Mutation Research 668 (2009) 117–132

Contents lists available at ScienceDirect

Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis

journal homepage: www.elsevier.com/locate/molmut
Community address: www.elsevier.com/locate/mutres

The Fanconi anemia/BRCA gene network in zebrafish: Embryonic expression and comparative genomics

Tom A. Titus, Yi-Lin Yan, Catherine Wilson, Amber M. Starks, Jonathan D. Frohnmayer, Ruth A. Bremiller, Cristian Cañestro, Adriana Rodriguez-Mari, Xinjun He, John H. Postlethwait*

Institute of Neuroscience, University of Oregon, 1425 E. 13th Avenue, Eugene, OR 97403, USA

ARTICLE INFO

Article history:
Received 2 October 2008
Accepted 23 November 2008
Available online 3 December 2008

Keywords:
Genome duplication
Disease model

ABSTRACT

Fanconi anemia (FA) is a genetic disease resulting in bone marrow failure, high cancer risks, and infertility, and developmental anomalies including microphthalmia, microcephaly, hypoplastic radius and thumb. Here we present cDNA sequences, genetic mapping, and genomic analyses for the four previously undescribed zebrafish FA genes (*fanci*, *fancj*, *fancm*, and *fancn*), and show that they reverted to single copy after the teleost genome duplication. We tested the hypothesis that FA genes are expressed during embryonic development in tissues that are disrupted in human patients by investigating *fanc* gene expression patterns. We found *fanc* gene maternal message, which can provide Fanc proteins to repair DNA damage encountered in rapid cleavage divisions. Zygotic expression was broad but especially strong in eyes, central nervous system and hematopoietic tissues. In the pectoral fin bud at hatching, *fanc* genes were expressed specifically in the apical ectodermal ridge, a signaling center for fin/limb development that may be relevant to the radius/thumb anomaly of FA patients. Hatching embryos expressed *fanc* genes strongly in the oral epithelium, a site of squamous cell carcinomas in FA patients. Larval and adult zebrafish expressed *fanc* genes in proliferative regions of the brain, which may be related to microcephaly in FA. Mature ovaries and testes expressed *fanc* genes in specific stages of oocyte and spermatocyte development, which may be related to DNA repair during homologous recombination in meiosis and to infertility in human patients. The intestine strongly expressed some *fanc* genes specifically in proliferative zones. Our results show that zebrafish has a complete complement of *fanc* genes in single copy and that these genes are expressed in zebrafish embryos and adults in proliferative tissues that are often affected in FA patients. These results support the notion that zebrafish offers an attractive experimental system to help unravel mechanisms relevant not only to FA, but also to breast cancer, given the involvement of *fancj* (*brip1*), *fancn* (*palb2*) and *fancd1* (*brca2*) in both conditions.

1. Introduction

Fanconi anemia (FA; MIM# 227650) is a rare autosomal recessive disorder appearing at a frequency of about 3 per million and affecting approximately 2000 families in the United States. This devastating disease is characterized by developmental abnormalities in a number of organ systems and catastrophic bone marrow failure, often by 5 years of age. Bone marrow transplantation has become an effective therapy [1–6], but survivors experience increased susceptibility to squamous cell carcinomas of the head and neck [7,8]. Thirteen FA complementation groups have been identified and their genes cloned (*FANCA*, *FANCB*, *FANCC*, *FANCD1 (BRCA2)*, *FANCD2*, *FANCE*, *FANCF*, *FANCG (XRCC9)*, *FANCI*, *FANCJ (BRIP1)*, *FANCL*, *FANCM*, and *FANCN (PALB2)*) [9–28]. FA proteins interact in a complex network that facilitates a DNA damage response leading to DNA repair. Cells exposed to DNA damaging agents or passing through the DNA synthesis phase of a normal cell cycle activate a nuclear core complex consisting of eight FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM). The nuclear core complex functions as an E3 ligase to trigger the monoubiquitination of the downstream proteins FANCI and FANCD2 (the ID complex) [21,29]. Monoubiquitination of the ID complex allows it to translocate to nuclear DNA repair foci containing BRCA1, histone H2AX, FANCD1 (BRCA2), FANCJ, FANCN and RAD51 [21,25,29,30]. In ways that are not fully understood, these protein complexes target to sites of DNA damage, which initiates DNA repair mechanisms. Biallelic mutation of any of the FA genes upstream of the ID complex disrupts ID monoubiquitination leading to the loss of nuclear foci and resulting genomic instability as reflected by the hypersensitivity of DNA to interstrand crosslinks (ICLs) caused by genotoxic agents such as Cisplatin, mitomycin C (MMC) and diepoxybutane (DEB) [31,32].

