Abstract

Background: Kinesins constitute a superfamily of microtubule motor proteins that are found in eukaryotic organisms. Members of the kinesin superfamily perform many diverse cellular functions such as transport of vesicles and organelles, spindle formation and elongation, chromosome segregation, microtubule dynamics and morphogenesis. Only a few kinesins have been characterized in plants including Arabidopsis thaliana. Because of the diverse cellular functions in which kinesins are involved, the number, types and characteristics of kinesins present in the Arabidopsis genome would provide valuable information for many researchers.

Results: Here we have analyzed the recently completed Arabidopsis genome sequence and identified sixty-one kinesin genes in the Arabidopsis genome. Among the five completed eukaryotic genomes the Arabidopsis genome has the highest percentage of kinesin genes. Further analyses of the kinesin gene products have resulted in identification of several interesting domains in Arabidopsis kinesins that provide clues in understanding their functions. A phylogenetic analysis of all Arabidopsis kinesin motor domain sequences with 113 motor domain sequences from other organisms has revealed that Arabidopsis has seven of the nine recognized subfamilies of kinesins whereas some kinesins do not fall into any known family.

Conclusion: There are groups of Arabidopsis kinesins that are not present in yeast, Caenorhabditis elegans and Drosophila melanogaster that may, therefore, represent new subfamilies specific to plants. The domains present in different kinesins may provide clues about their functions in cellular processes. The comparative analysis presented here provides a framework for future functional studies with Arabidopsis kinesins.
members to proteins in its own genome and the genomes of other organisms can lead to further characterization of the proteins and provide clues to their function. With the information gained from these analyses, experimental procedures could be designed to determine the functions of the proteins. Budding yeast has 113 conventional protein kinase genes [5]. Using multiple alignment and parsimony analysis of protein kinase catalytic domain sequences, the yeast protein kinases were categorized into subfamilies based on structural relatedness. This information led to identification of yeast specific kinases and gave hints about the function of unknown kinases. An analysis of proteins containing zinc finger domains in bacteria, yeast and C. elegans revealed that bacteria do contain some proteins that bind zinc but lack large families of zinc-binding proteins found in eukaryotes [6]. Between eukaryotes, the presence and size of the different zinc-binding families vary [6]. In a comparative analysis of the genomes of Drosophila, C. elegans and S. cerevisiae the number of proteins for each of the 200 most frequently occurring protein domains was identified for each organism [7]. Among these domains were various calcium-binding domains such as the EF-hand family. In Drosophila there are 130 proteins with this domain while in C. elegans there are 79 and only 16 in yeast. A BLASTP and TBLASTN search of the Drosophila genome with members of the cytoskeleton protein families identified 262 genes with moderate to completely convincing homology to cytoskeletal genes [8]. New members of the families were discovered and some proteins that are present in other genomes were missing in the Drosophila genome.

Kinesins constitute a superfamily of microtubule (MT) motor proteins found in all eukaryotic organisms. Members of the kinesin superfamily have a highly conserved motor domain. The first kinesin was identified in squid giant axons as a protein involved in transport of vesicles [9,10]. The conventional kinesin is a tetramer with two heavy chains and two light chains. The kinesin heavy chains (KHCs) contain the motor domain with ATPase activity, a central coiled-coil region and a tail that binds the light chains. Historically, proteins with homology to KHCs but falling in different subfamilies have been called kinesin-like proteins. However, KHCs are now recognized as a subfamily of the kinesin superfamily and all members of the superfamily are referred to as kinesins. We have followed that pattern in this paper. KHCs have been identified in fungi and animals and a large number of other kinesins from other subfamilies have been identified in all eukaryotes [11,12,13]. All kinesins have a domain with homology to the motor domain of KHC but little sequence similarity outside of this domain. Some kinesins have a coiled-coil region but others do not. The tail domain, which is believed to interact with specific cargoes, is nonconserved. Kinesins bind MTs and a variety of cargoes and perform force-generating tasks such as transport of vesicles and organelles, spindle formation and elongation, chromosome segregation and MT organization [12,13,14,15,16,17,18,19,20]. The motor domain of KHC is in the N-terminal region but in other kinesins it can also be located in the C-terminus or internally. Motility assays have been performed with a number of the kinesins. The C-terminal domain kinesins have been shown to have minus-end motility while the others have plus-end motility [21,22,23,24]. Nine subfamilies of kinesins have been identified by phylogenetic analysis using the conserved motor domain [25]. Not all kinesins fall into one of the nine subfamilies and may represent additional subfamilies.

The first two plant kinesins identified were found in tobacco pollen tube (pollen kinesin homologue, PKH) and tobacco phragmoplast (tobacco kinesin related protein, NtKRP125) [26,27,28,29]. Another kinesin isolated more recently from tobacco pollen tubes has also been characterized at the biochemical level [26]. Kinesins that have been characterized at the molecular level in plants include, NtKRP125 in tobacco [30], four Arabidopsis kinesins identified by PCR-based cloning (KatA, KatB, KatC, KatD) [31,32,33,34], KCBP (kinesin-like calmodulin binding protein) found in Arabidopsis, tobacco, potato, and maize [35,36,37,38], PAKRP in Arabidopsis [39] and DeKRP120-1 and DeKRP120-2 in carrot [40].

With the complete sequencing of the Arabidopsis genome, it has become possible to search the Arabidopsis database with the conserved motor domain of kinesins to identify the kinesins encoded in the Arabidopsis genome. We have identified 61 kinesin-like genes in Arabidopsis and their general location on the five Arabidopsis chromosomes. Surprisingly, the Arabidopsis genome contains the largest number of kinesins among all eukaryote genomes that have been sequenced. We have further analyzed the predicted protein sequences for the presence of domains that might lead to an understanding of their function. By definition all have a kinesin motor domain and several have a coiled-coil domain that might indicate dimerization. A phylogenetic analysis of the 61 Arabidopsis kinesin motor domain sequences with 113 other motor domain sequences has revealed that there are Arabidopsis kinesins that fall into seven of the nine recognized subfamilies of kinesins. Some do not fall into any family and there are some subfamilies unique to Arabidopsis and maybe to plants.

