



Review

The human phosphatase interactome: An intricate family portrait

Francesca Sacco^{a,1}, Livia Perfetto^{a,1}, Luisa Castagnoli^a, Gianni Cesareni^{a,b,*}

^a Department of Biology, University of Rome "Tor Vergata", Rome, Italy

^b Research Institute "Fondazione Santa Lucia", Rome, Italy

ARTICLE INFO

Article history:

Received 23 March 2012

Revised 8 May 2012

Accepted 8 May 2012

Available online 21 May 2012

Edited by Marius Sudol, Giulio Superti-Furga and Wilhelm Just

Keywords:

Human phosphatome

Phosphatase family classification

Substrate recognition specificity

ABSTRACT

The concerted activities of kinases and phosphatases modulate the phosphorylation levels of proteins, lipids and carbohydrates in eukaryotic cells. Despite considerable effort, we are still missing a holistic picture representing, at a proteome level, the functional relationships between kinases, phosphatases and their substrates. Here we focus on phosphatases and we review and integrate the available information that helps to place the members of the protein phosphatase super-families into the human protein interaction network. In addition we show how protein interaction domains and motifs, either covalently linked to the phosphatase domain or in regulatory/adaptor subunits, play a prominent role in substrate selection.

©2012 Federation of European Biochemical Societies. Published by Elsevier B.V.

Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Phosphorylation is a widespread post-translational modification governing signal propagation [1]. Indeed phosphorylation is an efficient mean to control cell response to internal and external cues: it is rapid, taking as little as a few seconds, it does not require new proteins to be synthesized or degraded and can be easily reverted. Protein phosphorylation plays a key role in controlling a variety of cellular processes, such as migration, proliferation, apoptosis, differentiation, metabolism, organelle trafficking, immunity, learning and memory [2–4]. Thus, it is not surprising that aberrant phosphorylation profiles correlate with disease conditions such as cancer, diabetes and neurodegenerative or inflammatory disorders [5–7]. Eukaryotic protein phosphorylation typically occurs on serine, threonine or tyrosine residues. Olsen et al. have found the distribution of pSer, pThr, pTyr sites in the human proteome to be around 79.3%, 16.9% and 3.8% respectively [8]. Furthermore approximately 17000 proteins have at least one annotated residue in the Phosphosite database [9]. Indeed protein kinases are one of the largest gene family in eukaryotes, making up about 2% of the genome [10,11]. The human genome encodes 518

protein kinases. 428 are known or predicted to phosphorylate serine and threonine residues, while the remaining 90 are members of the tyrosine kinase family [3,12]. By contrast, in the human genome there are only approximately 200 phosphatases, targeting phosphorylated proteins or lipids.

Over the past decades, much of the interest of the scientific community has focused on protein kinases, protein phosphatases being considered less interesting house-keeping enzymes playing a non-specific role in modulating phosphoprotein homeostasis. Recent findings, however, have led to the emerging recognition that protein phosphatases play key roles in setting the levels of tyrosine, serine and threonine phosphorylation in cells, thus participating in the regulation of many physiological processes, including cell growth, tissue differentiation and inter-cellular communication [5,13,14].

Many excellent comprehensive reviews have discussed the different phosphatase superfamilies, their evolution and the mechanisms underlying substrate recognition specificity [13–17,12]. Despite the progress, identification of functional *in vivo* substrates remains a challenge. As a consequence, we are still missing a holistic picture representing, at a proteome level, the functional relationships between phosphatases and substrates. In this short contribution, we propose a phosphatase classification and report a comprehensive analysis of our current understanding and coverage of the phosphatase interactome. Finally, we discuss how this intricate network of interactions can help to define the function of poorly characterized phosphatases and to identify their substrates.

Abbreviations: PTP, protein tyrosine phosphatase; LP, lipid phosphatase; PPP, phosphoprotein phosphatases; PPM, metallo-dependent protein phosphatase; HAD, haloacid dehalogenase; RS, regulatory subunit

* Corresponding author.

E-mail address: cesareni@uniroma2.it (G. Cesareni).

¹ These authors contributed equally.

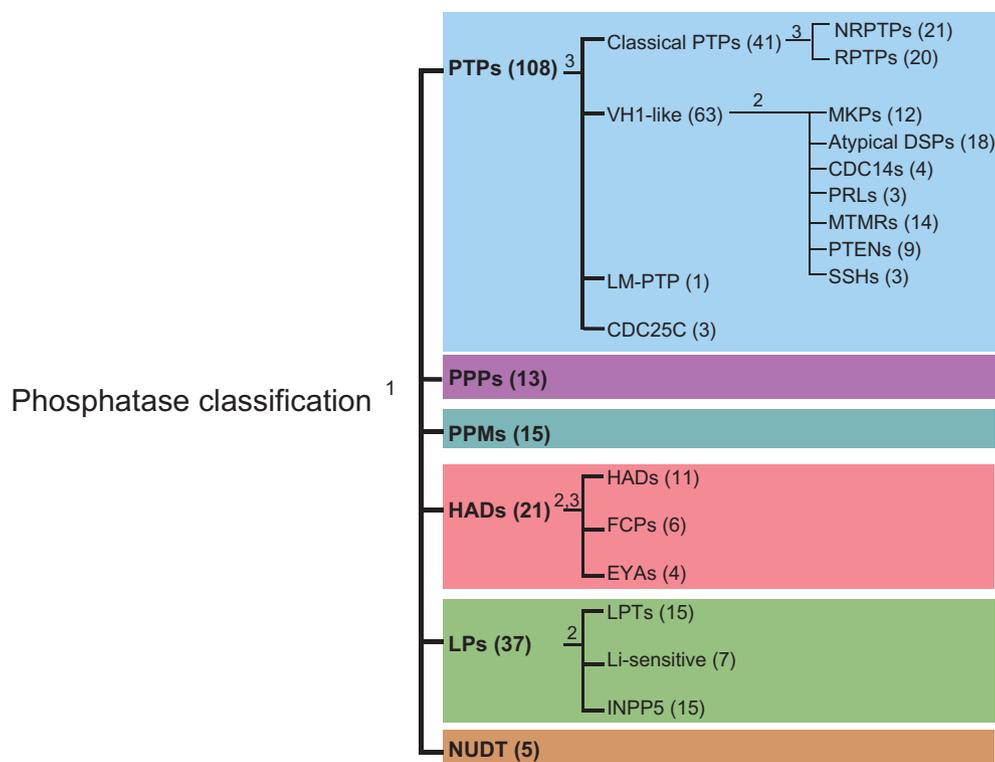


Fig. 1. Classification of protein phosphatase superfamilies. Protein phosphatases were first classified into six different families according to the catalytic domain InterPro annotation (1). Next each phosphatase family was further subdivided into different classes according to their preferred substrates (2) or literature annotation (3). The number of phosphatases in each family or class is in parenthesis.

