Mitochondrial complex I deficiency stratifies idiopathic Parkinson's disease

Irene H Flønes[1,2,3], Lilah Toker[1,2,3], Dagny Ann Sandnes[1,2], Martina Castelli[1], Sepideh Mostafavi[1,2], Njål Lura[4,5], Omnia Shadad[1,2], Erika Fernandez-Vizarra[6,7], Cèlia Painous[8], Alexandra Pérez-Soriano[8,9], Yaroslau Compta[8], Laura Molina-Porcel[10,11], Guido Alves[12,13], Ole-Bjørn Tysnes[1,2], Christian Dölle[1,2,3], Gonzalo S Nido[1,2,3] and Charalampos Tzoulis[1,2,3]*

1. Neuro-SysMed, Department of Neurology, Haukeland University Hospital, 5021 Bergen, Norway. 2. Department of Clinical Medicine, University of Bergen, Pb 7804, 5020 Bergen, Norway. 3. K.G. Jebsen Center for Translational Research in Parkinson's disease, University of Bergen, Pb 7804, 5020 Bergen, Norway. 4. Mohn Medical Imaging and Visualization Centre, Department of Radiology, Haukeland University Hospital, Bergen, Norway. 5. Section for Radiology, Department of Clinical Medicine, University of Bergen, Norway. 6. MRC-Mitochondrial Biology Unit, University of Cambridge, Hills Road, Cambridge, CB2 0XY, UK. 7. Veneto Institute of Molecular Medicine, 35131 Padova, Italy. 8. Parkinson's disease & Movement Disorders Unit, Neurology Service, Hospital Clínic I Universitari de Barcelona; IDIBAPS, CIBERNED (CB06/05/0018-ISCIII), ERN-RND, Institut Clínic de Neurociències (Maria de Maeztu excellence centre), Universitat de Barcelona. Barcelona, Catalonia. 9. UParkinson-Sinapsi Neurología, Centre Mèdic Teknon Grup Hospitalari Quirón Salud, Barcelona. 10. Alzheimer's disease and other cognitive disorders unit. Neurology Service, Hospital Clínic, Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS), 11. Neurological Tissue Bank, Biobanc-Hospital Clínic-IDIBAPS, Barcelona, Spain. 12. The Norwegian Centre for Movement Disorders and Department of Neurology, Stavanger University Hospital, Pb 8100, 4068 Stavanger, Norway. 13. Department of Mathematics and Natural Sciences, University of Stavanger, 4062 Stavanger, Norway.

***Correspondence to:**
Prof. Charalampos Tzoulis
Neuro-SysMed Center of Excellence for Clinical Research in Neurological Diseases
Department of Neurology, Haukeland University Hospital
Department of Clinical Medicine, University of Bergen
5021 Bergen, Norway.
Telephone: +47 55975061
E-mail: charalampos.tzoulis@nevro.uib.no

Word count main text (methods excluded): 5117

Running title: Complex I stratifies Parkinson's disease

Number of figures: 7

Number of tables: 2

Number of supplementary information: 13 figures, 13 tables

Keywords: neurodegeneration, neurodegenerative synucleinopathy, parkinsonism, subtypes, subgroups, mitochondrial PD, mitochondrial respiratory chain, mitochondrial DNA, gene expression.

Abstract

Idiopathic Parkinson's disease (iPD) is believed to have a heterogeneous pathophysiology, but molecular disease subtypes have not been identified. Here, we show that iPD can be stratified according to the severity of neuronal respiratory complex I (CI) deficiency, and identify two emerging disease subtypes with distinct molecular and clinical profiles. The CI deficient (CI-PD) subtype accounts for approximately a fourth of all cases, and is characterized by anatomically widespread neuronal CI deficiency, a distinct cell type-specific gene expression profile, increased load of neuronal mtDNA deletions, and a predilection for non-tremor dominant motor phenotypes. In contrast, the non-CI deficient (nCI-PD) subtype exhibits no evidence of mitochondrial impairment outside the dopaminergic substantia nigra and has a predilection for a tremor dominant phenotype. These findings constitute a step towards resolving the biological heterogeneity of iPD with implications for both mechanistic understanding and treatment strategies.

Introduction

Parkinson's disease (PD) is a heterogeneous, clinicopathologically defined neurodegenerative disorder, with an age-dependent incidence and a rapidly rising global prevalence[1]. More than 95% of affected individuals, in most populations, have idiopathic forms of PD (iPD), i.e., not of monogenic or other known etiology[1,2]. Available treatments for PD provide limited symptomatic relief, while attempts to develop neuroprotective agents have been consistently unsuccessful[3,4].

One of the major challenges in understanding and treating iPD lies in its biological heterogeneity. While all iPD cases exhibit degeneration of the dopaminergic neurons of the substantia nigra pars compacta (SNpc) in the presence of α-synuclein positive Lewy pathology (LP)[5], the disease shows significant clinicopathological variation beyond this common core. Widespread neuronal dysfunction and loss of variable anatomical distribution and severity gives rise to different constellations of motor and non-motor symptoms, with variable rates of progression, across affected individuals[1,4,6,7]. This phenotypical diversity has prompted the hypothesis that iPD may encompass multiple distinct disorders, each with its own underlying mechanisms and potential responses to treatments[4,6]. Previous efforts to identify subtypes of iPD have mainly focused on clinical features, with age of onset, and motor or cognitive phenotype being the most common. Several phenotypical subtypes have been proposed based on these classification features, including tremor dominant (TD), akinetic rigid (AR), postural instability and gait disorder (PIGD), and PD dementia (PDD)[8–10]. However, these classifications offer little insight into the pathogenic heterogeneity of iPD. Indeed, top-down approaches attempting to link consensus phenotypical subtypes to specific underlying molecular processes, have thus far been unsuccessful, indicating the need for alternative approaches[4]. It has been proposed that a bottom-up strategy, stratifying iPD according to

81 distinctive molecular features rather than clinical profiles, may prove more fruitful in defining
82 disease subtypes amenable to specific therapies[4].

83 Major biological processes which have been associated with iPD include mitochondrial
84 dysfunction, impaired proteostasis, and neuroinflammation[1,3,4]. Mitochondrial dysfunction is
85 considered a key feature of iPD, proposed to be integral to its pathogenesis and
86 pathophysiology[11]. Evidence of impaired mitochondrial DNA (mtDNA) maintenance, and
87 respiratory complex I (CI) deficiency has been consistently reported in the dopaminergic SNpc
88 of individuals with iPD[11–14]. While the causes of these mitochondrial defects are unknown, it is
89 hypothesized that they contribute to the neurodegenerative process, based on the facts that
90 chemical CI inhibition[15–17], genetic CI defects[18,19], and mutations in genes regulating mtDNA
91 homeostasis[20,21] can cause degeneration of the dopaminergic SNpc and other neuronal
92 populations in humans. In addition to the SNpc, CI deficiency has been reported in neurons
93 across multiple brain regions, as well as in extra-neural tissues, including muscle and blood[12,13].
94 However, reported findings are variable and contradictory, with nearly half the studies showing
95 no significant difference between individuals with PD and controls at the group level[13]. While
96 this variability may be in part methodological in nature, it also raises a valid question regarding
97 the pervasiveness of CI deficiency in iPD. In line with this notion, studies of CI activity in
98 platelets have shown a large inter-individual spread in PD, with some but not all individuals
99 below the range of controls [22,23] and studies in skin fibroblasts from individuals with iPD have
100 identified evidence of mitochondrial dysfunction in only a minority of the subjects[24,25].

101 In light of this evidence, we wished to explore whether mitochondrial dysfunction is a
102 universal feature of iPD, or one which occurs in a subpopulation of affected individuals,
103 potentially defining a molecular subtype of the disease. To test this hypothesis, we studied the
104 pervasiveness of neuronal CI deficiency in 92 individuals with genetically validated iPD and
105 24 neurologically healthy controls from two independent cohorts. We show that iPD can be

106 stratified according to the severity of neuronal respiratory complex I (CI) deficiency, and
107 identify two emerging disease subtypes with distinct molecular and clinical profiles.

108 ## Results

109 **CI deficiency shows specificity for iPD in the prefrontal cortex**

110 To select a robust marker for neuronal CI immunodetection, we tested antibodies against
111 different CI subunits across the functional domains of the complex[26] in the SNpc of 5
112 individuals with iPD and 3 controls. All tested subunits exhibited a similar pattern and severity
113 of neuronal CI deficiency, consistent with a quantitative reduction of the entire CI in affected
114 cells (Supplementary Fig. S1). Subunit NDUFS4 was chosen for downstream analyses due to
115 superior staining quality. Neuronal CI status was subsequently assessed, along with the total
116 mitochondrial mass marker VDAC1, in the dopaminergic SNpc and prefrontal cortex (PFC) of
117 two independent cohorts from Norway (NOR, n = 54) and Spain (ESP, n = 62). The
118 demographic and clinicopathological information of the cohorts, and experimental allocation
119 of the samples are shown in Table 1 and Supplementary Data 1 and 2.

120 Neurons were classified as either negative or positive (Fig. 1a-b) by two independent
121 investigators, and the percentage of positive neurons per region and individual was used as a
122 measure of CI status. There was excellent inter-investigator reliability (ICC coefficient = 0.91,
123 95% C.I (0.87,0.94), P = 1.9×10^{-29}) (Fig 1c-d, Supplementary Data 3). At the group level, the
124 SNpc exhibited pronounced CI deficiency in both iPD and control subjects, with considerable
125 overlap between the groups. iPD had a significantly lower proportion of CI positive
126 dopaminergic neurons in the combined cohorts (P = 0.003) and the NOR cohort (P = 8.2×10^{-4}),
127 but not in the ESP cohort (P = 0.26) (Fig. 1e, Table 2, Supplementary Data 3). In the PFC,
128 controls showed no or mild CI deficiency, whereas high variation was observed in the iPD
129 group. A significantly lower proportion of CI positive neurons among individuals with iPD was
130 observed in each cohort individually (NOR: P = 7.0×10^{-4}, ESP: P = 0.031) and in the combined
131 cohorts (P = 5.0×10^{-5}; Fig. 1f, Table 2, Supplementary Data 3). The neuronal CI status in the
132 PFC showed a significant association with disease state (i.e., iPD vs ctr), whereas that of the

133 SNpc did not (linear regression, PFC: P = 0.006, SNpc: P = 0.06). CI deficiency in the PFC was
134 not associated with post-mortem interval (PMI) (ρ(109) = -0.08, P = 0.42), age of death (ρ(110)
135 = 0.03, P = 0.73), or disease duration (ρ(79) = 0.06, P = 0.60) (Supplementary Data 1). The
136 VDAC1 staining showed no deficient neurons in iPD or controls of either cohort (Fig. 2a-b,
137 Supplementary Data 3).
138
139 **Neuronal CI deficiency stratifies iPD**
140 To assess the pervasiveness of CI deficiency in iPD, individuals were stratified based on the
141 proportion of CI positive neurons in the PFC. This region was chosen over the SNpc for two
142 reasons. First, neuronal CI deficiency was found to be more specific to PD in the PFC than in
143 the SNpc (Fig. 1e-f). This is in line with previous findings, showing that dopaminergic SNpc
144 neurons can exhibit severe CI deficiency and other markers of mitochondrial dysfunction with
145 healthy aging, making it challenging to distinguish between disease-specific and age-related
146 changes[27–29]. Second, the severe neuronal loss in the SNpc of iPD[30] complicates the comparison
147 between cases and controls. Specifically, surviving SNpc neurons are likely to be resilient cells,
148 introducing survival bias.
149 Individuals with iPD were classified by k-means clustering of the NOR and ESP cohorts
150 individually (cases and controls). The number of clusters was set to three based on visual
151 inspection of the data. In each cohort, two subgroups clustered away from the controls and were
152 termed mild and severe CI deficient PD (CI-PD mild; CI-PD severe), whereas one clustered
153 with the controls and was termed non-CI deficient PD (nCI-PD). The combined CI-PD clusters
154 accounted for 33% and 22% of the iPD cases, in the NOR and ESP cohorts, respectively. A
155 similar pattern was observed when the combined cohorts were clustered, resulting in three
156 clusters: CI-PD severe (*n* = 6/89, 6.7%), CI-PD mild (*n* = 18/89, 20.2%), and nCI-PD clustering
157 with the controls (*n* = 65/89, 73.0%) (Fig 1g, Supplementary Data 3). Since the severe CI-PD

158 cluster only contained six samples, the mild and severe clusters were combined into a single
159 CI-PD group for downstream statistical analyses.
160
161 **CI-PD is characterized by anatomically widespread CI deficiency**
162 To determine the anatomical distribution of neuronal CI deficiency in the iPD subgroups, we
163 studied 15 additional brain regions in representative individuals with CI-PD, nCI-PD and
164 controls from the NOR cohort. All individuals with iPD had Braak scores of 5-6 and the median
165 disease duration was 12 years in both subgroups (Supplementary Data 1 and 2). Subjects in the
166 CI-PD group showed an anatomically widespread neuronal CI-deficiency compared to controls,
167 including the temporal cortex, occipital cortex, CA1 region of the hippocampus, amygdala,
168 putamen, and the cingulate gyrus. Lower proportions of CI intact neurons were also seen in the
169 internal globus pallidus and the subiculum, but these did not reach statistical significance. No
170 difference from controls was observed in CA4-2 of the hippocampus, the external globus
171 pallidus, inferior olivary nucleus, cerebellar Purkinje cells and the dentate nucleus. The nCI-
172 PD group was similar to controls, with the single exception of the SNpc. The difference in
173 SNpc was, however, less pronounced compared to the CI-PD group. Finally, we noted that,
174 similar to the SNpc, the inferior olivary nucleus, Purkinje neurons and dentate nucleus showed
175 high proportions of CI negative neurons in several of the controls, suggesting that these
176 neuronal populations may be prone to age-dependent CI loss (Fig. 2, Fig. 3, Supplementary
177 Data 3 and 4).
178 To assess the state of the other mitochondrial respiratory chain (MRC) complexes in the
179 iPD groups, we performed IHC for selected subunits of complexes II-V (CII-V) and VDAC1
180 in 12 brain regions from two to four subjects from each of the CI-PD and nCI-PD groups
181 (Supplementary Table S2). Irrespective of group, individuals with iPD exhibited no or few
182 neurons staining negative for CIII, CIV and CV (generally ≤5%). No negative neurons were

183 observed in sections stained against VDAC1 or CII (Supplementary Data 3, Supplementary
184 Fig. S2).
185
186 **CI-PD is associated with a non-TD phenotype**
187 The CI-PD and nCI-PD groups were similar in terms of age of first reported symptom, age of
188 death, disease duration, and total UPDRS score at the final examination before death.
189 Individuals with severe CI-PD tended to have an earlier age of onset and age of death, but the
190 small size of that group did not allow any conclusions to be drawn (Fig. 4a-c, Supplementary
191 Data 5). The groups differed significantly in the distribution of motor phenotype classified as
192 AR/TD[9] (available from both cohorts), or PIGD/TD[10] (available from the NOR cohort only), at
193 the time of diagnosis. In the CI-PD group, the AR and PIGD phenotypes accounted for 88%
194 (14/16) and 75% (9/12) of the individuals, respectively. In the nCI-PD group, the AR or PIGD
195 phenotypes were significantly less prevalent, accounting for 60% (25/42) and 29% (6/21) of
196 the individuals, respectively. Notably, individuals with severe CI-PD exhibited exclusively AR
197 or PIGD phenotype (Fig. 4h-i). Furthermore, the CI-PD group exhibited a significantly higher
198 proportion of females (CI-PD: 54.2%, nCI-PD: 26.2%, P = 0.013), compared to the nCI-PD
199 group, and a significantly lower level of ethanol consumption than the nCI-PD group (Fig. 4e-
200 f). There was no significant difference between the groups in the prevalence of dementia at the
201 final examination before death (Fig. 4g). Following multiple testing correction (Benjamini
202 Hochberg), the difference in the distribution of the PIGD/TD phenotype (P = 0.04), sex (P =
203 0.04), and number of weekly alcoholic units (P = 0.04) remained significant. However, since
204 the tested variables were highly correlated with each other and the sample size was relatively
205 small, the risk of type 2 error by correcting for multiplicity, is higher than the risk of type 1
206 error by not correcting[31]. Individual clinical and demographic features are given in
207 Supplementary Data 1 and 5.

