Dev Genes Evol (2004) 214:277–284
DOI 10.1007/s00427-004-0401-9

ORIGINAL ARTICLE

Carmen Gloria Feijóo · Miguel Manzanares · Elisa de la Calle-Mustienes · José Luis Gómez-Skarmeta · Miguel L. Allende

The *Irx* gene family in zebrafish: genomic structure, evolution and initial characterization of *irx5b*

Received: 13 January 2004 / Accepted: 9 March 2004 / Published online: 3 April 2004

Abstract Genes of the *iroquois* (*Iro/Irx*) family are highly conserved from *Drosophila* to mammals and they have been implicated in a number of developmental processes. In flies, the *Iro* genes participate in patterning events in the early larva and in imaginal disk specification. In vertebrates, the *Irx* genes regulate developmental events during gastrulation, nervous system regionalization, activation of proneural genes and organ patterning. The *Iro* genes in *Drosophila* and the *Irx* genes of mammals show a clustered organization in the genome. Flies have a single cluster comprising three genes while mammals have two clusters also having three genes each. Moreover, experimental evidence in flies shows that transcriptional regulatory elements are shared among genes within the *Iro* cluster, suggesting that the same may be true in vertebrates. To date, the genomic organization of the *Irx* genes in non-mammalian species has not been studied. In this work, we have isolated the *irx5b* gene from zebrafish, *Danio rerio*, and have characterized its expression pattern. Furthermore, we have identified the complete set of *Irx* genes in two fish species, the zebrafish and pufferfish, *Takifugu rubripes*, and have determined the genomic organization of these genes. Our analysis indicates that early in fish evolutionary history, the *Irx* gene clusters have been duplicated and that subsequent events have maintained the clustered organization for some of the genes, while others have been lost. In total there are 11 existing *Irx* genes in zebrafish and 10 in pufferfish. We propose a new nomenclature for the zebrafish *Irx* genes based on the analysis of their sequences and their genomic relationships.

Keywords *Iroquois* · Zebrafish · Development · Gene duplication · Teleost

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00427-004-0401-9

Edited by D. Tautz

C. G. Feijóo · M. L. Allende (✉)
Millennium Nucleus in Developmental Biology
and Departamento de Biología, Facultad de Ciencias,
Universidad de Chile,
Casilla 653, Santiago, Chile
e-mail: allende@uchile.cl
Tel.: +56-2-6787390
Fax: +56-2-2763802

C. G. Feijóo
Facultad de Ciencias de la Salud,
Universidad Andrés Bello,
Santiago, Chile

M. Manzanares
Instituto de Investigaciones Biomédicas CSIC-UAM,
Madrid, Spain

E. de la Calle-Mustienes · J. L. Gómez-Skarmeta
Centro Andaluz de Biología del Desarrollo,
Universidad Pablo de Olavide,
Seville, Spain

Introduction

The *Iroquois* (*Irx/Iro*) genes were discovered in *Drosophila* because of their important role in the formation of sensory bristles in the dorsal mesothorax of the fly (Dambly-Chaudière and Leyns 1992). The *Irx* genes encode a family of highly conserved homeodomain-containing proteins of the TALE super-class (Burglin 1997), which contain a characteristic motif of 13 amino acids, called the IRO box. Analysis of *Irx* function in different model systems has defined a broad role for these genes during development (for reviews see Cavodeassi et al. 2001; Gómez-Skarmeta and Modolell 2002). Thus, during *Drosophila* early larval stages, *Irx* genes are required for the specification of the dorsal regions of the eye, the head and the mesothorax (Cavodeassi et al. 1999; Diez del Corral et al. 1999). Later, they are required for the activation of proneural and provein genes in a subset of proneural clusters and vein territories within the wing imaginal discs (Gómez-Skarmeta et al. 1996, Leyns et al. 1996). *Xenopus* and zebrafish Irx proteins, like their

Drosophila counterparts, also participate in the regulation of proneural genes (Bellefroid et al. 1998; Gómez-Skarmeta et al. 1998; Itoh et al. 2002). In addition, in *Xenopus*, they are essential for the development of the neural plate through repression of *Bmp-4* (Gómez-Skarmeta et al. 2001). This may also be true in other vertebrates since, in zebrafish, Irx proteins downregulate *BMP-4* and the BMP-4 downstream gene *GATA2* (Kudoh and Dawid 2001), while in chicken *Irx2* is expressed in the prospective neural plate in a pattern that is complementary to that of *Bmp-4* (Goriely et al. 1999). In addition, vertebrate *Irx* genes are also required to subdivide the spinal cord and the brain (Briscoe et al. 2000; Kobayashi et al. 2002).

