DETAILED RESEARCH PROPOSAL

Blood based epigenome-wide association study in Parkinson's disease: a search for population specific biomarkers and novel therapeutic targets

Summary: Aberrant DNA methylation plays a vital role in the onset and progression of Parkinson's disease (PD), a progressive neurodegenerative movement disorder with multifactorial aetiology. Despite several large-scale studies being conducted to characterise the integrated genetic and epigenetic landscape of PD, no studies have explored DNA methylation alterations in PD in the Indian subcontinent. As genetic heterogeneity has already been reported to vary in the Indian population, characterising blood DNA methylomic signatures, which might also differ between populations, will be useful in identifying population-specific biomarkers of disease progression, severity, and treatment response in a minimally invasive manner.

Background

Parkinson's disease (PD) is a progressive neurodegenerative movement disorder characterized by the loss of dopaminergic neurons in the *substantia nigra*. Extensive dopaminergic neuron loss occurs prior to PD diagnosis, and accurate diagnosis can only be achieved through post-mortem neuropathological analysis; therefore, identification of early pathological changes is paramount to target therapeutic treatments before major neuropathological damages occur (Horvath et al, 2015). DNA methylation involves the addition of a methyl group onto the C5 position of the cytosine residue to form 5-methylcytosine, and is responsible for regulating gene expression (Moore et al, 2013). Several DNA methylation (DNAm) biomarkers of cancer classification including those for brain-tumours already exist, and DNAm biomarkers for neurodegenerative diseases hold great promise. Concordant methylation alterations previously identified in brain and blood suggest that blood differential methylation might be a valuable surrogate for brain tissue in PD and can serve as an easily accessible and quantifiable biomarker to monitor disease progression and treatment responses (Masliah et al, 2013; Murthy et al, 2021). Previous studies have also demonstrated associations of DNAm alterations in blood with PD status, progression, and changes in blood cell composition (Henderson-Smith et al., 2019).

Unlike in the Caucasian population, well defined patient cohorts for PD are scarce in India despite a widespread prevalence (~0.77 million in 2019) (Singh et al, 2021). A few studies have explored genetic risk factors of PD, the most recent being the Gap-India project, and have demonstrated limited reproducibility in terms of risk-loci, with prominent variations identified in the Caucasian datasets posing negligible risk in the Indian population (Rajan et al, 2020). I hypothesize that given the genetic heterogeneity in the Indian subcontinent, similar to the genetic risk-loci, epigenetic drivers of PD pathology will also vary, thus necessitating population-specific studies. The objective of this study, therefore, would be to create a set of population specific DNAm biomarkers as well as to identify common global markers that can be used in the identification of disease onset and progression in a minimally invasive manner, which could be performed as a routine test on PD patients to evaluate and monitor PD progression, and on individuals at risk (family members), to efficiently combat PD.

Objectives

1. **Perform genome-wide DNA methylation (DNAm) profiling in the blood of individuals with PD and neurologically healthy controls**
    a. to identify population-specific DNAm alterations in PD as well as alterations shared between populations by comparing the generated dataset with publicly available datasets from other populations
    b. to investigate associations of these DNAm changes with disease traits, severity, and progression
    c. to investigate the presence of epigenetic age acceleration in PD
2. **Validate the DNA methylation alterations in independent samples**
    - using pyrosequencing (or methylation-sensitive restriction/qPCR)
3. **Investigate the regulatory consequences of the DNA methylation alterations identified**

a. on gene expression using transcriptomic analyses and RT-qPCR
b. on protein expression using proteomic analyses and western blotting

4. **Identify concordant signatures between blood and brain tissues**
   a. by comparing DNA methylation alterations in the blood with data from existing brain epigenome-wide association studies (EWAS)
   b. by exploring the regulatory consequences of the concordant alterations through gene (RT-qPCR) and protein (western blotting and immunohistochemistry) expression analyses in frozen and formalin fixed paraffin embedded (FFPE) brain tissues