2. Phosphatase classification

2.1. A complete compendium of phosphatase domains

A number of reports have discussed in detail the different phosphatase superfamilies. Here we aim at a catalogue of all the enzymes removing a phosphate group from proteins or lipids, including their accessory subunits. To obtain such a comprehensive list, we first screened the literature and protein databases to retrieve a list of 250 proteins 194 of which contain a phosphatase catalytic domain, while the remaining 56 were classified as regulatory subunits (RSs).

Next, to achieve proteome wide coverage of the proteins containing phosphatase domains we used a three step procedure: (i) we first recovered from InterPro [18] the sequences of the catalytic domains and aligned them with the ClustalW2 program [19]; (ii) clustering analysis allowed the classification of the phosphatase domains into 13 different subgroups; (iii) each subgroup alignment was used for a PSI-BLAST search, against the human proteome [20]. Hits with a *p-value* <0.001 were either retained in the group they were already assigned to, or added as new entries to our phosphatase compendium. This strategy resulted in a collection of 211 Phosphatase catalytic domains distributed in 199 proteins.

3. Phosphatase classification

The 211 phosphatase domains in the compendium were next assigned to 6 families defined by catalytic domain sequence similarity, after taking into account InterPro annotations:

- PTP membership is defined by the protein-tyrosine phosphatase domain (IPR016130);

- PPP by the serine/threonine-specific protein phosphatase domain (IPR006186);
- PPM by the Protein phosphatase 2C-like domain (IPR001932);
- HAD by the Haloacid dehalogenase-like hydrolase domain (IPR005834) or HD domain (IPR023279);
- LP by one of the following three domains: phosphatidic acid phosphatase (IPR000326), inositol monophosphatase (IPR000760) and inositol polyphosphate-related phosphatase (IPR000300);
- NUDT by the NUDIX hydrolase domain (IPR000086).

These six main families were further subdivided into classes, according to different criteria: sequence homology in the catalytic domain, substrate specificity and literature annotation (Fig. 1). The 211 catalytic domains captured by this procedure were aligned with ClustalW2 and the resulting sequence similarity tree is shown in Fig. 2. This graphic representation should not be interpreted as representing an evolutionary relationship between the different phosphatase superfamilies.

4. Catalytic specificity and substrate selection

Phosphatases have been considered promiscuous enzymes, displaying little intrinsic substrate specificity when assayed *in vitro*. Some published evidence, on the contrary, indicates that they may show remarkable preference for specific substrates *in vivo*. The reaction rate for a specific substrate is defined by the Michaelis Menten equation linking the rate of enzymatic reaction to substrate concentration and to the rate constant k_{cat} . Thus, enzymatic specificity may be achieved by increasing the ratio of the catalytic activity of the enzyme toward physiological substrates over the “background” catalytic activity for similar non-physiological substrates. Such intrinsic catalytic specificity can be obtained by shaping the

