A Drosophila-centric view of protein tyrosine phosphatases

Teri Hatzihristidis a,b, Nikita Desai a,c, Andrew P. Hutchins d, Tzu-Ching Meng e,f,g, Michel L. Tremblay a,b,c,* , Diego Miranda-Saavedra h,i,j,*

a Goodman Cancer Research Centre, McGill University, 1160 Pine Avenue, Montreal, Québec H3A 1A3, Canada
b Department of Medicine, Division of Experimental Medicine, McGill University, Montreal, Quebec, Canada
c Department of Biochemistry, McGill University, Montreal, Quebec, Canada
d Key Laboratory of Regenerative Biology and Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, South China Institute for Stem Cell Biology and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, Guangdong 510530, China
e Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan
f Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan
g Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan
h World Premier International (WPI) Immunology Frontier Research Center (IFReC), Osaka University, 3-1 Yamadaoka, Suita 565-0871, Osaka, Japan
i Centro de Biología Molecular Severo Ochoa, CSIC/Universidad Autónoma de Madrid, 28049 Madrid, Spain
j IE Business School, IE University, María de Molina 31 bis, 28006 Madrid, Spain

Abstract

Most of our knowledge on protein tyrosine phosphatases (PTPs) is derived from human pathologies and mouse knockout models. These models largely correlate well with human disease phenotypes, but can be ambiguous due to compensatory mechanisms introduced by paralogous genes. Here we present the analysis of the PTP complement of the fruit fly and the complementary view that PTP studies in Drosophila will accelerate our understanding of PTPs in physiological and pathological conditions. With only 44 PTP genes, Drosophila represents a streamlined version of the human complement. Our integrated analysis places the Drosophila PTPs into evolutionary and functional contexts, thereby providing a platform for the exploitation of the fly for PTP research and the transfer of knowledge onto other model systems.

Keywords:
Tyrosine phosphatase
Sequence analysis
PTP-central
Model system
Phosphorylation
Biochemical evolution
Drosophila

1. Introduction

Reversible protein phosphorylation is one of the most widespread mechanisms for controlling cellular functions. Tyrosine phosphorylation, in particular, has evolved into a sophisticated regulatory system in metazoans to control many animal-specific processes, ranging from development to cellular shape and motility, transcriptional regulation and proliferation versus differentiation decisions [1]. Out of 498 genes encoding protein kinases in the human genome, 91 encode tyrosine kinases (TKs) [2]. The set of TKs is complemented with 109 protein tyrosine phosphatase (PTP) genes [3], although TKs and PTPs do not seem to have overlapping targets.

Given the central importance of tyrosine phosphorylation, it is no surprise that its abnormal regulation is responsible for many human diseases: diabetes, obesity, cancer, inflammatory diseases, and many others have been associated with PTP over-expression or deficiencies in human [4]. Historically research on TKs has advanced at a faster rate than that on PTPs; not only were TKs identified nearly a decade earlier than PTPs, but also the intrinsic difficulties of investigating the “disappearance” of a phosphate moiety as opposed to the appearance of a radioactive phosphate represented a major burden for the PTP field. Major advances have...
been made with the development of substrate trapping techniques where specific mutations of the PTP catalytic domains allow the purification, detection and identification of their physiological substrates. Although PTPs are generally expressed at low levels, they are enzymatically very active and as a result their ectopic expression in cellular systems easily leads to off-target effects, usually resulting in cell death [5]. Therefore, the use of model organisms is indispensable for the study of PTPs.

Today we know that in most cases phosphatases play a dominant role over protein kinases in shaping the spatio-temporal dynamics of protein phosphorylation networks, which are characterized by their amplitude and duration [6]. Thus it has been proposed that in many instances protein phosphatases make better drug targets than protein kinases. Most of our current knowledge on PTPs is derived from mutations identified in human pathologies as well as loss-of-function studies in the mouse (embryonic gene targeting models, and siRNA and shRNA knockdown experiments). To a large extent, the murine genetic deletions correlate well with human disease phenotypes and have been instrumental in understanding the central importance of PTPs in cellular signaling, whose effects can range from embryonic-lethal to relatively mild and nearly unnoticeable likely due to the presence of compensatory mechanisms by paralogous PTP genes [4,5]. Whereas the mouse has so far been the favoured organism for the genetic dissection of PTP-controlled pathways, here we argue that Drosophila melanogaster with its streamlined version of the human tyrosine phosphatome is a complementary and powerful system for the dissection of PTP functions in vivo.

The fruit fly occupies a paramount position among model organisms: Drosophila research yielded the first observations on a wide variety of fundamental biological principles that are conserved in humans despite several hundred million years of independent evolution [7–9]. These include the principles of embryonic patterning by Hox genes, which apply to all bilaterian animals [10], and the functional conservation of other key signaling pathways, including among others: (i) the Notch signaling pathway [11–22]; (ii) the Wnt set of signal transduction pathways, which function in various developmental pathways (body-axis patterning, cell migration, cell fate specification, cell proliferation; and (iii) the identification of the Drosophila receptor tyrosine kinase gene Sevenless (sev), whose target protein Son of sevenless (Sos) unveiled the relationship between receptor TK and Ras signaling in eye development [23]. These conserved signaling pathways are frequently dysregulated in cancers and type II diabetes [24–27].

The main advantage of the fly over other model organisms is that once a suitable phenotype has been obtained, its genetic toolkit can be used to dissect the underlying disease pathways either by loss-of-function (by RNAi, transposon insertion, imprecise excision, or X-ray and chemical mutagenesis [28]) or gain-of-function in a tissue-specific manner (e.g. using the GAL4/UAS system [29]). Thus, Drosophila is a particularly attractive system for establishing models of human disease and studying genetic interactions for a number of reasons [8]: (i) many well-studied biological pathways are conserved between human and Drosophila, including pattern formation, endocrine and intracellular signaling, and cell death [30]. Furthermore, in recent times the fruit fly has emerged as an extraordinarily powerful model for the study of human metabolism, where a homolog of leptin (a hormone that regulates energy intake and expenditure, including appetite, hunger, metabolism and behavior) called unpaired 2 (upd2, a JAK-STAT pathway ligand) works similarly to human leptin [31]; (ii) over 75% of human disease-associated genes have a homolog in the fly [32], and as such fly models have been successfully established for many distinct diseases, including neurodegenerative disorders, cancer, cardiac, immunological and developmental disorders [33]; (iii) its short life cycle and abundant progeny facilitates genetic modeling, whereas its relatively short lifespan (~120 days) makes studies on ageing particularly feasible.

Here we describe the PTP complement of D. melanogaster, a model organism where studies on tyrosine phosphatases have previously shed light on key pan-metazoan functions. We draw similarities and differences between the tyrosine phosphatome of the fruit fly and those of human, mouse, worm and yeast, highlighting important conserved functions that underlie the potential of the fruit fly as a key model organism for the genetic and biochemical study of tyrosine phosphatases in higher animals.

