Phylogenetic Analyses of Proteins Coordinating G2 Size Control in Fission Yeast

Zsófia Nagy[1], Anna Medgyes-Horváth[1,2], Csilla Szalay[1], Matthias Sipiczki[3], Ákos Sveiczer[1*]

[1] Department of Applied Biotechnology and Food Science, Faculty of Chemical Technology and Biotechnology,
  Budapest University of Technology and Economics, H-1521 Budapest, P.O.B. 91, Hungary
[2] Department of Physics of Complex Systems, Institute of Physics, Faculty of Science, Eötvös Loránd University,
  H-1518 Budapest, P.O.B. 32, Hungary
[3] Department of Genetics and Applied Microbiology, Faculty of Science and Technology, University of Decrecen,
  H-4002 Debrecen, P.O.B. 400, Hungary
* Corresponding author, e-mail: asveiczer@mail.bme.hu

Abstract

Regulation of G2 phase is based on inhibition of MPF (M-phase Promoting Factor) through phosphorylation by Wee1-like kinases. Removal of the inhibiting phosphate group requires Cdc25-like phosphatases. In fission yeast, size control is achieved by monitoring cell length via interactions of Pom1, Nif1, Cdr1 and Cdr2 proteins, regulating MPF via the Wee1 kinase. Here, a search for homologues of these key proteins was performed in the genomes of several model organisms to analyze the evolution of G2 size control. Both the known upstream pathways regulating Wee1 protein (Pom1 → Cdr2, and Nif1 → Cdr1) have been found to be characteristic only in fission yeasts. Mik1, a backup copy of Wee1 kinase probably appeared in the common ancestor of the fission yeasts. The duplication resulting in Wee1A and Wee1B isoforms probably happened in a common ancestor of higher animals, while the Myt1 protein (found only in animals) could be a variant between an ancient serine / threonine kinase and the Wee1 tyrosine kinase. Probably both the ancestors of plants and that of fungi may have lost the *myt1* gene. In fission yeasts, Pyp3 is a backup phosphatase of Cdc25, also activating MPF in late G2. Interestingly, we found that the small Ibp1 phosphatase appeared to be a closer homologue of Cdc25, although its function is different. Moreover, Cdc25 homologues identified in plants were found to be more closely related to Ibp1 rather than to Cdc25 of fission yeast. In the Cdc25-like proteins, a novel conserved region was found with the consensus sequence LxxG(Y/F).

Keywords

G2 size control, phylogenetic analysis, fission yeast, Wee1, Cdc25

1 Introduction

G2 phase size control is based on the inhibition of a cyclin / cyclin-dependent kinase (Cdk) heterodimer called MPF (M-phase promoting factor), until a critical cell size is reached. In the G2 phase of the cell cycle, the Tyr-15 amino acid of the regulatory Cdk1 subunit of MPF is phosphorylated by the Wee1 kinase, generating the low-activity form of MPF, called pre-MPF. In late G2 phase, the inhibiting phosphate group is removed by the Cdc25 phosphatase, producing the high-activity MPF form (Fig. 1). MPF once directly phosphorylates several substrates to initiate mitosis, moreover regulates itself by activating the Cdc25 phosphatase (positive feedback loop), and also by inhibition of the Wee1 kinase (double negative feedback) to ensure a rapid entry into M phase

(Fig. 1) [1]. In the fission yeast *Schizosaccharomyces pombe* (fission yeast*)*, size control in G2 phase is of paramount importance, therefore backup copies of both Cdc25 phosphatase and Wee1 kinase exist, namely Pyp3 phosphatase and Mik1 kinase, respectively. In fission yeast, size control may probably be achieved via monitoring cell length through Pom1, Nif1, Cdr1 (Nim1) and Cdr2 proteins (Fig. 1), all being upstream regulators of Wee1 [2].

Cell length matters here since this unicellular eukaryote has cylindrically shaped cells, growing exclusively at the tips with a constant diameter during the mitotic cycle. By analyzing morphological mutant cells with abnormal widths, a recent study suggested that surface area is monitored in fission yeast rather than simply length [3].
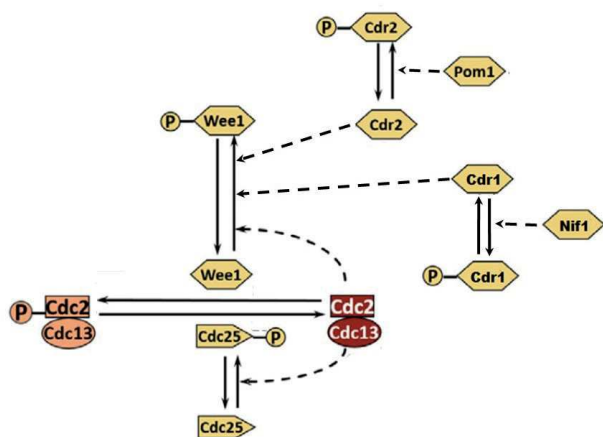
Fig. 1 A schematic view of the main proteins and their biochemical interactions, regulating G2 size control in fission yeast. Continuous lines with arrows represent biochemical reactions, while dashed ones represent enzymatic effects.

However, with the exception of the very rare cases of these mutants, cell diameter does not vary significantly among fission yeast strains, therefore cell surface can be considered to be proportional to length.

1.1 Characteristics of Wee1-like kinases in different organisms

The Wee1 protein and its homologs contain an approximately 280 amino acid long protein kinase domain on the C-terminus and a less conserved large N-terminal domain containing Ser-Pro or Thr-Pro dipeptides, which are potential inhibitory phosphorylation sites [4]. Wee1 is an atypical tyrosine kinase that exhibits a greater similarity to serine / threonine kinases by cluster analyses. The specificity to phosphorylate only the Tyr-15 amino acid of Cdk1 appears to be due to the presence of a glutamic acid (Glu-309) in its activation segment, which, due to steric inhibition, prevents phosphorylation of Thr-14 of Cdk1 [5]. Cyclin-Cdk complexes like MPF often recognize substrates based on short motifs in the sequence, such as the RxL motif. The human somatic Wee1 contains four RxL amino acid triplets that are conserved among the somatic Wee1 homologs of vertebrates. The most conserved of the four sequences overlap with the nuclear export signal (NES) sequence. The protein also contains a "Wee box" sequence, which increases the activity of the kinase domain. This region is also conserved in most eukaryotic cells, but is not found in the Wee1-like Swe1 protein of budding yeast [6].

In fission yeast, Mik1 (the backup copy of Wee1) also phosphorylates the side chain of Tyr-15 of the MPF

catalytic subunit [7]. In higher organisms, there are often two Wee1 homologs called Wee1A and Wee1B. In the frog *Xenopus laevis* (African clawed frog), Wee1A and Wee1B proteins are maternal and zygotic isoforms. Switching from expressing Wee1A kinase to that of Wee1B occurs during gastrulation, when the cell cycle is greatly prolonged depending on the zygotic transcription. Wee1 protein is a dose-dependent mitotic inhibitor, and the *Xenopus* Wee1B kinase has four times larger inhibitory effect than that of Wee1A. The Wee1B protein contains three tandem repeats of a seven amino-acid long region at the N-terminus, but the third repeat is incomplete. This repetitive region is absent in some other Wee1 homologs, including the Wee1A protein of *Xenopus*, but also occurs in Wee1B proteins of other species [8]. In contrast to *Xenopus*, the human Wee1B protein appears to be a maternal protein, whereas Wee1A is zygotic [9]. In mouses, the Wee1A protein is probably present in the somatic cells of the adult, while the Wee1B kinase is present in the mature egg cells [10]. Wee-1.1 isoform in the nematode *Caenorhabditis elegans* (nematode) plays a role in the early embryonic cycles, Wee-1.2 is not expressed at all [11], while Wee-1.3 is found in mature eggs [12].

In higher organisms, Myt1 (membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase) is present alongside Wee1 proteins. This protein is able to phosphorylate both the inhibitory Thr-14 and Tyr-15 positions of Cdk1 [8]. While Wee1 and Mik1 mainly localize in the nucleus, Myt1 protein is associated to the membranes of the endoplasmic reticulum and the Golgi apparatus [8]. Myt1 protein contains a domain that binds to the MPF complex. This domain in the human Myt1 protein is between amino acids 436-499, bearing a highly conserved RNL motif at position 486-488. Between this domain and the protein kinase one there is a 20 amino acid long motif, through which the Myt1 protein can associate with the membrane. In the human Myt1 protein, this region is between Arg-378 and His-399; it mainly consists of hydrophobic or uncharged amino acids, which forms an α-helical structure to be integrated into one phospholipid layer of the membrane. MPF is shuttling between the nucleus and the cytoplasm during interphase. Myt1 can inhibit MPF both by phosphorylating Tyr-15 and Thr-14 at the catalytic subunit, and by preventing nuclear import of the complex [13]. The Myt1 protein also plays a role at the end of mitosis, helping to assemble the endoplasmic reticulum and Golgi [14]. Both the *C. elegans* Wee1 homolog proteins, Wee-1.1 and Wee-1.3 appear to be more closely related to Myt1 than to Wee1 kinases [11].

