

# A White Paper Advocating Complete Sequencing of the Genome of the Rainbow Trout, *Oncorhynchus mykiss*

Gary Thorgaard <sup>a\*</sup>, George Bailey <sup>b</sup>, David Williams <sup>b</sup>, Donald Buhler <sup>b</sup>, Stephen Kaattari <sup>c</sup>, Sandra Ristow <sup>d</sup>, John Hansen <sup>e</sup>, James Winton <sup>f</sup>, Jerri Bartholomew <sup>g</sup>, James Nagler <sup>h</sup>, Patrick Walsh <sup>i</sup>, Matt Vijayan <sup>j</sup>, Robert Devlin <sup>k</sup>, Ronald Hardy <sup>l</sup>, Kenneth Overturf <sup>m</sup>, William Young <sup>n</sup>, Barrie Robison <sup>o</sup>, Caird Rexroad III <sup>p</sup>, Yniv Palti <sup>p</sup>, Bernie May <sup>q</sup>, Scott LaPatra <sup>r</sup>, Ruth Phillips <sup>s</sup>, Linda Park <sup>t</sup>, Takashi Sakamoto <sup>u</sup>, Nobuaki Okamoto <sup>u</sup>, Roy Danzmann <sup>v</sup>, Fred Allendorf <sup>v</sup>, Lars-Erik Holm <sup>x</sup>, Robert Bogden <sup>y</sup>, Patricia Iturra <sup>z</sup>, Rene Guyomard <sup>aa</sup>, Yann Guiguen <sup>aa</sup> (addresses of authors provided at end of paper)  
\*Corresponding author. Tel.: 509-335-7438; Fax: 509-335-3184; Email : thorglab @ wsu.edu

## I. Introduction

Rainbow trout (*Oncorhynchus mykiss*) are among the most intensively studied fish species. Sequencing the trout genome combines a set of diverse advantages that are not available in any other research organism. Such sequence information will provide a scaffold for detailed study of carcinogenesis, comparative immunology, toxicology and the evolutionary process (comparative genomics, evolutionary fate of duplicate genes, genetic architecture of adaptation).

Rainbow trout are a member of the family Salmonidae native to the Pacific coast of North America and Russia, and have been widely introduced around the world. Considerable basic biological knowledge has developed about this species as an outgrowth of their widespread cultivation as a food and sport fish. More is known about the physiology and biology of rainbow trout than any other fish species. Closely related species within the *Oncorhynchus*, *Salmo* and *Salvelinus* genera have also been studied extensively. In the past 20 years there have been over 40,000 reports on the ecology, behavior, physiology and genetics of these species, with rainbow trout specifically being used in half of these studies. The utility of genomic information to understanding the physiology and life history of an organism will depend in part on the applicability of that organism as a model species for experimentation. Rainbow trout have been widely used as model research species, and detailed knowledge of their biology is available with which genomic information may be correlated.

Rainbow trout excel as a physiological and genetic model organism. Although they are large relative to some other model fish species and have fairly long reproductive cycles (2-3 years), they are easy and inexpensive to culture. Trout are more amenable to surgical manipulations than smaller species, and their size allows specific tissues and cell types to be isolated for biochemical, immunological and/or molecular biological analyses. Rainbow trout reproduction is well understood, both under natural and culture conditions, allowing the collection of large numbers of gametes on a year-round basis. Eggs can be fertilized *in vitro* and cultured in simple hatchery environments. In contrast to other fish models, much is known about the natural populations of this species that can serve as a resource for addressing evolutionary questions (Behnke, 1992; Hershberger, 1992). These natural populations provide diverse genetic backgrounds for research.

Trout and salmon, in addition to being important animals in aquaculture and sport and commercial fisheries, are some of the most biologically fascinating vertebrates on earth. Their migrations, ability to move between fresh and salt water and home to specific rivers of origin constitute one of the great puzzles of nature. Genomic studies promise to help to address these and other interesting biological features of these fishes.

In addition to the widespread interest in the natural history and culture of this species, rainbow trout are being used very effectively as a model species to address problems in carcinogenesis, toxicology, comparative immunology and physiology, and evolutionary genetics. As will be discussed, those applications constitute the principal rationale for sequencing the trout genome.

Two fish species for which complete genome sequences will soon be available are the zebrafish, a freshwater species native to south Asia which is smaller than a goldfish, and the pufferfish (*Fugu rubripes*), a marine fish with an unusually small genome. The zebrafish (*Danio rerio*) is a superb experimental species for developmental and genetic research. Although the pufferfish is not widely used as an experimental model, its small genome allows interesting genomic comparisons with other species. However, the distinct advantages of the rainbow trout should not be dismissed. Trout are a very distinct evolutionary lineage from the zebrafish and fugu. Their large size and native distribution in North America have facilitated a wide range of research (Table 1). Sequence information for trout will complement rather than duplicate information from the zebrafish and pufferfish and will ultimately also be of considerable importance to issues related to conservation and agriculture.

TABLE 1. Comparison of numbers of publications in PubMed on rainbow trout, zebrafish and pufferfish during the period 1/1/1997-12/31/2001. Publications related to specific terms were found using the “AND” function. A search using the term “Fugu” in the place of “Pufferfish” yielded similar numbers of publications.

	<u>Rainbow trout</u>	<u>Zebrafish</u>	<u>Pufferfish</u>
Total publications	1426	1858	110
Physiology	1209	1742	101
Hormone	285	128	8
Nutrition	43	6	0
Genetics	524	1461	99
Transgenic	13	19	8
Biochemistry	88	69	3
Evolution	92	260	46
Behavior	51	74	1
Cancer	28	32	2
Carcinogenesis	11	4	0
Toxicology	10	3	0
Toxicity	192	60	4
Immunology	177	53	4
Antibody	154	65	2
Environment	215	86	3

## II. Specific biological rationales for the utility of new sequence data

### Improving human biology. How will the genomic sequence of rainbow trout lead to a better understanding of biological functions in the human?

Rainbow trout are one of the leading vertebrate experimental models in carcinogenesis, toxicology, immunology and evolutionary genetics. Their significance to human biology is in providing a cost-effective vertebrate model for studying fundamental biological processes.

### B. Informing the human sequence. How will the genomic sequence of rainbow trout lead to a better description of the functions of specific sequence features of the human genome?

Although not closely related to humans in evolutionary terms, a significant strength of the trout model is for investigating the effects of gene duplication, considered to be a fundamental process in evolution.

### C. Informing the sequences of model organisms used in the study of human biology. How will the genomic sequence of rainbow trout lead to a better description of the functions of specific sequence features of the genomes of particular model organisms?

As discussed, trout are a leading model organism themselves in the study of carcinogenesis, toxicology, immunology and evolutionary genetics. Their relatively larger size than some of the other significant model fish species (e.g., zebrafish, medaka, *Xiphophorus*) facilitates biochemical studies from isolated cells and tissues and increases the number of gametes available for reproductive and genetic studies

Sequencing the trout genome will also inform the sequence of two other fish species whose genomes will soon be completely sequenced, the fugu and the zebrafish. Comparative DNA sequence studies will provide insights into genome evolution among these species. These three bony fish (teleost) species are believed to have diverged from each other over 100 million years ago and represent three of the most important lineages of bony fishes: the superorder Acanthopterygii and the percomorph fishes (fugu), the superorder Ostariophysi and the cypriniform fishes (zebrafish), and the superorder Protacanthopterygii and the salmonid fishes (rainbow trout) (Helfman et al., 1997). Trout also present a distinctive model by having a relatively recent genome duplication, allowing study of the process of divergence and loss or silencing of genes following a genome-wide duplication. Rainbow trout are the best studied representative of the salmonid fishes that include a group of closely related species whose physiology has been examined in detail. Many of these species are studied both in the lab and the natural environment.

**D. Providing a better connection between the sequences of non-human organisms and the human sequence. How will the genomic sequence of rainbow trout increase our ability to identify orthologs in the sequences of well-studied model organisms and how will that deepen our understanding of the human sequence?**

As discussed above, genomic sequence information in trout will enhance our understanding of the trout model itself as well as the related fugu and zebrafish.