In *Drosophila* the Iro complex is composed of three clustered genes, *araucan*, *caupolican* and *mirror,* spanning 130 kb of genomic DNA (Gómez-Skarmeta et al. 1996; McNeil et al. 1997). In mouse and humans there exist six genes organised in two clusters, IrxA and IrxB (Bosse et al. 2000; Peters et al. 2000). The IrxA cluster contains the *Irx1*, *Irx2* and *Irx4* genes and is located in chromosome 8 in mouse and in chromosome 5 in human. IrxB contains *Irx3*, *Irx5* and *Irx6* and is located in chromosome 13 in mouse and in chromosome 16 in human. These two clusters most probably originated from the duplication of an ancestral cluster composed of three genes, as sequence comparisons have demonstrated that each gene in cluster A has a paralog in cluster B (*Irx1-Irx3*; *Irx2-Irx5*; *Irx4-Irx6*). Nonetheless, relationships between the vertebrate Irx clusters and the *Drosophila* Iro cluster are unclear, suggesting independent origins (Bosse et al. 2000; Peters et al. 2000). In non-mammalian vertebrates, our knowledge of the structure of this gene family is incomplete. In *Xenopus*, five *Irx* genes have been cloned, *Xiro1, Xiro2, Xiro3, Irx4* and *Irx5* (Bellefroid et al. 1998; Gómez-Skarmeta et al. 1998; Garriock et al. 2001) and in chicken four *Irx* genes have been described, *Irx1*, *Irx2*, *Irx3*, and *Irx4* (Goriely et al. 1999; Funayama et al. 1999; Bao et al. 1999). In zebrafish, four *Irx* genes (previously named *ziro*) have been identified to date: *ziro1* in chromosome 19, *ziro3* and *ziro5* in chromosome 7 and *ziro7* in chromosome 23 (Itoh et al. 2002; Lecaudey et al. 2001; Tan et al. 1999; Wang et al. 2001).

Given the fact that there are six *Irx* genes in mammals, it is expected that a similar complement must exist in other vertebrates. In teleost fish an even larger number could exist since, subsequent to their evolutionary divergence from other vertebrates, this lineage has undergone a genome duplication event, possibly involving complete tetraploidization (Taylor et al. 2001, 2003). Duplicated developmental genes in fish are often retained if each new paralog acquires a subset of the functions of the ancestral gene (Force et al. 1999; Prince and Pickett 2002). To explore this possibility among the *Irx* genes, we decided to search for additional members of this family in the zebrafish (*Danio rerio*). By screening a genomic library using a *Drosophila caupolican* homeobox probe, we were able to isolate a new *Irx* gene and characterize its expression pattern. In an effort to correctly classify this gene and to ascertain whether the zebrafish genome contains duplicated paralogs of the mammalian *Irx* genes, we searched the available zebrafish genomic sequence for other members of this gene family. In total, we have detected 11 zebrafish *Irx* gene sequences which, by radiation hybrid mapping, are found distributed among six linkage groups. A similar analysis in the *Takifugu rubripes* (pufferfish) genome identified 10 genes. Some of the zebrafish genes have maintained the clustered structure found in other vertebrates while others are found in isolation. We propose a new nomenclature, consistent with that used in mammals, for classification of these genes based on their sequence relationships. Thus, the newly isolated gene we have cloned and characterized has been named *irx5b*.

Materials and methods

Zebrafish embryos and whole-mount in situ hybridization

Wild-type fish were raised in our facility according to standard protocols (Westerfield 1995). Embryos were raised at 28°C in E3 medium and, when necessary, were anesthetized and fixed in 4% paraformaldehyde overnight at 4°C. Developmental timepoints are expressed as hours post-fertilization (hpf) or days post-fertilization (dpf).

Antisense RNA probes for zebrafish *irx5b* were synthesized from cDNA using digoxigenin (Roche) as a label. In situ hybridization was done as previously described (Jowett and Lettice 1994).

Cloning of *irx5b*

A 195-bp probe from the *Drosophila caupolican* gene (bases 1355–1550 of the *caupolican* cDNA), which encodes the homeodomain region, was used to screen a zebrafish genomic DNA library (gift of K. Zinn) at low stringency. Positive clones were purified and partially sequenced to determine their identities. A 1.1-kb *Sac*I/*Hind*III genomic fragment from one such positive clone encoded the homeodomain region of *irx5b*. This sequence was used to identify a partial cDNA of 800 bp in the database (clone 012-H09-2 from the PJR-Z1+Z2 *Danio rerio* library, kindly donated by Dr. Peng Jinrong) that, after sequencing, encoded the entire carboxy-terminus of the Irx5b protein. To obtain the *irx5b* 5′ cDNA fragment, 5′ RACE was performed using the SMART RACE cDNA Amplification Kit (Clontech); template cDNA was obtained by preparing RNA from 18-hpf embryos. Nested PCR was carried out using two reverse primers (R1 and R2) and the forward primers supplied by the manufacturer. The sequences of the primers (corresponding to the homeodomain of *irx5b*) are: R1: 5′ ctcgctccgcgtcctgggcgtcc 3′ and R2: 5′ tccccgaacgggtacgggcccag 3′. The irx5b cDNA sequence has been deposited in the GenBank database with accession number AY512587.

RT-PCR of *irx5b*

Analysis of the expression of *irx5b* by RT-PCR was performed using RNA from 0, 10, 18, 24 and 48 hpf and 5 dpf. The RT-PCR was carried out using primers F1 (5′ ggggctccccatatgacccttctcca 3′) and R3 (5′ gatttctgccagagaccagag 3′), which amplify a 200-bp fragment of the *irx5b* cDNA.