The amino acids that can be phosphorylated by Wee1 and Myt1 homologs, Thr-14 and Tyr-15 are located near the ATP binding site of the Cdk1. Thus, in phosphorylated state, the phosphate groups interfere with ATP, thereby inhibiting its binding. This reduces MPF activity [15], therefore Wee1, Mik1 and Myt1 kinases all inhibit MPF both sterically and enzymatically [2, 4].

1.2 The roles of Cdr1, Cdr2, Nif1 and Pom1 kinases in regulating Wee1 in fission yeast

Cdr1 and Cdr2 take part in the formation of large protein complexes in the middle of the cylindrical cells. These ones are called the medial cortical nodes, which can attract and inhibit the Wee1 kinase. Cdr2 phosphorylates Wee1 at the N-terminus, allowing the protein to bind to these nodes. When Wee1 is bound to one of these Cdr1 / Cdr2 nodes, Cdr1 phosphorylates Wee1 at the C terminal domain and thereby inactivates it. The number of these nodes is doubled as the cell grows during the cycle, and the residence time of Wee1 at these nodes increases by about 20 times [16]. The Pom1 protein has a negative effect on both Cdr1 and Cdr2 kinases, however, it directly phosphorylates only Cdr2. Pom1 protein forms a spatial gradient from the end of the cylindrical cell to the middle [17]. Pom1 has regions rich of arginine and lysine carrying positive charges. An electrostatic attraction occurs between Pom1 and the negatively charged lipids of the plasma membrane. Pom1 is first carried to the cell ends by a transport mechanism via microtubules, but then it laterally diffuses (along the cell membrane) away from the tips. At the same time, Pom1 is autophosphorylated at several places, causing it to dissociate from the membrane surface, because of the negative charge of the phosphate groups. The hyperphosphorylated Pom1 attaches to a microtubule again and is transported back to a cell pole, becoming dephosphorylated again at the same time. As a result, it is able to re-bind to the membrane at the cell end, and all these processes start to repeat [18]. While the cell size is small, the formed spatial gradient of Pom1 results in a Pom1 concentration, which is large enough even at the medial cortical nodes (containing Cdr2) to keep the Cdr2 proteins inactive, so that they are unable to phosphorylate Wee1, therefore MPF is kept in its inactive preMPF form and the cell remains in G2 phase. However, as the cell size increases, the concentration of Pom1 decreases in the middle of the cell, so that Cdr2 becomes active to phosphorylate and inactivate Wee1, thereby facilitating formation of the active form of MPF, leading to mitotic onset [17].

Although later a further paper denied the essential role of Pom1 in G2 size control [19], the interpretation of these results have been challenged recently [20, 21]. Similarly to Pom1, the Nif1 kinase also has a spatial cortical gradient. Nif1 can inactivate Cdr1 by phosphorylation, keeping Wee1 active while the cell is small [19].

1.3 Characteristics of Cdc25-like phosphatases in different organisms

Cdc25 protein is a dual-specifity phosphatase with an increasing concentration during G2 phase. It is mainly localized in the cytoplasm, but it exerts its effect in the nucleus [22]. Cdc25 shows a rhodanase-like three-dimensional structure and has only one characterized domain at the C-terminus. This conserved region is the catalytic domain, as the active site of the seven amino acid sequence CE(Y/F)SxxR forms a phosphate binding pocket [23]. The structural feature of the protein is that the α-helix of the phosphate binding site is located on the surface, while in all other known phosphatases it is located inside the structure. As a consequence, the substrate-binding pocket of Cdc25 is the lowest among the known phosphatases, which allows hydrolysis of the phosphorylated threonine-14 and tyrosine-15 residues of Cdk1, thus ensuring the dual-specificity of the protein [24]. The N-terminus of the protein is diverse and poorly characterized, however, there are sites responsible either for phosphorylation and/or ubiquitination for degradation, which regulate phosphatase activity, concentration and/or association with other proteins [25]. Two other conserved regions of Cdc25 have been identified in *C. elegans*. One consists of 21 amino acids with a consensus sequence IIDCRYPYEYxGGHIxGAxNL, in which the aspartic acid functions as a general acid, therefore it is found in all known Cdc25 proteins. The other region seems to be fully conserved in all Cdc25 proteins, and it can be described by the consensus sequence CxPxxYxxM [26].

Cdc25 is a dose-dependent activator of MPF. During G2 phase, Cdc25 is transported to the nucleus, and if the cell reaches a size required for mitosis, Cdc25 becomes active and dephosphorylates the inhibiting phosphorylation sites of Cdk1. In addition to the positive feedback loop between Cdc25 and MPF, a Polo kinase homolog in *S. pombe* also phosphorylates the Cdc25 protein and, on the other hand, promotes cyclin degradation after the onset of anaphase, thus ensuring a robust, but short-term activity of the MPF complex [27]. Protein expression increases as the cell grows, so Cdc25 becomes enriched during the cycle [28].

Under normal conditions, Cdc25 is essential in all eukaryotic cells, and the increasing number of isoforms in higher organisms is parallel to the diversity of Cdc25 substrates [24]. *C. elegans* has four known Cdc25 homologs, of which Cdc-25.1 is essential for divisions of germ line cells [26]. In *Drosophila melanogaster* (fruit fly) there are two isoforms, called string and twine. String isoform is required in mitosis after organogenesis [29], and twine is essential for meiosis of the germ line [30]. In *Arabidopsis thaliana* (thale cress) a dual-specificity phosphatase has been identified as a Cdc25 homolog, as it activates the Cdk complexes of the plant *in vitro* and also has some sequence similarity. However, the protein consists of only 146 amino acids and the complementation experiment in fission yeast was not successful. This protein may have probably been involved in events of genome duplication [31]. Three isoforms are found in vertebrates, probably due to gene duplication and divergence, called Cdc25A, Cdc25B and Cdc25C [32].

While Cdc25 plays a role in fission yeast only at the G2/M transition, it also participates in the G1/S transition in vertebrates. In the latter more complex regulatory system, the different isoforms have distinct functions and expression profiles. Cdc25A starts to be expressed in mid G1 phase; Cdc25B starts to accumulate during S phase and reaches its maximum during mitosis, while the Cdc25C protein level is relatively constant over the entire cell cycle. Cdc25A and Cdc25B proteins are degraded by ubiquitin-mediated proteolysis after mitosis [32]. Cdc25A and Cdc25B are likely to cooperate to regulate mitosis and are able to compensate slightly for each other. Based on experimental results, Cdc25B is assumed to be responsible for the activation of centrosomal MPF, while Cdc25A plays a role in the initiation of chromatin condensation by activating non-centrosomal Cdk–cyclin complexes [33]. The role of Cdc25C remains controversial in cell cycle regulation, as its absence does not inhibit G2/M transition and is not itself capable of initiating mitosis [33]. According to some assumptions, the protein must be thiophosphorylated by ATPγS to activate MPF. This phosphorylation only slightly increases the activity of the protein, but it may be enough to activate an autocatalytic positive feedback loop, resulting in the activation of additional Cdc25C proteins by active MPF, thereby further increasing its total activity [34].

In fission yeast, some further phosphatases are also able to dephosphorylate the tyrosine-15 residue of Cdk1 [35]. One such protein is Pyp3 that is localized in the cytoplasm.

This protein does not exhibit any significant similarity to the Cdc25 phosphatase, but it has a seven-residue domain responsible for its catalytic activity. It is not essential in wild-type cells, but its absence leads to a larger cell size, therefore it probably plays a role in mitotic onset, similarly to that of Cdc25. Overproduction of Pyp3 (similarly to that of Cdc25) accelerates the initiation of mitosis, resulting in a smaller cell size. Pyp3 is able to dephosphorylate and thereby activate MPF *in vivo* [35]. In 2003, another Cdc25-like protein was identified in fission yeast, called Ibp1 (Itsy Bitsy Phosphatase). Ibp1 is a catalytically active phosphatase, but its lack does not cause any phenotypic change. Its overproduction does not save the temperature-sensitive *cdc25-22* mutant of fission yeast [36].

2 Materials and methods
2.1 Searches for homologs
Amino acid sequences of the studied proteins of different species were obtained from several different databases. First, searches for the sequences of *S. pombe* Pom1, Nif1, Cdr1, Cdr2, Wee1 and Cdc25 proteins were performed in the PomBase [37] database. The sequences of *Schizosaccharomyces octosporus* and *Schizosaccharomyces cryophilus* Wee1 were downloaded from the database of the BROAD Institute [38]. The sequences of Wee1A and Wee1B proteins of zebrafish, the homologous corn protein, the Wee1 and Myt1 proteins of men, mouse and *Xenopus* were all collected from the UniProt database [39]. For the other species examined, Wee1 homologs were obtained from the NCBI [40] database. The Cdc25 protein sequences of men, mouse, *Xenopus* and *Arabidopsis* were also obtained from the UniProt database, and for the other species examined, search in the NCBI database was performed. The homologs of Pom1, Nif1, Cdr1 and Cdr2 were also searched for in the NCBI database. Homologs in protists were searched for to be used as outgroups on phylogenetic trees.

