironside et al, 2003a; Mautner et al, 2007). The fact that intersex *G. duebeni* were only found in the offspring produced by microsporidian infected females suggested intersexuality in amphipods was possibly induced by incomplete feminising effects of the microsporidians (Kelly et al, 2004).
Ginsburger-Vogel (1991) reported that intersex individuals could be introduced by grafting tissues from intersex amphipods onto animals presenting normal phenotypes, indicating that the factor causing intersexuality can be readily transmitted between hosts. Various tissues, such as muscle and gonads of Orchestia gammarellus could be used to induce intersexuality, while only testes can be utilised to induce intersexuality in O. aestuarensis or O. mediterranea.

To date, various techniques have been employed to identify the occurrence of microsporidian infection in amphipods, such as histology, DAPI staining as well as PCR based screening method (Baker et al, 1995; Terry et al, 1998; Garcia 2002; MacNeil et al, 2003; Kelly et al, 2004). Histology is utilised to identify the microsporidian infection in situ. Bulnheim and Vavra (1968) revealed that the microsporidian infection was restricted to the oocytes, follicle cells and surrounding ovarian tissues in female Gammarus duebeni by employing the histology based parasite screening method. DAPI staining method is usually carried out on hosts’ embryos, and is widely adopted to determine the microsporidian infection in amphipod broods (Haine et al, 2004; Kelly et al, 2004; Terry et al, 2004). The intensity of the parasite infection can also be quantified by counting the number of microsporidian nuclears in the embryo of the host (Weidner et al, 1994; Terry et al, 1998).

Molecular biological techniques provide a powerful parasite screening method based
on the PCR (Vossbrinck et al, 1993; Hogg et al, 2002; Terry et al, 2003). A fragment of some microsporidian gene, for instance the Small Subunit of ribosomal RNA (SSU rRNA) gene, is amplified by the PCR which is usually conducted on the genomic DNA or cDNA generated from the host tissue. The PCR products are loaded on the agarose gel (1%-2%) for electrophoresis, and the infection status can be determined by the present of a corresponding band. By purifying and sequencing the PCR product, and the species of the parasite can be identified by aligning the sequence to the existing database. A combination of restriction enzymes are used to digest the PCR products (PCR-RFLP, restriction fragment length polymorphism), and different parasite species can be identified by the pattern of bands (Hogg et al, 2002; Ironside et al, 2003a; 2003b).

This study aimed to investigate the relationship between microsporidian infection and intersexuality in several E. marinus populations. Strong correlation between microsporidian infection and intersexes has been reported in E. marinus, however the parasite species has not been identified (Ford et al, 2006). In this study, histology investigation was applied to determine the infected tissue of the host, and PCR based parasite screening technique was carried out to investigate the prevalence of microsporidian infection in various E. marinus populations. Species of microsporidians were identified by amplifying and sequencing the parasites’ Small Subunit ribosomal RNA (SSU rRNA) gene, and species-specific primers were developed based on the sequencing data. By using the species-specific primers, not
only the status of microsporidian infection, but also the species of microsporidian parasite can be determined by employing one PCR reaction. In order to quantify the microsporidian infection intensity in amphipods, Real-Time quantitative PCR (RT-QPCR) was applied to quantify the microsporidian SSU rRNA by using the host house-keeping genes, such as tubulin and GAPDH, as the internal control. By comparing the frequency as well as the intensity of microsporidian infection amongst various host sexual phenotypes, the relationship between the incidence of intersexuality and the occurrence of microsporidian infection was subsequently analysed.
3.2 Materials and methods

3.2.1 Specimen collection

*E. marinus* specimens were collected from the field, and sexed in the laboratory by using the same method as of Section 2.2.1. The relevant information for sampling carried out for each study, such as site, time and sample size, is interpreted in the following sections.

3.2.1.1 *E. marinus* samples for microsporidian identification by histology

*E. marinus* were collected from three Scottish sites, Inverkeithing, Loch Fleet and Thurso, in April and May 2007, as the same samplings carried out for sperm count analysis (Chapter 2.3.1). Eleven intersex males from Inverkeithing and approximately 20 normal males from each of the three sites were fixed in 80% methanol, and sent off to the University of Stirling for histological sectioning and staining.

3.2.1.2 Sampling for parasite screening in Inverkeithing *E. marinus* population

A total of 206 *E. marinus* specimens were collected from an industrially impacted site -- the southern dock side of Inverkeithing in Scotland, in August 2008. From a sub-sample of 112 amphipods including four sexual phenotypes, namely normal males, normal females, intersex males and intersex females, RNA was individually purified from the gonadal tissue and subsequently reverse transcribed into single-stranded cDNA. In order to assess the proportion of microsporidian infected
individuals in each sexual phenotype, a PCR based microsporidian parasite screening procedure was conducted on cDNA samples of the 112 amphipods.

3.2.1.3 Sampling for parasite screening in Portsmouth E. marinus populations

A total of 476 E. marinus were sampled from the intertidal zone nearby Portsmouth Harbour (50.826944, -1.095278), in summer 2009. A sub-sample of amphipods (N=48) were used to assess the prevalence of microsporidian infection using a PCR based parasite screening technique. Microsporidian infection identification was also carried out on the embryos of several female specimens in order to address the question whether the parasite could be vertically transmitted. Due to the low number of intersex female specimens found in Portsmouth Harbour, only four sexual phenotypes (NM, NF, IIM and EIM) were statistically investigated for the distribution of microsporidian infection in this study.

3.2.1.4 E. marinus for microsporidian infection quantification by QPCR

E. marinus were collected underneath seaweeds and stones from Inverkeithing (Scotland), in October 2007. A total of 75 male amphipods including 25 normal males, 25 internal intersex males and 25 external intersex males were sacrificed, and the testes were dissected out for RNA extraction as well as subsequent parasite infection intensity quantification. Five pairs of testes from the male specimens of the same phenotype were pooled together as a single sample which was then preserved in RNAlater ICE for future usage. In this study, five replicates were employed for
each male sexual phenotype.

3.2.2 Histology

The amphipod specimens were sacrificed and preserved in 80% methanol, and sent off to the University of Stirling for histology study. Samples were embedded in paraffin, and finely cut into layers with 5 µm thickness. Hematoxylin & Eosin was used to stain the tissue, and the slides were examined underneath a light microscope. The specimen was classified as “infected” if at least one microsporidian spore patch was observed in the muscle tissue, otherwise it was labeled as “uninfected”.

3.2.3 RNA extraction

In this chapter, all the RNA samples used for microsporidian parasite screening were purified from the gonadal tissue of *E. marinus*. For the QPCR quantification of microsporidian infection intensity, testes from five male amphipods were pooled together for the subsequent RNA extraction by using a Tri-reagent based method (Section 3.2.3.1). For the study of parasite screening in Inverkeithing *E. marinus* population, amphipod specimens were individually examined for microsporidian infection, and the total RNA purification for the individual pair of gonads was carried out by employing an improved RNA purification method combining the Tri-reagent and RNeasy Plus Micro Kit (Qiagen) (Section 3.2.3.2).
3.2.3.1 RNA extraction using Tri-reagent

Tissues were homogenised in 300 μl Tri-reagent using a plastic pestle after which the tissue debris was removed by centrifuging at 13,600 rpm for 5 minutes. The supernatant was subsequently transferred into a fresh 1.5 ml eppendorf tube, and mixed with 65 μl chloroform by vortexing for 10 seconds. The tube was then centrifuged at 13,600 rpm for 10 minutes to facilitate the separation of the aqueous and organic phases after which the upper aqueous layer was carefully transferred into a fresh tube. To precipitate the RNA, the upper aqueous was mixed with an equal volume of iso-propanol (~150 μl) and the sample mixed toughly by inverting the tube several times. After incubating at room temperature for 5 minutes, the tube was centrifuged at 14,000 rpm for 10 minutes. The supernatant was discarded, and the pellet was washed by pre-chilled 75% ethanol, and then dried at room temperature. A total of 20 μl RNase free water was added to each sample to dissolve the total RNA pellet. The RNA solution was subsequently quantified by a spectrophotometer (Nano-drop), and assessed for integrity and genomic DNA contamination by electrophoresis in 2% agarose gel. The RNA solutions were then preserved in -80 °C for future usage.
3.2.3.2 RNA extraction combing Tri-reagent and RNeasy Plus Micro Kit (Qiagen)

Gonadal tissue was homogenised in 250 µl Tri-reagent in an autoclaved eppendorf tube, by using a plastic pestle. Any un-homogenised tissue debris was removed by centrifuging at 13200 rpm for 10 min, and the aqueous was transferred into a fresh tube with 60 µl chloroform. After vortexing for 15 seconds, the tube was centrifuged at 13,000 rpm for 8 minutes. The upper aqueous layer was then transferred to a fresh sterile eppendorf tube where 120 µl (= 1 volume) of 70% ethanol as added. Subsequent isolation adhered to RNeasy Plus Micro Kit (Qiagen) protocol as defined by the manufacturer. The purified RNA was eluted from the hydrophobic affinity column by applying final volume of 20 µl RNase free water. The RNA solution was quantified by a spectrophotometry (Nanodrop, Labtech), and determined for integrity by a Bioanalyzer.

3.2.4 Single-stranded cDNA synthesis

For each RNA sample, approximately 1 µg total RNA was utilised to synthesize the first-strand cDNA by using the reverse transcriptase MMLV (Promega). Both oligo-dT and random hexamers were employed as primers. RNA solution containing 1 µg total RNA was mixed with 1 µl anchored oligo-dT (100 mM) and 1 µ random hexamers (100 mM), and the total volume was then brought up to 12 µl by adding ddH₂O. The mix was incubated at 70°C for 10 minutes and then immediately transferred on ice. A master mix containing 4 µl buffer (5X), 1 µl dNTP (10 mM), 1
μl MMLV, and 2 μl ddH₂O was added to the RNA-primers mix, and gently pippetted several times to achieve homology. The reverse transcription was carried out at 42 ºC for 2 hours.

For parasite screening of Inverkeithing *E. marinus* population, 1 μl aliquot out of the 20 μl first-strand cDNA synthesis reaction was directly used as the template for the Nested-PCR based microsporidian parasite screening. For the cDNA samples used for microsporidian infection intensity quantification by QPCR, QIAquick PCR Purification Kit (Qiagen) was employed to purify the cDNA following the manufacturer’s protocol. The first strand cDNA solution was then quantified by a spectrophotometer and subsequently stored in -20 ºC for further usage.

**3.2.5 Genomic DNA extraction**

The genomic DNA was purified from muscle or gonadal tissue of amphipod specimens by using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer’s protocol. A total of 180μl ATL buffer and 20μl proteinase K were added to the tissue in a sterile Eppendorf tube, which was then incubated in 56 ºC water bath for 3 hours to achieve thorough lyses. Aliquots of 200μl AL buffer and 200μl absolute ethanol were added into the lysis reaction, and the solution was mixed by vortexing for 15 seconds. The mixture was run through the DNeasy Mini spin column by centrifuging at 13600 rpm for one minute, and the flow-through was discarded. Two washing steps were carried out on the column by using 500μl wash
buffer AW1 and AW2, respectively. Genomic DNA was then eluted by 20μl distilled water, and the DNA solution was quantified by a spectrophotometer and then stored in 4 °C fridge for future usage.

3.2.6 Nested PCR based parasite screening method

A nested PCR was conducted on each cDNA sample to determine the microsporidian infection status of the corresponding individual. Primers designed for SSU rRNA of microsporidian were adapted from other studies (Weiss et al, 1994; Baker et al, 1995; MacNeil et al, 2003; McClymont et al, 2005) (Fig. 3.1, Table 3.1). V1f (5'-CACCAGGGTTGATTCTGCCTGAC-3') and 1342AC (5'-ACCGGCCGGTGTGTACAAGGTACAG-3') were used as the forward and reverse primer for the first round PCR, and 18sf (5'-TGTGATTCTGCCTGACGT-3') and 981r (5'-TGTAAGCTGTCCCGCGTTGAGTC-3') for the second round. In order to examine the quality of cDNA or genomic DNA specimens, PCR (LCO1490-HCO2198, Folmer et al, 1994, Table 3.1) amplifying the amphipod’s cytochrome oxidase I (COI) gene was conducted as a positive control.

For each PCR reaction, approximately 100ng DNA template (or a single colony) was amplified using 500nM forward and reverse primers, 50U/ml Taq DNA polymerase buffered with 1x Taq PCR Amplification Buffer (Promega, UK) and 0.2mM of each dNTP. Each reaction was supplemented with an optimised quantity of MgCl2 of between 2 and 2.5mM. The reaction was carried out in a Techne® Flexigene (Techne,
UK), an MWG Primus 25 or 96 thermal cycler (MWG Biotech, UK) using the following programs. PCR amplicons were analysed by 1.0% agarose gel electrophoresis.

**First Round PCR (V1f-1342r):**

<table>
<thead>
<tr>
<th>Temp (°C)</th>
<th>Time</th>
<th>Temp (°C)</th>
<th>Time</th>
<th>Temp (°C)</th>
<th>Time</th>
<th>Temp (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>5:00</td>
<td>94</td>
<td>40</td>
<td>72</td>
<td>1:00</td>
<td>4</td>
</tr>
<tr>
<td>62</td>
<td>1:00</td>
<td>72</td>
<td>1:30</td>
<td>72</td>
<td>7:00</td>
<td>Overnight</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Second Round PCR (18sf-981r):**

<table>
<thead>
<tr>
<th>Temp (°C)</th>
<th>Time</th>
<th>Temp (°C)</th>
<th>Time</th>
<th>Temp (°C)</th>
<th>Time</th>
<th>Temp (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>5:00</td>
<td>94</td>
<td>20</td>
<td>72</td>
<td>1:00</td>
<td>4</td>
</tr>
<tr>
<td>62</td>
<td>1:00</td>
<td>72</td>
<td>1:30</td>
<td>72</td>
<td>7:00</td>
<td>Overnight</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 3.1. The PCR program for microsporidian parasite screening, as well as the position and orientation of primers on microsporidian SSU rRNA sequence.**

Two sets of primers (V1f-1342AC and 18sf-981r) were used for the Nested PCR based microsporidian parasite screening in *E. marinus* populations.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1f</td>
<td>5'-CAC CAG GTT GAT TCT GCC TGAC-3'</td>
<td>Weiss et al, 1994</td>
</tr>
<tr>
<td>1342AC</td>
<td>5'-ACG GGC GGT GTG TAC AAG GTA CAG-3'</td>
<td>Adapted from McClymont et al, 2005</td>
</tr>
<tr>
<td>18sf</td>
<td>5'-GTT GAT TCT GCC TGA CGT-3'</td>
<td>Baker et al, 1995</td>
</tr>
<tr>
<td>981r</td>
<td>5'-TGG TAA GCT GTC CCG CGT TGA GTC-3'</td>
<td>MacNeil et al, 2003</td>
</tr>
<tr>
<td>LCO1490</td>
<td>5'- GGT CAA CAA ATC ATA AAG ATA TTG G-3'</td>
<td>Folmer et al, 1994</td>
</tr>
<tr>
<td>HCO2198</td>
<td>5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3'</td>
<td>Folmer et al, 1994</td>
</tr>
<tr>
<td>DMR</td>
<td>5'-GAT TTC TCT TCC GCA ATA CC AAT-3'</td>
<td>*</td>
</tr>
<tr>
<td>BMR</td>
<td>5'-GAT TTC TCT TCC GCA ATA CAG A-3'</td>
<td>*</td>
</tr>
<tr>
<td>ITSF</td>
<td>5'-AAA GGA AAT TGA CGG AGG AAC ACC-3'</td>
<td>*</td>
</tr>
<tr>
<td>ITS2R</td>
<td>5'-TTA AAG GAG TAY YCG ARC ATC-3'</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 3.1. The primers used for amplifying and sequencing microsporidian SSU rDNA.

* Primers designed by this study.
3.2.7 PCR products cloning

3.2.7.1 PCR products purification

The PCR products were purified from the reaction by employing the QIAGEN QIAquick PCR Purification Kit (QIAGEN, Crawley, UK). Buffer PB (100μl of 5X, supplied by the manufacturer) was added to the 20μl PCR reaction and transferred to a QIAquick column. The column was placed in the supplied collection tube and centrifuged at ~13,000g for 1 minute. The collection tube was emptied and phosphate wash buffer (750μl) was added to the column and centrifuged at 13,000g for 1 minute. The collection tube was emptied and the wash and centrifugation step repeated. The collection tube was emptied again and the column centrifuged for 1 minute at 16,000g. The column was transferred to a new 1.5ml microfuge tube and RNase free water (30μl) carefully added to the centre of the column membrane. The column was then incubated for one minute at room temperature and eluted by centrifugation at ~13,000g for 1 minute.

3.2.7.2 Ligation of PCR products to pGEMT Easy Vector

The PCR products were ligated to pGEMT easy vector (Promega, Southampton, UK) following the manufacturer’s instruction. The insert DNA quantity used was determined from a 3:1 insert:vector ratio calculated using the following equation:

\[
\frac{ng\ of\ vector \times size\ of\ insert\ (kb)}{vector\ size\ (kb)} \times \frac{3}{1} = ng\ of\ insert
\]
3.2.7.3 Plasmid transformation

Plasmid DNA with the ligated PCR product was transformed into *E. coli* DH5α™ competent cells (Invitrogen, UK) by hot shock in 42°C water-bath for 40 seconds, according to manufacturer’s instructions. The transformation was incubated in S.O.C. medium (supplied by the manufacturer) for an hour at 37°C in an incubator shaking at 250rpm. After gentle mixing ensuring cells were in suspension, 200μl was spread evenly onto LB agar plate containing IPTG, S-gal, as well as 100μg/ml ampicillin and incubated for 16 hours at 37°C. White colonies were then individually picked out to a liquid LB medium with ampicillin (100μg/ml), and incubated at 37 °C for 12 hours. Plasmids were extracted from the bacteria culture by using the Mini Prep Kit (Promega) following the manufacturer’s protocol.

3.2.8 DNA sequencing

The amplified SSU rRNA fragment was purified by using the QIAquick PCR purification kit (Qiagen). Samples of cloned plasmids or purified PCR products were sent to the DNA Sequencing Core, Cardiff University. DNA was sequenced using ABI PRISM® BigDye v3.1 Terminator Sequencing technology on the ABI PRISM® 3100 DNA Sequencer. DNA solutions were quantified, and the volume required for Sanger sequencing was calculated by the following equation:

\[
\text{Plasmid DNA volume (μl)} = \frac{290}{\text{Concentration (μg/ml)}}
\]

\[
\text{PCR product volume (μl)} = \frac{60}{\text{Concentration (μg/ml)}}
\]
3.2.9 Microsporidian infection intensity quantification by QPCR

The SYBR Green based quantitative-PCR was employed to quantify the parasite infection intensity. The primer set Msp16sF-Msp16sR was used to amplify a fragment of microsporidian SSU rRNA, and the GAPDH gene was used as the internal control (Table 3.2). Both PCR products were approximately 200 bp. The QPCR was conducted by mixing 12 ng single-stranded cDNA, 10 µl of QuantiFast Sybr Green PCR Master Mix (2X, Qiagen) and 1 µl of each of forward and reverse primer, and distilled water was added to the mixture to bring the total volume to 20 µl. The QPCR reaction was carried out on the ABI PRISM 7900HT QPCR machine using a two-step amplification method following the program in Figure 3.2.

![Fig. 3.2. The program for QPCR quantifying microsporidian parasite infection in E. marinus.](image-url)
The polymerase was activated at 95 °C for 5 minutes to initiate the reaction. DNA was denatured at 95 °C for 3 seconds, and annealed and extended at 62 °C for 30 seconds. A total of 40 cycles were applied for the QPCR, and the fluorescent intensity of the reaction was read after each cycle. A strategy of relative quantification was applied by utilising the Tubulin gene as the internal control, and the relative expression of target genes was normalised to the internal control gene by calculating the delta cycle threshold value (ΔC_T).

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDHF</td>
<td>5'-ATA GTG TCC AAC GCC TCC TG -3'</td>
<td>60.1</td>
</tr>
<tr>
<td>GAPDHR</td>
<td>5'-CCA GTG GAG GAT GGA ATG AT -3'</td>
<td>59.7</td>
</tr>
<tr>
<td>Msp16sF</td>
<td>5'-TCA AGG GCG AAT CCG ATG ATC -3'</td>
<td>58.9</td>
</tr>
<tr>
<td>Msp16sR</td>
<td>5'-CCC CAA AGC CTT ACT TGA TTT CTC -3'</td>
<td>61.0</td>
</tr>
</tbody>
</table>

Table 3.2. Primers of microsporidian SSU and amphipod GAPDH gene. The primers were used for quantifying microsporidian infection in *E. marinus* by QPCR.
3.3 Results

3.3.1 Microsporidian infection identified by histology

A total of 70 male specimens were examined by histology in this study. The microsporidian infection was identified by the present of pink dotted spore patches in the amphipods’ muscle tissue (Figure 3.3). The numbers of infected and uninfected amphipod specimens for each group are presented in the Table 3.3.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Number of Microsporidian infected individuals</th>
<th>Microsporidian infection ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Infected</td>
<td>Uninfected</td>
</tr>
<tr>
<td>Normal males from Thurso</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Normal males from Loch Fleet</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>Normal males from Inverkeithing</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Intersex males from Inverkeithing</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.3. The number and proportion of microsporidian infected *E. marinus* in different *E. marinus* populations. Microsporidian infection was identified by histology, in normal and intersex male *E. marinus* from Inverkeithing, Loch Fleet and Thurso.
Fig. 3.3. The microsporidian spore patches in the muscle tissue of *E. marinus*. The picture 3.2.B is an inset to the 3.2.A, in which the black square stands for the area of 3.2.B.
Approximately 20 normal males were collected from each of the three sites, Thurso, Inverkeithing and Loch Fleet, and microsporidian infection was not observed on all these 59 specimens. Normal and intersex males were collected from Inverkeithing, and 10 out of 11 intersex males were found to be infected. The proportion of microsporidian infected individuals was revealed to be significantly different between normal and intersex males from Inverkeithing (Fisher’s exact test, $p < 0.01$).
3.3.2 Microsporidian infection and parasite species identification in Inverkeithing *E. marinus* population

3.3.2.1 Microsporidian infection ratio in four sexual phenotypes

A total of 112 amphipod specimens collected from Inverkeithing in August 2008, including 36 normal males, 44 normal females, 16 external intersex males and 16 intersex females, were examined for microsporidian infection by a nested PCR based parasite screening method. The majority of specimens in normal female and two intersex groups were revealed to be infected by microsporidians, with 79.5% infection proportion in normal females, 100% in intersex males and 87.5% in intersex females (Table 3.4). The proportion of microsporidian infected individuals in normal males was relatively low, with only 25% of normal males showing the evidence of microsporidian infection.

