

Gene expression profiling in human neurodegenerative disease

Johnathan Cooper-Knock, Janine Kirby, Laura Ferraiuolo, Paul R. Heath, Magnus Rattray and Pamela J. Shaw

Abstract | Transcriptome study in neurodegenerative disease has advanced considerably in the past 5 years. Increasing scientific rigour and improved analytical tools have led to more-reproducible data. Many transcriptome analysis platforms assay the expression of the entire genome, enabling a complete biological context to be captured. Gene expression profiling (GEP) is, therefore, uniquely placed to discover pathways of disease pathogenesis, potential therapeutic targets, and biomarkers. This Review summarizes microarray human GEP studies in the common neurodegenerative diseases amyotrophic lateral sclerosis (ALS), Parkinson disease (PD) and Alzheimer disease (AD). Several interesting reports have compared pathological gene expression in different patient groups, disease stages and anatomical areas. In all three diseases, GEP has revealed dysregulation of genes related to neuroinflammation. In ALS and PD, gene expression related to RNA splicing and protein turnover is disrupted, and several studies in ALS support involvement of the cytoskeleton. GEP studies have implicated the ubiquitin–proteasome system in PD pathogenesis, and have provided evidence of mitochondrial dysfunction in PD and AD. Lastly, in AD, a possible role for dysregulation of intracellular signalling pathways, including calcium signalling, has been highlighted. This Review also provides a discussion of methodological considerations in microarray sample preparation and data analysis.

Cooper-Knock, J. *et al.* *Nat. Rev. Neurol.* **8**, 518–530 (2012); published online 14 August 2012; doi:10.1038/nrneurol.2012.156

Introduction

Since it was first introduced in 1995,¹ the complementary DNA microarray has been an important tool in biomedical research for the identification of dysregulated biological pathways and, thereby, potential therapeutic targets. Large numbers of probes can be used simultaneously, which allows ‘capture’ of the biological context in health and disease. The amount of information generated by microarray analysis is particularly suited to certain specialist tasks such as biomarker discovery.

Recently, next-generation sequencing (NGS) as a means to quantify the transcriptome has become more widely available. As a new era arrives, the aim of this Review is to examine gene expression profiling (GEP) studies in human tissue over the past 5 years in amyotrophic lateral sclerosis (ALS), Alzheimer disease (AD) and Parkinson disease (PD).

The primary application of GEP has been the use of an oligonucleotide/cDNA microarray to quantify the transcriptome of a particular cell type or tissue. The selected cell or tissue is isolated from a patient or control case either postmortem or from accessible peripheral tissue during life. RNA is extracted, fluorescently labelled and then hybridized to the microarray. A linear amplification step is often required prior to labelling to ensure an adequate quantity of RNA. The amount of labelled RNA binding to each oligonucleotide/cDNA sequence on the

microarray determines the intensity of fluorescence at that location and thereby allows quantification of the RNA transcripts in the sample.

This Review summarizes findings from GEP studies in ALS, PD and AD. Studies in each disease are described in turn, with discussion divided into GEP in postmortem CNS tissue and peripheral tissue, and further subdivided according to whether studies were conducted in mixed-cell samples, or in single cell types isolated using laser capture microdissection (LCM). Technical considerations and relative merits of available methods in microarray work are also described, followed by consideration of the application of GEP studies to biomarker development.

Amyotrophic lateral sclerosis

ALS is a disease characterized by degeneration of upper and/or lower motor neurons in the motor cortex, brainstem and ventral spinal cord, although evidence exists for involvement of other areas of the CNS and non-neural tissues.² GEP studies in peripheral cells and postmortem tissue have confirmed the findings of other lines of research, and identified novel pathogenic mechanisms, including a promising therapeutic target.

Studies in CNS tissue

Mixed-cell samples

Since 2005, three studies^{3–5} have used postmortem mixed-cell samples from patients with ALS and controls (Table 1 and Supplementary Table 1 online). The two

Academic Unit of Neurology, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK (J. Cooper-Knock, J. Kirby, L. Ferraiuolo, P. R. Heath, M. Rattray, P. J. Shaw).

Correspondence to: P. J. Shaw
pamela.shaw@sheffield.ac.uk

Competing interests

The authors declare no competing interests.

studies of motor cortex both discovered a predominant downregulation of gene expression, which included functional gene groups associated with the cytoskeleton, protein turnover and neurotransmission.^{4,5} By contrast, studies of spinal cord tissue identified upregulation of genes related to the cytoskeleton, protein turnover and neurotransmission, and downregulation of stress response genes, including those involved in the antioxidant response and neuroinflammation.³ These results are in agreement with other lines of research implicating alterations in the cytoskeleton, protein turnover and inflammation in the pathogenesis of ALS.⁶

Laser capture microdissection cell samples

Four analyses^{7–10} used motor neurons extracted by LCM from postmortem spinal cord of patients with ALS and controls (Table 1 and Supplementary Table 1 online). One of these studies used an exon-level platform and identified aberrant splicing in genes associated with the cytoskeleton,⁹ which were found to be downregulated in another of the studies⁸—a result that is in accordance with conclusions from the mixed-cell studies. In view of the discovery of aberrant splicing, it is noteworthy that a number of the other studies reported differential expression related to the process of RNA transcription.^{3,4,8} Pathogenic mutations in two RNA-processing genes have recently been discovered in ALS.^{11,12} In addition, we have found that aberrant splicing occurs in fibroblasts from patients with ALS who are carriers of mutations in the ALS risk gene *TARDBP* (which encodes the RNA-splicing protein TAR DNA-binding protein 43) and, to a lesser extent, in fibroblasts from patients with sporadic ALS, but is virtually absent in fibroblasts from patients with mutations in the ALS risk gene superoxide dismutase 1 (*SOD1*; J. R. Highley, personal communication).

Two studies involved LCM of motor neurons from the spinal cord of patients with ALS who were carriers of a mutation in charged multivesicular body protein 2B (*CHMP2B*) or *SOD1*.^{7,10} In patients with a *CHMP2B* mutation, differential gene expression was identified in genes associated with the cytoskeleton, inflammation and protein turnover,¹⁰ consistent with the mixed-cell studies. Differential gene expression identified in motor neurons from the spinal cord of patients with *SOD1* mutations⁷ showed concordance with the LCM study of sporadic ALS⁸ and with transcriptome studies of mutant *SOD1* models.^{13,14} This study⁷ highlighted altered expression of genes associated with the antiapoptotic signalling pathway involving phosphatidylinositol 3-kinase and protein kinase B (AKT1), such that activation of AKT1 is enhanced with concomitant reduced expression of the negative regulator of this pathway, phosphatase and tensin homologue (*PTEN*). The researchers hypothesized that these changes were observed because motor neurons extracted from postmortem tissue were those that survived the disease process. In support of this suggestion, reduced expression of *PTEN* in a motor neuron cell line and a primary motor neuron culture, both expressing mutant *SOD1*, increased cell survival.⁷ This study is a

Key points

- Gene expression profiling (GEP) has advanced considerably over the past 5 years, and has provided important insight into mechanisms underlying neurodegenerative disease
- In amyotrophic lateral sclerosis, GEP studies have consistently implicated certain biological structures and pathways, including the cytoskeleton, inflammation, protein turnover and RNA splicing
- GEP studies in Parkinson disease have highlighted dysfunction of the ubiquitin-proteasome system, RNA splicing, mitochondrial function and inflammation
- In Alzheimer disease, affected pathways identified by GEP analysis include neuroinflammation, mitochondrial function and calcium signalling
- GEP studies have investigated selective vulnerability to neurodegeneration between patients and in different anatomical areas of the CNS, in order to characterize disease mechanisms, identify therapeutic targets and potentially inform development of individualized treatments
- Technical aspects of GEP, including sample preparation, data analysis and validation, require careful consideration to optimize assays and yield reliable results

clear example of a GEP study leading to identification of a potential therapeutic target.

Studies in peripheral tissue

Since 2005, five studies involving GEP of peripheral cells in ALS have been conducted.^{2,15–18} GEP studies of whole blood from patients with ALS (Table 1 and Supplementary Table 1 online) used clustering analysis to look for genes with a similar pattern of expression, before determining which of these clusters showed altered expression in disease.² Interestingly, differentially expressed clusters exhibited better functional enrichment than a similar number of the most differentially expressed individual genes by *P*-value—that is, the differentially expressed clusters contained genes that were more functionally similar to each other and were, therefore, less likely to be false positives. This point illustrates how the interrelated nature of gene expression can be used to improve accuracy in GEP.

A study of purified lymphocytes from patients with ALS identified disease-specific differential expression of genes including *TARDBP*.¹⁶ This finding and the previously described study in whole blood² suggest that peripheral blood is a viable medium for the study of ALS. In addition, the findings of both of these studies were comparable to those of the CNS studies, including dysregulation of genes associated with protein processing, RNA post-transcriptional modification, and inflammation.^{2,16} Validation experiments using a proteasome inhibition assay in peripheral blood mononuclear cells from patients with ALS showed that expression of the proteasome-associated ubiquitin-protein ligase E3 component n-recogin 2 (*UBR2*) gene directly correlates with the degree of physical disability.¹⁶ Peripheral lymphocytes might, therefore, provide a functional assay for drug development and a biomarker of disease progression.

