

Protein Kinases and Phosphatases in the *Drosophila* Genome

Deborah K. Morrison, Monica S. Murakami, and Vaughn Cleghon

Regulation of Cell Growth Laboratory, National Cancer Institute, Frederick, Maryland 21702

The reversible phosphorylation of proteins on serine, threonine, and tyrosine residues represents a fundamental strategy used by eukaryotic organisms to regulate a host of biological functions, including DNA replication, cell cycle progression, energy metabolism, and cell growth and differentiation. Levels of cellular protein phosphorylation are modulated both by protein kinases and phosphatases. Although the importance of kinases in this process has long been recognized, an appreciation for the complex and fundamental role of phosphatases is more recent. Through extensive biochemical and genetic analysis, we now know that pathways are not simply switched on with kinases and off with phosphatases. Rather, it is the balance of phosphorylation that is often critical. Protein phosphorylation can regulate enzyme function, mediate protein-protein interactions, alter subcellular localization, and control protein stability. Furthermore, kinases and phosphatases may work together to modulate the strength of a signal. Adding further complexity to this picture is the fact that both kinases and phosphatases can function in signaling networks where multiple kinases and phosphatases contribute to the outcome of a pathway. To fully understand this complex and essential regulatory process, the kinases and phosphatases mediating the changes in cellular phosphorylation must be identified and characterized.

A variety of approaches, including biochemical purification, gene isolation by homology, and genetic screens, have been successfully used for the identification of putative protein kinases and phosphatases. Now, the genomic sequencing of organisms promises to be a major contributor to this field. Valuable insight into these important enzymes has already emerged from the analysis of the yeast and worm genomes. In particular, genomic sequencing of *Saccharomyces cerevisiae* and *Caenorhabditis elegans* has revealed the kinase and phosphatase gene families that have arisen during the evolution of multicellular eukaryotes (Plowman et al., 1999). With the recent determination of the *Drosophila* sequence, we can now survey the genome of a second multicellular eukaryote for its repertoire of kinases and phosphatases. In this review, we will present our findings on the protein kinase and phosphatase gene families identified in the fly, together with an examination of the kinase/phosphatase signaling pathways functioning in flies, worms, and humans.

Address correspondence to Deborah Morrison, NCI-FCRDC, P.O. Box B, Frederick, MD 21702. Tel.: (301) 846-1733. Fax: (301) 846-1666. E-mail: morrisod@nciexp.ncifcrf.gov

Vaughn Cleghon's current address is Beatson Institute CRC, Glasgow G61 1BD, UK. E-mail: v.cleghon@beatson.gla.ac.uk

Identification and Classification of Drosophila Protein Kinases and Phosphatases

Our survey of *Drosophila* protein kinases and phosphatases is based on the total set of predicted proteins that were identified in the *Drosophila* genome using automated gene predictor methods (Adams et al., 2000; available at <http://www.celera.com>). The 13,601 predicted fly proteins were surveyed for overall homology with known kinase and phosphatase sequences using BLASTP, and for the presence of polypeptide motifs using BLOCKS and InterPro databases (Rubin et al., 2000). Putative kinases and phosphatases identified by these means were further classified based on the presence of diagnostic amino acid residues in conserved motifs and by sequence similarities extending beyond conserved catalytic domains. Table I summarizes our survey of the *Drosophila* protein kinases and phosphatases. It is important to realize that this analysis represents the first tabulation of these enzymes in *Drosophila* and will be subject to revision as gaps in the genomic sequence are closed and methods for predicting and analyzing genes are improved. In particular, it is known that the Genie and Genscan programs used to annotate the fly genomic sequence make systematic errors with respect to intron-exon boundaries and gene borders, leading us to conclude that some kinase and phosphatase proteins may have been missed by these programs (Reese et al., 2000). These caveats notwithstanding, 251 kinases and 86 phosphatases were identified by our analysis of the predicted *Drosophila* protein set. Remarkably, more than half of these molecules had gone undetected in eight decades of *Drosophila* research.

Protein Kinases

Eukaryotic protein kinases are enzymes that catalyze the transfer of phosphate from ATP or GTP onto serine, threonine, or tyrosine residues of their appropriate substrates. They comprise a single protein superfamily having a common catalytic structure. However, these enzymes can be subdivided into distinct groups based on their structural and functional properties (Hanks and Hunter, 1995).

