

Figure 3. Flowchart of literature search, data extraction, and analysis strategies applied for PDGene.
doi:10.1371/journal.pgen.1002548.g003

GWAS-only meta-analyses. We obtained individual-level genotype data for all publicly available PD GWAS datasets from NCBI's "dbGAP" database (a total of three [10,12,13] at the time of the datafreeze, March 31st, 2011). Genotype data were cleaned using standard procedures, followed by imputation of untested genotypes (using reference panels from HapMap and the 1000 Genomes Project), and association analyses incorporating imputation uncertainty (case-control datasets only), age, sex, and population stratification. Overall, this procedure led to a total of 7,723,931 unique SNPs, 7,123,920 of which were present in at least two, and 711,271 in at least three datasets. Meta-analyses (either combining test-statistics and standard errors using random-effects models, or by combining P-values weighted by sample size, see Text S1 for more details) were performed on the 7,123,920 SNPs present in at least two of the GWAS datasets.

Online database

After completion of all data-management and analysis steps, all study-specific variables, genotype data (except for GWAS), and meta-analysis plots are posted on a dedicated, publicly available, online adaptation of the PDGene database using the same software and code as our databases for Alzheimer's disease [33] and schizophrenia [34]. The online database is hosted by the "Alzheimer Research Forum" and can be accessed via its own designated URL (<http://www.pdgene.org>).

Database code

The database software can easily be ported to other genetically complex diseases and will be made available on a collaborative basis to interested researchers upon request.

Supporting Information

Figure S1 QQ plots showing the distribution of expected versus observed P-values for the GWAS-only meta-analysis results. Analyses were performed using the METAL software (ref. [21] in Text S1). The excess of observed P-values (Figure S1, panel 1) is entirely due to association signals in the *SNCA*, *MAPT*, *LRRK2*, and *DGKQ/GAK* loci as can be seen in Figure S1, panel 2 that showcases the P-value distributions after removal of 18,622 SNPs in these regions ($\lambda = 1.007$). (TIF)

Figure S2 Forest plots of allelic meta-analyses for SNPs showing genome-wide significant association ($P < 5 \times 10^{-8}$) with PD susceptibility in the March 31st 2011 datafreeze. Study-specific allelic odds ratios (ORs, black squares) and 95% confidence intervals (CIs, lines) were calculated for each included dataset. The summary OR and CI was calculated using random-effects models (grey diamond). Whenever multiple polymorphisms showed genome-wide significant association in the same locus, only the variant with the smallest P-value is listed here for meta-

analysis results after stratification for Caucasian and Asian ancestries. For a complete list of meta-analyses performed for the datafreeze, see Table S1. Figure S1, panel 1-S1, panel 12 and S1, panel 13-S1, panel 16 display the SNP showing the most significant genome-wide association in datasets of Caucasian ancestry and Asian ancestry, respectively. Details and references of all included studies displayed here can be found on the PDGene database (<http://www.pdgene.org>). I₂ = estimate of percentage of between-study heterogeneity that is beyond chance, “excl initial” = summary OR and 95%CI after meta-analysis after exclusion of the initial study, C = Caucasian ancestry, A = Asian ancestry, H = Hispanic descent, D = African descent, “•” = initial study (applies to candidate-gene studies), “†” = no data provided or data was not eligible for inclusion in meta-analysis, “‡” = study excluded due to overlap, “#” = HWE violation in controls ($P < 0.05$, not applicable to quality-controlled GWAS datasets, see Text S1), “i” = SNP monomorphic in the respective dataset, “ø” = meta-analysis after excluding initial study not applicable.

(PDF)

Figure S3 Locus plot of the *ITGA8* region on chromosome 10p13 (15346353–15801533 bp, hg18). The figure displays association results for ~1,400 SNPs in the *ITGA8* region including at least four independent datasets. SNPs are color-coded based on linkage disequilibrium (r^2) estimates from the CEU 1000G dataset (release June 2010). All LD estimates refer to the most significantly associated SNP rs7077361. SNPs color-coded in grey indicate missing LD estimates in the CEU dataset. Recombination rates were estimated based on the CEU dataset, and are displayed as blue line in the background. Gene annotations are based on RefSeq and the UCSC Genome browser. Locus plots were generated using the LocusZoom Stand-alone package (http://genome.sph.umich.edu/wiki/LocusZoom_Standalone).

(TIF)

Figure S4 Forest plots of fixed-effect meta-analyses for SNP rs6723108 in the *ACMSD/TMEM163* locus and chr6:32609909 in the HLA locus. Symbols are the same as for Figure S2 (see above).

(TIF)

Table S1 Overview of all 867 polymorphisms meta-analyzed in the March 31st 2011 datafreeze using random-effects allelic models. Random-effects allelic meta-analyses were performed on polymorphisms for which four or more independent datasets were available. Meta-analyses after stratification for different ethnic descent were performed if at least three independent datasets were available in the respective stratum (applicable only to samples of European and Asian descent). Each nominally significant meta-analysis result ($P < 0.05$) was graded according to the HuGENet interim criteria. For details on how these criteria are applied, see Text S1. Meta-analysis results in this table are ordered by genomic location. OR = Odds Ratio, CI = confidence interval, N minor = number of minor alleles, Ethnicities: C = Caucasian, A = Asian, D = African Descent, H = Hispanic, O = Other/Mixed, Low OR = $OR < 1.15$ or ≥ 0.87 , respectively, F = loss of significance in the respective meta-analysis after exclusion of the first study, HWE = loss of significance after excluding studies violating HWE ($P < 0.05$), Regr = evidence for small-study/publication bias using a modified regression test (see Text S1), A = Grade A (‘strong’ epidemiologic credibility), B = Grade B (‘modest’ epidemiologic credibility), C = Grade C (‘weak’ epidemiologic credibility), logBF = Bayes Factor (see Text S1). “*” denotes SNPs that have been supplemented by additional data after the datafreeze (in total

this applies to 147 SNPs, see Text S1 for the description of included datasets).

(XLS)

Table S2 Investigation of the extent of statistical inflation assuming sample overlaps of 1%, 5%, and 10% across cases and controls in datasets originating from the same countries. Hypothetical sample overlap across datasets was assumed between different candidate-gene/replication studies and between candidate-gene/replication studies and GWAS datasets if they originated from the same country. These analyses were performed applying random-effects models and adding the sum of weighted co-variances of overlapping datasets to the overall study variance (see ref. [24] in the main text). Note that the assumption of undetected overlapping samples does not apply (and was therefore not modeled here) to overlap between individual GWAS as duplicate samples in these datasets were removed prior to meta-analysis. It also does not apply to independent datasets used in the same publication where duplicate samples had been removed by the authors prior to analysis and publication. We emphasize that this table describes hypothetical scenarios, because the geographical origin of each study had been investigated extensively and potentially overlapping datasets had been excluded as part of PDGene’s data inclusion protocol. Thus, the extent of overlap across geographically distinct datasets within the same countries is reduced to accidental recruitment of the same subjects more than once in different datasets throughout the respective countries, and can be expected to be less than ~1%. This estimate is based on data of the GEO-PD consortium, where sufficient data were centrally available of 6,072 subjects from 20 geographically distinct sites in 13 countries that had been investigated for potentially duplicate samples across sites, but no duplicate subjects (neither between not within countries) were identified when matching on ethnicity, birth, sex, and genotype. The investigation of overlap was not applicable here for Asian datasets, as they originated from different countries and/or were cleaned by the respective authors prior to publication.

(DOC)

Text S1 Supplementary material. This file includes supplementary methods and references as well as the list of members of the GWAS consortia, the GEO-PD Consortium, and consortia-specific acknowledgements.

(PDF)

Acknowledgments

23andMe acknowledges Elizabeth Dorfman, Amy K. Kiefer, Emily M. Drabant, Uta Francke, Joanna L. Mountain, David Hinds, and Anne Wojcicki from 23andMe, as well as Samuel M. Goldman, Caroline M. Tanner, and J. William Langston from the Parkinson’s Institute, Sunnyvale, CA, USA. We also acknowledge the contribution of Mitsutoshi Yamamoto, Nobutaka Hattori, and Miho Murata for sample collection in the Japanese GWAS 1.0 [14]. We are grateful to the Alzheimer Research Forum—in particular to June Kinoshita, Colin Knep, Paula Noyes, and Gabrielle Ströbel—for hosting PDGene on their website. We also thank the many PD researchers who have kindly provided us with genotype data and helpful information beyond those included in the original publications. Finally, we would like to thank the many PD patients and control subjects who volunteered to participate in the individual studies.

Author Contributions

Conceived and designed the experiments: CM Lill, MB McQueen, JPA Ioannidis, L Bertram. Performed the experiments: CM Lill, JT Roehr, S Bagade, B-M Schjeide, E Meissner, U Zauft, NC Allen, KJ Anderson, G Beecham, D Berg, JM Biernacka, A Brice, AL DeStefano, CB Do, N

Eriksson, SA Factor, MJ Farrer, T Foroud, T Gasser, T Hamzà, JA Hardy, P Heutink, C Klein, JC Latourelle, DM Maraganore, ER Martin, M Martinez, RH Myers, H Payami, WK Scott, M Sharma, AB Singleton, K Stefansson, T Toda, JY Tung, J Vance, NW Wood, CP Zabetian, 23andMe, GEO-PD, IPDGC, Parkinson's Disease GWAS, WTCC2. Analyzed the data: CM Lill, JT Roehr, MB McQueen, FK Kavvoura, L Bertram. Wrote the paper: CM Lill, JPA Ioannidis, L Bertram. Helped write the manuscript: E Meissner, MJ Farrer, T Foroud, T Gasser, C

Klein, DM Maraganore, H Payami, AB Singleton, M Sharma, F Zipp, H Lehrach. Helped analyze the data: S Bagade, T Liu, M Schilling, CB Do, N Eriksson, T Hamza, EM Hill-Burns, MA Nalls, N Pankratz, W Satake, M Sharma. Interpretation of results: CM Lill, JPA Ioannidis, L Bertram. Study coordination: CM Lill, T Foroud, JA Hardy, H Payami, AB Singleton, P Young, RE Tanzi, MJ Khoury, F Zipp, H Lehrach, JPA Ioannidis, L Bertram. Literature searches and data entry: CM Lill, S Bagade, B-M Schjerve, E Meissner, U Zauft, N Allen.

