Duplicated receptors for VIP and PACAP (VPAC₁R and PAC₁R) in a teleost fish, *Fugu rubripes*

J C R Cardoso¹,², D M Power²*, G Elgar¹ and M S Clark¹

¹Fugu Genomics group, MRC-HGMP Resource Centre, Genome Campus, Hinxton, Cambridge, UK
²Centre of Marine Sciences, Universidade do Algarve, Campus de Gambelas, 8000-117 Faro, Portugal

*These authors contributed equally to this work.

(Requests for offprints should be addressed to M S Clark who is currently at British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge, CB3 0ET, UK; Email: mscl@bas.ac.uk)

Abstract

Two principal groups of receptors orthologous with human PAC₁R and VPAC₁R and were identified and characterised at the genomic level in the teleost fish *Fugu rubripes*. An additional group orthologous with VPAC₂R was also identified and partially characterised. In *Fugu*, gene duplication of each of the PAC₁Rs, VPAC₁Rs and VPAC₂Rs appears to have occurred. The topology of the tree surrounding the *Fugu* duplications and other isolated piscine sequences indicates that the duplication events for these six genes clearly preceded the speciation event leading to the Cypriniformes and Tetraodontiformes and is probably teleost-specific. Overall, the combined pattern of gene expression for each pair of duplicated genes mirrored the expression in other vertebrates. However, within each pair of duplicates further specialisation had occurred, with each demonstrating differential tissue distribution profiles suggesting they may be responsible for the divergent action of the ligands, vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase-activating peptide (PACAP). The *Fugu* VPAC₁R gene regions showed conserved synteny with human chromosome 3p21·3 and also *C. elegans* chromosome X, indicating that the putative ancestral human chromosome 3 region may be equivalent to chromosome X in *Caenorhabditis elegans*.

*Journal of Molecular Endocrinology* (2004) 33, 411–428

Introduction

Vasoactive intestinal peptide (VIP) and the related peptide, pituitary adenylate cyclase-activating peptide (PACAP) belong to the glucagon hormone superfamily, which in humans includes secretin, growth hormone-releasing hormone (GHRH), glucagon, glucagon like-peptides 1 and 2 (GLP-1 and GLP-2), peptide histidine methionine (PHM) and glucose-dependent insulinotropic polypeptide (GIP) (Sherwood et al. 2000). These peptides have been grouped into a superfamily because the precursor molecules share a strikingly similar structural organisation, similar amino acid sequences and in general the hormones are related in terms of distribution and function (Bell 1986, Sherwood et al. 2000). They function by binding to the N-terminal extracellular domain of a family of specific G-protein-coupled receptors (GPCRs). Members of this hormone gene family are widespread in many phyla and have been identified in species as ancient as tunicates (McRory & Sherwood 1997). This has prompted the development of numerous models for this family’s evolution (Sherwood et al. 2000) from a single ancestral molecule by the processes of gene and exon duplication.

The genomic structures and evolution of the receptors for these small ligands are not so well characterised. The VIP and PACAP receptors belong to a sub-set of G-protein-coupled receptors (the class II family) that includes the receptors for GHRH, corticotrophin releasing factor (CRF), calcitonin, secretin, parathyroid hormone (PTH) and parathyroid hormone-related peptide (PTHrP) (Segre & Goldring 1993). The members of this family share amino acid sequence similarity in the seven transmembrane domains and a long N-terminal extracellular domain containing a number of cysteines that are important for ligand
binding (Segre & Goldring 1993, Laburthe et al. 1996). All the receptors mediate their action by activation of a G-protein that induces an intracellular signal cascade mediated by hetero-meric GTP-binding proteins (Gilman 1987, Bock- aert 1991, Ulrich et al. 1998). Ligand-binding studies have demonstrated, in humans, the existence of three VIP-interacting receptors: VPAC1R, VPAC2R and PAC1R (the PACAP type 1 receptor). The VPAC receptors bind both PACAP and VIP ligands with similar affinities, whilst the PACAP type 1 receptor preferentially binds PACAP (Vaudry et al. 2000c). The two human VIP receptors, VPAC1R and VPAC2R, share less than 42% amino acid sequence similarity and have a widespread tissue distribution with different expression patterns for each of the receptors, as demonstrated by RT-PCR and in situ hybridisation studies. VPAC1R is found mainly in the cerebral cortex, hippocampus, pineal gland, lungs and intestines and VPAC2R, which is proposed to be the neuro-endocrine receptor for VIP, is found in the hypothalamus, pituitary, adrenal glands, pancreatic islets, testes and ovaries (Usdin et al. 1994, Vaudry et al. 2000a,b). In mammals, PAC1R is expressed predominantly in the CNS, however the situation is more complex, with the production of alternative splice variants (hip-hop cassettes) which influence receptor selectivity for the PACAP38 and PACAP27 ligand isoforms (Spengler et al. 1993, Journot et al. 1994, Pantaloni et al. 1996, Chatterjee et al. 1997, Daniel et al. 1998, Dautzenberg et al. 1999). Isolation and functional characterisation of the PAC1R in goldfish indicates that the PACAP ligand acts as a novel GH-releasing factor in this organism (Wong et al. 2000).

The isolation and characterisation of two VIP receptor subtypes (VPAC1R and VPAC2R) has been reported only in humans (Sreedharan et al. 1993, Gagnon et al. 1994), rat (Ishihara et al. 1992, Lutz et al. 1993) and mouse (Inagaki et al. 1994, Hashimoto et al. 1999) which raises interesting questions about when the gene duplication that gave rise to the two forms occurred. Relatively few receptors of this superfamily have been isolated from non-mammalian vertebrates (Peeters et al. 1999, Alexandre et al. 1999, Hu et al. 2000). In teleosts, only goldfish VPAC1R and PAC1R cDNAs have been isolated (Chow et al. 1997, Wong et al. 1998) and no genomic structures for these receptors have been reported. In the present study the compact genome of the Japanese pufferfish (Fugu rubripes) was used to isolate and characterise the PAC1, VPAC1 and VPAC2 receptor genes at the genomic level. The gene sequences obtained in the present study were also used to search for related genes in the invertebrate databases Caenorhabditis elegans, Ciona intestinalis, Drosophila melanogaster and Saccharomyces cerevisiae with the aim of identifying a putative ancestral VIP/PACAP receptor.