208 In terms of neuropathology, there were no significant between-group differences in either Thal
209 phase staging of beta amyloid, Braak staging of neurofibrillary tangles, or Braak staging of LP.
210 Moreover, there was no difference in the load of LP in the PFC (available from the NOR cohort
211 only) between the groups (Supplementary Fig. S3).
212
213 **CI-PD and nCI-PD show different neuronal mtDNA profiles**
214 Neuronal mtDNA copy number and deletion burden were assessed by qPCR in single neurons
215 microdissected from the PFC (Fig. 5a-b) of individuals with CI-PD (n = 6), nCI-PD (n = 9) and
216 controls (n = 7; Supplementary Data 2). Eight to sixteen single neurons were assessed per
217 individual. There was a significant difference in neuronal mtDNA deletions between the groups
218 (ANOVA, P = 0.037, η^2 = 0.30). Post-hoc analyses showed that the CI-PD group had
219 significantly higher levels of mtDNA deletions compared to the nCI-PD group (P = 0.037; Fig.
220 5b). No differences were observed in neuronal mtDNA copy number (total or non-deleted)
221 between the groups (Fig. 5c-d). The mtDNA data is shown in Supplementary Data 6.
222
223 **CI-PD and nCI-PD exhibit distinct transcriptomic profiles**
224 Hierarchical clustering of RNA-seq data from the PFC of 98 of the individuals previously
225 assessed for CI deficiency (Supplementary Data 2) indicated that individuals with CI-PD tended
226 to cluster together (Fig. 6a), with a similar behavior observed in each of the cohorts
227 (Supplementary Fig. S4). The CI-PD group was associated with lower DV200 values, a
228 measure of RNA integrity indicating the percentage of RNA fragments with 200 nucleotides or
229 more, compared to both nCI-PD and controls (Fig. 6a). However, this association could not be
230 explained by PMI, since the PMI range of the CI-PD group was similar to that of the controls
231 within each cohort (Supplementary Fig. S5).

232 To control for between-group differences in the cell composition of the samples, we
233 assessed Marker Gene Profiles (MGPs)[32] of different cell types in each iPD group and the
234 controls. Linear regression adjusting for age of death, sex, PMI and cohort, indicated a
235 significant change in the MGP of multiple cell types in the CI-PD group compared to the
236 controls. However, step-wise regression analysis suggested that the observed changes were not
237 independent, but rather driven by a decrease in the MGP of GABAergic interneurons expressing
238 both vasoactive intestinal peptide-expressing (VIP) and reelin (GabaVIPReln) (Fig. 6b,
239 Supplementary Fig. S6). No difference was observed in the MGPs of the nCI-PD group
240 compared to the controls (Fig. 6b, Supplementary Fig. S6).

241 To account for the effects of different cell composition on the differential gene
242 expression (DE) profile of CI-PD, we performed DE analysis using two alternative models: 1)
243 Correcting for the confounding MGP of GabaVIPReln (Model_1) and 2) not correcting for
244 MGPs (Model_2). Both models included the demographic variables and oligodendrocyte
245 MGPs to adjust for any variation in the grey/white matter content between the samples[33,34].
246 Adjusting for GabaVIPReln MGPs resulted in a striking reduction of the number of
247 differentially expressed genes in the CI-PD group, from 11,138 to 10 (Supplementary Data 7).
248 Notably, oxidative phosphorylation (OXPHOS)-related genes were among the top differentially
249 downregulated genes in CI-PD, in the non-adjusted model. However, none of these genes
250 remained significant after adjusting for cell composition (Fig 6c). In contrast, MGP correction
251 had no major impact on the DE analysis of the nCI-PD group (Fig 6c).

252

253 **nCI-PD exhibits altered expression in multiple PD-linked pathways but no decrease in**
254 **OXPHOS genes**

255 We next performed functional enrichment analyses of the DE results. Since the DE analysis of
256 the CI-PD group was highly confounded by the underlying tissue composition, we were unable

257 to determine whether the results represented true regulatory changes or simply altered cell
258 composition. We therefore chose not to further focus on these results, which are given in
259 Supplementary Data 8. The transcriptomic signature of nCI-PD, which was not confounded by
260 estimated cell composition, was characterized by changes in numerous pathways, including the
261 downregulation of processes related to proteostasis (lysosome and proteasome), GABAergic
262 signaling, synaptic function and lipid metabolism. Upregulated processes included pathways
263 related to epigenomic regulation, steroid synthesis, and potassium channels. Notably, there
264 were no changes in genes or pathways related to mitochondrial respiration (Fig. 6d,
265 Supplementary Fig. S7 and S8, Supplementary Data 8).
266
267 **Single-nuclei RNA-seq shows distinct cell type-specific profiles in the iPD subtypes**
268 To mitigate the bias of cell composition indicated by the bulk RNA-seq data and assess cell
269 type-specific transcriptomic similarities between CI-PD and nCI-PD, we performed massively
270 parallel single-nuclei RNA-seq (snRNA-seq) in fresh-frozen PFC from n = 7 individuals with
271 CI-PD, n = 5 individuals with nCI-PD, and n = 6 controls (Supplementary Data 2). After
272 filtering and quality control, a total of 117,105 nuclei (mean: 6,506±2,952 nuclei per individual)
273 were included in the analyses. There was no significant difference in the number of captured
274 nuclei between groups (ANOVA: F =3.65, degrees of freedom = 2, P = 0.051).
275 Default clustering using Seurat identified a total of 28 clusters that could be
276 unequivocally assigned to a major cortical cell type, based on established cell type markers (12
277 excitatory neuron clusters, 8 inhibitory neuron clusters, 3 oligodendrocyte clusters, 2 astrocyte
278 clusters, 1 microglia cluster, 1 oligodendrocyte precursor cells (OPC) cluster, and 1 endothelial
279 cluster, Fig. 7a-b). Unique molecular identifier (UMI) counts and number of identified genes
280 per nucleus were not significantly different between groups in any of the identified cell type
281 clusters.

282 Differential gene expression analysis comparing each of the iPD groups to the controls
283 revealed distinct cell type-specific transcriptomic profiles between CI-PD and nCI-PD. After
284 FDR correction, we found a total of 148 significantly differentially expressed genes in CI-PD
285 and 466 significantly differentially expressed genes in nCI-PD, distributed among several cell
286 type clusters, with 47 genes overlapping between the groups (Fig. 7c). CI-PD showed
287 differential expression in most excitatory and inhibitory neuronal subtypes, and a notable
288 upregulation in microglia, including *PRKN* – encoding parkin, a key component of mitophagy
289 – in which mutations cause autosomal recessive PD[35]. nCI-PD exhibited multiple differentially
290 expressed genes, with most of the signal observed in neurons and oligodendrocytes, while no
291 differential expression was seen in microglia. The differentially expressed genes per cell type
292 and group are shown in Supplementary Data 9 and 10. A visual map of significantly
293 differentially expressed genes (FDR < 5%) with fold change ≥ 25%, in at least one cell type, is
294 shown in Supplementary Fig. S10.
295 To directly identify features differentiating CI-PD and nCI-PD, we performed DE
296 between the two groups. In this contrast, we identified a total of 909 genes differentially
297 expressed in at least one cell type, suggesting that differences between the iPD subtypes were
298 larger than between each subtype and the controls. We, therefore, focused further analyses on
299 this contrast. Significant differential expression was found across all cell types, neuronal and
300 glia, with the highest number of differentially expressed genes seen in excitatory neuron
301 clusters ex_09 and ex_19. According to the expression of known neuronal markers[36], ex_09 are
302 located in cortical layers L4c/L6, and express relatively high levels of *OPRK*, encoding kappa
303 opioid receptors. These neurons are, therefore, potentially involved in nociception. Neurons
304 classified as ex_19 are located in layers (L6/L6b), and have a high relative expression
305 of *ADRA2A*, encoding the alpha 2A adrenoreceptor, consistent with adrenergic afferentation.

306 Most differentially expressed genes were specific to 1-2 cell types, with few notable
307 exceptions, including *SRSF5*, *MAK*, and *RAB11A* which were downregulated in CI-PD across
308 several excitatory and inhibitory neurons, and/or glia. A focused analysis of the nuclear genes
309 encoding MRC proteins revealed a general trend for downregulation in CI-PD, compared to
310 nCI-PD, including significantly lower expression of multiple subunits of complexes I, III, IV,
311 and V. Similar to the overall differential expression signature, these changes were
312 predominantly seen in neuronal clusters ex_09 and ex_19 (Supplementary Fig. S11).
313 Transcriptional differences in MRC genes between either iPD subtype and controls did not
314 reach statistical significance. Nevertheless, there was a notable trend for upregulation in nCI-
315 PD (Supplementary Figs. S12 and S13).
316 Functional enrichment of the differentially expressed genes was carried out
317 independently for each cell type cluster by overrepresentation analysis of the Kyoto
318 Encyclopedia of Genes and Genomes (KEGG) pathways. Most cell types showed no significant
319 enrichment, with the exception of ex_19 which showed significant enrichment for pathways
320 relating to OXPHOS, the ribosome, and Parkinson's and Huntington's disease (the latter two
321 also driven by OXPHOS genes). Additionally, the ribosome pathway was enriched in clusters
322 ex_09 and ex_24 (Supplementary Data 12). Finally, we performed overrepresentation analysis
323 on the union of significantly differentially (FDR = 0.05) expressed genes in excitatory neurons,
324 inhibitory neurons, and glia. This analysis showed a highly significant enrichment in OXPHOS-
325 and ribosome related pathways in excitatory neurons (Supplementary Data 13).
326

327 **Discussion**

328 Identifying and exploring molecular subtypes of iPD is a critical step towards mechanistic and
329 therapeutic advancements. We show that iPD can be stratified according to the severity of
330 neuronal CI deficiency and identify two emerging disease subtypes with distinct molecular and
331 clinical features, indicating a divergent pathophysiology. CI-PD is characterized by widespread
332 neuronal CI deficiency across multiple brain regions. In contrast, nCI-PD, exhibits a neuronal
333 CI profile similar to that of demographically matched controls, with the single exception of the
334 dopaminergic SNpc, where higher levels of CI deficiency are observed, albeit to a lesser extent
335 than in CI-PD. Importantly, this difference is not driven by disease duration or age, indicating
336 that it does not reflect two different stages of the same process. The existence of the CI-PD and
337 nCI-PD subtypes is shown independently in two different populations, suggesting that this is
338 generalizable to- and representative of iPD.

339 It is likely that the observed CI deficiency contributes to neuronal dysfunction and death
340 in CI-PD. Neurons depend on CI function to maintain a high ATP turnover[37]. Brain
341 mitochondria exhibit impaired energy metabolism with major decreases in ATP production
342 upon 60% reduction of CI activity, and this threshold is much lower (~25%) for synaptic
343 mitochondria[38]. While functional assessment of CI activity is currently not feasible in
344 individual neurons from postmortem brain, it is reasonable to assume that neurons with no
345 detectable CI staining (i.e., CI negative neurons) in our samples must have substantially
346 decreased CI levels and, therefore, a compromised bioenergetic status. This is further
347 corroborated by the fact that isolated CI deficiency of either chemical[15,17] or genetic[18,39] etiology
348 is an established cause of neuronal dysfunction and death in humans and other animals.

349 With the single exception of mild deficiency in the dopaminergic SNpc, the nCI-PD
350 group exhibited a neuronal CI profile similar to that of age-matched controls. Thus, we
351 postulate that CI deficiency does not make an important contribution to the pathogenesis of PD-

352 specific neurodegeneration outside the SNpc in these individuals. CI deficiency may still
353 contribute to the degeneration of the dopaminergic SNpc neurons in nCI-PD. However, given
354 that this deficiency was significantly milder compared to CI-PD, and showed substantial
355 overlap with the aged-matched controls, its disease-specific pathophysiological significance
356 remains uncertain.

357 The CI-PD and nCI-PD groups differed in terms of motor disease subtype at the time of
358 diagnosis. Our data indicate that CI dysfunction is significantly more common in the non-TD
359 form of iPD, which has also been shown to harbor more severe neurodegeneration[40]. The
360 mechanism underlying this association is currently unknown but may be related to more
361 widespread and severe neuronal dysfunction and loss in CI-PD, induced by the CI deficiency.
362 This is corroborated by the bulk RNA-seq data indicating more pronounced neuronal loss in
363 CI-PD. Until clinically applicable biomarkers of CI-PD and nCI-PD can be developed,
364 enabling us to classify individuals in vivo, perhaps clinical studies could benefit by
365 distinguishing between TD and non-TD phenotypes. Since individuals with a non-TD
366 phenotype at the time of diagnosis are substantially more likely to belong to the CI-PD subtype,
367 they may be more likely to benefit from experimental treatments targeting mitochondria. While
368 there are currently no drugs specifically targeting CI, several agents targeting mitochondrial
369 biogenesis and/or function have been tested in clinical trials with negative results[3], while others
370 are still in the pipeline[41]. In light of our findings, it would be of interest to revisit the data of
371 previously conducted trials and assess how the treatment impacted individuals with a non-TD
372 or TD phenotype. In addition to the motor phenotype, we observed an association of CI-PD
373 with the female sex and a lower consumption of ethanol. The latter may be driven by the female
374 preponderance in this subtype.

375 Neuronal mtDNA profiles in the PFC corroborated the stratification of iPD into CI-PD
376 and nCI-PD groups. mtDNA deletions are an established cause of MRC deficiencies in post-

377 mitotic cells[42–44]. However, given the small effect size of the mtDNA deletion increase in the
378 CI-PD group, and the fact that neuronal deletion levels were generally below 60%, i.e., the
379 threshold at which MRC dysfunction has been proposed to occur[42], the biological significance
380 of the detected difference in mtDNA deletion levels remains uncertain. It should be noted that,
381 since most PFC neurons of individuals with CI-PD remain CI positive, it is possible that few or
382 none of the captured neurons in our experiments were CI negative. Thus, the possibility that
383 CI-negative neurons harbor higher deletion levels, driving the CI deficiency, cannot be ruled
384 out by this work. An alternative explanation is that mtDNA deletions in CI-PD individuals may
385 be the result, rather than the cause, of CI deficiency. Under this hypothesis, mtDNA deletions
386 may arise during the repair of double-strand breaks, caused by increased reactive oxygen
387 species (ROS) generation in CI deficient neurons[45,46].

388 RNA-seq revealed distinct transcriptomic profiles between the CI-PD and nCI-PD
389 groups. In bulk tissue data, the transcriptomic signature of CI-PD exhibited downregulation of
390 OXPHOS-related genes but was also severely confounded with cell-type-specific expression
391 profiles (i.e., MGPs), indicating differences in the cell type composition of the samples. The
392 observation that CI-PD exhibited decreased estimates for GABAergic interneurons is consistent
393 with a mitochondrial defect. GABAergic interneurons have high bioenergetic demands, due to
394 their fast-spiking firing pattern, and have been shown to be selectively lost in mitochondrial
395 disease characterized by neuronal CI deficiency disease [47]. While MGPs are not direct cell
396 counts, they are indicative of a simultaneous shift in multiple cell-type specific genes.
397 Regardless of whether this reflects a difference in the underlying cellular composition of the
398 tissue, synaptic loss, decreased transcriptional output, or phenotypic change[39], this association
399 indicates that any regulatory changes cannot be disentangled from differences due to altered
400 tissue composition[33,48]. Thus, no further conclusions should be drawn from transcriptomic
401 analyses of bulk brain tissue data from CI-PD individuals. In contrast, the nCI-PD group

402 exhibited no difference in MGPs compared to controls. Thus, we believe that the DE signature
403 of this group reflects regulatory changes associated with the disease. The transcriptional profile
404 of nCI-PD showed altered expression in multiple pathways, many of which have been
405 previously linked to PD[49–51].