Mapping of the newly identified *Irx* zebrafish genes

Genomic mapping was performed for all of the new *Irx* genes identified in the database search by using a zebrafish-hamster so-

Table 1 Nomenclature and genomic analysis of the fish *Irx* genes. We have grouped ten of the existing zebrafish *Irx* genes into four clusters according to linkage analysis and degree of sequence divergence; a single gene (*irx7*) is not assigned to any of these groups. Clone numbers (contigs) for *Danio rerio* and *Takifugu rubripes* scaffolds are as indicated in the Ensembl database (http://www.ensembl.org). Shown are the PCR primers used for radiation hybrid mapping (Kwok et al. 1998), the *D. rerio* chromosomes to which each gene was mapped and the nearest genetic marker found in each case with its corresponding LOD score

Cluster	New name	Previous name	*D. rerio* clone	Mapping primers	*D. rerio* chromosome	Nearest marker/ LOD score	*T. rubripes* scaffold
Aa	Irx1a	ziro1b[a]	ctg9347	F 5′ cctcgccaaactgcttggtgac 3′ R 5′ aaggctcgcccatgacatac 3′	16	Z1837/14.484	486
	Irx2a		NA11470	F 5′ cccaaccgctgtgtgtgtttata 3′ R 5′ ctccccgtccgactccaccga 3′	16	Z1642/17.326	486
	Irx4a		ctg9488	F 5′ tcctcaggatcacgtgtatgc 3′ R 5′ cgagctctgcagacgaaggac 3′	9	fa11f12.s1/7.636	1930
Ba	Irx3a	ziro3[b]	ctg30050		7		16
	Irx5a	ziro5[c]	ctg9730		7		29
	Irx6a		ctg9730	F 5′ gcaggacctgcagttcagtgac 3′ R 5′ caggcccagtccgtacatccc 3′	7	unp2038/15.223	29
Ab	Irx1b	ziro1[c]	ctg10320		19		1
	Irx4b		ctg10320	F 5′ atgcactgtcactatagcaatggc 3′ R 5′ agtcgtagggataatatgcagagc 3′	19	Z11159 /15.140	1
Bb	Irx3b		ctg30196	F 5′ tcctcaggatcacgtgtatgc 3′ R 5′ cgagctctgcagacgaaggac 3′	25	Z1213/15.357	Not present
	Irx5b		NA2600	F 5′ ggggtgatgaagatgatgagccc 3′ R 3′ cgccagtcccgataatgattagg 3′	25	z9044/14.110	433
	Irx7	ziro7[d]	Na13724		23		494

[a] Joseph EM and Fishman MC, unpublished. Cardiac and neural expression of *Irx1-b* in zebrafish embryos. GenBank accession number AY043241
[b] Tan et al. 1999
[c] Wang et al. 2001
[d] Lecaudey et al. 2001

matic cell radiation hybrid (RH) panel (Kwok et al. 1998). Specific primers for each gene were used and the RH panel was screened by PCR. Primer sequences and mapping results are detailed in Table 1.

Genomic sequence analysis

All gene structure predictions and scaffold locations were performed using sequence data from Ensembl release v16 (http://www.ensembl.org). Results have been checked against the latest release (v19) with no significant changes to what is reported herein. Similarity searches were performed using the BLAST server at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) or the BLASTView utility of the Ensembl server. A number of the zebrafish protein predictions were completed by searches in EST and HTGS databases at the NCBI, followed by translation and comparison to other mammalian and *T. rubripes* predicted proteins. Multiple alignments were created using Clustal X (Thompson et al. 1997), corrected by visual inspection and edited using the BioEdit tool (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). This alignment is available as Electronic Supplementary Material. The aligned regions of the Irx proteins used to construct the unrooted neighbour-joining bootstrapped tree (Saitou and Nei 1987; Felsenstein 1996), shown in Fig. 3 and drawn with Treeview (Page 1996), are also supplied as Electronic Supplementary Material. Parsimony (Felsenstein 1993) and maximum likelihood analysis (Adachi and Hasegawa 1996) provided identical tree topology and relationships between all of the Irx proteins.

Results

We screened a zebrafish genomic library at low stringency using a *Drosophila caupolican* homeobox probe. Among the positive clones that were recovered, we determined that three classes were present: two of these corresponded to the known genes *ziro1* (hereafter named *irx1b*) and *ziro3* (*irx3a*) while the third was a novel sequence belonging to the *Irx2/Irx5* paralog group. This gene was eventually named *irx5b* after extensive genome-wide analysis of the zebrafish *Irx* gene family (see later). The cloned *irx5b* genomic sequence was used to design oligonucleotide primers for cloning a partial cDNA. 5′ RACE experiments yielded a 900-bp cDNA that was sequenced to confirm its identity and which was subsequently used for in situ hybridization experiments. In addition, in the database, we identified an 800-bp cDNA fragment (clone 012-H09-2) which encoded the entire carboxy-terminus of the Irx5b protein. The full-length *irx5b* sequence has been deposited in the GenBank database with accession number AY512587.

We analyzed the expression of the *irx5b* gene. RT-PCR experiments (Fig. 1) show that *irx5b* mRNA is detected beginning at about 18 hpf, becoming stronger at 24 and 48 hpf. However, expression is not detected at 5 dpf suggesting transient embryonic expression of this gene.