References

- de Lau LML, Breteler MMB (2006) Epidemiology of Parkinson's disease. *Lancet Neurol* 5: 525–535. doi:10.1016/S1474-4422(06)70471-9.
- Hardy J, Lewis P, Revesz T, Lees A, Paisan-Ruiz C (2009) The genetics of Parkinson's syndromes: a critical review. *Curr Opin Genet Dev* 19: 254–265. doi:10.1016/j.gde.2009.03.008.
- Vilariño-Güell C, Wider C, Ross OA, Dachselt JC, Kachergus JM, et al. (2011) VPS35 mutations in Parkinson disease. *Am J Hum Genet* 89: 162–167. doi:10.1016/j.ajhg.2011.06.001.
- Zimprich A, Benet-Pagés A, Struhal W, Graf E, Eck SH, et al. (2011) A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *Am J Hum Genet* 89: 168–175. doi:10.1016/j.ajhg.2011.06.008.
- Chartier-Harlin M-C, Dachselt JC, Vilariño-Güell C, Lincoln SJ, Leprêtre F, et al. (2011) Translation initiator EIF4G1 mutations in familial Parkinson disease. *Am J Hum Genet* 89: 398–406. doi:10.1016/j.ajhg.2011.08.009.
- Maraganore DM, de Andrade M, Elbaz A, Farrer MJ, Ioannidis JP, et al. (2006) Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. *JAMA* 296: 661–670. doi:10.1001/jama.296.6.661.
- Zabetian CP, Yamamoto M, Lopez AN, Ujike H, Mata IF, et al. (2009) LRRK2 mutations and risk variants in Japanese patients with Parkinson's disease. *Mov Disord* 24: 1034–1041. doi:10.1002/mds.22514.
- Goris A, Williams-Gray CH, Clark GR, Foltynie T, Lewis SJG, et al. (2007) Tau and alpha-synuclein in susceptibility to, and dementia in, Parkinson's disease. *Ann Neurol* 62: 145–153. doi:10.1002/ana.21192.
- Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, et al. (2009) Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 361: 1651–1661. doi:10.1056/NEJMoa0901281.
- Maraganore DM, de Andrade M, Lesnick TG, Strain KJ, Farrer MJ, et al. (2005) High-resolution whole-genome association study of Parkinson disease. *Am J Hum Genet* 77: 685–693. doi:10.1086/496902.
- Fung H-C, Scholz S, Matarin M, Simón-Sánchez J, Hernandez D, et al. (2006) Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol* 5: 911–916. doi:10.1016/S1474-4422(06)70578-6.
- Pankratz N, Wilk JB, Latourelle JC, DeStefano AL, Halter C, et al. (2009) Genome-wide association study for susceptibility genes contributing to familial Parkinson disease. *Hum Genet* 124: 593–605. doi:10.1007/s00439-008-0582-9.
- Simón-Sánchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, et al. (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet* 41: 1308–1312. doi:10.1038/ng.487.
- Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, et al. (2009) Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet* 41: 1303–1307. doi:10.1038/ng.485.
- Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, et al. (2010) Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet* 74: 97–109. doi:10.1111/j.1469-1809.2009.00560.x.
- Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, et al. (2010) Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat Genet* 42: 781–785. doi:10.1038/ng.642.
- Spencer CCA, Plagnol V, Strange A, Gardner M, Paisan-Ruiz C, et al. (2011) Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum Mol Genet* 20: 345–353. doi:10.1093/hmg/ddq469.
- Saad M, Lesage S, Saint-Pierre A, Corvol J-C, Zelenika D, et al. (2011) Genome-wide association study confirms BST1 and suggests a locus on 12q24 as the risk loci for Parkinson's disease in the European population. *Hum Mol Genet* 20: 615–627. doi:10.1093/hmg/ddq497.
- Simón-Sánchez J, van Hilten JJ, van de Warrenburg B, Post B, Berendse HW, et al. (2011) Genome-wide association study confirms extant PD risk loci among the Dutch. *Eur J Hum Genet* 19: 655–661. doi:10.1038/ejhg.2010.254.
- Evangelou E, Maraganore DM, Ioannidis JPA (2007) Meta-analysis in genome-wide association datasets: strategies and application in Parkinson disease. *PLoS ONE* 2: e196. doi:10.1371/journal.pone.0000196.
- Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin U-M, et al. (2011) Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* 377: 641–649. doi:10.1016/S0140-6736(10)62345-8.
- Do CB, Tung JY, Dorfman E, Kiefer AK, Drabant EM, et al. (2011) Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet* 7: e1002141. doi:10.1371/journal.pgen.1002141.
- Sharma M, Ioannidis JPA, Aasly JO, Annesi G, Brice A, et al. (n.d.) Large-scale replication and heterogeneity in Parkinson disease genetic loci. *Neurology* in press.
- Lin D-Y, Sullivan PF (2009) Meta-analysis of genome-wide association studies with overlapping subjects. *Am J Hum Genet* 85: 862–872. doi:10.1016/j.ajhg.2009.11.001.
- Ioannidis JPA (2008) Effect of formal statistical significance on the credibility of observational associations. *Am J Epidemiol* 168: 374–383; discussion 384–390. doi:10.1093/aje/kwn156.
- Ioannidis JPA, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, et al. (2008) Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 37: 120–132. doi:10.1093/ije/dym159.
- Khoury MJ, Bertram L, Boffetta P, Butterworth AS, Chanock SJ, et al. (2009) Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. *Am J Epidemiol* 170: 269–279. doi:10.1093/aje/kwp119.
- Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D (2010) BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26: 2204–2207. doi:10.1093/bioinformatics/btq351.
- Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, et al. (2007) A survey of genetic human cortical gene expression. *Nat Genet* 39: 1494–1499. doi:10.1038/ng.2007.16.
- Varnum-Finney B, Venstrom K, Muller U, Kypta R, Backus C, et al. (1995) The integrin receptor alpha 8 beta 1 mediates interactions of embryonic chick motor and sensory neurons with tenascin-C. *Neuron* 14: 1213–1222.
- DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* 7: 177–188.
- Harbord RM, Egger M, Sterne JAC (2006) A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med* 25: 3443–3457. doi:10.1002/sim.2380.
- Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE (2007) Systematic meta-analysis of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 39: 17–23. doi:10.1038/ng1934.
- Allen NC, Bagade S, McQueen MB, Ioannidis JPA, Kavvoura FK, et al. (2008) Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 40: 827–834. doi:10.1038/ng.171.

国際共同研究における Genome-Wide Association Study (GWAS)

Genome-wide association study (GWAS) by international collaboration

戸田 達史

Key Words : GWAS, メタ解析, SNP, HapMap計画

■ Abstract ■

dbSNPやHapMap計画などの情報基盤や数十万種のSNPを解析できる技術基盤の整備により、ゲノムワイド関連解析 (GWAS) が2007年頃より実用的な戦略となり、疾患感受性遺伝子の発見が相次いだ。すでに165の疾患について、ゲノムワイド有意水準 $P < 5 \times 10^{-8}$ をクリアする904個の論文が発表された。国際共同研究によりGWASメタ解析を行い、よりeffect sizeの小さなものも同定するという試みが行われだしており、「第2世代のGWAS」とも言われている。

■ GWASの現状

2006年以降、common diseaseの感受性遺伝子の探索研究は新しい段階を迎えた。これをもたらしたおもな要因は二つの基盤整備である。まず情報基盤として、dbSNPやHapMap計画で代表されるように、ヒトゲノム全域にわたる膨大な多様性情報が集積されてきた。次に技術基盤として、数十万種のSNP (Single Nucleotide Polymorphism) を数千もの個体について並列解析できるプラットフォームが市販化された。

まずHapMap計画によると、日本人と白人は約25-30万個のタグSNP (これを調べれば連鎖不平衡で結ばれた近傍の多くのSNPの代表になる) で、ほぼ全ゲノムの遺伝子がカバーされる。つまり約30万個のタグSNPを患者と対照で調べれば、ほぼ全ゲノムの遺伝子を調べたことになる。そこで具体的には例えば患者1,000人、対照1,000人、計2,000人各人の50万個のSNPの遺伝子型を決定する。すなわちSNPチップとして2,000枚の実験を行う。それぞれSNP-1, ..., SNP-500000ずつ、患者、対照に

おけるそれぞれのアレルの出現頻度を合計し、偏りがないかどうかの検定を行うのである。

これらを活用することによって、ゲノムワイド関連解析 (GWAS: Genome-Wide Association Study) が実用的な戦略となり、2007年にはNature, Science誌などに立て続けに成果が発表されることとなった。その後のGWASによる疾患感受性遺伝子の発見ラッシュには目を見張るものがあり、すでに165の疾患について、ゲノムワイド有意水準 $P < 5 \times 10^{-8}$ をクリアする904個の論文が発表されている (図)。

■ GWASによる疾患感受性遺伝子発見

神経疾患のパーキンソン病を例にとると、我々のグループは患者の95%を占める孤発性PDのリスク遺伝子を同定するため、GWASを行い、PD発症に関わる2つの新しい遺伝子座*PARK16*, *BST1*を同定した。また、常染色体優性遺伝性PDの原因遺伝子 *α -synuclein*, *LRRK2*の孤発性PDへの関与を証明した。国際共同研究として次に、我々と併行してヨーロッパ起源の集団のPDのGWAS研究を行っていたグループと、データを交換した。彼らの研究では、 *α -synuclein*領域、*Tau*領域のみに、強い関連が検出されていた。そこで、我々の発見した*PARK16*・*BST1*・*LRRK2*の人数を増やして再現研究をしたところ、*PARK16*と*LRRK2*の関連は強く再現されたが、*BST1*は再現されなかった。逆に、我々は、彼らの検出した*Tau*の関連の再現を試みたが、我々の検体セットでは、再現されなかった。よって、 *α -synuclein*・*PARK16*・*LRRK2*は2人種に共通のPDリスクであり、*Tau*・*BST1*はのリスク多型の影響は人種特異的であると考えた¹⁾。

Tatsushi Toda

神戸大学大学院医学研究科神経内科学/分子脳科学
Division of Neurology/Molecular Brain Science, Kobe
University Graduate School of Medicine

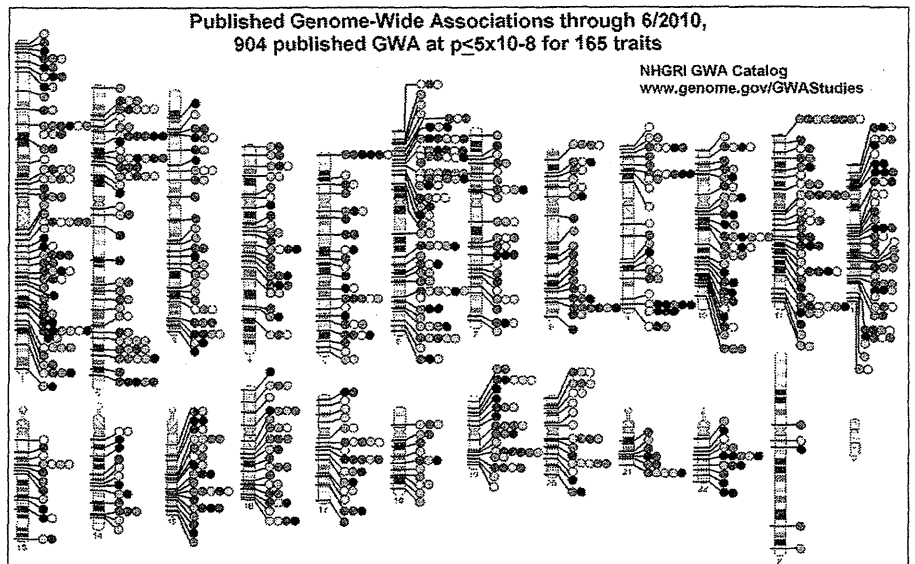
さらにパーキンソン病においては、さらなる国際共同研究として欧米の5つのグループがそれぞれ独立に行っていたGWASを合わせてメタ解析を行い（計患者5,333, 対照12,019）、有意なSNPをさらに患者7,053, 対照9,007で再現実験を行い、ゲノムワイド有意水準 $P < 5 \times 10^{-8}$ を超える遺伝子を従来の6個の他に、*ACMSD*, *STK39*, *LAMP3*, *SYT11*, *CCDC62*の5個を同定した²⁾。

