

Towards a pathway definition of Parkinson's disease: a complex disorder with links to cancer, diabetes and inflammation

Linda B. Moran · Manuel B. Graeber

Received: 5 December 2007 / Accepted: 12 December 2007 / Published online: 12 January 2008
© Springer-Verlag 2007

Abstract We have previously established a first whole genome transcriptomic profile of sporadic Parkinson's disease (PD). After extensive brain tissue-based validation combined with cycles of iterative data analysis and by focusing on the most comparable cases of the cohort, we have refined our analysis and established a list of 892 highly dysregulated priority genes that are considered to form the core of the diseased Parkinsonian metabolic network. The substantia nigra pathways, now under scrutiny, contain more than 100 genes whose association with PD is known from the literature. Of those, more than 40 genes belong to the highly significantly dysregulated group identified in our dataset. Apart from the complete list of 892 priority genes, we present pathways revealing PD 'hub' as well as 'peripheral' network genes. The latter include Lewy body components or interact with known PD genes. Biological associations of PD with cancer, diabetes and inflammation are discussed and interactions of the priority genes with several drugs are provided. Our study illustrates the value of rigorous clinico-pathological correlation when analysing high-throughput data to make optimal use of the histopathological phenome, or morphonome which currently serves as the key diagnostic reference for most human diseases. The need for systematic

human tissue banking, following the highest possible professional and ethical standard to enable sustainability, becomes evident.

Keywords Aposklesis · Expression analysis · Lewy bodies · Microarrays · Network medicine · Neurodegeneration · Synaptic dysfunction

Introduction

Our understanding of Parkinson's disease (PD) is largely incomplete. However, the pace of discovery in this field is rapidly accelerating. It took more than 100 years for the key region of neuronal damage, the substantia nigra to be identified [1], and it took almost 80 years for the first disease-causing mutation to be discovered [2]. Not even 10 years later, the first whole genome transcriptome analysis had been performed [3], and a number of other microarray studies focusing on known sequences were carried out (e.g. [4–6]). We now provide the complete list of 892 highly dysregulated PD nigral genes derived from a brain tissue-validated whole genome expression microarray data set. In addition, predicted interactions of a number of these genes are reported as potential drug targets. We would like to emphasise that the neurohistological validation that is so crucial for our work and which has already led to the identification of two novel Lewy body components predicted on the basis of this dataset [7, 8] could not have been performed without generous brain donations. In addition, the iterative analysis performed combining histological phenome (morphonome) data and clinical criteria within silico data mining would not have been possible without significant advances in computing, notably virtual machine technology. It is readily apparent that a publication of this

An internet connection is required to access the complete contents of this paper. A list of hyperlinks to the electronic supplemental material (ESM) of this article is provided at the end of the text.

L. B. Moran · M. B. Graeber (✉)
University Department of Neuropathology, Imperial College,
University of London, and Hammersmith Hospitals Trust,
Charing Cross Campus,
Fulham Palace Road,
London W6 8RF, UK
e-mail: manuel@graeber.net

format requires the Internet as it would not have been possible to publish its almost 3,200 hyperlinked files on paper, which are provided as electronic supplemental material.

PD is a severely disabling neurodegenerative disorder second in frequency only to Alzheimer's disease and has a significant socio-economic impact. Unlike in Alzheimer's disease, however, the brain region taking the brunt of the disease process is rather well circumscribed. In addition, there is widespread consensus on diagnostic criteria both clinically and neuropathologically (<http://www.ICDNS.org>). This is at least in part due to the fact that the leading motor symptoms are less complex and easier to recognize and define than the clinical signs in disorders mainly affecting higher brain functions such as cognition. Furthermore, there is symptomatic treatment for PD pointing to key pathways involved. All these are important prerequisites when working with high throughput technologies such as microarrays which require precise tissue sampling because the procedures employed are both laborious and expensive.

Major known pathways involved in PD include the ubiquitin-proteasome system dysfunction of which may lead to abnormal protein deposition, mitochondrial failure and decreased expression of synaptic proteins [6, 9–11]. Oxidative stress has been traditionally implicated in the aetiology of the disease but the changes observed could be secondary. The concept of 'neuroinflammation' has become very popular recently [12–14], but our own work in this field does not currently support a role for microglia as a driver of the disease process [15]. The effective failure of recent studies employing non-steroidal anti-inflammatory drugs supports this notion [16, 17]. Thus, there are leads and popular ideas but the big picture of PD pathogenesis is still missing. A true understanding of PD and its subtypes will require integrated knowledge from several system biological levels, ranging from genomics to proteomics and metabolomics as well as clinical data and neuroimaging. Through this study, we aim to contribute a validated transcriptomic data layer.

Materials and methods

Data set used

The 94.CEL files used for this study have been deposited at the National Center for Biotechnology Information, Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/projects/geo>) with GEO Series accession number GSE8397 (scheduled release date: 1 January 2008). The patient samples employed have been described previously [3]. This data set is based on Affymetrix HU_133A and HU_133B gene chips set and has been extensively

validated over a period of 2 years using qRT-PCR, immunocytochemistry and in situ hybridisation to cellularly 'back-map' sequences of interest [7, 8, 10, 18].

Data analysis procedures employed

In silico analyses were performed with the help of programme packages from different suppliers. Our initial microarray data analysis was performed using ArrayAssist 3.0 (Stratagene) but there were intermittent problems with some versions of the software (3.2–3.4). We have repeated our analysis using newer (4.0, 4.1; Linux and/or Windows) and the latest version of this program (ArrayAssist 5.5 for Windows, Stratagene). In addition, we have reproduced our results using an independent software package for microarray analysis, PathwayStudio 5.0. The (GC-)RMA algorithm which has become a gold standard for Affymetrix microarray normalisation was applied in all cases. In addition, PathwayAssist (Ariadne) and Pathway Architect (Stratagene) were employed during phases of our work. For ease of use, most software installations were performed in virtual machines (VMware Workstation and Fusion, respectively; <http://www.vmware.com>) and two virtual machines were frequently run in parallel on Windows, Linux, or Macintosh platforms. The ability to create virtual machine 'snap shots' of critical stages of the analysis proved invaluable for the backtracking of results and to allow comparability over time. It proved essential in the case of the most complex network analyses where software stability proved to be a factor.

We have used the microarray data in a hypothesis generating rather than hypothesis testing way (cf. [4, 10]). Since our original predictions based on the ArrayAssist 3.0 dataset turned out to be very reliable with respect to subsequent in situ tissue validation results, we have performed our refined analysis by means of this programme following extensive comparison with readings generated by the latest versions of ArrayAssist 5.5 and PathwayStudio 5.0. Anyone intending to reproduce our results is referred to the original CEL files (GEO ID GSE8397).

The following original cases were excluded from our refined analysis to obtain more homogeneous cohorts taking newly obtained histological and expression data into account: Con4, 8 and MS155 as well as the sample of medial substantia nigra from PD22. Thus, nine control nigra and 23 PD nigra samples remained in the study. To find differentially expressed genes, the *p* value cut off was kept at 0.001 (differential expression=1 log₂). Multiple testing corrections (FDR, Bonferroni) were carried out for comparative purposes. The 892 top genes identified on the basis of 1,145 probes are referred to as the 'priority genes'.

Hierarchical clustering was executed on both rows and columns using ArrayAssist 5.5 (Pearson centred distance

metric, centroid linkage rule). Similarity images were produced to visualise the quality of clustering results (SI_Figure_1). In addition, self-organising map clustering was performed on both rows and columns using a Euclidean distance metric in ArrayAssist 5.5 (maximum number of iterations 50, number of grid rows 3, number of grid columns 4, initial learning rate 0.03, initial neighbourhood radius 5, grid topology hexagonal, neighbourhood type bubble).

