

A New Evolutionary Paradigm for the Parkinson Disease Gene *DJ-1*

J. Ignasi Lucas and Ignacio Marín

Departamento de Genética, Universidad de Valencia, Burjassot, Spain

The *DJ-1* gene is extensively studied because of its involvement in familial Parkinson disease. *DJ-1* belongs to a complex superfamily of genes that includes both prokaryotic and eukaryotic representatives. We determine that many prokaryotic groups, such as proteobacteria, cyanobacteria, spirochaetes, firmicutes, or fusobacteria, have genes, often incorrectly called “Thij,” that are very close relatives of *DJ-1*, to the point that they cannot be clearly separated from the eukaryotic *DJ-1* genes by phylogenetic analyses of their sequences. In addition, and contrary to a previous study that suggested that *DJ-1* genes were animal specific, we show that *DJ-1* genes are found in at least 5 of the 6 main eukaryotic groups: opisthokonta (both animals and fungi), plantae, chromalveolata, excavata, and amoebozoa. Our results thus provide strong evidence for *DJ-1* genes originating before the origin of eukaryotes. Interestingly, we found that some fungal species, among them the model yeast *Schizosaccharomyces pombe*, have *DJ-1*-like genes, most likely orthologous to the animal genes. This finding opens new ways for the analysis of the functions of this group of genes.

Introduction

The gene *DJ-1* was originally described as an oncogene (Nagakubo et al. 1997) and, in parallel, found to encode a protein involved in male fertility in rats and other mammals (Wagenfeld et al. 1998; Welch et al. 1998) and able to work as a regulatory subunit of an RNA-binding complex (Hod et al. 1999). However, interest in this gene largely increased when it was found to be involved in familial Parkinson disease (Bonifati et al. 2003). Several recessive mutations of *DJ-1*, both deletions and missense mutations, have been shown to cause early-onset Parkinson disease (reviewed in Bonifati et al. 2004). Homozygosity for the disease-associated missense mutations is thought to functionally impair or inactivate DJ-1. In particular, it has been shown that DJ-1 proteins form dimers, and the crystal structures of these dimers suggested that the first missense mutation described, L166P (Bonifati et al. 2003), would interfere with the dimerization process (Honbou et al. 2003; Huai et al. 2003; Lee et al. 2003; Tao and Tong 2003; Wilson et al. 2003). Several groups have since then obtained experimental evidence confirming that the L166P mutation indeed interferes with dimer formation, favoring rapid degradation of the monomers by the proteasome (Macedo et al. 2003; Miller et al. 2003; Moore et al. 2003; Olzmann et al. 2004; Takahashi-Niki et al. 2004). There is evidence that another missense mutation associated to familial Parkinson disease, M26I (Abou-Sleiman et al. 2003), may also cause rapid proteasomal degradation (Takahashi-Niki et al. 2004; Blackinton et al. 2005; Xu et al. 2005; for negative results, however, see Moore et al. 2003; Baulac et al. 2004). Increased protein instability may also explain the effect of a third missense mutation, E64D (Hering et al. 2004). Changes in cellular localization may also contribute to the effects of these mutations (Xu et al. 2005; but see Zhang et al. 2005). Recently, individuals with a combination of 2 *DJ-1* mutations, a short duplication in its promoter plus another missense mutation, E163K, have been shown to suffer a complex syndrome with parkinsonism, dementia, and amyotrophic lateral sclerosis symptoms (Annesi et al. 2005). DJ-1 protein is also found

in inclusion bodies together with Tau in patients of different tauopathies including Alzheimer disease (Neumann et al. 2004; Rizzu et al. 2004). These results suggest that *DJ-1* may be involved in multiple neurodegenerative diseases.

With functions in the male germ line, the brain, and probably many other tissues (Nagakubo et al. 1997), the biological roles of *DJ-1* are expected to be diverse (see reviews by Bonifati et al. 2004; Abou-Sleiman et al. 2006). In the context of Parkinson disease, *DJ-1* has been implicated in response to oxidative stress. Loss of function of *DJ-1* leads to increased stress levels in cellular models (e.g., Yokota et al. 2003; Canet-Aviles et al. 2004; Martinat et al. 2004; Xu et al. 2005). *DJ-1* null mutants in mice and *Drosophila* also show exacerbated sensitivity to agents that increase oxidative stress (Goldberg et al. 2005; Kim, Smith, et al. 2005; Menzies et al. 2005; Meulener et al. 2005; Park et al. 2005), and the same is true for *Caenorhabditis* worms in which *DJ-1* is downregulated using RNA interference (Ved et al. 2005). Diverse functions in stress response have been already described for DJ-1. First, it has been reported to be a redox-dependent chaperone (Shendelman et al. 2004). Its chaperone action may contribute to inhibit the aggregation of α -synuclein (Shendelman et al. 2004; Zhou et al. 2006). This functional link may be very significant because α -synuclein and ubiquitin are the most abundant proteins in Lewy bodies, the characteristic cytoplasmic inclusions found in Parkinson disease. Moreover, some mutations in the α -synuclein gene such as duplications, triplications, and missense mutations that may contribute to increased aggregation are known to cause familial Parkinson disease (reviewed in Abou-Sleiman et al. 2006). DJ-1 may act also as a redox-sensitive negative regulator of apoptosis. Apoptosis inhibition is probably mediated by diverse independent actions. Thus, DJ-1 has been characterized as contributing to the activation of the PI3K/Akt survival signaling pathway (Kim, Peters, et al. 2005; Yang et al. 2005) and downregulates the DAXX-ASK1 proapoptotic pathway (Junn et al. 2005). Additional protective actions have been also reported (Xu et al. 2005; Zhou and Freed 2005).

Animal models have so far failed to recapitulate the effects of *DJ-1* mutations in humans. Null *DJ-1* mutant mice show anomalies in dopaminergic function and, as already indicated, increased sensitivity to oxidative stress but not lack of dopaminergic neurons (Goldberg et al. 2005; Kim, Smith, et al. 2005). Similarly, *Drosophila*

