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
E-mail-2: charalampos.tzoulis@helse-bergen.no

Word count main text: 7252

Running title: Complex I stratifies Parkinson's disease

Number of figures: 7

Number of tables: 2

Number of supplementary files: 10 figures, 10 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. 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

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

84 Major biological processes which have been associated with iPD include mitochondrial
85 dysfunction, impaired proteostasis, and neuroinflammation[1,3,4]. Mitochondrial dysfunction is
86 considered a key feature of iPD, proposed to be integral to its pathogenesis and
87 pathophysiology[11]. Evidence of impaired mitochondrial DNA (mtDNA) maintenance, and
88 respiratory complex I (CI) deficiency has been consistently reported in the dopaminergic SNpc
89 of individuals with iPD[11–13]. While the causes of these mitochondrial defects are unknown, it is
90 hypothesized that they contribute to the neurodegenerative process, based on the facts that
91 chemical CI inhibition[14–16], genetic CI defects[17,18], and mutations in genes regulating mtDNA
92 homeostasis[19,20] can cause degeneration of the dopaminergic SNpc and other neuronal
93 populations in humans. In addition to the SNpc, CI deficiency has been reported in neurons
94 across multiple brain regions, as well as in extra-neural tissues, including muscle and blood[12,21].
95 However, reported findings are variable and contradictory, with nearly half the studies showing
96 no significant difference between individuals with PD and controls at the group level[12]. While
97 this variability may be in part methodological in nature, it also raises a valid question regarding
98 the pervasiveness of CI deficiency in iPD. In line with this notion, studies in skin fibroblasts
99 from individuals with iPD have identified evidence of mitochondrial dysfunction in only a
100 minority of the subjects[22,23].

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.

106 ## Results

107 **Neuronal CI deficiency exhibits higher specificity for iPD in the prefrontal cortex than in**
108 **the SNpc**

109 To select a robust marker for neuronal CI immunodetection, we tested antibodies against
110 different CI subunits across the functional domains of the complex[24] in the SNpc of 5
111 individuals with iPD and 3 controls. All tested subunits exhibited a similar pattern and severity
112 of neuronal CI deficiency, consistent with a quantitative reduction of the entire CI in affected
113 cells (Supplementary Fig. S1). Subunit NDUFS4 was chosen for downstream analyses due to
114 superior staining quality. Neuronal CI status was subsequently assessed, along with the total
115 mitochondrial mass marker VDAC1, in the dopaminergic SNpc and prefrontal cortex (PFC) of
116 two independent cohorts from Norway (NOR, n = 54) and Spain (ESP, n = 62). The
117 demographic and clinicopathological information of the cohorts, and experimental allocation
118 of the samples are shown in Table 1 and Supplementary Tables S1 and S2.

119 Neurons were classified as either negative or positive (Fig. 1a-b) by two independent
120 investigators, and the percentage of positive neurons per region and individual was used as a
121 measure of CI status. There was excellent inter-investigator reliability (ICC coefficient = 0.91,
122 95% C.I (0.87,0.94), $P = 1.9x10^{-29}$) (Fig 1c-d, Supplementary Table S3). At the group level, the
123 SNpc exhibited pronounced CI deficiency in both iPD and control subjects, with considerable
124 overlap between the groups. iPD had a significantly lower proportion of CI positive
125 dopaminergic neurons in the combined cohorts (P = 0.003) and the NOR cohort ($P = 8.2x10^{-4}$),
126 but not in the ESP cohort (P = 0.26) (Fig. 1e, Table 2, Supplementary Table S3). In the PFC,
127 controls showed no or mild CI deficiency, whereas high variation was observed in the iPD
128 group. A significantly lower proportion of CI positive neurons among individuals with iPD was
129 observed in each cohort individually (NOR: $P = 7.0x10^{-4}$, ESP: P = 0.031) and in the combined
130 cohorts ($P = 5.0x10^{-5}$; Fig. 1f, Table 2, Supplementary Table S3). The neuronal CI status in the

131 PFC showed a significant association with disease state (i.e., iPD vs ctr), whereas that of the
132 SNpc did not (linear regression, PFC: P = 0.006, SNpc: P = 0.06). CI deficiency in the PFC was
133 not associated with post-mortem interval (PMI) (ρ(109) = -0.08, P = 0.42), age of death (ρ(110)
134 = 0.03, P = 0.73), or disease duration (ρ(79) = 0.06, P = 0.60) (Supplementary Table S1). The
135 VDAC1 staining showed no deficient neurons in iPD or controls of either cohort (Fig. 2a-b,
136 Supplementary Table S3).
137
138 **Neuronal CI deficiency stratifies iPD**
139 To assess the pervasiveness of CI deficiency in iPD, individuals were stratified based on the
140 proportion of CI positive neurons in the PFC. This region was chosen over the SNpc for two
141 reasons. First, neuronal CI deficiency was found to be more specific to PD in the PFC than in
142 the SNpc (Fig. 1e-f). This is in line with previous findings, showing that dopaminergic SNpc
143 neurons can exhibit severe CI deficiency and other markers of mitochondrial dysfunction with
144 healthy aging, making it challenging to distinguish between disease-specific and age-related
145 changes[25–27]. Second, the severe neuronal loss in the SNpc of iPD[28] complicates the comparison
146 between cases and controls. Specifically, surviving SNpc neurons are likely to be resilient cells,
147 introducing survival bias.
148 Individuals with iPD were classified by k-means clustering of the entire study
149 population (cases and controls). The number of clusters was set to three based on visual
150 inspection of the data. Of the three emerging iPD subgroups, two clustered away from the
151 controls and were termed mild (n = 18/89, 20.2%) and severe (n = 6/89, 6.7%) CI deficient PD
152 (CI-PD mild; CI-PD severe), whereas one (n = 65/89, 73.0%) clustered with the controls and
153 was termed non-CI deficient PD (nCI-PD) (Fig. 1g). A similar clustering pattern was observed
154 in each of the NOR and ESP cohorts individually, where the combined CI-PD clusters
155 accounted for 33% and 22% of the iPD cases, respectively (Fig. 1h, Supplementary Table S3).

156 Since the severe CI-PD cluster only contained six samples, the mild and severe clusters were
157 combined into a single CI-PD group for downstream statistical analyses.
158
159 **CI-PD is characterized by anatomically widespread CI deficiency**
160 To determine the anatomical distribution of neuronal CI deficiency in the iPD subgroups, we
161 studied 15 additional brain regions in representative individuals with CI-PD, nCI-PD and
162 controls from the NOR cohort. All individuals with iPD had Braak scores of 5-6 and the median
163 disease duration was 12 years in both subgroups (Supplementary Tables S1 and S2). Subjects
164 in the CI-PD group showed an anatomically widespread neuronal CI-deficiency compared to
165 controls, including the temporal cortex, occipital cortex, CA1 region of the hippocampus,
166 amygdala, putamen, and the cingulate gyrus. Lower proportions of CI intact neurons were also
167 seen in the internal globus pallidus and the subiculum, but these did not reach statistical
168 significance. No difference from controls was observed in CA4-2 of the hippocampus, the
169 external globus pallidus, inferior olivary nucleus, cerebellar Purkinje cells and the dentate
170 nucleus. The nCI-PD group was similar to controls, with the single exception of the SNpc. The
171 difference in SNpc was, however, less pronounced compared to the CI-PD group. Finally, we
172 noted that, similar to the SNpc, the inferior olivary nucleus, Purkinje neurons and dentate
173 nucleus showed high proportions of CI negative neurons in several of the controls, suggesting
174 that these neuronal populations may be prone to age-dependent CI loss (Fig. 2, Fig. 3,
175 Supplementary Tables S3 and S4).
176 To assess the state of the other mitochondrial respiratory chain (MRC) complexes in the
177 iPD groups, we performed IHC for selected subunits of complexes II-V (CII-V) and VDAC1
178 in 12 brain regions from two to four subjects from each of the CI-PD and nCI-PD groups
179 (Supplementary Table S2). Irrespective of group, individuals with iPD exhibited no or few
180 neurons staining negative for CIII, CIV and CV (generally ≤5%). No negative neurons were

observed in sections stained against VDAC1 or CII (Supplementary Table S3, Supplementary Fig. S2).

CI-PD is associated with a non-TD phenotype

The CI-PD and nCI-PD groups were similar in terms of age of first reported symptom, age of death, disease duration, and total UPDRS score at the final examination before death. Individuals with severe CI-PD tended to have an earlier age of onset and age of death, but the small size of that group did not allow any conclusions to be drawn (Fig. 4a-c, Supplementary Table S5). The groups differed significantly in the distribution of motor phenotype classified as AR/TD[9] (available from both cohorts), or PIGD/TD[10] (available from the NOR cohort only), at the time of diagnosis. In the CI-PD group, the AR and PIGD phenotypes accounted for 88% (14/16) and 75% (9/12) of the individuals, respectively. In the nCI-PD group, the AR or PIGD phenotypes were significantly less prevalent, accounting for 60% (25/42) and 29% (6/21) of the individuals, respectively. Notably, individuals with severe CI-PD exhibited exclusively AR or PIGD phenotype (Fig. 4h-i). Furthermore, the CI-PD group exhibited a significantly higher proportion of females (CI-PD: 54.2%, nCI-PD: 26.2%, P = 0.013), compared to the nCI-PD group, and a significantly lower level of ethanol consumption than the nCI-PD group (Fig. 4e-f). There was no significant difference between the groups in the prevalence of dementia at the final examination before death (Fig. 4g). Following multiple testing correction (Benjamini Hochberg), the difference in the distribution of the PIGD/TD phenotype (P = 0.04), sex (P = 0.04), and number of weekly alcoholic units (P = 0.04) remained significant. However, since the tested variables were highly correlated with each other and the sample size was relatively small, the risk of type 2 error by correcting for multiplicity, is higher than the risk of type 1 error by not correcting[29]. Individual clinical and demographic features are given in Supplementary Tables S1 and S5.

206 In terms of neuropathology, there were no significant between-group differences in either Thal
207 phase staging of beta amyloid, Braak staging of neurofibrillary tangles, or Braak staging of LP.
208 Moreover, there was no difference in the load of LP in the PFC (available from the NOR cohort
209 only) between the groups (Supplementary Fig. S3).
210
211 **CI-PD and nCI-PD show different neuronal mtDNA profiles**
212 Neuronal mtDNA copy number and deletion burden were assessed by qPCR in single neurons
213 microdissected from the PFC (Fig. 5a-b) of individuals with CI-PD (n = 6), nCI-PD (n = 9) and
214 controls (n = 7; Supplementary Table S2). Eight to sixteen single neurons were assessed per
215 individual. There was a significant difference in neuronal mtDNA deletions between the groups
216 (ANOVA, P = 0.037, η^2 = 0.30). Post-hoc analyses showed that the CI-PD group had
217 significantly higher levels of mtDNA deletions compared to the nCI-PD group (P = 0.037; Fig.
218 5b). No differences were observed in neuronal mtDNA copy number (total or non-deleted)
219 between the groups (Fig. 5c-d). The mtDNA data is shown in Supplementary Table S6.
220
221 **CI-PD and nCI-PD exhibit distinct transcriptomic profiles**
222 Hierarchical clustering of RNA-seq data from the PFC of 98 of the individuals previously
223 assessed for CI deficiency (Supplementary Table S2) indicated that individuals with CI-PD
224 tended to cluster together (Fig. 6a), with a similar behavior observed in each of the cohorts
225 (Supplementary Fig. S4). The CI-PD group was associated with lower DV200 values, a
226 measure of RNA integrity indicating the percentage of RNA fragments with 200 nucleotides or
227 more, compared to both nCI-PD and controls (Fig. 6a). However, this association could not be
228 explained by PMI, since the PMI range of the CI-PD group was similar to that of the controls
229 within each cohort (Supplementary Fig. S5).

