



肺鱼干扰素调节因子 1 基因(*irf-1*)的克隆及其在肉鳍类中的分子进化

甘小妮^{①②③}, 汪登强^②, 王绪祯^②, 陈新文^{①*}

① 中国科学院武汉病毒研究所, 武汉 430071;

② 中国科学院水生生物研究所, 武汉 430072;

③ 中国科学院研究生院, 北京 100049

* 联系人, E-mail: chenxw@wh.iov.cn

2010-12-20 收稿, 2011-02-28 接受

国家重点基础研究发展计划(HCV2009CB522500)资助项目

摘要 肉鳍类是脊椎动物的重要类群, 分为总鳍鱼类和四足类. 其中总鳍鱼类现生的物种包括腔棘鱼和肺鱼. 四足动物则包括两栖类、爬行类、鸟类和哺乳类. 为了比较肉鳍类 *irf-1* 基因的结构特点和系统发育意义, 本研究克隆了肺鱼 *irf-1* 基因 cDNA 全序列, 并与肉鳍类中其他物种的 *irf-1* 基因进行了比对分析. IRF 是与自然免疫相关的蛋白质, 目前已发现了 11 个家族成员, 其中 *irf-1* 和 *irf-2* 是最先被发现的 2 个成员, 在天然免疫中起着重要的作用. 但有关 IRF-1 的报道主要集中在哺乳动物, 其他类群中鲜有报道. 与已报道的 *irf-1* 基因一样, 肺鱼的 *irf-1* 基因可读框的前 345 核苷酸非常保守, 在其 C 端也含有 1 个反式激活区(transactivation domain)和 1 个 IRF 结合域 2 (IAD2)的模体. 虽然肺鱼与四足动物最近共同祖先距今有 417 个百万年之久, 但跨肉鳍类的序列分析显示, IRF-1 氨基酸及其基因序列在肉鳍类中都具有较强的保守性和显著的系统发育意义. 用 IRF-1 氨基酸序列所构建的肉鳍类系统发育与已有报道的结果也是一致的.

关键词

IRF-1 天然免疫
肉鳍类分子
进化系统发育

干扰素(interferon, IFN)是一类多功能的细胞因子, 在动物天然免疫过程中起着重要的作用. 干扰素调节因子(interferon regulation factor, IRF)在干扰素的产生过程中扮演着非常重要的角色, 它最先作为转录因子调节干扰素转录和作为干扰素诱导基因被发现^[1]. 实验进一步发现它还具有其他一些重要的功能, 如抗病毒防御、免疫调节、生长控制^[2]. 迄今为止, 干扰素调节因子家族在哺乳动物中已有 11 个成员被陆续报道. IRF-1 最先被作为人类干扰素- β (IFN- β)基因病毒诱导增强子样元件的调节子而被鉴定的^[3]. 进一步的研究证明, IRF-1 还可以作为 I 型干扰素基因和干扰素诱导基因的转录激活因子^[4]. 随后, 另外一个在结构上与其相似的因子被鉴定并命名为

IRF-2. 实验证明, IRF-2 在干扰素的产生过程中是转录抑制子, 与 IRF-1 有拮抗作用^[5]. 除此之外, IRF-2 还具有其他一些重要的代谢功能, 如在组蛋白 H4, VCAM-1, gp91phox, interleukin-7 转录过程中具转录活性, 是 Class II 的转录催化剂^[6-10]. 在结构上, 干扰素调节因子都具有相似度较高的 DNA 绑定区和变异性较高的 IAD1 或 IAD2 区域. 它的前 115 个氨基酸具较高同源性, 有色氨酸重复特征且包围在 DNA 绑定域周围, 在羧基末端区更具多变性. 除 IRF-1 和 IRF-2 以外, 所有干扰素调节因子家族成员都有 IRF 结合域 1 (IAD1), IAD1 是其家族成员间与其他因子之间相互作用的纽带. 而 IRF-1 和 IRF-2 包含了与其他家族成员不同的结合域(IAD2), IAD2 主要与 IRF-8