Searches were performed using BLASTp [41] with the default parameters and the appropriate *S. pombe* protein as the query sequence. The most similar proteins were then extracted from the searched databases and were used for a reciprocal search in the genomes of *S. pombe*, *Homo sapiens* (human) and *A. thaliana*. The proteins that gave the highest scores with the original *S. pombe* proteins in the reciprocal search were considered to be the putative homologs. In the case of the human genome, reciprocal BLAST was performed to see which human protein (Wee1 or Myt1) is more similar to the putative

homolog. In the case of *Arabidopsis*, the question was whether there is any similar protein at all.

Pairwise BLAST alignments of the putative homolog and the corresponding protein of *S. pombe* were performed in order to get comparable results. The conserved domains were localized in the proteins by scanning their sequences using the Pfam tool in the Pfam-A database [42], with the default threshold for the hidden Markov model search.

2.2 Multiple alignments and phylogenetic analyses

Multiple alignments were generated by both ClustalX [43] and PRANK [44] algorithms, differing in the scoring of gap insertions. For ClustalX, the matrix used (BLOSUM) and the gap costs (scoring penalty) for inserting (11) and extending (1) gaps have been adjusted to the ones applied by the BLAST searches. For PRANK alignments the default setting parameters were used. WebLogos [45] were generated from multiple alignments to analyze conserved regions.

Phylogenetic trees were generated from multiple alignment by MEGA6 [46] using neighbor joining, maximum parsimony, and maximum likelihood methods. Reproducibility of the trees was tested by bootstrap analyses (based on 100 replications). In the neighbor joining analyses, the Jones–Taylor–Thornton model of amino acid substitutions [47] was used for computing distance matrices, the pairwise deletion option and otherwise default parameters were used. For the maximum-likelihood analyses, an amino acid substitution model was chosen by the lowest BIC (Bayesian Information Criterion) value in MEGA. In case of the maximum parsimony and maximum likelihood methods, the "use all sites" option and otherwise default parameters were used.

Bayesian inference of phylogeny was performed using MrBayes [48]. The best substitution model was determined by the ProtTest [49] software. The phylogenetic trees were visualized with the FigTree [50] software. The Tracer [51] program was used to evaluate the Bayesian trees, and check the Effective Sample Size (ESS). If the ESS was below 200, the phylogenetic tree was re-edited with a different run time.

In the studies of both Cdc25- and Wee1-like proteins, several phylogenetic trees were constructed to test robustness. On the one hand, minor changes have been made to the list of studied species (Table S4), so that the different taxonomic units probably have the appropriate weights. Multiple alignments of the full proteins were performed with ClustalX for each species list, and in most cases also with Prank. Domain sequences (rather than full proteins) were also used for analyses in two cases with different

species included. The advantage of aligning the domain sequences is that they nearly have the same length, while the lengths of the entire protein sequences may be very different. However, the disadvantage of using domains is that they make up a relatively small portion of the entire protein, and therefore carry much less information, so it is more difficult to evaluate their evolutionary distances. Another change in the species list was the applied outgroup, which may also influence the structure of the phylogenetic tree. The two different identified Wee1-like *Dictyostelium discoideum* (cellular slime mold) homologs were used as outgroups, either one or the other. In case of one species list, the *S. pombe* Polo kinase was used as an outgroup. Table S5 shows the variations of the alignments. For each fit, phylogenetic trees were made with neighbor joining, maximum parsimony and maximum likelihood methods, and in several cases Bayesian trees were also constructed (run for 200 000 generations).

3 Results
3.1 Homologous proteins found in different databases

Pom1 is the only known DYRK kinase (Dual-specificity Tyrosine Regulated Kinase) in fission yeast, so BLAST and reciprocal BLAST searches have found sequence homology to several DYRK kinases in the tested organisms, however, their protein length is significantly different from that of Pom1. Moreover, although the matched section is the conserved domain, but the function of these proteins is different from that of Pom1, so they were not considered to be true homologs. Homolog of the Nif1 protein was found only in the members of the *Schizosaccharomyces* genus. Searching for Cdr1 or Cdr2 homologs both resulted in hits of BRSK (Brain-specific serine / threonine-protein kinases) 1 and 2 proteins in higher-order animals, but they were not considered to be true homologs. Wee1 and Cdc25 proteins generally have homologs in the wildlife, thus we examined them among the proteins of model organisms with fully sequenced genomes (the list of species and the results are shown in Tables S1 and S2).

3.1.1 Wee1-like homologs found in the tested species

All these results are shown in Table S1. In case of most fungi only one homologue was found in one species. In *Cryptococcus* and *Trichosporon* there were two homologs found, namely putative Wee1 and Myt1 ones. Mik1 in fission yeasts was found to be homologous to Wee1 by BLAST analyses. In case of plants, generally one result was found in one species, except in *Populus trichocarpa*

(black cottonwood) having two Wee1-like homologs. Since the alignment data of these two proteins show that they are only slightly different, it is likely that genome duplication might have occured within this species. In higher animals, besides Wee1, Myt1 kinase is also present, moreover in several many cases, two Wee1 homologs (called A and B) were found. The Wee1B protein of the *Xenopus* is more similar to the Wee1A orthologs based on reciprocal BLAST search, and this is confirmed by phylogenetic trees. The Wee-1.1 and Wee-1.3 proteins found among *C. elegans* proteins are closer to the Myt1 orthologs according to literature [11], and the phylogenetic trees confirmed this. All of the examined protists contained two homologs, probably putative Wee1 and Myt1. The protein kinase domain was present according to the Pfam analyses, but no other conserved domains were found.

3.1.2 Cdc25-like homologs found in the tested species
All these results are shown in Table S2. In members of the *Schizosaccharomyces* genus, the second best BLAST result (with the highest score) after Cdc25 was Ibp1, which really showed a significant similarity to Cdc25. It is surprising that Pyp3 does not show a close similarity to the Cdc25 protein, as it can dephosphorylate Cdk1 as does Cdc25. In most of the animals more then one Cdc25 isoforms were found. BLAST search showed no homologs in plants, either using the sequence of *S. pombe* or *H. sapiens* Cdc25. Landrieu et al. [31] identified a dualspecificity phosphatase in *A. thaliana*, which was thought to be a Cdc25 homolog. Further homologs of this protein were identified in the selected plants. In reciprocal BLAST, all these plant proteins were shown to be Ibp1 homologs, although they are very short ones compared to the Cdc25 protein identified in *S. pombe*. All of the examined protists contained one Cdc25 homolog, with a surprising exception of *Entamoeba histolytica*, which contained six homologous candidates. However, probably several of these hits are only the results of inappropriate annotations; therefore only one of them was used for further analyses, namely the one with the largest similarity to the *S. pombe* Cdc25 sequence.

The rhodanase-like domain at the C-terminus of *S. pombe* Cdc25 protein was found in the Pfam database in all the identified putative homologs, thus afterwards they can be considered as true homologs. In higher organisms, in addition to the rhodanase-like domain, a second characteristic motif was also identified, generally found in an M-phase inducer phosphatase family. This section of approximately 250 amino acids is located in the

N terminal direction from the rhodanase-like domain, but it is absent in Cdc25D proteins. The Ibp1 proteins in the four *Schizosaccharomyces* species also contain the rhodanase-like domain, but they are approximately twenty amino acids longer than that of the Cdc25 proteins.

3.2 Conserved regions
Multiple alignments were performed with the identified homologous proteins for generating phylogenetic trees and also for searching conserved motifs. Weblogos were generated and invariable regions were found in the orthologous and paralogous proteins.

3.2.1 Conserved motifs in Wee1-homologous sequences
Multiple alignments by Squire et al. [5] have shown the HxDxK(P/L)xN in the catalytic segment of Wee1 proteins and the K(I/L)(G/A)D(F/L)G in the activation segment to be conserved, but they have deduced their conclusions based on few aligned sequences. They found an EGD amino acid triplet in the activation segment characteristic of Wee1 kinases [5, 52]. Our involvement of several further species of animals, plants and fungi showed the conserved motif to be (E/D)GDxx(Y/F), rather than EGD. Besides the above mentioned three conserved motifs, we have also identified a fourth one with a general sequence D(I/V)(F/Y)(S/A)x(G/A), which has never been described previously to our best knowledge (Fig. 2).

In the Wee1 protein domain structure, a conserved RxL amino acid triplet has been found by multiple alignments in the human Wee1A, *Xenopus* Wee1A and Wee1B, and the Wee1A proteins of chicken and zebrafish. Since the motif is based on the somatic Wee1 proteins of vertebrates [6], the result meets our expectations. The Wee box region was found in human Wee1B; in mouse, *Xenopus*, chicken, and zebrafish Wee1A and Wee1B; and also in *Drosophila* Wee1, aligned to the human Wee1A protein. The Wee box region has been described to exist in most of the eukaryotic Wee1 proteins, while it is absent in the budding yeast ortholog [6]. Accordingly, we have identified this motif in Wee1-like proteins of animals, however, not in fungal homologs.