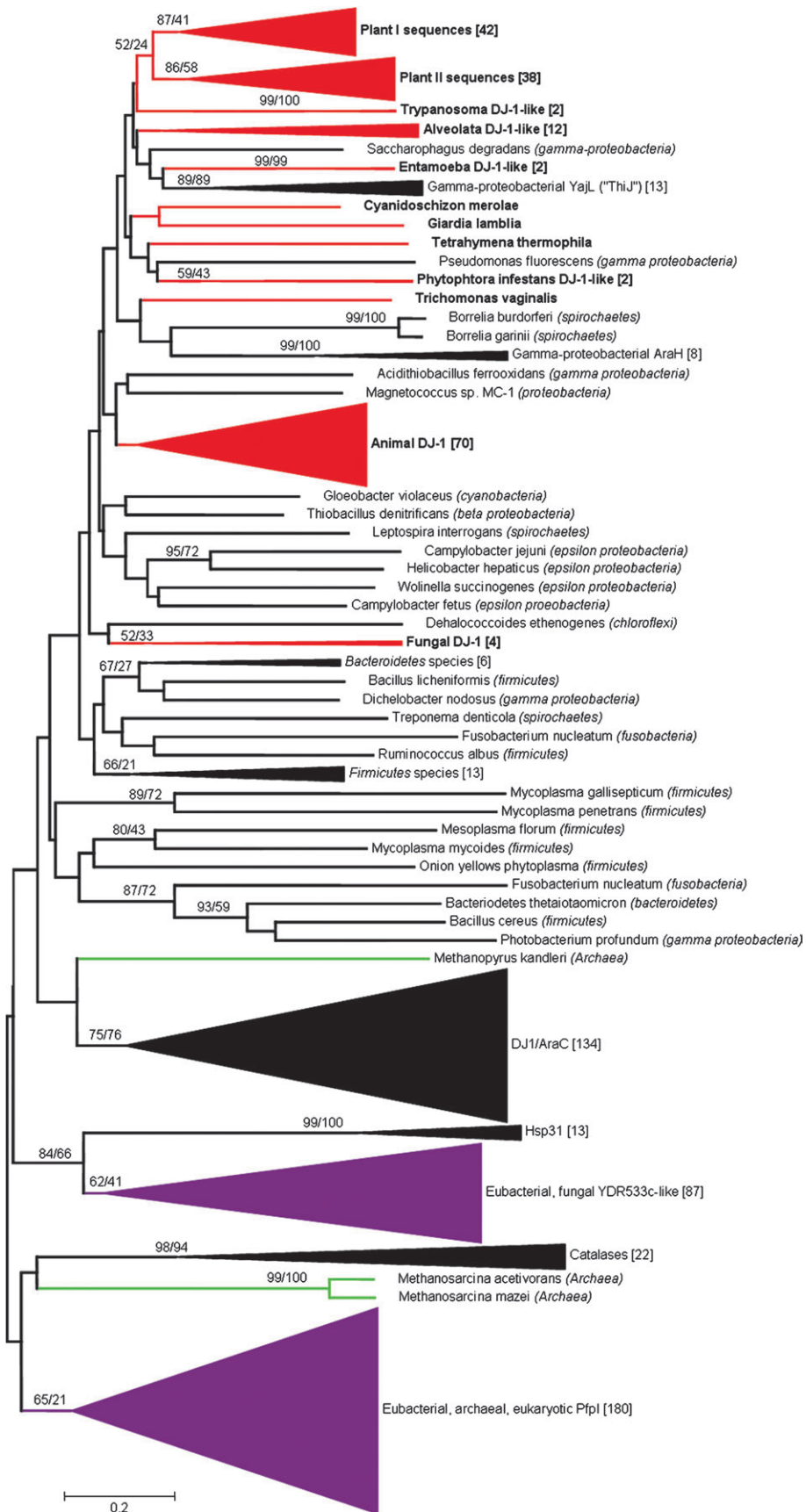
Key words: comparative genomics, Parkinson disease, *DJ-1* domain.

E-mail: ignacio.marin@uv.es.

Mol. Biol. Evol. 24(2):551–561. 2007

doi:10.1093/molbev/msl186

Advance Access publication November 30, 2006



has 2 recently duplicated *DJ-1* genes, and lack of both of them has no apparent effect on dopaminergic neurons (Meulener et al. 2005). The search for other possible models in which to analyze *DJ-1* genes function depends on the detection of orthologs of those genes. This is quite a complex task in this case because genes with sequences obviously related to that of *DJ-1* exist in many organisms, and to determine their precise relationships of orthology and paralogy, that is, whether they are true orthologs or just related paralogous genes, requires careful consideration of their similarities in a phylogenetic framework. Interestingly, Welch et al. (1998) determined long ago that some of the genes most similar to *DJ-1* are prokaryotic. In particular, they detected an *Escherichia coli* gene called “ThiJ,” supposedly involved in thiamine synthesis (Backstrom 1996), which was very similar to mammalian *DJ-1* genes. The puzzling idea that mammalian *DJ-1* genes would derive from prokaryotic thiamine synthesis genes has pervaded the literature until very recently, in spite of the fact that Mueller et al. (1998) already cited that the putative relationship of the *E. coli* gene similar to *DJ-1* to thiamine synthesis was indeed due to an experimental artifact, a fact detailed by many gene and protein databases (e.g., UniProt, EcoGene, EchoBase, EcoCyc, etc.). Even so, only recently a study has been published in which this fact has been finally fully taken into account (Wilson et al. 2005). The *DJ-1*-related gene of *E. coli* formerly called ThiJ is currently named *YajL*.

As part of an ongoing project to trace the evolutionary history of Parkinson disease genes in order to provide novel hints about their cellular functions (Marín and Ferrús 2002; Marín et al. 2004; Lucas et al. 2006; Marín 2006), we describe in this study our novel analyses of this significant group of genes. Our goals were to provide Parkinson disease researchers with a correct conceptual framework about the origin and evolutionary history of *DJ-1* genes and to suggest novel organisms in which to study their functions. As we will detail, our conclusions are in substantial disagreement with a superficial analysis of this family of genes published before.

