Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Common variants in the lysosomal pathway contribute to Parkinson's Disease risk

P. Dimartino
Department of Molecular Medicine, University of Pavia

A. Tenghe
Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen

S. Khodaee
Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen

A. A.K. Sreelatha
Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen

C. Domenighetti
Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805

P. E. Sugier
Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805

C. Schulte
Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen

B. Portugal
Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805

P. May
Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg

Z. Landoulsi
Transversal Translational Medicine, Luxembourg Institute of Health (LIH)

D. R. Bobbili
Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg

M. Radivojkov-Blagojevic
Institute of Human Genetics, Helmholtz Zentrum München

J. Krüger
Group of Applied Bioinformatics, University of Tübingen

P. Lichtner
Institute of Human Genetics, Helmholtz Zentrum München

D. G. Hernandez
Molecular Genetics Section, Laboratory of Neurogenetics, NIA, NIH, MD 20892
C. Edsall
Molecular Genetics Section, Laboratory of Neurogenetics, NIA, NIH, MD 20892
G. D. Mellick
Griffith Institute for Drug Discovery, Griffith University, Don Young Road, Nathan
A. Zimprich
Department of Neurology, Medical University of Vienna
W. Pirker
Department of Neurology, Klinik Ottakring
E. Rogaeva
Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto
A. E. Lang
Edmond J. Safra Program in Parkinson's Disease, Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, UHN, Toronto
S. Koks
Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University
P. Taba
Department of Neurology and Neurosurgery, University of Tartu
S. Lesage
Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, INSERM, CNRS, Assistance Publique Hôpitaux de Paris, Department of Neurologie
A. Brice
Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, INSERM, CNRS, Assistance Publique Hôpitaux de Paris, Department of Neurologie
J. C. Corvol
Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, INSERM, CNRS, Assistance Publique Hôpitaux de Paris, Department of Neurologie
M. C. Chartier-Harlin
Univ. Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Lille Neurosciences & Cognition, F-59000
E. Mutez
Univ. Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Lille Neurosciences & Cognition, F-59000
L. G. González Ricardo
Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen
K. Brockmann
Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen

A. B. Deutschlander
Department of Neurology, Ludwig Maximilians University of Munich
L. F. Burbulla
German Center for Neurodegenerative Diseases (DZNE) Munich
G. M. Hadjigeorgiou
Department of Neurology, Medical School, University of Cyprus
E. Dardiotis
Department of Neurology, Laboratory of Neurogenetics, University of Thessaly, University Hospital of Larissa
L. Stefanis
Center of Clinical Research, Experimental Surgery and Translational Research, Biomedical Research Foundation of the Academy of Athens
A. M. Simitsi
1st Department of Neurology, Eginition Hospital, Medical School, National and Kapodistrian University of Athens
S. Petrucci
Department of Clinical and Molecular Medicine, Sapienza University of Rome
L. Straniero
Department of Biomedical Sciences, Humanitas University
R. Asselta
Department of Biomedical Sciences, Humanitas University
L. Magistrelli
Parkinson Institute, Azienda Socio Sanitaria Territoriale (ASST) Gaetano Pini/CTO
I. U. Isaias
Parkinson Institute, Azienda Socio Sanitaria Territoriale (ASST) Gaetano Pini/CTO
A. Zecchinelli
Parkinson Institute, Azienda Socio Sanitaria Territoriale (ASST) Gaetano Pini/CTO
G. Pezzoli
Fondazione Pezzoli per la Malattia di Parkinson
L. Brighina
Department of Neurology, Fondazione IRCCS San Gerardo dei Tintori
C. Ferrarese
Department of Neurology, Fondazione IRCCS San Gerardo dei Tintori
G. Annesi
Institute for Biomedical Research and Innovation, National Research Council
A. Quattrone
Institute of Neurology, Department of Medical and Surgical Sciences, Magna Graecia University
M. Gagliardi

Department of Medical and Surgical Sciences, Neuroscience Research Center, Magna Graecia University

H. Matsuo
Department of Integrative Physiology and Bio-Nano Medicine, National Defense Medical College

A. Nakayama
Department of Integrative Physiology and Bio-Nano Medicine, National Defense Medical College

N. Hattori
Department of Neurology, Juntendo University School of Medicine, Bunkyo-ku

K. Nishioka
Department of Neurology, Juntendo University School of Medicine, Bunkyo-ku

S. J. Chung
Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine

Y. J. Kim
Department of Neurology, Yonsei University College of Medicine

L. Pavelka
Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg

I. Y. Bergua
Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg

B. P.C. Warrenburg
Radboud University Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Department of Neurology

B. R. Bloem
Radboud University Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Department of Neurology

M. Toft
Department of Neurology, Oslo University Hospital

L. Pihlstrøm
Department of Neurology, Oslo University Hospital

L. Correia Guedes
Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa

J. J. Ferreira
Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa

S. Bardien
Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University

J. Carr

Division of Neurology, Department of Medicine, Faculty of Medicine and Health Sciences, Stellenbosch University

A. B. Singleton
Molecular Genetics Section, Laboratory of Neurogenetics, NIA, NIH, MD 20892

E. Tolosa
Parkinson's disease & Movement Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona

M. Ezquerra
Parkinson's disease & Movement Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona

P. Pastor
Fundació per la Recerca Biomèdica i Social Mútua Terrassa, Terrassa

K. Wirdefeldt
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet

N. L. Pedersen
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet

C. Ran
Centre for Cluster Headache, Department of Neuroscience, Karolinska Institutet

A. C. Belin
Centre for Cluster Headache, Department of Neuroscience, Karolinska Institutet

A. Puschmann
Lund University, Skåne University Hospital, Department of Clinical Sciences Lund, Neurology, Getingevägen 4, 221 85

C. E. Clarke
University of Birmingham and Sandwell and West Birmingham Hospitals NHS Trust

K. E. Morrison
Faculty of Medicine, Health and Life Sciences, Queens University

M. Tan
Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London

D. Krainc
Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago

M. J. Farrer
Department of Neurology, McKnight Brain Institute, University of Florida

A. Elbaz
Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805

R. Kruger
Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg

T. Gasser
Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen
E. M. Valente
Department of Molecular Medicine, University of Pavia
M. Sharma
manu.sharma@uni-tuebingen.de

Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen

---

Article

Keywords: Lysosomal Pathway, Polygenic Risk Score, Parkinson's Disease

Posted Date: August 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10552418/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Common variants in the lysosomal pathway contribute to Parkinson's Disease risk

P. Dimartino[1,2], A. Tenghe[3], S. Khodaee[3], A.A.K. Sreelatha[3], C. Domenighetti[4], P.E. Sugier[4], C. Schulte[5,6], B. Portugal[4], P. May[7], Z. Landoulsi[8], D.R. Bobbili[9], M. Radivojkov-Blagojevic[10], J. Krüger[11], P. Lichtner[10], D.G. Hernandez[12], C. Edsall[12], G.D. Mellick[13], A. Zimprich[14], W. Pirker[15], E. Rogaeva[16], A.E. Lang[17,18,19], S. Koks[20,21], P. Taba[22,23], S. Lesage[24], A. Brice[24], J.C. Corvol[24], M.C. Chartier-Harlin[25], E. Mutez[25], L.G. González Ricardo[3], K. Brockmann[5,6], A.B. Deutschlander[26,27,28], L.F. Burbulla[29,30,31,32], G.M. Hadjigeorgiou[33,34], E. Dardiotis[34], L. Stefanis[35,36], A.M. Simitsi[36], S. Petrucci[37,38], L. Straniero[39,40], R. Asselta[39,40], L. Magistrelli[41], I.U. Isaias[41], A. Zecchinelli[41], G. Pezzoli[42], L. Brighina[43,44], C. Ferrarese[43,44], G. Annesi[45], A. Quattrone[46], M. Gagliardi[47], H. Matsuo[48], A. Nakayama[48], N. Hattori[49], K. Nishioka[49], S.J. Chung[50], Y.J. Kim[51], L. Pavelka[7,8,9,52], I.Y. Bergua[7,9], B.P.C. Van de Warrenburg[53], B.R. Bloem[53], M. Toft[54], L. Pihlstrøm[54], L. Correia Guedes[55,56], J.J. Ferreira[55,57], S. Bardien[58,59], J. Carr[60], A.B. Singleton[12,61], E. Tolosa[62,63], M. Ezquerra[62], P. Pastor[64,65], K. Wirdefeldt[66], N.L. Pedersen[66], C. Ran[67], A.C. Belin[67], A. Puschmann[68], C.E. Clarke[69], K.E. Morrison[70], M. Tan[71], D. Krainc[30], M.J. Farrer[72], A. Elbaz[4], R. Kruger[7,8,9,52], T. Gasser[5,6], E.M. Valente[1,2], M. Sharma[3] on behalf of the COURAGE-PD Consortium

1. Department of Molecular Medicine, University of Pavia, Pavia, Italy
2. Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Mondino Foundation, Pavia, Italy
3. Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen, Germany
4. Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805, Villejuif, France
5. Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tubingen, Germany
6. German Center for Neurodegenerative Diseases (DZNE), Tubingen, Germany
7. Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-Belval, Luxembourg
8. Transversal Translational Medicine, Luxembourg Institute of Health (LIH), Strassen, Luxembourg
9. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-Belval, Luxembourg
10. Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany
11. Group of Applied Bioinformatics, University of Tübingen, Tübingen, Germany
12. Molecular Genetics Section, Laboratory of Neurogenetics, NIA, NIH, Bethesda, MD 20892, USA
13. Griffith Institute for Drug Discovery, Griffith University, Don Young Road, Nathan, Queensland, Australia