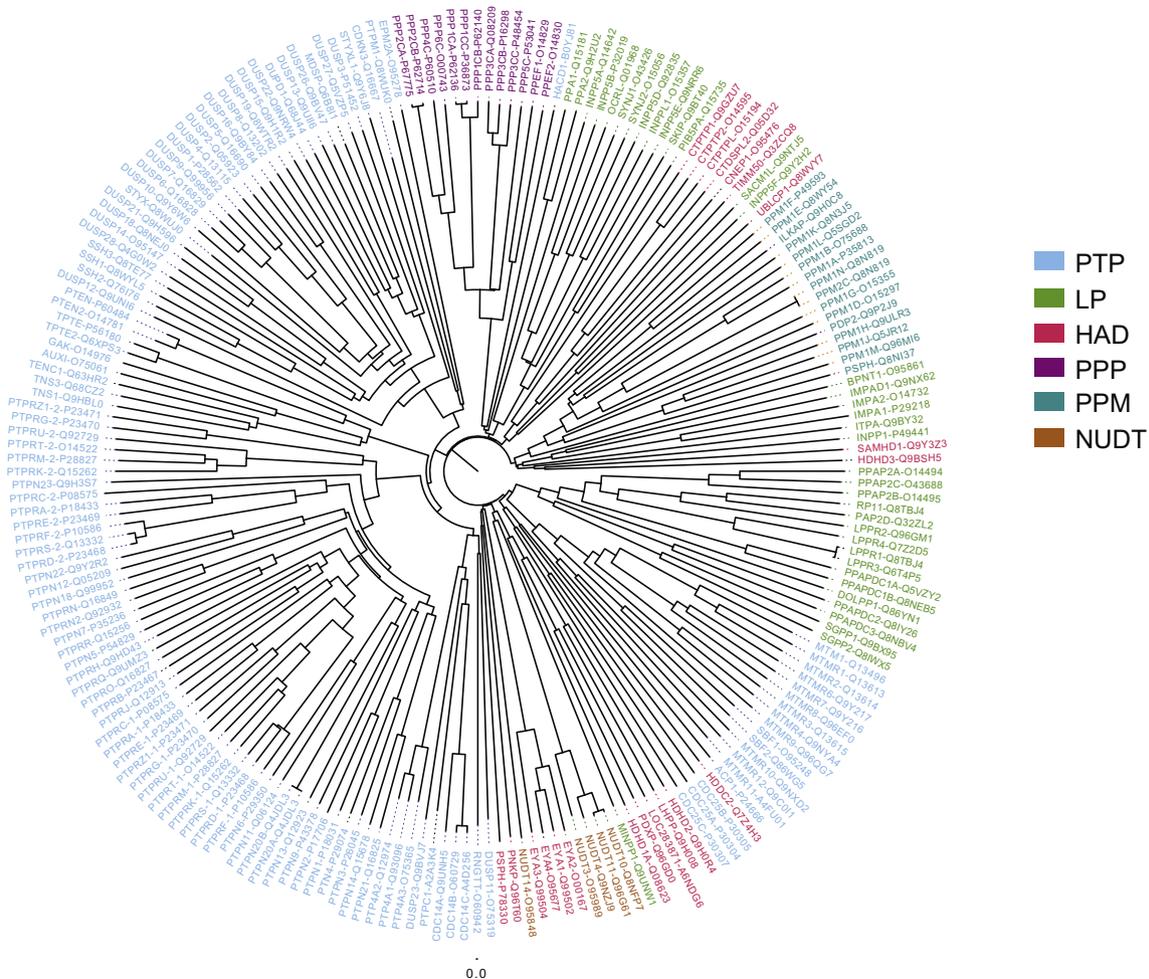


Fig. 2. Sequence similarity tree of the phosphatase catalytic domains. The tree illustrates the sequence similarity of domain families that do not have a common ancestor and should not be interpreted as an evolutionary tree. The phosphatase names are colored according to the classification in Fig. 1 and the classification of phosphatase catalytic domain. The tree chart was created by the FigTree software using as input a multiple sequence alignment, generated by ClustalW2 [19] software using a PAM protein weight matrix, a 25 gap open penalty value and a 0.20 gap extension penalty value.

chemical environment of the substrate binding pocket via side chain substitutions. Alternatively, protein interactions mediated either by adaptor proteins or accessory domains covalently linked to the phosphatase domain, might favor physiological substrates selection by increasing their relative local concentration.

The two highly homologous proteins PTPN11/SHP-2 and PTPN6/SHP-1 offer a clear example of this combined strategy. PTPN11 uses its amino-terminal SH2 domains to regulate activity and to increase concentration in specific cellular compartments. As a consequence, a catalytically active PTPN11 phosphatase, carrying a 65 amino acids deletion in the region encoding for the SH2 domains, cannot activate MAPK signalling under a variety of stimulation conditions [21]. However, domain-swapping experiments have provided evidence that the main determinant of substrate specificity resides in the catalytic domain. For instance, PTPN11 promotes FGF-2 induced animal cap elongation in *Xenopus* oocytes, but a chimeric derivative, where the catalytic domain is replaced by the PTP domain of PTPN6, fails to do so [22]. On the same line, only chimeras containing the PTPN6 PTP domain can dephosphorylate the EGFR [23].

To date, no consensus identifying the preferred phosphopeptide sequence contexts for Ser/Thr phosphatases could be identified. We have carried out an extensive curation of the phosphopeptide substrates reported in the literature without being able to show any statistically relevant position specific sequence preference (Liberti

et al., submitted). Several reports, on the other hand, have indicated that the catalytic domains of classical tyrosine phosphatases, when probed in vitro, display an intrinsic, albeit somewhat weak, preference for phosphorylated tyrosine residues embedded in specific sequence contexts [13,24–27]. We have recently used a new approach based on high density phosphopeptide chips to probe with trapping mutants [28] the substrate preference of the classical PTP against most of the phosphopeptides in the human proteome (Palma et al., in preparation). By this approach we have been able to show that most classical tyrosine phosphatases display an intrinsic substrate preference. From the alignment of the in vitro peptide substrates, it was possible to derive position specific scoring matrices that were used to infer putative substrates of the PTPRJ and PTPN1 phosphatases [29,30]. However, not all the peptides matching these weak consensi are targeted by the phosphatases in vivo, supporting the notion that the enzymatic domains must be guided to their functional substrates via a network of interactions [13].

5. Adaptors and docking sites

Phosphatases use three main strategies to physically target their substrates as schematically illustrated in Fig. 3. The three mechanisms contribute to phosphatase substrate recognition, although the relative importance of the three in different phosphatase families remains to be established.