すべては網羅できないが他の例では、クローン病においては、6個のGWASを合わせてメタ解析を行い（計患者6,333, 対照15,056）、さらに有意なSNPを患者15,694, 対照14,026で再現実験を行い、ゲノムワイド有意水準を超える遺伝子を新規に30個同定し、計71個となった³⁾。

また2型糖尿病においては、8個のGWASを合わせてメタ解析を行い（計患者8,130, 対照38,987）、さらに有意なSNPを患者34,412, 対照59,925で再現実験を行い、ゲノムワイド有意水準を超える遺伝子を新規に12個同定した⁴⁾。

さらに驚くべきことに、GIANT Consortiumは国際共同研究として46個のGWASを合わせてメタ解析を行い（計133,653人）、50,074人で再現実験を行い、ゲノムワイド有意水準を超える身長に関わる遺伝子を少なくとも180個同定した⁵⁾。

GWASによって多数の疾患感受性遺伝子が同定されたものの、それらは遺伝要因全体の一部しか説明できないことから（missing heritability）、このような圧倒的な数の試料を各地から集めてゲノムワイドメタ解析を行いよりeffect sizeの小さなものも同定するという試みが行われだしており、「第2世代のGWAS」とも言われている。



<ul style="list-style-type: none"> ○ Acute lymphocytic leukemia ● Adhesion molecules ● Adiponectin levels ● Age-related macular degeneration ○ AIDS progression ○ Alcohol dependence ○ Alzheimer disease ○ Amyotrophic lateral sclerosis ○ Angiotensin-converting enzyme activity ● Antihypertensive therapy ● Arterial stiffness ● Asthma ● Atherosclerosis in IHTV ● Atrial fibrillation ● Mitral aortic hyperactivity disorder ○ Autism ● Basal cell cancer ● Bipolar disorder ● Bilirubin ● Blood pressure ● Blue or green eyes ● BMI waist circumference ○ Bone density ● Breast cancer ● C-reactive protein ● Cardiac structure/function ● Carotid intima ● Catecholamine levels ○ Celiac disease ● Chronic lymphocytic leukemia 	<ul style="list-style-type: none"> ○ Cleft lip/palate ○ Cognitive function ○ Conduct disorder ○ Colorectal cancer ○ Corneal thickness ○ Coronary disease ● Creutzfeldt-Jakob disease ● Crohn's disease ● Cutaneous melanoma ○ Dermatitis ● Drug-induced liver injury ● Eosinophil count ● Esophageal esophagitis ● Erythrocyte parameters ● Esophageal cancer ● Essential tremor ○ Exfoliation glaucoma ● Eye color traits ● F cell distribution ● Fibrinogen levels ● Folate pathway vitamins ○ Fracture and burning ○ Galactose ● Glioma ● Glycemic traits ● Hair color ● Hair morphology ○ HDL cholesterol ● Heart failure ● Heart rate ● Height ● Hemostasis parameters ● Hepatitis ● Hirschsprung's disease ○ HIV-1 control 	<ul style="list-style-type: none"> ○ Homocysteine levels ● Idiopathic pulmonary fibrosis ● IgE levels ● Inflammatory bowel disease ● Intracranial aneurysm ● Iris color ● Iron status markers ● Ischemic stroke ○ Juvenile idiopathic arthritis ● Kidney stones ● LDL cholesterol ● Leprosy ● Leptin receptor levels ● Liver enzymes ● LP (a) levels ● LpPLA2 activity and mass ● Lung cancer ● Major mood disorders ● Malacia ● Male pattern baldness ● Male melanocytosis levels ● MCP-1 ● Melanoma ● Menarche & menopause ● Multiple sclerosis ○ Myeloproliferative neoplasms ○ Narcolepsy ○ Nasopharyngeal cancer ○ Neuroblastoma ● Nicotine dependence ● Obesity ● Open angle glaucoma ● Open personality ● Optic disc parameters ● Osteoarthritis 	<ul style="list-style-type: none"> ○ Osteoporosis ● Osteosclerosis ● Other metabolic traits ● Ovarian cancer ● Pancreatic cancer ● Pain ● Paget's disease ● Panic disorder ● Parkinson's disease ○ Psoriasis ● Psoriasis ● Peripheral arterial disease ● Phosphatidylcholine levels ● Phytosterol levels ○ Platelet count ○ Primary biliary cirrhosis ● PR interval ● Prostate cancer ● Protein levels ○ Psoriasis ● Pulmonary function COFD ● QT interval ● Quercetin levels ● Resonance rate ● Red vs non-red hair ● Renal function ● Response to antidepressants ● Response to antipsychotic therapy ● Response to hepatitis C treat ● Response to statin therapy ● Restless leg syndrome ● Rheumatoid arthritis ● Schizophrenia 	<ul style="list-style-type: none"> ● Serum metabolites ● Skin pigmentation ● Smoking behavior ● Speech perception ○ Sphingolipid levels ● Statin-induced myopathy ● Stroke ● Systemic lupus erythematosus ○ Systemic sclerosis ○ Testosterone length ○ Testicular germ cell tumor ● Thyroid cancer ● Tooth development ● Total cholesterol ● Trichotillomania ● Type 1 diabetes ● Type 2 diabetes ● Uterovaginal calcitis ● Urate ● Venous thromboembolism ● Vertical cup-disc ratio ● Vitamin B12 levels ● Vitamin D insufficiency ● Vitiligo ● Warfarin dose ● Weight ○ White cell count ● YKL-40 levels
--	---	--	--	--

図 ゲノムワイド有意水準 $P < 5 \times 10^{-8}$ をクリアする多因子疾患の遺伝子座

一方でCommon Disease-Multiple Rare Variant仮説による、頻度は低いが発症へのEffect sizeが大きいrare variantも、missing heritabilityとして重要な位置を占める。

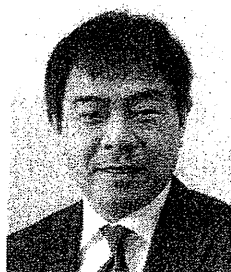
文献

- 1) Satake W, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. Nat Genet. 2009; 41: 1303-7.
- 2) International Parkinson Disease Genomics Consortium. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet. 2011; 377: 641-9.
- 3) Franke A, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010; 42: 1118-25.
- 4) Voight BF, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 2010; 42: 579-89.
- 5) Lango Allen H, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010; 467: 832-8.

6 パーキンソン病の分子遺伝学—ゲノム関連解析研究—

とだ たつし さたけわたる
戸田 達史・佐竹 渉

神戸大学大学院医学研究科 神経内科学



戸田 達史
 1985年東京大学医学部卒業，神経内科入局。94年東京大学人類遺伝学助手。96年東京大学医科学研究所助教授。2000年大阪大学臨床遺伝学教授。2008年神戸大学神経内科学／分子脳科学教授。日本人類遺伝学会賞，日本神経学会賞，朝日賞，文部科学大臣表彰。研究テーマは神経内科学，神経疾患の分子遺伝学・治療。趣味はワイン，グルメ。

Key words : パーキンソン病， α シヌクレイン，PARK16，GWAS

Abstract

患者の95%を占める孤発性PDは多因子疾患である。孤発性PDのリスク遺伝子を同定するため，ゲノムワイド関連解析を行い，PD発症に関わる2つの新しい遺伝子座PARK16，BST1を同定した。また，常染色体優性遺伝性PDの原因遺伝子SNCA，LRRK2の孤発性PDへの関与を証明した。さらに，人種差が，PDの遺伝的不均一性に，寄与していることを示唆した。国際共同研究によりGWASメタ解析を行いより多くの感受性遺伝子を同定することが行われだしており，「第2世代のGWAS」とも言われている。さらなる遺伝子の解明が期待される。

はじめに

パーキンソン病 (PD) 症例の90%以上は孤発性発症であるが、5-10%は家族性 (その一部はメンデル遺伝性) に発症する。メンデル遺伝性パーキンソン病 (PD) 家系の連鎖解析などから6つのメンデル遺伝性PD原因遺伝子 (α シヌクレイン、パーキン、LRRK2遺伝子など) が明らかにされた。孤発性PD、メンデル遺伝性PDとも、一部共通の発症メカニズムが存在していると考えられ、それらを切り口にして孤発性PDの病態解明が進んでおり、ミトコンドリア障害、酸化ストレス障害の病態

への関与に加え、新たにユビキチン・プロテアソーム系の機能低下、つまり蛋白分解異常からドパミン細胞死に至る経路の重要性が示されたり。

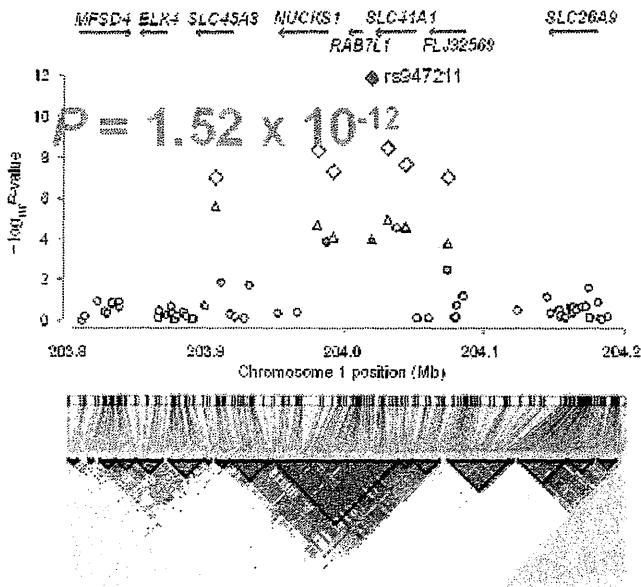
一方dbSNPやHapMap計画などの情報基盤や数十万種のSNPを解析できる技術基盤の整備により、ゲノムワイド関連解析 (GWAS) が2007年頃より実用的な戦略となり、現在多数の疾患でGWAS (ゲノムワイド関連解析) が行われており、疾患感受性遺伝子の発見がいくつかある。我々はパーキンソン病の大規模なGWASを行い、4つの感受性遺伝子を見つけた²⁾。本稿ではパーキンソン病のGWASについて述べる。

1. パーキンソン病は多因子遺伝性疾患

症例的には大多数 (95%) の孤発性PDの原因は、現時点では不明であるが、環境因子と1つ1つは影響力の弱い遺伝因子 (おそらく数10個) によってなり、その総和が、ある閾値を超えたとき発症するという多因子疾患であると考えられている。アイスランド国民を対象とした大規模な疫学的調査の結果が発表され、同胞再発危険率は6.7で、PD発症には遺伝因子が影響していることが示された。

Molecular genetics for Parkinson's disease -genomewide association study- : Tatsushi Toda, Wataru Satake, Division of Neurology, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-chou, Chuo-ku, Kobe 650-0017, Japan

a *PARK16* (1q32)



b *α-synuclein* (4q22)