Two GEP studies have been conducted in tissue obtained by muscle biopsy from patients with ALS.^{17,18} Transcriptome changes in muscle were largely distinct from those in blood and neuronal tissue. Expression of a 198-gene panel correlated with severity of degeneration in the biopsied muscle, and the researchers suggested that

Table 1 | Selected gene expression profiling studies in patients with ALS

Study	Patients	Cell sample
Wang <i>et al.</i> (2006) ⁴	5 ALS; 3 controls	Mixed-cell samples from primary motor and sensory cortex
Lederer <i>et al.</i> (2007) ⁵	11 ALS; 9 controls	Mixed-cell samples from motor cortex
Offen <i>et al.</i> (2009) ³	4 ALS; 4 controls	Mixed-cell samples from cervical spinal cord
Jiang <i>et al.</i> (2005) ⁸	14 ALS; 13 controls	LCM of motor neurons from lumbar spinal cord
Rabin <i>et al.</i> (2010) ⁹	12 ALS; 10 controls	LCM of motor neurons from lumbar spinal cord
Cox <i>et al.</i> (2010) ¹⁰	3 ALS with <i>CHMP2B</i> mutation; 7 controls	LCM of motor neurons from cervical spinal cord
Kirby <i>et al.</i> (2011) ⁷	3 ALS with <i>SOD1</i> mutation; 7 controls	LCM of motor neurons from cervical spinal cord
Saris <i>et al.</i> (2009) ²	123 ALS; 123 controls	Venous blood
Zhang <i>et al.</i> (2011) ¹⁵	20 ALS; 22 controls	Mononuclear cells from venous blood
Mougeot <i>et al.</i> (2011) ¹⁶	11 sporadic ALS; 11 controls	Lymphocytes purified from venous blood
Pradat <i>et al.</i> (2012) ¹⁷	5 late ALS; 4 early ALS; 10 controls	Myocytes from deltoid muscle
Shtilbans <i>et al.</i> (2011) ¹⁸	3 ALS; 3 MMN; 3 controls	Myocytes

Abbreviations: ALS, amyotrophic lateral sclerosis; *CHMP2B*, charged multivesicular body protein 2B; LCM, laser capture microdissection; MMN, multifocal motor neuropathy; *SOD1*, superoxide dismutase 1. For further details, see Supplementary Table 1 online.

this panel could be an effective biomarker for measuring disease progression.¹⁷ An important question is whether these changes represent the primary disease process or a downstream effect, the answer to which will determine how well this panel is able to quantify severity of disease beyond the biopsied muscle, and whether it can differentiate ALS from other causes of muscle wasting.

Parkinson disease

The main pathological feature of PD is selective degeneration of dopaminergic neurons of the substantia nigra pars compacta. Evidence also exists, however, for more widespread involvement of the CNS and other tissues.¹⁹ In PD, GEP studies have repeatedly implicated dysfunction of mitochondria and protein processing, consistent with other research; moreover, GEP has produced novel insights into the mechanism of this dysfunction and a potential biomarker of disease.

In a landmark paper, Zheng *et al.*²⁰ conducted a biological-pathway-level comparison of 17 microarray studies, an approach that goes some way towards overcoming the interstudy variability of gene-level analyses.²¹ 10 biological pathways were initially identified as being differentially expressed in postmortem GEP studies of the substantia nigra of patients with PD. The identified pathways were confirmed in GEP studies of non-nigral tissue, including other brain areas and antemortem peripheral tissue, and in a study of patients with subclinical Lewy body pathology of the substantia nigra. The 10 pathways identified included genes controlled by peroxisome proliferator-activated receptor- γ coactivator 1 α (PGC1 α), a master regulator of mitochondrial function. Activation of PGC1 α was demonstrated to ameliorate the phenotype in cell models of PD. Other highlighted pathways included mitochondrial function and pyruvate metabolism, consistent with an energy deficit and attempted compensation. Parkin inactivation, which is associated with familial and sporadic PD, has subsequently been shown to cause repression of PGC1 α expression, thereby validating the results of this GEP study.²²

Studies in CNS tissue

Mixed-cell samples

Since 2005, 10 groups^{23–32} have conducted GEP studies in postmortem mixed-cell samples from various brain areas, including parts of the basal ganglia, in patients with PD (Table 2 and Supplementary Table 2 online). Seven of the studies identified PD-associated differential gene expression in brain areas including the substantia nigra, and showed good consensus relating to key gene expression changes,^{23–26,30–32} particularly with regard to dysregulation of protein processing and mitochondrial pathways. GEP analysis of 21 brain areas related to PD revealed that gene expression changes related to mitochondrial function occur throughout the brain, to varying degrees among the different regions.²³

Duke and colleagues³³ compared gene expression between the medial and lateral substantia nigra, in order to investigate the relative susceptibility of the lateral substantia nigra in PD.³⁴ In the lateral substantia nigra of control and PD cases, proinflammatory genes and genes encoding components of mitochondrial complex I were upregulated, and genes involved in glutathione synthesis and function were downregulated, compared with the corresponding medial regions. The researchers suggested that increased energy demand and lack of glutathione function would render neurons in the lateral substantia nigra more susceptible to oxidative stress, a mechanism that has been strongly implicated in the pathophysiology of PD.³⁵ Similarly, Bossers *et al.* selectively studied relatively spared areas of the substantia nigra, as determined by neuronal density.³¹ As well as confirming dysregulation of mitochondrial function and protein-processing genes, they highlighted involvement of biological pathways related to neurotrophic signalling and axon guidance. Both of these studies have identified potential therapeutic targets for PD by characterizing selective neuronal vulnerability in this disease.

Implication of dysregulation of protein processing and mitochondrial pathways is consistent with other research on the pathophysiology of PD.^{36,37} Several disease-causing

Table 2 | Selected gene expression profiling studies in patients with PD

Study	Patients	Cell sample
Moran <i>et al.</i> (2006) ²⁴	15 PD; 1 multiple sclerosis; 7 controls	Mixed-cell sample from lateral and medial SN and superior frontal gyrus
Vogt <i>et al.</i> (2006) ²⁷	8 PD; 8 multiple system atrophy; 8 controls	Mixed-cell sample from putamen, cerebellum and occipital cortex
Hauser <i>et al.</i> (2005) ²⁵	6 PD; 2 progressive supranuclear palsy; 1 frontotemporal dementia with parkinsonism; 5 controls	Mixed-cell sample from SN and surrounding midbrain
Naydenov <i>et al.</i> (2010) ²⁸	15 PD with dyskinesia; 16 PD without dyskinesia; 32 controls	Mixed-cell sample from putamen
Zhang <i>et al.</i> (2005) ²⁶	15 PD; 15 controls	Mixed-cell sample from SN, putamen and Brodmann area BA9
Botta-Orfila <i>et al.</i> (2012) ²⁹	5 idiopathic PD; 3 LRRK2-PD; 1 asymptomatic LRRK2 mutation carrier; 5 controls	Mixed-cell sample from putamen
Durrenberger <i>et al.</i> (2012) ³⁰	12 idiopathic PD; 7 controls	Mixed-cell sample from SN _{pc}
Miller <i>et al.</i> (2006) ³²	6 PD; 8 controls	Mixed-cell sample from SN and/or striatum
Papapetropoulos <i>et al.</i> (2006) ²³	22 PD; 23 controls	Mixed-cell sample from various brain areas
Bossers <i>et al.</i> (2009) ³¹	4 PD; 4 controls	Mixed-cell sample from spared areas of SN
Elstner <i>et al.</i> (2011) ²¹	11 PD; 11 age-matched controls; 8 young controls	LCM of dopaminergic neurons from SN _{pc}
Cantuti-Castelvetri <i>et al.</i> (2007) ⁴⁴	8 PD; 8 controls	LCM of dopaminergic neurons from SN _{pc}
Simunovic <i>et al.</i> (2009) ⁴⁵	10 PD; 9 controls	LCM of dopaminergic neurons from SN _{pc}
Stamper <i>et al.</i> (2008) ⁴⁶	13 PD with dementia; 15 PD without dementia; 14 controls	LCM of layer V–VI pyramidal neurons from posterior cingulate
Scherzer <i>et al.</i> (2007) ¹⁹	50 PD; 55 healthy and disease controls	Venous blood
Shehadeh <i>et al.</i> (2010) ⁴⁹	17 PD; 11 controls	Venous blood
Mutez <i>et al.</i> (2011) ⁵²	10 LRRK2-PD; 1 asymptomatic LRRK2 mutation carrier; 7 controls plus sample from 40 pooled controls	Monocytes removed from whole blood
Matigian <i>et al.</i> (2010) ⁵¹	19 PD; 9 schizophrenia; 14 controls	Olfactory-neurosphere-derived cells
Mar <i>et al.</i> (2011) ⁵⁰	13 PD; 9 schizophrenia; 11 controls	Olfactory-neurosphere-derived cells

Abbreviations: LCM, laser capture microdissection; LRRK2, leucine-rich repeat kinase 2; PD, Parkinson disease; SN, substantia nigra; SN_{pc}, substantia nigra pars compacta. For further details, see Supplementary Table 2 online.

mutations in PD impair mitochondrial complex I function³⁸ or are part of the ubiquitin–proteasome system.³⁹ Moreover, the susceptibility gene *DJ1* encodes a chaperone protein that is also involved in proteolytic stress.⁴⁰ In fact, many genes and associated pathways implicated in familial PD are differentially expressed in the substantia nigra of sporadic PD cases compared with controls.⁴¹

Duke *et al.*⁴² used a clustering technique to show that downregulation of mitochondrial and ubiquitin–proteasomal gene clusters correlate with each other and with clinical phenotype, suggesting a close relationship between impairments of these two systems in PD. These pathways could, therefore, contain a common therapeutic target. By contrast, a study comparing gene expression in the putamen of PD patients with either a mutation in leucine-rich repeat kinase 2 (*LRRK2*) or idiopathic PD concluded that the transcriptome, and thus the pathogenesis, of *LRRK2*-associated PD was distinct, despite evidence that *LRRK2* is involved in mitochondrial function.⁴³ The study involved only a small number of cases, however, which limited the description of the pathways involved in *LRRK2*-associated PD.

