# Elucidating Molecular Networks Underpinning Heterogeneity in Parkinson's Disease Progression Across Clinical Manifestation Spectrum

Manqi Zhou[1,2], Alison Ke[1,2], Xingbo Wang[2,3], Kun Chen[4], Fei Wang[2,3*], Chang Su[2,3*]

[1]Department of Computational Biology, Cornell University, Ithaca, New York, 14853, USA.
[2]Institute of Artificial Intelligence for Digital Health, Weill Cornell Medicine, New York, New York, 10021, USA.
[3]Division of Health Informatics, Department of Population Health Sciences, Weill Cornell Medicine, New York, New York, 10021, USA.
[4]Department of Statistics, University of Connecticut, Storrs, Connecticut, 06269, USA.

*Corresponding author(s). E-mail(s): few2001@med.cornell.edu; chs4001@med.cornell.edu;

**Abstract**

In this study, we applied statistical and machine learning techniques to identify molecular mechanisms underlying the heterogeneity in individual Parkinson's Disease (PD) progression. Leveraging data from the Parkinson's Progression Markers Initiative (PPMI) cohort, we analyzed genetic and clinical data for patients with PD, focusing on traits including motor symptoms, non-motor symptoms, and biomarkers. Our method identified significant single-nucleotide polymorphisms (SNPs) associated with each PD trait, revealing key genetic factors and their impact on disease progression. Furthermore, through network medicine approaches, we delineated disease modules, uncovering unique gene clusters and their roles in PD pathology. The integration of pathway enrichment analysis further enhanced our understanding of the functional implications of these genetic variations, notably highlighting the significance of cellular stress response and protein aggregation pathways in PD. Overall, our findings offer a comprehensive view of the genetic landscape of PD progression, highlighting the potential of personalized medicine in managing this complex disease.

## 1 Introduction

Parkinson' disease (PD) is the 2nd most prevalent age-related neurodegenerative disease, affecting around 1 million people in the United States[1]. The severity of disability in PD patients has been proven to be parallel with the course of disease progression plus their refractory responses to treatment, accompanying with symptoms of postural and ambulatory instability, speech and swallowing dysfunctions, autonomic failure, dementia, psychological symptoms, among others[2–5]. Additionally, it has been increasingly recognized that these diverse manifestation dimensions progress distinctly from patient to patient, yielding differentiated progression trajectories throughout PD course[6]. Investigating biological underpinnings that govern such heterogeneous progression procedures are crucial for better understanding of PD etiology and importantly, might suggest novel therapeutic targets for preventing or slow down PD progression.

Prior genome-wide association studies (GWASs) in PD have discovered a collection of genetic loci associated with PD[7–9]. Some recent studies have also shifted their focus towards identifying genetic factors of PD progression. For instance, Manuela et al. conducted a GWAS study with a focus on motor and cognition progression in PD[10]. They found that, rs429358, known as the APOE E4 tagging variant, is associated with cognition progression in PD; while the *ATP8B2* gene is associated with motor progression. Using a hypothesis-free machine learning approach, a PD specific novel genetic interaction between an intronic LINGO2 (9p21.1) SNP (rs9298897) and the 2q14.1 variant (rs17710829) was identified in the Parkinson's Progression Markers Initiative (PPMI) and the Longitudinal and Biomarker Study in Parkinson's disease (LABS-PD) cohorts[11]. The LRRK2 G2019S mutation has been associated with a slower progression in motor manifestation[12], and

the an interaction between LRRK2 G2019S and GBA has been found in cognition deline in PD[13]. Though significant progress has been made, these existing works generally focus on single or a few progression traits of PD. There remains the need of systematic investigation of biological bases of full spectrum of PD progression traits.

Additionally, translating genetic findings to therapeutic development is challenging due to the limited impact size of single mutation[7, 14]. Most molecular components, including genes, don't function along; rather, they interact with some others to contribute to a specific biochemical process or disease[15, 16]. There has been the "disease module hypothesis" – cellular components (e.g., genes) linked to a specific disease phenotype show a tendency to cluster in the same network neighborhood. A gene module is a group of network components that together contribute to a molecular function and disruption of which results in a particular phenotype[15, 17]. Systematic identification of such gene modules paves the way for an in-depth understanding of the biological underpinnings of traits, facilitates the discovery of network-targeting therapeutics, and enhances the precision of clinical (endo)phenotype[15, 17, 18].

In this study, we conducted a comprehensive analysis to investigate the associations between genetic factors and progression rates of a wide spectrum of clinical manifestations in PD. Using a network medicine approach, we aimed at discovery of gene modules associated with progression rates in different PD traits, revealing biological bases governing the heterogeneous progression procedures. We posited that systematic identification of underlying progression trait-specific modules would provide a foundation for generating predictive models to characterize PD pathogenesis for efficacious therapeutic development to slow down PD progression.

# 2 Results

## 2.1 Study Cohort

In compliance with the PPMI Data Use Agreement, we utilized data from the PPMI cohort, which was accessed in September 2023. The data included in our analysis were derived from individuals diagnosed with PD who met specific criteria: possession of genetic data, a minimum of two clinical visits, and an inter-visit period exceeding six months (**Methods**). A total of 29 variables across eight domains were examined: demographic, medicine, motor, cognition, mood, sleep, autonomic, and biomarker (**Methods**). In the demographic domain, the mean age of the cohort was found to be 59.7 years with a standard deviation of 9.76. The sample was predominantly male (58.2%) and white (96.4%). The average duration of PD symptoms among the participants was 38.77 months, with 29% having at least one family member diagnosed with PD. Within the motor domain, the average scores for the Unified Parkinson's Disease Rating Scale Part II (UPDRSII) and Part III (UPDRSIII), as well as the Schwab and England Activities of Daily Living Scale, were recorded as 9.11, 24.84, and 87.07 respectively, with standard deviations of 6.54, 11.85, and 11.68. Detailed information for other domains is presented in **Table** 1.