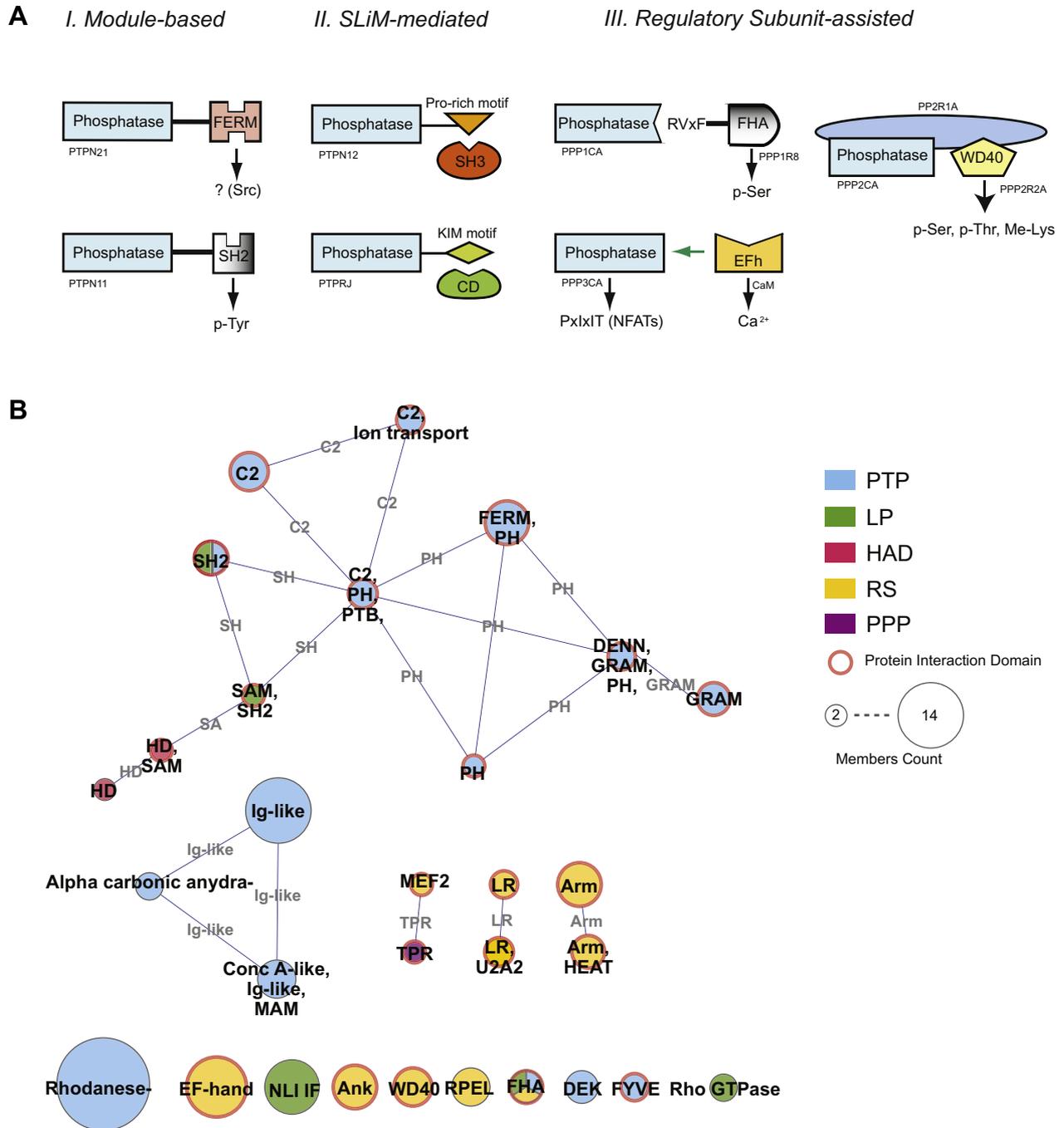


Fig. 3. Accessory module organization in the phosphatase proteins. (A) Cartoon of substrate recognition via three main strategies. *Module-based*: the catalytic domain is fused to an accessory domain that targets the substrate; *SLiM-mediated*: the phosphatase contains one (or more) short linear motif (SLiM), which is recognized by a binding domain in the substrate; *Regulatory Subunit-assisted*: the catalytic and the targeting domains are in separate chains and act in concert to target substrate dephosphorylation. (B) The nodes in the graph representation indicate different domain architectures, that is, proteins that have the same domain composition. Node size is proportional to the class member-count, node color matches the classification of class members according to the legend in the lower part of the Figure. Red borders identify classes containing at least a protein interaction domain. Domain architecture classes that share one domain are linked with an edge labeled with the shared domain.

One strategy employs targeting domains that are covalently linked to the catalytic domains (Fig. 3A, strategy I). For instance, as already mentioned, PTPN11 function depends on the integrity of the amino terminal SH2 domains [21]. In addition, PTPN11 binds its substrate GAB1 via SH2-mediated recognition of phosphorylated tyrosines on the C-terminus of GAB1. Removal of the PTPN11 SH2 binding sites on GAB1 results in hyperphosphorylation of GAB1 [31]. As another example, the FERM domain of the PTPN21 phosphatase is necessary to recruit the SRC kinase. After formation of the FERM-mediated complex SRC is activated via dephosphorylation

of its carboxyterminal inhibitory phosphotyrosine. Carlucci et al. have demonstrated that the expression of PTPN21 lacking either catalytic activity or the FERM domain cannot promote SRC activation and cell motility is reduced [32]. As another example, the TC48 and TC45 isoforms of PTPN2 are characterized by different carboxy-termini that are responsible for subcellular localization and, as a consequence, modulate substrate targeting [33].

Some phosphatases exploit a complementary strategy and use Small Linear Motifs (SLiM) to dock into a substrate binding pocket (Strategy II in Fig. 3A). As an example the interaction of the SH3

domain of p130Cas with a proline rich motif in PTPN12 [34] favors its dephosphorylation by PTPN12 and by this mechanism modulates cell motility [35]. A similar strategy is exploited by a class of dual specificity phosphatases (DSPs) that target MAPKs by direct and specific interaction of a short segment of the non-catalytic amino-terminal domain with a receptor surface in the MAPK. For example, the phosphatase MKP-3/Pyst1 forms a stable complex only with the ERKs, but not the JNKs or p38 [36]. The specificity of this interaction correlates with the selectivity of MKP-3 for dephosphorylation of the ERKs. The interaction between the two proteins also stimulates the catalytic activity of the MKP-3 phosphatase. One of the sites critical for the interaction of ERKs with MKP-3 is a 14-residues motif, dubbed KIM, which is also found in two classical tyrosine phosphatases PTP-SL and STEP, both regulating the phosphorylation status of ERK1/2 [37]. Also in this case, selectivity was ascribed to the 14-residues KIM that directly interacted with ERK proteins. More recently, we have shown that the trans-membrane receptor protein-tyrosine phosphatase Density-enhanced phosphatase-1 (DEP-1/PTPRJ) also targets ERKs via a juxta-membrane KIM motif [29]. Mutations either in the conserved ERK docking domain or in the PTPRJ KIM motif prevent ERKs dephosphorylation. The functional relevance of this interaction was confirmed by the observation that PTPRJ KIM-motif point mutations are frequent in breast cancer metastasis where ERK kinases are hyper-activated [38].

Finally, some phosphatases bind non-covalently to regulatory subunits or scaffold proteins that mediate substrate docking (strategy III in Fig. 3A). For instance, Cong and colleagues have shown that an adaptor protein PST-PIP1 bridges the PTPN12 phosphatase to its substrate, the ABL kinase, thereby promoting its dephosphorylation [39]. In another variation of the same theme, the catalytic subunits of the protein phosphatase 1 (PP1) contain docking sites that host short linear motifs (SLiM) of regulatory subunits that in turn direct the phosphatase to specific cellular locations and determine substrate specificity [16]. Also PPP3CA/Calcineurin requires a regulatory subunit to bind its substrates albeit by a somewhat different mechanism. In stimulatory conditions, both calcineurin regulatory protein B subunit and calmodulin bind calcium ions, displacing the auto-inhibitory fragment of the phosphatase domain from the catalytic cleft thereby activating the enzyme [40]. The PPP3CA substrate specificity has been extensively analyzed and two consensus recognition motif (PxIxIT and LxVP) have been identified in most of its substrates [41].

