Comparative Genomics of Phylogenetically Diverse Unicellular Eukaryotes Provide New Insights into the Genetic Basis for the Evolution of the Programmed Cell Death Machinery

Aurora M. Nedelcu

Abstract Programmed cell death (PCD) represents a significant component of normal growth and development in multicellular organisms. Recently, PCD-like processes have been reported in single-celled eukaryotes, implying that some components of the PCD machinery existed early in eukaryotic evolution. This study provides a comparative analysis of PCD-related sequences across more than 50 unicellular genera from four eukaryotic supergroups: Unikonts, Excavata, Chromalveolata, and Plantae. A complex set of PCD-related sequences that correspond to domains or proteins associated with all main functional classes—from ligands and receptors to executors of PCD—was found in many unicellular lineages. Several PCD domains and proteins previously thought to be restricted to animals or land plants are also present in unicellular species. Noteworthy, the yeast, *Saccharomyces cerevisiae*—used as an experimental model system for PCD research, has a rather reduced set of PCD-related sequences relative to other unicellular species. The phylogenetic distribution of the PCD-related sequences identified in unicellular lineages suggests that the genetic basis for the evolution of the complex PCD machinery present in extant multicellular lineages has been established early in the evolution of eukaryotes. The shaping of the PCD machinery in multicellular lineages involved the duplication, co-option, recruitment, and shuffling of domains already present in their unicellular ancestors.

A. M. Nedelcu (✉)
Department of Biology, University of New Brunswick,
P.O. Box 4400, Fredericton, NB, Canada E3B 5A3
e-mail: anedelcu@unb.ca

Introduction

Programmed cell death (PCD)—an active process resulting in the controlled elimination of unwanted or damaged cells—has long been recognized as a significant component of normal growth and development in multicellular organisms, both animals and plants (e.g., Jacobson et al. 1997; Lam 2004). Recently, PCD-like processes (i.e., diagnostic features such as protoplast shrinking, accumulation of reactive oxygen species, DNA-laddering, externalization of phosphatidylserine, caspase-like activity) have been reported in several unicellular groups – including dinoflagellates, green algae, diatoms, yeasts, kinetoplastids, apicomplexans, and amoebozoans (e.g., Cornillon et al. 1994; Ameisen et al. 1995; Madeo et al. 1999; Vardi et al. 1999; Al-Olayan et al. 2002; Arnoult et al. 2002; Segovia et al. 2003; Bidle and Falkowski 2004; Nedelcu 2006; Moharikar et al. 2006; Bidle et al. 2007; Zuppini et al. 2007; Deponte 2008a; Bidle and Bender 2008), suggesting that some components of the PCD machinery existed early in the evolution of the eukaryotic lineage. Nevertheless, the mechanistic basis for PCD in single-celled organisms and the evolutionary relationships between these PCD-like processes and the better understood forms of PCD in multicellular lineages are still to be deciphered. Several homologues of genes involved in the most studied form of animal PCD, apoptosis, have been identified in unicellular lineages, and their involvement in PCD-like processes addressed (e.g., Madeo et al. 2002; Fahrenkrog et al. 2004; Wissing et al. 2004; Walter et al.

2006; Buettner et al. 2007; Moharikar et al. 2007), yet many others are reportedly missing (Koonin and Aravind 2002).

The increasing availability of genome sequences from various eukaryotic groups provides an opportunity (i) to explore the degree of conservation of the PCD machinery across evolutionarily distant lineages, (ii) to infer which elements might have been present early in evolution, and (iii) to investigate the evolutionary processes (e.g., gene duplication and diversification, co-option, loss or replacement, domain recruitment and shuffling, lateral gene transfer) responsible for the shaping of the PCD machinery in specific lineages. Several recent studies have addressed the early evolution of the PCD machinery and the potential bacterial origin of some of the genes involved in PCD (e.g., Aravind et al. 1999; Koonin and Aravind 2002). However, these studies were based on information from a very limited number of unicellular lineages—mostly yeast and several unicellular lineages thought to be "early-branching eukaryotes"—and suggested that unicellular lineages possess a very limited PCD-related gene toolkit. For instance, of the 33 domains and proteins involved in apoptosis and related pathways investigated by Koonin and Aravind (2002), only 13 are indicated as having potential homologues in unicellular lineages—either eukaryotes or prokaryotes. Furthermore, because some PCD-related sequences appeared to be missing in unicellular eukaryotes but had potential homologues in prokaryotes, it was proposed that an "influx" of bacterial genes occurred in the multicellular ancestor of the "eukaryotic crown group" (note that the term "crown group" is obsolete; current evidence supports the independent evolution of multicellular fungi, plants, and animals from distinct unicellular ancestors; see, e.g., Embley and Martin 2006).

To address the issues discussed above, this study (i) provides a comparative analysis of PCD-related sequences [employing a domain-centered approach (Aravind et al. 2001)] from phylogenetically diverse unicellular lineages and (ii) indicates potential mechanisms involved in the early evolution of the eukaryotic PCD machinery. Although our understanding of the eukaryotic tree has improved greatly in the last decade, the exact relationships among the major eukaryotic groups (in particular, those including unicellular lineages) are uncertain; in addition, the monophyly of some of these groups as well as the root of the eukaryotic tree are still debated (e.g., Keeling et al. 2005; Embley and Martin 2006; Yoon et al. 2008). Five or six major eukaryotic supergroups that diverged from each other early in the evolution of eukaryotes are recognized to date: the Unikonts (i.e., Opisthokonta and Amoebozoa), Chromalveolata, Plantae, Rhizaria, and Excavata—the latter four groups also being known as Bikonts (Cavalier-Smith 2002; Stechmann and Cavalier-Smith 2003; Keeling et al. 2005).

As many as 37 PCD-related sequences have been identified in unicellular species from at least one of the four eukaryotic supergroups investigated in this study (i.e., Unikonts, Excavates, Chromalveolates, and Plantae; at this time, genomic information from the Rhizaria is not available). The phylogenetic distribution of these sequences suggests that the potential (i.e., the genetic basis) for the evolution of the complex PCD machinery present in multicellular lineages was established early in the evolution of eukaryotes. Compared to their counterparts in multicellular lineages, many PCD-related domains in single-celled eukaryotes are found in single-domain proteins or in unique domain combinations, indicating that the early shaping of the PCD machinery in multicellular lineages involved the duplication, co-option, recruitment, and shuffling of domains already present in their unicellular ancestors.

Methods

Several protein databases (e.g., Interpro—http://www.ebi.ac.uk/interpro/; Pfam—http://www.sanger.ac.uk/Software/Pfam/; Prosite—http://www.expasy.org/prosite/; Uniprot—http://www.pir.uniprot.org/; Superfamily—http://supfam.cs.bris.ac.uk/), as well as genome and EST databases (Joint Genome Institute—http://www.jgi.doe.gov/; NCBI—http://www.ncbi.nlm.nih.gov/; Protist EST Program—http://amoebidia.bcm.umontreal.ca/pepdb/), were searched for PCD-related sequences [in particular, the "domains of death" described and used in previous comparative analyses (Aravind et al. 1999, 2001; Koonin and Aravind 2002)]. Initial searches employed (i) text searches using PCD-related keywords and Interpro/Pfam accession numbers corresponding to PCD-related domains, and (ii) Blast searches [tblastn, blastp, psi-Blast (Altschul et al. 1990, 1997)] using sequences from the closest species as queries. Gene and protein sequences retrieved in this manner were checked for the presence of the corresponding PCD-specific domains using SMART, Pfam, and InterProScan (http://smart.embl-heidelberg.de/; http://www.sanger.ac.uk/Software/Pfam/search.shtml; http://www.ebi.ac.uk/InterProScan/); only sequences with domains confidently predicted [using the default cutoffs specific for each domain; see http://www.ebi.ac.uk/interpro/documentation.html and Schultz et al. (1998)] were included in this study. Sequences were aligned with Muscle [http://www.drive5.com/muscle/ (Edgar 2004)]. Phylogenetic analyses (gaps and unalignable regions excluded) were performed using MrBayes v3.0B4 (http://mrbayes.csit.fsu.edu/; mixed amino acid model; 3,500,000 generations; 100 sample frequency; 5000 burn-in) and PhyML (http://atgc.lirmm.fr/phyml/; 200 replicates; four-category gamma distribution; proportion of variable sites

estimated from the data; best-fit amino acid model indicated by ProtTest).

