21

Genome Profiling and Potential Biomarkers in Neurodegenerative Disorders

Luca Lovrečić, Aleš Maver and Borut Peterlin
Clinical Institute of Medical Genetics, University Medical Center Ljubljana, Slovenia

1. Introduction

Neurodegenerative disorders (NDG) are incurable, progressive and debilitating conditions resulting from progressive degeneration and death of nerve cells. They are among the most serious health problems faced by modern society. Most of these disorders become more common with advancing age, including Alzheimer's disease and Parkinson's disease. The burden of these neurodegenerative diseases is growing inexorably as the population ages, with incalculable economic and human costs. According to the Global Burden of Disease Study, a collaborative study of the World Health Organization, the World Bank and the Harvard School of Public Health, dementia and other neurodegenerative diseases will be the eighth cause of disease burden for developed regions in 2020 [1, 2]. Also, according to the WHO, neurodegenerative diseases will become the world's second leading cause of death by 2050, overtaking cancer [2]. True, such estimates and predictions need to be taken with caution, but they definitely confirm that neurodegenerative diseases are of an increasing public concern.

Most NDG diseases are characterized by the aggregation of intracellular proteins. Majority of neurodegenerative disorders occur sporadically and are believed to arise through interactions between genetic and environmental factors. Only a small minority belong to familial forms where certain disease occurs due to a mutation of the gene coding for the abnormally aggregating protein.

We differentiate many types of NDG disease, but the lines that separate one from another are often unclear. For instance, symptoms such as motor impairment and dementia may occur in many different types of NDG disease. Motor impairment similar to that seen in Parkinson's disease is not enough to rule out other diagnoses, especially when both motor and cognitive impairment are present. At the time being, there is no such diagnostic test that can clearly indicate the presence, absence, or category of a NDG disease. Individual diagnosis is based on clinical evaluation of the symptoms, with the exception of monogenic NDG diseases, such as Huntington's disease (HD). HD is a single gene disorder and cause is invariably trinucleotide expansion mutation [3].

Definitive diagnosis of certain NDG diseases still relies on neuropathological evaluation. But it has been demonstrated that brain pathology can show marked overlap among the syndromes of age-related cognitive and motor impairment [4]. Also, previous research reports have shown that pathological markers do not always correlate optimally with clinical findings. Some individuals with extensive neuropathology may retain relatively

intact neurological function while others with less extensive pathology may be significantly impaired [5, 6]. The neuropathological findings may be the response to other antecedent disease processes and are not necessarily the cause of the underlying disease at the early disease stages. Later, as disease progresses, they probably contribute to disease progression in a positive feedback loop.

Analysis of whole genome transcriptome in brain might give us insights into the disturbed pathways and processes involved in disease onset and progression. Many different mechanisms have been proposed to be dysregulated in NDG diseases. We collected all reported studies to date on brain transcriptome in Parkinson's disease, Alzheimer disease, Huntington disease and Down syndrome and performed an integrated meta-analysis.

2. Background

2.1 Common neurodegenerative disorders – Alzheimer and Parkinson disease

Two most common neurodegenerative diseases, Parkinson's disease (PD) and Alzheimer disease(AD) are believed to be heterogeneous based on the causes - combination of genetic and environmental factors, vast variety in the age at onset, variability in leading symptoms and presenting clinical manifestations, disease progression and responses to different therapies employed. Definitive diagnosis of both, AD and PD still relies on a 'gold standard' post mortem neuropathological evaluation, although a number of clinical and neuropsychological tests are often employed when making a clinical diagnosis. AD is detected with approximately 85–90% accuracy and PD with approximaly 75% accuracy. The pathogenesis of both AD and PD are complex and still remain unexplained in worldwide research community.

It has been recently estimated [7] that 24 million people have dementia worldwide and majority is attributable to AD. The authors emphasized the urgency of better understanding of pathophysiology of the disease in order to improve development of disease-modifying treatment. Due to the age-dependent incidence rate of AD and due to the population ageing, it is foreseen that more than 80 million people will have AD by 2040 [8]. It is a progressive neurologic disease affecting particularly cortical and hippocampal neurons, leading to their irreversible loss [9]. Major clinical signs and symptoms are progressive impairment in memory, judgment, decision making, orientation to physical surroundings, and language. The key pathological characteristics are neuronal loss, β amyloid containing extracellular senile plaques, and neurofibrillary tangles, which are composed of a hyperphosphorylated form of the microtubular protein tau.

PD is the second most prevalent NDG disease after AD. According to available data of European Parkinson's Disease Association (EPDA), there are 6.3 million people with PD worldwide. Prevalence is age-dependent - there are approximately 0.5 to 1 percent of individuals with PD in the age group 65 to 69 years, and 1 to 3 percent of individuals with PD in the group of people older than 80 years [10]. Typical clinical sign is parkinsonism - resting tremor, bradykinesia, rigidity, and postural instability. Neuropathological characteristics are the loss of neurons in the substantia nigra and the presence of neuronal inclusions termed Lewy bodies and Lewy neurites whose main component is aggregated and phosphorylated alpha-synuclein [11].

Important futuristic challenge in the management AD and PD remains the establishment of early diagnosis or even identification of individuals prior to the onset of dementia in AD or resting tremor in PD. This implicates advancement in understanding disease

pathogenesis and development of diagnostic approaches, including disease/process specific biomarkers.

2.2 Huntington disease – A model of genetic neurodegenerative disorder

Huntington disease is a late onset, single gene disorder and its cause is invariably trinucleotide expansion mutation, known for almost 2 decades [3]. Clinical characteristics of the disease include progressive motor impairment, cognitive decline and various psychiatric symptoms with the typical age of onset in the third to fifth decade. The disease is fatal after 15-20 years of progressive neurodegeneration [12]. So far, no effective treatment has been available to cure the disease or to slow its progression. Hyperkinesias and psychiatric symptoms may respond well to pharmacotherapy, but neuropsychological deficits and dementia remain untreatable [13]. We are unable to predict the age at onset and to follow the disease progression over short time periods due to the unsensitivity of rating scales. Even more, no useful measures to follow response to symptomatic treatment over short time periods are known. In addition, in the presymptomatic period when preventive treatment and slowing of neurodegeneration might be most effective, we have no measures/markers to monitor those responses and benefits.