For finding cell processes regulated by the differentially expressed genes, PathwayStudio software (version 5.0) was used. The 164 top up-regulated priority genes showing a differential expression $>1 \log_2$ were selected and the *find common targets algorithm* was employed in the *build pathway tool* setting the *cell process filter option*. The procedure was repeated for *diseases entity type*. Cell processes and disease conditions showing the highest number of biological associations, i.e. the strongest probabilistic relationship based on literature evidence to the group of 164 top up-regulated priority genes were selected. Furthermore, known interactions between all 892 dysregulated genes were identified using the ResNet 5.0 database of molecular interactions which has been derived from the published literature by means of a natural language processing technology called Medscan [19]. A filter for the parameters, promoter binding, expression and regulation was applied in the latter case.

For additional validation, the commercial PathArt database (PathwayAssist plug-in, Jubilant Biosys) containing a large number of manually curated pathways was queried. A number of GEO datasets (e.g. astrocytes and microglia in culture) were downloaded to evaluate cell-type specificity of expression of individual priority genes. A search for novel secreted biomarkers was also performed of which NPTX2 [8] may be one.

To retrieve known drug interactions of the 892 priority genes, more than 9,000,000 database objects were checked in ResNet 5.0 (PathwayStudio 5.0). Subsequently, all known interactions of individual relevant drugs contained in ResNet 5.0 were retrieved and the results overlaid with the set of priority genes. Drugs scoring high in terms of the number of their known interactions with the 892 priority genes were identified and the relationship to PD analysed employing the PubMed database (<http://www.ncbi.nlm.nih.gov/sites/entrez>).

Results

The extended set of 892 PD priority genes

Screening of histological and clinical parameters combined with iterative cycles of data mining enabled us to identify

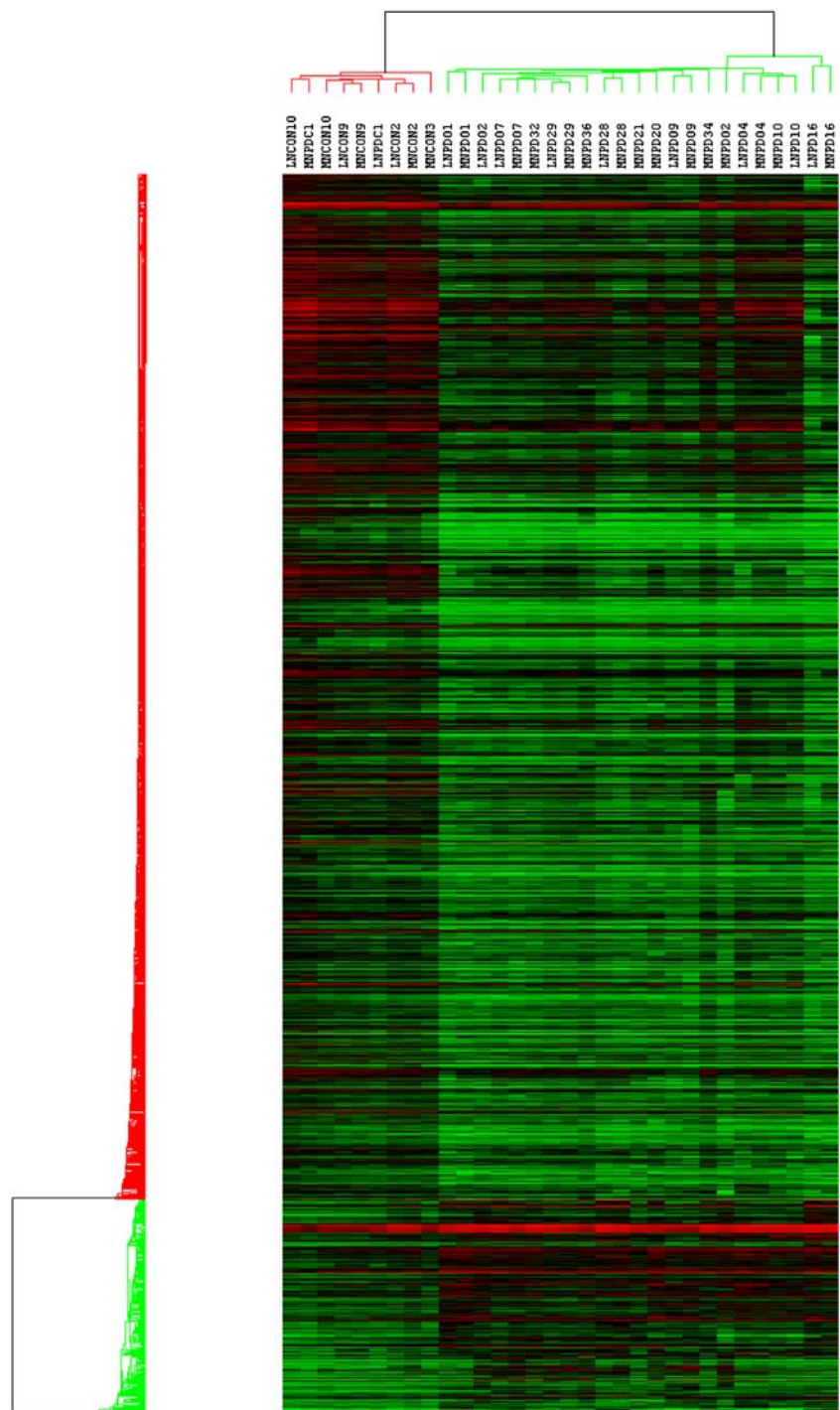
the most comparable cases in the cohorts which are detailed in SI_Figure_1. The number of highly significantly dysregulated genes ($p < 0.001$) increased from 429 (570 probes) to 892 (1,145 probes) while 97% of the originally identified genes [3] were confirmed. The complete list of all 892 PD priority genes is provided in SI_Table_1a. It should be mentioned that there are still some sequences that show highly significant dysregulation in PD on which there is no information available from ResNet 5.0, the IDconverter database [20] or NetAffx. However, all 1,145 probes (SI_Tables_1b&c) were used for the clustering (Fig. 1 and SI_Figure_1). Information on the 1,145 Affymetrix probes available from the public IDconverter database is provided in SI_Table_1d. SI_Table_1e shows the probe information not found in IDConverter or the ResNet database, respectively, that was obtained via NetAffx (<http://www.affymetrix.com/analysis/index.affx>). Thirty-one priority genes have been previously designated PD candidate genes (<http://www.pdgene.org/>) (SI_Table_1f), and it is worth noting that 33 are considered AD candidate genes as well (SI_Table_1g) with HSPA1A, HSPA1B, HTR2A, KCNJ6, SLC6A3, SNCA, SNCG, TF and UCHL1 being present in both lists. A PubMed search revealed that another 10 priority genes have known functional links to PD (SI_Table_1h). Finally, ‘whole genome/whole human body’ clustering using an independent data set derived from 64 different organ samples which is based on Affymetrix Human Genome U133 Plus 2.0 arrays allowed categorical separation of all nervous tissue samples including three different ganglia (SI_Figure_2).

Gene ontology and pathway analysis of the differentially expressed genes

The statistically most significant overlap (Fisher Exact test) between the 892 priority genes and GO groups was found for GO ID 0016020 (membrane) and GO IDs 0005515/0045308 (protein binding or protein degradation tagging activity, respectively). Further details are provided in the supplemental material (SI_Table_2). The main cell processes predicted to be influenced by the top up-regulated priority genes (164 genes showing differential expression $>1 \log_2$) are illustrated in Fig. 2a. A hyperlinked version of this figure providing details on all genes and their interactions is available online (SI_Figure_3a). A search of 192 canonical pathways and 555 signalling pathways in PathwayStudio 5.0 yielded the RET_HSF1 signalling pathway which shares six priority genes as the highest ranking result (SI_Table_3, SI_Pathway_1).

