BIOINFORMATICS ORIGINAL PAPER

Vol. 26 no. 7 2010, pages 946–952
doi:10.1093/bioinformatics/btq073

Data and text mining

Advance Access publication February 19, 2010

A meta-analysis of two-dimensional electrophoresis pattern of the Parkinson's disease-related protein DJ-1

Massimo Natale[1], Dario Bonino[1], Paolo Consoli[1], Tiziana Alberio[2], Rivka G. Ravid[3], Mauro Fasano[2] and Enrico M. Bucci[1,4,*]

[1]BioDigitalValley S.r.l., Via Carlo Viola 78, 11026 Pont Saint Martin (AO), [2]Department of Structural and Functional Biology and Center of Neuroscience, University of Insubria, Via Alberto da Giussano 12, 21052, Busto Arsizio (VA), Italy, [3]Brainbank Consultants, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands and [4]Istituto di Biostrutture e Bioimmagini, Via Mezzocannone 16, 80134, Naples, Italy

Associate Editor: Jonathan Wren

ABSTRACT

Motivation: The two-dimensional electrophoresis (2-DE) pattern of proteins is thought to be specifically related to the physiological or pathological condition at the moment of sample preparation. On this ground, most proteomic studies move to identify specific hallmarks for a number of different conditions. However, the information arising from these investigations is often incomplete due to inherent limitations of the technique, to extensive protein post-translational modifications and sometimes to the paucity of available samples.

The meta-analysis of proteomic data can provide valuable information pertinent to various biological processes that otherwise remains hidden.

Results: Here, we show a meta-analysis of the PD protein DJ-1 in heterogeneous 2-DE experiments. The protein was shown to segregate into specific clusters associated with defined conditions.

Interestingly, the DJ-1 pool from neural tissues displayed a specific and characteristic molecular weight and isoelectric point pattern. Moreover, changes in this pattern have been related to neurodegenerative processes and aging. These results were experimentally validated on human brain specimens from control subjects and PD patients.

Availability: ImageJ is a public domain image processing program developed by the National Institutes of Health and is freely available at http://rsbweb.nih.gov/ij. All the ImageJ macros used in this study are available as supplementary material and upon request at info@biodigitalvalley.com. XLSTAT can be purchased online at http://www.xlstat.com/en/home/ at a current cost of ∼300 EUR.

Contact: enrico.bucci@biodigitalvalley.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received and revised on February 12, 2010; accepted on February 16, 2010

1 INTRODUCTION

DJ-1/PARK7 is an ubiquitous, highly conserved protein that was originally identified because of its ability to transform NIH3T3 mouse cells in cooperation with Ras (Nagakubo *et al.*, 1997). Starting from the association between loss-of-function mutations in the DJ-1 gene and PARK7, a monogenic, autosomal-recessive form of Parkinson's disease (PD) (Bonifati *et al.*, 2003), an accumulating body of evidence pinpointed the important role of DJ-1 in this neurodegenerative condition. Very recently, it has been shown that oxidized dopamine can covalently modify DJ-1 (Van Laar *et al.*, 2009); however, whether this can affect dopamine cell degeneration is unknown. Some hints may come from the involvement of DJ-1 into many cellular functions, including evidence linking this protein to oxidative stress response—a fact well known even before the association of DJ-1 with PD (Mitsumoto *et al.*, 2001)—mitochondrial function (Zhang *et al.*, 2005) and transcription (Zhong and Xu, 2008), nuclear transcription (Xu *et al.*, 2005), mRNA binding and protein interaction (Hod *et al.*, 1999; van der Brug *et al.*, 2008) and protein degradation (Xiong *et al.*, 2009). Mirroring the involvement of DJ-1 in multiple cellular activities, this protein has been found in complex with multiple molecular partners, including DJ-1 itself (Tao and Tong, 2003), PINK-1 and Parkin (Xiong *et al.*, 2009), alpha-synuclein (Meulener *et al.*, 2005), HSP70 (Li *et al.*, 2005), DJBP (Niki *et al.*, 2003), PIASx alpha (Takahashi *et al.*, 2001), ASK1 (Waak *et al.*, 2009), histone deacetylase 6 (Olzmann *et al.*, 2007), androgen receptor (Tillman *et al.*, 2007), DAXX (Junn *et al.*, 2005) and Abstrakt (Sekito *et al.*, 2005).

DJ-1 has been shown to modulate dopamine toxicity in cellular models of oxidative stress with reference to PD (Fasano *et al.*, 2008). Dopamine exposure leads to upregulation of DJ-1 that in turn increases cell resistance to dopamine itself and reduces intracellular oxidants (Lev *et al.*, 2008; Lev *et al.*, 2009). On the other hand, α-synuclein overexpression leads to upregulation of DJ-1, and DJ-1 overexpression protects cells from α-synuclein toxicity (Batelli *et al.*, 2009; Colapinto *et al.*, 2006; Zhou and Freed, 2005).

Besides being in complex with a number of different partners, DJ-1 is often post-translationally modified. DJ-1 modifications mainly include oxidations at different sites, which is related to its antioxidant role (Choi *et al.*, 2006; Kinumi *et al.*, 2004; Mitsumoto *et al.*, 2001), but there are also evidences of ubiquitination (Olzmann *et al.*, 2004, 2007) and SUMOylation (small ubiquitin-like modifier) (Shinbo *et al.*, 2006).