14. Department of Neurology, Medical University of Vienna, Austria
15. Department of Neurology, Klinik Ottakring, Vienna Austria
16. Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Ontario, Canada
17. Edmond J. Safra Program in Parkinson's Disease, Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, UHN, Toronto, Ontario, Canada
18. Division of Neurology, University of Toronto, Toronto, Ontario, Canada
19. Krembil Brain Institute, Toronto, Ontario, Canada
20. Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Murdoch, Australia
21. Perron Institute for Neurological and Translational Science, Nedlands, Western Australia, Australia
22. Department of Neurology and Neurosurgery, University of Tartu, Estonia
23. Neurology Clinic, Tartu University Hospital, Tartu, Estonia
24. Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, INSERM, CNRS, Assistance Publique Hôpitaux de Paris, Department of Neurologie, Paris, France
25. University Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Lille Neurosciences & Cognition, F-59000 Lille, France
26. Department of Neurology, Ludwig Maximilians University of Munich, Germany
27. Department of Neurology, Max Planck Institute of Psychiatry, Munich, Germany
28. Department of Neurology and Department of Clinical Genomics, Mayo Clinic Florida, Jacksonville, FL, USA
29. German Center for Neurodegenerative Diseases (DZNE), Munich, Germany
30. Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611, United States
31. Metabolic Biochemistry, Biomedical Center (BMC), Faculty of Medicine, LMU Munich, Munich, Germany
32. Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
33. Department of Neurology, Medical School, University of Cyprus, Nicosia, Cyprus
34. Department of Neurology, Laboratory of Neurogenetics, University of Thessaly, University Hospital of Larissa, Larissa, Greece
35. Center of Clinical Research, Experimental Surgery and Translational Research, Biomedical Research Foundation of the Academy of Athens, Athens, Greece
36. 1st Department of Neurology, Eginition Hospital, Medical School, National and Kapodistrian University of Athens, Athens, Greece
37. Department of Clinical and Molecular Medicine, Sapienza University of Rome, Italy
38. UOC Medical Genetics and Advanced Cell Diagnostics, S. Andrea University Hospital, Rome, Italy
39. Department of Biomedical Sciences, Humanitas University, Pieve Emanuele (Milan), Italy
40. IRCCS Humanitas Research Hospital, Rozzano (Milan), Italy
41. Parkinson Institute, Azienda Socio Sanitaria Territoriale (ASST) Gaetano Pini/CTO, Milano, Italy
42. Fondazione Pezzoli per la Malattia di Parkinson, Milan, Italy
43. Department of Neurology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy
44. Department of Medicine and Surgery and Milan Center for Neuroscience, University of Milano Bicocca, Milano, Italy
45. Institute for Biomedical Research and Innovation, National Research Council, Cosenza, Italy
46. Institute of Neurology, Department of Medical and Surgical Sciences, Magna Graecia University, Catanzaro, Italy
47. Department of Medical and Surgical Sciences, Neuroscience Research Center, Magna Graecia University, Catanzaro, Italy

48. Department of Integrative Physiology and Bio-Nano Medicine, National Defense Medical College, Saitama 359-8513, Japan
49. Department of Neurology, Juntendo University School of Medicine, Bunkyo-ku, Tokyo 113-8421, Japan
50. Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea
51. Department of Neurology, Yonsei University College of Medicine, Seoul, South Korea
52. Parkinson's Research Clinic, Centre Hospitalier de Luxembourg, Luxembourg
53. Radboud University Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Department of Neurology, Nijmegen, The Netherlands
54. Department of Neurology, Oslo University Hospital, Oslo, Norway
55. Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
56. Department of Neurosciences and Mental Health, Neurology, Hospital de Santa Maria, Centro Hospitalar Universitario Lisboa Norte (CHULN), Lisbon, Portugal
57. Laboratory of Clinical Pharmacology and Therapeutics, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
58. Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa
59. South African Medical Research Council/Stellenbosch University Genomics of Brain Disorders Research Unit, Stellenbosch University, Cape Town, South Africa
60. Division of Neurology, Department of Medicine, Faculty of Medicine and Health Sciences, Stellenbosch University, South Africa
61. Center For Alzheimer's and Related Dementias, NIA, NIH, Bethesda, MD 20892, USA
62. Parkinson's disease & Movement Disorders Unit, Neurology Service, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona, Spain
63. Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED: CB06/05/0018-ISCIII) Barcelona, Spain
64. Fundació per la Recerca Biomèdica i Social Mútua Terrassa, Terrassa, Barcelona, Spain
65. Movement Disorders Unit, Department of Neurology, Hospital Universitari Mutua de Terrassa, Terrassa, Barcelona, Spain
66. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
67. Centre for Cluster Headache, Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden
68. Lund University, Skåne University Hospital, Department of Clinical Sciences Lund, Neurology, Getingevägen 4, 221 85, Lund, Sweden
69. University of Birmingham and Sandwell and West Birmingham Hospitals NHS Trust, United Kingdom
70. Faculty of Medicine, Health and Life Sciences, Queens University, Belfast, United Kingdom
71. Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London, London, UK
72. Department of Neurology, McKnight Brain Institute, University of Florida, Gainesville, FL, USA

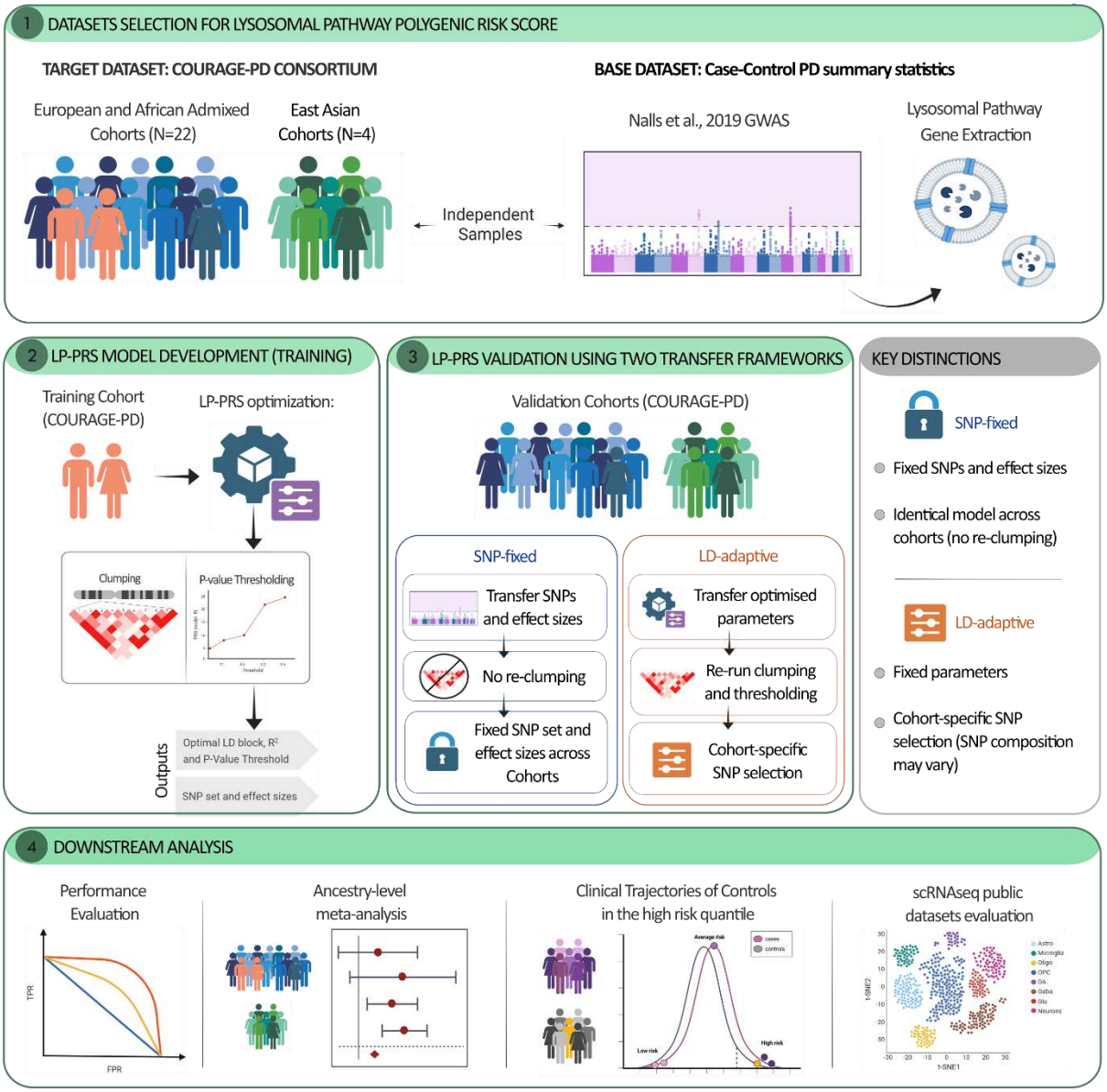
*Corresponding author

manu.sharma@uni-tuebingen.de

Keywords

1. Lysosomal Pathway
2. Polygenic Risk Score
3. Parkinson's Disease

Graphical abstract



Abstract

Objective

Despite the progress in identifying an ever-increasing number of genetic loci implicated in Parkinson's disease (PD), it has been increasingly recognised that pathway-based polygenic risk approaches could provide an alternative strategy to identify a subset of PD cases that might be more prone to pathway-specific susceptibility. We hypothesised that lysosomal-based pathway approaches would identify novel genetic risk loci that could explain part of lysosomal-dependent neurodegeneration in PD.