Predicted interactions of a subset of the 892 priority genes are shown in Fig. 3. A total of 417 known connections were retrieved from ResNet. This figure is

Fig. 1 Dendrogram showing the results of unsupervised hierarchical clustering of the 1,145 probes encompassing the 892 priority genes (corresponding Affymetrix probe IDs and expression data are provided online in SI_Figure_1). The algorithm separates the control (*left*) from the Parkinson's disease group (*right*) (rows: *red*, expression above average; *black*, average expression; *green*, below average). An explanation of sample designations, patient data and a quantitative colour code are also provided in SI_Figure_1



provided for orientation purposes. Hyperlinked permutations of this figure showing details of all genes and their interactions are available online (SI_Figure_4).

Known components of Lewy bodies and how they relate to the group of priority genes are presented in Fig. 4. The probes for most of these components were found to cluster together in the self-organising map shown in SI_Figure_5. A hyperlinked version of Fig. 4 with expression values

overlaid is available online (SI_Figure_6). Interactions of known PD genes with the priority genes of this study are demonstrated in Fig. 5 (SI_Figure_7).

We did not observe an effect for gender (cf. [21]). Hierarchical clustering of the male and female PD patients on the basis of the expression values of the 892 priority genes did not separate the groups nor did a whole genome clustering omitting sex chromosomal sequences.

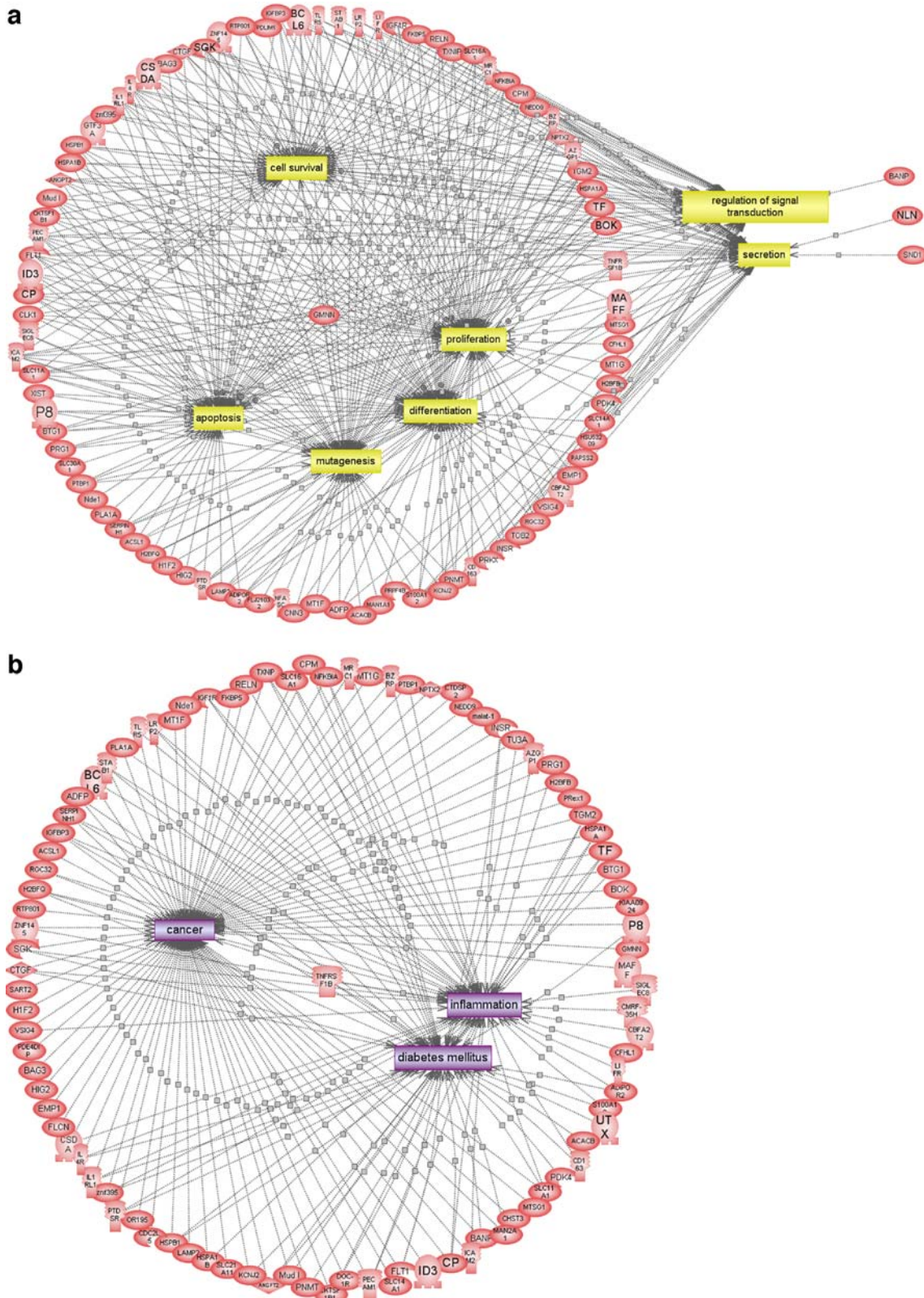


Fig. 2 a Cell processes predicted to be influenced by the top up-regulated priority genes (164 genes showing a differential expression $>1 \log_2$) based on *in silico* analysis employing the ResNet database (PathwayStudio 5.0, Ariadne). *P* and expression values of all 1,145 probes can be found in SI_Table_1c. An online version of this figure with hyperlinks is provided as SI_Figure_3a. **b** The three disease

conditions showing the strongest biological association with the group of top up-regulated priority genes which may serve as drivers of the disease process underlying PD. *In silico* analysis was performed employing the ResNet database of molecular interactions (PathwayStudio 5.0, Ariadne). An online version of this figure with hyperlinks is provided as SI_Figure_3b