To decide whether any obtained sequences was either a Wee1 or a Myt1 type kinase, the environment of human Wee1 Glu-309 were analysed (which sterically inhibits the phosphorylation of Thr-14 of Cdk1 [5]) (Fig. 3).

Glutamic acid (E) can be found in the GxGEF conserved environment in almost all of the animal and fungal Wee1 proteins tested (Fig. 3), except for the proteins of zebrafish Wee1A, budding yeast Swe1, *Candida albicans* and
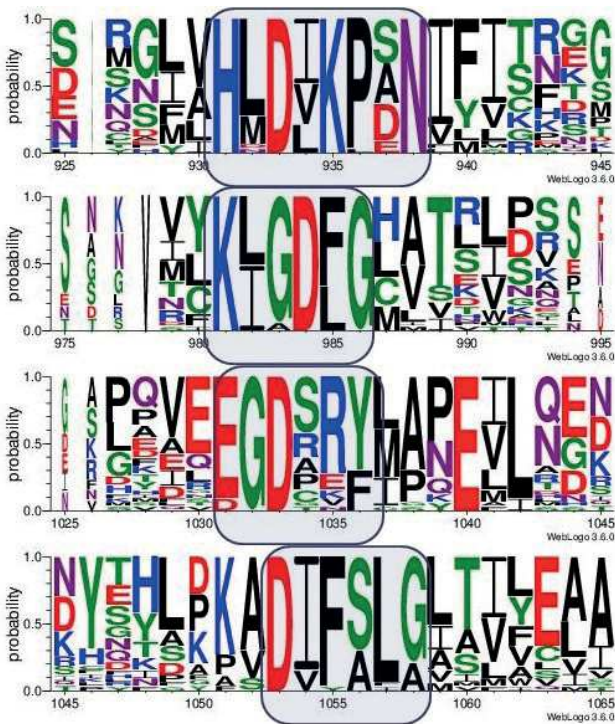
Fig. 2 Conserved motifs in Wee1-homolog sequences: HxDxK(P/L)xN; K(I/L)(G/A)D(F/L)G; (E/D)GDxx(Y/F); D(I/V)(F/Y)(S/A)x(G/A). The letters are one-letter codes for the amino acids. Two amino acids in the same position are separated by / and are in parentheses. x means any amino acid (more than two options in a position).The height of a letter is proportional to the relative frequency of occurrence of that amino acid. The width of the letters is the relative frequency of the occurrence of an amino acid (and not a gap) at the given position. (Color letters represent different types of amino acids, as black, apolar; green, non-charged, polar and glycine; purple, diamino acids Q and N; blue, positively charged; red, negatively charged amino acids.)
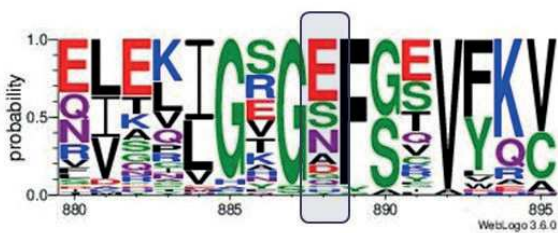

Fig. 3 Weblogo of the environment of the Wee1 Glu-309 position. Numbering is slightly different from that of Fig. 2, because of the inserted gaps. For further details, see also the legend to Fig. 2.

Cryptococcus neoformans Wee1. In the latter two cases, and in the case of shorter proteins of *Trichosporon asahii* and *C. neoformans* Myt1, Alanin (A) was found in this conserved environment.

Animal Myt1 proteins and *C. elegans* Wee-1.1 and Wee-1.3 (they are more similar to Myt1 then to Wee1

homologs), and also most plant Wee1 proteins contained either serine (S) or asparagine (N) instead of glutamic acid (E-309). In some cases, however, plant proteins contained totally different amino acids, like *Arabidopsis* (H) and *Medicago* (barrelclover) (Y) in this position. The generally conserved neighbor phenylalanine (F-310) also changed to other amino acids (Y or S) in some plant proteins.

In the case of the protist proteins examined, no similar regularity could be detected; in the *Ostreococcus tauri* Wee1 protein, only the phenylalanine was present in this conserved region, similarly to the fission yeasts' Mik1 proteins.

3.2.2 Conserved motifs in the Cdc25-homologous sequences

The IIDCRYPYEYxGGHIxGAxNL consensus sequence identified in *C. elegans* is positioned between positions 917 and 937 of the alignment (Fig. 4). If the Cdc25 sequence of all species and the Ibp1 sequence of fission yeasts are aligned, only the aspartic acid (D) at position 919 and arginine (R) at position 921 are conserved completely. Of the two amino acids sharing position 920, valine (V) is present in plant Cdc25 and Ibp1 proteins, while all other sequences contain cysteine (C). Histidine (H) is highly conserved at position 930, since only the sequences of *O. tauri*, *Ostreococcus lucimarinus* Cdc25 and *S. pombe* Ibp1 are different, which contain arginine (R). Similarly, isoleucine (I) at position 931 is also highly conserved, with the exceptions of a few protists' Cdc25 sequences and also with the Cdc-25.1 of *C. elegans*.

The CxxSxxR consensus sequence of the catalytic domain required for the phosphatase activity is between positions 986 and 992 (Fig. 5). According to the data of
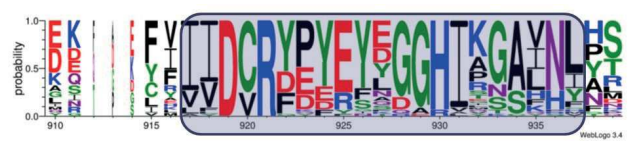

Fig. 4 Weblogo of the conserved IIDCRYPYEYxGGHIxGAxNL sequence in Cdc25 homologs. For further details, see also the legend to Fig. 2.

Fig. 5 Weblogo of CxxSxxR consensus sequence of the catalytic domain of Cdc25 homologs. For further details, see also the legend to Fig. 2.

Bordo and Bork, the second position of the conserved region should be a conserved glutamic acid (E) [23], but plants contain alanine (A), while the Ibp1 sequences contain threonine (T) at this position. Some papers also consider histidine (H) located in front of the catalytic cysteine (C) as part of the conserved region [23]. In this position, the four proteins of *C. elegans*, like the sequence of *O. tauri* and *O. lucimarinus*, contain tyrosine (Y), while the Cdc25A protein of *Mus musculus* (mouse) contains leucine (L). It can be assumed that the phenylalanine (F) at position 984 and the proline (P) at position 994 may play a role in forming the substrate binding pocket.

The third conserved region described in the literature [26] is between positions 1061 and 1071 (Fig. 6). Cysteine (C) at position 1061 is conserved in all species and proteins studied. However, the methionine (M-1071) described to be conserved in the literature, is absent from most Ibp1 proteins, the *Danio rerio* (zebrafish) Cdc25D and the *Paramecium tetraurelia* Cdc25 proteins. Some plant sequences have lysine (K) or glycine (G) in this position. The originally conceived consensus sequence was extended with a gap in positions 1066-1067 by our multiple alignments. These were primarily due to the extra amino acids found in the *S. pombe* Ibp1 protein and some plant sequences.

A fourth conserved region has been found during this work, where leucine (L) at position 1037 and glycine (G) at position 1040 are totally conserved in all sequences analysed (except *Populus trichocarpa*, which lacks this region). The exact function of these amino acids is not yet known, and to our knowledge, there is no former literature mentioning it. In the Cdc25-like proteins, the consensus sequence of this region is probably LxxG(Y/F) (Fig. 7).
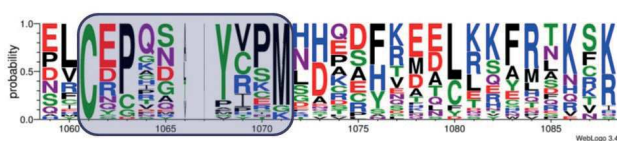

Fig. 6 Weblogo of CxPxxYxxM consensus sequence of the catalytic domain of Cdc25 homologs. For further details, see also the legend to Fig. 2.
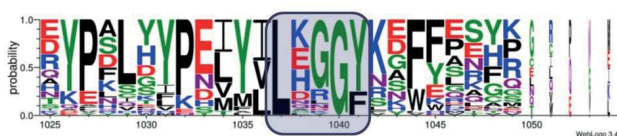

Fig. 7 Weblogo of the novel conserved region LxxG(Y/F) of Cdc25 homologs. For further details, see also the legend to Fig. 2.

