

1 Transcriptomic signatures of brain regional 2 vulnerability to Parkinson's disease

3 Arlin Keo^{1,2}, Ahmed Mahfouz^{1,2}, Angela M.T. Ingrassia³, Jean-Pascal Meneboe^{4,5}, Celine Villenet⁴,
4 Eugénie Mutez^{6,7,8}, Thomas Comptdaer^{6,7}, Boudewijn P.F. Lelieveldt^{1,2,9}, Martin Figeac^{4,5}, Marie-Christine
5 Chartier-Harlin^{6,7*}, Wilma D.J. van de Berg^{3*}, Jacobus J. van Hilt^{10*}, and Marcel J.T. Reinders^{1,2*}.

6 1. Leiden Computational Biology Center, Leiden University Medical Center, Leiden, The
7 Netherlands.

8 2. Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands.

9 3. Department of Anatomy and Neurosciences, Amsterdam Neuroscience, Amsterdam UMC,
10 location VUmc, Amsterdam, The Netherlands.

11 4. Univ. Lille, Plate-forme de génomique fonctionnelle et Structurale, F-59000 Lille, France.

12 5. Univ. lille. Bilille, F-59000 Lille, France

13 6. Univ. Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Jean-Pierre AUBERT
14 Neurosciences et Cancer, F-59000 Lille, France

15 7. Inserm, UMR-S 1172, Early Stages of Parkinson's Disease, F-59000 Lille, France.

16 8. Univ. Lille, Service de Neurologie et Pathologie du mouvement, centre expert Parkinson, F-59000
17 Lille, France

18 9. Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands.

19 10. Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands

20 *Shared corresponding authors. E-mail: marie-christine.chartier-harlin@inserm.fr;
21 wdj.vandeberg@amsterdamumc.nl; j.j.van_hilt@lumc.nl; m.j.t.reinders@tudelft.nl.

22 Keywords: Parkinson's disease, Braak Lewy body stages, transcriptomics, selective vulnerability, gene
23 co-expression.

24

25 **Abstract**

26 The molecular mechanisms underlying the caudal-to-rostral progression of Lewy body pathology in
27 Parkinson's disease (PD) remain poorly understood. Here, we aimed to unravel transcriptomic signatures
28 across brain regions involved in Braak Lewy body stages in non-neurological controls and PD donors.
29 Using human postmortem brain datasets of non-neurological adults from the Allen Human Brain Atlas, we
30 identified expression patterns related to PD progression, including genes found in PD genome-wide
31 associations studies: *SNCA*, *ZNF184*, *BAP1*, *SH3GL2*, *ELOVL7*, and *SCARB2*. We confirmed these
32 patterns in two datasets of non-neurological subjects (Genotype-Tissue Expression project and UK Brain
33 Expression Consortium) and found altered patterns in two datasets of PD patients. Additionally, co-
34 expression analysis across vulnerable regions identified two modules associated with dopamine
35 synthesis, the motor and immune system, blood-oxygen transport, and contained microglial and
36 endothelial cell markers, respectively. Alterations in genes underlying these region-specific functions may
37 contribute to the selective regional vulnerability in PD brains.

38

39 **Background**

40 Parkinson's disease (PD) is characterized by a temporal caudal-rostral progression of Lewy body (LB)
41 pathology across a selected set of nuclei in the brain¹. The distribution pattern of LB pathology is divided
42 into six Braak stages based on accumulation of the protein α -synuclein – the main component of LBs and
43 Lewy neurites – in the brainstem, limbic and neocortical regions¹. Different hypotheses have been
44 brought forward to explain the evolving LB pathology across the brain, including: retrograde transport of
45 pathological α -synuclein via neuroanatomical networks, α -synuclein's prion-like behavior, and cell- or
46 region-autonomous factors^{2,3}. Yet, the mechanisms underlying the selective vulnerability of brain regions
47 to LB pathology remains poorly understood, limiting the ability to diagnose and treat PD.

48 Multiplications of the *SNCA* gene encoding α -synuclein are relatively common in autosomal dominant PD
49 and *SNCA* dosage has been linked to the severity of PD^{4,5}. For other PD-associated variants, e.g. *GBA*
50 and *LRRK2*, their role in progressive α -synuclein accumulation is less clear, although they have been
51 associated with mitochondrial (dys)function and/or protein degradation pathways⁶⁻⁸. On the other hand,
52 transcriptomic changes between PD and non-neurological controls of selected brain regions, e.g. the
53 substantia nigra, have identified several molecular mechanisms underlying PD pathology, including
54 synaptic vesicle endocytosis⁹⁻¹¹. However, post-mortem human brain tissue of well-characterized PD
55 patients and controls is scarce, usually focuses on a select number of brain regions, and have a limited
56 coverage of patients with different Braak LB stages, resulting in low concordance of findings across
57 different studies¹².

58 Spatial gene expression patterns in the human brain have been studied to unravel the pathogenic
59 mechanisms underlying amyloid- β and tau pathology progression in Alzheimer's disease, revealing
60 proteins that co-aggregate with amyloid- β and tau, and protein homeostasis components^{13,14}.
61 Interestingly, by integrating Allen Human Brain Atlas (AHBA) gene expression data¹⁵ with magnetic
62 resonance imaging of PD patients, the regional expression pattern of *MAPT* and *SNCA* was associated
63 with loss of functional connectivity in PD¹⁶, and regional expression of synaptic transfer genes was related
64 to regional gray matter atrophy in PD¹⁷. This combined gene-MRI analysis illustrates the importance of
65 local gene expression changes on functional brain networks. More detailed knowledge about the spatial

66 organization of transcriptomic changes in physiological and pathological conditions may aid in
67 understanding these changes on a functional level during disease progression in PD.

68 In the present study, we aim to unravel the molecular factors underlying selective vulnerability to LB
69 pathology during PD progression. We used transcriptomic postmortem human brain data of non-
70 neurological adult donors from the AHBA and analyzed brain regions involved in Braak LB stages¹⁸. We
71 identified genes whose expression patterns increase or decrease across these regions, referred to as
72 Braak stage-related genes (BRGs) and hypothesized that they might contribute to higher vulnerability to
73 LBs in PD brains based on the sequence of events as postulated by Braak et al.¹ (Figure 1). We validated
74 the relevance of the identified BRGs to PD progression with two datasets of non-neurological controls
75 (Genotype-Tissue Expression project (GTEx)¹⁹ and UK Brain Expression Consortium (UKBEC)²⁰) and two
76 datasets of PD donors. We found altered expression patterns in patients with PD and incidental Lewy
77 body disease (iLBD), who are assumed to represent the pre-clinical stage of PD^{11,21}. In non-neurological
78 brains from the AHBA, we further identified modules of co-expressed genes across Braak LB stage
79 involved regions, and characterized them functionally by assessing their enrichment for cell-type markers,
80 gene ontology (GO)-terms, and disease-associated genes. Using our region-based approach, we
81 identified genes known to harbor PD-associated variants (e.g. SNCA, ZNF184, BAP1, SH3GL2, ELOVL7,
82 and SCARB2) and pathways linked to PD (e.g. dopamine biosynthetic process and immune response)
83 with expression signatures associated with progressive LB pathology. In non-neurological controls,
84 several dopaminergic genes were highly expressed in brain regions related to the preclinical stages of PD
85 progression, highlighting their importance in maintaining motor functions. These observed transcriptomic
86 signatures provide insights into the molecular mechanisms underlying brain regional vulnerability to PD,
87 enabling the development of new and improved methods for diagnosis and treatment of PD.

88 **Results**

89 **Parkinson's disease Braak stage-related genes**

90 The PD Braak staging scheme defines a temporal order of brain regions affected during the progression
91 of the disease¹⁸. We analyzed these brain regions using a microarray data set of 3,702 anatomical brain

92 regions from six individuals without any known neuropsychiatric or neurological background from the
93 Allen Human Brain Atlas (AHBA)¹⁵. We first assigned brain samples to Braak stage-related regions R1-
94 R6¹⁸: myelencephalon (medulla, R1), pontine tegmentum including locus coeruleus (R2), substantia
95 nigra, basal nucleus of Meynert, CA2 of hippocampus (R3), amygdala, occipito-temporal gyrus (R4),
96 cingulate gyrus, temporal lobe (R5), frontal lobe including the olfactory area, and parietal lobe (R6)
97 (Figure 1, Supplementary Table 1, and Supplementary Figure 1). To identify genes with expression
98 patterns that are associated with selective vulnerability to PD, Braak stage-related genes (BRGs), we
99 correlated gene expression with the label of these vulnerable regions as defined by Braak stage. To focus
100 on genes with large expression differences across regions, we assessed differential expression between
101 all pairs of Braak stage-related regions R1-R6, and found most significant changes between regions
102 related to the most distant stages: R1 versus R5 and R1 versus R6 ($|fold\text{-}change (FC)| > 1$; Benjamini-
103 Hochberg (BH)-corrected $P < 0.05$; Supplementary Figure 2).

104 BRGs were selected based on (i) the highest absolute Braak label correlation (r), (ii) highest absolute FC
105 between R1 and R6 (FC_{R1-R6}), and (iii) smallest BH-corrected P -values of the FC (P_{FC} ; two-sided t-test;
106 Figure 2a and b). The top 10% (2,001) ranked genes for each criterion resulted in genes with $|r| > 0.66$,
107 $|FC_{R1-R6}| > 1.33$, and $P_{FC} < 0.00304$. The overlap of the three sets of top 10% ranked genes resulted in 960
108 BRGs, with 348 negatively and 612 positively correlated genes showing a decreasing ($r < 0$) or increasing
109 ($r > 0$) expression pattern across regions R1-R6, respectively (Figure 2c and Supplementary Table 2).
110 Negatively correlated BRGs were significantly enriched for gene ontology (GO) terms like anatomical
111 structure morphogenesis and blood vessel morphogenesis (Supplementary Table 3), while positively
112 correlated BRGs were significantly enriched for functions like anterograde trans-synaptic signaling and
113 nervous system development (Supplementary Table 4).

