The circular and linear transcriptome in Parkinson's disease

Benjamin Whittle

Thesis submitted to Newcastle University for the degree of Doctor of Philosophy

Biosciences Institute

Faculty of Medical Sciences

Newcastle University

July 2024

Abstract

Neurodegeneration in Parkinson's disease (PD) primarily affects dopaminergic neurons in the substantia nigra, often preceding clinical diagnosis by years. Biomarkers that reflect this pathology are needed for diagnosis, tracking disease progression and identifying at-risk individuals. RNA levels in PD patients might serve as measurable biomarkers in easily accessible tissues. While RNA produced at genetic loci is typically linear, circular RNA (circRNA) can also be expressed. CircRNAs are potential diagnostic biomarkers due to their increased stability and distinct regulation compared to cognate linear RNA. However previous RNA studies lack concordance, partly due to small cohort sizes, disease heterogeneity, and variable analysis methodologies.

In this thesis, I analysed RNA sequencing data from blood samples of PD patients and controls from the PPMI and ICICLE-PD studies to measure gene, circRNA and cognate linear RNA expression. Using circRNA as a diagnostic biomarker in PD showed no clear improvement over linear RNA, minimising its potential clinical utility. Early-stage idiopathic PD patients exhibited a global reduction in circRNA expression, alongside increased expression of *RNASEL* and genes linked to the innate immune response. Similar global reductions in circRNA expression were observed in PPMI PD patients harbouring pathogenic *LRRK2* and *GBA* variants. Asymptomatic PPMI participants at increased risk of developing PD, due to pathogenic *LRRK2* and *GBA* variants or hyposmia, also had globally reduced circRNA expression. Conversely, RNA sequencing of dopaminergic neurons isolated from the substantia nigra of individuals from the BRAINcode study showed increased global circRNA expression in PD patients compared to controls.

Overall, this work suggests that changes in global circRNA levels, potentially driven by the innate antiviral response, may play an important role in PD development and pathobiology.

Dedicated to my grandad, Ken.

Acknowledgements

Firstly, thank you to my supervisors Professor Gavin Hudson, Dr Mauro Santibanez-Koref, Dr Michael Jackson and Dr Angela Pyle, for your support and guidance throughout this project. I am grateful for all for your advice and for helping me develop as a scientist. Gavin, it's been a pleasure working for you the past few years - thank you for all the hours spent discussing science and importantly for making day to day life in the lab (desk) fun. Mauro, thank you for all your advice on pretty much everything, apologies for having to put up with my northern dialect sneaking into my writing!

This work would not have been possible without the PPMI and ICICLE-PD studies. Thank you to all study participants for giving up their time and to the respective study teams for producing these fantastic scientific resources. Furthermore, I am grateful for the Michael J Fox Foundation and MRC DiMeN DTP for allowing me to carry out this project. Thank you to Dr Osagie Izuogu and Dr Dasha Deen for guidance regarding computational analyses and members of my panel, Professor Alison Yarnall and Dr Ian Wilson, for discussions regarding my work.

Thank you to the Mitochondrial Research Group for providing a fantastic environment to carry out research. I've accidently learnt a lot about mitochondria and in return I appreciate you all listening to me witter on about RNA unrelated to OXPHOS or heteroplasmy. Thank you to the M4047 lab and wider group for the scientific discussions and members past and present of the Big Office for the (mainly) non-scientific discussions. Thank you to everyone that made time at work feel a bit less like working, ranging from Supermarket Chat™ at lunchtimes to unwinding at Friday Club™ after a long week - I'm grateful to be able to call you friends.

Outside of work, thank you to all the Lancashire lot for the welcome distractions of hikes, festivals and holidays. Especially the lads from Accy for trekking up the A59/A1 to come visit and for keeping me grounded (Another degree? When are you moving back? When are you getting a proper job?). In Newcastle, thanks to Matt and Bentley (and Spareroom) for being my best mates since I first moved to Newcastle and providing hours of fun and a negative bank balance. I also don't think I'd have finished this thesis without a few inanimate objects including a pair of Soundcore Liberty 4 NC headphones, my slow cooker and Aphex Twin's ambient albums.

Lastly, thank you to Lucy and my family for your love and support. Lucy, cheers for putting up with me and being my best friend for the past few years. Your constant enthusiasm and

positivity managed to balance out the slog of thesis writing. I'm thankful for the supply of poke bowls, killing spiders and always being up for a glass of wine after a tough day. To my family, thank you for the facetimes and finding any excuse to come visit Newcastle. I also appreciate the attempts to understand what I'm doing - Mum this is my big book!

Author's declaration

This thesis is submitted for the degree of Doctor of Philosophy at Newcastle University. I, Benjamin Whittle, declare that the work described in this thesis is my own, unless clearly acknowledged and stated otherwise. I certify that I have not submitted any of the material in this thesis for a degree qualification at this or any other university.

Benjamin Whittle

Publications arising from this work

Whittle, B.J., Izuogu, O.G., Lowes, H., Deen, D., Pyle, A., Coxhead, J., Lawson, R.A., Yarnall, A.J., Jackson, M.S., Santibanez-Koref, M., Hudson, G., 2024. Early-stage idiopathic Parkinson's disease is associated with reduced circular RNA expression. *npj Parkinsons Dis*. 10, 1–14. https://doi.org/10.1038/s41531-024-00636-y

The publication above contains aspects of **Chapters 3**, **4** and **5**.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Full name
(MDS)-UPDRS	(Movement Disorder Society) Unified Parkinson's Disease Rating Scale
95% CI	95% Confidence Interval
AMP-PD	Accelerating Medicine Partnership Parkinson's disease program
AUC	Area under the curve
BRAINcode	BRAIN Cell encyclOpeDia of transcribed Elements Consortium
BSJ	Back-spliced junction
ceRNA	Competitive endogenous RNA
circRNA	Circular RNA
CNS	Central Nervous System
CSF	Cerebrospinal fluid
DAT	Dopamine transporters
DLB	Dementia with Lewy bodies
DNA	Deoxyribonucleic acid
dsDNA	double-stranded DNA
dsRNA	Double-stranded RNA
FDR	False Discovery Rate
FSJ	Forward-spliced junction
GD	Gaucher's disease
GI	Gastrointestinal tract
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
GWAS	Genome wide association study
HLA	Human leukocyte antigen
IBD	Inflammatory bowel disease
ICICLE-PD	Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-PD
ILB	Incidental Lewy body disease
iPSC	Induced pluripotent stem cell
IQR	Interquartile range
IR*Alus*	Inverted repeat sequences in *Alu* elements
IRES	Internal ribosome entry site
IRF	Interferon regulatory factors
ISR	Integrated Stress Response
KEGG	Kyoto Encyclopedia of Genes and Genomes
L-DOPA	Levodopa
LEDD	Levodopa equivalent daily dose
lncRNA	Long non-coding RNA
log_2FC	log_2 Fold Change
m^6A	N6-methyladenosine
MCI	Mild cognitive impairment
miRNA	microRNA
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
mRNA	Messenger RNA
NES	Normalised Enrichment Score
NfL	Neurofilament light chain
PBMC	Peripheral blood mononuclear cells

PC	Prinipal Component
PCA	Principal Component Analysis
PD	Parkinson's disease
PDBP	Parkinson’s Disease Biomarker Program
PMCA	Protein misfolding cyclic amplification assay
Pol II	RNA polymerase II
poly(A)	Polyadenylated
poly(I:C)	Polyinosinic:polycytidylic acid
PPMI	Parkinson's Progression Marker Iniative
QC	Quality control
qPCR	Quantitative polymerase chain reaction
RBD	Rapid eye movement sleep behaviour disorder
RBP	RNA-binding protein
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROC	Receiver operator characteristic
rRNA	Ribosomal RNA
RT-QuIC	Real-time quaking-induced conversion assay
SAA	α-synuclein seed amplification assay
scRNA-seq	Single cell RNA sequencing
SD	Standard deviation
TMM	Trimmed Mean of M-values
TPM	Transcripts per million
UPSIT	University of Pennsylvania Smell Identification Test
USSLB	The Unified Staging System for Lewy body disorders
UTR	Untranslated region

Chapter 1. Introduction

1.1 Parkinson's disease (PD)

Parkinson's disease (PD) is a complex neurodegenerative condition characterised by, among other symptoms, progressive motor impairment. Following Dr. James Parkinson's 1817 description in *An Essay on the Shaking Palsy* (Parkinson, 2002), PD is now recognised as the second most common neurodegenerative condition after Alzheimer's disease, with increasing incidence (Dorsey, Sherer, et al., 2018). PD imposes significant societal, economic, and individual burdens. Currently, clinical diagnosis occurs only after motor deficits appear due to neurodegeneration. There are no approved disease-modifying treatments; existing therapies focus on symptom management and quality of life improvement. Non-invasive and accurate diagnostic biomarkers are needed to support earlier clinical diagnosis, which is crucial for enrolling participants in trials for potential treatments.

1.1.1 Clinical features

PD is associated with both motor and non-motor symptoms (**Figure 1.1**). The hallmark motor symptoms include bradykinesia, tremors, and rigidity, often presenting unilaterally. Non-motor symptoms can encompass mood disorders, gastrointestinal issues, sensory and sleep dysfunctions (Jankovic, 2008). Motor symptoms are typically the first to be recognised before diagnosis, but there is evidence that non-motor symptoms can precede motor symptoms by up to 10 years (Postuma et al., 2009) (**Figure 1.1**). PD patients have an increased mortality compared to control populations (Macleod et al., 2014). However, there is considerable heterogeneity based on disease progression, for example, cognitive impairment is associated with increased mortality (Macleod et al., 2014), with ~80% of PD patients eventually developing dementia (Buter et al., 2008; Hely et al., 2008).

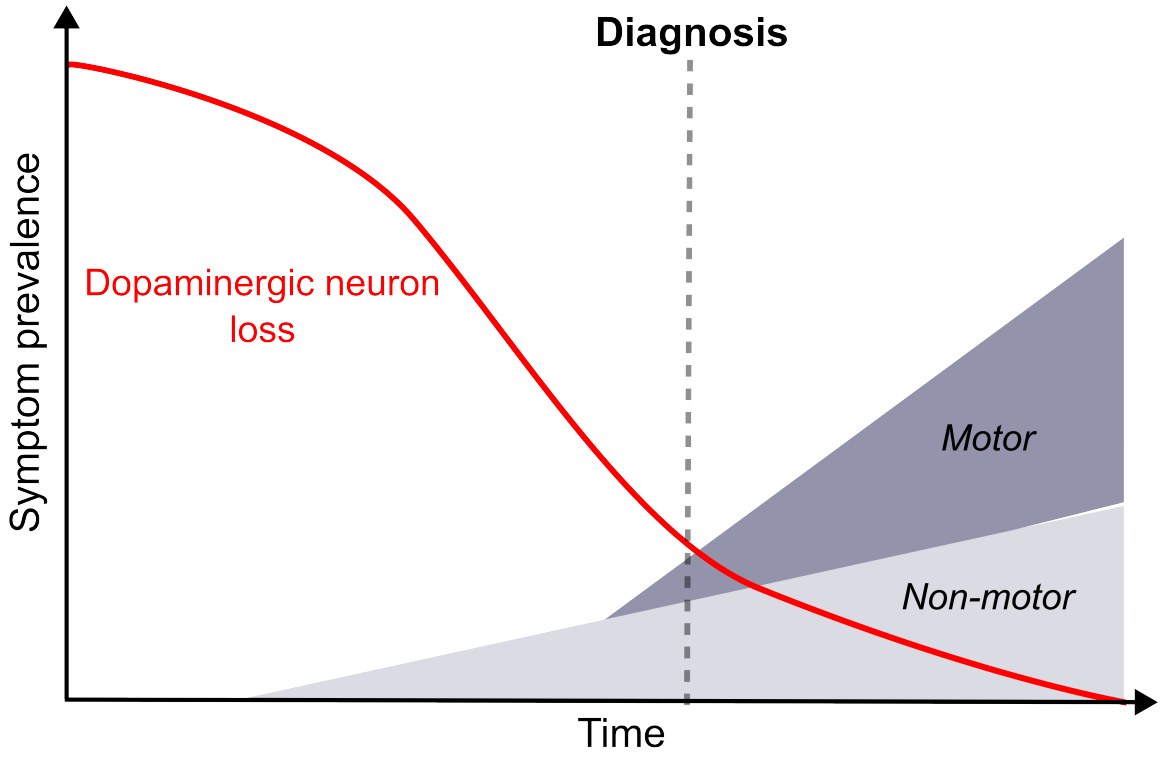


Figure 1.1. Development of symptoms following the onset of neurodegeneration associated with Parkinson's disease. PD occurs due to dopaminergic neuron loss. Diagnosis is made after the presentation of motor symptoms. Before diagnosis, there is a prodromal phase in which non-motor symptoms may present. Adapted from (Kalia & Lang, 2015).

The variability in clinical phenotypes, symptoms, age of onset, progression, and prognosis among PD patients suggests the existence of multiple clinical subtypes. Early classifications grouped PD patients based on motor symptoms such as postural instability gait dysfunction/akinetic-rigid (PIGD/AR) or tremor-dominant (TD) (Zetusky et al., 1985; Jankovic et al., 1990; Schiess et al., 2000). PD patients categorised as TD generally exhibit more benign disease progression than other subtypes (Rajput et al., 2009). Later subgroups such as mild-motor predominant, diffuse malignant and intermediate, were proposed based on data relating to demographics, genetic information, and symptomatic features (Fereshtehnejad, Zeighami, et al., 2017). PD patients classed as diffuse-malignant exhibit faster disease progression (Fereshtehnejad, Zeighami, et al., 2017; Johansson et al., 2023).

PD can only be confirmed upon post-mortem neuropathological investigation. The antemortem clinical diagnosis of PD is based on symptom presentation and medical history. Traditionally, the most common diagnostic criteria used were the UK Brain Bank criteria (Hughes et al., 1992). These criteria required the presence of bradykinesia and at least one other major motor symptom (rigidity, resting tremor or postural instability), along with supporting features and the absence of exclusionary criteria. In 2015, new diagnostic criteria were proposed by the

Movement Disorder Society (Postuma, Berg, et al., 2015). While the motor symptom prerequisites remained largely the same, the new criteria reflected a greater recognition of non-motor symptoms (Postuma, Berg, et al., 2015). In cases where a diagnosis is inconclusive, imaging methods can be used (**Section 1.2.3**) (Pagano et al., 2016). In the UK, the National Institute for Health and Care Excellence (NICE) guidelines do not recommend the use of imaging methods when diagnosing PD; however, they may be considered to distinguish between essential tremor or other parkinsonism disorders.

1.1.2 Epidemiology and aetiology

Estimates from the Global Burden of disease study place the prevalence of PD at 11,800,000 globally, an increase of 273.9% since 1990 (Steinmetz et al., 2024). This makes PD the fastest growing neurological condition in the world and second most common behind Alzheimer's disease (Dorsey, Sherer, et al., 2018). The biggest risk factor for PD development is age (Reeve et al., 2014), with the age of diagnosis peaking between the ages of 85-89 (Dorsey, Elbaz, et al., 2018). There are also notable sex differences in PD; males have around 1.5x the risk of developing PD (Wooten et al., 2004), yet females typically manifest faster disease progression and poorer prognosis (Cerri et al., 2019).

Most PD cases are sporadic with no family history of the disease (Tran et al., 2020). In cases where there is no known genetic or environmental cause, the disease is termed idiopathic. Despite this, there are significant genetic contributions to PD risk. Monozygotic twins have higher PD diagnosis concordance rates than dizygotic twins (Piccini et al., 1999; Wirdefeldt et al., 2011) and ~15% of patients report a family history of PD (Tran et al., 2020). Those with a family history of the disease are around 3-4x the risk compared with individuals with no familial link (Noyce et al., 2012). Genetic PD cases are linked to the co-segregation of pathogenic variants, with up to ~30% of familial PD cases consistent with Mendelian inheritance patterns (Klein & Westenberger, 2012). These highly penetrant variants are sometimes referred to as monogenic PD genes, however variable penetrance and clinical presentations attributed to the same variants implicate the involvement of other factors (Schulte & Gasser, 2011).

PD segregation within families was first recognised in the 1990s (Golbe et al., 1990), with the *SNCA* locus, encoding α-synuclein, eventually identified as the causal gene (Polymeropoulos et al., 1997). Several genes are now firmly linked to familial PD including *LRRK2*, *PINK1*, *PRKN*, *VPS35* and *PARK7* (Kim & Alcalay, 2017). Pathogenic variants in *PINK1*, *PRKN* and *PARK7* result in PD transmission patterns consistent with autosomal recessive inheritance, while pathogenic variants in *SNCA*, *LRRK2* and *VPS35* typically lead to autosomal dominant

PD (Reed et al., 2019). Additionally, variants in *GBA* have been linked to PD (Gegg & Schapira, 2018). As the inheritance of *GBA*-associated PD does not follow Mendelian inheritance patterns, it is commonly thought of as a major genetic risk factor. Individuals with genetic PD exhibit clinical heterogeneity. Autosomal recessive inherited PD is associated with an earlier age of onset than autosomal dominant inherited PD (Vollstedt et al., 2023). Despite this, autosomal recessive PD is usually associated with mild disease progression (Aasly, 2020).

Other genes linked to familial PD in several families have been reported such as *TMEM230*, *DNAJC6* and *CBCHD2* (Funayama et al., 2015; Olgiati et al., 2016; Deng et al., 2016). Recent work leveraging large sequencing datasets has identified rare variants in known PD genes such as *GBA* and *LRRK2* (Makarious et al., 2023; Pitz et al., 2024), as well as uncovering new potentially causative genes such as *RAB32* (Gustavsson et al., 2024; Hop et al., 2024).

Genome wide association studies (GWAS) have provided insights into the role of common genetic variation in shaping complex disease risk. In PD, GWAS have mostly been performed using individuals of European genetic ancestry (Fung et al., 2006; Simón-Sánchez et al., 2009; Hamza et al., 2010; The UK Parkinson's Disease Consortium and The Wellcome Trust Case Control Consortium 2, 2011; Chang et al., 2017; International Parkinson Disease Genomics Consortium, 2011; Nalls et al., 2019). Nalls et al identified 90 independent risk loci within 78 genomic regions accounting for an estimated 16-36% of heritable risk (Nalls et al., 2019). Risk loci in the genes *SNCA*, *LRRK2* and *GBA*, show pleiotropic and mechanistic overlap between idiopathic and familial PD. Recent GWAS have identified novel PD risk loci by including more diverse genetic ancestries (Foo et al., 2020; Rajan et al., 2020; Loesch et al., 2021; Pan et al., 2023; Rizig et al., 2023).

Alongside genetic variation, environmental factors contribute to an individual's susceptibility to PD. Epidemiological studies have identified an association between pesticide exposure and PD risk (van der Mark et al., 2012). Pesticide screening revealed the toxic and additive effects of pesticide exposure on dopaminergic neurons derived from PD patients induced pluripotent stem cells (iPSCs) (Paul et al., 2023). Other environmental factors associated with PD risk include heavy metals, pollution and head trauma (Crane et al., 2016; Cole-Hunter et al., 2023; Zhao et al., 2023). On the other hand, several environmental factors have been associated with protective effects against PD. Smoking has been widely reported to be protective against PD, with recent studies identifying a dose-response relationship (Gallo et al., 2019; Mappin-Kasirer et al., 2020; Rose et al., 2024). Additionally, inverse associations between caffeine intake and PD susceptibility have been reported (Costa et al., 2010; Zhao et al., 2024). Overall, PD is a multifactorial and complex disorder with extensive evidence of genetic and environmental

contributions to risk. Environmental exposures can be integrated with genomic information to provide a model by which gene-environment interactions shape PD susceptibility (Tsalenchuk et al., 2023; Reynoso et al., 2024; K. J. Ngo et al., 2024; Huang et al., 2024).

The molecular mechanisms leading to PD are complex. Genetic and environmental factors influencing PD risk have highlighted common dysregulated processes such as mitochondrial function, synaptic function, endolysosomal function and immune-related responses (Morris et al., 2024). These pathways are not mutually exclusive, with considerable mechanistic overlaps (Coukos & Krainc, 2024).

1.1.3 Pathology

PD is associated with degeneration of the nigrostriatal pathway, a key component of the basal ganglia motor circuit. Dopaminergic neurons axons in the *substantia nigra pars compacta* project to the striatum. In PD, dopaminergic neurons in the *substantia nigra pars compacta* are preferentially lost, leading to dysregulation of the motor circuitry. Here, dopaminergic neurons release dopamine which can bind to receptors on GABAergic medium spiny neurons to control movement. Neurodegeneration can also occur in other areas of the brain such as the locus coeruleus, affecting the adrenergic system (Gesi et al., 2000) and the *dorsal raphe nucleus*, dysregulating serotonergic signalling (Politis & Loane, 2011).

Lewy body pathology is also often observed during neuropathological inspection. Lewy bodies (Lewy neurites when located in neural projections) are toxic aggregates of the protein, α-synuclein. α-synuclein is involved in various cellular functions including neurotransmitter release, synaptic vesicle regulation, RNA processing and immune responses (Burré, 2015; Hallacli et al., 2022; Alam et al., 2022). Misfolded α-synuclein forms oligomers that can aggregate into amyloid fibrils and Lewy pathology. The reason for the tendency of α-synuclein to misfold and form toxic aggregates is unclear, yet pathological α-synuclein can spread from cell to cell (Volpicelli-Daley et al., 2011; Luk et al., 2012), contributing to the progressive development of Lewy pathology in PD.

Several models explain the onset and progression of Lewy pathology. The Braak model suggests that Lewy pathology originates in the olfactory bulb or the brain stem from the gut via the vagus nerve, at which point it begins to spread into the surrounding brain regions (Braak et al., 2002; Braak, Del Tredici, et al., 2003; Braak, Rüb, et al., 2003). The Unified Staging System for Lewy body disorders (USSLB) proposes that Lewy pathology typically starts in the olfactory bulb and spreads through the brainstem, limbic system and neocortical regions (Beach

et al., 2009). A newer model unifies the onset of Lewy pathology proposed in both models through two subtypes, brain-first and body-first (Borghammer & Van Den Berge, 2019).

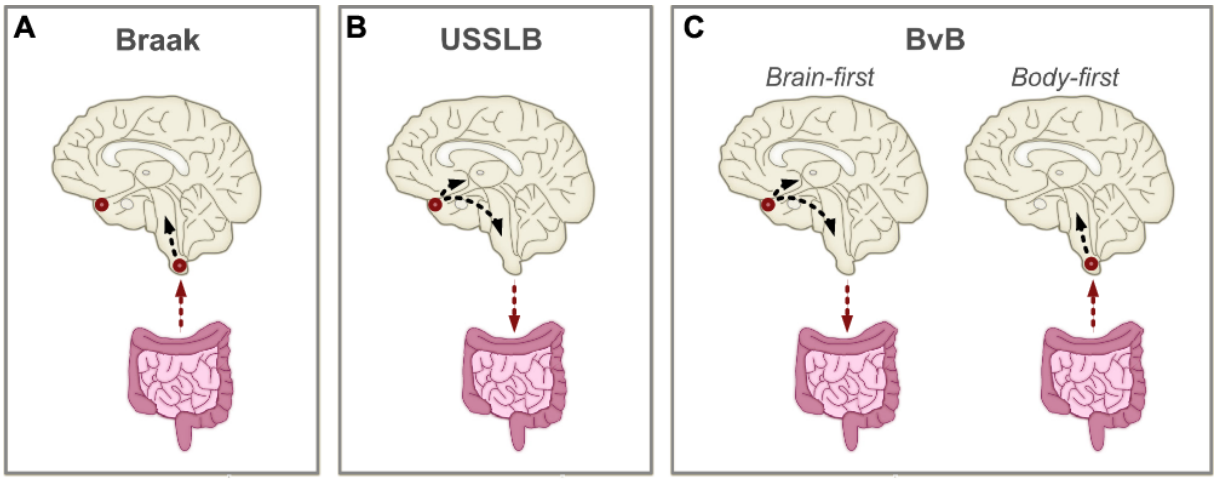


Figure 1.2. Models of Lewy pathology onset and progression. The red dot indicates the onset of Lewy pathology according to each model. USSLB = Unified Staging System for Lewy Body disorders. BvB = Brain-first versus Body-first. Taken from (Mastenbroek et al., 2024).

1.1.4 Treatment

Current pharmacological PD treatments address symptom management. Treatments often target the dopaminergic system at multiple points to address motor dysfunction, raising dopamine levels to counter the loss of dopaminergic neurons (**Figure 1.3**). The current gold standard treatment is levodopa (L-DOPA), first introduced in the 1960s (Connolly & Lang, 2014). The precursor to dopamine is L-DOPA. Tyrosine hydroxylase (TH) catalyses the endogenous synthesis of L-DOPA from tyrosine. Alternatively, L-DOPA is provided therapeutically where it traverses the blood-brain barrier supplying dopamine precursors to the central nervous system (CNS). The conversion of L-DOPA to dopamine is catalysed by DOPA decarboxylase (DDC). Alternatively, catechol-O-methyltransferase (COMT) catalyses the conversion of L-DOPA to 3-O-Methyldopa (3-OMD). As such, L-DOPA is usually administered with peripheral DDC and COMT inhibitors to increase L-DOPA availability in the CNS and side effects caused by increased dopamine abundance to peripheral dopaminergic neurons (Diederich et al., 2008). Additionally, Monoamine Oxidase-B (MAOB) inhibitors can also be given as adjuncts to L-DOPA therapy, whereby they reduce the MAO-B mediated metabolism of synaptically released dopamine to 3,4-Dihydroxyphenylacetic acid (DOPAC) (Dezsi & Vecsei, 2017). Additionally, dopamine agonists can bind to and activate post-synaptic dopamine receptors located on

medium-spiny neurons, thereby continuing normal striatal neuronal signalling (Isaacson et al., 2023). Long-term dopaminergic replacement therapies can lead to fluctuations in efficacy and dyskinesia (Thanvi et al., 2007). In these situations, deep brain stimulation of the subthalamic nucleus by electrical impulses has emerged to manage motor symptoms (Groiss et al., 2009). Treatment for non-motor symptoms is again largely pharmacological and specific to each symptom (e.g., selective serotonin reuptake inhibitors for depression and anxiety).

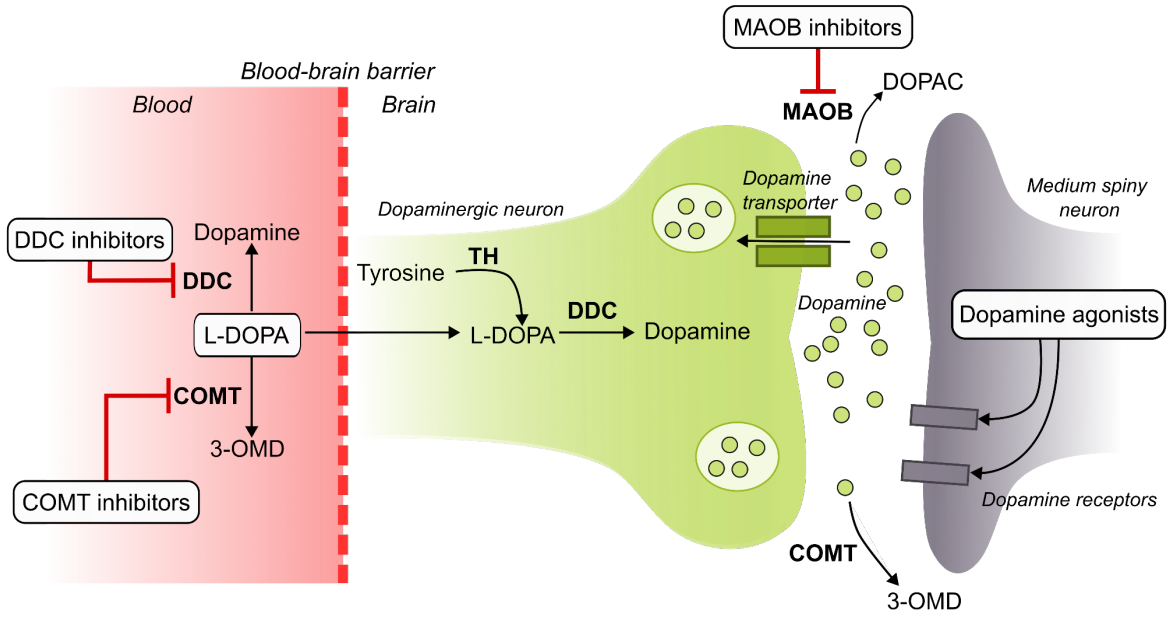


Figure 1.3. Pharmacological PD treatments targeting the dopaminergic system. DDC = DOPA decarboxylase, COMT = Catechol-O-methyltransferase, 3-OMD = 3-O-Methyldopa, TH = Tyrosine hydroxylase, MAOB = Monoamine Oxidase-B. Adapted from (Poewe et al., 2017).

Disease-modifying treatments aim to delay, slow, or revert disease progression. There are currently no disease-modifying treatments in clinical use for PD. In a 2023 review of registered clinical trials, 63/139 (45.3%) aimed to assess disease-modifying treatments (McFarthing et al., 2023). Prasinezumab and Cinpanemab are monoclonal antibodies targeting aggregated α-synuclein (Schenk et al., 2017; Jankovic et al., 2018; Brys et al., 2019). Phase 2 studies investigating the effects of Prasinezumab and Cinpanemab showed no difference in their primary outcomes compared to the placebo group (Lang Anthony E. et al., 2022; Pagano Gennaro et al., 2022). In the case of Prasinezumab, there was a slight improvement in the MDS-Unified Parkinson's Disease Rating Scale Part III (MDS-UPDRS III), assessing motor symptom severity (Pagano Gennaro et al., 2022). In a post-hoc analysis, Prasinezumab

demonstrated greater effects in slowing motor symptom progression (based on MDS-UPDRS III scores) in rapidly progressing PD subgroups (Pagano et al., 2024).

1.1.5 Peripheral involvement in PD

Whilst considered a neurodegenerative disease, a large body of work supports the presence of peripheral dysfunction in PD. Aggregated α-synuclein has been detected in areas outside of the CNS such as the peripheral nervous system and skin (Beach et al., 2010; Comi et al., 2014; Z. Wang et al., 2021). Aggregations of α-synuclein have also been found in the gastrointestinal tract (GI) (Shannon et al., 2012; Sánchez-Ferro et al., 2015), consistent with prevalent GI symptoms in PD. In some cases, GI symptoms are present decades before PD diagnosis. According to the Braak hypothesis of Lewy body staging, α-synuclein aggregation may originate in the GI (Braak, Del Tredici, et al., 2003). The proposal of "body-first" PD subtypes is again consistent with a possible GI origin of Lewy body pathology (Borghammer & Van Den Berge, 2019). Associations between inflammatory bowel disease (IBD) and PD have been reported (Lin et al., 2016; Weimers et al., 2019), although a later mendelian randomisation study found no putative causal relationship between IBD and PD (Freuer & Meisinger, 2022). GWAS have revealed that *LRRK2* is a risk locus for both PD and IBD (Hui et al., 2018; Herrick & Tansey, 2021). IBD may be accompanied by increased pathological α-synuclein aggregation in the GI, also observed in IBD rodent models exhibiting dopaminergic neuron degeneration (Espinosa-Oliva et al., 2024). Inflammation, altered permeability of the intestinal epithelial barrier and subsequent gut dysbiosis have been hypothesised as contributing factors in both pathologies (Houser & Tansey, 2017; Romano et al., 2021).

Immune related loci have been identified as influencing PD risk through GWAS. Several risk loci map to the human leukocyte antigen (HLA) region (Hamza et al., 2010; Hill-Burns et al., 2011; Chang et al., 2017; Nalls et al., 2019), with fine mapping implicating HLA-DRB1*04 as driving a protective effect in PD susceptibility (Yu et al., 2021). Initial work described reactive microglia and elevated inflammatory cytokine levels in sites of PD-associated neurodegeneration (McGeer et al., 1988; Mogi et al., 1994). There is also evidence identifying changes in the peripheral immune system in PD. Blood inflammatory cytokine levels are increased in PD patients and linked to more rapid disease progression (Qin et al., 2016; Williams-Gray et al., 2016). Alterations in immune cell proportions and functions may drive increased inflammation. Classical monocytes are associated with pro-inflammatory effects (Yang et al., 2014). PD patients exhibit increased classical monocyte proportions (Grozdanov et al., 2014; Wijeyekoon et al., 2020). Classical monocytes also exhibit transcriptional

differences in PD patients, particularly in genes related to mitochondrial and lysosomal functioning and monocyte-specific expression quantitative trait loci (Y. I. Li et al., 2019; Navarro et al., 2021).

Regarding other types of blood cells, lower lymphocyte levels have also been reported in PD patients, including CD4+ T cells and CD19+ B cells (Stevens et al., 2012). Reduced general lymphocyte levels are associated with an increased risk of developing PD (Jensen et al., 2021). Conversely, neutrophil levels are increased in PD patients (Craig et al., 2021). The neutrophil-to-lymphocyte ratio has been proposed as a measure of the balance between innate and adaptive immune responses (Song et al., 2021). PD patients have higher neutrophil-to-lymphocyte ratios (Muñoz-Delgado et al., 2021; Grillo et al., 2023; Hosseini et al., 2023) with additional work suggesting increased neutrophil-to-lymphocyte ratios in idiopathic and *GBA*-PD, but not *LRRK2*-PD (Muñoz-Delgado et al., 2023). α-synuclein may be an antigenic target for blood T-cell subsets in some PD patients (Sulzer et al., 2017). In these patients, T-cell α-synuclein reactivity peaked around PD diagnosis and could be detected before diagnosis (Lindestam Arlehamn et al., 2020). However, not all PD patients show this T-cell reactivity, as transcriptional differences exist between those with T-cells responsive to specific α-synuclein antigens and those without (Dhanwani et al., 2022).

1.2 Biomarker use in PD

Biomarkers are measurable signals that reflect a biological process or state. Biomarkers provide clinical utility in the context of disease by indicating the presence of disease (diagnostic biomarkers), the progression of disease (progression biomarkers) or the expected prognosis (prognosis biomarkers).

In situations where there is a binary outcome (i.e., presence or absence of disease), biomarker performance is assessed by comparing the predicted class produced by the biomarker against the actual class. When investigating PD, the class will often refer to whether a sample is from a PD patient or a control. Each prediction is grouped (true positive, true negative, false positive or false negative), depending on the predicted and actual class (**Figure 1.4a**). From this, the sensitivity (ability of the test to correctly classify the sample as PD) and specificity (ability of the test to correctly classify the sample as control) of a test are calculated (**Figure 1.4a**). There is often a trade-off between the sensitivity and specificity of a test. To assess performance, receiver operator characteristic (ROC) curves can be constructed (**Figure 1.4b**). ROC curves show the relationship between the sensitivity and specificity at different prediction thresholds. The area under the ROC curve (AUC) is used to summarise the ability of the test to distinguish PD samples from controls. In this case, an AUC of 0.5 represents random chance while an AUC of 1 represents a perfect ability to distinguish PD from controls. Reporting AUCs provides a single numerical value to compare results across different tests and study populations. Performance expectations of a test vary by clinical context. For example, recent guidelines for fluid biomarkers in Alzheimer's disease, recommend higher specificity for confirmatory tests than triaging tests (Schindler et al., 2024).

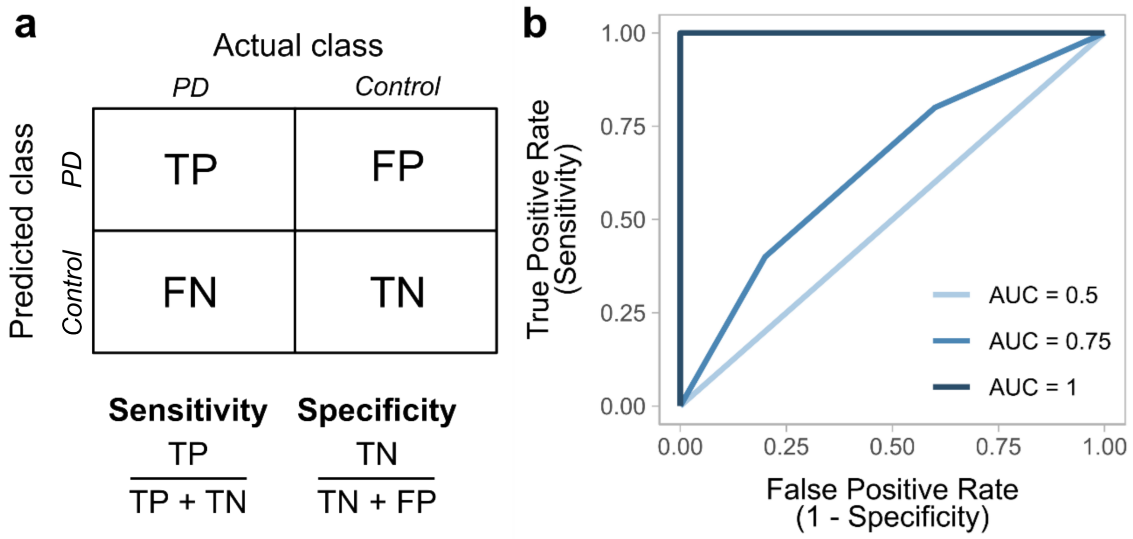


Figure 1.4. Evaluating diagnostic PD biomarker performance. (a) A confusion matrix visualising test predictions with binary outcomes. **(b)** Receiver-operator characteristic (ROC) curves with varying performance based on the area under the curve (AUC). TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

Identifying diagnostic, progression and prognostic biomarkers is crucial for advancing PD research and clinical management. These biomarker categories can overlap, offering multiple benefits. Currently, PD diagnosis relies heavily on motor symptoms which appear after significant neurodegeneration (Jankovic, 2008) (**Section 1.1.1**). Diagnosis accuracy varies with clinical expertise and disease stage (Rizzo et al., 2016). Movement disorder specialists diagnose PD more accurately (Schiess et al., 2022). Early-stage misdiagnosis can occur due to phenotypically similar conditions such as progressive supranuclear palsy and multiple system atrophy (Beach & Adler, 2018).

Biomarkers that can specifically identify PD patients would improve diagnostic confidence and enable earlier diagnosis. Neuropathological studies suggest about 30% of dopaminergic neurons in the substantia nigra are lost by the time motor symptoms appear, indicating neurodegeneration starts years earlier (Fearnley & Lees, 1991; Greffard et al., 2006). Early diagnosis would allow earlier enrolment in clinical trials, potentially in the prodromal stage, allowing the assessment of disease-modifying treatment efficacy earlier in neurodegeneration. Once developed, the efficacy of disease-modifying treatments may be improved by providing a larger therapeutic window (Vijiaratnam et al., 2021). Identification of neurodegeneration in the prodromal stage also informs disease risk, allowing clinical management and planning to be initiated sooner.

PD patients exhibit clinical heterogeneity in motor and cognitive outcomes (Greenland et al., 2019). Progression biomarkers providing objective measures of underlying disease processes over time would help inform PD heterogeneity. This would be particularly beneficial for assessing treatment efficacy, for example as a clinical endpoint when assessing disease-modifying interventions. Progression biomarkers may also assist in PD patient stratification. Disease heterogeneity has likely contributed to the failure to identify effective PD-modifying treatments (Athauda & Foltynie, 2016). For example, testing the therapeutic use of LRRK2 inhibition (Jennings et al., 2022) can be targeted in PD patients with increased LRRK2 activity (Naaldijk et al., 2024). Patient stratification, for example by the prediction of clinical trajectories at diagnosis, would enable future precision medicine approaches.

A single biomarker would ideally possess the capabilities needed for diagnosing, tracking progression, and prognosticating PD. However, as exemplified by Alzheimer's disease, multiple biomarkers may be required (Salvadó et al., 2024). Integrating various data modalities (such as those described below) may be also necessary for clinical utility (Caspell-Garcia et al., 2017; Schrag et al., 2017; Makarious et al., 2022). Ultimately, biomarkers are key to emerging PD biological definitions (Höglinger et al., 2024; Simuni et al., 2024).

1.2.1 Clinical biomarkers

Clinical measures are currently often used as proxies for underlying PD pathology, particularly when assessing disease progression. The Unified Progression Disease Rating Scale (UPDRS) was introduced to measure impairment and disability in PD patients and is among the most widely used clinical rating scales (Ramaker et al., 2002). Later updates to the UPDRS sponsored by the Movement Disorder Society (MDS) were published (MDS-UPDRS) comprising four sections (non-motor experiences of daily living, motor experiences of daily living, motor examination and motor complications) (Goetz et al., 2008). UPDRS and MDS-UPDRS scores generally increase over the disease course (Simuni et al., 2018; Holden et al., 2018; Arshad et al., 2023). MDS-UPDRS scores are often used as an indication of disease progression in research, such as clinical endpoints in clinical trials assessing disease-modifying treatments (Lang Anthony E. et al., 2022; Pagano Gennaro et al., 2022) or when evaluating novel progression biomarkers (Sotirakis et al., 2023).

Given the prevalence of cognitive impairment in PD (Aarsland et al., 2021), clinical measures of general cognitive function are often utilised, such as the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA). Compared to MMSE, MoCA has been shown to offer improved predictive ability for identifying individuals with mild cognitive

impairment (MCI) compared to cognitively unimpaired controls (Ciesielska et al., 2016; Siqueira et al., 2019). The identification of MCI in PD patients can inform future prognosis, with increased dementia risk for patients with MCI (Pigott et al., 2015; Hoogland et al., 2017).

A positive response to L-DOPA is usually a part of clinical PD diagnostic criteria (Hughes et al., 1992; Postuma, Berg, et al., 2015). However, the response to L-DOPA can vary among PD patients, with one systematic review finding that among 119 PD patients, 26.9% had no response to L-DOPA, making it unsuitable as a diagnostic biomarker (Pitz et al., 2020).

1.2.2 Phenotypic biomarkers

As a potential induction site for α-synuclein pathology (Braak, Del Tredici, et al., 2003), the olfactory bulb and its role in processing odorant stimuli may be key to identifying PD earlier. Indeed, olfactory dysfunction has long been recognised as a feature of PD (Doty et al., 1988). In one multi-centre study involving 400 PD patients, >96% had olfactory dysfunction (Haehner et al., 2009). Objective measurement is essential as PD patients often have poor self-assessment of their olfactory function (White et al., 2016). The University of Pennsylvania Smell Identification Test (UPSIT) is the most widely used measurement of olfactory function (Doty et al., 1984). UPSIT scores have been evaluated for their ability to distinguish PD patients from controls with one systematic review of 20 studies identifying sensitivities ranging from 62-92% and specificities ranging from 65-96% (Nielsen et al., 2018). There have been further suggestions that responses to specific odours in the test may be more informative (Gerkin et al., 2017; Bestwick et al., 2021). However, a later study found these specific responses offered no predictive benefit in a prodromal cohort where conversion to PD had been tracked longitudinally (Vaswani et al., 2023). In multivariable PD classification models, hyposmia (based on UPSIT scores) has emerged as one of the most predictive variables (Nalls et al., 2015; Li et al., 2022). Total UPSIT scores had AUCs ranging from 0.88 to 0.99 across 6 different cohorts (Nalls et al., 2015). A later study confirmed these findings in PPMI (AUC = 0.9) and extended them to an additional cohort, DeNoPa (AUC = 0.89) (Li et al., 2022).

Rapid eye movement sleep behaviour disorder (RBD) involves abnormal sleep behaviour, such as the physical enactment of dreams, due to a loss of normal muscle paralysis during sleep (Dauvilliers et al., 2018). RBD frequently occurs in PD patients with up to 86% of patients eventually developing it (Sixel-Döring et al., 2023). It is also a common prodromal marker of synucleinopathy risk (PD, Dementia with Lewy bodies (DLB), multiple system atrophy (MSA)) (Postuma, 2014). A longitudinal study of 44 RBD patients found that 92.5% had a neurodegenerative disease 14 years after RBD diagnosis (Iranzo et al., 2013). One study of 121

individuals with RBD found that the presence of RBD among other prodromal criteria (Berg et al., 2015) identified future conversion to Lewy body disease (after 4 years) with a sensitivity of 81.3% and specificity of 67.9% (Fereshtehnejad, Montplaisir, et al., 2017).

1.2.3 Imaging biomarkers

Imaging techniques visualise structural and functional changes associated with PD neurodegeneration. T1-based structural magnetic resonance imaging (MRI) showed atrophy in the midbrain and basal ganglia in early-stage PD patients compared to controls, which also correlated to motor and cognitive scores in follow-up work (Zeighami et al., 2015, 2019). Neuromelanin-sensitive MRI visualises neurodegeneration in the substantia nigra and locus coeruleus, which normally contain abundant neuromelanin. A meta-analysis of 12 studies reported a sensitivity of 89% and a specificity of 83% for using neuromelanin-sensitive MRI to distinguish PD patients from controls (Cho et al., 2021). Additionally, reduced neuromelanin-MRI signals in individuals with RBD indicate a potential prognostic use (Knudsen et al., 2018). Longitudinal studies have reported a reduction in neuromelanin signal in PD patients, suggesting that neuromelanin-sensitive MRI may be useful for tracking disease progression (Matsuura et al., 2016; Gaurav et al., 2021).

Dopaminergic function can be assessed using radioligand signals with single-photon emission computed tomography (SPECT) or positron emission tomography (PET). Radioligands often target dopamine transporters (DAT) to signal presynaptic dopaminergic dysfunction (Thobois et al., 2019). The use of Ioflupane (^{123}I) is used in SPECT imaging, known as DaT scan, to distinguish PD from vascular or drug-induced PD, or essential tremor (Jennings et al., 2004; Benamer et al., 2000; Brigo et al., 2014). However, DAT binding cannot distinguish PD from other parkinsonism disorders with similar nigrostriatal degeneration (Kim et al., 2002; Cilia et al., 2005; Cummings et al., 2011). Reductions in DAT binding are observed in individuals with RBD and non-manifesting carriers of pathogenic *LRRK2* variants (Pont-Sunyer et al., 2017; Iranzo et al., 2020; Simuni, Uribe, et al., 2020). Although longitudinal reductions in DAT binding are seen in PD patients, these reductions may plateau around four years post-diagnosis, limiting the use of DAT imaging as a progression biomarker (Kordower et al., 2013; Simuni et al., 2018). Changes in DAT binding over time may also be confounded by dopamine replacement therapy (Merchant et al., 2019).

Imaging techniques have also focused on areas other than the brain. Cardiac sympathetic nerve dysfunction in PD can be visualized using ^{123}I-metaiodobenzylguanidine (MIBG) scintigraphy (Orimo et al., 2007). A meta-analysis of 19 studies found the sensitivity and specificity of

MIBG scintigraphy for distinguishing PD patients from controls to be 88% and 85% respectively (Treglia et al., 2012). A multicentre study later suggested that MIBG scintigraphy is better at distinguishing PD patients from controls than DaT Scan (De Feo et al., 2023).

The utility of incorporating imaging biomarkers into research and clinical use has been recognised. Striatal DAT binding assessments were used for participant recruitment in the PPMI (Marek et al., 2018), and several imaging criteria are included in the MDS clinical diagnostic criteria (Postuma, Berg, et al., 2015). However, imaging biomarkers often cannot distinguish between PD and other parkinsonism disorders, and their high cost and need for specialised personnel and equipment limit widespread use. In the US, DaT Scan was approved by the Food and Drug Administration in 2011, but NICE guidelines in the UK do not recommend the routine clinical use of imaging methods for PD diagnosis. Combining multiple imaging methods may improve differentiation between PD and controls, but compounds existing limitations (Depierreux et al., 2021). Cost effectiveness analyses of DaT scan have been performed in the context of the German, Belgian, Italian and healthcare systems, with only the Italian study identifying a cost saving (Dodel et al., 2003; Van Laere et al., 2008; Antonini et al., 2008). As such, biomarkers from accessible biological fluids are attractive for PD research due to their increased cost-effectiveness and scalability.

1.2.4 Biological fluid biomarkers

Biomarkers measured in bodily fluids typically offer a more cost effective and accessible source of biomarkers. The cerebrospinal fluid (CSF) flows within the brain ventricles and the subarachnoid spaces of the cranium and spine (Sakka et al., 2011). It provides essential components for brain function, removes waste products, and protects the brain (Spector et al., 2015). CSF is replenished about four times daily (Sakka et al., 2011), allowing it to reflect physiological changes in the brain, making it a key source of neurodegenerative disease biomarkers. In Alzheimer's disease, several biomarkers detectable in the CSF such as Aβ42/Aβ40 (ratio of β-amyloid 42 to β-amyloid 40) and p-tau181/Aβ42 (ratio of phosphorylated tau at residue 181 to β-amyloid 42) have shown clinical utility in distinguishing patients from controls (Mattsson et al., 2017; Hansson et al., 2018). CSF samples are collected via lumbar puncture, an invasive procedure with potential complications, limiting its use (Evans, 1998). Consequently, more accessible biomarker sources, like blood or saliva, are often explored for widespread use or repeat measurements. Emerging evidence has supported a role for plasma levels of p-tau217 (phosphorylated tau at residue 217) to possess similar performance to CSF Alzheimer's disease biomarkers (Barthélemy et al., 2024). Successfully

identifying diagnostic CSF biomarkers for Alzheimer's disease and the evolution of more accessible measurements in blood raises the hope of finding similar biomarkers for PD.

Rare, pathogenic variants linked to PD development (**Section 1.1.2**) can inform an individual's risk of developing PD but due to incomplete penetrance, do not always lead to PD. Aggregating common variants associated with PD into polygenic risk scores can identify differences in PD risk at the population level, yet lack utility for individual prediction (Nalls et al., 2019; Koch et al., 2021). Molecules reflecting dynamic cellular processes, such as proteins, metabolites and RNA offer potentially exploitable biomarkers.

Given its pathological aggregation in PD, α-synuclein has been extensively assessed as a biomarker. α-synuclein is detectable in various tissues such as the CSF, blood, saliva and skin (Fjorback et al., 2007; Devic et al., 2011; Mollenhauer et al., 2019; Han et al., 2022), indicating secretion mechanisms (Marques & Outeiro, 2012). However, measuring α-synuclein in CSF is challenging due to blood contamination (Barbour et al., 2008; Barkovits et al., 2020), leading to variability across studies (Magalhães & Lashuel, 2022). Meta-analyses have reported reduced total α-synuclein levels in the CSF in PD patients (Eusebi et al., 2017; Magalhães & Lashuel, 2022). In blood, measurements are often limited to serum or plasma because of high erythrocyte abundance (Barbour et al., 2008), with studies showing conflicting results regarding total α-synuclein levels in PD patients (Magalhães & Lashuel, 2022). Specific α-synuclein species have also been evaluated as diagnostic PD biomarkers. Oligomeric α-synuclein levels are higher in PD patients in both CSF and blood samples (Zhou et al., 2015; Wang et al., 2015; Williams et al., 2016; Eusebi et al., 2017; Daniele et al., 2018). Phosphorylated α-synuclein at residue 129, which constitutes over 90% of Lewy body α-synuclein (Oueslati, 2016), shows increased levels in CSF but not in serum or plasma of PD patients (Eusebi et al., 2017; Zubelzu et al., 2022).

Parallels have been drawn between pathogenic α-synuclein and prions due to their self-propagating abilities (Steiner et al., 2018). Prions propagation has been exploited to develop diagnostic tests for Creutzfeldt-Jakob disease (Moda Fabio et al., 2014), such as the protein misfolding cyclic amplification assay (PMCA) and real-time quaking-induced conversion assay (RT-QuIC) (Saborio et al., 2001; Atarashi et al., 2008). When RT-QuIC was first used to detect α-synuclein in CSF, it identified PD patients with 95% sensitivity and 100% specificity (Fairfoul et al., 2016). A larger study using PMCA detected α-synuclein in CSF with 88.5% sensitivity and 96.9% specificity against controls with other neurological and neurodegenerative disorders (Shahnawaz et al., 2017). These and other assays, such as HANdai

Amyloid Burst Inducer (Kakuda et al., 2019), are collectively referred to as α-synuclein seed amplification assays (SAAs), which vary in protocols but follow a similar premise (**Figure 1.5**).

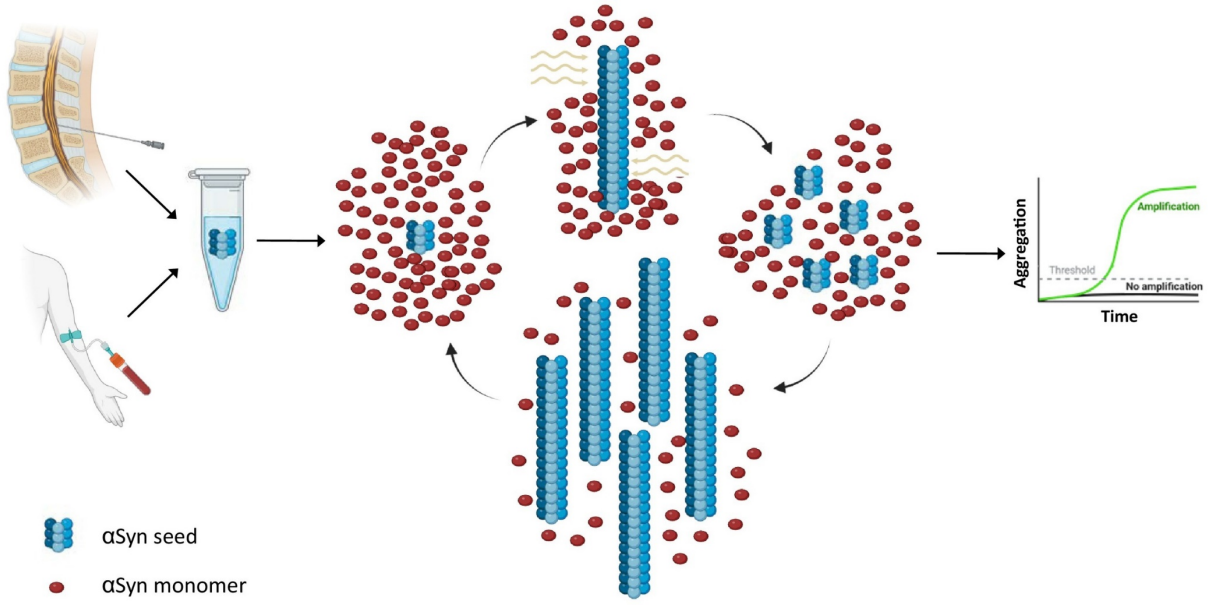


Figure 1.5. α-synuclein seed amplification assay. Biofluid samples are collected (cerebrospinal fluid and blood are represented here). Seed competent α-synuclein induces seeding of α-synuclein monomers added to the sample. This process is amplified through an incubation phase and periodic fragmentation, typically by sonication (PMCA) or shaking (RT-QuIC), to promote the formation of new α-synuclein seeds. α-synuclein aggregation kinetics are measured during this process by fluorescence. Taken from (Soto, 2024).

Subsequent studies have confirmed the diagnostic potential of CSF α-synuclein SAAs in PD across numerous cohorts (Fairfoul et al., 2016; Shahnawaz et al., 2017; Groveman et al., 2018; Rossi et al., 2020; Shahnawaz et al., 2020; Brockmann et al., 2021; Russo et al., 2021; Concha-Marambio et al., 2021; Siderowf et al., 2023). A meta-analysis of 22 studies found α-synuclein SAAs had a pooled sensitivity of 91% and specificity of 96% for distinguishing Lewy body synucleinopathies from non-synucleinopathies (Grossauer et al., 2023). The largest single-cohort study, involving 1123 Parkinson's Progression Marker Initiative (PPMI) participants, reported an overall sensitivity of 87.7% and a specificity of 96.3% for distinguishing PD patients from controls (Siderowf et al., 2023). However, subgroup analysis within the PPMI study showed reduced sensitivity in *LRRK2*-PD patients (67.5%) and PD patients without olfactory dysfunction (63.0%) (Siderowf et al., 2023).

Studies have reported positive α-synuclein SAA results in individuals with RBD, suggesting its potential for early PD identification (Iranzo et al., 2021; Siderowf et al., 2023; Concha-Marambio et al., 2023). SAA kinetics may offer quantitative insights into α-synuclein

aggregation (Shahnawaz et al., 2017; Groveman et al., 2018; Gilboa et al., 2024), demonstrating associations with cognitive impairment (Bräuer et al., 2023; Brockmann et al., 2024) and more general PD severity (Shahnawaz et al., 2017; Vivacqua et al., 2023). Longitudinal studies are needed to establish the clinical utility of SAA as a prognostic or progression PD biomarker. Additionally, α-synuclein SAAs have been tested in more accessible tissues like skin and the olfactory mucosa (De Luca et al., 2019; Manne et al., 2020; Mammana et al., 2021; Donadio et al., 2021; Bargar et al., 2021). In serum samples, α-synuclein SAA distinguished PD patients from controls with AUCs of 0.96 and 0.86 in two cohorts (Okuzumi et al., 2023). Preprocessing to isolate neuronally derived extracellular vesicles in blood has shown promise in identifying individuals at high risk of developing PD (Kluge et al., 2022; Shijun Yan et al., 2023; Kluge et al., 2024).

Proteins other than α-synuclein have also been assessed as PD biomarkers. Neurofilament light chain (NfL) release, associated with axonal damage (Gaetani et al., 2019), is higher in the CSF of PD patients, correlating with motor and cognitive impairment (Lerche et al., 2020). Elevated blood NfL levels may also predict future cognitive impairment risk in PD (Aamodt et al., 2021; Batzu et al., 2022). NfL levels in CSF and blood can help distinguish PD from atypical parkinsonian conditions (Bäckström et al., 2015; Magdalinou et al., 2015; Hansson et al., 2017). However, since increased NfL is a general marker of neurodegeneration, its specificity is limited in the presence of other neurodegenerative diseases (Ashton et al., 2021). Consistent with the role of inflammation in PD, levels of specific inflammatory cytokines and chemokines such as TNF-α, IL-6, IL-1B, MCP-1 and CRP, are increased in the blood and CSF of PD patients (Qu et al., 2023). The ability of inflammatory markers to distinguish PD patients from controls was inconsistent, potentially due to mixed methods of measuring inflammatory marker levels (Qu et al., 2023). Inflammatory processes have been associated with motor and cognitive outcomes and thus may have potential clinical utility as progression or prognostic PD biomarkers (Williams-Gray et al., 2016).

Several proteomic studies have shown increased DDC levels in the CSF of PD patients compared to controls (del Campo et al., 2023; Paslawski et al., 2023; Pereira et al., 2023; Rutledge et al., 2024; Appleton et al., 2024). As shown in **Figure 1.3**, DDC catalyses the conversion of L-Dopa to dopamine. DDC distinguished patients with Lewy body disease (defined as dementia with Lewy bodies and PD patients) from controls with an AUC of 0.89 in CSF and 0.92 in plasma (Pereira et al., 2023). However, a later study indicated that the ability of DDC to distinguish between PD patients and controls in plasma was influenced by dopaminergic treatment status (Appleton et al., 2024).

Metabolic dysfunction has been described in PD and metabolites are commonly exploited as potential biomarkers (Mamas et al., 2011; Anandhan et al., 2017). Blood urate levels are inversely associated with PD risk (Wen et al., 2017; Chang et al., 2022), and higher urate levels correlate with slower disease progression (Schwarzschild et al., 2008; Ascherio et al., 2009). Reduced serum levels of caffeine and associated metabolites have been described (Hatano et al., 2016), consistent with the inverse between caffeine intake and PD risk (**Section 1.1.2**), with the potential ability of caffeine and its metabolite levels to distinguish PD patients from controls (AUC = 0.87) (Fujimaki et al., 2018). Metabolite panels have shown AUCs of 0.80-0.90 in distinguishing PD patients from controls (Burté et al., 2017; Stoessel et al., 2018; Klatt et al., 2021), though require validation in independent cohorts (Shao & Le, 2019). In two studies that evaluated the ability of metabolites to distinguish between PD patients and controls in independent cohorts, AUCs of 0.83 and 0.846 were reported (Saiki et al., 2017; Mallet et al., 2022).

Gene expression changes in CSF, blood, and skin samples between PD patients and controls suggest that mRNA (messenger RNA) could be a source of PD biomarkers (reviewed in (Borrageiro et al., 2018)). However, not all studies tested the ability of mRNA to distinguish PD patients from controls (described in more detail in **Section 3.1**). Initial studies using microarray and quantitative polymerase chain reaction (qPCR) to measure mRNA levels reported AUCs from 0.63 to 0.96 with fewer than 100 PD patients (Scherzer et al., 2007; Molochnikov et al., 2012; Santiago & Potashkin, 2013, 2015; Santiago et al., 2016). Later, larger studies assessed gene expression as a PD biomarker using microarrays (Shamir et al., 2017) and RNA-seq data from large-scale projects such as the PPMI and Parkinson's Disease Biomarker Program (PDBP), involving hundreds of PD patients and controls (Marek et al., 2011). Using gene expression to distinguish PPMI PD patients from controls with AUCs from 0.72 to 0.81 (Makarious et al., 2022; Pantaleo et al., 2022; Li et al., 2023), but an AUC of 0.69 when replicated in PDBP (Li et al., 2023).

The progressive annotation of the human genome has uncovered thousands of non-protein coding genes, including those for encoding microRNAs (miRNAs) (Snyder et al., 2020). miRNAs regulate gene expression by forming the RNA-induced silencing complex, which degrades complementary mRNAs (Hammond et al., 2001; Lai, 2002; Shang et al., 2023). Altered levels of hsa-miR-221-3p, hsa-miR-214-3p and hsa-miR-29c-3p have been reported in the blood of PD patients in multiple studies (Schulz et al., 2019). A large scale study of small non-coding RNAs, predominantly miRNAs, showed dysregulation of molecular PD hallmarks including chronic inflammatory responses, and identified miRNAs associated with disease

progression (Kern et al., 2021). A meta-analysis of reported sensitivities and specificities for miRNA to distinguish PD patients from controls across a variety of biological fluids produced a sensitivity and specificity of 0.82 and 0.80 respectively (Guévremont et al., 2023). miRNAs are seen as attractive biomarker candidates due to their inherent stability (Xi et al., 2007; Mitchell et al., 2008). Another class of RNA, circular RNA (circRNA, **Section 1.3**) is also noted for its inherent stability (Jeck et al., 2013; Enuka et al., 2016) and has been proposed as a source of biomarkers for a range of diseases (Verduci et al., 2021).

1.3 Circular RNA

CircRNAs are largely non-coding, covalently closed loop molecules formed by a back-splicing reaction in which a downstream splice site donor is spliced to an upstream splice site acceptor. The field emerged in the early 2010s (Salzman et al., 2012; Jeck et al., 2013; Hansen et al., 2013; Memczak et al., 2013) and has since expanded rapidly (**Figure 1.6**). However, there is concern over the number of retracted publications and links to 'paper mills' (Bricker-Anthony & Herzog, 2023) (**Figure 1.6**). Given the delay between publication dates and article retractions, the number of retractions will unfortunately grow over the coming years. Thus, developing and following robust research methods for circRNA studies is crucial for reproducibility and consistency (Dodbele et al., 2021; Nielsen et al., 2022).

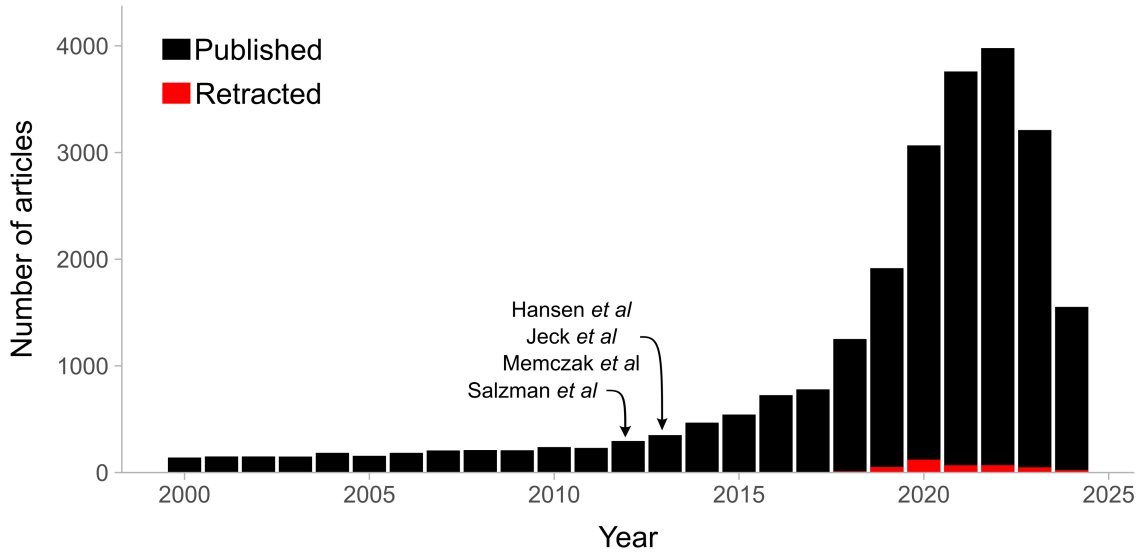


Figure 1.6. Expansion of circular RNA research. Number of articles containing the term *circRNA* deposited in PubMed. Retraced articles were identified using the search term *circRNA retracted*.

Initially, studies utilising electron microscopy reported the presence of circular RNA molecules (Sanger et al., 1976; Kolakofsky, 1976; Hsu & Coca-Prados, 1979). Further work identified mitochondrial-derived circRNAs in yeast (Arnberg et al., 1980) and circular RNA genomes contained within the hepatitis delta virus (Kos et al., 1986). Beginning in the 1990s, circRNAs were first linked to specific human genes. Nigro et al discovered transcripts within the gene *DCC* that exhibited exon ordered differentially to that of the primary transcript, thus providing the first description of "scrambled exons" (Nigro et al., 1991). Scrambled exons were also identified in c-*ets*-1 (now known as *ETS1*), in combination with long flanking introns, non-polyadenylation and low expression (Cocquerelle et al., 1992); aligning with the current

paradigm of circRNA properties (Jeck et al., 2013). In a follow up study, c-*ets*-1 was subsequently shown to be circular (Cocquerelle et al., 1993). At the same time, circular mouse *Sry* transcripts were identified (Capel et al., 1993). Over the next few years, numerous genes were confirmed to also produce circular transcripts including the cytochrome P450 gene in rats and humans (Zaphiropoulos, 1996, 1997), *DMD* (Surono et al., 1999) and *ANRIL* (Burd et al., 2010). Later work identified hundreds of loci producing transcripts in which the spliced exon order did not match the canonical genomic order (Dixon et al., 2005; Al-Balool et al., 2011). The development of RNA-seq library preparations capable of amplifying non-polyadenylated transcripts further clarified the genome-wide production of non-linearly spliced transcripts and the appreciation that many are circular (Al-Balool et al., 2011; Salzman et al., 2012; Jeck et al., 2013; Hansen et al., 2013; Memczak et al., 2013).

1.3.1 Biogenesis

Back-spliced junctions (BSJs) are the key defining feature distinguishing circRNAs from the cognate linear RNA (**Figure 1.7a**). CircRNA production utilises the canonical spliceosome machinery (Starke et al., 2015) so is recognised as a form of alternative splicing. However, the splicing of distal splice junctions to produce circRNAs requires a distinct set of mechanisms beyond the canonical splicing of linear transcripts.

The circRNA life cycle begins with the RNA polymerase II (Pol II) mediated transcription of DNA to form precursor mRNA. Given the links between the transcription elongation rate and splicing (Fong et al., 2014), there are indications at this early stage of factors influencing circRNA production. Faster rates of transcription have been correlated with increased circRNA production (Y. Zhang et al., 2016; Debès et al., 2023). Furthermore, pathogenic variants in Pol II that reduce transcription speed result in reduced circRNA expression (Y. Zhang et al., 2016). Numerous studies have subsequently explored whether circRNA production occurs co- or post-transcriptionally. The detection of circRNAs in sequencing of chromatin-bound RNA suggested some circRNAs are produced co-transcriptionally (Ashwal-Fluss et al., 2014). Supporting this view, disrupting poly(A) sequences at specific circRNA loci diminished circRNA production (Liang & Wilusz, 2014). In follow up work, disrupting polyA sequences at separate circRNA loci had no effect on circRNA production, suggesting that cis-regulatory elements, such as long flanking introns, may promote the co-transcriptional production of circRNA (Kramer et al., 2015). Later work using tagging of newly transcribed RNA identified a proportion of co-transcriptionally produced circRNAs, yet most were produced post-transcriptionally (Y. Zhang et al., 2016). CircRNA production can also be induced by the depletion of spliceosome

components and other splicing factors (Kramer et al., 2015; Liang et al., 2017) further demonstrating how pre-mRNA processing shapes the production of circRNAs.

To produce most circRNAs, distal splice sites need to come into proximity, which can occur through several mechanisms (**Figure 1.7b, c**). Early work recognised that long flanking introns facilitate the production of circular human *c-ets-1* and mouse *Sry* transcripts (Cocquerelle et al., 1992; Dubin et al., 1995). Genome wide profiling of circRNA expression confirmed that circRNA loci were enriched with long flanking introns containing reverse complementary sequences (Jeck et al., 2013). In humans, reverse complementary sequences flanking circRNA loci often reside in primate-specific *Alu* elements (Jeck et al., 2013; Zhang et al., 2014). When positioned in reverse orientations, *Alu* elements base-pair intramolecular due to inverted repeat sequences (IR*Alu*). CircRNA biogenesis is promoted by the formation of IR*Alus* across flanking introns, but not by IR*Alus* within individual introns (Zhang et al., 2014). The competition of IR*Alu* pairing within and across introns contributes to variable circRNA production (Zhang et al., 2014). Additionally, IR*Alu* pairings compete to produce circRNAs with alternative back-spliced junctions (Zhang et al., 2014; X.-O. Zhang et al., 2016). The lack of conserved inverted repeats flanking BSJ loci may explain differences in circRNA expression between humans and mice (Guo et al., 2014; Ji et al., 2019). This is consistent with work supporting that circRNA development is linked to convergent evolution through species-specific transposable elements in flanking regions (Gruhl et al., 2021; Santos-Rodriguez et al., 2021), explaining the increasing complexity of circRNA expression throughout species evolution (Dong et al., 2016).

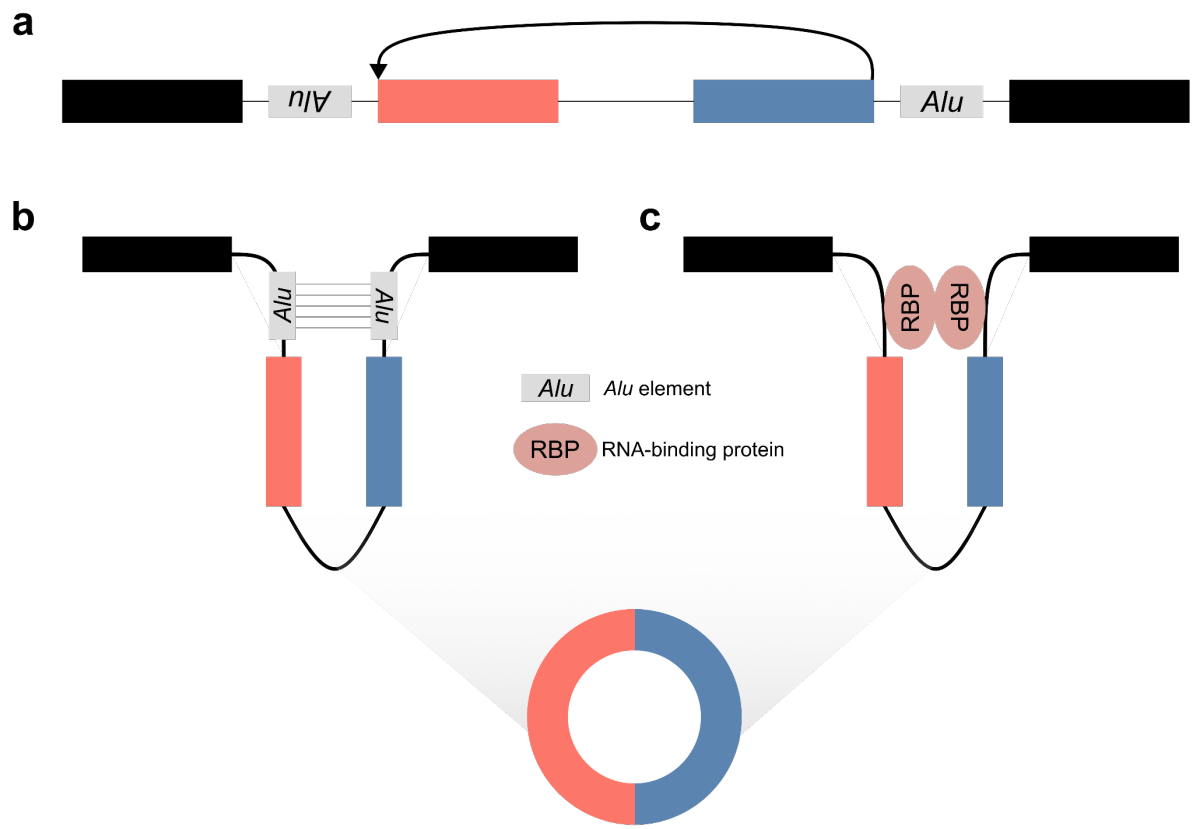


Figure 1.7. Common circular RNA biogenesis mechanisms. (a) Back-splicing of a 3' splice site to an upstream 5' splice site. **(b)** Back-splicing mediated by reverse complementary sequences. In humans, *Alu* elements are a common source of repeat sequences. **(c)** Back-splicing mediated through RNA-binding proteins.

CircRNA expression changes in a cell (Salzman et al., 2013; Wu et al., 2022), tissue (Maass et al., 2017; Xia et al., 2017) and development stage-specific manner (Conn et al., 2015; Mellough et al., 2019) suggest the involvement of additional *trans* regulatory factors. For example, the binding of *trans*-acting RNA-binding proteins (RBPs) can influence circRNA production. The first RBP identified to potentiate circRNA biogenesis was the splicing factor muscleblind (encoded by *mbl*) (Ashwal-Fluss et al., 2014). In *Drosophila melanogaster*, a circRNA produced from the *mbl* locus contained muscleblind motifs, promoting the production of this circRNA (Ashwal-Fluss et al., 2014). Knockdown of the *mbl* locus also affected levels of circRNAs produced from loci other than *mbl*, supporting a wider role of muscleblind, and potentially other RBPs, in circRNA biogenesis (Ashwal-Fluss et al., 2014). Subsequently, quaking was shown to increase general circRNA abundances during epithelial-mesenchymal transition through binding to quaking motifs in the flanking introns of circRNAs (Conn et al., 2015). Other RBPs have been reported to influence circRNA biogenesis through binding to double-stranded RNA (dsRNA) structures, such as those produced by IR*Alus* pairing (Ahmad

et al., 2018). ADAR1 is an enzyme that binds dsRNA, catalysing the conversion of adenine to inosine bases, termed A-I editing (Hogg et al., 2011). In humans, most A-I editing occurs at dsRNA formed by IR*Alus* (Levanon et al., 2004). ADAR1 depletion leads to increased levels of specific circRNAs, possibly through ADAR1-mediated A-I editing destabilising dsRNA structures formed by IR*Alus* (Ivanov et al., 2015; Rybak-Wolf et al., 2015). Another RBP, DHX9, is also thought to repress circRNA biogenesis by destabilising dsRNA structures (Aktaş et al., 2017). However, later work also associated ADAR1 A-I editing with increased biogenesis of certain circRNAs, possibly through stabilising dsRNA structures or editing of RBP binding sites (Shen et al., 2022).

Additional, less well-characterised mechanisms of circRNA biogenesis also exist. In some cases, lariats produced after exon skipping may undergo a back-splicing reaction to form circRNAs (Zaphiropoulos, 1996; Surono et al., 1999; Kelly et al., 2015; Barrett et al., 2015). Furthermore, circular intronic RNAs may be formed when RNA motifs located near the 5' splice site and branchpoint lead to intron lariats escaping debranching and subsequently forming covalently closed molecules via a 2', 5' phosphodiester bond (Zhang et al., 2013). Finally, circRNAs can be produced due to chromosomal translocations (Guarnerio et al., 2016).

1.3.2 Subcellular localisation

Following their formation, subcellular circRNA localisation can depend on the characteristics of the circRNA. CircRNAs that possess retained introns (exon-intron circRNAs, intronic circRNAs) are typically retained in the nucleus (Zhang et al., 2013; Z. Li et al., 2015). Exonic circRNAs, however, are typically localised to the cytoplasm (Jeck et al., 2013; Zhang et al., 2019). Canonical mRNA export from the nucleus is dependent on mRNA processing steps during transcription (Wickramasinghe & Laskey, 2015), so the lack of a 5' cap and polyA tail in circRNA suggests alternative forms of transport out of the nucleus. CircRNA transport out of the nucleus may be passive, such as through the disassembly of the nuclear envelope during mitosis (Güttinger et al., 2009), and maintained in the cytoplasm through their stability (Enuka et al., 2016). Initial work identified that the depletion of two RNA helicases, URH49 (encoded by *DDX39A*) and UAP56 (encoded by *DDX39B*) impaired circRNA export in a size dependent manner (Huang et al., 2018). When URH49 was depleted, circRNAs with a length <356 nucleotides were enriched in the nucleus, while UAP56 depletion led to an enrichment of circRNAs with a length >1298 nucleotides. However, no mechanism for the export of circRNAs of intermediate length was suggested. Furthermore, a separate study failed to recapitulate consistent size-dependent nuclear circRNA accumulation upon URH49 and UAP56 depletion

(L. H. Ngo et al., 2024). Later studies have supported roles for members of the exportin family, exportin-4 (Chen et al., 2022) and exportin-2 (L. H. Ngo et al., 2024) in nuclear circRNA export. Differences in circRNA localisation may reflect specific physiological contexts or circRNA characteristics (Zhang et al., 2019). The binding of the splicing factor SRSF1 leads to the nuclear retention of certain circRNAs (Ron & Ulitsky, 2022). Furthermore, adenosine-rich circRNAs are exported from nuclei to the cytosol by PABPC1 during neural differentiation of H9 pluripotent stem cells (Cao et al., 2024). Finally, the detection of circRNAs within extracellular vesicles additionally suggests they can be released from cells via exocytosis (Y. Li et al., 2015; Lasda & Parker, 2016).

1.3.3 Degradation

Degradation by exoribonucleases begins at the terminal ends of an RNA molecule (Garneau et al., 2007). Owing to their covalently closed structure, circRNAs have increased resistance to exoribonuclease activity (Jeck et al., 2013). This resistance is commonly exploited to experimentally enrich circRNAs through the addition of the 3'-5' exoribonuclease RNase R. CircRNA stability manifests as increased half-lives compared to the cognate linear RNA molecules (Enuka et al., 2016), and may contribute to the enrichment of circRNAs in post-mitotic tissue such as the brain (Westholm et al., 2014; Rybak-Wolf et al., 2015).

However, there are multiple known mechanisms by which circRNA turnover occurs. CircRNAs can be degraded through endoribonuclease activity, which degrades RNAs internally (Tomecki & Dziembowski, 2010). One such endoribonuclease, RNase L, is activated as part of an innate immune response and leads to widespread turnover of multiple RNA species (Chakrabarti et al., 2011; Burke et al., 2019), including circRNAs (C.-X. Liu et al., 2019). The N^6-methyladenosine (m^6A) modification of circRNA has been implicated in several downstream effects such as promoting translation (Yang et al., 2017) and recognising endogenous circRNAs (Chen et al., 2019). Additionally, there is evidence that m^6A-circRNAs are degraded by the endoribonucleases, RNase P and RNase MRP (Park et al., 2019). Another endoribonuclease, G3BP1 (Tourrière et al., 2001), has been reported to degrade highly structured circRNAs in tandem with UPF1 (Fischer et al., 2020). The early characterised circRNA CDR1as/ciRS-7 (Memczak et al., 2013; Hansen et al., 2013), contains an miRNA binding site (miR-671) allowing for the argonaute-2 associated degradation (Hansen et al., 2011). This method of turnover is likely to be highly specific given the lack of circRNAs highly populated with miRNA binding sites (Guo et al., 2014). Finally, the processing-body localised proteins

TNRC6A, TNRC6B, and TNRC6C have been implicated in circRNA degradation, yet the exact mechanism is still unclear (Jia et al., 2019).

1.3.4 Biological relevance

Despite the identification of millions of circRNAs across species (Wu et al., 2024), the function, if any, of the vast majority has not been determined.

RNAs with multiple miRNA binding sites may act as competitive endogenous RNAs (ceRNAs) (Salmena et al., 2011). Initial work recognised binding sites and subsequent cleaving of a circular lncRNA located antisense to *CDR1*, *CDR1as/ciRS*-7, by miR-671 (Hansen et al., 2011). In 2013, two independent labs identified >60 binding sites for another miRNA, miR-7, within *CDR1as/ciRS-7* (Hansen et al., 2013; Memczak et al., 2013). The abundant expression of *CDR1as/ciRS*-7 relative to other circRNAs and the presence of multiple miR-7 binding sites lead to the proposal that circRNAs could act as miRNA sponges and hundreds of publications describing this mechanism for various circRNAs and miRNAs have been published (reviewed in Panda, 2018; Jarlstad Olesen and S Kristensen, 2021). However, the role of circRNAs acting as miRNA sponges is controversial. Stoichiometry is important as some ceRNAs require thousands of target sites to induce physiological changes to miRNA function (Denzler et al., 2014; Bosson et al., 2014). Examination of circRNA sequences showed that only one circRNA besides *CDR1as/ciRS-7*, contained more miRNA binding sites than expected by chance (Guo et al., 2014). Given a lack of miRNA binding site enrichment in circRNAs, they would need to be expressed at high levels to affect miRNA expression, yet most circRNAs are expressed at low levels (Salzman et al., 2013; Guo et al., 2014). Later work showed that *CDR1as/ciRS-7* knockdown reduced miR-7 levels (Piwecka et al., 2017) consistent with a regulatory network other than miRNA sponging (Kleaveland et al., 2018). Furthermore, correlating expression between circRNAs and miRNAs is not sufficient evidence supporting sponging, and spatial localisation of expression must be considered (Kristensen et al., 2020).

The potential ability of circRNAs to sponge proteins may lead to functional consequences. In some cases, protein sponging by circRNAs assists in regulating gene expression. This was first identified in a negative feedback loop at the *mbl* locus in *Drosophila* (Ashwal-Fluss et al., 2014). A circRNA at the *mbl* locus, *circMBL*, contains binding sites for the muscleblind protein, preventing muscleblind binding to the flanking introns of *circMBL* to promote circularisation thus favouring canonical splicing and the production of more muscleblind protein (Ashwal-Fluss et al., 2014). A similar mechanism was recently described at the paralog *RpL22* and *RpL22-like* loci, also in *Drosophila* (Ng et al., 2024). A circRNA produced from the RpL22

(*circRpL22*) locus controls the RpL22-mediated repression of *RpL22* and *RpL22-like* (Ng et al., 2024). *circRpL22* binding to RpL22 therefore mediates RpL22 repression of these during spermatogenesis (Ng et al., 2024).

circRpL22 is formed from an intronic sequence of its cognate gene, a characteristic that has previously been shown to be important for transcriptional regulation (Zhang et al., 2013; Z. Li et al., 2015). Intron containing circRNAs can be found in the nucleus (**Section 1.3.2**) whereby they localise to promotors and interact with Pol II and the spliceosome component U1 to promote cognate gene transcription and processing (Z. Li et al., 2015). Alternatively, intron-containing circRNAs may form DNA-RNA hybrids, termed R-loops, to influence cognate gene transcription and splicing. A circRNA produced from *ANKRD52* forms an R-loop at its cognate locus promoting transcription through the RNase H1-mediated degradation of the R-loop (Zhang et al., 2013; Li et al., 2021). This mechanism does not appear to be specific to intron-containing circRNAs as an exonic circRNA produced from the *SEPALLATA3* locus in *Arabidopsis*, promotes an exon exclusion in its cognate pre-mRNA through R-loop formation (Conn et al., 2017).

CircRNA protein sponging may also be important for immune responses. Some circRNAs form dsRNA secondary structures, interacting with immune-surveillance dsRNA-binding proteins (Li et al., 2017; C.-X. Liu et al., 2019). The binding of circRNAs to the dsRNA-binding protein kinase R (PKR) results in inhibition (Li et al., 2017; C.-X. Liu et al., 2019). PKR, one of four kinases responding to cellular stress, induces the integrated stress response (ISR) by phosphorylating the alpha subunit of eukaryotic translation initiation factor 2 (eIF2α) (Pakos-Zebrucka et al., 2016). PKR specifically mediates the antiviral aspect of the ISR. During viral infection, circRNA degradation by the endoribonuclease RNase L allows PKR activation (C.-X. Liu et al., 2019; Liu et al., 2022). Thus, regulation of circRNA levels appears essential for innate antiviral responses. Similarly, a circRNA from the *cGAS* locus suppresses the activation of the double-stranded DNA (dsDNA) binding immune sensor cyclic GMP-AMP Synthase (cGAS) (Xia et al., 2018). Additionally, extracellular circRNAs can be taken up by macrophages, potentially acting as danger-associated molecular patterns (Amaya et al., 2024).

While often described as non-coding, at least a subset of circRNAs may produce translatable products (Schneider & Bindereif, 2017). The presence of internal ribosome entry sites (IRES) enables cap-independent translation (Yang & Wang, 2019). IRES can be engineered into circRNA to initiate protein synthesis (Chen & Sarnow, 1995; Wesselhoeft et al., 2018, 2019; R. Chen et al., 2023). Whether circRNAs are endogenously translated is controversial (Dodbele et al., 2021; Hansen, 2021). Some studies report no evidence of endogenous circRNA

translation (Jeck et al., 2013; Guo et al., 2014; Stagsted et al., 2019). However other studies have reported endogenous circRNA expression of m^{6}A-modified circRNAs (Yang et al., 2017) and specific loci such as *ZNF609*, *Mbl*, *Sfl* and *FGFR1* (Legnini et al., 2017; Pamudurti et al., 2017; Weigelt et al., 2020; C.-K. Chen et al., 2021). Notably, the *circZNF609* translation reported by Legnini et al may be an artefact of the trans-spliced linear RNA from the overexpression plasmid used (Ho-Xuan et al., 2020). Recent work has suggested that circRNAs uptaken by macrophages undergo translation to be presented as antigens to stimulate immune responses (Amaya et al., 2024). Cap-independent translation often occurs during periods of cellular stress (Advani & Ivanov, 2019), suggesting a potential mechanism behind the reported widespread translation of circRNA products (van Heesch et al., 2019; C.-K. Chen et al., 2021).

The biological roles of circRNAs offer opportunities for translational applications, with artificial circRNA synthesis crucial for this purpose. Some synthetic circRNAs, such as those produced through permuted group I phage T4 introns (Puttaraju & Been, 1992), trigger innate antiviral immune responses via cytosolic dsRNA sensors PKR and retinoic acid-inducible gene I (RIG-I) (Chen et al., 2017; Zhang et al., 2018; Wesselhoeft et al., 2019). The differentiation between endogenous and exogenous circRNAs remains unclear, but it may involve synthetic circRNA biogenesis artefacts or a lack of m^{6}A modification (Wesselhoeft et al., 2019; Chen et al., 2019; Liu et al., 2022). Endogenous circRNAs appear to suppress certain aspects of the immune response, exemplified through PKR inhibition (C.-X. Liu et al., 2019; Liu et al., 2022). CircRNAs with lower immunogenicity, generated through phage T4 RNA ligase (Liu et al., 2022), have been used to inhibit PKR and applied to diseases models of systemic lupus erythematosus, psoriasis and Alzheimer's disease (C.-X. Liu et al., 2019; Liu et al., 2022; Guo et al., 2024; Feng et al., 2024). Immunogenic circRNAs that activate innate and adaptive immune responses can act as adjuvants for boosting immune response (Chen et al., 2019; Amaya et al., 2023). This, combined with engineering and delivering translatable circRNAs, has been explored for vaccines targeting SARS-CoV-2 and tumours (Meganck et al., 2018; Wesselhoeft et al., 2018, 2019; Qu et al., 2022; Amaya et al., 2023).

1.3.5 Genome-wide detection and quantification of circular RNAs

Detecting circRNAs is challenging because the circular transcript is almost identical to the cognate linear transcript. High throughput computational detection of circRNAs relies on detecting the BSJ. Microarrays containing probes corresponding to BSJ have been developed (S. Li et al., 2019). However, this approach suffers from inflexibility over which BSJs are included in the array. For example, a circRNA array produced by Arraystar contains 13,617

human circRNAs, a fraction of the circRNAs reported in circRNA databases (Vromman et al., 2023).

As an alternative, circRNAs can be computationally detected in short-read RNA-seq data. Given their lack of 3' polyA tails, circRNAs were initially missed in early RNA-seq datasets due to the prevalence of sequencing libraries prepared with oligo(dT) primers, enriching poly(A)+ transcripts. Widespread detection of circRNAs in RNA-seq data occurred after the development of library preparation methods which retain non-polyA transcripts e.g., ribo-depleted RNA-seq (Salzman et al., 2012; Hansen et al., 2013; Jeck et al., 2013; Memczak et al., 2013). Additionally, some studies have sequenced sample fractions left over after poly(A) enrichment has taken place and can be described as poly(A)- selection (Yang et al., 2011; Zhang et al., 2013, 2014). Unlike poly(A)- RNA-seq, ribo-depleted RNA-seq preserves expression data of the cognate linear RNA, which can be useful for comparing relative abundances of circRNAs (Salzman et al., 2013; Ma et al., 2019; Zhang et al., 2020). For optimum circRNA detection, treatment with RNase R to deplete linear RNAs can be performed before library preparation (Jeck et al., 2013; Xiao & Wilusz, 2019). Again however, this has the downside of losing expression data regarding linear RNA transcripts.

CircRNAs are detected by sequencing reads that span the BSJ (**Figure 1.8a**). These reads map divergently to the genomic reference, making them undetectable by standard RNA-seq workflows. Specialised computational tools have been developed to detect BSJ-spanning reads (reviewed in Rebolledo et al., 2023). Many tools begin by aligning reads to a genomic reference and extracting the unmapped reads, which may contain BSJs. For example, CIRI2 (Gao et al., 2018), identifies read segments that do not fully align to the reference genome. These segments are split into seeds, which are then matched to adjacent regions. Maximum likelihood estimation is used to determine if the segments fit the patterns expected from BSJs, canonical splice junctions, or neither. Other methods utilise the aligner STAR (Dobin et al., 2013), to detect chimeric transcripts, some of which may originate from BSJs. Tools like CIRCexplorer2 (X.-O. Zhang et al., 2016) and DCC (Cheng et al., 2016) further filter and annotate these chimeric junctions to identify BSJs. Some tools (such as KNIFE and NCLScan (Szabo et al., 2015; Chuang et al., 2016)) construct pseudo-reference sequences by concatenating sequences to represent BSJ sequences. A pseudo-reference approach has also been utilised by CIRIquant (Zhang et al., 2020), by constructing sequences of BSJs detected by CIRI2 before mapping sequencing reads to this pseudo-reference.

Several tools also include the detection and quantification of the cognate linear RNA. This can be estimated using reads that span canonical splice junctions, sometimes termed forward-

spliced junctions (FJSs) (**Figure 1.8b**). The ratio between circRNA and linear RNA expression is expressed as the circular:linear RNA ratio and has previously been used to identify changes in circRNA expression (Ma et al., 2019; Zhang et al., 2020). Reads that do not span any junction are uninformative for estimating circRNA or linear RNA expression (**Figure 1.8c**)

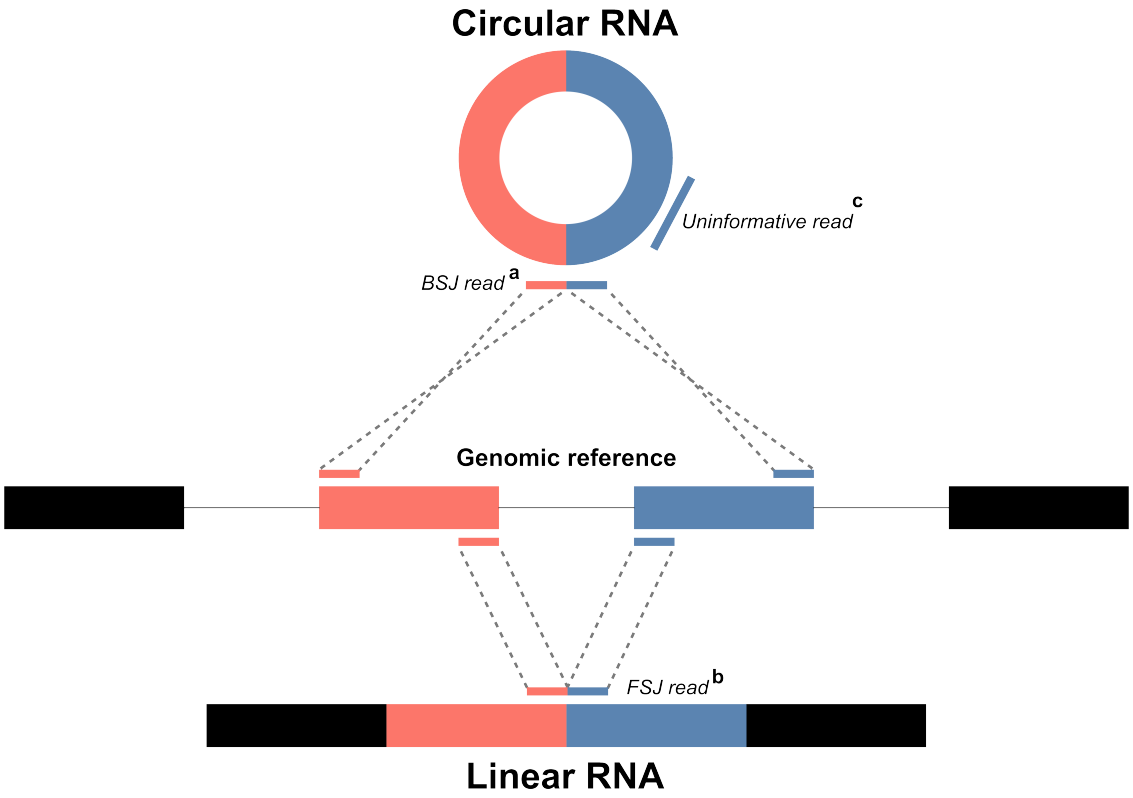


Figure 1.8. Detection of circular and linear RNA using mapping of short-read RNA sequencing reads. (a) A read mapping across the back-spliced junction (BSJ) of a circular RNA. **(b)** A read mapping across a forward-spliced junction (FSJ) of a linear RNA. **(c)** A read mapping within an exon boundary that it uninformative for delineating between circular or linear RNA molecules.

CircRNA reporting in the literature is inconsistent. The development of circRNA databases has led to conflicting nomenclatures, confusing the reporting of identical circRNAs based on the BSJ (Vromman et al., 2021). Most circRNA detection tools report the BSJ location (e.g. *chromosome:BSJ start–BSJ end:strand*). Recently, a consensus nomenclature was proposed (L.-L. Chen et al., 2023), to simplify circRNA structure summarisation. For instance, (*circX*(2, 3)) denotes a circRNA from the second and third exons of gene *X*, distinct from *circX*(3, 4) produced from the third and fourth exons of gene *X*. Given the recency of the new nomenclature, widespread use has not yet been adopted by circRNA detection tools.

1.3.6 Circular RNAs as PD biomarkers

CircRNAs are often discussed as potential biomarkers owing to several characteristics (Verduci et al., 2021). Notably, their covalently closed structure results in increased resistance to exoribonuclease activity (Jeck et al., 2013). Consequently, circRNAs have longer half-lives compared to cognate linear mRNA molecules (Enuka et al., 2016). CircRNAs show tissue- and cell-type specific expression (Salzman et al., 2013; Alhasan et al., 2016; Maass et al., 2017; Xia et al., 2017; Wu et al., 2022). CircRNA expression profiles can differ from cognate linear RNA, and at certain loci, are more abundantly expressed than linear RNA (Salzman et al., 2012; Izuogu et al., 2018; Zhang et al., 2020; Dong et al., 2023). These findings suggest that circRNAs may exhibit unique expression patterns compared to linear RNA which may be exploitable as biomarkers.

CircRNA levels are higher in the brain than in other tissues, and their production is enriched in genes involved in synaptic function (Westholm et al., 2014; Rybak-Wolf et al., 2015; You et al., 2015). As such, circRNAs are of particular interest to researchers studying neurodegenerative disorders.

For use as biomarkers, circRNAs should be detectable in accessible and non-invasive sources. Fortunately, circRNAs are detectable in blood, saliva and urine (Bahn et al., 2015; Memczak et al., 2015; Vo et al., 2019; Hutchins, Reiman, et al., 2021). Furthermore, circRNAs have been detected in extracellular vesicles (Y. Li et al., 2019). As extracellular vesicles can cross the blood-brain barrier (Ramos-Zaldívar et al., 2022), the isolation and detection of circRNAs within neuronal-derived vesicles may offer a source of biomarkers relevant to neurodegenerative disorders.

Initial studies on circRNA expression in PD used cellular and animal models to identify differential expression from loci such as *zip-2*, *SNCA* and *DLGAP4* (Kumar et al., 2018; Sang et al., 2018; Feng et al., 2020; Jia et al., 2020). Later work identified circRNAs differentially expressed in human substantia nigra and blood (Hanan et al., 2020; Ravanidis et al., 2021; Kong et al., 2021; Zhong et al., 2021; Xiao et al., 2022), reinforcing circRNA transcription differences in PD. The identification of differentially expressed circRNAs in blood suggested they could serve as accessible PD biomarkers. In these studies, specific circRNAs could distinguish PD patients from controls with AUCs ranging from 0.84 to 0.98, indicating potential clinical utility in PD diagnosis (Ravanidis et al., 2021; Kong et al., 2021; Zhong et al., 2021; Xiao et al., 2022). However, concerns about sample sizes, analysis methods, and replicability (**Section 4.1**) necessitate further investigation with larger cohorts and rigorous methodology.

1.4 Research hypotheses and aims

1.4.1 Research hypotheses

I hypothesise that patients in the early stages of idiopathic PD will exhibit transcriptional differences when compared to controls. Specifically, there will be specific genes and circRNAs that are differentially expressed in PD patients compared to controls. I also hypothesise that the expression of individual genes or circRNAs will offer the ability to distinguish between idiopathic PD patients and controls. Given prominent differences detectable in the blood of PD patients, the identification of dysregulated genes and pathways would provide disease-related changes in the blood, providing novel insights into PD development or progression.

1.4.2 Research aims

To address my hypotheses, I use RNA-seq data from several large, well-phenotyped cohorts. The first cohort involves participants enrolled in the Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-PD (ICICLE-PD). I supplement these RNA-seq data with publicly available and protected-access RNA-seq data from the Parkinson's Progression Marker Initiative (PPMI) and BRAIN Cell encyclOpeDia of transcribed Elements Consortium (BRAINcode).

In **Chapter 3**, I compare blood gene expression between patients recently diagnosed with idiopathic PD and controls from the PPMI and ICICLE-PD studies. By focusing specifically on individuals recently diagnosed with idiopathic PD, I aim to discover genes differentially expressed in the early stages of PD. As such, I then aim to assess the ability of gene expression to distinguish recently diagnosed idiopathic PD patients from controls, in the hope gene expression could be used as a diagnostic biomarker of PD. Furthermore, I aim to identify dysregulated genes that relate to our current understanding of the genetic architecture of PD and may be used as a proxy for current clinical measures of disease progression. Finally, I aim to uncover dysregulated biological processes that could inform our current understanding regarding peripheral systematic changes that are occurring in the early stages of idiopathic PD.

In **Chapter 4**, I compare blood circRNA expression between patients recently diagnosed with idiopathic PD and controls from the PPMI and ICICLE-PD studies. Again, I aim to discover circRNAs differentially expressed specifically in the early stages of PD. I then aim to evaluate the ability of circRNAs to distinguish idiopathic PD patients from controls. I am particularly interested in whether circRNA expression offers an improved ability to identify PD patients over gene expression (**Chapter 3**).

Following the identification of a global reduction in circRNA expression in recently diagnosed idiopathic PD patients in PPMI and ICICLE-PD studies (**Chapter 4**), in **Chapter 5**, I aim to explore the reasons for this reduction. I aim to assess a variety of possible technical and biological explanations.

In **Chapter 6**, I compare blood circRNA expression between controls and patients with genetic forms of PD enrolled in the PPMI. I also aim to compare blood circRNA expression between controls and individuals at increased risk of developing PD enrolled in the PPMI. Finally, I aim to assess circRNA expression changes in dopaminergic neurons isolated post-mortem from the substantia nigra of individuals with varying levels of Lewy body pathology ranging from controls (no pathology), incidental Lewy body cases (Lewy body pathology but no clinical neurodegenerative diagnosis) and PD patients from the BRAINcode study.

Chapter 2. General methods

This section describes data relating to the PPMI and ICICLE-PD study cohorts. All participant enrolment and data collection were performed externally from my PhD by the respective study teams. This thesis concerns the analysis of sequencing data generated using samples collected as part of the ICICLE-PD study and protected-access data provided by the PPMI. Other contributions (e.g. library preparation and sequencing) are clarified in the respective sections.

2.1 Cohort descriptions

2.1.1 The Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-PD (ICICLE-PD)

The Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-PD (ICICLE-PD) is an observational and longitudinal study jointly run by groups from Newcastle University and the University of Cambridge. Its primary aim goal is to characterise the development of cognitive impairment in PD. Data generated in this thesis is based on participants recruited to the Newcastle arm of the study.

Details of the study, including inclusion and exclusion criteria, have been previously reported (Yarnall et al., 2014). Briefly, patients recently diagnosed with PD (mean disease duration [SD] = 5.5 [5.0] months) were recruited between 2009-2011 (Yarnall et al., 2014). PD diagnosis was based on the UK Brain Bank criteria (Hughes et al., 1992). Patients were excluded if a diagnosis of alternative parkinsonism was made (Dementia with Lewy bodies, Drug-induced parkinsonism, Vascular parkinsonism, Progressive supranuclear palsy, Multiple system atrophy, Corticobasal degeneration). Control participants of similar sex and age profiles were also recruited. Participants with no known neurological disorders (movement or mood disorders) were enrolled as controls. Given ICICLE-PD's overarching focus on cognitive impairment, participants were excluded if cognitive impairment was recorded at the first visit. All data from ICICLE-PD participants analysed in this thesis was based on samples and data collected at the initial (baseline) visit. Ultimately, 381 participants were enrolled in ICICLE-PD (219 PD and 99 controls). Analysis in this thesis includes 96 participants (48 PD and 48 controls) as described in **Section 2.2**.

2.1.2 Parkinson's Progression Markers Initiative (PPMI)

The Parkinson's Progression Markers Initiative (PPMI) is a global observational and longitudinal study. Detailed information regarding the study has been previously published

(Marek et al., 2011, 2018). Briefly, patients recently diagnosed with PD (<2 years) were recruited from 24 study sites globally. Inclusion criteria included the presence of established clinical features and imaging-validated dopaminergic deficits (Marek et al., 2018). Notably, all PD patients were not receiving dopaminergic treatment at enrolment. Control participants of similar age and sex profiles were also recruited. Control participants were free of neurological disorders and without a first-degree relative diagnosed with PD. In this initial phase, 423 PD patients and 196 controls were recruited. An additional 64 PD patients were enrolled but were found to have no evidence of dopaminergic deficit when imaged. In the initial phase, several PD patients were found to harbour pathogenic variants in *LRRK2* and *GBA* (Nalls et al., 2016). Expansion of the PPMI study in 2013-2019 resulted in the recruitment of individuals harbouring pathogenic variants in *GBA*, *LRRK2* and *SNCA*. Further expansion included the recruitment of individuals with established risk factors such as hyposmia and REM sleep behaviour disorder (RBD). Genetic status was based on results from clinical genetic screening or whole genome sequencing as previously described (Hutchins, Craig, et al., 2021; Craig et al., 2021).

Analysis in **Chapters 3, 4, and 5** concerns PPMI participants categorised as idiopathic PD and controls in metadata provided by the PPMI. Further filtering criteria are outlined in **Section 2.2**.

Analysis concerning a wider set of PPMI participants is described in **Chapter 6**. For these participants, selection criteria are outlined in **Section 0**.

2.1.3 Clinical measures

PD-related clinical measures were collected by trained examiners during the participation of PPMI and ICICLE-PD studies. Clinical assessments were used to measure the severity of symptoms associated with PD. In this thesis, I primarily utilise two clinical scoring systems: The Movement Disorders Society Unified Parkinson Disease Rating Scale Part III (MDS-UPDRS III) and the Montreal Cognitive Assessment (MoCA). The MDS-UPDRS III assesses motor ability (Goetz et al., 2008). The MDS-UPDRS III is scored from 0-132 with proposed cutoff points including 32/33 for mild/moderate and 58/59 for moderate/severe motor dysfunction (Martínez-Martín et al., 2015). MoCA scores assess cognitive impairment (Nasreddine et al., 2005). MoCA is scored from 0-30 with scores >=26 typically used to reflect normal cognition (Nasreddine et al., 2005). To enable comparisons of different PD associated medications, the dosage is typically compared to the gold standard levodopa dosage, providing a levodopa equivalent daily dose (LEDD). The PPMI and ICICLE-PD studies reported LEDD using established methodology (Tomlinson et al., 2010).

2.2 Defining discovery (PPMI) and replication (ICICLE-PD) cohorts of individuals with early-stage idiopathic PD and matched controls

Analyses presented in **Chapters 3, 4, and 5** use the same samples from the PPMI and ICICLE-PD studies. In these analyses, the PPMI cohort is used as a discovery cohort due to its larger size, and the ICICLE-PD cohort is used for replication. To increase comparability between cohorts, additional inclusion criteria were enacted. Broadly, the analyses concern individuals with early-stage idiopathic PD and matched controls. Here, idiopathic was defined as a diagnosis of PD in the absence of a known genetic or environmental cause. Early-stage was defined based on the time following diagnosis; both PPMI and ICICLE-PD patients were included if the time from sample collection was within 13 months of a clinical PD diagnosis. 13 months was selected as a cutoff based on the maximum disease duration in the analysed ICICLE-PD samples. Both PPMI and ICICLE-PD are longitudinal studies with follow up visits, yet the data in this thesis includes only data from the baseline visit.

An outline of the steps used to create the PPMI analysis cohort is shown in **Figure 2.1**. Participants taking medications i.e., neuroleptics/anti-psychotics, and anticoagulants, as detailed in the PPMI protocol were excluded. Study groups (idiopathic PD or control) were as recorded in metadata provided with RNA sequencing data (Hutchins, Craig, et al., 2021; Craig et al., 2021). Updates to participant subgroup status (e.g., a change in diagnosis) are reported by the PPMI periodically. Reported study groups were then matched against the study groups reported as of 28th October 2021. RNA sequencing quality control (QC) performed by PPMI has been previously described (Hutchins, Craig, et al., 2021; Craig et al., 2021). The number of samples processed in **Figure 2.1** refers to samples which were successful in QC steps and downstream analysis workflows (**Sections 2.4, 2.4.3, 2.6**). In total, 420 PPMI samples (259 iPD, 161 controls) were used in the analyses described in **Chapters 3, 4 and 5**.

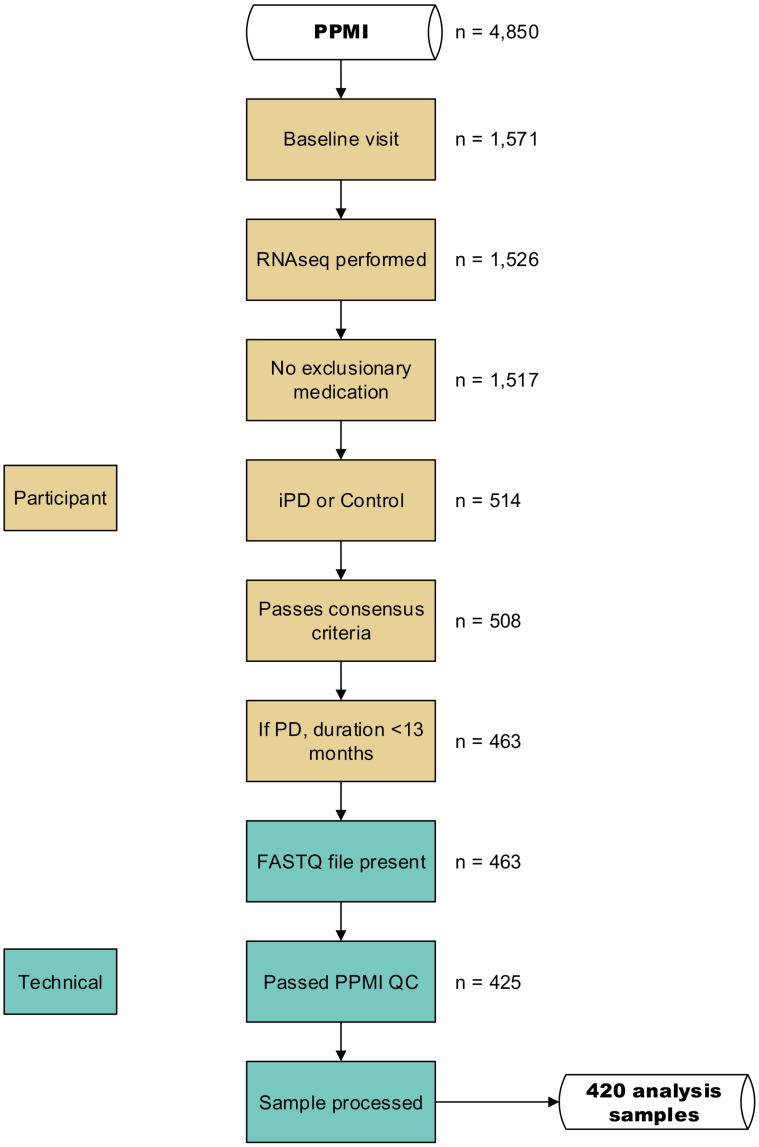


Figure 2.1. PPMI cohort filtering. Schematic outlining the workflow for selecting PPMI samples for analysis. At each step, the number of samples remaining is shown on the right-hand side.

From the ICICLE-PD cohort, 96 samples were selected for RNA sequencing. These comprised 48 PD patients and 48 controls matched for age and sex. Detailed cohort breakdowns are given in **Table 2.1**.

	PPMI				***ICICLE-PD***		
Variable	*PD (n = 259)*	*Controls (n = 161)*	*P*-value		*PD (n = 48)*	*Controls (n = 48)*	*P*-value
Sex (M/F)	172/87	105/56	0.83		29/19	23/25	0.31
Age at sample collection (Years)	62.9 ± 9.44	60.8 ± 11.7	0.058		64.8 ± 10.5	67.9 ± 7.36	0.094
Age at diagnosis (Years)	62.1 ± 9.43	N/A	N/A		64.3 ± 10.5	N/A	N/A
Disease duration (Months)	4.54 ± 2.90	N/A	N/A		5.04 ± 3.41	N/A	N/A
UPDRS-III	20.2 ± 8.26	1.26 ± 2.25	**$< 2.2 \times 10^{-16}$**		23.5 ± 11.8	N/A	N/A
MoCA	27.2 ± 2.39	28.3 ± 1.11	**5.1×10^{-9}**		26.0 ± 3.24	27.1 ± 2.57	0.096
Dopaminergic treatment (Y/N)	0/259	N/A	N/A		44/4	N/A	N/A

Table 2.1. PPMI and ICICLE-PD analysis participant characteristics. Table showing relevant demographic and clinical characteristics of individuals involved in analysis following sample QC. Where possible, the mean value for each group is given along with the standard deviation (SD). Differences in the proportion of Males/Females between PD and Controls were assessed using a two-sided Fisher's exact test. All other group differences were assessed using a two-sided two-sample T-test. MoCA = Montreal Cognitive Assessment. UPDRS-III = 3rd part of the Unified Parkinson's Disease Rating Scale. NA indicates not available.

