On the use of Genomics to Assess Environmental Risks of Pharmaceuticals

Lina Gunnarsson

UNIVERSITY OF GOTHENBURG

Department of Physiology / Endocrinology
Institute of Neuroscience and Physiology
The Sahlgrenska Academy
University of Gothenburg
Sweden, 2009

A doctoral thesis at a university in Sweden is produced either as a monograph or as a collection of papers. In the latter case, the introductory part constitutes the formal thesis, which summarizes the accompanying papers. These papers have already been published or are in manuscript at various stages (in press, submitted or in manuscript)

The cover picture was designed by Fredrik Jutfelt and Lina Gunnarsson. The left microarray picture shows the two-channel cDNA salmonid microarray used in paper II and the right picture shows the one-channel Geniom oligonucleotide microarray used in paper IV.
Göteborg, Sweden, 2009
ISBN 978-91-628-7749-1
Printed by Intellecta Infolog AB, 2009

ABSTRACT

Many drugs are found in the aquatic environment and are therefore a cause for concern. Low concentrations of active ingredients from human pharmaceuticals reach the environment via sewage treatment plants, mainly as a result of excretion. However, other routes, such as incorrect disposal and direct releases from manufacture, could also be of importance.

The effects of residual drugs in the environment are not well understood. This thesis addresses the issue by using different genomic techniques. The evolutionary conservation of 1,318 human drug targets were predicted in 16 species from different taxonomic groups. We show that the majority of the drug targets are conserved in aquatic vertebrates, while invertebrates and plants lack orthologs to many of the targets. The presented predictions can serve as a basis for identifying potentially sensitive (and insensitive) species that are used for the environmental risk assessment of pharmaceuticals.

The effects on fish of exposure to a single pharmaceutical (ethinylestradiol) as well as a complex industrial effluent that contains high levels of many drugs were explored using microarray analysis. We identified two sensitive and potentially robust biomarkers of estrogen exposure by performing a meta-analysis that combined our results with data from the literature. The identified biomarkers were also used to evaluate the ability of different sewage treatment technologies to remove estrogenic substances. Several treatment technologies reduced the levels of estrogenic substances, but ozonation was required to remove all measured biological effects. The fish that were exposed to a high dilution of the industrial effluent showed increased hepatic Cyp1a enzyme activity and altered expression of several genes that are involved in the detoxification of chemicals and drugs. Although the gene expression pattern did not clearly point to any specific group of substances, it could serve as a basis for hypothesizing mechanisms of toxicity and possible causative agents in the effluent.

More research is needed to understand the risks of residual drugs in the environment, and the presented results show that genomic approaches are useful for this purpose.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Många av de läkemedel vi tar bryts inte ner fullständigt av våra kroppar. Istället transporteras resterna genom reningsverken och hamnar slutligen i våra sjöar och vattendrag. Låga koncentrationer av flera olika läkemedel har hittats i miljön, men vi vet väldigt lite om vilka effekter de har på djur- och växtliv. Det finns dock en risk att djur och växter kan påverkas av läkemedelsresterna

Målet med denna avhandling har dels varit att plocka fram data som kan förbättra regelverket kring miljöriskbedömningen av läkemedel, dels utveckla verktyg som kan öka möjligheterna att upptäcka miljöeffekter av läkemedel. För att göra detta har vi använt olika storskaliga molekylärbiologiska metoder.

Läkemedel är ett viktigt verktyg inom sjukvården som används för att förebygga, lindra och bota olika sjukdomstillstånd. Ett läkemedel är designat eller utvalt för att specifikt kunna påverka ett visst biologiskt system och samtidigt minimera påverkan på kroppens andra funktioner. Lite förenklat kan man säga att läkemedel binder och påverkar främst den målmolekyl, till exempel en receptor eller ett enzym, som har den funktion i kroppen som läkemedlet är designat att påverka. Läkemedelssubstanser är i många fall mycket kraftfulla ämnen. Det behövs endast en låg koncentration av läkemedlet för att dess målmolekyl ska påverkas. Detta innebär att djur som har målmolekyler som liknar våra löper en större risk att påverkas av de låga koncentrationer av läkemedelsrester som hittas i miljön. Det är exempelvis känt att det syntetiska östrogenet i p-piller binder till östrogenreceptorn i fisk. Det kan leda till att reproduktionen hos fisk påverkas redan vid väldigt låga koncentrationer av det syntetiska östrogenet i vattendrag.

I den första artikeln i den här avhandlingen har vi kartlagt vilka grupper av organismer som har målmolekyler som är lika människans. På så sätt har vi kunnat se vilka organismer som det är troligt att läkemedelsresterna påverkar. Fisk och groda har en motsvarande molekyl till 80% av de 1318 undersökta målmolekylerna. Vattenloppan saknade däremot målmolekyler till många läkemedel och algen saknade i stort sett målmolekyler till alla läkemedel. Idag görs dock mycket av miljörisk-bedömningen av läkemedel inte på fisk, utan på ryggradslösa djur, såsom vattenloppor eller på alger eller växter. Vår studie visar att vattenloppan i många fall inte är representativt som försöksdjur för miljöriskbedömning av humana läkemedel. Baserat på våra resultat föreslår vi att miljöriskbedömningen av humana

läkemedel i högre grad bör baseras på försök med fisk eller groda. Kartläggningen av målmolekyler kan också användas som en guide för att avgöra om det är relevant att använda andra organismer än fisk och groda för en miljöriskbedömning av ett läkemedel.

I de andra tre studierna som den här avhandlingen bygger på har vi tittat på hur läkemedel kan påverka genuttrycksmönstret i levern hos fisk. Genuttrycksmönstret förändras hela tiden hos alla organismer som ett svar på att miljön förändras. Vardagliga förändringar så som dag/natt eller kallt/varmt påverkar genuttrycket, men en exponering för en kemikalie, exempelvis ett läkemedel, leder i många fall också till genuttrycksförändringar. Ett förändrat genuttryck är alltså inget negativt i sig självt. Genom att studera uttrycket hos flera tusentals gener samtidigt så kan vi få en förståelse för vilka biologiska processer som satts igång i fisk som exponerats för ett läkemedel. Vi kan också få indikationer på om förändringarna i genuttryck innebär något negativt för fisken. Metoden som vi har använt kallas microarray eller gen-chip-analys och med den kan vi mäta uttrycket av tiotusentals gener samtidigt.

Tidigare studier har visat att det renade avloppsvattnet från reningsverk kan innehålla tillräckligt höga koncentrationer av östrogena ämnen för att feminisera hanfisk och på så sätt störa reproduktionen. Det syntetiska östrogenet i p-piller är en substans som starkt bidrar till att hanfiskar blir feminiserade. Vår målsättning med den andra studien artikel nummer två var att hitta förändringar i genuttryck som kan användas för att visa att hanfisk utsatts för östrogena ämnen. Regnbågslax exponerades för antingen en hög eller en låg koncentration av det syntetiska östrogenet och vi analyserade fiskens genuttrycksmönster med microarrayer. Vi identifierade två genuttrycks-markörer som kan användas för att visa att hanfisk utsatts för östrogena ämnen. Markörerna använde vi sedan i den tredje studien för att utvärdera hur bra olika reningsverkstekniker är på att ta bort östrogena ämnen från avloppsvatten. Det ingående avloppsvattnet till ett kommunalt reningsverk renades med sex olika parallella tekniker och fisk exponerades för samtliga renade vattnen. Alla de testade reningsverksteknikerna minskade koncentrationerna av östrogena ämnen, men för att ta bort alla de analyserade effekterna i fisken krävdes en avancerad teknik, ozonering.

I den fjärde studien använde vi microarraymetoden för att studera vilka biologiska processer som påverkas hos fisk som exponerats för ett avloppsvatten från läkemedelsindustrier. Tidigare har vår forskargrupp visat att ett reningsverk som tar hand om processvatten från 90 olika läkemedelsindustrier i Patancheru, Indien, släpper ut mycket höga nivåer av läkemedel. Ett annat försök visade att

vattnet från reningsverket påverkar grodors tillväxt redan vid stor utspädning. Genom att exponera fisk för samma spädning av avloppsvattnet som i det ovannämnda grodförsöket kunde vi visa på att många biologiska processer som bryter ner läkemedel och andra kemikalier hade startats i fiskens lever. Fisken hade också förhöjda nivåer av fosfat och kolesterol i blodet. En förhöjd fosfatnivå i blodet hos människa kan betyda att njuren sviktar. Däremot kunde vi inom ramen för denna studie inte identifiera vilka grupper av läkemedel eller andra kemikalier som orsakar vattnets giftighet, då de vatten som fisken exponerades för innehåller tusentals olika substanser. Våra resultat kan dock utgöra en grund för fortsatt forskning inom området och det är viktigt att vi får reda på om och i så fall hur de höga koncentrationerna av läkemedelsrester kan påverka djur, växter och även människor. Många av de läkemedelsubstanser som tillverkas i Patancheru i Indien används i produkter som säljs över hela värden, inklusive Sverige. Ansvaret för att föroreningssituationen i det här området förbättras är därför också global.

Sammanfattningsvis är kunskap om vilka risker som finns med läkemedel som släpps ut i vattendrag och sjöar begränsad. I den här avhandlingen har vi visat att storskaliga molekylärbiologiska metoder kan hjälpa till att fylla de kunskapsluckor som finns, som i sin tur kan leda till förbättrat regelverk kring miljöriskbedömningen av läkemedel.

LIST OF PUBLICATIONS

This thesis is based on the following articles and manuscripts:

I Evolutionary Conservation of Human Drug Targets in Organisms used for Environmental Risk Assessments.
[*]Gunnarsson L, [*]Jauhiainen A, Kristiansson E, Nerman O, Larsson D G J
Environ Sci Technol. 2008. 42(15):5807-5813. [*] *=equal contribution*

II Sensitive and Robust Gene Expression Changes in Fish Exposed to Estrogen - a Microarray Approach.
Gunnarsson L, Kristiansson E, Förlin L, Nerman O, Larsson D G J
BMC Genomics. 2007. 8 (149).

III Comparison of six different sewage treatment technologies - reduction of estrogenic substances and gene expression changes in exposed fish.
Gunnarsson L, Adolfsson-Erici M, Björlenius B, Rutgersson C, Förlin L, Larsson D G J *Submitted.*

IV Pharmaceutical industry effluent diluted 1:500 affects global gene expression, CYP1A activity and plasma phosphate in fish.
Gunnarsson L, Kristiansson E, Rutgersson C, Sturve J, Fick J, Förlin L, Larsson D G J *Submitted.*

Parts of the introduction are based on a submitted book chapter with the title *Environmental Comparative Pharmacology: Theory and Application*, authored by Gunnarsson L, Kristiansson E and Larsson D G J*, in review*.

CONTENTS

A note on gene and protein nomenclature

Throughout the thesis the rainbow trout (*Oncorhynchus mykiss*) genes are named according to the ZFIN zebrafish nomenclature, i.e. the gene symbols of the ortholog gene in zebrafish have been used. Gene symbols are three or more lowercase letters and are in italics and correspond to the human gene name in cases of established orthology between zebrafish and human. The protein symbol is the same as the gene symbol, but non-italic and the first letter is uppercase. Human genes and proteins are named according to their official HGNC symbol, three or more uppercase letters that are italicized when referring to the gene.

1. INTRODUCTION

Pharmaceuticals are vital tools for preventing, treating or mitigating diseases, and, as such, medicines greatly contribute to the well-being of humans. Unfortunately, active pharmaceutical ingredients (APIs) can also become environmental pollutants if they reach the environment. Pharmaceuticals are specifically designed or selected to have a potent biological action and to not be too easily degraded. They are, in most cases, also able to cross cell membranes. Furthermore, the specific molecular targets of a drug can be well conserved in wildlife species, which may result in pharmaceuticals becoming potential environmental hazards, even at low concentrations. The extensive use of drugs in human and veterinary medicine is steadily increasing, and awareness of the presence and effects of pharmaceuticals in the environment has been growing since the mid 1990s.

1.1. Occurrence and emission routes

More than 150 APIs have been detected in the environment, and monitoring studies have demonstrated the presence of pharmaceutical residues in treated waste water and surface water in Sweden (Bendz et al., 2005; Zorita et al., 2009), in Europe (Thomas and Hilton, 2004; Wiegel et al., 2004; Zuccato et al., 2006), in North America (Kolpin et al., 2002) and elsewhere (for reviews, see Heberer (2002) and Kümmerer (2008)). Pharmaceutical residues are typically detected at ng/L to low μg/L concentrations in treated effluents and surface waters downstream from sewage treatment plants (STP). The objectives for municipal STPs are, for example, to remove organic substances, phosphorus and nitrogen. They are not, however, designed to remove pharmaceuticals, even if many APIs happen to be degraded or separated from the effluent by the treatment processes. There is great variability in the removal efficiency, which depends both upon the STP and the drug (Kümmerer, 2008). Certain drugs, such as ibuprofen and salicylic acid, are easily degraded and commonly have removal rates above 95%, while carbamazepine and diclofenac are considerably more persistent, with the great majority passing the treatment plants (Joss et al., 2005; Kasprzyk-Hordern et al., 2009). The removal of the synthetic estrogen, 17-α-ethinylestradiol (EE_2), varies between 64% - 100%, depending on the type of STP treatment (Ternes et al., 1999; Muller et al., 2008).

Drug residues that originate from human use end up in sewage effluents as a result of excretion and, to some extent, from inappropriate disposal of unused drugs. Veterinary drugs, on the other hand, reach the terrestrial or aquatic environment through a variety of sources and routes. The manufacture of bulk drugs has recently been identified as a point source for the release of extraordinarily high concentrations of APIs to the aquatic environment in India and China (Cui et al., 2006; Larsson et al., 2007; Li et al., 2008). The effluent from a treatment plant, which regularly receives process water from about 90 bulk drug manufacturers, contains very high levels (up to several mg/L) of a variety of drugs (Larsson et al., 2007). This treatment plant (Patancheru Environ Tech Ltd, PETL) is located in an industrial center just outside Hyderabad, India and the APIs that are produced in this area are primarily exported to the world market. Larsson and Fick (2009) showed that pharmaceutical products sold on the Swedish market contain APIs that are produced in this area. They investigated the origin of the APIs in all products on the Swedish market containing any of nine preselected substances. In 31% of the Swedish products, the API originated from manufacturers that regularly send wastewater to the PETL treatment plant.

1.2. Environmental effects

Today, we know of a few examples where the presence of APIs in the environment has been indisputably linked to adverse effects on wildlife species. Relevant, chronic toxicity data are currently lacking for most APIs, and the effects and risks of the presence of drug residues in the environment are not well understood.

The widely used analgesic and anti-inflammatory drug, diclofenac, and the synthetic estrogen in contraceptives, EE_2, are two examples of drugs present in the environment that have been clearly linked to adverse effects in wildlife species. Diclofenac caused a dramatic decline in a vulture population in India and Pakistan (Oaks et al., 2004). Diclofenac is both a human and veterinary drug, and it was extensively used to treat livestock in India and Pakistan. Vultures are natural scavengers that feed on the carcasses of wildlife and domestic livestock. The dead vultures showed signs of renal failure, which is a known side effect in mammals that are given high doses of diclofenac. Several lines of evidence, including both experimental and epidemiological studies, have helped establish the causality between diclofenac exposure and the decline in the vulture population in India and Pakistan (Oaks et al., 2004; Shultz et al., 2004; Swan et al., 2006; Cuthbert et al.,

2007; Green et al., 2007). The synthetic estrogen EE_2 is thought to strongly contribute to the feminization of fish that are found downstream of STPs. In the beginning of the 1990s, roaches (*Rutilus rutilus*) with intersex characters were observed close to municipal STPs in England, and caged fish downstream from the STPs showed strong signs of exposure to estrogenic compounds (Purdom et al., 1994). Fractionation of STP effluents suggested that natural estrogens and EE_2 were the main contributors to the estrogenic activity of STP effluents (Desbrow et al., 1998), although sewage effluents can contain a number of endocrine disrupting compounds, such as the natural hormones 17β-estradiol (E_2) and estrone (E_1) and the industrial phenols nonylphenol, octylphenol and bisphenol A. Since then, many studies have been published that provide extensive evidence for a causal link between EE_2 exposure and feminization of fish in the environment (Routledge et al., 1998; Larsson et al., 1999; Jobling et al., 2002b; Parrott and Blunt, 2005; Kidd et al., 2007).

There are a few more examples of studies suggesting that pharmaceuticals affect aquatic organisms at or around measured environmental concentrations. Laboratory experiments with the non-steroidal anti-inflammatory drug (NSAID) diclofenac show that it may affect aquatic organisms at environmentally relevant concentrations. Diclofenac is frequently detected in at concentrations up to the µg/l-level in investigations of sewage effluents and surface waters (Heberer, 2002). A concentration 1µg/L cause a variety of effects in fish, including cytological alterations in the liver, kidney and gills (Schwaiger et al., 2004; Triebskorn et al., 2004; Hoeger et al., 2005). Neuroactive drugs such as SSRIs have also been proposed to affect non-target organisms at environmentally realistic concentrations (De Lange et al., 2006; Kreke and Dietrich, 2008). The fibrate gemfibrozil may also pose a risk for aquatic organisms, as a low concentration (1.5µg/L) was reported to decrease the circulatory levels of testosterone in goldfish (*Carassius auratus*) (Mimeault et al., 2005).

The above mentioned studies all describe examples of drugs that primarily are designed to interact with human drug-targets. The release of drugs that target different parasites and bacteria can also affect wildlife species. For example, veterinary parasiticides, such as ivermectin, affect non-pest dung-feeding flies and beetles (for reviews, see Lumaret and Errouissi (2002) and Floate (2006)) . Ivermectin is also highly toxic to aquatic crustaceans (Garric et al., 2007). The release of antibiotics into the environment is also of concern (Larsson et al., 2007). Residual antibiotics can potentially affect environmental bacteria, fungi and

microalgea, and, even more alarmingly, the presence of antibiotics in surface waters might contribute to the spread of antibiotic resistance (Kümmerer, 2008). The issue will, however, not be discussed further, since the focus of this thesis has been on pharmaceuticals with human drug targets and/or the effects of pharmaceuticals on aquatic vertebrates.

1.3. Environmental risk assessment

An environmental risk assessment is defined as the procedure by which the likely or actual adverse effect of pollutants on ecosystems and their components are estimated with a known degree of certainty using scientific methodologies (van der Oost et al., 2003). The papers that this thesis is based upon are primarily focused on the assessment of effects of pharmaceuticals and to a much lesser extent on the exposure. However, the potential implications of our data on gene expression and/or conservation of drug targets on the environmental risk assessment of pharmaceuticals will be discussed throughout the thesis.