114 To validate the decreasing or increasing expression patterns of 960 BRGs observed in six non-
115 neurological brains from the AHBA, we used two independent datasets from non-neurological controls
116 and assessed differential expression between the two most distant regions in each dataset. First, using
117 microarray data from 134 individuals in the UKBEC²⁰, we selected brain samples corresponding to the
118 myelencephalon (R1), substantia nigra (R3), temporal cortex (R5), and frontal cortex (R6). For the 885

119 BRGs present in UKBEC, 139 out of 317 (43.8%) negatively correlated BRGs and 400 out of 571 (70.1%)
120 positively correlated BRGs were differentially expressed between R1 and R6 ($|FC_{R1-R6}| > 1$, BH-corrected
121 $P < 0.05$, and Figure 2d). Second, we used RNA-sequencing (RNA-seq) data from 88-129 individuals in
122 the GTEx consortium¹⁹ and selected samples of the substantia nigra (R3), amygdala (R4), anterior
123 cingulate cortex (R5), and frontal cortex (R6). For the 883 BRGs present in the GTEx consortium, 204 out
124 of 318 (64.2%) negatively correlated BRGs and 475 out of 565 (84.1%) positively correlated BRGs were
125 differentially expressed between R3 and R6 in this dataset ($|FC_{R3-R6}| > 1$, BH-corrected $P < 0.05$, and
126 Figure 2e). Together, this indicates that the expression patterns of BRGs in the brain are consistent
127 across non-neurological individuals.

128 We next hypothesized that if the identified BRGs are associated with vulnerability to PD, they are also
129 indicative of vulnerability differences between PD patients and age-matched controls. To test this
130 hypothesis, we used two datasets with transcriptomic measurements from brain regions covering most
131 Braak stage-related regions sampled from PD, iLBD and non-demented age-matched controls (PD
132 microarray¹¹ (Supplementary Table 5 and 6) and PD RNA-seq datasets (Supplementary Table 7 and 8);
133 see Methods). First, we found larger differences between brain regions within the same group of
134 individuals (PD, iLBD, and control) than between conditions within the same region (number of
135 differentially expressed genes in Supplementary Figure 3). This observation further highlights the
136 importance of assessing expression patterns across regions rather than disease conditions²². Next, we
137 validated the expression patterns of BRGs, which we identified in non-neurological adults, in non-
138 demented age-matched controls in both the PD microarray and PD RNA-seq datasets and observed
139 similar patterns (Figure 2f and g). Interestingly, the increasing and decreasing expression patterns of
140 BRGs were diminished in iLBD patients and totally disrupted in PD patients across regions involved in
141 preclinical stages R1-R3 (Figure 2f). Across regions R3 and R4/R5 however, these expression patterns
142 were preserved in PD patients (Figure 2g). In addition to the changes across brain regions, we found that
143 BRGs also captured changes across conditions PD, iLBD, and control within the substantia nigra (R3) for
144 both PD datasets (Supplementary Figure 4). Negatively correlated BRGs that had higher expression in
145 the most vulnerable brain regions R1-R3 also had higher expression in PD patients compared to controls.
146 Vice versa, positively correlated BRGs that have higher expression in the least vulnerable brain regions

147 R4-R6 also have higher expression in controls compared to PD patients. These findings thus support the
148 relation of BRGs with PD vulnerability encountered in brain regions of non-neurological individuals and
149 show how their expression may influence the vulnerability at a region-specific level as well as between
150 patients and controls.

151 **Braak stage-related co-expression modules**

152 In addition to the expression of individual genes, we analyzed non-neurological brains from the AHBA to
153 examine the expression of gene sets that may jointly affect the vulnerability of brain regions to PD. To
154 study genetic coherence in vulnerable brain regions, we clustered all 20,017 genes into modules based
155 on their pairwise co-expression across regions R1-R6. The module eigengene, which summarizes the
156 overall expression of genes within a module, was correlated with the labels of regions as defined by
157 Braak stages (Figure 3a and Supplementary Table 9). Whether or not the modules showed expression
158 patterns that correlated with Braak stages, their expression in the arcuate nucleus of medulla, locus
159 coeruleus and CA2-field was consistently low (Figure 3b and Supplementary Figure 5). For the CA2-field
160 this might be explained by the presence of Lewy neurites rather than LBs¹⁸. Correlations with Braak
161 stages were mostly driven by the expression change between regions involved in preclinical stages (R1-
162 R3) and clinical stages (R4-R6), which were most evidently reflected in R3 or R4. In addition, regions R1-
163 R3 showed more extreme expression values (high and low) than in regions R4-R6.

164 We selected 23 co-expression modules for which the eigengene was significantly correlated with Braak
165 stages ($P < 0.0001$, BH-corrected). Module M39 showed the lowest correlation with Braak stages ($r = -$
166 0.87 and BH-corrected $P = 3.65e-7$), while M50 showed the highest correlation ($r = 0.92$ and BH-
167 corrected $P = 4.42e-7$). Most modules were significantly enriched for BRGs that were similarly correlated
168 with Braak stages (Figure 3c). For functional characterization, modules were further assessed for
169 enrichment of cell-type markers²³, and gene sets associated with functional GO-terms or diseases.

170 We found that modules that were negatively correlated with Braak stages were enriched for markers for
171 all different cell-types, and linked to various functions and diseases. M39 was enriched for markers of
172 astrocytes and endothelial cells, and the function membrane raft which plays a role in neurotransmitter
173 signaling. M127 was enriched for microglia and neurons, and associated with functional GO-terms such

174 as locomotory behavior and dopamine biosynthetic process, as well as diseases including
175 dopa-responsive dystonia, dystonia – limb, and tremor, highlighting their role in motor circuitry. M47 was
176 enriched for endothelial cell markers and genes involved in immune response, blood coagulation,
177 interferon-gamma-mediated signaling pathway, and oxygen transport. This module was also enriched for
178 genes involved in auto-inflammatory or auto-immunity disorders, e.g. hypersensitivity, infection, and
179 inflammation. These modules and their associated pathways were associated with the preclinical stages
180 of PD, because of their higher expression in regions R1-R3.

181 Modules that were positively correlated with Braak stages were specifically enriched for neuronal markers
182 and related functions (e.g. axon, cell junction, and chemical synaptic transmission) reflecting higher
183 expression of these modules in the synapse-dense cerebral cortex. Furthermore, M157 was enriched for
184 the function olfactory bulb development, M105 for functions such as cell junction, postsynaptic density,
185 calcium ion binding, and genes linked to bipolar disorder, M153 for functions cell junction and
186 postsynaptic membrane, and both M153 and M50 were linked to tobacco use disorder. Overall, gene co-
187 expression across Braak stage-related regions R1-R6 revealed interesting modules that highlight
188 pathways and potential gene interactions involved in preclinical and clinical stages of PD.

189 **Variable expression of BRGs is not fully explained by variations in cellular
190 composition**

191 We validated whether the identification of BRGs was confounded by variations in cellular compositions
192 across the six Braak stage-related regions R1-R6. We applied population-specific expression analysis
193 (PSEA)²⁴ to the AHBA to validate the cell-type specificity of each of the 960 BRGs. We found all 960
194 BRGs to be differentially expressed (BH-corrected $P < 0.05$ and $\beta \neq 0$) between regions R1 and R6 after
195 correcting for five major cell-types (neurons, astrocytes, oligodendrocytes, microglia, and endothelial
196 cells). For example, the neuronal marker *ADCY1* which was identified as a BRG remains differentially
197 expressed between regions R1 and R6 when corrected for neurons or other cell-types (Figure 4).
198 Similarly as for BRGs, PSEA analysis on all 23 Braak stage related co-expression modules showed
199 significant differential expression between regions R1 and R6 which cannot be fully explained by
200 differences in cellular composition.

201 In the PD datasets, not all BRGs were found significant after correction for cellular composition, however
202 smaller changes can be expected when comparing regions that are less distant (R1-R3 and R3-R4/R5).
203 Similar to the differential expression analysis without correction for cellular composition (Supplementary
204 Figure 3), PSEA revealed most changes between brain regions than between patients and controls
205 (Supplementary Figure 6).

206 **Expression of several Parkinson's disease-implicated genes are related to**
207 **Braak staging**

208 We found that the expression patterns of several PD-implicated genes, identified in the two most recent
209 PD genome-wide association studies^{7,8}, were correlated with the Braak LB staging scheme. These
210 included BRGs (*SCARB2*, *ELOVL7*, *SH3GL2*, *SNCA*, *BAP1*, and *ZNF184*; Table 1 and Supplementary
211 Figure 7) or genes present in Braak stage-related co-expression modules (*GCH1*, *ITIH3*, *ITPKB*, *RAB7L1*,
212 *BIN3*, *SATB1*, *ASHL1*, *MAPT*, *DLG2*, and *DNAH1*; Figure 3).

213 We further explored the relationship between *SNCA* expression and PD vulnerability in more detail.
214 *SNCA* was positively correlated with Braak stages in non-neurological brains from the AHBA, with a lower
215 expression in regions R1-R2 and higher expression in R3-R6 (Figure 5a, Figure 5b, and Figure 5c), which
216 was replicated in larger cohorts of non-neurological individuals (Figure 5d and Figure 5e). This
217 observation suggests that lower *SNCA* expression indicates high vulnerability of brain regions to develop
218 LB pathology. We further validated this concept in two cohorts of PD patients in which *SNCA* expression
219 similarly increased across the medulla oblongata (R1), locus ceruleus (R2), and substantia nigra (R3) of
220 PD-, iLBD patients, and age-matched controls. *SNCA* was significantly lower expressed in region R1
221 compared to R2 and R3 in PD- and iLBD patients, but not in controls (Figure 5f). In the RNA-seq dataset,
222 *SNCA* was significantly lower expressed in the substantia nigra (R3) compared to the medial temporal
223 gyrus (R4/R5) in PD patients, but again not in controls (Figure 5g). Altogether, *SNCA* expression patterns
224 could be replicated in brain regions of age-matched controls, however changes were larger between brain
225 regions in PD- and iLBD cases. We further assessed *SNCA* expression using PSEA in the AHBA (Figure
226 5h) and found that changes were independent of neuronal or other cell-type densities when comparing
227 different brain regions. In PD datasets results were scattered and did not align between both PD

228 microarray and RNAseq datasets because of the small sample sizes and the comparison of different
229 brain regions (Supplementary Figure 8).