<table>
<thead>
<tr>
<th>Sex-Phenotype</th>
<th>Number of Specimens Examined</th>
<th>Microsporidian Infection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Normal male</td>
<td>36</td>
<td>9</td>
</tr>
<tr>
<td>Normal female</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>Intersex male</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Intersex female</td>
<td>16</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 3.4. Proportion of microsporidian infected *E. marinus* for each phenotype.

Four sexual phenotypes of *E. marinus* were collected from Inverkeithing, Scotland. The microsporidian infection was identified by a PCR based screening method.
A significant difference on the proportion of microsporidian infected individuals was observed amongst the four sex phenotypes (Fisher’s exact test, p<0.001). Pairwise comparisons were made amongst any of the four groups, with the Bonferroni corrected p value (p < 0.0083). The differences between normal males and any one of the three other groups were all found to be significant (Fisher’s exact test, p<0.0083). There was no significant difference observed between any two groups of normal females and two intersex phenotypes (Fisher’s exact test, p>0.0083).

3.3.2.2 Stochasticism of Nested PCR

The Nested-PCR was found to perform inconsistently when the input DNA template was in a low concentration. A certain degree of randomness was observed when the parasite screening technique was applied on normal males. When the Nested-PCR was repeated on the parasite screened samples, the nine normal males that were initially determined as “microsporidian infected” did not consistently give out positive result. A series of Nested-PCR were repeated on 6 out of the 9 “microsporidian infected” normal males, by using different quantities of input template (Table 3.5). Only 1 out of 6 “infected” normal males gave out positive results, when 1 µl first-stranded cDNA synthesis reaction (approximately 3ng cDNA) was input. Four out of six showed positive when the amount of cDNA template was doubled (6 ng). When the template was increased to 15 ng, the number of specimens showing positive results was brought to five.
Table 3.5. Results of parasite screening on six normal male *E. marinus*.

Three different amounts of input DNA template were used to conduct the screening on six normal male *E. marinus*, which were previously determined as “microsporidian infected” in the initial parasite screening procedure. The results indicate a certain level of stochasticity which is possibly caused by low copy number of microsporidian genes in the template.
 Nested PCR products were purified by using the QIAquick PCR purification kit (Qiagen), following the manufacturer’s protocol. For each sex phenotype (normal male, normal female, intersex male and intersex female), three samples were sequenced in Cardiff University by using the Sanger method. All the three purified PCR products from normal males were found to contain only one sequence, but mixed alignment signals indicating multiple amplified sequences in the same PCR reaction were observed in the specimens from other three groups (Fig. 3.4). Before we could identify the species of the microsporidian infecting this E. marinus population, we temporarily named the two microsporidia parasites Dictyocoela marinum 1 and Dictyocoela marinum 2, following the conventional naming rule of other microsporidians in the Dictyocoela genus.

The raw sequencing data suggested there was an insertion (base A) in one sequence (D. marinum 2) and lead to the one-base shift of the following bases. The extra base A in D. marinum 2 happened to introduce a restriction enzyme digestion site, which was able to be utilised to discriminate the two sequences. The restriction enzyme Mfe I successfully digested D. marinum 2 into two fragments at 37 °C for 1 hour. The digested PCR products were loaded onto a 2% agarose gel for electrophoresis. The intensity of the bands was quantified by the gel image software. Based on the four digested samples, D. marinum 1 was found to be dominant in all the four sex phenotypes, and the ratio of D. marinum 1/D. marinum 2 in normal female, intersex
male and intersex female was between 4:1 and 10:1 (Fig 3.5).

Fig. 3.4. The absorbance curve for base call of Sanger sequencing.
Raw sequencing data were read by the software Finch TV. PCR products of microsporidian SSU from normal males (NM) gave out pure signals. For the other three sexual phenotypes, the first 150 base were pure, whilst mixed peaks were observed in the following sequences. NM: normal male; NF: normal female; IM: intersex male; IF: intersex female.
**Fig. 3.5. Gel picture of digested PCR products.**

PCR products of microsporidian SSU from two normal female *E. marinus*, one intersex male and one intersex female were digested by Mfe I. Band 1, intact sequence (890 bp) after Mfe I digestion; band 2, long fragment (709) after Mfe I digestion; band 3, short fragment (182) after digestion. NF: normal female; IF: intersex female; IM: intersex male.
3.3.2.4 PCR products cloning and sequencing

In order to separate the sequence *D. marinum* 1 from *D. marinum* 2, the PCR products from one normal female specimen were cloned by pGEMT Easy Vector (Promega). Eight colonies were picked out for sequencing. To identify the colony type from *D. marinum* 1 to *D. marinum* 2, PCR using the second round primer set (18sf-981r) was carried out on each colony. The PCR products were subsequently purified, and digested by the restriction enzyme Mfe I. After loading on a 2% agarose gel for electrophoreses, the two types of sequence could be differentiated by the pattern of their bands (Fig. 3.6). *D. marinum* 1 retained intact after Mfe I treatment, and gave out a single band with the length of 890bp. *D. marinum* 2 was digested into two fragments, with the length of 184bp and 707 bp, respectively. The weak band at the size of 890bp in *D. marinum* 2 suggested the purified PCR products were not completely digested. Of the eight colonies, four were found to be *D. marinum* 1 and the others were *D. marinum* 2. Sanger method sequencing was carried out on 6 of the 8 screened colonies, with three replicates for each type.
Mfe I digested PCR products from eight colonies inserted with microsporidian SSU sequence. The bands with longer length (in lane 1, 3, 5 and 6) were *D. marinum* 1, which was intact after digestion. *D. marinum* 2 (in lane 2, 4, 7 and 8) was digested into two short length sequences (182 bp and 709 bp).

*D. marinum* 1 and *D. marinum* 2 were found to have only one base pair difference (an extra A in *D. marinum* 2) during the 890/891 base pair length. A phylogenetic tree of the Dictyocoea genus was generated by Mega 4, by aligning all the available microsporidian SSU sequences on GenBank and the 890/891 SSU fragments, *D. marinum* 1 and *D. marinum* 2, sequenced in this study (Fig. 3.7). *D. duebenum* was found to relate to the sequenced species the most, with 98.4% identity based on the 890/891bp sequence.
Fig. 3.7. Phylogenetic tree of the Dictyocoela genus.

Phylogenetic relationship amongst microsporidian parasites (Inverkeithing microsporidian 1 and 2, which are temporarily named as *D. marinum* 1 and *D. marinum* 2, respectively) infecting *E. marinus* in Inverkeithing and putative microsporidian species from previous studies, based on their SSU rRNA sequence.
3.3.3 Microsporidian infection in an *E. marinus* population in Portsmouth Harbour

3.3.3.1 Sex-phenotype ratio of *E. marinus* in Portsmouth

A total of 476 *E. marinus* were collected from Portsmouth Harbour, by two sampling trips carried out in June and August, 2009. Of these sampled amphipods, three intersex phenotypes were revealed, namely intersex female, internal intersex and external intersex male. The number of each sex-phenotype is shown in Table 3.6. There was no significant difference observed on proportions of the five sex-phenotypes between the two sampling occasions (Chi-square=4.75, p=0.31).

<table>
<thead>
<tr>
<th></th>
<th>Normal Males</th>
<th>Intersex males</th>
<th>Normal Females</th>
<th>Intersex Females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>EIM</td>
<td></td>
<td>IIM</td>
<td></td>
</tr>
<tr>
<td>June</td>
<td>117 (59.7%)</td>
<td>10 (5.1%)</td>
<td>13 (6.6%)</td>
<td>54 (27.5%)</td>
<td>196</td>
</tr>
<tr>
<td>August</td>
<td>143 (51.1%)</td>
<td>11 (3.9%)</td>
<td>26 (9.3%)</td>
<td>97 (34.6%)</td>
<td>280</td>
</tr>
<tr>
<td>Total</td>
<td>260 (54.6%)</td>
<td>21 (4.4%)</td>
<td>39 (8.2%)</td>
<td>151 (31.7%)</td>
<td>476</td>
</tr>
</tbody>
</table>

**Table 3.6. The sex ratio of Portsmouth *E. marinus* population in June and August.**

Amphipod specimens were collected from Portsmouth Harbour (England) in summer 2009. EIM: external intersex male; IIM: internal intersex males.
Male amphipods outnumbered females in both samplings. Intersex females were found to be very rare in this site, making up only 1% of 476 amphipods collected from the field. The proportion of internal intersex male (8.2%) was higher than external intersex male (4.4%). Amongst the 12 external intersex males examined in this study, 11 of which lacked an oviduct on their testes, and there was only one intersex specimen possessing both internal (oviduct) and external (brood plates) intersex characters.

3.3.3.2 Microsporidian infection in normal and intersex E. marinus

A sub-sample of 48 amphipods from Portsmouth Harbour, with 12 specimens for each 4 phenotypes (normal male, normal female, internal intersex and external intersex male), were examined for microsporidian infection by a PCR based screening method. Due to the low number collected from the field, intersex females were not statistically investigated in this study. The initial screening applied the non-species-specific primer set (18sf-981r). According to the band intensity, the infection status of amphipods was classified into three groups, namely high, low infection intensity and uninfected. Examples of various band intensities are showed in Figure 3.8. The band for the muscle tissue of the sample NF1 present the highest intensity as “low infection”, and other bands were classified as “low infection” or “high infection” if the intensity was less or more than the muscle tissue of NF1 (Figure 3.7). Both high and low infection were categorised as “infected” when the microsporidian infection ratio was calculated for each group (Table 3.6).
Fig. 3.8. Gel picture for microsporidian parasite screening in *E. marinus*
Microsporidian SSU sequence was amplified from gonadal (above) and muscle (below) tissue of *E. marinus*. Eight samples were investigated for each of the four sexual phenotypes. NM: normal male; NF: normal female; IM: intersex male; IF: intersex female.
For the three male phenotypes, the initial parasite screening on muscle and gonadal tissues gave out consistent results. The majority of external intersex males (83.3%) were found to be infected by microsporidians, and the microsporidian infection ratio in normal males and internal intersex males was 33.3% and 50%, respectively (Table 3.7). For normal females, a greater proportion of microsporidian infection rate was observed in gonads (50%) than in muscle tissue (25%), but the additional infections in gonads were contributed by individuals with low infection level (Table 3.7). All the samples showing negative result for microsporidian infection gave out substantial bands when the CO1 gene amplification was conducted on them.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Sex phenotype</th>
<th>Microsporidian infection</th>
<th>Infection rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Gonads***</td>
<td>NM</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>IIM</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>EIM</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NF</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Muscle**</td>
<td>NM</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>IIM</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>EIM</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>NF</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.7. The number of microsporidian infected *E. marinus* in 4 sex phenotypes

*E. marinus* were collected from Portsmouth Harbour (England). NM: normal male; IIM: internal intersex male; EIM: external intersex male; NF: normal female. ** p<0.01; *** p<0.001.
Significant difference of proportions of microsporidian infection was observed amongst the four sex phenotypes, in both gonadal (Fisher’s exact test; p<0.001) and muscle tissue (Fisher’s exact test; p<0.01). Proportions of parasite infected individuals were also revealed to be significantly different amongst the three male phenotypes (Fisher’s exact test; p<0.0125, Bonferroni correction). Significant differences were observed in the proportion of microsporidian infection between normal males and external intersex males, as well as between normal females and external intersex males, in both tissues (Fisher’s exact test; p<0.0083, Bonferroni correction). No significant difference was found in the proportion of microsporidian infection between normal males and normal females (Fisher’s exact test; p>0.0083).

3.3.3.3 Microsporidian species identification

Six PCR products from the initial parasite screening were sequenced by the Sanger method, and two types of SSU sequence were revealed, sharing 92.9% identity out of the 800bp length sequence. For each sequence, the entire SSU rDNA gene including the ITS region was sequenced by PCR amplification (V1f-1342AC, ITSF-ITS2R) followed by Sanger method sequencing of the purified fragments. The two overlapped fragments were then clustered and the sequences were blasted against GenBank (NCBI). One sequence was found to have 98.7% similarity to Dictyocoela berillonum, and the other shared 98.6% identity to Dictyocoela duebenum. In the following sections, the two microsporidian species infecting Portsmouth E. marinus population were called D. berillonum and D. duebenum, respectively. A phylogenetic
tree was generated by aligning the whole SSU rDNA sequences of the two microsporidian parasites to other microsporidian species from the genus Dictyocoela on GenBank. The two microsporidian species revealed in this study were found to be closely related to *D. berillonum* and *D. duebenum* respectively (NCBI, MEGA 4, Fig. 3.9).

**Fig. 3.9. A phylogenetic tree of genus Dictyocoela.**
The phylogenetic relationship of putative microsporidians from previous studies, and the two parasites revealed in Portsmouth *E. marinus* population (*D. duebenum* Portsmouth and *D. berillonum* Portsmouth), based on their SSU rDNA sequence. Scale bar = 1% sequence divergence.
3.3.3.4 Differential distribution of the two microsporidian parasites

This study designed two species-specific primers (BMR, DMR) which were successfully used to discriminate the two parasites species, *D. duebenum* and *D. berillonum*, by a single PCR reaction (Fig. 3.10).

![Fig. 3.10. Alignment of the SSU rRNA sequence of two microsporidiants.](image)

Two microsporidian species were revealed in Portsmouth *E. marinus* population. Species-specific primers were designed on the underlined divergence sequence of the two parasites, *D. duebenum* and *D. berillonum*.

The 48 amphipods initially examined for microsporidian infection were re-screened for parasite species by the species-specific PCR. Both gonadal and muscle tissues were assessed for each specimen. Examples of the species-specific parasite screening were showed in Figure 3.11, and the results were elucidated in Table 3.8. Of these 48 amphipods, only one microsporidian species, *D. berillonum*, was found in three male phenotypes, whilst both *D. berillonum* and *D. duebenum* were revealed in normal females.
Fig 3.11. Gel pictures for species-specific microsporidian parasite screening.
Examples of species-specific parasite screening on the muscle tissue of the 48 sub-samples of Portsmouth E. marinus population. Each specimen was screened by the two species-specific primers, DMR (D, for D. duebenum) and BMR (B, for D. berillonum), respectively. NM: normal male; NF: normal female; IM: intersex male; IF: intersex female.
<table>
<thead>
<tr>
<th>Microsporidia Infection</th>
<th>Tissue</th>
<th>Sex phenotype</th>
<th>High</th>
<th>Low</th>
<th>Uninfected</th>
<th>Infection rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>D. berillonum</em></td>
<td>Gonads***</td>
<td>NM</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>66.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IIM</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>66.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EIM</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NF</td>
<td>0</td>
<td>7</td>
<td>5</td>
<td>58.3%</td>
</tr>
<tr>
<td></td>
<td>Muscle***</td>
<td>NM</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>41.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IIM</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EIM</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>83.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NF</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>16.7%</td>
</tr>
<tr>
<td><em>D. duebenum</em></td>
<td>Gonads</td>
<td>NM</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IIM</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EIM</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NF</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Muscle</td>
<td>NM</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IIM</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EIM</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NF</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>33.3%</td>
</tr>
</tbody>
</table>

Table 3.8. The number of *D. berillonum* and *D. duebenum* infected *E. marinus* in four sexual phenotypes.

*E. marinus* were collected from Portsmouth Harbour (England). NM: normal male; IIM: internal intersex male; EIM: external intersex male; NF: normal female. ** *p*<0.01; *** *p*<0.001.

Both tissues showed significant difference on the *D. berillonum* infection amongst the four sex phenotypes. No significant difference was observed between normal male and normal female (Fisher’s exact test, *p*>0.05). The difference in the *D. berillonum* infection amongst the three male phenotypes was also revealed to be significant (Fisher’s exact test, *p*<0.0125, Bonferroni correction) in both muscle and gonadal tissues. Significant differences were found between normal and external intersex male in both tissues (Fisher’s exact test, *p*<0.0083, Bonferroni correction).
No significant difference was observed between internal intersex males and either normal males or external intersex males (Fisher’s exact test, p>0.05).

*D. duebenum* was only found in normal females, and the infection rate was slightly higher in muscle (33.3%) than in gonads (25%), and the additional infected sample had low infection intensity. The difference of *D. duebenum* infection rate between normal male and normal female failed to be significant (Fisher’s exact test, p>0.05). In addition to the statistically studied 48 specimens, three intersex females were screened for parasite infection. All of the three intersex females were found to be infected by *D. berillonum*, and no *D. duebenum* was revealed.

In order to assess the transmission route of the two microsporidian species, PCR employing species-specific primers was conducted on the genomic DNA of eggs from one *D. berillonum* infected intersex female, one *D. duebenum* infected and one uninfected normal female. Eggs were found to have the same parasite infection status as their mothers, with *D. berillonum* infection in the eggs from intersex female and *D. duebenum* in those from the infected normal female. No microsporidian infection was detected in the eggs from the uninfected normal female.
3.3.4 Microsporidian bi-infection in Inverkeithing and Portsmouth *E. marinus*

The species-specific primers DMR and BMR were employed to identify the microsporidian parasites infecting Inverkeithing *E. marinus* population. Three out of six Inverkeithing amphipod specimens which were screened by the species-specific primers DMR and BMR showed bi-infection of microsporidians. To validate the evidence of microsporidian bi-infection, two Inverkeithing specimens (one external intersex male and one intersex female) which showed evidence of bi-infection by microsporidian parasites and four Portsmouth samples (normal male, internal intersex male, external intersex male, and normal female) were screened for microsporidians using the species-specific primers, and the PCR products were purified and sequenced (Fig. 3.12).
Species-specific microsporidian parasite screening was carried out on four Portsmouth *E. marinus* samples and two Inverkeithing *E. marinus* samples. NM: normal male; IIM: internal intersex male; EIM: external intersex male; NF: normal female; IF: intersex female; P: Portsmouth; I: Inverkeithing. D: PCR product of *D. duebenum* SSU rRNA sequence; B: PCR product of *D. berillonum* SSU rRNA sequence.

The three male phenotypes from Portsmouth were found to be infected by *D. berillonum*, and the normal female was infected by *D. duebenum*. No evidence of bi-infection was observed on the muscle tissue of these four Portsmouth amphipods, whilst both of the two Inverkeithing intersexes had two bands indicating they were infected by at least two distinct microsporidian species. PCR products of the Inverkeithing external intersex male and intersex female were then purified and sequenced by the Sanger method. The sequencing results confirmed that the PCR products amplified by the primers BMR and DMR were different sequences.

The sequence amplified by the primer v1f-BMR from the two Inverkeithing
intersexes was subsequently blasted again GenBank. Although *D. duebenum* was revealed as the closest related species, the two shared only 90% identity (Fig. 3.13). A blast search was also carried out on the sequence amplified by the primer DMR, and *D. duebenum* was found to be the closest species sharing 95% similarity. The microsporidian with the SSU sequence amplified by the primer BMR is temporarily named *D. marinum* 3, and was used for the following sections.

**Fig. 3.13.** BLAST search results of SSU rRNA sequence of *D. marinum* 3. *D. marinum* 3 was found to be a rare microsporidian infecting the *E. marinus* population in Inverkeithing. The SSU rRNA sequence was amplified by the species-specific primer DMR, and the BLAST search was carried out against the nucleotide database of GenBank (NCBI).
Bi-infection was also observed in the ovaries of some normal females from Portsmouth *E. marinus* population. All the infected males, including internal and external intersex males, were found to harbour *D. berillonum* only. For normal females, 2 (NF1 and NF7) out of 8 specimens were found be infected by *D. duebenum* in the muscle tissue and bi-infection by *D. duebenum* and *D. berillonum* in ovaries (Fig. 3.14).

**Fig 3.14. Gel picture for species-specific microsporidian parasite screening.**
Eight normal female *E. marinus* from Portsmouth population were examined for microsporidian infection by using two species-specific primers. Both the muscle (A) and gonadal tissues of the eight normal females were assessed. The presence of microsporidian was indicated by its corresponding band. NF: normal females; D: PCR product of *D. duebenum* SSU rRNA sequence; B: PCR product of *D. berillonum* SSU rRNA sequence.
The species-specific microsporidian parasite screening was repeated on four microsporidian infected individuals from Portsmouth, representing four sexual phenotypes, namely normal male, internal intersex male, external intersex male and normal female. The three males (normal male, internal intersex male and external intersex male) were found to be infected by *D. berillonum* only. Bi-infection was observed in the gonadal tissue of the normal female, whilst only *D. duebenum* infection was detected in its muscle tissue (Fig. 3.15).
Fig. 3.15. Gel picture for species-specific microsporidian screening.

Four individuals from Portsmouth *E. marinus* population were examined for microsporidian infection by using two species-specific primers. Both gonadal (A) and muscle tissues (B) of the four samples were investigated. NM: normal male; IIM: internal intersex male; EIM: external intersex male; NF: normal female; D: PCR product of *D. duebenum* SSU rRNA sequence; B: PCR product of *D. berillonum* SSU rRNA sequence.
3.3.5 Microsporidian infection intensity quantification by QPCR

A total of 15 *E. marinus* gonadal cDNA samples (each sample was pooled from five individuals of the same phenotype) representing five replicates for each of the three male phenotypes (normal male, internal and external intersex male), were assessed for microsporidian infection quantification by QPCR. Microsporidian infection intensity was assessed by quantifying the SSU rRNA sequence in cDNA samples. GAPDH was used as the internal control for the relative quantification of targeted genes. According to their C\textsubscript{T} value, the variation of the internal control gene GDPDH amongst 15 samples was found to be within 3 times difference (Fig. 3. 16).