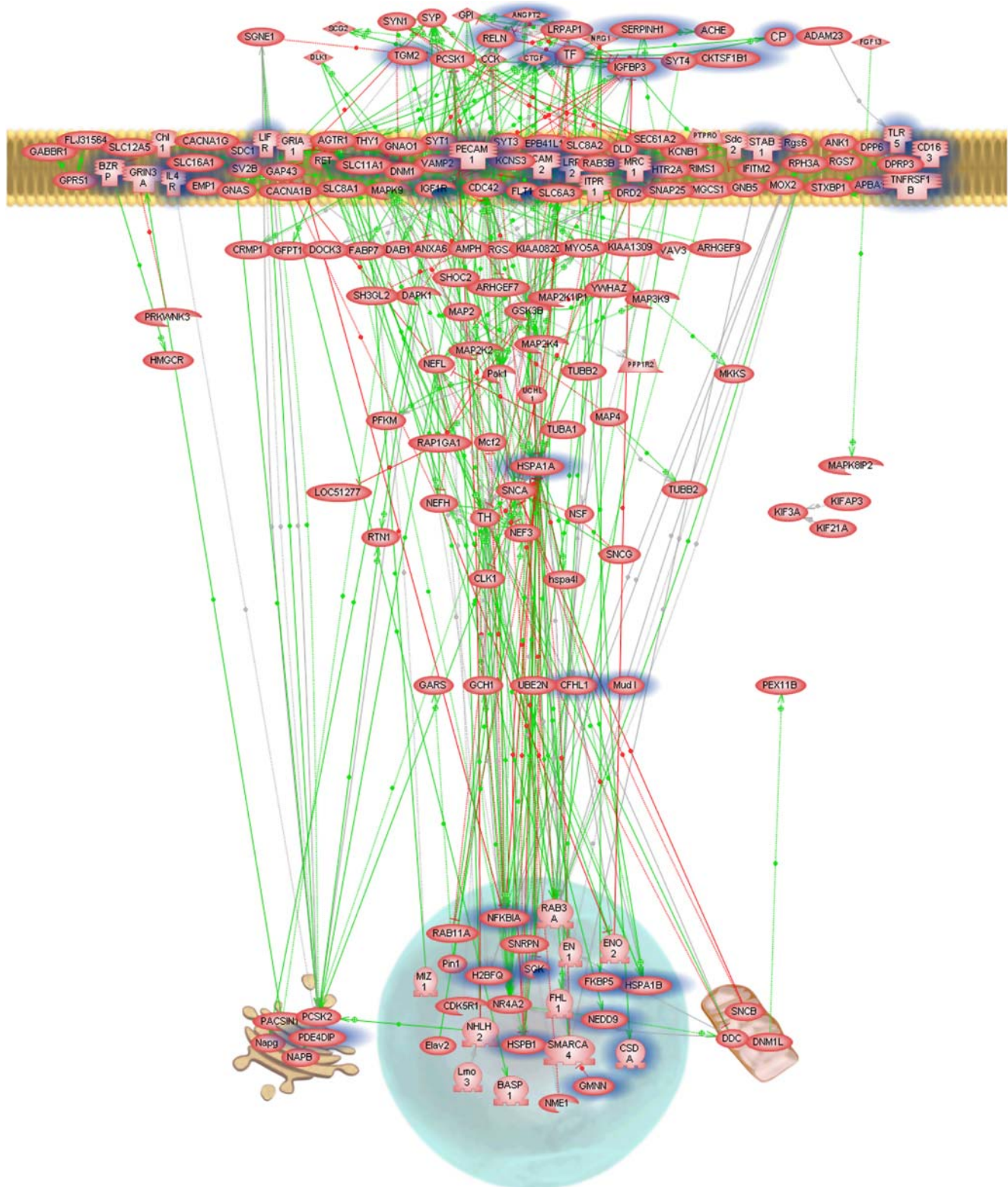


Fig. 3 Hypothetical ‘super pathway’ not stratified for cell type illustrating known direct interactions between the 892 PD priority genes (regulation, expression and promoter binding only). A total of 417 interactions (relations) are shown and any unlinked entities were removed. Display style: by effect; cellular layout; colour codes: promoter binding, violet; green, positive regulation; red, negative

regulation; grey and/or broken lines, unknown (Resnet 5.0 database, unedited). Blue shading around selected genes indicates their involvement in the cellular process and disease conditions depicted in Fig. 2a and b, respectively. Permutations of this figure with hyperlinks are provided online as SI_Figures_4a-c

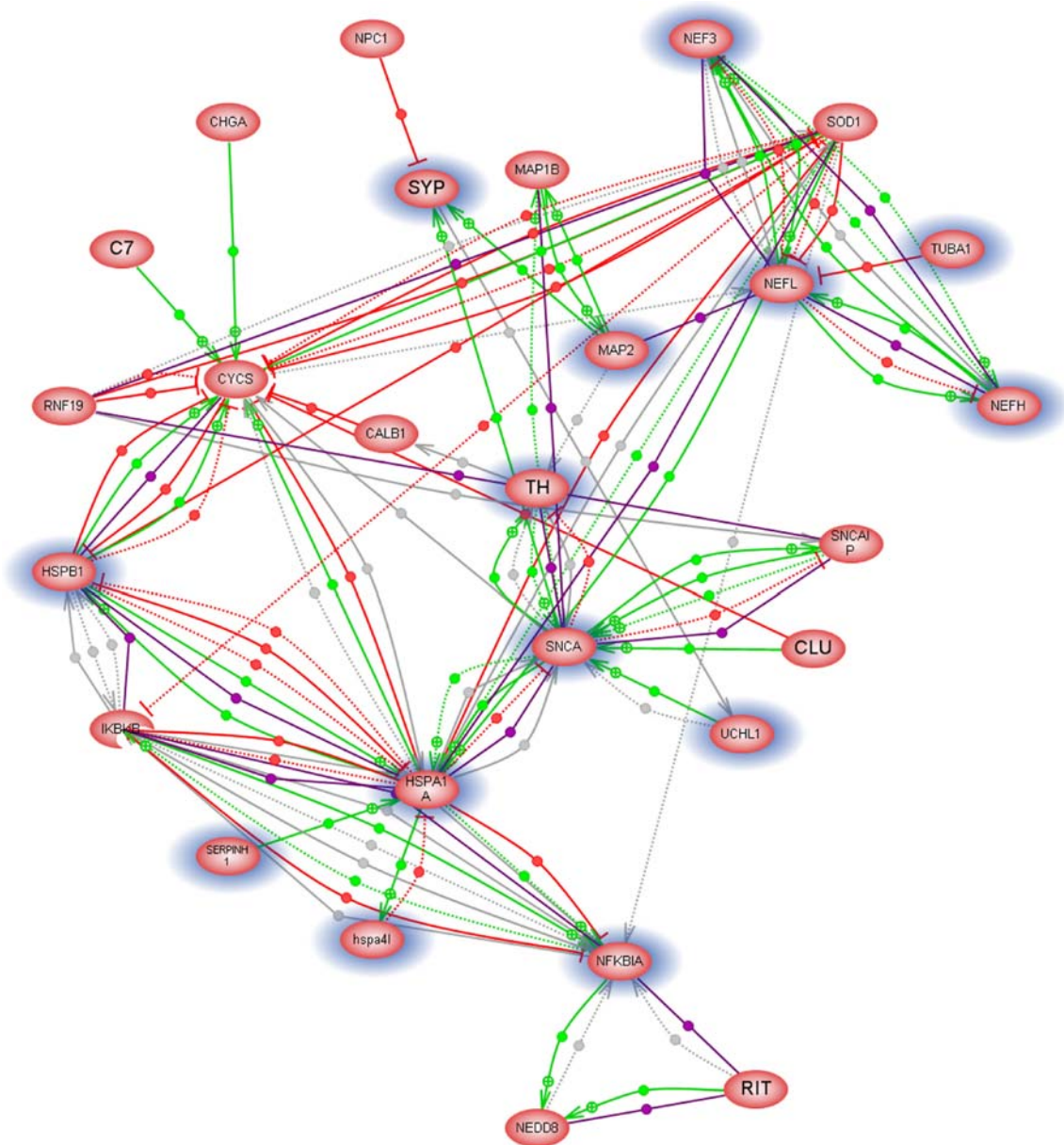


Fig. 4 Neuronal pathway containing proteins found in Lewy bodies [11]. Priority genes of this study are marked by the blue shading. An online version of this figure with hyperlinks and an overlay of

expression values is provided as SI_Figure_6. Symmetrical layout, display by effect (PathwayStudio)

Relationship to diseases and drug interactions

A search of the ResNet database identified three disease conditions that showed the strongest probabilistic relationship to the group of top up-regulated priority genes, cancer, diabetes and inflammation as illustrated in Fig. 2b. A hyperlinked version of this figure providing details on all genes and their interactions is available online (SI_Figure_3b). Known drug interactions of some of the priority genes were retrieved from the ResNet 5.0 database through checking of more

than 9,000,000 database objects. It is noteworthy that drugs such as clozapine, cocaine and haloperidol, which are used in the treatment of PD or which cause Parkinsonian side effects, appear to interact with a large number of PD priority genes (SI_Figure_8). A search of the ResNet database also yielded information on the interactions of two cytostatic drugs, paclitaxel and vincristine with the priority genes identified in this study (SI_Pathways_2&3). Both paclitaxel and vincristine have been reported to induce parkinsonian side effects [22, 23].