Not unexpectedly, DJ-1 binding to its molecular counterparts, and thus its pleiotropic effects, are affected by DJ-1 post-translational modification. For example, oxidation regulates homodimerization (Ito *et al.*, 2007) and affects DJ-1 binding to mRNA (van der Brug *et al.*, 2008), ASK1 (Waak *et al.*, 2009), HSP70 (Li *et al.*, 2005),

*To whom correspondence should be addressed.

parkin (Moore *et al.*, 2005), while ubiquitination mediates the binding to histone deacetylase 6 and thus to the dynein–dynactin complex (Olzmann *et al.*, 2007).

Multiple DJ-1 modified forms are simultaneously present, so that DJ-1 can be better considered as a pool of different forms, with different modifications and in different amounts. It is very likely that, instead of the total amount of DJ-1 or of a defined DJ-1 form, the composition of this pool and the precise balance between different forms is the main factor determining DJ-1 global activity. In particular, alterations of this pool, instead of DJ-1 mutations, are expected to play a role in those non-genetic conditions correlated to DJ-1 activity, including idiopathic PD. Indeed, since many different DJ-1 forms can be separated on the basis of their isoelectric point (pI), it is a common finding that DJ-1 oxidation, correlated to aging, Parkinson's and Alzheimer's diseases, produces an increase in DJ-1 species of acidic pH, and a decrease in basic species, so that these conditions are characterized by a pool of DJ-1 forms different from that observed in controls (Canet-Avilés *et al.*, 2004; Choi *et al.*, 2006; Meulener *et al.*, 2006). Ubiquitination and SUMOylation of DJ-1, on the other side, affects also the molecular weight (MW) of the modified species, and are thus separated by mono-dimensional electrophoresis accordingly (Olzmann *et al.*, 2004, 2007; Shinbo *et al.*, 2006).

While two-dimensional electrophoresis (2-DE) is able in principle to separate DJ-1 modified forms on the basis of both pI and MW changes, to the best of our knowledge there is no experimental evaluation of DJ-1 modifications, which takes into account both dimensions simultaneously, so as to completely describe the pool of DJ-1 possible forms and to evaluate the pool composition in a given condition.

To this purpose, a robust procedure was developed to reevaluate 2-DE experiments available in the literature by automatic alignment. Since detailed information on experimental conditions, protocols and samples was available, it was possible to segregate different DJ-1 forms into pools associated with a defined condition independently by the particular experimental protocols used. Interestingly, the DJ-1 pool from neural tissues, in particular from brain, displayed a specific and characteristic MW and pI pattern. Moreover, changes in this pattern might reflect neurodegenerative processes and aging. These results were experimentally validated by 2-DE western blotting on human post-mortem brain specimens from control subjects and PD patients.

2 METHODS

2.1 Search strategy and 2D-gel data extraction

A set of relevant data to be included into the analysis, i.e. 2D-gel experiments where the protein DJ-1 was clearly identified, was obtained by extracting the relevant 2D-gel images from Pubmed-indexed Portable Document Format documents through July 2009. Briefly, we started by querying Medline for publications indexed with the subject heading DJ-1 or its alias: "PARK7"[All Fields] OR "CTA-215D11.1"[All Fields] OR "DJ-1"[All Fields] OR "DJ1"[All Fields] OR "FLJ27376"[All Fields] OR "FLJ34360"[All Fields] OR "RP23-272N19.5"[All Fields]).

As of July, 2009, this query identified 474 papers. We then proceed to parsing using Big Faceless Organization software (KSM Ltd, London, UK), so as to extract the figures and the figures captions of the considered papers.

Those figures containing 2D-gel representations were identified by a text-matching procedure, and able to evaluate each caption for the occurrence of at least one term identified as connected to a 2D-gel experiment (for the list of terms used, see Supplementary Table 2). This step led to the identification of 53 figures from 37 papers, which were automatically segmented to isolate 160 2D-gel images. The 2D-gel sources included in the analysis are listed in Supplementary Table 2.

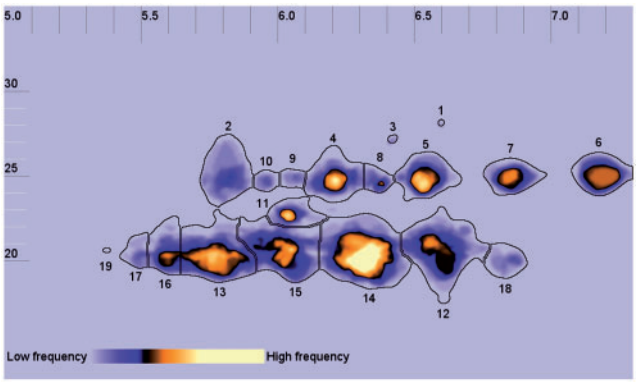

Fig. 1. DJ-1 metagel, derived from 160 different bidimensional experiments, representing the overall pool of DJ-1 possible forms. Conserved spots are numbered and bordered in black. The color scale range from blue (spots poorly represented in the considered experiments) to white (spots very represented in the considered experiments).

2.2 2D-gel alignment and 'metagel' production

The 2D-gel images were processed using ImageJ (National Institutes of Health) (Girish and Vijayalakshmi, 2004), including DJ-1 spots that could not be directly aligned, given the differences in resolution of the original figures and the different pI/MW range represented. Therefore, gel images were rescaled by calibrating the pixel-to-pI ratio using the following factor:

$$F = \mathrm{pix}_{pI} \times \frac{\mathrm{pI}_B - \mathrm{pI}_A}{\mathrm{pix}_B - \mathrm{pix}_A}$$

where pix_{pI} was the desired resolution in the final image (i.e. 400 pixels per pI unit), pI_B and pI_A were the pI values of the center of mass of the most basic and of the most acid spot, respectively, and pI_B and pI_A the pixel number of the center of mass of the most basic and of the most acid spot, respectively. For MW calibration, the average position of the mass center of the four most intense DJ-1 spots was assumed to correspond to the mass reported for DJ-1.

