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]

[1]Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104 USA; [2]Abteilung Biochemie, Universität Ulm, D-89081 Ulm, Germany; [3]Department 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 (HNF3γ), 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 scientists[1] 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 transcription 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-

[1]The 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 Zannini, and Ken Zaret.

[4]Fox Nomenclature Committee.
[5]Corresponding authors.
E-MAIL kaestner@mail.med.upenn.edu; FAX (215) 573-5892.
E-MAIL dmartinez@pomona.edu; FAX (909) 621-8878.

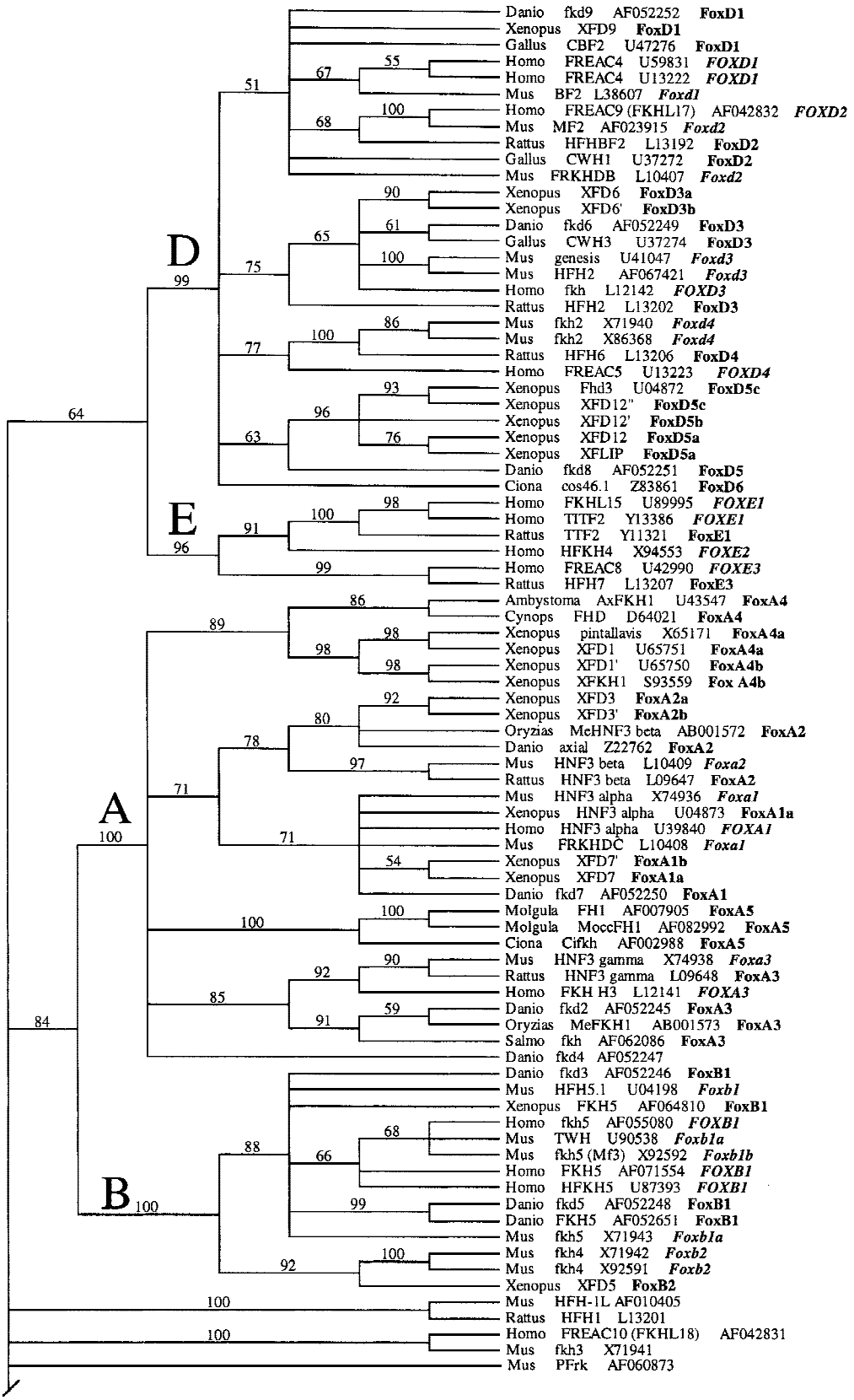

Figure 1. (*Continued on p. 144*) Neighbor-joining phylogeny of chordate Fox proteins based on the amino acid sequence of the Fox domain. The distance measure used was mean character difference. The tree was rooted using *Homo* QRF1 (AF086040) as the outgroup. The numbers in the interior branches are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

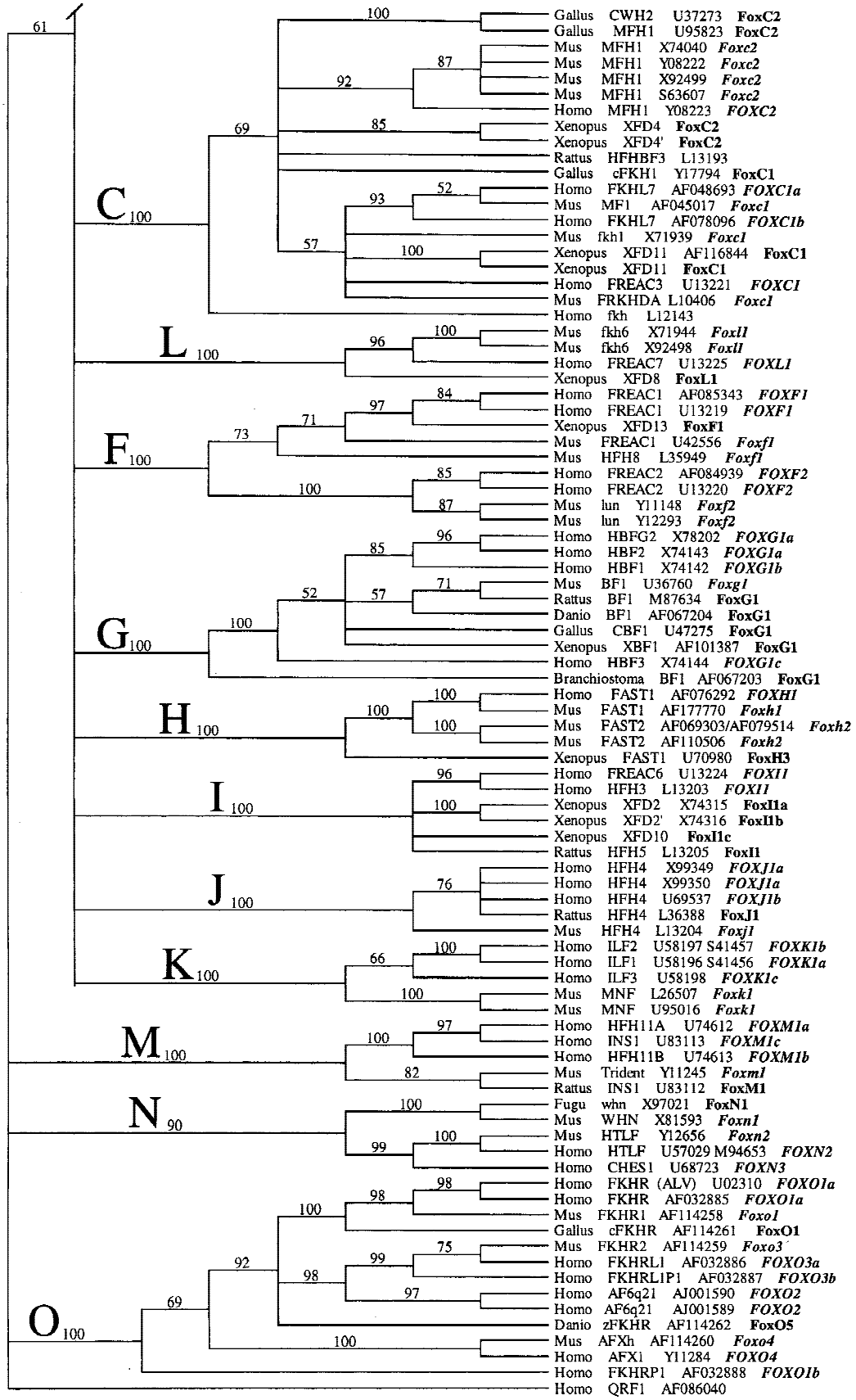
C 100
L 100
F 100
G 100
H 100
I 100
J 100
K 100
M 100
N 90
O 100
Gallus CWH2 U37273 FoxC2
Gallus MFH1 U95823 FoxC2
Mus MFH1 X74040 Foxc2
Mus MFH1 Y08222 Foxc2
Mus MFH1 X92499 Foxc2
Mus MFH1 S63607 Foxc2
Homo MFH1 Y08223 FOXC2
Xenopus XFD4 FoxC2
Xenopus XFD4' FoxC2
Rattus HFHBF3 L13193
Gallus cFKH1 Y17794 FoxC1
Homo FKHL7 AF048693 FOXC1a
Mus MF1 AF045017 Foxc1
Homo FKHL7 AF078096 FOXC1b
Mus fkh1 X71939 Foxc1
Xenopus XFD11 AF116844 FoxC1
Xenopus XFD11 FoxC1
Homo FREAC3 U13221 FOXC1
Mus FRKHDA L10406 Foxc1
Homo fkh L12143
Mus fkh6 X71944 Foxl1
Mus fkh6 X92498 Foxl1