![Fig. 3.16. The amplification curves of the internal control gene](image)

Each amplification curve represents one pooled cDNA sample of *E. marinus*, and different samples were displayed by different colours. Rn is the fluorescence signal normalised to the control signal of ROX.
The relative abundance of microsporidian rRNA SSU sequence in each cDNA sample was represented by the delta C\textsubscript{T} value normalised to internal control gene GAPDH. The level of microsporidian rRNA SSU were revealed to be significantly higher in external intersex males than normal and intersex intersex males (p<0.05) (Figure 3.17). Only one out of five (25%) internal intersex male samples showed high level of microsporidian infection, whilst the other four internal intersex male samples as well as all the five normal male samples had extremely low level of microsporidian rRNA SSU sequence, indicating they are not infected by microsporidian parasites.

**Fig. 3.17. Microsporidian infection intensity in three male phenotypes.**
The microsporidian infection was assessed by quantifying the microsporidian SSU rRNA and normalised to the internal control gene GAPDH. The star in the internal intersex male group is an outlier sample having high microsporidian infection intensity. NM: normal male; IIM: internal intersex male; EIM: external intersex male.
3.4 Discussion

Microsporidian parasites have been reported in a variety of amphipod species, such as *Gammarus duebeni*, *Gammarus pulex*, *Corophium volutator* and *Orchestia mediterrane* (Gingsburger-Vogel, 1991; Dunn et al, 1993a; Terry et al, 2004; Mautner et al, 2007). During this investigation, a preliminary study used a histology-based technique to identify the prevalence of microsporidian parasites in male *E. marinus* from three Scottish sites, Inverkeithing, Loch Fleet and Thurso. No evidence of microsporidian infection was observed in any of the 59 normal males, whilst 10 out of 11 intersex males from Inverkeithing were found to be infected by microsporidian parasites. The results were accord with previously studies on the *E. marinus* population in Inverkeithing, which showed greater proportion of microsporidian infected individuals in intersexes than in normal amphipods (Ford et al, 2006).

*E. marinus* populations in Inverkeithing and Portsmouth Harbour were further investigated for the distribution of microsporidian infection and the parasite species were also identified by employing a PCR based parasite screening method. Three microsporidian species, which were named as *D. marinum* 1, *D. marinum* 2 and *D. marinum* 3, were revealed in the Inverkeithing population based on the SSU rRNA sequence. *D. marinum* 1 and *D. marinum* 2 were found to co-occur in a large number of *E. marinus* samples, and have only one base pair difference (one base pair
insertion in *D. marinum* 2) based on the 890/891 SSU rRNA sequence. Both showed the similarity of 98.4% with *D. duebenum*, which has been reported as a feminising parasite in a variety of amphipods, such as *Gammarus duebeni*, *Gammarus tigrinus* and *Echinogammarus berilloni* (Terry et al, 2004). Both *D. marinum* 1 and *D. marinum* 2 were then classified as *D. duebenum* since their SSU rRNA sequence divergence is below 2% (Terry et al, 2004). According to the BLAST result to NCBI GenBank, *D. marinum* 3 was also found to relate to *D. duebenum* the most, but with the identity of only 90% based on the SSU rRNA. The divergence of SSU sequence amongst microsporidian species in the genus Dictyocoela is 4-11% (Terry et al, 2004). It is possible that *D. marinum* 3 is not only a potential new species, but also could be in a new genus. Two microsporidian species were revealed in the population from Portsmouth Harbour. Based on the entire SSU rRNA sequence including the ITS region, the two parasites were discovered to closely relate to *D. duebenum* (98.6%) and *D. berillonum* (98.7%), respectively, and are subsequently referred as *D. duebenum* or *D. berillonum* in this study.

For Inverkeithing amphipods, direct sequencing on Nested PCR products of microsporidian parasite screening revealed that PCR products from three normal males were purely *D. marinum* 1, but *D. marinum* 1 and *D. marinum* 2 were observed to coexist in the intersex males and females plus normal females. *D. marinum* 1 was found to be 4-10 times more abundant than *D. marinum* 2. The difference (with only one base pair difference out of 890/891 bp length sequence)
between the two strains was too small to claim them as two distinct species. Although rRNA sequence is conservative and is supposed to homogenise during the evolution (Arnheim, 1983), the polymorphism of SSU rRNA sequence has been reported in some microsporidian species (Tay et al, 2005). O’Mahony et al (2007) reported that there were 2 or 3 rRNA sequence variants in a single *Nosema bombi* spore. The two sequences *D. marinum* 1 and *D. marinum* 2 are possibly two formats of SSU rRNA sequence of the same species.

Both the histology and PCR based parasite screenings revealed significantly greater proportions of microsporidian infected individuals in intersexes than their normal counterparts. However, a certain level (25%) of microsporidian infection was observed in normal males from Inverkeithing when applying the molecular technique, whilst no evidence of microsporidian infection was found in 20 normal males from the same site using histology. The difference between proportions of microsporidian infected normal males determined by the two techniques was found to be statistically significant (Fisher’s exact test, p<0.05). Microsporidian identification by molecular techniques has been reported to be more sensitive than light microscope based histology. Webber et al (1999) reported the threshold of microsporidian detection for PCR based screening method was $10^2$ spores per ml of diluted stool sample from microsporidian infected patients, whilst at least $10^4$ spores per ml was required for light microscope to detect the microsporidian parasite. The intensity of microsporidian infection in normal males was possibly in relatively low levels.
therefore it could be detected by the sensitive molecular technique but not by histology. The stochasticism observed in normal male infection status determined by PCR also indicated that the microsporidian infection intensity was very low in normal males. It is possible that parasite burden in normal males was not sufficient enough to cause substantial feminising effects (Kelly et al, 2002). The observation that only *D. marinum* 1 was detected in normal males could be simply due to the low parasite infection intensity in normal males as well as the dominant prevalence of *D. marinum* 1 over *D. marinum* 2.

Some vertically transmitted microsporidia have been discovered to exert feminising effects on their amphipod host, and they increase their survival chance by distorting the host’s sex ratio in the favour of females, since the majority of vertically transmitted parasites use the maternal line and male hosts are biological “dead-end” (Terry et al, 1998; 2004; Ironside et al, 2003a; Haine et al, 2004). There has been strong evidence showing the association between microsporidian infection and the intersexuality in amphipods, and intersexes are presumed as the result of incomplete feminisation by microsporidian parasites (Ginsburger-Vogel, 1991; Kelly et al, 2002; Rodgers-Gray et al, 2004). The significant difference in the proportion of microsporidian infection between normal and intersex male *E. marinus* suggests high co-relationship between microsporidian infection and the occurrence of intersexuality in this amphipod species. Sterile intersex isopods also have been reported induced by incomplete feminising effects of *Wolbachia* at high temperature
(Rigaud and Juchault, 1998). Ginsburger-Vogel (1991) reported intersexuality in amphipod could be induced by grafting microsporidian infected tissues into an uninfected individual. Kelly et al (2004) reported that intersex *Gammarus duebeni* were only found in the offspring produced by females infected by the feminising parasite *N. granulosis*, which suggested that the intersexes were induced by incomplete feminisation of the microsporidian parasite.

In order to further study the relationship between microsporidian infection and the two male intersex phenotypes in amphipods, an *E. marinus* population in Portsmouth was examined for the prevalence of microsporidians in normal male and females, as well as in two male intersex phenotypes -- internal and external intersex males. Significant difference was observed in the proportion of microsporidian infection between normal and external intersex males, which was consistence with results of the studies carried out in Inverkeithing *E. marinus* population. Two microsporidian species, *D. berillonum* and *D. duebenum*, were revealed coexisting in this *E. marinus* population. Both microsporidians have been reported in various amphipods, and the two parasites shared a list of host species, for example *Echinogammarus berelloni*, *Gammarus duebeni* and *Gammarus tigrinus* (Terry et al, 2004). Both parasites were found to apply vertical transmission, which is consistent with other studies (Terry et al, 2004). The microsporidian parasite infecting Inverkeithing population was found to be closely related to *D. duebenum*, whilst it only showed 97.2% identity to the *D. duebenum* revealed in Portsmouth *E. marinus* population, based on the sequenced
1149 bp SSU rRNA sequence. Although the difference on *D. duebenum* infection rate between normal male and normal female from Portsmouth population was not significant, the fact that *D. duebenum* was only found in normal females suggested this parasite is a potential sex-distorter in this amphipod population. However, further studies would need to be conducted to confirm this.

According to a theoretical model, multiple vertical transmission parasites are not able to stably coexist in a single host population, since a less efficient species will be excluded by others (Ironside et al, 2003a). In fact, co-occurrence of different microsporidian species in the same amphipod population has been reported by several studies (Hogg et al, 2002; Ironside et al, 2003a; Haine et al, 2004). The empirical evidence of multiple vertical transmission parasite coexistence in the same population suggested some assumptions of the theoretical model might have been violated, for example the ratio of infected and uninfected hosts might be prevented from equilibrium by seasoning variations (Ironside et al, 2003a). It is also possible that the assumption of pure vertical transmission applied by parasites was not met, and microsporidian parasites revealed in this study might be able to apply both vertical and horizontal transmissions. Microsporidian from the genus Dictyocoela were reported to infect only gonadal tissue of their amphipod hosts (Terry et al, 2004). In contrast, the study on the Portsmouth *E. marinus* population revealed that the two microsporidian species, *D. berillonum* and *D. duebenum*, infected both muscle and gonadal tissues of *E. marinus*. Generally, a higher proportion of *D.*
berillonum infection was observed in hosts’ gonadal tissue than in muscle.

* D. berillonum also has been reported in a variety of amphipods, but it has never been reported as a sex-distorter. *D. berillonum* has been reported to infect *E. marinus* which were sampled from Northern Ireland in the UK, and no significant difference was observed on the microsporidian infection rate between male and female specimens (Terry et al, 2004). In this study, a correlation between *D. berillonum* infection and male intersexuality has been observed, but the relationship between the two was found to be not conclusive. There were two normal males infected by *D. berillonum* and showed no evidence of female characters. The fact that some intersex males lacked microsporidian infection indicated that there might be other factors causing intersexuality in this *E. marinus* population. The later exclusion of microsporidi ans by their intersex hosts could not be ruled out.

In this study, bi-infection of microsporidian parasites in a single amphipod host was observed in both Inverkeithing and Portsmouth *E. marinus* populations. Parasites infecting the same host could have different interests in the host, and those different needs from the host will usually conflict (Dunn et al, 1993a). Two microsporidian species, *D. duebenum* and *D. berillonum*, were discovered to co-infect ovaries of a single normal female host in Portsmouth. In Inverkeithing, bi-infection was observed in both gonadal and muscle tissue of three *E. marinus*. The proportion of *E. marinus* infected by two microsporidian species (*D. duebenum* and *D. marinum 3*) have not
been investigated in large scale, but the ratio that three out of the six specimens showed bi-infection indicated that it is not an exceptional incidence.

To date, a variety of techniques have been used to identify the microsporidian infection, such as histology, DAPI staining, and PCR based parasite screening (Terry et al, 1998; Garcia 2002; MacNeil et al, 2003; Kelly et al, 2004). The majority of microsporidian detection methods identify the occurrence of infection, and techniques for infection intensity quantification are limited. Given that the extreme sensitivity, the PCR technique allows even a trace amount of microsporidians to be detected. It is important to ascertain what level of infection is biologically relevant since trace amount of microsporidians may be introduced by recent intake of spore contaminated food, or attachment of spores from the environment. It is especially essential to quantified the parasite infection intensity, when pooled samples are compared, since the conventional PCR could not discriminate the difference from one infected individual in the pool to all infected individuals pooled together.

In this chapter, QPCR was employed to quantify the microsporidian infection intensity in pooled specimens. The microsporidian infection intensity was assessed by quantifying the microsporidian SSU rRNA in cDNA sample generated from amphipods’ gonads. Three male phenotypes -- normal male, internal and external intersex males -- were compared in this study, with five replicates in each group. All of the five external intersex male samples showed high level of microsporidian SSU
rRNA. The majority of internal intersex males and normal males had low amount of microsporidian SSU rRNA, with only one internal intersex male sample had the level of microsporidian SSU rRNA as high as the external intersex males. The result is consistent with the data of microsporidian parasite screenings carried out in Inverkeithing and Portsmouth E. marinus populations. It suggests that microsporidian infection is closely related to external intersex males but not necessary to the internal intersex males. There are possibly other factors, for instance pollutants that compromise the androgenic gland activities (Brown et al, 1999; Rodriguez et al, 2007), might cause the phenotype of intersex intersex males in E. marinus.

This study adds weight on the strong association between microsporidian infection and intersexuality in both E. marinus populations in Inverkeithing and Portsmouth. In both sites, the majority of external intersex males were found to be infected, whilst the relationship between microsporidian infection and internal intersex males was found to be much less convincing, which suggests the two male intersex phenotypes might be induced by different mechanisms. Although a certain proportion of normal males in Inverkeithing were found to be infected by microsporidians, there was evidence showing the infection intensity was relatively low. By sequencing the microsporidian SSU rRNA, the parasite species were identified. The microsporidian infecting the Inverkeithing population was found to be D. duebenum, whilst two microsporidian species -- D. berillonum and D.
*duebenum* -- were observed in the Portsmouth population. Both microsporidian species were found to apply vertical transmission, and infect both muscle and gonadal tissues of *E. marinus*. SYBR green based QPCR was successfully applied to quantify the microsporidian infection intensity in three male phenotypes, and provided a more accurate method to determined the microsporidian infection status comparing to qualitative techniques.
Chapter 4

Transcriptomic Comparison of Normal and Intersex Male *E. marinus* Using a Cross-Species DNA Microarray
4.1. Introduction

In the previous chapters (Chapter 2 and 3), two intersex male phenotypes were identified from the *E. marinus* populations differentiated by the internal or external appearance of secondary female characteristics, and thus classified as internal intersex (IIM) and external intersex males (EIM), respectively. Physiological differences, such as impaired fertility, increased body weight and elevated moult frequency were observed in intersex *E. marinus* compared to normal specimens (Ford et al, 2003a; 2004b; Yang et al, 2008; Chapter 2). It is not clear whether these observed intersex phenotypes are driven by de-masculinisation or through alternative perturbations of the endocrine axis. In order to elucidate the mechanism of intersexuality in this amphipod species, we determine to investigate the alterations in gene expression that characterised the dysfunction in sexual phenotype by comparing the transcriptomic profile between normal and intersex specimens.

A variety of powerful tools have been identified allowing assessment of differential gene expression, with Shimkets (2004) reporting in excess of 27 different techniques ranging from variety of medium and high density solid phase arraying techniques (described as macro- and micro-arraying) to sequence based counting approaches using Sanger and Next Generation Sequencing technology. A DNA microarray enables the comparison of expression levels of thousands of genes in parallel, and has been widely exploited in ecotoxicology studies to derive mode and mechanism
of action of toxicants in a wide range of receptor organisms, such as rats, fish and earthworm (Hamadch et al, 2002; Hoyt et al, 2003; Bundy et al, 2008). A brief history and the working principle of the microarray is provided in the general introduction of this thesis (Chapter 1, Section 1.5.1).

Since the ability to fabricate a microarray is dependent on the availability, either physically or informatically, of reporter sequences which define the target transcriptome, these techniques were initially limited to those organisms where significant genetic information or resources had been determined. However, studies have revealed that adequate sequence identity shared by phylogenetic-related species allows cross-species microarray hybridization to be successfully carried out (Buckley, 2007). This approach of applying a microarray constructed from the cDNA sequences of one species onto the transcriptomic reporters generated from a second species was termed as “heterologous hybridisation” (Renn et al, 2004). Heterologous hybridisation saves the time and expenses allowing a microarray platform constructed for a species to be applied to close relatives without similar level of genetic knowledge or resources being required therefore making it a feasible approach to study gene expression patterns in non-model species (Buckley, 2007).

A cross-species DNA microarray was employed to compare gene expression between normal males and the two distinct intersex males. The DNA microarray used in this study was originally designed for Gammarus pulex (Sambles 2008) and represented
12644 ESTs isolated from a mixed cDNA library was generated from approximately 200 *G. pulex* individuals of various genders, life stages and moult cycles, together with male- and female-specific transcripts isolated using Suppression Hybridisation. Bioinformatical analysis of these ESTs showed them to represent 3917 gene objects. The cloned cDNAs and cDNA fragments were amplified and the resultant PCR products purified and spotted on epoxysilane coated glass slide, Schott Nexterion Slide E (SCHOTT Jenaer Glas GmbH) by employing a SpotArray72 Microarray Printing System (PerkinElmer LAS (UK) Ltd.). A total of 13728 spots (for 12644 *G. pulex* ESTs, as well as blanks, landmarks and ratio controls for Lucidea Universal ScoreCard spike) were printed onto an array, with 100 μm diameter of each spot (Sambles, 2008). The current study aims to use data generated from deploying the cross-hybridisation to determine:

i) EIM and IIM have the same molecular signature.

ii) The molecular pathways underlying the intersex phenotype and

iii) To develop potential molecular biomarkers for male intersexuality in crustaceans

This analysis was designed to use multi-variant analysis to compare the change in global gene expression profiles of the three populations (normal males, internal intersex males and external intersex males), to identify genes displaying statistically significant changes in gene expression amongst the three male phenotypes and map these differentially expressed genes on known pathways and biological processes.
Finally genes showing discriminatory expression profiles will be validated using an independent transcript quantification methodology.
4.2. Materials and Methods

4.2.1 Sampling

*Echinogammarus marinus* specimens were collected underneath seaweed and stones in an industrially impacted site, Inverkeithing, Scotland, in the middle of October, 2007. Animals were anaesthetised in the carbonated seawater, and their sex-phenotype was identified using visual inspection under a stereo microscope. According to their secondary sex characteristics and internal testicular structure, male specimens were further classified into three male phenotype groups, namely normal (NM), internal (IIM) and external intersex males (EIM) (Chapter 2, Section 2.1). Male specimens were sacrificed, and the testes were dissected out and preserved in RNAlater ICE at -80 °C. Five pairs of testes from the same male phenotype were pooled together, as a single sample for the following microarray experiment. Five replicates were employed for each male phenotype group.

4.2.2 RNA purification

All the plastic wares used for RNA extraction were double-autoclaved to deactivate RNase. The testes samples which had been preserved at -80 °C in RNAlater ICE (Chapter 4, Section 4.2.1), were taken out from the solution and any excessive buffer was removed by dapping the sample gently on a sterile tissue. Total RNA was then purified from the testes using Tri-reagent (Ambion) employing the procedure recommended by the supplier with the refinements described in Chapter 3 (Section
3.2.3.1). 

4.2.3 DNA Microarray

4.2.3.1 First-stranded cDNA Synthesis and Labeling

A post-labeling strategy was exploited to generate fluorescence labeled cDNA probes. The generation of the aminoallyl labeling first-stranded cDNA was achieved by mixing following components, anchored oligo(dT)$_{18}$ (1μl of 100mM), random hexamers (2μl of 100mM), total RNA (13.8μl containing 10μg) and Lucidea ScoreCard spike (1μl, GE Healthcare, Little Chalfont, Bucks., UK) to make a total volume of 17.8μl. The mixture was vortexed and centrifuged briefly and heated in a thermocycler to 70°C for 10 minutes. The reverse transcriptase mix was prepared with 5x First Strand Buffer (6μl, supplied by Invitrogen), DTT (3μl of 0.1M), aa dUTP/dNTP mix (1.2μl, freshly prepared at a ratio U:T:ACG=4:1:5) to a total volume of 10.2μl per sample. RT mix (10.2μl) was added to each samples and vortexed; the tube contents pulsed down and incubated at 25°C for 2 minutes for optimal nucleotide binding of random hexamers.

An aliquot of 2μl reverse transcriptase (Superscript II, Invitrogen) was individually added to each reaction tube to give a total volume of 30μl. The contents were gently mixed by flicking the tube and pulsing down the contents. After an incubation in the thermocycler at 42°C for 3 hours, NaOH (10μl of 1M) and EDTA (10μl of 0.5M, pH 8.0) were added to each reaction, and the 50μl mixture was briefly vortexed and
pulsed down by centrifuging. The EDTA compromises the activity of the reverse transcriptase and the following incubation completely inactivates the enzyme. The samples were heated to 65°C and incubated for 15 minutes. To neutralise the solution, HEPES buffer (25μl of 1M, pH 7.0) was added, and the contents briefly vortexed and centrifuged giving a total volume of 75μl.

4.2.3.2 cDNA purification

This method is a modification of the QIAGEN QIAquick PCR purification kit (QIAGEN, Crawley, UK). The phosphate and wash buffers are substituted for the QIAGEN supplied buffers supplied by the manufacturer because they contain free amines which compete with the CyDye in coupling reaction. Buffer PB (375μl of 5X, supplied by the manufacturer) was added to the 75μl cDNA solution and transferred to a QIAquick column. The column was placed in the supplied collection tube and centrifuged at ~13,000g for 1 minute. The collection tube was emptied and phosphate wash buffer (750μl) was added to the column and centrifuged at 13,000g for 1 minute. The collection tube was emptied and the wash and centrifugation step repeated. The collection tube was emptied again and the column centrifuged for 1 minute at 16,000g. The column was transferred to a new 1.5ml microfuge tube and RNase free water (30μl) carefully added to the centre of the column membrane. The column was then incubated for 1 minute at room temperature and eluted by centrifugation at ~13,000g for 1 minute. A second elution was performed into the same tube by repeating the elution with fresh RNase free water to give a final elution
volume of 60μl. The samples were dried down in the SpeedVac (Eppendorf UK Ltd., Cambridge, UK) to a final volume of 3μl. If cDNA solution was completely dried out, 3 μl RNase-free water was added into the tube, and resuspension was assisted by heat shocking at 65°C for 1 minute and then returning the samples to ice for 5 minutes.