Once rescaled and registered on the reference pI/MW grid, gel images were background subtracted, segmented for the spots and binarized. To this aim, a rolling-ball procedure was used first (rolling ball radius = 50 pixel, performing a sliding paraboloid), then a watershed segmentation procedure was applied (smooth radius 5.0, dark objects on bright background, neighborhood of 8 pixels, minimum level '0', maximum level '120', display outputs 'binary image') and finally an automatic thresholding to get the binarized image (Pleissner *et al.*, 1999).

The 160 binarized images were used to get a 'metagel', i.e. an average intensity projection image, wherein each pixel stores the average intensity over all images in the 160 original images at corresponding pixel location (Fig. 1). The color scale of this metagel accounts for the occurrence frequency of each spot in the considered experiments, with less frequent spot positions in blue and more frequent spot positions in white.

The same approach was used to build two additional 'metagels' which represent all the experiments of Alzheimer's/PD patient brain ($n = 11$) and all the experiments of control human brain origin ($n = 12$).

The entire procedure is represented as a flowchart in Supplementary Figure 5.

2.3 'Spot matrix' generation

In order to gain insight on possible correlations between specific spot patterns on 2D-gel experiments and biologically relevant information, an objective matrix was generated to correlate each spot in a given 2D-gel to one of the reference spot of the metagel, i.e. to one of the DJ-1 forms occurring in at least 5% of the considered experiments. Briefly, the following, fully automated procedure was performed: (i) the 'metagel' was segmented as described above for the experimental gels; (ii) the Euclidean distance of the center of mass of each experimental spot from the center of mass of each spot in the 'metagel' was calculated; and (iii) a 'spot matrix' was build, where rows correspond to two dimensional gel (2D-gel) experiments, and columns to the spots in the 'metagel'.

A value of 1 was assigned if the Euclidean distance between any spot of the considered experiment and the 'metagel' spot defining the matrix column was minor or equal to the Feret diameter of the 'metagel' spot (i.e. the considered experiment contains a spot that is covered at least partially by a specific spot of the 'metagel').

The 'spot matrix' so obtained is thus a synthetic representation of the occurrence of each spot, identified in the 'metagel' and corresponding to a form of DJ-1 in a given experiment, allowing the description of each experiment in terms of which DJ-1 forms were present.

2.4 Statistics

The spot matrix was analyzed using the software XLSTAT (Addinsoft, Paris, France) as described hereafter.

The co-occurrence of different spots was measured to group those spots which most often are present simultaneously (i.e. to get pools of DJ-1 forms) according to Pearson's correlation method applied row by row in the 'Spot matrix': the linear correlation was measured between the presence of a given spot in a given experiment (expressed by a value of 1 in the corresponding cell of the spot matrix) and the presence of the other spots in the same experiment (expressed by the value, 0 or 1, of the other spots in the same row). The correlation alpha significance threshold to include a spot in a pool was set to 0.01.

Grouping of the considered experiments was achieved by agglomerative hierarchical clustering in order to put together different experiments on the basis of the similarity of their vectorial representation in the 'spot matrix', i.e. on the basis of the occurrence of similar spots. The similarity between experiments was computed as Euclidean distances for each pair of rows in the 'spot matrix'. The Euclidean distances (d) are given for each pair of rows, $A = (a_1,\ a_2, \ldots, a_p)$ and $B = (b_1,\ b_2, \ldots, b_p)$, by the equation:

$$d(A,B) = \sqrt{\sum_{j=1}^{p} (a_j - b_j)^2}$$

where p is the number of columns/variables (19 in our case). The resulting distance matrix was used to cluster the experiments according to the Ward's method (Ward, 1963). Here, the clustering criterion is based on the error sum of squares, E, which is defined as the sum of the squared distances of individuals from the center of gravity of the cluster to which they have been assigned. Initially, E is 0, since every individual is in a cluster of its own. At each stage, the link created is the one that makes the least increase to E. Aggregation was stopped by automatic truncation, based on the entropy level, so to get groups as homogeneous as possible.

2.5 Human substantia nigra specimens

Specimens of human substantia nigra pars compacta (SNpc) from four patients having known sporadic PD history and from three unaffected controls were provided by the Nederlandse Hersenbank (the Netherland's Brain Bank, NBB), Amsterdam. Each specimen was accompanied by histopathological report, showing depletion of pigmented dopaminergic neurons and the occurrence of extracellular neuromelanin granules as well as Lewy bodies in surviving neurons of the PD patient. Human brain material was obtained via the rapid autopsy system of the NBB that supplies post-mortem specimens from clinically well documented and neuropathologically confirmed cases. For the staging of pathological hallmarks, a combination of parameters was applied (Braak *et al.*, 2003; Ravid *et al.*, 1998). Details are given in Table 1. Autopsies were performed on donors from whom written informed consent was obtained either from the donor or direct next of kin. In agreement with Dutch law, no ethics statement is required. All research involving human brain material, however, was conducted according to the ethical declaration of the NBB. Human brain tissue specimens (0.5–0.7 g) were homogenized (20 up-and-down strokes with a Teflon homogenizer followed by five passages through a 22-gauge needle) in 3 µl UTC (7M Urea, 2M Thiourea, 4% CHAPS) lysis solution per mg tissue. After centrifugation at 21 000*g* for 30 min at 20°C, supernatants were recovered; protein content was estimated with the Bradford assay.