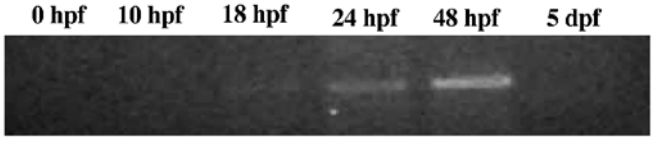

Fig. 1 Temporal expression of *irx5b*. RT-PCR was carried out using primers specific for the *irx5b* gene and using template RNA prepared from the indicated stages; timepoints are expressed in hours post-fertilization (*hpf*) or days post-fertilization (*dpf*). The specific amplified band (200 bp) can be seen at 18, 24 and 48 hpf

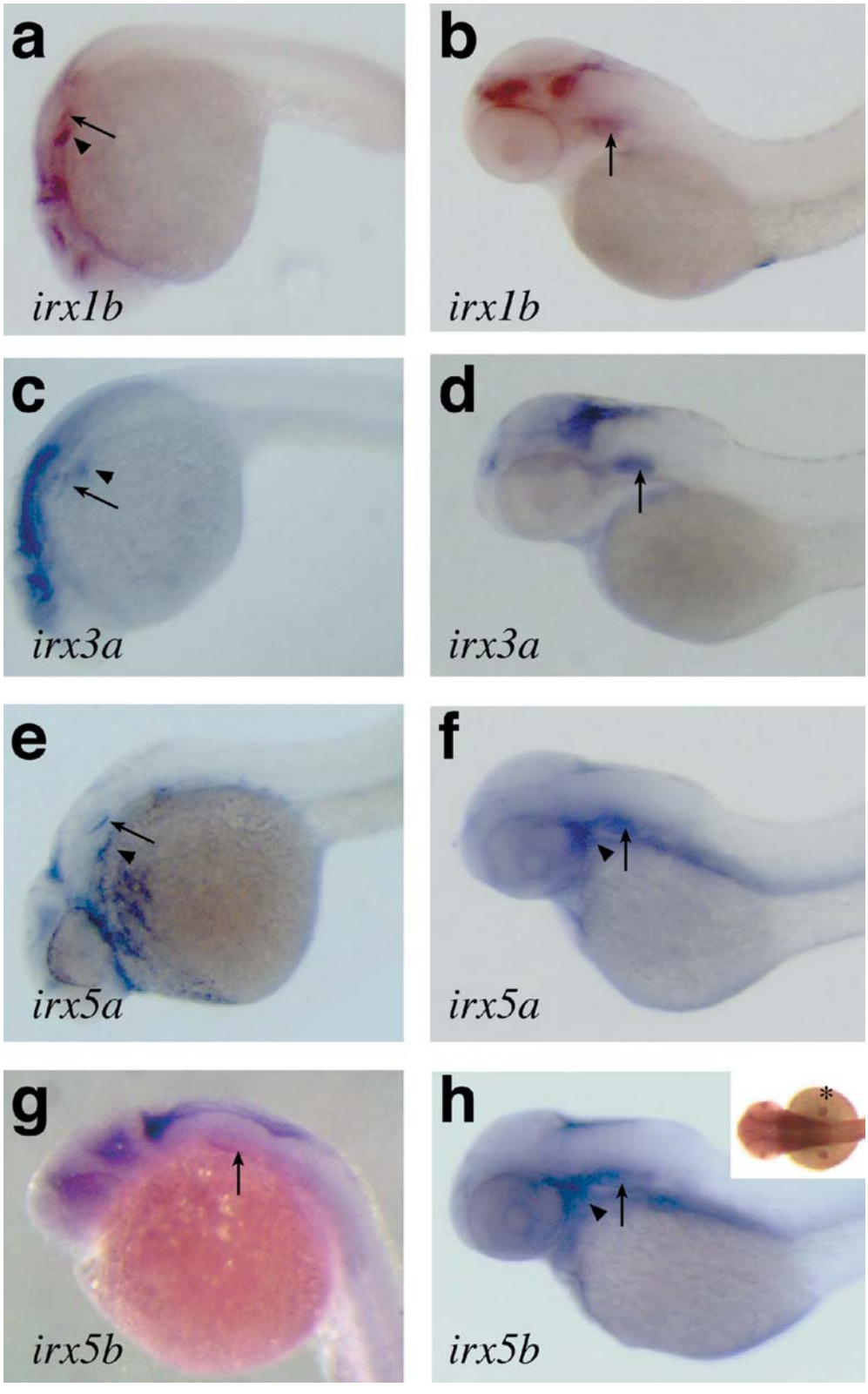

Fig. 2a–h Spatial expression of zebrafish *Irx* genes. Detection of *Irx* RNA was carried out by whole-mount in situ hybridization using 24-hpf embryos (**a**, **c**, **e**, **g**) or 48 hpf embryos (**b**, **d**, **f**, **h**). All images are lateral views in which anterior is to the *left* and dorsal is *up* except for the *inset* in **h** which is a dorsal view. Probes used were specific for *irx1b* (**a**, **b**), *irx3a* (**c**, **d**), *irx5a* (**e**, **f**) and of *irx5b* (**g**, **h**). *Arrows* indicate expression in the developing otic vesicle while *arrowheads* denote expression in the area of the cranial ganglia. *Asterisk* in **h**, *inset*, labels the developing fin bud