Methodology

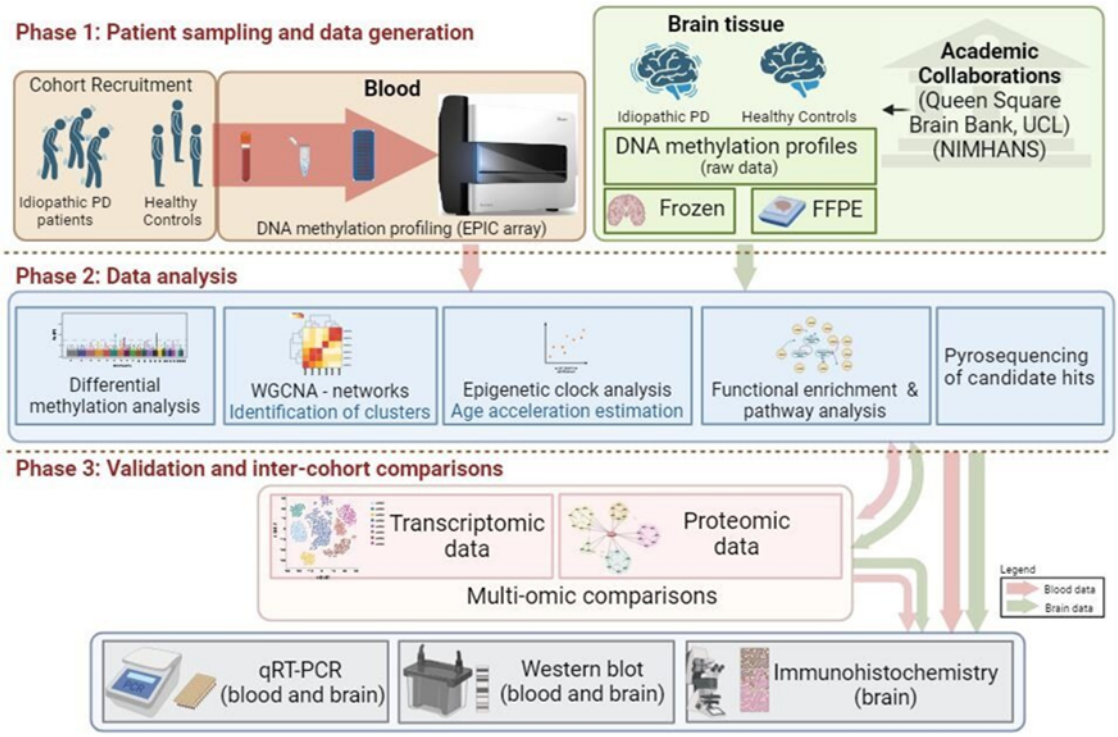

Figure 1: Overview of proposed project (Created using Biorender).

DNA methylation (DNAm) profiling and analysis (Aim 1): Whole blood collected from clinically diagnosed idiopathic PD (iPD) patients (n=50) and age and sex matched neurologically healthy controls (n=50) (collected in collaboration with neurologist) will be subjected to DNA extraction followed by bisulfite conversion and generation of DNAm profiles using Illumina EPIC array containing ~850,000 methylation sites across the genome (discovery cohort). According to the online "EPIC Array Power Calculations" tool (https://epigenetics.essex.ac.uk/shiny/EPICDNAmPowerCalcs/), a study with 50 cases and 50 controls (total n=100 samples) is well powered (>80%) to detect a 5% mean difference at a genome-wide significance (P<9.42x10-8) in 62.1% (~527,850) of the ~850,000 EPIC array CpG sites.

Raw data obtained will be subjected to a series of pre-processing and quality checks using several R and Bioconductor packages using an already established data analysis pipeline (**Fig 2**) [during my postdoc at the Queen Square Brain Bank (QSBB), Institute of Neurology, University College London (UCL-QSIoN), UK]; the processed data will be used to perform differential methylation analysis and functional enrichment to identify molecular pathways (Aim-1a) (Murthy et al, 2023(a)). Using a systems biology approach, clusters of highly correlated methylation sites will be identified by creating weighted gene co-methylation networks, and these clusters will be correlated to phenotypic traits such as disease onset and duration to dissect signatures indicative of disease severity (Aim-1b) (Fodder et al, 2023). DNAm ages will be determined using multi-tissue and blood-specific epigenetic clocks to determine if PD patients show epigenetic age acceleration compared to controls (Aim-1c) (Murthy et al, 2023(b)). In addition, presence of these differentially methylated probes (DMPs) in the global populations will be investigated by comparing with other publicly available blood EWAS (e.g., GSE165081, PPMI datasets) to identify common global markers as well as population-specific signatures.

Validation of the significant DNAm alterations (Aim 2): An additional independent replication cohort of iPD (n=100) and control (n=100) individuals will be recruited (validation cohort). Primers specific to the top significant DMPs identified in Aim 1 will be designed and pyrosequencing will be performed (**Fig 2)**.

