Evolution of the Neuropeptide Y System in Vertebrates with Focus on Fishes

TOMAS LARSSON
Dissertation presented at Uppsala University to be publicly examined in B22, BMC, Husargatan 3, Uppsala, Friday, September 28, 2007 at 10:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

Gene families in vertebrates often contain more duplicates (paralogs) than in invertebrates. This has been attributed to genome duplications, i.e., tetraploidizations. Two of the gene families that have expanded in vertebrate evolution are the neuropeptide Y (NPY) family of peptides and the neuropeptide Y receptors (NPYR) that are involved in many brain functions including appetite regulation.

Two NPYR genes, Y2 and Y7, were cloned in the rainbow trout. Although they arose from a common ancestral gene in early vertebrate evolution, their ligand-binding properties are very similar. Two NPYR genes were cloned in the coelacanth Latimeria chalumnae and found to be orthologs of Y5 and Y6 discovered in mammals.

Analyses of gene families close to the NPY genes in the pufferfishes T. nigroviridis and T. rubripes showed that at least 25 additional gene families had an evolutionary history similar to the NPY family, thereby providing evidence for fish specific-duplications of these chromosomes. Cloning and phylogenetic analysis of 22 NPYR gene fragments from several ray-finned fishes showed that basal species seem to have the same repertoire as tetrapods. Despite the tetraploidization in the teleost fish lineage, many teleosts seem to have fewer genes than the gnathostome ancestor due to gene loss. Only one duplicate seems to have survived.

The NPY peptide family was found to have expanded in the teleost tetraploidization with duplicates of both NPY and PYY (peptide YY) in some teleosts. Fourteen neighboring gene families were found to have evolved in a similar manner as the NPY-family genes. Positional information facilitated orthology assignment of peptide genes in teleost fishes and allowed correction of previously misidentified genes.

In summary, the evolutionary history of the NPY and NPYR gene families involve large-scale duplication events coinciding with the proposed tetraploidizations. The appearance of new genes in early vertebrates and in teleost fishes probably had important implications for the evolution of new functions in this system.

Keywords: neuropeptide Y, GPCR, evolution, gene duplication, vertebrates, ray-finned fishes

Tomas Larsson, Department of Neuroscience, Box 593, Uppsala University, SE-75124 Uppsala, Sweden

© Tomas Larsson 2007

ISSN 1651-6206
ISBN 978-91-554-6958-0
urn:nbn:se:uu:diva-8189 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8189)
“I was taught a month ago
To bide my time and take it slow
But then I learned just yesterday
To rush and never waste the day
Now I’m convinced the whole day long
That all I learn is always wrong
And things are true that I forget
But no one taught that to me yet…”

“Character zero” by Phish from the 1996 album “Billy Breathes”
List of publications


VI. Sundström G, Larsson TA and Larhammar D. Phylogenetic analyses of multiple gene families syntenic with the Hox clusters add further support for two rounds of block/chromosome duplication in early vertebrate evolution. Manuscript.
## Contents

**Introduction**...................................................................................................11  
Neurotransmitters and receptors ....................................................................12  
The NPY family of peptides .......................................................................12  
The problem of analyzing short peptide sequences ....................................13  
G-protein coupled receptors ....................................................................14  
The NPY receptor family .......................................................................16  
Functions of the NPY system ..................................................................20  
Vertebrate evolution and evolutionary novelties .......................................21  
Duplication of genes ..............................................................................21  
The 2R and 3R hypotheses ......................................................................22  
Evolution of new functions after duplication ...........................................25  
Genome duplications and diversity of vertebrates ...................................27  
Species examined ....................................................................................28  

**Aims** ..............................................................................................................30  

**Materials and Methods** ..................................................................................31  
  Degenerate PCR, cloning and sequencing of gene fragments ....................31  
  Library screening ....................................................................................31  
  Binding studies ......................................................................................31  
  RT-PCR ..................................................................................................31  
  Database mining and BLAST searches ....................................................32  
  Phylogenetic analysis ............................................................................33  
  Statistical testing of regions ..................................................................33  

**Results and discussion** ..................................................................................34  
  Paper I ..................................................................................................34  
  Paper II .................................................................................................35  
  Paper III ...............................................................................................35  
  Paper IV ...............................................................................................36  
  Paper V .................................................................................................37  
  Paper VI ...............................................................................................38  

**Summary and Conclusions** .............................................................................39  

**Ongoing work and future perspectives** .........................................................41  

**Acknowledgements** .......................................................................................43  

**References** ....................................................................................................45
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2R</td>
<td>Two rounds of genome duplication</td>
</tr>
<tr>
<td>3R</td>
<td>Third round of genome duplication</td>
</tr>
<tr>
<td>AC</td>
<td>Adenylyl cyclase</td>
</tr>
<tr>
<td>ARC</td>
<td>Arcuate nucleus of the hypothalamus</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>DMN</td>
<td>Dorsomedial nucleus of the hypothalamus</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxy ribonucleic acid</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptor</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum likelihood</td>
</tr>
<tr>
<td>NJ</td>
<td>Neighbor-joining</td>
</tr>
<tr>
<td>MP</td>
<td>Maximum parsimony</td>
</tr>
<tr>
<td>NPY</td>
<td>Neuropeptide Y (Y=tyrosine)</td>
</tr>
<tr>
<td>PYY, 125I-pPYY</td>
<td>Peptide YY (Y = tyrosine), iodinated PYY</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PP</td>
<td>Pancreatic polypeptide</td>
</tr>
<tr>
<td>PVN</td>
<td>Paraventricular nucleus of the hypothalamus</td>
</tr>
<tr>
<td>QP</td>
<td>Quartet puzzling</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase PCR</td>
</tr>
<tr>
<td>VMN</td>
<td>Ventromedial nucleus of the hypothalamus</td>
</tr>
<tr>
<td>Y1, Y2, …Y7, Y8</td>
<td>Neuropeptide Y receptor Yn</td>
</tr>
</tbody>
</table>
Introduction

All living organisms are related to each other by common ancestry as evident from the shared use of the same type of genetic material, the DNA, with the ability to store genetic information as well as passing it on to the next generation. Despite the many underlying similarities when it comes to the basic molecular and cellular make up, evolution has produced an overwhelming diversity and complexity of organisms. Which processes and mechanisms are involved in generating this diversity and complexity? How can we explain the change from seemingly simple organisms to more complex ones? What are the underlying differences between an early branching chordate compared to vertebrates?

Recent progress in molecular biology has provided biologists with an enormous amount of information on the genetic material from a large number of organisms making it possible to address the questions posed above with greater accuracy than ever before. This can be achieved thanks to improvements of the techniques for DNA sequencing and the development of software for assembly of the obtained sequences into larger continuous stretches of DNA, and also thanks to the development of new techniques for the subsequent analysis of these sequences and the information and proteins they encode.

The observation that the genomes of closely related organisms contain many similar genes and that the genome of any particular organism usually contains many genes with sequence similarity to each other opens for the possibility of using comparative methods to unravel the evolutionary history of species, gene families and ultimately whole genomes. The understanding of gene family evolution is crucial for understanding origin of complex systems such as for example the vertebrate nervous system, because related genes usually codes for proteins performing similar functions. Thus, comparative evolutionary approaches can be used to infer functions of newly found genes, the relationships of species or genes to each other, and greatly aid in the design of future experiments in physiology, behavior and other disciplines.
Neurotransmitters and receptors

The nervous systems of all animals rely on the conversion of electrical impulses within nerve cells to chemical signals that can be passed on to other cells. The multiplicity and specificity of this inter-cellular communication is made possible by a large variety of signaling molecules (referred to as neurotransmitters), each interacting with one or several types of preferred target molecules (receptors) on the receiving cell. One class of such signaling molecules are made up of relatively short amino acid sequences (peptides) encoded by several families of genes. After release, these peptides bind to their specific target receptors in the postsynaptic cell membrane of the receiving cell. Families of related genes, like the ones coding for the peptides, code for the receptor proteins present on the cell surface.

In this thesis, I have studied the evolution of one such family of genes coding for the neurotransmitter neuropeptide Y (NPY) and related peptides with not only neuronal but also endocrine (i.e. hormonal) functions and the genes coding for the receptor proteins that receive these chemical messages, the neuropeptide Y receptors. This work was mainly concentrated on this system in vertebrates, particularly various species of fish.

The NPY family of peptides

Neuropeptide Y (NPY) (Tatemoto, 1982b), peptide YY (PYY) (Tatemoto & Mutt, 1980; Tatemoto, 1982a) and in tetrapods pancreatic polypeptide (PP) (Kimmel et al., 1968; Kimmel et al., 1975) constitute a family of related peptides. They act by binding to G-protein coupled receptors (see below). Almost all known members of the NPY-family are 36 amino acids long (Larhammar, 1996; Larhammar et al., 2004) and have a similar three-dimensional structure, the PP-fold (Allen et al., 1987; Darbon et al., 1992; Li et al., 1992; Keire et al., 2000), see Figure 1. NPY is predominantly expressed in neurons with regulatory effects on food intake, reproduction, blood pressure, circadian rhythms, learning and memory (Dumont et al., 1992). Both PYY and PP are endocrine peptides with their primary effects in the gut (Hazelwood, 1993). The naturally occurring truncated form of PYY, PYY 3-36 has been found to have an inhibitory effect on food intake (Batterham et al., 2003a) by binding Y2 receptors in the arcuate nucleus (ARC) in the hypothalamus of mammals (Batterham et al., 2002). This is interesting because the opposite effect is seen by NPY in the hypothalamus. In fact NPY is one of the most effective endogenous inducers of feeding with effect in all vertebrates studied. The increased food intake as observed when administering NPY is mediated via two other receptor subtypes, the Y1 and the Y5 receptors. Thus, the different peptides and receptors of the neuropeptide Y system have evolved to have opposing effects on food intake.
in mammals. This is an example of how duplication of genes creating several paralogs (genes originated by duplication of an ancestral gene) can facilitate the divergence of functions as seen both in the peptide gene family and the receptor gene family.