**E. Facilitating our ability to do experiments involving direct genetics or positional mapping.**

Detailed genetic maps have already been developed for rainbow trout (Young et al., 1998; Sakamoto et al., 2000) and these maps have been integrated through the use of common markers to produce a map with over 1000 total genetic markers (Nichols and co-authors, submitted to the journal *Animal Genetics*). Rainbow trout have 58-64 chromosomes and 104 chromosome arms, with the number differences related to Robertsonian rearrangements (Thorgaard, 1983). A high-density genetic map and complete sequence of the genome will be extremely useful resources for associating traits with specific genes. Considerable effort has already been directed toward mapping quantitative trait loci (QTL) using these genetic linkage maps (e.g., Sakamoto et al., 1999; Ozaki et al., 2001; Perry et al., 2001; Robison et al., 2001). High density maps that can be developed from the genome sequencing project are needed for fine mapping of those QTL and for identifying biomedically and economically important QTL (e.g. immune response loci, loci affecting response to carcinogenesis, loci affecting stress response, growth rate, feed conversion, and osmoregulation). These results will also be relevant to other salmonid fishes, including Pacific and Atlantic salmon and brown trout.

An excellent example of the application of such approaches was the recent identification of a single major QTL associated with differences in nonspecific cytotoxic cell (natural killer equivalent) response to mouse YAC-1 cells between two clonal lines of rainbow trout (Zimmerman and Ristow, Washington State University, unpublished results). With the availability of sequence information in the region of this QTL, it should be possible to ascertain homologies and ultimately to identify the specific genes responsible for the difference. This could lead to a better and more fundamental understanding of the origins of innate immunity in vertebrates. Similar differences in other traits are likely to be identified among clonal and outbred lines, can be mapped using QTL analyses and, ultimately, assigned to specific genes using the sequence information.

**F. Expanding our understanding of basic biological processes relevant to human health e.g. developmental biology, neurobiology, cancer biology, stem cell biology.**

The strongest biomedical applications of the rainbow trout model lie in its use for carcinogenesis and toxicology research, and as a comparative immunology model. Excellent and complementary knowledge is also available in the related fields of physiology, nutrition (Hardy, 2002) and disease pathogenesis (Ozaki et al., 2001) as a result of extensive research conducted with rainbow trout in those fields. Fundamental processes such as vision (Julian et al., 1998), olfaction (Laberge and Hara, 2001), exercise physiology (Kieffer, 2000), excretion (Wood, 2001), osmoregulation (Perry et al., 2000) and stress response (Iwama et al., 1998) have been intensively studied. Those fields will also be greatly enhanced by the availability of genome sequence data.

**Carcinogenesis and toxicology:** An epizootic of liver cancer in Pacific Northwest trout hatcheries in the early 1960s contributed to the discovery of aflatoxin B1 (AFB1) as a potential human hepatocarcinogen, and led to the development of the rainbow trout as a sensitive alternative model for cancer research (reviewed in Bailey et al., 1996). The attributes of this model for cancer research include its non-mammalian comparative status, well established husbandry and nutritional requirements, availability of multiple exposure routes, a well defined tumor pathology, externalized gametes and embryos for experimental manipulation, tissue accessibility from animals in the milligram to kilogram size range, low spontaneous tumor background, and high sensitivity. For instance, a single microinjection of as little as 0.5 ng of AFB1 per embryo yields a 40% incidence of hepatocellular/cholangiocellular carcinoma nine months later (Dashwood et al., 1994) – this is 1 billion-fold less aflatoxin than was required to elicit the same incidence in the monkey.

Perhaps the attribute most extensively exploited in the past decade has been the trout's very low husbandry/per diem cost. This feature allows fundamental dose-response issues in carcinogenesis and chemoprevention to be addressed with the trout, using statistically rigorous study designs unaffordable or impossible with traditional rodent models. It was practical, for example, to employ pioneering study designs of up to 10,000 animals each to investigate the quantitative interrelationships between increasing carcinogen dose, increasing anti-carcinogen dose, level of target organ DNA adduction, and eventual

tumor outcome (Dashwood et al, 1989; Breinholt et al., 1995). The latter study was important not only in providing the first demonstration of chlorophyllin (CHL) as an effective blocking agent against AFB1 exposure, but also in demonstrating the conditions under which CHL-mediated reductions in AFB1-DNA adducts were quantitatively predictive biomarkers of eventual reduction in tumor risk. These studies in trout were pivotal to the planning and design of a recent human clinical intervention trial, in which CHL proved equally effective in reducing biomarkers of effective AFB1 uptake and damage in human volunteers in Daxin, China (Egner et al., 2001). Similar trout study designs showed that post-initiation exposure to the dietary supplement indole-3-carbinol provided dose-response promotion of hepatocarcinogenesis, at doses relevant to human exposure, and of a magnitude at least equal to its efficacy as a chemopreventive blocking agent (Dashwood et al, 1991; Oganessian et al., 1999). These findings were eventually proven equally applicable to the rat (Stoner et al., 2002), and continue to influence National Cancer Institute policy regarding human clinical trials with this promising but problematic agent. Perhaps the ultimate application of the low-cost trout model is the recent study utilizing 42,000 trout to establish the relationships between dose of carcinogen (dibenzo[a,l]pyrene, DBP), quantity and spectrum of target organ DNA adducts, cell proliferation, oncogene activation, and tumor response down to an above-background incidence of 0.02% (reviewed in Bailey, 2000; not yet published in detail). This largest-ever cancer dose-response study provides the most extensive low-dose tumorigenesis data yet generated. The central findings of this study are that DBP-DNA adducts decreased linearly over the 500-fold DBP dose range, that PCNA-based liver cell proliferation did not decrease with dose, that the percentage of neoplasms bearing oncogenic Ki-ras mutations was high and invariable with dose, and that above-background tumor incidence showed a strong, statistically significant negative departure from the default EPA LED10 linear extrapolation model for risk assessment.

The rainbow trout is also very useful model for biomonitoring of numerous chemicals in the aquatic environment. Rainbow trout is the standard cold water test species used for regulatory 96-hour acute lethality testing for chemicals entering freshwater aquatic ecosystems. An ASTM standard protocol was established and has been used in the US and elsewhere for over 20 years. Many of these chemicals are highly lipophilic and exhibit bioaccumulation and biomagnification (Kleinow et al., 1987). Examples of such chemical contaminants include polycyclic aromatic hydrocarbons (PAHs), polyhalogenated biphenyls (e.g., PCBs), dioxins (e.g., TCDD) and dibenzofurans (TCDF) and organochlorine pesticides e.g., DDT). One sensitive indicator of such exposures has been the induction of cytochrome P450 (CYP) 1A, particularly in trout liver (Kleinow et al., 1987; Otto et al., 1994; Engwall et al., 1994). The trout CYP1A gene is regulated by the Aryl Hydrocarbon Receptor (AHR) signaling pathway in much the same way as it is in humans (Ma, 2001). A number of P450 genes have been identified and cloned in trout (Buhler and Wang-Buhler, 1998).

Trout exposed to metals respond with the induction of the metal binding protein, metallothionein (Samson and Gedamu, 1995). The assay of metallothionein mRNA or protein has served as a reliable and sensitive biomarker for exposure of trout to metals, a serious class of environmental pollutants (Lange et al., 2002).

In recent years, a great deal of interest has been focused on the potential for environmental chemicals to act as “endocrine disrupters”. Exposures to such chemicals are thought to perhaps play a role in breast cancers in women and declines in sperm quality in men, although the evidence for these effects is equivocal (Safe, 1995). It is well established, however, that feral populations of some aquatic organisms exposed to endocrine disrupters exhibit a number of adverse reproductive responses including lack of development of secondary sex characteristics or even sex reversal (Sumpter, 1995). The class of compounds known as xenoestrogens include *o,p'*-DDT, nonylphenol, hydroxylated PCBs and a number of other significant aquatic pollutants. Trout make an excellent sentinel and model for studying xenoestrogens in the environment as they express a protein in liver and blood that is induced in juveniles of either sex or in adult males. This protein is vitellogenin, an egg yolk precursor protein (Donohoe and Curtis, 1996).

The trout model has made, and will continue to make, important contributions in the study of cancer and its prevention, and in environmental toxicology research. At present there are sufficient gene sequences available to contemplate gene expression profiling using targeted genes in limited macroarray analysis. However, data-rich exploration of unsuspected interactions within and among various developmental and regulatory pathways must await sequencing of the genome.