To detect mRNA in specific embryonic tissues we used the *irx5b* cDNA to generate an antisense riboprobe and we performed whole-mount in situ hybridization. We could not detect expression at stages before 18 hpf. At 24 hpf, strong expression is seen in the ventral diencephalon, midbrain, cerebellum, hindbrain and in the otic vesicle (Fig. 2g). At 48 hpf, expression becomes diminished in the brain but can be detected strongly in fin buds (Fig. 2h, inset). Expression is restricted to the distal portion of the developing fin bud at this stage. Expression

in the otic vesicle has not been described previously for *Irx* genes in the zebrafish though *irx1b* is strongly expressed in the cranial placodes—including the otic placode—during late gastrulation (C.G. Feijóo and M.L. Allende, unpublished results). We examined the expression of the other known *Irx* genes in the otic placode and we detected the presence of *irx1b*, *irx3a* and *irx5a* transcripts in the ventral domain of this structure at 24 hpf (Fig. 2a, c, e, arrows). In addition, we detect expression of these genes in tissues adjacent to the otic placode, which, by their position, appear to be where cranial ganglia develop. In particular, *irx1b* and *irx5a* are expressed anterior to the ear (Fig. 2a, e, arrowheads) while *irx3a* is expressed slightly posterior (Fig. 2c, arrowheads). These areas most likely correspond to the otic and lateral line ganglia. At 48 hpf, expression of all *Irx* genes analyzed becomes strongly reduced. Expression can still be detected in the otic vesicle at this stage, though it is now apparent throughout this structure (Fig. 2b, d, f, h, arrows). Expression in the trigeminal ganglia can be seen for *irx5a* and *irx5b* which have almost identical expression patterns at this stage (Fig. 2f, h, arrowheads).

To correctly classify and name the gene we have described here, we performed a genome-wide search for all *Irx* sequences present in zebrafish. Searches using as a query sequence any of the *Irx* genes described in *Xenopus* or *Danio*, detected 11 different genomic sequences from which at least the homeobox region and conserved IRO box could be unambiguously found. Four of the recovered sequences correspond to the previously described *Iroquois* genes: *irx1b* (*ziro1*), *irx3a* (*ziro3*), *irx5a* (*ziro5*) and *irx7* (*ziro7*; Lecaudey et al. 2001; Tan et al. 1999; Wang et al. 2001). For all of the newly identified sequences, specific primers were designed that amplify small (150–300 bp) DNA fragments that were used for radiation hybrid mapping (Kwok et al. 1998). The mapping results were unambiguous (LOD scores above 7 in all cases), allowing us to place the *Irx* genes on the zebrafish genetic map (Table 1). Some pairs of genes mapped to identical map locations and, in two cases, *Irx* genes were found on the same contig as published in the finished genomic sequence.

Given the large number of *Irx* sequences found in zebrafish compared to mammals, we performed pair-wise sequence comparisons to identify orthology and paralogy relationships. We used the ClustalX program to perform the alignments and phylogenetic trees were derived from this data (Fig. 3). In all cases but one (*irx7*), we were able to clearly assign the zebrafish *Irx* genes to orthology groups corresponding to the six mammalian *Irx* genes. For each of four mammalian *Irx* genes we find two orthologs in zebrafish; the remaining two have a single ortholog in the fish genome. This result can be expected from the mounting evidence that indicates that a total or partial genome duplication event occurred in the teleost lineage and that many of the duplicated genes that were generated have persisted in the fish genome. Considering this evidence, and in keeping with the mammalian *Irx* gene cluster names as A (*Irx1*, *Irx2*, *Irx4*) and B (*Irx3*, *Irx5*, *Irx6*), we have classified the ten clearly orthologous zebrafish genes as belonging to four groups: Aa, Ab, Ba and Bb (Table 1; Fig. 4a, b). To correctly name individual genes, we assigned the gene of the pair of duplicates that is phylogenetically closest to the mammalian ortholog as belonging to the "a" group (Aa and Ba). The more divergent genes were classified as belonging to the "b" group (Ab and Bb). Moreover, this grouping of *Irx* genes by sequence is fully corroborated by the genetic linkage between paralogs in each of the four groups that we report here.

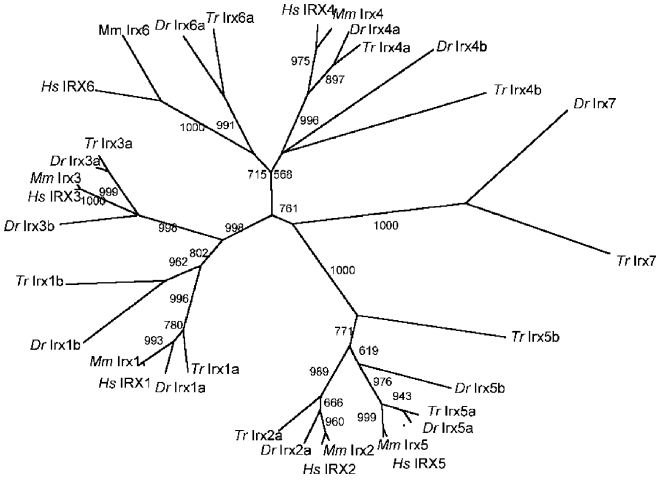

Fig. 3 Unrooted neighbour-joining phylogenetic tree of human (*Hs*), mouse (*Mm*), fugu (*Tr*) and zebrafish (*Dr*) Irx predicted peptides. Those fish protein sequences that group more closely with mammalian proteins were named *a*, and those more divergent *b*. Fish Irx7 proteins are highly diverged and cannot be related by sequence to any of the mammalian classes. Bootstrap values higher than 50% are shown

During evolution of the teleosts, genes of the *Irx* Aa cluster have been split into two chromosomal locations: *irx1a* and *irx2a* are found in LG16, while *irx4a* is found by itself in LG9. Cluster Ab is found in LG19 though only the linked *irx1b* and *irx4b* genes are present; we could not find a second *irx2* sequence either between *irx1b* and *irx4b* or elsewhere in the genome. Cluster Ba appears to be the only intact cluster that resembles the mammalian three-gene complex, with *irx3a* (previously *ziro3*), *irx5a* (*ziro5*) and *irx6a* in LG7. Cluster Bb has only two genes, *irx3b* and *irx5b* on LG25, as we could not find another *Irx6*-type sequence. Finally, *irx7* (previously *ziro7*) is found alone on LG23 (Itoh et al. 2002) and could not be unambiguously assigned to any of the orthology groups either by maximum likelihood or parsimony analysis.