Results and Discussion

A Complex Set of PCD-Related Sequences in Phylogenetically Diverse Unicellular Lineages

Several protein as well as genome and EST databases (see Supplementary Table 1) have been searched for domains and proteins known to be associated with PCD in animals and/or land plants (see Methods). As genomic information for many unicellular groups is still limited, the inability to detect a particular PCD-related protein or domain in the available sequence data cannot be taken, at this time, as indicating that the sequence is absent in that group. On the other hand, because most of the PCD-related domains included in Fig. 1 were inferred using domain prediction tools (see Methods), they require functional confirmation.

Nevertheless, many PCD-related sequences were found in more than 50 unicellular genera from four eukaryotic supgroups: Unikonts (amoebozoans, choanoflagellates, fungi), Chromalveolata (cryptomonads, pelagophytes, oomycetes, diatoms, haptophytes, ciliates, dinoflagellates, apicomplexans), Plantae (glaucophytes, red algae, green algae), and Excavata (kinetoplastids, euglenoids, jakobids, diplomonads, trichomonads, heteroloboseans). Figure 1 provides a list of 37 PCD-related domains [i.e., "domains of death" (Koonin and Aravind 2002)] and proteins associated with all main functional classes (from ligands and receptors to executors of PCD) found in unicellular species from at least one eukaryotic supergroup; a succinct discussion of their role in PCD and their phylogenetic distribution is provided later in this section.

Overall, among the unicellular lineages investigated in this study, the choanoflagellates, Amoebozoa, and Excavata appear to posses the largest number of PCD-related sequences (Fig. 1). Of the unicellular species with an available genome sequence, the choanoflagellate *Monosiga brevicollis*, considered to be a close relative of Metazoa (King 2004), the amoebozoan *Dictyostelium discoideum*, the excavate *Naegleria gruberi*, and the green alga *Chlamydomonas reinhardtii* have the largest PCD gene complements.

Noteworthy, the yeast, *Saccharomyces cerevisiae*—used as a model system for PCD research—has a rather reduced set of PCD-related sequences relative to other unicellular lineages. This finding is consistent with earlier reports (Aravind et al. 2000) that ca. 300 genes—most of which belong to functionally connected groups—have been lost (and ca. 300 other genes have diverged beyond recognition) in the lineage leading to *S. cerevisiae* (a hemiascomycete/ budding yeast) after its divergence from the lineage leading

to *Schizosaccharomyces pombe* (an archiascomycete/fission yeast). Losses of sets of genes in some lineages can be understood in terms of lineage-specific differences in their biology and/or ecology; it is possible that some of the PCD-related sequences missing in *S. cerevisiae* (and other budding yeasts) were involved in pathways that have been lost (or reshaped) during this lineage's adaptation to its unique lifestyle and/or mode of growth and reproduction.

The finding that the closest unicellular relatives of multicellular animals and plants—the choanoflagellates and the green algae, respectively—have such a complex PCD-related set of sequences (including some sequences thought to be restricted to animals or plants; see discussion in next sections) suggests that the evolution of the complex PCD machinery known in multicellular lineages involved the co-option of sequences already present in their unicellular ancestors. Because (i) genomic information from many unicellular eukaryotic groups is still limited, and (ii) the relationships among the four eukaryotic supergroups—as well as the monophyly of some of the major eukaryotic groups—are still debated (e.g., Keeling et al. 2005; Yoon et al. 2008), the eukaryotic ancestral set of PCD-related sequences cannot be inferred at this time. However, based on the available information, several conclusions can be drawn.

Of the 37 entries in Fig. 1, as many as 23 PCD-related sequences appear to be shared by all four eukaryotic supergroups and, thus, are likely to have been present in their last common ancestor. In addition, eight other PCD-related sequences are shared by three of the four supergroups. These include the BAG and API5 domains—shared by Unikonts, Excavates, and Plantae, to the exclusion of Chromalveolates; the NB-ARC, NACHT, and MDM35 domains—shared by Unikonts, Chromalveolates, and Plantae, to the exclusion of Excavates; and the DEATH, DED, and sestrin domains—shared by Unikonts, Excavates, and Chromalveolates, to the exclusion of Plantae. Thus, depending on the phylogenetic relationships among and within these four eukaryotic supergroups—and in the absence of lateral gene transfer—between 23 and 31 PCD-related sequences can be hypothesized to have been present in their last common ancestor. For instance, if the root of the eukaryotic tree were between the Unikonts and the Bikonts—as proposed by Stechmann and Cavalier-Smith (2003)—the 31 sequences that are shared between Unikonts and Bikonts would be hypothesized to have been present in the last common ancestor of eukaryotes.

The remaining six PCD-related sequences included in Fig. 1 appear to have specifically evolved in the Unikont or Plantae lineages, from sequences already present in their unicellular ancestors. These include (i) the tumor suppressor p53—to date reported only in animals, choanoflagellates, and the amoebozoan *Entamoeba histolytica* (Mendoza et al.
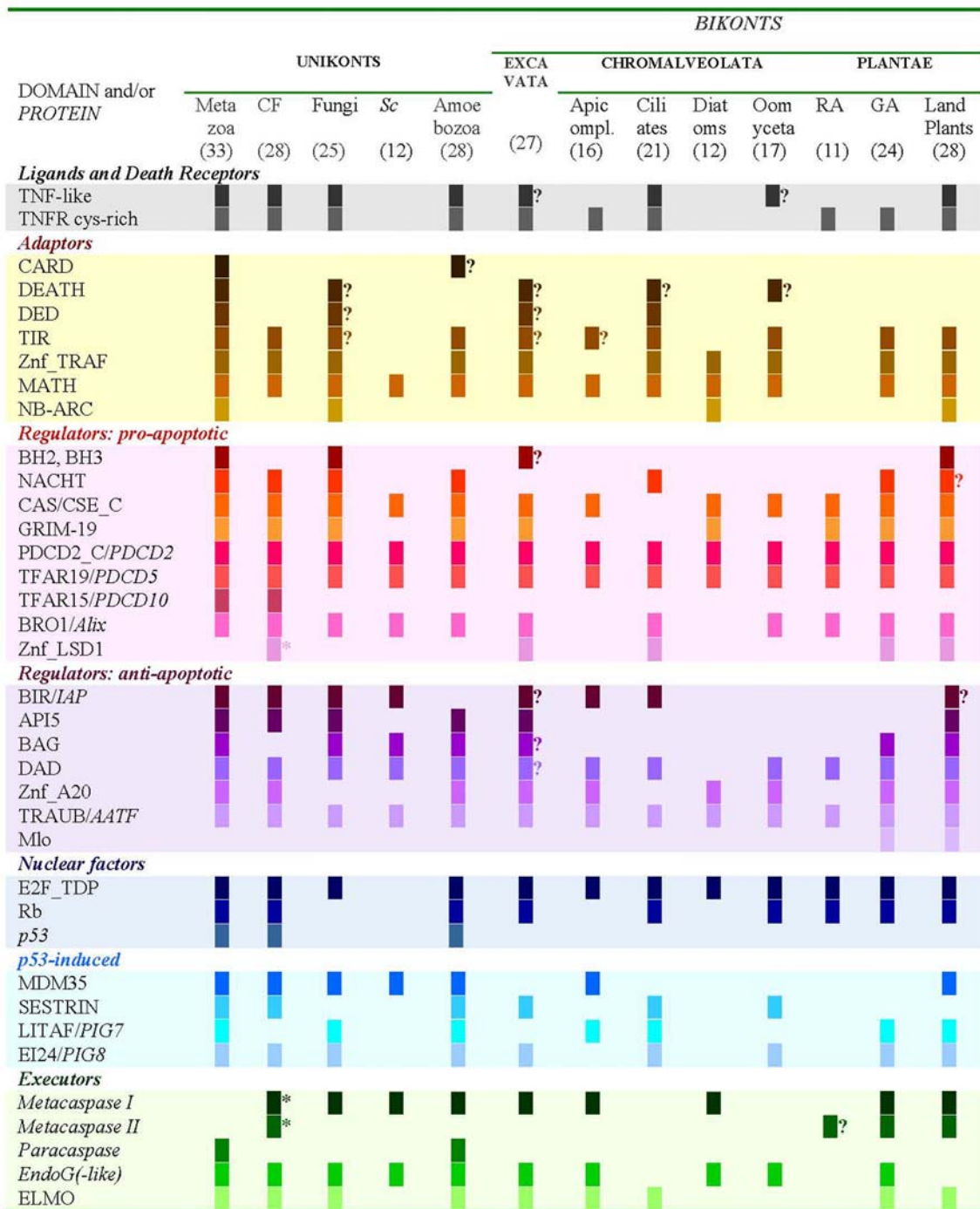
Fig. 1 Comparative analysis of PCD-related domains and proteins (in *italics*) across four eukaryotic supergroups: Unikonts, Excavata, Chromalveolata, and Plantae (see Supplementary Table 1 for species). Only sequences identified in at least one unicellular lineage are included. CF, choanoflagellates; *Sc*, *Saccharomyces cervisiae*; Apicompl., Apicomplexa; RA, red algae; GA, green algae. Numbers in brackets indicate the number of PCD-related domains or proteins (of the total 37 PCD-related sequences included in this analysis) identified in each lineage. A question mark (?) denotes cases in which the finding of a domain/protein is restricted to one instance (although sequence information from several species is available), thus allowing for the possibility of a prediction artifact, contamination, or lateral gene transfer event; asterisks indicate known cases of lateral gene transfer from the Plantae lineage (see text for discussion). See text for full names, descriptions, and references to Interpro accession numbers for each domain or protein