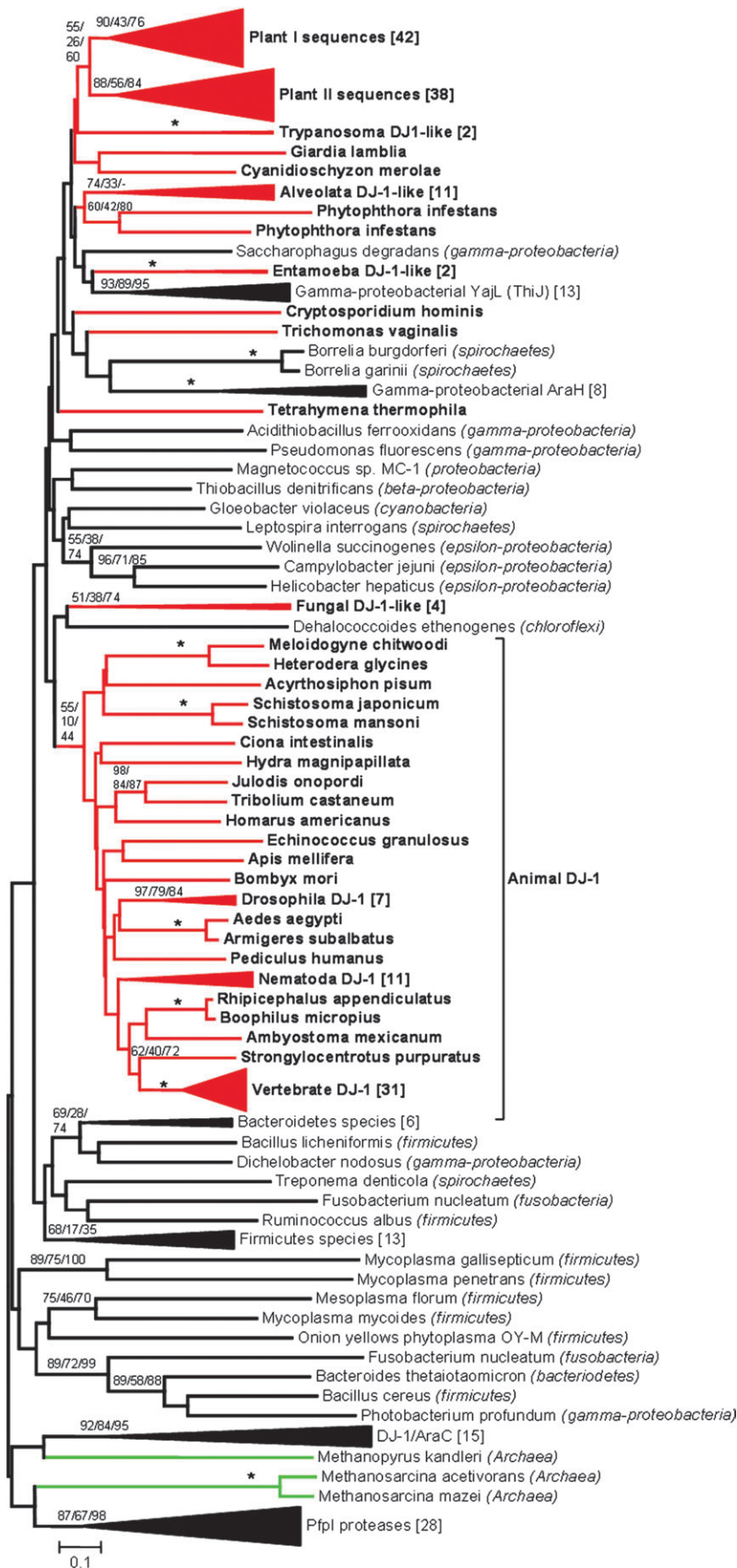
Methods

All genes of the DJ-1 superfamily are characterized by having a common protein domain that in the structural Pfam database has been named “DJ-1_PfpI” and that we will call here, in abbreviated form, DJ-1 domain. We found that this highly conserved domain extends from amino acids 5–172 in the 189 amino acids-long human DJ-1 protein. To obtain a representative sample of genes of the DJ-1 superfamily, we first performed extensive BLASTP and TBlastN searches against the National Center for Biotechnology Information databases (<http://www.ncbi.nlm.nih.gov/>) and those compiled in the GOLD database ([\[www.genomesonline.org/\]\(http://www.genomesonline.org/\)\) using the DJ-1 domains of several *DJ-1* and *YajL* sequences as queries. Additional specific searches were performed to obtain a representative sample of genes distantly related to *DJ-1*. These genes included those belonging to the PfpI class, genes that contain both a DJ-1 and a DNA-binding AraC domain \(that we will call DJ-1/AraC from now on\), genes encoding *DJ-1*-related catalases, and genes encoding Hsp31-related and YDR533c-related proteins \(see descriptions in Gallegos et al. 1997; Du et al. 2000; Horvath and Grishin 2001; Quigley et al. 2003; Graille et al. 2004; Wilson et al. 2004\). Results obtained from all those searches were merged and the protein sequences aligned using ClustalX version 1.83 \(Thompson et al. 1997\). We then generated a preliminary phylogenetic tree using the Neighbor-Joining \(NJ, Saitou and Nei 1987\) routine available in ClustalX 1.83. That tree was used to detect duplicates and partial sequences, which were eliminated. After these corrections, our final database of *DJ-1*-like genes contained 686 sequences.](http://</p>
</div>
<div data-bbox=)

We then used this database to generate a final multiple protein alignment using again ClustalX 1.83, which was manually corrected with GeneDoc version 2.6 (Nicholas et al. 1997). Phylogenetic trees were obtained from this alignment both by the NJ and the maximum-parsimony (MP) methods, using the routines available in MEGA 3.1 (Kumar et al. 2004) and PAUP*, beta 10 version (Swofford 2003), respectively. For NJ, sites with gaps were included and Kimura’s (also known as Poisson) correction was used, whereas for MP, the parameters were as follows: 1) all sites included, 2) randomly generated trees used as seeds, 3) maximum number of tied trees saved equal to 20, and 4) heuristic search using the subtree pruning–regrafting algorithm. Support for the topologies obtained with those 2 methods was determined using the bootstrap routines also available in MEGA 3.1 and PAUP*. One thousand replicates were performed for both NJ and MP bootstrap analyses. For the more limited analyses that include only part of the DJ-1 superfamily sequences (see Results and figs. 2 and 3 below), we used a third method of phylogenetic inference, namely, maximum likelihood (ML), as implemented in the PHYML program (Guindon and Gascuel 2003). We took the BIONJ tree as starting point for the iterative ML searches, and calculations were performed using the Blosum62 matrix of amino acidic similarity. This method is computer intensive, so only 200 bootstrap replicates were performed to check for the reliability of the topologies obtained. The figures that show the phylogenetic trees were generated using the tree editor of MEGA 3.1. Three-dimensional structures were predicted with Swiss-Model (Peitsch 1996; <http://swissmodel.expasy.org/>) using different crystal models of human *DJ-1* protein as templates (Protein Data Bank codes 1PE0, 1UCF, 1Q2U, and 1PS4). Swiss-Pdb viewer version 3.7 (Guex

←

FIG. 1.—Phylogenetic tree for sequences of the DJ-1 superfamily. Results for NJ and MP analyses were similar enough as to be shown in a single tree. We show bootstrap values for all branches supported by both methods of phylogenetic reconstruction and in which NJ results were higher than 50%. Values are ordered as NJ/MP. Numbers of species for the condensed branches are shown in brackets. Notice the overlapping positions of prokaryotic (black) and eukaryotic (red) branches around animal *DJ-1* genes. The branches that contain YDR533c-like and PfpI proteins are shown in purple because they contain both prokaryotic and eukaryotic sequences. Three archaeal sequences of unclear phylogenetic position are shown in green.



and Peitsch 1997) was used to generate the 3-dimensional image shown below.

Results

Genes Closely Related to *DJ-1* Are Found in Both Prokaryotes and Eukaryotes

Figure 1 shows the results for the phylogenetic analyses based on our complete protein sequence database. NJ and MP results were congruent enough as to be shown in a single tree. Several likely monophyletic groups that contain well-known prokaryotic proteins, such as DJ-1/AraC, DJ-1 domain-containing catalases, and Hsp31-related proteins, were detected. In addition, 3 groups containing both prokaryotic and eukaryotic sequences were also found. One of them contains both eubacterial and fungal sequences related to the protein encoded by the *Saccharomyces cerevisiae* YDR533c gene. The second one mostly contains eubacterial PfpI-related proteins but also includes a few archaeal and eukaryotic proteins. Finally, the third one, that is shown in full in figure 1, is a poorly supported ensemble that contains the animal *DJ-1* genes together with a mixture of prokaryotic and eukaryotic genes. We found that CHG (or, occasionally, CHA) putative catalytic triads, which suggest a hydrolytic function, were restricted to 3 of these groups—Hsp31-related proteins, YDR533c-related proteins, and PfpI proteins—in agreement with the results presented by other authors (Du et al. 2000; Quigley et al. 2003; Graille et al. 2004; Wilson et al. 2004).

We detected that the sequences of the DJ-1 domains of several of those groups, in spite of their obvious similarity, cannot be fully reliably aligned along their whole sequences. This is due either to the fact that sequences contain long stretches of group-specific amino acids (e.g., the protein encoded by the genes most related to *S. cerevisiae* YDR533C contains 2 extra regions in the middle of the DJ-1 domain) or to low similarity in particular regions (e.g., the C-terminal part of the DJ-1 domain of the catalases cannot be unambiguously aligned to *DJ-1* genes). Thus, to avoid serious distortions in our phylogenetic trees that might obscure the true evolutionary relationships with *DJ-1* genes, we decided to select only those sequences similar enough as to be aligned without ambiguity with animal *DJ-1* genes along the whole DJ-1 domain. This eliminated from the analysis the catalase, Hsp31, and YDR533c-like monophyletic groups detected before. We thus aligned all DJ-1-related sequences plus representatives of the DJ-1/AraC and PfpI classes. This final selection included 288 sequences. However, after repeating our analyses, and even including a third method of phylogenetic reconstruction, ML, we determined that leaving only the sequences with the most similar, fully alignable DJ-1 domains did not significantly improve the resolution of the topology (fig. 2). All 3 methods generated very similar trees, but the bootstrap support for the inner branches of those trees was al-

ways very low. We conclude that the information contained in the alignment of *DJ-1*-related genes is insufficient to completely determine the true topology of the tree. Therefore, the relationships among these *DJ-1*-related sequences are uncertain, and, considering our large sample size, probably they will remain so no matter the number or variety of sequences of *DJ-1*-related genes analyzed.