For about ten years, environmental risk assessments have been compulsory for the approval of new pharmaceutical products in the United States and the European Union (FDA, 1998; EMEA, 2006; EMEA, 2007). The risk assessment for the aquatic environment is, in principal, based upon a predicted exposure concentration (PEC) and a predicted no-effect concentration (PNEC). In the USA, the PNEC value may be based on acute responses (lethality) only, and no tests on fish are mandatory. However, the possible environmental effects of APIs are not expected to be acute and general, such as narcosis, but rather more subtle and specific, since most APIs act through specific modes of action. In addition, the no-observed effect concentration (NOEC) of the acute and chronic responses for the few investigated APIs show a very low degree of correlation (EMEA, 2006). Acute toxicity can be presented as median lethal concentration (LC_{50}) and measured LC_{50} values for EE_2 in fish are typically around 1 mg/L. Often an assessment (uncertainty) factor, usually 1,000 can be applied to the lowest acute toxicity data when chronic data is missing and such a factor is also used to extrapolate results between species. However, the PNEC value for chronic toxicity data for EE_2 in fish is below 1 ng/L which would correspond to an assessment factor of more than a million for the acute value to be protective for chronic effects in fish. In 2006, acute tests were abandoned from the European risk assessment procedures for human drugs and were replaced by chronic toxicity tests with Daphnia and algae and semichronic, early life-stage tests with fish (EMEA, 2006). Relevant and comprehensive toxicity

data are, however, lacking for APIs that were approved before 2006. Despite the new EU guidelines, there is still little focus on targeted test strategies that use the known pharmacological properties of the API in the selection of species, tests, and endpoints.

1.4. Comparative pharmacology

The molecular mechanisms behind the uptake, distribution, metabolism, excretion and pharmacological effects can be rather well conserved between the organism that the API is intended to affect and non-target organisms in the environment. Indeed, all environmental effects that have been clearly assigned to residual drugs are consistent with high affinity interactions with conserved drug targets in the affected wildlife species rather than with a general toxic effect. The vast knowledge base that has been derived during the development of a drug in humans and other mammalian models can therefore provide a basis for an expanded understanding of the potential action of residual pharmaceuticals in exposed wildlife species. It has been suggested by several authors that such information can be used to develop more efficient test strategies (Seiler, 2002; Ankley et al., 2007; Kostich and Lazorchak, 2008).

The physiochemical, pharmacological and toxicological properties of an API is extensively studied during the development of a new drug. For most approved pharmaceuticals, these kinds of data are easily accessible in different public databases. The most frequently used drug databases in this thesis are listed below:

- *The **DrugBank** database (http://www.drugbank.ca/) is an example of a bioinformatics and chemoinformatics resource that combines detailed drug data with metabolizing enzyme and drug target information for the majority of all FDA approved drugs.*
- ***KEGG DRUG** (http://www.genome.jp/kegg/drug/) is an information resource for all approved drugs in Japan and the USA based upon the drugs' chemical structures. It also contains information about drug targets and pathways.*
- ***The Pharmacogenetic and Pharmacogenomics Knowledge base** (PharmGKB; http://www.pharmgkb.org/) has information about the relationships between drugs, their affected pathways and the genes therein, although it contains information on only a few drugs and drug targets. It*

aims to aid researchers in understanding how genetic variation among individuals contributes to differences in their reactions to drugs.

- ***RxList*** *(http://www.rxlist.com/script/main/hp.asp) is a comprehensive drug information database that aims to assist and support clinical decisions.*

However, to make well-founded predictions, a comprehensive understanding of the physiology of the exposed wildlife species is equally important, and this is currently a hampering factor for the vast majority of species. A few comparative pharmacology reviews that focus on a certain drug and/or a group of species have been published. The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) has published a review on the use of intelligent test strategies in ecotoxicology (ECETOC, 2007). The ECETOC's review demonstrates how information about the mode-of-action for specifically acting chemicals could be used in the environmental risk assessment. Many of the examples and case studies include pharmaceuticals. Other reviews that have focused on the comparative pharmacology in fish for selective serotonin reuptake inhibitors (SSRIs) (Kreke and Dietrich, 2008) and adrenoreceptor antagonists (beta-blockers) (Owen et al., 2007) have been published. Brain et al. (2008) reviewed the effects and risks of exposure to pharmaceuticals in aquatic plants. A number of target sites for antibiotic drugs are evolutionarily conserved in plants because their plastid organelles have bacterial ancestry. The statin type of blood lipid regulators belongs to a group of pharmaceuticals that have a human target, which is also conserved in plants. Measurements of the downstream metabolites (sterols) of the target enzyme (HMG-CoA reductase) provided a sensitive endpoint in the aquatic plant *Lemna gibba*. Apart from antibiotics and statins, there are few other classes of pharmaceuticals that we know exert a strong toxicity in plants. The different reviews all conclude that future toxicological testing should encompass and reflect the known pharmacological effects of the substances studied and should, therefore, focus more strongly on specific molecular targets.

The number of attempts to perform comparative pharmacology in a large scale manner are few (Huggett et al., 2003; Kostich and Lazorchak, 2008). The recent advances in DNA sequence analysis, and the characterization of genomes have created new possibilities to advance the field of comparative pharmacology with an ecotoxicological focus. Information that is gained from deeper comparative efforts has the potential to aid in prioritizing which drugs need further attention for assessment of their environmental risks, and which organisms could be prioritized for testing, and appropriate endpoints.

1.5. Genomics

The increasing number of sequencing efforts of environmentally relevant species has enabled the use of different genomic approaches in the assessment of the environmental risks of pharmaceuticals.

Since there is no straightforward definition of *genomics*, the word, in this thesis, will be used to illustrate large scale methods to study the genome, its different building-blocks and their functions. The mouse geneticist Thomas Roderick is believed to have coined the term genomics to describe an approach to study DNA at the level of chromosomes, entire genomes or large clusters of genes (McKusick and Ruddle, 1987). The purpose was to distinguish traditional genetic methods that focused on one gene or a family of related genes from more large scale methods. Another common definition of genomics is "*the study of genes and their function*" (Hocquette, 2005).

There are a multitude of advances that have been made in molecular biology, statistics and computer science that led to the term genomics being coined. One important step forward was the development of the chain termination method of DNA sequencing, also known as the Sanger method (Sanger et al., 1977b). In 1977 Frederick Sanger used this technique to, for the first time, fully sequence a genome (Phage Φ-X174) (Sanger et al., 1977a). Since then, there has been rapid progress in automated sequencing methods. Today, hundreds of eukaryotic genomes have been extensively sequenced (Peregrin-Alvarez and Parkinson, 2007) (http://genome.jgi-psf.org/euk_home.html and http://www.ensembl.org/index.html), and the majority of the sequenced vertebrate species are mammals (http://www.ncbi.nlm.nih.gov/genomes/static/gpstat.html). The sequencing efforts for species that represent more ecotoxicologically relevant taxonomic groups were, until very recently, rather few. However, in the last few years, several environmentally relevant species have become fully sequenced, including the green alga (*Chlamydomonas reinhardtii*) (Merchant et al., 2007), the water flea (*Daphnia pulex*) (http://genome.jgi-psf.org/Dappu1/Dappu1.home.html) and the stickleback (*Gasterosteus aculeatus*) (http://www.ensembl.org/Gasterosteus_aculeatus/Info/Index), and there are numerous additional, extensive EST libraries that are available for other species, such as the rainbow trout (*Oncorhynchus mykiss*) (http://compbio.dfci.harvard.edu/tgi/). This has opened up new possibilities to perform microarray analysis and large scale bioinformatic comparisons to assess the environmental risks of pharmaceuticals.

1.5.1. DNA Microarrays

The modes of action for many APIs in mammalian species are well known, and this information could be used to create hypotheses of potential effects in wildlife species. However, interactions with unexpected targets are possible, and other endpoints, which are different from the mammalian ones, could be more sensitive. In addition, even if a drug target is evolutionary well conserved, the target protein might have different biological functions in different organisms. Microarrays provide an efficient tool to study thousands of potential gene responses simultaneously and thus may be used to find sensitive responses to an exposure as well as information about the mode of action.

DNA microarray technology was introduced in the mid 1990s. The first use of a spotted cDNA array was published in 1995 (Schena et al., 1995), and the first use of an oligonucleotide array was published in 1997 (Wodicka et al., 1997). In 1999, the journal Nature Genetics dedicated a whole supplement to this up and coming technology (Cohen, 1999). Since then, microarrays have been successfully applied to various areas within biology, such as the categorization of cell cycle genes in yeast (Spellman et al., 1998), the classification of breast cancer (Sorlie et al., 2001) and toxicant profiling (Waring et al., 2002).

Several commercial array platforms are available through companies such as Affymetrix (http://www.affymetrix.com) and Agilent (http://www.agilent.com). The strengths of commercial platforms are their readymade protocols and support, as well as their overall high quality. Few commercial array platforms are, however, available for environmentally relevant species. Non-commercial arrays have been developed by academia, such as a small spotted cDNA array for the European flounder (*Platichthus flesus*) (Williams et al., 2003), different cDNA arrays for salmonid species (Rise et al., 2004; von Schalburg et al., 2005), a cDNA array for *Daphnia magna* (Soetaert et al., 2006), an oligonucleotide microarray for the fathead minnow (*Pimephales promelas*) (Larkin et al., 2007) and an oligonucleotide microarray for the eelpout (*Zoarces viviparus*) (Kristiansson et al., 2009a). The usage of microarrays to generate or test an ecotoxicological hypothesis is a rather new concept, as shown in Figure 1.

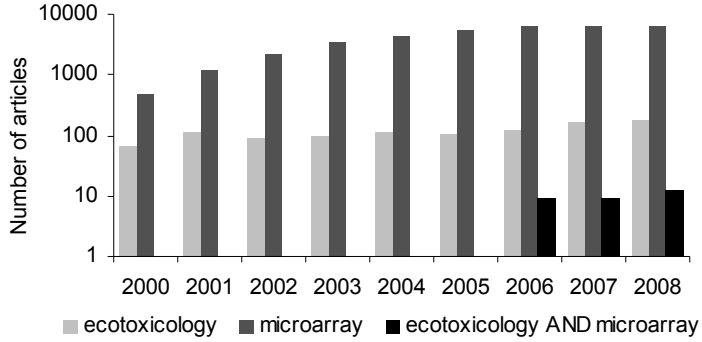

Figure 1. The number of articles with the different topic words ecotoxicology, microarray, or ecotoxicology and microarray that are present in the Science Citation Index Expanded (SCI-EXPANDED) database at the ISI Web of Science.

Several articles that describe the potential of integrating genomics and microarrays into ecotoxicology have been published (Snape et al., 2004; Ankley et al., 2006; Ju et al., 2006; Lettieri, 2006; Steinberg et al., 2008). A microarray analysis could assist in the identification of potential biomarkers of exposure or effect and increase the understanding of the (toxicological) mode of action of chemicals or mixtures. Microarray analysis could also be used to categorize individual chemicals and mixtures into different classes for their modes of action. Furthermore, the obtained gene response patterns might aid in the identification of substances that are responsible for the toxic effects that are caused by a mixture.

1.6. Biomarkers

A biomarker can be defined as "*a biological response to a chemical or chemicals which gives a measure of exposure and sometimes, also, of toxic effect*" (Peakall and Walker, 1994). The biological response can range from molecular to cellular and physiological responses and to behavioral changes. According to WHO (1993), biomarkers can be subdivided into three classes, which are the biomarkers of exposure, effect and susceptibility. The definitions of biomarkers of exposure and effect are not clearly separated. Biomarkers of exposure can be used to confirm and assess the exposure to a particular substance or to a group of substances. Biomarkers of effect can be used to document either activational, reversible alterations or more permanent adverse effects. This could become important when relevant chemical measurements cannot be achieved or full life cycle tests are not feasible. Biomarkers of susceptibility can be genetic differences that can explain

variations in the degree of response to toxicant exposure that are observed between different individuals or species (see section 4.1.1.).

The induction of vitellogenin mRNA (*vtg*) or protein (Vtg) in male and juvenile fish has been established as a commonly used biomarker for estrogen exposure. Vtg is a precursor of yolk proteins in oviparous vertebrates. It is usually produced in the liver of sexually maturing female fish and transported to the ovaries via the circulation. Endogenous estrogen regulates its expression and it is normally not expressed in male or juvenile fish. However, Vtg can be induced in male and juvenile fish if they are exposed to estrogen or estrogen-like chemicals (Sumpter and Jobling, 1995). The discovery that EE_2 is an important contributor to the endocrine disruption that is observed in fish downstream from municipal STPs was greatly facilitated by the use of Vtg as a biomarker (Purdom et al., 1994; Sumpter and Jobling, 1995; Desbrow et al., 1998; Routledge et al., 1998; Larsson et al., 1999; Jobling et al., 2002b). However, it has been shown that estrogens can have an organizational, permanent effect, such as on gonadal sex differentiation, at a concentration that is not sufficient to give rise to a measurable Vtg response (Örn et al., 2003). The induction of zona radiata proteins has been suggested to be more sensitive biomarkers of estrogen exposure (Thomas-Jones et al., 2003), but their specificity for estrogens is not as clear (Larsson et al., 2002; Berg et al., 2004).

In paper II, III and IV, we have focused on the identification of gene expression changes that can be used as biomarkers of exposure and/or effect. Good biomarkers are sensitive, and the response should preferably also correlate with the magnitude of the exposure. It is also desirable that the biomarker is specific for the substance or group of substances of concern (or the adverse effect of concern). Likewise, it should preferably be robust when considering both different exposure scenarios and different measuring techniques. The regulation of a single gene rarely fulfills all the above mentioned criteria, but a combination of responses could together potentially increase the degree of evidence.

Sometimes the questions are complicated and the answers are simple
Dr.Seuss

2. SCIENTIFIC AIMS

There is growing concern about the environmental effects of pharmaceuticals. Many different APIs are detected in the environment, but their effects are not well understood. We have addressed this issue by using experimental and computational genomics techniques.

The aims of this thesis have been:

- To investigate and exemplify how the use of available mammalian pharmacological data can aid in the assessment of ecotoxicological risks that are posed by pharmaceuticals.

- To explore how exposure to a single pharmaceutical, as well as exposure to a complex mixture of APIs and other chemicals, affects the global hepatic gene expression pattern in fish, with a special focus on the identification of biomarkers of exposure.

- To evaluate the abilities of different sewage treatment technologies to reduce estrogenic substances and gene expression changes in exposed fish.

3. METHODOLOGICAL CONSIDERATIONS

3.1. Fish exposures

All exposure experiments that are described in this thesis were performed with farmed rainbow trout under controlled laboratory conditions. Rainbow trout is a salmonid species, the basic physiology of which is rather well understood. An extensive expressed sequence tag (EST) library for this species is also available at the DFCI rainbow trout gene index database (http://compbio.dfci.harvard.edu/tgi/). Trout is easily obtained and thrives well in both laboratory conditions and in Swedish fresh water. Our group often exposes rainbow trout in the field to various contaminant sources (e.g., caged in a river), and it was important that microarrays were developed for a species that would tolerate Swedish waters and temperatures.

The abundance of thousands of mRNAs can be simultaneously measured in a microarray experiment but the number of biological replicates is usually small (less than 10) due to the relatively high costs of the analyses. The identification of genes with an altered expression due to differences in exposure is therefore particularly sensitive to both technical and biological variation. Controlled laboratory experiments, with well-acclimatized farmed fish that were not fed during the exposure, are examples of the efforts that were taken in order to minimize the biological differences between individuals and aquaria. Rainbow trout cope well without food for two weeks, and longer exposure time periods allow chemicals to bioconcentrate and approach a steady state between the concentrations in the fish and the water. In papers II and III, the fish were exposed in a flow-through system for two weeks. In paper IV, the fish were exposed for five days due to a limited amount of effluent.

The liver was used for gene expression analysis in all three studies. The liver is the major detoxifying organ and many pharmaceuticals, including EE_2, affects the hepatic gene expression pattern. Effects on mRNA abundance in several organs, at different time points and at different exposure concentrations would have given a more comprehensive picture but such studies were economically not feasible. In paper II, the aim was to find sensitive responses and therefore we exposed the fish to a low, environmentally relevant concentration and a high concentration. The high concentration facilitated the identification of the more subtle responses that were

induced by the low concentration. In paper IV the aim was to increase the understanding of the effluent's toxicological mode(s) of action. A previous study had shown that 0.2% of the effluent strongly reduced the growth of tadpoles (Carlsson et al., 2009). This dilution was therefore used and the limited amount of effluent decided the duration of the exposure.

3.2. Sequence analysis and annotation assignment

As mentioned above, an extensive EST library is available for rainbow trout, but relatively few genes are fully sequenced and characterized in trout. The transcripts, ESTs and corresponding probes that were used in this thesis have been functionally annotated using different automated methods for sequence analysis. The basic local alignment search tool (BLAST) (Altschul et al., 1997) is a method to search and compare protein and DNA sequences against query databases i.e. BLAST searches for subsequences in a database that match subsequences of the query transcript/protein by optimizing a local similarity measure. Functionally similar genes and proteins can be linked by identifying the best BLAST hit with an E-value cut-off. ESTs are typically annotated using a tblastx (comparison of a translated nucleic acid query versus a translated nucleic acid database) with a non-stringent E-value cut-off, such as 10^{-5}. This was the approach that was taken to annotate the transcripts of the probes on the array that was used in paper IV. Different BLAST strategies were also used in paper II, both to cluster the ESTs that were available on the salmonid cDNA microarray and to link functionally similar genes between the different species that were used in the meta-analysis.

3.2.1. Ortholog prediction

In paper I, a more sophisticated BLAST-based method was used to predict functionally similar proteins between species. One way to link functionally similar proteins is to look at the evolutionary relationship between proteins. Homologous proteins share a common ancestry and can be further characterized as orthologs or paralogs (Figure 2). Orthologs are homologous proteins that exist in different species as a result of speciation. Orthologous proteins in different species generally, but not necessarily, retain the same function. Paralogs, on the other hand, originate from gene duplication events within a genome and can be further divided into in-paralogs and out-paralogs, depending on whether the duplication event occurred before or after a speciation event. Although paralogous proteins often retain similar biochemical functions, they generally diverge after the duplication event.