Although the responsible gene and mutation were already identified and characterized in 1993, the function of normal huntingtin and the mutation mechanism that leads to neurodegeneration are still not clear. Basic research has demonstrated that the pathogenesis of HD involves recruitment of multiple biochemical pathways like protein degradation, apoptosis, accumulation of misfolded mutated proteins, intracelular signaling, oxidative stress, mitochondrial involvement and in the last years also transcription [14, 15].

2.3 Dementia and Down syndrome

Dementia, common symptom of all three already mentioned neurodegenerative diseases is also a common symptom in individuals with Down syndrome (DS). Most of individuals with DS after about age of 30 have the characteristic plaques and neurofibrillary tangles, associated with AD. As in general population, the prevalence of AD in people with DS increases significantly with age. On the other hand, age-related cognitive decline and dementia in people with DS occurs 30–40 years earlier than in the general population, reaching almost 40% in the 50s [16]. Life expectancy of people with DS continues to increase and therefore, dementia is becoming an important issue.

2.4 Biomarkers

Research in the field of biomarkers is a rapidly growing and developing area in medicine. Everyday advances in genomic, proteomic, metabolomic and epigenomic knowledge and technologies have made their way also in the neuroscientific research area. Biomarkers are very important indicators of normal and abnormal biological processes. By definition, biological marker or biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [17]. Despite the fact that enormous effort and extensive research have been concentrated on this area, there is still a major lack of biomarkers for diagnosis, progression monitoring, response to treatment evaluation, etc. in neurodegenerative disorders such as Alzheimer's disease (AD), Parkinson's disease (PD) and Huntington's disease (HD).

Biomarkers have many valuable applications, such as identification of major neuropathological processes in specific disease, disease detection and monitoring of health status, early efficacy and safety evaluations in *in vitro* studies in tissue samples, *in vivo* studies in animal models, and early-phase clinical trials. They are invaluable as a diagnostic tool for identification of patients with a disease or abnormal condition, as a tool in staging the disease or classification of the extent of disease, as an indicator of disease prognosis and in predicting and monitoring of a clinical response to treatment.
Biomarkers are of extreme relevance in chronic NDG diseases - there are no cures for these diseases, as neurons of the central nervous system cannot regenerate on their own after cell death or damage. Tremendous efforts have been made in recent years to identify the neuropathological, biochemical, and genetic biomarkers of these diseases aiming to establish the diagnosis in earlier stages, to survey the rate of progression, or response to treatment. Currently, the neuropathologic diagnosis is a gold standard, but it can only be made in the form of an autopsy after the patient's death. On the other hand, biomarkers may improve the early diagnosis at a stage when disease-modifying therapies are likely to be most effective, the monitoring of disease progression and the efficacy of any therapeutic intervention [18].

2.5 Brain transcriptome in neurodegenerative disorders

Many different research groups have tried to solve the neuropathophysiological puzzle in PD, AD, HD and DS. Human brain has been extensivelly studied using many approaches, in the last decade also variety of »omic« technologies. Whole-genome gene expression studies in brain of each of four diseases individually have shown changes in transcription of number of genes when compared to normal human brain.
We investigated, reviewed and collected data from all reported studies to date on brain transcriptome in Parkinson's disease, Alzheimer disease, Huntington disease and Down syndrome and performed integrated meta-analysis.

3. Methods

In an attempt to present the alterations consistently reported by studies of brain transcriptome in neurodegenerative diseases, we initially searched for such reports in literature databases, then obtained raw and processed experimental data from microarray data repositories, after which we performed probe level meta-analyses of datasets originating from various studies. In addition, to reveal possible commonalities and shared pathways across various neurodegenerative diseases, we inspected the similarities and differences in gene expression dysregularities occurring in these conditions.

3.1 Study inclusion

Initially, we have searched Medline database (http://www.ncbi.nlm.nih.gov/pubmed) for reports from studies of interest using the search string (transcriptom* OR microarray OR profiling OR Affymetrix OR Agilent OR Illumina OR array) AND (Parkinson's disease OR Parkinsons disease OR Parkinson disease AND Alzheimer's disease OR Alzheimers disease OR Alzheimer disease OR dementia OR Down's syndrome OR Downs syndrome OR Down syndrome OR trisomy 21 OR Huntington's disease OR Huntingtons disease OR Huntington disease) to obtain the complete list of studies reporting results relating to transcriptional alterations in brain tissues affected by neurodegenerative processes.

As we were primarily interested in the studies with microarray experimental results accessible from biological repositories, we then searched Gene Expression Omnibus (GEO) repository (*http://www.ncbi.nlm.nih.gov/geo/*), ArrayExpress database (*http://www.ebi.ac.uk/arrayexpress/*) and Stanford Microarray database (*http://smd.stanford.edu*) for studies with data available in the raw or processed form. As most of the gene expression profiling experiments were performed on Affymetrix platform and to avoid difficulties due to different probe annotations utilized by different microarray manufacturers, only results from experiments performed on the Affymetrix U133 platform were included to facilitate further steps in probe level meta-analysis of microarray data. The detailed information on datasets included in the analyses may be observed in Table 1.

3.2 Microarray data pre-processing and preparation for meta-analysis

All the integration and statistical steps described were performed in R statistical environment version 2.13.1 (http://cran.r-project.org), using Bioconductor version 2.8 packages (available at http://bioconductor.org) [19]. Raw data from all microarray experiments listed in Table 1 was obtained directly from Gene Expression Omnibus (GEO) repository (http://www.ncbi.nlm.nih.gov/geo/) utilizing the GEOquery package for R [20, 21].