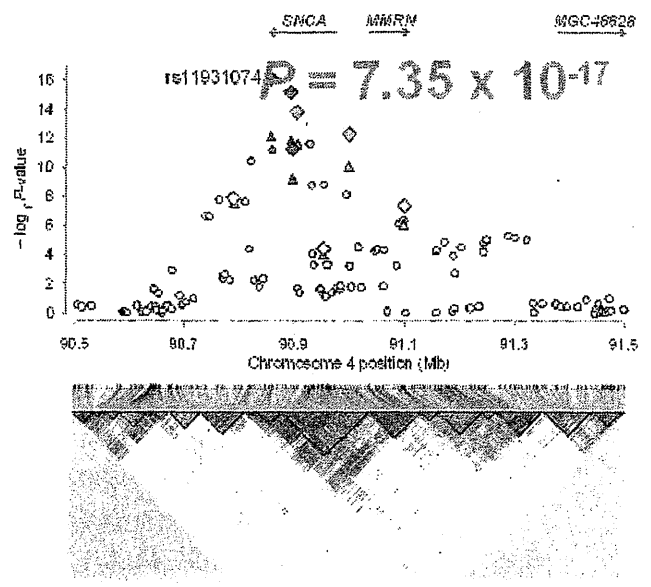


図1 日本人パーキンソン患者2011検体, 対照18381検体のゲノムワイド関連解析より同定された4つのパーキンソン病遺伝子座の抜粋²⁾

PDの感受性遺伝子を発見するため、ここ10年間、多くの研究がなされてきたが、アルツハイマー病におけるApoE4多型のような確実に発症リスクを高める遺伝因子はなかなか確認されていなかった。ゲノムワイド有意水準 ($P < 5 \times 10^{-8}$) を満たす確実なものは、 α シヌクレインの3' 非翻訳領域SNP (Gasserら³⁾, 筆者ら⁴⁾が同定) とゴーシェ病遺伝子*GBA*のrare variant⁵⁾⁶⁾の2つの遺伝子のみであった。

2. パーキンソン病の GWAS (ゲノムワイド関連解析)

パーキンソン病に関しても30~50万SNPチップのGWASが数個論文報告されたが、小~中人数の解析であり、これらの論文では確実なことは言えなかった。我々は、大規模の患者対照集団と、56万個のSNPを搭載したイルミナHap550アレイを用いて、GWASと、2つの独立な再現研究を行い、PDの遺伝リスク因

子を明らかにしようとした。患者検体は、11施設より提供され、総数としては、患者2,011検体, 対照18,381検体を用いた。まず、GWASステージとして、患者1,078検体, 対照2,628検体について、それぞれ56万個のSNP型を決定した²⁾。

SNPや検体のジェノタイプ成功率などのよい患者988検体, 対照2,521検体の、435,470 SNPのジェノタイプをもちいて、GWASステージのtrend検定では、 $P < 5 \times 10^{-8}$ のゲノムワイド有意水準を超える遺伝子は α -synucleinのみであったものの、その一方で、Quantile-Quantile解析からは、 α -synuclein以外の領域のSNPにおいても、帰無仮説のもとで予想されるc2統計量分布からの、有意、かつ強力なインフレーションを観察した。このことは、これらインフレートしたSNPの中に、真なる感受性座が存在していること意味している。

そこで、GWASステージで関連を認めたp値

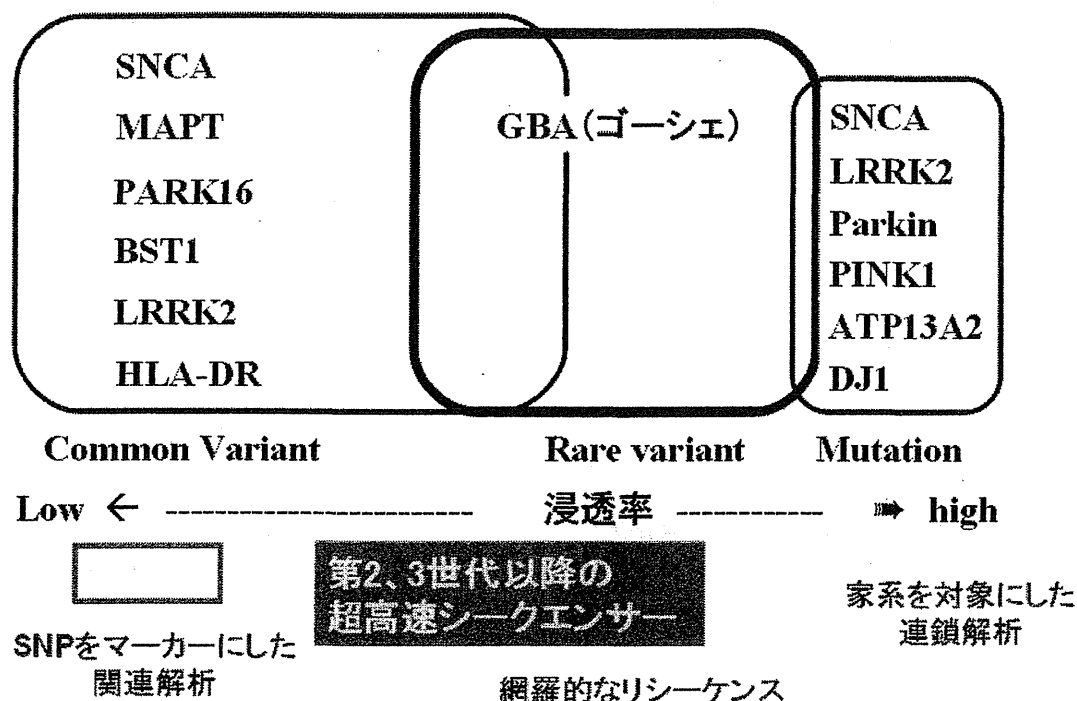


図2 パーキンソン病の遺伝メンデル遺伝性変異以外に、SNPとしてシヌクレイン (SNCA), PARK16, BST1, LRRK2, rare variantとしてゴースェ遺伝子 (GBA) が重要。

上位から337 SNPそれぞれが、2つの独立した患者・対照検体セットでも有意か、を検証する再現研究を行った。さらに全検体 (PD 2,011検体, 対照18,381検体) のジェノタイプデータを用いてメタ解析を行ったところ、絶対的な有意水準 $P < 5 \times 10^{-8}$ をクリアする4つのPD感受性遺伝子座を見いだした²⁾。

まず2つの新しいPD感受性座を、1q32 (PARK16と命名, $P = 1.52 \times 10^{-12}$) と4p15 ($P = 3.94 \times 10^{-9}$) に発見した。1q32領域は、3つの遺伝子 (NUCKS1, RAB7L1, SLC41A1) を含む連鎖不平衡ブロックであるが、発現量的形質座(eQTL)解析から、NUCKS1が、最も有力な責任遺伝子であると考えた (図1a)。また、4p15領域は、BST1のみを含んでいた。さらに、常染色体優性遺伝性PDの原因遺伝子である、 α -synuclein (4q22, $P = 7.35 \times 10^{-17}$, 図1b) とLRRK2 (12q12, $P = 2.72 \times 10^{-8}$) の領域を同定した。 α -synucleinは、我々も以前報告した確

実なPD感受性遺伝子である。また、本研究で初めて、LRRK2領域のありふれた多型が、ゲノムワイド水準でPDと関連することが示された。

次にヨーロッパ起源の集団のPDのGWAS研究を行っていたグループと、データを交換した。彼らの研究では、 α -synuclein領域、Tau領域のみに、強い関連が検出されていた³⁾。そこで、我々の発見したPARK16・BST1・LRRK2の再現研究を依頼したところ、PARK16とLRRK2の関連は強く再現されたが、BST1は再現されなかった。逆に、我々は、彼らの検出したTauの関連の再現を試みたが、我々の検体セットでは、再現されなかった。よって、 α -synuclein・PARK16・LRRK2は2人種に共通のPDリスクであり、Tau・BST1はのリスク多型の影響は人種特異的であると考えた²⁾。

また孤発性PDの発症に、常染色体優性遺伝性PDの原因遺伝子が、密に関係していること

表 日本人集団における各遺伝リスクの大きさ (寄与度)

遺伝リスクマーカー	日本人集団における リスクアレル頻度	オッズ比	人口寄与危険度 (もしこのリスクが、 日本人集団になかったら、 何%の患者が減るか)	
PARK16	rs947211	0.50	1.30	13.0
BST1	rs11931532	0.42	1.24	9.1
SNCA	rs11931074	0.58	1.37	17.6
LRRK2	rs1994090	0.08	1.39	3.0
Tau	rs393152	1.00	1.30	23.0
GBA	RecNcil	0.0018	7.21	1.1 (GBA内の11個の変異を 全てたし合わせると、4.6)

人口リスク寄与度 (もしこのリスクが日本人集団になかったら、何%の患者が減るかの指標) を示す。SNPはオッズ比は低いと人口リスク寄与度は高い。GBAなどのrare variantはその逆である。

を明示した。原因変異により常染色体優性遺伝性PDを引き起こす遺伝子が、そのSNPを通して、孤発性PDの遺伝リスクとなったことは興味深い。このような関係は、PDを超えて、ほかの疾患に対しても、あてはまると思われる⁹⁾。

PARK16領域には、3つの遺伝子が存在するが、発現量的形質座解析から、**NUCKS1**が、最も有力な責任遺伝子であると考えた (図1a)。**NUCKS1**は、リン酸化部位をふくむ核タンパクであるが、神経系における機能は、未知であり、新たなPD発症のパスウェイを開拓する可能性がある⁸⁾。

BST1は、細胞内Ca²⁺貯蔵からのCa²⁺放出を誘発するサイクリックADPリボースの形成を触媒する酵素であり⁹⁾、最近提唱されている、ドパミン細胞死のCa²⁺ストレス説を想起させ、興味深い。これらは従来のPD病態説からは全

く新規な遺伝子であり、従来にとらわれず、新規なものを同定できるところにGWASの強みがある。

さらに最近アメリカの別グループのGWASから、HLA-DR抗原の領域との関連が報告された ($P = 1.9 \times 10^{-10}$)。PD脳ではDR抗原陽性のミクログリアが検出されること、NSAIDsは疫学的にPDのリスクを減少させることなどもあり、PDと炎症の関係を示しており興味深い¹⁰⁾。

3. 第二世代のGWAS

さらにパーキンソン病においては、さらなる国際共同研究として欧米の5つのグループがそれぞれ独立に行っていたGWASを合わせてメタ解析を行い (計患者5333, 対照12019)、有意なSNPをさらに患者7053, 対照9007で再現実験を行い、ゲノムワイド有意水準 $P < 5 \times$

10⁻⁸を超える遺伝子を従来の6個の他に, *ACMSD*, *STK39*, *LAMP3*, *SYT11*, *CCDC62*の5個を同定した¹¹⁾。

GWASによって多数の疾患感受性遺伝子が同定されたものの, それらは遺伝要因全体の一部しか説明できないことから (missing heritability), SNPとしては, このような圧倒的な数の試料を各地から集めてゲノムワイドメタ解析を行いより多くの感受性遺伝子を同定することが行われだしており, 「第2世代のGWAS」とも言われている。

おわりに

おそらく数10個あるパーキンソン病の疾患感受性遺伝子としては, α シヌクレイン, GBA, 今回のPARK16, BST1, LRRK2など以外に確立されたものは少ないのが現状であり, 今後のGWASからさらなる遺伝子の解明, そこから新たな疾患パスウェイとそこからの治療薬開発が期待される (図2)。これら4座位 (1q32, 4p15, 4q22, 12q12) の, 人口リスク寄与度 (もしこのリスクが日本人集団になかったら, 何%の患者が減るかの指標) を示すが, それぞれ13%, 8%, 18%, 3%と見積もられ, SNPはオッズ比は低いが人口リスク寄与度は高い。GBAなどのrare variantはその逆であり, どちらも重要である (表)²⁾。

しかし現在のSNPチップによるGWASでは, アレル頻度の低いものは同定不可能であり (搭載されていない), rare variantは見逃されてしまう。エクソンキャプチャー, 次世代シーケンサーによるエキソームリシーケンシングが行われており, 期待される。

文献

- 1) Farrer MJ : Genetics of Parkinson disease: paradigm shifts and future prospects. *Nat Rev Genet* 7:306-318, 2006.
- 2) Satake W, *et al*: Genome-wide association study identifies

common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet* 41:1303-1307, 2009.