2003; Nedelcu and Tan 2007; see discussion below); (ii) the programmed cell death protein 10 (PDCD10)—found only in animals and choanoflagellates; (iii) the paracaspases,

present only in animals and amoeobozoans, and the CARD domain, present only in animals and possibly Amoebozoa; and (iv) type II metacaspase and a family of plant-specific

cell death inhibitors—the Mlo family, found only in plants and their algal relatives (Fig. 1).

Several additional proteins that are involved in PCD but function in other conserved vital cellular activities as well—and thus are ubiquitous among eukaryotic lineages (Ekert and Vaux 2005; Modjtahedi et al. 2006)—were not incorporated in this analysis; these include, for instance, Bit1, a Bcl-2 inhibitor of transcription/precursor of mitochondrial peptidyl-tRNA hydrolase 2; PIG3, a p53-induced gene/proline oxidase; Beclin1, a Bcl-2 interacting protein/autophagy protein; PDCD11, programmed cell death protein 11/ribosomal RNA biogenesis protein; and several mitochondrial proteins that are also involved in bioenergetic and redox metabolism, such as cytochrome *c*, the apoptosis inducing factor (a FAD-dependent oxidoreductase), and the components of the permeability transition pore complex. Also, Bax Inhibitor-1 (BI-1), a cell death suppressor in animals and plants, might also be added to this list, as BI-1-like sequences (though without a canonical BI-1 domain) have been found in many unicellular lineages [e.g., yeast, green algae, amoebozoans, apicomplexans (Huckelhoven 2004)], and at least the yeast BI-1 is able to block Bax-induced cell death (Chae et al. 2003).

Ligands, Receptors, and Adaptors

In mammals, apoptosis can be induced via the activation of death-inducing signaling complexes at the plasma membrane. These include: ligands (e.g., the tumor necrosis factor, TNF); death receptors, such as Fas, TNFR1, and TNFR2 (which contain multiple copies of a cysteine-rich extracellular domain, TNFR, and an intracellular Death domain); and adaptors, such as, TRAF, MATH, CARD, DED, DD, TIR, TRADD, and FADD.

TNF and TNFR-like proteins are thought to be specific to metazoans (Koonin and Aravind 2002). Nevertheless, putative TNF-like domains or signatures (IPR008983) and TNFR/NGFR cysteine-rich region signatures (IPR001368) have been predicted by Superfamily or Prosite in several unicellular lineages (e.g., the excavates, *Giardia lamblia*, *Trichomonas vaginalis*, *Leishmania* spp.; the amoebozoans, *D. discoideum* and *Entamoeba dispar*; the choanoflagellate, *M. brevicollis*; the ciliates, *Paramecium tetraurelia* and *Tetrahymena thermophila*; the apicomplexan, *Theileria parva*; the oomycetes, *Phytophptora* spp.; the green algae, *Ostreococcus* spp., *Chlorella* sp., and *C. reinhardtii*; and the red alga, *Cyanidioschyzon merolae*) (Fig. 1).

TRAFs (TNF receptor-associated factors) are adaptor proteins that interact with TNF receptors; they comprise three structural domains—a RING-type Zn finger (IPR001841), one to seven TRAF-type zinc fingers (Znf_TRAF; IPR001293), and a MATH (Meprin and TRAF homology; IPR002083) domain. Although legitimate TRAF

proteins (i.e., possessing a recognizable TRAF domain; IPR012227) could not be found outside Metazoa, TRAF zinc finger and MATH domains were predicted by Pfam in multiple instances in all four eukaryotic supergroups in Fig. 1 (in some cases together, e.g., in fungi and in the amoebozoan *D. discoideum*, and even coupled with a RING domain—in *D. discoideum* and the excavates, *Leishmania* spp.).

Because of their sequence similarity and similar α-helical fold, it was suggested that three of the other adaptors found only in animal apoptosis proteins, CARD (caspase recruitment domain; IPR001315), DED (death effector domain; IPR001875), and DD (death domain; IPR000488) have evolved from a common ancestor before the divergence of the extant animal lineages (Aravind et al. 2001). Noteworthily, proteins containing putative CARD (IPR001315), DED (IPR001875), or Death (IPR000488) domains were found in unicellular taxa from four of the five eukaryotic supergroups (Fig. 1), suggesting a possible earlier origin and diversification for this family. These include putative CARD domains predicted by Prosite in the amoebozoan *E. histolytica* and the excavate *Leishmania major*; putative Death domains predicted by Prosite and ProfileScan in the ciliate *P. tetraurelia*, the excavates, *L. major* and *Trypanosoma cruzi*, and the oomycete *P. ramorum*; and putative DED domains, predicted by Prosite and ProfileScan, in the ciliates *T. thermophila* and *P. tetraurelia* and the excavates *T. vaginalis* and *Naegleria gruberi*.

Similarly, as the TIR (Toll/IL-1R homologous region; IPR000157) domain was not detected in fungi or "early-branching eukaryotes," it was hypothesized that TIR could have been acquired either from the mitochondrial precursor (and later lost in multiple eukaryotic lineages) or through lateral gene transfer from bacteria to the multicellular ancestor of the "crown" eukaryotes (Koonin and Aravind 2002). Nevertheless, putative TIR domains are predicted by Pfam, Prosite, or Smart in several unicellular lineages, including the excavate *T. vaginalis*, the ciliate *P. tetraurelia*, the apicomplexan *P. falciparum*, the choanoflagellate *M. brevicollis*, the amoebozoans *Dictyostelium* spp., the oomycetes *Phytophtora* spp., the green algae *C. reinhardtii* and *Micromonas* spp., and two chromalveolate species that belong to lineages not included in Fig. 1 (i.e., the haptophyte *Emiliania huxleyi* and the pelagophyte *Aureococcus anophagefferens*). The identification of TIR-containing proteins among unicellular lineages from phylogenetically diverse groups (Fig. 1) is consistent with an early acquisition for this domain. Last, although not associated with CARD and/or DD domains as in metazoans, the NB-ARC domain (IPR002182)—a signaling motif shared by plant resistance gene products and regulators of cell death in animals (van der Biezen and Jones 1998)—was also

predicted by Pfam in several fungal and chromalveolate (i.e., the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, the pelagophyte *A. anophagefferens*, and the haptophyte *E. huxleyi*) proteins (Fig. 1).

PCD Regulators: Proapoptotic

Cell death is controlled by many regulators, which either have an inhibitory effect on PCD (antiapoptotic) or block the protective effect of inhibitors (proapoptotic). Bcl-2 is a family that contains both pro (e.g., Bax)- and anti (e.g., Bcl-2)-apoptotic proteins with BH1, BH2, BH3, or BH4 motifs. Noteworthy, while Bcl-2-like proteins have not been found outside Metazoa, BH2 and BH3 motif signatures (PS01258, PS01259) are predicted by Prosite in several fungi, dynoflagellates (*Gonyaulx polyedra*, *Alexandrium* spp., *Pyrocystis* spp.), excavates, and plants.

