The Journal of Immunology

Ikaros Family Members from the Agnathan *Myxine glutinosa* and the Urochordate *Oikopleura dioica*: Emergence of an Essential Transcription Factor for Adaptive Immunity

Pauline M. Cupit, John D. Hansen, Aaron S. McCarty, Greg White, Mariacristina Chioda, Fabio Spada, Stephen T. Smale and Charles Cunningham

J Immunol 2003; 171:6006-6013; ;
doi: 10.4049/jimmunol.171.11.6006
http://www.jimmunol.org/content/171/11/6006

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Ikaros Family Members from the Agnathan *Myxine glutinosa* and the Urochordate *Oikopleura dioica*: Emergence of an Essential Transcription Factor for Adaptive Immunity[1]

Pauline M. Cupit,[2]* John D. Hansen,[2,3†] Aaron S. McCarty,[‡] Greg White,* Mariacristina Chioda,* Fabio Spada,* Stephen T. Smale,* and Charles Cunningham*

The Ikaros multigene family encodes a number of zinc finger transcription factors that play key roles in vertebrate hemopoietic stem cell differentiation and the generation of B, T, and NK cell lineages. In this study, we describe the identification and characterization of an Ikaros family-like (IFL) protein from the agnathan hagfish *Myxine glutinosa* and the marine urochordate *Oikopleura dioica*, both of which lie on the evolutionary boundary between the vertebrates and invertebrates. The IFL molecules identified in these animals displayed high conservation in the zinc finger motifs critical for DNA binding and dimerization in comparison with those of jawed vertebrates. Expression of the IFL gene in hagfish was strongest in blood, intestine, and gills. In *O. dioica*, transcription from the IFL gene was initiated at or around the time of hatching and maintained throughout the life span of the animal. In situ hybridization localized *O. dioica* IFL expression to the Fol cells, which are responsible for generating the food filter of the house. Biochemical analysis of the DNA binding and dimerization domains from *M. glutinosa* and *O. dioici* IFLs showed that *M. glutinosa* behaves as a true Ikaros family member. Taken together, these results indicate that the properties associated with the Ikaros family preceded the emergence of the jawed vertebrates and thus adaptive immunity. *The Journal of Immunology*, 2003, 171: 6006–6013.

Lower vertebrates and invertebrates exist in an environment surrounded by pathogenic organisms, yet survive with an immune system that does not exhibit the complexity of higher vertebrates. The adaptive immune system of gnathostomes or jawed vertebrates shows both Ag specificity and memory and is characterized by the presence of Ig, TCR, and MHC (1). In addition, gnathostomes benefit from a diverse leukocyte population, including B and T lymphocytes, NK cells, macrophages, and granulocytes. In almost all vertebrates, including bony fish and elasmobranchs, Ig, TCR, and MHC appear to be present, but no homologues of these genes have been identified in agnathans (jawless vertebrates), despite exhaustive searches (1, 2). Agnathans represent the most primitive living vertebrates that include lamprey (Petromyzontiformes) and hagfish (Myxiniformes), which diverged from a common ancestor at least 550 million years ago (3). Although hagfish lack the classical hemopoietic tissues, lymphoid-like cells have been identified that are particularly abundant in the intestine, suggesting that this organ may function as a primordial spleen (4). Other sites implicated in hemopoiesis are the pronephrotic mass, gills, liver, and blood, which all have circulating lymphocyte-like cells. Similarly in lamprey, the closest living vertebrate to hagfish, lymphoid-like cells were recognized in the gut epithelium, pronephros, and the gonads (4–6).

The protochordate *Oikopleura dioica* is a pelagic tunicate (Urochordata) that occupies a phylogenetic niche close to the transition between vertebrates and invertebrates (7). Oikopleura is an appendicularia, a term that refers to their house appendices. The filter-feeding house of appendicularia enables them to use a wide range of food particles and thus become an important component of marine zooplankton. To maintain adequate filtration rates, the house must frequently be replaced. For *O. dioica*, a new house is synthesized every 3–4 h over a life cycle of 5–6 days at 15°C (8). To date, there is no evidence for any molecules that may be ascribed to an immune function in these animals. Because no lymphocyte-like cells have been identified in other deuterosomes, the earliest period for the emergence of lymphocytes is concomitant with the divergence of cartilaginous fish, >420 million years ago. Around this time point, two possible gene duplication or partial duplications are thought to have occurred, creating a large gene pool that provided opportunities to further gene diversification and adopt new roles peculiar to vertebrates (9, 10). Recent evidence of the existence of transcription factors involved in lymphocyte differentiation in skate (cartilaginous fish) and lamprey indicates that they duplicated around the time of the appearance of Ig, TCR, and MHC (11–15). Therefore, the isolation of regulatory transcription factors can be used to identify homologues and putative hemopoietic sites, which may improve our understanding about the evolution of lymphopoiesis in vertebrates. One important family of transcription factors is the Ikaros family, whose members function as key regulators of hemopoiesis.

*Sars International Centre for Marine Molecular Biology, High Technology Centre, Bergen, Norway; †Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Baltimore, MD 21202; and ‡Howard Hughes Medical Institute and Molecular Biology Institute, University of California, Los Angeles, CA 90095

Received for publication June 26, 2003. Accepted for publication September 26, 2003.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This study was supported by Public Health Service Training Grant to A.S.M. S.T.S. is an investigator of the Howard Hughes Medical Institute. J.D.H. is supported by the United States Department of Agriculture (National Research Initiative Competitive Grants Program 2002-03472). This is University of Maryland Biotechnology Institute/Center of Marine Biotechnology technical paper 03-585. Sequences reported in this paper have been deposited in the European Molecular Biology Laboratory database (accession AY237104-107).

[2] P.M.C. and J.D.H. contributed equally to this work and share joint first authorship.

[3] Address correspondence and reprint requests to Dr. John Hansen, Center of Marine Biotechnology, 701 E. Pratt Street, Baltimore, MD 21202. E-mail address: hansenj@umbi.umd.edu

0022-1767/03/$02.00

The development of the lymphoid system, at least in higher vertebrates, is regulated by a network of transcription factors in hemopoietic stem cells and signaling mediated through cell contact and growth factor receptors. Gene inactivation experiments in mice have identified several transcription factors, including the GATA family, Pax5, PU.1, and Ikaros that are crucial for early lymphocyte lineage development and the development of other hemopoietic cell types (16–20). The expression of Ikaros is restricted to hemopoietic cells and is essential for hemopoietic stem cell differentiation to the lymphocyte lineages. It is expressed in the earliest hemopoietic progenitors and throughout the life of the lymphocyte as well as in all stages of B, T, and NK cell lineage development (18). Ikaros is a member of a small, closely related gene family, two members of which, Aiolos and Helios, are also implicated in lymphocyte development (21–23). These Ikaros family proteins can contain up to six C2H2 zinc fingers (ZFs)[4] organized into two distinct domains. The four N-terminal ZF motifs participate in DNA binding, whereas the two C-terminal ZFs are involved in homo- and heterodimerization between members of the Ikaros clan (i.e., Ikaros, Helios, and Aiolos). The Ikaros multigene family function as regulators of hemopoiesis with proposed roles in gene activation and silencing during lymphocyte development (24, 25); in particular, its association with heterochromatin containing transcriptionally inert genes suggests that it may act as a negative regulator (26, 27).