The NPY family of peptides has expanded during vertebrate evolution via block or genome duplication as well as local duplications (Larhammar et al., 2004). The vertebrate ancestor is supposed to have possessed one ancestral NPY/PYY gene that was duplicated to give rise to NPY and PYY. This duplication most probably included large parts of the genome since NPY and PYY are situated close to two clusters of homeobox genes, HOXA and HOXB (located on chromosome 7 and 17 in the human genome) together with several other gene families also represented on these two chromosomes (Paper V and VI). PP seems to be the result of a local tetrapod-specific duplication based on its tandem location close to PYY in several species (Hort et al., 1995; Larhammar et al., 2004). Additional lineage specific duplications of these peptides are known in many species, including local events as well as block duplications or tetraploidizations (Larhammar et al., 2004). In actinopterygians (ray-finned fishes), the tetraploidization proposed to have occurred before the radiation of teleosts seems to have resulted in additional copies of both NPY and PYY, referred to as NPYb and PYYb (Paper V). The evolutionary rates of the peptides vary; NPY is the most slowly evolving, with 20 amino acids shared among all known sequences while PP is the fastest with only 6 conserved positions.

Figure 1. Schematic representation of the structure of NPY-family peptides represented by human NPY. The characteristic PP-fold is shown in the figure. Residues 1-8 (light grey) constitutes a proline helix followed by a β-turn (residues 9-14, black), the α-helix composed of residues 15-30 (indicated in dark grey) and the amidated C-terminal tail of six residues.

The problem of analyzing short peptide sequences

Comparative analyses of DNA or amino acid sequences can be used as mentioned above in order to infer the relationships among genes or species using methods of phylogenetic reconstruction. All phylogenetic methods are based
on the general idea that homologous sequences (sequences related by common descent), separated due to speciation (orthologs) or duplication (paralogs), will accumulate an increasing number of differences due to mutations largely proportional to the time separating them. This simplified view of sequence evolution is complicated by several factors. One common complication is purifying selection acting against certain substitutions thus slowing down the evolutionary rate due to functional constraints on the molecule that the gene is coding for. Another problem is underestimation of the true number of substitutions due to large evolutionary distances or increased evolutionary rates in one or several lineages. Reliable phylogenetic reconstruction is also dependent on a sufficient number of informative characters being compared, making the analysis of short sequences such as the ones coding for the NPY family of peptides or other short peptides particularly unreliable (Dores et al., 1996). These problems can partly be overcome by analyzing sequences from many species carefully selected based on their relative evolutionary positions in the tree of life as well as including information on chromosomal location for the genes in species where large parts of the genome has been sequenced or mapped.

G-protein coupled receptors

The family of G-protein coupled receptors (GPCR) is one of the largest gene families in vertebrate genomes (Civelli et al., 2001) with more than 800 members in the human genome (approximately 3% of all genes based on a total gene count of 25000) (Bockaert & Pin, 1999; Lander et al., 2001; Venter et al., 2001; Takeda et al., 2002; Fredriksson et al., 2003; Vassilatis et al., 2003; Fredriksson & Schioth, 2005). Members of this receptor superfamily share the common feature of having seven hydrophobic alpha helices spanning the cell membrane with the N-terminal part of the receptor molecule located outside the cell and the C-terminal part located in the cytosol. Ligands for these receptors bind specifically to extracellular parts of the receptor, causing a conformational change of the intracellular parts. In this way, the extracellular signal is converted to an intracellular signal by interaction of the receptor with downstream targets (see Fig. 2). Different types of GPCRs react to a variety of stimuli such as light, ions, peptides, odorants and hormones (Bockaert & Pin, 1999) and with an increasing number of new ligand-receptor interactions being identified (Civelli, 2005). GPCRs have been divided into different families by several authors based on “fingerprints” of TM regions (Attwood & Findlay, 1993; Attwood & Findlay, 1994), their binding to G-proteins and sequence identity (Kolakowski, 1994), ligand binding and receptor size (Bockaert & Pin, 1999) and most recently based on phylogenetic analysis (Fredriksson et al., 2003; Fredriksson & Schioth, 2005). Because of the diversity of signals received by GPCRs they are extremely important drug targets (Hopkins & Groom, 2002;
Ma & Zemmel, 2002; Wise et al., 2002). So far, the 3D structure has only been determined for one GPCR, bovine rhodopsin (Palczewski et al., 2000).

The downstream targets for many GPCRs are most often trimeric G-proteins (hence the name) consisting of one alpha, one beta and one gamma subunit. G-proteins can be divided into four families based on the similarity of their alpha subunits and effects on targets further downstream (Hamm, 1998; Cabrera-Vera et al., 2003; Milligan & Kostenis, 2006; Oldham & Hamm, 2006). In mammals, several different genes have been found coding for the three different subunits (Hamm, 1998; Downes & Gautam, 1999; Kristiansen, 2004). Currently there are 16 known genes coding for alpha subunits, 5 genes for beta subunits and 12 genes for gamma subunits (Downes & Gautam, 1999; Oldham & Hamm, 2006). One of the main effects of G-protein activation is regulation of any of several isoforms of adenylyl cyclases (AC), which in turn influence the levels of the second messenger cyclic adenosine monophosphate (cAMP) (Hanoune & Defer, 2001; Sunahara & Taussig, 2002).

![Figure 2](image.jpg)

Figure 2. Schematic view of a GPCR in a patch of cell membrane with the transmembrane helices depicted in green. The N-terminal part of the receptor is located on the outside of the cell. On the cytosolic side the intracellular loops and the C-terminal interact with downstream targets, in this case a trimeric G-protein (alpha subunit depicted in light purple, beta subunit in pale yellow and gamma subunit in darker yellow). The ligand (exemplified by NPY) binds to extracellular parts of the receptor to cause a conformational change that is propagated to the inside of the cell.
The NPY receptor family

NPY receptors form a family within the superfamily of rhodopsin-like GPCRs. Currently there are five different subtypes cloned in mammals with additional subtypes found in various species of fish, amphibians and birds (Larhammar & Salaneck, 2004). In addition to the receptors cloned in vertebrates, there are a few NPY receptor-like genes sequenced in invertebrates, for example the gastropod mollusc *Lymnaea stagnalis* (Tensen et al., 1998), the fruit fly *Drosophila melanogaster* (Garczynski et al., 2002) and the malaria mosquito *Anopheles gambiae* (Hill et al., 2002). The vertebrate NPY family of receptors can be divided into three different clades, the Y1, Y2 and Y5 subfamilies. These subfamilies are approximately 30% identical to each other. The presence of NPY receptor-like genes in many different animal groups and the low similarity between subfamilies suggest an ancient origin of the NPY system. The fact that the NPY peptide and several of the receptors are highly conserved over long evolutionary time periods implies important functions of the NPY system. The signalling of NPY receptors is predominantly mediated via the inhibitory G-proteins G\(_i\) and G\(_o\), thus inhibiting adenylyl cyclase and thereby causing a decrease of cAMP in the cell. This signalling seems to be conserved among all NPY receptors although additional pathways are known to be involved in NPY signalling as well (Holliday et al., 2004).

Classification of receptors is based on sequence similarity at the gene (nucleotide) or gene product (amino acid) level, followed by phylogenetic analysis. However, prior to sequencing of their genes NPY receptors were classified by their different binding properties. In species where detailed chromosomal maps are available, the position of a certain gene in the genome also gives important information regarding its identity. The genes encoding the Y1, Y2 and Y5 receptors are present on the corresponding chromosome in several species (Paper III). This suggests local duplications early in vertebrate evolution followed by chromosome (or at least block) duplications, giving rise to the additional members in each subfamily (Fig. 3). The Y1, Y2 and Y5 receptors are particularly interesting since they have been implied in the regulation of feeding (see below). The Y3 subtype was hypothesised to exist based on pharmacological data but no sequence has been found in any species (Larhammar et al., 2001) hence it probably does not exist as a separate gene. One explanation for the pharmacological finding could be the interaction of one of the other receptors with regulating proteins that alter the receptor’s properties.

The Y1 receptor subfamily

The Y1, Y4 and y6 receptors belong to the Y1 subfamily and have arisen by duplication of an ancestral Y1 like gene (Wraith et al., 2000). The receptors Y8a and Y8b found in teleosts and previously named Yc and Yb (Lundell et
also belong to this subfamily. The first Y1 receptor was cloned from rat in 1990 (Eva et al., 1990) and is the best characterized among all NPY-receptors. It is highly conserved when comparing the sequence of different species and in tetrapods it has an intron of 80-120 bp inserted after the part of the gene encoding TM5 (Herzog et al., 1993). Pharmacologically it is characterised by an equal affinity for NPY and PYY and lover affinity for PP and truncated fragments of the peptides (Herzog et al., 1992; Larhammar et al., 1992). It is expressed both in the central nervous system and in the periphery and has been shown to have a wide variety of effects in for example decreasing anxiety (Wahlestedt et al., 1993), depression (Kask et al., 2001; Redrobe et al., 2002a; Redrobe et al., 2002b), regulation of ethanol (Thiele et al., 2002) and food intake (Larsen et al., 1999; Kanatani et al., 2001; Mullins et al., 2001) (and see below) and blood pressure regulation (Capurro & Hudobro-Toro, 1999). The receptor initially found in zebrafish and named Ya (Starback et al., 1999) was later shown to be the fish ortholog of Y4 (Salaneck et al., 2003) (and Paper III). The Y4-receptor has evolved the fastest of the NPY receptors, in line with its preferred binding to the fast-evolving peptide PP in mammals. The Y6 subtype is a pseudogene in many mammalian species including human (this is why it is designated y6 and not Y6) (Michel et al., 1998; Starback et al., 2000).