To provide additional data to support our findings in the zebrafish, we performed a similar search for *Irx* genes in the *Takifugu rubripes* (pufferfish) genome sequence. We found ten pufferfish *Irx* genes all of which are clearly orthologous to the zebrafish *Irx* genes (Fig. 4c). Genomic linkage could be established between the *irx1a* and *irx2a*, *irx1b* and *irx4b* and *irx5a* and *irx6a* genes, as they are found on single contigs in the finished sequence. We could not find an ortholog of the zebrafish *irx3b* gene and,

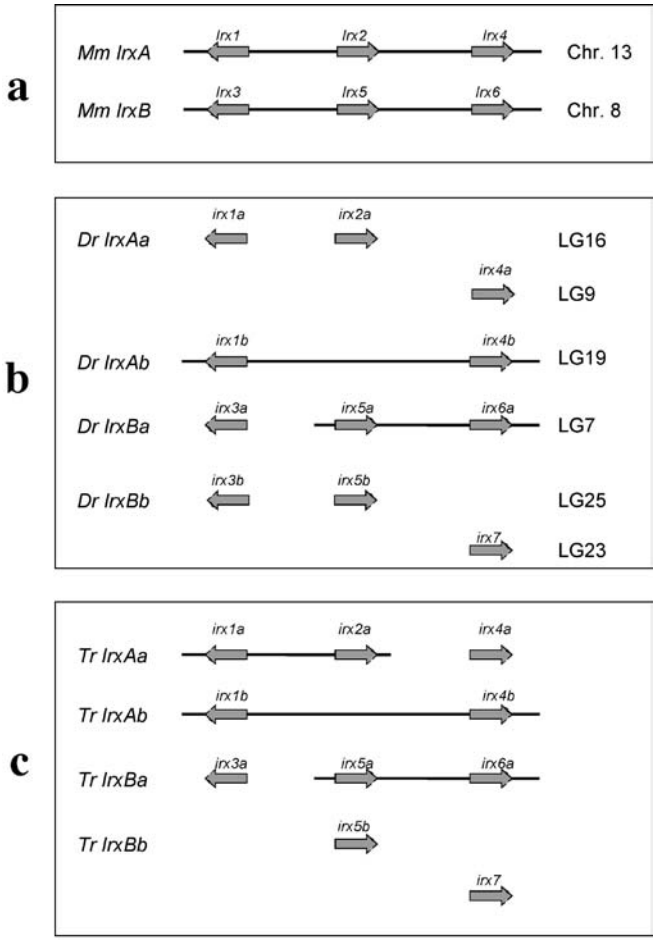

Fig. 4 *Irx* gene organization in vertebrates. *Horizontal lines* denote those genes that lie in the same contig in the zebrafish and fugu genome assemblies. (a), mouse, *Mm*, (b), zebrafish *Dr*, (c), pufferfish *Tr*

as in zebrafish, there are single orthologs of the mammalian *Irx2* and *Irx6* genes; in consequence, the pufferfish *irx5b* gene is unlinked to other *Irx* genes. An *irx7* gene is also found in pufferfish and, together with its zebrafish ortholog, appears as an outlier in the phylogenetic tree that represents the vertebrate *Irx* gene family (Fig. 3).

Discussion

In this work, we have isolated a novel member of the *Irx* gene family and we have determined the entire repertoire of *Irx* genes present in the zebrafish. This family of genes has been shown to be important in the development of a number of tissues, both in invertebrates and vertebrates. Several human diseases map to genomic loci corresponding to *Irx* genes, though precise assignment of the genes responsible for these syndromes remains to be elucidated. Until now, the genomic structure of the vertebrate *Irx* complexes had been determined solely in mammals. In this animal group, two clusters of three genes each are present. The sequence relationships between the six genes suggest duplication of an original cluster of three genes during evolution (Peters et al. 2000), an event most likely simultaneous with the genome-wide duplication that occurred early in vertebrate history (Aparicio 2000). Therefore, it was of interest to determine if this structure is conserved in other vertebrates, allowing for a more precise reconstruction of the evolutionary history of this gene family. We are also interested in providing evidence for a possible functional significance of the conserved clustered structure of the *Irx* genes.

We find that the zebrafish genome contains 11 *Irx* genes while the *T. rubripes* genome has 10. The high number of genes compared to mammals is expected, given the high incidence of duplicated genes in teleosts. It is widely believed that early in the evolution of this group, a partial or whole genome duplication event occurred (Taylor et al. 2001). The evidence we have gathered from our search indicates that the last common ancestor to fish and mammals had two clusters of three genes each (A and B) and that each of these clusters was duplicated in the early fish lineage to produce four clusters (Aa, Ab, Ba and Bb) most likely totalling 12 genes. When the zebrafish and mammalian *Irx* genes are compared, we can assign duplicated genes for four of the six members of this family. We could not find duplicated paralogs of *Irx2* or *Irx6* in *D. rerio*. A parallel analysis in *T. rubripes* shows a similar structure, though, in this species, a third gene has been lost secondarily to the duplication event as there is only a single *Irx3* gene.