2.3 RNA sequencing

2.3.1 ICICLE-PD

2.3.1.1 Sample collection

Blood samples were collected and RNA was extracted as previously described (Martin-Ruiz et al., 2020). Briefly, total RNA was extracted from blood as per the manufacturer's instructions (PAXgene Blood RNA Kit, Qiagen).

2.3.1.2 Library preparation

Library preparation of ICICLE-PD samples was performed by the Genomics Core Facility (Newcastle University). RNA quality was assessed using an Agilent 2100 Bioanalyzer Instrument. All samples had an RNA integrity number (RIN) > 8. All samples underwent ribosomal and haemoglobin RNA reduction (Globin-Zero Gold rRNA Removal Kit, Illumina) followed by total RNA-seq library preparation (Illumina TruSeq Stranded Total RNA, Illumina).

2.3.1.3 Sequencing

Sequencing of ICICLE-PD samples was performed by the Genomics Core Facility (Newcastle University) on an Illumina NovaSeq 6000 generating 150bp paired-end reads. Sequencing was performed over three sequencing runs using a block design allowing potential batch effects to be accounted for in downstream analyses (Leek et al., 2010). Samples were sequenced to a median depth of 89.2 (IQR=15.7) million paired-end reads.

2.3.2 PPMI

All PPMI sample collection and sequencing were performed externally to this project. Raw FASTQ files were downloaded from the PPMI's online repository (https://www.ppmi-info.org/).

2.3.2.1 Sample collection

Due to the involvement of various contributing sites to the PPMI cohort, sample collection was subject to publicly available standard operating protocols (https://www.ppmi-info.org/study-design/research-documents-and-sops). The detailed protocol for whole blood collection is given in the PPMI Biologics Manual and outlined in the PPMI's RNAseq publication (Craig et al., 2021). Briefly, 2.5mL of whole blood was collected into PAXgene tubes (Qiagen), incubated at room temperature for 24 hours and then frozen at -80°C. Samples were then

shipped to the Biorepository Core at Indiana University (US). Total RNA was extracted from blood as per the manufacturer's instructions (PAXgene Blood RNA Kit, Qiagen).

2.3.2.2 Library preparation

Library preparation was performed at the HudsonAlpha Institute for Biotechnology, with detailed information given in the PPMI's RNAseq manuscript (Craig et al., 2021). Briefly, ribosomal and globin RNA were depleted using the Globin-Zero Gold Kit (Illumina). A stranded total RNA library was then prepared using a NEB/KAPA based kit.

2.3.2.3 Sequencing

Sequencing was performed at HudsonAlpha Institute for Biotechnology on an Illumina NovaSeq 6000. This generated paired-end reads with lengths 125-150bp. The PPMI report the median sequencing depth for the entire project as approximately 100 million paired-end reads (Craig et al., 2021).

2.4 RNA sequencing quality control

Outputs from the QC steps below were parsed and collated using MultiQC v1.18 (Ewels et al., 2016).

2.4.1 Sequencing quality control

Sample contamination and species of origin was assessed using FastQ screen v0.14.1 (Wingett & Andrews, 2018) with default parameters. Reads were aligned with bowtie2 v2.3.4.2 (Langmead & Salzberg, 2012) against a set of commonly used reference genomes (*Homo sapien*, *Mus musculus*, *Rattus norvegicus*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana*, *Escherichia coli*) and contaminants (ribosomal RNA, PhiX, Lambda, vectors and sequencing adapters). General sequencing metrics were obtained using FastQC v0.11.7 (www.bioinformatics.babraham.ac.uk/projects/fastqc/).

2.4.2 Alignment quality control

Reads were aligned to the Ensembl GRCh38 reference with HISAT2 v2.1.0 (D. Kim et al., 2019). A pre-generated HISAT2 index based on Ensembl GRCh38 release 84 was downloaded from the AWS Public Dataset Program (https://registry.opendata.aws/jhu-indexes/). Both ICICLE-PD and PPMI underwent stranded and paired-end sequencing which was specified during alignment (*--rna-strandness RF*).

Alignment metrics were collected from HISAT2 log files, *stats* from SAMtools v1.17 (Danecek et al., 2021) and *CollectRnaSeqMetrics* from Picard v2.27.5. As *CollectRnaSeqMetrics* requires a genomic reference in the refFlat format, the Ensembl GRCh38 (release 101) annotation GTF was converted to the genePred format using the UCSC *gtfToGenePred* tool before being reformatted into the refFlat format. Ribosomal RNA (rRNA) loci were acquired from the GTF annotation based on the *rrna* tag and then converted to BED format using gff2bed. A rRNA loci interval list was then generated using Picard *BedToIntervalList*, utilising a Ensembl GRCh38 FASTA sequence dictionary created using Picard *CreateSequenceDictionary*. Again, strand specificity was explicitly defined when running Picard on PPMI and ICICLE-PD samples (*--STRAND_SPECIFICITY SECOND_READ_TRANSCRIPTION_STRAND)*.

2.4.3 Validating clinically recorded sex

Clinically recorded sex is typically available as metadata with sequencing data, yet errors may occur during data input. Sex can be inferred using the expression of genes located on the sex chromosomes. For example, comparing the expression of *XIST* against the expression of Y chromosomal genes (’t Hoen et al., 2013). The PPMI analysis performed dimensionality reduction on the expression of six genes (*XIST*, *RPS4Y1*, *RPS4Y2*, *KDM5D*, *DDX3Y* and *USP9Y*). XIST, a long non-coding RNA (lncRNA), was not present in the transcriptome reference used for transcript quantification (**Section 2.5.1**). In addition, *RPS4Y2* was not present in either the PPMI or ICICLE-PD expression data after count filtering (raw count > 10 counts in the smallest sample group). Principal Component analysis (PCA) was performed on the counts of Y chromosomal genes (*RPS4Y1*, *KDM5D*, *DDX3Y* and *USP9Y*) as described in **Section 2.10.1**. Scatter plots of the first and second principal components were then visually inspected to identify mismatches between the plotted sample position and reported sex.

2.5 Gene quantification

2.5.1 Transcript quantification

Transcriptome indexing and quantification were performed using Salmon v1.3.0 (Patro et al., 2017) using scripts written by Dr Dasha Deen (Newcastle University). Indexing was performed on the Ensembl GRCh38 transcriptome reference FASTA (release 101). As ICICLE-PD and PPMI reads are longer than 75bp, the default k-mer length (31) was used to generate the index as recommended (https://salmon.readthedocs.io/en/stable/salmon.html, accessed 03/06/2024). Transcripts were quantified using the selective mapping strategy (*--validateMappings*), which

has shown improved performance compared to standard quasi-mapping methods (Srivastava et al., 2020). Sequencing biases (*--seqBias*), potentially caused by random hexamer priming, and non-uniform coverage biases (*--posBias*) were corrected (Roberts et al., 2011). GC content biases were also corrected for (*--gcBias*) which may otherwise affect transcript expression estimates ('t Hoen et al., 2013; Love et al., 2016).

2.5.2 Gene expression summarisation

Transcript annotations were extracted from the Ensembl GRCh38 release 101 annotation using the *makeTxDbFromGFF* function from GenomicFeatures v.52.2 (Lawrence et al., 2013). Transcript IDs were matched to their host gene IDs and gene-level counts were calculated using tximport v1.28.0 (Soneson et al., 2015).

2.5.3 Count normalisation

Further information regarding the normalisation of gene expression is provided in the respective chapters. Transcript per million (TPM) counts were collected from the transcript estimates produced by Salmon (**Section 2.5.1**) and collapsed to the gene-level using tximport (**Section 2.5.2**).

Information regarding the normalisation methods implemented in the differential expression tools, accounting for both differences in sequencing depth and RNA composition (Evans et al., 2017), is given in **Section 2.8**.

2.6 Circular RNA

2.6.1 Detection and quantification

Circular RNAs (circRNAs) were detected based on reads containing a back-spliced junction (BSJ). A circRNA processing workflow was built to ensure consistent circRNA detections across cohorts (**Figure 2.2**, https://github.com/bj-w/PD-lin-circ-RNA-paper/tree/main/circRNA_detection). This workflow was run separately for PPMI and ICICLE-PD samples in **Chapters 4 and 5**. In **Chapter 6**, an alternative workflow was developed which is described in **Section 6.2.6**.

As some circRNA detection methods work on unmapped reads, low quality (MAPQ <15) and sequencing adapters were removed from raw sequencing data using Trim Galore v0.67 running cutadapt v4.2 (Martin, 2011). CircRNAs were detected using multiple tools as recommended

(Hansen et al., 2016; Hansen, 2018; Gaffo et al., 2021; Vromman et al., 2023). Tools with differing methodologies (i.e., read alignment and BSJ identification) were selected to increase the diversity of the detected circRNAs.

CircRNAs were detected using CIRI v2.0.6, which uses a maximum likelihood estimation to identify whether unmapped segments of reads are consistent with a BSJ (Gao et al., 2018). Reads were aligned with BWA v0.7.17 (Li & Durbin, 2009), restricting alignments to reads with scores above 19 (*-T 19*) as recommended by the authors of CIRI2 (https://ciri-cookbook.readthedocs.io/en/latest/CIRI2.html, accessed 03/06/2024). When running CIRI2, non-default parameters included restricting the maximum BSJ span (*--max_span 1000000*) and BSJ expression relative to the corresponding FSJ expression (*--rel_exp 0.05*).

CircRNAs were also detected with PTESfinder v2.0.0 (PFv2, https://github.com/osagiei/pfv2), an updated version of PTESfinder (Izuogu et al., 2016). Paired end FASTQ files were interleaved using BBtools v38.92 (https://sourceforge.net/projects/bbmap/). Bowtie v2.3.4 (Langmead & Salzberg, 2012) and STAR v2.7.10a (Dobin et al., 2013) were used as aligners in this step using parameters outlined in the PFv2 script (https://github.com/osagiei/pfv2). Briefly, PFv2 identifies putative BSJs based on chimeric alignments detected by STAR. Using Bowtie 2, sequencing reads are then mapped to the generated sequence constructs that are consistent with BSJs.

Finally, circRNAs were detected using CIRCexplorer v2.3.8 (X.-O. Zhang et al., 2016), which can parse chimeric alignments produced by a range of aligners. For efficiency, chimeric junctions output by STAR in the PFv2 step were used for this purpose.

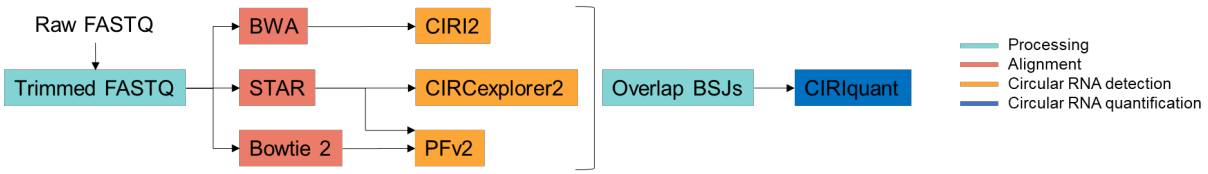


Figure 2.2. Circular RNA detection and quantification pipeline. Schematic outlining the individual steps from raw sequencing files to circRNA quantification.

Further information on identified candidate circRNAs in PPMI and ICICLE-PD samples (**Chapters 4 and 5**) is outlined in **Section 4.3.1**. BSJ loci detected by each tool were then overlapped. Similar to others (Memczak et al., 2015), false positives were limited by retaining BSJs with a read count >1 in at least two individuals. Then, BSJs detected by at least two tools were kept.

Higher confidence BSJs were used as input for quantification with CIRIquant v1.1.2 (Zhang et al., 2020). CIRIquant quantifies circRNA expression by constructing a pseudo-reference sequence including sequences that would be present due to the input BSJs, and aligning raw reads to this sequence using HISAT v2.2.0 (D. Kim et al., 2019). Paired-end reads that align concordantly across 10bps of the predicted BSJ site are classed as circRNA reads. Reads aligning around the BSJ site but do not pass the criteria to be deemed as circRNA reads are recorded as canonical splice junction reads (forward-spliced junctions, FSJs), to give an estimation of the cognate linear RNA expression. BSJ and FSJ counts were used to calculate the circular to linear ratio as implemented in CIRIquant (**Equation 2.1**). BSJs counts were multiplied by two to account for the fact that FSJ reads are detected at two loci while BSJ reads are detected at one locus.

$$Circular\ to\ linear\ ratio = \frac{2 \times BSJ}{2 \times BSJ + FSJ}$$

Equation 2.1. Calculation of the circular to linear ratio at back-spliced junction loci.

2.6.2 Count normalisation

There is no current consensus on the optimum circRNA expression normalisation denominator (L. Chen et al., 2021; Rebolledo et al., 2023; Ma et al., 2023). As the number of circRNAs detected in a sample can vary it is an inappropriate normalisation denominator. Based on the work carried out in **Section 4.3.2**, I normalised junction counts (BSJ and FSJ) to the number of transcriptome-mapped reads in each sample as reported by Salmon (**Section 2.5.1**).

2.7 Identifying sources of extraneous variation in the expression of RNAs

Biological and technical factors are known to impact the quantification of gene expression (’t Hoen et al., 2013; Su et al., 2014). Like previous large-scale transcriptomic studies, I quantified sources of expression variation at both the sample and gene level (Craig et al., 2021; M. Wang et al., 2021; Lopes et al., 2022).

Firstly, I identified technical sources that influenced sample-level expression variation. Technical metrics were collected from alignment quality control (**Section 2.4.2**) and comprised of RNA-seq sequencing metrics as reported by Picard *CollectRnaSeqMetrics* v2.27.5 and the

average insert size as reported by SAMtools *stats* v1.17. I then constructed univariate linear regression models to explore the relationship between each technical metric and the first 10 principal components (**Section 2.10.1**) of RNA expression (i.e., gene expression and circRNA expression). Metrics with R^2 >0.5 were highlighted unless otherwise stated. Pairwise correlations were assessed and visualised as described in **Section 2.10.2**. To reduce redundancy, highly correlated (absolute Spearman's rho >0.9) technical metrics were excluded by keeping the metric with the highest correlation to the PC.

Individual RNA-specific sources of expression variation were identified using variancePartition v1.30.2 (Hoffman & Schadt, 2016). Using this method, linear mixed models were created for each individual RNA feature (gene/circRNA) to calculate the proportion of variation explained by a range of potential covariates. At this stage, biological sources of variation (age, sex, study group) were also included. Covariates that generally explained more variance than the study group (based on the median proportion of variance explained) were included in the analyses. Specific covariates identified for each cohort and RNA type are described in the respective chapters.

2.8 Differential expression of RNAs

Differential expression testing was performed using both DESeq2 v1.40.2 (Love et al., 2014) (**Chapters 3, 4 and 5**) and limma v3.56.2 (Ritchie et al., 2015) (**Chapters 4 and 5**). Specific uses of each tool are described in the respective chapters. In **Chapter 6**, limma was used for differential expression due to its increased computational speed compared to DESeq2.

Within DESeq2, normalised expression was calculated using the default Median of Ratios method (Anders & Huber, 2010) using the function counts(dds, normalized = TRUE). Differential expression was performed using the *DESeq()* function with default settings.

Before differential expression using limma, normalised expression was calculated using the default Trimmed Mean of M-values (TMM) method (Robinson & Oshlack, 2010) using the cpm() function implemented in edgeR v3.42.4 (Robinson et al., 2010). TMM-normalised counts underwent voom transformation (*voom()*) to account for heteroskedasticity (Law et al., 2014). Linear models were fitted for each gene (*lmFit()*) before robust empirical Bayes moderation (*eBayes(robust = TRUE)*) was used to compute moderated t-statistics (Phipson et al., 2016).

P-values reported by each tool were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). Similar to previous work performed on PPMI RNA-seq data (Craig et al., 2021), genome-wide significantly differentially expressed genes and junctions were defined as those that had an FDR <0.05 and a $\log_2$ fold change ($\log_2$FC) <-0.1/ >0.1.

Notably, when testing for differential expression of junctions (both BSJ and FSJs), sample size factors (DESeq2) and normalisation factors (edgeR) were calculated from gene expression data (**Section 2.5.3**). This method avoids the issue of variable circRNA abundances (**Section 2.6.2**); by normalising against the number of transcriptome mapped reads.

2.9 PD classification

Various models to classify the PD status of samples are reported in this thesis. The classification models all follow to same general framework (outlined below) yet differ in the predictors used. Details of the predictors used in the respective models are outlined in each Chapter.

Regularised logistic regression classification models described in **Chapters 3 and 4** were created using glmnet v4.1-8 (Friedman et al., 2010). Models were trained and parameters optimised using PPMI samples. Model training and optimising parameters on the same data through k-fold cross-validation can lead to biased estimates of performance (Varma & Simon, 2006; Vabalas et al., 2019; Lewis et al., 2023). Furthermore, performing a single splitting of data into training and test splits has been shown to provide unstable estimates of performance, particularly with smaller sample sizes associated with biological data (An et al., 2021). Nested cross-validation has been shown to address these problems (Varma & Simon, 2006; Vabalas et al., 2019; Lewis et al., 2023).

Nested cross-validation was carried out using the *nestcv.glmnet* function implemented in nestedcv v0.7.3 (Lewis et al., 2023) (**Figure 2.3**). Briefly, PPMI samples were divided into 10 cross-validation folds, forming an outer loop. Within the training data of the outer loop, an inner loop was formed comprised of 10 cross-validation folds. Optimum model parameters were identified using training data of the inner loop based on predictions made in the test data of the inner loop. For this, the alpha parameter was assessed using a grid search with the values 0-1 in increments of 0.1. Predictions from each outer loop test set were combined to produce an unbiased estimate of model performance. Lastly, cross-validation was performed on the entire PPMI dataset to identify final model parameters, with the final model fit using all PPMI

samples. The final model was then used to predict PD status in ICICLE-PD samples to assess the generalisability of the model to an external data source.

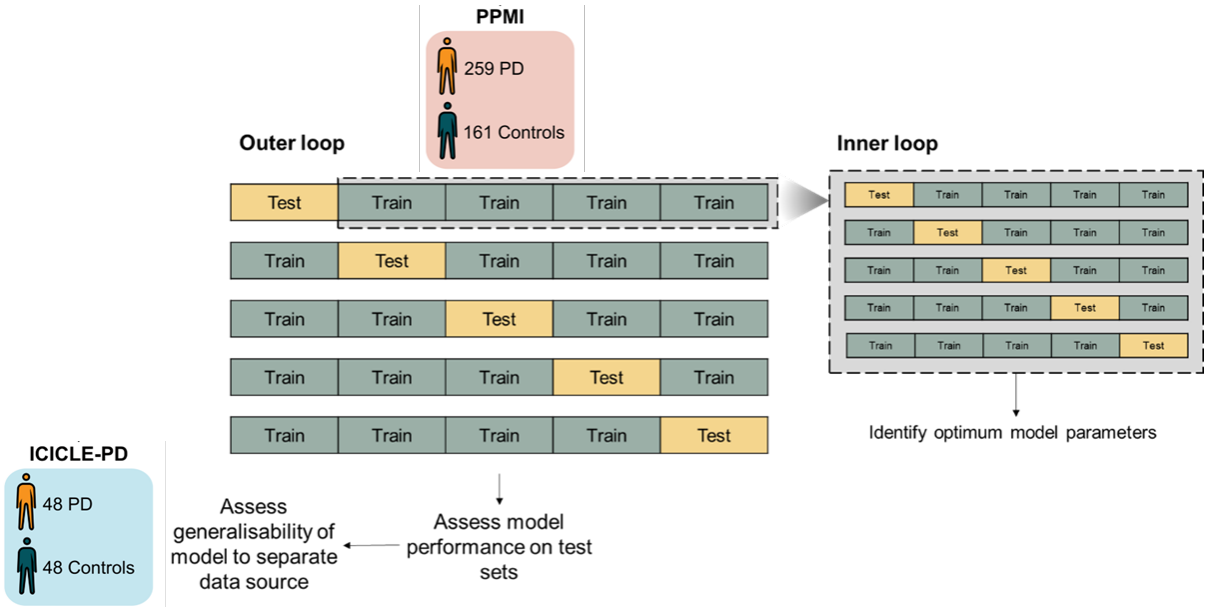


Figure 2.3. Nested cross-validation. Schematic outlining how nested cross-validation was performed to create classification models trained using PPMI samples.

Receiver-operator characteristic (ROC) curves were created using the *roc* function from pROC v1.18.4 (Robin et al., 2011). This returns the sensitivity and specificity at various thresholds i.e., each prediction value returned by a classification model. To summarise classification performance, the area under the ROC curve (AUC) was calculated using the *auc* function as implemented in pROC. 95% confidence intervals (95% CI) were calculated using Delong's method (DeLong et al., 1988) as implemented in the *auc* function.

2.10 Statistical analyses

All statistical analyses were performed in R v4.3.1. Statistical tests are described in the text where necessary and performed using the implementations in the base R *stats* package unless otherwise stated. Detailed information regarding specific statistical analyses is outlined in the respective Chapters.

2.10.1 Principal component analysis (PCA)

RNA expression counts were transformed to account for heteroskedasticity using the variance-stabilising transformation (Anders & Huber, 2010), as implemented in the *vst()* function from

DESeq2 v1.40.2 (Love et al., 2014). Principal components were then calculated using the *prcomp()* function.

2.10.2 Visualisation of pairwise correlations

Pairwise correlation matrices were visualised using ComplexHeatmap v2.16.0 (Gu et al., 2016). Hierarchical clustering of heatmap rows and columns was performed using the complete linkage method of clustering based on Euclidean distances as is default in ComplexHeatmap.

2.10.3 Identifying global changes in RNA expression

Global changes in RNA expression were identified based on systematic biases in the direction of the effect of individual features (gene or junction expression depending on the context). The direction of effect of each feature was determined by the fold change returned from comparing the expression of each feature in PD patients compared to controls. For example, if a $\log_2$ fold change threshold was set at 1.5, a $\log_2$ fold change > 1.5 would indicate increased expression in PD while a $\log_2$ fold change < -1.5 would indicate decreased expression in PD. The term 'imbalance' is used to describe an excess of increased or decreased features. For simplicity, I report the imbalance estimate based on **Equation 2.2**. For example, an imbalance of 0.5 would represent an equal number of increased and decreased features, an imbalance of 0 indicates a decrease for all features and an imbalance of 1 an increase for all features. Significance was assessed using a two-sided exact binomial test (*binom.test*) with the probability for the null hypothesis set at 0.5.

$$\frac{Number\ of\ loci\ increased\ in\ PD}{Number\ of\ loci\ increased\ in\ PD + Number\ of\ loci\ decreased\ in\ PD}$$

Equation 2.2. Calculation of RNA expression imbalances.

Chapter 3. Gene expression in early-stage idiopathic Parkinson's disease

3.1 Background

Dysregulated biological processes during neurodegeneration leads to measurable gene expression changes (Noori et al., 2021). Most PD transcriptomic studies have been performed on postmortem substantia nigra (reviewed in Lewis and Cookson, 2012; Borrageiro et al., 2018). Dysregulated processes highlighted by these studies converge on processes extensively linked to PD, such as mitochondrial and lysosomal function, proteostasis and inflammation (Borrageiro et al., 2018; Morris et al., 2024).

Profiling gene expression in postmortem tissues has several inherent issues. RNA quality is an important consideration for robust gene quantification (Gallego Romero et al., 2014). The amount of time following death (postmortem intervals) may influence expression changes (Zhu et al., 2017), yet the effect of postmortem interval on RNA quality is not clear (White et al., 2018). Death can induce transcriptional changes (Ferreira et al., 2018), further compounded by the manner of death (Harrison et al., 1991; Preece & Cairns, 2003). Furthermore, postmortem PD patient tissue is typically biased towards advanced disease stages (Halliday & McCann, 2010) and age-associated transcriptomic changes potentially limit the applicability of findings to other age demographics (Ham & Lee, 2020).

The acquisition of brain tissue requires invasive interventions, curtailing its ability to provide easily accessible and actionable biomarkers. Although profiling transcriptomes in the living brain is possible (Benoit et al., 2020), it is not routinely performed, limiting sample sizes. Blood represents an ideal biomarker source: it is non-invasive and easily accessible. Samples can be obtained imminently following diagnosis, allowing insights into early changes in disease onset and progression. There is considerable peripheral involvement in PD, including measurable changes in blood (**Section 1.1.5**). Therefore, changes in blood gene expression in PD patients may be exploitable as diagnostic PD biomarkers.

Early work investigating gene expression changes in the blood of PD patients evaluated the expression of select genes using microarray (Scherzer et al., 2007) or qPCR (quantitative polymerase chain reaction) measurements (Grünblatt et al., 2010; Molochnikov et al., 2012; Santiago & Potashkin, 2013). More recent work used RNA-seq to measure blood gene expression changes in PD (Infante et al., 2016; Garofalo et al., 2020; Henderson et al., 2021; Kurvits et al., 2021) (**Table 3.1**).

Publication	Sample size		
	PD	Control	Number of differentially expressed genes
Infante et al., 2016	20	20	297
Garofalo et al., 2020	6	14	5
Henderson et al., 2021	15	15	30
Kurvits et al., 2021	12	12	25
Craig et al., 2021	825	825	2,019
Riboldi et al., 2022	743	527	1,285
Li et al., 2023	691	594	2,432
Irmady et al., 2023	479	195	1,584

Table 3.1. Blood PD RNA-seq studies investigating differences in gene expression between PD patients and controls. The number of differentially expressed genes is shown as reported in the respective publications.

The sample sizes of initial studies were generally quite small (ranging from 6-20 PD patients, **Table** 3.1), potentially limiting their ability to detect gene expression changes in PD. Large scale projects, such as the Parkinson's Progression Marker Initiative (PPMI, Marek et al., 2011), have collected larger sample sizes in combination with rich phenotypic information and high-throughput molecular measurement for the detection of disease related changes and biomarker identification. Analysis of PPMI RNA-seq data identified over 2,000 genes differentially expressed in PD patients compared to controls considering all longitudinal data (Craig et al., 2021). Analysis of PPMI RNA-seq data from other groups has also detected widespread changes in blood gene expression in PD patients (Riboldi et al., 2022; Li et al., 2023; Irmady et al., 2023) (**Table 3.1**). Analyses of PPMI RNA-seq data included PD patients of varying subtypes (some harbour known pathogenic PD risk variants) and disease durations. Therefore, there is a need to understand if there are blood gene expression changes in idiopathic PD patients in the early disease stages.

The ability of blood gene expression to distinguish PD patients from controls has also been assessed (Scherzer et al., 2007; Grünblatt et al., 2010; Molochnikov et al., 2012; Santiago & Potashkin, 2013, 2015; Santiago et al., 2016; Shamir et al., 2017; Makarious et al., 2022; Pantaleo et al., 2022; Li et al., 2023) (**Table 3.2**). Initial studies assessed the expression of multiple genes in their ability to distinguish PD patients from controls (Scherzer et al., 2007; Grünblatt et al., 2010; Molochnikov et al., 2012; Santiago & Potashkin, 2013, 2015). The best performance was reported by Molochnikov et al, with a five-gene panel exhibiting an AUC of 0.96 when distinguishing early-stage PD patients from controls. However, these genes were

selected, and performance was evaluated in the same cohort which may lead to overoptimistic performance estimates. Overall, despite promising performance in distinguishing PD from patients, there was a lack of consistently identified informative genes. Amongst these studies, only *HIP2* (Scherzer et al., 2007; Molochnikov et al., 2012) and *ALDH1A1* (Grünblatt et al., 2010; Molochnikov et al., 2012) were identified in multiple studies. This may be due to limited sample numbers, with these studies including <100 PD patients. Additionally, clinically heterogeneous PD patients may have been assessed.

Using a larger sample size of 175 PD patients and 193 controls, Shamir et al identified a panel of 87 genes that distinguished PD from controls with an AUC of 0.79 using unseen data and an AUC of 0.74 in an additional validation dataset (Shamir et al., 2017). Whilst Shamir et al included idiopathic PD patients, not all PD patients were early stage so differences in gene expression may reflect those in later stages of disease progression. Additionally, the use of microarray technology to measure gene expression may miss low abundance or rare transcripts. RNA-seq data from PPMI participants have identified the ability of gene expression to distinguish PD patients from controls with AUCs ranging from 0.72-0.81 (Makarious et al., 2022; Pantaleo et al., 2022; Li et al., 2023). Notably, some studies restricted PPMI PD patients to those enrolled as idiopathic, which due to PPMI enrolment criteria, means they were early-stage idiopathic PD patients (Makarious et al., 2022; Pantaleo et al., 2022). Li et al also assessed the ability of gene expression to distinguish PD patients from controls in an external dataset (Parkinson's Disease Biomarker Program, PDBP) with an AUC of 0.69, which may reflect differences in cohort enrolment demographics (Li et al., 2023).

While gene expression shows potential clinical utility for detecting PD, it is unclear if these genes consistently reflect the earliest stages of PD across different cohorts. Here, I used RNA-seq data to measure whole blood gene expression in patients recently diagnosed with idiopathic PD and matched controls. Participants were drawn from two independent cohorts (PPMI (Marek et al., 2011) and ICICLE-PD (Yarnall et al., 2014)) to replicate findings. I compared gene expression between early-stage idiopathic PD patients and controls to identify genes differentially expressed and dysregulated gene sets in early-stage idiopathic PD. I also evaluated the use of gene expression to distinguish early-stage idiopathic PD patients from controls in the PPMI and ICICLE-PD cohorts.

Publication	Genes	Discovery			Replication		
		PD	*Control*	*AUC*	*PD*	*Control*	*AUC*
Scherzer et al., 2007	*VDR*, *HIP2*, *CLTB*, *FPRL2*, *CA12*, *CEACAM4*, *ACRV1*, and *UTX*	31	35	NA	19	20	NA
Grünblatt et al., 2010	*PSMA2*, *LAMB2*, *ALDH1A1*, and *HIST1H3E*	94	34	0.92	22	33	0.93
Molochnikov et al., 2012	*HIP2*, *KP1A*, *ALDH1A1*, *PSMC4* and *HSPA8*	62	64	0.96			
Santiago and Potashkin, 2013	*APP*	50	46	0.80	51	45	0.81
Santiago and Potashkin, 2015	*PTBP1* and *HNF4A*	95	91	0.90			
Santiago et al., 2016	*NAMPT*	99	101	0.63			
Shamir et al., 2017	87 genes	175	193	0.79	30	40	0.74
Makarious et al., 2022	597 genes	427	171	0.80			
Pantaleo et al., 2022	493 genes	390	189	0.72			
Li et al., 2023	257 genes	691	594	0.81	702	458	0.69

Table 3.2. Blood-based gene expression PD biomarkers. Results from a series of publications that have assessed the ability of gene expression to distinguish between PD patients and controls. For each publication, the genes assessed are given along with the sample size and performance of the biomarkers. AUC = Area under the receiver operator characteristic curve.

3.2 Methods

3.2.1 Cohorts

Samples used in analyses presented in this chapter have been described in **Section 2.2**. Briefly, two independent cohorts were used from the Parkinson's Progression Marker Initiative (PPMI) and the Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-PD study (ICICLE-PD). Within each cohort, I have restricted to investigating patients with early-stage idiopathic PD and matched controls (**Section 2.2**).

3.2.2 RNA sequencing

Sample collection, RNA extraction, library preparation and sequencing of PPMI and ICICLE-PD samples are described in **Section 2.3**.

3.2.3 Gene quantification

Transcripts were quantified with Salmon v1.3.0 (Patro et al., 2017) as described (**Section 2.5.1**). The median number of reads mapped to the transcriptome was 23.8 million (IQR = 9.8) and 35.4 million (IQR = 9.1) for PPMI and ICICLE-PD samples respectively, corresponding to 22.3% (IQR = 3.9%) and 39.9% (IQR = 6.4%) of all reads.

Transcript counts were summarised to the gene level using tximport v1.28.0 for downstream use with DESeq2 as described (**Section 2.5.2**)

3.2.4 RNA sequencing quality control

A series of RNA sequencing quality control steps were performed as described in **Section 2.4**.

3.2.5 Identifying sources of extraneous variation in gene expression

I quantified the contribution of biological and technical sequencing measures on gene expression using the methods described in **Section 2.7**. Based on the information described in **Section 3.3.1**, I included the percentage of usable bases (PPMI), the percentage of intronic bases (ICICLE-PD) and the percentage of coding bases (ICICLE-PD) as explanatory variables in linear mixed models predicting individual gene expression using variancePartition (Hoffman & Schadt, 2016). I also included sequencing batch, age, and sex as explanatory variables in the

linear mixed models. Categorical variables (study group, sequencing batch, and sex) were modelled as random effects as recommended (Hoffman & Schadt, 2016).

3.2.6 Differential expression

Testing for genes differentially expressed between PD patients and controls was performed using DESeq2 v1.40.2 (Love et al., 2014) using the methodology in **Section 2.8**.

Biological and technical sources of gene expression variation were included as covariates when testing for differential expression (**Section 3.3.1**). For both cohorts, this included age at sample collection (binned into <55, 55-65, >65 years), sex and sequencing batch. In PPMI, I included the percentage of usable bases. In ICICLE-PD, I included the percentages of coding and intronic bases.

3.2.7 Gene set enrichment analysis

Gene set enrichment analysis (GSEA) (Subramanian et al., 2005) was performed using clusterProfiler v4.8.3 (Wu et al., 2021). For each cohort, all genes included in differential expression testing were ranked by fold change. Gene sets were obtained as Gene Ontologies (Ashburner et al., 2000) (including biological processes, molecular functions, and cellular compartments) and KEGG pathways (Kanehisa et al., 2016) using the *gseGO* and *gseKEGG* functions, respectively.

Based on their Normalised Enrichment Score (NES), gene sets may be either positively (NES >0) or negatively (NES <0) enriched depending on whether set members are more likely to show increased or reduced expression respectively in idiopathic PD patients compared to controls.

3.2.8 Effect of dopaminergic PD treatment dosage on gene expression

Idiopathic PD patients enrolled in PPMI were not receiving PD medication (e.g., L-DOPA, dopamine agonists, monoamine oxidase inhibitors etc – detailed at PPMI.org) at recruitment (i.e., baseline visit). Conversely, 91.7% (44/48) of ICICLE-PD PD patients were receiving PD medication at the baseline visit (**Section 2.1.1**). Testing the effect of PD treatment dosage on RNA expression was subsequently performed using ICICLE-PD PD patients. A linear regression model was used to model variance-stabilised gene expression counts (Anders & Huber, 2010), incorporating treatment dosage (given as the levodopa equivalent daily dose, LEDD), age at sample collection, sex, percentage of coding bases and percentage of intronic

bases. The coefficient and *P*-value for the LEDD term were extracted for each gene. False discovery rates were calculated using the Benjamini-Hochberg procedure (*p.adjust*).

3.2.9 Correlating expression with PD-related clinical measures

PD-related clinical measures were collected from study files produced by the respective studies (**Section 2.2**), including the age at diagnosis (years), MDS-UPDRS part III scores (assessing motor function) and MoCA scores (representing cognitive impairment). Where possible, comparisons between PD patients and controls are shown in **Section 2.2**.

To test the association of RNA expression with clinical measures, linear regression models were used to model variance-stabilised transformed gene expression counts (Anders & Huber, 2010), incorporating the clinical measure along sources of biological and technical sources of variation (**Section 3.3.1**). The coefficients and *P*-value for the expression of each gene tested were extracted. *P*-values were adjusted for multiple testing (where necessary) using a Bonferroni correction.

3.2.10 Replication of previously reported differentially expressed genes in blood

A total of four studies were identified that quantified gene expression in the blood of PD patients using RNA-seq and for which summary results were available (Infante et al., 2016; Garofalo et al., 2020; Henderson et al., 2021; Kurvits et al., 2021), The PPMI has published their analysis (Craig et al., 2021) yet the results from this analysis were not included due to the overlap of samples. A total of 390 unique genes (as determined by unique Ensembl gene IDs) were assessed. Of these, 256 were detected in PPMI and 253 in ICICLE-PD.

3.2.11 Genes underlying the genetic architecture of PD

Loci associated with PD risk were collected from the current largest GWAS (Nalls et al., 2019). This identified 99 unique genes in proximity to these risk loci. 70 were detected in PPMI, and 69 in ICICLE-PD.

35 genes with established pathogenic variants, including *SNCA*, *LRRK2* and *GBA*, were identified in the Genomics England Parkinson Disease and Complex Parkinsonism gene panel v1.111 (accessed 10/2022) (Martin et al., 2019). Of these, 30 genes were included in differential expression analysis in PPMI and ICICLE-PD.

3.2.12 Classification of PD status using gene expression

Normalised gene expression counts were calculated based on the median of ratios method as implemented in DESeq2 v1.40.2. Genes with altered expression (*P*-value < 0.05) between idiopathic PD patients and controls in PPMI were included as predictors. A regularised logistic regression model was then created, and performance assessed as described in **Section 2.9**.

3.3 Results

3.3.1 Sources of extraneous variation in gene expression

Gene expression was quantified from whole blood total RNA-sequencing (RNA-seq) data from individuals with early-stage idiopathic Parkinson's disease (PD) (PPMI n = 259, ICICLE-PD n = 48) along with matched controls (PPMI n = 161; ICICLE-PD n = 48) (detailed information regarding samples is provided in **Section 2.2**).

Biological and technical variation can impact the estimation of gene expression ('t Hoen et al., 2013). To assess structure within the gene expression data, I projected gene expression values into the first two principal components (PCs) (**Figure 3.1**). This revealed no clear segregation of samples by study group (**Figure 3.1**). I evaluated other sources of variation contributing to the quantification of gene expression using the methods outlined in **Section 3.2.5.**

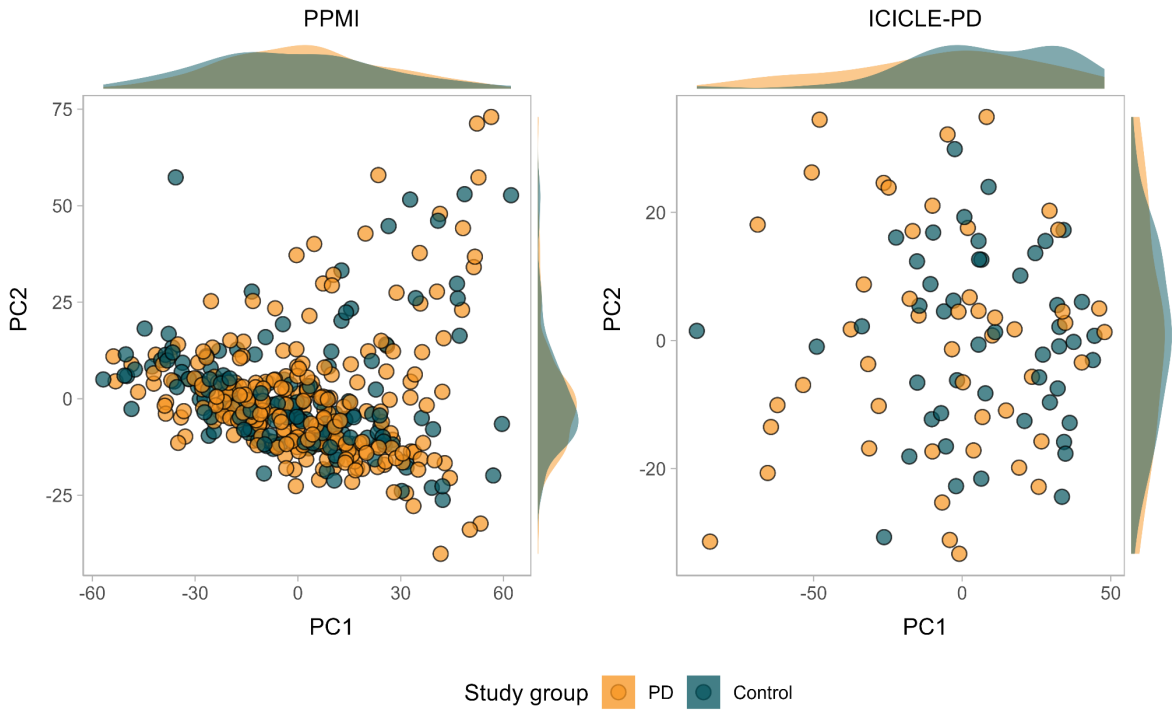


Figure 3.1. The first two principal components of gene expression in PPMI and ICICLE-PD. The marginal densities of the points are shown on the edges of the plot.

I identified technical sources of variation based on the relationship between various sequencing metrics and PCs of gene expression. The first 10 PCs accounted for most of the variance in PPMI (56.9%) and ICICLE-PD (61.4%) (**Figure 3.2a-b**). PC1 generally correlated with

multiple sequencing variables and explained 22.3% and 30.2% of the total variance in PPMI and ICICLE-PD respectively (**Figure 3.2a-b)**. Specifically in PPMI, the percentage of reads mapping to usable bases correlated to PC1 (R^2 = 0.51) (**Figure 3.2a**), in line with previously published data from the entire cohort (Craig et al., 2021). Sequencing batch and PC9 were also correlated (R^2 = 0.61) (**Figure 3.2a**). In ICICLE-PD, multiple sequencing metrics could individually explain most of the variance (R^2 >0.5) in PC1 (**Figure 3.2b**). These included the percentage of untranslated region (UTR) bases (R^2 = 0.52), percentage of usable bases (R^2 = 0.63), percentage of mRNA bases (R^2 = 0.66), percentage of intronic bases (R^2 = 0.7) and percentage of coding bases (R^2 = 0.63) (**b**).

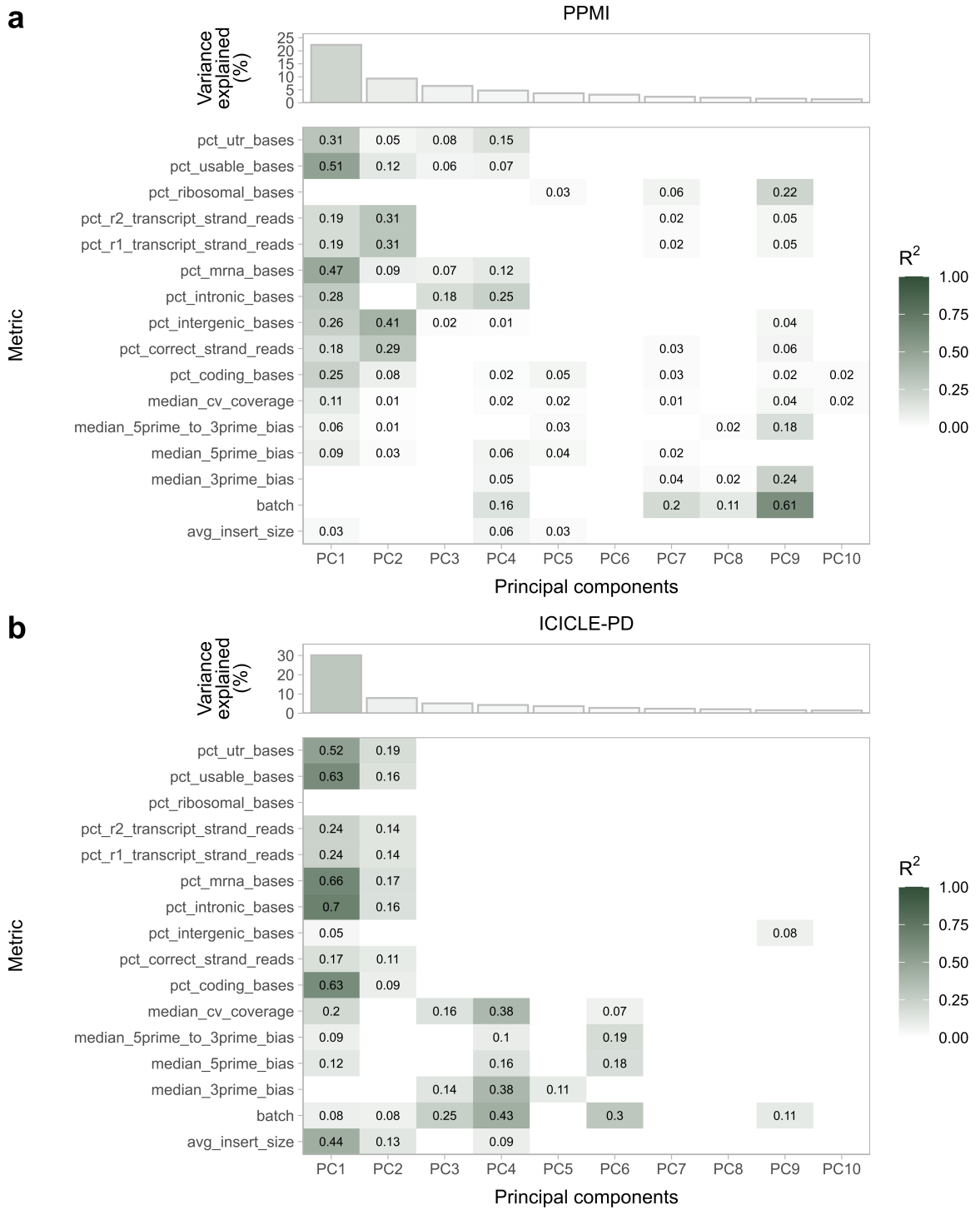


Figure 3.2. Identifying sources of extraneous technical variation when quantifying gene expression. (a, b) Results from univariate linear regression of technical metrics in the PPMI (**a**) and ICICLE-PD (**b**) cohorts against each of the first 10 principal components (PCs) of gene expression. The top panel of each figure shows the percentage of variation explained by each PC. The amount of variation explained (R^2) by each technical metric is only shown for metrics which passed multiple testing correction (FDR <0.05).

Many measurable sequencing metrics were correlated (**Figure 3.3a-b**). To avoid redundancy, I measured the correlation between informative metrics (R^2 >0.5) from **Figure 3.2a-b**. In PPMI, the percentage of usable bases was the only sequencing metric highlighted (**Figure 3.2a, Figure 3.3c**). In ICICLE-PD, I assessed the pairwise correlations between the percentages of UTR bases, usable bases, mRNA bases, intronic bases and coding bases. I was primarily concerned with other metrics correlated with the percentage of intronic bases as this metric explained the most variation in PC1. All the selected measures except for the percentage of coding bases were negatively correlated (absolute Spearman's rho >0.9) with the percentage of intronic bases (**Figure 3.3d**). As such, I flagged the percentages of intronic and coding bases in each ICICLE-PD sample as possible technical covariates. Overall, inter-cohort differences in technical variation were expected considering the different library preparation kits used (**Section 3.2.2**)

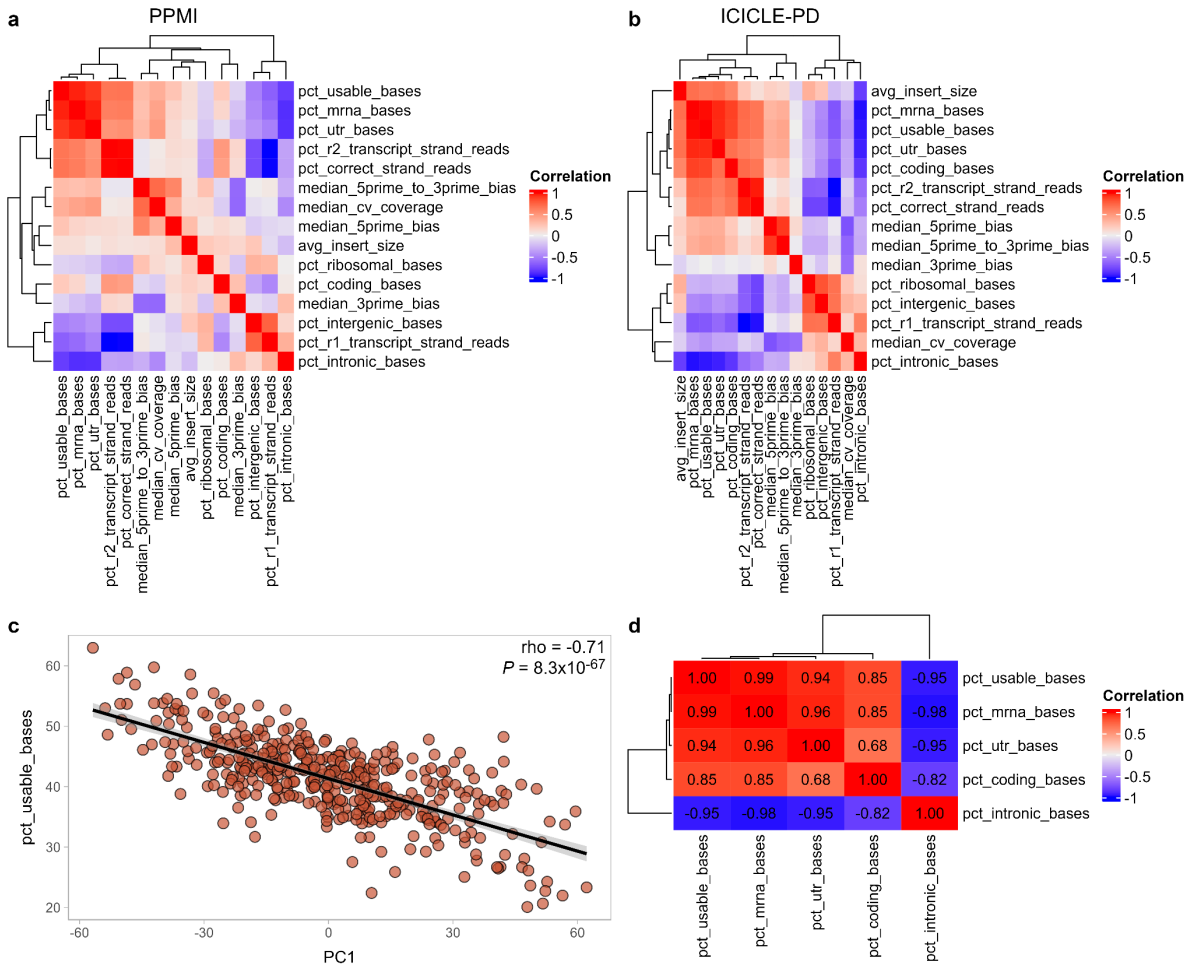


Figure 3.3. Correlations of technical sequencing metrics. (a, b) Pairwise correlations between technical sequencing metrics in PPMI (**a**) and ICICLE-PD (**b**). (**c**) Negative correlation between the percentage of usable bases and PC1 in PPMI. (**d**) Pairwise correlations between potential covariates in ICICLE-PD. All correlations are given as Spearman's rho coefficient. Metric ordering was determined by hierarchical clustering based on Euclidean distance.

Next, I determined how sources of technical variation identified in the previous step (PPMI: percentage of usable bases; ICICLE-PD: percentages of intronic and coding bases), as well as common sources of technical (i.e., sequencing batch) and biological variation (i.e., age, sex, study group) influenced the expression of individual genes. Using these biological and technical covariates, I fitted linear mixed models to the expression of individual genes using variancePartition (Hoffman & Schadt, 2016) (**Section 3.2.5**). The selected covariates tended to explain more variance in gene expression than the study group (labelled condition) (**Figure 3.4**), highlighting the importance of their inclusion in downstream analyses.

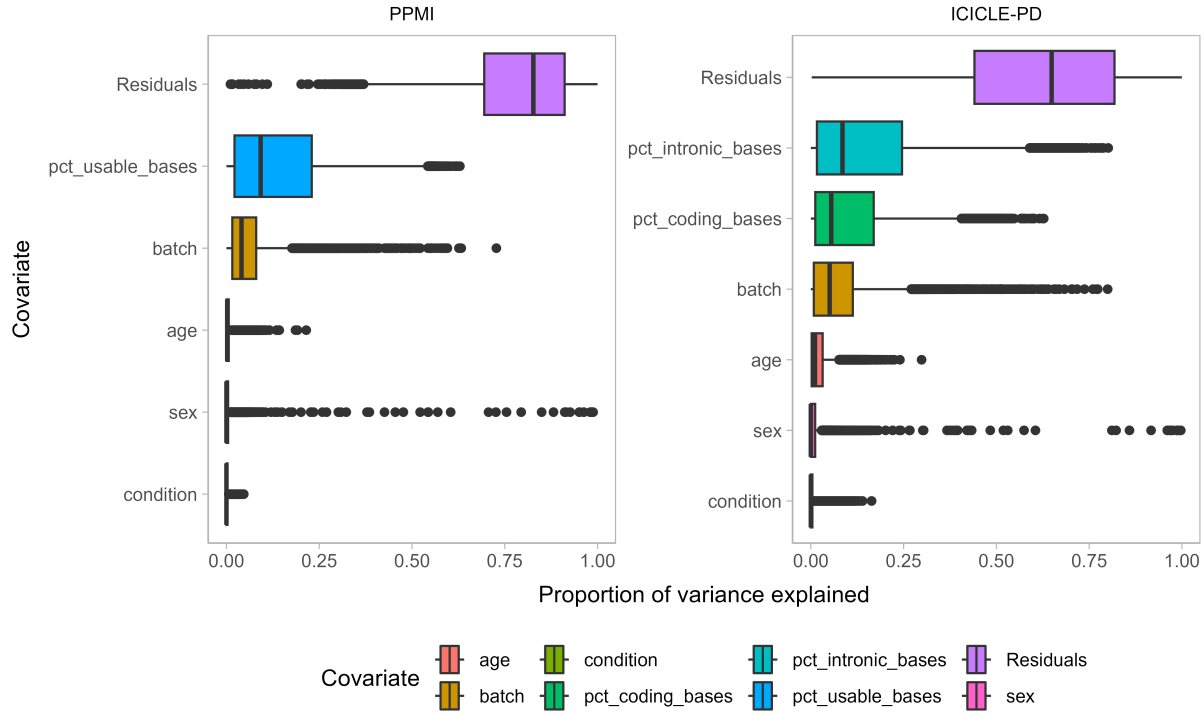


Figure 3.4. Contribution of covariates to individual gene expression. Each point represents a gene, showing the proportion of variation explained by each covariate on the x-axis.

3.3.2 Identification of differentially expressed genes in early-stage idiopathic PD patients from PPMI and ICICLE-PD studies

I performed differential gene expression analysis on 16,191 (PPMI) and 15,852 (ICICLE-PD) genes respectively (overlap of 15,655 genes or 95.53%), adjusting for sources of biological and technical variation (**Section 3.3.1**).

Using the PPMI as a discovery cohort, I identified 44 genes differentially expressed in idiopathic PD patients compared to controls (log_2 fold change (log_2FC) >0.1/<-0.1, false discovery rate (FDR) < 0.05, Wald test) (**Figure 3.5a**). This included 28 upregulated genes and 16 downregulated genes (**Figure 3.5a**). ICICLE-PD data was then used as a replication cohort to identify significant DEGs with the same direction of effect. Of the 44 PPMI differentially expressed genes, *TMEM252* and *LMNB1* were similarly increased in ICICLE-PD idiopathic PD patients compared to controls (log_2FC >0.1, FDR < 0.05, Wald test) (**Figure 3.5b**).

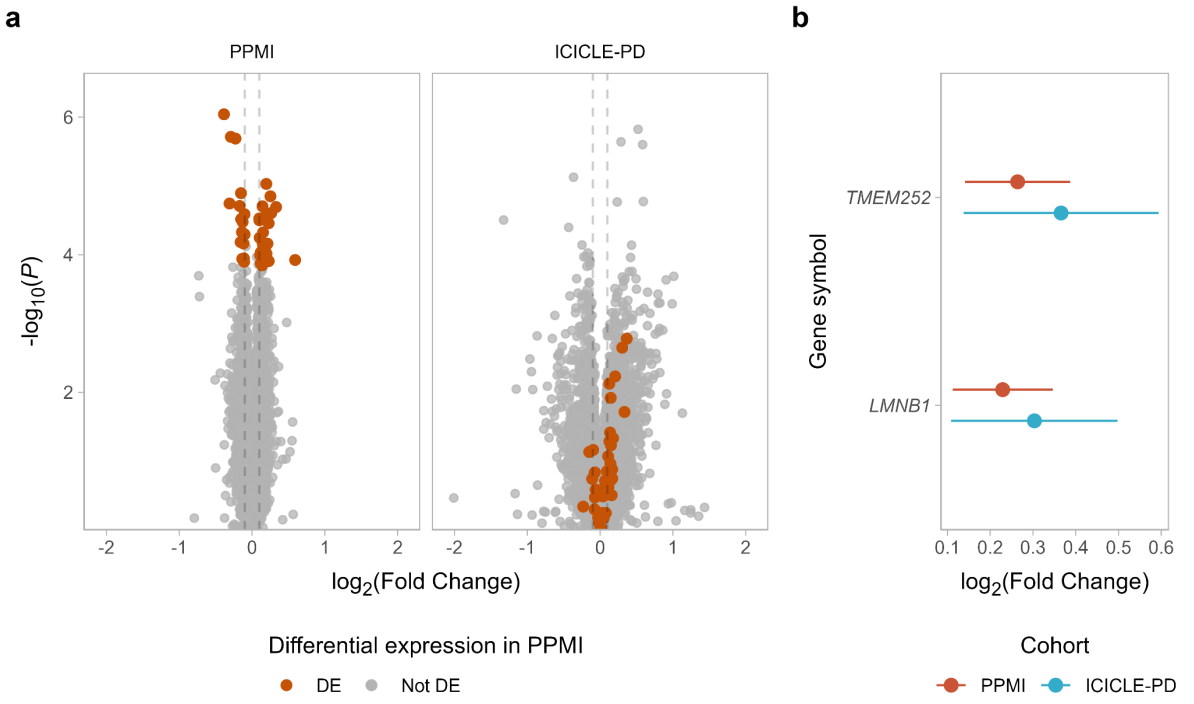


Figure 3.5. Differential expression of genes in early-stage idiopathic PD. (a) Volcano plot showing the differential expression of genes in PPMI and ICICLE-PD. Highlighted genes are those significantly differentially expressed (log_2 fold change >0.1/<-0.1, FDR <0.05) in PPMI. **(b)** Graph showing the comparative log_2 fold change and 95% CI of genes (*TMEM252* and *LMNB1*) significant in both PPMI and ICICLE-PD.

3.3.3 Correlating gene expression to PD clinical measures

As most ICICLE-PD PD patients were receiving PD medication at enrolment (**Section 2.2**), I first explored the association between gene expression and dopaminergic treatment dosage adjusting for biological and technical sources of variation (**Section 3.2.8**). Ultimately, treatment dosage (levodopa equivalent daily dose, LEDD) was not associated with the expression of any PPMI differentially expressed genes (FDR > 0.05, linear regression) (**Table 3.3**).

Gene	Ensembl ID	β (95% CI)	*P*-value	FDR
CEP19	ENSG00000174007	0.0013 (4.4e-05, 0.0026)	0.05	0.91
PRRG4	ENSG00000135378	0.0014 (-1.7e-05, 0.0028)	0.061	0.91
ZNF33A	ENSG00000189180	0.00061 (-4.6e-05, 0.0013)	0.075	0.91
POLD2	ENSG00000106628	-0.00081 (-0.0017, 8.5e-05)	0.086	0.91
SYTL3	ENSG00000164674	-0.00077 (-0.0018, 0.00022)	0.13	0.91
ISG20	ENSG00000172183	0.00077 (-0.00031, 0.0019)	0.17	0.91
KIAA0040	ENSG00000235750	0.00062 (-0.00026, 0.0015)	0.18	0.91
CREBRF	ENSG00000164463	0.00041 (-0.00023, 0.001)	0.22	0.91
RIT1	ENSG00000143622	0.00047 (-0.00032, 0.0013)	0.25	0.91
MTND4P30	ENSG00000227454	-0.00045 (-0.0012, 0.00032)	0.26	0.91
ARHGAP19	ENSG00000213390	5e-04 (-0.00037, 0.0014)	0.27	0.91
CISD2	ENSG00000145354	-0.00087 (-0.0025, 0.00075)	0.3	0.91
HECTD2	ENSG00000165338	0.00055 (-0.00056, 0.0017)	0.34	0.91
TIMM9	ENSG00000100575	-0.00051 (-0.0016, 0.00053)	0.34	0.91
FPR2	ENSG00000171049	7e-04 (-0.00076, 0.0022)	0.35	0.91
RPL15	ENSG00000174748	-0.00041 (-0.0014, 0.00053)	0.4	0.91
RNASEL	ENSG00000135828	0.00029 (-0.00043, 0.001)	0.43	0.91
DPP7	ENSG00000176978	0.00022 (-0.00034, 0.00078)	0.44	0.91
HSDL2	ENSG00000119471	0.00037 (-0.00056, 0.0013)	0.44	0.91
MCTP2	ENSG00000140563	0.00036 (-0.00055, 0.0013)	0.45	0.91
TSTA3	ENSG00000104522	-0.00067 (-0.0025, 0.0012)	0.49	0.91
CLEC4E	ENSG00000166523	0.00068 (-0.0013, 0.0026)	0.5	0.91
ZNF517	ENSG00000197363	-2e-04 (-0.00082, 0.00042)	0.52	0.91
AIDA	ENSG00000186063	-0.00031 (-0.0013, 0.00071)	0.55	0.91
SLA	ENSG00000155926	0.00026 (-0.00062, 0.0011)	0.56	0.91
TMCC3	ENSG00000057704	0.00036 (-0.0011, 0.0018)	0.62	0.91
AC092032.2	ENSG00000288620	-0.00013 (-0.00068, 0.00042)	0.63	0.91
DCK	ENSG00000156136	0.00019 (-0.00058, 0.00096)	0.64	0.91
ROPN1L	ENSG00000145491	0.00026 (-0.00081, 0.0013)	0.64	0.91
RNF149	ENSG00000163162	0.00021 (-0.00075, 0.0012)	0.67	0.91
FAHD2A	ENSG00000115042	-0.00016 (-0.00089, 0.00057)	0.67	0.91
PPIAP8	ENSG00000225185	0.00017 (-0.00064, 0.00098)	0.67	0.91
GSTZ1	ENSG00000100577	0.00014 (-0.00053, 0.00081)	0.69	0.91
BMX	ENSG00000102010	-0.00035 (-0.0022, 0.0015)	0.7	0.91
ZNF555	ENSG00000186300	-0.00011 (-0.00091, 0.00069)	0.79	0.93
PPIAP40	ENSG00000230593	9.8e-05 (-0.00063, 0.00083)	0.79	0.93
TMEM252	ENSG00000181778	0.00019 (-0.0013, 0.0016)	0.8	0.93
LMNB1	ENSG00000113368	0.00017 (-0.0013, 0.0016)	0.82	0.93
SLC25A20	ENSG00000178537	9.1e-05 (-0.00073, 0.00091)	0.83	0.93
KPNB1	ENSG00000108424	8.5e-05 (-0.00087, 0.001)	0.86	0.93
THUMPD2	ENSG00000138050	6.8e-05 (-0.00089, 0.001)	0.89	0.93
YJU2	ENSG00000105248	-3.7e-05 (-0.00062, 0.00055)	0.9	0.93
AC011816.1	ENSG00000234073	6.9e-05 (-0.0011, 0.0012)	0.9	0.93
CD300LD	ENSG00000204345	7.3e-05 (-0.0018, 0.0019)	0.94	0.94

Table 3.3. Effect of dopaminergic PD treatment on gene expression in ICICLE-PD. A linear regression model was fit as described in **Section 3.2.8**. β coefficients show the modelled change in variance-stabilised gene expression based on per unit change in treatment dosage (mg/day). *P*-values were adjusted for multiple testing based on the Benjamini-Hochberg method.

Given increased *TMEM252* and *LMNB1* expression in early-stage idiopathic PD patients (**Figure 3.5b**), I then compared the expression of these genes to PD-related clinical measures. I used linear regression to test the association between gene expression and the age of PD

diagnosis, motor severity (assessed using MDS-UPDRS III scores) and cognitive impairment (assessed using MoCA scores) (**Section 3.2.9**). I found no significant associations between *TMEM252* and *LMNB1* expression and age of PD diagnosis or motor symptom severity (*P*-value > 0.05, linear regression) (**Table 3.4**). I detected a negative association between *TMEM252* expression and cognitive impairment (β = -2.5, *P*-value = 0.04, linear regression); however, the association was not significant after adjustment for multiple testing (Bonferroni *P*-value = 0.09, two tests, linear regression) (**Table 3.4**).

			β (95% CI)		*P*-value		Adjusted *P*-value	
Response variable	**Gene symbol**	**Gene ID**	*PPMI*	*ICICLE-PD*	*PPMI*	*ICICLE-PD*	*PPMI*	*ICICLE-PD*
Age at PD diagnosis								
	LMNB1	ENSG00000113368	-0.086 (-0.99, 0.82)	0.5 (-2.6, 3.6)	0.85	0.74	1.00	1.00
	TMEM252	ENSG00000181778	0.017 (-1.4, 1.4)	0.2 (-2.9, 3.3)	0.98	0.90	1.00	1.00
MDS-UPDRS III								
	LMNB1	ENSG00000113368	0.92 (-0.82, 2.7)	-4.2 (-12, 3.6)	0.30	0.29	0.60	0.57
	TMEM252	ENSG00000181778	1 (-1.7, 3.7)	1.8 (-6.2, 9.7)	0.46	0.66	0.91	1.00
MoCA								
	LMNB1	ENSG00000113368	-0.22 (-0.74, 0.29)	-0.38 (-2.9, 2.2)	0.39	0.77	0.79	1.00
	TMEM252	ENSG00000181778	-0.25 (-1.1, 0.56)	-2.5 (-4.9, -0.084)	0.55	0.04	1.00	0.09

Table 3.4. Associations between differentially expressed genes and PD-related clinical measures. Table giving the output from linear regression of various clinical measures as the response variables and the corresponding gene expression as an explanatory variable (**Section 3.2.9**). MDS-UPDRS III = Movement Disorder's Society Unified Parkinson's Disease Rating Scale Part 3, MoCA = Montreal Cognitive Assessment.

3.3.4 Assessing systematic changes in gene expression in early-stage idiopathic PD patients

To gain an overview of global alterations in gene expression patterns, I performed gene set enrichment analysis (GSEA) (Subramanian et al., 2005) of Gene Ontology (GO) terms (Ashburner et al., 2000). All GO terms significantly enriched (FDR < 0.05) in PPMI and subsequently replicated in ICICLE-PD (FDR < 0.05 with the same direction of effect) are given in **Table 8.1**.

Multiple gene groups broadly related to ribosome function were downregulated in idiopathic PD patients in both cohorts (**Figure 3.6a, Table 8.1**). In PPMI, the most downregulated biological process was *cytoplasmic translation* (NES = -3.1, FDR < 2.2×10^{-16}) which was also significantly downregulated in ICICLE-PD (NES = -2.2, FDR = 2.3×10^{-6}) (**Figure 3.6a**, **Table 8.1**). Likewise, *mitochondrial translation* was also negative enriched (PPMI NES = -2.1, FDR = 1.4×10^{-4}; ICICLE-PD NES = -1.9, FDR = 1.3×10^{-4}) (**Figure 3.6a**, **Table 8.1**). In PPMI, the most downregulated molecular function was *structural constituent of ribosome* (NES = -2.9, FDR < 2.2×10^{-16}), again replicated in ICICLE-PD (NES = -2.2, FDR = 4.8×10^{-8}) (**Figure 3.6a, Table 8.1**). Together, the collective downregulation of ribosomal-related genes may reflect reduced cytoplasmic and mitochondrial protein translation in the blood of idiopathic PD patients.

Conversely, when ranked by PPMI *P*-values, I found that gene sets associated with activation of the immune response, particularly antiviral responses, were among the most significantly upregulated in idiopathic PD patients (**Figure 3.6a, Table 8.1**). These included *negative regulation of viral process* (PPMI NES = 2.6, FDR = 1.9×10^{-7}; ICICLE-PD NES = 2.2, FDR = 1.4×10^{-7}) and *regulation of innate immune response* (PPMI NES = 2.0, FDR = 8.9×10^{-6}; ICICLE-PD NES = 1.7, FDR = 2.8×10^{-4}) (**Figure 3.6a, Table 8.1**). The induction of an innate immune response may be due to the activation of the *pattern recognition receptor signaling pathway* (PPMI NES = 1.9, FDR = 4.0×10^{-4}; ICICLE-PD NES = 2.2, FDR = 1.9×10^{-9}, **Table 8.1**) which recognise pathogen- or damage-associated molecular patterns (Li & Wu, 2021). The expression of genes associated with signalling pathways downstream of pattern recognition receptor signalling such as the *type I interferon signaling pathway* (PPMI NES = 2.2, FDR = 9.7×10^{-4}; ICICLE-PD NES = 2.1, FDR = 3.9×10^{-5}) and *regulation of inflammatory response* (PPMI NES = 1.6, FDR = 6.2×10^{-3}; ICICLE-PD NES = 1.5, FDR = 2.1×10^{-3}) were also increased in idiopathic PD patients (**Table 8.1**).

I also performed GSEA using KEGG pathways (Kanehisa & Goto, 2000; Kanehisa et al., 2023) as an alternative source of curated gene sets. Focusing on pathways related to those identified

through GSEA of GO terms, I also found enrichment of the KEGG pathways *Ribosome* (PPMI NES = -3.0, FDR< $2.2x10^{-16}$; ICICLE-PD NES = -2.5, FDR = $5.6x10^{-10}$) and NOD-like receptor signalling (PPMI NES = 1.8, FDR = 0.002; ICICLE-PD NES = 2.0, FDR = $2.3x10^{-6}$), part of the innate immune response (Zhong et al., 2013) (**Figure 3.6a, b**).

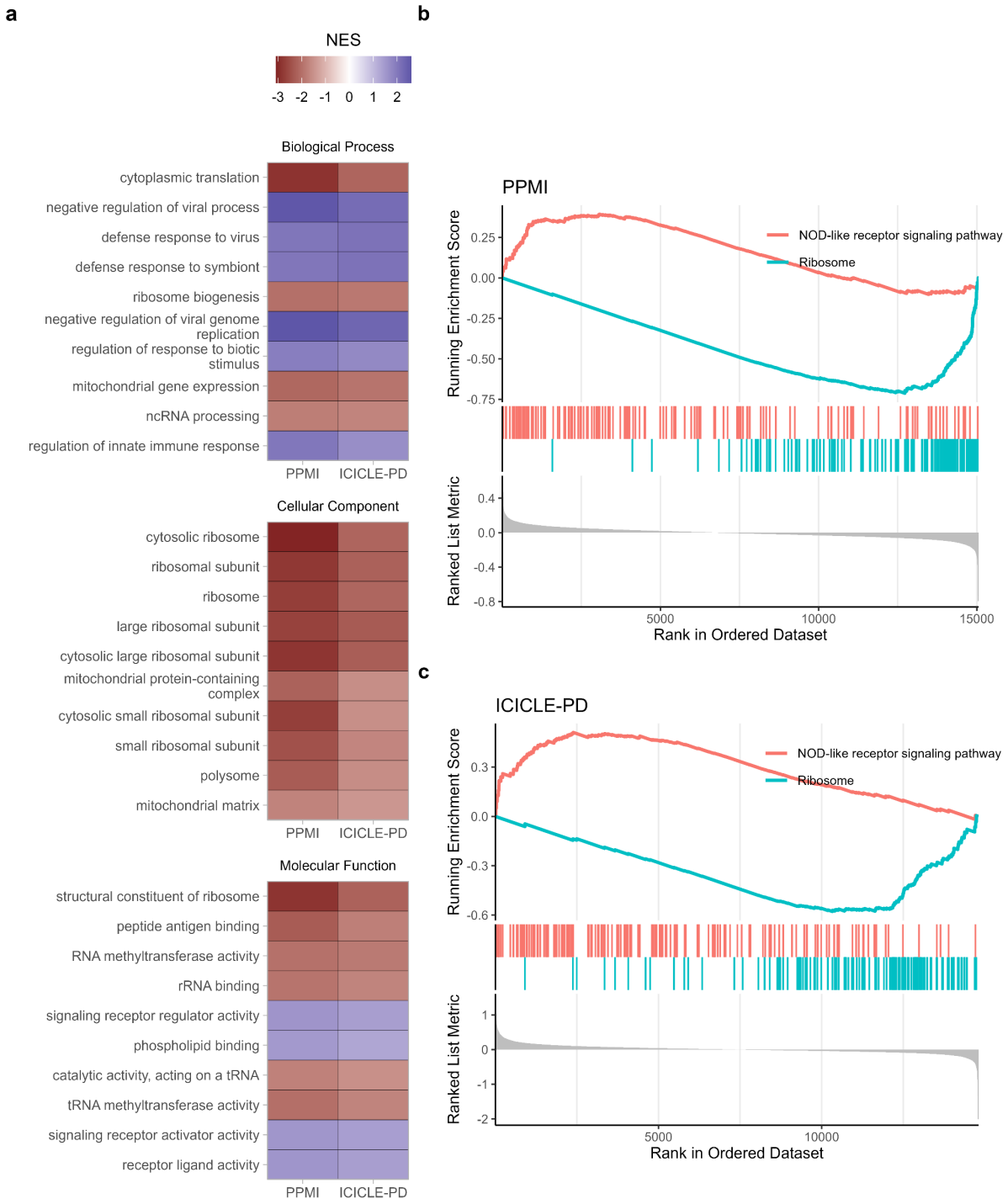


Figure 3.6. Gene sets enriched in early-stage idiopathic PD patients. (a) Gene set enrichment of Gene Ontology terms in PPMI and ICICLE-PD cohorts. Enrichment was performed with ranked fold changes calculated by comparing the expression of genes between idiopathic PD patients and controls. The top ten replicated terms per gene ontology are arranged according to the PPMI *P*-value. NES = Normalised Enrichment Score. **(b-c)** Gene set enrichment of two selected KEGG in PPMI (**b**) and ICICLE-PD (**c**).

3.3.5 *Replication of previous blood RNA-seq studies in PD*

Previous studies have also leveraged RNA-seq of blood samples to identify genes differentially expressed in PD patients (Infante et al., 2016; Garofalo et al., 2020; Henderson et al., 2021; Kurvits et al., 2021). After collating summary statistics from these studies (**Section 3.2.10**), I examined their expression in PPMI and ICICLE-PD. I independently replicated (FDR <0.05, Wald test) several previously reported genes in PPMI (*LSMEM1*, *TPST1* and *SLED1*) and ICICLE-PD (*IFIT1*, *RSAD2*, *IFI44L* and *OLIG2*) (**Figure 3.7**).

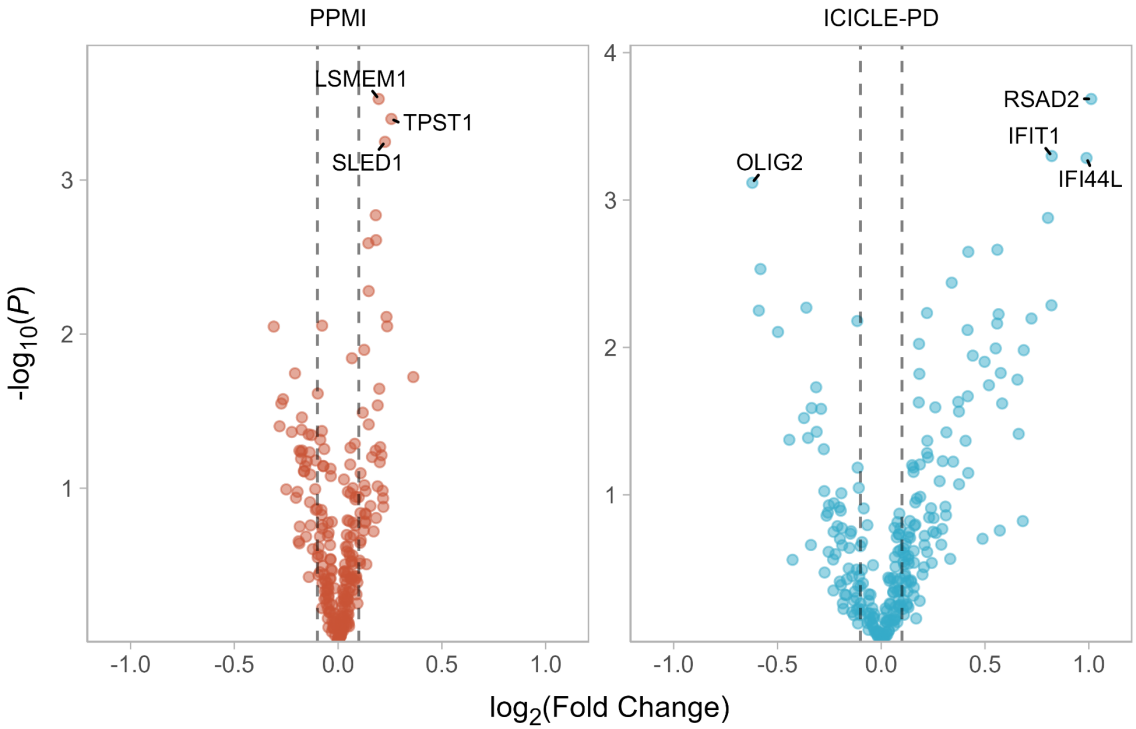


Figure 3.7. Replication of previously reported differentially expressed genes in the blood of PD patients. Volcano plots showing the differential expression of genes reported as differentially expressed in PD from previous blood RNA-seq studies. Genes significantly differentially expressed (FDR <0.05 and $\log_2$ fold change >0.1/<-0.1) in each cohort are labelled.

3.3.6 *Differential expression of genes underlying the genetic architecture of PD*

Genetic variants have been linked to the risk of developing PD (Blauwendraat et al., 2020). The current largest genome-wide association study (GWAS) assessing PD risk in individuals of European genetic ancestry, identified 90 variants within 78 loci (Nalls et al., 2019). Examining the expression of the nearest genes to the identified PD risk loci (70 in PPMI, 69 in ICICLE-PD), five were differentially expressed ($\log_2$FC >0.1/<-0.1, FDR < 0.05, Wald test) in PPMI (*VAMP4*, *BST1*, *FCGR2A*, *SIPA1L2* and *NOD2*) **Figure 3.8a**). In ICICLE-PD, the direction of expression change was concordant for *BST1*, *FCGR2A*, *NOD2* and *SIPA1L2* (**Figure 3.8b**).

BST1 showed suggestive evidence of differential expression (*P*-value < 0.05, Wald test) although this was not significant after correcting for multiple testing (FDR = 0.09)

I then examined expression levels of 30 genes associated with monogenic forms of PD and other parkinsonism disorders from Genomics England's PanelApp (**Section 3.2.11**) (Martin et al., 2019) (**Figure 3.8c**). Of these genes, *PTRHD1* expression was reduced in PPMI PD patients (log_2FC = -0.13, *P*-value = 5.1x10-4, Wald test), yet did not reach significance in ICICLE-PD (log_2FC = -0.14, *P*-value = 0.09, Wald test) (**Figure 3.8d**).

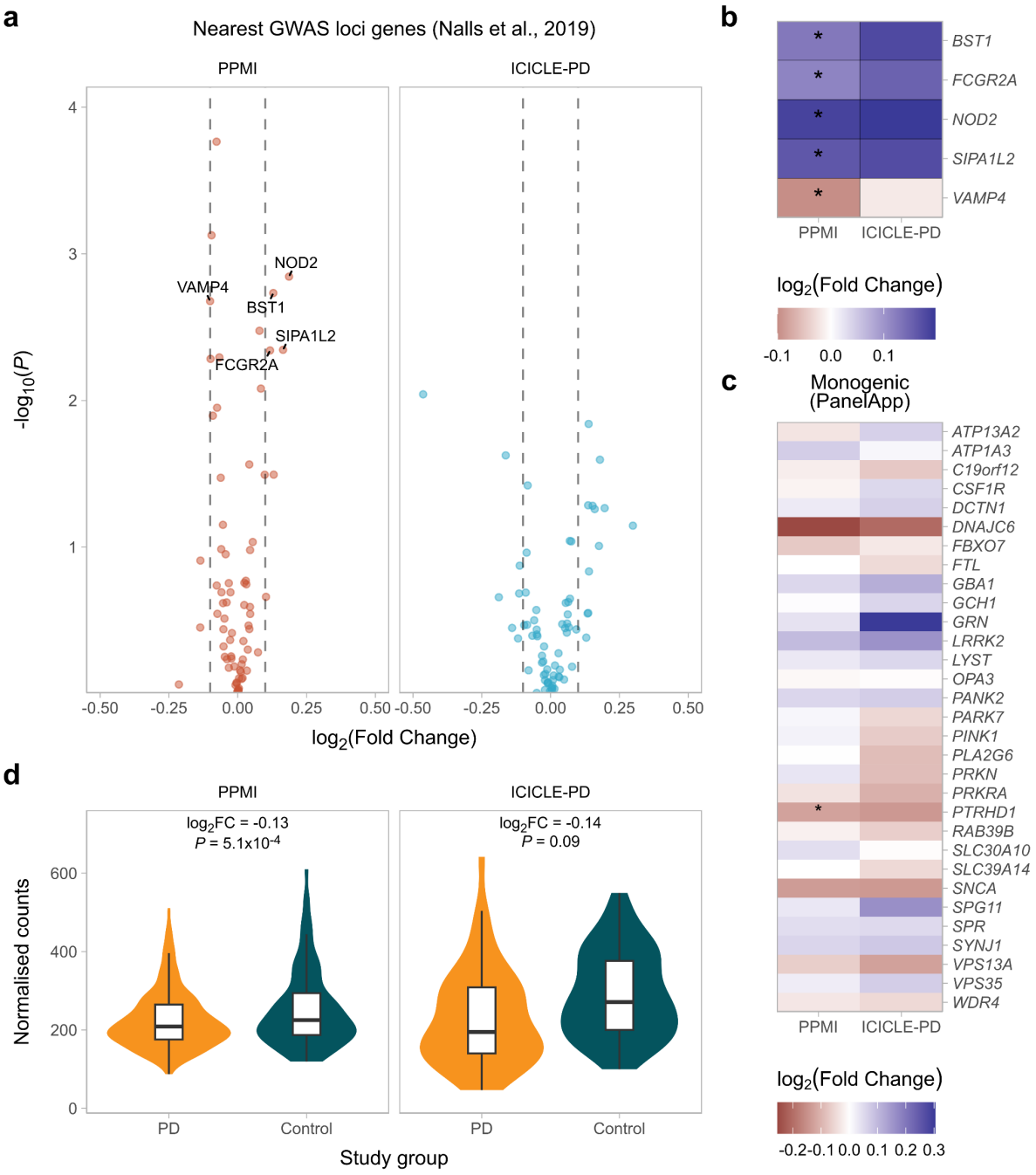


Figure 3.8. Differential expression of genes underlying the genetic architecture of PD. Volcano plot showing the differential expression of genes identified as the nearest genes to PD GWAS hit loci (Nalls et al., 2019) (70 tested in PPMI, 69 in ICICLE-PD) between idiopathic PD patients and controls in PPMI and ICICLE-PD cohorts. Labelled genes were differentially expressed (log2FC >0.1/<-0.1, FDR < 0.05). **(b)** Expression of five differentially expressed genes (out of 30 tested in PPMI and ICICLE-PD)located near PD risk loci (identified in **a**) in PPMI and ICICLE-PD cohorts. * = log2FC >0.1/-0.1, FDR < 0.05. **(c)** Comparison of expression for genes listed in Genomics England *Parkinson Disease and Complex Parkinsonism* gene panel between idiopathic PD patients and controls. * = log2FC >0.1/<-0.1, FDR <0.05. **(d)** Expression of *PTRHD1* in idiopathic PD patients and controls. *P*-value derived from a Wald test. $Log_2FC = log_2$ fold change.

3.3.7 Evaluating gene expression as a diagnostic biomarker of early-stage idiopathic PD

To evaluate the potential of gene expression to classify PD status, I used regularised logistic regression classification (**Section 2.9**). Gene expression had an AUC of 0.84 (95% CI: 0.81-0.88) when distinguishing early-stage idiopathic PD patients from controls in PPMI samples, using separate samples to train models (**Figure 3.9**). After fitting a model on all PPMI samples, gene expression had an AUC of 0.59 (95% CI: 0.48-0.71) when distinguishing early-stage idiopathic PD patients from controls in ICICLE-PD (**Figure 3.9**).

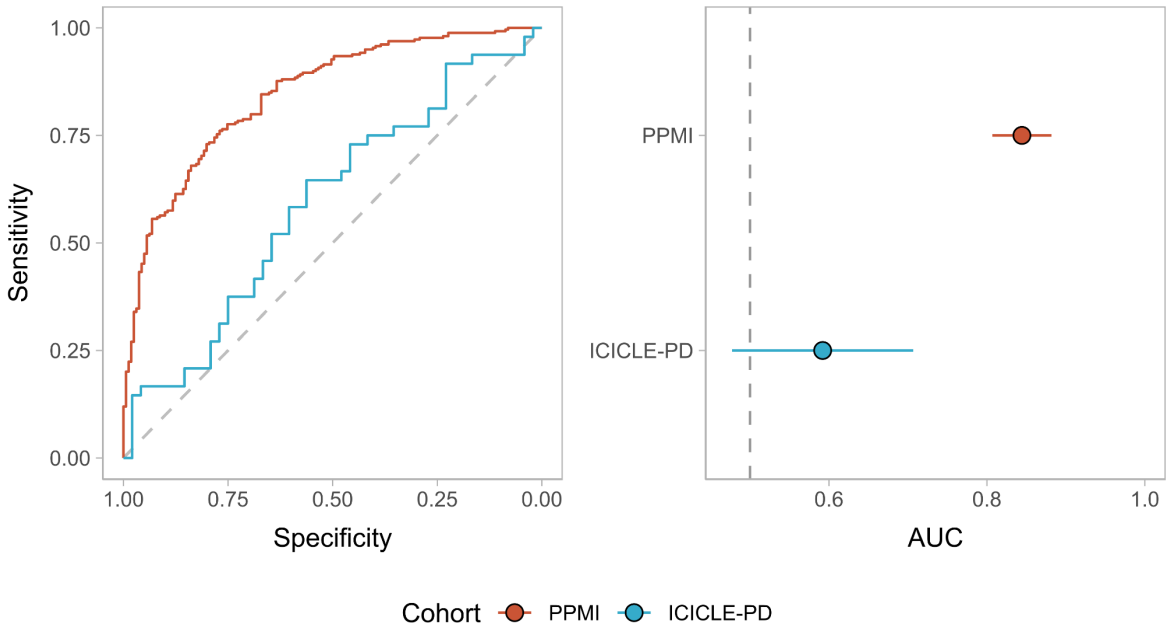


Figure 3.9. Classifying PD status based on gene expression. Performances of the gene PD classification model in PPMI and ICICLE-PD cohorts. The left panels display the ROC curve. The right panels show the area under the ROC curve (AUC). Error bars indicate the 95% CI of the AUC.

3.4 Discussion

Peripheral changes in gene expression in PD patients may serve as easily accessible and detectable biomarkers to aid in the diagnosis of and measure the progression of PD. Previous investigations into the PD blood transcriptome using RNA-seq data from the PPMI have revealed peripheral disease-related changes (Craig et al., 2021; Irmady et al., 2023; Li et al., 2023). However, these studies examined gene expression across different genetic backgrounds (harbouring pathogenic variants in *SNCA*, *LRRK2*, and *GBA*) and different disease durations. To identify changes in gene expression in the early stages of disease, I examined gene expression in the blood of recently diagnosed idiopathic PD patients in both the PPMI and ICICLE-PD cohorts. I identified increased expression of the genes *LMNB1* and *TMEM252* in early-stage idiopathic PD patients compared to controls. Commonly dysregulated processes related to ribosomal function and activation of an innate immune response. The ability to distinguish early-stage idiopathic PD patients from controls in PPMI participants using gene expression was consistent with previous work using the same cohort (Makarious et al., 2022; Pantaleo et al., 2022; Li et al., 2023), yet showed reduced ability in ICICLE-PD participants.

3.4.1 Gene expression changes in early-stage idiopathic PD patients

Comparing gene expression between early-stage idiopathic PD patients and controls in PPMI and ICICLE-PD, I found increased expression of both *LMNB1* and *TMEM252* in PD patients from both cohorts (**Figure 3.5).** *LMNB1* encodes Lamin B1, a component of the nuclear lamina (Dechat et al., 2008). *LMNB1* duplications lead to the progressive neurological condition autosomal dominant adult-onset demyelinating leukodystrophy (Padiath et al., 2006; Schuster et al., 2011). Morphological changes in the nuclear lamina have been reported in neurodegenerative disorders (Chou et al., 2018; Alcalá-Vida et al., 2021). In PD, alterations in the nuclear lamina have been attributed to dysfunction of the PD-risk protein LRRK2 and α-synuclein inclusions (Shani et al., 2019; Mansuri et al., 2024). Increased *LMNB1* expression has been shown in dopaminergic neurons isolated from the substantia nigra of idiopathic PD patients (Simunovic et al., 2009). A single nucleotide polymorphism mapped to *LMNB1* has been associated with cognitive decline in PD patients (Chung et al., 2012), yet I did not detect any association between *LMNB1* expression and cognitive impairment in PPMI or ICICLE-PD participants (**Table 3.4**).

TMEM252 encodes for a transmembrane protein of unknown function. *TMEM252* expression is enriched in the gastrointestinal tract, kidney, and urinary bladder (Uhlén et al., 2015); areas of the body which frequently present with dysfunction in PD patients (Sakakibara et al., 2012;

Warnecke et al., 2022). *TMEM252* expression was altered and provided prognostic information in a study investigating triple-negative breast cancer (Lv et al., 2019), yet has no link to PD or neurodegeneration at present.

By analysing recently diagnosed idiopathic PD patients (blood samples collected within 13 months of diagnosis), I was able to identify gene expression changes occurring early in disease progression. Previous investigations using PPMI RNA-seq data identified >1,000 differentially expressed genes in PD patients (Craig et al., 2021; Riboldi et al., 2022; Irmady et al., 2023; Li et al., 2023). In comparison, I identified fewer differentially expressed genes (44 in PPMI), indicative of modest transcriptomic changes in the blood shortly after idiopathic PD diagnosis following motor symptoms onset. Notably, previous studies collected PD patients across different subtypes, including those harbouring highly penetrant PD risk variants (Craig et al., 2021; Riboldi et al., 2022; Li et al., 2023; Irmady et al., 2023), which may exhibit transcriptional differences. In some comparisons, RNA-seq data from later time points were used (Craig et al., 2021; Irmady et al., 2023) and thus may reflect expression changes related to dopaminergic neuron loss in the years following motor symptom onset (Nandhagopal et al., 2009; Kordower et al., 2013).

I replicated the differential expression of genes (*LSMEM1*, *TPST1* and *SLED1*, **Figure 3.7**) previously reported in blood RNA-seq studies of PD patients (Infante et al., 2016; Garofalo et al., 2020; Henderson et al., 2021; Kurvits et al., 2021). As previous studies were not limited to early-stage PD, these genes may reflect those with altered expression throughout PD progression. Of the three genes replicated in PPMI samples, *SLED1* and *LSMEM1* encode proteins of unknown function. *TPST1* encodes tyrosylprotein sulfotransferase 1, an enzyme required for post-translational modification of proteins that plays important roles in the inflammatory process, leukocyte movement and cytosis, viral cell entrance, and other cell-cell and protein-protein interactions (Stone et al., 2009). In addition, my data provide a potential mechanistic link between PD-associated genetic variation, identifying differential expression among genes associated with PD risk. These included *PTRHD1* in PPMI PD patients (**Figure 3.8c, d**), which encodes a peptidyl-tRNA hydrolase that has been associated with recessive parkinsonism (Khodadadi et al., 2017; Al-Kasbi et al., 2021), and genes identified by through large-scale association testing (*BST1*, *FCGR2A*, *SIPA1L2*, *NOD2* and *VAMP4*, **Figure 3.8a**) (Nalls et al., 2019).

One limitation of this work was that gene expression was measured in bulk blood tissue. Whilst this has the advantage of reducing preprocessing steps, it is important to consider that different cell types contribute to measured gene expression. Given the reported changes in blood cell

proportions in PD patients (Craig et al., 2021; Muñoz-Delgado et al., 2021, 2023), differences in gene expression may be partly due to changes in blood cell composition. Studies investigating gene expression at a single cell resolution have typically focused on individual blood cell populations (Navarro et al., 2021; P. Wang et al., 2021; Riboldi et al., 2022; P. Wang et al., 2022; Shi Yan et al., 2023). As such, cell type specific changes in gene expression in the context of all blood cell populations have not yet been investigated.

Furthermore, a notable difference between PD patients enrolled in PPMI and ICICLE-PD studies was treatment status, in which PPMI participants were enrolled untreated with dopaminergic PD treatment while most ICICLE-PD cases were receiving dopaminergic PD treatment (**Section 2.2**). PPMI differentially expressed genes were not associated with treatment dosage in ICICLE-PD PD patients (**Table 3.3**). This finding is consistent with other work failing to identify broad changes in gene expression associated with PD treatment dosage (Navarro et al., 2021; Irmady et al., 2023). However, due to the small number of untreated ICICLE-PD cases (4/48), I did not assess how the commencement of treatment may affect gene expression, as previously described with cell-free mitochondrial DNA levels (Lowes et al., 2020).

3.4.2 Coordinated changes in gene expression reveal dysregulated processes related to protein translation and immune activation in early-stage idiopathic PD patients

Based on coordinated changes in gene expression, I identified the downregulation of genes associated with protein synthesis in the blood of early-stage idiopathic PD patients in both PPMI and ICICLE-PD cohorts (**Figure 3.6**). Expression of genes encoding ribosomal proteins is reduced in the substantia nigra of PD patients (Garcia-Esparcia et al., 2015), which may be secondary to neurodegeneration in this region. Protein synthesis is also reduced in idiopathic PD-patient derived fibroblasts (Deshpande et al., 2020; Flinkman et al., 2023), providing further evidence of detectable peripheral changes in translation in idiopathic PD.

Protein synthesis regulation is a common cellular mechanism activated during innate immune responses (Argüello et al., 2015). PD is now viewed as a multi-system disorder, potentially mediated by immune dysfunction (reviewed in Tan et al., 2020; Tansey et al., 2022). Consistent with the potential role of immune dysfunction in the susceptibility to or progression of PD, I identified an upregulation of genes associated with the innate immune response in the blood of idiopathic PD patients across PPMI and ICICLE-PD cohorts (**Figure 3.6**). In the periphery, increased levels of proinflammatory cytokines have also been reported in PD patients (Qin et al., 2016), including those enrolled in ICICLE-PD (Williams-Gray et al., 2016). Specifically, I

identified the upregulation of genes related to antiviral responses in idiopathic PD patients (**Figure 3.6**). The link between viral exposure and PD has been extensively hypothesised (reviewed in Olsen et al., 2018; Leta et al., 2022). Analyses of population-level cohorts have described an association between influenza infection and the development of PD (Cocoros et al., 2021; Levine et al., 2023). Increased α-synuclein levels have been shown in both central (Beatman et al., 2016; Monogue et al., 2022) and peripheral neurons (Stolzenberg et al., 2017) during viral infections, proposing a role for α-synuclein in the innate antiviral response which may be affected when pathological aggregations of α-synuclein are present. Genes related to type I interferon signalling, a key component of the cellular antiviral defence (McNab et al., 2015), were also upregulated (**Table 8.1**). Interestingly, three genes that have previously been reported as differentially expressed in PD replicated in ICICLE-PD (*IFIT1*, *RSAD2*, *IFI44l*, **Figure 3.7**) are induced by type I interferons (Kyogoku et al., 2013). Studies investigating the role of type I interferon signalling in PD have produced conflicting findings. Functional α-synuclein may be required for proper type I interferon signalling in response to viral infection of neurons (Monogue et al., 2022). Mice lacking interferon β, a type I interferon, or the type I interferon receptor (interferon-α/β receptor) develop neuronal Lewy body pathology and PD-like phenotypes (Ejlerskov et al., 2015; Villanueva et al., 2021). In mice and cellular PD models, however, the ablation of the interferon-α/β receptor was neuroprotective (Main et al., 2016, 2017). This work suggests that type I interferon signalling is likely increased in the blood of patients with idiopathic PD. However, it is unclear whether this is linked to the pathogenesis or the systemic response to PD.

3.4.3 Inconsistent performance of gene expression to classify early-stage idiopathic PD patients across PPMI and ICICLE-PD cohorts

Gene expression differentiated between early-stage idiopathic PD patients from controls in the PPMI cohort (AUC = 0.84, **Figure 3.9**). The ability of gene expression to identify PD patients was similar to previous work also used PPMI RNA-seq data with AUCs of 0.80 (Makarious et al., 2022), 0.81 (Li et al., 2023) and 0.72 (Pantaleo et al., 2022). Of these estimates, Li et al report a reduced ability of gene expression to classify PD (AUC = 0.69) when assessed by an independent group of samples (Parkinson's Disease Biomarker Program, PDBP). Likewise, gene expression had a reducibility ability when classifying early-stage idiopathic PD patients from the ICICLE-PD cohort (AUC = 0.59, **Figure 3.9**). Notably, PPMI and PDBP blood samples underwent library preparation and sequencing at a centralised location as part of the Accelerating Medicine Partnership Parkinson's disease program (AMP-PD, https://amp-

pd.org/transcriptomics-data, accessed 28/05/2024), minimising the impact of technical differences between these cohorts.

The lack of generalisability of gene expression to predict PD status across data sources limits its potential utility as a diagnostic biomarker, especially considering its worse performance compared to α-synuclein seed amplification and DOPA decarboxylase levels in multiple cohorts (Grossauer et al., 2023; Pereira et al., 2023; Siderowf et al., 2023; Rutledge et al., 2024).

Alternative RNA species may provide improved classification performance compared to gene expression. Circular RNA (circRNA) is a covalently closed RNA species that has increased resistance to exonuclease activity resulting in longer half-lives compared to the cognate linear RNA (Jeck et al., 2013; Enuka et al., 2016). As such, the ability of circRNAs to act as disease biomarkers has been extensively discussed (reviewed in Verduci et al., 2021). By assessing the ability of gene expression to classify early-stage idiopathic PD status in PPMI and ICICLE-PD datasets, I have created a reference for evaluating the performance of future predictors. This will be explored further in **Chapter 4**, primarily focusing on circular RNA expression.

3.5 Conclusion

To summarise, the work presented in this chapter analysed gene expression in idiopathic PD. By measuring gene expression in recently diagnosed idiopathic PD patients, I aimed to identify changes in gene expression that reflected the earliest stages of disease progression. Overall, I identified subtle changes in individual gene expression, with two genes *LMNB1* and *TMEM252*, showing increased expression in idiopathic PD patients in both PPMI and ICICLE-PD cohorts. Measuring systematic changes in gene expression identified changes in genes associated with ribosomal function and activation of an innate immune response across both cohorts, expanding knowledge about potentially dysregulated processes in the periphery of PD patients. The observed classification performance of gene expression is consistent with published literature and a reference to assess the performance of other predictors in this thesis.

Chapter 4. Circular RNA expression in early-stage idiopathic Parkinson's disease

4.1 Background

Initial work used cellular and animal PD models to investigate differences in circRNA expression. *Circzip-2* was downregulated in a PD *Caenorhabditis elegans* model (Kumar et al., 2018). *circSNCA* and *circDLGAP4* were downregulated in PD cellular models, with *circDLGAP4* also downregulated in a PD mouse model (Sang et al., 2018; Feng et al., 2020). RNA-seq of the cerebral cortex, hippocampus, striatum and cerebellum tissue in a PD mouse model identified 24, 66, 71 and 121 differentially expressed circRNAs respectively (Jia et al., 2020). Human substantia nigra tissue showed lower circRNA abundance compared to controls, contrasting with the increases shown in amygdala and temporal gyrus tissue (Hanan et al., 2020). Hanan et al also reported 24 differentially expressed circRNAs in substantia tissue from PD patients.

Several studies have now investigated circRNA expression in the blood of PD patients. Quantitative PCR (qPCR) measurement of 87 pre-selected circRNAs in peripheral blood mononuclear cells (PBMCs), identified six circRNAs from *SLAIN1*, *DOP1B*, *REPS1*, *MAPK9*, *PSEN1* and *HOMER1* loci to be reduced in PD patients ($n = 60$) compared to controls ($n = 60$) (Ravanidis et al., 2021). Four circRNAs (produced from *SLAIN1*, *SLAIN2*, *ANKRD12* and *PSEN1*) could distinguish PD patients from controls with an AUC of 0.84 (Ravanidis et al., 2021). However, the approach of Ravanidis et al is limited to the selected circRNAs and the ability to distinguish PD from controls was not evaluated in an additional set of samples. RNA-seq of blood samples from PD patients ($n = 4$) and controls ($n = 4$) identified 411 differentially expressed circRNAs (Kong et al., 2021). Whilst the use of RNA-seq to measure circRNA expression allows the unbiased detection of circRNAs, it is not clear how Kong et al detected circRNAs from RNA-seq data and the ability of circRNA expression to distinguish between PD patients and controls was not assessed. Using microarray measurements of circRNA expression, six circRNAs were increased in PD patients ($n = 20$) compared to controls ($n = 20$) (Zhong et al., 2021). Four of these circRNAs could distinguish between PD patients and controls in this initial cohort with an AUC of 0.98, with an AUC of 0.85 in an independent cohort (early-stage PD = 80, controls = 80) (Zhong et al., 2021). Another study also used microarray technology to measure circRNA expression, identifying 139 differentially expressed circRNAs in three PD patients compared to three controls (Xiao et al., 2022). 10 of these circRNAs were selected for replication, of which four were validated and could distinguish PD patients ($n = 63$) from controls ($n = 60$) with an AUC of 0.94 (Xiao et al., 2022). For the studies that used circRNA microarrays (Zhong et al., 2021; Xiao et al., 2022), without manifest files, it is unclear which BSJs the probes correspond to limiting the ability to replicate

these findings. Overall, these results indicate that blood circRNA expression may have a clinical utility to distinguish PD patients from controls. As these studies typically use small cohorts and highlight specific circRNAs that are not replicated, a large-scale, unbiased assessment of circRNAs in PD is warranted.

Circular RNAs (circRNAs) can be identified and defined through the presence of a back-spliced junction (BSJ), whereas linear RNAs possess canonical, or forward-spliced junctions (FSJ) (**Figure 4.1**). Here, I utilised total RNA sequencing to quantify linear and circRNA expression in the blood across two large unrelated cohorts, The Michael J Fox Foundation Parkinson's Progression Markers Initiative (PPMI) (Marek et al., 2011, 2018) and The Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation-PD (ICICLE-PD) (Yarnall et al., 2014). I selected early-stage (diagnosed <13 months) idiopathic Parkinson's disease (PD) patients comparing BSJ and FSJ abundance to matched controls using PPMI as a discovery cohort and ICICLE-PD as a replication cohort (**Figure 4.1**).

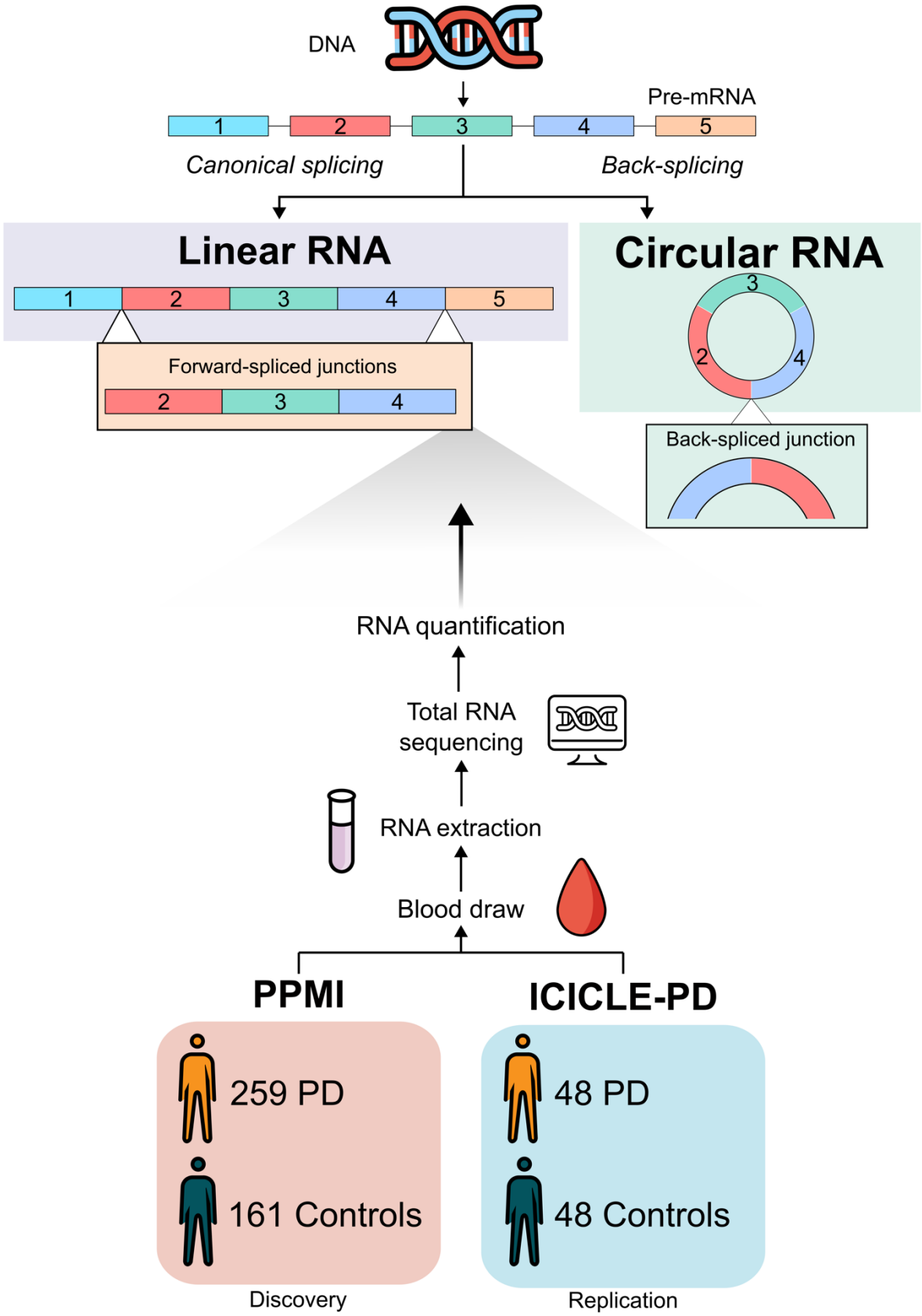


Figure 4.1. Circular RNA formation and study overview. The canonical splicing of pre-mRNA produces a linear RNA molecule containing forward-spliced junctions (FSJ). Circular RNA molecules are formed by a back-splicing reaction and can be identified through the presence of a back-spliced junction (BSJ). Individuals from our discovery (PPMI, PD = 259, Controls = 161) and replication (ICICLE-PD, PD = 48, Controls = 48) cohorts underwent whole blood total RNA sequencing to detect and quantify linear and circular RNAs.

4.2 Methods

4.2.1 Cohorts

RNA sequencing (RNAseq) data presented here is the same individuals as **Chapter 3**. Individuals were selected for sequencing based on the criteria outlined in **Section 2.2**. Briefly, patients with early-stage (diagnosed <13 months) idiopathic Parkinson's disease (PD) and control individuals were obtained from the PPMI and ICICLE-PD cohorts. Clinical information regarding the samples was taken from each cohort's respective metadata file.

4.2.2 RNA sequencing

Sample collection, RNA extraction, library preparation and sequencing of PPMI and ICICLE-PD samples have been outlined in **Section 2.3**. Quality control steps of the sequencing data were described in **Section 2.4**.

4.2.3 Circular RNA detection and quantification

Circular RNAs were detected and quantified as described in **Section 2.6.1**.

4.2.4 Assessing circular RNA normalisation methods

Raw BSJ counts must be normalised before analyses to account for differing sequencing depths between samples. However, there is currently no consensus on how to normalise circRNA counts. To identify the optimal method for circRNA count normalisation in the PPMI and ICICLE-PD datasets, I first collated the number of total reads, genome and transcriptome mapped reads in each sample using FastQC v0.11.9 and MultiQC v1.9. The number of genome mapped reads was reported by CIRIquant based on alignment to the Ensembl GRCh38 "genome_tran" index (https://registry.opendata.aws/jhu-indexes/, downloaded 13/10/2020) using HISAT2 v2.2.0 (D. Kim et al., 2019). The amount of transcriptome-mapped reads was obtained from log files produced by Salmon v1.3.0 (Patro et al., 2017) using MultiQC v1.9. I correlated the number of reads mapping to each BSJ with the number of total sequenced reads, total number of genome-mapped reads or total number of transcriptome-mapped reads. I also correlated these factors to the number of unique circRNAs detected in each sample.

4.2.5 Identifying sources of extraneous variation in circular RNA quantification

The contribution of various biological and technical sequencing measures on circRNA expression was measured using the methods outlined in **Section 2.7**. Based on the information

outlined in **Section 4.3.3**, I included the percentage of intronic bases (PPMI and ICICLE-PD) and the median coefficient of variation of coverage (ICICLE-PD) as explanatory variables in linear mixed models predicting individual circRNA expression. I also included sequencing batch, age, and sex as explanatory variables in the models. Categorical variables (study group, sequencing batch, and sex) were modelled as random effects as recommended (Hoffman & Schadt, 2016).

4.2.6 Circular RNA annotation

CircRNAs can be annotated by resolving their BSJ location to the genomic reference and identifying genetic features. In this instance, CIRIquant reports junction annotations relative to the Ensembl GRCh38 reference (release 101). Annotating circRNAs in this fashion may be ambiguous, for example when overlapping genes are present. Ambiguous annotation occurred at a similar frequency in both cohorts (PPMI = 11.0%, ICICLE-PD 11.1%). While this does not impact circRNA quantification, these circRNAs have been omitted from summary statistics describing circRNA host-gene annotation.

Circular RNAs recorded in CircAtlas v3.0 (Wu et al., 2024) and CIRCpedia v2.0 (R. Dong et al., 2018) were downloaded from their respective websites (accessed 09/01/2024). BSJ coordinates from CircAtlas were converted to 0-based indexing, in line with the BED format specification and to ensure consistency when describing coordinates throughout this chapter. BSJ coordinates were reported as chr:start-end:strand with regards to the GRCh38 reference (e.g. 16:85633913-85634132:+).

4.2.7 Gene set overrepresentation

Gene set enrichment analysis (GSEA) is not suitable for assessing functional enrichment of circRNA host genes as each gene can host multiple circRNAs and GSEA requires unique genes. As such, functional circRNA host gene enrichment was performed using an over-representation test as implemented in clusterProfiler v4.6.2 (Wu et al., 2021). CircRNA host genes were grouped into Gene Ontology (GO) terms and enrichment was calculated using the *enrichGO* function. To detect GO terms enriched in abundant BSJs, those included in differential expression testing (and thus defined as abundant) were compared against all high confidence BSJs detected in that cohort.