Homo FREAC7 U13225 FOXL1
Xenopus XFD8 FoxL1
Homo FREAC1 AF085343 FOXF1
Homo FREAC1 U13219 FOXF1
Xenopus XFD13 FoxF1
Mus FREAC1 U42556 Foxf1
Mus HFH8 L35949 Foxf1
Homo FREAC2 AF084939 FOXF2
Homo FREAC2 U13220 FOXF2
Mus lun Y11148 Foxf2
Mus lun Y12293 Foxf2
Homo HBFG2 X78202 FOXG1a
Homo HBF2 X74143 FOXG1a
Homo HBF1 X74142 FOXG1b
Mus BF1 U36760 Foxg1
Rattus BF1 M87634 FoxG1
Danio BF1 AF067204 FoxG1
Gallus CBF1 U47275 FoxG1
Xenopus XBF1 AF101387 FoxG1
Homo HBF3 X74144 FOXG1c
Branchiostoma BF1 AF067203 FoxG1
Homo FAST1 AF076292 FOXH1
Mus FAST1 AF177770 Foxh1
Mus FAST2 AF069303/AF079514 Foxh2
Mus FAST2 AF110506 Foxh2
Xenopus FAST1 U70980 FoxH3
Homo FREAC6 U13224 FOXI1
Homo HFH3 L13203 FOXI1
Xenopus XFD2 X74315 FoxI1a
Xenopus XFD2' X74316 FoxI1b
Xenopus XFD10 FoxI1c
Rattus HFH5 L13205 FoxI1
Homo HFH4 X99349 FOXJ1a
Homo HFH4 X99350 FOXJ1a
Homo HFH4 U69537 FOXJ1b
Rattus HFH4 L36388 FoxJ1
Mus HFH4 L13204 Foxj1
Homo ILF2 U58197 S41457 FOXK1b
Homo ILF1 U58196 S41456 FOXK1a
Homo ILF3 U58198 FOXK1c
Mus MNF L26507 Foxk1
Mus MNF U95016 Foxk1
Homo HFH11A U74612 FOXM1a
Homo INS1 U83113 FOXM1c
Homo HFH11B U74613 FOXM1b
Mus Trident Y11245 Foxm1
Rattus INS1 U83112 FoxM1
Fugu whn X97021 FoxN1
Mus WHN X81593 Foxn1
Mus HTLF Y12656 Foxn2
Homo HTLF U57029 M94653 FOXN2
Homo CHES1 U68723 FOXN3
Homo FKHR (ALV) U02310 FOXO1a
Homo FKHR AF032885 FOXO1a
Mus FKHR1 AF114258 Foxo1
Gallus cFKHR AF114261 FoxO1
Mus FKHR2 AF114259 Foxo3
Homo FKHRL1 AF032886 FOXO3a
Homo FKHRL1P1 AF032887 FOXO3b
Homo AF6q21 AJ001590 FOXO2
Homo AF6q21 AJ001589 FOXO2
Danio zFKHR AF114262 FoxO5
Mus AFXh AF114260 Foxo4
Homo AFX1 Y11284 FOXO4
Homo FKHRP1 AF032888 FOXO1b
Homo QRF1 AF086040

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