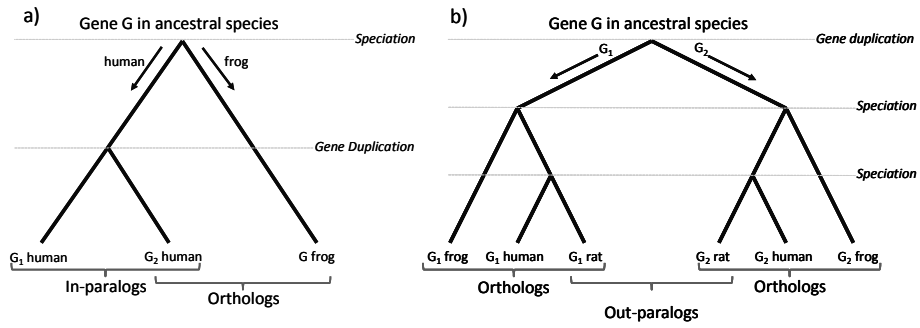

Figure 2. Homologous proteins have a shared common ancestry. In (a), an ancestral gene G undergoes one speciation event, which is followed by a gene duplication event in human. The different variants of G that appear in human and frog are called orthologs (G frog and G_1 or G_2 human). The different variants within each species are called paralogs. Paralogs that appear as a result of a gene duplication event that occurred after the latest speciation event are called recent or in-paralogs (G_1 human and G_2 human). In (b), the ancestral gene G undergoes one gene duplication event followed by two speciation events. The different variants (G_1 frog, G_1 human and G_1 rat) are referred to as orthologs while; for example, G_1 frog and G_2 rat are not orthologs. The paralogs in (b) are called ancient or out-paralogs.

There are many different computational approaches to identify homologous proteins between organisms, but two basic schemes underlie these different approaches: phylogenetic analysis and BLAST-based methods. Phylogenetics is the study of evolutionary relatedness based upon molecular sequence data. A phylogenetic analysis is a computational approach that uses phylogenetic trees to infer groups of orthologous proteins. To build phylogenetic trees, high quality multiple sequence alignments are usually required, and the methods are often computationally expensive and generally seem to also require manual intervention (Dolinski and Botstein, 2007).

BLAST-based approaches require much less computation and allow for full automation. In its simplest form, a one-way BLAST is performed, and a sequence similarity above a pre-defined threshold defines two proteins as orthologs. For example, a BLASTp is performed for a query protein against a protein database, such as Uniprot (http://www.uniprot.org/). However, one-way sequence comparisons often create many false positives. Chen et al. (2007) showed that a one-way BLAST strategy can have as much as 50% false positives and 4% false negatives in their assignment of the performance of different orthology detection strategies. The one-way BLAST method cannot distinguish between in-paralogs and out-paralogs and assigns them inappropriately as ortholog pairs, which leads to many false positives. In contrast, a reciprocal best BLAST hit strategy displays few false positives (estimated at 8%), but many false negatives (estimated at 30%)

(Chen et al., 2007). This strategy cannot recognize many-to-many or many-to-one ortholog relationships, which results in a high false negative error rate. Despite this problem, it is often suitable as a first step for advanced BLAST-based ortholog detection methods.

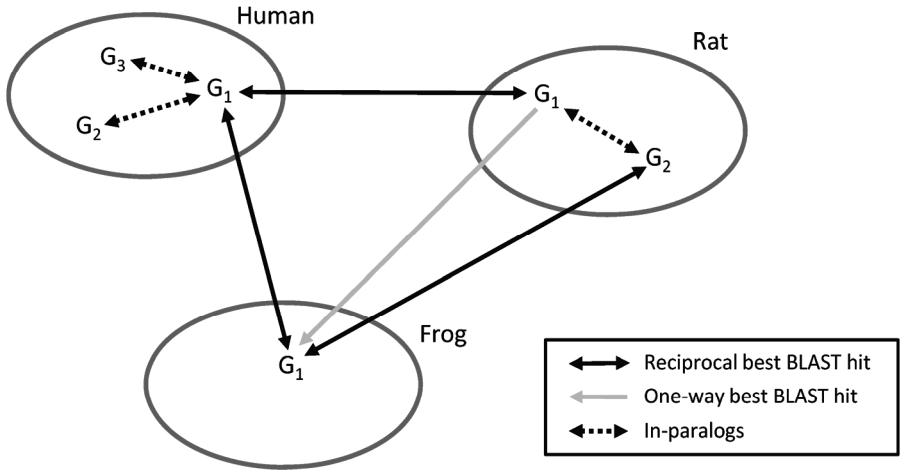

Figure 3. Identification of both paralogs and orthologs is necessary to understand the evolutionary history of a gene. The figure shows an example of an ancestral gene G, which appears as only one gene in frog, while rat has two in-paralogous genes and human has three. Since the reciprocal best BLAST hit (RBH) of G_1 in frog is G_1 in human and G_2 in rat, while the RBH of G_1 in human is G_1 in rat, algorithms that solely use information from the RBHs would not identify orthologs for G in all three species. However, the clustering approach utilized by OrthoMCL can detect all these non-trivial relationships and correctly assign all variants within a single cluster that contains both the in-paralogs and the orthologs.

The OrthoMCL algorithm (Li et al., 2003), which was used in paper I, is an example of a BLAST-based method that aims to separate the in-paralogs from the out-paralogs. The OrthoMCL algorithm distinguishes between the two types of paralogs by first using all-against-all BLAST comparisons, both within and between species. The result is interpreted as a graph where the proteins are nodes and the weighted edges are their sequence similarity. This graph is then partitioned into sub-graphs using the Markov Cluster (MCL) algorithm (Enright et al., 2002). In the final result, the ortholog groups contain only those paralogs that are more similar to each other than to any other sequence from the other species. The algorithm was estimated to perform well (16% false positives and 7% false negatives) on a divergent set of eukaryotic species (Chen et al., 2007).

There are other methods and considerations that can further improve the accuracy of orthology predictions. For example, considerations of synteny, which

are the evolutionarily preserved chromosomal positions of genes, can be added to the predictions in order to further improve the approach. The NCBI Homologene database (http://www.ncbi.nlm.nih.gov/homologene) combines BLAST-based sequence similarity measurements with synteny information. For comprehensive reviews of existing approaches for ortholog prediction, please refer to Chen et al. (2007) and Dolinski and Botstein (2007).

3.2.2. Functional annotation and characterization of genes

Different bioinformatic resources that aim to aid in the biological analysis of microarray data are available. In papers I and IV, Gene Ontology (GO) (Ashburner et al., 2000; The Gene Ontology Consortium, 2008) analysis was used. The GO project is a collaborative effort for the controlled and consistent description of protein attributes. GO has three main ontologies that describe molecular functions, biological processes and cellular components. The GO project is an important resource that describes proteins in a systematic way, but the management and merging of ontologies (e.g., between species) is a complex task and an ongoing research area.

Pathway analysis is another method that can be used to understand and interpret gene expression data. Many proteins in mammalian model species have been associated with one or more pathways, and there are several ongoing efforts to systematically characterize pathways, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2008).

Both pathway and GO analysis can be performed in order to test whether genes from a given pathway or genes with a certain GO term have higher tendency to be differentially expressed than other genes on a microarray. The procedure is usually repeated for all pathways or GO terms in a database, and the models best describing the gene expression pattern can thus be identified. In paper IV, different top lists of the most significantly regulated genes were analyzed for overrepresentation of GO terms and KEGG pathways.

Few rainbow trout proteins have GO annotations assigned to them, and, therefore, GO annotations that have been assigned to orthologous proteins in reference species need to be used instead. For example, the estrogen receptor in rainbow trout does not have any GO assignments in the database AmiGO (http://amigo.geneontology.org/cgi-bin/amigo/go.cgi), but the human estrogen receptor 1 (ESR1) has 11 GO term associations, which include four biological process terms (GO:0030520: estrogen receptor signaling pathway, GO:0048386:

positive regulation of retinoic acid receptor signaling pathway, GO:0006355: regulation of transcription and DNA-dependent, GO:0007165: signal transduction). On the other hand, the zebrafish (*Danio rerio*) Esr1 has only three GO term associations, which, are all biological process terms (GO:0042221: response to chemical stimulus, GO:0043627: response to estrogen stimulus and GO:0009410: response to xenobiotic stimulus). To my knowledge, all the biological process terms mentioned above describe the estrogen receptor 1 both in humans and in zebrafish, but the discrepancy between the annotations implies that the output from a rainbow trout GO term analysis will heavily depend on the chosen reference species.

Since rather few proteins have been functionally characterized in rainbow trout (181 reviewed proteins at http://www.uniprot.org/), the function of the majority of the proteins is described via high sequence similarities to proteins that have been characterized in other species. There are many challenges in the correct assignment of orthologous proteins in divergent species, and, even if the assignment is correct from an evolutionary perspective, their functions might still differ. Another weakness with automated annotations is that genes are annotated according to already discovered functions. Accordingly, biological processes that are implicated in a microarray experiment seem to confirm the functions of proteins and are seldom used to discover new pathways or protein functions.

3.3. DNA microarray experiments

3.3.1. General

The microarray technology enables measurement of the abundance of several thousands of transcripts simultaneously. Simplified, total RNA is isolated from the animal/tissues/cells of interest, and the mRNA in the sample is converted into fluorescently-labeled complementary DNA (cDNA) or a type of modified amplified RNA (aRNA). A hybridization solution, which contains the cDNA, is then incubated on a surface, such as a glass slide, that contains thousands of regularly spaced probes. Finally, the unbound cDNA is washed away, and the amount of hybridized cDNA is visualized, typically by a laser scanner or a CCD camera (Figure 4). The spots containing the different probes are localized, and intensity data are extracted from the fluorescent images with software for image analysis (for this purpose, ImaGene (http://www.biodiscovery.com/index/imagene) was used in paper II and Geniom Wizard (http://www.febit.com/go/en/products/geniom-rt-analyzer) in paper IV).

Several technological variants of microarray chips exist, and the probes can be of different length and origin. A microarray that uses cDNAs as probes is typically created by spotting amplified clones, which are a few hundred nucleotides long, from cDNA libraries onto a glass slide. On a spotted cDNA microarray, neither the length of the probe sequence nor the amount of signal per transcript is standardized. Oligonucleotide arrays, on the other hand, contain probes that are either synthesized from the ground up while attached to a surface, such as those produced by Affymetrix and febit, or pre-synthesized and then spotted onto a glass slide. The oligonucleotide probes are typically between 25 and 70 bases long, and the design is done *in silico*. Spotted microarrays are typically of the two-channel type, in which the cDNA from two samples are labeled differently and hybridized to the same array, while *in situ* synthesized arrays are of the one-channel type. In general, the results from oligonucleotide arrays compare more favorably to qPCR results than the results from cDNA microarrays (Woo et al., 2004; Kuo et al., 2006).

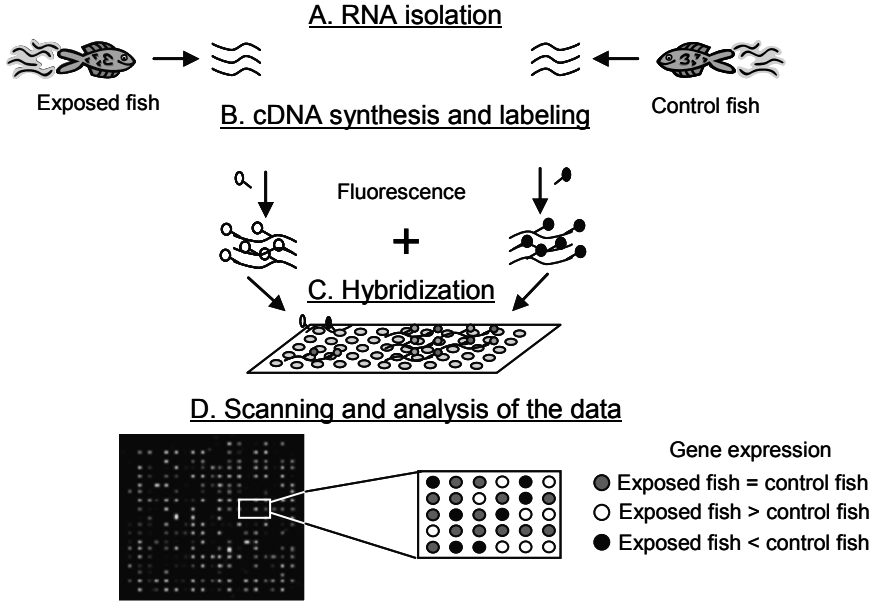

Figure 4. The principle of microarray analysis for spotted cDNA microarrays, and the experimental set-up used in paper II. Total RNA was isolated from control and exposed fish, cDNA was synthesized and labeled with the different fluorescent dyes cyanine 3 (Cy3) and 5 (Cy5). Differently labeled cDNA from one control fish and one exposed fish was hybridized to the same microarray. See also the left microarray picture on the cover of the thesis.

3.3.2. The 16k salmonid cDNA microarray

The spotted cDNA microarray used in paper II was developed by a Canadian research group belonging to the Consortium for Genomic Research on All Salmon Project (GRASP) (von Schalburg et al., 2005). The array contained 16,006 cDNAs from EST libraries that were sequenced from different tissues from the Atlantic salmon (*Salmon salar*) and the rainbow trout, in addition to a few cDNAs from other salmonid species. The long probes (~400 base pairs) allow some cross hybridization to occur which enables cross species usage. The principle of the microarray experiment in paper I is shown in Figure 4.

When the GRASP microarrays were purchased in 2004, the Canadian group had just started to re-evaluate its cDNA synthesis kit and hybridization strategy and, therefore, did not have any protocol to recommend. We decided to use Invitrogen's SuperScript Indirect cDNA labeling System kit for cDNA synthesis and labeling and tried out different hybridization buffers and wash protocols. No matter how stringent the hybridization buffer or protocol was, we were not able to solve problems with apparent cross hybridization (low specificity). The later addition of a

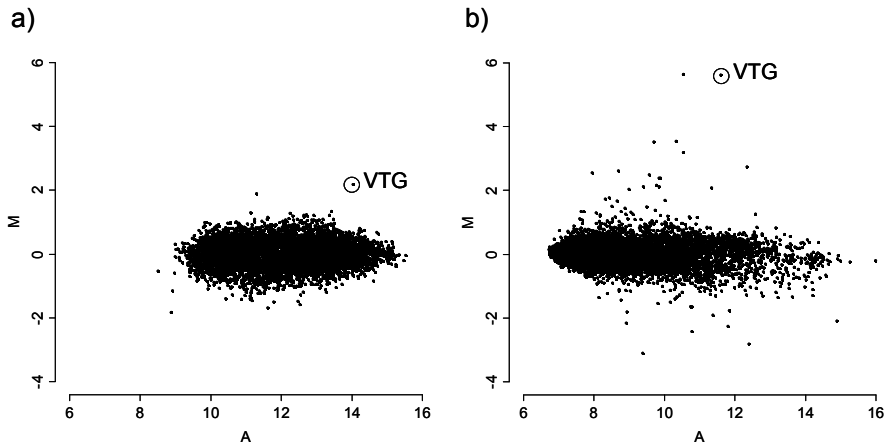

Figure 5. MA plots showing normalized data from microarray experiments with fish exposed to a high concentration of 17-α-ethinylestradiol (EE_2) (R) and control fish (G). The log ratios (M = $\log_2 R/G$) are plotted on the y axis against the log of the geometric mean of the signal strength (A = $(\log_2 R + \log_2 G)/2$) for each spot on the chip. a) Data from a non-stringent hybridization where LNA dT blocker has not been added to the hybridization buffer are compared to b) data from the optimized hybridization protocol used in paper II. *Vtg* is expected to be up-regulated several hundredfold in the EE_2 exposed fish. The absolute fold regulation of *vtg* is roughly 4 in plot a, while it is almost 60 in plot b. Notice also the differences in A values between the two plots. Many of the transcripts that are represented on the GRASP array are not expected to be expressed in liver tissue, therefore, their corresponding cDNA spot should have an A-value below 8.

locked nucleic acid (LNA) dT blocker to the hybridization buffer was found to be absolutely crucial for a stringent hybridization (Figure 5). The LNA dT blocker is designed to block poly (dA) sequences that are present in the cDNAs spotted on the array and prevent them from hybridizing to poly (dT) segments of the labeled targets. With the optimized hybridization protocol, several sensitive estrogen responsive genes could be identified in paper II.

Nevertheless, we were not completely satisfied with the platform for several reasons. The labeling and hybridization methods typically used for spotted cDNA microarrays in general are long and labor intensive and, therefore, easily accumulate a high technical variability. In addition, the selection of the ESTs for the microarray was not entirely satisfactory. There were also major problems with the red channel (Cy5) during the late spring and summer of 2005 and 2006. The signal to background ratio was usually above four in both the red (Cy5) and green (Cy3) channel on our scanned microarrays. However, in May 2006, the signal to background ratio in the red channel decreased below three. One reason for this might have been environmental ozone that rapidly can degrade Cy5 while Cy3 is much less affected (Fare et al., 2003; Branham et al., 2007). The level of tropospheric ozone, which is generated from car exhaust and factory emissions that are exposed to sunlight (http://www.naturvardsverket.se/sv/Tillstandet-i-miljon/Luftkvalitet/Marknara-ozon/), increases during spring and summer in Sweden, since high temperatures increase its generation. The highest ozone levels are usually measured during the afternoon on warm spring and early summer days, and the measured levels in Gothenburg during May 2006 (> 40ppb) were well above the levels shown to degrade Cy5 (Fare et al., 2003; Branham et al., 2007).

Regardless of the problem with Cy5, spotted cDNA arrays in general show less agreement with quantitative PCR data compared to oligo-based arrays (Woo et al., 2004; Kuo et al., 2006). During the autumn of 2006, we had the opportunity to start using the highly flexible oligonucleotide-based platform Geniom (Baum et al., 2003) from febit (www.febit.com) with *in situ* synthesized probes and several automated steps in the hybridization and washing process.

3.3.3. The Geniom oligonucleotide microarray

Geniom opened up new possibilities to determine which transcripts and probes to put on the array and to also create microarrays for different species. The technical variability could be reduced, and the probe design could be optimized.

We started with an evaluation project using febit's own probe design (25bases) and recommended hybridization and washing conditions. The expression levels of hepatic estrogen-regulated genes were compared between female pools and male pools of zebrafish, in which the regulation of six genes was measured by quantitative PCR (qPCR). The obtained microarray data unfortunately had low consistency with the qPCR data. Therefore, we implemented our own probe design procedure. We also had the opportunity to use pools of total RNA from mouse, in which the expression of 165 genes was already characterized by qPCR (Kuo et al., 2006). These pools had previously also been used to compare gene expression measurements across different microarray platforms. With our own probe design procedure, using OligoArray 2.1 (Rouillard et al., 2003) and optimized hybridization and washing conditions, Geniom performed as well as the best commercial platforms (Figure 6). The optimized probe design procedure and hybridization and wash conditions were used in paper IV (see also the right microarray picture on the cover of the thesis).