* Corresponding author. Tel.: +1 541 346 4538; fax: +1 541 346 4538.
E-mail address: jpostle@uoneuro.uoregon.edu (J.H. Postlethwait).

0027-5107/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrfmmm.2008.11.017

The sensitivity of FA cells to ICLs reflects defects in DNA repair mechanisms that lead to aberrant apoptosis, genomic instability, and cancer. Because the FA pathway intersects pathways for breast and other cancers [29,30], and is involved in the evolution of resistance to cancer chemotherapies [33,34], a better understanding of the mechanisms that lead to Fanconi anemia will be broadly applicable to the biology of malignancy. Despite major advances in our knowledge of the biochemistry of FA proteins, we still know little about the mechanisms by which loss-of-function mutations in the FA network impact the DNA damage response pathway and contribute to the heterogeneous clinical features found in FA patients (developmental defects, progressive onset of aplastic anemia, and increased predisposition to hematological malignancies and the formation of solid tumors). Furthermore, FA cells have other defects, including poor resistance to oxidative damage, premature telomere shortening, abnormal cell cycle kinetics, interaction with inflammation pathways, and hyperactivation of the MAPK pathway leading to overproduction of TNF-alpha [35–38].

People with mutations in some complementation groups that alter the nuclear core complex experience greater cancer risks than others, which would not be expected if the only function of these proteins were to activate the ID complex. For some FA genes, even heterozygotes suffer an elevated risk of breast cancer (*FANCD1 (BRCA2)*, *FANCN (PALB2)* and *FANCJ (BRIP1)*). Understanding the network of interactions that unites different complementation groups into a single disease will benefit from studies that exploit model systems. The zebrafish model shares cellular, developmental, and genetic features with humans and provides advantages that facilitate experimentation. Clinically relevant studies designed to understand the complex web of interactions that unite the different complementation groups into a single disease could exploit a convenient, small vertebrate model such as zebrafish.

In this paper, we first report the identification and isolation of the 4 previously unidentified zebrafish orthologs of human FA genes, and then present gene expression data for all 13 complementation groups to address the mechanisms that cause FA to affect some organs but not others and to understand why different complementation groups have different phenotypes. Finally, we evaluate zebrafish as a model for future FA research.

2. Materials and methods

2.1. Genomic analysis

The Zv7 version of the zebrafish genome assembly (http://www.ensembl.org/Danio_rerio/index.html) was searched using human Fanconi protein sequences. We further investigated sequences satisfying the reciprocal best BLAST "hit" (RBH) method for orthology [39]. Sequences from the genome assembly were used to design primers for 5′ and 3′ rapid amplification of cDNA ends (RACE, Advantage cDNA PCR Kit, Clontech, Inc., Palo Alto) using as template 5′ or 3′ first strand zebrafish cDNA synthesized from mRNA of pooled embryos at 12, 24, and 48 h post-fertilization (hpf). Some genes were too large for full-coverage by RACE, and so we performed standard amplification with forward and reverse gene-specific primers, using as template second strand cDNA from 60 hpf embryos. Amplified gene fragments were cloned using the TOPO Cloning Kit for Sequencing (Invitrogen, Carlsbad, CA). FANC proteins were aligned using ClustalW [40]. Genomic structure was inferred using Genomescan (http://genes.mit.edu/genomescan/). Zebrafish *fanc* genes were mapped by single strand conformation analysis (SSCP) on the heat shock doubled haploid mapping panel [41]. For comparative mapping, putative orthologs were defined by RBH [39]. The human map locations of putative orthologs were obtained from Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi). Nomenclature rules specify human, mouse, and zebrafish genes as *FANCA*, *Fanca*, and *fanca*, respectively and human, mouse, and zebrafish proteins as FANCA, FANCA, and Fanca, respectively (http://zfin.org/zf_info/nomen.html). When indicating a vertebrate protein without respect to species, we use the nomenclature Fanca.