AGC Family

The AGC serine/threonine kinases function in many intracellular signaling pathways and were first classified based on their tendency to phosphorylate sites surrounded by basic amino acids. *Drosophila* contains ~30 AGC kinases, including members of the cyclic nucleotide-dependent ki-

Table I. Summary of Protein Kinases and Phosphatases in Flies, Worms, and Humans

Group	Fly	Worm*	Humans*
Protein kinase			
AGC	30 (8)	30	100
CaMK	25 (13)	32	83
CKI	8 (6)	87	5
CMGC	24 (7)	42	62
STE	21 (12)	28	63
PTK	32 (8)	92	100
OPK	56 (28)	62	163
Atypical	3 (2)	4	11
Fragment/unknown	18		
Protein kinase like			
Gcyc	11 (6)	26	8
PIK	13 (8)	12	20
DAG	8 (5)	7	8
Choline K	2 (1)	7	2
Phosphatase			
STP	28 (14)	65	21
RPTP, CPTP, LMW-PTP	20 (12)	83	47
DSP	18 (11)	26	51
IPP	20 (18)	11	7

Fly numbers in parentheses represent the proteins newly identified by the fly genome project.

*These numbers are taken from the review by Plowman et al. (1999).

nases, protein kinase C (PKC),¹ AKT, NDR, MNK, MAST, ribosomal S6 kinase, and G protein-coupled receptor kinase families. The majority of the fly AGC kinases had been identified previously by molecular and genetic analysis; however, eight members were uncovered in the fly genome project. Interestingly, four of the new genes encode PKC or PKC-related proteins, including the first atypical PKC isoforms identified in *Drosophila*. Also identified by the fly genome project were additional PKA and PKG proteins, as well as kinases related to mammalian MAST205 and Citron.

CaMK Family

The CaMK serine/threonine kinases also tend to have substrate recognition motifs containing basic amino acids, and some but not all members of this family are regulated by calcium or calmodulin. Approximately 25 CaMKs are present in *Drosophila*, including representatives of the calcium/calmodulin-regulated kinase, SNF1/AMP-dependent kinase, EMK, CHK2, myosin light chain kinase (MLCK), phosphorylase kinase, death-associated protein kinase, and MAPKAP kinase families (the last four of which are found in *C. elegans* but not yeast). Like worms, flies do not encode a complete ortholog of the mammalian Trio kinase, but do have a protein that is related to the entire Trio regulatory domain. CaMK members revealed by the fly genome project include proteins related to calcium/calmodulin-regulated kinases, MLCK, EMK, and mammalian DRAK1. Of the 13 newly identified CaMKs, 6 be-

¹Abbreviations used in this paper: CDK, cyclin-dependent kinase; CKI, casein kinase I; CTK, cytoplasmic tyrosine kinase; DSP, dual specificity phosphatase; LMW, low molecular weight; MKP, MAPK phosphatase; PKC, protein kinase C; PTP, protein tyrosine phosphatase; RTK, receptor tyrosine kinase; STP, serine/threonine protein phosphatase.

long to the EMK family, making this the largest CaMK group in flies. Mammalian and *C. elegans* EMK proteins have been implicated in the regulation of cell polarity and microtubule stability (Drewes et al., 1998).

Casein Kinase I Family

The casein kinase I (CKI) proteins originally were characterized as ubiquitous serine/threonine kinases with a preference for acidic substrates such as casein. Although members of this family were among the first kinases purified, elucidating their function and regulation has been difficult. Recently, however, CKI isoforms have been found to play a role in DNA repair and cell division (Gross and Anderson, 1998), in the Wnt signaling pathway (Peters et al., 1999), and in circadian rhythm regulation (Lowrey et al., 2000). *Drosophila* contains at least eight CKI proteins, only two of which were known previously. Intriguingly, CKI is one of the kinase families that is significantly expanded in the worm, with 87 members identified in *C. elegans* (Plowman et al., 1999). The biological significance of the worm-specific expansion is currently unknown.

CMGC Family

CMGC family members are primarily proline-directed serine/threonine kinases. The major subfamilies of this group play key roles in cell cycle regulation and intracellular signal transduction, and, not surprisingly, are conserved from yeast to humans. Approximately 24 CMGC kinases are found in *Drosophila*, including members of the cyclin-dependent kinase (CDK), CDC-like kinase (CLK), glycogen synthase kinase 3 (GSK3), and MAPK families. Although extensive genetic analysis had revealed many of the *Drosophila* CMGC kinases, seven novel proteins were uncovered by the fly genome project. These include additional CDK (CDK7-like, CDC2-related KKIALRE, CHED-related), GSK3, and MAPK (ERK7) members, as well as an RCK family member (MAK). Also uncovered in the fly genome were proteins related to the MP1 and JIP-1 scaffolding proteins. These molecules function to localize MAPK proteins with their upstream activators and provide signaling specificity (Whitmarsh and Davis, 1998). Although MAPK scaffolding proteins are present in yeast, they are structurally different from the ones found in flies, worms, and mammals, perhaps indicating the evolution of these molecules in multicellular eukaryotes.