3.3 Phylogenetic trees of homologous proteins
3.3.1 Relationships of Wee1-like proteins

Animal Wee1, animal Myt1, plant and fungal Wee1 proteins can clearly be distinguished on the phylogenetic trees. The separation of Wee1A and Wee1B proteins is also obvious, although the Wee1B protein of zebrafish is not always on the same branch as other Wee1B ones. Both proteins of *Xenopus* resemble the Wee1A proteins, while both proteins of *C. elegans* are on a branch with Myt1 kinases. There were slight differences in the order of branching of the proteins of the species within each protein group and between the plants and the fungi. The phylogenetic trees also containing protists' proteins showed that these sequences were not placed on a separate branch. Five typical topologies could have been derived from the phylogenetic analyses, and their frequencies are shown in Table 1 (see also Section 2.2 for methods).

Sorrell et al. [52] found a homolog of the Wee1 protein in *Arabidopsis thaliana*, and generated a phylogenetic tree based on the amino acid sequences of the catalytic domains of 5 animal Wee1, 4 animal Myt1 and 2 plant proteins, as well as of the budding yeast Swe1 and fission yeast Wee1 and Mik1. In correspondance with their results, we also found the Mik1 protein to be separated on the tree before branching of the budding and fission yeast proteins. In contrast to Sorrell's phylogenetic tree, our results show that the *Xenopus* proteins are similar to the human Wee1A protein, although only one of the Wee1 homologs was used (Figs. 8, 9).

Based on the Wee1 homologs of the studied model organisms, a total of 75 phylogenetic trees were made. In Table 1, topology D can be explained by the least evolutionary steps, however, this appeared in only four out of 75 cases. Based on the much more abundant topologies A-C (Table 1), the Myt1 protein may be more ancient than Wee1. This idea is also supported by the fact that Myt1 is able to phosphorylate both potential inhibitory phosphorylation sites of Cdk1, however, its activity is significantly lower than that of Wee1 [13]. We can conclude that appearance of the higher-activity Wee1 protein (partly replacing the role of Myt1) was evolutionarily profitable. Since Wee1 is a tyrosine kinase originated from a serine / threonine kinase [5], and Myt1 is a dual-specific (serine / threonine and tyrosine) kinase [13], Myt1 may be the transient variant between an ancient serine / threonine kinase and the modern Wee1 tyrosine kinase. If Myt1 were more ancient than Wee1, it is likely that after the separation of animals, plants and fungi, both genes remained in the common ancestor

Table 1 Main topology types of phylogenetic trees and their frequencies acquired by examination of major groups of Wee1 homologs. The most common topology was confirmed with the Bayesian method. (More details are shown in Table S3.)

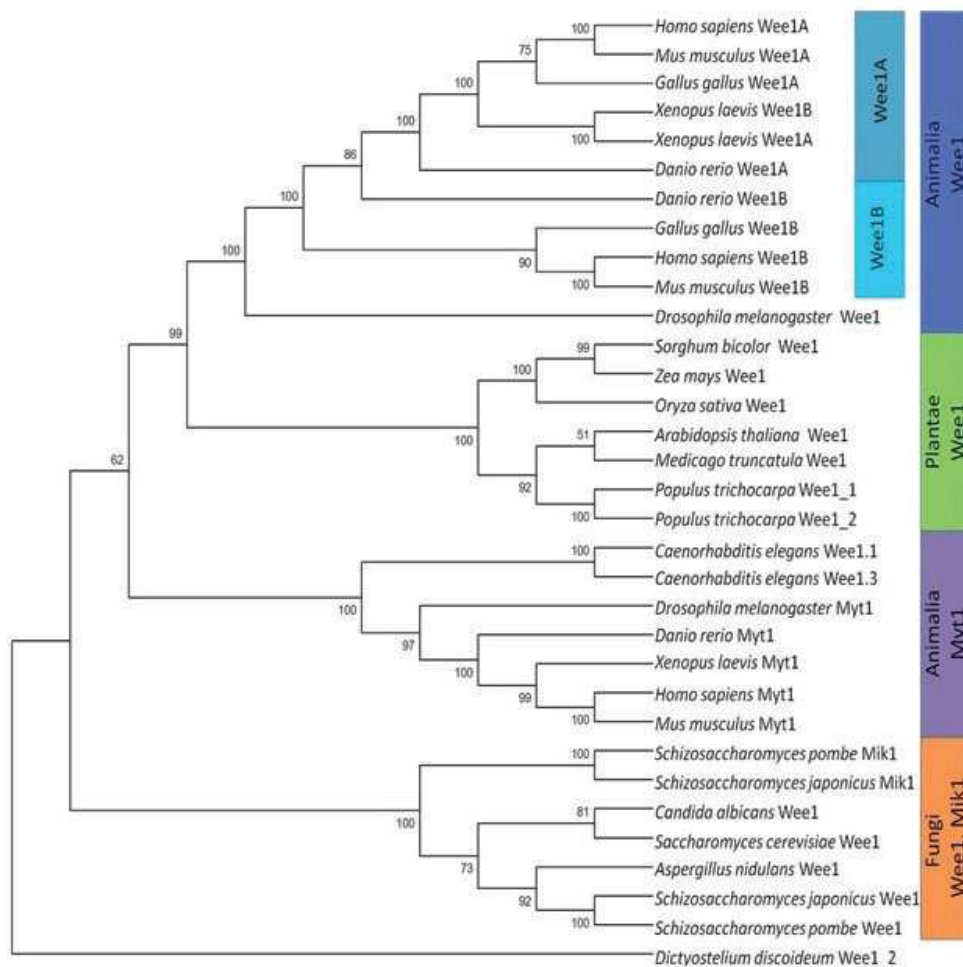| Topology	A	B	C	D	E	Σ
Σ	29	22	18	4	2	75

Fig. 8 Maximum likelihood phylogenetic tree of Wee1 homologs (full protein sequences; generated with the LG + G + I model).

of animals, the *myt1* gene could have been lost by the ancestors of both plants and fungi. Based on phylogenetic trees, it is likely that the Wee1A and Wee1B proteins have been diverged somewhere in the ancestor of higher animals. The Mik1 kinase occurs only in fission yeasts, so it may have been developed in the common ancestor of the *Schizosaccharomyces* genus (Figs. 8, 9).

3.3.2 Relationships of Cdc25-like proteins

There were differences between the phylogenetic trees of the Cdc25-like proteins generated by different methods,

but an approximately consensus topology was drawn, which can be seen in Fig. 10.

The evolution of the Cdc25 homologs identified in the animals (Cdc25A, Cdc25B and Cdc25C) shows a high degree of similarity among various methods. Cdc25D proteins identified in zebrafish and *Xenopus* show the greatest similarity to each other, rather than to any of Cdc25A, Cdc25B or Cdc25C [53], but, however, they rarely occur on the same branch. It can be concluded that the common ancestor of Cdc25A and Cdc25B proteins has been separated from other animal Cdc25 homologs first.
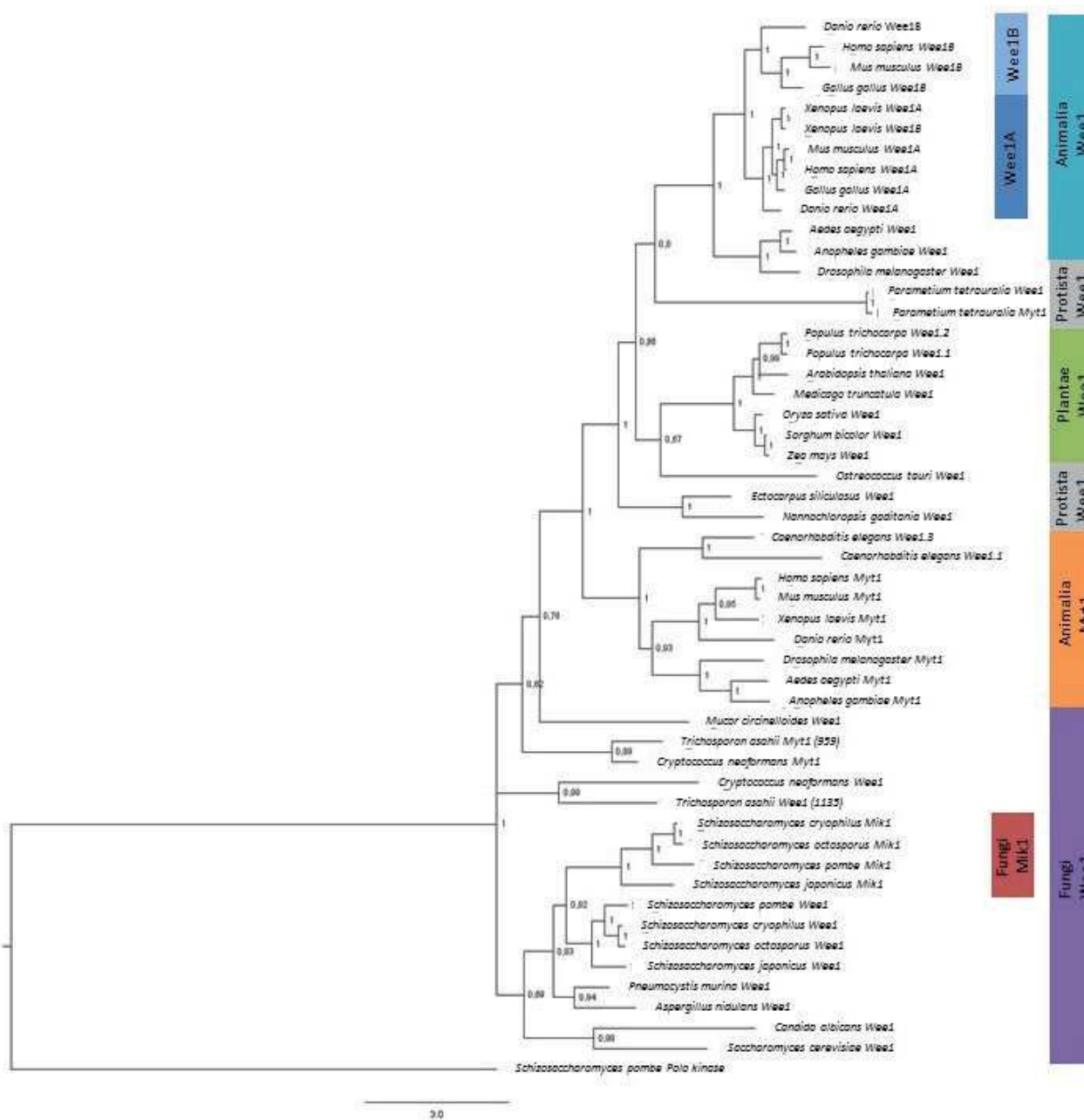
Fig. 9 Bayesian phylogenetic tree of Wee1 homologs (full protein sequences; generated with the VT + F + G + I model)