Table 1. Demographic and autopsy data of patients and controls enrolled in the present study

Label	Sex	Age	PMD	BW	YAD	Fam	ApoE	Tangles
CO235	M	56	5 h 25 m	1522	n.a.	n.a.	$\varepsilon3/\varepsilon3$	0
CO280	M	85	5 h 13 m	1394	n.a.	n.a.	$\varepsilon2/\varepsilon3$	2
CO296	M	49	9 h 50 m	1730	n.a.	n.a.	$\varepsilon3/\varepsilon3$	0
PD145	F	74	4 h 35 m	1061	16	Yes[a]	$\varepsilon3/\varepsilon3$	1
PD152	F	79	5 h 20 m	1158	15	No	$\varepsilon3/\varepsilon4$	1
PD184	M	76	5 h 15 m	1283	10	No	$\varepsilon2/\varepsilon3$	1

PMD, post-mortem delay; BW, brain weight; YAD, years after diagnosis; Fam, familiarity; ApoE, apolipoprotein E polymorphism; n.a., not applicable.
[a]Familiarity found for dementia and PD.

2.6 2D-electrophoresis experiments

Immobilized pH gradients (IPG) strips, IPG buffer, acrylamide and polyvinyldifluoridene membranes were from GE Healthcare, Uppsala, Sweden. The Bradford assay for protein concentration determination was from Bio-Rad Laboratories (Hercules, CA, USA). The chemiluminescence western blotting detection system (SuperSignal West Pico Chemiluminescent) was from Pierce Biotechnology, Inc. (Rockford, IL, USA). Anti-DJ-1 monoclonal antibody was from Medical and Biological Laboratories, Nagoya, Japan. All other reagents were cell-culture grade from Sigma (St Louis, MO, USA) and were used without further purification.

2-DE was performed according to Jacobs (Jacobs *et al.*, 2001), with some modifications. Samples (100 µg) were diluted to 250 µl with UTC containing 0.5% immobilized pH gradient (IPG) buffer 3–10, 2 mM tributylphosphine and traces of bromophenol blue, and loaded on 13 cm IPG DryStrips with a non-linear 3–10 pH gradient by in-gel rehydration. Isoelectrofocusing was performed at 22°C on IPGphor (GE Healthcare) according to the following schedule: 1 h at 0 V, 10 h at 30 V, 1 h at 200 V, 30 min linear gradient to 3500 V, 3 h at 3500 V, 2 h and 30 min of linear gradient to 8000 V and 6 h at 8000 V. Prior to sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE), the IPG strips were equilibrated for 2 × 30 min in 50 mM Tris–HCl pH 8.8, 6 M urea, 30% glycerol, 2% SDS and traces of bromophenol blue containing 1% dithiothreitol for the first equilibration step and 2.5% iodoacetamide for the second one. SDS–PAGE was performed using 12.5% polyacrylamide gels without stacking gel with the Hoefer SE 600 system (GE Healthcare).

The second dimension was carried out at 15 mA/gel for 15 min and at 45 mA/gel, at 16°C, until the bromophenol dye front had migrated to the lower end of the gels. Carbamylated pI and MW markers (GE Healthcare) were used for 2-DE calibration. Gels were transferred to PVDF membranes at 1.5 mA/cm^2 for 1 h, 4°C.

Membranes were saturated in 5% bovine serum albumin (BSA) in tris buffered saline and tween (TBS-T) (0.5 M Tris, pH 7.4, 2 M NaCl and 0.1% Tween-20) and incubated with anti-DJ-1 monoclonal antibody (1 μg/ml in 1% BSA TBS-T) for 1 h at room temperature. Membranes were then washed with TBS-T and incubated with peroxidase-conjugated anti-mouse-IgG antibody (1:80000 in 1% BSA TBS-T) for chemiluminescence detection.

3 RESULTS

3.1 Two pools of DJ-1 forms are alternatively present

Figure 1 reports the distribution of all the DJ-1 major forms represented in the scientific literature considered in the present work. From this Figure, it becomes immediately evident that the protein was reported with different MWs (20 and 25 kDa).

Correlated DJ-1 forms (i.e. DJ-1 forms that appear together in different experiments) were identified by building the correlation matrix reported in Figure 2A, where the cells of the matrix are colored according to the Pearson's correlation coefficient of the corresponding spot pair (see Section 2). Two groups of spots are positively correlated: a first group, including spots 4, 5, 6, 7, 8, and a second group, including spots 12, 13, 14, 15, 16, 17. Out of 160 experiments considered in the present work, 68 contained at least one spot of the correlation group 1, and 54 contained at least one spot from correlation group 2.

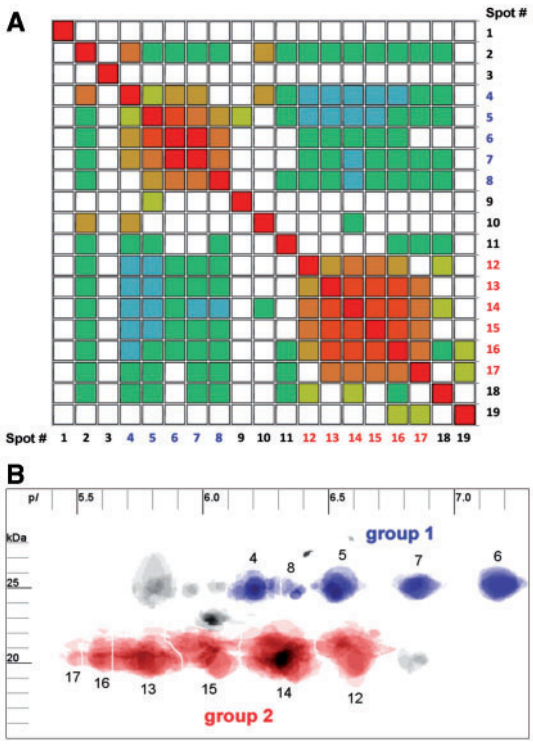

Fig. 2. (**A**) Spot correlation matrix. Cells of the matrix are colored according to the Pearson's coefficient of correlation, from blue (negatively correlated spots) to red (positively correlated spots). White cells represent absence of correlation between the corresponding spots. (**B**) The correlation analysis identified two groups of positively correlated spots, colored in red and blue, respectively, which are mutually exclusive. The average MW and pI of the two groups is different (heavier and more basic in group 1, lighter and more acidic in group 2).