2. A revised annotation of the Drosophila tyrosine phosphatome

The PTP superfamily is divided into 4 distinct classes that differ both in their catalytic mechanisms and phosphatase catalytic domain sequences [34]. Class I are cysteine-based PTPs including the classical tyrosine-specific phosphatases (both receptor and non-receptor), and the dual-specificity phosphatases (DSP, or VH1-like). DSPs are the most promiscuous type of PTPs in terms of substrate specificity, with some members dephosphorylating mRNA 5’-triphosphate while other enzymes dephosphorylate lipids. Class II PTPs are a small but evolutionarily highly conserved group of PTPs with only one member in humans (ACP1); they are also found in bacteria (which display tyrosine phosphorylation [35]) and are structurally related to bacterial arsenate reductases. Class III PTPs, like the Class I and II, are also cysteine-based enzymes displaying specificity towards tyrosine and threonine residues. The human enzymes (CDC25A, CDC25B and CDC25C) control cell cycle progression by dephosphorylating cyclin-dependent kinases. Despite sharing a cysteine-based catalytic mechanism, Class I, II, and III PTPs are believed to have evolved independently. A fourth class of PTPs displays an aspartic acid-based catalytic mechanism with dependence on a cation, and is represented by the developmentally important EyA (‘Eyes Absent’) genes, of which only one member is found in the fruit fly compared to four genes in mouse and human.

We recently developed a highly sensitive and specific method for the automatic classification of proteins into the various PTP classes and families (‘Y-Phosphatomer’). Y-Phosphatomer relies on a specific collection of protein domain models drawn from InterPro member databases [3]. Upon evaluation, Y-Phosphatomer reported perfect coverage and classification rates. Then, as proof of principle we reannotated the human tyrosine phosphatome and showed that the human genome harbors 109 PTP genes instead of 105 genes as originally reported in a hallmark paper 10 years ago [34]. We subsequently used Y-Phosphatomer to annotate the PTP complements of 65 eukaryotic genomes (including Drosophila), which are available through the PTP-central database (http://www.PTP-central.org/) [3]. The D. melanogaster genome contains 44 PTP genes (Fig. 1A), with RNA-seq data from FlyBase [36] supporting both the robust expression and intron-exon structures of all PTPs genes. The curated tyrosine phosphatome of the fly presented here extends the previous catalogue of 16 (classical) tyrosine-specific PTPs [37] by 21 dual-specificity phosphatases (DSPs), 4 Class II phosphatases (LMWPs), 2 Class III phosphatases (CDC25s) and 1 eyes-absent (EyA) homolog (Table 1).

Previous efforts at characterizing the PTP complement of D. melanogaster include those by Andersen et al. [37] and Morrison et al. [38], While the former study exclusively concentrates on Class I tyrosine-specific PTPs (and correctly identifies all 16 genes in Drosophila), the latter is much broader in scope and attempts to identify all protein kinases and phosphatases in the fly genome. To do so, Morrison and colleagues mined the fly genome using
collections are both especially sensitive and specific, as we previously demonstrated for protein kinases [2, 40] and ubiquitinating and deubiquitinating enzymes [41]. The protein models combined into the Y-Phosphatomer library were specifically built to represent the diversity of the PTP repertoire and Y-Phosphatomer outperforms search efforts that use standard sequence-analysis tools.

3. The Drosophila tyrosine phosphatome is a streamlined version of the human complement

The D. melanogaster genome contains 44 PTP genes (Table 1, Fig. 1A). All phylogenetic relationships reported here are those available on the PTP-central site, and as originally fetched from the MetaPhOrs database [42]. MetaPhOrs is a public repository of orthologs and paralogs derived from phylogenetic trees from five distinct databases (PhylomeDB [43], Ensembl [44], EggNOG [45], Hogenom [46] and TreeFAM [47]). Thus, MetaPhOrs is the most comprehensive database of homology relationships currently available, and provides specific metrics to assess the quality of the PTP homology predictions reported here.

Interrogation of the PTP complements of other model organisms in PTP-central shows that the proportion of genes encoding PTPs is kept in strict bounds from yeast to human (~0.2% of all genes). The fly’s PTP complement is larger than that of yeast (17 genes), but smaller than those of C. elegans (125 genes) or human (109 genes) (Table 1, Fig. 1). The C. elegans genome, however, is known to have undergone large protein family expansions in comparison with other nematodes of similar phenotypic complexity [48], including 11 receptor PTPs harboring unique extracellular domains that are absent in other species [49].

Whereas 75% of human PTPs are multidomain proteins [34], only about half of the fly PTPs harbor accessory domains that enhance the enzymes’ capabilities. In striking contrast, some serine/threonine phosphatases, such as PP1, mainly consist of small catalytic subunits that bind regulatory or targeting subunits to form holoenzymes with different functions. PP1 Ser/Thr phosphatases may in principle generate more flexibility and functional diversity, but strict specificity must be provided by the PP1-interacting proteins that serve as targeting subunits, substrates and/or inhibitors [50]. Hence one may argue that the tight regulation typical of single-chain multidomain PTPs may be better suited to the specific regulation of signaling pathways. We describe 17 accessory domains associated with the fly PTPs, which can be classified into four main functional categories: (i) cellular localization (e.g. FERM and PH); (ii) protein–protein interaction (e.g. SH2, immunoglobulin); (iii) catalytic (e.g. the catalytic domain of mRNA-capping enzyme); and (iv) small-molecule binding (e.g.

### Table 1
Comparison of the human and Drosophila tyrosine phosphatomes split by PTP class and family.

<table>
<thead>
<tr>
<th>PTP classification</th>
<th>Human</th>
<th>Fruit fly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classical PTPs</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Classical PTPs</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Dual-specificity PTPs</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Atypical</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>Slingshots</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>PRD</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>CDC14</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>PTEN</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Myotubulinins</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>Class II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LMWP</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Class III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDC25</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Asp-based PTPs</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>109</td>
<td>44</td>
</tr>
</tbody>
</table>
It must be noted, however, that although comparatively more human PTPs harbor accessory domains, the types of such domains found both in human and fly PTPs are almost identical.

3.1. Drosophila pseudophosphatases

Besides regulation by accessory domains and interacting proteins, PTPs that lack catalytic activity (‘pseudophosphatases’) provide an additional level of complexity. Proposed functions of pseudophosphatases include: (i) being part of scaffolding signaling complexes; (ii) acting as ‘anti-PTPs’ that protect specific substrates from other active PTPs; and (iii) working as functional inhibitors of their substrates to stoichiometrically remove phosphorylated proteins from their normal signaling functions, as has been suggested for the C. elegans pseudophosphatases EGG-4 and EGG-5 [51,52].

The importance of catalytically inactive enzymes in regulating key cellular processes was probably first understood in the case of pseudokinases, which act as signal transducers by integrating signaling network components, as well as being allosteric activators of active protein kinases [53,54]. In Drosophila, 11/44 (25%) of PTPs harbor catalytically inactive domains (labeled ‘PTPi’ in Fig. 2), including 4 receptor (Lar, Ptp69D, Ptp99A and IA-2) and 2 non-receptor PTPs (CG7180 and Mop), 3 myotubulinars (CG5026, CG14411 and Sbf), and 2 LMWPs (Primo-1 and CG31469). The classification of a PTP as a pseudophosphatase is not as straightforward as it is for protein kinases: a kinase should lack at least one of three essential motifs in the catalytic domain (the VAIK, HRD or the DFG motif) to be classified as a pseudokinase [53]. In 2001, Andersen et al. [55] described 10 conserved motifs in the catalytic domain of PTP1B and highlighted specific amino acids common to all PTPs that play critical roles in catalysis. We used both information from the literature and this set of conservation criteria to predict the catalytic activity of Drosophila PTPs. For instance, the second domains (D2) of all receptor PTPs lack a critical amino acid in motif I (Tyr46 in PTP1B), and mutations of the general acid donor of the WPD loop (Asp181 in PTP1B) also render D2 domains catalytically inert. However, D2 domains in an increasing number of pseudophosphatases appear to have retained their capacity to bind substrates, potentially acting as important regulators of signaling pathways.