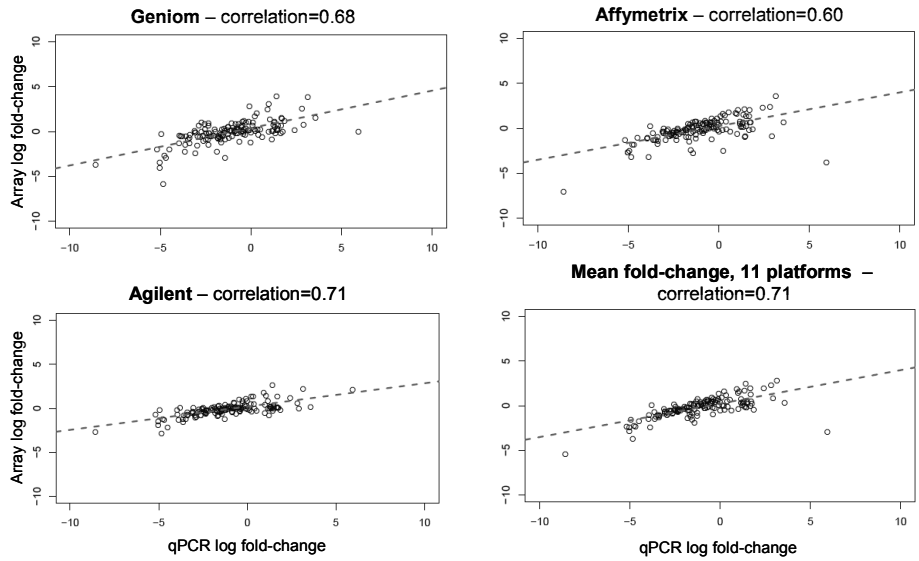

Figure 6. Pre-processed microarray data from 11 platforms compared to qPCR results for 165 genes (Kuo et al., 2006; Kristiansson et al., 2009b). Geniom performs as well as the commercial platforms Agilent and Affymetrix. The correlation is calculated according to the Pearson's measure (Kristiansson et al., 2009b).

3.3.4. Microarray data processing and analysis

Microarray data need to be processed and statistically analyzed before a biological interpretation of the results can be performed. The intensity data that are obtained from the image analysis (see section 3.3.1.) need to be processed in order to remove technical artifacts and systematic errors. The processing of microarray data typically involves background correction and normalization. Background correction removes any intensity that stems from wash residue and/or noise from the scanning process. Subtraction is the most commonly used method, where the background is simply subtracted from the intensity in the spot. Several background correction methods have been compared, and the performance of subtraction was inferior to all other methods that were investigated, including using no background correction at all (Ritchie et al., 2007). No background correction was performed on the array data that were presented in papers II and IV. Next, the data are normalized to enable comparison between the intensity signals from different samples. Intensity-dependent trends due to unequal quantities of labeled cDNA, different efficiencies in the labeling process, different properties of the dye and other systematic errors and biases can be minimized by applying proper normalization algorithms (Kristiansson, 2007). Data from two-channel microarrays (differently labeled cDNAs hybridized to the same array) need within-array normalization. In paper I, the loess normalization method (Yang et al., 2002) was used for this purpose. Loess normalization is the most commonly used within-array normalization method. The algorithm fits a robust weighted regression line to the MA-plot, and the M-values are then centered around zero. In paper IV, we used a one channel microarray, and the quantile-quantile algorithm (Bolstad et al., 2003) was used for normalization.

Statistical analysis is typically performed to identify regulated genes. The moderated t-statistic (Smyth, 2004) was used in both papers II and IV. In a microarray analysis, a high number of genes are studied in parallel and generally, there are few replicates (less than ten). The estimation of the variance can therefore become unreliable and standard statistical methods, such as t-tests, can perform inadequately. Thus, different methods moderating the variance have been suggested. Today, penalized and moderated methods are preferred when analyzing microarray datasets (Kristiansson, 2007; Sjögren et al., 2007). The list of potential differentially expressed genes can still contain false positives, since thousands of responses are tested simultaneously. The threshold for the p-value is therefore commonly corrected, and one popular way is to calculate a false discovery rate (FDR) (Dudoit et al., 2003). For example, by setting the FDR to 0.2, the list of

potentially regulated genes is estimated to contain no more than 20% false positives. However, depending on the downstream analysis, different numbers of false positives and, thus, false negatives can be allowed in the input data. If all genes on a list were to be considered differentially regulated, few false positives should be allowed, while a longer list with more false positives can be used in, for example, pathway or GO analysis (also discussed in section 3.2.2.). The pathway or GO analysis can accommodate a larger number of false positives, since the "falsely regulated genes" are less prone to have the same GO annotation or be involved in the same pathway.

In general, other methods are needed to quantify the real difference of the mRNA abundance of genes that are identified as differentially regulated by microarray experiments. Cross-hybridization and other artifacts are problematic issues which, to date, have not been completely solved. Also, the dynamic range of microarrays is small, which requires that other methods needs to be applied if a more exact quantification is crucial. In the future, ultra high-throughput sequencing methods might replace microarray analysis, although costs need to be significantly reduced. Sequencing methods have the advantage of elucidating the more exact nucleotide content of the target cDNA sequences, and, with the ultra high-throughput massively parallel sequencing technology, a vast number of sequence reads (several hundreds of thousands of single reads per run) can be generated relatively inexpensively (Shendure and Ji, 2008; Kristiansson et al., 2009a).

3.3.5. Principal limitations of gene expression analysis

A rather large number of the rainbow trout ESTs are not similar to any known proteins and, therefore, do not have any annotation. As mentioned in section 3.2.2, the fish genomes are not as extensively studied as the mammalian genomes, and the presence of unique genes is plausible. However, poly-adenylated non-coding transcripts are expressed at low levels and there is a growing understanding of the importance of translational/transcriptional regulation by for example microRNAs (Kapranov et al., 2002; Bartel, 2009). This implies that measured differences on a microarray might not correspond to differences in levels of protein coding transcripts.

It is also important to point out that differences in the abundance of protein coding transcripts do not necessary imply differences in protein levels. In general, however, protein expression levels do correlate, if only moderately, with transcript abundance (de Hoog and Mann, 2004; Lu et al., 2007).

Data from microarray experiments is primarily used for hypothesis generation. If the analysis, for example, is performed with the aim to find potential biomarkers, the usefulness of the identified biomarker needs to be tested. If the microarray analysis is instead performed with the aim to obtain information about (toxicological) mode of action, the created hypothesis might need to be validated at a higher level of biological organization. Furthermore, the gene expression pattern for the same chemical exposure can appear very different depending on the phase of the organism's stress response (alarm, resistance or exhaustion phase) (Steinberg et al., 2008). It is, therefore, important to have an experimental design that suits the aim of the study.

3.4. Quantitative PCR

Quantitative PCR is one of the most sensitive methods for the detection and quantification of mRNAs, and it is often used for the validation of microarray data. It requires an optimized assay that is characterized by well-designed primers that amplify only the desired product, an efficient amplification (95-105% efficiency) and good reference gene(s). In papers II and IV, qPCR was used to confirm a subset of the microarray results, and, in paper III, the method was used to measure the gene expression of estrogen-regulated genes. Quantitative PCR analysis typically requires consistently expressed reference gene(s) as an internal control for differences in the amount of starting material (cDNA) between the samples. An optimal reference gene should have 1) a low variability across different biological samples (e.g., between different individuals and subtypes of normal tissue), 2) a level of expression that is not too far from the expression level of the regulated gene of interest and 3) consistency between the different treatments. Normalization can involve either just one or several reference genes. In paper II, the two reference genes that were selected were *β-actin* (*bactin*), which was a frequently used housekeeping gene at the time, and *ubiquitin* (*uba*), which was selected based upon literature data and our own microarray results. *Bactin* had a clear tendency to be regulated in the fish that were exposed to the high EE_2 concentration, while *uba* fulfilled the abovementioned criteria. Accordingly, only *uba* was used as a reference gene. It was later shown that the expression of hepatic bactin is regulated in fish that are exposed to estrogens (Filby and Tyler, 2007). In paper III, *hypoxantine phosphoribosyltransferase* (*hprt*) and a tata *box binding protein* (*tbp*) were evaluated in addition to ubiquitin, as they were suggested to be appropriate housekeeping genes for expression profiling of estrogens in the fathead minnow

(Filby and Tyler, 2007). The assay for *tbp* could not be optimized due to too low amplification efficiency. *Hprt* was clearly regulated between several of the treatment groups, while the expression of *uba* was consistent. Consequently, only *uba* was used as a reference gene. In paper IV, again two reference genes were evaluated, uba and *tubulin alpha* (*tuba*), which were selected based upon our microarray results and literature data. Both fulfilled the criteria for good reference genes.

3.6. Enzymatic assays

In paper IV, the analysis of enzymatic activities and glutathione levels was performed in order to validate hypothesis that were created from the microarray results. The activities of the hepatic antioxidant enzymes glutathione S-transferase, glutathione reductase and catalase, the activity of cytochrome P450 1A (Cyp1a) and the levels of glutathione were assessed through the use of well-established methods that had been previously applied in rainbow trout. All measured enzyme activities and glutathione levels were measured spectrophotometrically, with the samples diluted in a physiological buffer and the addition of a substrate or electron donor. The accumulation of product or the consumption of reagents was then monitored spectrophotometrically over time (Carney Almroth, 2008).

The regulation of antioxidant enzyme activity is complex, and hepatic antioxidant enzymes are often less responsive than, for example, Cyp1a (van der Oost et al., 2003). The low responsiveness can be partly explained by the rather high activities of glutathione S-transferase and catalase in contrast to the low Cyp1a activity in the liver under normal conditions. Measurements of antioxidant enzyme activities are, therefore, usually combined with the analysis of oxidative damage products, such as protein carbonyls, lipid peroxidation products or oxidative DNA damage (Carney Almroth, 2008). Analysis of protein carbonyls and lipid peroxidation products in the exposed fish in paper IV were considered, but we decided no to proceed with this since there was no tendency towards altered glutathione levels or antioxidant enzyme activities.

In the places I go there are things that I see
That I never could spell if I stopped with the Z.
I'm telling you this 'cause you're one of my friends.
My alphabet starts where your alphabet ends!
Dr Seuss

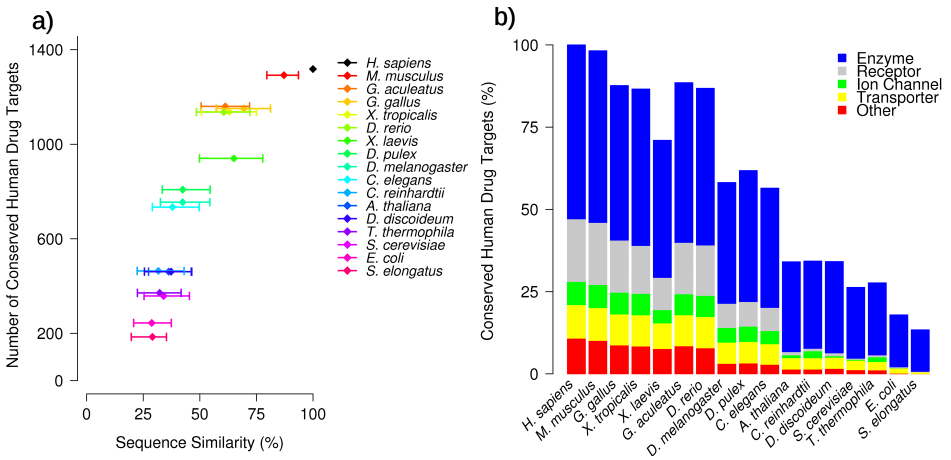

Figure 7. a) Median sequence similarity and the number of predicted orthologs in 16 non-target species compared to the human drug targets. The boxes indicate 25% and 75% quantiles. b) GO categories for the drug targets in human and in 16 non-target species. Some drug targets were annotated to more than one functional category. Note that the ortholog prediction in *X. laevis* is based on comprehensive EST data and not on a fully sequenced genome. The figures are reproduced with permission from (Gunnarsson et al., 2008) copyright 2009 American Chemical Society.

Receptors constitute an important group of pharmacological targets, with as many as 40% of all FDA approved drugs elicit their effects through receptors (Overington et al., 2006). Thus, the choice of environmental test species is particularly important for drugs that target receptors. Most drugs from this group are therefore predicted to have a low potency in non-vertebrates. Enzyme drug targets, on the other hand, were more ubiquitously present. Generally speaking, the drugs that target enzymes could be expected to affect a wider range of species, including invertebrates.

The presence of a drug target ortholog in a non-target species does not guarantee that a functional interaction with the drug can occur. However, we also presented data from the literature that supports the idea that an ortholog prediction can often indicate the ability of a conserved drug target protein to interact with the human drug. For example, the predicted well-conserved enzyme aldehyde dehydrogenase is inhibited by the drug disulfiram (which in humans is used as a alcohol deterrent), in humans as well as in a bacterium and a plant (Velasco-Garcia et al., 2006). Similarly, the 5HT transporter (sodium dependent serotonin transporter) is inhibited by SSRI drugs in humans, zebrafish (Wang et al., 2006), *D. melanogaster* (Demchyshyn et al., 1994) and a flatworm (Patocka and Ribeiro, 2007). The plant, the alga and the fungi were predicted to not have any orthologs to the 5HT

accordingly, a very high concentration of fluoxetine (1 mg/L) had no effect on either wet weight, frond number, chlorophyll or cartenoids in the plant *Lemna gibba* (Brain et al., 2004).

Based on our results, we propose that aquatic environmental risk assessment for drugs with human targets should include comprehensive studies on aquatic vertebrates. These results also suggests that toxicity data that are generated from a species that is missing a drug target ortholog might not be able to predict the toxicity of that same drug in another species that has the drug target ortholog. The copepode, *Nitocra spinipes*, does not have an ortholog to the estrogen receptor. Consequently, the NOEC value of 0.05 mg/L for chronic toxicity of EE_2 in the copepode (Breitholtz and Bengtsson, 2001) is not protective for fish, even if an assessment factor of 1000 was applied.

In Sweden, data from the environmental risk assessments that are required for the approval of new drugs in the EU and USA (se also section 1.3) are made easily available to prescribers and users at the web page www.fass.se. Although this is a good initiative, most of the available data are based upon results from acute tests, and data regarding effects on fish are often lacking. It is therefore important to recognize the limitations of this dataset for a complete environmental risk assessment perspective. The full list of all orthology predictions that were presented in paper I is available as supplementary material (Gunnarsson et al., 2008). The orthology data, however, should be interpreted carefully on an individual basis. The prediction of orthologs can be very difficult, especially when many paralogs are present and, even if the orthology prediction is correct, it does not guarantee ligand binding or conserved function. A more precise prediction of a potential drug target interaction might be possible with better knowledge about drug-binding domains. Sakharkar et al. (2007) have summarized protein family domains (Pfam) that are related to drug binding and called them druggable domains . Unfortunately, many of these domains are too general to provide additional useful information to predict the ability of a protein to interact with a drug in an evolutionarily distant species. The database Supersite (http://bioinformatics.charite. de/supersite) was recently released. It contains three-dimensional protein structures and ligand-binding site information for 1,300 drugs, as well as some evolutionary information (Bauer et al., 2009). Such information on drug binding domains could improve the current predictions.

4.1.2. Future perspectives - extending the orthology prediction

Accurate predictions of evolutionary conservation of all proteins, as well as their downstream pathways, that are known to interact with drugs may increase our understanding of both pharmacokinetics and pharmacodynamics in non-target species. For example, predictions of evolutionary conserved transporter proteins may enhance our understanding of the uptake and distribution of drugs. Passive diffusion through the lipid bilayer of the cell is often considered the dominant process by which a drug is taken up by the cell (Lipinski et al., 2001). However, Dobsen and Kell (2008) argue that carrier-mediated and active uptake of pharmaceuticals may be more common than traditionally assumed. Indeed, the evidence is growing for transporter-mediated drug uptake in those specific cases where it has been studied. Carrier-mediated and active drug uptake, for example, can lead to drugs concentrating in specific tissues. The presence of a conserved transporter protein might also explain why certain drugs bioaccumulate in some aquatic organisms to a higher degree than predicted from the drugs' lipophilicity (e.g., LogP). Data on the uptake and distribution of pharmaceuticals, as well as the molecular characterization of transporters in wildlife species, are needed to validate the importance of conserved transporter proteins.

Another possible extension of the ortholog prediction approach could be to predict evolutionary conservation of pathways downstream from the drug targets. Such information may provide input for an increased understanding of drugs' mode of action and means for the identification of affected physiological endpoints in wildlife species. The conserved counterparts of human pathways have been predicted in 22 species by the Reactome Project (www.reactome.org). Most of the species are mammalian model species or human pathogens, but the representatives for a few, perhaps more ecotoxicologically relevant, taxonomic groups are also included, such as the spotted green pufferfish (*Tetraodon nigroviridis*). The pathways that are affected by pharmaceuticals have not been as extensively studied as drug targets in mammals, and information about pathways in non-traditional model species is still very sparse. In general, a better understanding of the physiology of the wildlife species that are exposed to pharmaceuticals is needed. Nevertheless, we believe that large-scale comparisons of conserved proteins and pathways can still aid in the prioritization of which drugs need further attention for assessment of their environmental risks and which organisms should be prioritized for testing and what endpoints are appropriate.

4.2. Gene expression changes induced by estrogens

4.2.1. Microarray analysis of fish exposed to EE_2 (paper II)

This paper aimed to identify sensitive and robust hepatic gene expression changes that could potentially be used as additional biomarkers for exposure to estrogens in fish, which is an organism group that has functionally conserved estrogen receptors. Gene expression profiles were characterized in juvenile rainbow trout that were exposed to a measured concentration of 0.87 and 10 ng EE_2/L using the salmonid cDNA microarray from GRASP. The higher concentration was used to guide the identification of the more subtle responses that were induced by the low concentration. Twenty-nine potentially sensitive gene responses, all of which responded in a dose dependent manner, were identified. Transcripts encoding zona pellucida (*zp*) proteins were among these sensitive estrogen responsive genes, while *vtg* was not up-regulated by the low concentration of EE_2. Induction of the *zp* genes has been previously suggested to be more sensitive than *vtg* (Celius et al., 2000; Thomas-Jones et al., 2003).