The ZF motifs of the Ikaros gene family have been conserved throughout vertebrate evolution, allowing genes from phylogenetically distant vertebrate species to be cloned (11, 18, 28–31). In this study, we describe an Ikaros family-like (IFL) transcription factor from the most primitive vertebrate, the agnathan Atlantic hagfish *Myxine glutinosa*, and from two marine urochordates, *O. dioica* and *Ciona intestinalis*, which lends insight into the evolution of this essential gene family and that of adaptive immunity of vertebrates.

Materials and Methods

Animals

Atlantic hagfish were collected in baited traps offshore from the Marine Research Institute field station at Espegrend, Norway, and then transferred to a darkened aquarium containing fresh seawater at 8°C. *O. dioica* were collected from the fjords around Bergen, Norway, and cultured, as previously described (8).

Isolation and analysis of IFL cDNA and genomic clones

A cDNA clone encoding a pseudo-Ikaros gene (in-frame stop codon) was isolated from a Pacific hagfish (*Eptatretus stouti*) cDNA library (kind gift from M. Flajnik, University of Maryland School of Medicine) by cross-hybridization using a probe encoding rainbow trout Ikaros ZF1–4. The pseudo-Ikaros transcript was used as a template to generate a cDNA probe encompassing ZF1–4. This probe was used to screen a *M. glutinosa* peripheral blood leukocyte λ ZAP Express (Stratagene, La Jolla, CA) cDNA library. Two identical full-length hagfish Ikaros-like (HIL) clones with uninterrupted open reading frames (ORFs) were isolated. HIL cDNA was used as a template to generate a probe by PCR recognizing the HIL ZFs homologous to the IFL ZF5 and 6. This probe was used to screen a λ DASH II *M. glutinosa* genomic DNA library. Four overlapping λ clones were identified, and Southern analyses of the clones with probes recognizing HIL 5′ and 3′ untranslated regions (UTRs) revealed that only a 3′ portion of the gene (exons Eδ and Eϵ in Fig. 2) was present. Subsequent efforts to clone the missing 5′ genomic DNA from the library using a 5′UTR cDNA probe were unsuccessful. The intron-exon boundaries of the genomic DNA encoding ZF1–4 were subsequently determined by PCR. The position of the forward and reverse primers (see Fig. 2) was chosen based on the known position of intron-exon boundries in lamprey, trout, and mouse. PCR experiments were conducted using genomic DNA as template to generate amplicons both across these predicted boundries and within predicted exons. All amplicons were cloned and sequenced to verify that they corresponded to the expected products.

Analysis of a database of randomly cloned *O. dioica* genomic DNA fragments revealed four clones with high homology to vertebrate IFL genes. One of these clones was used as a template to generate a probe by PCR recognizing IFL ZF3 that was used to screen an *O. dioica* λ DASH II genomic library. Two overlapping clones were obtained that were found to encode an apparently complete IFL gene. Genomic DNA was then used as a template to produce a probe by PCR recognizing the *O. dioica* ZF homologous by sequence and position to vertebrate IFL ZF5. This probe was used to isolate a full-length *O. dioica* Ikaros-like (OIL) cDNA from a λ ZAP Express day 3 library (8).

In situ hybridization using OIL

A probe derived from the full-length OIL clone was labeled with digoxigenin-11-dUTP (Roche MB, Mannheim, Germany) by nick translation. Three-day-old animals were fixed with 4% paraformaldehyde in 0.1 M MOPS, pH 7.5, and 0.5 M NaCl for 1 h at room temperature, washed in 2× SSC containing 0.1% Tween 20 (2× SSCT), and then stored in methanol at −20°C overnight. The probe was resuspended in hybridization buffer (50% formamide, 2× SSCT, 25 mM sodium phosphate, pH 7.2, 10 mM EDTA, and 15% dextran sulfate) at a concentration of 5 ng/μl and applied to the animals. Samples and probe were then denatured at 90°C for 2 min and immediately cooled at 42°C. Probes were detected by overnight incubation with FITC-conjugated sheep anti-digoxigenin Fab (Roche MB). After washing, samples were counterstained with 1 μM TO-PRO-3 (Molecular Probes, Eugene, OR) and mounted in Vectashield mounting medium (Vector Laboratories, Burlingame, CA). Images were collected with a TCS SP laser-scanning confocal microscope (Leica, Deerfield, IL) equipped with Leica Confocal Software. When the probe was omitted from control samples otherwise subjected to the same procedure, no signal was detected over a low intensity diffuse background common to all samples. Also, a number of Fab Abs unrelated to that used in this study showed no nonspecific staining of any *O. dioica* structures when used under the same conditions.

GST fusion proteins and shift assay

cDNAs encoding the N-terminal ZF of OIL (ZF1–4) and HIL (ZF1–3) DNA binding regions were cloned into pGEX-4T (Pharmacia, Peapack, NJ) for the production of Ikaros-GST fusion products for EMSA analysis using a consensus Ikaros target sequence as the probe. Briefly, small scale cultures were induced with isopropyl β-D-thiogalactoside (1.5 h) in the presence of 50 mM ZnCl. Fusion proteins were then purified from the lysates using glutathione-Sepharose spin columns (Amersham, Arlington Heights, IL). All buffers postinduction contained 10 mM ZnCl. A double-stranded probe conforming to Ikaros binding site (IKBS4) (32) was end labeled with [γ-^{32}P]ATP. Approximately 25,000 cpm of labeled probe was incubated with the purified fusion proteins (~5 μg) at room temperature. Samples were loaded directly onto 10% PAGE gels containing glycerol, electrophoresed, dried, and exposed to film for 16 h. Specificity was determined by adding a 50 molar excess of cold IKBS4.

Dimerization assay

Flag-tagged and untagged murine, OIL, and HIL1 constructs were generated using the murIK I parental vector in conjunction with PCRSoeing, as described (25). Briefly, the murine dimerization ZF (DZF) domain was replaced by OIL and HIL1 aa 488–562 and 384–447, respectively. Human embryo kidney 293T cells were transfected and harvested 48 h later. The cytoplasmic fraction was used for all DZF assays. Coimmunoprecipitation and chemical cross-linking assays have been described (33).

Sequence analysis

DNA sequencing was performed using Applied Biosystems (Foster City, CA) cycle-sequencing chemistry. GCG software package (Genetics Computing Group, Madison, WI), BLAST (NCIB), and Pfam (version 7.7b) were used for sequence assembly and analysis. Amino acid sequences were aligned using Clustal X version 1.81 and visual inspection. From this alignment, a phylogenetic tree was constructed by the neighbor-joining method and bootstrapped 1000 times.

Results and Discussion

Isolation of Ikaros family-related cDNAs

Two identical clones were isolated from an Atlantic hagfish PBL cDNA library that contained an apparently full-length sequence.

[4] Abbreviations used in this paper: ZF, zinc finger; DZF, dimerization ZF; HIL, hagfish Ikaros-like; IFL, Ikaros family-like; OIL, *O. dioica* Ikaros-like; ORF, open reading frame; UTR, untranslated region.

