CORRESPONDENCE

Unified nomenclature for the winged helix/forkhead transcription factors

Klaus H. Kaestner,^{1,4,5} Walter Knöchel,^{2,4} and Daniel E. Martínez^{3,4,5}

¹Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104 USA;

²Abteilung Biochemie, Universität Ulm, D-89081 Ulm, Germany; ³Department of Biology, Pomona College, Claremont, California 91711 USA

The winged helix/forkhead class of transcription factors is characterized by a 100-amino-acid, monomeric DNA-binding domain. The structure of the DNA-binding domain of one of the class members, hepatocyte nuclear factor 3 γ (HNF3 γ), in a complex with a DNA target has been solved (Clark et al. 1993). The DNA-binding domain folds into a variant of the helix–turn–helix motif and is made up of three α helices and two characteristic large loops, or “wings.” Therefore, the DNA-binding motif has been named the winged helix DNA-binding domain.

Over the past 9 years since the identification of the first member of this class, the *Drosophila melanogaster* gene *Fork head*, >100 members of this gene family have been identified (for review, see Kaufmann and Knöchel 1996) in species ranging from yeast to human. The rapid accumulation of sequences by many different laboratories has led to the use of multiple names and classification systems, making it very difficult to follow the literature and to name newly characterized winged helix/forkhead transcription factors. This problem was recognized and discussed at the first International Meeting on Forkhead/Winged Helix Proteins, held in La Jolla, California, in November 1998. At that time a proposal was developed to standardize the nomenclature for these proteins. Fox (Forkhead box) was adopted as the unified symbol for all chordate winged helix/forkhead transcription factors. A winged helix/forkhead nomenclature committee was elected to implement this proposal, in consultation with the community at large. This final proposal has been endorsed by >20 scientists¹ as well as the Human and Mouse Gene Nomenclature Committees.

The Fox subclasses

All Fox proteins contain the characteristic 100-amino-acid winged helix domain, that defines this class of tran-

scription factors. Other portions of the Fox proteins, which encode, for instance, transactivation or trans-repression domains, are highly divergent. We have utilized phylogenetic analysis to delineate 15 subclasses for all known chordate Fox proteins. The analysis included chordate sequences obtained from GenBank and sequences submitted directly to the nomenclature committee. The Fox domains of the proteins were aligned using Clustal W (Thompson et al. 1994), and a neighbor-joining tree was generated using PAUP* 4.0 (Swofford 1999); (Fig. 1). This phylogenetic tree will be updated regularly as new sequences are discovered and may be downloaded from <http://www.biology.pomona.edu/fox.html>. A complete phylogenetic analysis of all known forkhead proteins will be published elsewhere (D. Martínez and J.E. Signorovitch, pers. comm.).

Numbering

Fox proteins were assigned to individual subclasses based on the phylogenetic analysis described above. Subclasses were designated by a letter, and within each subclass proteins were given an Arabic numeral. Therefore, the actual name of any Fox protein is “Fox, subclass N, member X”, or for example, Foxd3. Abbreviations for the chordate Fox proteins will contain all uppercase letters for human (e.g., FOXD3); only the first letter capitalized for mouse (e.g., Foxd3); and the first and subclass letters capitalized for all other chordates (e.g., FoxD3). Current assignments for the chordate Fox proteins are listed in Table 1, together with previously used names. Whenever possible we have assigned the same name to ortholog proteins from different species. In a few cases where the phylogenetic affinities were not well resolved by the tree in Figure 1 (e.g., Foxd1 and Foxd2), a within-class phylo-

⁴Fox Nomenclature Committee.

⁵Corresponding authors.

E-MAIL kaestner@mail.med.upenn.edu; FAX (215) 573-5892.

E-MAIL dmartinez@pomona.edu; FAX (909) 621-8878.

¹The following scientists have endorsed the use of this nomenclature system: Frederic G. Barr, William Biggs, Peter Carlsson, James E. Darnell, Sven Enerbäck, Peter Gruss, Brigid Hogan, Andrew D. Hollenbach, Robert Hromas, Tsutomu Kume, Trish Labosky, Eseng Lai, Suzanne C. Li, Naoyuki Miura, Sally A. Moody, Sharon Plon, Hiroshi Sasaki, Günther Schütz, Mathias Treier, Malcolm Whitman, Jeffrey Whitsett, Stella Zanini, and Ken Zaret.

Fox nomenclature system

Figure 1. (Continued on p. 144) Neighbor-joining phylogeny of chordate Fox proteins based on the amino acid sequence of the Fox domain. The distance measure used was mean character difference. The tree was rooted using *Homo* QRF1 (AF086040) as the outgroup. The numbers in the interior branches are bootstrap percentages. For each protein we indicate the organism (genus), the name, the accession number, and the proposed Fox name.