The Y2 receptor subfamily

The Y2 and Y7 receptors are the only two members of the Y2 subfamily. The first member of this subfamily was cloned in 1995 (Gerald et al., 1995). The mammalian members of this subfamily are special in that they bind truncated fragments on NPY and PYY with high affinity (Michel et al., 1998). This has not been observed for Danio rerio Y2 or Y7 (Fredriksson et al., 2004; Fredriksson et al., 2006), Oncorhynchys mykiss Y2 or Y7 (Paper I) or chicken Y7 (Fredriksson et al., 2006). The Y2 receptor has been found to function as a presynaptic autoreceptor leading to inhibition of neurotransmitter release (Caberlotto et al., 2000). Y2 receptors are mainly expressed in the brain (Kaga et al., 2001) and often have opposite effects to that seen for Y1 receptors, for example increased anxiety (Nakajima et al., 1998) and increased blood pressure (Morton et al., 1999). The binding of truncated fragments of PYY to Y2 receptors in the arcuate nucleus (ARC) of the hypothalamus in mammals has been shown to negatively regulate food intake (Batterham et al., 2002; Challis et al., 2003) (and see below). Y2 receptors have also been shown to be important for memory formation (Redrobe et al., 2004). Compared to Y1 and Y5 receptors the Y2 subfamily differs in that it is not internalized upon agonist stimulation (Parker et al., 2001; Gicquiaux et al., 2002). The roles of the Y7 receptor found in teleost fishes (Fredriksson et al., 2004) (Paper I and III), amphibians (Fredriksson et al., 2004) and chicken (Bromee et al., 2006) is still unknown.
The Y5 receptor subfamily
The Y5 receptor is the sole member of the Y5 subfamily and was discovered by cloning in 1996 (Gerald et al., 1996). The extended third intracellular loop is one feature of this subfamily that distinguishes it from the Y1 and Y2 subfamilies. The gene for Y5 is located close to the Y1 gene in a “head to head” fashion (the two genes are transcribed in opposite directions) in several species. These two genes have been shown to be transcriptionally co-regulated (Herzog et al., 1997). Y5 is mainly expressed in the brain (Parker & Herzog, 1998; Parker & Herzog, 1999; Durkin et al., 2000) but also found in the periphery (Statnick et al., 1998). The most prominent functions for Y5 receptors are its effects on food intake.

A summary of the NPY receptor repertoire in some vertebrate lineages is shown in Fig. 3 illustrating the proposed evolution by duplication of chromosomal segments, partially based on the results presented in this thesis. The relationships between the different receptor subtypes based on phylogenetic analysis of a representative set of receptor sequences are shown in Fig. 4.

Figure 3. Schematic drawing of the NPY receptor repertoire found in different vertebrate lineages and that inferred for the ancestral gnathostome (Paper III). Multiple duplication events as well as lineage-specific losses have led to the different subtypes seen in extant species. 3R marks the duplication of the chromosomal region harbouring the Y8b gene duplicated via the fish-specific tetraploidization (see below). Crossed out boxes indicate lost genes or genes yet to be found. Picture modified from Larhammar et al. 2004.
Figure 4. Neighbor-joining tree constructed with MEGA 3.1 with standard settings showing the relationships of NPY receptor subtypes. Numbers at each branch are percent bootstrap support for each node. Nodes with bootstrap values less than 70% have been collapsed. Human bradykinin receptor B1 was used to root the tree.
Functions of the NPY system

Feeding regulation
As mentioned above the physiological functions of the NPY system are diverse. Among the most well studied effects of this system is the regulation of food intake. In the hypothalamus of mammals several different appetite regulating signals, both derived from the CNS and from peripheral tissues, are integrated to modulate feeding behavior. The NPY-producing neurons most important in this context are located in the brainstem and the arcuate (ARC) and dorsomedial nuclei (DMN) of the hypothalamus. The NPY neurons originating in the brainstem project to sites in the hypothalamus including ARC and DMN as well as to the ventromedial nucleus (VMN) and paraventricular nucleus (PVN). The NPY neurons of the ARC in turn project to both the PVN and DMN. These parts of the hypothalamus play central roles in food intake regulation. In addition to this central regulation, peripheral signals such as insulin, leptin and ghrelin also regulate this signaling (Schwartz & Morton, 2002). Interestingly, the NPY system itself regulates the activity of NPY signaling in the hypothalamus. For example, Y2 receptors in the hypothalamus down-regulate NPY release by responding to the circulating truncated peptide PYY 3-36 (Batterham et al., 2002; Batterham & Bloom, 2003; Batterham et al., 2003a; Challis et al., 2003; Chelikani et al., 2005). Similar effects have been proposed for circulating PP signaling via Y4 receptors (Batterham et al., 2003b). Y2 receptors also function as autoreceptors inhibiting the release of NPY onto proopiomelanocortin (POMC) neurons. The interplay between NPY and POMC neurons is well characterized and has interconnections at multiple levels. The NPY and POMC neurons of the ARC express receptors for leptin and insulin whereby they sense the energy levels in the body. Activation of NPY-containing neurons leads to the expression of agouti gene-related peptide (AgRP) that functions as a natural antagonist on melanocortin 4 receptors on the PVN neurons (Schwartz & Morton, 2002). The Y1 and Y5 receptors both have stimulatory effects on food intake, however each receptor influences different aspects of it. The Y1 receptor was found to stimulate the food-seeking behavior while Y5 stimulates the consummatory behavior in rodents (Lecklin et al., 2002; Lecklin et al., 2003; Day et al., 2005).

Blood pressure regulation
NPY is often co-released with noradrenaline (NA) in the periphery (Ekblad et al., 1984). It was shown at an early timepoint that NPY causes vasoconstriction via the Y1 receptor (Lundberg & Tatemoto, 1982). Several studies have shown the importance of the NPY system in blood pressure regulation, see for example (Shine et al., 1994), (Matsumura et al., 2003) and (Pedrazzini et al., 1998; Pedrazzini, 2004)
Effects on learning and memory

Another physiological function of the NPY system is the involvement in learning and memory. Region-specific effects have been observed after injection of NPY in different brain areas of rodents (Flood et al., 1989). These effects are predominantly attributed to signaling through Y2 receptors. This has been further corroborated by studies of Y2 knockout mice (Redrobe et al., 2004). The Y1 receptor has also been implicated in memory regulation by influencing glutamate release (Whittaker et al., 1999; St-Pierre et al., 2000). Rats overexpressing NPY in hippocampus were shown to have a reduced ability for spatial learning in a Morris water maze test (Thorsell et al., 2000).

Vertebrate evolution and evolutionary novelties

The idea that duplication of genetic material could be a powerful evolutionary force was put forward before it was established that DNA was the molecule carrying the genetic information, for examples see (Metz, 1947; Taylor & Raes, 2004). In 1970, Susumu Ohno published the book titled “Evolution by gene duplication” in which he gave numerous examples of how duplication of genetic material was central for evolutionary novelties to appear (Ohno, 1970). He argued that big leaps in evolution were only possible after the formation of redundant gene copies. Ohno also suggested that this frequently happened through duplication of whole genomes (polyploidization). He based these ideas on the observation of chromosome numbers of related species, frequently showing an increase by even multiples of the diploid chromosome number, the presence of duplicated gene loci, as well as studies of DNA content. He proposed that at least two whole genome duplications had occurred in the lineage leading to vertebrates. Since then, gene and genome duplication have been shown to be an important mechanism for generating new genes and providing the material for evolution of new functions. In the simplest case, two copies provide the opportunity for one copy to evolve new functions while the other copy maintains the original function.

Duplication of genes

Various mechanisms contribute to the generation of new gene copies. These can range from single, local gene duplications to duplications of several genes at the same time or even duplication of whole chromosomes or, as Ohno proposed, the whole genome of an organism. Single gene duplications are usually the result of unequal crossing over or retrotransposition. Tandem arrays of paralogous genes with similar intron-exon structures are usually the result of unequal crossing over while retrotransposed genes are intronless
and not necessarily linked to each other (Zhang, 2003). It has been shown that large-scale segmental duplications (also called low-copy repeats) are common in primate genomes (Bailey et al., 2002; Samonte & Eichler, 2002; Bailey & Eichler, 2006). These duplications do not seem to be the result of unequal crossing over and the mechanisms for generating these duplications are not clear. It has been shown that repeated sequences such as Alu-repeats are overrepresented in the boundaries of this type of duplication and that double strand breaks probably also plays a role in generating duplications of this type (Bailey & Eichler, 2006).