**Results and Discussion**

**Identification of Arabidopsis kinesins**

The recent completion of the Arabidopsis genome has allowed the analysis of the first plant genome for kinesins.
Using the sequence of the conserved motor domain of kinesins database searches were performed using BLASTP and TBLASTN at TAIR (The Arabidopsis Information Resource, [http://www.arabidopsis.org/blast/]). The sequences obtained from TAIR were compared to the kinesin-like sequences identified by the Munich Information Centre for Protein Sequences (MIPS). Analysis of both these databases resulted in identification of 61 unique sequences that contain a kinesin motor domain as identified by the SMART ([www.smart.embl-heidelberg.de/]) program (1Table I). In our analysis of these databases, we found some discrepancies between the two. For example, some kinesins in TAIR are not identified as kinesins in the MIPS database. All 61 sequences had BLASTP scores greater than or equal to 87 and E values equal to or less than 2e-017 and the next identified sequence had a score of only 34 and an E value of 0.13. A TBLASTN search did not identify any other kinesins nor did searches with a representative member of each of the subfamilies of kinesins. In comparison, S. cerevisiae, S. pombe, C. elegans and Drosophila have 6, 9, 19 and 24 kinesins respectively (Figure 1A). Arabidopsis has the highest percentage (0.24%) of the total number of genes as compared to S. cerevisiae and S. pombe with 0.1% and 0.17% respectively, C. elegans with 0.11% and Drosophila with 0.18% (See Figure 1A).

Only six of the 61 Arabidopsis kinesins have been reported in the literature [12]. The other 55 sequences obtained from the Arabidopsis Database or from MIPS have been sequenced as part of the Arabidopsis Genome Sequencing Project. The sequences are, therefore, predicted sequences that have not been verified by complete cDNAs. The three AtKRP125 kinesins (AtKRP125a,b,c; 1Table I) show homology to the kinesin (NtKRP125) isolated from phragmoplasts of tobacco [30]. AtKRP125b has 68% identity with NtKRP125 over the 1000 residues they have in common (NtKRP has 56 additional residues). The average sequence length of the Arabidopsis kinesins is just over 1000 residues with the shortest sequence prediction being 425 (AtF15A18.10) amino acids and the longest are 2158 (AtMGD8.20) and 2756 (AtK13E13.17). Besides the above two sequences, no other predicted sequence is over 1400 amino acids. Some of the intron/exon predictions may not be correct which could reduce or increase the size of predicted proteins in the databank and so the sizes may change as more characterization is done for each kinesin. A case in point is the sequence that was isolated by Lee, et al. [39] for AtPAKRP which is 1292 amino acids while the predicted protein has 1662 residues with very few substitutions in these residues. V was often substituted by I as is also found in non-plant kinesins. Three Arabidopsis kinesins did not have a highly conserved NXXSSRSH sequence that is completely conserved in other kinesins. The predicted amino acid sequence of AtMAA21.110 has only part of the ATP binding site and lacks the NXXSSRSH domain. However, analysis of the genomic sequence indicates that the genes contain the coding region for the conserved domains but they are not present in the deduced sequence due to inaccurate prediction of introns. Isolation of the cDNA for this clone is needed to determine the correct sequence.

Using the Arabidopsis Sequence Map Overview of TAIR, the location of each kinesin was determined (Figure 2). The kinesins are distributed throughout the genome. In four cases, pairs of kinesins were sequenced in the same clone (F19H22, F14P13, F3K23 and F15H). One to 12 other predicted proteins in the clones separate the members of the pairs. Interestingly, two pairs of clones show closest homology to a member of the other pair, AtF3K23.6-AtF19H22.50 and AtF3K23.14-AtF19H22.150 (Figure 3) which suggests that this is a result of gene duplication. Analysis of the total Arabidopsis genome revealed that a whole genome duplication occurred followed by subsequent gene loss and extensive local gene duplications [4]. The duplicated segments represent 58% of the Arabidopsis genome. The S. cerevisiae genome has also had a complete ancient genome duplication and 30% of the genes form duplicate pairs. Duplicated genes account for 48% of the total genes of C. elegans and Drosophila [7].

Phylogenetic analysis
Using ScSMY1, a highly divergent kinesin [25], as an outgroup, the motor domain sequences of the 61 Arabidopsis kinesins and 113 kinesins from other organisms were analyzed for phylogenetic relationships using PAUP [44] (Figures 3 and 4). In other organisms nine subfamilies of kinesins have been identified by phylogenetic analysis using the conserved motor domain [25]. Functional studies with members of three of the subfamilies (KHC,
KRP85/95, and Unc104/KIF1) indicate that they are involved in transport [20]. Members from the other subfamilies (C-terminal, Kip3, MKLP1, BimC, chromokinesin/KIF4, and MCAK/KIF4) have been shown to function in nuclear movement, chromosome segregation, spindle formation and stability, and other cytoskeletal processes associated with cell division [20]. Seven of the nine families are represented in Arabidopsis. However, several Arabidopsis kinesins do not fall into any of the nine subfamilies and are likely to represent additional subfamilies that are unique to plants (Figure 1B, 3 and 4). Most of the Arabidopsis kinesins are more closely related to another Arabidopsis or another plant kinesin than to any other kinesin used in the comparison.