406 snRNA-seq allowed us to compare the iPD subgroups within each cell-type, thus
407 circumventing the confounder of different cell composition. These analyses revealed a highly
408 distinct cell type-specific transcriptomic profile for each of CI-PD and nCI-PD, supporting that
409 the two subtypes harbor distinct molecular processes. While single-nucleus transcriptomics are
410 not sufficient to adequately characterize these processes, since they only capture a small part of
411 the cell's transcriptional state, the data provides us with some hints. Neurons and astrocytes
412 were affected in both groups, albeit by largely different genes. However, microglia and
413 oligodendrocytes showed subgroup-specific involvement. The CI-PD-specific changes in
414 microglia are suggestive of mitophagy activation, which indicates the presence of
415 mitochondrial dysfunction and stress in these cells[52]. In contrast, nCI-PD showed no changes
416 in microglia, but harbored substantial transcriptional alterations in oligodendrocytes.
417 Oligodendrocyte involvement has been previously reported in PD[53]. Our findings suggest that
418 this may be a feature of nCI-PD.

419 Direct comparison of the CI-PD and nCI-PD groups further highlighted their distinct
420 molecular profiles, as it revealed substantially more differentially expressed genes than in each
421 of the subtypes compared to controls. At the single-nucleus level, CI-PD was characterized by
422 a significantly lower expression of MRC-encoding genes compared to nCI-PD. This difference
423 was largely driven by an upregulation trend in nCI-PD, compared to controls. One possible
424 interpretation of this observation is that, in the majority of individuals with PD (i.e., nCI-PD),
425 neurons exhibit an upregulation of mitochondrial respiration and/or biogenesis, potentially
426 indicative of a compensatory response to disease-related stress. Conversely, individuals with

427 CI-PD may represent cases where this compensatory response fails, possibly due to underlying
428 mitochondrial pathology.

429 Our findings must be interpreted considering certain limitations. Although we defined
430 subgroups of iPD, the exact threshold between them remains arbitrary, and the severity of
431 neuronal CI deficiency across individuals can also be seen as a continuum. Nevertheless, this
432 does not alter the fact that only a subset of individuals with iPD show neuronal CI deficiency
433 outside the SNpc, and this subpopulation exhibits distinct clinical and molecular features.

434 Functional assessment of specific CI activity was not undertaken in this work, because
435 it is currently not feasible to assess in individual cells. Measurement of CI activity in bulk tissue
436 would be unreliable because of both low sensitivity due to the mosaic distribution of the CI
437 defect, and low specificity in the sense that it cannot adequately distinguish between cell-
438 specific reduction in CI activity or a decrease in neuronal and synaptic content in the sample.
439 We have previously shown that the PFC of iPD exhibits markers of decreased neuronal
440 content[33], and our current data suggests that these differences may be more pronounced in CI-
441 PD, thereby biasing any comparison in bulk tissue. In terms of the gene expression data, the
442 cell-type composition bias is mitigated by the snRNA-seq. However, when interpreting this
443 data, it is important to consider that nuclear transcriptomes do not reflect post-transcriptional
444 RNA turnover and subcellular segregation and do not, therefore, necessarily capture the full
445 transcriptomic state of the cell. This is particularly true in neurons, where RNA composition
446 varies considerably between the soma, processes, and synapses. Furthermore, snRNA-seq
447 cannot capture the mitochondrial transcriptome.

448 In summary, this work establishes that anatomically widespread neuronal CI deficiency
449 is not a pervasive feature of iPD, but one that occurs in a subset of individuals, accounting for
450 approximately a fourth of all cases. These results have important scientific and clinical
451 implications for the field. First, the existence of the CI-PD and nCI-PD subtypes provides an

452 explanation for the variable and conflicting findings in previous studies of mitochondrial
453 pathology in iPD, which were performed on unstratified samples. Second, given the major
454 molecular differences between the two iPD subtypes, we recommend that future studies
455 employing iPD brain tissue should stratify their material according to CI subtype. This would
456 decrease the biological heterogeneity of the sample and allow for more consistent and
457 biologically meaningful observations to be made. Finally, it is plausible that the CI-PD and
458 nCI-PD subtypes exhibit different susceptibility to therapeutic interventions targeting
459 mitochondria or other subtype-associated pathways. Thus, we believe that harnessing this
460 classification in clinical practice may be an essential step towards implementing precision
461 medicine in iPD. While our findings suggest that the non-TD motor phenotype may be used as
462 a proxy for CI-PD to some extent, the accuracy of this approach is limited. To efficiently
463 translate these findings into clinical practice requires the development of clinically applicable
464 biomarkers, allowing us to classify individuals in vivo.
465
466
467

468 # Methods

469 **Ethics approval**

470 Ethics approval for this work was obtained from The Regional Committee for Medical and
471 Health Research Ethics, Western Norway (REK Vest; REK 2017/2082, 2010/1700,
472 2016/1592).

473

474 **Subject cohorts and tissue samples**

475 Brain tissue was obtained from four independent cohorts of European origin. The Norwegian
476 cohort (NOR, *n* = 55) comprised 41 individuals with iPD from the ParkWest study, a
477 prospective population-based cohort followed since 2004[54], and 13 neurologically healthy
478 controls (i.e., no clinical diagnosis of neurological disease or pathological evidence of α-
479 synucleinopathy), who had been prospectively collected by our group. The NOR samples were
480 stored at our in-house brain bank. The Spanish cohort (ESP, *n* = 62) comprised 51 individuals
481 with iPD and 11 neurologically healthy controls from the Neurological Tissue Bank, Hospital
482 Clínic – IDIBAPS Biobank, Barcelona, Spain. For the snRNA analyses, two additional controls
483 were added from the Medical Research Council London Neurodegenerative Disease Brain
484 Bank, King's college London, Great Britain, and one additional control from the Netherlands
485 Brain Bank, Amsterdam, The Netherlands.

486 Analysis of whole-exome sequencing data from the NOR cohort [55] and RNA-sequencing
487 (RNA-seq) data from both cohorts (described below), revealed no pathogenic mutations in
488 nuclear genes linked to monogenic PD or mitochondrial diseases, thus ruling out monogenic
489 causes. Prospectively collected systematic clinical and environmental exposure data were
490 available for the iPD cases of the NOR cohort, as previously described[54]. In the NOR cohort,
491 probable PDD was diagnosed according to the Movement Disorder Society Level 1 criteria[8].
492 Cognitive deficiency severe enough to impair daily life was defined based on the Unified

493 Parkinson's Disease Rating Scale (UPDRS) score 1-1[56]. Impairment in cognitive domains was
494 determined based on tests of attention ("serial 7's" test), visuo-constructive ability (correct
495 drawing of the pentagons), and memory impairment (3-word recall), from the Mini-Mental
496 State Examination (MMSE)[57]. Appropriate tests on executive function were not available, thus
497 the diagnosis of PDD in the NOR cohort was based on data from three, rather than four
498 cognitive domains, as required by the MDS Level 1 criteria[8]. Two individuals had impairment
499 in only one cognitive domain in addition to fulfilling the remaining criteria for PDD. These two
500 individuals may thus have been potentially misclassified and as non-PDD cases. Individuals
501 were classified by motor phenotype as TD or AR according to the Schiess classification[9], and
502 TD or PIGD according to the Jankovic classification[10]. Detailed retrospective clinical data were
503 available on the controls of the NOR cohort from the hospital records. Retrospective clinical
504 data on disease duration, dementia (yes/no) and PD phenotype (TD/AR/mixed) were available
505 from the hospital records of individuals of the ESP cohort.

506 There was no significant difference between controls and iPD subjects in PMI or
507 assigned sex distribution in either of the cohorts, or in the combined cohort. Both sexes were
508 included in the study as available from our cohort. Age of death was similar between controls
509 and iPD subjects in the NOR cohort, while the control subjects were older in the ESP cohort
510 and in the combined cohort. Subsets of samples used in downstream analyses were matched for
511 age of death, sex and PMI-range, with the exception of the transcriptomic analyses, where the
512 controls were older as reflected by the age distribution in the ESP and combined cohorts. The
513 demographic data of the cohorts and descriptive statistics of each cohort included in the IHC
514 experiments are shown in Table 1.

515 Brains were collected at autopsy and processed according to either in-house[27] or
516 Neurological Tissue Bank protocol[58]. For the neuropathological evaluation, Lewy and tau
517 pathology were staged according to the Braak criteria[59–62], and β-amyloid phases were

518 evaluated according to the Thal criteria[62,63]. All cases fulfilled the essential pathological criteria
519 for PD: loss of the dopaminergic neurons of the SNpc in the presence of LP[64]. Subject
520 demographics, pathology scores and clinical data are summarized in Supplementary Data 1.
521 Experimental allocation of subjects is summarized in Supplementary Data 2.
522
523 **Immunohistochemistry**
524 Immunohistochemistry (IHC) for mitochondrial markers was carried out on formalin-fixed,
525 paraffin-embedded brain tissue from individuals of both the NOR and ESP cohorts. Serial 3 µm
526 thick sections were stained with primary antibodies against the following epitopes; NDUFS4
527 (ab137064, polyclonal, dilution 1:1500, Abcam), NDUFB8 (ab110242, monoclonal, isotype
528 IgG1, dilution 1:300, Abcam, validated in frozen IHC and tested in human samples), NDUFA9
529 (ab14713, monoclonal, isotype IgG1, dilution 1:750, Abcam), NDUFS1 (ab169540, dilution
530 1:300, Abcam), SDHA (ab14715, monoclonal, isotype IgG1, dilution 1:10,000, Abcam),
531 UQCRC2 (ab14745, monoclonal, isotype IgG1, dilution 1:10,000, Abcam), MTCOI (459600,
532 monoclonal, isotype IgG2a, dilution 1:15,000, Thermo Fisher Scientific, validated by Cell
533 treatment), ATP5A (ab14748, dilution 1:20,000, Abcam), and VDAC1 (ab14734, monoclonal,
534 isotype IgG2b, dilution 1:20,000, Abcam). Antibodies against NDUFS4, NDUFA9, NDUFS1,
535 SDHA, UQCRC2, ATP5a and VDAC1 were all validated in IHC in paraffin embedded tissues
536 and tested in human samples according to the supplier (Abcam).
537 Sections were stained in the intelliPATH FLX Automated Slide Stainer (Biocare).
538 Deparaffinization and antigen retrieval was performed in the low pH EnVision FLEX Target
539 Retrieval Solution at 98 °C for 24 minutes in the DAKO PT link from Agilent. Antibodies were
540 diluted in DaVinci Green Diluent (Biocare, PD900) and incubated for 1 hour at room
541 temperature. MACH4 Universal HRP-polymer (Biocare, M4U534) and DAB chromogen kit
542 (Biocare DB801) were used for visualization. Tacha's Automated Hematoxylin (Biocare, NM-

543 HEM) was used for visualization of nuclei. Sections were washed in the TBS Automation Wash
544 buffer from Biocare (TWB945M). Experimental allocation is given in Supplementary Data 2.
545
546 **Image analysis**
547 All sections were scanned using the NanoZoomer XR (Hamamatsu). Visual analysis was
548 performed on all sections, and neurons with clearly visible nucleus and nucleolus were
549 classified as either negative or positive (Fig. 1a). After classification was performed, the total
550 percentage of positive neurons was calculated for each individual and area. From the prefrontal
551 cortex (PFC) in the NOR cohort, three regions of 3.5 mm^2 spanning cortical layers 2-6 were
552 chosen at a low magnification (1.5X), before analyses were carried out in the selected area at
553 40X magnification. Area selection from one region spanning 3.5 mm^2 was performed in a
554 similar manner in the PFC from the ESP cohort, and in the occipital and temporal cortex, the
555 cingulate gyrus, subiculum, putamen and the amygdala. All neurons per section were analyzed
556 in the CA-areas of the hippocampus (CA4-1), globus pallidus, dentate nucleus, inferior olivary
557 nucleus, and Purkinje cell layer of the cerebellar cortex. In the SNpc, all neuromelanin
558 containing neurons were analyzed per section. SNpc sections with fewer than 25 identifiable
559 neuromelanin containing neurons were excluded from downstream analyses of neuronal CI
560 status because data resulting from these sections would entail high levels of statistical
561 uncertainty, at the range of the expected group differences (i.e., any misclassified neuron as CI
562 positive or negative would result in at least 4% (1/25) difference in the estimated proportion of
563 positive neurons.) An uncertainty greater or equal to 4% was deemed unacceptable since it is
564 close to the difference between the medians of the groups being compared in our analyses.
565 Independent evaluation of the PFC from the NOR cohort was performed by two observers (IHF
566 and CT). All analyses were carried out at 40X magnification. Visual analysis was carried out
567 using the NDP.view2plus version(v)2.7.25 software (Hamamatsu).

568

569 **Single neuron laser-microdissection studies**

570 Qualitative and quantitative mtDNA studies were performed in single laser-microdissected
571 neurons from 20 µm thick sections of fresh-frozen tissue from the PFC, which had been stained
572 with 0.25% cresyl violet in ddH_2O as described[28]. Neurons were recognized based on their
573 morphology. Laser microdissection was performed on a LMD7 microscope (Leica
574 microsystems). A total of 162 cells (8-16 from each individual) were cut, of which 55 were
575 from controls (n = 7) and 107 were from individuals with iPD (n = 14). Experimental allocation
576 is shown in Supplementary Data 2. Single neurons were collected in the cap of 0.2 ml PCR
577 tubes and lysed at 56°C overnight in 15 μL lysis buffer (50mM Tris pH 8.0, 0.5% Tween
578 20, 200 μg/mL proteinase K). Following lysis, cells were centrifuged at 2250 g for 10 minutes
579 before proteinase K was deactivated at 95°C for 10 minutes, and the samples were centrifuged
580 again. Total mtDNA copy number and the fraction of major arc deletion were determined using
581 a duplex real-time PCR assay which targeted the mitochondrial genes *MTND1* and *MTND4* as
582 described[28]. Amplification was performed on a Step One Plus Real-Time PCR System (Applied
583 Biosystems by Thermo Fisher Scientific) using TaqMan Fast Advanced Master Mix (Thermo
584 Fisher Scientific). Thermal cycling consisted of one cycle at 95°C for 20 seconds, and 40 cycles
585 at 95°C for 1 second and 60°C for 20 seconds. The following primers and probes were used:
586 MTND1: forward primer: 5'CCCTAAAACCCGCCACATCT-3', reverse primer: 5'-
587 GAGCGATGGTGAGAGCTAAGGT-3', TaqMan MGB Probe: 5'-FAM-
588 CCATCACCCTCTACATCACCGCCC-3'. MTND4: forward primer: 5'-
589 CCATTCTCCTCCTATCCCTCAAC-3', reverse primer: 5'-
590 CACAATCTGATGTTTTGGTTAAACTATATTT-3', TaqMan MGB probe: 5'-VIC-
591 CCGACATCATTACCGGGTTTTCCTCTTG-3'. For absolute quantification, target
592 amplification was compared with a standard curve made form a serial dilution containing equal