The spots in the two groups are separated on the basis of the apparent MW, because the first group includes only DJ-1 forms with MW ≈ 25 kDa, while in the second group only forms with MW ≈ 20 kDa are present. Therefore, DJ-1 appears either at the expected MW (MW ≈ 20 kDa) or as an heavier, modified form. The latter cannot be explained by a different sequence, since all the species considered in the present study (human, mouse, ovine and bovine) express a protein of predicted MW ≈ 20 kDa. Moreover, DJ-1 forms with MW ≈ 25 kDa are widely represented in the considered papers, and therefore they are not correlated with a particular protocol or a particular laboratory.

The existence of a 25 kDa DJ-1 form is documented in the literature also by a number of one-dimensional western blot experiments. For example, human DJ-1 of a similar MW appears in the works of Shinbo (Shinbo *et al.*, 2006) and Bieler (Bieler *et al.*, 2009), and in rat a similarly heavier DJ-1 was observed in the work of Yanagida (Yanagida *et al.*, 2006). Further support to the existence of a heavier form of DJ-1 comes from the one-dimensional western blot reported as quality control by DJ-1 commercial antibody manufacturers (Supplementary Table 3). Again, it appears that, independently from the particular antibody used, the protein appears alternatively as a 25 kDa or 20 kDa band. Like the MW, the pI of the two identified pools is also different. The average pI of the 25 kDa pool tends to be more basic, if compared with the average pI of the 20 kDa pool, so that the increase in MW appears to be correlated to a corresponding increase in pI.

Interestingly, spots that are positively correlated in one group tend to be negatively correlated with the spots of the other group, so that the two identified pools are mutually exclusive. This fact could be due to at least one kind of post-translational modification of DJ-1, changing the mass of the protein, which acts like an alternative switch. The functional meaning of this switch will be investigated in future studies.

3.2 DJ-1 patterns in 2D-gel experiments are reminiscent of the tissutal origin of the sample

To ascertain whether the identified pools have any biological relevance, all the experiments (i.e. all the 2-DE patterns) were grouped according to their similarity in terms of Euclidean distances between the corresponding rows in the 'spot matrix', and then the experiments were aggregated on the basis of the described procedure. The resulting dendrogram is reported in Figure 3. Most of the experiments segregates into two groups: one contains mainly

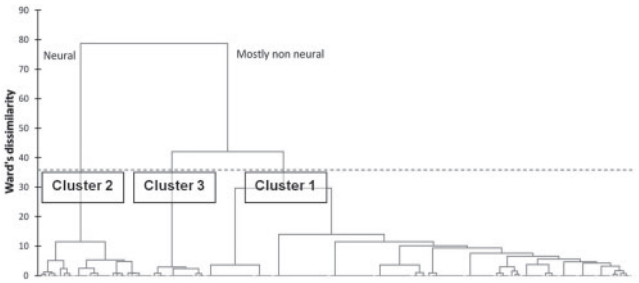

Fig. 3. Cluster analysis of the considered 2D experiments, based on the presence/absence of spots shown in Figure 1. Three clusters are formed according to the method explained in the text; cluster 2 contains only experiments on DJ-1 of neural origin.

2-DE gels from 'human, brain' samples (cluster 2, including neural cell lines), while the other contains experiments of heterogeneous, mostly non-neural and/or non-human origin (cluster 1). Moreover, there is a separate 'non-neural' cluster (cluster 3), which contains experiments coming from a single paper, with an overall DJ-1 pattern not supported by any other paper and consequently suspicious of experimental biases. For this reason, we did not consider this cluster further. Interestingly, the 'human, neural' cluster 2 contains only three samples originating from cellular lines (dopaminergic neuroblastoma M17 line), while all the other 23 experiments in this cluster are 2-DE conducted on proteins from human brain.

Once the experiments were grouped, it was possible to verify whether the clusters formed on the basis of the similarity between the images could be supported also by some non-biological bias (such as the paper of origin of the images, the authors of the experiments and the particular protocol used). The latter hypothesis would imply either that the results are strongly dependent on the particular author presenting them (i.e. they are not reproduced by others) or they are extremely sensitive to the experimental conditions used (i.e. they are not reproducible). Beside the aforementioned bias in cluster 3, both clusters 1 and 2 contain experiments from different papers, with different staining procedures, so that neither of these bias factors could explain the different DJ-1 patterns responsible for the observed clustering.

Therefore, it is possible to distinguish, at least in humans, the brain origin of DJ-1 on the basis of its 2-DE pattern, i.e. there is a specific experimental pattern associated with DJ-1 of human brain origin.