6. Association of phosphatase and protein interaction domains

These and other observations emphasize that phosphatases, similarly to kinases, use protein interaction domains and docking motifs to modulate the function of their catalytic domains and to select substrates. To estimate the generality of this strategy we analyzed the domain architecture of the phosphatase family.

To this purpose, we retrieved the list of InterPro domains associated with either phosphatases or regulatory subunits. To simplify the classification, similar InterPro domains sharing a common “parent” term were collapsed into the parent term. Approximately 50% (134 out of 255) of the members of the phosphatase family display one or more InterPro domains, while 121 have only the catalytic domain. The vast majority of members of this latter group are members of the PPP, PPM and HAD classes. These differences in domain association reflect different molecular mechanisms of substrate recognition. For instance, PPPs have evolved an adaptor-protein based strategy to target their substrates.

Following this mapping, phosphatases displaying the same domain architecture were grouped into classes. This module-based

Table 1

Functional Enrichment of the 24 most represented accessory domains. Domains were classified as “protein interaction domain” (according to Jin J. et al. [42]), “binding” (according to InterPro) or “other”; *p*-value was calculated by the DAVID web tool [49], using as input the list of all phosphatases and as background the whole proteome.

Domain type	Count	<i>p</i> -Value	Class	Role
Rhodanese-like	14	3,05E-18	PTP	Other
NLI interacting factor	7	2,94E-10	HAD	Binding
Fibronectin, type III	15	7,47E-04	PTP	Protein interaction domain
RPEL repeat	4	1,12E-04	RS	Binding
GRAM	5	1,61E-04	PTP	Protein interaction domain
U2A'/phosphoprotein	4	6,65E-04	RS	Other
FERM	5	9,98E-04	PTP	Protein interaction domain
DEK	3	1,32E-03	PTP	Protein interaction domain
MAM	4	2,31E-03	PTP	Protein interaction domain
SH2	7	4,52E-02	PTP,LP	Protein interaction domain
EF-HAND	8	8,74E-02	RS	Protein interaction domain
PH	9		PTP	Binding
Ig-like	15		PTP	Protein interaction domain
C2	9		PTP	Protein interaction domain
Arm repeat	8		RS	Protein interaction domain
LR repeat	6		RS	Protein interaction domain
Ank repeat	5		RS	Other
Concavalin A-like	4		PTP	Protein interaction domain
WD40	4		RS	Protein interaction domain
FHA	3		PTP,HAD,RS	Protein interaction domain
HEAT	3		RS	Protein interaction domain
Nucleic acid binding	3		PTP,LP	Binding
PTB	3		PTP	Protein interaction domain

classification defined 40 groups, 10 of which have only one member. In Fig. 3B, we have represented the phosphatase domain relationships as a graph where domain architectures, that are common to more than one phosphatase, are represented as nodes and nodes that share at least one domain are linked by an edge. The node size is proportional to the number of phosphatases that are grouped in the architectural category, while the node-color describes class affiliation.

In Table 1, we have reported for each domain the total number of occurrences in the phosphatase family and the *p*-value of observing such an occurrence by random sampling. Interestingly, most phosphatase-associated domains are annotated as linked to protein interaction by domain experts [42] or to “binding” by InterPro and their occurrence in the family is often higher than random. This observation further supports the role of these additional structures in targeting potential substrates.

The most populated class of domain architectures is represented by the 14 MAPK phosphatases of the DSP family that are associated to a Rhodanese-like domain. Although Rhodanese like motifs are not annotated as “protein binding” modules, it is tempting to speculate that, in this enzyme family, they have evolved into MAPK docking modules. Interestingly, despite being one of the most numerous protein binding domain in the human proteome, the SH3 domain is not found associated to phosphatases or regulatory subunits.