230 Co-expression analysis in non-neurological brains from the AHBA revealed several dopaminergic genes
231 present in module M127. Their expression patterns were further investigated together with *SNCA* which is
232 also known to regulate dopamine homeostasis²⁵. *GCH1*, *TH*, and *SLC6A3* (*DAT*) are related to the
233 functional term dopamine biosynthetic process, and *SLC18A2* (*VMAT2*) is known to store dopamine into
234 synaptic vesicles²⁶. Unlike *SNCA*, the expression of *GCH1*, *TH*, *SLC6A3*, and *SLC18A2* was higher
235 expressed in regions involved at preclinical stages than those involved at clinical stages. Furthermore, all
236 these dopaminergic genes and *SNCA* showed a clear peak of expression in region R3 which includes the
237 substantia nigra, basal nucleus of Meynert, and CA2-field (Figure 6 and Supplementary Figure 9).

238 **Discussion**

239 In PD, the progressive accumulation of LB pathology across the brain follows a characteristic pattern,
240 which starts in the brainstem and subsequently evolves to more rostral sites of the brain (Braak
241 ascending scheme)¹. Using transcriptomic data of non-neurological brains, we identified genes (e.g.
242 *SNCA*, *SCARB2*, and *ZNF184*) and modules of co-expressed genes for which the expression decreased
243 or increased across brain regions defined by the Braak ascending scheme. Interestingly, these patterns
244 were disrupted in brains of patients with PD across regions that are preclinically involved in the
245 pathophysiology of PD. One gene co-expression module that showed higher expression in preclinically
246 involved regions was related to dopamine synthesis, locomotory behavior, and microglial and neuronal
247 activity. Another module was related to blood oxygen transport, the immune system, and may involve
248 endothelial cells. Our results highlight the complex genetic architecture of PD in which the combined
249 effects of genetic variants and co-expressed genes may underlie the selective regional vulnerability of the
250 brain.

251 Multiple studies suggests that a cytotoxic role and prion-like transfer of α -synuclein may contribute to its
252 progressive spread across the brain in PD, assuming a gain-of-function^{3,27,28}. In line with this assumption
253 are reports of familial PD caused by *SNCA* multiplications, suggesting a *SNCA* dosage effect in causing

254 PD^{5,29}. Interestingly, in contrast to the temporal and spatial pattern of the α -synuclein distribution
255 associated with the ascending Braak scheme in PD, the *SNCA* expression signature across brain regions
256 R1-R6 in non-neurological brains followed a reverse pattern with lowest expression in preclinically
257 involved regions (brainstem) and highest expression in clinically involved regions (limbic system and
258 cortex). Expression changes between regions were larger in PD and iLBD brains, because of lower
259 expression in preclinically involved regions compared to age-matched controls. The abundance of
260 physiological *SNCA* in non-neurological brains suggests a protective role, while at the same time it may
261 impact vulnerability to LB pathology in PD brains as demonstrated in earlier studies detecting both
262 proteins and mRNA levels (literature overview in Supplementary Table 10). Cell lines or animal models
263 without *SNCA* showed a synaptic deficit, increased susceptibility to viruses, sensitivity to reward, and
264 resulted in nigrostriatal neurodegeneration underscoring the importance of the presence of α -synuclein for
265 neuronal function. Mutant α -synuclein accelerated cell death induced by various stimuli (staurosporine,
266 serum deprivation, trypsin, or oxidative stress by H₂O₂), while wild-type α -synuclein exerted anti-apoptotic
267 effects. In contrast to the suggested neuroprotective role of α -synuclein, other studies suggest a
268 deleterious effect when overexpressed and that removing *SNCA* mediates resistance to LB pathology.
269 Collectively, our findings suggest that low *SNCA* expression in preclinically involved regions may increase
270 the vulnerability of brain regions to LB pathology.

271 Next to *SNCA*, the expression of several other genes known as genetic risk factors for PD^{7,8} were related
272 to the Braak staging scheme (Table 1 and Supplementary Figure 7). Two genes *ZNF184* (zinc finger
273 protein 184) and *ELOVL7* (fatty acid elongase 7) have recently been associated with early onset PD in a
274 Chinese population³⁰. *SCARB2* (scavenger receptor class B member 2) encodes for the lysosomal
275 integral membrane protein-2 (*LIMP2*), the specific receptor for glucocerebrosidase (GCase), and is
276 important for transport of GCase from the endoplasmic reticulum via Golgi to lysosomes³¹. *SCARB2*-
277 deficiency in mice brains led to α -synuclein accumulation mediating neurotoxicity in dopaminergic
278 neurons³¹. Overexpression in murine and human cell lines improved lysosomal activity of this enzyme
279 and enhanced α -synuclein clearance³¹. *SH3GL2* (SH3 Domain Containing GRB2 Like 2, Endophilin A1) is
280 thought to act downstream of *LRRK2* to induce synaptic autophagosome formation and may be
281 deregulated in PD³². *BAP1* (ubiquitin carboxyl-terminal hydrolase) is a deubiquitinase that acts as a tumor

282 suppressor. Cancer-associated mutations within this gene were found to destabilize protein structure
283 promoting β -amyloid aggregation in vitro, which is the pathological hallmark in Alzheimer's disease³³.

284 A number of functional pathways have been suggested to play a role in the pathogenesis of PD, such as
285 lysosomal function, immune system response, and neuroinflammation^{6–8,34}. We identified modules of
286 genes that co-express across the six Braak stage-related regions and found they were enriched for genes
287 related to molecular processes that have been linked to the (pre)clinical symptoms and functional deficits
288 in PD.

289 One negatively correlated module M127 was enriched for genes related to functions and diseases
290 involving dopamine synthesis and motor functions. This module also contains the PD variant-associated
291 gene *GCH1* (GTP cyclohydrolase 1) that is known to co-express with *TH* (tyrosine hydroxylase, the
292 enzyme responsible for converting tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA) in the dopamine
293 synthesis pathway) to enhance dopamine production and enable recovery of motor function in rat models
294 of PD³⁵. In this study, both *GCH1* and *TH* occur in M127 and thus were co-expressed across brain
295 regions involved in Braak stages supporting their interaction (Figure 6 and Supplementary Figure 9). The
296 higher expression in the more vulnerable brain regions R1-R3 indicates that *GCH1*, *TH*, and possibly
297 other genes within module M127 are essential to maintain dopamine synthesis that is affected in the early
298 Braak stages of PD. Indeed, by inhibiting *TH* activity, α -synuclein can act as a negative regulator of
299 dopamine release^{26,36}. In this module, *SLC18A2* (VMAT2; vesicular monoamine transporter 2) and
300 *SLC6A3* (DAT; dopamine transporter) were also present, which are important for storage of dopamine
301 and transport in the cell²⁶. Interestingly, dopamine may increase neuronal vulnerability, as was suggested
302 by an earlier study showing that α -synuclein is selectively toxic in dopaminergic neurons, and
303 neuroprotective in non-dopaminergic cortical neurons³⁷. Cell-type marker enrichment showed that module
304 M127 was enriched for microglia- and neuronal markers, suggesting a role in neuroinflammation. α -
305 Synuclein aggregates evoke microglia activation which in turn promotes aggregated protein propagation
306 to other brain regions, possibly even from the gut or periphery to the brain^{27,34}. The higher expression of
307 microglial genes within module M127 may contribute to the higher vulnerability of brain regions affected
308 during preclinical stages to form protein aggregates. Further investigation of genes within module M127

309 will provide a better understanding of the molecular mechanisms underlying microglia activation,
310 dopaminergic pathways and motor functions.

311 Another negatively correlated module M47 was enriched for endothelial cell markers and genes involved
312 in functions and disorders that relate to the immune response and oxygen transport in blood. One
313 previous case-control study showed that anemia or low hemoglobin levels may precede the onset of
314 PD³⁸. Several studies using blood transcriptomic meta-analysis revealed genes associated with
315 hemoglobin and iron metabolism were downregulated in PD patients compared to controls³⁹⁻⁴¹. In our
316 study, several hemoglobin genes (*HBD*, *HBB*, *HBA1*, *HBA2*, and *OASL*) were also present in module
317 M47 of which *HBD* and *HBB* have been described to be highly interconnected with *SNCA*⁴¹. We also
318 found an association between the interferon-gamma-mediated signaling pathway and M47 in which *OASL*
319 also plays a role. Module M47 was negatively co-expressed with *SNCA*. Notably, a significant loss of
320 negative co-expression between *SNCA* and interferon-gamma genes in the substantia nigra has been
321 demonstrated in PD patients as compared to controls⁴². This loss may result from a downregulation of
322 genes within M47 in the substantia nigra of PD patients, similarly as was observed in PD blood
323 transcriptomics³⁹⁻⁴¹. Therefore, these genes have the potential to serve as blood biomarkers for PD
324 vulnerability. Overall, these studies suggest that dysregulation of genes within module M47 involved in
325 blood-oxygen transport and the immune system influence brain regions to be selectively vulnerable to
326 PD.