Before the meta-analysis of data from selected studies was performed, all the datasets obtained in such manner were inspected for significant inter-array differences in distribution of probe intensities. For this reason, raw datasets were initially examined using arrayQualityMetrics package and where necessary the straightforward quantile normalization functions in the affyPLM package was utilized [30, 31]. Non-specific intensity and interquartile variation filters were applied using methods in genefilter package [19]. Log_2 transformations were applied where discrepancies in data reporting format were observed.

Data collections for each individual neurodegenerative disease were then merged using probeset annotations as the common denominator. Using this approach we avoided potential statistical issues originating from averaging probe intensity values to obtain a single mean intensity value for each gene, possibly disregarding distinct expression of different transcripts from the same gene.

These steps resulted in generation of 4 separate data matrices, each carrying data for a single disease, originating from multiple studies – Alzheimer disease (AD), Down syndrome (DS), Huntington disease (HD) and Parkinson disease (PD) datasets.

3.3 Meta-analysis

Summarized differential expression of genes in each merged dataset was calculated using meta-analysis algorithms incorporated in the RankProd package for R [32]. RankProd uses a non-parametric statistical algorithm that facilitates detection of genes that are consistently highly ranked across microarray datasets originating from various microarray experiments in various studies perfomed on the same condition (ie. disease). As this approach is based on rank statistics in contrast to approaches requiring analyzing absolute intensity values, it allows for inclusion of data originating from different laboratories, differing platforms and potentially studies performed under differing conditions [32].

For analyses of such multi-study data, RPadvance function was utilized in our analyses, with origin parameter set to account for data originating from number of different sources corresponding to the number of different originating study [32]. Here it is important to

GEO Accession	Disease name	Platform	Number of probesets*	Number of array experiments		Tissue	Ref
				Affected tissue	Unaffected tissue		
GSE5281	Alzheimer's disease	Affymetrix HG-U133Plus2	54,675	87	74	Entorhinal cortex Hippocampus Medial temporal gyrus Posterior cingulate cortex Primary visual cortex Superior frontal gyrus	[22]
GSE1297	Alzheimer's disease	Affymetrix HG-U133A	22,283	22	9	Hippocampus	[23]
†GSE16759	Alzheimer's disease	Affymetrix HG-U133Plus2	54,675	4	4	Parietal lobe tissue	[24]
†GSE7307	Parkinson's disease	Affymetrix HG-U133Plus2	54,675	22	45	Caudate Gloubus pallidum Putamen Substantia nigra Subthalamic nucleus Thalamus lateral nuclei Thalamus subthalamic nucleus	NA‡
GSE8397	Parkinson's disease	Affymetrix HG-U133A and Affymetrix HG-U133B	22,283 and 22,645	29 and 29	18 and 18	Substantia nigra Frontal cortex	[25]
GSE7621	Parkinson's disease	Affymetrix HG-U133Plus2	54,675	16	9	Substantia nigra	[26]
GSE3790	Huntington's disease	Affymetrix HG-U133A and Affymetrix HG-U133B	22,283 and 22,645	114	87	Cerebellum Frontal cortex Caudate nucleus	[27]
†GSE1397	Down syndrome	Affymetrix HG-U133A	22,283	9	9	Cerebrum Cerebellum Astrocyte samples	[28]
GSE5390	Down Syndrome	Affymetrix HG-U133A	22,283	7	8	Dorsolateral prefrontal cortex	[29]

* According to data obtained from the GEO site

† The dataset included some microarray experiments not related to the scope of this study and those were omitted from the analyses

‡ The study related to listed GEO entry was not yet published

Table 1. Detailed information on studies included in meta-analysis

stress that we have faced the issue of multiple studies simultaneously reporting differential expression in several different anatomical brain parts. As we wanted to facilitate the discovery of differentially expressed genes in diseased tissue in comparison to control samples, we set the origin parameter to take into account these considerations and regard such data as originating from different sources, thereby avoiding comparisons of gene expression between different brain regions rather than between affected and unaffected samples. Afterwards, P-values and q-values were obtained by performing 100 permutation cycles of complete originating datasets. An arbitrary P-value cut-off for significance of differential gene expression was then set at $P<0.05$.

3.4 Investigating intersections between datasets and gene set enrichment analyses

Resulting ordered lists of differentially expressed probesets were subsequently investigated for overlap between AD, DS, HD and PD datasets. Top 1000 genes from each dataset were used and intersections between combinations of two, three and four datasets were obtained. Venn diagrams in the results section were produced using Venny utility available at *http://bioinfogp.cnb.csic.es/tools/venny/index.html*. Furthermore, to gain insight in functional properties of genes in the intersections, gene set enrichment analyses (GSEA) were performed, utilizing GOstats package for R and investigating significant (uncorrected $p<0.05$) over- or underrepresentation of GeneOntology (GO) and KEGG terms annotating genes occurring in the intersections [33-36]. Additionally, DAVID tool (http://david.abcc.ncifcrf.gov/) was used to reveal the functional annotation clusters related to intersecting genes [37]. Required annotation conversions were performed using the hgu133plus.db package from Bioconductor annotation package collection and using biomaRt package for R in combination with Ensembl Biomart service (http://www.biomart.org/) [38, 39].

4. Results

Alltogether, our data collection comprised of data from 9 whole-genome expression studies, performed on samples from 4 neurodegenerative conditions (AD, DS, HD and PD). Collectively, 200, 33, 201, and 186 microarray analysed samples were included in the investigations of AD, DS, HD and PD, respectively, which accounted for 620 separate experiments included overall. A slight predominance of experiments performed on case tissues was noted in most of the experiements with summary case:control ratio amounting to 1,2:1 (339 affected tissues and 281 unaffected tissues included).

Separate analyses of datasets for each NDG disorder have revealed significant perturbances in expression profiles of several genes. When arbitrary permutation p-value cut-off was set at 0.05 for upregulated genes, 5701 probesets attained significance in the AD dataset, 3291 in DS dataset, 4174 in the HD dataset and 3043 in the PD dataset. In the downregulated gene group the $p<0.05$ significance was reached for 5496 probesets in the AD dataset, 2983 probesets in the DS dataset, 4079 in the HD dataset and 3410 in the PD dataset. A detailed view of the distribution of significance values of the top 10,000 ordered differentially expressed genes may be observed in Figure 1 for each of the NDG disorders.