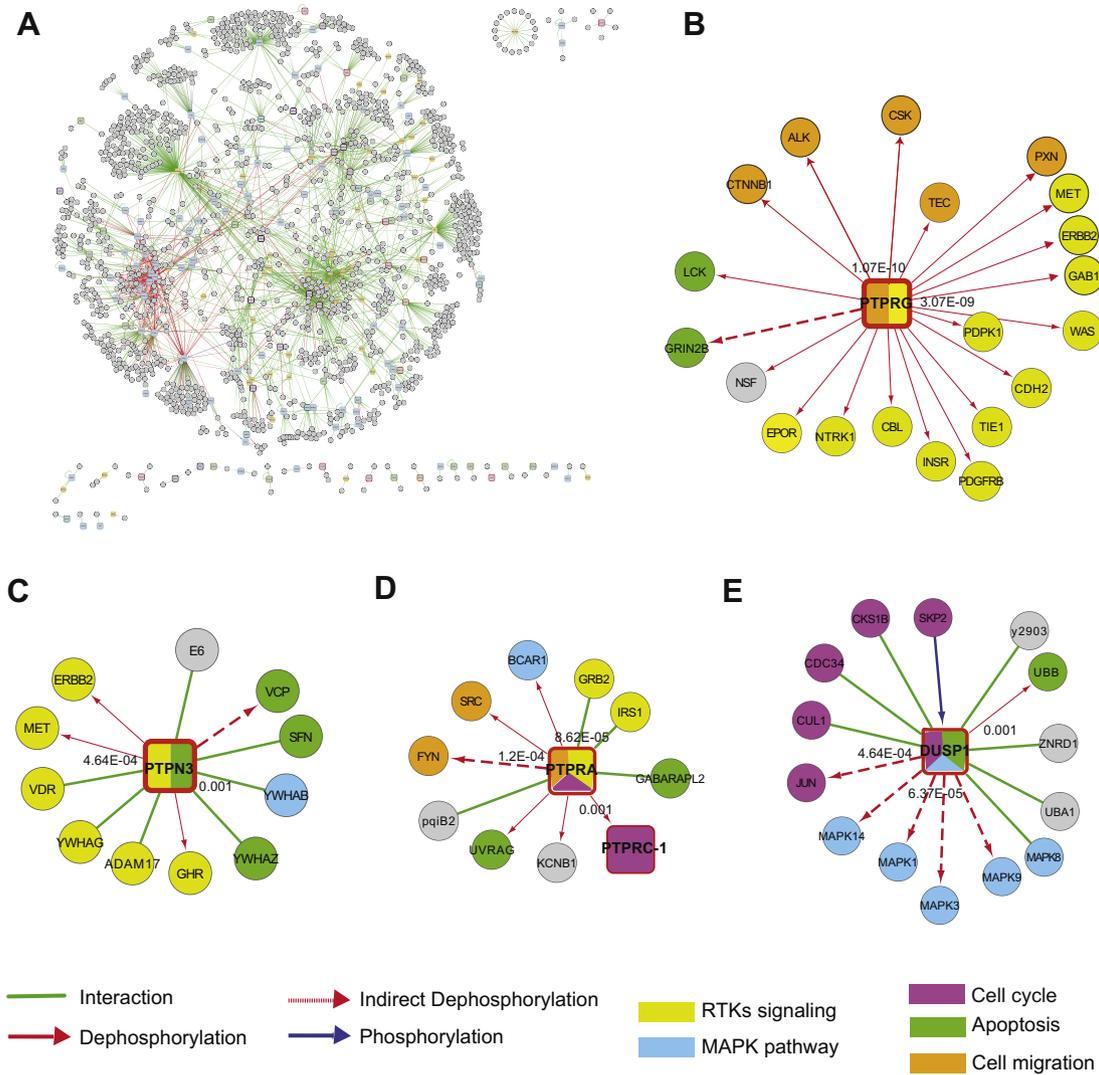


Fig. 4. The human phosphatase interactome. (A) Graph representation of the human phosphatase interactome. 2496 protein–protein interaction pairs involving one of the phosphatase proteins were retrieved from the HomoMINT and IntAct databases [43,44]. The network was generated with the Cytoscape graphic software. Squared nodes represent phosphatases, while rounded grey nodes represent their targets. The phosphatase color code reflects class affiliation according to the color code convention used throughout this review. Interactors (circles) of “cancer PTPs” whose biological process is not annotated in UniProtKB with GO terms (B) PTPRG-B, (C) PTPN3, (D) PTPRA and (E) DUSP1 were extracted from the global phosphatase interactome and marked in different colors, according to their functional association with GO-Biological Process terms. The phosphatase square nodes are labeled with sectors whose color matches that of the GO-Biological Process term that was significantly overrepresented in the phosphatase interactors and substrates (p -value < 0.002). Edges are colored according to the functional relationships between the nodes they connect: physical associations, dephosphorylations in red and phosphorylation reactions in green and blue respectively. Continuous and dashed lines represent direct (demonstrated in vitro with purified proteins) and indirect dephosphorylation (demonstrated in vivo), respectively.

7. Interactome

Thus, phosphatases and their regulatory subunits are well equipped with domains specialized in protein/peptide binding that integrate them in the functional protein interaction network. Exploration of this intricate web is likely to provide information about phosphatase function.

To estimate the coverage of the protein interaction network achieved so far, we undertook a curation effort to capture from the literature a large fraction of the interactions in which at least one of the partners is one of the 255 phosphatases or regulatory subunits in our compendium. This interaction list was merged with those that could be downloaded from the HomoMINT and Intact databases [43,44]. The interaction information we collected contributed to assemble an intricate and highly connected network consisting of 1555 nodes and 2496 edges (Fig. 4A).

While the majority of the phosphatases could be integrated by this strategy in the human interaction network, for 25 PTPs, 1 PPP

and 9 RSs we were not able to retrieve any interaction information. 380 of the 2496 edges in the interactome graph represent “dephosphorylation” reactions. Most of these (345) link tyrosine phosphatases to their substrates, while we could only recover information about 35 dephosphorylation reactions mediated by PPPs.

8. Guilt by association

Next we asked whether the phosphatase interaction network could be used to obtain information about the function of poorly characterized phosphatases. Recently, Julien et al. reviewed the functional role of protein tyrosine phosphatases in cancer [45]. For 35 members of this family there is evidence that they affect transformation and could be classified as oncogenic or oncosuppressor phosphatases, but for 18 of them we hardly have any information about the fulfilled role or molecular mechanism [45]. To try to fill this gap, we asked whether the interactors and the substrates of these phosphatases could provide some hint about the pathways

that are affected by phosphatase activity. The subnetwork of four of these (PTPRG, PTPN3, PTPRA and DUSP1), once extracted from the human phosphatase interactome, displayed a sufficient number of protein partners to encourage a GO term enrichment analysis (Fig. 4). Remarkably, these four phosphatases have been described as both oncogenes and oncosuppressors, depending on the cellular context. As illustrated in panel B and C of Fig. 4, PTPRG and PTPN3 partners are both significantly associated to the RTKs signaling GO-term, suggesting a possible involvement in cancer onset and progression. Interestingly, we also found that PTPRG partners, as well as those of PTPRA, are significantly associated to cell migration pathways, which are highly correlated to metastasis progression [46]. Indeed PTPRA has been previously described as an oncogene, since it is over-expressed both in colon and head and neck cancer [45]. The oncogenic properties of PTPRA are mediated by its direct dephosphorylation of the SRC kinase on Y527, which increases FAK activity and consequently cell migration [47]. It has been shown that PTPRA is significantly down-regulated in breast cancer, where it acts as tumor suppressor, affecting signaling pathways that are still unknown and attenuating GRB2 mediated signaling [45,48]. Our analysis reveals that PTPRA is also significantly associated to RTK signaling as well as apoptosis, suggesting that these could be the potential mechanisms by which this tyrosine phosphatase acts as tumor suppressor. Finally, DUSP1 has been correlated with both oncogenic and onco-suppressor functions depending on the cellular context: in ovarian and liver cancers it is down-regulated, while in a variety of tumors (Head and neck, breast, gastric, pancreatic, bladder, ovarian, prostate, lung and CRC) it is over-expressed [45]. The DUSP1 onco-suppressor property relies on its ability to dephosphorylate the MAPK family members ERK and JNK, as confirmed by our analysis. In addition, as shown in panel E of Fig. 4, DUSP1 partner annotation is significantly enriched in terms pertinent to cell cycle regulation.