Ensemble Transcript	Annotation		*O. mykiss*, juvenile, EE_2, 10 ng/L (Gunnarsson et al., 2007)	*O. mykiss*, male, EE_2 50 ng/L (Hook et al., 2006)	*O. mykiss*, juvenile, E_2 Dietary 5ppm (Tilton et al., 2006)	*O. latipes*, male, E_2 100 ng/L (Kishi et al., 2006)
61165	Vitellogenin	↑				
61744/ 61751/ 65820	Zona pellucida protein 3	↑				
49240	Transducer of ERBB2, 1a	↑				
56088/ 77745/ 8439/ 24598	Zona pellucida protein 2	↑				
26180/ 33724	Fatty acid binding protein 3 or 7	↑				
55579/ 74814	Peptidylprolyl isomerase B	↑				
59139/ 64339	Non-metastatic cells 2 (Nme2)	↑				
56095	Fatty acid binding protein 10	↓				
2842	Cytochrome P450	↓				

Figure 9. Transcripts or groups of transcripts affected by estrogen treatment (17-α-ethinylestradiol, EE_2 or 17β-estradiol, E_2) in at least three out of four microarray studies. Dark grey refers to an estrogen-induced mRNA expression, light grey refers to suppression whereas white refers to either missing value or not significantly regulated. Note that different microarray platforms and statistical methods were used in each study.

A meta-analysis was also performed in order to identify robust estrogen-responsive genes (Figure 9). The microarray results from the present study and three similar studies using different fish species, estrogens and microarray platforms were compared using the zebrafish genome as a reference. Transcripts that presumably corresponded to nine genes or groups of paralogous genes were differentially regulated by estrogens in at least three of the four studies. Two of these genes, which were confirmed by qPCR, fulfilled the criteria of high sensitivity and robustness. One was the previously known sensitive estrogen responsive gene, *zp3*, and the other was a novel estrogen responsive gene in fish, *non-metastatic cells 2* (nm23, *nme2*).

4.2.1.1. Nme2

Nme2 is a multifunctional protein that belongs to a larger class of nucleoside diphosphate kinases. The protein is highly conserved throughout evolution, and the salmon Nme2 shows high similarity to the human NME1 (NM23-H1) and NME2 (NM23-H2) proteins. A phylogenetic analysis suggests that the NME1 and NME2 are in-paralogs with respect to the teleost fish genomes (i.e., they have arisen by gene duplication after the speciation event that separated modern teleost fish from the human lineage) (Murphy et al., 2000). It is assumed, therefore, that the rainbow trout genome only has one ortholog of the NME1 and NME2 proteins. In humans, the *NME2* encodes the c-MYC transcription factor (Postel et al., 1993). NME1 has DNAse and histidine protein kinase activities and is also known to limit metastasis (Curtis et al., 2007). Furthermore, *NME1* gene expression is induced by estrogens (Lin et al., 2002), and NME1 also interacts with the estrogen receptor 1 (estrogen receptor α, ESR1) and can influence ESR1-mediated gene expression (Curtis et al., 2007). It is suggested that the metastasis-suppressive effects of NME1 is mediated, at least in part, by its capacity to limit expression of *cathepsin D* (*CTSD*) and the proto-oncogene *BCL2*. Both these genes are estrogen-inducible, and their proteins are involved in cell migration, apoptosis and angiogenesis (Rochefort et al., 1990; Biroccio et al., 2000).

The meta-analysis revealed that *nme2* was induced in all three microarray studies with rainbow trout. The study with Japanese medaka (*Oryzias latipes*) did not report *nme2* as an estrogen-responsive gene, but it is unclear if *nme2* was represented on the microarray (Kishi et al., 2006). Recently, Tobias Österlund in our lab performed an extended meta-analysis with five microarray studies on fish that were exposed to estrogens (Österlund, 2009). He identified *nme2* as slightly up-

regulated in estradiol-treated stickleback (Geoghegan et al., 2008), which supports the hypothesis that estrogen also regulates *nme2* in fish. The physiological function of Nme2 in fish, however, remains to be determined.

4.2.2. Induction of estrogen-responsive genes by different STP effluents (paper III)

In paper III, we investigated the gene expression of *nme2*, *zp3* and *vtg*, along with four other estrogen-responsive genes, with the aim to compare six different sewage treatment technologies' ability to reduce estrogen-induced gene expression changes in exposed male fish. STP effluents often contain a mixture of estrogenic compounds such as EE_2, natural hormones and industrial phenols. The total combined activity of the estrogenic substances may be sufficiently high to cause feminization of fish in the environment (Purdom et al., 1994; Sumpter and Jobling, 1995; Larsson et al., 1999; Jobling et al., 2002a). Improvements to and/or the implementation of new sewage treatment technologies have been suggested to remove micro-contaminants, such as pharmaceuticals and estrogenic substances (Clara et al., 2005; Esplugas et al., 2007; Koh et al., 2008; Laven et al., 2009).

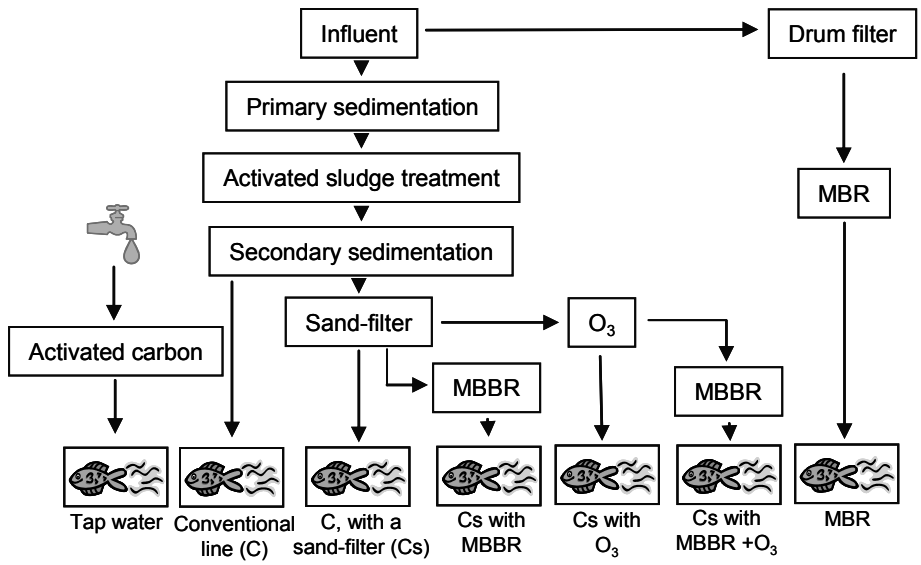

Figure 10. Principal outline of the different sewage treatment lines and the reference fish. Fish were exposed to tap water or effluent from the conventional line (C), C with a sand-filter (Cs), Cs with a moving bed biofilm reactor (MBBR), Cs with ozonation (O_3), Cs with both O_3 and MBBR (O_3 + MBBR) or the membrane bioreactor line (MBR).

In this study, one municipal sewage influent was treated with six different treatment technologies in parallel (Figure 10), and fish were exposed to the different effluents. The gene expression of *zp3*, *esr1*, *nme2*, *cyp1a*, *estrogen receptor 2a* and *2b* (*esr2a and esr2b*) were then analyzed in the exposed fish.

The STP lines with the advanced treatment technologies ozonation and/or moving bed biofilm reactor (MBBR) removed the majority of the measured estrogenic chemicals from the effluents. For example, EE_2 and E_2 were removed to a level below detection limits. However, the biological data suggested that the effluent from the MBBR line still induced some estrogenic responses in the exposed male fish. Ozonation was required to remove all measured biological effects.

Advanced treatment technologies, such as ozonation, can be very effective in removing pharmaceuticals and estrogenic substances (Esplugas et al., 2007; Laven et al., 2009), however, they are generally expensive to implement (Koh et al., 2008). The chemical oxidation, which is produced by ozonation, can also potentially create hazardous by-products (Petala et al., 2008), and this risk needs to be further studied.

In this study, the fish were exposed to 100% effluent, while a substantial dilution commonly occurs when STP effluents reach surface waters. It may not be necessary to remove so much of the estrogenic substances if an effluent is heavily diluted when it reaches surface waters, especially if the substances of concern are degradable in nature. The addition of a sand-filter to the conventional line (Cs), which is a solution that is quite common in Swedish STPs, reduced the gene expression of *vtg*, *esr1* and *zp3* the most dramatically (Table 1). If only the removal of estrogenic substances is considered, the addition of a sand-filtering step might be a suitable and relatively inexpensive improvement for many STPs.

The membrane bioreactor (MBR) technology reduced the expression of estrogen-responsive genes to baseline levels (Table 1). However, the fish exposed to this effluent or the Cs effluent had enlarged livers, which could be caused by a variety of chemicals. The cause behind the enlarged livers could be important to investigate further. The fish exposed to the Cs effluent had a slightly induced expression of *cyp1a* in addition to enlarged livers. Induction of *cyp1a* gene expression is a biomarker for contaminant exposure. Many chemicals, such as polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), some pesticides and pharmaceuticals, can induce *cyp1a* (Whyte et al., 2000).

Cost is an important aspect to consider when an implementation of a new sewage treatment technology is discussed, but the costs need to be weighed against the possible benefits. Decreasing the risk of effects on effluent-exposed wildlife species is clearly one benefit that should be evaluated for any candidate treatment technology. Unfortunately, determining what biological effect that may pose a risk to cause detrimental effects in the environment and what price we are willing to pay to avoid this is difficult.

4.2.3. Biomarkers of estrogen exposure (papers II, III and IV)

We have analyzed several estrogen responsive genes with the argument that a combination of responses increases the evidence that the fish have indeed been exposed to estrogenic substances to such a degree as to affect estrogen-regulated physiological processes (Table 1). We have also been able to evaluate the usefulness of *nme2* as a biomarker. *Vtg*, *zp3*, *nme2* and *esr1* are all genes that are induced by estrogens and have estrogen-responsive elements (EREs) in their promoters (Lin et al., 2002; Arukwe and Goksoyr, 2003; Yu et al., 2006). The ligand-activated estrogen receptor regulates transcription by binding to the EREs, interacting with other transcription factors and interacting with co–regulators, such as activators or repressors. *Nme2* did not have the same pattern of regulation as the other genes when the gene expression results from papers II, III and IV were compared (Table 1). The estrogenic chemicals present in, for example, the effluent from the conventional line did not significantly induce gene expression of *nme2* in exposed fish. Many factors may regulate the expression of *nme2*, and the estrogen-induced gene expression of *nme2* does not seem to follow a traditional dose-response pattern. Thus, the regulation of *nme2* is probably not specific or robust enough to work as a biomarker for estrogen exposure in a complex exposure situation. However, *nme2* could still have some value in the assessment of single estrogenic chemicals, if used together with other biomarkers of exposure, such as *zp3* and *vtg*.

Zp3 and *esr1* showed a similar pattern of induction in all studies, while *vtg* was not induced by the low concentration of EE_2 in paper II or by exposure to the effluent in paper IV. The zona pellucida genes are also referred to as vitelline envelope proteins or zona radiata proteins in fish (Larsson et al., 1994). Similar to Vtg, they are, in many fish species, produced in the liver under the influence of estrogen in maturing females and transported to the growing oocytes. There, they are deposited to form the extracellular coat that surrounds the oocyte of all

vertebrates. In mammals, as well as in some fish species, however, the zona pellucida proteins are synthesized within the ovary rather than the liver. Zona pellucida genes are not expressed in the liver of zebrafish, for example, and their expression in the ovaries is not under the regulation of estrogen (Liu et al., 2006). Thus, in fish that lack hepatic Zp synthesis, the induction of *zp* genes in the liver will not be useful as biomarkers of exogenous estrogen exposure. In salmonids, which are a group of fish that have hepatic Zp synthesis, the induction of *zp* genes in the liver is known to be more sensitive to estrogen exposure than *vtg* (Celius et al., 2000; Thomas-Jones et al., 2003; Gunnarsson et al., 2007) , but the specificity of the Zp induction for estrogens has been questioned (Larsson et al., 2002; Rotchell and Ostrander, 2003; Berg et al., 2004).

Table 1. Summary of the fold change of the different estrogen-regulated genes *vitellogenin* (*vtg*), *zona pellucida protein 3* (*zp3*), *estrogen receptor 1* (*esr1*) and *non-metastatic cells 2* (*nme2)*.

	Treatment (ng/L)	***vtg*** fold change	***zp3*** fold change	***esr1*** fold change	***nme2*** fold change
Paper II	**EE_2** (10)[a]	**540** **(p=0.001)**	**84** **(p=0.002)**	NA	**1.5** **(p=0.01)**
	EE_2 (0.87)[b]	1.0 (p=0.9)	**3.5** **($p=8 \times 10^{-4}$)**	NA	**1.4** **(p=0.01)**
Paper III	**C** (6.3 E_1, 0.3 E_2, 0.2 EE_2)[c]	**1500** **($p= 6 \times 10^{-12}$)**	**5.1** **($p=3 \times 10^{-6}$)**	**2.3** **($p=6 \times 10^{-7}$)**	1.2 (p=0.2)
	Cs (0.7 E_1, 0.2 E_2, 0.2 EE_2)[d]	**23** **($p=1 \times 10^{-4}$)**	**2.1** **(p=0.04)**	**1.8** **($p=5 \times 10^{-4}$)**	**1.5** **(p=0.02)**
	MBR (0.5 E_1, <0.1 E_2, 0.1 EE_2)[e]	1.8 (p=0.3)	1.6 (p=0.1)	0.29 (p=0.6)	**1.5** **(p=0.03)**
	O_3 (0.2 E_1, <0.1 E_2, <0.1 EE_2)[f]	2.3 (p=0.3)	1.4 (p=0.7)	0.7 (p=0.2)	1.2 (p=0.1)
Paper IV	**PETL effluent** (0.6 E_3)[g]	1.4 (p=0.2)	**1.9** **(p=0.005)**	**1.7** **(p=0.006)**	**1.2** **(p=0.009)**

To allow comparisons between studies, all of the p-values were re-calculated with a two sided t-test, and no correction for multiple testing was applied. Bold indicates a significant regulation (p value < 0.05). NA = data not available. The fold changes were calculated as the ratio between gene expression in the exposed fish and the expression in the reference fish. The gene expression levels in the reference fish in paper III were excluded from the statistical analysis in the original study, since 30 out of the 44 fish were killed by an unannounced raise in the Cl_2-concentration in the drinking water early during the exposure. Based on similar expression levels in the reference fish and the fish that were exposed to effluent with low levels of estrogens and on follow-up experiments in 2009 (unpublished data), we believe that the data from the reference fish nevertheless provides some information on the basal expression levels. The gene expression in those reference fish was thus used to enable comparisons between the different studies that are presented in the table. The different treatment lines in paper III were: conventional line (C), conventional line with a sand filter (Cs), the line with a membrane bioreactor (MBR) and the Cs line with ozonation (O_3) (see Figure 10). The measured estrogens and their potencies: 17-α-ethinylestradiol (EE_2), >17β-estradiol (E_2), >estrone (E_1) and >estriol (E_3). The number of biological replicates in each study was: [a] (n=4, n=4),[b] (n=8, n=8),[c] (n=8, n=6),[d] (n=6, n=6), [e] (n=6, n=6),[f] (n=6, n=6) and [g] (n=9, n=10).

In paper IV, it is unclear what chemicals induced *zp3* and *esr1*. The concentrations of EE_2, estrone (E_1), estradiol (E_2) and estriol (E_3) were analyzed in the effluent, but only E_3 was detected. Given the high dilution of the effluent (1:500), the calculated exposure concentration of E_3 was not sufficient to explain the observed induction. The potential presence of estrogenic industrial phenols, such as nonylphenol or other estrogenic chemicals that were not measured could explain the weak estrogenic effect.

Estrogens can induce *vtg* several thousand-fold, and it has been shown in other studies that the induction corresponds fairly well to the magnitude of the exposure. In our studies, however, the gene expression of *zp3* agrees better with the measured concentrations of estrogen-acting substances in papers II and III. The three-fold difference between the *vtg* induction at the high concentration of EE_2 in paper II and the induction of *vtg* by the conventional effluent in paper III, however, is marginal considering that a *vtg* induction could span several orders of magnitude, and could potentially be explained by technical variability in the qPCR measurements or by biological variability between the reference fish in the two studies.

4.2.3.1. The connection to adverse effect

The utility of *vtg* and, consequently, also *zp3* in male fish as biomarkers of potentially adverse effects is somewhat controversial. I agree with Hutchinson et al. (2006) that biomarkers for endocrine disruption should be viewed as signposts and not traffic lights. An induction of *vtg* or *zp3* in the field situation does not necessarily imply an adverse effect for the individual or the population. However, in a controlled lab experiment, an induction of *vtg* and/or *zp3* can serve as a biomarker of potential adverse effects. In other words, reproductive effects may occur if the fish were to be continuously exposed to the same concentration of the investigated estrogen for a prolonged time period or briefly exposed during a particularly sensitive time window. Such time windows exist during the very sensitive sexual differentiation period of the brain and gonads, such as very early during development or during the reproductively active phase when a negative feedback of estrogen on the hypothalamic-pituitary axis may interfere with the hormonal signaling that is required for producing viable gametes.

Manufacturers of pharmaceuticals have also showed an interest in illuminating how to use or possibly not use molecular biomarkers for estrogenic effects. For

example, Caldwell et al. (2008) conclude in their paper that the induction of Vtg in male fish should be considered as a biomarker of exposure and not of effect. They evaluated all papers investigating EE_2 and reproductive effects that were published before mid-2007 with the aim to derive an aquatic predicted no-effect concentration. Twelve of the studies that passed their selection criteria reported both a reproductive NOEC and Vtg NOEC in fish. Two studies showed that Vtg induction was less sensitive than reproductive endpoints were (Länge et al., 2001; Parrott and Blunt, 2005), three studies showed the opposite pattern and seven studies reported the same NOEC concentration for the Vtg induction and reproductive endpoints (Caldwell et al., 2008). My conclusion from the presented data is that Vtg induction appears to be a rather good biomarker of potentially adverse effects. In addition, several of the included studies report reproductive effects at concentrations around 1 ng/L. Thus, specific gene expression changes that are induced by estrogens at this concentration could serve as signposts/biomarkers for potentially adverse effects caused by estrogens.

Caldwell et al. (2009) further suggested 0.35 ng/L as the aquatic PNEC for EE_2, based up on an analysis of 39 papers that covered 26 species. More than half of the included species are not believed to have a functionally conserved estrogen receptor, which means that those species will naturally have much higher NOEC values than species that have a functional estrogen receptor (see also section 4.1.1.). Consequently, the derived aquatic PNEC value will depend on how many of the species included in the analysis are missing estrogen receptor orthologs. Nevertheless, the suggested PNEC value of 0.35 ng/L is below most reported LOEC values in fish and is, therefore, likely to be protective for the majority of fish species if no other estrogens are present, which could add to the total effect.