Laser capture microdissection cell samples

Four studies^{21,44–46} used LCM to study populations of either dopaminergic or pyramidal neurons in post-mortem tissue from patients with PD (Table 2 and Supplementary Table 2 online). Two studies were comparable to and broadly in agreement with the mixed-cell studies: they found disease-specific downregulation of genes related to mitochondrial function and the ubiquitin–proteasome system, and dysregulation of many of the genes implicated in familial PD.^{21,45} One study showed that these changes were specific to PD and did not occur in normal ageing.²¹ In addition, this study examined known biological interactions between genes that are differentially expressed in PD, in order to identify a disease-associated gene network. The genes included those related to energy metabolism and nutrient deprivation, consistent with the findings of Zheng and colleagues.²⁰

Two studies^{44,47} investigated the basis for male susceptibility to PD in the transcriptome of dopaminergic neurons of the substantia nigra pars compacta. Both found upregulation of genes related to mitochondrial

function in male controls compared with female controls, consistent with evidence from other studies of an accelerated metabolic rate in male dopaminergic neurons. This feature might predispose the neurons to development of PD,⁴⁸ highlighting a potential target for protection against the disease. Notably, this suggestion is similar to one made by Duke and colleagues, that increased susceptibility of the lateral substantia nigra is attributable to an energy deficit.³³ Moreover, a sex-specific comparison of gene expression in PD found enrichment of similar functional categories of genes in both sexes, but the specific genes involved in each sex showed little overlap, suggesting that mechanisms of pathogenesis may differ between the sexes.⁴⁷

GEP analysis of cortical neurons in PD⁴⁶ indicated that development of dementia in this disease involves progressive aberrant expression of genes associated with alternative splicing. This result is supported by those of two mixed-cell studies, which implicated dysregulation of RNA processing in the putamen²⁷ and substantia nigra²⁴ in PD.

Studies in peripheral tissue

A GEP study in peripheral blood from a large number of patients with sporadic PD aimed to generate a gene signature for this disease (Table 2 and Supplementary Table 2 online).¹⁹ Samples were predominantly taken from patients with early-stage disease and compared with control samples from healthy individuals and patients with other neurodegenerative diseases. The patient groups were chosen to facilitate development of a biomarker for diagnosis in early PD. A molecular marker consisting of eight genes (*VDR*, *HIP2*, *CLTB*, *FPRL2*, *CA12*, *CEACAM4*, *ACRV1* and *UTX*) was then validated in an independent test set of blood samples from a different group of patients.¹⁹ Another GEP study in peripheral blood used exon-level probes,⁴⁹ and showed altered transcript splicing in venous blood from patients with PD. The researchers suggested this result could be related to altered expression of *SRRM2*—a splicing factor that was found to be differentially expressed in a previous study.¹⁹

Two studies have evaluated the use of olfactory neurosphere-derived cells obtained via biopsy of patient nasal mucosa.^{50,51} This novel source of peripheral cells showed GEP changes in neurological disease, suggesting that it might constitute a viable peripheral model of CNS disease. A comparison with fibroblasts showed more functionally uniform transcriptome changes in the olfactory neurosphere-derived cells,⁵¹ suggesting that there might be less 'noise' associated with GEP in these cells. In PD, the neurosphere-derived cells showed GEP changes similar to some of those obtained in the CNS studies, including downregulation of glutathione-related genes⁵¹ and dysregulation of neurotrophic signalling.⁵⁰

GEP analysis of peripheral blood mononuclear cells from *LRRK2*-PD cases found dysregulation of similar pathways to those identified in CNS studies in idiopathic PD, including mitochondrial function and the ubiquitin–proteasome system.⁵² This result conflicts with that of another study, which suggested that *LRRK2*-PD

is distinct,²⁹ but a direct comparison is difficult because the peripheral blood work did not include samples from patients with idiopathic PD.

Alzheimer disease

AD is the most common neurodegenerative disease and is characterized by progressive dementia, initially presenting with short-term memory impairment. GEP studies in AD have identified dysfunction of mitochondrial activity, intracellular signalling and neuroinflammation across different tissues and brain areas. Elegant practical and statistical work has further developed these findings towards therapeutic targets and potential biomarkers.

Studies in CNS tissue

Mixed-cell samples

Since 2005, 16 GEP studies of mixed-cell samples from postmortem tissue in AD have been published (Table 3 and Supplementary Table 3 online).^{53–68} Broadly speaking, these studies have investigated two anatomical regions: the temporal lobe–hippocampus and the frontal–prefrontal cortex. Two studies highlighted that these areas show the greatest number of differentially expressed genes in AD.^{59,60} Interestingly, the greatest increase in aberrant gene expression occurs during progression from mild to moderate dementia,⁵⁸ suggesting that therapeutic strategies should be implemented early in the disease course.

Functional categories of genes identified as being differentially expressed in AD are numerous and varied, but some common themes arise: notably, intracellular signalling pathways—particularly calcium signalling^{54,56,57,60,62,66}—and neuroinflammation.^{53,54,56,60,66,67}

Other lines of evidence implicate disturbance of calcium signalling in AD: for example, amyloid- β plaques disrupt calcium signalling within neurons,⁶⁹ and presenilin-1 mutations cause abnormal accumulation of calcium in neuronal endoplasmic reticulum.⁷⁰ Disruption of calcium signalling has also been linked to other proposed pathological mechanisms in AD, including mitochondrial dysfunction, and may represent an upstream therapeutic target.⁷¹

Age-related differences exist in the dysregulation of pathways identified, including decreased expression, in older patients with AD, of neuroinflammatory genes that are upregulated in older controls.⁶⁰ This difference may represent a neuroprotective response that fails in the development of AD. Such a hypothesis is consistent with studies of normal ageing: GEP of mixed-cell cortical tissue from cognitively normal individuals demonstrated changes in neuroinflammation-associated gene expression with increasing age.⁷² Similarly, changes in calcium signalling-related gene expression that occur with ageing are accelerated in AD.⁷³ GEP analysis in normal ageing gives context to transcriptome changes in neurodegenerative disease: the transcriptome in normal ageing is dynamic, not static, and needs to be understood to correctly interpret changes in disease. Applying this correctly can identify potential therapeutic strategies, such as blockade of excessive upregulation of calcium signalling genes and reconstitution of age-related neuroinflammation.

Table 3 | Selected gene expression profiling studies in patients with AD

Study	Patients	Cell sample
Xu <i>et al.</i> (2006) ⁵⁶	16 AD (Braak stage 4–5); 4 controls	Mixed-cell sample from hippocampal areas CA1–4
Parachikova <i>et al.</i> (2007) ⁵³	10 AD (MMSE 17–22, Braak stage 4–5); 14 controls	Mixed-cell sample from hippocampus and prefrontal cortex
Emilsson <i>et al.</i> (2006) ⁵⁷	61 AD (CERAD-positive); 53 controls	Mixed-cell samples from Brodmann areas BA8 and BA9
Katsel <i>et al.</i> (2009) ⁶⁰	Aged <87 years: 6 mild AD (CDR 0.5–1.0); 13 severe AD (CDR 4–5); 15 controls Aged ≥87 years: 15 mild AD; 15 severe AD; 7 controls	Mixed-cell samples from 14 cortical areas and hippocampus
Haroutunian <i>et al.</i> (2009) ⁵⁸	104 mild to severe AD (CDR 0.5–5.0); 26 controls	Mixed-cell samples from 14 cortical areas and hippocampus
Katsel <i>et al.</i> (2007) ⁵⁹	98 AD (CERAD-positive) with mild to severe dementia (CDR 0.5–5.0); 19 controls	Mixed-cell samples from 14 cortical areas, hippocampus, caudate and putamen
Umemura <i>et al.</i> (2006) ⁶¹	7 AD; 3 amyotrophic lateral sclerosis	Mixed-cell sample from frontal lobe
Weeraratna <i>et al.</i> (2007) ⁶²	6 AD (Braak stage ≥5); 6 non-Alzheimer dementia (Braak stage ≤3); 6 controls	Mixed-cell sample from inferior parietal lobe
Tan <i>et al.</i> (2010) ⁵⁴	12 AD (CERAD-positive, CAMCOG <80, Braak stage >3); 8 controls	Mixed-cell sample from temporal cortex
Bronner <i>et al.</i> (2009) ⁶³	5 AD (Braak stage 6); 5 progressive supranuclear palsy; 5 Pick disease; 5 frontotemporal dementia; 5 controls	Mixed-cell sample from medial temporal cortex
Youn <i>et al.</i> (2007) ⁶⁶	19 AD; 15 controls	Mixed-cell sample from hippocampus and cerebellum
Bossers <i>et al.</i> (2010) ⁶⁴	7 for each Braak stage 0–6	Mixed-cell sample from prefrontal cortex
Horesh <i>et al.</i> (2011) ⁶⁵	55 AD; 28 schizophrenia; 22 controls	Mixed-cell samples from frontal lobe, cingulate, temporal cortex, parietal cortex, occipital cortex and basal ganglia
Williams <i>et al.</i> (2009) ⁵⁵	6 early AD (MMSE 21–26); 8 controls	Synaptoneuroosomes isolated from prefrontal cortex
Tollervey <i>et al.</i> (2011) ⁶⁸	6 AD; 7 FTLT-TDP; 3 FTLT-tau; 9 controls	Mixed-cell sample from temporal cortex
Wang <i>et al.</i> (2012) ⁶⁷	12 AD; 12 controls	Pooled microvessels from temporal, parietal and frontal cortex
Liang <i>et al.</i> (2008) ⁷⁶	33 AD (Braak stage 3–4, CERAD plaque density moderate–frequent); 14 controls	LCM of unaffected neurons from entorhinal cortex, hippocampus, medial temporal cortex, posterior cingulate, superior frontal gyrus and visual cortex
Dunckley <i>et al.</i> (2006) ⁷⁸	19 AD; 14 controls	LCM of entorhinal cortical neurons with or without neurofibrillary tangles
Simpson <i>et al.</i> (2011) ⁸⁶	6 AD for each Braak stage group 0–2, 3–4, 5–6	LCM of astrocytes from lateral temporal cortex
Maes <i>et al.</i> (2007) ⁸⁸	14 mild AD; 14 controls	Monocytes from whole blood
Nagasaka <i>et al.</i> (2005) ⁸⁹	21 familial AD from three families; 12 wild-type siblings	Cultured fibroblasts from skin biopsy
Fehlbaum-Beurdeley <i>et al.</i> (2010) ⁹³	80 AD; 70 controls	Venous blood
Booij <i>et al.</i> (2011) ⁹¹	Training set: 94 AD; 94 controls Test set: 31 AD; 25 age-matched controls; 7 young controls; 27 PD	Venous blood
Calciano <i>et al.</i> (2010) ⁹⁰	28 AD receiving no treatment or donepezil, galantamine or rivastigmine	Venous blood
Kalman <i>et al.</i> (2005) ⁹²	8 AD (MMSE 16.0 ± 5.1); 8 controls	Lymphocytes from venous blood
Chen <i>et al.</i> (2011) ⁹⁴	5 AD; 4 mild cognitive impairment; 4 controls	Lymphocytes from venous blood