Materials and methods
Gene identification and isolation
During routine Fugu cosmid library sequence scanning (within the confines of the Fugu Landmark Mapping Project (Elgar et al. 1999)) a cosmid (C139D21) was identified as having a gene fragment with high sequence similarity to goldfish VPAC1R (probability value in excess of 9e-19). PCR primers were designed for a conserved sequence motif of this family identified after a multiple sequence alignment of all the members of this family and the Fugu cosmid C139D21 partial sequence. A homologous PCR product was obtained with Fugu genomic DNA and used to screen the Fugu cosmid library (available from http://www.hgmp.mrc.ac.uk) at low stringency using 50 °C hybridization temperature in Church Gilbert buffer (1984), followed by a one minute wash at room temperature in Church Gilbert wash buffer. Twenty four further cosmids were isolated and characterised by sequence scanning as described in http://fugu.hgmp.mrc.ac.uk/Protocols/Biology/fugu_section2.html#2.1. The cosmids were sorted into different groups based on DNA sequence similarity, secondary screening of SacI restriction enzyme digest Southern blots and analysis of short-range linkage relationships. Gene content was determined by searching the individual sequence fragments using BLAST v2.0 (Altschul et al. 1997) against the SPTR (Bairoch et al. 1997) and Unigene databases (http://www.ncbi.nlm.nih.gov/Unigene/Hs.home.html).

Cosmids were identified for a single PAC1R gene and for each of two distinct VPAC1R-like and two VPAC2R-like genes. Cosmids for each of the genes were subjected to full-depth sequencing. The cosmid containing PAC1RA lacked the first exon, so this gene was completed using RACE PCR.
Duplicate VIPRs in a teleost fish · J C R CARDOSO, D M POWER and others

(RACE Ex4, 5’actctggacagttgacga3’). A further PAC1R gene was identified from Fugu Consortium data (Aparicio et al. 2002) and mapped to two scaffolds: exons 2 and 3 mapped to scaffold S011651 and S004929 contained exon 6 to the 3’ end. The first exon was identified using RACE PCR (PACAPP2 RACE2 5’accagatgggacactctggtcag3’; PACAP2Ex6R, 5’tatgtatgaagctgcctgt3’). Probing of a Southern blot, revealed the genomic distance of these two exons to cover approximately 1 kb, but the gene was completed by RT-PCR using Fugu brain cDNA. Fugu database mining revealed no further VIP or PACAP receptors.

For gene assembly, all sequences were transferred to a UNIX environment and quality clipped using a modified Pregap script (Bonfield and Staden 1996). Sequences were screened against sequencing and cloning vectors and matching regions were masked prior to further analysis. Edited sequences were then assembled using the Pint Assembly Programme (Weston P, HGMP, unpublished data, http://menu.hgmp.mrc.ac.uk/cgi-bin/pint/). Contiguation and finishing of the gene was performed using PCR walks and sequencing of SacI cosmid sub-clones. All genes, unless stated in the results, were fully sequenced at the genomic level. Exons were amplified using primers designed to exons 2 and 6 (PACAP2Ex3F, 5’caggcacacgctctagat3’; PACAP2Ex6R, 5’tatgtatgaagctgcctgt3’). Probings of a Southern blot, revealed the genomic distance of these two exons to cover approximately 1 kb, but the gene was completed by RT-PCR using Fugu brain cDNA. Fugu database mining revealed no further VIP or PACAP receptors.

For gene assembly, all sequences were transferred to a UNIX environment and quality clipped using a modified Pregap script (Bonfield and Staden 1996). Sequences were screened against sequencing and cloning vectors and matching regions were masked prior to further analysis. Edited sequences were then assembled using the Pint Assembly Programme (Weston P, HGMP, unpublished data, http://menu.hgmp.mrc.ac.uk/cgi-bin/pint/). Contiguation and finishing of the gene was performed using PCR walks and sequencing of SacI cosmid sub-clones. All genes, unless stated in the results, were fully sequenced at the genomic level. Exons were determined by comparison with other vertebrate VPACR/PAC1R sequences and searching for the GT/AG consensus sequences. Anomalous exon/intron boundaries were confirmed by RT-PCR.

The Fugu VPAC1R, VPAC2R and PAC1R genes have been submitted to the EMBL database with accession numbers AJ296144, AJ296143, AJ408877, AJ40887, AJ490804 and AJ494861 for VPAC1R, VPAC2R, VPAC1R, VPAC2R, VPAC1R and PAC1R respectively.

Fugu VIPR/PACAPR gene analysis and database mining

Analysis of the finished sequence was carried out using the HGMP Nix interface (G Williams, P Woollard & P Hingamp, unpublished data, http://www.hgmp.mrc.ac.uk/NIX/). Percentage of identities of the different genes was calculated with the Blosum 62 score table using the Genedoc program (Multiple Sequence Alignment Editor and Shading Utility v2.6.001, (www.psc.edu/biomed/genedoc) and also using the needle programme from the EMBOSS suite of programmes (Rice et al. 2000) (http://www.uk.embnet.org/Software/EMBOSS/Apps/) with a Gap opening penalty of 10-0 and a Gap extension penalty of 0-5. Comparative human gene map positions were determined using the Ensembl viewer (v1.0.0) (http://www.ensembl.org/) and gene names assigned using the HUGO nomenclature accessed via the Ensembl link to LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink/).