## 2.2 Identification of SNPs associated with single PD progression traits

We first investigated genetic factors associated with progression rates of different PD progression traits. Specifically, we focused on a total of 21 PD traits, including motor manifestations (International Parkinson's disease and Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS)-Part II and III, tremor score, postural instability–gait difficulty (PIGD) score[19], and Schwab-England Activities of daily living score), general non-motor assessment (MDS-UPDRS Part I), cognitive decline (Montreal Cognitive Assessment (MoCA)[20], Benton Judgment of Line Orientation (Bento)[21], Symbol Digit Modalities Test (symbol digit)[22], Hopkins Verbal Learning Test (HVLT)[23] for total recall, delayed recall, retention and recognition-discrimination, and semantic verbal-language fluency test (semantic fluency)[24] and letter-number sequencing (LNS)[25]), mood assessment (Geriatric Depression Scale (GDS) score[26], State-Trait Anxiety Inventory (STAI) score[27]), sleep assessment (REM Sleep Behavior Questionnaire (RBD)[28], Epworth Sleepiness Score (ESS)[29]), autonomic assessment (Scales for Outcomes in Parkinson's Disease-Autonomic questionnaire (SCOPA)[30]), blood biomarkers (p tau181p, total tau, abeta 42, alpha syn), and a global assessment (Global Composite Outcome (GCO)[31]).

Identifying statistically and biologically significant signals from whole genome is challenging due to small effect size of single mutation and limited sample size of individuals with well-phenotyped progression information of PD. Therefore, we focused on 2,972 loci reported in previous GWAS of PD, PD progression, other neurodegenerative disorders, including Alzheimer's disease, Huntington's disease, dementia, and Lewy body dementia, as well as PD-related pathology like alpha-synuclein measurement, collected from GWAS Catalog[32]. For each clinical trait and each SNP, a linear mixed-effect model was fitted using up to 13-years follow-up data. Specifically, each clinical score was regressed on a specific SNP and SNP$\times$Time terms and a set of covariates including age, sex, race, family history, symptom duration, education level (for cognition trait only), levodopa equivalent daily dose (LEDD), and LEDD$\times$Time. We reported coefficient and P-value

of the SNP×Time term, indicating the association between progression rate and the SNP. Multiple correction was conducted based on FDR.

As a result, we identified an average of 157 significant SNPs for each trait (**Table** 2, **Supplementary Table** 1). Specifically, we found 141 significant SNPs associated with the UPDRSII based on P values. To visually represent these findings, we plotted the changes in trait scores over time, contrasting SNP-carrier patients with non-SNP-carrier patients (**Fig.** 1). For example, the UPDRSII score changes between patients carrying rs190780914 mutation and patients not carrying the mutation are at different speed rates. This difference could be attributed to the gene *MARK2P15*, mapped to rs190780914 and functions as a pseudogene of *MARK2*. *MARK2* has been identified as an activating kinase for *PINK1*, and *PINK1* plays a crucial role in maintaining mitochondrial health. The loss of *PINK1* function due to mutations can lead to impaired mitochondrial function. This impairment can contribute to the degeneration of dopaminergic neurons in the substantia nigra, which in turn leads to the characteristic motor symptoms[33].

### 2.3 Discovery of gene modules driving single PD progression traits using network medicine

We next turned to expand the genetic findings to identify gene modules associated with progression rate of each PD trait. Specifically, we used the Genome-wide Positioning Systems network (GPSnet) algorithm[17], a network medicine approach that greedily constructing gene module of a disease phenotype by simultaneously enhancing gene connectivity within module and module-level risk score. Here, we initialized gene risk score based on coefficients and P values derived in genetic analysis. See detailed information in the **Methods**section.

As a result, we identified gene modules of varying sizes corresponding to each clinical trait. Specifically, we delineated a gene module related to the MoCA comprising 232 genes, which included notable genes such as *CCNT2* and *MEX3C*. Additionally, we combined gene modules associated with all motor traits to form a comprehensive motor gene module, which encompassed 449 genes, including *PINK1* and *PIMS1* (**Supplementary Table** 2). The interrelationships within the overall network were illustrated in **Fig.** 2a, with distinct colors representing different traits and purple color denoting areas of trait overlap. To further analyze the interconnectivity between gene modules, we employed the Jaccard Index[34] (**Fig.** 2b). Our analysis revealed that gene modules associated with traits within the same category exhibited a higher degree of overlap.

In total, 26 genes were identified as overlapping across all clinical traits, with some exhibiting a broad association with Parkinson's Disease, and others demonstrating trait specificity (**Fig.** 2c). Genes such as *GBA*, *APOE*, *LRRK2*, and *FYN* were ubiquitous across all traits, indicating a general role in PD pathology. Conversely, genes like *PINK1* were specific to motor features, *MCCC1* to mood features, and *ITPKB* to sleep features. Notably, mutations in the *MCCC1* gene are known to cause an autosomal recessive metabolic disorder affecting leucine metabolism[35]. Research has shown that leucine supplementation may positively influence depressive states and reduce social avoidance behaviors in animal models[36], underscoring its potential relevance in mood-related aspects of PD.