593 amounts of PCR-generated and purified full length *MTND1* and *MTND4* template (10^2-10^6).
594 The following primers were used for the standard curve; *MTND1*: forward primer: 5′-
595 CAGCCGCTATTAAAGGTTCG- 3', reverse primer: 5′- AGAGTGCGTCATATGTTGTTC-
596 3′; *MTND4*: forward primer: 5′- TCCTTGTACTATCCCTATGAG-3′, reverse primer: 5′-
597 GTGGCTCAGTGTCAGTTCG-3′. All primers and probes were custom made and purchased
598 from ThermoFisher Scientific.
599
600 **RNA sequencing**
601 RNA-seq was performed on frozen PFC tissue samples from 98 individuals from both cohorts
602 (Supplementary Data 2). Total RNA was extracted from tissue homogenates using RNeasy plus
603 mini kit (Qiagen) with on-column DNase treatment according to manufacturer's protocol. Final
604 elution was made in 65 µl of ddH2O. The concentration and integrity of the total RNA were
605 estimated by Ribogreen assay (Thermo Fisher Scientific), and Fragment Analyzer (Advanced
606 Analytical), respectively, and 500 ng of total RNA was used for library prep using the TruSeq™
607 Stranded Total RNA With Ribo-Zero™ Plus rRNA Depletion protocol (Illumina). Library
608 quantity was assessed by Picogreen Assay (Thermo Fisher Scientific), and the library quality
609 was estimated by utilizing a DNA High Sense chip on a Caliper Gx (Perkin Elmer). Accurate
610 quantification of the final libraries for sequencing applications was determined using the qPCR-
611 based KAPA Biosystems Library Quantification kit (Kapa Biosystems, Inc.). Each library was
612 diluted to a final concentration of 1.4 nM and pooled equimolar prior to clustering. Sequencing
613 was performed on a NovaSeq 6000 (Illumina, Inc.), to yield 100 million, 100 base pair paired-
614 end reads per sample. Sequencing was performed at HudsonAlpha Institute for Biotechnology,
615 AL, USA. RNA integrity number (RIN) and DV200 index (percentage of fragments > 200
616 nucleotides) were lower in the NOR cohort compared to the ESP cohort: (mean [min-max]),
617 NOR_{RIN}: 4.3 [1.7-7.7], ESP_{RIN}: 6.9 [2.7–9.1]; NOR_{DV200}: 85.3 [71.8-93.2], ESP_{DV200}: 92.2

618 [80.1–96.2]. No significant difference in the RNA quality measures between individuals with
619 iPD and controls was observed in either of the cohorts (linear model, NOR_{RIN}: P = 0.26,
620 NOR_{DV200}: P = 0.33; ESP_{RIN}: P = 0.11, ESP_{DV200}: P = 0.11).
621
622 **RNA-seq analysis**
623 Salmon v1.3.0 [58] was used to quantify the abundance at the transcript level with the fragment-
624 level GC bias correction option (--gcBias) against GENCODE v35. Transcript-level
625 quantification was collapsed onto gene-level quantification using the tximport R package
626 v1.8.0[65] according to the gene definitions provided by the same Ensembl release. Non-canonical
627 chromosomes and scaffolds, and genes encoded by the mitochondrial genome were excluded
628 from the analysis. In addition, we removed genes accounting for >0.6% of the reads in more
629 than half of the samples, resulting in the removal of 4 transcripts (*RN7SL2*, *RN7SL1*, *RN7SK*,
630 *LINC00632*; Supplementary Fig. S9, as well as low-expressed genes (expression <5 reads in
631 80% of the samples). At the final step, we used the expression level of the sex-specific genes
632 *XIST, KDM5D* and *RPS4Y1* to define the noise threshold for our data[32]. Relative cell abundance
633 was estimated using cell-type specific marker gene profiles (MGPs), as previously
634 described[32,66]. Difference in the cell type MGPs between the groups were assessed through
635 linear regression adjusting for age, sex, PMI and cohort, using "lm" function from the R "stats"
636 function v4.1.0. Differential expression (DE) analysis between each of iPD clusters and controls
637 was performed using "DESeq2" R package v1.32, using the protein-coding genes to calculate
638 the normalization factor, while adjusting for sex, age, PMI, cohort and the relevant MGPs.
639 Functional enrichment analysis was performed based on the one-sided P-value, previously
640 described[32,33], using the gene score resampling method implemented in the ermineR package
641 v1.0.1[67], an R wrapper package for ermineJ[68]. Enrichment was calculated for the combined list
642 of gene-sets of Gene Ontology (GO) database [69] and Kyoto Encyclopedia of Genes and

643 Genomes (KEGG) pathways [70,71] v7.2 as well as An Inventory of Mammalian Mitochondrial
644 Genes v3 (MitoCarta3.0)[72]. Only gene-sets of size >=20 and <=500 were included in the
645 analysis, resulting in 5,397 testable gene-sets. To visualize the enrichment results, pair-wise
646 similarity between significant gene-sets (excluding gene-sets from the MitoCarta database) was
647 first assessed using Jaccard Index. Next, the obtained similarity matrix was used to cluster the
648 gene-sets using the "hclust" function from the "stats" R package v4.1.2, with the "ward.D2"
649 method. The phylo tree was then created using the "ape" R package v5.5. The count matrix and
650 the code necessary to reproduce these results are available through github repository:
651 github.com: https://github.com/gsnido/pd_complex-i_stratification. All the information
652 regarding the version of the software and the packages used for the analysis are provided in the
653 SessionInfo.Rds file.
654
655 **Single nucleotide variant discovery**
656 RNA-seq reads were aligned against the hg38 reference genome using Hisat2 v2.2.1[73]. Genome
657 Analysis Toolkit (GATK) v4.1.9.0[74] was used to mark duplicate reads and split the CIGAR
658 signatures using the *MarkDuplicates* and *SplitNCigarReads* tools, respectively. Base quality
659 score recalibration was then performed using the *BaseRecalibrator* and *ApplyBQRS* tools to
660 generate the final alignment files. For each alignment, *HaplotypeCaller* was used in GVCF
661 mode (option *-ERC GVCF*) with optional arguments *-dont-use-soft-clipped-bases* and *--max-*
662 *alternate-alleles 4*. Joint genotyping was carried out for all resulting sample gVCF files using
663 the *GenomicsDBImport* and *GenotypeGVCFs* tools. The final analysis-ready multisample VCF
664 consisted only of single-nucleotide variants filtered using *VariantFiltration* with the conditions
665 QD < 5; DP < 10; FS > 60; MQ < 40. Samples were classified into the major mitochondrial
666 haplogroups using HaploGrep2 [75] based on the Phylotree Build 17[76].

667 Genotypes with < 10 reads were removed, and the genetic data was recoded into binary
668 PLINK format using VCFTools v0.1.15[77]. PLINK v1.90 was used for individual and variant-
669 level quality control (QC)[78]. Only autosomes were considered, and variants within the sex
670 chromosomes and mtDNA were removed. Individuals were excluded if their genomic data
671 showed a missing rate > 20%, abnormal heterozygosity (± 3 standard deviations) or cryptic
672 relatedness (identity by descent (IBD) > 0.2). Population stratification was studied using
673 multidimensional scaling (MDS) against the HapMap populations[79].
674 Single nucleotide variants (SNVs) were excluded if the missing genotype rate was >
675 20%, the call rates between PD and controls differed significantly ($P < 0.02$) or they displayed
676 departure from the Hardy-Weinberg equilibrium ($P < 1.0x10^{-5}$). Monomorphic variants were
677 removed, and the dataset contained no multiallelic variants after QC. Principal component (PC)
678 analysis was performed using Eigensoft[80,81].
679
680 **Single-nuclei isolation**
681 Nuclei isolation procedure was performed as described in[82] with minor modifications. For each
682 sample, 50 - 85 mg of frozen human prefrontal cortex were homogenized. All procedures were
683 performed on ice. Tissue was dissociated in 2 mL of homogenization buffer (320 mM sucrose,
684 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl [pH 7.8], 0.1 mM EDTA [pH 8.0],
685 0.1% IGEPAL CA-630, 1 mM DTT, and 0.4 U/µL RNase Inhibitor (ThermoFisher Scientific)
686 inside a 7 mL KIMBLE Dounce tissue grinder set (Merck), using 10 strokes with loose pestle
687 followed by 10 strokes with tight pestle. Homogenized tissue was filtered through a 30 µm cell
688 strainer and centrifuged at 1000 x g for 8 min at 4°C. Supernatant was discarded, and pellet was
689 resuspended in 450 µL of 2% BSA (in 1x PBS) containing 0.12 U/µL RNase Inhibitor. Cleanup
690 of the nuclei suspension was achieved with density gradient centrifugation. Equal volumes (450
691 µL) of nuclei suspension and working solution (50% OptiPrep density gradient medium

692 (Merck), 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl [pH 7.8], 0.1 mM EDTA [pH
693 8.0], and 1 mM DTT) were mixed. The density gradient was prepared by carefully layering 750
694 µL of 30% OptiPrep Solution (134 mM sucrose, 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM
695 Tris HCl [pH 7.8], 0.1 mM EDTA [pH 8.0], 1 mM DTT, 0.04% IGEPAL CA-630, and 0.17
696 U/µL RNase Inhibitor) on top of 300 µL of 40% OptiPrep Solution (96 mM sucrose, 5 mM
697 CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl [pH 7.8], 0.1 mM EDTA [pH 8.0], 1 mM
698 DTT, 0.03% IGEPAL CA-630, and 0.12 U/µL RNase Inhibitor) inside a 2 mL Sorenson
699 Dolphin microcentrifuge tube (Merck). The mixture containing the nuclei (800 µL) was slowly
700 pipetted onto the top of the OptiPrep density gradient and samples were centrifuged at 10,000
701 x g for 5 min at 4°C using a fixed angle rotor (FA-45-24-11-Kit). Nuclei were recovered by
702 withdrawing a volume of 200 µL from the 30 – 40% interface of the OptiPrep density gradient
703 and transferring to a new 2 mL Sorenson Dolphin microcentrifuge tube (Merck). Nuclei were
704 washed with 2% BSA (in 1x PBS) containing 0.12 U/µL RNase Inhibitor and then pelleted by
705 centrifugation at 300 x g for 3 min at 4°C using a swing-bucket rotor (S-24-11-AT) for a total
706 of three times. The nuclear pellet was resuspended in a minimum volume of 100 µL of 2% BSA
707 (in 1x PBS) containing 0.12 U/µL RNase Inhibitor and carefully mixed by pipetting 20 times
708 on ice. The nuclei suspension was filtered through a 40 µm FlowMe cell strainer (Merck) to
709 remove any remaining nuclei aggregates. Nuclei were stained with Trypan blue and counted
710 manually on a hemocytometer.
711 10,000 nuclei per sample were targeted for droplet-based snRNA sequencing. cDNA synthesis
712 and libraries were prepared using the Chromium Next GEM Single Cell 3′ Reagent Kit v3.1
713 Dual Index (10x Genomics) according to the manufacturer's protocol. Libraries from each
714 sample were pooled and run using paired-end (300 million read pairs per sample) sequencing
715 on the NovaSeq 6000 platform (Illumina) at Novogene (Cambridge).
716

717 **snRNA-seq analyses**

718 Raw FASTQ files were aligned to the human genome (GRCh38) using CellRanger v.6.1.2 (10x
719 Genomics) accounting also for intronic reads to include all nuclear transcripts (include-introns
720 option) and excluding secondary mappings (nosecondary option). In order to estimate
721 intronic:exonic ratio, an additional alignment was carried only accounting for reads mapped to
722 exonic reads. Version 35 of the ENCODE transcriptome assembly was used (Ensembl 101).

723 To minimize the confounding effect of non-nuclear ambient RNA and discard empty
724 droplets, we employed CellBender[83] on the raw data prior to importing the counts into R.
725 Percentage of mitochondrially-encoded genes and ribosomal genes were then calculated for
726 each droplet using the Seurat R Package (*PercentageFeatureSet*)[84]. Each biological sample was
727 independently normalized, and automatic feature selection was carried out using Seurat
728 (*NormalizeData*, *FindVariableFeatures*). In order to collectively analyze the datasets, samples
729 were integrated using the standard Seurat integration strategy. Briefly, features (genes) were
730 automatically selected using *SelectIntegrationFeatures* to scale each sample and obtain a
731 principal component analysis (PCA) of gene expression. Using the first 50 principal
732 components, samples were integrated using reciprocal PCA with automatically detected
733 anchors (*FindIntegrationAnchors*) and using controls to build the reference (*IntegrateData*).
734 Clusters were first identified using the standard Seurat analysis with excess barcodes (i.e.,
735 including twice as many barcodes in each sample to ensure the inclusion of a large population
736 of empty droplets). On each sample the largest cluster always corresponded to the cluster with
737 the largest proportion of empty droplets (as defined by DropletUtils[85]). Visual inspection of the
738 UMAP plots allowed the identification of other smaller ambient clusters which were removed
739 together with the main empty droplet cluster for downstream analyses. Resulting non-empty
740 droplets were then further screened and only droplets with at least 500 transcripts, at least 1,000
741 UMIs, and below 3% mitochondrially-encoded transcripts were analyzed. Transcripts detected

742 in fewer than 10 cells were also filtered out. Doublets were assessed independently for each
743 sample using the DoubletFinder R package[86] assuming 7% doublet rate. Standard clustering
744 using Seurat with 0.5 resolution resulted in a total of 28 clusters (Fig. 7a) which could be
745 assigned to the main cortical cell types based on known markers (SATB2 and SLC17A7 for
746 excitatory neurons, GAD1 and SLC6A1 for inhibitory neurons, MOG and MOBP for
747 oligodendroglia, GFAP and AQP4 for astrocytes, CLDN5 and SLC2A1 for endothelial cells,
748 OLIG1 and VCAN for oligodendrocyte precursor cells, and CSF1R and CSF3R for microglia[87],
749 Fig. 7b). Finally, to increase the within-cluster cell type homogeneity and minimize the number
750 of doublets, we used two alternative approaches. First, the Euclidean distance in PCA space
751 was calculated between each single cell and the average of each cluster using the first 5 PCs.
752 Barcodes for which the shortest distance did not correspond to their cluster membership were
753 removed from the analyses. We then calculated the first PC of the cell type specific markers for
754 each barcode (marker gene profiles, MGP[66]). MGPs correspond to the first PC of the cell type
755 specific markers. In this way, we obtained for each droplet a value for each major cell type (i.e.,
756 neuron, oligodendrocyte, endothelial cell, microglia, astrocyte, OPC) scaled across all droplets.
757 We then calculated a score for each droplet, corresponding to the MGP for the annotated cell
758 type divided by the sum of the other MGPs. Droplets with poor scores (i.e., below 0.4) were
759 filtered out. The filtering procedure resulted in a total of 117,105 nuclei. The cluster labels were
760 used to compare cell type composition of the diseased and control groups using ANOVA.

761 To assess the statistical significance of the difference in the number of captured nuclei
762 between groups (nCI-PD, CI-PD and Ctrl) we used an ANOVA test at $\alpha = 0.05$. Differences
763 between groups in the number of genes per nucleus were assessed for each cell type cluster
764 with linear mixed-effects models (LMM) with disease status as fixed effect and individual as
765 random effect. P-values for each LMM were estimated using the Satterthwaite's method. The
766 resulting p-values were considered significant above a false discovery rate (FDR) of 5%. The

767 same statistical approach was used to assess differences in the number of unique molecular
768 (UMI) counts between disease groups.

769 Differential expression within clusters between diseased groups and controls was
770 carried out using the MAST R package[88] using sex, post-mortem interval and age as fixed
771 effects, and individual as random effect[89] (Fig. 7c). For each cluster, only autosomal, protein-
772 coding genes that were expressed were considered (i.e., with counts > 0 in 25% of the barcodes).
773 The same methodology was used to compare between the CI-PD and nCI-PD subtypes.
774 Overrepresentation analysis was carried out using the fora function from the fgsea R package[90]
775 independently for each cell type cluster. The background ("universe") for each
776 overrepresentation test consisted of the set of expressed genes for each specific cell type cluster.

777 UMI counts and code necessary to reproduce these results is available through a github
778 repository: https://github.com/gsnido/pd_complex-i_stratification.

779

780 **Statistical analyses and reproducibility of the demographic, clinical, IHC and mtDNA**
781 **data**

782 Subject demographics, clinical data, and pathology scores are given in Supplementary Data 1.
783 Experimental allocation is given in Supplementary Data 2. The sample size was not
784 predetermined. No data were excluded from the analyses. Apart from the IHC staining of the
785 tissues from the ESP cohort, the experiments were not randomized. Observer 1 was blinded to
786 allocation during IHC staining and outcome assessments of the CI positivity in the ESP cohort.
787 Observer 2 in the IHC analyses of CI positive staining was blinded to group specificity.
788 Investigators were blinded during the preprocessing of the RNAseq data. Apart from this, the
789 investigators were not blinded to allocation during experiments and outcome assessments. The
790 visual analysis of the IHC stainings for CI was performed by two independent observers (IHC

791 and CT). IHC findings from the NOR cohort were replicated in the ESP cohort. The samples
792 from single cell qPCR were run in triplicates on the same 96 well plate.