4.2.3.3 CyDye and cDNA Coupling

The coupling of CyDye and amine-labeled cDNA (Fig. 4.1.) requires mildly alkaline condition. Sodium bicarbonate (1.5μl, 0.3M, pH9.0, NaHCO3) was added to the 3μl cDNA solution to give a total volume of 4.5μl and the sample vortexed followed by a brief centrifugation.

The CyDye pack was warmed to room temperature and after which 38μl dimethyl sulfoxide (DMSO) was added. This was done immediately before adding it to the cDNA. The CyDye was fully resuspended by vortexing combined with brief centrifugation to make sure no dye remained in the lid of the tube. CyDye (4.5μl) was added to the cDNA resuspension to give a total volume of 9μl. The mixture was briefly vortexed and centrifuged to gather all reactants to the bottom of the tube, this was followed by an overnight incubation in the dark.
4.2.3.4 CyDye labeled cDNA purification and quality assessment

The DNA-affinity of the column of Qia-Quick PCR Purification Kit requires high salt concentration and non-alkaline PH (PH <7.5). To each reaction, NaOAc (35μl of 100mM, pH5.2) was added followed by Buffer PB (250μl of 5X, supplied by the manufacturer). A QIAquick PCR spin column was placed in a collection tube (supplied by the manufacturer) and the sample added to the column and centrifuged at 13,000g for 1 minute. The collection tube was emptied and the column washed by the addition of Buffer PE (750μl, supplied by the manufacturer QIAGEN) which was subsequently centrifuged at 13,000g for 1 minute. The collection tube was emptied and the column centrifuged for an additional 1 minute at 16,000g. The column was transferred to a fresh microfuge tube, and then Buffer EB (30μl) was added to the top of the column followed by incubation at room temperature for 1 minute. After centrifugation at 13,000g for 1 minute, the eluent was retained.
After purification, fluorescence labeled cDNA was assessed for quantity, purity and incorporation efficiency, using stringent specifications to assess its suitability for hybridisation. Labeled cDNA quantification was performed on a spectrophotometer (Nano-drop). Spectrophotometer readings provided an indication of the amount of cDNA present (260nm) and the amount of Cy5 (650nm) present. The frequency of incorporation (FOI) was calculated which should ideally have been between 20 and 50. Further calculation was performed to determine the quantity of sample required to add to the hybridisation for 50pmols of label. Quality was assessed using minigel electrophoresis and the resulting gel was scanned on a GeneTAC™ LS IV scanner (Genomic Solutions, Huntingdon, Cambs., UK).

A “minigel” was prepared by pouring 1.5% (w/v) agarose into the mould and then once set turned out onto a microscope slide. The gel was placed in an electrophoresis tank and covered with 1x TAE (all tanks were rinsed thoroughly first on order to remove all traces of ethidium bromide). Labeled cDNA (1μl) was combined with glycerol (1μl, 50% (v/v)) and loaded into the well of the gel which was run for 30 minutes at 100V. The excess liquid was drained from the gel and scanned in a GeneTAC™ LS IV scanner (560 nm) to assess the size distribution of the labeled product. At this stage the label was assessed as good if it has a large size distribution through the gel and was not degraded and didn’t contain unincorporated product which appears as a smear at the base of the gel (Figure 4.2). Additionally the amount of protein contamination could also be assessed as this appears stuck in the well. If
the entire label was protein and not a cDNA distribution then the sample should be abandoned and the RNA further purified before attempting to label again.

**Fig. 4.2. Minigel electrophoresis of CyDye labeled cDNA**
A: an example of quality CyDye labeled cDNA which shows a large size distribution of fluorescence overlapped with the cDNA smear. B: Un-incorporated CyDye could be determined by the blob at the base of minigel.
4.2.3.5 Preparing probes for microarray hybridisation and array pre-wash

All samples were labeled with Cy5 and combined with Cy3 labeled oligo nucleotide mixture (pGEMT & pBluescript mix). The appropriate amount of label was added to a fresh microfuge tube and dried down in the SpeedVac at 65°C in the dark. Samples were dried down to a volume of 29.5μl or less and made up to 29.5μl with HPLC grade water if required. A blocker, poly-A (2μl of 100μM) and reference oligonucleotide mixture (3.5μl of 3.5pmols) were added to the probe to give a total volume of 35μl. The contents were mixed by gentle vortexing and pulsed to the bottom of the tube.

The cDNA microarrays were washed to remove unbound DNA molecules and buffer substances to avoid interference with subsequent hybridisation experiments. The slides were rinsed at room temperature, once in Triton®X 100 (5 minutes, 0.1% (v/v)), twice in HCl (2 x 2 minutes, 1mM), once in KCl (10 minutes, 10mM) and finally once in deionised water (1 minute). The recipe for pre-wash, blocking and post-wash buffer is listed in Table 4.1. The pre-wash procedure was followed immediately by blocking. Blocking buffer was prepared freshly each time (Table 4.1). The slides were incubated in Coplin jars (3 slides maximum per jar) or 200 ml glass troughs (6 slides maximum), firstly in Nexterion Block E (1x, 15 minutes, 50°C), rinsed in deionised water (1 minute) and then in a BSA based buffer (5x SSC, 0.1% (w/v) SDS, 1% (w/v) BSA) for 30 minutes at 42°C. The slides were finally washed by dipping in four changes of filtered sterile water (4 x 1 minute). If white streaks or
marks were seen on the slide then the water washes were repeated. The slides were quickly dried using compressed air and kept in the dark until required for hybridisation (less than 1 hour).

4.2.3.6 Microarray hybridisation

The hybridisation buffer (50% (v/v) formamide, 10x SSC, 0.2% (w/v) SDS) was prepared freshly each time and warmed at 37°C until the SDS was in solution. The labeled probes (29.5μl) were heat denatured in a heat block at 95°C for 3 minutes and centrifuged for 1 minute to remove condensation from the lid. Hybridisation buffer (35μl) was added to the labeled probe and mixed by pipetting up and down briefly. A LifterSlip (Erie Scientific 24x60I 2 4733, VWR, Poole, UK) was placed on soft tissue on the bench and the full 70μl of the sample pipetted in a line down the longest side. The blocked microarray was placed facing down (side with etching or label downwards) onto the slide so that the hybridisation mixture was slowly drawn across the slide. The slide sandwich was turned over so that the microarray slide was on the bottom and the LifterSlip on the top. They were placed in a humidity chamber at 38°C for 24 hours. There was always some water present in the bottom of the humidity chamber and the platform for the slides was completely dry before the slides were placed inside.
4.2.3.7 Microarray Post-wash

The slides were separated in Wash Buffer 1 (1xSSC, 0.2% (w/v) SDS, Table 4.1) at room temperature before being incubated in a fresh aliquot of Wash Buffer 1 in Coplin jars (3 slides maximum per jar) or 200ml glass troughs (6 slides maximum) at 55°C for 10 minutes. The slides were agitated 5 times and transferred to Wash Buffer 2 (0.1x SSC, 0.1% (w/v) SDS, Table 4.1) at 55°C for 10 minutes. After agitation, the slides were transferred to fresh at 55°C for 10 minutes; this was repeated once more to give a total of 3 washes in Wash Buffer 2. After further agitation the slides were transferred to Wash Buffer 3 (0.1x SSC, Table 4.1) 1 minute at room temperature and this step was then repeated. The slides were dried quickly using compressed air and stored in the dark until ready for scanning.
<table>
<thead>
<tr>
<th>Buffer</th>
<th>Recipe</th>
<th>Working temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wash Buffer A (0.1% Triton X-100)</td>
<td>0.1ml Triton X-100, make up to 100ml with filtered distilled autoclaved water.</td>
<td>Room temperature</td>
</tr>
<tr>
<td>Wash Buffer B (1mM HCl)</td>
<td>10μl HCl (32% w/w) in 101.8ml filtered distilled autoclaved water.</td>
<td>Room temperature</td>
</tr>
<tr>
<td>Wash Buffer C (10mM KCl)</td>
<td>0.75g KCl, make up to 100ml with filtered autoclaved water.</td>
<td>Room temperature</td>
</tr>
<tr>
<td>Blocking Buffer A (Schott Block E 1X)</td>
<td>Combine 25ml Schott Block E 4X with 75ml distilled filtered autoclaved water. Acidify with 24μl HCl (32% w/w).</td>
<td>50 °C</td>
</tr>
<tr>
<td>Blocking Buffer B (5X SSC, 0.1% SDS and 1% w/v BSA)</td>
<td>50ml 20X SSC, 1ml 20% SDS, 2g BSA Make up to 200ml with sterile filtered water.</td>
<td>42 °C</td>
</tr>
<tr>
<td>Hybridisation buffer (50% formamide, 10X SSC, 0.2% SDS)</td>
<td>500μl formamide, 500μl 20x SSC, 10μl 20% SDS</td>
<td>38 °C</td>
</tr>
<tr>
<td>Wash buffer 1 (1X SSC, 0.2% SDS)</td>
<td>50ml 20X SSC, 10ml 20% SDS, 940ml Sterile Filtered Water</td>
<td>55 °C</td>
</tr>
<tr>
<td>Wash buffer 2 (0.1X SSC,0.1% SDS)</td>
<td>5ml 20x SSC, 5ml 20% SDS, 990ml Sterile Filtered Water</td>
<td>55 °C</td>
</tr>
<tr>
<td>Wash buffer 3 (0.1X SSC)</td>
<td>5ml 20X SSC, 995ml Sterile Filtered Water</td>
<td>Room temperature</td>
</tr>
</tbody>
</table>

Table 4.1. The recipe for buffers used for pre-washing, blocking, and post-washing *G. pulex* cDNA microarrays.

### 4.2.3.8 Scanning and fluorescent quantification

Slides were scanned using a ScanArray™ Express HT scanner (PerkinElmer LAS (UK) Ltd., Bucks., UK ) using a fixed PMT and laser power at 66% for the Cy3 oligo nucleotide channel and 70% for the Cy5 sample channel. The resolution was set to 5μm and saved as a tiff file for subsequent image analysis. The pixels on the
tiff image were proportional to the fluorescent intensity of the microarray, hence the amount of cDNA hybridised to the target on the array. Image Gene 5 was employed to quantify the fluorescent intensity of each spot on the array, and the data were then exported into a spreadsheet for further analysis.

4.2.3.9 Microarray Data Analysis
The data files from Image Gene 5 were imported into GeneSpring (Version GX 7.3; Agilent Technologies UK Ltd., Cheshire, UK) and data was normalised per spot by dividing the test channel (Cy5) by the control channel (Cy3). The LOWESS method was applied to normalise the data with a raw data cutoff of 100. Slide data was compared using box plots and those with poor distribution removed; this was performed as a screening process for bad slides. Raw data (log ratio) was filtered between 100 and the maximum in the number of conditions that equaled the minimum sample size for each group within the experiment. Using the screened raw data, genes showing 2 fold changes (up- and down-regulated) by analysis of normalised data were isolated and probabilities assessed by statistics suited to the experiment. Principal component analysis (PCA) was performed on the statistically valid data. This allows visualisation of a data set within a coordinate system, with axis based on variance, selecting the characteristics that contribute the most variance. Dendrograms (hierarchical clustering) of statistically relevant genes with a 2 fold change in expression were created. Potential biomarkers for intersexuality or de-masculinisation were identified from this dataset.
4.2.4 Microarray validation by Quantitative-PCR

SYBR Green based QPCR was employed to validate the DNA microarray data. Five ESTs showing differential expression levels amongst three male phenotypes were chosen based on their biological functions (three genes related to reproductive process, one for immuno-activity and one random chosen gene without any blast hit in Flybase) according to their annotated GO terms. The sequences of the five ESTs (sequences of G. pulex) were blasted against the E. marinus cDNA libraries, which were generated by employing GS FLX sequencer (Chapter 5), to retrieve the orthologues of the five ESTs in E. marinus. Primers were then designed on each E. marinus cDNA sequence using a webpage based tool (DNA Tool), with the amplicons at approximately 200 bp and the Tm at around 60 °C. A total of 12 ng single-stranded cDNA was used to carry out the QPCR, following the protocol of Chapter 3 (Section 3.2.9).
4.3 Results

4.3.1 Total RNA and labeled probes purification

Approximately 5 μg total RNA was extracted from the pooled testes of five male *E. marinus* specimens. All of the RNA solutions had the 260/280 ratio between 1.80 and 2.00, indicating little protein contamination in these RNA samples (Fig. 4.3). Agarose electrophoresis was employed to examine the integrity of the purified total RNA (Fig. 4.4). Neither RNA degradation nor genomic DNA contamination was observed.

The purified CyDye labeled cDNA probes were assessed for quality by the minigel, and no evidence of un-incorporated CyDye or protein contamination was observed. The fluorescent cDNA probes were subsequently quantified by a spectrophotometre (Nana-drop), and the FOI (frequency of incorporation) values of all samples were within the range of 20 and 50.
**Fig. 4.3. The spectral pattern of a total RNA sample.**

The absorbance curve was generated by a spectrophotometer (Nano-drop), and the total RNA was extracted from pooled testes of five male *E. marinus*.

**Fig. 4.4. Agarose gel of four total RNA samples.**

RNA samples were purified from *E. marinus* testes, and electrophoresis was carried out on 2% agarose gel. Neither genomic DNA contamination nor evidence of RNA degradation was observed.
4.3.2 Microarray calibration and data analysis

Lucidea Universal Scorecard containing a set of artificial mRNA was applied as internal control to calibrate the DNA microarray data. The Lucidea Universal Scorecard was reverse transcribed and labeled simultaneously with the *E. marinus* RNA. These internal control probes include a series of ratio controls, and each of them has ten identical complementary targets printed on the microarray. A calibration curve can be generated based on the concentrations of each ratio control and its corresponding fluorescent intensity. Two examples of calibration curve were shown in Figure 4.5.

The DNA microarray data was normalised by LOWESS method with a raw data cutoff of 100. After filtering, a total number of 13383 (out of 13824) elements were revealed to have the substantial fluorescent intensity against background. Box plot depicting the fluorescent intensity distribution of each slide was generated for poor slides screening (Figure 4.6A). No bad slide with poor data distribution was identified. The spread of normalised data for each array was also elucidated by the box plot (Figure 4.6B). Reduced systematic bias was observed by comparing the box plots of raw and normalised data (Figure 4.6).
Fig. 4.5. Universal ScoreCard calibration of two microarrays utilised used to generate the *E. marinus* datasets

Ten replicates of each ratio control were printed on each slide; the mean signal intensity was used to generate calibration curves which are displayed with standard error bars.
Fig. 4.6. Box plots of the fluorescent intensities of the 15 microarrays utilised used to generate the *E. marinus* datasets.
A: Raw microarray data before normalisation; B: LOWESS normalised microarray data.
4.3.3 Genes showed differential expression amongst three male phenotypes

4.3.3.1 ANOVA, Volcano plot and PCA analysis

ANOVA (Analysis of Variance) analysis was conducted on the normalised microarray data, and a list of genes (349 ESTs) having significantly different (p>0.05) expression levels amongst three male groups were filtered. The global expression profiles of the 349 ESTs were visualised by generating the hierarchical clustering tree using a distance algorithm (Fig. 4.7). In order to identify the key variables of the gene expression profiles amongst three male phenotypes, the PCA analysis was carried out on the 349 ESTs filtered by ANOVA (Fig. 4.8). Fifteen samples were clearly differentiated into three isolated patches which were formed by five specimens of the same male phenotype. The two intersex male groups were closer to each other than to the normal male group (Fig 4.8).
Fig. 4.7. Global gene expression profiles for three male phenotypes in *E. marinus*.

Hierarchical tree of 349 ESTs with differential expressions amongst normal males (NMs), internal intersex males (IIMs) and external intersex males (EIMs).
Fig. 4.8. Principal component analysis of individual *E. marinus* transcriptomes. PCA analysis was based on 349 ESTs showing statistically relevant differential expression genes amongst normal males (blue), internal intersex males (yellow) and external intersex males (red).
Volcano plot was applied to identify genes with significantly different expression levels (p < 0.05) as well as the substantial fold change (above two-fold change) amongst various male groups. Pairwise comparisons were conducted on the three male phenotypes by the Volcano analysis, and three gene lists were then generated, respectively (Fig. 4.9). The numbers of ESTs filtered by Volcano plot for each pairwise comparison were listed in Table 4.2

<table>
<thead>
<tr>
<th>Pairwise Comparison</th>
<th>Number of ESTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM VS IIM</td>
<td>20</td>
</tr>
<tr>
<td>NM VS EIM</td>
<td>13</td>
</tr>
<tr>
<td>IIM VS EIM</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table 4.2 The number of ESTs filtered by Volcano plot.* Pairwise comparisons amongst the three male phenotypes by the Volcano analysis (p < 0.05; above 2 fold change). NM: normal males; IIM: internal intersex males; EIM: external intersex male.
Volcano plot was employed to filter genes had statistically significant (Y-axis, p < 0.05) difference and relative fold change (X-axis, above two fold) of their expression levels between groups. A: external intersex males VS normal males; B: internal intersex males VS external intersex males; C: internal intersex males VS normal males.
In order to generate gene lists with more ESTs, pairwise Volcano plot analysis was repeated on the three male phenotypes by using the parameter of \( p < 0.05 \) and at least one fold change. The numbers of ESTs filtered by Volcano plot for each pairwise comparison were listed in Table 4.3.

<table>
<thead>
<tr>
<th>Pairwise Comparison</th>
<th>Number of ESTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: NM VS IIM</td>
<td>453</td>
</tr>
<tr>
<td>Group 2: NM VS EIM</td>
<td>351</td>
</tr>
<tr>
<td>Group 3: IIM VS EIM</td>
<td>230</td>
</tr>
</tbody>
</table>

Table 4.3 The number of ESTs filtered by Volcano plot.
Pairwise comparisons amongst the three male phenotypes by the Volcano analysis (\( p < 0.05 \); above one fold change). NM: normal males; IIM: internal intersex males; EIM: external intersex male.

Venn diagram was generated to compare the three EST lists (Group 1: normal males VS internal intersex males; Group 2: normal males VS external intersex males; Group 3: internal intersex males VS external intersex males, Table 4.3) produced by Volcano plot analysis (Fig. 4.10). Only one EST (Gp_mxA_68D02) was found to be shared by all the three lists. There were 88 elements shared by Group 1 and Group 2, and 42 and 20 ESTs were shared by Group 1 and Group 3 as well as Group 2 and 3, respectively (Fig. 4.10).
Fig. 4.10. The Venn diagram on the gene lists generated by Volcano plot.

Three gene lists of *E. marinus* (Group 1, 2, and 3) generated by the Volcano plot (p < 0.05; more than one fold change) were compared by the Venn circles. Group 1: normal males VS internal intersex males; Group 2: normal males VS external intersex males; Group 3: internal intersex males VS external intersex males.
4.3.3.2 Annotation of the genes showing differential expression amongst the three male phenotypes

The 349 ESTs showing significantly (ANOVA, p < 0.05) different expression levels amongst the three male phenotypes were annotated by employing Blast2GO. The majority (275/349) of these genes were unknown and without any assigned GO terms (Fig. 4.11). Annotated genes with the similar function in the biological process were grouped into the same category, and pie charts were generated to depict the distribution of biological processes these genes involved (Fig. 4.12, 4.13 and 4.14). Figure 4.12 showed the genes function distribution of all the 13383 ESTs. EST groups produced by ANOVA analysis and Volcano plot were also annotated and the pie charts were presented in Figure 4.13. No significant bias was observed on the gene function distribution of 349 ESTs as well as the other three EST groups, compared to all the 13383 transcripts. Figure 4.14 interprets the distribution of 349 ESTs’ gene functions on biological process and its child category of metabolic process. The DAVID Bioinformatics Resource was used to identify the functions of the 349 ESTs, and those involved in the reproductive processes were illustrated in the Figure 4.15.
Fig. 4. The distribution of annotated sequences.
X-axis stands for the number of Gene Ontology (GO) terms assigned for each gene. Y-axis represents the gene numbers.
Fig 4.12. The pie chart of the distribution of gene functions of all the ESTs on the microarray used for characterising *E. marinus* transcriptome.
Fig 4.13. Pie charts of gene function distributions of four *E. marinus* EST groups.

1: The 349 ESTs from ANVOA analysis; 2: ESTs which were significantly differently expressed between internal and external intersex males; 3: ESTs from Volcano Analysis which had significant expression levels (with at least 2 fold change) between normal and internal intersex males; 4: ESTs from Volcano Analysis which had significant expression levels (with at least 2 fold change) between normal and external intersex males. There was no significant differences observed amongst the four pie charts.
Fig. 4.14. The distribution of biological processes of the genes differentially expressed amongst three male phenotypes in *E. marinus*.
Pie charts of 349 ESTs’ gene functions on biological processes (above) and its child category of metabolic process (below).
Fig. 4.15. Genes involved in the reproductive process.

Gene functions of 349 ESTs, showing significantly different expressions amongst three male groups in *E. marinus*, were identified by the DAVID Bioinformatics Resource.