3.2. Class I PTPs: receptor tyrosine-specific phosphatases (RPTPs)

Nearly 20% of human PTPs are predicted to contain a transmembrane domain [34], a figure similar to the fly genome where we have identified 8 RPTPs (~18% of the tyrosine phosphatome), 7 of which are likely to be active enzymes (Fig. 2). The functions of RPTPs were first described in the development of the fly’s nervous system where these enzymes are essential for the proper regulation of axon guidance and synapse formation. Mutations in Lar, Ptp69D, Ptp99A, Ptp4E, Ptp10D and Ptp52F affect axon guidance, either alone or in combination as each guidance decision made by embryonic motor axons during outgrowth to their muscle targets requires a specific subset of the neural RPTPs. Highly penetrant defects in the CNS and motor axon guidance are typically observed only when specific combinations of two or more RPTPs are removed. Functionally the RPTPs can be divided into two classes: (a) lethals, which have a clear phenotype upon deletion [Lar, Ptp69D and Ptp52F]; and (b) viables, which present no phenotype on their own but synergize with each other and with
mutations in the lethal RPTPs to produce stronger phenotypes (Ptp99A, Ptp10D and Ptp4E). Structurally, 3 RPTPs present tandem PTP domains (Lar, Ptp69D and Ptp99A, where the second domain, D2, is catalytically inactive), and a protein architecture reminiscent of the R1/R6 RPTP subtype. The proposed function of the inactive D2 domain is to serve as a dynamic modulator of the D1 domain, which mechanistically could happen in several ways: (i) through the direct inhibition of the D1 domain by physical hindrance (hinge vs. head-and-toe models) [56]; (ii) by serving as a redox sensor for modulating D1 activity [57]; and (iii) possibly D2 could function as a pseudophosphatase domain that interacts with substrates to prevent them from performing their functions [52]. Five RPTPs in the fly genome (Ptp4E, Ptp10D, Ptp52F, CG42327 and IA-2) harbor a single PTP domain.

3.2.1. Tandem domain RPTPs: Lar, Ptp69D and Ptp99A
Drosophila Lar is the fly ortholog of human and mouse R2B sub-family members PTPRD, PTPRF and PTPRS. Like the human enzymes, Lar is strongly expressed in the brain and also in tissues of endodermal origin. Lar has a crucial role in regulating synapse growth and maturation at the neuromuscular junction, relying on its ligands Syndecan (Sdc) and Dally-like. Sdc is a heparan sulfate proteoglycan that contributes to Lar function in motor axon guidance. The overexpression of Sdc in muscles produces the same phenotype as the overexpression of Lar in neurons, and the genetic deletion of Lar suppresses the effect produced by ectopic muscle Sdc. In muscles Sdc can interact with neuronal Lar and binding to Sdc increases Lar's signaling efficiency [58]. In contrast, the family of extracellular matrix molecules known as chondroitin sulfate proteoglycans (CSPGs) negatively regulate axonal growth by interacting with neuronal Lar and inactivating the downstream Akt and RhoA signals. Mice lacking the Lar orthologue do overcome the neurite growth restrictions imposed by CSPGs in neuronal cultures [59]. However, the function of Lar in controlling R7 photoreceptor axon targeting in the visual system differs in a number of respects: (i) the extracellular domain of Lar is different and might be regulated by other ligands; and (ii) R7 targeting does not require Lar's phosphatase activity but relies on that of Ptp69D. Although Lar's phosphatase activity is not required for R7, Lar dimerisation probably leads to the assembly of downstream effectors [60]. Thus, the genetic deletion of Lar leads to axon elongation and fasciculation deficiency, a phenotype that is partially mimicked in compound mouse knock-outs of the Ptprd and Ptprs genes [61]. In particular, the mouse Ptprs knockout promotes axonal regeneration and even increases memory and the plasticity of neural interactions in the brain [62]. Similarly, the C. elegans ortholog, ptp-3, is most highly expressed in the nervous system, from the late embryo to the adult stages [63]. It thus seems that the original function of Lar (neuron targeting) has become diversified in the course of evolution into additional functions of brain patterning and synaptic interaction formation in mammals. Not surprisingly, both the mammalian ligand CSPG and the cytoplasmic signaling networks downstream of the receptors have diverged from those of the fly [64,65]. Interestingly, recent work suggests that the inhibition of Ptprs might constitute a novel approach in the treatment of spinal cord, brain injuries and neurodegenerative diseases [62,66–69].

The two other tandem domain-containing receptor PTPs, Ptp69D and Ptp99A, are also involved in neural patterning. Their extracellular domains contain the typical immunoglobulin-like and fibronectin-like domains that mediate cell–cell interactions. In contrast to Lar, Ptp69D and Ptp99A have a much smaller number of fibronectin repeats (3 and 4, respectively). Whereas Ptp99A has two detectable orthologues in human (PTPRG and PTPRZ1), Ptp69D is the phylogenetic ortholog of human CD45 (PTPRC), a key regulator of lymphocyte antigen receptor signaling. However, adaptive immune mechanisms are absent in the fly. While being somewhat functionally redundant, differing expression patterns in the Drosophila CNS (Ptp69D in young axons, Ptp99A and Lar in older axons) as well as the study of mutant Lar, Ptp69D and Ptp99A, indicate that these enzymes may have specialized roles in CNS development and in the adult brain of the fly [70]. In embryos lacking Ptp69D motor neuron growth cones cease to grow before reaching their muscle targets, or follow incorrect pathways that bypass these muscles. Embryos lacking Ptp99A have no distinguishable phenotype, but Ptp99A Ptp69D double mutants present a much more severe phenotype. Therefore, Ptp69D and Ptp99A are required for motor axon guidance and have partially redundant roles during the development of the neuromuscular system [71]. Furthermore, Ptp69D seems to function in conjunction with Ptp10D in bundle formation and is necessary for axonal growth to the lobes from the peduncle [70], whereas Lar mutants display a phenotype of axon overgrowth across the midline of the brain in both larval and adult brains [70]. Given the parallel between neuronal and immune synapses, it is interesting to note that these enzymes have acquired additional roles in immunity in higher animals [72]. For instance the human orthologs of Lar possess an extracellular carbonic anhydrase domain (absent in flies), responsible for catalyzing the interconversion of carbon dioxide and bicarbonate to maintain the acid-base balance in blood and tissues. Despite Ptp69D and CD45 (PTPRC) being strict phylogenetic orthologs, their proteins' extracellular domains are slightly different, which is understandable since the fly has evolved separately for over 400 million years to possibly recognize different extracellular stimuli while maintaining the same intracellular signaling domain architecture. Finally, the worm ortholog of Ptp69D (clr-1) has non-neural roles but is also broadly involved in developmental processes as it has been shown to be a negative regulator of the FGF receptor (EGL-15) by diphosphorylation [73].

3.2.2. Ptp4E, Ptp10D and Ptp52F
Ptp4E, Ptp10D and Ptp52F are all structurally similar with a string of N-terminal fibronectin-like domains and a single, intracellular, active catalytic domain, and all are homologous to enzymes of the R3 subtype (PTPRB/PTPRO).