If we center our attention in the group in which animal *DJ-1* genes are included and that therefore must contain the most likely candidates to be considered true orthologs of the human *DJ-1*, some significant features are detected. First, regarding the prokaryotic sequences, we can see that they belong to taxa that cover essentially all the eubacterial phylogenetic range, such as proteobacteria, spirochaetes, chloroflexi, firmicutes, fusobacteria, cyanobacteria, etc. (details in figs. 1 and 2, black branches). This result strongly suggests an ancient origin for this type of genes in eubacteria. Second, there are many eukaryotic groups in which we found species that contain sequences closely related to *DJ-1* (figs. 1 and 2; shown in red). We concluded that *DJ-1*-like genes are present in 5 of the 6 main eukaryotic groups (see Simpson and Roger 2004): Opisthokonta, including both animals and fungi, Plantae, Chromoalveolata, Excavata, and Amoebozoa. Within these eukaryotic groups, the range of species is normally very wide. For example, animals ranging from cnidarians to vertebrates, including all model organisms with completely sequenced genomes, contain *DJ-1* genes. Similarly, in plants, both angiosperms and gymnosperms are found to contain *DJ-1*-related genes. Only in fungi, we found that independent losses of *DJ-1* genes must have occurred. Although they are present in both ascomycetes (*Schizosaccharomyces*, *Alternaria*) and basidiomycetes (*Ustilago*, *Coprinopsis*), most fully or almost fully sequenced fungi (e.g., *Saccharomyces*, *Candida*, *Aspergillus*, etc.) lack *DJ-1*-related genes. As it can be noticed from examining figures 1 and 2, we cannot exclude that *DJ-1*-like genes may also be present in some archaea. A few archaeal sequences with ambiguous positions in the trees were found that cannot be clearly included in any known group (green branches in figs. 1 and 2).

With these results, we may ask which would be the simplest hypothesis to explain such a close relationship among prokaryotic and eukaryotic sequences as to be intermingled in our trees. A first option would be similarity due to common descent. Alternatively, we can envisage similarity due to horizontal transmission. In our opinion, our results, taken as a whole, clearly favor the hypothesis that prokaryotic and eukaryotic *DJ-1*-like genes are orthologous and that the high similarity of some eukaryotic and prokaryotic sequences is due to parallel or convergent evolution. The reason why we favor this hypothesis is the broad phylogenetic range of these genes in both prokaryotes and eukaryotes. This broad range means that, although we cannot discard horizontal transfer events, they are unnecessary

←

FIG. 2.—Dendrogram obtained for *DJ-1*-related sequences once the most diverged groups are excluded. Results follow the same conventions that in figure 1, but here, besides the NJ and MP bootstrap values, a third number indicates the ML bootstrap results. Asterisks indicate branches for which bootstrap values higher than 95% were obtained in the 3 methods of phylogenetic reconstruction. Dashes refer to lack of support for one of the methods.

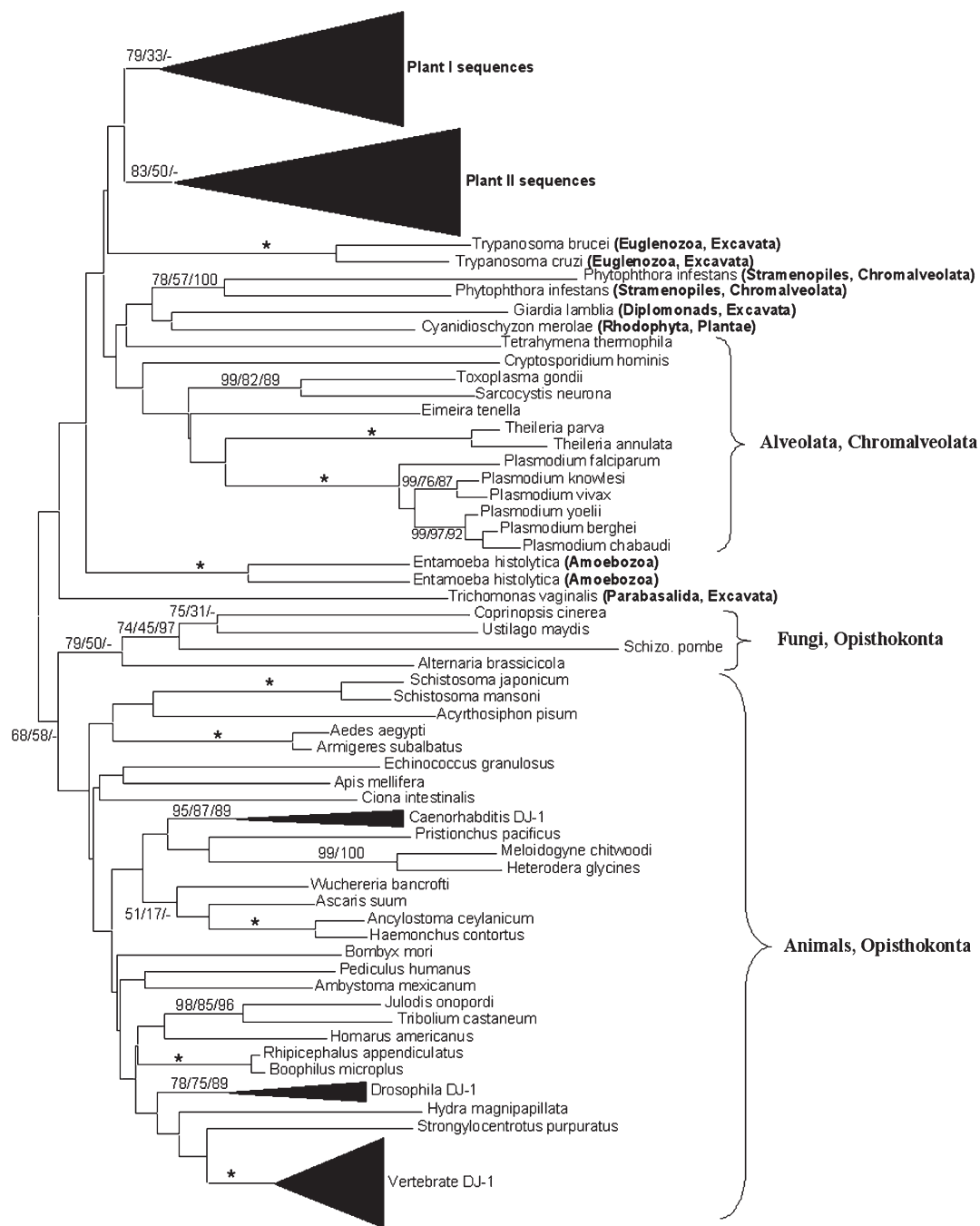


FIG. 3.—Phylogenetic tree for eukaryotic *DJ-1*-like genes. As in the previous figures, numbers refer to bootstrap values (NJ/MP/ML).

to explain our results. We therefore suggest that eukaryotic *DJ-1*-like genes may have derived from prokaryotic genes and were already present when eukaryotes arose. If this hypothesis is correct, the group including all the eukaryotic *DJ-1*-like genes shown in figure 1 should be monophyletic, and therefore, they should have related functions.