Table 1. *A proposed nomenclature for chordate winged helix proteins*

Anhydrotoma mexicanus	am AFXKH1	FoxA4	U43547	Homo sapiens	hs FREAC9 (FKHL17)	FOXO2	Mus musculus	mm lun	Y12383	FoxF2
Branchiostoma floridae	bf BF1	FoxG1	AF067203	Homo sapiens	hs FREAC10 (FKHL16)	FOXG61b	Mus musculus	mm lun (DNA)	Y12294	FoxF2
Ciona intestinalis	ci Cikh	FoxA5	AF002988	Homo sapiens	hs HBF-1	FOXG61a	Mus musculus	mm MF1	AF045017	FoxF1
Ciona testudinella	ct cox46.1	FoxD6	Z63861	Homo sapiens	hs HBF2	FOXG1c	Mus musculus	mm M2	AF023915	FoxD2
Cynops pyrrhogaster	cp FHD	FoxA4	D64021	Homo sapiens	hs HBF3	FOXG1c	Mus musculus	mm MFH1	S63607	FoxF2
Danio rerio	dr axial	FoxA2	Z22762	Homo sapiens	hs HBF3	FOXG1c	Mus musculus	mm MFH1	X92499	FoxF2
Danio rerio	dr BF1	FoxA3	AF067204	Homo sapiens	hs HFH3	FOX1	Mus musculus	mm MFH1	X74040	FoxF2
Danio rerio	dr Rcd2	FoxA3	AF062246	Homo sapiens	hs HFH4	FOX11b	Mus musculus	mm MFH1	X08222	FoxF2
Danio rerio	dr Rcd3	FoxB1	AF052246	Homo sapiens	hs HFH4	FOXJ1a	Mus musculus	mm MNF	L26507	FoxF1
Danio rerio	dr Rcd4	FoxA5	AF052247	Homo sapiens	hs HFH4	FOXJ1a	Mus musculus	mm MNF	U95016	FoxF1
Danio rerio	dr Rcd5	FoxB1	AF052248	Homo sapiens	hs HFH4 (DNA)	FOXJ1a	Mus musculus	mm PFR	AF060873	FoxF1
Danio rerio	dr Rcd6	FoxD3	AF052249	Homo sapiens	hs HFH11A	FOXK1a	Mus musculus	mm Trident	Y11245	Foxm1
Danio rerio	dr Rcd7	FoxA1	AF052250	Homo sapiens	hs HFH11B	FOXK1b	Mus musculus	mm TWH	X05538	Foxm1
Danio rerio	dr Rcd8	FoxD5	AF052251	Homo sapiens	hs HFKH4	FOXK1c	Mus musculus	mm WHN	X81593	Foxm1
Danio rerio	dr Rcd9	FoxD1	AF052252	Homo sapiens	hs HFKH5	FOXK1d	Mus musculus	mo FH1	AF007895	FoxA5
Danio rerio	dr FKHS	FoxB1	AF052253	Homo sapiens	hs HNF3a	FOXK1e	Mus musculus	mo MocoFH1	AF082902	FoxA5
Danio rerio	dr zFKHR	FoxO5	AF052254	Homo sapiens	hs HNF3a	FOXK1f	Mus musculus	al Merkh1	AB001573	FoxA3
Danio rerio	f-whn	FoxN1	X97021	Homo sapiens	hs ILF1	FOXK1g	Oryzias latipes	al MeHNFS3a	AB001572	FoxA3
Fugu rubripes	fb CBF1	FoxG1	U47275	Homo sapiens	hs ILF2	FOXK1b	Rattus norvegicus	m BF1	M87634	FoxG1
Gallus gallus	gg CBF2	FoxD1	U47276	Homo sapiens	hs ILF3	FOXK1b	Rattus norvegicus	m flk2 (DNA)	AA956527	FoxG1
Gallus gallus	gg gFKH1	FoxC1	Y17794	Homo sapiens	hs INS1	FOXK1c	Rattus norvegicus	m FREAC1 (DNA)	AA817785	FoxG1
Gallus gallus	gg gFKH2	FoxO1	AF114261	Homo sapiens	hs FREAC1 (DNA)	FOXK1c	Rattus norvegicus	m FREAC1 (DNA)	AA819437	FoxG1
Gallus gallus	gg gFKH3	FoxD2	U37272	Homo sapiens	hs TTF2	FOXK1d	Rattus norvegicus	m FH1	L13201	FoxD2
Gallus gallus	gg CWH2	FoxC2	U37273	Homo sapiens	mm AFFXh	FOXK1e	Rattus norvegicus	m FH2	L13202	FoxD3
Gallus gallus	gg CWH3	FoxD3	U37274	Homo sapiens	mm AFFXh	FOXK1e	Rattus norvegicus	m FH4	L136388	FoxD3
Gallus gallus	gg MFH1	FoxC2	U95623	Mus musculus	mm BF1 (Hnf1b)	FOXK1e	Rattus norvegicus	m FH5	L13205	FoxD1
Homo sapiens	hs AF6q2.1	FoxO2	AJ001589	Mus musculus	mm BF2	FOXK1e	Rattus norvegicus	m FH6	L13206	FoxD4
Homo sapiens	hs AF6q2.1 (exon 1)	FoxO2	AJ001590	Mus musculus	mm FAST1	FOXK1e	Rattus norvegicus	m FH7	L13207	FoxE3
Homo sapiens	hs AFX1	FoxO4	Y11284	Mus musculus	mm FAST2	FOXK1e	Rattus norvegicus	m FH8	L13192	FoxD2
Homo sapiens	hs AFX1 (DNA exon 2)	FoxO4	Y11285	Mus musculus	mm FAST2	FOXK1e	Rattus norvegicus	m FH9	L13193	FoxA2
Homo sapiens	hs AFX1 (partial)	FoxO4	U10072	Mus musculus	mm FAST2	FOXK1e	Rattus norvegicus	m HNF3a	L09647	FoxA3
Homo sapiens	hs AFX1 (DNA exon 3)	FoxO4	Y11286	Mus musculus	mm fkl1	FOXK1e	Rattus norvegicus	m HNF3y	L09648	FoxM1