Ptp4E and Ptp10D have important roles in neural patterning in the fly, although their expression pattern is not limited to neurons. Ptp4E is broadly expressed and its mutants are viable and fertile, but Ptp4E Ptp10D double mutants have a mild CNS phenotype and the resulting embryos die before the larval stage. Whereas Ptp4E Ptp69D double mutants have no phenotype, Ptp10D Ptp69D double mutants present a strong CNS phenotype, whereas axons cross the midline, and the outer and middle longitudinal bundles are fused to the inner bundle. Therefore both Ptp10D and Ptp69D are necessary for repulsion of growth cones from the midline of the embryonic CNS [74,75]. Ptp10D cooperates with Lar, Ptp69D and Ptp99A to facilitate the outgrowth and bifurcation of the Snail nerve (one of the five main nerve branches), but acts in opposition to the others in regulating the extension of ISN motor axons past intermediate targets [76]. A GAL4-based screening method developed by the Zinn laboratory to identify orphan receptor ligands identified Stranded at second (Sas) as a Ptp10D ligand required to prevent longitudinal axons from crossing the midline. Whereas Sas is expressed both in neurons and glia, Ptp10D expression is restricted to CNS axons [77]. Moreover, in the adult brain, signaling through Ptp10D is required for long-term memory but not for learning, early memory or anesthesia-resistant memory [78].

Studies on compound mutations of both Ptp4E and Ptp10D by Jeon and Zinn unveiled a complementary role in embryonic tracheal tube formation, which is partly controlled by Ptp4E and Ptp10D through the down-regulation of the FGFR ortholog (Btl), Breathless and Prv receptor tyrosine kinases (which in turn have
compensatory activities) [79,80]. This non-neuronal function is in line with other reports on murine members of the R3 subtype, such as PTPRJ, which also controls receptor tyrosine kinase activity during organ development [81,82]. The C. elegans ortholog of Ptp10D (dep-1) functions through a similar mechanism by negatively regulating EGFR signaling, but instead controls vulval fate during development, and duct cell and excretory pore development [83].

Ptp52F is selectively expressed in the CNS of late embryos, where Ptp52F mRNA knockdown mutants present motor axon and CNS axon guidance phenotypes. However, this phenotype is suppressed in Lar Ptp52F double mutants, indicating that these RPTPs compete in the regulation of CNS axon guidance decisions [84]. An associated cell surface receptor, Tartan, displays a similar phenotype upon interference and it has been proposed to be a ligand of Ptp52F [85]. The expression pattern of Ptp52F is very similar to that of its mouse ortholog, Ptpro (also known as stomach-associated phosphatase or SAP-1) [86]. Therefore, in common with their mouse orthologs, this set of fly RPTPs have functional roles that extend beyond the nervous system, and are clearly complementing each other in controlling phosphotyrosine-based signaling pathways involved in pattern formation and cell-to-cell interactions.

Furthermore, RNAi-based strategies have shown that Ptp52F plays an indispensable role in the destruction of the larval midgut during the larva-to-pupa transition [86]. The destruction of the larval midgut is a critical developmental process triggered by the molting hormone ecdysone. Ecdysone induces the expression of Ptp52F, which in turn dephosphorylates Transitional ER ATPase (Ter94), a regulator of the ubiquitin proteasome system. Dephosphorylated Ter94 leads to the fast degradation of ubiquitinated proteins, including the Drosophila inhibitor of apoptosis 1 (Diap1) whose degradation is essential for the onset of apoptosis and the initiation of autophagy [87].

3.2.3. CG42327

Almost nothing is known about the RPTP CG42327 since no mutants have so far been characterized. The CG42327 gene achieves expression peaks 12–24 h post-fertilisation, during the late larval stages and at various stages throughout the pupal period. In the adult fly CG42327 is particularly highly expressed in the heart, as well as the eye, the thoraco-abdominal ganglion, the midgut and the larval/adult hindgut and carcass [36].

3.2.4. IA-2

The eighth RPTP, Islet antigen-2 (IA-2) is the ortholog of the mammalian genes insulinoma-associated antigen (PTPRN and PTPRN2), initially described upon detection of high levels of anti-IA-2 antibodies in type I diabetes patients. Moreover, sera from type I diabetic patients has been reported to recognize Drosophila IA-2. Drosophila IA-2 is predicted to be catalytically inactive due to the lack of the catalytic cysteine (mutated to a glycine residue), and the replacement of the essential aspartate by an alanine residue. In situ hybridization showed that ia-2 is expressed in the CNS (the neuronal pattern of expression being similar to that of mammals), and the midgut region where it plays a critical role in gut development during metamorphosis [88]. Given the evolutionary conservation of ia-2 genes in Drosophila and C. elegans (ida-1), both of which lack a pancreas and an adaptive immune system, this suggests that essential functions of this pseudophosphatase remain to be discovered.

3.3. Class I PTPs: non-receptor tyrosine-specific phosphatases

Eight non-receptor PTPs are found in the fly genome, which map to 7 of the 10 human subtypes.

3.3.1. CG7180

CG7180 is a PTP gene that we describe here in detail for the first time: CG7180 harbors 2 PTP domains (thus resembling an RPTP) but lacks any signal peptide or transmembrane segments, and it is expressed in the late pupal stages, in the central nervous system of the adult, as well as in other tissues of endodermal origin. Interestingly only a single transcript encoding 2 exons is reported, with the 5′-terminal exon encoding the entire protein. This is important, as RPTPs are known to have multiple splice isoforms, some of which encode only cytoplasmic PTP domains. Therefore it appears that CG7180 has a unique evolutionary history by encoding a dual catalytic intracellular PTP. Therefore all evidence points to CG7180 being a cytoplasmic enzyme. Although no clear orthology relationships can be ascribed to CG7180, it bears considerable similarity to a number of receptor PTPs of the R2A subtype (PTPRK, PTPRM, PTPRT and PTPRU). Sequence similarity is found throughout the first and second PTP domains, and even in the junction sequence linking them. Two possible hypotheses are that either the CG7180 gene lost its extracellular domain secondarily, or that such an extracellular domain was never gained. In the light of the work by Muller et al. [89], the latter possibility is more plausible. Upon cloning RPTPs in the sponge Geodia cydonium (sponges being examples of primitive metazoans as they lack true tissues or organs, and present no body symmetry), Muller and colleagues proposed that the fibronectin-containing extracellular domain of receptor PTPs, and each of their two PTP domains, evolved independently. The PTP domain would have appeared before the lineage leading to yeasts, but the extracellular fibronectin domain likely appeared immediately prior to the emergence of metazoans. CG7180 thus presents an exciting opportunity to study the evolution of receptor PTPs with tandem catalytic domains. Examination of the amino acid sequence of the two catalytic domains suggests that the D2 domain is catalytically inactive (as with other D2 domains, the conserved tyrosine residue-Y46 in PTP1B-present in Motif 1 of the NXXKNRY motif is mutated to an Aspartate (NREKNRD)). Additionally, the WPD loop motif of the second domain harbors a methionine instead of the general acid donor aspartate. Moreover, motif 1 of the D1 domain of CG7180 harbors the sequence NLEKNQ, where the important RY amino acid pair is missing. Besides, in the WPD loop of D1, a WYS motif is found with the proline and the crucial aspartate missing. Therefore, CG7180 is the first identified cytoplasmic enzyme harboring 2 pseudophosphatase domains.

3.3.2. Corkscrew

Corkscrew (csw) is the fly ortholog of the SH2 domain-containing PTPs SHP-1 (PTPN6) and SHP-2 (PTPN11). The structural and functional characteristics of Corkscrew have been extremely well conserved in evolution: for instance, the csw mutant phenotype mimicks gain-of-function mutations typical of the LEOPARD syndrome (an autosomal dominant Ras/MAPK multisystem condition caused by mutations in PTPN11) [90]. Both SHP-1 and SHP-2 play critical roles in hematopoiesis as positive and negative regulators, respectively [91]. Moreover, corkscrew is involved in oogenesis and is a player in terminal cell fate determination [92], a function mirrored by the C. elegans ortholog ptp-2 [93].