All other Cdc25 homologs of *X. laevis* show the greatest similarity to Cdc25C proteins. The protein identified as Cdc25B in zebrafish is sometimes situated on the branch of Cdc25A proteins. The homolog identified in *Gallus gallus* (chicken) is clearly one of the Cdc25A isoforms. There are multiple branches seen in the Bayesian tree topology that can be caused by small differences between sequences, probably not containing enough information to clearly reconstruct their relationships (Fig. 11).

The homologs identified in green algae (*Chlorella variabilis*, *O. tauri*, *O. lucimarinus*) are usually not on the same

branch with other protists, but together with the Cdc25 proteins identified in animals and fungi instead. The four Cdc25-like proteins of *C. elegans* are often located elsewhere on different trees, but it can be assumed that they were the first sequences to be separated from other animal ones. The insect proteins, *Anopheles gambiae* (mosquitoe) Cdc25 and *Drosophila melanogaster* string and twine are always situated on the same branch.

Because plant Cdc25 sequences showed greater similarities to Ibp1 proteins in fission yeasts than to Cdc25 proteins, phylogenetic trees were generated for Ibp1 homologs
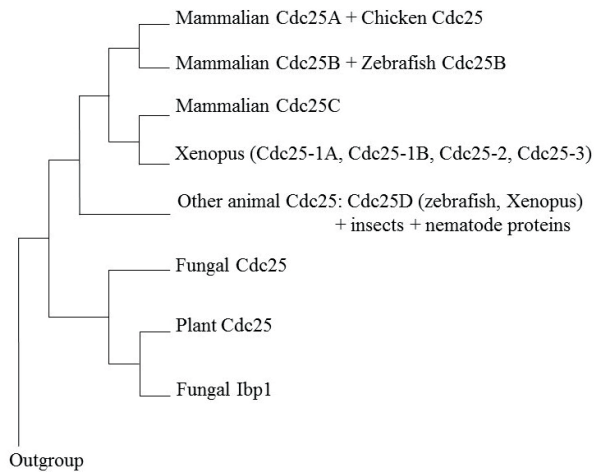
Fig. 10 Approximately consensus topology of phylogenetic trees of Cdc25-like proteins.

and for Cdc25 homologs detached. To represent the plant Cdc25 and Ibp1 sequences, a new outgroup, the Ibp1 protein of *C. variabilis*, was applied. The topology of

the obtained phylogenetic trees shows a great similarity to those ones containing all the sequences. For example, one of *Sorghum bicolor*'s (great millet) Cdc25-like proteins exhibited a larger similarity to the *Oryza sativa* (rice) protein, meanwhile the other one to that of *Zea mays* (corn).

4 Discussion and conclusions

In this work, we have generated phylogenetic trees of Wee1-like proteins of general model organisms in cell cycle studies. The homologs have been obtained from their fully sequenced genomes, and the phylogenetic trees showed a total of five typical but different topologies. The topology (Table 1, D) that can be explained with the fewest evolutionary steps, appeared only in 5 % of cases. Moreover, these rare results were obtained by the maximum parsimony method, which is less reliable for amino acid sequences than either the neighbor joining or the maximum likelihood methods. Based on the three most frequent topologies (39 %, 29 % and 24 %; Table 1, A, B
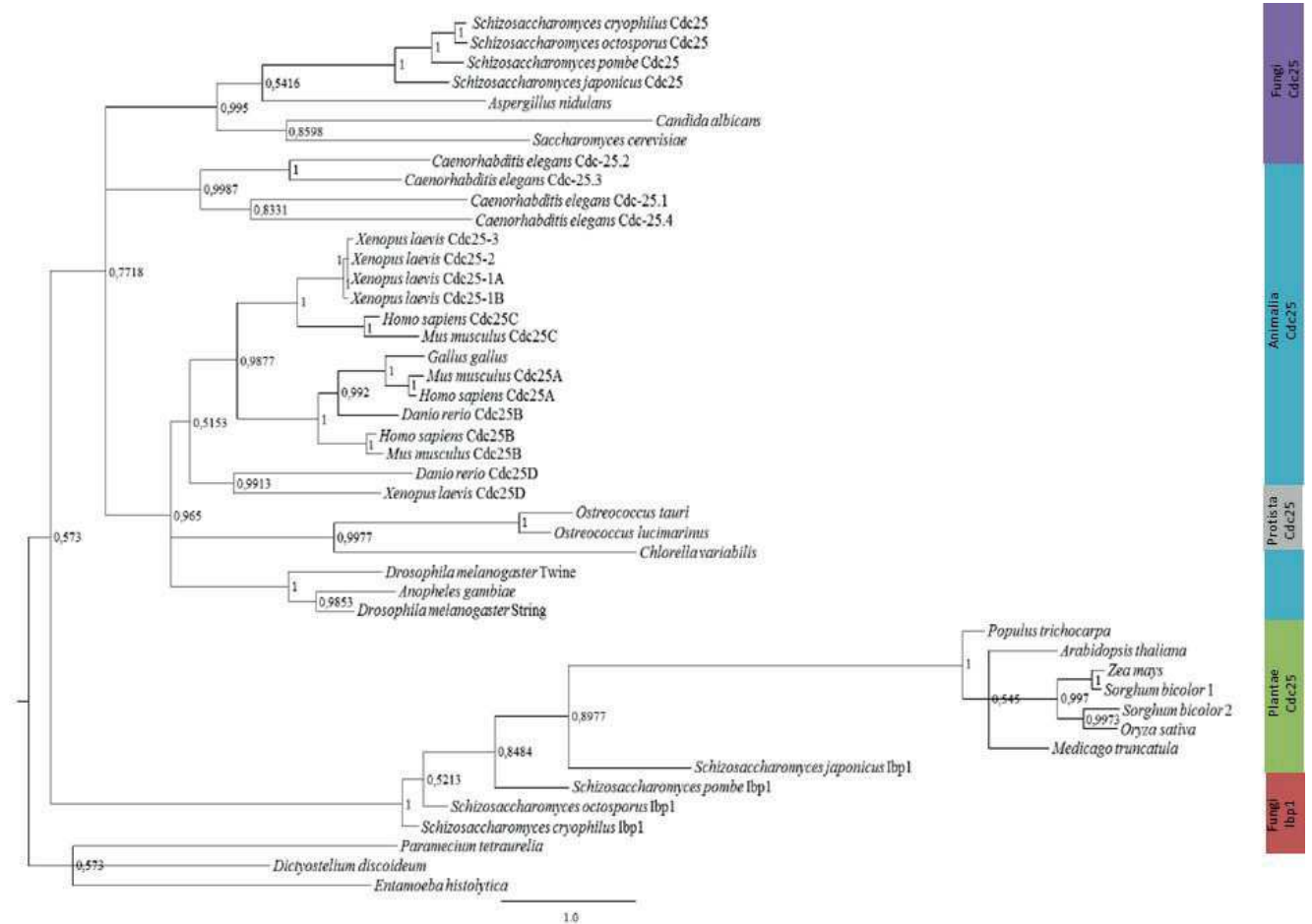

Fig. 11 Bayesian phylogenetic tree of Cdc25 and Ibp1 homologs generated with the WAG + G + F + I model.

and C, respectively), it can be concluded that the Myt1 double-specificity kinase may be the transient variant between an ancient serine / threonine kinase and the Wee1 tyrosine kinase. In that case it is likely that after the separation of animals, plants and fungi, both *wee1* and *myt1* genes remained only in the common ancestor of the animals, meanwhile the ancestors of both plants and fungi could have lost the *myt1* gene. Wee1A and Wee1B proteins probably have been separated in the ancestor of higher animals. The Mik1 kinase may have developed in the common ancestor of the *Schizosaccharomyces* genus, as this protein is specific to fission yeasts. The formerly published conserved motifs in Wee1-like proteins have been verified by our present work, moreover we have significantly extended the number of studied sequences.