Homo sapiens	hs ALV (PAX3-FKHR)	FoxO1a	U02368	Mus musculus	mm flk2	FOXK1e	Rattus norvegicus	m INS1	X83112	FoxE1
Homo sapiens	hs ALV (PAX3-FKHR)	FoxO1a	U02368	Mus musculus	mm flk2	FOXK1e	Rattus norvegicus	m ITT2	Y11321	FoxE1
Homo sapiens	hs CHEST1	FoxG3	U68723	Mus musculus	mm flk3	FOXK1e	Salmo salar	ss ltk	AF062086	FoxA3
Homo sapiens	hs FAS1	FoxH1	AF076292	Mus musculus	mm flk4	FOXK1e	Xenopus laevis	xl FAD11	U70980	FoxH3
Homo sapiens	hs Fkl1 (DNA)	FoxD3	L12142	Mus musculus	mm flk4	FOXK1e	Xenopus laevis	xl Fhd3	U04872	FoxH3
Homo sapiens	hs FKH H3	FoxA3	L12141	Mus musculus	mm flk5	FOXK1b	Xenopus laevis	xl FKH5	AF064810	FoxB1
Homo sapiens	hs FKH5	FoxB1	AF050580	Mus musculus	mm flk5 (MG)	FOXK1b	Xenopus laevis	xl HNF3a	U04873	FoxA1a
Homo sapiens	hs FKL7	FoxB1	AF071654	Mus musculus	mm flk6	FOXK1e	Xenopus laevis	xl phtalavlis	X65171	FoxA4a
Homo sapiens	hs FKL7	FoxC1a	AF048693	Mus musculus	mm flk6	FOXK1e	Xenopus laevis	xl XBF1	AF101387	FoxG1
Homo sapiens	hs FKL7	FoxC1b	AF078096	Mus musculus	mm FKH10	FOXK1e	Xenopus laevis	xl XFD1	U65751	FoxA4a
Homo sapiens	hs FKLH15	FoxE1	U89995	Mus musculus	mm FKH2	FOXK1e	Xenopus laevis	xl XFD1	U65750	FoxA4b
Homo sapiens	hs FKHR (ALV)	FoxO1a	U02310	Mus musculus	mm FREAC1	FOXK1e	Xenopus laevis	xl XFD1	X74315	FoxF1a
Homo sapiens	hs FKHR	FoxO1a	U37332	Mus musculus	mm FREAC1	FOXK1e	Xenopus laevis	xl XFD2	X74316	FoxF1b
Homo sapiens	hs FKHR	FoxO1a	AF032885	Mus musculus	mm FRKHDA	FOXK1e	Xenopus laevis	xl XFD3	*	FoxA2a
Homo sapiens	hs FKHR (DNA)	FoxO1a	U37335	Mus musculus	mm FRKHDB	FOXK1e	Xenopus laevis	xl XFD3	*	FoxA2b
Homo sapiens	hs FKHRL1	FoxO3a	AF032886	Mus musculus	mm FRKHDC	FOXK1e	Xenopus laevis	xl XFD4	*	FoxC2
Homo sapiens	hs FKHRL1P1 (DNA)	FoxO3b	AF032887	Mus musculus	mm HFH1L	FOXK1e	Xenopus laevis	xl XFD4	*	FoxC2
Homo sapiens	hs FKHRLP1 (DNA)	FoxO3b	AF032888	Mus musculus	mm HFH2	FOXK1e	Xenopus laevis	xl XFD5	*	FoxB2
Homo sapiens	hs FREAC1	FoxF1	U13219	Mus musculus	mm HFH4	FOXK1e	Xenopus laevis	xl XFD6	*	FoxD3a
Homo sapiens	hs FREAC1	FoxF1	AF085343	Mus musculus	mm HFH4 (DNA)	FOXK1e	Xenopus laevis	xl XFD7	*	FoxD3b
Homo sapiens	hs FREAC1 (DNA exon1)	FoxF1	AF085342	Mus musculus	mm HFH5.1	FOXK1e	Xenopus laevis	xl XFD7	*	FoxA1b
Homo sapiens	hs FREAC2	FoxF2	U13220	Mus musculus	mm HFH5	FOXK1e	Xenopus laevis	xl XFD7	*	FoxA1b
Homo sapiens	hs FREAC2	FoxF2	AF084939	Mus musculus	mm HNF3a	FOXK1e	Xenopus laevis	xl XFD8	*	FoxL1
Homo sapiens	hs FREAC2 (DNA exon1)	FoxF2	AF084938	Mus musculus	mm HNF3a	FOXK1e	Xenopus laevis	xl XFD8	*	FoxD1
Homo sapiens	hs FREAC3	FoxF2	AF084939	Mus musculus	mm HNF3a (DNA)	FOXK1e	Xenopus laevis	xl XFD11	*	FoxF1c
Homo sapiens	hs FREAC3	FoxF2	U13221	Mus musculus	mm HNF3a (DNA)	FOXK1e	Xenopus laevis	xl XFD11	*	FoxC1
Homo sapiens	hs FREAC4	FoxD1	U13222	Mus musculus	mm HNF3a (DNA)	FOXK1e	Xenopus laevis	xl XFD12	*	FoxC1
Homo sapiens	hs FREAC4	FoxD1	U98981	Mus musculus	mm HNF3b	FOXK1e	Xenopus laevis	xl XFD12	*	FoxD5a
Homo sapiens	hs FREAC4	FoxD1	U98982	Mus musculus	mm HNF3b	FOXK1e	Xenopus laevis	xl XFD12	*	FoxD5b
Homo sapiens	hs FREAC5	FoxD1	U13223	Mus musculus	mm HNF3y	FOXK1e	Xenopus laevis	xl XFD12	*	FoxD5c
Homo sapiens	hs FREAC6	FoxF1	U13224	Mus musculus	mm HNF3y	FOXK1e	Xenopus laevis	xl XFD13	*	FoxF1
Homo sapiens	hs FREAC7	FoxF1	U13225	Mus musculus	mm HTLF	FOXK1e	Xenopus laevis	xl XFKH1	S93559	FoxA4b
Homo sapiens	hs FREAC8	FoxE3	U42980	Mus musculus	mm lun	FOXK1e	Xenopus laevis	xl FKLIP	**	FoxD5a

