Multiplying myosins

Holly V. Goodson*† and Scott C. Dawson‡

**Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556; and ‡Department of Molecular and Cellular Biology, University of California, Berkeley, CA 94720*

Myosins are a diverse family of actin-based molecular motors that appeared early in eukaryotic evolution. Just how early, and how diverse, has begun to become clear from work that appears in this issue of PNAS (1) and recent work from *Nature* (2). For most of its existence, the term "myosin" applied only to the actin-activated ATPase that forms the bipolar thick filaments of muscle and the cytokinetic furrow. This biochemical definition has given way to a bioinformatic one based on presence of a canonical 80-kDa motor domain related to that of other myosins.

Although most myosins have the same general body plan (an N-terminal motor domain, a "neck" consisting of varying numbers of IQ motifs, and a C-terminal tail involved in protein–protein interactions), the nonmotor regions vary enormously. This structural diversity reflects functional diversity (3). Humans have at least 11 different classes of myosin, each with a clearly distinguishable motor domain and a unique constellation of other sequence motifs (3, 4) Other organisms have their own complement of myosins, some similar to those in humans, but many are unique or found in only small groups of species (1, 2, 4).

How does one begin to make sense of such complexity? The now standard approach has been to perform a "phylogenetic analysis," basically, to determine the evolutionary relationships between members of a gene family on the basis of amino acid sequence similarities (homology) between conserved regions (5, 6). The derived relationships are presented visually as a "phylogenetic tree" [see figure 1 of ref. 1]. The evolutionary history reflected in these trees contains a wealth of information. Most practically, it provides a logical structure for classifying and naming the proteins. However, these trees also contain functional information: orthologs (homologs separated by species divergence) are predicted to have similar functions, whereas paralogs (homologs separated by gene duplication) likely have divergent functions. Phylogenetic trees can aid in predicting which homologs exist in a given organism (useful for unsequenced organisms) or inferring which ones existed in ancestral organisms (important for considering the evolution of cell biological processes).

Of course, the utility of a tree depends on its accuracy, and like any scientific endeavor, phylogenetic analyses are subject to interpretation. There are many reasons that a tree can fail to properly reflect the actual path of evolution. One common issue is taxon sampling: trees with more "leaves" (more sequences from more diverse organisms) are often more robust (7). Until now, attempts at comprehensive analyses of the myosin family have focused on sequences from a relatively small number of organisms from the "crown" of the eukaryotic tree (animals, plants, and fungi) (4, 8). Although this group includes most of the common experimental models, it encompasses only a small fraction of the diversity of the eukaryotic world, which is mainly microbial (9, 10).

Foth *et al.* (1) have taken advantage of the recent explosion of genome data from eukaryotic microbes (mainly parasitic protists) to examine the myosin family from a less "crown-centric" perspective. A major goal of this work was to provide information about the complement of myosins in protists. In addition, Foth *et al.* (1) made a significant effort to improve the resolution of the deeper parts of the tree to gain insight into the relationships between myosin classes. To this end, their approaches included adding sequences from recently sequenced crown organisms, using more sensitive phylogenetic inference methods, comparing the results of different methods, and carefully controlling for artifacts generated by the alignment procedure (1).

This thorough analysis comes to several conclusions. First, it provides evidence for six novel myosin classes, increasing the number of named myosin classes from 18 to 24. One new class is in chordate metazoans (XIX), one is in insects (XX), one is in kinetoplastid protozoans (trypanosomes and *Leishmania*), and three are found in alveolate protists (a group including apicomplexan organisms such as malaria and *Toxoplasma*; XXII, XXIII, and XXIV) (see ref. 11 for an overview of eukaryotic relationships). The justification for assigning these groups of proteins to "class status" is that all have been found in the genomes of multiple species, and the motor domains cannot be convincingly assigned to any other class. Some proteins in these new myosin classes contain tail motifs not previously recognized in myosins (FYVE, WW, UBA, ATS1-like, WD40) (see also ref. 2).

Second, their analysis begins to provide hints of the superstructure that must exist in the myosin family (Table 1). They provide compelling statistical evidence from multiple phylogenetic methods that the vertebrate-specific myosin XVIII is a divergent myosin II (conventional myosin). Although not discussed in detail, their tree also provides strong bootstrap support for a relationship between myosin III (a metazoan photoreceptor protein) and vertebrate myosin XVI (a neuronal myosin with N-terminal ankyrin repeats). They find moderate phylogenetic support for the hypothesis that the nematode-specific myosin XII is the myosin XV of the worm lineage, as might be inferred from the strikingly similar tail structures of these proteins (1). They also find strong support for the existence of a group (Class XIV) containing myosins from both apicomplexans and ciliates, including some Tetrahymena sequences that have been suggested previously to establish Class XX (12).