3.3.3. l(1)G0232

The Drosophila l(1)G0232 gene is orthologous to mammalian PTPN9 (also known as PTP-MEG2). Eight splice isoforms of l(1)G0232 mRNA have been reported in the fly (compared to six in human), all of which encode an active PTP domain plus 1 or 2 CRAL-TRIO domains, except for the shorter isoforms which only harbor the PTP domain. CRAL-TRIO are ~170 amino acid domains
that form hydrophobic lipid-binding pockets and are generally present in membrane-associated proteins such as GTPases. Although l(1)G0232 has not been thoroughly studied, mammalian PTP-MEG2 has been reportedly associated with secretory vesicles through its Sec14p domain [94]. Several studies have implicated PTPN9 in the negative regulation of signaling downstream of important tyrosine kinase receptors, such as VEGFR in endothelial cells [95], and ErbB2 and EGFR in breast cancer cell lines [96] by dephosphorylating STAT3 [97]. STAT3 also appears to be a target of PTPN9 during primitive hematopoiesis in zebrafish [98].

3.3.4. Mop

The myopic (mop) gene is orthologous to PTPN23 (HD-PTP) in mammals, and the only PTP in the fly genome harboring a BRO protein–protein interaction domain. The BRO domain is found in proteins associated with endosomal trafficking through potential interactions with the Escort III complex. Out of 73 PTPN23 orthologs analyzed, all but 6 harbor a conserved serine residue (VHCSGKXG) instead of an alanine residue (VHCSAGXG) in the PTP signature motif; it has been suggested that this Ala to Ser mutation in PTPN23 may influence the entry and positioning of the substrate in the catalytic pocket [99]. Besides, the D → K mutation in the WPD loop appears to contribute to Mop’s lack of observed catalytic activity [100]. Despite the abrogation of its catalytic activity, Mop has been reported to act as a tumor suppressor by influencing different signaling pathways, for example through the inhibition of receptor protein kinase signaling during their endocytic internalization [101], and by interfering with the transcription factor Kette [121]. More recently, a negative role for Ptp61F has been found in stem cell maintenance, as well as its inhibition by the transcription factor Ken [122], and a similar role has been reported for PTPN2 [123]. The deletion of Ptp61F produces a multitude of phenotypes reminiscent of a combination of those of mammalian PTPN1 and PTPN2 [124,125].

3.3.5. Ptpmeg and Pez

Early analysis of non-receptor PTP sequences identified three specific subtypes in the human genome, each exemplified by the MEG, PEZ and BAS enzymes [104]. In addition to their active catalytic PTP domain, all three contain an N-terminal FERM domain that is known to anchor proteins onto membrane-associated proteins involved in actin rearrangement and cytoskeleton dynamics. This classification scheme was mainly based on the absence of a PDZ domain in the PEZ enzymes, the presence of one PDZ domain in the MEG subfamily, and of multiple PDZ domains in the BAS subfamily. Such classification was retained in the annotation of the human PTPs by Andersen and colleagues with the subtype names NT5 (MEG), NT6 (PEZ) and NT7 (BAS) [55]. However, Edwards et al. noted that no orthologs of human BAS exist in Drosophila, a finding that we have confirmed here. Recent work has uncovered an important function for Drosophila Pez in intestinal stem cells by acting on the Hippo signaling pathway, which controls organ growth by regulating cell proliferation and apoptosis. In vitro studies have shown that the human ortholog (PTPN14) is an evolutionarily conserved regulator of this pathway [105,106], as well as a regulator of the lymphatic system and cholangiocyte development [107]. PTPN21, the second human ortholog of Drosophila pez, controls proliferation [108]. Recently, the group of Norbert Perrimon have produced a high-resolution interactome of the Hippo pathway by using existing components of the pathway as TAP-tagged baits, analysed by mass spectrometry [109].

Ptpmeg is involved in neuronal circuit formation in the central brain of the fly, regulating both the establishment and stabilization of axonal projection patterns, and its vertebrate orthologs (PTPN3 and PTPN4) are expressed in the nervous system too [110]. The human orthologs of Ptpmeg (PTPN3 and PTPN4) are also negative regulators of T-cell receptor (TCR) signaling (a function absent in the fly), both dephosphorylating the TCRβ chain [111]. Additionally, PTPN4 is anti-apoptotic, as was shown by its promotion of the survival of glioblastoma cells, and has been proposed as a therapeutic target [112].

3.3.6. Ptp61F

Ptp61F is the fly ortholog of mammalian PTPN1 (PTP1B) [113] and PTPN2 (TC-PTP) [114], and plays fundamental roles in growth, life span and fecundity [115]. Structurally, Ptp61F contains a PTP catalytic domain, 2 proline-rich motifs (PPXXPXP) and, depending on the splice isoform, a C-terminus with either a nuclear-targeting motif (isoform Ptp61Fm), or with a hydrophobic stretch that localizes Ptp61F to the cytoplasm (isoform Ptp61Fs) [116]. Interestingly, the multiple functions attributed to Ptp61F relate to many of the composite functions of mammalian PTPN1 and PTPN2. For instance, Ptp61F has been found to be both a negative modulator of the JAK-STAT pathway in Drosophila [117,118] and mouse [119], and an important modulator of insulin signaling through dock, the fly ortholog of NCK [120]. Ptp61F has also been shown to control the organization of F-actin via the regulation of Kette [121]. Further analysis of its role in the regulation of MAP kinase signaling [103,126,127].

3.4. Class I PTPs: dual-specificity phosphatases (DSPs)

The DSPs constitute the largest and most diverse group of PTPs whose defining feature is their ability to dephosphorylate both tyrosine and serine/threonine residues within the same substrate. However, we now know that DSPs target a much larger and diverse set of substrates, including phospho-serine, phospho-threonine and phospho-tyrosine residues, phosphoinositide lipids, RNA 5′-triphosphate, and carbohydrates. Although both classical PTPs and DSPs rely on an essential catalytic cysteine residue located at the base of the catalytic cleft, the catalytic pocket of DSPs is generally broader and shallower than that of classical PTPs, which explains why DSPs can accommodate more than one phosphorylated residue [128]. The 21 DSP genes of Drosophila can be classified into the same seven distinct human families (MKP, Atypical, Slingshot, PRL, CDC14, PTEN and Myotubularins), and therefore are a streamlined version of the 61 human DSP genes.

3.4.1. MKPs (mitogen-activated protein kinase phosphatases)

MKP family enzymes are defined by their specific ability to dephosphorylate both the phospho-serine and phospho-tyrosine residues in the activation loop (TxY) of MAP kinases (MAPK), a dual phosphorylation requirement that is essential for the catalytic activation of MAPKs by MAP kinase kinases (MEK). MAPKs are typically part of three-tiered pathways (MAPK/MEK/ERK) activated by a variety of stimuli (e.g. growth factors, cytokines, mitogens, osmotic stress) to enact a diversity of responses (e.g. gene expression changes, cellular proliferation/differentiation, inflammation and apoptosis). The human genome harbors 14 genes encoding...
MAPks (divided into the ERK, JNK, p38 and NLK sub-families) compared to only 7 in Drosophila [2]. The dephosphorylation of both phospho-residues of the MAPK activation loop by MKPs results in the complete abolition of MAPK catalytic activity. However, the partial deactivation of MAPks can be achieved by dephosphorylating the phosphotyrosine residue of the activation loop of MAPks by classical PTPs, including PTPN5 (STEP), PTPN7 (HepPTP) and PTPRR (PTP-SL) [129]. Interestingly, the PTPNS/PTP7/PTPRR group maps to a single gene in Drosophila (CG2327). Therefore, both classical PTPs and DUSPs are essential regulators that functionally overlap to fine-tune the activities of MAPK-dependent signaling pathways.