Each of the Fugu VIP/PACAP receptor genes were used to search for similar genes in the human, zebrafish, Tetraodon, medaka, C. elegans, D. melanogaster, C. intestinalis and S. cerevisiae databases: (http://www.ensembl.org/perl/blastview; http://www.ensembl.org/Danio_rerio/; http://www.genoscope.cns.fr/cgi-bin/recherche.cgi; http://shigen.lab.nig.ac.jp/medaka/genome/top.jsp; http://www.sanger.ac.uk/Projects/C_elegans/blast_server.shtml; http://www.fruitfly.org/blast; aluminium.jgi-psf.org/prod/bin/RunBlast.pl?db=ciona4 and http://genome-www.stanford.edu/Saccharomyces/. To establish if linkage groups had been maintained during the evolution of the VIP/PACAP receptor genes in Fugu, human and C. elegans, orthologous genes for the genes surrounding Fugu VPAC1R, VPAC2R and PAC1R were searched and mapped in the human and C. elegans genomes using the http://www.ncbi.org/LocusLink and http://www.wormbase.org/db/searches/basic#searchagain web sites.

Phylogenetic analyses

Multiple sequence alignments were carried out using Clustal X (Thompson et al. 1997) and in the Genedoc program (Multiple Sequence Alignment Editor and Shading Utility v2.6.001, (www.psc.edu/biomed/genedoc)). Phylogenetic analysis was carried out from the Clustal X multiple alignment output using both the neighbour joining method (Saitou and Nei 1987) via the PHYLO_WIN interface v1.2 (Galtier et al. 1996) and maximum likelihood method (Felsenstein 1985) via Phylip v3.57c. Multiple alignment parameters for PHYLO_WIN: gap opening 15; gap extension 0-05; delay divergent sequences 40%; DNA transition weight 0-50, with 1000 bootstrap replicates.

Maximum likelihood analysis was carried out using
the dnaml programme with default parameters and global rearrangements. The human GHRH clone (accession number, P43220) was used as the outgroup for the tree construction.

Tissue-specific expression

Tissue-specific expression of each of the genes was carried out by RT-PCR on eight tissues from different sexually immature juvenile fish (brain, gonads, gut, spleen, liver, kidney, heart and gills) generated by first strand synthesis of mRNA using the Promega Reverse Transcriptase kit according to the manufacturer's instructions. PCR was carried out on 1µg cDNA using specific primer pairs designed for individual genes and spanning an intron as a control check for potential genomic contamination. Primers pairs for each gene: VPAC1RA (V1AE4 – 5’gtgaacctctcaagcttgc3’; V1AE5 – 5’tgctgatgaaacgcctggag3’); VPAC1RB (V1BE2 – 5’tttctctctgggaaattcgca3’; V1BE8 – 5’tagtggctccaaatgggtg3’); VPAC2RA (V2AE3 – 5’agcgcgcgctggaggtta3’; V2AE7 – 5’gccatgacaaagtagttgac3’); VPAC2RB (V2BE4 – 5’atcttccaaacatcacc3’; V2BE7 – 5’gcggttctcagagaagattac3’); PAC1RA (P1AE2F – 5’gtggcgtcagagttgaatcac3’); PAC1RB (P1AE7R – 5’atcaccctggttgcaagcgc3’); PAC2RA (P1BE3F – 5’ctgctgcagcagctggt3’); PAC2RB (P1BE3R – 5’ctgttgaggccatcagctg3’). PCR was performed with an initial denaturing step at 96 °C for 2 min, cycled 35 times (96 °C for 1 min, 62 °C for 1 min, 72 °C for 1 min) with a final extension of 72 °C for 5 min. The VPAC1RA PCR was performed at 67 °C with the same conditions described above. All PCR products of the correct size were sequenced to confirm their identity. To ensure that similar quantities of cDNA were used in all RT-PCR reactions, a further PCR with primers designed for ubiquitously expressed β-actin genes was also performed (data not shown).

Results

Screening of the Fugu genomic cosmids library with a probe generated by PCR to a conserved sequence motif of the family resulted in the identification of four different VIP receptors. They were named VPAC1RA, VPAC1RB, VPAC2RA and VPAC2RB, according to their sequence similarity with the other vertebrate orthologous genes. Due to the difficulty in identifying the duplicated VIP receptor genes at the level of sequence scanning, the individual genes were characterised by their cosmid short-range linkage relationships (Table 1), sequence similarity and restriction digestion patterns. The A and B nomenclature was created to facilitate the distinction between the duplicated genes.

Fugu VPAC1RA and VPAC1RB genes

The Fugu VPAC1RA and VPAC1RB genes were completely sequenced at the genomic level and both code for a 419 amino acid protein, sharing 68% sequence identity. Comparison of the deduced amino acid sequence of VPAC1RA and VPAC1RB revealed they shared 51% and 49% sequence similarity with human VPAC1R and 75% and 64% similarity to goldfish VPAC1R, the only teleost in which the sequence is available. Structurally the Fugu VPAC1R genes are composed of 12 exons and are 14-5 kb (VPAC1RA) and 8-5 kb (VPAC1RB) in length. Exon/intron positioning, splice site sequences and exon/intron boundary phases are identical between the two Fugu sequences. The two genes in Fugu, in common with the goldfish VPAC1R gene (Chow et al. 1997), lack the first exon present in mammals that encodes the signal peptide sequence. To date, only two mammalian VPAC1R genomic sequences from mouse (Hashimoto et al. 1999) and human (Sreedharan et al. 1993) have been characterised. These two genes are both composed of 13 exons and are 16 kb and 23 kb in length respectively. In comparative terms, the genes in Fugu are of a similar size (Fig. 1) and do not show the characteristic gene compaction of this organism (Brenner et al. 1993, Elgar et al. 1999). In general, with the exception of the absent first exon in both Fugu VPAC1Rs, the gene organisation between human and Fugu is very similar, sharing the same splice sites, intron/exon boundary phases and location of the transmembrane domains. Sequence alignment of VPAC1Rs from various species (Fig. 1B) indicate that they are very similar and signature motifs such as the RLAK motif between TM5 and TM6 which has been indicated for the functional coupling to Gsa (Okamoto et al. 1991) and the PDI motif typical of VIP binding receptors, were identified. Other motifs such as, the conserved R and L residues localised within TM2 which have
Table 1 VIP receptor genes in *Fugu*, humans (VPAC1R and VPAC2R) and *C. elegans* (AL031620) and their short-range linkage relationships. *Fugu* genes are listed according to their order in the *Fugu* genome. Human chromosome positions are indicated in brackets and the putative *C. elegans* homologous genes are characterised by their probability value obtained by sequence comparison and chromosome position.