英文版见: Gan X N, Wang D Q, Wang X Z, et al. Cloning the interferon regulatory factor 1 gene in lungfish (*Protopterus annectens*) and its molecular evolution among sarcopterygians. Chinese Sci Bull, 2011, 56, doi: 10.1007/s11434-011-4455-6

发生相互作用^[11,12]. 为了比较总鳍鱼类与四足动物中 *irf-1* 基因的结构, 并探讨其系统发育意义, 我们克隆了肺鱼 *irf-1* 基因的 cDNA 序列. 目前已报道的四足动物的系统发育分析主要是基于线粒体基因和少数核糖体基因, 使用功能基因特别是天然免疫相关基因重建肉鳍类的系统发育关系还未见报道. 肉鳍类 *irf-1* 基因也缺乏完整的数据集. 本研究使用简并引物扩增出 *irf-1* 基因的保守片段, 通过 RACE 的方法获得肺鱼的 *irf-1* 基因 cDNA 序列, 完成了四足动物 *irf-1* 基因序列的数据集, 从而获得更加可靠的系统发育结果. 以辐鳍鱼类为外类群的系统发育分析表明该基因具有明显的系统发育意义.

1 材料与方法

(i) 标本采集和鉴定. 非洲肺鱼(*Protopterus annectens*)的标本购自广州花鸟市场, 标本的鉴定在中国科学院水生生物研究所水生生物博物馆进行.

(ii) RNA 提取、cDNA 合成、基因扩增和测序. 总 RNA 提取使用 Trizol 试剂盒(Invitrogen, 美国). 第一链 cDNA 合成使用逆转录 M-MLV 试剂盒(Invitrogen, 美国), 具体操作按试剂盒说明书操作. *irf-1* 基因保守片段扩增和 RACE 所使用的引物见表 1. 首先, 使用简并引物从 cDNA 中扩增 *irf-1* 基因保

守片段. 简并引物(Lf-IRF-1F1 和 Lf-IRF-1R1)的设计比对了已知 *irf-1* 基因序列. PCR 反应的总体积为 50 μ L, 含 100 ng cDNA, 10 $\times$ 缓冲液 5 μ L (TaKaRa, 大连), 引物(10 μ mol/L)各 1 μ L, dNTPs (每种浓度为 2.5 μ mol/L) (TaKaRa, 大连) 2 μ L, *Taq* 聚合酶(TaKaRa, 大连) 2.0 U, 最后补足灭菌双蒸水至终体积. PCR 反应条件为: 94 $^{\circ}$ C 预变性 3 min; 94 $^{\circ}$ C 变性 30 s, 52 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 1 min, 34 个循环; 72 $^{\circ}$ C 延伸 7 min. 每次反应设不含 cDNA 模板的空白对照组. 为了获得 *irf-1* 基因 cDNA 全长, 我们在扩增的保守片段上设计引物. 所设计的引物扩增序列从可读框起始密码子起约 240 bp 的片段. 非洲肺鱼引物序列分别为 Lf-IRF-1F2 和 Lf-IRF-1R2. PCR 反应条件为: 退火温度 54 $^{\circ}$ C, 其他条件同上. 3'RACE 用 Ambion 公司 FirstChoice RLM- RACE Kit (Cat # AM1700)试剂盒. 引物序列见表 1. PCR 反应条件为: 94 $^{\circ}$ C 预变性 3 min; 94 $^{\circ}$ C 变性 30 s, 57 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 1 min 40 s, 35 个循环; 72 $^{\circ}$ C 延伸 7 min. 所有扩增产物经 1.2%的琼脂糖凝胶电泳后用 DNA 凝胶回收试剂盒(Omega 公司)割胶回收, 回收的目的片段用 pMD18-T Vector 试剂盒(TaKaRa, 大连)为载体进行连接, 以 DH5 α (top 10)作为宿主菌进行转化和克隆, 对所得到的克隆进行测序, 引物采用 M13 通用测序引物.