327 Identification of transcriptomic features in regions or disease conditions may be confounded by changes
328 in cell-type composition. We used PSEA²⁴ to examine the impact of this confounding factor and found that
329 all 960 BRGs remained differentially expressed between regions R1 and R6 in AHBA. We also applied
330 PSEA in the two PD datasets that allowed us to examine cell-type specificity between regions as well as
331 between disease conditions. Although it is known that gene expression varies more between regions than
332 between disease conditions²², it is less clear how cell-type composition contributes to this variation. Here,
333 we found that regional comparisons yielded more significant results than when comparing disease
334 conditions. Therefore, BRGs also captured expression changes between patients and controls, but

335 changes were less dependent on cell-type abundance between regions than between patients and
336 controls.

337 In conclusion, we identified genes and pathways that may be important to maintain biological processes
338 in specific brain regions, but may also contribute to a higher selective vulnerability to PD. Our results
339 suggest that interactions between microglial genes and genes involved in dopamine synthesis and motor
340 functions, as well as between genes involved in blood-oxygen transport and the immune system may
341 contribute to the early involvement of specific brain regions in PD progression. Our observations highlight
342 a potential complex interplay of pathways in healthy brains and provide clues for future genetic targets
343 concerning the pathobiology in PD brains.

344 **Methods**

345 Methods and any associated references are available in the online version of the paper.

346 **References**

- 347 1. Braak, H. *et al.* Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol.*
348 *Aging* **24**, 197–211 (2003).
- 349 2. Surmeier, D. J., Obeso, J. A. & Halliday, G. M. Selective neuronal vulnerability in Parkinson
350 disease. *Nat Rev Neurosci.* **3**, 973–982 (2016).
- 351 3. Borghammer, P. How does parkinson's disease begin? Perspectives on neuroanatomical
352 pathways, prions, and histology. *Mov. Disord.* **33**, 48–57 (2018).
- 353 4. Chartier-Harlin, M.-C. *et al.* α -synuclein locus duplication as a cause of familial Parkinson's
354 disease. *Lancet* **364**, 1167–1169 (2004).
- 355 5. Singleton, A. B. *et al.* α -Synuclein Locus Triplication Causes Parkinson's Disease. *Science* **302**,
356 841 (2003).
- 357 6. Bonifati, V. Genetics of Parkinson's disease – state of the art, 2013. *Parkinsonism Relat. Disord.*
358 **20**, S23–S28 (2014).

359 7. Chang, D. *et al.* A meta-analysis of genome-wide association studies identifies 17 new
360 Parkinson's disease risk loci. *Nat. Genet.* 1–6 (2017). doi:10.1038/ng.3955

361 8. Nalls, M. A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new
362 risk loci for Parkinson's disease. *Nat. Publ. Gr.* **46**, 989–993 (2014).

363 9. Glaab, E. & Schneider, R. Neurobiology of Disease Comparative pathway and network analysis of
364 brain transcriptome changes during adult aging and in Parkinson's disease. *Neurobiol. Dis.* **74**, 1–
365 13 (2015).

366 10. Riley, B. E. *et al.* Systems-based analyses of brain regions functionally impacted in Parkinson's
367 disease reveals underlying causal mechanisms. *PLoS One* **9**, 1–14 (2014).

368 11. Dijkstra, A. A. *et al.* Evidence for immune response, axonal dysfunction and reduced endocytosis
369 in the substantia nigra in early stage Parkinson's disease. *PLoS One* **10**, 1–21 (2015).

370 12. Oerton, E. & Bender, A. Concordance analysis of microarray studies identifies representative gene
371 expression changes in Parkinson's disease: a comparison of 33 human and animal studies. *BMC*
372 *Neurol.* **17**, 1–14 (2017).

373 13. Freer, R. *et al.* A protein homeostasis signature in healthy brains recapitulates tissue vulnerability
374 to Alzheimer's disease. *Sci. Adv.* **2**, 1–8 (2016).

375 14. Sepulcre, J. *et al.* Neurogenetic contributions to amyloid beta and tau spreading in the human
376 cortex. *Nat. Med.* **24**, 1910–1918 (2018).

377 15. Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat. Neurosci.* **18**,
378 1832–1844 (2015).

379 16. Rittman, T. *et al.* Regional expression of the *MAPT* gene is associated with loss of hubs in brain
380 networks and cognitive impairment in Parkinson disease and progressive supranuclear palsy.
381 *Neurobiol. Aging* **48**, 153–160 (2016).

382 17. Freeze, B. S., Acosta, D., Pandya, S., Zhao, Y. & Raj, A. Regional expression of genes mediating

383 trans-synaptic alpha-synuclein transfer predicts regional atrophy in Parkinson disease.

384 *NeuroImage Clin.* **18**, 456–466 (2018).

385 18. Alafuzoff, I. *et al.* Staging/typing of Lewy body related α -synuclein pathology: A study of the
386 BrainNet Europe Consortium. *Acta Neuropathol.* **117**, 635–652 (2009).

387 19. Carithers, L. J. *et al.* A novel approach to high-quality postmortem tissue procurement: The GTEx
388 project. *Biopreserv Biobank* **13**, 311–319 (2015).

389 20. Trabzuni, D. *et al.* Quality control parameters on a large dataset of regionally dissected human
390 control brains for whole genome expression studies. *J. Neurochem.* **119**, 275–282 (2011).

391 21. Hurley, M. J., Durrenberger, P. F., Gentleman, S. M., Walls, A. F. & Dexter, D. T. Altered
392 Expression of Brain Proteinase-Activated Receptor-2, Trypsin-2 and Serpin Proteinase Inhibitors
393 in Parkinson's Disease. *J. Mol. Neurosci.* **57**, 48–62 (2015).

394 22. Melé, M. *et al.* The human transcriptome across tissues and individuals. *Science* **348**, 660–665
395 (2015).

396 23. Zhang, Y. *et al.* An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and
397 Vascular Cells of the Cerebral Cortex. *J. Neurosci.* **34**, 11929–11947 (2014).

398 24. Kuhn, A., Thu, D., Waldvogel, H. J., Faull, R. L. M. & Luthi-Carter, R. Population-specific
399 expression analysis (PSEA) reveals molecular changes in diseased brain. *Nat. Methods* **8**, 945–
400 947 (2011).

401 25. Deng, H. & Yuan, L. Genetic variants and animal models in SNCA and Parkinson disease. *Ageing*
402 *Res. Rev.* **15**, 161–176 (2014).

403 26. Venda, L. L., Cragg, S. J., Buchman, V. L. & Wade-, R. α -Synuclein and dopamine at the
404 crossroads of Parkinson's disease. *Trends Neurosci.* **33**, 559–568 (2013).

405 27. Steiner, J. A., Quansah, E. & Brundin, P. The concept of alpha-synuclein as a prion-like protein :
406 ten years after. *Cell Tissue Res.* (2018). doi:10.1007/s00441-018-2814-1

407 28. Sulzer, D. *et al.* T cells from patients with Parkinson's disease recognize α -synuclein peptides.
408 *Nature* **546**, 656–661 (2017).

409 29. Chartier-Harlin, M.-C. *et al.* α -synuclein locus duplication as a cause of familial Parkinson's
410 disease. *Lancet* **364**, 1167–1169 (2004).

411 30. Li, G. *et al.* Association of *GALC*, *ZNF184*, *IL1R2* and *ELOVL7* With Parkinson's Disease in
412 Southern Chinese. *Front. Aging Neurosci.* **10**, 1–6 (2018).

413 31. Rothenberg, M., Zunke, F., Mazzulli, J. R., Schweizer, M. & Altmeppen, H. LIMP-2 expression is
414 critical for β -glucocerebrosidase activity and α -synuclein clearance. *Proc. Natl. Acad. Sci.* **111**,
415 15573–15578 (2014).

416 32. Soukup, S. & Verstreken, P. EndoA/Endophilin-A creates docking stations for autophagic proteins
417 at synapses. *Autophagy* **13**, 971–972 (2017).

418 33. Bhattacharya, S., Hanpude, P. & Maiti, T. K. Cancer associated missense mutations in BAP1
419 catalytic domain induce amyloidogenic aggregation: A new insight in enzymatic inactivation. *Sci.*
420 *Rep.* **5**, doi: 10.1101/18462 (2015).

421 34. Sampson, T. R. *et al.* Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of
422 Parkinson's Disease. *Cell* **167**, 1469–1480 (2016).

423 35. Cederfjäll, E., Sahin, G., Kirik, D. & Björklund, T. Design of a Single AAV Vector for Coexpression
424 of TH and GCH1 to Establish Continuous DOPA Synthesis in a Rat Model of Parkinson's Disease.
425 *Mol. Ther.* **20**, 1315–1326 (2012).

426 36. Abeliovich, A. *et al.* Mice Lacking α -Synuclein Display Functional Deficits in the Nigrostriatal
427 Dopamine System. *Cell Neuron* **25**, 239–252 (2000).

428 37. Xu, J. *et al.* Dopamine-dependent neurotoxicity of α -synuclein : A mechanism for selective
429 neurodegeneration in Parkinson disease. *Nat. Med.* **8**, 600–606 (2002).

430 38. Savica, R. Anemia or low hemoglobin levels preceding Parkinson disease. *Neurology* **73**, 1381–

431 1388 (2009).

432 39. Mutez, E. *et al.* Transcriptional profile of Parkinson blood mononuclear cells with LRRK2 mutation.

433 *NBA* **32**, 1839–1848 (2011).

434 40. Mutez, E. *et al.* Neurobiology of Disease Involvement of the immune system, endocytosis and

435 EIF2 signaling in both genetically determined and sporadic forms of Parkinson's disease.

436 *Neurobiol. Dis.* **63**, 165–170 (2014).

437 41. Santiago, J. A. & Potashkin, J. A. Blood Transcriptomic Meta-analysis Identifies Dysregulation of

438 Hemoglobin and Iron Metabolism in Parkinson' Disease. *Front. Aging Neurosci.* **9**, 1–8 (2017).

439 42. Liscovitch, N. & French, L. Differential co-expression between α -synuclein and IFN- γ signaling

440 genes across development and in Parkinson's disease. *PLoS One* **9**, 1–13 (2014).