793 Data were tested for normality and equality of variances, using the Shapiro-Wilk test
794 and Levene's test. Clustering was performed in R (v4.1.2), using k-means with three centers,
795 20 random starting partitions, 100 iterations, and the Hartigan-Wong algorithm. Between
796 group-comparisons were consequently performed using either independent student's t-test or
797 Mann-Whitney U-test, depending on data normality, and descriptive data are given either as
798 mean (standard deviation) or median [25th-percentile-75th-percentile], respectively.
799 Comparisons of grouped data were performed using either Chi-Square (X^2) or Fisher's exact
800 test depending on the group size of each variable. Data is given as X^2 and/or odds ratio (95%
801 confidence interval). Correlation analyses were performed using Spearman's rank correlation
802 due to non-normality of the data, Spearman's rank correlation coefficient is given as $\rho(n-2)$.
803 Between-group (Ctrl, CI-PD and nCI-PD) comparison of neuronal mtDNA deletion proportion
804 and copy number was performed by one-way ANOVA on the individual median values. Due
805 to the unequal sample sizes, Welch's test for unequal variance was used. Eta squared (η^2) was
806 used to determine the effect size of any significant difference. The Games-Howell post hoc test
807 was used to assess the statistical significance of pairwise differences between groups, due to
808 their unequal sample sizes and variance. Hierarchical clustering analysis was performed in R
809 v4.1.2 with the "ward.D2" method. Correlation analyses and between-group comparisons were
810 done in SPSS v28.0.0. Plots were created in R v4.1.2 or GraphPad Prism v9.3.0. A P-value <
811 0.05 was considered statistically significant. All tests are 2-sided unless else is noted in the text.
812 Statistical tests performed in the context of omics data analyses are described in the
813 corresponding methods sections.

Data availability statement

The IHC and mtDNA qPCR data generated in this study are provided in the Supplementary Information/Source Data file. The raw and filtered counts from the transcriptomic data are available at Synapse.org: (https://doi.org/10.7303/syn53502566.1)[91]. The raw transcriptomic data of the bulk tissue samples (accession number : EGAD50000000432; https://ega-archive.org/datasets/EGAD50000000432) and the single nuclei samples (accession number : EGAD50000000433; https://ega-archive.org/datasets/EGAD50000000433) are available in the Federated European Genome-phenome Archive (FEGA) Norway.

Code availability statement

The code required to reproduce the results of the transcriptomic analyses are available in github.com: https://github.com/gsnido/pd_complex-i_stratification (https://doi.org/10.5281/zenodo.10715023)[92].

References

1. Bloem BR, Okun MS, Klein C. Parkinson's disease. The Lancet 2021;397(10291):2284–2303.
2. Kalia LV, Lang AE. Parkinson's disease. Lancet 2015;386(9996):896–912.
3. Athauda D, Foltynie T. The ongoing pursuit of neuroprotective therapies in Parkinson disease. Nat. Rev. Neurol. 2015;11(1):25–40.
4. Espay AJ, Schwarzschild MA, Tanner CM, et al. Biomarker-driven phenotyping in Parkinson disease: a translational missing link in disease-modifying clinical trials. Mov. Disord. Off. J. Mov. Disord. Soc. 2017;32(3):319–324.
5. Dickson DW. Parkinson's Disease and Parkinsonism: Neuropathology. Cold Spring Harb. Perspect. Med. 2012;2(8):a009258–a009258.
6. Chen-Plotkin AS, Albin R, Alcalay R, et al. Finding Useful Biomarkers for Parkinson's Disease. Sci. Transl. Med. 2018;10(454):eaam6003.

842 7. Greenland JC, Williams-Gray CH, Barker RA. The clinical heterogeneity of Parkinson's
843 disease and its therapeutic implications. Eur J Neurosci 2019;49(3):328–338.

844 8. Dubois B, Burn D, Goetz C, et al. Diagnostic procedures for Parkinson's disease
845 dementia: Recommendations from the movement disorder society task force. Mov.
846 Disord. 2007;22(16):2314–2324.

847 9. Schiess MC, Zheng H, Soukup VM, et al. Parkinson's disease subtypes: clinical
848 classification and ventricular cerebrospinal fluid analysis. Parkinsonism Relat. Disord.
849 2000;6(2):69–76.

850 10. Jankovic J, McDermott M, Carter J, et al. Variable expression of Parkinson's disease: a
851 base-line analysis of the DATATOP cohort. The Parkinson Study Group. Neurology
852 1990;40(10):1529–34.

853 11. Grünewald A, Kumar KR, Sue CM. New insights into the complex role of mitochondria in
854 Parkinson's disease. Prog. Neurobiol. 2019;177:73–93.

855 12. Schapira AH, Cooper JM, Dexter D, et al. Mitochondrial complex I deficiency in
856 Parkinson's disease. Lancet 1989;1(8649):1269.

857 13. Subrahmanian N, LaVoie MJ. Is there a special relationship between complex I activity
858 and nigral neuronal loss in Parkinson's disease? A critical reappraisal. Brain Res.
859 2021;147434.

860 14. Flønes IH, Tzoulis C. Mitochondrial Respiratory Chain Dysfunction—A Hallmark
861 Pathology of Idiopathic Parkinson's Disease? Front. Cell Dev. Biol. 2022;10:874596.

862 15. Ramsay RR, Salach JI, Singer TP. Uptake of the neurotoxin 1-methyl-4-phenylpyridine
863 (MPP+) by mitochondria and its relation to the inhibition of the mitochondrial oxidation
864 of NAD+-linked substrates by MPP+. Biochem. Biophys. Res. Commun.
865 1986;134(2):743–748.

866 16. Langston JW, Ballard P, Tetrud JW, Irwin I. Chronic Parkinsonism in humans due to a
867 product of meperidine-analog synthesis. Science 1983;219(4587):979–980.

868 17. Betarbet R, Sherer TB, MacKenzie G, et al. Chronic systemic pesticide exposure
869 reproduces features of Parkinson's disease. Nat. Neurosci. 2000;3(12):1301–1306.

870 18. Reinson K, Õunap K. Complex I deficiency and Leigh syndrome through the eyes of a
871 clinician. EMBO Mol. Med. 2020;12(11):e13187.

872 19. Nikoskelainen EK, Marttila RJ, Huoponen K, et al. Leber's "plus": neurological
873 abnormalities in patients with Leber's hereditary optic neuropathy. J. Neurol.
874 Neurosurg. Psychiatry 1995;59(2):160–4.

875 20. Balafkan N, Tzoulis C, Muller B, et al. Number of CAG repeats in POLG1 and its
876 association with Parkinson disease in the Norwegian population. Mitochondrion
877 2012;12(6):640–3.

878 21. Gaweda-Walerych K, Safranow K, Maruszak A, et al. Mitochondrial transcription factor
879 A variants and the risk of Parkinson's disease. Neurosci Lett 2010;469(1):24–9.

880 22. Benecke R, Strümper P, Weiss H. Electron transfer complexes I and IV of platelets are
881 abnormal in Parkinson's disease but normal in Parkinson-plus syndromes. Brain
882 1993;116(6):1451–1463.

883 23. Yoshino H, Nakagawa-Hattori Y, Kondo T, Mizuno Y. Mitochondrial complex I and II
884 activities of lymphocytes and platelets in Parkinson's disease. J. Neural Transm. Park.
885 Dis. Dement. Sect. 1992;4(1):27–34.

886 24. Carling PJ, Mortiboys H, Green C, et al. Deep phenotyping of peripheral tissue facilitates
887 mechanistic disease stratification in sporadic Parkinson's disease. Prog. Neurobiol.
888 2020;187:101772.

889 25. Winkler-Stuck K, Wiedemann FR, Wallesch C-W, Kunz WS. Effect of coenzyme Q10 on
890 the mitochondrial function of skin fibroblasts from Parkinson patients. J. Neurol. Sci.
891 2004;220(1–2):41–48.

892 26. Stroud DA, Surgenor EE, Formosa LE, et al. Accessory subunits are integral for assembly
893 and function of human mitochondrial complex I. Nature 2016;538(7623):123–126.

894 27. Flønes IH, Fernandez-Vizarra E, Lykouri M, et al. Neuronal complex I deficiency occurs
895 throughout the Parkinson's disease brain, but is not associated with neurodegeneration
896 or mitochondrial DNA damage. Acta Neuropathol. (Berl.) 2018;135(3):409–425.

897 28. Dölle C, Flønes I, Nido GS, et al. Defective mitochondrial DNA homeostasis in the
898 substantia nigra in Parkinson disease. Nat. Commun. 2016;7(1):13548.

899 29. Bender A, Krishnan KJ, Morris CM, et al. High levels of mitochondrial DNA deletions in
900 substantia nigra neurons in aging and Parkinson disease. Nat Genet 2006;38(5):515–7.

901 30. Zarow C, Lyness SA, Mortimer JA, Chui HC. Neuronal Loss Is Greater in the Locus
902 Coeruleus Than Nucleus Basalis and Substantia Nigra in Alzheimer and Parkinson
903 Diseases. Arch. Neurol. 2003;60(3):337–341.

904 31. Streiner DL. Best (but oft-forgotten) practices: the multiple problems of multiplicity-
905 whether and how to correct for many statistical tests. Am. J. Clin. Nutr.
906 2015;102(4):721–728.

907 32. Toker L, Mancarci BO, Tripathy S, Pavlidis P. Transcriptomic Evidence for Alterations in
908 Astrocytes and Parvalbumin Interneurons in Subjects With Bipolar Disorder and
909 Schizophrenia. Biol. Psychiatry 2018;84(11):787–796.

910 33. Nido GS, Dick F, Toker L, et al. Common gene expression signatures in Parkinson's
911 disease are driven by changes in cell composition. Acta Neuropathol. Commun.
912 2020;8(1):55.

913 34. Toker L, Tran GT, Sundaresan J, et al. Genome-wide histone acetylation analysis reveals
914 altered transcriptional regulation in the Parkinson's disease brain. Mol. Neurodegener.
915 2021;16(1):31.

916 35. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal
917 recessive juvenile parkinsonism. Nature 1998;392(6676):605–8.

918 36. Lake BB, Ai R, Kaeser GE, et al. Neuronal subtypes and diversity revealed by single-
919 nucleus RNA sequencing of the human brain. Science 2016;352(6293):1586–1590.

920 37. Shulman RG, Rothman DL, Behar KL, Hyder F. Energetic basis of brain activity:
921 implications for neuroimaging. Trends Neurosci. 2004;27(8):489–495.

922 38. Pathak RU, Davey GP. Complex I and energy thresholds in the brain. Biochim. Biophys.
923 Acta 2008;1777(7–8):777–782.

924 39. González-Rodríguez P, Zampese E, Stout KA, et al. Disruption of mitochondrial complex
925 I induces progressive parkinsonism. Nature 2021;599(7886):650–656.

926 40. Boonstra JT, McGurran H, Temel Y, Jahanshahi A. Nigral neuropathology of Parkinson's
927 motor subtypes coincide with circuitopathies: a scoping review. Brain Struct. Funct.
928 2022;227(7):2231–2242.

929 41. Brakedal B, Dölle C, Riemer F, et al. The NADPARK study: A randomized phase I trial of
930 nicotinamide riboside supplementation in Parkinson's disease. Cell Metab.
931 2022;34(3):396-407.e6.

932 42. Rossignol R, Faustin B, Rocher C, et al. Mitochondrial threshold effects. Biochem J
933 2003;370(Pt 3):751–62.

934 43. Kraytsberg Y, Kudryavtseva E, McKee AC, et al. Mitochondrial DNA deletions are
935 abundant and cause functional impairment in aged human substantia nigra neurons.
936 Nat Genet 2006;38(5):518–20.

937 44. Rocha MC, Rosa HS, Grady JP, et al. Pathological mechanisms underlying single large-
938 scale mitochondrial DNA deletions. Ann. Neurol. 2018;83(1):115–130.

939 45. Verkaart S, Koopman WJH, van Emst-de Vries SE, et al. Superoxide production is
940 inversely related to complex I activity in inherited complex I deficiency. Biochim.
941 Biophys. Acta BBA - Mol. Basis Dis. 2007;1772(3):373–381.

942 46. Zsurka G, Peeva V, Kotlyar A, Kunz WS. Is There Still Any Role for Oxidative Stress in
943 Mitochondrial DNA-Dependent Aging? Genes 2018;9(4):175.

944 47. Lax NZ, Grady J, Laude A, et al. Extensive respiratory chain defects in inhibitory
945 interneurones in patients with mitochondrial disease. Neuropathol. Appl. Neurobiol.
946 2016;42(2):180–193.

947 48. Toker L, Nido GS, Tzoulis C. Not every estimate counts – evaluation of cell composition
948 estimation approaches in brain bulk tissue data. Genome Med. 2023;15:41.

949 49. Borrageiro G, Haylett W, Seedat S, et al. A review of genome-wide transcriptomics
950 studies in Parkinson's disease. Eur. J. Neurosci. 2018;47(1):1–16.

951 50. Alecu I, Bennett SAL. Dysregulated Lipid Metabolism and Its Role in α-Synucleinopathy
952 in Parkinson's Disease. Front. Neurosci. 2019;13:328.

953 51. Shao Y, Li T, Liu Z, et al. Comprehensive metabolic profiling of Parkinson's disease by
954 liquid chromatography-mass spectrometry. Mol. Neurodegener. 2021;16(1):4.

955 52. Picca A, Faitg J, Auwerx J, et al. Mitophagy in human health, ageing and disease. Nat.
956 Metab. 2023;5(12):2047–2061.

957 53. Agarwal D, Sandor C, Volpato V, et al. A single-cell atlas of the human substantia nigra
958 reveals cell-specific pathways associated with neurological disorders. Nat. Commun.
959 2020;11(1):4183.

960 54. Alves G, Muller B, Herlofson K, et al. Incidence of Parkinson's disease in Norway: the
961 Norwegian ParkWest study. J. Neurol. Neurosurg. Psychiatry 2009;80(8):851–857.

962 55. Gaare JJ, Nido GS, Sztromwasser P, et al. Rare genetic variation in mitochondrial
963 pathways influences the risk for Parkinson's disease: Mitochondrial Pathways In PD.
964 Mov. Disord. 2018;33(10):1591–1600.

965 56. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision
966 of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and
967 clinimetric testing results: MDS-UPDRS: Clinimetric Assessment. Mov. Disord.
968 2008;23(15):2129–2170.

969 57. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for
970 grading the cognitive state of patients for the clinician. J. Psychiatr. Res.
971 1975;12(3):189–198.

972 58. Borrego-Écija S, Turon-Sans J, Ximelis T, et al. Cognitive decline in amyotrophic lateral
973 sclerosis: Neuropathological substrate and genetic determinants. Brain Pathol. Zurich
974 Switz. 2021;31(3):e12942.

975 59. Braak H, Del Tredici K, Rüb U, et al. Staging of brain pathology related to sporadic
976 Parkinson's disease. Neurobiol. Aging 2003;24(2):197–211.

977 60. Braak H, Alafuzoff I, Arzberger T, et al. Staging of Alzheimer disease-associated
978 neurofibrillary pathology using paraffin sections and immunocytochemistry. Acta
979 Neuropathol. (Berl.) 2006;112(4):389–404.

980 61. Attems J, Toledo JB, Walker L, et al. Neuropathological consensus criteria for the
981 evaluation of Lewy pathology in post-mortem brains: a multi-centre study. Acta
982 Neuropathol. (Berl.) 2021;141(2):159–172.

983 62. Alafuzoff I, Arzberger T, Al-Sarraj S, et al. Staging of Neurofibrillary Pathology in
984 Alzheimer's Disease: A Study of the BrainNet Europe Consortium. Brain Pathol.
985 2008;18(4):484–496.