230 To control for between-group differences in the cell composition of the samples, we
231 assessed Marker Gene Profiles (MGPs)[30] of different cell types in each iPD group and the
232 controls. Linear regression adjusting for age of death, sex, PMI and cohort, indicated a
233 significant change in the MGP of multiple cell types in the CI-PD group compared to the
234 controls. However, step-wise regression analysis suggested that the observed changes were not
235 independent, but rather driven by a decrease in the MGP of GABAergic interneurons expressing
236 both vasoactive intestinal peptide-expressing (VIP) and reelin (GabaVIPReln) (Fig. 6b,
237 Supplementary Fig. S6). No difference was observed in the MGPs of the nCI-PD group
238 compared to the controls (Fig. 6b, Supplementary Fig. S6).

239 To account for the effects of different cell composition on the differential gene
240 expression (DE) profile of CI-PD, we performed DE analysis using two alternative models: 1)
241 Correcting for the confounding MGP of GabaVIPReln (Model_1) and 2) not correcting for
242 MGPs (Model_2). Both models included the demographic variables and oligodendrocyte
243 MGPs to adjust for any variation in the grey/white matter content between the samples[31,32].
244 Adjusting for GabaVIPReln MGPs resulted in a striking reduction of the number of
245 differentially expressed genes in the CI-PD group, from 11,138 to 10 (Supplementary Table
246 S7). Notably, while oxidative phosphorylation (OXPHOS)-related genes were among the top
247 differentially downregulated genes in the non-adjusted model, none of these genes remained
248 significant after adjusting for cell composition (Fig 6c). In contrast, MGP correction had no
249 major impact on the DE analysis of the nCI-PD group (Fig 6c).

250

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

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

255 to determine whether the results represented true regulatory changes or simply altered cell
256 composition. We therefore chose not to further focus on these results, which are given in
257 Supplementary Table S8. The transcriptomic signature of nCI-PD, which was not confounded
258 by estimated cell composition, was characterized by changes in numerous pathways, including
259 the downregulation of processes related to proteostasis (lysosome and proteasome),
260 GABAergic signaling, synaptic function and lipid metabolism. Upregulated processes included
261 pathways related to epigenomic regulation, steroid synthesis, and potassium channels. Notably,
262 there were no changes in genes or pathways related to mitochondrial respiration (Fig. 6d,
263 Supplementary Fig. S7 and S8, Supplementary Table S8).
264
265 **Single-nuclei RNA sequencing reveals distinct cell type-specific profiles between CI-PD**
266 **and nCI-PD**
267 To mitigate the bias of cell composition indicated by the bulk RNA-seq data and assess cell
268 type-specific transcriptomic similarities between CI-PD and nCI-PD, we performed massively
269 parallel single-nuclei RNA-seq (snRNA-seq) in fresh-frozen PFC from n = 7 individuals with
270 CI-PD, n = 5 individuals with nCI-PD, and n = 6 controls (Supplementary Table S2). After
271 filtering and quality control, a total of 117,105 nuclei (mean: 6,506±2,952 nuclei per individual)
272 were included in the analyses. There was no significant difference in the number of captured
273 nuclei between groups (ANOVA P = 0.051).
274 Default clustering using Seurat identified a total of 28 clusters that could be unequivocally
275 assigned to a major cortical cell type, based on established cell type markers (12 excitatory
276 neuron clusters, 8 inhibitory neuron clusters, 3 oligodendrocyte clusters, 2 astrocyte clusters, 1
277 microglia cluster, 1 oligodendrocyte precursor cells (OPC) cluster, and 1 endothelial cluster,
278 Figure 7A-B). Unique molecular identifier (UMI) counts and number of identified genes per

279 nucleus were not significantly different between groups in any of the identified cell type
280 clusters.
281 Differential gene expression analysis comparing each of the iPD groups to the controls revealed
282 distinct cell type-specific transcriptomic profiles between CI-PD and nCI-PD. After FDR
283 correction, we found a total of 189 significantly differentially expressed genes in CI-PD and
284 569 significantly differentially expressed genes in nCI-PD, distributed among several cell type
285 clusters, with 66 genes overlapping between the groups (Figure 7C). CI-PD showed differential
286 expression in most excitatory and inhibitory neuronal subtypes, and a notable upregulation in
287 microglia, including *PRKN* – encoding parkin, a key component of mitophagy – in which
288 mutations cause autosomal recessive PD[33]. nCI-PD exhibited multiple differentially expressed
289 genes, with most of the signal observed in neurons and oligodendrocytes, while no differential
290 expression was seen in microglia. The differentially expressed genes per cell type and group
291 are shown in Supplementary Tables S9 and S10. A visual map of significantly differentially
292 expressed genes (FDR < 5%) with fold change ≥ 25%, in at least one cell type, is shown in
293 Supplementary Fig. S10.

294 **Discussion**

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

306 It is likely that CI deficiency contributes to neuronal dysfunction and death in CI-PD.
307 Neurons depend on CI function to maintain a high ATP turnover[34]. Brain mitochondria exhibit
308 impaired energy metabolism with major decreases in ATP production upon 60% reduction of
309 CI activity, and this threshold is much lower (~25%) for synaptic mitochondria[35]. While
310 functional assessment of CI activity is currently not feasible in individual neurons from
311 postmortem brain, it is reasonable to assume that neurons with no detectable CI staining (i.e.,
312 CI negative neurons) must have severely decreased CI activity and, therefore, a compromised
313 bioenergetic status. This is corroborated by the fact that isolated CI deficiency of either
314 chemical[14,16] or genetic[17,36] etiology is an established cause of neuronal dysfunction and death
315 in humans and other animals.

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

319 specific neurodegeneration outside the SNpc in these individuals. CI deficiency may still
320 contribute to the degeneration of the dopaminergic SNpc neurons in nCI-PD. However, given
321 that this deficiency was significantly milder compared to CI-PD, and showed substantial
322 overlap with the aged-matched controls, its disease-specific pathophysiological significance
323 remains uncertain.
324 The CI-PD and nCI-PD groups differed in terms of motor disease subtype at the time of
325 diagnosis. Our data indicate that CI dysfunction is significantly more common in the non-TD
326 form of iPD, which has also been shown to harbor more severe neurodegeneration[37]. The
327 mechanism underlying this association is currently unknown but may be related to more
328 widespread and severe neuronal dysfunction and loss in CI-PD, induced by the CI deficiency.
329 This is corroborated by the bulk RNAseq data indicating more neuronal loss in CI-PD. Until
330 clinically applicable biomarkers of CI-PD and nCI-PD can be developed, enabling us to classify
331 individuals in vivo, we advocate that clinical iPD studies would benefit by distinguishing
332 between TD and non-TD phenotypes. Since individuals with a non-TD phenotype at the time
333 of diagnosis are substantially more likely to belong to the CI-PD subtype, they may be more
334 likely to benefit from experimental treatments targeting mitochondria. While there are currently
335 no drugs specifically targeting CI, several agents targeting mitochondrial biogenesis and/or
336 function have been tested in clinical trials with negative results[3], while others are still in the
337 pipeline[38]. In light of our findings, it would be of interest to revisit the data of previously
338 conducted trials, this time with a stratified approach, and assess how the treatment impacted
339 individuals with a non-TD or TD phenotype. Moreover, stratifying future clinical trials by
340 motor phenotype may improve the chances of success by addressing biological heterogeneity
341 in the cohort. In addition to the motor phenotype, we observed an association of CI-PD with
342 the female sex and a lower consumption of ethanol. The latter may be driven by the female
343 preponderance in this subtype.

344 Neuronal mtDNA profiles in the PFC corroborated the stratification of iPD into CI-PD
345 and nCI-PD groups. mtDNA deletions are an established cause of MRC deficiencies in post-
346 mitotic cells[39–41]. However, given the small effect size of the mtDNA deletion increase in the
347 CI-PD group, and the fact that neuronal deletion levels were generally below 60%, i.e., the
348 threshold at which MRC dysfunction has been proposed to occur[39], the biological significance
349 of the detected difference in mtDNA deletion levels remains uncertain. It should be noted that,
350 since most PFC neurons of individuals with CI-PD remain CI positive, it is possible that few or
351 none of the captured neurons in our experiments were CI negative. Thus, the possibility that
352 CI-negative neurons harbor higher deletion levels, driving the CI deficiency, cannot be ruled
353 out by this work. An alternative explanation is that mtDNA deletions in CI-PD individuals may
354 be the result, rather than the cause, of CI deficiency. Under this hypothesis, mtDNA deletions
355 may arise during the repair of double-strand breaks, caused by increased reactive oxygen
356 species (ROS) generation in CI deficient neurons[42,43].

357 RNA-seq revealed distinct transcriptomic profiles between the CI-PD and nCI-PD
358 groups. In bulk tissue data, the transcriptomic signature of CI-PD was severely confounded
359 with cell-type-specific expression profiles (i.e., MGPs), indicating differences in the cell type
360 composition of the samples. The observation that CI-PD exhibited decreased estimates for
361 GABAergic interneurons is consistent with a mitochondrial defect. GABAergic interneurons
362 have high bioenergetic demands, due to their fast-spiking firing pattern, and have been shown
363 to be selectively lost in mitochondrial disease characterized by neuronal CI deficiency disease
364 [44]. While MGPs are not direct cell counts, they are indicative of a simultaneous shift in multiple
365 cell-type specific genes. Regardless of whether this reflects a difference in the underlying
366 cellular composition of the tissue, synaptic loss, or phenotypic change[36], this association
367 indicates that any regulatory changes cannot be disentangled from differences due to altered
368 tissue composition[31,45]. Thus, no further conclusions should be drawn from transcriptomic

369 analyses of bulk brain tissue data from CI-PD individuals. In contrast, the nCI-PD group
370 exhibited no difference in MGPs compared to controls. Thus, we believe that the DE signature
371 of this group reflects regulatory changes associated with the disease. The transcriptional profile
372 of nCI-PD showed altered expression in multiple pathways, many of which have been
373 previously linked to PD[46–48].

374 snRNA-seq allowed us to compare the iPD subgroups within each cell-type, thus
375 circumventing the confounder of different cell composition. These analyses revealed a highly
376 distinct cell type-specific transcriptomic profile for each of CI-PD and nCI-PD, supporting that
377 the two subtypes harbor distinct molecular processes. While single-nucleus transcriptomics are
378 not sufficient to adequately characterize these processes, since they only captures a small part
379 of the cell's transcriptional state, the data provides us with some hints. Neurons and astrocytes
380 were affected in both groups, albeit by largely different genes. However, microglia and
381 oligodendrocytes showed subgroup-specific involvement. The CI-PD-specific changes in
382 microglia are suggestive of mitophagy activation, which indicates the presence of
383 mitochondrial dysfunction and stress in these cells. In contrast, nCI-PD showed no changes in
384 microglia, but harbored substantial transcriptional alterations in oligodendrocytes.
385 Oligodendrocyte involvement has been previously reported in PD[49]. Our findings suggest that
386 this may be specific to nCI-PD.