- 3) Mueller JC, *et al*: Multiple regions of alpha-synuclein are associated with Parkinson's disease. *Ann Neurol* 57: 535-541, 2005.
- 4) Mizuta I, *et al*: Multiple candidate gene analysis identifies a-synuclein as a susceptibility gene for sporadic Parkinson's disease. *Hum Mol Genet* 15:1151-1158, 2006.
- 5) Mitsui J, *et al*: Mutations for Gaucher disease confer a high susceptibility to Parkinson disease. *Arch Neurol* 66:571-576, 2009.
- 6) Sidransky E, *et al*: Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 361: 1651-1661, 2009.
- 7) Simon-Sanchez J, *et al*: Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet* 41: 1308-1312, 2009.
- 8) Ostvold AC, *et al*: Molecular cloning of a mammalian nuclear phosphoprotein NUCKS, which serves as a substrate for Cdk1 *in vivo*. *Eur J Biochem* 268: 2430-2440, 2001.
- 9) Yamamoto-Katayama S, *et al*: Crystallographic studies on human BST-1/CD157 with ADP-ribosyl cyclase and NAD glycohydrolase activities. *J Mol Biol* 316: 711-723, 2002.
- 10) Hamza TH, *et al*: Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat Genet* 42:781-785, 2010.
- 11) International Parkinson Disease Genomics Consortium: Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* 377:641-649, 2011.

第19回日本乳癌学会学術総会のご案内

本学会総会は下記日程で開催します。

会期: 2011年9月2日 (金) ~4日 (日)

会場: 仙台国際センター 東北大学百周年記念会館
〒980-0856 仙台市青葉区青葉山無番地
022-265-2211 (代)

会長: 大内憲明 (東北大学病院がんセンター長 東北大学腫瘍外科学分野教授)

テーマ: 「未来のために、今できること-Challenge to the Future」

連絡先: 事務局 東北大学大学院医学系研究科外科病態学講座腫瘍外科学分野
〒980-8574 仙台市青葉区星陵町1番1号
TEL. 022-717-7214 FAX. 022-717-7217
運営準備室

日本コンベンションサービス株式会社 東北支社
〒980-0824 仙台市青葉区支倉町4-34 丸金ビル6階
TEL. 022-722-1311 FAX. 022-722-1178
E-mail: 19jbcsc@convention.co.jp

<Special Article>

孤発性パーキンソン病の分子病態機序はどこまで解明されたか

戸田 達史*

要 旨

- 孤発性パーキンソン病の発症機序としては、ミトコンドリア呼吸系酵素の障害、炎症反応、酸化ストレス障害が一因とされてきたが、遺伝性パーキンソン病家系の解析などから6つのメンデル遺伝性原因遺伝子が明らかにされ、蛋白分解異常の重要性が示された。
- ゲノムワイド関連解析による新たな疾患感受性遺伝子の同定や、ゴーシェ病変異も rare variant として重要である。
- α -シヌクレインのミスフォールド蛋白質が近隣の神経細胞に伝播していくシヌクレインプリオン仮説も提唱されている。
- ミクログリアの活性化を伴う継続的な炎症状態も注目される。

はじめに●

パーキンソン病(PD)は、臨床的には、振戦、筋固縮、寡動、姿勢反射障害を主徴とし、認知症、自律神経障害などのさまざまな随伴症状を呈する神経変性疾患である。わが国には16万人以上の患者が存在するが、今後社会の高齢化に伴いさらなる患者数増加が予想されている。一部は遺伝性であるが、患者の約95%は孤発性であり、本稿では孤発性PDの分子病態機序について、現在主流となっている仮説に加え、最近のトピックスも含めて概観する。

孤発性PDは多因子遺伝性疾患である●

症例的には大多数(90%以上)の孤発性PDの原因は、現時点では不明であるが、加齢および複

数の影響力の弱い遺伝因子(おそらく数十個からなる)と環境要因の組み合わせにより発症する多因子遺伝性疾患と考えられている。Fig.1に示すように各遺伝要因と環境要因の総和が閾値を超えたときに発症すると考えられている(Fig.1)。遺伝要因が関係するということの根拠は、①患者の約10%に家族内発症がみられる、②患者の同胞における有病率の一般集団の有病率に対する比(λ_s)は6.7(アイスランドの報告)、③一卵性双生児での疾患一致率(55%)が二卵性双生児での一致率(18%)の約3倍であった、ということから示唆されている。

一方、PD症例の90%以上は孤発性発症であるが、5~10%は家族性(その一部はメンデル遺伝性)に発症する。メンデル遺伝性PD家系の連鎖解析などから、6つのメンデル遺伝性PD原因遺伝子(α -シヌクレイン、パーキン、*LRRK2* 遺伝子

* T. Toda(教授): 神戸大学大学院医学研究科神経内科学。

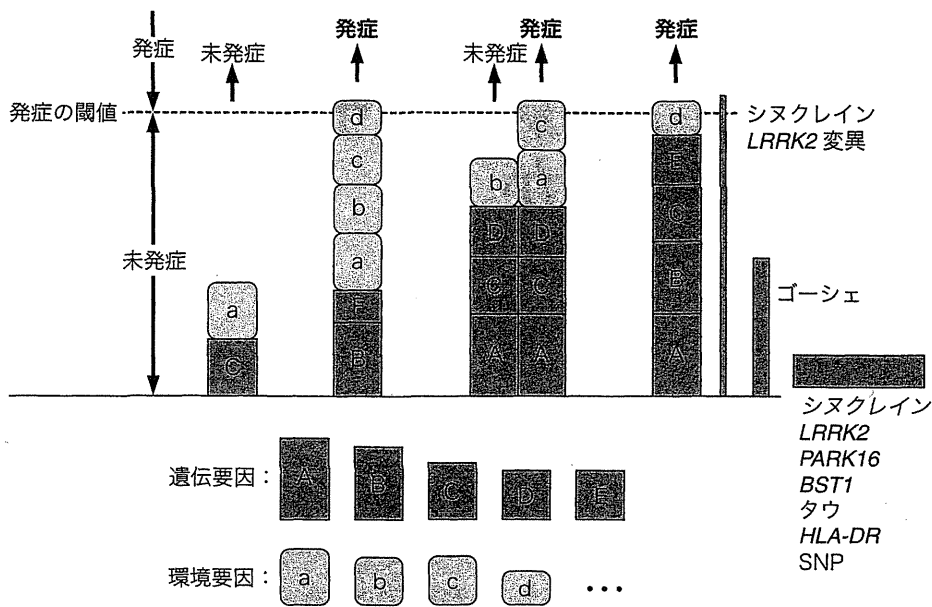


Fig. 1. 孤発性 PD は多因子遺伝性疾患

PD, アルツハイマー病, または生活習慣病を含むほとんどの疾患は, 複数の遺伝的要因と複数の環境要因の積み木の総和が, ある閾値を超えたとき発症すると考えられている。メンデル遺伝性変異以外に, common variant として α -シヌクレイン, PARK16, BST1, LRRK2, タウ, HLA-DR の SNP が, また rare variant としてゴーシェ病遺伝子が重要である。

など)が明らかにされた。これまでにレビー小体の主要構成成分である α -シヌクレインのほか, プロテアソーム系に関係する遺伝子(パーキン, UCH-L1), ミトコンドリアに関係する遺伝子(PINK1, DJ-1), 酸化ストレスに関係する遺伝子(DJ-1)などが同定されている。孤発性 PD, メンデル遺伝性 PD とも, 一部共通の発症メカニズムが存在していると考えられ, それらを切り口にして孤発性 PD の病態解明が進んでいる¹⁾。

メンデル遺伝性 PARK 遺伝子の孤発性 PD 病態における立ち位置

1) PARK1, PARK4(α -シヌクレイン)について→1997年に PARK1(α -シヌクレイン)遺伝子が同定されたのを皮切りに, 次々に PARK 遺伝子の同定が進むと, そのほとんどが孤発性 PD の病態メカニズムにも関係する蛋白であることが判明した(Table 1, Fig. 2)¹⁾。もっともインパクトが強かったのは, PARK1 遺伝子産物 α -シヌクレイン

蛋白が, メンデル遺伝性 PD, 孤発性 PD 両方のレビー小体の主要成分であったという報告である。 α -シヌクレインはメンデル遺伝性 PD, 孤発性 PD の病態に共通した分子であることにより, α -シヌクレイン変異の解析は, 孤発性 PD のモデルにもなると考えられる。アミノ酸変異をもつ α -シヌクレインは *in vitro* で重合化が促進することが知られており, 野生型 α -シヌクレイン蛋白もリン酸化やニトロ化を受けて患者脳で蓄積していることが報告されている²⁾。また, PARK4 が, 野生型 α -シヌクレインを含む 1.6~2.0 Mb の三重重複であったことが報告され, 正常 α -シヌクレインの mRNA, 蛋白発現レベルが増加すれば発症にいたることが示された。

現在, α -シヌクレイン蛋白は生理的には可溶性単量体(monomer)として存在し, 特定の構造をとらないが, 種々の要因で重合化, 不溶化, 凝集する過程で何らかの神経毒性をもたらすと考えられている。そして, 孤発例では, α -シヌクレインの

Table 1. 孤発性 PD の機序に関する主な仮説とエビデンス

機序	エビデンス
ミトコンドリア機能障害	患者にミトコンドリア機能低下 パーキンソン様症状を起こす神経毒(MPTP, rotenone など)はミトコンドリア機能を障害 ミトコンドリア機能に関係する遺伝子がメンデル遺伝型家系で同定(PINK1, DJ-1 など)
プロテアソーム機能低下	プロテアソーム機能に関係する遺伝子がメンデル遺伝型家系で同定(パーキン, UCH-L1 など) PD 動物モデルでプロテアソーム活性低下
酸化ストレス	患者で酸化ストレスマーカーの増加 パーキンソン様症状を起こす神経毒(MPTP, rotenone, 6-OHDA など)の作用過程でフリーラジカル産生 メンデル遺伝型家系遺伝子の 1 つ DJ-1 は酸化ストレス防御作用
α -シヌクレイン蓄積	レビー小体の主成分は α -シヌクレイン メンデル遺伝型家系でみられるアミノ酸変異をもつ合成蛋白の凝集性促進 メンデル遺伝型の正常 α -シヌクレイン 遺伝子重複家系で, 蛋白・遺伝子発現レベルが上昇 培養細胞, 動物モデルで, 封入体を伴った神経細胞死

