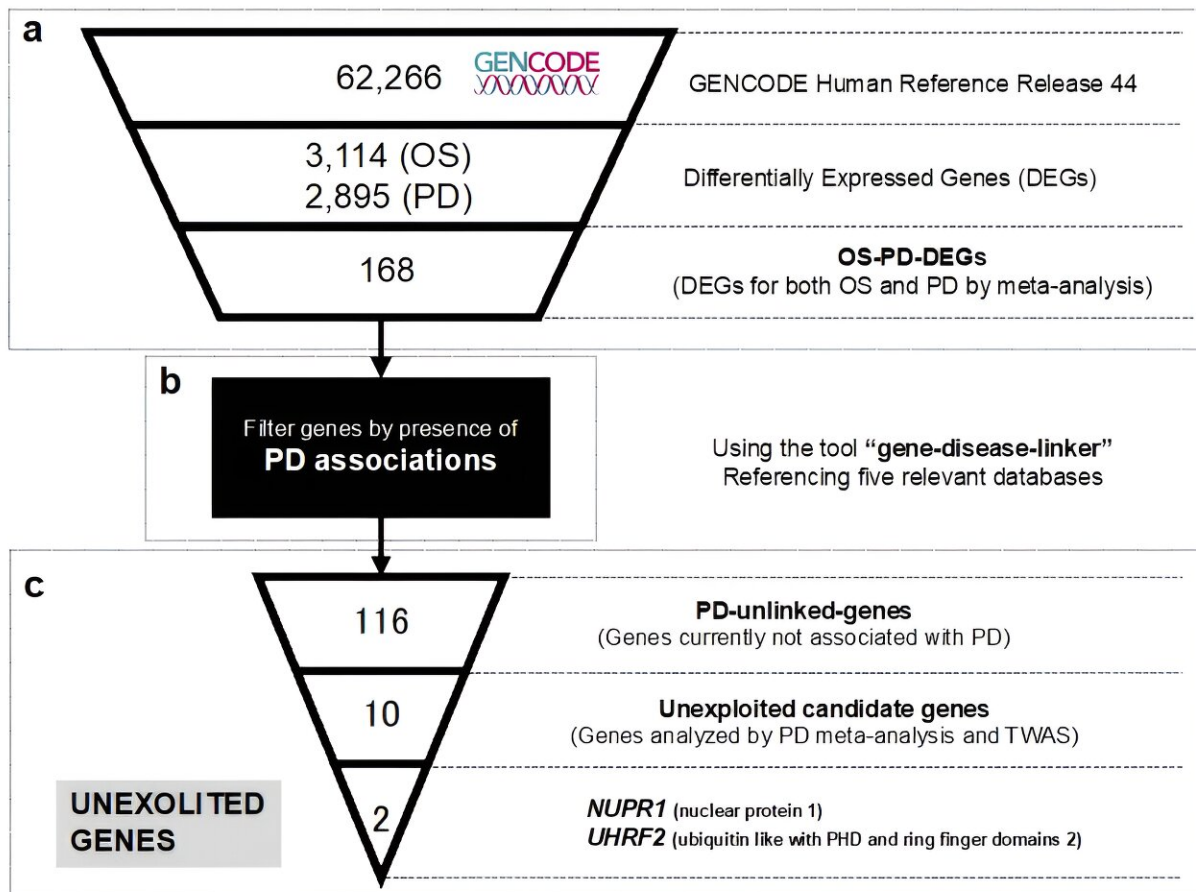# Parkinson's disease study develops analytical pipeline to identify unexploited genes

September 9 2024

A pipeline to explore candidate target genes using transcriptomics and unknomics. Credit: Takayuki Suzuki/Hiroshima University

A research team from Hiroshima University has developed an analysis

pipeline to identify unexploited genes for a given disease against five databases that provide gene-disease associations. They used their pipeline to study oxidative stress and its related disease, Parkinson's disease, as a case study.

Their results were published in the journal *npj Parkinson's Disease* on August 17, 2024.

Scientists can access human disease-associated gene data through databases, such as the Open Targets Platform, DisGeNET, miRTex, RNADisease, and PubChem. However, these databases are missing some data entries because of curational errors, biases, and text-mining failures. In addition to the missing entries, because of the extensive research on human diseases, challenges have occurred in registering comprehensive data.

Scientists recognize that the lack of essential data in databases has a negative impact on knowledge sharing and needs to be addressed. So, the research team proposed an analysis pipeline to explore missing entries of unexploited genes in the human disease-associated gene databases.

"We proposed a method to explore novel candidate genes involved in oxidative stress related to diseases by utilizing public databases," said Hidemasa Bono, a professor at the Laboratory of Genome Informatics and the Laboratory of Bio-DX, Hiroshima University. Their method is divided into three major parts.

In part 1, they identified genes that respond to both Parkinson's disease and oxidative stress by analyzing gene expression data and by referencing transcriptome-wide association study results.

In part 2, they accessed multiple public databases that archive the association between genes and diseases, narrowing down to genes that

have not been reported in association with Parkinson's disease.

In part 3, they refined the selection to those genes deemed more critical by revisiting the gene expression and transcriptome-wide association study data.

"In this study, using oxidative stress and its related disease, Parkinson's disease, as a case study, we sought to explore methods for identifying understudied genes that also hold research value," said Bono.

Their analysis revealed two unexploited genes: nuclear protein 1 (NUPR1) and ubiquitin-like with PHD and ring finger domains 2 (UHRF2). This pipeline will allow researchers to identify underrepresented disease-associated genes and gain easier access to potential disease-related functional genes.

In their search for genes in Parkinson's disease with oxidative stress, the researchers used their pipeline to filter 62,226 genes down to 168 genes that exhibited dysregulation of gene expression in both Parkinson's disease and oxidative stress contexts. Next, they classified the 168 genes into Parkinson's disease-unlinked and Parkinson's-linked genes, based on existing evidence of their involvement in Parkinson's disease.

The researchers further narrowed down the Parkinson's disease-unlinked genes to 12 unexploited candidate genes. Following a manual search of these 12 genes, they identified NUPR1and UHRF2 as unexploited genes that were absent from the current gene-disease associations databases.

The researchers note several reasons why missing entries in the gene databases may be occurring. It could be that the databases simply have not been updated recently.

A second cause for missing entries might be because of text-mining

extraction failures. A third possible cause is that some of the databases rely on expert manual curation for new data entry, allowing for the possibility of human errors or biases.

However, the newly developed analytical pipeline to identify unexploited genes mitigates these limitations in the databases.

The 12 genes the researchers identified in this study are responsive to both Parkinson's disease and oxidative stress. They have not been extensively studied in the field of Parkinson's disease research. These genes have the potential to help researchers better understand the mechanisms of oxidative stress and Parkinson's disease. The genes might also be useful in the development of new therapeutic approaches.

"Furthermore, the method proposed in this study can be applied to narrow down candidate genes involved in oxidative stress for other related diseases as well. It is expected that this will contribute to advancing research on oxidative stress and various diseases," said Bono.

The research team includes Takayuki Suzuki from Hiroshima University and Hidemasa Bono from Hiroshima University and Research Organization of Information and Systems (RIOS).

More information: Takayuki Suzuki et al, A systematic exploration of unexploited genes for oxidative stress in Parkinson's disease, *npj Parkinson's Disease* (2024). DOI: 10.1038/s41531-024-00776-1

Provided by Hiroshima University