The resulting numbers of significant results are inflated by the effect of multiple testing and therefore the q- values were also estimated as described in the article by Breitling et al [40]. The numbers of upregulated probesets with estimated q-values below 0.05 were 3775 for AD, 1496 for DS, 3182 for HD and 1894 for PD datasets. The numbers of downregulated probesets meeting this criterion were 3624 in AD, 652 in DS, 3065 in HD and 2541 probesets in the PD dataset.

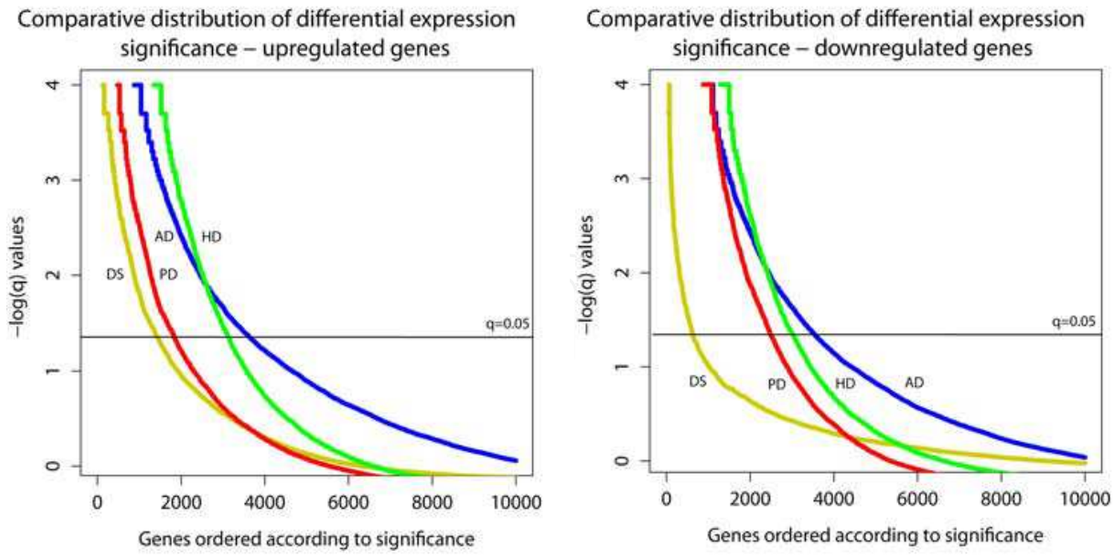


Fig. 1. Distribution of significance estimations for differential expression in 4 neurodegenerative disorders

An extent of global perturbation of the transciptome may be compared, with AD displaying the greatest extent of differentially expressed genes (blue line) and DS displaying the lowest extent, especially in the case of genes displaying downregulation.

4.1 Common patterns of differential expression in neurodegenerative disorders

Comparisons of comformity between profiles of transcriptome perturbations in four neurodegenerative diseases was initially performed by inspecting lists of top 1000 DE (differentially expressed) probesets for each condition and subsequently obtaining probesets (and genes) found to be differentially expressed simultaneously in several conditions.

The numbers of overlapping probesets may be observed in Figure 2. The largest overlap was observed between between the PD and HD lists, with altogether 338 (33.8%) upregulated and 267 (26.7%) downregulated genes differentially expressed in both conditions. Detailed overview of the extent of overlap between pairs of top DE gene list may be observed in Figure 3. A notable number of probesets was DE in all four conditions: 44 upregulated and 16 downregulated as presented in Figure 2a and 2b.

4.2 Comparative functional analyses of differential expression profile in neurodegenerative diseases

Calculations of gene set enrichment profile of upregulated and downregulated sets of genes presented here, were performed using hypergeometric test in the GOstats package. The profiles of DE genes were first calculated for each disorder separately, and afterwards every intersection between combinations of four sets of DE genes was evaluated.

Results of interests from separate GSEA analyses are presented in Table2(a-d) for top 1000 downregulated DE gene sets (the data for upregulated GSEA are not shown). Several GO biological process annotations appeared in all of the four analyses, most notably terms related to synaptic transmission and to cognitive processes.

We have also investigated the extent of similarity of GSEA profiles across four diseases. Top 200 enriched GO terms were inspected in each neurodegenerative disorder and compared for matching terms in pair with other three disorders. Greatest similarity was observed between GSEA terms annotating downregulated genes in all four disorders, which may be observed in more detail in Figure 4. As previously observed for overlapping genes, greatest overlap was observed between PD and HD GO profiles in the upregulated (40.0% overlap) and downregulated sets (59.5% overlap).

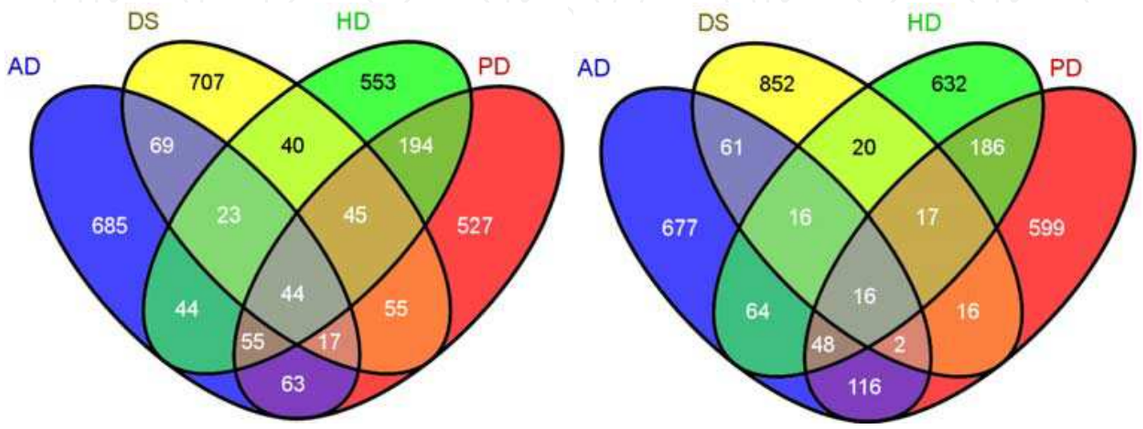


Fig. 2. Number of probesets overlapping between four sets of top 1000 DE upregulated (2a) and of top 1000 DE downregulated (2b) genes

Please note the abbreviations: Alzheimer disease (AD), Down syndrome (DS), Huntington disease (HD) and Parkinson disease (PD).