The *irx7* gene is clearly an outlier and to date it has only been found in fish. It is not possible to discriminate by sequence comparisons from which of the original *Irx* genes it could have arisen. Since both pufferfish and zebrafish have single copies of *Irx2* and *Irx6*, it is possible that *irx7* is a highly divergent paralog of one of these genes. It is also possible that it arose in an ancient (prevertebrate) duplication event rendering it very divergent compared to the remaining clustered *Irx* genes; in this case it must have been lost in mammals.

Interestingly, cluster coherence is not strongly maintained in fish; only the IrxBa complex retains three genes. However, at least one of the *Irx1*/*Irx2* and *Irx3*/*Irx5* pairs of paralogs maintain their close linkage in the genome suggesting that functional constraints (perhaps shared regulatory elements) keep these genes from becoming segregated.

We isolated a zebrafish cDNA sequence corresponding to a novel member of the *Iroquois* gene family, *irx5b*. Expression of *irx5b* shows some differences from that described for other *Irx* genes in fish (Wang et al. 2001; Tan et al. 1999). For example, *irx5b* begins to be expressed late in development, at 18 hpf, and is no longer detected by 5 dpf. In contrast, *irx1b*, *irx3a* and *irx5a* begin transcription during mid to late gastrulation (Wang et al. 2001; Kudoh and Dawid 2001). There appears to be no maternally inherited transcripts for any of the described *Irx* genes in fish. Spatial distribution of *Irx* gene products has been described mainly in pharyngula-stage embryos.

At 24 hpf, *irx5b* is detected in the ventral diencephalon similar to the expression of *irx3a* (Tan et al. 1999) while *irx1b* is found in the dorsal diencephalon (Wang et al. 2001). These three genes are found in the midbrain, cerebellum and hindbrain, in similar widespread patterns. We detected a patch of *irx5b* label in the ventral otic vesicle at 24 hpf, a tissue where expression of other *Irx* genes had not been described. However, upon further examination of the previously isolated *Irx* genes from fish, we have detected a similar pattern for *irx1b*, *irx3a* and *irx5a* at 24–30 hpf. Among the four *Irx* genes analyzed here, the only one of this group found to be expressed in fin buds is *irx5b*. This expression is interesting given the fact that a deletion of the IrxB cluster in the mouse (the *Fused toes* mutant) produces, among multiple severe defects, a limb phenotype (Peters et al. 2000).

In all vertebrate species analyzed, the expression patterns of the *Irx3/Irx5* gene pairs are largely equivalent, coinciding with their conserved genomic linkage. It will be interesting to learn whether these facts reflect the presence of shared regulatory elements among these genes as has been the case for other clustered genes such as those of the Hox complex (Duboule 1998).

Note

Dildrop and Rüther (2004) have independently carried out a similar analysis of the fish *Irx* genes. Their conclusions are identical to ours with the exception of the genomic location of the *irx4a* gene, which they place on LG16 together with *irx1a* and *irx2a*.

Acknowledgements We thank Renate Dildrop and Ulrich Rüther for providing a copy of their manuscript prior to publication. X. Wang kindly provided probes. We are also grateful to Peng Jinrong for clone 012-H09-2. We wish to thank Gabino Sánchez-Pérez for help with phylogenetic analysis. Carolina Achondo and Florencio Espinoza provided technical help. C.G.F. and M.A. were supported by grants from Fondecyt (1031003), the ICM (P02-050) and by a grant from the UNAB (DI 10-03). M.M. and J.L.G.-S. were supported by grants from the Spanish Ministry of Science and Technology (BMC2002-03558 and BMC2001-2122, respectively). In addition, M.A. and J.L.G.-S. were supported by a CONICYT-CSIC collaboration grant (2003CL00021).

References

Adachi J, Hasegawa M (1996) MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. Comput Sci Monogr Inst Stat Math 28:1–50

Aparicio S (2000) Vertebrate evolution: recent perspectives from fish. Trends Genet 16:54–56

Bao ZZ, Bruneau BG, Seidman JG, Seidman CE, Cepko CL (1999) Regulation of chamber-specific gene expression in the developing heart by *Irx4*. Science 283:1161–1164

Bellefroid EJ, Kobbe A, Gruss P, Pieler T, Gurdon JB, Papalopulu N (1998) *Xiro3* encodes a *Xenopus* homolog of the *Drosophila* Iroquois genes and functions in neural specification. EMBO J 17:191–203

Bosse A, Stoykova A, Nieselt-Struwe K, Chowdhury K, Copeland N, Jenkins NA, Gruss P (2000) Identification of a novel mouse *Iroquois* homeobox gene, *Irx5*, chromosomal localization of all members of the mouse *Iroquois* gene family. Dev Dyn 218:160–174

Briscoe J, Pierani A, Jessell TM, Ericson J (2000) A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell 101:435–445

Burglin TR (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Res 25:4173–4180

Cavodeassi F, Diez del Corral R, Campuzano S, Dominguez M (1999) Compartments and organizing boundaries in the *Drosophila* eye: the role of the homeodomain Iroquois proteins. Development 126:4833–4942

Cavodeassi F, Modolell J, Gómez-Skarmeta JL (2001) The *Iroquois* family of genes: from body building to neural patterning. Development 128:2847–2855

Dambly-Chaudière C, Leyns L (1992) The determination of sense organs in *Drosophila*: a search for interacting genes. Int J Dev Biol 36:85–91