Finally, Foth *et al.* (1) find strong statistical support for a relationship between animal myosin VI (the "backward" myosin) and a group of apicomplexan myosins. This grouping is somewhat surprising, given the apparent lack of tail homology between these proteins and absence of clear members of this group in other organisms. However, the conclusion that these proteins are related appears to be justified. This case appears to be the first time that myosins without detectable tail homology have been placed into the same class.

The structure of the tree hints at additional relationships. Myosins containing the "dilute" tail domain cluster on the same part of the tree, suggesting the existence of a greater dilute supergroup containing myosin V, plant myosin XI, and unclassified *Dictyostelium* and microsporidian myosins. This hypothesis is attractive because it suggests that actin-based intracellular transport is ancient and is driven by a related set of motors. Myosins containing MyTH and FERM domains in their tails (VII, X, XV, and XII; see ref. 13) cluster together with the two *Dictyostelium* MyTH/FERM myosins, supporting the existence of an ancient MyTH/FERM supergroup. Although the relationships implied by these tree topologies are provocative, statistical support for them is

Conflict of interest statement: No conflicts declared.

See companion article on page 3681.

†To whom correspondence should be addressed. E-mail: hgoodson@nd.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0600045103

Table 1. Likely group relationships

Related classes (?)	Phylogenetic support	Common nonmotor element
II, XVIII	Strong	Long coiled-coil
III, XVI	>90% NJ bootstrap	None apparent
VII, X, XII, XV, IV, XIV, XXIII	Weakly grouped to ungrouped	MyTH and/or FERM
V, XI, XIII	Strong to weakly grouped	Dilute motif
VI, apicomplexan VI	Strong	None apparent

Bold type indicates both phylogenetic support (sometimes weak) and common tail element; regular type indicates strong phylogenetic support or possession of common tail element. Additional relationships will likely become apparent as structural and functional characterization continues. For phylogenetic support, see table 3 and figure 3 in ref. 1.

Table 2. Myosin distribution according to Foth *et al.* (1)

Myosin class	Anim.	Fungi	Dict.	Plants	Apic.	Kinet.
I	x	x	x			x
II	x	x	x			
III	x					
IV			*			
V	x	x	?			
VI	x		?		x	
VII	x		?			
VIII				x		
IX	x					
X	*Vt*					
XI				x		
XII	*Ca*					
XIII				*Ab*		
XIV					x	
XV	x					
XVI	*Vt*					
XVII		*As*				
XVIII	x					
XIX	*Ch*					
XX	*In*					
XXI						x
XXII					x	
XXIII					x	
XIV					x	

Additional unclassified sequences have been excluded. Breadth of distribution within a group is indicated either by "x" (wide) or a subgroup designation (distribution apparently limited to that subgroup). Anim., animals; Dict., *Dictyostelium*; Apic., apicomplexans; Kinet., kinetoplastids; *Ab*, *Acetabularia*; *As*, ascomycetes; *Ca*, *Caenorhabditis*; *Ch*, chordates; *In*, insects; *Vt*, vertebrates.
*Class IV myosin is found in the *Dictyostelium* relative *Acanthamoeba*.

weak, so caution is required. This supergroup analysis is consistent with the hypothesis, advanced by Richards and Cavalier-Smith (2), that the ancestral eukaryote contained Myosin I, a dilute myosin, and a MyTH/FERM myosin. The detection of an apparent relationship between myosin VI and apicomplexan myosins raises the possibility that this organism also had myosin VI.

A note on comparison between Foth *et al.* (1) and Richards and Cavalier-Smith (2): these two analyses use a similar set of sequences and an overlapping set of methods. They come to some common conclusions. However, whereas Foth *et al.* (1) focus on using phylogenetic analysis of the myosin motor domain to gain insight into myosins themselves, Richards and Cavalier-Smith (2) investigate myosin evolution to address questions about the evolution of early eukaryotes, basing some of their conclusions on shared tail similarities. Because of these different goals and somewhat different approaches, and because Richards and Cavalier-Smith (2) introduce a unique myosin nomenclature based only on tail domains, it seems likely that the Foth analysis will be of greater utility to those researchers interested in myosins themselves.

