Reduced frequency of clonal hematopoiesis in Parkinson's disease

Authors: Shixiang Sun[1], Daisy Sproviero[2], César Payán-Gómez[3], Zhenqiu Huang[1], Jan H.J. Hoeijmakers[4], Alexander Y. Maslov[1,5], Pier G. Mastroberardino[2,4,6], Jan Vijg[1,7]*

Affiliations:

[1]Department of Genetics, Albert Einstein College of Medicine, Bronx 10461, USA

[2]IFOM-ETS, The AIRC Institute of Molecular Oncology, Milan 20139, Italy

[3]Universidad Nacional de Colombia, Sede de La Paz, La Paz 111321, Colombia

[4]Erasmus MC, Rotterdam 3000 DR, the Netherlands

[5]Laboratory of Applied Genomic Technologies, Voronezh State University of Engineering Technologies, Voronezh 394000, Russia

[6]Universita' degli Studi dell'Aquila, L'Aquila 67100, Italy

[7]Center for Single-Cell Omics, School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.

*Corresponding author. Email: jan.vijg@einsteinmed.edu

Abstract:

To test if clonal hematopoiesis of indeterminate potential (CHIP) is associated with the incidence of Parkinson's Disease (PD), we analyzed blood whole exome sequencing data from 181 healthy controls and 391 PD subjects in the Parkinson's Progression Markers Initiative. We observed an age-related increase of CHIP carriers in the healthy controls. Surprisingly, the percentage of CHIP carriers was significantly lower in PD patients than in age-matched controls.

Main:

Parkinson's disease (PD) is the second most common neurodegenerative disorder, affecting 2~3% of the population over 65 years old[1]. The major neuropathological hallmarks of PD are the loss of dopaminergic neurons in the substantia nigra, and the aggregation of Lewy bodies containing α-synuclein. Evidence has been obtained that the molecular basis of PD involves mitochondrial dysfunction leading to an increased production of reactive oxygen species and oxidative DNA damage[2]. This could increase somatic mutation frequency during aging and clonally amplified somatic mutations have indeed recently been reported for different regions of the brain, affecting genes enriched in synaptic and neuronal functions as well as in blood of PD patients[3].

In blood, clonal expansion of somatic mutations in hematopoietic stem cells, known as Clonal Hematopoiesis of Indeterminate Potential (CHIP), increases with age and has been associated with cancer, cardiovascular disease and mortality in general[4]. CHIP clones often carrying cancer-associated somatic mutations are also linked with non-malignant disease states of aging[5], including neurodegenerative disease[6].

We tested the hypothesis that CHIP could be associated with PD by assessing CHIP frequencies in whole exome (WES) data from blood of PD patients and control subjects available in the Parkinson's Progression Markers Initiative (PPMI) (**Fig. 1A**)[7]. In this cohort, there is no significant difference in age between PD patients and healthy controls, with the ages at enrollment ranging from 31 to 85.2 years (P = 0.467; **Table S1**). Only PD patients diagnosed with PD and demonstrated dopaminergic deficits were enrolled in the study.

We examined known CHIP variants in 73 genes[8] with pre-applied filters[9] in the WES data involving 181 healthy controls and 391 PD subjects. The sequencing depth was 47.7±20.6 and 50.6±18.3 in healthy controls and PD patients, respectively. A variant was considered a CHIP variant if its allele frequency (VAF) was larger than 2% and the variant was supported by at least 3 reads (Methods). Of the 572 subjects, we identified 86 CHIP events in 77 subjects (**Table S2**). The top three affected genes were DNA methyltransferase 3A (*DNMT3A*), tet methylcytosine dioxygenase 2 (*TET2*) and tumor protein p53 (*TP53*) (**Fig. 1B**), which are the major affected genes that have been associated with CHIP[8].