### 2.4 Enriched pathways associated with single PD progression traits

To elucidate the functional mechanisms of the genes and gene modules identified, we conducted pathway enrichment analysis utilizing g:Profiler[37]. This analysis incorporated data from established pathway databases including Gene Ontology (GO)[38], KEGG Pathway[39], and Reactome Pathway[40]. To refine the results and reduce redundancy, we employed tools GOxploreR[41] and REVIGO[42], which streamlined the set of GO terms by removing those with overlapping biological information (**Methods**).

The analysis revealed an average of 668 significantly enriched pathways for each trait (**Supplementary Table** 3). Specifically, 947 pathways were enriched for UPDRSII and 906 for MoCA. These pathways commonly pertained to functions implicated in Parkinson's Disease pathogenesis, including cellular stress responses, immune inflammation, and protein misfolding (**Fig.** 3). For instance, both MoCA and UPDRSII exhibited enrichment in cell-stress related pathways, such as cellular responses to stress/stimuli as highlighted in orange in **Fig.** 3. Oxidative stress, Oxidative stress, a specific type of cellular stress, is known to play a critical role in the degeneration of dopaminergic neurons in PD[43]. Moreover, protein aggregation related pathways (highlighted in green in **Fig.** 3) like ubiquitin protein ligase binding, tau protein binding, peptidyl-serine modification, and protein stabilization, were identified. The impairment of Parkin's ubiquitin–protein ligase function in certain familial-linked mutations has been suggested as a potential cause of familial autosomal recessive Parkinson's disease[44].

## 3 Conclusion

We yielded significant insights into the genetic and clinical landscape of Parkinson's Disease progression heterogeneity. By identifying key SNPs associated with PD progression traits and constructing gene modules

through network medicine approaches, we have revealed multifaceted genetic influence on PD pathology. Genes with both broad associations and trait-specific implications, such as *PINK1* for motor features and *MCCC1* for mood features, highlight the intricate genetic interplay within PD progression. Furthermore, the discovery of enriched pathways, especially those related to cellular stress responses and protein misfolding, aligns with and expands our current understanding of PD pathogenesis.

The implications of our findings are profound for the future of PD research and treatment. The diverse genetic factors and pathways identified offer potential targets for therapeutic intervention and deepen our understanding of the disease's progression. By integrating genetic data with clinical traits and employing advanced analytical tools, we can tailor treatment strategies based on an individual's genetic profile. Overall, this study not only enhances our understanding of PD but also serves as a model for investigating the genetic basis of other complex diseases.

# 4 Method

## 4.1 Study cohort and data preprocess

In PPMI, a comprehensive set of clinical features is assessed. We used the following data (**Table** 1):
(1) Demographics: age, sex, race, family history, symptom duration, education level.
(2) Medicine: levodopa equivalent daily dose (LEDD).
(3) Blood biomarkers: biochemical tests (p tau181p, total tau, abeta 42, alpha syn).
(4) Motor manifestations: International Parkinson's disease and Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS)-Part II, MDS-UPDRS-Part III, total tremor score, postural instability–gait difficulty (PIGD) score[19], Schwab-England Activities of daily living score.
(5) Cognition manifestations: Montreal Cognitive Assessment (MoCA)[20], Benton Judgment of Line Orientation (Bento)[21], Symbol Digit Modalities Test (symbol digit)[22], Hopkins Verbal Learning Test (HVLT)[23] for total recall, delayed recall, retention and recognition-discrimination, and semantic verbal-language fluency test (semantic fluency)[24] and letter-number sequencing (LNS)[25].
(6) Mood manifestations: Geriatric Depression Scale (GDS) score[26], State-Trait Anxiety Inventory (STAI) score[27].
(7) Sleep manifestations: REM Sleep Behavior Questionnaire (RBD)[28], Epworth Sleepiness Score (ESS)[29].
(8) Autonomic manifestations: Scales for Outcomes in Parkinson's Disease-Autonomic questionnaire (SCOPA)[30].
(9) Global manifestations: we calculated a global composite outcome (GCO) to indicate global prognosis based on a previous study[31]. The GCO equally weighted non-motor symptoms (UPDRS-I), motor symptoms (UPDRS-II), motor signs (UPDRS-III), overall activities of daily living (Schwab and England ADL), and global cognition (MoCA), standardized by averaging the z-scores of each component.

Data preprocess steps involved converting raw scores to percentages, subsequently standardizing them relative to the baseline population mean and standard deviation to ensure the different measures were comparable. Firstly, we converted raw scores into a percentage of the maximum score for that scale, with higher scores indicating worse symptoms. The MoCA scores underwent inverse scoring to represent the count of incorrect responses out of a total possible score of 30. Similarly, semantic fluency was inversely scored against the highest individual score recorded at the baseline. Subsequently, each measure was normalized to the population baseline mean and standard to adjust for any differences in the scales.

## 4.2 Linear mixed effect model

We adopted the linear mixed-effects model framework to identify SNPs that are significantly associated with PD progression speed for each trait. For patient $i$ at time $t$, the clinical trait score $y_{it}$ for one trait is modeled as follows:
$y_{it} = \beta_0 + \gamma_{0i} + \beta_1 t + \gamma_{1i} t + \beta_{02} m_i + \beta_{12} m_i t + \beta_{03} med_t + \beta_{13} med_t t + \beta_a age + \beta_r race + \beta_d duration + \beta_s sex + \beta_f family + \beta_b basline + \epsilon_{it}$

Here, $\gamma_{0i}$ and $\gamma_{1i}$ are random effects of the patient $i$, and $m_i$ denotes the genotype of the SNP $m$ in patient $i$. Covariates such as age (in years), race, symptom duration (in years), and the levodopa equivalent daily dose (med) were included in the model, with additional consideration of education years for cognitive traits. All covariate values, except for baseline trait measurements, were standardized to a z-score to ensure comparability. The risk locus whose $\beta_{12}$ (i.e. the interaction effect of the genotype with time) is smaller than 0.05 would be considered as progression speed-associated genetic factors.