9. Conclusions

We have reviewed published evidence suggesting that phosphatases select their substrates by a combination of enzymatic specificity (e.g., tyrosine phosphatases) and spatial organization, mediated by the protein interaction network that modulates subcellular compartmentalization and substrate docking. This emphasizes the importance of compiling a reliable and complete list of phosphatase interacting partners in order to correctly place the phosphatase family in the context of the cell interactome. By combining the data annotated in protein interaction databases and a new curation campaign, we have been able to recover ~2500 interactions involving as many as 185 phosphatases out of the 255 encoded in the human proteome. These data can now be explored in a publicly available database (hupho.uniroma2.it). Systematic studies, which are underway, will soon contribute to extend the coverage of the phosphatase interactome. This effort will be instrumental in advancing our understanding of the rules governing substrate selection by this enzyme family and provide a unique tool for making sense of phosphoprotein homeostasis in the cell. Furthermore, a better understanding of the interactions, governed by non enzymatic domains, may enable the design of selective inhibitors to disrupt the physical interaction of a phosphatase or a targeting subunit with a specific substrate.

We have taken advantage of the current incomplete version of the phosphatase interactome and combined it with the characterization of the peptides substrate preference of PTP, as determined by a high density peptide chip approach, to develop a strategy to infer new phosphatase substrates [30]. This is based on the Bayesian integration of two types of evidence: (i) substrate preference obtained by probing with phosphatase domains a large

number of phosphopeptides arrayed on a glass chip and (ii) WID (Whole Interactome Distance), a distance matrix detailing the distance between any pair of proteins in a weighted protein interaction graph. The approach was successfully used to infer new substrates of the PTPN1/PTP1B phosphatase. Some of these new substrates have been validated in vivo [30].

The full characterization of the peptide recognition specificity of the different members of the PTP family and the completion of projects aimed at a higher coverage of the interactions mediated by phosphatases will allow the application of this strategy to a larger number of tyrosine phosphatases.

Acknowledgments

This work was supported by Telethon (GGP09243), the Italian Association for Cancer Research (AIRC) and the FIRB project Oncodiet.

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.febslet.2012.05.008>.

References

- [1] Mann, M. and Jensen, O.N. (2003) Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21, 255–261.
- [2] Graves, J.D. and Krebs, E.G. (1999) Protein phosphorylation and signal transduction. *Pharmacol. Ther.* 82, 111–121.
- [3] Manning, G., Whyte, D.B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science* 298, 1912–1934.
- [4] Manning, G., Plowman, G.D., Hunter, T. and Sudarsanam, S. (2002) Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* 27, 514–520.
- [5] Tonks, N.K. (2006) Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* 7, 833–846.
- [6] Easty, D., Gallagher, W. and Bennett, D.C. (2006) Protein tyrosine phosphatases, new targets for cancer therapy. *Curr. Cancer Drug Targets* 6, 519–532.
- [7] Gee, C.E. and Mansuy, I.M. (2005) Protein phosphatases and their potential implications in neuroprotective processes. *Cell. Mol. Life Sci.* 62, 1120–1130.
- [8] Olsen, J.V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P. and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648.
- [9] Hornbeck, P.V., Kornhauser, J.M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V. and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270.
- [10] Caenepael, S., Charyczak, G., Sudarsanam, S., Hunter, T. and Manning, G. (2004) The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proc. Natl. Acad. Sci. USA* 101, 11707–11712.
- [11] Zhu, H. et al. (2000) Analysis of yeast protein kinases using protein chips. *Nat. Genet.* 26, 283–289.
- [12] Alonso, A. et al. (2004) Protein tyrosine phosphatases in the human genome. *Cell* 117, 699–711.
- [13] Tiganis, T. and Bennett, A.M. (2007) Protein tyrosine phosphatase function: the substrate perspective. *Biochem. J.* 402, 1–15.
- [14] Andersen, J.N., Jansen, P.G., Echwald, S.M., Mortensen, O.H., Fukada, T., Del Vecchio, R., Tonks, N.K. and Møller, N.P. (2004) A genomic perspective on protein tyrosine phosphatases: gene structure, pseudogenes, and genetic disease linkage. *FASEB J.* 18, 8–30.
- [15] Tonks, N.K. and Neel, B.G. (2001) Combinatorial control of the specificity of protein tyrosine phosphatases. *Curr. Opin. Cell Biol.* 13, 182–195.
- [16] Roy, J. and Cyert, M.S. (2009) Cracking the phosphatase code: docking interactions determine substrate specificity. *Sci. Signal.* 2, re9.
- [17] Moorhead, G.B., De Wever, V., Templeton, G. and Kerk, D. (2009) Evolution of protein phosphatases in plants and animals. *Biochem J.* 417, 401–409.
- [18] Hunter, S. et al. (2012) InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res.* 40, D306–D312.
- [19] Larkin, M.A. et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
- [20] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- [21] Shi, Z.Q., Lu, W. and Feng, G.S. (1998) The Shp-2 tyrosine phosphatase has opposite effects in mediating the activation of extracellular signal-regulated