4.2.8 Modelling the diversity of the circular RNA transcriptome

The number of unique circRNAs (based on distinct BSJs) and circRNA-host genes were calculated for each sample. This value was then normalised by the number of reads that mapped to the transcriptome in that sample to account for differences in sequencing depth. A linear model was fit predicting the number of unique circRNAs or circRNA-host genes based on age at sample collection, reported sex and sequencing batch. The marginal mean number of unique circRNAs and circRNA host-genes for each study group were identified using the *modelbased* v0.8.6 package.

4.2.9 Differential expression

Abundant circRNAs with raw counts over 10 in more than *n* samples (where *n* is the size of the smallest sample group in each cohort) were retained. Differential expression of circRNAs between PD patients and controls was tested as described in **Section 2.8**.

I used the method described in **Section 2.10.2** to determine whether circRNA expression exhibits a global imbalance based on the calculated fold changes between PD patients and controls obtained from differential expression testing.

4.2.10 Correlating circular RNA expression to relevant measures

PD-related treatment dosage was assessed using the levodopa equivalent daily dose (LEDD) in ICICLE-PD cases as reported in the ICICLE-PD metadata. I fit linear models to predict normalised circRNA expression (log_2 transformed BSJ counts per million transcriptome mapped reads) by LEDD adjusting for sources of biological and technical circRNA expression variation (**Section 3.2.5**). Coefficient *P*-values for each circRNA were adjusted for multiple testing using the Benjamini-Hochberg method.

Relevant measures associated with circRNA expression in both PPMI and ICICLE-PD cases included the age at diagnosis (years), motor symptom severity (The Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale Part III, MDS-UPDRS III) and cognitive impairment (Montreal Cognitive Assessment, MoCA). CircRNAs identified as significantly differentially expressed in PPMI PD patients were used for all associations. I fit linear models to test the association of each clinical measure with circRNA expression (log_2 transformed BSJ counts per million transcriptome mapped reads) adjusting for sources of biological and technical circRNA expression variation (**Section 3.2.5**). Coefficient *P*-values for each circRNA were adjusted for multiple tests using the Bonferroni correction.

4.2.11 Classification of PD status using RNA expression

The BSJ and FSJ loci of all BSJs detected both PPMI and ICICLE-PD cohorts and included in differential expression analysis were used as potential predictors of PD status. BSJ and FSJ expression were transformed to stabilise their variance (*varianceStabilizingTransformation* in DESeq2 v1.40.2) (Anders & Huber, 2010). BSJ:FSJ expression ratios were obtained from the output of CIRIquant.

The classification potential of individual junctions was assessed initially based on the expression of BSJs, FSJs and the ratio between these values (BSJ:FSJ ratio). Receiver operator characteristic (ROC) curves were generated for each junction using pROC v1.18.4 (Robin et al., 2011). The classification potential of each junction was summarised based on the area under the ROC curve (AUC) with 95% confidence intervals reported using Delong's method (DeLong et al., 1988) as implemented in pROC.

Regularised logistic regression was used as described (**Section 2.9)**. As predictors, I included BSJ expression (BSJ), FSJ expression (FSJ), and the ratio between BSJ and FSJ expression (BSJ:FSJ ratio). I also extended the gene expression model created in **Chapter 3** (Gene) with BSJ counts to create a combined model (Gene and BSJ). As with the individual junctions, I created ROC curves and calculated the AUC based on the predictions returned by each respective model.

4.3 Results

4.3.1 Detection of blood circular RNAs in PPMI and ICICLE-PD participants

PPMI samples had a significantly higher percentage of reads containing adapter sequences (median = 36.3%, IQR = 3.7%), compared to ICICLE-PD samples (median = 20.9%, IQR = 3.8%) (*P*-value < $2.2x10^{-16}$, Wilcoxon rank-sum test, **Figure 4.2a**). Furthermore, PPMI samples had a significantly higher percentage of reads that were trimmed due to low quality (median = 2.07%, IQR = 0.88%) compared to ICICLE-PD samples (median = 0.56%, IQR = 0.76%) (*P*-value < $2.2x10^{-16}$, Wilcoxon rank-sum test, **Figure 4.2b**), reflecting lower general sequencing quality.

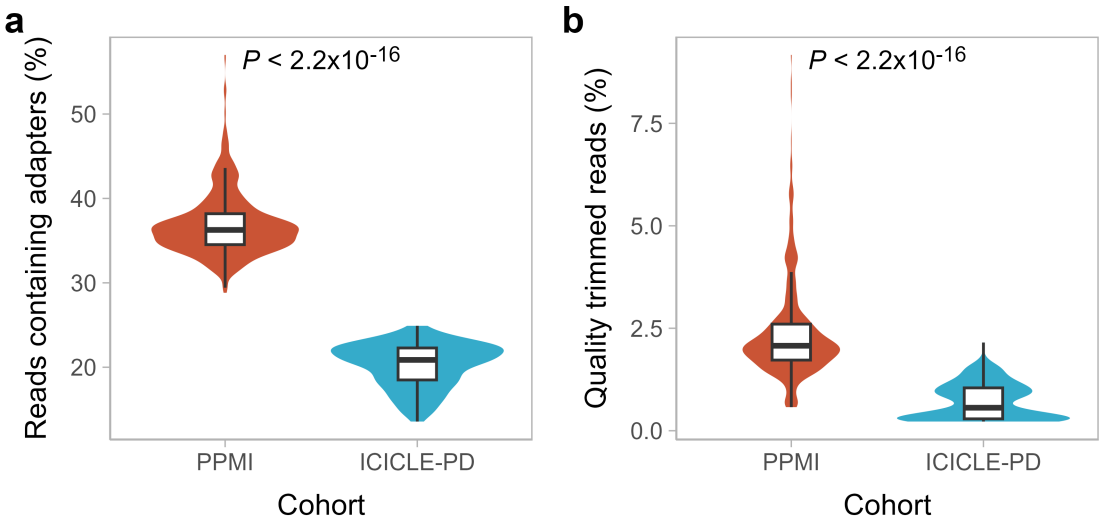


Figure 4.2. Quality and adapter trimming of PPMI and ICICLE-PD samples. (a) The percentage of reads that contained adapter sequences. **(b)** The percentage of reads that required trimming due to low quality sequences. In both panels, results are shown for each paired-end FASTQ file (R1 and R2). *P*-values produced by a Wilcoxon rank-sum test.

CircRNAs were detected by the presence of a BSJ as described in **Section 2.6.1**. The number of BSJ detections and subsequent overlaps reported by each tool are shown in **Figure 4.3**. Overall, a total of 438,189 and 222,133 unique BSJs were detected in PPMI and ICICLE-PD datasets respectively (**Figure 4.3a, b**). Notably, most detected BSJs were uniquely identified by CIRCexplorer2 (PPMI = 63.5%, ICICLE-PD = 57.9%), which may reflect false positive detections. 2.0-2.6% of the BSJs detected in PPMI and ICICLE-PD cohorts respectively were detected by all three tools. When considering just BSJs detected in both cohorts, similar

proportions of BSJ overlaps were present (**Figure 4.3c**). For example, 3.2% of common BSJs were detected by all three tools, indicating consistent inter-cohort detection (**Figure 4.3c**).

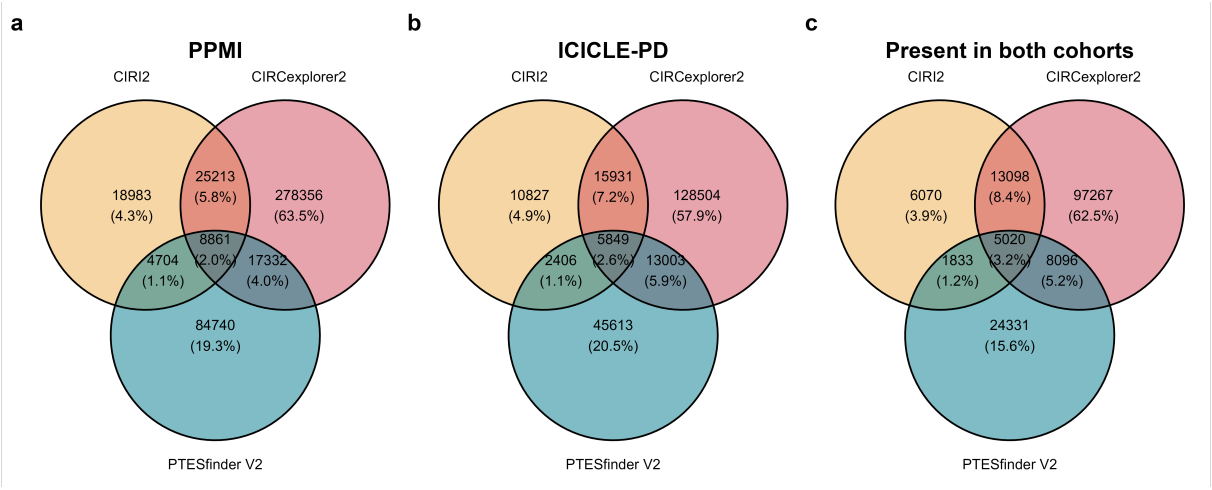


Figure 4.3. Circular RNA detection in PPMI and ICICLE-PD samples. (a, b) Venn diagrams showing the number of circRNAs (unique back-spliced junctions) detected by CIRI2, CIRCexplorer2 and PTESfinder V2 in PPMI (**a**) and ICICLE-PD samples (**b**). **(c)** Number of circRNAs that were present in both cohorts. For each diagram, the contribution of each segment is given as the percentage of the total circRNAs detected in each cohort and is shown in brackets.

'High confidence circRNAs' were defined by overlapping BSJ loci reported by each tool (**Section 2.6.1**). This filtering step reduced the initial detection sets (PPMI = 23,454 from 438,198 or 5.35% and ICICLE-PD = 15,345 from 222,133 or 6.91%). When restricted to BSJs detected by all three tools, circRNA counts reported by CIRI2 and CIRCexplorer2 were highly correlated in both cohorts (Pearson's $r > 0.9$) (**Figure 4.4**). Of the three tools, PTESfinder v2 showed the lowest correlation with the other tools in both cohorts (**Figure 4.4**). I subsequently quantified the majority of BSJ loci in each cohort (PPMI = 23,447/23,454, ICICLE-PD = 15,336/15,345) using CIRIquant, producing junction counts for both circRNA (spanning the BSJ) and linear RNA (spanning the FSJ) at each locus.

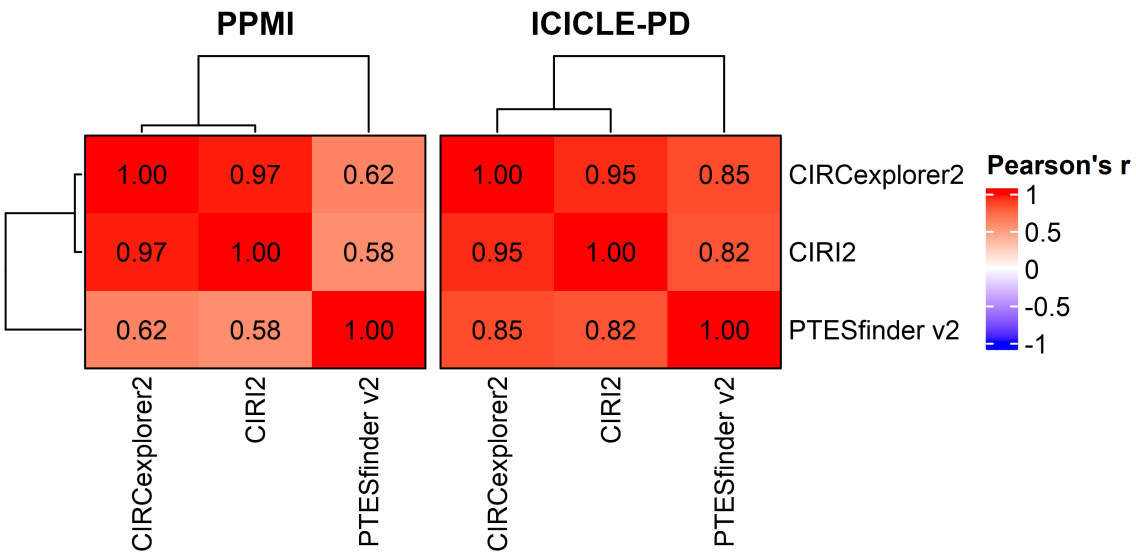


Figure 4.4. Pairwise correlation of raw circular RNA counts between detection methods. The raw counts of circRNAs detected by all three tools within each cohort were correlated in a pairwise manner. Correlations between the counts reported by each tool are reported as Pearson's *r* coefficient.

4.3.2 Identifying the optimum circRNA normalisation strategy

Several possible normalisation factors were assessed (**Section 4.2.4**). There were no significant differences in normalisation factors between PD and control samples in either cohort (*P*-values >0.05, Wilcoxon rank-sum test) (**Figure 4.5**).

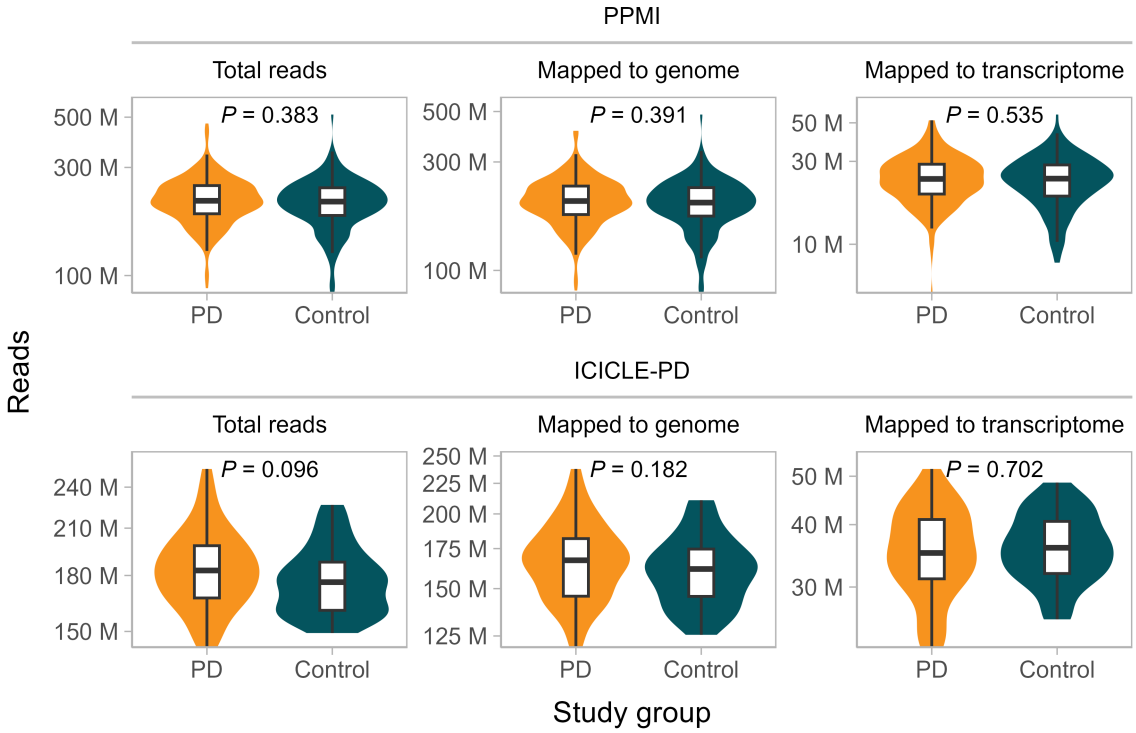


Figure 4.5. Comparison of potential circular RNA normalisation denominators between study groups. The number of reads (total, genome-mapped and transcriptome-mapped) in PD patients and controls. *P*-values from a Wilcoxon rank-sum test.

In PPMI, I observed a positive correlation between total circRNA expression (**Figure 4.6a**) and the number of unique circRNAs (**Figure 4.6b**) in each sample with each normalisation factor. In ICICLE-PD samples, the number of total and genome-mapped reads was not significantly correlated to total expression (**Figure 4.6a**, *P*-value >0.05, linear regression) or number of circRNAs (**Figure 4.6b**, *P*-value >0.05, linear regression). However, the number of transcriptome-mapped reads was positively correlated to both total expression **(Figure 4.6a**) and number of circRNAs (**Figure 4.6b**, *P*-value >0.05, linear regression). In PPMI, the number of transcriptome-mapped reads explained a higher proportion of the variation in the total expression (R^2 = 0.25) or number of circRNAs (R^2 = 0.26) (**Figure 4.6a, b**). As such, normalising the expression and number of circRNAs in each sample against the number of transcriptome-mapped reads was selected as the most appropriate strategy.

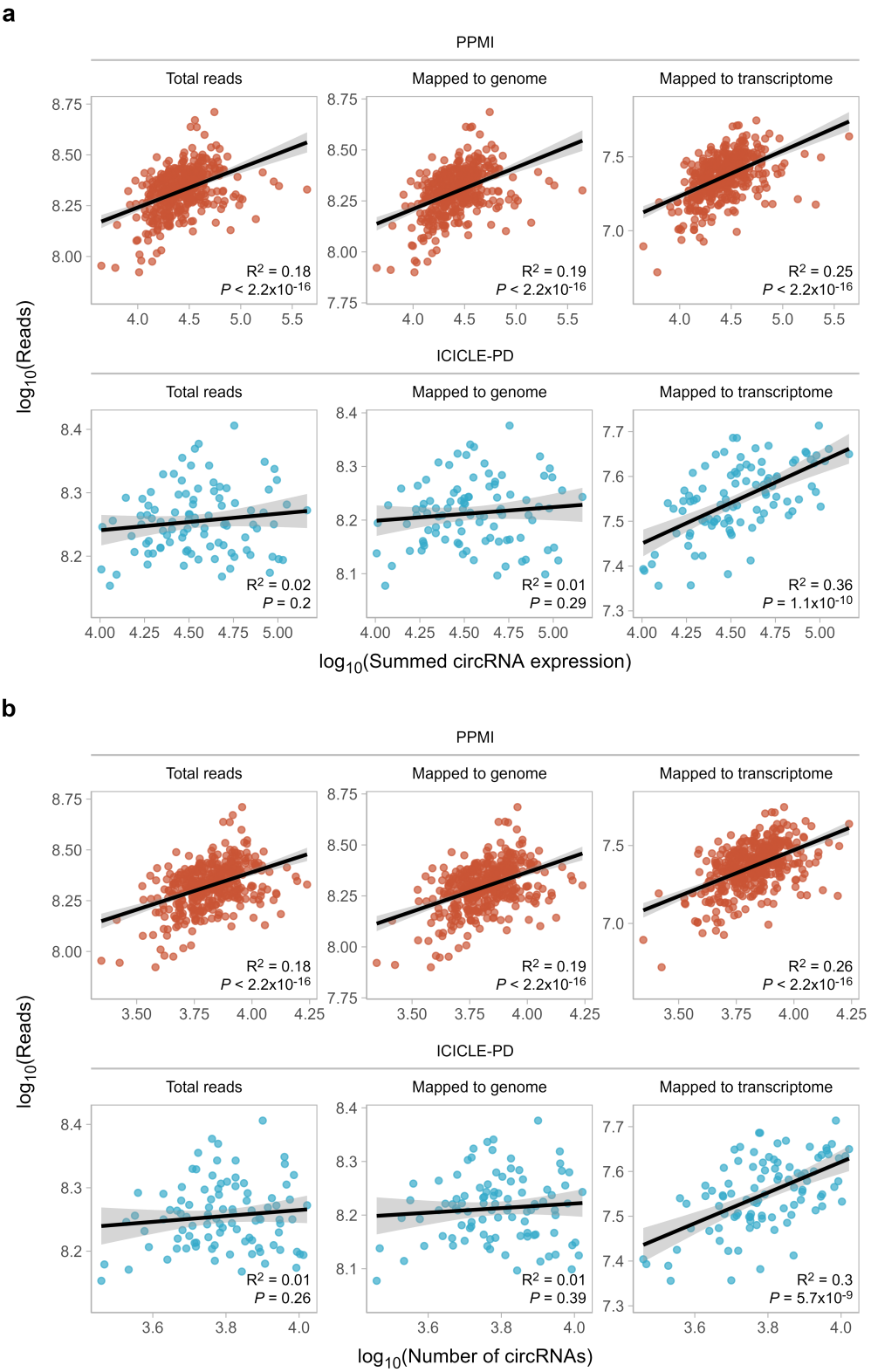


Figure 4.6. Identifying normalising against transcriptome mapped reads as the optimum circular RNA normalisation strategy. Scatter plots showing the number of total, genome-mapped and transcriptome-mapped reads in each sample compared to the number of total circRNA mapped reads (spanning the back-spliced junction) **(a)** or number of unique circRNAs (unique back-spliced junction loci) in each sample **(b)**. In each comparison, the regression line resulting from linear regression is shown with shaded regions indicating the 95% confidence interval. Also shown is the amount of variance explained by each circRNA normalisation strategy (R^2) and the *P*-value of the F-statistic from the linear model.

4.3.3 Identifying sources of extraneous variation in circular RNA quantification

Sources of biological and technical variation influence general RNA quantification (’t Hoen et al., 2013), yet the extent to which circRNA quantification may be affected is not well established. Previous work measuring the variance of circRNA expression suggests that circRNA expression is more robust to biological and technical variation than gene expression (Z. Liu et al., 2019). Projection of circRNA expression into the first and second principal components showed no clear segregation of individuals by study group (**Figure 4.7**). As such, other sources of variation contributing to the quantification of circRNA expression were measured (**Section 4.2.5**).

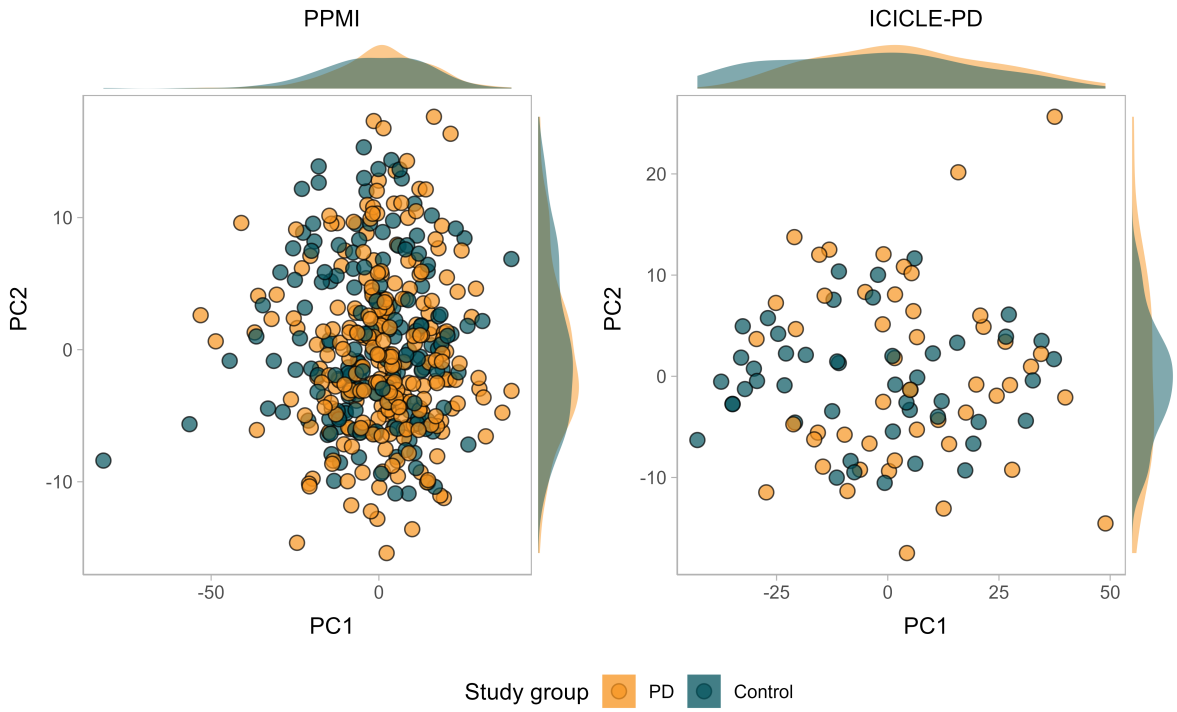


Figure 4.7. First two principal components of circular RNA expression in PPMI and ICICLE-PD. Scatter plot of the first and second principal components of circRNA expression in PPMI (left panel) and ICICLE-PD (right panel). The density of the points is shown on the edges of the plot. Points and densities are coloured based on which study group each individual belonged to (PD or Control).

Firstly, sources of technical variation were identified based on their associations with circRNA expression principal components (PCs). Cumulatively, the first 10 circRNA PCs explained 66.2% and 72.8% of the variance in PPMI and ICICLE-PD samples respectively. In PPMI, the percentage of intronic bases was the only technical metric identified as explaining a large proportion of the variance (R^2 >0.5) with any of the first ten PCs (**Figure 4.8a**). In ICICLE-PD,

the percentages of intronic and untranslated region (UTR) bases explained a large proportion of PC1 (R^2 >0.5). Furthermore, the median coefficient of variance of coverage explained a large proportion (R^2 >0.5) of PC3 in ICICLE-PD (**Figure 4.8b**). In PPMI, associations were observed between the median coefficient of variance of coverage and several PCs yet were below the R^2 threshold (R^2 <0.5) (**Figure 4.8a)**. Inter-cohort differences in technical variation were expected considering the different library preparation kits used (**Section 2.3**).

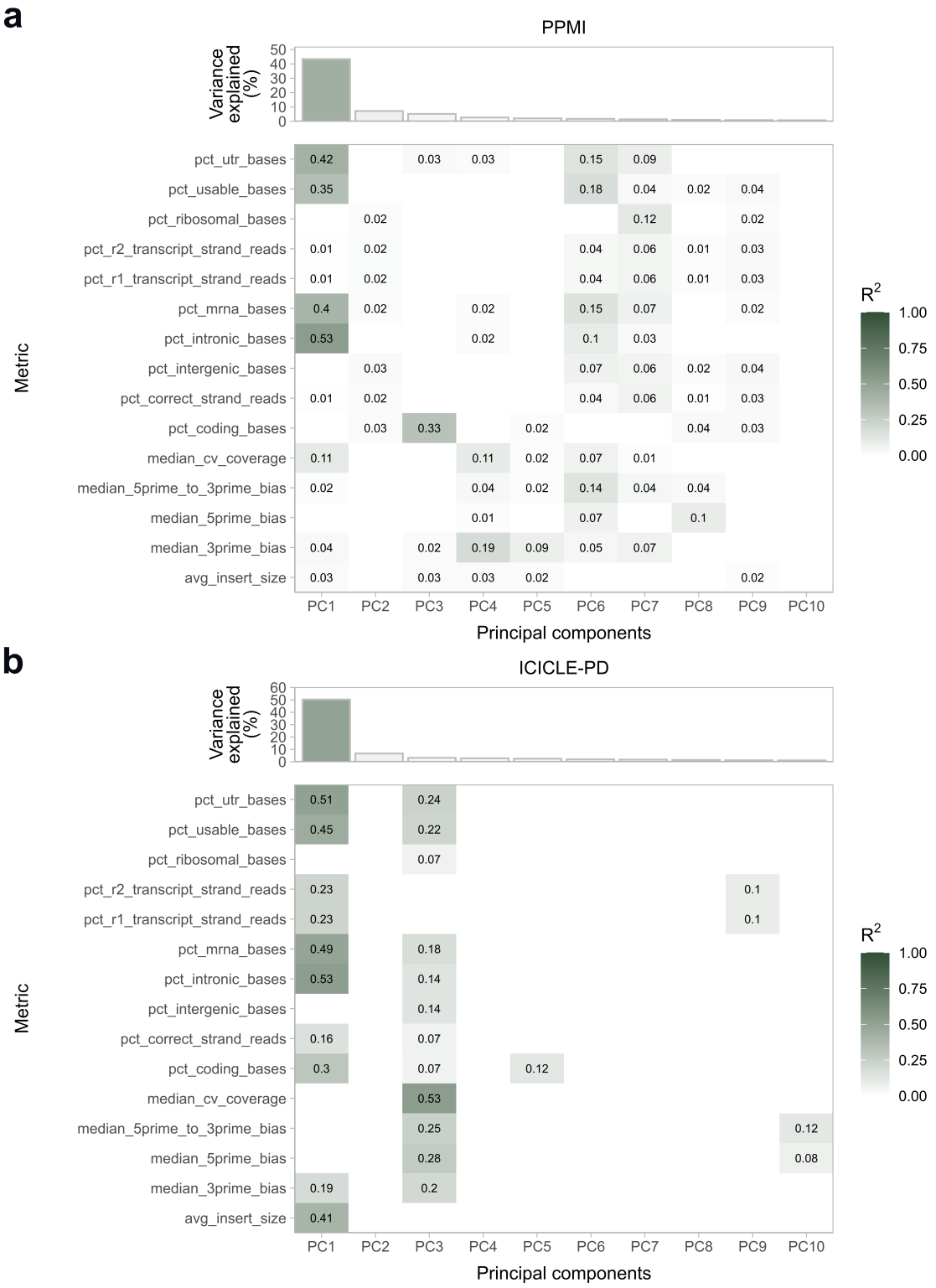


Figure 4.8. Sources of extraneous technical variation when quantifying circular RNA expression. (a, b) Results from univariate linear regression of technical metrics in the PPMI (**a**) and ICICLE-PD (**b**) cohorts against each of the first 10 principal components (PCs) of variance-stabilised transformed circRNA counts. The top panel of each figure shows the percentage of variation explained by each PC. The amount of variation explained (R^2) by each technical metric is only shown for metrics which passed multiple testing correction (FDR <0.05).

Next, I determined how sources of technical variation identified in the previous step, as well as other common sources of technical (i.e., sequencing batch) and biological variation (i.e., age, sex, study group) influenced the expression of individual circRNAs. Where multiple technical metrics were identified, highly correlated metrics (Spearman's rho >0.9) were removed to reduce collinearity when included in linear mixed models (Hoffman & Schadt, 2016). In PPMI, I identified the percentage of intronic bases as the only technical covariate (**Figure 4.9a**). As multiple technical covariates were identified in ICICLE-PD, I assessed the pairwise correlation between the percentage of intronic bases, the percentage of UTR bases and the median coefficient of variance of coverage (**Figure 4.9b**). There was a negative correlation between the percentage of intronic and UTR bases (Spearman's rho = -0.95) (**Figure 4.9b**). As the percentage of intronic bases explained a larger proportion of the variance in PC1 (R^2 = 0.53) compared with UTR bases (R^2 = 0.51) (**Figure 4.8b**), I retained the percentage of intronic bases as a covariate. Further, the relationship between the percentage of intronic bases and the median coefficient of variation of coverage was not colinear (<0.9 Spearman's rho) (**Figure 4.9b**).

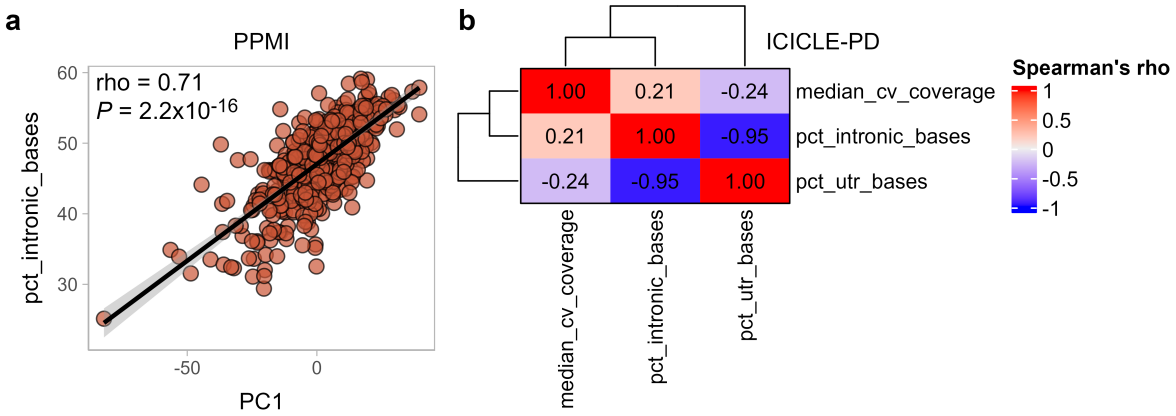


Figure 4.9. Colinear sources of extraneous technical circular RNA expression variation. (a) Positive correlation between the percentage of intronic bases and the first circRNA expression principal component in the PPMI cohort. Correlation reported as Spearman's rho coefficient. **(b)** Pairwise correlations between sources of extraneous technical circRNA expression variation in the ICICLE-PD cohort. Correlations reported as Spearman's rho coefficient.

Examination of the proportion of variance explained by the included covariates revealed they generally explained more variance than condition (representing the study group; PD or control), justifying their inclusion when modelling circRNA expression (**Figure 4.10**).

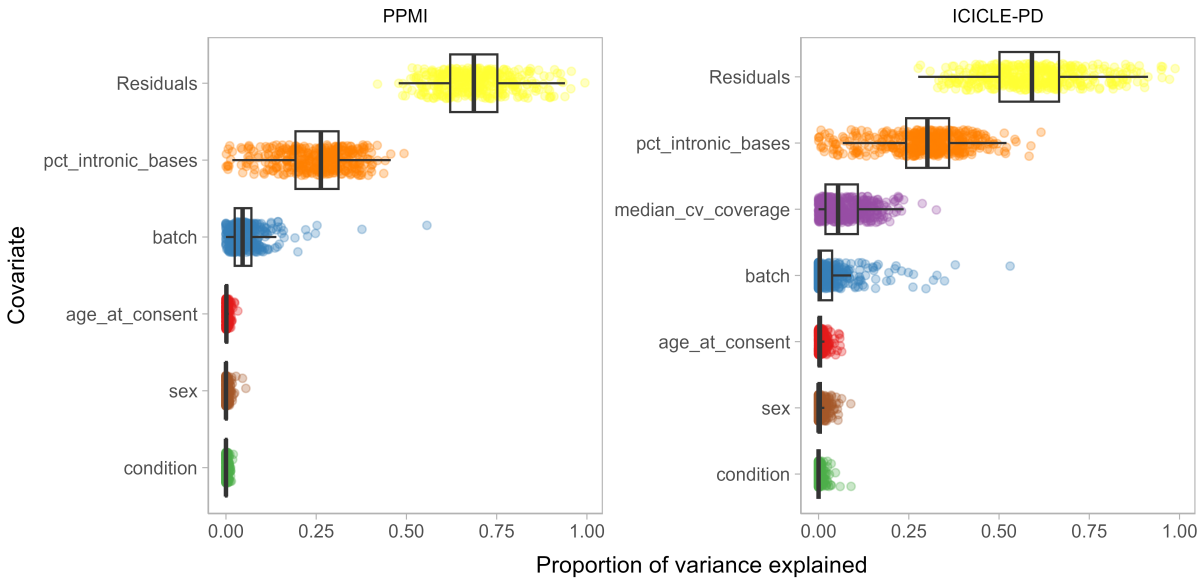


Figure 4.10. Contribution of covariates to individual circular RNA expression. Each point represents a circRNA, giving the proportion of variation explained by the covariates on the x-axis.

4.3.4 Characterising circular RNAs in whole blood

CircRNA properties have previously been examined in whole blood using a small number of samples (Memczak et al., 2015). I expand upon this knowledge by describing circRNA characteristics within the context of the PPMI and ICICLE-PD cohorts.

To assess the novelty of the detected BSJs in PPMI and ICICLE-PD datasets, I compared the circRNAs detected here against those reported in CircAtlas v3.0 (Wu et al., 2024) and CIRCpedia v2.0 (R. Dong et al., 2018), based on BSJ coordinates. Over 99.9% of the circRNAs had been previously reported in public repositories (**Figure 4.11a, b**). 54.8% of all circRNAs detected across both cohorts were detected in both cohorts (**Figure 4.11c**). 38.8% of all circRNAs were detected in PPMI samples compared to 6.4% detected in ICICLE-PD samples, demonstrating the increased ability of larger sample sizes to detect circRNAs (**Figure 4.11c**).

Most circRNAs were detected in a minority of the samples of each cohort and generally expressed at low levels (**Figure 4.11d**). Specifically, I detected 91.5% and 87.6% of circRNAs in PPMI and ICICLE-PD cohorts in less than half the respective cohorts with a median normalised expression < 0.1 (BSJ counts per million transcriptome mapped reads). However, certain circRNAs, such as 1:95143890-95151419:+ (*TLCD4-RWDD3*, *TLCD4*), 16:85633913-85634132:+ (*GSE1*), 17:20204332-20205912:+ (*AC004702.1*, *SPECC1*), 6:4891712-4892379:+ (*CDYL*), were among the most highly and commonly expressed in both cohorts (**Figure 4.11d**). The 13,728 circRNAs detected in both cohorts exhibited significantly

correlated expression (Pearson's r = 0.94) (**Figure 4.11e**). Together, these results indicate consistent detection and quantification of circRNAs across PPMI and ICICLE-PD samples.

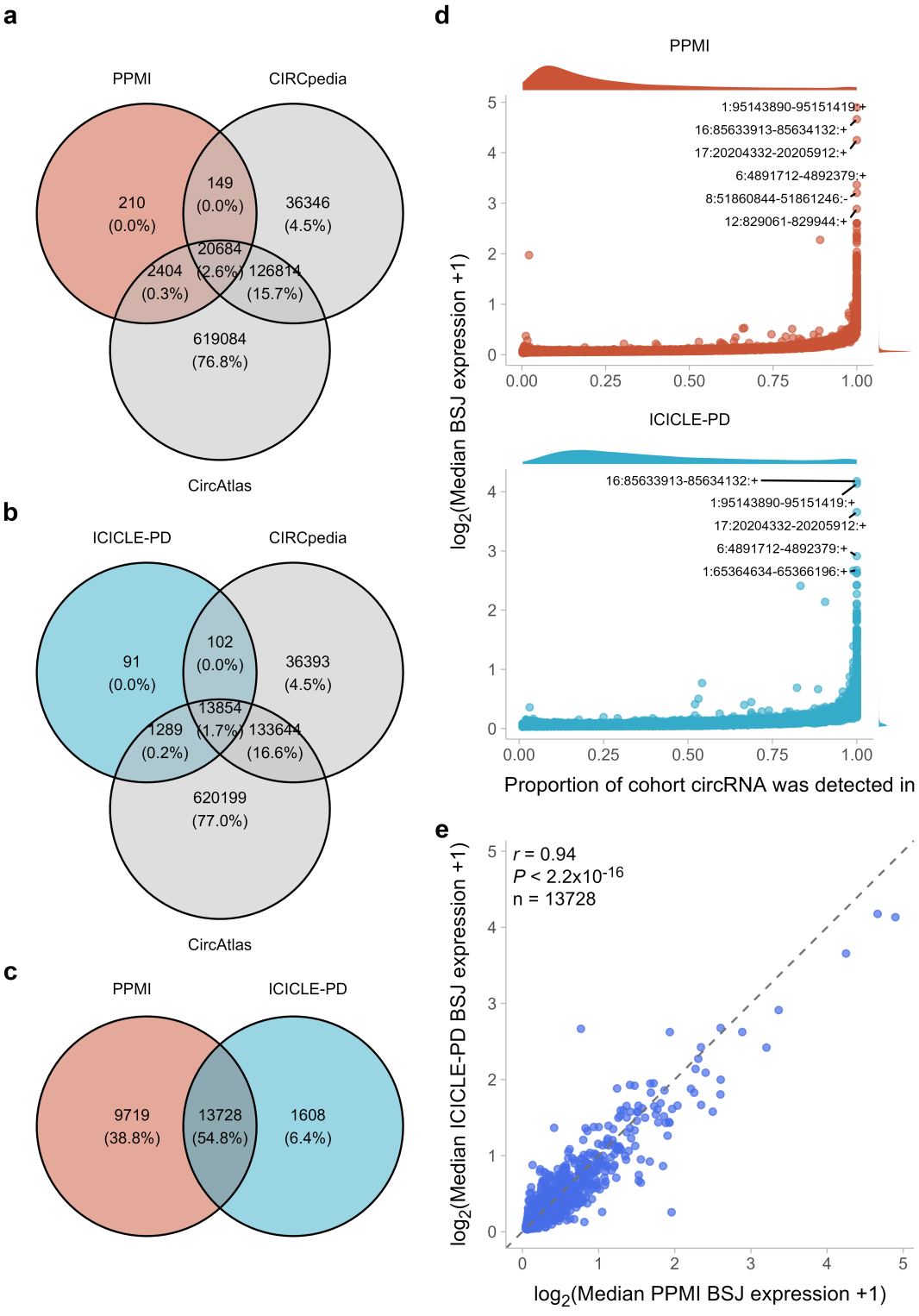


Figure 4.11. Characteristics of circular RNA detection and expression in PPMI and ICICLE-PD samples. (a, b) Overlap between the circRNAs detected in PPMI **(a)** and ICICLE-PD **(b)** samples compared to circRNAs present in the CircAtlas and CIRCpedia databases. (**c**) Overlap between circRNAs detected in PPMI and ICICLE-PD cohorts. **(d)** Some circRNAs exhibit ubiquitous and robust expression. **(e)** Comparison of circRNA expression in PPMI and ICICLE-PD samples. Correlation given as Pearson's *r* coefficient. Grey dashed line indicates equivalent expression between cohorts. circRNAs expression is based on the back-spliced junction count normalised against the number of transcriptome-mapped reads in each sample. CircRNA expression is summarised based on the median expression in each cohort, $\log_2$ transformed with a pseudo count of 1 added.

The proportions of BSJ genomic locations were similar in PPMI and ICICLE-PD datasets (**Figure 4.12a, b**). The largest proportion of BSJ coordinates mapped to exons (PPMI = 97.3%, ICICLE-PD = 97.6%), followed by introns (PPMI = 2.4%, ICICLE-PD = 2.1%), intergenic regions (0.2% in both cohorts) and finally antisense to genomic features (<0.1% in both cohorts) (**Figure 4.12a**). Of BSJs located within genes, most mapped to protein-coding genes (PPMI = 97.3%, ICICLE-PD = 97.4%). A smaller proportion mapped to lncRNAs (PPMI = 1.7%, ICICLE-PD = 1.6%); other genomic features had BSJ mapping rates <1%. (**Figure 4.12b**). Circular RNA host genes generally produced one circRNA (PPMI = 37.0%, ICICLE-PD = 44.4%) (**Figure 4.12c**). Several circRNA-host genes, such as *HECTD4*, *SOX6*, and *ABCC4* had comparatively higher production of different circRNAs, producing >20 circRNAs compared to a median of two in both cohorts (**Figure 4.12d**). The ability to produce unique BSJs was positively correlated to the number of exons within the gene (PPMI, Pearson's $r = 0.25$, P-value $<2.2\text{x}10^{-16}$; ICICLE-PD, Pearson's $r = 0.23$, P-value $< 2.2\text{x}10^{-16}$) (**Figure 4.12d**).

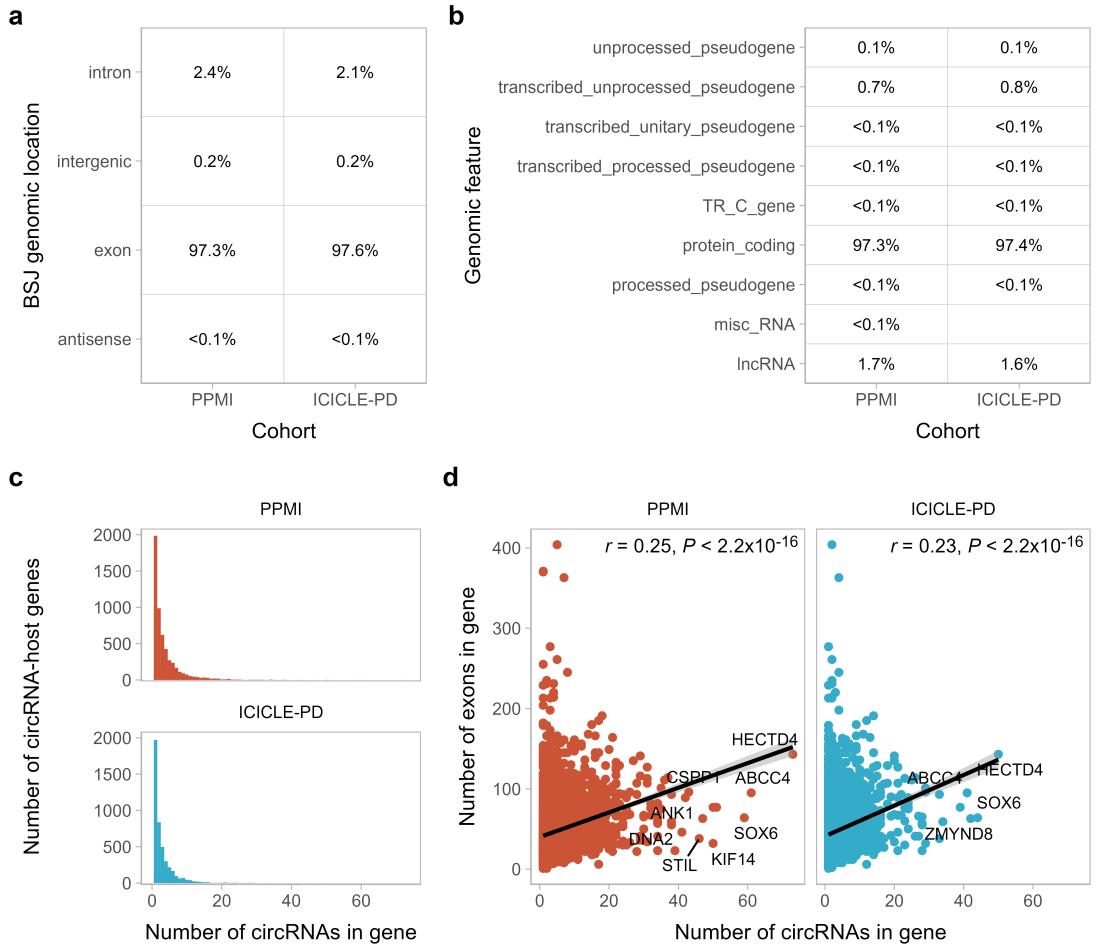


Figure 4.12. Characteristics of circular RNA-host genes. (a) Annotation of back-spliced junction based on genomic location. **(b)** Gene biotypes based on back-spliced junctions located within genes. **(c)** Histogram showing the distribution of the number of circRNAs produced by circRNA-host genes. **(d)** The relationship between the number of circRNAs produced by circRNA host-genes and the number of exons within that gene. Correlation assessed using Pearson's *r* coefficient.

Reads spanning canonical, or forward splice junctions (FSJs) can be used to estimate corresponding linear RNA expression (**Figure 4.13a**). The relationship between the amount of circular and linear RNA at a BSJ can then be summarised as the circular:linear (or BSJ:FSJ ratio). The circular:linear ratio at a locus may depend on gene function. For example, there is a tendency for circRNAs produced from lncRNAs to show higher circular:linear ratios than circRNAs produced from protein-coding genes (Izuogu et al., 2018), as also shown in **Figure 4.13b**.

I then examined the relationship between circular and linear RNA expression in PPMI and ICICLE-PD samples. There was generally no correlation between circular and linear RNA expression within samples (PPMI median Spearman's rho = 0.081, IQR = 0.03, ICICLE-PD median Spearman's rho = 0.079, IQR = 0.044) (**Figure 4.13c**). Similar results were also

observed when stratified by PD patients (PPMI median = 0.079, IQR = 0.029, ICICLE-PD median = 0.077, IQR = 0.043) and controls (PPMI median = 0.083, IQR = 0.033, ICICLE-PD median = 0.083, IQR = 0.054) (**Figure 4.13c**). At the individual junction level, a subset of junction loci had higher circRNA expression compared to the corresponding linear RNA (PPMI = 19.1%, ICICLE-PD = 17.2%) and was largely composed of junctions where no corresponding linear RNA was detected (PPMI = 15.5%, ICICLE-PD = 12.2%) (**Figure 4.13d**).

However, this only describes the contributions of the circular to linear ratio at BSJ loci. As the location of BSJs within genes is not random (Ragan et al., 2019) and host-genes can give rise to numerous BSJs (**Figure 4.12c**), I quantified the circular:linear expression ratio within genes. Where possible, I collapsed the circular and linear RNA expression to the host-gene and calculated the expression ratio. Some genes expressed circRNAs at higher levels than the comparative linear RNA expression at BSJ loci (PPMI = 5.9%, ICICLE-PD = 6.1%) (**Figure 4.13d**). In both cohorts, genes such as *GSE1*, *CDYL* and *STIL*, produced higher levels of circRNAs than cognate linear RNA (**Figure 4.13d**). Finally, certain circRNA host genes produced no detectable linear RNA expression at BSJ loci (PPMI = 1.0%, ICICLE-PD = 0.9%) (**Figure 4.13d**).

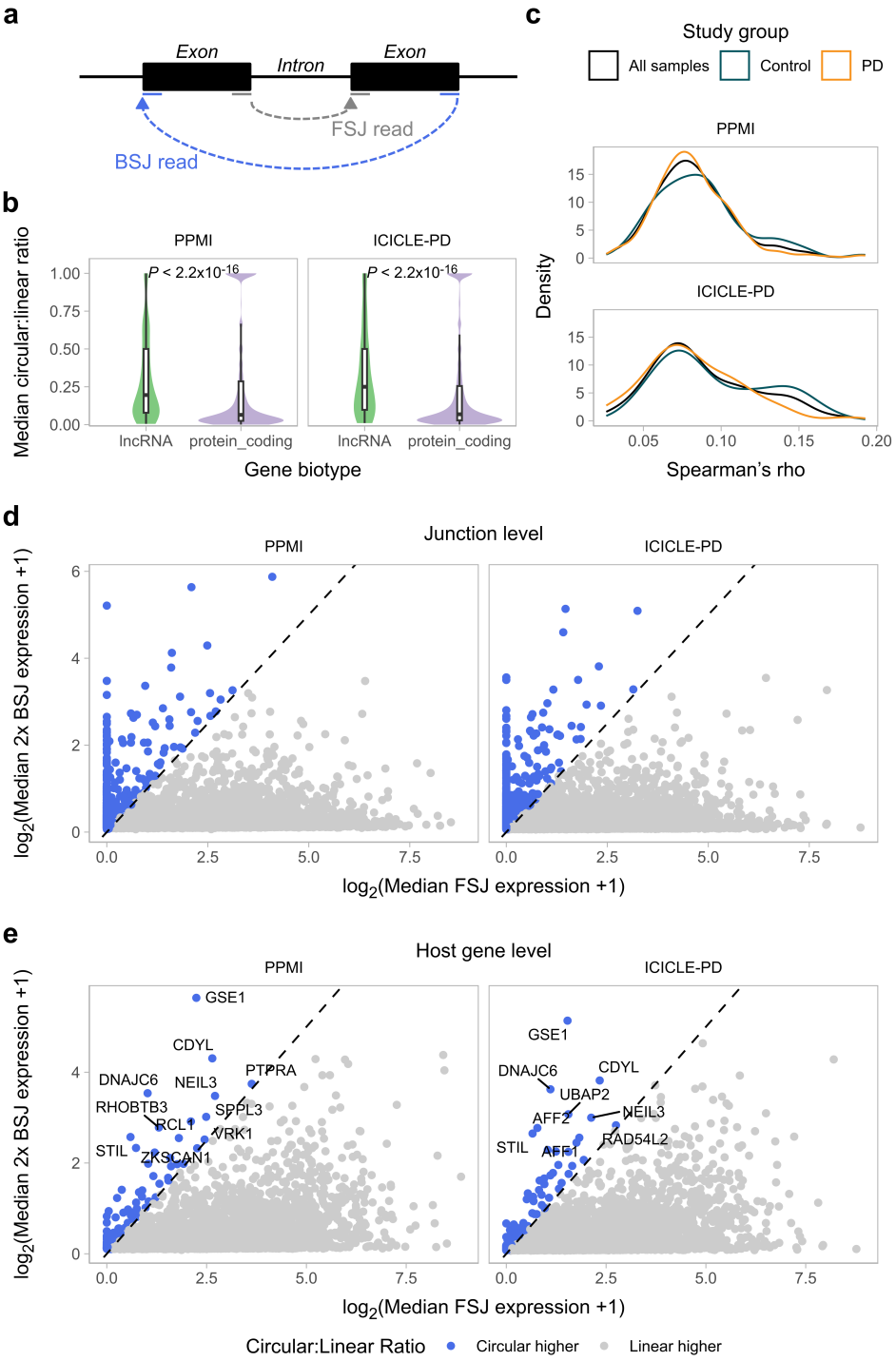


Figure 4.13. Circular RNA expression relative to linear RNA expression. (a) Schematic showing the detection of circular (back-spliced junction, BSJ) and linear (forward-spliced junction, FSJ) RNA expression by the mapping of RNA sequencing reads. **(b)** Comparison of the median circular:linear expression ratios of circRNAs produced by lncRNAs and protein-coding genes. Group differences were assessed using a Wilcoxon rank-sum test. **(c)** Density plot of the correlation between circular and linear RNA expression in each sample in the PPMI and ICICLE-PD datasets. Correlation was measured using Spearman's rho coefficient. Shown are the correlations across all samples, and when stratified by sample group (PD or Control). **(d, e)** Comparisons of the circular and linear expression at each junction locus **(d)** or when collapsed to the host gene **(e)**. BSJ expression was multiplied by two to account for the fact that BSJ expression is measured at one position, whereas FSJ expression is measured at two positions, following the method of CIRIquant (Zhang et al., 2020). The black dotted line shows the slope at which the circular and linear RNA expression is equal. Loci at which the circRNA expression is higher relative to the linear RNA expression are highlighted in blue.

4.3.5 Diversity of the blood circular RNA transcriptome in PD

CircRNA diversity, or the number of unique circRNAs detectable in an individual, is partially explained by circRNA biogenesis and turnover (Yang et al., 2022). Previous work has demonstrated increased circRNA diversity in the substantia nigra, but decreased diversity in the medial temporalis gyrus and amygdala (Hanan et al., 2020).

To assess the circRNA and circRNA-host gene diversity in blood, I collated the number of normalised circRNA and circRNA-host genes detected (**Section 4.2.8**). Different sequencing batches produced variability in the levels of circRNA (PPMI *P*-value = 6.5×10^{-9}, ICICLE-PD *P*-value = 0.01) and circRNA-host genes detected (PPMI *P*-value = 1.0×10^{-15}, ICICLE-PD *P*-value = 1.8×10^{-4}, Kruskal-Wallis test) (**Figure 4.14a, b**). As such, I included sequencing batch as a covariate when modelling the detection of circRNAs. Furthermore, circRNA or circRNA-host gene diversity was not associated with age or sex after adjusting for sequencing batch (*P*-value >0.05, linear regression, **Table 4.1**).

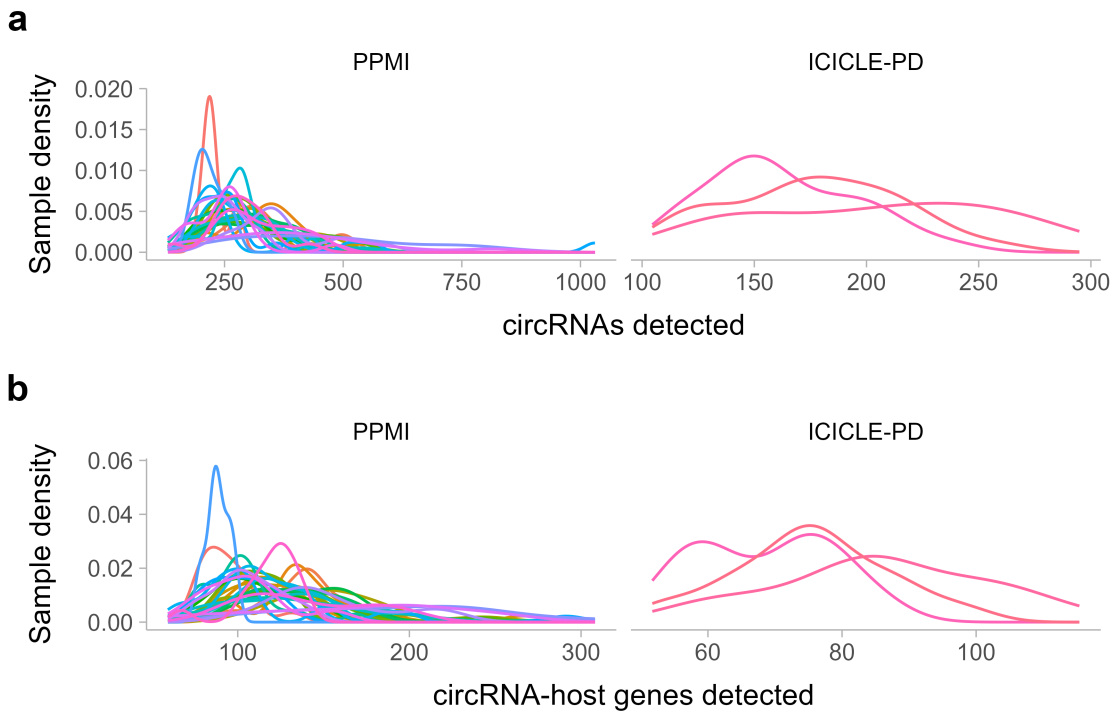


Figure 4.14. Variation in the number of circular RNAs and circular RNA-host genes across sequencing batches. (a, b) The distribution of circRNA (**a**) and circRNA-host gene diversity (**b**) in each sample across sequencing batches in the PPMI and ICICLE-PD cohorts. The number of circRNAs and circRNA-host genes in each sample was normalised against the number of transcriptome-mapped reads. Each coloured line represents the distribution for a different sequencing batch.

		β (95% CI)		*P*-value	
Response variable	**Explanatory variable**	*PPMI*	*ICICLE-PD*	*PPMI*	*ICICLE-PD*
Unique circRNAs					
	Sex	-8.4 (-29, 12)	15 (-2.5, 32)	0.42	0.09
	Age	-0.2 (-1.1, 0.72)	0.13 (-0.8, 1.1)	0.67	0.78
Unique circRNA-host genes					
	Sex	-3.2 (-10, 3.8)	3.6 (-1.5, 8.7)	0.37	0.16
	Age	0.011 (-0.31, 0.33)	-0.09 (-0.37, 0.19)	0.95	0.52

Table 4.1. Sex and age are not significantly associated with the number of circular RNAs or circular RNA-host genes detected. Output of multiple linear regression testing if age and sex were associated with the number of unique circRNAs or circRNA-host genes following adjustment for sequencing batch. The beta coefficient and 95% confidence interval show the effect of sex (Male) and age (per year) on the number of unique circRNAs and circRNA-host genes detected.

Following linear modelling taking covariates into account, PD patients in PPMI and ICICLE-PD cohorts showed significant reductions in circRNA diversity (PPMI, β = -20.2, *P*-value = 0.039; ICICLE-PD, β = -18.5, *P*-value = 0.029, linear regression) (**Figure 4.15a, b**). However, there were no significant reductions in circRNA-host gene diversity between PD patients and controls (*P*-value >0.05, linear regression) (**Figure 4.15c, d**). Overall, these findings indicate that circRNA biogenesis or turnover may be altered in early-stage idiopathic PD.

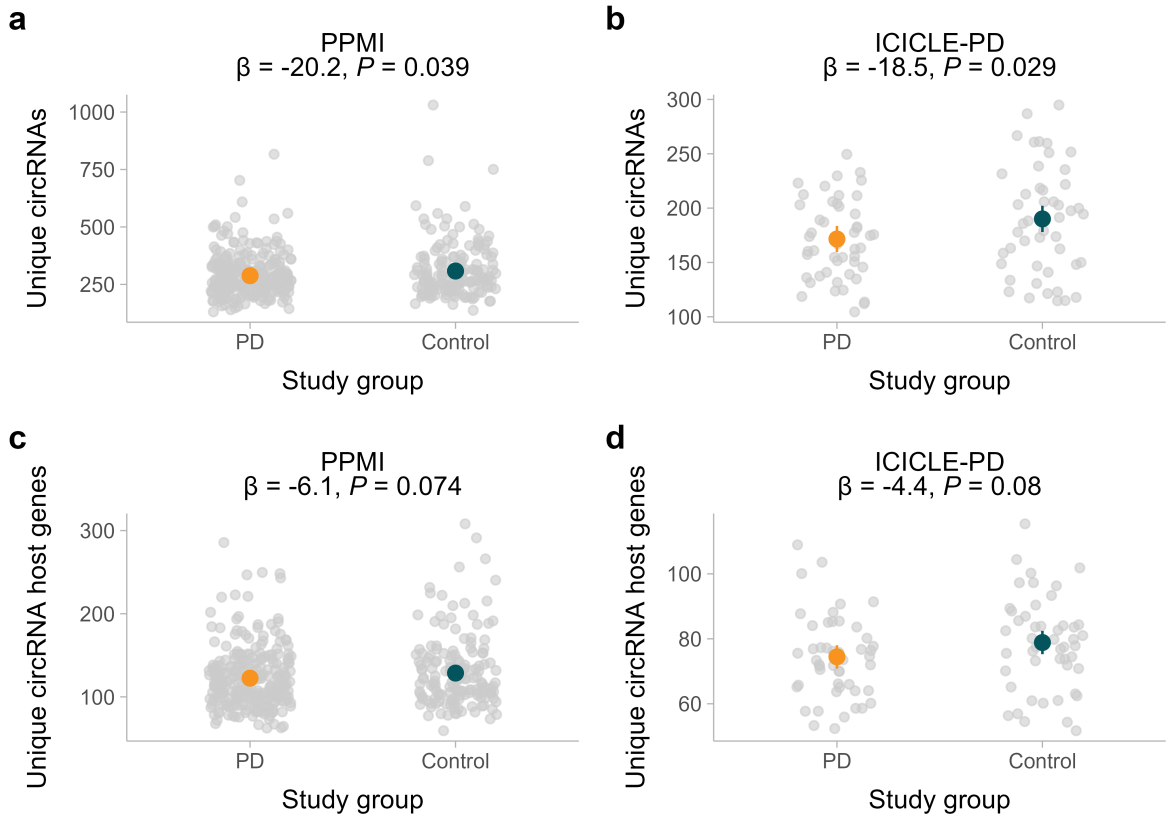


Figure 4.15. Circular RNA diversity in early-stage idiopathic PD. (a, b) CircRNA diversity in PD patients compared to controls. **(c, d)** CircRNA-host gene diversity in PD patients compared to controls. For all comparisons, a linear model was fit predicting the number of circRNAs/circRNA-host genes per million mapped reads depending on the study group, adjusted for sequencing batch. Above each panel, the β coefficient (showing the change in the normalised amount of unique circRNA/circRNA host genes in PD patients) for the study group term is shown along with the *P*-value of the test statistic for the study group term. Highlighted on each plot is the estimated marginal mean of each study group and 95% confidence intervals derived from the linear model.

4.3.6 Identification of differentially expressed circular RNAs in early-stage idiopathic PD

Motivated by reduced circRNA diversity in early-stage idiopathic PD, I next asked whether specific circRNAs were differentially expressed in PD and thus potential diagnostic biomarkers. Diagnostic biomarkers should be reliably detectable (Califf, 2018), yet for many circRNAs, this was not the case (**Figure 4.11c**). In keeping with the gene differential expression analysis, I imposed a minimum expression criterion (**Section 4.2.9**). Post-filtering, 403 circRNAs in PPMI and 457 in ICICLE-PD remained, with an overlap of 331 (62.6% of unique, abundant circRNAs).

Enriched circRNA expression may correspond to the functionality of their host genes (Dong et al., 2023). As such, I investigated the enrichment of host gene functions for highly expressed circRNAs in comparison to circRNAs with lower expression levels. Several Gene Ontology

(GO) terms showed enrichment (*P*-value <0.05) however, none were significant after multiple testing correction (FDR >0.05) (**Table 8.2**).

I identified three circRNAs significantly reduced in early-stage idiopathic PD versus controls in PPMI (log_2 fold change >0.1/<-0.1, FDR < 0.05, Wald test) (**Figure 4.16a).** These circRNAs, derived from the genes *ESYT2* (7:158759485-158764853), *BMS1P1* (10:46795805-46798168:+), and *CCDC9* (19:47264602-47264946:+) were similarly decreased in ICICLE-PD but did not reach statistical significance (*P*-value >0.05, Wald test) (**Figure 4.16a, b**). Intriguingly, there appeared to be a global reduction of circRNA expression in early-stage idiopathic PD (**Figure 4.16a**). Comparison of the fold changes (**Section 4.2.9**) revealed a reduction in circRNAs expression in both PPMI (imbalance = 0.09, *P*-value < $2.2x10^{-16}$, Exact binomial test) and ICICLE-PD (imbalance = 0.29, *P*-value = $2.3x10^{-10}$, Exact binomial test).

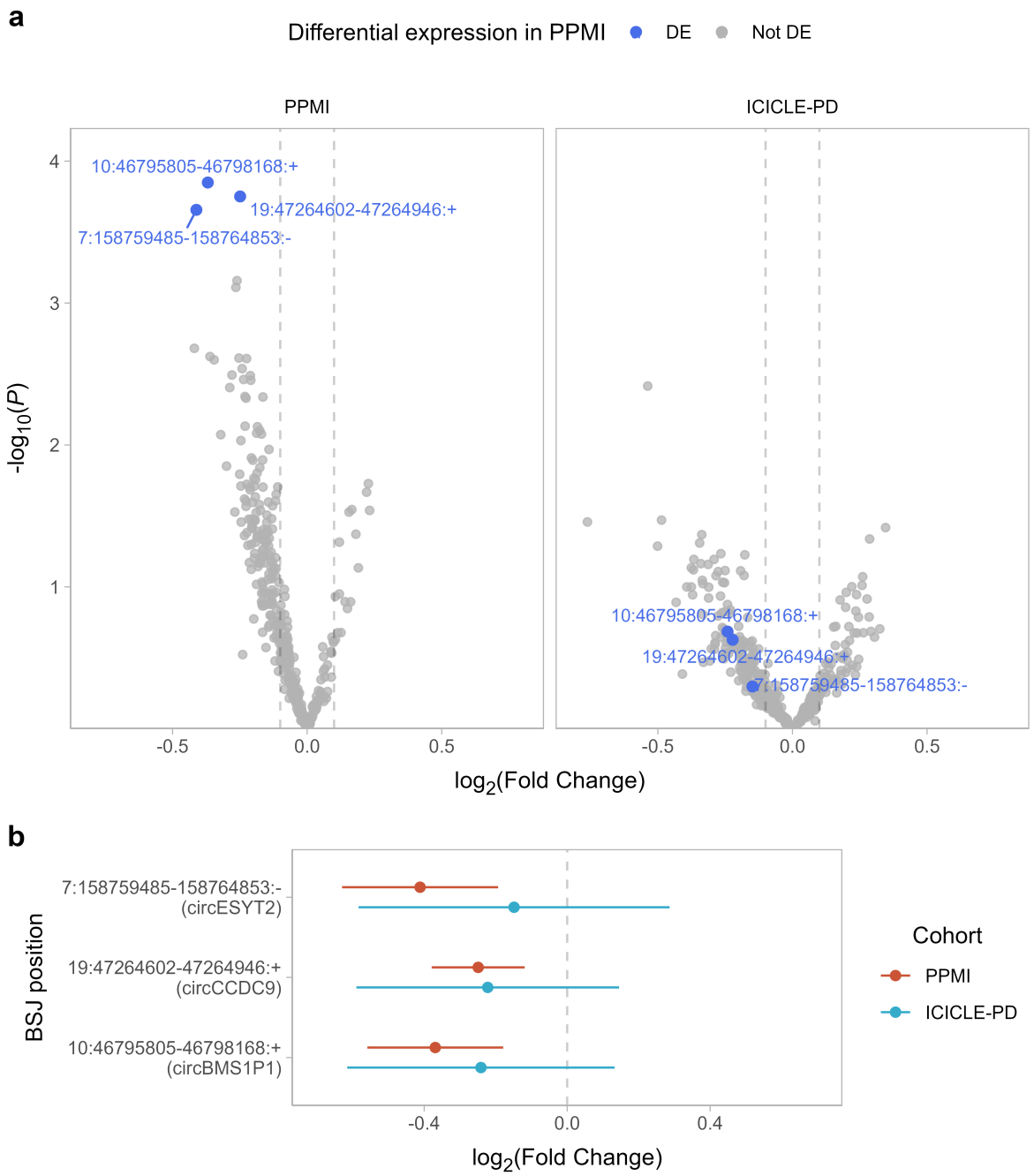


Figure 4.16. Circular RNA differential expression in early-stage idiopathic PD. (a) Volcano plot of differential expression in PPMI and ICICLE-PD, highlighted (blue) are the three circRNAs significantly differentially expressed in PPMI (FDR <0.05 and $\log_2$ fold change >0.1/<-0.1) and the corresponding results in ICICLE-PD. Grey dashed lines indicate the $\log_2$ fold changes of 0.1 and -0.1. **(b)** Log_2 fold changes of the three circRNAs that were significantly differentially expressed in PPMI. None reached significance in ICICLE-PD (*P*-values >0.05). Grey dashed line indicates a $\log_2$ fold change of 0.

4.3.7 Correlating circular RNA expression with PD-related clinical measures

Due to the difference in treatment status of PD patients in the PPMI and ICICLE-PD cohorts, I first used linear regression to assess the association of treatment dosage (given as the levodopa

equivalent daily dose, LEDD), with circRNA expression, adjusting for sources of biological and technical variation (**Section 4.2.5***)*. I did not detect any circRNAs that were significantly associated with LEDD (FDR >0.05, linear regression) in ICICLE-PD cases (**Figure 4.17**).

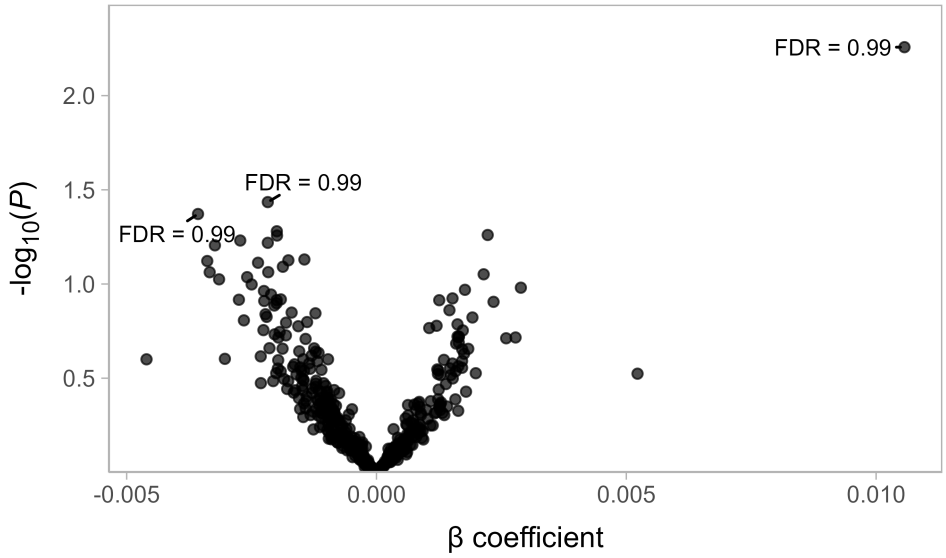


Figure 4.17. Influence of PD treatment dosage on circular RNA expression in ICICLE-PD. A linear regression model was fit as described in **Section 4.2.10**). The β coefficient shows the modelled change in log2-transformed normalised circRNA expression based on per unit change in treatment dosage. The FDR-adjusted *P*-value is shown for the circRNAs in which the treatment dosage term was significantly associated with circRNA expression (*P*-value <0.05).

Given the identification of three circRNAs differentially expressed in PPMI PD patients, I then compared the expression of these circRNAs to PD-related clinical measures. I tested the association between circRNA expression and the age of diagnosis, motor severity (assessed using MDS-UPDRS III scores) and cognitive impairment (assessed using MoCA scores) in PD patients using linear regression models (**Section 4.2.10**)

None of the three circRNAs tested were significantly associated with age at diagnosis, motor symptom severity or cognitive impairment in PD patients in the PPMI and ICICLE-PD cohorts (*P*-value >0.05, linear regression, **Table 4.2**). The lack of association suggests that these specific circRNAs do not warrant further study as progression biomarkers in PD.

			β (95% CI)		*P*-value		Adjusted *P*-value	
Response variable	**BSJ position**	**Gene symbol**	*PPMI*	*ICICLE-PD*	*PPMI*	*ICICLE-PD*	*PPMI*	*ICICLE-PD*
Age at PD diagnosis								
	10:46795805-46798168:+	*BMS1P1*	-0.054 (-0.61, 0.5)	-0.12 (-1.7, 1.5)	0.85	0.87	1.00	1.00
	19:47264602-47264946:+	*CCDC9*	-0.44 (-1.3, 0.45)	-1.3 (-2.9, 0.32)	0.33	0.11	0.99	0.34
	7:158759485-158764853:-	*ESYT2*	-0.28 (-0.77, 0.21)	0.18 (-1.1, 1.5)	0.26	0.79	0.77	1.00
MDS-UPDRS III								
	10:46795805-46798168:+	*BMS1P1*	0.16 (-0.9, 1.2)	0.69 (-3.5, 4.8)	0.76	0.74	1.00	1.00
	19:47264602-47264946:+	*CCDC9*	0.18 (-1.5, 1.9)	-1.1 (-5.8, 3.5)	0.84	0.62	1.00	1.00
	7:158759485-158764853:-	*ESYT2*	0.77 (-0.17, 1.7)	-0.87 (-3.9, 2.2)	0.11	0.57	0.32	1.00
MoCA								
	10:46795805-46798168:+	*BMS1P1*	-0.066 (-0.38, 0.25)	1.1 (-0.14, 2.3)	0.68	0.08	1.00	0.24
	19:47264602-47264946:+	*CCDC9*	-0.48 (-0.99, 0.033)	0.8 (-0.74, 2.3)	0.07	0.30	0.20	0.89
	7:158759485-158764853:-	*ESYT2*	0.026 (-0.26, 0.31)	0.3 (-0.86, 1.5)	0.86	0.60	1.00	1.00

Table 4.2. Associations between PPMI differentially expressed circular RNAs and PD-related clinical measures. Table giving the output from linear regression of various clinical measures as the response variables and the corresponding circRNA expression as an explanatory variable. Linear models were fit adjusting for sources of biological and technical variation. For each model, the β coefficient and 95% confidence intervals show the effect on the response variable per unit increase of $\log_2$ normalised circRNA expression. *P*-values were adjusted for multiple tests using the Bonferroni procedure based on three tests (*P*-value threshold < 0.017). BSJ positions are reported as chromosome:start-end:strand (GRCh38). MDS-UPDRS III = Movement Disorder's Society Unified Parkinson's Disease Rating Scale Part 3, MoCA = Montreal Cognitive Assessment.

4.3.8 Replication of previously reported PD circular RNAs

Where available, I collected summary statistics from studies on individual circRNAs (based on BSJ position) that are differentially expressed in PD in the substantia nigra and peripheral blood mononuclear cells (PBMCs) (Hanan et al., 2020; Ravanidis et al., 2021) (**Table 8.3**). Seven circRNAs previously reported as differentially expressed in PD were sufficiently expressed in the PPMI or ICICLE-PD cohorts (**Table 8.3**). Two, derived from *DOP1B* and *INTS6L*, showed altered expression in PPMI (*P*-value <0.05, Wald test). While no circRNA reached statistical significance after multiple testing correction within each cohort (FDR > 0.05, Wald test), I was able to replicate the direction of change of four circRNAs in PPMI and three in ICICLE-PD (**Figure 4.18**).

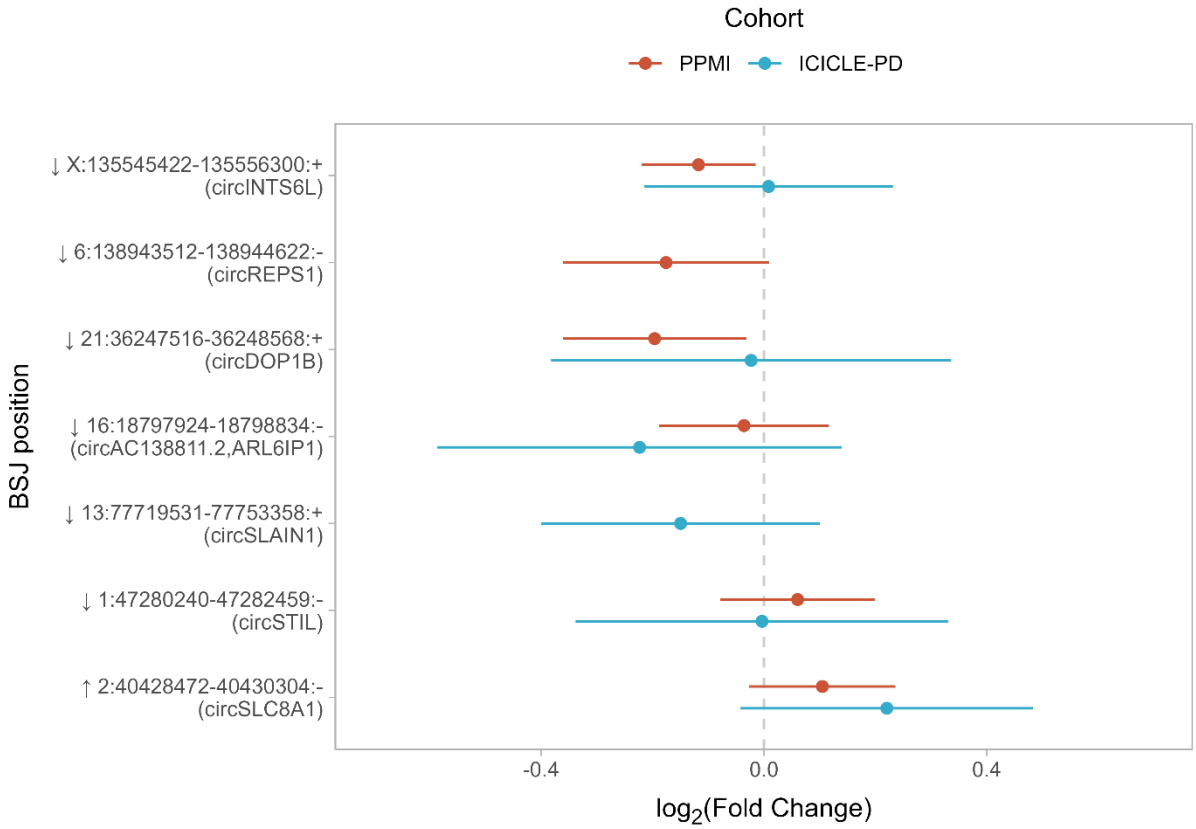


Figure 4.18. Fold changes of previously reported PD circular RNAs. Back-spliced junction (BSJ) positions are reported as chromosome:start-end:strand (GRCh38) and the hosting gene. The arrows indicate the previously reported direction of change in PD relative to controls based on summary statistics (Hanan et al., 2020; Ravanidis et al., 2021). Error bars show the 95% confidence interval of the fold change.

4.3.9 Overlap with known PD and parkinsonism risk loci

I compiled genomic loci associated with PD risk from the summary statistics of the largest genome-wide association study (GWAS) in PD to date (Nalls et al., 2019). A total of four (PPMI) and five (ICICLE-PD) circRNAs included in differential expression testing were located within the nearest genes to risk loci. In PPMI, three circRNAs were located within the gene *UBAP2*, the nearest gene to rs6476434 (chr9: 34046391, GRCh37). The other circRNA, produced by *DYRK1A*, is proximal to rs2248244 (chr21: 38852361, GRCh37). Of these circRNAs, 21:37420298-37472880:+ located in *DYRK1A*, was reduced in PD patients in PPMI (**Figure 4.19a, b**). In ICICLE-PD, I observed a corresponding reduction in 21:37420298-37472880:+ expression, yet this did not reach statistical significance (*P*-value >0.05, Wald test) (**Figure 4.19a, b**).

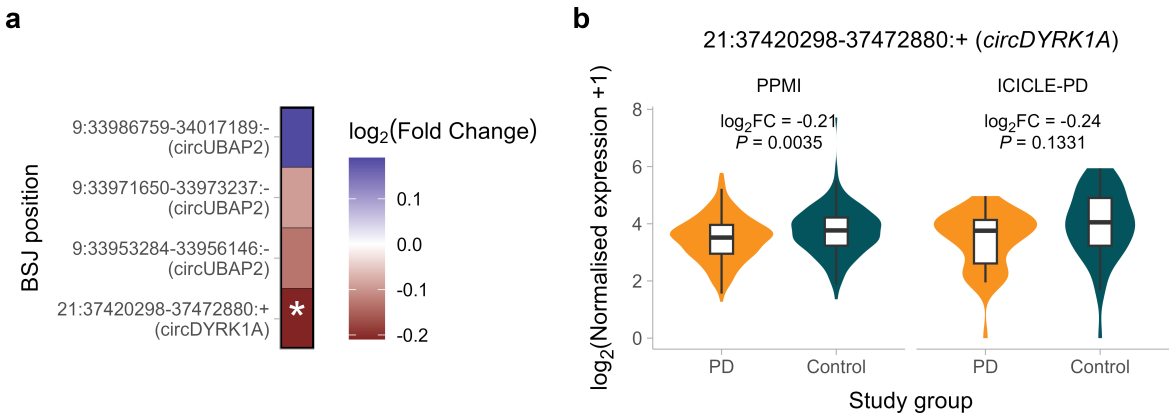


Figure 4.19. Differential expression of circular RNAs produced by genes proximal to PD GWAS risk loci. (a) Heatmap showing the expression changes of four circRNAs located within genes proximal to PD GWAS loci that were detected in the PPMI cohort. For each circRNA, fold change is reported in PD patients relative to controls. CircRNAs differentially expressed following FDR correction (based on the four circRNAs detected) are highlighted (*). **(b)** Expression of the circRNA located at 21:37420298-37472880:+ between PD patients and controls in PPMI and ICICLE-PD cohorts. Fold change and Wald test *P*-values were obtained from the results of testing for differential expression between PD patients and controls adjusting for sources of biological and technical variation. CircRNA expression was normalised using the median of ratios method in DESeq2. All BSJ positions are reported as chromosome:start-end:strand with respect to the Ensembl GRCh38 reference.

Subsequently, I investigated the expression of circRNAs produced by genes that are established, causative parkinsonism genes (Martin et al., 2019). For instance, circRNAs produced by *SNCA*, the gene encoding α-synuclein, have been reported to be elevated in PD models *in vitro* (Sang et al., 2018). However, I was unable to detect any circRNAs produced by *SNCA* in PPMI or ICICLE-PD cohorts regardless of expression filtering.

Of the 35 parkinsonism genes, three (*DNAJC6*, *FBXO7* and *SPG11*) produced circRNAs could be assessed by differential expression analysis (the circRNA produced by *SPG11* did not pass the expression threshold in PPMI samples). However, none of these circRNAs were differentially expressed in PD patients in either the PPMI or ICICLE-PD cohorts (*P*-value >0.05, Wald test) (**Table 4.3**).

		Log_2 Fold Change		***P*-value**	
BSJ position	Gene symbol	PPMI	ICICLE-PD	PPMI	ICICLE-PD
1:65364634-65366196:+	*DNAJC6*	-0.15	-0.19	0.09	0.35
22:32478980-32479275:+	*FBXO7*	-0.16	-0.06	0.10	0.76
15:44615362-44620403:-	*SPG11*	N/A	-0.09	N/A	0.63

Table 4.3. Differential expression of circular RNAs produced by genes causative for Parkinson's disease and complex parkinsonism disorders. For each circRNA, the BSJ position is reported as chromosome:start-end:strand (GRCh38). The log2 fold change shows the change in PD patients relative to controls. Wald test *P*-values obtained from differential expression testing after adjusting for sources of biological and technical variation.

4.3.10 Evaluating circular RNA as a diagnostic biomarker of early-stage idiopathic PD

To evaluate the use of selected circRNAs as a predictor of early-stage idiopathic PD, I first examined the potential of individual junction loci to classify PD from controls. I assessed BSJ expression at each locus, in addition to the corresponding FSJ expression as well as the ratio between BSJ:FSJ expression. For each of these values, I constructed receiver operator characteristic (ROC) curves and then calculated the area under the ROC curves (AUC) to summarise its potential PD classification potential. AUCs across the cohorts were positively correlated yet some variability was present (**Figure 4.20a)**. For example, FSJ expression at 15:41668827-41669958:+ had an AUC of 0.77 (95% CI = 0.67-0.86) in ICICLE-PD yet an AUC of 0.54 (95% CI = 0.48-0.59) in PPMI. Generally, predictors in the ICICLE-PD cohort performed better (median AUC = 0.57) than in the PPMI cohort (AUC = 0.53).

Based on the AUCs displayed in **Figure 4.20a**, there were no junctions with an AUC >0.8 in both PPMI and ICICLE-PD. To improve performance, I included multiple features (BSJ, FSJs, BSJ:FSJ ratios) in multivariable logistic regression models (**Section 4.2.11**). I trained and evaluated the classification performance of models using the PPMI cohort, then used the ICICLE-PD cohort as an independent test cohort to assess the performance of classifiers trained using the PPMI cohort.

As a baseline comparator, the gene expression model described in **Chapter 3** is also shown in **Figure 4.20**. Notably, classification using gene expression performed better than any of the circRNA-related predictors in the PPMI dataset. In the ICICLE-PD dataset, the performance of all predictors was much more uniform, in line with the reduced performance of the classification using gene expression in this dataset (**Chapter 3**). Of the standalone circRNA-related predictors evaluated, the classification based on BSJ expression had the highest AUC in the PPMI dataset (AUC = 0.61, 95% CI = 0.55-0.66) with a similar performance in the ICICLE-PD (AUC = 0.59, 95% CI = 0.48-0.71). Using FSJ expression or BSJ:FSJ expression ratio as classifiers produced AUCs <0.6 in the PPMI dataset. In the ICICLE-PD dataset, the BSJ:FSJ expression ratio showed the highest classification potential of all assessed predictors (AUC = 0.63, 95% CI = 0.51-0.74). Using both gene and BSJ expression marginally increased performance in both PPMI (AUC = 0.85, 95% CI = 0.81-0.89) and ICICLE-PD (AUC = 0.60, 95% CI = 0.48-0.71) datasets over the use of gene expression alone.

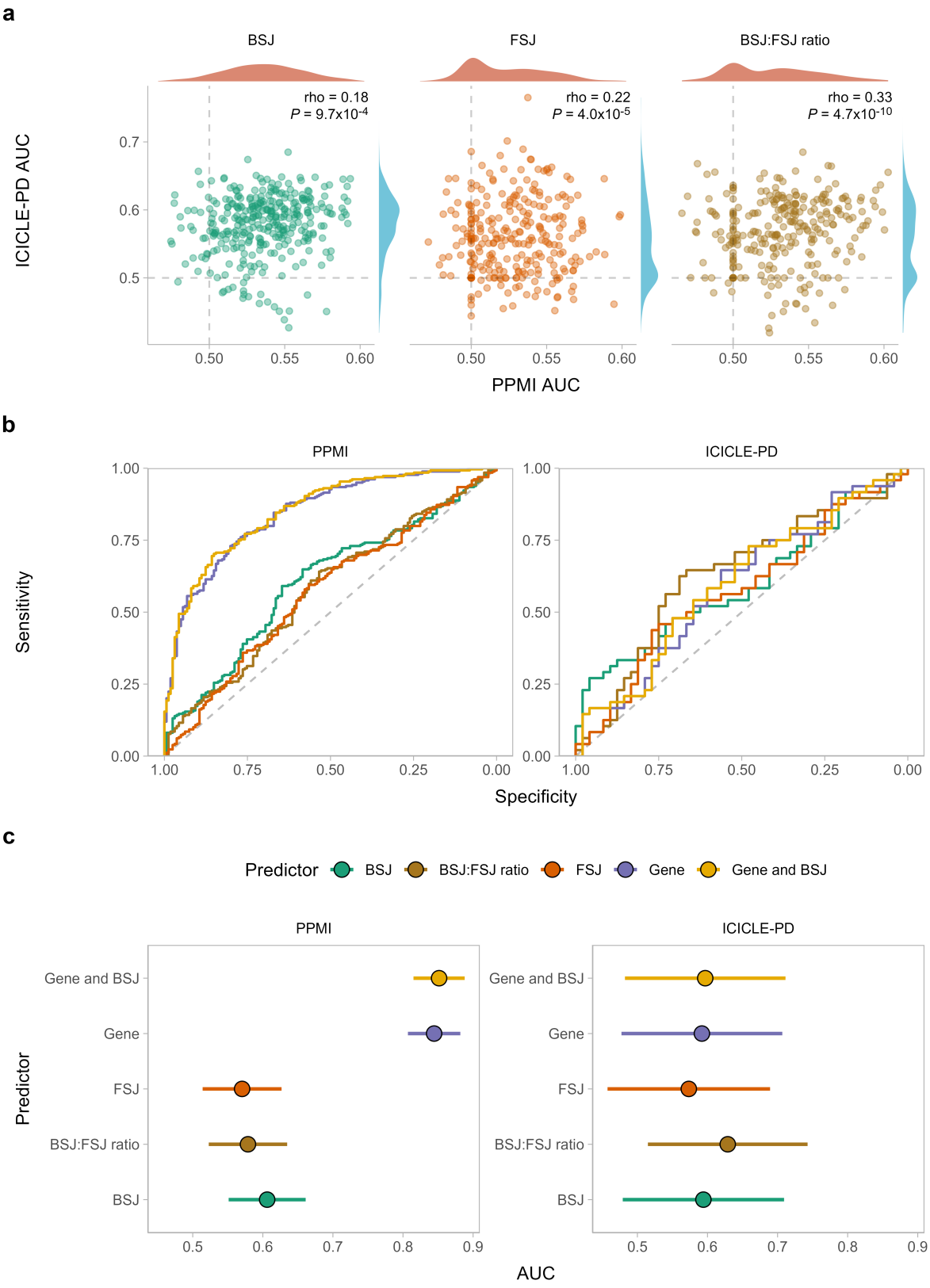


Figure 4.20. Using RNA expression to classify early-stage idiopathic PD status. (a) Comparison of AUCs calculated in PPMI and ICICLE-PD cohorts using the expression of individual junctions (BSJ and FSJ) and BSJ:FSJ ratios to classify PD status. Spearman's rho demonstrates the correlation between AUC in each cohort. Grey dashed lines indicate an AUC of 0.5. **(b, c)** Ability of each predictor to classify PD status. ROC curves are shown in **(b)** with the corresponding AUC shown in **(c)**. Error bars give the 95% confidence intervals of the AUC.

4.4 Discussion

Due to their increased stability and specific expression patterns relative to linear RNA, circRNAs have been proposed as a potential diagnostic biomarker in a range of diseases, including PD (Verduci et al., 2021; Doxakis, 2022). In this chapter, I presented work describing the detection and quantification of blood circRNAs in two cohorts (PPMI and ICICLE-PD) comprised of both PD patients and matched controls. CircRNAs showed poor performance as diagnostic PD biomarkers in both PPMI (AUC = 0.61) and ICICLE-PD (AUC = 0.59). Testing for differences between PD patients and controls identified three circRNAs that were significantly differentially expressed in the PPMI cohort. Despite a small number of differentially expressed circRNAs, there was a global reduction in circRNA expression in early-stage idiopathic PD patients in both PPMI and ICICLE-PD cohorts, which may be indicative of systematic dysregulation related to circRNA biogenesis or turnover. In conclusion, circRNAs have limited clinical utility as a standalone diagnostic biomarker for idiopathic PD. Future investigations into the mechanisms underpinning global circRNA expression may uncover new processes related to the pathobiology of PD.

4.4.1 Increased understanding of computational circular RNA analysis methods

Given the novelty of the circRNA field, methodologies are expected to be refined as the field develops. Publications on circRNA best practices have predominantly focused on experimental techniques (Dodbele et al., 2021; Nielsen et al., 2022). Discussions regarding the bioinformatic analyses of circRNAs often relate to challenges in their detection (Szabo & Salzman, 2016; Gao & Zhao, 2018; Jakobi & Dieterich, 2019; L. Chen et al., 2021; Ma et al., 2023; Rebolledo et al., 2023). The lack of discussion regarding the downstream analyses of circRNAs following their detection has led to conflicting methodologies.

Overlapping multiple circRNA detection algorithms followed consensus within the field to reduce false positives (Hansen et al., 2016; Hansen, 2018; Vromman et al., 2023). How to address the varying circRNA expression estimates produced by each tool (**Figure 4.4**) remains an unresolved question (Gaffo et al., 2017, 2021). Here, the use of a single, high-performing tool (Zhang et al., 2020; Vromman et al., 2023) to quantify circRNA expression from raw sequenced reads led to a unified count at each locus.

CircRNA expression required normalisation to account for inter-sample differences in sequencing depth. Typically, a measure of the total sequencing depth or the number of mapped reads is used (Zhang et al., 2014; Venø et al., 2015; Zhang et al., 2020). In some cohorts, these measures may not be appropriate, as exemplified by the lack of significant association between

total sequenced reads and genome-mapped reads to circRNA levels in ICICLE-PD samples (**Figure 4.6b**). I propose that other measures, such as the number of transcriptome-mapped reads, may be more informative. Regardless of the method used, this finding highlights the importance of evaluating the applicability of circRNA normalisation methods before downstream analysis.

In **Chapter 3**, I assessed the impact of biological and technical factors on gene expression. In this chapter, I investigated their influence on circRNA expression. Consistent with the drivers of gene expression variation discussed in **Chapter 3**, technical sources were among the top drivers of circRNA expression variation (**Figure 4.8** and **Figure 4.10**). Differences in identified technical factors between gene and circRNAs emphasise the need to identify gene and circRNA-specific technical factors. Notably, the proportion of bases mapped to intronic regions correlated to circRNA expression variation across both PPMI and ICICLE-PD cohorts, highlighting its importance as a key source of variation in circRNA studies.

4.4.2 Consistent detection and quantification of circular RNAs in PPMI and ICICLE-PD cohorts

Describing the characteristics of detected circRNAs served two purposes; a) as a quality control step to compare the detection of circRNAs in PPMI and ICICLE-PD cohorts, and b) to place these characteristics into the context of current literature regarding circRNA properties. Overall, reported circRNA characteristics were generally similar across cohorts, indicative of consistent detection and quantification. Notably, the expression of shared circRNAs was highly correlated across cohorts (**Figure 4.11e**), demonstrating the suitability of using the PPMI and ICICLE-PD cohorts in analyses.

Previous experiments profiling circRNA expression have resulted in several circRNA databases (reviewed in Vromman et al., 2021). Nearly all detected circRNAs were previously reported in public repositories (R. Dong et al., 2018; Wu et al., 2024), increasing confidence in the detection of bona-fide circRNAs. Most BSJs were located within gene boundaries (**Figure 4.12a**), consistent with the role of the canonical spliceosome machinery in the production of circRNAs (Ashwal-Fluss et al., 2014; Starke et al., 2015). BSJs mapping to intronic regions may reflect both exon-intron (Z. Li et al., 2015) and intronic circRNAs (Zhang et al., 2013; Panda et al., 2017). A small number of circRNAs were classed as intergenic, which have been previously detected adjacent to canonical splicing motifs and validated by RNase R resistance (Gao et al., 2015; Maass et al., 2017).

In most cases, circRNA host-genes produced one dominant circular isoform in the blood (**Figure 4.12c**), consistent with previous work (Memczak et al., 2015). When multiple circRNA isoforms were detected, this was correlated to the number of exons in the host-gene (**Figure 4.12d**) as described in tissues other than blood (Vo et al., 2019; Dong et al., 2023).

Within samples, the lack of consistent concordance between circRNA and corresponding linear RNA expression (**Figure 4.13c**) supports the notion that circRNA expression can be regulated independently of cognate linear RNA (Salzman et al., 2013; Memczak et al., 2015; Vo et al., 2019) and further outlines the need for studying circRNAs as a separate entity from genes. In some cases, at both the junction and gene level, no corresponding linear RNA was detected at BSJ loci (**Figure 4.12d, e**), possibly due to exon skipping (Kelly et al., 2015). Expression of circRNAs has previously been related to host-gene function (Memczak et al., 2015; You et al., 2015).

Differences in sample sizes between the PPMI and ICICLE-PD cohorts will have affected my ability to detect circRNAs and identify those differentially expressed in PD. Given the larger sample size, more circRNAs were detected in PPMI samples. Depletion of linear RNAs through the addition of RNase R during library preparation (Xiao & Wilusz, 2019), would enrich circRNAs, allowing the detection of rarer transcripts. However, not depleting linear RNAs allowed me to quantify both linear and circular RNAs simultaneously.

4.4.3 Three novel circRNAs differentially expressed in early-stage idiopathic PD in PPMI samples

My comparison of circRNA expression identified three that were significantly reduced in PPMI PD cases relative to controls (within the genes *BMS1P1*, *CCDC9* and *ESYT2*) with similar trends observed in ICICLE-PD (**Figure 4.16a-b**). These specific circRNAs, or their host genes, have not previously been associated with PD.

The circRNA produced by *CCDC9* is located at 19:47264602-47264946:+ between exons 6 and 7 (*circCCDC9(6,7),* ENST00000221922.11). The protein product of *CCDC9* is believed to be a member of the exon junction complex involved in RNA splicing (Drew et al., 2021). The circRNA produced between exons 6 and 7 of *CCDC9* is reportedly reduced in gastric cancer whereby it acts as a sponge for miR-6792-3p (Luo et al., 2020). In stroke, an unspecified circRNA produced by *CCDC9* suppressed NOTCH signalling in murine models of ischaemia (Wu et al., 2020). Interestingly, alterations in NOTCH signalling have been linked to PD through the function of LRRK2 (Imai et al., 2015).

The circRNA produced by ESYT2 is located at 7:158759485-158764853:- between exons 9 and 13 (*circESYT2(9,10,11,12,13)*, ENST00000275418.13). *ESYT2* encodes extended synaptotagmin 2, a member of the E-Syt family, which are endoplasmic reticulum (ER) localised proteins involved in tethering the ER to the cellular plasma membrane (Giordano et al., 2013). There is suggestive evidence that circRNAs produced by *ESYT2* may be upregulated upon viral infection (Tagawa et al., 2021), in contrast to the decreased expression we observed in PD patients (**Figure 4.16a-b**).

The circRNA produced by *BMS1P1* is located at 10:46795805-46798168:+ between exons 6 and 8 (circ*BMS1P1(6,7,8),* ENST00000580094.5) or exons 5 and 8 (*circBMS1P1(5,6,7,8), ENST00000374336.5)* depending on the transcript. *BMS1P1* encodes a pseudogene of ribosome biogenesis factor pseudogene 1 (*BMS1*), and there are no known disease associations for circRNAs produced by this gene.

Several limitations of this differential expression analysis must be considered, some of which are shared with **Chapter 3**. CircRNAs were detected and quantified in whole blood, thus expression estimates reflect underlying blood cell composition. However, I sought to identify potential circRNA biomarkers in blood with minimal preprocessing steps. Furthermore, the PPMI cohort has increased statistical power to detect changes in circRNA expression due to the larger sample size (**Figure 4.16a**). As I used as many samples as were accessible at the time, sample sizes were not predetermined before analysis. However, the PPMI was used as the discovery cohort due to the greater number of participants. PPMI patients were not receiving pharmaceutical treatment for PD at the time of enrolment, in contrast with most PD patients in the ICICLE-PD study. I found no significant associations between treatment dosage and circRNA expression in ICICLE-PD cases (**Figure 4.17**), suggesting the impact of treatment is minimal. However, as only four PD patients were untreated in the ICICLE-PD participants, this comparison is underpowered to detect changes in circRNA expression due to the onset of treatment.

4.4.4 Reduced circular RNA detection in early-stage idiopathic PD

Intriguingly, I found a reduction in the number of circRNAs detected in PD patients compared to control samples (**Figure 4.15a**). Reduced circRNA abundance has been previously reported in the substantia nigra of PD patients (Hanan et al., 2020). As a primary site of PD-associated neurodegeneration, circRNA abundances may be influenced by cell loss. As a similar reduction in circRNA levels is observed in the blood, it suggests that decreased circRNA levels in substantia nigra may reflect disease pathology rather than cell loss. Reduced circRNA levels in

the blood of PD patients are in contrast however to the increased circRNA levels reported in the medial temporalis gyrus and amygdala (Hanan et al., 2020).

Variation in the detection of circRNAs will be driven by both technical and biological factors. I addressed biological and technical variation by analysing two independent cohorts, normalising circRNA levels for sequencing depth and adjusting for sequencing batch effects within each cohort. Some biological factors driving circRNA levels are known, like those involved in circRNA biogenesis and turnover (Yang et al., 2022). However, additional factors will undoubtedly be identified. Considering the current knowledge of circRNA regulation, investigating circRNA diversity in human samples is likely to be difficult. A more robust method may be to measure the global expression of circRNAs based on the individual change in circRNA expression between PD patients and controls. Indeed, preliminary work using this idea revealed a systemic bias in circRNA expression in PD patients (**Figure 4.16a**). The work presented in **Chapter 5** will investigate this expression bias.

4.4.5 Inconsistent findings with previously reported PD differentially expressed circRNAs

I aimed to replicate previously published studies reporting differentially expressed circRNAs in PD (Hanan et al., 2020; Kong et al., 2021; Ravanidis et al., 2021; Zhong et al., 2021; Xiao et al., 2022). However, identification of the specific reported circRNAs was not possible for several studies (Kong et al., 2021; Zhong et al., 2021; Xiao et al., 2022). Several circRNAs reported by Hanan et al and Ravanidis et al demonstrated consistent directions of change in PPMI and ICICLE-PD samples (**Figure 4.18b**). I identified altered expression of circRNAs produced by *DOP1B* and *INTS6L* in PPMI PD patients. These genes have not been functionally linked to PD. However, copy number variation in *DOP1B* has previously been linked to Alzheimer's disease (Swaminathan et al., 2012), while deletion of *INTS6L* leads to a cardiomyopathic phenotype (Spielmann et al., 2022).

Several factors may contribute to differences between the findings reported by Hanan et al and Ravanidis et al compared to the work presented here. Differences in the tissue studied will impact circRNA detection and quantification. Notably, 23 of the 30 previously reported circRNAs (20 from the substantia nigra and three detected in PBMCs) were not tested for differential expression in either PPMI or ICICLE-PD, either because they were not detected or not expressed above the filtering threshold. Notably, most of the brain expressed circRNAs reported by Hanan et al were not expressed in blood. The inability to detect these circRNAs at sufficient levels is consistent with the existence of brain-specific and -enriched circRNAs (Rybak-Wolf et al., 2015; You et al., 2015; Xia et al., 2017; Ji et al., 2019; Dong et al., 2023).

The exclusion of three circRNAs reported in Ravanidis et al and discrepancies in those successfully tested may reflect differences in the cellular composition and thus transcriptomic profiles between purified PBMCs (Ravanidis et al) and whole blood (this chapter) (He et al., 2019). CircRNA expression varies among blood cell types (Alhasan et al., 2016; Nicolet et al., 2018; Gaffo et al., 2019; Grassi et al., 2021), thus different cellular compositions will be reflected in altered circRNA expression profiles. Furthermore, methodological differences in the detection (qPCR vs RNAseq) and quantification of circRNA may impact differential expression analysis. For instance, in this chapter, I identified the optimum circRNA normalisation strategy (**Figure 4.6a, b**, yet other strategies could result in different circRNA expression estimates.

4.4.6 Circular RNA is of limited clinical utility as a predictor of early-stage idiopathic PD

CircRNA expression has emerged as a potential biomarker in a range of diseases (Verduci et al., 2021). Several studies have described the ability of blood circRNA expression to distinguish PD from controls at sensitivity and specificity values that warranted investigations in larger, independent cohorts (Ravanidis et al., 2021; Zhong et al., 2021; Xiao et al., 2022). Using the larger sample size of the PPMI dataset described in this chapter, the classification ability of circRNA was poor (AUC = 0.61, **Figure 4.20c**) compared to previously published estimates. Likewise, similar performance was also observed in an independent cohort (ICICLE-PD AUC = 0.59, **Figure 4.20c**). The disparity in the ability of circRNA to distinguish PD was evident compared to the use of gene expression in PPMI (circRNA AUC = 0.61, Gene AUC = 0.84, **Figure 4.20b-c**), although similar performance was seen in ICICLE-PD (circRNA AUC = 0.59, Gene AUC = 0.59, **Figure 4.20b-c**). Overall, the performance of circRNA expression to distinguish PD from controls was lower than previously described fluid biomarker candidates (**Section 1.2.4**) (Parnetti et al., 2019). The observed performance was also worse than a recent study reporting the ability of circRNA to distinguish PPMI PD patients from controls (AUC = 0.825) using a model trained on Parkinson's Disease Biomarker Program (PDBP) participants (Beric et al., 2024). Notably, Beric et al summarised circRNA expression to the gene level, suggesting this method of quantification may result in improved PD classification.

Similar to previous work assessing circRNA expression as a lung cancer biomarker (D'Ambrosi et al., 2023), combining circRNA expression with gene expression showed a small improvement in predictive ability over gene expression alone in both PPMI (AUC = 0.85 vs 0.84, **Figure 4.20b, c**) and ICICLE-PD (AUC = 0.60 vs 0.59, **Figure 4.20b, c**). Combining clinical, demographic, genetic and transcriptomic data in multi-modal machine learning models

has previously been shown to be predictive of PD status (Makarious et al., 2022). Based on my results, including circRNA expression may increase the predictive power of the multi-modality model described by Makarious and colleagues. Assuming sequencing of non-polyadenylated transcriptomes, the quantification of circRNA could be simultaneously generated. Despite this, the performance of the multi-modal model is heavily driven by scores from the University of Pennsylvania smell identification test (UPSIT) (Makarious et al., 2022). Indeed, previous work has shown that the standalone use of UPSIT scores can be highly predictive of PD status (Nalls et al., 2015), questioning the additional benefit of including transcriptomic data, and by extension, circRNA expression data.

4.5 Conclusion

Overall, the work presented in this chapter represents a comprehensive investigation into blood circRNA expression in the context of PD. By measuring circRNA expression in PD patients in the earliest stages after diagnosis, I aimed to identify circRNAs differentially expressed at a key point in disease progression. However, my investigation into the performance of circRNAs would suggest they are of limited clinical utility as a clinical PD biomarker when compared to existing biomarkers. Interestingly, their global expression appeared to be reduced in early-stage idiopathic PD patients. As such, reduced expression may be indicative of pathology underpinning PD and will be investigated in **Chapter 5**.

Chapter 5. Global reduction of circRNA expression in early-stage idiopathic PD

5.1 Background

In **Chapter 4**, I reported three differentially expressed circRNAs in early-stage idiopathic PD patients from the PPMI, yet these circRNAs did not reach statistical significance when tested in ICICLE-PD. Furthermore, circRNAs showed limited ability to distinguish PD patients from controls in these cohorts. However, there was a general reduction in circRNA expression in early-stage idiopathic PD patients in both cohorts.

CircRNA regulation leads to tissue, cell and development-stage specific expression (Salzman et al., 2013; Wu et al., 2022; Xia et al., 2017; Mellough et al., 2019). CircRNA abundance reflects the combined effects of biogenesis, stability, and degradation (**Sections 1.3.1, 1.3.3**). As changes in circRNA expression can highlight physiological changes, there is a growing recognition that global circRNA expression may be altered in disease. This view is particularly prevalent in cancer, with reduced circRNA expression changes reported in several cancer types (Vo et al., 2019; Hansen et al., 2022; C. Wang et al., 2022; Fuchs et al., 2023; Korsgaard et al., 2024). Reduced global circRNA abundances have also been observed in lupus (C.-X. Liu et al., 2019), chronic inflammatory skin conditions (Moldovan et al., 2019, 2021; Seeler et al., 2022; Guo et al., 2024), schizophrenia (Huang et al., 2023), Alzheimer's disease (Lo et al., 2020) and Parkinson's disease (Hanan et al., 2020).

In several cancer types, reduced circRNA levels are linked to increased tumour proliferation (Bachmayr-Heyda et al., 2015). In *MYCN*-amplified neuroblastoma tumours specifically, reduced circRNA levels have been attributed to increased DHX9 activity (Fuchs et al., 2023). When DHX9, an RNA helicase (Koh et al., 2014), is depleted, increased circRNA levels are reported, potentially by IR*Alu* unwinding (**Section 1.3.1**) (Aktaş et al., 2017). Aktaş et al also demonstrated that ADAR1 and DHX9 depletion increased circRNA levels more than individual depletion (Aktaş et al., 2017). ADAR1 belongs to the adenosine deaminase acting on RNA (ADAR) family. ADAR family members play key roles in a post-transcriptional RNA modification in which adenosine (A) is converted to inosine (I), termed A-I editing (Hogg et al., 2011). Humans possess three ADAR family members: ADAR1 and ADAR2 are enzymatically active and catalyse A-I editing while ADAR3 acts as a brain-specific negative regulator of A-I editing (Chen et al., 2000). Intramolecular inverted repeats formed across pairs of transcribed *Alu* elements can produce dsRNA structures which act as binding sites for ADAR proteins (Athanasiadis et al., 2004). In humans, most A-I editing occurs in dsRNA structures formed by intramolecular inverted repeats across pairs of highly repetitive *Alu* elements (IR*Alus*), which act as binding sites for ADAR proteins (Athanasiadis et al., 2004; Levanon et al., 2004). The A-I editing of *Alu* dsRNA structures in this manner is primarily driven by

ADAR1, which leads to widespread A-I editing at low levels (<1%) (Bazak et al., 2014). This form of A-I editing is thought to label these dsRNA structures as 'self', a lack of this A-I editing leads to aberrant immune activation (Mannion et al., 2014; Liddicoat et al., 2015). ADAR2 is associated with high levels of editing of specific sites (Hoopengardner et al., 2003), usually in coding regions leading to potential protein-coding changes (Pullirsch & Jantsch, 2010; Tan et al., 2017).

IR*Alus* play a role in circRNA biogenesis (Zhang et al., 2014) (detailed in **Section 1.3.1**). Alterations to dsRNA structure due to A-I editing could impact circRNA biogenesis (Bass & Weintraub, 1988). qPCR analysis identified numerous circRNAs upregulated upon *ADAR* (ADAR1) knockdown (Ivanov et al., 2015; Rybak-Wolf et al., 2015), although this upregulation appeared to be circRNA specific (Aktaş et al., 2017). CircRNA quantification via RNA-seq identified a global increase in circRNA expression after dual *ADAR* and *ADARB2* (ADAR2) knockdown (Ivanov et al., 2015). *ADAR* knockdown both promoted and repressed specific circRNA biogenesis (Shen et al., 2022), with similar, yet tissue specific effects shown in mice (Kapoor et al., 2020). Whilst the exact mechanism for the bidirectional regulation of circRNA biogenesis by ADAR1 has not been identified, A-I editing may both stabilise or destabilise dsRNA structures formed by IR*Alus* depending on the paired bases (Shen et al., 2022). Additionally, A-I editing of RNA binding protein (RBP) binding sites may enhance or diminish the binding of RBPs associated with circRNA biogenesis (Tang et al., 2020; Shen et al., 2022). *ADARB1* knockdown, however, is consistently associated with mostly repressing circRNA biogenesis (Shen et al., 2022; Kokot et al., 2022).