## 4.3 Network medicine approach

After the derivation of individual risk genes, their associated disease modules can be identified within the protein-protein interaction (PPI) network[45] as densely connected subnetworks using the Genome-wide Positioning Systems network (GPSnet) algorithm. For each risk gene, we calculate a score $s(i) = -log(p) \times \beta$ with p as its corresponding p-value and $\beta$ is the coefficient derived from the linear mixed-effects model. Then we can propagate the scores in the whole PPI network to get a smoothed $s(i)$ for each gene $i$.

In constructing a gene module, we begin with a module $M$ containing only a seed gene identified from the linear mixed-effects model. The score of $M$, denoted as $\Gamma_M$, is defined as $\Gamma_M = \sum_{i \in M}(s(i) - \mu)/\sqrt{m}$, where $\mu$ is the averaged score of all genes and $m$ is the number of genes in $M$. The module $M$ is then expanded incorporating genes that are not only highly connected to genes within M but also serve to increase $\Gamma_M$. This expansion process is halted when no new gene inclusion can further augment $\Gamma_M$.

By iterating these procedures multiple times, we can get a collection of modules. From this collection, we select modules with the top 0.5% highest $\Gamma_M$ scores and define the largest connected component after discarding isolated genes.

## 4.4 Pathway enrichment analysis

In our study, we incorporated three well-established databases: Gene Ontology (GO)[38], Kyoto Encyclopedia of Genes and Genomes (KEGG)[39], and Reactome[40], for pathway analysis. Each database was independently analyzed using the g:Profiler tool[37]. To determine the statistical significance of our findings, we conducted hypergeometric tests. Subsequently, to account for the issue of multiple testing, all P values were adjusted using Benjamini-Hochberg correction[46]. We identified pathways as significantly enriched if they have an adjusted P value below 0.05 and a size encompassing between 5 and 1000 genes. Additionally, to eliminate redundancy in the GO terms, we utilized the GOxploreR[41] and REVIGO[42] tools to filter out overlapping or redundant biological terms.

# 5 Tables

**Table 1**: Cohort Statistics. Demographic and clinical information with mean, standard deviation, and percentage.

| | | | |
|---|---|---|---|
| # patient | | | 500 |
| #visit, mean(SD) | | | 14.23(3.73) |
| visit time gap, year, mean(SD) | | | 8.72(2.54) |
| Demographic | age onset, mean(SD) | | 59.7(9.76) |
| | Sex male, N(%) | | 291 (58.2) |
| | Race white, N(%) | | 482(96.4) |
| | symptom duration, month, mean(SD) | | 38.77(43.22) |
| | Family history, N(%) | | 145(29) |
| | Education history group, N(%) | <12 | 48 (9.6) |
| | | 12-16 | 260 (52) |
| | | >16 | 192 (38.4) |
| Medicine | LEDD | | 523.94 |
| Motor | motor | updrs2, mean(SD) | 9.11(6.54) |
| | | updrs3, mean(SD) | 24.84(11.85) |
| | | Schwad England score, mean(SD) | 87.07(11.68) |
| | combined motor | Tremor score, mean (SD) | 0.51(0.41) |
| | | PIGD score, mean(SD) | 0.39(0.44) |
| Non-motor | Cognition | moca, mean(SD) | 26.51(3.44) |
| | | Bento, mean(SD) | 25(4.80) |
| | | HVLT, mean(SD) | 47.07(11.46) |
| | | LNS, mean(SD) | 11.26(2.97) |
| | | semantic fluency, mean(SD) | 51.5(10.99) |
| | | symbol digit test, mean(SD) | 44.83(10.70) |
| | Mood | gds, mean(SD) | 2.9(2.98) |
| | | stai, mean(SD) | 66.42(19.04) |
| | Autonomic | SCOPA, mean(SD) | 19.84(13.13) |
| | Sleep | REM, mean(SD) | 5.01(3.35) |
| | | ESS, mean(SD) | 7.07(4.43) |
| biomarker | | p_tau181p(SD) | 14.07(5.21) |
| | | total_tau(SD) | 166.49(59.25) |
| | | abeta 42(SD) | 914.45(370.77) |
| | | alpha syn(SD) | 1472.17(608.31) |

**Table 2**: Result summary. Counts of identified SNPs, gene module sizes, and enriched pathways.

| | Manifestation | Trait | #SNP (p) | #gene | gene module size | #pathway |
|---|---|---|---|---|---|---|
| motor | motor | updrs2 | 141 | 188 | 274 | 947 |
| | | updrs3 | 179 | 235 | 170 | 702 |
| | | schwab England | 125 | 167 | 241 | 904 |
| | combined motor | pigd | 172 | 230 | 132 | 552 |
| | | tremor score | 144 | 208 | 233 | 884 |
| non-motor | Cognition | moca | 154 | 220 | 232 | 906 |
| | | Bento | 173 | 260 | 344 | 1053 |
| | | HVLT | 448 | 599 | 530 | 1099 |
| | | LNS | 137 | 199 | 116 | 696 |
| | | semantic fluency | 137 | 202 | 195 | 652 |
| | | symbol digit test | 134 | 198 | 94 | 432 |
| | Mood | geriatric depression scale | 98 | 169 | 193 | 777 |
| | | state trait anxiety index | 104 | 143 | 118 | 242 |
| | Autonomic | SCOPA | 283 | 370 | 230 | 870 |
| | Sleep | REM | 108 | 154 | 119 | 800 |
| | | ESS | 108 | 165 | 125 | 622 |
| Global | | global composite outcome | 118 | 164 | 105 | 490 |
| Biomarker | | p_tau181p | 166 | 209 | 34 | 190 |
| | | total_tau | 126 | 176 | 34 | 172 |
| | | abeta_42 | 125 | 183 | 112 | 472 |
| | | alpha_syn | 111 | 155 | 94 | 581 |