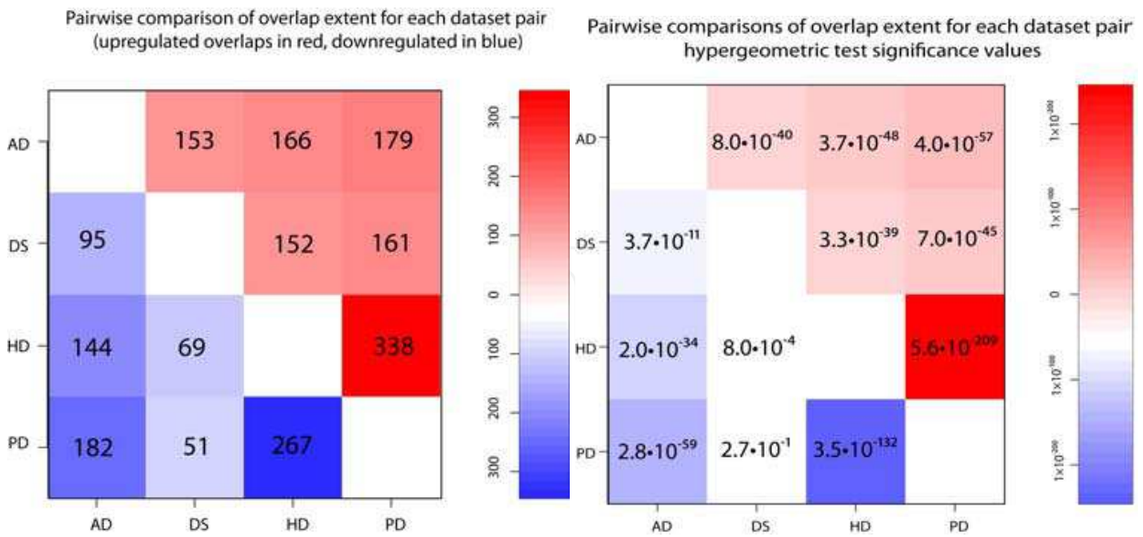


Fig. 3. a) Pairwise overlaps between lists of top DE upregulated (in red) and downregulated genes (in blue). Color intensiy of each square is proportional to size of overlap between a pair of DE gene lists. b) Pairwise overlaps between lists of top DE upregulated (in red) and downregulated genes (in blue). Color intensiy of each square is proportional to the value of –logp value obtained by performing hypergeometric test

GOBPID Accession	P-value	Count of genes annotated	Term
GO:0007268	1,01E-10	61	synaptic transmission
GO:0019226	1,49E-09	63	transmission of nerve impulse
GO:0035637	1,49E-09	63	multicellular organismal signaling
GO:0044282	2,61E-07	68	small molecule catabolic process
GO:0051443	5,09E-07	17	positive regulation of ubiquitin-protein ligase activity
GO:0019752	9,12E-07	70	carboxylic acid metabolic process
GO:0009144	1,82E-06	46	purine nucleoside triphosphate metabolic process
GO:0051438	5,03E-06	17	regulation of ubiquitin-protein ligase activity
GO:0007017	9,55E-06	35	microtubule-based process
GO:0007611	1,65E-05	18	learning or memory
GO:0030330	4,23E-05	16	DNA damage response, signal transduction by p53 class mediator
GO:0031398	4,70E-05	17	positive regulation of protein ubiquitination

Table 2a Alzheimer disease (downregulated genes). GOBPID stands for GeneOntology biological process ID

GOBPID Accession	P-value	Count of genes annotated	Term
GO:0007268	3,47E-37	90	synaptic transmission
GO:0019226	2,02E-35	93	transmission of nerve impulse
GO:0007267	5,12E-29	110	cell-cell signaling
GO:0007399	8,47E-15	113	nervous system development
GO:0007611	6,38E-13	25	learning or memory
GO:0007610	1,09E-12	46	behavior
GO:0050890	5,29E-12	25	cognition
GO:0048666	2,05E-10	59	neuron development
GO:0006836	2,24E-10	23	neurotransmitter transport
GO:0006811	9,51E-10	73	ion transport
GO:0001505	1,00E-09	21	regulation of neurotransmitter levels
GO:0031175	3,23E-09	52	neuron projection development
GO:0032940	9,52E-09	50	secretion by cell
GO:0048667	1,66E-08	46	cell morphogenesis involved in neuron differentiation
GO:0022008	3,13E-08	67	neurogenesis

Table 2b Huntington's disease (downregulated genes). GOBPID stands for GeneOntology biological process ID

GOBPID Accession	P-value	Count of genes annotated	Term
GO:0007268	4,16E-17	68	synaptic transmission
GO:0051234	1,99E-16	229	establishment of localization
GO:0019226	1,34E-15	70	transmission of nerve impulse
GO:0035637	1,34E-15	70	multicellular organismal signaling
GO:0006836	9,40E-14	29	neurotransmitter transport
GO:0009259	8,03E-13	58	ribonucleotide metabolic process
GO:0009144	1,47E-12	55	purine nucleoside triphosphate metabolic process
GO:0007399	8,20E-11	115	nervous system development
GO:0001505	8,41E-11	24	regulation of neurotransmitter levels
GO:0072521	9,91E-11	69	purine-containing compound metabolic process
GO:0007269	1,05E-10	20	neurotransmitter secretion
GO:0006753	2,95E-10	73	nucleoside phosphate metabolic process
GO:0009117	2,95E-10	73	nucleotide metabolic process
GO:0007267	3,23E-10	82	cell-cell signaling
GO:0015980	9,19E-10	39	energy derivation by oxidation of organic compounds