Methods

A total of 7,668 patients with PD and 7,146 controls from 26 COURAGE-PD sites of European and East Asian ancestries were included. Lysosomal pathway-specific polygenic risk scores (LP-PRS) were generated using two complementary transfer strategies: a linkage-disequilibrium (LD)-adaptive approach using optimised clumping and thresholding parameters within validation cohorts, and a SNP-fixed approach directly transferring the optimised SNP set across cohorts. PRS associations were evaluated independently within each cohort and then combined in a meta-analysis. Single-cell RNA sequencing datasets (58 cases and 61 controls), including substantia nigra and midbrain tissues, were used to assess expression patterns of top LP-PRS-associated genes. Individuals within the top and bottom 10% of the LP-PRS distribution were classified as high-risk and low-risk, respectively. Clinical trajectories were assessed in three large cohorts with available clinical follow-up data.

Results

LP-PRS optimisation identified 49 independent SNPs spanning 35 lysosomal pathway genes, including 14 variants not previously reported in PD case-control GWAS meta-analyses. LP-PRS was significantly associated with increased PD risk across ancestries, with a pooled odds ratio (OR) estimate of 1.42 (95% CI=[1.36,1.47]) in the European cohorts and OR=1.15 (95% CI=[1.03,1.29]) in the East Asian cohorts, with minimal between-cohort heterogeneity. Single-cell RNA sequencing analyses revealed reproducible differential expression of several LP-PRS-associated genes across multiple brain regions and cell types, particularly oligodendrocytes and microglia. Longitudinal analyses over approximately 10 years identified four individuals within the high-risk LP-PRS control group who prospectively developed PD, highlighting the potential use of pathway-informed genetic stratification.

Interpretation

These findings support the contribution of common variants within lysosomal pathway genes to PD susceptibility across ancestries and further support a central role for endolysosomal biology in PD pathogenesis. This study also demonstrates the potential utility of biologically informed pathway-specific PRS approaches for genetic stratification in PD opening novel strategies for precision medicine in the future.

Introduction

Parkinson's disease (PD) is influenced by the complex interplay of environmental and genetic factors, with the latter having provided a novel insight to understand the genetic aetiology of the disease risk, onset and progression and the pathophysiological changes leading to neurodegeneration(1). The availability of cost-effective array-based approaches has revolutionised our understanding of the underlying genetic causes for the vast majority of complex diseases, including neurodegenerative diseases. Complementary to the array-based approach, the application of next-generation sequencing technologies has also expedited gene discovery in families where the mode of inheritance is either dominant or recessive (2–4). Thus far, a total number of 9 PARK genes have been identified (the dominant genes *SNCA*, *LRRK2*, *VPS35*, *CHCHD2*, *RAB32*, the three recessive genes *PRKN*, *PINK1*, *PARK7* as well as the strong risk factor *GBA1* (5)) and the number of risk loci within the span of 15 years has increased from 2 to 90 (1). Despite the observed contribution of genetics in PD pathogenesis, the reported heritable risk of 22% denotes a considerable proportion of missing heritability, which could be partially explained by a polygenic cumulative inheritance of both common and rare variants (6).

Over the last decade, several studies have highlighted the involvement of a dysregulated lysosomal pathway in PD risk (7–10). The role of lysosomal function and vesicle trafficking in PD has long been established, given that monoallelic variants in *GBA1* and *LRRK2* are the most frequent risk factors for PD (11–13). Additionally, numerous lysosomal genes are emerging as PD susceptibility loci from large-scale genome-wide associations studies (GWAS; i.e., *SCARB2*, *SMPD1, CTSB* and *GALC*) (14,15). Complementary to the findings emerged from GWAS, a previously published study provided evidence that a burden of rare variants in genes that are causative of lysosomal storage disorders (LSD) has been reported to increase PD risk (7). Recently, the dysregulation of lysosomal function and the involvement of the Commander complex and its components, COMMD/coiled-coil domain–containing protein 22 (CCDC22) and CCDC93 (CCC), have also been established as a modifiers of GCase activity in PD (9,10). Although genome-wide PRS approaches have improved genetic risk stratification in PD on a larger scale, pathway-specific PRS models may provide additional biological insight by capturing cumulative genetic variation within disease-relevant pathways(16). To date, however, the polygenic contribution of common variants within lysosomal pathway genes to PD susceptibility has not been systematically evaluated on a large scale. In the present study, we performed a comprehensive multiancestry lysosomal pathway polygenic risk score (LP-PRS) analysis to investigate whether common variants within lysosomal pathway genes contribute to PD susceptibility across ancestries.

Materials and Methods

Study cohorts and genotyping

This study used individual-level data generated by the COmprehensive Unbiased Risk factor Assessment for Genetics and Environment in Parkinson's Disease (COURAGE-PD) Consortium, a collaborative project with a significant contribution from the Genetic Epidemiology of Parkinson's Disease Consortium (GEo-PD) comprising 35 cohorts of primarily European ancestry, as well as representation from an East Asian cohort. The technical details, including data characteristics, quality control (QC) procedures, and the corresponding sample and variant counts after each QC step, are described elsewhere(17–19).

Selection of genes involved in the lysosomal pathway

We generated a comprehensive map of genes pivotal to the mapping of the lysosomal pathway by querying the leading databases for gene function annotation. We searched and extracted genes from the KEGG pathway database (search term: "GO:0005764"), the Gene Ontology database through the AmiGO 2 web application (search terms: "Lysosomal" and "Lysosome") and Gene Set Enrichment Analysis (GSEA) (MsigDB human Gene Set: "GOCC_VACUOLE"). Additionally, we included the list of lysosomal genes and modifiers of GCase and lysosomal activity found in two recent studies (7,10) to evaluate their polygenic involvement. Gene coordinates were retrieved from the GENCODE v47 comprehensive gene annotation dataset (GRCh37 reference build). The final list of genes was curated to exclude duplicates, and only autosomal chromosomes were considered for further analyses. Finally, by joining the information from the above-mentioned databases, we generated a comprehensive list of unique lysosomal genes to build our LP-PRS.

Quality control of base and target datasets

GWAS summary statistics from the published European PD meta-analysis (1) were used as the base dataset for all PRS analyses. QC of the base dataset was performed following recommended procedures for PRS analyses(20). Variants with minor allele frequency < 0.01, imputation quality score < 0.8, ambiguous SNPs, and mismatched variants were excluded. Only autosomal variants located within the curated lysosomal gene set were retained for LP-PRS construction.

For the target datasets, each COURAGE-PD cohort underwent independent QC. This included removal of ambiguous SNPs and harmonisation of genomic coordinates, strand orientation, SNP identifiers, and effect alleles between the base and target datasets. To minimise confounding from highly penetrant monogenic forms of PD, known carriers of pathogenic variants in *GBA1* or *LRRK2* were excluded as previously described(17–19). Individuals overlapping between the base GWAS and target cohorts were also excluded to avoid inflation of PRS associations.

To improve ancestry-specific LD estimation during clumping, individuals of European and East Asian ancestries were extracted from the 1000 Genomes Phase 3 reference panel (21) to generate ancestry-specific LD reference panels. These datasets were used exclusively as external LD reference panels for ancestry-specific PRS clumping.

Lysosomal PRS construction and optimisation strategy

To generate lysosomal pathway-specific PRS, we implemented two complementary prediction strategies: a SNP-fixed approach and a Linkage Disequilibrium (LD) adaptive approach. Both models used the same base dataset comprising summary statistics from published PD meta-GWAS and were implemented using the clumping-and-thresholding method in PRSice-2 (22). These complementary strategies were designed to evaluate the trade-off between cohort-specific optimisation and PRS transferability across cohorts.

The largest COURAGE-PD cohort (Duga dataset) was used as the training dataset for model optimisation. PRS optimisation was performed across combinations of LD-clumping window sizes (250 kb and 500 kb), LD r^2 thresholds (0.1, 0.2, and 0.5), and p-value thresholds. Sex, age, the first ten principal components (PCs) and PD prevalence (0.005 in Europe(23)) were included in all models. The best-fit model was defined as the combination of parameters maximising incremental Nagelkerke's R^2 in the training cohort. The optimised model identified in the training dataset was subsequently evaluated independently across the validation cohorts. Validation was performed within ancestry-specific datasets, setting a PD prevalence of 0.005 for the European cohorts and of 0.007 for the East Asian cohorts (23).

SNP-fixed approach

In the SNP-fixed approach, optimisation was first performed in the training dataset to identify the optimal LD-clumping and p-value thresholding parameters. Consequently, the SNPs retained in the best-performing training model were selected. Both the SNP composition and corresponding effect size weights derived from the base GWAS summary statistics were then transferred directly to each validation cohort. This strategy preserved identical SNP composition and effect size weighting across cohorts, enabling direct transferability of the optimised PRS model derived from the training dataset.

LD-adaptive approach

In the LD-adaptive approach, PRS optimisation was performed similarly in the training dataset using the same clumping-and-thresholding procedure. These parameters were subsequently applied independently across the validation cohorts. Under this strategy, LD clumping and SNP selection were repeated separately within each validation cohort using cohort-specific genotype data. Consequently, although the same optimisation parameters were retained, the specific SNPs included in the final PRS models could differ between cohorts because of differences in LD structure, allele frequencies, and variant availability. This distinction enabled assessment of whether cohort-specific LD structure improved PRS performance relative to the fixed SNP approach.

PRS calculation and meta-analysis

For both approaches, PRS were calculated as the weighted sum of risk alleles carried by each individual, using effect size estimates derived from the base GWAS summary statistics. Associations between PRS and case-control status were evaluated within each COURAGE-PD cohort using logistic regression models, in which case-control status was modelled as a function of PRS while adjusting for sex, age, and 10 PCs computed from

target genotype data. Incremental Nagelkerke's R^2 was computed as the difference between the full model (case-control status ~ PRS + covariates) and the null model (case-control status ~ covariates). Predictive performance was evaluated using receiver operating characteristic (ROC) curves and corresponding area under the curve (AUC) estimates.