986 63. Thal DR, Rüb U, Orantes M, Braak H. Phases of A beta-deposition in the human brain
987 and its relevance for the development of AD. Neurology 2002;58(12):1791–1800.

988 64. Dickson DW, Braak H, Duda JE, et al. Neuropathological assessment of Parkinson's
989 disease: refining the diagnostic criteria. Lancet Neurol. 2009;8(12):1150–1157.

990 65. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level
991 estimates improve gene-level inferences. F1000Research 2016;4:1521.

992 66. Mancarci BO, Toker L, Tripathy SJ, et al. Cross-Laboratory Analysis of Brain Cell Type
993 Transcriptomes with Applications to Interpretation of Bulk Tissue Data. eNeuro
994 2017;ENEURO.0212-17.2017.

995 67. Mancarci O. ErmineR (R package) [Internet]. [date unknown];Available from:
996 https://github.com/PavlidisLab/ermineR

997 68. Gillis J, Mistry M, Pavlidis P. Gene function analysis in complex data sets using ErmineJ.
998 Nat. Protoc. 2010;5(6):1148–1159.

999 69. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology.
1000 The Gene Ontology Consortium. Nat. Genet. 2000;25(1):25–29.

1001 70. Kanehisa M. The KEGG resource for deciphering the genome. Nucleic Acids Res.
1002 2004;32(90001):277D – 280.

1003 71. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB)
1004 3.0. Bioinformatics 2011;27(12):1739–1740.

1005 72. Rath S, Sharma R, Gupta R, et al. MitoCarta3.0: an updated mitochondrial proteome
1006 now with sub-organelle localization and pathway annotations. Nucleic Acids Res.
1007 2021;49(D1):D1541–D1547.

1008 73. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory
1009 requirements. Nat. Methods 2015;12(4):357–360.

1010 74. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence
1011 variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc
1012 Bioinforma. 2013;43:11 10 1-33.

1013 75. Weissensteiner H, Pacher D, Kloss-Brandstätter A, et al. HaploGrep 2: mitochondrial
1014 haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res.
1015 2016;44(W1):W58-63.

1016 76. Dür A, Huber N, Parson W. Fine-Tuning Phylogenetic Alignment and Haplogrouping of
1017 mtDNA Sequences. Int. J. Mol. Sci. 2021;22(11):5747.

1018 77. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools.
1019 Bioinformatics 2011;27(15):2156–8.

1020 78. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of
1021 larger and richer datasets. Gigascience 2015;4:7.

1022 79. Gibbs RA, Belmont JW, Hardenbol P, et al. The International HapMap Project. Nature
1023 2003;426(6968):789–796.

1024 80. Patterson N, Price AL, Reich D. Population Structure and Eigenanalysis. PLOS Genet.
1025 2006;2(12):e190.

1026 81. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for
1027 stratification in genome-wide association studies. Nat Genet 2006;38(8):904–9.

1028 82. Lee H, Fenster RJ, Pineda SS, et al. Cell Type-Specific Transcriptomics Reveals that
1029 Mutant Huntingtin Leads to Mitochondrial RNA Release and Neuronal Innate Immune
1030 Activation. Neuron 2020;107(5):891-908.e8.

1031 83. Fleming SJ, Chaffin MD, Arduini A, et al. Unsupervised removal of systematic
1032 background noise from droplet-based single-cell experiments using CellBender
1033 [Internet]. 2022;791699.[cited 2023 Jun 17] Available from:
1034 https://www.biorxiv.org/content/10.1101/791699v2

1035 84. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell
1036 data. Cell 2021;184(13):3573-3587.e29.

1037 85. Griffiths JA, Richard AC, Bach K, et al. Detection and removal of barcode swapping in
1038 single-cell RNA-seq data. Nat. Commun. 2018;9(1):2667.

1039 86. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell
1040 RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. 2019;8(4):329-
1041 337.e4.

1042 87. Lake BB, Ai R, Kaeser GE, et al. Neuronal subtypes and diversity revealed by single-
1043 nucleus RNA sequencing of the human brain. Science 2016;352(6293):1586–1590.

1044 88. Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing
1045 transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing
1046 data. Genome Biol. 2015;16(1):278.

1047 89. Zimmerman KD, Espeland MA, Langefeld CD. A practical solution to pseudoreplication
1048 bias in single-cell studies. Nat. Commun. 2021;12(1):738.

1049 90. Korotkevich G, Sukhov V, Budin N, et al. Fast gene set enrichment analysis [Internet].
1050 2021;060012.[cited 2024 Jan 2] Available from:
1051 https://www.biorxiv.org/content/10.1101/060012v3

1052 91. Nido G. CI deficiency PD stratification [Internet]. 2024;[cited 2024 Apr 10] Available
1053 from: https://repo-
1054 prod.prod.sagebase.org/repo/v1/doi/locate?id=syn53502566&type=ENTITY&version=1

92. gsnido. gsnido/pd_complex-i_stratification: Code release for publication [Internet]. 2024;[cited 2024 Apr 10] Available from: https://zenodo.org/records/10715023

93. Edlow BL, Mareyam A, Horn A, et al. 7 Tesla MRI of the ex vivo human brain at 100 micron resolution. Sci. Data 2019;6(1):244.

Acknowledgements

We are deeply grateful to the study participants and their families involved in the study for their unique contribution. Furthermore, we would like to thank the entire Park West study consortium for their diligent efforts characterizing and following the Park West cohort; Dr. Yamila Torres Cleuren for the inspirational feedback and critical review of the manuscript; the Neurological Tissue Bank of the IDIBAPS-Hospital Clinic Biobank for providing data and samples; Dr. Brian L Edlow and colleagues for allowing us to use images from their 7 Tesla MRI atlas in Fig. 3; Dr. Romain Guitton, Dr. Kristoffer Haugarvoll and Janani Sundaresan for helping dissect and prepare the brain tissue for RNA-seq; Gry Hilde Nilsen for technical support.

This work is supported by grants from The Research Council of Norway (288164; CT), Bergen Research Foundation (BFS2017REK05; CT), The KG Jebsen Foundation (SKGJ-MED-023; CT), and the Western Norway Regional Health Authority (F-10229-D11661; IHF).

Authors' contributions

IHF: participated in the study conception and design, designed and performed the IHC and qPCR experiments, assessed the neuropathology of the NOR cohort, analyzed the IHC, mtDNA, clinical and pathology data, and drafted the manuscript. LT: participated in the study design, analysed and interpreted the transcriptomics data, participated in, and advised on statistical approaches, and drafted parts of the manuscript. EFV: critical revision of the manuscript. DAS and OS: performed part of the IHC stainings. MG and SM: nuclei isolation from brain tissue and library preparation for snRNA-seq. NL: contributed to the preparation of Figure 2, advised on the statistical approach, and provided critical input for the manuscript. GA

1086 and OBT: contributed biological material and clinical data and provided critical input to the
1087 manuscript. AP-S, CP, YC: contributed clinical data and input to the manuscript. LMP:
1088 contributed biological material, pathological data and critical input to the manuscript. CD:
1089 contributed to the qPCR analyses and provided critical input for the manuscript. GSN:
1090 participated in the study design, performed the prepossessing of all transcriptomic and genetic
1091 data, performed the mtDNA variant analyses, participated in the interpretation of the
1092 transcriptomic and genetic data, and drafted parts of the manuscript. CT: conceived, designed
1093 and directed the study, contributed to data analyses and interpretation, drafted the manuscript,
1094 and acquired funding for the study. All authors have read and approved the manuscript.
1095

1096 **Competing interests**

1097 The authors declare no competing interests.
1098

1099 **Table 1. Demographic information**

Variable	**Group**	***n***	**Median**	**IQR**	**U**	**P-value** (Two-sided **Mann-Whitney U test)**
PMI	NOR: Ctrl	13	43.0	31.5-60.0	188.0	0.16
	NOR: PD	39	30.0	24.0-48.0		
	ESP: Ctrl	11	7.5	6.0-10.3	242.0	0.48
	ESP: PD	51	8.0	6.0-14.0		
	Combined: Ctrl	24	27.0	7.5-46.8	933.5	0.31
	Combined: PD	90	16.0	7.0-30.0		
Variable	Group	*n*	Mean	SD	t	P-value (Two-sided Student's t-test)
Age of death	NOR: Ctrl	13	82.8	10.5	-1.20	0.24
	NOR: PD	41	79.6	7.3		
	ESP: Ctrl	11	83.5	8.9	-2.10	0.038
	ESP: PD	51	78.0	7.4		
	Combined: Ctrl	24	83.1	9.6	-2.40	0.018
	Combined: PD	92	78.8	7.4		
Variable	Group	*n*		OR male (PD/Ctrl)	95% C.I.	P-value (Two-sided Fisher's exact test)
Sex	NOR: Ctrl, male	4		3.90	1.0-14.9	0.056
	NOR: Ctrl, female	9				
	NOR: PD, male	26				
	NOR: PD, female	15				
	ESP: Ctrl, male	7		1.25	0.3-4.9	0.74
	ESP: Ctrl, female	4				
	ESP: PD, male	35				
	ESP: PD, female	16				
	Combined: Ctrl, male	11		2.33	0.9-5.9	0.097
	Combined: Ctrl, female	13				
	Combined: PD, male	61				
	Combined: PD, female	31				

Demographic information and group comparisons for individuals included in the immunohistochemistry experiments. PMI: post-mortem interval, IQR: interquartile range, SD: standard deviation, OR: odds ratio, C.I.: confidence interval. NOR: Norwegian cohort, ESP: Spanish cohort, PD: Parkinson's disease, Ctrl: controls.

1107 **Table 2. Between group comparisons of percentage CI positive neurons**
1108

Tissue	Cohort	Group	*n*	Median	IQR	U-value	**P-value (**Two-sided **Mann-Whitney U)**
SNpc	NOR	iPD	25	78.6	67.1-83.9	44.5	$8.2x10^{-4}$
	NOR	Ctrl	11	90.8	83.6-97.8		
	ESP	iPD	35	75.0	65.7-87.1	148.5	0.26
	ESP	Ctrl	11	80.1	60.7-89.4		
	Combined	iPD	60	76.3	66.7-84.2	372.5	0.003
	Combined	Ctrl	22	87.1	77.1-94.7		
PFC	NOR	iPD	40	96.6	90.3-98.1	84.0	$7.0x10^{-4}$
	NOR	Ctrl	12	98.6	98.0-99.1		
	ESP	iPD	49	97.1	94.1-98.1	138.0	0.031
	ESP	Ctrl	10	98.3	97.0-99.7		
	Combined	iPD	89	96.8	92.8-98.1	431.0	$5.0x10^{-5}$
	Combined	Ctrl	22	98.6	97.8-99.2		

1109
1110 IQR: interquartile range, SNpc: substantia nigra pars compacta, PFC: prefrontal cortex, NOR:

1111 Norwegian cohort, ESP: Spanish cohort, iPD: idiopathic Parkinson's disease, Ctrl: controls

1112 **Fig. 1. Neuronal CI deficiency stratifies iPD**

1113 (a-b) SNpc (a) and PFC (b) sections immunohistochemically stained for CI (NDUFS4). The
1114 stippled lines in (a) indicate neuromelanin in dopaminergic neurons of the SNpc. Arrowheads
1115 show examples of CI deficient (green) and -intact (black) neurons. Scalebar: 50 µm. (c-d) Inter-
1116 rater and inter-region replicability of the proportion of CI-positive neurons in the PFC. Bland-
1117 Altmann plots showing the difference between the measurements of observer 1 (O1) and
1118 observer 2 (O2) (c), and the difference between measurements by observer 1 in three different
1119 regions of the same PFC sections: region 1 (R1) vs region 2 (R2), or region 3 (R3) (d). The Y-
1120 axes show the respective average percentages of CI-positive neurons. (e-f) Scatter plots
1121 showing the proportion (in %) of CI positive neurons (y-axes) of controls and individuals with
1122 iPD (x-axes) in the SNpc (controls NOR: n =11, PD NOR: n =25, controls ESP: n =11, PD
1123 ESP: n =35; e) and PFC (controls NOR: n =12, PD NOR: n =40, controls ESP: n =10, PD ESP:
1124 n =49; f). Bars show median and interquartile range. The groups were compared using a two-
1125 sided the Mann Whitney U-test. The test statistics are given in Table 2. Multiple comparison
1126 adjustments were not made. (g) Scatter plots showing the proportion (in %) of CI positive
1127 neurons (y-axes) in the PFC of controls and individuals with iPD (x-axes). K-means clustering
1128 classified the iPD subjects into three groups: nCI-PD (purple), CI-PD mild (dark green) and CI-
1129 PD severe (light green), in each of the NOR and ESP cohorts and in the combined cohorts.

1130 **Fig. 2. CI and VDAC1 immunostaining in selected brain regions**

1131 Representative images of immunostaining against CI (NDUFS4) and VDAC1, in areas showing
1132 CI deficiency in iPD. Green arrowheads show examples of CI negative neurons. No negative
1133 neurons were observed in the VDAC1 staining. Scalebar: 50 µm.

1134

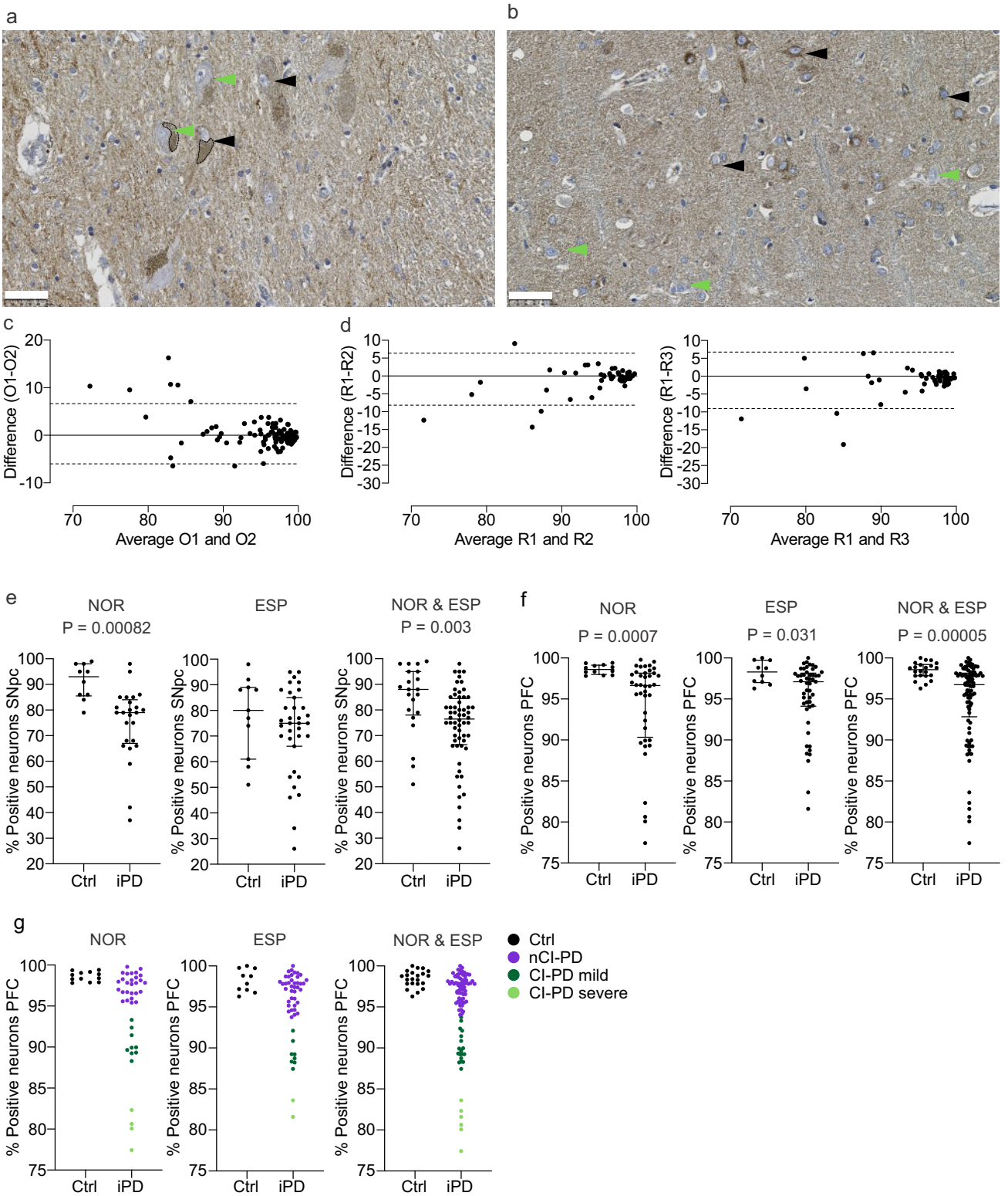
1135 **Fig. 3. Anatomical distribution of neuronal CI deficiency in the nCI-PD and CI-PD groups**