Four ESTs (Figure 4.16, Table 4.4), showing significantly differential expressions amongst male groups based on microarray data, and biological importance according to the assigned GO terms were considered as the potential biomarkers for sexual phenotypes in *E. marinus*. The gene functions of their homologues in *Drosophila* were listed in Table 4.4. Orthologues in *E. marinus* were also obtained by blasting (Blastn) against the *E. marinus* 454 sequencing dataset (Chapter 5). QPCR validation was conducted on five ESTs (the four ESTs with annotated biological functions as well as a random EST from Volcano Plot), and the results were included in the Appendix.
<table>
<thead>
<tr>
<th>EIMs</th>
<th>IIMs</th>
<th>NMs</th>
<th>MV1: Gp_mxAA_69A03</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Green]</td>
<td>![Green]</td>
<td>![Green]</td>
<td>MV2: Gp_mxAA_16C10</td>
</tr>
<tr>
<td>![Green]</td>
<td>![Green]</td>
<td>![Green]</td>
<td>MV3: Gp_mxAA_40B09</td>
</tr>
<tr>
<td>![Red]</td>
<td>![Red]</td>
<td>![Red]</td>
<td>MV5: Gp_maSA_28F02</td>
</tr>
</tbody>
</table>

**Fig. 4.16.** Genes showing relative differential expression levels amongst three male phenotypes in *E. marinus.*

Genes with relatively high expression levels were marked by red, and low by green. The ESTs which were considered as potential biomarkers for de-masculinisation in crustaceans were named from MV1 to MV5, followed by their EST ID.
<table>
<thead>
<tr>
<th><strong>G. pulex</strong> ESTs</th>
<th><strong>Homologue in Flybase</strong></th>
<th>**Homologue in <em>E. marinus</em>[^<em>]</em></th>
<th><strong>Identity</strong></th>
<th><strong>Identity</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MV1</strong></td>
<td>FBgn0001233: Molecular chaperone, developmentally expressed during oogenesis.</td>
<td></td>
<td>77%</td>
<td>Contig00683 79%</td>
</tr>
<tr>
<td><strong>MV2</strong></td>
<td>FBgn0040321: Involved in the recognition of invading micro-organisms.</td>
<td></td>
<td>50%</td>
<td>Contig13882 82%</td>
</tr>
<tr>
<td><strong>MV3</strong></td>
<td>FBgn0011217: Involved in postmeiotic stages of spermatogenesis.</td>
<td></td>
<td>55%</td>
<td>Contig11575 92%</td>
</tr>
<tr>
<td><strong>MV4</strong></td>
<td>FBgn0013770: Important for the overall degradation of proteins in lysosomes. Essential for adult male and female fertility.</td>
<td></td>
<td>58%</td>
<td>Contig01470 87%</td>
</tr>
<tr>
<td><strong>MV5</strong></td>
<td>NA</td>
<td>NA</td>
<td>Contig01343 92%</td>
<td></td>
</tr>
</tbody>
</table>

[^*]: *E. marinus* sequences from 454 GS FLX sequencing (Chapter 5).

Table 4.4. Homologues of five *G. pulex* ESTs in Drosophila and *E. marinus*, respectively.

The BLAST hits of five *G. pulex* ESTs in Flybase, and homologue identities as well as their gene functions were listed in the table. Five *G. pulex* ESTs were also blasted against *E. marinus* sequences, and homologues contigs and corresponding identity were elucidated. *E. marinus* sequences from 454 GS FLX sequencing (Chapter 5).
4.4 Discussion

This study applied microarrays to compare the gene expression patterns between normal and intersex male amphipods, with aims to determine the mechanism of intersexuality in *E. marinus*, differentiate internal and external intersex males on the genetic level, and develop molecular biomarkers for de-masculinisation in crustaceans. Intersexuality has been reported in a variety of crustaceans, and there have been studies investigating the morphology, behaviour as well as the cost of intersexuality in crustaceans (Sagi et al, 1996; Ford et al, 2004b; McCurdy et al, 2004), however investigations on the genetic level are sparse (Ford et al, 2008). The genetic knowledge on *E. marinus* was extremely limited. This study applied a *G. pulex* microarray to characterise the transcriptome profiles in *E. marinus* by heterologous hybridisation (Renn et al, 2004; Ford et al, 2008).

Cross-species microarray has been successfully applied on various organisms groups, for example between salmon fish species from different orders (von Schalburg et al, 2005). *G. pulex* and *E. marinus* are not as phylogenetically distant as organisms investigated by previous studies (Buckley 2007). It is reasonable to expect the heterologous hybridisation between the two amphipod species to produce viable data (Sambles, 2008). Apart from the phylogenetic relationship between species, the type of the microarray platform also affects the efficiency of cross-species hybridisation (Buckley 2007). Microarrays apply short length targets are more
sensitive than those with long sequence targets (Lockhart et al., 1996). The targets on the microarray employed by this study were generated from the cDNA libraries of *G. pulex*, and the long-length sequences kept the destabilization caused by sequence pair mis-match into minimum. The *G. pulex* microarray applied by this work contains approximately 13,000 elements (Sambles, 2008). In this study, more than 95% of the elements gave out substantial fluorescent signals against background, suggesting vital cross-species hybridisation. Calibration curves generated by the Lucidea Universal ScoreCard have also confirmed the linear correlation between probe concentrations and fluorescent intensities.

Variations are inevitably introduced during multiple steps of microarray experiment, such as sampling, fluorescence incorporation and “probe-target” hybridisation, therefore the raw microarray data need to be normalised before any further analysis being carried out (Bammler et al., 2005; Irizarry et al., 2005). However, no universal normalisation method could be systematically applied, and it is not always obvious which method is the most suitable for a certain microarray experiment (Kreil and Russell, 2005; Lemoine et al., 2006). In this study, the microarray data was normalised by the LOWESS method (Limma statistics package), which is a widely adapted approach to control the global fluorescent intensity amongst microarrays. The ANOVA analysis identified 349 ESTs having significantly different expression levels amongst the three male phenotypes in *E. marinus*. The PCA analysis was then employed to clearly differentiate the three male groups into three distinct patches,
based on the expression levels of the 349 ESTs.

Previously, it was not clear whether internal and external intersex male were two distinct intersex phenotypes, or they belong to the same phenotype but at different stages. The PCA demonstrated that there was greater variation between normal and each intersex male group than between the two intersex male groups. Also the two intersex male groups were observed to be equally away from the normal males, suggesting the two male intersexes are two distinct phenotypes. A strong relationship between microsporidian infection and external intersex males has been revealed in Inverkeithing *E. marinus* population, whilst less than 25% internal intersex males were found to be infected (see Chapter 3). Feminising parasite has been presumed as a potential factor to induce intersexuality in crustacean (Rigaud and Juchault, 1998; Kelly et al, 2002, 2004), and the lack of microsporidian infection in the majority internal intersex males suggesting there were other mechanisms causing the intersex phenotype in *E. marinus*, and internal and external intersex males were possibly caused by different factors. Both the transcriptomic pattern and incidence of microsporidian infection suggest that internal and external intersex males were distinct phenotypes. Further investigation into the causes of intersexuality in these two different intersex phenotypes may reveal interesting insights into sexual differentiation in these crustaceans.

This study employed a cross-species cDNA microarray to characterise the
transcriptome profiles in *E. marinus*, and the global gene expression patterns of normal males and two male intersex phenotypes was differentiated by the PCA analysis. However, the QPCR validation suggested the fluctuation of the microarray data when focused on individual genes. Two ESTs (MV3 and MV4) filtered by Volcano plot had the same expression pattern in the three male groups identified by both microarray and QPCR, whilst QPCR results for ESTs (MV1 and MV2, from the ANOVA gene list) were not consistent with their microarray data (Appendix). The divergence on the cDNA sequences of the two species can reduce the sensitivity of microarray. Without knowing the sequences of the transcriptome of *E. marinus*, it is impossible to estimate the variation caused by sequence mis-match. In order to clearly review the transcriptomic difference of normal and intersex specimens in *E. marinus*, we need new techniques to sequence the *E. marinus* transcriptome and carry out further studies based on its own DNA sequence information.
Chapter 5

Transcriptome Sequencing in Normal and Intersex *E. marinus* by GS FLX (454)
5.1 Introduction

Intersexuality in *E. marinus* was investigated on both population and individual levels in previous chapters. Environmental pollutions as well as microsporidian parasites have been found to consistently associate with intersexes. A preliminary study employing a cross-species microarray was conducted with the aim to understand the mechanism of intersexuality in amphipods on the genetic level. However, in order to clearly identify the genetic difference between normal and intersex specimens, the DNA sequence information of *E. marinus* is needed. The conventional method for sequencing the genomic or complementary DNA of a non-model species involves recombining fragments of the target DNA within a bacterial vector, isolation of DNA from the individual recombinants followed Sanger sequencing process which is manually intensive, time-consuming and expensive. These limitations have therefore restricted the spectrum of organisms from which substantive bodies of genetic information have been generated to a handful model species (Ellegren, 2008). The advent of massive parallel sequencing, otherwise termed next generation sequencing (NGS), (Chapter 1, Section 1.5.2), have challenged this dogma and facilitating the global genetic analysis of a much wider spectrum of organism by providing the ultra-high throughput sequencing capability at a low cost and without the requirement for manual handling of thousands if not millions of individual target DNAs (Margulies et al, 2005; Ellegren, 2008).
GS FLX (454) sequencing is based on the concept of pyrosequencing (Ronaghi et al, 1998). At the time of writing this thesis, the newest sequencing platform, GS FLX Titanium, is able to sequence 400-600 million bases per run (approximately 10 hours), with the average sequence length of 400 bp and the sequencing accuracy above 99%. By employing sample-specific multiplex identifiers (MIDs), which are short-nucleotide sequences attached to the end of DNA strand for identification, GS FLX sequencing is able to sequence up to 2304 samples in a single run (Droegea and Hill, 2008). The mechanism of 454 sequencing is illustrated in the general introduction of this thesis (Chapter 1, Section 1.5.2.2). The 454 sequencing technology makes it possible to retrieve the nucleic acid sequence information from a non-model species organism at a reasonable cost, therefore largely facilitates the biological studies which used to be constrained by the genetics knowledge. This powerful high-throughput sequencing technique has been employed by studies from various fields to gather DNA sequences information for a variety of purposes, such as whole genome and transcriptome sequencing (Rothberg & Leamon, 2008; Vera et al, 2008; Wheeler et al, 2008). When this thesis was written up (September 2011), there have been 1070 peer-reviewed academic papers published by applying 454 sequencing techniques.

The aim of the current chapter was to indentify genetic difference between normal and intersex _E. marinus_ by sequencing and comparing the transcriptomes in gonads of _E. marinus_ employing the GS FLX sequencer. Sex-specific genes, as well as the
genes related to intersexuality could be identified and utilised to develop molecular biomarkers for endocrine disruption in crustaceans. Gonad transcriptomes from four sexual phenotypes of *E. marinus* were sequenced and assembled, and four libraries were developed based on the corresponding reads from the sequencing run. The *E. marinus* cDNA datasets generated by *de novo* transcriptome sequencing will represent a substantial supplement to the existing genetic information of crustaceans.

The transcript information provides descriptive information as to the transcript content, i.e. the sequence of the *E. marinus* transcripts, but also allows estimates of relative transcript abundance to be calculated by counting the relative number of times a specific gene object is observed (Trapnell et al, 2010). Identification of differentially expressed genes amongst different sexual phenotypes provides valuable information to investigate the mechanism of intersexuality in crustaceans. They can also be used to develop potential molecular biomarkers for differentiating sexual phenotypes in amphipods.
5.2 Materials and Methods

5.2.1 Sampling

In August 2008, *Echinogammarus marinus* specimens were collected underneath seaweed and stones in an industrially impacted site, Inverkeithing, Scotland. Animals were brought back to the laboratory and kept in filtered seawater from Sully, Cardiff, for recovery. Amphipods were anaesthetised by carbonated seawater, and phenotypically examined by visual inspection with the aid of a stereo microscope. According to their external and internal secondary sex characters, amphipod specimens were classified into four sex phenotypes, namely normal male, normal female, intersex male and intersex female (Chapter 2, Section 2.1). Gonadal tissue was dissected out from each specimen by using the micro-scissors under a stereo microscope, and total RNA was immediately extracted from the fresh tissues.

5.2.2 RNA purification

Total RNA of the gonads of each amphipod was purified by employing an improved RNA extraction method combining Tri-reagent and RNeasy Plus Micro Kit (Chapter 3, Section 3.2.3.2).
5.2.3 Microsporidian screening and RNA samples preparation for GS FLX sequencing

Individual amphipod specimens were examined for microsporidian infection before material was pooled in preparation for GS FLX sequencing. An aliquot of 1μl RNA solution containing approximately 50 ng total RNA was used to synthesize single-stranded cDNA following the reverse transcription protocol described in Chapter 3 (Section 3.2.4). Complementary DNA was then utilised as the template for microsporidian parasite screening by a PCR based method (Chapter 3, Section 3.2.6). For normal males and normal females, un-infected samples were used for the GS FLX transcriptome sequencing, whilst infected samples were used to represent the two intersex phenotypes. RNA samples of nine individuals from the same sexual phenotypes with the same infection status were pooled together, and subsequently concentrated by using a Speed Vac. Concentrated RNA samples were quantified by a spectrophotometer, and the integrity of the RNA samples was examined by a Bioanalyzer. The four pooled RNA samples were amplified (see below) to obtain an adequate amount of cDNA as a template for 454 FLX sequencing.

5.2.4 RNA amplification

5.2.4.1 First-strand cDNA synthesis

The MessageAmp™ aRNA Amplification Kit (Ambion) was employed to amplify the mRNA of the pooled total RNA samples. Total RNA, approximately 4 μg for each sample, was used for aRNA amplification. The RNA solution was mixed with
1μl T7 Oligo dT primer, and the final volume was brought up to 12μl by adding the nuclease free water. The RNA-primer mix was then incubated at 70 °C for 10 minutes, and immediately transferred on ice afterwards. An aliquot of 8μl master mix, containing 2μl 10X First strand buffer, 1μl Ribonuclease inhibitor, 4μl dNTP mix and 1μl Reverse transcriptase, was added to the RNA-primer solution, and then well mixed by gently pipetting for several times. The reverse transcription reaction (20μl) was incubated at 42 º C for 2 hours, and the tube was then placed on ice for the subsequent second-strand cDNA synthesis.

5.2.4.2 Second-strand cDNA synthesis

Fresh aliquot of second-strand cDNA synthesis master mix was made out of 63μl nuclease-free water, 10μl 10X Second Strand Buffer, 4μl dNTP mix, 2μl DNA Polymerase and 1μl RNase H. The second-strand cDNA synthesis mix, 80μl was added to the 20μl single-stranded cDNA synthesis reaction, and well mixed by gently pipetting for several times. The resultant 100μl second-strand cDNA synthesis reaction was incubated at 16°C for 2 hours.

5.2.4.3 cDNA purification

The nuclease free water was incubated at 50 °C for later use. An aliquot of 250μl cDNA binding buffer was added to the 100μl second-strand cDNA synthesis reaction. The mixture was then transferred into a cDNA filter cartridge, and was centrifuged at 10,000g for 1 minute. The flow-through was discarded, and the cDNA filter cartridge
was washed by adding 500μl cDNA wash buffer, and centrifuged at 10,000g for 1 minute. The cDNA filter cartridge was transferred onto a fresh autoclaved eppendorf tube, and was eluted by 20μl 50 °C pre-warmed nuclease free water. After being centrifuged at 10,000g for 1 minute, the flow-through purified cDNA solution was kept on ice for the following amplification step.

5.2.4.4 \textit{In Vitro} RNA amplification

\textit{In Vitro} Transcription Mix was made out of 4μl T7 ATP Solution (75 mM), 4μl T7 CTP Solution (75 mM), 4μl T7 GTP Solution (75 mM), 4μl T7 UTP Solution (75 mM), 4μl T7 10x Reaction Buffer and 4μl T7 Enzyme Mix. An aliquot of 16μl double-stranded cDNA solution (Section 5.2.4.3) was mixed with 24μl of \textit{In Vitro} Transcription Mix. The reaction was incubated at 37°C for 6 hours.

5.2.4.5 Purification of aRNA

The nuclease free water was incubated at 50 °C for later use. After 6 hours incubation, the volume of the RNA amplification reaction was brought up to 100μl by adding 60μl nuclease free water. An aliquot of 350μl aRNA binding buffer was added to each RNA amplification reaction. After being well mixed by pipetting for several times, 250μl ACS grade 100% EtOH was added to the mixture, and the solution was then transferred into an aRNA purification cartridge. The cartridge was centrifuged at 10,000 g for 1 minute, and the flow-through was discarded. A total of 650μl aRNA Wash Buffer was used to wash the cartridge. After centrifuging the
cartridge and discarding the flow-through, the column was subsequently transferred onto a fresh Eppendorf tube. To elute the aRNA, an aliquot of 100μl pre-warmed nuclease free water was added to the centre of the cartridge. After being incubated at room temperature for 2 minutes, the cartridge was centrifuged at 8000prm for 1 minute, and the flow-through aRNA solution was then qualified by a spectrophotometer and examined for integrity by a Bioanalyzer.

5.2.5 Double-stranded cDNA synthesis

5.2.5.1 First strand cDNA synthesis

Approximately 10μg aRNA of each sample was used to synthesis the first strand cDNA, using recombinant M-MLV reverse transcriptase (Promega). Random hexamers (final concentration 10 mM) were added to the aRNA solution, and the total volume was brought up to 12μl by adding ddH2O. The aRNA-primer mixture was incubated at 70 º C for 10 minutes and then immediately transferred on ice. A master mix containing 5μl buffer (5X), 1.25μl dNTP (10 mM), 1.25μl MMLV, and 5.5μl ddH2O was added to the RNA-primers mix, and gently pipetted several times to achieve homology. The reaction was then incubated at 42 º C for 2 hours.

5.2.5.2 Second-strand cDNA synthesis:

To synthesis the second strand of the cDNA, 20μl RT reaction (5.2.3) was added to the second-strand cDNA synthesis master mix, which included 4μl dNTP (10 mM), 10μl 10X buffer, 3μl (30 U) DNA Polymerase I, 0.5μl (2.5 U) RNase H and 62.5μl
nuclease free water. The reaction was incubated at 16 °C for 2.5 hours. Double-stranded DNA was subsequently purified by Qiagen PCR purification kit following manufacturer’s protocol. Double-stranded DNA solution was quantified by PicoGreen assay.

5.2.5.3 Double-stranded cDNA quantification by PicoGreen assay

PicoGreen (Invitrogen) exclusively binds to the double-stranded nucleic acid, and the concentration of double stranded DNA or DNA:RNA hybrid is proportional to the fluorescent intensity of PicoGreen. A series of DNA marker dilutions (λ HindIII DNA marker, Promega) was prepared for generating a standard curve. By quantifying the fluorescent intensity and fitting into the standard curve, the concentration of an unknown DNA sample could then be estimated. The preparation of standard curve, blanks and unknown DNA samples were elucidated in Table 5.1.

The standards (50µl) and samples (50µl) were transferred to a black OptiPlate 96 (PerkinElmer, UK). PicoGreen® was diluted 1:400 with 1x TE in an amber tube. Diluted PicoGreen® (50µl) was added to each well to make a total volume of 100µl per well. The plate was shaken in the DPC MicroMix® 5 plate shaker (Diagnostic Products Corporation, USA) for 1 minute (Form 50, amplitude 9) and then pulsed in a centrifuge to bring the contents to the bottom of the wells. The plate was read using a Flourocounter (FluoroCount™, PerkinElmer, UK) at the following wavelengths; excitation, 485nm and emission, 530nm. The standard curve was used to determine
the original concentration of the cDNA.

<table>
<thead>
<tr>
<th></th>
<th>DNA Marker (μl)</th>
<th>1 x TE (μl)</th>
<th>*PicoGreen (μl)</th>
<th>DNA Sample (μl)</th>
<th>Final DNA Concentration (ng/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard Curve</strong></td>
<td>50</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>0</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>12.5</td>
<td>37.5</td>
<td>50</td>
<td>0</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>6.25</td>
<td>43.75</td>
<td>50</td>
<td>0</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>3.13</td>
<td>46.87</td>
<td>50</td>
<td>0</td>
<td>62.5</td>
</tr>
<tr>
<td></td>
<td>1.56</td>
<td>48.43</td>
<td>50</td>
<td>0</td>
<td>31.25</td>
</tr>
<tr>
<td></td>
<td>0.78</td>
<td>49.22</td>
<td>50</td>
<td>0</td>
<td>15.62</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Blanks</strong></td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Sample</strong></td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>1</td>
<td>TBD</td>
</tr>
</tbody>
</table>

Table 5.1. The preparation of standard curve, blanks and unknown DNA samples for the PicoGreen assay.

The concentration of DNA marker is 2μg/ml. TBD = To be determined. * PicoGreen was diluted 1:400 with 1x TE before being added to DNA solutions.
5.2.6 GS FLX sequencing

For each of the four *E. marinus* sexual phenotypes, approximately 4μg double-stranded cDNA was sent to the Centre for Genomic Research of The University of Liverpool, for the GS FLX sequencing. The cDNA samples were fragmented into 300-800 bp length sequences, and short adaptors were attached to the fragments for the subsequent purification, amplification and sequencing procedures. The single-stranded DNA sequences were then individually immobilised into DNA Capture Beads, and emulsion PCR was carried out for each bead in the water-in-oil mixture. The amplified fragments were then transferred into a PicoTiterPlate for sequencing.

5.2.7 Sequencing data assembly and annotation

GS FLX sequencing reads from all libraries were pooled, trimmed to remove T7 adapter sites and poor quality bases then assembled into contigs by using the software 454/Roche GS de novo Assembler (Newbler, Version 2.3) under default parameters (Chaisson and Pevzner, 2008). Sex-specific and intersex-specific groups of contigs were generated based on their reads distribution in different phenotypes of *E. marinus*. All contigs as well as sex-specific and intersex-specific groups were annotated by applying Blast2Go (Conesa et al, 2005; Götz et al, 2008).

5.2.8 GS FLX sequencing results validation

The relative abundance of transcripts was determined by deriving the contribution of
library specific reads to each assembled contig. This was performed by interrogating
the assembly using the common assembly file (.ace file – Kumar and Blaxter, 2010)
and cross referencing each contributing read with its originating library. This was
used to determine list of clusters which had great variations on the expression levels
amongst four sexual phenotypes, normal males, normal females, intersex males and
intersex females. Amongst these clusters, five genes were chosen to validate the 454
sequencing results by conventional PCR. Primers were designed on the
corresponding sequences using a webpage based tool (DNA tool), with the Tm
temperature at around 60 º C and the amplicon of 200 bp. The leftover cDNA samples
which were sequenced by GS FLX sequencer were applied as the templates for the
PCR. An aliquot of 8 ng cDNA was applied as the template, and the reactions were
carried out using various cycle numbers to optimise the band intensities. The PCR
products were loaded on 1.5% agarose gel for electrophoresis. The gene’s expression
levels in different sexual phenotypes were visualised and compared to their read
numbers from 454 sequencing.
5.3 Results

5.3.1 Qualitative and quantitative assessment of total RNA, aRNA and cDNA

Total RNA was individually extracted from gonads of *E. marinus* specimens representing each sexual phenotype. Individual RNAs were pre-screened for microsporidian infection (see Chapter 5, Section 5.2.3) this allowed us to isolated uninfected normal males and normal females for analysis whilst infected animals were used to represent the two intersex phenotypes, intersex males and intersex females. All the four sexual phenotypes pool of RNA was generated representing nine individuals and then concentrated by using a Speed Vac. Both before and after concentrating, the samples were spectrophotometrically analysed to confirm observed 260/280 ratios were between 1.8 and 2.0 (Figure 5.1).