# 6 Figures

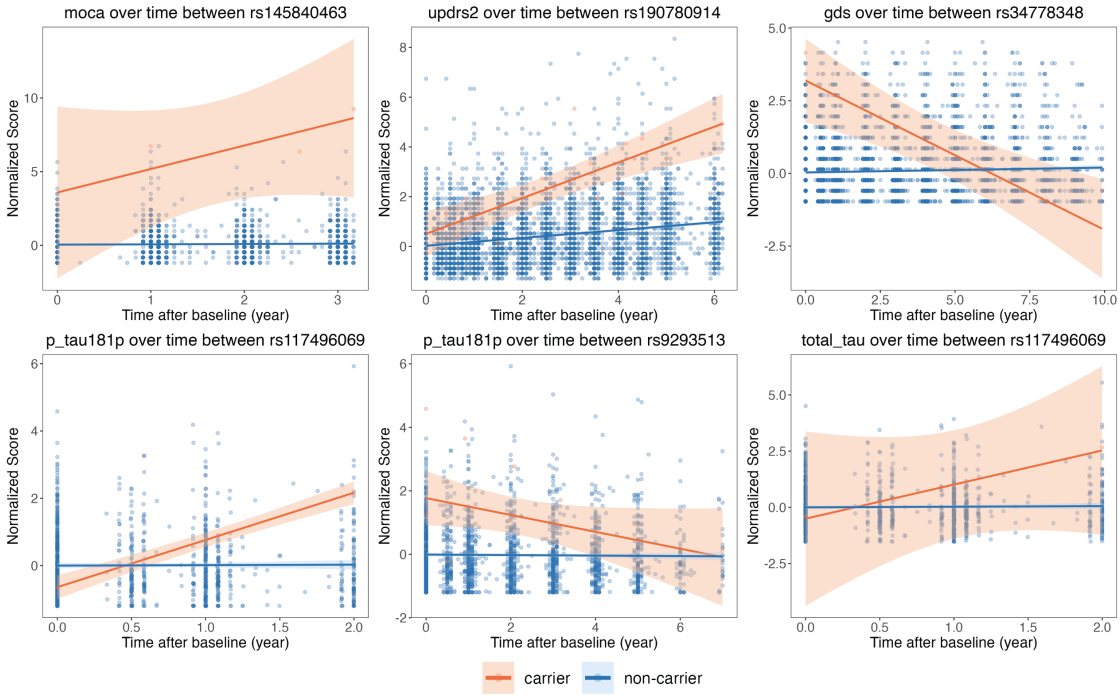


**Fig. 1**: **Clinical score changes over time between the SNP-carrier group and the SNP-non-carrier group**. The x-axis represents the visit time in years since baseline, while the y-axis shows the normalized clinical scores. Blue points represent the SNP-non-carrier group and orange points represent the SNP-carrier group. For each group, a linear model fitted line and its 95% confidence interval are depicted.