<table>
<thead>
<tr>
<th>Fugu VIPR and linked genes</th>
<th>Gene description</th>
<th>Human homologue gene</th>
<th>Putative <em>C. elegans</em> homologue genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothetical</td>
<td></td>
<td>KIAA1173 (3p24-3-p22·1)</td>
<td>Z93395 (0·995) II</td>
</tr>
<tr>
<td>LOC131377</td>
<td>Protein of unknown function</td>
<td>LOC131377 (3p21-32)</td>
<td>AF067946 (6·1e–32) V</td>
</tr>
<tr>
<td>NKTR</td>
<td>Kelch-related protein 1 (Sarcosin)</td>
<td>NKTR (3p23-p21)</td>
<td>U40938 (1·61e–30) X</td>
</tr>
<tr>
<td>SEC22C</td>
<td>Natural killer tumor recognition protein</td>
<td>SEC22C (3p24-3-p22·1)</td>
<td>Not found</td>
</tr>
<tr>
<td>VPAC1RA</td>
<td>Vesicle trafficking protein</td>
<td>VPAC1R (3p22)</td>
<td>AL31620 (7·7e–27) X</td>
</tr>
<tr>
<td>CASQ</td>
<td>Vasoactive intestinal peptide receptor 1A</td>
<td>CASQ (1q21)</td>
<td>Z69792 (1·5e–09) X</td>
</tr>
<tr>
<td>E2BE</td>
<td>Calsequestrin muscle skeletal isoform percursor</td>
<td>E2BE (3q27-1-q24)</td>
<td>Z54284 (7·8e–12) II</td>
</tr>
<tr>
<td>VPAC1RB</td>
<td>Translation initiation factor EIF-2B epsilon subunit</td>
<td>VPAC1R (3p22)</td>
<td>AL31620 (2·5e–27) X</td>
</tr>
<tr>
<td>LOC131377</td>
<td>Vasoactive intestinal peptide receptor 1B</td>
<td>LOC131377 (3p21-32)</td>
<td>AF067946 (6·1e–32) V</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Kelch-related protein 1 (Sarcosin)</td>
<td>KIAA1173 (3p24-3-p22·1)</td>
<td>Z93395 (0·995) II</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Protein of unknown function</td>
<td>KIAA0934 (10p15·3)</td>
<td>Not found</td>
</tr>
<tr>
<td>VPAC1RA</td>
<td>Hypothetical protein of unknown function</td>
<td>VPAC1RA (7q36-3)</td>
<td>AL31620 (1·5e–24) X</td>
</tr>
<tr>
<td>KIAA1585</td>
<td>Vasoactive intestinal peptide receptor 2A</td>
<td>KIAA1585 (3p12-3-q24)</td>
<td>Not found</td>
</tr>
<tr>
<td>TNSF10</td>
<td>Protein fragment of unknown function</td>
<td>TNSF10 (3q26-31)</td>
<td>Not found</td>
</tr>
<tr>
<td>BS69</td>
<td>Tumor necrosis factor ligand superfamily member 10</td>
<td>BS69 (10p14)</td>
<td>AF003389 (2·0e–10) II</td>
</tr>
<tr>
<td>Hypothetical</td>
<td>Adenovirus 5 E1-A binding protein</td>
<td>KIAA0934 (10p15·3)</td>
<td>Not found</td>
</tr>
<tr>
<td>KIAA1228</td>
<td>Hypothetical protein of unknown function</td>
<td>KIAA1228 (7q36-3)</td>
<td>U13019 (3·6e–29) III</td>
</tr>
<tr>
<td>VPAC2RB</td>
<td>Protein Kinase conserved region 2-protein fragment</td>
<td>VPAC2RB (7q36·3)</td>
<td>AL31620 (7·7e–30) X</td>
</tr>
<tr>
<td>PAC1RA</td>
<td>Vasoactive intestinal peptide receptor 2B</td>
<td>PAC1RA (7p14)</td>
<td>AL31620 (3·8e–30)</td>
</tr>
<tr>
<td>GHRHR</td>
<td>Pituitary adenylate-cyclase activating polypeptide receptor 1</td>
<td>GHRHR (7p14)</td>
<td>AL31620 (8·0e–23)</td>
</tr>
<tr>
<td>PAC1RB</td>
<td>Growth Hormone Releasing Factor Receptor</td>
<td>PAC1RB (7p14)</td>
<td>AL31620 (3·8e–30)</td>
</tr>
<tr>
<td>GHRHR</td>
<td>Growth Hormone Releasing Factor Receptor</td>
<td>GHRHR (7p14)</td>
<td>AL31620 (8·0e–23)</td>
</tr>
</tbody>
</table>
been suggested as being responsible for the receptor activation (Solano et al. 2001) and a highly conserved region at TM7 FQGGBVXXBYCFXNXEXQ which has been designated the consensus signature motif for the mammalian class II family members (Lok et al. 1994), where X and B represent any amino acid residue and any hydrophobic amino acid residue respectively, were also found in Fugu. In addition to these motifs, seven conserved cysteines at the 5’ end and two conserved and three additional putative glycosylation sites were also identified in the Fugu genes (Fig. 1B).