A comparison of the five eukaryotic genomes that have been sequenced shows that not all organisms have all types of kinesins (Figure 1B). The subfamilies that are involved in transport are underrepresented in Arabidopsis. There are no members of the KRP85/95 or Unc104/KIF1 subfamilies in Arabidopsis while Drosophila and C. elegans both have a few members in each of these families. C. elegans does not have a Kip3-type kinesin, and yeasts do not have KHC, MKLP1, chromokinesin, or MCAK/KIF2 representatives. The C-terminal and BimC subfamilies are the only ones having at least one representative in each of the five organisms (Figure 1B). Each organism has ungrouped kinesins. However, Arabidopsis has a larger number (24 out of 61) than in any of the other sequenced organisms (Figure 1B).

The motor domain in kinesins is located either in the N terminus, C terminus or in the middle of the protein. Members of BimC subfamily, which are N-terminal plus-end motors, are present in all five eukaryotic organisms that have been sequenced. The three NtKRP125-like Arabidopsis kinesins (AtKRP125a, b, and c) and AtF16L2.60 are grouped with the tobacco homologue in the BimC subfamily which are involved in cross-linking and antiparallel sliding of MTs (Figure 3) [20].

Twenty-one Arabidopsis kinesins fell into the C-terminal subfamily (Figure 3). This is an unusually large number compared to the other organisms. C. elegans has five and S. pombe has two while Drosophila and S. cerevisiae have only one (Fig. 1). It is also unusual because phylogenetically they group with other C-terminal proteins but, structurally, 11 have internal motors and five have N-terminal motors. The internal kinesins have a motor domain that is closer to the C-terminus than the N-terminus but each has some sequence C-terminal to the motor domain and could be called internal depending on the parameters used to define an internal motor. Kat D, At30B22.20 and At32N15.10 which have earlier been classified as C-terminal kinesins [25] and six other Arabidopsis kinesins form a subgroup within the C-terminal family (See Fig. 3). The other two kinesins (AtF15A18.10 and AtT9N14.6) having an internal motor which fall into the C-terminal subfamily form a subgroup with one of the kinesins with an N-terminal motor (AtF15A18.10). These three are most closely related to a group of animal C-terminal kinesins including HsKIFC3 and MMFIFC2. The other four kinesins with N-terminal motors that fall into the C-terminal subfamily form a subgroup that is most closely related to three C. elegans C-terminal kinesins.
As the phylogenetic tree was based on the motor domain alone, it will be interesting to find out the direction of movement of these kinesins whose motor domains are N-terminal or internal but are most closely related to the C-terminal subfamily. Kinesins in the C-terminal subfamily translocate toward the minus-end of MTs [45]. The sequence responsible for the direction of movement was recently determined [23,46,47]. C-terminal kinesins have a conserved sequence at the neck/motor core junction (Fig. 5A) [47]. Endow and Waligora [47] determined that the GN residues at the neck/motor core junction (Fig. 5A) are necessary for minus end directed movement. Examination of the neck/motor core junction of the 21 C-terminal class Arabidopsis kinesins shows conservation of these residues in most of the C-terminal Arabidopsis kinesins.

KCBP, a C-terminal calcium/calmodulin-regulated kinesin, forms a distinct group within the C-terminal subfamily (Figure 3). A few of the C-terminal Arabidopsis kinesins (e.g., KatA, KCBP) have been localized to mitotic MT arrays (spindle, spindle poles, and phragmoplast) [33,48], suggesting a role for these kinesins in cell division. The motor activity of only two Arabidopsis C-terminal kinesins has been demonstrated experimentally and both, as expected of C-terminal motors, showed minus-end motor activity [33,49].
Figure 3
3 and 4 Phylogenetic tree. The tree shown above was built from a kinesin motor domain sequence alignment using the heuristic search method of PAUP v4.0b6, a maximum parsimony program, with random stepwise addition and tree bisection-reconnection (TBR). The tree is the consensus tree from 68 trees built from 100 replicates. It is arbitrarily rooted using ScSmy1 as an outgroup. Vertical dashes indicate ungrouped kinesins (light dashes - those grouped with other kinesins, bold dashes - exclusively Arabidopsis kinesins). The Arabidopsis kinesins are in bold. Kinesins from the following organisms were used: An, Aspergillus nidulans; Bm, Bombyx mori; Ce, Caenorhabditis elegans; Cf, Cylindrotheca fusiformis; Cg, Cricetulus griseus; Cr, Chlamydomonas reinhardtii; Dd, Dictyostelium discoideum; Dm, Drosophila melanogaster; Gg, Gallus gallus; Hs, Homo sapiens; Lc, Leishmania chagasi; Lm, Leishmania major; Lp, Loloipepali; Mm, Mus musculus; Ms, Morone saxatilis; Nh, Nectria haematococca; N, Nicotiana tabacum; St, Solanum tuberosum; Rn, Rattus norvegicus; Sc, Saccharomyces cerevisiae; Sp, Strongylocentrotus purpuratus; Spo, Schizosaccharomyces pombe; Sr, Syncephalastrum racemosum; Um, Ustilago maydis; Xi, Xenopus laevis.
Figure 4