Similarly, the NACHT (NAIP, CIIA, HET-E and TPI) domain; (IPR007111) is a nucleoside triphosphatase (NTPase) domain found in animal apoptosis proteins (both antiapoptotic—the neuronal apoptosis inhibitor protein, NAIP; and proapoptotic—CARD4) as well as in a protein, HET-E, responsible for vegetative incompatibility (a form of PCD) in the fungus *Podospora anserina* (Koonin and Aravind 2000). The presence of NACHT domains in PCD-related proteins in both animals and fungi was seen as evidence for an ancient role of NACHT in PCD—preceding the radiation of animals and fungi (Koonin and Aravind 2000). Interestingly, putative NACHT domains (associated with WD40 repeats, as in nematode proteins) have been predicted by Prosite in ciliates, amoebozoans, and green algae (Fig. 1), suggesting that the NACHT domain could be even older than previously proposed, possibly preceding the Chromalveolata/Plantae/Unikont divergence. Noteworthily, NACHT NTPases are a sister group of another family of ATPases, the AP-ATPases, which include the human apoptotic effector APAF-1 and numerous plant proteins involved in stress and disease responses (Koonin and Aravind 2002).

Another group of pro-apoptotic regulators consists of the mammalian CAS (cellular apoptosis susceptibility; IPR005043) proteins, which are homologous to the yeast chromosome-segregation protein, CSE1 (Brinkmann et al. 1995); they are involved in both cellular apoptosis and proliferation, presumably by facilitating the nuclear import of proteins (such as p53 and other transcription factors) (Brinkmann et al. 1995). A conserved function for these proteins is supported by their presence in all four eukaryotic supergroups in Fig. 1. Likewise, GRIM-19 (gene associated with retinoic-interferon-induced mortality 19; IPR009346)—described as a death regulator that interacts with Stat3 (a transcription factor with important roles in cell growth and antiapoptosis in humans) (Lufei et al.

2003)—was found to be widely distributed across eukaryotes, though apparently missing in yeast (Fig. 1).

Several programmed cell death (PCD) proteins have also been shown to be expressed or up-regulated during apoptosis in animals. Programmed cell death protein 2 (PDCD2) is expressed during apoptosis of lymphoid and myeloid cells and thus may play an important role in cell death and/or in regulation of cell proliferation (Vaux and Hacker 1995). PDCD2 proteins contain a PDCD2_C terminal domain (IPR007320) and a zinc finger, MYND-type (IPR002893). Interestingly, proteins containing the PDCD-2 C terminal domain were found in all four eukaryotic supergroups (Fig. 1). However, in most unicellular lineages, only the PDCD-2 C terminal domain is present in these proteins (for exceptions, see discussion below).

Another programmed cell death protein, PDCD5 or TFAR19 (TF-1 cell apoptosis-related gene 19 protein) was shown to be up-regulated in tumor cells undergoing apoptosis (Liu et al. 1999). Notably, the DNA-binding TFAR19 domain (IPR002836) was found in lineages from all four major eukaryotic groups in Fig. 1. PDCD protein 6, or ALG-2 (apoptosis-linked gene 2), is a calcium-binding protein of the penta-EF-hand family that is also essential for the execution of apoptosis (Jung et al. 2001; Krebs et al. 2002); Alix/AIP1 (ALG-2-interacting protein X/apoptosis-linked gene 2-interacting protein 1)—an adaptor protein that contains a BRO1 domain—can bind to ALG-2 and regulate caspase-dependent and -independent cell death (Sadoul 2006). ALG2-like proteins and proteins containing the BRO1 domain (IPR004328) were found in all major eukaryotic groups (Fig. 1). In contrast, PDCD protein 10 (PDCD10 or TFAR15; IPR009652), of unknown function, was found only in metazoans and their unicellular relative, the choanoflagellates (Fig. 1).

Finally, LSD1 is a putative Zn finger (IPR005735) thought to play a role in the regulation of transcription (via either repression of a prodeath pathway or activation of an antideath pathway) in response to signals emanating from cells undergoing pathogen-induced hypersensitive cell death (a form of PCD) in plants (Lam 2004). Although previously thought to be specific to land plants, proteins containing one or more LSD1 Zn fingers were also found in excavates, ciliates, green algae, and choanoflagellates (for the latter, see discussion below).

PCD Regulators: Antiapoptotic

Among the many described antiapoptotic regulators, the defender against death (DAD) proteins can cause apoptosis if mutated (Nakashima et al. 1993). Proteins with a putative DAD domain (IPR003038) were predicted by Pfam in unicellular lineages from all four eukaryotic supergroups (Fig. 1). Notably, the *dad1* homologue in *C. reinhardtii*

was recently shown to be down-regulated with the onset of PCD in UV-exposed cells (Moharikar et al. 2007). Another rather conserved antiapoptotic factor is the apoptosis antagonizing transcription factor (AATF)—a protein that contains a Traub domain (IPR012617), which also appears to be widely distributed across eukaryotes (Fig. 1).

Several other proteins that act as antiapoptotic regulators are known. The inhibitors of apoptosis proteins (IAPs) are a family of polypeptides that contain the BIR domain (baculovirus inhibitor of apoptosis protein repeat; or proteinase inhibitor I32, inhibitor of apoptosis—IPR001370). Although initially described in metazoans, putative BIR domains were also predicted by Pfam and Smart in fungi (including yeast), choanoflagellates, ciliates, excavates (*N. gruberi*), and apicomplexans (*Plasmodium* spp.) (Fig. 1). Likewise, BAG proteins also have antiapoptotic activity, by increasing the anti-cell-death function of Bcl-2 (Doong et al. 2002). Notably, while Bcl-2 proteins have not been identified outside metazoans, putative BAG domains (IPR003103) are predicted by Pfam in plants, fungi (including yeast), excavates (*N. gruberi*) and green algae, and by Prosite in the amoeba, *E. dispar* (Fig. 1).

The apoptosis inhibitory protein 5 (API5) is an additional antiapoptotic factor, which in humans prevents PCD induced by the deprivation of growth factors (Tewari et al. 1997). Interestingly, API5 domains (IPR008383) were also predicted by Pfam in plants, as well as in several unicellular lineages from the Excavata and Unikonts groups (Fig. 1). Similarly, A20 is known as an inhibitor of cell death in animals (DeValck et al. 1996); its N-terminal half interacts with the conserved C-terminal TRAF domain of TRAF1 and TRAF2, while its C-terminal domain mediates inhibition of NF-κB activation (Song et al. 1996). Putative A20-type zinc fingers (IPR002653) were also found (alone or in association with another zinc finger, AN1; IPR000058) not only in animals, but also in plants and among unicellular lineages from all four eukaryotic groups (Fig. 1).

Finally, several plant-specific cell death inhibitors are known. The Mlo family includes integral membrane proteins whose deficiency is thought to lower the threshold required to trigger the cascade of events that result in plant cell death (Devoto et al. 1999; Kim et al. 2002). While not reported in animals, fungi, excavates, and chroamalveolates, proteins with predicted Mlo domains (IPR004326) were, nevertheless, found in unicellular green algae (Fig. 1).

Nuclear Factors

The tumor suppressor, p53, is a transcription factor that plays the leading role in malignancy and in maintaining the genome's integrity and stability, by orchestrating various responses to DNA damage, including cell cycle arrest and PCD (Helton and Chen 2007). Although believed to be restricted to metazoans, two tumor suppressor p53-like sequences were found in the choanoflagellate, *M. brevicollis* (Nedelcu and Tan 2007). As one of the two *M. brevicollis* p53-like sequences contains a SAM domain (IPR001660)—which is associated with the p63/73 members of the p53 family (Yang et al. 2002)—these findings suggest an early duplication and diversification of this gene family, before the evolution of Metazoa (Nedelcu and Tan 2007). Furthermore, a diverged p53-like sequence has also been reported in *E. histolytica* (Mendoza et al. 2003), tracing back the origin of this important tumor suppressor family to Amoebozoa. Noteworthily, although no p53-like sequences have been identified outside unikonts, p53-like-mediated responses have been reported in green algae (Nedelcu 2006), and homologues of several p53-induced genes are found in many unicellular lineages (see below).