表 1 肺鱼 *irf-1* 基因保守片段扩增和 RACE 所使用的引物

引物名称	引物序列	引物用途
Lf-IRF-1F1	5'-CARATHCCNTGGATGCAYGC-3'	IRF-1 保守区扩增
Lf-IRF-1R1	5'-GCYTTCCKAGTYTTNGGRTC-3'	IRF-1 保守区扩增
Lf-IRF-1F2	5'-ATGCCNGTVGARMGRATGMG-3'	IRF-1 保守区扩增
Lf-IRF-1R2	5'-GCATTCCAAGTATTAGGTTCT-3'	IRF-1 保守区扩增
3'-Lf-IRF-1-RACE outer	5'-GCTTGCATTTTCCAGAGATGGGCG-3'	3'RACE
3'-Lf-IRF-1-RACE inner	5'-CTGGGCGGTACAAACCTGGAGAG-3'	3'RACE

表 2 用于比对和系统发育分析的基因序列

No.	物种	中文名	拉丁名	GenBank 登录号
1	Snakehead	乌鳢	<i>Channa argus</i>	EF067849
2	Mandarin fish	鳊鱼	<i>Siniperca chuatsi</i>	AY647431
3	Grasscarp	草鱼	<i>Ctenopharynodon idellus</i>	GU9970981
4	Human	人	<i>Homo sapiens</i>	NM_002198
5	Sheep	羊	<i>Ovis aries</i>	AF3319701
6	Mouse	鼠	<i>Mus musculus</i>	NM_001159396
7	Rat	大鼠	<i>Rattus norvegicus</i>	NM_012591
8	Chicken	鸡	<i>Gallus gallus</i>	NM_245015
9	Claw frog	爪蛙	<i>Xenopus laevis</i>	NM_001089781
10	West African lungfish	肺鱼	<i>Protopterus annectens</i>	HQ873639

(iii) 序列比对和系统发育分析. 从 GenBank 下载的 *irf-1* 基因序列信息见表 2. 序列比对用 Clustal X 软件, 参数设置使用缺省值. 序列编辑和蛋白质翻译使用 BioEdit 和 SEAVIEW. 系统发育分析使用邻近聚类法(neighborhood joint)、最大似然法(maximum likelihood)和贝叶斯法(Bayesian analysis). 最大似然法利用 Model-test 进行模型和参数的估计. 分支节点的置信度使用 Bootstrap 1000 次检验. 主要使用的分析软件包括 PAUP 10.1, MrBayes 3.1.2, Mega 4.0. 草鱼(*Ctenopharyngodon idellus*)、鳊鱼(*Siniperca chuatsi*)和乌鳢(*Channa argus*)作为辐鳍鱼类的代表和系统发育分析的外类群.

2 结果

2.1 肺鱼 *irf-1* 基因序列特征

克隆测序结果显示, 非洲肺鱼 *irf-1* 基因的 mRNA 含有 1 个包括 942 个核苷酸的可读框, 编码 314 个氨基酸, 3'UTR 区域包含 399 bp 核苷酸, 终止密码子为 TAG (图 1), GenBank 登录号为 HQ873639. 通过比对我们还在非洲肺鱼 IRF-1 氨基酸序列中辨别出 NSL, C 端含反式激活和 IAD2.

2.2 肉鳍类 *irf-1* 基因序列比较

虽然肉鳍类的 *irf-1* 基因序列在不同种类具有不同长度, 但是其可读框却大致相同. 肉鳍类的 *irf-1* 基因序列(哺乳类、鸟类、爬行类、两栖类、肺鱼及作为外类群的辐鳍鱼类)比对后的结果显示, 它的前 345 核苷酸(DBD)具有较高的同源性, 而其余部分(IAD2)呈现较高的变异度. *irf-1* 基因可读框翻译成的氨基酸序列比对结果显示他们具更高的同源性, 特别是前 115 氨基酸所构成的 DNA 绑定域(DBD), 色氨酸重复特征非常明显, 与其余种类的高度一致(图 1). 但在非 DBD 区域, 我们仍然可以辨别出明显的 IAD2 和 C 端含反式激活域的模体.

2.3 分子进化和系统发育分析

肉鳍类与辐鳍类的 IRF-1 氨基酸序列比对分析(图 1)显示, 肺鱼的 *irf-1* 基因序列与四足动物的存在一定差异. 四足动物中爬行类与鸟类较为相似, 而两栖类与肺鱼相似. 尽管如此, *irf-1* 前 115 个氨基酸序列还是具有高度的保守性. 除了 DBD 以外,

我们还在 C 端含反式激活域和 IAD2 上发现许多保守位点.