The 4933-bp sequence included a 628-bp 5′UTR, an ORF of 1341 bp encoding 447 aa, and a 2964-bp 3′UTR. We have named this clone HIL1. BLASTX analysis of the nucleotide sequence showed a significant alignment of HIL1 with Ikaros family members from jawed vertebrates, especially Ikaros and Helios. Highest similarity (60%), however, was with a lamprey (*Petromyzon marinus*) Ikaros family member (e-102). The inferred protein from HIL1 was 41 and 29% identical with the lamprey and human Ikaros proteins, respectively. Pfam analyses suggested the presence of four C2H2 ZFs (Fig. 1; HIL1 ZF1, 2, 3, and 6) within the inferred HIL1 protein; however, manual inspection suggested a fifth C2H2 ZF to be present (Fig. 1; HIL1 ZF5). The alignment suggests that the five HIL1 ZFs are equivalent to mammalian Ikaros ZFs 1, 2, 3, 5, and 6. For DNA binding, Ikaros family members must possess at least three of the four N-terminal ZFs (18).

While cloning and sequencing HIL1 products generated by RT-PCR, we identified a 579-bp fragment whose nucleotide sequence was identical with that of HIL1 (Fig. 1), with the exception of a 180-bp insertion encoding ZF4. This partial clone was named HIL2. When hagfish genomic DNA was digested with a series of restriction enzymes with 6 base recognition sites and probed with either DNA common to HIL1 and HIL2 or novel HIL2 DNA, identical banding patterns were observed (data not shown). This confirmed that HIL1 and HIL2 were almost certainly the result of differentially spliced products from the same gene.

BLASTX analyses of a library of shotgun clones derived from *O. dioica* genomic DNA revealed ZF containing clones with high homology to vertebrate IFL proteins. This sequence was used to amplify a probe from an *O. dioica* cDNA library that showed a relatively high level of identity with vertebrate IFL transcripts, and was used to screen the *O. dioica* cDNA library. Only one apparently full-length cDNA clone encoding an *O. dioica* IFL protein (OIL) was obtained (Fig. 1). This 1890-bp transcript contained an 83-bp 5′UTR, an ORF of 1686 bp (562 aa), and a 228-bp 3′UTR. BLASTX scores showed highest homology to mouse and skate Helios (e-40) and lamprey Ikaros (e-38). Pfam analysis suggested OIL to encode six C2H2 ZFs (OILZF1–6). The OIL protein shared only 19% identity with the hagfish and *P. marinus* Ikaros. Alignment of the amino acid sequence of Ikaros family members from four diverse vertebrate species as well as the IFL proteins from agnathans and *O. dioica* showed that the length of each ZF was conserved (Fig. 1).

Finally, a tBLASTn search of the recently updated *C. intestinalis* (ascidian tunicate) database using OIL as bait revealed the presence of two *Ciona* IFL genes that showed a high degree of similarity to the OIL sequence (*Ciona* IFL1) and murine/human Pegasus (*Ciona* IFL2) (Fig. 1). The *Ciona* IFL2 gene encoded two N-terminal ZFs corresponding to the first and third ZFs for mammalian Pegasus (38% similarity for the three Pegasus-like sequences), which contains three N-terminal ZFs (34), while the *Ciona* IFL1 encoded four N-terminal ZFs. *Ciona* IFL1 is 39% similar to OIL, with almost perfect identity for ZF1–4. Both *Ciona* genes encode the two C-terminal ZFs that form the dimerization domain. Finally, it appears that ZF4 differs among the IFL members in that some members possess the standard C2H2 finger (i.e., murine-IK and Pegasus), while others most likely use C3H fingers (i.e., Oiko-IFL, murine-AI) for DNA recognition and binding.

Exon organization of hagfish and O. dioica *IFL genes*

The exon boundaries of mouse and lamprey *Ikaros* genes are shown in Fig. 2. Both of these genes contain seven exons (E1–7). Four identical recombinant λ phage clones were isolated from the hagfish genomic library. Southern analysis of one of these clones (gHIL) using HIL1 5′ and 3′ UTR cDNA probes indicated that part of the 5′ end of the gene was missing. Nucleotide sequencing revealed that gHIL began within the equivalent of mouse/lamprey intron 5. Two exons were observed in gHIL (Fig. 2); the first (Eδ) contains the equivalent of mouse/lamprey exon 6. The second exon (Eϵ) is similar to mouse/lamprey exon 7 and encoded ZF5 and ZF6. The presence of exons Eα-Eγ (mouse/lamprey exons 3–5) was confirmed by PCR.

Two overlapping clones were isolated from the *O. dioica* genomic library, and the exon organization of the *O. dioica OIL* gene is shown in Fig. 2. The gene contained nine exons (Ea-i). The positions of only two intron-exon boundaries were conserved between *O. dioica* and the mouse/lamprey *Ikaros* genes. These were the boundaries between Eb/c (mouse/lamprey E3/4) and Ec/d (mouse/lamprey E4/5). One other notable feature of the gene structure was the observation that ZF5 was split between exons Eh and Ei. The amino termini of both HIL1 and OIL are shorter than those of their mouse/lamprey Ikaros family counterparts. It is possible that the sequence equivalent to that contributed by mouse/lamprey exons E1 and E2 has been spliced out of HIL1. The 5′ sequence of the *OIL* gene did not suggest the presence of an equivalent to mouse/lamprey exon E1.

Phylogenetic analyses

The amino termini of the agnathan and urochordate proteins appeared to be of variable length, so phylogenetic analysis was conducted on the amino acid sequences encompassing the first to sixth ZFs of vertebrate IFL members Ikaros, Helios, and Aiolos. The two agnathan IFLs clustered tightly, forming their own separate lineage just before the branch leading to rest of the Ikaros family members from vertebrates that possess an adaptive immune system (i.e., skates-humans). Although the branch length leading to the Pegasus IFLs is deep, it does indicate that two IFL genes are present within nonvertebrate deuterostomes and that the Pegasus clade itself is most likely the result of a separate, older duplication event that occurred before the emergence of protochordates. All of these observations were supported by high bootstrap values (Fig. 3). Finally, it should be noted that a second alignment was generated (data not shown), in which the second ZF of *Ciona* IFL2 was aligned with the ZF3 region instead of the ZF4 region. The topology of the NJ-tree resulting from the second alignment was nearly identical with that shown in Fig. 3.

Expression analysis of HIL and OIL

In higher vertebrates, expression of *Ikaros* is confined to hemopoietic sites. To determine the tissues that express the gene encoding HIL1, RT-PCR was performed on total RNA extracted from a number of potential hemopoetic tissues. Fig. 4*A* shows that HIL1 was expressed in all tissues tested, with strongest expression in blood, gills, and intestine. Coincidentally, intestinal cells from lamprey with lymphocyte-like morphology were shown to express both Spi-B, a transcription factor involved in hemopoiesis, and IFL transcripts, lending support that the gut is an important site for lymphocyte-like cell development in agnathan fish (6, 14). RT-PCR also revealed the presence of two other transcripts that were present in all tissues and mirrored the expression of the predominant transcript. These bands were cloned and sequenced and revealed the isoform HIL2 and pseudo-IFL sequences that showed 80 and 90% amino acid identity with each other and the original *E. stouti* pseudo-transcript, respectively, which may represent additional IFL genes in hagfish, as suggested from the phylogenetic analysis. The psuedo-IFL transcripts possessed either in-frame stop codons or frame shifts (data not shown).

