Identification of the actc1c cardiac actin gene in zebrafish

Matiyo Ojehomon a, b, Sarah L. Alderman c, Love Sandhu a, b, Sierra Sutcliffe a, Terry Van Raay a, Todd E. Gillis b, c, John F. Dawson a, b, * 

a Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada 
b The Centre for Cardiovascular Investigations, University of Guelph, Guelph, Ontario, N1G 2W1, Canada 
c Department of Integrative Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada

A R T I C L E   I N F O

Article history: 
Received 31 January 2018 
Received in revised form 12 May 2018 
Accepted 13 June 2018 
Available online 18 June 2018

Keywords: 
Actin protein 
Sequence conservation 
Actin orthologs 
In situ hybridization 
Zebrafish

A B S T R A C T

Zebrafish is rapidly becoming a key model organism for studying a variety of biological processes from molecules to organisms. Interactions involving actin, a contractile protein and part of the cytoskeleton, are regulated by actin binding proteins in the majority of physiological processes in eukaryotic cells. To understand the contribution of actin proteins to the physiological processes of zebrafish, it is important to identify the diverse isoforms of actin encoded by its genome; however, significant sequence identity complicates isoform assignments. Through a combination of human-directed sequence and functional analysis, we have assigned and performed localization of actc1c, a previously undesigned cardiac actin gene, and propose an updated assignment of α-actin protein isoform identities in zebrafish.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

The use of zebrafish (Danio rerio) as a developmental model organism for human disease has emerged as a powerful tool. Research regarding the dysregulation or mutation of human genes requires a strong assignment of the corresponding zebrafish gene to translate results between species. An example is our research of the role of cardiac actin in the development of human cardiovascular disease. Zebrafish are an excellent model for cardiovascular work because zebrafish embryos, unlike the mouse, do not require a fully functional cardiovascular system to survive as they can receive oxygen through diffusion from the surrounding environment (Bakkers, 2011). This initial reduced reliance on a functional heart allows us to assess the development of cardiovascular defects, such as a non-beating heart. Finally, the optical transparency of zebrafish embryos permits non-invasive visualization of the developing cardiovascular system using a simple light microscope (Bakkers, 2011).

Our goal is to develop zebrafish as a model to study alterations in cardiac actin proteins; however, the high sequence identity among actin isoforms creates many challenges. In humans, there are six primary actin isoforms (Bertola et al., 2008): α-skeletal (ACTA1) and α-cardiac actin (ACTC1) classified as striated muscle isoforms; α-smooth (ACTA2), and gamma smooth muscle (ACTG2) classified as smooth muscle isoforms and lastly beta cytoplasmic (ACTB) and gamma cytoplasmic (ACTG1) classified as cytoplasmic isoforms (Bertola et al., 2008; Müller et al., 2012). A major distinguishing difference among human isoforms exists at the N-terminus with the β-, γ-, and α-isoforms possessing three, three and four acidic residues, respectively. Importantly, these differences are exploited by commercially-available isoform-specific antibodies. Studies have shown that the number of acidic residues at the N-terminus of actin is correlated with function; for instance, the binding activity of myosin subfragment-1 is increased by the addition of acidic residues (reviewed in (Doyle and Reisler, 2002)). At the same time, the sequence of α-isoforms N-termini in humans possesses a few differences: a conserved threonine to serine change distinguishes the α-smooth isoform, and a variance in the sequence of the acidic amino acids between α-skeletal and α-cardiac isoforms (Table 1).

In our work described below, we found the zebrafish gene associated with zgc:86709 was not given a gene name; therefore, we have assigned the name actc1c to this gene. Taking the whole set of genomes available into consideration, we searched for the sequence of the actc1c gene using Ensembl to view homologs in a