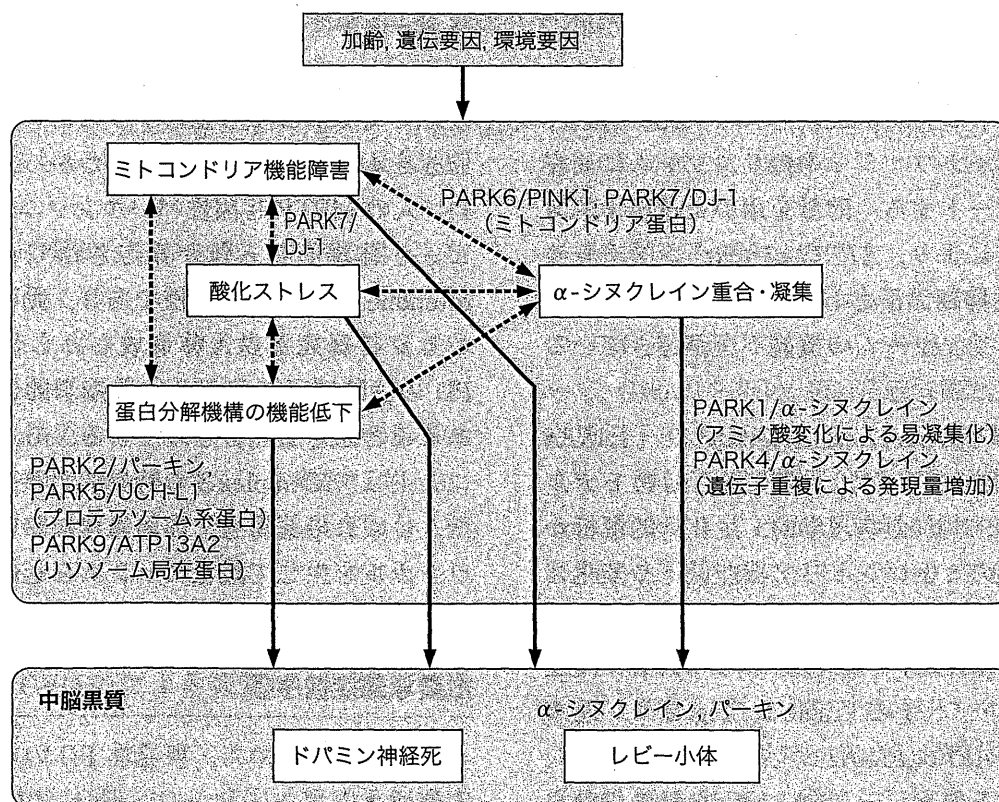


Fig. 2. 孤発性 PD 発症機序と PARK 遺伝子

重合が促進されるような、変異以外の要因が働いて、発症にいたると考えられている³⁾。その際には α -シヌクレインが小胞体-ゴルジ体の小胞輸送を障害し、神経変性を起こすことが報告されている⁴⁾。

2) PARK2 (パーキン), PARK9 (ATP13A2) について→孤発性 PD の病理では、レビー小体がユビキチン抗体で染まる、脳でオートファジーの亢進がみられる、という報告がある。このことと、蛋白分解機構に関する PARK 遺伝子の存在によ

り、プロテアソーム系、オートファジー系の蛋白分解機構の破綻がPDの病態に関係するという仮説がたてられた(Table 1, Fig. 2)^{5,6)}。PARK2(パーキン)蛋白の機能については、ユビキチン・プロテアソーム系蛋白分解において、基質にユビキチンを結合させるE3ユビキチンリガーゼであることが明らかにされた。ユビキチンリガーゼはプロテアソームにより分解を受けるべき蛋白質にユビキチンという目印をつける酵素であり、ユビキチン・プロテアソーム系の破綻によって本来分解される蛋白質(すなわちパーキンの基質)が蓄積することで神経変性を引き起こしていると考えられている。このパーキンの基質はさまざまな報告があるが、真の基質はどれであるか結論が出ていない。

PARK9(ATP13A2)は近年発見され、進行が早く核上性上方注視麻痺を呈するなど非典型的である。原因遺伝子はlysosomal type 5 P-type ATPaseと呼ばれ、オートファジー系のリソソームに局在する蛋白をコードする。酵母を用いた遺伝学的スクリーニングにより、ATP13A2が α -シヌクレインの細胞毒性抑制因子として同定された⁷⁾。劣性遺伝であり、リソソームの機能不全が起きていると考えられる。

3) PARK6(PINK1), PARK7(DJ-1), PARK8(LRRK2)について→また、剖検脳や動物モデル(薬物性PDモデル)の病理学的・生化学的解析から、孤発例でのミトコンドリア機能低下や酸化ストレスマーカーの増加が報告されていた。PARK6(PINK1), PARK7(DJ-1)は、これらの機構に関連する蛋白をコードする(Table 1, Fig. 2)⁸⁾。近年パーキンとPARK6(PINK1)が共同してミトコンドリアのオートファジー(ミトファジー)の役割を担うことが明らかにされ、変異があると機能低下した異常ミトコンドリアの除去ができなくなると考えられる⁹⁾。PARK7の原因はDJ-1蛋白であり、ミトコンドリアに移行し抗酸化作用をもちドパミン細胞死との関係が注目される。またPARK8は常染色体優性遺伝形式を呈するものの中で最多であり、原因蛋白はLRRK2というキナーゼであり、

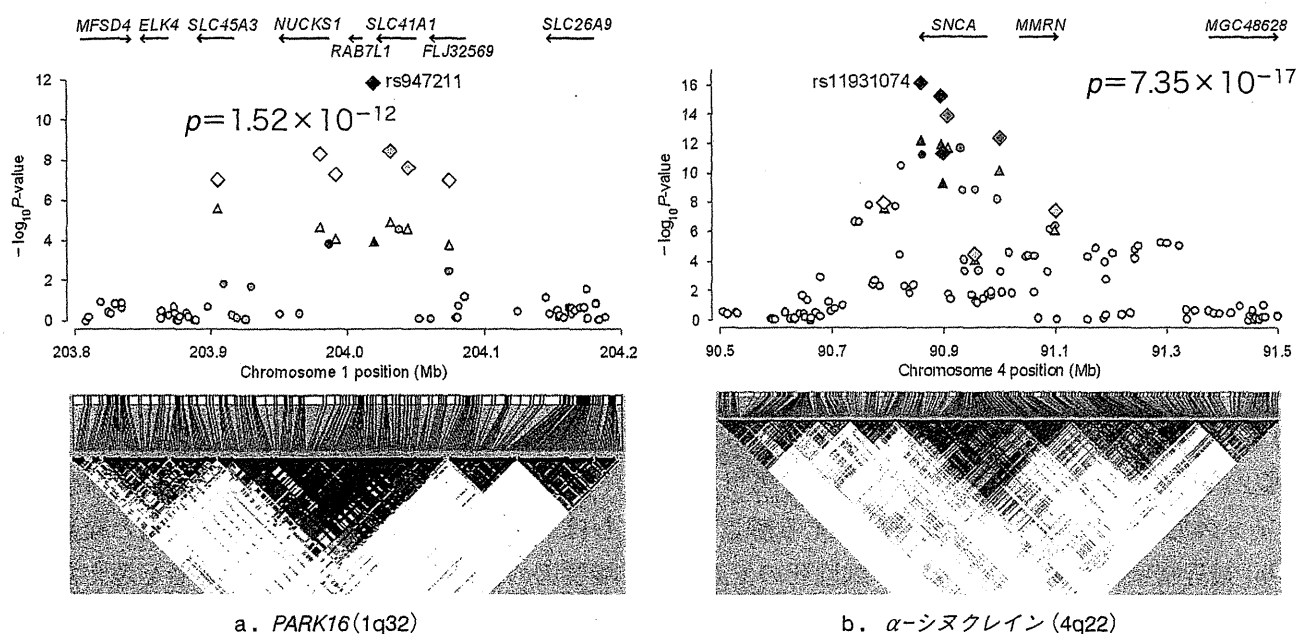
疾患型変異はキナーゼ活性の亢進をもたらすことから、そのリン酸化基質がPDにおいて重要な役割を担うと考えられる。シヌクレインに対するキナーゼ活性が注目されているが定かではない。

神経毒の解析

PDの症状を引き起こす薬物として、MPTP(1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine), ロテノン(rotenone), 6-OHDAなどが知られており、これらの作用機序からも、病態を考察することができる。これらの神経毒のうち、病理学的にも孤発性PDと類似する所見を示すのが、MPTPとrotenoneである。MPTPは麻薬の合成過程の副産物であり、脳内でMPP⁺となりドパミントランスポーターを介してドパミン神経に取り込まれ、エネルギー産生場であるミトコンドリア電子伝達系のComplex Iを阻害することにより、神経細胞死をもたらすということがわかっている(Table 1)。人のMPTP起因性パーキンソンニズムでは、レービ小体様の封入体が報告されている。また、rotenoneは、農薬の一種であり、動物モデルでドパミン神経死と封入体形成がみられる。農薬がPDの危険因子であるという疫学的研究の報告があることから注目されている。rotenoneもミトコンドリアComplex Iの特異的阻害薬である。また、いずれの薬物も、作用過程において、フリーラジカルを生じて酸化ストレスを起こすことも知られている(Table 1)。

主要な発症機序

これらをまとめて、孤発性PDの神経死にいたる機序として、①ミトコンドリア機能障害により、細胞内エネルギー産生が低下する、②蛋白分解機構(プロテアソーム系、オートファジー系など)機能低下により、神経細胞内に構造異常をきたした蛋白が貯まって、細胞毒性をもつ、③酸化ストレス(活性酸素物質: reactive oxygen species (ROS))が細胞毒性をもつ、などがさまざまな割合で関与していること、そして④ α -シヌクレイン



a. *PARK16* (1q32)

b. α -シヌクレイン (4q22)

Fig. 3. 日本人 PD 患者 2011 検体, 対照 18381 検体のゲノムワイド関連解析より同定された 4 つの PD 遺伝子座の抜粋

このうち 2 つはまったく新規の領域であった。とくに(a)は非常に強い関連を示し、白人集団でも関連が再現されたことから *PARK16* と名づけた。また残りの 2 つは、常染色体優性遺伝性パーキンソンニズムの原因遺伝子 α -シヌクレイン (b) と *LRRK2* を含む領域が同定された。縦軸は $-\log P$, すなわち上に行くほど関連が強い。 [文献 10) より引用]

蛋白がこれらの経路あるいは別の経路に関与していることについては、あまり異論がない。さらに各機序の関与を裏づける報告も剖検脳の病理学的・生化学的解析から得られている (Table 1)。これらの機序は複数の遺伝要因や環境要因の影響のもとで、それぞれ、あるいは互いに影響しあって、孤発性 PD の病態を形成していくと現在のところ考えられている (Fig. 2)。

ゲノムワイド関連解析 (Genomewide Association Study : GWAS) による孤発性 PD 感受性遺伝子の同定●

ここ 10 年間孤発性 PD の疾患感受性遺伝子の発見を目指した多くの研究がなされてきたが、アルツハイマー病における ApoE4 多型のような確実に発症リスクを高める遺伝因子はなかなか確認されていなかった。ゲノムワイド有意水準 ($p < 5 \times 10^{-8}$) を満たす確実なものは、 α -シヌクレインの 3' 非翻訳領域 SNP (Gasser ら, 筆者らが同定) とゴーシェ病遺伝子 *GBA* の rare variant の 2 つ