Polyploidizations can occur either by allopolyploidization (the fusion of the genomes from two related species with the same ploidy level) or autopolyploidization (a doubling of every set of homologous chromosomes) (Wolfe, 2001; Panopoulou & Poustka, 2005). Both these processes are known to be common in a variety of species and also known to have occurred quite recently in evolutionary time. In the amphibian lineage examples of both allo- and auto-polyploidization are known; allopolyploidization has been responsible for generating several polyploid species ranging from tetraploid to dodecaploid in the family Pipidae (i.e. different Xenopus species) (Evans et al., 2004) while autopolyploidy has been observed in the families Leptodactylidae and Hylidae resulting in tetraploid and octaploid species (Becak & Kobashi, 2004). Another lineage where polyploidization has occurred is the salmonid fishes. In this group, one autopolyploidization event has been proposed in early salmonid evolution (Allendorf & Thorgaard, 1984). There have also been several whole genome duplications in the cyprinid fish lineage, giving rise to several polyploid species (Stellwag, 1999; David et al., 2003).

In general polyploidization seems to be much more widespread in plants than in animals with as much as 70% of all angiosperm species being polyploid (Otto & Whitton, 2000). Among the best characterized polyploidization events are the ones in yeast (Wolfe & Shields, 1997; Seoighe & Wolfe, 1998; Seoighe & Wolfe, 1999a; Wagner, 2002; Kellis et al., 2004; Conant & Wolfe, 2006; Scannell et al., 2006; Byrne & Wolfe, 2007), Arabidopsis (The Arabidopsis Genome Initiative, 2000; Vision et al., 2000; Simillion et al., 2002; Vandepoele et al., 2002; Blanc & Wolfe, 2004; Wang et al., 2006) and teleost fishes (Meyer & Schartl, 1999; Taylor et al., 2001; Taylor et al., 2003; Christoffels et al., 2004; Jaillon et al., 2004; Van de Peer, 2004b).

The 2R and 3R hypotheses
Mammals, birds, fishes, and amphibians have more genes and usually have more members of any particular gene family than more basal vertebrates and invertebrates. It has also been shown that vertebrate genomes contain simi-
larly looking chromosomal segments (Lundin, 1993) referred to as paralogons (Coulier et al., 2000). These paralogons have been shown to be distributed over many chromosomes and to cover large parts of the genome in several species (Pebusque et al., 1998; Popovici et al., 2001a; Popovici et al., 2001b; Abi-Rached et al., 2002; Larhammar et al., 2002; McLysaght et al., 2002; Lundin et al., 2003; Vienne et al., 2003; Dehal & Boore, 2005; Olinski et al., 2006). Several examples of paralogons containing quartets of similar chromosomes have been taken as evidence for at least two large-scale duplications of chromosomes or two rounds of tetraploidization early in vertebrate evolution. This has been referred to as the 2R or the "one-to-four theory" (Hughes, 1999; Meyer & Schartl, 1999; Hokamp et al., 2003; Panopoulou & Poustka, 2005). The 2R theory is supported by studies of a number of gene families, the prime example being the developmentally important homeobox (Hox) genes in many species (Amores et al., 1998; Hoegg & Meyer, 2005). The early branching chordate Amphioxus has only one Hox cluster (Garcia-Fernandez & Holland, 1994; Holland et al., 1994) while sarcopterygians usually have four clusters (Hoegg & Meyer, 2005).

Several authors have expressed criticism of the 2R theory based on the fact that many gene families in vertebrates contain two or three gene copies instead of the four predicted. Furthermore, estimated time points of duplication differ between gene families (Hughes, 1999; Wang & Gu, 2000; Hughes et al., 2001). In addition to these complications, phylogenetic trees that do contain four members do not show the topology consistent with two consecutive genome duplication events (Hughes & Friedman, 2003) (Fig. 5).

![Figure 5](image)

Figure 5. Panel A shows the (A,B),(C,D) topology which is the topology that Hughes et al. proposed to be the only topology compatible with the 2R theory (Hughes, 1999). Others have argued that topologies as seen in panel B and C also are compatible with the 2R hypothesis if the evolutionary rates varies between paralogs or if the two duplication events occurred in a short time window as proposed by Nadeu and Sankoff (1997), Gibson and Spring (2000) and Aburomia et al. (2003).
The observation that some gene families have less than four members could be accounted for by gene loss after duplication. This has been investigated in detail for Hox genes, reviewed by (Hoegg & Meyer, 2005). In fact, most duplicates seem to be deleted in the first million year following duplication (Lynch & Conery, 2000). Variations in selection pressure and evolutionary rates in genes belonging to the same family can sometimes make the estimation of divergence times and construction of reliable phylogenetic trees difficult. In addition, it has been argued (Nadeau & Sankoff, 1997; Gibson & Spring, 2000; Furlong & Holland, 2002; Larhammar et al., 2002; Aburomia et al., 2003) that two duplication events in a close time window will scramble the phylogenetic signal making it hard to recover the actual duplication history using only phylogenetic methods. Because of these limitations of phylogenetic analyses, some authors have focused on combining relative dating of duplicates with statistical testing of positional information of paralogs from several gene families (Abi-Rached et al., 2002; Vienne et al., 2003; Danchin & Pontarotti, 2004). The debate over the validity of 2R has also been centered on the methods used to detect duplications and the predictions of the theory (Skrabanek & Wolfe, 1998; Hokamp et al., 2003; Simillion et al., 2004; Van de Peer, 2004a; Durand & Hoberman, 2006). It has been argued that earlier studies based on small selected regions of the genome were unrepresentative of the genome as a whole and that the methods used were inadequate to prove or disprove 2R (Skrabanek & Wolfe, 1998; McLysaght et al., 2002; Hokamp et al., 2003). The combination of phylogenetic studies of gene families and map-based approaches has been the best evidence in favor of 2R (Dehal & Boore, 2005).

In the ray-finned fishes (Actinopterygia) support for one additional genome duplication has accumulated for teleosts (Meyer & Schartl, 1999; Postlethwaite et al., 2000; Taylor et al., 2001; Van de Peer et al., 2003). The complete genome sequences from teleost model species, especially the pufferfish Takifugu rubripes and Tetraodon nigroviridis, have further strengthened the idea of a third round of genome duplications or “3R” early in this lineage (Christoffels et al., 2004; Jaillon et al., 2004; Vandepoele et al., 2004). An overview of the evolution of chordates with the postulated genome duplications is shown in Fig. 6.
Figure 6. Schematic overview of chordate evolution. Black arrows indicate proposed time points for the two tetraploidizations "2R" according to (Panopoulou & Poustka, 2005) and the fish specific tetraploidization "3R" (Christoffels et al., 2004; Jaillon et al., 2004; Panopoulou & Poustka, 2005). Gray lines indicate phylogenetic relationships between taxa that are still debated i.e. the position of cephalochordates or urochordates as the sister group to the rest of the chordate tree (Winchell et al., 2002; Delsuc et al., 2006; Vienne & Pontarotti, 2006) and that of the jawless fishes being monophyletic or polyphyletic (Kuraku et al., 1999; Delarbre et al., 2000; Gursoy et al., 2000; Delarbre et al., 2002; Takezaki et al., 2003). The branching order as shown here is based on recent studies of molecular data. Proposed divergence times for protostomes/deuterostomes (a), sarcopterygians/actinopterygians (d) reptilomorphs/batrachomorphs (f) and synapsids/sauropsids (g) are based on paleontological data as recently reviewed (Benton & Donoghue, 2007) (Indicated by triangles in the figure). The remaining time points (b, c, e and h) (indicated by diamonds) are fossil minimum times according to Blair and Hedges 2005 and references therein (Blair & Hedges, 2005). The dashed line for cyclostomes (Lampreys and Hagfishes) indicates lack of fossil representation of hagfishes in their early evolution. Time scale indicates million years before present.

Evolution of new functions after duplication

The rate of duplication and fate of duplicated genes have been investigated in several eukaryotic model species. These studies suggest that non-functionalization occurs more frequently than partitioning of the ancestral function (sub-functionalization) or appearance of completely new functions (neo-functionalization). Based on data from nine eukaryotic species Lynch and Connery (Lynch & Conery, 2000) have estimated the duplication rate to be on average 0.01 per gene per million years but with considerable variation (between 0.02 to 0.002) in the species investigated. With this rate of duplica-
tion as many as 50% of the genes in a genome could duplicate in 35 million years even without large-scale events such as tetraploidizations. The half-life for duplicated genes was estimated to be approximately 4 million years. Retention of duplicate genes in the genome over long time usually requires divergence of function (Nadeau & Sankoff, 1997; Krakauer & Nowak, 1999). It is generally believed that there is a period of relaxed selection pressure following gene duplication due to the redundancy created (Lynch & Conery, 2000). This has been suggested to play an important role in speciation since redundant copies could be differentially silenced in geographically isolated populations and therefore by purely random events lead to reproductive barriers (Lynch & Conery, 2000). The rate of fixation of duplicates is influenced by both selection and random events i.e. genetic drift. Because of this the effective population size is crucial in determining the fixation rate of gene duplicates (Nadeau & Sankoff, 1997; Lynch et al., 2001). Models for the retention of duplicated genes indicate that subfunctionalization is very important in the early stages followed by longer periods where neofunctionalization can occur (Force et al., 1999; Lynch & Force, 2000; Lynch et al., 2001; He & Zhang, 2005; Rastogi & Liberles, 2005). In cases where an increase in the amount of gene product gives an immediate selective advantage, this by itself might be enough to retain the extra copy in the genome (Seoighe & Wolfe, 1999b; Krylov et al., 2003). This would be most important for already highly expressed genes such as histones or rRNAs (Zhang, 2003). The importance of dosage effects has been put forward as one factor that increases the rate of nonfunctionalization after single duplication events just because of instability between interacting gene products. The retention of duplicates seems to be higher after polyploidization and this has been explained by the fact that the ratio of interaction partners stays the same as in the ancestral species (Shimeld, 1999). A simplified illustration of the different fates of duplicates is shown in Fig. 7. The extra genome duplication in actinopterygians compared to for example mammals has made it possible to test models of functional divergence using gene families with two copies in fishes and only one copy in mammals, reviewed in (Prince & Pickett, 2002). Several studies have shown that the duplicated genes have changed their expression as compared to the ancestral expression pattern (Altschmied et al., 2002; McClintock et al., 2002; Yu et al., 2003). This has been interpreted such that the genes are retained due to partitioning of the original function, very much as predicted by the models mentioned above.
Figure 7. Schematic description of gene fates after gene duplication. R1 and R2 depict two regulatory regions while C1 depicts the coding region of the gene. In A the two copies are retained simply due to increased amount of gene product. Panel B depicts the origin of a new function by neofunctionalization creating a new regulatory domain R3 or a new coding domain C2. C shows partitioning of the ancestral function due to complementary degenerate mutations generating two non-functional regulatory regions R0, giving rise to a partitioning of the ancestral function. In D the most common fate for the duplicate is illustrated, mutations have rendered one copy non-functional. Picture modified from Force et al., 1999.