CircRNA levels decrease globally upon poly(I:C) (Polyinosinic:polycytidylic acid) treatment, which mimics dsRNA present during viral infections (Li et al., 2017; C.-X. Liu et al., 2019). Pathogenic dsRNA is detected by dsRNA-binding sensors as part of the innate antiviral immune response (Hur, 2019). These sensors typically converge upon downstream activation of type I interferon responses and induce interferon stimulated genes (Schlee & Hartmann, 2016). Regulation of circRNA levels is linked to the functionality of the dsRNA sensor protein kinase R (PKR). CircRNAs contain short hairpin dsRNA secondary structures that bind to and inhibit PKR (C.-X. Liu et al., 2019; Liu et al., 2022; Guo et al., 2024). RNase L-mediated circRNA degradation augments PKR phosphorylation and activation in response to dsRNA (C.-X. Liu et al., 2019). PKR phosphorylation leads to the activation of the integrated stress response (ISR) through the phosphorylation of the alpha subunit of eIF2 (eIF2α) (Donnelly et al., 2013). ISR activation reduces global protein synthesis and upregulates specific stress-related genes (Pakos-Zebrucka et al., 2016). ISR responses are mediated by translating specific mRNAs encoding

proteins such as ATF3, ATF4, ATF5, CHOP and GADD34 (Pakos-Zebrucka et al., 2016; Costa-Mattioli & Walter, 2020). PKR activation is also linked to downstream inflammatory and apoptotic signalling (Gil & Esteban, 2000; Lu et al., 2012).

Reduced circRNA levels have been observed in the substantia nigra of PD patients compared to controls (Hanan et al., 2020). While circRNA levels were inversely correlated to global A-I editing levels, no significant difference in global A-I editing between PD patients and controls was found in the substantia nigra (Hanan et al., 2020).

In **Chapter 3**, I reported a decrease in global circRNA expression in the blood of early-stage idiopathic PD patients in the PPMI and ICICLE-PD cohorts. This preliminary finding raises several questions: Is this decrease a true biological effect or a technical artefact? If biological, is the reduction specific to circRNA or a general decrease in RNA expression? If specific can global circRNA expression be used as a diagnostic biomarker for PD? Why is circRNA reduced? In this chapter, I use previously generated circRNA, linear RNA and gene expression data (**Chapters 3 and 4**) of early-stage idiopathic PD patients and controls to address these questions.

5.2 Methods

5.2.1 Cohorts

Participants from the PPMI and ICICLE-PD studies included in this chapter are the same as those outlined in **Chapters 3 and 4**. Selection criteria were described in **Section 2.2**. Briefly, PD patients were recently diagnosed (<13 months) and have no known underlying cause (idiopathic). Controls were matched for age, and sex and were free of neurological disorders.

5.2.2 RNA sequencing

Blood samples from participants underwent RNA sequencing as described in **Section 2.3**.

5.2.3 Differential expression

This chapter uses gene expression data generated in **Chapter 3** and circRNA expression data generated in **Chapter 4**. CircRNAs were identified from reads containing a back-spliced junction (BSJ), with the number of BSJ reads used as an absolute count of circRNA expression. The cognate linear RNA expression is estimated as the number of reads consistent with forward-spliced junctions (FSJs). Further information detailing BSJ and FSJ detection is outlined in **Chapter 4**. In total, 403 (PPMI) and 457 (ICICLE-PD) BSJ loci were abundantly expressed and described here.

Differential expression was performed using DESeq2 as described in **Section 2.8**.

Differential expression testing of circRNAs was also performed using limma v3.56.2 (Ritchie et al., 2015) as described in **Section 2.8**. For this approach, circRNA expression was normalised using the Trimmed Mean of M-values (TMM) approach (Robinson & Oshlack, 2010) as implemented in edgeR v3.42.4 (Robinson et al., 2010). Like the method used with DESeq2, normalisation factors and library sizes for each sample were determined from gene expression data to account for differences in circRNA detection, similar to the implementation in CIRIquant (Zhang et al., 2020). Gene expression was estimated as described in **Section 2.8**. Variation in transcript length was accounted for as recommended for downstream analysis with limma (*countsFromAbundance = "lengthScaledTPM")* (Soneson et al., 2015). Lowly expressed genes were removed using the *filterByExpr* function with default settings, retaining 16,930 (PPMI) and 16,374 (ICICLE-PD) genes respectively. Model designs accounting for sources of biological technical variation were the same as those described in **Chapters 3 and 4**.

5.2.4 Identifying global changes in RNA expression

Global changes in RNA expression were quantified using the methodology described in **Section 2.10.3**. In this chapter, the $\log_2$ fold change threshold was set at >0.1/<-0.1 unless otherwise stated in the text.

5.2.5 Blood cell deconvolution

RNA expression in bulk blood samples reflects the combination of different blood cell populations. As such, RNA expression can be influenced by heterogeneity in the relative proportions of each blood cell population (Monaco et al., 2019). Blood cell counts for ICICLE-PD participants were not available. Blood cell counts in PPMI participants were measured at an initial screening visit but not at the baseline visit. Thus, to include cell count as a covariate in regression analysis, proportions were estimated using CIBERSORTx (Newman et al., 2019). CIBERSORTx has shown consistent performance when benchmarked on RNAseq samples (Avila Cobos et al., 2020; Jin & Liu, 2021; Nadel et al., 2021; Sutton et al., 2022). The protocol used was based on recommendations from the CIBERSORTx developers (Steen et al., 2020). To create a mixture matrix, gene expression values were normalised for sequencing depth and gene length based on transcripts per million (TPM). Genes with a TPM >0.5 in at least half the samples in each cohort were retained. A reference cell proportion matrix was created based on the LM22 dataset (Newman et al., 2015). CIBERSORTx implements a B-mode batch correction when the reference matrix (microarray) and mixture matrix (RNA-seq) are from different quantification platforms (Newman et al., 2019). Quantile normalisation was disabled as recommended for RNAseq data. Differences in the proportion of each blood cell type were assessed using a Wilcoxon rank-sum test.

5.2.6 Creating a classifier based on reduced circular RNA expression

The mean variance-stabilised transformed expression (Anders & Huber, 2010) of each abundant circRNA included in differential expression analysis was calculated using PPMI control samples. Each circRNA was defined depending on whether it had higher ($\log_2$ fold change >0) or lower ($\log_2$ fold change <0) expression in PPMI cases relative to controls. circRNAs were scored per sample on whether the directional deviation from the mean (i.e., over- or under-expression) was consistent with the direction observed in the PPMI. When the directions agreed, a score of 1 was given for that circRNA and a score of 0 when the directions disagreed. Scores were summed per sample and the ability of this sum to discriminate between

cases and controls was assessed by constructing receiver operator characteristic (ROC) curves and calculating the area under the ROC curves (AUC) using pROC v1.18.4 (Robin et al., 2011).

5.2.7 Circular RNA regulators

Factors associated with changes in the expression of multiple circRNAs were identified in the literature (regulators and evidence are shown in **Table 5.1**). To assess differences in circRNA regulator expression between PD patients and controls, I retrieved summary statistics from **Chapter 3**. FDR values (Benjamini-Hochberg procedure) were recalculated in PPMI based on just the retrieved circRNA regulator results. For validation in ICICLE-PD, FDR values were calculated based on significantly differentially expressed circRNA regulators in PPMI.

Gene ID	Gene symbol	Effect on circRNA levels	Evidence
ENSG00000160710	*ADAR*	Conflicting reports	(Ivanov et al., 2015; Rybak-Wolf et al., 2015; Shen et al., 2022)
ENSG00000197381	*ADARB1*	Decrease	(Shen et al., 2022; Kokot et al., 2022)
ENSG00000135829	*DHX9*	Decrease	(Aktaş et al., 2017)
ENSG00000089280	*FUS*	Bidirectional	(Errichelli et al., 2017)
ENSG00000092199	*HNRNPC*	Bidirectional	(Di Liddo et al., 2019)
ENSG00000129351	*ILF3*	Increase	(Li et al., 2017)
ENSG00000152601	*MBNL1*	Increase	(Ashwal-Fluss et al., 2014)
ENSG00000104967	*NOVA2*	Increase	(Knupp et al., 2021)
ENSG00000167005	*NUDT21*	Increase	(Li et al., 2020)
ENSG00000112531	*QKI*	Increase	(Conn et al., 2015)
ENSG00000135828	*RNASEL*	Decrease	(C.-X. Liu et al., 2019)
ENSG00000090905	*TNRC6A*	Decrease	(Jia et al., 2019)
ENSG00000100354	*TNRC6B*	Decrease	(Jia et al., 2019)
ENSG00000078687	*TNRC6C*	Decrease	(Jia et al., 2019)

Table 5.1. Circular RNA regulators. For each regulator, the Ensembl gene ID, gene symbol and the publications which describe the evidence for the regulation of circular RNA are given.

5.2.8 RNA editing

Reads were aligned against the Ensembl GRCh38 genome (release 101) using STAR v.2.7.10a (Dobin et al., 2013). Aligned reads with >95% of matched bases between read and reference were retained as recommended for measuring RNA editing (*--outFilterMatchNminOverLread 0.95*) (Roth et al., 2019). Uniquely mapped reads were kept (*-q 255*) using samtools v1.17 with optical and sequencing duplicate reads removed using sambamba markdup v1.0.0 (Tarasov et al., 2015). RNA editing levels (RNA editing index) were measured using RNAEditingIndexer

(Roth et al., 2019). RNA editing events were inferred based on base substitutions between the sequenced bases and the genomic reference. Base substitutions at loci containing common single nucleotide polymorphisms (based on dbSNP150) or sequencing errors (PHRED quality ≤30 or unknown base called) were excluded by RNAEditingIndexer. The RNA editing index for each base mismatch is calculated based on the ratio of edited bases over the total coverage of that base. A-I editing was inferred as A-G mismatches as inosine is recognised as a guanine during sequencing (Levanon et al., 2004). As such A-I editing levels are represented by the A>G index which is the number of A-G substitutions divided by all A-A matches and A-G substitutions multiplied by 100. The signal:noise ratio was estimated as the A>G index divided by the C>T index as recommended (Roth et al., 2019)

RNA editing was measured in several defined regions: A) Human *Alu* regions as included in RNAEditingIndexer based on hg38 repetitive elements from RepeatMasker (Tarailo-Graovac & Chen, 2009). B) the genomic span within all BSJs of circRNAs included in differential expression analysis of PPMI and ICICLE-PD datasets (labelled *circRNAs* in **Figure 5.10).** C) 1500bp upstream of the BSJ start and 1500bp downstream of the BSJ end given enriched A-I editing 1500bp around BSJs (Ivanov et al., 2015) (labelled *flanking circRNAs* in **Figure 5.10**). Five PPMI samples were excluded as RNA editing levels could not be measured in all three regions.

5.2.9 Gene set enrichment analysis

Summary statistics of gene differential expression results between early-stage idiopathic PD patients and controls were generated in **Chapter 3**. Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) was performed using clusterProfiler v4.8.3 (Wu et al., 2021). Genes were sorted in decreasing order based on fold change (PD vs controls). Genes were grouped based on Hallmark gene sets (Liberzon et al., 2015).

To identify putative downstream regulatory targets, sets of genes with common transcription factor motifs were collected from MSigDB v2023.2.Hs (Liberzon et al., 2011). These gene sets (total = 1115) comprised 610 gene sets from (Xie et al., 2005) and TRANSFAC v.7.4 (Matys et al., 2006) based on motifs within a 4kb window centred on each gene's respective transcription start site. A further 505 gene sets consisted of genes predicted to contain the transcription factor binding sequencing in the promotor region (-1000bp and +100bp around the transcription start site) (Kolmykov et al., 2021).

5.3 Results

5.3.1 Reduced circular RNA expression in early-stage idiopathic PD

In **Chapter 4**, I tested the differential expression of circRNAs in early-stage idiopathic PD patients compared to controls. As shown in **Chapter 4**, BSJ expression was globally reduced in early-stage idiopathic PD patients based on the observed fold changes (PPMI imbalance = 0.09, *P*-value < 2.2×10^{-16}, ICICLE-PD imbalance = 0.29, *P*-value = 2.3×10^{-10}) (**Figure 5.1a, b**).

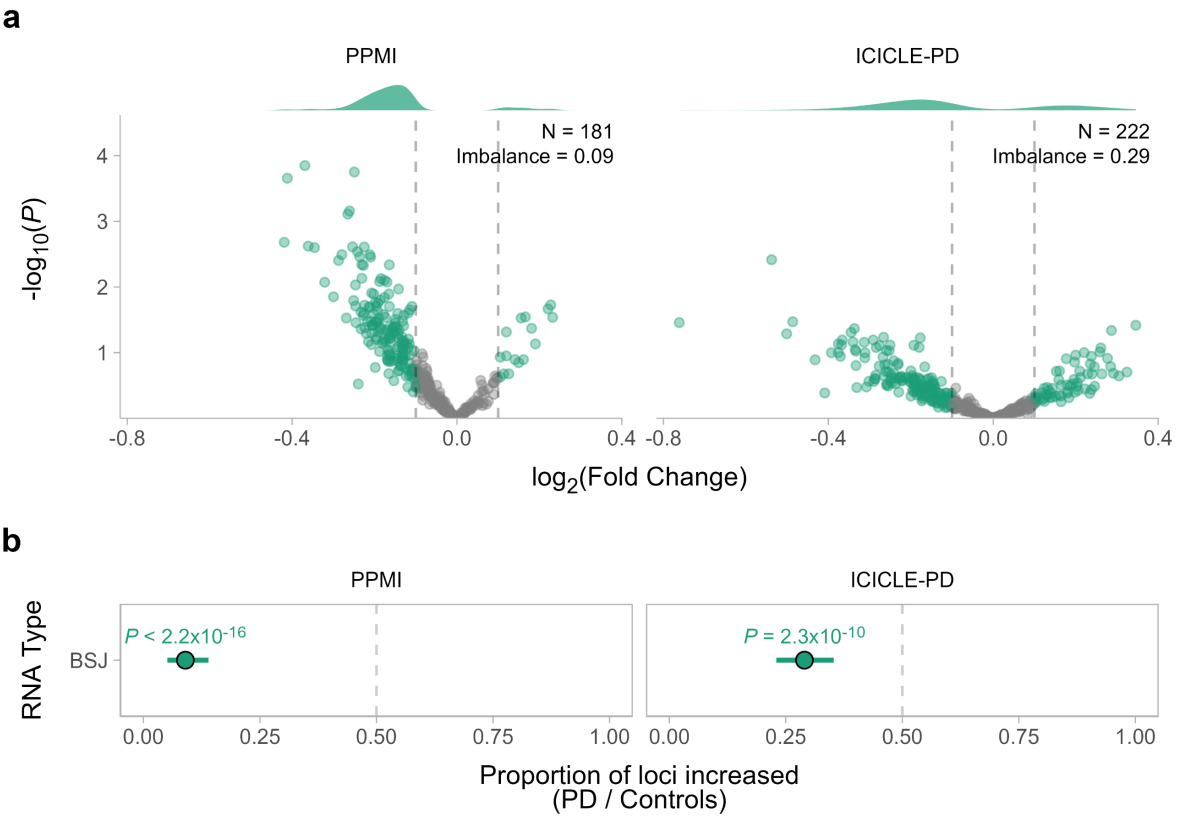


Figure 5.1. Globally reduced circular RNA expression in early-stage idiopathic PD patients. (a) Volcano plots showing the differential expression of back-spliced junctions (BSJ) in early-stage idiopathic PD patients and controls. Coloured points indicate BSJs above the fold change threshold ($\log_2$ fold change >0.1/<-0.1, indicated by the dashed grey lines). N gives the number of BSJs above or below the fold change thresholds. Density plots show the distribution of BSJs above or below the fold change thresholds. **(b)** Expression imbalances were identified based on the proportion of BSJs increased in PD patients relative to controls. *P*-values obtained from a two-sided exact binomial test. Error bars show the 95% confidence interval of the imbalance estimate.

Next, I carried out a series of tests to assess the possibility that the reduction of circRNA expression in PD samples was a technical artefact. A comparison of values used to normalise

circRNA expression (DESeq2 size factors, **Section 5.2.3**) revealed no significant differences between PD and control samples (*P*-values > 0.05, Wilcoxon rank-sum test) (**Figure 5.2a**). I also compared circRNA expression between PD patients and controls with alternative count normalisation and differential expression methods (TMM normalisation with testing using limma, see **Section 5.2.3**). Using this method, I again observed reduced BSJ expression in the PPMI (imbalance = 0.13, *P*-value < $2.2x10^{-16}$) and ICICLE-PD (imbalance = 0.12, *P*-value < $2.2x10^{-16}$) cohorts (**Figure 5.2b, c**).

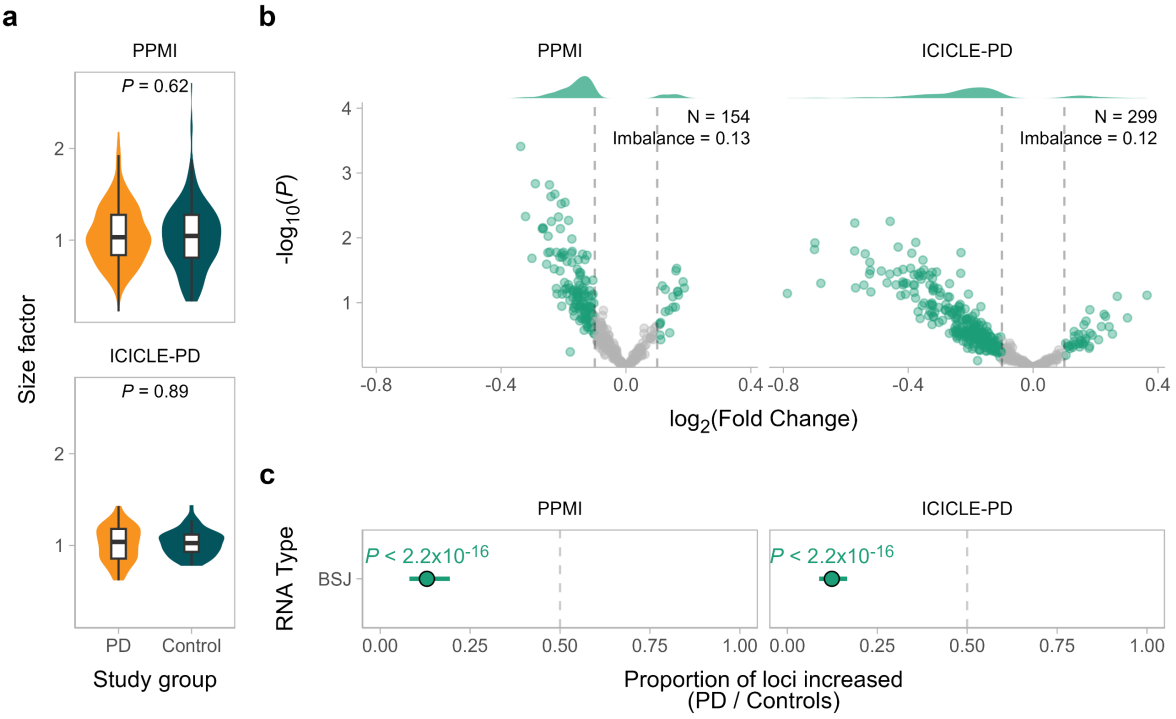


Figure 5.2. Reduced circular RNA expression is not an artefact of expression normalisation methods. (a) Comparison of DESeq2 size factors between PD and control samples. *P*-values derived from a Wilcoxon rank-sum test. **(b)** Volcano plots showing the results produced by Limma comparing back-spliced junction (BSJ) expression in PD relative to control samples. Coloured points indicate BSJs above the fold change threshold (log2 fold change >0.1/<-0.1, indicated by the dashed grey lines). N shows the number of BSJs above or below the fold change thresholds. Density plots show the distribution of BSJs above or below the fold change thresholds. **(c)** Expression imbalances were identified based on the proportion of BSJs increased in PD patients relative to controls based on differential expression results produced by limma. *P*-values obtained from a two-sided exact binomial test. Error bars show the 95% confidence interval of the imbalance estimate.

5.3.2 Comparing circular RNA expression to corresponding linear RNA expression and general gene expression

To examine whether a reduction in circRNA expression is reflective of a general reduction in RNA expression, I also compared the expression of the corresponding forward splice junctions (FSJs), in addition to leveraging the output of summary statistics regarding the differential expression of genes (**Chapter 3**) (**Figure 5.3**).

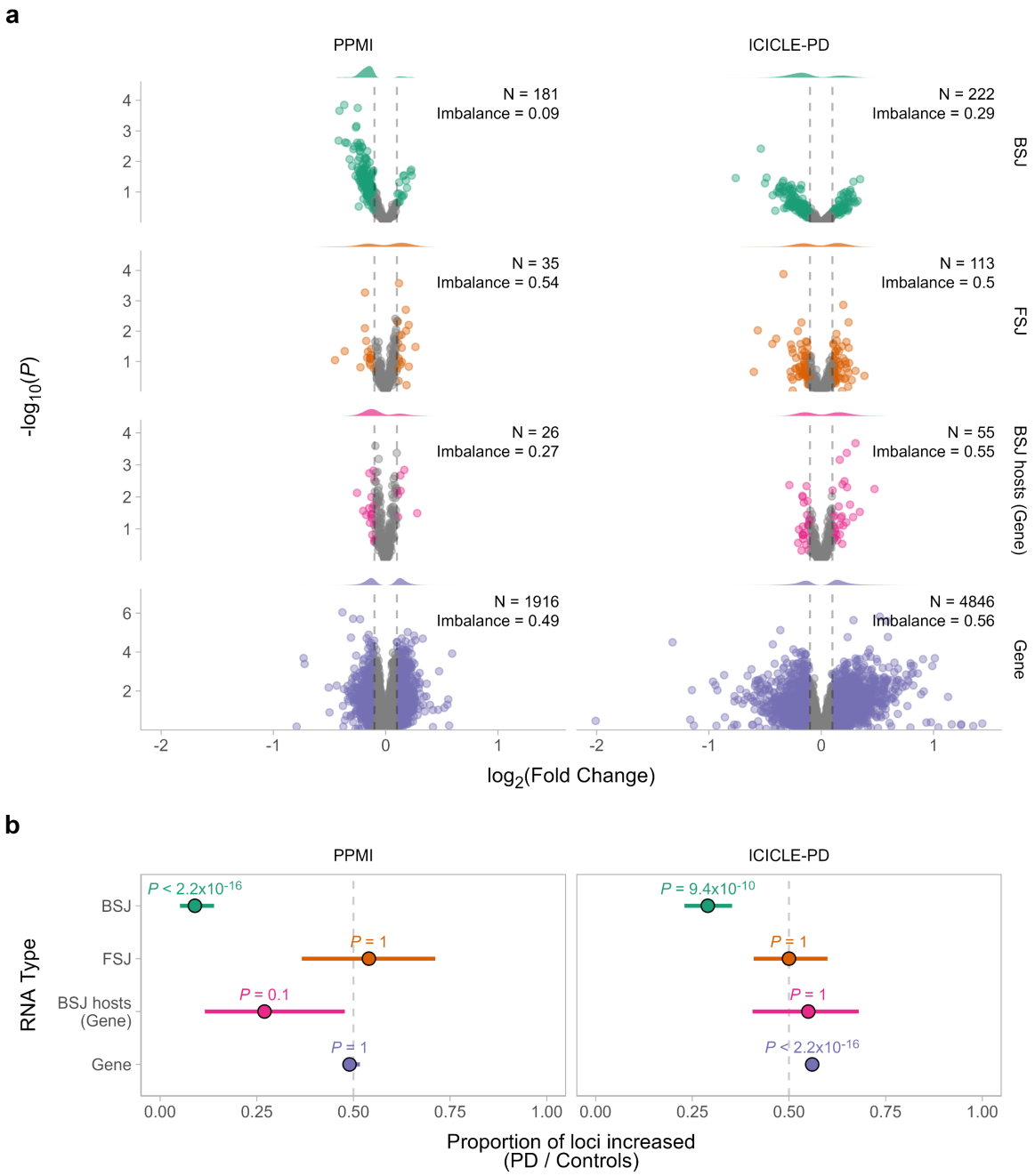


Figure 5.3. Global changes in RNA expression. (a) Volcano plots showing the differential expression of RNA types between PD and controls in PPMI and ICICLE-PD. Differential expression was carried out on back-spliced junctions (BSJs), forward-spliced junctions (FSJ), expression of genes which host an abundant BSJ (BSJ hosts) and expression of all genes (Gene). Coloured points indicate RNAs above the fold change threshold (log2 fold change >0.1 or <-0.1, indicated by the dashed grey lines). N shows the number of loci above or below the fold change thresholds. Density plots show the distribution of loci above or below the fold change thresholds. **(b)** Expression imbalances based on the fold changes shown in (**a**). *P*-values obtained from a two-sided exact binomial test corrected for multiple testing (Bonferroni correction, four tests). Error bars show the 95% confidence interval of the imbalance estimate.

FSJ counts provide the best approximation of the linear RNA as gene expression estimates will inevitably include circRNA-derived reads that do not span the BSJ and thus cannot be

distinguished. Crucially, a reduction was not evident when examining the fold changes of FSJs (PPMI imbalance = 0.54, Bonferroni corrected *P*-value = 1.0, ICICLE-PD imbalance = 0.50, Bonferroni corrected *P*-value = 1.0, four tests, Exact binomial test) (**Figure 5.3**). A comparison of fold changes based only on genes that host circRNAs identified no reduction (PPMI imbalance = 0.27, Bonferroni corrected *P*-value = 0.1, ICICLE-PD imbalance = 0.54, Bonferroni corrected *P*-value = 1.0, four tests) (**Figure 5.3**). Global gene expression also showed no reduction in PPMI (imbalance = 0.49, Bonferroni corrected *P*-value = 1.0, four tests), yet was significantly increased in ICICLE-PD (imbalance = 0.56, Bonferroni corrected *P*-value = 3.6×10^{-18}, four tests) (**Figure 5.3**). Together, these results suggest that reduced circRNA expression in early-stage idiopathic PD is not due to a general decrease in RNA expression.

So far, I have demonstrated globally reduced circRNA expression based on circRNAs with a difference in expression between PD and controls above a threshold (absolute $\log_2$ fold change >0.1). However, in some cases, this may result in a low number of tested loci. For example, in the PPMI cohort, 26 genes that hosted abundant circRNAs passed the fold change threshold (**Figure 5.3**). As a sensitivity analysis, I also calculated RNA expression imbalances with a relaxed fold change threshold (**Figure 5.4a**). Here, I classified circRNAs with a recorded log2 fold change >0 as *increased*, and those with a log2 fold change <0 as *decreased*. BSJ expression was again globally reduced in PPMI (imbalance = 0.19, Bonferroni corrected *P*-value < 2.2×10^{-16}, four tests) and ICICLE-PD PD patients (imbalance = 0.36, Bonferroni corrected *P*-value = 2.1×10^{-8}, four tests). There were no significant reductions when looking at FSJ expression or BSJ host gene expression (Bonferroni corrected *P*-values > 0.05, four tests). All gene expression was significantly reduced in PPMI PD patients (imbalance = 0.46, Bonferroni corrected *P*-value < 2.2×10^{-16}, four tests) however no significant reduction was identified in ICICLE-PD PD patients (imbalance = 0.50, Bonferroni corrected *P*-value = 0.8, four tests). Furthermore, a comparison of the absolute normalised expression of BSJs between PD and control individuals showed a reduction in PD patients in both PPMI and ICICLE-PD cohorts (**Figure 5.4c**). Subsequently, there were no significant differences (*P*-value >0.05, Wilcoxon rank-sum test) in the expression of FSJs (**Figure 5.4d**), circRNA host genes (**Figure 5.4e**) or all genes (**Figure 5.4f**) in PD patients.

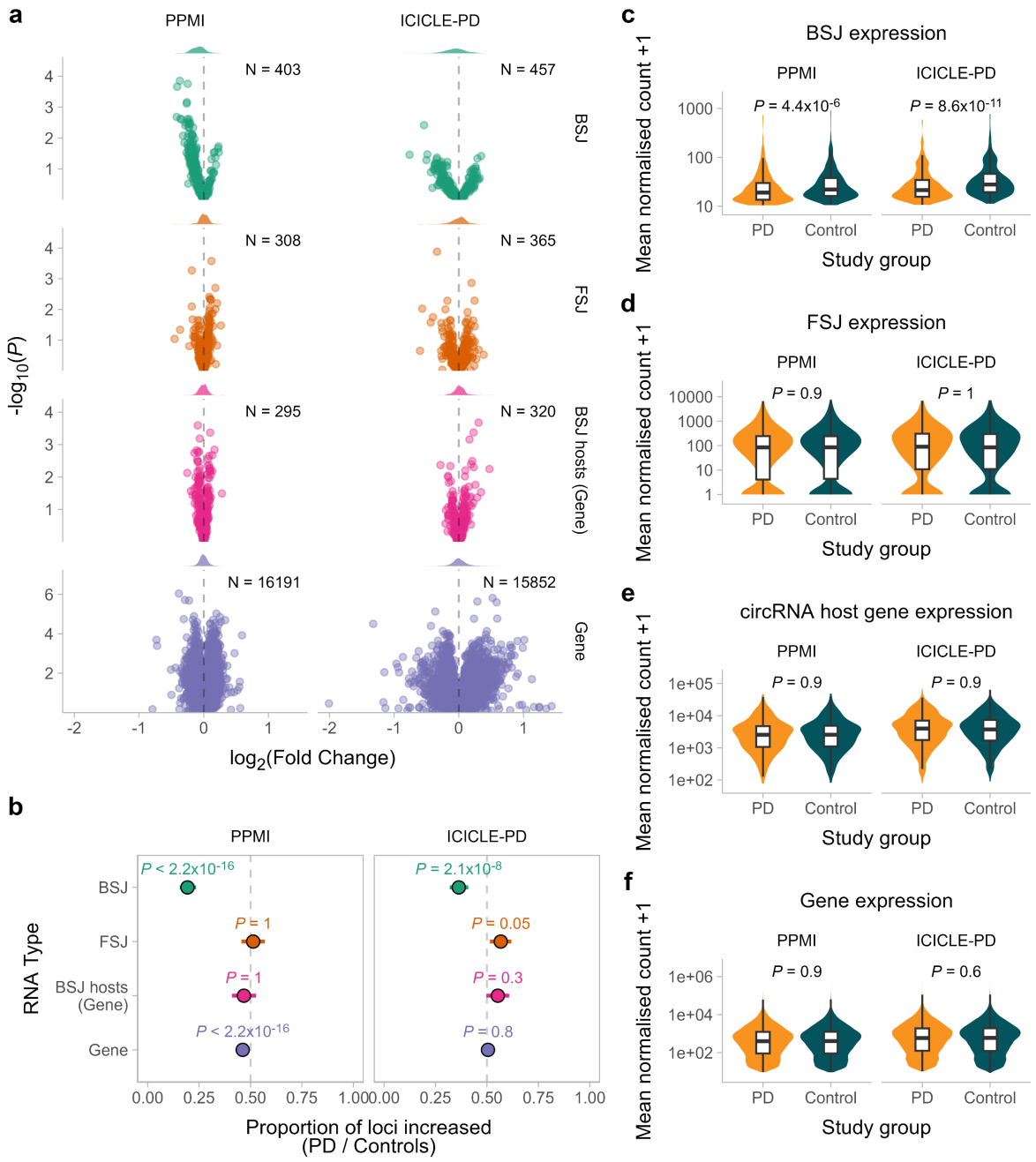


Figure 5.4. Reduced circular RNA expression in PD irrespective of fold change threshold. (a) Volcano plots showing the differential expression of back-spliced junctions (BSJs), forward-spliced junctions (FSJ), genes which host an abundant BSJ (BSJ hosts) and all genes (Gene) between PD patients and controls in PPMI and ICICLE-PD. N shows the number of loci included when estimating the expression imbalance. Density plots show the fold change distribution of loci. **(b)** Imbalances were identified based on the proportion of loci increased in PD relative to controls based on differential expression testing of each RNA type. *P*-values obtained from a two-sided exact binomial test corrected for multiple testing (Bonferroni correction, four tests). Error bars show the 95% confidence interval of the imbalance estimate. **(c-f)** Comparison of mean normalised expression of BSJs (**c**), FSJs (**d**), genes which host an abundant BSJ (**e**) and all genes (**f**). Means were calculated on DESeq2 normalised expression for each study group separately. Group differences were assessed with a Wilcoxon-rank sum test.

To test whether a global reduction in expression is specific to BSJs, I compared fold changes between BSJ expression and their flanking FSJs. For a subset of junctions, this was not possible as no corresponding FSJs were detected (PPMI = 95/403 or 23.6%, ICICLE-PD = 92/457 or 20.1%). BSJ fold changes were consistently lower, albeit generally modest, across both cohorts (PPMI *P*-value = 2.9×10^{-34}, ICICLE-PD *P*-value = $1.6\text{x}10^{-9}$, Wilcoxon rank-sum test). I observed a weak correlation in fold change differences between BSJs and corresponding FSJs, in which differences in FSJ expression explained 7-8% of the variation in differences of BSJ expression (**Figure 5.5**). These findings suggest that reduced circRNA expression is not driven by comparative reductions in the cognate linear RNA expression.

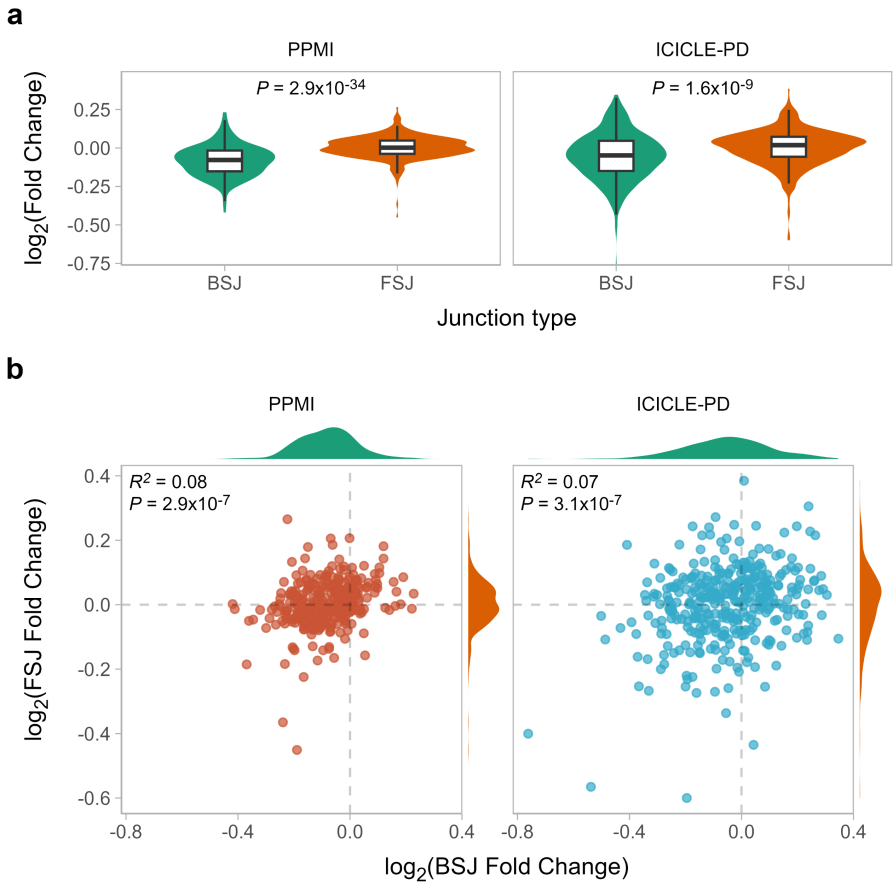


Figure 5.5. Reduced circRNA expression in early-stage idiopathic PD is not driven by concomitant changes in linear RNA expression. (a) Comparing the expression of all BSJs included in differential expression testing between PD and controls in PPMI and ICICLE-PD. *P*-values are derived from a Wilcoxon rank-sum test comparing BSJ expression between PD and controls. **(b)** The relationship between BSJ and FSJ expression fold changes in both PPMI and ICICLE-PD. The strength of the relationship is reported as R^2. Density plots show the distribution of the fold changes for each respective junction type.

5.3.3 Assessing reduced circular RNA expression as a predictor of early-stage idiopathic PD

Given the distinctive pattern of global circRNA reduction observed in early-stage idiopathic PD, I next sought to establish whether this reduction could distinguish PD patients from controls. For each sample, I calculated an expression score based on the expression of abundant circRNA (**Section 5.2.3**). I then evaluated the potential of this expression score (termed the imbalance model) to discriminate between early-stage idiopathic PD patients and controls in PPMI and ICICLE-PD participants. In the PPMI cohorts, the imbalance model had an AUC of 0.59 (95% CI = 0.54-0.65), while in ICICLE-PD participants, the imbalance model also had an AUC of 0.59 (95% CI = 0.48-0.71) (**Figure 5.6a, b**). As a comparison, the respective gene models (**Chapter 3**) are also shown in **Figure 5.6a, b**.

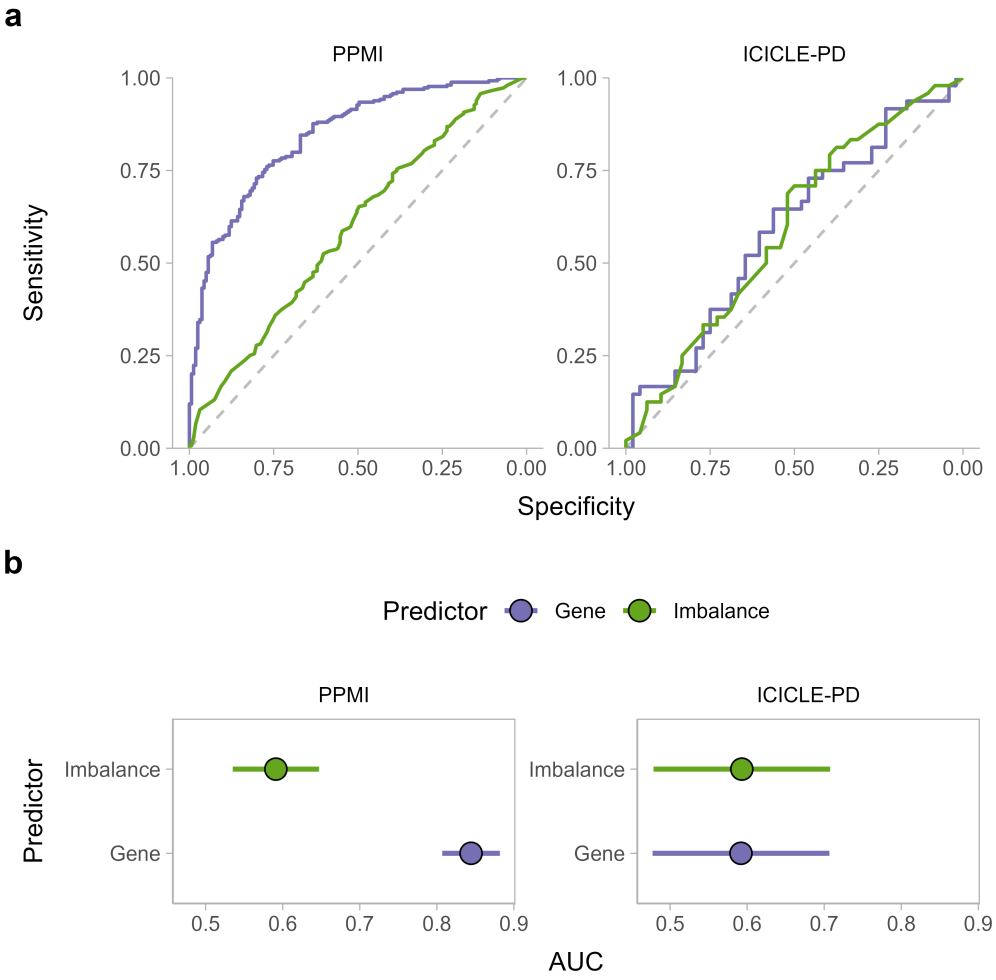


Figure 5.6. Performance of a classification model based on reduced circular RNA expression. ROC curves (**a**) and the corresponding AUCs (**b**) for the Imbalance (**Section 5.2.6**) and Gene models (**Chapter 3**). Error bars denote the 95% confidence intervals of the AUC.

5.3.4 Blood cell deconvolution

Given that blood cell lineages exhibit characteristic circRNA expression profiles (Grassi et al., 2021), I hypothesised that altered proportions of blood cells in PD patients could influence global circRNA expression. In the absence of available blood cell counts for PPMI and ICICLE-PD participants (**Section 5.2.4**), I estimated cell-type proportions using CIBERSORTx (Newman et al., 2019). Reassuringly, neutrophils were generally the largest detected contributors to blood composition (median proportion [IQR], PPMI = 0.45 [0.11], ICICLE-PD = 0.42 [0.15], followed by various subtypes of lymphocytes (Coffelt et al., 2016).

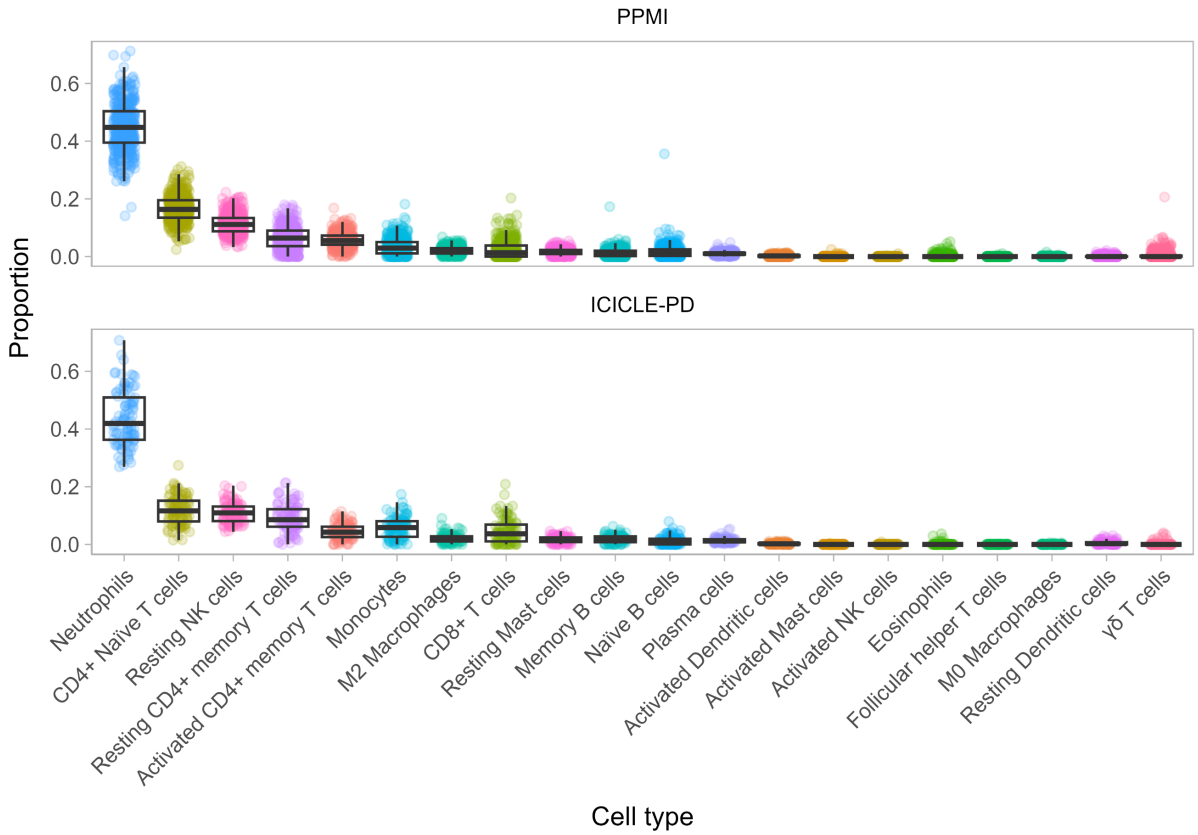


Figure 5.7. Estimated blood cell proportions. Proportions of specific leukocyte populations estimated by CIBERSORTx in PPMI and ICICLE-PD samples. Proportions deconvoluted based on the LM22 signature matrix (Newman et al., 2015).

I subsequently limited my analyses to the most abundant blood cell types (median proportion >0.01) leaving a set of 11 cell types (Neutrophils, CD4+ Naïve T cells, Resting NK cells, Resting CD4+ memory T cells, Activated CD4+ memory T cells, Monocytes, M2 Macrophages, Resting Mast cells, CD8+ T cells, Memory B cells and Plasma cells). In PPMI participants, Neutrophil (P = 0.02, Wilcoxon rank-sum test) and resting CD4+ memory T cell

(P = 0.007, Wilcoxon rank-sum test) proportions were significantly altered in PD patients yet were not significant after multiple testing correction (Bonferroni P-values > 0.05, 11 tests) (**Figure 5.8**). Furthermore, there were no significant differences (all P-values > 0.05, Wilcoxon rank-sum test) in any blood cell type proportions in ICICLE-PD participants (**Figure 5.8**). These findings suggest that there were no consistent differences in the estimated blood cell proportions across PD patients and controls in both PPMI and ICICLE-PD cohorts, and thus are unlikely to be contributing to a global reduction in circRNA expression.

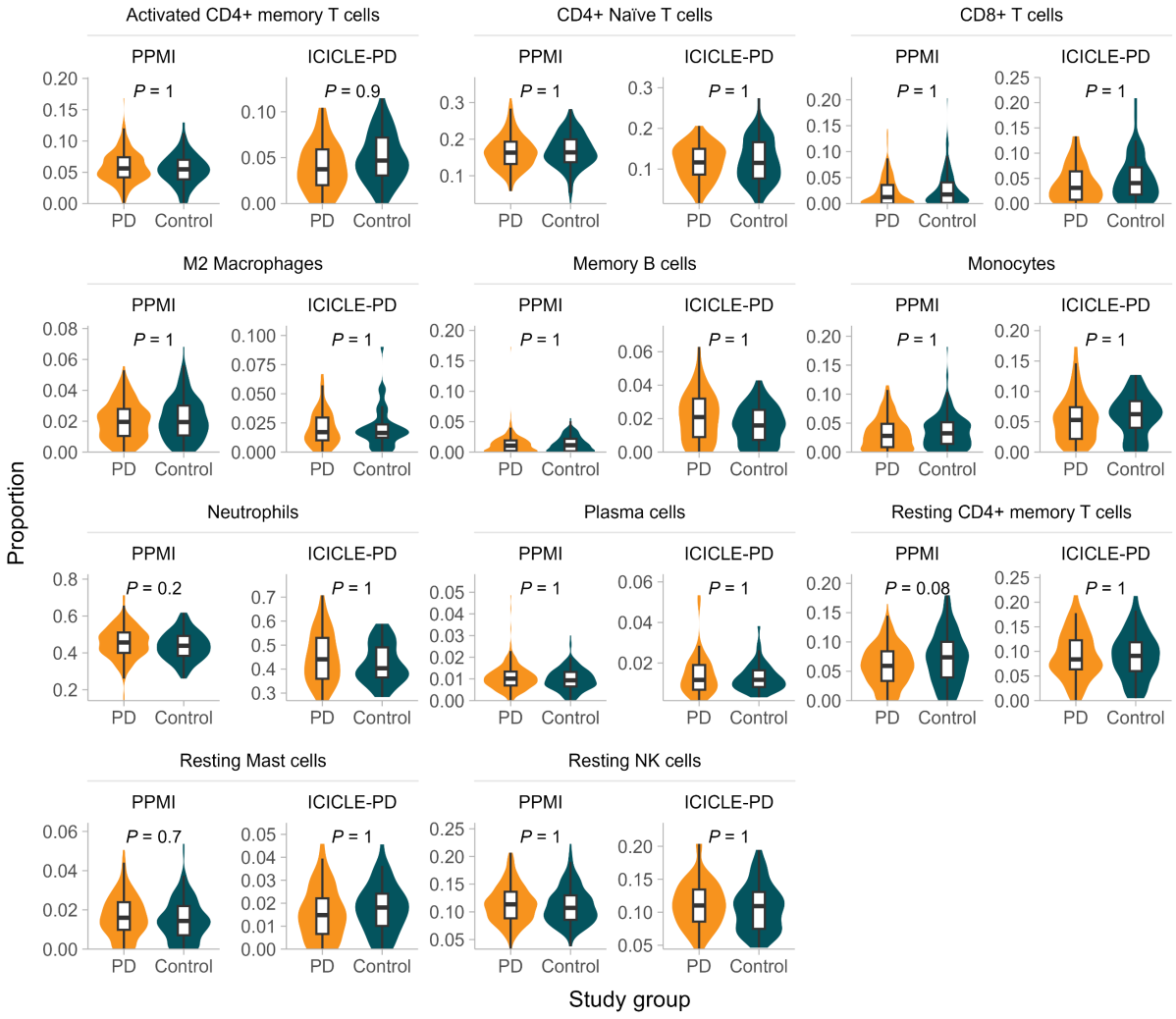


Figure 5.8. Differences in blood cell proportions between idiopathic PD patients and controls. Blood cell proportions were estimated using CIBERSORTx. Group differences were assessed using a Wilcoxon rank-sum test. *P*-values adjusted for 11 tests using Bonferroni correction.

5.3.5 *Identifying differentially expressed circular RNA regulators*

CircRNA biogenesis and degradation are linked to several *cis*- and *trans*-acting factors (Kristensen et al., 2019). I therefore explored whether differential expression of previously reported circRNA regulators (Ashwal-Fluss et al., 2014; Conn et al., 2015; Ivanov et al., 2015; Aktaş et al., 2017; Errichelli et al., 2017; Li et al., 2017; Di Liddo et al., 2019; Jia et al., 2019; C.-X. Liu et al., 2019; Li et al., 2020; Knupp et al., 2021; Shen et al., 2022) may be contributing to the global reduction of circRNA expression in PD (**Section 5.2.7, Table 5.1**).

Comparison of the expression of circRNA regulators revealed two genes that were significantly differentially expressed in PPMI (log_2 fold change >0.1/<-0.1, FDR <0.05). The genes *ADAR* (log_2 fold change = 0.12, FDR = 0.03) and *RNASEL* (log_2 fold change = 0.10, FDR = $7.9x10^{-4}$) had increased expression in PD patients. Comparing the expression of *ADAR* and *RNASEL* in ICICLE-PD revealed that *RNASEL* was also significantly increased in PD patients (log_2 fold change = 0.13, FDR = 0.01).

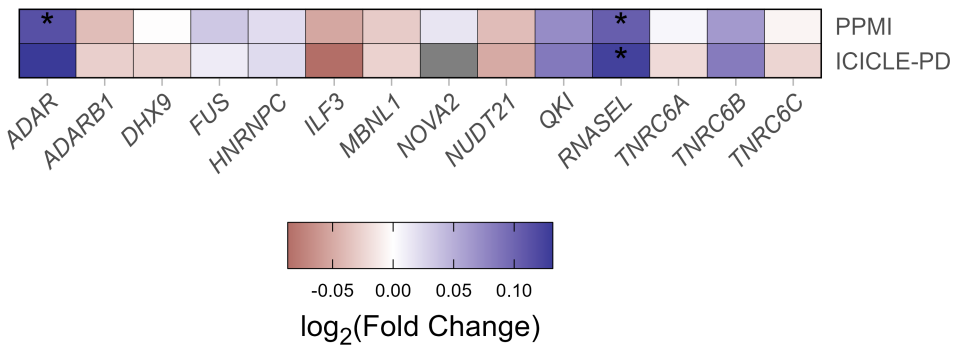


Figure 5.9.Differential expression of circular RNA regulators in early-stage idiopathic PD patients compared to controls. * = log2 fold change >0.1/ <-0.1 and FDR < 0.05.

5.3.6 *Quantifying blood A-I RNA editing levels*

Due to the use of total RNA sequencing library preparations, most aligned bases in PPMI (median = 59.8%, IQR = 4.8%) and ICICLE-PD (median = 48.7%, IQR = 9.0%) samples were resolved to intronic regions (**Figure 5.10a**). This, given the enrichment of Alu elements within introns (Sela et al., 2007), allowed me to capture A-I editing events within these regions.

I measured RNA editing events across *Alu* elements and circRNA coordinates inclusive and exclusive of flanking regions (**Section 5.2.8**) in 415 PPMI samples (PD = 255, Control = 160) and 96 ICICLE-PD samples (PD = 48, Control = 48). As expected, A-G substitutions were the

highest recorded base mismatch across each region (**Figure 5.10b, Table 8.4**) (Levanon et al., 2004; Roth et al., 2019). Whilst the A-G mismatch signal was clear in *Alu* elements, for circRNAs and their flanking regions, the signal was less pronounced (**Figure 5.10b, Table 8.4**). The signal:noise ratio (**section 5.2.8**) was higher in *Alu* elements (Mean signal:noise [SD], PPMI = 24.0 [4.7], ICICLE-PD = 27.3 [3.2], than in circRNAs (PPMI = 2.3 [0.6), ICICLE-PD = 1.8 [0.6]) or circRNAs with their flanking regions (PPMI = 2.4 [0.6), ICICLE-PD = 1.9 [0.6]) (**Figure 5.10c**). The reported signal:noise ratio in *Alu* elements was higher than previously reported estimates in GTEx whole blood samples (mean = 15.36) (Roth et al., 2019). The reported signal:noise ratios for circRNAs and regions flanking circRNAs were lower than the lowest measured tissue in GTEx samples (skeletal muscle, mean = 6.69), indicating a lack of detection power in these regions. As such, I exclusively focused on A-I editing within *Alu* elements in downstream analyses.

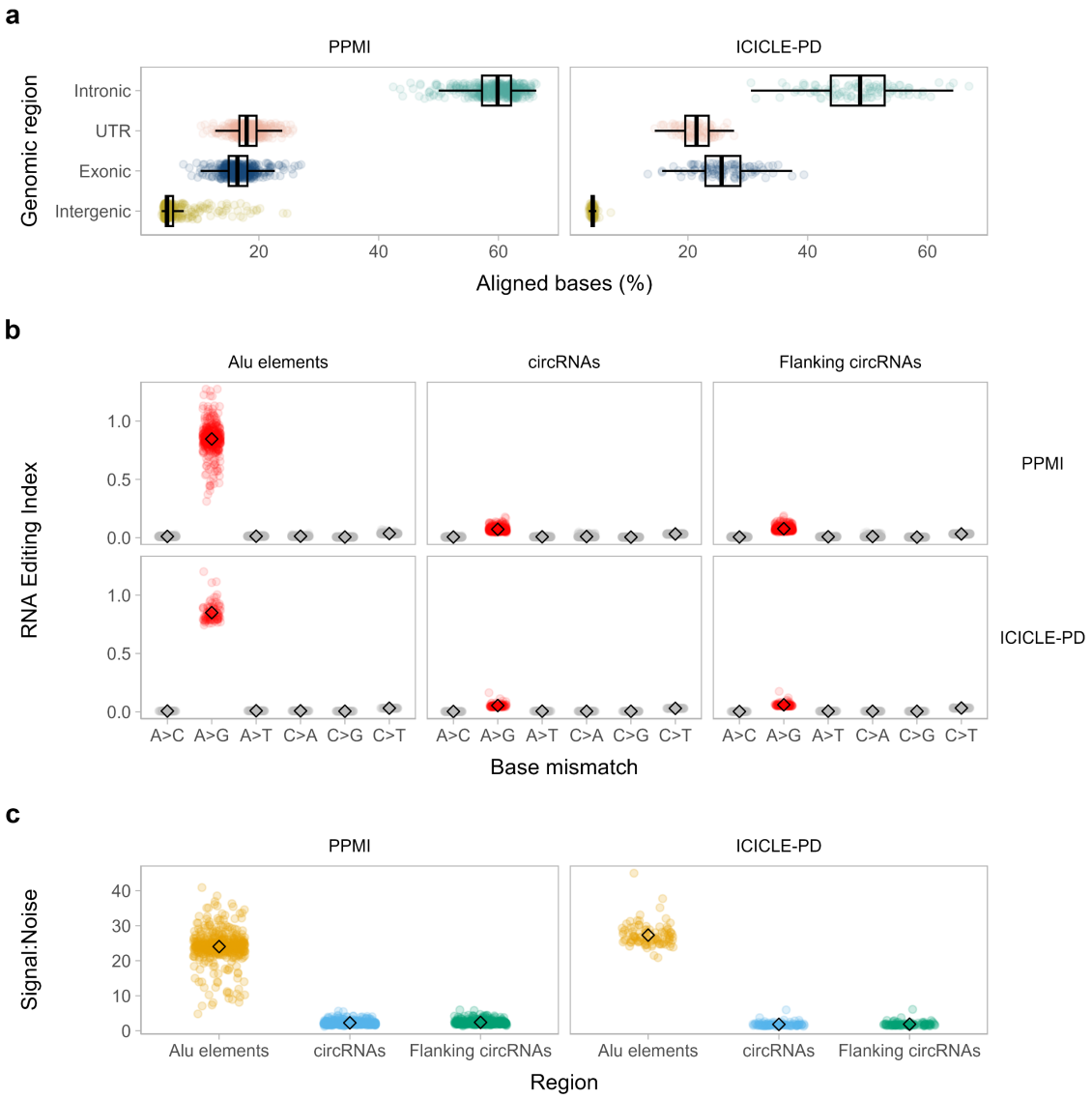


Figure 5.10. Detection of RNA editing events. (a) Bases aligned to genomic regions. **(b)** RNA editing indices across various regions based on base mismatches. A-G mismatches represent A-I editing, while other mismatches are used as control indices. **(c)** Signal: Noise ratio across various regions. This ratio shows the number of A-G mismatches relative to C-T mismatches. The diamond on each panel shows the mean value.

Following the measurement of A-I editing within *Alu* elements (defined as the *Alu* Editing Index), I next questioned the relationship between the expression of genes encoding ADAR family members and levels of A-I editing. Across PPMI, ICICLE-PD and GTEx (Lonsdale et al., 2013) samples, *ADAR* exhibited higher expression relative to *ADARB1* and *ADARB2* (**Figure 5.11a**). *ADAR* expression was significantly positively correlated to the *Alu* Editing Index in both PPMI and ICICLE-PD samples (**Figure 5.11b**). For the rest of the comparisons, only *ADARB2* was correlated to the *Alu* Editing Index in PPMI samples (**Figure 5.11b**). Furthermore, as *ADAR* expression was increased in PPMI PD patients (**Figure 5.9**), I also

stratified by study group, showing *ADAR* expression was positively correlated to the *Alu* editing index in both PD patients and controls (**Figure 5.11c**).

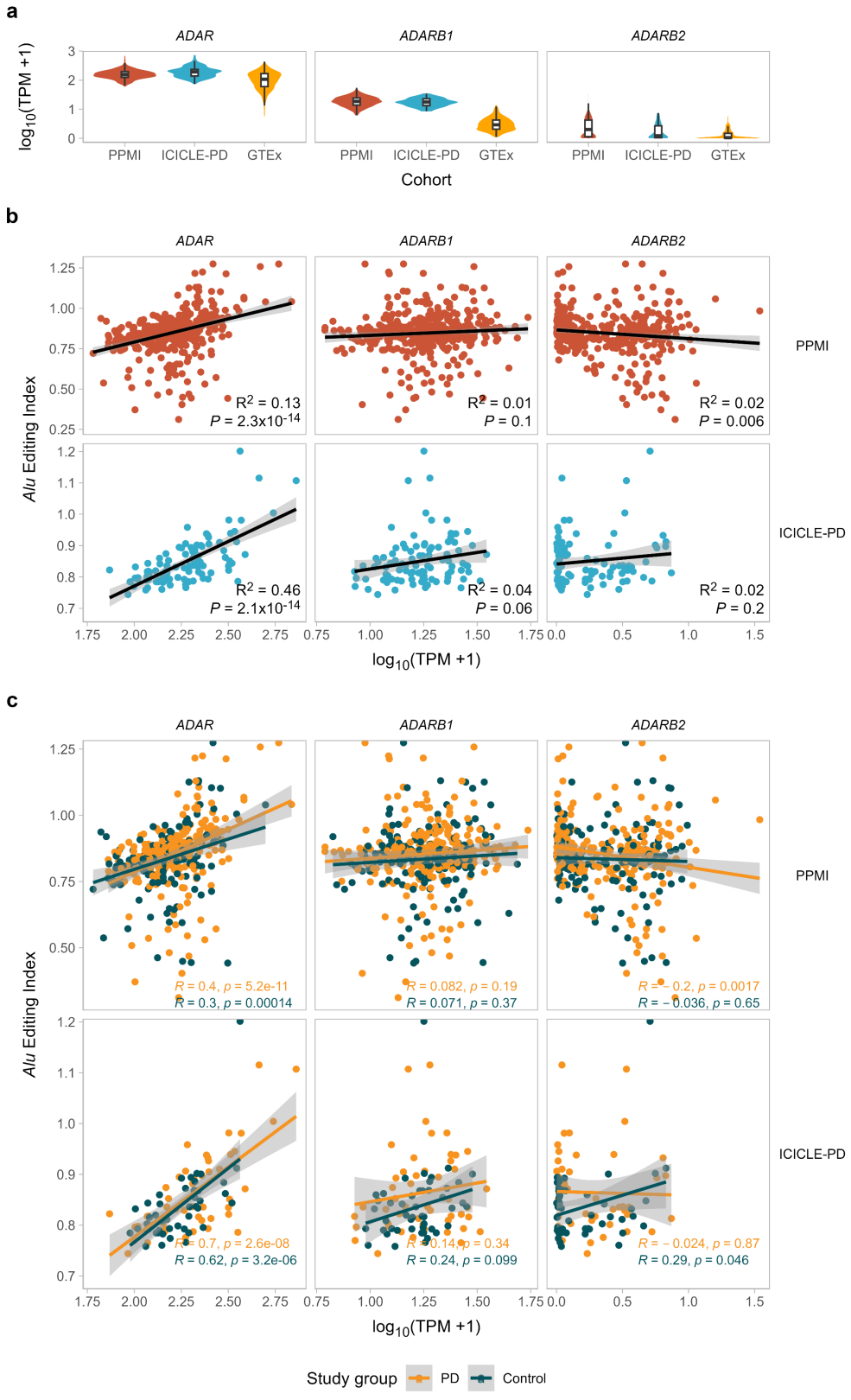


Figure 5.11. Relationship between ADAR family gene expression and A-I editing levels. (a) Expression of the ADAR family genes *ADAR*, *ADARB1* and *ADARB2* in whole blood samples from the PPMI, ICICLE-PD and GTEx (v8) studies. Normalised gene expression is given as the number of transcripts per million (TPM). **(b)** Correlation between the expression of *ADAR*, *ADARB1* and *ADARB2* with levels of A-I editing within *Alu* elements. R^2 and *P*-values derived from linear regression. **(c)** Correlation between the expression of *ADAR*, *ADARB1* and *ADARB2* with levels of A-I editing within *Alu* elements stratified by study group. Correlation assessed using Pearson's correlation coefficient.

Across study groups, A-I editing was significantly increased in PPMI PD patients (median Alu editing index = 0.86) relative to controls (median Alu editing index = 0.84, *P*-value = 0.02, Wilcoxon rank-sum test) (**Figure 5.12a**). However, although A-I editing was also increased in ICICLE-PD PD patients (median Alu editing index = 0.84) relative to controls (median Alu editing index = 0.83), this difference did not reach significance (*P*-value = 0.07, Wilcoxon rank-sum test) (**Figure 5.12a**).

To assess if A-I editing levels could account for the reduction in global circRNA expression in PD patients, I included the *Alu* editing index as an additional covariate when comparing BSJ and FSJ expression between PD patients and controls (**Figure 5.12b**). Global BSJ expression was significantly reduced in PPMI (imbalance = 0.08, Bonferroni *P*-value < 2.2×10^{-16}, two tests) and ICICLE-PD (imbalance = 0.11, Bonferroni *P*-value < 2.2×10^{-16}, two tests), with no significant changes in global FSJ expression observed (Bonferroni corrected *P*-values > 0.05, two tests) (**Figure 5.12c**).

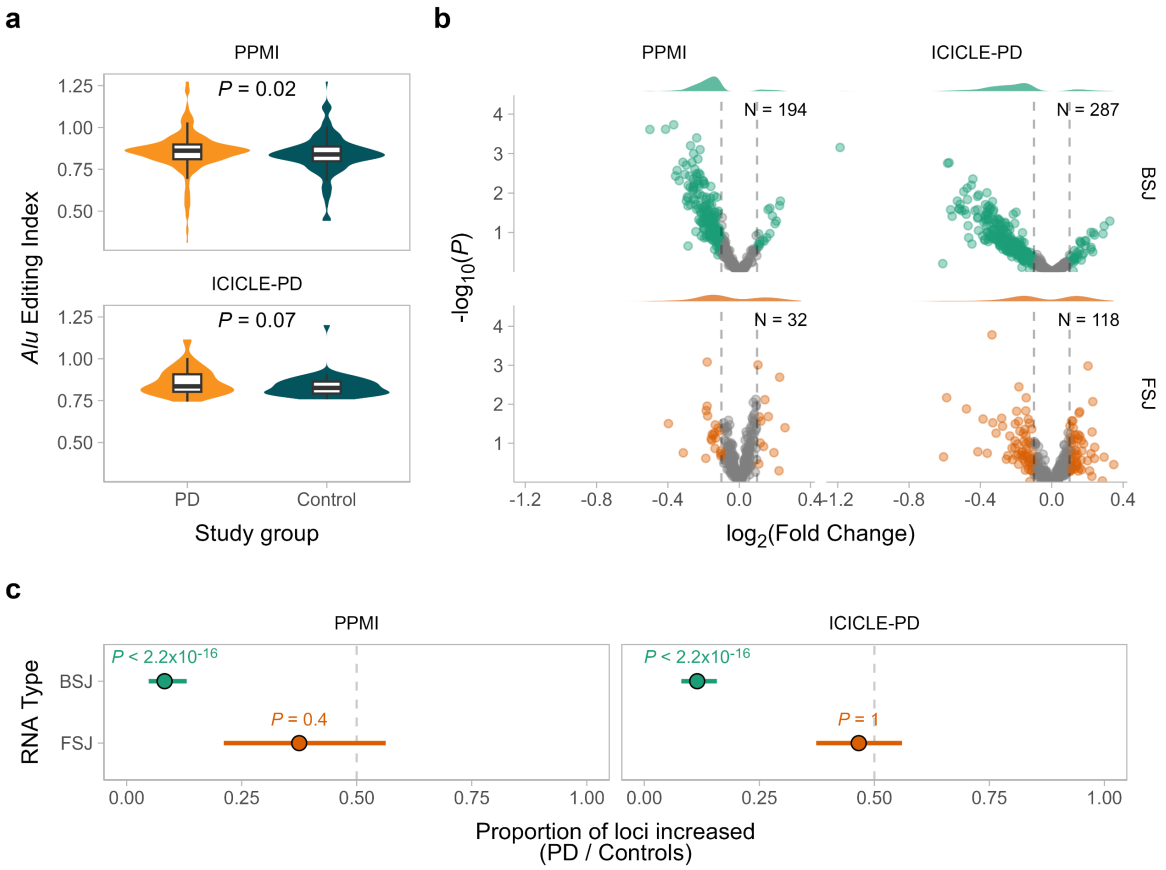


Figure 5.12. Reduced circular RNA expression in early-stage idiopathic PD patients irrespective of A-I editing levels. (a) *Alu* editing indices in early-stage idiopathic PD patients and controls. **(b)** Volcano plots showing the differential expression of back-spliced junctions (BSJs) and forward-spliced junctions (FSJ) between PD patients and controls. N shows the number of loci included when estimating the expression imbalance. Density plots show the fold change distribution of loci. **(c)** Imbalances were identified based on the proportion of loci increased in PD relative to controls based on differential expression testing of each RNA type. *P*-values obtained from a two-sided exact binomial test corrected for multiple testing (Bonferroni correction, two tests). Error bars show the 95% confidence interval of the imbalance estimate.

5.3.7 Investigating increased RNASEL expression in early-stage idiopathic PD

RNASEL was significantly increased in PD patients relative to controls in both PPMI and ICICLE-PD cohorts (**Figure 5.13b**). Consistent with increased PKR activation due to RNase L-mediated circRNAs degradation (**Figure 5.13a**) (C.-X. Liu et al., 2019), I observed significantly increased expression of *EIF2AK2*, encoding PKR, in PD patients from both PPMI and ICICLE-PD cohorts (PPMI $\log_2$ fold change = 0.1, *P*-value = $5.x10^{-5}$; ICICLE-PD $\log_2$ fold change = 0.13, *P*-value = 0.007, Wald test) (**Figure 5.13b**).

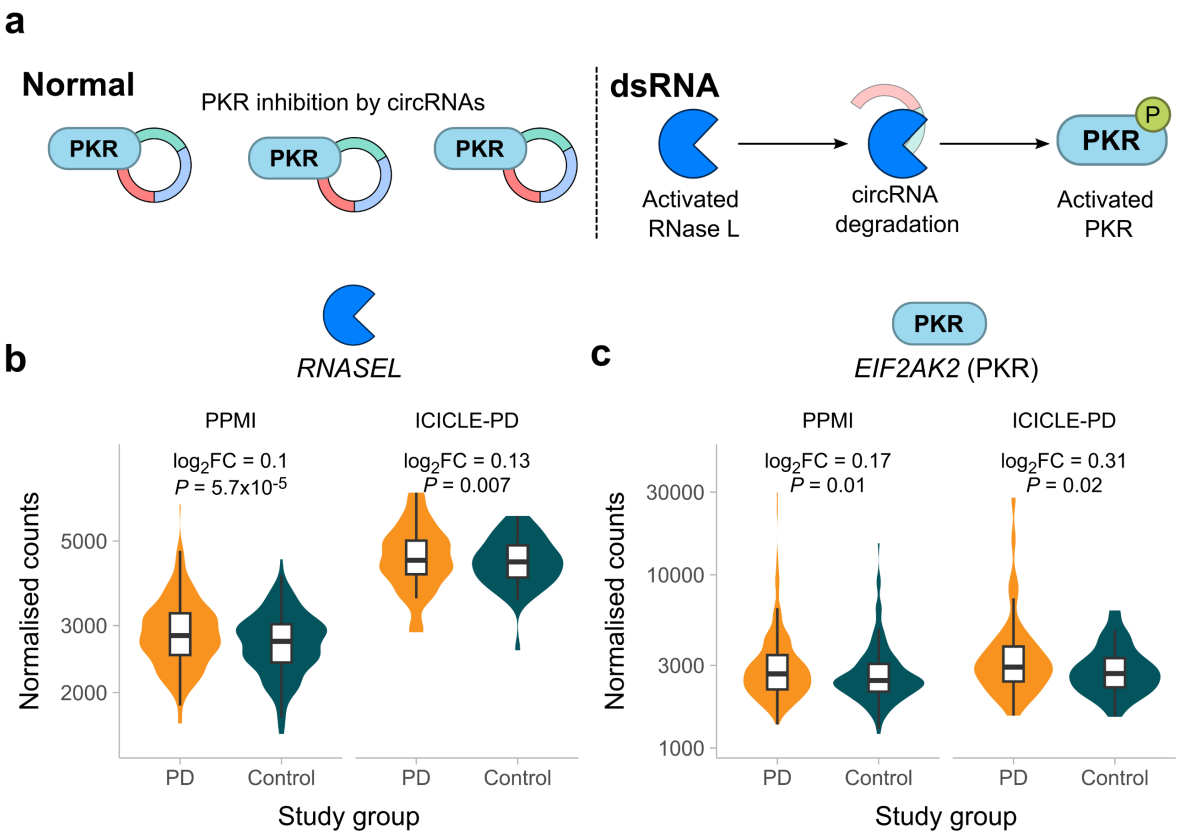


Figure 5.13. Increased expression of the genes encoding RNase L and PKR in early-stage idiopathic PD. (a) Schematic outlining the relationship between RNase L, circRNAs and PKR. Under normal cellular conditions, circRNAs bind to PKR resulting in inhibition. In the presence of dsRNA, activated RNAse L degrades circRNA allowing the activation of PKR. **(b, c)** Expression of *RNASEL* (**b**) and *EIF2AK2* (encoding PKR) (**c**) in PD patients and controls. Fold changes and *P*-values taken from the results of differential gene expression analysis (**Chapter 3**). Counts were normalised using the median of ratios method (DESeq2).

Given the role of PKR as a sensor of the integrated stress response (ISR) (Pakos-Zebrucka et al., 2016), I explored the gene expression of several ISR-associated genes (*ATF4*, *ATF5*, *DDIT3*, *PPP1R15A*, *ATF3*). There were no significant differences in the expression of these ISR-associated genes in PD patients (FDR > 0.05, Wald test, **Figure 5.14a-e**). As *ATF3*, *ATF4*, *ATF5* and *CHOP* encode transcription factors, I next examined the enrichment of genes possessing *cis*-regulatory motifs specific to these transcription factors (**Section 5.2.9**). However, no gene sets for these transcription factors were significantly enriched in PD patients in PPMI or ICICLE-PD (*P*-value > 0.05, **Figure 5.14f**).

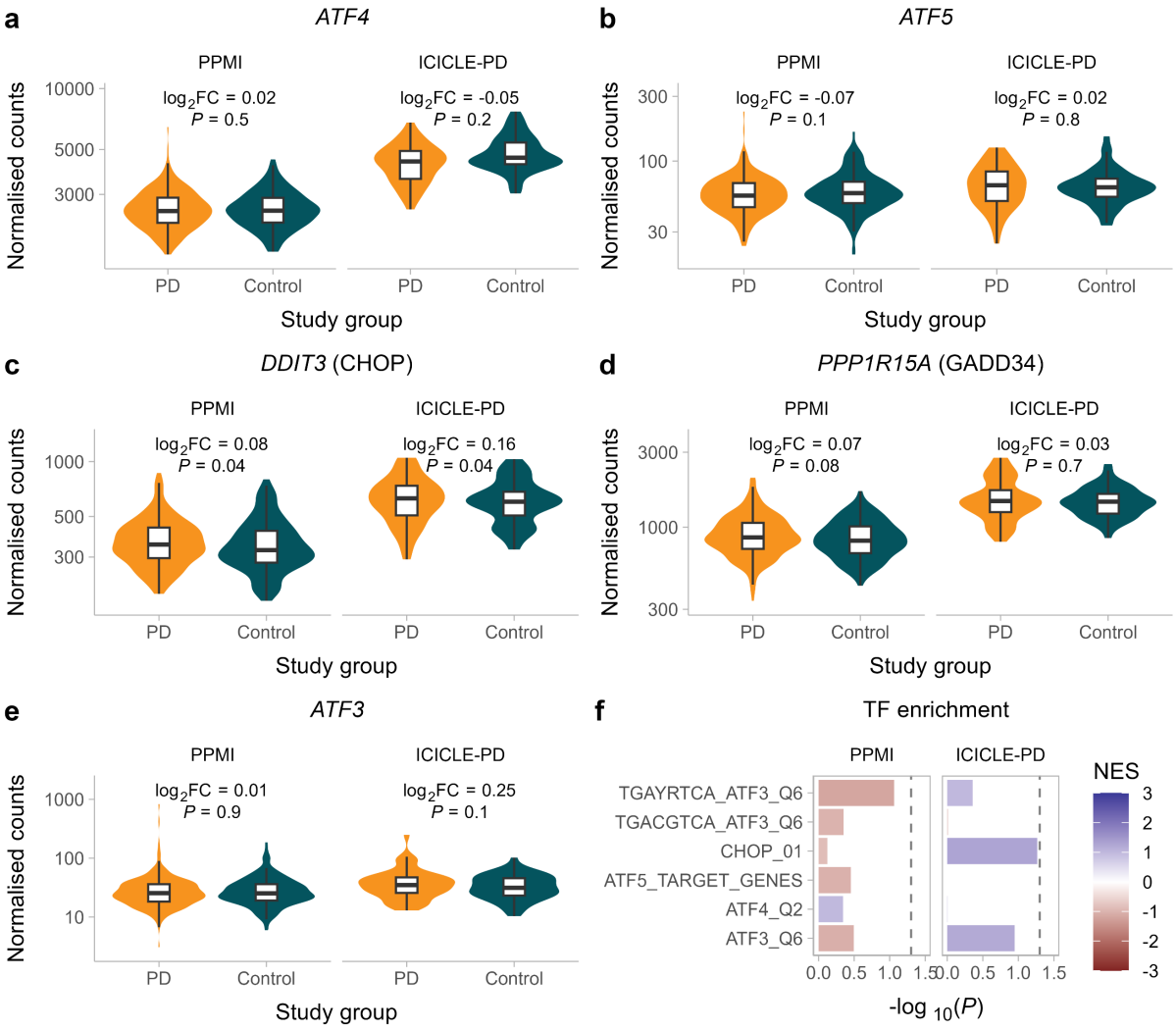
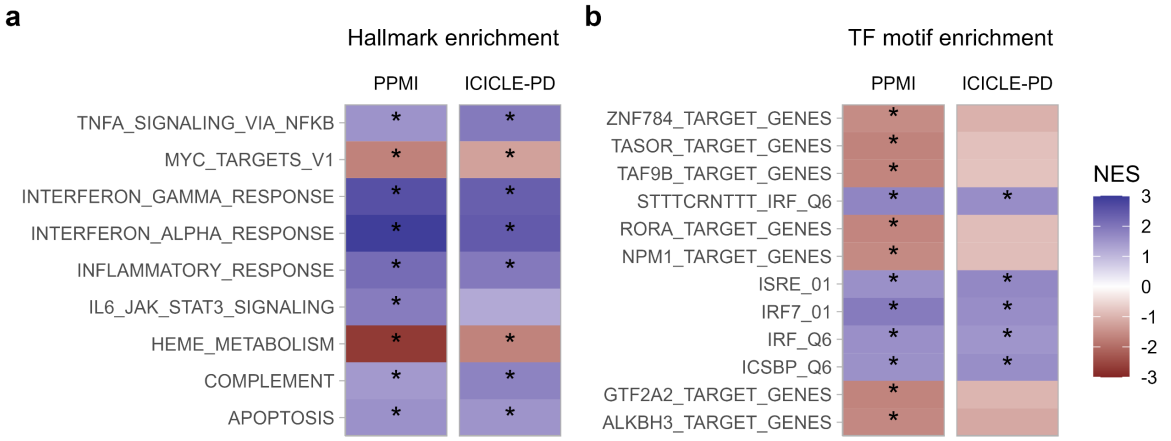


Figure 5.14. Integrated stress response-related gene expression. (a-f) Expression of integrated stress response associated genes *ATF4* (**a**), *ATF5* (**b**), *DDIT3* (encoding CHOP, (**c**)), *PPP1R15A* (encoding GADD34, (**d**)) and *ATF3* (**e**) in PD patients and controls. Fold changes and *P*-values taken from the results of differential gene expression analysis (**Chapter 3**). Counts were normalised using the median of ratios method (DESeq2). **(f)** Enrichment of genes containing *cis* ATF3, ATF4, ATF5 and CHOP binding motifs. Dashed line denotes a *P*-value of 0.05. TF = Transcription Factor, NES = Normalised Enrichment Score.

To identify systematic changes in gene expression in early-stage idiopathic PD, I performed gene set enrichment analysis (GSEA, **Section 5.2.9**). In contrast to my previous use of GSEA (**Chapter 3**), I used MsigDB Hallmark gene sets to reduce redundancy in gene sets (Liberzon et al., 2015). A total of nine gene sets were significantly enriched (FDR < 0.05) in PPMI PD cases compared to controls (**Figure 5.15a**). I subsequently replicated the dysregulation of eight gene sets (FDR < 0.05) with concordant directions of effect in ICICLE-PD samples (**Figure 5.15a**). Upregulated processes in early-stage idiopathic PD included those related to anti-viral and inflammatory activity (INTERFERON_ALPHA_RESPONSE, INTERFERON_GAMMA_RESPONSE, INFLAMMATORY RESPONSE,

TNFA_SIGNALING_VIA_NFKB, COMPLEMENT) in addition to apoptotic genes (APOPTOSIS) (**Figure 5.15a**).

To discern factors potentially involved in gene induction, I assessed *cis*-regulatory motif enrichment of all MsigDB transcription factor gene sets (**Section 5.2.9**). Gene sets corresponding to twelve transcription factor motifs showed significantly altered expression (FDR < 0.05) in PPMI PD patients (**Figure 5.15b**). In ICICLE-PD, I was able to replicate the significant positive upregulation of gene sets relating to five transcription factor motifs (**Figure 5.15b**). As shown in **Figure 5.15c**, potentially dysregulated transcription factors included IRF7, IRF1, STAT1/STAT2 and IRF8. Collectively, systematic changes in gene expression implicate activation of an antiviral and inflammatory response in early-stage idiopathic PD consistent with the work described in **Chapter 3**.



c

Gene set ID	Transcription factor	Motif
IRF7_01	IRF7	TNSGAAWNCGAAANTNNN
STTTCRNTTT_IRF_Q6	IRF1	STTTCRNTTT
IRF_Q6	IRF1	BNCRSTTTCANTTYY
ISRE_01	STAT1, STAT2	CAGTTTCWCTTTYCC
ICSBP_Q6	IRF8	RAARTGAAACTG

Figure 5.15. Activation of PKR-associated pathways in early-stage idiopathic PD. (a, b) Commonly dysregulated genes based on Hallmark gene sets (**a**) and transcription factor (TF) motifs (**b**). * = FDR < 0.05. **(c)** Description of significantly enriched TF motifs in both PPMI and ICICLE-PD datasets. NES = Normalised Enrichment Score

5.4 Discussion

Preliminary work in **Chapter 4** identified a global reduction in circRNA expression in early-stage idiopathic PD. Subsequently, I have confirmed that this reduction is not a technical or methodological artefact. Globally reduced expression in early-stage idiopathic PD was specific to circRNAs and could not be explained by concomitant reductions in the corresponding linear RNA expression. Differences in gene expression of factors involved in the circRNA life cycle indicated that processes involving the gene products of *ADAR* and *RNASEL* may be altered in PD. Subsequently, RNA editing was increased in the PPMI cohort, yet could not fully account for the reduction of circRNA expression. Increased *RNASEL* expression indicated activation of an innate immune response. This was further supported by systematic changes in gene expression related to an antiviral and inflammatory response in the blood of early-stage idiopathic PD patients.

5.4.1 Circular RNA expression is reduced in early-stage idiopathic PD

I report for the first time that circRNA expression in blood is globally reduced in early-stage idiopathic PD. I validated this finding in two independent cohorts (PPMI and ICICLE-PD) and using numerous different methodologies. The global decrease in circRNA expression appears independent of changes in linear RNA expression (**Figure 5.5**), further highlighting the distinct layer of regulation acting on circRNA (Vo et al., 2019).

Reduced blood circRNA levels is consistent with a reduction observed in substantia nigra tissue of PD patients possibly suggesting a shared underlying mechanism (Hanan et al., 2020). Results from a recent preprint would suggest a global circRNA reduction in the blood of PD patients enrolled in the Parkinson's Disease Biomarker (PDBP) study, albeit the authors do not comment on this finding (Beric et al., 2024). These findings place PD among other diseases in which reduced circRNA expression has been reported such as several cancer types (Vo et al., 2019; Hansen et al., 2022; C. Wang et al., 2022; Fuchs et al., 2023; Korsgaard et al., 2024), lupus (C.-X. Liu et al., 2019), chronic inflammatory skin conditions (Moldovan et al., 2019, 2021; Seeler et al., 2022; Guo et al., 2024), schizophrenia (Huang et al., 2023) and Alzheimer's disease (Lo et al., 2020).

The ability of global circRNA expression profiles to identify early-stage idiopathic PD patients was consistent across cohorts (AUCs = 0.59). The performance in PPMI was below that observed by gene expression (**Chapter 3**) and other biological fluid markers discussed in **Section 1.2.4**.

5.4.2 Altered proportions blood cell proportions early stage idiopathic PD

Characteristic blood cell-type circRNA expression profiles have been reported, with platelets and red blood cells in particular exhibiting the highest circRNA abundances (Alhasan et al., 2016; Nicolet et al., 2018; Gaffo et al., 2019; Grassi et al., 2021). As such, altered cell proportions may present as changes in global blood circRNA expression profiles. This is especially pertinent given that peripherally altered lymphocyte and neutrophil proportions have been reported in PD (Baba et al., 2005; Stevens et al., 2012; Jensen et al., 2021; Muñoz-Delgado et al., 2021; Craig et al., 2021). In one sense this effect is minimised as a global reduction in circRNA expression was present in early-stage idiopathic PD patients regardless of blood cell proportions. As demonstrated in systemic lupus erythematosus patients, differences in global circRNA expression can vary among blood cell types (C.-X. Liu et al., 2019).

In place of available blood cell proportions, I estimated blood cell proportions using gene expression values with the in silico deconvolution tool CIBERSORTx (Newman et al., 2015, 2019), based on its widespread use in estimating cell proportions in transcriptomic data and consistent performance across multiple independent benchmarks (Avila Cobos et al., 2020; Jin & Liu, 2021; Nadel et al., 2021; Sutton et al., 2022). Neutrophil proportions were increased in PPMI cases (unadjusted *P*-value = 0.02), consistent with previous findings (Craig et al., 2021; Pike et al., 2024). However, no blood cell proportions were significantly altered in PD patients compared to controls after multiple testing correction (**Figure 5.8**). Also using CIBERSORT, Li et al identified changes in the proportions of neutrophils and CD4+ T cells in PPMI PD cases, including those with PD linked to genetic risk variants. However, these changes were only observed in a subgroup that was classified as high-risk for faster motor progression, based on a prognostic gene expression model (Li et al., 2023), suggesting that the clinical heterogeneity of PD patients contributes to variability in blood cell proportions.

5.4.3 ADAR1 and RNase L as regulators of circular RNAs

CircRNA abundance reflects formation, stability, and degradation which is mediated by both *cis*- and *trans*-acting factors (Kristensen et al., 2019). Among regulators, those with experimental evidence of influencing levels of multiple circRNAs represent likely candidates modulating global circRNA expression (Ashwal-Fluss et al., 2014; Conn et al., 2015; Ivanov et al., 2015; Aktaş et al., 2017; Errichelli et al., 2017; Li et al., 2017; Di Liddo et al., 2019; Jia et al., 2019; C.-X. Liu et al., 2019; Li et al., 2020; Knupp et al., 2021; Shen et al., 2022). Among the list of circRNA regulator candidates, the expression of *ADAR* and *RNASEL* were significantly increased in PPMI PD cases with *RNASEL* expression also significantly increased

in ICICLE-PD cases (**Figure 5.9b**). Increased activity of *ADAR* and *RNASEL* gene products can reduce circRNA expression (Ivanov et al., 2015; Rybak-Wolf et al., 2015; C.-X. Liu et al., 2019), consistent with my observations in early-stage idiopathic PD patients (**Figure 5.1**).

My ability to form hypotheses regarding factors impacting global circRNA expression is limited to current knowledge regarding the process. To ensure stringency, I restricted the list of candidates to those with experimentally validated effects on numerous circRNAs. Several studies have utilised high-throughput screens in cell lines to identify factors involved in circRNA formation (Li et al., 2017; Liang et al., 2017). However, these screens rely on a single circRNA construct, potentially limiting their applicability to global circRNA levels and changes in circRNA expression that occur due to physiological changes (C.-X. Liu et al., 2019). Prediction based methods have proposed new RNA-binding protein candidates (Shao et al., 2022), yet further work is needed to expand knowledge of new regulators and validate their effect.

5.4.4 Increased A-I editing levels within the blood in early-stage idiopathic PD

ADAR was significantly increased in PPMI PD cases and showed a similar trend in ICICLE-PD (**Figure 5.9b**). Interestingly, the 150kDa isoform of ADAR1 (ADAR1p150) is induced by type-I interferons (Patterson & Samuel, 1995), consistent with increased expression of interferon-related genes in early-stage idiopathic PD patients (**Figure 5.15**). Given the primary role of ADAR1 in catalysing A-I RNA editing (Nishikura, 2016), increased *ADAR* expression indicated that A-I editing levels may be altered in early-stage idiopathic PD.

The expression of ADAR family genes cannot fully explain A-I editing levels (Tan et al., 2017; Roth et al., 2019), necessitating the use of computational methods to detect and quantify editing events within RNA-seq data. I observed the strongest correlation between editing in *Alu* elements and *ADAR* expression (**Figure 5.11b**), consistent with work showing that *ADAR* expression primarily correlates with editing within repetitive regions while *ADARB1* expression correlates to editing of coding regions (Tan et al., 2017).

Alu editing levels in PPMI cases were significantly increased compared to controls, with a similar trend noted in ICICLE-PD cases (**Figure 5.12a**). Using the same methodology, Hanan et al identified reduced *Alu* editing in medial temporal gyrus and amygdala samples from PD patients, albeit no difference in substantia nigra samples (Hanan et al., 2020). Using an alternative method that quantifies A-I editing at individual sites (local A-I editing) (Picardi & Pesole, 2013), other studies have reported reduced frequencies of A-I editing events in the

prefrontal cortex and blood of PD patients (Pozdyshev et al., 2022; Wu et al., 2023). Of note, Wu et al analysed A-I editing in blood samples from PPMI participants yet reported a reduced amount of A-I editing events (Olsen et al., 2018). The *Alu* editing index method of quantifying A-I editing used in this chapter was developed to address several limitations of existing techniques (Roth et al., 2019). For instance, the detection of local A-I editing events scales with coverage, limiting the comparison across samples without normalisation (Roth et al., 2019).

Global A-I editing levels could not fully account for the reduction in circRNA expression in early-stage idiopathic PD (**Figure 5.12c**). The dual role of ADAR1 in promoting and repressing circRNA biogenesis complicates its role in affecting circRNA levels (Kapoor et al., 2020; Shen et al., 2022). Identifying site-specific A-I editing events in BSJ flanking regions may therefore provide insights into their regulation by A-I editing.

5.4.5 Patterns of gene expression in early-stage idiopathic PD implicate activation of an innate antiviral response.

RNASEL expression was increased in early-stage idiopathic PD patients in both PPMI and ICICLE-PD (**Figure 5.13a**). *RNASEL* expression regulation is not well understood yet its promotor contains an interferon gamma (IFNγ) activation site and its 3' untranslated region contains AU-rich elements which influence mRNA stability (Zhou et al., 2005; Li et al., 2007). *EIF2AK2* (encoding PKR) expression was also increased in early-stage idiopathic PD patients in both PPMI and ICICLE-PD (**Figure 5.13c**). *EIF2AK2* expression is induced in response to type I interferons (interferon alpha (IFNα) and interferon beta (IFNβ)) due to an interferon-stimulated response element in its promotor (Kuhen & Samuel, 1997). Increased levels of phosphorylated PKR have previously been observed in hippocampal tissue and blood lymphocytes of PD patients (Bando et al., 2005; Pain et al., 2020). PKR can also phosphorylate α-synuclein, influencing the formation of insoluble aggregates and linking PKR to the pathogenesis of PD (Reimer et al., 2018, 2022). Given phosphorylated PKR levels are also elevated in Alzheimer's (Chang et al., 2002; Peel & Bredesen, 2003; Paccalin et al., 2006; Mouton-Liger et al., 2012), there may be a general role of PKR in the onset or progression of neurodegenerative disease.

PKR is a signalling hub for a diverse set of pathways (Kang & Tang, 2012). It is best characterised as one of the four kinases acting on the alpha subunit of the Eukaryotic Initiation Factor 2 (eIF2α) (Donnelly et al., 2013). Phosphorylation of serine 51 on eIF2α triggers the activation of the integrated stress response (ISR). ISR induction results in broadly reduced protein translation, emphasising the preferential translation of ISR-associated mRNAs. ISR-

associated mRNAs were not differentially expressed in PD patients (**Figure 5.14a-e**), consistent with translational regulation of these genes (Harding et al., 2000; Lu et al., 2004). There was also no evidence of altered gene expression of genes containing *cis*-regulatory motifs of ISR-associated transcription factors (**Figure 5.14f**). Without measuring transcription factor activity, however, I cannot exclude this possibility.

In PD patients across the PPMI and ICICLE-PD cohorts, there was an upregulation of genes associated with antiviral, inflammatory and apoptotic responses (**Figure 5.15a**). Upregulated processes, such as NF-κB signalling and apoptosis, can occur downstream of PKR activation (Jesús Gil & Esteban, 1999; Bonnet et al., 2000). Genes associated with INFα and IFNγ responses were also upregulated (**Figure 5.15a**), consistent with increased levels of these cytokines in PD (Mogi et al., 2007; Barcia et al., 2011; Main et al., 2016, 2017; Eidson et al., 2017; Quan et al., 2024). Complementarily, expression of genes containing *cis*-regulatory motifs corresponding to the transcription factors IRF1, IRF7, IRF8 and STAT1/STAT2 were increased in PD patients (**Figure 5.15a**). Both interferon regulatory factors (IRFs) and STAT1/STAT2 are key mediators of interferon signalling (Li et al., 1996; Ozato et al., 2007), further supporting the activation of interferon signalling in the blood of early-stage idiopathic PD patients.

Overall, my findings support a role of the peripheral innate immune system in the development and progression of PD (Tansey et al., 2022; Castro-Gomez & Heneka, 2024). I propose that reduced circRNA levels in early-stage idiopathic PD are related to the activation of the antiviral immune response, reflected through increased expression of *RNASEL* and *EIF2AK2* (PKR). Reduced circRNA expression would allow activation of PKR, consistent with the increased activity of PKR in lymphocytes of PD patients (Pain et al., 2020). Additionally, ADAR1-mediated A-I editing of RNA may contribute to reduced expression of specific circRNAs (Rybak-Wolf et al., 2015; Ivanov et al., 2015; Shen et al., 2022).

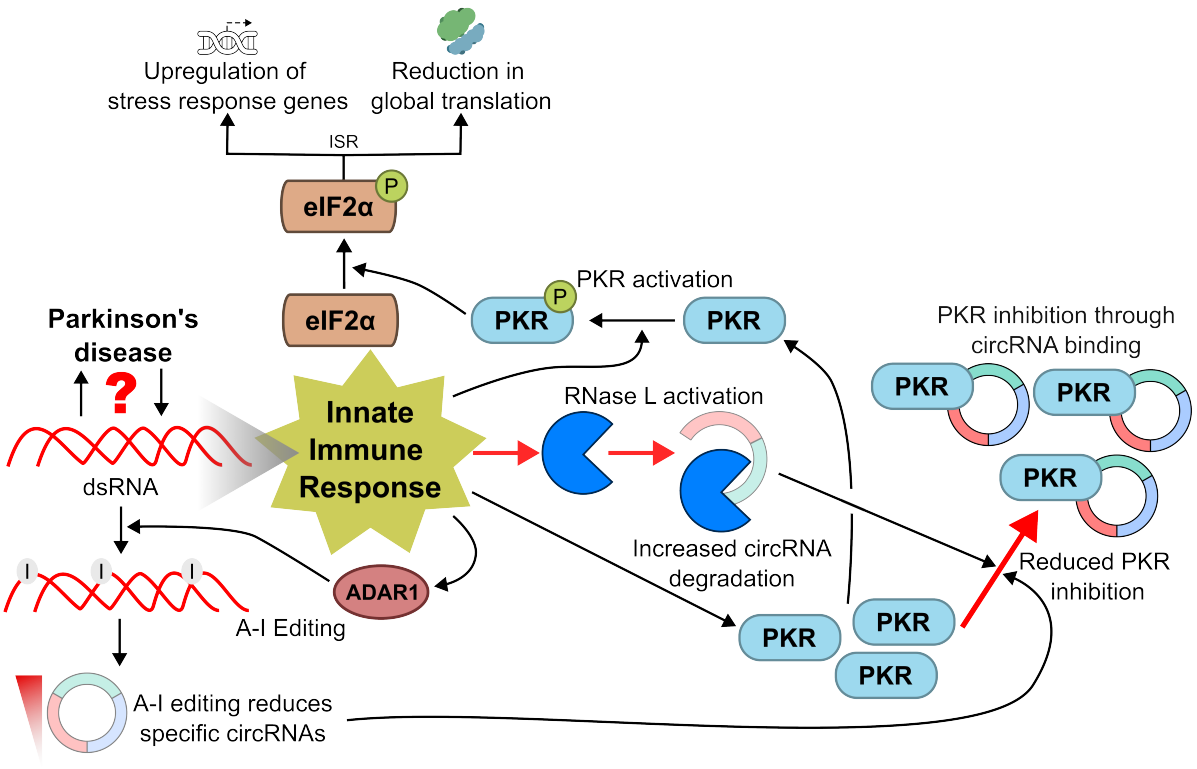


Figure 5.16. A mechanism of circular RNA reduction in Parkinson's disease. Innate immune responses can be induced by double-stranded RNA (dsRNA). A-I RNA editing by ADAR1 suppresses immune detection of endogenous dsRNA (Mannion et al., 2014; Liddicoat et al., 2015) and influences circRNA synthesis(Rybak-Wolf et al., 2015; Ivanov et al., 2015; Shen et al., 2022; Kokot et al., 2022). Under normal conditions, circRNAs bound to PKR (encoded by *EIF2AK2*) act as inhibitors, undergoing degradation when activation is required (C.-X. Liu et al., 2019). Activation of the endonuclease RNase L as part of an antiviral immune response leads to circRNA degradation (C.-X. Liu et al., 2019). Decreased circRNAs in idiopathic PD could result in increased PKR, available for activation. Once activated, PKR has numerous downstream effects including the phosphorylation of eIF2α, activating the integrated stress response (ISR).

Finally, my proposal for the activation of an innate antiviral immune response leading to reduced circRNA expression is based on work described in the literature combined with gene expression data. Levels of gene expression correlate with corresponding protein levels to a certain extent (Lundberg et al., 2010; Kosti et al., 2016). Follow up work will be required to confirm changes at the protein level. In previous work, increased PKR abundance in systemic lupus erythematosus patients was not statistically significant (C.-X. Liu et al., 2019). Employing the larger sample sizes of PPMI and ICICLE-PD datasets, I was able to detect changes in *EIF2AK2* (PKR) gene expression. Illustrated by previous work, differences in activated PKR may be even greater in PD patients relative to controls (C.-X. Liu et al., 2019). However, without these data on the levels of phosphorylated PKR in PPMI and ICICLE-PD samples, I cannot confirm PKR activation.

5.5 Conclusions

Overall, I have demonstrated that global circRNA expression is reduced in early-stage idiopathic PD. This effect was not due to technical or methodological artefacts and was detected across two independent cohorts. Reduced expression was specific to circRNAs, indicative of specific regulation among this class of RNA. As potential explanatory factors, I found altered proportions of specific blood cell types and RNA editing levels in PD patients in one cohort. However, these factors could not explain the reduction in circRNA expression. Increased expression of *RNASEL* and *EIF2AK2* suggests that circRNAs may be degraded through activation of the innate antiviral immune response. Gene expression patterns further support the activation of the antiviral and inflammatory response in the blood of patients with early-stage idiopathic PD.

These findings contribute to the growing appreciation of peripheral immune system involvement in PD. Importantly, it represents a biological change that can be detected in the blood in the early stages of disease. Elucidation of the causal mechanism may uncover deeper insights into immune dysfunction in PD and possible biomarkers signalling the onset and progression of neurodegeneration. Moving forward, it will be crucial to determine whether a reduction in global circRNA expression is represented among other PD subtypes and detectable before diagnosis.

Chapter 6. Opposing global circular RNA expression in blood and dopaminergic neurons in Parkinson's disease

6.1 Background

Most cases of PD are idiopathic, with around 15% classed as familial (Tran et al., 2020). The segregation of PD within families has uncovered genes harbouring rare variants influencing individual susceptibility to PD (**Section 1.1.2**). These pathogenic variants exhibit variable penetrance (Vollstedt et al., 2023). PD in the presence of a highly penetrant risk variant is often designated as genetic PD.

Variants in *LRRK2* are the most common cause of genetic PD (*LRRK2*-PD) (Healy et al., 2008). *LRRK2* encodes leucine-rich repeat kinase 2 (LRRK2), a protein with multiple functional domains including serine-threonine kinase and GTPase domains, implicated in the endolysosomal system, cytoskeleton dynamics and immune responses (Civiero et al., 2018; Wallings & Tansey, 2019; Erb & Moore, 2020). The G2019S variant is the most common and best characterised pathogenic PD-linked *LRRK2* variant (Simpson et al., 2022), affecting both kinase and GTPase activity (Nguyen et al., 2020). G2019S is most prevalent in individuals with Ashkenazi Jewish and North African Berbers genetic ancestry (Bar-Shira et al., 2009; Ben El Haj et al., 2017). Familial studies of individuals harbouring G2019S show an autosomal dominant pattern of inheritance and incomplete penetrance (Biskup & West, 2009). The cumulative incidence of PD in G2019S carriers varies across studies (Healy et al., 2008; Marder et al., 2015; Lee et al., 2017). A recent study estimated that 49% of G2019S carriers developed PD by age 80 (Kmiecik et al., 2024). Like G2019S, other pathogenic variants in *LRRK2* such as R1441C, R1441G, and I2020T (Richards et al., 2015; Simpson et al., 2022), are gain of function mutations (Henry et al., 2015).

GBA encodes the lysosomal enzyme glucocerebrosidase. Biallelic pathogenic variants in *GBA* cause Gaucher's disease (GD), an autosomal recessive lysosomal storage disorder leading to glucocerebroside accumulation. Increased PD rates were first observed among GD patients and their family members carrying heterozygous GD variants (Beavan & Schapira, 2013). Larger studies confirmed the increased PD incidence in individuals with GD-associated *GBA* variants (Sato et al., 2005; Nichols et al., 2009; Neumann et al., 2009). Ashkenazi Jewish populations have a ~1:17 carrier rate of GD-associated variants, with N370S being the most common (Horowitz et al., 1998). A multi-centre study found N370S and L444P variants in 15.3% of Ashkenazi Jewish PD patients and 3.2% of non-Ashkenazi Jewish PD patients (Sidransky et al., 2009). Hundreds of *GBA* variants are now linked to PD (Parlar et al., 2023), with PD penetrance estimated at ~10-30% by age 80 (Anheim et al., 2012; Rana et al., 2013; Balestrino et al., 2020). Due to a lack of Mendelian inheritance patterns, *GBA* is considered a significant genetic risk factor for PD (*GBA*-PD). *GBA* variants are classified based on their

association with mild or severe GD phenotypes (Beutler et al., 2005). Severe *GBA* variants lead to earlier onset with worse motor and cognitive outcomes compared to mild variants (Gan-Or et al., 2015; Cilia et al., 2016; Liu et al., 2016).

There is heterogeneity between idiopathic-, *GBA*- and *LRRK2*-PD patients (Kestenbaum & Alcalay, 2017; Thaler et al., 2018). *LRRK2*-associated PD (*LRRK2*-PD) patients often present with milder motor symptoms and a slightly earlier onset than idiopathic PD (Healy et al., 2008). Some *LRRK2*-PD cases lack classical Lewy body pathology (Kalia et al., 2015; Schneider & Alcalay, 2017), while Lewy body pathology is common in *GBA*-PD (Neumann et al., 2009). *GBA*-PD patients have similar initial motor presentations to idiopathic PD patients, but faster motor dysfunction progression and worse cognitive outcomes (Brockmann et al., 2015; Cilia et al., 2016). Notably, PD patients harbouring both pathogenic *LRRK2* and *GBA* variants experience a less severe cognitive decline than those harbouring only *GBA* variants, suggesting an interaction between these genes (Ortega et al., 2021). Understanding the heterogeneity among idiopathic-, *GBA*- and *LRRK2*-PD patients may reveal common dysregulated pathways.

PD motor symptoms often appear years after neurodegeneration begins (Fearnley & Lees, 1991). Detecting early-stage neurodegeneration is crucial for developing disease-modifying treatments and evaluating them in clinical trials (Mahlknecht et al., 2015). Without biomarkers for early neurodegeneration, research participant enrolment relies on risk-factor based prognostic scoring (Berg et al., 2015; Heinzel et al., 2019). Key risk factors include pathogenic variants in genes like *LRRK2* and *GBA*, in addition to rapid eye movement sleep behaviour disorder (RBD) and hyposmia. Estimates place the prevalence of olfactory dysfunction at 96% of PD patients (Haehner et al., 2009), and RBD in over 86% (Sixel-Döring et al., 2023). Individuals with poor olfactory function are at higher risk of developing PD (Ross et al., 2008) and 80% of individuals with RBD develop neurodegenerative disorders, typically synucleinopathies (Iranzo et al., 2013). Studies like the PPMI use such risk factors to create prodromal cohorts, in an attempt to recruit PD patients before diagnosis (Marek et al., 2018). Identifying differences in these prodromal individuals can reveal early dysregulated processes, offering novel biomarker targets.

In **Chapter 5**, I identified a specific reduction in global circRNA expression in the blood of early-stage idiopathic PD patients from the PPMI and ICICLE-PD cohorts. However, whether a reduction is also present in genetic PD patients is currently unknown. Additionally, the detection of globally reduced circRNA could be detectable before diagnosis. Given the lower circRNA levels in the substantia nigra of PD patients (Hanan et al., 2020), similar reductions

in blood could indicate a shared mechanism. However, the impact of PD on global circRNA levels in dopaminergic neurons remains to be explored.

To address these questions, I measured circRNA expression in the blood of PD patients with pathogenic *GBA* and *LRRK2* variants from the PPMI cohort, as well as in individuals at high risk for developing PD, comparing both to controls. To explore a potential mechanistic link between circRNA expression and PD, I compared circRNA levels in dopaminergic neurons from the substantia nigra of PD patients and matched controls from the BRAINcode cohort. Integrating gene expression data from these comparisons provided insights into potentially dysregulated processes affecting global circRNA expression in the brain and blood of PD patients and those at heightened risk.

6.2 Methods

This chapter refers to RNA sequencing data generated externally to this PhD project by the PPMI and BRAINcode studies. All sample metadata was collected from previous publications (X. Dong et al., 2018; Craig et al., 2021; Dong et al., 2023). Raw sequencing data were downloaded as described in the relevant sections.

6.2.1 Parkinson's Progression Marker Initiative (PPMI) cohort

PPMI sample study groups were in accordance with a 2023 data freeze of the PPMI cohort (dated 23rd October 2023). Samples were assigned to three cohorts: PD, Controls and Prodromal, based on a centralised review by the PPMI. These reviews highlight changes in diagnoses compared to initial enrolment status. They also take account of updated biological information regarding participants. For example, controls with positive α-synuclein seed amplification assay results are excluded (Siderowf et al., 2023).

Sample metadata were taken from the most recent curated data set (dated 29th January 2024). Some of the samples originally included in the genetic registry subgroup do not have any follow-up information and may lack information at baseline (disease duration and treatment status for example). For these samples, metadata was taken from the information provided in previous publications (Craig et al., 2021). Analysis samples were selected according to the participant- and technical-level criteria outlined in **Figure 6.1**.

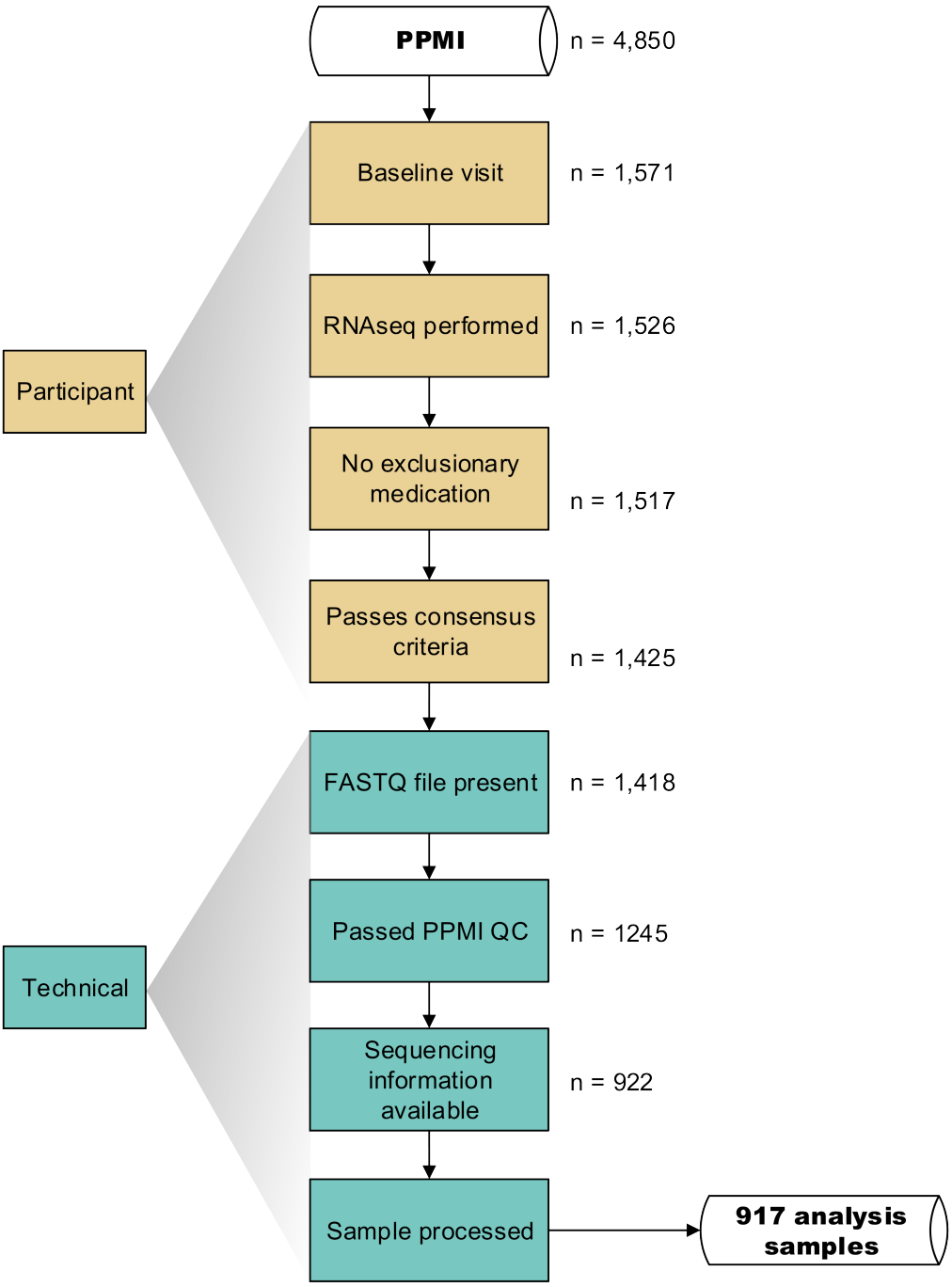


Figure 6.1. PPMI sample selection. Schematic outlining the workflow for selecting PPMI samples for analysis. At each step, the number of samples remaining is shown on the right-hand side.

A breakdown of demographic and clinical information for each PPMI sub-cohort is shown below (**Table 6.1**).

Variable	PD (n = 491)	Prodromal (n = 264)	Control (n = 162)	*P*-value
Sex (M/F)	292/199	119/145	103/59	$7.8x10^{-5}$
Age at consent (years)	63.4 ± 10.1	59.9 ± 12.8	60.5 ± 11.7	$5.0x10^{-5}$
Disease duration (months)	13.8 ± 18.8	NA	NA	NA
LEDD (mg/day)	130.8 ± 315.5	2.6 ± 14.1	0.0 ± 0.0	$9.9x10^{-13}$

Table 6.1. Sample characteristics of PPMI analysis samples. Where possible, the mean value for each group is given ± standard deviation. The difference in the proportion of Males/Females between cohorts was assessed using a two-sided Fisher's exact test. All other group differences were assessed using a one-way ANOVA.