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

392 Functional assessment of specific CI activity was not undertaken in this work, because
393 it is currently not feasible to assess in individual cells. Measurement of CI activity in bulk tissue

394 would be unreliable because of both low sensitivity due to the mosaic distribution of the CI
395 defect, and low specificity in the sense that it cannot adequately distinguish between cell-
396 specific reduction in CI activity or a decrease in neuronal and synaptic content in the sample.
397 We have previously shown that the PFC of iPD exhibits markers of decreased neuronal
398 content[31], and our current data suggests that these differences may be more pronounced in CI-
399 PD, thereby biasing any comparison in bulk tissue. In terms of the gene expression data, the
400 cell-type composition bias is mitigated by the snRNA-seq. However, when interpreting this
401 data, it is important to consider that nuclear transcriptomes do not reflect post-transcriptional
402 RNA turnover and subcellular segregation and do not, therefore, necessarily capture the full
403 transcriptomic state of the cell. This is particularly true in neurons, where RNA composition
404 varies considerably between the soma, processes, and synapses. Furthermore, snRNA-seq
405 cannot capture the mitochondrial transcriptome.

406 In summary, this work establishes that anatomically widespread neuronal CI deficiency
407 is not a pervasive feature of iPD, but one that occurs in a subset of individuals, accounting for
408 approximately a fourth of all cases. These results have important scientific and clinical
409 implications for the field. First, the existence of the CI-PD and nCI-PD subtypes provides an
410 explanation for the variable and conflicting findings in previous studies of mitochondrial
411 pathology in iPD, which were performed on unstratified samples. Second, given the major
412 molecular differences between the two iPD subtypes, we recommend that future studies
413 employing iPD brain tissue should stratify their material according to CI subtype. This would
414 decrease the biological heterogeneity of the sample and allow for more consistent and
415 biologically meaningful observations to be made. Finally, it is plausible that the CI-PD and
416 nCI-PD subtypes exhibit different susceptibility to therapeutic interventions targeting
417 mitochondria or other subtype-associated pathways. Thus, we believe that harnessing this
418 classification in clinical practice may be an essential step towards implementing precision

medicine in iPD. While our findings suggest that the non-TD motor phenotype may be used as a proxy for CI-PD to some extent, the accuracy of this approach is limited. To efficiently translate these findings into clinical practice requires the development of clinically applicable biomarkers, allowing us to classify individuals in vivo.

426 ## Materials and Methods

427 **Subject cohorts and tissue samples**

428 Brain tissue was obtained from two independent cohorts of European origin. The Norwegian
429 cohort (NOR, n = 55) comprised 41 individuals with iPD from the ParkWest study, a
430 prospective population-based cohort followed since 2004[50], and 13 neurologically healthy
431 controls (i.e., no clinical diagnosis of neurological disease or pathological evidence of α-
432 synucleinopathy), who had been prospectively collected by our group. The NOR samples were
433 stored at our in-house brain bank. The Spanish cohort (ESP, n = 62) comprised 51 individuals
434 with iPD and 11 neurologically healthy controls from the Neurological Tissue Bank, Hospital
435 Clínic – IDIBAPS Biobank, Barcelona, Spain. Analysis of whole-exome sequencing data from
436 the NOR cohort [51] and RNA-sequencing (RNAseq) data from both cohorts (described below),
437 revealed no pathogenic mutations in nuclear genes linked to monogenic PD or mitochondrial
438 diseases, thus ruling out monogenic causes. Prospectively collected systematic clinical and
439 environmental exposure data were available for the iPD cases of the NOR cohort, as previously
440 described[50]. In the NOR cohort, probable PDD was diagnosed according to the Movement
441 Disorder Society Level 1 criteria[8]. Cognitive deficiency severe enough to impair daily life was
442 defined based on the Unified Parkinson's Disease Rating Scale (UPDRS) score 1-1[52].
443 Impairment in cognitive domains was determined based on tests of attention ("serial 7's" test),
444 visuo-constructive ability (correct drawing of the pentagons), and memory impairment (3-word
445 recall), from the Mini-Mental State Examination (MMSE)[53]. Appropriate tests on executive
446 function were not available, thus the diagnosis of PDD in the NOR cohort was based on data
447 from three, rather than four cognitive domains, as required by the MDS Level 1 criteria[8]. Two
448 individuals had impairment in only one cognitive domain in addition to fulfilling the remaining
449 criteria for PDD. These two individuals may thus have been potentially misclassified and as
450 non-PDD cases. Individuals were classified by motor phenotype as TD or AR according to the

451 Schiess classification[9], and TD or PIGD according to the Jankovic classification[10]. Detailed
452 retrospective clinical data were available on the controls of the NOR cohort from the hospital
453 records. Retrospective clinical data on disease duration, dementia (yes/no) and PD phenotype
454 (TD/AR/mixed) were available from the hospital records of individuals of the ESP cohort.

455 There was no significant difference between controls and iPD subjects in PMI or sex
456 distribution in either of the cohorts, or in the combined cohort. Age of death was similar
457 between controls and iPD subjects in the NOR cohort, while the control subjects were older in
458 the ESP cohort and in the combined cohort. Subsets of samples used in downstream analyses
459 were matched for age of death, sex and PMI-range, with the exception of the transcriptomic
460 analyses, where the controls were older as reflected by the age distribution in the ESP and
461 combined cohorts. The demographic data of the cohorts and descriptive statistics of each cohort
462 are shown in Table 1.

463 Brains were collected at autopsy and processed according to either in-house[25] or
464 Neurological Tissue Bank protocol[54]. For the neuropathological evaluation, LP was staged
465 according to the Braak criteria[55,56], and β-amyloid phases were evaluated according to the Thal
466 criteria[57]. All cases fulfilled the essential pathological criteria for PD: loss of the dopaminergic
467 neurons of the SNpc in the presence of LP[58]. Subject demographics, pathology scores and
468 clinical data are summarized in Supplementary Table S1. Experimental allocation of subjects
469 is summarized in Supplementary Table S2.

470

471 **Immunohistochemistry**

472 Immunohistochemistry (IHC) for mitochondrial markers was carried out on formalin-fixed,
473 paraffin-embedded brain tissue from individuals of both the NOR and ESP cohorts. Serial 3 µm
474 thick sections were stained with primary antibodies against the following epitopes; NDUFS4
475 (ab137064, dilution 1:1500, Abcam), NDUFB8 (ab110242, dilution 1:300, Abcam), NDUFA9

476 (ab14713, dilution 1:750, Abcam), NDUFS1 (ab169540, dilution 1:300, Abcam), SDHA
477 (ab14715, dilution 1:10,000, Abcam), UQCRC2 (ab14745, dilution 1:10,000, Abcam), MTCOI
478 (459600, dilution 1:15,000, Thermo Fisher Scientific), ATP5A (ab14748, dilution 1:20,000,
479 Abcam), and VDAC1 (ab14734, dilution 1:20,000, Abcam).

480 Sections were stained in the intelliPATH FLX Automated Slide Stainer (Biocare).
481 Deparaffinization and antigen retrieval was performed in the low pH EnVision FLEX Target
482 Retrieval Solution at 98 °C for 24 minutes in the DAKO PT link from Agilent. Antibodies were
483 diluted in DaVinci Green Diluent (Biocare, PD900) and incubated for 1 hour at room
484 temperature. MACH4 Universal HRP-polymer (Biocare, M4U534) and DAB chromogen kit
485 (Biocare DB801) were used for visualization. Tacha's Automated Hematoxylin (Biocare, NM-
486 HEM) was used for visualization of nuclei. Sections were washed in the TBS Automation Wash
487 buffer from Biocare (TWB945M). Experimental allocation is given in Supplementary Table
488 S2.

489

490 **Image analysis**

491 All sections were scanned using the NanoZoomer XR (Hamamatsu). Visual analysis was
492 performed on all sections, and neurons with clearly visible nucleus and nucleolus were
493 classified as either negative or positive (Fig. 1a). After classification was performed, the total
494 percentage of positive neurons was calculated for each individual and area. From the prefrontal
495 cortex (PFC) in the NOR cohort, three regions of 3.5 mm^2 spanning cortical layers 2-6 were
496 chosen at a low magnification (1.5X), before analyses were carried out in the selected area at
497 40X magnification. Area selection from one region spanning 3.5 mm^2 was performed in a
498 similar manner in the PFC from the ESP cohort, and in the occipital and temporal cortex, the
499 cingulate gyrus, subiculum, putamen and the amygdala. All neurons per section were analyzed
500 in the CA-areas of the hippocampus (CA4-1), globus pallidus, dentate nucleus, inferior olivary

501 nucleus, and Purkinje cell layer of the cerebellar cortex. In the SNpc, all neuromelanin
502 containing neurons were analyzed per section. SNpc sections with fewer than 25 identifiable
503 neuromelanin containing neurons were excluded from downstream analyses of neuronal CI
504 status because data resulting from these sections would entail high levels of statistical
505 uncertainty, at the range of the expected group differences (i.e., any misclassified neuron as CI
506 positive or negative would result in at least 4% (1/25) difference in the estimated proportion of
507 positive neurons.) An uncertainty greater or equal to 4% was deemed unacceptable since it is
508 close to the difference between the medians of the groups being compared in our analyses.
509 Independent evaluation of the PFC from the NOR cohort was performed by two observers (IHF
510 and CT). All analyses were carried out at 40X magnification. Visual analysis was carried out
511 using the NDP.view2plus version(v)2.7.25 software (Hamamatsu).
512
513 **Single neuron laser-microdissection studies**
514 Qualitative and quantitative mtDNA studies were performed in single laser-microdissected
515 neurons from 20 µm thick sections of fresh-frozen tissue from the PFC, which had been stained
516 with 0.25% cresyl violet in ddH_2O as described[26]. Neurons were recognized based on their
517 morphology. Laser microdissection was performed on a LMD7 microscope (Leica
518 microsystems). A total of 162 cells (8-16 from each individual) were cut, of which 55 were
519 from controls (n = 7) and 107 were from individuals with iPD (n = 14). Experimental allocation
520 is shown in Supplementary Table S2. Single neurons were collected in the cap of 0.2 ml PCR
521 tubes and lysed at 56°C overnight in 15 μL lysis buffer (50mM Tris pH 8.0, 0.5% Tween
522 20, 200 μg/mL proteinase K). Following lysis, cells were centrifuged at 3,700 rpm for 10
523 minutes before proteinase K was deactivated at 95°C for 10 minutes, and the samples were
524 centrifuged again. Total mtDNA copy number and the fraction of major arc deletion were
525 determined using a duplex real-time PCR assay which targeted the mitochondrial genes