の遺伝子のみであった。

われわれは、大規模の患者対照集団と、56 万個の SNP を搭載したイルミナ Hap550 アレイを用いて、GWAS を行い、PD 発症に関わる 2 つの新しい遺伝子座 *PARK16*; *BST1* を同定した (Fig. 3a)。また、常染色体優性遺伝性 PD の原因遺伝子 α -シヌクレイン、*LRRK2* の孤発性 PD への関与を証明した (Fig. 3b)。白人の PD の GWAS 研究を行っていたグループと共同し、 α -シヌクレイン・*PARK16*・*LRRK2* は 2 人種に共通の PD リスクであり、*タウ*・*BST1* のリスク多型の影響は人種特異的であることを示した¹⁰⁾。

また孤発性 PD の発症に、常染色体優性遺伝性 PD の原因遺伝子が密に関係していることを示した。原因変異により常染色体優性遺伝性 PD を引き起こす遺伝子が、その SNP を通して、孤発性 PD の遺伝リスクとなったことは興味深い。すなわち、 α -シヌクレイン、*LRRK2* が関与する病理経路が、孤発性を含めた PD 全体の病因の根幹をなすことを示している。一方、微小管の安定化に

働くタウは、アルツハイマー病、前頭側頭型認知症などの原因産物として知られており、PD と他神経変性疾患との共通の病理経路の存在をも示唆している¹⁰⁾。

PARK16 領域には、3つの遺伝子が存在するが、発現量的形質座解析から、*NUCKS1* が、もっとも有力な責任遺伝子であると考えた (Fig. 3a)。*NUCKS1* は、リン酸化部位を含む核蛋白であるが、神経系における機能は、未知であり、新たなPD 発症病態を開拓する可能性がある。*BST1* は、細胞内 Ca^{2+} 貯蔵からの Ca^{2+} 放出を誘発するサイクリック ADP リボースの形成を触媒する酵素であり、最近提唱されているドパミン細胞死の Ca^{2+} ストレス説を想起させ、興味深い。これらは従来のPD 病態説からはまったく新規な遺伝子であり、従来説にとらわれず、新規なものを同定できるところにGWAS の強みがある。

さらに最近別グループのGWAS から、HLA-DR 抗原の領域との関連が報告された。PD 脳ではDR 抗原陽性のミクログリア(後述)が検出されること、非ステロイド抗炎症薬(NSAIDs)は疫学的にPD のリスクを減少させることなどもあり、PD と炎症の関係を示しており興味深い¹¹⁾。

PD とゴーシェ病とまれな多型 (rare variant) ●

リピーシスの常染色体劣性遺伝のユダヤ人ゴーシェ病家系内にPD 患者が多いことから、PD では*GBA* (glucocerebrosidase, 1q21) 変異のヘテロ保因者が有意に多いことが報告された¹²⁾。東京大学神経内科とわれわれの共同研究グループは、*GBA* 遺伝子の全11エクソンとその近傍をPD 患者534人、対照544人リシーケンスして塩基配列変化の有無を調べ、11種類の疾患原性点変異が同定された。ヘテロでもつ保因者はPD 患者534人中50人(9.4%)、対照544人中2人(0.37%)であり、PD と*GBA* 変異は強く関連していた($p = 6.9 \times 10^{-14}$, オッズ比28.0)¹²⁾。

さらに全世界での米国人、フランス人、ポルトガル人、台湾人など計約10,000人の患者対照集団

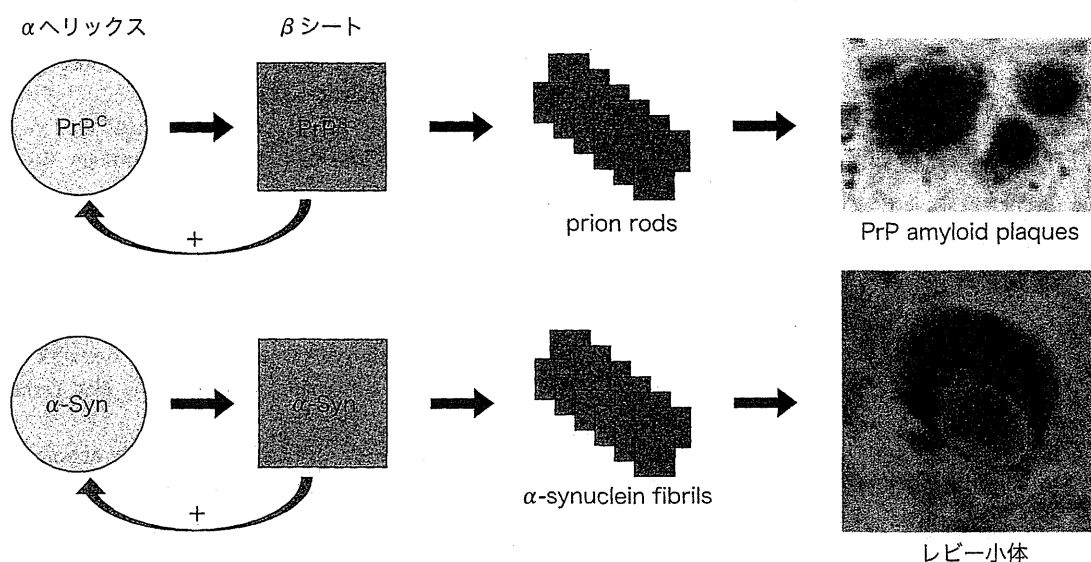
のメタ解析により、どの人種にても*GBA* 遺伝子はリスクとなり平均オッズ比は5である¹³⁾。*GBA* 変異は確実なPD リスク因子であり、SNP によるCommon Disease-Common Variants 仮説と異なり、Common Disease-Multiple Rare Variant 仮説によるものである。すなわち、頻度は低いが発症へのeffect size が大きい。なぜリスクとなるかの原因は不明であるが、単なる酵素活性低下でなく、*GBA* 変異の蛋白分解機構への負担によることも推定されている。

PD、アルツハイマー病、または生活習慣病を含むほとんどの疾患は、複数の遺伝要因と複数の環境要因の積み木の総和が、ある閾値を超えたとき発症すると考えられている。これまでみてきたように α -シヌクレインや*LRRK2* の変異はそれ1つだけで閾値に到達し発症するが、対象患者はほとんど存在しないので積み木の幅はととても狭い。ゴーシェ病遺伝子などのrare variant は中等度の高さをもつが10%以下の患者にしか当てはまらないため幅は狭い。一方SNP は、それ自体のオッズは低いがほとんどの患者に当てはまるため、積み木の幅は広い。いずれも重要である (Fig. 1)。

シヌクレインプリオン仮説●

本特集別稿を参照されたいが、Braak らは α -シヌクレインの蓄積部位を多数例のPD 脳で詳細に検討し、この蛋白質の異常蓄積はまず嗅球、延髄から始まり、徐々に上行して橋、中脳に及び、最終的に大脳皮質にいたるという仮説を提唱し、Braak 仮説として有名である。

一方、PD の治療を目的として胎児の中脳腹側組織を移植し10年以上経過した患者脳を病理学的に検索したところ、驚くべきことに移植された神経細胞にレビー小体が認められた¹⁴⁾。また α -シヌクレインを高発現する神経細胞とマウスの幹細胞を共培養する*in vitro* 実験や、 α -シヌクレインを発現するトランスジェニックマウスの海馬に幹細胞を移植する*in vivo* 実験にて、いずれの場合にも幹細胞の細胞質に α -シヌクレインの凝集体が



未知の環境要因→嗅覚器や腸管(外界)→自律神経叢や嗅球で α -シヌクレインのミスフォールディング→放出伝播?

Fig. 4. PD のシヌクレインプリオンモデル

[文献 16) より引用]

形成されること、 α -シヌクレインがエンドサイトーシスによって幹細胞に取り込まれることが示された。これらの知見は、 α -シヌクレインが近接する神経細胞に「伝染」する可能性を示している。またごく最近、細胞内の α -シヌクレインが、他の細胞に移動し凝集し、その凝集体の量が継時的に増加すること、 α -シヌクレイン蛋白の細胞内取り込みに関してエンドサイトーシス阻害薬の同時注入により、取り込みの減少が明らかとなった¹⁵⁾。

α -シヌクレインもプリオン蛋白質と同様に、 α ヘリックス構造と β シート構造をとりうることから、何らかのきっかけで β シート化した α -シヌクレインがシードとなって凝集を開始するとともに、細胞外に放出されて近隣の細胞に取り込まれ、プリオン病と同様の機序で病変が進展するという、PD のプリオンモデルが提唱されている (Fig. 4)。さらにこのモデルと Braak 仮説を合わせて PD の発症機序として、未知の環境要因が嗅覚器や腸管など外界に接した神経組織に作用し、自律神経叢や嗅球において最初に α -シヌクレインのミスフォールド蛋白質が形成され、逆行性に細

胞体まで輸送され、中枢神経内で放出されて近隣の神経細胞に連続的に伝播していく、という大胆な仮説もある (Fig. 4)¹⁶⁾。中枢神経内における α -シヌクレインの伝播を防ぐことで (例：エンドサイトーシス阻害薬) PD の進行を抑制するという、新たな治療戦略が考えられる。

ミクログリアと炎症反応●

孤発性 PD 患者剖検脳や家族性 PD 患者剖検脳、さらに MPTP 誘導性 PD モデル動物脳の黒質線条体部において共通にみられる病理変化として、ミクログリアの活性化と集積がある。ミクログリアは活性化すると HLA-DR や iNOS や COX-2 を発現し、インターロイキン (IL)-1 β や IL-6 などの炎症性サイトカインを含む種々のサイトカインを産生し、炎症反応の特徴の多くを示す。また、活性化したミクログリアは炎症性サイトカイン以外に、活性酸素や NO, プロテアーゼ類などの神経傷害性因子も産生する。とくに活性酸素や NO などのフリーラジカルは PD における神経変性の進行機序に関与するため、ミクログリ

アの活性化や脳内炎症反応は PD の発症や進行に密接に関わっている¹⁷⁾。MPTP 曝露 16 年後の死後、脳においても活性化ミクログリアが存在し、ミクログリアの活性化を伴う長期間にわたる継続的な炎症状態が PD の病態の進行において重要な意味をもつ可能性を示唆している。GWAS で HLA-DR 抗原の領域が同定されたことも興味深い¹¹⁾。

おわりに

PD のメカニズムの解明にあたっては、メンデル遺伝型 PD の原因遺伝子の研究成果が大きく貢献してきた。今後は未知のメンデル遺伝型遺伝子の同定のみならず、孤発性 PD において大規模な GWAS, rare variant 解析から、新規の感受性遺伝子が同定され、新しい機序、創薬の手がかりがみつかる可能性がある。次世代シーケンサーによるエクソームリシーケンスも開始され、今後はパーソナルゲノム解析が医学研究に応用されよう。

また、プリオン仮説は異常蛋白凝集を特徴とする神経変性疾患に共通した病態機序である可能性がある。同時に、異常蛋白の細胞間伝播は神経変性疾患に共通の治療ターゲットとなる可能性を秘めており、今後に期待する。