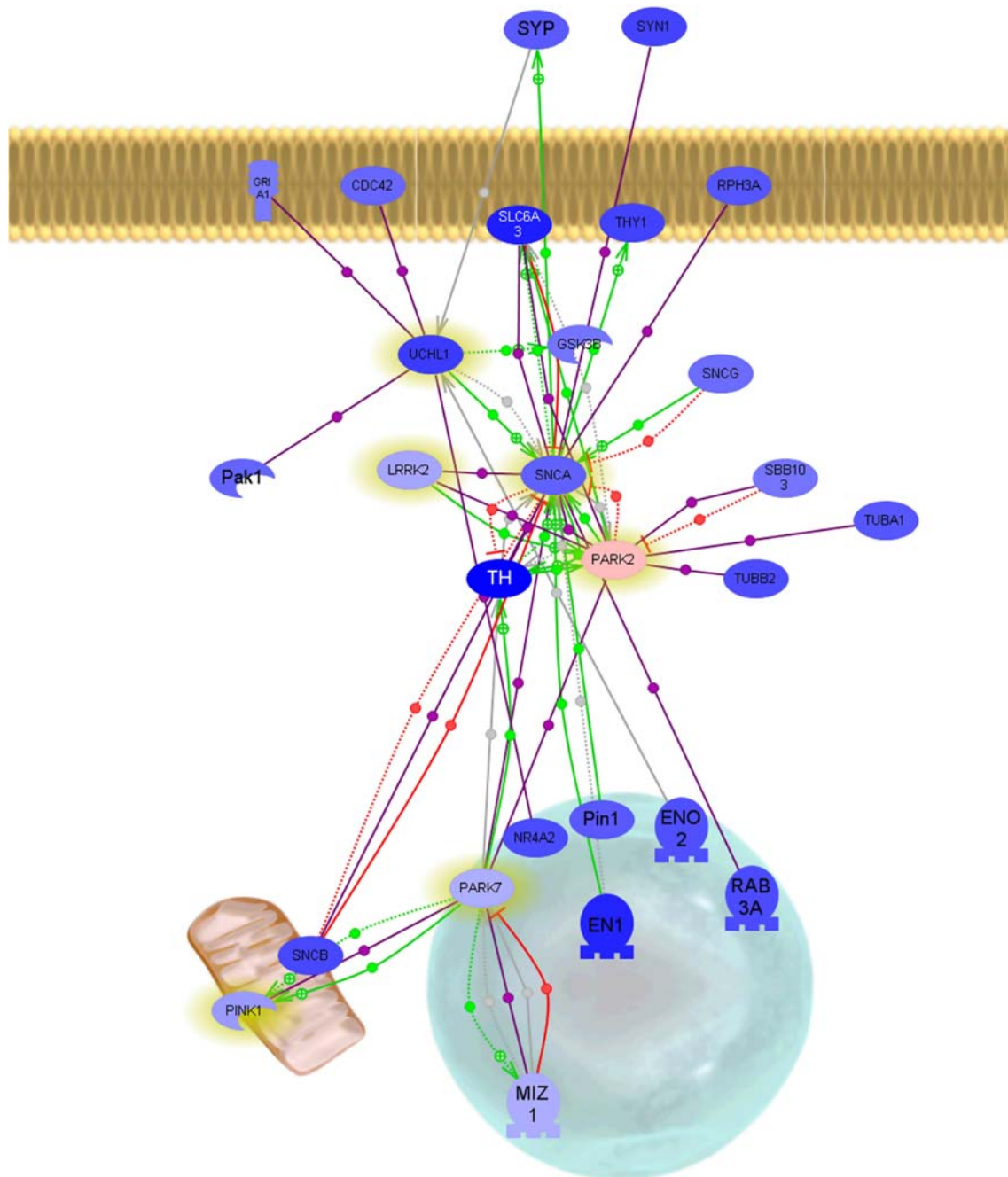


Fig. 5 Interactions of known PD genes (marked yellow) with the priority genes of this study. Expression value overlay: red indicates up-regulation and blue indicates down-regulation in PD nigra. An

online version of this figure with hyperlinks and an overlay of significance values is provided as SI_Figure_7. Cellular layout, display by effect (ResNet)

‘Hub’ vs ‘peripheral’ genes

A number of genes known to have numerous interactions with other genes were found amongst the priority genes. These represent so-called network ‘hubs’ and include HSPA1A, NFKBIA, CDC42, GSK3B, ACHE, AGTR1, IGF1R and TH as well as about 200 others (50 to >1,600 connectivities). Figure 2a and b contain a number of hubs.

However, almost 30% of the priority genes have no known interactions according to the ResNet 5.0 database although the cellular localisation is known in some instances. Such genes may be called ‘peripheral’ genes of the human interactome [24, 25]. Apparently, peripheral priority genes with still unknown pathway connections and cellular localisation are ACTR10, ANKRD29, ANKRD34, ANKRD50, ARM CX4, ARRDC2, ASMTL, ATAD1,

BLOC1S2, CAP2, CCDC4, CCDC85A, CKMT1A, CNIH3, DBNDD1, DCUN1D4, DIRAS2, DOPEY1, EHBP1, ELMOD1, FARSLB, FBXO9, FHOD3, GABARAPL3, GARNL4, GPRASP2, GPRIN3, GUSBP1, HISPPD1, HNRPUL2, IPW, KLHL1, KRT222P, LRRC49, LRRC55, LYNX1, MANEAL, MAP1LC3A, MAP9, MGC22265, MGC39606, MGC4677, MIA3, MRC1L1, NCDN, NGFRAP1L1, NRIP3, NUDT11, OCIAD2, OGDHL, OSBPL10, PCYOX1L, PFAAP5, PGM2L1, PLCXD3, PNMA6A, PRKY, PRMT8, PRPS1, RFPL1S, RUTBC2, SLC35F1, SNX10, SNX25, TAGLN3, TBC1D24, TBC1D9, TMEM130, TMEM132B, TMEM35, TRIM4, TRIM9, TSGA14, TTC7B, TUBB2B, TUBB3, UBPH, USP34, WDR47, XKR4, ZNF204.

Up-regulated peripheral genes are shown in Table 1. Following the above criteria, a number of Lewy body components and genes known to interact with the established PD genes seem peripheral to main brain metabolic pathways but their exact status remains to be determined.

Discussion

Our study reveals a significant up-regulation of substantia nigra genes in PD which have known biological associations with cancer, diabetes and inflammation. This includes major ‘hub’ genes [24–26] such as p53, somatic mutations of which can cause cancer. This is of note as p53 forms part of a molecular network that integrates tumour suppression and ageing [27]. DJ-1 is another cancer- and Parkinson’s disease-associated protein [28], and it is of special interest in the present context that the ubiquitin-proteasomal pathway has an established role in neoplastic processes [29]. Furthermore, both parkin and PINK1 might be tumour suppressor genes, and it has been suggested that although cancer is rare in PD, unraveling the link between PD and cancer [30–31] may open a therapeutic window for both diseases [32]. The finding of a molecular biological association between diabetes and PD is not truly surprising either [33–37]. Thirdly, the link of PD with inflammation which emerges from our unsupervised analysis seems almost expected considering the very lively debate of this topic in the literature. The whole genome transcriptome data presented here certainly justify additional scrutiny of the underlying mechanisms in relation to PD pathogenesis.

The problem of defining what causes PD at a system level has become more complex with the recent finding that disease-relevant genes may reside at the periphery of disease networks. It is interesting to note that the neighbour of a disease node appears more likely to be another disease protein, which also preferentially interacts with other disease nodes [38]. Proteins that are associated with the same disease show a 10-fold increased tendency to interact

with each other than those not associated with the same disease [26]. This should direct our attention also to genes that do not form major network hubs but which are either likely to be involved in PD on cell biological grounds (e.g. Lewy body components) or which interact with PD causing genes. A significant fraction of the genes identified in this study still represent functionally ill-characterised entities.

The on-line material of this manuscript illustrates some of the pathways and biological association networks that emerge from our analysis. Networks are now recognized to pervade all aspects of human biology and the question where function lies within a cell is shifting from a simple focus on genes to the understanding that behind each cellular function there is a discernible network module consisting of genes, transcription factors, RNAs, enzymes and metabolites [39]. However, ‘network medicine’ is still in its infancy and the present study may be the first where an iterative multidimensional tissue analysis approach, <http://www.neurogenetics.net/Multidimensional.html>, has been applied to a human neurological disorder. The ultimate goal of such analyses is the precise cellular localisation of all expressed human disease genes in their affected tissues. The present PD dataset has so far yielded two novel components of Lewy bodies [7, 8] but much more back-mapping work will need to be performed.