MKPs have been shown to shape the duration, magnitude and spatial organization of MAPK activities, with some MKPs being specific towards a single MAPK while other MKPs are able to regulate multiple MAPK pathways within the same cell [130]. Whereas 11 MKP genes have been described in human, only two (Mkp3 and pucked) exist in the fly, and thus Drosophila might be an ideal system to investigate the mechanisms that endow MKPs with spatiotemporal and quantitative control over MAPK responses and cellular outcomes. Mkp3 (the fly ortholog of human DUSP6, DUSP7 and DUSP9) was originally characterized as a negative feedback regulator of EGFR signaling [131], although today we know that Mkp3 is one of many such negative regulators of this essential pathway [132]. Physiologically, Mkp3 is a negative modulator of innate immunity in the fly gut, without which a strong immune response to the host gut microflora would occur [133]. The Mkp3 ortholog DUSP6 has been shown to play roles in both the promotion and inhibition of apoptosis. The pro-apoptotic function of DUSP6 was shown in a human colorectal cancer cell line where the authors demonstrated that DUSP6 acts by dephosphorylating ERK upon activation by P53 [134]. Conversely, expression studies on human thyroid tumor samples point to a role in the promotion of tumorigenesis [135]. Overexpression of human DUSP7 is observed in leukemias, and a role in maintaining pluriotropy has recently been identified in mouse embryonic stem (ES) cells [136–138]. Similarly, DUSP9 also maintains pluripotency in ES cells [139], and additionally DUSP9 polymorphisms are associated with type 2 diabetes [140,141]. Pucked (Puc) is a negative regulator of JNK [142] and was named after its mutant phenotype in the fly’s eye. Its human ortholog, DUSP10, is also known as Jun N-terminal kinase phosphatase in human, an important down-regulator of the JNK pathway. DUSP10 is also indirectly a positive regulator of ERK in a JNK/p38 dependent manner [143]. In inflammation, DUSP10 inhibits chemokine expression and is downregulated by ASC/PYCARD, a key adaptor protein in the inflammatory response [144].

3.4.2. Atypical DUSPs

Drosophila CG7378, CG10089, CG13197 and CG15528 are atypical DSPs that we describe here for the first time. The human orthologs of CG7378 (DUSP3, DUSP13, DUSP26, DUSP27 and DUSP1) are involved in a multitude of processes from cell cycle regulation to spermatogenesis [145,146]. Therefore ascribing a clear function to CG7378 is not straightforward. CG10089 is the fly ortholog of human DUSP15 and DUSP22. DUSP15 displays high levels of expression in human testis, as does CG10089 in the adult fly testis [145]. DUSP22 is expressed in the testis as well as in other many other tissues [145]. Since CG10089 is paralogous to Mkp, it might have a similar function in the negative regulation of JNK activity. CG13197, the fly ortholog of human DUSP11, is a 5′-triophosphatase that preferentially binds RNA, like its paralog mRNA-cap [147,148]. DUSP11 also binds splicing complexes, its expression is controlled by P53 [149], and may even play a role in inflammatory bowel disease [150].

MAP kinase-specific phosphatase (Mkp) is a negative regulator of the JNK pathway [151] and the fly ortholog of DUSP19. DUSP19 has also been shown to regulate JNK signaling by binding the activator MKK7, which in turn binds JNK, thus enabling DUSP19 to dephosphorylate JNK [145]. It has been suggested that DUSP19 might act as a scaffold protein given that DUSP19-mediated MKK7 inhibition is not dependent on the catalytic activity of DUSP19 [145]. Aberrant regulation of DUSP19 may also play a role in malignant pleural mesothelioma as syndecan-1 (a heparin sulfate proteoglycan), which is aberrantly expressed in malignant pleural mesothelioma and regulates proliferation in a highly complex way, is known to regulate negative DUSP19 expression [152].

MAPK Phosphatase 4 (MKP-4), like puc, is involved in the inhibition of JNK activity upon Gam-negative peptidoglycan-induced activation, thus being part of the Drosophila innate immune system [153]. Overexpression of MKP-4 has also been shown to negatively regulate ERK and p38 MAPks in Drosophila S2 cells, where MKP-4 is constitutively expressed [153]. The human ortholog, DUSP12, harbors a C-terminal C2H2 zinc finger domain, which plays a role in ribosome biosynthesis and cell cycle regulation [154,155]. The yeast ortholog, YVH1, lacks the zinc finger domain like MKP-4 but is equally involved in the assembly of the 60s ribosomal subunit, glycogen metabolism, gametogenesis and vegetative growth [156,157]. Furthermore, human DUSP12 overexpression has been reported in chronic myeloid leukemia, and DUSP12 genetic polymorphisms have been associated with type 2 diabetes [158,159]. Interestingly, the C. elegans ortholog, C24F3.2, participates in reproduction and fat storage [160,161]. Therefore, Drosophila MKP-4 might also be involved in ribosome biosynthesis, as well as being relevant for the study of type 2 diabetes, and its study might provide additional insights into the versatile human DUSP12.

The mRNA-capping phosphatase (mRNA-cap) is a highly conserved bi-functional enzyme with N-terminal 5′-triphosphatase catalytic activity and a C-terminal guanylyl transferase function. mRNA-cap recognizes the 5′ end of eukaryotic mRNA to add the 7-methylguanosine cap which is essential for translation to occur [162]. The human (RNGCT) and C. elegans (cel-1) mRNA-cap orthologs function similarly to their Drosophila counterpart. However, in fungi, the triphosphatase and guanylyltransferase activities are carried out by two separate proteins, Cet1p and Cgt1p respectively, with a study showing that they can be replaced by the single mammalian bi-functional enzyme [163].

PTEN-like phosphatase (Plip) dephosphorylates the lipid signaling molecule phosphatidylinositol 5-phosphate [164]. In Drosophila, Plip is localized to the mitochondria and plays a role in ATP production [165]. Its human ortholog, PTPM1, shows some functional similarities in that PTPM1 is also a mitochondrial phosphatase participating in the regulation of ATP production [164]. However, in humans this enzyme also regulates insulin release [166]. The C. elegans ortholog, F28C6.8, is implicated in growth and reproduction based on an RNAi screen that resulted in sterile worms that display a slow growth phenotype [167].

3.4.3. Slingshot phosphatases

Ssh is a cofillin phosphatase of the Slingshot family, with an important role in eye morphogenesis during the assembly of ommatidia (the basic units of the compound insect eye), where it works to inhibit actin polymerization by activating coflin [168]. In addition to its catalytic domain, Ssh, like its human orthologs (SSH1, SSH2 and SSH3), also contains a DEK domain N-terminal to the phosphatase catalytic domain, and C-terminal serine- and glutamine-rich regions. The human orthologs act through similar mechanisms, but they seem to be involved in wound healing by affecting keratinocyte motility [169], and therefore the original function of ssh might have diverged in metazoans, but maybe not the regulatory mechanisms as phosphorylation by Protein Kinase D (PKD) has recently been identified as a regulator of both human SSH1 and Drosophila Ssh [170].
3.4.4. PRLs (phosphatases of regenerating liver)

The Phosphatase of Regenerating Liver (PRL) is an oncogenic PTP family that was first identified in the regenerating rat liver [171]. Members of the PRL family can be prenylated, making it the sole PTP family containing the CAAX C-terminal prenylation motif, which is responsible for targeting PRLs to the cell and inner membranes [172]. This prenylation motif is also conserved in the human and C. elegans orthologs. PRL-1 is the only family member in Drosophila, and has been reported in the hearts of adult flies at days 7 and 14 in a recent proteome analysis [173]. The human orthologs of Drosophila PRL-1 (PTP4A1, PTP4A2 and PTP4A3) have been associated with cancer cell proliferation, metastasis, decreased survival time and poor post-operative outcome [174–177]. The recent characterization of PRL family members (by whole-mount immunostaining and in situ hybridization) during the early embryonic development of Drosophila, amphioxus and zebrafish showed that PRLs are uniformly expressed in the developing CNS and thus might be involved in early neural development [178].