In multicellular organisms, PCD and cell cycle regulation reflect the two opposing options faced by a cell during development: death and proliferation (Aravind et al. 1999). In animals, in addition to p53, this decision is mediated by the transcription factors E2F-1/DP-1 and the retinoblastoma (Rb) protein, with the latter being antiapoptotic and sequestering the former. Until recently, the transcription factors that link cell cycle control to apoptosis were thought to be restricted to animals (Aravind et al. 1999). However, although missing in yeast, the E2F_TDP domain (IPR003316) was identified in many unicellular lineages (Fig. 1). Likewise, the two domains associated with retinoblastoma-like and retinoblastoma-associated proteins (IPR002719 and IPR002720)—also missing in yeast, were nevertheless found in unicellular groups (Fig. 1).

p53-Induced Genes

Although p53 homologues have only been reported in two unicellular lineages (discussed above), several genes that are known to be p53 targets in animals are found in many unicellular lineages. For instance, the human p53CSV (a member of the MDM35 family: mitochondrial distribution and morphology family 35) is a transcriptional target for p53 that mediates cell survival in response to genotoxic stress, by inhibiting the activation of procaspase-3 and -9 (Park and Nakamura 2005). In addition to the MDM35 protein reported in yeast [which is essential for maintenance of normal mitochondrial distribution and morphology (Dimmer et al. 2002)], proteins with putative MDM35 domains (IPR007918) were also predicted by Pfam in apicomplexans, amoebozoans, and choanoflagellates (Fig. 1). Similarly, the LPS-induced tumor necrosis factor α factor (LITAF) is known as a p53 target (p53-induced gene 7, or PIG7) in mammalian cells following treatment with lipopolysaccharide, and proteins with a LITAF domain (IPR006629) were predicted in several unicellular lineages (Fig. 1).

Sestrin (PA26 p-53 induced protein; IPR006730) was described as a novel p53 target gene, differentially induced by genotoxic stress (UV, γ-irradiation, and cytotoxic drugs) in a p53-dependent manner (Dimmer et al. 2002); interestingly, although apparently missing in yeast, proteins with predicted sestrin domains were found in several unicellular lineages from three eukaryotic supergroups (Fig. 1). Likewise, PIG8 (p53-induced gene 8; a.k.a. EI24, etoposide-induced 2.4) is induced by p53 in cells treated with the cytotoxic drug etoposide (Lehar et al. 1996). Notably, putative EI24 domains (IPR009890) were predicted in fungi (though missing in yeast), land plants, and many unicellular lineages (Fig. 1), and an EI24-like protein appears to be induced during PCD in green algae (Nedelcu 2006).

Executors

The essential executors in metazoan apoptosis are caspases—a class of cysteine proteases that catalyze peptide bond cleavage at aspartyl residues in their substrates. While homologues of caspases have not been found outside Metazoa, two related cysteine protease families have been described previously: paracaspases—present in metazoans and the amoebozoan *Dictyostelium*; and metacaspases—reported in plants, fungi, and some protozoans (Uren et al. 2000). Metacaspases share with caspases the presence of a conserved catalytic dyad composed of a cysteine and a histidine residue [although several exceptions have been reported (Mottram et al. 2003)]. However, in contrast to caspases, which are specific for acidic residues, metacaspases appear to prefer basic residues (Gonzalez et al. 2007; Vercammen et al. 2007; Deponte 2008b). Remarkably, in addition to fungi and protozoans, metacaspase-like sequences were found in many unicellular lineages (Fig. 1), and the presence of the conserved catalytic dyad argues for their performing similar proteolytic activities (Fig. 2a). Nevertheless, as metacaspases are known also to be involved in PCD-unrelated functions (e.g., Helms et al. 2006; Vercammen et al. 2007; Ambit et al. 2008), functional studies are needed to address the involvement of these sequences in PCD-like processes.

Two types of metacaspases, types I and II, have been reported in land plants—the main difference being the presence of an N-terminal extension in type I metacaspases and of an insertion between the p20- and the p10-like subunits in type II metacaspases (Uren et al. 2000). Notably, metacaspases displaying the insertion characteristic of type II metacaspases (Fig. 2a) were also found in green algae; furthermore, although no metacaspase sequences could be found in the available red algal genomes, a putative red algal-type II metacaspase (based on several *Porphyra* ESTs in GenBank) was also identified (Fig. 2a).

Fig. 2 a Partial alignment of representative type I and type II ▶ metacaspase predicted sequences from red algae (*Porphyra yezoensis*; Py), green algae (*Chlamydomonas reinhardtii,* Cr; *Volvox carteri,* Vc), vascular plants (*Arabidopsis thaliana*; At), excavates (*Trypanosoma cruzi*, Tc; *Leishmania braziliensis*, Lb), diatoms (*Thalassiosira pseudonana*, Tp; *Phaeodactylum tricornutum*, Pt), haptophytes (*Emiliania huxleyi*; Eh), pelagophytes (*Aureococcus anaphagefferens*; Aa), yeasts (*Schizosaccharomyces pombe*, Sp; *Saccharomyces cerevisiae*, Sc) showing the conservation of the cysteine-histidine dyad and the insertion characteristic of plant type II metacaspases (for more sequences and a full alignment see Supplementary Fig. 1). Numbers following species abbreviations are Uniprot IDs, if composed of both letters and numbers, or JGI IDs, if consisting of only numbers; the *Porphyra yezoensis* cluster is based on several GenBank overlapping ESTs (AU189679, AU189520, AU186857, AU188368, AU194902, AV433034). **b** Bayesian analysis (58 taxa; 122 sites; numbers represent posterior probability distributions of trees) of selected type I and II metacaspases from Plantae (red algae, in red; green algae, in dark green; plants, in light green), Chromalveolata (diatoms, in purple; haptophytes, in orange; pelagophytes, in pink), Excavata (in blue), and Unikonts (fungi, in brown). Maximum likelihood analyses predict similar relationships (bootstrap values for key nodes are indicated in *italics*, below the posterior probability values)

Phylogenetic analyses do support the inclusion of these algal sequences in the type II metacaspase group (Fig. 2), indicating that the diversification of the metacaspase family started early in the evolution of the Plantae lineage. These analyses also indicate that independent lineage-specific expansions involving type I metacaspases took place in several unicellular groups, including trypanosomatids, diatoms, and haptophytes (Fig. 2b).

Interestingly, type I and type II metacaspases have also been found in the closest unicellular relative of animals, the choanoflagellates (Fig. 1), but are thought to have been acquired from a green algal lineage early in the evolution of the choanoflagellates (Nedelcu et al. 2008). This scenario is supported by the presence of an LSD1-type Zn finger (discussed above) in the N-terminal of the *Monosiga* type I metacaspase; this specific association is only known in land plants (and possibly their close green algal ancestors), and although both LSD1 Zn fingers and type I metacaspases are present in many unicellular lineages (Fig. 1), the two domains are found together only in *Monosiga* (Nedelcu et al. 2008). A lateral gene transfer event is also consistent with the absence of type II metacaspases as well as LSD1 Zn fingers from the Unikont lineage (Fig. 1). If the unicellular ancestors of Metazoa possessed metacaspases, they must have been lost and/or replaced by caspases early in the evolution of Metazoa, as an early-diverged metazoan—the cnidarian, *Nematostella vectensis*—already contains a diversified family of caspases (as well as a putative paracaspase; see http://genome.jgi-psf.org/Nemve1/Nemve1. home.html).