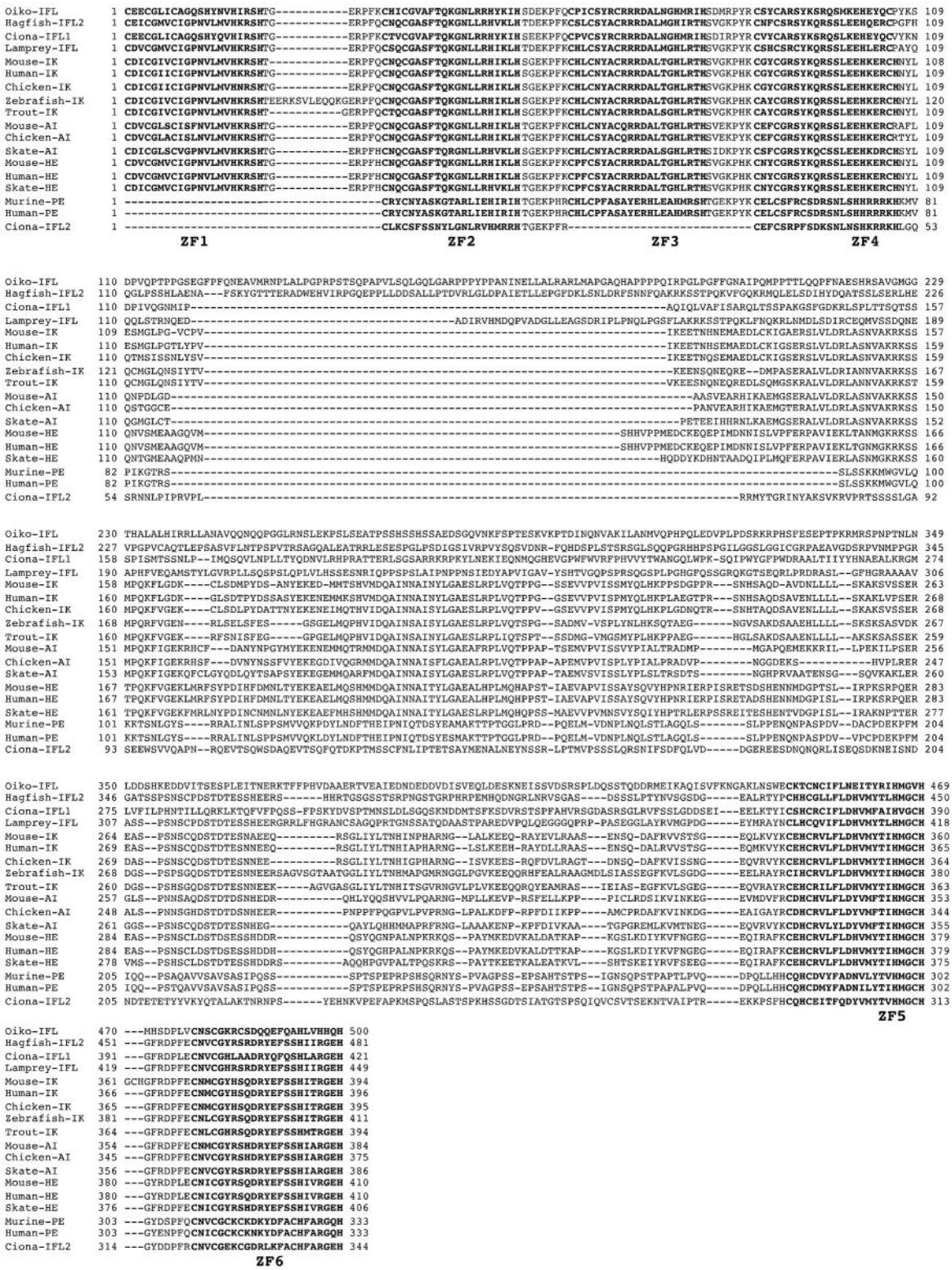

FIGURE 1. Clustal X amino acid alignment of IFL genes from chordates. Accession numbers are as follows: *O. dioica* IFL (AY237106), lamprey (*P. marinus*; AF192380), mouse Ikaros (IK)(L03547), human Ikaros (U40462), chicken Ikaros (Y11833), zebrafish Ikaros (AF092175), trout Ikaros (U92201), mouse Aiolos (AI)(AF001293), chicken Aiolos (AJ005933), skate Aiolos (AF163850), mouse Helios (HE)(AF044257), human Helios (AF130863), skate Helios (AF163847), mouse Pegasus (PE)(AK088289.1), human Pegasus (AF230808). The hagfish IFL is a composite of HIL1 (AY237104) and HIL2 (AY237105). Two potential Ikaros orthologues were identified in the *C. intestinalis* database version 1.0 by searching with the OIL sequence. These sequences have the model identity numbers 144428 (*Ciona* IFL1) and 131901 (*Ciona* IFL2). ZF motifs are in bold and labeled ZF1–6. A dash indicates an alignment gap.

In vertebrates, each of the Ikaros family members is capable of producing alternatively spliced variants that have different DNA-binding potentials as well as dominant-negative forms. HIL1 is the agnathan equivalent to IKx (35), a recently described Ikaros isoform that is selectively expressed in myeloid cells, raising the possibility that HIL1 may be involved in myelopoiesis in hagfish.

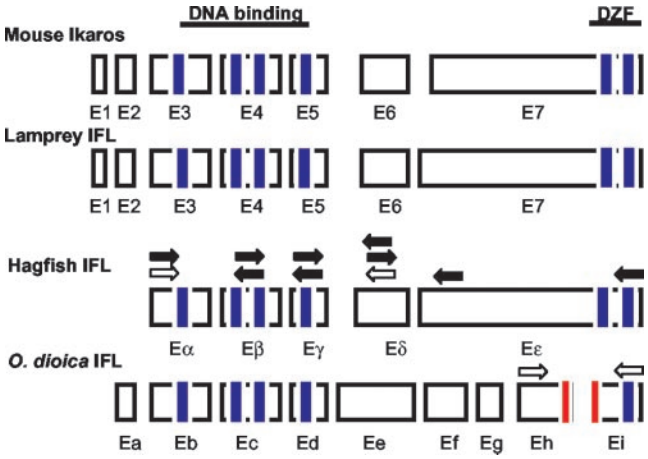

FIGURE 2. A schematic diagram showing a comparison of the intron-exon organization of mouse (18, 19) and lamprey (*P. marinus*) (29) *Ikaros* genes and a partial hagfish and complete *O. dioica* ILF genes. Boxes representing the exons (E) are shown. ZFs are shown in blue. One ZF that is split between exons Eh and Ei in *O. dioica* is shown in red. Filled arrows represent the positions of primers used to define the hagfish intron-exon boundaries by PCR. Open arrows represent the position of primers used to examine the expression of OIL and HIL1 by RT-PCR.