STE Family

The STE family is composed of the STE7 (MEK), STE11 (MEKK), and STE20 (MEKKK) kinases that function upstream of MAPK proteins. *Drosophila* contains ~21 members of this family, only 9 of which were known previously. Remarkably, 9 members of the PAK/STE20 group were uncovered by the fly genome project, including proteins related to mammalian PAK3, GLK1, NIK, MST2, STK3, TAO1, and CDC7. Although PAK proteins containing PH domains are found in yeast (Sells et al., 1999), no PH-domain-containing PAKs have been identified in higher eukaryotes and none are present in *Drosophila*. MEKK- and NEK-related kinases were also revealed by the genome project. It is worth noting that even with the discovery of additional MEK and MAPK proteins in the fly, *C. elegans*

contains over twice as many of these kinases, suggesting an expansion of MAPK signaling modules in the worm.

PTK Family

The PTK group consists of receptor (RTK) and cytoplasmic (CTK) tyrosine kinases. Although yeasts contain no conventional PTKs, 92 have been identified in the worm and ~32 are present in the fly. A major function of PTKs is in intercellular communication, perhaps explaining why these enzymes have only been identified in multicellular eukaryotes. In comparison to *Drosophila*, the much larger number of PTKs found in *C. elegans* is due primarily to expansions of the worm-specific Kin-15/16 RTK and FER CTK families. The majority of the fly PTKs had been identified previously by genetic approaches, reflecting the involvement of these proteins in critical growth and developmental pathways. RTKs encoded in the fly genome include the fly-specific Torso and Sevenless kinases, as well as kinases related in sequence if not function to the mammalian EGFR, FGFR, insulin receptor, EPH, RET, ROR, RYK, ALK, and TRK kinases. Of the five newly identified RTKs, two are related to mammalian PDGFR/VEGFR, two are DDR receptors, and one shares homology with FGFR1. In the CTK group, fly members include the JAK, FAK, SYK/SHARK, ACK, ABL, and FPS kinases. Of the newly identified CTKs, one is related to mammalian ACK2 and one is an ortholog of CSK, a kinase that negatively regulates the activity of mammalian SRC kinases. Interestingly, several members of the PTK class are not found in worms, including representatives of the SYK, JAK, TRK, and RET families.

OPK Group

This group is comprised of other protein kinase (OPK) families that do not belong to the six major groups described above. It is the largest class of kinases found in flies and consists of both serine/threonine and dual specificity kinases. Approximately 56 of these enzymes are present in the fly genome, only half of which were known previously. Representatives of this group are extremely diverse and include members of the following families: Aurora, BUB1, CHK1, DYRK, WEE-1, PLK, EIF2, TGF β , and activin receptor, TAK, IKK kinases, CKII, and RAF kinase. Notable in the novel group are additional BUB1 and TAK members and enzymes related to *C. elegans* UNC 51 and mammalian ALK3, DLK, GAK, MLK2, SRPK, IRE, ILK, TLK1, LIM-domain kinase, and LKB1/Peutz-Jeghers kinase.

Atypical, Lipid, and Unknown Kinases

Several protein groups that are structurally related to the eukaryotic protein kinases are also found in the *Drosophila* genome. These include the atypical kinases, guanylyl cyclases, and the eukaryotic lipid kinases. Flies contains at least three atypical kinase members, pyruvate dehydrogenase kinase, A6, and a newly identified BCR protein. Although worms lack BCR, they do contain a protein related to the atypical *Dictyostelium* myosin heavy chain kinase, which appears to be missing in flies. Also absent in both *Drosophila* and *C. elegans* are representatives of the classical prokaryotic histidine kinases. In the lipid kinase group,

Drosophila encodes at least 8 diacylglycerol kinases, 2 choline/ethanolamine kinases, and 13 phosphatidylinositol kinases (PI3-, PI4-, PIP5-, and PIP3-related kinases), the majority of which were unknown previously. In mammalian cells, members of the PIP3-related kinase family participate in the cellular response to DNA damage and have authentic protein kinase activity (for review see Fruman et al., 1998). The fly genome project has revealed three kinases of this group, namely ATM, FRAP-related protein (FRP), and FRAP/TOR; however, as is true for worms, flies do not contain a DNA-PK. Finally, ~18 proteins were identified that represent either partial kinase fragments or kinases with no significant homology to the groups listed above. Since errors have been identified in the transcript annotation of several protein kinases, such as the DDR receptors, Citron, and a PKC isoform, some of the partial kinase sequences may represent intact enzymes that have been improperly annotated. Further analysis will be required to confirm their identity.