**Fugu VPAC_{2}R and VPAC_{2}RB genes**

Two different putative VPAC_{2}R genes were identified in *Fugu*, VPAC_{2}RA and VPAC_{2}RB. Genomic sequencing of each encompassed 7·4 kb and 12·2 kb respectively, to the 5’ end of each cosmid. Neither of the cosmids containing the two VPAC_{2}R-like genes appeared to contain the 5’ end of the genes, using homology searching and exon prediction software. Despite repeated attempts at RACE PCR, screening of the Fugu Consortium database (http://fugu.hgmp.mcc.ac.uk) and heterologous screening of a sea bream lambda cDNA library, it was not possible to identify the 5’ terminal two exons of either VPAC_{2}R gene (12 exons were characterised). A sea bream clone for a VPAC_{2}R gene was isolated from the cDNA library, but this was truncated at the 5’ end. Comparative analysis of the partial Fugu VPAC_{2}R genes with the VPAC_{2}R sequences from mouse, human, and rat revealed, as already observed for the VPAC_{1}R genes, the presence of seven conserved transmembrane domains (Fig. 2). The conserved motif characteristic of the VIP binding receptors, P/A-D-V and the RLAK motif for the functional coupling to Gsα were also identified. As for Fugu VPAC_{1}R, the consensus signature identified for the mammalian class II family members at the TM7 FQGGBVXXBYCFXNXEXQ was also identified (Fig. 2).

**Fugu PAC_{1}R RA and PAC_{1}R RB genes**

Two PAC_{1}R genes have been identified in *Fugu*. The first gene (PAC_{1}R RA) was identified from heterologous probing of the *Fugu* cosmid filters and a further gene (PAC_{1}R RB) was identified from the Fugu Consortium data (Aparicio et al. 2002). To date only one PAC_{1}R RA receptor has been found in other vertebrates and previously only the genomic structure of rat and mouse have been characterised. The estimated sizes of each of the genes in *Fugu* was determined as 11·2 kb (PAC_{1}R RA) and 5·7 kb (PAC_{1}R RB), compared with 40 kb and 50+ kb for rat and mouse respectively (Hashimoto et al. 1993, Aino et al. 1995), representing a varying compaction rate in *Fugu* (Fig. 3). Comparative analysis of the deduced amino acid sequence of *Fugu* PAC_{1}R RA and PAC_{1}RB revealed that they are the most conserved of the three family duplications, sharing 74% sequence identity with each other and both share 69% sequence identity with human PAC_{1}R. However, such calculations are more complex with this receptor as the length of the PAC_{1}Rs varies between species from 465 amino acids in *Xenopus* to 523 amino acids in rat, due to the presence of the hip-hop cassettes which have only been sequenced in some species and were not identified in *Fugu*. The *Fugu* PAC_{1}Rs are both 444 amino acids in length and share the same genomic organisation. Comparative analysis of the *Fugu* PAC_{1}R genes with the PAC_{1}R sequences from other vertebrates.
revealed the presence of seven conserved transmembrane domains (Fig. 4). All potential conserved cysteines were present in the Fugu sequences and two conserved and two non-conserved N-glycosylation sites were also identified. The conserved motif characteristic of the PACAP binding receptors, PDM and the RLARS motif for the functional coupling to Gs were also identified and a putative signal peptide sequence, albeit shorter than the sequence reported in mammals. As for the Fugu VIP receptors, the consensus signature identified for the mammalian class II family members at the TM7 FOQGBBVXXBYCFXNX EVXQ was also identified (Fig. 4).

Gene analysis and short-range linkage of Fugu VPAC/PAC1 receptor genes

The gene environments linked to each of the Fugu VPACRs were different (Table 1) confirming the identification of four different VIP receptor genes in Fugu. Human VPAC1 is located on chromosome region 3p22. Fugu VPAC1 was linked to four genes, all of which map in humans to chromosome region 3p22. The gene environments linked to each of the Fugu VPACRs were different (Table 1).
Figure 3 Comparison between rat PAC1R and the two orthologous genes in Fugu (PAC1RA; PAC1RB). The gene sizes are indicated next to each gene on the right, exons are numbered and are represented by closed boxes and the introns by a black line.
Figure 4 Clustal X alignment of the Fugu PAC1R protein coding regions with other known PAC1Rs. The transmembrane domains are annotated above the sequence. Conserved cysteines are annotated by a circle and potential conserved N-glycosylation sites are marked with a triangle. The alternatively spliced hop cassette is shown with a dotted box. Transmembrane domains (TM) and signal peptide sequences were identified by sequence comparison and according to the GPCR database (www.cmbi.ni/7tm/) using the Swiss-Prot annotation. Accession numbers of the sequences used are mouse, P70205; rat, P32215; human, P41586; bovine, D17290 and goldfish, AF048820. The chicken gene is taken from Peeters et al. 1999.

J C R CARDOSO, D M POWER and others · Duplicate VIPRs in a teleost fish

www.endocrinology-journals.org

Journal of Molecular Endocrinology (2004) 33, 411–428

Downloaded from Bioscientifica.com at 03/26/2019 10:58:55PM via free access
similarity to the *Fugu* VIP receptor sequences. These genes mapped within 6 Mb of each other. The *C. elegans* clone AL31620, located on chromosome X, was the only one with the highest identity for all of the *Fugu* VPAC/PAC1 receptors (probability in excess of $e^{-24}$) and a multiple sequence alignment (data not shown) of this clone with the other *Fugu* VPAC/PAC1 receptors revealed that they shared high sequence similarity for the TM domains. This clone also shared conserved linkage for the calsequestrin gene with the *Fugu* VPAC1RB. *Fugu* VPAC1RA was also linked to a gene (NKTR), which in *C. elegans* mapped to the same chromosome (X) as the putative VPAC1R clone and calsequestrin.

**Phylogenetic analysis**

Phylogenetic analysis was carried out on VIP/PACAP receptor genes using both the Neighbour Joining and Maximum Likelihood methods. Both methods produced the same relationships, of which the results from the Neighbour Joining method are shown in Fig. 5. The analysis clearly indicates that the ‘extra’ genes in *Fugu* clustered with the other VPAC1R, VPAC2R and PAC1R genes and not with the closest family members (GHRH and secretin) and confirmed the initial identification made by sequence similarity and linkage. Of the genes considered, only a single PAC1R and VPAC1R have been previously identified and entered in the database from a teleost (goldfish). The topology of the tree surrounding the *Fugu* duplications and the goldfish sequences indicates that these duplication events preceded the speciation event leading to divergence of the Cypriniformes and Tetraodontiformes. Therefore it is expected that duplicate VPAC1R, VPAC2R and PAC1R genes will be found in other teleosts. The relationship between the three duplications is difficult to determine with low bootstrap values and it is not possible to determine which of the receptors arose first.