**Comparative immunology:** Immunogenetic studies of rainbow trout have served two very important functions. The comparable design and function of human and salmonid immune systems make the trout an excellent biomedical model while, as an important agricultural commodity, the same

information becomes of direct and immediate importance to the aquaculture industry. Furthermore, the excellent understanding of disease ecology and disease progression in rainbow trout arising from their widespread rearing in government and commercial hatcheries allows basic studies of the immune system to be complemented by functional studies of response to pathogens. Trout present the best fish model for disease studies because of the well-defined nature of their bacterial and viral pathogens and well-established disease challenge protocols. Immunologically, the availability of clonal/ syngeneic lines of trout, as well as an ample repertoire of immunoreagents and probes, provides investigators with highly advantageous tools to begin extensive genomic analysis of immune system structure and function. Understanding the organization of the immunoglobulin, T cell receptor, major histocompatibility, and nonspecific cytotoxic cell (NCC) gene complexes / loci will enable us to precisely envision the basic repertoires by which an individual recognizes and responds to the gamut of pathogens in its environment. This knowledge of immune capabilities will be required to exploit the unique capabilities of this simpler immune system.

Innate immunity is increasingly recognized as being a central phenomenon in the immune response. Natural killer cells mediate important parts of this response in humans. The natural killer cell equivalent in fishes is known as a nonspecific cytotoxic cell (NCC). The NCC of the rainbow trout appear to be small agranular lymphocytes that are capable of lysing various xenogeneic targets including YAC-1, K-562, and IM-9 by both apoptotic and necrotic mechanisms (Greenlee et al. 1992). Since that original study, Ristow and co-workers have incorporated homozygous clones of rainbow trout into these studies revealing that there are strains of rainbow trout possessing both high and low levels of NCC activity in the peripheral blood (Ristow et al., 1995). Current work has shown that a single QTL may be responsible for this phenomenon in trout (Zimmerman and Ristow, unpublished results). This capability is now being mapped to a specific genetic linkage group of the rainbow trout. We are now well poised with the trout system to dissect one of the most critical elements of innate immunity. Further, much of that which can be modeled within the murine system can be studied with this alternate vertebrate model. Such an alternate model, with a potentially simpler immune system and a slower rate of response, could provide fundamental insight into the innate immune response of humans and other vertebrate species.

The MHC of trout has also been studied in a comparative manner. Class I and II sequences (mainly cDNAs) have been cloned from a variety of teleost species including zebrafish (Takeuchi et al., 1995), salmonids (Grimholt et al., 1993; Hansen et al., 1996; Shum et al., 1999), cod (Persson et al., 1999), catfish (Godwin et al., 2000; Antao et al., 2001) and pufferfish (Timon et al., 1998). Fine mapping experiments for the teleost MHC are limited to date but this constitutes an active area of research. The most unexpected discovery for MHC architectural arrangements came from studies in bony fish where the MHC class I and II loci do not co-segregate and, in fact, reside on different chromosomes altogether (Takami et al., 1997; Hansen et al., 1999) which differs from all other vertebrate classes including cartilaginous fish (sharks, rays and chimeras) (Flajnik and Kasahara, 2001; Ohta et al., 2002). The take home message for the teleost MHC is that the MHC class Ia antigen processing and presentation machinery have been maintained as a solid genetic linkage group during the course of vertebrate evolution. The tight linkage of members of the MHC Ia antigen presentation pathway and the lack of class II linkage in fish present an ideal model for selecting class I and II haplotypic combinations possessing enhanced disease resistance qualities. In support of this idea, zebrafish and fugu show conserved synteny (gene content but not strict order) between the human and teleost MHC class I region for >30 different loci, not including duplicate genes (Michalova et al., 2000; Clark et al., 2001). Clearly, our understanding of the MHC and other components of the immune system of trout and other fishes will advance dramatically with detailed sequence information in those regions.

Investigation of other aspects of the specific immune response in fish will also be aided by genomic sequence data. To date, there have been no reported studies that attempted to determine the  $V_H$  genes utilized in the production of specific antibodies in fish. All cloning and sequencing efforts, thus far, have employed primers derived from various consensus sequences, without regard to their probable specificity, and which are often based on sequences from phylogenetically distant species (Roman and Charlemagne, 1994; Andersson and Matsunaga, 1995; Roman et al., 1996). Lewis (2000) determined that specific gene families are employed in the production of specific antibodies to TNP (Family V) and the G protein of infectious hematopoietic necrosis virus (Family IV). Of particular interest was the observation that Family V was very small (approximately 3 - 5 members), while Family IV was much larger ~ 30 members, prompting consideration of the possibility that intense selection by a virulent virus may promote the expansion of antibody gene families. Thus the complete structure of these gene complexes can do much to increase understanding of the evolutionary processes that mold the capacity of an individual's immune response.

### **G. Expanding our understanding of evolutionary processes (biological innovation, selection) in general, and human evolution in particular.**

The rainbow trout provides an exceptional opportunity to study genome evolution following tetraploidy. The rainbow trout, and all other species in the family Salmonidae, are descended from a single tetraploid event that occurred approximately 25 million years ago (Allendorf and Thorgaard 1984). Postlethwait et al. (1998) have discovered large conserved chromosome segments in the linkage maps of zebrafish and mammals and concluded that two polyploidization events occurred in a common ancestor before the divergence of fish and mammals, resulting in four paralogous copies of each chromosome segment in each lineage. Amores et al. (1998) have suggested that an additional polyploid event is responsible for the 5-8 *HOX* gene clusters in teleost fish in comparison to the 4 *HOX* gene clusters in mammals. Extensive sequencing supports this view of the importance of gene duplication in the evolution of the vertebrate genome (Henikoff et al. 1997; Rubin et al. 2000). Thus, chromosomal and gene duplication by polyploidy also has played an essential role in evolution of the genome of vertebrates. The evolutionarily recent tetraploid event specific to the salmonid lineage provides an exceptional opportunity to study the evolution of chromosomes and genes following polyploidy in a vertebrate.

The presence of duplicate genes can complicate the interpretation of genetic information because very similar sequences and gene products may, in fact, represent different genes. Genic inheritance is complicated in rainbow trout relative to the elegant simplicity of diploid Mendelian inheritance because of their polyploid ancestry (Allendorf and Danzmann 1997). The positive aspect of the complexity of salmonid genomes is that it provides a natural laboratory for following the process of evolution by gene duplication. This process, which is generally acknowledged to be a central one in evolution (Ohno, 1970; Force et al., 1999), is still taking place in these fishes. Only disomic inheritance has been detected in females. However, some loci are inherited tetrasomically in males from some populations but disomically in males from other populations. The salmonid duplication is more extensive and recent than the similar event which has been studied in the zebrafish (Amores et al., 1998). Many of the duplicate pairs of protein coding loci produced by the salmonid-specific polyploid event show divergent patterns of tissue-specific expression (Allendorf and Thorgaard, 1984). In addition, polymorphisms in tissue-specific patterns of gene expression (Allendorf et al., 1983; Danzmann et al., 1985) and gene silencing (Ferguson et al., 1988; Leary et al. 1993) occur in rainbow trout at some of these duplicate genes. Duplicate gene copies are beginning to be studied at the DNA sequence level (e.g., Brunelli et al., 2001). For example, in the duplicate growth hormone genes in rainbow trout, the only estrogen response element has been inactivated in one of the duplicates, so there is a sex difference in expression of these genes (Yang et al., 1997; Phillips, unpublished results). This difference also occurs in the Pacific salmon, but not in Atlantic salmon. Salmonid fishes are an ideal group to study evolution of regulatory elements in a phylogenetic context because there is a large body of information available on their physiology and natural history. Wider-scale sequence information for rainbow trout will provide an excellent and distinctive system for studying the aftermath of a genome-wide duplication event and the associated structural and regulatory gene changes. The extension of current bioinformatics approaches (e.g., the Lek algorithm of Venter et al. 2001) will likely be needed to identify and characterize the internal genomic homologies resulting from this tetraploid ancestry and obtain full benefit from this dataset.