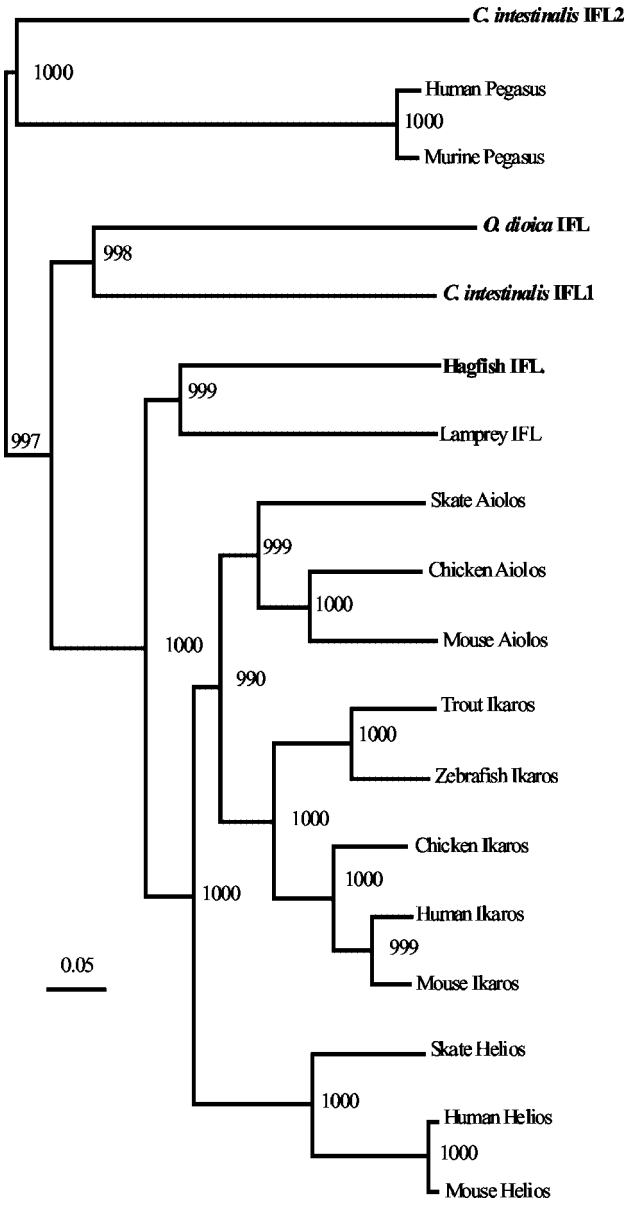

FIGURE 3. Neighbor-joining tree of the Ikaros-like transcription factors. The accession numbers and the alignment used to generate this tree are found in Fig. 1. The numbers at the branch nodes indicate percentage of recovery of the branch in 1000 bootstrap replications.

Intestinal myeloid-like and circulating lymphoid-like cells have been described in hagfish (4). HIL2 is homologous to IK1, the predominant mammalian isoform involved in erythropoiesis and lymphopoiesis. Thus, alternative splicing within the Ikaros family is found in all vertebrates.

Expression of the *OIL* gene (Fig. 4*B*) was determined at different developmental stages of *O. dioica*. Very weak expression was observed in the oocyte. Strongest expression was in early tadpole (2 h posthatching) and at 4/5 days posthatching. As well as the major transcript, there was also a slightly larger transcript amplified, whose strength of expression seemed to mirror that of the OIL transcript. Cloning and sequencing of this fragment revealed that it contained exons Eh and Ei as well as the intervening intron. Whole mount in situ hybridization was then performed on 20 day 3 animals using the full-length OIL cDNA as the probe (Fig. 5) that showed that the anterior Fol cells of the oikoplastic epithelium were strongly positive (yellow). This site is mainly involved in the formation of food concentration filters in the *O. dioica* house. Interestingly, the Fol cells undergo a gene amplification event (polyploidization) that most likely assists in quick manufacturing of the house (36). In mammals, Ikaros has been shown to interact with nucleosome-remodeling deacetylase complexes that include the SNF2-related (sucrose nonfermenting) helicase-ATPase Mi-2 and histone deacetylases (reviewed in Ref. 27). These points raise the possibility that OIL may be involved in chromatin-remodeling events during the development of the *O. dioica* filter house.

HIL1, but not OIL, can bind to the consensus Ikaros target sequence

To examine whether OIL and HIL1 behave as true members of the Ikaros gene family, we examined the DNA-binding potential of OIL and HIL1-GST fusion proteins. Ikaros and its closely related family members each contain up to four N-terminal C2H2 ZFs. Ikaros proteins containing at least three N-terminal ZFs can recognize and bind the consensus Ikaros target site (GGGA) found in the promoter regions of immunologically relevant genes such as TdT, CD3, CD8, and λ5 (37–39). GST fusion proteins containing the N-terminal ZFs in OIL and HIL1 were tested in gel shift mobility assays using a consensus (IKBS4) Ikaros target (32).

The results (Fig. 6, *lane 2*) indicate that the N-terminal ZFs from HIL1 were capable of weakly associating with the probe and that the interaction was specific (Fig. 6, *lane 3*). An interaction was not observed for the OIL GST fusion protein with the consensus target. The murine IK DNA binding region (data not shown) displays a stronger association with the IKBS4 probe, which is most likely due to specific amino acids found in murine Ikaros as compared with HILI. Thus, it appears that although the OIL N-terminal ZFs are similar to that of the true Ikaros clan, they do not possess the same overall binding specificity. This later finding is consistent with the fact that key residues found in the N-terminal ZFs involved in specific base recognition have diverged from the residues found in the vertebrate proteins. Exactly what role OIL plays in the development of *O. dioica* awaits further investigation.

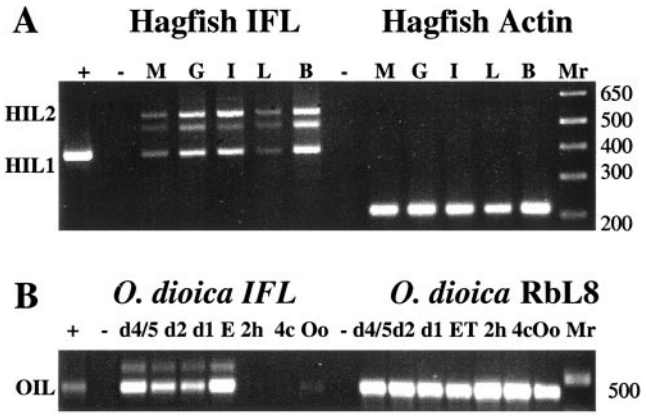

FIGURE 4. Expression patterns of hagfish and *O. dioica* IFL transcripts by RT PCR. Total RNA extracted from tissue obtained from three hagfish and during different developmental stages of *O. dioica* was reversed transcribed and amplified by PCR using HIL1- and OIL-specific primers. Lanes labeled + and − are positive controls (HIL1 or OIL cDNA was used as template) and negative (no template) controls, respectively. Numbers refer to the size, in base pairs, of the marker bands. *A*, Expression of the HIL1 transcript in muscle (M), gill (G), intestine (I), liver (L), and blood (B). The two larger transcripts present were found to contain HIL2 and an Ikaros-like transcript containing in-frame stop codons. Hagfish β actin was used to standardize the amount of mRNA in the initial reaction. This experiment was repeated on two other occasions with the same results. *B*, The expression of the OIL transcript during the life cycle of *O. dioica*. Oocyte (Oc), 2–4 cell stage (4c), 2 h postfertilization (2 h), 2 h posthatching (early tadpole (ET)), day 2 (d2), and day 4/5 (d4/5). Ribosomal protein RbL8 was used to standardize the amount of cDNA in the PCR. This experiment was repeated once with the same result.