Genome duplications and diversity of vertebrates

The idea that gene duplication is responsible for creating raw material for the subsequent evolution of new functions inevitably leads to the prediction that genome duplications should be linked to an increase in species diversity and complexity. However, it has been argued that the increased complexity seen in extant vertebrates as compared to chordates is an illusion due to the fact that the diversity of extinct species is not taken into account (Donoghue & Purnell, 2005). It has also been pointed out that there is no direct connection between the duplication of Hox clusters and complexity as the species rich clade of sarcopterygians and the species-poor cartilaginous fishes have the same number of Hox clusters (Robinson-Rechavi et al., 2004). It has been shown that the genome duplication in actinopterygians (3R) is specific to teleosts, which comprise approximately half of all vertebrate species (Hoegg
et al., 2004; Crow et al., 2006; Crow & Wagner, 2006). Again, this correlation between species diversity and complexity was challenged when paleontological data were taken into account (Donoghue & Purnell, 2005). It should also be noted that there is a large time gap between 3R and the origin of several species-rich groups of teleosts (Crow & Wagner, 2006; Hurley et al., 2007). However, it was estimated that the probability of extinction was greatly reduced in the lineages that underwent 3R and that this may have allowed the species diversity to increase (Crow & Wagner, 2006).

Species examined

**Oncorhynchus mykiss**
The rainbow trout *Oncorhynchus mykiss* is a commonly cultivated salmonid species. It is also well studied from a physiological viewpoint and genetic and genomic information is beginning to accumulate although there is no whole genome sequencing project on this species. It is a representative of the ray-finned fishes (actinopterygia), a group of animals comprising well over 23000 species (Taylor et al., 2001; Taylor et al., 2003; Venkatesh, 2003) and probably as many as 30000 species (www.fishbase.org). In addition to the fish specific genome duplication, there is evidence for additional genome duplications in the salmonid lineage (Allendorf & Thorgaard, 1984).

**Latimeria chalumnae**
The coelacanth, *Latimeria chalumnae*, is an early branch of the sarcopterygian lineage. Recently, a second species of coelacanth was discovered in Indonesia, *Latimeria menadoensis* (Holder et al., 1999). Together with lungfishes the coelacanths are the only extant species separating actinopterygians from tetrapods (Zardoya & Meyer, 1997; Zardoya et al., 1998; Venkatesh et al., 2001; Brinkmann et al., 2004a; Brinkmann et al., 2004b).

**Tetraodon nigroviridis and Takifugu rubripes**
These two species of pufferfish, just as rainbow trout, belong to the actinopterygian lineage. The evolutionary distance between these two species has been estimated to be 32-56 million years (Benton & Donoghue, 2007). The interest in these species stems from their exceptionally compact genomes (only about 10-12% of the human genome). This feature was the primary reason why these two species were selected for whole genome sequencing (Crnogorac-Jurcevic et al., 1997; Aparicio et al., 2002; Neafsey & Palumbi, 2003; Jaillon et al., 2004). Despite their small genomes the pufferfishes, like
other teleosts, have undergone 3R and therefore probably have even more genes than humans and other mammals.

**Danio rerio**

The zebrafish has long been used as a model in developmental biology because it has transparent embryos and is easily kept and bred in captivity. Many genetic tools have been developed for use in *Danio rerio*. Zebrafish is a relatively early branching actinopterygian but still is a member of the teleost lineage (Chen et al., 2004; Nelson, 2006; Steinke et al., 2006; Yamaneou et al., 2006).

**Basal actinopterygians**

The basal actinopterygians studied in Paper IV, namely *Acipenser baeri, Amia calva, Anguilla anguilla, Clupea harengus, Osteoglossus bichiriosum, Lepisosteus osseus* and *Polypterus senegalus* are of interest because of their phylogenetic position in relation to the proposed tetraploidization in the actinopterygian lineage (Hurley et al., 2007). The fact that the early, non-polyploidized branches constitute rather few species as compared to the species rich group of teleosts has been put forward as an argument for linkage between genome duplication and evolutionary success (Hoegg et al., 2004).

**Ciona intestinalis**

*Ciona intestinalis* is a member of the urochordate lineage. Being an early branching chordate, it was selected for whole genome sequencing with hopes to shed light on the evolution of chordates (Dehal et al., 2002; Canestro et al., 2003). Recently, the traditional phylogenetic position of urochordates as the earliest diverging group of extant chordates has been questioned. The latest analysis of molecular data indicates that urochordates are more closely related to vertebrates than are cephalochordates (Delsuc et al., 2006; Vienne & Pontarotti, 2006).

**Other important species**

In the phylogenetic analysis, we have included several vertebrates including human and mouse as well as some invertebrate species in addition to *Ciona intestinalis* mentioned above. The inclusion of invertebrate species other than *Ciona intestinalis* was due to lack of orthologs in the *Ciona intestinalis* genome, possibly due to lineage-specific deletions. The inclusion of invertebrate sequences is crucial for relative dating of duplication events.
Aims

One aim of this thesis was to determine sequences and to analyze available sequence information from several species of vertebrates in order to better understand the origin and evolution of the NPY system in vertebrates. This approach was combined with comparisons of chromosomal locations of genes in different species. The overall idea is that comparative approaches like these can provide not only basic information on the evolution of the repertoire of peptides and receptors but also help determining the interaction properties and functional roles of different ligands and receptor subtypes.

Specific aims of the different papers:

- In Papers I and II, the aim was to clone as many NPY receptor genes as possible from *Oncorhyncus mykiss* (Paper I) and *Latimeria chalumnae* (Paper II) and characterize them phylogenetically as well as compare their binding properties.

- Paper III aimed to phylogenetically characterize the total NPY receptor repertoire found in the sequenced genomes of the two pufferfish species *Takifugu rubripes* and *Tetraodon nigroviridis*. In addition to the phylogenetic analysis of the receptor genes, the aim was to analyze additional genes situated on the same chromosomes as the receptor genes in order to test the earlier notion that several of these receptor genes originated through large scale duplication events involving large regions or possibly the whole genome of the vertebrate ancestor.

- The main objective of Paper IV was to shed light on the evolution of the NPY receptor genes in the early evolution of actinopterygians in relation to the proposed tetraploidization in this lineage. Degenerate PCR and sequencing was used to obtain as many receptor sequences as possible for the Y1 and Y5 subfamilies in several basal actinopterygian species. This was done in order to compare the receptors present in early diverging euteleost species as well as those found in tetrapods.

- In Papers V and VI, positional and phylogenetic information were combined to characterize the chromosomal regions where the NPY peptide genes reside in order to dissect their evolutionary history. An additional aim was to use the information derived in this way to infer the identity of peptide genes where phylogenetic sequence analysis alone is not sufficient to define orthology.
Materials and Methods

Degenerate PCR, cloning and sequencing of gene fragments
Alignments of previously sequenced NPY receptors were used to construct degenerate primers with specificity for the Y1 subfamily (Paper IV), Y2 subfamily (Paper I), Y5 (paper II and paper IV) and Y6 (Paper II). Touch-down PCR with genomic DNA from *Oncorhynchus mykiss* (Paper I) and *Latimeria chalumnae* (Paper II) and several different species of actinopterygians (Paper IV) was performed in order to amplify part of the receptor gene sequences. The products were separated on 1% agarose gels and bands of the expected sizes were excised and retrieved DNA extracted using the Qiaquick gel extraction kit (Qiagen). Fragments were cloned using the TOPO-TA cloning kit (Invitrogen) and thereafter sequenced using the BigDye v.3 terminator sequencing kit (Applied biosystems) on an ABI 310 automatic sequencer.

Library screening

$^{32}$P-labelled gene fragments were used to screen λ-phage genomic libraries for the full-length sequences of NPY receptor genes from *Oncorhynchus mykiss* (Paper I) and *Latimeria chalumnae* (Paper II). Positive clones were cleaved using restriction enzymes and fragments were subcloned into pUC18 or pUC19 vectors and thereafter sequenced as described above.