**Tissue distribution**

RT-PCR was carried out using specific primers for each receptor in order to characterise the expression of each of the receptors in different *Fugu* tissues (data not shown). Although most of the receptors were expressed in gonads, brain and gills, there was clearly a difference in the expression patterns between each of the duplicated genes.

**Discussion**

*Fugu* is the first organism in which duplication of PAC1R, VPAC1R and VPAC2R has been identified. Comparison of the deduced amino acid sequences of the VIP/PACAP receptors from *Fugu* with other family members isolated from vertebrates revealed that they exhibited greatest sequence similarity to VIP/PACAP receptors and contained all the characteristic features and signature motifs of the class II GPCR family members. Structural comparison of *Fugu* receptors with the orthologous VPAC1R and PAC1R in humans shows that gene organisation has been highly conserved (as indeed was the case for the partial sections of the *Fugu* VPAC2Rs). The main difference observed is that the first exon, which in mammalian VPAC1R contains the signal peptide, is absent in both VPAC1R receptors in *Fugu*. This situation is identical to that of goldfish, in which VPAC1R was cloned as a cDNA and this exon is also absent, although the 5’ UTR is present and the receptor is functional. Studies were carried out using goldfish VPAC1R expressed in a mammalian cell line (Chow 1997) which showed that in goldfish the receptor stimulated the production of cAMP in the transfected cells, suggesting that the amino acids encoded by the first exon are not essential for receptor activity. It remains to be established if the fish VPAC1R is integrated into the cell membrane and the existence of an alternative non-consensus signal peptide sequence in fish remains to be explored. Of the three duplicate receptors in *Fugu*, the PAC1R genes are the most conserved in terms of sequence similarity and identity, they both encode a putative signal peptide, albeit shorter than that identified in other vertebrates so far. It was not possible to determine whether the *Fugu* VPAC2R genes contained a signal peptide as the 5’ end remains uncharacterised.

It is thought that all three receptors under study originated from a single ancestral gene which duplicated and subsequently diverged during the course of evolution (Ishihara et al. 1992, Lutz et al. 1993, Pisegna & Wank 1993, Inagaki et al. 1994).
However, the topology of the initial branching between the PAC$\textsubscript{1}$Rs, VPAC$\textsubscript{1}$Rs and VPAC$\textsubscript{2}$Rs is not clearly defined, with low bootstraps, and it is not possible to determine which gene represents the more ancestral form of these receptors.

One theory about the evolution of the receptors is based on the chromosomal localisation of the
receptors in humans and rats. Both VPAC$_2$R and PAC$_1$R map to human chromosome 7 and rat chromosome 4, whereas VPAC$_1$R is located on human chromosome 3 and rat chromosome 8. A first duplication would produce the VPAC$_1$R gene and a common ancestor for VPAC$_2$R/PAC$_1$R. A second duplication event acting on the ancestral VPAC$_2$R/PAC$_1$R gene would produce the two separate genes (Cai et al. 1995, Sreedharan et al. 1995, Brabet et al. 1996, Mackay et al. 1996, Vaudry et al. 2000c).

Short-range linkage of the Fugu VIP receptor genes with human genes shows conserved synteny for some of the genes, especially for VPAC$_1$R-linked genes on human chromosome region 3p. This situation may be an indicator of a specific gene duplication event that occurred in this particular chromosome in the teleost lineage and so far the locus site for these receptors in fish has still to be mapped, so it is not known whether the Fugu VIP receptors linked to human chromosome 3 genes are very close on the same chromosome or on different chromosomes. VPAC$_1$RB shows conserved synteny for three genes LOC131377, E2BE and KIAA1173, the latter two of which in human are on opposite arms of chromosome 3, megabases apart. However, this is not surprising as studies in both zebrafish (Postlethwait et al. 2000) and Fugu (Bouchireb et al. 2001) have demonstrated that whilst fish may share large blocks of conserved synteny with humans, potentially extending over a whole human chromosome, the gene order can be radically altered by numerous inversion events. Fugu VPAC$_2$RA also shows similar linkage data to the Fugu VPAC$_1$R genes since most of the linked genes identified map to human chromosome 3. Whilst Fugu VPAC$_2$RB shares the same locus as the human VPAC$_2$R gene, as regards linkage to the KIAA1228 gene (7q36-3), there is stronger linkage to human chromosome 10p in the form of two genes linked to VPAC$_2$RB (BS69 and KIAA0934) and additionally VPAC$_2$RA is linked to a further duplicated KIAA0934. It is therefore clear that the duplication event that produced VPAC$_2$RA and VPAC$_2$RB also encompassed these KIAA0934 genes. These data support the idea that VPAC$_1$R and VPAC$_2$R arose from a gene duplication event of a common ancestor and suggest that in teleosts this duplication event occurred within the same chromosome followed by translocation of VPAC$_2$RB away from the putative ancestral 3p region. This was then followed by a further duplication to produce two separate genes VPAC$_2$R and PAC$_1$R, within an environment similar to human chromosome 7. This Fugu linkage data further supports the hypothesis of vertebrate VIP/PACAP receptor evolution based on human chromosomal locations described earlier (Cai et al. 1995, Sreedharan et al. 1995, Brabet et al. 1996, Mackay et al. 1996, Vaudry et al. 2000c).