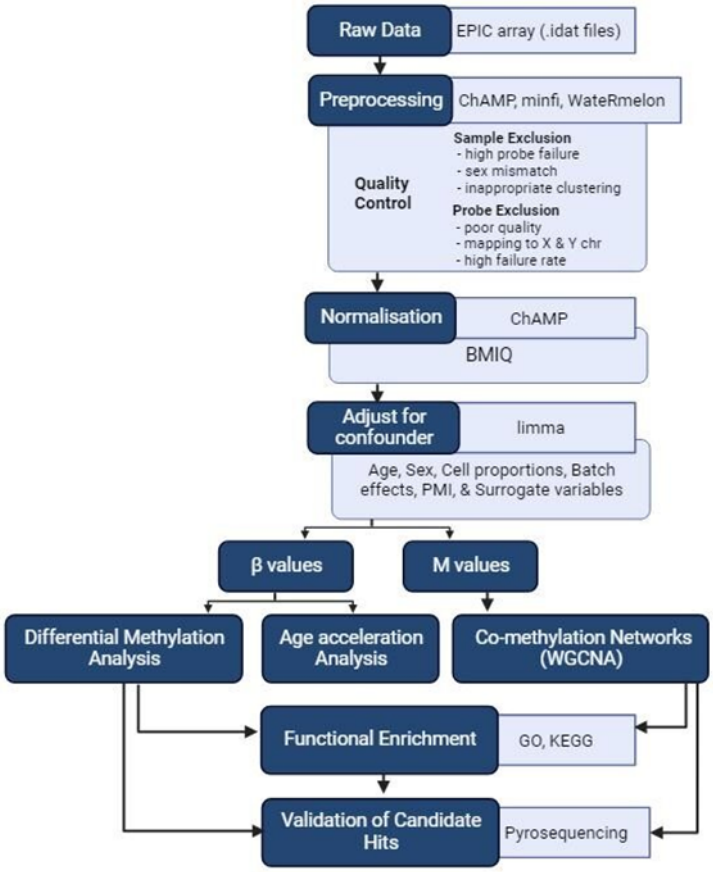

Figure 2: Detailed workflow of methodology to be employed for Aims 1 and 2.

Analysis of regulatory consequences of the alterations (Aim 3): Regulatory effects of the DNAm alterations will be investigated by performing RNA extraction from blood followed by gene expression (RT-qPCR) analysis using transcript-specific primers in both the discovery and validation cohorts. Additionally, presence of these gene expression differences will also be explored in publicly available transcriptomic datasets (e.g., GSE165082, PPMI). Similarly, complementary changes in protein expression will also be examined using protein extraction and western blotting for the proteins of the candidate hits (**Fig 3**).

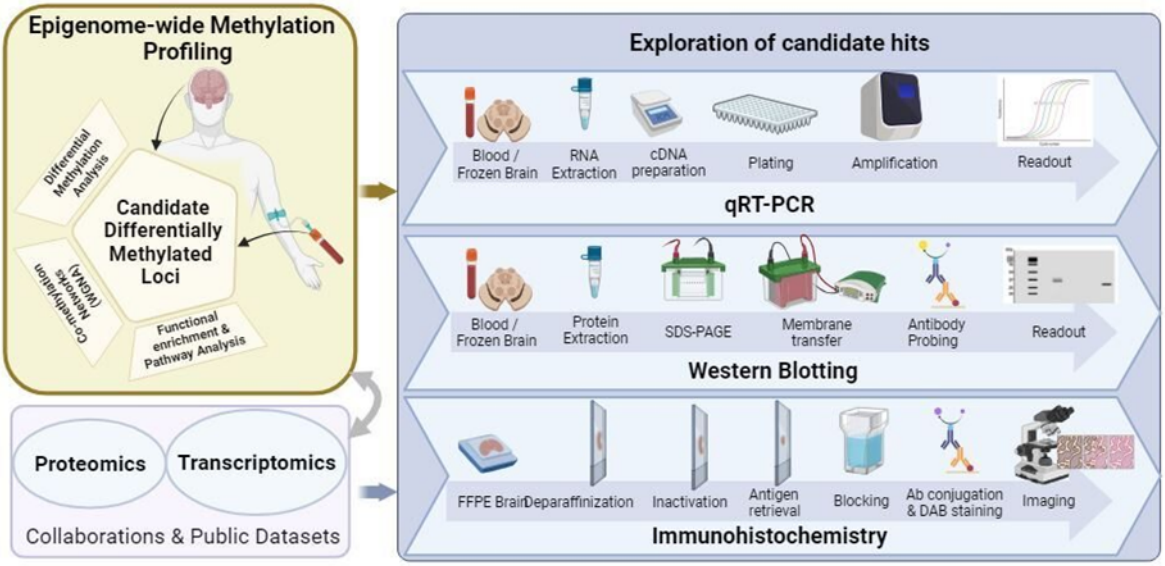

Figure 3: Detailed methodology for Aims 3 and 4 (created using Biorender).