HIL1, but not OIL, is capable of both homo- and heterodimerization

Previous studies showed that the C-terminal ZFs of Ikaros represent a bona fide dimerization domain referred to as a DZF domain (33, 40). This domain (Fig. 7*A*) supports homodimerization as well as heterodimerization with the corresponding domains of the closely related family members Aiolos and Helios (22). However, the Ikaros DZF could not form heterodimers with the *Drosophila* Hunchback DZF domain, although this domain supported homodimerization. To determine whether the putative DZF domains from the hagfish HIL1 and *O. dioica* OIL proteins were capable of supporting homodimerization or heterodimerization with murine Ikaros, chemical cross-linking and coimmunoprecipitation assays were performed. For these experiments, the sequences encoding

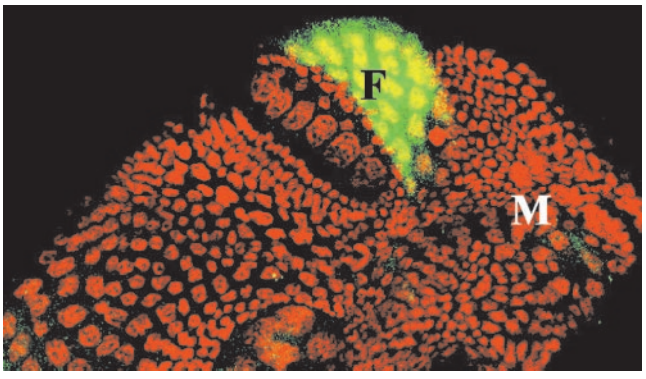

FIGURE 5. Detection of OIL expression (yellow) by in situ hybridization in whole mount preparations of day 3 *O. dioica*. The mouth (M) and anterior part of the Fol (F) region are indicated.

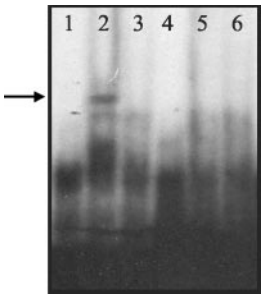

FIGURE 6. HIL1 binds to the consensus Ikaros target site. Gel shift analysis was performed in duplicate with recombinant HIL1 and OIL GST fusion proteins in which the N-terminal ZFs of HIL1 and OIL were fused to GST and then tested for their ability to bind a ^{32}P-labeled probe conforming to the IKBS4 consensus target. *Lane 1*, Free probe (no protein); *lane 2*, HIL1-GST plus probe; *lane 3*, HIL1-GST with probe and cold competitor; *lane 4*, GST plus probe; *lane 5*, OIL-GST plus probe; and *lane 6*, OIL-GST with probe and cold competitor.

the C-terminal fingers from the HIL1 and OIL proteins were substituted for the corresponding sequences of murine Ikaros in the context of mammalian expression plasmids. These expression plasmids encoded small untagged or Flag epitope-tagged proteins after transfection into HEK 293T cells. For the cross-linking assay, extracts from the transfected cells were treated with the cross-linker dithiobis[succinimidyl suberate] for a limited time, followed by Western blot analysis using Abs against an N-terminal domain of Ikaros that is retained in all proteins. The results of this analysis revealed efficient homodimerization of a protein containing the murine Ikaros DZF, but no cross-linked homodimers or heterodimers containing the HIL1 or OIL DZF domains were observed (Fig. 7*B*, and data not shown). These domains were also unable to form detectable heterodimers with the Ikaros DZF (data not shown).

One possible reason for our failure to detect dimers in the cross-linking assay is that the lysine residues that are covalently cross-linked might be missing from the HIL and OIL proteins or might be incorrectly oriented. To address this possibility, a coimmunoprecipitation assay was used. For this assay, untagged and Flag-tagged proteins were coexpressed in HEK 293T cells. Extracts were then analyzed by immunoprecipitation with anti-Flag Abs, followed by Western blot analysis using anti-Ikaros Abs. The results revealed that a tagged protein containing the Ikaros DZF efficiently coimmunoprecipitated untagged proteins containing either the Ikaros DZF or the hagfish DZF (Fig. 7*C*, *lanes 1* and *6*). Furthermore, a tagged protein containing the hagfish DZF coimmunoprecipitated untagged proteins containing either the Ikaros or hagfish DZFs (*lanes 5* and *7*). These results demonstrate that the hagfish protein contains a functional DZF that exhibits the same dimerization specificity as the murine Ikaros DZF (Fig. 7*C*). In striking contrast, a protein containing the putative OIL DZF did not yield a detectable interaction with itself or with the DZF from murine Ikaros (Fig. 7*C*, *lanes 2–4*). The failure of this protein to interact with the murine and hagfish DZF domains is perhaps not surprising, as key residues within the OIL domain diverge from the residues found to be critical for dimerization of the murine DZF (see Fig. 4) (33). Our inability to detect OIL homodimers was unexpected, however. One possibility is that these fingers contribute an alternative function in this primitive protein, with the dimerization function emerging only in vertebrates. This possibility seems unlikely, however, because the DZF domain from *Drosophila* hunchback supports homodimerization. An alternative possibility, which is perhaps more likely, is that the OIL domain does not fold properly when expressed in mammalian cells, and therefore is unable to dimerize. The hypothesis that the OIL DZF does