For instance, the exact mechanism of cell death in PD is still unknown [40]. Recent evidence has suggested that one mechanism linked to the death of terminally differentiated neurons is aberrant re-entry into the cell cycle, and possible connections between oxidative stress and unscheduled cell cycle re-entry in PD have been proposed [41]. However, as neuroscientists, we may have to move beyond the description of the cell cycle that has been propagated by those in the cancer field because the regulation of the cell cycle in the neuron is much more nuanced (K Herrup, <http://www.alzforum.org/new/detailprint.asp?id=1688>). This raises the possibility that some of the data supporting cell processes such as mutagenesis in this study may have to be re-read and interpreted in a modified way. It is worth noting in this context that absence of RET signalling in mice causes progressive and late degeneration of the nigrostriatal system [42]. We would also like to point out that the present study provides additional evidence for the importance of changes in the neuronal cytoskeleton in PD [43–44] because neurofilament subunit as well as microtubuli-associated protein genes were found to be highly dysregulated. Dysregulation of signal transduction, heat shock and synaptic proteins also featured very prominently.

The view that the 892 nigral genes presented here are relevant for sporadic PD is supported by the finding that their pattern of expression is characteristic of nervous tissue (SI_Figure_2). It is further clear from our data that the human substantia nigra in PD does not represent dead tissue

Table 1 Up-regulated 'peripheral' priority genes

Name	Description	Cell Localization
ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8b (predicted)	Plasma membrane
ARHGEF10	Rho guanine nucleotide exchange factor (GEF) 10	
ARRDC2	arrestin domain containing 2	
BBX	bobby sox homolog (Drosophila) (predicted) biregional cell adhesion molecule-related/down-regulated by oncogenes	Nucleus
BOC	(Cdon) binding protein (predicted)	Plasma membrane
C10orf104	DNA segment, Chr 10, ERATO Doi 641, expressed	
C10orf128	similar to novel protein (predicted)	
C10orf54	hypothetical protein LOC690899	
C22orf9	similar to CG9646-PA cysteine and histidine-rich domain (CHORD)-containing, zinc-binding protein 1 (predicted)	
CHORDC1		
CLEC2B	C-type lectin domain family 2, member B	Plasma membrane
CRLF3	cytokine receptor-like factor 3	Plasma membrane
DDA3	similar to differential display and activated by p53	
DEPP	similar to hypothetical protein MGC6835	
DKFZp667G2110	similar to Hypothetical protein 5031404N19 (predicted)	
DNAJB6	DnaJ (Hsp40) homolog, subfamily B, member 6	Nucleus
DOCK5	dedicator of cytokinesis 5 (predicted)	Golgi
ELTD1	EGF, latrophilin seven transmembrane domain containing 1	Plasma membrane
FAM123A	similar to hypothetical protein FLJ25477 isoform 2 (predicted)	
FLJ11280	hypothetical protein LOC310665	
FLJ40092	FLJ40092 protein	
GUSBP1	glucuronidase, beta pseudogene 1	
H2AFL	histone 1, H2aa	Nucleus
HIGD1B	HIG1 domain family, member 1B (predicted)	Plasma membrane
IFITM2	interferon induced transmembrane protein 2 (1-8D) potassium large conductance calcium-activated channel, subfamily M, beta member 4	Plasma membrane
KCNMB4		Plasma membrane
KIAA0500	KIAA0500 protein	
KST1	solute carrier family 5 (sodium/glucose cotransporter), member 11	Plasma membrane
LOC399959	hypothetical gene supported by BX647608	
LOC402573	hypothetical LOC402573	
LOC440248	hypothetical LOC440248	
MGC22265	hypothetical protein MGC22265	
MKNK2	MAP kinase-interacting serine/threonine kinase 2	
mn7	hect domain and RLD 2 pseudogene 2	
MRC1L1	mannose receptor, C type 1-like 1	
MT1H	metallothionein 1H	
MT1K	metallothionein 1M	
MT1X	metallothionein 1X	
NSUN6	NOL1/NOP2/Sun domain family, member 6 (predicted)	Nucleus
PFAAP5	similar to hypothetical protein from BCRA2 region	
PPAP2C	phosphatidic acid phosphatase type 2c	Plasma membrane
PRKY	protein kinase, Y-linked	
RPS11	ribosomal protein S11	
SEC14L1	similar to SEC14-like 1 (predicted)	Plasma membrane
SLCO4A1	solute carrier organic anion transporter family, member 4a1	Plasma membrane
TncRNA	trophoblast-derived noncoding RNA	
TRIM4	tripartite motif-containing 4	
TTID	myotilin (predicted)	
USP34	ubiquitin specific peptidase 34	
USP52	ubiquitin specific protease 52	Nucleus
USP54	ubiquitin specific peptidase 54	Plasma membrane
ZDHC11	similar to DHHC-containing protein 10	Plasma membrane
ZNF302	zinc finger protein 302	Nucleus

but that there is an active ongoing disease process understanding of which may hold the key to halting PD. The dysregulated priority genes may reside at the core of the disease process and could serve as novel targets for therapeutic intervention. This idea is supported by the observation that a number of priority genes interact with drugs whose actions are associated with a Parkinsonian clinical phenotype.

The uncertainty whether inflammatory processes truly represent a causative factor in the aetiology of PD [45] requires an answer. Our own work and that of others suggests a direct role of primary glial degeneration in the pathogenetic process underlying PD [10, 46]. This means that PD extends beyond the neuron. The disease is also not confined to the substantia nigra anatomically. Detailed cellular back-mapping of all priority genes to brain tissue will help to settle these questions. New algorithms are required to explain the links of PD as defined in the living and in the microscope with the underlying high-throughput datasets.

Collection of links to Electronic Supplemental Material (ESM)

SI_Table_1a List of all 892 highly dysregulated genes ('priority genes') contained in the ResNet 5.0 database (Ariadne) http://www.morphonom.net/ng/ESM/t/SI_Table_1a.xls

SI_Table_1b List of the 1,145 Affymetrix probes identifying highly dysregulated sequences including all 892 'priority genes' http://www.morphonom.net/ng/ESM/t/SI_Table_1b.txt

SI_Table_1c P and differential expression values for the 1,145 Affymetrix probes http://www.morphonom.net/ng/ESM/t/SI_Table_1c.xls

SI_Table_1d Publicly available information on the 1,145 Affymetrix probes from IDconverter (<http://idconverter.bioinfo.cnio.es/>) http://www.morphonom.net/ng/ESM/t/SI_Table_1d.xls

SI_Table_1e Information on gene sequences not found in the ResNet database obtained via NetAffx (<http://www.affymetrix.com/analysis/index.affx>) http://www.morphonom.net/ng/ESM/t/SI_Table_1e.xls

SI_Table_1f List of priority genes that are published PD candidate genes (<http://www.pdgene.org/>) http://www.morphonom.net/ng/ESM/t/SI_Table_1f.xls

SI_Table_1g List of priority genes that are published AD candidate genes http://www.morphonom.net/ng/ESM/t/SI_Table_1g.rtf

SI_Table_1h List of priority genes with known functional links to PD (PubMed) http://www.morphonom.net/ng/ESM/t/SI_Table_1h.xls

The literature references for this table can be found in http://www.morphonom.net/ng/ESM/r/SI_References_1.rtf

SI_Table_2 Gene Ontology analysis of the 892 genes http://www.morphonom.net/ng/ESM/t/SI_Table_2.xls

SI_Table_3 Signalling pathway analysis of the 892 genes (PathwayStudio 5.0) http://www.morphonom.net/ng/ESM/t/SI_Table_3.xls

SI_Figures_1a Hierarchical clustering of the 1,145 probes (dendrogram description: clustering on rows and columns, Pearson centred distance metric, centroid linkage rule); the clustering separates PD cases (green, right) from controls (red, left); designation of samples as indicated in the dataset submitted to GEO (see 'Materials and methods' section). http://www.morphonom.net/ng/ESM/f/SI_Figure_1a.png

Abbreviations used: F, female; M, male; LN, lateral substantia nigra; MN, medial substantia nigra; CON, PDC, controls; PD, Parkinson's disease; followed by the number indicating the age of each subject (cf. Moran et al. [3]); LN, MNCON10, M, 71; LN, MNCON2, M, 77; MNCON3, M, 81; LN, MNCON9, M, 57; LN, MNPD01, F, 87; LN, MNPD02, M, 83; LN, MNPD04, M, 68; LN, MNPD07, M, 78; LN, MNPD09, F, 86; LN, MNPD10, F, 81; LN, MNPD16, F, 85; MNPD20, M, 75; MNPD21, M, 76; LN, MNPD28, M, 82; LN, MNPD29, M, 76; MNPD32, M, 89; MNPD34, F, 84; MNPD36, M, 76; LN, MNPDC1, M, 76.