Phylogenetic analysis suggests that VPAC$_1$RA and VPAC$_1$RB; VPAC$_2$RA and VPAC$_2$RB; and PAC$_1$RA and PAC$_1$RB in Fugu are duplicated forms of the VPAC$_1$R, VPAC$_2$R and PAC$_1$R in other vertebrates. Despite the relatively restricted number of species from which VIP receptors have been isolated, the topology of the phylogenetic tree suggests that duplication of the receptors in Fugu occurred prior to the branch that gave rise to mammals and, in line with current theories, is teleost-specific. In the goldfish and salmon, RT-PCR for VPAC$_1$R with degenerate primers amplified two partial cDNAs for VPAC$_1$R in both species. However, the authors only fully characterised one cDNA as the second fragment was assumed to arise because of the tetraploid nature of the genome in Salmoniformes and Cypriniformes (Chow et al. 1997). However, sequence comparison of the two distinct VPAC$_1$R cDNAs isolated from both salmon and goldfish with the Fugu VPAC$_1$Rs using phylogenetic methods clustered one of each of the goldfish and salmon cDNAs with Fugu VPAC$_1$RA and the others with Fugu VPAC$_1$RB (data not shown). Data-mining of other sequenced fish genomes further supports the timing of these teleost-specific duplication events, as orthologues of all the duplicated Fugu VIP receptors were identified in Tetraodon and a further two for VPAC$_2$A and PAC$_1$A were identified in zebrafish.

It is now a well-established fact that there are ‘extra’ genes in fish (Wittbrodt et al. 1998). However, the process by which these arose is still under debate. One proposal suggests that a further round of whole genome duplication has occurred in the fish lineage (Force et al. 1999, Taylor et al. 2001, 2003, Naruse et al. 2004, Vandepoele et al. 2004, Christoffels et al. 2004). Alternatively, these may have arisen due to a whole series of minor duplication events and not the single large-scale duplication originally proposed (Robinson-Rechavi et al. 2001a,b,c, Hughes et al. 2001). This data (together with the data mining and anecdotal...
evidence in other fish species) clearly shows that there has been a minimum of several large chromosomal segmental duplications in teleosts involving the VIP receptors, producing further support for the hypothesis of an extra whole genome duplication having taken place in the teleost ancestor. Identification and characterisation of duplicated genes, particularly gene family members, provides important data towards the more large-scale, class-level (Teleostei) phylogenetic analyses cited above. Analysis of these large data sets, whilst necessary for a complete understanding of genome evolution, can contain weaknesses, as the history of specific gene families may not be completely understood or investigated, rendering gene loss and local gene duplications difficult to take into consideration. Analysis of a particular gene family (similar to that presented here) can identify these small-scale events. Whilst the issue of an extra whole genome duplication event in fish remains a contentious issue, the preponderance of polyploid genomes in fish certainly complicates genome analysis and a resolution to this issue. It also raises a further issue, as to why these events occurred and the benefits for the organisms involved. It is generally acknowledged that gene duplication has played a significant role in the metazoan radiation. Further duplications in the fish could have fuelled their incredible speciation, as they comprise over half of all vertebrate species and unlike mammals their genomes (and the ploidy levels) are not constrained by a rigid sex chromosome system. The procession of duplicated genes with differential expression patterns may also, to a certain extent, obviate the need for alternative splicing.

The four VPAC and 2 PACI receptors identified in *Fugu* were used to search for receptor homologues in the *C. elegans* genome (Fig. 6). A single gene localised on chromosome X that shared a high level of sequence similarity with all the *Fugu* VPAC1Rs was identified and assumed to be an ancestral form of VIP/PACAP receptor (AL31620, probability value higher than e-24 using BLASTX). It was

---

**Figure 6** Comparative maps showing conserved linkage between *Fugu*, human and *C. elegans*. The chromosome positions of the genes in human and *C. elegans* are given in Mb and annotated next to the gene name. KIAA1173, LOC131377, NKTR, SEC22C and VPAC1R in human all map to a gene cluster on 3p21·32 which contains only two other genes.
considered that short-range gene linkage analysis might help to substantiate the phylogenetic model of how the VIP receptors evolved from the ancestral form in *C. elegans*. *Fugu* VPAC₁RA is linked to the gene calsequestrin, which is also present in *C. elegans* (Z69792) on the same chromosome as the putative ancestral VPAC₁R. Additionally, VPAC₁RA shows linkage to another gene (NKTR) which maps in *C. elegans* (U40938) to chromosome X. Taken as a whole, this data potentially indicates that the ancestral VPAC₁R gene and immediate gene environment in *C. elegans* was on chromosome X (Fig. 6) and that this may be equivalent to the putative ancestral human chromosome 3 region.

The widespread tissue distribution of the two VIP receptors in mammals may explain why VIP has such a wide spectrum of activity. In mammals it is an important regulator of the digestive tract (Ulrich et al. 1998), it is also involved in the immune (Pozo et al. 2000), reproductive (Usdin et al. 1994), circulatory (Fahrenkrug 1993) and neuro-endocrine systems (Vijayan et al. 1979, Biancani et al. 1985, Sherwood et al. 2000).

In fish, far fewer studies of VIP function exist but it has been proposed to be of importance in osmoregulation (Maino and Bern 1984, Chow 1997). The tissue distribution of the receptor in *Fugu* should provide clues about the potential activities of VIP in fish. In fact, *Fugu* VIP receptors have a tissue distribution reminiscent of that observed for VPAC₁R and VPAC₂R in mammals (Hashimoto et al. 1993, Usdin et al. 1994). The *Fugu* VPAC receptors are all present in gonads with varied expression profiles in the other different tissues (VPAC₁RA is additionally present in spleen, brain, kidney and gills; VPAC₁RB in brain; VPAC₂RA in heart and gut and VPAC₂RB present in all tissues tested except gonads). Both VPAC₁R and VPAC₂R were identified in mammal heart and gut (Ishihara et al. 1992, Gagnon et al. 1994, Usdin et al. 1994) but in *Fugu* only VPAC₂R was present. RT-PCR in *Fugu*, using primers specific for each receptor form, demonstrated that there is some overlap in tissue distribution of the duplicated VPAC₁R genes in brain and gonads and of the duplicated VPAC₂R genes in gonads, heart and gut. The only other study in fish is on goldfish VPAC₁R (Chow 1997) but it is unclear if the primers used in the study hybridised to other forms of the receptor, making it difficult to directly compare *Fugu* and goldfish VPAC₁R tissue distribution. Surprisingly although VPAC₁R has been detected in the gut of human, rat and goldfish (Sreedharan et al. 1993, Usdin et al. 1994, Chow 1997) it was absent in *Fugu*. This may merely reflect the position of the sample of *Fugu* gut collected, as VIP receptors have a differential distribution in this tissue (Sayadi et al. 1988, Korman et al. 1989). It should be noted that further overlap of tissue expression may occur between these six receptors, as the RT-PCRs were only carried out on single tissues from juvenile fish and developmental differences may also occur. In view of the fact that *Fugu* tissue availability is limited, expression and functional studies of these receptors is being further characterised in a more amenable teleost, *Sparus aurata* (sea bream).