3 and 4 Phylogenetic tree. The tree shown above was built from a kinesin motor domain sequence alignment using the heuristic search method of PAUP v4.0b6, a maximum parsimony program, with random stepwise addition and tree bisection-reconnection (TBR). The tree is the consensus tree from 68 trees built from 100 replicates. It is arbitrarily rooted using ScSmy1 as an outgroup. Vertical dashes indicate ungrouped kinesins (light dashes - those grouped with other kinesins, bold dashes - exclusively Arabidopsis kinesins). The Arabidopsis kinesins are in bold. Kinesins from the following organisms were used: An, Aspergillus nidulans; Bm, Bombyx mori; Ce, Caenorhabditis elegans; Cf, Cylindrotheca fusiformis; Cg, Cricetulus griseus; Nr, Chlamydomonas reinhardtii; Dd, Dicyostelium discoideum; Dm, Drosophila melanogaster; Gg, Gallus gallus; Hs, Homo sapiens; Lc, Leishmania chagasi; Lm, Leishmania major; Lp, Loligo pealii; Mm, Mus musculus; Ms, Morone saxatilis; Nc, Neurospora crassa; Nh, Nectria haematococca; Nt, Nicotiana tabacum; St, Solanum tuberosum; Rn, Rattus norvegicus; Sc, Saccharomyces cerevisiae; Sp, Strongylocentrotus purpuratus; Spo, Schizosaccharomyces pombe; Sr, Synechocystis racemosum; Um, Ustilago maydis; Xi, Xenopus laevis.
Figure 5
Alignments of neck/motor core region and kinesin light chain binding site of KHC. Alignments were done using the Clustal method in DNA STAR MEGALIGN. A. Alignment of the neck/motor core regions from the 21 kinesins falling in the C-terminal subfamily. White letters on black are identical residues, white on dark gray are strongly similar and black on light gray are weakly similar. Asterisks mark the two residues shown to confer minus end directed movement [47]. B. Alignment of the kinesin light chain binding site in KHCs. The small letters indicate the heptad positions in the heptad repeats as given by Diefenbach et al. [52]. Hs, human KHC; Hsn, human neuronal KHC; Sp, sea urchin KHC, Dm, Drosophila melanogaster KHC; Nc, Neurospora crassa KHC; Um, Ustilago maydis, At, AtMAA21.110 kinesin.
The KHC subfamily (conventional kinesins) is made up of two subgroups, one animal and one fungus. To date, no KHC has been found in any plant. The phylogenetic tree indicates that there is possibly one KHC-type kinesin in Arabidopsis. AtMAA21.110 falls into the KHC group with a closer relationship to KHCs found in fungi (Figure 3). Both Drosophila and C. elegans have one KHC whereas the yeast genomes do not have any KHCs (Figure 1B). Some kinesin light chain sequences have been predicted in the Arabidopsis genome but none have been studied experimentally. Experimental data suggests that fungal conventional kinesins do not have light chains [50,51]. The KHC binding site for kinesin light chain proteins has been identified [52]. The binding site consists of four highly conserved contiguous heptad repeats which are predicted to form a tight α-helical coiled-coil interaction with the heptad repeat-containing N-terminal of the light chain. The tail domain of AtMAA21.110, four KHCs from vertebrates and invertebrates and two fungi were aligned and compared for the presence of this region (Fig. 4). The four vertebrate and invertebrate KHCs have a very conserved sequence in this region while the fungal sequences are conserved in respect to these KHCs in only a few short stretches. The AtMAA21.110 tail shows little similarity and in fact, has a very short tail that does not show similarity to any other kinesin (data not shown). However, the binding domain is a coiled coil and the tail of AtMAA21.110 does have a region of coiled-coil from residues 377-415. As stated above, there are problems with the predicted amino acid sequence of AtMAA21.110 and further work needs to be done in order to determine if this is a KHC homolog.

One Arabidopsis kinesin groups with the MKLP1 subfamily. Two internal motor Arabidopsis kinesins group with the MCAK/KIF2 subfamily (also called the internal family) which have members involved in vesicle transport, chromosome movement and MT catastrophe [20]. Three Arabidopsis kinesins fall into a group with Kip3 subfamily members in which ScKip3 is involved in nuclear movement [53]. Three Arabidopsis kinesins form a branch off of the chromokinesin/KIF4 subfamily members, some of which are involved in vesicle transport (HsKIF) and spindle organization and chromosome positioning (XIKlp1) [20,25].

Several Arabidopsis kinesins show some similarity to other ungrouped kinesins (Figures 3 and 4). The ungrouped centromeric proteins (HsCENPE and UmKin1) cluster with seven Arabidopsis kinesins. CENP-E binds to the kinetochore throughout mitosis and to MTs of the spindle mid-zone during late stages of mitosis [20]. One Arabidopsis kinesin is paired with HsKid, an ungrouped kinesin. HsKid is a kinesin-like DNA-binding protein that is involved in spindle formation and the movements of chromosomes during mitosis and meiosis [20]. Arabidopsis PAKRP along with five other Arabidopsis kinesins are grouped with XIKlp2. XIKlp2 is required for centrosome separation and maintenance of spindle bipolarity [54] whereas PAKRP associates with the phragmoplast [39] and is expected to function in cytokinesis. Three Arabidopsis kinesins form a subgroup separate from any other kinesin but share a branch with CeLF22F4 and DmNOD. Eight other Arabidopsis kinesins form a subgroup separate from kinesins of any other organism. In many cases a group of Arabidopsis kinesins forms a separate branch within the major subgroup in which they fall. Since there are many processes unique to plants, it is likely that these novel kinesins function in plant-specific processes.

**Domain analysis**

Analysis of Arabidopsis kinesins with domain analysis programs identified different domains in many of the kinesins (Figure 6 and (1Table I). In addition to the motor domain, most kinesins have a coiled-coil region. Only three Arabidopsis kinesins do not contain some coiled-coil domain (AtT1E22.130, AtMG6L.9, AtT20H2.17). However, as can be seen in Fig. 5, some have very short coiled-coil domains (i.e. AtKRP125b, AtKatD, T9C5.240). AtKatD was reported [34] to lack a coiled-coil domain but the SMART program identified a small coiled-coil domain starting just inside the C-terminal end of the motor domain.

A few Arabidopsis kinesins have predicted domains that are not present in other kinesins. The SMART predicted domains were listed as either confidently predicted based on the E-value or less significant than the required threshold. The confidently predicted domains may provide some clues about the possible functions of plant kinesins. While the predictions that are below the significant threshold may not be reliable, some consideration of them may be warranted.