Regression coefficients and corresponding standard errors from each cohort were subsequently meta-analysed using inverse-variance weighted fixed-effects and restricted maximum likelihood (REML)-based random-effects models implemented in the R package metafor (24). Random-effects models incorporated both within-cohort variance and between-cohort heterogeneity. Between-cohort heterogeneity was assessed using Cochran's Q and I^2 statistics. Effect estimates were exponentiated to obtain odds ratios (OR), and results were visualised using forest plots.

To assess risk stratification, PRS were standardised within each cohort using the cohort-specific mean and standard deviation. Individuals in the top and bottom 10% of the cohort's PRS distribution were identified and labelled as having high-risk and low-risk of developing PD.

Cross-cohort analysis of single-nucleus RNA-seq data

To examine transcriptional changes across diverse neuronal and glial populations, single-nucleus RNA sequencing data were analysed from four postmortem human cohorts deposited in GEO GSE178265 (25), GSE243639 (26), GSE161045 (27), and GSE157783 (28–30), spanning the substantia nigra, putamen, and midbrain of PD cases and controls. The primary dataset (GSE178265) had been sorted by fluorescence-activated nuclei sorting (FANS) into an $NR4A2^+$ fraction enriched for dopaminergic nuclei, the population selectively vulnerable in PD, and a complementary $NR4A2^-$ fraction representing mostly glial cells. All datasets were processed using a unified pipeline in R v4.5.0 with Seurat (31). After standard quality filtering and doublet removal with Scrublet (32), 98,220 $NR4A2^+$ and 145,376 $NR4A2^-$ nuclei were retained from GSE178265, together with 61,106 nuclei from GSE243639, 17,844 from GSE161045, and 37,530 from GSE157783. Counts were normalised with SCTransform, and datasets were integrated using Harmony (33). Nuclei were clustered on the first 30 PCs, and cell types were annotated manually using established neuronal and non-neuronal marker genes. Differential expression was assessed by Wilcoxon rank-sum testing with Bonferroni correction, and results were visualised as heatmaps and dot plots (significance threshold: expressing fraction > 0.1, |log fold-change| > 0.25, adjusted $p < 0.01$) (details: Supplementary Methods).

Data Availability

The source code to run the analysis is available in https://github.com/Centre-for-Genetic-Epidemiology/Lysosomal_PRS, and data is available via de.NBI cloud.

Results

Study Cohorts

The COURAGE-PD dataset include 15,849 PD patients and 11,444 controls recruited from 35 sites of European and East Asian ancestries. Following the GWAS QC steps, imputation and post-imputation QC already described(17–19), we additionally removed samples with missing age at collection, known to have been included in the PD Case-Control GWAS summary statistics used in this paper(1) and known to be carrying a *LRRK2* or *GBA1* monoallelic variant. After this extensive data preprocessing, the final dataset was well-balanced and included 6,980 cases (60% males, average age at onset: 59.23±11.16 years) and 6,318 controls (45% males) derived from 22 European sites, and 688 cases (52% males, average age at onset: 57.4±10.3 years) and 828 controls (50% males) from 4 sites of East Asian ancestry. The final list of COURAGE-PD cases, controls and SNPs is reported in Supplementary Table 1.

Lysosomal gene set and LP-PRS variant selection

To establish whether an additive contribution of common variants in genes involved in the lysosomal pathway could be associated with an increased risk of PD, our integrative annotation strategy identified 1001 genes known to be involved in the formation and regulation of the lysosomal pathway (Supplementary Table 2). Many of the previously published genes were included in our list, thus supporting the validity of our approach. Starting from this list of genes, we extracted from the base dataset a total of 376,360 common variants passing all quality control checks.

LP-PRS model optimisation

LP-PRS optimisation was performed using an out-of-sample strategy. The optimal model, defined as the model yielding the highest incremental Nagelkerke's R^2, was obtained using an LD clumping window size of 500 kb, an r^2 threshold of 0.2, and a p-value threshold of 3×10^{-4} (Supplementary Fig.1). This model included 49 independent SNPs and explained 1.7% of the variance in PD status (Supplementary Table 3).

The list of 49 independent SNPs encompasses 35 lysosomal pathway genes, of which 14 SNPs are novel and not reported in previous PD case-control GWAS (*CDC42, LHCGR, BRI3, CLVS1, DENND3, KCNQ1, DRAM1, AP5M1, PLEKHM1, SPPL2C, CLTCL1, MAPK1, ARSA*; as detailed in Supplementary Table 4 and 5). Out of the 35 not novel SNPs, three were already known to be associated with PD risk (RAB29/rs11557080, SNCA/rs1372518 and LRRK2/rs76904798), whereas 28 variants were in LD with previously reported PD GWAS signals based on LDpair analysis (34). Three SNPs within known PD-associated genes were not in LD with previously reported GWAS loci (GAK/rs6964, *SLC2A13/* rs141927107*, SLC2A13/* rs12296462 and *LRRK2*/ rs4567538). For each novel SNP, we meta-analysed effect sizes across the 22 European cohorts and combined the resulting estimate with the original effect size reported in the GWAS summary statistics(1). This confirmed a significant effect for all novel SNPs identified, although considerable heterogeneity was observed, likely reflecting the varying availability of genotype data for each SNP across the COURAGE-PD cohorts (Supplementary Table 5).

SNP-fixed approach

The SNP-fixed approach retained, on average, 45 of the 49 optimised SNPs across the 21 European cohorts (Table 1; Supplementary Table 6). Inclusion of LP-PRS modestly improved prediction of PD case-control status relative to covariate-only models (Table 1, Supplementary Table 6, Supplementary Fig. 2a). Meta-analysis demonstrated a significant association between LP-PRS and PD risk (OR = 1.40, 95% CI = [1.35,1.46]) with low between-cohort heterogeneity (Q = 21.84, p = 0.41; I^2 = 3.9%; τ^2 = 0) (Fig. 1a). Similar effect estimates were observed under random-effects models (Supplementary Fig. 3a). Sensitivity analyses excluding *LRRK2* and *GBA1* yielded comparable results (Supplementary Fig. 4-7a). Likewise, analyses performed without covariates showed similar overall effect estimates, although heterogeneity modestly increased (Supplementary Figs. 8-15a).

LD-adaptive approach

Application of the optimised LD clumping and thresholding parameters across cohorts retained a similar number of variants to the SNP-fixed approach (Supplementary Table 7), although cohort-specific differences in SNP selection were observed. In several cohorts, the LD-adaptive approach identified different (i.e. *TMEM163*) or additional (i.e. *SNCA*) nearby SNPs, or variants in other genes: *UBQLN4*, *HLA-DQA1* and *CUBN* picked in more than half of the cohorts, and *MARCHF1*, *SYK, SPNS1* and *CACNG2* in less than 30% of the European COURAGE-PD cohorts (Supplementary Fig. 16).

Predictive performance was broadly comparable to that observed under the SNP-fixed approach, with modestly higher AUC values in several cohorts (Table 1, Supplementary Fig. 2b). Meta-analysis demonstrated a significant association between LP-PRS and PD risk (OR = 1.42, 95% CI = [1.36,1.47]) with low between-cohort heterogeneity (Q = 22.38, p = 0.38; I^2 = 6.1%; τ^2 = 0) (Fig. 1b). Similar effect estimates were observed following exclusion of *LRRK2* and *GBA1* and under random-effects models (Supplementary Figs. 3-15b).

Transferability of LP-PRS in East Asian cohorts

To evaluate cross-ancestry transferability, LP-PRS was assessed in four independent East Asian cohorts using both the SNP-fixed and LD-adaptive approaches. Relative to the European cohorts, approximately half of the optimised variants were retained in the East Asian datasets (Supplementary Tables 8 and 9).

Despite reduced SNP retention, both approaches demonstrated significant associations between LP-PRS and PD risk, although with smaller effect sizes than those observed in European cohorts (OR = 1.15, 95% CI = [1.03,1.29]). Between-cohort heterogeneity remained low across analyses with and without covariate adjustment (Fig. 2; Supplementary Figs. 17-19). These findings support the partial transferability of lysosomal pathway-associated PRS across ancestries.

Evaluation of previously reported lysosomal pathway genes

Previous studies have implicated rare variants in lysosomal pathway genes as contributors to PD risk (7,10). Because the NeuroChip array provides limited coverage of rare variants, direct replication of prior rare variant findings was not feasible(7). However, several previously implicated lysosomal genes, including *SCARB2*, *CLN8*, and *ARSA*, were represented within the LP-PRS dataset.

To further evaluate the contribution of previously reported lysosomal modifier genes, PRS analyses were restricted to genes within the Commander pathway gene set (Supplementary Table 2). All genes within this set were represented in the summary statistics dataset. Optimisation identified 19 SNPs encompassing 10 genes (Supplementary Tables 10 and 11). Five of these genes (*REL*, *RAB7A*, *VPS41*, *VPS39*, and *IGF1R*) had previously been reported as candidate modifiers of lysosomal activity (10)**.** Associations between Commander pathway PRS and PD risk were modest and observed inconsistently across cohorts (Supplementary Table 12; Supplementary Figs. 20–27). Notably, none of the Commander pathway genes were retained in the optimised LP-PRS models derived from the genome-wide lysosomal pathway analyses.