3.4.5. CDC14 phosphatases

cdc14 is the fly ortholog of human CDC14A and CDC14B, which have been implicated in a diversity of cellular processes, including DNA repair, DNA damage checkpoint control, centrosome splitting, chromosome segregation and spindle assembly [179]. In Drosophila, Cdc14 has been reported to play a role in cytokinesis [180], and displays an affinity for proteins that have been phosphorylated by proline-dependent kinases such as CDC2, CDK5, ERK1 and P38 [181]. The catalytic site of Cdc14 is located at the C-terminus whereas the substrate specificity is determined by an N-terminal recognition pocket [181,182]. The yeast (CDC14) and C. elegans (cdc-14) orthologs also have functions in cell cycle control [183,184].

3.4.6. PTEN phosphatases

The phosphatase and tensin homologue deleted from chromosome ten (PTEN) is one of the most frequently mutated tumor suppressor genes, and 7 splice variant mRNAs and encoded polypeptides of the single Drosophila Pten gene have been described [185]. Since PTEN is such a powerful tumour suppressor, Pten mutants have recently been engineered to derive new Drosophila cell lines [186]. PTEN not only targets acidic substrates in proteins, but can also dephosphorylate the lipid second messenger phosphatidylinositol-3,4,5 (PIP3), thereby regulating the PI3K/AKT pathway since PIP3 activates AKT/catalytically active myotubularins, MTMR3 localizes to the ER and the Golgi [208], whereas MTMR4 is involved in early endosome sorting and also regulates BMP signaling [208,210].

3.4.7. Myotubularin phosphatases

Myotubularins are lipid phosphatases that act upon phosphatidylinositol 3-phosphate and phosphatidylinositol (3,5)-bisphosphate. Besides the catalytic domain, myotubularins are characterized by N-terminal PH-GRAM domains (gleckstrin homology glycosyltransferases, Gab-like GTPase activators and myotubularins), and each mammalian member of the myotubularin family acts on its own subset of lipid molecules and has its own specialized function [193]. Of the seven myotubularins identified in Drosophila (compared to 15 genes in human), three are most likely pseudophosphatases [194] (CG5026, CG14411 and Sbf, Fig. 2). Despite the important involvement of myotubularins in a multitude of human diseases, only three have been studied in detail (EDTP, Mtm and Sbf) [195]. EDTP was first identified in the egg of the flesh fly Sarcophaga peregrina and subsequently found to play multiple roles starting from oogenesis until early embryogenesis in the fly [196,197]. The mammalian ortholog of EDTP (MTMR14) is devoid of protein domains accessory to the catalytic domain (like EDTP), and is highly expressed in muscle tissue and implicated in the regulation of autophagy [195,198]. MTMR14 mutations have been linked to centronuclear myopathy [199]. Mtm and Sbf play important roles in macrophage cortical remodeling, endolysosomal homeostasis and wound healing, and are mutually dependent on each other, with Mtm being recruited and stabilized by the pseudophosphatase Sbf [200,201]. Similarly, human SBF1 interacts with MTMR2 to modulate its activity [202]. Drosophila mtm has three human orthologs (MTM1, MTM1R and MTMR2), which are believed to have undergone neo-functionalization events. For instance, MTMR2 is a regulator of late endocytosis, whereas MTM1 is more important in the early endosome [193,203,204]. Myotubularins display a propensity for dimerization. In mammals the orthologs of Drosophila (pseudophosphatase) myotubularins CG5026 (MTMR9) and CG14411 (MTMR12) have also been shown to bind catalytically active myotubularin family members indicating that these proteins likely function as adaptors [193]. Perhaps CG5026 and CG14411 also interact with catalytically active myotubularins, functioning as adaptors in a manner similar to Sbf. CG3530 is the fly ortholog of human MTMR6, MTMR7 and MTMR8. MTMR6 plays a role in apoptosis and in immunity by negatively regulating CD4+ T cell activation by inhibiting KCA3.1, a Ca++-activated potassium channel [193,195,205,206]. MTMR8 is involved in autophagy [207], CG3632 is the ortholog of human MTMR3 and MTMR4: unlike other myotubularins, MTMR3 localizes to the ER and the Golgi [208], whereas MTMR4 is involved in early endosome sorting and also regulates BMP signaling [209,210].

3.5. Class II PTPs: low molecular weight phosphatases (LMWPs)

First identified in the human liver [211], LMWPs harbor a bi-functional catalytic domain with both tyrosine phosphatase and arsenate reductase activities, and consequently target both phosphotyrosines and aryl phosphates [212,213]. Although yeast and mammals possess a single LMWP enzyme (ACP1) (Fig. 1A), 4 LMWP genes exist in Drosophila (primo-1, primo-2, CG14297 and CG31469), of which only 2 (primo-1 and primo-2) have been reported before (Fig. 2, Supplementary Table 1). Additionally, only two of these four genes are active phosphatases since both Primo-1 and CG31469 are pseudophosphatases. The primo genes are so named because of their expression in primary pigment cells of the pupal retina [214], although primo-2 is also highly expressed in the developing gut, head and central nerve cord axons, both in developing and adult flies. Primo-1 is expressed in all developmental stages except early embryos, and both primo genes are believed to play crucial roles in Drosophila neurogenesis [214]. CG14297 is highly expressed in...
the testis of the adult male fly, whereas the pseudophosphatase gene CG31469 has relatively low expression levels throughout the fly’s life (although its expression levels peak in the adult midgut and adult male head). All Drosophila LMWPs (primo-1, primo-2, CG14297 and CG31469) are orthologous to the ubiquitously expressed human enzyme red cell acid phosphatase gene (ACP1). ACP1 polymorphisms have been associated with hemolytic conditions such as favism and hemolytic anemia, which manifests in systemic lupus erythematosus [215,216]. In one colorectal cancer study, an ACP1 polymorphism was associated with cancer [217]. Given the divergent functions of the mammalian and Drosophila LMWPs, it is reasonable to assume that the functions of the fly enzymes are carried out by other phosphatases in mammals.

3.6. Class III PTPs: CDC25 phosphatases

The CDC25 family is represented in Drosophila by the paralogous genes string (stg) and twine (twe). string plays a role at the start of mitotic events during embryogenesis, oogenesis and regulates proliferation and stem cell maintenance, whereas twine plays a role in meiosis and oogenesis [218–222]. Furthermore, the degradation of the Twine protein as triggered by the onset of zygotic transcription has recently been found to be a critical switch that allows embryos to progress from the mid-blastula transition stage onto gastrulation. At this stage the cell cycle is remodeled from a short S phase and mitosis to the introduction of a longer S phase and a G2 phase [223]. Accordingly, the human orthologs CDC25A, CDC25B and CDC25C also play critical roles in mitosis [224], and their involvement in cancers has been demonstrated through their association with angiogenesis, invasion and survival time [225–228].
specialization within this subfamily: CDC25A has been shown to play a role in the G1/S transition, with CDC25A-blocked cells arresting in G1 [230,231]. Moreover CDC25A and CDC25B cooperatively regulate G2/M, with CDC25B being essential for the initiation of mitosis [232]. Expression of CDC25C is seen in all phases of the cell cycle [232]. Given the complexity of multicellular organisms, such specialization may provide additional checkpoints to maintain synchrony in cell division, with the different genes potentially complementing each other.