Novel Model Organisms in Which to Study *DJ-1* Function

One of the main goals of our study was to determine whether novel eukaryotic model species could be found in which significant functional analyses of *DJ-1* genes may be

performed. When we detected that *DJ-1*-like genes were present in some fungi, and especially in *Schizosaccharomyces pombe*, a species that is broadly used in genetic, cell biology, and biochemical studies, we decided to perform additional analyses to confirm whether these genes are true *DJ-1* orthologs. We thus selected from our database all eukaryotic sequences and performed new phylogenetic analyses to avoid any potential distortion due to the presence of homoplastic prokaryotic sequences. Results are shown in figure 3. All 3 methods generated very similar topologies. However, the ML analysis showed a few anomalies: 2 plant II sequences appeared as abnormally long branches inside

the plant I group and one fungal sequence, from *Alternaria brassicicola*, again appeared as a very long branch, this time inside the plant II group (data not shown). These obvious misplacements, which are in contradiction with all the previous results (including the ML results obtained with the larger data set that we showed before; fig. 2), must have been caused by this particular analysis being trapped in a local likelihood maximum (Chor et al. 2000). In any case, either 3 (ML) or 4 (NJ, MP) fungal sequences, among them the *S. pombe* gene (called *SPAC22E12.03c*), are found as the closest relatives of animal *DJ-1* genes, in good agreement with the classification of animals and fungi as sister groups within the Opisthokonta. The fungal gene most similar to animal *DJ-1* genes was found to be the one in the basidiomycete *Coprinopsis cinerea* (identity with the human gene along the DJ-1 domain: 35%, similarity: 56%). This similarity is high enough as to allow the modeling of the 3-dimensional structure of the fungal protein using the available crystallographic data for animal *DJ-1* genes. The model is shown in figure 4. Fungal sequences have a characteristic short extra loop of 2–12 amino acids located close to the *N*-terminus of the DJ-1 domain. However, this loop is in a position that would not interfere with the formation of the characteristic DJ-1 domain fold (asterisk in fig. 4). It is also opposite to the dimerization surface, and therefore, it would not affect the dimer formation that is characteristic of *DJ-1* gene products. Amino acids known to be critical for DJ-1 redox sensitivity (e.g., a cysteine residue that in human DJ-1 protein is located at position 106; Canet-Aviles et al. 2004; Zhou et al. 2006) are also conserved in the fungal sequences. The alignment in figure 5 shows additional evidence for all genes that we have hypothesized to be orthologous to human *DJ-1* actually being true orthologs and for the products of all of them being able to dimerize in the way described for human DJ-1 and *E. coli* YajL proteins. In that figure, we have included a canonical member of each of the main type of DJ-1-related sequences detected in our analyses plus canonical examples of human, fungal, and plant DJ-1 sequences (in this last case, the 2 DJ-1 domains characteristic of those proteins are included). Two results are noteworthy. First, as it could be predicted from our phylogenetic trees and is obvious observing the alignment, eukaryotic DJ-1 proteins are more similar among them and with prokaryotic YajL proteins (identity in the DJ-1 domain: 27–38%, similarity: 48–56%) than with the rest of DJ-1-related sequences (identity: 12–20%, similarity: 24–38%). Second, and very important, all eukaryotic DJ-1 proteins (both DJ-1 domains in plant sequences) and prokaryotic YajL proteins have an additional alpha helix at the C-terminus of their DJ-1 domain (see fig. 5). This helix is known to be part of the dimerization surface of human DJ-1 and *Escherichia* YajL proteins (Honbou et al. 2003; Tao and Tong 2003; Wilson et al. 2003, 2005). It is absent from other proteins, such as PfpI, Hsp31, or YDR533c, that use different modes of dimer and/or multimer formation (Du et al. 2000; Quigley et al. 2003; Tao and Tong 2003; Graille et al. 2004; Wilson et al. 2004, 2005). The presence of this additional helix in the DJ-1-like plant and fungal sequences is an independent confirmation that of all of them are closely related to the animal DJ-1 proteins. Significantly, the only other type of DJ-1-related

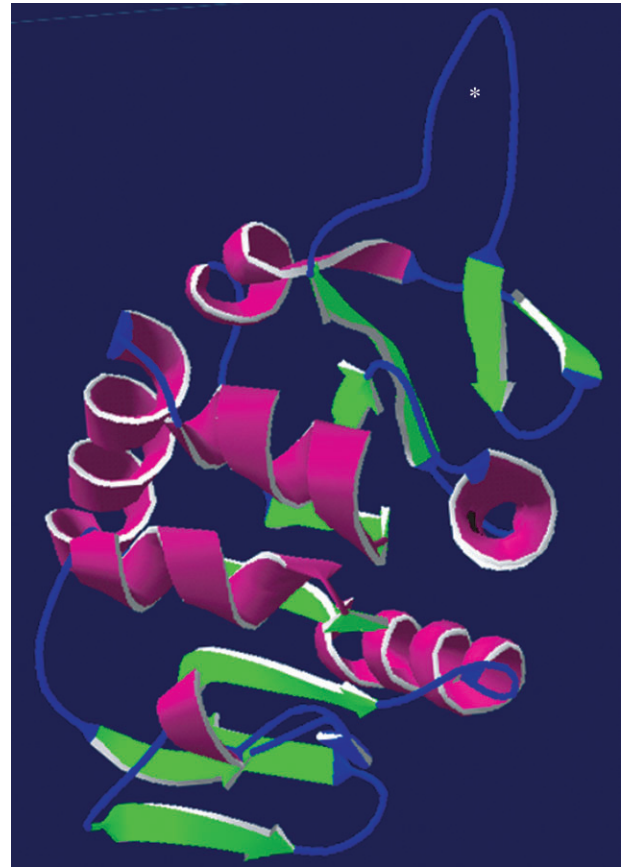


FIG. 4.—Model for the 3-dimensional structure of the product of the *DJ-1*-like gene of the fungus *Coprinopsis cinerea*. This figure has been oriented similar to figure 1a in Wilson et al. (2003), which shows the structure of a human DJ-1 dimer. Dimerization surfaces would be situated on the right of this figure. The only significant difference between human DJ-1 and this model for the fungal protein is the loop marked with an asterisk, located externally to the dimerization surface.

proteins that may contain this additional helix are those in the DJ-1/AraC group, which are characterized by having an additional AraC domain that can bind DNA (reviewed in Gallegos et al. 1997). In figure 5, the sequence of one of these proteins (from *Pseudomonas aeruginosa*) is shown after eliminating the AraC domain. Notice the similarity in the region that would correspond to DJ-1 helix 8.

Discussion

A precise knowledge of the origin and evolution of a complex family of genes may provide useful functional perspectives. As we have shown in previous works, a detailed consideration of the phylogenetic framework and structural characteristics of the products of genes involved in human diseases may often unearth significant findings (e.g., Marín and Ferrús 2002; Marco et al. 2004; Lucas et al. 2006; Marín 2006). In this study, we have performed detailed analyses in order to establish the origin and evolutionary history of the *DJ-1* genes, known to be involved in Parkinson disease. The main conclusions of our study are as follows: 1) the existence of a group of prokaryotic and eukaryotic *DJ-1*-like genes that are so similar that they cannot