Protein Phosphatases

Unlike protein kinases, which share a common catalytic structure, protein phosphatases have different basic structures, use distinct catalytic mechanisms, and comprise at least three separate protein families. Phosphatases are typically classified into two main groups, the serine/threonine protein phosphatases (STPs) and protein tyrosine phosphatases (PTPs).

STPs

STPs can be subdivided into the PPP and PPM families based on distinct amino acid sequences and crystal structures (for review see Cohen, 1997). Both families are widely distributed across phyla with representatives found in yeast, flies, worms, and mammals. Before the *Drosophila* sequencing project, almost all known fly STPs had been identified by molecular cloning approaches. Very few STPs have been isolated by genetic analysis, indicating that shared substrate specificity and/or functional redundancy may have prevented the recovery of such mutants. *Drosophila* contains ~28 STPs, whereas >65 are encoded in the *C. elegans* genome. The increased number of worm STPs appears to be due to an expansion of the PPP family. Members of the PPP family, such as PP1, PP2A, and PP2B, have been implicated in numerous biological processes and signal transduction pathways. The diverse functions of this family are accomplished by a relatively small number of highly conserved catalytic subunits that complex with a wide variety of regulatory proteins, thus targeting the enzyme to specific intracellular locations and substrates. The *Drosophila* genome encodes ~17 PPP catalytic proteins, 8 PP1-related enzymes (including PP1s, PPN, and PPY), 4 PP2A members (including PP2A, PP4, and PPV), 3 PP2B-like molecules, and 2 PP5 proteins. Additional PPP catalytic subunits uncovered by the fly genome project include members of the PP1, PP4, and PP2B groups. In regard to PPP regulatory subunits, *Drosophila* contains at least 3 PP1, 5 PP2A, and 2 PP2B proteins. However, because the regulatory subunits are so diverse, these numbers are likely to be low.

The PPM family includes PP2C and mitochondrial pyruvate dehydrogenase phosphatase. Due to their highly divergent primary sequences, few PPM members have been isolated by homology-based methods and none have been identified by genetic analysis. The only *Drosophila* PP2C protein that had been previously known was identified by genomic walking (Dick et al., 1997). Remarkably, the genome project has uncovered at least 11 new PP2C-related sequences, including one that closely resembles pyruvate dehydrogenase phosphatase. The biological function of the PPM family has been difficult to assess in mammalian cells due to the lack of specific inhibitors that target these enzymes. Recently, however, a PP2C protein has been found to dephosphorylate CDC2 on Thr161 in yeast (Cheng et al., 1999). Whether any of the PP2Cs perform a similar function in *Drosophila* waits to be determined.

PTPs

PTPs are found in all eukaryotic organisms, and are defined by the catalytic signature motif Cys-X5-Arg (for review see Neel and Tonks, 1997). The PTP superfamily consists of classical PTPs (RPTP, CPTP), dual specificity phosphatases (DSPs), and low molecular weight (LMW) PTPs. Approximately 38 PTPs are encoded in the fly genome, including representatives of each class. Again, many more PTPs are found in the worm (109 total). It is interesting to note that the expansion of serine/threonine and tyrosine kinase families in worms has been accompanied by a corresponding expansion of both serine/threonine and tyrosine phosphatases.