526 *MTND1* and *MTND4* as described[26]. Amplification was performed on a Step One Plus Real-
527 Time PCR System (Applied Biosystems by Thermo Fisher Scientific) using TaqMan Fast
528 Advanced Master Mix (Thermo Fisher Scientific). Thermal cycling consisted of one cycle at
529 95°C for 20 seconds, and 40 cycles at 95°C for 1 second and 60°C for 20 seconds. For absolute
530 quantification, target amplification was compared with a standard curve made form a serial
531 dilution containing equal amounts of PCR-generated and purified full length *MTND1* and
532 *MTND4* template (10^2-10^6). The following primers were used for the standard curve; *MTND1*:
533 forward primer: 5´-CAGCCGCTATTAAAGGTTCG- 3', reverse primer: 5´-
534 AGAGTGCGTCATATGTTGTTC-3´; *MTND4*: forward primer: 5´-
535 TCCTTGTACTATCCCTATGAG-3´, reverse primer: 5´- GTGGCTCAGTGTCAGTTCG-3´.
536

537 **RNA sequencing**

538 RNA-seq was performed on frozen PFC tissue samples from 98 individuals from both cohorts
539 (Supplementary Table S2). Total RNA was extracted from tissue homogenates using RNeasy
540 plus mini kit (Qiagen) with on-column DNase treatment according to manufacturer's protocol.
541 Final elution was made in 65 µl of ddH2O. The concentration and integrity of the total RNA
542 were estimated by Ribogreen assay (Thermo Fisher Scientific), and Fragment Analyzer
543 (Advanced Analytical), respectively, and 500 ng of total RNA was used for library prep using
544 the TruSeq™ Stranded Total RNA With Ribo-Zero™ Plus rRNA Depletion protocol
545 (Illumina). Library quantity was assessed by Picogreen Assay (Thermo Fisher Scientific), and
546 the library quality was estimated by utilizing a DNA High Sense chip on a Caliper Gx (Perkin
547 Elmer). Accurate quantification of the final libraries for sequencing applications was
548 determined using the qPCR-based KAPA Biosystems Library Quantification kit (Kapa
549 Biosystems, Inc.). Each library was diluted to a final concentration of 1.4 nM and pooled
550 equimolar prior to clustering. Sequencing was performed on a NovaSeq 6000 (Illumina, Inc.),

551 to yield 100 million, 100 base pair paired-end reads per sample. Sequencing was performed at
552 HudsonAlpha Institute for Biotechnology, AL, USA. RNA integrity number (RIN) and DV200
553 index (percentage of fragments > 200 nucleotides) were lower in the NOR cohort compared to
554 the ESP cohort: (mean [min-max]), NOR_{RIN}: 4.3 [1.7-7.7], ESP_{RIN}: 6.9 [2.7–9.1]; NOR_{DV200}:
555 85.3 [71.8-93.2], ESP_{DV200}: 92.2 [80.1–96.2]. No significant difference in the RNA quality
556 measures between individuals with iPD and controls was observed in either of the cohorts
557 (linear model, NOR_{RIN}: P = 0.26, NOR_{DV200}: P = 0.33; ESP_{RIN}: P = 0.11, ESP_{DV200}: P = 0.11).
558

559 **RNA-seq analysis**

560 Salmon v1.3.0 [58] was used to quantify the abundance at the transcript level with the fragment-
561 level GC bias correction option (--gcBias) against GENCODE v35. Transcript-level
562 quantification was collapsed onto gene-level quantification using the tximport R package
563 v1.8.0[59] according to the gene definitions provided by the same Ensembl release. Non-canonical
564 chromosomes and scaffolds, and genes encoded by the mitochondrial genome were excluded
565 from the analysis. In addition, we removed genes accounting for >0.6% of the reads in more
566 than half of the samples, resulting in the removal of 4 transcripts (*RN7SL2*, *RN7SL1*, *RN7SK*,
567 *LINC00632*; Supplementary Fig. S9, as well as low-expressed genes (expression <5 reads in
568 80% of the samples). At the final step, we used the expression level of the sex-specific genes
569 *XIST, KDM5D* and *RPS4Y1* to define the noise threshold for our data[30]. Relative cell abundance
570 was estimated using cell-type specific marker gene profiles (MGPs), as previously
571 described[30,60]. Difference in the cell type MGPs between the groups were assessed through
572 linear regression adjusting for age, sex, PMI and cohort, using "lm" function from the R "stats"
573 function v4.1.0. Differential expression (DE) analysis between each of iPD clusters and controls
574 was performed using "DESeq2" R package v1.32, using the protein-coding genes to calculate
575 the normalization factor, while adjusting for sex, age, PMI, cohort and the relevant MGPs.

576 Functional enrichment analysis was performed based on the one-sided P-value, previously
577 described[30,31], using the gene score resampling method implemented in the ermineR package
578 v1.0.1[61], an R wrapper package for ermineJ[62]. Enrichment was calculated for the combined list
579 of gene-sets of Gene Ontology (GO) database [63] and Kyoto Encyclopedia of Genes and
580 Genomes (KEGG) pathways [64,65] v7.2 as well as An Inventory of Mammalian Mitochondrial
581 Genes v3 (MitoCarta3.0)[66]. Only gene-sets of size >=20 and <=500 were included in the
582 analysis, resulting in 5,397 testable gene-sets. To visualize the enrichment results, pair-wise
583 similarity between significant gene-sets (excluding gene-sets from the MitoCarta database) was
584 first assessed using Jaccard Index. Next, the obtained similarity matrix was used to cluster the
585 gene-sets using the "hclust" function from the "stats" R package v4.1.2, with the "ward.D2"
586 method. The phylo tree was then created using the "ape" R package v5.5. The count matrix and
587 the code necessary to reproduce these results are available through github repository:
588 https://git.app.uib.no/neuromics/PD_Complex-I_stratification. All the information regarding
589 the version of the software and the packages used for the analysis are provided in the
590 SessionInfo.Rds file.
591

592 **Single nucleotide variant discovery**

593 RNA-seq reads were aligned against the hg38 reference genome using Hisat2 v2.2.1[67]. Genome
594 Analysis Toolkit (GATK) v4.1.9.0 [68] was used to mark duplicate reads and split the CIGAR
595 signatures using the *MarkDuplicates* and *SplitNCigarReads* tools, respectively. Base quality
596 score recalibration was then performed using the *BaseRecalibrator* and *ApplyBQRS* tools to
597 generate the final alignment files. For each alignment, *HaplotypeCaller* was used in GVCF
598 mode (option *-ERC GVCF*) with optional arguments *-dont-use-soft-clipped-bases* and *--max-*
599 *alternate-alleles 4*. Joint genotyping was carried for all resulting sample gVCF files using the
600 *GenomicsDBImport* and *GenotypeGVCFs* tools. The final analysis-ready multisample VCF

601 consisted only of single-nucleotide variants filtered using *VariantFiltration* with the conditions
602 QD < 5; DP < 10; FS > 60; MQ < 40. Samples were classified into the major mitochondrial
603 haplogroups using HaploGrep2 [69] based on the Phylotree Build 17[70].

604 Genotypes with < 10 reads were removed, and the genetic data was recoded into binary
605 PLINK format using VCFTools v0.1.15[71]. PLINK v1.90 was used for individual and variant-
606 level quality control (QC)[72]. Only autosomes were considered, and variants within the sex
607 chromosomes and mtDNA were removed. Individuals were excluded if their genomic data
608 showed a missing rate > 20%, abnormal heterozygosity (± 3 standard deviations) or cryptic
609 relatedness (identity by descent (IBD) > 0.2). Population stratification was studied using
610 multidimensional scaling (MDS) against the HapMap populations[73].

611 Single nucleotide variants (SNVs) were excluded if the missing genotype rate was >
612 20%, the call rates between PD and controls differed significantly ($P < 0.02$) or they displayed
613 departure from the Hardy-Weinberg equilibrium ($P < 1.0x10^{-5}$). Monomorphic variants were
614 removed, and the dataset contained no multiallelic variants after QC. Principal component (PC)
615 analysis was performed using Eigensoft[74,75].

616

617 **Single-nuclei isolation**

618 Nuclei isolation procedure was performed as described in[76] with minor modifications. For each
619 sample, 50 - 85 mg of frozen human prefrontal cortex were homogenized. All procedures were
620 performed on ice. Tissue was dissociated in 2 mL of homogenization buffer (320 mM sucrose,
621 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl [pH 7.8], 0.1 mM EDTA [pH 8.0],
622 0.1% IGEPAL CA-630, 1 mM DTT, and 0.4 U/µL RNase Inhibitor (ThermoFisher Scientific)
623 inside a 7 mL KIMBLE Dounce tissue grinder set (Merck), using 10 strokes with loose pestle
624 followed by 10 strokes with tight pestle. Homogenized tissue was filtered through a 30 µm cell
625 strainer and centrifuged at 1000 x g for 8 min at 4°C. Supernatant was discarded, and pellet was

626 resuspended in 450 µL of 2% BSA (in 1x PBS) containing 0.12 U/µL RNase Inhibitor. Cleanup
627 of the nuclei suspension was achieved with density gradient centrifugation. Equal volumes (450
628 µL) of nuclei suspension and working solution (50% OptiPrep density gradient medium
629 (Merck), 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl [pH 7.8], 0.1 mM EDTA [pH
630 8.0], and 1 mM DTT) were mixed. The density gradient was prepared by carefully layering 750
631 µL of 30% OptiPrep Solution (134 mM sucrose, 5 mM CaCl2, 3 mM Mg(CH3COO)2, 10 mM
632 Tris HCl [pH 7.8], 0.1 mM EDTA [pH 8.0], 1 mM DTT, 0.04% IGEPAL CA-630, and 0.17
633 U/µL RNase Inhibitor) on top of 300 µL of 40% OptiPrep Solution (96 mM sucrose, 5 mM
634 CaCl2, 3 mM Mg(CH3COO)2, 10 mM Tris HCl [pH 7.8], 0.1 mM EDTA [pH 8.0], 1 mM
635 DTT, 0.03% IGEPAL CA-630, and 0.12 U/µL RNase Inhibitor) inside a 2 mL Sorenson
636 Dolphin microcentrifuge tube (Merck). The mixture containing the nuclei (800 µL) was slowly
637 pipetted onto the top of the OptiPrep density gradient and samples were centrifuged at 10,000
638 x g for 5 min at 4°C using a fixed angle rotor (FA-45-24-11-Kit). Nuclei were recovered by
639 withdrawing a volume of 200 µL from the 30 – 40% interface of the OptiPrep density gradient
640 and transferring to a new 2 mL Sorenson Dolphin microcentrifuge tube (Merck). Nuclei were
641 washed with 2% BSA (in 1x PBS) containing 0.12 U/µL RNase Inhibitor and then pelleted by
642 centrifugation at 300 x g for 3 min at 4°C using a swing-bucket rotor (S-24-11-AT) for a total
643 of three times. The nuclear pellet was resuspended in a minimum volume of 100 µL of 2% BSA
644 (in 1x PBS) containing 0.12 U/µL RNase Inhibitor and carefully mixed by pipetting 20 times
645 on ice. The nuclei suspension was filtered through a 40 µm FlowMe cell strainer (Merck) to
646 remove any remaining nuclei aggregates. Nuclei were stained with Trypan blue and counted
647 manually on a hemocytometer.
648 10,000 nuclei per sample were targeted for droplet-based snRNA sequencing. cDNA synthesis
649 and libraries were prepared using the Chromium Next GEM Single Cell 3′ Reagent Kit v3.1
650 Dual Index (10x Genomics) according to the manufacturer's protocol. Libraries from each