The evolution of sex chromosomes and sex determination is another fruitful research area with the trout model. The X/Y sex determination system of rainbow trout is distinctive because of the existence of among- population variation in sex chromosome morphology, suggesting that they are in the early stages of differentiation (Thorgaard, 1983). Analyzing this variation at the sequence level will provide important knowledge about the earliest stages of sex chromosome evolution in vertebrates, something that is not possible in the highly derived mammalian models. Similarly, locating and isolating the factors responsible for sex determination in an easily manipulated model species such as rainbow trout will lead to beneficial comparative research into this important but inadequately studied aspect of vertebrate development.

## **H. Providing additional surrogate systems for human experimentation. (e.g., new disease models, improved opportunities for drug testing, or other medical procedures, such as transplantation).**

As discussed elsewhere in this white paper, the trout presents an excellent experimental system for studies of carcinogenesis, toxicology and comparative immunology. It provides a cost-effective system for screening chemicals for toxic and carcinogenic effects. In addition, because natural populations of salmonids are widespread in North America and other temperate parts of the world, these fishes can serve as environmental monitors. The recent observation of apparent phenotypic sex reversal in natural populations of chinook salmon (Nagler et al. 2001) illustrates this potential.

## **III. Strategic issues in acquiring new sequence data**

### **A. The demand for the new sequence data. What is the size of the research community that will use it? What is the community's enthusiasm for having the sequence? Will the new sequence data stimulate the expansion of the research community?**

Rainbow trout are one of the most intensively studied of fish species. They were the first fish species selected for a species symposium by the journal "Aquaculture" (Gall, 1992). Researchers in a number of countries, including the US, Canada, United Kingdom, Japan, Norway, Denmark, France and Chile are currently conducting genetic, carcinogenesis, toxicology, reproductive physiology and immunology research with this species. Many of those researchers are co-authors of this white paper. A group of researchers meets annually at the Plant and Animal Genome meeting in San Diego to discuss coordination of trout genomic research. Discussions at the January 2002 meeting led to preparation of this white paper. All of the co-authors have had the opportunity to comment on drafts of this white paper and many provided substantive contributions. We anticipate that the availability of full genome sequence data for rainbow trout will expand the community and lead to increased research with this attractive model species.

### **B. The suitability of the organism for experimentation. How will the new sequence data enhance the experimental use of rainbow trout? What genomic resources and technologies (e.g., gene transfer, ability to go from molecule to mutation) are available that will allow the new sequence information to be effectively used?**

The rainbow trout is the leading model fish species inhabiting cool and cold waters. Major advantages for using rainbow trout for genetic research include: (1) High fecundity. - A single 3 to 4 kg female produces 3000 eggs, and a very large amount of sperm (10 ml) is available from a single male. (2) Ease of manipulation and handling of gametes. - This enables transportation (air shipment world-wide) of gametes before and after fertilization from a broodstock holding facility to grow out farms or research institutions. The hardiness of rainbow trout gametes also facilitates chromosomal set manipulations by means of heat shock or hydrostatic pressure shock. These methods are widely utilized by farmers and scientists. Gynogenesis is used for producing all-female populations and to prevent deterioration of flesh quality due to early sexual maturation of males, and induced triploidy is employed for producing sterile fish to reduce genetic contamination of wild populations by farmed fish (Thorgaard, 1992; Palti et al., 1997). Androgenesis produces true homozygous, clonal lines in two generations (Young et al., 1996). (3) Year round supply of gametes due to a wide natural variability in spawning time (Siitonen and Gall, 1989; Sakamoto et al., 1999) and the use of photoperiod control. (4) Well-developed and widely-used methods for rainbow trout semen cryopreservation (e.g. Wheeler and Thorgaard, 1991), which allows for crosses between strains that spawn at different seasons and between different generations (e.g. for backcrossing mating design). It is also used for conservation of germplasm from wild populations and from economically important strains.

The resources for effectively utilizing rainbow trout for genomic studies are in place or being developed. Five clonal lines of rainbow trout have recently been established using the chromosome set manipulation methods of androgenesis and gynogenesis. These lines provide valuable experimental uniformity (well-established as an advantage by research with inbred lines of mice) and will provide opportunities for analysis and genetic dissection of traits as differences among the lines are identified (Ristow et al., 1995; Robison et al., 1999) and genetically characterized (Robison et al., 2001). Detailed genetic linkage maps are also developing for this species (Young et al., 1998; Sakamoto et al., 2000), with one of the maps (Young et al., 1998) being based on a cross of two of the clonal lines.

The mitochondrial genome of rainbow trout has been fully sequenced (Zardoya et al., 1995). There are likely to be excellent opportunities to examine the effect of mitochondrial variation on development and physiology because, using androgenesis, it is possible to create lines with the same nuclear but different mitochondrial genomes (Brown and Thorgaard, 2002).

Over 250 polymorphic microsatellite markers have been developed for rainbow trout at the USDA/ARS National Center for Cool and Cold Water Aquaculture (NCCCWA) in the past 2 years (e.g. Rexroad et al., 2001). Approximately half of the microsatellites developed for Atlantic salmon work for rainbow trout (Moira Ferguson, personal communication). Microsatellites developed for the Pacific salmon (*Oncorhynchus*) species also generally work for rainbow trout. This brings the total number of microsatellites currently available to well over 500.

The total number of nucleotide sequences for rainbow trout, which were deposited on the NCBI databases as of April 30, 2002, was 2,200, of which 632 were expressed sequence tags (ESTs). These figures are far behind other fish model organisms such as zebrafish and medaka (245,000 and 46,700 nucleotide sequences, respectively). A major effort is now focused on increasing the rainbow trout EST database. Forty five thousand clones from a normalized cDNA library, which was derived from gill, liver, brain, kidney, spleen and muscle are currently being sequenced to produce additional ESTs in another efforts at the NCCCWA. The INRA (Institut National de la Recherche Agronomique) trout genome and transcriptome project is expected to complete a sequencing project of 120,000 ESTs (60,000 clones in the 5' and 3' directions) by mid 2003. Their cDNA library is derived from liver, interrenal, brain, muscle, blood, intestine, ovary, testis, differentiating gonads, adipose tissue, gills, pituitary and kidney. A radiation hybrid panel is being constructed by the INRA group and is expected to be completed by the end of this year. It should be useful for rapid linkage mapping of the ESTs (Rene Guyomard, personal communication).

Four BAC libraries of the rainbow trout genome have been constructed to date. Two were constructed in Japan by Katagiri et al. (2001). These contain an average insert size of 58 kb and 110 kb, and provide haploid genome coverage of 6.7 fold and 5.3 fold, respectively. Two BAC libraries from the OSU female homozygous line and the Swanson male homozygous line were commercially constructed by Amplicon Express Inc.. The OSU BAC library has 96,768 clones arrayed in 384 well plates with an average insert size of 110 kb (haploid genome coverage of 4.5 fold). The Swanson BAC library has 184,704 clones arrayed in 384 well plates with an average insert size of 137 kb (haploid genome coverage of over 10 fold). BACs from the OSU library have been used as probes in fluorescence in situ hybridization to anchor the genetic linkage map to specific chromosomes (Philips, 2001). The Swanson library is now being characterized with 20 genes of interest. Both libraries will be used for mapping of type I markers (ESTs), and can be used to produce a sequence-ready BAC contig map.

Gene transfer is well-developed for the trout model. The transfer of new genetic information into species has proven and will continue to be a powerful tool for studying the physiological, phenotypic, and fitness consequences of specific genes (Pereira, 2000; Tyagi and Mohanty, 2000). The first transgenic fish was reported in 1985 (Zhu et al., 1985). Currently, over 3000 reports of transgenic aquatic organisms exist in the literature, testimony to the large academic and commercial interest in this field. A great deal of research has been undertaken with model fish species, including zebrafish and medaka, as well as with salmonids, carps, catfish, tilapia, loach, and shellfish (Bachere et al., 1997; Sin, 1997; Ivetac et al., 2000; Maclean and Laight, 2000; Hew and Fletcher, 2001). Among fish species used in transgenic studies, rainbow trout is among the most commonly-cultured and widely geographically-distributed species, being grown worldwide for academic study and for aquaculture. To date, germ-line transgenic trout have been produced by microinjection of gene constructs into the fertilized egg shortly after fertilization in numerous studies (see Devlin et al., 2001; Hew and Fletcher, 2001 for references). Other approaches utilizing sperm-mediated incubations, biolistic (gene gun), or electroporation methods have been explored in fish, but have produced equivocal or negative results (Chourrout and Perrot, 1992; Sin, 1997), and no confirmed cases of germ-line transformation of trout to date. Approximately  $10^6$  to  $10^7$  copies of linear DNA are microinjected in a volume of 1-2nL into germinal disk of the egg using ground glass needles of approximately 5 $\mu$ m in diameter. For microinjection, DNA is injected cytoplasmically, and is incorporated into nuclei following mitotic division and nuclear membrane dissolution and reforming. The reported frequency of DNA integration into the trout genome varies widely (see Devlin, 1997), in part due to different criteria used to define a transgenic organism. Although mosaic integration in founder animals is common, germline integration events occur at reasonable frequencies which allow individual families to be readily formed for subsequent analysis of defined transgene inserts.