使用不同方法构建的系统发育树都具有相似的拓扑结构. 邻近聚类法所构建的 IRF-1 氨基酸序列系统发育树如图 2 所示. 当使用 3 种辐鳍鱼类作为外类群时, 可以看出, 肉鳍类是一个单系类群. 肺鱼作为总鳍鱼类的代表, 位于系统树的基部. 系统树的大部分节点都得到很高的重抽样支持率. 四足动物中, 代表两栖类的非洲爪蟾最为原始, 然后依次是爬行类的蜥蜴和鸟类的鸡, 最后是哺乳类. 这些节点大都得到 90% 以上的支持. 基于线粒体细胞色素 *b* 基因序列构建的四足类动物系统发育树与 *irf-1* 基因序列构建的系统树拓扑结构也是相似的.

3 讨论

irf 基因在物种间的变异较大, 外显子不太集中, 另外它们的内含子普遍较长且在不同物种中差异很大. 长期以来, 有关 IRF 因子的报道都集中在哺乳类. 有关鱼类^[13]和低等四足类报道较少. 肺鱼 *irf-1* 基因的克隆测序使我们对总鳍类 IRF-1 有一个完整的数据集. 化石记录和分析表明, 肺鱼与四足动物最近的共同祖先距今 417 百万年^[14,15]. 序列比对的结果显示, 包括非洲肺鱼在内的肉鳍类 IRF-1 氨基酸序列都有较高的相似性. 在如此大的时间尺度上, *irf-1* 基因的可读框能保持这样高的保守性, 说明其二级结构在其功能上的重要性, 也说明以 IRF1 为基础的天然免疫功能起码要追溯到硬骨类起源的初期.

肉鳍类包括总鳍鱼类和四足类. 总鳍鱼类除存在一些化石记录外, 现生的种类中只有空棘鱼和肺鱼. 现生的空棘鱼有 2 种, 肺鱼 3 种, 有关它们的分子系统发育研究未见报道. 化石记录和形态解剖的研究一致表明, 四足动物起源于早期的总鳍鱼类. 最原始的四足动物可能是两栖类的鱼石螈, 而爬行类是从两栖类派生出的一大类群并在侏罗纪得到迅速的发展, 鸟类是由恐龙的一支特化而来. 哺乳类也是爬行动物中另一支系演化而来. 而分子系统发育的证据在鸟类与哺乳类和爬行类的关系上长期没有一致的结果. Xia 等人^[16]、Hedges 等人^[17]和 Bininda-Emonds 等人^[18]都以 18S rRNA 基因序列为基础, 从不同的角度讨论了四足动物的系统发育关系. 基于 *irf-1* 基因序列所构建的系统发育树与他们的结果大致相同, 肺鱼的系统位置在肉鳍类的基部且得到很