An explanation for the colour coding is provided in http://www.morphonom.net/ng/ESM/f/SI_Figures_1a&Color_range.jpg

The labels used for the clustering and the corresponding Affymetrix probe set IDs are explained in http://www.morphonom.net/ng/ESM/f/SI_Figures_1a&Labels_used_for_clustering.xls

The four probes at the very bottom of the figure (XIST, X (inactive)-specific transcript) in PD01, 9, 10, 16 and 34 identify the female patients in our refined cohort. They served as an internal control.

SI_Figures_1b Column similarity image for SI_Figure_1a.png http://www.morphonom.net/ng/ESM/f/SI_Figure_1b.png

SI_Figures_1c Self-organising map corresponding to SI_Figure_1a (dendrogram description: clustering on rows and columns, Euclidean distance metric, maximum number of iterations 50, number of grid rows 3, number of grid columns 4, initial learning rate 0.03, initial neighbourhood radius 5, grid topology hexagonal, neighbourhood type bubble) http://www.morphonom.net/ng/ESM/f/SI_Figure_1c.png

SI_Figure_2 Hierarchical clustering of the 1,145 probes using 64 whole genome array datasets (Affymetrix Human Genome U133 Plus 2.0 arrays) representing individual organ samples including 20 brain regions and three ganglia [47]. GSM numbers refer to the respective file names in the complete dataset which comprises 353 whole genome arrays (GEO database, GSE ID GSE3526). There is a

complete separation of nervous tissue (left) from other organs on the basis of the 1,145 probes. http://www.morphonom.net/ng/ESM/f/SI_Figure_2.png

The designations of all samples and their code numbers are provided in http://www.morphonom.net/ng/ESM/r/SI_References_2.rtf

SI_Figure_3a Online version (with links) of Fig. 2a http://www.morphonom.net/ng/ESM/f/Cell_Processes.html

SI_Figure_3b Online version (with links) of Fig. 2b. (XIST was not included in this analysis). <http://www.morphonom.net/ng/ESM/f/Disease.html>

SI_Figures_4a-c Online permutations of Fig. 3

Layout by cellular localisation with links <http://www.morphonom.net/ng/ESM/f/892.html>

Symmetrical layout with links <http://www.morphonom.net/ng/ESM/f/892s.html>

Hierarchical layout with links <http://www.morphonom.net/ng/ESM/f/892h.html>

SI_Figure_5 Known components of Lewy bodies are indicated by the blue highlighting (this figure is identical to SI_Figure_1c except that it has a lower resolution) http://www.morphonom.net/ng/ESM/f/SI_Figure_5.png

SI_Figure_6 Online version of Fig. 4 with expression values overlaid (red indicates up- and blue indicates down-regulation in PD nigra, respectively). The blue shading of the priority genes has been replaced with yellow. http://www.morphonom.net/ng/ESM/f/Lewy_Body.html

SI_Figure_7 Hyperlinked online version of Fig. 5. The gene colours range in this figure indicates significance: high, white; low, red. The blue shading indicates that the respective gene is a priority gene. http://www.morphonom.net/ng/ESM/f/PD_genes_interactions_with_892_direct_no_DorCP.html

SI_Figures_8a-c Known drug interactions of some of the priority genes as derived from the ResNet 5.0 database (more than 9,000,000 database objects were checked). It is noteworthy that drugs such as clozapine, cocaine and haloperidol which are used in the treatment of PD or which cause Parkinsonian side effects appear to interact with an especially large number of PD priority genes.

Clozapine http://www.morphonom.net/ng/ESM/f/SI_Figure_8a.png

Cocaine http://www.morphonom.net/ng/ESM/f/SI_Figure_8b.png

Haloperidol http://www.morphonom.net/ng/ESM/f/SI_Figure_8c.png

SI_Pathway_1 This signalling pathway was identified based on a search of 192 canonical pathways and 555 signalling pathways in PathwayStudio 5.0 and ranked highest (also see SI_Table_3). Six priority genes are represented in this pathway and are marked by the blue shading. http://www.morphonom.net/ng/ESM/p/RET_HSF1_signaling_pathway.html

SI_Pathway_2 Interactions of the cytostatic drug, paclitaxel with a total of 13 priority genes (blue shading) are shown. Display by effect. Hierarchical layout. http://www.morphonom.net/ng/ESM/p/paclitaxel_interactions.html

SI_Pathway_3 Interactions of the cytostatic drug, vincristine with 3 priority genes (violet shading, bottom of figure) are illustrated. Red shaded genes also showed dysregulation in the PD nigra ($p < 0.001$). Display by references count (darker blue indicates a larger number of references supporting the respective connection). Hierarchical layout. http://www.morphonom.net/ng/ESM/p/vincristine_interactions.html

SI_References_1 References for SI_Table_1h http://www.morphonom.net/ng/ESM/r/SI_References_1.rtf

SI_References_2 Designations and code numbers for SI_Figure_2 http://www.morphonom.net/ng/ESM/r/SI_References_2.rtf

Acknowledgements We express our deepest appreciation to the donors and their families for donating human brain tissue for research to the UK Parkinson's Disease Society Brain Bank. This work was funded by the Parkinson's Disease Society of the United Kingdom, registered charity 948776. We are also most grateful to BUG. LBM would like to thank the UK Parkinson's Disease Society for special support.

References

1. Tretiakoff C (1919) Contribution à l'étude de l'anatomie du locus niger de Soemmering avec quelques déductions relatives à la pathogenie des troubles du tonus musculaire et de la maladie de Parkinson. Thèse de Paris
2. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, Stenroos ES, Chandrasekharappa S, Athanassiadou A, Papapetropoulos T, Johnson WG, Lazzarini AM, Duvoisin RC, Di Iorio G, Golbe LI, Nussbaum RL (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 276:2045–2047
3. Moran LB, Duke DC, Deprez M, Dexter DT, Pearce RK, Graeber MB (2006) Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. *Neurogenetics* 7:1–11
4. Hauser MA, Li YJ, Xu H, Noureddine MA, Shao YS, Gullans SR, Scherzer CR, Jensen RV, McLaurin AC, Gibson JR, Scott BL, Jewett RM, Stenger JE, Schmechel DE, Hulette CM, Vance JM (2005) Expression profiling of substantia nigra in Parkinson disease, progressive supranuclear palsy, and frontotemporal dementia with parkinsonism. *Arch Neurol* 62:917–921
5. Zhang Y, James M, Middleton FA, Davis RL (2005) Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. *Am J Med Genet B Neuropsychiatr Genet* 137:5–16
6. Miller RM, Kiser GL, Kaysser-Kranich TM, Lockner RJ, Palaniappan C, Federoff HJ (2006) Robust dysregulation of gene expression in substantia nigra and striatum in Parkinson's disease. *Neurobiol Dis* 21:305–313
7. Moran LB, Durrenberger PF, Pearce RK, Graeber MB (2007) Two new molecular markers of Lewy bodies in Parkinson's disease.