Table 2c Parkinson's disease (downregulated genes). GOBPID stands for GeneOntology biological process ID

GOBPID Accession	P-value	Count of genes annotated	Term
GO:0048856	6,65E-11	173	anatomical structure development
GO:0007267	2,83E-08	65	cell-cell signaling
GO:0050877	6,86E-07	69	neurological system process
GO:0050789	1,60E-06	318	regulation of biological process
GO:0022008	3,01E-06	58	neurogenesis
GO:0030182	3,71E-06	53	neuron differentiation
GO:0007399	4,91E-06	82	nervous system development
GO:0048839	8,26E-06	14	inner ear development
GO:0009887	1,69E-05	42	organ morphogenesis
GO:0007186	1,76E-05	41	G-protein coupled receptor protein signaling pathway
GO:0051716	2,70E-05	194	cellular response to stimulus
GO:0003001	4,34E-05	24	generation of a signal involved in cell-cell signaling
GO:0010903	5,15E-05	3	negative regulation of very-low-density lipoprotein particle remodeling
GO:0007268	7,44E-05	35	synaptic transmission
GO:0048667	7,44E-05	35	cell morphogenesis involved in neuron differentiation
GO:0007165	9,36E-05	165	signal transduction
GO:0048666	1,73E-04	41	neuron development

Table 2d Down's syndrome (downregulated genes). GOBPID stands for GeneOntology biological process ID

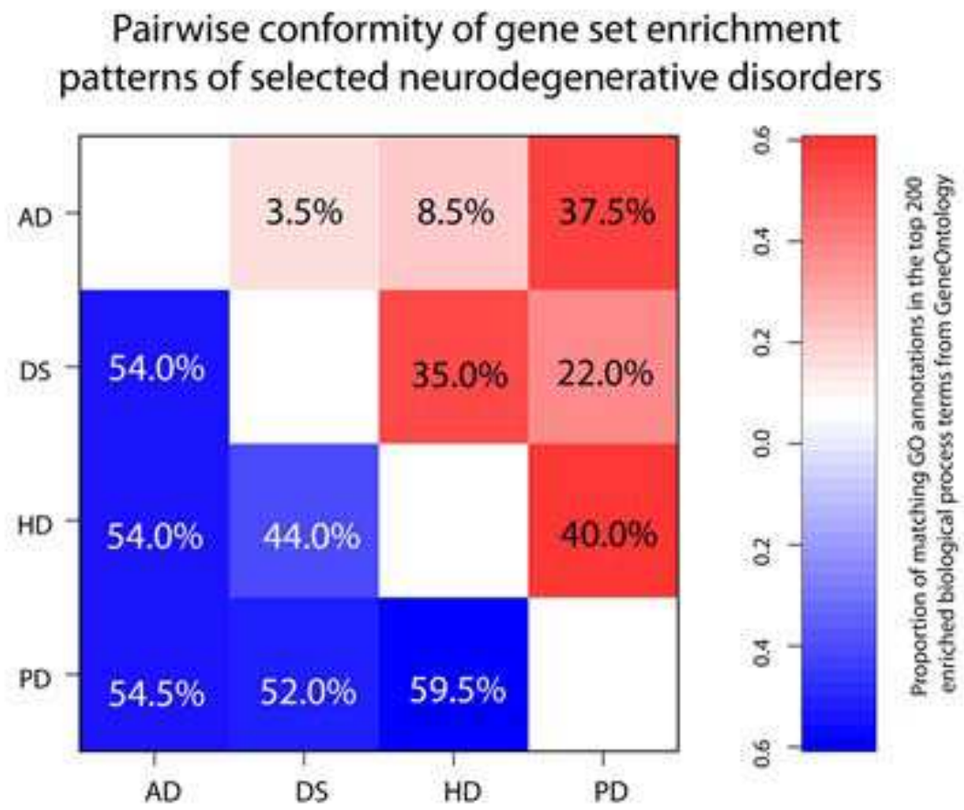


Fig. 4. Pairwise comparison of GO terms between pairs of datasets representing four neurodegenerative diseases. Percentages were calculated by dividing the number of GO terms overlapping by the number of all GO terms included in the overlapping analysis (N=200). GO terms annotating upregulated genes are presented in shades of red color and those annotating downregulated genes in blue

5. Conclusion

We have shown that whole-genome transcription analysis might be useful for identification and clarification of pathophysiological mechanisms in neurodegenerative diseases. We have used innovative approach of comparing and integrating experiment results from different NDG diseases and provided new important insights into the common NDG processes. Elucidation of these mechanisms holds important potential for future prediction and development of new useful treatments as well as for identification of biomarkers of neurodegeneration.

When comparisons of intersections between groups of top DE genes were performed, the greatest overlap was found between DE genes in brain samples of patients with HD and PD, which is possibly in accordance with their primary manifestation in movement disturbances related to function of basal ganglia. On the other hand, this similarity is surprising, as the known etiological agents in HD and PD differ significantly, one disorder being a consequence of monogenic disruption and other being a complex disorder with heterogeneous combination of genetic and environmental factors [41]. Surprisingly high is

also the profile overlap between AD and PD, which present as clinically somewhat distinct entities. Recently however, it has been becoming progressively more obvious that the two disorders share not only a significant proportion of clinical elements (movement disorder, cognitive decline, mood and psychiatric disorders) but also share common pathophysiological pathways [42]. These results potentially suggest that clinical distinction between disease entities may not be perfect projection of actual processes at cellular and molecular level. Additionally, in contrast to expectation, however, the lowest overlap was observed between samples from patients with DS and AD, especially as these conditions have been known to share NDG pathways related to amyloid beta deposition in neurons. Reasons for lower extent of overlap may be found in significant differences in the age of patients from whom the brain samples were obtained for studies of DS in comparison with AD. Additionally, it is important that in most instances, a complete triplication of genes located on chromosome 21 may dominate genes commonly dysregulated in DS and AD [29]. Also, the number of brain tissue samples samples profiled in microarray experiments was by far the lowest among other types of NDG diseases investigated in our survey. Therefore, before final answer regarding this finding is obtained, more studies investigating transcriptional alterations in DS brain samples must be performed.