4.3. Microarrays and exposure to complex mixtures

Microarray analysis could be useful for the development of biomarkers of exposure (paper II) as well as to increase the knowledge about the toxicological mode of action. The mechanism of sub-lethal toxicity of a complex mixture, such as an effluent, could also be explored with microarray analysis.

4.3.1. Effluent from bulk drug manufacturers (paper IV)

Larsson et al. (2007) and Carlsson et al. (2009) recently showed that an effluent from bulk drug manufacturers contained a mixture of a wide range of APIs in high

concentrations along with a variety of other chemicals (see also section 1.1.). The investigated effluent came from a treatment plant, PETL, that regularly receives process water from about 90 pharmaceutical industries in Patancheru, India. Many of the APIs that were detected in the effluent in concentrations higher than 100 µg/L were antibiotics. This raises concerns about the development of antibiotic resistance, and our research group has several ongoing studies that address that matter.

Carlsson et al. (2009) investigated the potency of the effluent from PETL to aquatic vertebrates. As little as 0.2% of the effluent was sufficient to strongly reduce the growth of tadpoles (Carlsson et al., 2009). However, the underlying toxicological mode(s) of action and which substances in the mixture are the most biologically potent were not known. The aim of paper IV was to start to unravel the mechanism behind this toxicity to aquatic vertebrates. Therefore, we exposed fish to 0.2% of the effluent and analyzed the effects on hepatic global gene expression along with other physiological endpoints.

Cyp1a was the most significantly regulated gene, and the exposed fish also showed a four-fold increase in Cyp1a activity. The detoxification enzyme Cyp1a is involved in the metabolism of many different organic chemicals, such as polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs) (Whyte et al., 2000) and pharmaceuticals (Lee et al., 2006). Many of the chemicals that are substrates for Cyp1a metabolism are agonists to the ligand-activated transcription factor, aryl hydrocarbon receptor (AhR), which can induce the expression of genes with xenobiotic responsive elements (Köhle and Bock, 2007). *Cyp1a* can be induced via the Ah-receptor or via other mechanisms (Hu et al., 2007). A few genes that are known to be regulated via the Ah-receptor were affected, while other AhR responsive genes had no altered gene expression in the exposed fish. The observed ethoxyresorufin-O-deethylase (EROD) induction may be an indication of adverse effects that are mediated via AhR, but a longer exposure time or exposure to a higher concentration is needed to elucidate if such a mechanism contributes to the toxicity of the effluent.

The concentration of more than 200 different chemicals, such as APIs, organic solvents, metals and pesticides, have been analyzed in the effluent (Larsson et al., 2007; Carlsson et al., 2009) and, to our knowledge, no individual chemical was detected at levels that could explain the observed EROD induction in the exposed fish. This does not exclude, however, the possibility that an individual API or non-pharmaceutical chemical was responsible, as many of these have not been tested in

fish. Several fluoroquinolones were measured in very high concentrations, but they are known to inhibit rather than induce hepatic CYP1A activity in mammals (McLellan et al., 1996; Regmi et al., 2005). Inhibition of Cyp1a enzyme activity could lead to an inefficient metabolism of its substrates, which would imply that low concentrations of Cyp1a substrates, which are often inducers, could be sufficient to up-regulate *cyp1a* transcription, if any inhibitors are present in the mixture. Data from a pilot study (data not shown) showed that the PETL effluent first starts to inhibit EROD activity in β-naphthoflavone (BNF)-stimulated cells at a dilution with roughly 2% effluent. The inhibitory effect of the fluoroquinolones in fish that are exposed to 0.2% effluent, therefore, is assumed to be low.

Several oxidative stress-related genes, such as *glutathione S-transferase π (gstp), glutathione peroxidase (gpx)* and *superoxide dismutase (sod1, SOD [Cu/Zn])*, were slightly up-regulated in the exposed fish. In order to test if the exposed fish were suffering from oxidative damage, the hepatic levels of glutathione and the enzyme activities of glutathione S-transferase, glutathione reductase and catalase were measured. The exposed fish had no altered levels of glutathione or antioxidant enzyme activities, which suggested that they were not suffering from oxidative damage. Oxidative damage cannot be excluded at a higher concentration or a longer exposure time.

The blood chemistry analysis showed that the exposed fish had significantly increased plasma phosphate and cholesterol levels when compared to the control fish. Elevated phosphate blood levels are associated with renal failure and hypoparathyroidism in mammals (Berner and Shike, 1988). High concentrations of the fluoroquinolone, ciprofloxacin, have been associated with oxidative stress in the liver and kidney in mammals (Gurbay and Hincal, 2004; Weyers et al., 2008) but it is unknown how ciprofloxacin affects fish. The increase in cholesterol levels has some support in the pattern of regulated genes. Regulated transcripts with, GO-terms such as cellular metabolic process (GO:0044237) and triacylglycerol metabolic process (GO:0006641), were overrepresented in the list of the top 300 most significantly regulated transcripts. Transcripts with the GO-term cholesterol catabolic process (GO:0006707) were overrepresented in the top 600 list, according to an analysis that was performed with the GOrilla web service (Eden et al., 2009) (http://cbl-gorilla.cs.technion.ac.il/). Further studies are needed to evaluate the physiological importance of the increased levels of plasma phosphate and cholesterol in the exposed fish.

In conclusion, the results from this study show that exposure of fish to a high dilution (1:500) of the PETL effluent affects the expression of hepatic genes that are involved in the biotransformation and detoxification of chemicals and pharmaceuticals. We were not able to link the pattern of regulated genes to any specific group of chemicals or pharmaceuticals, but the observed alterations could serve as basis for hypothesizing which pharmaceuticals or compounds in the effluent that are the most biologically potent to aquatic vertebrates.

The combined conclusion from this and previous studies (Larsson et al., 2007; Carlsson et al., 2009) is that the wastewater treatment at PETL is insufficient, which implies that one of the world's largest centers for bulk drug production pollutes surface water to an extent that strongly affect aquatic life. PETL serves bulk drug manufacturer that exports their APIs to the world market. Many pharmaceutical products, including products on the Swedish market, contains APIs from this area (Larsson and Fick, 2009). Thus, Western countries, such as Sweden, have a responsibility to deal with the environmental problems that are related to the production of APIs in this area. However, information regarding the origin and the production conditions of the APIs in different pharmaceutical products is not easily available, which prevents prescribers and consumers from making informed decisions. An increased transparency would enable consumers and large purchasers, such as county councils, to put pressure on the pharmaceutical industry sector to reduce the environmental impacts throughout the production chain.

4.4. Microarrays and environmental risk assessments of pharmaceuticals

The potential of microarray analysis to aid in environmental risk assessment has been described in numerous articles (see also section 1.5.1) (Snape et al., 2004; Ankley et al., 2006; Ju et al., 2006; Lettieri, 2006; Steinberg et al., 2008). Ankley et al. (2006) described a conceptual timeline for the development and integration of toxicogenomic data into ecotoxicological regulatory programs. They suggest that we have just entered the phase of validation and "phenotypic anchoring" of the results from microarray studies.

A large proportion of the published ecotoxicologically relevant microarray data explore the effects of pharmaceuticals. Pharmaceuticals make up a group of chemicals for which a "proof of concept" is feasible since much is known about the molecular and toxicological responses to drugs in mammals (see section 1.4). In

addition, extensive sets of mammalian microarray data that describe responses to individual drugs are available (Fostel, 2008; Natsoulis et al., 2008).

Microarray analysis of drug-exposed, environmentally relevant species has been used to increase our understanding of toxicological modes of action. Heckmann et al. (2008) suggested a molecular mode of action of ibuprofen in *D. magna* that was based upon microarray data, and he linked the molecular responses to specific adverse effects. The analysis was, however, conducted with very high concentrations of ibuprofen, which only included levels that were more than 1000 times higher than those detected in the environment. Another example involved the analysis of the effects of the neuropharmaceutical mianserin in zebrafish (van der Ven et al., 2006), again using considerably higher concentrations than those normally found in the environment. Other studies have shown that different classes of chemicals induce different patterns of expression (Poynton et al., 2007; Williams et al., 2008). However, to my knowledge, there is no study that describes how a certain gene expression pattern has been used to pinpoint the effects of a chemical or group of chemicals within an unknown mixture. Falciani et al. (2008) started to approach this aim in a recent paper on the microarray analysis of European flounder (*Platichthys flesus*) that were sampled from field sites that were either clean or polluted in different ways. The paper described how a computational approach, which included data from previous field samplings and controlled lab-exposure experiments using single compounds (Williams et al., 2008), could identify gene expression patterns that predict which one of three sites the flounder was sampled from (Falciani et al., 2008).

There are very few published papers that use microarrays to characterize the mechanism of toxicity of complex mixtures, such as effluents (Steinberg et al., 2008). The microarray data in the study of Moens et al. (2007) with regard to an industrial effluent provided clues about the toxicological mode of action, although the chemicals that actually resulted in the toxicity remained unknown (Moens et al., 2007). Identifying causative agents and unraveling the molecular mode of action that underlies various toxic effects that result from the exposure to a complex mixture is a highly challenging task.

The use of microarray analysis on environmentally relevant species has just begun to emerge, and many issues remain to be solved. More studies that connect a specific gene expression pattern to a certain adverse effect and/or a certain class of chemicals are the first steps towards the development of predictive gene expression

patterns. Requirements for such studies include a vast understanding of the time and dose dependency of the gene response and a correlation to particular phenotypes.

The development of databases and tools for predictive toxicogenomics is part of the ongoing research within the pharmaceutical industry (Maggioli et al., 2006; Fostel, 2008; Natsoulis et al., 2008). In drug development, predictive toxicogenomics holds promise to aid in the preclinical evaluation of drug candidates by predicting toxic effects and identifying the mechanisms of toxicity. Information from such efforts may be useful for cross-species extrapolation and environmental risk assessment. However, the development of bioinformatical and statistical methods is needed for reliable interpretations of the microarray data and cross-species comparisons. Equally important is an increased understanding of the molecular responses and the general physiology of ecotoxicologically relevant species.

Don't give up! I believe in you all.
A person's a person, no matter how small!
Dr. Seuss

5. SUMMARY AND CONCLUSIONS

To summarize, different genomic approaches have been used to increase the understanding of the effects of pharmaceutical residues in the environment. The conclusions from the papers, upon which this thesis is based, are summarized below:

- **Comparative pharmacology (paper I)**

  We show that the majority of the specific targets for drugs are evolutionarily conserved in aquatic vertebrates, while invertebrates and plants lack orthologs to many of the drug targets for human pharmaceuticals. Therefore, we propose that aquatic environmental risk assessments for human pharmaceutical drugs should include comprehensive studies on aquatic vertebrates. The presented predictions can also serve as a basis for the interpretation of the relevance of existing ecotoxicity data and the identification of potentially sensitive wildlife species that could be used in the environmental risk assessment of pharmaceuticals.

- **Microarray analysis of fish exposed to the API, EE_2 (paper II)**

  Two potentially robust and sensitive biomarkers of exposure were identified by exploring the hepatic global gene expression pattern in fish that were exposed to estrogens. The results from our own microarray analysis, which used both a high and a very low concentration of EE_2, were combined with results from a meta-analysis of microarray data from three additional studies. Two identified, sensitive biomarkers, *nme2* and *zp3*, may increase the possibility to detect an exposure to low levels of estrogenic compounds in fish. In paper III, however, we showed that the regulation of *nme2* is not specific or robust enough to use as a biomarker in a complex exposure situation.

- **Microarray analysis of fish exposed to effluent from manufacturers of APIs (paper IV)**

  Fish exposed to the effluent from PETL, which was diluted 1:500, had an altered expression of hepatic genes involved in the biotransformation and detoxification of chemicals and pharmaceuticals, increased Cyp1a enzyme activity and increased plasma phosphate levels. The pattern of differentially expressed genes did not clearly point to any specific group of substances, but the observed gene

regulations could serve as the basis from which to hypothesize mechanisms of sub-lethal toxicity and possible causative agents. The presented data support previously published data that showed that the wastewater treatment at PETL is insufficient and that there is an urgent need for improvement.

- **The expression of estrogen responsive genes in fish exposed to municipal effluents, processed with different treatment technologies (paper III)**

  The advanced treatment technology ozonation was required during the STP process to remove all measured biological effects. However, advanced treatment technologies, such as ozonation, are generally expensive to implement. The addition of a sand-filtering step reduced the estrogen-induced response in male fish dramatically. In conclusion, this study provides valuable input into the assessment of advanced treatment technologies for the removal of estrogenic substances, given that the same influent was treated in parallel with a broad set of technologies and that the chemical analyses were combined with an *in vivo* assessment of estrogenic responses.

This thesis contains examples of how genomic techniques can be helpful in the environmental risk assessment of pharmaceuticals. In this context comparative genomics could aid in the prioritization of drugs, species and endpoints, and techniques, such as microarray analysis, could be useful to increase the knowledge regarding the toxicological mode of action and to develop biomarkers of exposure. These biomarkers could then be used to evaluate the ability of different STP technologies to decrease the effects of residual pharmaceuticals in the environment.

More knowledge is clearly needed to understand the risks that low concentrations of residual drugs pose to the aquatic environment. The risks that are connected to the release of very high concentrations of APIs at certain point sources are also important to assess and this might be the most urgent matter to resolve. High concentrations of APIs pose a significant threat to the local environment surrounding certain production sites and, in addition, high environmental concentrations of antibiotics could potentially be a more global threat.

6. ACKNOWLEDGEMENTS

Snart fem års doktorandstudier avslutas nu med ett gäng krumelurer i en bok. Ni är många som bidragit till att krumelurerna blev till och jag vill ödmjukt tacka för allt vetenskapligt stöd jag fått men framförallt för er vänskap och kärlek. Ett särskilt tack till:

Joakim Larsson, min handledare och vetenskapliga guide, för din outtömliga optimism och förmåga att entusiasmera. Tack för att du gav mig möjligheten att börja doktorera och för att du alltid har tagit dig tid att diskutera stora som små vetenskapliga problem.

Lars Förlin, min biträdande handledare, för att du har varit en lugn motpol och för att jag alltid har känt mig välkommen i ÅL-gruppen.

Erik Kristiansson, min främste samarbetspartner, för ditt stora vetenskapliga bidrag, för alla diskussioner som även om vi inte är överens alltid slutar med att jag förstått lite till (hoppas att du känner samma sak). Framförallt tack för din vänskap.

Eva Albertsson, utan dig hade jag nog inte genomfört de här fem åren. Tack för alla diskussioner, både de vetenskapliga och de om viktigare saker i livet. Tack också för att miljöfrågorna är så självklart viktiga för dig.

Carolin Rutgersson och **Linda Augustsson**, för gott samarbete och för ert oändliga tålamod med mig när jag har försökt att vara pedagogisk. Carolin, jag vet att du kommer att fortsätta skapa superbra forskning, håll flaggan högt!

Jag vill också tacka mina andra medförfattare som kämpat tillsammans med mig för att producera bra artiklar. Tack **Alexandra Jauhiainen**, **Olle Nerman**, **Margaretha Adolfsson-Erici**, **Berndt Björlenius**, **Joachim Sturve** och **Jerker Fick**.

Det finns många viktiga personer i zoologihuset. Bara att gå in genom dörren till det ganska fula huset har gjort mig på bättre humör så många gånger när något har känts tungt. Ett stort tack till **Jari Parkkonen**, för att du är en fiskexponeringshjälte och en grymt bra tågvän. Jag vill också tacka **Noomi Asker, Bethanie Carney Almroth, Anna Lennquist, Inger Holmqvist, Britt Wassmur** och **Johanna Gräns** för att ni alltid är glada och hjälpsamma.

Alla personer som gjort endokrin till en trevlig arbetsplats! Speciellt vill jag tacka **Magdalena Nutu** och **Linda Samuelsson** för att ni har varit grymt bra rumskamrater. **Birgitta Weidegård** för att du alltid tar dig an alla uppgifter med stort kunnande, entusiasm och ett leende. **Lena Olofsson** för att du är endokrins

stora administrativa hjältinna. **Louise Nilsson, Anders Friberg** och **Emilia Rung** för att ni har varit allvetande guruer bl.a. med kunskap om hur man skriver en avhandling. **Filip Cuklev** (tack för korrekturläsningen av kappan) och **Tobias Österlund** för trevliga och givande diskussioner, ni kommer att bli grymt bra forskare!

Alla trevliga människor som gick på forskarskolan i bioinformatik och genomik. Jag vill också särskilt tacka **Anders Blomberg** för att jag fick möjligheten att prova på forskning, **Homan Alipour** och **Jonas Warringer** för många goda råd och för att ni visste att jag skulle börja forska långt innan mig.

Mina lunchkompisar, **Michael Thorsen** och **Gabriella Arne** för trevligt sällskap och nyttiga microarraytips, och **Henrik Antonsson** och **Jakob Åsberg** för många bra diskussioner och för er vänskap.

Klara Stenqvist, **Marin Persson** och **Silja Marit Zetterqvist** (ett särskilt tack för språkförbättringen av den populärvetenskapliga sammanfattningen), tack för att ni är fantastiska vänner, jag är så glad att jag har er.

Alla fina människor som uppskattar att klättra på stenar. Ett särskilt tack till **Lina Lundkvist** för stor hjälp med den populärvetenskapliga sammanfattningen och för att du är en fantastiskt fin vän. Stort tack till **Josephine, Kalle, Frida, Fredrik** (särskilt tack för hjälp med bilden) **Martin m.fl.** för att ni drar med mig ut och för att ni påminner mig om att livet är ganska gött i solen med händerna mot en klippa.

My American family, that came along with the love of my life. Thank you for being the best "extra" family imaginable, for support and for giving me the opportunity to have great long vacations at beautiful places such as Pie Town and Malibu.

Mina föräldrar, Ingela och Håkan och min bror Peter, för att ni alltid har uppmuntrat mig och gett mig stöd, för att ni finns där när jag behöver det och för er kärlek. Nu kommer jag förhoppningsvis att ha mer tid att vara i Skåne och Blekinge.

Hammer strike, för trevligt sällskap under avhandlingsskrivandet och för att du gav mig en bra anledning att ta det lite lugnare!

Walker, det finns inte tillräckligt med ord för att beskriva hur glad och tacksam jag är för att du vill dela livet med mig. Tack för hjälpen med avhandlingen, tack för att du har skött markservicen de senaste månaderna och framförallt tack för att du påminner mig om vad som är viktigt i livet. I love you!

6. REFERENCES

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25 (17), 3389-3402.

Ankley, G.T., Brooks, B.W., Huggett, D.B., Sumpter, J.P. 2007. Repeating history: pharmaceuticals in the environment. Environ Sci Technol 41 (24), 8211-8217.