Variable	iPD (n = 325)	*LRRK2*-PD (n = 125)	*GBA*-PD (n = 36)	*LRRK2* & *GBA*-PD (n = 5)	*LRRK2*-Control (n = 136)	*GBA*-Control (n = 76)	*LRRK2* & *GBA*-Control (n = 10)	Hyposmia (n = 18)	RBD (n = 24)	Control (n = 162)
Sex (M/F)	210/115	61/64	21/15	0/5	59/77	25/51	3/7	12/6	20/4	103/59
Age at consent (years)	62.6 ± 9.55	65.7 ± 10.6	63.4 ± 11.3	60.9 ± 12.9	58.7 ± 12.6	56.5 ± 13.9	62.8 ± 7.12	67.6 ± 6.55	70.2 ± 6.35	60.5 ± 11.7
Disease duration (months)	6.6 ± 6.5	35.1 ± 25.0	28.2 ± 29.4	43.4 ± 25.4	N/A	N/A	N/A	N/A	N/A	N/A
LEDD (mg/day)	0 ± 0	548.9 ± 454.0	338.6 ± 400.2	363.2 ± 277.3	N/A	N/A	N/A	N/A	N/A	N/A

Table 6.2. Sample characteristics of PPMI analysis subgroups. Where possible, the mean value for each group is given ± standard deviation.

Identification of PD risk variants in PPMI participants was performed by the PPMI (dated 20th December 2023). Briefly, variants were called using data from several technologies: genotyping array, whole genome sequencing, whole exome sequencing, sanger sequencing (*GBA*) and Clinical Laboratory Improvement Amendments approved genetic testing. Variants were extracted from selected genes included in the PD GENEration study (*LRRK2*, *GBA*, *VPS35*, *SNCA*, *PRKN*, *PARK7*, and *PINK1*) (Cook et al., 2023). All PD-risk variants met the American College of Medical Genetics and Genomics criteria for pathogenicity at the time of analysis (protocol document dated 20th December 2023) (Richards et al., 2015).

This chapter focuses on participants with PD risk variants in *LRRK2* or *GBA* (**Table 6.3**). As previously described, G2019S was the most common LRRK2 variant while N409S (previously reported as N370S) was the most common GBA variant (**Table 6.3**) (Simuni, Uribe, et al., 2020; Simuni, Brumm, et al., 2020).

Gene	Variant	Frequency
LRRK2 (*n* = 261)	G2019S	178
	N1437H	0
	R1441C	0
	R1441G	0
	R1441H	0
	I2020T	0
	NA	83
GBA (*n* = 108)	N409S (N370S)	53
	L483P (L444P)	1
	115+1G>A (IVS2+1G>A)	1
	R159W (R120W)	1
	R535H (R496H)	0
	R502C (R463C)	0
	D448H (D409H)	0
	S212* (S173*)	0
	L29Afs*18 (84GG)	0
	NA	52
LRRK2 & *GBA* (*n* = 15)	G2019S & N409S (N370S)	9
	G2019S & R535H (R496H)	3
	G2019S & R159W (R120W)	2
	G2019S & L29Afs*18 (84GG)	1

Table 6.3. Pathogenic GBA and LRRK2 variants included in the 2023 PPMI data freeze. Frequencies of each variant are provided where possible. Specific variant information was not available for samples enrolled in the PPMI genetic registry subgroup. Previous *GBA* variant nomenclature is shown in parentheses.

A total of 166 PPMI samples with a clinical diagnosis of PD harbouring known pathogenic variants (Genetic PD) passed the selection criteria, with the breakdown according to gene shown in **Table 6.4**.

Genetic PD subgroup	n
LRRK2	125
GBA	36
LRRK2 & *GBA*	5

Table 6.4. Number of processed samples in PPMI genetic PD subgroups.

PPMI enrolment strategies for genetic PD subgroups differed from idiopathic PD patients, resulting in different sample characteristics. Participants in genetic PD subgroups had longer

disease duration and, as a function of disease progression, were receiving dopaminergic treatment upon enrolment (**Figure 6.2a, b**). PD patients harbouring dual pathogenic *LRRK2* and *GBA* variants were excluded.

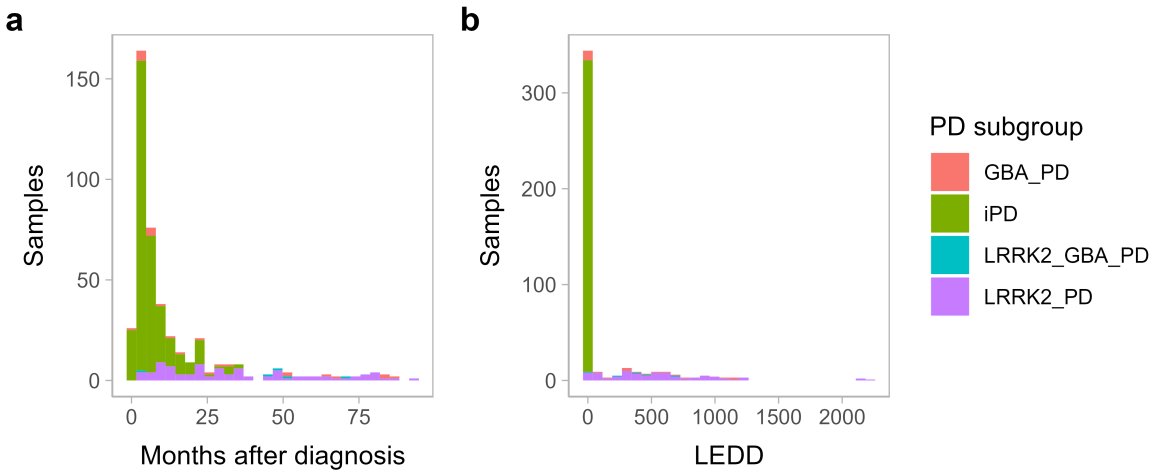


Figure 6.2. Differences in disease duration and dopaminergic treatment among PPMI PD subgroups. (a, b) Distribution of disease duration (**a**) and dopaminergic treatment dosage (**b**) among PPMI PD subgroups. Each bar is coloured according to subgroup i.e., presence of pathogenic variants in *LRRK2* or *GBA*. LEDD = Levodopa Equivalent Daily Dose.

Individuals were recruited to the Prodromal cohort if they harboured at least one known PD-risk factor: pathogenic genetic risk variants (*LRRK2*, *GBA*, *SNCA*), hyposmia with a deficit based on dopamine transporter imaging, or REM sleep behaviour disorder (RBD). Non-manifest *LRRK2* and *GBA* carriers (i.e., harbouring pathogenic *LRRK2* and *GBA* variants but have no clinical PD diagnosis) are referred to as *LRRK2*-Control and *GBA*-Control respectively. 10 participants with no diagnosis of PD had pathogenic variants in both *LRRK2* and *GBA* yet were excluded. After selection criteria and processing of total RNAseq data, 254 Prodromal samples were available (Table **6.5**)

At-risk subgroup	**n**
LRRK2-Control	136
GBA-Control	76
Hyposmia	18
RBD	24

Table 6.5. Number of PPMI participants at increased risk of developing PD. Subgroups are stratified by specific PD risk category.

6.2.2 BRAINcode cohort

Samples from the BRAIN Cell encyclOpeDia of transcribed Elements Consortium (BRAINcode), were collected and processed by BRAINcode members as previously described (X. Dong et al., 2018; Dong et al., 2023). Briefly, frozen postmortem brain samples were collected from individuals diagnosed with PD, AD and those free of neurodegenerative disease after neuropathological inspection. Individuals without a neurodegenerative diagnosis but evidence of Lewy body pathology at autopsy were termed Incidental Lewy body cases (ILB). Specific neuron types were isolated using laser-capture microdissection. For PD and ILB samples, dopaminergic neurons were extracted from the substantia nigra, for AD samples, pyramidal neurons from the temporal cortex. Both neuron types were also isolated in control samples.

A total of 197 BRAINcode samples were available (i.e. passed QC checks in previously published work) (X. Dong et al., 2018; Dong et al., 2023). Given the scope of this chapter, I focused on a set of 104 samples comprising dopaminergic neurons isolated from the substantia nigra of PD and ILB cases alongside controls. Sample characteristics are shown in **Table 6.6**.

Variable	PD (n = 18)	ILB (n = 27)	Control (n = 59)	*P*-value
Sex (Male/Female)	14/4	18/9	39/20	0.68
Age	76.9 ± 7.7	84.6 ± 6.3	79.4 ± 11.3	**0.02**
Postmortem interval (hours)	5.2 ± 5.8	4.0 ± 4.8	5.3 ± 7.2	0.68
Unified Lewy Body staging system	3.1 ± 1.0	2.3 ± 1.3	0.0 ± 0.0	**$< 2.2 \times 10^{-16}$**

Table 6.6. BRAINcode sample characteristics. Table showing the relevant demographic, clinical and technical characteristics of BRAINcode samples for which total RNAseq was performed on dopaminergic neurons isolated from the substantia nigra. Numeric values are given as mean ± standard deviation. The difference in the proportion of Males/Females between cohorts was assessed using a two-sided Fisher's exact test. All other group differences were assessed using a one-way ANOVA.

6.2.3 RNA sequencing

RNAseq methods of the PPMI cohort were described in **Section 2.3.2**.

BRAINcode sequencing data were downloaded from the Gene Expression Omnibus from the project accession *PRJNA902741*. Methods describing the sequencing have been previously published (X. Dong et al., 2018; Dong et al., 2023). Briefly, neuronal RNA was isolated with the Arcturus PicoPure RNA Isolation Kit (Applied Biosystems) and treated with DNase (Qiagen). RNA was amplified with the Ovation RNA-Seq System (NuGen). Library preparation used the TruSeq RNA Library Prep Kit v2 (Illumina). Sequencing was performed Illumina HiSeq 2000 and 2500 generating 50-75bp paired end reads.

6.2.4 RNA sequencing quality control

A series of quality control steps were performed as described in **Section 2.4**. Due to the use of an unstranded library preparation when sequencing BRAINcode samples, several strand-specific parameters were changed. For HISAT2 alignment, no *--rna-strandness* flag was set. When running Picard CollectRnaSeqMetrics, *--STRAND_SPECIFICITY* was set to *NONE*.

BRAINcode samples had a median sequencing depth of 80.1 million paired-end reads (IQR = 44.9) (**Figure 6.3a**). Sequencing quality was generally above PHRED 30 (**Figure 6.3b**). The last base of samples sequenced at 50bp length reads showed a drop in quality (**Figure 6.3b**), however downstream soft clipping during alignment and transcript quantification accounted for this. Across the cohort, a median of 85.1% (IQR = 2.8%) of sequenced reads aligned to the genome, and a median of 52.1% of all reads aligned uniquely (IQR = 7.4%) (**Figure 6.3c**). In BRAINcode samples, a median of 9.6% (IQR = 3.3%) of aligned bases corresponded to the protein-coding regions of genes (**Figure 6.3d**). Overall, aligned bases mostly mapped to regions encoding ribosomal RNA (rRNA) (median 31.0%, IQR = 14.1%) (**Figure 6.3d**). However, the detection of circRNAs is likely to be impacted by the presence of rRNA (Nielsen et al., 2022).

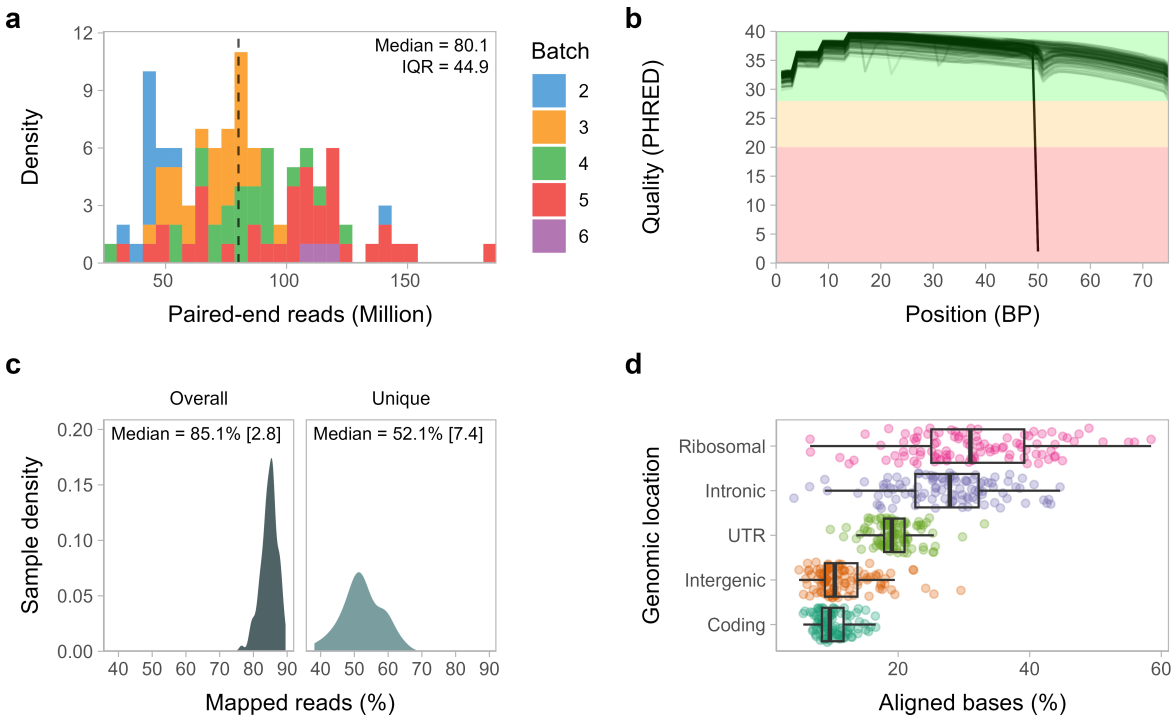


Figure 6.3. Sequencing QC of BRAINcode samples. (a) The distribution of the amount of paired-end reads, coloured based on a sequencing batch. The dotted line indicates the median sequencing depth of all sequenced samples. **(b)** Sequencing quality along the read length. At each base position, the mean quality (represented as a PHRED score) is shown. Quality scores are coloured to reflect the categories in FastQC. **(c)** The distribution of reads mapping to the genome. Mapped reads have been categorised as a percentage of all reads that were mapped (Overall) and reads that were mapped uniquely (Unique). Summary statistics are given as the median percentage of reads in each category [Interquartile range]. **(d)** Distribution of aligned bases according to the genomic location of the annotated feature. UTR = Untranslated Region.

6.2.5 Gene quantification

Transcripts were quantified with Salmon v1.3.0 (Patro et al., 2017) (**Section 2.5.1**). Transcript counts were summarised to the gene level using tximport v1.28.0 for downstream use with both DESeq2 and limma as described (**Section 2.8**).

6.2.6 Circular RNA detection and quantification

The circRNA detection workflow (described in **Section 2.6**) was modified to account for FSJ detection batch effects **(Section 8.3**). Instead of overlapping circRNA detections from multiple tools, circRNAs were detected using CIRI2 v2.0.6 (Gao et al., 2018) and then quantified from raw sequencing reads using CIRIquant v1.1.2 (Zhang et al., 2020). In a recent benchmark, this approach was shown to have the highest precision of all tested methods (Vromman et al., 2023). From this, back-spliced junction (BSJ) and forward-spliced junction (FSJ) counts were

estimated representing circRNA and the corresponding cognate linear RNA expression respectively. I retained circRNAs mapped to at least five raw reads in more than one individual in line with increased precision of detection (Vromman et al., 2023).

In line with previous **Chapters 3, 4 and 5**, circRNA expression was normalised to the number of reads mapped by Salmon (i.e., transcriptome-mapped reads). When performing differential expression, this method of normalisation was achieved by providing the corresponding library size and normalisation factors calculated for each sample based on gene expression.

Circular:linear RNA ratios at each BSJ locus were reported by CIRIquant based on the formula described in **Equation 2.1**.

6.2.7 Modelling biological and technical sources of RNA expression variation

The contributions of biological and technical sources of variation to RNA expression were assessed based on the method outlined in **Section 2.7**. Principal components of gene and circRNA expression were calculated as described in **Section 2.10.1**.

As described in **Section 2.7**, a reduced set of technical and biological covariates (age, sex) were used to construct linear mixed models predicting individual gene and circRNA expression using variancePartion v1.30.2 (Hoffman & Schadt, 2016).

6.2.8 RNA differential expression

Before differential expression analyses, additional filtering of genes and circRNAs based on expression was performed to ensure consistent detection across different comparison groups. For PPMI and BRAINcode samples, genes were filtered using the filterByExpr function in edgeR v3.42.4 (Robinson et al., 2010). In PPMI samples, circRNAs were retained that had an expression > 0.1 counts per million in at least half of all samples. In BRAINcode samples, circRNAs were retained that were detected in at least 10% of all samples. For both sets of samples, the corresponding cognate linear RNAs (based on FSJ loci) were also kept.

Differential expression was performed using limma as outlined in **Section 2.8**. For comparisons of RNA expression across PPMI subgroups, I adjusted for age (binned into <55, 55-65, >65), sex and sequencing batch. For gene differential expression, I also adjusted for the percentage of usable bases. For circRNA and cognate linear RNA differential expression, I also adjusted for the percentage of intronic bases. For comparisons of RNA expression across BRAINcode subgroups, I adjusted for age, sex, sequencing batch, RNA integrity number and post-mortem

interval. For gene differential expression, I also adjusted for the median coefficient of variation of coverage.

6.2.9 Gene Set Enrichment Analysis

Systematic changes in gene expression were identified using Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005). GSEA was performed using the *GSEA()* function in clusterProfiler v4.8.3 (Wu et al., 2021). Hallmark gene sets (Liberzon et al., 2015) were collected using msigdbr v7.5.1. GSEA was performed on all genes ranked by fold change using 10,000 permutations running the fgsea v1.26.0 backend (Korotkevich et al., 2021). *P*-values were corrected for multiple testing using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995).

6.2.10 Statistical analysis

All statistical analysis was performed in R v4.3.1, with relevant statistical tests detailed in the text. Cumulative proportions were calculated and plotted using the *stat_ecdf()* function implemented in ggplot v3.4.4.

Imbalances in RNA expression were determined based on the distribution of fold changes obtained by differential expression testing as described in **Section 2.10.3**. In this chapter however, no fold change threshold was set meaning $\log_2$ fold changes > 0 were classed as increased while $\log_2$ fold changes < 0 were classed as decreased.

6.3 Results

6.3.1 Circular RNA detection and expression in PPMI blood samples

I used total RNAseq data from 917 PPMI participants (see **Section 0** for selection criteria) to detect gene, circRNA (BSJ) and cognate linear RNA (FSJ) expression (**Figure 6.4a**). Participants were categorised into three top-level study cohorts: those with clinically diagnosed PD (n = 491), prodromal participants that possess at least one PD-related risk factor with no PD symptoms (n = 264) and controls free of neurological disease and known risk factors (n = 162). After processing, 18,751 unique circRNAs were retained based on distinct BSJ coordinates. CircRNAs were consistently detected across the entire dataset with 18,635 (99.4%) detectable in at least one individual in each study cohort (**Figure 6.4b**).

Previously, I normalised the number of circRNAs detected and their expression against the number of reads mapping to the transcriptome in each sample (**Chapters 3, 4 and 5**). To test the suitability of this normalisation method in this set of PPMI samples, I collected sample-level measures of sequencing depth including the number of total reads sequenced, the number of genome-mapped reads and the number of transcriptome-mapped reads. Both the number of total reads and genome-mapped reads varied across study cohorts (P-value < 0.05, Kruskal-Wallis test), however no study cohort differences were identified in the number of transcriptome-mapped reads (P-value = 0.4, Kruskal-Wallis test) (**Figure 6.4c**). Compared to the other measures of sequencing depth, transcriptome-mapped reads had the highest correlation with total circRNA expression (R^2 = 0.28) and the number of unique circRNAs (R^2 = 0.26) in each sample (**Figure 6.4d, e**). As such, I normalised the amount and abundance of circRNA against the number of transcriptome-mapped reads in each sample in this set of PPMI samples.

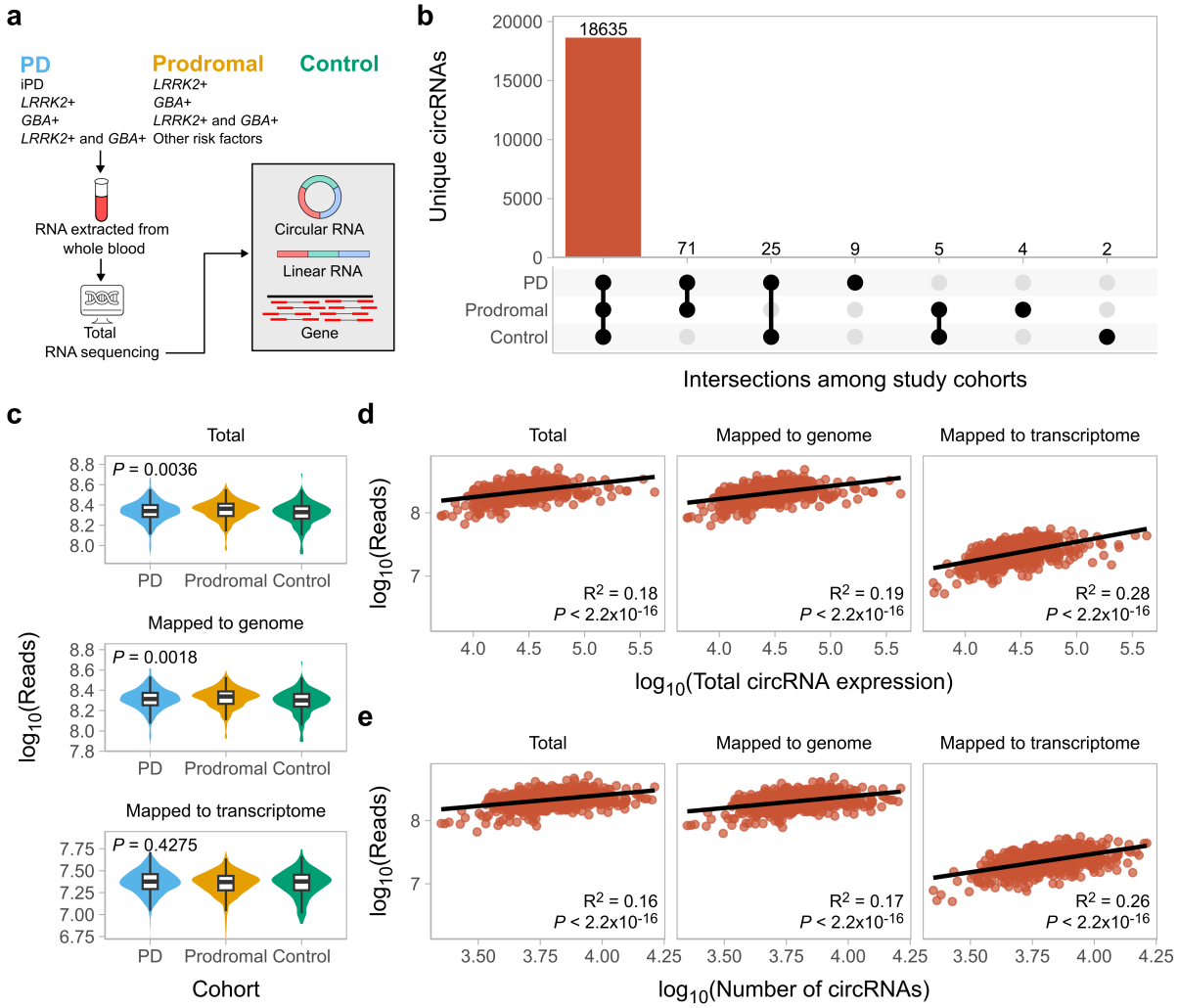


Figure 6.4. Overview of RNA quantification in PPMI samples and normalisation of circular RNA expression. (a) Overview of the RNA quantification workflow. The study cohorts present within the PPMI are shown along with examples of subgroups belonging to each cohort. iPD = idiopathic PD **(b)** Intersection of detected circRNAs across PPMI study cohorts. The number above each bar shows the size of the intersection. **(c)** Comparison of sequencing depth across study cohorts. *P*-value calculated from a Kruskal-Wallis test. **(d, e)** Correlation between sequencing depth to total circRNA expression (**d**) and number of unique circRNAs in each sample (**e**). Shaded area indicates the 95% confidence interval of the linear regression line.

To improve comparisons of gene and circRNA across PPMI study cohorts and subgroups, I applied additional expression-based filtering (**Section 6.2.8**) resulting in 20,562 genes and 4,546 circRNAs remaining. Projection of the first two PCs of gene and circRNA expression (**Figure 6.5a**) indicated sources of variation contributing to RNA expression other than the study cohort.

To identify sources of variation contributing to RNA expression, I first assessed the relationship between technical sequencing metrics to gene and circRNA quantification (**Section 6.2.7**). The first ten PCs of gene and circRNA expression cumulatively explained 57.9% and 33.3% of the

variation respectively (**Figure 6.5b, c**). The first PCs of both gene and circRNA expression were significantly correlated to several technical sequencing metrics (FDR < 0.05, linear regression) (**Figure 6.5b, c**). For gene expression, the percentage of usable bases explained >52% of the variation in PC1 and was highlighted as a potential technical covariate (**Figure 6.5b**). For circRNA expression, no technical metric explained >50% of the variation in any of the first ten PCs (**Figure 6.5c**). Given the variation explained by PC1 relative to PC2 (23.2% vs 3.8%) and ensure consistency with previous analyses (**Chapters 4 and 5**), I selected the percentage of intronic bases (accounting for 47% of the variation in PC1), as a potential technical covariate to capture variation attributed to PC1 (**Figure 6.5c**).

Next, I fit linear mixed models using variancePartition (Hoffman & Schadt, 2016) to the expression of individual gene and circRNA expression. As explanatory variables, I included sources of technical variation identified in the previous step (percentage of usable bases for gene expression and percentage of intronic bases for circRNA expression) alongside sequencing batch and common sources of biological variation (age and sex). Apart from sex, all included covariates generally explained more variance than the PPMI study cohort, indicating I should account for these covariates when modelling differences in expression (**Figure 6.5d**). Due to the presence of outliers when considering the contributions of sex to the expression of specific genes and circRNAs (**Figure 6.5d**), I also included sex as an explanatory variable when modelling differences in expression.

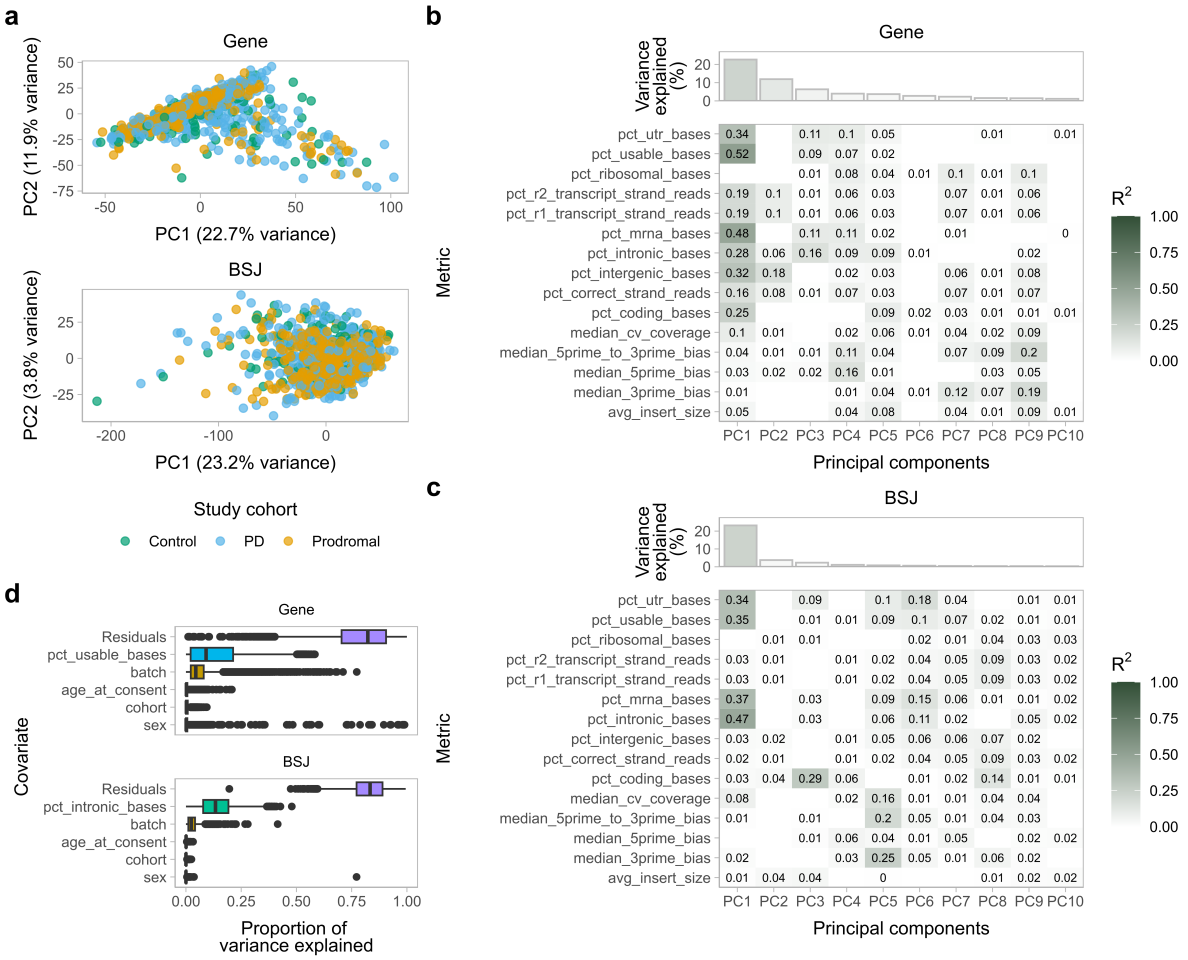


Figure 6.5. Sources of extraneous biological and technical variation when quantifying gene and circular RNA expression in PPMI samples. (a) Scatter plot of the first two PCs of gene and circRNA (BSJ) expression. Each point indicates a sample coloured according to the study cohort. **(b, c)** Results from univariate linear regression of technical sequencing metrics against each of the first 10 PCs of gene (**b**) and BSJ (**c**) expression. The top panel of each figure shows the percentage of variation explained by each PC. The amount of variation explained (R^2) by each metric is only shown when it passed multiple testing correction (FDR <0.05). **(d)** Contribution of covariates to individual gene and BSJ expression. Each point represents a single gene or BSJ, with the proportion of variation explained by the covariates shown on the x-axis.

6.3.2 Circular RNA expression in genetic PD

Chapters 4 and 5 focused on idiopathic PD. To determine if the changes in global circRNA expression I observed were detectable in genetic PD (i.e., where PD patients carried a genetic risk allele), I compared BSJ and FSJ expression between genetic PD patients and controls.

Based on BSJ and FSJ fold change differences between *LRRK2*-PD and *GBA*-PD PD patients to controls, circRNA expression was globally reduced (*LRRK2*-PD imbalance = 0.35, 95% CI: 0.33-0.38, *GBA*+ imbalance = 0.39, 95% CI: 0.37-0.42) (**Figure 6.6a, b**), consistent with earlier observations in iPD (**Chapter 5**). However, unlike idiopathic PD patients, linear RNA

expression was globally increased in both comparisons in *LRRK2*-PD and *GBA*-PD patients (*LRRK2*-PD imbalance = 0.58, 95% CI: 0.55-0.60, *GBA*-PD imbalance = 0.54, 95% CI: 0.51-0.56) (**Figure 6.6a, b**). Comparing BSJ and FSJ fold changes confirmed differences between circRNA and linear RNA expression changes in genetic PD patients compared to controls (*P*-value < $2.2x10^{-16}$, Kolmogorov–Smirnov test) (**Figure 6.6c**). FSJ fold changes could explain 12% and 23% of the variation in BSJ fold changes in *LRRK2*-PD and *GBA*-PD patients respectively (**Figure 6.6d**). Overall, these findings show that, like iPD patients, circRNA is also globally reduced in *LRRK2*-PD and *GBA*-PD patients.

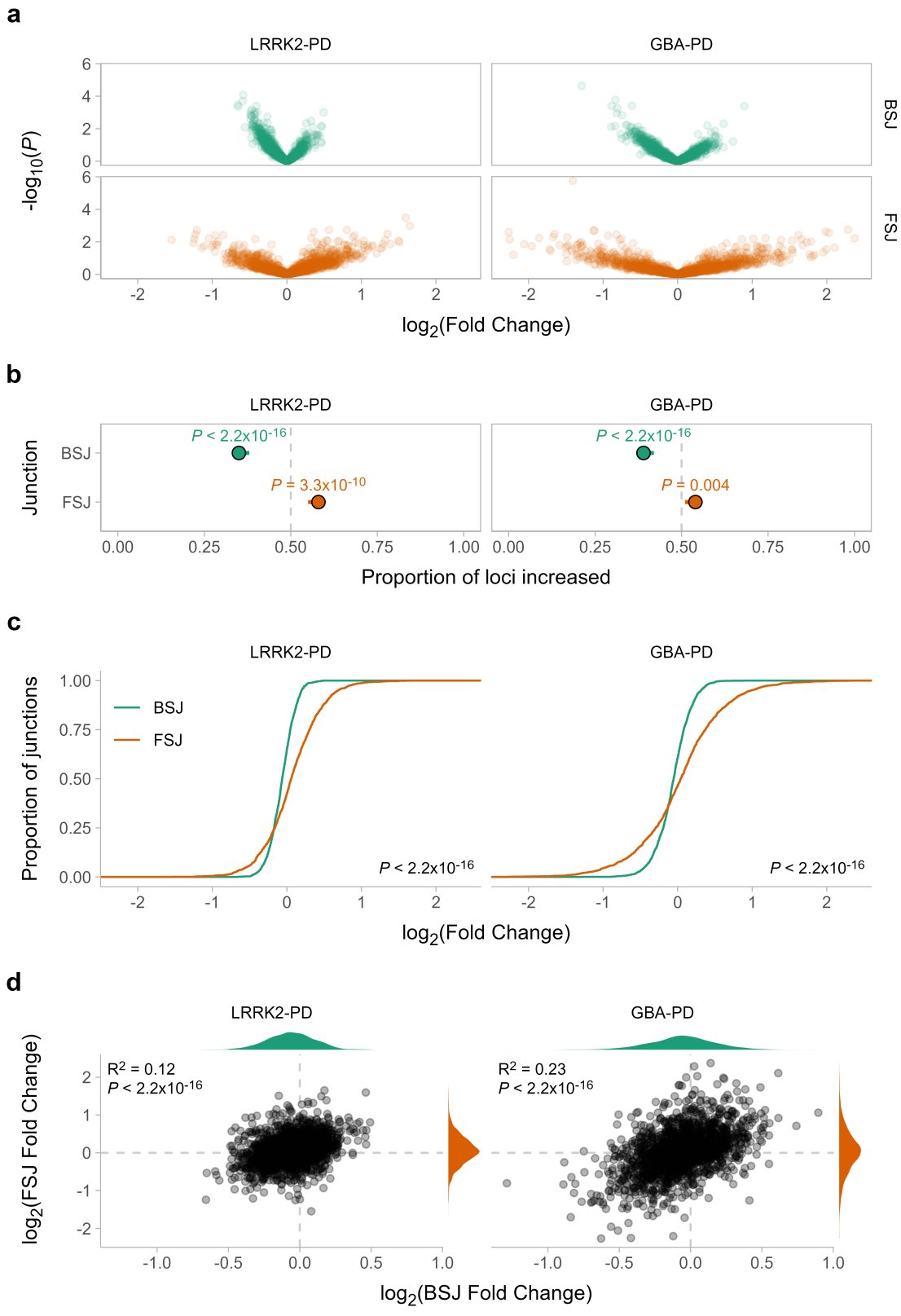


Figure 6.6. Globally reduced circular RNA expression in genetic PD patients. (a) Volcano plot showing the comparison of circRNA (BSJ) and linear RNA (FSJ) expression between patients with genetic PD and controls stratified by pathogenic gene (*LRRK2* or *GBA*). **(b)** The proportion of loci (both BSJs and FSJs) increased in genetic PD patients relative to controls. *P*-values obtained from a two-sided exact binomial test corrected for multiple testing (Bonferroni correction, two tests). **(c)** Cumulative distribution of BSJ and FSJ fold changes. *P*-values derived from a Kolmogorov-Smirnov test. **(d)** Correlation between BSJ and cognate FSJ fold changes.

6.3.3 Circular RNA expression in individuals at increased risk of developing PD

Individuals at a heightened risk of developing PD were enrolled into the PPMI Prodromal study cohort, enabling the monitoring of potential biomarkers throughout disease trajectory. After comparing BSJ fold changes, *LRRK2*-Controls (imbalance = 0.39, 95% CI: 0.37-0.41), *GBA*-Controls (imbalance = 0.15, 95% CI: 0.13-0.16) and individuals with hyposmia (imbalance = 0.23, 95% CI: 0.21-0.25) all showed varying but consistent reductions in global circRNA expression (**Figure 6.7a, b**). Of the prodromal subgroups, only individuals with RBD showed a global increase in circRNA expression (imbalance = 0.61, 95% CI: 0.59-0.64), combined with a concomitant global increase in the cognate linear RNA expression (imbalance = 0.61, 95% CI: 0.59-0.63) (**Figure 6.7a, b**). BSJ and FSJ fold changes showed varied distributions in all comparisons (*P*-value < $2.2x10^{-16}$, Kolmogorov–Smirnov test) (**Figure 6.7c**). For all prodromal subgroups, FSJ fold changes explained between 11-20% of the variance in BSJ fold changes (**Figure 6.7d**), again highlighting the unique expression changes in circRNA compared to linear RNA. Together, these results indicate that individuals with pathogenic *LRRK2* variants, pathogenic *GBA* variants or hyposmia exhibit a global reduction in circRNA expression that is not solely driven by a reduction in linear RNA expression. Therefore, in certain PD risk factor groups, reduced circRNA expression may reflect systematic responses to neurodegeneration.

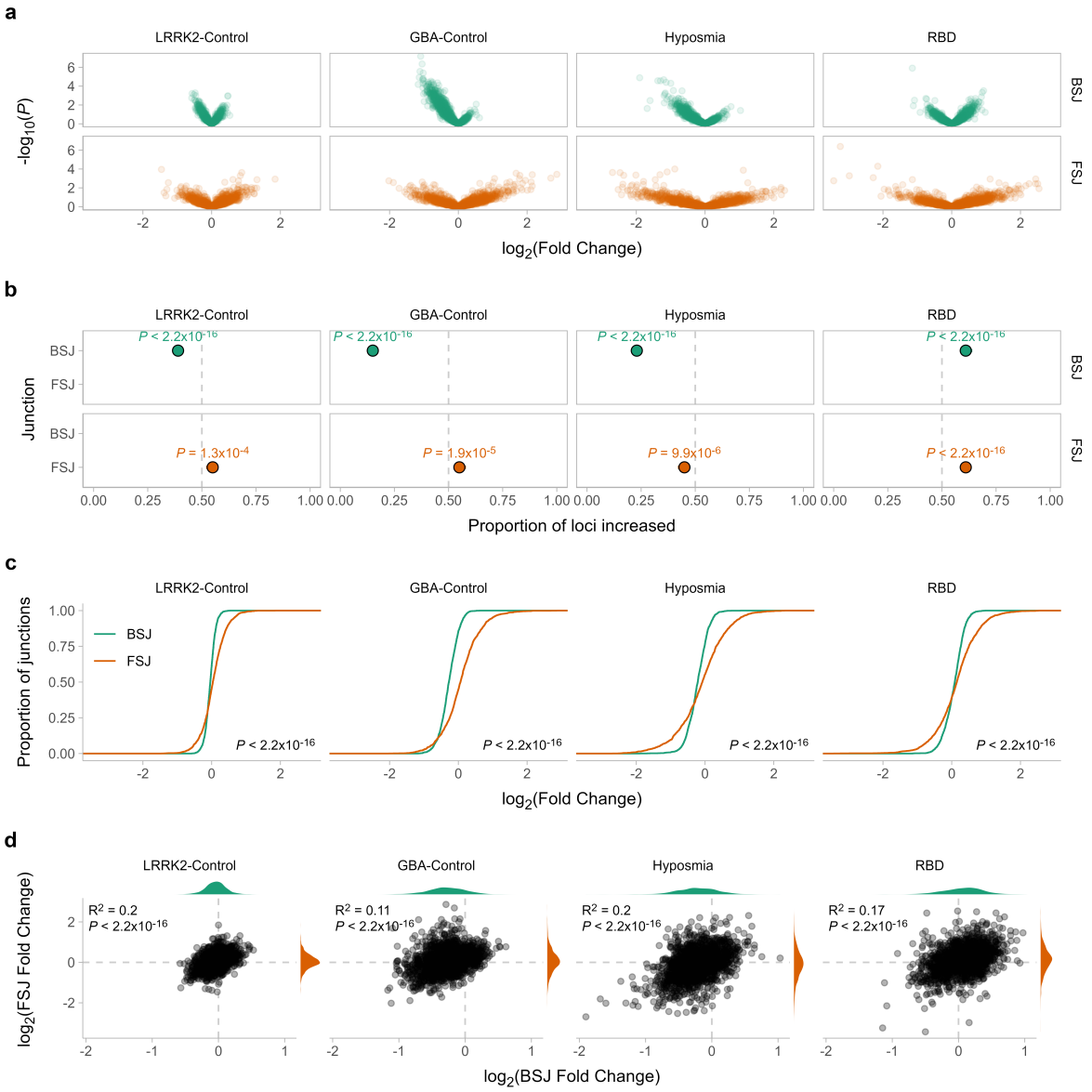


Figure 6.7. Globally reduced circular RNA expression in at-risk individuals when stratified by PD risk factor. (a) Volcano plot showing the comparison of circRNA (BSJ) and linear RNA (FSJ) expression between subgroups of individuals at increased risk of developing PD (*LRRK2*-Control, *GBA*-Control, Hyposmia, RBD) and controls. **(b)** The proportion of loci (both BSJs and FSJs) increased in at-risk individuals relative to controls. *P*-values obtained from a two-sided exact binomial test corrected for multiple testing (Bonferroni correction, two tests). **(c)** Cumulative proportions of BSJ and FSJ fold changes for each subgroup comparison. *P*-values derived from a Kolmogorov-Smirnov test. **(d)** Correlation between BSJ and the cognate FSJ fold changes for each subgroup comparison. RBD = REM (rapid eye movement) sleep behaviour disorder.

6.3.4 Gene expression changes in PPMI subgroups that exhibit reduced circular RNA expression

Previously I implicated known innate immune regulators of circRNA expression, such as ADAR1 and RNase L, as potential drivers of reduced circRNA expression in iPD (**Chapter 5**). The comparative reductions in circRNA expression in genetic PD patients and among

individuals with specific PD-risk factors raise the hypothesis that *ADAR* (ADAR1) and *RNASEL* (RNase L) expression may also be altered in these individuals. *ADAR* expression was significantly increased in *LRRK2*-PD and *GBA*-PD patients relative to controls (*P*-values < 0.05, moderated *t*-test) (**Figure 6.8a**). *RNASEL* expression was also significantly increased in *LRRK2*-PD and *GBA*-PD patients (*P*-values < 0.05, moderated *t*-test) (**Figure 6.8a**). Contrary to iPD patients (**Chapter 5**), there were no significant expression differences in the gene encoding PKR (*EIF2AK2*), in *LRRK2*-PD or *GBA*-PD patients compared to controls (*P*-values > 0.05, moderated *t*-test) (**Figure 6.8a**). In prodromal subgroups that exhibit reduced circRNA expression (*LRRK2*-Control, *GBA*-Control, Hyposmia), increased expression of *ADAR* and *RNASEL* was also observed but did not reach significance (*P*-values > 0.05, moderated *t*-test) (**Figure 6.8a**).

Next, I employed gene set enrichment analysis (GSEA) of Hallmark gene sets (Liberzon et al., 2015) to identify commonly dysregulated processes across subgroups that exhibited globally reduced circRNA expression. Across all surveyed subgroups, no gene sets were commonly enriched (**Figure 6.8b**). In genetic PD patients, genes involved in inflammation (IL6_JAK_STAT3_SIGNALING, INFLAMMATORY_RESPONSE, TNFA_SIGNALING_VIA_NFKB) were significantly upregulated (**Figure 6.8b**). Interferon related gene sets (INTERFERON_ALPHA_RESPONSE, INTERFERON_GAMMA_RESPONSE) were significantly upregulated in *GBA*-PD patients. Among prodromal groups, interferon related gene sets (INTERFERON_ALPHA_RESPONSE, INTERFERON_GAMMA_RESPONSE) were significantly downregulated in *LRRK2*-Controls (**Figure 6.8b**). Overall, despite all subgroups exhibiting globally reduced circRNA expression, their gene expression profiles highlight considerable heterogeneity within subgroups.

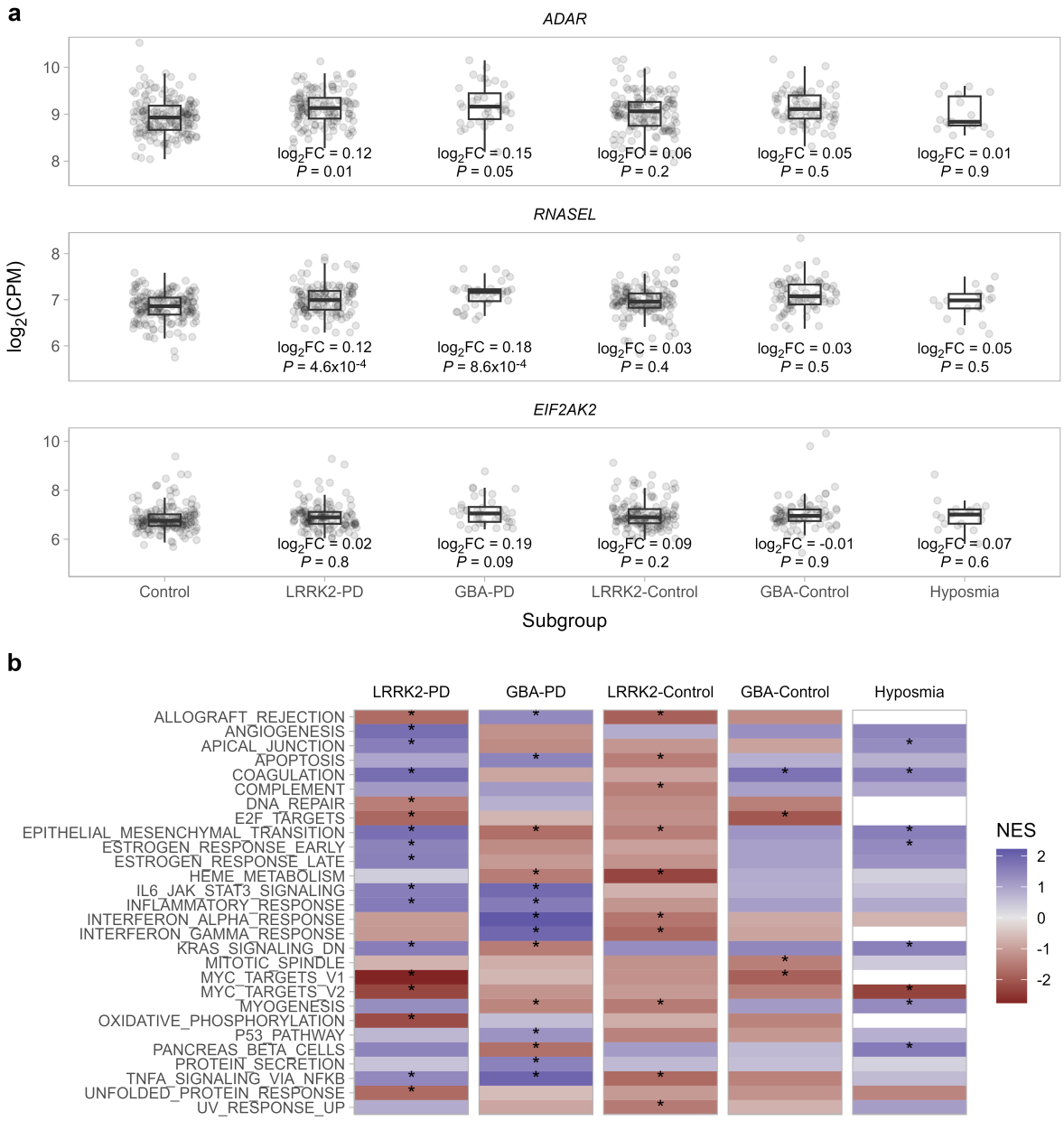


Figure 6.8. Increased expression of inflammatory gene signatures in PD patients harbouring *LRRK2* and *GBA* variants. (a) Differential expression of *ADAR* (encoding ADAR1), *RNASEL* (encoding RNase L) and *EIF2AK2* (encoding PKR) in subgroups that show globally reduced circRNA expression compared to controls. log_2FC = log_2 Fold Change. **(b)** Gene set enrichment analysis of Hallmark gene sets. All comparisons were made in comparison to control samples. * indicates FDR < 0.05 based on all Hallmark gene sets. White tile = NA. NES = Normalised Enrichment Score.

6.3.5 *Circular RNA detection and expression in BRAINcode samples*

Globally reduced circRNA expression in blood appeared to be a ubiquitous feature within PPMI PD patients, raising the question of whether circRNA expression in the periphery is reflected in the brain. To explore circRNA expression in the brain, I leveraged published total RNA sequencing of post-mortem substantia nigral tissue (X. Dong et al., 2018; Dong et al., 2023).

Laser-capture microdissection of vulnerable dopaminergic neurons allowed the profiling of circRNA expression at a single-cell resolution. In total, sequencing data were available for 104 samples comprising 18 PD cases, 27 incidental Lewy Body cases (ILB) and 59 controls (**Section 6.2.2**). Notably, ILB cases had neuropathological evidence of Lewy bodies at autopsy and are thought to reflect preclinical PD (DelleDonne et al., 2008; Dickson et al., 2008).

I processed BRAINcode samples through RNA quantification pipelines to produce estimates of gene, circRNA and cognate linear RNA expression. Overall, 936 unique circRNAs were detected across the samples including 479 (51.2%) that were detected across all PD, ILB and Control groups (**Figure 6.9a**).

Previously I have normalised against the number of transcriptome-mapped reads (**Chapters 4 and 5**) due to its consistent performance across cohorts and to ensure parity with the method of normalisation when performing differential expression. To confirm the utility of normalising against the number of transcriptome-mapped reads in BRAINcode samples, I correlated measures of sequencing depth (total sequenced reads, genome-mapped reads and transcriptome-mapped reads) against the number of unique circRNAs (unique BSJs) and the total amount of circRNA expression (total BSJ reads) in each sample. I did not identify any differences in measures of sequencing across the sample groups ($P > 0.05$, Kruskal-Wallis test) (**Figure 6.9b**), so were appropriate for normalisation. I found that all measures of sequencing depth were positively correlated to the total amount of circRNA expression in each sample (all P-values < 0.05, linear regression) (**Figure 6.9c**). The number of transcriptome-mapped and genome-mapped reads explained 37% of the variance in total circRNA expression in each sample. The number of transcriptome-mapped and genome-mapped reads were also correlated to the number of unique circRNAs detected in each sample (all P-values < 0.05, linear regression) (**Figure 6.9d**), albeit explained 7% and 5% of the variance respectively. These results indicate that normalising the number of circRNAs and their expression against the number of transcriptome-mapped reads was a suitable method in BRAINcode samples.

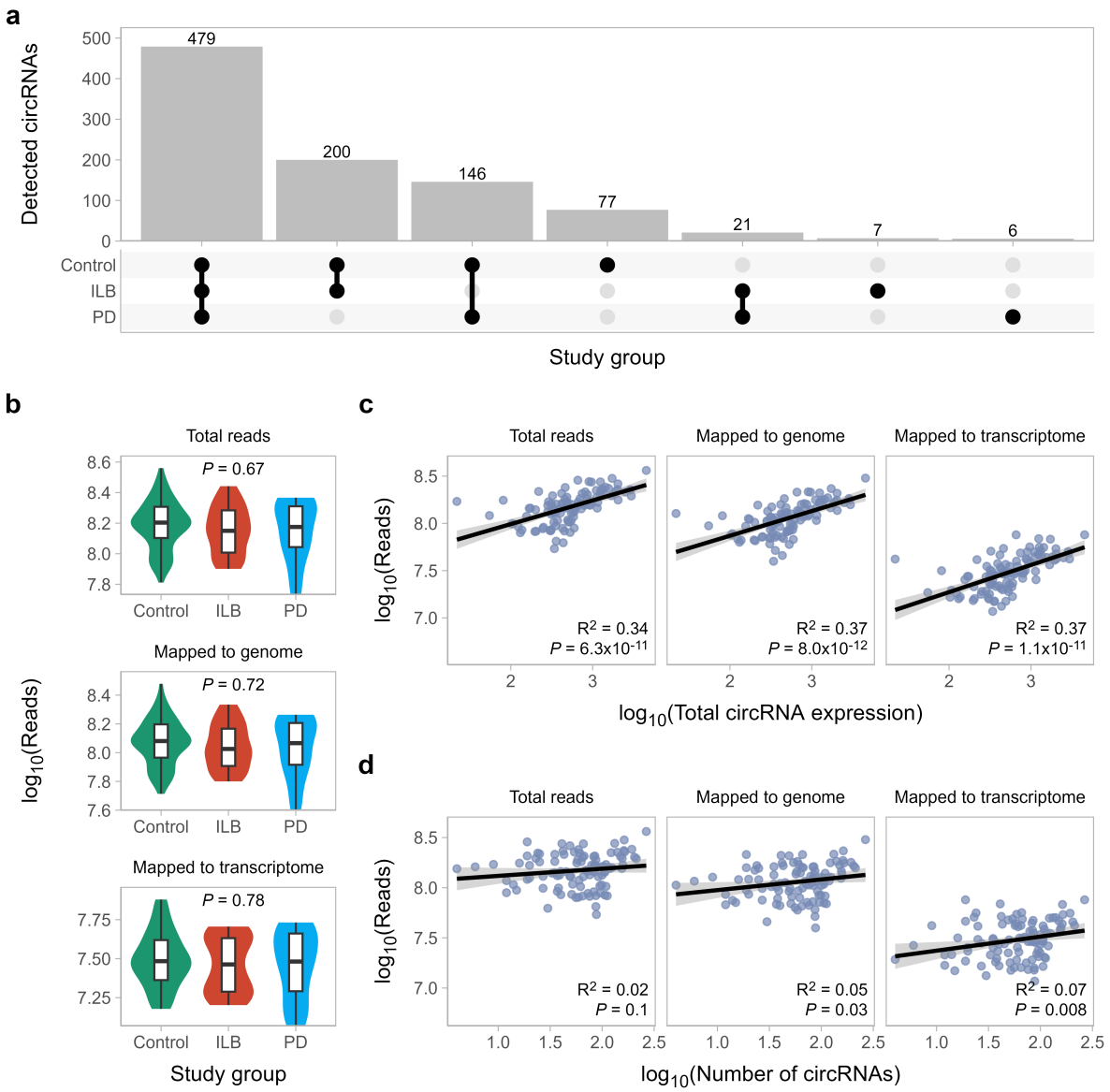


Figure 6.9. Circular RNA detection and normalisation in BRAINcode sequencing. (a) Intersection of detected circRNAs across PD, ILB and Control study groups. For each combination of study groups, the intersection of detected circRNAs is shown. The number above each bar shows the size of the intersection. **(b)** Comparison of sequencing depth across sample groups. *P* values calculated from Kruskal-Wallis tests. **(c, d)** Relationship of sequencing depth to total circRNA expression (**c**) and number of unique circRNAs in each sample (**d**). Shaded area indicates the 95% confidence interval of the linear regression line.

I detected a 20-fold reduction in the number of unique BSJs detected across BRAINcode samples (936) compared to PPMI samples (18,751). Whilst the PPMI has a larger sample size, the number of unique BRAINcode BSJs is still fewer than was detected in the similarly sized ICICLE-PD cohort (n = 96, 15,336 BSJs detected) (**Chapter 4**). In BRAINcode samples, BSJs also appeared to be not widely detected, with 99.7% of the 936 unique BSJs detected in less than half of all samples.

In their original publication, Dong *et al* identified 6,699 unique BSJs same BRAINcode samples described here, which may be due to differences in the circRNA detection algorithm used (Dong et al., 2023). To compare BSJ detection between the BSJs detected in this chapter and those in the associated publication, I collected summarised circRNA count tables provided with the deposited sequencing data (Gene Expression Omnibus ID: GSE218203). Among the 6,699 BSJs detected by Dong et al, 99.9% were detected in less than half the samples. The BSJs detected in my analysis (median percentage of cohort = 5.8%) were more widely detected across the cohort (*P*-value < $2.2x10^{-16}$, Wilcoxon rank-sum test) compared to Dong et al (median percentage of cohort = 0.96%), likely due to the inclusion of rarer BSJs in the original publication. Despite the higher number of absolute BSJs reported by Dong et al, the normalised number of BSJs detected (per million transcriptome-mapped reads) were highly correlated across analyses (Spearman's rho = 0.96, *P*-value < 2.2x10-16) (**Figure 6.10c**).

The presence of rRNA can impair circRNA detection by sequestering sequencing capacity (Nielsen et al., 2022). Subsequently, I found a negative correlation between the number of mapped rRNA bases in each sample and the number of normalised BSJs detected, both for the BSJs I detected (Spearman's rho = -0.42, *P*-value = $9.6x10^{-6}$) and those identified by Dong et al (Spearman's rho = -0.35, *P*-value = $2.5x10^{-4}$) (**Figure 6.10d**). This suggests that the presence of rRNA may have impaired circRNA detection. There were no significant differences in the relative amounts of mapped rRNA bases across sample groups (*P*-value = 0.11, Kruskal-Wallis test) (**Figure 6.10e**).

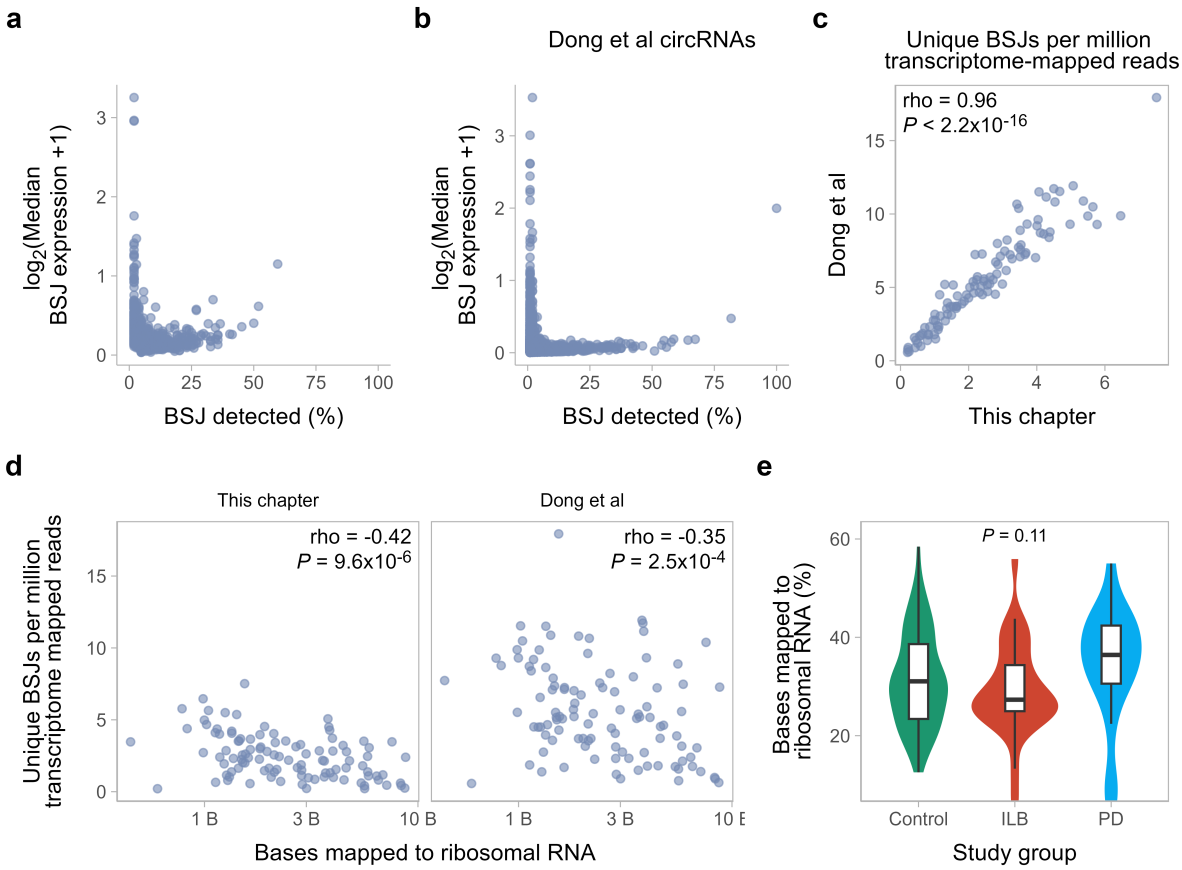


Figure 6.10. Impaired circular RNA detection in BRAINcode samples due to presence of ribosomal RNA. (a) Detection of 936 unique circRNAs in BRAINcode samples. **(b)** Detection of 6699 unique circRNAs as previously reported in Dong et al (2023). CircRNA expression was provided as "*normalized reads per million (RPM) at the back-splicing sites for each sample*" (Dong et al., 2023). **(c)** Correlation between the number of unique circRNAs detected in each sample as reported by Dong et al and in this chapter. Correlation reported as Spearman's rho based on 104 samples. **(d)** Correlation between the number of unique circRNAs in each sample and the number of bases mapped to ribosomal RNA loci. Correlation reported as Spearman's rho. **(e)** Percentage of bases mapped to ribosomal RNA loci across sample groups. *P*-value derived from a Kruskal-Wallis test.

Given previously reported differences in total circRNA abundance found in bulk substantia nigra tissue from PD patients (Hanan et al., 2020), I investigated how total circRNA abundance may be altered within dopaminergic neurons. I did not detect any variation in the number of unique BSJs (normalised to the number of transcriptome-mapped reads in each sample) across sample groups (*P*-value = 0.8, Kruskal-Wallis test). There was, however, variation in the total expression of BSJs (normalised to the number of transcriptome-mapped reads in each sample) across sample groups (*P*-value = 1.5×10^{-4}, Kruskal-Wallis test) (**Figure 6.11**). When comparing individual study groups, I found that PD samples had a greater number of total normalised BSJ reads than control (Holm corrected *P*-value = 7.0×10^{-4}, Dunn's test) or ILB samples (Holm corrected *P*-value = 1.6×10^{-4}, Dunn's test) (**Figure 6.11**). I did not detect any changes in total

normalised BSJ reads in ILB samples compared to controls (Holm corrected *P*-value = 0.3, Dunn's test) (**Figure 6.11**). Overall, these results suggest that dopaminergic neurons located in the substantia nigra of PD patients have increased expression of circRNAs.

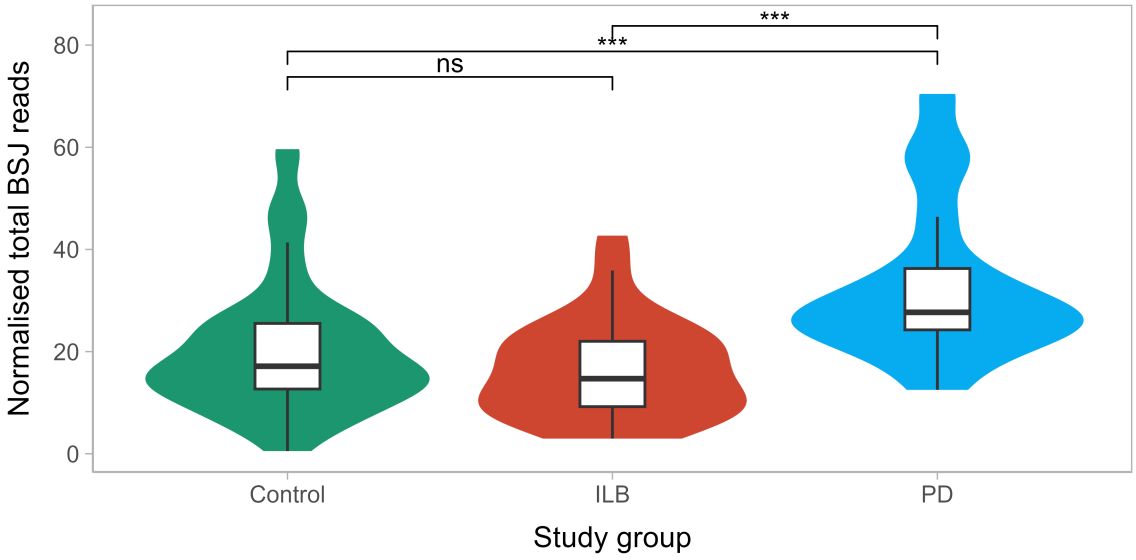


Figure 6.11. Increased total circular RNA expression in BRAINcode PD samples. Total BSJ reads were normalised against the number of transcriptome-mapped reads in each sample. Post-hoc pairwise comparisons were performed using Dunn's test with *P*-values corrected using Holm's method. ns = *P*-value > 0.05, *** = *P*-value < 0.001.

I projected and visualised gene and circRNA expression into the first and second PCs, finding no clear segregation of individuals by study group (**Figure 6.12a**), suggesting other sources of variation are contributing to RNA expression. To assess the impact of technical sequencing factors on RNA expression in BRAINcode samples, I correlated sequencing metrics with the first ten expression principal components of gene and circRNA expression (**Section 6.2.7**). Cumulatively, the first ten PCs of gene and circRNA expression explained 42.5% and 25.9% of the variance respectively (**Figure 6.12b, c**). PC1 of gene and circRNA expression was significantly associated with several technical sequencing metrics (FDR < 0.05, linear regression) (**Figure 6.12b, c**). For gene expression, the median coefficient of variance of coverage explained 54% of the variance of PC1, possibly reflecting variability in the coverage across genes. For circRNA expression, no technical metric explained >50% of the variation in any of the first ten PCs. The proportion of variance explained by circRNA PC1 (5.7%, **Figure 6.12a, c**) was lower than observed in PPMI samples (23.2%, **Figure 6.5a, c**), suggesting that technical variation had less influence on BRAINcode circRNA quantification and as such no technical factors were flagged.

I then included select metrics and sequencing batch, coupled with covariates used in a previous analysis of the BRAINcode samples (sex, age, post-mortem interval, RNA integrity number) (Dong et al., 2023), in linear mixed models to estimate their contribution to the variance in individual gene and circRNA expression (**Section 6.2.7**, **Figure 6.5d**). When ranked according to the median proportion of gene and circRNA expression variance explained, the included covariates accounted for more variance than condition (representing the study group; PD, ILB or control), justifying their inclusion when modelling RNA expression (**Figure 6.5d**).

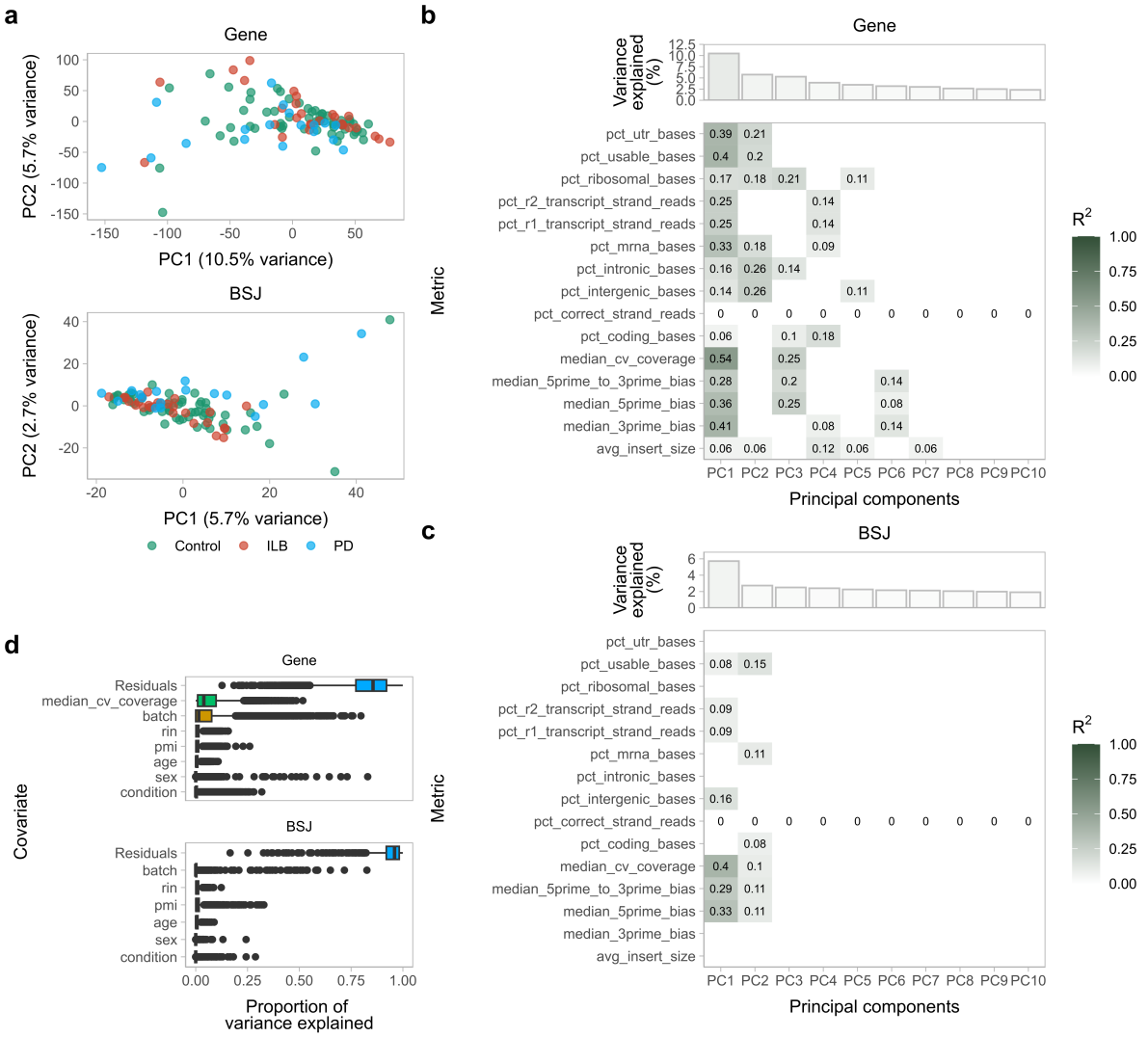


Figure 6.12. Identifying sources of extraneous technical variation when quantifying circRNA expression in BRAINcode samples. (a) Scatter plot of the first two PCs of gene (top panel) and BSJ (circRNA) expression (bottom panel). Each point indicates a sample coloured according to the study cohort. **(b, c)** Results from univariate linear regression of technical sequencing metrics against each of the first 10 PCs of gene (**b**) and BSJ (**c**) expression. The top panel of each figure shows the percentage of variation explained by each PC. The amount of variation explained (R^2) by each metric is only shown when it passed multiple testing correction (FDR <0.05). **(d)** Contribution of covariates to individual gene (top panel) and BSJ (bottom panel) expression. Each point represents a single gene or circRNA, with the proportion of variation explained by the covariates shown on the x-axis. PCs = Principal Components.

6.3.6 Circular RNA expression in dopaminergic neurons

To assess global differences in circRNA expression between PD and ILB samples to controls, I performed differential expression of both BSJ and FSJ counts from 241 loci (**Figure 6.13a**). Based on BSJ fold changes, I found opposing directions of change including a global reduction in circRNA expression in ILB samples (imbalance = 0.33, 95% CI: 0.27-0.39) and a global increase of circRNA expression in PD samples (imbalance = 0.62, 95% CI = 0.55-0.68) (**Figure**

6.13a, b). In ILB samples, I saw a comparative reduction in linear RNA expression (imbalance = 0.31, 95% CI: 0.25-0.37), yet no imbalance in linear RNA expression for PD samples (imbalance = 0.52, 95% CI: 0.45-0.58) (**Figure 6.13a, b**). In both comparisons, the cumulative distribution of circRNA and linear RNA expression changes varied (ILB *P*-value = 4.5x10-4, PD *P*-value = 9.4x10-5, Kolmogorov–Smirnov test) (**Figure 6.13c**). Together, these results support earlier findings (**Figure 6.11**) that circRNA expression in globally increased in dopaminergic neurons isolated from PD patients. In this case however, the reduction in circRNA expression in dopaminergic neurons isolated from ILB samples cannot be uncoupled from a concurrent reduction in linear RNA expression.

I observed a stronger correlation between changes in circRNA and linear RNA expression in BRAINcode samples (ILB R^2 = 0.74, PD R^2 = 0.64, *P*-values < $2.2x10^{-16,}$ linear regression) compared to PPMI samples (**Figure 6.6d**, **Figure 6.7d**). One possibility is that changes in circRNA expression reflected differences in the detection of individual circRNAs.

To resolve potential detection differences across study groups and the observed association between circRNA and linear RNA expression in ILB samples, I used the circular:linear ratio of each circRNA to normalise against the cognate linear RNA background, similar to previous work (Ma et al., 2019). The circular:linear ratios varied across study groups (*P*-value = 1.6x10-16, Kruskal-Wallis test) (**Figure 6.13e**). Specifically, circular:linear ratios were increased in PD patients compared to both ILB (Holm corrected *P*-value = $2.4x10^{-9}$, Dunn's test) and control samples (Holm corrected *P*-value = $2.4x10^{-9}$, Dunn's test) (**Figure 6.13e**). In comparison, there was no significant difference in circular:linear ratios between ILB and control samples (Holm corrected *P*-value = 0.3, Dunn's test) (**Figure 6.13e**). These findings suggest that ILB samples did not exhibit reduced circRNA expression relative to the comparative linear RNA expression and confirmed increased circRNA expression in PD patients disparate from changes in linear RNA expression.

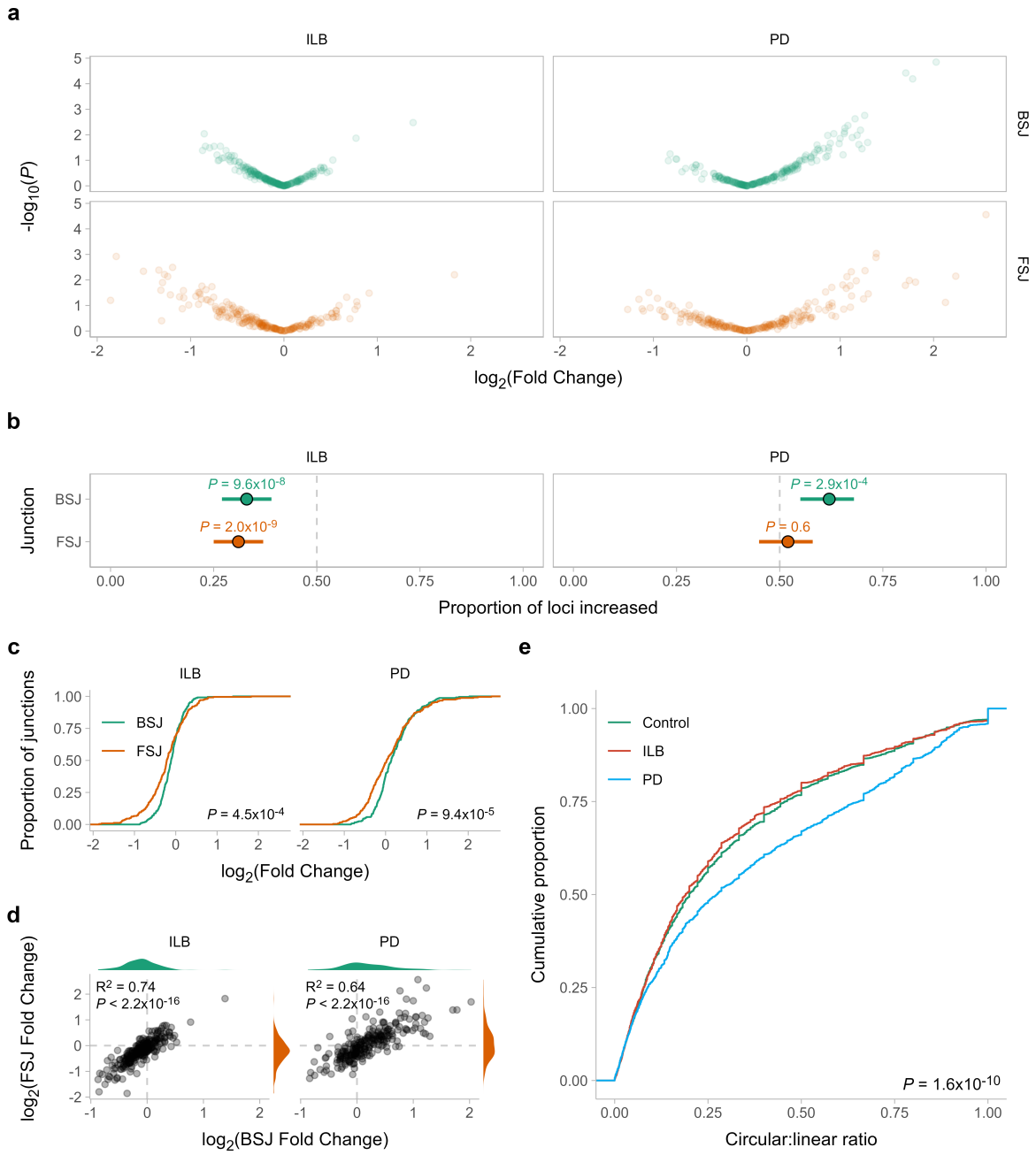


Figure 6.13. Globally increased circular RNA expression in dopaminergic neurons from PD patients. (a) Volcano plot showing the comparison of circRNA (BSJ) and linear RNA (FSJ) expression between ILB and PD samples relative to controls. **(b)** The proportion of loci (both BSJs and FSJs) increased in ILB or PD samples compared to controls. *P*-values obtained from a two-sided exact binomial test corrected for multiple testing (Bonferroni correction, two tests). **(c)** Cumulative distribution of BSJ and FSJ fold changes. *P*-values derived from a Kolmogorov-Smirnov test. **(d)** Correlation between BSJ and cognate FSJ fold changes. **(e)** Cumulative distribution of the circular:linear ratio of all 936 circRNAs detected in BRAINcode samples. Distributions are shown for each study group separately. *P*-value calculated from a Kruskal-Wallis test.

As circRNA expression was globally increased in PD patients, I hypothesised that expression of the previously described circRNA regulators ADAR1 (*ADAR*) and RNase L (*RNASEL*)

(**Chapter 5**) may be altered in these samples. However, I did not identify any significant change in *ADAR*, *RNASEL* or *EIF2AK2* (PKR) expression in ILB or PD samples when compared to controls (*P*-values > 0.05, moderated *t*-test) (**Figure 6.14a**).

To identify systematic gene expression changes in ILB and PD samples, I performed GSEA of Hallmark gene sets (**Section 6.2.9**). Dysregulated gene sets were largely distinct across ILB and PD samples (**Figure 6.14b**). Commonly dysregulated gene sets among ILB and PD samples included genes involved in allograft rejection (ALLOGRAFT_REJECTION) and KRAS signalling (KRAS_SIGNALING_DN, KRAS_SIGNALING_UP) (**Figure 6.14b**). Compared to controls, several gene sets related to proinflammatory processes (COMPLEMENT, IL6_JAK_STAT3_SIGNALING, INFLAMMATORY RESPONSE, TNFA_SIGNALING_VIA_NFKB) and innate antiviral responses (INTERFERON_ALPHA_RESPONSE, INTERFERON_GAMMA_RESPONSE) were upregulated in PD patients (**Figure 6.14b**), reflecting previously identified changes in the blood of early-stage idiopathic PD patients (**Chapter 5**).

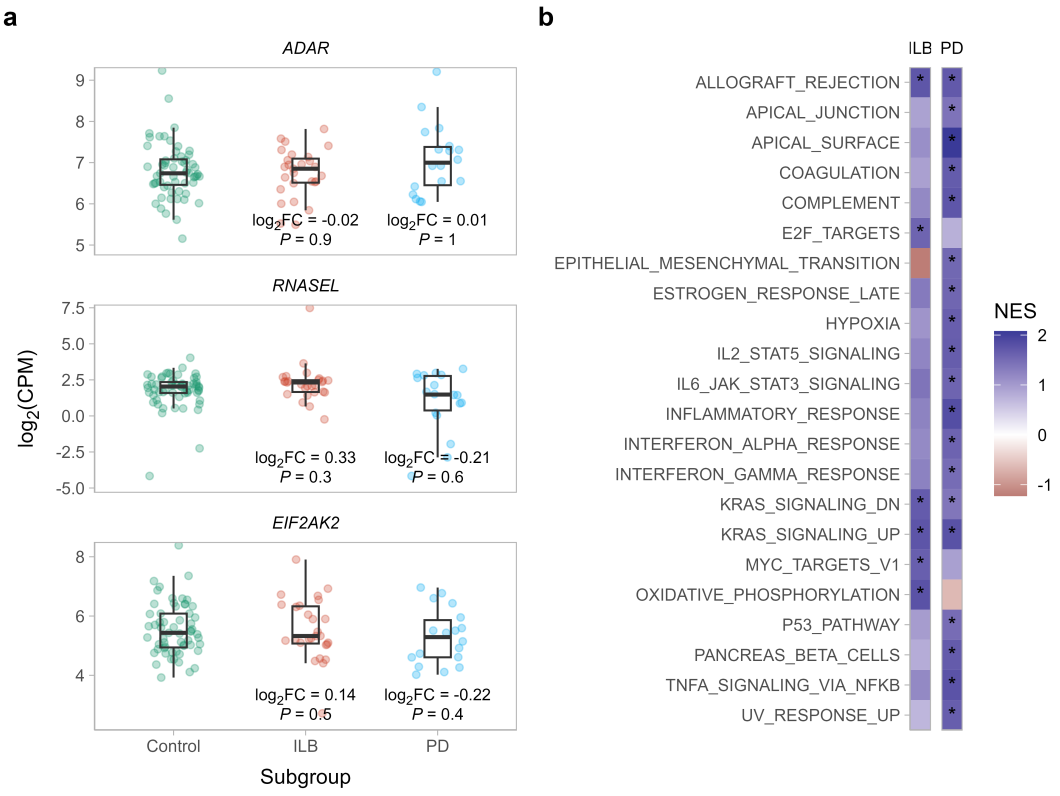


Figure 6.14. Distinct processes altered in dopaminergic neurons from individuals with incidental Lewy body pathology and pathologically confirmed PD. (a) Expression of *ADAR* (encoding ADAR1), *RNASEL* (encoding RNase L) and *EIF2AK2* (encoding PKR) in individuals converted to PD compared to controls. Fold changes and *P*-values were calculated using limma. $log_2FC = log_2$ Fold Change. **(b)** Gene set enrichment analysis of Hallmark gene sets. All comparisons were made in comparison to control samples. Only gene sets which are significantly enriched are shown (* = FDR < 0.05). NES = Normalised Enrichment Score.

6.4 Discussion

CircRNA expression varies based on the cell or tissue and their responses to physiological changes (Yang et al., 2022). Alterations in global circRNA expression can signal underlying mechanistic changes, such as the activation of an innate antiviral response (C.-X. Liu et al., 2019). Global changes in circRNA expression are increasingly reported in a range of diseases. (Bachmayr-Heyda et al., 2015; C.-X. Liu et al., 2019; Moldovan et al., 2019, 2021; Seeler et al., 2022; Fuchs et al., 2023; Guo et al., 2024). In this chapter, I leveraged total RNAseq data from the PPMI cohort demonstrating how global circRNA expression is reduced in the blood of PD patients harbouring genetic risk variants in *GBA* or *LRRK2*. I further demonstrated how a reduction in circRNA expression can also be detected in groups of individuals that possess specific PD risk factors such as genetic risk variants in LRRK2 or GBA and hyposmia without a PD diagnosis. Conversely, global circRNA expression was increased in dopaminergic neurons from PD patients.

6.4.1 Circular RNA expression is globally reduced in the blood of genetic PD patients

My identification of reduced circRNA expression in idiopathic PD patients (**Chapters 4 and 5**) raised the question of whether reduced circRNA expression would be a feature in PD patients harbouring risk alleles. PD patients harbouring pathogenic *LRRK2* and *GBA* variants also showed reduced circRNA expression in the periphery (**Figure 6.6**). Due to different patient demographics in genetic PD patients enrolled in PPMI (**Figure 6.2**), the reduction in circRNA expression appears to be present in a wider range of disease duration and considering treatment status. Reduced circRNA expression in the blood of idiopathic, *LRRK2*-PD and *GBA*-PD patients indicates there may be a common mechanism leading to reduced circRNA expression.

Reduced circRNA expression within the blood of PD patients is notable considering established phenotypic and pathological heterogeneity among PD patients, including those specifically harbouring *GBA* and *LRRK2* variants (Greenland et al., 2019). In PPMI PD patients, the detection of pathological α-synuclein in cerebrospinal fluid varied among subgroups, with worse sensitivity reported in *LRRK2*-PD patients compared to GBA/idiopathic PD patients (Siderowf et al., 2023). Earlier work also reported variable postmortem presentations of Lewy body pathology in *LRRK2*-PD (Kalia et al., 2015). Identification of the causal mechanism driving a reduction in circRNA expression provides an opportunity to identify novel blood-based markers of PD. One limitation of this work is that global circRNA expression was assessed in groups of PD patients. Reduced circRNA expression may be concentrated in a

subset of PD patients, driving changes in the whole group. Identification of PD subgroups based on global circRNA expression may inform future disease stratification.

6.4.2 Circular RNA expression is globally reduced in the blood of individuals with specific PD risk factors

I found that asymptomatic *LRRK2* and *GBA* carriers show reduced global circRNA expression compared to controls (**Figure 6.7**). Motor and non-motor changes have previously been reported in asymptomatic *LRRK2* and *GBA* carriers enrolled in the PPMI in addition to elevated urinary bis(monoacylglycerol) phosphate levels compared to controls (Simuni, Uribe, et al., 2020; Merchant et al., 2023). Other studies have also described neurophysiological alterations in asymptomatic *LRRK2* and *GBA* mutation carriers (Droby et al., 2022).

I also found reduced circRNA expression in individuals with hyposmia (**Figure 6.7**). Olfactory dysfunction is common in PD, with over 96% of PD patients experiencing olfactory loss according to one multicentre study (Haehner et al., 2009). Although the mechanism is debated, Lewy body pathology has been observed in the olfactory bulb before motor symptoms appear (Braak et al., 2002; Braak, Del Tredici, et al., 2003; Beach et al., 2009). Therefore, reduced circRNA expression in individuals with hyposmia may indicate early stages of Lewy body pathology.

Of all PD risk factor subgroups, only individuals with RBD did not show reduced circRNA expression, in fact demonstrating a global increase in circRNA expression (**Figure 6.7**). RBD is a frequent feature of PD, with one study showing more than 86% of patients had RBD (Sixel-Döring et al., 2023). RBD is considered a risk factor not only for PD but also for other synucleinopathies, such as dementia with Lewy bodies and multiple system atrophy (Postuma, Gagnon, et al., 2015; Postuma et al., 2019). Therefore, the different global circRNA expression changes in the RBD subgroup compared to other prodromal subgroups, might indicate the development of synucleinopathies other than PD. Although circRNAs have been studied in patients with Lewy body dementia and multiple system atrophy (Chen et al., 2016; Puri et al., 2023), the global circRNA expression patterns in the blood of patients with such disorders are currently unexplored. In PD, the onset of RBD varies in disease progression; in some patients, RBD is present years before diagnosis, while in others RBD may manifest towards later disease stages (Sixel-Döring et al., 2023; Cicero et al., 2023). It is hypothesised that RBD is associated with body-first PD, in which pathological α-synuclein first aggregates in the enteric or peripheral nervous systems before propagating to the central nervous system (Borghammer & Van Den Berge, 2019). Differences in global circRNA expression patterns among different PD

risk subgroups may therefore provide an insight into the origin of reduced circRNA expression. Increased circRNA expression in individuals with RBD may indicate that reduced circRNA expression in other PD risk subgroups may be driven by a response to neuroinflammation rather than peripheral inflammation.

6.4.3 Heterogeneity among gene expression changes between groups that exhibit reduced circular RNA expression

This chapter established that reduced blood circRNA expression is a feature of genetic PD. Reduced global circRNA expression has previously been reported in several cancers, systemic lupus erythematosus, psoriasis and atopic dermatitis (Bachmayr-Heyda et al., 2015; C.-X. Liu et al., 2019; Moldovan et al., 2019, 2021; Seeler et al., 2022; Fuchs et al., 2023; Guo et al., 2024). Inflammation is a common theme among diseases associated with reduced circRNA expression such as atopic dermatitis, psoriasis and systemic lupus erythematosus (C.-X. Liu et al., 2019; Moldovan et al., 2019; Seeler et al., 2022; Moldovan et al., 2021; Guo et al., 2024). In systemic lupus erythematosus, the associated inflammation was linked to aberrant PKR activation due to diminished inhibition resulting from decreased circRNA expression (C.-X. Liu et al., 2019). Preclinical studies have explored using circRNA to suppress PKR in systemic lupus erythematosus, psoriasis and Alzheimer's disease (C.-X. Liu et al., 2019; Liu et al., 2022; Guo et al., 2024; Feng et al., 2024).

Previously I identified upregulated expression of *RNASEL* and *EIF2AK2* (PKR) in the blood of early-stage idiopathic PD patients (**Chapter 5**). Altered *RNASEL* and *EIF2AK2* expression in combination with coordinated changes in related genes, implicated the activation of innate antiviral and inflammatory response in the blood of early-stage idiopathic PD patients (**Chapter 5**). Consistent with known circRNA regulation (C.-X. Liu et al., 2019), *RNASEL* expression was increased in PD patients harbouring either *LRRK2* or *GBA* risk alleles. I did not detect any changes in *RNASEL* expression in PD risk subgroups that showed reduced circRNA expression such as asymptomatic *LRRK2/GBA* risk allele carriers or individuals with hyposmia. *RNASEL* levels are increased in certain physiological contexts (Li et al., 2007), suggesting that sustained activation of an innate antiviral response due to PD pathology may lead to an increase in steady state levels of *RNASEL* over time. Despite this, I did not detect any significant changes in *EIF2AK2* expression in *LRRK2*-PD, *GBA*-PD, *LRRK2*-Controls, *GBA*-Controls, or individuals with hyposmia. I also detected increased expression of *ADAR* in *LRRK2*-PD and *GBA*-PD patients (**Figure 6.8a**), which may reflect increased A-I editing as shown in early-stage idiopathic PD PPMI patients (**Chapter 5**).

Genes associated with inflammation were upregulated in *LRRK2*-PD and *GBA*-PD patients (**Figure 6.8b**). Several strands of evidence link *LRRK2* and *GBA* to immune activation and inflammation. Common variants in *LRRK2* and *NOD2* are associated with both PD and inflammatory bowel disease (Huang et al., 2017; Nalls et al., 2019). LRRK2 and GCase have both been implicated in immune regulation (Liu et al., 2012; Wallings & Tansey, 2019). Elevated circulating peripheral inflammatory proteins have been reported in both *LRRK2*-PD and *GBA*-PD patients compared to controls (Chahine et al., 2013; Brockmann et al., 2016; Dzamko et al., 2016), with increased inflammatory markers also found in asymptomatic LRRK2 carriers (Dzamko et al., 2016). However, a recent study did not find altered levels of inflammatory cytokines in the blood or cerebrospinal fluid of symptomatic and asymptomatic *LRRK2* and *GBA* carriers (Thaler et al., 2021). In asymptomatic *LRRK2* or *GBA* carriers, there was no significant upregulation of genes related to inflammatory responses (**Figure 6.8**) in contrast to symptomatic *LRRK2* or *GBA* carriers (**Figure 6.8**). Thus, the induction of these genes may occur later in the disease, after motor symptoms appear.

6.4.4 Circular RNA expression is increased in dopaminergic neurons isolated from the substantia nigra of PD patients

Previous work identified varied circRNA abundance in PD patients in a brain-region specific manner (Hanan et al., 2020). In dopaminergic neurons isolated from PD samples, I found a global increase in circRNA expression (**Figure 6.11, Figure 6.13a, b**), which was maintained after normalising to cognate linear RNA expression (**Figure 6.13d**). This finding contrasts with earlier findings in the blood of idiopathic (**Chapter 5**) and genetic PD patients (**Figure 6.6**). This finding also contrasted with previous work reporting reduced circRNA abundances in bulk tissue derived from the substantia nigra of PD patients (Hanan et al., 2020). However, this study did find increased circRNA abundance in the amygdala and medial temporalis gyrus of PD patients (Hanan et al., 2020). As Hanan et al utilised bulk tissue, differences in global circRNA expression in dopaminergic neurons and bulk substantia nigra tissue may reflect changes in cellular composition or dysfunction in other brain cell types. Estimation of cell types in bulk brain tissue was performed, yet the results were not reported (Hanan et al., 2020). Among resident brain cell types, circRNAs appear to be enriched in neurons (Wu et al., 2022). However, samples from the substantia nigra contain varying cellular proportions with several recent single cell RNA-seq (scRNA-seq) studies reporting immunomodulatory oligodendrocytes (Zeis et al., 2016), as the most abundant cell type detected (~41-51%) (Smajić et al., 2022; Wang et al., 2024; Martirosyan et al., 2024).

In dopaminergic neurons isolated from ILB samples, I found a reduction in both circRNA and cognate linear RNA expression (**Figure 6.13**), suggestive of a reduction in general RNA expression in these samples. RNA degradation can be induced by acute activation of RNase L (C.-X. Liu et al., 2019). Upon RNA degradation by RNase L however, greater reductions in circRNA relative to linear RNA have been reported (C.-X. Liu et al., 2019; Guo et al., 2024), owing to the reduced efficiency of back-splicing compared to canonical splicing to replace the degraded RNA (Y. Zhang et al., 2016). In this work however, I did not identify any differences in circRNA expression relative to cognate linear RNA expression (**Figure 6.13e**) nor any changes in *RNASEL* expression in dopaminergic neurons from ILB samples (**Figure 6.14a**). As such, it is unclear why a general reduction in RNA expression is present in these samples.

The reason for the preferential loss of dopaminergic neurons within the substantia nigra in PD is debated (Surmeier, 2018). Even among dopaminergic neurons, some are more susceptible to neurodegeneration (Yamada et al., 1990; Chung et al., 2009). Dopaminergic neurons that exhibit reduced circRNA expression may have been lost. As expression of genes related to antiviral and inflammatory responses are upregulated in PD patients (**Figure 6.14b**), neurons with higher circRNA levels may be more resilient to neurodegeneration. Recent work has highlighted a specific dopaminergic neuron subtype prone to neurodegeneration in PD (Kamath et al., 2022). This neuronal subtype, typified by the expression of *AGTR1*, exhibited diverse transcriptional profiles compared to other subtypes, raising the possibility that transcriptional differences can reflect neuronal vulnerability to neurodegeneration (Kamath et al., 2022). As neurons isolated from BRAINcode samples underwent bulk RNA-seq, I cannot assess inter-neuron differences in circRNA expression. The development of scRNA-seq library preparation methods that do not enrich for polyadenylated transcripts provides a future platform from which to elucidate inter-cell differences in circRNA expression (Fan et al., 2015; Hayashi et al., 2018; Verboom et al., 2019; Isakova et al., 2021; Wu et al., 2022). scRNA-seq could also help reveal how differences in cell populations might contribute to the differences in global circRNA abundances between bulk substantia nigra tissue (Hanan et al., 2020) and isolated dopaminergic neurons, as described in this chapter. scRNA-seq may also determine if the isolation of dopaminergic neurons using laser-capture microdissection affects RNA expression compared to analysing bulk tissue.

6.5 Conclusion

In conclusion, I have identified a common pattern of reduced circRNA expression in the blood of patients with idiopathic, *LRRK2*-PD and *GBA*-PD. As circRNA expression was also reduced in specific groups at increased risk of developing PD, it may reflect PD pathology in the early disease stages before motor symptom onset. In contrast, I found increased global circRNA expression in dopaminergic neurons isolated from the substantia nigra. As the primary location and cell-type affected by neurodegeneration in PD, increased circRNA expression may highlight a method of resilience of neuronal subtypes to neurodegeneration.

Chapter 7. General Discussion

Parkinson's disease (PD) is currently diagnosed after the onset of characteristic motor symptoms resulting from the loss of dopaminergic neurons. However, these motor symptoms present following substantial neurodegeneration. Identifying accessible biomarkers with high sensitivity and specificity is crucial for early detection, accurate diagnoses and tracking disease progression over time. Earlier diagnoses allow for enrolling participants in research studies targeting neurodegeneration sooner, providing more time to assess disease-modifying treatments. Monitoring disease progression over time would also benefit research studies by providing a clinical endpoint in which to assess treatment efficacy. While PD is usually associated with brain neurodegeneration, there is increasing recognition of peripheral dysfunction. Examining changes in peripheral tissues like blood therefore provides insights into PD pathogenesis and progression.

7.1 RNA as a diagnostic biomarker for PD

Chapters 3 and **4** investigated specific forms of RNA as a diagnostic biomarker for early-stage idiopathic PD. **Chapter 3** focuses on gene expression as a diagnostic biomarker, previously studied using microarray-derived and quantitative polymerase chain reaction (qPCR) gene expression measurements (Molochnikov et al., 2012; Scherzer et al., 2007; Santiago & Potashkin, 2015; Santiago et al., 2016; Shamir et al., 2017). Work published during this project also using Parkinson's Progression Marker Initiative (PPMI) participants also provided estimates for distinguishing PD patients from controls (Makarious et al., 2022; Pantaleo et al., 2022; Li et al., 2023). Reassuringly, the ability of gene expression to distinguish PD patients from controls (**Chapter 3**) was consistent with these publications. Generalisability in a separate cohort (Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation–PD (ICICLE-PD), was limited, similar to previous work (Li et al., 2023). Establishing a predictor in **Chapter 3** set a foundation for comparing future predictors tested in **Chapters 4** and **5**.

Circular RNA (circRNA) has characteristics that make it a potential diagnostic biomarker (Verduci et al., 2021), such as a longer half-life than the corresponding linear RNA (Enuka et al., 2016), and detectability in easily accessible tissues (Memczak et al., 2015; Wen et al., 2021). During this project, several studies were published suggesting that specific circRNAs could distinguish PD patients from controls (Ravanidis et al., 2021; Zhong et al., 2021; Xiao et al., 2022). **In Chapter 4**, I investigated whether they could be used for this purpose using the larger sample sizes of the PPMI and ICICLE-PD cohorts. Overall, this work suggested that circRNA

expressed was of limited clinical utility in distinguishing early-stage idiopathic PD patients from controls.

Recent work by Beric et al showed improved discriminatory ability with larger sample sizes from the Parkinson's Disease Biomarker Program (PDBP) and additional PPMI participants (Beric et al., 2024). A key difference between the work presented in this thesis and that of Beric et al is the circRNA quantification method. Beric et al. measured circRNA expression at the gene level, collapsing all back-spliced junction (BSJ) reads to the host gene, thus it is not clear which BSJs and how many contribute to each gene's circular transcript output. In contrast, I reported circRNA expression based on individual BSJs in anticipation of circRNA biomarker candidate validation through qPCR amplification of divergent primers spanning the BSJs (Nielsen et al., 2022).

The identification of potential gene and circRNA biomarkers was based on differential expression between early-stage idiopathic PD patients and controls. My analyses found fewer differentially expressed genes (**Chapter 3**) and circRNAs (**Chapter 4**) compared to similar studies also using PPMI RNA-seq data (Craig et al., 2021; Riboldi et al., 2022; Li et al., 2023; Irmady et al., 2023; Beric et al., 2024). Unlike some of these studies, I specifically tested PD patients with no known genetic causes shortly after diagnosis. Including patients with genetic risk alleles may result in more pronounced RNA expression (Craig et al., 2021; Riboldi et al., 2022; Li et al., 2023). Additionally, blood transcriptome differences may become more pronounced at later timepoints (Craig et al., 2021; Irmady et al., 2023; Beric et al., 2024). As such, future work could examine how gene expression changes over time in early-stage idiopathic PD patients using longitudinal PPMI RNA-seq data.

7.2 Reduced blood circular RNA expression in PD patients and subgroups at increased risk

In **Chapters 4, 5 and 6**, I identified global reductions in circRNA expression in early-stage idiopathic PD patients, PD patients with pathogenic risk variants (*LRRK2* and *GBA*) and individuals with specific risk factors (pathogenic *LRRK2* and *GBA* variants, hyposmia with dopamine transporter deficiency) compared to controls. Importantly, this reduction was specific to circRNAs and did not affect corresponding linear RNAs (**Chapters 5 and 6**), supporting the idea that, for a large proportion of circRNAs, expression is regulated independently of cognate gene expression (Salzman et al., 2013; Memczak et al., 2015; Vo et al., 2019).

Although global circRNA expression measurement did not reliably distinguish early-stage idiopathic PD patients from controls (**Chapter 5**), uncovering the causal mechanism could offer

novel insights into PD pathogenesis or progression. This is important as those considered at risk due to having rapid eye movement sleep behaviour disorder (RBD) showed increased circRNA expression compared to controls (**Chapter 6**). As group-level comparisons identified reduced circRNA expression in PD patients and at-risk subgroups, unsupervised clustering could highlight subgroups with this reduction and may be associated with clinical measures of motor or cognitive outcomes.

7.3 Activation of an innate antiviral response in the blood of PD patients

A large body of work supports a role for the immune system dysfunction in PD, including dysfunction in the periphery (Tansey et al., 2022). Results from across **Chapters 3-6** converge on the activation of an innate immune response in the blood of PD patients. In **Chapter 3**, genes related to the innate immune response were upregulated in early-stage idiopathic PD patients from PPMI and ICICLE-PD cohorts, including genes implicated in the activation of the innate antiviral immune response. Many of these gene categories were quite broad, such as the upregulation of genes related to interferon signalling, a common consequence of numerous convergent antiviral pathways (Katze et al., 2002).

Immune-mediated inflammatory conditions like atopic dermatitis, psoriasis, and systemic lupus erythematosus have been linked to reduced circRNA expression (Liu et al., 2019; Moldovan et al., 2019; Seeler et al., 2022; Moldovan et al., 2021; Guo et al., 2024). Activation of antiviral responses can lead to reduced circRNA expression through several mechanisms (Li et al., 2017; C.-X. Liu et al., 2019). Like previous studies (Moldovan et al., 2019; Fuchs et al., 2023), I examined the expression of known regulators of circRNA biogenesis or degradation (**Chapter 5**). This showed that reduced blood circRNA in early-stage idiopathic PD patients (**Chapter 5**) may be due to activation of an antiviral response related to double-stranded RNA. In cellular models, circRNA degradation was induced by stressors mimicking dsRNA and viral infection, but not stressors mimicking double-stranded DNA (dsDNA), bacterial infection and inflammatory cytokines (C.-X. Liu et al., 2019). This RNase L-mediated degradation of circRNAs abrogates their inhibitory effect on protein kinase R (PKR) (C.-X. Liu et al., 2019), a well characterised activator of the integrated stress response (ISR) (Ma et al., 2006). Activation of the ISR leads to broad translational shutdown through phosphorylation of the alpha subunit of translation initiation factor eIF2 protein complex (eIF2α) (Pakos-Zebrucka et al., 2016) and RNase L degrades ribosomal-associated mRNAs (Andersen et al., 2009), possibly explaining why I observed a downregulation of genes associated with translation and ribosomal function in early-stage idiopathic PD patients (**Chapter 3**). Gene expression of

RNase L (*RNASEL*) was increased in idiopathic PD patients from PPMI and ICICLE-PD cohorts (**Chapter 5**) and in PD patients harbouring pathogenic *LRRK2* and *GBA* variants (**Chapter 6**). Gene expression of PKR (*EIF2AK2*) was upregulated in idiopathic PD patients (**Chapter 5**). Ideally, I would have been able to measure the activity of RNase L and PKR in blood from PD patients and controls. RNase L activation can be inferred from the identification of RNA cleavage products (Donovan et al., 2017; C.-X. Liu et al., 2019) while PKR and ISR activation is typically assessed by measuring phosphorylated PKR and eIF2α levels respectively (Dey et al., 2014).

The inhibitory effect of specific circRNAs on PKR function has now been characterised in multiple studies (C.-X. Liu et al., 2019; Liu et al., 2022; Guo et al., 2024; Feng et al., 2024). In the initial publication, the inhibitory circRNA (*circPOLR2A*) preferentially interacted with PKR (C.-X. Liu et al., 2019). However, the binding of circRNAs was also reported for dsRNA binding proteins such as NF90, OAS1 and ADAR1 (150kDa isoform) (C.-X. Liu et al., 2019), which also play roles in the innate immune response (Silverman, 2007; Mannion et al., 2014; Patiño et al., 2015). As such, it remains to be explored whether circRNAs have a similar inhibitory effect on other dsRNA-binding proteins.

Why might an innate antiviral response be activated in PD? Neurodegenerative disorders have been linked to viral exposure (Olsen et al., 2018). Population studies link influenza infection to increased PD risk (Cocoros et al., 2021; Levine et al., 2023), but distinguishing between the consequences of viral exposure and immune dysfunction associated with PD is challenging (Tansey et al., 2022). Neuronal α-synuclein plays a role in responding to viral infections (Beatman et al., 2016; Stolzenberg et al., 2017). In neuron-astrocyte co-cultures, α-synuclein oligomers activate antiviral responses (D'Sa et al., 2024). In blood, α-synuclein contributes to haematopoiesis and immune cell functioning (Gardai et al., 2013; Xiao et al., 2014). Abnormal α-synuclein expression in blood cells may therefore impact PD development and progression through its antiviral functions.

While the presence of dsRNA is usually linked to viral infection, endogenous dsRNA can also form within cells due to the bidirectional transcription of mitochondrial RNA and the transcription of repetitive regions of the genome, such as transposable elements (S. Kim et al., 2019). Endogenous dsRNA can trigger innate immune responses by binding to cytosolic nucleic acid receptors such as PKR (Kim et al., 2014, 2018; Dhir et al., 2018). Neurons have higher levels of endogenous dsRNA compared to other cell types (Dorrity et al., 2023), and immune responses to dsRNAs have been implicated in Amyotrophic Lateral Sclerosis, Alzheimer's and Huntington's disease (Lee et al., 2020; Rodriguez et al., 2021; Ochoa et al., 2023). In mice,

synthetic dsRNA has been shown to induce α-synuclein aggregation and dopaminergic neuron loss, suggesting a possible role of dsRNA in PD pathogenesis (Xu et al., 2020). In future, it would be useful to examine RNA-seq data described in this thesis for aberrant transposable element expression which could lead to an innate immune response activation (Lanciano & Cristofari, 2020; Gázquez-Gutiérrez et al., 2021).

7.4 Global circular RNA expression in dopaminergic neurons

Since dopaminergic neurons are primarily affected in PD, examining changes in these cells can provide insights into the disease. Previous research indicated region-specific differences in circRNA levels, with lower circRNA abundances in the substantia nigra of PD patients compared to controls (Hanan et al., 2020). In contrast, I found higher global circRNA expression in dopaminergic neurons from PD patients compared to controls. Neurons are notable for their enriched circRNA expression compared to other brain cell-types and expression of neuron-specific circRNAs (Dong et al., 2023). Differences in global circRNA expression between dopaminergic neurons and bulk substantia nigra tissue would indicate a potential dopaminergic neuron specific effect in PD, that has previously been missed in the analysis of bulk tissue. It's important to consider that the neurons isolated from PD patients are those that have not been lost, therefore creating an artificial selection bias for factors that may improve neuron survivability. As such, it is plausible that higher global circRNA expression may confer protective benefits against neurodegeneration. Despite lacking physiological context that may induce circRNA expression changes, it would be useful to examine circRNA expression in induced pluripotent stem cells (iPSCs) derived from PD patients and controls, which may reveal genetic predispositions to circRNA changes in dopaminergic neurons *in vitro* (Bressan et al., 2023). Furthermore, it would be interesting to quantify circRNA expression in PD patients among other resident brain cell-types, particularly those involved in neuroimmune responses like microglia, astrocytes, and oligodendrocytes. Given the activation of antiviral immune responses in neuron-astrocyte co-cultures (D'Sa et al., 2024), it would be interesting to characterise global circRNA expression changes during this process, and potentially give an insight into early neurodegenerative processes. Emerging single-cell RNA-seq (scRNA-seq) methods, which do not enrich polyadenylated transcripts (Fan et al., 2015; Hayashi et al., 2018; Verboom et al., 2019; Isakova et al., 2021; Wu et al., 2022), would allow global circRNA expression profiles to be assigned to individual brain cell-types and in the context of the other brain-cell types.

7.5 Final conclusions

This thesis explored RNA expression as a diagnostic biomarker for Parkinson's disease (PD) and its potential to highlight peripheral PD-related changes. While circRNA expression lacked clinical utility to be used as a diagnostic biomarker for idiopathic PD, I identified a consistent global reduction in circRNA expression in two independent cohorts (PPMI and ICICLE-PD). Integrating this circRNA signature with gene expression data suggested activation of an innate antiviral response in the blood of idiopathic PD patients. Reduced circRNA expression was also observed in other PD subgroups, including those with pathogenic genetic risk variants and higher PD risk, suggesting it may reflect underlying pathology in these groups. Conversely, dopaminergic neurons from PD patients exhibited increased circRNA expression, thereby linking circRNA expression to the cell-type commonly lost in PD. Overall, assessment of global circRNA expression revealed PD-associated changes in both blood and at-risk cell populations.

Chapter 8. Appendix

8.1 Quality control of PPMI and ICICLE-PD RNA sequencing

PPMI RNA-seq samples have previously been independently and externally quality-checked as part of the data release process (Hutchins, Craig, et al., 2021; Craig et al., 2021). ICICLE-PD samples were sequenced in-house, so I performed several quality control steps (**Section 2.4**). As a point of reference, I also performed equivalent checking of sequenced PPMI samples.

To identify possible contaminants, I aligned sequencing reads against a panel of reference and contamination sequences. Most reads mapped solely to the human genome (median proportion of mapped reads [IQR], PPMI = 0.69 [0.09], ICICLE-PD = 0.72 [0.09) (**Figure 8.1a**). The second largest proportion of reads mapped to multiple genomes (median proportion of mapped reads [IQR], PPMI = 0.31 [0.09], ICICLE-PD = 0.27 [0.08]) (**Figure 8.1a**), likely reflecting orthologous sequences across species. All other sequences assessed had a median proportion of mapped reads <0.01 (**Figure 8.1a**).

PPMI and ICICLE-PD samples were sequenced in 29 and three batches respectively. The overall sequence depths of each dataset were high (median million paired-end reads [IQR], PPMI = 106.9 [31.3], ICICLE-PD = 89.2 [15.7]) (**Figure 8.1b**). The deep sequencing of PPMI and ICICLE-PD cohorts was therefore suitable for the detection of even lowly expressed genes (~80 million reads) (Consortium, 2011; Sims et al., 2014). The sequencing quality was mostly PHRED >30 along the length of the reads, with a decrease towards the 3' end as is well documented (Conesa et al., 2016) (**Figure 8.1c**).