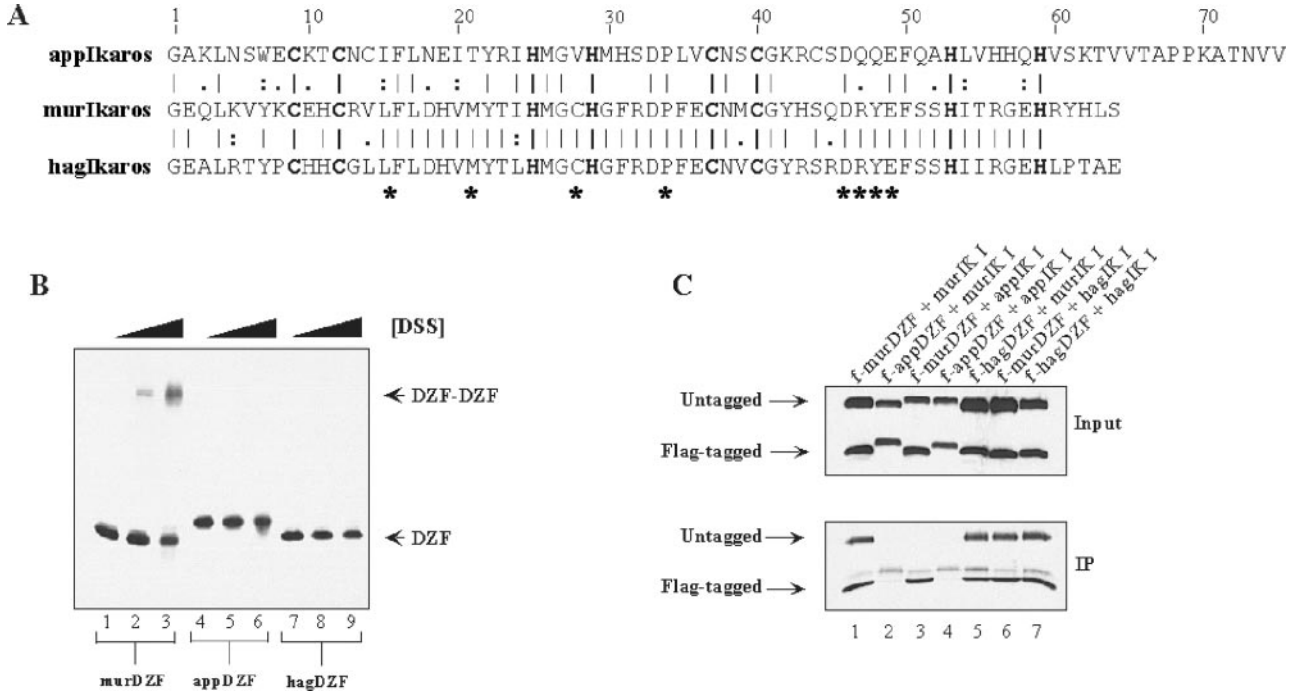

FIGURE 7. Functional conservation of the Ikaros DZF domain is found in the most primitive vertebrate. *A*, Amino acid alignment displaying the DZF regions of OIL, murine Ikaros, and HIL1. The HIL1 and murine DZFs displayed 72% identity, whereas the OIL DZF was only 36% identical with the two vertebrate DZF regions. Critical residues involved in murine DZF dimerization potential are indicated with asterisks (∗). *B*, Chemical cross-linking assay using dithiobis[succinimidyl suberate]. The murine DZF forms typical dimerization pairs (*lanes 1–3*) upon increasing amounts of the chemical cross-linker. DZF-DZF homodimerization was not observed for either the OIL (*lanes 4–6*) or HIL1 (*lanes 7–9*) DZF domains. This assay was performed in duplicate. *C*, Interaction potential for homo- and heterodimerization among the murine, OIL, and HIL1 was assessed (in duplicate) using Flag-tagged and untagged proteins via coimmunoprecipitation. Flag-tagged murine DZF (f-murDZF) can form dimer pairs with both untagged murine (*lane 1*) and HIL1 (hagIK I, *lane 6*), but not with an untagged OIL IKI isoform (appIK I, *lane 3*). Flag-tagged HIL1 DZF (f-hagDZF) dimerizes with both untagged murine (mur Ik I, *lane 3*) and HIL1 (hagIK I, *lane 7*). Flag-tagged OIL (f-appDZF) cannot form dimers with either untagged OIL (appIK I) or murine (murIK I).

not fold properly in mammalian cells is supported by the fact that immunoprecipitation of the F-AppDZF protein was inefficient, despite the fact that it is efficiently expressed (Fig. 7*C*, *lanes 2* and *4*). The inefficient immunoprecipitation suggests that the Flag epitope in the F-AppDZF protein may not be accessible to the Flag Ab, possibly due to improper folding.

Concluding remarks

Gene inactivation experiments have clearly shown the central importance of Ikaros family members in the development of B and T lymphocyte and NK cell and dendritic cell lineages. Specifically, Ikaros has been implicated in the activation of the CD8α gene and potential silencing of the TdT and λ5 genes (37–39) during lymphocyte development. In both fish and higher vertebrates, Ikaros expression coincides with the known sites for hemopoiesis (41, 42). Our present study indicates that Ikaros, which is an essential component for the regulatory network orchestrating the development and maintenance of the adaptive immune system, was in place before the emergence of jawed vertebrates and that this is an ancient gene family. This argument is based upon: 1) the existence of IFL genes in agnathan fish (presented in this study and in Refs. 11 and 29) that have similar gene structures to Ikaros; 2) high sequence conservation of the ZF domains; 3) expression of HIL1 in tissues housing lymphocyte-like cells in hagfish; 4) evidence for alternatively spliced isoforms in both lamprey and hagfish; 5) phylogenetic analysis and the presence of two protochordate IFLs; 6) ability to bind the Ikaros DNA binding sites; and 7) the capacity to homo- and heterodimerize through the DZF domain, all of which are characteristics of true Ikaros family members. Based upon our phylogenetic analysis and biochemical inspection, it appears that the OIL and *Ciona* IFL genes seem to have preserved some of the properties of the ancestral IFL gene. Thus, based upon their presence and lack of evidence for lymphocyte-like cells in protochordates, we suspect that the OIL and CIL gene products most likely have a role outside of any hemopoietic duty, although this awaits further analysis. Future studies will address the role that Ikaros plays in agnathan fish, which to date have been reported to possess lymphocyte-like cells, but lack true components of adaptive immunity, including major histocompatibility Ags, Igs, TCRs, recombination-activating genes, or immunological memory.

References

1. Flajnik, M. F., and M. Kasahara. 2001. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity 13:351.*
2. Klein, J., A. Sato, and W. E. Mayer. 2000. *Major Histocompatibility Complex: Evolution, Structure, and Function.* M. Kasahara, ed. Springer, Tokyo, pp. 3–26.
3. Forey, P., and P. Janvier. 1993. Agnathans and the origin of jawed vertebrates. *Nature 361:129.*
4. Zapata, A., and C. T. Amemiya. 2000. Phylogeny of lower vertebrates and their immunological structures. *Curr. Top. Microbiol. Immunol. 248:67.*
5. Kilarski, W., and B. Plytycz. 1981. The presence of plasma cells in the lamprey (Agnatha). *Dev. Comp. Immunol. 5:361.*
6. Mayer, W. E., T. Uinuk-ool, H. Tichy, L. A. Gartland, J. Klein, and M. D. Cooper. 2002. Isolation and characterization of lymphocyte-like cells from a lamprey. *Proc. Natl. Acad. Sci. USA 99:14350.*
7. Wada, H., and N. Satoh. 1994. Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl. Acad. Sci. USA 91:1801.*
8. Spada, F., H. Steen, C. Troedsson, T. Kallesøe, E. Spriet, M. Mann, and E. M. Thompson. 2001. Molecular patterning of the oikoplastic epithelium of the larvacean tunicate *Oikopleura dioica*. *J. Biol. Chem. 276:20624.*
9. Holland, P. W., J. Garcia-Fernandez, N. A. Williams, and A. Sidow. 1994. Gene duplications and the origins of vertebrate development. *Development (Suppl.): 125.*