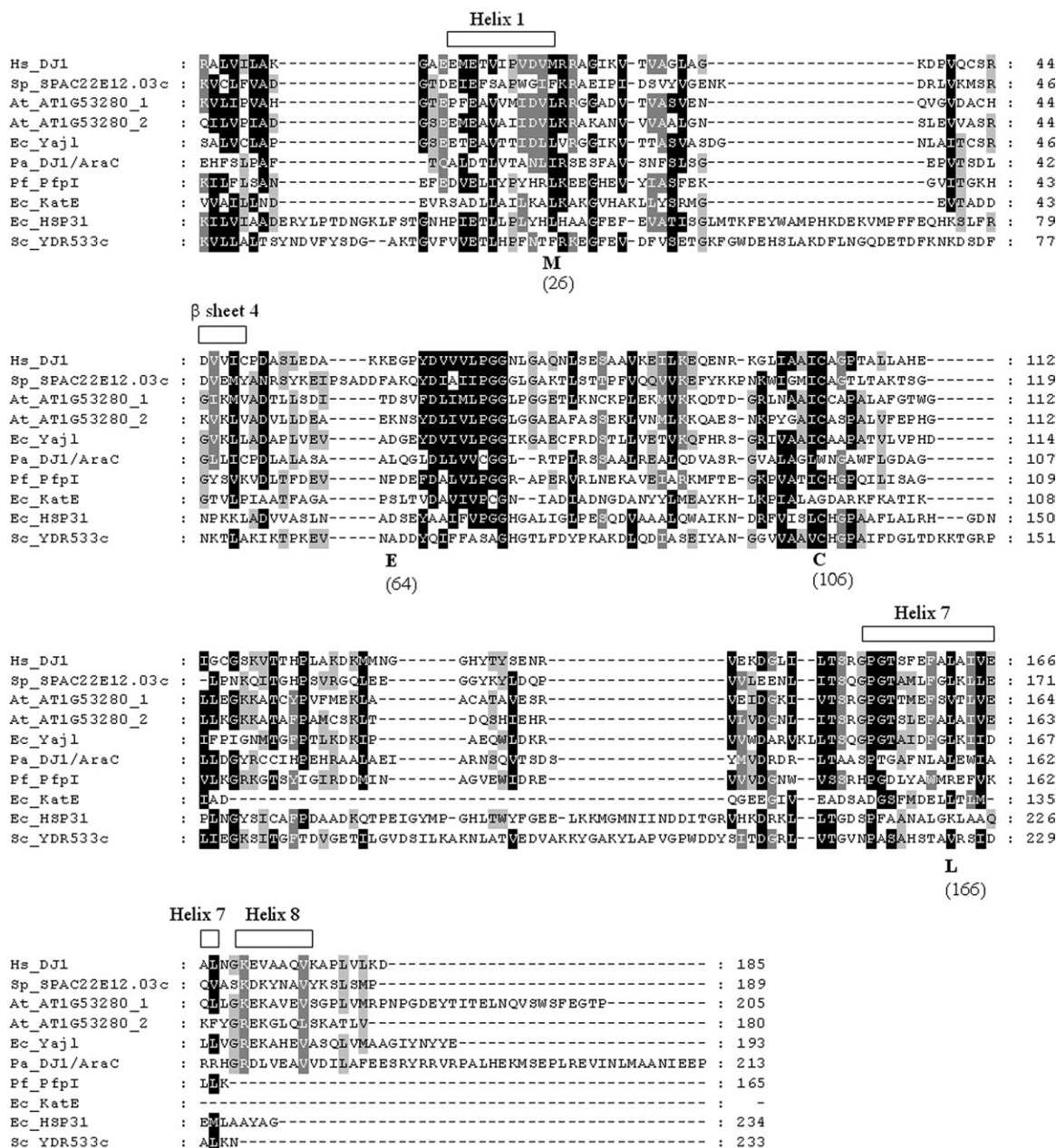


FIG. 5.—Alignment of DJ-1 domains of canonical representatives of the main classes of proteins found in our phylogenetic trees. The positions of alpha helices 1, 7, and 8 and beta sheet 4, which have been implicated in dimerization, are detailed (data from Honbou et al. 2003; Lee et al. 2003; Tao and Tong 2003; Wilson et al. 2003, 2005). Notice the presence of an additional alpha helix (helix 8) in the DJ-1 and Yaj1 sequences that does not appear in most of the other DJ-1-related proteins. The 2 first letters in the names refer to the species from which the sequences derive, as follows: Hs: *Homo sapiens*; Sp: *Schizosaccharomyces pombe*; At: *Arabidopsis thaliana*; Ec: *Escherichia coli*. Pa: *Pseudomonas aeruginosa*; Pf: *Pyrococcus furiosus*; Sc: *Saccharomyces cerevisiae*. For the *Arabidopsis* protein, both DJ-1 domains are shown. The amino acid positions referred in the text, either those in which missense mutations have been found in Parkinson disease patients (M26, E64, L166) or the one involved in DJ-1 redox sensitivity (C106) are also shown.

be separated by sequence analyses of their protein products. At least 5 of the 6 main eukaryotic groups possess this type of genes and the phylogenetic range in prokaryotes is also very wide. These results suggest that they have an ancient origin. 2) The presence in all these sequences of a particular structural feature, an additional helix at the end of the DJ-1 domain, which supports their close evolutionary relationship. Notable is the fact that plant proteins, which have 2 DJ-1 domains, conserve this feature in both of them, suggesting that these domains may interact in the same way as

the monomers of human DJ-1 interact in composing the homodimer that has been observed in crystal structures of the protein. This plant-specific domain duplication must be quite recent, after the plant/red alga split, because the DJ-1-like gene found in the red alga *Cyanidioschyzon* lacks it (see fig. 3). The finding of a similar helix in DJ-1/AraC proteins is interesting, especially considering that this helix is involved in dimerization in other types of AraC-containing proteins (Soisson et al. 1997a, 1997b). A possible, although highly speculative, explanation would be that current DJ-1

genes evolved from a *DJ-1/AraC* gene, by a fission in which the AraC domain was almost completely eliminated, but left behind a single alpha helix. 3) The finding of *DJ-1*-like genes in fungi, and most especially in *Schizosaccharomyces*, suggests new ways of examining the function of *DJ-1* genes. We think that the characterization of the endogenous roles of the fungal *DJ-1* genes by both biochemical analyses of *S. pombe DJ-1* functions and genetic screenings for suppressors or enhancers of *DJ-1* mutations in the yeast may provide many significant insights of the in vivo roles of this interesting class of genes. Such experiments have already been successfully performed in a different yeast, *S. cerevisiae*, for determining the potential cellular roles of another Parkinson-related protein, α -synuclein (Outeiro and Lindquist 2003; Cooper et al. 2006), although this protein is not even found in yeasts.

Bandyopadhyay and Cookson (2004) examined these same genes, arriving to qualitatively different conclusions. The most important difference is that their analyses concluded that *DJ-1* genes were animal specific and that a few prokaryotic *YajL*-like genes (that they still called Thij and considered related to thiamine biosynthesis) were their only close relatives. Part of the discrepancies may be due to them considering just a small sample of *DJ-1*-related sequences. However, it must also be pointed out that their study is technically deficient (superficial phylogenetic analyses, ML is confused with MP, it is stated that bootstrap is used as a measure of branch length, etc.) and contains several factual mistakes, the most striking being that they explicitly depicted 2 independent times *Plasmodium falciparum*, *Plasmodium yoelii*, and *Giardia lamblia* as prokaryotes (see their figs. 2 and 3). In summary, their conclusion of a close proximity of eubacterial and animal genes, with *DJ-1* genes being absent in all the other eukaryotes, is an artifact. We hope that our study will contribute to debunk this mistaken paradigm and thus to open new venues of research.

Acknowledgments

Our group is supported by grants SAF2003-09506 and SAF2006-08977 (Ministerio de Educación y Ciencia, Spain). J.I.L. is supported by Generalitat Valenciana.