* Corresponding author. Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada. 
E-mail address: jdawso01@uoguelph.ca (J.F. Dawson).
gene tree generated by the Gene Orthology/Paralogy prediction method pipeline. We were able to view the evolutionary history of the zgc:86709 gene. According to these results, orthologs exist between two species: the cave fish and the spotted gar. A BlastP (NCBI Blast) amino acid database against protein database search of the cave fish zgc:86709 protein sequence compared to zebrasfish revealed 100% protein sequence conservation. The remaining z-isoforms expressed significant percentage identity (>98%). Specifically, actc1a, acta1a, acta1b, and acta2 had percentage identities of 99.73%, 99.20%, 99.73%, 98.67%, 98.94%, respectively. A BlastP search was also performed between the spotted gar and zebrasfish. Similar to the cave fish, the percentage identity results were also significant (>98%) between the spotted gar and zebrasfish. Specifically, actc1a, acta1b, acta1a, actc1b, and acta2 had percentage identities of 99.47%, 99.47%, 99.47%, 98.94%, 99.20%, respectively.

The partially duplicated genome in zebrasfish complicates research aimed at analyzing the effect of knocking-out actin isoforms using clustered regularly interspaced short palindromic repeats (CRISPR) technology, due to potential functional redundancies (Bakkers, 2011). The Zebrasfish Information Network (ZFIN) database has assigned isoforms to zebrasfish actin genes on the basis of alignments with mammalian orthologs (Sprague et al., 2006); however, caution is necessary as the significant sequence identity between actin isoforms creates challenges with isoform assignments. Also, even slight divergence of the actin sequence from the human canonical sequences of actin isoforms can produce misassignments during machine annotation of actin gene isoforms.

To develop a more accurate picture of the actin genes in zebrasfish, we combined analysis of the actin genes in the zebrasfish genome with available experimental evidence to suggest the most probable zebrasfish z-actin gene. Our intention is to propose the genes that most likely produce actin isoform orthologs in zebrasfish as a starting point for further functional and quantitative experimental testing, recognizing that the significant sequence identity among actin sequences can mislead both computers and researchers.

2. Methods and results

A Basic Local Alignment Search Tool (BLASTP) search of homologous sequences of actin proteins in the zebrasfish compared with that of the human yielded 18 different proteins, with sequence identities up to 99%. BLAST-like Alignment Tool (BLAT) analysis revealed that several of these sequences result from the same gene. After merging duplicates, we identified nine distinct zebrasfish actin gene encoding the previous 18 different proteins. An actin sequence located on chromosome 20 that had not been previously assigned was found associated with the zgc: 86709 entry and was designated as actc1c (Table 1).

Duplicate genes on chromosome 19 and 20, both called actc1a, encode the same zebrasfish z-actin protein. The actc1a genes are almost identical. Both genes are 6.37 kb long with 9 exons. An alignment of the sequences between the start and stop codons of the genes reveals only 71 nucleotide changes and 7 gaps: five 1 bp, one 4 bp and one 20 bp long. Seven nucleotide changes are found among the 1134 nucleotides of the coding regions (99.4% identity).

The nine zebrasfish actin gene product sequences, together with human actin isoforms, were organized into clades in a phylogenetic tree using the Neighbor-Joining method corresponding to cytoplasmic actin proteins, smooth actin isoforms, and z-actin, paralogous (z-cytoplasmic and z-skeletal) isoforms (Fig. 1). The zebrasfish cytoplasmic actb1 and actb2 genes are distinguished from the smooth and striated actin isoforms, being most similar to human β-cytoplasmic actin, including the lack of a codon encoding cysteine after the initiator methionine seen among z-actin genes. Of the nine zebrasfish genes identified, no clear ortholog of human γ-cytoplasmic actin was included; however, the acta2 gene product was grouped into a clade with the human z- and γ-smooth actin isoforms. The remaining six zebrasfish actin gene products were associated with z-striated actin isoforms in a clade that was separated from the rest with a high level of confidence from the bootstrapping.

An alignment of the human z-actin protein sequences with the six zebrasfish putative z-actin proteins revealed significant levels of sequence identity (Table 1). Interestingly, none of the zebrasfish z-actin sequences possess the canonical human z-skeletal N-terminal acidic amino acid motif DEDE, while two zebrasfish z-actin proteins possess the corresponding DDEE motif found in human z-actin. The most common zebrasfish N-terminal acidic residue motif is DDDE, which is not found in any human sequence.