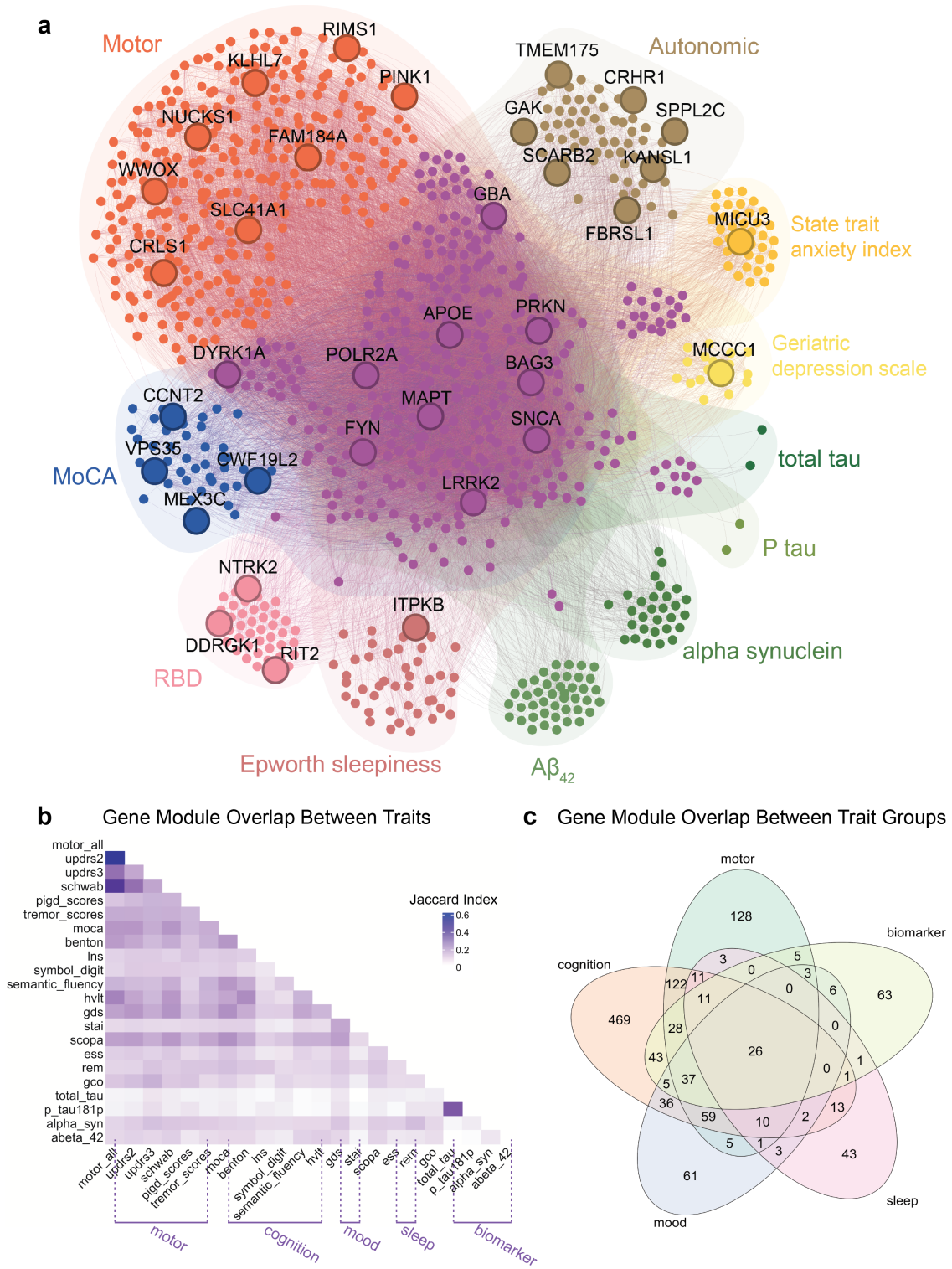


**Fig. 2**: **Visualization and Overlap of Gene Modules Across Traits. a.** Network visualization of gene modules associated with each trait. Each trait is denoted by a unique color, with overlaps indicated in purple. Genes identified from previous studies related to Parkinson's disease are highlighted. **b.** Heatmap of gene module overlap between traits. The x-axis and y-axis label the different traits, and each cell represents the Jaccard index, quantifying the overlap between the gene modules of corresponding traits. **c.** Venn diagram of gene module overlap among trait categories. Traits within each category are consolidated, with gene modules for each category differentiated by distinct colors.