Some of the confidently predicted domains have a connection with actin cytoskeleton in some way. KCBP has MyTH4 (1.15e-49) and talin-like domains (6.34e-34) present in some myosins, suggesting that it has domains of both MT- and actin-based motors. Such motors may be involved in cross talk between MT and actin cytoskeleton [12]. KCBP also has been shown to have a calmodulin-binding domain [35,55,56]. KCBP is unique among kinesins in having these domains. However, a sea urchin kinesin, kinesin C, has also been reported to have a calmodulin binding domain but not the MyTH4 and talin-like domains of KCBP [57]. Six of the Arabidopsis kinesins have a calponin homology domain (2.88e-49 to 2.89e-10), which is an actin-binding domain present in the N-termini of spectrin-like proteins. The CH domain is a
Proteins with a tandem pair of CH domains cross-link F-actin, bundle actin or connect intermediate filaments to cytoskeleton. Proteins with a single copy are involved in signal transduction [58,59]. Perhaps the kinesins containing a CH domain bind actin and are involved in signal transduction or linking of actin and MTs.

Some domains that are involved in protein-protein or protein-DNA interactions have been identified. Three Arabidopsis kinesins with armadillo repeats (tandem repeats that form a superhelix of helices) form a group on a separate branch from other kinesins except for a C. elegans kinesin that does not have an armadillo domain. The E-values range from $4.84 \times 10^{-3}$ to $1.51 \times 10^{-1}$ and are above the significant threshold. Kinesin-interacting protein KAP3 that has armadillo repeats is thought to mediate cargo binding as part of a heterotrimeric complex with two kinesins in the KRP85/95 subfamily [18]. One kinesin (AtMRO11.5) has a helix-hairpin-helix DNA binding domain ($2.48 \times 10^{-00}$) and also a nuclear localization signal. This kinesin may be involved in signaling.

Domains with less significant E-values include six kinesins which contain spectrin repeats ($1.91 \times 10^{00}$ to $2.97 \times 10^{01}$). There is some evidence which suggests that spectrin facilitates Golgi membrane association with motor proteins, including cytoplasmic dynein, kinesin and myosin [60,61]. Perhaps these kinesins may be related to motor functions involving the Golgi complex. In some cases the coiled-coil and/or SPEC repeats overlapped with basic region leucine zipper domains (AtZCF125, AtT15B3.190, AtK13E13.17) and/or homeobox associated leucine zipper domains (AtK13E13.17, AtMCA23.6).

Other domains with less significant E-values include domains that are associated with the Rho family of GTPases, the RhoGEF domain ($1.49 \times 10^{01}$), HR1 domain ($3.09 \times 10^{01}$ and $3.35 \times 10^{01}$) and FH2 domain ($1.23 \times 10^{01}$). Rho GTPases control cellular processes including cytoskeletal reorganization and transcriptional activation [62]. A RhoGEF protein was isolated that not only activates RhoA but also directly interacts with MTs [63]. The HR1 repeat has been identified as a binding site for Rho [64]. The FH2 domain ties to both Rho and the actin cytoskeleton. FH proteins control rearrangements of the actin cytoskeleton and members of this family have been found to interact with Rho-GTPases. Kinesins having these domains may be involved in actin-MT cooperation during processes such as vesicle and organelle transport, spindle rotation and nuclear migration [65].

**Figure 6**
Schematic diagram of all Arabidopsis kinesins. Motor domain and coiled-coil domains are marked in red and green, respectively. CH, calponin homology domain; MyTH4, myosin tail homology domain; Talin-like, talin-like domain; CBD, calmodulin binding domain; ARM, armadillo/beta-catenin-like repeats; HhH1, helix-hairpin-helix domain. Bar = 100 aa.
**Why do plants have so many kinesins?**

Why do plants have so many kinesins? What is the function of each of these kinesins in plant growth and development? There are many MT-associated processes that are unique to plants [12]. For example, during cell division in plants several plant-specific MT arrays such as the preprophase band and the phragmoplast are formed that are important in determining the future location of the cell wall and cell wall formation, respectively. These unique processes are likely to require additional plant-specific motors. In addition, centrosomes play an important role in MT organization in animals whereas plants have no well-defined centrosomes. Hence, MT organization and dynamics in plants may also require additional MT motors. Also, in plants there is cell to cell transport of macromolecules such as RNA through plasmodesmata and such transport may also require MT motors [12]. It is possible that kinesins may participate in functions other than transport, and MT dynamics and organization [20]. Functional redundancy may also explain the large number of kinesins. Three different *S. cerevisiae* kinesins (KAR3, KIP2, and KIP3) are involved in nuclear positioning [66,67,68]. The loss of one gene alone does not alter viability but double mutants are lethal. Each has a particular role in positioning but lack of one step is not essential whereas loss of more leads to nonviability [20]. Six subfamilies of kinesins have been shown to be involved in some aspect of cell division [12,20]. Three Arabidopsis kinesins, Kat A, KCBP and PAKRP, have been shown to be associated with the phragmoplast [39,48,69] and so are expected to be necessary in some way for cytokinesis. Immunolocalization and microinjection studies have shown the involvement of KCBP in cell division [48,70]. However, KCBP mutants, (*ZWICHEL*), grow normally with no apparent defects in cell division except that they contain abnormal trichomes [71,72]. It is likely that KCBP function in cell division in *ZWICHEL* mutants is compensated by other C-terminal kinesin(s).

**Conclusions**

In summary, Arabidopsis has a surprisingly large number of kinesins among the five completed eukaryotic genomes. Many Arabidopsis kinesins do not fall into any known subfamilies of kinesins and several Arabidopsis kinesins are not present in yeast, *C. elegans* and *D. sophila* and are likely to represent new subfamilies specific to plants. Further analyses of kinesins have resulted in identification of several interesting domains in Arabidopsis kinesins that provide clues in understanding their functions.