Then, we examined if the percentage of CHIP carriers is associated with age or PD status (**Fig. 2**). First, we summarized the results by age bins (**Fig. 2A**) and found that the percentages of CHIP carriers are significantly increased with age, in both controls (P = 9.90e-7) and PD samples (P = 3.19e-4), confirming the well-documented age-related increase of CHIP in blood[4,8]. However, the increase was slower in PD samples than in controls (P = 2.79e-4). Since PD risk is exponentially increased since age 55 years[10], we then separated both controls and PD samples into younger (<55 years old; 42 controls and 95 PD patients) and older (≥55 years old; 139 controls and 296 PD patients) groups (**Fig. 2B**). In the older control group the percentage of subjects carrying CHIP variants was significantly elevated as compared to the younger group (younger vs. older: 2.38% vs. 22.30%; P = 2.06e-3), which was true also for PD patients (younger vs. older: 5.26% vs. 13.51%; P = 2.71e-2). There was no significant difference between younger controls and younger PD patients (P = 0.665). However, contrary to what we expected, the percentage of older PD patients carrying CHIP variants was significantly lower than the older age-matched controls (P = 2.58e-2).

Next, we compared the most highly mutated CHIP genes between controls and PD patients. The percentages of CHIP carriers with variants in *DNMT3A*, *TET2* and protein phosphatase,

Mg2+/Mn2+ dependent 1D (*PPM1D*) were higher in controls than in PD patients (9.94% vs. 4.86%, P = 2.77e-2), while carriers with variants in *TP53* and ASXL transcriptional regulator 1 (*ASXL1*) were lower in controls than in PD patients (0.55% vs. 2.56%, P = 0.115), suggesting a different mutational pattern on genes affected by PD.

In summary, our results indicate that the percentage of subjects with CHIP is increased in blood from older controls or PD patients. Surprisingly, the percentage of CHIP carriers was found to be lower in the older PD patients than in older controls. These results suggest that either PD protects against CHIP or CHIP protects against PD. Interestingly, CHIP has recently been reported to have a negative association with diabetic peripheral neuropathy in Type 2 Diabetes[11]. Additionally, non-peer-reviewed data suggest that in Alzheimer's disease (AD) patients, the frequency of CHIP is lower, rather than higher, compared to age-matched control subjects[12]. It should be noted that PD has previously been found negatively associated with multiple cancers, including leukemia[13]. Thus, it is possible that individuals without CHIP have a lower cancer risk and a higher likelihood of aging and developing PD. Alternatively, the higher percentage of CHIP carriers with variants in *DNMT3A*, *TET2* and *PPM1D* among controls compared to PD patients could contribute to this protective effect, e.g., the inactivation of *TET2* protects against inflammation-induced dopaminergic neuronal loss[14]. Our current findings provide a rational basis to further evaluate such possible cause and effect relationships.

Methods

Data source

Basic information of subjects is obtained from PPMI (file name: meta_data.11192021.csv; https://ida.loni.usc.edu/web/ppmi-rnaseq-app#Accessdata). Missed age information were found in[15]. Raw exome sequencing data was obtained from PPMI's data server.

Examining CHIP variants

Raw sequencing data were trimmed for adapter and low-quality nucleotides removals with Trim Galore (version 0.6.7). The minimal length of reads for alignment is required at least 30 bp. Alignments were performed according to the best practices of the Genome Analysis Toolkit (GATK version: 4.2.5.0)[16]. In general, trimmed reads were aligned with BWA (mem mode; version 0.7.17)[17]. PCR duplications and base recalibration were performed in the alignment files using GATK. Mutect2 (version 4.2.5.0)[18] was applied to called variants in known CHIP associated genes obtained from[8]. Candidates were kept if they passed Mutect2 default filtering, reads depth ≥20x, number of reads supporting variant alleles ≥3, variant allele fraction (VAF) ⩾ 2%, gnomAD allele frequency <0.001 and at least one read in both forward and reverse direction supporting the reference and variant alleles.

Statistical analyses

To compare the distribution of age between PD and healthy controls, two-tailed student's t-test was used, respectively. A generalized linear regression model was applied to examine the effect of age or PD status, with Chi-square tests to assess the statistical significance. Fisher's exact test was used to compare the percentage of CHIP carriers between younger and older groups and between healthy control and PD patients.