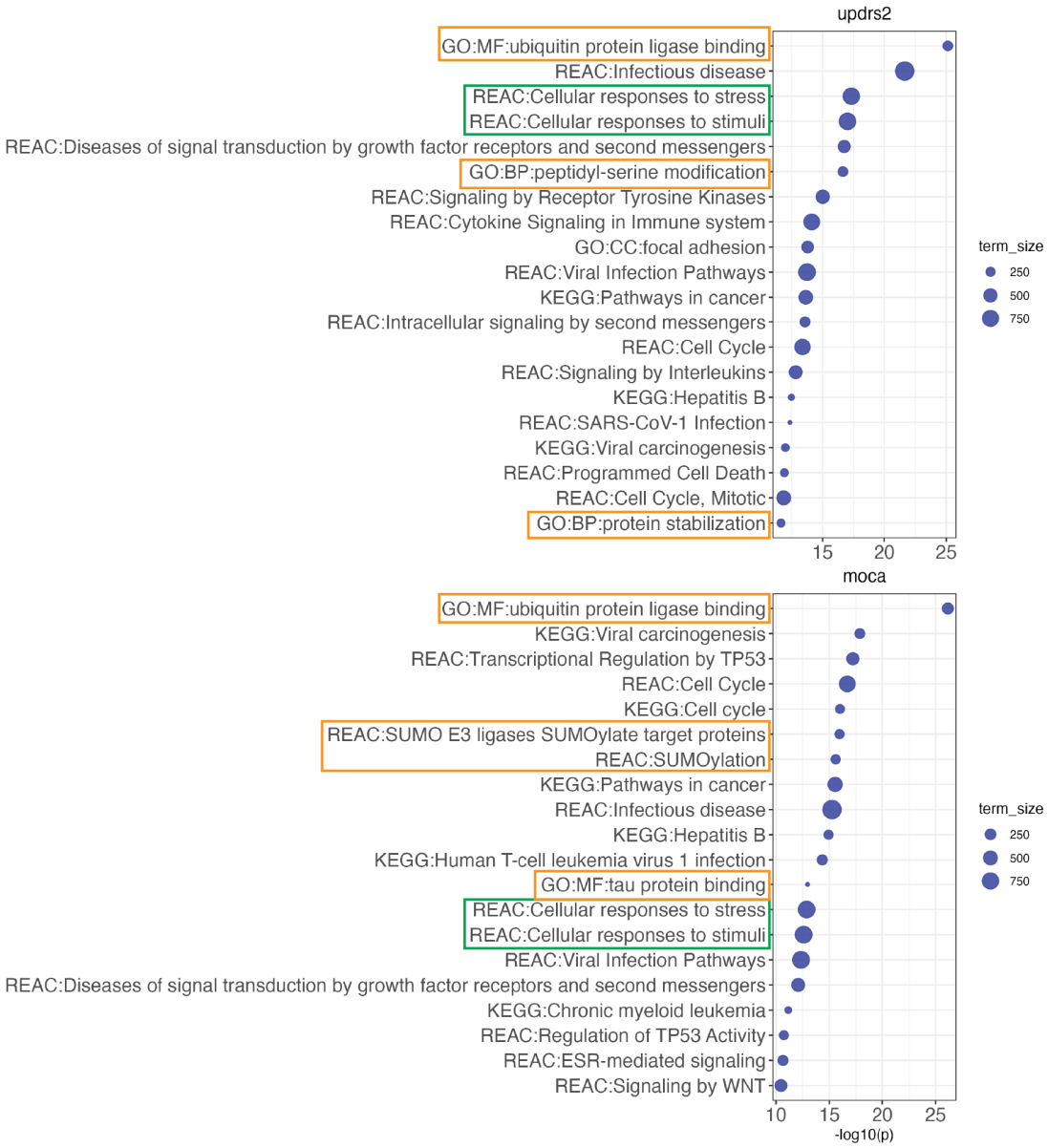


**Fig. 3**: **Bubble plot of enriched pathways in traits**. Each plot displays the enriched pathways, with the x-axis indicating the $-log(p_{adjust})$ value and the y-axis listing the top 20 enriched pathways. The bubble size corresponds to the size of the pathway.

# References


[1] Marras, C., Beck, J., Bower, J., Roberts, E., Ritz, B., Ross, G., Abbott, R., Savica, R., Van Den Eeden, S., Willis, A.: Prevalence of parkinson's disease across north america. NPJ Parkinson's disease **4**(1), 21 (2018)

[2] Xu, J., Gong, D., Man, C., Fan, Y.: Parkinson's disease and risk of mortality: meta-analysis and systematic review. Acta Neurologica Scandinavica **129**(2), 71–79 (2014)

[3] Macleod, A.D., Taylor, K.S., Counsell, C.E.: Mortality in parkinson's disease: a systematic review and meta-analysis. Movement Disorders **29**(13), 1615–1622 (2014)

[4] Lang, A.E., Espay, A.J.: Disease modification in parkinson's disease: current approaches, challenges, and future considerations. Movement Disorders **33**(5), 660–677 (2018)

[5] Politis, M., Wu, K., Molloy, S., G. Bain, P., Chaudhuri, K.R., Piccini, P.: Parkinson's disease symptoms: the patient's perspective. Movement Disorders **25**(11), 1646–1651 (2010)

[6] Farrow, S.L., Cooper, A.A., O'Sullivan, J.M.: Redefining the hypotheses driving parkinson's diseases research. npj Parkinson's Disease **8**(1), 45 (2022)

[7] Nalls, M.A., Blauwendraat, C., Vallerga, C.L., Heilbron, K., Bandres-Ciga, S., Chang, D., Tan, M., Kia, D.A., Noyce, A.J., Xue, A., *et al.*: 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), 1091–1102 (2019)

[8] Chang, D., Nalls, M.A., Hallgrímsdóttir, I.B., Hunkapiller, J., Van Der Brug, M., Cai, F., Consortium, I.P.D.G., Research Team, Kerchner, G.A., Ayalon, G., *et al.*: A meta-analysis of genome-wide association studies identifies 17 new parkinson's disease risk loci. Nature genetics **49**(10), 1511–1516 (2017)

[9] 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., *et al.*: 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), 746–754 (2020)

[10] Tan, M.M., Lawton, M.A., Jabbari, E., Reynolds, R.H., Iwaki, H., Blauwendraat, C., Kanavou, S., Pollard, M.I., Hubbard, L., Malek, N., *et al.*: Genome-wide association studies of cognitive and motor progression in parkinson's disease. Movement Disorders **36**(2), 424–433 (2021)

[11] Latourelle, J.C., Beste, M.T., Hadzi, T.C., Miller, R.E., Oppenheim, J.N., Valko, M.P., Wuest, D.M., Church, B.W., Khalil, I.G., Hayete, B., *et al.*: Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed parkinson's disease: a longitudinal cohort study and validation. The Lancet Neurology **16**(11), 908–916 (2017)

[12] Saunders-Pullman, R., Mirelman, A., Alcalay, R.N., Wang, C., Ortega, R.A., Raymond, D., Mejia-Santana, H., Orbe-Reilly, M., Johannes, B.A., Thaler, A., *et al.*: Progression in the lrrk2-associated parkinson disease population. JAMA neurology **75**(3), 312–319 (2018)

[13] Ortega, R.A., Wang, C., Raymond, D., Bryant, N., Scherzer, C.R., Thaler, A., Alcalay, R.N., West, A.B., Mirelman, A., Kuras, Y., *et al.*: Association of dual lrrk2 g2019s and gba variations with parkinson disease progression. JAMA network open **4**(4), 215845–215845 (2021)

[14] Blauwendraat, C., Nalls, M.A., Singleton, A.B.: The genetic architecture of parkinson's disease. The Lancet Neurology **19**(2), 170–178 (2020)

[15] Barabási, A.-L., Gulbahce, N., Loscalzo, J.: Network medicine: a network-based approach to human disease. Nature reviews genetics **12**(1), 56–68 (2011)

[16] Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W.: From molecular to modular cell biology. Nature **402**(Suppl 6761), 47–52 (1999)

[17] Cheng, F., Lu, W., Liu, C., Fang, J., Hou, Y., Handy, D.E., Wang, R., Zhao, Y., Yang, Y., Huang, J.: A genome-wide positioning systems network algorithm for in silico drug repurposing. Nature communications **10**(1), 3476 (2019)