651 sample were pooled and run using paired-end (300 million read pairs per sample) sequencing
652 on the NovaSeq 6000 platform (Illumina) at Novogene (Cambridge).
653
654 **snRNA-seq analyses**
655 Raw FASTQ files were aligned to the human genome (GRCh38) using CellRanger v.6.1.2 (10x
656 Genomics) accounting also for intronic reads to include all nuclear transcripts (include-introns
657 option) and excluding secondary mappings (nosecondary option). In order to estimate
658 intronic:exonic ratio, an additional alignment was carried only accounting for reads mapped to
659 exonic reads. Version 35 of the ENCODE transcriptome assembly was used (Ensembl 101).
660 To minimize the confounding effect of non-nuclear ambient RNA and discard empty
661 droplets, we employed CellBender[77] on the raw data prior to importing the counts into R.
662 Percentage of mitochondrially-encoded genes and ribosomal genes were then calculated for
663 each droplet using the Seurat R Package (*PercentageFeatureSet*)[78]. Each biological sample was
664 independently normalized, and automatic feature selection was carried out using Seurat
665 (*NormalizeData*, *FindVariableFeatures*). In order to collectively analyze the datasets, samples
666 were integrated using the standard Seurat integration strategy. Briefly, features (genes) were
667 automatically selected using *SelectIntegrationFeatures* to scale each sample and obtain a
668 principal component analysis (PCA) of gene expression. Using the first 50 principal
669 components, samples were integrated using reciprocal PCA with automatically detected
670 anchors (*FindIntegrationAnchors*) and using controls to build the reference (*IntegrateData*).
671 Clusters were firstly identified using the standard Seurat analysis with excess barcodes (i.e.,
672 including twice as many barcodes in each sample to ensure the inclusion of a large population
673 of empty droplets). On each sample the largest cluster always corresponded to the cluster with
674 the largest proportion of empty droplets (as defined by DropletUtils[79]]). Visual inspection of
675 the UMAP plots allowed the identification of other smaller ambient clusters which were

676 removed together with the main empty droplet cluster for downstream analyses. Resulting non-
677 empty droplets were then further screened and only droplets with at least 500 transcripts, at
678 least 1,000 UMIs, and below 3% mitochondrially-encoded transcripts were analyzed.
679 Transcripts detected in fewer than 10 cells were also filtered out. Doublets were assessed
680 independently for each sample using the DoubletFinder R package[80] assuming 7% doublet rate.
681 Standard clustering using Seurat with 0.5 resolution resulted in a total of 28 clusters (Fig. 7A)
682 which could be assigned to the main cortical cell types based on known markers (SATB2 and
683 SLC17A7 for excitatory neurons, GAD1 and SLC6A1 for inhibitory neurons, MOG and MOBP
684 for oligodendroglia, GFAP and AQP4 for astrocytes, CLDN5 and SLC2A1 for endothelial
685 cells, OLIG1 and VCAN for oligodendrocyte precursor cells, and CSF1R and CSF3R for
686 microglia[81], Fig. 7B. Finally, to increase the within-cluster cell type homogeneity and minimize
687 the number of doublets, we used two alternative approaches. First, the Euclidean distance in
688 PCA space was calculated between each single cell and the average of each cluster using the
689 first 5 PCs. Barcodes for which the shortest distance did not correspond to their cluster
690 membership were removed from the analyses. We then calculated the first PC of the cell type
691 specific markers for each barcode (marker gene profiles, MGP[60]). MGPs correspond to the first
692 PC of the cell type specific markers. In this way, we obtained for each droplet a value for each
693 major cell type (i.e., neuron, oligodendrocyte, endothelial cell, microglia, astrocyte, OPC)
694 scaled across all droplets. We then calculated a score for each droplet, corresponding to the
695 MGP for the annotated cell type divided by the sum of the other MGPs. Droplets with poor
696 scores (i.e., below 0.4) were filtered. The filtering procedure resulted in a total of 117,105
697 nuclei. The cluster labels were used to compare cell type composition of the diseased and
698 control groups using ANOVA.

699 To assess the statistical significance of the difference in the number of captured nuclei
700 between groups (nCI-PD, CI-PD and Ctrl) we used an ANOVA test at $\alpha = 0.05$. Differences

701 between groups in the number of genes per nucleus were assessed for each cell type cluster
702 with linear mixed-effects models (LMM) with disease status as fixed effect and individual as
703 random effect. P-values for each LMM were estimated using the Satterthwaite's method. The
704 resulting p-values were considered significant above a false discovery rate (FDR) of 5%. The
705 same statistical approach was used to assess differences in the number of unique molecular
706 (UMI) counts between disease groups.

707 Differential expression within clusters between diseased groups and controls was
708 carried out using the MAST R package[82] using sex, post-mortem interval and age as fixed
709 effects, and individual as random effect[83] (Fig 7C).

710

711 **Statistical analyses of the demographic, clinical, IHC and mtDNA data**

712 Subject demographics, clinical data, and pathology scores are given in Supplementary Table
713 S1. Experimental allocation is given in Supplementary Table S2. Data were tested for normality
714 and equality of variances, using the Shapiro-Wilk test and Levene's test. Clustering was
715 performed in R (v4.1.2), using k-means with three centers, 20 random starting partitions, 100
716 iterations, and the Hartigan-Wong algorithm. Between group-comparisons were consequently
717 performed using either independent student's t-test or Mann-Whitney U-test, depending on data
718 normality, and descriptive data are given either as mean (standard deviation) or median [25th-
719 percentile-75th-percentile], respectively. Comparisons of grouped data were performed using
720 either Chi-Square (X^2) or Fisher's exact test depending on the group size of each variable. Data
721 is given as X^2 and/or odds ratio (95% confidence interval). Correlation analyses were performed
722 using Spearman's rank correlation due to non-normality of the data, Spearman's rank
723 correlation coefficient is given as $\rho(n-2)$. Between-group (Ctrl, CI-PD and nCI-PD)
724 comparison of neuronal mtDNA deletion proportion and copy number was performed by one-
725 way ANOVA on the individual median values. Due to the unequal sample sizes, Welch's test

726 for unequal variance was used. Eta squared (η^2) was used to determine the effect size of any
727 significant difference. The Games-Howell post hoc test was used to assess the statistical
728 significance of pairwise differences between groups, due to their unequal sample sizes and
729 variance. Hierarchical clustering analysis was performed in R v4.1.2 with the "ward.D2"
730 method. Correlation analyses and between-group comparisons were done in SPSS v27.0.0.
731 Plots were created in R v4.1.2 or GraphPad Prism v9.3.0. A P-value < 0.05 was considered
732 statistically significant. Statistical tests performed in the context of omics data analyses are
733 described in the corresponding methods sections.

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. 2; 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.

Competing interests

The authors declare no competing interests.

Availability of data and materials

The datasets and code required to reproduce the results of these analyses are available in the Neuromics Group repository, https://git.app.uib.no/neuromics/PD_Complex-I_stratification.

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: performed part of the IHC stainings. NL: contributed to the preparation of Figure 2, advised on the statistical approach, and provided critical input for the manuscript. GA

759 and OBT: contributed biological material and clinical data and provided critical input to the
760 manuscript. AP-S, CP, YC: contributed clinical data and input to the manuscript. LMP:
761 contributed biological material, pathological data and critical input to the manuscript. CD:
762 contributed to the qPCR analyses and provided critical input for the manuscript. GSN:
763 participated in the study design, performed the prepossessing of all transcriptomic and genetic
764 data, performed the mtDNA variant analyses, participated in the interpretation of the
765 transcriptomic and genetic data, and drafted parts of the manuscript. CT: conceived, designed
766 and directed the study, contributed to data analyses and interpretation, drafted the manuscript,
767 and acquired funding for the study. All authors have read and approved the manuscript.
768

769 **Funding sources**

770 This work is supported by grants from The Research Council of Norway (288164), Bergen
771 Research Foundation (BFS2017REK05), The KG Jebsen Foundation (SKGJ-MED-023), and
772 the Western Norway Regional Health Authority (F-10229-D11661).
773

774 **Ethics declarations**

775 Ethics approval for these studies was obtained from our regional ethics committee (REK
776 2017/2082, 2010/1700, 2016/1592). Consent for publication was provided by all participants.
777 The authors declare that they have no competing interests.

References

1. Bloem BR, Okun MS, Klein C. Parkinson's disease [Internet]. The Lancet 2021;0(0)[cited 2021 Apr 17] Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)00218-X/abstract

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 [Internet]. Cold Spring Harb Perspect Med 2012;2(8)Available from: http://www.ncbi.nlm.nih.gov/pubmed/22908195

6. Chen-Plotkin AS, Albin R, Alcalay R, et al. Finding Useful Biomarkers for Parkinson's Disease [Internet]. Sci. Transl. Med. 2018;10(454)[cited 2020 May 24] Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6097233/

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

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

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

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

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

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

13. Flønes IH, Tzoulis C. Mitochondrial Respiratory Chain Dysfunction—A Hallmark Pathology of Idiopathic Parkinson's Disease? [Internet]. Front. Cell Dev. Biol.

813 2022;10[cited 2022 Aug 11] Available from:
814 https://www.frontiersin.org/articles/10.3389/fcell.2022.874596

815 14. Ramsay RR, Salach JI, Singer TP. Uptake of the neurotoxin 1-methyl-4-phenylpyridine
816 (MPP+) by mitochondria and its relation to the inhibition of the mitochondrial oxidation
817 of NAD+-linked substrates by MPP+. Biochem. Biophys. Res. Commun.
818 1986;134(2):743–748.

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

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

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

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

828 19. Balafkan N, Tzoulis C, Muller B, et al. Number of CAG repeats in POLG1 and its
829 association with Parkinson disease in the Norwegian population. Mitochondrion
830 2012;12(6):640–3.

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

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

835 22. Carling PJ, Mortiboys H, Green C, et al. Deep phenotyping of peripheral tissue facilitates
836 mechanistic disease stratification in sporadic Parkinson's disease. Prog. Neurobiol.
837 2020;187:101772.

838 23. Winkler-Stuck K, Wiedemann FR, Wallesch C-W, Kunz WS. Effect of coenzyme Q10 on
839 the mitochondrial function of skin fibroblasts from Parkinson patients. J. Neurol. Sci.
840 2004;220(1–2):41–48.

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

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

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

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

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

853 29. Streiner DL. Best (but oft-forgotten) practices: the multiple problems of multiplicity-
854 whether and how to correct for many statistical tests. Am. J. Clin. Nutr.
855 2015;102(4):721–728.

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

859 31. Nido GS, Dick F, Toker L, et al. Common gene expression signatures in Parkinson's
860 disease are driven by changes in cell composition. Acta Neuropathol. Commun.
861 2020;8(1):55.

862 32. Toker L, Tran GT, Sundaresan J, et al. Genome-wide histone acetylation analysis reveals
863 altered transcriptional regulation in the Parkinson's disease brain. Mol. Neurodegener.
864 2021;16(1):31.

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

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

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

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

873 37. Boonstra JT, McGurran H, Temel Y, Jahanshahi A. Nigral neuropathology of Parkinson's
874 motor subtypes coincide with circuitopathies: a scoping review. Brain Struct. Funct.
875 2022;227(7):2231–2242.