References

1. Poewe, W. *et al.* Parkinson disease. *Nat Rev Dis Primers* **3**, 17013 (2017).
2. Bose, A. & Beal, M.F. Mitochondrial dysfunction in Parkinson's disease. *J Neurochem* **139 Suppl 1**, 216-231 (2016).
3. Lobon, I. *et al.* Somatic Mutations Detected in Parkinson Disease Could Affect Genes With a Role in Synaptic and Neuronal Processes. *Front Aging* **3**, 851039 (2022).
4. Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* **371**, 2488-98 (2014).
5. Jaiswal, S. Clonal hematopoiesis and nonhematologic disorders. *Blood* **136**, 1606-1614 (2020).
6. Cochran, J.N. *et al.* Non-coding and Loss-of-Function Coding Variants in TET2 are Associated with Multiple Neurodegenerative Diseases. *Am J Hum Genet* **106**, 632-645 (2020).
7. Parkinson Progression Marker, I. The Parkinson Progression Marker Initiative (PPMI). *Prog Neurobiol* **95**, 629-35 (2011).
8. Bick, A.G. *et al.* Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763-768 (2020).
9. Niroula, A. *et al.* Distinction of lymphoid and myeloid clonal hematopoiesis. *Nat Med* **27**, 1921-1927 (2021).
10. Driver, J.A., Logroscino, G., Gaziano, J.M. & Kurth, T. Incidence and remaining lifetime risk of Parkinson disease in advanced age. *Neurology* **72**, 432-8 (2009).
11. Oh, T.J., Song, H., Koh, Y. & Choi, S.H. The Presence of Clonal Hematopoiesis Is Negatively Associated with Diabetic Peripheral Neuropathy in Type 2 Diabetes. *Endocrinol Metab (Seoul)* **37**, 243-248 (2022).
12. Bouzid, H. *et al.* Clonal hematopoiesis is associated with protection from Alzheimer's disease. *medRxiv*, 2021.12.10.21267552 (2021).
13. Feng, D.D., Cai, W. & Chen, X. The associations between Parkinson's disease and cancer: the plot thickens. *Transl Neurodegener* **4**, 20 (2015).
14. Marshall, L.L. *et al.* Epigenomic analysis of Parkinson's disease neurons identifies Tet2 loss as neuroprotective. *Nat Neurosci* **23**, 1203-1214 (2020).
15. Kern, F. *et al.* Deep sequencing of sncRNAs reveals hallmarks and regulatory modules of the transcriptome during Parkinson's disease progression. *Nature Aging* **1**, 309-322 (2021).
16. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-303 (2010).
17. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-60 (2009).
18. Benjamin, D. *et al.* Calling Somatic SNVs and Indels with Mutect2. *bioRxiv*, 861054 (2019).

Acknowledgments: We thank all members of the Jan Vijg's laboratory for helpful discussions related to this project. We thank the staffs in the PPMI for assisting the data access.

Funding: This study was supported by The Michael J. Fox Foundation (J.V., P.G.M.); NIH grants P01 AG017242, P01 AG047200, P30 AG038072 and U01 ES029519, the Glenn Foundation for Medical Research (J.V.).

Author contributions: J.V. and P.G.M. conceived and supervised the study. S.S. analyzed the data. S.S., D.S., C.P.G., Z.H., J.H.H, A.Y.M, P.G.M., and J.V. wrote the manuscript.

Competing interests: A.Y.M and J.V. are the co-founders of SingulOmics, Corp.Other authors declare no competing interests.

Data and materials availability: All raw sequencing data were obtained from https://www.ppmi-info.org. Additional data and script related to this paper may be requested from the authors.