441 43. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis.

442 *BMC Bioinformatics* **9**, (2008).

443 44. Smedley, D. *et al.* The BioMart community portal: An innovative alternative to large, centralized

444 data repositories. *Nucleic Acids Res.* **43**, W589–W598 (2015).

445 45. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-

446 seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).

447 46. Piñero, J. *et al.* DisGeNET: A comprehensive platform integrating information on human disease-

448 associated genes and variants. *Nucleic Acids Res.* **45**, D833–D839 (2017).

449 **Acknowledgements**

450 We thank S.M.H. Huisman and prof. J.J. Goeman for their support in the statistical analyses. We also

451 want to thank G. Bonvicini for her help in accessing the RNA-seq data from PD patients and V. Bonifati

452 for critical discussions during the preparation of the manuscript.

453 **Author contributions**

454 AK, AM, BL, JvH, and MR designed the study. AK wrote the scripts and performed the data analyses.
455 EM, MCCH, MF and WvB designed the RNA-seq experiments, designed, and realized the RNA-seq
456 sample comparisons and validations. WvB selected brain tissue samples for RNA-seq and AI processed
457 the brain tissue samples. TC and CV prepared the sequencing libraries. JPM performed the analysis of
458 the PD RNA-seq data (quality control, alignment, normalization) under the supervision of MF. AK, AM,
459 MCCH, WvB, JvH, and MR interpreted the data and wrote the manuscript with input from all authors. AM
460 and MR supervised the overall project. The manuscript was read and approved by all authors.

461 **Grants**

462 This research received funding from The Netherlands Technology Foundation (STW), as part of the STW
463 project 12721 (Genes in Space, PI Lelieveldt).

464 WvB and MCCH received funding from Alzheimer Netherlands and LECMA to collect the RNA-
465 sequencing datasets that are used in this study. WvB was financially supported by grants from
466 Amsterdam Neuroscience, Dutch Research council (ZonMW), Stichting Parkinson Fonds, Alzheimer
467 association, and Rotary Aalsmeer-Uithoorn. WvB performed contract research and consultancy for Roche
468 Pharma, Lysosomal Therapeutics, CHDR, Cross beta Sciences and received research consumables from
469 Roche and Prothena. MCCH was financially supported by grants from INSERM, CHU de Lille, Université
470 de Lille, BiLille, Vaincre Alzheimer, French Health Ministry for the PHRCs, French National Research
471 Agency, and the Michael J. Fox Foundation for Parkinson's Research.

472 **Competing interests**

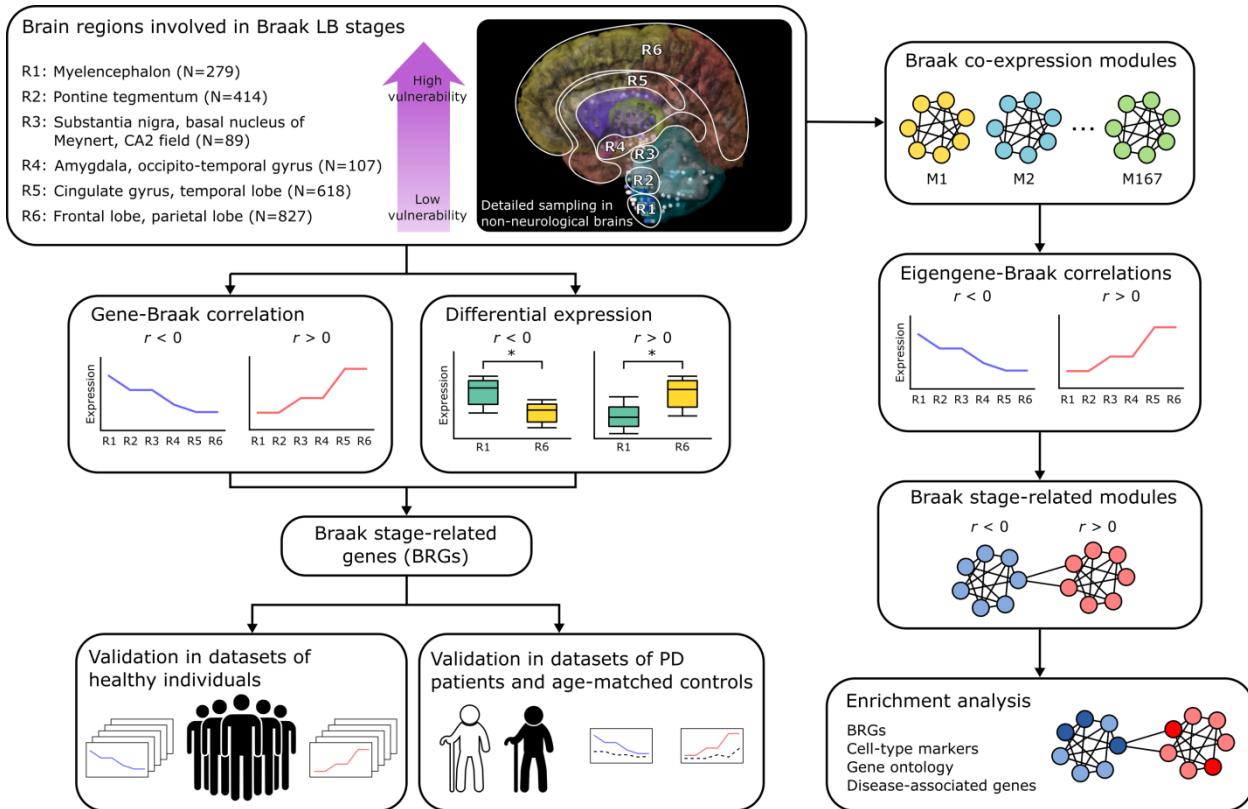
473 The authors declare no competing interests.

474 **Additional information**

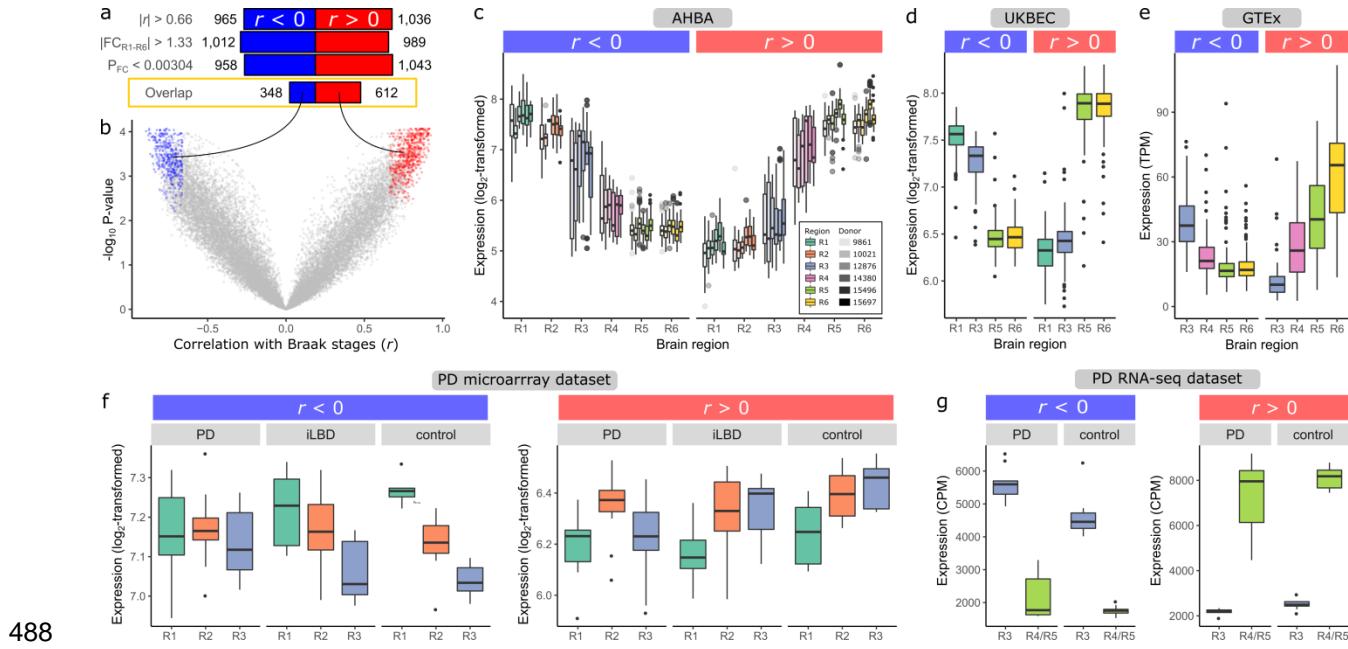
475 Supplementary information is available for this paper.