3.3 The pool containing lighter, more acidic forms of DJ-1 is prevalent in neural tissues

We further proceeded to verify whether the 2-DE gel patterns responsible for clusterization in Figure 3 corresponded to the pools of correlated DJ-1 forms identified in Figure 2. Interestingly, all the experiments included in cluster 2 contained spots from correlation group 2, characterized by a MW $\approx$ 20 kDa and a more acid average pI (see Fig. 2B); moreover, none of these experiments contained any spot from correlation group 1 (MW $\approx$ 25 kDa). Conversely, all the experiments included in cluster 1 contained spots from correlation group 1, which has a MW $\approx$ 25 kDa and does not contain any spot of MW $\approx$ 20 kDa. Since cluster 2 groups together with experiments on human brain samples, there is a pool of DJ-1 isoforms, identified in Figure 2B as group 2, that is typical of neural samples and particularly of DJ-1 from brain. Thus, it appears that the post-translational modification responsible for the 25 kDa forms is at best only marginally active in most of the samples of human brain origin.

3.4 Meta-analysis of DJ-1 modifications induced by neurodegeneration and aging

To check whether any described modification of the DJ-1 pool associated with neurodegeneration could be supported by our meta-analysis, a metagel obtained by all documented experiments on control human brains (12 experiments) was compared with the corresponding metagel obtained from patients with neurodegenerative conditions (11 experiments) (Figure 4A and B, respectively). According to the source literature, these experiments are representative of the DJ-1 status in 53 control brains

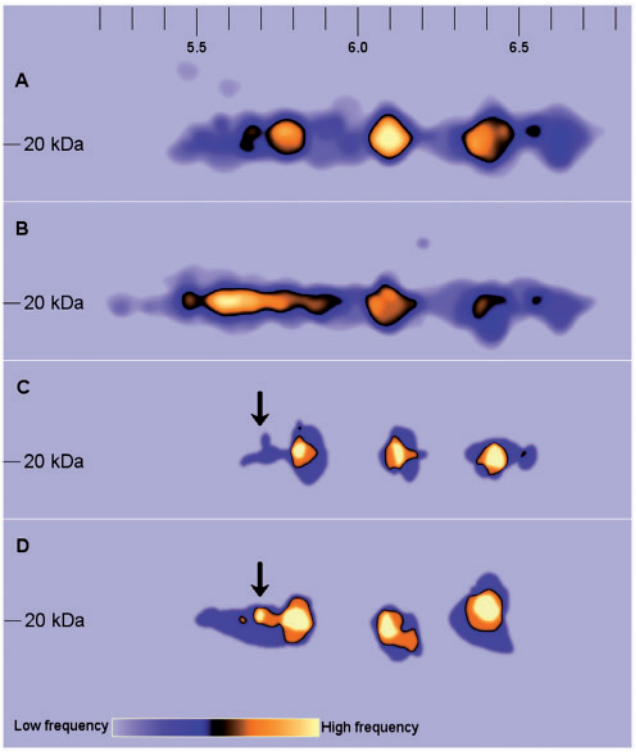

Fig. 4. Analysis of DJ-1 forms in human brain. (**A** and **B**) Represent metagel obtained for control and neurodegenerative individuals as documented in the considered literature. (**C** and **D**) Represent the metagel obtained from experiments on three control individuals and three PD patients as described in the text. Both the literature-derived and the experimentally derived metagel support the increase in DJ-1 acidic forms for neurodegenerative conditions.

(33 healthy, 20 astrocytoma) and 70 brains from individuals showing neurodegenerative condition (30 PD, 11 Alzheimer's disease, 13 progressive supranuclear palsy, 4 corticobasal degeneration and 12 frontotemporal lobar degeneration). Human brain tissues display a subset of DJ-1 forms that correlate with the occurrence of neurodegenerative processes, including PD. In particular, the frequency of acid forms increases, while the frequency of more basic forms decreases. Noteworthy, if a single individual is considered, no specific form of DJ-1 could be connected to any neurodegenerative condition. On the contrary, the data reported herein are referred to the overall patient population, considered in the source literature.

These findings are expected on the basis of a specific accumulation of DJ-1 acidic forms in the brain of Parkinson's and Alzheimer's disease patients, mentioned originally by Bandopadhyay and coworkers (Bandopadhyay *et al.*, 2004). Indeed, 10 different DJ-1 forms, including the acidic forms (pI 5.5 and 5.7) of 20 kDa DJ-1 were observed to be selectively accumulated in frontal cortex of patients compared with age-matched controls (Choi *et al.*, 2006).

In order to take into account the role of aging, we considered a couple of DJ-1 patterns from the brains of two old subjects (age = 92 and 98, respectively) (Meulener *et al.*, 2006). Also in this case, the literature hypothesis is confirmed, in that both subjects possess a DJ-1 pool similar to that of PD patients, thus suggesting

an accumulation of oxidized, acidic forms of the protein, which is related to aging, similar to that observed in PD patients.

3.5 Occurrence of specific DJ-1 forms associated with neurodegeneration is confirmed by western blot analysis of post-mortem specimens of human substantia nigra

The meta-analysis so far described led us to predict that: (i) the pool of DJ-1 forms in brain tends to have a weight of ∼20 kDa; (ii) PD patients are expected to have an unbalance toward more acidic forms of the protein, as previously found in literature.

To experimentally verify this prediction, anti-DJ-1 immunoblots for three controls and three PD patients were performed. Among them, two were clinically and histopathologically negative for PD diagnosis, whereas one of the control subjects showed histological evidence of mild degeneration although he was clinically healthy (CO280). Three patients were included in the PD group after careful evaluation of clinical history and neuropathological report.

Two metagels were built for control cases and PD subjects (Figure 4C and D, respectively). Only 20 kDa species were detected in both metagels, confirming the predominance of this form in the brain. For all extracts, three main DJ-1 forms are observed at pI = 5.8, 6.1 and 6.4. Despite the limited number of cases, due to difficulty in obtaining human post-mortem samples of substantia nigra, even in the small population considered, there is some tendency toward more acidic forms into the PD subjects (arrows in Figure 4C and D).