[18] Lee, L.Y.-H., Loscalzo, J.: Network medicine in pathobiology. The American Journal of Pathology **189**(7), 1311–1326 (2019)

[19] Stebbins, G.T., Goetz, C.G., Burn, D.J., Jankovic, J., Khoo, T.K., Tilley, B.C.: How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified parkinson's disease rating scale: comparison with the unified parkinson's disease rating scale. Movement Disorders **28**(5), 668–670 (2013)

[20] Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H.: The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society **53**(4), 695–699 (2005)

[21] Benton, A.L., Varney, N.R., Hamsher, K.d.: Visuospatial judgment: A clinical test. Archives of neurology **35**(6), 364–367 (1978)

[22] Smith, A.: Symbol Digit Modalities Test. Western psychological services Los Angeles, ??? (1973)

[23] Shapiro, A.M., Benedict, R.H., Schretlen, D., Brandt, J.: Construct and concurrent validity of the hopkins verbal learning test–revised. The Clinical Neuropsychologist **13**(3), 348–358 (1999)

[24] Gladsjo, J.A., Schuman, C.C., Evans, J.D., Peavy, G.M., Miller, S.W., Heaton, R.K.: Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. Assessment **6**(2), 147–178 (1999)

[25] Wechsler, D.: Wechsler adult intelligence scale. Frontiers in Psychology (1981)

[26] Yesavage, J.A., Sheikh, J.I.: 9/geriatric depression scale (gds) recent evidence and development of a shorter version. Clinical gerontologist **5**(1-2), 165–173 (1986)

[27] Spielberger, C.D.: Manual for the state-trait anxiety inventory (self-evaluation questionnaire). (No Title) (1970)

[28] Stiasny-Kolster, K., Mayer, G., Schäfer, S., Möller, J.C., Heinzel-Gutenbrunner, M., Oertel, W.H.: The rem sleep behavior disorder screening questionnaire—a new diagnostic instrument. Movement disorders **22**(16), 2386–2393 (2007)

[29] Johns, M.W.: A new method for measuring daytime sleepiness: the epworth sleepiness scale. sleep **14**(6), 540–545 (1991)

[30] Visser, M., Marinus, J., Stiggelbout, A.M., Van Hilten, J.J.: Assessment of autonomic dysfunction in parkinson's disease: the scopa-aut. Movement disorders: official journal of the Movement Disorder Society **19**(11), 1306–1312 (2004)

[31] Fereshtehnejad, S.-M., Zeighami, Y., Dagher, A., Postuma, R.B.: Clinical criteria for subtyping parkinson's disease: biomarkers and longitudinal progression. Brain **140**(7), 1959–1976 (2017)

[32] Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J.: The nhgri-ebi gwas catalog: knowledgebase and deposition resource. Nucleic acids research **51**(D1), 977–985 (2023)

[33] Matenia, D., Mandelkow, E.M.: Emerging modes of pink1 signaling: another task for mark2. Frontiers in molecular neuroscience **7**, 37 (2014)

[34] Jaccard, P.: The distribution of the flora in the alpine zone. 1. New phytologist **11**(2), 37–50 (1912)

[35] Grünert, S.C., Stucki, M., Morscher, R.J., Suormala, T., Bürer, C., Burda, P., Christensen, E., Ficicioglu, C., Herwig, J., Kölker, S.: 3-methylcrotonyl-coa carboxylase deficiency: clinical, biochemical, enzymatic and molecular studies in 88 individuals. Orphanet journal of rare diseases **7**(1), 1–24 (2012)

[36] Wang, Q., Wu, Z., Xiang, H., Zhou, Y., Qin, X., Tian, J.: Revealing the role of leucine in improving the social avoidance behavior of depression through a combination of untargeted and targeted metabolomics. Food & Function **14**(14), 6397–6409 (2023)

[37] Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., Peterson, H.: g: Profiler—interoperable web

service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Research, 347 (2023)

[38] Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S., Hub, A., Group, W.P.W.: Amigo: online access to ontology and annotation data. Bioinformatics **25**(2), 288–289 (2009)

[39] Kanehisa, M., Goto, S.: Kegg: kyoto encyclopedia of genes and genomes. Nucleic acids research **28**(1), 27–30 (2000)

[40] Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B.: The reactome pathway knowledgebase. Nucleic acids research **46**(D1), 649–655 (2018)

[41] Manjang, K., Tripathi, S., Yli-Harja, O., Dehmer, M., Emmert-Streib, F.: Graph-based exploitation of gene ontology using goxplorer for scrutinizing biological significance. Scientific reports **10**(1), 16672 (2020)

[42] Supek, F., Bošnjak, M., Škunca, N., Šmuc, T.: Revigo summarizes and visualizes long lists of gene ontology terms. PloS one **6**(7), 21800 (2011)

[43] Dias, V., Junn, E., Mouradian, M.M.: The role of oxidative stress in parkinson's disease. Journal of Parkinson's disease **3**(4), 461–491 (2013)

[44] Zhang, Y., Gao, J., Chung, K.K., Huang, H., Dawson, V.L., Dawson, T.M.: Parkin functions as an e2-dependent ubiquitin–protein ligase and promotes the degradation of the synaptic vesicle-associated protein, cdcrel-1. Proceedings of the National Academy of Sciences **97**(24), 13354–13359 (2000)

[45] Kovács, I.A., Luck, K., Spirohn, K., Wang, Y., Pollis, C., Schlabach, S., Bian, W., Kim, D.-K., Kishore, N., Hao, T.: Network-based prediction of protein interactions. Nature communications **10**(1), 1240 (2019)

[46] Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) **57**(1), 289–300 (1995)