Literature Cited

- Abou-Sleiman PM, Healy DG, Quinn N, Lees AJ, Wood NW. 2003. The role of pathogenic *DJ-1* mutations in Parkinson's disease. *Ann Neurol*. 54:283–286.
- Abou-Sleiman PM, Muqit MMK, Wood NW. 2006. Expanding insights of mitochondrial dysfunction in Parkinson's disease. *Nat Rev Neurosci*. 7:207–219.
- Annesi G, Savettieri G, Pugliese P, et al. (15 co-authors). 2005. *DJ-1* mutations and parkinsonism-dementia-amyotrophic lateral sclerosis complex. *Ann Neurol*. 58:803–807.
- Backstrom AD. 1996. The biosynthesis of thiamin in *Escherichia coli* K-12: structural genes for thiamin biosynthetic enzymes (*thiCEFGH* and *thiJ*) and function of the *thiE* gene product (thiamin phosphate synthase [E.C. 2.5.1.3]) [PhD thesis]. Ithaca (NY): Cornell University.
- Bandyopadhyay S, Cookson MR. 2004. Evolutionary and functional relationships within the DJ-1 superfamily. *BMC Evol Biol*. 4:6.
- Baulac S, LaVoie MJ, Strahle J, Schlossmacher MG, Xia W. 2004. Dimerization of Parkinson's disease-causing DJ-1 and formation of high molecular weight complexes in human brain. *Mol Cell Neurosci*. 27:236–246.
- Blackinton J, Ahmad R, Miller DQ, van der Brug MP, Canet-Avilés RM, Hague SM, Kaleem M, Cookson MR. 2005. Effects of *DJ-1* mutations and polymorphisms on protein stability and subcellular localization. *Mol Brain Res*. 134:76–83.
- Bonifati V, Oostra BA, Heutink P. 2004. Linking *DJ-1* to neurodegeneration offers novel insights for understanding the pathogenesis of Parkinson's disease. *J Mol Med*. 82:163–174.
- Bonifati V, Rizzu P, van Baren MJ, et al. (18 co-authors). 2003. Mutation in the *Dj-1* gene associated with autosomal recessive early-onset parkinsonism. *Science*. 299:256–259.
- Canet-Aviles RM, Wilson MA, Miller DW, Ahmad R, McLendon C, Bandyopadhyay S, Baptista MJ, Ringe D, Petsko GA, Cookson MR. 2004. The Parkinson's disease protein *DJ-1* is neuroprotective due to cysteine-sulfenic acid-driven mitochondrial localization. *Proc Natl Acad Sci USA*. 101:9103–9108.
- Chor B, Hendy MD, Holland BR, Penny D. 2000. Multiple maxima of likelihood in phylogenetic trees: an analytic approach. *Mol Biol Evol*. 17:1529–1541.
- Cooper AA, Gitler AD, Cashikar A, et al. (19 co-authors). 2006. Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science*. 313:324–328.
- Du X, Choi IG, Kim R, Wang W, Jancarik J, Yokota H, Kim SH. 2000. Crystal structure of an intracellular protease from *Pyrococcus horikoshii* at 2-Å resolution. *Proc Natl Acad Sci USA*. 97:14079–14084.
- Gallegos MT, Schleif R, Bairoch A, Hofmann K, Ramos JL. 1997. AraC/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev*. 61:393–410.
- Goldberg MS, Pisani A, Haburcak M, et al. (15 co-authors). 2005. Nigrostriatal dopaminergic deficits and hypokinesia caused by inactivation of the familial parkinsonism-linked gene *DJ-1*. *Neuron*. 45:489–496.
- Graille M, Quevillon-Cheruel S, Leulliot N, et al. (11 co-authors). 2004. Crystal structure of the YDR533c *S. cerevisiae* protein, a class II member of the Hsp31 family. *Structure*. 12:839–847.
- Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 18:2714–2723.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 52:696–704.
- Hering R, Strauss KM, Tao X, et al. (19 co-authors). 2004. Novel homozygous p.E64D mutation in *DJ-1* in early onset Parkinson disease (PARK7). *Hum Mutat*. 24:321–329.
- Hod Y, Pentyala SN, Whyard TC, El-Maghrabi MR. 1999. Identification and characterization of a novel protein that regulates RNA-protein interaction. *J Cell Biochem*. 72:435–444.
- Honbou K, Suzuki NN, Horiuchi M, Niki T, Taira T, Ariga H, Inagaki F. 2003. The crystal structure of DJ-1, a protein related to male fertility and Parkinson's disease. *J Biol Chem*. 278:31380–31384.
- Horvath MM, Grishin NV. 2001. The C-terminal domain of HPII catalase is a member of the type I glutamine amidotransferase superfamily. *Proteins*. 42:230–236.
- Huai Q, Sun YJ, Wang HC, Chin LS, Li L, Robinson H, Ke HM. 2003. Crystal structure of DJ-1/RS and implication on familial Parkinson's disease. *FEBS Lett*. 549:171–175.
- Junn E, Taniguchi H, Jeong BS, Zhao X, Ichijo H, Mouradian MM. 2005. Interaction of DJ-1 with Daxx inhibits apoptosis signal-regulating kinase 1 activity and cell death. *Proc Natl Acad Sci USA*. 102:9691–9696.

