

A Practical Systematic Nomenclature for Kinesin Superfamily Proteins (KIFs) Based on Phylogenetic Classification

Nobutaka Hirokawa, Dept Cell Biol, Grad Sch Med, Univ Tokyo

We have systematically searched for the kinesin superfamily genes using mouse as the model system. We have recently completed the search and genomic information confirmed that now we indeed have identified the complete set of all kinesin genes in mouse and human. From the beginning we have adopted a systematic nomenclature. This system is the de-facto standard for the kinesin superfamily gene names in mammals. For example, our KIF –naming system is adopted by the HUMAN Genome Organization (HUGO), Mouse Genome Sequencing Consortium (MGSC) and at the National Center for Biotechnology Information (NCBI) in the Reference Sequence Collection database (RefSeq). However, kinesin genes in other organisms are named independently and non-systematically, leading to the current confusing situation. To straighten out the situation, here we propose a practical and systematic nomenclature based on phylogenetic classification. At present, only information of the primary sequence can provide a solid, reliable basis for the classification. Therefore, we have classified the sequences of all kinesins available in public databases. A total of 603 kinesins from a wide variety of species were classified into 15 classes based on their phylogenetic relationship. Amino acid sequence analysis was conducted using catalytic core domains and features specific to each class. Based on this phylogenetic classification, we propose a unified, systematic nomenclature system (see Table).

- 1) Unification of each gene name is impractical. It leads to further confusion and loses the continuity from past literature.
- 2) Instead, we propose to cite the class or family name when referring to each family member. Then, we can easily recognize the relation of that kinesin with others within or outside that family.
- 3) For the naming of classes, we see four possibilities.
 - a) Numerical names. Eg. Kinesin I, Kinesin II
 - Initial confusion and conflict between the gene name and class name is inevitable. Eg. A member of class Kinesin I, KIF5B.
 - b) Adopt the founding member name. Eg. BimC family, CENP-E family...
 - Easiest way, but the resulting list of class names lack unity.
 - c) Base nomenclature on model species such as human or mouse. Eg. KIF1 family, KIF2 family ...
 - Only systematic nomenclature that covers all kinesin classes.
 - d) Use one-letter mnemonic system. Eg. K-type Kinesin (for KHC family) or Kinesin B family (for BimC or Bipolar), U Kinesin (for Unc104/KIF1 family) ...
 - A compromise of a) and b) and also systematic, though not to the extent of c).
- 4) I would also like to stress several points that should be considered for a better naming convention.
 - a) Avoid misleading or confusing lettering or naming. Eg. Avoid “I” or Roman numerals.
 - b) Use full length sequence information for classification. Class-specific features are often found outside the catalytic core.

c) Avoid names that have direct functional implication. Eg. Unc=UNCoordinated movement, Chromokinesin, Rab6kinesin etc.

Table. Proposed family names, member examples and numbers.

Mnemonic	KIF-name	Family member examples	Member no.*
U (Unc104)	KIF1	KIF1A, KIF1B, KIF1C, KIF13A, KIF13B, KIF14, KIF16A, KIF16B, GAKIN, UNC104 , UmKIN3, DdUNC104	8/4/2/0/1
M (Middle)	KIF2	KIF2A, KIF2B, KIF2C, KIF24, MCAK, XKCM1, PfkINI	4/3/2/1/1
T (trimeric)	KIF3	KIF3A, KIF3B, KIF3C, KIF17, KRP85, KRP95, KLP64D, KLP68D, OSM3, FLA10	4/3/3/0/0
D (4 th)	KIF4	KIF4A , KIF7, KIF21A, KIF21B, KIF27, Chromokinesin, INVA	5/3/2/3/1
K (KHC)	KIF5	KIF5A, KIF5B, KIF5C, KHC , NKIN, DdK3, DdK5, DdK7	3/1/1/4/3
F (6 th)	KIF9	KIF6 , KIF9, KRP3	2/0/0/0/0
E (CENP-E)	KIF10	KIF10, CENP-E , CMET, CANA, KIP2	1/2/0/14/2
B (BimC)	KIF11	KIF11, Eg5, BIMC , CIN8, KIP1, CUT7	1/1/1/4
L (12 th)	KIF15	KIF12 , KIF15, HKLP2, KLP54D, PAKRPd	2/1/0/6/0
P (Kip3)	KIF19	KIF18A, KIF18B, KIF19A, KIF19B, KLP67A, KIP3	3/2/1/2/0
R (Rab6Kin)	KIF23	KIF20A, KIF20B, Rab6Kinesin , CHO1, MKLP1, PAV, ZEN	5/2/1/0/1
H (HhH ^{**})	KIF22	KIF22, KID, NOD	1/1/0/1/0
V (Vab8)	KIF26	KIF26A, KIF26B, VAB8 , SMY1	2/1/1/2/0
C (KIFC)	KIFC1	KIFC1 , CHO2, NCD, KAR3, KATA, KATB, KATC	1/1/4/4/1
	KIFC2	KIFC2 , KIFC3, KATD, KCBP, KIF25	2/0/1/16/0
X	orphan	CeKLP10, CeKLP18, DdK9	1/0/2/2/1
Total		45/25/19/60/10	

* : Numbers of members in Human/Drosophila/C. elegans/Arabidopsis/Dictyostelium

** : Helix-hairpin-Helix DNA binding motif class 1, a specific feature of this class