Although the functions of most of the Arabidopsis kinesins remains to be determined, phylogenetic analysis of kinesins and identification of functional domains in these proteins provide clues to their function which can be tested empirically. Several knockout mutant libraries obtained by T-DNA insertions are available to screen for mutations in kinesins ([www.arabidopsis.org](http://www.arabidopsis.org)). An alignment of the motor domains of the Arabidopsis kinesins revealed a few very conserved stretches of amino acids in the motor domain that could be used for designing a universal degenerate primer set to screen for mutations in all kinesins. Once an insertion of T-DNA is detected in a kinesin, the sequence of the amplified product could be obtained to identify the kinesin involved. Due to the redundancy in function that has been seen with the non-plant kinesins, other strategies such as overexpression of the kinesins may also be needed. Protein-protein interaction studies using the yeast two-hybrid trap and expression analysis of all kinesins in different tissues and cell types with microarrays can also provide valuable information about the function of the kinesins.

**Methods**

**Identification of Arabidopsis kinesin-like proteins**

The motor domain sequence of AtKCBP (a plant kinesin) was used to do BLAST™ Similarity Searches at TAIR ([http://www.arabidopsis.org/blast/](http://www.arabidopsis.org/blast/)). Both BLASTP and TBLASTN searches were done. The sequences showing homology to the motor domain were obtained from the *Arabidopsis thaliana* Database (AtDB) and analyzed using SMART (Simple Modular Architecture Research Tool, [http://smart.embl-heidelberg.de/](http://smart.embl-heidelberg.de/)) which identifies putative domains within the sequence. Sequences containing a kinesin motor domain were identified as kinesin-like proteins. The sequences from TAIR were also compared to the proteins containing kinesin-like motor domains as listed at MIPS ([http://www.mips.biochem.mpg.de/proj/taic/db/index.html](http://www.mips.biochem.mpg.de/proj/taic/db/index.html)). Motor domains from other subfamilies were also used in BLAST™ searches to identify any possible kinesins missed using the AtKCBP motor domain sequence. The subsequence containing the kinesin motor domain was delineated by the SMART program was generated from each kinesin and compared to all other Arabidopsis motor domain sequences to see if the kinesin was unique. In some cases, two groups had sequenced the gene and the predicted intron/exon status was not the same. We included the sequence with the closest homology to other kinesin motor domains and eliminated the other sequence. Domains were identified using the SMART program. Location of the genes on the chromosome was identified using the Arabidopsis Sequence Map Overview of TAIR ([http://www.arabidopsis.org/cgi-bin/maps/Schrom](http://www.arabidopsis.org/cgi-bin/maps/Schrom)).

**Phylogenetic analysis**

The motor domains of the Arabidopsis kinesins and 113 kinesins from animals and yeast with the myosin motor domain of ScMMY1 as an outgroup were aligned using
the Clustal method in the DNA STAR program MEGALIGN. The alignment was saved as a PAUP file and used to generate a phylogenetic tree using PAUP 4.0b6. A heuristic search method with tree-bisection-reconnection branch swapping and one hundred replicates was used. The tree used in Figures 3 and 4 is the consensus tree of 68 trees retained in the search. The score of the best trees was 15849.

Note Added in Proof

It has been recently reported that there are 45 kinesin genes in the human genome (H Miki, M Setou, K Kanehiro, N Hirokawa; All kinesin superfamily protein, KIF, genes in mouse and human. Pro Natl Acad Sci USA 2001 98:7004-7011). It is estimated that there are 35,000 - 45,000 genes in the human genome. Kinesin genes make up 0.1 to 0.13% of the total number of human genes whereas they represent 0.24% of Arabidopsis genes (61 out of 25,498 predicted Arabidopsis genes are kinesins). Among all sequenced organisms arabidopsis has the highest number of kinesins.

Additional material

Table I

Table I

[http://www.biomedcentral.com/content-supplementary/1471-2164-2-2-s1.doc]

Acknowledgments

We thank Drs. V.S. Reddy and Maxim Golovkin for their comments on the manuscript. This work was supported by a grant from the National Science Foundation to ASNR.