876 38. Brakedal B, Dölle C, Riemer F, et al. The NADPARK study: A randomized phase I trial of
877 nicotinamide riboside supplementation in Parkinson's disease. Cell Metab.
878 2022;34(3):396-407.e6.

879 39. Rossignol R, Faustin B, Rocher C, et al. Mitochondrial threshold effects. Biochem J
880 2003;370(Pt 3):751–62.

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

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

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

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

891 44. Lax NZ, Grady J, Laude A, et al. Extensive respiratory chain defects in inhibitory
892 interneurones in patients with mitochondrial disease. Neuropathol. Appl. Neurobiol.
893 2016;42(2):180–193.

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

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

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

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

902 49. Agarwal D, Sandor C, Volpato V, et al. A single-cell atlas of the human substantia nigra
903 reveals cell-specific pathways associated with neurological disorders. Nat. Commun.
904 2020;11(1):4183.

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

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

910 52. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision
911 of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and
912 clinimetric testing results: MDS-UPDRS: Clinimetric Assessment. Mov. Disord.
913 2008;23(15):2129–2170.

914 53. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for
915 grading the cognitive state of patients for the clinician. J. Psychiatr. Res.
916 1975;12(3):189–198.

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

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

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

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

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

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

931 60. Mancarci BO, Toker L, Tripathy SJ, et al. Cross-Laboratory Analysis of Brain Cell Type
932 Transcriptomes with Applications to Interpretation of Bulk Tissue Data. eNeuro
933 2017;ENEURO.0212-17.2017.

934 61. Mancarci O. ErmineR (R package) [Internet]. [date unknown];Available from:
935 https://github.com/PavlidisLab/ermineR

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

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

940 64. Kanehisa M. The KEGG resource for deciphering the genome. Nucleic Acids Res.
941 2004;32(90001):277D – 280.

942 65. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB)
943 3.0. Bioinformatics 2011;27(12):1739–1740.

944 66. Rath S, Sharma R, Gupta R, et al. MitoCarta3.0: an updated mitochondrial proteome
945 now with sub-organelle localization and pathway annotations. Nucleic Acids Res.
946 2021;49(D1):D1541–D1547.

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

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

952 69. Weissensteiner H, Pacher D, Kloss-Brandstätter A, et al. HaploGrep 2: mitochondrial
953 haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res.
954 2016;44(W1):W58-63.

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

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

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

961 73. Gibbs RA, Belmont JW, Hardenbol P, et al. The International HapMap Project. Nature
962 2003;426(6968):789–796.

963 74. Patterson N, Price AL, Reich D. Population Structure and Eigenanalysis. PLOS Genet.
964 2006;2(12):e190.

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

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

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

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

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

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

981 81. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the
982 human brain | Science [Internet]. [date unknown];[cited 2023 Jun 4] Available from:
983 https://www.science.org/doi/10.1126/science.aaf1204

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

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

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

Table 1. Demographic information

Variable	Group	n	Median	IQR	U	P-value (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 (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 (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				

995 PMI: post-mortem interval, IQR: interquartile range, SD: standard deviation, OR: odds ratio,

996 C.I.: confidence interval. NOR: Norwegian cohort, ESP: Spanish cohort, PD: Parkinson's

997 disease, Ctrl: controls.

998

999

1000 **Table 2. Between group comparisons of percentage CI positive neurons**
1001

Tissue	Cohort	Group	n	Median	IQR	U-value	P-value (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		

1002
1003 IQR: interquartile range, SNpc: substantia nigra pars compacta, PFC: prefrontal cortex, NOR:

1004 Norwegian cohort, ESP: Spanish cohort, iPD: idiopathic Parkinson's disease, Ctrl: controls

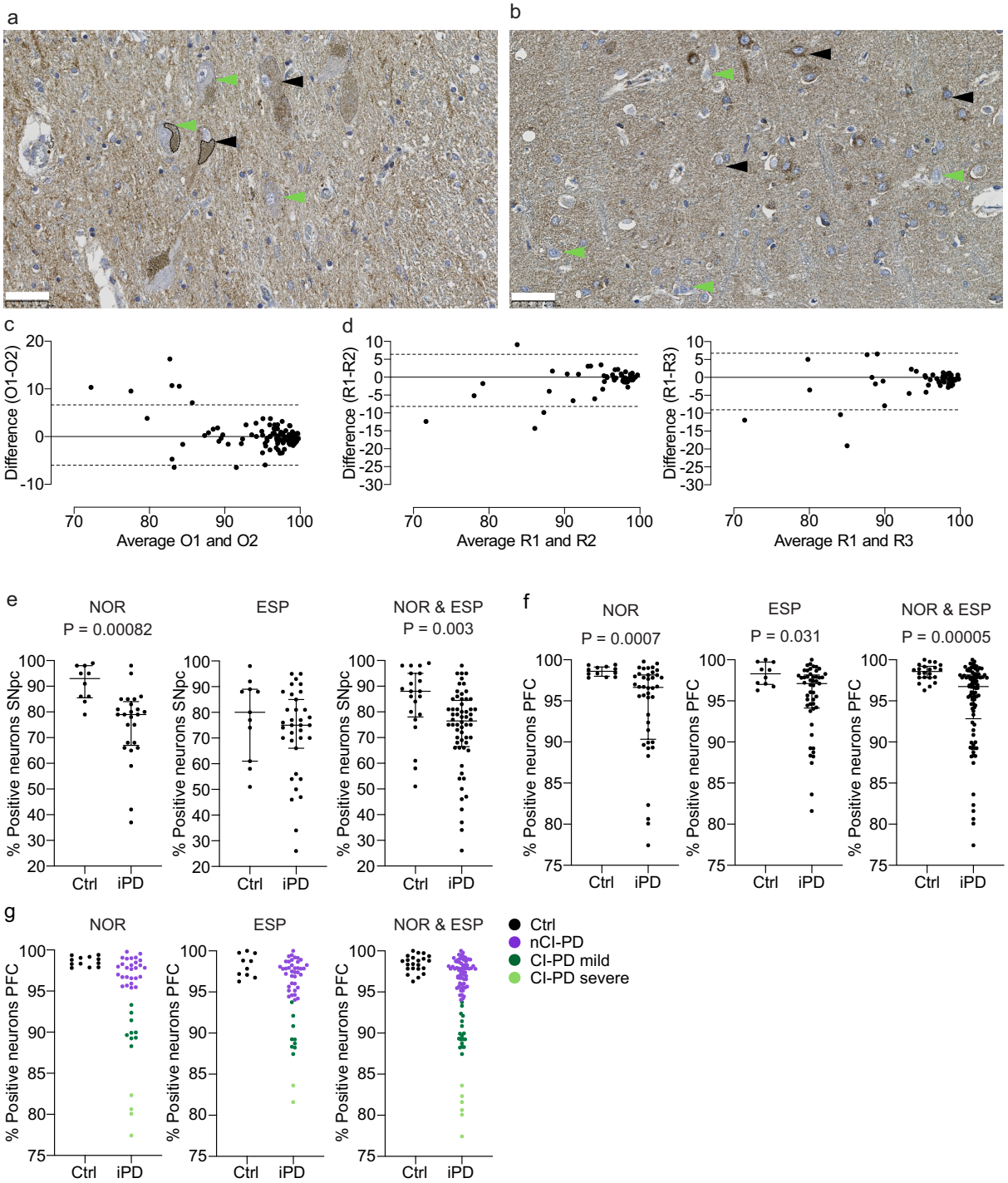
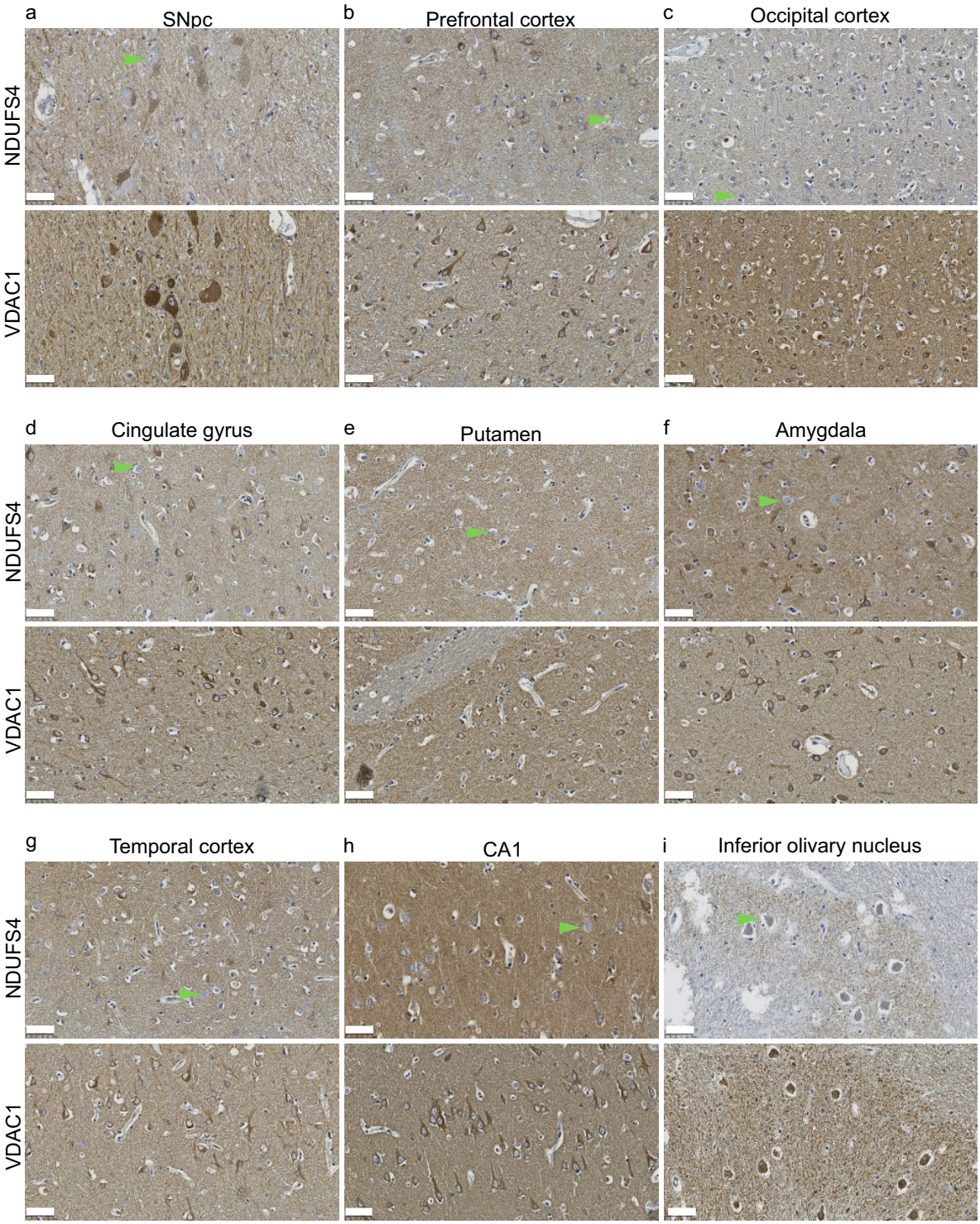


Fig. 1. Neuronal CI deficiency stratifies iPD

(a-b) SNpc (a) and PFC (b) sections immunohistochemically stained for CI (NDUFS4). The stippled lines in (a) indicate neuromelanin in dopaminergic neurons of the SNpc. Arrowheads show examples of CI deficient (green) and -intact (black) neurons. Scalebar: 50 µm. (c-d) Inter-rater and inter-region replicability of the proportion of CI-positive neurons in the PFC. Bland-Altmann plots showing the difference between the measurements of observer 1 (O1) and

1012 observer 2 (O2) (c), and the difference between measurements by observer 1 in three different
1013 regions of the same PFC sections: region 1 (R1) vs region 2 (R2), or region 3 (R3) (d). The Y-
1014 axes show the respective average percentages of CI-positive neurons. (e-f) Scatter plots
1015 showing the proportion (in %) of CI positive neurons (y-axes) in the SNpc (e) and PFC (f) of
1016 controls and individuals with iPD (x-axes). Bars show median and interquartile range. (g)
1017 Scatter plots showing the proportion (in %) of CI positive neurons (y-axes) in the PFC of
1018 controls and individuals with iPD (x-axes). K-means clustering classified the iPD subjects into
1019 three groups: nCI-PD (purple), CI-PD mild (dark green) and CI-PD severe (light green), in each
1020 of the NOR and ESP cohorts and in the combined cohorts.