[†]CERAD-positive[†] describes patients who met CERAD neuropathological criteria for a diagnosis of AD. Abbreviations: AD, Alzheimer disease; CAMCOG, Cambridge Cognitive Examination; CDR, Clinical Dementia Rating; CERAD, Consortium to Establish a Registry for Alzheimer Disease; FTLT, frontotemporal lobar degeneration; LCM, laser capture microdissection; MMSE, Mini-Mental State Examination; TDP, TAR DNA-binding protein. For further details, see Supplementary Table 3 online.

Several GEP studies have discovered aberrant expression of synapse-related genes in AD, and have started to define the mechanisms involved. The neuropathologically defined Braak stages of AD have been used as a proxy for

disease course in identification of gene clusters that show a consistent change in expression with increasing Braak stage.⁶⁴ Synapse-related genes were upregulated in low Braak stages and downregulated in higher Braak stages,

suggesting a neuroprotective response in early disease and a therapeutic target in later disease. Further developing the study of synaptic dysfunction in AD,⁷⁴ GEP of the synaptoneurosome⁵⁵ revealed aberrant gene expression related to synaptic function, and found that expression of the glutamate receptor 2 (*GRIA2*) gene in synaptoneurosome, but not in whole-cell homogenates, correlated with declining cognition. This interesting finding supports other work suggesting that axons and nerve terminals might display a distinct transcriptome from the neuronal cell body, by virtue of transfer of mRNA to the neuron from neighbouring glial cells.⁷⁵

An exon-level study of temporal cortex from patients with AD or frontotemporal lobar degeneration (FTLD) and controls⁶⁸ demonstrated alternative splicing associated with AD and FTLD. An enrichment analysis indicated that the transcriptome changes might be related to increased activity of polypyrimidine tract binding protein 1 (PTBP1) and reduced activity of neuro-oncological ventral antigen 1 (NOVA1), two RNA-splicing proteins that represent potential therapeutic targets. Although AD and FTLD are clinically and pathologically distinct, this analysis was unable to differentiate the two conditions. Nevertheless, the findings are interesting and should be developed with a larger number of samples.

Laser capture microdissection cell samples

LCM studies in AD have uncovered changes in gene expression that were not apparent from mixed-cell studies (Table 3 and Supplementary Table 3 online). A series of studies^{76–78} investigated the transcriptome of neurons from six cortical areas that were chosen to represent different stages of AD according to the pattern of disease spread. The researchers showed that the entorhinal cortex and hippocampus—the two areas that are most susceptible to accumulation of neurofibrillary tangles—shared differential expression related to glycolysis, which could reflect an increased energy demand.⁷⁶ Consistent with this finding, another study showed that expression of essential components of mitochondrial function is downregulated in brain areas affected by AD.⁷⁷ Interestingly, the same brain regions show a reduced metabolic rate, as measured by fluorodeoxyglucose PET, in AD.^{79,80} Provision of support to meet the energy demand in these areas could be a neuroprotective strategy.

Several studies have developed these microarray data sets using pathway-based approaches.^{81–83} One report used a coexpression network analysis⁸³ to demonstrate dysregulation of genes that are controlled by transcription factors involved in cardiovascular disease⁸³—an association that is the subject of ongoing research.^{84,85} Notably, GEP analysis of pooled microvessels from various brain areas in AD identified similar transcriptome changes to those seen in other postmortem AD studies.⁶⁷ These findings indicate that targeting of cardiovascular risk factors might be a useful therapy for AD.

Transcriptome analysis of astrocytes from lateral temporal cortex in patients with AD explored patterns of differential expression with increasing Braak stage.⁸⁶ Findings included dysregulation of pathways related to

intracellular signalling and the cytoskeleton, which could represent progressive dysfunction of astrocyte connectivity during disease progression. Notably, these transcriptome changes occurred at a lower Braak stage in patients carrying the apolipoprotein E ϵ 4 allele, which is a risk allele for AD.⁸⁷ Disrupted signalling pathways included calcium signalling, suggesting that aberrant calcium signalling is not limited to or even primarily within neurons. An important role for astrocytes in AD pathogenesis might explain synaptoneurosome dysfunction in postmortem AD tissue,⁵⁵ and support of astrocyte function could have therapeutic potential.

Studies in peripheral tissue

Seven GEP studies of peripheral samples from patients with AD have been carried out since 2005.^{88–94} Study of blood mononuclear cells⁸⁸ and peripheral leukocytes^{92,94} highlighted involvement of similar functional categories of genes to those identified in studies of CNS tissue, including aberrant gene expression related to neuroinflammation and intracellular signalling. Such consistency supports the suitability of peripheral cells for studies of AD pathogenesis and for biomarker development. Expression of ATP-binding cassette subfamily B member 1 (*ABCB1*) in peripheral leukocytes correlated with scores of cognitive ability in patients with AD or mild cognitive impairment and in controls.⁹⁴ Further work is needed to characterize the relationship between *ABCB1* function and AD pathogenesis, but expression levels of this protein may represent a biomarker for AD.

Three studies^{89,91,93} identified transcriptome biomarkers that can differentiate between disease and control samples. Two of the studies used an independent test set of samples to validate their panel, which was originally developed in a distinct training set,^{91,93} and one group demonstrated that their biomarker was independent of medication use and disease severity, and could classify AD and mild cognitive impairment cases according to eventual diagnosis.⁹¹ A subset of this gene panel was further validated using a customized reverse transcription PCR (rtPCR) array⁹⁵ to illustrate efficacy of the biomarker in a platform that is more suited than a microarray to application on a larger scale in the clinic. Of note, one of the components of this panel is *SORL1*, which was originally shown to be downregulated in a GEP study of lymphocytes from patients with AD.⁹⁶ Genetic variation in *SORL1* is associated with risk of AD, and the *SORL1* protein directs amyloid- β trafficking.⁹⁷

Technical considerations

The processes of sample acquisition and preparation before microarray work and subsequent data analysis (Figure 1) can have a profound impact on the results of GEP studies. Over time, consensus has developed on a number of these processes, as presented below.

Sample preparation

The ability to use postmortem tissue for microarray studies is essential, because the CNS is largely inaccessible before death. However, the extent to which neurons

obtained at the end stage of disease, which have survived the disease process, are representative of the disease in general is not known.⁹⁸ Postmortem specimens from patients with neurodegenerative disease are often collected as homogenized tissue from a specific brain or spinal cord area, containing multiple cell types that might each reflect the underlying pathophysiology to varying degrees. Gene expression in numerous cell types can dilute and potentially mask changes in less-numerous cell types.⁸ LCM allows isolation of individual cells by carbon dioxide laser pulse under light microscopy,⁹⁹ which partially overcomes this problem.^{8,21} Another potential problem with sample preparation is that case–control comparisons of mixed-cell tissue in neurodegenerative disease are likely to compare a neuron-rich sample with a sample depleted of neurons, unless a suitable correction is made.¹⁰⁰

RNA quality can affect the results of a transcriptome analysis. Freezing of RNA samples at -80°C or storage of samples in a specialized 'RNAlater' are preferable to alternative storage methods for maintaining RNA quality.¹⁰¹ Perimorbid events, including agonal status, can also affect RNA quality and can be difficult to match between groups. The BrainNet Europe Consortium recently concluded that brain pH, which can be more easily matched across samples, is a good proxy of RNA quality.¹⁰² RNA integrity number is a measure of RNA degradation that is based on evaluation of an electropherogram trace.^{103,104} A threshold value of 7.8 indicates RNA quality that is sufficient for reproducible microarray work.¹⁰⁵

Microarray platform and data analysis

Different microarray platforms can generate different results.¹⁰⁶ However, a recent multicentre study evaluated the reproducibility of microarray results between centres, platforms and analysis techniques, and concluded that determination of differentially expressed transcripts was acceptably consistent.¹⁰⁷

The threshold used to define a significant difference in gene expression will clearly influence the results of a GEP study. An ideal threshold would yield only results that are reproducibly linked to pathophysiology. The optimal statistical test for calculation of significance is also debated. Parametric tests are limited, as many of the genes represented on a microarray are not normally distributed,¹⁰⁸ but nonparametric tests can result in loss of power.¹⁰⁹ The large number of statistical tests performed when determining differentially expressed genes can lead to false-positive results, a so-called multiple-testing problem. Conversely, commonly used multiple-testing corrections can lead to a high false-negative rate.¹¹⁰