Diez del Corral R, Aroca P, Gómez-Skarmeta JL, Cadoveassi F, Modolell J (1999) The iroquois homeodomain proteins are required to specify body wall identify in *Drosophila*. Genes Dev 13:1754–1761

Dildrop R, Rüther U (2004) Organization of *iroquois* genes in fish. Dev Genes Evol DOI 10.1007/s00427-004-0402-8

Duboule D (1998) Vertebrate Hox gene regulation: clustering and/or colinearity? Curr Opin Genet Dev 8:514–518

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle

Felsenstein J (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. Methods Enzymol 266:418–427

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545

Funayama N, Sato Y, Matsumoto K, Ogura T, Takahashi Y (1999) Coelom formation: binary decision of the lateral plate mesoderm is controlled by the ectoderm. Development 126:4129–4138

Garriock RJ, Vokes SA, Small EM, Larson R, Krieg PA (2001) Developmental expression of the *Xenopus* Iroquois-family homeobox genes, *Irx4* and *Irx5*. Dev Genes Evol 211:257–260

Gómez-Skarmeta JL, Modolell J (2002) Iroquois genes: genomic organization and function in vertebrate neural development. Curr Opin Genet Dev 12:403–408

Gómez-Skarmeta JL, Diez del Corral R, de la Calle-Mustienes E, Ferrés-Marcó D, Modolell J (1996) *araucan* and *caupolican*, two members of the novel Iroquois complex, encode homeoproteins that control proneural and vein forming genes. Cell 85:95–105

Gómez-Skarmeta JL, Glavic A, de la Calle-Mustienes E, Modollel J, Mayor R (1998) *Xiro*, a *Xenopus* homolog of the *Drosophila* Iroquois complex genes, controls development al the neural plate. EMBO J 17:181–190

Gómez-Skarmeta JL, de la Calle-Mustienes E, Modollel J (2001) The Wnt-activated *Xiro1* gene encodes a repressor that is essential for neural development and down regulates BMP4. Development 128:551–560

Goriely A, Diez del Corral R, Storey KG (1999) *c-Irx2* expression reveals an early subdivision of the neural plate in the chick embryo. Mech Dev 87:203–206

Houweling AC, Dildrop R, Peters T, Mummenhoff J, Moorman AFM, Rüther U, Christoffels VM (2001) Gene and cluster-specific expression of the Iroquois family members during mouse development. Mech Dev 107:169–174

Itoh M, Kudoh T, Dedekian M, Kim CH, Chitnis AB (2002) A role for *iro1* and *iro7* in the establishment of an anteroposterior compartment of the ectoderm adjacent to the midbrain-hindbrain boundary. Development 129:2317–2327

Jowett T, Lettice L (1994) Whole-mount in situ hybridizations on zebrafish embryos using a mixture of digoxigenin-and fluorescein-labeled probes. Trends Genet 10:73–74

Kobayashi D, Kobayashi M, Matsumoto K, Ogura T, Nakafuku M, Shimamura K (2002) Early subdivisions in the neural plate define distinct competence for inductive signals. Development 129:83–93

Kudoh T, Dawid IB (2001) Role of the *Iroquois3* homeobox gene in organizer formation. Proc Natl Acad Sci USA 98:7852–7857

Kwok C, Korn R, Davis M, Burt D, Paw B, Zon L, Goodfellow P, Schmitt K (1998) Characterization of whole genome radiation hybrid mapping resources for non-mammalian vertebrates. Nucleic Acids Res 26:3562–3566

Lecaudey V, Thisse C, Thisse B, Schneider-Maunoury S (2001) Sequence and expression pattern of *ziro7*, a novel, divergent zebrafish iroquois homeobox gene. Mech Dev 109:383–388

Leyns L, Gómez-Skarmeta JL, Dambly-Chaudière C (1996) *iroquois*: a prepattern gene that controls the formation of bristles on the thorax of *Drosophila*. Mech Dev 59:63–72

McNeil H, Yang CH, Brosky M, Ungos J, Simon MA (1997) mirror encodes a novel PBX-class homeoprotein that functions in the definition of dorso-ventral border of *Drosophila* eye. Genes Dev 11:1073–1082

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Peters T, Dildrop R, Ausmeier K, Ruther U (2000) Organization of the mouse Iroquois homeobox genes in two clusters suggest a conserved regulation and function in vertebrate development. Genome Res 10:1453–1462

Prince VE, Pickett FB (2002) Splitting pairs: the diverging fates of duplicated genes. Nat Rev Genet 3:827–837

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Tan JT, Korzh V, Gong Z (1999) Expression of a zebrafish iroquois homeobox gene, *Ziro3*, in the midline axial structures and central nervous system. Mech Dev 87:165–168

Taylor JS, Van de Peer Y, Braasch I, Meyer A (2001) Comparative genomics provides evidence for an ancient genome duplication event in fish. Philos Trans R Soc Lond B Biol Sci 356:1661–1679

Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y (2003) Genome duplication, a trait shared by 22,000 species of ray-finned fish. Genome Res 13:382–390

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by its quality analysis tools. Nucleic Acids Res 25:4876–4882

Wang X, Emelyanov A, Sleptsova-Friedrich I, Korzh V, Gong Z (2001) Expression of two novel zebrafish Iroquois homologues (*ziro1* and *ziro5*) during early development of axial structures and central nervous system. Mech Dev 105:191–195

Westerfield M (1995) The zebrafish book. University of Oregon Press, Eugene, Ore.