476

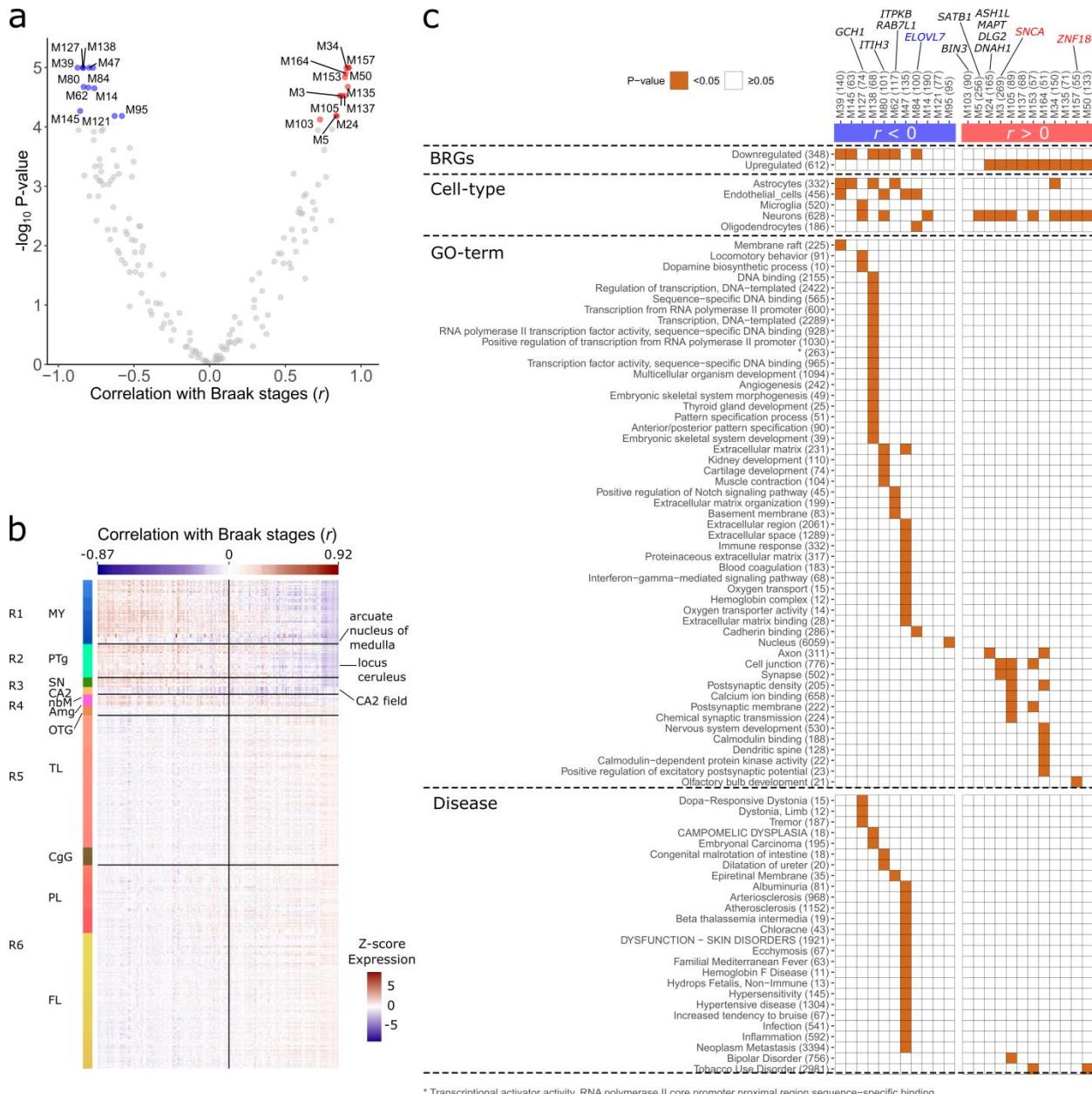
477 Figures & Tables



479 **Figure 1 Study overview.** Differential vulnerability to PD was examined across brain regions R1-R6. N is the tissue
480 sample size across all six non-neurological donors from the Allen Human Brain Atlas (AHBA), which are involved in
481 the six PD Braak stages as they sequentially accumulate Lewy bodies during disease progression (Supplementary
482 Table 1 and Supplementary Figure 1). Through correlation and differential expression analysis, we identified Braak
483 stage-related genes (BRGs) with expression patterns that are either positively ($r > 0$) or negatively ($r < 0$) correlated
484 with Braak stages in the non-neurological brain. These were validated in cohorts of non-neurological individuals and
485 subsequently in PD patients and age-matched controls. To obtain a more global view of BRG expression signatures,
486 we focused on co-expression modules of genes and correlated the module eigengene expression with Braak stages.
487 The resulting modules of genes were subsequently analyzed to detect common biologically meaningful pathways.



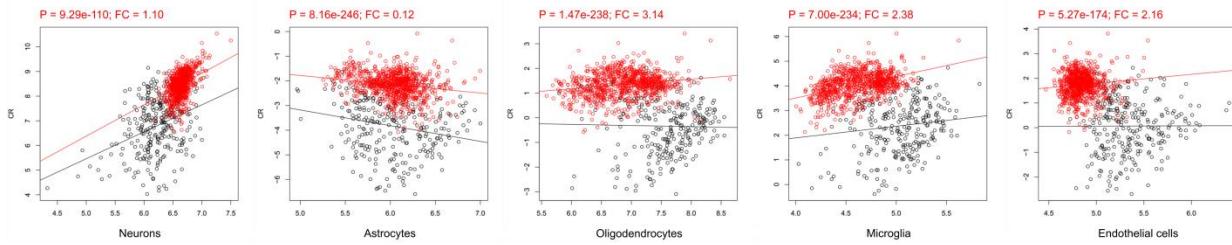
489 **Figure 2 Expression patterns of Braak stage-related genes across brain regions of non-neurological, iLBD**
490 **and PD brains.** (a) Selection of BRGs that were either negatively (blue) or positively (red) correlated with Braak
491 stages. Genes were selected based on 1) highest absolute correlation of gene expression and Braak stage labels, 2)
492 highest absolute fold-change (FC) between R1 and R6, and 3) lowest P -value of FC in the differential expression
493 analysis (P_{FC} ; BH-corrected), for which the top 10% (2,001) genes resulted in the shown thresholds. The overlap
494 between the three sets of top 10% genes resulted in 960 BRGs. (b) Correlation r of BRGs (red and blue points) with
495 Braak stages (x-axis) and $-\log_{10}$ BH-corrected P -value (y-axis). (c) Mean expression of BRGs for each region
496 (colors) and donor (opacity) in AHBA (number of samples in Supplementary Table 1). (d) Validation across 134 non-
497 neurological individuals in UK Brain Expression Consortium (UKBEC; R1: medulla, R3: substantia nigra, R5: temporal
498 cortex, R6: frontal cortex), and (e) 88-129 non-neurological individuals in Genotype-Tissue Expression consortium
499 (GTEx; R3: substantia nigra, R4: amygdala, R5: anterior cingulate cortex, R6: frontal cortex). Each data point is the
500 mean expression of a BRG across brain donors in one brain region. (f) Validation in PD microarray dataset (R1:
501 medulla oblongata, R2: locus ceruleus, R3: substantia nigra; number of samples in Supplementary Table 5) and (g)
502 PD RNA-seq dataset (R3: substantia nigra, R4/R5: medial temporal gyrus; number of samples in Supplementary
503 Table 7). Boxplots are shown (f, g) per patient group (PD, incidental Lewy body disease (iLBD), and control) and
504 (Supplementary Figure 4) per brain region.



505

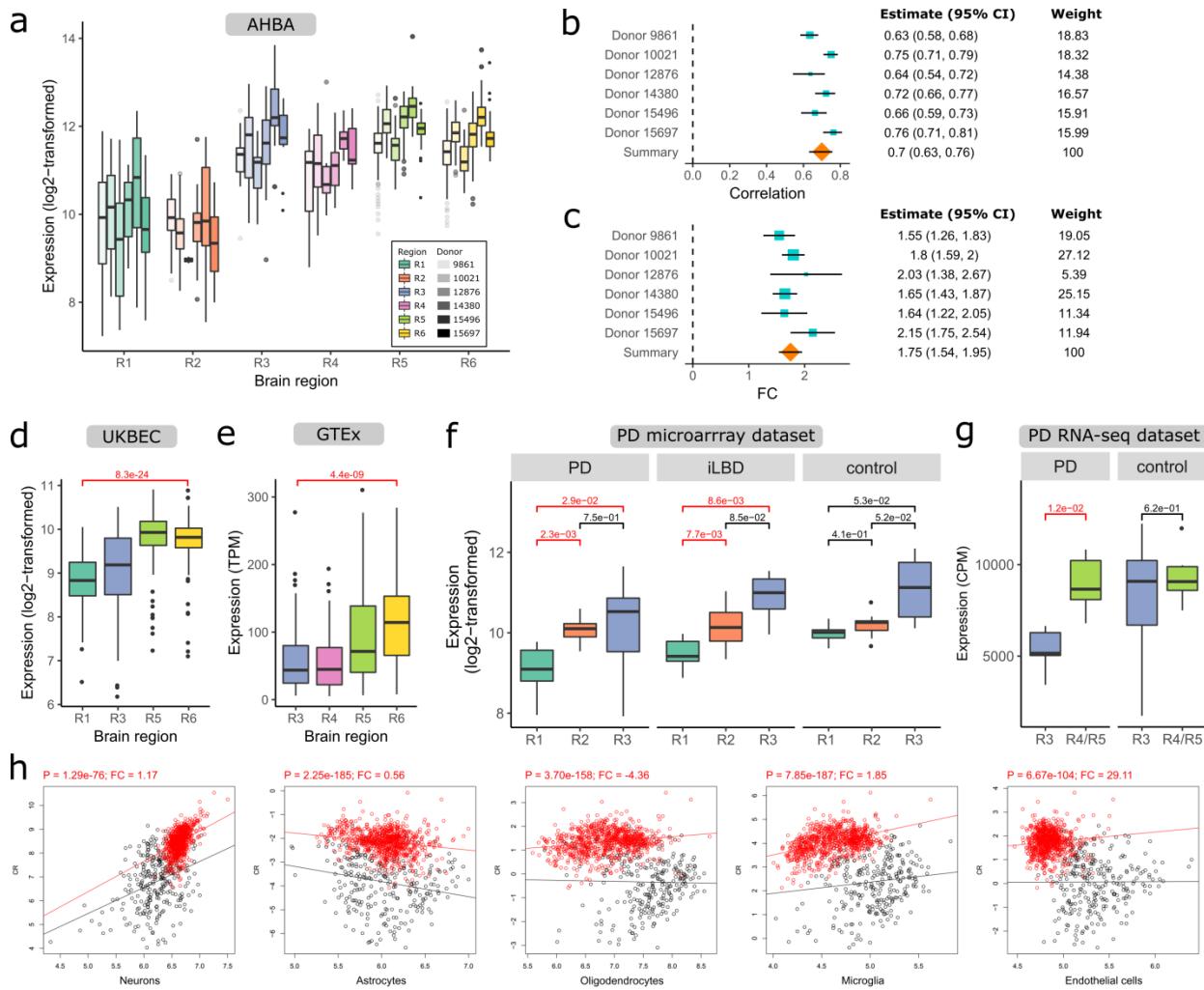
506 **Figure 3 Braak co-expression modules.** Genes were analyzed for co-expression across regions R1-R6 in the Allen
 507 Human Brain Atlas. (a) Module eigengene correlation with Braak. Each point reflects a module showing its correlation
 508 r with Braak stages (x-axis) and \log_{10} -transformed P -values (BH-corrected; y-axis); 23 significant modules (BH-
 509 corrected $P < 0.0001$; $H_0: r = 0$) were selected for further analysis (blue and red). (b) Eigengene expression of all 167
 510 modules across brain regions (rows) of donor 9861 sorted by their correlation with Braak stages (column colors). The
 511 vertical line separates negatively and positively correlated modules, and correlations are shown for the modules with
 512 the lowest and highest correlation. Brain regions involved in Braak includes the following anatomical structures:
 513 myelencephalon (MY), pontine tegmentum (PTg), substantia nigra (SN), CA2-field (CA2), basal nucleus of Meynert