Other approaches including targeted integration (Shashikant et al., 1998; Copeland et al., 2001) and gene knockouts (Harris, 2001) are possible in other germline transgenic vertebrates, and are currently being experimentally explored but not yet achieved in fish. However, recent progress in the development of embryonic stem (ES) cells and blastomere transplantation in trout and other fish species (Nilsson and Cloud, 1992; Takeuchi et al., 2001) suggests that targeted gene replacement in the genome

of trout will be available soon and provide a powerful technology for functional genomic analysis.

Transgenic approaches will allow detailed examination of gene function in many areas, including endocrinology, physiology, neurobiology, metabolism, nutrition, immunology, and developmental biology (Abe et al., 1998; Waller et al., 1998; Ko, 2001; Trethewey, 2001; Fiehn, 2002). To date, transgenic trout have already been produced in several of these research areas, and a great deal of information has been gathered regarding methodology and the function of a few well-studied gene systems. However, historically, such activities have been hampered by the availability of genes to study, relying on those isolated by traditional molecular biology approaches. The emergence of genomic technology will remove this limitation, and will make available a plethora of gene sequences for study. The function of candidate genes identified in this way can be directly tested in many cases using transgenic technologies by examining phenotypes resulting from the altered expression of coding regions. Coding regions identified by proteomic approaches can similarly be tested for specific functions using transgenic approaches, as can coding regions whose function remains obscure after traditional molecular biological analyses.

**C. The rationale for the complete sequence of the organism. Why would the complete sequence be more useful than the sequences of specific regions, or only the coding sequences, or only ESTs? Are there alternative ways to get the necessary information?**

The complete sequence of the genome is needed for effective positional cloning and identification of novel candidate genes for QTL. Complete sequencing would also allow detailed exploration of unsuspected interactions of pathways related to carcinogenesis, response to toxicants and the immune response through gene expression and microarray methods. Detailed analysis of the process of evolution following full genome duplication, including comparison of rates of evolution of non-coding, regulatory and coding regions, must also await the availability of a full genome sequence.

A complete genome dataset will be also needed for effective genomic comparisons with the zebrafish and fugu, which are currently being sequenced. In plants, comparison of sequenced segments of the tomato and *Arabidopsis* genomes indicated that duplications and gene loss creates a network of synteny (Ku et al., 2000). On the other hand, large regions of conserved synteny between mammalian genomes were used to enhance genomic mapping of mammals (O'Brien et al., 1999). Due to the large evolutionary distance between trout, zebrafish and fugu, it is likely that comparative genomics between those species will not be very useful for fine mapping, but rather for identifying conserved regulatory elements and general genomic organization.

**D. The cost of sequencing the genome and the state of readiness of the organism's DNA for sequencing. What is the size of the genome? What quality of sequence product is needed (finished sequence? draft? full shotgun)? What sequencing strategy will be used? Is suitable DNA readily available?**

Rainbow trout have a genome size which is 80% that of humans ( $2.4 \times 10^9$  bp). We anticipate utilizing the existing BAC libraries to develop an arrayed BAC contig library. This would then be utilized to develop a finished product. BAC contigs will be very important to develop because the high homology between homeologous chromosome pairs will be a challenge in the sequence assembly, and assembly will be greatly facilitated if there is a BAC contig backbone to establish the linear relationship between the sequences. The availability of defined clonal lines and BAC libraries from two clonal lines should allow progress to be made rapidly. A target cost for this effort could be \$80,000,000.

**E. Are there other (partial) sources of funding available or being sought for this sequencing project?**

An NIH-funded effort could benefit from complementary genomics efforts being conducted with the support of NIH, NSF and USDA extramural funds and intramural USDA funds in the US, and government funding in Japan, Canada, United Kingdom, Norway, Denmark, France and Chile. EST databanks in the US and France and a radiation hybrid panel developed in France are available as complementary resources. This represents a significant "value-added" research opportunity.

## References

- Abe, K., Ko, M.S.H., Macgregor, G.R., 1998. A systematic molecular genetic approach to study mammalian germline development. *Int. J. Dev. Biol.* 47, 1051-1065.
- Allendorf, F. W., Thorgaard, G.H., 1984. Tetraploidy and the evolution of salmonid fishes. In: *The Evolutionary Genetics of Fishes*, Turner, B.J. (Ed.) Plenum Press, New York, pp. 1-53.
- Allendorf, F.W., Danzmann, R.G., 1997. Secondary tetrasomic segregation of MDH-B and preferential pairing of homeologues in rainbow trout. *Genetics* 145, 1083-1092.
- Allendorf, F.W., Knudsen, K.L., Leary, R.F., 1983. Adaptive significance of differences in the tissue-specific expression of a phosphoglucosyltransferase gene in rainbow trout. *Proc. Nat. Acad. Sci. U.S.A.* 80, 1397-1400.
- Amores, A. Force, A., Yan, Y.L., Joly, L., Amemiya, C. Fritz, A., Ho, R.K., Langeland, J. Prince, V., Wang, Y.L., Westerfield, M., Ekker, M. and Postlethwait, J., 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science* 282, 1711-1714.
- Andersson, E., Matsunaga, T., 1995. Evolution of immunoglobulin heavy chain variable region genes: a  $V_H$  family can last for 150- 200 million years or longer. *Immunogen.* 41, 210 – 216.
- Antao, A. B., Wilson, M., Wang, J., Bengten, E., Miller, N.W., Clem, L.W., Chinchar.V.G., 2001. Genomic organization and differential expression of channel catfish MHC class I genes. *Dev Comp Immunol* 25, 579-595.
- Bachere, E., Cedeno, V., Rousseau, C., Destoumieux, D., Boulo, V., Cadoret, J.P., Mialhe, E., 1997. Transgenic crustaceans. *World Aquacult.* 28, 51-55.
- Bailey, G.S., 2000. Mechanisms of carcinogenesis and chemoprevention at low dose. In Cohen, S.M., Hino, O., Ishikawa, T., Nakagama, H., Shirai, T., Sugimura, T (Eds.) *New Frontiers in Mechanistic Cancer Research in Animal Models*, Princess Takamatsu Cancer Research Fund, Tokyo. pp. 67-71.
- Bailey, G.S., Williams, D.E., Hendricks, J.D., 1996. Fish models for environmental carcinogenesis: the rainbow trout. *Environ. Health Persp.* 104 suppl. 1, 5-21.
- Behnke, R.J., 1992. *Native Trout of Western North America*. American Fisheries Society Monograph 6. American Fisheries Society, Bethesda, MD.
- Breinholt, V., Hendricks, J., Pereira, C., Arbogast, D., Bailey, G., 1995. Dietary chlorophyllin is a potent inhibitor of aflatoxin B1 hepatocarcinogenesis in rainbow trout. *Cancer Research* 55, 57-62.
- Brown, K.H., Thorgaard, G.H., 2002. Mitochondrial and nuclear inheritance in an androgenetic line of rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 204, 323-335.
- Brunelli, J.P., Robison, B.D., Thorgaard, G.H., 2001. Ancient and recent duplications of the rainbow trout Wilms' tumor gene. *Genome* 44, 455-462.
- Buhler, D. R., Wang-Buhler, J.-L., 1998. Rainbow trout cytochrome P450s : purification, molecular aspects, metabolic activity, induction, and role in environmental monitoring. *Comp. Biochem. Physiol.* C121, 107-137.
- Chourrout, D., Perrot, E., 1992. No transgenic rainbow trout produced with sperm incubated with linear DNA. *Mol. Mar. Biol. Biotech.* 1, 282-285.
- Clark, M. S., L. Shaw, A. Kelly, P. Snell, and G. Elgar. 2001. Characterization of the MHC class I region of the Japanese pufferfish (*Fugu rubripes*). *Immunogenetics* 52, 174-185.
- Copeland, N.G., Jenkins, N.A., Court, D.L., 2001. Recombineering: A powerful new tool for mouse functional genomics. *Nature Reviews Genetics* 2, 769-779.
- Danzmann, R. G., M. M. Ferguson, and F. W. Allendorf. 1985. Allelic differences in initial expression of paternal alleles at an isocitrate dehydrogenase locus in rainbow trout (*Salmo gairdneri*). *Develop Genet.* 5:117-128.
- Dashwood, R., Arbogast, D.N., Fong, A.T., Pereira, C., Hendricks, J.D., Bailey, G.S., 1989. Quantitative inter-relationships between aflatoxin B1 carcinogen dose, indole-3-carbinol anti-carcinogen dose, target organ DNA adduction and final tumor response. *Carcinogenesis* 10, 175-181.
- Dashwood, R.H., Fong, A.T., Williams, D.E., Hendricks, J.D., Bailey, G.S., 1991. Promotion of aflatoxin B1 carcinogenesis by the natural tumor modulator indole-3-carbinol (I3C): influence of dose, duration, and intermittent exposure on I3C promotional potency. *Cancer Research* 51, 2362-2365.
- Dashwood, R., Fong, A., Arbogast, D., Bjeldanes, L., Hendricks, J., Bailey, G., 1994. Anticarcinogenic activity of indole-3-carbinol acid products: ultrasensitive bioassay by trout embryo