Members of the classical PTP family contain a conserved catalytic domain that is often fused to a large non-catalytic region. The PTP noncatalytic domains are quite diverse and can function to regulate enzyme activity and/or mediate protein interactions. Like PTKs, classical PTPs can be divided into two groups, receptor PTPs (RPTPs) and cytoplasmic PTPs (CPTPs). Genetic studies in *Drosophila* have been instrumental to our understanding of both groups. In particular, experiments in the fly were among the first to demonstrate the involvement of RPTPs in neuronal axon guidance (for review see Desai et al., 1997; den Hertog, 1999). *Drosophila* encodes ~8 RPTKs, at least 5 of which function in this capacity. Of the newly identified RPTPs, one is related to mammalian RPTP- κ and two share homology with RPTP-X/1A2, a type 1 transmembrane PTP implicated in nervous system development and insulin-mediated pancreatic function. In regard to the CPTP class, *Drosophila* studies on the CSW phosphatase were pivotal in demonstrating that a CPTP could function as a positive effector of cell signaling (Perkins et al., 1992). CSW is a member of the SH2-domain containing PTPs (SHP subclass). Mammals are known to have at least two SHPs, whereas no additional SHP proteins were found in *Drosophila*, indicating that flies, like worms, possess a single SHP molecule. Overall the fly genome encodes at least 5 CPTPs, namely CSW, PTP-ER, and newly identified CPTPs related to the mammalian MEG1, MEG2, and PTPD1 phosphatases. Finally, *Drosophila* contains four additional PTP-related proteins which are either difficult to classify or represent incomplete phosphatase fragments.

DSPs are a diverse collection of phosphatase subgroups

that share little sequence homology outside of the conserved Cys-X5-Arg motif with other DSP subgroups or with members of the larger PTP family. DSPs were originally characterized by their ability to dephosphorylate both serine/threonine and tyrosine residues; however, some of the DSP subgroups, namely PTEN and myotubularin, also possess lipid phosphatase activity (Maehama and Dixon, 1999). Approximately 18 DSPs are found in *Drosophila*, including representatives of the MAPK phosphatase (MKP), PTEN, nuclear prenylated PRL, myotubularin, PIR1, CDC14, and CDC25 phosphatase groups. Of the nine DSPs uncovered by the fly genome project, six belong to the MKP group, a remarkable finding considering the extraordinary effort spent studying MAPK pathways in *Drosophila*. Only Puckered, a negative regulator of the JNK pathway, previously had been identified by genetic techniques (Martin-Blanco et al., 1998). The failure of the new MKPs to be uncovered by genetic analysis may indicate that they participate in MAPK pathways controlling subtle or unappreciated phenotypes. Alternatively, their functions may have been obscured by redundancy within the MKP group or with other phosphatases. Additional DSPs revealed by the genome project include enzymes related to CDC14 and myotubularin. Interestingly, flies also contain three myotubularin-related sequences that lack the active site Cys and Arg residues. As has been suggested for similar mammalian myotubularin-related molecules, these proteins may function as antiphosphatases by binding to and protecting substrates from dephosphorylation by myotubularin or related phosphatase (Hunter, 1998; for review see Laporte et al., 1998).

LMW-PTPs are ~150-amino acid residue cytoplasmic enzymes that have been shown to possess tyrosine phosphatase activity (Ostanin et al., 1995). Other than a strictly conserved Cys-X5-Arg catalytic motif, LMW-PTPs bear little resemblance to the other PTP members. Mammalian LMW-PTPs have been implicated to function in EPH (Stein et al., 1998) and PDGF receptor signaling (Chiarugi et al., 2000); however, much remains to be learned regarding the biological activity of these enzymes. Although two putative LMW-PTPs are revealed by the *Drosophila* genome project, both predicted proteins are larger than would be expected (424 and 250 amino acids, respectively). The smaller protein contains a complete LMW-PTP domain but lacks the conserved Arg residue in the catalytic motif. Intriguingly, the larger protein has two complete LMW-PTP domains. Although the first domain has a mutation in the active site Cys residue and is likely to be inactive, the second domain contains an intact PTP catalytic motif and presumably has catalytic activity. If this protein is made in vivo, it would represent a new type of LMW-PTP having a tandem catalytic domain structure similar to that observed in many RPTPs. Whether this molecule is an authentic LMW-PTP and whether it has a human counterpart remains to be determined.

Lipid Phosphatases

Lipid inositol phosphatases play an important role in mediating the intracellular balance of second messenger phospholipids. *Drosophila* encodes approximately 20 inositol phosphatases (IPP), only 2 of which were known pre-

viously. Six inositol-1,4,5-triphosphatase phosphatase-like enzymes are contained in the fly genome; yet as is true for worms, no ortholog of the mammalian SH2-domain-containing inositol 5' phosphatase (SHIP) appears to be present. *Drosophila* does encode eight PPAP enzymes, which dephosphorylate phosphatidic acid to generate diacylglycerol. The prototype member of this class, Wunen, was first identified in a genetic screen for factors controlling germ cell migration in the early *Drosophila* embryo (Zhang et al., 1996). Related proteins were subsequently identified in yeast, worms, and mammals. Remarkably, the fly genome project reveals seven additional Wunen-like phosphatases. Also uncovered by the genome project are six members of the inositol monophosphate phosphatase (IMP) group. Both the Wunen-like and inositol monophosphate phosphatases are characterized by small tandem gene arrangements, suggesting a limited expansion of these phosphatase families in *Drosophila*. The large number of newly identified inositol phosphatases underscores the hitherto unappreciated importance of lipid phosphoregulation in the fly.