- and c-Jun NH2-terminal mitogen-activated protein kinases. *J. Biol. Chem.* 273, 4904–4908.
- [22] O'Reilly, A.M. and Neel, B.G. (1998) Structural determinants of SHP-2 function and specificity in *Xenopus* mesoderm induction. *Mol. Cell. Biol.* 18, 161–177.
- [23] Tenev, T. et al. (1997) Both SH2 domains are involved in interaction of SHP-1 with the epidermal growth factor receptor but cannot confer receptor-directed activity to SHP-1/SHP-2 chimera. *J. Biol. Chem.* 272, 5966–5973.
- [24] Walchli, S., Espanel, X., Harrenga, A., Rossi, M., Cesareni, G. and Hooft van Huijsduijnen, R. (2004) Probing protein-tyrosine phosphatase substrate specificity using a phosphotyrosine-containing phage library. *J. Biol. Chem.* 279, 311–318.
- [25] Barr, A.J. et al. (2009) Large-scale structural analysis of the classical human protein tyrosine phosphatome. *Cell* 136, 352–363.
- [26] Zhang, Z.Y., Thieme-Seffler, A.M., Maclean, D., McNamara, D.J., Dobrusin, E.M., Sawyer, T.K. and Dixon, J.E. (1993) Substrate specificity of the protein tyrosine phosphatases. *Proc. Natl. Acad. Sci USA* 90, 4446–4450.
- [27] Zhang, Z.Y., Maclean, D., McNamara, D.J., Sawyer, T.K. and Dixon, J.E. (1994) Protein tyrosine phosphatase substrate specificity: size and phosphotyrosine positioning requirements in peptide substrates. *Biochemistry* 33, 2285–2290.
- [28] Blanchetot, C., Chagnon, M., Dube, N., Halle, M. and Tremblay, M.L. (2005) Substrate-trapping techniques in the identification of cellular PTP targets. *Methods* 35, 44–53.
- [29] Sacco, F. et al. (2009) Tumor suppressor density-enhanced phosphatase-1 (DEP-1) inhibits the RAS pathway by direct dephosphorylation of ERK1/2 kinases. *J. Biol. Chem.* 284, 22048–22058.
- [30] Ferrari, E. et al. (2011) Identification of new substrates of the protein-tyrosine phosphatase PTP1B by Bayesian integration of proteome evidence. *J. Biol. Chem.* 286, 4173–4185.
- [31] Holgado-Madruga, M. and Wong, A.J. (2004) Role of the Grb2-associated binder 1/SHP-2 interaction in cell growth and transformation. *Cancer Res.* 64, 2007–2015.
- [32] Carlucci, A. et al. (2008) Protein-tyrosine phosphatase PTPD1 regulates focal adhesion kinase autophosphorylation and cell migration. *J. Biol. Chem.* 283, 10919–10929.
- [33] Tiganis, T., Bennett, A.M., Ravichandran, K.S. and Tonks, N.K. (1998) Epidermal growth factor receptor and the adaptor protein p52Shc are specific substrates of T-cell protein tyrosine phosphatase. *Mol. Cell. Biol.* 18, 1622–1634.
- [34] Garton, A.J., Burnham, M.R., Bouton, A.H. and Tonks, N.K. (1997) Association of PTP-PEST with the SH3 domain of p130Cas; a novel mechanism of protein tyrosine phosphatase substrate recognition. *Oncogene* 15, 877–885.
- [35] Angers-Loustau, A., Cote, J.F., Charest, A., Dowbenko, D., Spencer, S., Lasky, L.A. and Tremblay, M.L. (1999) Protein tyrosine phosphatase-PEST regulates focal adhesion disassembly, migration, and cytokinesis in fibroblasts. *J. Cell Biol.* 144, 1019–1031.
- [36] Camps, M., Nichols, A., Gillieron, C., Antonsson, B., Muda, M., Chabert, C., Boschert, U. and Arkinstall, S. (1998) Catalytic activation of the phosphatase MKP-3 by ERK2 mitogen-activated protein kinase. *Science* 280, 1262–1265.
- [37] Pulido, R., Zuniga, A. and Ullrich, A. (1998) PTP-SL and STEP protein tyrosine phosphatases regulate the activation of the extracellular signal-regulated kinases ERK1 and ERK2 by association through a kinase interaction motif. *EMBO J.* 17, 7337–7350.
- [38] Ding, L. et al. (2010) Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 464, 999–1005.
- [39] Cong, F., Spencer, S., Cote, J.F., Wu, Y., Tremblay, M.L., Lasky, L.A. and Goff, S.P. (2000) Cytoskeletal protein PSTPIP1 directs the PEST-type protein tyrosine phosphatase to the c-Abl kinase to mediate Abl dephosphorylation. *Mol. Cell* 6, 1413–1423.
- [40] Perrino, B.A., Ng, L.Y. and Soderling, T.R. (1995) Calcium regulation of calcineurin phosphatase activity by its B subunit and calmodulin. Role of the autoinhibitory domain. *J. Biol. Chem.* 270, 340–346.
- [41] Roy, J., Li, H., Hogan, P.G. and Cyert, M.S. (2007) A conserved docking site modulates substrate affinity for calcineurin, signaling output, and in vivo function. *Mol. Cell* 25, 889–901.
- [42] Jin, J. et al. (2009) Eukaryotic protein domains as functional units of cellular evolution. *Sci. Signal.* 2, ra76.
- [43] Persico, M., Ceol, A., Gavrila, C., Hoffmann, R., Florio, A. and Cesareni, G. (2005) HomoMINT: an inferred human network based on orthology mapping of protein interactions discovered in model organisms. *BMC Bioinformatics* 6 (Suppl 4), S21.
- [44] Kerrien, S. et al. (2012) The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* 40, D841–D846.
- [45] Julien, S.G., Dube, N., Hardy, S. and Tremblay, M.L. (2011) Inside the human cancer tyrosine phosphatome. *Nat. Rev. Cancer* 11, 35–49.
- [46] Gabarra-Niecko, V., Schaller, M.D. and Dunty, J.M. (2003) FAK regulates biological processes important for the pathogenesis of cancer. *Cancer Metastasis Rev.* 22, 359–374.
- [47] Harder, K.W., Moller, N.P., Peacock, J.W. and Jirik, F.R. (1998) Protein-tyrosine phosphatase alpha regulates Src family kinases and alters cell-substratum adhesion. *J. Biol. Chem.* 273, 31890–31900.
- [48] den Hertog, J., Tracy, S. and Hunter, T. (1994) Phosphorylation of receptor protein-tyrosine phosphatase alpha on Tyr789, a binding site for the SH3-SH2-SH3 adaptor protein GRB-2 in vivo. *EMBO J.* 13, 3020–3032.
- [49] Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.