Binding studies

Full-length PCR primers containing restriction sites for HindIII and XhoI were constructed for rainbow trout Y2 and Y7 (Paper I). The entire coding region was amplified and ligated into a modified pCEP4 vector (Marklund et al., 2002). Plasmid DNA was purified and used for transfection of HEK 293 cells. Cells were harvested and frozen at -80°C. Thawed aliquots of transfected cells were homogenized and used for binding studies. All binding experiments were performed with 2h incubation at room temperature with $^{125}$I-pPYY as radioligand. Saturation was performed using twofold serial dilutions of the radioligand. Competition binding experiments were performed with serial dilutions of nine agonists and one antagonist previously used for
characterization of NPY receptors. Ligands used were: pNPY, pPYY, zfNPYa, zfPYYa, zfPYYb, pNPY2-36, pNPY3-36, pNPY13-36, pNPY18-36 and the antagonist BIEE0246.

RT-PCR

Paper I
Three adult rainbow trout were killed after anesthesia with benzocaine. Twelve different tissues were dissected: gill, spleen, heart, muscle, testis, ovary, kidney, head kidney, intestine, brain, eye and liver. Total RNA from each tissue was prepared from one single individual. RNA was used for cDNA synthesis. PCR using primers specific for Y2 and Y7 were used with cDNA as template. Products were separated on agarose gels and blotted onto nylon filters for Southern hybridization.

Papers III and V
Total RNA was isolated from eleven *T. rubripes* tissues using TRIzol reagent (Life Technologies, USA) according to the manufacturer’s protocol. Purified total RNA was reverse transcribed to cDNA using AMV reverse transcriptase first-strand cDNA synthesis kit (Gibco BRL, USA). The PCR was performed according to the following protocol: a denaturation step at 95°C for 2 min, 35 cycles of 95°C 30 sec, 55°C for 1 min, 72°C for 1 min followed by a final elongation step at 72°C for 5 min. Identity of representative RT-PCR products was confirmed by sequencing on an Applied Biosystems 3700 DNA Analyzer using dye-terminator chemistry.

Database mining and BLAST searches
All NPY receptor genes (Paper III) and NPY peptide family gene sequences (Paper V) were identified in the genome databases of the two pufferfishes *Takifugu rubripes* (Japanese pufferfish) and *Tetraodon nigroviridis* (Green spotted pufferfish) [http://www.ensembl.org/Fugu_rubripes/index.html](http://www.ensembl.org/Fugu_rubripes/index.html) and [http://www.ensembl.org/Tetraodon_nigroviridis/index.html](http://www.ensembl.org/Tetraodon_nigroviridis/index.html) by BLAST searches (Altschul et al., 1990) with previously known sequences. In Paper III, the genes neighbouring NPY receptor genes on the *Takifugu rubripes* scaffolds and *Tetraodon nigroviridis* chromosomes were investigated and used to construct phylogenetic trees. All genes surrounding the receptor gene entries were downloaded from the Ensembl database and sorted according to gene family identity. Gene families with members in proximity to three or more of the *Tetraodon* receptor genes were used for initial phylogenetic analysis using the neighbor-joining method (see below). A similar approach was taken in Paper V and Paper VI for the genes located near the peptide
coding genes. Included in the initial datasets were sequences from Homo sapiens, Mus musculus, Tetraodon nigroviridis, Takifugu rubripes, and Danio rerio as well as outgroup sequences from Ciona intestinalis, Ciona savigny or Drosophila melanogaster. This selection was made in order to achieve relative dating of the duplication of the genes under study. All sequences were included based on the Ensembl family prediction and downloaded from the Ensembl-database: http://www.ensembl.org/. The pfam database, http://www.sanger.ac.uk/Software/Pfam/, was used in Paper III, V and VI to identify shared domains and to aid in the alignment of sequences (see below).

Phylogenetic analysis

Several different software packages were used for alignment of sequences and phylogenetic inference. Multiple alignments were constructed using the Windows version of ClustalX 1.81 (Thompson et al., 1994; Jeanmougin et al., 1998) and the implementation of clustal W in MEGA 3.1 (Kumar et al., 2004). Phylogenetic trees were constructed using Clustal X and MEGA 3.1 (for NJ analyses in Papers I, II, III, V and IV), MEGA 3.1 (for MP in Paper II), Treepuzzle 5.2. (Schmidt et al., 2002) (for QP maximum-likelihood in Paper II, III, V and VI), Phylip 3.63 (Felsenstein, 2004) (ML in Paper I), PHYML (ML in Paper IV) (Guindon & Gascuel, 2003) and PAUP (NJ in paper IV) (Swofford, 2003).

Statistical testing of regions

In Paper III, we used a binomial test as described earlier by Vienne et al. (Vienne et al., 2003) to test if the distribution of investigated paralogs were randomly distributed in the genome based on the assumption that the probability for a randomly chosen gene to be present in a particular region of the genome is proportional to the number of genes in that region.
Results and discussion

Paper I

Degenerate PCR was performed using genomic DNA from rainbow trout as template. Two different clones were obtained, one with sequence similarity to previously sequenced Y2 receptor genes and one similar to Y7 genes. These two clone inserts were labelled with $^{32}$P and used to screen a genomic λ-phage library. Six genomic clones containing Y2 and three containing Y7 sequence was isolated. Phylogenetic analysis indicated that both sequences were unambiguously grouped with other members of the Y2 subfamily. The sequences have been submitted to GenBank with the accession numbers DQ231509 (Y2) and DQ231510 (Y7). The coding regions of the two genes were cloned into a pCEP4 vector and expressed in HEK-293 cells. Saturating binding experiments were performed with $^{125}$I-pPYY in order to determine the Kd of the receptors (Y2 = 193 ± 22 pM and Y7 = 240 ± 49 pM). Ten different ligands were used in competition binding experiments to characterize the binding properties of the two receptors. Both receptors bound full-length zebrafish and porcine NPY and PYY with affinities similar to mammalian and chicken Y2. Truncated fragments of porcine NPY had lower affinity by several orders of magnitude.

The phylogenetic position of these two receptors further corroborated that Y2 and Y7 are the result of an ancient duplication in early vertebrate evolution as proposed before (Fredriksson et al., 2004). The chromosomal organization of the genes coding for these two receptors in several species further supports this notion (see paper III and paper IV). The rainbow trout Y2 and Y7 receptors share 50% sequence identity. This is similar to their orthologs in zebrafish and also in the same range as for members of the Y1 subfamily. RT-PCR was performed on 12 different tissues revealing a broad expression of both receptors. The roles for Y2 and Y7 are not well studied in fishes. Since the fishes seem to have another repertoire of peptide genes and receptor genes as compared to mammals it would be very interesting to see if Y7 has any effects on feeding behaviour. We suggest that this is likely to be the case based on the high expression of Y7 in brain and gut. However, this interpretation should be made with caution since the RT-PCR was not quantitative. An effect on food intake is also interesting because Y2 has been shown to inhibit feeding in mammals (Batterham et al., 2002).
Paper II

Two sequences with similarity to NPY receptors were identified in *Latimeria chalumnae* by degenerate PCR and screening of a genomic phage library. The sequences were found to be similar to the Y5 and Y6 subtypes found in other species. The Y5-like sequence was approximately 62% identical to tetrapod Y5 sequences while the Y6 was approximately 60% identical to mammalian Y6 sequences. Phylogenetic analysis with three different methods showed that the receptors unambiguously belong to the Y5 and Y6 clades. The Latimeria Y5 receptor contained several motifs conserved in other Y5 receptors indicating important functions. The well defined role of Y5 in feeding regulation in mammals (Lecklin et al., 2002; Lecklin et al., 2003; Day et al., 2005) show differences in receptor repertoire between mammals and teleosts since Y5 has so far not been found in euteleost fishes (Larsson et al., 2005) (and Paper III). The comparison between Y6 in *Latimeria chalumnae* and other Y6 sequences is complicated by uneven evolutionary rates in the Y6 clade. The Y6 gene is a pseudogene in many species and the function of it is presently unclear. It seems like the mammalian Y6 receptor has undergone a more rapid change while chicken (Bromee et al., 2006), *Squalus acanthias* (Salaneck et al., 2003) and *Latimeria chalumnae* have lower evolutionary rates.

Paper III

In this paper we phylogenetically characterized the NPY receptor repertoire in the two pufferfish species *Tetraodon nigroviridis* and *Takifugu rubripes*. These species have five receptor genes orthologous to those earlier described in zebrafish (Lundell et al., 1997; Ringvall et al., 1997; Starback et al., 1999; Fredriksson et al., 2004; Fredriksson et al., 2006). In addition, we compared the chromosomal regions harbouring these genes and made phylogenetic analysis of 44 gene families with members on the NPY receptor gene chromosomes in *T. nigroviridis* and *T. rubripes*. A total of 25 gene families in addition to the NPY family were found to be compatible with an expansion early in the vertebrate lineage. Statistical analysis indicates that these paralogs were not randomly distributed in the genomes of human and *T. nigroviridis*. In addition to families having members linked to three or four of the NPY receptor gene chromosomes in the *T. nigroviridis* genome, we found 127 gene families with members on two of the chromosomes. All ten combinations of chromosome pairs were represented in these families.

RT-PCR was performed on eleven tissues from *Takifugu rubripes* to investigate the expression patterns. All five receptors were found to be ex-
pressed in the brain and eye. Expression patterns were different for all five receptors in the peripheral tissues examined. RT-PCR products were sequenced to verify intron positions. Y4 was found to have an extension of the second extracellular loop by 74 amino acids in both T. rubripes and T. nigroviridis as compared to Y4 in zebrafish as well as other species. Y8a had three additional introns in both T. rubripes and T. nigroviridis as compared to zebrafish with one extra short (63 bp) exon in the middle of the third intron in T. rubripes, resulting in an extension of the second extracellular loop by 21 amino acids.