Investigating concordant signatures between blood and brain tissues (Aim 4): Blood differential methylation will be compared with that of brain tissues using frontal lobe white matter DNAm profiles [from PD cases (n=17); controls (n=17)] previously generated by Dr. Conceicao Bettencourt (collaborator) and using a publicly available frontal cortex grey matter dataset [PD (n=139); controls (n=73)] from the Netherlands Brain Bank and Normal Aging Brain Collection, Amsterdam. Gene and protein expression changes will be explored in the genes showing concordant changes in the brain by RT-qPCR, western blotting, and immunohistochemistry using frozen and FFPE brain tissues provided by Prof. Tammaryn Lashley (collaborator) at QSBB, UCL, UK and comparing them with tissues obtained from the Human Brain Tissue Repository, NIMHANS, Bengaluru, India (**Fig 3**).

Impact statement in relation to current national and international status

This study will be the first population study, to my knowledge, to perform DNAm profiling in the blood of PD patients in India by leveraging the high-throughput EPIC array platform. The study will involve careful matching of population-based PD cases and suitable controls, account for possible confounders including blood-cell compositions, and perform rigorous control of false positives resulting from multiple comparisons (Bonferroni correction). In addition, my previous work with Dr. Bettencourt at QSBB identified several DNAm alterations in frontal lobe white matter and grey matter (public dataset) of PD cases, this data would be useful in comparing alterations in the blood to that in the brain and aid in dissecting pathogenic mechanisms and molecular pathways.

As DNAm is a dynamic process that can be influenced by physiological and environmental cues, studying blood DNAm is an excellent way to continually identify and measure cues corresponding to disease progression and monitor molecular changes as a proxy to measure efficiencies of different therapeutic modalities. Therefore, I plan to execute the objectives of the study using my experience in omic technologies and bioinformatics along with the mentorship, expertise, and support of collaborators Dr. Bettencourt (bioinformatics), Prof. Lashley (neuropathology), and Prof. Patrick Lewis, Royal Veterinary College, London (protein biology) with an overarching aim to ultimately discover novel biomarkers and therapeutic targets to be explored using gene-editing technologies in future projects.

Expected output and outcome of the proposed research work

Expected outcomes of the project include

i. discovery of robust **population-specific molecular signatures** of PD and identification of signatures shared with other populations to arrive at a set of biomarkers specifically designed for PD **diagnosis, progression, and treatment response** in the Indian subcontinent (Aims 1, 2)
ii. finding **concordant changes in other molecular markers of blood** by studying downstream regulatory changes of the alterations (Aim 3)
iii. **dissecting the molecular mechanisms** of PD pathophysiology by comparing DNA methylation alterations in the blood and brain and **finding targetable genes** that can be used **to design new therapies as well as in drug repurposing** in future projects.

Epigenomic mechanisms of neurodegenerative diseases in humans in the Indian scenario are heavily understudied and have remained largely unexplored. The proposed project will be an excellent platform to lay a foundation towards establishing collaborations with consortiums in India such as the GAP-India project (~10,000 individuals), in addition to global consortiums like Parkinson's Progression Markers Initiative (PPMI), Parkinson's Genetics Program (GP2) and the Aligning Science Across Parkinson's (ASAP) projects, to design comprehensive biomarker panels. These can further be put to use in Indian cohorts of aging to identify at-risk individuals and determine the severity and rate of progression in these individuals.

APPENDIX

Timeline for the proposed project including feasibility of individual aims and contingency plans.