Cell-type-specific differential expression of PD risk genes across snRNA-Seq datasets

To characterise the cellular distribution and disease-associated regulation of the subset of novel genes selected in the LP-PRS (Supplementary Table 5), four single-nucleus RNA sequencing datasets derived from the substantia nigra, midbrain, and putamen (regions critically affected in PD) were integrated and analysed. The set of genes evaluated demonstrated robust and reproducible differential expression across datasets and cell types. Among these, *CLVS1* showed the most consistent cross-cohort signal, exhibiting significant downregulation in oligodendrocytes as well as differential expression within neuronal populations and most prominently in GABAergic neurons across the majority of cohorts, indicating both glial and synaptic involvement in PD-associated transcriptional changes (Fig.3, Supplementary Fig.28, Supplementary Table 13). Microglial transcriptional alterations were likewise reproducible but exhibited regional specificity. *DENND3* was significantly downregulated in microglia across the substantia nigra and midbrain datasets. Notably, these microglial signals were not detected in the putamen dataset (Fig.3, Supplementary Fig.28, Supplementary Table 13). The dopaminergic neuron–enriched fraction of dataset GSE178265 ($NR4A2^{+}$) provided enhanced resolution for neuron-specific effects. Given the selective vulnerability and frequent underrepresentation of dopaminergic neurons in standard single-cell and single-nucleus datasets, this enrichment enabled the detection of transcriptional changes that were less apparent in other cohorts. Within this population, several genes including *CDC42***,** *CLVS1***,** *DENND3*, *KCNQ1* and *CLTCL1* showed significant differential expression. The reduced detectability of these signals in other datasets is consistent with limited dopaminergic cell representation rather than the absence of disease-related effects (Fig.3; Supplementary Fig.28; Supplementary Table 13). Importantly, a subset of these neuron-associated genes, most notably *CLTCL1* and *CLVS1,* also demonstrated significant differential expression in additional neuronal subclasses across independent cohorts, supporting

broader neuronal involvement beyond dopaminergic populations (Fig.3, Supplementary Fig.28; Supplementary Table 13).

Longitudinal risk stratification analyses

Longitudinal follow-up data were available for three COURAGE-PD cohorts (23% of the 2998 subjects in the low and high-risk quantiles), allowing exploratory evaluation of clinical outcomes across PRS deciles (Fig. 4). Among controls classified in the high-risk quantile, four individuals developed PD over approximately 10 years of follow-up (Age at onset: 67.8 ± 7.7; Supplementary Table 14), corresponding to a conversion rate of 3% with the SNP-fixed approach and 2% with the LD-adaptive method. Six additional controls received new diagnoses unrelated to PD (Alzheimer's disease: 2, mild cognitive impairment: 1, other diagnoses: 3), and one subject developed restless legs syndrome (RLS), who was classified as high-risk only by the SNP-fixed approach. Among PD cases, 22 individuals had a revised diagnosis during follow-up (Supplementary Table 15). Of the nine subjects with PD (including PD with peripheral neuropathy, PD with epilepsy, and Parkinson's disease dementia), five were correctly classified as high-risk, two were identified as high-risk by only one PRS method, and two were classified as low-risk. Among ten subjects diagnosed with Parkinsonism, three were classified as low-risk, four as high-risk, and three had differing classifications between the two methods. All three subjects re-evaluated as having other neurodegenerative, non-PD diseases were either classified as low-risk or not included in any risk quantile. Although limited by small sample size, these findings suggest that LP-PRS stratification may capture clinically relevant differences in longitudinal disease risk and diagnostic trajectories.

Discussion

To date, the contribution of common variants in lysosomal genes to explain part of the missing heritability of PD has not been explored at a large scale. In this study, we investigated the polygenic involvement of the lysosomal pathway in PD risk in a large, multiethnic case–control cohort, finding a significant lysosomal gene contribution to PD susceptibility. Our study unequivocally established that, independent of risk driven by *GBA1* and *LRRK2*, there are many common variants within the lysosomal pathway that are significant risk factors for PD.

We developed an LP-PRS approach integrating curated lysosomal pathway genes with two complementary PRS strategies: a SNP-fixed approach based on direct transfer of the optimised SNP set across cohorts, and an LD-adaptive approach that retained cohort-specific variants using fixed LD clumping and thresholding parameters. Both strategies yielded highly consistent effect estimates and low between-cohort heterogeneity across European cohorts, supporting the robustness of the observed associations. Importantly, the LP-PRS signal remained significant after exclusion of *LRRK2* and *GBA1*, suggesting that the association is not driven solely by established high-effect PD genes but instead reflects a broader cumulative contribution of common variants across lysosomal pathway genes.

Pathway-informed PRS models may provide advantages beyond conventional genome-wide PRS approaches by capturing biologically coherent genetic architecture. While genome-wide PRS models maximise predictive performance by incorporating variants

across the genome, pathway-specific PRS approaches enable direct investigation of biological systems implicated in disease pathogenesis. As a prove of concept, a recent study on mitochondrial PRS (mitoPRS) functionally validated for the first time that high mitoPRS translates into impaired mitochondrial function in neurons from PD patients that was not observed in low mitoPRS using patient-derived induced pluripotent stem cells (iPSC) from genetically defined individuals(16). In PD, mounting evidence supports a central role for endolysosomal trafficking, autophagy, and vesicular transport pathways in α-synuclein accumulation and neuronal degeneration (7,35,36). Our findings suggest that common variants distributed across lysosomal pathway genes collectively contribute to PD susceptibility and further support the hypothesis that endolysosomal biology represents a core component of PD genetic architecture.

The underlying allelic and genetic heterogeneity, driven by many factors, including population and LD differences, is an important contributor to a disease (37). Taking this into account, our study aimed to implement a pathway-based PRS approach using two different transfer frameworks to incorporate the impact of allelic and genetic heterogeneity. Indeed, the different selection of SNPs spanning genes associated with the LP-PRS in the Asian population highlighted the validity of our model for PRS-based approaches. Although most currently available PD GWAS summary statistics have been derived predominantly from European populations(1,38), the LP-PRS remained significantly associated with PD risk in East Asian cohorts. Effect sizes were attenuated, and fewer variants were retained relative to European cohorts, likely reflecting differences in LD structure and allele frequency distributions between ancestries. Nevertheless, the persistence of significant associations across independent East Asian cohorts supports the transferability of lysosomal pathway-associated polygenic risk across populations. These observations are consistent with broader PRS literature demonstrating reduced but retained predictive performance when European-derived PRS models are applied to non-European populations (39,40).

Assessment of SNPs selected during LP-PRS parameter optimisation showed that the majority are located within well-established PD genes (i.e. *SNCA*, *LRRK2*, *SCARB2*, *GAK*, *TMEM175* and *RAB29*). Interestingly, several genes included in the LP-PRS (*CLN8*, *TBC1D5*, *CTSB*, *SLC2A13*, *USP8*, and *LMBRD1*) have been reported as GWAS loci only in more recent studies. Despite this, they were identified using the older summary statistics employed here, highlighting the ability of our approach to prioritize biologically relevant genes even before they reached genome-wide significance in larger and more recent GWAS analyses. Our LP-PRS also picked SNPs in genes not previously associated with PD risk. Out of the novel 13 genes, ARSA is the most promising candidate, as a contribution of rare variants has already been reported (41). Of interest*, CDC42, BRI3, CLVS1, DENND3* and *CLTCL1* were found differentially expressed in PD against controls in our functional validation using scRNA-seq datasets. Cell Division Control Protein 42 (*CDC42*) involvement in PD has been suggested. This small GTPase regulates the growth and degeneration of dopaminergic neurons and interacts with α-synuclein (29). Of note, its downregulation has been reported in several areas of postmortem brains derived from PD patients(42). Moreover, deleterious loss-of-function variants in its activator, *ITSN1*, have also been recently associated with an increased PD risk (43). Brain Protein I3 (*BRI3*)

is a transmembrane protein containing the endoplasmic reticulum luminal chaperone domain BRICHOS. Proteins with BRICHOS domains (i.e. *BRI2*) have been reported to directly bind to oligomeric species, preventing α-synuclein secondary nucleation reactions (44,45). DENN domain-containing protein 3 (*DENND3*) specifically activates Rab12, a GTPase recently reported to inhibit astrocytic primary ciliogenesis when binding *LRRK2*(46). Clavesin 1 (*CLVS1*) is a neuron-specific membrane-associated protein that interacts with endosomes and Golgi vesicle network with clathrin coat components. The clavesin family is essential for the regulation of lysosomal morphology (46,47). However, to date, there is no evidence linking clavesins disregulation to PD. Clathrin Heavy polypeptide-like 1 (*CLTCL1*) encodes for the minor clathrin heavy chain and is also involved in endosomal trafficking(48). These results strongly suggest that the cumulative dysregulation of lysosomal pathway proteins may cooperatively impair the machinery required for proper lysosomal trafficking and homeostasis.

Considering the limited availability of longitudinal trajectories of COURAGE-PD cohorts, the LP-PRS mostly classified subjects as high-risk when having PD or atypical Parkinsonism, while placing individuals having other non-PD diseases in the lower risk quantile. When evaluating controls with a high risk of developing PD, we saw a conversion rate of 3%. Although limited by a small sample size, as a proof of principle our approach highlighted that the longitudinal evaluation of high-risk controls is a viable and effective strategy not only to understand the progression of a disease trajectory, but will also allow the development of intervention strategies at the “crucial window” to either delay or stop the disease onset or progression. Thus, the stratification of patients and high-risk controls, as definite proof of absolute genetic measure, solely based on the individual’s genetic background, would allow a robust, cost-effective screening to identify an important at-risk cohort for trials of selective targeted experimental disease modifying therapies. These observations suggest that pathway-informed genetic stratification may identify clinically relevant differences in longitudinal disease risk. Larger prospective studies will be necessary to determine whether pathway-specific PRS models will contribute to PD risk stratification and early disease prediction.

A limitation of this study is the reduced accuracy of array-based techniques in detecting *GBA1* variants, owing to the presence of a highly homologous pseudogene. Therefore, we cannot exclude that undetected *GBA1* variants may still be present in high-risk individuals and contribute to the phenotype. To minimise this effect, we excluded all previously known *GBA1* carriers from the analysis and computed the LP-PRS excluding the gene, as also done for *LRRK2*. As the cost of next-generation sequencing (NGS) continues to decrease, future studies employing these techniques may provide a clearer picture of the frequency and contribution of *GBA1* variants to lysosomal susceptibility(49–51).