The specific internucleosomal fragmentation of DNA (DNA-laddering) is considered to be a diagnostic feature of PCD. Several endonucleases involved in apoptotic DNA

fragmentation have been identified. Among them, mitochondrial endonucleases of the EndoG type (containing a DNA/RNA nonspecific endonuclease domain; IPR001604) have been shown to participate in this process in both mammals and yeast (Li et al. 2001; Buettner et al. 2007), and a proapoptotic nuclease activity for EndoG was recently reported in trypanosomatids. Interestingly, EndoG-like sequences appear to be absent in the land plant lineage, and alternative endonucleases are responsible for DNA fragmentation in plants (Balk et al. 2003). However, proteins with a predicted DNA/RNA nonspecific endonuclease domain were found in green algae, suggesting that EndoG-like sequences were present in the unicellular ancestors of Viridiplantae and were later lost or replaced in the lineage leading to land plants. Putative mitochondrial endonucleases of the EndoG-type were also found in many other unicellular lineages (Fig. 1). Noteworthily, a DNA-laddering effect during PCD-like processes was observed in some of these unicellular lineages [e.g., in *Chlamydomonas* (Nedelcu 2006; Moharikar et al. 2006)].

Finally, among the proteins involved in the cytoskeletal rearrangements required for phagocytosis of apoptotic cells, the mammalian ELMO1 and its *Caenorhabditis elegans* orthologue, CED-12, are required for cell migration and engulfment of dying cells (Gumienny et al. 2001). While missing in yeast, proteins with engulfment and cell motility (ELM) domains (IPR006816) were found in many unicellular lineages (Fig. 1).

The Early Evolution of the PCD Genetic Toolkit: Gene Duplications, Domain Shuffling, and Recruitment

Similarly to proteins that are implicated in other signaling and regulatory networks, many of the proteins involved in PCD in multicellular organisms are composed of multiple domains (signaling, protein-protein interaction, DNA binding) whose complex patterns of interactions define specific pathways (Aravind et al. 1999). Interestingly, several domains found in complex multidomain PCD-related proteins in metazoans are found as single-domain proteins in unicellular lineages (and in some early-diverged metazoans), indicating that they have been later recruited into complex multidomain proteins. For instance, the BIR domain, present in combination with the RING, NACHT, and CARD domains in several inhibitors of apoptosis in vertebrates, is found alone in unicellular and simple metazoans (as a single domain in ciliates, apicomplexans, and choanoflagellates; in tandem in yeast; and both singly and in tandem in nematodes). Likewise, the Znf_LSD1 domain—present in combination with the peptidase C14 domain in plant type I metacaspases—is found alone in unicellular lineages (as one or two copies—in *P. tetraurelia*,

Trypanosoma spp., and *Leishmania* spp. or as three copies in *C. reinhardtii*).

In some cases, the domains present in complex multidomain PCD-related proteins are present in unicellular lineages both as single-domain proteins and in multidomain proteins encompassing some or all the domains found in specific PCD proteins. For instance, human TRAF proteins are composed of three domains—RING, TRAF, and MATH—and all three domains are present in unicellular lineages, either as single-domain proteins or in multidomain proteins with a TRAF-like domain organization (e.g., TRAF in green algae and ciliates; RING-TRAF in choanoflagellates; RING-TRAF, RING-TRAF-TRAF, and RING-TRAF-TRAF-TRAF in ciliates; RING-TRAF-MATH and RING-TRAF-TRAF-MATH combinations in some *Leishmania* and *Dictyostelium* proteins).

In other cases, PCD domains found in complex PCD proteins in multicellular lineages are found in unique combinations in unicellular lineages, suggesting that, in addition to duplication and recruitment, domain shuffling was also important in the early evolution of the PCD machinery. This is the case for the NB-ARC domain, found in combination with TIR and LRR domains in plant disease resistance proteins, and with CARD and WD40 domains in apoptosis protease activating factors in animals, but in combination with TPR repeats in fungi and diatoms. Likewise, the NACHT domain, found together with Pyrin, CARD, and LRR domains in vertebrate apoptotic proteins, is found in combination with WD40 and/or EF_hand (Calcium-binding) domains in green algae, ciliates, *D. discoideum*, choanoflagellates, fungi, and the nematode *Caenorhabditis elegans*.

Notably, in many instances, in the same species a domain is found both alone and as part of a multidomain protein—suggesting that its recruitment into multidomain proteins followed a gene duplication event. In this context, an interesting case is provided by PDCD2. In metazoans and land plants, PDCD2 proteins comprise both a PDCD2_C (discussed earlier) and a Znf_MYND (IPR002893) domain. Interestingly, although proteins containing a PDCD2_C domain are present in all four major eukaryotic groups (Fig. 1), in most unicellular lineages for which information is available, the PDCD2_C domain is found only as single-domain proteins. The exceptions are the oomycetes, the amoebozoan *D. discoideum*, the red alga *Galdieria sulphuraria*, and the choanoflagellate *M. brevicollis*; noteworthily, in the former two lineages, the PDCD2_C domain is present alone as well as in association with a MYND-type zinc finger.

Nevertheless, several PCD-related proteins share the same domain organization across diverse eukaryotic groups, in both unicellular and multicellular lineages. For instance, cellular apoptosis susceptibility (CAS) proteins

exhibit the same domain organization (importin-β N-terminal–Cse1–CAS/CSE_C) in metazoans, land plants, fungi, Amoebozoa, Apicomplexa, and Excavata. Similarly, A20-like zinc fingers are associated with AN1 zinc fingers in proteins from Amoebozoa, Metazoa, Apicomplexa, green algae, and land plants, suggesting that this domain organization was achieved early in the evolution of eukaryotes.

An increase in the number and types of domains in PCD proteins has been noted previously in vertebrates, compared to invertebrates (Aravind et al. 2001). At this time, the available information on the domain organization of PCD-related sequences in unicellular and multicellular lineages allows us to infer some of the processes underlying the shaping of the eukaryotic PCD machinery during an earlier landmark in the evolution of life; that is, the transition from unicellular to multicellular life—such as during the evolution of multicellular animals from their unicellular ancestors. Figure 3 compares the domain organizations of representative PCD-related sequences in Metazoa and their unicellular relative, *Monosiga*. As during the evolution of vertebrates, a clear increase in complexity in terms of number and type of domains in PCD proteins can also be inferred during the unicellular-multicellular transition (Fig. 3). Consistent with the overall analysis of domain organization in unicellular and multicellular lineages discussed above, the evolution of complex PCD-related proteins in the metazoan lineage appears to have involved the domain recruitment of single-domain proteins into multidomain proteins, the duplication of single domains in conjunction with recruitment of new domains, and domain shuffling.
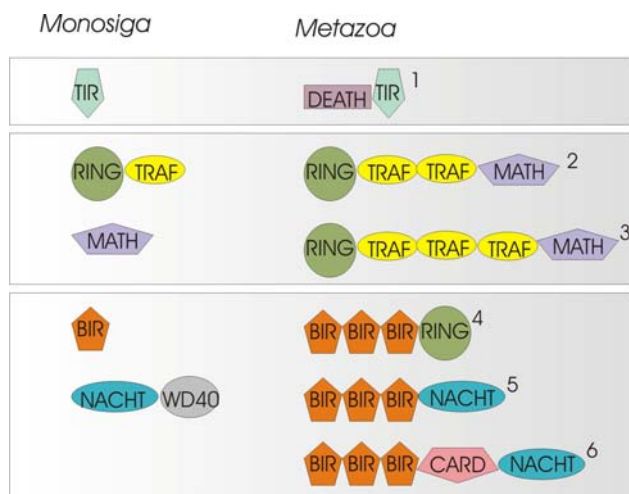
Fig. 3 Representative domain organizations of PCD-related sequences in Metazoa and their closest unicellular relative, *Monosiga* (see text for discussion). Superscripts: [1]human myeloid differentiation primary response protein myd88; [2]TNF receptor associated factor (TRAF) 4; [3]TRAF2, -3, -5, and -6; [4]inhibitor of apoptosis (IAP) 3; [5]neuronal inhibitor of apoptosis; [6]IAP1 and IAP2

Conclusion

The current comparative analysis reveals a very complex set of PCD-related domains and proteins among evolutionarily distant unicellular lineages. Although the exact relationships among the five main eukaryotic supergroups are not fully resolved (for a discussion see, e.g., Keeling et al. 2005; Embley and Martin 2006), and the monophyly of some of these groups is still under debate (e.g., Yoon et al. 2008), the phylogenetic distribution of the PCD-related sequences available to date (many of which are present in at least three of the four eukaryotic supergroups; Fig. 1) implies that a rather large repertoire of PCD-related sequences was likely present early in the evolution of eukaryotes. These findings (i) suggest that the potential (i.e., the genetic basis) for the complex eukaryotic PCD machinery present in extant multicellular lineages has been established in their unicellular ancestors and (ii) eliminate the need to invoke multiple later lateral gene transfers from bacterial sources to account for the presence of components of the PCD machinery thought to be restricted to "crown eukaryotes" (Koonin and Aravind 2002). On the other hand, the inability to detect many components of the metazoan apoptosis machinery (e.g., Bax, Bcl-2, SMAC/Diablo) in unicellular lineages indicates that the shaping of the PCD machinery in Metazoa also involved the evolution of new, metazoan-specific proteins.