- microinjection. *Cancer Res.* 54, 3617-3619.
- Devlin, R.H., 1997. Transgenic salmonids. In: Houdebine, L.M. (Ed.), *Transgenic animals: generation and use*. Harwood Academic Publishers, Amsterdam, The Netherlands, pp. 105-117.
- Devlin, R.H., Biagi, C.A., Yesaki, T.Y., Smailus, D.E., Byatt, J.C., 2001. Growth of domesticated transgenic fish. *Nature* 409, 781-782.
- Donohoe, R.M., Curtis, L.R., 1996. Estrogenic activity of chlordecone, o,p'-DDT and o,p'-DDE in juvenile rainbow trout: induction of vitellogenesis and interaction with hepatic estrogen binding sites. *Aquatic Toxicol.* 36, 31-52.
- Egner, P., Wang, J., Zhu, Y., Zhang, B., Wu, Y., Zhang, Q., Qian, G., Kuang, S., Gange, S., Jacobson, L., Munoz, A., Helzlsouer, K., Bailey, G., Groopman, J., Kensler, T., 2001. Chlorophyllin reduces urinary levels of a carcinogen-DNA adduct biomarker in residents of Qidong, Peoples' Republic of China in a randomized, double-blind, placebo-controlled trial. *Proc. Natl. Acad. Sci. USA* 98, 14601-14606.
- Engwall, M., Brunstrom, B., Brewer, A., Norrgren, L., 1994. Cytochrome P450IA induction by a coplanar PCB, a PAH mixture and PCB-contaminated sediment extracts following microinjection of rainbow trout sac-fry. *Aquatic Toxicol.* 30, 311-324.
- Ferguson, M.M., Knudsen, K.L., Danzmann, R.G. Allendorf, F.W., 1988. Developmental rate and viability of rainbow trout with a null allele at a lactate dehydrogenase locus. *Biochem.Genet.* 26, 177-189.
- Fiehn, O., 2002. Metabolomics: The link between genotypes and phenotypes. *Plant Molec. Biol.* 48, 155-171.
- Flajnik, M. F., and M. Kasahara. 2001. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity* 15, 351-362.
- Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, U.L. Postlethwait, J., 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531-1545.
- Gall, G.A.E., ed. 1992. *The Rainbow Trout. The Proceedings of a Symposium held at Stirling, Scotland. Aquaculture, Volume 100.* Elsevier, New York.
- Godwin, U. B., M. Flores, S. Quiniou, M. R. Wilson, N. W. Miller, L. W. Clem, and T. J. McConnell. 2000. MHC class II A genes in the channel catfish (*Ictalurus punctatus*). *Dev. Comp. Immunol.* 24, 609-622.
- Greenlee, A. R., Ristow, S. S. 1992. Detection of apoptotic killing by trout nonspecific cytotoxic cells using the DNA fragmentation assay and agarose gel electrophoresis In : Stolen , J. S et al.(Eds.). *Techniques in Fish Immunology.* SOS Publications, Fair Haven, NJ, pp. 97 - 105.
- Grimholt, U., Hordvik, I., Fosse, V.M., Olsaker, I., Endresen, C., Lie, O., 1993. Molecular cloning of major histocompatibility complex class I cDNAs from Atlantic salmon (*Salmo salar*). *Immunogenetics* 37, 469-473.
- Hansen, J. D., Strassburger, P., Du Pasquier, L., 1996. Conservation of an alpha 2 domain within the teleostean world, MHC class I from the rainbow trout *Oncorhynchus mykiss*. *Dev. Comp. Immunol.* 20, 417-425.
- Hansen, J. D., Strassburger, P., Thorgaard, G.H., Young, W.P., Du Pasquier, L., 1999. Expression, linkage, and polymorphism of MHC-related genes in rainbow trout, *Oncorhynchus mykiss*. *J Immunol* 163, 774-786.
- Hardy, R.W. 2002. Rainbow trout, *Oncorhynchus mykiss*. In: Webster, C.D., Lin, C.E. (Eds.), *Nutrient Requirements and Feeding of Finfish for Aquaculture*, CABI Publishing, New York, NY., pp. 184-202.
- Harris, S., 2001. Transgenic knockouts as part of high-throughput, evidence-based target selection and validation strategies. *Drug Discovery Today* 6, 628-636.
- Helfman, G.S., Collette, B.B., Facey, D.E., 1997. *The Diversity of Fishes.* Blackwell, London.
- Henikoff, S., E. A. Greene, S. Pietrokovski, P. Bork, T. K. Attwood, and L. Hood. 1997. Gene families: The taxonomy of protein paralogs and chimeras. *Science* 278:609-614.
- Hershberger, W.K., 1992. Genetic variability in rainbow trout populations. *Aquaculture* 100, 51-72.
- Hew, C.L., Fletcher, G.L., 2001. The role of aquatic biotechnology in aquaculture. *Aquaculture* 197, 191-204.
- Ivetac, I., Becanovic, J., Krishnapillai, V., 2000. Zebrafish: genetic tools and genomics. *Asia-Pacific J. Mol. Biol. Biotechnol.* 8, 1-11.
- Iwama G.K, Thomas, P.T., Forsyth, R.B., Vijayan, M.M., 1998 Heat shock protein expression in fish. *Rev. Fish Biol. Fisher.* 8, 35-56.
- Julian, D., Ennis, K., Korenbrot, J.I., 1998. Birth and fate of proliferative cells in the inner nuclear layer of the mature fish retina. *J. Comp. Neurol.* 394(3), 271-82.