Ankley, G.T., Daston, G.P., Degitz, S.J., Denslow, N.D., Hoke, R.A., Kennedy, S.W., Miracle, A.L., Perkins, E.J., Snape, J., Tillitt, D.E., et al. 2006. Toxicogenomics in regulatory ecotoxicology. Environ Sci Technol 40 (13), 4055-4065.

Arukwe, A., Goksoyr, A. 2003. Eggshell and egg yolk proteins in fish: hepatic proteins for the next generation: oogenetic, population, and evolutionary implications of endocrine disruption. Comp Hepatol 2 (1), 4.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25 (1), 25-29.

Bartel, D.P. 2009. MicroRNAs: target recognition and regulatory functions. Cell 136 (2), 215-233.

Bauer, R.A., Gunther, S., Jansen, D., Heeger, C., Thaben, P.F., Preissner, R. 2009. SuperSite: dictionary of metabolite and drug binding sites in proteins. Nucleic Acids Res 37 (Database issue), D195-200.

Baum, M., Bielau, S., Rittner, N., Schmid, K., Eggelbusch, K., Dahms, M., Schlauersbach, A., Tahedl, H., Beier, M., Guimil, R., et al. 2003. Validation of a novel, fully integrated and flexible microarray benchtop facility for gene expression profiling. Nucleic Acids Res 31 (23), e151.

Bendz, D., Paxeus, N.A., Ginn, T.R., Loge, F.J. 2005. Occurrence and fate of pharmaceutically active compounds in the environment, a case study: Hoje River in Sweden. J Hazard Mater 122 (3), 195-204.

Berg, A.H., Westerlund, L., Olsson, P.E. 2004. Regulation of Arctic char (Salvelinus alpinus) egg shell proteins and vitellogenin during reproduction and in response to 17beta-estradiol and cortisol. Gen Comp Endocrinol 135 (3), 276-285.

Berner, Y.N., Shike, M. 1988. Consequences of phosphate imbalance. Annu Rev Nutr 8, 121-148.

Biroccio, A., Candiloro, A., Mottolese, M., Sapora, O., Albini, A., Zupi, G., Del Bufalo, D. 2000. Bcl-2 overexpression and hypoxia synergistically act to modulate vascular endothelial growth factor expression and in vivo angiogenesis in a breast carcinoma line. Faseb J 14 (5), 652-660.

Bolstad, B.M., Irizarry, R.A., Astrand, M., Speed, T.P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19 (2), 185-193.

Brain, R.A., Johnson, D.J., Richards, S.M., Sanderson, H., Sibley, P.K., Solomon, K.R. 2004. Effects of 25 pharmaceutical compounds to Lemna gibba using a seven-day static-renewal test. Environ Toxicol Chem 23 (2), 371-382.

Branham, W.S., Melvin, C.D., Han, T., Desai, V.G., Moland, C.L., Scully, A.T., Fuscoe, J.C. 2007. Elimination of laboratory ozone leads to a dramatic improvement in the reproducibility of microarray gene expression measurements. BMC Biotechnol 7, 8.

Breitholtz, M., Bengtsson, B.E. 2001. Oestrogens have no hormonal effect on the development and reproduction of the harpacticoid copepod Nitocra spinipes. Mar Pollut Bull 42 (10), 879-886.

Caldwell, D.J., Mastrocco, F., Hutchinson, T.H., Lange, R., Heijerick, D., Janssen, C., Anderson, P.D., Sumpter, J.P. 2008. Derivation of an aquatic predicted no-effect concentration for the synthetic hormone, 17 alpha-ethinyl estradiol. Environ Sci Technol 42 (19), 7046-7054.

Carlsson, G., Örn, S., Larsson, D.G.J. 2009. Effluent from Bulk Drug Production is Toxic to Aquatic Vertebrates. Environ Toxicol Chem, DOI: 10.1897/1808-1524.1891

Carney Almroth, B. 2008. Oxidative Damage in Fish Used as Biomarkers in Field and Laboratory Studies [PhD Thesis]. Gothenburg: University of Gothenburg.

Celius, T., Matthews, J.B., Giesy, J.P., Zacharewski, T.R. 2000. Quantification of rainbow trout (Oncorhynchus mykiss) zona radiata and vitellogenin mRNA levels using real-time PCR after in vivo treatment with estradiol-17 beta or alpha-zearalenol. J Steroid Biochem Mol Biol 75 (2-3), 109-119.

Chen, F., Mackey, A.J., Vermunt, J.K., Roos, D.S. 2007. Assessing performance of orthology detection strategies applied to eukaryotic genomes. PLoS ONE 2 (4), e383.

Clara, M., Strenn, B., Gans, O., Martinez, E., Kreuzinger, N., Kroiss, H. 2005. Removal of selected pharmaceuticals, fragrances and endocrine disrupting compounds in a membrane bioreactor and conventional wastewater treatment plants. Water Res 39 (19), 4797-4807.

Cohen, B. 1999. The Chipping Forcast. Nature genetics 20 (no.1 supplement).

Cui, C.W., Ji, S.L., Ren, H.Y. 2006. Determination of steroid estrogens in wastewater treatment plant of a controceptives producing factory. Environ Monit Assess 121 (1-3), 409-419.

Curtis, C.D., Likhite, V.S., McLeod, I.X., Yates, J.R., Nardulli, A.M. 2007. Interaction of the tumor metastasis suppressor nonmetastatic protein 23 homologue H1 and estrogen receptor alpha alters estrogen-responsive gene expression. Cancer Res 67 (21), 10600-10607.

Cuthbert, R., Parry-Jones, J., Green, R.E., Pain, D.J. 2007. NSAIDs and scavenging birds: potential impacts beyond Asia's critically endangered vultures. Biol Lett 3 (1), 90-93.

de Hoog, C.L., Mann, M. 2004. Proteomics. Annu Rev Genomics Hum Genet 5, 267-293.

De Lange, H.J., Noordoven, W., Murk, A.J., Lurling, M., Peeters, E.T. 2006. Behavioural responses of Gammarus pulex (Crustacea, Amphipoda) to low concentrations of pharmaceuticals. Aquat Toxicol 78 (3), 209-216.

Demchyshyn, L.L., Pristupa, Z.B., Sugamori, K.S., Barker, E.L., Blakely, R.D., Wolfgang, W.J., Forte, M.A., Niznik, H.B. 1994. Cloning, expression, and localization of a chloride-facilitated, cocaine-sensitive serotonin transporter from Drosophila melanogaster. Proc Natl Acad Sci U S A 91 (11), 5158-5162.

Desbrow, C., Rutledge, E.J., Brighty, G.C., Sumpter, J.P., Waldock, M. 1998. Identification of estrogenic chemicals in STW effluent: 1. Chemical fractionation and in vitro biological screening. Environ Sci Technol 32 (11), 1549-1558.

Dobson, P.D., Kell, D.B. 2008. Carrier-mediated cellular uptake of pharmaceutical drugs: an exception or the rule? Nat Rev Drug Discov 7 (3), 205-220.

Dolinski, K., Botstein, D. 2007. Orthology and functional conservation in eukaryotes. Annu Rev Genet 41, 465-507.

Dudoit, S., Shaffer, J.P., Boldrick, J.C. 2003. Multiple Hypothesis Testing in Microarray Experiments. Statistical Science 18 (1), 71-103.

ECETOC. 2007. Intelligent testing strategies in ecotoxicology: mode-of action approach for specifically acting chemicals. Bryssel, Belgium: Technical report from the European center for the Ecotoxicology and Toxicology of Chemicals (ECETOC).

Eden, E., Navon, R., Steinfeldt, I., Lipson, D., Yakhini, Z. 2009. GOrilla: a tool for discovering and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10 (48).

EMEA. 2006. Guideline on the environmental Risk Assessment of Medical Products for Human Use. London, UK:The Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency, CHMP/SWP/4447/00, 01 June 2006.

EMEA. 2007. Guideline on Environmental Impact Assessment for Veterinary Medicinal Products in support of the VICH Guidelines GL6 and GL 38. London: EMEA/CVMP/ERA/418282/05

Enright, A.J., Van Dongen, S., Ouzounis, C.A. 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res 30 (7), 1575-1584.

Esplugas, S., Bila, D.M., Krause, L.G., Dezotti, M. 2007. Ozonation and advanced oxidation technologies to remove endocrine disrupting chemicals (EDCs) and pharmaceuticals and personal care products (PPCPs) in water effluents. J Hazard Mater 149 (3), 631-642.

Falciani, F., Diab, A.M., Sabine, V., Williams, T.D., Ortega, F., George, S.G., Chipman, J.K. 2008. Hepatic transcriptomic profiles of European flounder (Platichthys flesus) from field sites and computational approaches to predict site from stress gene responses following exposure to model toxicants. Aquat Toxicol 90 (2), 92-101.

Fare, T.L., Coffey, E.M., Dai, H., He, Y.D., Kessler, D.A., Kilian, K.A., Koch, J.E., LeProust, E., Marton, M.J., Meyer, M.R., et al. 2003. Effects of atmospheric ozone on microarray data quality. Anal Chem 75 (17), 4672-4675.

FDA. 1998. Guidance for Environmental Assessment of Human Drug and Biologics Applications. Rockville, MD: Food and Drug Administartion, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research.

Filby, A.L., Tyler, C.R. 2007. Appropriate 'housekeeping' genes for use in expression profiling the effects of environmental estrogens in fish. BMC Mol Biol 8, 10.

Floate, K.D. 2006. Endectocide use in cattle and fecal residues: environmental effects in Canada. Can J Vet Res 70 (1), 1-10.

Fostel, J.M. 2008. Towards standards for data exchange and integration and their impact on a public database such as CEBS (Chemical Effects in Biological Systems). Toxicol Appl Pharmacol 233 (1), 54-62.

Garric, J., Vollat, B., Duis, K., Pery, A., Junker, T., Ramil, M., Fink, G., Ternes, T.A. 2007. Effects of the parasiticide ivermectin on the cladoceran Daphnia magna and the green alga Pseudokirchneriella subcapitata. Chemosphere 69 (6), 903-910.

Geoghegan, F., Katsiadaki, I., Williams, T.D., Chipman, J.K. 2008. A cDNA microarray for the three-spined stickleback, Gasterosteus aculeatus L., and analysis of the interactive effects of oestradiol and dibenzanthracene exposures Journal of Fish Biology 72 (9), 2133-2153.

Green, R.E., Taggart, M.A., Senacha, K.R., Raghavan, B., Pain, D.J., Jhala, Y., Cuthbert, R. 2007. Rate of decline of the oriental white-backed vulture population in India estimated from a survey of diclofenac residues in carcasses of ungulates. PLoS ONE 2 (1), e686.

Gunnarsson, L., Jauhiainen, A., Kristiansson, E., Nerman, O., Larsson, D.G.J. 2008. Evolutionary conservation of human drug targets in organisms used for environmental risk assessments. Environ Sci Technol 42 (15), 5807-5813.

Gunnarsson, L., Kristiansson, E., Förlin, L., Nerman, O., Larsson, D.G.J. 2007. Sensitive and robust gene expression changes in fish exposed to estrogen--a microarray approach. BMC Genomics 8, 149.

Gurbay, A., Hincal, F. 2004. Ciprofloxacin-induced glutathione redox status alterations in rat tissues. Drug Chem Toxicol 27 (3), 233-242.

Heberer, T. 2002. Occurrence, fate, and removal of pharmaceutical residues in the aquatic environment: a review of recent research data. Toxicol Lett 131 (1-2), 5-17.

Heckmann, L.H., Sibly, R.M., Connon, R., Hooper, H.L., Hutchinson, T.H., Maund, S.J., Hill, C.J., Bouetard, A., Callaghan, A. 2008. Systems biology meets stress ecology: linking molecular and organismal stress responses in Daphnia magna. Genome Biol 9 (2), R40.

Hocquette, J.F. 2005. Where are we in genomics? J Physiol Pharmacol 56 Suppl 3, 37-70.
Hoeger, B., Kollner, B., Dietrich, D.R., Hitzfeld, B. 2005. Water-borne diclofenac affects kidney and gill integrity and selected immune parameters in brown trout (Salmo trutta f. fario). Aquat Toxicol 75 (1), 53-64.
Hook, S.E., Skillman, A.D., Small, J.A., Schultz, I.R. 2006. Gene expression patterns in rainbow trout, Oncorhynchus mykiss, exposed to a suite of model toxicants. Aquat Toxicol 77 (4), 372-385.
Hu, W., Sorrentino, C., Denison, M.S., Kolaja, K., Fielden, M.R. 2007. Induction of cyp1a1 is a nonspecific biomarker of aryl hydrocarbon receptor activation: results of large scale screening of pharmaceuticals and toxicants in vivo and in vitro. Mol Pharmacol 71 (6), 1475-1486.
Huggett, B.D., Cook, C.J., Ericson, F.J., Williams, R.T. 2003. A Theoretical Model for Utilizing Mammalian Pharmacology and Safety Data to Prioritize Potential Impacts of Human Pharmaceuticals to Fish. Human and Ecological Risk Assessment 9, 1789-1799.
Hutchinson, T.H., Ankley, G.T., Segner, H., Tyler, C.R. 2006. Screening and testing for endocrine disruption in fish-biomarkers as "signposts," not "traffic lights," in risk assessment. Environ Health Perspect 114 Suppl 1, 106-114.
Jobling, S., Beresford, N., Nolan, M., Rodgers-Gray, T., Brighty, G.C., Sumpter, J.P., Tyler, C.R. 2002a. Altered sexual maturation and gamete production in wild roach (Rutilus rutilus) living in rivers that receive treated sewage effluents. Biol Reprod 66 (2), 272-281.
Jobling, S., Coey, S., Whitmore, J.G., Kime, D.E., Van Look, K.J., McAllister, B.G., Beresford, N., Henshaw, A.C., Brighty, G., Tyler, C.R., et al. 2002b. Wild intersex roach (Rutilus rutilus) have reduced fertility. Biol Reprod 67 (2), 515-524.
Joss, A., Keller, E., Alder, A.C., Gobel, A., McArdell, C.S., Ternes, T., Siegrist, H. 2005. Removal of pharmaceuticals and fragrances in biological wastewater treatment. Water Res 39 (14), 3139-3152.
Ju, Z., Wells, M.C., Walter, R.B. 2006. DNA microarray technology in toxicogenomics of aquatic models: Methods and applications. Comp Biochem Physiol C Toxicol Pharmacol.
Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., et al. 2008. KEGG for linking genomes to life and the environment. Nucleic Acids Res 36 (Database issue), D480-484.
Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P., Gingeras, T.R. 2002. Large-scale transcriptional activity in chromosomes 21 and 22. Science 296 (5569), 916-919.
Kasprzyk-Hordern, B., Dinsdale, R.M., Guwy, A.J. 2009. The removal of pharmaceuticals, personal care products, endocrine disruptors and illicit drugs during wastewater treatment and its impact on the quality of receiving waters. Water Res 43 (2), 363-380.
Kidd, K.A., Blanchfield, P.J., Mills, K.H., Palace, V.P., Evans, R.E., Lazorchak, J.M., Flick, R.W. 2007. Collapse of a fish population after exposure to a synthetic estrogen. Proc Natl Acad Sci U S A 104 (21), 8897-8901.
Kishi, K., Kitagawa, E., Onikura, N., Nakamura, A., Iwahashi, H. 2006. Expression analysis of sex-specific and 17beta-estradiol-responsive genes in the Japanese medaka, Oryzias latipes, using oligonucleotide microarrays. Genomics 88 (2), 241-251.
Koh, Y.K., Chiu, T.Y., Boobis, A., Cartmell, E., Scrimshaw, M.D., Lester, J.N. 2008. Treatment and removal strategies for estrogens from wastewater. Environ Technol 29 (3), 245-267.
Köhle, C., Bock, K.W. 2007. Coordinate regulation of Phase I and II xenobiotic metabolisms by the Ah receptor and Nrf2. Biochem Pharmacol 73 (12), 1853-1862.
Kolpin, D.W., Furlong, E.T., Meyer, M.T., Thurman, E.M., Zaugg, S.D., Barber, L.B., Buxton, H.T. 2002. Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999-2000: a national reconnaissance. Environ Sci Technol 36 (6), 1202-1211.

Kostich, M.S., Lazorchak, J.M. 2008. Risks to aquatic organisms posed by human pharmaceutical use. Sci Total Environ 389 (2-3), 329-339.

Kreke, N., Dietrich, D.R. 2008. Physiological endpoints for potential SSRI interactions in fish. Crit Rev Toxicol 38 (3), 215-247.

Kristiansson, E. 2007. Statistical Analysis of Gene Expression Data [PhD Thesis]. Göteborg, Sweden: Chalmers University of Thechnology.

Kristiansson, E., Asker, N., Förlin, L., Larsson, D.G.J. 2009a. Characterization of the Zoarces viviparus liver transcriptome using massively parallel pyrosequencing. submitted.

Kristiansson, E., Asker, N., Gunnarsson, L., Kuo, W., Förlin, L., Larsson, D.G.J. 2009b. Large-scale gene expression analysis in non-model organisms – de novo transcriptome sequencing and microarray design In: Swedish Bioinformatics Workshop2009b, Umeå.

Kümmerer, K., 2008. Pharmaceuticals in the Environment - Sources, Fate, Effects and Risks Springer Berlin Heidelberg.

Kuo, W.P., Liu, F., Trimarchi, J., Punzo, C., Lombardi, M., Sarang, J., Whipple, M.E., Maysuria, M., Serikawa, K., Lee, S.Y., et al. 2006. A sequence-oriented comparison of gene expression measurements across different hybridization-based technologies. Nat Biotechnol 24 (7), 832-840.

Länge, R., Hutchinson, T.H., Croudace, C.P., Siegmund, F., Schweinfurth, H., Hampe, P., Panter, G.H., Sumpter, J.P. 2001. Effects of the synthetic estrogen 17 alpha-ethinylestradiol on the life-cycle of the fathead minnow (Pimephales promelas). Environ Toxicol Chem 20 (6), 1216-1227.

Larkin, P., Villeneuve, D.L., Knoebl, I., Miracle, A.L., Carter, B.J., Liu, L., Denslow, N.D., Ankley, G.T. 2007. Development and validation of a 2,000-gene microarray for the fathead minnow (Pimephales promelas). Environ Toxicol Chem 26 (7), 1497-1506.

Larsson, D.G.J., Adolfsson-Erici, M., Parkkonen, J., Pettersson, M., Berg, A.H., Olsson, P.E., Förlin, L. 1999. Ethinyloestradiol - an undesired fish contraceptive? Aquat Toxicol 45, 91-97.