Discerning the relationships between supergroups should help in the prediction and analysis of the function of these proteins. However, it also raises the difficult issue of myosin nomenclature: When should a "group" be considered a "class"? Historically, myosin nomenclature has been defined by a very practical test: "if we cannot detect a convincing relationship of a protein to an established myosin class, it must be part of a new class." However, this approach causes several problems, beyond the obvious issues with semantics. First, it raises the risk that the number of myosin classes will simply explode as more genomes, and thus more organism-specific myosins, are identified: several classes of myosin are restricted to one or a small group of organisms (Table 2). Second, how do we address the issue that some classes are now detectably related? Should some class names eventually be decommissioned? Finally, how does one assign a classification to the protistan orthologs of proteins that gave rise to multiple myosin classes in animals (e.g., the *Dictyostelium* and animal myosins with MyTH/FERM tails)?

These issues will require careful consideration over the next few years as the amount of sequence, structural, and functional information grows. It may become necessary to reevaluate myosin nomenclature, as was recently done for the kinesin family (14). Foth *et al.* (1) include an insightful discussion of nomenclature and phylogenetic issues in their supporting online information. We agree with many of their suggestions, in particular that a myosin be given new a class name only if convincing evidence is found for this protein in more than one (preferably divergent) species. Perhaps it would be useful to follow the guidelines used for kinesins, which stipulate that proteins are given subfamily status only if identified in more than one kingdom, and the remaining proteins or groups are classified as "orphans" (14). Some of these orphans would likely gain class status as the tree is filled out. To avoid needless reclassification, we suggest that new classes be proposed only after a comprehensive phylogenomic analysis of myosins from diverse eukaryotes.

Examination of the phylogenetic tree presented in figure 1 of Foth *et al.* (1) raises a final consideration: how conserved is myosin function in diverse eukaryotes? Do all myosins share actin-activated ATPase motor activity? The extreme sequence divergence (evidenced in the disparity in the length of different branches) suggests that the answer is "no." Some myosins (e.g., vertebrate class XVI, animal class XVIII, and the insect branch of class III) have undergone extreme changes in primary sequence in a relatively short span of time. These rapid changes suggest that the motor domain of these proteins must have undergone significant loss or alteration of activity. Discerning the function of these and the myriad of other myosins requires biochemical characterization, which will keep researchers busy for a long time to come.

We thank Meg Titus for helpful insights, discussions, and comments on this manuscript.

1. Foth, B. J., Goedecke, M. C. & Soldati, D. (2006) *Proc. Natl. Acad. Sci. USA* **103,** 3681–3686.
2. Richards, T. A. & Cavalier-Smith, T. (2005) *Nature* **436,** 1113–1118.
3. Kieke, M. C. & Titus, M. A. (2003) in *Molecular Motors*, ed. Schliwa, M. (Wiley–VCH, Weinheim, Germany), pp. 3–44.
4. Berg, J. S., Powell, B. C. & Cheney, R. E. (2001) *Mol. Biol. Cell* **12,** 780–794.
5. Goodson, H. V. & Spudich, J. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 659–663.
6. Cheney, R. E., Riley, M. A. & Mooseker, M. S. (1993) *Cell Motil. Cytoskeleton* **24,** 215–223.
7. Zwickl, D. J. & Hillis, D. M. (2002) *Syst. Biol.* **51,** 588–598.
8. Hodge, T. & Cope, M. J. (2000) *J. Cell Sci.* **113,** 3353–3354.
9. Moreira, D. & Lopez-Garcia, P. (2002) *Trends Microbiol.* **10,** 31–38.
10. Dawson, S. C. & Pace, N. R. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 8324–8329.
11. Adl, S. M., Simpson, A. G., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., Bowser, S. S., Brugerolle, G., Fensome, R. A., Fredericq, S., *et al.* (2005) *J. Eukaryot. Microbiol.* **52,** 399–451.
12. Williams, S. A. & Gavin, R. H. (2005) *Cell Motil. Cytoskeleton* **61,** 237–243.
13. Sousa, A. D. & Cheney, R. E. (2005) *Trends Cell Biol.* **15,** 533–539.
14. Lawrence, C. J., Dawe, R. K., Christie, K. R., Cleveland, D. W., Dawson, S. C., Endow, S. A., Goldstein, L. S., Goodson, H. V., Hirokawa, N., Howard, J., *et al.* (2004) *J. Cell Biol.* **167,** 19–22.