All the upstream regulators of Wee1, namely Pom1, Nif1, Cdr1, and Cdr2, were found to be unique in the *Schizosaccharomyces* genus among eukaryotic organisms. This is a valuable, but not surprising result, since cylindrical cell shape is very rare in the universe of unicellular eukaryots.

The other objects of this study were the Cdc25-like phosphatases. Ibp1 proteins of fission yeasts appeared to be homologs of the Cdc25 cell cycle regulator phosphatase, although their biochemical function is completely different. In contrast, Pyp3 phosphatase has the same biochemical function, but it has been found to be less similar to Cdc25 (than Ibp1). In the Cdc25 proteins, a novel conserved region was found with the consensus sequence LxxG(Y/F) in positions 1037-1041 (Fig. 7), which, to our best knowledge, has never been described previously in the literature.

The Cdc25 homologs identified in plants show the largest similarity to Ibp1 proteins. These proteins have a very short sequence (about 150 amino acids), and presumably the plant Cdc25, like Ibp1 proteins, may have played a role in ancient genome doubling events [31, 36]. Some formerly published phylogenetic trees of Cdc25-like proteins have been extended here by studying much more sequences, and their previously suggested topology has been verified.

Size control is a mechanism evolved during billions of years to ensure size homeostasis from one generation to the next in populations. In fission yeast, the players of this mechanism in late G2 phase have been characterized by genetic, biochemical and molecular biological methods during the last 40 years [21, 54, 55]. These actors are the direct regulators of MPF, i.e. the Wee1 kinase and the Cdc25 phosphatase, and also their upstream regulators (Fig. 1). By contrast, we still do not know exactly how cells sense their sizes, it is challenged whether this mechanism is mainly based on dilution of inhibitor molecules, or rather on accumulation of activator ones, or both [3, 28, 56–59]. In this study, we have performed a phylogenetic study of these proteins, which may also help to solve this fundamental problem of cell biology.

References

[1] Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. "Essential Cell Biology", 4th ed., Garland Publishing, Inc., New York, USA, 2014.

[2] Gould, K. L. "Protein Kinases Driving the Cell Cycle", In: Egel, R. (ed.) The Molecular Biology of *Schizosaccharomyces pombe*, 1st ed., Springer-Verlag, Berlin, Heidelberg, Germany, 2014, pp. 27–40.
https://doi.org/10.1007/978-3-662-10360-9_3

[3] Facchetti, G., Knapp, B., Flor-Parra, I., Chang, F., Howard, M. "Reprogramming Cdr2-Dependent Geometry-Based Cell Size Control in Fission Yeast", Current Biology, 29(2), pp. 350–358, 2019.
https://doi.org/10.1016/j.cub.2018.12.017

[4] Aligue, R., Wu, L., Russell, P. "Regulation of *Schizosaccharomyces pombe* Wee1 Tyrosine Kinase", The Journal of Biological Chemistry, 272(20), pp. 13320–13325, 1997.
https://doi.org/10.1074/jbc.272.20.13320

[5] Squire, C. J., Dickson, J. M., Ivanovic, I., Baker, E. N. "Structure and Inhibition of the Human Cell Cycle Checkpoint Kinase, Wee1A Kinase: An Atypical Tyrosine Kinase with a Key Role in CDK1 Regulation", Structure, 13(4), pp. 541–550, 2005.
https://doi.org/10.1016/j.str.2004.12.017

[6] Enders, G. H. "Gauchos and ochos: a Wee1-Cdk tango regulating mitotic entry", Cell Division, 5(12), pp. 1–7, 2010.
https://doi.org/10.1186/1747-1028-5-12

[7] Lundgren, K., Walworth, N., Booher, R., Dembski, M., Kirschner, M., Beach, D. "mik1 and wee1 cooperate in the inhibitory tyrosine phosphorylation of cdc2", Cell, 64(6), pp. 1111–1122, 1991.
https://doi.org/10.1016/0092-8674(91)90266-2

[8] Okamoto, K., Nakajo, N., Sagata, N. "The existence of two distinct Wee1 isoforms in *Xenopus*: implications for the developmental regulation of the cell cycle", The EMBO Journal, 21(10), pp. 2472–2484, 2002.
https://doi.org/10.1093/emboj/21.10.2472

[9]   Nakanishi, M., Ando, H., Watanabe, N., Kitamura, K., Ito, K., Okayama, H., Miyamoto, T., Agui, T., Sasaki, M. "Identification and characterization of human Wee1B, a new member of the Wee1 family of Cdk-inhibitory kinases", Genes to Cells, 5(10), pp. 839–847, 2000.
https://doi.org/10.1046/j.1365-2443.2000.00367.x

[10] Oh, J. S., Han, S. J., Conti, M. "Wee1B, Myt1, and Cdc25 function in distinct compartments of the mouse oocyte to control meiotic resumption", Journal of Cell Biology, 188(2), pp. 199–207, 2010.
https://doi.org/10.1083/jcb.200907161

[11] Wilson, M. A., Hoch, R. V., Ashcroft, N. R., Kosinski, M. E., Golden, A. "A *Caenorhabditis elegans wee1* homolog is expressed in a temporally and spatially restricted pattern during embryonic development", Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression, 1445(1), pp. 99–109, 1999.
https://doi.org/10.1016/S0167-4781(99)00027-5

[12] Yoon, S., Kawasaki, I., Shim, Y.-H. "CDC-25.1 controls the rate of germline mitotic cell cycle by counteracting WEE-1.3 and by positively regulating CDK-1 in *Caenorhabditis elegans*", Cell Cycle, 11(7), pp. 1354–1363, 2012.
https://doi.org/10.4161/cc.19755

[13] Liu, F., Rothblum-Oviatt, C., Ryan, C. E., Piwnica-Worms, H. "Overproduction of Human Myt1 Kinase Induces a G_2 Cell Cycle Delay by Interfering with the Intracellular Trafficking of Cdc2-Cyclin B1 Complexes", Molecular and Cellular Biology, 19(7), pp. 5113–5123, 1999.
https://doi.org/10.1128/MCB.19.7.5113

[14] Nakajima, H., Yonemura, S., Murata, M., Nakamura, N., Piwnica-Worms, H., Nishida, E. "Myt1 protein kinase is essential for Golgi and ER assembly during mitotic exit", Journal of Cell Biology, 181(1), pp. 89–103, 2008.
https://doi.org/10.1083/jcb.200708176

[15] Morgan, D. O. "The Cell Cycle: Principles and Control", 1st ed., New Science Press Ltd, London, UK, 2007.

[16] Allard, C. A. H., Opalko, H. E., Liu, K., Medoh, U., Moseley, J. B. "Cell size-dependent regulation of Wee1 localization by Cdr2 cortical nodes", Journal of Cell Biology, 217(5), pp. 1589–1599, 2018.
https://doi.org/10.1083/jcb.201709171

[17] Martin, S. G., Berthelot-Grosjean, M. "Polar gradients of the DYRK-family kinase Pom1 couple cell length with the cell cycle", Nature, 459(7248), pp. 852–856, 2009.
https://doi.org/10.1038/nature08054

[18] Hachet, O., Berthelot-Grosjean, M., Kokkoris, K., Vincenzetti, V., Moosbrugger, J., Martin, S. G. "A Phosphorylation Cycle Shapes Gradients of the DYRK Family Kinase Pom1 at the Plasma Membrane", Cell, 145(7), pp. 1116–1128, 2011.
https://doi.org/10.1016/j.cell.2011.05.014

[19] Wood, E., Nurse, P. "Pom1 and cell size homeostasis in fission yeast", Cell Cycle, 12(19), pp. 3228–3236, 2013.
https://doi.org/10.4161/cc.26462

[20] Horváth, A., Rácz-Mónus, A., Buchwald, P., Sveiczer, Á. "Cell length growth patterns in fission yeast reveal a novel size control mechanism operating in late G2 phase", Biology of the Cell, 108(9), pp. 259–277, 2016.
https://doi.org/10.1111/boc.201500066

[21] Sveiczer, Á., Horváth, A. "How do fission yeast cells grow and connect growth to the mitotic cycle?", Current Genetics, 63(2), pp. 165–173, 2017.
https://doi.org/10.1007/s00294-016-0632-0

[22] Ducommun, B., Draetta, G., Young, P., Beach, D. "Fission yeast CDC25 is a cell-cycle regulated protein", Biochemical and Biophysical Research Communications, 167(1), pp. 301–309, 1990.
https://doi.org/10.1016/0006-291X(90)91765-K