References

1. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, et al: Life with 6000 genes Science 1996, 274:563-567
2. TCeS Consortium: Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium Science 1998, 282:1202-1218
3. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, AAronasdes PG, Scherer SE, et al: The genome sequence of Drosophila melanogaster Science 2000, 287:2185-2195
4. TAG Initiative: Analysis of the genome sequence of the flowering plant Arabidopsis thaliana Nature 2000, 408:796-815
5. Hunter T, Flowman GD: The protein kinases of budding yeast: six score and more Trends Biochem Sci 1997, 22:18-22
6. Clarke ND, Berg JM: Zinc fingers in Caenorhabditis elegans: finding families and probing pathways [see comments] Science 1998, 282:2018-2022
7. Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR, Harsharan IK, Fortini ME, et al: Comparative genomics of the eu-karyotes Science 2000, 287:2204-2215
8. Goldstein LS, Gunawardena S: Flying through the drosophila cytoskeletal genome J Cell Biol 2000, 150:F63-8
9. Vale RD, Reese TS, Sheuz MP: Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility Cell 1985, 42:39-50
10. Brady ST: A novel brain ATPase with properties expected for the fast axonal transport motor Nature 1985, 317:73-75
11. Bloom GS, Endow SA: Motor proteins 1: Kinesins Protein profile 1994, 1:1059-1116
12. Reddy ASN: Molecular motors and their functions in plants Intl Rev Cytol & Cell Biol 2001, 204:97-178
13. Leopold PL, McDowall AW, Pfister KK, Bloom GS, Brady ST: Association of kinesin with characterized membrane-bounded organelles Cell Motil Cytoskeleton 1992, 23:19-33
14. Sawin KE, LeGuillerc K, Philippe M, Mitchison TJ: Mitotic spindle organization by a plus-end-directed microtubule motor Nature 1992, 359:540-543
15. Barton NR, Goldstein LSB: Going mobile: Microtubule motors and chromosome seggregation Proc Natl Acad Sci USA 1996, 93:1735-1742
16. Wein H, Foss M, Brady B, Cande WZ: DSK1, a novel kinesin-related protein from the diatom Cylindrotheca fusiformis that is involved in anaphase spindle elongation J Cell Biol 1996, 133:595-604
17. Raich WB, Moran AN, Rothman JH, Hardin J: Cytokinesis and midzone microtubule organization in Caenorhabditis elegans require the kinesin-like protein ZEN-4 Mol Cell Biol 1998, 18:207-219
18. Manning BD, Snyder M: Drivers and passengers wanted! The role of kinesin-associated proteins Trends in Cell Biol 1999, 10:281-289
19. Moore JD, Endow SA: Kinesin proteins: a phylum of motors for microtubule-based motility BioEssays 1998, 18:207-219
20. Goldstein LSB, Philip AV: The road less traveled: Emerging principles of kinesin motor utilization Annu Rev Cell Dev Biol 1999, 15:141-183
21. McDonald HB, Steward RJ, Goldstein LSB: The kinesin-like ncd protein of Drosophila is a minus end-directed microtubule motor Cell 1990, 63:1159-1165
22. Walker RA, Salmon ED, Endow SA: The Drosophila claret segregation protein is a minus-end-directed motor molecule Nature 1990, 347:780-782
23. Endow SA: Determinants of motor protein directionality Nature Cell Biol 1999, 1:E163-E167
24. Kuriyama R, Kofron M, Essner R, Kato T, Dragas-Granic S, Omoto CK, Khodjakov A: Characterization of a minus end-directed kinesin-like motor protein from cultured mammalian cells J Cell Biol 1995, 129:1049-1059
25. AJ Kim, SA Endow: A kinesin family tree J Cell Sci 2000, 113:3681-3682
26. Cai G, Bartalesi A, Del Casino C, Moscatelli A, Tiezzi A, Cresti M: The kinesin-immunoreactive homologue from Nicotiana tabacum pollen tubes: Biochemical properties and subcellular localization Planta 1993, 191:496-506
27. Tiezzi A, Moscatelli A, Cai G, Bartalesi A, Cresti M: An immunoreactive motor resembling Arakidopsis thaliana kinesin in Nicotiana tabacum pollen tubes Cell Motil Cytoskeleton 1992, 21:132-137
28. Asada T, Sonobe S, Shibaoka H: Microtubule translocation in the cytokinetic apparatus of cultured tobacco cells Nature 1991, 350:238-241
29. Asada T: A kinesin-related motor associated with plant specific microtubule systems Plant Cell Physiol 1996, 37 Suppl.:S08
30. Asada T, Kuriyama R, Shibaoka H: TKRP125, a kinesin-related protein involved in the centrosome-independent organization of the cytokinetic apparatus in tobacco BY-2 cells J Cell Sci 1997, 110:179-189
31. Mitsui H, Yamaguchi-Shinozaki K, Shinozaki K, Nishikawa K, Takahashi H: Identification of a gene family (kat) encoding kinesin-like proteins in Arabidopsis thaliana and the characterization of secondary structure of KatA Mol Gen Genet 1993, 238:362-368
32. Mitsui H, Nakatani K, Yamaguchi-Shinozaki K, K, Shinozaki, K, Nishikawa, H Takahashi: Sequencing and characterization of the kinesin-related genes katB and katC of Arabidopsis thaliana Plant Mol Biol 1994, 25:865-876
33. Liu B, Cyr RJ, Palevitz BA: A kinesin-related protein, KatAp in the cells of Arabidopsis and other plants Plant Cell 1996, 8:119-132
34. Tamura K, Nakatani K, Mitus H, Ohashi Y, Takahashi H: Characterization of katD, a kinesin-like protein gene specifically expressed in floral tissues of Arabidopsis thaliana Gene 1999, 230:23-32
35. Reddy ASN, Safadi F, Narasimulu SB, Golovkin M, Hu X: A novel plant calmodulin-binding protein with a kinesin heavy chain motor domain. J Biol Chem 1996, 271:7052-7060.

36. Reddy ASN, Narasimulu SB, Safadi F, Golovkin M: A plant kinesin heavy chain-like protein is a calmodulin-binding protein. Plant J 1996, 10:9-21.

37. Abdel-Ghany SE, Reddy ASN: A novel calcium/calmodulin-regulated kinesin-like protein is highly conserved between monocots and dicots. DNA Cell Biol 2000, 19:567-578.

38. Wang W, Takezawa D, Narasimulu SB, Reddy ASN, Pooovahy BW: A novel kinesin-like protein with a calmodulin-binding domain. Plant Mol Biol 1996, 31:87-100.

39. Y-RJ Lee, Liu S: Identification of a phragmoplast-associated kinesin-related protein in higher plants.Curr Biol 2000, 10:797-800.

40. Barroso C, Chan J, Allan V, Doonan J, Hussey P, Lloyd C: Characterization of a novel kinesin, DcKRP120-2, in carrot cells: characterization of a novel kinesin, DcKRP120-2. Nature 1997, 380:87-100.

41. Reddy ASN: Nuclear pre-mRNA splicing in plants. CRC Crit Rev Pl Sci 2001.

42. Yang JT, Laymon RA, Goldstein LSB: Nuclear region of the stalk domain of ubiquitous human kinesin-C, a calmodulin-binding carboxy-terminal kinesin in animal (Strongylocentrotus purpuratus) cells. J Biol Chem 1998, 273:2698-2705.

43. Abdel-Ghany SE, Reddy ASN: A novel calcium/calmodulin-regulated kinesin-like protein is highly conserved between monocots and dicots. DNA Cell Biol 2000, 19:567-578.

44. Swofford DL: Identifying the calponin homology (CH) domain of actin-binding proteins. Mol Biol Cell 2000, 11:2235-2246.