![Fig. 5.1. The spectral pattern of total RNA samples.](image)
The absorbance curve was generated by a spectrophotometer (Nano-drop). Left: The pooled total RNA samples purified from gonads of nine *E. marinus*. Right: the same RNA sample concentrated by a Speed vac.
In order to generate sufficient cDNA for sequencing, it was required to amplify the mRNA before reverse transcription. This was achieved by generating double-stranded cDNA incorporating a T7 RNA polymerase site 3’ of the poly-A tail and then exploiting T7 RNA polymerase to generate multiple copies of each transcript (aRNA) (see Chapter 5, Section 5.2.4). The yield of aRNA and amplification efficiency was calculated for each sexual phenotype (Table 5.2). Total RNA and aRNA samples were examined for integrity by using a Bioanalyzer (Fig. 5.2). No evidence of degradation was observed in total RNA samples, and the apparent smears including a multitude of discrete bands, as observed for all the four aRNA specimens, indicated wide size distribution of aRNA sequences.

<table>
<thead>
<tr>
<th>Sexual Phenotype</th>
<th>Input total RNA (μg)</th>
<th>Output aRNA (μg)</th>
<th>*Amplification Efficiency (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal male</td>
<td>5</td>
<td>40</td>
<td>400</td>
</tr>
<tr>
<td>Normal female</td>
<td>5</td>
<td>200</td>
<td>2000</td>
</tr>
<tr>
<td>Intersex male</td>
<td>5</td>
<td>80</td>
<td>800</td>
</tr>
<tr>
<td>Intersex female</td>
<td>5</td>
<td>200</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Table 5.2. The yield of aRNA samples amplified from *E. marinus* gonadal total RNA**

The aRNA was generated by using MessageAmp™ aRNA Amplification Kit (Ambion). * The amplification efficiency was calculated by presuming mRNA made up 2% of total RNA in gonads of *E. marinus*. 
Fig. 5.2. Integrity analysis of Total and aRNA

The sequence distribution of four amplified aRNA (lane 1-4) and four total RNA (lane 5-8) samples generated from *E. marinus* gonads. Analysis was performed using a bioanalyzer (Agilent 2100 Bioanalyzer) by applying Agilent RNA 6000 Nano Kit.

Picogreen assay was applied to quantify the double stranded cDNA samples for each sexual phenotype. A standard curve was generated by a series dilutions of DNA marker (Fig. 5.3), and the concentration of each double-stranded cDNA sample was determined (Table 5.3).
Fig. 5.3. The standard curve for the Picogreen assay.
Standard curve was generated by a series of DNA marker dilutions their corresponding Picogreen absorbance. X-axis: concentration of double stranded DNA; Y-axis: Picogreen fluorescent intensity

<table>
<thead>
<tr>
<th>Samples</th>
<th>Input aRNA (μg)</th>
<th>Concentration of Double-stranded cDNA samples (ng/μl)</th>
<th>Yield of Double-stranded cDNA (μg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal male</td>
<td>~20</td>
<td>81</td>
<td>4.9</td>
</tr>
<tr>
<td>Normal female</td>
<td>~20</td>
<td>82</td>
<td>4.9</td>
</tr>
<tr>
<td>Intersex male</td>
<td>~20</td>
<td>70.8</td>
<td>4.2</td>
</tr>
<tr>
<td>Intersex female</td>
<td>~20</td>
<td>88.4</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Table 5.3. The concentration and yield of double-stranded cDNA generated from 20 μg aRNA for each sample.
5.3.2 454 sequencing data of *E. marinus*

5.3.2.1 GS FLX sequence reads

The four un-normalised cDNA libraries representing gonadal transcriptome profiles of four sexual phenotypes in *E. marinus* were sequenced by 454 GS FLX platform. A total number of 213,212 ESTs were generated from the four libraries, with the average read length at 200 bp. The number of sequence reads and the mean read length for each library are shown in Table 5.4.

<table>
<thead>
<tr>
<th>Library</th>
<th>Sequenced reads</th>
<th>Mean EST length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal male</td>
<td>62,288</td>
<td>200</td>
</tr>
<tr>
<td>Normal female</td>
<td>37,459</td>
<td>201</td>
</tr>
<tr>
<td>Intersex male</td>
<td>60,395</td>
<td>200</td>
</tr>
<tr>
<td>Intersex female</td>
<td>53,070</td>
<td>202</td>
</tr>
<tr>
<td>Total</td>
<td>213,212</td>
<td>200</td>
</tr>
</tbody>
</table>

*Table 5.4. The number of GS FLX sequencing reads and mean EST length for cDNA libraries of four *E. marinus* sexual phenotypes.*
5.3.2.2 GS FLX sequences assembly

ESTs from the four libraries were assembled by using the software Newbler, and a total of 12,645 contigs were generated from the 213,212 ESTs. On average, each contig was represented by 9 raw sequence reads. In addition, there were 96,878 orphan quality reads, these were considered as singletons during subsequent analysis.

The distribution of the number of ESTs per contig is showed in Table 5.5. The average length of all contigs is 283 bp, with the minimum and maximum length at 100 bp and 2874 bp, respectively. Figure 5.4 describes the distribution of contigs with different sequence lengths.

<table>
<thead>
<tr>
<th>From (Number of ESTs)</th>
<th>To (Number of ESTs)</th>
<th>Number of Contigs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>12548</td>
</tr>
<tr>
<td>201</td>
<td>400</td>
<td>34</td>
</tr>
<tr>
<td>401</td>
<td>600</td>
<td>10</td>
</tr>
<tr>
<td>601</td>
<td>800</td>
<td>4</td>
</tr>
<tr>
<td>801</td>
<td>1000</td>
<td>0</td>
</tr>
<tr>
<td>1001</td>
<td>1200</td>
<td>1</td>
</tr>
<tr>
<td>1201</td>
<td>1400</td>
<td>0</td>
</tr>
<tr>
<td>1401</td>
<td>1600</td>
<td>0</td>
</tr>
<tr>
<td>1601</td>
<td>1800</td>
<td>2</td>
</tr>
<tr>
<td>1801</td>
<td>1999</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5.5. The distribution of *E. marinus* contigs with different numbers of ESTs

*E. marinus* contigs were generated by 454 sequencing and de novo assembly.
Fig. 5.4. The distribution of *E. marinus* contig lengths.

*E. marinus* contigs were generated by 454 sequencing and *de novo* assembly.

For each library, the numbers of sequence reads and assembled contigs are listed in Table 5.6. In addition, library-specific contigs were filtered based on their presence in each phenotypic group. Normal-male-specific genes were exclusively observed in the normal male libraries, and not in any of the other three libraries. The same standard was applied for generating normal-female-specific, intersex-male-specific and intersex-female-specific genes. There are 934, 1051, 1166 and 1534 contigs which are found to be specific to normal males, normal females, intersex males and intersex females, respectively (Table 5.6).
<table>
<thead>
<tr>
<th>Library</th>
<th>Reads</th>
<th>Contigs</th>
<th>Library-specific Contigs</th>
<th>Novelty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal male</td>
<td>62,288</td>
<td>5,059</td>
<td>934</td>
<td>18</td>
</tr>
<tr>
<td>Normal female</td>
<td>37,459</td>
<td>6,047</td>
<td>1051</td>
<td>17</td>
</tr>
<tr>
<td>Intersex male</td>
<td>60,395</td>
<td>6,295</td>
<td>1166</td>
<td>19</td>
</tr>
<tr>
<td>Intersex female</td>
<td>53,070</td>
<td>7,252</td>
<td>1534</td>
<td>21</td>
</tr>
</tbody>
</table>

**Table 5.6. The number of reads and contigs in each library.**

Reads of *E. marinus* were generated by GS FLX sequencing, and contigs were produced by *de novo* assembly.

According to the presence of contigs in each of the four libraries, lists of phenotype-specific genes were identified (Table 5.7). Male-specific genes are found in both of the two male libraries, but not in either normal or intersex females. Female-specific and intersex-specific genes were identified following the same logic as determining male-specific genes. A Venn Diagram (Figure 5.5) was made to elucidate the various genes lists, and the gene numbers of each list are shown in Table 5.7.
Fig. 5.5. The Venn Diagram interpreting phenotype-specific genes of *E. marinus*. *E. marinus* contigs were generated by 454 sequencing and their presence in each library were used to identify phenotype-specific genes. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females.

<table>
<thead>
<tr>
<th>Gene List</th>
<th>Description</th>
<th>Contig Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Normal-male-specific genes</td>
<td>934</td>
</tr>
<tr>
<td>2</td>
<td>Normal-female-specific genes</td>
<td>1051</td>
</tr>
<tr>
<td>3</td>
<td>Intersex-male-specific genes</td>
<td>1166</td>
</tr>
<tr>
<td>4</td>
<td>Intersex-female-specific genes</td>
<td>1534</td>
</tr>
<tr>
<td>5</td>
<td>Male-specific genes</td>
<td>1206</td>
</tr>
<tr>
<td>6</td>
<td>Female-specific genes</td>
<td>1745</td>
</tr>
<tr>
<td>7</td>
<td>Intersex-specific genes</td>
<td>782</td>
</tr>
<tr>
<td>8</td>
<td>Present in all the four libraries</td>
<td>1078</td>
</tr>
</tbody>
</table>

Table 5.7. The number of contigs in each phenotype-specific gene list of *E. marinus*. 
By presuming that the read counts are able to reflect the expression level of the corresponding contigs, genes showed the trend of masculinisation (with the expression levels following the pattern of normal males > intersex males > intersex females > normal females) and de-masculinisation (normal females > intersex females > intersex males > normal males) were classified into two gene lists (Table 5.8). For all the ten gene lists generated by the above analysis, the cut-off values of 10 reads/contig and 30 reads/contigs were used in order to filter out those genes with low expression levels (Table 5.8).

<table>
<thead>
<tr>
<th>Gene List</th>
<th>&gt;0 reads/contig</th>
<th>&gt;10 reads/contig</th>
<th>&gt;30 reads/contig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal-male-specific genes</td>
<td>934</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Normal-female-specific genes</td>
<td>1051</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Intersex-male-specific genes</td>
<td>1166</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Intersex-female-specific genes</td>
<td>1534</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Male-specific genes</td>
<td>1206</td>
<td>160</td>
<td>45</td>
</tr>
<tr>
<td>Female-specific genes</td>
<td>1745</td>
<td>185</td>
<td>34</td>
</tr>
<tr>
<td>Intersex-specific genes</td>
<td>782</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td>Genes present in all libraries</td>
<td>1078</td>
<td>754</td>
<td>305</td>
</tr>
<tr>
<td>Genes showed the trend of masculinisation</td>
<td>194</td>
<td>170</td>
<td>103</td>
</tr>
<tr>
<td>Genes showed the trend of de-masculinisation</td>
<td>132</td>
<td>103</td>
<td>39</td>
</tr>
</tbody>
</table>

Table 5.8. The number of contigs in each phenotype-specific gene list of *E. marinus*.

According to the read number of each contig, eight gene lists are classified into three categories, with the total read number of each contig above 0, 10 and 30, respectively.
A matrix including 12,645 rows and 4 columns was generated, and the row and column represents contigs and libraries, respectively (Figure 5.6). According to the presence or absence of a contig in each library, the number 1 or 0 were assigned to each element, respectively. The Euclidean distances amongst the four libraries were then calculated (Table 5.9). The distance between normal and intersex male libraries (72.01) was found to be closer than the intersex male to any of the two female libraries (84.21 for intersex male and normal female, and 84.17 for intersex male and intersex female). For intersex female, normal female was observed to be the closest one (72.73), followed by intersex male (84.17) and normal male (87.40).

**Fig. 5.6.** A matrix presenting the presence and absence of each *E. marinus* contigs in each of the four libraries. The number 1 or 0 indicate the presence of absence of a contig in the corresponding library. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females.
The Euclidean distance amongst four *E. marinus* libraries.

The Euclidean distance is calculated based on the presence and absence of each contig in each library, NM: normal males; NF: normal females; IM: intersex males; IF: intersex females.

<table>
<thead>
<tr>
<th></th>
<th>NM</th>
<th>NF</th>
<th>IM</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM</td>
<td>0.00</td>
<td>84.44</td>
<td>72.01</td>
<td>87.40</td>
</tr>
<tr>
<td>NF</td>
<td>84.44</td>
<td>0.00</td>
<td>84.21</td>
<td>72.73</td>
</tr>
<tr>
<td>IM</td>
<td>72.01</td>
<td>84.21</td>
<td>0.00</td>
<td>84.17</td>
</tr>
<tr>
<td>IF</td>
<td>87.40</td>
<td>72.73</td>
<td>84.17</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 5.9. The Euclidean distance amongst four *E. marinus* libraries.
5.3.2.3 Functional annotation of *E. marinus* 454 sequences

*E. marinus* contigs were annotated by employing Blast2Go using a stringency cut of E-3. By the Blast against NCBI GenBank, the majority of the 12,645 contigs (11,128 out of 12,645, 88%) were found to be without of any substantial blast hits (Blastx, expect value < 0.001) (Figure 5.7), indicating they are novel sequences. Only 850 (6.7%) contigs were subsequently annotated, and the distributions of their biological processes, molecular functions as well as cellular components were presented by pie charts in Figure 5.8.

![Pie chart showing distribution of Blast and annotation results for 12,645 E. marinus contigs](image)

**Fig. 5.7.** The blast and annotation results for the 12,645 *E. marinus* contigs. The *E. marinus* contigs were from 454 sequencing and de novo assembly. BLAST and annotation were carried out by using Blast2Go.
Fig. 5.8. The GO term distributions of 12,645 *E. marinus* contigs.
A: biological processes; B: molecular functions; C: cellular components.
The 12,645 contigs were then Blast (Blastx) against the Drosophila database (Flybase). A spreadsheet including each contig sequence, and its corresponding blast hit as well as the e-value, were subsequently generated. Of these 12,645 contigs, a total of 1596 genes, presenting 12.6% of the whole list, were found to have the blast hit with the e-value below 0.05.

Based on the read counts of contigs in each of the four sexual phenotypes, three groups (male-, female- and intersex-specific genes) of contigs with differential expressions amongst sexual phenotypes were generated. The three contig groups were annotated by employing Blast2Go, and their distributions on biological processes, molecular functions as well as cellular components were elucidated by Fig. 5.9 through 5.11. In order to assess the gene ontology terms to see if they were enrich in any phenotype-specific gene list, Fisher’s exact test was employed to pairwise compared the male-, female- and intersex-specific contigs. However, no significant difference was observed on the gene ontology terms amongst the three gene groups (Fisher’s exact test, p > 0.05).
Fig. 5.9. The GO term distributions of male-specific contigs of *E. marinus*. A: biological processes; B: molecular functions; C: cellular components.
Fig. 5.10. The GO term distributions of female-specific contigs of *E. marinus*. A: biological processes; B: molecular functions; C: cellular components.
Fig. 5.11. The GO term distributions of intersex-specific contigs of *E. marinus*. A: biological processes; B: molecular functions; C: cellular components.
From the genes which were annotated by GO terms, the ones showing phenotype specificity and being revealed to involve in important biological processes, such as reproduction, immune-system activities, stress response and parasite-host interactions, which related to the corresponding sexual phenotype were listed as potential molecular biomarkers for the sexual phenotype or intersexuality in *E. marinus* (Table 5.10 through 5.12).

<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
<th>Gene Annotation</th>
<th>eValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig12140</td>
<td>26 0 18 0</td>
<td>Reproduction; steroid hormone receptor activity</td>
<td>4.56E-05</td>
</tr>
<tr>
<td>Contig10502</td>
<td>3 0 2 0</td>
<td>Cuticle development</td>
<td>2.83E-16</td>
</tr>
<tr>
<td>Contig11347</td>
<td>19 0 10 0</td>
<td>Structural constituent of cuticle</td>
<td>2.00E-06</td>
</tr>
<tr>
<td>Contig00956</td>
<td>77 0 31 0</td>
<td>Immune system respond</td>
<td>4.82E-10</td>
</tr>
<tr>
<td>Contig01573</td>
<td>8 0 0 0</td>
<td>Male germ cell nucleus</td>
<td>3.94E-34</td>
</tr>
<tr>
<td>Contig10410</td>
<td>22 8 28 9</td>
<td>Reproduction; growth</td>
<td>8.09E-17</td>
</tr>
<tr>
<td>Contig12391</td>
<td>40 27 70 23</td>
<td>Inter-male aggressive behaviour</td>
<td>3.53E-103</td>
</tr>
<tr>
<td>Contig12245</td>
<td>13 0 12 4</td>
<td>Multicellular organismal development</td>
<td>5.97E-51</td>
</tr>
<tr>
<td>Contig10276</td>
<td>16 0 12 6</td>
<td>Embryonic development ending in birth or egg hatching</td>
<td>6.12E-19</td>
</tr>
<tr>
<td>Contig12047</td>
<td>6 0 3 0</td>
<td>Angiogenesis; organelle organization</td>
<td>7.76E-36</td>
</tr>
</tbody>
</table>

**Table 5.10.** Annotated male genes of *E. marinus*.

Contigs were generated from gonads of *E. marinus*, and sequenced by 454 pyrosequencing. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females. * The read number of each contig was normalised to the total read of each library, and their relative abundance were displayed as X/100,000 reads.
<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
<th>Gene Annotation</th>
<th>eValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig00825</td>
<td>NM 0</td>
<td>NF 69</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig05418</td>
<td>NM 0</td>
<td>NF 3</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig05016</td>
<td>NM 0</td>
<td>NF 8</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig01785</td>
<td>NM 0</td>
<td>NF 8</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig06138</td>
<td>NM 0</td>
<td>NF 8</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig06193</td>
<td>NM 0</td>
<td>NF 5</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig09346</td>
<td>NM 0</td>
<td>NF 3</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig09409</td>
<td>NM 0</td>
<td>NF 5</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig05386</td>
<td>NM 0</td>
<td>NF 3</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig03089</td>
<td>NM 2</td>
<td>NF 11</td>
<td>IM 2</td>
</tr>
<tr>
<td>Contig06273</td>
<td>NM 0</td>
<td>NF 16</td>
<td>IM 3</td>
</tr>
<tr>
<td>Contig06196</td>
<td>NM 0</td>
<td>NF 3</td>
<td>IM 0</td>
</tr>
<tr>
<td>Contig00555</td>
<td>NM 0</td>
<td>NF 32</td>
<td>IM 2</td>
</tr>
</tbody>
</table>

Table 5.11. Annotated female genes of *E. marinus*.
Contigs were generated from gonads of *E. marinus*, and sequenced by 454 pyrosequencing. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females. * The read number of each contig was normalised to the total read of each library, and their relative abundance were displayed as X/100,000 reads.
<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
<th>Gene Annotation</th>
<th>eValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig00606</td>
<td>0 0 17 8</td>
<td>Response to stress</td>
<td>2.16E-36</td>
</tr>
<tr>
<td>Contig06741</td>
<td>0 0 2 4</td>
<td>Interspecies interaction between organisms</td>
<td>1.62E-27</td>
</tr>
<tr>
<td>Contig12426</td>
<td>0 0 2 4</td>
<td>Positive regulation of cell proliferation</td>
<td>7.50E-17</td>
</tr>
<tr>
<td>Contig07595</td>
<td>0 0 2 9</td>
<td>Initiation of viral infection; interspecies interaction between organisms;</td>
<td>9.72E-41</td>
</tr>
<tr>
<td>Contig12393</td>
<td>0 0 2 4</td>
<td>Regulation of growth</td>
<td>1.98E-14</td>
</tr>
<tr>
<td>Contig07397</td>
<td>0 0 2 2</td>
<td>Oogenesis</td>
<td>3.15E-15</td>
</tr>
<tr>
<td>Contig07137</td>
<td>0 0 2 2</td>
<td>Collagen and cuticulin-based cuticle development</td>
<td>1.03E-10</td>
</tr>
<tr>
<td>Contig07412</td>
<td>0 0 3 4</td>
<td>Reproduction; embryonic development ending in birth or egg hatching</td>
<td>4.01E-10</td>
</tr>
<tr>
<td>Contig06919</td>
<td>0 0 2 6</td>
<td>Response to stress</td>
<td>7.78E-06</td>
</tr>
<tr>
<td>Contig07598</td>
<td>0 0 3 2</td>
<td>Response to stress</td>
<td>1.76E-10</td>
</tr>
<tr>
<td>Contig07281</td>
<td>0 0 2 8</td>
<td>Positive regulation of epithelial cell proliferation</td>
<td>1.74E-05</td>
</tr>
<tr>
<td>Contig02744</td>
<td>0 0 10 2</td>
<td>Growth; embryonic development</td>
<td>4.60E-08</td>
</tr>
<tr>
<td>Contig08644</td>
<td>0 0 0 4</td>
<td>Gonad development;</td>
<td>4.17E-28</td>
</tr>
<tr>
<td>Contig09131</td>
<td>0 0 0 4</td>
<td>Embryonic development ending in birth or egg hatching</td>
<td>3.60E-11</td>
</tr>
<tr>
<td>Contig06691</td>
<td>0 0 5 0</td>
<td>Ovarian follicle cell migration; ovarian follicular epithelium;</td>
<td>9.86E-09</td>
</tr>
<tr>
<td>Contig09628</td>
<td>0 0 22 4</td>
<td>Response to stress</td>
<td>3.53E-53</td>
</tr>
<tr>
<td>Contig00489</td>
<td>0 0 5 0</td>
<td>Negative regulation of growth of symbiont in host; defense response to Gram-positive bacterium</td>
<td>1.21E-06</td>
</tr>
</tbody>
</table>

Table 5.12. Annotated intersex genes of *E. marinus*.
Contigs were generated from gonads of *E. marinus*, and sequenced by 454 pyrosequencing. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females. * The read number of each contig was normalised to the total reads of each library, and their relative abundance were displayed as X/100,000 reads.
Apart from the genes which can be annotated by the current GO databases, there are many other genes showing phenotype specificity but without available information on their functions in current GO database (Table 5.13 through 5.15). Although their gene functions are not known yet, they may involve important biological processes in crustaceans and can be developed as potential molecular biomarkers of sexual phenotype in crustaceans. Examples of these genes are listed in Table 5.13, 5.14 and 5.15.