1021

1022 **Fig. 2. CI and VDAC1 immunostaining in selected brain regions**

1023 Representative images of immunostaining against CI (NDUFS4) and VDAC1, in areas showing
1024 CI deficiency in iPD. Green arrowheads show examples of CI negative neurons. No negative
1025 neurons were observed in the VDAC1 staining. Scalebar: 50 µm.
1026

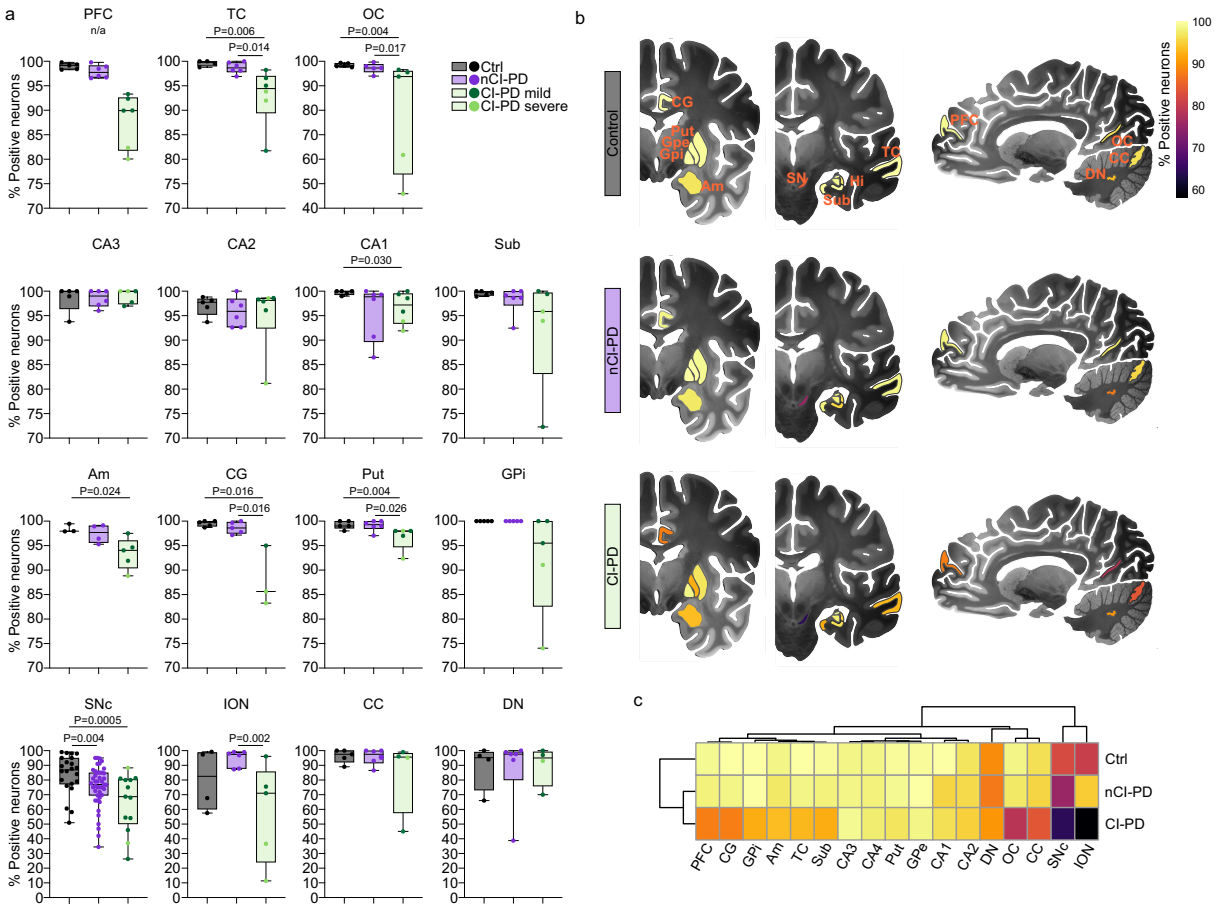


Fig. 3. Anatomical distribution of neuronal CI deficiency in the nCI-PD and CI-PD groups (a) Box plots showing the proportion (in %) of CI positive neurons (Y-axes) in 16 brain regions from controls (grey), and individuals with nCI-PD (purple) and CI-PD (green). The boxplot from PFC represents the spread in all individuals, dots show individuals included in downstream analyses for regional CI staining. Box plots show individual values (dots), median and interquartile range (box) and minimum-maximum range (fences). Statistical analysis was not performed in the PFC, since this region was used to classify the samples, and the samples were selected to represent the full spectrum of CI status for each group. Nominal p-values are shown. Multiple testing correction was not performed for these analysis, due to the relatively small sample size and high correlation between the CI state of different anatomical regions per individual, as illustrated in the heatmap (c). All individuals with CI-PD goup are shown in the same green box, but differentiated by the color of the points into mild (dark green) and severe

1041 (light green) CI-PD. (b) Structural brain maps illustrating the anatomical distribution of
1042 neuronal CI deficiency in the control, nCI-PD and CI-PD groups. The median proportion of CI
1043 positive neurons is shown as a heatmap. Brain illustrations are modified images from the 7MRI
1044 atlas published by Edlow et al[84]. (c) Hierarchical clustering of the median proportion of CI
1045 positive neurons in each group. SNpc: substantia nigra pars compacta, PFC: prefrontal cortex,
1046 OC: occipital cortex, TC: temporal cortex, Sub: subiculum, Am: amygdala, CG: cingulate
1047 gyrus, Put: putamen, GPe: external globus pallidus, GPi: internal globus pallidus internus, ION:
1048 inferior olivary nucleus, CC: cerebellar cortex, DN: dentate nucleus, Hi: hippocampal regions
1049 CA4-CA1, Ctrl: controls. Statistical test: Mann-Whitney U-test, n/a: statistical comparison not
1050 performed.
1051

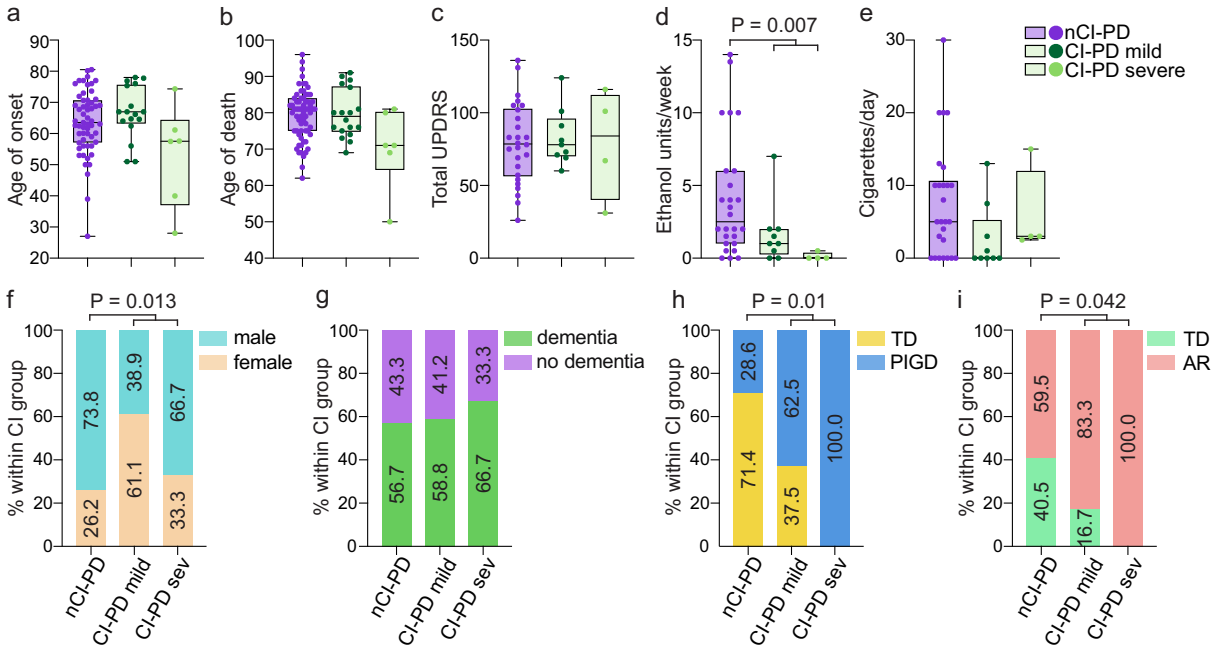
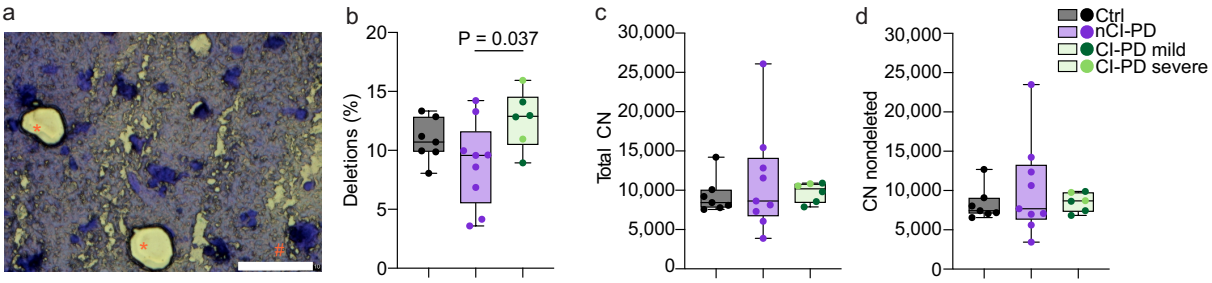