In summary, the results show that the NPY receptor family and several other gene families have expanded through block (chromosome) duplications or tetraploidizations early in the vertebrate lineage and in the teleost lineage.

Paper IV

A total of 22 sequences were obtained representing members of the Y1 and Y5 subfamilies of NPY receptors in several species of basal actinopterygians using degenerate PCR. The obtained partial gene sequences were subjected to phylogenetic analysis using the neighbor-joining and maximum likelihood methods. This analysis confirmed that Y1 orthologs are present in several species of the basal actinopterygians as well as the euteleost Danio rerio. Sequences with higher identity to Y4 than to any other Y1 subfamily sequences were obtained from all investigated species except Clupea harengus and Anguilla anguilla. The lack of these two sequences could be explained by lineage specific loss or simply failure of our PCR approach to amplify these sequences. However, Y4 orthologs have been identified in euteleosts (see (Starback et al., 1999) and paper III). Despite the increased evolutionary rate for this clade as seen earlier (Salaneck et al., 2003) and also for the Os- teoglossus bichirosum sequence cloned here, all obtained sequences included in this analysis were reliably located within the Y4 clade irrespective of the inclusion or exclusion of the rapidly evolving Y4 sequences from Takifugu rubripes and Danio rerio. The single combination of degenerate primers for the Y5 subfamily amplified fragments from Squalus acantbias, Polypterus senegalus, Acipenser baeri, Amia calva and Osteoglossus bichirosum. So far no Y5 sequence has been found in any of the euteleosts for which whole genome sequences are available. From four of the eight species investigated, namely Polypterus senegalus, Acipenser baeri, Lepisosteus osseus and Os- teoglossus bichirosum Y6 orthologs were found. Y6 forms a sister clade to Y1 but lacks the intron within the ORF found in all described Y1 genes. The primer combination used for Y6 sequences spanned the gene region containing the intron in Y1 (TM5/IL3). This was used as an additional criterion to distinguish these sequences from the Y1 sequences. No product from Clupea
Harengus was found using this primer combination. Thus, this leaves open the possibility that the obtained product from Clupea is actually Y6 and not Y1, although this seems unlikely based on its position in the phylogenetic trees. The Y1-subfamily members initially called Yb and Yc were renamed Y8b and Y8a to reflect their common origin from an ancestral Y8 receptor (see Paper III). Three sequences with a high degree of identity to Danio rerio and Takifugu rubripes Y8a and Y8b were acquired from Amia calva, Osteoglossum bicirrhosum and Clupea harengus. These three sequences were clearly identified as Y8. However, the NJ and ML trees could not clearly define these sequences as either Y8a or Y8b.

In summary, this study was undertaken to understand when the loss of several neuropeptide Y subtypes occurred in teleost fishes and also to investigate the evolutionary history of the Y8 subtype so far only found in euteleosts and frogs but not in mammals or birds. Taken together with the results obtained in Paper III, these data suggest that the Y8a and Y8b subtypes arose in 3R. The differences in receptor repertoire between fishes have interesting implications for feeding regulation in teleosts and are also interesting for understanding the evolution of new functions after duplication. Rather surprisingly, it seems that teleosts in general have fewer receptor genes as compared to the inferred number in the ancestral gnathostome despite their additional tetraploidization.

Paper V

In this paper, we used bioinformatic methods to identify and classify members of the NPY peptides in several species by combining phylogeny and positional information. Initially, BLAST searches were used to find the genes for the NPY family peptides in the genome databases of the five teleost fish species Tetraodon nigroviridis, Takifugu rubripes, Danio rerio, Gasterosteus aculeatus and Oryzias latipes for which whole genome data are available. Three genes each were identified in Danio rerio and Oryzias latipes and four genes each were found in the remaining three species. Additional database and literature searches were performed to obtain sequence representation of as many lineages as possible. In total, 82 sequences were aligned and compared. Four additional gene families (DLX, MPP, OSBPL and UPP), known to have members located close to the NPY family members and the Hox clusters in several species, were analyzed in detail. This analysis showed that the extra gene family members of the NPY family most likely arose due to duplication of large parts of the chromosomes involving the four above mentioned gene families as well as the Hox cluster and several other gene families. This result is in line with the proposed 3R tetraploidization in the actinopterygian lineage. The linkage of NPY family members to other genes also facilitated the verification of the identity of the
peptide genes found in the various fish species, leading us to propose that the
d Peptide should be named NPYa, NPYb, PYYa and PYYb to reflect their
origin from NPY and PYY ancestors shared with the genes found in other
(non-teleost) vertebrate species. This information combined with sequence
comparisons led us to suggest that the sequences previously described as
PYY and PY in *Dicentrarchus labrax* and *Paralichthys olivaceus* (Cerdá-
Reverter et al., 2000; Kurokawa & Suzuki, 2002) should be renamed NPYb
and PYYb respectively. The anatomical distribution of mRNA was investi-
gated in *Takifugu rubripes* and showed that NPYa and NPYb were expressed
predominantly in brain while the two PYY peptides had a broader tissue
distribution. The similar expression patterns of the two NPY as well as the
two PYY paralogs, respectively, also are in line with the proposed large-
scale duplication since duplication of this type would lead to duplication of
intact regulatory sequences.

**Paper VI**

In order to further characterize the chromosome regions harboring the NPY
family peptide genes and the Hox clusters, 14 unrelated adjacent gene fami-
lies were analyzed. The starting point for this analysis was the position of the
peptide genes in the genomes of *Takifugu rubripes* and *Danio rerio* com-
pared to the positions in the human genome. The inclusion of invertebrate
sequences allowed relative dating of duplication events for 13 of the ana-
alyzed families. The analyses of these families, combining positional and
phylogenetic information for *Homo sapiens, Mus musculus, Takifugu ru-
bripes, Tetraodon nigroviridis* and *Danio rerio* show that there are multiple
additional gene families with evolutionary histories similar to that of the Hox
clusters and the NPY peptide family. Our analysis further strengthens the
hypothesis that these gene families originated by block or chromosomes
duplications in line with the 2R and 3R hypotheses.
Summary and Conclusions

In Paper I, two receptor genes were cloned from the rainbow trout *Oncorhynchus mykiss*. Phylogenetic analysis of these two receptors identified them as the *Oncorhynchus mykiss* orthologs to the Y2 and Y7 receptors found in other species. Binding experiments performed with $^{125}$I-pPYY indicated similar binding affinities for the two receptors; Kd of the receptors $193 \pm 22$ pM and $240 \pm 49$ pM for Y2 and Y7, respectively. Competition binding experiments with ten different ligands also showed the two receptors to be very similar. Both receptors had affinities for full-length zebrafish and porcine NPY and PYY in the same range as mammalian and chicken Y2 receptors. Truncated fragments of porcine NPY had several orders of magnitude lower affinity, which makes these receptors different from mammalian Y2.

In Paper II, two receptor genes from the Coelacanth *Latimeria chalumnae* were cloned and phylogenetically characterized. These two receptor sequences were shown to be the Y5 and Y6 orthologs in *Latimeria chalumnae*, displaying approximately 62% and 60% identity to their mammalian orthologs, respectively.

In Paper III, the NPY receptor repertoire in the two pufferfish species *Tetraodon nigroviridis* and *Takifugu rubripes* was characterized. Both species were found to have five receptor genes. In addition, analyses of neighbouring gene families with members on these chromosomes in *Tetraodon nigroviridis* and *Takifugu rubripes* were performed. A total of 25 gene families were found to be compatible with an expansion in the vertebrate lineage, similar to that seen for the NPY receptors. Expression of all five receptors was detected in the brain and eye of *Takifugu rubripes*. Intron positions in *Takifugu rubripes* were verified by sequencing of RT-PCR products. Y4 had one unique intron and Y8a had three introns in both *T. rubripes* and *T. nigroviridis* with one short (63 bp) extra exon in the middle of the third intron in *Takifugu rubripes*.

In Paper IV, a total of 22 sequences were obtained from the Y1 and Y5 subfamilies of NPY receptors from several actinopterygian species. Phylogenetic analysis of the obtained sequences together with sequences obtained earlier showed that there are considerable differences between the actinopterygian and tetrapod repertoires. It also showed that teleosts have a similar
number of NPY receptor genes as tetrapods despite the extra genome duplication in early actinopterygians.

In Papers V and VI, the chromosomal regions where the NPY peptides are located were characterized. 14 gene families with members close to the NPY family were shown to have a similar evolutionary history as the NPY family and the Hox clusters. The analyses showed that the extra family members of the NPY family in teleosts as compared to tetrapods most likely arose due to duplication of large parts of the chromosomes, in line with the proposed 3R tetraploidization in the actinopterygian lineage. In paper V the NPY gene family was investigated in *Tetraodon nigroviridis*, *Takifugu rubripes*, *Gasterosteus aculeatus*, *Danio rerio* and *Oryzias latipes*. Four genes were found in the three former species and three genes were identified in the two latter species. The anatomical distribution of mRNA was investigated in *Takifugu rubripes* and showed that NPYa and NPYb were expressed predominantly in brain while the two PYY peptides had a broader tissue distribution.