Accession nos. submitted to the Fox Nomenclature Committee by W. Knöchel (*) and S.A. Sullivan and S. Molly (**).

Kaestner et al.

genetic analysis was performed using the full sequence of the proteins (data not shown). We have used lowercase letters to distinguish between virtually identical proteins (e.g., Foxa4a and Foxa4b), presumably derived from duplicated genes, a case commonly found in polyploid species like *Xenopus laevis*. Please note that the phylogenetic tree includes several proteins that have not received a Fox designation because as yet, their phylogenetic relationships remain unclear due to limited sequence information.

Naming new sequences

A new Fox protein is defined as a fully sequenced gene, cDNA, or protein that belongs to the Fox family of proteins based on sequence homology of its winged helix DNA-binding domain. We have established a Fox Nomenclature web site (<http://www.biology.pomona.edu/fox.html>) that provides a form for submitting protein sequences to the Fox Nomenclature Committee. We encourage investigators who have discovered new Fox sequences to submit them to the committee for assignment of the proper Fox name. These sequences will be kept confidential until publication of the sequence by the investigators. We recommend that this new system of nomenclature be used in all future publications.

References

- Clark, K.L., E.D. Halay, E. Lai, and S.K. Burley. 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**: 412–420.
- Kaufmann, E. and W. Knöchel. 1996. Five years on the wings of fork head. *Mech. Dev.* **57**: 3–20.
- Swofford, D.L. 1999. *PAUP*: Phylogenetic analysis using parsimony and other methods*. Sinauer, Sunderland, MA.
- Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighing, position specific gap penalties and weight matrix choice. *Nucleic Acid Res.* **22**: 4673–4680.

Unified nomenclature for the winged helix/forkhead transcription factors

Klaus H. Kaestner, Walter Knöchel and Daniel E. Martínez

Genes Dev. 2000, **14**:

Access the most recent version at doi:[10.1101/gad.14.2.142](https://doi.org/10.1101/gad.14.2.142)

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).