Aside from the four acidic amino acids at the N-terminus, only

---

### Table 1

**N-terminal sequences of human z-actin isoforms and zebrasfish actin translated genes.** Sequence positions are those found before N-terminal processing of the mature actin protein. The human z-actin isoform protein sequences (top) refer to UniProt database entries. Sequence names in **bold** possess the canonical z-cardiac N-terminal acidic residue sequence. The clusters of N-terminal acidic amino acids are underlined. The position of variable amino acids outside of the N-termini are listed across the top. Distinguishing variances are highlighted in **bold.** The zebrasfish gene names (bottom) correspond those in the ZFIN database (accessed June 14, 2016).

<table>
<thead>
<tr>
<th>Isoform</th>
<th>N-terminal sequence</th>
<th>UniProt</th>
</tr>
</thead>
<tbody>
<tr>
<td>z-skeletal</td>
<td>MCDDDETTAL</td>
<td>Q</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>z-skeletal</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>P68133</td>
</tr>
<tr>
<td>z-cardiac</td>
<td>MCDDDETTAL</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>z-cardiac</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>P68032</td>
</tr>
<tr>
<td>z-smooth</td>
<td>MCDDDETTAL</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>P62736</td>
</tr>
</tbody>
</table>

### Table 2

**Gene-specific primers for generating the probes used for in situ hybridization experiments (ISH), and for quantifying relative gene expression (RT-qPCR).**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actc1c F1</td>
<td>5′-ATGTTTAATGGTCTTCTGG-3′</td>
<td>ISH</td>
</tr>
<tr>
<td>Actc1c F1</td>
<td>5′-AAAAAATGCAGCTTTGG-3′</td>
<td>ISH</td>
</tr>
<tr>
<td>Acta1b F1</td>
<td>5′-TGCATCTGTTCTGCTG-3′</td>
<td>ISH</td>
</tr>
<tr>
<td>Acta1b F1</td>
<td>5′-ATGCCTGAATGGCTTCTG-3′</td>
<td>ISH</td>
</tr>
<tr>
<td>Acta1b F2</td>
<td>5′-TCCATCTGTTCTTCCATAGG-3′</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Actc1c F2</td>
<td>5′-AAATGGTCACTTTGGAACAT-3′</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Acta1b F2</td>
<td>5′-ATGGGGTCATCGTGCTTCT-3′</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Acta1b F2</td>
<td>5′-ATGGGGTCATCGTGCTTCT-3′</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Acta1b F2</td>
<td>5′-ATGGGGTCATCGTGCTTCT-3′</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Rpl8 F</td>
<td>5′-ATGCCAAGAATGGAGG-3′</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Rpl8 R</td>
<td>5′-TGGGCCAAACCGGCGGCC-3′</td>
<td>RT-qPCR</td>
</tr>
</tbody>
</table>
eight other positions among the 375 amino acids in the zebrafish or human α-actin protein sequences exhibit substitutions (Table 1), demonstrating the significant conservation of these actin protein sequences. For half of these changes, a change is found in only one of the nine sequences. Of those four changes, three are conserved (T162S, I167V, and L301), whereas one is less conserved (Q362P).

The *acta2* zebrafish α-actin possesses the α-cardiac actin-like DDEE N-terminal acidic residue motif with a characteristic threonine-to-serine change found in α-smooth actin and an α-smooth actin-like Ser-91. That zebrafish gene product was grouped in the α-smooth actin clade, but it is important to note that it is close to the α-cardiac actin. Alternatively, all of the non-N-terminal amino acids of the *actc1c* zebrafish α-actin match those of α-cardiac actin, excluding the presence of the common DDDE N-terminal motif in zebrafish actin sequences. The result is that the *actc1c* gene product is located close to the human α-cardiac actin sequence on the phylogenetic tree.