Several GO categories appeared to be consistently singled out in GSEA analyses of separate and overlapping genes DE in NDG disorders. Interestingly several terms were related to processes previously associated with neuron degeneration [42], most prominently GO terms: synaptic transmission (GO:0007268), neurogenesis (GO:0022008) and terms related to higher cognitive processes (GO:0007611). Dysfunctional synaptic transmission (as in glutamate exitotoxicity) and defects in neurogenesis have been previously repeatedly shown to be related to various NDG diseases [42-44]. It is interesting that although disturbances in neuroinflammatory mechanisms have been proposed as a possible causative factor in a number of NDG diseases, our analysis of intersecting genes dysregulated in brain samples of these conditions did not single out a particular common inflammatory pathogenetic pathway. This notion may be interpreted in the light of previously recognized differences in complement-activating immunogenic activity of plaques in different NDG diseases, resulting in absence of commonly overlapping inflammatory genes and GO terms [42].

When we investigated the compatibility of functional profiles between four NDG diseases, we have found greatest overlaps between sets of GO terms annotating genes characterized by downregulation in NDG diseases, where an overlap greater than 40% was observed in all of the pairwise comparisons of the sets of top 200 enriched GO terms. Again, the greatest functional conformance was noted between top downregulated genes in HD and PD as well as AD and PD dataset pairs. Notable overlap was also observed in the functional profiles of upregulated genes, where we noted good functional conformity between DS and HD datasets in addition to HD-PD and AD-PD functional overlaps.

It is important to stress that genome-wide expression studies included in this survey are inherently burdened by important statistical issues that predominantly originate from the issue of testing a large number of variables on a relatively small population of biological replicates (ie. study subjects) [45]. For this reason we attempted to gain a more complete account of biological alterations in neurodegenerative diseases by merging data from several different studies investigating transcriptional changes in brain samples of distinct neurological conditions (AD, DS, HD and PD) [46]. This increased the number of biological replicates considerably, allowing for potentially more reliable calling of DE genes in these conditions. There are, however, important downsides to this approach: the studies included

were performed under differing conditions in different institutions and by different research staff. Even more important is the great heterogeneity between brain tissue samples investigated. We have attempted to circumvent these issues by using appropriate RankProd meta-analysis methods, nevertheless these results must be interpreted in light of these considerations.

Nevertheless it is still difficult to differentiate between the causal changes in transcriptome in contrast to changes resulting from previous damage to neural tissue. It is possible, however, that the similarities in transcriptome profile between clinically and pathologically distinct entities suggest a common response to an unknown initial damaging stimulus. We propose that in future, integration of various data such as genomic in combination with transcriptomic data should provide a way to delineate possible mechanisms, where genetic predisposition results in manifestation of transcriptional imbalances, consequently resulting in observed phenotype. Genome-wide expression profiling may however direct further research attempts into a particular direction. Also, there are other "omics" approaches besides transcriptomics and integrating all of them is future challenge.

6. References

[1] C.J.L. Murray ADL, *The Global Burden of Disease*. 1996, World Health Organisation: Geneva.

[2] Menken M, Munsat TL, Toole JF. The global burden of disease study: implications for neurology. *Arch Neurol* 2000; 57(3): p: 418-20.

[3] A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 1993; 72(6): p: 971-83.

[4] Schneider JA, Arvanitakis Z, Bang W, Bennett DA. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology* 2007; 69(24): p: 2197-204.

[5] Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet* 2001; 357(9251): p: 169-75.

[6] Prohovnik I, Perl DP, Davis KL, et al. Dissociation of neuropathology from severity of dementia in late-onset Alzheimer disease. *Neurology* 2006; 66(1): p: 49-55.

[7] Ballard C, Gauthier S, Corbett A, et al. Alzheimer's disease. *Lancet*; 377(9770): p: 1019-31.

[8] Forlenza OV, Diniz BS, Gattaz WF. Diagnosis and biomarkers of predementia in Alzheimer's disease. *BMC Med*; 8: p: 89.

[9] McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34(7): p: 939-44.

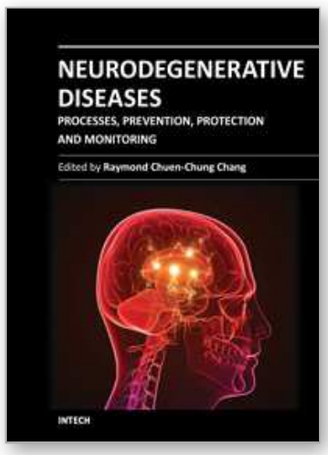
[10] Nussbaum RL, Ellis CE. Alzheimer's disease and Parkinson's disease. *N Engl J Med* 2003; 348(14): p: 1356-64.

[11] Spillantini MG, Schmidt ML, Lee VM, et al. Alpha-synuclein in Lewy bodies. *Nature* 1997; 388(6645): p: 839-40.

[12] Harper PS. Huntington's disease. Major Problems in Neurology. 1996, London: W.B. Saunders Company Ltd.