Blood based epigenome-wide association study in Parkinson's disease: a search for population-specific biomarkers and novel therapeutic targets											
Objective	**Year 1**		**Year 2**		**Year 3**		**Year 4**		**Year 5**		**Feasibility**
	1-6 months	7-12 months	1-6 months	7-12 months	1-6 months	7-12 months	1-6 months	7-12 months	1-6 months	7-12 months	
Genome-wide DNA methylation profiling in PD blood	Blood sample collection (DNAm profiling) [iPD (n=50); CTR (n=50)]										**Sample collection** - Collaboration with Neurologist; **acheivable sample size**
			Analysis of the DNAm profiles generated								**DNAm profiling** - Previous experience present, hence low risk of failing, even if fails, publicly available data can be leveraged to shortlist candidate hits for Aims2-4; **Data analysis** - Previous experience present, pipelines already established
	Analysis of publicly available datasets (Blood-EWAS)				Compare datasets						
Validating candidate DNA methylation alterations	Blood sample collection (validation cohort) [iPD (n=100); CTR (n=100)]										**Sample collection** - subject to availability of PD cases; **Pyrosequencing** - Low cost, troubleshooting possible or alternative methods (PCR) can be used
					Pyrosequencing						
Investigating regulatory consequences of the DNA methylation alterations identified							Comparison with transcriptomic & proteomic data (Publicly available)				**Transcriptomic and proteomic datasets**- Publicly available, hence no risk;
								RT-qPCR & WB of candidate hits in blood			**RT-qPCR & WB** - Experienced in experimental & analysis methods (No risk)
Identification of concordant signatures between blood and brain tissues					Brain sample collection (Frozen & FFPE)						**Brain tissue collection** - Collaborations already established with QSBB, UCL, UK - Consented to ship tissues and share data (No risk); HBTR, NIMHANS - to be aprroached
					Analysis of publicly available datasets (Brain-EWAS) / Blood-Brain concordance						**Public datasets:** Preliminary data already analysed (EWAS)/readily available (No risk)
							Comparison with transcriptomic & proteomic data (Publicly available)				
								RT-qPCR, WB, & IHC of concordant hits in brain tissues			**RT-qPCR, WB, IHC** - Experienced in experimental & analysis methods (No risk)

iPD - idiopahic Parkinson's disease, CTR - control, EWAS - epigenome-wide associaton study, RT-qPCR - reverse transcriptase qualitative polymerase chain reaction, WB - western blotting, FFPE - formalin fixed paraffin embedded, IHC - immunohistochemistry, QSBB - Queen Square Brain Bank, UCL - University College London, HBTR - Human Brain Tissue Repository, NIMHANS - National Institute of Mental Health and NeuroScience, Bengaluru

REFERENCES

Henderson-Smith, A, et al. "DNA methylation changes associated with Parkinson's disease progression: outcomes from the first longitudinal genome-wide methylation analysis in blood." Epigenetics 14.4 (2019): 365-382.

Horvath, S, and Beate R.R. "Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients." Aging (Albany NY) 7.12 (2015): 1130.

Masliah, E, et al. "Distinctive patterns of DNA methylation associated with Parkinson disease: identification of concordant epigenetic changes in brain and peripheral blood leukocytes." Epigenetics 8.10 (2013): 1030-1038.

Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013 Jan;38(1):23-38

Murthy M, Cheng YY, Holton JL, Bettencourt C. Neurodegenerative movement disorders: An epigenetics perspective and promise for the future. Neuropathol Appl Neurobiol. 2021 Dec;47(7):897-909. doi: 10.1111/nan.12757.

Murthy M, Shireby G, Miki Y, Viré E, Lashley T, Warner TT, Mill J, Bettencourt C. Epigenetic age acceleration is associated with oligodendrocyte proportions in MSA and control brain tissue. Neuropathol Appl Neurobiol. 2023 (a) Feb;49(1): e12872. doi: 10.1111/nan.12872.

Murthy M, Rizzu P, Heutink P, Mill J, Lashley T, Bettencourt C. Epigenetic Age Acceleration in Frontotemporal Lobar Degeneration: A Comprehensive Analysis in the Blood and Brain. Cells. 2023 (b) Jul 24;12(14):1922.

Fodder K, **Murthy M,** Rizzu P, Toomey CE, Hasan R, Humphrey J, Raj T, Lunnon K, Mill J, Heutink P, Lashley T, Bettencourt C. Brain DNA methylomic analysis of frontotemporal lobar degeneration reveals OTUD4 in shared dysregulated signatures across pathological subtypes. Acta Neuropathol. 2023 Jul;146(1):77-95.

Perlmutter, J. S. Assessment of Parkinson disease manifestations. Curr Protoc Neurosci. 2009; Chapter 10, Unit 10 11.

Rajan, R, et al. "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 (2020): 524.

Singh, G, et al. "The burden of neurological disorders across the states of India: the Global Burden of Disease Study 1990–2019." The Lancet Global Health 9.8 (2021): e1129-e1144.