Whether all the PCD-related sequences identified in unicellular lineages are involved in death pathways remains to be addressed. Indeed, some of the PCD-related sequences with counterparts in multicellular lineages are known to be involved in distinct or additional cellular processes (e.g., Brinkmann et al. 1995; Uren et al. 1999; Dimmer et al. 2002). In this context, this study provides a comparative framework to address the potential involvement of these sequences in PCD-like processes in single-celled organisms (which will argue for the ancestry of some of the components of the PCD machinery) or, alternatively, uncover their specific roles in unicellular lineages (which will shed light on the cellular pathways that have been co-opted during the evolution of PCD in more complex lineages). As genomic information accumulates and our understanding of the evolutionary relationships among the main eukaryotic lineages improves, it will be possible to reconstruct the ancestral eukaryotic PCD machinery and infer the specific evolutionary events and processes (e.g., gene duplications, losses and replacements, domain recruitment and shuffling, lateral gene transfer) that contributed to the shaping of the PCD machinery in various lineages.

References

Al-Olayan EM, Williams GT, Hurd H (2002) Apoptosis in the malaria protozoan, *Plasmodium berghei*: a possible mechanism for limiting intensity of infection in the mosquito. Int J Parasitol 32:1133–1143

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Ambit A, Fasel N, Coombs GH, Mottram JC (2008) An essential role for *Leishmania major* metacaspase in cell cycle progression. Cell Death Diff 15:113–122

Ameisen JC, Idziorek T, Billautmulot O, Loyens M, Tissier JP, Potentier A, Ouaissi A (1995) Apoptosis in a unicellular eukaryote (*Trypanosoma cruzi*)—implications for the evolutionary origin and role of programmed cell-death in the control of cell-proliferation, differentiation and survival. Cell Death Differ 2:285–300

Aravind L, Dixit VM, Koop B (1999) The domains of death: evolution of the apoptosis machinery. Trends Biochem Sci 24:47–53

Aravind L, Watanabe H, Lipman DJ, Koonin EV (2000) Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proc Natl Acad Sci USA 97:11319–11324

Aravind L, Dixit VM, Koonin EV (2001) Apoptotic molecular machinery: vastly increased complexity in vertebrates revealed by genome comparisons. Science 291:1279–1284

Arnoult D, Akarid K, Grodet A, Petit PX, Estaquier J, Ameisen JC (2002) On the evolution of programmed cell death: apoptosis of the unicellular eukaryote *Leishmania major* involves cysteine proteinase activation and mitochondrion permeabilization. Cell Death Differ 9:65–81

Balk J, Chew SK, Leaver CJ, McCabe PF (2003) The intermembrane space of plant mitochondria contains a DNase activity that may be involved in programmed cell death. Plant J 34:573–583

Bidle KD, Bender SJ (2008) Iron starvation and culture age activate metacaspases and programmed cell death in the marine diatom *Thalassiosira pseudonana*. Eukaryot cell 7:223–236

Bidle KD, Falkowski PG (2004) Cell death in planktonic, photosynthetic microorganisms. Nat Rev Microbiol 2:643–655

Bidle KD, Haramaty L, Barcelos E, Ramos J, Falkowski P (2007) Viral activation and recruitment of metacaspases in the unicellular coccolithophore, *Emiliania huxleyi*. Proc Natl Acad Sci USA 104:6049–6054

Brinkmann U (1998) CAS, the human homologue of the yeast chromosome-segregation gene CSE1, in proliferation, apoptosis, and cancer. Am J Hum Genet 62:509–513

Brinkmann U, Brinkmann E, Gallo M, Pastan I (1995) Cloning and characterization of a cellular apoptosis susceptibility gene, the human homologue to the yeast chromosome segregation gene CSE1. Proc Natl Acad Sci USA 92:10427–10431

Buettner S, Eisenberg T, Carmona-Gutierrez D, Ruli D, Knauer H, Ruckenstuhl C, Sigrist C, Wissing S, Kollroser M, Froehlich KU, Sigrist S, Madeo F (2007) Endonuclease G regulates budding yeast life and death. Mol Cell 25:233–246

Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol 52:297–354

Chae HJ, Ke N, Kim HR, Chen SR, Godzik A, Dickman M, Reed JC (2003) Evolutionarily conserved cytoprotection provided by Bax Inhibitor-1 homologs from animals, plants, and yeast. Gene 323:101–113

Cornillon S, Foa C, Davoust J, Buonavista N, Gross JD (1994) Programmed cell death in *Dictyostelium*. J Cell Sci 107:2691–2704

Deponte M (2008a) Programmed cell death in protists. Biochim Biophys Acta Mol Cell Res 1783:1396–1405

Deponte M (2008b) In search of Atropos' scissors: severing the life-thread of *Plasmodium*. In: Pérez Martin JM (ed) Programmed cell death in protozoa. Landes Bioscience, Austin, pp 91–96

DeValck D, Heyninck K, VanCriekinge W, Contreras R, Beyaert R, Fiers W (1996) A20, an inhibitor of cell death, self-associates by its zinc finger domain. FEBS Lett 384:61–64

Devoto A, Piffanelli P, Nilsson I, Wallin E, Panstruga R, von Heijne G, Schulze-Lefert P (1999) Topology, subcellular localization, and sequence diversity of the Mlo family in plants. J Biol Chem 274:34993–35004

Dimmer KS, Fritz S, Fuchs F, Messerschmitt M, Weinbach N, Neupert W, Westermann B (2002) Genetic basis of mitochondrial function and morphology in *Saccharomyces cerevisiae*. Mol Biol Cell 13:847–853

Doong H, Vrailas A, Kohn EC (2002) What's in the 'BAG'? A functional domain analysis of the BAG-family proteins. Cancer Lett 188:25–32

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:1–19

Ekert PG, Vaux DL (2005) The mitochondrial death squad: hardened killers or innocent bystanders? Curr Opin Cell Biol 17:626–630

Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. Nature 440:623–630

Fahrenkrog B, Sauder U, Aebi U (2004) The S-cerevisiae HtrA-like protein Nma111p is a nuclear serine protease that mediates yeast apoptosis. J Cell Sci 117:115–126

Gonzalez IJ, Desponds C, Schaff C, Mottram JC, Fasel N (2007) *Leishmania major* metacaspase can replace yeast metacaspase in programmed cell death and has arginine-specific cysteine peptidase activity. Int J Parasitol 37:161–172

Gumienny TL, Brugnera E, Tosello-Trampont AC, Kinchen JM, Haney LB, Nishiwaki K, Walk SF, Nemergut ME, Macara IG, Francis R, Schedl T, Qin Y, Van Aelst L, Hengartner MO, Ravichandran KS (2001) CED-12/ELMO, a novel member of the crkII/dock180/rac pathway, is required for phagocytosis and cell migration. Cell 107:27–41

Helms MJ, Ambit A, Appleton P, Tetley L, Coombs GH, Mottram JC (2006) Bloodstream form *Trypanosoma brucei* depend upon multiple metacaspases associated with RAB11-positive endosomes. J Cell Sci 119:1105–1117

Helton ES, Chen XB (2007) p53 modulation of the DNA damage response. J Cell Biochem 100:883–896

Huckelhoven R (2004) BAX Inhibitor-1, an ancient cell death suppressor in animals and plants with prokaryotic relatives. Apoptosis 9:299–307

Jacobson MD, Weil M, Raff MC (1997) Programmed cell death in animal development. Cell 88:347–354

Jung YS, Kim KS, Kim KD, Lim JS, Kim JW, Kim E (2001) Apoptosis-linked gene 2 binds to the death domain of Fas and dissociates from Fas during Fas-mediated apoptosis in Jurkat cells. Biochem Biophys Res Commun 288:420–426