- Kim RH, Peters M, Jang YJ, et al. (16 co-authors). 2005. Dj-1, a novel regulator of the tumor suppressor PTEN. *Cancer Cell*. 7:263–273.
- Kim RH, Smith PD, Aleyasin H, et al. (15 co-authors). 2005. Hypersensitivity of *DJ-1*-deficient mice to 1-methyl-4-phenyl-1,2,3,4-tetrahydropyridine (MPTP) and oxidative stress. *Proc Natl Acad Sci USA*. 102:5215–5220.
- Kumar S, Tamura K, Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform*. 5:150–163.
- Lee SJ, Kim SJ, Kim IK, et al. (11 co-authors). 2003. Crystal structures of human DJ-1 and *Escherichia coli* Hsp31, which share an evolutionarily conserved domain. *J Biol Chem*. 278:44552–44559.
- Lucas JI, Arnau V, Marín I. 2006. Comparative genomics and protein domain graph analyses link ubiquitination and RNA metabolism. *J Mol Biol*. 357:9–17.
- Macedo MG, Anar B, Bronner IF, Cannella M, Squitieri F, Bonifati V, Hoogeveen A, Heutink P, Rizzu P. 2003. The DJ-1L166P mutant protein associated with early onset Parkinson's disease is unstable and forms higher-order protein complexes. *Hum Mol Genet*. 12:2807–2816.
- Marco A, Cuesta A, Pedrola L, Palau F, Marín I. 2004. Evolutionary and structural analyses of *GDAP1*, involved in Charcot-Marie-Tooth disease, characterize a novel class of glutathione transferase-related genes. *Mol Biol Evol*. 21:176–187.
- Marín I. 2006. The Parkinson disease gene *LRRK2*: evolutionary and structural insights. *Mol Biol Evol*. 23:2423–2433.
- Marín I, Ferrús A. 2002. Comparative genomics of the RBR family, including the Parkinson's disease-related gene *parkin* and the genes of the ariadne subfamily. *Mol Biol Evol*. 19:2039–2050.
- Marín I, Lucas JI, Gradilla AC, Ferrús A. 2004. Parkin and relatives: the RBR family of ubiquitin ligases. *Physiol Genomics*. 17:253–263.
- Martinat C, Shendelman S, Jonason A, Leete T, Beal MF, Yang L, Floss T, Abeliovich A. 2004. Sensitivity to oxidative stress in *DJ-1*-deficient dopamine neurons: an ES-derived cell model of primary parkinsonism. *PLoS Biol*. 2:1754–1763.
- Menzies FM, Yenissetti SC, Min KT. 2005. Roles of *Drosophila Dj-1* in survival of dopaminergic neurons and oxidative stress. *Curr Biol*. 15:1578–1582.
- Meulener M, Whitworth AJ, Armstrong-Gold CE, Rizzu P, Heutink P, Wes PD, Pallanck LJ, Bonini NM. 2005. *Drosophila DJ-1* mutants are selectively sensitive to environmental toxins associated with Parkinson's disease. *Curr Biol*. 15:1572–1577.
- Miller DW, Ahmad R, Hague S, et al. (13 co-authors). 2003. L166P mutant *DJ-1*, causative for recessive Parkinson's disease, is degraded through the ubiquitin-proteasome system. *J Biol Chem*. 278:36588–36595.
- Moore DJ, Zhang L, Dawson TM, Dawson VL. 2003. A missense mutation (L166P) in *DJ-1*, linked to familial Parkinson's disease, confers reduced protein stability and impairs homooligomerization. *J Neurochem*. 87:1558–1567.
- Mueller EG, Buck CJ, Palenchar PM, Barnhart LE, Paulson JL. 1998. Identification of a gene involved in the generation of 4-thiouridine in tRNA. *Nucleic Acids Res*. 26:2606–2610.
- Nagakubo D, Taira T, Kiatura H, Ikeda M, Tamai K, Iguchi-Arigo SMM, Ariga H. 1997. *DJ-1*, a novel oncogene which transforms mouse NIH3T3 cells in cooperation with ras. *Biochem Biophys Res Commun*. 231:509–513.
- Neumann M, Müller V, Görner K, Kretschmar HA, Haass C, Kahle PJ. 2004. Pathological properties of the Parkinson's disease-associated protein DJ-1 in α -synucleinopathies and tauopathies: relevance for multiple system atrophy and Pick's disease. *Acta Neuropathol*. 107:489–496.
- Nicholas KB, Nicholas HB Jr, Deerfield DW. 1997. GeneDoc: analysis and visualization of genetic variation. *Embnew.News*. 4:14.
- Olzmann JA, Brown K, Wilkinson KD, Rees HD, Huai Q, Ke H, Levey AI, Li L, Chin LS. 2004. Familial Parkinson's disease-associated L166P mutation disrupts DJ-1 protein folding and function. *J Biol Chem*. 279:8506–8515.
- Outeiro TF, Lindquist S. 2003. Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science*. 302:1772–1775.
- Park J, Kim SY, Cha GH, Lee SB, Kim S, Chung J. 2005. *Drosophila DJ-1* mutants show oxidative stress-sensitive locomotive dysfunction. *Gene*. 361:133–139.
- Peitsch MC. 1996. ProMod and Swiss-Model: internet-based tools for automated comparative protein modelling. *Biochem Soc Trans*. 24:274–279.
- Quigley PM, Korotkov K, Bañéis F, Hol WGJ. 2003. The 1.6-Å crystal structure of the class of chaperones represented by *Escherichia coli* Hsp31 reveals a putative catalytic triad. *Proc Natl Acad Sci USA*. 100:3137–3142.
- Rizzu P, Hinkle DA, Zhukareva V, et al. (12 co-authors). 2004. DJ-1 colocalizes with Tau inclusions: a link between Parkinsonism and dementia. *Ann Neurol*. 55:113–118.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 4:406–425.
- Shendelman S, Jonason A, Martinat C, Leete T, Abeliovich A. 2004. DJ-1 is a redox-dependent molecular chaperone that inhibits α -synuclein aggregate formation. *PLoS Biol*. 2:1764–1773.
- Simpson AG, Roger AJ. 2004. The real 'kingdoms' of eukaryotes. *Curr Biol*. 14:R693–R696.
- Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. 1997a. Structural basis for ligand-regulated oligomerization of AraC. *Science*. 276:421–425.
- Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. 1997b. The 1.6 Å crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. *J Mol Biol*. 273:226–237.
- Swofford DL. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland (MA): Sinauer Associates.
- Takahashi-Niki K, Niki T, Taira T, Iguchi-Arigo SMM, Ariga H. 2004. Reduced anti-oxidative stress activities of DJ-1 mutants found in Parkinson's disease patients. *Biochem Biophys Res Commun*. 320:389–397.
- Tao X, Tong L. 2003. Crystal structure of human DJ-1, a protein associated with early onset Parkinson's disease. *J Biol Chem*. 278:31372–31379.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 24:4876–4882.
- Ved R, Saha S, Westlund B, et al. (13 co-authors). 2005. Similar patterns of mitochondrial vulnerability and rescue induced by genetic modification of α -synuclein, parkin and DJ-1 in *Caenorhabditis elegans*. *J Biol Chem*. 280:42655–42668.
- Wagenfeld A, Gromoll J, Cooper TG. 1998. Molecular cloning and expression of rat contraception associated protein 1 (CAP1), a protein putatively involved in fertilization. *Biochem Biophys Res Commun*. 251:545–549.
- Welch JE, Barbee RR, Roberts NL, Suarez JD, Klinefelter GR. 1998. SP22: a novel fertility protein from a highly conserved gene family. *J Androl*. 19:385–393.
- Wilson MA, Collins JL, Hod Y, Ringe D, Petsko GA. 2003. The 1.1-Å resolution crystal structure of DJ-1, the protein mutated

- in autosomal recessive early onset Parkinson's disease. *Proc Natl Acad Sci USA*. 100:9256–9261.
- Wilson MA, Ringe D, Petsko GA. 2005. The atomic resolution crystal structure of the YajL (ThiJ) protein from *Escherichia coli*: a close prokaryotic homologue of the Parkinsonism-associated protein DJ-1. *J Mol Biol*. 353:678–691.
- Wilson MA, St. Amour CV, Collins JL, Ringe D, Petsko GA. 2004. The 1.8-Å resolution crystal structure of YDR533Cp from *Saccharomyces cerevisiae*: a member of the DJ-1/Thij/Pfpl superfamily. *Proc Natl Acad Sci USA*. 101:1531–1536.
- Xu J, Zhong N, Wang H, Elias JE, Kim CY, Woldman I, Pifl C, Gygi SP, Geula C, Yankner BA. 2005. The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis. *Hum Mol Genet*. 14:1231–1241.
- Yang Y, Gehrke S, Haque ME, et al. (12 co-authors). 2005. Inactivation of *Drosophila DJ-1* leads to impairments of oxidative stress response and phosphatidylinositol 3-kinase/Akt signaling. *Proc Natl Acad Sci USA*. 102:13670–13675.
- Yokota T, Sugawara K, Ito K, Takahashi R, Ariga H, Mizusawa H. 2003. Down regulation of *DJ-1* enhances cell death by oxidative stress, ER stress, and proteasome inhibition. *Biochem Biophys Res Commun*. 312:1342–1348.
- Zhang L, Shimoji M, Thomas B, et al. (11 co-authors). 2005. Mitochondrial localization of the Parkinson's disease related protein DJ-1: implications for pathogenesis. *Hum Mol Genet*. 14:2063–2073.
- Zhou W, Freed CR. 2005. DJ-1 up-regulates glutathione synthesis during oxidative stress and inhibits A53T α -synuclein toxicity. *J Biol Chem*. 280:43150–43158.
- Zhou W, Zhu M, Wilson MA, Petsko GA, Fink AL. 2006. The oxidation state of DJ-1 regulates its chaperone activity toward α -synuclein. *J Mol Biol*. 356:1036–1048.

Claudia Schmidt-Dannert, Associate Editor

Accepted November 15, 2006