The remaining four zebrafish α-actin protein sequences are more closely related to human α-skeletal actin in the phylogenetic tree. All four contain the same zebrafish DDDE N-terminus motif. Subtle conserved changes seem to distinguish this group; namely the presence of Ala-280 and/or Thr-360. Since Thr-360 is present in α-skeletal actin and the product of the zebrafish *acta1a* gene contains a Thr-360 and Thr-280, *acta1a* is situated closer to α-skeletal actin than any other zebrafish actin gene. At the same time, α-skeletal actin does not encode Ala-280 and thus the remaining three genes possessing Ala-280 are defined on a separate branch. Similarly, cardiac and skeletal actin proteins, which exhibit N-terminal acidic amino acids identical to zebrafish muscle actins (DDDE and DDEE) have been identified in medaka *Oryzias latipes* (Kusakabe et al., 1999).

Interestingly, the remaining three actin genes — *actc1a, acta1b,* and *actc1b* — are the best candidate α-cardiac actin genes in zebrafish. The database naming of these genes as *actc* or *acta* highlights the need for thoughtful designation of actin genes that incorporates more information than sequence similarity alone. Research has shown cardiac-related effects with mutations in *actc1a* or *acta1b*. The S434 mutant of the *actc1a* gene results from a
Y169S amino acid substitution, exhibiting serious cardiac defects. In situ hybridization reveals that the actc1a gene products are expressed in the heart and somatic muscle (Glenn et al., 2012). The actc1b mutant known as cardiofunk (cfk), resulting from a R177H substitution, also exhibits cardiac defects and in situ hybridization shows gene expression in myocardial cells (Bartman et al., 2004; Wen and Rubenstein, 2003). A recent analysis of the zebrafish cardiac transcriptome confirms that actc1a and actc1b are expressed in the hearts of zebrafish, where actc1b is expressed at a higher level in early development and actc1a is expressed at higher levels during adulthood (Shih et al., 2015), suggesting development modulation of some zebrafish cardiac actin genes. Of note, the transcriptome results demonstrate that the third gene in this group, actc1b, is expressed at higher levels in the skeletal muscle of zebrafish, suggesting that this isoform does not function primarily in cardiac muscle. Ultimately, functional characterization like those obtained for actc1a and actc1b are needed to assign an isoform designation with high confidence.

As noted above, nearly identical actc1a genes were found on chromosomes 19 and 20. The S434 mutant studied by Glenn et al. (2012) was mapped only to the actc1a gene on chromosome 20. If there are two copies of the actc1a gene, one might expect compensation from the S434 mutation. One possibility is that the actc1a gene on chromosome 19 is an artifact of genome assembly. To address this issue and provide additional data regarding the assignment of ortholog isoforms, we compared the synteny of the nine zebrafish actin genes shown in Fig. 1 with the three major human α-actin isoforms (Fig. 2).

The syntenic analysis suggests that the β-actin isoforms in zebrafish resulted from a gene duplication involving both the actin gene and the neighbouring fcn genes encoding the actin-binding fascin protein. The syntenic analysis supports the assignment of the zebrafish acta1a and acta2 genes as orthologs of human α-skeletal (ACTA1) and α-smooth muscle actin (ACTA2), respectively. Interestingly, we see a gene encoding the centriole, cilia and spindle associated protein (CCSAP) associated with human ACTA1 and two actin genes in zebrafish; the ccsap gene linked to acta1a and ccsapb beside acta1b, suggesting that these genes may diverge from a common ancestor, with the acta1b gene perhaps resulting from duplication of a region around the actc1a gene that included the ccsap gene. In situ hybridization experiments localizing the products of the acta1b gene in both the heart and tail muscle of zebrafish (Glenn et al., 2012) suggests that this actin isoform functions in both tissues.

Are there duplicate acta1a genes on chromosome 19 and 20, or is this a result of a genome assembly error? Both acta1a genes are between a downstream rm2 gene, encoding the ribonucleotide reductase regulatory subunit M2, and an upstream scg5 gene, encoding secretogranin V. These regions encompass 46,612 bp and 26,412 bp on chromosome 19 and 20, respectively. The structures of these regions are very similar, with a single 20 kb insert starting 257 bp after the acta1a stop codon on chromosome 19. However, the common regions are not identical, with sufficient gaps and nucleotide changes to suggest that these are two different regions. A similar study was performed by Venkatesh et al. in which the actin genes were characterized in the Japanese pufferfish Fugu rubripes (Venkatesh et al., 1996). In this study, the cardiac actin genes encoded the same protein while differing in their nucleotide sequences and genomic organization. Given the high degree of similarity in the 5′ untranslated regions between the two acta1a genes, it is likely that both are expressed. Further experiments are required, however, to determine the expression patterns of the two proposed acta1a genes and how the presence of the acta1a gene on chromosome 19 affects the S434 mutation of the chromosome 20 gene.