Keeling PJ, Burger G, Durnford DG, Lang BF, Lee RW, Pearlman RE, Roger AJ, Gray MW (2005) The tree of eukaryotes. Trends Ecol Evol 20:670–676

Kim MC, Lee SH, Kim JK, Chun HJ, Choi MS, Chung WS, Moon BC, Kang CH, Park CY, Yoo JH, Kang YH, Koo SC, Koo YD, Jung JC, Kim ST, Schulze-Lefert P, Lee SY, Cho MJ (2002) Mlo, a modulator of plant defense and cell death, is a novel calmodulin-binding protein—isolation and characterization of a rice Mlo homologue. J Biol Chem 277:19304–19314

King N (2004) The unicellular ancestry of animal development. Dev Cell 7:313–325

Koonin EV, Aravind L (2000) The NACHT family—a new group of predicted NTPases implicated in apoptosis and MHC transcription activation. Trends Biochem Sci 25:223–224

Koonin EV, Aravind L (2002) Origin and evolution of eukaryotic apoptosis: the bacterial connection. Cell Death Differ 9:394–404

Krebs J, Saremaslani P, Caduff R (2002) ALG-2: a Ca^{2+}-binding modulator protein involved in cell proliferation and in cell death. Biochim Biophys Acta 1600:68–73

Lam E (2004) Controlled cell death, plant survival and development. Nat Rev Mol Cell Biol 5:305–315

Lehar SM, Nacht M, Jacks T, Vater CA, Chittenden T, Guild BC (1996) Identification and cloning of EI24, a gene induced by p53 in etoposide-treated cells. Oncogene 12:1181–1187

Li LY, Luo L, Wang XD (2001) Endonuclease G is an apoptotic DNase when released from mitochondria. Nature 412:95–99

Liu HT, Wang YG, Zhang YM, Song QS, Di CH, Chen GH, Tang J, Ma DL (1999) TFAR19, a novel apoptosis-related gene cloned from human leukemia cell line TF-1, could enhance apoptosis of some tumor cells induced by growth factor withdrawal. Biochem Biophys Res Commun 254:203–210

Lufei CC, Ma J, Huang GC, Zhang T, Novotny-Diermayr V, Ong CT, Cao XM (2003) GRIM-19, a death-regulatory gene product, suppresses Stat3 activity via functional interaction. EMBO J 22:1325–1335

Madeo F, Frohlich E, Ligr M, Grey M, Sigrist SJ, Wolf DH, Frohlich KU (1999) Oxygen stress: a regulator of apoptosis in yeast. J Cell Biol 145:757–767

Madeo F, Herker E, Maldener C, Wissing S, Lachelt S, Herian M, Fehr M, Lauber K, Sigrist SJ, Wesselborg S, Frohlich KU (2002) A caspase-related protease regulates apoptosis in yeast. Mol Cell 9:911–917

Mendoza L, Orozco E, Rodriguez MA, Garcia-Rivera G, Sanchez T, Garcia E, Gariglio P (2003) Ehp53, an Entamoeba histolytica protein, ancestor of the mammalian tumour suppressor p53. Microbiology 149:885–893

Modjtahedi N, Giordanetto F, Madeo F, Kroemer G (2006) Apoptosis-inducing factor: vital and lethal. Trends Cell Biol 16:264–272

Moharikar S, D'Souza JS, Kulkarni AB, Rao BJ (2006) Apoptotic-like cell death pathway is induced in unicellular chlorophyte Chlamydomonas reinhardtii (Chlorophyceae) cells following UV irradiation: detection and functional analyses. J Phycol 42:423–433

Moharikar S, D'Souza JS, Rao BJ (2007) A homologue of the defender against the apoptotic death gene (dad1) in UV-exposed Chlamydomonas cells is downregulated with the onset of programmed cell death. J Biosci 32:261–270

Mottram JC, Helms MJ, Coombs GH, Sajid M (2003) Clan CD cysteine peptidases of parasitic protozoa. Trends Parasitol 19:182–187

Nakashima T, Sekiguchi T, Kuraoka A, Fukushima K, Shibata Y, Komiyama S, Nishimoto T (1993) Molecular-cloning of a human Cdna-encoding a novel protein, Dad1, whose defect causes apoptotic cell-death in hamster Bhk-21-cells. Mol Cell Biol 13:6367–6374

Nedelcu AM (2006) Evidence for p53-like-mediated stress responses in green algae. FEBS Lett 580:3013–3017

Nedelcu A, Tan C (2007) Early diversification and complex evolutionary history of the p53 tumor suppressor gene family. Dev Genes Evol 216:801–806

Nedelcu AM, Miles IH, Fagir AM, Karol K (2008) Adaptive eukaryote-to-eukaryote lateral gene transfer: stress-related genes of algal origin in the closest unicellular relatives of animals. J Evol Biol 21:1852

Park WR, Nakamura Y (2005) p53CSV, a novel p53-inducible gene involved in the p53-dependent cell-survival pathway. Cancer Res 65:1197–1206

Sadoul M (2006) Do Alix and ALG-2 really control endosomes for better or for worse? Biol Cell 98:69–77

Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. Proc Natl Acad Sci USA 95:5857–5864

Segovia M, Haramaty L, Berges JA, Falkowski PG (2003) Cell death in the unicellular chlorophyte Dunaliella tertiolecta. A hypothesis on the evolution of apoptosis in higher plants and metazoans. Plant Physiol 132:99–105

Song HY, Rothe M, Goeddel DV (1996) The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappa B activation. Proc Natl Acad Sci USA 93:6721–6725

Stechmann A, Cavalier-Smith T (2003) The root of the eukaryote tree pinpointed. Curr Biol 13:R665–R666

Tewari M, Yu M, Ross B, Dean C, Giordano A, Rubin R (1997) AAC-11, a novel cDNA that inhibits apoptosis after growth factor withdrawal. Cancer Res 57:4063–4069

Uren AG, Beilharz T, O'Connell MJ, Bugg SJ, van Driel R, Vaux DL, Lithgow T (1999) Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell division. Proc Natl Acad Sci USA 96:10170–10175

Uren AG, O'Rourke K, Aravind L, Pisabarro MT, Seshagiri S, Koonin EV, Dixit VM (2000) Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. Mol Cell 6:961–967

van der Biezen EA, Jones JDG (1998) The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol 8:R226–R227

Vardi A, Berman-Frank I, Rozenberg T, Hadas O, Kaplan A, Levine A (1999) Programmed cell death of the dinoflagellate Peridinium gatunense is mediated by CO_2 limitation and oxidative stress. Curr Biol 9:1061–1064

Vaux DL, Hacker G (1995) Cloning of mouse Rp-8 Cdna and its expression during apoptosis of lymphoid and myeloid cells. DNA Cell Biol 14:189–193

Vercammen D, Declercq W, Vandenabeele P, Van Breusegem F (2007) Are metcaspases caspases? J Cell Biol 179:375–380

Walter D, Wissing S, Madeo F, Fahrenkrog B (2006) The inhibitor-of-apoptosis protein Bir1p protects against apoptosis in S. cerevisiae and is a substrate for the yeast homologue of Omi/HtrA2. J Cell Sci 119:1843–1851

Wissing S, Ludovico P, Herker E, Buttner S, Engelhardt SM, Decker T, Link A, Proksch A, Rodrigues F, Corte-Real M, Frohlich KU, Manns J, Cande C, Sigrist SJ, Kroemer G, Madeo F (2004) An AIF orthologue regulates apoptosis in yeast. J Cell Biol 166:969–974

Yang A, Kaghad M, Caput D, McKeon F (2002) On the shoulders of giants: p63, p73, and the rise of p53. Trends Genet 18:90–95

Yoon HS, Grant J, Tekle YI, Wu M, Chaon BC, Cole JC, Logsdon JM Jr, Patterson DJ, Bhattacharya D, Katz LA (2008) Broadly sampled multigene trees of eukaryotes. BMC Evol Biol 8:14

Zuppini A, Andreoli C, Baldan B (2007) Heat stress: an inducer of programmed cell death in Chlorella saccharophila. Plant Cell Physiol 48:1000–1009