Another problem in microarray analysis is that the most significant changes by *P*-value or fold change are not always the most important. For example, a small change in expression of a gene such as a transcription factor can have a large pathophysiological effect. Additionally, genes with a low level of expression, regardless of their functional importance, show increased variability and are, therefore, liable to become false negatives, particularly if a stringent multiple-testing correction is applied. To counteract this limitation, a clustering or enrichment

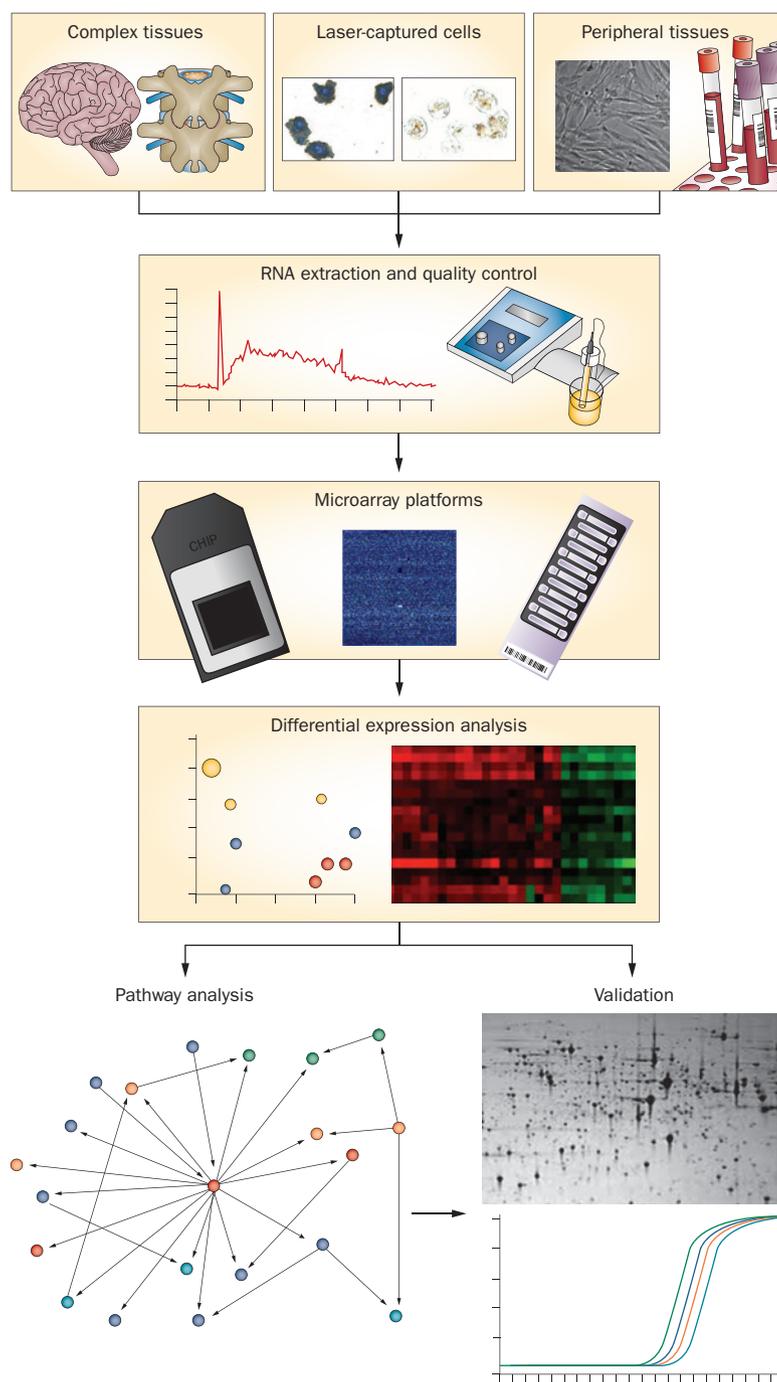


Figure 1 | Schematic illustration of the various methodological stages of a microarray study and analysis.

analysis,¹¹¹ which uses the interrelated nature of gene expression to reduce the false-positive rate without increasing the number of false negatives, can be conducted. This approach also allows better comparison of different studies.²¹ The problem of genes with a low level of expression can be addressed using a penalized *t*-statistic¹¹² or the Propagating Uncertainty in Microarray Analysis (PUMA) package, which uses a quantification of noise in probe set measurements in subsequent analyses, rather than simply assigning an expression value per probe set at normalization.¹¹³

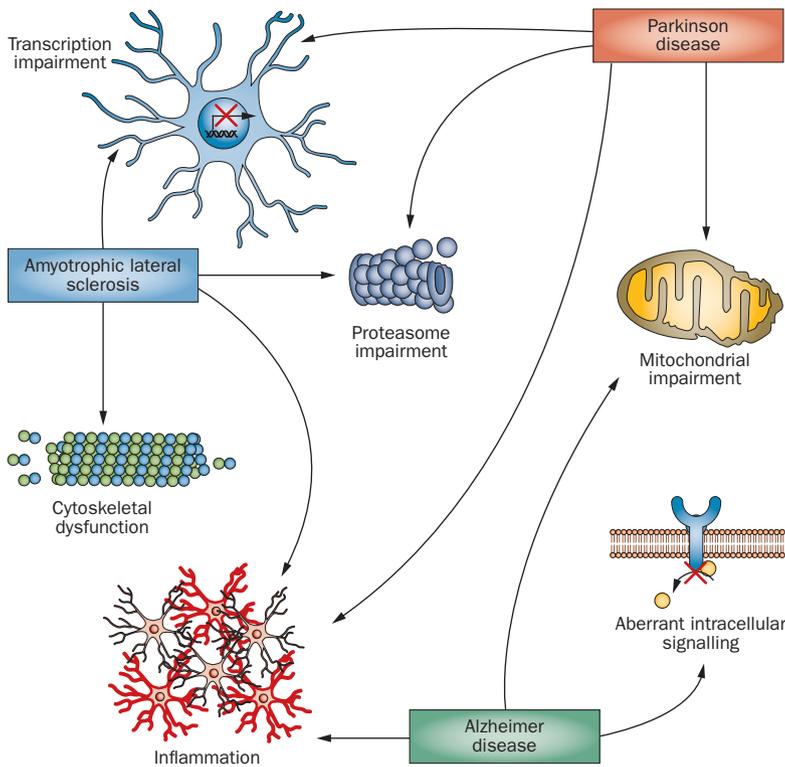


Figure 2 | Summary of the biological pathways that are consistently identified by gene expression profiling of human tissue samples from patients with amyotrophic lateral sclerosis, Alzheimer disease or Parkinson disease. Crossover between findings in the different diseases is also highlighted.

Validation

Validation of microarray data that show gene expression changes is a way to avoid false-positive conclusions. This can be achieved through assay of gene expression by another method, such as quantitative rtPCR or *in situ* hybridization. Alternatively, results can be validated by measuring downstream effects resulting from changes in gene expression—for example, by western blotting, immunohistochemistry or assays of protein function.

Exon-level microarrays

Many studies described in this Review have used microarray platforms with probes for a particular gene, irrespective of the specific transcript isoform. More than 90% of multi-exon genes are estimated to undergo alternative splicing¹¹⁴ and, therefore, important physiological and perhaps pathological variation is not captured by a gene-level analysis. Indeed, evidence discussed in this Review has implied that pathological gene changes related to RNA splicing do occur in neurodegeneration. A difficulty associated with exon-level analysis using a microarray platform with probes for every exon is that, necessarily, there are fewer probes per exon than probes per gene on a gene-level microarray,¹¹⁵ which means that the measurements of individual exon expression may be subject to excessive variability.

Commonly used analysis methods try to overcome this hurdle by measuring expression at the level of known transcripts, such that probes from a number of exons can be combined. Alternatively, Affymetrix AltSplice arrays

—which contain probes specific for all splicing events in the University of California, Santa Cruz and Ensembl databases¹¹⁵—can be used. Importantly neither of these approaches can detect novel splice variants, which might be particularly relevant to disruption of splicing function.

Next-generation sequencing of RNA

NGS refers to new technologies that allow sequencing of large amounts of DNA or RNA at high speed and relatively low cost compared with dideoxy sequencing. NGS therefore enables compilation of a substantial amount of data, comparable to GEP by microarray analysis. NGS of RNA has several advantages over microarray platforms, largely resulting from a lack of reliance on probes. Rather, every base of every component of the transcriptome is sequenced, improving the detection rate of known transcripts and splicing events,¹¹⁶ and enabling detection of novel transcripts in the absence of specific probes. The problem of nonspecific binding to probes, a significant source of technical variability, is bypassed.¹¹⁷ In addition, the single-nucleotide resolution of NGS allows detection of sequence variability within an RNA molecule.

A substantial challenge in GEP via NGS is analysis of the large volume of data produced. Such NGS applications include mapping of RNA sequences to the genome, which is particularly difficult with degraded RNA. Another barrier is the amount of starting material required. Whereas microarray techniques often require only 100 ng or less, clonal amplification prior to NGS requires 3–20 µg of RNA of suitable quality.¹¹⁸ These difficulties are particularly relevant in neurodegenerative disease, in which LCM of a selected population of CNS cells often produces small amounts of relatively poor-quality RNA. However, a recent study using an isothermal linear amplification technique reported successful RNA NGS using only 500 pg of starting material.¹¹⁹

A comparison of the two methods concluded that microarray analysis performed better than NGS in the quantification of known low-abundance transcripts, and the researchers proposed a hybrid approach using NGS to definitively determine which transcripts are present, and a custom microarray to probe identified transcripts.¹²⁰

Biomarker development

The capacity of GEP to simultaneously assay a large number of biological pathways is suited to devising a biomarker of complex neurodegenerative disease, which could be used to facilitate diagnosis and subclassification of heterogeneous disease states, predict prognosis, and adjust treatment dosage by titration. A biomarker panel of genes for early PD that can be tested in peripheral blood samples is showing promise.¹⁹ Similarly, in AD, several studies^{89,91,121} have attempted to separate disease and control samples according to expression of a panel of genes. One group has developed this approach into a test that could be used on a large scale in the clinic.⁹⁵ Similar panels are already in clinical use in other disease areas.¹²² The main challenges facing the use of GEP in biomarker development for neurodegenerative disease are the choice of tissue and analysis method.