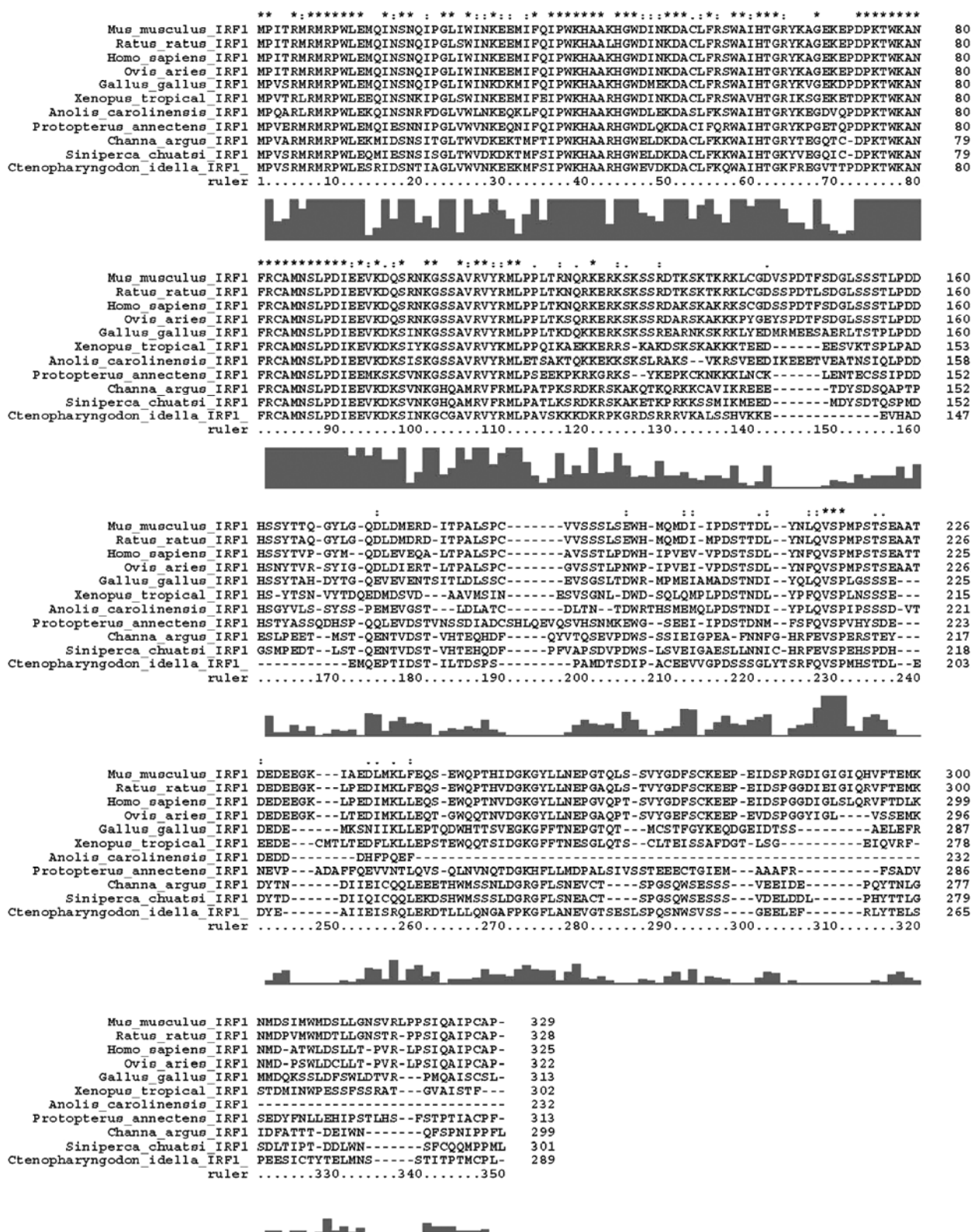


图 1 肉鳍类 IRF-1 氨基酸序列比对

DBD, DNA 结合域(DNA binding domain); IAD, 干扰素相关域(interferon associate domain)

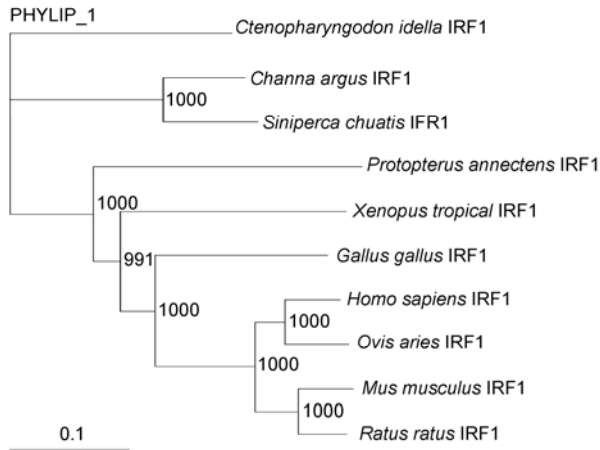


图2 基于IRF-1氨基酸序列构建的肉鳍类系统发育树
遗传距离使用 Kimura 2 parameter. 树的构建使用邻近聚类法, 节点的置信度使用 bootstrap 1000 次重抽样检测. *Protopterus annectens* 为肺鱼