Comparative Analysis of Phosphorylation-dependent Signaling Pathways

With the completion of both the *Drosophila* and *C. ele-*

gans genome projects, together with our current knowledge of mammalian signaling pathways, we can begin to draw conclusions regarding the regulatory complexity of protein phosphorylation mechanisms across the evolutionary spectrum. For example, in flies, worms, and humans, there is a high degree of structural and functional conservation between the components of the RTK and stress-activated signaling pathways, with the major difference being the number of isoforms present for individual pathway members. In higher organisms, the number of isoforms is increased, presumably providing greater potential for tissue- or stage-specific functions, signaling cross-talk, and regulatory complexity (Fig. 1). Significantly, differences in phosphorylation-mediated signaling cascades between worms, flies and humans become apparent when examining the pathways involved in hematopoiesis and immunity. The JAK/STAT cascade, which has been implicated in hematopoiesis and cytokine signaling, is present in humans and flies. Worms, however, lack JAK kinases but do possess STAT proteins that are regulated by tyrosine phosphorylation. Like humans, flies also contain the Toll/IKK/NFκB pathway, which plays a role in the immune response to microbial organisms. No evidence of an inducible host defense system has been demonstrated in worms, consistent with the lack of this pathway in *C. elegans*. Also miss-

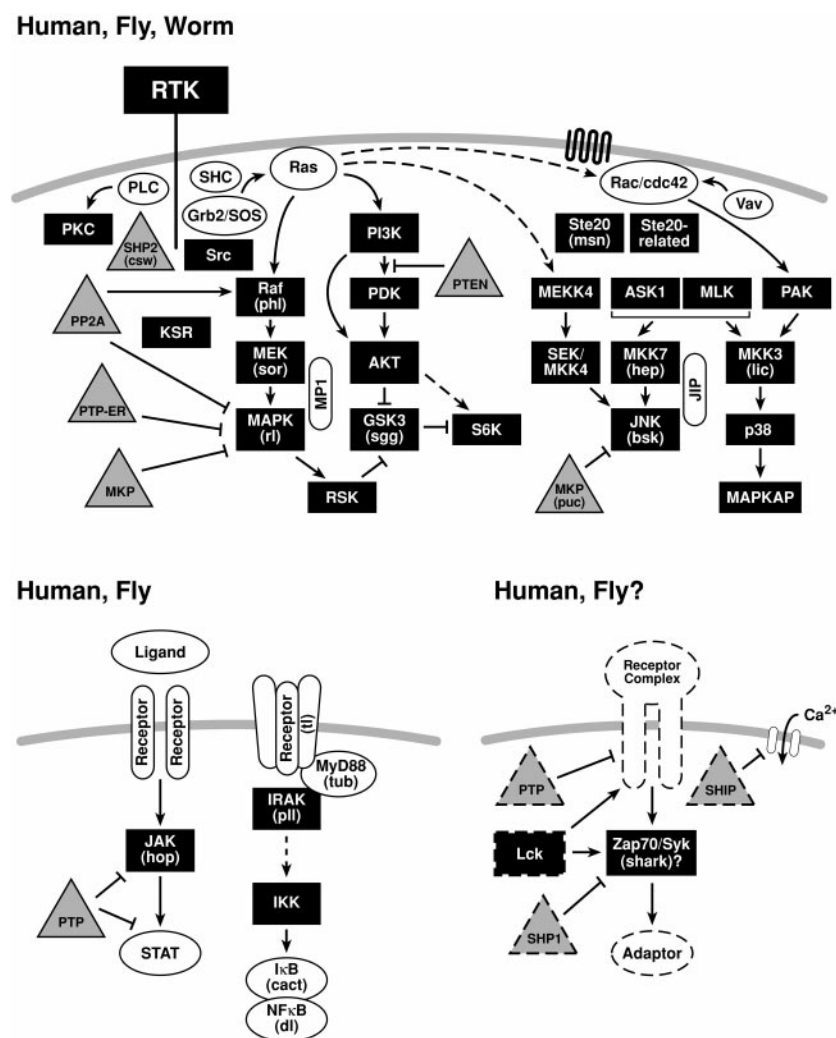


Figure 1. Comparison of the protein kinase/phosphatase signaling pathways in flies, worms, and humans (see text for description). Kinases are depicted as black rectangles, phosphatases are gray triangles, and other signaling components are in white. Shapes in dotted lines indicate mammalian proteins with no clear fly homologue; however, the function of these components in the pathway may be provided by other *Drosophila* proteins with related biochemical activities. *Drosophila* gene names are listed in parentheses.