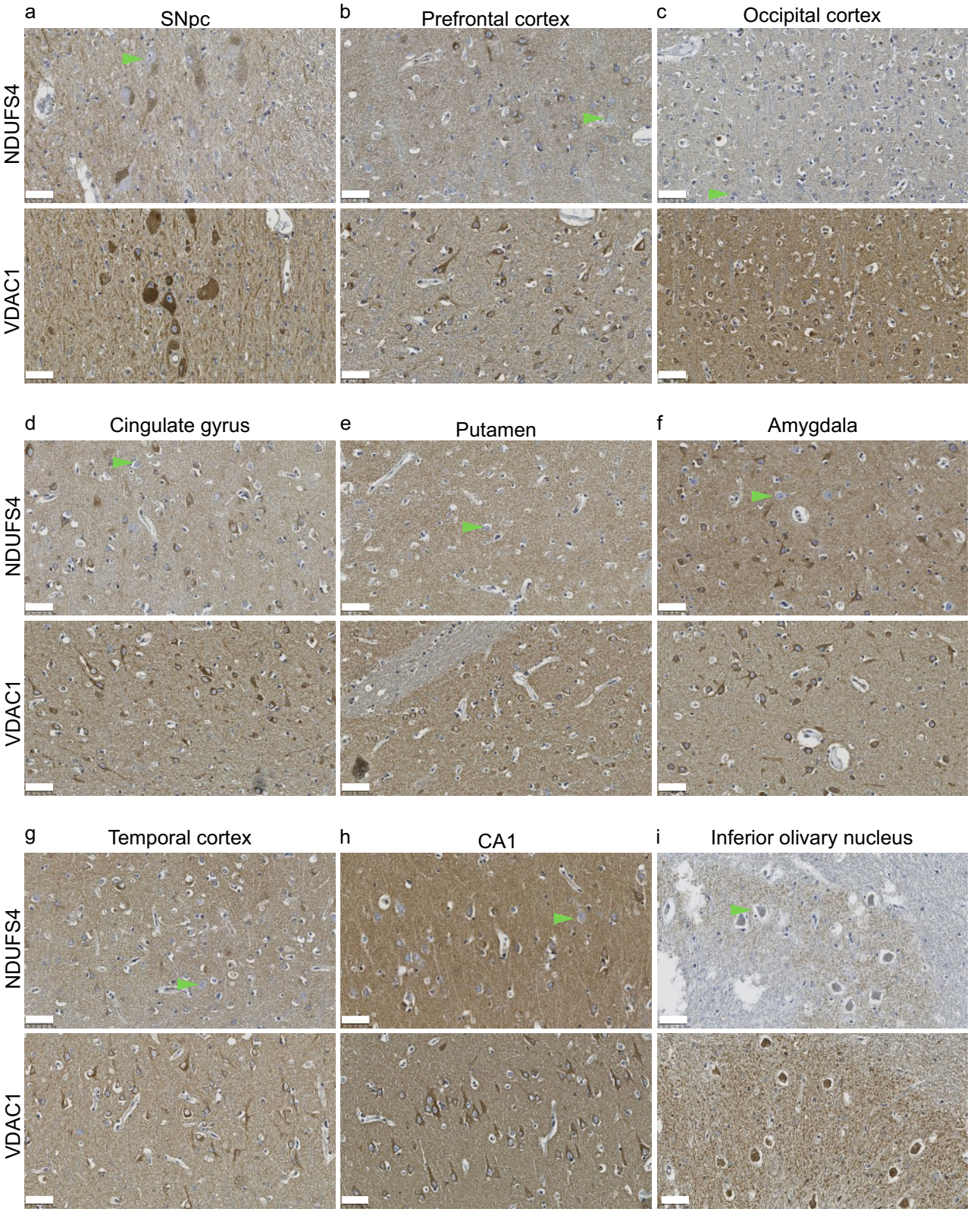
1136 (a) Box plots showing the proportion (in %) of CI positive neurons (Y-axes) in 16 brain regions
1137 from controls (grey), and individuals with nCI-PD (purple) and CI-PD (green). The boxplot
1138 from PFC represents the spread in all individuals, dots show individuals included in
1139 downstream analyses for regional CI staining. Box plots show individual values (dots), median
1140 and interquartile range (box) and minimum-maximum range (whiskers). Statistical analysis was
1141 not performed in the PFC, since this region was used to classify the samples, and the samples
1142 were selected to represent the full spectrum of CI status for each group. Nominal p-values of
1143 two-sided tests are shown. Multiple testing correction was not performed for these analysis, due
1144 to the relatively small sample size and high correlation between the CI state of different
1145 anatomical regions per individual, as illustrated in the heatmap (c). All individuals with CI-PD
1146 goup are shown in the same green box, but differentiated by the color of the points into mild
1147 (dark green) and severe (light green) CI-PD. (b) Structural brain maps illustrating the
1148 anatomical distribution of neuronal CI deficiency in the control, nCI-PD and CI-PD groups.
1149 The median proportion of CI positive neurons is shown as a heatmap. Brain illustrations are
1150 modified images from the 7MRI atlas published by Edlow et al[93]. (c) Hierarchical clustering of
1151 the median proportion of CI positive neurons in each group. SNpc: substantia nigra pars
1152 compacta, PFC: prefrontal cortex, OC: occipital cortex, TC: temporal cortex, Sub: subiculum,
1153 Am: amygdala, CG: cingulate gyrus, Put: putamen, GPe: external globus pallidus, GPi: internal
1154 globus pallidus internus, ION: inferior olivary nucleus, CC: cerebellar cortex, DN: dentate

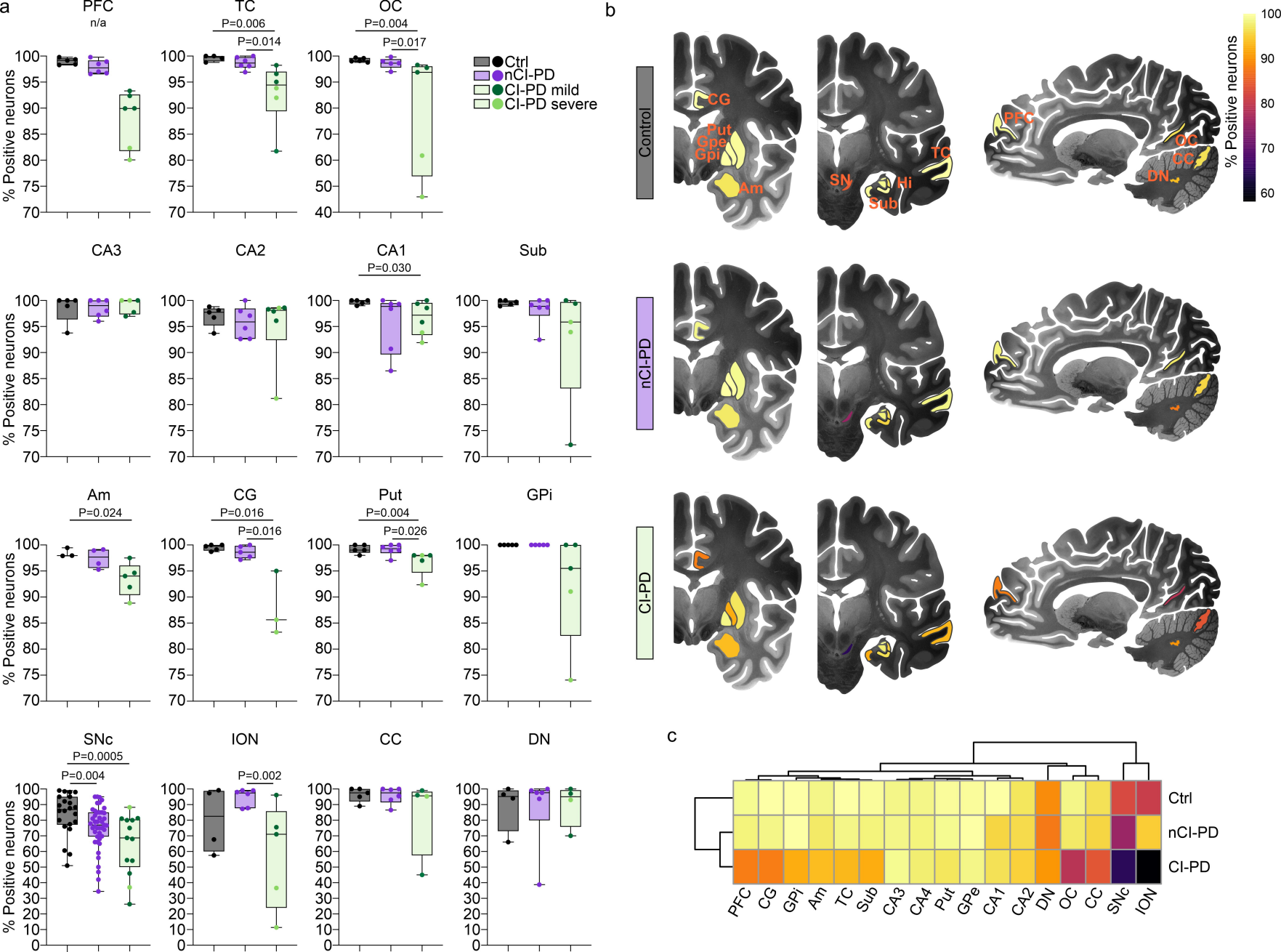
1155 nucleus, Hi: hippocampal regions CA4-CA1, Ctrl: controls. Statistical test: Mann-Whitney U-
1156 test, n/a: statistical comparison not performed.
1157
1158 **Fig. 4. Clinical differences between CI-PD and nCI-PD**
1159 Box plots show individual values (dots), median and interquartile range (box) and minimum-
1160 maximum range (whiskers) of clinical data (y-axes) for the three iPD groups (nCI-PD: purple,
1161 CI-PD mild: dark green, CI-PD severe: light green). Statistical significance was tested between
1162 nCI-PD and the entire CI-PD group (i.e., not differentiating between mild and severe), due to
1163 the low number of individuals in the severe group, using the two-sided Mann-Whitney U-test.
1164 Bar plots show percentages of individuals within each iPD group (y-axis). Statistical
1165 significance was tested using the X^2 test. Demographic data, data on dementia at time of death,
1166 and the AR/TD-clinical phenotypes were available from both cohorts. All other clinical
1167 variables were available only from the NOR cohort. The number of included individuals is not
1168 equal in the different comparisons and is shown in Supplementary Data 5. AR: akinetic rigid,
1169 TD: tremor dominant, PIGD: postural instability and gait disorder, UPDRS: Unified
1170 Parkinson's Disease Rating Scale.
1171
1172 **Fig. 5. The nCI-PD and CI-PD groups show different neuronal mtDNA profiles**
1173 Single neurons were collected from frozen PFC sections using laser-microdissection and
1174 mtDNA was analysed for copy number (CN) and deletion proportion using qPCR. Eight to
1175 sixteen single neurons were assessed per individual from individuals with CI-PD (n = 6), nCI-
1176 PD (n = 9) and controls (n = 7). (a) Microphotograph of a representative PFC section, stained
1177 with cresyl violet and showing neurons before (#) and after (*) microdissection. Scale bar 10
1178 µm. (b) Total neuronal mtDNA copy number. (c) Neuronal mtDNA deletion levels. (d) Non-
1179 deleted neuronal mtDNA levels. Box plots show median and interquartile range (box) and

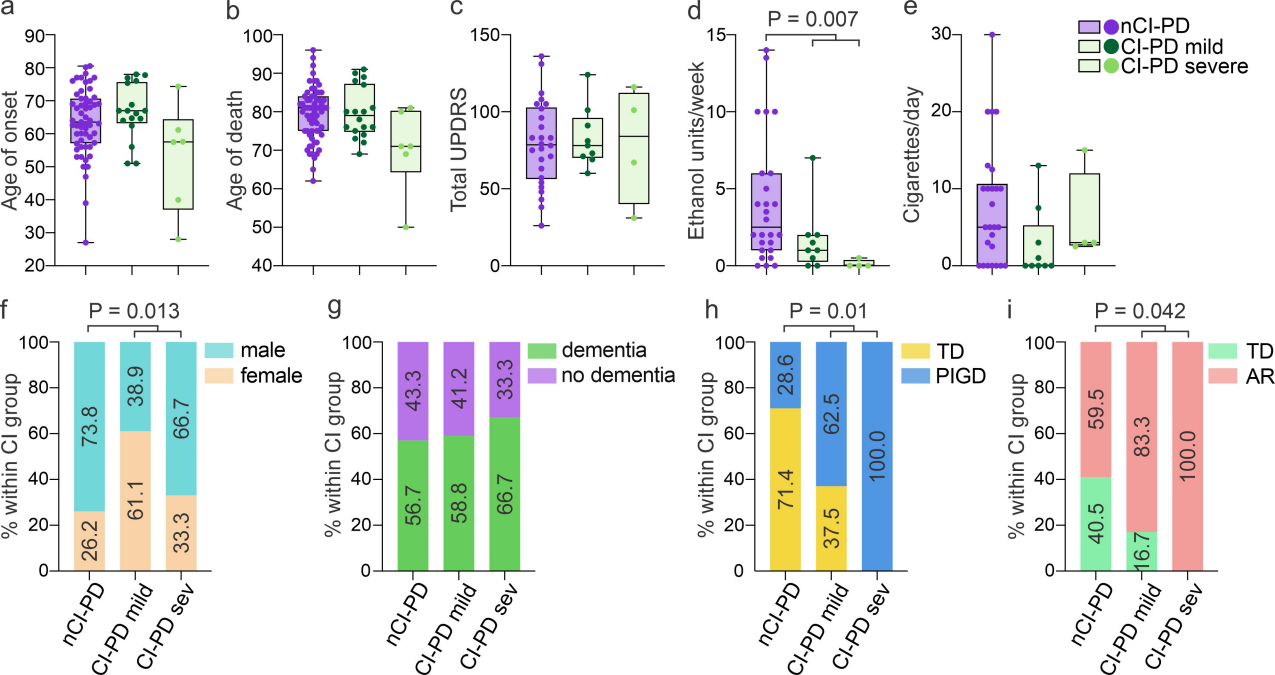
1180 minimum-maximum range (whiskers) of the mean values for each individual (points) per
1181 group. Statistical significance of group difference was tested using two-sided
1182 ANOVA. CN: copy number, Ctrl: control.
1183
1184 **Fig. 6. nCI-PD and CI-PD groups exhibit distinct transcriptomic profiles in the PFC**
1185 (a) Hierarchical clustering of the samples based on sample-sample correlation. The heatmap
1186 colors indicate Pearson's correlation values between each pair of samples. AGE: age of death.
1187 (b) The estimated group difference (y-axis) in MGPs of the cell types (x-axis) based on linear
1188 regression using the three different models are shown. The cell types are described in Mancarci
1189 et al [56]. The estimate and the 95%CI are indicated by the point and the whiskers, respectively
1190 (c) Significance (shown as -log10 of the adjusted p-value) of genes included in the KEGG
1191 OXIDATIVE_PHOSPHORYLATION pathways (blue) or all other genes (grey) among the
1192 downregulated/upregulated genes based on a model not adjusting (Model_2) or adjusting
1193 (Model_1) for the confounded MGPs, in each of the iPD subtypes. Dashed line indicates
1194 adjusted pvalue of 0.05. Boxplots represent the interquartile range (IQR) with the median
1195 indicated by the bold line. Whiskers are extensions from the top and the bottom of the boxplot
1196 by a value of 1.5*IQR. Values outside of the whisker range are indicated by circles (d)
1197 Enrichment analysis outcome of the nCI-PD group, based on DE analysis using Model_2. Each
1198 leaf of the tree represents one significantly enriched gene-set. Jaccard similarity was used to
1199 cluster the gene-sets based on semantic similarity. Each cluster was annotated to represent the
1200 main function represented by its members. The names of the gene-sets are shown in
1201 Supplementary Fig. S7 (downregulated) and Supplementary Fig. S8 (upregulated).
1202
1203 **Fig. 7. Single-nuclei RNA-seq shows distinct cell type-specific transcriptional profiles for**
1204 **CI-PD and nCI-PD**