A confounding factor is due to the inclusion in the control group of CO280. The DJ-1 pattern of this subject is different from other controls, showing the same acidic forms observed in PD subjects, and rendering the control metagel somehow more similar to the PD metagel. Remarkably, this 'control' appears positive for the presence in its brain of tangles (Table 1), a typical hallmark of oxidative stress and damage in brain, although Lewy bodies were not observed. DJ-1 pool modification in this subject might thus be related to some form of brain damage, without clinical symptoms.

More specimens would be needed to confirm the outlined picture of the modifications of DJ-1 under PD, nevertheless it should be noted here that the reported pattern is sustained by a number of publications in a larger population, as previously outlined. Although we could imagine that in SNpc pf PD patients dopaminergic neurons are mainly lost, DJ-1 oxidative modifications witness the occurrence of oxidative stress conditions in the affected area, either it comes from residual neurons or from astrocytes.

4 DISCUSSION AND CONCLUSIONS

The meta-analysis of proteomic data can provide valuable information pertinent to various biological processes (or methods involved) that otherwise remains hidden if a single paper is considered. The occurrence of frequently identified proteins, for instance, may represent common cellular stress responses or simply reflect technical limitations of 2-DE, as recently stated by Petrak and coworkers in their visionary paper on déjà-vu in proteomic reports (Petrak *et al.*, 2008). We developed therefore an innovative procedure to examine the status of DJ-1, a protein extensively and heterogeneously modified, by comparing its 2-DE pattern documented in 160 experiments from 37 different papers. Despite the presence of many different forms of the protein, with different pI and MW, it was found that DJ-1 main forms consistently associated with two different pools, with a different MW and average pI. These two pools are mutually exclusive, and the pool with a MW ≈ 20 kDa and a more acid average pI exhibits strong tissue specificity for brain and neural models. Moreover, we could show that this pool is indeed found experimentally in the SNpc of both PD patients and control subjects and, in agreement with the previous literature, there is a tendency at the level of population toward more acid forms in PD patients.

Altogether, these findings support the value of a proteomic meta-analysis for the unraveling of complex proteomic signatures and biomarker analysis.

ACKNOWLEDGEMENTS

Authors wish to thank Drs Chiara Abrescia and Enzo Medico for their helpful discussion. The Internationale Stichting Parkinson Onderzoek, who subsidizes the NBB for collecting brains of PD donors, is gratefully acknowledged.

Funding: Valle d'Aosta Regional Government (http://www.regione.vda.it/) in the frame of the regional law n.84-07/12/1993 (grant 30Apr09 - project IMAGE - Image Meta Analysis Generation and Exploitation 2009/2010 partially).

Conflict of Interest: none declared.

REFERENCES

Bandopadhyay,R. *et al.* (2004) The expression of DJ-1 (PARK7) in normal human CNS and idiopathic Parkinson's disease. *Brain*, **127**, 420–430.

Batelli,S. *et al.* (2008) DJ-1 modulates alpha-synuclein aggregation state in a cellular model of oxidative stress: relevance for Parkinson's disease and involvement of HSP70. *PLoS One*, **23**, e1884.

Bieler,S. *et al.* (2009) Comprehensive proteomic and transcriptomic analysis reveals early induction of a protective anti-oxidative stress response by low-dose proteasome inhibition. *Proteomics*, **9**, 3257–3267.

Bonifati,V. *et al.* (2003) Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science*, **299**, 256–259.

Braak *et al.* (2003) Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging*, **24**, 197–211.

Canet-Avilés,R.M. *et al.* (2004) The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfinic acid-driven mitochondrial localization. *Proc. Natl Acad. Sci. USA*, **101**, 9103–9108.

Choi,J. *et al.* (2006) Oxidative damage of DJ-1 is linked to sporadic Parkinson and Alzheimer diseases. *J. Biol. Chem.*, **281**, 10816–10824.

Colapinto,M. *et al.* (2006) alpha-Synuclein protects SH-SY5Y cells from dopamine toxicity. *Biochem. Biophys. Res. Commun.*, **349**, 1294–1300.

Fasano,M. *et al.* (2008) Proteomics as a tool to investigate cell models for dopamine toxicity. *Parkinsonism Relat. Disord.*, **14**(Suppl. 2), S135–S138.

Girish,V. and Vijayalakshmi,A. (2004) Affordable image analysis using NIH Image/ImageJ. *Indian J. Cancer*, **41**, 47.

Hod,Y. *et al.* (1999) Identification and characterization of a novel protein that regulates RNA-protein interaction. *J. Cell Biochem.*, **72**, 435–444.

Ito,G. *et al.* (2007) Roles of distinct cysteine residues in S-nitrosylation and dimerization of DJ-1. *Biochem. Biophys. Res. Commun.*, **339**, 667–672.

Jacobs *et al.* (2001) Sequential solubilization of proteins precipitated with trichloroacetic acid in acetone from cultured Catharanthus roseus cells yields 52% more spots after two-dimensional electrophoresis. *Proteomics*, **1**, 1345–1350.

Junn,E. *et al.* (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.

Kinumi,T. *et al.* (2004) Cysteine-106 of DJ-1 is the most sensitive cysteine residue to hydrogen peroxide-mediated oxidation in vivo in human umbilical vein endothelial cells. *Biochem. Biophys. Res. Commun.*, **317**, 722–728.

Lev,N. *et al.* (2008) Oxidative insults induce DJ-1 upregulation and redistribution: implications for neuroprotection. *Neurotoxicology*, **29**, 397–405.