<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NM</td>
</tr>
<tr>
<td>Contig12625</td>
<td>906</td>
</tr>
<tr>
<td>Contig00873</td>
<td>505</td>
</tr>
<tr>
<td>Contig12229</td>
<td>382</td>
</tr>
<tr>
<td>Contig10010</td>
<td>291</td>
</tr>
<tr>
<td>Contig12175</td>
<td>257</td>
</tr>
<tr>
<td>Contig12599</td>
<td>206</td>
</tr>
<tr>
<td>Contig11446</td>
<td>279</td>
</tr>
<tr>
<td>Contig12215</td>
<td>229</td>
</tr>
<tr>
<td>Contig12254</td>
<td>317</td>
</tr>
<tr>
<td>Contig10439</td>
<td>266</td>
</tr>
<tr>
<td>Contig12532</td>
<td>257</td>
</tr>
<tr>
<td>Contig09904</td>
<td>182</td>
</tr>
<tr>
<td>Contig00182</td>
<td>224</td>
</tr>
</tbody>
</table>

Table 5.13. Male genes of *E. marinus*.

Contigs were generated from gonads of *E. marinus*, and sequenced by 454 pyrosequencing. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females. * The read number of each contig were normalised to the total read count of each library, and their relative abundance were displayed as X/100,000 reads.
<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NM</td>
</tr>
<tr>
<td>contig12542</td>
<td>1</td>
</tr>
<tr>
<td>contig11612</td>
<td>1</td>
</tr>
<tr>
<td>contig09823</td>
<td>0</td>
</tr>
<tr>
<td>contig00675</td>
<td>0</td>
</tr>
<tr>
<td>contig09067</td>
<td>2</td>
</tr>
<tr>
<td>contig10267</td>
<td>0</td>
</tr>
<tr>
<td>contig12286</td>
<td>1</td>
</tr>
<tr>
<td>contig11785</td>
<td>0</td>
</tr>
<tr>
<td>contig10027</td>
<td>0</td>
</tr>
<tr>
<td>contig10533</td>
<td>0</td>
</tr>
<tr>
<td>contig10857</td>
<td>0</td>
</tr>
<tr>
<td>contig09721</td>
<td>0</td>
</tr>
<tr>
<td>contig11678</td>
<td>0</td>
</tr>
<tr>
<td>contig04287</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.14. Female genes of *E. marinus*.
Contigs were generated from gonads of *E. marinus*, and sequenced by 454 pyrosequencing. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females. * The read number of each contig were normalised to the total read count of each library, and their relative abundance were displayed as X/100,000 reads.

<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NM</td>
</tr>
<tr>
<td>contig11108</td>
<td>0</td>
</tr>
<tr>
<td>contig10158</td>
<td>0</td>
</tr>
<tr>
<td>contig01175</td>
<td>0</td>
</tr>
<tr>
<td>contig01179</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.15. Intersex genes of *E. marinus*.
Contigs were generated from gonads of *E. marinus*, and sequenced by 454 pyrosequencing. NM: normal males; NF: normal females; IM: intersex males; IF: intersex females. * The read number of each contig were normalised to the total read count of each library, and their relative abundance were displayed as X/100,000 reads.
5.3.3 GS FLX sequencing validation by semi-quantitative PCR

The relative abundance of all the 12645 contigs in the four libraries were listed on a spreadsheet, based on the total sequencing reads of each contig in each GS FLX sequenced library. According to their expression difference amongst the four sexual phenotypes, six contigs representing three male-specific genes (M1, M2 and M3), two female-specific genes (F1 and F2) and one intersex-specific gene (I1) were chosen for validating the 454 GS FLX sequencing data. The relative abundance together with the identified Drosophila homologues of the 6 target contigs is shown in Table 5.16 and Table 5.17, respectively.

<table>
<thead>
<tr>
<th>Unigene</th>
<th>Relative Abundance *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NM</td>
</tr>
<tr>
<td>Contig11446 (M1)</td>
<td>448</td>
</tr>
<tr>
<td>Contig00873 (M2)</td>
<td>811</td>
</tr>
<tr>
<td>Contig09845 (M3)</td>
<td>75</td>
</tr>
<tr>
<td>Contig12507 (F1)</td>
<td>0</td>
</tr>
<tr>
<td>Contig12542 (F2)</td>
<td>2</td>
</tr>
<tr>
<td>Contig10158 (I1)</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.16. The relative abundance of six contigs in four E. marinus libraries. The six contigs from 454 sequencing had significant differential expression amongst four sexual phenotypes of E. marinus (NM, NF, IM and IF) .NM: normal males; NF: normal females; IM: intersex females; IF: intersex males. * The read number of each contig were normalised to the total read count of each library, and their relative abundance were displayed as X/100,000 reads.
Unigene | HitName     | % ID | eValue     | HSP_length |
---------|-------------|------|------------|------------|
M1       | FBgn0061469 | 52   | 1.20E+00   | 25         |
M2       | FBgn0005677 | 4.9  | 1.80E+00   | 61         |
M3       | FBgn0037348 | 21.7 | 2.10E-02   | 152        |
I1       | FBgn0033361 | 31.0 | 3.70E-01   | 58         |
F1       | FBgn0026083 | 31.2 | 2.00E+00   | 77         |
F2       | FBgn0038385 | 27.2 | 4.00E-04   | 136        |

Table 5.17. Blast hits of six *E. marinus* contigs against Flybase.
The six contigs showed significantly different expression levels amongst four sexual phenotypes (normal males, intersex males, normal females and intersex females).

Conventional PCR was employed to validate the GS FLX sequencing results. The first attempt applied 32 cycles to conduct the PCR validation. Artifacts were observed in the PCR products of the contig M1 (Figure 5.12, 32 cycles, M1). For the other five validated genes, the present of bands were consistent with the present of EST counts in the four sexual phenotypes (Figure 5.12, 32 cycles). However, all positive bands in 32-cycle PCR products had the equivalent intensity therefore it was deduced that the PCR reaction had exceeded the logarithmic phase for all samples reaching the plateau phase at which point relative abundance could not be determined. The PCR optimisation was then repeated on the same specimens and the cycle number of PCR reduced to 25. Various band intensities were observed, and the expression patterns of the five genes were found to be consistent to their EST read counts from 454 sequencing (Figure 5.12, 25 cycles).
Fig. 5.12. PCR products of six sex-specific genes and GAPDH gene in the four sexual phenotypes in *E. marinus*.

The first figure showed the PCR products after 32 cycles amplification, and 25 cycles for the second. M1, M2 and M3 are male-specific genes; I1 is intersex-specific gene; F1 and F2 are female-specific genes. The same cDNA samples (NM: normal male; NF: normal female; IM: intersex male; IF: intersex female) were used for 454 sequencing and PCR validation. NM+ and NM- were not the 454 cDNA samples, and were generated from pooled infected and uninfected normal males, respectively.
5.4 Discussion

In this study, we characterised the transcriptome in gonads of *E. marinus* by applying 454 pyrosequencing and *de novo* assembly. In order to identify the genes which are differentially expressed in different sexual phenotypes, non-normalised cDNA libraries representing transcripts from each of the four *E. marinus* groups (normal males, intersex males, normal females and intersex females) were sequenced in the same GS FLX sequencing run. Each library was pooled from nine individuals of the same phenotype with the aim to reduce the transcriptomic fluctuation caused by biological variations, such as the moult stage and life cycle. A total number of 12,645 contigs were assembled based on the 213,212 trimmed sequencing reads of the four libraries. The majority of unigenes were found to be novel sequences, which will be supplements to the existing genetic information in crustaceans.

The Roche GS FLX sequencer employed by this study is the second generation of 454 sequencing instrument (after GS20) of the 454 technology. The accuracy of GS FLX sequencing is reported to above 99.5% for the first 200 bp (Droegea and Hill, 2008). The majority of sequencing errors were deletions or insertions caused by the presence of homopolymers, and substitution errors have been found to be as low as $10^{-6}$ (Droegea and Hill, 2008). The single run of GS FLX sequencing takes as long as 10 hours, and is expected to generate above 400,000 reads with the average length of
250 bp for genomic DNA sequencing. In this study, a total of clean 213,212 reads, with the average length of 200 bp, were obtained from a half run of GS FLX sequencing. The read length of transcriptome sequencing could be shorter than genomic DNA sequencing due to the short fragments that were being sequenced, and the average read length for transcriptome sequencing has been reported to be approximately 200 bp (Jarvie and Harkins, 2008), which is consistent with the results of this study.

This study amplified the RNA samples before being reverse transcribed and sequenced, due to the limited availability of intersex specimens in the field and consequently the inadequate amount of cDNA specimens for 454 sequencing. The introduction of aRNA amplification process exclusively enriched mRNA by the complementary hybridisation of T7 primers and poly-A tails of mRNA therefore rRNA, non-coding RNA as well as small RNAs were filtered out of the libraries for 454 sequencing, and there will be possible bias of sequencing reads to the 3’ end of the genes (Phillips and Eberwine, 1996). The T7 RNA polymerases based in vitro transcription was presumed to be a linear amplification therefore the relative abundance of each expressed genes in the whole transcriptome should not be substantially influenced (Baugh et al, 2001; Pabon et al, 2001). Five 454 contigs showing significantly different expression levels among the four libraries were used to validate the 454 sequencing results by employing the semi-quantitative PCR. The
band intensities of the contigs were found to match their EST read counts in the corresponding phenotypes, indicating that relative abundance of transcripts was not compromised by RNA amplification therefore the sequencing read counts are able to reflect the expression level of the corresponding contig.

Non-normalised cDNA libraries were sequenced, and the relative abundance of each EST or contig could be estimated based on their corresponding sequencing read counts. According to the presence or absence of the contigs in four sexual phenotypes, four library-specific and three phenotype-specific gene lists were created. Euclidean distance was calculated and applied to assess the correlation amongst the four libraries (Ewing et al, 1999). Both normal male and normal female library were observed to closely relate to their corresponding intersex counterparts, which is consistent with their biological resemblance. However, the two intersex phenotypes were not found to closely relate to each other. For example, intersex males showed almost the same distance to intersex females (Euclidean distance = 84.17) as to normal females (Euclidean distance = 84.21), indicating that intersex female is not necessary the intermediate form between intersex male and normal female. Intersexes have been presumed as the result of incomplete feminisation, with the intermediate biological features between male and female (Sexton, 1924; Kelly et al, 2002). Data of this study suggested that the transcriptomic divergence between intersex male and intersex female (Euclidean distance = 84.17) is similar to the
difference between normal male and normal female (Euclidean distance = 84.44). It is likely that the two intersex phenotypes are under distinct pathways, instead of preliminary or advanced biological stages induced by a single mechanism.

The three phenotype-specific contig lists (male-, female- and intersex-specific contigs) were annotated by Blast2Go on their biological processes, molecular functions or cellular components, respectively. However, no statistically significant bias was observed on assigned gene ontology terms of the three gene lists. The majority (~80%) of contigs from 454 sequencing were found to be novel sequences that without any substantial blast hits to NCBI GenBank, and less than 10% of contigs were subsequently annotated. The comparison on the gene functions amongst the three gene lists was restricted to the small proportion (~10%) of genes which could be annotated by the putative databases. In other words, the gene annotation is biased to limited number of currently known genes, and not able to truly demonstrate the overall distribution of gene functions of the whole gene list. In order to interpret the gene functions of the contigs sequenced by this study, comprehensive databases are urgently needed.

By annotating the *E. marinus* contigs using GO terms, the functions of the transcripts sequenced by this study were interpreted, and a number of interesting genes were
revealed. Contig12140 which was only found in the two male groups (normal males and intersex males) was discovered to have equavents of androgen receptor genes in human (De Vos et al, 1994), and this gene possibly involves in the steroid hormone relevant pathways in *E. marinus*. Counterparts of several female-specific *E. marinus* genes, such as contig05418, contig09409 and contig01785, were found to involve in the biological process of “hermaphrodite genitalia development” in *Caenorhabditis elegans* (Ceron et al, 2007). This progression is part of the process of “sex differentiation” (www.ebi.ac.uk/QuickGO) and these three genes may contribute to the development of female characteristics. Also, Contig00825, which was annotated to relate to oogenesis in fruit fly and nemotode (Stapleton et al, 2002; Maddox et al, 2005), was observed to be highly expressed in the two female groups and no read was found in either normal or intersex males. Three intersex genes (contig06741, contig07595 and contig00489) were presumed to relate to processes of interspecies interaction, regulation of viral genome replication, or negative regulation of growth of symbiont in host. The intersex specimens applied for 454 sequencing were infected by microsporidian, and clean amphipod samples were utilised for both normal males and normal females. The three intersex-specific genes might be expressed to interact with microsporidian parasites which have been reported to cause intersexes in amphipods (Kelly et al, 2004).

In this study, a huge number of genes which expressed in gonads of various *E.*
marinus sexual phenotypes were sequenced by GS FLX platform. The majority of contigs revealed by this study were found to be novel sequences, which enriched the availability of genetics information for crustaceans. Three phenotype-specific contig lists were generated based on the read counts in different libraries, and the PCR validation on five genes demonstrated that the RNA amplification did not bias the relative abundance of gonadal transcripts in E. marinus therefore the read counts of each contig in the four libraries are able to reflect their expression levels in various sexual phenotypes. A number of phenotype-specific genes were annotated by GO terms, and were revealed to involve in important biological functions, such as reproduction and sex differentiation. The E. marinus gonadal transcripts generated by this study provide a great dataset to conduct future molecular studies in this amphipod species, and biomarkers for intersexuality and sexual phenotypes can be developed on those phenotype-specific genes with annotated gene functions of biological importance.
Chapter 6

General Discussion
This study applied an intertidal amphipod species, *Echinogammarus marinus*, as the model organism, to investigate reproductive dysfunction in crustaceans. Intersexuality has been presumed as the result of dys-functioned endocrine system therefore intersexes provide a good model to study endocrine disruption in crustaceans. This study compared normal and intersex *E. marinus* on both physiological and transcriptomic levels, with the ultimate goal to develop biomarkers for endocrine disruption in crustaceans.

Following on previous studies which observed male and female phenotypes of intersexuality in *E. marinus*, this study found two phenotypes of male intersexuality, described as internal intersex males and external intersex males. The incidence of these different phenotypes varied in *E. marinus* populations from various sites and an increased incidence of intersexuality was found to associate with industrial pollution. The mechanism of intersexuality in crustaceans has not been fully understood, but several factors such as pollutants and feminising parasites infections have been reported as potential causes (Moore and Stevenson, 1991; Rigaud and Juchault, 1998; Terry et al, 2004). Further studies revealed strong correlation between microsporidian infection and external intersex males in *E. marinus*. The relationship amongst feminising parasites, environmental pollution as well as intersexuality in *E. marinus* is discussed in the following section (Section 6.5).

In order to explore sex-specific genes and develop potential molecular biomarkers
for de-masculinisation in crustaceans, the difference of transcriptomic profiling between normal and intersex *E. marinus* were compared in this study. A preliminary study applying a cross-species DNA microarray spotted 349 genes which had significantly different expression levels between normal and intersex males. A further study employing the GS-FLX platform (Roche) was conducted to sequence the whole transcriptome in gonadal tissues of normal and intersex amphipods. Sex-specific as well as intersex-related genes were revealed, which could be candidate biomarkers indicating sex-phenotypes or the occurrence of de-masculinisation in crustaceans.

In this chapter, key findings of previous chapters were reviewed, and the relationship amongst intersexuality, environmental pollutions and microsporidian parasite infection was discussed. Two male intersex phenotypes were also compared on their prevalence in field, proportions of microsporidian infection as well as gene expression patterns, with the aim to shed some light on the mechanism of intersexuality in amphipods. The microarray and 454 sequencing studies have discovered a huge number of genes which were found to associate with sexual phenotypes in amphipods. To understand the functions of these genes will facility the genetic studies on crustaceans, and molecular biomarkers could be developed to indicate the endocrine disruption in crustaceans.
6.1 Evidence of endocrine disruption in *E. marinus*

Endocrine disruption in wildlife has recently been a concern amongst scientists, with impaired male fertility or fecundity being reported in a lot of animal groups, such as alligators, fish, amphibians and molluscs (Guillette et al, 1994; Jiménez 1997; Lye et al, 1998; Hayes et al, 2002; Nice, 2005). In this study, the fecundity of male amphipods from clean and polluted sites was compared based on their sperm counts. Normal males from an industrially impacted site – Inverkeithing were found to have significantly less sperm counts than ones from the two reference sites, Loch Fleet and Thurso (Chapter 2). The relationship between lower sperm counts and smaller broods in normal amphipods has been reported by previous studies (Dunn et al, 2006a). Therefore, the decreased sperm counts in amphipods could conceivably have adverse effects at the population level.

The androgenic gland is known to regulate male differentiation and spermatogenesis in amphipods (Charniaux-Cotton, 1954). Charniaux-Cotton (1962) reported that male amphipod *Orchestia gammarella* had reduced feminised gnathopods and very limited number of spermatozoa left in testes four or five months after removal of androgenic gland, and some males even totally transferred their testes into ovaries. Androgenic gland ablation has also been reported to lead to regression of spermatogenesis in other crustaceans (Nagamine et al, 1980; Khalaila et al, 1999). Within the crustacean endocrine system, the androgenic gland is under a tier of control from other endocrine organs. For example, Khalaila et al (1999) reported
androgenic gland hypertrophy in eyestalk ablated crayfish, which indicates that X-organ in eyestalk is capable of inhibiting the activities of androgenic gland. Other studies have demonstrated the role of methyl farnesoate in the regulation of gonadal development in prawns (Nagaraju et al, 2003).

Evidence of endocrine disruption in Inverkeithing E. marinus population has been discovered in previous studies. Ford et al (2004a) reported that morphometric characters, for example the size of gnathopods and coxal plate depth and width, of normal males from Inverkeithing were found to resemble intersex males more than normal males from reference sites. Both the reduced size of gnathopods (Ford et al, 2004a), as well as the decreased number of spermatozoa revealed in normal males (Yang et al, 2008; Chapter 2, Section 2.3.4.1), indicates the possible occurrence of de-masculinisation in this amphipod population. Environmental contamination may be directly impacting the activities of the androgenic glands of E. marinus, or possibly inducing endocrine disruption by affecting the function of other glands, such as X-organ and mandibular organ. Further studies are required to reveal the association between the industrial pollutions in Inverkeithing and their targeted organs of male E. marinus.
6.2 Prevalence of intersexuality in *E. marinus* populations from clean and polluted sites

The levels of intersexuality in *E. marinus* populations from five sites in the UK (Loch Fleet, Thurso and Inverkeithing in Scotland, as well as Portsmouth Harbour and Langstone Harbour in Portsmouth, England) were examined in Chapter 2. Three phenotypes of intersexuality were recorded, internal intersex males, external intersex males and intersex females, based on their internal and external sexual characteristics. Intersexes have been revealed in a variety of amphipod species, such as *G. duebeni* and *O. gammarallus* (Ginsburger-Vogel and Carre-Lecuyer, 1976; Kelly et al, 2004). Some amphipods were found to have only one phenotype of intersexuality, such as intersex females in *G. duebeni* as well as intersex males in *Corophium volutator* (Dunn et al, 1990; Barbeau and Grecian, 2003), whilst both intersex males and intersex females have been observed in other species, for example *O. gammarellus* (Ginsburger-Vogel, 1975).

In the three Scottish sites, a significantly higher proportion of intersexes was revealed in the industrially polluted site than the two reference sites. Similarly, in a parallel investigation carried out in southern England, the proportion of total male intersexes (external and internal) was also found to be higher in an industrially impacted site - Portsmouth Harbour (20.7%) than in the reference site - Langstone Harbour (10.1%), whilst the difference was not statistically significant (Fisher’s exact test, p>0.05). Inverkeithing was classified as “Seriously Polluted” site by
SEPA, and upgraded as “Unsatisfactory” in 2002. It was impacted by the heavy metals and PCBs, which were mainly contributed by a shipbreaker’s yard and a decommissioned paper mill. Portsmouth Harbour was affected by the historic ship painting industry, and has been reported to have increased level of anti-fouling chemicals, such as TBT and Irgarol 1051 (Gough et al, 1994; Zhou, 2008).

Elevated incidence of intersexuality in crustaceans has been reported to associate with environmental contaminations by other studies (Moore and Stevenson, 1991; 1994). This study also revealed increased incidence of intersexes in polluted sites. However, it is not clear whether intersexuality was directly induced by environmental contaminations, or it was caused by the feminisation of microsporidian parasites, and pollutions merely exerted indirect effects by increasing the incidence of microsporidian infection through suppressing the immune system of amphipods. Further studies were then conducted with the aim to identify the relationship between microsporidian infection and intersexuality in *E. marinus*. 
6.3 Microsporidian infection in *E. marinus* populations

This study revealed a strong correlation between microsporidian infection and intersexuality in various *E. marinus* populations. This consistent relationship between intersexes and microsporidian infection has been reported in other amphipod species, such as *Gammarus duebeni* and *Gammarus roeseli* (Haine et al., 2004; Kelly et al., 2004), and our observations were in accord with those studies. The specific species of microsporidian infecting *E. marinus* populations were identified for the first time by this study (Chapter 3). Geographical difference of dominant parasite species were observed -- *D. duebenum* and *D. berillonum* was found to dominant in Inverkeithing and Portsmouth *E. marinus* populations, respectively (Chapter 3).