In summary, the studies performed in this thesis show that the repertoires of both the NPY family peptides and receptors have expanded in the same time window in early vertebrate evolution. Furthermore, the peptide family underwent duplications in the teleost lineage due to duplication of large portions of chromosomes. These results are compatible with the 2R and 3R hypothesis.
To further understand the evolution of the NPY system in early chordates, we are currently searching the preliminary genome release of the cephalochordate *Branchiostoma floridae* for genes with similarity to NPY receptor genes. So far, 6 candidates have been cloned into expression vectors and tagged with GFP in order to test their expression, binding properties and functional coupling to second messenger systems. Based on fluorescence microscopy, five of the six candidates seemed to be functionally expressed in vitro. These five receptors have been tested for binding with $^{125}$I-pNPY in order to confirm their identity as NPY receptors. Initial experiments testing cAMP signaling have also been performed for these five receptors.

The full repertoire of *Xenopus tropicalis* NPY receptors and peptides as found in the genome database has been identified and are being tested in competition binding experiments. We wish to see if the receptor subtypes differ in their preferences for the tree endogenous peptides indicating sub-functionalization of these receptors and peptides. A similar characterization is also in progress for the six *Danio rerio* receptors. The characterization of the zebrafish Y1 receptor is of special interest since it is the only Y1 receptor found in euteleosts so far and as such it is important for understanding feeding regulation in this lineage. The binding properties may also give clues to important points of interaction between the peptides and receptors, as has previously been observed for chicken Y2 (Berglund et al., 2002). This can be tested by site-directed mutagenesis with the hope of improving the structural and three-dimensional models of the NPY receptors.

The approach used in Paper III, V and VI is being used to investigate the evolutionary history of several other peptide-receptor systems. We are currently investigating if the opioid peptides and receptors also have expanded in vertebrate evolution much in the same way as the NPY system.

The approach of combining phylogenetic data and positional information as done in Paper III, V and VI heavily relies on the availability of genome sequences of high quality. To be able to understand the details of the early evolution of chordates, this type of data from several additional lineages is crucial. Besides cephalochordates represented by *Branchiostoma floridae* mentioned above, several other important species have had or will have their genomes sequenced in the near future. The draft sequences of the cartilaginous fish elephant shark, *Callorhinchus milii* (Venkatesh et al., 2007) and
the sea lamprey *Petromyzon marinus* are already available and we are in the process of characterizing their NPY systems.

The identification of several gene families with similar evolutionary histories as described in Paper III, V and VI opens for the possibility to use these families that were verified to have expanded in vertebrate evolution to search for conserved synteny in several other more distantly related species. Among the candidates for this type of studies are, in addition to the species already mentioned, the urochordate *Oikopleura dioica* and the sea urchin *Strongylocentrotus purpuratus* as well as more distantly related species such as for example *Drosophila melanogaster* and *Anopheles gambiae*.

Two other interesting avenues for future studies are the comparison of regulatory regions and functional coupling of genes located in the same chromosomal regions. The available genome sequences from several species make it possible to identify noncoding sequences that have been conserved over long evolutionary timescales suggesting important regulatory functions.
I would like to thank several people that all had a big part in this. Read on! Your name might be in here somewhere!

If you don’t find your name, I apologize for this and hope that you do understand that all help and support meant a lot to me during these years, despite that I might have forgotten to mention you here.

First and foremost my supervisor Dan “nån som vill ha en Dumlekola?” Larhammar. I especially want to thank you for being very patient with all my mistakes. And also extremely kind and supportive when things didn’t go my way (not only regarding science). Without this support I would never have finished this. Thanks. My co-supervisor Svante Winberg, thanks for the trouts, for being a good and inspirational teacher during basic biology courses and for getting me into contact with Danlab in the first place. Lars-Gustaf “Ludde” Lundin, living database and good friend. I especially would like to thank you for your enthusiasm for diverse aspects of biology and also for other interesting discussions. My co-authors for helping me finishing this. Michael Conlon, Byrappa Venkatesh and Sydney Brenner.

Special thanks to Earl and Robert F for teaching me a lot when I was new in the lab and to Earl also for guiding me to the best beer in Boston and for help with reading and improving this thesis. Erik Salaneck for a few important beer sessions and for advice on how not to go crazy when finalizing a thesis. A special greeting goes to Görel “catwoman” Sundström for all the help and collaboration and for putting up with a disorganized person like me, especially during the last months. Frida Olsson for doing a lot of the hard work on one of the papers and for good company during the time in the lab, I wish you all the best.

I would like to thank Christina Bergqvist for all the help in the lab and for constantly keeping track of things that I constantly forget. Thanks a million! Ingrid Lundell for help in the lab, especially for introducing me to binding studies. And also for promoting feel-good activities such as barbecues, canoeing trips and ice-skating. Helena Ä: a walk in the woods is sometimes enough to make life feel like a walk in the park. Thanks for the support and for the swimming competition in “Linnéträdgårdens kortbanebassäng”.

43
Helena S: “-Varför ska man sjunga i moll? Det blir väl ingen jävel glad av!”
No comments needed…

Ulla, Marita, Birgitta, Lena, Emma och Maria for making all things run smoothly at the department.

Past and present Ph. D. students and co-workers at the department, especially: Tobias (a good friend for “grötfrukost” is bigger than the Internet), Robert “the little professor” Olinski for the collaboration, fishing trips and for being a good friend. Nadine, especially for sharing some weekend lunches at BMC (yes, weekends at BMC can be nice if you have the right company) and for your support and really good friendship, Torun and Paula (for help when I was new in the lab and for some memorable parties, unfortunately I don’t remember all of them), Karin N (for the nicest “disputationsfest” so far), Susanne (thanks for all the help during these years, you have a big part in this), Camilla for being a party-animal (“-Va! År klockan 2:15 jag som lovade att vara hemma 21:00”), Hanna-Linn (especially for organizing “sällskapsresan”), Janis (for darts and fishing), Charlotte I for encouragement, smiles and for the “lathund”. Bengt for the nice island trips. Finn for good discussions and collaboration. Daniel, Michalina, Lisa, Marie, Britta, Aneta, Jenny, Mattias, Cecilia, Niklas, Sigrid, Malin, Tatjana, Marlene, Erika, Henrik R and all other co-workers that makes this department a nice and/or interesting place to work.

The students I have met during courses and projects. Especially the students in the lab: Theresa, Pasi, Johan, and Ulrika. You definetly taught me more than I taught you.

Och så finns det ju ett annat liv också, har jag för mig…

Därför vill jag också passa på att framföra ett varmt tack till Finspångs godtemplareförening med förgreningar runt om i Sverige. Särskilt bör nämnas f.d. ordföranden i Halmstads kretsen Mats Lantz samt alla representanter i Östergötland, Martin Ö, Jon L, Johan “Psycho”, “Rabbi-Ronni” och alla övriga hangarounds. “I’m a godtempler, yes I am!”.

Claes “diskofil” Palmblad, för att du alltid ställer upp.


Storebroor Martin och lillebroor John.

Och till sist, Mamma och Pappa, Farmor och Mormor. Tack för både arv och miljö.
References

Ensembl database. [http://www.ensembl.org]
Fisbase. [http://www.fishbase.org]
The Pfam database. [http://www.sanger.ac.uk/Software/Pfam/]

bloc duplication in vertebrate genomes. — Nat Genet 31, 100-5.
of vertebrates or when genomic complexity was wagging its morphological tail. —
J Struct Funct Genomics 3, 45-52.
ian neuropeptide Y: Analysis by molecular cloning and computer-aided comparison
with crystal structure of avian homologue. — Proc. Natl. Acad. Sci. USA 84, 2532-
2536.
Fishes. — In: The Evolutionary Genetics of Fishes (B. J. Turner, ed). Plenum Press,
ALTSCHMIED, J., DELFGAAUW, J., WILDE, B., DUSCHL, J., BOUENEAU, L., VOLFF, J. N. &
SCHARTL, M. (2002). Subfunctionalization of duplicate mitf genes associated with
ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. (1990). Basic local
AMORES, A., FORCE, A., YAN, Y. L., JOLY, L., AMEMIYA, C., FRITZ, A., HO, R. K., LANGE-
LAND, J., PRINCE, V., WANG, Y. L., WESTERFIELD, M., EKKER, M. & POSTLETHWAIT,
282, 1711-4.
APARICIO, S., CHAPMAN, J., STUPKA, E., PUTNAM, N., CHIA, J. M., DEHAL, P., CHRISTOFFELS,
A., RASH, S., HOON, S., SMIT, A., GELPKE, M. D., ROACH, J., OH, T., HO, I. Y.,
WONG, M., DETTER, C., VERHOEF, F., PREDKI, P., TAY, A., LUCAS, S., RICHARDSON,
P., SMITH, S. F., CLARK, M. S., EDWARDS, Y. J., DOGGETT, N., ZHARKIKH, A., TAV-
TIGIAN, S. V., PRUSS, D., BARNSTEAD, M., EVANS, C., BADEN, H., POWELL, J.,
GLUSMAN, G., ROWEN, L., HOOD, L., TAN, Y. H., ELGAR, G., HAWKINS, T.,
VENKATESH, B., ROKHSAR, D. & BRENNER, S. (2002). Whole-genome shotgun as-
sembly and analysis of the genome of Fugu rubripes. — Science 297, 1301-10.
ing of a neuropeptide Y/peptide YY receptor from the Atlantic cod: the Yb receptor.
— Regul Pept 75-76, 39-43.
ATTWOOD, T. K. & FINDLAY, J. B. (1993). Design of a discriminating fingerprint for G-
protein-coupled receptors. — Protein Eng 6, 167-76.


LARHAMMAR, D., LUNDIN, L. G. & HALLBOOK, F. (2002). The human Hox-bearing chromosome regions did arise by block or chromosome (or even genome) duplications. — Genome Res 12, 1910-20.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 272

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)