Ambystoma mexicanus	am AxFKH1	U43547	**FoxA4**	*Homo sapiens*	hs FREAC9 (FKHL17)	AF042832	***FOXD2***	*Mus musculus*	mm lun	Y12293	***Foxf2***
Branchiostoma floridae	bf BF1	AF067203	**FoxG1**	*Homo sapiens*	hs FREAC10 (FKHL18)	AF042831		*Mus musculus*	mm lun (DNA)	Y12294	***Foxf2***
Ciona intestinalis	ci Cifkh	AF002988	**FoxA5**	*Homo sapiens*	hs HBF1	X74142	***FOXG1b***	*Mus musculus*	mm MF1	AF045017	***Foxc1***
Ciona intestinalis	ci cos46.1	Z83861	**FoxD6**	*Homo sapiens*	hs HBF2	X74143	***FOXG1a***	*Mus musculus*	mm Mf2	AF023915	***Foxd2***
Cynops pyrrhogaster	cp FHD	D64021	**FoxA4**	*Homo sapiens*	hs HBF3	X74144	***FOXG1c***	*Mus musculus*	mm MFH1	S63607	***Foxc2***
Danio rerio	dr axial	Z22762	**FoxA2**	*Homo sapiens*	hs HBFG2	X78202	***FOXG1a***	*Mus musculus*	mm MFH1	X92499	***Foxc2***
Danio rerio	dr BF1	AF067204	**FoxG1**	*Homo sapiens*	hs HFH3	L13203	***FOXI1***	*Mus musculus*	mm MFH1	X74040	***Foxc2***
Danio rerio	dr fkd2	AF052245	**FoxA3**	*Homo sapiens*	hs HFH4	U69537	***FOXJ1b***	*Mus musculus*	mm MFH1	Y08222	***Foxc2***
Danio rerio	dr fkd3	AF052246	**FoxB1**	*Homo sapiens*	hs HFH4	X99349	***FOXJ1a***	*Mus musculus*	mm MNF	L26507	***Foxk1***
Danio rerio	dr fkd4	AF052247		*Homo sapiens*	hs HFH4	X99350	***FOXJ1a***	*Mus musculus*	mm MNF	U95016	***Foxk1***
Danio rerio	dr fkd5	AF052248	**FoxB1**	*Homo sapiens*	hs HFH4 (DNA)	X99351	***FOXJ1a***	*Mus musculus*	mm PFrk	AF060873	
Danio rerio	dr fkd6	AF052249	**FoxD3**	*Homo sapiens*	hs HFH11A	U74612	***FOXM1a***	*Mus musculus*	mm Trident	Y11245	***Foxm1***
Danio rerio	dr fkd7	AF052250	**FoxA1**	*Homo sapiens*	hs HFH11B	U74613	***FOXM1b***	*Mus musculus*	mm TWH	U90538	***Foxb1a***