2.2. Expression analysis

For reverse transcriptase (RT) PCR, total RNA was extracted following instructions in the TRI REAGENT kit (Molecular Research Center Inc., TR-118) using pools of 40–50 embryos for each developmental stage. First strand of cDNA was generated from 2 μg of total RNA, using Superscript III Rnase H-reverse transcriptase (Invitrogen, #18080-044) and oligo(dT) primer. After deactivation of the reverse transcriptase by incubating the sample at 70 °C for 15 min, RNA was degraded with RNase H (Biolabs, M0297S). A dilution 1:10 or 1:20 of the first strand cDNA (1:200 for actin) was then used to assay for gene expression by PCR. Gene-specific oligonucleotide primers were *fanca*, GCAGACCCGGAACAGCCCACAC/CAGCGCTGAATAATCCCGCAGACA; *fancb*, CGGCCCGTCTGCGGTGAAGA/GCCGCTGGAGAACTGAAGCCACAC; *fancc*, TGGAGGGGGCAGCAGTGAGC/TTGGTGGGGTGGTGGAAGAACG; *fancd1*, GGGCCAGAAAACACAGCAACTCAAA/GCACAGGCCCAGATAGCACTCG; *fancd2*, GCAGCGGGCGATCCACAAAGTC/GGCCATCCTCACTCGCTCCTTCAA, *fance*, CGGCCTCTCGCTGTTTGGTGAC/GCGGCCTGCAGTGATTTCTTGAG; *fancf*, CTGCTGCAGCGGAGCGTCTG/ATTTCACTGACACAATTTATTACTAAG; *fancg*, GCGCACTTTTGGCTGCTCTGTGTA/ACCACCGAGCAGCATAGCAGGAGA; *fancj*, CCAGAGCATCCAAACCACCTACAG/GTTGTTCCCCCGTGTCTTGTTCTC; *fancl*, GACGGCTTCATCACAGTGCTGGAAAA/GCCTTCAGCTGGAGTGTGCGAAACT; *fancm*, GAGCCCCGAGGACCAGGAG/GTGGGCGCCATGAAGACGA; *actin*, GAGAAGATCTGGCATCACACCTTC/GGTCTGTGGATACCGCAAGATTC.

In situ hybridization to mRNA was used to detect *fanc* gene expression in whole-mounted embryos. For RNA probes, we in vitro transcribed clones linearized with TOPO/Not-1 and labeled them with digoxigenin-UTP using T3 RNA polymerase. Wild-type embryos (AB strain) were fixed in 4% paraformaldehyde at 41 °C for at least 2 days before dechorionating by hand using a dissecting microscope. In situ hybridizations were performed as described [42] with several individuals for each developmental stage [43]. The University of Oregon IACUC approved experiments for this study. Probes were—*fanca*: nucleotides 1–1529 of NM_001040635, including 94 nt of 5′UTR and exons 1–10; *fancb*: nucleotides 399–1127 on NM_001040636, exons 1–3; *fancc*: nucleotides 240–822 on NM_001040637, exons 1–7, excluding exon 4 (probe constructed from splice variant without exon 4); *fancd1*: nucleotides 681–1012 of NM_001110394, exons 7–10; *fancd2*: nucleotides 1928–2445 of NM_201341, exons 20–25; *fance*: nucleotides 643–1012 on NM_001040634, exons 2–10 plus 64 nt of 3′UTR; *fancf*: nucleotides 45–1128 of NM_001045234, including 20 nt of the 5′UTR, the entire single exon, and 41 nt of the 3′UTR; *fancg*: nucleotides 678–1064 of NM_205639, exons 5–7; *fanci*: nucleotides 3364–3910 of XM_001921104, exons 37–42; *fancj*: nucleotides 2048–2931 of NM_001110296, exons 12–18; *fancl*: nucleotides 90–809 of NM_212982, 54 nt of 5′UTR and exons 1–8; *fancm*: nucleotides 423–2346 of NM_001113660, including exons 1–10 and additional exons not yet annotated because of incomplete genomic sequence; *fancn*: nucleotides 225–975 of XM_001919731, exons 2–8.

3. Results

3.1. Does zebrafish contain an intact FA network?

To exploit zebrafish for FA research, we must identify similarities and differences between the FA networks of zebrafish and human. The first evidence that zebrafish possesses an FA network functionally similar to human was the isolation and sequencing of *fancd2* cDNA [44], followed by the identification of *fancg* with its multiple TPR motifs conserved with human [45]. These results suggested that zebrafish may possess a functional FA network.

Evidence that zebrafish has a full FA genetic system came with the isolation and characterization of cDNAs and/or genomic BAC clones for zebrafish orthologs for the remaining human FA genes known at the time (*FANCA*, *FANCB*, *FANCC*, *FANCD1*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, and *FANCL*) [46]. The sequence of coding regions showed that zebrafish FA proteins often share low levels of overall identity with the human genes—for example, Fancd1, Fancb, Fancf, and Fance proteins have only 21–28% amino acid identities between the zebrafish and human proteins, but some regions within these proteins were more highly conserved, suggesting regions of functional significance. Proof that the zebrafish genes are in fact orthologs of the human genes comes from three independent data sets. First, the intron/exon organization of the zebrafish/human ortholog pairs is nearly identical. Second, amino acid hydrophobicity plots are extensively correlated, showing overall conservation of protein shape. Third, comparative synteny analysis showed that the neighbors of zebrafish *fanc* genes are generally orthologs of human *FANC* gene neighbors. This shows that chromosome regions containing *fanc/FANC* genes have been evolutionarily conserved for the 450 million years that separate zebrafish and human from their last common ancestor [47].