高的重抽样支持率. 此外, 两栖类和爬行类的系统位置相反, 爬行类位于更原始的位置. 从其氨基酸序列比对来看, 蜥蜴的 *irf-1* 基因序列在羧基末端不完整导致可分析序列太短是主要原因, 最后我们在图中删去了蜥蜴的序列. 虽然 *irf-1* 基因除前 345 个碱基较为保守外, 其余部分都具有较高的变异度, 但使用 *irf-1* 全基因序列和 345 个位点的 DBD 保守区所构建的系统发育树大部分的节点均得到较高的支持率, 而且各代表种类间的关系也与过去的研究结果高度一致. 使用功能基因构建较长进化历史类群的系统树比线粒体基因能够获得更高的支持率, 而且推导的氨基酸序列所构建的系统发育树获得了更高的支持率. 值得我们思考的是, 虽然 IRF-1 及其基因行使非常重要的功能, 但其变异点在这样大的尺度上仍然显现明显的系统发育信号.

参考文献

- Mamane Y, Heylbroeck C, Genin P, et al. Interferon regulatory factors: The next generation. *Gene*, 1999, 237: 1–14
- Barnes B, Lubyova B, Pitha P M. On the role of IRF in host defense. *J Interferon Cytokine Res*, 2002, 22: 59–71
- Miyamoto M, Fujita T, Kimura Y, et al. Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifically binds to IFN- β regulatory elements. *Cell*, 1988, 54: 903–913
- Kroger A, Koster M, Schroeder K, et al. Activities of IRF-1. *J Interferon Cytokine Res*, 2002, 11: 5–14
- Harada H, Takahashi E, Itoh S, et al. Structure and regulation of the human interferon regulatory factor 1 (IRF-1) and IRF-2 genes: Implications for a gene network in the interferon system. *Mol Cell Biol*, 1994, 14: 1500–1509
- Vaughan P S, Azia F, Van Wijnen A J, et al. Activation of a cell cycle regulated histone gene by the oncogenic transcription factor IRF-2. *Nature*, 1995, 377: 362–365
- Luo W, Skalnik D G. Interferon regulatory factor-2 directs transcription from the gp91phox promoter. *J Biol Chem*, 1996, 271: 23445–23451
- Jesse T L, LaChance R, Iademarco M F, et al. Interferon regulatory factor-2 is a transcriptional activator in muscle where it regulates expression of vascular cell adhesion molecular-1. *J Cell Biol*, 1998, 140: 1265–1276
- Xi H, Blanck G. The IRF-2 DNA binding domain facilitates the activation of the Class II transactivator (CII TA) type IV promoter by IRF-1. *Mol Immunol*, 2003, 39: 677–684
- Oshima S, Nakamura T, Mamiki S, et al. Interferon regulatory factor-1 (IRF-1) and IRF-2 distinctively up-regulate gene expression and production of interleukin-7 in human intestinal epithelial cells. *Mol Cell Biol*, 2004, 24: 6298–6310
- Schaper F, Kirchhoff S, Posern G, et al. Functional domains of interferon regulatory factor 1 (IRF-1). *Biochem J*, 1998, 335: 147–157
- Meraro D, Hashmueli S, Koren B, et al. Protein-protein and DNA-protein interactions affect the activity of lymphoid-specific IFN regulatory factors. *J Immunol*, 1999, 162: 6468–6478
- Jia W, Guo Q. Gene structures and promoter characteristics of interferon regulatory factor 1 (IRF-1), IRF-2 and IRF-7 from snakehead *Channa argus*. *Mol Immunol*, 2008, 45: 2419–2428
- Zhu M, Yu X. A primitive fish close to the common ancestor of tetrapods and lungfish. *Nature*, 2002, 418: 767–770
- Ahlberg P E, Milner A R. The origin and early diversification of tetrapods. *Nature*, 1994, 368: 507–514
- Xia X H, Xie Z, Kjer K M. 18S ribosomal RNA and tetrapod phylogeny. *System Biol*, 2003, 52: 283–295
- Hedges S B, Moberg K D, Maxson L R. Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Mol Biol Evol*, 1990, 7: 607–633
- Bininda-Emonds O R, Jeffery J E, Sánchez-Villagra M R, et al. Forelimb-hindlimb developmental timing changes across tetrapod phylogeny. *BMC Evol Biol*, 2007, 7: 182