- Journal of Neurology, Neurosurgery and Psychiatry 78(9) (Abs ABN Spring Scientific Meeting, Cambridge 2007)
8. Moran LB, Hickey L, Michael GJ, Derkacs M, Christian LM, Kalaitzakis ME, Pearce RK, Graeber MB (2007) Neuronal pentraxin II is highly upregulated in Parkinson's disease and a novel component of Lewy bodies. *Acta Neuropathol* DOI 10.1007/s00401-007-0309-3
 9. West AB, Dawson VL, Dawson TM (2005) To die or grow: Parkinson's disease and cancer. *Trends Neurosci* 28:348–352
 10. Duke DC, Moran LB, Kalaitzakis ME, Deprez M, Dexter DT, Pearce RK, Graeber MB (2006) Transcriptome analysis reveals link between proteasomal and mitochondrial pathways in Parkinson's disease. *Neurogenetics* 7:139–148
 11. Shults CW (2006) Lewy bodies. *Proc Natl Acad Sci USA* 103:1661–1668
 12. Kim YS, Joh TH (2006) Microglia, major player in the brain inflammation: their roles in the pathogenesis of Parkinson's disease. *Exp Mol Med* 38:333–347
 13. Wersinger C, Sidhu A (2006) An inflammatory pathomechanism for Parkinson's disease? *Curr Med Chem* 13:591–602
 14. McGeer PL, McGeer EG (2007) Glial reactions in Parkinson's disease. *Mov Disord* DOI 10.1002/mds.21751
 15. Croisier E, Moran LB, Dexter DT, Pearce RK, Graeber MB (2005) Microglial inflammation in the parkinsonian substantia nigra: relationship to alpha-synuclein deposition. *J Neuroinflammation* 2:14
 16. Ton TG, Heckbert SR, Longstreth WT, Rossing MA, Kukull WA, Franklin GM, Swanson PD, Smith-Weller T, Checkoway H (2006) Nonsteroidal anti-inflammatory drugs and risk of Parkinson's disease. *Mov Disord* 21:964–969
 17. Bornebroek M, de Lau LM, Haag MD, Koudstaal PJ, Hofman A, Stricker BH, Breteler MM (2007) Nonsteroidal anti-inflammatory drugs and the risk of Parkinson disease. *Neuroepidemiology* 28:193–196
 18. Moran LB, Croisier E, Duke DC, Kalaitzakis ME, Roncaroli F, Deprez M, Dexter DT, Pearce RK, Graeber MB (2007) Analysis of alpha-synuclein, dopamine and parkin pathways in neuropathologically confirmed parkinsonian nigra. *Acta Neuropathol* 113:253–263
 19. Yuryev A, Mulyukov Z, Kotelnikova E, Maslov S, Egorov S, Nikitin A, Daraselia N, Mazo I (2006) Automatic pathway building in biological association networks. *BMC Bioinformatics* 7:171
 20. Alibes A, Yankilevich P, Canada A, Diaz-Uriarte R (2007) IDconverter and IDClight: conversion and annotation of gene and protein IDs. *BMC Bioinformatics* 8:9
 21. Cantuti-Castelvetri I, Keller-McGandy C, Bouzou B, Asteris G, Clark TW, Frosch MP, Standaert DG (2007) Effects of gender on nigral gene expression and Parkinson disease. *Neurobiol Dis* 26:606–614
 22. Bowyer JH, Muentner MD (1995) Temporary worsening of parkinsonism in a patient with Parkinson's disease after treatment with paclitaxel for a metastatic grade IV adenocarcinoma. *Mov Disord* 10:681–682
 23. Boranic M, Raci F (1979) A Parkinson-like syndrome as side effect of chemotherapy with vincristine and adriamycin in a child with acute leukaemia. *Biomedicine* 31:124–125
 24. Jonsson PF, Bates PA (2006) Global topological features of cancer proteins in the human interactome. *Bioinformatics* 22:2291–2297
 25. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL (2007) The human disease network. *Proc Natl Acad Sci USA* 104:8685–8690
 26. Loscalzo J, Kohane I, Barabasi AL (2007) Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Mol Syst Biol* 3:124
 27. Papazoglu C, Mills AA (2007) p53: at the crossroad between cancer and ageing. *J Pathol* 211:124–133
 28. Clements CM, McNally RS, Conti BJ, Mak TW, Ting JP (2006) DJ-1, a cancer- and Parkinson's disease-associated protein, stabilizes the antioxidant transcriptional master regulator Nrf2. *Proc Natl Acad Sci USA* 103:15091–15096
 29. Mani A, Gelmann EP (2005) The ubiquitin-proteasome pathway and its role in cancer. *J Clin Oncol* 23:4776–4789
 30. Kim RH, Mak TW (2006) Tumours and tremors: how PTEN regulation underlies both. *Br J Cancer* 94:620–624
 31. Zanetti R, Loria D, Rosso S (2006) Melanoma, Parkinson's disease and levodopa: causal or spurious link? A review of the literature. *Melanoma Res* 16:201–206
 32. Inzelberg R, Jankovic J (2007) Are Parkinson disease patients protected from some but not all cancers? *Neurology* 69:1542–1550
 33. Sandyk R (1993) The relationship between diabetes mellitus and Parkinson's disease. *Int J Neurosci* 69:125–130
 34. Ristow M (2004) Neurodegenerative disorders associated with diabetes mellitus. *J Mol Med* 82:510–529
 35. Powers KM, Smith-Weller T, Franklin GM, Longstreth WT, Swanson PD, Checkoway H (2006) Diabetes, smoking, and other medical conditions in relation to Parkinson's disease risk. *Parkinsonism Relat Disord* 12:185–189
 36. Arvanitakis Z, Wilson RS, Bienias JL, Bennett DA (2007) Diabetes and parkinsonian signs in older persons. *Alzheimer Dis Assoc Disord* 21:144–149
 37. Hu G, Jousilahti P, Bidel S, Antikainen R, Tuomilehto J (2007) Type 2 diabetes and the risk of Parkinson's disease. *Diabetes Care* 30:842–847
 38. Xu J, Li Y (2006) Discovering disease-genes by topological features in human protein-protein interaction network. *Bioinformatics* 22:2800–2805
 39. Barabasi AL (2007) Network medicine—from obesity to the “diseasome”. *N Engl J Med* 357:404–407
 40. Graeber MB, Moran LB (2002) Mechanisms of cell death in neurodegenerative diseases: fashion, fiction, and facts. *Brain Pathol* 12:385–390
 41. Klein JA, Ackerman SL (2003) Oxidative stress, cell cycle, and neurodegeneration. *J Clin Invest* 111:785–793
 42. Kramer ER, Aron L, Ramakers GM, Seitz S, Zhuang X, Beyer K, Smidt MP, Klein R (2007) Absence of Ret signaling in mice causes progressive and late degeneration of the nigrostriatal system. *PLoS Biol* 5:e39
 43. Ren Y, Liu W, Jiang H, Jiang Q, Feng J (2005) Selective vulnerability of dopaminergic neurons to microtubule depolymerization. *J Biol Chem* 280:34105–34112
 44. Jiang Q, Yan Z, Feng J (2006) Neurotrophic factors stabilize microtubules and protect against rotenone toxicity on dopaminergic neurons. *J Biol Chem* 281:29391–29400
 45. Whitton PS (2007) Inflammation as a causative factor in the aetiology of Parkinson's disease. *Br J Pharmacol* 150:963–976
 46. Croisier E, Graeber MB (2006) Glial degeneration and reactive gliosis in alpha-synucleinopathies: the emerging concept of primary gliodegeneration. *Acta Neuropathol* 112:517–530
 47. Roth RB, Hevezi P, Lee J, Willhite D, Lechner SM, Foster AC, Zlotnik A (2006) Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. *Neurogenetics* 7:67–80