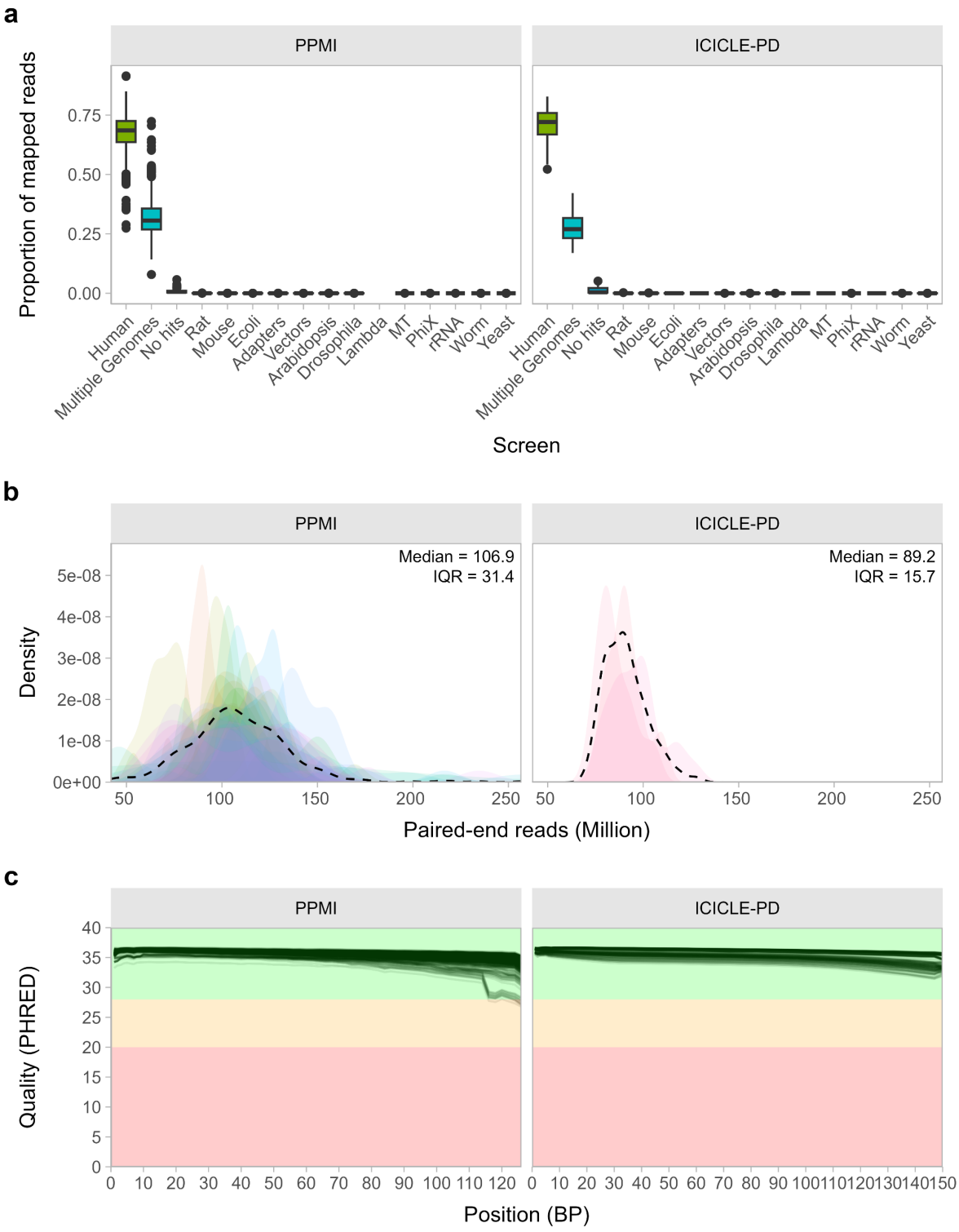


Figure 8.1. Sequencing quality control of PPMI and ICICLE-PD samples. (a) Screening for contamination. Boxplots show the median and IQR of the proportion of reads that mapped to a range of possible contaminants. Outliers greater than 1.5x IQR are shown as dots. **(b)** Distribution of sequencing depths. Densities are coloured according to unique sequencing batches. The dotted line indicates the overall density of sequencing depths for PPMI and ICICLE-PD datasets respectively. **(c)** Quality of sequencing. The mean quality (represented as a PHRED score) at each base pair along the length of the reads. Red (0-20), amber (20-28) and green (28-40) shadings provide a reference of quality thresholds as in FastQC.

After aligning to the GRCh38 genomic reference, the overall mapping rate was >90% in both cohorts, primarily driven by uniquely mapped reads (median percentage of reads [IQR], PPMI = 75.4% [6.8], ICICLE-PD = 80.8% [5.8] (**Figure 8.2a**). Consistent with total RNA-seq library preparations (Zhao et al., 2014, 2018), the highest percentage of bases mapped to intronic regions (median aligned bases [IQR], PPMI = 47.7% [7.5], ICICLE-PD = 40.3% [10.4]) across both cohorts (**Figure 8.2b**). This was followed by untranslated region (UTR) (median aligned bases [IQR], PPMI = 33.4% [8.8], ICICLE-PD = 30.1% [6.3]), coding bases (median aligned bases [IQR], PPMI = 11.7% [2.2], ICICLE-PD = 23.4 [4.5]) and intergenic bases (median aligned bases [IQR], PPMI = 6.0% [2.4], ICICLE-PD = 5.0% [1.6]). After depletion, levels of ribosomal RNA were <1.5% in all samples (**Figure 8.2b**).

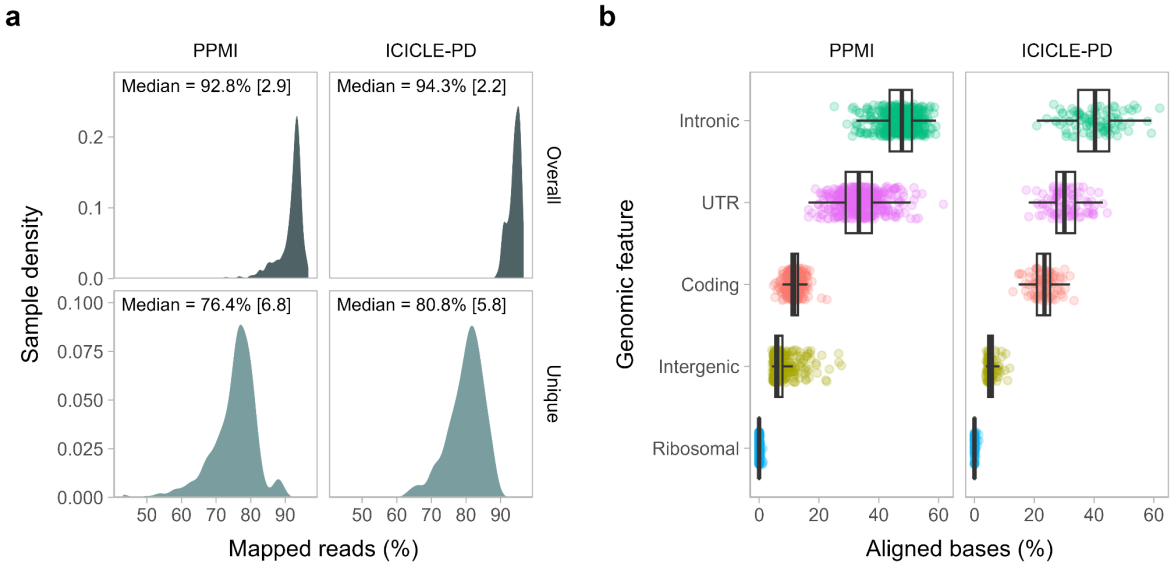


Figure 8.2. Genomic alignment quality control. (a) The distribution of reads mapping to the genome. Mapped reads were categorised as a percentage of all reads that were mapped (Overall) and reads that were mapped uniquely (Unique). Summary statistics are given as the median percentage of reads in each category [IQR]. **(b)** Distribution of aligned bases according to the genomic location of the annotated feature. UTR = Untranslated Region.

The clinically recorded sex of each sample was provided with the sequencing metadata associated with each study. The recorded sex was checked against the sex-biased gene expression as described in **Section 2.4.3**. Visual inspection revealed a mismatch between gene expression and clinically reported sex in one ICICLE-PD participant (**Figure 8.3**). The metadata for this individual was adjusted for downstream analysis.

Figure 8.3. Confirming clinically reported sex in PPMI and ICICLE-PD datasets. Scatter plot showing the first and second principal components of Y chromosomal gene expression (*RPS4Y1, KDM5D, DDX3Y* and *USP9Y*). Points are coloured and shaped according to the clinically reported sex.

8.2 Evaluating classification methods in ICICLE-PD samples

8.2.1 Methods

Logistic regression models were created using glmnet v4.1-8 (Friedman et al., 2010). The alpha value was determined by investigating values from 0.1 to 0.9 in increments of 0.05 (Elastic Net logistic regression). The value that minimised the mean cross-validation error was chosen. In other models, the alpha value was set at 0 (ridge regression) and 1 (LASSO regression). The lambda parameter was determined using five-fold cross-validation, with the optimum value identified based on the highest area under the ROC curve.

Random Forest classification models were created using ranger v0.15.1 (Wright & Ziegler, 2017). A grid search was used to determine the parameters: Number of randomly sampled features to split at each node (1 to 4 in increments of 1), Minimum node size (1 to 3 in increments of 1), and the number of trees (0 to 500 in increments of 50). The combination of parameters which minimised the out-of-bag prediction error was selected.

Support Vector Machine (SVM) classification used e1071 v1.7-13. Cost and Gamma values were chosen from a series of values (0.00, 0.01, 0.1, 1, 10, 100), using five-fold cross-validation as implemented in the *tune.svm* function. The parameters from the model that minimised the misclassification error were chosen and used to create a final model using all the training data.

8.2.2 Results

To identify a suitable method for classifying PD status using RNA expression data, I evaluated a series of supervised machine learning methods (using gene expression data from ICICLE-PD samples (**Section 2.1.1, 2.2**). Before model building, I carried out preliminary feature selection to reduce the computational burden. I performed differential expression on 15,852 genes identifying genes significantly differentially expressed (FDR < 0.05, Wald test) in idiopathic PD patients compared to controls. 608 genes were identified as predictors. I then included variance-stabilised counts (Anders & Huber, 2010) of these 608 genes as predictive variables in the selected classification methods (**Section 8.2.1**)

To train classification models and evaluate performance, I split the ICICLE-PD cohort into training and set sets. Particularly with small sample sizes, performing one split tends to overestimate the test error rate (James et al., 2013). As such, I used leave-one-out cross-validation to train classification models and estimate the test performance (**Figure 8.4**). For

each classification method, I generated 96 models with a different sample acting as the training set in each iteration.

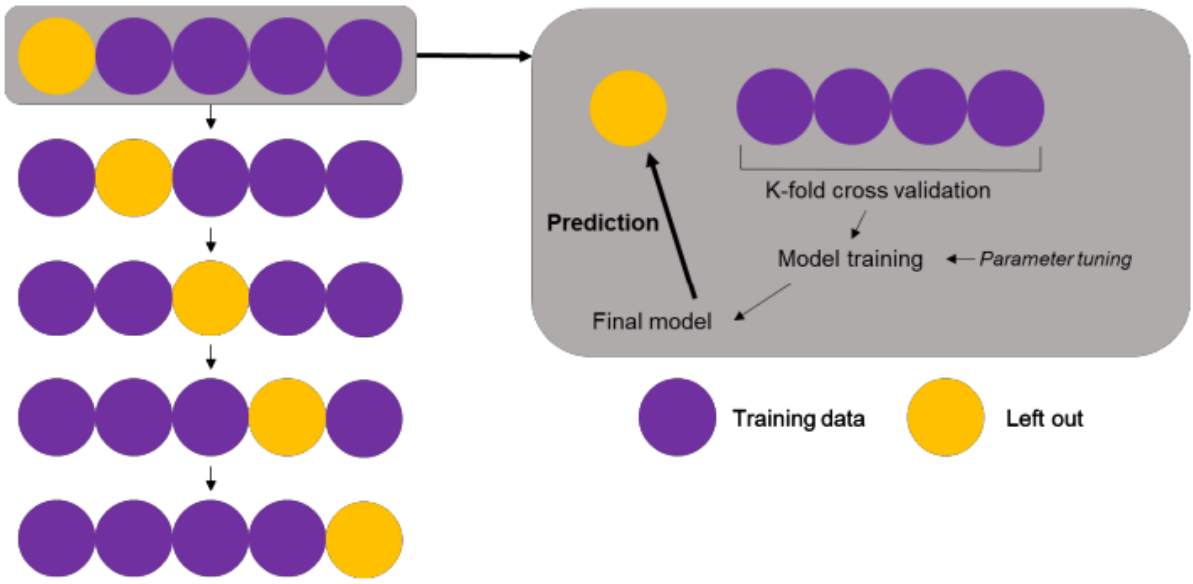


Figure 8.4. Leave-one-out cross-validation schematic. An example case is shown comprising five samples to illustrate how leave-one-out cross-validation works. One sample is excluded from the model building to serve as the left-out sample. On the remaining samples, models are trained to produce a final model. The performance of this model is then assessed on the left-out sample. This process occurs until all samples have acted as the left-out sample.

Reassuringly, all classification methods consistently produced higher probabilities for PD patients compared to controls (all *P*-values < 0.05, Wilcoxon rank-sum test, **Figure 8.5a-f**), indicating these methods could capture meaningful information from gene expression. I then summarised the classification performance by calculating the area under each ROC curve as described in **Section 3.2.12** (**Figure 8.5g, h**). The two SVM models had the highest classification potential (Linear kernel AUC = 0.95, 95% CI: 0.92-0.99; Radial kernel AUC = 0.91, 95% CI: 0.85-0.97) (**Figure 8.5g, h**). The regularised logistic regression models showed the next best performance (Elastic Net AUC = 0.86, 95% CI: 0.79-0.94; Ridge AUC = 0.83 95% CI: 0.75- 0.91; LASSO AUC = 0.83, 95% CI: 0.75- 0.91) (**Figure 8.5g, h**). The Random Forest method showed the lowest classification potential (Random Forest AUC = 0.74, 95% CI: 0.64- 0.84) (**Figure 8.5g, h**).

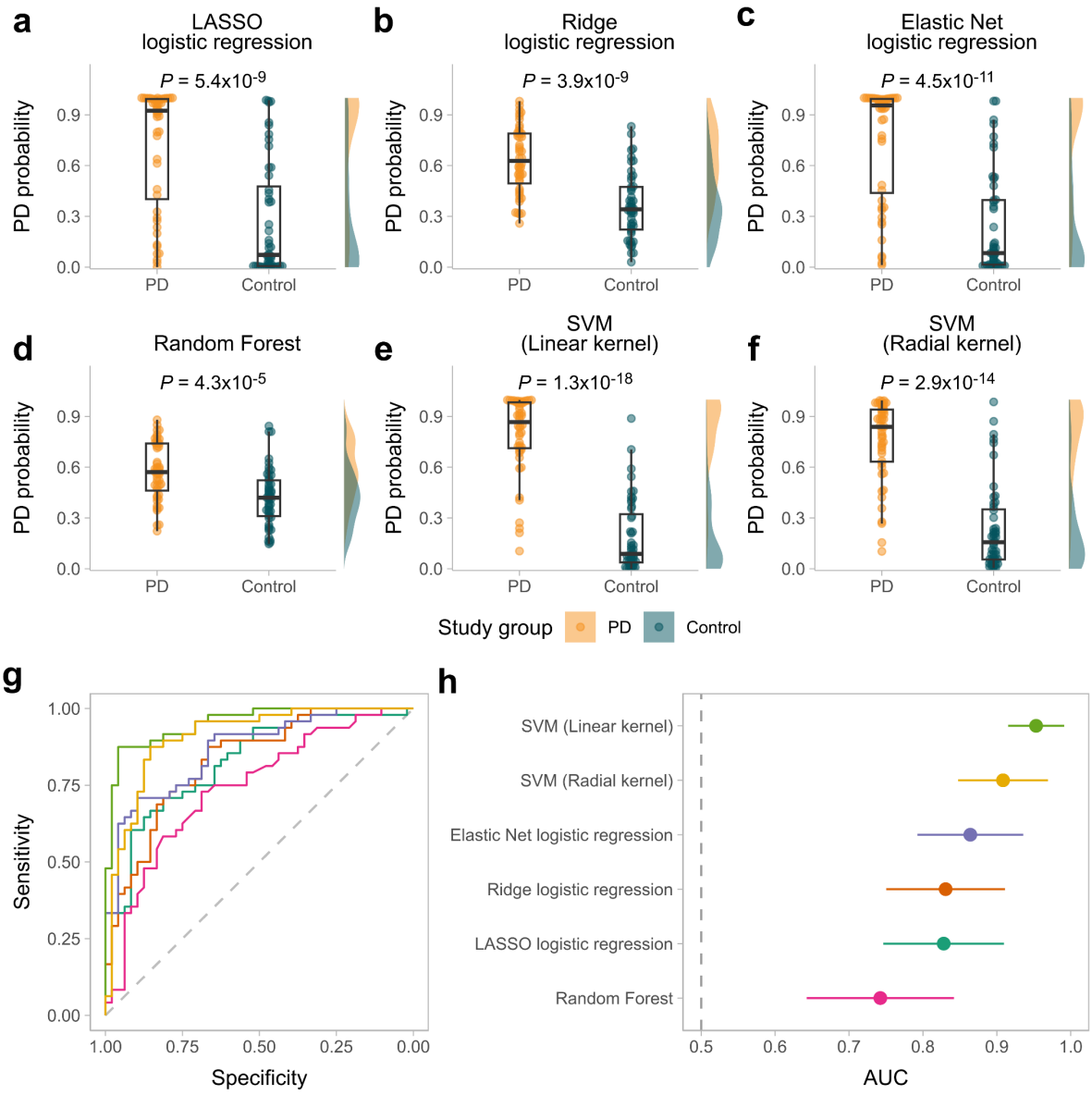


Figure 8.5. Performance of each classification strategy in ICICLE-PD. (a-f) Predicted PD probabilities from each classification method. Probabilities were obtained from predictions using that sample as an independent test dataset. Group differences were assessed with a Wilcoxon rank-sum test. Point densities are shown on the right edge of each plot. **(g)** Receiver-operator characteristic (ROC) curves are based on the specificity and sensitivity at each prediction threshold. **(h)** Areas under the ROC curves. Error bars indicate 95% CIs. An AUC of 0.5, equivalent to random chance, is shown by the grey dotted line.

A key consideration when identifying a suitable classification method is the interpretability of the included predictors. SVM classifiers lack an inherent measure of feature importance (Sanz et al., 2018). The number of support vectors used in an SVM classifier can indicate the bias-variance trade-off; a classifier with a high number of support vectors likely has low variance but high bias (James et al., 2013). Out of 95 possible support vectors (minus one sample acting as the test set), the 96 radial and linear kernel SVM classifiers used a mean of 80.7 (SD = 1.0)

and 67.9 (SD = 1.0) support vectors respectively, potentially signalling overfitting to the training data (**Figure 8.6a**). The next highest-performing classifier was elastic net logistic regression in which feature importance can be inferred from the fitted β coefficients of the predictors. Because of elastic net regularisation, I identified genes excluded from the models where β = 0. Many of the input genes (366/608, 60.2%), were not included in any of the models (**Figure 8.6b**), highlighting the successful feature selection ability of regularisation. In addition, the frequency at which a gene appears in the models can be used as a proxy of feature importance. For example, 21 genes were included in all 96 models created, indicating that they are highly informative predictors (**Figure 8.6a**). Given the increased interpretability of elastic net logistic regression classifiers compared to SVM classifiers, I opted to use elastic net logistic regression as a classification method in this thesis.

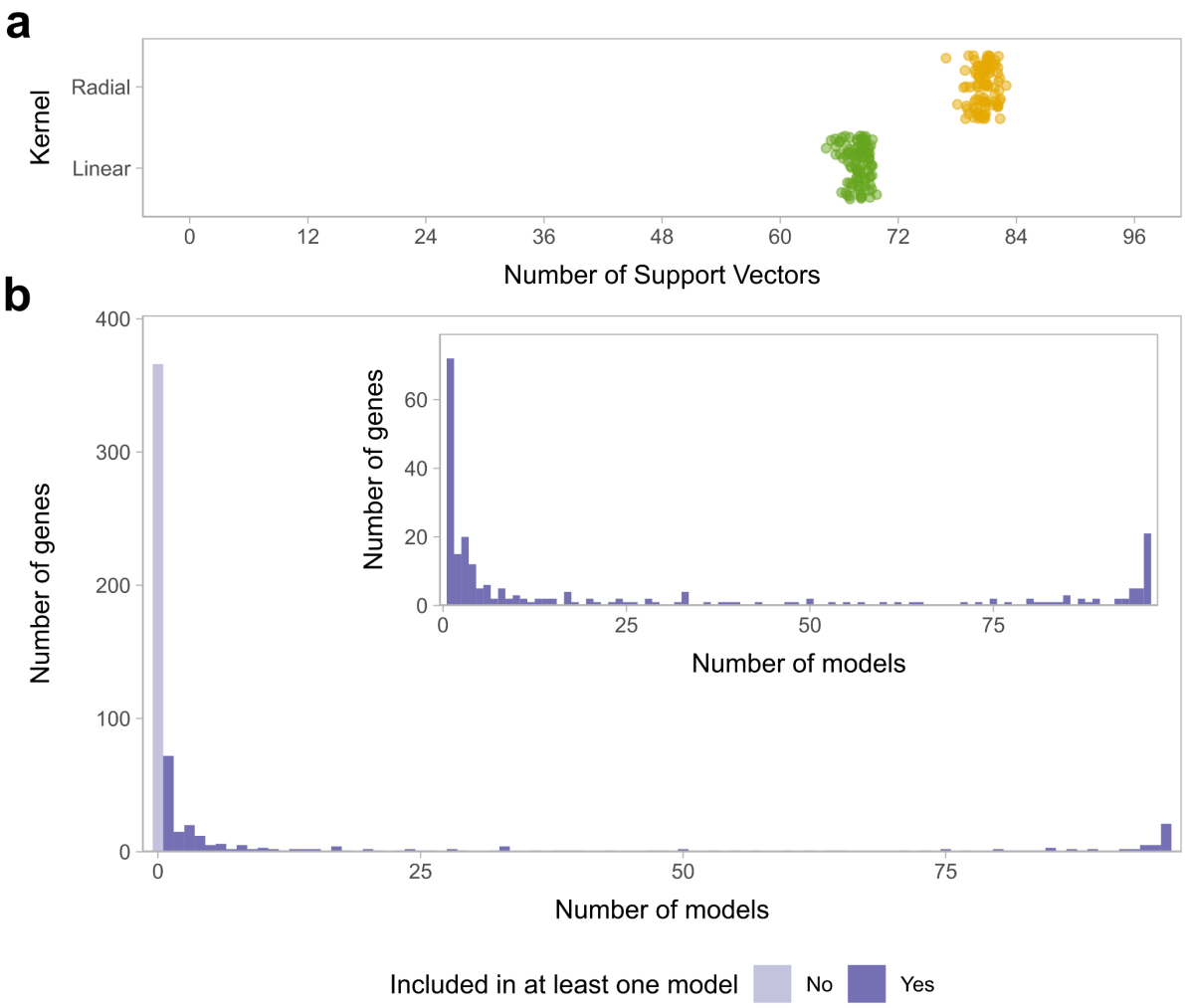


Figure 8.6. Interpretability of SVM and elastic net logistic regression models. (a) Number of support vectors used in each SVM model. **(b)** Effect of regularisation in Elastic Net logistic regression models. Histogram showing the distribution of model inclusion for genes. Inset shows the distribution of model inclusion for genes included in at least one model.

8.3 Overlapping circular RNA predictions from multiple tools produces batch effects when estimating canonical splice junction expression

8.3.1 Methods

PPMI sample information was taken from the baseline visit. Information regarding control samples and patients with idiopathic PD was described in **Section 2.2**. Study groups were provided in sequencing metadata provided with a publication detailing the PPMI's analysis of RNA sequencing data (Craig et al., 2021). A total of 415 individuals with *GBA* (*GBA*_PD = 61, *GBA*_Controls = 78) and *LRRK2* (*LRRK2*_PD = 134, *LRRK2*_Controls = 142) variants were processed. In addition, 258 idiopathic PD patients and 155 controls were also previously processed (**Chapter 4**).

8.3.2 Results

I detected 28,398 unique BSJs across 828 samples. 1,545 highly expressed BSJs were retained (**Section 6.2.6**). To assess global structure within the circular and linear RNA expression data, I calculated principal components (PCs) of circRNA (BSJ counts) and the cognate linear RNA (FSJ counts) expression. Study groups separated along PC1 and PC2 of FSJ expression yet there was no comparable separation using BSJ expression (**Figure 8.7a, b**). Specifically, these groups corresponded to both the presence of a known PD-associated genetic variant, and the batches the samples were processed in.

To discern whether widespread differences in FSJ expression across study groups were driven by technical or biological effects, I detected BSJs using one tool (CIRI2) before quantifying BSJ and FSJ expression using CIRIquant (**Section 6.2.6**). In the same 828 samples, I detected 18,751 BSJs. Projection of BSJ and FSJ expression into the first two PCs showed no segregation of study groups (**Figure 8.7a, b**). Comparison of the reported expression of 3,080 shared loci between the original workflow (overlapping BSJ detection) and the rerun workflow (detecting using one tool) showed a higher correlation between BSJ counts ($R^2 = 0.96$) compared to FSJ counts ($R^2 = 0.7$) (**Figure 8.7d**). Notably, 6.9% (212/3080) of the FSJs had reported expression in the rerun workflow yet no reported expression in the original workflow (**Figure 8.7d**). Together, these results indicate that differences in the BSJ background provided to CIRIquant can produce changes in BSJ and FSJ expression estimates with the changes more pronounced when quantifying FSJ expression.

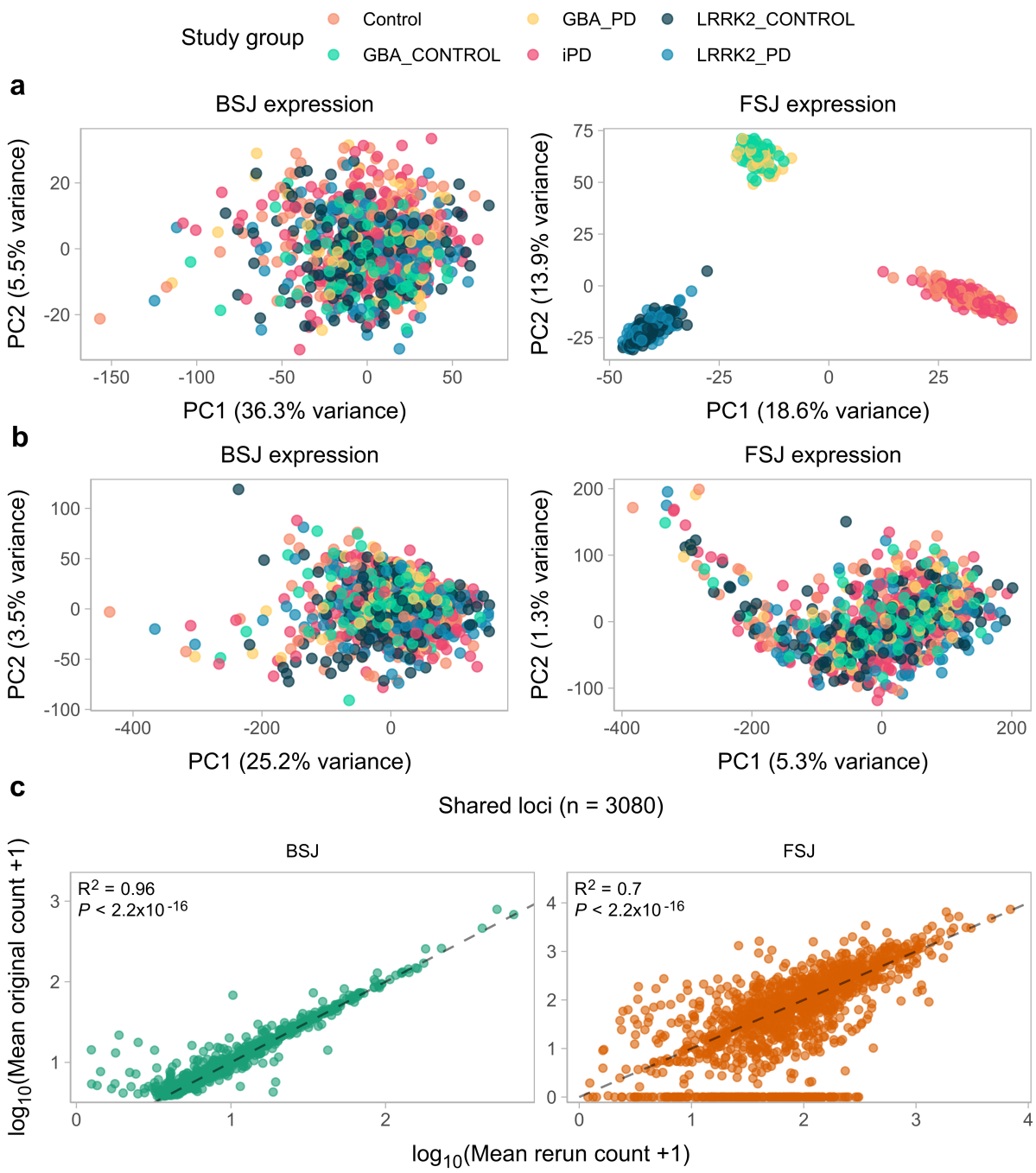


Figure 8.7. Batch effects when quantifying forward-spliced junction expression after overlapping circular RNA detections from multiple tools. (a) First two principal components of BSJ and FSJ expression after overlapping detections. **(b)** The first two principal components of BSJ and FSJ expression after detection with CIRI2 before quantification with CIRIquant. **(c)** Comparison of raw counts estimated from the original BSJ workflow (overlapping BSJ detections) and the rerun workflow (CIRI2 and CIRIquant). iPD = idiopathic PD.

8.4 Supplementary tables

					PPMI			ICICLE-PD		
Ontology	**Name**	**Significant in PPMI**	**Replicated in ICICLE-PD**	**Direction**	**NES**	**P-value**	**FDR**	**NES**	**P-value**	**FDR**
CC	cytosolic ribosome	Y	Y	Agree	-3.08	2.76E-25	2.01E-22	-2.16	4.86E-07	4.03E-06
CC	ribosomal subunit	Y	Y	Agree	-2.81	3.81E-24	1.38E-21	-2.22	2.52E-10	8.08E-09
BP	cytoplasmic translation	Y	Y	Agree	-2.88	9.74E-23	5.57E-19	-2.16	4.34E-09	6.56E-08
MF	structural constituent of ribosome	Y	Y	Agree	-2.86	1.30E-22	1.40E-19	-2.21	2.55E-09	4.68E-08
CC	ribosome	Y	Y	Agree	-2.72	3.22E-22	7.79E-20	-2.14	4.38E-10	1.25E-08
CC	large ribosomal subunit	Y	Y	Agree	-2.67	2.21E-15	4.02E-13	-2.31	2.44E-09	4.68E-08
CC	cytosolic large ribosomal subunit	Y	Y	Agree	-2.85	5.84E-15	8.49E-13	-2.27	4.10E-06	2.51E-05
CC	mitochondrial protein-containing complex	Y	Y	Agree	-2.21	6.71E-13	8.12E-11	-1.57	6.85E-05	2.67E-04
CC	cytosolic small ribosomal subunit	Y	Y	Agree	-2.72	2.87E-12	2.97E-10	-1.56	1.17E-02	2.20E-02
BP	negative regulation of viral process	Y	Y	Agree	2.57	6.69E-11	1.91E-07	2.23	9.32E-09	1.26E-07
BP	defense response to virus	Y	Y	Agree	2.06	4.72E-10	8.56E-07	2.12	1.81E-12	2.33E-10
BP	defense response to symbiont	Y	Y	Agree	2.06	5.99E-10	8.56E-07	2.12	3.79E-12	3.25E-10
CC	small ribosomal subunit	Y	Y	Agree	-2.45	7.37E-10	5.94E-08	-1.69	2.39E-03	5.75E-03
BP	ribosome biogenesis	Y	Y	Agree	-2.01	9.46E-10	1.08E-06	-1.88	1.98E-08	2.31E-07
BP	negative regulation of viral genome replication	Y	Y	Agree	2.59	2.06E-09	1.96E-06	2.33	1.08E-08	1.38E-07
BP	regulation of response to biotic stimulus	Y	Y	Agree	1.93	3.01E-09	2.46E-06	1.76	1.49E-06	1.01E-05
BP	mitochondrial gene expression	Y	Y	Agree	-2.13	7.27E-09	5.19E-06	-1.94	1.07E-06	7.87E-06
BP	ncRNA processing	Y	Y	Agree	-1.81	1.37E-08	8.71E-06	-1.71	6.20E-08	6.38E-07
BP	regulation of innate immune response	Y	Y	Agree	2.04	1.56E-08	8.89E-06	1.68	8.57E-05	3.24E-04
BP	regulation of viral genome replication	Y	Y	Agree	2.40	1.71E-08	8.89E-06	2.20	1.54E-08	1.89E-07
BP	chemotaxis	Y	Y	Agree	1.77	2.69E-08	1.18E-05	1.52	2.08E-04	6.68E-04
BP	taxis	Y	Y	Agree	1.77	2.69E-08	1.18E-05	1.52	2.08E-04	6.68E-04
BP	regulation of viral life cycle	Y	Y	Agree	2.15	4.70E-08	1.87E-05	2.21	1.29E-09	3.01E-08
BP	ribonucleoprotein complex biogenesis	Y	Y	Agree	-1.77	4.91E-08	1.87E-05	-1.61	1.96E-06	1.26E-05
BP	response to virus	Y	Y	Agree	1.84	6.17E-08	2.21E-05	2.04	6.83E-12	4.39E-10
CC	polysome	Y	Y	Agree	-2.30	1.00E-07	6.37E-06	-1.59	1.12E-02	2.14E-02
CC	mitochondrial matrix	Y	Y	Agree	-1.73	1.05E-07	6.37E-06	-1.45	1.09E-04	4.07E-04
BP	viral genome replication	Y	Y	Agree	2.11	1.30E-07	4.38E-05	2.13	2.46E-08	2.75E-07
CC	polysomal ribosome	Y	Y	Agree	-2.41	1.54E-07	8.59E-06	-1.98	9.44E-04	2.56E-03
CC	secretory granule membrane	Y	Y	Agree	1.85	1.74E-07	9.02E-06	1.97	1.51E-09	3.24E-08
BP	cell chemotaxis	Y	Y	Agree	1.92	1.78E-07	5.66E-05	1.36	2.47E-02	4.04E-02
BP	regulation of viral process	Y	Y	Agree	2.07	1.95E-07	5.87E-05	2.15	4.70E-09	6.71E-08
BP	leukocyte chemotaxis	Y	Y	Agree	1.99	2.06E-07	5.89E-05	1.41	1.35E-02	2.45E-02
BP	cytokine-mediated signaling pathway	Y	Y	Agree	1.75	3.93E-07	9.82E-05	1.66	6.01E-06	3.43E-05
BP	positive regulation of response to external stimulus	Y	Y	Agree	1.74	4.23E-07	9.82E-05	1.80	1.24E-07	1.22E-06

CC	ficolin-1-rich granule	Y	Y	Agree	1.99	5.27E-07	2.55E-05	1.44	1.28E-02	2.35E-02
BP	rRNA processing	Y	Y	Agree	-1.90	5.90E-07	1.30E-04	-1.75	4.47E-06	2.67E-05
BP	immune response-regulating signaling pathway	Y	Y	Agree	1.69	6.29E-07	1.33E-04	1.89	1.19E-09	3.01E-08
BP	mitochondrial translation	Y	Y	Agree	-2.09	6.98E-07	1.37E-04	-1.87	3.49E-05	1.52E-04
BP	viral life cycle	Y	Y	Agree	1.77	7.84E-07	1.49E-04	1.84	1.41E-07	1.34E-06
BP	leukocyte migration	Y	Y	Agree	1.78	9.42E-07	1.74E-04	1.31	2.70E-02	4.31E-02
MF	peptide antigen binding	Y	Y	Agree	-2.29	1.82E-06	6.49E-04	-1.79	6.35E-03	1.37E-02
BP	positive regulation of response to biotic stimulus	Y	Y	Agree	1.96	2.18E-06	3.89E-04	1.77	5.77E-05	2.34E-04
BP	rRNA metabolic process	Y	Y	Agree	-1.80	2.35E-06	4.03E-04	-1.56	1.57E-04	5.60E-04
BP	pattern recognition receptor signaling pathway	Y	Y	Agree	1.90	2.40E-06	4.03E-04	2.21	3.62E-11	1.86E-09
BP	positive regulation of defense response	Y	Y	Agree	1.80	2.68E-06	4.38E-04	1.80	1.49E-06	1.01E-05
BP	phagocytosis	Y	Y	Agree	1.80	2.93E-06	4.66E-04	1.89	5.32E-08	5.69E-07
BP	response to interferon-alpha	Y	Y	Agree	2.28	3.70E-06	5.56E-04	1.75	8.67E-03	1.74E-02
BP	positive regulation of cytokine production	Y	Y	Agree	1.65	3.94E-06	5.77E-04	1.75	1.48E-07	1.36E-06
BP	tRNA processing	Y	Y	Agree	-2.02	4.16E-06	5.94E-04	-1.72	2.84E-04	8.65E-04
CC	organellar ribosome	Y	Y	Agree	-2.07	4.17E-06	1.44E-04	-1.86	3.10E-04	9.16E-04
CC	mitochondrial ribosome	Y	Y	Agree	-2.07	4.17E-06	1.44E-04	-1.86	3.10E-04	9.16E-04
BP	response to type I interferon	Y	Y	Agree	2.18	4.65E-06	6.32E-04	2.15	1.37E-06	9.80E-06
BP	positive regulation of innate immune response	Y	Y	Agree	1.94	5.17E-06	6.78E-04	1.69	5.84E-04	1.65E-03
BP	positive regulation of locomotion	Y	Y	Agree	1.59	5.22E-06	6.78E-04	1.43	1.49E-03	3.92E-03
BP	tRNA modification	Y	Y	Agree	-2.07	5.47E-06	6.94E-04	-1.79	4.73E-04	1.37E-03
BP	activation of innate immune response	Y	Y	Agree	2.17	6.17E-06	7.67E-04	1.72	1.85E-03	4.81E-03
BP	ribosomal large subunit biogenesis	Y	Y	Agree	-2.10	7.72E-06	9.26E-04	-1.91	1.53E-04	5.55E-04
BP	viral process	Y	Y	Agree	1.65	7.78E-06	9.26E-04	1.73	1.06E-06	7.87E-06
BP	type I interferon signaling pathway	Y	Y	Agree	2.16	8.47E-06	9.69E-04	2.08	7.28E-06	3.90E-05
BP	positive regulation of immune response	Y	Y	Agree	1.57	9.07E-06	9.98E-04	1.64	1.57E-06	1.04E-05
BP	myeloid leukocyte migration	Y	Y	Agree	1.86	9.55E-06	1.03E-03	1.47	7.00E-03	1.46E-02
BP	positive regulation of cell migration	Y	Y	Agree	1.61	1.05E-05	1.12E-03	1.40	2.88E-03	6.72E-03
BP	positive regulation of cell motility	Y	Y	Agree	1.59	1.09E-05	1.14E-03	1.40	2.49E-03	5.92E-03
BP	tRNA metabolic process	Y	Y	Agree	-1.81	1.29E-05	1.31E-03	-1.73	5.82E-05	2.34E-04
CC	organellar large ribosomal subunit	Y	Y	Agree	-2.08	1.49E-05	4.71E-04	-1.75	6.51E-03	1.38E-02
CC	mitochondrial large ribosomal subunit	Y	Y	Agree	-2.08	1.49E-05	4.71E-04	-1.75	6.51E-03	1.38E-02
BP	RNA methylation	Y	Y	Agree	-1.99	1.66E-05	1.60E-03	-1.95	7.41E-05	2.84E-04
BP	cellular response to type I interferon	Y	Y	Agree	2.12	1.71E-05	1.63E-03	2.14	9.37E-07	7.30E-06
BP	regulation of response to cytokine stimulus	Y	Y	Agree	1.86	1.75E-05	1.64E-03	1.76	2.04E-04	6.68E-04
BP	regulation of cytokine-mediated signaling pathway	Y	Y	Agree	1.88	2.52E-05	2.29E-03	1.74	1.76E-04	6.02E-04
BP	activation of immune response	Y	Y	Agree	1.60	4.55E-05	3.76E-03	1.58	1.89E-04	6.38E-04
MF	RNA methyltransferase activity	Y	Y	Agree	-1.98	4.88E-05	1.31E-02	-1.86	5.52E-04	1.58E-03

BP	response to wounding	Y	Y	Agree	1.52	5.02E-05	4.04E-03	1.46	4.66E-04	1.36E-03
BP	defense response to bacterium	Y	Y	Agree	1.70	5.14E-05	4.08E-03	2.29	2.10E-14	5.40E-12
BP	tRNA methylation	Y	Y	Agree	-2.07	6.23E-05	4.81E-03	-1.76	4.21E-03	9.41E-03
BP	exocytosis	Y	Y	Agree	1.59	6.68E-05	5.02E-03	1.37	1.17E-02	2.20E-02
BP	positive regulation of pattern recognition receptor signaling pathway	Y	Y	Agree	2.08	7.73E-05	5.59E-03	2.10	1.10E-05	5.24E-05
BP	cellular extravasation	Y	Y	Agree	2.01	8.11E-05	5.72E-03	1.52	1.25E-02	2.33E-02
BP	regulation of inflammatory response	Y	Y	Agree	1.60	8.97E-05	6.18E-03	1.53	7.52E-04	2.08E-03
BP	ribosomal small subunit biogenesis	Y	Y	Agree	-1.91	9.75E-05	6.63E-03	-2.08	1.82E-05	8.20E-05
CC	inflammasome complex	Y	Y	Agree	2.11	1.05E-04	3.04E-03	1.88	9.47E-04	2.56E-03
MF	rRNA binding	Y	Y	Agree	-1.94	1.18E-04	2.11E-02	-1.72	2.81E-03	6.62E-03
BP	negative regulation of cytokine-mediated signaling pathway	Y	Y	Agree	1.95	1.30E-04	8.76E-03	1.90	1.61E-04	5.67E-04
CC	tertiary granule	Y	Y	Agree	1.74	1.31E-04	3.65E-03	2.23	4.39E-11	1.88E-09
BP	positive regulation of interleukin-8 production	Y	Y	Agree	1.97	1.42E-04	9.42E-03	1.80	2.09E-03	5.33E-03
MF	signaling receptor regulator activity	Y	Y	Agree	1.58	1.59E-04	2.28E-02	1.43	9.15E-03	1.81E-02
MF	phospholipid binding	Y	Y	Agree	1.51	1.70E-04	2.28E-02	1.31	1.70E-02	2.96E-02
BP	response to interferon-beta	Y	Y	Agree	2.02	2.01E-04	1.28E-02	1.76	4.62E-03	1.02E-02
BP	regulation of type I interferon-mediated signaling pathway	Y	Y	Agree	2.04	2.16E-04	1.36E-02	2.03	1.68E-04	5.83E-04
MF	catalytic activity, acting on a tRNA	Y	Y	Agree	-1.76	2.17E-04	2.41E-02	-1.54	2.17E-03	5.37E-03
MF	tRNA methyltransferase activity	Y	Y	Agree	-2.02	2.25E-04	2.41E-02	-1.72	7.36E-03	1.53E-02
BP	regulated exocytosis	Y	Y	Agree	1.64	2.52E-04	1.52E-02	1.52	4.16E-03	9.38E-03
BP	cellular response to peptide hormone stimulus	Y	Y	Agree	1.56	2.67E-04	1.59E-02	1.35	2.87E-02	4.55E-02
CC	ficolin-1-rich granule membrane	Y	Y	Agree	1.90	2.82E-04	7.07E-03	1.58	1.26E-02	2.33E-02
BP	leukocyte mediated immunity	Y	Y	Agree	1.46	3.19E-04	1.88E-02	1.77	2.63E-07	2.33E-06
BP	mitochondrial RNA metabolic process	Y	Y	Agree	-1.90	3.70E-04	2.14E-02	-1.72	8.01E-03	1.65E-02
BP	pyroptosis	Y	Y	Agree	2.05	4.40E-04	2.41E-02	1.84	2.20E-03	5.40E-03
BP	regulation of secretion	Y	Y	Agree	1.43	4.63E-04	2.50E-02	1.41	2.40E-03	5.75E-03
BP	positive regulation of MAPK cascade	Y	Y	Agree	1.47	5.05E-04	2.70E-02	1.31	2.63E-02	4.25E-02
CC	immunoglobulin complex	Y	Y	Agree	1.85	5.06E-04	1.11E-02	2.38	9.19E-11	3.37E-09
BP	porphyrin-containing compound biosynthetic process	Y	Y	Agree	-1.94	5.36E-04	2.79E-02	-1.71	8.34E-03	1.69E-02
BP	tetrapyrrole biosynthetic process	Y	Y	Agree	-1.94	5.36E-04	2.79E-02	-1.71	8.34E-03	1.69E-02
BP	superoxide metabolic process	Y	Y	Agree	1.85	5.47E-04	2.81E-02	1.56	1.64E-02	2.87E-02
BP	cellular response to peptide	Y	Y	Agree	1.51	5.52E-04	2.81E-02	1.41	6.96E-03	1.46E-02
CC	extracellular matrix	Y	Y	Agree	1.48	5.76E-04	1.23E-02	1.68	1.14E-05	5.24E-05
BP	interleukin-1 production	Y	Y	Agree	1.77	6.35E-04	3.10E-02	1.97	8.98E-06	4.52E-05

BP	regulation of interleukin-1 production	Y	Y	Agree	1.77	6.35E-04	3.10E-02	1.97	8.98E-06	4.52E-05
BP	positive regulation of fatty acid transport	Y	Y	Agree	2.00	6.41E-04	3.11E-02	1.61	2.01E-02	3.41E-02
MF	signaling receptor activator activity	Y	Y	Agree	1.56	6.49E-04	4.34E-02	1.42	8.94E-03	1.78E-02
BP	positive regulation of superoxide anion generation	Y	Y	Agree	2.02	6.83E-04	3.24E-02	1.66	9.54E-03	1.84E-02
BP	trans-synaptic signaling	Y	Y	Agree	1.41	6.85E-04	3.24E-02	1.62	1.07E-05	5.20E-05
BP	positive regulation of toll-like receptor signaling pathway	Y	Y	Agree	1.92	7.23E-04	3.36E-02	1.99	2.04E-04	6.68E-04
BP	positive regulation of response to cytokine stimulus	Y	Y	Agree	1.83	7.79E-04	3.52E-02	1.52	2.18E-02	3.64E-02
BP	regulation of vesicle-mediated transport	Y	Y	Agree	1.42	7.81E-04	3.52E-02	1.50	2.28E-04	7.14E-04
BP	negative regulation of response to cytokine stimulus	Y	Y	Agree	1.82	7.98E-04	3.53E-02	1.84	7.07E-04	1.97E-03
CC	blood microparticle	Y	Y	Agree	1.75	8.59E-04	1.64E-02	1.94	1.97E-05	8.75E-05
MF	receptor ligand activity	Y	Y	Agree	1.57	8.63E-04	4.90E-02	1.43	9.35E-03	1.83E-02
BP	heme biosynthetic process	Y	Y	Agree	-1.89	8.67E-04	3.80E-02	-1.71	1.37E-02	2.48E-02
MF	glycosaminoglycan binding	Y	Y	Agree	1.66	8.69E-04	4.90E-02	1.79	1.37E-04	5.03E-04
BP	synaptic signaling	Y	Y	Agree	1.39	8.79E-04	3.81E-02	1.60	5.72E-06	3.34E-05
BP	phagocytosis, engulfment	Y	Y	Agree	1.78	9.48E-04	3.95E-02	2.04	2.31E-06	1.45E-05
BP	ribosome assembly	Y	Y	Agree	-1.85	9.51E-04	3.95E-02	-2.12	5.06E-05	2.10E-04
CC	membrane microdomain	Y	Y	Agree	1.48	9.77E-04	1.79E-02	1.34	2.19E-02	3.64E-02
CC	external encapsulating structure	Y	Y	Agree	1.47	9.87E-04	1.79E-02	1.69	1.12E-05	5.24E-05
BP	toll-like receptor signaling pathway	Y	Y	Agree	1.66	1.03E-03	4.19E-02	2.04	4.60E-07	3.94E-06
BP	protein targeting to mitochondrion	Y	Y	Agree	-1.67	1.05E-03	4.22E-02	-1.57	3.94E-03	8.97E-03
BP	circulatory system process	Y	Y	Agree	1.41	1.06E-03	4.25E-02	1.40	5.43E-03	1.18E-02
BP	regulation of superoxide metabolic process	Y	Y	Agree	1.89	1.07E-03	4.25E-02	1.70	9.99E-03	1.92E-02
BP	regulation of fatty acid transport	Y	Y	Agree	1.89	1.11E-03	4.37E-02	1.65	1.64E-02	2.87E-02
BP	negative regulation of cell differentiation	Y	Y	Agree	1.40	1.18E-03	4.57E-02	1.24	3.01E-02	4.74E-02
BP	interleukin-1 beta production	Y	Y	Agree	1.74	1.20E-03	4.57E-02	1.91	3.93E-05	1.65E-04
BP	regulation of interleukin-1 beta production	Y	Y	Agree	1.74	1.20E-03	4.57E-02	1.91	3.93E-05	1.65E-04
CC	membrane raft	Y	Y	Agree	1.46	1.56E-03	2.62E-02	1.35	1.50E-02	2.67E-02
CC	collagen-containing extracellular matrix	Y	Y	Agree	1.50	1.62E-03	2.62E-02	1.68	6.14E-05	2.43E-04
CC	vesicle lumen	Y	Y	Agree	1.45	1.70E-03	2.62E-02	1.95	3.06E-09	4.91E-08
CC	cytoplasmic vesicle lumen	Y	Y	Agree	1.45	1.70E-03	2.62E-02	1.95	3.06E-09	4.91E-08

Table 8.1. Gene Set Enrichment Analysis of Gene Ontology terms. Table showing the results of Gene set enrichment analysis of genes (based on ranked fold changes) grouped into Gene Ontology terms (GO). Ontologies include Biological Processes (BP), Molecular Functions (MF) and Cellular Components (CC). For each term, the corresponding ontology, unique GO ID, the name of the term, whether it is significant in PPMI and replicated in ICICLE-PD (FDR <0.05) and whether the direction of effect agrees between PPMI and ICICLE-PD cohorts (based on the Normalised Enrichment Score, NES). FDR values in ICICLE-PD are based on terms significantly enriched in PPMI.

			PPMI					*ICICLE-PD*		
ONTOLOGY	**ID**	**Description**	**GeneRatio**	**BgRatio**	**P-value**	**FDR**	**GeneRatio**	**BgRatio**	**P-value**	**FDR**
MF	GO:0048156	tau protein binding	6/276	18/4877	3.2E-04	1.4E-01	5/299	17/4051	6.3E-03	5.3E-01
MF	GO:0003682	chromatin binding	27/276	257/4877	1.2E-03	2.5E-01	28/299	219/4051	2.5E-03	4.9E-01
BP	GO:0006325	chromatin organization	29/273	282/4784	1.2E-03	8.2E-01	27/296	251/3982	3.0E-02	6.7E-01
BP	GO:2000058	regulation of ubiquitin-dependent protein catabolic process	11/273	70/4784	1.8E-03	8.2E-01	10/296	65/3982	2.0E-02	6.2E-01
BP	GO:1903201	regulation of oxidative stress-induced cell death	5/273	17/4784	2.0E-03	8.2E-01	4/296	15/3982	2.1E-02	6.2E-01
BP	GO:0006898	receptor-mediated endocytosis	13/273	94/4784	2.4E-03	8.2E-01	11/296	81/3982	3.6E-02	7.1E-01
BP	GO:2000369	regulation of clathrin-dependent endocytosis	4/273	1053666	2.5E-03	8.2E-01				
BP	GO:0009880	embryonic pattern specification	5/273	18/4784	2.7E-03	8.2E-01	5/296	18/3982	8.4E-03	6.2E-01
BP	GO:0006479	protein methylation	13/273	97/4784	3.2E-03	8.2E-01	13/296	88/3982	1.2E-02	6.2E-01
BP	GO:0008213	protein alkylation	13/273	97/4784	3.2E-03	8.2E-01	13/296	88/3982	1.2E-02	6.2E-01
MF	GO:0008276	protein methyltransferase activity	8/276	45/4877	3.3E-03	4.7E-01	8/299	44/4051	1.4E-02	5.3E-01
BP	GO:0071168	protein localization to chromatin	5/273	19/4784	3.5E-03	8.2E-01	5/296	15/3982	3.5E-03	6.2E-01
BP	GO:1903050	regulation of proteolysis involved in protein catabolic process	12/273	88/4784	3.9E-03	8.2E-01	11/296	81/3982	3.6E-02	7.1E-01
BP	GO:2000060	positive regulation of ubiquitin-dependent protein catabolic process	8/273	46/4784	4.0E-03	8.2E-01	7/296	41/3982	2.9E-02	6.7E-01
BP	GO:0043401	steroid hormone mediated signaling pathway	9/273	56/4784	4.0E-03	8.2E-01	6/296	45/3982	1.1E-01	8.5E-01
BP	GO:0043414	macromolecule methylation	17/273	151/4784	5.1E-03	8.2E-01	16/296	132/3982	3.4E-02	6.8E-01
BP	GO:1900407	regulation of cellular response to oxidative stress	5/273	21/4784	5.6E-03	8.2E-01	4/296	20/3982	5.6E-02	7.8E-01
BP	GO:0009798	axis specification	5/273	23/4784	8.4E-03	8.2E-01	5/296	20/3982	1.3E-02	6.2E-01
BP	GO:0035282	segmentation	5/273	23/4784	8.4E-03	8.2E-01	5/296	16/3982	4.8E-03	6.2E-01
BP	GO:0090311	regulation of protein deacetylation	5/273	23/4784	8.4E-03	8.2E-01	5/296	19/3982	1.1E-02	6.2E-01
BP	GO:0030518	intracellular steroid hormone receptor signaling pathway	8/273	52/4784	8.5E-03	8.2E-01	5/296	42/3982	2.0E-01	8.8E-01
BP	GO:0018027	peptidyl-lysine dimethylation	4/273	15/4784	8.6E-03	8.2E-01	4/296	760741	6.5E-03	6.2E-01
BP	GO:0001841	neural tube formation	7/273	42/4784	8.8E-03	8.2E-01	6/296	36/3982	4.7E-02	7.5E-01
MF	GO:0008270	zinc ion binding	25/276	269/4877	8.9E-03	4.9E-01	22/299	233/4051	1.3E-01	7.9E-01
MF	GO:0019902	phosphatase binding	11/276	87/4877	9.4E-03	4.9E-01	13/299	71/4051	1.7E-03	4.9E-01
MF	GO:0042054	histone methyltransferase activity	6/276	33/4877	9.4E-03	4.9E-01	5/299	31/4051	7.4E-02	7.9E-01
BP	GO:0016570	histone modification	22/273	227/4784	9.5E-03	8.2E-01	23/296	196/3982	1.8E-02	6.2E-01
BP	GO:0006476	protein deacetylation	8/273	53/4784	9.6E-03	8.2E-01	7/296	45/3982	4.6E-02	7.5E-01

BP	GO:0009791	post-embryonic development	6/273	33/4784	9.8E-03	8.2E-01	4/296	31/3982	2.0E-01	8.8E-01
BP	GO:0001838	embryonic epithelial tube formation	7/273	43/4784	1.0E-02	8.2E-01	6/296	37/3982	5.3E-02	7.8E-01
BP	GO:0010821	regulation of mitochondrion organization	7/273	43/4784	1.0E-02	8.2E-01	9/296	41/3982	2.6E-03	6.2E-01
BP	GO:0030324	lung development	7/273	43/4784	1.0E-02	8.2E-01	7/296	39/3982	2.3E-02	6.3E-01
BP	GO:0032259	methylation	17/273	162/4784	1.0E-02	8.2E-01	16/296	139/3982	5.1E-02	7.8E-01
CC	GO:0016234	inclusion body	6/285	33/4929	1.0E-02	9.8E-01	5/306	27/4103	4.6E-02	9.5E-01
BP	GO:1903052	positive regulation of proteolysis involved in protein catabolic process	8/273	54/4784	1.1E-02	8.2E-01	7/296	49/3982	6.8E-02	8.0E-01
BP	GO:0072583	clathrin-dependent endocytosis	6/273	34/4784	1.1E-02	8.2E-01	6/296	26/3982	1.0E-02	6.2E-01
BP	GO:1902532	negative regulation of intracellular signal transduction	17/273	164/4784	1.2E-02	8.2E-01	14/296	129/3982	9.5E-02	8.5E-01
BP	GO:0036473	cell death in response to oxidative stress	5/273	25/4784	1.2E-02	8.2E-01	4/296	22/3982	7.6E-02	8.2E-01
MF	GO:0008134	transcription factor binding	22/276	235/4877	1.3E-02	4.9E-01	24/299	199/4051	1.0E-02	5.3E-01
BP	GO:0030323	respiratory tube development	7/273	45/4784	1.3E-02	8.2E-01	7/296	41/3982	2.9E-02	6.7E-01
BP	GO:0072175	epithelial tube formation	7/273	45/4784	1.3E-02	8.2E-01	6/296	39/3982	6.5E-02	7.9E-01
BP	GO:0009755	hormone-mediated signaling pathway	9/273	67/4784	1.3E-02	8.2E-01	6/296	53/3982	2.0E-01	8.8E-01
BP	GO:0035601	protein deacylation	8/273	56/4784	1.3E-02	8.2E-01	7/296	48/3982	6.2E-02	7.9E-01
BP	GO:0098732	macromolecule deacylation	8/273	56/4784	1.3E-02	8.2E-01	7/296	48/3982	6.2E-02	7.9E-01
BP	GO:1904888	cranial skeletal system development	4/273	17/4784	1.4E-02	8.2E-01	4/296	15/3982	2.1E-02	6.2E-01
MF	GO:0018024	histone-lysine N-methyltransferase activity	5/276	26/4877	1.4E-02	4.9E-01	5/299	24/4051	2.8E-02	7.9E-01
BP	GO:0031331	positive regulation of cellular catabolic process	18/273	181/4784	1.4E-02	8.2E-01	16/296	156/3982	1.1E-01	8.5E-01
BP	GO:0016331	morphogenesis of embryonic epithelium	7/273	46/4784	1.4E-02	8.2E-01	6/296	40/3982	7.2E-02	8.1E-01
BP	GO:0051093	negative regulation of developmental process	21/273	222/4784	1.4E-02	8.2E-01	22/296	192/3982	2.6E-02	6.3E-01
BP	GO:0032886	regulation of microtubule-based process	11/273	92/4784	1.5E-02	8.2E-01	12/296	73/3982	6.8E-03	6.2E-01
BP	GO:0001843	neural tube closure	6/273	36/4784	1.5E-02	8.2E-01	5/296	31/3982	7.6E-02	8.2E-01
BP	GO:0060606	tube closure	6/273	36/4784	1.5E-02	8.2E-01	5/296	31/3982	7.6E-02	8.2E-01
MF	GO:0061629	RNA polymerase II-specific DNA-binding transcription factor binding	15/276	144/4877	1.5E-02	4.9E-01	16/299	119/4051	1.3E-02	5.3E-01
BP	GO:0040029	epigenetic regulation of gene expression	10/273	81/4784	1.6E-02	8.2E-01	8/296	69/3982	1.4E-01	8.5E-01
BP	GO:0009952	anterior/posterior pattern specification	7/273	47/4784	1.6E-02	8.2E-01	8/296	39/3982	6.8E-03	6.2E-01
MF	GO:0031593	polyubiquitin modification-dependent protein binding	5/276	27/4877	1.6E-02	4.9E-01	4/299	28/4051	1.5E-01	7.9E-01
MF	GO:0140030	modification-dependent protein binding	11/276	94/4877	1.6E-02	4.9E-01	8/299	87/4051	3.1E-01	8.8E-01
BP	GO:0010452	histone H3-K36 methylation	3/273	1053635	1.6E-02	8.2E-01	3/296	760741	4.3E-02	7.3E-01

BP	GO:1901984	negative regulation of protein acetylation	3/273	1053635	1.6E-02	8.2E-01				
MF	GO:0070577	lysine-acetylated histone binding	4/276	18/4877	1.6E-02	4.9E-01	2/299	16/4051	3.3E-01	8.8E-01
MF	GO:0140033	acetylation-dependent protein binding	4/276	18/4877	1.6E-02	4.9E-01	2/299	16/4051	3.3E-01	8.8E-01
BP	GO:0048488	synaptic vesicle endocytosis	5/273	27/4784	1.7E-02	8.2E-01	5/296	22/3982	2.0E-02	6.2E-01
BP	GO:0140238	presynaptic endocytosis	5/273	27/4784	1.7E-02	8.2E-01	5/296	22/3982	2.0E-02	6.2E-01
BP	GO:1902882	regulation of response to oxidative stress	5/273	27/4784	1.7E-02	8.2E-01	4/296	25/3982	1.1E-01	8.5E-01
BP	GO:0001756	somitogenesis	4/273	18/4784	1.7E-02	8.2E-01	4/296	760771	9.2E-03	6.2E-01
BP	GO:0098751	bone cell development	4/273	18/4784	1.7E-02	8.2E-01	4/296	15/3982	2.1E-02	6.2E-01
BP	GO:1901532	regulation of hematopoietic progenitor cell differentiation	4/273	18/4784	1.7E-02	8.2E-01	4/296	15/3982	2.1E-02	6.2E-01
BP	GO:0014020	primary neural tube formation	6/273	37/4784	1.7E-02	8.2E-01	5/296	31/3982	7.6E-02	8.2E-01
BP	GO:0021700	developmental maturation	9/273	70/4784	1.7E-02	8.2E-01	8/296	59/3982	6.8E-02	8.0E-01
BP	GO:0080135	regulation of cellular response to stress	25/273	282/4784	1.7E-02	8.2E-01	32/296	236/3982	5.0E-04	6.2E-01
BP	GO:0016571	histone methylation	10/273	82/4784	1.7E-02	8.2E-01	9/296	72/3982	8.3E-02	8.5E-01
CC	GO:0000785	chromatin	31/285	366/4929	1.9E-02	9.8E-01	33/306	313/4103	2.4E-02	9.5E-01
MF	GO:0008757	S-adenosylmethionine-dependent methyltransferase activity	9/276	72/4877	1.9E-02	4.9E-01	9/299	67/4051	5.6E-02	7.9E-01
BP	GO:0006897	endocytosis	19/273	201/4784	2.0E-02	8.3E-01	19/296	158/3982	2.4E-02	6.3E-01
BP	GO:0060541	respiratory system development	7/273	49/4784	2.0E-02	8.3E-01	7/296	44/3982	4.1E-02	7.3E-01
BP	GO:0022008	neurogenesis	37/273	463/4784	2.0E-02	8.3E-01	39/296	389/3982	2.9E-02	6.7E-01
MF	GO:0042393	histone binding	14/276	136/4877	2.1E-02	4.9E-01	13/299	120/4051	1.0E-01	7.9E-01
BP	GO:0048699	generation of neurons	33/273	404/4784	2.1E-02	8.3E-01	35/296	337/3982	2.4E-02	6.3E-01
MF	GO:0003712	transcription coregulator activity	21/276	232/4877	2.1E-02	4.9E-01	20/299	197/4051	8.7E-02	7.9E-01
CC	GO:0035577	azurophil granule membrane	4/285	19/4929	2.1E-02	9.8E-01	4/306	16/4103	2.7E-02	9.5E-01
BP	GO:0021915	neural tube development	8/273	61/4784	2.1E-02	8.3E-01	7/296	51/3982	8.1E-02	8.3E-01
BP	GO:0009948	anterior/posterior axis specification	3/273	1053666	2.2E-02	8.3E-01				
BP	GO:0060339	negative regulation of type I interferon-mediated signaling pathway	3/273	1053666	2.2E-02	8.3E-01				
BP	GO:0072673	lamellipodium morphogenesis	3/273	1053666	2.2E-02	8.3E-01	3/296	760710	3.3E-02	6.7E-01
BP	GO:1902036	regulation of hematopoietic stem cell differentiation	3/273	1053666	2.2E-02	8.3E-01				
BP	GO:1903202	negative regulation of oxidative stress-induced cell death	3/273	1053666	2.2E-02	8.3E-01	2/296	760710	1.7E-01	8.5E-01
MF	GO:0016278	lysine N-methyltransferase activity	5/276	29/4877	2.2E-02	4.9E-01	6/299	27/4051	1.2E-02	5.3E-01
MF	GO:0016279	protein-lysine N-methyltransferase activity	5/276	29/4877	2.2E-02	4.9E-01	6/299	27/4051	1.2E-02	5.3E-01

BP	GO:0003002	regionalization	9/273	73/4784	2.2E-02	8.3E-01	10/296	62/3982	1.5E-02	6.2E-01
BP	GO:0035295	tube development	24/273	274/4784	2.2E-02	8.3E-01	26/296	233/3982	2.2E-02	6.3E-01
BP	GO:0070828	heterochromatin organization	7/273	50/4784	2.2E-02	8.3E-01	5/296	42/3982	2.0E-01	8.8E-01
BP	GO:0035196	miRNA processing	5/273	29/4784	2.2E-02	8.3E-01	4/296	27/3982	1.4E-01	8.5E-01
MF	GO:0001221	transcription coregulator binding	7/276	51/4877	2.3E-02	4.9E-01	6/299	39/4051	6.4E-02	7.9E-01
BP	GO:0061136	regulation of proteasomal protein catabolic process	9/273	74/4784	2.4E-02	8.3E-01	7/296	69/3982	2.5E-01	9.0E-01
BP	GO:0071383	cellular response to steroid hormone stimulus	9/273	74/4784	2.4E-02	8.3E-01	6/296	59/3982	2.7E-01	9.0E-01
BP	GO:0030219	megakaryocyte differentiation	4/273	20/4784	2.4E-02	8.3E-01	4/296	16/3982	2.7E-02	6.3E-01
BP	GO:0060218	hematopoietic stem cell differentiation	4/273	20/4784	2.4E-02	8.3E-01	4/296	17/3982	3.3E-02	6.7E-01
BP	GO:0061008	hepaticobiliary system development	6/273	40/4784	2.4E-02	8.3E-01	5/296	37/3982	1.4E-01	8.5E-01
BP	GO:0031109	microtubule polymerization or depolymerization	7/273	51/4784	2.4E-02	8.3E-01	8/296	45/3982	1.6E-02	6.2E-01
BP	GO:0030182	neuron differentiation	31/273	380/4784	2.5E-02	8.3E-01	33/296	316/3982	2.6E-02	6.3E-01
BP	GO:0048562	embryonic organ morphogenesis	8/273	63/4784	2.6E-02	8.3E-01	11/296	51/3982	1.0E-03	6.2E-01
BP	GO:0043009	chordate embryonic development	20/273	221/4784	2.6E-02	8.3E-01	19/296	185/3982	9.1E-02	8.5E-01
MF	GO:0019903	protein phosphatase binding	8/276	64/4877	2.7E-02	4.9E-01	8/299	50/4051	2.8E-02	7.9E-01
BP	GO:0034502	protein localization to chromosome	7/273	52/4784	2.7E-02	8.3E-01	6/296	45/3982	1.1E-01	8.5E-01
BP	GO:0035148	tube formation	7/273	52/4784	2.7E-02	8.3E-01	6/296	44/3982	1.0E-01	8.5E-01
BP	GO:0045814	negative regulation of gene expression, epigenetic	7/273	52/4784	2.7E-02	8.3E-01	5/296	43/3982	2.1E-01	8.9E-01
MF	GO:0030507	spectrin binding	3/276	1087664	2.7E-02	4.9E-01				
MF	GO:0035173	histone kinase activity	3/276	1087664	2.7E-02	4.9E-01	3/299	785973	5.3E-02	7.9E-01
MF	GO:0050321	tau-protein kinase activity	3/276	1087664	2.7E-02	4.9E-01	2/299	785973	2.2E-01	7.9E-01
BP	GO:0048598	embryonic morphogenesis	15/273	153/4784	2.7E-02	8.3E-01	17/296	127/3982	1.2E-02	6.2E-01
BP	GO:0006517	protein deglycosylation	3/273	1053696	2.8E-02	8.3E-01	2/296	760710	1.7E-01	8.5E-01
BP	GO:0021872	forebrain generation of neurons	3/273	1053696	2.8E-02	8.3E-01				
BP	GO:0032211	negative regulation of telomere maintenance via telomerase	3/273	1053696	2.8E-02	8.3E-01				
BP	GO:0035855	megakaryocyte development	3/273	1053696	2.8E-02	8.3E-01	3/296	760710	3.3E-02	6.7E-01
BP	GO:0048701	embryonic cranial skeleton morphogenesis	3/273	1053696	2.8E-02	8.3E-01	3/296	760710	3.3E-02	6.7E-01
MF	GO:0140658	ATP-dependent chromatin remodeler activity	4/276	21/4877	2.8E-02	4.9E-01	5/299	20/4051	1.3E-02	5.3E-01
BP	GO:0000902	cell morphogenesis	27/273	324/4784	2.8E-02	8.3E-01	26/296	261/3982	7.2E-02	8.1E-01
BP	GO:0048666	neuron development	27/273	324/4784	2.8E-02	8.3E-01	30/296	269/3982	1.4E-02	6.2E-01

MF	GO:0043130	ubiquitin binding	7/276	53/4877	2.8E-02	4.9E-01	6/299	46/4051	1.2E-01	7.9E-01
BP	GO:0030520	intracellular estrogen receptor signaling pathway	4/273	21/4784	2.9E-02	8.3E-01	1/296	18/3982	7.5E-01	9.6E-01
BP	GO:0051148	negative regulation of muscle cell differentiation	4/273	21/4784	2.9E-02	8.3E-01	3/296	17/3982	1.3E-01	8.5E-01
BP	GO:0031329	regulation of cellular catabolic process	27/273	325/4784	2.9E-02	8.3E-01	28/296	280/3982	6.1E-02	7.9E-01
BP	GO:0070918	small regulatory ncRNA processing	5/273	31/4784	2.9E-02	8.3E-01	4/296	29/3982	1.6E-01	8.5E-01
BP	GO:0001501	skeletal system development	13/273	128/4784	3.0E-02	8.3E-01	14/296	109/3982	3.0E-02	6.7E-01
MF	GO:0046914	transition metal ion binding	26/276	314/4877	3.0E-02	5.0E-01	23/299	274/4051	2.9E-01	8.5E-01
BP	GO:0048469	cell maturation	6/273	42/4784	3.0E-02	8.4E-01	5/296	37/3982	1.4E-01	8.5E-01
BP	GO:0006605	protein targeting	12/273	116/4784	3.2E-02	8.5E-01	11/296	106/3982	1.6E-01	8.5E-01
BP	GO:0009792	embryo development ending in birth or egg hatching	20/273	226/4784	3.2E-02	8.5E-01	19/296	189/3982	1.1E-01	8.5E-01
MF	GO:0003684	damaged DNA binding	6/276	43/4877	3.2E-02	5.1E-01	5/299	36/4051	1.2E-01	7.9E-01
BP	GO:0031110	regulation of microtubule polymerization or depolymerization	5/273	32/4784	3.3E-02	8.5E-01	6/296	26/3982	1.0E-02	6.2E-01
BP	GO:0033143	regulation of intracellular steroid hormone receptor signaling pathway	5/273	32/4784	3.3E-02	8.5E-01	3/296	29/3982	3.7E-01	9.4E-01
BP	GO:0006338	chromatin remodeling	15/273	157/4784	3.3E-02	8.5E-01	15/296	136/3982	7.8E-02	8.3E-01
BP	GO:0009894	regulation of catabolic process	31/273	389/4784	3.3E-02	8.5E-01	31/296	337/3982	1.2E-01	8.5E-01
BP	GO:0060998	regulation of dendritic spine development	4/273	22/4784	3.4E-02	8.5E-01	4/296	20/3982	5.6E-02	7.8E-01
BP	GO:0071539	protein localization to centrosome	4/273	22/4784	3.4E-02	8.5E-01	4/296	19/3982	4.8E-02	7.5E-01
BP	GO:1901216	positive regulation of neuron death	4/273	22/4784	3.4E-02	8.5E-01	5/296	17/3982	6.4E-03	6.2E-01
BP	GO:0045620	negative regulation of lymphocyte differentiation	3/273	13/4784	3.4E-02	8.6E-01	3/296	760741	4.3E-02	7.3E-01
CC	GO:0030135	coated vesicle	11/285	103/4929	3.4E-02	9.8E-01	9/306	76/4103	1.1E-01	9.5E-01
BP	GO:0009968	negative regulation of signal transduction	30/273	375/4784	3.5E-02	8.6E-01	28/296	298/3982	1.1E-01	8.5E-01
CC	GO:0005791	rough endoplasmic reticulum	4/285	22/4929	3.5E-02	9.8E-01	4/306	21/4103	6.6E-02	9.5E-01
BP	GO:0002244	hematopoietic progenitor cell differentiation	7/273	55/4784	3.5E-02	8.6E-01	8/296	49/3982	2.6E-02	6.3E-01
BP	GO:0008360	regulation of cell shape	7/273	55/4784	3.5E-02	8.6E-01	6/296	44/3982	1.0E-01	8.5E-01
BP	GO:0071695	anatomical structure maturation	7/273	55/4784	3.5E-02	8.6E-01	6/296	48/3982	1.4E-01	8.5E-01
CC	GO:0000932	P-body	6/285	43/4929	3.6E-02	9.8E-01	6/306	40/4103	7.3E-02	9.5E-01
BP	GO:0007389	pattern specification process	11/273	105/4784	3.6E-02	8.6E-01	12/296	90/3982	3.3E-02	6.7E-01
MF	GO:0140297	DNA-binding transcription factor binding	17/276	188/4877	3.6E-02	5.5E-01	19/299	155/4051	1.8E-02	6.5E-01
BP	GO:0021953	central nervous system neuron differentiation	6/273	44/4784	3.7E-02	8.8E-01	7/296	36/3982	1.5E-02	6.2E-01
BP	GO:0032434	regulation of proteasomal ubiquitin-dependent protein catabolic process	7/273	56/4784	3.8E-02	8.9E-01	5/296	53/3982	3.6E-01	9.4E-01

BP	GO:0055072	iron ion homeostasis	4/273	23/4784	3.9E-02	8.9E-01	2/296	21/3982	4.7E-01	9.5E-01
BP	GO:0061053	somite development	4/273	23/4784	3.9E-02	8.9E-01	4/296	16/3982	2.7E-02	6.3E-01
BP	GO:1905508	protein localization to microtubule organizing center	4/273	23/4784	3.9E-02	8.9E-01	4/296	19/3982	4.8E-02	7.5E-01
BP	GO:1905897	regulation of response to endoplasmic reticulum stress	4/273	23/4784	3.9E-02	8.9E-01	4/296	20/3982	5.6E-02	7.8E-01
CC	GO:0000151	ubiquitin ligase complex	14/285	146/4929	4.2E-02	9.8E-01	12/306	118/4103	1.7E-01	9.5E-01
BP	GO:0036465	synaptic vesicle recycling	5/273	34/4784	4.2E-02	9.2E-01	5/296	28/3982	5.2E-02	7.8E-01
BP	GO:0018193	peptidyl-amino acid modification	36/273	473/4784	4.2E-02	9.2E-01	42/296	395/3982	9.2E-03	6.2E-01
BP	GO:0006801	superoxide metabolic process	3/273	14/4784	4.2E-02	9.2E-01	3/296	13/3982	6.7E-02	7.9E-01
BP	GO:0043627	response to estrogen	3/273	14/4784	4.2E-02	9.2E-01	2/296	13/3982	2.5E-01	9.0E-01
CC	GO:0016604	nuclear body	30/285	377/4929	4.3E-02	9.8E-01	38/306	318/4103	2.0E-03	6.9E-01
BP	GO:0048863	stem cell differentiation	8/273	70/4784	4.4E-02	9.5E-01	9/296	56/3982	2.1E-02	6.2E-01
BP	GO:0030865	cortical cytoskeleton organization	4/273	24/4784	4.5E-02	9.5E-01	3/296	23/3982	2.4E-01	9.0E-01
BP	GO:0034340	response to type I interferon	4/273	24/4784	4.5E-02	9.5E-01	3/296	19/3982	1.6E-01	8.5E-01
BP	GO:0045596	negative regulation of cell differentiation	15/273	164/4784	4.6E-02	9.5E-01	15/296	141/3982	9.9E-02	8.5E-01
MF	GO:0008022	protein C-terminus binding	9/276	84/4877	4.6E-02	6.8E-01	11/299	63/4051	5.6E-03	5.3E-01
BP	GO:0061515	myeloid cell development	5/273	35/4784	4.6E-02	9.5E-01	4/296	31/3982	2.0E-01	8.8E-01
BP	GO:0070936	protein K48-linked ubiquitination	5/273	35/4784	4.6E-02	9.5E-01	3/296	32/3982	4.3E-01	9.5E-01
BP	GO:0030097	hemopoiesis	25/273	310/4784	4.7E-02	9.5E-01	25/296	264/3982	1.2E-01	8.5E-01
CC	GO:0016323	basolateral plasma membrane	8/285	70/4929	4.7E-02	9.8E-01	6/306	56/4103	2.4E-01	9.5E-01
MF	GO:0008170	N-methyltransferase activity	6/276	47/4877	4.7E-02	6.8E-01	6/299	45/4051	1.1E-01	7.9E-01
BP	GO:0030099	myeloid cell differentiation	13/273	137/4784	4.8E-02	9.5E-01	11/296	115/3982	2.3E-01	9.0E-01
BP	GO:0016358	dendrite development	9/273	84/4784	4.9E-02	9.5E-01	9/296	74/3982	9.5E-02	8.5E-01
BP	GO:1901800	positive regulation of proteasomal protein catabolic process	6/273	47/4784	4.9E-02	9.5E-01	5/296	44/3982	2.3E-01	8.9E-01

Table 8.2. Gene ontology enrichment of abundant circular RNA-host genes. Ontologies include Biological Processes (BP), Molecular Functions (MF) and Cellular Components (CC). For each term, the corresponding ontology, unique GO ID, the name of the term is shown. For each cohort, I give the Gene ratio (number of genes in the GO term that host abundant circRNAs), Background ratio (number of genes in the GO term that did not host abundant circRNAs), overrepresentation *P*-value and the false discovery rate of the P-value (FDR). Blank indicates not available.

				Detected		**Abundant**			

BSJ position	Gene symbol	Tissue	Direction	*PPMI*	*ICICLE-PD*	*PPMI*	*ICICLE-PD*
21:36247516-36248568:+	*DOP1B*	Blood	Decreased in PD	YES	YES	YES	YES
X:135545422-135556300:+	*INTS6L*	Brain	Decreased in PD	YES	YES	YES	YES
6:138943512-138944622:-	*RESP1*	Blood	Decreased in PD	YES	YES	YES	NO
2:40428472-40430304:-	*SLC8A1*	Brain	Increased in PD	YES	YES	YES	YES
1:47280240-47282459:-	*STIL*	Brain	Decreased in PD	YES	YES	YES	YES
16:18797924-18798834:-	*ARL6IP1*	Brain	Decreased in PD	YES	YES	YES	YES
13:77719531-77753358:+	*SLAIN1*	Blood	Decreased in PD	YES	YES	NO	YES
14:73147794-73148106:+	*PSEN1*	Blood	Decreased in PD	YES	YES	NO	NO
5:180261683-180280608:-	*MAPK9*	Blood	Decreased in PD	YES	YES	NO	NO
5:79439009-79457018:-	*HOMER1*	Blood	Decreased in PD	YES	YES	NO	NO
9:84702158-84710791:+	*NTRK2*	Brain	Increased in PD	NO	NO	NO	NO
2:226906792-226914351:+	*RHBDD1*	Brain	Decreased in PD	YES	YES	NO	NO
5:132927120-132934941:-	*AFF4*	Brain	Decreased in PD	YES	YES	NO	NO
8:38819521-38820635:+	*TACC1*	Brain	Decreased in PD	YES	YES	NO	NO
20:41267462-41269083:-	*ZHX3*	Brain	Increased in PD	YES	YES	NO	NO
4:112562370-112585725:-	*ZGRF1*	Brain	Decreased in PD	YES	YES	NO	NO
15:87929190-87940753:-	*NTRK3*	Brain	Decreased in PD	NO	NO	NO	NO
10:68393137-68394451:-	*RUFY2*	Brain	Decreased in PD	YES	YES	NO	NO
12:40028300-40092933:-	*SLC2A13*	Brain	Decreased in PD	NO	NO	NO	NO
10:20868663-20899450:-	*NEBL*	Brain	Decreased in PD	NO	NO	NO	NO
6:68930558-68956809:+	*ADGRB3*	Brain	Decreased in PD	NO	NO	NO	NO
3:11807832-11809682:-	*TAMM41*	Brain	Decreased in PD	YES	YES	NO	NO
9:83659779-83686155:-	*UBQLN1*	Brain	Increased in PD	NO	NO	NO	NO
1:243545509-243615161:-	*AKT3*	Brain	Decreased in PD	NO	NO	NO	NO
2:206841106-206881891:+	*NA*	Brain	Decreased in PD	NO	NO	NO	NO
9:6880011-6893232:+	*KDM4C*	Brain	Decreased in PD	YES	NO	NO	NO
4:150761782-150808398:-	*LRBA*	Brain	Increased in PD	YES	YES	NO	NO
1:8495062-8557523:-	*RERE*	Brain	Increased in PD	YES	YES	NO	NO
10:20880793-20899450:-	*NEBL*	Brain	Increased in PD	NO	NO	NO	NO

5:168488601-168494650:+	*RARS1*	Brain	Increased in PD	YES	YES	NO	NO

Table 8.3. Previously reported differentially expressed circular RNAs in PD. Summary statistics of were taken from the results of Hanan et al (Brain) and Ravanidis et al (Blood). Direction describes the change in circRNA expression in PD patients according to the respective publications. For each circRNA, I report whether it was detected in each cohort, and subsequently whether it was abundantly expressed and therefore included in differential expression analysis. CircRNAs are given as unique back-spliced junction positions reported according the GRCh38 reference.

		Mean RNA Editing Index (SD)	
Region	**Base mismatch**	*PPMI*	*ICICLE-PD*
Alu elements	A-C	0.01 (0)	0.01 (0)
	A-G	0.85 (0.12)	0.85 (0.08)
	A-T	0.01 (0)	0.01 (0)
	C-A	0.01 (0)	0.01 (0)
	C-G	0.01 (0)	0 (0)
	C-T	0.04 (0)	0.03 (0)
circRNAs	A-C	0.01 (0)	0 (0)
	A-G	0.07 (0.02)	0.05 (0.02)
	A-T	0.01 (0)	0.01 (0)
	C-A	0.01 (0.01)	0 (0)
	C-G	0.01 (0)	0 (0)
	C-T	0.03 (0)	0.03 (0)
Flanking circRNAs	A-C	0.01 (0)	0 (0)
	A-G	0.08 (0.02)	0.06 (0.02)
	A-T	0.01 (0)	0.01 (0)
	C-A	0.01 (0.01)	0.01 (0)
	C-G	0.01 (0)	0 (0)
	C-T	0.03 (0)	0.03 (0)

Table 8.4. RNA Editing indices.

Chapter 9. References

Aamodt, W.W., Waligorska, T., Shen, J., Tropea, T.F., Siderowf, A., Weintraub, D., Grossman, M., Irwin, D., Wolk, D.A., Xie, S.X., Trojanowski, J.Q., Shaw, L.M. & Chen-Plotkin, A.S. (2021) 'Neurofilament Light Chain as a Biomarker for Cognitive Decline in Parkinson Disease', *Movement Disorders*, 36(12), pp. 2945–2950.

Aarsland, D., Batzu, L., Halliday, G.M., Geurtsen, G.J., Ballard, C., Ray Chaudhuri, K. & Weintraub, D. (2021) 'Parkinson disease-associated cognitive impairment', *Nature Reviews Disease Primers*, 7(1), pp. 1–21.

Aasly, J.O. (2020) 'Long-Term Outcomes of Genetic Parkinson's Disease', *Journal of Movement Disorders*, 13(2), pp. 81–96.

Advani, V.M. & Ivanov, P. (2019) 'Translational Control under Stress: Reshaping the Translatome', *BioEssays*, 41(5), p. 1900009.

Ahmad, S., Mu, X., Yang, F., Greenwald, E., Park, J.W., Jacob, E., Zhang, C.-Z. & Hur, S. (2018) 'Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation', *Cell*, 172(4), pp. 797-810.e13.

Aktaş, T., Avşar Ilık, İ., Maticzka, D., Bhardwaj, V., Pessoa Rodrigues, C., Mittler, G., Manke, T., Backofen, R. & Akhtar, A. (2017) 'DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome', *Nature*, 544(7648), pp. 115–119.

Alam, M.M., Yang, D., Li, X.-Q., Liu, J., Back, T.C., Trivett, A., Karim, B., Barbut, D., Zasloff, M. & Oppenheim, J.J. (2022) 'Alpha synuclein, the culprit in Parkinson disease, is required for normal immune function', *Cell Reports*, 38(2), p. 110090.

Al-Balool, H.H., Weber, D., Liu, Y., Wade, M., Guleria, K., Nam, P.L.P., Clayton, J., Rowe, W., Coxhead, J., Irving, J., Elliott, D.J., Hall, A.G., Santibanez-Koref, M. & Jackson, M.S. (2011) 'Post-transcriptional exon shuffling events in humans can be evolutionarily conserved and abundant', *Genome Research*, 21(11), pp. 1788–1799.

Alcalá-Vida, R., Garcia-Forn, M., Castany-Pladevall, C., Creus-Muncunill, J., Ito, Y., Blanco, E., Golbano, A., Crespí-Vázquez, K., Parry, A., Slater, G., Samarajiwa, S., Peiró, S., Di Croce, L., Narita, M. & Pérez-Navarro, E. (2021) 'Neuron type-specific increase in lamin B1 contributes to nuclear dysfunction in Huntington's disease', *EMBO Molecular Medicine*, 13(2), p. e12105.

Alhasan, A.A., Izuogu, O.G., Al-Balool, H.H., Steyn, J.S., Evans, A., Colzani, M., Ghevaert, C., Mountford, J.C., Marenah, L., Elliott, D.J., Santibanez-Koref, M. & Jackson, M.S. (2016) 'Circular RNA enrichment in platelets is a signature of transcriptome degradation', *Blood*, 127(9), pp. e1–e11.

Al-Kasbi, G., Al-Saegh, A., Al-Qassabi, A., Al-Jabry, T., Zadjali, F., Al-Yahyaee, S. & Al-Maawali, A. (2021) 'Biallelic PTRHD1 Frameshift Variants Associated with Intellectual Disability, Spasticity, and Parkinsonism', *Movement Disorders Clinical Practice*, 8(8), pp. 1253–1257.

Amaya, L., Abe, B., Liu, J., Zhao, F., Zhang, W.L., Chen, R., Li, R., Wang, S., Kamber, R.A., Tsai, M.-C., Bassik, M.C., Majeti, R. & Chang, H.Y. (2024) 'Pathways for macrophage uptake of cell-free circular RNAs', *Molecular Cell*,

Amaya, L., Grigoryan, L., Li, Z., Lee, A., Wender, P.A., Pulendran, B. & Chang, H.Y. (2023) 'Circular RNA vaccine induces potent T cell responses', *Proceedings of the National Academy of Sciences*, 120(20), p. e2302191120.

An, C., Park, Y.W., Ahn, S.S., Han, K., Kim, H. & Lee, S.-K. (2021) 'Radiomics machine learning study with a small sample size: Single random training-test set split may lead to unreliable results', *PLOS ONE*, 16(8), p. e0256152.

Anandhan, A., Jacome, M.S., Lei, S., Hernandez-Franco, P., Pappa, A., Panayiotidis, M.I., Powers, R. & Franco, R. (2017) 'Metabolic Dysfunction in Parkinson's Disease: Bioenergetics, Redox Homeostasis and Central Carbon Metabolism', *Brain Research Bulletin*, 133pp. 12–30.

Anders, S. & Huber, W. (2010) 'Differential expression analysis for sequence count data', *Genome Biology*, 11(10), p. R106.

Andersen, J.B., Mazan-Mamczarz, K., Zhan, M., Gorospe, M. & Hassel, B.A. (2009) 'Ribosomal protein mRNAs are primary targets of regulation in RNase-L-induced senescence', *RNA Biology*, 6(3), pp. 305–315.

Anheim, M., Elbaz, A., Lesage, S., Durr, A., Condroyer, C., Viallet, F., Pollak, P., Bonaïti, B., Bonaïti-Pellié, C., Brice, A., & French Parkinson Disease Genetic Group (2012) 'Penetrance of Parkinson disease in glucocerebrosidase gene mutation carriers', *Neurology*, 78(6), pp. 417–420.

Antonini, A., Berto, P., Lopatriello, S., Tamma, F., Annemans, L. & Chambers, M. (2008) 'Cost-effectiveness of 123I-FP-CIT SPECT in the differential diagnosis of essential tremor and Parkinson's disease in Italy', *Movement Disorders: Official Journal of the Movement Disorder Society*, 23(15), pp. 2202–2209.

Appleton, E., Khosousi, S., Ta, M., Nalls, M., Singleton, A.B., Sturchio, A., Markaki, I., Paslawski, W., Iwaki, H. & Svenningsson, P. (2024) 'DOPA-decarboxylase is elevated in CSF, but not plasma, in prodromal and de novo Parkinson's disease', *Translational Neurodegeneration*, 13(1), p. 31.

Argüello, R.J., Rodrigues, C.R., Gatti, E. & Pierre, P. (2015) 'Protein synthesis regulation, a pillar of strength for innate immunity?', *Current Opinion in Immunology*, 32pp. 28–35.

Arnberg, A.C., Van Ommen, G.J., Grivell, L.A., Van Bruggen, E.F. & Borst, P. (1980) 'Some yeast mitochondrial RNAs are circular', *Cell*, 19(2), pp. 313–319.

Arshad, U., Rahman, F., Hanan, N. & Chen, C. (2023) 'Longitudinal Meta-Analysis of Historical Parkinson's Disease Trials to Inform Future Trial Design', *Movement Disorders*, 38(9), pp. 1716–1727.

Ascherio, A., LeWitt, P.A., Xu, K., Eberly, S., Watts, A., Matson, W.R., Marras, C., Kieburtz, K., Rudolph, A., Bogdanov, M.B., Schwid, S.R., Tennis, M., Tanner, C.M., Beal, M.F., Lang, A.E., Oakes, D., Fahn, S., Shoulson, I., Schwarzschild, M.A., et al. (2009) 'Urate as a Predictor of the Rate of Clinical Decline in Parkinson Disease', *Archives of Neurology*, 66(12), pp. 1460–1468.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L.,

Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., et al. (2000) 'Gene Ontology: tool for the unification of biology', *Nature genetics*, 25(1), pp. 25–29.

Ashton, N.J., Janelidze, S., Al Khleifat, A., Leuzy, A., van der Ende, E.L., Karikari, T.K., Benedet, A.L., Pascoal, T.A., Lleó, A., Parnetti, L., Galimberti, D., Bonanni, L., Pilotto, A., Padovani, A., Lycke, J., Novakova, L., Axelsson, M., Velayudhan, L., Rabinovici, G.D., et al. (2021) 'A multicentre validation study of the diagnostic value of plasma neurofilament light', *Nature Communications*, 12(1), p. 3400.

Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N. & Kadener, S. (2014) 'circRNA Biogenesis Competes with Pre-mRNA Splicing', *Molecular Cell*, 56(1), pp. 55–66.

Atarashi, R., Wilham, J.M., Christensen, L., Hughson, A.G., Moore, R.A., Johnson, L.M., Onwubiko, H.A., Priola, S.A. & Caughey, B. (2008) 'Simplified ultrasensitive prion detection by recombinant PrP conversion with shaking', *Nature Methods*, 5(3), pp. 211–212.

Athanasiadis, A., Rich, A. & Maas, S. (2004) 'Widespread A-to-I RNA Editing of Alu-Containing mRNAs in the Human Transcriptome', *PLOS Biology*, 2(12), p. e391.

Athauda, D. & Foltynie, T. (2016) 'Challenges in detecting disease modification in Parkinson's disease clinical trials', *Parkinsonism & Related Disorders*, 32pp. 1–11.

Avila Cobos, F., Alquicira-Hernandez, J., Powell, J.E., Mestdagh, P. & De Preter, K. (2020) 'Benchmarking of cell type deconvolution pipelines for transcriptomics data', *Nature Communications*, 11(1), p. 5650.

Baba, Y., Kuroiwa, A., Uitti, R.J., Wszolek, Z.K. & Yamada, T. (2005) 'Alterations of T-lymphocyte populations in Parkinson disease', *Parkinsonism & Related Disorders*, 11(8), pp. 493–498.

Bachmayr-Heyda, A., Reiner, A.T., Auer, K., Sukhbaatar, N., Aust, S., Bachleitner-Hofmann, T., Mesteri, I., Grunt, T.W., Zeillinger, R. & Pils, D. (2015) 'Correlation of circular RNA abundance with proliferation – exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues', *Scientific Reports*, 5p. 8057.

Bäckström, D.C., Eriksson Domellöf, M., Linder, J., Olsson, B., Öhrfelt, A., Trupp, M., Zetterberg, H., Blennow, K. & Forsgren, L. (2015) 'Cerebrospinal Fluid Patterns and the Risk of Future Dementia in Early, Incident Parkinson Disease', *JAMA Neurology*, 72(10), pp. 1175–1182.

Bahn, J.H., Zhang, Q., Li, F., Chan, T.-M., Lin, X., Kim, Y., Wong, D.T.W. & Xiao, X. (2015) 'The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva', *Clinical Chemistry*, 61(1), pp. 221–230.

Balestrino, R., Tunesi, S., Tesei, S., Lopiano, L., Zecchinelli, A.L. & Goldwurm, S. (2020) 'Penetrance of Glucocerebrosidase (GBA) Mutations in Parkinson's Disease: A Kin Cohort Study', *Movement Disorders*, 35(11), pp. 2111–2114.

Bando, Y., Onuki, R., Katayama, T., Manabe, T., Kudo, T., Taira, K. & Tohyama, M. (2005) 'Double-strand RNA dependent protein kinase (PKR) is involved in the extrastriatal

degeneration in Parkinson's disease and Huntington's disease', *Neurochemistry International*, 46(1), pp. 11–18.

Barbour, R., Kling, K., Anderson, J.P., Banducci, K., Cole, T., Diep, L., Fox, M., Goldstein, J.M., Soriano, F., Seubert, P. & Chilcote, T.J. (2008) 'Red Blood Cells Are the Major Source of Alpha-Synuclein in Blood', *Neurodegenerative Diseases*, 5(2), pp. 55–59.

Barcia, C., Ros, C.M., Annese, V., Gómez, A., Ros-Bernal, F., Aguado-Llera, D., Martínez-Pagán, M.E., de Pablos, V., Fernandez-Villalba, E. & Herrero, M.T. (2011) 'IFN-γ signaling, with the synergistic contribution of TNF-α, mediates cell specific microglial and astroglial activation in experimental models of Parkinson's disease', *Cell Death & Disease*, 2(4), pp. e142–e142.

Bargar, C., De Luca, C.M.G., Devigili, G., Elia, A.E., Cilia, R., Portaleone, S.M., Wang, W., Tramacere, I., Bistaffa, E., Cazzaniga, F.A., Felisati, G., Legname, G., Di Fonzo, A., Xu, R., Gunzler, S.A., Giaccone, G., Eleopra, R., Chen, S.G. & Moda, F. (2021) 'Discrimination of MSA-P and MSA-C by RT-QuIC analysis of olfactory mucosa: the first assessment of assay reproducibility between two specialized laboratories', *Molecular Neurodegeneration*, 16p. 82.

Barkovits, K., Kruse, N., Linden, A., Tönges, L., Pfeiffer, K., Mollenhauer, B. & Marcus, K. (2020) 'Blood Contamination in CSF and Its Impact on Quantitative Analysis of Alpha-Synuclein', *Cells*, 9(2), p. 370.

Barrett, S.P., Wang, P.L. & Salzman, J. (2015) 'Circular RNA biogenesis can proceed through an exon-containing lariat precursor' Douglas L Black (ed.), *eLife*, 4p. e07540.

Bar-Shira, A., Hutter, C.M., Giladi, N., Zabetian, C.P. & Orr-Urtreger, A. (2009) 'Ashkenazi Parkinson's disease patients with the LRRK2 G2019S mutation share a common founder dating from the second to fifth centuries', *neurogenetics*, 10(4), pp. 355–358.

Barthélemy, N.R., Salvadó, G., Schindler, S.E., He, Y., Janelidze, S., Collij, L.E., Saef, B., Henson, R.L., Chen, C.D., Gordon, B.A., Li, Y., La Joie, R., Benzinger, T.L.S., Morris, J.C., Mattsson-Carlgren, N., Palmqvist, S., Ossenkoppele, R., Rabinovici, G.D., Stomrud, E., et al. (2024) 'Highly accurate blood test for Alzheimer's disease is similar or superior to clinical cerebrospinal fluid tests', *Nature Medicine*, 30(4), pp. 1085–1095.

Bass, B.L. & Weintraub, H. (1988) 'An unwinding activity that covalently modifies its double-stranded RNA substrate', *Cell*, 55(6), pp. 1089–1098.

Batzu, L., Rota, S., Hye, A., Heslegrave, A., Trivedi, D., Gibson, L.L., Farrell, C., Zinzalias, P., Rizos, A., Zetterberg, H., Chaudhuri, K.R. & Aarsland, D. (2022) 'Plasma p-tau181, neurofilament light chain and association with cognition in Parkinson's disease', *npj Parkinson's Disease*, 8(1), pp. 1–7.

Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., Isaacs, F.J., Rechavi, G., Li, J.B., Eisenberg, E. & Levanon, E.Y. (2014) 'A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes', *Genome Research*, 24(3), pp. 365–376.

Beach, T.G. & Adler, C.H. (2018) 'Importance of low diagnostic Accuracy for early Parkinson's disease', *Movement Disorders*, 33(10), pp. 1551–1554.

Beach, T.G., Adler, C.H., Lue, L., Sue, L.I., Bachalakuri, J., Henry-Watson, J., Sasse, J., Boyer, S., Shirohi, S., Brooks, R., Eschbacher, J., White, C.L., Akiyama, H., Caviness, J., Shill, H.A., Connor, D.J., Sabbagh, M.N., Walker, D.G., & the Arizona Parkinson's Disease Consortium (2009) 'Unified staging system for Lewy body disorders: correlation with nigrostriatal degeneration, cognitive impairment and motor dysfunction', *Acta Neuropathologica*, 117(6), pp. 613–634.

Beach, T.G., Adler, C.H., Sue, L.I., Vedders, L., Lue, L., White III, C.L., Akiyama, H., Caviness, J.N., Shill, H.A., Sabbagh, M.N., Walker, D.G., & Arizona Parkinson's Disease Consortium (2010) 'Multi-organ distribution of phosphorylated α-synuclein histopathology in subjects with Lewy body disorders', *Acta Neuropathologica*, 119(6), pp. 689–702.

Beatman, E.L., Massey, A., Shives, K.D., Burrack, K.S., Chamanian, M., Morrison, T.E. & Beckham, J.D. (2016) 'Alpha-Synuclein Expression Restricts RNA Viral Infections in the Brain', *Journal of Virology*, 90(6), pp. 2767–2782.

Beavan, M.S. & Schapira, A.H.V. (2013) 'Glucocerebrosidase mutations and the pathogenesis of Parkinson disease', *Annals of Medicine*, 45(8), pp. 511–521.

Ben El Haj, R., Salmi, A., Regragui, W., Moussa, A., Bouslam, N., Tibar, H., Benomar, A., Yahyaoui, M. & Bouhouche, A. (2017) 'Evidence for prehistoric origins of the G2019S mutation in the North African Berber population', *PLoS ONE*, 12(7), p. e0181335.

Benamer, H.T.S., Patterson, J., Grosset, D.G., Booij, J., de Bruin, K., van Royen, E., Speelman, J.D., Horstink, M.H.I.M., Sips, H.J.W.A., Dierckx, R.A., Versijpt, J., Decoo, D., Van Der Linden, C., Hadley, D.M., Doder, M., Lees, A.J., Costa, D.C., Gacinovic, S., Oertel, W.H., et al. (2000) 'Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: The [123I]-FP-CIT study group', *Movement Disorders*, 15(3), pp. 503–510.

Benjamini, Y. & Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), pp. 289–300.

Benoit, S.M., Xu, H., Schmid, S., Alexandrova, R., Kaur, G., Thiruvahindrapuram, B., Pereira, S.L., Jog, M. & Hebb, M.O. (2020) 'Expanding the search for genetic biomarkers of Parkinson's disease into the living brain', *Neurobiology of Disease*, 140p. 104872.

Berg, D., Postuma, R.B., Adler, C.H., Bloem, B.R., Chan, P., Dubois, B., Gasser, T., Goetz, C.G., Halliday, G., Joseph, L., Lang, A.E., Liepelt-Scarfone, I., Litvan, I., Marek, K., Obeso, J., Oertel, W., Olanow, C.W., Poewe, W., Stern, M., et al. (2015) 'MDS research criteria for prodromal Parkinson's disease', *Movement Disorders*, 30(12), pp. 1600–1611.

Beric, A., Sun, Y., Sanchez, S., Martin, C., Powell, T., Pardo, J.A., Sanford, J., Botia, J.A., Cruchaga, C. & Ibanez, L. (2024) *Circulating blood circular RNA in Parkinson's Disease; a systematic study*. p.p. 2024.01.22.24301623.

Bestwick, J.P., Auger, S.D., Schrag, A.E., Grosset, D.G., Kanavou, S., Giovannoni, G., Lees, A.J., Cuzick, J. & Noyce, A.J. (2021) 'Optimising classification of Parkinson's disease based on motor, olfactory, neuropsychiatric and sleep features', *npj Parkinson's Disease*, 7(1), pp. 1–11.

Beutler, E., Gelbart, T. & Scott, C.R. (2005) 'Hematologically important mutations: Gaucher disease', *Blood Cells, Molecules, and Diseases*, 35(3), pp. 355–364.

Biskup, S. & West, A.B. (2009) 'Zeroing in on LRRK2-linked pathogenic mechanisms in Parkinson's disease', *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1792(7), pp. 625–633.

Blauwendraat, C., Nalls, M.A. & Singleton, A.B. (2020) 'The genetic architecture of Parkinson's disease', *The Lancet. Neurology*, 19(2), pp. 170–178.

Bonnet, M.C., Weil, R., Dam, E., Hovanessian, A.G. & Meurs, E.F. (2000) 'PKR Stimulates NF-κB Irrespective of Its Kinase Function by Interacting with the IκB Kinase Complex', *Molecular and Cellular Biology*, 20(13), pp. 4532–4542.

Borghammer, P. & Van Den Berge, N. (2019) 'Brain-First versus Gut-First Parkinson's Disease: A Hypothesis', *Journal of Parkinson's Disease*, 9(s2), pp. S281–S295.

Borrageiro, G., Haylett, W., Seedat, S., Kuivaniemi, H. & Bardien, S. (2018) 'A review of genome-wide transcriptomics studies in Parkinson's disease', *European Journal of Neuroscience*, 47(1), pp. 1–16.

Bosson, A.D., Zamudio, J.R. & Sharp, P.A. (2014) 'Endogenous miRNA and Target Concentrations Determine Susceptibility to Potential ceRNA Competition', *Molecular Cell*, 56(3), pp. 347–359.

Braak, H., Del Tredici, K., Bratzke, H., Hamm-Clement, J., Sandmann-Keil, D. & Rüb, U. (2002) 'Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages)', *Journal of Neurology*, 249(3), pp. iii1–iii5.

Braak, H., Del Tredici, K., Rüb, U., de Vos, R.A.I., Jansen Steur, E.N.H. & Braak, E. (2003) 'Staging of brain pathology related to sporadic Parkinson's disease', *Neurobiology of Aging*, 24(2), pp. 197–211.

Braak, H., Rüb, U., Gai, W.P. & Del Tredici, K. (2003) 'Idiopathic Parkinson's disease: possible routes by which vulnerable neuronal types may be subject to neuroinvasion by an unknown pathogen', *Journal of Neural Transmission*, 110(5), pp. 517–536.

Bräuer, S., Rossi, M., Sajapin, J., Henle, T., Gasser, T., Parchi, P., Brockmann, K. & Falkenburger, B.H. (2023) 'Kinetic parameters of alpha-synuclein seed amplification assay correlate with cognitive impairment in patients with Lewy body disorders', *Acta Neuropathologica Communications*, 11(1), p. 162.

Bressan, E., Reed, X., Bansal, V., Hutchins, E., Cobb, M.M., Webb, M.G., Alsop, E., Grenn, F.P., Illarionova, A., Savytska, N., Violich, I., Broeer, S., Fernandes, N., Sivakumar, R., Beilina, A., Billingsley, K.J., Berghausen, J., Pantazis, C.B., Pitz, V., et al. (2023) 'The Foundational Data Initiative for Parkinson Disease: Enabling efficient translation from genetic maps to mechanism', *Cell Genomics*, 3(3), .

Bricker-Anthony, C. & Herzog, R.W. (2023) 'Distortion of journal impact factors in the era of paper mills', *Molecular Therapy*, 31(6), pp. 1503–1504.

Brigo, F., Matinella, A., Erro, R. & Tinazzi, M. (2014) '[^{123}I]FP-CIT SPECT (DaTSCAN) may be a useful tool to differentiate between Parkinson's disease and vascular or drug-

induced parkinsonisms: a meta-analysis', *European Journal of Neurology*, 21(11), pp. 1369-e90.

Brockmann, K., Apel, A., Schulte, C., Schneiderhan-Marra, N., Pont-Sunyer, C., Vilas, D., Ruiz-Martinez, J., Langkamp, M., Corvol, J.-C., Cormier, F., Knorpp, T., Joos, T.O., Gasser, T., Schüle, B., Aasly, J.O., Foroud, T., Marti-Masso, J.F., Brice, A., Tolosa, E., et al. (2016) 'Inflammatory profile in LRRK2-associated prodromal and clinical PD', *Journal of Neuroinflammation*, 13(1), p. 122.

Brockmann, K., Lerche, S., Baiardi, S., Rossi, M., Wurster, I., Quadalti, C., Roeben, B., Mammana, A., Zimmermann, M., Hauser, A.-K., Deuschle, C., Schulte, C., Liepelt-Scarfone, I., Gasser, T. & Parchi, P. (2024) 'CSF α-synuclein seed amplification kinetic profiles are associated with cognitive decline in Parkinson's disease', *npj Parkinson's Disease*, 10(1), pp. 1–9.

Brockmann, K., Quadalti, C., Lerche, S., Rossi, M., Wurster, I., Baiardi, S., Roeben, B., Mammana, A., Zimmermann, M., Hauser, A.-K., Deuschle, C., Schulte, C., Waniek, K., Lachmann, I., Sjödin, S., Brinkmalm, A., Blennow, K., Zetterberg, H., Gasser, T., et al. (2021) 'Association between CSF alpha-synuclein seeding activity and genetic status in Parkinson's disease and dementia with Lewy bodies', *Acta Neuropathologica Communications*, 9(1), p. 175.

Brockmann, K., Srulijes, K., Pflederer, S., Hauser, A.-K., Schulte, C., Maetzler, W., Gasser, T. & Berg, D. (2015) 'GBA-associated Parkinson's disease: Reduced survival and more rapid progression in a prospective longitudinal study', *Movement Disorders*, 30(3), pp. 407–411.

Brys, M., Fanning, L., Hung, S., Ellenbogen, A., Penner, N., Yang, M., Welch, M., Koenig, E., David, E., Fox, T., Makh, S., Aldred, J., Goodman, I., Pepinsky, B., Liu, Y., Graham, D., Weihofen, A. & Cedarbaum, J.M. (2019) 'Randomized phase I clinical trial of anti–α-synuclein antibody BIIB054', *Movement Disorders*, 34(8), pp. 1154–1163.

Burd, C.E., Jeck, W.R., Liu, Y., Sanoff, H.K., Wang, Z. & Sharpless, N.E. (2010) 'Expression of Linear and Novel Circular Forms of an INK4/ARF-Associated Non-Coding RNA Correlates with Atherosclerosis Risk', *PLOS Genetics*, 6(12), p. e1001233.

Burke, J.M., Moon, S.L., Matheny, T. & Parker, R. (2019) 'RNase L Reprograms Translation by Widespread mRNA Turnover Escaped by Antiviral mRNAs', *Molecular Cell*, 75(6), pp. 1203-1217.e5.

Burré, J. (2015) 'The Synaptic Function of α-Synuclein', *Journal of Parkinson's Disease*, 5(4), pp. 699–713.

Burté, F., Houghton, D., Lowes, H., Pyle, A., Nesbitt, S., Yarnall, A., Yu-Wai-Man, P., Burn, D.J., Santibanez-Koref, M. & Hudson, G. (2017) 'metabolic profiling of Parkinson's disease and mild cognitive impairment', *Movement Disorders*, 32(6), pp. 927–932.

Buter, T.C., van den Hout, A., Matthews, F.E., Larsen, J.P., Brayne, C. & Aarsland, D. (2008) 'Dementia and survival in Parkinson disease: a 12-year population study', *Neurology*, 70(13), pp. 1017–1022.

Califf, R.M. (2018) 'Biomarker definitions and their applications', *Experimental Biology and Medicine*, 243(3), pp. 213–221.

del Campo, M., Vermunt, L., Peeters, C.F.W., Sieben, A., Hok-A-Hin, Y.S., Lleó, A., Alcolea, D., van Nee, M., Engelborghs, S., van Alphen, J.L., Arezoumandan, S., Chen-Plotkin, A., Irwin, D.J., van der Flier, W.M., Lemstra, A.W. & Teunissen, C.E. (2023) 'CSF proteome profiling reveals biomarkers to discriminate dementia with Lewy bodies from Alzheimer´s disease', *Nature Communications*, 14(1), p. 5635.

Cao, S.-M., Wu, H., Yuan, G.-H., Pan, Y.-H., Zhang, J., Liu, Y.-X., Li, S., Xu, Y.-F., Wei, M.-Y., Yang, L. & Chen, L.-L. (2024) 'Altered nucleocytoplasmic export of adenosine-rich circRNAs by PABPC1 contributes to neuronal function', *Molecular Cell*,

Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., Goodfellow, P. & Lovell-Badge, R. (1993) 'Circular transcripts of the testis-determining gene Sry in adult mouse testis', *Cell*, 73(5), pp. 1019–1030.

Caspell-Garcia, C., Simuni, T., Tosun-Turgut, D., Wu, I.-W., Zhang, Y., Nalls, M., Singleton, A., Shaw, L.A., Kang, J.-H., Trojanowski, J.Q., Siderowf, A., Coffey, C., Lasch, S., Aarsland, D., Burn, D., Chahine, L.M., Espay, A.J., Foster, E.D., Hawkins, K.A., et al. (2017) 'Multiple modality biomarker prediction of cognitive impairment in prospectively followed de novo Parkinson disease', *PLOS ONE*, 12(5), p. e0175674.