Danio rerio	dr fkd8	AF052251	**FoxD5**	*Homo sapiens*	hs HFKH4	X94553	***FOXE2***	*Mus musculus*	mm WHN	X81593	***Foxn1***
Danio rerio	dr fkd9	AF052252	**FoxD1**	*Homo sapiens*	hs HFKH5	U87393	***FOXB1***	*Molgula oculata*	mo FH1	AF007905	**FoxA5**
Danio rerio	dr FKH5	AF052651	**FoxB1**	*Homo sapiens*	hs HNF3α	U39840	***FOXA1***	*Molgula oculata*	mo MoccFH1	AF082992	**FoxA5**
Danio rerio	dr zFKHR	AF114262	**FoxO5**	*Homo sapiens*	hs HTLF	U57029 M94653	***FOXN2***	*Oryzias latipes*	ol Mefkh1	AB001573	**FoxA3**
Fugu rubripes	fr whn	X97021	**FoxN1**	*Homo sapiens*	hs ILF1	U58196 S41456	***FOXK1a***	*Oryzias latipes*	ol MeHNF3ß	AB001572	**FoxA2**
Gallus gallus	gg CBF1	U47275	**FoxG1**	*Homo sapiens*	hs ILF2	U58197 S41457	***FOXK1b***	*Rattus norvegicus*	m BF1	M87634	**FoxG1**
Gallus gallus	gg CBF2	U47276	**FoxD1**	*Homo sapiens*	hs ILF3	U58198	***FOXK1c***	*Rattus norvegicus*	m fkh2 (DNA)	AA956527	
Gallus gallus	gg cFKH1	Y17794	**FoxC1**	*Homo sapiens*	hs INS1	U83113	***FOXM1c***	*Rattus norvegicus*	m FREAC1 (DNA)	AA817785	
Gallus gallus	gg cFKHR	AF114261	**FoxO1**	*Homo sapiens*	hs MFH1	Y08223	***FOXC2***	*Rattus norvegicus*	m FREAC1 (DNA)	AA819437	
Gallus gallus	gg CWH1	U37272	**FoxD2**	*Homo sapiens*	hs QRF1 (DNA)	AF086040		*Rattus norvegicus*	m HFH1	L13201	
Gallus gallus	gg CWH2	U37273	**FoxC2**	*Homo sapiens*	hs TITF2	Y13386	***FOXE1***	*Rattus norvegicus*	m HFH2	L13202	**FoxD3**
Gallus gallus	gg CWH3	U37274	**FoxD3**	*Mus musculus*	mm AFXh	AF114260	***Foxo4***	*Rattus norvegicus*	m HFH4	L36388	**FoxJ1**
Gallus gallus	gg MFH1	U95823	**FoxC2**	*Mus musculus*	mm BF1 (Hfhbf1)	U36760	***Foxg1***	*Rattus norvegicus*	m HFH5	L13205	**FoxI1**