[13] Bonelli RM, Hofmann P. A systematic review of the treatment studies in Huntington's disease since 1990. *Expert Opin Pharmacother* 2007; 8(2): p: 141-53.
[14] Harjes P, Wanker EE. The hunt for huntingtin function: interaction partners tell many different stories. *Trends Biochem Sci* 2003; 28(8): p: 425-33.
[15] Sugars KL, Rubinsztein DC. Transcriptional abnormalities in Huntington disease. *Trends Genet* 2003; 19(5): p: 233-8.
[16] Holland AJ, Hon J, Huppert FA, Stevens F, Watson P. Population-based study of the prevalence and presentation of dementia in adults with Down's syndrome. *Br J Psychiatry* 1998; 172: p: 493-8.
[17] Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001; 69(3): p: 89-95.
[18] Henley SM, Bates GP,Tabrizi SJ. Biomarkers for neurodegenerative diseases. *Curr Opin Neurol* 2005; 18(6): p: 698-705.
[19] Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004; 5(10): p: R80.
[20] Barrett T, Edgar R. Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. *Methods Enzymol* 2006; 411: p: 352-69.
[21] Sean D, Meltzer PS. GEOquery: A bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 2007; 23(14): p: 1846-7.
[22] Liang WS, Dunckley T, Beach TG, et al. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol Genomics* 2007; 28(3): p: 311-22.
[23] Blalock EM, Geddes JW, Chen KC, et al. Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc Natl Acad Sci U S A* 2004; 101(7): p: 2173-8.
[24] Nunez-Iglesias J, Liu CC, Morgan TE, Finch CE, Zhou XJ. Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer's disease cortex reveals altered miRNA regulation. *PLoS One* 2010; 5(2): p: e8898.
[25] Moran LB, Duke DC, Deprez M, et al. Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. *Neurogenetics* 2006; 7(1): p: 1-11.
[26] Lesnick TG, Papapetropoulos S, Mash DC, et al. A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet* 2007; 3(6): p: e98.
[27] Hodges A, Strand AD, Aragaki AK, et al. Regional and cellular gene expression changes in human Huntington's disease brain. *Hum Mol Genet* 2006; 15(6): p: 965-77.
[28] Mao R, Wang X, Spitznagel EL, Jr., et al. Primary and secondary transcriptional effects in the developing human Down syndrome brain and heart. *Genome Biol* 2005; 6(13): p: R107.
[29] Lockstone HE, Harris LW, Swatton JE, et al. Gene expression profiling in the adult Down syndrome brain. *Genomics* 2007; 90(6): p: 647-60.
[30] Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. *Bioinformatics* 2009; 25(3): p: 415-6.
[31] Heber S,Sick B. Quality assessment of Affymetrix GeneChip data. *OMICS* 2006; 10(3): p: 358-68.
[32] Hong F, Breitling R, McEntee CW, et al. RankProd: A bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* 2006; 22(22): p: 2825-7.

[33] Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. *Bioinformatics* 2007; 23(2): p: 257-8.
[34] Ashburner M, Ball CA, Blake JA, et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; 25(1): p: 25-9.
[35] Ogata H, Goto S, Sato K, et al. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 1999; 27(1): p: 29-34.
[36] Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 2003; 34(3): p: 267-73.
[37] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; 4(1): p: 44-57.
[38] Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009; 4(8): p: 1184-91.
[39] Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* 2005; 21(16): p: 3439-40.
[40] Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 2004; 573(1-3): p: 83-92.
[41] Feany MB. ASIP Outstanding Investigator Award Lecture. New approaches to the pathology and genetics of neurodegeneration. *Am J Pathol* 2010; 176(5): p: 2058-66.
[42] Grunblatt E. Commonalities in the genetics of Alzheimer's disease and Parkinson's disease. *Expert Rev Neurother* 2008; 8(12): p: 1865-77.
[43] Palomo T, Archer T, Beninger RJ, Kostrzewa RM. Gene-environment interplay in neurogenesis and neurodegeneration. *Neurotox Res* 2004; 6(6): p: 415-34.
[44] Shankar GM, Walsh DM. Alzheimer's disease: synaptic dysfunction and Abeta. *Mol Neurodegener* 2009; 4: p: 48.
[45] Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 2006; 7(1): p: 55-65.
[46] Cahan P, Rovegno F, Mooney D, et al. Meta-analysis of microarray results: challenges, opportunities, and recommendations for standardization. *Gene* 2007; 401(1-2): p: 12-8.



Neurodegenerative Diseases - Processes, Prevention, Protection and Monitoring

Edited by Dr Raymond Chuen-Chung Chang

ISBN 978-953-307-485-6
Hard cover, 558 pages
Publisher InTech
Published online 09, December, 2011
Published in print edition December, 2011

Neurodegenerative Diseases - Processes, Prevention, Protection and Monitoring focuses on biological mechanisms, prevention, neuroprotection and even monitoring of disease progression. This book emphasizes the general biological processes of neurodegeneration in different neurodegenerative diseases. Although the primary etiology for different neurodegenerative diseases is different, there is a high level of similarity in the disease processes. The first three sections introduce how toxic proteins, intracellular calcium and oxidative stress affect different biological signaling pathways or molecular machineries to inform neurons to undergo degeneration. A section discusses how neighboring glial cells modulate or promote neurodegeneration. In the next section an evaluation is given of how hormonal and metabolic control modulate disease progression, which is followed by a section exploring some preventive methods using natural products and new pharmacological targets. We also explore how medical devices facilitate patient monitoring. This book is suitable for different readers: college students can use it as a textbook; researchers in academic institutions and pharmaceutical companies can take it as updated research information; health care professionals can take it as a reference book, even patients' families, relatives and friends can take it as a good basis to understand neurodegenerative diseases.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Luca Lovrečić, Aleš Maver and Borut Peterlin (2011). Genome Profiling and Potential Biomarkers in Neurodegenerative Disorders, Neurodegenerative Diseases - Processes, Prevention, Protection and Monitoring, Dr Raymond Chuen-Chung Chang (Ed.), ISBN: 978-953-307-485-6, InTech, Available from: http://www.intechopen.com/books/neurodegenerative-diseases-processes-prevention-protection-and-monitoring/genome-profiling-and-potential-biomarkers-in-neurodegenerative-disorders

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen

IntechOpen