Castro-Gomez, S. & Heneka, M.T. (2024) 'Innate immune activation in neurodegenerative diseases', *Immunity*, 57(4), pp. 790–814.

Cerri, S., Mus, L. & Blandini, F. (2019) 'Parkinson's Disease in Women and Men: What's the Difference?', *Journal of Parkinson's Disease*, 9(3), pp. 501–515.

Chahine, L.M., Qiang, J., Ashbridge, E., Minger, J., Yearout, D., Horn, S., Colcher, A., Hurtig, H.I., Lee, V.M.-Y., Van Deerlin, V.M., Leverenz, J.B., Siderowf, A.D., Trojanowski, J.Q., Zabetian, C.P. & Chen-Plotkin, A. (2013) 'Clinical and Biochemical Differences in Patients Having Parkinson Disease With vs Without GBA Mutations', *JAMA Neurology*, 70(7), pp. 852–858.

Chakrabarti, A., Jha, B.K. & Silverman, R.H. (2011) 'New Insights into the Role of RNase L in Innate Immunity', *Journal of Interferon & Cytokine Research*, 31(1), pp. 49–57.

Chang, D., Nalls, M.A., Hallgrímsdóttir, I.B., Hunkapiller, J., van der Brug, M., Cai, F., International Parkinson's Disease Genomics Consortium, 23andMe Research Team, Kerchner, G.A., Ayalon, G., Bingol, B., Sheng, M., Hinds, D., Behrens, T.W., Singleton, A.B., Bhangale, T.R. & Graham, R.R. (2017) 'A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci', *Nature Genetics*, 49(10), pp. 1511–1516.

Chang, H., Wang, B., Shi, Y. & Zhu, R. (2022) 'Dose-response meta-analysis on urate, gout, and the risk for Parkinson's disease', *npj Parkinson's Disease*, 8(1), pp. 1–9.

Chang, R.C.C., Wong, A.K.Y., Ng, H.-K. & Hugon, J. (2002) 'Phosphorylation of eukaryotic initiation factor-2α (eIF2α) is associated with neuronal degeneration in Alzheimer's disease', *NeuroReport*, 13(18), p. 2429.

Chen, B.J., Mills, J.D., Takenaka, K., Bliim, N., Halliday, G.M. & Janitz, M. (2016) 'Characterization of circular RNAs landscape in multiple system atrophy brain', *Journal of Neurochemistry*, 139(3), pp. 485–496.

Chen, C.-K., Cheng, R., Demeter, J., Chen, J., Weingarten-Gabbay, S., Jiang, L., Snyder, M.P., Weissman, J.S., Segal, E., Jackson, P.K. & Chang, H.Y. (2021) 'Structured elements drive extensive circular RNA translation', *Molecular Cell*,

Chen, C.X., Cho, D.S., Wang, Q., Lai, F., Carter, K.C. & Nishikura, K. (2000) 'A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains.', *RNA*, 6(5), pp. 755–767.

Chen, C.Y. & Sarnow, P. (1995) 'Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs', *Science (New York, N.Y.)*, 268(5209), pp. 415–417.

Chen, L., Wang, C., Sun, H., Wang, J., Liang, Y., Wang, Y. & Wong, G. (2021) 'The bioinformatics toolbox for circRNA discovery and analysis', *Briefings in Bioinformatics*, 22(2), pp. 1706–1728.

Chen, L., Wang, Yucong, Lin, J., Song, Z., Wang, Q., Zhao, W., Wang, Yan, Xiu, X., Deng, Y., Li, X., Li, Q., Wang, X., Li, J., Liu, X., Liu, K., Zhou, J., Li, K., Liu, Y., Liao, S., et al. (2022) 'Exportin 4 depletion leads to nuclear accumulation of a subset of circular RNAs', *Nature Communications*, 13(1), p. 5769.

Chen, L.-L., Bindereif, A., Bozzoni, I., Chang, H.Y., Matera, A.G., Gorospe, M., Hansen, T.B., Kjems, J., Ma, X.-K., Pek, J.W., Rajewsky, N., Salzman, J., Wilusz, J.E., Yang, L. & Zhao, F. (2023) 'A guide to naming eukaryotic circular RNAs', *Nature Cell Biology*, 25(1), pp. 1–5.

Chen, R., Wang, S.K., Belk, J.A., Amaya, L., Li, Z., Cardenas, A., Abe, B.T., Chen, C.-K., Wender, P.A. & Chang, H.Y. (2023) 'Engineering circular RNA for enhanced protein production', *Nature Biotechnology*, 41(2), pp. 262–272.

Chen, Y.G., Chen, R., Ahmad, S., Verma, R., Kasturi, S.P., Amaya, L., Broughton, J.P., Kim, J., Cadena, C., Pulendran, B., Hur, S. & Chang, H.Y. (2019) 'N6-Methyladenosine Modification Controls Circular RNA Immunity', *Molecular Cell*, 76(1), pp. 96-109.e9.

Chen, Y.G., Kim, M.V., Chen, X., Batista, P.J., Aoyama, S., Wilusz, J.E., Iwasaki, A. & Chang, H.Y. (2017) 'Sensing Self and Foreign Circular RNAs by Intron Identity', *Molecular Cell*, 67(2), pp. 228-238.e5.

Cheng, J., Metge, F. & Dieterich, C. (2016) 'Specific identification and quantification of circular RNAs from sequencing data', *Bioinformatics*, 32(7), pp. 1094–1096.

Cho, S.J., Bae, Y.J., Kim, J.-M., Kim, D., Baik, S.H., Sunwoo, L., Choi, B.S. & Kim, J.H. (2021) 'Diagnostic performance of neuromelanin-sensitive magnetic resonance imaging for patients with Parkinson's disease and factor analysis for its heterogeneity: a systematic review and meta-analysis', *European Radiology*, 31(3), pp. 1268–1280.

Chou, C.-C., Zhang, Y., Umoh, M.E., Vaughan, S.W., Lorenzini, I., Liu, F., Sayegh, M., Donlin-Asp, P.G., Chen, Y.H., Duong, D.M., Seyfried, N.T., Powers, M.A., Kukar, T., Hales, C.M., Gearing, M., Cairns, N.J., Boylan, K.B., Dickson, D.W., Rademakers, R., et al. (2018) 'TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD', *Nature Neuroscience*, 21(2), pp. 228–239.

Chuang, T.-J., Wu, C.-S., Chen, C.-Y., Hung, L.-Y., Chiang, T.-W. & Yang, M.-Y. (2016) 'NCLscan: accurate identification of non-co-linear transcripts (fusion, trans-splicing

and circular RNA) with a good balance between sensitivity and precision', *Nucleic Acids Research*, 44(3), p. e29.

Chung, C.Y., Koprich, J.B., Hallett, P.J. & Isacson, O. (2009) 'Functional enhancement and protection of dopaminergic terminals by RAB3B overexpression', *Proceedings of the National Academy of Sciences*, 106(52), pp. 22474–22479.

Chung, S.J., Armasu, S.M., Biernacka, J.M., Anderson, K.J., Lesnick, T.G., Rider, D.N., Cunningham, J.M., Eric Ahlskog, J., Frigerio, R. & Maraganore, D.M. (2012) 'Genomic determinants of motor and cognitive outcomes in Parkinson's disease', *Parkinsonism & Related Disorders*, 18(7), pp. 881–886.

Cicero, C.E., Luca, A., Mostile, G., Donzuso, G., Giuliano, L., Zappia, M. & Nicoletti, A. (2023) 'Influence of RBD onset on the clinical characteristics of Parkinson's disease patients: a retrospective study', *Journal of Neurology*, 270(6), pp. 3171–3178.

Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A. & Kędziora-Kornatowska, K. (2016) 'Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis', *Psychiatria Polska*, 50(5), pp. 1039–1052.

Cilia, R., Marotta, G., Benti, R., Pezzoli, G. & Antonini, A. (2005) 'Brain SPECT imaging in multiple system atrophy', *Journal of Neural Transmission*, 112(12), pp. 1635–1645.

Cilia, R., Tunesi, S., Marotta, G., Cereda, E., Siri, C., Tesei, S., Zecchinelli, A.L., Canesi, M., Mariani, C.B., Meucci, N., Sacilotto, G., Zini, M., Barichella, M., Magnani, C., Duga, S., Asselta, R., Soldà, G., Seresini, A., Seia, M., et al. (2016) 'Survival and dementia in GBA-associated Parkinson's disease: The mutation matters', *Annals of Neurology*, 80(5), pp. 662–673.

Civiero, L., Cogo, S., Biosa, A. & Greggio, E. (2018) 'The role of LRRK2 in cytoskeletal dynamics', *Biochemical Society Transactions*, 46(6), pp. 1653–1663.

Cocoros, N.M., Svensson, E., Szépligeti, S.K., Vestergaard, S.V., Szentkúti, P., Thomsen, R.W., Borghammer, P., Sørensen, H.T. & Henderson, V.W. (2021) 'Long-term Risk of Parkinson Disease Following Influenza and Other Infections', *JAMA Neurology*, 78(12), pp. 1461–1470.

Cocquerelle, C., Daubersies, P., Majérus, M.A., Kerckaert, J.P. & Bailleul, B. (1992) 'Splicing with inverted order of exons occurs proximal to large introns.', *The EMBO Journal*, 11(3), pp. 1095–1098.

Cocquerelle, C., Mascrez, B., Hétuin, D. & Bailleul, B. (1993) 'Mis-splicing yields circular RNA molecules', *The FASEB Journal*, 7(1), pp. 155–160.

Coffelt, S.B., Wellenstein, M.D. & de Visser, K.E. (2016) 'Neutrophils in cancer: neutral no more', *Nature Reviews Cancer*, 16(7), pp. 431–446.

Cole-Hunter, T., Zhang, J., So, R., Samoli, E., Liu, S., Chen, J., Strak, M., Wolf, K., Weinmayr, G., Rodopolou, S., Remfry, E., de Hoogh, K., Bellander, T., Brandt, J., Concin, H., Zitt, E., Fecht, D., Forastiere, F., Gulliver, J., et al. (2023) 'Long-term air pollution exposure

and Parkinson's disease mortality in a large pooled European cohort: An ELAPSE study', *Environment International*, 171p. 107667.

Comi, C., Magistrelli, L., Oggioni, G.D., Carecchio, M., Fleetwood, T., Cantello, R., Mancini, F. & Antonini, A. (2014) 'Peripheral nervous system involvement in Parkinson's disease: Evidence and controversies', *Parkinsonism & Related Disorders*, 20(12), pp. 1329–1334.

Concha-Marambio, L., Farris, C.M., Holguin, B., Ma, Y., Seibyl, J., Russo, M.J., Kang, U.J., Hutten, S.J., Merchant, K., Shahnawaz, M. & Soto, C. (2021) 'Seed Amplification Assay to Diagnose Early Parkinson's and Predict Dopaminergic Deficit Progression', *Movement Disorders*, 36(10), pp. 2444–2446.

Concha-Marambio, L., Weber, S., Farris, C.M., Dakna, M., Lang, E., Wicke, T., Ma, Y., Starke, M., Ebentheuer, J., Sixel-Döring, F., Muntean, M.-L., Schade, S., Trenkwalder, C., Soto, C. & Mollenhauer, B. (2023) 'Accurate Detection of α-Synuclein Seeds in Cerebrospinal Fluid from Isolated Rapid Eye Movement Sleep Behavior Disorder and Patients with Parkinson's Disease in the DeNovo Parkinson (DeNoPa) Cohort', *Movement Disorders*, 38(4), pp. 567–578.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X. & Mortazavi, A. (2016) 'A survey of best practices for RNA-seq data analysis', *Genome Biology*, 17(1), p. 13.

Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A. & Goodall, G.J. (2015) 'The RNA Binding Protein Quaking Regulates Formation of circRNAs', *Cell*, 160(6), pp. 1125–1134.

Conn, V.M., Hugouvieux, V., Nayak, A., Conos, S.A., Capovilla, G., Cildir, G., Jourdain, A., Tergaonkar, V., Schmid, M., Zubieta, C. & Conn, S.J. (2017) 'A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation', *Nature Plants*, 3(5), pp. 1–5.

Connolly, B.S. & Lang, A.E. (2014) 'Pharmacological Treatment of Parkinson Disease: A Review', *JAMA*, 311(16), pp. 1670–1683.

Consortium, T.E.P. (2011) 'A User's Guide to the Encyclopedia of DNA Elements (ENCODE)', *PLOS Biology*, 9(4), p. e1001046.

Cook, L., Verbrugge, J., Schwantes-An, T.-H., Schulze, J., Beck, J.C., Naito, A., Hall, A., Chan, A.K., Casaceli, C.J., Marder, K., Nance, M., Schwarzschild, M.A., Simuni, T., Wills, A.-M. & Alcalay, R.N. (2023) 'Providing genetic testing and genetic counseling for Parkinson's disease to the community', *Genetics in Medicine*, 25(10), p. 100907.

Costa, J., Lunet, N., Santos, C., Santos, J. & Vaz-Carneiro, A. (2010) 'Caffeine Exposure and the Risk of Parkinson's Disease: A Systematic Review and Meta-Analysis of Observational Studiess', *Journal of Alzheimer's Disease*, 20(s1), pp. S221–S238.

Costa-Mattioli, M. & Walter, P. (2020) 'The integrated stress response: From mechanism to disease', *Science*, 368(6489), p. eaat5314.

Coukos, R. & Krainc, D. (2024) 'Key genes and convergent pathogenic mechanisms in Parkinson disease', *Nature Reviews Neuroscience*, pp. 1–21.

Craig, D.W., Hutchins, E., Violich, I., Alsop, E., Gibbs, J.R., Levy, S., Robison, M., Prasad, N., Foroud, T., Crawford, K.L., Toga, A.W., Whitsett, T.G., Kim, S., Casey, B., Reimer, A., Hutten, S.J., Frasier, M., Kern, F., Fehlman, T., et al. (2021) 'RNA sequencing of whole blood reveals early alterations in immune cells and gene expression in Parkinson's disease', *Nature Aging*, 1(8), pp. 734–747.

Crane, P.K., Gibbons, L.E., Dams-O'Connor, K., Trittschuh, E., Leverenz, J.B., Keene, C.D., Sonnen, J., Montine, T.J., Bennett, D.A., Leurgans, S., Schneider, J.A. & Larson, E.B. (2016) 'Association of Traumatic Brain Injury With Late-Life Neurodegenerative Conditions and Neuropathologic Findings', *JAMA Neurology*, 73(9), pp. 1062–1069.

Cummings, J.L., Henchcliffe, C., Schaier, S., Simuni, T., Waxman, A. & Kemp, P. (2011) 'The role of dopaminergic imaging in patients with symptoms of dopaminergic system neurodegeneration', *Brain*, 134(11), pp. 3146–3166.

D'Ambrosi, S., Giannoukakos, S., Antunes-Ferreira, M., Pedraz-Valdunciel, C., Bracht, J.W.P., Potie, N., Gimenez-Capitan, A., Hackenberg, M., Fernandez Hilario, A., Molina-Vila, M.A., Rosell, R., Würdinger, T. & Koppers-Lalic, D. (2023) 'Combinatorial Blood Platelets-Derived circRNA and mRNA Signature for Early-Stage Lung Cancer Detection', *International Journal of Molecular Sciences*, 24(5), p. 4881.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. & Li, H. (2021) 'Twelve years of SAMtools and BCFtools', *GigaScience*, 10(2), p. giab008.

Daniele, S., Frosini, D., Pietrobono, D., Petrozzi, L., Lo Gerfo, A., Baldacci, F., Fusi, J., Giacomelli, C., Siciliano, G., Trincavelli, M.L., Franzoni, F., Ceravolo, R., Martini, C. & Bonuccelli, U. (2018) 'α-Synuclein Heterocomplexes with β-Amyloid Are Increased in Red Blood Cells of Parkinson's Disease Patients and Correlate with Disease Severity', *Frontiers in Molecular Neuroscience*, 11.

Dauvilliers, Y., Schenck, C.H., Postuma, R.B., Iranzo, A., Luppi, P.-H., Plazzi, G., Montplaisir, J. & Boeve, B. (2018) 'REM sleep behaviour disorder', *Nature Reviews Disease Primers*, 4(1), pp. 1–16.

De Feo, M.S., Frantellizzi, V., Locuratolo, N., Di Rocco, A., Farcomeni, A., Pauletti, C., Marongiu, A., Lazri, J., Nuvoli, S., Fattapposta, F., De Vincentis, G. & Spanu, A. (2023) 'Role of Functional Neuroimaging with 123I-MIBG and 123I-FP-CIT in De Novo Parkinson's Disease: A Multicenter Study', *Life*, 13(8), p. 1786.

De Luca, C.M.G., Elia, A.E., Portaleone, S.M., Cazzaniga, F.A., Rossi, M., Bistaffa, E., De Cecco, E., Narkiewicz, J., Salzano, G., Carletta, O., Romito, L., Devigili, G., Soliveri, P., Tiraboschi, P., Legname, G., Tagliavini, F., Eleopra, R., Giaccone, G. & Moda, F. (2019) 'Efficient RT-QuIC seeding activity for α-synuclein in olfactory mucosa samples of patients with Parkinson's disease and multiple system atrophy', *Translational Neurodegeneration*, 8p. 24.

Debès, C., Papadakis, A., Grönke, S., Karalay, Ö., Tain, L.S., Mizi, A., Nakamura, S., Hahn, O., Weigelt, C., Josipovic, N., Zirkel, A., Brusius, I., Sofiadis, K., Lamprousi, M., Lu, Y.-X., Huang, W., Esmaillie, R., Kubacki, T., Späth, M.R., et al. (2023) 'Ageing-associated changes in transcriptional elongation influence longevity', *Nature*, 616(7958), pp. 814–821.

Dechat, T., Pfleghaar, K., Sengupta, K., Shimi, T., Shumaker, D.K., Solimando, L. & Goldman, R.D. (2008) 'Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin', *Genes & Development*, 22(7), pp. 832–853.

DelleDonne, A., Klos, K.J., Fujishiro, H., Ahmed, Z., Parisi, J.E., Josephs, K.A., Frigerio, R., Burnett, M., Wszolek, Z.K., Uitti, R.J., Ahlskog, J.E. & Dickson, D.W. (2008) 'Incidental Lewy Body Disease and Preclinical Parkinson Disease', *Archives of Neurology*, 65(8), pp. 1074–1080.

DeLong, E.R., DeLong, D.M. & Clarke-Pearson, D.L. (1988) 'Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach', *Biometrics*, 44(3), pp. 837–845.

Deng, H.-X., Shi, Y., Yang, Y., Ahmeti, K.B., Miller, N., Huang, C., Cheng, L., Zhai, H., Deng, S., Nuytemans, K., Corbett, N.J., Kim, M.J., Deng, H., Tang, B., Yang, Z., Xu, Y., Chan, P., Huang, B., Gao, X.-P., et al. (2016) 'Identification of TMEM230 mutations in familial Parkinson's disease', *Nature Genetics*, 48(7), pp. 733–739.

Denzler, R., Agarwal, V., Stefano, J., Bartel, D.P. & Stoffel, M. (2014) 'Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA and Target Abundance', *Molecular Cell*, 54(5), pp. 766–776.

Depierreux, F., Parmentier, E., Mackels, L., Baquero, K., Degueldre, C., Balteau, E., Salmon, E., Phillips, C., Bahri, M.A., Maquet, P. & Garraux, G. (2021) 'Parkinson's disease multimodal imaging: F-DOPA PET, neuromelanin-sensitive and quantitative iron-sensitive MRI', *npj Parkinson's Disease*, 7(1), pp. 1–10.

Deshpande, P., Flinkman, D., Hong, Y., Goltseva, E., Siino, V., Sun, L., Peltonen, S., Elo, L.L., Kaasinen, V., James, P. & Coffey, E.T. (2020) 'Protein synthesis is suppressed in sporadic and familial Parkinson's disease by LRRK2', *The FASEB Journal*, 34(11), pp. 14217–14233.

Devic, I., Hwang, H., Edgar, J.S., Izutsu, K., Presland, R., Pan, C., Goodlett, D.R., Wang, Y., Armaly, J., Tumas, V., Zabetian, C.P., Leverenz, J.B., Shi, M. & Zhang, J. (2011) 'Salivary α-synuclein and DJ-1: potential biomarkers for Parkinson's disease', *Brain*, 134(7), p. e178.

Dey, M., Mann, B.R., Anshu, A. & Mannan, M.A. (2014) 'Activation of Protein Kinase PKR Requires Dimerization-induced cis-Phosphorylation within the Activation Loop', *The Journal of Biological Chemistry*, 289(9), pp. 5747–5757.

Dezsi, L. & Vecsei, L. (2017) 'Monoamine Oxidase B Inhibitors in Parkinson's Disease', *CNS & neurological disorders drug targets*, 16(4), pp. 425–439.

Dhanwani, R., Lima-Junior, J.R., Sethi, A., Pham, J., Williams, G., Frazier, A., Xu, Y., Amara, A.W., Standaert, D.G., Goldman, J.G., Litvan, I., Alcalay, R.N., Peters, B., Sulzer, D., Arlehamn, C.S.L. & Sette, A. (2022) 'Transcriptional analysis of peripheral memory T cells reveals Parkinson's disease-specific gene signatures', *NPJ Parkinson's disease*, 8(1), p. 30.

Dhir, A., Dhir, S., Borowski, L.S., Jimenez, L., Teitell, M., Rötig, A., Crow, Y.J., Rice, G.I., Duffy, D., Tamby, C., Nojima, T., Munnich, A., Schiff, M., de Almeida, C.R., Rehwinkel, J., Dziembowski, A., Szczesny, R.J. & Proudfoot, N.J. (2018)

'Mitochondrial double-stranded RNA triggers antiviral signalling in humans', *Nature*, 560(7717), pp. 238–242.

Di Liddo, A., de Oliveira Freitas Machado, C., Fischer, S., Ebersberger, S., Heumüller, A.W., Weigand, J.E., Müller-McNicoll, M. & Zarnack, K. (2019) 'A combined computational pipeline to detect circular RNAs in human cancer cells under hypoxic stress', *Journal of Molecular Cell Biology*, 11(10), pp. 829–844.

Dickson, D.W., Fujishiro, H., DelleDonne, A., Menke, J., Ahmed, Z., Klos, K.J., Josephs, K.A., Frigerio, R., Burnett, M., Parisi, J.E. & Ahlskog, J.E. (2008) 'Evidence that incidental Lewy body disease is pre-symptomatic Parkinson's disease', *Acta Neuropathologica*, 115(4), pp. 437–444.

Diederich, C., Milakofsky, L., Hare, T.A., Hofford, J.M., Dadmarz, M. & Vogel, W.H. (2008) 'Effects of L-DOPA/Carbidopa Administration on the Levels of L-DOPA, Other Amino Acids and Related Compounds in the Plasma, Brain and Heart of the Rat', *Pharmacology*, 55(3), pp. 109–116.

Dixon, R.J., Eperon, I.C., Hall, L. & Samani, N.J. (2005) 'A genome-wide survey demonstrates widespread non-linear mRNA in expressed sequences from multiple species', *Nucleic Acids Research*, 33(18), pp. 5904–5913.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T.R. (2013) 'STAR: ultrafast universal RNA-seq aligner', *Bioinformatics (Oxford, England)*, 29(1), pp. 15–21.

Dodbele, S., Mutlu, N. & Wilusz, J.E. (2021) 'Best practices to ensure robust investigation of circular RNAs: pitfalls and tips', *EMBO reports*, p. e52072.

Dodel, R.C., Höffken, H., Möller, J.C., Bornschein, B., Klockgether, T., Behr, T., Oertel, W.H. & Siebert, U. (2003) 'Dopamine transporter imaging and SPECT in diagnostic work-up of Parkinson's disease: a decision-analytic approach', *Movement Disorders: Official Journal of the Movement Disorder Society*, 18 Suppl 7pp. S52-62.

Donadio, V., Wang, Z., Incensi, A., Rizzo, G., Fileccia, E., Vacchiano, V., Capellari, S., Magnani, M., Scaglione, C., Stanzani Maserati, M., Avoni, P., Liguori, R. & Zou, W. (2021) 'In Vivo Diagnosis of Synucleinopathies', *Neurology*, 96(20), pp. e2513–e2524.

Dong, R., Ma, X.-K., Chen, L.-L. & Yang, L. (2016) 'Increased complexity of circRNA expression during species evolution', *RNA Biology*, 14(8), pp. 1064–1074.

Dong, R., Ma, X.-K., Li, G.-W. & Yang, L. (2018) 'CIRCpedia v2: An Updated Database for Comprehensive Circular RNA Annotation and Expression Comparison', *Genomics, Proteomics & Bioinformatics*, 16(4), pp. 226–233.

Dong, X., Bai, Y., Liao, Z., Gritsch, D., Liu, X., Wang, T., Borges-Monroy, R., Ehrlich, A., Serrano, G.E., Feany, M.B., Beach, T.G. & Scherzer, C.R. (2023) 'Circular RNAs in the human brain are tailored to neuron identity and neuropsychiatric disease', *Nature Communications*, 14(1), p. 5327.

Dong, X., Liao, Z., Gritsch, D., Hadzhiev, Y., Bai, Y., Locascio, J.J., Guennewig, B., Liu, G., Blauwendraat, C., Wang, T., Adler, C.H., Hedreen, J.C., Faull, R.L.M., Frosch, M.P., Nelson, P.T., Rizzu, P., Cooper, A.A., Heutink, P., Beach, T.G., et al. (2018) 'Enhancers

active in dopamine neurons are a primary link between genetic variation and neuropsychiatric disease', *Nature Neuroscience*, 21(10), pp. 1482–1492.

Donnelly, N., Gorman, A.M., Gupta, S. & Samali, A. (2013) 'The eIF2α kinases: their structures and functions', *Cellular and Molecular Life Sciences*, 70(19), pp. 3493–3511.

Donovan, J., Rath, S., Kolet-Mandrikov, D. & Korennykh, A. (2017) 'Rapid RNase L–driven arrest of protein synthesis in the dsRNA response without degradation of translation machinery', *RNA*, 23(11), pp. 1660–1671.

Dorrity, T.J., Shin, H., Wiegand, K.A., Aruda, J., Closser, M., Jung, E., Gertie, J.A., Leone, A., Polfer, R., Culbertson, B., Yu, L., Wu, C., Ito, T., Huang, Y., Steckelberg, A.-L., Wichterle, H. & Chung, H. (2023) 'Long 3′UTRs predispose neurons to inflammation by promoting immunostimulatory double-stranded RNA formation', *Science Immunology*, 8(88), p. eadg2979.

Dorsey, E.R., Elbaz, A., Nichols, E., Abd-Allah, F., Abdelalim, A., Adsuar, J.C., Ansha, M.G., Brayne, C., Choi, J.-Y.J., Collado-Mateo, D., Dahodwala, N., Do, H.P., Edessa, D., Endres, M., Fereshtehnejad, S.-M., Foreman, K.J., Gankpe, F.G., Gupta, R., Hankey, G.J., et al. (2018) 'Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016', *The Lancet Neurology*, 17(11), pp. 939–953.

Dorsey, E.R., Sherer, T., Okun, M.S. & Bloem, B.R. (2018) 'The Emerging Evidence of the Parkinson Pandemic', *Journal of Parkinson's Disease*, 8(Suppl 1), pp. S3–S8.

Doty, R.L., Deems, D.A. & Stellar, S. (1988) 'Olfactory dysfunction in parkinsonism', *Neurology*, 38(8), pp. 1237–1237.

Doty, R.L., Shaman, P., Kimmelman, C.P. & Dann, M.S. (1984) 'University of Pennsylvania Smell Identification Test: a rapid quantitative olfactory function test for the clinic', *The Laryngoscope*, 94(2 Pt 1), pp. 176–178.

Doxakis, E. (2022) 'Insights into the multifaceted role of circular RNAs: implications for Parkinson's disease pathogenesis and diagnosis', *NPJ Parkinson's disease*, 8(1), p. 7.

Drew, K., Wallingford, J.B. & Marcotte, E.M. (2021) 'hu.MAP 2.0: integration of over 15,000 proteomic experiments builds a global compendium of human multiprotein assemblies', *Molecular Systems Biology*, 17(5), p. e10016.

Droby, A., Artzi, M., Lerman, H., Hutchison, R.M., Bashat, D.B., Omer, N., Gurevich, T., Orr-Urtreger, A., Cohen, B., Cedarbaum, J.M., Sapir, E.E., Giladi, N., Mirelman, A. & Thaler, A. (2022) 'Aberrant dopamine transporter and functional connectivity patterns in LRRK2 and GBA mutation carriers', *npj Parkinson's Disease*, 8(1), pp. 1–7.

D'Sa, K., Choi, M.L., Wagen, A.Z., Setó-Salvia, N., Kopach, O., Evans, J.R., Rodrigues, M., Lopez-Garcia, P., Ghareeb, A., Bayne, J., Grant-Peters, M., Garcia-Ruiz, S., Chen, Z., Rodriques, S., Athauda, D., Gustavsson, E., Taliun, S.A.G., Reynolds, R.H., Young, G., et al. (2024) *Alpha-synuclein aggregates trigger anti-viral immune pathways and RNA editing in human astrocytes*. p.p. 2024.02.26.582055.

Dubin, R.A., Kazmi, M.A. & Ostrer, H. (1995) 'Inverted repeats are necessary for circularization of the mouse testis Sry transcript', *Gene*, 167(1), pp. 245–248.

Dzamko, N., Rowe, D.B. & Halliday, G.M. (2016) 'Increased peripheral inflammation in asymptomatic leucine-rich repeat kinase 2 mutation carriers', *Movement Disorders*, 31(6), pp. 889–897.

Eidson, L.N., Kannarkat, G.T., Barnum, C.J., Chang, J., Chung, J., Caspell-Garcia, C., Taylor, P., Mollenhauer, B., Schlossmacher, M.G., Ereshefsky, L., Yen, M., Kopil, C., Frasier, M., Marek, K., Hertzberg, V.S. & Tansey, M.G. (2017) 'Candidate inflammatory biomarkers display unique relationships with alpha-synuclein and correlate with measures of disease severity in subjects with Parkinson's disease', *Journal of Neuroinflammation*, 14(1), p. 164.

Ejlerskov, P., Hultberg, J.G., Wang, J., Carlsson, R., Ambjørn, M., Kuss, M., Liu, Y., Porcu, G., Kolkova, K., Friis Rundsten, C., Ruscher, K., Pakkenberg, B., Goldmann, T., Loreth, D., Prinz, M., Rubinsztein, D.C. & Issazadeh-Navikas, S. (2015) 'Lack of Neuronal IFN-β-IFNAR Causes Lewy Body- and Parkinson's Disease-like Dementia', *Cell*, 163(2), pp. 324–339.

Enuka, Y., Lauriola, M., Feldman, M.E., Sas-Chen, A., Ulitsky, I. & Yarden, Y. (2016) 'Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor', *Nucleic Acids Research*, 44(3), pp. 1370–1383.

Erb, M.L. & Moore, D.J. (2020) 'LRRK2 and the Endolysosomal System in Parkinson's Disease', *Journal of Parkinson's Disease*, 10(4), pp. 1271–1291.

Errichelli, L., Dini Modigliani, S., Laneve, P., Colantoni, A., Legnini, I., Capauto, D., Rosa, A., De Santis, R., Scarfò, R., Peruzzi, G., Lu, L., Caffarelli, E., Shneider, N.A., Morlando, M. & Bozzoni, I. (2017) 'FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons', *Nature Communications*, 8(1), p. 14741.

Espinosa-Oliva, A.M., Ruiz, R., Soto, M.S., Boza-Serrano, A., Rodriguez-Perez, A.I., Roca-Ceballos, M.A., García-Revilla, J., Santiago, M., Serres, S., Economopoulus, V., Carvajal, A.E., Vázquez-Carretero, M.D., García-Miranda, P., Klementieva, O., Oliva-Martín, M.J., Deierborg, T., Rivas, E., Sibson, N.R., Labandeira-García, J.L., et al. (2024) 'Inflammatory bowel disease induces pathological α-synuclein aggregation in the human gut and brain', *Neuropathology and Applied Neurobiology*, 50(1), p. e12962.

Eusebi, P., Giannandrea, D., Biscetti, L., Abraha, I., Chiasserini, D., Orso, M., Calabresi, P. & Parnetti, L. (2017) 'Diagnostic utility of cerebrospinal fluid α-synuclein in Parkinson's disease: A systematic review and meta-analysis', *Movement Disorders*, 32(10), pp. 1389–1400.

Evans, C., Hardin, J. & Stoebel, D.M. (2017) 'Selecting between-sample RNA-Seq normalization methods from the perspective of their assumptions', *Briefings in Bioinformatics*, 19(5), pp. 776–792.

Evans, R.W. (1998) 'COMPLICATIONS OF LUMBAR PUNCTURE', *Neurologic Clinics*, 16(1), pp. 83–105.

Ewels, P., Magnusson, M., Lundin, S. & Käller, M. (2016) 'MultiQC: summarize analysis results for multiple tools and samples in a single report', *Bioinformatics*, 32(19), pp. 3047–3048.

Fairfoul, G., McGuire, L.I., Pal, S., Ironside, J.W., Neumann, J., Christie, S., Joachim, C., Esiri, M., Evetts, S.G., Rolinski, M., Baig, F., Ruffmann, C., Wade-Martins, R., Hu, M.T.M., Parkkinen, L. & Green, A.J.E. (2016) 'Alpha-synuclein RT-QuIC in the CSF of patients with alpha-synucleinopathies', *Annals of Clinical and Translational Neurology*, 3(10), pp. 812–818.

Fan, X., Zhang, X., Wu, X., Guo, H., Hu, Y., Tang, F. & Huang, Y. (2015) 'Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos', *Genome Biology*, 16(1), p. 148.

Fearnley, J.M. & Lees, A.J. (1991) 'AGEING AND PARKINSON'S DISEASE: SUBSTANTIA NIGRA REGIONAL SELECTIVITY', *Brain*, 114(5), pp. 2283–2301.

Feng, X., Jiang, B.-W., Zhai, S.-N., Liu, C.-X., Wu, H., Zhu, B.-Q., Wei, M.-Y., Wei, J., Yang, L. & Chen, L.-L. (2024) *Circular RNA aptamers ameliorate AD-relevant phenotypes by targeting PKR*. p.p. 2024.03.27.583257.

Feng, Z., Zhang, L., Wang, S. & Hong, Q. (2020) 'Circular RNA circDLGAP4 exerts neuroprotective effects via modulating miR-134-5p/CREB pathway in Parkinson's disease', *Biochemical and Biophysical Research Communications*, 522(2), pp. 388–394.

Fereshtehnejad, S.-M., Montplaisir, J.Y., Pelletier, A., Gagnon, J.-F., Berg, D. & Postuma, R.B. (2017) 'Validation of the MDS research criteria for prodromal Parkinson's disease: Longitudinal assessment in a REM sleep behavior disorder (RBD) cohort', *Movement Disorders*, 32(6), pp. 865–873.

Fereshtehnejad, S.-M., Zeighami, Y., Dagher, A. & Postuma, R.B. (2017) 'Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression', *Brain*, 140(7), pp. 1959–1976.

Ferreira, P.G., Muñoz-Aguirre, M., Reverter, F., Sá Godinho, C.P., Sousa, A., Amadoz, A., Sodaei, R., Hidalgo, M.R., Pervouchine, D., Carbonell-Caballero, J., Nurtdinov, R., Breschi, A., Amador, R., Oliveira, P., Çubuk, C., Curado, J., Aguet, F., Oliveira, C., Dopazo, J., et al. (2018) 'The effects of death and post-mortem cold ischemia on human tissue transcriptomes', *Nature Communications*, 9(1), p. 490.

Fischer, J.W., Busa, V.F., Shao, Y. & Leung, A.K.L. (2020) 'Structure-Mediated RNA Decay by UPF1 and G3BP1', *Molecular Cell*, 78(1), pp. 70-84.e6.

Fjorback, A.W., Varming, K. & Jensen, P.H. (2007) 'Determination of alpha-synuclein concentration in human plasma using ELISA', *Scandinavian Journal of Clinical and Laboratory Investigation*, 67(4), pp. 431–435.

Flinkman, D., Hong, Y., Gnjatovic, J., Deshpande, P., Ortutay, Z., Peltonen, S., Kaasinen, V., James, P. & Coffey, E. (2023) 'Regulators of proteostasis are translationally repressed in fibroblasts from patients with sporadic and LRRK2-G2019S Parkinson's disease', *npj Parkinson's Disease*, 9(1), pp. 1–13.

Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X.-D. & Bentley, D.L. (2014) 'Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate', *Genes & Development*, 28(23), pp. 2663–2676.

Foo, J.N., Chew, E.G.Y., Chung, S.J., Peng, R., Blauwendraat, C., Nalls, M.A., Mok, K.Y., Satake, W., Toda, T., Chao, Y., Tan, L.C.S., Tandiono, M., Lian, M.M., Ng, E.Y., Prakash, K.-M., Au, W.-L., Meah, W.-Y., Mok, S.Q., Annuar, A.A., et al. (2020) 'Identification of Risk Loci for Parkinson Disease in Asians and Comparison of Risk Between Asians and Europeans: A Genome-Wide Association Study', *JAMA Neurology*, 77(6), pp. 746–754.

Freuer, D. & Meisinger, C. (2022) 'Association between inflammatory bowel disease and Parkinson's disease: A Mendelian randomization study', *npj Parkinson's Disease*, 8(1), pp. 1–5.

Friedman, J.H., Hastie, T. & Tibshirani, R. (2010) 'Regularization Paths for Generalized Linear Models via Coordinate Descent', *Journal of Statistical Software*, 33pp. 1–22.

Fuchs, S., Danßmann, C., Klironomos, F., Winkler, A., Fallmann, J., Kruetzfeldt, L.-M., Szymansky, A., Naderi, J., Bernhart, S.H., Grunewald, L., Helmsauer, K., Rodriguez-Fos, E., Kirchner, M., Mertins, P., Astrahantseff, K., Suenkel, C., Toedling, J., Meggetto, F., Remke, M., et al. (2023) 'Defining the landscape of circular RNAs in neuroblastoma unveils a global suppressive function of MYCN', *Nature Communications*, 14(1), p. 3936.

Fujimaki, M., Saiki, S., Li, Y., Kaga, N., Taka, H., Hatano, T., Ishikawa, K.-I., Oji, Y., Mori, A., Okuzumi, A., Koinuma, T., Ueno, S.-I., Imamichi, Y., Ueno, T., Miura, Y., Funayama, M. & Hattori, N. (2018) 'Serum caffeine and metabolites are reliable biomarkers of early Parkinson disease', *Neurology*, 90(5), pp. e404–e411.

Funayama, M., Ohe, K., Amo, T., Furuya, N., Yamaguchi, J., Saiki, S., Li, Y., Ogaki, K., Ando, M., Yoshino, H., Tomiyama, H., Nishioka, K., Hasegawa, K., Saiki, H., Satake, W., Mogushi, K., Sasaki, R., Kokubo, Y., Kuzuhara, S., et al. (2015) 'CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study', *The Lancet. Neurology*, 14(3), pp. 274–282.

Fung, H.-C., Scholz, S., Matarin, M., Simón-Sánchez, J., Hernandez, D., Britton, A., Gibbs, J.R., Langefeld, C., Stiegert, M.L., Schymick, J., Okun, M.S., Mandel, R.J., Fernandez, H.H., Foote, K.D., Rodríguez, R.L., Peckham, E., Vrieze, F.W.D., Gwinn-Hardy, K., Hardy, J.A., et al. (2006) 'Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data', *The Lancet Neurology*, 5(11), pp. 911–916.

Gaetani, L., Blennow, K., Calabresi, P., Filippo, M.D., Parnetti, L. & Zetterberg, H. (2019) 'Neurofilament light chain as a biomarker in neurological disorders', *Journal of Neurology, Neurosurgery & Psychiatry*, 90(8), pp. 870–881.

Gaffo, E., Boldrin, E., Dal Molin, A., Bresolin, S., Bonizzato, A., Trentin, L., Frasson, C., Debatin, K.-M., Meyer, L.H., te Kronnie, G. & Bortoluzzi, S. (2019) 'Circular RNA differential expression in blood cell populations and exploration of circRNA deregulation in pediatric acute lymphoblastic leukemia', *Scientific Reports*, 9(1), p. 14670.

Gaffo, E., Bonizzato, A., Kronnie, G.T. & Bortoluzzi, S. (2017) 'CirComPara: A Multi-Method Comparative Bioinformatics Pipeline to Detect and Study circRNAs from RNA-seq Data', *Non-Coding RNA*, 3(1), p. 8.

Gaffo, E., Buratin, A., Dal Molin, A. & Bortoluzzi, S. (2021) 'Sensitive, reliable and robust circRNA detection from RNA-seq with CirComPara2', *Briefings in Bioinformatics*, p. bbab418.

Gallego Romero, I., Pai, A.A., Tung, J. & Gilad, Y. (2014) 'RNA-seq: impact of RNA degradation on transcript quantification', *BMC Biology*, 12(1), p. 42.

Gallo, V., Vineis, P., Cancellieri, M., Chiodini, P., Barker, R.A., Brayne, C., Pearce, N., Vermeulen, R., Panico, S., Bueno-de-Mesquita, B., Vanacore, N., Forsgren, L., Ramat, S., Ardanaz, E., Arriola, L., Peterson, J., Hansson, O., Gavrila, D., Sacerdote, C., et al. (2019) 'Exploring causality of the association between smoking and Parkinson's disease', *International Journal of Epidemiology*, 48(3), pp. 912–925.

Gan-Or, Z., Amshalom, I., Kilarski, L.L., Bar-Shira, A., Gana-Weisz, M., Mirelman, A., Marder, K., Bressman, S., Giladi, N. & Orr-Urtreger, A. (2015) 'Differential effects of severe vs mild GBA mutations on Parkinson disease', *Neurology*, 84(9), pp. 880–887.

Gao, Y., Wang, J. & Zhao, F. (2015) 'CIRI: an efficient and unbiased algorithm for de novo circular RNA identification', *Genome Biology*, 16(1), p. 4.

Gao, Y., Zhang, J. & Zhao, F. (2018) 'Circular RNA identification based on multiple seed matching', *Briefings in Bioinformatics*, 19(5), pp. 803–810.

Gao, Y. & Zhao, F. (2018) 'Computational Strategies for Exploring Circular RNAs', *Trends in Genetics*, 34(5), pp. 389–400.

Garcia-Esparcia, P., Hernández-Ortega, K., Koneti, A., Gil, L., Delgado-Morales, R., Castaño, E., Carmona, M. & Ferrer, I. (2015) 'Altered machinery of protein synthesis is region- and stage-dependent and is associated with α-synuclein oligomers in Parkinson's disease', *Acta Neuropathologica Communications*, 3(1), p. 76.

Gardai, S.J., Mao, W., Schüle, B., Babcock, M., Schoebel, S., Lorenzana, C., Alexander, J., Kim, S., Glick, H., Hilton, K., Fitzgerald, J.K., Buttini, M., Chiou, S.-S., McConlogue, L., Anderson, J.P., Schenk, D.B., Bard, F., Langston, J.W., Yednock, T., et al. (2013) 'Elevated Alpha-Synuclein Impairs Innate Immune Cell Function and Provides a Potential Peripheral Biomarker for Parkinson's Disease', *PLOS ONE*, 8(8), p. e71634.

Garneau, N.L., Wilusz, J. & Wilusz, C.J. (2007) 'The highways and byways of mRNA decay', *Nature Reviews Molecular Cell Biology*, 8(2), pp. 113–126.

Garofalo, M., Pandini, C., Bordoni, M., Pansarasa, O., Rey, F., Costa, A., Minafra, B., Diamanti, L., Zucca, S., Carelli, S., Cereda, C. & Gagliardi, S. (2020) 'Alzheimer's, Parkinson's Disease and Amyotrophic Lateral Sclerosis Gene Expression Patterns Divergence Reveals Different Grade of RNA Metabolism Involvement', *International Journal of Molecular Sciences*, 21(24), .

Gaurav, R., Yahia-Cherif, L., Pyatigorskaya, N., Mangone, G., Biondetti, E., Valabrègue, R., Ewenczyk, C., Hutchison, R.M., Cedarbaum, J.M., Corvol, J.-C., Vidailhet, M. & Lehéricy, S. (2021) 'Longitudinal Changes in Neuromelanin MRI Signal in Parkinson's Disease: A Progression Marker', *Movement Disorders*, 36(7), pp. 1592–1602.

Gázquez-Gutiérrez, A., Witteveldt, J., Heras, S.R. & Macias, S. (2021) 'Sensing of transposable elements by the antiviral innate immune system', *RNA*, 27(7), pp. 735–752.

Gegg, M.E. & Schapira, A.H.V. (2018) 'The role of glucocerebrosidase in Parkinson disease pathogenesis', *The FEBS journal*, 285(19), pp. 3591–3603.

Gerkin, R.C., Adler, C.H., Hentz, J.G., Shill, H.A., Driver-Dunckley, E., Mehta, S.H., Sabbagh, M.N., Caviness, J.N., Dugger, B.N., Serrano, G., Belden, C., Smith, B.H., Sue, L., Davis, K.J., Zamrini, E. & Beach, T.G. (2017) 'Improved diagnosis of Parkinson's disease from a detailed olfactory phenotype', *Annals of Clinical and Translational Neurology*, 4(10), pp. 714–721.

Gesi, M., Soldani, P., Giorgi, F.S., Santinami, A., Bonaccorsi, I. & Fornai, F. (2000) 'The role of the locus coeruleus in the development of Parkinson's disease', *Neuroscience & Biobehavioral Reviews*, 24(6), pp. 655–668.

Gil, J. & Esteban, M. (2000) 'The interferon-induced protein kinase (PKR), triggers apoptosis through FADD-mediated activation of caspase 8 in a manner independent of Fas and TNF-α receptors', *Oncogene*, 19(32), pp. 3665–3674.

Gilboa, T., Swank, Z., Thakur, R., Gould, R.A., Ooi, K.H., Norman, M., Flynn, E.A., Deveney, B.T., Chen, A., Borberg, E., Kuzkina, A., Ndayisaba, A., Khurana, V., Weitz, D.A. & Walt, D.R. (2024) 'Toward the quantification of α-synuclein aggregates with digital seed amplification assays', *Proceedings of the National Academy of Sciences*, 121(3), p. e2312031121.

Giordano, F., Saheki, Y., Idevall-Hagren, O., Colombo, S.F., Pirruccello, M., Milosevic, I., Gracheva, E.O., Bagriantsev, S.N., Borgese, N. & De Camilli, P. (2013) 'PI(4,5)P2-Dependent and Ca2+-Regulated ER-PM Interactions Mediated by the Extended Synaptotagmins', *Cell*, 153(7), pp. 1494–1509.

Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M.B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A.E., Lees, A., Leurgans, S., LeWitt, P.A., Nyenhuis, D., et al. (2008) 'Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results', *Movement Disorders*, 23(15), pp. 2129–2170.

Golbe, L.I., Di Iorio, G., Bonavita, V., Miller, D.C. & Duvoisin, R.C. (1990) 'A large kindred with autosomal dominant Parkinson's disease', *Annals of Neurology*, 27(3), pp. 276–282.

Grassi, L., Izuogu, O.G., Jorge, N.A.N., Seyres, D., Bustamante, M., Burden, F., Farrow, S., Farahi, N., Martin, F.J., Frankish, A., Mudge, J.M., Kostadima, M., Petersen, R., Lambourne, J.J., Rowlston, S., Martin-Rendon, E., Clarke, L., Downes, K., Estivill, X., et al. (2021) 'Cell type-specific novel long non-coding RNA and circular RNA in the BLUEPRINT hematopoietic transcriptomes atlas', *Haematologica*, 106(10), pp. 2613–2623.

Greenland, J.C., Williams-Gray, C.H. & Barker, R.A. (2019) 'The clinical heterogeneity of Parkinson's disease and its therapeutic implications', *The European Journal of Neuroscience*, 49(3), pp. 328–338.

Greffard, S., Verny, M., Bonnet, A.-M., Beinis, J.-Y., Gallinari, C., Meaume, S., Piette, F., Hauw, J.-J. & Duyckaerts, C. (2006) 'Motor Score of the Unified Parkinson Disease

Rating Scale as a Good Predictor of Lewy Body–Associated Neuronal Loss in the Substantia Nigra', *Archives of Neurology*, 63(4), pp. 584–588.

Grillo, P., Sancesario, G.M., Bovenzi, R., Zenuni, H., Bissacco, J., Mascioli, D., Simonetta, C., Forti, P., Degoli, G.R., Pieri, M., Chiurchiù, V., Stefani, A., Mercuri, N.B. & Schirinzi, T. (2023) 'Neutrophil-to-lymphocyte ratio and lymphocyte count reflect alterations in central neurodegeneration-associated proteins and clinical severity in Parkinson Disease patients', *Parkinsonism & Related Disorders*, 112p. 105480.

Groiss, S.J., Wojtecki, L., Südmeyer, M. & Schnitzler, A. (2009) 'Deep Brain Stimulation in Parkinson's Disease', *Therapeutic Advances in Neurological Disorders*, 2(6), pp. 20–28.

Grossauer, A., Hemicker, G., Krismer, F., Peball, M., Djamshidian, A., Poewe, W., Seppi, K. & Heim, B. (2023) 'α-Synuclein Seed Amplification Assays in the Diagnosis of Synucleinopathies Using Cerebrospinal Fluid—A Systematic Review and Meta-Analysis', *Movement Disorders Clinical Practice*, 10(5), pp. 737–747.

Groveman, B.R., Orrù, C.D., Hughson, A.G., Raymond, L.D., Zanusso, G., Ghetti, B., Campbell, K.J., Safar, J., Galasko, D. & Caughey, B. (2018) 'Rapid and ultra-sensitive quantitation of disease-associated α-synuclein seeds in brain and cerebrospinal fluid by αSyn RT-QuIC', *Acta Neuropathologica Communications*, 6(1), p. 7.

Grozdanov, V., Bliederhaeuser, C., Ruf, W.P., Roth, V., Fundel-Clemens, K., Zondler, L., Brenner, D., Martin-Villalba, A., Hengerer, B., Kassubek, J., Ludolph, A.C., Weishaupt, J.H. & Danzer, K.M. (2014) 'Inflammatory dysregulation of blood monocytes in Parkinson's disease patients', *Acta Neuropathologica*, 128(5), pp. 651–663.

Gruhl, F., Janich, P., Kaessmann, H. & Gatfield, D. (2021) 'Circular RNA repertoires are associated with evolutionarily young transposable elements' Juan Valcárcel, Detlef Weigel, & Juan Valcárcel (eds.), *eLife*, 10p. e67991.

Grünblatt, E., Zehetmayer, S., Jacob, C.P., Müller, T., Jost, W.H. & Riederer, P. (2010) 'Pilot study: peripheral biomarkers for diagnosing sporadic Parkinson's disease', *Journal of Neural Transmission*, 117(12), pp. 1387–1393.

Gu, Z., Eils, R. & Schlesner, M. (2016) 'Complex heatmaps reveal patterns and correlations in multidimensional genomic data', *Bioinformatics*, 32(18), pp. 2847–2849.

Guarnerio, J., Bezzi, M., Jeong, J.C., Paffenholz, S.V., Berry, K., Naldini, M.M., Lo-Coco, F., Tay, Y., Beck, A.H. & Pandolfi, P.P. (2016) 'Oncogenic Role of Fusion-circRNAs Derived from Cancer-Associated Chromosomal Translocations', *Cell*, 165(2), pp. 289–302.

Guévremont, D., Roy, J., Cutfield, N.J. & Williams, J.M. (2023) 'MicroRNAs in Parkinson's disease: a systematic review and diagnostic accuracy meta-analysis', *Scientific Reports*, 13(1), p. 16272.

Guo, J.U., Agarwal, V., Guo, H. & Bartel, D.P. (2014) 'Expanded identification and characterization of mammalian circular RNAs', *Genome Biology*, 15(7), p. 409.

Guo, S.-K., Liu, C.-X., Xu, Y.-F., Wang, X., Nan, F., Huang, Y., Li, S., Nan, S., Li, L., Kon, E., Li, C., Wei, M.-Y., Su, R., Wei, J., Peng, S., Ad-El, N., Liu, J., Peer, D., Chen, T.,

et al. (2024) ‘Therapeutic application of circular RNA aptamers in a mouse model of psoriasis’, *Nature Biotechnology*, pp. 1–11.

Gustavsson, E.K., Follett, J., Trinh, J., Barodia, S.K., Real, R., Liu, Z., Grant-Peters, M., Fox, J.D., Appel-Cresswell, S., Stoessl, A.J., Rajput, A., Rajput, A.H., Auer, R., Tilney, R., Sturm, M., Haack, T.B., Lesage, S., Tesson, C., Brice, A., et al. (2024) ‘*RAB32* Ser71Arg in autosomal dominant Parkinson’s disease: linkage, association, and functional analyses’, *The Lancet Neurology*,

Güttinger, S., Laurell, E. & Kutay, U. (2009) ‘Orchestrating nuclear envelope disassembly and reassembly during mitosis’, *Nature Reviews Molecular Cell Biology*, 10(3), pp. 178–191.

Haehner, A., Boesveldt, S., Berendse, H.W., Mackay-Sim, A., Fleischmann, J., Silburn, P.A., Johnston, A.N., Mellick, G.D., Herting, B., Reichmann, H. & Hummel, T. (2009) ‘Prevalence of smell loss in Parkinson’s disease – A multicenter study’, *Parkinsonism & Related Disorders*, 15(7), pp. 490–494.

Hallacli, E., Kayatekin, C., Nazeen, S., Wang, X.H., Sheinkopf, Z., Sathyakumar, S., Sarkar, S., Jiang, X., Dong, X., Di Maio, R., Wang, W., Keeney, M.T., Felsky, D., Sandoe, J., Vahdatshoar, A., Udeshi, N.D., Mani, D.R., Carr, S.A., Lindquist, S., et al. (2022) ‘The Parkinson’s disease protein alpha-synuclein is a modulator of processing bodies and mRNA stability’, *Cell*, 185(12), pp. 2035-2056.e33.

Halliday, G.M. & McCann, H. (2010) ‘The progression of pathology in Parkinson’s disease’, *Annals of the New York Academy of Sciences*, 1184(1), pp. 188–195.

Ham, S. & Lee, S.-J.V. (2020) ‘Advances in transcriptome analysis of human brain aging’, *Experimental & Molecular Medicine*, 52(11), pp. 1787–1797.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R. & Hannon, G.J. (2001) ‘Argonaute2, a Link Between Genetic and Biochemical Analyses of RNAi’, *Science*, 293(5532), pp. 1146–1150.

Hamza, T.H., Zabetian, C.P., Tenesa, A., Laederach, A., Montimurro, J., Yearout, D., Kay, D.M., Doheny, K.F., Paschall, J., Pugh, E., Kusel, V.I., Collura, R., Roberts, J., Griffith, A., Samii, A., Scott, W.K., Nutt, J., Factor, S.A. & Payami, H. (2010) ‘Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson’s disease’, *Nature Genetics*, 42(9), pp. 781–785.

Han, Y., Wu, D., Wang, Y., Xie, J. & Zhang, Z. (2022) ‘Skin alpha-synuclein deposit patterns: A predictor of Parkinson’s disease subtypes’, *eBioMedicine*, 80.

Hanan, M., Simchovitz, A., Yayon, N., Vaknine, S., Cohen-Fultheim, R., Karmon, M., Madrer, N., Rohrlich, T.M., Maman, M., Bennett, E.R., Greenberg, D.S., Meshorer, E., Levanon, E.Y., Soreq, H. & Kadener, S. (2020) ‘A Parkinson’s disease CircRNAs Resource reveals a link between circSLC8A1 and oxidative stress’, *EMBO Molecular Medicine*, 12(9), p. e11942.

Hansen, E.B., Fredsøe, J., Okholm, T.L.H., Ulhøi, B.P., Klingenberg, S., Jensen, J.B., Kjems, J., Bouchelouche, K., Borre, M., Damgaard, C.K., Pedersen, J.S., Kristensen, L.S. & Sørensen, K.D. (2022) ‘The transcriptional landscape and biomarker potential of circular RNAs in prostate cancer’, *Genome Medicine*, 14(1), p. 8.

Hansen, T.B. (2018) 'Improved circRNA Identification by Combining Prediction Algorithms', *Frontiers in Cell and Developmental Biology*, 6.

Hansen, T.B. (2021) 'Signal and noise in circRNA translation', *Methods*, 196pp. 68–73.

Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K. & Kjems, J. (2013) 'Natural RNA circles function as efficient microRNA sponges', *Nature*, 495(7441), pp. 384–388.

Hansen, T.B., Venø, M.T., Damgaard, C.K. & Kjems, J. (2016) 'Comparison of circular RNA prediction tools', *Nucleic Acids Research*, 44(6), p. e58.

Hansen, T.B., Wiklund, E.D., Bramsen, J.B., Villadsen, S.B., Statham, A.L., Clark, S.J. & Kjems, J. (2011) 'miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA', *The EMBO Journal*, 30(21), pp. 4414–4422.

Hansson, O., Janelidze, S., Hall, S., Magdalinou, N., Lees, A.J., Andreasson, U., Norgren, N., Linder, J., Forsgren, L., Constantinescu, R., Zetterberg, H., Blennow, K., & For the Swedish BioFINDER study (2017) 'Blood-based NfL', *Neurology*, 88(10), pp. 930–937.

Hansson, O., Seibyl, J., Stomrud, E., Zetterberg, H., Trojanowski, J.Q., Bittner, T., Lifke, V., Corradini, V., Eichenlaub, U., Batrla, R., Buck, K., Zink, K., Rabe, C., Blennow, K. & Shaw, L.M. (2018) 'CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts', *Alzheimer's & Dementia*, 14(11), pp. 1470–1481.

Harding, H.P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M. & Ron, D. (2000) 'Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells', *Molecular Cell*, 6(5), pp. 1099–1108.

Harrison, P.J., Procter, A.W., Barton, A.J.L., Lowe, S.L., Najlerahim, A., Bertolucci, P.H.F., Bowen, D.M. & Pearson, R.C.A. (1991) 'Terminal coma affects messenger RNA detection in post mortem human temporal cortex', *Molecular Brain Research*, 9(1), pp. 161–164.

Hatano, T., Saiki, S., Okuzumi, A., Mohney, R.P. & Hattori, N. (2016) 'Identification of novel biomarkers for Parkinson's disease by metabolomic technologies', *Journal of Neurology, Neurosurgery & Psychiatry*, 87(3), pp. 295–301.

Hayashi, T., Ozaki, H., Sasagawa, Y., Umeda, M., Danno, H. & Nikaido, I. (2018) 'Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs', *Nature Communications*, 9(1), p. 619.

He, D., Yang, C.X., Sahin, B., Singh, A., Shannon, C.P., Oliveria, J.-P., Gauvreau, G.M. & Tebbutt, S.J. (2019) 'Whole blood vs PBMC: compartmental differences in gene expression profiling exemplified in asthma', *Allergy, Asthma & Clinical Immunology*, 15(1), p. 67.

Healy, D.G., Falchi, M., O'Sullivan, S.S., Bonifati, V., Durr, A., Bressman, S., Brice, A., Aasly, J., Zabetian, C.P., Goldwurm, S., Ferreira, J.J., Tolosa, E., Kay, D.M., Klein, C., Williams, D.R., Marras, C., Lang, A.E., Wszolek, Z.K., Berciano, J., et al. (2008)

‘Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson’s disease: a case-control study’, *The Lancet Neurology*, 7(7), pp. 583–590.

van Heesch, S., Witte, F., Schneider-Lunitz, V., Schulz, J.F., Adami, E., Faber, A.B., Kirchner, M., Maatz, H., Blachut, S., Sandmann, C.-L., Kanda, M., Worth, C.L., Schafer, S., Calviello, L., Merriott, R., Patone, G., Hummel, O., Wyler, E., Obermayer, B., et al. (2019) ‘The Translational Landscape of the Human Heart’, *Cell*, 178(1), pp. 242-260.e29.

Heinzel, S., Berg, D., Gasser, T., Chen, H., Yao, C., Postuma, R.B. & Disease, the M.T.F. on the D. of P. (2019) ‘Update of the MDS research criteria for prodromal Parkinson’s disease’, *Movement Disorders*, 34(10), pp. 1464–1470.

Hely, M.A., Reid, W.G.J., Adena, M.A., Halliday, G.M. & Morris, J.G.L. (2008) ‘The Sydney multicenter study of Parkinson’s disease: the inevitability of dementia at 20 years’, *Movement Disorders: Official Journal of the Movement Disorder Society*, 23(6), pp. 837–844.

Henderson, A.R., Wang, Q., Meechoovet, B., Siniard, A.L., Naymik, M., De Both, M., Huentelman, M.J., Caselli, R.J., Driver-Dunckley, E. & Dunckley, T. (2021) ‘DNA Methylation and Expression Profiles of Whole Blood in Parkinson’s Disease’, *Frontiers in Genetics*, 12.

Henry, A.G., Aghamohammadzadeh, S., Samaroo, H., Chen, Y., Mou, K., Needle, E. & Hirst, W.D. (2015) ‘Pathogenic LRRK2 mutations, through increased kinase activity, produce enlarged lysosomes with reduced degradative capacity and increase ATP13A2 expression’, *Human Molecular Genetics*, 24(21), pp. 6013–6028.

Herrick, M.K. & Tansey, M.G. (2021) ‘Is LRRK2 the missing link between inflammatory bowel disease and Parkinson’s disease?’, *npj Parkinson’s Disease*, 7(1), pp. 1–7.

Hill-Burns, E.M., Factor, S.A., Zabetian, C.P., Thomson, G. & Payami, H. (2011) ‘Evidence for More than One Parkinson’s Disease-Associated Variant within the HLA Region’, *PLOS ONE*, 6(11), p. e27109.

’t Hoen, P.A.C., Friedländer, M.R., Almlöf, J., Sammeth, M., Pulyakhina, I., Anvar, S.Y., Laros, J.F.J., Buermans, H.P.J., Karlberg, O., Brännvall, M., Johan T. den Dunnen, van Ommen, G.-J.B., Gut, I.G., Guigó, R., Xavier Estivill, Syvänen, A.-C., Gut, I.G., Dermitzakis, E.T. & Lappalainen, T. (2013) ‘Reproducibility of high-throughput mRNA and small RNA sequencing across laboratories’, *Nature Biotechnology*, 31(11), pp. 1015–1022.

Hoffman, G.E. & Schadt, E.E. (2016) ‘variancePartition: interpreting drivers of variation in complex gene expression studies’, *BMC Bioinformatics*, 17(1), p. 483.

Hogg, M., Paro, S., Keegan, L.P. & O’Connell, M.A. (2011) ‘RNA editing by mammalian ADARs’, *Advances in Genetics*, 73pp. 87–120.

Höglinger, G.U., Adler, C.H., Berg, D., Klein, C., Outeiro, T.F., Poewe, W., Postuma, R., Stoessl, A.J. & Lang, A.E. (2024) ‘A biological classification of Parkinson’s disease: the SynNeurGe research diagnostic criteria’, *The Lancet Neurology*, 23(2), pp. 191–204.

Holden, S.K., Finseth, T., Sillau, S.H. & Berman, B.D. (2018) 'Progression of MDS-UPDRS Scores Over Five Years in De Novo Parkinson Disease from the Parkinson's Progression Markers Initiative Cohort', *Movement Disorders Clinical Practice*, 5(1), pp. 47–53.

Hoogland, J., Boel, J.A., de Bie, R.M.A., Geskus, R.B., Schmand, B.A., Dalrymple-Alford, J.C., Marras, C., Adler, C.H., Goldman, J.G., Tröster, A.I., Burn, D.J., Litvan, I., Geurtsen, G.J. & Disease", on behalf of the M.S.G. "Validation of M.C.I. in P. (2017) 'Mild cognitive impairment as a risk factor for Parkinson's disease dementia', *Movement Disorders*, 32(7), pp. 1056–1065.

Hoopengardner, B., Bhalla, T., Staber, C. & Reenan, R. (2003) 'Nervous System Targets of RNA Editing Identified by Comparative Genomics', *Science*, 301(5634), pp. 832–836.

Hop, P.J., Lai, D., Keagle, P.J., Baron, D.M., Kenna, B.J., Kooyman, M., Shankaracharya, Halter, C., Straniero, L., Asselta, R., Bonvegna, S., Soto-Beasley, A.I., Wszolek, Z.K., Uitti, R.J., Isaias, I.U., Pezzoli, G., Ticozzi, N., Ross, O.A., Veldink, J.H., et al. (2024) 'Systematic rare variant analyses identify RAB32 as a susceptibility gene for familial Parkinson's disease', *Nature Genetics*, pp. 1–6.

Horowitz, M., Pasmanik-Chor, M., Borochowitz, Z., Falik-Zaccai, T., Heldmann, K., Carmi, R., Parvari, R., Beit-Or, H., Goldman, B., Peleg, L., Levy-Lahad, E., Renbaum, P., Legum, S., Shomrat, R., Yeger, H., Benbenisti, D., Navon, R., Dror, V., Shohat, M., et al. (1998) 'Prevalence of glucocerebrosidase mutations in the Israeli Ashkenazi Jewish population', *Human Mutation*, 12(4), pp. 240–244.

Hosseini, S., Shafiabadi, N., Khanzadeh, M., Ghaedi, A., Ghorbanzadeh, R., Azarhomayoun, A., Bazrgar, A., Pezeshki, J., Bazrafshan, H. & Khanzadeh, S. (2023) 'Neutrophil to lymphocyte ratio in parkinson's disease: a systematic review and meta-analysis', *BMC Neurology*, 23(1), p. 333.

Houser, M.C. & Tansey, M.G. (2017) 'The gut-brain axis: is intestinal inflammation a silent driver of Parkinson's disease pathogenesis?', *npj Parkinson's Disease*, 3(1), pp. 1–9.

Ho-Xuan, H., Glažar, P., Latini, C., Heizler, K., Haase, J., Hett, R., Anders, M., Weichmann, F., Bruckmann, A., Van den Berg, D., Hüttelmaier, S., Rajewsky, N., Hackl, C. & Meister, G. (2020) 'Comprehensive analysis of translation from overexpressed circular RNAs reveals pervasive translation from linear transcripts', *Nucleic Acids Research*, 48(18), pp. 10368–10382.

Hsu, M.-T. & Coca-Prados, M. (1979) 'Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells', *Nature*, 280(5720), pp. 339–340.

Huang, C., Liang, D., Tatomer, D.C. & Wilusz, J.E. (2018) 'A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs', *Genes & Development*, 32(9–10), pp. 639–644.

Huang, H., Fang, M., Jostins, L., Umićević Mirkov, M., Boucher, G., Anderson, C.A., Andersen, V., Cleynen, I., Cortes, A., Crins, F., D'Amato, M., Deffontaine, V., Dmitrieva, J., Docampo, E., Elansary, M., Farh, K.K.-H., Franke, A., Gori, A.-S., Goyette, P., et al. (2017) 'Fine-mapping inflammatory bowel disease loci to single-variant resolution', *Nature*, 547(7662), pp. 173–178.

Huang, H., Luo, J., Qi, Y., Wu, Y., Qi, J., Yan, X., Xu, G., He, F. & Zheng, Y. (2023) 'Comprehensive analysis of circRNA expression profile and circRNA-miRNA-mRNA network susceptibility to very early-onset schizophrenia', *Schizophrenia*, 9(1), pp. 1–10.

Huang, Y.-M., Ma, Y.-H., Gao, P.-Y., Cui, X.-H., Hou, J.-H., Chi, H.-C., Fu, Y., Wang, Z.-B., Feng, J.-F., Cheng, W., Tan, L. & Yu, J.-T. (2024) 'Genetic susceptibility modifies the association of long-term air pollution exposure on Parkinson's disease', *npj Parkinson's Disease*, 10(1), pp. 1–7.

Hughes, A.J., Daniel, S.E., Kilford, L. & Lees, A.J. (1992) 'Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases', *Journal of Neurology, Neurosurgery, and Psychiatry*, 55(3), pp. 181–184.

Hui, K.Y., Fernandez-Hernandez, H., Hu, J., Schaffner, A., Pankratz, N., Hsu, N.-Y., Chuang, L.-S., Carmi, S., Villaverde, N., Li, X., Rivas, M., Levine, A.P., Bao, X., Labrias, P.R., Haritunians, T., Ruane, D., Gettler, K., Chen, E., Li, D., et al. (2018) 'Functional variants in the LRRK2 gene confer shared effects on risk for Crohn's disease and Parkinson's disease', *Science Translational Medicine*, 10(423), p. eaai7795.

Hur, S. (2019) 'Double-Stranded RNA Sensors and Modulators in Innate Immunity', *Annual Review of Immunology*, 37pp. 349–375.

Hutchins, E., Craig, D., Violich, I., Alsop, E., Casey, B., Hutten, S., Reimer, A., Whitsett, T.G., Crawford, K.L., Toga, A.W., Levy, S., Robinson, M., Prasad, N., Gibbs, J.R., Cookson, M.R., Keuren-Jensen, K.V. & Initiative, P.P.M. (2021) 'Quality Control Metrics for Whole Blood Transcriptome Analysis in the Parkinson's Progression Markers Initiative (PPMI)', *medRxiv*, p. 2021.01.05.21249278.

Hutchins, E., Reiman, R., Winarta, J., Beecroft, T., Richholt, R., De Both, M., Shahbander, K., Carlson, E., Janss, A., Siniard, A., Balak, C., Bruhns, R., Whitsett, T.G., McCoy, R., Anastasi, M., Allen, A., Churas, B., Huentelman, M. & Van Keuren-Jensen, K. (2021) 'Extracellular circular RNA profiles in plasma and urine of healthy, male college athletes', *Scientific Data*, 8(1), p. 276.

Imai, Y., Kobayashi, Y., Inoshita, T., Meng, H., Arano, T., Uemura, K., Asano, T., Yoshimi, K., Zhang, C.-L., Matsumoto, G., Ohtsuka, T., Kageyama, R., Kiyonari, H., Shioi, G., Nukina, N., Hattori, N. & Takahashi, R. (2015) 'The Parkinson's Disease-Associated Protein Kinase LRRK2 Modulates Notch Signaling through the Endosomal Pathway', *PLoS Genetics*, 11(9), p. e1005503.

Infante, J., Prieto, C., Sierra, M., Sánchez-Juan, P., González-Aramburu, I., Sánchez-Quintana, C., Berciano, J., Combarros, O. & Sainz, J. (2016) 'Comparative blood transcriptome analysis in idiopathic and LRRK2 G2019S-associated Parkinson's disease', *Neurobiology of Aging*, 38p. 214.e1-214.e5.

International Parkinson Disease Genomics Consortium (2011) 'Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies', *Lancet*, 377(9766), pp. 641–649.

Iranzo, A., Fairfoul, G., Ayudhaya, A.C.N., Serradell, M., Gelpi, E., Vilaseca, I., Sanchez-Valle, R., Gaig, C., Santamaria, J., Tolosa, E., Riha, R.L. & Green, A.J.E. (2021) 'Detection of α-synuclein in CSF by RT-QuIC in patients with isolated rapid-eye-

movement sleep behaviour disorder: a longitudinal observational study', *The Lancet Neurology*, 20(3), pp. 203–212.

Iranzo, A., Stefani, A., Niñerola-Baizan, A., Stokner, H., Serradell, M., Vilas, D., Holzknecht, E., Gaig, C., Pavia, J., Lomeña, F., Reyes, D., Seppi, K., Santamaria, J., Högl, B., Tolosa, E., Poewe, W., & the Sleep Innsbruck Barcelona (SINBAR) Group (2020) 'Left-hemispheric predominance of nigrostriatal deficit in isolated REM sleep behavior disorder', *Neurology*, 94(15), pp. e1605–e1613.

Iranzo, A., Tolosa, E., Gelpi, E., Molinuevo, J.L., Valldeoriola, F., Serradell, M., Sanchez-Valle, R., Vilaseca, I., Lomeña, F., Vilas, D., LLadó, A., Gaig, C. & Santamaria, J. (2013) 'Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study', *The Lancet Neurology*, 12(5), pp. 443–453.

Irmady, K., Hale, C.R., Qadri, R., Fak, J., Simelane, S., Carroll, T., Przedborski, S. & Darnell, R.B. (2023) 'Blood transcriptomic signatures associated with molecular changes in the brain and clinical outcomes in Parkinson's disease', *Nature Communications*, 14(1), p. 3956.

Isaacson, S.H., Hauser, R.A., Pahwa, R., Gray, D. & Duvvuri, S. (2023) 'Dopamine agonists in Parkinson's disease: Impact of D1-like or D2-like dopamine receptor subtype selectivity and avenues for future treatment', *Clinical Parkinsonism & Related Disorders*, 9p. 100212.

Isakova, A., Neff, N. & Quake, S.R. (2021) 'Single-cell quantification of a broad RNA spectrum reveals unique noncoding patterns associated with cell types and states', *Proceedings of the National Academy of Sciences*, 118(51), p. e2113568118.

Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H.T., Orejuela, M.R., Piechotta, M., Levanon, E.Y., Landthaler, M., Dieterich, C. & Rajewsky, N. (2015) 'Analysis of Intron Sequences Reveals Hallmarks of Circular RNA Biogenesis in Animals', *Cell Reports*, 10(2), pp. 170–177.

Izuogu, O.G., Alhasan, A.A., Alafghani, H.M., Santibanez-Koref, M., Elliott, D.J. & Jackson, M.S. (2016) 'PTESFinder: a computational method to identify post-transcriptional exon shuffling (PTES) events', *BMC Bioinformatics*, 17(1), p. 31.

Izuogu, O.G., Alhasan, A.A., Mellough, C., Collin, J., Gallon, R., Hyslop, J., Mastrorosa, F.K., Ehrmann, I., Lako, M., Elliott, D.J., Santibanez-Koref, M. & Jackson, M.S. (2018) 'Analysis of human ES cell differentiation establishes that the dominant isoforms of the lncRNAs RMST and FIRRE are circular', *BMC Genomics*, 19(1), p. 276.

Jakobi, T. & Dieterich, C. (2019) 'Computational approaches for circular RNA analysis', *WIREs RNA*, 10(3), p. e1528.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013) *An introduction to statistical learning*. Vol. 112. Springer.

Jankovic, J. (2008) 'Parkinson's disease: clinical features and diagnosis', *Journal of Neurology, Neurosurgery & Psychiatry*, 79(4), pp. 368–376.

Jankovic, J., Goodman, I., Safirstein, B., Marmon, T.K., Schenk, D.B., Koller, M., Zago, W., Ness, D.K., Griffith, S.G., Grundman, M., Soto, J., Ostrowitzki, S., Boess, F.G., Martin-Facklam, M., Quinn, J.F., Isaacson, S.H., Omidvar, O., Ellenbogen, A. & Kinney, G.G. (2018) 'Safety and Tolerability of Multiple Ascending Doses of PRX002/RG7935, an Anti–α-Synuclein Monoclonal Antibody, in Patients With Parkinson Disease: A Randomized Clinical Trial', *JAMA Neurology*, 75(10), pp. 1206–1214.

Jankovic, J., McDermott, M., Carter, J., Gauthier, S., Goetz, C., Golbe, L., Huber, S., Koller, W., Olanow, C. & Shoulson, I. (1990) 'Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort. The Parkinson Study Group', *Neurology*, 40(10), pp. 1529–1534.

Jarlstad Olesen, M.T. & S Kristensen, L. (2021) 'Circular RNAs as microRNA sponges: evidence and controversies', *Essays in Biochemistry*, 65(4), pp. 685–696.

Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F. & Sharpless, N.E. (2013) 'Circular RNAs are abundant, conserved, and associated with ALU repeats', *RNA (New York, N.Y.)*, 19(2), pp. 141–157.

Jennings, D., Huntwork-Rodriguez, S., Henry, A.G., Sasaki, J.C., Meisner, R., Diaz, D., Solanoy, H., Wang, X., Negrou, E., Bondar, V.V., Ghosh, R., Maloney, M.T., Propson, N.E., Zhu, Y., Maciuca, R.D., Harris, L., Kay, A., LeWitt, P., King, T.A., et al. (2022) 'Preclinical and clinical evaluation of the LRRK2 inhibitor DNL201 for Parkinson's disease', *Science Translational Medicine*, 14(648), p. eabj2658.

Jennings, D.L., Seibyl, J.P., Oakes, D., Eberly, S., Murphy, J. & Marek, K. (2004) '(123I) β-CIT and Single-Photon Emission Computed Tomographic Imaging vs Clinical Evaluation in Parkinsonian Syndrome: Unmasking an Early Diagnosis', *Archives of Neurology*, 61(8), pp. 1224–1229.

Jensen, M.P., Jacobs, B.M., Dobson, R., Bandres-Ciga, S., Blauwendraat, C., Schrag, A., Noyce, A.J. & Consortium (IPDGC), T.I.P.D.G. (2021) 'Lower Lymphocyte Count is Associated With Increased Risk of Parkinson's Disease', *Annals of Neurology*, 89(4), pp. 803–812.

Jesús Gil, J.A. & Esteban, M. (1999) 'Induction of Apoptosis by Double-Stranded-RNA-Dependent Protein Kinase (PKR) Involves the α Subunit of Eukaryotic Translation Initiation Factor 2 and NF-κB', *Molecular and Cellular Biology*, 19(7), pp. 4653–4663.

Ji, P., Wu, W., Chen, S., Zheng, Y., Zhou, L., Zhang, J., Cheng, H., Yan, J., Zhang, S., Yang, P. & Zhao, F. (2019) 'Expanded Expression Landscape and Prioritization of Circular RNAs in Mammals', *Cell Reports*, 26(12), pp. 3444-3460.e5.

Jia, E., Zhou, Y., Liu, Z., Wang, L., Ouyang, T., Pan, M., Bai, Y. & Ge, Q. (2020) 'Transcriptomic Profiling of Circular RNA in Different Brain Regions of Parkinson's Disease in a Mouse Model', *International Journal of Molecular Sciences*, 21(8), p. 3006.

Jia, R., Xiao, M.-S., Li, Z., Shan, G. & Huang, C. (2019) 'Defining an evolutionarily conserved role of GW182 in circular RNA degradation', *Cell Discovery*, 5(1), pp. 1–4.

Jin, H. & Liu, Z. (2021) 'A benchmark for RNA-seq deconvolution analysis under dynamic testing environments', *Genome Biology*, 22(1), p. 102.

Johansson, M.E., van Lier, N.M., Kessels, R.P.C., Bloem, B.R. & Helmich, R.C. (2023) 'Two-year clinical progression in focal and diffuse subtypes of Parkinson's disease', *npj Parkinson's Disease*, 9(1), pp. 1–11.

Kakuda, K., Ikenaka, K., Araki, K., So, M., Aguirre, C., Kajiyama, Y., Konaka, K., Noi, K., Baba, K., Tsuda, H., Nagano, S., Ohmichi, T., Nagai, Y., Tokuda, T., El-Agnaf, O.M.A., Ogi, H., Goto, Y. & Mochizuki, H. (2019) 'Ultrasonication-based rapid amplification of α-synuclein aggregates in cerebrospinal fluid', *Scientific Reports*, 9(1), p. 6001.

Kalia, L.V. & Lang, A.E. (2015) 'Parkinson's disease', *The Lancet*, 386(9996), pp. 896–912.

Kalia, L.V., Lang, A.E., Hazrati, L.-N., Fujioka, S., Wszolek, Z.K., Dickson, D.W., Ross, O.A., Van Deerlin, V.M., Trojanowski, J.Q., Hurtig, H.I., Alcalay, R.N., Marder, K.S., Clark, L.N., Gaig, C., Tolosa, E., Ruiz-Martínez, J., Marti-Masso, J.F., Ferrer, I., López de Munain, A., et al. (2015) 'Clinical Correlations With Lewy Body Pathology in LRRK2-Related Parkinson Disease', *JAMA Neurology*, 72(1), pp. 100–105.

Kamath, T., Abdulraouf, A., Burris, S.J., Langlieb, J., Gazestani, V., Nadaf, N.M., Balderrama, K., Vanderburg, C. & Macosko, E.Z. (2022) 'Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease', *Nature Neuroscience*, 25(5), pp. 588–595.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. (2023) 'KEGG for taxonomy-based analysis of pathways and genomes', *Nucleic Acids Research*, 51(D1), pp. D587–D592.

Kanehisa, M. & Goto, S. (2000) 'KEGG: Kyoto Encyclopedia of Genes and Genomes', *Nucleic Acids Research*, 28(1), pp. 27–30.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. (2016) 'KEGG as a reference resource for gene and protein annotation', *Nucleic Acids Research*, 44(D1), pp. D457–D462.

Kang, R. & Tang, D. (2012) 'PKR-Dependent Inflammatory Signals', *Science Signaling*, 5(247), pp. pe47–pe47.

Kapoor, U., Licht, K., Amman, F., Jakobi, T., Martin, D., Dieterich, C. & Jantsch, M.F. (2020) 'ADAR-deficiency perturbs the global splicing landscape in mouse tissues', *Genome Research*, 30(8), pp. 1107–1118.

Katze, M.G., He, Y. & Gale, M. (2002) 'Viruses and interferon: a fight for supremacy', *Nature Reviews Immunology*, 2(9), pp. 675–687.

Kelly, S., Greenman, C., Cook, P.R. & Papantonis, A. (2015) 'Exon Skipping Is Correlated with Exon Circularization', *Journal of Molecular Biology*, 427(15), pp. 2414–2417.

Kern, F., Fehlmann, T., Violich, I., Alsop, E., Hutchins, E., Kahraman, M., Grammes, N.L., Guimarães, P., Backes, C., Poston, K.L., Casey, B., Balling, R., Geffers, L., Krüger, R., Galasko, D., Mollenhauer, B., Meese, E., Wyss-Coray, T., Craig, D.W., et al. (2021) 'Deep sequencing of sncRNAs reveals hallmarks and regulatory modules of the transcriptome during Parkinson's disease progression', *Nature Aging*, 1(3), pp. 309–322.

Kestenbaum, M. & Alcalay, R.N. (2017) 'Clinical Features of LRRK2 Carriers with Parkinson's Disease', in Hardy J. Rideout (ed.) *Leucine-Rich Repeat Kinase 2 (LRRK2)*. [Online]. Cham: Springer International Publishing. pp. 31–48.

Khodadadi, H., Azcona, L.J., Aghamollaii, V., Omrani, M.D., Garshasbi, M., Taghavi, S., Tafakhori, A., Shahidi, G.A., Jamshidi, J., Darvish, H. & Paisán-Ruiz, C. (2017) 'PTRHD1 (C2orf79) mutations lead to autosomal-recessive intellectual disability and parkinsonism', *Movement Disorders*, 32(2), pp. 287–291.

Kim, C.Y. & Alcalay, R.N. (2017) 'Genetic Forms of Parkinson's Disease', *Seminars in Neurology*, 37pp. 135–146.

Kim, D., Paggi, J.M., Park, C., Bennett, C. & Salzberg, S.L. (2019) 'Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype', *Nature Biotechnology*, 37(8), pp. 907–915.

Kim, S., Ku, Y., Ku, J. & Kim, Y. (2019) 'Evidence of Aberrant Immune Response by Endogenous Double-Stranded RNAs: Attack from Within', *BioEssays*, 41(7), p. 1900023.

Kim, Y., Lee, J.H., Park, J.-E., Cho, J., Yi, H. & Kim, V.N. (2014) 'PKR is activated by cellular dsRNAs during mitosis and acts as a mitotic regulator', *Genes & Development*, 28(12), pp. 1310–1322.

Kim, Y., Park, J., Kim, S., Kim, M., Kang, M.-G., Kwak, C., Kang, M., Kim, B., Rhee, H.-W. & Kim, V.N. (2018) 'PKR Senses Nuclear and Mitochondrial Signals by Interacting with Endogenous Double-Stranded RNAs', *Molecular Cell*, 71(6), pp. 1051-1063.e6.

Kim, Y.J., Ichise, M., Ballinger, J.R., Vines, D., Erami, S.S., Tatschida, T. & Lang, A.E. (2002) 'Combination of dopamine transporter and D2 receptor SPECT in the diagnostic evaluation of PD, MSA, and PSP', *Movement Disorders*, 17(2), pp. 303–312.

Klatt, S., Doecke, J.D., Roberts, A., Boughton, B.A., Masters, C.L., Horne, M. & Roberts, B.R. (2021) 'A six-metabolite panel as potential blood-based biomarkers for Parkinson's disease', *npj Parkinson's Disease*, 7(1), pp. 1–14.

Kleaveland, B., Shi, C.Y., Stefano, J. & Bartel, D.P. (2018) 'A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain', *Cell*, 174(2), pp. 350-362.e17.

Klein, C. & Westenberger, A. (2012) 'Genetics of Parkinson's Disease', *Cold Spring Harbor Perspectives in Medicine*, 2(1), .

Kluge, A., Bunk, J., Schaeffer, E., Drobny, A., Xiang, W., Knacke, H., Bub, S., Lückstädt, W., Arnold, P., Lucius, R., Berg, D. & Zunke, F. (2022) 'Detection of neuron-derived pathological α-synuclein in blood', *Brain*, p. awac115.

Kluge, A., Schaeffer, E., Bunk, J., Sommerauer, M., Röttgen, S., Schulte, C., Roeben, B., von Thaler, A.-K., Welzel, J., Lucius, R., Heinzel, S., Xiang, W., Eschweiler, G.W., Maetzler, W., Suenkel, U. & Berg, D. (2024) 'Detecting Misfolded α-Synuclein in Blood Years before the Diagnosis of Parkinson's Disease', *Movement Disorders*, n/a(n/a), .

Kmiecik, M.J., Micheletti, S., Coker, D., Heilbron, K., Shi, J., Stagaman, K., Filshtein Sonmez, T., Fontanillas, P., Shringarpure, S., Wetzel, M., Rowbotham, H.M., Cannon, P.,

Shelton, J.F., Hinds, D.A., Tung, J.Y., 23andMe Research Team, Holmes, M.V., Aslibekyan, S. & Norcliffe-Kaufmann, L. (2024) 'Genetic analysis and natural history of Parkinson's disease due to the LRRK2 G2019S variant', *Brain*, p. awae073.

Knudsen, K., Fedorova, T.D., Hansen, A.K., Sommerauer, M., Otto, M., Svendsen, K.B., Nahimi, A., Stokholm, M.G., Pavese, N., Beier, C.P., Brooks, D.J. & Borghammer, P. (2018) 'In-vivo staging of pathology in REM sleep behaviour disorder: a multimodality imaging case-control study', *The Lancet Neurology*, 17(7), pp. 618–628.

Knupp, D., Cooper, D.A., Saito, Y., Darnell, R.B. & Miura, P. (2021) 'NOVA2 regulates neural circRNA biogenesis', *Nucleic Acids Research*, 49(12), pp. 6849–6862.

Koch, S., Laabs, B.-H., Kasten, M., Vollstedt, E.-J., Becktepe, J., Brüggemann, N., Franke, A., Krämer, U.M., Kuhlenbäumer, G., Lieb, W., Mollenhauer, B., Neis, M., Trenkwalder, C., Schäffer, E., Usnich, T., Wittig, M., Klein, C., König, I.R., Lohmann, K., et al. (2021) 'Validity and Prognostic Value of a Polygenic Risk Score for Parkinson's Disease', *Genes*, 12(12), p. 1859.

Koh, H.R., Xing, L., Kleiman, L. & Myong, S. (2014) 'Repetitive RNA unwinding by RNA helicase A facilitates RNA annealing', *Nucleic Acids Research*, 42(13), pp. 8556–8564.

Kokot, K.E., Kneuer, J.M., John, D., Rebs, S., Möbius-Winkler, M.N., Erbe, S., Müller, M., Andritschke, M., Gaul, S., Sheikh, B.N., Haas, J., Thiele, H., Müller, O.J., Hille, S., Leuschner, F., Dimmeler, S., Streckfuss-Bömeke, K., Meder, B., Laufs, U., et al. (2022) 'Reduction of A-to-I RNA editing in the failing human heart regulates formation of circular RNAs', *Basic Research in Cardiology*, 117(1), p. 32.

Kolakofsky, D. (1976) 'Isolation and characterization of Sendai virus DI-RNAs', *Cell*, 8(4), pp. 547–555.

Kolmykov, S., Yevshin, I., Kulyashov, M., Sharipov, R., Kondrakhin, Y., Makeev, V.J., Kulakovskiy, I.V., Kel, A. & Kolpakov, F. (2021) 'GTRD: an integrated view of transcription regulation', *Nucleic Acids Research*, 49(D1), pp. D104–D111.

Kong, F., Lv, Z., Wang, L., Zhang, K., Cai, Y., Ding, Q., Sun, Z., Zhen, H., Jiao, F., Ma, Q., Nie, C. & Yang, Y. (2021) 'RNA-sequencing of peripheral blood circular RNAs in Parkinson disease', *Medicine*, 100(23), p. e25888.

Kordower, J.H., Olanow, C.W., Dodiya, H.B., Chu, Y., Beach, T.G., Adler, C.H., Halliday, G.M. & Bartus, R.T. (2013) 'Disease duration and the integrity of the nigrostriatal system in Parkinson's disease', *Brain*, 136(8), pp. 2419–2431.

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N. & Sergushichev, A. (2021) *Fast gene set enrichment analysis*. p.p. 060012.

Korsgaard, U., García-Rodríguez, J.L., Jakobsen, T., Ahmadov, U., Dietrich, K.-G., Vissing, S.M., Paasch, T.P., Lindebjerg, J., Kjems, J., Hager, H. & Kristensen, L.S. (2024) 'The Transcriptional Landscape of Coding and Noncoding RNAs in Recurrent and Nonrecurrent Colon Cancer', *The American Journal of Pathology*,

Kos, A., Dijkema, R., Arnberg, A.C., van der Meide, P.H. & Schellekens, H. (1986) 'The hepatitis delta (delta) virus possesses a circular RNA', *Nature*, 323(6088), pp. 558–560.

Kosti, I., Jain, N., Aran, D., Butte, A.J. & Sirota, M. (2016) 'Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues', *Scientific Reports*, 6(1), p. 24799.

Kramer, M.C., Liang, D., Tatomer, D.C., Gold, B., March, Z.M., Cherry, S. & Wilusz, J.E. (2015) 'Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins', *Genes & Development*, 29(20), pp. 2168–2182.

Kristensen, L.S., Andersen, M.S., Stagsted, L.V.W., Ebbesen, K.K., Hansen, T.B. & Kjems, J. (2019) 'The biogenesis, biology and characterization of circular RNAs', *Nature Reviews Genetics*, 20(11), pp. 675–691.

Kristensen, L.S., Ebbesen, K.K., Sokol, M., Jakobsen, T., Korsgaard, U., Eriksen, A.C., Hansen, T.B., Kjems, J. & Hager, H. (2020) 'Spatial expression analyses of the putative oncogene ciRS-7 in cancer reshape the microRNA sponge theory', *Nature Communications*, 11(1), p. 4551.

Kuhen, K.L. & Samuel, C.E. (1997) 'Isolation of the Interferon-Inducible RNA-Dependent Protein Kinase*Pkr*Promoter and Identification of a Novel DNA Element within the 5′-Flanking Region of Human and Mouse*Pkr*Genes', *Virology*, 227(1), pp. 119–130.

Kumar, L., Shamsuzzama, Jadiya, P., Haque, R., Shukla, S. & Nazir, A. (2018) 'Functional Characterization of Novel Circular RNA Molecule, circzip-2 and Its Synthesizing Gene zip-2 in C. elegans Model of Parkinson's Disease', *Molecular Neurobiology*, 55(8), pp. 6914–6926.

Kurvits, L., Lättekivi, F., Reimann, E., Kadastik-Eerme, L., Kasterpalu, K.M., Kõks, S., Taba, P. & Planken, A. (2021) 'Transcriptomic profiles in Parkinson's disease', *Experimental Biology and Medicine*, 246(5), pp. 584–595.

Kyogoku, C., Smiljanovic, B., Grün, J.R., Biesen, R., Schulte-Wrede, U., Häupl, T., Hiepe, F., Alexander, T., Radbruch, A. & Grützkau, A. (2013) 'Cell-Specific Type I IFN Signatures in Autoimmunity and Viral Infection: What Makes the Difference?', *PLoS ONE*, 8(12), p. e83776.

Lai, E.C. (2002) 'Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation', *Nature Genetics*, 30(4), pp. 363–364.

Lanciano, S. & Cristofari, G. (2020) 'Measuring and interpreting transposable element expression', *Nature Reviews Genetics*, 21(12), pp. 721–736.

Lang Anthony E., Siderowf Andrew D., Macklin Eric A., Poewe Werner, Brooks David J., Fernandez Hubert H., Rascol Olivier, Giladi Nir, Stocchi Fabrizio, Tanner Caroline M., Postuma Ronald B., Simon David K., Tolosa Eduardo, Mollenhauer Brit, Cedarbaum Jesse M., Fraser Kyle, Xiao James, Evans Karleyton C., Graham Danielle L., et al. (2022) 'Trial of Cinpanemab in Early Parkinson's Disease', *New England Journal of Medicine*, 387(5), pp. 408–420.

Langmead, B. & Salzberg, S.L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature methods*, 9(4), pp. 357–359.

Lasda, E. & Parker, R. (2016) 'Circular RNAs Co-Precipitate with Extracellular Vesicles: A Possible Mechanism for circRNA Clearance', *PLOS ONE*, 11(2), p. e0148407.

Law, C.W., Chen, Y., Shi, W. & Smyth, G.K. (2014) 'voom: precision weights unlock linear model analysis tools for RNA-seq read counts', *Genome Biology*, 15(2), p. R29.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T. & Carey, V.J. (2013) 'Software for Computing and Annotating Genomic Ranges', *PLOS Computational Biology*, 9(8), p. e1003118.

Lee, A.J., Wang, Y., Alcalay, R.N., Mejia-Santana, H., Saunders-Pullman, R., Bressman, S., Corvol, J.-C., Brice, A., Lesage, S., Mangone, G., Tolosa, E., Pont-Sunyer, C., Vilas, D., Schüle, B., Kausar, F., Foroud, T., Berg, D., Brockmann, K., Goldwurm, S., et al. (2017) 'Penetrance estimate of LRRK2 p.G2019S mutation in individuals of non-Ashkenazi Jewish ancestry', *Movement Disorders*, 32(10), pp. 1432–1438.

Lee, H., Fenster, R.J., Pineda, S.S., Gibbs, W.S., Mohammadi, S., Davila-Velderrain, J., Garcia, F.J., Therrien, M., Novis, H.S., Gao, F., Wilkinson, H., Vogt, T., Kellis, M., LaVoie, M.J. & Heiman, M. (2020) 'Cell Type-Specific Transcriptomics Reveals that Mutant Huntingtin Leads to Mitochondrial RNA Release and Neuronal Innate Immune Activation', *Neuron*, 107(5), pp. 891-908.e8.

Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K. & Irizarry, R.A. (2010) 'Tackling the widespread and critical impact of batch effects in high-throughput data', *Nature Reviews Genetics*, 11(10), pp. 733–739.

Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., Fatica, A., Santini, T., Andronache, A., Wade, M., Laneve, P., Rajewsky, N. & Bozzoni, I. (2017) 'Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis', *Molecular Cell*, 66(1), pp. 22-37.e9.

Lerche, S., Wurster, I., Röben, B., Zimmermann, M., Machetanz, G., Wiethoff, S., Dehnert, M., Rietschel, L., Riebenbauer, B., Deuschle, C., Stransky, E., Lieplt-Scarfone, I., Gasser, T. & Brockmann, K. (2020) 'CSF NFL in a Longitudinally Assessed PD Cohort: Age Effects and Cognitive Trajectories', *Movement Disorders*, 35(7), pp. 1138–1144.

Leta, V., Urso, D., Batzu, L., Lau, Y.H., Mathew, D., Boura, I., Raeder, V., Falup-Pecurariu, C., van Wamelen, D. & Ray Chaudhuri, K. (2022) 'Viruses, parkinsonism and Parkinson's disease: the past, present and future', *Journal of Neural Transmission*, 129(9), pp. 1119–1132.

Levanon, E.Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z.Y., Shoshan, A., Pollock, S.R., Sztybel, D., Olshansky, M., Rechavi, G. & Jantsch, M.F. (2004) 'Systematic identification of abundant A-to-I editing sites in the human transcriptome', *Nature Biotechnology*, 22(8), pp. 1001–1005.

Levine, K.S., Leonard, H.L., Blauwendraat, C., Iwaki, H., Johnson, N., Bandres-Ciga, S., Ferrucci, L., Faghri, F., Singleton, A.B. & Nalls, M.A. (2023) 'Virus exposure and neurodegenerative disease risk across national biobanks', *Neuron*, 111(7), pp. 1086-1093.e2.

Lewis, M.J., Spiliopoulou, A., Goldmann, K., Pitzalis, C., McKeigue, P. & Barnes, M.R. (2023) 'nestedcv: an R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data', *Bioinformatics Advances*, 3(1), p. vbad048.

Lewis, P.A. & Cookson, M.R. (2012) 'Gene expression in the Parkinson's disease brain', *Brain Research Bulletin*, 88(4), pp. 302–312.

Li, D. & Wu, M. (2021) 'Pattern recognition receptors in health and diseases', *Signal Transduction and Targeted Therapy*, 6(1), pp. 1–24.

Li, H. & Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics (Oxford, England)*, 25(14), pp. 1754–1760.

Li, J., Mestre, T.A., Mollenhauer, B., Frasier, M., Tomlinson, J.J., Trenkwalder, C., Ramsay, T., Manuel, D. & Schlossmacher, M.G. (2022) 'Evaluation of the PREDIGT score's performance in identifying newly diagnosed Parkinson's patients without motor examination', *npj Parkinson's Disease*, 8(1), pp. 1–12.

Li, S., Teng, S., Xu, J., Su, G., Zhang, Y., Zhao, J., Zhang, S., Wang, H., Qin, W., Lu, Z.J., Guo, Y., Zhu, Q. & Wang, D. (2019) 'Microarray is an efficient tool for circRNA profiling', *Briefings in Bioinformatics*, 20(4), pp. 1420–1433.

Li, W., Shen, J., Wu, H., Lin, L., Liu, Y., Pei, Z. & Liu, G. (2023) 'Transcriptome Analysis Reveals a Two-Gene Signature Links to Motor Progression and Alterations of Immune Cells in Parkinson's Disease', *Journal of Parkinson's Disease*, 13(1), pp. 25–38.

Li, X., Ding, J., Wang, X., Cheng, Z. & Zhu, Q. (2020) 'NUDT21 regulates circRNA cyclization and ceRNA crosstalk in hepatocellular carcinoma', *Oncogene*, 39(4), pp. 891–904.

Li, X., Leung, S., Qureshi, S., Darnell, J.E. & Stark, G.R. (1996) 'Formation of STAT1-STAT2 Heterodimers and Their Role in the Activation of *IRF-1* Gene Transcription by Interferon-α(∗)', *Journal of Biological Chemistry*, 271(10), pp. 5790–5794.

Li, X., Liu, C.-X., Xue, W., Zhang, Y., Jiang, S., Yin, Q.-F., Wei, J., Yao, R.-W., Yang, L. & Chen, L.-L. (2017) 'Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection', *Molecular Cell*, 67(2), pp. 214-227.e7.

Li, X., Zhang, J.-L., Lei, Y.-N., Liu, X.-Q., Xue, W., Zhang, Y., Nan, F., Gao, X., Zhang, J., Wei, J., Yang, L. & Chen, L.-L. (2021) 'Linking circular intronic RNA degradation and function in transcription by RNase H1', *Science China Life Sciences*, 64(11), pp. 1795–1809.

Li, X.-L., Andersen, J.B., Ezelle, H.J., Wilson, G.M. & Hassel, B.A. (2007) 'Post-transcriptional Regulation of RNase-L Expression Is Mediated by the 3′-Untranslated Region of Its mRNA*', *Journal of Biological Chemistry*, 282(11), pp. 7950–7960.

Li, Y., Zheng, Q., Bao, C., Li, S., Guo, W., Zhao, J., Chen, D., Gu, J., He, X. & Huang, S. (2015) 'Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis', *Cell Research*, 25(8), pp. 981–984.

Li, Y.I., Wong, G., Humphrey, J. & Raj, T. (2019) 'Prioritizing Parkinson's disease genes using population-scale transcriptomic data', *Nature Communications*, 10(1), p. 994.

Li, Y., Zhao, J., Yu, S., Wang, Z., He, Xigan, Su, Y., Guo, T., Sheng, H., Chen, J., Zheng, Q., Li, Yan, Guo, W., Cai, X., Shi, G., Wu, J., Wang, L., Wang, P., He, Xianghuo & Huang, S. (2019) 'Extracellular Vesicles Long RNA Sequencing Reveals Abundant mRNA,

circRNA, and lncRNA in Human Blood as Potential Biomarkers for Cancer Diagnosis', *Clinical Chemistry*, 65(6), pp. 798–808.

Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., Zhong, G., Yu, B., Hu, W., Dai, L., Zhu, P., Chang, Z., Wu, Q., Zhao, Y., Jia, Y., Xu, P., Liu, H. & Shan, G. (2015) 'Exon-intron circular RNAs regulate transcription in the nucleus', *Nature Structural & Molecular Biology*, 22(3), pp. 256–264.

Liang, D., Tatomer, D.C., Luo, Z., Wu, H., Yang, L., Chen, L.-L., Cherry, S. & Wilusz, J.E. (2017) 'The Output of Protein-Coding Genes Shifts to Circular RNAs When the Pre-mRNA Processing Machinery Is Limiting', *Molecular Cell*, 68(5), pp. 940-954.e3.

Liang, D. & Wilusz, J.E. (2014) 'Short intronic repeat sequences facilitate circular RNA production', *Genes & Development*, 28(20), pp. 2233–2247.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P. & Tamayo, P. (2015) 'The Molecular Signatures Database (MSigDB) hallmark gene set collection', *Cell systems*, 1(6), pp. 417–425.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P. & Mesirov, J.P. (2011) 'Molecular signatures database (MSigDB) 3.0', *Bioinformatics*, 27(12), pp. 1739–1740.

Liddicoat, B.J., Piskol, R., Chalk, A.M., Ramaswami, G., Higuchi, M., Hartner, J.C., Li, J.B., Seeburg, P.H. & Walkley, C.R. (2015) 'RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself', *Science*, 349(6252), pp. 1115–1120.

Lin, J.-C., Lin, C.-S., Hsu, C.-W., Lin, C.-L. & Kao, C.-H. (2016) 'Association Between Parkinson's Disease and Inflammatory Bowel Disease: a Nationwide Taiwanese Retrospective Cohort Study', *Inflammatory Bowel Diseases*, 22(5), pp. 1049–1055.

Lindestam Arlehamn, C.S., Dhanwani, R., Pham, J., Kuan, R., Frazier, A., Rezende Dutra, J., Phillips, E., Mallal, S., Roederer, M., Marder, K.S., Amara, A.W., Standaert, D.G., Goldman, J.G., Litvan, I., Peters, B., Sulzer, D. & Sette, A. (2020) 'α-Synuclein-specific T cell reactivity is associated with preclinical and early Parkinson's disease', *Nature Communications*, 11(1), p. 1875.

Liu, C.-X., Guo, S.-K., Nan, F., Xu, Y.-F., Yang, L. & Chen, L.-L. (2022) 'RNA circles with minimized immunogenicity as potent PKR inhibitors', *Molecular Cell*, 82(2), pp. 420-434.e6.

Liu, C.-X., Li, X., Nan, F., Jiang, S., Gao, X., Guo, S.-K., Xue, W., Cui, Y., Dong, K., Ding, H., Qu, B., Zhou, Z., Shen, N., Yang, L. & Chen, L.-L. (2019) 'Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity', *Cell*, 177(4), pp. 865-880.e21.

Liu, G., Boot, B., Locascio, J.J., Jansen, I.E., Winder-Rhodes, S., Eberly, S., Elbaz, A., Brice, A., Ravina, B., van Hilten, J.J., Cormier-Dequaire, F., Corvol, J.-C., Barker, R.A., Heutink, P., Marinus, J., Williams-Gray, C.H., Scherzer, C.R. & Consortium, for the I.G. of P.D.P. (IGPP) (2016) 'Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's', *Annals of Neurology*, 80(5), pp. 674–685.

Liu, J., Halene, S., Yang, M., Iqbal, J., Yang, R., Mehal, W.Z., Chuang, W.-L., Jain, D., Yuen, T., Sun, L., Zaidi, M. & Mistry, P.K. (2012) 'Gaucher disease gene GBA functions in immune regulation', *Proceedings of the National Academy of Sciences*, 109(25), pp. 10018–10023.

Liu, Z., Ran, Y., Tao, C., Li, S., Chen, J. & Yang, E. (2019) 'Detection of circular RNA expression and related quantitative trait loci in the human dorsolateral prefrontal cortex', *Genome Biology*, 20p. 99.

Lo, Ij., Hill, J., Vilhjálmsson, B.J. & Kjems, J. (2020) 'Linking the association between circRNAs and Alzheimer's disease progression by multi-tissue circular RNA characterization', *RNA Biology*, 17(12), pp. 1789–1797.

Loesch, D.P., Horimoto, A.R.V.R., Heilbron, K., Sarihan, E.I., Inca-Martinez, M., Mason, E., Cornejo-Olivas, M., Torres, L., Mazzetti, P., Cosentino, C., Sarapura-Castro, E., Rivera-Valdivia, A., Medina, A.C., Dieguez, E., Raggio, V., Lescano, A., Tumas, V., Borges, V., Ferraz, H.B., et al. (2021) 'Characterizing the Genetic Architecture of Parkinson's Disease in Latinos', *Annals of Neurology*, 90(3), pp. 353–365.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., et al. (2013) 'The Genotype-Tissue Expression (GTEx) project', *Nature Genetics*, 45(6), pp. 580–585.

Lopes, K. de P., Snijders, G.J.L., Humphrey, J., Allan, A., Sneeboer, M.A.M., Navarro, E., Schilder, B.M., Vialle, R.A., Parks, M., Missall, R., van Zuiden, W., Gigase, F.A.J., Kübler, R., van Berlekom, A.B., Hicks, E.M., Böttcher, C., Priller, J., Kahn, R.S., de Witte, L.D., et al. (2022) 'Genetic analysis of the human microglial transcriptome across brain regions, aging and disease pathologies', *Nature Genetics*, 54(1), pp. 4–17.

Love, M.I., Hogenesch, J.B. & Irizarry, R.A. (2016) 'Modeling of RNA-seq fragment sequence bias reduces systematic errors in transcript abundance estimation', *Nature Biotechnology*, 34(12), pp. 1287–1291.

Love, M.I., Huber, W. & Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), p. 550.

Lowes, H., Pyle, A., Santibanez-Koref, M. & Hudson, G. (2020) 'Circulating cell-free mitochondrial DNA levels in Parkinson's disease are influenced by treatment', *Molecular Neurodegeneration*, 15.

Lu, B., Nakamura, T., Inouye, K., Li, J., Tang, Y., Lundbäck, P., Valdes-Ferrer, S.I., Olofsson, P.S., Kalb, T., Roth, J., Zou, Y., Erlandsson-Harris, H., Yang, H., Ting, J.P.-Y., Wang, H., Andersson, U., Antoine, D.J., Chavan, S.S., Hotamisligil, G.S., et al. (2012) 'Novel role of PKR in inflammasome activation and HMGB1 release', *Nature*, 488(7413), pp. 670–674.

Lu, P.D., Harding, H.P. & Ron, D. (2004) 'Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response', *Journal of Cell Biology*, 167(1), pp. 27–33.

Luk, K.C., Kehm, V., Carroll, J., Zhang, B., O'Brien, P., Trojanowski, J.Q. & Lee, V.M.-Y. (2012) 'Pathological α-Synuclein Transmission Initiates Parkinson-like Neurodegeneration in Nontransgenic Mice', *Science*, 338(6109), pp. 949–953.

Lundberg, E., Fagerberg, L., Klevebring, D., Matic, I., Geiger, T., Cox, J., Älgenäs, C., Lundeberg, J., Mann, M. & Uhlen, M. (2010) 'Defining the transcriptome and proteome in three functionally different human cell lines', *Molecular Systems Biology*, 6(1), p. 450.

Luo, Z., Rong, Z., Zhang, J., Zhu, Z., Yu, Z., Li, T., Fu, Z., Qiu, Z. & Huang, C. (2020) 'Circular RNA circCCDC9 acts as a miR-6792-3p sponge to suppress the progression of gastric cancer through regulating CAV1 expression', *Molecular Cancer*, 19(1), p. 86.

Lv, X., He, M., Zhao, Y., Zhang, L., Zhu, W., Jiang, L., Yan, Y., Fan, Y., Zhao, H., Zhou, S., Ma, H., Sun, Y., Li, X., Xu, H. & Wei, M. (2019) 'Identification of potential key genes and pathways predicting pathogenesis and prognosis for triple-negative breast cancer', *Cancer Cell International*, 19p. 172.

Ma, G., J, G., I, V., S, G., E, D., C, R. & M, E. (2006) 'Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action', *Microbiology and molecular biology reviews : MMBR*, 70(4), .

Ma, X.-K., Wang, M.-R., Liu, C.-X., Dong, R., Carmichael, G.G., Chen, L.-L. & Yang, L. (2019) 'CIRCexplorer3: A CLEAR Pipeline for Direct Comparison of Circular and Linear RNA Expression', *Genomics, Proteomics & Bioinformatics*, 17(5), pp. 511–521.

Ma, X.-K., Zhai, S.-N. & Yang, L. (2023) 'Approaches and challenges in genome-wide circular RNA identification and quantification', *Trends in Genetics*, 39(12), pp. 897–907.

Maass, P.G., Glažar, P., Memczak, S., Dittmar, G., Hollfinger, I., Schreyer, L., Sauer, A.V., Toka, O., Aiuti, A., Luft, F.C. & Rajewsky, N. (2017) 'A map of human circular RNAs in clinically relevant tissues', *Journal of Molecular Medicine*, 95(11), pp. 1179–1189.

Macleod, A.D., Taylor, K.S.M. & Counsell, C.E. (2014) 'Mortality in Parkinson's disease: A systematic review and meta-analysis', *Movement Disorders*, 29(13), pp. 1615–1622.

Magalhães, P. & Lashuel, H.A. (2022) 'Opportunities and challenges of alpha-synuclein as a potential biomarker for Parkinson's disease and other synucleinopathies', *npj Parkinson's Disease*, 8(1), pp. 1–26.

Magdalinou, N.K., Paterson, R.W., Schott, J.M., Fox, N.C., Mummery, C., Blennow, K., Bhatia, K., Morris, H.R., Giunti, P., Warner, T.T., Silva, R. de, Lees, A.J. & Zetterberg, H. (2015) 'A panel of nine cerebrospinal fluid biomarkers may identify patients with atypical parkinsonian syndromes', *Journal of Neurology, Neurosurgery & Psychiatry*, 86(11), pp. 1240–1247.

Mahlknecht, P., Seppi, K. & Poewe, W. (2015) 'The Concept of Prodromal Parkinson's Disease', *Journal of Parkinson's Disease*, 5(4), pp. 681–697.

Main, B.S., Zhang, M., Brody, K.M., Ayton, S., Frugier, T., Steer, D., Finkelstein, D., Crack, P.J. & Taylor, J.M. (2016) 'Type-1 interferons contribute to the neuroinflammatory response and disease progression of the MPTP mouse model of Parkinson's disease', *Glia*, 64(9), pp. 1590–1604.