Fig. 4. Clinical differences between CI-PD and nCI-PD

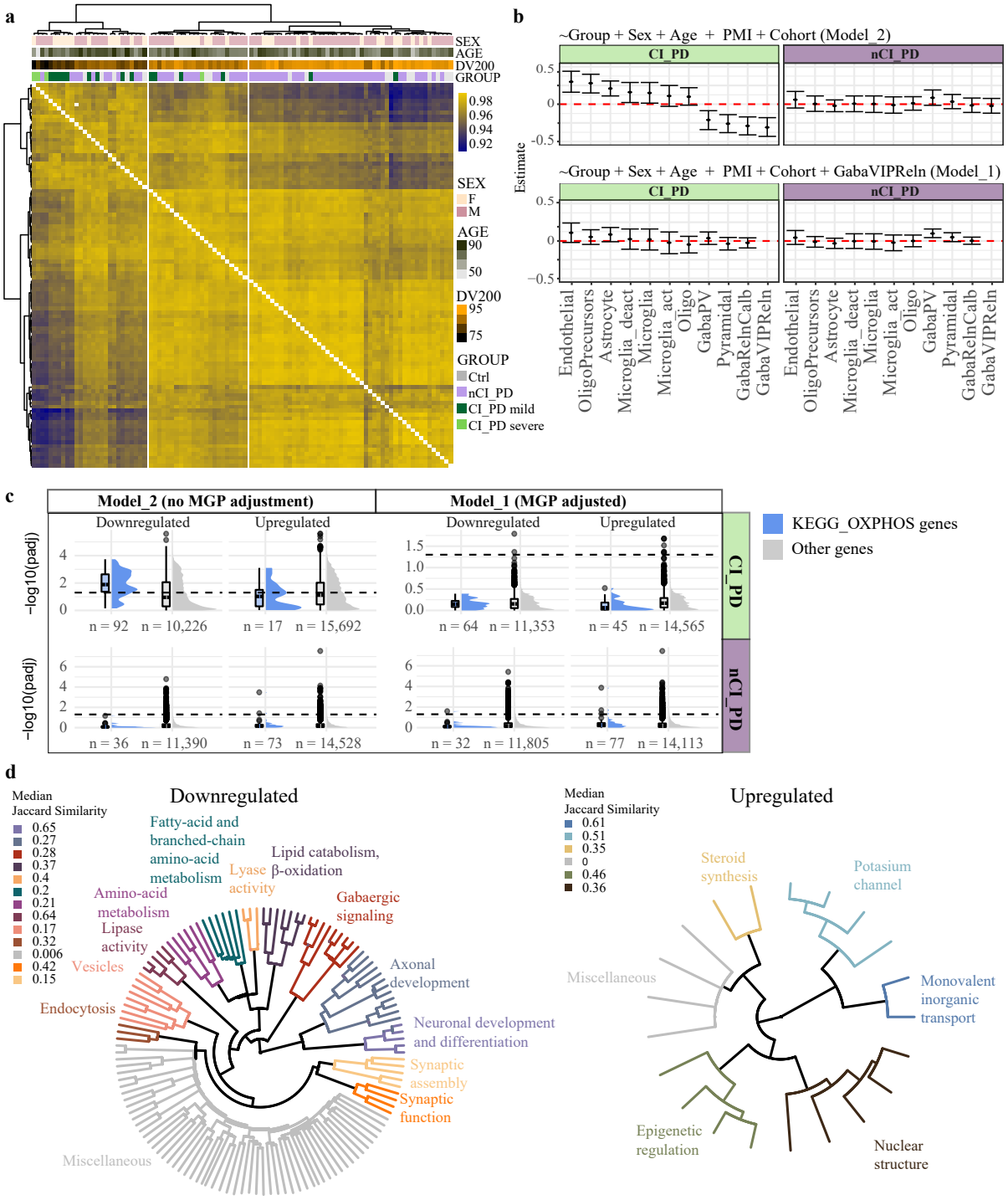
Box plots show individual values (dots), median and interquartile range (box) and minimum-maximum range (fences) of clinical data (y-axes) for the three iPD groups (nCI-PD: purple, CI-PD mild: dark green, CI-PD severe: light green). Statistical significance was tested between nCI-PD and the entire CI-PD group (i.e., not differentiating between mild and severe), due to the low number of individuals in the severe group, using the Mann-Whitney U-test. Bar plots show percentages of individuals within each iPD group (y-axis). Statistical significance was tested using the X^2 test. Demographic data, data on dementia at time of death, and the AR/TD-clinical phenotypes were available from both cohorts. All other clinical variables were available only from the NOR cohort. AR: akinetic rigid, TD: tremor dominant, PIGD: postural instability and gait disorder, UPDRS: Unified Parkinson's Disease Rating Scale.

1066 

1067 **Fig. 5. The nCI-PD and CI-PD groups show different neuronal mtDNA profiles**

1068 Single neurons were collected from frozen PFC sections using laser-microdissection and
1069 mtDNA was analysed for copy number (CN) and deletion proportion using qPCR. (a)
1070 Microphotograph of a representative PFC section, stained with cresyl violet and showing
1071 neurons before (#) and after (*) microdissection. Scale bar 10 µm. (b) Total neuronal mtDNA
1072 copy number. (c) Neuronal mtDNA deletion levels. (d) Non-deleted neuronal mtDNA levels.
1073 Box plots show median and interquartile range (box) and minimum-maximum range (fences)
1074 of the median values for each individual (points) per group. Statistical significance of group
1075 difference was tested using Student's t-test. CN: copy number, Ctrl: control.

1076



1077

1078 **Fig. 6. nCI-PD and CI-PD groups exhibit distinct transcriptomic profiles in the PFC**

1079 (a) Hierarchical clustering of the samples based on sample-sample correlation. The heatmap

1080 colors indicate Pearson's correlation values between each pair of samples. AGE: age of death.

1081 (b) The estimated group difference (y-axis) in MGPs of the cell types (x-axis) based on linear

1082 regression using the three different models are shown. The cell types are described in Mancarci

et al [56]. (c) Significance (shown as -log10 of the adjusted p-value) of genes included in the KEGG OXIDATIVE_PHOSPHORYLATION pathways (blue) or all other genes (grey) among the downregulated/upregulated genes based on a model not adjusting (Model_2) or adjusting (Model_1) for the confounded MGPs, in each of the iPD subtypes. Dashed line indicates adjusted pvalue of 0.05. (d) Enrichment analysis outcome of the nCI-PD group, based on DE analysis using Model_2. Each leaf of the tree represents one significantly enriched gene-set. Jaccard similarity was used to cluster the gene-sets based on semantic similarity. Each cluster was annotated to represent the main function represented by its members. The names of the gene-sets are shown in Supplementary Fig. S7 (downregulated) and Supplementary Fig. S8 (upregulated).

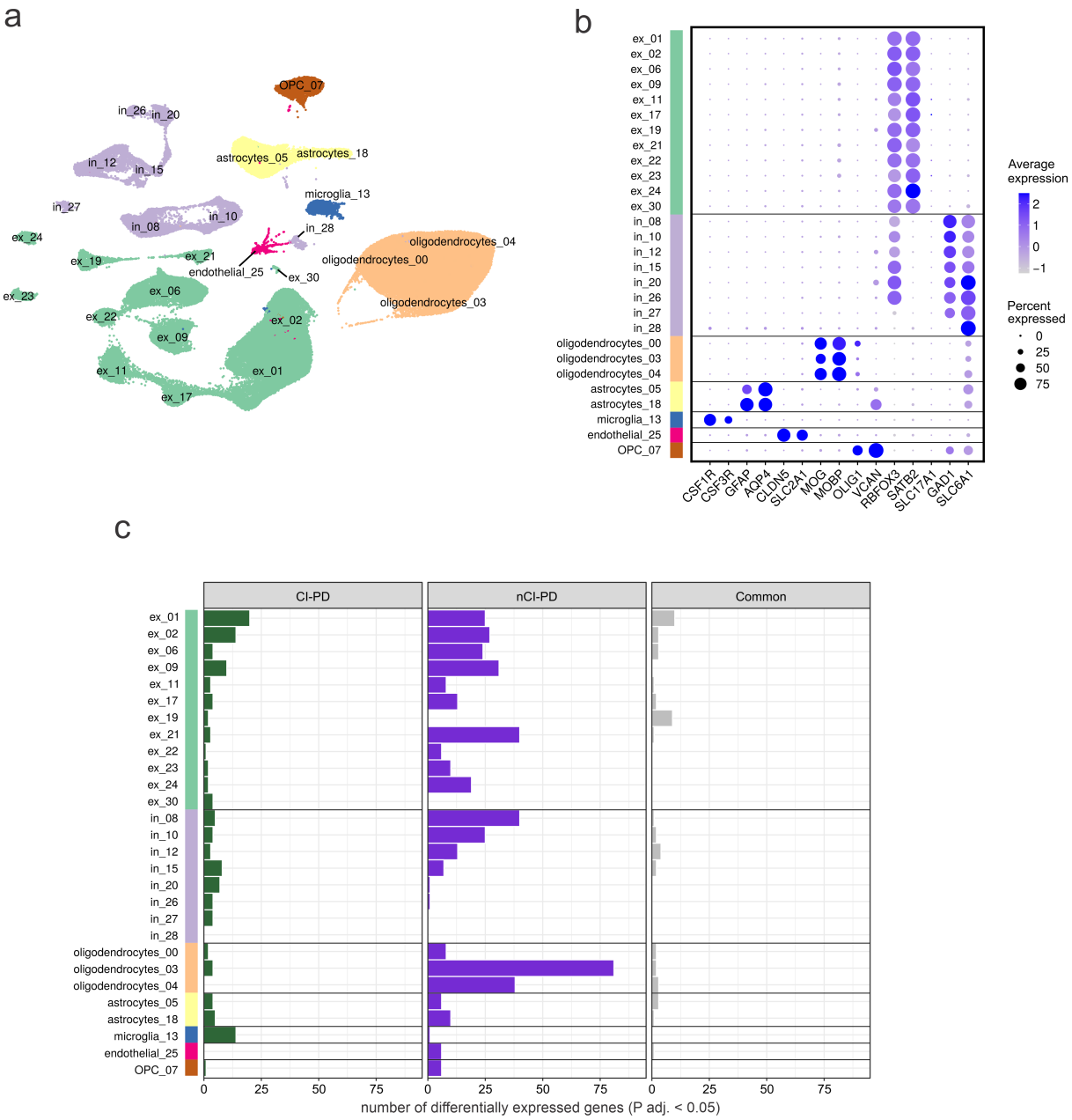


Fig. 7. Single-nuclei RNA-seq shows distinct cell type-specific transcriptional profiles for CI-PD and nCI-PD

(a) UMAP plot of the filtered snRNA-seq dataset (117,105 nuclei) showing the clustering in the main cortical cell types (ex: excitatory neurons, in: inhibitory neurons, oligodendrocytes, astrocytes, microglia, oligodendrocyte precursor cells (OPC) and endothelial cells). (b) Dotplot showing the average cell type markers' expression in each of the 28 subclusters identified by Seurat, unequivocally mapping the nuclei to the main cortical known cell types. The color of the dots indicates average scaled expression, the size representing the proportion of nuclei in

1103 the cluster that express the marker. (c) Summary of the number of differentially expressed genes

1104 (at FDR < 5%) within each cluster for the CI-PD *vs* controls (left hand side panel, green bars),

1105 for the nCI-PD *vs* controls (center panel, purple bars), and common between the two contrasts

1106 (right hand side panel, grey bars).