45. Reddy ASN: Kinesin-related proteins associated with the cold-stable cytoskeleton of carrot cells: characterization of a novel kinesin, KIP3 of Saccharomyces cerevisiae is required for a distinct step in nuclear migration. J Cell Biol 1997, 138:1023-1040.

46. Steinberg G, Schliwa M: The Neurasplosa organelle motor: a distant relative of conventional kinesin with unconventional properties. Mol Biol Cell 1995, 6:1605-1618.

47. Goode BL, Drubin DG, Barnes G: Mitotic spindle positioning in Saccharomyces cerevisiae is accomplished by antagonistically acting microtubule motor proteins. J Cell Biol 1997, 138:41-53.

48. Steinberg G, Schliwa M, Lehmler C, Bolker M, Kahmann R, McIntosh J: Kinesin from the plant pathogenic fungus Ustilago maydis is involved in vacuole formation and cytoplasmic migration. J Cell Sci 1998, 111:2235-2246.

49. Reddy ASN, Day IS: The role of the cytoskeleton and a molecular motor in trichome morphogenesis Trends Plant Sci 2000, 5:503-505.

50. Goode BL, Drubin DG, Barnes G: Functional cooperation between the microtubule and actin cytoskeletons. Curr Opin Cell Biol 2000, 12:63-71.

51. Steinberg G, Schliwa M, Lehmler C, Bolker M, Kahmann R, McIntosh J: Kinesin from the plant pathogenic fungus Ustilago maydis is involved in vacuole formation and cytoplasmic migration. J Cell Sci 1998, 111:2235-2246.

52. Oppenheimer DG, Pollack MA, Vazk J, Szymanski DB, Ericson B, Feldmann K, Marks D: Essential role of a kinesin-like protein in Arabidopsis trichome morphogenesis. Proc Natl Acad Sci USA 1997, 94:6261-6266.

53. DeZwaan TM, Ellington E, Pellman D, Roof DM: Kinesin-related KIP3 of Saccharomyces cerevisiae is required for a distinct step in nuclear migration. J Cell Biol 1997, 138:1023-1040.

54. Bolleti H, Karstens E, Vermos I: Xklp2, a novel Xenopus centrosomal kinesin-like protein required for centrosome separation during mitosis. Cell 1996, 84:49-59.

55. Reddy V, Safadi F, Zielinski RE, Reddy ASN: Interaction of a kinesin-like protein with calmodulin isoforms from Arabidopsis. Mol Biol Cell 1997, 8:3175-3183.

56. Reddy ASN, Narasimulu SB, Day IS: Structural organization of a gene encoding a novel calmodulin-binding kinesin-like protein from Arabidopsis. Genes Dev 1998, 12:195-200.

57. Akl D, Hart CL, Wiedman K, Scholey JM: Identification of kinesin-C, a calmodulin-binding carboxy-terminal kinesin in animal (Strongylocentrotus purpuratus) cells. J Biol Chem 1999, 274:1419-1431.

58. Banuelos S, Saraste M, Carugo KD: Structural comparisons of calponin homology domains: implications for actin binding. Structure 1998, 6:1419-1431.

59. Leinweber BD, Leavis PC, Grabarek Z, Wang CL, Morgan KG: Extra-cellular regulated kinase (ERK) interaction with actin and the calponin homology (CH) domain of actin-binding proteins. Biochem J 1999, 344:117-23.

60. Lippincott-Schwartz J: Cytoskeletal proteins and Golgi dynamics. Curr Opin Cell Biol 1998, 10:52-59.

61. De Matteis MA, Morrow JS: The role of ankyrin and spectrin in membrane transport and domain formation. Curr Opin Cell Biol 1998, 10:542-549.

62. Mockay DJG, Hail A: Rho GTPases. J Biol Chem 1998, 273:20685-20688.

63. van Horck FP, Ahmadian MR, Haessler LC, Moolenaar WH, Kranenburg O: Characterization of p190RhoGEF, a RhoA-specific guanine nucleotide exchange factor that interacts with microtubules. J Biol Chem 2000.

64. Flynn P, Mellor H, Palmer R, Panayotou G, Parker PJ: Multiple interactions of PRK1 with RhoA. Functional assignment of the HR1 repeat motif. J Biol Chem 1998, 273:2698-2705.

65. Goode BL, Drubin DG, Barnes G: Functional cooperation between the microtubule and actin cytoskeletons. Curr Opin Cell Biol 2000, 12:63-71.

66. Corttingham FR, Hoyt MA: Mitotic spindle positioning in Saccharomyces cerevisiae is accomplished by antagonistically acting microtubule motor proteins. J Cell Biol 1997, 138:41-53.

67. DeZwaan TM, Ellington E, Pellman D, Roof DM: Kinesin-related KIP3 of Saccharomyces cerevisiae is required for a distinct step in nuclear migration. J Cell Biol 1997, 138:1023-1040.

68. Miller RK, Heller KK, Frisen L, Wallack DL, Loayza D, Gammie AE, Rose MD: The kinesin-related proteins, Kip2p and Kip3p, function differently in nuclear migration in yeast. Mol Biol Cell 1998, 9:2051-2068.

69. Burgova ST, Wilson PB, Exton JH: Phosphatidate-dependent protein phosphorylation. Proc Natl Acad Sci USA 1991, 88:6210-6213.

70. Yoss JW, Safadi F, Reddy ASN, Hepler PK: The kinesin-like calmodulin binding protein is differentially involved in cell division. Plant Cell 2000, 12:979-990.

71. Oppenheimer DG, Pollack MA, Vazk J, Szymanski DB, Ericson B, Feldmann K, Marks D: Essential role of a kinesin-like protein in Arabidopsis trichome morphogenesis. Proc Natl Acad Sci USA 1997, 94:6261-6266.