Lev,N. *et al.* (2009) DJ-1 protects against dopamine toxicity. *J. Neural. Transm.*, **116**, 151–160.

Li,H.M. *et al.* (2005) Association of DJ-1 with chaperones and enhanced association and colocalization with mitochondrial Hsp70 by oxidative stress. *Free Radic. Res.*, **39**, 1091–1099.

Meulener,M.C. *et al.* (2005) DJ-1 is present in a large molecular complex in human brain tissue and interacts with alpha-synuclein. *J. Neurochem.*, **93**, 1524–1532.

Meulener,M.C. *et al.* (2006) Mutational analysis of DJ-1 in Drosophila implicates functional inactivation by oxidative damage and aging. *Proc. Natl Acad. Sci. USA*, **103**, 12517–12522.

Mitsumoto,A. *et al.* (2001) Oxidized forms of peroxiredoxins and DJ-1 on two-dimensional gels increased in response to sublethal levels of paraquat. *Free Radic Res.*, **35**, 301–310.

Moore,D.J. *et al.* (2005) Association of DJ-1 and parkin mediated by pathogenic DJ-1 mutations and oxidative stress. *Hum. Mol. Genet.*, **14**, 71–84.

Nagakubo,D. *et al.* (1997) DJ-1, a novel oncogene which transforms mouse NIH3T3 cells in cooperation with ras. *Biochem. Biophys. Res. Commun.*, **231**, 509–513.

Niki,T. *et al.* (2003) DJBP: a novel DJ-1-binding protein, negatively regulates the androgen receptor by recruiting histone deacetylase complex, and DJ-1 antagonizes this inhibition by abrogation of this complex. *Mol. Cancer Res.*, **1**, 247–261.

Olzmann,J.A. *et al.* (2004) Familial Parkinson's disease-associated L166P mutation disrupts DJ-1 protein folding and function. *J. Biol. Chem.*, **279**, 8506–8515.

Olzmann,J.A. *et al.* (2007) Parkin-mediated K63-linked polyubiquitination targets misfolded DJ-1 to aggresomes via binding to HDAC6. *J. Cell. Biol.*, **178,** 1025–1038.

Petrak,J. *et al.* (2008) Déjà vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins. *Proteomics*, **8**, 1744–1749.

Pleissner,K.P. *et al.* (1999) New algorithmic approaches to protein spot detection and pattern matching in two-dimensional electrophoresis gel databases. *Electrophoresis*, **20**, 755–765.

Ravid,R. *et al.* (1998) Brain banking in aging and dementia research - The Amsterdam experience. In Fisher,A., Yorshida,M., Hanin,I. (eds) *Progress in Alzheimer's and Parkinson's disease*. Plenum Press, New York, pp. 277–286.

Sekito,A. *et al.* (2005) Stimulation of transforming activity of DJ-1 by Abstrakt, a DJ-1-binding protein. *Int. J. Oncol.*, **26**, 685–689.

Shinbo,Y. *et al.* (2006) Proper SUMO-1 conjugation is essential to DJ-1 to exert its full activities. *Cell Death Differ.*, **13**, 96–108.

Takahashi,K. *et al.* (2001) DJ-1 positively regulates the androgen receptor by impairing the binding of PIASx alpha to the receptor. *J. Biol. Chem.*, **276**, 37556–37563.

Tao,X. and Tong,L. (2003) Crystal structure of human DJ-1, a protein associated with early onset Parkinson's disease. *J. Biol. Chem.*, **278**, 31372–31379.

Tillman,J.E. *et al.* (2007) DJ-1 binds androgen receptor directly and mediates its activity in hormonally treated prostate cancer cells. *Cancer Res.*, **67**, 4630–4637.

van der Brug,M.P. *et al.* (2008) RNA binding activity of the recessive parkinsonism protein DJ-1 supports involvement in multiple cellular pathways. *Proc. Natl Acad. Sci. USA*, **105**, 10244–10249.

Van Laar,V.S. *et al.* (2009) Proteomic identification of dopamine-conjugated proteins from isolated rat brain mitochondria and SH-SY5Y cells. *Neurobiol. Dis.*, **34**, 487–500.

Waak,J. *et al.* (2009) Oxidizable residues mediating protein stability and cytoprotective interaction of DJ-1 with apoptosis signal-regulating kinase 1. *J. Biol. Chem.*, **284**, 14245–14257.

Ward,J.H. Jr (1963) Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.*, **58**, 236–244.

Xiong,H. *et al.* (2009) Parkin, PINK1, and DJ-1 form a ubiquitin E3 ligase complex promoting unfolded protein degradation. *J. Clin. Invest.*, **119**, 650–660.

Xu,J. *et al.* (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.

Yanagida,T. *et al.* (2006) Distribution of DJ-1, Parkinson's disease-related protein PARK7, and its alteration in 6-hydroxydopamine-treated hemiparkinsonian rat brain. *J. Pharmacol. Sci.*, **102**, 243–247.

Zhang,L. *et al.* (2005) Mitochondrial localization of the Parkinson's disease related protein DJ-1: implications for pathogenesis. *Hum. Mol. Genet.*, **14**, 2063–2073.

Zhong,N. and Xu,J. (2008) Synergistic activation of the human MnSOD promoter by DJ-1 and PGC-1alpha: regulation by SUMOylation and oxidation. *Hum. Mol. Genet.*, **17**, 3357–3367.

Zhou and Freed (2005) DJ-1 up-regulates glutathione synthesis during oxidative stress and inhibits A53T alpha-synuclein toxicity. *J. Biol. Chem.*, **280**, 43150–43158.