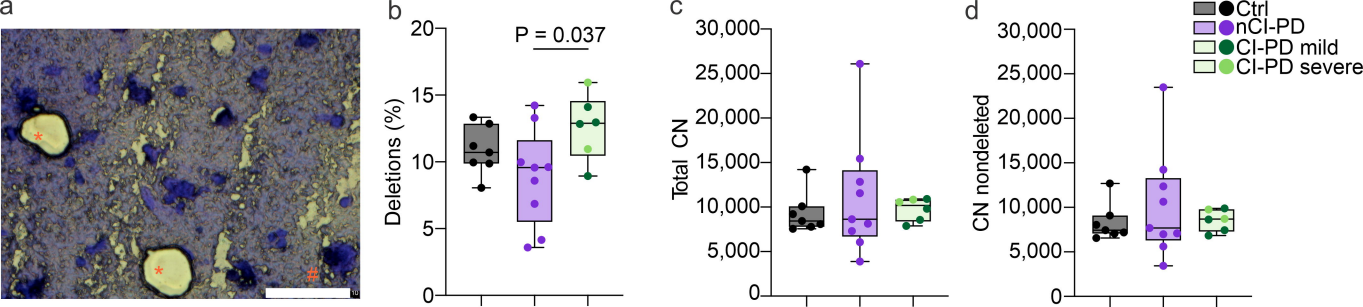
1205 (a) UMAP plot of the filtered snRNA-seq dataset (117,105 nuclei) showing the clustering in
1206 the main cortical cell types (ex: excitatory neurons, in: inhibitory neurons, oligodendrocytes,
1207 astrocytes, microglia, oligodendrocyte precursor cells;OPC, and endothelial cells). (b) Dotplot
1208 showing the average cell type markers' expression in each of the 28 subclusters identified by
1209 Seurat, unequivocally mapping the nuclei to the main cortical known cell types. The color of
1210 the dots indicates average scaled expression, the size representing the proportion of nuclei in
1211 the cluster that express the marker. (c) Summary of the number of differentially expressed genes
1212 in autosomes (at FDR < 5%) within each cluster for the CI-PD *vs* controls (left hand side panel,
1213 green bars), for the nCI-PD *vs* controls (center panel, purple bars), and common between the
1214 two contrasts (right hand side panel, grey bars).

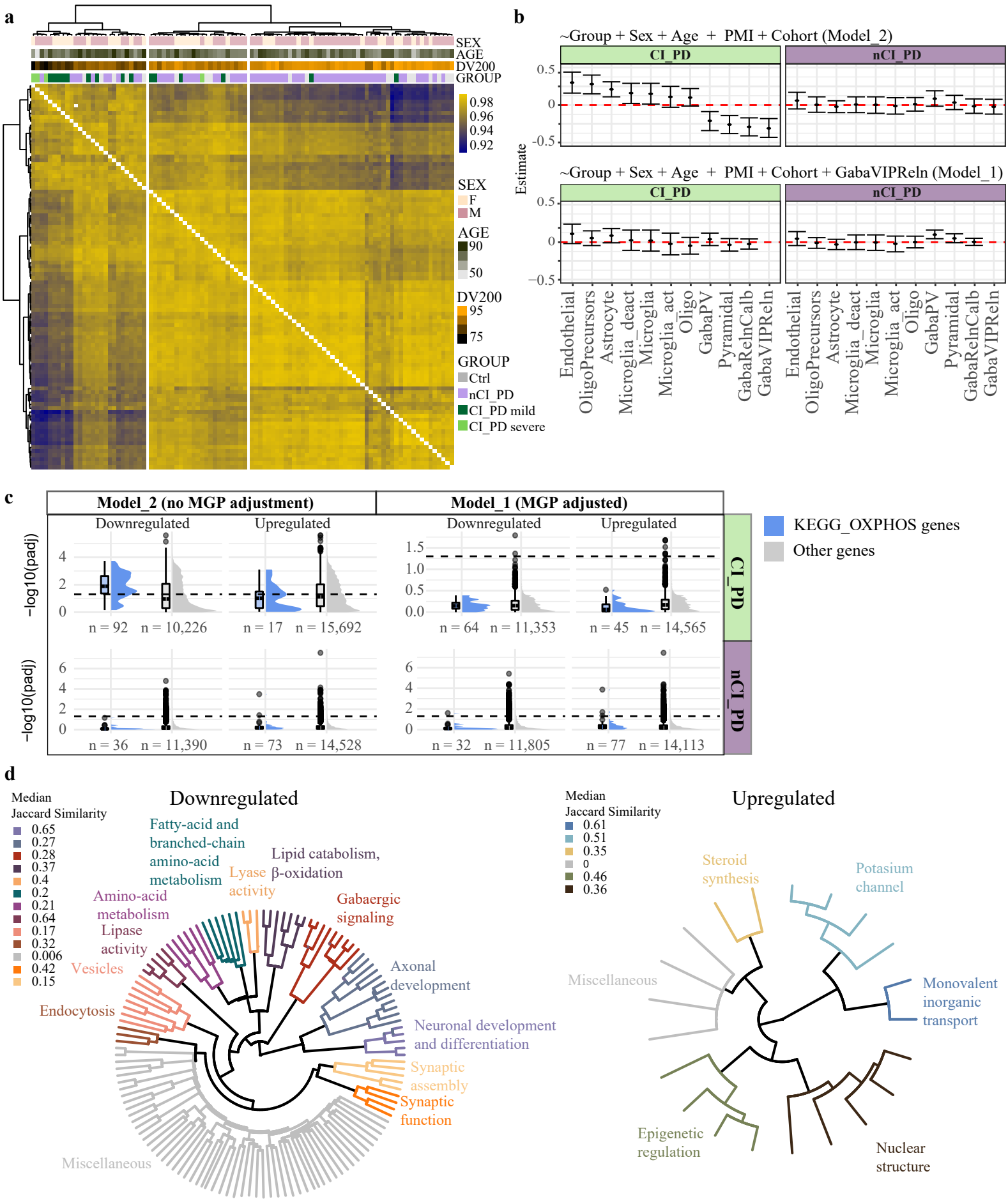

a
b
c
Difference (O1-O2)
Average O1 and O2
d
Difference (R1-R2)
Average R1 and R2
Difference (R1-R3)
Average R1 and R3
e
NOR
P = 0.00082
ESP
NOR & ESP
P = 0.003
% Positive neurons SNpc
Ctrl
iPD
f
NOR
P = 0.0007
ESP
P = 0.031
NOR & ESP
P = 0.00005
% Positive neurons PFC
g
Ctrl
nCI-PD
CI-PD mild
CI-PD severe

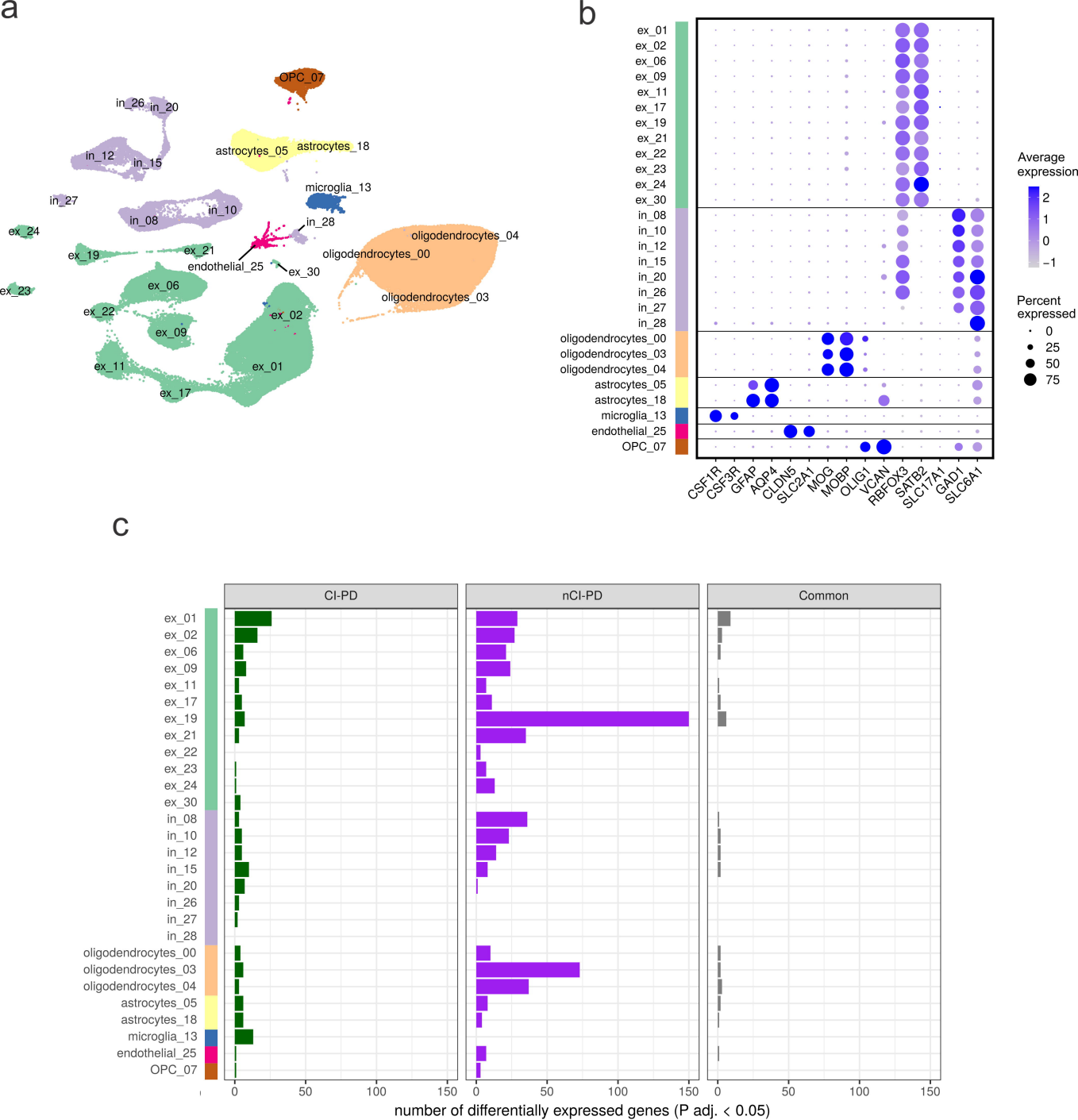

a
SNpc
b
Prefrontal cortex
c
Occipital cortex
NDUFS4
VDAC1
d
Cingulate gyrus
e
Putamen
f
Amygdala
NDUFS4
VDAC1
g
Temporal cortex
h
CA1
i
Inferior olivary nucleus
NDUFS4
VDAC1


a
PFC
n/a
TC
P=0.006
P=0.014
OC
P=0.004
P=0.017
Ctrl
nCI-PD
CI-PD mild
CI-PD severe
% Positive neurons
CA3
CA2
CA1
P=0.030
Sub
Am
P=0.024
CG
P=0.016
P=0.016
Put
P=0.004
P=0.026
GPi
SNc
P=0.0005
P=0.004
ION
P=0.002
CC
DN
b
Control
nCI-PD
CI-PD
CG
Put
Gpe
Gpi
Am
SN
Hi
Sub
TC
PFC
OC
CC
DN
% Positive neurons
c
Ctrl
nCI-PD
CI-PD
PFC
CG
GPi
Am
TC
Sub
CA3
CA4
Put
GPe
CA1
CA2
DN
OC
CC
SNc
ION


a
Age of onset
b
Age of death
c
Total UPDRS
d
P = 0.007
Ethanol units/week
e
Cigarettes/day
nCI-PD
CI-PD mild
CI-PD severe
f
P = 0.013
male
female
% within CI group
73.8
26.2
38.9
61.1
66.7
33.3
nCI-PD
CI-PD mild
CI-PD sev
g
dementia
no dementia
43.3
56.7
41.2
58.8
33.3
66.7
h
P = 0.01
TD
PIGD
28.6
71.4
62.5
37.5
100.0
i
P = 0.042
TD
AR
59.5
40.5
83.3
16.7
100.0


a
b
P = 0.037
Deletions (%)
20
15
10
5
0
c
Total CN
30,000
25,000
20,000
15,000
10,000
5,000
0
d
CN nondeleted
30,000
25,000
20,000
15,000
10,000
5,000
0
Ctrl
nCI-PD
CI-PD mild
CI-PD severe


a
SEX
AGE
DV200
GROUP
0.98
0.96
0.94
0.92
SEX
F
M
AGE
90
50
DV200
95
75
GROUP
Ctrl
nCI_PD
CI_PD mild
CI_PD severe
b
~Group + Sex + Age + PMI + Cohort (Model_2)
CI_PD
nCI_PD
Estimate
~Group + Sex + Age + PMI + Cohort + GabaVIPReln (Model_1)
Endothelial
OligoPrecursors
Astrocyte
Microglia_deact
Microglia
Microglia_act
Oligo
GabaPV
Pyramidal
GabaRelnCalb
GabaVIPReln
c
Model_2 (no MGP adjustment)
Model_1 (MGP adjusted)
Downregulated
Upregulated
−log10(padj)
n = 92
n = 10,226
n = 17
n = 15,692
n = 64
n = 11,353
n = 45
n = 14,565
n = 36
n = 11,390
n = 73
n = 14,528
n = 32
n = 11,805
n = 77
n = 14,113
KEGG_OXPHOS genes
Other genes
d
Median Jaccard Similarity
0.65
0.27
0.28
0.37
0.4
0.2
0.21
0.64
0.17
0.32
0.006
0.42
0.15
Downregulated
Fatty-acid and branched-chain amino-acid metabolism
Lipid catabolism, β-oxidation
Lyase activity
Amino-acid metabolism
Gabaergic signaling
Lipase activity
Vesicles
Axonal development
Endocytosis
Neuronal development and differentiation
Synaptic assembly
Synaptic function
Miscellaneous
Median Jaccard Similarity
0.61
0.51
0.35
0
0.46
0.36
Upregulated
Steroid synthesis
Potasium channel
Miscellaneous
Monovalent inorganic transport
Epigenetic regulation
Nuclear structure


a
OPC_07
in_26
in_20
in_12
in_15
astrocytes_05
astrocytes_18
microglia_13
in_27
in_10
in_08
in_28
ex_24
oligodendrocytes_04
ex_19
ex_21
oligodendrocytes_00
endothelial_25
ex_30
ex_23
ex_06
oligodendrocytes_03
ex_22
ex_02
ex_09
ex_11
ex_01
ex_17
b
Average expression
2
1
0
−1
Percent expressed
0
25
50
75
CSF1R
CSF3R
GFAP
AQP4
CLDN5
SLC2A1
MOG
MOBP
OLIG1
VCAN
RBFOX3
SATB2
SLC17A1
GAD1
SLC6A1
c
CI-PD
nCI-PD
Common
number of differentially expressed genes (P adj. < 0.05)