Conclusion

Our findings demonstrate that common variants within lysosomal pathway genes contribute to PD susceptibility across ancestries and further support a central role for

endolysosomal biology in PD pathogenesis. More broadly, this study highlights the utility of biologically informed pathway-specific PRS approaches for investigating disease mechanisms and genetic stratification in complex neurodegenerative disorders.

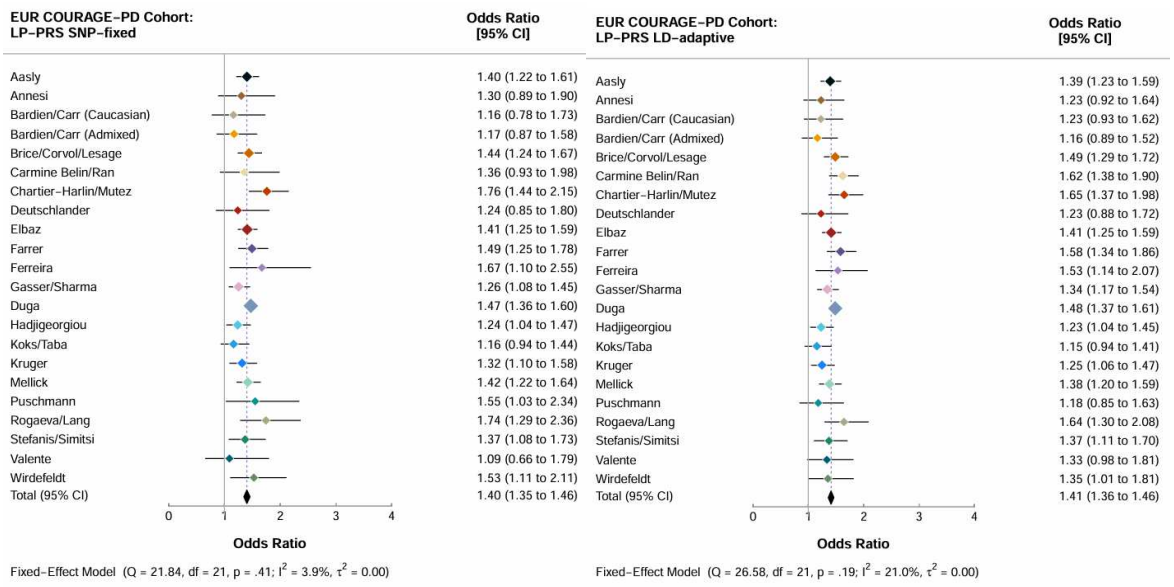
Table 1: LP-PRS Performance Metrics Across the European COURAGE-PD Cohorts Using the SNP-Fixed and LD-Adaptive Approaches

Cohort	SNP-fixed approach						LD-adaptive approach					
	Num selected SNP	Effect Size (SE)	AUC_{FULL} [0.95 C.I.]	AUC_{NULL} [0.95 C.I.]	AUC_{DELTA}	P-value	Num selected SNP	Effect Size (SE)	AUC_{FULL} [0.95 C.I.]	AUC_{NULL} [0.95 C.I.]	AUC_{DELTA}	P-value
Aasly	46	0.339 (0.071)	0.661 [0.626–0.695]	0.627 [0.592–0.662]	0.034	2.10E-03	54	0.364 (0.072)	0.663 [0.628–0.697]	0.627 [0.592–0.662]	0.036	1.90E-03
Annesi	46	0.265 (0.192)	0.849 [0.794–0.904]	0.841 [0.785–0.897]	0.008	1.94E-01	47	0.258 (0.194)	0.849 [0.794–0.905]	0.841 [0.785–0.897]	0.009	1.55E-01
Bardien/Carr	46	0.149 (0.203)	0.895 [0.853–0.938]	0.895 [0.852–0.938]	0.001	7.45E-01	53	0.183 (0.204)	0.895 [0.853–0.938]	0.895 [0.852–0.938]	0	8.78E-01
Bardien/Carr	47	0.159 (0.153)	0.731 [0.664–0.797]	0.727 [0.660–0.794]	0.004	5.98E-01	54	0.157 (0.153)	0.735 [0.668–0.801]	0.727 [0.660–0.794]	0.008	3.00E-01
Brice/Corvol/Lesage	42	0.364 (0.075)	0.615 [0.577–0.653]	0.565 [0.526–0.605]	0.05	9.20E-03	45	0.395 (0.075)	0.620 [0.582–0.658]	0.565 [0.526–0.605]	0.055	4.40E-03
Carmine Belin/Ran	46	0.304 (0.194)	0.741 [0.648–0.834]	0.724 [0.622–0.825]	0.017	2.68E-01	54	0.299 (0.198)	0.743 [0.651–0.835]	0.724 [0.622–0.825]	0.019	2.65E-01
Chartier-Harlin/Mutez	46	0.566 (0.101)	0.703 [0.659–0.747]	0.647 [0.599–0.696]	0.056	1.30E-03	52	0.561 (0.101)	0.703 [0.658–0.747]	0.647 [0.599–0.696]	0.056	1.20E-03
Deutschländer	36	0.215 (0.191)	0.736 [0.653–0.820]	0.733 [0.645–0.822]	0.003	7.29E-01	42	0.109 (0.190)	0.733 [0.647–0.819]	0.733 [0.645–0.822]	0	9.86E-01
Elbaz	48	0.341 (0.061)	0.610 [0.578–0.643]	0.558 [0.525–0.591]	0.052	2.50E-03	51	0.367 (0.062)	0.617 [0.584–0.649]	0.558 [0.525–0.591]	0.059	9.00E-04
Farrer	46	0.401 (0.089)	0.740 [0.703–0.778]	0.726 [0.687–0.764]	0.015	1.12E-01	52	0.436 (0.090)	0.742 [0.705–0.780]	0.726 [0.687–0.764]	0.017	8.44E-02
Ferreira	49	0.514 (0.215)	0.885 [0.826–0.943]	0.876 [0.817–0.935]	0.009	3.51E-01	49	0.514 (0.215)	0.885 [0.826–0.943]	0.876 [0.817–0.935]	0.009	3.51E-01
Gasser/Sharma	46	0.227 (0.075)	0.679 [0.643–0.715]	0.670 [0.634–0.707]	0.008	1.99E-01	51	0.265 (0.075)	0.682 [0.647–0.718]	0.670 [0.634–0.707]	0.012	1.08E-01
Duga	**49**	**0.388 (0.043)**	**0.710 [0.690–0.729]**	**0.688 [0.668–0.708]**	**0.021**	**2.46E-05**	**49**	**0.388 (0.043)**	**0.710 [0.690–0.729]**	**0.688 [0.668–0.708]**	**0.021**	**2.46E-05**
Hadjigeorgiou	46	0.213 (0.087)	0.606 [0.560–0.652]	0.591 [0.545–0.637]	0.015	2.85E-01	47	0.215 (0.087)	0.607 [0.561–0.652]	0.591 [0.545–0.637]	0.015	2.63E-01
Koks/Taba	45	0.151 (0.109)	0.609 [0.552–0.665]	0.598 [0.541–0.654]	0.011	3.47E-01	53	0.162 (0.109)	0.610 [0.554–0.667]	0.598 [0.541–0.654]	0.013	3.23E-01
Kruger	44	0.274 (0.093)	0.750 [0.711–0.789]	0.736 [0.697–0.776]	0.014	2.49E-02	49	0.272 (0.094)	0.751 [0.712–0.790]	0.736 [0.697–0.776]	0.014	1.91E-02
Mellick	46	0.350 (0.075)	0.642 [0.604–0.679]	0.611 [0.573–0.649]	0.031	2.43E-02	51	0.350 (0.075)	0.640 [0.602–0.677]	0.611 [0.573–0.649]	0.028	3.75E-02
Puschmann	45	0.440 (0.208)	0.811 [0.740–0.883]	0.791 [0.715–0.868]	0.02	1.15E-01	52	0.449 (0.210)	0.816 [0.746–0.886]	0.791 [0.715–0.868]	0.024	7.49E-02
Rogaeva/Lang	47	0.556 (0.154)	0.857 [0.816–0.897]	0.842 [0.799–0.885]	0.015	7.96E-02	53	0.578 (0.155)	0.859 [0.819–0.900]	0.842 [0.799–0.885]	0.017	4.43E-02
Stefanis/Simitsi	44	0.315 (0.120)	0.718 [0.665–0.771]	0.703 [0.648–0.757]	0.015	1.46E-01	44	0.315 (0.120)	0.718 [0.665–0.771]	0.703 [0.648–0.757]	0.015	1.46E-01
Valente	43	0.085 (0.255)	0.722 [0.612–0.832]	0.721 [0.611–0.830]	0.001	8.82E-01	44	0.073 (0.254)	0.722 [0.612–0.832]	0.721 [0.611–0.830]	0.001	8.68E-01
Wirdefeldt	45	0.424 (0.164)	0.682 [0.610–0.754]	0.655 [0.580–0.731]	0.027	2.88E-01	52	0.380 (0.162)	0.683 [0.612–0.755]	0.655 [0.580–0.731]	0.028	2.43E-01

Evaluation of the association between LP-PRS and PD status across the European COURAGE-PD cohorts. Logistic regression was performed independently for each cohort to assess the association between LP-PRS and PD status, adjusting for sex, age, and the first 10 principal components (PCs). Null models were generated by excluding the LP-PRS from the regression model. The predictive performance of the LP-PRS was evaluated using receiver operating characteristic (ROC) curves and the corresponding area under the curve (AUC) estimates. Results obtained from the training dataset are shown in bold.