文献

- 1) Farrer MJ : Genetics of Parkinson disease : paradigm shifts and future prospects. *Nat Rev Genet* 7 : 306, 2006
- 2) Fujiwara H et al : alpha-Synuclein is phosphorylated in synucleinopathy lesions. *Nat Cell Biol* 4 : 160, 2002
- 3) Uversky VN : Neuropathology, biochemistry, and biophysics of alpha-synuclein aggregation. *J Neurochem* 103 : 17, 2007
- 4) Cooper AA et al : Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* 313 : 324, 2006
- 5) McNaught KS et al : Proteasomal dysfunction in sporadic Parkinson's disease. *Neurology* 66 [Suppl 4] : S37, 2006
- 6) Martinez-Vicente M, Cuervo AM : Autophagy and neurodegeneration : when the cleaning crew goes on strike. *Lancet Neurol* 6 : 352, 2007
- 7) Gitler AD et al : Alpha-synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nat Genet* 41 : 308, 2009
- 8) Schapira AH : Mitochondria in the aetiology and pathogenesis of Parkinson's disease. *Lancet Neurol* 7 : 97, 2008
- 9) Narendra DP et al : PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. *PLoS Biol* 28 : e1000298, 2010
- 10) Satake W et al : Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet* 41 : 1303, 2009
- 11) Hamza TH et al : Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat Genet* 42 : 781, 2010
- 12) Mitsui J et al : Mutations for Gaucher disease confer a high susceptibility to Parkinson disease. *Arch Neurol* 66 : 571, 2009
- 13) Sidransky E et al : Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 361 : 1651, 2009
- 14) Li JY et al : Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nat Med* 14 : 501, 2008
- 15) Hansen C et al : α -Synuclein propagates from mouse brain to grafted dopaminergic neurons and seeds aggregation in cultured human cells. *J Clin Invest* 121 : 715, 2011
- 16) Olanow CW, Prusiner SB : Is Parkinson's disease a prion disorder? *Proc Natl Acad Sci U S A* 106 : 12571, 2009
- 17) Sawada M et al : Effects of aging on neuroprotective and neurotoxic properties of microglia in neurodegenerative diseases. *Neurodegener Dis* 5 : 254, 2008

A functional variant in *ZNF512B* is associated with susceptibility to amyotrophic lateral sclerosis in Japanese

Aritoshi Iida¹, Atsushi Takahashi², Michiaki Kubo³, Susumu Saito³, Naoyo Hosono³, Yoza Ohnishi³, Kazuma Kiyotani⁴, Taisei Mushiroda⁴, Masahiro Nakajima¹, Kouichi Ozaki⁵, Toshihiro Tanaka⁵, Tatsuhiko Tsunoda⁶, Shuichi Oshima⁸, Motoki Sano⁹, Tetsumasa Kamei¹⁰, Torao Tokuda¹¹, Masashi Aoki¹², Kazuko Hasegawa¹³, Koichi Mizoguchi¹⁴, Mitsuya Morita¹⁵, Yuji Takahashi¹⁶, Masahisa Katsuno^{17,18}, Naoki Atsuta¹⁷, Hirohisa Watanabe¹⁷, Fumiaki Tanaka¹⁷, Ryuji Kaji¹⁹, Imaharu Nakano¹⁵, Naoyuki Kamatani², Shoji Tsuji¹⁶, Gen Sobue¹⁷, Yusuke Nakamura^{7,20} and Shiro Ikegawa^{1,*}

¹Laboratory for Bone and Joint Diseases and ²Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Tokyo 108-8639, Japan, ³Laboratory for Genotyping Development, ⁴Laboratory for Pharmacogenetics, ⁵Laboratory for Cardiovascular Diseases, ⁶Laboratory for Medical Informatics and ⁷Laboratory for International Alliance, Center for Genomic Medicine, RIKEN, Yokohama 230-0045, Japan, ⁸Department of Neurosurgery, Chiba Tokushukai Hospital, Funabashi 274-8503, Japan, ⁹Department of Neurology, Chibanishi General Hospital, Matsudo 270-2251, Japan, ¹⁰Department of Neurology, Chigasaki Tokushukai General Hospital, Chigasaki 253-8558, Japan, ¹¹Tokushukai Group, Tokyo 102-0093, Japan, ¹²Department of Neurology, Tohoku University School of Medicine, Sendai 980-8574, Japan, ¹³Department of Neurology, National Hospital Organization Sagami-hara National Hospital, Sagami-hara 228-8522, Japan, ¹⁴Department of Neurology, Shizuoka Institute of Epilepsy and Neurological Disorders, Shizuoka 420-8688, Japan, ¹⁵Division of Neurology, Department of Medicine, Jichi Medical University, Shimotsuke 329-0498, Japan, ¹⁶Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan, ¹⁷Department of Neurology, Graduate School of Medicine and ¹⁸Institute for Advanced Research, Nagoya University, Nagoya 466-8550, Japan, ¹⁹Department of Neurology, Graduate School of Medicine, The University of Tokushima, Tokushima 770-8503, Japan and ²⁰Department of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan

Received January 19, 2011; Revised May 26, 2011; Accepted June 6, 2011

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease characterized by the selective loss of motor neurons. Several susceptibility genes for ALS have been reported; however, ALS etiology and pathogenesis remain largely unknown. To identify further ALS-susceptibility genes, we conducted a large-scale case-control association study using gene-based tag single-nucleotide polymorphisms (SNPs). A functional SNP (rs2275294) was found to be significantly associated with ALS through a stepwise screening approach (combined $P = 9.3 \times 10^{-10}$, odds ratio = 1.32). The SNP was located in an enhancer region of *ZNF512B*, a transcription factor of unknown biological function, and the susceptibility allele showed decreased activity and decreased binding to nuclear proteins. *ZNF512B* over-expression increased transforming growth factor- β (TGF- β) signaling, while knockdown had the opposite effect. *ZNF512B* expression was increased in the anterior horn motor neurons of the spinal cord of ALS patients when compared with controls. Our results strongly suggest that *ZNF512B* is an important positive regulator of TGF- β signaling and that decreased *ZNF512B* expression increases susceptibility to ALS.

*To whom correspondence should be addressed at: Laboratory of Bone and Joint Diseases, Center for Genomic Medicine, RIKEN, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. Tel/Fax: +81 354495393; Email: sikegawa@ims.u-tokyo.ac.jp

© The Author 2011. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a heterogeneous motor neuron disease that results from selective death of motor neurons in the brain and spinal cord (1). The predominant clinical feature of ALS is progressive wasting and weakness of limb, bulbar and respiratory muscles. The mean survival of patients after onset of symptoms is 3–5 years. Its worldwide incidence and prevalence are 0.3–2.4 and 0.7–7.0 per 100 000 each year (2). The heritability of ALS is high, with twin studies estimating it at 0.61 and the unshared environment component at 0.39 (3).

Approximately 10% of ALS cases are familial (fALS), and the remaining 90% are sporadic (sALS). Genetic factors have been reported in ALS. Detailed information regarding ALS-related genes is available via amyotrophic lateral sclerosis online genetics database and the ALS mutation database (4,5). Most fALS is monogenic in origin. At least 15 fALS loci, under various modes of inheritance, have been identified by linkage studies, and pathogenic mutations have been described in 11 genes, *SOD1*, *NEFH*, *ALS2*, *DCTN1*, *VAPB*, *SETX*, *ANG*, *TARDBP*, *FUS*, *OPTN* and *DAO*, in fALS (6–19). Despite the abundance of genes and loci identified in fALS, mutations in these genes explain only a small minority of sALS (20).

Regarding susceptibility genes for sALS, >30 association studies based on the candidate-gene approach have been reported (21,22). Among them, *NEFH*, *APEX* and *ANG* have the most evidence; associations of these genes have been found in Caucasians (23–25) and replicated in several studies (7,22,26). However, many of the reported genes are still controversial. For example, the association of non-synonymous substitution (P413L) in the chromogranin B gene (*CHGB*) is reported in French, French-Canadian and Scandinavian ALS populations (27), but has not been found in a Dutch and another French population (28,29).

The genome-wide association study (GWAS) has identified five ALS-susceptibility genes (*FGGY*, *ITPR2*, *DPP6*, *KIFAP3* and *UNC13A*) and two loci (9p21.2 and 10q26.3) in Caucasian (30–35). These results are promising, but remain slightly controversial (36–39). The association of the 9p21.2 locus has been independently replicated in three studies (34,40,41), but is not found in all populations, including those from Japan and China (42). More studies are necessary to evaluate and confirm these previously reported ALS-susceptibility genes.

To identify novel susceptibility genes for ALS, we conducted a large-scale genetic association study in Japanese ALS patients using gene-based single-nucleotide polymorphisms (SNPs) (43). We identified a functional SNP that was significantly associated with ALS. The SNP was located in an enhancer region of *ZNF512B*, a previously uncharacterized transcription factor, and the susceptibility allele of the SNP had decreased enhancer activity for the *ZNF512B* promoter and decreased binding capacity to nuclear proteins. We found that in neuron cells, *ZNF512B* acts as a positive regulator of transforming growth factor- β (TGF- β) signaling, which is known to be neuroprotective and critical for maintenance and/or survival of neurons (44–46). We demonstrated the localization of *ZNF512B* in the spinal cord of ALS patients and it showed enhanced expression in motor neuron cells of the anterior horn when compared with controls.

RESULTS

Genome screening

We carried out a stepwise case–control association study (Supplementary Material, Fig. S1) as previously described (47–51). In stage 1 of the discovery series, 92 ALS and 233 control subjects were analyzed at 52 608 gene-based SNP loci selected from the JSNP database (43). Genotype information was successfully obtained for 48 939 SNPs on autosomal chromosomes passed after the quality control. Either the Chi-square test or Fisher's exact test was performed for three genetic models: dominant, recessive and allelic. Comparison of observed and expected distributions showed no evidence for inflation of the trend test statistics (inflation factor, $\lambda = 1.04$; Supplementary Material, Fig. S2). Also, principal component analysis (52) in stage 1 and HapMap samples showed no evidence of population stratification between the case and control groups (Supplementary Material, Fig. S3). In stage 2 of the discovery series, 893 SNPs that showed P -values of ≤ 0.01 in stage 1 were genotyped for an additional 1087 subjects (362 ALS cases and 725 controls). Subsequently, 10 SNPs with P -values < 0.001 were identified by the Chi-square test for the three models (Supplementary Material, Table S1).

Identification of genetic association between rs2275294 and ALS

We validated the association of these SNPs using independent subjects from Biobank Japan (sample set 1). In all, 249 ALS cases and 1030 controls were genotyped and validated the association in rs2275294 (allele model, $P = 1.8 \times 10^{-3}$). The SNP was then genotyped in an independent Japanese population consisting of 602 ALS cases and 2256 controls (sample set 2). Significant association was replicated in this population (allele model $P = 5.6 \times 10^{-5}$). The combined P -values for the stepwise association study calculated by the Mantel–Haenszel method and the joint analysis were 9.3×10^{-10} and 6.7×10^{-10} , respectively (Table 1). The combined P -values remained significant after Bonferroni correction ($9.3 \times 10^{-10} \times 52\,608 \times 3 = 1.47 \times 10^{-4}$). The P -values from the Mantel–Haenszel method and the joint analysis were very similar, supporting the fact that there is no hidden confounder in our population. The minor allele frequency (MAF) of rs2275294 in 744 samples of the Japan Biological Informatics Consortium (JBIC)-genotyping data deposited in the dbSNP database was similar to that of our controls (0.414).

Evaluation of rs2275294

We assessed the stratification using principal component analysis (52). The top six principal components were associated with case–control status. The association of rs2275294 with the top six principal components included as covariates (trend model $P = 0.00287$) was similar to that in stage 1 (trend model $P = 0.00246$), suggesting no stratification. Population stratification was also assessed by evaluating differences in population structure among all case and control sample sets using Wright's F statistics (53). There was no difference in the population structure among these groups (Supplementary Material, Table S2). Potential confounding factors were also