3.7. Aspartic acid-based phosphatases: Eye-Absent homologues (Eya)

Eyes-Absent Phosphatases (EyA) are haloacid dehalogenases that require a metal for catalytic activity and use aspartic acid as a nucleophile instead of the classic cysteine residue. *Drosophila Eya* is involved in numerous processes including the regulation of the anti-DNA immune response, retinal determination, spermatocyte development and oogenesis [233–238]. The human EYA phosphatases (EYA1, EYA2, EYA3 and EYA4) play crucial developmental roles in the eye, muscle and ear [239], and similar regulatory networks have been identified in *Drosophila*. For instance, the *Pax-Six-Eya-Dach* network is well-characterized in humans, and orthologs of these proteins exist in the fly [239]. Human Eya phosphatases can shift from the cytoplasm to the nucleus to regulate post-translational modifications through histone dephosphorylation [240], but *Drosophila Eya* appears to be cytoplasmic only [239]. The *C. elegans* ortholog (EYA-1) has been implicated in the regulation of apoptosis and embryogenesis, as well as in mesoderm determination into either striated muscle or a coelomocyte fate [241]. Worm EYA-1 is regulated by factors also present in the fly and human genomes [241].

4. Conclusions and future perspectives

Despite recent efforts on the characterization of the human tyrosine phosphatome [3,34], we lack detailed descriptions of the phosphatomes of simpler but equally relevant model organisms. The genome of *D. melanogaster* encodes 44 PTP genes, representing a streamlined version of the human tyrosine phosphatome. None of the fruit fly PTPs is lineage-specific and all genes have orthologs in human. Conversely, over 70% (78/109) of human PTPs have an ortholog in the fruit fly. Most importantly, fly and human PTPs share a high degree of functional conservation despite millions of years of independent evolution. *Drosophila* harbors enzymes of the same PTP families found in human: Transmembrane Classical PTPs (including all eight subtypes except R4), Non-Receptor PTPs (including all ten subtypes except NT4, NT7 and NT8), Dual-Specificity Phosphatases (MKP, Atypical, Slingshot, PRL, CDC14, PTEN, Myotubularin), Class II Cys-based PTPs (LMWP), Class III Cys-based PTPs (CDC25) and Asp-based PTPs (Eya). Of the fly PTP genes, 12 and 13 are in 1:1 and 1:2 orthology relationships with both mouse and human (i.e. 25/44 PTPs, or ~57%); 15 additional PTPs are in a 1:1 many relationship, and four are in a many:1 relationship. Moreover, five yeast PTPs and 30 worm PTPs have homologs in the fly, as well as in mouse and human (Supplementary Table 1).

Our analysis of the functional similarities and divergences, and orthology relationships between the PTPs of the fly and those of other model organisms suggests that *Drosophila* is an ideal modern organism for the study of mammalian PTPs. *Drosophila*’s streamlined tyrosine phosphatome should be more straightforward to study than those of mouse or human (where many PTP knockouts display mild effects, suggesting compensatory mechanisms), as it is much easier to generate compound knockout models in the fly. Over 75% of human disease-associated genes have a homolog in the fly [32], and this figure also holds true for human PTP-associated diseases [3]. Since the publication of the fly genome, *Drosophila* has emerged as a very successful model for a number of human diseases, including cancer, immunological, developmental and neurodegenerative disorders [242], as well as cardiac conditions [8].

Zebrafish has been favoured by a number of groups for the study of PTPs as this model fish shares many fundamental processes with human [243] and has the great advantage of producing transparent embryos (whose development can be followed by time-lapse microscopy). A fundamental disadvantage of zebrafish is that, although homologs of most human PTPs are found in the fish genome, the whole-genome duplication event that occurred in the teleost lineage ca. 320 million years ago resulted in a large number of duplicated PTP genes [244,245]. The implication is that compound knockouts and complicated analyses will be necessary to dissect sub-functionalization and neo-functionalization events, as well as compensatory mechanisms. Another popular model organism, the sea urchin, is a marine invertebrate widely used in marine, developmental, gene regulation and evolutionary biology studies [246]. Its PTP complement (n = 66 genes) is 50% larger than that of the fly, with substantial lineage-specific expansions unrelated to human enzymes [247], and thus is not as ideal as the fly for investigating human tyrosine phosphatases (see Fig. 3).

In the last 15 years protein kinases have arguably become the most popular class of drug targets for the pharmaceutical industry, not only because of their essential roles in regulating cellular processes but also because their druggability principles are well understood [248]. The old concept that the role of phosphatases is to counteract kinases in linear signaling pathways has become superseded by the emerging view that phosphatases play a dominant role in the dynamics of phosphorylation networks by introducing a multitude of negative and positive controls, which in turn control the amplitude and duration of the signal, thus influencing different sets of target genes [6]. Despite the recognition of the importance of PTPs in disease and the dominant effect of phosphatases over kinases, the development of PTP inhibitors has been hindered because PTPs are extremely difficult to drug with small molecules given their close structural relationship and the general hydrophilicity of the compounds identified by large pharmacological screens. The failures of the brute-force approach of high-throughput screening (HTS) for identifying small molecule inhibitors are well known since most hits thus identified cannot be used in vivo as they typically fall short of exhibiting all the desirable characteristics (i.e. low toxicity, absorption, metabolism, distribution and excretion). In this context, the degree of conserved biology and physiology between flies and humans makes *Drosophila* an invaluable model system to overcome some of the technical difficulties associated with the development of PTP inhibitors in disease models. Since *Drosophila* has high face validity with respect to human physiology and disease, primary drug screening could be performed directly in live flies to select high-quality hits that display desirable features such as bioavailability, metabolic stability and low toxicity [8]. This approach will enhance the rate of discovery of PTP inhibitors by reducing the time needed to identify a small collection of potentially more effective leads for subsequent validation by traditional HTS methods. This should be used in combination with genome-wide functional screens with cultured cells carrying pathway-specific reporters, which in the past have successfully identified new cellular pathways and drug targets [249]. Ultimately the chemical inhibition of phosphatases must be understood in terms of phosphatase-substrate networks to gain a systems-level understanding of the effect of specific inhibitors: in this context the group of Maja Köhn have applied a computational approach (by integrating information on phosphatases and their substrates, crystal structures, co-localization and co-
expression data) to derive protein phosphatase–substrate networks on which testable hypotheses can be generated [250]. Despite over 400 million years of independent evolution [251], the tiny fruit fly once again comes to the rescue of biochemists, geneticists and systems biologists to help us understand the essential roles of PTPs in controlling cellular networks and how these can be used for therapeutic purposes.

Acknowledgements

We are very grateful to Drs. Talia Vollk (Weizmann Institute) and Acaimo González-Reyes (CABD, Spain) for their critical reading of the manuscript. We also thank Mr. Shaq Liu for excellent system support and Ms. Mineko Tanimoto for secretarial assistance. TH is a holder of the Samuel Lupovitch Graduate Excellence Fellowship. This work was supported by the WPI Immunology Frontier Research Center. MLT holds the Jeanne and Jean-Louis Lévesque Chair in Cancer Research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jfebslet.2015.03.005.

References


Sun, Z.H. and Bu, P. (2012) Downregulation of phosphatase of regenerating liver-3 is involved in the inhibition of proliferation and apoptosis induced by emodin 3, 1077–1081.