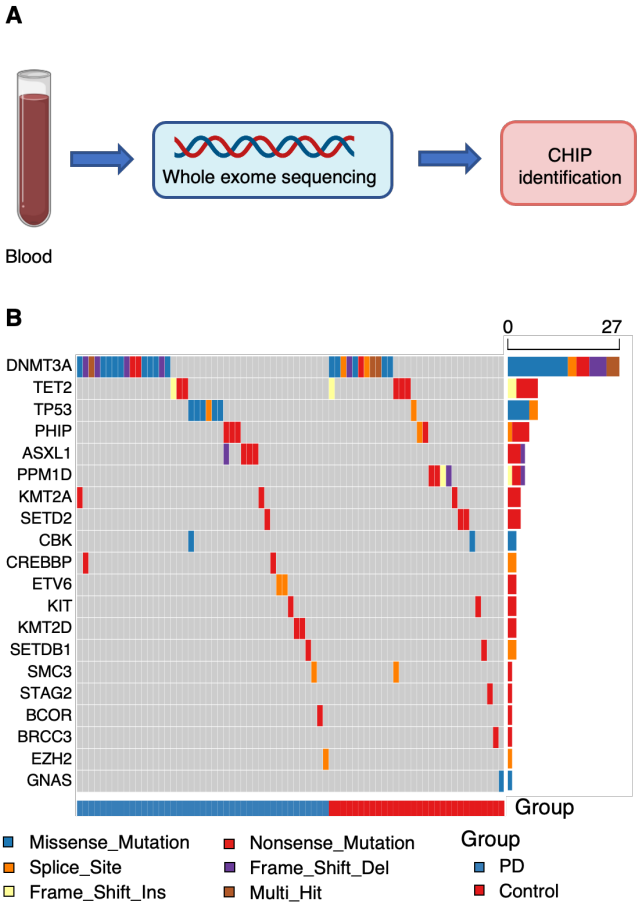

Fig. 1. Study overview and CHIP variants identified in WES data. (A) Workflow of sequencing and data processing. Sequencing data was obtained from PPMI. This figure is created using BioRender. (B) Co-mutation plot with mutations across the 78 subjects from healthy controls and PD patients.

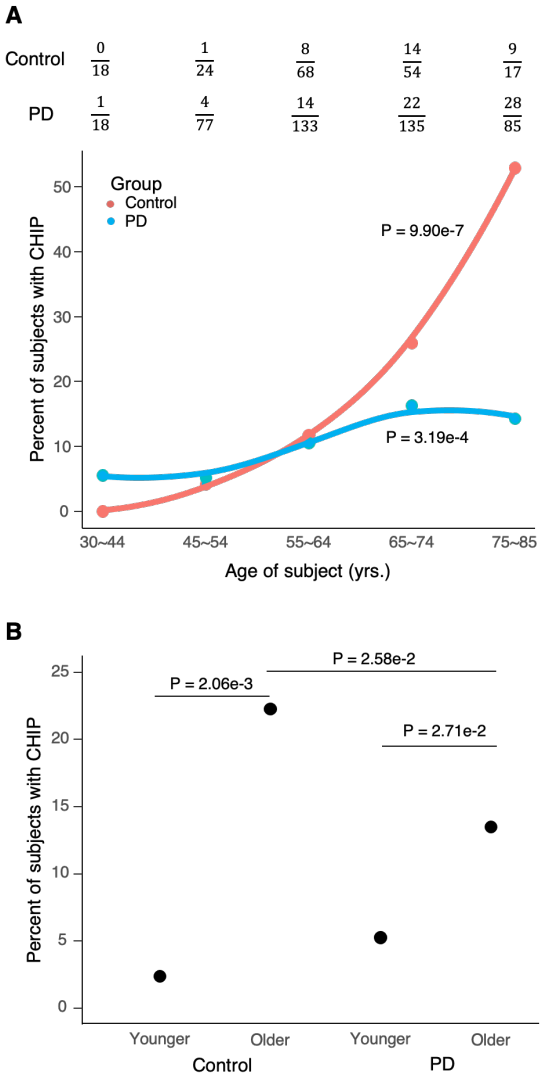

Fig. 2. The percentage of CHIP carriers. (A) Distribution of percentages of CHIP carriers across different age bins. The numbers in the top panel indicated the numbers of CHIP carriers and total subjects in the age bin for healthy controls and PD patients, separately. (B) Percentages of CHIP carriers from younger (≤55 years old) and older (>55 years old) subjects using WES data, in healthy controls and in PD patients.

Supplementary Table 1. List of PPMI samples analyzed in this study.

Supplementary Table 2. List of CHIP variants detected in PPMI cohort.