10. Pebusque, M. J., F. Coulier, D. Birnbaum, and P. Pontarotti. 1998. Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. *Mol. Biol. Evol. 15:1145*.
11. Haire, R. N., A. L. Miracle, J. P. Rast, and G. W. Litman. 2000. Members of the Ikaros family gene family are present in early representative vertebrates. *J. Immunol. 165:306*.
12. Shintani, S., J. Terzic, A. Sato, M. Saraga-Babic, C. O'huigin, H. Tichy, and J. Klein. 2000. Do lampreys have lymphocytes? The Spi evidence. *Proc. Natl. Acad. Sci. USA 97:7417*.
13. Anderson, M. K., X. Sun, A. L. Miracle, G. W. Litman, and E. V. Rothenberg. 2001. Evolution of haematopoiesis: three members of the PU.1 transcription factor family in a cartilaginous fish, *Raja eglanteria*. *Proc. Natl. Acad. Sci. USA 98:553*.
14. Uinuk-ool, T., W. E. Mayer, A. Sato, R. Dongak, M. D. Cooper, and J. Klein. 2002. Lamprey lymphocyte-like cells express homologs of genes involved in immunologically relevant activities of mammalian lymphocytes. *Proc. Natl. Acad. Sci. USA 99:14356*.
15. Nagata, T., T. Suzuki, Y. Ohta, M. F. Flajnik, and M. Kasahara. 2002. The leukocyte common antigen (CD45) of the Pacific hagfish, *Eptatretus stoutii*: implications for the primordial function of CD45. *Immunogenetics 54:286*.
16. Ting, C. N., M. C. Olson, K. P. Barton, and J. M. Leiden. 1996. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature 384:474*.
17. Urbanek, P., Z. Q. Wang, I. Fetka, E. F. Wagner, and M. Busslinger. 1994. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell 79:901*.
18. Georgopoulos, K., M. Bigby, J. H. Wang, A. Molnar, P. Wu, S. Winandy, and A. Sharpe. 1994. The Ikaros gene is required for the development of all lymphoid lineages. *Cell 79:143*.
19. Wang, J. H., A. Nichogiannopoulou, L. Wu, L. Sun, A. Sharpe, M. Bigby, and K. Georgopoulos. 1996. Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. *Immunity 5:537*.
20. Scott, E. W., M. C. Simon, J. Anastasi, and H. Singh. 1994. Requirement of the transcription factor PU.1 in the development of multiple haematopoietic lineages. *Science 265:1573*.
21. Morgan, B., L. Sun, N. Avitahl, K. Andrikopoulos, T. Ikeda, E. Gonzales, P. Wu, S. Neben, and K. Georgopoulos. 1997. Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. *EMBO J. 16:2004*.
22. Hahm, K., B. S. Cobb, A. S. McCarty, K. E. Brown, C. A. Klug, R. Lee, K. Akashi, I. L. Weissman, A. G. Fisher, and S. T. Smale. 1998. Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. *Genes Dev. 12:782*.
23. Kelley, C. M., T. Ikeda, J. Koipally, N. Avitahl, L. Wu, K. Georgopoulos, and B. A. Morgan. 1998. Helios, a novel dimerization partner of Ikaros expressed in the earliest haematopoietic progenitors. *Curr. Biol. 8:508*.
24. Singh, H. 1996. Gene targeting reveals a hierarchy of transcription factors regulating specification of lymphoid cell fates. *Curr. Opin. Immunol. 8:160*.
25. Clevers, H. C., and R. Grosschedl. 1996. Transcriptional control of lymphoid development: lessons from gene targeting. *Immunol. Today 17:336*.
26. Fisher, A., and M. Merkenschlager. 2002. Gene silencing, cell fate and nuclear organization. Curr. *Opin. Genet. Dev. 12:193*.
27. Georgopoulos, K. 2002. Haematopoietic cell-fate decisions, chromatin regulation and Ikaros. *Nat. Rev. Immunol. 2:162*.
28. Hansen, J. D., P. Strassburger, and L. Du Pasquier. 1997. Conservation of a master hematopoietic switch gene during vertebrate evolution: isolation and characterization of Ikaros from teleost and amphibian species. *Eur. J. Immunol. 27:3049*.
29. Mayer, W. E., C. O'Huigin, H. Tichy, J. Terzic, and M. Saraga-Babic. 2002. Identification of two Ikaros-like transcription factors in lamprey. *Scand. J. Immunol. 55:162*.
30. Durand, C., J. Charlemagne, and J. S. Fellah. 1999. Structure and developmental expression of Ikaros in the Mexican axolotl. *Immunogenetics 50:336*.
31. Liipo, J., and O. Lassila. 1997. Avian Ikaros gene is expressed early in embryogenesis. *Eur. J. Immunol. 27:1853*.
32. Molnar, A., and K. Georgopoulos. 1994. The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. *Mol. Cell. Biol. 14:8292*.
33. McCarty, A. S., G. Kleiger, D. Eisenberg, and S. T. Smale. 2003. Selective dimerization of a C2H2 zinc finger subfamily. *Mol. Cell 11:459*.
34. Perdomo, J., M. Holmes, B. Chong, and M. Crossley. 2000. Eos and pegasus, two members of the Ikaros family of proteins with distinct DNA binding activities. *J. Biol. Chem. 275:38347*.
35. Payne, K. J., G. Huang, E. Sahakian, J. Y. Zhu, N. S. Barteneva, L. W. Barsky, M. A. Payne, and G. M. Crooks. 2003. Ikaros isoform x is selectively expressed in myeloid differentiation. *J. Immunol. 170:3091*.
36. Ganot, P., and E. M. Thompson. 2002. Patterning through differential endoreduplication in epithelial organogenesis of the chordate, *Oikopleura dioica*. *Dev. Biol. 252:59*.
37. Harker, N., T. Naito, M. Cortes, A. Hostert, S. Hirschberg, M. Tolaini, K. Roderick, K. Georgopoulos, and D. Kioussis. 2002. The CD8α gene locus is regulated by the Ikaros family of proteins. *Mol. Cell 10:1403*.
38. Trinh, L. A., R. Ferrini, B. S. Cobb, A. S. Weinmann, K. Hahm, P. Ernst, I. P. Garraway, M. Merkenschlager, and S. T. Smale. 2001. Down-regulation of TDT transcription in $CD4^+CD8^+$ thymocytes by Ikaros proteins in direct competition with an Ets activator. *Genes Dev. 151817*.
39. Sabbattini, P., M. Lundgren, A. Georgiou, C. Chow, G. Warnes, and N. Dillon. 2001. Binding of Ikaros to the λ5 promoter silences transcription through a mechanism that does not require heterochromatin formation. *EMBO J. 20:2812*.
40. Sun, L., A. Liu, and K. Georgopoulos. 1996. Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. *EMBO J. 15:5358*.
41. Hansen, J. D., and A. G. Zapata. 1998. Lymphocyte development in fish and amphibians. *Immunol. Rev. 166:199*.
42. Rumfelt, L. L., E. C. McKinney, E. Taylor, and M. F. Flajnik. 2002. The development of primary and secondary lymphoid tissues in the nurse shark *Ginglymostoma cirratum*: B-cell zones precede dendritic cell immigration and T-cell zone formation during ontogeny of the spleen. *Scand. J. Immunol. 56:130*.