Homo sapiens	hs AF6q21	AJ001589	***FOXO2***	*Mus musculus*	mm BF2	L38607	***Foxd1***	*Rattus norvegicus*	m HFH6	L13206	**FoxD4**
Homo sapiens	hs AF6q21 (exon 1)	AJ001590	***FOXO2***	*Mus musculus*	mm FAST1	AF177770	***Foxh1***	*Rattus norvegicus*	m HFH7	L13207	**FoxE3**
Homo sapiens	hs AFX1	Y11284	***FOXO4***	*Mus musculus*	mm FAST2	AF069303	***Foxh2***	*Rattus norvegicus*	m HFHBF2	L13192	**FoxD2**
Homo sapiens	hs AFX1 (DNA exon 2)	Y11285	***FOXO4***	*Mus musculus*	mm FAST2	AF079514	***Foxh2***	*Rattus norvegicus*	m HFHBF3	L13193	
Homo sapiens	hs AFX1 (partial)	U10072	***FOXO4***	*Mus musculus*	mm FAST2	AF110506	***Foxh2***	*Rattus norvegicus*	m HNF3ß	L09647	**FoxA2**
Homo sapiens	hs AFX1(DNA exon 3)	Y11286	***FOXO4***	*Mus musculus*	mm fkh1	X71939	***Foxc1***	*Rattus norvegicus*	m HNF3γ	L09648	**FoxA3**
Homo sapiens	hs ALV (PAX3-FKHR)	U02368	***FOXO1a***	*Mus musculus*	mm fkh2	X86368	***Foxd4***	*Rattus norvegicus*	m INS1	U83112	**FoxM1**
Homo sapiens	hs ALV (PAX3-FKHR)	U02308	***FOXO1a***	*Mus musculus*	mm fkh2	X71940	***Foxd4***	*Rattus norvegicus*	m TTF2	Y11321	**FoxE1**
Homo sapiens	hs CHES1	U68723	***FOXN3***	*Mus musculus*	mm fkh3	X71941		*Salmo salar*	ss fkh	AF062086	**FoxA3**
Homo sapiens	hs FAST1	AF076292	***FOXH1***	*Mus musculus*	mm fkh4	X71942	***Foxb2***	*Xenopus laevis*	xl FAST1	U70980	**FoxH3**
Homo sapiens	hs fkh (DNA)	L12142	***FOXD3***	*Mus musculus*	mm fkh4	X92591	***Foxb2***	*Xenopus laevis*	xl Fhd3	U04872	**FoxD5c**
Homo sapiens	hs fkh (DNA)	L12143		*Mus musculus*	mm fkh5	X71943	***Foxb1a***	*Xenopus laevis*	xl FKH5	AF064810	**FoxB1**
Homo sapiens	hs FKH H3	L12141	***FOXA3***	*Mus musculus*	mm fkh5 (Mf3)	X92592	***Foxb1b***	*Xenopus laevis*	xl HNF3α	U04873	**FoxA1a**