Larsson, D.G.J., de Pedro, C., Paxeus, N. 2007. Effluent from drug manufactures contains extremely high levels of pharmaceuticals. J Hazard Mater 148 (3), 751-755.

Larsson, D.G.J., Fick, J. 2009. Transparency throughout the production chain - a way to reduce pollution from the manufacturing of pharmaceuticals? Regulatory Toxicology and Pharmacology 53, 161-163.

Larsson, D.G.J., Hyllner, S.J., Haux, C. 1994. Induction of vitelline envelope proteins by estradiol-17 beta in 10 teleost species. Gen Comp Endocrinol 96 (3), 445-450.

Larsson, D.G.J., Mayer, I., Hyllner, S.J., Förlin, L. 2002. Seasonal variations of vitelline envelope proteins, vitellogenin, and sex steroids in male and female eelpout (Zoarces viviparus). Gen Comp Endocrinol 125 (2), 184-196.

Laven, M., Alsberg, T., Yu, Y., Adolfsson-Erici, M., Sun, H. 2009. Serial mixed-mode cation- and anion-exchange solid-phase extraction for separation of basic, neutral and acidic pharmaceuticals in wastewater and analysis by high-performance liquid chromatography-quadrupole time-of-flight mass spectrometry. J Chromatogr A 1216 (1), 49-62.

Lee, M.D., Ayanoglu, E., Gong, L. 2006. Drug-induced changes in P450 enzyme expression at the gene expression level: a new dimension to the analysis of drug-drug interactions. Xenobiotica 36 (10-11), 1013-1080.

Lettieri, T. 2006. Recent applications of DNA microarray technology to toxicology and ecotoxicology. Environ Health Perspect 114 (1), 4-9.

Li, D., Yang, M., Hu, J., Ren, L., Zhang, Y., Li, K. 2008. Determination and fate of oxytetracycline and related compounds in oxytetracycline production wastewater and the receiving river. Environ Toxicol Chem 27 (1), 80-86.

Li, L., Stoeckert, C.J., Jr., Roos, D.S. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13 (9), 2178-2189.

Lin, K.H., Wang, W.J., Wu, Y.H., Cheng, S.Y. 2002. Activation of antimetastatic Nm23-H1 gene expression by estrogen and its alpha-receptor. Endocrinol 143 (2), 467-475.

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46 (1-3), 3-26.

Liu, X., Wang, H., Gong, Z. 2006. Tandem-repeated Zebrafish zp3 genes possess oocyte-specific promoters and are insensitive to estrogen induction. Biol Reprod 74 (6), 1016-1025.

Lu, P., Vogel, C., Wang, R., Yao, X., Marcotte, E.M. 2007. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol 25 (1), 117-124.

Lumaret, J.P., Errouissi, F. 2002. Use of anthelmintics in herbivores and evaluation of risks for the non target fauna of pastures. Vet Res 33 (5), 547-562.

Maggioli, J., Hoover, A., Weng, L. 2006. Toxicogenomic analysis methods for predictive toxicology. J Pharmacol Toxicol Methods 53 (1), 31-37.

McKusick, V.A., Ruddle, F.H. 1987. Toward a complete map of the human genome. Genomics 1 (2), 103-106.

McLellan, R.A., Drobitch, R.K., Monshouwer, M., Renton, K.W. 1996. Fluoroquinolone antibiotics inhibit cytochrome P450-mediated microsomal drug metabolism in rat and human. Drug Metab Dispos 24 (10), 1134-1138.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., et al. 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318 (5848), 245-250.

Mimeault, C., Woodhouse, A.J., Miao, X.S., Metcalfe, C.D., Moon, T.W., Trudeau, V.L. 2005. The human lipid regulator, gemfibrozil bioconcentrates and reduces testosterone in the goldfish, Carassius auratus. Aquat Toxicol 73 (1), 44-54.

Moens, L.N., Smolders, R., van der Ven, K., van Remortel, P., Del-Favero, J., De Coen, W.M. 2007. Effluent impact assessment using microarray-based analysis in common carp: a systems toxicology approach. Chemosphere 67 (11), 2293-2304.

Muller, M., Rabenoelina, F., Balaguer, P., Patureau, D., Lemenach, K., Budzinski, H., Barcelo, D., de Alda, M.L., Kuster, M., Delgenes, J.P., et al. 2008. Chemical and biological analysis of endocrine-disrupting hormones and estrogenic activity in an advanced sewage treatment plant. Environ Toxicol Chem 27 (8), 1649-1658.

Murphy, M., Harte, T., McInerney, J., Smith, T.J. 2000. Molecular cloning of an Atlantic salmon nucleoside diphosphate kinase cDNA and its pattern of expression during embryogenesis. Gene 257 (1), 139-148.

Natsoulis, G., Pearson, C.I., Gollub, J., B, P.E., Ferng, J., Nair, R., Idury, R., Lee, M.D., Fielden, M.R., Brennan, R.J., et al. 2008. The liver pharmacological and xenobiotic gene response repertoire. Mol Syst Biol 4, 175.

Oaks, J.L., Gilbert, M., Virani, M.Z., Watson, R.T., Meteyer, C.U., Rideout, B.A., Shivaprasad, H.L., Ahmed, S., Chaudhry, M.J., Arshad, M., et al. 2004. Diclofenac residues as the cause of vulture population decline in Pakistan. Nature 427 (6975), 630-633.

Örn, S., Holbech, H., Madsen, H.T., Norrgren, L., Petersen, I.G. 2003. Gonad development and vitellogenin production in zebrafish (*Danio reio*) exposed to ethinylestradiol and methyltestosterone. Aquat Toxicol 65, 397-411.

Österlund, T. 2009. A Novel Framework for Cross-Species Gene Expression Analysis [Master Thesis]. Göteborg, Sweden: Chalmers University of Thechnology.

Overington, J.P., Al-Lazikani, B., Hopkins, A.L. 2006. How many drug targets are there? Nat Rev Drug Discov 5 (12), 993-996.

Owen, S.F., Giltrow, E., Huggett, D.B., Hutchinson, T.H., Saye, J., Winter, M.J., Sumpter, J.P. 2007. Comparative physiology, pharmacology and toxicology of beta-blockers: mammals versus fish. Aquat Toxicol 82 (3), 145-162.

Parrott, J.L., Blunt, B.R. 2005. Life-cycle exposure of fathead minnows (Pimephales promelas) to an ethinylestradiol concentration below 1 ng/L reduces egg fertilization success and demasculinizes males. Environ Toxicol 20 (2), 131-141.

Patocka, N., Ribeiro, P. 2007. Characterization of a serotonin transporter in the parasitic flatworm, Schistosoma mansoni: cloning, expression and functional analysis. Mol Biochem Parasitol 154 (2), 125-133.

Peakall, D.B., Walker, C.H. 1994. The role of biomarkers in environmental assessment (3). Vertebrates. Ecotoxicology 3, 173-179.

Peregrin-Alvarez, J.M., Parkinson, J. 2007. The global landscape of sequence diversity. Genome Biol 8 (11), R238.

Petala, M., Samaras, P., Zouboulis, A., Kungolos, A., Sakellaropoulos, G.P. 2008. Influence of ozonation on the in vitro mutagenic and toxic potential of secondary effluents. Water Res 42 (20), 4929-4940.

Postel, E.H., Berberich, S.J., Flint, S.J., Ferrone, C.A. 1993. Human c-myc transcription factor PuF identified as nm23-H2 nucleoside diphosphate kinase, a candidate suppressor of tumor metastasis. Science 261 (5120), 478-480.

Poynton, H.C., Varshavsky, J.R., Chang, B., Cavigiolio, G., Chan, S., Holman, P.S., Loguinov, A.V., Bauer, D.J., Komachi, K., Theil, E.C., et al. 2007. Daphnia magna ecotoxicogenomics provides mechanistic insights into metal toxicity. Environ Sci Technol 41 (3), 1044-1050.

Purdom, C.E., Hardiman, P.A., Bye, V.J., Eno, N.C., Tyler, C.R., Sumpter, J.P. 1994. Estrogenic effects of effluents from sewage treatment works. Chem Ecol 8, 275-285.

Regmi, N.L., Abd El-Aty, A.M., Kuroha, M., Nakamura, M., Shimoda, M. 2005. Inhibitory effect of several fluoroquinolones on hepatic microsomal cytochrome P-450 1A activities in dogs. J Vet Pharmacol Ther 28 (6), 553-557.

Rise, M.L., von Schalburg, K.R., Brown, G.D., Mawer, M.A., Devlin, R.H., Kuipers, N., Busby, M., Beetz-Sargent, M., Alberto, R., Gibbs, A.R., et al. 2004. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics. Genome Res 14 (3), 478-490.

Ritchie, M.E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A., Smyth, G.K. 2007. A comparison of background correction methods for two-colour microarrays. Bioinformatics 23 (20), 2700-2707.

Rochefort, H., Capony, F., Garcia, M. 1990. Cathepsin D: a protease involved in breast cancer metastasis. Cancer Metastasis Rev 9 (4), 321-331.

Rotchell, J.M., Ostrander, G.K. 2003. Molecular markers of endocrine disruption in aquatic organisms. J Toxicol Environ Health B Crit Rev 6 (5), 453-496.

Rouillard, J.M., Zuker, M., Gulari, E. 2003. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res 31 (12), 3057-3062.

Routledge, E.J., Sheahan, D., Desbrow, C., Brighty, G.C., Waldock, M., Sumpter, J.P. 1998. Identification of estrogenic chemicals in STW effluent. 2. In vivo responses in trout and roach. Environ Sci Technol 32 (11), 1559-1565.

Sakharkar, M.K., Sakharkar, K.R., Pervaiz, S. 2007. Druggability of human disease genes. Int J Biochem Cell Biol 39 (6), 1156-1164.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, C.A., Hutchison, C.A., Slocombe, P.M., Smith, M. 1977a. Nucleotide sequence of bacteriophage phi X174 DNA. Nature 265 (5596), 687-695.

Sanger, F., Nicklen, S., Coulson, A.R. 1977b. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74 (12), 5463-5467.

Schena, M., Shalon, D., Davis, R.W., Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270 (5235), 467-470.

Schwaiger, J., Ferling, H., Mallow, U., Wintermayr, H., Negele, R.D. 2004. Toxic effects of the non-steroidal anti-inflammatory drug diclofenac. Part I: histopathological alterations and bioaccumulation in rainbow trout. Aquat Toxicol 68 (2), 141-150.

Seiler, J.P. 2002. Pharmacodynamic activity of drugs and ecotoxicology--can the two be connected? Toxicol Lett 131 (1-2), 105-115.

Shendure, J., Ji, H. 2008. Next-generation DNA sequencing. Nat Biotechnol 26 (10), 1135-1145.

Shultz, S., Baral, H.S., Charman, S., Cunningham, A.A., Das, D., Ghalsasi, G.R., Goudar, M.S., Green, R.E., Jones, A., Nighot, P., et al. 2004. Diclofenac poisoning is widespread in declining vulture populations across the Indian subcontinent. Proc Biol Sci 271 Suppl 6, S458-460.

Sjögren, A., Kristiansson, E., Rudemo, M., Nerman, O. 2007. Weighted analysis of general microarray experiments. BMC Bioinformatics 8, 387.

Smyth, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3 (1), Article3.

Snape, J.R., Maund, S.J., Pickford, D.B., Hutchinson, T.H. 2004. Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. Aquat Toxicol 67 (2), 143-154.

Soetaert, A., Moens, L.N., Van der Ven, K., Van Leemput, K., Naudts, B., Blust, R., De Coen, W.M. 2006. Molecular impact of propiconazole on Daphnia magna using a reproduction-related cDNA array. Comp Biochem Physiol C Toxicol Pharmacol 142 (1-2), 66-76.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98 (19), 10869-10874.

Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., Futcher, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell 9 (12), 3273-3297.

Steinberg, C.E., Sturzenbaum, S.R., Menzel, R. 2008. Genes and environment - striking the fine balance between sophisticated biomonitoring and true functional environmental genomics. Sci Total Environ 400 (1-3), 142-161.

Sumpter, J.P., Jobling, S. 1995. Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. Environ Health Perspect 103 Suppl 7, 173-178.

Swan, G.E., Cuthbert, R., Quevedo, M., Green, R.E., Pain, D.J., Bartels, P., Cunningham, A.A., Duncan, N., Meharg, A.A., Oaks, J.L., et al. 2006. Toxicity of diclofenac to Gyps vultures. Biol Lett 2 (2), 279-282.

Ternes, T.A., Kreckel, P., Mueller, J. 1999. Behaviour and occurrence of estrogens in municipal sewage treatment plants--II. Aerobic batch experiments with activated sludge. Sci Total Environ 225 (1-2), 91-99.

The Gene Ontology Consortium. 2008. The Gene Ontology project in 2008. Nucleic Acids Res 36 (Database issue), D440-444.

Thomas-Jones, E., Thorpe, K., Harrison, N., Thomas, G., Morris, C., Hutchinson, T., Woodhead, S., Tyler, C. 2003. Dynamics of estrogen biomarker responses in rainbow trout exposed to 17beta-estradiol and 17alpha-ethinylestradiol. Environ Toxicol Chem 22 (12), 3001-3008.

Thomas, K.V., Hilton, M.J. 2004. The occurrence of selected human pharmaceutical compounds in UK estuaries. Mar Pollut Bull 49 (5-6), 436-444.

Tilton, S.C., Givan, S.A., Pereira, C.B., Bailey, G.S., Williams, D.E. 2006. Toxicogenomic profiling of the hepatic tumor promoters indole-3-carbinol, 17beta-estradiol and beta-naphthoflavone in rainbow trout. Toxicol Sci 90 (1), 61-72.

Triebskorn, R., Casper, H., Heyd, A., Eikemper, R., Kohler, H.R., Schwaiger, J. 2004. Toxic effects of the non-steroidal anti-inflammatory drug diclofenac. Part II: cytological effects in liver, kidney, gills and intestine of rainbow trout (Oncorhynchus mykiss). Aquat Toxicol 68 (2), 151-166.

van der Oost, R., Beyer, J., Vermeulen, N.P.E. 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: a reveiw. Environmental Toxicology and Pharmacology 13, 57-149.

van der Ven, K., Keil, D., Moens, L.N., Van Leemput, K., van Remortel, P., De Coen, W.M. 2006. Neuropharmaceuticals in the environment: mianserin-induced neuroendocrine disruption in zebrafish (Danio rerio) using cDNA microarrays. Environ Toxicol Chem 25 (10), 2645-2652.

Velasco-Garcia, R., Zaldivar-Machorro, V.J., Mujica-Jimenez, C., Gonzalez-Segura, L., Munoz-Clares, R.A. 2006. Disulfiram irreversibly aggregates betaine aldehyde dehydrogenase--a potential target for antimicrobial agents against Pseudomonas aeruginosa. Biochem Biophys Res Commun 341 (2), 408-415.

von Schalburg, K.R., Rise, M.L., Cooper, G.A., Brown, G.D., Gibbs, A.R., Nelson, C.C., Davidson, W.S., Koop, B.F. 2005. Fish and chips: various methodologies demonstrate utility of a 16,006-gene salmonid microarray. BMC Genomics 6, 126.

Wang, Y., Takai, R., Yoshioka, H., Shirabe, K. 2006. Characterization and expression of serotonin transporter genes in zebrafish. Tohoku J Exp Med 208 (3), 267-274.

Waring, J.F., Gum, R., Morfitt, D., Jolly, R.A., Ciurlionis, R., Heindel, M., Gallenberg, L., Buratto, B., Ulrich, R.G. 2002. Identifying toxic mechanisms using DNA microarrays: evidence that an experimental inhibitor of cell adhesion molecule expression signals through the aryl hydrocarbon nuclear receptor. Toxicology 181-182, 537-550.

Weyers, A., Ugnia, L.I., Ovando, H.G., Gorla, N.B. 2008. Antioxidant capacity of vitamin C in mouse liver and kidney tissues. Biocell 32 (1), 27-31.

WHO. 1993. Biomarkers and risk assessment: concepts and principles International Programme on Chemical Safety (IPCS), Environmental Health Criteria 155. Geneva.

Whyte, J.J., Jung, R.E., Schmitt, C.J., Tillitt, D.E. 2000. Ethoxyresorufin-O-deethylase (EROD) activity in fish as a biomarker of chemical exposure. Crit Rev Toxicol 30 (4), 347-570.

Wiegel, S., Aulinger, A., Brockmeyer, R., Harms, H., Loffler, J., Reincke, H., Schmidt, R., Stachel, B., von Tumpling, W., Wanke, A. 2004. Pharmaceuticals in the river Elbe and its tributaries. Chemosphere 57 (2), 107-126.

Williams, T.D., Diab, A., Ortega, F., Sabine, V.S., Godfrey, R.E., Falciani, F., Chipman, J.K., George, S.G. 2008. Transcriptomic responses of European flounder (Platichthys flesus) to model toxicants. Aquat Toxicol 90 (2), 83-91.

Williams, T.D., Gensberg, K., Minchin, S.D., Chipman, J.K. 2003. A DNA expression array to detect toxic stress response in European flounder (Platichthys flesus). Aquat Toxicol 65 (2), 141-157.

Wishart, D.S., Knox, C., Guo, A.C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B., Hassanali, M. 2008. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res 36 (Database issue), D901-906.

Wishart, D.S., Knox, C., Guo, A.C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., Woolsey, J. 2006. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 34 (Database issue), D668-672.

Wodicka, L., Dong, H., Mittmann, M., Ho, M.H., Lockhart, D.J. 1997. Genome-wide expression monitoring in Saccharomyces cerevisiae. Nat Biotechnol 15 (13), 1359-1367.

Woo, Y., Affourtit, J., Daigle, S., Viale, A., Johnson, K., Naggert, J., Churchill, G. 2004. A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. J Biomol Tech 15 (4), 276-284.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., Speed, T.P. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 30 (4), e15.

Yu, R.M., Wong, M.M., Kong, R.Y., Wu, R.S., Cheng, S.H. 2006. Induction of hepatic choriogenin mRNA expression in male marine medaka: a highly sensitive biomarker for environmental estrogens. Aquat Toxicol 77 (4), 348-358.

Zorita, S., Martensson, L., Mathiasson, L. 2009. Occurrence and removal of pharmaceuticals in a municipal sewage treatment system in the south of Sweden. Sci Total Environ.

Zuccato, E., Castiglioni, S., Fanelli, R., Reitano, G., Bagnati, R., Chiabrando, C., Pomati, F., Rossetti, C., Calamari, D. 2006. Pharmaceuticals in the environment in Italy: causes, occurrence, effects and control. Environ Sci Pollut Res Int 13 (1), 15-21.