Figure 1: Fixed-effect meta-analysis of LP-PRS across all the European COURAGE-PD Cohorts.

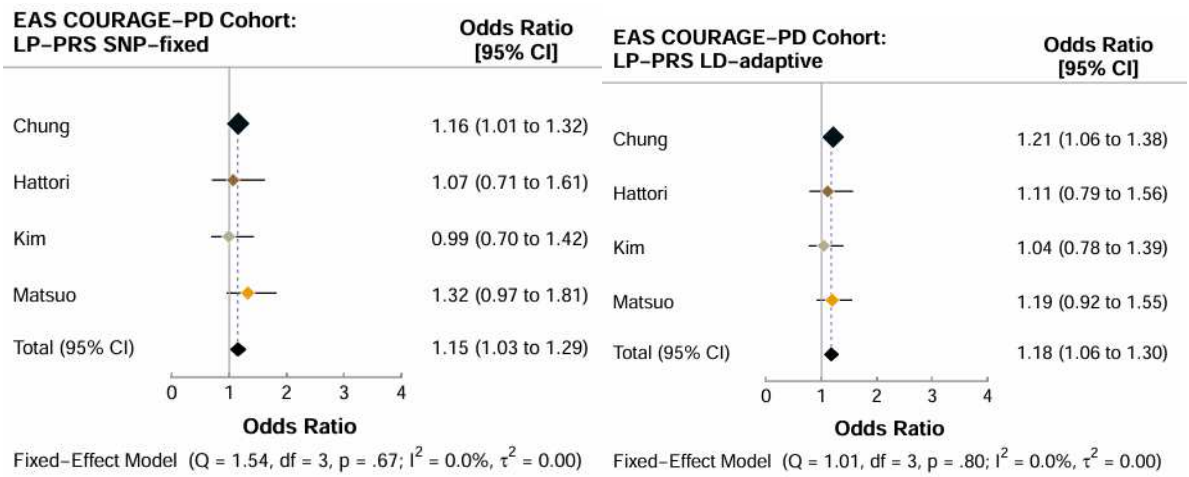
a. **b.**



LP-PRS meta-analysis of all European COURAGE-PD effect sizes. Effect sizes and standard errors were obtained by fitting a binomial glm model with each cohort standardized PRS and adjusting for age, sex and the first ten principal components. Meta-analysis was performed using the R package metafor, using the inverse-variance method to fit a fixed-effect model. Results were plotted using the forest() function provided in metafor.

Figure 2: Fixed-effect meta-analysis of LP-PRS across all the East Asian COURAGE-PD Cohorts.

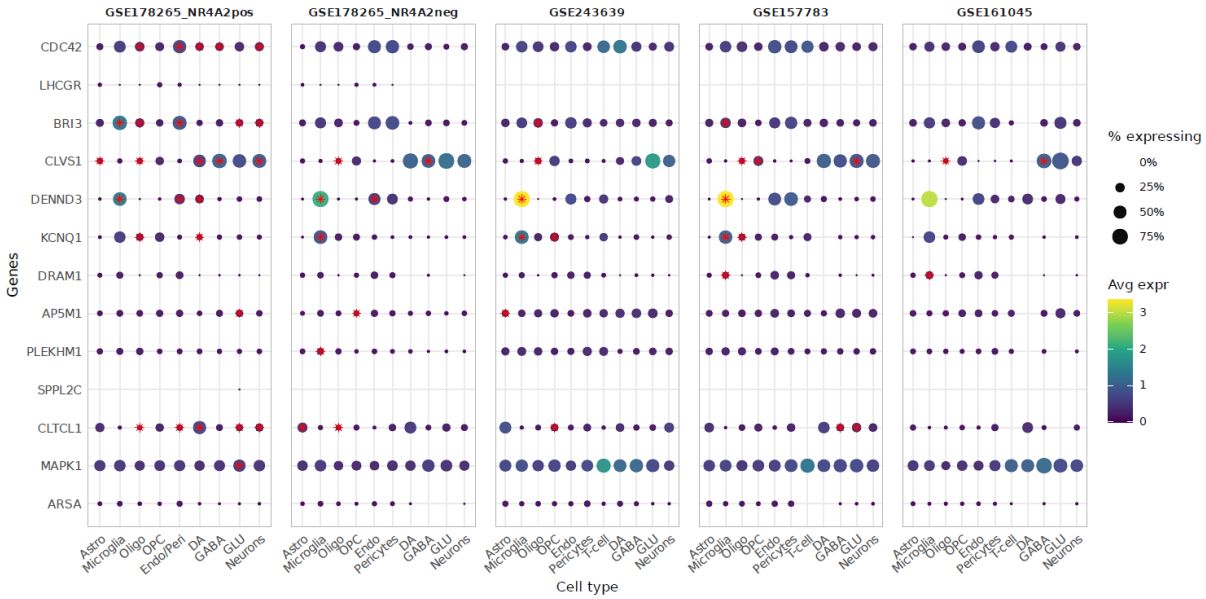
a. **b.**



LP-PRS meta-analysis of all East Asian COURAGE-PD effect sizes. Effect sizes and standard errors were obtained by fitting a binomial glm model with each cohort standardized PRS and adjusting for age, sex and the first ten principal components. Meta-

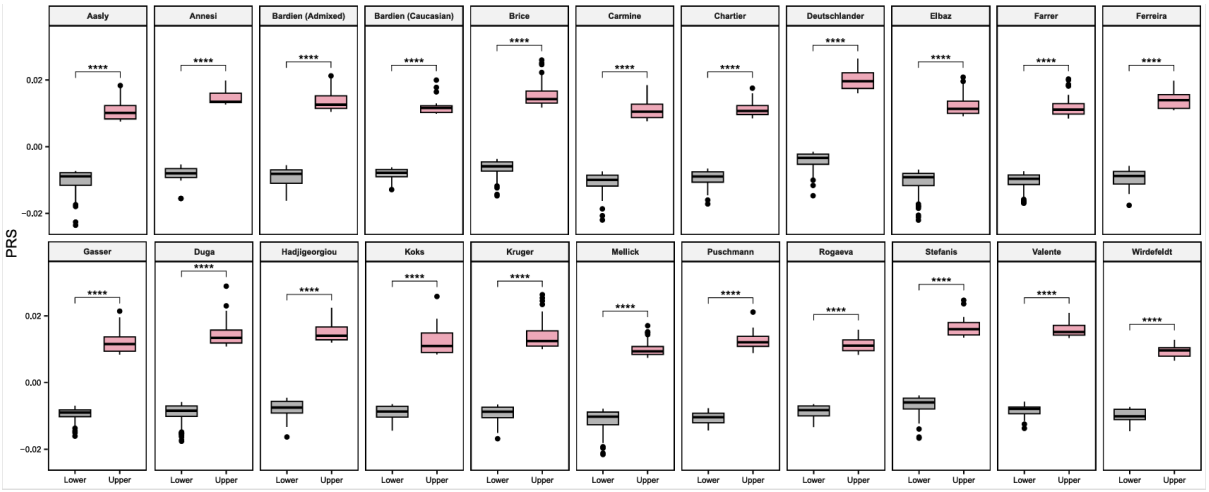
analysis was performed using the R package metafor, using the inverse-variance method to fit a fixed-effect model. Results were plotted using the forest() function provided in metafor.

Figure 3: Cell-type–resolved expression of PD risk genes across brain snRNA-Seq datasets



Dot plot showing gene expression of the novel genes across major brain cell types in multiple snRNA-seq datasets. Each facet represents an individual cohort. Shared cell types are displayed first to enable cross-cohort comparison, followed by dataset-specific populations. Dot colour denotes mean expression level, dot size indicates the proportion of expressing cells, and red asterisks mark significant differential expression (adjusted p<0.01). **Abbreviations:** Astro, astrocytes; Microglia, microglial cells; Oligo, oligodendrocytes; OPC, oligodendrocyte precursor cells; Endo, endothelial cells; Pericytes, pericytes; T cell, T lymphocytes; DA, dopaminergic neurons; GABA, GABAergic neurons; GLU, glutamatergic neurons; Neurons, unspecified neuronal subtypes.

Figure 4: Distribution of individuals of each European COURAGE-PD Cohort in the Lower and Upper PRS risk quantiles



Individuals in the top and bottom 10% of each European COURAGE-PD cohort's PRS distribution were identified and labelled as having high-risk and low-risk of developing PD. The boxplots show a significant difference in distribution between high and low risk quantiles. Boxplots were drawn using the R package ggpubr.

Acknowledgements

PD was supported by grants from the Italian Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006) – A multiscale integrated approach to the study of the nervous system in healthy and disease (DN. 1553 11.10.2022).

References

1. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol. 2019 Dec;18(12):1091–102.

2. Zimprich A, Benet-Pagès A, Struhal W, Graf E, Eck SH, Offman MN, et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. Am J Hum Genet. 2011 July 15;89(1):168–75.

3. Gustavsson EK, Follett J, Trinh J, Barodia SK, Real R, Liu Z, et al. RAB32 Ser71Arg in autosomal dominant Parkinson's disease: linkage, association, and functional analyses. Lancet Neurol. 2024 June;23(6):603–14.

4. Quadri M, Fang M, Picillo M, Olgiati S, Breedveld GJ, Graafland J, et al. Mutation in the SYNJ1 gene associated with autosomal recessive, early-onset Parkinsonism. Hum Mutat. 2013 Sept;34(9):1208–15.

5. Westenberger A, Brüggemann N, Klein C. Genetics of Parkinson's disease: From causes to treatment. Cold Spring Harb Perspect Med. 2025 July 1;15(7):a041774.