Homo sapiens	hs fkh5	AF055080	***FOXB1***	*Mus musculus*	mm fkh6	X71944	***Foxl1***	*Xenopus laevis*	xl pintallavis	X65171	**FoxA4a**
Homo sapiens	hs FKH5	AF071554	***FOXB1***	*Mus musculus*	mm fkh6	X92498	***Foxl1***	*Xenopus laevis*	xl XBF1	AF101387	**FoxG1**
Homo sapiens	hs FKHL7	AF048693	***FOXC1a***	*Mus musculus*	mm Fkh10	AF093879		*Xenopus laevis*	xl XFD1	U65751	**FoxA4a**
Homo sapiens	hs FKHL7	AF078096	***FOXC1b***	*Mus musculus*	mm FKHR1	AF114258	***Foxo1***	*Xenopus laevis*	xl XFD1'	U65750	**FoxA4b**
Homo sapiens	hs FKHL15	U89995	***FOXE1***	*Mus musculus*	mm FKHR2	AF114259	***Foxo3***	*Xenopus laevis*	xl XFD2	X74315	**FoxI1a**
Homo sapiens	hs FKHR (ALV)	U02310	***FOXO1a***	*Mus musculus*	mm FREAC1	U42556	***Foxf1***	*Xenopus laevis*	xl XFD2'	X74316	**FoxI1b**
Homo sapiens	hs FKHR	U37332	***FOXO1a***	*Mus musculus*	mm FRKHDA	L10406	***Foxe1***	*Xenopus laevis*	xl XFD3	*	**FoxA2a**
Homo sapiens	hs FKHR	AF032885	***FOXO1a***	*Mus musculus*	mm FRKHDB	L10407	***Foxd2***	*Xenopus laevis*	xl XFD3'	*	**FoxA2b**
Homo sapiens	hs FKHR (DNA)	U37335	***FOXO1a***	*Mus musculus*	mm FRKHDC	L10408	***Foxa1***	*Xenopus laevis*	xl XFD4	*	**FoxC2**
Homo sapiens	hs FKHRL1	AF032886	***FOXO3a***	*Mus musculus*	mm genesis	U41047	***Foxd3***	*Xenopus laevis*	xl XFD4'	*	**FoxC2**
Homo sapiens	hs FKHRL1P1 (DNA)	AF032887	***FOXO3b***	*Mus musculus*	mm HFH1L	AF010405		*Xenopus laevis*	xl XFD5	*	**FoxB2**
Homo sapiens	hs FKHRP1 (DNA)	AF032888	***FOXO1b***	*Mus musculus*	mm HFH2	AF067421	***Foxd3***	*Xenopus laevis*	xl XFD6	*	**FoxD3a**
Homo sapiens	hs FREAC1	U13219	***FOXF1***	*Mus musculus*	mm HFH4	L13204	***Foxj1***	*Xenopus laevis*	xl XFD6'	*	**FoxD3b**
Homo sapiens	hs FREAC1	AF085343	***FOXF1***	*Mus musculus*	mm HFH4 (DNA)	AF006200	***Foxj1***	*Xenopus laevis*	xl XFD7	*	**FoxA1a**