ing in the worm are the SYK/ZAP70 kinases which play an important role in human T and B cell signaling. *Drosophila* may possess some form of this pathway as indicated by the presence of the fly SHARK kinase. The *Drosophila* SHARK kinase is a member of the SYK/ZAP70 family; however, it is most closely related to the HTK16 kinase of Hydra based on the presence of ANK repeats which are not found in any of the known mammalian SYK/ZAP70 family members (Chan et al., 1994; Ferrante et al. 1995). Exact homologues of proteins functioning with SYK/ZAP70 in the mammalian hematopoietic cascade, including the SLP-76, LAT, and BLNK adaptor proteins, the LCK and LYN kinases, and the SHP-1 and SHIP phosphatases were not revealed by the fly genome project; however, *Drosophila* proteins with related biological activities are found, namely SHP-2, inositol-1,4,5-triphosphate phosphatase, and other SRC-kinase members. Thus, further studies are required to determine whether a rudimentary form of the SYK/ZAP70 pathway does function in flies.

The completion of the *Drosophila* genome project also allows us to look globally at the pathways in which many of the newly identified fly enzymes may function. In particular, many of the proteins revealed in the *Drosophila* genome are orthologs of kinases and phosphatases known to function in the Rac/Rho/CDC42 signaling pathway (Citron, ACK2, MLK2, MEKK4, LIM-domain kinase, PAK/STE20, and DSPs members), in cell cycle regulation (CDK7, BUB1, NEK1, NEK2, CDC14, CDC7, and PP2C), and in pathways establishing asymmetry and cell polarity (LKB1, SLK1, and EMK kinases). Whether these enzymes went undetected for so many years because of functional redundancy or unappreciated phenotypes has yet to be determined.

In conclusion, ~251 protein kinases and 86 phosphatases have been identified in the *Drosophila* genome. Although the overall number of fly enzymes is lower than that found *C. elegans*, the difference is largely due to the worm-specific expansion of certain gene families. Interestingly, no large expansions or deletions of particular kinase or phosphatase gene families were uncovered by the *Drosophila* genome project. All of the previously known *Drosophila* kinases and phosphatases were detected in our analysis, confirming the relative completeness of the genome sequence data. Remarkably, almost 170 new protein kinases and phosphatases were identified by the fly genome project (Table I). The next challenge for scientists will be to determine the role of these enzymes in *Drosophila* development and physiology.

We are grateful to Celera Genomics and the Berkeley *Drosophila* Genome Project for providing access to the *Drosophila* genome sequence and predicted protein dataset. We thank Greg Plowman for his advice and insights into *Drosophila* kinase and phosphatase gene families.

Submitted: 8 June 2000

Revised: 27 June 2000

Accepted: 27 June 2000

References

Adams, M.D., S.E. Celniker, R.A. Holt, C.A. Evans, J.D. Gocayne, P.G.