514 (nbM), amygdala (Amg), occipito-temporal gyrus (OTG), temporal lobe (TL), cingulate gyrus (CgG), parietal lobe (PL),
515 and frontal lobe (FL). Modules were low expressed in the arcuate nucleus of medulla, locus ceruleus and CA2-field,
516 independently of their correlation with Braak stages. (c) Significant modules were sorted based on their correlation
517 with Braak stages (columns) and assessed for significant overlap with BRGs, cell-type markers, and gene sets
518 associated with functional GO-terms or diseases using a hypergeometric test (brown squares; BH-corrected $P <$
519 0.05). Additionally, these modules reveal the presence of genes associated with PD variants (annotated at the top)
520 that have (blue and red) and have not (black) been identified as BRGs.



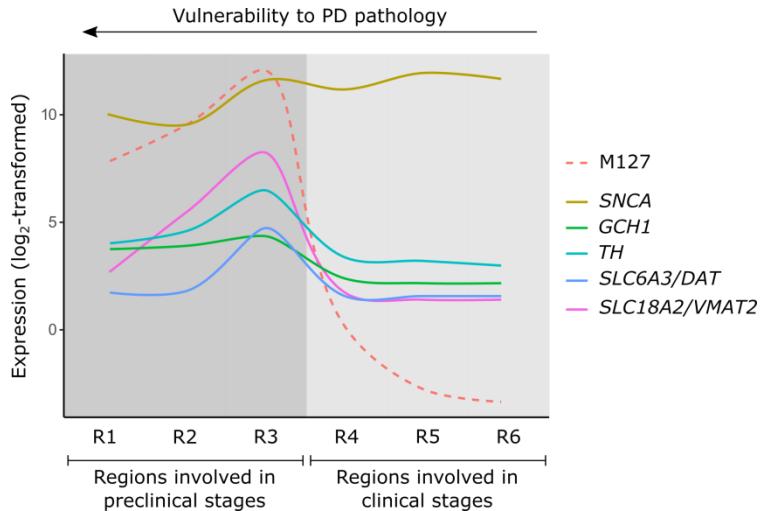
521

522 **Figure 4 Differential expression of neuronal marker ADCY1 in AHBA corrected for cell-type abundance.**
523 ADCY1 is a neuronal marker identified as one of the 960 Braak stage-related genes (BRGs). We found it is still
524 significantly differentially expressed between samples from region R1 (black) and R6 (red) when correcting for one of
525 the five main cell-types with PSEA²⁴ (BH-corrected $P < 0.05$ for 960 BRGs). Significant BH-corrected P -values are
526 highlighted in red text together with cell-type specific fold-changes (FC; slope change of red line).



528 **Figure 5 SNCA expression in brain regions R1-R6 of non-neurological individuals (a-e, h) and PD patients (f, 529 g).** (a) Boxplots of SNCA expression in regions R1-R6 (colored) for each donor (opacity) in AHBA (number of 530 samples in Supplementary Table 1). Meta-analysis of (b) SNCA correlation with Braak across the six donors in AHBA 531 and (c) SNCA expression fold-change between region R1 and R6. To calculate the summary effect size, the 95% 532 confidence intervals (CI) and weights are taken into account. The positive correlation with Braak was validated in 533 datasets from two healthy cohorts, (d) UK Brain Expression Consortium (UKBEC; 134 donors) and (e) Genotype- 534 Tissue Expression consortium (GTEx; 88-129 donors), and (f, g) two PD cohorts with PD patients, incidental Lewy 535 body disease (iLBD) and non-demented age-matched controls (number of samples in Supplementary Table 5 and 7). 536 In the PD datasets, SNCA expression was tested for differential expression between (f, g) regions and conditions 537 (Supplementary Figure 6) and was considered significant when BH-corrected $P < 0.05$ (red). (h) SNCA was still 538 significantly differentially expressed between region R1 (black) and R6 (red) when correcting for five main cell-types 539 with PSEA in the AHBA (BH-corrected $P < 0.05$ for 960 BRGs). Significant BH-corrected P -values are highlighted in

540 red text together with cell-type specific fold-changes (FC; slope change of red line). PSEA results for PD data are
541 shown in Supplementary Figure 8.



543 **Figure 6 Schematic overview of molecular activity of dopaminergic genes in module M127 and SNCA across**
544 **brain regions of the Braak staging scheme.** Lines across regions R1-R6 were based on transcriptomic data from
545 the Allen Human Brain Atlas (Figure 1 and Supplementary Figure 9). Expression of module M127 is in the eigengene
546 space. Genes showed peak activity in region R3 that includes the substantia nigra, basal nucleus of Meynert, and
547 CA2-field. While *SNCA* was generally high expressed in all regions, other genes were low or not expressed in other
548 regions than R3. *SNCA*: responsible for dopamine release, *GCH1*: together with *TH* required for production of
549 dopamine, *TH*: catalyzes tyrosine to the dopamine precursor L-3,4-dihydroxyphenylalanine (L-DOPA), *SLC6A3/DAT*:
550 dopamine transporter; transports dopamine from the synaptic cleft back to the cytosol, *SLC18A2/VMAT2*: stores
551 dopamine into synaptic vesicles.

552

553 **Table 1 Braak stage-related genes that have previously been associated with PD.** Several PD variant-
554 associated genes showed expression profiles that are correlated with Braak stages. The correlation r with Braak, fold-
555 change, and P -value of fold-change (t-test, BH-corrected) are within the selection thresholds for BRGs.

Gene symbol	Entrez ID	Correlation with Braak (r)	P -value (BH-corrected)	Fold- change	P -value (BH-corrected)	Module	Reference
						member	
SCARB2	950	-0.78	4.4e-04	-1.44	1.7e-03	M130	Nalls et al. 2014
<i>ELOVL7</i>	79993	-0.67	7.2e-04	-1.35	1.4e-03	M84	Chang et al. 2017
<i>SH3GL2</i>	6456	0.70	4.5e-04	1.4	2.3e-03	-	Chang et al. 2017
SNCA	6622	0.70	3.3e-04	1.75	4.3e-04	M3	Nalls et al. 2014; Chang et al. 2017
<i>BAP1</i>	8314	0.77	3.2e-03	1.99	1.6e-03	M85	Chang et al. 2017
<i>ZNF184</i>	7738	0.81	4.6e-04	2.34	2.9e-03	M157	Chang et al. 2017

556

557 Online methods

558 **Allen Human Brain Atlas.** To examine gene expression patterns across brain regions involved in
559 Parkinson's disease (PD), we used normalized gene expression data from the Allen Human Brain Atlas
560 (AHBA), a microarray data set of 3,702 anatomical brain regions from six non-neurological individuals (5
561 males and 1 female, mean age 42, range 24–57 years¹⁵). We downloaded the data from
562 <http://human.brain-map.org/>. To filter and map probes to genes, the data was concatenated across the six
563 donors. We removed 10,521 probes with missing Entrez IDs, and 6,068 probes with low presence as they
564 were expressed above background in <1% of samples (PA-call containing presence/absence flag¹⁵). The
565 remaining 44,072 probes were mapped to 20,017 genes with unique Entrez IDs using the *collapseRows*-
566 function in R-package WGCNA v1.64.1⁴³ as follows: i) if there is one probe, that one probe is chosen, ii) if
567 there are two probes, the one with maximum variance across all samples is chosen
568 (method="maxRowVariance"), iii) if there are more than two probes, the probe with the highest
569 connectivity (summed adjacency) is chosen (connectivityBasedCollapsing=TRUE). Based on the
570 anatomical labels given in AHBA, samples were mapped to Braak stage-related regions R1-R6 as
571 defined by the BrainNet Europe protocol¹⁸ and each region corresponds to one or multiple anatomical
572 structures. The locus ceruleus and pontine raphe nucleus are both part of the pontine tegmentum in R2.

573 **Braak stage-related genes (BRGs).** Two analysis methods were used to find genes for which the spatial
574 expression in AHBA is related to the spread of the disease: i) correlations between gene expression and
575 labels 1-6 according to their assignment to one of the Braak stage-related regions R1-R6, and ii)
576 differential expression between Braak stage-related regions R1 and R6. As the expression values were
577 log₂-transformed, the mean difference between two regions was interpreted as the fold-change (FC).
578 Genes were assigned as BRGs based on the overlap of the top 10% (2,001) ranked genes with: i) highest
579 absolute correlation of gene expression and Braak stage labels, ii) highest absolute FC between R1 and
580 R6, and iii) lowest Benjamini-Hochberg (BH) corrected *P*-value of the FC.