- Katagiri, T., Asakawa, S., Minagawa, S., Shimizu, N., Hirono, I., Aoki, T., 2001. Construction and characterization of BAC libraries for three fish species; rainbow trout, carp and tilapia. *Animal Genetics* 32, 200-204.
- Kieffer, J.D. 2000. Limits to exhaustive exercise in fish. *Comp. Biochem. Physiol.* 126A: 161-179.
- Kleinow, K.M., Melancon, M.J, Lech, J.J. 1987. Biotransformation and induction: implications for toxicity, bioaccumulation and monitoring of environmental xenobiotics in fish. *Environm. Hlth. Perspec.* 71, 105-119.
- Ko, M.S.H., 2001. Embryogenomics: Developmental biology meets genomics. *Trends Biotech.* 19, 511-518.
- Ku, H.M., Vision, T., Liu, J.P., Tanksley, S.D., 2000. Comparing sequenced segments of the tomato and *Arabidopsis* genomes: Large-scale duplication followed by selective gene loss creates a network of synteny. *Proc. Natl. Acad. Sci. USA* 97, 9121-9126.
- Laberge, F., Hara, T.J., 2001. Neurobiology of fish olfaction: a review. *Brain Res. Rev.* 36, 46-59.
- Lange, A., Ausseil, O., Segner, H., 2002. Alterations of tissue glutathione levels and metallothionein mRNA in rainbow trout during single and combined exposure to cadmium and zinc. *Comp. Biochem. Physiol.* 131C, 231-243.
- Leary, R. F., F. W. Allendorf, and K. L. Knudsen. 1993. Null allele heterozygotes at two lactate dehydrogenase loci in rainbow trout are associated with decreased developmental stability. *Genetica* 89:3-13.
- Lewis, T. D. 2000. The specific immune response in rainbow trout. Somatic hypermutation and V<sub>H</sub> gene utilization. Ph.D. Thesis. College of William and Mary. 223 pp
- Ma, Q., 2001. Induction of CYP1A1. The AhR/DRE paradigm: transcription, receptor regulation, and expanding biological roles. *Curr. Drug Metabol.* 2, 149-164.
- Macleay, N., Laight, R.J., 2000. Transgenic fish: an evaluation of benefits and risks. *Fish Fish.* 1, 146-172.
- Michalova, V., Murray, B.W., Sultmann, H., Klein, J., . 2000. A contig map of the Mhc class I genomic region in the zebrafish reveals ancient synteny. *J Immunol* 164, 5296-5305.
- Nagler, J.J., Bouma, J., Thorgaard, G.H., Dauble, D.D., 2001. High incidence of a male-specific genetic marker in phenotypic female chinook salmon from the Columbia River. *Environ. Health Perspect.* 109: 67-69.
- Nilsson, E.E., Cloud, J.G., 1992. Rainbow trout chimeras produced by injection of blastomeres into recipient blastulae. *Proc. Natl. Acad. Sci. USA* 89, 9425-9428.
- O'Brien S.J., Menotti-Raymond M., Murphy W.J., Nash W.G., Wienberg J., Stanyon R., Copeland N.G., Jenkins N.A., Womack J.E., Graves J.A.M., 1999. The promise of comparative genomics in mammals. *Science* 286, 458-462.
- Oganesian, A., Hendricks, J.D., Pereira, C.B., Orner, G.A., Bailey, G.S. and Williams, D.E. , 1999. Potency of dietary indole-3-carbinol as a promoter of aflatoxin B1-initiated hepatocarcinogenesis: Results from a 9000 animal tumor study. *Carcinogenesis* 20, 453-458.
- Ohno, S., 1970. *Evolution By Gene Duplication.* Springer Verlag, Berlin/ Heidelberg/ New York.
- Ohta, Y., McKinney, E.C., Criscitiello, M.F. Flajnik, M.F. 2002. Proteasome, transporter associated with antigen processing, and class I genes in the nurse shark *Ginglymostoma cirratum*: evidence for a stable class I region and MHC haplotype lineages. *J. Immunol.* 168, 771.
- Otto, D.M., Linstrom-Seppa, P. and Sen, C., 1994. Cytochrome P450-dependent enzymes and oxidant-mediated responses in rainbow trout exposed to contaminated sediments. *Ecotoxicol. Environm. Safety* 27:265-280.
- Ozaki, A., Sakamoto, T., Khoo, S., Nakamura, N., Coimbra, M. R. M., Akutsu, T., Okamoto, N., 2001. Quantitative trait loci (QTLs) associated with resistance /susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Mol. Genet. Genomics* 265, 23-31.
- Palti, Y., Li J.J., Thorgaard, G.H, 1997. Improved efficiency of heat and pressure shocks for producing gynogenetic rainbow trout. *The Progressive Fish-Culturist* 59, 1-13.
- Pereira, A., 2000. A transgenic perspective on plant functional genomics. *Transgenic Research* 9, 245-260.
- Perry, G.M.L., Danzman, R.G., Ferguson, M.M., Gibson, J.P., 2001. Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*). *Heredity* 86, 333-341.
- Perry, S.F., Beyer, M.L., Johnson, D.A., . 2000. Cloning and molecular characterization of the trout vacuolar H<sup>+</sup>-ATPase B subunit. *J. Exp. Biol.* 203, 459-470.
- Persson, A. C., Stet, R.J., Pilstrom., L., 1999. Characterization of MHC class I and beta(2)-microglobulin sequences in Atlantic cod reveals an unusually high number of expressed class I

- genes. *Immunogenetics* 50, 49-59.
- Phillips, R., 2001. Application of fluorescence in situ hybridization (FISH) to fish genetics and genome mapping. *Mar. Biotech.* 3, S145-S152.
- Postlethwait, J. H., Y-L. Yan, M. A. Gates, and 26 additional authors. 1998. Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics* 18:345-349.
- Rexroad CE III, Coleman, RL, Martin, AM, Hershberger, WK, Killefer, J (2001). Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). *Animal Genetics* 32, 316-331.
- Ristow, S.S., Grabowski, L.D., Wheeler, P.A., Prieur, D.J., Thorgaard, G.H., 1995. Arlee line of rainbow trout (*Oncorhynchus mykiss*) exhibits a low level of nonspecific cytotoxic cell activity. *Dev. Comp. Immunol.* 19: 497-505.
- Robison, B.D., Wheeler, P.A., Thorgaard, G.H., 1999. Variation in development rate among clonal lines of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 173, 131-141.
- Robison, B.D., Wheeler, P.A., Sundin, K., Sikka, P., Thorgaard, G.H., 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *J. Hered.* 92, 16-22.
- Roman, T., Charlemagne, J. 1994. The immunoglobulin repertoire of the rainbow trout (*Oncorhynchus mykiss*): definition of nine Igh-V families. *Immunogen.* 40, 210 - 216.
- Roman, T., Andersson, E. Hansen, J., Kaattari, S. Pilstrom, L. Charlemagne, J. Matsunaga, T. 1996. Unified nomenclature of Ig VH genes in rainbow trout (*Oncorhynchus mykiss*): definition of 11 V<sub>H</sub> families. *Immunogen.* 43, 325 - 326.
- Rubin, G. M., M. D. Yandell, J. R. Wortman, G. L. G. Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, P. W. Li, R. Apweiler, W. Fleischmann, J. M. Cherry, S. Henikoff, M. P. Skupski, S. Misra, M. Ashburner, E. Birney, M. S. Boguski, T. Brody, P. Brokstein, S. E. Celniker, S. A. Chervitz, D. Coates, A. Cravchik, A. Gabrielian, R. F. Galle, W. M. Gelbart, R. A. George, L. S. B. Goldstein, F. C. Gong, P. Guan, N. L. Harris, B. A. Hay, R. A. Hoskins, J. Y. Li, Z. Y. Li, R. O. Hynes, S. J. M. Jones, P. M. Kuehl, B. Lemaitre, J. T. Littleton, D. K. Morrison, C. Mungall, P. H. Ofarrell, O. K. Pickeral, C. Shue, L. B. Vosshall, J. Zhang, Q. Zhao, X. Q. H. Zheng, and et al. 2000. Comparative genomics of the eukaryotes. *Science* 287:2204-2215
- Safe, S.H., 1995. Environmental and dietary estrogens and human health: is there a problem? *Environm. Hlth. Perspect.* 103, 346-351.
- Sakamoto, T., Danzmann, R.G., Okamoto, N., Ferguson, M.M., Ihssen, P.E., 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 173, 33-43.
- Sakamoto, T., Danzmann, R.G., Gharbi, K., Howard, P., Ozaki, A., Khoo, S.K., Woram, R.A., Okamoto, N., Ferguson, M.M., Holm, L.-E., Guyomard, R., Hoyheim, B., 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155, 1331-1345.
- Samson S.L., Gedamu L., 1995. Metal-responsive elements of the rainbow trout metallothionein-B gene function for basal and metal-induced activity. *J. Biol. Chem.* 270, 6864-6871.
- Shashikant, C.S., Carr, J.L., Bhargava, J., Bentley, K.L., Ruddle, F.H., 1998. Recombinogenic targeting: A new approach to genomic analysis-a review. *Gene* 223, 9-20.
- Shum, B. P., Rajalingam, R., Magor, K.E., Azumi, K., Carr, W.H., Dixon, B., Stet, R.J., Adkison, M.A., Hedrick, R.P., Parham, P., 1999. A divergent non-classical class I gene conserved in salmonids. *Immunogenetics* 49, 479-490.
- Siitonen, L., Gall, G.A.E., 1989. Response to selection for early spawn data in rainbow trout, *Salmo gairdneri*. *Aquaculture* 78, 153-161.
- Sin, F.Y.T., 1997. Transgenic fish. *Rev. Fish Biol. Fish.* 7, 417-441.
- Stoner, G., Casto, B., Ralston, S., Roebuck, B., Pereira, C., Bailey, G., 2002. Development of a multi-organ rat model for evaluating chemopreventive agents: efficacy of indole-3-carbinol. *Carcinogenesis* 23, 265-272
- Sumpter, J.P., 1995. Feminized responses of fish to environmental estrogens. *Toxicol. Lett.* 82-83, 737-742.
- Takami, K., Zaleska-Rutczynska, Z., Figueroa, F., Klein, J., 1997. Linkage of LMP, TAP, and RING3 with Mhc class I rather than class II genes in the zebrafish. *J. Immunol.* 159, 6052-6060.
- Takeuchi, H., Figueroa, F., O'HUigin, C., Klein, J., 1995. Cloning and characterization of class I Mhc genes of the zebrafish, *Brachydanio rerio*. *Immunogenetics* 42, 77-84.
- Takeuchi, Y., Yoshizaki, G., Takeuchi, T., 2001. Production of germ-line chimeras in rainbow trout by