The PACAP ligand is remarkably well conserved in evolution between protochordates and mammals indicating that it is involved in important biological functions (McRory & Sherwood 1997). It shows a widespread tissue distribution in mammals, notably in the brain and in peripheral organs such as the pancreas, gonads and respiratory and urogenital tracts (Vaudry et al. 2000c). Similarly, the PACAP type 1 receptor shows a broad spectrum of tissue expression, however it is predominantly expressed in the CNS (D’Agata et al. 1996, Basille et al. 2000) where it is thought to act not only as a hypophysiotropic hormone, but also as a neurotransmitter and/or neuromodulator (Hashimoto et al. 1993, Ogi et al. 1993, Peeters et al. 1999). The correlation between ligand and receptor distribution is more difficult to make in the case of PACAP compared with the VIPs, as the PACAP ligand can also interact with the VIP receptors. The distribution in *Fugu* is interesting as the two receptors show quite different distribution patterns, although overall the expression patterns of the two receptors combined in *Fugu* mirror the tissue distribution in mammals. PAC₁RB shows a universal tissue distribution on all tissues tested except for the gonads, which is very similar to the situation found in mammals. Whereas PAC₁RA is much more restricted in that it only shows expression in the brain, gonads and gills, with predominant expression in the former two tissues. Clearly there has been a redefinition of expression patterns between these two receptors, with PAC₁RA developing the more specialised role. The only other study in fish is on goldfish PAC₁R (Wong et al. 1998). RT experiments revealed a
generalised low level of expression throughout all 11 tissues tested, with predominant expression in the heart, brain and pituitary. However, the identities of the RT products were only confirmed by hybridisation and therefore they could have potentially cross-hybridised with both receptor forms, making it difficult to directly compare the Fugu and goldfish PAC₁R tissue distribution. The phylogenetic analysis of the receptors indicates that the goldfish sequence is the orthologue of the Fugu PAC₁R and therefore, based on the results presented here, expression of the goldfish receptor would be expected to have a restricted tissue distribution. Further work will be required on a wider range of teleosts to determine if there is species-specific tissue distribution of the PAC₁Rs, which reflects physiological differences between fish.

The reason why VPAC₁R, VPAC₂R and PAC₁R genes have persisted in duplicate in Fugu and have overlapping expression patterns is still unknown. Under the classical model of evolution, duplicate genes survive by altering function, either by one becoming a pseudogene, by amino acid alterations or complementary degenerate mutation of the promotor regions to produce alternate and maybe very different functions (sub-functionalisation) (Force et al. 1999, Lynch & Conery 2000). Promotor analysis of these genes in Fugu has not yet been carried out, due to lack of 5′ genomic data for some of the genes, but considering the tissue distribution, whilst each gene may perform a similar overall function, they may have alternate distinct tissue-specific functions specific to the fish lineage.

Gene evolution is a complex issue, which in fish is further complicated by the plasticity of the genome and capacity to accommodate ‘extra’ genes. However, comparison of genomic regions in humans, fish and more ancient organisms can help to reconstruct ancestral genomic regions. Although fish diverged from the tetrapod lineage approximately 450 Ma, it would be wrong to assume that their development has been static. Genes in fish, which are also present in human, have been adapted throughout that 450 Ma span to accommodate their different lifestyles in aquatic and land environments. Interestingly, despite extensive database searching, heterologous probing and degenerate PCR experiments, it has not been possible to identify either a secretin receptor or ligand in Fugu and therefore we propose that the duplication that gave rise to secretin and its receptor occurred after the divergence of fish from the tetrapod lineage 450 Ma. The exact nature of the function of the four VIP and two PACAP receptor genes in Fugu described here, and their relationship to the mammalian orthologues remains to be determined, but no doubt will produce interesting results about the adaptation to the aquatic environment and osmoregulation.

Acknowledgements

This work was supported by an MRC grant to MSC and GE and an FCT pluriannual grant to CCMAR, Portugal. DMP was in receipt of a sabbatical grant (BL58, FCT, Portugal). JCRC was funded by FCT, Portugal (grant, BD/19925/99).

References


Cai Y, Xin X, Yamada T, Muramatsu Y, Szpirer C & Matsumoto K 1995 Assignment of the genes for rat pituitary adenylate cyclase activating polypeptide (Adcyap1) and its receptor subtypes (Adcyap1r1, Adcyap1r2, and Adcyap1r3). _Cyto genetics Cell Genetics_ 71 193–196.


Felzenstein J 1985 Confidence limits on phylogenies an approach using the bootstrap _Evolution_ 39 783–791.


Mainoiva JR & Bern HA 1984 Influence of vasoactive intestinal polypeptide urotensin II on the absorption of water and NaCl by the anterior intestine of the tilapia, _Sarotherodon mossambicus_. _Zeological Science_ 1 100–105.


Robinson-Rechavi M, Marchand O, Escriva H & Laudet V 2001c An ancestral whole-genome duplication may not have been responsible for the abundance of duplicated fish genes. *Current Biology* **11** 438–459.


Received 29 June 2004
Accepted 5 July 2004