Homo sapiens	hs FREAC1(DNA exon1)	AF085342	***FOXF1***	*Mus musculus*	mm HFH5.1	U04198	***Foxb1***	*Xenopus laevis*	xl XFD7'	*	**FoxA1b**
Homo sapiens	hs FREAC2	U13220	***FOXF2***	*Mus musculus*	mm HFH8	L35949	***Foxf1***	*Xenopus laevis*	xl XFD8	*	**FoxL1**
Homo sapiens	hs FREAC2	AF084939	***FOXF2***	*Mus musculus*	mm HNF3α	X74936	***Foxa1***	*Xenopus laevis*	xl XFD9	*	**FoxD1**
Homo sapiens	hs FREAC2 (DNA exon1)	AF084938	***FOXF2***	*Mus musculus*	mm HNF3α (DNA)	X76684	***Foxa1***	*Xenopus laevis*	xl XFD10	*	**FoxI1c**
Homo sapiens	hs FREAC3	U13221	***FOXC1***	*Mus musculus*	mm HNF3α (DNA)	X76685	***Foxa1***	*Xenopus laevis*	xl XFD11	AF116844	**FoxC1**
Homo sapiens	hs FREAC4	U13222	***FOXD1***	*Mus musculus*	mm HNF3α (DNA)	X76686	***Foxa1***	*Xenopus laevis*	xl XFD11	*	**FoxC1**
Homo sapiens	hs FREAC4	U59831	***FOXD1***	*Mus musculus*	mm HNF3ß	U04197	***Foxa2***	*Xenopus laevis*	xl XFD12	*	**FoxD5a**
Homo sapiens	hs FREAC4	U59832	***FOXD1***	*Mus musculus*	mm HNF3ß	L10409	***Foxa2***	*Xenopus laevis*	xl XFD12'	*	**FoxD5b**
Homo sapiens	hs FREAC5	U13223	***FOXD4***	*Mus musculus*	mm HNF3ß	X74937	***Foxa2***	*Xenopus laevis*	xl XFD12"	*	**FoxD5c**
Homo sapiens	hs FREAC6	U13224	***FOXI1***	*Mus musculus*	mm HNF3γ	X74938	***Foxa3***	*Xenopus laevis*	xl XFD13	*	**FoxF1**
Homo sapiens	hs FREAC7	U13225	***FOXL1***	*Mus musculus*	mm HTLF	Y12656	***Foxn2***	*Xenopus laevis*	xl XFKH1	S93559	**FoxA4b**
Homo sapiens	hs FREAC8	U42990	***FOXE3***	*Mus musculus*	mm lun	Y11148	***Foxf2***	*Xenopus laevis*	xl XFLIP	**	**FoxD5a**

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

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.