- blastomere transplantation. *Mol. Reprod. Dev.* 59, 380-389.
- Thorgaard, G.H., 1983. Chromosomal differences among rainbow trout populations. *Copeia* 1983, 650-662.
- Thorgaard, G.H., 1992. Application of genetic technologies to rainbow trout. *Aquaculture* 100, 85-97.
- Timon, M., Elgar, G., Habu, S., Okumura, K., Beverley, P.C., 1998. Molecular cloning of major histocompatibility complex class I cDNAs from the pufferfish Fugu rubripes. *Immunogenetics* 47, 170-173.
- Trethewey, R.N., 2001. Gene discovery via metabolic profiling. *Curr. Opin. Biotech.* 12, 135-138.
- Tyagi, A.K., Mohanty, A., 2000. Rice transformation for crop improvement and functional genomics. *Plant Science* 158, 1-18.
- Venter, J.C. and co-authors, 2001. The sequence of the human genome. *Science* 291, 1304-1351.
- Waller, S.J., Ratty, A., Burbach, J.P.H., Murphy, D., 1998. Transgenic and transcriptional studies on neurosecretory cell gene expression. *Cellular & Molecular Neurobiology* 18, 149-171.
- Wheeler, P.A., Thorgaard, G.H., 1991. Cryopreservation of rainbow trout semen in large straws. *Aquaculture* 93, 95-100.
- Wood, C.M. 2001. Influence of feeding, exercise, and temperature on nitrogen metabolism and excretion. In: Wright, P.A., Anderson, P.M., (Eds.) *Nitrogen Excretion, Fish Physiology*, Vol. 20, Hoar, W.S., Randall, D.J., Farrell, A.P., (Series Eds.), Academic Press: New York, pp. 201-238.
- Yang, B.-Y., Chan, K.-M., Lin, C.-M., Chen, T.T., 1997. Characterization of rainbow trout (Oncorhynchus mykiss) growth hormone 1 gene and the promoter region of growth hormone 2 gene. *Arch. Biochem. Biophys.* 340, 359-368.
- Young, W.P., Wheeler, P.A., Fields, R.D., Thorgaard, G.H., 1996. DNA fingerprinting confirms isogenicity of androgenetically-derived rainbow trout lines. *J. Hered.* 87, 77-81.
- Young, W.P., Wheeler, P.A., Coryell, V.H., Keim, P., Thorgaard, G.H., 1998. A detailed linkage map of rainbow trout produced using doubled haploids. *Genetics* 148, 839-850.
- Zardoya, R. Garrido-Perteierra, A. Bautista, J.M., 1995. The complete nucleotide sequence of the mitochondrial genome of the rainbow trout, Oncorhynchus mykiss. *J. Mol. Evol.* 41, 942-951.
- Zhu, Z., Li, G. He, L., Chen, S., 1985. Novel gene transfer into the fertilized eggs of goldfish (Carassius auratus L. 1758). *Z. angew. Ichthyol.* 1, 31-34.

Addresses of co-authors of this white paper

- <sup>a</sup>School of Biological Sciences and Center for Reproductive Biology, Washington State University, Pullman WA 99164-4236, USA
- <sup>b</sup>Department of Environmental and Molecular Toxicology, Oregon State University, Corvallis OR 97331, USA
- <sup>c</sup>Virginia Institute of Marine Science, College of William & Mary, Gloucester Point, VA 23062, USA
- <sup>d</sup>Agricultural Research Center, Washington State University, Pullman WA 99164-6240, USA
- <sup>e</sup>Center of Marine Biotechnology, University of Maryland, Columbus Center, Suite 236, 701 E. Pratt Street, Baltimore, MD 21202, USA
- <sup>f</sup>Western Fisheries Research Center, U.S. Geological Survey- Biological Resources Division, 6505 NE 65<sup>th</sup> Street, Seattle, WA 98115, USA
- <sup>g</sup>Center for Fish Disease Research, Department of Microbiology, Nash Hall 220, Oregon State University, Corvallis, OR 97331-3804, USA
- <sup>h</sup>Department of Biological Sciences and Center for Reproductive Biology, University of Idaho, Moscow ID 83844-3051, USA
- <sup>i</sup>NIEHS Marine and Freshwater Biomedical Science Center, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149-1098, USA
- <sup>j</sup>Department of Biology, 200 University Ave. West, University of Waterloo, Waterloo ON N2L 3G1, CANADA
- <sup>k</sup>Fisheries and Oceans Canada, 4160 Marine Drive, West Vancouver, BC V7V 1N6, CANADA
- <sup>l</sup>Department of Animal and Veterinary Sciences, University of Idaho, Hagerman Fish Culture Experiment Station, 3059-F National Fish Hatchery Road, Hagerman ID 83332, USA
- <sup>m</sup>U.S. Department of Agriculture, Agricultural Research Service, Hagerman Fish Culture Experiment Station, 3059-F National Fish Hatchery Road, Hagerman ID 83332, USA
- <sup>n</sup>Department of Biological Sciences, Northern Arizona University, Flagstaff AZ 86011-5640, USA
- <sup>o</sup>Department of Biology, Indiana University, Bloomington IN 47405, USA
- <sup>p</sup>U.S. Department of Agriculture, Agricultural Research Service, National Center for Cool and Cold Water Aquaculture, 11876 Leetown Road, Kearneysville, WV 25430, USA
- <sup>q</sup>Department of Animal Science, University of California, Davis CA 95616 USA
- <sup>r</sup>Clear Springs Foods, PO Box 712, Buhl ID 83316
- <sup>s</sup>School of Biological Sciences, Washington State University, Vancouver WA 98686-9600
- <sup>t</sup>National Marine Fisheries Service, 2725 Montlake Blvd. E., Seattle WA 98112
- <sup>u</sup>Department of Aquatic Biosciences, Tokyo University of Fisheries, Minato, Tokyo 108, JAPAN
- <sup>v</sup>Department of Zoology, University of Guelph, Guelph ON N1G 2W1 CANADA
- <sup>w</sup>Division of Biological Sciences, University of Montana, Missoula MT 59812 USA
- <sup>x</sup>Department of Animal Breeding and Genetics, Danish Institute of Agricultural Science, DK-8830, Tjele DENMARK
- <sup>y</sup>Amplicon Express Inc., 1610 NE Eastgate Blvd Suite #880, Pullman, WA 99163 USA
- <sup>z</sup>Programa de Genetica Humana, I.C.B.M. Facultad de Medicina, Universidad de Chile, Santiago, CHILE
- <sup>aa</sup>Institut Nationale de la Recherche Agronomique, Laboratoire de génétique des poissons CRJJ, 78352 Jouy-en-Josas cedex, FRANCE