[23] Bordo, D., Bork, P. "The rhodanese / Cdc25 phosphatase superfamily: Sequence-structure-function relations", EMBO Reports, 3(8), pp. 741–746, 2002.
https://doi.org/10.1093/embo-reports/kvf150

[24] Hofmann, K., Bucher, P., Kajava, A. V. "A model of Cdc25 phosphatase catalytic domain and Cdk-interaction surface based on the presence of a rhodanese homology domain", Journal of Molecular Biology, 282(1), pp. 195–208, 1998.
https://doi.org/10.1006/JMBI.1998.1998

[25] Rudolph, J. "Cdc25 Phosphatases: Structure, Specificity, and Mechanism", Biochemistry, 46(12), pp. 3595–3604, 2007.
https://doi.org/10.1021/bi700026j

[26] Ashcroft, N. R., Kosinski, M. E., Wickramasinghe, D., Donovan, P. J., Golden, A. "The four *cdc25* genes from the nematode *Caenorhabditis elegans*", Gene, 214(1–2), pp. 59–66, 1998.
https://doi.org/10.1016/S0378-1119(98)00228-5

[27] Mulvihill, D. P., Petersen, J., Ohkura, H., Glover, D. M., Hagan, I. M. "Plo1 Kinase Recruitment to the Spindle Pole Body and Its Role in Cell Division in *Schizosaccharomyces pombe*", Molecular Biology of the Cell, 10(8), pp. 2771–2785, 1999.
https://doi.org/10.1091/mbc.10.8.2771

[28] Keifenheim, D., Sun, X.-M., D'Souza, E., Ohira, M. J., Magner, M., Mayhew, M. B., Marguerat, S., Rhind, N. "Size-Dependent Expression of the Mitotic Activator Cdc25 Suggests a Mechanism of Size Control in Fission Yeast", Current Biology, 27(10), pp. 1491–1497, 2017.
https://doi.org/10.1016/j.cub.2017.04.016

[29] Edgar, B. A., O'Farrell, P. H. "Genetic Control of Cell Division Patterns in the Drosophila Embryo", Cell, 57(1), pp. 177–187, 1989.
https://doi.org/10.1016/0092-8674(89)90183-9

[30] Courtot, C., Fankhauser, C., Simanis, V., Lehner, C. F. "The Drosophila cdc25 homolog twine is required for meiosis", Development, 116(2), pp. 405–416, 1992.

[31] Landrieu, I., da Costa, M., De Veylder, L., Dewitte, F., Vandepoele, K., Hassan, S., Wieruszeski, J.-M., Faure, J.-D., Van Montagu, M., Inzé, D., Lippens, G. "A small CDC25 dual-specifity tyrosine-phosphatase isoform in *Arabidopsis thaliana*", Proceedings of the National Academy of Sciences of the United States of America (PNAS), 101(36), pp. 13380–13385, 2004.
https://doi.org/10.1073/pnas.0405248101

[32] Boutros, R., Dozier, C., Ducommun, B. "The when and wheres of CDC25 phosphatases", Current Opinion in Cell Biology, 18(2), pp. 185–191, 2006.
https://doi.org/10.1016/J.CEB.2006.02.003

[33] Lindqvist, A., Källström, H., Lundgren, A., Barsoum, E., Rosenthal, C. K. "Cdc25B cooperates with Cdc25A to induce mitosis but has a unique role in activating cyclin B1-Cdk1 at the centrosome", Journal of Cell Biology, 171(1), pp. 35–45, 2005.
https://doi.org/10.1083/jcb.200503066

[34] Hoffmann, I., Clarke, P. R., Marcote, M. J., Karsenti, E., Draetta, G. "Phosphorylation and activation of human cdc25-C by cdc2-cyclin B and its involvement in the self-amplification of MPF at mitosis", The EMBO Journal, 12(1), pp. 53–63, 1993.
https://doi.org/10.1002/j.1460-2075.1993.tb05631.x

[35] Millar, J. B., Lenaers, G., Russel, P. "Pyp3 PTPase acts as a mitotic inducer in fission yeast", The EMBO Journal, 11(13), pp. 4933–4941, 1992.
https://doi.org/10.1002/j.1460-2075.1992.tb05600.x

[36] Snaith, H. A., Marlett, J., Forsburg, S. L. "Ibp1p, a novel Cdc25-related phosphatase, suppresses *Schizosaccharomyces pombe hsk1* (*cdc7*)", Current Genetics, 44(1), pp. 38–48, 2003.
https://doi.org/10.1007/s00294-003-0424-1

[37] PomBase "The scientific resource for fission yeast", [online] Available at: https://www.pombase.org/ [Accessed: 28 January 2019]

[38] BROAD Institute "Data, Software and Tools", [online] Available at: https://www.broadinstitute.org/ [Accessed: 28 January 2019]

[39] UniProt "UniProt Knowledgebase", [online] Available at: https://www.uniprot.org/ [Accessed: 11 February 2019]

[40] National Center for Biotechnology Information (NCBI) "National Center for Biotechnology Information", [online] Available at: https://www.ncbi.nlm.nih.gov/ [Accessed: 11 February 2019]

[41] Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Research, 25(17), pp. 3389–3402, 1997.
https://doi.org/10.1093/nar/25.17.3389

[42] Pfam "The Pfam protein families database", [online] Available at: https://pfam.xfam.org/ [Accessed: 18 February 2019]

[43] Clustal W / Clustal X "Multiple alignment of nucleic acid and protein sequences", [online] Available at: http://www.clustal.org/clustal2/ [Accessed: 18 February 2019]

[44] Prank "webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser", [online] Available at: https://www.ebi.ac.uk/Tools/msa/prank/ [Accessed: 18 February 2019]

[45] WebLogo 3 "A sequence logo generator", [online] Available at: http://weblogo.threeplusone.com [Accessed: 19 February 2019]

[46] Molecular Evolutionary Genetics Analysis (MEGA) "Molecular Evolutionary Genetics Analysis", [online] Available at: https://www.megasoftware.net/ [Accessed: 20 February 2019]

[47] Jones, D. T., Taylor, W. R., Thornton, J. M. "The rapid generation of mutation data matrices from protein sequences", Bioinformatics, 8(3), pp. 275–282, 1992.
https://doi.org/10.1093/bioinformatics/8.3.275

[48] MrBayes "Bayesian Inference of Phylogeny", [computer program] Available at: http://mrbayes.sourceforge.net/ [Accessed: 20 February 2019]

[49] ProtTest "Fast selection of best-fit models of protein evolution", [computer program] Available at: https://github.com/ddarriba/prottest3 [Accessed: 20 February 2019]

[50] FigTree "Molecular evolution, phylogenetics and epidemiology", [computer program] Available at: http://tree.bio.ed.ac.uk/software/figtree/ [Accessed: 20 February 2019]

[51] Tracer "A graphical tool for visualization and diagnostics of MCMC output", [computer program] Available at: http://beast.bio.ed.ac.uk/tracer [Accessed: 20 February 2019]

[52] Sorrell, D. A., Marchbank, A., McMahon, K., Dickinson, R. J., Rogers, H. J., Francis, D. "A *WEE1* homologue from *Arabidopsis thaliana*", Planta, 215(3), pp. 518–522, 2002.
https://doi.org/10.1007/s00425-002-0815-4

[53] Nakajo, N., Deno, Y., Ueno, H., Kenmochi, C., Shimuta, K., Sagata, N. "Temporal and spatial expression patterns of Cdc25 phosphatase isoforms during early *Xenopus* development", The International Journal of Developmental Biology, 55(6), pp. 627–632, 2011.
https://doi.org/10.1387/ijdb.113287nn

[54] Navarro, F. J., Weston, L., Nurse, P. "Global control of cell growth in fission yeast and its coordination with the cell cycle", Current Opinion in Cell Biology, 24(6), pp. 833–837, 2012.
https://doi.org/10.1016/j.ceb.2012.10.015

[55] Wood, E., Nurse, P. "Sizing up to Divide: Mitotic Cell-Size Control in Fission Yeast", Annual Review of Cell and Developmental Biology, 31(1), pp. 11–29, 2015.
https://doi.org/10.1146/annurev-cellbio-100814-125601

[56] Sveiczer, A., Novak, B., Mitchison, J. M. "The size control of fission yeast revisited", Journal of Cell Science, 109(12), pp. 2947–2957, 1996.

[57] Sveiczer, A., Novak, B., Mitchison, J. M. "Mitotic control in the absence of cdc25 mitotic inducer in fission yeast", Journal of Cell Science, 112(7), pp. 1085–1092, 1999.

[58] Schmoller, K. M., Skotheim, J. M. "The Biosynthetic Basis of Cell Size Control", Trends in Cell Biology, 25(12), pp. 793–802, 2015.
https://doi.org/10.1016/j.tcb.2015.10.006

[59] Facchetti, G., Chang, F., Howard, M. "Controlling cell size through sizer mechanisms", Current Opinion in Systems Biology, 5, pp. 86–92, 2017.
https://doi.org/10.1016/j.coisb.2017.08.010