Choice of tissue

Peripheral cells, which can be acquired noninvasively during life at any stage of disease, are a good source of material, particularly for biomarker studies. A key question in neurodegenerative disease is the extent to which peripheral cells are representative of a process primarily occurring in the CNS. In ALS, PD and AD, evidence of systemic involvement exists,^{2,19,88} and the majority of genes implicated in familial neurodegenerative diseases are ubiquitously expressed. Several studies reviewed in this article have shown that gene expression changes in peripheral tissues from patients with neurodegenerative disease are directly comparable to changes in the CNS.^{2,16,20,49,52}

Analysis method

The 2010 MicroArray Quality Control study compared multiple classification methodologies that are suitable for biomarker analysis.¹²³ This study concluded that the primary determinant of performance of a classifier algorithm was the end point under consideration, and that various techniques for data analysis, including normalization and panel selection, largely produced similar results. Difficulty in identification of a reproducible biomarker has been attributed to disease heterogeneity, even in oncology where, in contrast to neurodegenerative disease, the tissue of interest is usually accessible during life.¹²⁴ Alternatively, other studies suggest that a large number of genes can have a small but comparable association with a particular end point.^{125,126} This means that when sample numbers are small, ranking of genes by their association with the end point, as is common in many biomarker studies, relies on an order that could vary significantly each time the study is repeated. In the case of breast cancer prognosis, for example, several thousand samples are thought to be needed to produce a reliable biomarker panel.¹²⁶ Disease heterogeneity is undoubtedly a problem, but the numbers of samples in biomarker development might also need to be vastly increased.

Conclusions

In each of the diseases that have been considered, the GEP studies are numerous and varied. Direct comparison of results is often difficult, because studies use different tissues and/or different microarray platforms. It is interesting that, despite these barriers, common themes arise in each disease. In ALS, several studies implicate alterations in gene expression related to the cytoskeleton, inflammation, protein turnover and RNA splicing. In PD, the pathological transcriptome supports disruption of the ubiquitin–proteasome system, RNA splicing, mitochondrial dysfunction and neuroinflammation. In AD, neuroinflammation, mitochondrial dysfunction

and various intracellular signalling pathways, including calcium signalling, are repeatedly implicated. These changes are summarized in Figure 2.

Notably, similarities exist between the diseases: for example, mitochondrial dysfunction seems to play a part in both AD and PD, and disruption of RNA splicing and protein processing are implicated in ALS and PD. Neuroinflammation is observed in all three diseases, although this process might be expected to occur in the context of degenerating neurons and reactive gliosis, and could, therefore, represent a downstream effect that is not a viable therapeutic target. Studies in animal models of neurodegenerative disease have provided insight into pathogenesis in early-stage disease, and support similarity between the diseases.^{14,127} Similarities between neurodegenerative diseases are the subject of ongoing research¹²⁸ and, in certain cases, a common genetic basis has been discovered for multiple neurodegenerative diseases.^{129,130}

As well as identifying dysfunctional pathways, GEP studies have highlighted a number of potential therapeutic strategies, including reduction of PTEN activity in ALS,⁷ activation of PGC1 α in PD,²⁰ and prevention of decline in synaptic function in AD.^{55,64} In many cases, the conclusions are based on exploration of selective vulnerability between sexes,^{44,47} ages,^{60,73} brain areas²⁴ and even cellular subcompartments.⁵⁵

A number of studies have highlighted aberrant gene splicing in ALS, PD and AD. Many of the platforms used in the studies reviewed are not capable of measuring exon-level expression, and currently even exon-level microarrays are not suitable for identification of novel splicing events. Post-transcriptional regulation of gene expression, including alternative splicing, is an area that seems to be much more complex than first supposed.⁹⁸ With its inherent advantages, NGS of the transcriptome, building on the foundation of microarray studies, might be the technique that substantially advances our understanding of how alternative splicing contributes to the pathogenesis of neurodegenerative disease.

Review criteria

The scope of this Review was limited to mRNA profiling by microarray. Studies using customized arrays were not included; instead the focus was on the findings and methodologies of studies that profiled large numbers of genes to identify molecular mechanisms of disease. References for this Review were identified through searches of PubMed with the search terms “gene expression profiling” and “amyotrophic lateral sclerosis”, “Parkinson” or “Alzheimer” for papers published from January 2005 to April 2012. Only papers published in English were reviewed.

- Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).
- Saris, C. G. *et al.* Weighted gene co-expression network analysis of the peripheral blood from amyotrophic lateral sclerosis patients. *BMC Genomics* **10**, 405 (2009).
- Offen, D. *et al.* Spinal cord mRNA profile in patients with ALS: comparison with transgenic mice expressing the human SOD-1 mutant. *J. Mol. Neurosci.* **38**, 85–93 (2009).
- Wang, X. S., Simmons, Z., Liu, W., Boyer, P. J. & Connor, J. R. Differential expression of genes in amyotrophic lateral sclerosis revealed by profiling the post mortem cortex. *Amyotroph. Lateral Scler.* **7**, 201–210 (2006).
- Lederer, C. W., Torrisi, A., Pantelidou, M., Santama, N. & Cavallaro, S. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC Genomics* **8**, 26 (2007).
- Ferraiuolo, L., Kirby, J., Grierson, A. J., Sendtner, M. & Shaw, P. J. Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis. *Nat. Rev. Neurol.* **7**, 616–630 (2011).