Although the association between male intersexuality and microsporidian infection has been revealed in both *E. marinus* populations – Inverkeithing and Portsmouth Harbour, the parasite species related to male intersexes were found to be different in two sites. In Inverkeithing, male intersexes were found to be related to *D. duebenum*, whilst *D. berillonum* was revealed to correlate to male intersexuality in Portsmouth Harbour. *D. duebenum* has been discovered as a feminising parasite in several amphipod species, such as *Gammarus tigrinus, Gammarus duebeni* and *Echinogammarus berilloni* (Terry et al., 2004). Although a variety of amphipods have been observed to harbour *D. berillonum*, this parasite has never been reported as a sex distorter in these amphipod populations. Fenimising parasites have been reported
to associate with female biased population in amphipods (Bulnheim, 1965). Inverkeithing *E. marinus* population has been revealed to be female biased (Ford et al, 2004a), and the microsporidian infecting this population is found to be a feminising parasite. The fact that *E. marinus* population in Portsmouth was male biased indicates the feminising effect of the parasite infecting this amphipod population is not strong.

Amphipods infected by different microsporidian parasites in different populations have been reported by previous studies. For example, Ironside et al (2003b) screened *G. duebeni* populations from three sites for parasite infection. In Fintray Bay and Stinking Bay (Scotland), two microsporidian species were observed in *G. duebeni* populations, whilst only one microsporidian parasite was revealed in Ballochmartin Bay (Scotland). Similarly, *Nosema granulosis* was found to dominant over *D. duebenum* in *G. duebeni* populations from Isle of Cumbrae and Isle of Man, whilst *Nosema granulosis* was not observed in *G. duebeni* from Ballochmartin Bay (Hogg et al, 2002; Ironside et al, 2003b).

Various reasons could potentially cause the different composition of microsporidian parasites in different amphipod populations, for instance environmental factors. High salinity has been reported to impede the transmission of microsporidians, and lead to low proportion of infected offspring in *G. duebeni* (Dunn and Hatcher, 1997). Elevated level of microsporidian infection has also been found in amphipods
exposed to endocrine disrupting chemicals (Jacobson et al, 2010; 2011). The combination of environment fluctuations is able to affect the fitness of microsporidian parasites, and finally contributes to the different distribution of microsporidian parasites in various amphipods populations.
6.4 Internal and external intersex males

Depending on whether the male intersexuality could be determined from external appearance or only internally, two male intersex phenotypes were differentiated. In Inverkeithing, the majority of external intersex males were also found to have the internal intersex character – oviduct, which indicated that an internal intersex male was possibly a prerequisite for an external intersex male (Yang et al, 2008; Chapter 2). In contrast, in Portsmouth Harbour, 11 out of the 12 examined external intersex males did not possess oviducts, and internal and external intersex characteristics were rarely observed on the same male amphipod, suggesting that internal intersex male is not necessarily the early stage of the external intersex male phenotype (Yang et al, 2010; Chapter 3).

Differences in the proportion of microsporidian infection were also observed in internal and external intersex males. For example, in Inverkeithing Bay, all examined external intersex males (n=16) were found to be infected by microsporidian, whilst the proportion of infected individuals in internal intersex males was less than 25% (n=25). Intersexes have been presumed to be induced by the incomplete feminisation by microsporidians (Kelly et al, 2002). The difference between internal and external intersex males on the relationship with microsporidian suggests that the two male intersexes are possibly caused by different mechanisms.

The internal female character -- oviduct has been reported in intersex males of
various amphipod species (Rodger-Gray et al, 2004; Ford et al, 2005b). Juvenile amphipods have undifferentiated gonads with the anlagen of both oviduct and vas deferens (Lockwood, 1968). Oviducts in normal male *Orchestia gammarella* were reported to degenerate quickly after sex differentiation, whilst they persisted for 2 or 3 moult cycles in young intersex males (Ginsburger-Vogel, 1975). Oviducts observed in internal intersex males *E. marinus* was likely due to delayed degeneration, but specimens with very well developed oviducts indicated the later development was also possible. Charniaux-cotton (1962) proposed the development of brood plates (oostegites) in amphipod was controlled by ovaries in *O. gammarella*, although ovarian hormones have not been identified in Crustacea (Suzuki, 1999). Brood plates observed in external intersex male *E. marinus* were probably induced by ovarian-hormone-like effects which are not over-ridden by androgenic gland, possibly from the environmental pollution or feminising parasite infection.

The microarray study on normal male and two male intersex phenotypes revealed that internal and external intersex males had different expression patterns, suggesting that the two male intersexes are likely two distinct phenotypes. The PCA analysis clearly differentiated the three male groups based on the genes which were significantly differently expressed in the three male phenotypes (Chapter 4). The genes having differential expression levels (p<0.05) between the two male intersex groups (230 genes) were found to be less than the differently expressed genes between normal males and any of the intersex males (453 and 351 genes...
differentially expressed between normal males VS internal intersex males, and normal males VS external intersex males, respectively). It suggests that there were less gene expression difference between the two male intersexes than between normal males and any of the intersex groups. In other words, although the PCA analysis indicates that the two male intersexes are distinct phenotypes, internal and external males do have the similarity on the transcriptomic level to some extent.
6.5 Relationship amongst intersexuality, parasite and pollution

The mechanism of intersexuality in crustaceans has not been fully understood, however several factors such as environmental pollution and feminising parasite infection have been identified to correlate with intersex prevalence (Moore and Stevenson, 1991; 1994; Kelly et al, 2004). This study also revealed a high incidence of intersex *E. marinus* in industrially impacted site as well as strong correlation between intersexuality and microsporidian infection. Although high incidence of intersexuality has been revealed in polluted site, it is not clear whether the intersexes could be directly caused by the pollution. Elevated level of microsporidian parasites by man-made chemicals has been reported (Jacobson et al, 2010; 2011), and it is possible that the high proportion of intersexes in Inverkeithing might be due to the increased incidence of microsporidian infection in industrially polluted site.

Intersex male *E. marinus* from various sites were observed to have significantly higher body weight than normal male specimens (Chapter 2, Section 2.3.2), and the result was consistent with previous findings (Ford et al, 2003a). An investigation on the moult frequencies of normal and intersex *E. marinus* revealed that higher proportion of adult intersexes moulted than their normal counterparts (Chapter 2, Section 2.3.3). The inter-moult duration has been reported to increase with the age of amphipods (Sexton, 1928). Observed higher moult frequency in intersexes indicates the constantly progressive growth which leads the intersexes have larger body weight than their normal counterparts. Intersexes were presumed to suffer some level
of endocrine disruption, and the increased body weight in intersexes was also possibly caused by the delayed maturation (Ford et al, 2004b). Reduced fecundity or fertility has also been reported in intersex *E. marinus* (Ford et al, 2003a; Yang et al, 2008), and the high body weight may be contributed by the energy transferred from the impaired reproductive system.

Sexton (1924) reported that male and female *Gammarus chevreuxi* could not be differentiated from each other until the sixth moult, whilst intersex individuals can be distinguished from normal amphipods at as early stage as the third moult, by their relatively small body size. In addition, intersex *G. chevreuxi* were found to need more mouls to display sexual characteristics than normal amphipods (Sexton 1924). The fact that intersex individuals are significantly smaller than their normal counterparts suggests the occurrence of delayed development or maturation. Following this hypothesis, the fact that external intersex males were heavier than internal intersex males indicates the former might suffer even severer endocrine dys-function than internal intersexes.

Parasite-induced gigantism has been reported in some animal groups, such as mammals and fish (Cheng, 1971; Pearre, 1976). Wilson and Denison (1980) reported that snails infected by *Fasciola hepatica*, a parasitic flatworm species, were found to be significantly heavier than control samples. However, the total biomasses (tissue mass plus cumulative egg weight) between infected and clean females were not
significantly different, suggesting that the gigantism in infected individuals was caused by the nutrition supply switch from reproductive system onto somatic tissue (Wilson and Denison, 1980). In this study, normal male *E. marinus* from Loch Fleet were revealed to be significantly heavier than the ones from Inverkeithing and Thurso (Chapter 2, Section 2.3.2). A wide prevalence of trematode parasite was observed in amphipods from Loch Fleet, however, the proportion of infected specimens as well as the intensity of infections were not quantified in this study. Trematode parasite induced gigantism has been reported in snails (Mouritsen, 1994; Gorbushin, 1997; Chapuis, 2009), and is a potential explanation for the increased body weight observed in males from Loch Fleet. However, further studies are needed to validate the relationship.

Reduced size of gnathopods as well as decreased sperm counts have also been reported in normal male *E. marinus* from Inverkeithing (Ford et al, 2004a; Chapter 2, Section 2.3.4.1), suggesting the occurrence of endocrine disruption possibly induced by the industrial contamination. The reduced fertility and fecundity in female intersexes have been reported in *E. marinus* (Ford et al, 2003a, 2004b). Intersex females were found to produce 20% less eggs compared to normal females, and they lost 30% of the eggs in the late development stage when their normal counterparts lost only 20% (Ford et al, 2003a). In addition, male amphipods have been found to allocate more sperm to uninfected females, and female productivity was reported to be impaired under sperm limitation (Dunn et al, 2006a). In *Armadillidium vulgare*,
*Wolbachia* infected females were also revealed to suffer reduced fertility under sperm depletion (Rigaud & Moreau 2004). The reduced fertility in intersex females might be contributed by both reduced spermatozoa in males as well as the male sperm-allocation strategy, and the high level of intersexes in Inverkeithing could lead to the reduction of the population.
6.6 Genetic studies in *E. marinus*

Constrained by the information of genomic DNA or cDNA sequences, the genetic studies in crustaceans is relatively sparse comparing to the model species from other animal groups, for example mammals, fish or insects. In this study, differential gene expression patterns amongst three male phenotypes of *E. marinus* have been revealed by employing a cross-species microarray (Chapter 4). However, the sequence divergence between *G. pulex* and *E. marinus* was able to reduce the sensitivity of microarrays therefore decrease the reliability of the results, for example the cross-species microarray is not able to distinguish differential hybridisations of mismatch from expression variations (Gilad et al, 2005). Also, it could be a challenge to apply the *G. pulex* cDNA sequences onto *E. marinus* for further investigation without knowing the actual DNA sequence of *E. marinus* itself. The transcriptome sequences of *E. marinus* were required in order to develop molecular biomarkers in this amphipod species.

Thanks to the recently introduced next-generation sequencing technologies, the DNA sequence of non-model species can be sequenced with substantially reduced cost and time compared to the traditional BAC and Sanger based sequencing technique (Rothberg and Leamon, 2008; Schuster, 2008; Metzker, 2010). Currently, there are three main next-generation sequencing technologies available on the market, which are 454 (Roche) and Solexa (Illumina) applying the sequencing-by-synthesis concept,
and SOLiD (Life Technologies) employing the sequencing-by-ligation strategy (Bennett, 2004; Marguelis et al, 2005; Shendure et al, 2005; Ellegren, 2008). This study employed 454 sequencing to carry out the de novo transcriptome sequencing for gonadal tissue of *E. marinus*. Read length together with fragments number has been considered as crucial factors for sequence assembly (Chaisson et al, 2009). The 454 sequencing has the longest read length amongst the three next-generation sequencing platforms, and makes it the ideal technology for de novo sequencing and assembly.

This study generated a total number of 213,212 ESTs and 12,645 contigs by sequencing the transcripts from gonadal tissues of *E. marinus* presenting four sexual phenotypes (Chapter 5). This is the first study to conduct de novo transcriptome sequencing in an amphipod species, and the huge volume of cDNA sequences provide good supplements to the current genetic information for crustaceans. Although a huge number of cDNA sequences in *E. marinus* become available, it is a challenge to annotate and interpret the gene functions. The majority (~88%) of contigs were found to be novel sequences therefore without any substantial blast hits to the NCBI GenBank (Chapter 5, Section 5.3.2.3). To date, Drosophila is the best documented genus that most closely related to amphipods, however, the phylogenetic relationship between these two groups is still quite distinct (Averof and Akam, 1995). In 2011, *Daphnia pulex* became the first Crustacea species that had its
whole genome sequence available (Colbourne et al, 2011). Because of the increasingly dropped cost for high throughput sequencing, it is not surprising to see more and more DNA sequence information from crustaceans. However, well studied model species and intensively annotated gene database of crustaceans are urgently needed in order to interpret the gene functions and apply the data for further studies.

A number of genes showing phenotypic specificity were annotated and found to involve in important biological functions. For example, the female-specific gene Contig00825 was annotated to relate to oogenesis, and the male gene Contig12140 and Contig01573 were found to play a role in steroid hormone receptor activity and male germ cell nucleus, respectively. In addition, three of intersex-specific genes are discovered to interact to the parasites, and their expressions are possibly induced by the microsporidian infection. The expression patterns of these contigs are accord with their gene functions, and they are good candidates for indicating sexual phenotype in *E. marinus*. Based on the relative expression levels and annotated biological functions of contigs sequenced by this study, potential molecular biomarkers for intersexuality and sexual phenotypes can be developed and utilised for future genetic studies onto this amphipod species. This study has generated a wealth of sex specific or biased sequences from comparisons of males, females and intersexes. Due to the lack of current annotation of crustacean genes, an important next step in this field of science is to conduct a large scale series of gene function studies to fully take advantage of the resource.
Main Conclusions of This Thesis

1. Following the previous studies which revealed male and female intersexes in an intertidal amphipod species *Echinogammarus marinus*, this study further identified two phenotypes of male intersexes (internal intersex males and external intersex males) in this amphipod species.

2. This study investigated the proportion of intersexes in various *E. marinus* populations, and revealed increased incidence of intersexuality in an industrially polluted site, Inverkeithing, compared to reference sites, Thurso and Loch Fleet.

3. By comparing the sperm counts in males from clean and polluted sites, male amphipods from industrially impacted site were found to have approximately 20% less spermatozoa than the ones from clean sites.

4. Intersex *E. marinus* were revealed to have significantly higher body weights than their normal counterparts.

5. Statistically significant difference was observed on the moult frequency of four sexual phenotypes in *E. marinus*, with intersex males having the highest moult frequency, followed by normal males, intersex females and normal females.

6. Microsporidian infection was identified in *E. marinus*, and the parasite species, *Dictyocoela berillonum* and *Dictyocoela duebenum*, were determined for the first time in this amphipod. Further studies revealed strong association between the occurrence of microsporidian infection and the incidence of intersexuality in *E. marinus*. 
7. A cross-species cDNA microarray was employed to characterise the gene expressions in gonads of three male phenotypes of *E. marinus*, and PCA analysis clearly differentiated the three male groups -- normal males, internal intersex and external intersex males.

8. The *de novo* transcriptome sequencing was carried out on the gonads of four sexual phenotypes of *E. marinus*, by employing the next generation sequencing platform-Roche GS FLX sequencer. A total of 12,645 contigs were assembled and several phenotypic-specific gene lists were generated and subsequently annotated by GO terms. A number of phenotype-specific genes, for example female gene Contig00825, male genes Contig12140 and Contig01573, and intersex genes Contig06741, Contig07595 and Contig00489, were found to involve in important biological functions such as sex differentiation and reproduction, and can be developed as molecular biomarkers for indicating sexual phenotype and endocrine disruption in amphipods.

9. Based on the differential gene expression patterns, as well as the different proportions of microsporidian infected individuals between internal and external intersex males, the two male intersex phenotypes of *E. marinus* are presumed as two distinct phenotypes.
Acknowledgement

This studentship was funded by the UHI Millennium Institute ARC Programme (an initiative jointly supported by Highlands and Islands Enterprise, the Scottish Funding Council and the European Regional Development Fund) and supported by a Natural Environmental Research Council (NERC UK) molecular facilities grant (MGF 223), University of Portsmouth and Cardiff University.

I am very grateful to my supervisors, Dr. Alex Ford and Dr. Peter Kille. It was a wonderful experience working with Alex and Pete for my three-year PhD project. Many thanks for equipping me with the statistical and biological skills, and more importantly, the scientific logic and morality. Thank you for the supervision and guidance on my work, and I really appreciate your help and encouragement when I am writing up this thesis.

Thanks for the advice provided by Nina Wedell and Christopher Tudge in the analysis and quantification of amphipod spermatozoa. Thank Dr. Christine Samples for developing the protocol of cross-species microarray hybridisation. I am grateful to Dr. Jane Andre, Dr. Eleftheria Pervolaraki and Dr. Victoria Workman for their help when I was doing molecular biological experiments in Cardiff University. And thanks to Dr. Stephen Short for his advice on the microsporidian chapter, and also thank Dr. John Morgan, Dr. Jane Andre, Dr. Stephen Short, Yasmin Guler and Robert Mansergh for their help on *E. marinus* sample collection. I am also very grateful to Prof. Yang Wanxi, Wang Dahui, Wang Wei and Yu Keming for providing protocol, reagents and instruments on the work of amphipod sperm ultrastructure identification in Zhejiang University in China.

I would like to thank Cardiff University, Portsmouth University and Zhejiang University for offering office and instruments for my PhD project. I am thankful to Dr. Kenneth Boyd, Mona Larsen and other staff in ERI (Environmental Research Institute) for their help on my PhD project. Finally, I want to thank my wife Xiaoyan Su for her assistance and support on my work.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGH</td>
<td>Androgenic Gland Hormone</td>
</tr>
<tr>
<td>ANCOVA</td>
<td>Analysis of Covariance</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DDT</td>
<td>Dichlorodiphenyltrichloroethane</td>
</tr>
<tr>
<td>ED</td>
<td>Endocrine Disruption</td>
</tr>
<tr>
<td>EDC</td>
<td>Endocrine Disruption Chemicals</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed Sequence Tag</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein Isothiocyanate</td>
</tr>
<tr>
<td>FOI</td>
<td>Frequency of Incorporation</td>
</tr>
<tr>
<td>LOWESS</td>
<td>Locally Weighted Scatterplot Smoothing</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>SEPA</td>
<td>Scottish Environment Protection Agency</td>
</tr>
<tr>
<td>SNP</td>
<td>Single-Nucleotide Polymorphism</td>
</tr>
<tr>
<td>SSU</td>
<td>Small Sub-Unit</td>
</tr>
<tr>
<td>TBT</td>
<td>Tributyltin</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission Electron Microscopy</td>
</tr>
</tbody>
</table>
Appen

Appendix

QPCR validation for the microarray study characterising three male phenotypes (normal males, internal and external intersex males) in *E. marinus*.

<table>
<thead>
<tr>
<th></th>
<th>MV1</th>
<th>MV2</th>
<th>MV3</th>
<th>MV4</th>
<th>MV5</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM1</td>
<td>271.7</td>
<td>2.1</td>
<td>761.7</td>
<td>2539.8</td>
<td>0.68</td>
</tr>
<tr>
<td>NM2</td>
<td>237.2</td>
<td>0.63</td>
<td>709.6</td>
<td>3465.5</td>
<td>0.94</td>
</tr>
<tr>
<td>NM3</td>
<td>213.3</td>
<td>1.01</td>
<td>666.9</td>
<td>2565.2</td>
<td>83.3</td>
</tr>
<tr>
<td>NM4</td>
<td>712.4</td>
<td>1.81</td>
<td>1481.3</td>
<td>8007.5</td>
<td>1.37</td>
</tr>
<tr>
<td>NM5</td>
<td>591.6</td>
<td>1.50</td>
<td>1439.5</td>
<td>6017.7</td>
<td>0.32</td>
</tr>
<tr>
<td>IIM1</td>
<td>603.5</td>
<td>4.15</td>
<td>1676.0</td>
<td>6225.8</td>
<td>0.29</td>
</tr>
<tr>
<td>IIM2</td>
<td>685.9</td>
<td>3.00</td>
<td>1800.3</td>
<td>7403.3</td>
<td>0.43</td>
</tr>
<tr>
<td>IIM3</td>
<td>123.0</td>
<td>1.76</td>
<td>280.9</td>
<td>1019.1</td>
<td>0.053</td>
</tr>
<tr>
<td>IIM4</td>
<td>730.8</td>
<td>0.93</td>
<td>1944.7</td>
<td>10597.7</td>
<td>0.32</td>
</tr>
<tr>
<td>IIM5</td>
<td>767.8</td>
<td>2.68</td>
<td>1748.7</td>
<td>7219.8</td>
<td>0.40</td>
</tr>
<tr>
<td>EIM1</td>
<td>677.8</td>
<td>1.58</td>
<td>1796.4</td>
<td>10424.2</td>
<td>0.30</td>
</tr>
<tr>
<td>EIM2</td>
<td>841.2</td>
<td>9.39</td>
<td>1645.6</td>
<td>5172.0</td>
<td>0.63</td>
</tr>
<tr>
<td>EIM3</td>
<td>669.9</td>
<td>1.85</td>
<td>1733.0</td>
<td>8175.2</td>
<td>0.52</td>
</tr>
<tr>
<td>EIM4</td>
<td>986.9</td>
<td>0.88</td>
<td>1351.6</td>
<td>5186.0</td>
<td>0.36</td>
</tr>
<tr>
<td>EIM5</td>
<td>742.2</td>
<td>13.3</td>
<td>1560.3</td>
<td>10947.0</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 7.1. The relative expression levels of the five genes chosen for the microarray data validation using the SYBR green based QPCR. The beta-tubulin gene was used as the internal control. NM: normal male; IIM: internal intersex male; EIM: external intersex male.

259
Fig. 7.1. The relative expression levels of the five validated genes in three male phenotypes. The expression levels were quantified by microarray (above) and QPCR (below), respectively. The beta-tubulin was used as the internal control for the both methods. NM: normal male; IIM: internal intersex male; EIM: external intersex male.
Reference


261


Gammarus duebeni. Invertebr. Reprod. Dev. 25, 139-142.


Maguire, T.L., Grimmond, S., Forrest, A., Iturbe-Ormaetze, I., Meksem, K.,


Metzker, M.L., 2005. Emerging technologies in DNA sequencing. Genome Res. 15,
1767-1776.


gammaridean amphipods: variation of life history traits with latitude, depth, salinity and superfamily. Hydrobiologia. 223, 189-227.


Yang, G., Kille, P., Ford, A.T., 2008. Infertility in a marine crustacean: Have we been ignoring pollution impacts on male invertebrates? Aquatic Toxicology. 88, 81-87