581 To avoid capturing donor-specific changes, we applied the correlation and differential expression
582 analyses for each of the six brain donors separately, and effect sizes were then combined by meta-
583 analysis (metafor R-package 2.0). A random effects model was applied which assumes that each brain is
584 considered to be from a larger population of brains and therefore takes the within-brain and between-
585 brain variance into account. The between-brain variance (τ^2) was estimated with the Dersimionian-
586 Delaird model. Variances and confidence intervals were obtained using the *escalc*-function. Correlations
587 were Fisher-transformed ($z = \text{arctanh}(r)$) to obtain summary estimates, which were then back-transformed
588 to correlation values ranging between -1 and +1. *P*-values were BH-corrected for all 20,017 genes. The
589 significance of summary effect sizes (correlations and FCs) was assessed through a two-sided t-test (H_0 :
590 $FC=0$; unequal variances). The weights used in the meta-analysis are based on the non-pooled
591 expression variance in R1-R6.

592 **UK Brain Expression Consortium (UKBEC).** UKBEC²⁰ (<http://www.braineac.org>) contains microarray
593 expression data from 10 brain regions of 134 non-neurological donors (74.5% males, mean age 59, range
594 16–102 years) for which their control status was confirmed by histology. We used the biomaRt R-package
595 version 2.38⁴⁴ to map Affymetrix probe IDs from UKBEC to gene Entrez IDs; 262,134 out of 318,197
596 probes could be mapped. Similar as for AHBA, expression data for all probes and samples was
597 concatenated across the 10 brain regions before mapping probes to 18,333 genes with unique Entrez IDs
598 using the *collapseRows*-function.

599 **Genotype-Tissue expression consortium (GTEx).** From GTEx¹⁹ (<https://gtexportal.org>), we obtained
600 RNA-sequencing (RNA-seq) samples from four brain tissues from multiple non-neurological subjects
601 (65.7% males, range 20–79 years): substantia nigra (88 samples), amygdala (121 samples), anterior
602 cingulate cortex (100 samples), and frontal cortex (129 samples). These brain regions corresponded to
603 Braak stage-related regions R3-R6, respectively. We downloaded gene read counts (v7) for differential
604 expression analysis and gene transcript per million (TPM) expression values (v7) for visualization. From
605 56,202 genes, we selected 19,820 protein coding genes and removed 405 genes with zero counts in one
606 of the four regions of interest; 19,415 genes were left for analysis.

607 **PD microarray dataset.** In the PD microarray dataset, samples were collected from the medulla
608 oblongata (R1), locus ceruleus (R2), and substantia nigra (R3) from PD- (67.6% males, mean age 78,
609 range 61–87 years) and, incidental Lewy body disease (iLBD) patients (42.4% males, mean age 80,
610 range 56–98 years), and non-demented controls (54.5% males, mean age 77, range 60–91 years)
611 (Supplementary Table 5 and 6). The PD microarray data of the substantia nigra (R3) was previously
612 published in Dijkstra et al.¹¹. Based on pathological examination, PD patients in the microarray dataset
613 revealed Lewy body (LB) pathology in accordance with Braak stages 4-6, and iLBD patients showed LB
614 pathology in the brainstem (Braak stages 1-3), and therefore represent the early stages of PD. Additional
615 samples from the medulla oblongata (R1) and locus ceruleus (R2) were collected and processed of the
616 same cohort in the same manner for hybridization on GeneChip® Human Genome U 133 Plus 2.0 arrays.
617 Probe IDs were mapped to Entrez IDs with the mapIds-function in the hgu133plus2.db R-package v3.2.3.
618 We removed 10,324 out of 54,675 probes with missing Entrez IDs. The remaining 44,351 probes were
619 mapped to 20,988 genes with unique Entrez IDs using the *collapseRows* function similarly as was done
620 for AHBA.

621 **PD RNA-sequencing dataset.** In the PD RNA-seq dataset, samples from the substantia nigra (R3) and
622 medial temporal gyrus (R4/R5) were collected from PD patients (61.1% males, mean age 79, range 57–
623 88 years), and non-demented age-matched controls (48.0% males, mean age 78, range 59–93 years)
624 (Supplementary Table 7 and 8). The extracted RNA was quantified using an Ozyme NanoDrop System,
625 of which 500 ng of total RNA from each sample was further processed for purification of ribosomal RNA
626 (rRNA) using human Illumina Ribo-Zero™ rRNA Removal Kit. Then Illumina TruSeq stranded total RNA
627 protocol was used for library preparation. The library was sequenced on a Hiseq4000. RNA-seq reads
628 were aligned to the human genome (GRCh 38) with TopHat software (version: 2.1.1) using reference
629 gene annotations (Ensembl GRCh38.p3) to guide the alignment. The count of reads per gene were
630 determined from the alignment file (bam) and reference gene annotations (Ensembl) using
631 FeatureCounts software (version: 1.5.3), resulting in 52,411 transcripts with Ensembl IDs. Entrez IDs of
632 20,017 genes in AHBA were mapped to Ensembl IDs using biomaRt R-package version 2.38.

633 For both the PD microarray and PD RNA-seq dataset, a written informed consent for a brain autopsy and
634 the use of the material and clinical information for research purposes had been obtained from all donors
635 by The Netherlands Brain bank (NBB), Amsterdam. Autopsy was performed using a standardized
636 protocol by NBB (open access: www.brainbank.nl). All procedures of NBB were approved by the local
637 VUmc medical ethics committee.

638 **Differential gene expression between brain regions and conditions.** A two-sided unpaired t-test was
639 used to assess expression differences between conditions (PD, iLBD, and controls) and brain regions
640 (R1-R6) in the AHBA, UKBEC, and PD microarray dataset. For GTEx, we used DESeq2 version 1.22.2⁴⁵.
641 For the PD RNA-seq dataset, normalization and differential expression was done with 'DESeq2' R-
642 package version 1.10.1, with age and sex introduced in the statistical model to take into account possible
643 biases. *P*-values were BH-corrected across all genes. The cut-off for differentially expressed genes was *P*
644 < 0.05 (BH-corrected). For microarray experiments, the FC was interpreted as the difference in mean
645 expression $\mu_B - \mu_A$, with μ as the mean expression in either group A and B. For RNA-seq experiments, FC
646 is the log₂ fold-change obtained from DESeq2.

647 **Cell type-specific analysis.** To assess whether results were confounded by cell-type composition in
648 different brain regions and conditions, we applied population-specific expression analysis (PSEA)²⁴ in the
649 AHBA, PD microarray, and PD RNA-seq datasets. Data from AHBA were first concatenated across the
650 six donors before applying PSEA. This method applies linear regression to examine whether the
651 expression between two groups of samples is different while correcting for cell-type composition
652 estimated from cell-type markers. To define cell-type markers, we used gene expression data from sorted
653 cells of the mouse cerebral cortex²³. Genes were selected as markers when they had a 20-fold higher
654 expression compared to the mean of the other cell-types. All genes were analyzed while correcting for
655 five main cell-types for which the cell-type signal was estimated by taking the mean expression of
656 markers: 628 neurons, 332 astrocytes, 186 oligodendrocytes, 520 microglia, and 456 endothelial cells. *P*-
657 values were BH-corrected across all genes in a dataset.

658 **Gene co-expression modules in Braak stage-related regions R1-R6.** Co-expression matrices
659 (pairwise Pearson's correlation, *r*) were calculated for each one of the six brain donors of the AHBA

660 separately, and then combined into one consensus matrix based on the element-wise mean across all
661 donors. Co-expression was converted to dissimilarity based on $1 - r$, in this way only positively co-
662 expressed genes are taken into account. All genes were hierarchically clustered using either single,
663 complete, average linkage and co-expression modules were obtained with the *cutreeDynamicTree*-
664 function in the WGCNA R-package version 1.64.1⁴³; minimum module size was set to 50 by default.
665 Hierarchical clustering by average linkage resulted in an acceptable number of missing genes while
666 retaining the maximum number of modules (Supplementary Figure 10; 167 modules with sizes up to 297
667 genes). For each module, the eigengene was obtained for each brain donor separately based on the first
668 principle component and thus summarizes the expression of all genes within a module across all samples
669 in Braak stage-related regions R1-R6. The sign of the eigengene expression was corrected based on the
670 sign of its correlation with the mean expression of all genes within the module. Similar to the BRGs, the
671 eigengene of each module was correlated with Braak labels for each donor separately and correlations
672 were combined across donors using meta-analysis.

673 **Gene set enrichment analysis of Braak stage-related modules.** The hypergeometric test was used to
674 identify modules that are significantly enriched for BRGs, cell-type markers²³, gene ontology- (GO), and
675 disease-associated genes from DisGeNET⁴⁶. A table of 561,119 gene-disease associations were
676 obtained from DisGeNET version 5.0 (May, 2017) from <http://www.disgenet.org/>. Genes associated with
677 GO-terms were obtained from the Ensembl dataset *hsapiens_gene_ensembl* version 92 through biomaRt
678 R-package version 2.38. All gene sets were filtered to contain only genes matching the 20,017 genes in
679 AHBA and at least 10 genes. Modules were significantly enriched when P -value < 0.05 (BH-corrected for
680 number of modules) using all 20,017 genes from AHBA as background genes.

681 **Data availability.** Data from healthy subjects used in this study are publicly available at brain-map.org,
682 braineac.org, and gtexportal.org. Microarray data from PD-, and iLBD patients, and controls were
683 collected and shared by Amsterdam University Medical Center, the Netherlands.

684 **Code availability.** Scripts were run in R version 3.5 and can be found online:
685 <https://github.com/arlinkeo/PD>. Scripts to analyze the PD microarray and PD RNA-seq dataset were run
686 in R version 3.4.