7. Kirby, J. *et al.* Phosphatase and tensin homologue/protein kinase B pathway linked to motor neuron survival in human superoxide dismutase 1-related amyotrophic lateral sclerosis. *Brain* **134**, 506–517 (2011).
8. Jiang, Y. M. *et al.* Gene expression profile of spinal motor neurons in sporadic amyotrophic lateral sclerosis. *Ann. Neurol.* **57**, 236–251 (2005).
9. Rabin, S. J. *et al.* Sporadic ALS has compartment-specific aberrant exon splicing and altered cell-matrix adhesion biology. *Hum. Mol. Genet.* **19**, 313–328 (2010).
10. Cox, L. E. *et al.* Mutations in *CHMP2B* in lower motor neuron predominant amyotrophic lateral sclerosis (ALS). *PLoS ONE* **5**, e9872 (2010).
11. Hewitt, C. *et al.* Novel FUS/TLS mutations and pathology in familial and sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* **67**, 455–461 (2010).
12. Kirby, J. *et al.* Broad clinical phenotypes associated with TAR-DNA binding protein (*TARDBP*) mutations in amyotrophic lateral sclerosis. *Neurogenetics* **11**, 217–225 (2010).
13. Kirby, J. *et al.* Mutant SOD1 alters the motor neuronal transcriptome: implications for familial ALS. *Brain* **128**, 1686–1706 (2005).
14. Ferraiuolo, L. *et al.* Microarray analysis of the cellular pathways involved in the adaptation to and progression of motor neuron injury in the SOD1 G93A mouse model of familial ALS. *J. Neurosci.* **27**, 9201–9219 (2007).
15. Zhang, R. *et al.* Gene expression profiling in peripheral blood mononuclear cells from patients with sporadic amyotrophic lateral sclerosis (sALS). *J. Neuroimmunol.* **230**, 114–123 (2011).
16. Mougeot, J.-L., Li, Z., Price, A., Wright, F. & Brooks, B. Microarray analysis of peripheral blood lymphocytes from ALS patients and the SAFE detection of the KEGG ALS pathway. *BMC Med. Genomics* **4**, 74 (2011).
17. Pradat, P. F. *et al.* Muscle gene expression is a marker of amyotrophic lateral sclerosis severity. *Neurodegen. Dis.* **9**, 38–52 (2012).
18. Shtilbans, A. *et al.* Differential gene expression in patients with amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.* **12**, 250–256 (2011).
19. Scherzer, C. R. *et al.* Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc. Natl Acad. Sci. USA* **104**, 955–960 (2007).
20. Zheng, B. *et al.* PGC-1 α , a potential therapeutic target for early intervention in Parkinson's disease. *Sci. Transl. Med.* **2**, 52ra73 (2010).
21. Elstner, M. *et al.* Expression analysis of dopaminergic neurons in Parkinson's disease and aging links transcriptional dysregulation of energy metabolism to cell death. *Acta Neuropathol.* **122**, 75–86 (2011).
22. Shin, J.-H. *et al.* PARIS (ZNF746) repression of PGC-1 α contributes to neurodegeneration in Parkinson's disease. *Cell* **144**, 689–702 (2011).
23. Papapetropoulos, S. *et al.* Multiregional gene expression profiling identifies *MRPS6* as a possible candidate gene for Parkinson's disease. *Gene Expr.* **13**, 205–215 (2006).
24. Moran, L. *et al.* Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. *Neurogenetics* **7**, 1–11 (2006).
25. Hauser, M. A. *et al.* Expression profiling of substantia nigra in Parkinson disease, progressive supranuclear palsy, and frontotemporal dementia with parkinsonism. *Arch. Neurol.* **62**, 917–921 (2005).
26. Zhang, Y., James, M., Middleton, F. A. & Davis, R. L. 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.* **137B**, 5–16 (2005).
27. Vogt, I. R. *et al.* Transcriptional changes in multiple system atrophy and Parkinson's disease putamen. *Exp. Neurol.* **199**, 465–478 (2006).
28. Naydenov, A., Vassoler, F., Luksik, A., Kaczmarzka, J. & Konradi, C. Mitochondrial abnormalities in the putamen in Parkinson's disease dyskinesia. *Acta Neuropathol.* **120**, 623–631 (2010).
29. Botta-Orfila, T. *et al.* Microarray expression analysis in idiopathic and *LRKK2*-associated Parkinson's disease. *Neurobiol. Dis.* **45**, 462–468 (2012).
30. Durrenberger, P. *et al.* Inflammatory pathways in Parkinson's disease; a BNE microarray study. *Parkinson Dis.* **2012**, 214714 (2012).
31. Bossers, K. *et al.* Analysis of gene expression in Parkinson's disease: possible involvement of neurotrophic support and axon guidance in dopaminergic cell death. *Brain Pathol.* **19**, 91–107 (2009).
32. Miller, R. M. *et al.* Robust dysregulation of gene expression in substantia nigra and striatum in Parkinson's disease. *Neurobiol. Dis.* **21**, 305–313 (2006).
33. Duke, D., Moran, L., Pearce, R. & Graeber, M. The medial and lateral substantia nigra in Parkinson's disease: mRNA profiles associated with higher brain tissue vulnerability. *Neurogenetics* **8**, 83–94 (2007).
34. Fearnley, J. M. & Lees, A. J. Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain* **114**, 2283–2301 (1991).
35. Zhou, C., Huang, Y. & Przedborski, S. Oxidative stress in Parkinson's disease. *Ann. NY Acad. Sci.* **1147**, 93–104 (2008).
36. Rideout, H. J., Larsen, K. E., Sulzer, D. & Stefanis, L. Proteasomal inhibition leads to formation of ubiquitin/ α -synuclein-immunoreactive inclusions in PC12 cells. *J. Neurochem.* **78**, 899–908 (2001).
37. Vila, M. & Przedborski, S. Targeting programmed cell death in neurodegenerative diseases. *Nat. Rev. Neurosci.* **4**, 365–375 (2003).
38. Ved, R. *et al.* Similar patterns of mitochondrial vulnerability and rescue induced by genetic modification of α -synuclein, parkin, and DJ-1 in *Caenorhabditis elegans*. *J. Biol. Chem.* **280**, 42655–42668 (2005).
39. Dawson, T. M. & Dawson, V. L. Molecular pathways of neurodegeneration in Parkinson's disease. *Science* **302**, 819–822 (2003).
40. Quigley, P. M., Korotkov, K., Baneyx, F. & Hol, W. G. The 1.6-Å crystal structure of the class of chaperones represented by *Escherichia coli* Hsp31 reveals a putative catalytic triad. *Proc. Natl Acad. Sci. USA* **100**, 3137–3142 (2003).
41. Moran, L. *et al.* Analysis of alpha-synuclein, dopamine and parkin pathways in neuropathologically confirmed parkinsonian nigra. *Acta Neuropathol.* **113**, 253–263 (2007).
42. Duke, D. *et al.* Transcriptome analysis reveals link between proteasomal and mitochondrial pathways in Parkinson's disease. *Neurogenetics* **7**, 139–148 (2006).
43. Mortiboys, H., Johansen, K., Aasly, J. & Bandmann, O. Mitochondrial impairment in patients with Parkinson disease with the G2019S mutation in *LRKK2*. *Neurology* **75**, 2017–2020 (2010).
44. Cantuti-Castelvetri, I. *et al.* Effects of gender on nigral gene expression and Parkinson disease. *Neurobiol. Dis.* **26**, 606–614 (2007).
45. Simunovic, F. *et al.* Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson's disease pathology. *Brain* **132**, 1795–1809 (2009).
46. Stamper, C. *et al.* Neuronal gene expression correlates of Parkinson's disease with dementia. *Mov. Disord.* **23**, 1588–1595 (2008).
47. Simunovic, F., Yi, M., Wang, Y., Stephens, R. & Sonntag, K. C. Evidence for gender-specific transcriptional profiles of nigral dopamine neurons in Parkinson disease. *PLoS ONE* **5**, e8856 (2010).
48. Surmeier, D. J. Calcium, ageing, and neuronal vulnerability in Parkinson's disease. *Lancet Neurol.* **6**, 933–938 (2007).
49. Shehadeh, L. A. *et al.* SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. *PLoS ONE* **5**, e9104 (2010).
50. Mar, J. C. *et al.* Variance of gene expression identifies altered network constraints in neurological disease. *PLoS Genet.* **7**, e1002207 (2011).
51. Matigian, N. *et al.* Disease-specific, neurosphere-derived cells as models for brain disorders. *Dis. Model. Mech.* **3**, 785–798 (2010).
52. Mutez, E. *et al.* Transcriptional profile of Parkinson blood mononuclear cells with *LRKK2* mutation. *Neurobiol. Aging* **32**, 1839–1848 (2011).
53. Parachikova, A. *et al.* Inflammatory changes parallel the early stages of Alzheimer disease. *Neurobiol. Aging* **28**, 1821–1833 (2007).
54. Tan, M. G. *et al.* Genome wide profiling of altered gene expression in the neocortex of Alzheimer's disease. *J. Neurosci. Res.* **88**, 1157–1169 (2010).
55. Williams, C. *et al.* Transcriptome analysis of synaptoneurosomes identifies neuroplasticity genes overexpressed in incipient Alzheimer's disease. *PLoS ONE* **4**, e4936 (2009).
56. Xu, P.-T. *et al.* Differences in apolipoprotein E3/3 and E4/4 allele-specific gene expression in hippocampus in Alzheimer disease. *Neurobiol. Dis.* **21**, 256–275 (2006).
57. Emilsson, L., Saetre, P. & Jazin, E. Alzheimer's disease: mRNA expression profiles of multiple patients show alterations of genes involved with calcium signaling. *Neurobiol. Dis.* **21**, 618–625 (2006).
58. Haroutunian, V., Katsel, P. & Schmeidler, J. Transcriptional vulnerability of brain regions in Alzheimer's disease and dementia. *Neurobiol. Aging* **30**, 561–573 (2009).
59. Katsel, P., Li, C. & Haroutunian, V. Gene expression alterations in the sphingolipid metabolism pathways during progression of dementia and Alzheimer's disease: a shift toward ceramide accumulation at the earliest recognizable stages of Alzheimer's disease? *Neurochem. Res.* **32**, 845–856 (2007).
60. Katsel, P., Tan, W. & Haroutunian, V. Gain in brain immunity in the oldest-old differentiates cognitively normal from demented individuals. *PLoS ONE* **4**, e7642 (2009).
61. Umemura, K. *et al.* Autotaxin expression is enhanced in frontal cortex of Alzheimer-type dementia patients. *Neurosci. Lett.* **400**, 97–100 (2006).
62. Weeraratna, A. T. *et al.* Alterations in immunological and neurological gene expression patterns in Alzheimer's disease tissues. *Exp. Cell Res.* **313**, 450–461 (2007).
63. Bronner, I. F. *et al.* Comprehensive mRNA expression profiling distinguishes tauopathies