6. Ohnmacht J, May P, Sinkkonen L, Krüger R. Missing heritability in Parkinson's disease: the emerging role of non-coding genetic variation. J Neural Transm (Vienna). 2020 May;127(5):729–48.

7. Robak LA, Jansen IE, van Rooij J, Uitterlinden AG, Kraaij R, Jankovic J, et al. Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. Brain. 2017 Dec 1;140(12):3191–203.

8. Blauwendraat C, Reed X, Krohn L, Heilbron K, Bandres-Ciga S, Tan M, et al. Genetic modifiers of risk and age at onset in GBA associated Parkinson's disease and Lewy body dementia. Brain. 2020 Jan 1;143(1):234–48.

9. Straniero L, Rimoldi V, Monfrini E, Bonvegna S, Melistaccio G, Lake J, et al. Role of lysosomal gene variants in modulating GBA-associated Parkinson's disease risk. Mov Disord. 2022 June;37(6):1202–10.

10. Minakaki G, Safren N, Bustos BI, Lubbe SJ, Mencacci NE, Krainc D. Commander complex regulates lysosomal function and is implicated in Parkinson's disease risk. Science. 2025 Apr 11;388(6743):204–11.

11. Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med. 2009 Oct 22;361(17):1651–61.

12. Lesage S, Anheim M, Condroyer C, Pollak P, Durif F, Dupuits C, et al. Large-scale screening of the Gaucher's disease-related glucocerebrosidase gene in Europeans with Parkinson's disease. Hum Mol Genet. 2011 Jan 1;20(1):202–10.

13. Ross OA, Soto-Ortolaza AI, Heckman MG, Aasly JO, Abahuni N, Annesi G, et al. Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: a case-control study. Lancet Neurol. 2011 Oct;10(10):898–908.

14. Chang D, Nalls MA, Hallgrímsdóttir IB, Hunkapiller J, van der Brug M, Cai F, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nat Genet. 2017 Oct;49(10):1511–6.

15. Do CB, Tung JY, Dorfman E, Kiefer AK, Drabant EM, Francke U, et al. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. PLoS Genet. 2011 June;7(6):e1002141.

16. Arena G, Landoulsi Z, Grossmann D, Payne T, Vitali A, Delcambre S, et al. Polygenic risk scores validated in patient-derived cells stratify for mitochondrial subtypes of Parkinson's disease. Ann Neurol. 2024 July;96(1):133–49.

17. Grover S, Ashwin AKS, Pihlstrom L, Domenighetti C, Schulte C, Sugier P-E, et al. Genome-wide association and meta-analysis of age-at-onset in Parkinson disease:

Evidence from COURAGE-PD consortium. Neurology. 2022 May 26;99(7):e698–710.

18. Domenighetti C, Sugier P-E, Sreelatha AAK, Schulte C, Grover S, Mohamed O, et al. Mendelian randomisation study of smoking, alcohol, and coffee drinking in relation to Parkinson's disease. J Parkinsons Dis. 2022;12(1):267–82.

19. Domenighetti C, Sugier P-E, Ashok Kumar Sreelatha A, Schulte C, Grover S, Portugal B, et al. Association of body mass index and Parkinson disease: A bidirectional Mendelian randomization study. Neurology. 2024 Aug 13;103(3):e209620.

20. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020 Sept;15(9):2759–72.

21. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015 Oct 1;526(7571):68–74.

22. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience. 2019 July 1;8(7):giz082.

23. Pereira GM, Teixeira-Dos-Santos D, Soares NM, Marconi GA, Friedrich DC, Saffie Awad P, et al. A systematic review and meta-analysis of the prevalence of Parkinson's disease in lower to upper-middle-income countries. NPJ Parkinsons Dis. 2024 Sept 30;10(1):181.

24. Viechtbauer W. Conducting Meta-Analyses inRwith themetaforPackage. J Stat Softw [Internet]. 2010;36(3). Available from: http://dx.doi.org/10.18637/jss.v036.i03

25. Kamath T, Abdulraouf A, Burris SJ, Langlieb J, Gazestani V, Nadaf NM, et al. Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. Nat Neurosci. 2022 May;25(5):588–95.

26. Martirosyan A, Ansari R, Pestana F, Hebestreit K, Gasparyan H, Aleksanyan R, et al. Unravelling cell type-specific responses to Parkinson's Disease at single cell resolution. Mol Neurodegener. 2024 Jan 20;19(1):7.

27. Xu J, Farsad HL, Hou Y, Barclay K, Lopez BA, Yamada S, et al. Human striatal glia differentially contribute to AD- and PD-specific neurodegeneration. Nat Aging. 2023 Mar;3(3):346–65.

28. Badanjak K, Mulica P, Smajic S, Delcambre S, Tranchevent L-C, Diederich N, et al. IPSC-derived microglia as a model to study inflammation in idiopathic Parkinson's disease. Front Cell Dev Biol. 2021 Nov 5;9:740758.

29. Schnack C, Danzer KM, Hengerer B, Gillardon F. Protein array analysis of oligomerization-induced changes in alpha-synuclein protein-protein interactions

points to an interference with Cdc42 effector proteins. Neuroscience. 2008 July 17;154(4):1450–7.

30. Aleknonytė-Resch M, Trinh J, Leonard H, Delcambre S, Leitão E, Lai D, et al. Genome-wide case-only analysis of gene-gene interactions with known Parkinson's disease risk variants reveals link between LRRK2 and SYT10. NPJ Parkinsons Dis. 2023 June 29;9(1):102.

31. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024 Feb;42(2):293–304.

32. Wolock SL, Lopez R, Klein AM. Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. Cell Syst. 2019 Apr 24;8(4):281-291.e9.

33. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019 Dec;16(12):1289–96.

34. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015 Nov 1;31(21):3555–7.

35. Dehay B, Martinez-Vicente M, Caldwell GA, Caldwell KA, Yue Z, Cookson MR, et al. Lysosomal impairment in Parkinson's disease. Mov Disord. 2013 June;28(6):725–32.

36. Klein AD, Mazzulli JR. Is Parkinson's disease a lysosomal disorder? Brain. 2018 Aug 1;141(8):2255–62.

37. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet. 2001 Sept;17(9):502–10.

38. Leonard HL, Global Parkinson's Genetics Program (GP2). Novel Parkinson's disease genetic risk factors within and across European populations [Internet]. medRxiv. 2025. Available from: http://dx.doi.org/10.1101/2025.03.14.24319455

39. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019 Apr;51(4):584–91.

40. Duncan L, Shen H, Gelaye B, Meijsen J, Ressler K, Feldman M, et al. Analysis of polygenic risk score usage and performance in diverse human populations. Nat Commun. 2019 July 25;10(1):3328.

41. Makarious MB, Lake J, Pitz V, Ye Fu A, Guidubaldi JL, Solsberg CW, et al. Large-scale rare variant burden testing in Parkinson's disease. Brain. 2023 Nov 2;146(11):4622–32.

42. Zhang Y, James M, Middleton FA, Davis RL. Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. Am J Med Genet B Neuropsychiatr Genet. 2005 Aug 5;137B(1):5–16.

43. Skuladottir AT, Tragante V, Sveinbjornsson G, Helgason H, Sturluson A, Bjornsdottir A, et al. Loss-of-function variants in ITSN1 confer high risk of Parkinson's disease. NPJ Parkinsons Dis. 2024 Aug 15;10(1):140.

44. Adam L, Kumar R, Arroyo-Garcia LE, Molenkamp WH, Nowak JS, Klute H, et al. Specific inhibition of α-synuclein oligomer generation and toxicity by the chaperone domain Bri2 BRICHOS. Protein Sci. 2024 Aug;33(8):e5091.

45. Ghosh D, Torres F, Schneider MM, Ashkinadze D, Kadavath H, Fleischmann Y, et al. The inhibitory action of the chaperone BRICHOS against the α-Synuclein secondary nucleation pathway. Nat Commun. 2024 Nov 20;15(1):10038.

46. Li X, Zhu H, Huang BT, Li X, Kim H, Tan H, et al. RAB12-LRRK2 complex suppresses primary ciliogenesis and regulates centrosome homeostasis in astrocytes. Nat Commun. 2024 Sept 29;15(1):8434.

47. Katoh Y, Ritter B, Gaffry T, Blondeau F, Höning S, McPherson PS. The clavesin family, neuron-specific lipid- and clathrin-binding Sec14 proteins regulating lysosomal morphology. J Biol Chem. 2009 Oct 2;284(40):27646–54.

48. Nahorski MS, Borner GHH, Shaikh SS, Davies AK, Al-Gazali L, Antrobus R, et al. Clathrin heavy chain 22 contributes to the control of neuropeptide degradation and secretion during neuronal development. Sci Rep. 2018 Feb 5;8(1):2340.

49. Toffoli M, Chen X, Sedlazeck FJ, Lee C-Y, Mullin S, Higgins A, et al. Comprehensive short and long read sequencing analysis for the Gaucher and Parkinson's disease-associated GBA gene. Commun Biol. 2022 July 6;5(1):670.

50. Cuconato G, Palmieri I, Percetti M, Pagliarani S, Tenace S, Morelli MJ, et al. LONG-NEXT: A new accurate and efficient NGS-based method for GBA1 analysis in Parkinson disease. Parkinsonism Relat Disord. 2025 May;134(107780):107780.

51. Pachchek S, Landoulsi Z, Pavelka L, Schulte C, Buena-Atienza E, Gross C, et al. Accurate long-read sequencing identified GBA1 as major risk factor in the Luxembourgish Parkinson's study. NPJ Parkinsons Dis. 2023 Nov 23;9(1):156.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.xlsx
- supplementarymaterialandfigures.pdf
- githubreposcripts.zip