- Amanatides, S.E. Scherer, P.W. Li, R.A. Hoskins, R.F. Galle, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science*. 287:2185–2195.
- Chan, T.A., C.A. Chu, K.A. Rauen, M. Kroiher, S.M. Tatarewicz, and R.E. Steele. 1994. Identification of a gene encoding a novel protein-tyrosine kinase containing SH2 domains and ankyrin-like repeats. *Oncogene*. 9:1253–1259.
- Cheng, A., K.E. Ross, P. Kaldis, and M.J. Solomon. 1999. Dephosphorylation of cyclin-dependent kinases by type 2C protein phosphatases. *Genes Dev*. 13:2946–2957.
- Chiariugi, P., P. Cirri, L. Taddei, E. Giannoni, G. Camici, G. Manao, G. Raugeri, and G. Ramponi. 2000. The low M(r) protein-tyrosine phosphatase is involved in Rho-mediated cytoskeleton rearrangement after integrin and platelet-derived growth factor stimulation. *J. Biol. Chem.* 275:4640–4646.
- Cohen, P.T.W. 1997. Novel protein serine/threonine phosphatases: variety is the spice of life. *Trends Biochem. Sci.* 22:245–251.
- den Hertog, J. 1999. Protein-tyrosine phosphatases in development. *Mech. Dev.* 85:3–14.
- Desai, C.J., Q. Sun, and K. Zinn. 1997. Tyrosine phosphorylation and axon guidance: of mice and flies. *Curr. Opin. Neurobiol.* 7:70–74.
- Dick, T., S.M. Bahri, and W. Chia. 1997. *Drosophila* DPP2C1, a novel member of the protein phosphatase 2C (PP2C) family. *Gene*. 199:139–143.
- Drewe, G., A. Ebneth, and E.M. Mandelkow. 1998. MAPs, MARKs, and microtubule dynamics. *Trends Biochem. Sci.* 23:307–311.
- Ferrante, A.W., Jr., R. Reinke, and E.R. Stanley. 1995. Shark, a Src homology 2, ankyrin repeat, tyrosine kinase, is expressed on the apical surfaces of ectodermal epithelia. *Proc. Natl. Acad. Sci. USA*. 92:1911–1915.
- Fruman, D.A., R.E. Meyers, and L.C. Cantley. 1998. Phosphoinositide kinases. *Annu. Rev. Biochem.* 67:481–507.
- Gross, S.D., and R.A. Anderson. 1998. Casein kinase I: spatial organization and positioning of a multifunctional protein kinase family. *Cell Signal*. 10:699–711.
- Hanks, S.K., and T. Hunter. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 9:576–596.
- Hunter, T. 1998. Anti-phosphatases take the stage. *Nat. Genet.* 18:303–305.
- Laporte, J., F. Blondeau, A. Buj-Bello, K. Virdee, N. Kirov, A.M. Tolkovsky, and A. Martinez-Arias. 1998. *puckered* encodes a phosphatase that mediates a feedback loop regulating JNK activity during dorsal closure in *Drosophila*. *Genes Dev.* 12:557–570.
- Neel, B.G., and N.K. Tonks. 1997. Protein tyrosine phosphatases in signal transduction. *Curr. Opin. Cell Biol.* 9:193–204.
- Ostani, K., C. Pokalsky, S. Wang, and R.L. Van Etten. 1995. Cloning and characterization of a *Saccharomyces cerevisiae* gene encoding the low molecular weight protein-tyrosine phosphatase. *J. Biol. Chem.* 270:18491–18499.
- Perkins, L.A., I. Larsen, and N. Perrimon. 1992. *corkscrew* encodes a putative tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase *torso*. *Cell*. 70:225–236.
- Peters, J.M., R.M. McKay, J.P. McKay, and J.M. Graff. 1999. Casein kinase I transduces Wnt signals. *Nature*. 401:345–350.
- Plowman, G.D., S. Sudarsanam, J. Bingham, D. Whyte, and T. Hunter. 1999. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc. Natl. Acad. Sci. USA*. 96:13603–13610.
- Reese, M.G., G. Hartzell, N.L. Harris, U. Ohler, J.F. Abril, and S.E. Lewis. 2000. Genome annotation assessment in *Drosophila melanogaster*. *Genome Res.* 10:483–501.
- Rubin, G.M., M.D. Yandell, J.R. Wortman, G.L. Gabor Miklos, C.R. Nelson, I.K. Hariharan, M.E. Fortini, P.W. Li, R. Apweiler, W. Fleischmann, et al. 2000. Comparative genomics of the eukaryotes. *Science*. 287:2204–2215.
- Sells, M.A., and J. Chernoff. 1997. Emerging from the PAK: the p21-activated protein kinase family. *Trends Cell Biol.* 7:162–167.
- Stein, E., A.A. Lane, D.P. Cerretti, H.O. Shloekmann, A.D. Schroff, R.L. Van Etten, and T.O. Daniel. 1998. Eph receptors discriminate specific ligand oligomers to determine alternative signaling complexes, attachment, and assembly responses. *Genes Dev.* 12:667–678.
- Whitmarsh, A.J., and R.J. Davis. 1998. Structural organization of MAP-kinase signaling modules by scaffold proteins in yeast and mammals. *Trends Biochem. Sci.* 23:481–485.
- Zhang, N., J. Zhang, Y. Cheng, and K. Howard. 1996. Identification and genetic analysis of *wunen*, a gene guiding *Drosophila melanogaster* germ cell migration. *Genetics*. 143:1231–1241.