- and identifies shared molecular pathways. *PLoS ONE* **4**, e6826 (2009).
64. Bossers, K. *et al.* Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease. *Brain* **133**, 3699–3723 (2010).
 65. Horesh, Y., Katsel, P., Haroutunian, V. & Domany, E. Gene expression signature is shared by patients with Alzheimer's disease and schizophrenia at the superior temporal gyrus. *Eur. J. Neurol.* **18**, 410–424 (2011).
 66. Youn, H. *et al.* Kalirin is under-expressed in Alzheimer's disease hippocampus. *J. Alzheimers Dis.* **11**, 385–397 (2007).
 67. Wang, S., Qaisar, U., Yin, X. & Grammas, P. Gene expression profiling in Alzheimer's disease brain microvessels. *J. Alzheimers Dis.* <http://dx.doi.org/10.3233/JAD-2012-120454>.
 68. Tollervy, J. R. *et al.* Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. *Genome Res.* **21**, 1572–1582 (2011).
 69. Kuchibhotla, K. V. *et al.* A β plaques lead to aberrant regulation of calcium homeostasis *in vivo* resulting in structural and functional disruption of neuronal networks. *Neuron* **59**, 214–225 (2008).
 70. Guo, Q. *et al.* Alzheimer's PS-1 mutation perturbs calcium homeostasis and sensitizes PC12 cells to death induced by amyloid peptide. *Neuroreport* **8**, 379–383 (1996).
 71. Camandola, S. & Mattson, M. P. Aberrant subcellular neuronal calcium regulation in aging and Alzheimer's disease. *Biochim. Biophys. Acta* **1813**, 965–973 (2011).
 72. Berchtold, N. C. *et al.* Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proc. Natl Acad. Sci. USA* **105**, 15605–15610 (2008).
 73. Saetre, P., Jazin, E. & Emilsson, L. Age-related changes in gene expression are accelerated in Alzheimer's disease. *Synapse* **65**, 971–974 (2011).
 74. Masliah, E. *et al.* Altered expression of synaptic proteins occurs early during progression of Alzheimer's disease. *Neurology* **56**, 127–129 (2001).
 75. Giuditta, A. *et al.* Local gene expression in axons and nerve endings: the glia–neuron unit. *Physiol. Rev.* **88**, 515–555 (2008).
 76. Liang, W. S. *et al.* Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: a reference data set. *Physiol. Genomics* **33**, 240–256 (2008).
 77. Liang, W. S. *et al.* Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *Proc. Natl Acad. Sci. USA* **105**, 4441–4446 (2008).
 78. Dunckley, T. *et al.* Gene expression correlates of neurofibrillary tangles in Alzheimer's disease. *Neurobiol. Aging* **27**, 1359–1371 (2006).
 79. Alexander, G. E., Chen, K., Pietrini, P., Rapoport, S. I. & Reiman, E. M. Longitudinal PET evaluation of cerebral metabolic decline in dementia: a potential outcome measure in Alzheimer's disease treatment studies. *Am. J. Psychiatry* **159**, 738–745 (2002).
 80. Minoshima, S. *et al.* Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann. Neurol.* **42**, 85–94 (1997).
 81. Ray, M. & Zhang, W. Analysis of Alzheimer's disease severity across brain regions by topological analysis of gene co-expression networks. *BMC Syst. Biol.* **4**, 136 (2010).
 82. Liu, Z.-P., Wang, Y., Zhang, X.-S. & Chen, L. Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC Syst. Biol.* **4**, S11 (2010).
 83. Ray, M., Ruan, J. & Zhang, W. Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. *Genome Biol.* **9**, R148 (2008).
 84. Gorelick, P. B. *et al.* Vascular contributions to cognitive impairment and dementia. *Stroke* **42**, 2672–2713 (2011).
 85. Helbecque, N. & Amouyel, P. Commonalities between genetics of cardiovascular disease and neurodegenerative disorders. *Curr. Opin. Lipidol.* **15**, 121–127 (2004).
 86. Simpson, J. E. *et al.* Microarray analysis of the astrocyte transcriptome in the aging brain: relationship to Alzheimer's pathology and APOE genotype. *Neurobiol. Aging* **32**, 1795–1807 (2011).
 87. Strittmatter, W. J. *et al.* Apolipoprotein E: high-avidity binding to β -amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 1977–1981 (1993).
 88. Maes, O. C. *et al.* Transcriptional profiling of Alzheimer blood mononuclear cells by microarray. *Neurobiol. Aging* **28**, 1795–1809 (2007).
 89. Nagasaka, Y. *et al.* A unique gene expression signature discriminates familial Alzheimer's disease mutation carriers from their wild-type siblings. *Proc. Natl Acad. Sci. USA* **102**, 14854–14859 (2005).
 90. Calciano, M. A., Zhou, W., Snyder, P. J. & Einstein, R. Drug treatment of Alzheimer's disease patients leads to expression changes in peripheral blood cells. *Alzheimers Dement.* **6**, 386–393 (2010).
 91. Booij, B. B. *et al.* A gene expression pattern in blood for the early detection of Alzheimer's disease. *J. Alzheimers Dis.* **23**, 109–119 (2011).
 92. Kálmán, J. *et al.* Gene expression profile analysis of lymphocytes from Alzheimer's patients. *Psychiatr. Genet.* **15**, 1–6 (2005).
 93. Fehlbaum-Beurdeley, P. *et al.* Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimers Dement.* **6**, 25–38 (2010).
 94. Chen, K.-D. *et al.* Gene expression profiling of peripheral blood leukocytes identifies and validates ABCB1 as a novel biomarker for Alzheimer's disease. *Neurobiol. Dis.* **43**, 698–705 (2011).
 95. Rye, P. D. *et al.* A novel blood test for the early detection of Alzheimer's disease. *J. Alzheimers Dis.* **23**, 121–129 (2011).
 96. Scherzer, C. *et al.* Loss of apolipoprotein E receptor LR11 in Alzheimer disease. *Arch. Neurol.* **61**, 1200–1205 (2004).
 97. Lee, J., Barral, S. & Reitz, C. The neuronal sortilin-related receptor gene *SORL1* and late-onset Alzheimer's disease. *Curr. Neurol. Neurosci. Rep.* **8**, 384–391 (2008).
 98. Sutherland, G. T., Janitz, M. & Kril, J. J. Understanding the pathogenesis of Alzheimer's disease: will RNA-Seq realize the promise of transcriptomics? *J. Neurochem.* **116**, 937–946 (2011).
 99. Emmert-Buck, M. R. *et al.* Laser capture microdissection. *Science* **274**, 998–1001 (1996).
 100. Sutherland, G. T. *et al.* A cross-study transcriptional analysis of Parkinson's disease. *PLoS ONE* **4**, e4955 (2009).
 101. Casale, V., Oneda, R., Lavezzi, A. M. & Matturri, L. Optimisation of postmortem tissue preservation and alternative protocol for serotonin transporter gene polymorphisms amplification in SIDS and SIUD cases. *Exp. Mol. Pathol.* **88**, 202–205 (2010).
 102. Durrenberger, P. F. *et al.* Effects of antemortem and postmortem variables on human brain mRNA quality: a BrainNet Europe study. *J. Neuropathol. Exp. Neurol.* **69**, 70–81 (2010).
 103. Strand, C., Enell, J., Hedenfalk, I. & Ferno, M. RNA quality in frozen breast cancer samples and the influence on gene expression analysis—a comparison of three evaluation methods using microcapillary electrophoresis traces. *BMC Mol. Biol.* **8**, 38 (2007).
 104. Schroeder, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* **7**, 3 (2006).
 105. Copois, V. *et al.* Impact of RNA degradation on gene expression profiles: assessment of different methods to reliably determine RNA quality. *J. Biotechnol.* **127**, 549–559 (2007).
 106. Tan, P. K., *et al.* Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res.* **31**, 5676–5684 (2003).
 107. MAQC Consortium *et al.* The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotech.* **24**, 1151–1161 (2006).
 108. Posekany, A., Felsenstein, K. & Sykacek, P. Biological assessment of robust noise models in microarray data analysis. *Bioinformatics* **27**, 807–814 (2011).
 109. Stekel, D. *Microarray Bioinformatics* 110–138 (Cambridge University Press, Cambridge, UK, 2003).
 110. Storey, J. D., Dai, J. Y. & Leek, J. T. The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. *Biostatistics* **8**, 414–432 (2007).
 111. Hosack, D., Dennis, G., Sherman, B., Lane, H. & Lempicki, R. Identifying biological themes within lists of genes with EASE. *Genome Biol.* **4**, R70 (2003).
 112. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article 3 (2004).
 113. Liu, X., Milo, M., Lawrence, N. D. & Rattray, M. Probe-level measurement error improves accuracy in detecting differential gene expression. *Bioinformatics* **22**, 2107–2113 (2006).
 114. Blencowe, B. J., Ahmad, S. & Lee, L. J. Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. *Genes Dev.* **23**, 1379–1386 (2009).
 115. Yamamoto, M. L. *et al.* Alternative pre-mRNA splicing switches modulate gene expression in late erythropoiesis. *Blood* **113**, 3363–3370 (2009).
 116. Sultan, M. *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956–960 (2008).
 117. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
 118. Metzker, M. L. Sequencing technologies—the next generation. *Nat. Rev. Genet.* **11**, 31–46 (2010).
 119. Tariq, M. A., Kim, H. J., Jejelowo, O. & Pourmand, N. Whole-transcriptome RNAseq analysis from minute amount of total RNA. *Nucleic Acids Res.* <http://dx.doi.org/10.1093/nar/gkr547>.
 120. Łabaj, P. P. *et al.* Characterization and improvement of RNA-Seq precision in quantitative transcript expression profiling. *Bioinformatics* **27**, i383–i391 (2011).

121. Pascale, F.-B. *et al.* Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimers Dement.* **6**, 25–38 (2010).
122. Kaklamani, V. A genetic signature can predict prognosis and response to therapy in breast cancer: oncotype DX. *Expert Rev. Mol. Diagn.* **6**, 803–809 (2006).
123. Shi, L. *et al.* The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat. Biotechnol.* **28**, 827–838 (2010).
124. Rottenberg, S. *et al.* Impact of intertumoral heterogeneity on predicting chemotherapy response of BRCA1-deficient mammary tumors. *Cancer Res.* **72**, 2350–2361 (2012).
125. Ein-Dor, L., Kela, I., Getz, G., Givol, D. & Domany, E. Outcome signature genes in breast cancer: is there a unique set? *Bioinformatics* **21**, 171–178 (2005).
126. Ein-Dor, L., Zuk, O. & Domany, E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc. Natl Acad. Sci. USA* **103**, 5923–5928 (2006).
127. Cabeza-Arvelaiz, Y. *et al.* Analysis of striatal transcriptome in mice overexpressing human wild-type alpha-synuclein supports synaptic dysfunction and suggests mechanisms of neuroprotection for striatal neurons. *Mol. Neurodegen.* **6**, 83 (2011).
128. Bredeisen, D. E., Rao, R. V. & Mehlen, P. Cell death in the nervous system. *Nature* **443**, 796–802 (2006).
129. Grünblatt, E. Commonalities in the genetics of Alzheimer's disease and Parkinson's disease. *Expert Rev. Neurother.* **8**, 1865–1877 (2008).
130. van Es, M. A. *et al.* Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis. *Ann. Neurol.* **70**, 964–973 (2011).

Acknowledgements

The work of this group is supported by the Motor Neurone Disease Association, the Wellcome Trust,

the Medical Research Council, and by a European Community 7th Framework Programme (FP7/2007-2013) under grant agreement number 259867 Euromotor to P. J. Shaw and J. Kirby; by a David Peake fellowship awarded to J. Cooper-Knock; and by BBSRC and EPSRC funding awarded to M. Rattray.

Author contributions

J. Cooper-Knock researched the data for the article. J. Cooper-Knock, J. Kirby, P. R. Heath, M. Rattray and P. J. Shaw made substantial contributions to discussion of the article content. J. Cooper-Knock and L. Ferraiuolo wrote the article. J. Kirby, P. R. Heath, M. Rattray and P. J. Shaw contributed to review and/or editing of the manuscript before submission.

Supplementary information

Supplementary information is linked to the online version of the paper at www.nature.com/nrneuro