Main, B.S., Zhang, M., Brody, K.M., Kirby, F.J., Crack, P.J. & Taylor, J.M. (2017) 'Type-I interferons mediate the neuroinflammatory response and neurotoxicity induced by rotenone', *Journal of Neurochemistry*, 141(1), pp. 75–85.

Makarious, M.B., Lake, J., Pitz, V., Ye Fu, A., Guidubaldi, J.L., Solsberg, C.W., Bandres-Ciga, S., Leonard, H.L., Kim, J.J., Billingsley, K.J., Grenn, F.P., Jerez, P.A., Alvarado, C.X., Iwaki, H., Ta, M., Vitale, D., Hernandez, D., Torkamani, A., Ryten, M., et al. (2023) 'Large-scale rare variant burden testing in Parkinson's disease', *Brain*, 146(11), pp. 4622–4632.

Makarious, M.B., Leonard, H.L., Vitale, D., Iwaki, H., Sargent, L., Dadu, A., Violich, I., Hutchins, E., Saffo, D., Bandres-Ciga, S., Kim, J.J., Song, Y., Maleknia, M., Bookman, M., Nojopranoto, W., Campbell, R.H., Hashemi, S.H., Botia, J.A., Carter, J.F., et al. (2022) 'Multi-modality machine learning predicting Parkinson's disease', *NPJ Parkinson's disease*, 8(1), p. 35.

Mallet, D., Dufourd, T., Decourt, M., Carcenac, C., Bossù, P., Verlin, L., Fernagut, P.-O., Benoit-Marand, M., Spalletta, G., Barbier, E.L., Carnicella, S., Sgambato, V., Fauvelle, F. & Boulet, S. (2022) 'A metabolic biomarker predicts Parkinson's disease at the early stages in patients and animal models', *The Journal of Clinical Investigation*, 132(4), .

Mamas, M., Dunn, W.B., Neyses, L. & Goodacre, R. (2011) 'The role of metabolites and metabolomics in clinically applicable biomarkers of disease', *Archives of Toxicology*, 85(1), pp. 5–17.

Mammana, A., Baiardi, S., Quadalti, C., Rossi, M., Donadio, V., Capellari, S., Liguori, R. & Parchi, P. (2021) 'RT-QuIC Detection of Pathological α-Synuclein in Skin Punches of Patients with Lewy Body Disease', *Movement Disorders*, 36(9), pp. 2173–2177.

Manne, S., Kondru, N., Jin, H., Serrano, G.E., Anantharam, V., Kanthasamy, A., Adler, C.H., Beach, T.G. & Kanthasamy, A.G. (2020) 'Blinded RT-QuIC Analysis of α-Synuclein Biomarker in Skin Tissue from Parkinson's Disease Patients', *Movement disorders : official journal of the Movement Disorder Society*, 35(12), pp. 2230–2239.

Mannion, N.M., Greenwood, S.M., Young, R., Cox, S., Brindle, J., Read, D., Nellåker, C., Vesely, C., Ponting, C.P., McLaughlin, P.J., Jantsch, M.F., Dorin, J., Adams, I.R., Scadden, A.D.J., Öhman, M., Keegan, L.P. & O'Connell, M.A. (2014) 'The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA', *Cell Reports*, 9(4), pp. 1482–1494.

Mansuri, S., Jain, A., Singh, R., Rawat, S., Mondal, D. & Raychaudhuri, S. (2024) 'Widespread nuclear lamina injuries defeat proteostatic purposes of α-synuclein amyloid inclusions', *Journal of Cell Science*, 137(7), p. jcs261935.

Mappin-Kasirer, B., Pan, H., Lewington, S., Kizza, J., Gray, R., Clarke, R. & Peto, R. (2020) 'Tobacco smoking and the risk of Parkinson disease', *Neurology*, 94(20), pp. e2132–e2138.

Marder, K., Wang, Y., Alcalay, R.N., Mejia-Santana, H., Tang, M.-X., Lee, A., Raymond, D., Mirelman, A., Saunders-Pullman, R., Clark, L., Ozelius, L., Orr-Urtreger, A., Giladi, N., Bressman, S., & For the LRRK2 Ashkenazi Jewish Consortium (2015) 'Age-specific penetrance of LRRK2 G2019S in the Michael J. Fox Ashkenazi Jewish LRRK2 Consortium', *Neurology*, 85(1), pp. 89–95.

Marek, K., Chowdhury, S., Siderowf, A., Lasch, S., Coffey, C.S., Caspell-Garcia, C., Simuni, T., Jennings, D., Tanner, C.M., Trojanowski, J.Q., Shaw, L.M., Seibyl, J., Schuff, N., Singleton, A., Kieburtz, K., Toga, A.W., Mollenhauer, B., Galasko, D., Chahine, L.M., et al. (2018) 'The Parkinson's progression markers initiative (PPMI) – establishing a PD biomarker cohort', *Annals of Clinical and Translational Neurology*, 5(12), pp. 1460–1477.

Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., Flagg, E., Chowdhury, S., Poewe, W., Mollenhauer, B., Klinik, P.-E., Sherer, T., Frasier, M., Meunier, C., Rudolph, A., Casaceli, C., Seibyl, J., et al. (2011) 'The Parkinson Progression Marker Initiative (PPMI)', *Progress in Neurobiology*, 95(4), pp. 629–635.

van der Mark, M., Brouwer, M., Kromhout, H., Nijssen, P., Huss, A. & Vermeulen, R. (2012) 'Is Pesticide Use Related to Parkinson Disease? Some Clues to Heterogeneity in Study Results', *Environmental Health Perspectives*, 120(3), pp. 340–347.

Marques, O. & Outeiro, T.F. (2012) 'Alpha-synuclein: from secretion to dysfunction and death', *Cell Death & Disease*, 3(7), pp. e350–e350.

Martin, A.R., Williams, E., Foulger, R.E., Leigh, S., Daugherty, L.C., Niblock, O., Leong, I.U.S., Smith, K.R., Gerasimenko, O., Haraldsdottir, E., Thomas, E., Scott, R.H., Baple, E., Tucci, A., Brittain, H., de Burca, A., Ibañez, K., Kasperaviciute, D., Smedley, D., et al. (2019) 'PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels', *Nature Genetics*, 51(11), pp. 1560–1565.

Martínez-Martín, P., Rodríguez-Blázquez, C., Mario Alvarez, Arakaki, T., Arillo, V.C., Chaná, P., Fernández, W., Garretto, N., Martínez-Castrillo, J.C., Rodríguez-Violante, M., Serrano-Dueñas, M., Ballesteros, D., Rojo-Abuin, J.M., Chaudhuri, K.R. & Merello, M. (2015) 'Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale', *Parkinsonism & Related Disorders*, 21(1), pp. 50–54.

Martin-Ruiz, C., Williams-Gray, C.H., Yarnall, A.J., Boucher, J.J., Lawson, R.A., Wijeyekoon, R.S., Barker, R.A., Kolenda, C., Parker, C., Burn, D.J., Von Zglinicki, T. & Saretzki, G. (2020) 'Senescence and Inflammatory Markers for Predicting Clinical Progression in Parkinson's Disease: The ICICLE-PD Study', *Journal of Parkinson's Disease*, 10(1), pp. 193–206.

Martirosyan, A., Ansari, R., Pestana, F., Hebestreit, K., Gasparyan, H., Aleksanyan, R., Hnatova, S., Poovathingal, S., Marneffe, C., Thal, D.R., Kottick, A., Hanson-Smith, V.J., Guelfi, S., Plumbly, W., Belgard, T.G., Metzakopian, E. & Holt, M.G. (2024) 'Unravelling cell type-specific responses to Parkinson's Disease at single cell resolution', *Molecular Neurodegeneration*, 19(1), p. 7.

Mastenbroek, S.E., Vogel, J.W., Collij, L.E., Serrano, G.E., Tremblay, C., Young, A.L., Arce, R.A., Shill, H.A., Driver-Dunckley, E.D., Mehta, S.H., Belden, C.M., Atri, A., Choudhury, P., Barkhof, F., Adler, C.H., Ossenkoppele, R., Beach, T.G. & Hansson, O. (2024) 'Disease progression modelling reveals heterogeneity in trajectories of Lewy-type α-synuclein pathology', *Nature Communications*, 15(1), p. 5133.

Matsuura, K., Maeda, M., Tabei, K., Umino, M., Kajikawa, H., Satoh, M., Kida, H. & Tomimoto, H. (2016) 'A longitudinal study of neuromelanin-sensitive magnetic resonance imaging in Parkinson's disease', *Neuroscience Letters*, 633pp. 112–117.

Mattsson, N., Lönneborg, A., Boccardi, M., Blennow, K. & Hansson, O. (2017) 'Clinical validity of cerebrospinal fluid Aβ42, tau, and phospho-tau as biomarkers for Alzheimer's disease in the context of a structured 5-phase development framework', *Neurobiology of Aging*, 52pp. 196–213.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., Voss, N., Stegmaier, P., Lewicki-Potapov, B., Saxel, H., Kel, A.E. & Wingender, E. (2006) 'TRANSFAC® and its module TRANSCompel®: transcriptional gene regulation in eukaryotes', *Nucleic Acids Research*, 34(Database issue), pp. D108–D110.

McFarthing, K., Buff, S., Rafaloff, G., Fiske, B., Mursaleen, L., Fuest, R., Wyse, R.K. & Stott, S.R.W. (2023) 'Parkinson's Disease Drug Therapies in the Clinical Trial Pipeline: 2023 Update', *Journal of Parkinson's Disease*, 13(4), pp. 427–439.

McGeer, P.L., Itagaki, S., Boyes, B.E. & McGeer, E.G. (1988) 'Reactive microglia are positive for HLA-DR in the substantia nigra of Parkinson's and Alzheimer's disease brains', *Neurology*, 38(8), pp. 1285–1285.

McNab, F., Mayer-Barber, K., Sher, A., Wack, A. & O'Garra, A. (2015) 'Type I interferons in infectious disease', *Nature Reviews Immunology*, 15(2), pp. 87–103.

Meganck, R.M., Borchardt, E.K., Rivera, R.M.C., Scalabrino, M.L., Wilusz, J.E., Marzluff, W.F. & Asokan, A. (2018) 'Tissue-Dependent Expression and Translation of Circular RNAs with Recombinant AAV Vectors In Vivo', *Molecular Therapy - Nucleic Acids*, 13pp. 89–98.

Mellough, C.B., Bauer, R., Collin, J., Dorgau, B., Zerti, D., Dolan, D.W.P., Jones, C.M., Izuogu, O.G., Yu, M., Hallam, D., Steyn, J.S., White, K., Steel, D.H., Santibanez-Koref, M., Elliott, D.J., Jackson, M.S., Lindsay, S., Grellscheid, S. & Lako, M. (2019) 'An integrated transcriptional analysis of the developing human retina', *Development*, 146(2), p. dev169474.

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., le Noble, F. & Rajewsky, N. (2013) 'Circular RNAs are a large class of animal RNAs with regulatory potency', *Nature*, 495(7441), pp. 333–338.

Memczak, S., Papavasileiou, P., Peters, O. & Rajewsky, N. (2015) 'Identification and Characterization of Circular RNAs As a New Class of Putative Biomarkers in Human Blood', *PLOS ONE*, 10(10), p. e0141214.

Merchant, K.M., Cedarbaum, J.M., Brundin, P., Dave, K.D., Eberling, J., Espay, A.J., Hutten, S.J., Javidnia, M., Luthman, J., Maetzler, W., Menalled, L., Reimer, A.N., Stoessl, A.J., Weiner, D.M. & Group, and T.M.J.F.F.A.S.C.P.W. (2019) 'A Proposed Roadmap for Parkinson's Disease Proof of Concept Clinical Trials Investigating Compounds Targeting Alpha-Synuclein', *Journal of Parkinson's Disease*, 9(1), pp. 31–61.

Merchant, K.M., Simuni, T., Fedler, J., Caspell-Garcia, C., Brumm, M., Nudelman, K.N.H., Tengstrandt, E., Hsieh, F., Alcalay, R.N., Coffey, C., Chahine, L., Foroud, T., Singleton, A., Weintraub, D., Hutten, S., Sherer, T., Mollenhauer, B., Siderowf, A., Tanner, C., et al. (2023) 'LRRK2 and GBA1 variant carriers have higher urinary bis(monacylglycerol) phosphate concentrations in PPMI cohorts', *npj Parkinson's Disease*, 9(1), pp. 1–10.

Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L., Peterson, A., Noteboom, J., O'Briant, K.C., Allen, A., Lin, D.W., Urban, N., Drescher, C.W., Knudsen, B.S., Stirewalt, D.L., Gentleman, R., Vessella, R.L., Nelson, P.S., Martin, D.B., et al. (2008) 'Circulating microRNAs as stable blood-based markers for cancer detection', *Proceedings of the National Academy of Sciences*, 105(30), pp. 10513–10518.

Moda Fabio, Gambetti Pierluigi, Notari Silvio, Concha-Marambio Luis, Catania Marcella, Park Kyung-Won, Maderna Emanuela, Suardi Silvia, Haïk Stéphane, Brandel Jean-Philippe, Ironside James, Knight Richard, Tagliavini Fabrizio, & Soto Claudio (2014) 'Prions in the Urine of Patients with Variant Creutzfeldt–Jakob Disease', *New England Journal of Medicine*, 371(6), pp. 530–539.

Mogi, M., Harada, M., Kondo, T., Riederer, P., Inagaki, H., Minami, M. & Nagatsu, T. (1994) 'Interleukin-1β, interleukin-6, epidermal growth factor and transforming growth factor-α are elevated in the brain from parkinsonian patients', *Neuroscience Letters*, 180(2), pp. 147–150.

Mogi, M., Kondo, T., Mizuno, Y. & Nagatsu, T. (2007) 'p53 protein, interferon-γ, and NF-κB levels are elevated in the parkinsonian brain', *Neuroscience Letters*, 414(1), pp. 94–97.

Moldovan, L.-I., Hansen, T.B., Venø, M.T., Okholm, T.L.H., Andersen, T.L., Hager, H., Iversen, L., Kjems, J., Johansen, C. & Kristensen, L.S. (2019) 'High-throughput RNA sequencing from paired lesional- and non-lesional skin reveals major alterations in the psoriasis circRNAome', *BMC Medical Genomics*, 12(1), p. 174.

Moldovan, L.I., Tsoi, L.C., Ranjitha, U., Hager, H., Weidinger, S., Gudjonsson, J.E., Kjems, J. & Kristensen, L.S. (2021) 'Characterization of circular RNA transcriptomes in psoriasis and atopic dermatitis reveals disease-specific expression profiles', *Experimental Dermatology*, 30(8), pp. 1187–1196.

Mollenhauer, B., Bowman, F.D., Drake, D., Duong, J., Blennow, K., El-Agnaf, O., Shaw, L.M., Masucci, J., Taylor, P., Umek, R.M., Dunty, J.M., Smith, C.L., Stoops, E., Vanderstichele, H., Schmid, A.W., Moniatte, M., Zhang, J., Kruse, N., Lashuel, H.A., et al. (2019) 'Antibody-based methods for the measurement of α-synuclein concentration in human cerebrospinal fluid - method comparison and round robin study', *Journal of Neurochemistry*, 149(1), pp. 126–138.

Molochnikov, L., Rabey, J.M., Dobronevsky, E., Bonuccelli, U., Ceravolo, R., Frosini, D., Grünblatt, E., Riederer, P., Jacob, C., Aharon-Peretz, J., Bashenko, Y., Youdim, M.B. & Mandel, S.A. (2012) 'A molecular signature in blood identifies early Parkinson's disease', *Molecular Neurodegeneration*, 7(1), p. 26.

Monaco, G., Lee, B., Xu, W., Mustafah, S., Hwang, Y.Y., Carré, C., Burdin, N., Visan, L., Ceccarelli, M., Poidinger, M., Zippelius, A., Pedro de Magalhães, J. & Larbi, A. (2019) 'RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types', *Cell Reports*, 26(6), pp. 1627-1640.e7.

Monogue, B., Chen, Y., Sparks, H., Behbehani, R., Chai, A., Rajic, A.J., Massey, A., Kleinschmidt-Demasters, B.K., Vermeren, M., Kunath, T. & Beckham, J.D. (2022) 'Alpha-synuclein supports type 1 interferon signalling in neurons and brain tissue', *Brain*, 145(10), pp. 3622–3636.

Morris, H.R., Spillantini, M.G., Sue, C.M. & Williams-Gray, C.H. (2024) 'The pathogenesis of Parkinson's disease', *The Lancet*, 403(10423), pp. 293–304.

Mouton-Liger, F., Paquet, C., Dumurgier, J., Lapalus, P., Gray, F., Laplanche, J.-L. & Hugon, J. (2012) 'Increased Cerebrospinal Fluid Levels of Double-Stranded RNA-Dependant Protein Kinase in Alzheimer's Disease', *Biological Psychiatry*, 71(9), pp. 829–835.

Muñoz-Delgado, L., Macías-García, D., Jesús, S., Martín-Rodríguez, J.F., Labrador-Espinosa, M.Á., Jiménez-Jaraba, M.V., Adarmes-Gómez, A., Carrillo, F. & Mir, P. (2021) 'Peripheral Immune Profile and Neutrophil-to-Lymphocyte Ratio in Parkinson's Disease', *Movement Disorders*, 36(10), pp. 2426–2430.

Muñoz-Delgado, L., Macías-García, D., Periñán, M.T., Jesús, S., Adarmes-Gómez, A.D., Bonilla Toribio, M., Buiza Rueda, D., Jiménez-Jaraba, M. del V., Benítez Zamora, B., Díaz Belloso, R., García-Díaz, S., Martín-Bórnez, M., Pineda Sánchez, R., Carrillo, F., Gómez-Garre, P. & Mir, P. (2023) 'Peripheral inflammatory immune response differs among sporadic and familial Parkinson's disease', *npj Parkinson's Disease*, 9(1), pp. 1–9.

Naaldijk, Y., Fernández, B., Fasiczka, R., Fdez, E., Leghay, C., Croitoru, I., Kwok, J.B., Boulesnane, Y., Vizeneux, A., Mutez, E., Calvez, C., Destée, A., Taymans, J.-M., Aragon, A.V., Yarza, A.B., Padmanabhan, S., Delgado, M., Alcalay, R.N., Chatterton, Z., et al. (2024) 'A potential patient stratification biomarker for Parkinson´s disease based on LRRK2 kinase-mediated centrosomal alterations in peripheral blood-derived cells', *npj Parkinson's Disease*, 10(1), pp. 1–13.

Nadel, B.B., Oliva, M., Shou, B.L., Mitchell, K., Ma, F., Montoya, D.J., Mouton, A., Kim-Hellmuth, S., Stranger, B.E., Pellegrini, M. & Mangul, S. (2021) 'Systematic evaluation of transcriptomics-based deconvolution methods and references using thousands of clinical samples', *Briefings in Bioinformatics*, 22(6), p. bbab265.

Nalls, M., Blauwendraat, C., Vallerga, C.L., Heilbron, K., Bandres-Ciga, S., Chang, D., Tan, M., Kia, D.A., Noyce, A.J., Xue, A., Bras, J., Young, E., von Coelln, R., Simón-Sánchez, J., Schulte, C., Sharma, M., Krohn, L., Pihlstrøm, L., Siitonen, A., et al. (2019) 'Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies', *The Lancet Neurology*, 18(12), pp. 1091–1102.

Nalls, M.A., Keller, M.F., Hernandez, D.G., Chen, L., Stone, D.J., Singleton, A.B. & Investigators, on behalf of the P.P.M.I. (PPMI) (2016) 'Baseline genetic associations in the Parkinson's Progression Markers Initiative (PPMI)', *Movement Disorders*, 31(1), pp. 79–85.

Nalls, M.A., McLean, C.Y., Rick, J., Eberly, S., Hutten, S.J., Gwinn, K., Sutherland, M., Martinez, M., Heutink, P., Williams, N.M., Hardy, J., Gasser, T., Brice, A., Price, T.R., Nicolas, A., Keller, M.F., Molony, C., Gibbs, J.R., Chen-Plotkin, A., et al. (2015) 'Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: a population-based modelling study', *The Lancet Neurology*, 14(10), pp. 1002–1009.

Nandhagopal, R., Kuramoto, L., Schulzer, M., Mak, E., Cragg, J., Lee, C.S., McKenzie, J., McCormick, S., Samii, A., Troiano, A., Ruth, T.J., Sossi, V., de la Fuente-Fernandez, R., Calne, D.B. & Stoessl, A.J. (2009) 'Longitudinal progression of sporadic

Parkinson's disease: a multi-tracer positron emission tomography study', *Brain*, 132(11), pp. 2970–2979.

Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L. & Chertkow, H. (2005) 'The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment', *Journal of the American Geriatrics Society*, 53(4), pp. 695–699.

Navarro, E., Udine, E., de Paiva Lopes, K., Parks, M., Riboldi, G., Schilder, B.M., Humphrey, J., Snijders, G.J.L., Vialle, R.A., Zhuang, M., Sikder, T., Argyrou, C., Allan, A., Chao, M.J., Farrell, K., Henderson, B., Simon, S., Raymond, D., Elango, S., et al. (2021) 'Dysregulation of mitochondrial and proteolysosomal genes in Parkinson's disease myeloid cells', *Nature aging*, 1(9), pp. 850–863.

Neumann, J., Bras, J., Deas, E., O'Sullivan, S.S., Parkkinen, L., Lachmann, R.H., Li, A., Holton, J., Guerreiro, R., Paudel, R., Segarane, B., Singleton, A., Lees, A., Hardy, J., Houlden, H., Revesz, T. & Wood, N.W. (2009) 'Glucocerebrosidase mutations in clinical and pathologically proven Parkinson's disease', *Brain*, 132(7), pp. 1783–1794.

Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M. & Alizadeh, A.A. (2015) 'Robust enumeration of cell subsets from tissue expression profiles', *Nature Methods*, 12(5), pp. 453–457.

Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., Diehn, M. & Alizadeh, A.A. (2019) 'Determining cell type abundance and expression from bulk tissues with digital cytometry', *Nature Biotechnology*, 37(7), pp. 773–782.

Ng, A.Y.E., Chan, S.N. & Pek, J.W. (2024) 'Genetic compensation between ribosomal protein paralogs mediated by a cognate circular RNA', *Cell Reports*, 43(5), p. 114228.

Ngo, K.J., Paul, K.C., Wong, D., Kusters, C.D.J., Bronstein, J.M., Ritz, B. & Fogel, B.L. (2024) 'Lysosomal genes contribute to Parkinson's disease near agriculture with high intensity pesticide use', *npj Parkinson's Disease*, 10(1), pp. 1–10.

Ngo, L.H., Bert, A.G., Dredge, B.K., Williams, T., Murphy, V., Li, W., Hamilton, W.B., Carey, K.T., Toubia, J., Pillman, K.A., Liu, D., Desogus, J., Chao, J.A., Deans, A.J., Goodall, G.J. & Wickramasinghe, V.O. (2024) 'Nuclear export of circular RNA', *Nature*, pp. 1–9.

Nguyen, A.P.T., Tsika, E., Kelly, K., Levine, N., Chen, X., West, A.B., Boularand, S., Barneoud, P. & Moore, D.J. (2020) 'Dopaminergic neurodegeneration induced by Parkinson's disease-linked G2019S LRRK2 is dependent on kinase and GTPase activity', *Proceedings of the National Academy of Sciences*, 117(29), pp. 17296–17307.

Nichols, W.C., Pankratz, N., Marek, D.K., Pauciulo, M.W., Elsaesser, V.E., Halter, C.A., Rudolph, A., Wojcieszek, J., Pfeiffer, R.F. & Foroud, T. (2009) 'Mutations in GBA are associated with familial Parkinson disease susceptibility and age at onset', *Neurology*, 72(4), pp. 310–316.

Nicolet, B.P., Engels, S., Aglialoro, F., van den Akker, E., von Lindern, M. & Wolkers, M.C. (2018) 'Circular RNA expression in human hematopoietic cells is widespread and cell-type specific', *Nucleic Acids Research*, 46(16), pp. 8168–8180.

Nielsen, A.F., Bindereif, A., Bozzoni, I., Hanan, M., Hansen, T.B., Irimia, M., Kadener, S., Kristensen, L.S., Legnini, I., Morlando, M., Jarlstad Olesen, M.T., Pasterkamp, R.J., Preibisch, S., Rajewsky, N., Suenkel, C. & Kjems, J. (2022) 'Best practice standards for circular RNA research', *Nature Methods*, pp. 1–13.

Nielsen, T., Jensen, M.B., Stenager, E. & Andersen, A.D. (2018) 'The use of olfactory testing when diagnosing Parkinson's disease - a systematic review', *Danish Medical Journal*, 65(5), p. A5481.

Nigro, J.M., Cho, K.R., Fearon, E.R., Kern, S.E., Ruppert, J.M., Oliner, J.D., Kinzler, K.W. & Vogelstein, B. (1991) 'Scrambled exons', *Cell*, 64(3), pp. 607–613.

Nishikura, K. (2016) 'A-to-I editing of coding and non-coding RNAs by ADARs', *Nature Reviews Molecular Cell Biology*, 17(2), pp. 83–96.

Noori, A., Mezlini, A.M., Hyman, B.T., Serrano-Pozo, A. & Das, S. (2021) 'Systematic review and meta-analysis of human transcriptomics reveals neuroinflammation, deficient energy metabolism, and proteostasis failure across neurodegeneration', *Neurobiology of Disease*, 149p. 105225.

Noyce, A.J., Bestwick, J.P., Silveira-Moriyama, L., Hawkes, C.H., Giovannoni, G., Lees, A.J. & Schrag, A. (2012) 'Meta-analysis of early nonmotor features and risk factors for Parkinson disease', *Annals of Neurology*, 72(6), pp. 893–901.

Ochoa, E., Ramirez, P., Gonzalez, E., De Mange, J., Ray, W.J., Bieniek, K.F. & Frost, B. (2023) 'Pathogenic tau–induced transposable element–derived dsRNA drives neuroinflammation', *Science Advances*, 9(1), p. eabq5423.

Okuzumi, A., Hatano, T., Matsumoto, G., Nojiri, S., Ueno, S., Imamichi-Tatano, Y., Kimura, H., Kakuta, S., Kondo, A., Fukuhara, T., Li, Y., Funayama, M., Saiki, S., Taniguchi, D., Tsunemi, T., McIntyre, D., Gérardy, J.-J., Mittelbronn, M., Kruger, R., et al. (2023) 'Propagative α-synuclein seeds as serum biomarkers for synucleinopathies', *Nature Medicine*, 29(6), pp. 1448–1455.

Olgiati, S., Quadri, M., Fang, M., Rood, J.P.M.A., Saute, J.A., Chien, H.F., Bouwkamp, C.G., Graafland, J., Minneboo, M., Breedveld, G.J., Zhang, J., International Parkinsonism Genetics Network, Verheijen, F.W., Boon, A.J.W., Kievit, A.J.A., Jardim, L.B., Mandemakers, W., Barbosa, E.R., Rieder, C.R.M., et al. (2016) 'DNAJC6 Mutations Associated With Early-Onset Parkinson's Disease', *Annals of Neurology*, 79(2), pp. 244–256.

Olsen, L.K., Dowd, E. & McKernan, D.P. (2018) 'A role for viral infections in Parkinson's etiology?', *Neuronal Signaling*, 2(2), p. NS20170166.

Orimo, S., Takahashi, A., Uchihara, T., Mori, F., Kakita, A., Wakabayashi, K. & Takahashi, H. (2007) 'Degeneration of Cardiac Sympathetic Nerve Begins in the Early Disease Process of Parkinson's Disease', *Brain Pathology*, 17(1), pp. 24–30.

Ortega, R.A., Wang, C., Raymond, D., Bryant, N., Scherzer, C.R., Thaler, A., Alcalay, R.N., West, A.B., Mirelman, A., Kuras, Y., Marder, K.S., Giladi, N., Ozelius, L.J., Bressman, S.B. & Saunders-Pullman, R. (2021) 'Association of Dual LRRK2 G2019S and GBA Variations With Parkinson Disease Progression', *JAMA Network Open*, 4(4), p. e215845.

Oueslati, A. (2016) ‘Implication of Alpha-Synuclein Phosphorylation at S129 in Synucleinopathies: What Have We Learned in the Last Decade?’, *Journal of Parkinson’s Disease*, 6(1), pp. 39–51.

Ozato, K., Tailor, P. & Kubota, T. (2007) ‘The Interferon Regulatory Factor Family in Host Defense: Mechanism of Action*’, *Journal of Biological Chemistry*, 282(28), pp. 20065–20069.

Paccalin, M., Pain-Barc, S., Pluchon, C., Paul, C., Besson, M.-N., Carret-Rebillat, A.-S., Rioux-Bilan, A., Gil, R. & Hugon, J. (2006) ‘Activated mTOR and PKR Kinases in Lymphocytes Correlate with Memory and Cognitive Decline in Alzheimer’s Disease’, *Dementia and Geriatric Cognitive Disorders*, 22(4), pp. 320–326.

Padiath, Q.S., Saigoh, K., Schiffmann, R., Asahara, H., Yamada, T., Koeppen, A., Hogan, K., Ptáček, L.J. & Fu, Y.-H. (2006) ‘Lamin B1 duplications cause autosomal dominant leukodystrophy’, *Nature Genetics*, 38(10), pp. 1114–1123.

Pagano, G., Niccolini, F. & Politis, M. (2016) ‘Imaging in Parkinson’s disease’, *Clinical Medicine*, 16(4), pp. 371–375.

Pagano, G., Taylor, K.I., Anzures Cabrera, J., Simuni, T., Marek, K., Postuma, R.B., Pavese, N., Stocchi, F., Brockmann, K., Svoboda, H., Trundell, D., Monnet, A., Doody, R., Fontoura, P., Kerchner, G.A., Brundin, P., Nikolcheva, T. & Bonni, A. (2024) ‘Prasinezumab slows motor progression in rapidly progressing early-stage Parkinson’s disease’, *Nature Medicine*, 30(4), pp. 1096–1103.

Pagano Gennaro, Taylor Kirsten I., Anzures-Cabrera Judith, Marchesi Maddalena, Simuni Tanya, Marek Kenneth, Postuma Ronald B., Pavese Nicola, Stocchi Fabrizio, Azulay Jean-Philippe, Mollenhauer Brit, López-Manzanares Lydia, Russell David S., Boyd James T., Nicholas Anthony P., Luquin María R., Hauser Robert A., Gasser Thomas, Poewe Werner, et al. (2022) ‘Trial of Prasinezumab in Early-Stage Parkinson’s Disease’, *New England Journal of Medicine*, 387(5), pp. 421–432.

Pain, S., Deguil, J., Belin, J., Barraud, P., Ragot, S. & Houeto, J.-L. (2020) ‘Regulation of Protein Synthesis and Apoptosis in Lymphocytes of Parkinson Patients: The Effect of Dopaminergic Treatment’, *Neurodegenerative Diseases*, 19(5–6), pp. 178–183.

Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A. & Gorman, A.M. (2016) ‘The integrated stress response’, *EMBO reports*, 17(10), pp. 1374–1395.

Pamudurti, N.R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., Hanan, M., Wyler, E., Perez-Hernandez, D., Ramberger, E., Shenzis, S., Samson, M., Dittmar, G., Landthaler, M., Chekulaeva, M., Rajewsky, N. & Kadener, S. (2017) ‘Translation of CircRNAs’, *Molecular Cell*, 66(1), pp. 9-21.e7.

Pan, H., Liu, Z., Ma, J., Li, Y., Zhao, Y., Zhou, Xiaoxia, Xiang, Y., Wang, Y., Zhou, Xun, He, R., Xie, Y., Zhou, Q., Yuan, K., Xu, Q., Sun, Q., Wang, J., Yan, X., Zhang, H., Wang, C., et al. (2023) ‘Genome-wide association study using whole-genome sequencing identifies risk loci for Parkinson’s disease in Chinese population’, *npj Parkinson’s Disease*, 9(1), pp. 1–11.

Panda, A.C. (2018) ‘Circular RNAs Act as miRNA Sponges’, *Advances in Experimental Medicine and Biology*, 1087pp. 67–79.

Panda, A.C., De, S., Grammatikakis, I., Munk, R., Yang, X., Piao, Y., Dudekula, D.B., Abdelmohsen, K. & Gorospe, M. (2017) 'High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs', *Nucleic Acids Research*, 45(12), p. e116.

Pantaleo, E., Monaco, A., Amoroso, N., Lombardi, A., Bellantuono, L., Urso, D., Lo Giudice, C., Picardi, E., Tafuri, B., Nigro, S., Pesole, G., Tangaro, S., Logroscino, G. & Bellotti, R. (2022) 'A Machine Learning Approach to Parkinson's Disease Blood Transcriptomics', *Genes*, 13(5), p. 727.

Park, O.H., Ha, H., Lee, Y., Boo, S.H., Kwon, D.H., Song, H.K. & Kim, Y.K. (2019) 'Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex', *Molecular Cell*, 74(3), pp. 494-507.e8.

Parkinson, J. (2002) 'An Essay on the Shaking Palsy', *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14(2), pp. 223–236.

Parlar, S.C., Grenn, F.P., Kim, J.J., Baluwendraat, C. & Gan-Or, Z. (2023) 'Classification of GBA1 Variants in Parkinson's Disease: The GBA1-PD Browser', *Movement Disorders*, 38(3), pp. 489–495.

Parnetti, L., Gaetani, L., Eusebi, P., Paciotti, S., Hansson, O., El-Agnaf, O., Mollenhauer, B., Blennow, K. & Calabresi, P. (2019) 'CSF and blood biomarkers for Parkinson's disease', *The Lancet Neurology*, 18(6), pp. 573–586.

Paslawski, W., Khosousi, S., Hertz, E., Markaki, I., Boxer, A. & Svenningsson, P. (2023) 'Large-scale proximity extension assay reveals CSF midkine and DOPA decarboxylase as supportive diagnostic biomarkers for Parkinson's disease', *Translational Neurodegeneration*, 12(1), p. 42.

Patiño, C., Haenni, A.-L. & Urcuqui-Inchima, S. (2015) 'NF90 isoforms, a new family of cellular proteins involved in viral replication?', *Biochimie*, 108pp. 20–24.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. (2017) 'Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference', *Nature methods*, 14(4), pp. 417–419.

Patterson, J.B. & Samuel, C.E. (1995) 'Expression and Regulation by Interferon of a Double-Stranded-RNA-Specific Adenosine Deaminase from Human Cells: Evidence for Two Forms of the Deaminase', *Molecular and Cellular Biology*, 15(10), pp. 5376–5388.

Paul, K.C., Krolewski, R.C., Lucumi Moreno, E., Blank, J., Holton, K.M., Ahfeldt, T., Furlong, M., Yu, Y., Cockburn, M., Thompson, L.K., Kreymerman, A., Ricci-Blair, E.M., Li, Y.J., Patel, H.B., Lee, R.T., Bronstein, J., Rubin, L.L., Khurana, V. & Ritz, B. (2023) 'A pesticide and iPSC dopaminergic neuron screen identifies and classifies Parkinson-relevant pesticides', *Nature Communications*, 14(1), p. 2803.

Peel, A.L. & Bredesen, D.E. (2003) 'Activation of the cell stress kinase PKR in Alzheimer's disease and human amyloid precursor protein transgenic mice', *Neurobiology of Disease*, 14(1), pp. 52–62.

Pereira, J.B., Kumar, A., Hall, S., Palmqvist, S., Stomrud, E., Bali, D., Parchi, P., Mattsson-Carlgren, N., Janelidze, S. & Hansson, O. (2023) 'DOPA decarboxylase is an emerging

biomarker for Parkinsonian disorders including preclinical Lewy body disease', *Nature Aging*, pp. 1–9.

Phipson, B., Lee, S., Majewski, I.J., Alexander, W.S. & Smyth, G.K. (2016) 'Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression', *The Annals of Applied Statistics*, 10(2), pp. 946–963.

Picardi, E. & Pesole, G. (2013) 'REDItools: high-throughput RNA editing detection made easy', *Bioinformatics*, 29(14), pp. 1813–1814.

Piccini, P., Burn, D.J., Ceravolo, R., Maraganore, D. & Brooks, D.J. (1999) 'The role of inheritance in sporadic Parkinson's disease: evidence from a longitudinal study of dopaminergic function in twins', *Annals of Neurology*, 45(5), pp. 577–582.

Pigott, K., Rick, J., Xie, S.X., Hurtig, H., Chen-Plotkin, A., Duda, J.E., Morley, J.F., Chahine, L.M., Dahodwala, N., Akhtar, R.S., Siderowf, A., Trojanowski, J.Q. & Weintraub, D. (2015) 'Longitudinal study of normal cognition in Parkinson disease', *Neurology*, 85(15), pp. 1276–1282.

Pike, S.C., Havrda, M., Gilli, F., Zhang, Z. & Salas, L.A. (2024) 'Immunological shifts during early-stage Parkinson's disease identified with DNA methylation data on longitudinally collected blood samples', *npj Parkinson's Disease*, 10(1), pp. 1–16.

Pitz, V., Makarious, M.B., Bandres-Ciga, S., Iwaki, H., Singleton, A.B., Nalls, M., Heilbron, K. & Blauwendraat, C. (2024) 'Analysis of rare Parkinson's disease variants in millions of people', *npj Parkinson's Disease*, 10(1), pp. 1–10.

Pitz, V., Malek, N., Tobias, E.S., Grosset, K.A., Gentleman, S. & Grosset, D.G. (2020) 'The Levodopa Response Varies in Pathologically Confirmed Parkinson's Disease: A Systematic Review', *Movement Disorders Clinical Practice*, 7(2), pp. 218–222.

Piwecka, M., Glažar, P., Hernandez-Miranda, L.R., Memczak, S., Wolf, S.A., Rybak-Wolf, A., Filipchyk, A., Klironomos, F., Cerda Jara, C.A., Fenske, P., Trimbuch, T., Zywitza, V., Plass, M., Schreyer, L., Ayoub, S., Kocks, C., Kühn, R., Rosenmund, C., Birchmeier, C., et al. (2017) 'Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function', *Science*, 357(6357), p. eaam8526.

Poewe, W., Seppi, K., Tanner, C.M., Halliday, G.M., Brundin, P., Volkmann, J., Schrag, A.-E. & Lang, A.E. (2017) 'Parkinson disease', *Nature Reviews Disease Primers*, 3(1), pp. 1–21.

Politis, M. & Loane, C. (2011) 'Serotonergic Dysfunction in Parkinson's Disease and Its Relevance to Disability', *TheScientificWorldJOURNAL*, 11pp. 1726–1734.

Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E.S., Chandrasekharappa, S., Athanassiadou, A., Papapetropoulos, T., Johnson, W.G., Lazzarini, A.M., Duvoisin, R.C., Di Iorio, G., Golbe, L.I., et al. (1997) 'Mutation in the α-Synuclein Gene Identified in Families with Parkinson's Disease', *Science*, 276(5321), pp. 2045–2047.

Pont-Sunyer, C., Tolosa, E., Caspell-Garcia, C., Coffey, C., Alcalay, R.N., Chan, P., Duda, J.E., Facheris, M., Fernández-Santiago, R., Marek, K., Lomeña, F., Marras, C., Mondragon, E., Saunders-Pullman, R., Waro, B. & Consortium, on behalf of L.C. (2017) 'The

prodromal phase of leucine-rich repeat kinase 2–associated Parkinson disease: Clinical and imaging Studies', *Movement Disorders*, 32(5), pp. 726–738.

Postuma, R.B. (2014) 'Prodromal Parkinson's disease – Using REM sleep behavior disorder as a window', *Parkinsonism & Related Disorders*, 20pp. S1–S4.

Postuma, R.B., Berg, D., Stern, M., Poewe, W., Olanow, C.W., Oertel, W., Obeso, J., Marek, K., Litvan, I., Lang, A.E., Halliday, G., Goetz, C.G., Gasser, T., Dubois, B., Chan, P., Bloem, B.R., Adler, C.H. & Deuschl, G. (2015) 'MDS clinical diagnostic criteria for Parkinson's disease', *Movement Disorders: Official Journal of the Movement Disorder Society*, 30(12), pp. 1591–1601.

Postuma, R.B., Gagnon, J.-F., Bertrand, J.-A., Génier Marchand, D. & Montplaisir, J.Y. (2015) 'Parkinson risk in idiopathic REM sleep behavior disorder', *Neurology*, 84(11), pp. 1104–1113.

Postuma, R.B., Gagnon, J.F., Vendette, M., Fantini, M.L., Massicotte-Marquez, J. & Montplaisir, J. (2009) 'Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder', *Neurology*, 72(15), pp. 1296–1300.

Postuma, R.B., Iranzo, A., Hu, M., Högl, B., Boeve, B.F., Manni, R., Oertel, W.H., Arnulf, I., Ferini-Strambi, L., Puligheddu, M., Antelmi, E., Cochen De Cock, V., Arnaldi, D., Mollenhauer, B., Videnovic, A., Sonka, K., Jung, K.-Y., Kunz, D., Dauvilliers, Y., et al. (2019) 'Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study', *Brain*, 142(3), pp. 744–759.

Pozdyshev, D.V., Zharikova, A.A., Medvedeva, M.V. & Muronetz, V.I. (2022) 'Differential Analysis of A-to-I mRNA Edited Sites in Parkinson's Disease', *Genes*, 13(1), p. 14.

Preece, P. & Cairns, N.J. (2003) 'Quantifying mRNA in postmortem human brain: influence of gender, age at death, postmortem interval, brain pH, agonal state and inter-lobe mRNA variance', *Molecular Brain Research*, 118(1), pp. 60–71.

Pullirsch, D. & Jantsch, M.F. (2010) 'Proteome diversification by adenosine to inosine RNA-editing', *RNA Biology*, 7(2), pp. 205–212.

Puri, S., Hu, J., Sun, Z., Lin, M., Stein, T.D., Farrer, L.A., Wolozin, B. & Zhang, X. (2023) 'Identification of circRNAs linked to Alzheimer's disease and related dementias', *Alzheimer's & Dementia*, 19(8), pp. 3389–3405.

Puttaraju, M. & Been, M. (1992) 'Group I permuted intron-exon (PIE) sequences self-splice to produce circular exons', *Nucleic Acids Research*, 20(20), pp. 5357–5364.

Qin, X.-Y., Zhang, S.-P., Cao, C., Loh, Y.P. & Cheng, Y. (2016) 'Aberrations in Peripheral Inflammatory Cytokine Levels in Parkinson Disease: A Systematic Review and Meta-analysis', *JAMA Neurology*, 73(11), pp. 1316–1324.

Qu, L., Yi, Z., Shen, Y., Lin, L., Chen, F., Xu, Y., Wu, Z., Tang, H., Zhang, X., Tian, F., Wang, C., Xiao, X., Dong, X., Guo, L., Lu, S., Yang, C., Tang, C., Yang, Y., Yu, W., et al. (2022) 'Circular RNA vaccines against SARS-CoV-2 and emerging variants', *Cell*, 185(10), pp. 1728-1744.e16.

Qu, Y., Li, Jiangting, Qin, Q., Wang, D., Zhao, J., An, K., Mao, Z., Min, Z., Xiong, Y., Li, Jingyi & Xue, Z. (2023) 'A systematic review and meta-analysis of inflammatory biomarkers in Parkinson's disease', *npj Parkinson's Disease*, 9(1), pp. 1–14.

Quan, P., Li, X., Si, Y., Sun, L., Ding, F.F., Fan, Y., Liu, H., Wei, C., Li, R., Zhao, X., Yang, F. & Yao, L. (2024) 'Single cell analysis reveals the roles and regulatory mechanisms of type-I interferons in Parkinson's disease', *Cell Communication and Signaling*, 22(1), p. 212.

Ragan, C., Goodall, G.J., Shirokikh, N.E. & Preiss, T. (2019) 'Insights into the biogenesis and potential functions of exonic circular RNA', *Scientific Reports*, 9p. 2048.

Rajan, R., Divya, K.P., Kandadai, R.M., Yadav, R., Satagopam, V.P., Madhusoodanan, U.K., Agarwal, P., Kumar, N., Ferreira, T., Kumar, H., Sreeram Prasad, A.V., Shetty, K., Mehta, S., Desai, S., Kumar, S., Prashanth, L.K., Bhatt, M., Wadia, P., Ramalingam, S., et al. (2020) 'Genetic Architecture of Parkinson's Disease in the Indian Population: Harnessing Genetic Diversity to Address Critical Gaps in Parkinson's Disease Research', *Frontiers in Neurology*, 11.

Rajput, A.H., Voll, A., Rajput, M.L., Robinson, C.A. & Rajput, A. (2009) 'Course in Parkinson disease subtypes', *Neurology*, 73(3), pp. 206–212.

Ramaker, C., Marinus, J., Stiggelbout, A.M. & van Hilten, B.J. (2002) 'Systematic evaluation of rating scales for impairment and disability in Parkinson's disease', *Movement Disorders*, 17(5), pp. 867–876.

Ramos-Zaldívar, H.M., Polakovicova, I., Salas-Huenuleo, E., Corvalán, A.H., Kogan, M.J., Yefi, C.P. & Andia, M.E. (2022) 'Extracellular vesicles through the blood–brain barrier: a review', *Fluids and Barriers of the CNS*, 19(1), p. 60.

Rana, H.Q., Balwani, M., Bier, L. & Alcalay, R.N. (2013) 'Age-specific Parkinson disease risk in GBA mutation carriers: information for genetic counseling', *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 15(2), pp. 146–149.

Ravanidis, S., Bougea, A., Karampatsi, D., Papagiannakis, N., Maniati, M., Stefanis, L. & Doxakis, E. (2021) 'Differentially Expressed Circular RNAs in Peripheral Blood Mononuclear Cells of Patients with Parkinson's Disease', *Movement Disorders: Official Journal of the Movement Disorder Society*,

Rebolledo, C., Silva, J.P., Saavedra, N. & Maracaja-Coutinho, V. (2023) 'Computational approaches for circRNAs prediction and in silico characterization', *Briefings in Bioinformatics*, 24(3), p. bbad154.

Reed, X., Bandrés-Ciga, S., Blauwendraat, C. & Cookson, M.R. (2019) 'The role of monogenic genes in idiopathic Parkinson's disease', *Neurobiology of disease*, 124pp. 230–239.

Reeve, A., Simcox, E. & Turnbull, D. (2014) 'Ageing and Parkinson's disease: Why is advancing age the biggest risk factor?', *Ageing Research Reviews*, 14(100), pp. 19–30.

Reimer, L., Gram, H., Jensen, N.M., Betzer, C., Yang, L., Jin, L., Shi, M., Boudeffa, D., Fusco, G., De Simone, A., Kirik, D., Lashuel, H.A., Zhang, J. & Jensen, P.H. (2022) 'Protein kinase R dependent phosphorylation of α-synuclein regulates its membrane binding and aggregation', *PNAS Nexus*, 1(5), p. pgac259.

Reimer, L., Vesterager, L.B., Betzer, C., Zheng, J., Nielsen, L.D., Kofoed, R.H., Lassen, L.B., Bølcho, U., Paludan, S.R., Fog, K. & Jensen, P.H. (2018) 'Inflammation kinase PKR phosphorylates α-synuclein and causes α-synuclein-dependent cell death', *Neurobiology of Disease*, 115pp. 17–28.

Reynoso, A., Torricelli, R., Jacobs, B.M., Shi, J., Aslibekyan, S., Norcliffe-Kaufmann, L., Noyce, A.J. & Heilbron, K. (2024) 'Gene–Environment Interactions for Parkinson's Disease', *Annals of Neurology*, 95(4), pp. 677–687.

Riboldi, G.M., Vialle, R.A., Navarro, E., Udine, E., de Paiva Lopes, K., Humphrey, J., Allan, A., Parks, M., Henderson, B., Astudillo, K., Argyrou, C., Zhuang, M., Sikder, T., Oriol Narcis, J., Kumar, S.D., Janssen, W., Sowa, A., Comi, G.P., Di Fonzo, A., et al. (2022) 'Transcriptome deregulation of peripheral monocytes and whole blood in GBA-related Parkinson's disease', *Molecular Neurodegeneration*, 17(1), p. 52.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K. & Rehm, H.L. (2015) 'Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology', *Genetics in Medicine*, 17(5), pp. 405–423.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. & Smyth, G.K. (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*, 43(7), p. e47.

Rizig, M., Bandres-Ciga, S., Makarious, M.B., Ojo, O.O., Crea, P.W., Abiodun, O.V., Levine, K.S., Abubakar, S.A., Achoru, C.O., Vitale, D., Adeniji, O.A., Agabi, O.P., Koretsky, M.J., Agulanna, U., Hall, D.A., Akinyemi, R.O., Xie, T., Ali, M.W., Shamim, E.A., et al. (2023) 'Identification of genetic risk loci and causal insights associated with Parkinson's disease in African and African admixed populations: a genome-wide association study', *The Lancet. Neurology*, 22(11), pp. 1015–1025.

Rizzo, G., Copetti, M., Arcuti, S., Martino, D., Fontana, A. & Logroscino, G. (2016) 'Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis', *Neurology*, 86(6), pp. 566–576.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L. & Pachter, L. (2011) 'Improving RNA-Seq expression estimates by correcting for fragment bias', *Genome Biology*, 12(3), p. R22.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C. & Müller, M. (2011) 'pROC: an open-source package for R and S+ to analyze and compare ROC curves', *BMC Bioinformatics*, 12(1), p. 77.

Robinson, M.D., McCarthy, D.J. & Smyth, G.K. (2010) 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics*, 26(1), pp. 139–140.

Robinson, M.D. & Oshlack, A. (2010) 'A scaling normalization method for differential expression analysis of RNA-seq data', *Genome Biology*, 11(3), p. R25.

Rodriguez, S., Sahin, A., Schrank, B.R., Al-Lawati, H., Costantino, I., Benz, E., Fard, D., Albers, A.D., Cao, L., Gomez, A.C., Evans, K., Ratti, E., Cudkowicz, M., Frosch, M.P., Talkowski, M., Sorger, P.K., Hyman, B.T. & Albers, M.W. (2021) 'Genome-encoded

cytoplasmic double-stranded RNAs, found in C9ORF72 ALS-FTD brain, propagate neuronal loss', *Science Translational Medicine*, 13(601), p. eaaz4699.

Romano, S., Savva, G.M., Bedarf, J.R., Charles, I.G., Hildebrand, F. & Narbad, A. (2021) 'Meta-analysis of the Parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation', *npj Parkinson's Disease*, 7(1), pp. 1–13.

Ron, M. & Ulitsky, I. (2022) 'Context-specific effects of sequence elements on subcellular localization of linear and circular RNAs', *Nature Communications*, 13(1), p. 2481.

Rose, K.N., Schwarzschild, M.A. & Gomperts, S.N. (2024) 'Clearing the Smoke: What Protects Smokers from Parkinson's Disease?', *Movement Disorders*, 39(2), pp. 267–272.

Ross, G.W., Petrovitch, H., Abbott, R.D., Tanner, C.M., Popper, J., Masaki, K., Launer, L. & White, L.R. (2008) 'Association of olfactory dysfunction with risk for future Parkinson's disease', *Annals of Neurology*, 63(2), pp. 167–173.

Rossi, M., Candelise, N., Baiardi, S., Capellari, S., Giannini, G., Orrù, C.D., Antelmi, E., Mammana, A., Hughson, A.G., Calandra-Buonaura, G., Ladogana, A., Plazzi, G., Cortelli, P., Caughey, B. & Parchi, P. (2020) 'Ultrasensitive RT-QuIC assay with high sensitivity and specificity for Lewy body-associated synucleinopathies', *Acta Neuropathologica*, 140(1), pp. 49–62.

Roth, S.H., Levanon, E.Y. & Eisenberg, E. (2019) 'Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity', *Nature Methods*, 16(11), pp. 1131–1138.

Russo, M.J., Orru, C.D., Concha-Marambio, L., Giaisi, S., Groveman, B.R., Farris, C.M., Holguin, B., Hughson, A.G., LaFontant, D.-E., Caspell-Garcia, C., Coffey, C.S., Mollon, J., Hutten, S.J., Merchant, K., Heym, R.G., Soto, C., Caughey, B. & Kang, U.J. (2021) 'High diagnostic performance of independent alpha-synuclein seed amplification assays for detection of early Parkinson's disease', *Acta Neuropathologica Communications*, 9p. 179.

Rutledge, J., Lehallier, B., Zarifkar, P., Losada, P.M., Shahid-Besanti, M., Western, D., Gorijala, P., Ryman, S., Yutsis, M., Deutsch, G.K., Mormino, E., Trelle, A., Wagner, A.D., Kerchner, G.A., Tian, L., Cruchaga, C., Henderson, V.W., Montine, T.J., Borghammer, P., et al. (2024) 'Comprehensive proteomics of CSF, plasma, and urine identify DDC and other biomarkers of early Parkinson's disease', *Acta Neuropathologica*, 147(1), p. 52.

Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., Herzog, M., Schreyer, L., Papavasileiou, P., Ivanov, A., Öhman, M., Refojo, D., Kadener, S. & Rajewsky, N. (2015) 'Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed', *Molecular Cell*, 58(5), pp. 870–885.

Saborio, G.P., Permanne, B. & Soto, C. (2001) 'Sensitive detection of pathological prion protein by cyclic amplification of protein misfolding', *Nature*, 411(6839), pp. 810–813.

Saiki, S., Hatano, T., Fujimaki, M., Ishikawa, K.-I., Mori, A., Oji, Y., Okuzumi, A., Fukuhara, T., Koinuma, T., Imamichi, Y., Nagumo, M., Furuya, N., Nojiri, S., Amo, T., Yamashiro, K. & Hattori, N. (2017) 'Decreased long-chain acylcarnitines from

insufficient β-oxidation as potential early diagnostic markers for Parkinson's disease', *Scientific Reports*, 7(1), p. 7328.

Sakakibara, R., Tateno, F., Kishi, M., Tsuyuzaki, Y., Uchiyama, T. & Yamamoto, T. (2012) 'Pathophysiology of bladder dysfunction in Parkinson's disease', *Neurobiology of Disease*, 46(3), pp. 565–571.

Sakka, L., Coll, G. & Chazal, J. (2011) 'Anatomy and physiology of cerebrospinal fluid', *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 128(6), pp. 309–316.

Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P.P. (2011) 'A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?', *Cell*, 146(3), pp. 353–358.

Salvadó, G., Horie, K., Barthélemy, N.R., Vogel, J.W., Pichet Binette, A., Chen, C.D., Aschenbrenner, A.J., Gordon, B.A., Benzinger, T.L.S., Holtzman, D.M., Morris, J.C., Palmqvist, S., Stomrud, E., Janelidze, S., Ossenkoppele, R., Schindler, S.E., Bateman, R.J. & Hansson, O. (2024) 'Disease staging of Alzheimer's disease using a CSF-based biomarker model', *Nature Aging*, 4(5), pp. 694–708.

Salzman, J., Chen, R.E., Olsen, M.N., Wang, P.L. & Brown, P.O. (2013) 'Cell-Type Specific Features of Circular RNA Expression', *PLOS Genetics*, 9(9), p. e1003777.

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N. & Brown, P.O. (2012) 'Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types', *PLoS ONE*, 7(2), .

Sánchez-Ferro, Á., Rábano, A., Catalán, M.J., Rodríguez-Valcárcel, F.C., Díez, S.F., Herreros-Rodríguez, J., García-Cobos, E., Álvarez-Santullano, M.M., López-Manzanares, L., Mosqueira, A.J., Desojo, L.V., López-Lozano, J.J., López-Valdés, E., Sánchez-Sánchez, R. & Molina-Arjona, J.A. (2015) 'In vivo gastric detection of α-synuclein inclusions in Parkinson's disease', *Movement Disorders*, 30(4), pp. 517–524.

Sang, Q., Liu, X., Wang, L., Qi, L., Sun, W., Wang, W., Sun, Y. & Zhang, H. (2018) 'CircSNCA downregulation by pramipexole treatment mediates cell apoptosis and autophagy in Parkinson's disease by targeting miR-7', *Aging*, 10(6), pp. 1281–1293.

Sanger, H.L., Klotz, G., Riesner, D., Gross, H.J. & Kleinschmidt, A.K. (1976) 'Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures.', *Proceedings of the National Academy of Sciences*, 73(11), pp. 3852–3856.

Santiago, J.A., Littlefield, A.M. & Potashkin, J.A. (2016) 'Integrative transcriptomic meta-analysis of Parkinson's disease and depression identifies NAMPT as a potential blood biomarker for de novo Parkinson's disease', *Scientific Reports*, 6(1), p. 34579.

Santiago, J.A. & Potashkin, J.A. (2013) 'Integrative Network Analysis Unveils Convergent Molecular Pathways in Parkinson's Disease and Diabetes', *PLOS ONE*, 8(12), p. e83940.

Santiago, J.A. & Potashkin, J.A. (2015) 'Network-based metaanalysis identifies HNF4A and PTBP1 as longitudinally dynamic biomarkers for Parkinson's disease', *Proceedings of the National Academy of Sciences*, 112(7), pp. 2257–2262.

Santos-Rodriguez, G., Voineagu, I. & Weatheritt, R.J. (2021) 'Evolutionary dynamics of circular RNAs in primates' Juan Valcárcel & Patricia J Wittkopp (eds.), *eLife*, 10p. e69148.

Sanz, H., Valim, C., Vegas, E., Oller, J.M. & Reverter, F. (2018) 'SVM-RFE: selection and visualization of the most relevant features through non-linear kernels', *BMC Bioinformatics*, 19(1), p. 432.

Sato, C., Morgan, A., Lang, A.E., Salehi-Rad, S., Kawarai, T., Meng, Y., Ray, P.N., Farrer, L.A., St George-Hyslop, P. & Rogaeva, E. (2005) 'Analysis of the glucocerebrosidase gene in Parkinson's disease', *Movement Disorders*, 20(3), pp. 367–370.

Schenk, D.B., Koller, M., Ness, D.K., Griffith, S.G., Grundman, M., Zago, W., Soto, J., Atiee, G., Ostrowitzki, S. & Kinney, G.G. (2017) 'First-in-human assessment of PRX002, an anti–α-synuclein monoclonal antibody, in healthy volunteers', *Movement Disorders*, 32(2), pp. 211–218.

Scherzer, C.R., Eklund, A.C., Morse, L.J., Liao, Z., Locascio, J.J., Fefer, D., Schwarzschild, M.A., Schlossmacher, M.G., Hauser, M.A., Vance, J.M., Sudarsky, L.R., Standaert, D.G., Growdon, J.H., Jensen, R.V. & Gullans, S.R. (2007) 'Molecular markers of early Parkinson's disease based on gene expression in blood', *Proceedings of the National Academy of Sciences*, 104(3), pp. 955–960.

Schiess, M.C., Zheng, H., Soukup, V.M., Bonnen, J.G. & Nauta, H.J.W. (2000) 'Parkinson's disease subtypes: clinical classification and ventricular cerebrospinal fluid analysis', *Parkinsonism & Related Disorders*, 6(2), pp. 69–76.

Schiess, N., Cataldi, R., Okun, M.S., Fothergill-Misbah, N., Dorsey, E.R., Bloem, B.R., Barretto, M., Bhidayasiri, R., Brown, R., Chishimba, L., Chowdhary, N., Coslov, M., Cubo, E., Di Rocco, A., Dolhun, R., Dowrick, C., Fung, V.S.C., Gershanik, O.S., Gifford, L., et al. (2022) 'Six Action Steps to Address Global Disparities in Parkinson Disease: A World Health Organization Priority', *JAMA Neurology*, 79(9), pp. 929–936.

Schindler, S.E., Galasko, D., Pereira, A.C., Rabinovici, G.D., Salloway, S., Suárez-Calvet, M., Khachaturian, A.S., Mielke, M.M., Udeh-Momoh, C., Weiss, J., Batrla, R., Bozeat, S., Dwyer, J.R., Holzapfel, D., Jones, D.R., Murray, J.F., Partrick, K.A., Scholler, E., Vradenburg, G., et al. (2024) 'Acceptable performance of blood biomarker tests of amyloid pathology — recommendations from the Global CEO Initiative on Alzheimer's Disease', *Nature Reviews Neurology*, 20(7), pp. 426–439.

Schlee, M. & Hartmann, G. (2016) 'Discriminating self from non-self in nucleic acid sensing', *Nature Reviews Immunology*, 16(9), pp. 566–580.

Schneider, S.A. & Alcalay, R.N. (2017) 'Neuropathology of genetic synucleinopathies with parkinsonism: Review of the literature', *Movement Disorders: Official Journal of the Movement Disorder Society*, 32(11), pp. 1504–1523.

Schneider, T. & Bindereif, A. (2017) 'Circular RNAs: Coding or noncoding?', *Cell Research*, 27(6), pp. 724–725.

Schrag, A., Siddiqui, U.F., Anastasiou, Z., Weintraub, D. & Schott, J.M. (2017) 'Clinical variables and biomarkers in prediction of cognitive impairment in patients with newly

diagnosed Parkinson's disease: a cohort study', *The Lancet Neurology*, 16(1), pp. 66–75.

Schulte, C. & Gasser, T. (2011) 'Genetic basis of Parkinson's disease: inheritance, penetrance, and expression', *The Application of Clinical Genetics*, 4pp. 67–80.

Schulz, J., Takousis, P., Wohlers, I., Itua, I.O.G., Dobricic, V., Rücker, G., Binder, H., Middleton, L., Ioannidis, J.P.A., Perneczky, R., Bertram, L. & Lill, C.M. (2019) 'Meta-analyses identify differentially expressed microRNAs in Parkinson's disease', *Annals of Neurology*, 85(6), pp. 835–851.

Schuster, J., Sundblom, J., Thuresson, A.-C., Hassin-Baer, S., Klopstock, T., Dichgans, M., Cohen, O.S., Raininko, R., Melberg, A. & Dahl, N. (2011) 'Genomic duplications mediate overexpression of lamin B1 in adult-onset autosomal dominant leukodystrophy (ADLD) with autonomic symptoms', *neurogenetics*, 12(1), pp. 65–72.

Schwarzschild, M.A., Schwid, S.R., Marek, K., Watts, A., Lang, A.E., Oakes, D., Shoulson, I., Ascherio, A., & Parkinson Study Group PRECEPT Investigators (2008) 'Serum Urate as a Predictor of Clinical and Radiographic Progression in Parkinson Disease', *Archives of Neurology*, 65(6), pp. 716–723.

Seeler, S., Moldovan, L.-I., Bertelsen, T., Hager, H., Iversen, L., Johansen, C., Kjems, J. & Kristensen, L.S. (2022) 'Global circRNA expression changes predate clinical and histological improvements of psoriasis patients upon secukinumab treatment', *PLOS ONE*, 17(9), p. e0275219.

Sela, N., Mersch, B., Gal-Mark, N., Lev-Maor, G., Hotz-Wagenblatt, A. & Ast, G. (2007) 'Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome', *Genome Biology*, 8(6), p. R127.

Shahnawaz, M., Mukherjee, A., Pritzkow, S., Mendez, N., Rabadia, P., Liu, X., Hu, B., Schmeichel, A., Singer, W., Wu, G., Tsai, A.-L., Shirani, H., Nilsson, K.P.R., Low, P.A. & Soto, C. (2020) 'Discriminating α-synuclein strains in Parkinson's disease and multiple system atrophy', *Nature*, 578(7794), pp. 273–277.

Shahnawaz, M., Tokuda, T., Waragai, M., Mendez, N., Ishii, R., Trenkwalder, C., Mollenhauer, B. & Soto, C. (2017) 'Development of a Biochemical Diagnosis of Parkinson Disease by Detection of α-Synuclein Misfolded Aggregates in Cerebrospinal Fluid', *JAMA Neurology*, 74(2), pp. 163–172.

Shamir, R., Klein, C., Amar, D., Vollstedt, E.-J., Bonin, M., Usenovic, M., Wong, Y.C., Maver, A., Poths, S., Safer, H., Corvol, J.-C., Lesage, S., Lavi, O., Deuschl, G., Kuhlenbaeumer, G., Pawlack, H., Ulitsky, I., Kasten, M., Riess, O., et al. (2017) 'Analysis of blood-based gene expression in idiopathic Parkinson disease', *Neurology*, 89(16), pp. 1676–1683.

Shang, R., Lee, S., Senavirathne, G. & Lai, E.C. (2023) 'microRNAs in action: biogenesis, function and regulation', *Nature Reviews Genetics*, 24(12), pp. 816–833.

Shani, V., Safory, H., Szargel, R., Wang, N., Cohen, T., Elghani, F.A., Hamza, H., Savyon, M., Radzishevsky, I., Shaulov, L., Rott, R., Lim, K.-L., Ross, C.A., Bandopadhyay, R.,

Zhang, H. & Engelender, S. (2019) 'Physiological and pathological roles of LRRK2 in the nuclear envelope integrity', *Human Molecular Genetics*, 28(23), pp. 3982–3996.

Shannon, K.M., Keshavarzian, A., Mutlu, E., Dodiya, H.B., Daian, D., Jaglin, J.A. & Kordower, J.H. (2012) 'Alpha-synuclein in colonic submucosa in early untreated Parkinson's disease', *Movement Disorders*, 27(6), pp. 709–715.

Shao, M., Hao, S., Jiang, L., Cai, Y., Zhao, X., Chen, Q., Gao, X. & Xu, J. (2022) 'CRIT: Identifying RNA-binding protein regulator in circRNA life cycle via non-negative matrix factorization', *Molecular Therapy - Nucleic Acids*, 30pp. 398–406.

Shao, Y. & Le, W. (2019) 'Recent advances and perspectives of metabolomics-based investigations in Parkinson's disease', *Molecular Neurodegeneration*, 14.

Shen, H., An, O., Ren, X., Song, Y., Tang, S.J., Ke, X.-Y., Han, J., Tay, D.J.T., Ng, V.H.E., Molias, F.B., Pitcheshwar, P., Leong, K.W., Tan, K.-K., Yang, H. & Chen, L. (2022) 'ADARs act as potent regulators of circular transcriptome in cancer', *Nature Communications*, 13(1), p. 1508.

Siderowf, A., Concha-Marambio, L., Lafontant, D.-E., Farris, C.M., Ma, Y., Urenia, P.A., Nguyen, H., Alcalay, R.N., Chahine, L.M., Foroud, T., Galasko, D., Kieburtz, K., Merchant, K., Mollenhauer, B., Poston, K.L., Seibyl, J., Simuni, T., Tanner, C.M., Weintraub, D., et al. (2023) 'Assessment of heterogeneity among participants in the Parkinson's Progression Markers Initiative cohort using α-synuclein seed amplification: a cross-sectional study', *The Lancet Neurology*, 22(5), pp. 407–417.

Sidransky, E., Nalls, M.A., Aasly, J.O., Aharon-Peretz, J., Annesi, G., Barbosa, E.R., Bar-Shira, A., Berg, D., Bras, J., Brice, A., Chen, C.-M., Clark, L.N., Condroyer, C., De Marco, E.V., Dürr, A., Eblan, M.J., Fahn, S., Farrer, M., Fung, H.-C., et al. (2009) 'Multi-center analysis of glucocerebrosidase mutations in Parkinson disease', *The New England journal of medicine*, 361(17), pp. 1651–1661.

Silverman, R.H. (2007) 'Viral Encounters with 2′,5′-Oligoadenylate Synthetase and RNase L during the Interferon Antiviral Response', *Journal of Virology*, 81(23), pp. 12720–12729.

Simón-Sánchez, J., Schulte, C., Bras, J.M., Sharma, M., Gibbs, J.R., Berg, D., Paisan-Ruiz, C., Lichtner, P., Scholz, S.W., Hernandez, D.G., Krüger, R., Federoff, M., Klein, C., Goate, A., Perlmutter, J., Bonin, M., Nalls, M.A., Illig, T., Gieger, C., et al. (2009) 'Genome-wide association study reveals genetic risk underlying Parkinson's disease', *Nature Genetics*, 41(12), pp. 1308–1312.

Simpson, C., Vinikoor-Imler, L., Nassan, F.L., Shirvan, J., Lally, C., Dam, T. & Maserejian, N. (2022) 'Prevalence of ten *LRRK2* variants in Parkinson's disease: A comprehensive review', *Parkinsonism & Related Disorders*, 98pp. 103–113.

Sims, D., Sudbery, I., Ilott, N.E., Heger, A. & Ponting, C.P. (2014) 'Sequencing depth and coverage: key considerations in genomic analyses', *Nature Reviews Genetics*, 15(2), pp. 121–132.

Simuni, T., Brumm, M.C., Uribe, L., Caspell-Garcia, C., Coffey, C.S., Siderowf, A., Alcalay, R.N., Trojanowski, J.Q., Shaw, L.M., Seibyl, J., Singleton, A., Toga, A.W., Galasko, D., Foroud, T., Nudelman, K., Tosun-Turgut, D., Poston, K., Weintraub, D.,

Mollenhauer, B., et al. (2020) ‘Clinical and Dopamine Transporter Imaging Characteristics of Leucine Rich Repeat Kinase 2 (LRRK2) and Glucosylceramidase Beta (GBA) Parkinson’s Disease Participants in the Parkinson’s Progression Markers Initiative: A Cross-Sectional Study’, *Movement Disorders*, 35(5), pp. 833–844.

Simuni, T., Chahine, L.M., Poston, K., Brumm, M., Buracchio, T., Campbell, M., Chowdhury, S., Coffey, C., Concha-Marambio, L., Dam, T., DiBiaso, P., Foroud, T., Frasier, M., Gochanour, C., Jennings, D., Kieburtz, K., Kopil, C.M., Merchant, K., Mollenhauer, B., et al. (2024) ‘A biological definition of neuronal α-synuclein disease: towards an integrated staging system for research’, *The Lancet Neurology*, 23(2), pp. 178–190.

Simuni, T., Siderowf, A., Lasch, S., Coffey, C.S., Caspell-Garcia, C., Jennings, D., Tanner, C.M., Trojanowski, J.Q., Shaw, L.M., Seibyl, J., Schuff, N., Singleton, A., Kieburtz, K., Toga, A.W., Mollenhauer, B., Galasko, D., Chahine, L.M., Weintraub, D., Foroud, T., et al. (2018) ‘Longitudinal Change of Clinical and Biological Measures in Early Parkinson’s Disease: Parkinson’s Progression Markers Initiative Cohort’, *Movement Disorders*, 33(5), p. 771.

Simuni, T., Uribe, L., Cho, H.R., Caspell-Garcia, C., Coffey, C.S., Siderowf, A., Trojanowski, J.Q., Shaw, L.M., Seibyl, J., Singleton, A., Toga, A.W., Galasko, D., Foroud, T., Tosun, D., Poston, K., Weintraub, D., Mollenhauer, B., Tanner, C.M., Kieburtz, K., et al. (2020) ‘Clinical and dopamine transporter imaging characteristics of non-manifest LRRK2 and GBA mutation carriers in the Parkinson’s Progression Markers Initiative (PPMI): a cross-sectional study’, *The Lancet Neurology*, 19(1), pp. 71–80.

Simunovic, F., Yi, M., Wang, Y., Macey, L., Brown, L.T., Krichevsky, A.M., Andersen, S.L., Stephens, R.M., Benes, F.M. & Sonntag, K.C. (2009) ‘Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson’s disease pathology’, *Brain*, 132(7), pp. 1795–1809.

Siqueira, G.S.A., Hagemann, P. de M.S., Coelho, D. de S., Santos, F.H.D. & Bertolucci, P.H.F. (2019) ‘Can MoCA and MMSE Be Interchangeable Cognitive Screening Tools? A Systematic Review’, *The Gerontologist*, 59(6), pp. e743–e763.

Sixel-Döring, F., Muntean, M.-L., Petersone, D., Leha, A., Lang, E., Mollenhauer, B. & Trenkwalder, C. (2023) ‘The Increasing Prevalence of REM Sleep Behavior Disorder with Parkinson’s Disease Progression: A Polysomnography-Supported Study’, *Movement Disorders Clinical Practice*, 10(12), pp. 1769–1776.

Smajić, S., Prada-Medina, C.A., Landoulsi, Z., Ghelfi, J., Delcambre, S., Dietrich, C., Jarazo, J., Henck, J., Balachandran, S., Pachchek, S., Morris, C.M., Antony, P., Timmermann, B., Sauer, S., Pereira, S.L., Schwamborn, J.C., May, P., Grünewald, A. & Spielmann, M. (2022) ‘Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state’, *Brain*, 145(3), pp. 964–978.

Snyder, M.P., Gingeras, T.R., Moore, J.E., Weng, Z., Gerstein, M.B., Ren, B., Hardison, R.C., Stamatoyannopoulos, J.A., Graveley, B.R., Feingold, E.A., Pazin, M.J., Pagan, M., Gilchrist, D.A., Hitz, B.C., Cherry, J.M., Bernstein, B.E., Mendenhall, E.M., Zerbino, D.R., Frankish, A., et al. (2020) ‘Perspectives on ENCODE’, *Nature*, 583(7818), pp. 693–698.

Soneson, C., Love, M.I. & Robinson, M.D. (2015) ‘Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences’, *F1000Research*, 4p. 1521.

Song, M., Graubard, B.I., Rabkin, C.S. & Engels, E.A. (2021) 'Neutrophil-to-lymphocyte ratio and mortality in the United States general population', *Scientific Reports*, 11(1), p. 464.

Sotirakis, C., Su, Z., Brzezicki, M.A., Conway, N., Tarassenko, L., FitzGerald, J.J. & Antoniades, C.A. (2023) 'Identification of motor progression in Parkinson's disease using wearable sensors and machine learning', *npj Parkinson's Disease*, 9(1), pp. 1–8.

Soto, C. (2024) 'α-Synuclein seed amplification technology for Parkinson's disease and related synucleinopathies', *Trends in Biotechnology*,

Spector, R., Robert Snodgrass, S. & Johanson, C.E. (2015) 'A balanced view of the cerebrospinal fluid composition and functions: Focus on adult humans', *Experimental Neurology*, 273pp. 57–68.

Spielmann, N., Miller, G., Oprea, T.I., Hsu, C.-W., Fobo, G., Frishman, G., Montrone, C., Haseli Mashhadi, H., Mason, J., Munoz Fuentes, V., Leuchtenberger, S., Ruepp, A., Wagner, M., Westphal, D.S., Wolf, C., Görlach, A., Sanz-Moreno, A., Cho, Y.-L., Teperino, R., et al. (2022) 'Extensive identification of genes involved in congenital and structural heart disorders and cardiomyopathy', *Nature Cardiovascular Research*, 1(2), pp. 157–173.

Srivastava, A., Malik, L., Sarkar, H., Zakeri, M., Almodaresi, F., Soneson, C., Love, M.I., Kingsford, C. & Patro, R. (2020) 'Alignment and mapping methodology influence transcript abundance estimation', *Genome Biology*, 21(1), p. 239.

Stagsted, L.V., Nielsen, K.M., Daugaard, I. & Hansen, T.B. (2019) 'Noncoding AUG circRNAs constitute an abundant and conserved subclass of circles', *Life Science Alliance*, 2(3), .

Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L.-H. & Bindereif, A. (2015) 'Exon Circularization Requires Canonical Splice Signals', *Cell Reports*, 10(1), pp. 103–111.

Steen, C.B., Liu, C.L., Alizadeh, A.A. & Newman, A.M. (2020) 'Profiling Cell Type Abundance and Expression in Bulk Tissues with CIBERSORTx', in Benjamin L. Kidder (ed.) *Stem Cell Transcriptional Networks: Methods and Protocols*. Methods in Molecular Biology. [Online]. New York, NY: Springer US. pp. 135–157.

Steiner, J.A., Quansah, E. & Brundin, P. (2018) 'The concept of alpha-synuclein as a prion-like protein: ten years after', *Cell and Tissue Research*, 373(1), pp. 161–173.

Steinmetz, J.D., Seeher, K.M., Schiess, N., Nichols, E., Cao, B., Servili, C., Cavallera, V., Cousin, E., Hagins, H., Moberg, M.E., Mehlman, M.L., Abate, Y.H., Abbas, J., Abbasi, M.A., Abbasian, M., Abbastabar, H., Abdelmasseh, M., Abdollahi, Mohammad, Abdollahi, Mozhan, et al. (2024) 'Global, regional, and national burden of disorders affecting the nervous system, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021', *The Lancet Neurology*, 23(4), pp. 344–381.

Stevens, C.H., Rowe, D., Morel-Kopp, M.-C., Orr, C., Russell, T., Ranola, M., Ward, C. & Halliday, G.M. (2012) 'Reduced T helper and B lymphocytes in Parkinson's disease', *Journal of Neuroimmunology*, 252(1), pp. 95–99.

Stoessel, D., Schulte, C., Teixeira dos Santos, M.C., Scheller, D., Rebollo-Mesa, I., Deuschle, C., Walther, D., Schauer, N., Berg, D., Nogueira da Costa, A. & Maetzler, W. (2018)

'Promising Metabolite Profiles in the Plasma and CSF of Early Clinical Parkinson's Disease', *Frontiers in Aging Neuroscience*, 10.

Stolzenberg, E., Berry, D., Yang, D., Lee, E.Y., Kroemer, A., Kaufman, S., Wong, G.C.L., Oppenheim, J.J., Sen, S., Fishbein, T., Bax, A., Harris, B., Barbut, D. & Zasloff, M.A. (2017) 'A Role for Neuronal Alpha-Synuclein in Gastrointestinal Immunity', *Journal of Innate Immunity*, 9(5), pp. 456–463.

Stone, M.J., Chuang, S., Hou, X., Shoham, M. & Zhu, J.Z. (2009) 'Tyrosine sulfation: an increasingly recognised post-translational modification of secreted proteins', *New Biotechnology*, 25(5), pp. 299–317.

Su, Z., Łabaj, P.P., Li, S., Thierry-Mieg, J., Thierry-Mieg, D., Shi, W., Wang, C., Schroth, G.P., Setterquist, R.A., Thompson, J.F., Jones, W.D., Xiao, W., Xu, W., Jensen, R.V., Kelly, R., Xu, J., Conesa, A., Furlanello, C., Gao, H., et al. (2014) 'A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium', *Nature Biotechnology*, 32(9), pp. 903–914.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. & Mesirov, J.P. (2005) 'Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles', *Proceedings of the National Academy of Sciences*, 102(43), pp. 15545–15550.

Sulzer, D., Alcalay, R.N., Garretti, F., Cote, L., Kanter, E., Agin-Liebes, J., Liong, C., McMurtrey, C., Hildebrand, W.H., Mao, X., Dawson, V.L., Dawson, T.M., Oseroff, C., Pham, J., Sidney, J., Dillon, M.B., Carpenter, C., Weiskopf, D., Phillips, E., et al. (2017) 'T cells from patients with Parkinson's disease recognize α-synuclein peptides', *Nature*, 546(7660), pp. 656–661.

Surmeier, D.J. (2018) 'Determinants of dopaminergic neuron loss in Parkinson's disease', *The FEBS journal*, 285(19), pp. 3657–3668.

Surono, A., Takeshima, Y., Wibawa, T., Ikezawa, M., Nonaka, I. & Matsuo, M. (1999) 'Circular Dystrophin RNAs Consisting of Exons That Were Skipped by Alternative Splicing', *Human Molecular Genetics*, 8(3), pp. 493–500.

Sutton, G.J., Poppe, D., Simmons, R.K., Walsh, K., Nawaz, U., Lister, R., Gagnon-Bartsch, J.A. & Voineagu, I. (2022) 'Comprehensive evaluation of deconvolution methods for human brain gene expression', *Nature Communications*, 13(1), p. 1358.

Swaminathan, S., Huentelman, M.J., Corneveaux, J.J., Myers, A.J., Faber, K.M., Foroud, T., Mayeux, R., Shen, L., Kim, S., Turk, M., Hardy, J., Reiman, E.M. & Saykin, A.J. (2012) 'Analysis of Copy Number Variation in Alzheimer's Disease in a Cohort of Clinically Characterized and Neuropathologically Verified Individuals', *PLOS ONE*, 7(12), p. e50640.

Szabo, L., Morey, R., Palpant, N.J., Wang, P.L., Afari, N., Jiang, C., Parast, M.M., Murry, C.E., Laurent, L.C. & Salzman, J. (2015) 'Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development', *Genome Biology*, 16(1), p. 126.

Szabo, L. & Salzman, J. (2016) 'Detecting circular RNAs: bioinformatic and experimental challenges', *Nature Reviews Genetics*, 17(11), pp. 679–692.

Tagawa, T., Oh, D., Santos, J., Dremel, S., Mahesh, G., Uldrick, T.S., Yarchoan, R., Kopardé, V.N. & Ziegelbauer, J.M. (2021) 'Characterizing Expression and Regulation of Gamma-Herpesviral Circular RNAs', *Frontiers in Microbiology*, 12.

Tan, E.-K., Chao, Y.-X., West, A., Chan, L.-L., Poewe, W. & Jankovic, J. (2020) 'Parkinson disease and the immune system — associations, mechanisms and therapeutics', *Nature Reviews Neurology*, 16(6), pp. 303–318.

Tan, M.H., Li, Q., Shanmugam, R., Piskol, R., Kohler, J., Young, A.N., Liu, K.I., Zhang, R., Ramaswami, G., Ariyoshi, K., Gupte, A., Keegan, L.P., George, C.X., Ramu, A., Huang, N., Pollina, E.A., Leeman, D.S., Rustighi, A., Goh, Y.P.S., et al. (2017) 'Dynamic landscape and regulation of RNA editing in mammals', *Nature*, 550(7675), pp. 249–254.

Tang, S.J., Shen, H., An, O., Hong, H., Li, J., Song, Y., Han, J., Tay, D.J.T., Ng, V.H.E., Bellido Molias, F., Leong, K.W., Pitcheshwar, P., Yang, H. & Chen, L. (2020) 'Cis- and trans-regulations of pre-mRNA splicing by RNA editing enzymes influence cancer development', *Nature Communications*, 11(1), p. 799.

Tansey, M.G., Wallings, R.L., Houser, M.C., Herrick, M.K., Keating, C.E. & Joers, V. (2022) 'Inflammation and immune dysfunction in Parkinson disease', *Nature Reviews Immunology*, 22(11), pp. 657–673.

Tarailo-Graovac, M. & Chen, N. (2009) 'Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences', *Current Protocols in Bioinformatics*, 25(1), p. 4.10.1-4.10.14.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. (2015) 'Sambamba: fast processing of NGS alignment formats', *Bioinformatics*, 31(12), pp. 2032–2034.

Thaler, A., Bregman, N., Gurevich, T., Shiner, T., Dror, Y., Zmira, O., Gan-Or, Z., Bar-Shira, A., Gana-Weisz, M., Orr-Urtreger, A., Giladi, N. & Mirelman, A. (2018) 'Parkinson's disease phenotype is influenced by the severity of the mutations in the *GBA* gene', *Parkinsonism & Related Disorders*, 55pp. 45–49.

Thaler, A., Omer, N., Giladi, N., Gurevich, T., Bar-Shira, A., Gana-Weisz, M., Goldstein, O., Kestenbaum, M., Shirvan, J.C., Cedarbaum, J.M., Orr-Urtreger, A., Regev, K., Shenhar-Tsarfaty, S. & Mirelman, A. (2021) 'Mutations in GBA and LRRK2 Are Not Associated with Increased Inflammatory Markers', *Journal of Parkinson's Disease*, 11(3), pp. 1285–1296.

Thanvi, B., Lo, N. & Robinson, T. (2007) 'Levodopa-induced dyskinesia in Parkinson's disease: clinical features, pathogenesis, prevention and treatment', *Postgraduate Medical Journal*, 83(980), pp. 384–388.

The UK Parkinson's Disease Consortium and The Wellcome Trust Case Control Consortium 2 (2011) 'Dissection of the genetics of Parkinson's disease identifies an additional association 5′ of SNCA and multiple associated haplotypes at 17q21', *Human Molecular Genetics*, 20(2), pp. 345–353.

Thobois, S., Prange, S., Scheiber, C. & Broussolle, E. (2019) 'What a neurologist should know about PET and SPECT functional imaging for parkinsonism: A practical perspective', *Parkinsonism & Related Disorders*, 59pp. 93–100.

Tomecki, R. & Dziembowski, A. (2010) 'Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism', *RNA*, 16(9), pp. 1692–1724.

Tomlinson, C.L., Stowe, R., Patel, S., Rick, C., Gray, R. & Clarke, C.E. (2010) 'Systematic review of levodopa dose equivalency reporting in Parkinson's disease', *Movement Disorders*, 25(15), pp. 2649–2653.

Tourrière, H., Gallouzi, I.E., Chebli, K., Capony, J.P., Mouaikel, J., van der Geer, P. & Tazi, J. (2001) 'RasGAP-associated endoribonuclease G3Bp: selective RNA degradation and phosphorylation-dependent localization', *Molecular and Cellular Biology*, 21(22), pp. 7747–7760.

Tran, J., Anastacio, H. & Bardy, C. (2020) 'Genetic predispositions of Parkinson's disease revealed in patient-derived brain cells', *npj Parkinson's Disease*, 6(1), pp. 1–18.

Treglia, G., Cason, E., Stefanelli, A., Cocciolillo, F., Di Giuda, D., Fagioli, G. & Giordano, A. (2012) 'MIBG scintigraphy in differential diagnosis of Parkinsonism: a meta-analysis', *Clinical Autonomic Research*, 22(1), pp. 43–55.

Tsalenchuk, M., Gentleman, S.M. & Marzi, S.J. (2023) 'Linking environmental risk factors with epigenetic mechanisms in Parkinson's disease', *npj Parkinson's Disease*, 9(1), pp. 1–12.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C.A.-K., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., et al. (2015) 'Tissue-based map of the human proteome', *Science*, 347(6220), p. 1260419.

Vabalas, A., Gowen, E., Poliakoff, E. & Casson, A.J. (2019) 'Machine learning algorithm validation with a limited sample size', *PLOS ONE*, 14(11), p. e0224365.

Van Laere, K., Everaert, L., Annemans, L., Gonce, M., Vandenberghe, W. & Vander Borght, T. (2008) 'The cost effectiveness of 123I-FP-CIT SPECT imaging in patients with an uncertain clinical diagnosis of parkinsonism', *European Journal of Nuclear Medicine and Molecular Imaging*, 35(7), pp. 1367–1376.

Varma, S. & Simon, R. (2006) 'Bias in error estimation when using cross-validation for model selection', *BMC Bioinformatics*, 7(1), p. 91.

Vaswani, P.A., Morley, J.F., Jennings, D., Siderowf, A. & Marek, K. (2023) 'Predictive value of abbreviated olfactory tests in prodromal Parkinson disease', *npj Parkinson's Disease*, 9(1), pp. 1–6.

Venø, M.T., Hansen, T.B., Venø, S.T., Clausen, B.H., Grebing, M., Finsen, B., Holm, I.E. & Kjems, J. (2015) 'Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development', *Genome Biology*, 16(1), p. 245.

Verboom, K., Everaert, C., Bolduc, N., Livak, K.J., Yigit, N., Rombaut, D., Anckaert, J., Lee, S., Venø, M.T., Kjems, J., Speleman, F., Mestdagh, P. & Vandesompele, J. (2019) 'SMARTer single cell total RNA sequencing', *Nucleic Acids Research*, 47(16), p. e93.

Verduci, L., Tarcitano, E., Strano, S., Yarden, Y. & Blandino, G. (2021) 'CircRNAs: role in human diseases and potential use as biomarkers', *Cell Death & Disease*, 12(5), pp. 1–12.

Vijiaratnam, N., Simuni, T., Bandmann, O., Morris, H.R. & Foltynie, T. (2021) 'Progress towards therapies for disease modification in Parkinson's disease', *The Lancet. Neurology*, 20(7), pp. 559–572.

Villanueva, E.B., Tresse, E., Liu, Y., Duarte, J.N., Jimenez-Duran, G., Ejlerskov, P., Kretz, O., Loreth, D., Goldmann, T., Prinz, M. & Issazadeh-Navikas, S. (2021) 'Neuronal TNFα, Not α-Syn, Underlies PDD-Like Disease Progression in IFNβ-KO Mice', *Annals of Neurology*, 90(5), pp. 789–807.

Vivacqua, G., Mason, M., De Bartolo, M.I., Węgrzynowicz, M., Calò, L., Belvisi, D., Suppa, A., Fabbrini, G., Berardelli, A. & Spillantini, M. (2023) 'Salivary α-Synuclein RT-QuIC Correlates with Disease Severity in de novo Parkinson's Disease', *Movement Disorders*, n/a(n/a), .

Vo, J.N., Cieslik, M., Zhang, Yajia, Shukla, S., Xiao, L., Zhang, Yuping, Wu, Y.-M., Dhanasekaran, S.M., Engelke, C.G., Cao, X., Robinson, D.R., Nesvizhskii, A.I. & Chinnaiyan, A.M. (2019) 'The Landscape of Circular RNA in Cancer', *Cell*, 176(4), pp. 869-881.e13.

Vollstedt, E.-J., Schaake, S., Lohmann, K., Padmanabhan, S., Brice, A., Lesage, S., Tesson, C., Vidailhet, M., Wurster, I., Hentati, F., Mirelman, A., Giladi, N., Marder, K., Waters, C., Fahn, S., Kasten, M., Brüggemann, N., Borsche, M., Foroud, T., et al. (2023) 'Embracing Monogenic Parkinson's Disease: The MJFF Global Genetic PD Cohort', *Movement Disorders*, 38(2), pp. 286–303.

Volpicelli-Daley, L.A., Luk, K.C., Patel, T.P., Tanik, S.A., Riddle, D.M., Stieber, A., Meany, D.F., Trojanowski, J.Q. & Lee, V.M.-Y. (2011) 'Exogenous α-Synuclein Fibrils Induce Lewy Body Pathology Leading to Synaptic Dysfunction and Neuron Death', *Neuron*, 72(1), pp. 57–71.

Vromman, M., Anckaert, J., Bortoluzzi, S., Buratin, A., Chen, C.-Y., Chu, Q., Chuang, T.-J., Dehghannasiri, R., Dieterich, C., Dong, X., Flicek, P., Gaffo, E., Gu, W., He, C., Hoffmann, S., Izuogu, O., Jackson, M.S., Jakobi, T., Lai, E.C., et al. (2023) 'Large-scale benchmarking of circRNA detection tools reveals large differences in sensitivity but not in precision', *Nature Methods*, 20(8), pp. 1159–1169.

Vromman, M., Vandesompele, J. & Volders, P.-J. (2021) 'Closing the circle: current state and perspectives of circular RNA databases', *Briefings in Bioinformatics*, 22(1), pp. 288–297.

Wallings, R.L. & Tansey, M.G. (2019) 'LRRK2 regulation of immune-pathways and inflammatory disease', *Biochemical Society Transactions*, 47(6), pp. 1581–1595.

Wang, C., Liu, W.-R., Tan, S., Zhou, J.-K., Xu, X., Ming, Y., Cheng, J., Li, J., Zeng, Z., Zuo, Y., He, J., Peng, Y. & Li, W. (2022) 'Characterization of distinct circular RNA signatures in solid tumors', *Molecular Cancer*, 21(1), p. 63.

Wang, M., Li, A., Sekiya, M., Beckmann, N.D., Quan, X., Schrode, N., Fernando, M.B., Yu, A., Zhu, L., Cao, J., Lyu, L., Horgusluoglu, E., Wang, Q., Guo, L., Wang, Y., Neff, R., Song, W., Wang, E., Shen, Q., et al. (2021) 'Transformative Network Modeling of Multi-omics Data Reveals Detailed Circuits, Key Regulators, and Potential Therapeutics for Alzheimer's Disease', *Neuron*, 109(2), pp. 257-272.e14.

Wang, P., Luo, M., Zhou, W., Jin, X., Xu, Z., Yan, S., Li, Y., Xu, C., Cheng, R., Huang, Y., Lin, X., Yao, L., Nie, H. & Jiang, Q. (2022) 'Global Characterization of Peripheral B Cells in Parkinson's Disease by Single-Cell RNA and BCR Sequencing', *Frontiers in Immunology*, 13.

Wang, P., Yao, L., Luo, M., Zhou, W., Jin, X., Xu, Z., Yan, S., Li, Y., Xu, C., Cheng, R., Huang, Y., Lin, X., Ma, K., Cao, H., Liu, H., Xue, G., Han, F., Nie, H. & Jiang, Q. (2021) 'Single-cell transcriptome and TCR profiling reveal activated and expanded T cell populations in Parkinson's disease', *Cell Discovery*, 7(1), pp. 1–16.

Wang, Q., Wang, M., Choi, I., Sarrafha, L., Liang, M., Ho, L., Farrell, K., Beaumont, K.G., Sebra, R., De Sanctis, C., Crary, J.F., Ahfeldt, T., Blanchard, J., Neavin, D., Powell, J., Davis, D.A., Sun, X., Zhang, B. & Yue, Z. (2024) 'Molecular profiling of human substantia nigra identifies diverse neuron types associated with vulnerability in Parkinson's disease', *Science Advances*, 10(2), p. eadi8287.

Wang, X., Yu, S., Li, F. & Feng, T. (2015) 'Detection of α-synuclein oligomers in red blood cells as a potential biomarker of Parkinson's disease', *Neuroscience Letters*, 599pp. 115–119.

Wang, Z., Becker, K., Donadio, V., Siedlak, S., Yuan, J., Rezaee, M., Incensi, A., Kuzkina, A., Orrú, C.D., Tatsuoka, C., Liguori, R., Gunzler, S.A., Caughey, B., Jimenez-Capdeville, M.E., Zhu, X., Doppler, K., Cui, L., Chen, S.G., Ma, J., et al. (2021) 'Skin α-Synuclein Aggregation Seeding Activity as a Novel Biomarker for Parkinson Disease', *JAMA Neurology*, 78(1), pp. 30–40.

Warnecke, T., Schäfer, K.-H., Claus, I., Del Tredici, K. & Jost, W.H. (2022) 'Gastrointestinal involvement in Parkinson's disease: pathophysiology, diagnosis, and management', *NPJ Parkinson's Disease*, 8p. 31.

Weigelt, C.M., Sehgal, R., Tain, L.S., Cheng, J., Eßer, J., Pahl, A., Dieterich, C., Grönke, S. & Partridge, L. (2020) 'An Insulin-Sensitive Circular RNA that Regulates Lifespan in *Drosophila*', *Molecular Cell*, 79(2), pp. 268-279.e5.

Weimers, P., Halfvarson, J., Sachs, M.C., Saunders-Pullman, R., Ludvigsson, J.F., Peter, I., Burisch, J. & Olén, O. (2019) 'Inflammatory Bowel Disease and Parkinson's Disease: A Nationwide Swedish Cohort Study', *Inflammatory Bowel Diseases*, 25(1), pp. 111–123.

Wen, G., Zhou, T. & Gu, W. (2021) 'The potential of using blood circular RNA as liquid biopsy biomarker for human diseases', *Protein & Cell*, 12(12), pp. 911–946.

Wen, M., Zhou, B., Chen, Y.-H., Ma, Z.-L., Gou, Y., Zhang, C.-L., Yu, W.-F. & Jiao, L. (2017) 'Serum uric acid levels in patients with Parkinson's disease: A meta-analysis', *PLOS ONE*, 12(3), p. e0173731.

Wesselhoeft, R.A., Kowalski, P.S. & Anderson, D.G. (2018) 'Engineering circular RNA for potent and stable translation in eukaryotic cells', *Nature Communications*, 9(1), p. 2629.

Wesselhoeft, R.A., Kowalski, P.S., Parker-Hale, F.C., Huang, Y., Bisaria, N. & Anderson, D.G. (2019) 'RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration In Vivo', *Molecular Cell*, 74(3), pp. 508-520.e4.

Westholm, J.O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S.E., Graveley, B.R. & Lai, E.C. (2014) 'Genome-wide Analysis of *Drosophila* Circular RNAs Reveals Their Structural and Sequence Properties and Age-Dependent Neural Accumulation', *Cell Reports*, 9(5), pp. 1966–1980.

White, K., Yang, P., Li, L., Farshori, A., Medina, A.E. & Zielke, H.R. (2018) 'Effect of Postmortem Interval and Years in Storage on RNA Quality of Tissue at a Repository of the NIH NeuroBioBank', *Biopreservation and Biobanking*, 16(2), pp. 148–157.

White, T.L., Sadikot, A.F. & Djordjevic, J. (2016) 'Metacognitive knowledge of olfactory dysfunction in Parkinson's disease', *Brain and Cognition*, 104pp. 1–6.

Wickramasinghe, V.O. & Laskey, R.A. (2015) 'Control of mammalian gene expression by selective mRNA export', *Nature Reviews Molecular Cell Biology*, 16(7), pp. 431–442.

Wijeyekoon, R.S., Kronenberg-Versteeg, D., Scott, K.M., Hayat, S., Kuan, W.-L., Evans, J.R., Breen, D.P., Cummins, G., Jones, J.L., Clatworthy, M.R., Floto, R.A., Barker, R.A. & Williams-Gray, C.H. (2020) 'Peripheral innate immune and bacterial signals relate to clinical heterogeneity in Parkinson's disease', *Brain, Behavior, and Immunity*, 87pp. 473–488.

Williams, S.M., Schulz, P. & Sierks, M.R. (2016) 'Oligomeric α-synuclein and β-amyloid variants as potential biomarkers for Parkinson's and Alzheimer's diseases', *European Journal of Neuroscience*, 43(1), pp. 3–16.

Williams-Gray, C.H., Wijeyekoon, R., Yarnall, A.J., Lawson, R.A., Breen, D.P., Evans, J.R., Cummins, G.A., Duncan, G.W., Khoo, T.K., Burn, D.J., Barker, R.A., & ICICLE-PD study group (2016) 'Serum immune markers and disease progression in an incident Parkinson's disease cohort (ICICLE-PD)', *Movement Disorders: Official Journal of the Movement Disorder Society*, 31(7), pp. 995–1003.

Wingett, S.W. & Andrews, S. (2018) 'FastQ Screen: A tool for multi-genome mapping and quality control', *F1000Research*, 7.

Wirdefeldt, K., Gatz, M., Reynolds, C.A., Prescott, C.A. & Pedersen, N.L. (2011) 'Heritability of Parkinson disease in Swedish twins: a longitudinal study', *Neurobiology of Aging*, 32(10), p. 1923.e1–8.

Wooten, G.F., Currie, L.J., Bovbjerg, V.E., Lee, J.K. & Patrie, J. (2004) 'Are men at greater risk for Parkinson's disease than women?', *Journal of Neurology, Neurosurgery & Psychiatry*, 75(4), pp. 637–639.

Wright, M.N. & Ziegler, A. (2017) 'ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R', *Journal of Statistical Software*, 77pp. 1–17.

Wu, L., Xu, H., Zhang, W., Chen, Z., Li, W. & Ke, W. (2020) 'Circular RNA circCCDC9 alleviates ischaemic stroke ischaemia/reperfusion injury via the Notch pathway', *Journal of Cellular and Molecular Medicine*, 24(24), pp. 14152–14159.

Wu, S., Xue, Q., Qin, X., Wu, X., Kim, P., Chyr, J., Zhou, X. & Huang, L. (2023) 'The Potential Regulation of A-to-I RNA Editing on Genes in Parkinson's Disease', *Genes*, 14(4), p. 919.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X. & Yu, G. (2021) 'clusterProfiler 4.0: A universal enrichment tool for interpreting omics data', *The Innovation*, 2(3), p. 100141.

Wu, W., Zhang, J., Cao, X., Cai, Z. & Zhao, F. (2022) 'Exploring the cellular landscape of circular RNAs using full-length single-cell RNA sequencing', *Nature Communications*, 13(1), p. 3242.

Wu, W., Zhao, F. & Zhang, J. (2024) 'circAtlas 3.0: a gateway to 3 million curated vertebrate circular RNAs based on a standardized nomenclature scheme', *Nucleic Acids Research*, 52(D1), pp. D52–D60.

Xi, Y., Nakajima, G., Gavin, E., Morris, C.G., Kudo, K., Hayashi, K. & Ju, J. (2007) 'Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples', *RNA*, 13(10), pp. 1668–1674.

Xia, P., Wang, S., Ye, B., Du, Y., Li, C., Xiong, Z., Qu, Y. & Fan, Z. (2018) 'A Circular RNA Protects Dormant Hematopoietic Stem Cells from DNA Sensor cGAS-Mediated Exhaustion', *Immunity*, 48(4), pp. 688-701.e7.

Xia, S., Feng, J., Lei, L., Hu, J., Xia, L., Wang, J., Xiang, Y., Liu, L., Zhong, S., Han, L. & He, C. (2017) 'Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes', *Briefings in Bioinformatics*, 18(6), pp. 984–992.

Xiao, M.-S. & Wilusz, J.E. (2019) 'An improved method for circular RNA purification using RNase R that efficiently removes linear RNAs containing G-quadruplexes or structured 3′ ends', *Nucleic Acids Research*, 47(16), pp. 8755–8769.

Xiao, W., Shameli, A., Harding, C.V., Meyerson, H.J. & Maitta, R.W. (2014) 'Late stages of hematopoiesis and B cell lymphopoiesis are regulated by α-synuclein, a key player in Parkinson's disease', *Immunobiology*, 219(11), pp. 836–844.

Xiao, Y., Chen, H., Liao, J., Zhang, Q., He, H., Lei, J., Huang, J., Ouyang, Q., Shen, Y. & Wang, J. (2022) 'The Potential Circular RNAs Biomarker Panel and Regulatory Networks of Parkinson's Disease', *Frontiers in Neuroscience*, 16p. 893713.

Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S. & Kellis, M. (2005) 'Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals', *Nature*, 434(7031), pp. 338–345.

Xu, W., Wang, Y., Quan, H., Liu, D., Zhang, H., Qi, Y., Li, Q., Liao, J., Gao, H.-M., Huang, J. & Zhou, H. (2020) 'Double-stranded RNA-induced dopaminergic neuronal loss in the

substantia nigra in the presence of Mac1 receptor', *Biochemical and Biophysical Research Communications*, 533(4), pp. 1148–1154.

Yamada, T., McGeer, P.L., Baimbridge, K.G. & McGeer, E.G. (1990) 'Relative sparing in Parkinson's disease of substantia nigra dopamine neurons containing calbindin-D28K', *Brain Research*, 526(2), pp. 303–307.

Yan, Shijun, Jiang, C., Janzen, A., Barber, T.R., Seger, A., Sommerauer, M., Davis, J.J., Marek, K., Hu, M.T., Oertel, W.H. & Tofaris, G.K. (2023) 'Neuronally Derived Extracellular Vesicle α-Synuclein as a Serum Biomarker for Individuals at Risk of Developing Parkinson Disease', *JAMA Neurology*,

Yan, Shi, Si, Y., Zhou, W., Cheng, R., Wang, P., Wang, D., Ding, W., Shi, W., Jiang, Q., Yang, F. & Yao, L. (2023) 'Single-cell transcriptomics reveals the interaction between peripheral CD4+ CTLs and mesencephalic endothelial cells mediated by IFNG in Parkinson's disease', *Computers in Biology and Medicine*, 158p. 106801.

Yang, J., Zhang, L., Yu, C., Yang, X.-F. & Wang, H. (2014) 'Monocyte and macrophage differentiation: circulation inflammatory monocyte as biomarker for inflammatory diseases', *Biomarker Research*, 2p. 1.

Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G. & Chen, L.-L. (2011) 'Genomewide characterization of non-polyadenylated RNAs', *Genome Biology*, 12(2), p. R16.

Yang, L., Wilusz, J.E. & Chen, L.-L. (2022) 'Biogenesis and Regulatory Roles of Circular RNAs', *Annual Review of Cell and Developmental Biology*, 38pp. 263–289.

Yang, Y. & Wang, Z. (2019) 'IRES-mediated cap-independent translation, a path leading to hidden proteome', *Journal of Molecular Cell Biology*, 11(10), pp. 911–919.

Yang, Yun, Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Yi, Chen, L.-L., Wang, Y., Wong, C.C., Xiao, X. & Wang, Z. (2017) 'Extensive translation of circular RNAs driven by N6-methyladenosine', *Cell Research*, 27(5), pp. 626–641.

Yarnall, A.J., Breen, D.P., Duncan, G.W., Khoo, T.K., Coleman, S.Y., Firbank, M.J., Nombela, C., Winder-Rhodes, S., Evans, J.R., Rowe, J.B., Mollenhauer, B., Kruse, N., Hudson, G., Chinnery, P.F., O'Brien, J.T., Robbins, T.W., Wesnes, K., Brooks, D.J., Barker, R.A., et al. (2014) 'Characterizing mild cognitive impairment in incident Parkinson disease: the ICICLE-PD study', *Neurology*, 82(4), pp. 308–316.

You, X., Vlatkovic, I., Babic, A., Will, T., Epstein, I., Tushev, G., Akbalik, G., Wang, M., Glock, C., Quedenau, C., Wang, X., Hou, J., Liu, H., Sun, W., Sambandan, S., Chen, T., Schuman, E.M. & Chen, W. (2015) 'Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity', *Nature Neuroscience*, 18(4), pp. 603–610.

Yu, E., Ambati, A., Andersen, M.S., Krohn, L., Estiar, M.A., Saini, P., Senkevich, K., Sosero, Y.L., Sreelatha, A.A.K., Ruskey, J.A., Asayesh, F., Spiegelman, D., Toft, M., Viken, M.K., Sharma, M., Blauwendraat, C., Pihlstrøm, L., Mignot, E. & Gan-Or, Z. (2021) 'Fine mapping of the HLA locus in Parkinson's disease in Europeans', *npj Parkinson's Disease*, 7(1), pp. 1–7.

Zaphiropoulos, P.G. (1996) ‘Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping.’, *Proceedings of the National Academy of Sciences*, 93(13), pp. 6536–6541.

Zaphiropoulos, P.G. (1997) ‘Exon Skipping and Circular RNA Formation in Transcripts of the Human Cytochrome P-450 2C18 Gene in Epidermis and of the Rat Androgen Binding Protein Gene in Testis’, *Molecular and Cellular Biology*, 17(6), pp. 2985–2993.

Zeighami, Y., Fereshtehnejad, S.-M., Dadar, M., Collins, D.L., Postuma, R.B. & Dagher, A. (2019) ‘Assessment of a prognostic MRI biomarker in early de novo Parkinson’s disease’, *NeuroImage: Clinical*, 24p. 101986.

Zeighami, Y., Ulla, M., Iturria-Medina, Y., Dadar, M., Zhang, Y., Larcher, K.M.-H., Fonov, V., Evans, A.C., Collins, D.L. & Dagher, A. (2015) ‘Network structure of brain atrophy in de novo Parkinson’s disease’ David C Van Essen (ed.), *eLife*, 4p. e08440.

Zeis, T., Enz, L. & Schaeren-Wiemers, N. (2016) ‘The immunomodulatory oligodendrocyte’, *Brain Research*, 1641pp. 139–148.

Zetusky, W.J., Jankovic, J. & Pirozzolo, F.J. (1985) ‘The heterogeneity of Parkinson’s disease’, *Neurology*, 35(4), pp. 522–522.

Zhang, J., Chen, S., Yang, J. & Zhao, F. (2020) ‘Accurate quantification of circular RNAs identifies extensive circular isoform switching events’, *Nature Communications*, 11(1), p. 90.

Zhang, J., Zhang, X., Li, C., Yue, L., Ding, N., Riordan, T., Yang, L., Li, Y., Jen, C., Lin, S., Zhou, D. & Chen, F. (2019) ‘Circular RNA profiling provides insights into their subcellular distribution and molecular characteristics in HepG2 cells’, *RNA Biology*, 16(2), pp. 220–232.

Zhang, S.-Y., Clark, N.E., Freije, C.A., Pauwels, E., Taggart, A.J., Okada, S., Mandel, H., Garcia, P., Ciancanelli, M.J., Biran, A., Lafaille, F.G., Tsumura, M., Cobat, A., Luo, J., Volpi, S., Zimmer, B., Sakata, S., Dinis, A., Ohara, O., et al. (2018) ‘Inborn Errors of RNA Lariat Metabolism in Humans with Brainstem Viral Infection’, *Cell*, 172(5), pp. 952-965.e18.

Zhang, X.-O., Dong, R., Zhang, Y., Zhang, J.-L., Luo, Z., Zhang, J., Chen, L.-L. & Yang, L. (2016) ‘Diverse alternative back-splicing and alternative splicing landscape of circular RNAs’, *Genome Research*, 26(9), pp. 1277–1287.

Zhang, X.-O., Wang, H.-B., Zhang, Y., Lu, X., Chen, L.-L. & Yang, L. (2014) ‘Complementary Sequence-Mediated Exon Circularization’, *Cell*, 159(1), pp. 134–147.

Zhang, Y., Xue, W., Li, X., Zhang, J., Chen, S., Zhang, J.-L., Yang, L. & Chen, L.-L. (2016) ‘The Biogenesis of Nascent Circular RNAs’, *Cell Reports*, 15(3), pp. 611–624.

Zhang, Y., Zhang, X.-O., Chen, T., Xiang, J.-F., Yin, Q.-F., Xing, Y.-H., Zhu, S., Yang, L. & Chen, L.-L. (2013) ‘Circular Intronic Long Noncoding RNAs’, *Molecular Cell*, 51(6), pp. 792–806.

Zhao, S., Zhang, Y., Gamini, R., Zhang, B. & von Schack, D. (2018) ‘Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA depletion’, *Scientific Reports*, 8(1), p. 4781.

Zhao, W., He, X., Hoadley, K.A., Parker, J.S., Hayes, D.N. & Perou, C.M. (2014) 'Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling', *BMC Genomics*, 15(1), p. 419.

Zhao, Y., Lai, Y., Konijnenberg, H., Huerta, J.M., Vinagre-Aragon, A., Sabin, J.A., Hansen, J., Petrova, D., Sacerdote, C., Zamora-Ros, R., Pala, V., Heath, A.K., Panico, S., Guevara, M., Masala, G., Lill, C.M., Miller, G.W., Peters, S. & Vermeulen, R. (2024) 'Association of Coffee Consumption and Prediagnostic Caffeine Metabolites With Incident Parkinson Disease in a Population-Based Cohort', *Neurology*, 102(8), p. e209201.

Zhao, Y., Ray, A., Portengen, L., Vermeulen, R. & Peters, S. (2023) 'Metal Exposure and Risk of Parkinson Disease: A Systematic Review and Meta-Analysis', *American Journal of Epidemiology*, 192(7), pp. 1207–1223.

Zhong, L., Ju, K., Chen, A. & Cao, H. (2021) 'Circulating CircRNAs Panel Acts as a Biomarker for the Early Diagnosis and Severity of Parkinson's Disease', *Frontiers in Aging Neuroscience*, 13.

Zhong, Y., Kinio, A. & Saleh, M. (2013) 'Functions of NOD-Like Receptors in Human Diseases', *Frontiers in Immunology*, 4.

Zhou, A., Molinaro, R.J., Malathi, K. & Silverman, R.H. (2005) 'Mapping of the Human RNASEL Promoter and Expression in Cancer and Normal Cells', *Journal of Interferon & Cytokine Research*, 25(10), pp. 595–603.

Zhou, B., Wen, M., Yu, W.-F., Zhang, C.-L. & Jiao, L. (2015) 'The Diagnostic and Differential Diagnosis Utility of Cerebrospinal Fluid α-Synuclein Levels in Parkinson's Disease: A Meta-Analysis', *Parkinson's Disease*, 2015(1), p. 567386.

Zhu, Y., Wang, L., Yin, Y. & Yang, E. (2017) 'Systematic analysis of gene expression patterns associated with postmortem interval in human tissues', *Scientific Reports*, 7(1), p. 5435.

Zubelzu, M., Morera-Herreras, T., Irastorza, G., Gómez-Esteban, J.C. & Murueta-Goyena, A. (2022) 'Plasma and serum alpha-synuclein as a biomarker in Parkinson's disease: A meta-analysis', *Parkinsonism & Related Disorders*, 99pp. 107–115.