Is the product of the actc1c gene identified here located in cardiac tissue? To answer this question, we performed whole mount in situ hybridization experiments to compare the localization of actc1c mRNA to the confirmed cardiac isoform acta1b in zebrafish embryos (Fig. 3). We found that both acta1b and actc1c were localized to the developing heart tissue in the embryos, supporting our sequence-based prediction that the actc1c gene encodes a cardiac actin isoform. Next, we compared the relative mRNA abundances of acta1b and actc1c in cardiac tissue from 36 hpf zebrafish embryos using a previously described heart isolation protocol (Burns and Mackae, 2006) and reverse-transcription quantitative polymerase chain reaction (RT-qPCR; Fig. 4). Results confirmed that both isoforms are expressed at quantifiable levels in embryonic hearts, with approximately 20-fold more acta1b than actc1c.

3. Discussion

The hallmark N-terminal acidic residue motifs distinguishing the α-actin protein isoforms in mammals are different in zebrafish. Since ZFIN relies on homology with mammalian sequences to assign a gene name, some of the assignments of the zebrafish actin genes may require additional information to be accurate. We have examined the actin sequences and considered experimental evidence to re-assign the most-likely isoform identity for zebrafish genes. In particular, assigning the actc1c name to a previously undesignated gene and confirming that this gene is expressed in the heart at early developmental stages. While there are no direct orthologs in zebrafish, sequence and syntenic analyses suggest the
We can target isoform sequencing. The Topo vector containing the USA. Insertion of the sequence into the vector was con
csequences were ampli
shing localization in the developing heart.

Fig. 3. Localization of isoform designations of the
CHRISPRs that are locus-
ners. Future studies include experimental evidence to
dcoding sequence-disrupting CRISPRs cannot target the
and exons encompassing the coding region are nearly identical. As
such, coding sequence-disrupting CRISPRs cannot target the
ence of zebrafish actin isoforms that are more closely related to
human isoforms. In particular, acta2 appears to be a smooth muscle
isoform and acta1a is closely related to \( \alpha \)-skeletal actin. Combined
with direct research, the actc1a and acta1b gene products are likely
\( \alpha \)-cardiac actin isoforms, while acta1b is likely an \( \alpha \)-skeletal

With this information in hand, we can now test the proposed
isoform assignments with a combination of functional and
expression analyses involving techniques such as comparative
fluorescence in situ hybridization, RT-qPCR, and CRISPR knockouts.
We can target \( \alpha \)-actin genes in zebrafish with greater confidence; however, the repeating of actc1a on two chromosomes presents
challenges when designing morpholinos or CRISPRs that are locus-
specific. For actc1a, the genomic sequences do not begin to diverge
until about 600 bp upstream of the start codon, while the introns
and exons encompassing the coding region are nearly identical. As
such, coding sequence-disrupting CRISPRs cannot target the
different loci. Future studies include experimental evidence to
accompany the functional assignment of actin proteins; this is
necessary due to the high conservation of the amino acid sequences
in the actin proteins. The work presented here permits clarification
of the isoform designations of the \( \alpha \)-actin genes in zebrafish to aid

in studies involving actin genes and proteins.

Editors’ note

Please see also related communications in this issue by Segert et al. (2018) and Rafferty et al. (2018).
Acknowledgements

We thank Dr. Andreas Heyland, Department of Integrative Biology at the University of Guelph, for access to equipment for in situ hybridization experiments. This work was funded by a Grant-in-Aid to the Heart and Stroke Foundation of Canada to JD (G-15-0008961).

References


