Human Molecular Genetics, 2017, Vol. 26, No. 1 226–232

doi: 10.1093/hmg/ddw379
Advance Access Publication Date: 22 December 2016
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Genome-wide association study of Parkinson's disease in East Asians

Jia Nee Foo[1,2], Louis C. Tan[3], Ishak D. Irwan[2], Wing-Lok Au[3], Hui Qi Low[2], Kumar-M. Prakash[3], Azlina Ahmad-Annuar[4], Jinxin Bei[5], Anne YY Chan[6], Chiung Mei Chen[7], Yi-Chun Chen[7], Sun Ju Chung[8], Hao Deng[9], Shen-Yang Lim[10], Vincent Mok[6], Hao Pang[11], Zhong Pei[12], Rong Peng[13], Hui-Fang Shang[13], Kyuyoung Song[14], Ai Huey Tan[10], Yih-Ru Wu[7], Tin Aung[15,16], Ching-Yu Cheng[15,16,17], Fook Tim Chew[18], Soo-Hong Chew[19], Siow-Ann Chong[20], Richard P. Ebstein[21], Jimmy Lee[17,20], Seang-Mei Saw[15,16,17,22], Adeline Seow[22], Mythily Subramaniam[20], E-Shyong Tai[23], Eranga N. Vithana[15,16,17], Tien-Yin Wong[15,16,17], Khai Koon Heng[2], Wee-Yang Meah[2], Chiea Chuen Khor[2,15,24], Hong Liu[25], Furen Zhang[25], Jianjun Liu[2,†,*] and Eng-King Tan[3,17,†,*]

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, [2]Human Genetics, Genome Institute of Singapore, A*STAR, Singapore, [3]Department of Neurology, National Neuroscience Institute, Singapore, [4]Department of Biomedical Science, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia, [5]Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangzhou, Guangdong, People's Republic of China, [6]Neurology Division, Department of Medicine and Therapeutics, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong SAR, People's Republic of China, [7]Department of Neurology, Chang Gung Memorial Hospital, Chang Gung University, Taipei, Taiwan, Republic of China, [8]Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea, [9]Department of Neurology and Center for Experimental Medicine, the Third Xiangya Hospital, Central South University, Changsha, People's Republic of China, [10]Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia, [11]School of Forensic Medicine, China Medical University, Shenyang, People's Republic of China, [12]Department of Neurology, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, People's Republic of China, [13]Department of Neurology, West China Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China, [14]Department of Biochemistry and Molecular Biology, University of Ulsan College of Medicine, Seoul, Korea, [15]Singapore Eye Research Institute, Singapore, [16]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University Health System, National University of Singapore, Singapore, [17]Duke-National University of Singapore Medical School, Singapore, [18]Department of Biological Sciences, National University of Singapore, Singapore, [19]Department of Economics,

[†]These authors jointly directed this work.

Received: July 22, 2016. **Revised:** October 28, 2016. **Accepted:** November 2, 2016

National University of Singapore, Singapore, [20]Institute of Mental Health, Singapore, [21]Department of Psychology, National University of Singapore, Singapore, [22]Saw Swee Hock School of Public Health, National University Health System, National University of Singapore, Singapore, [23]Department of Medicine, Yong Loo Lin School of Medicine, National University Health System, National University of Singapore, Singapore, [24]Department of Biochemistry, Yong Loo Lin School of Medicine, National University Health System, National University of Singapore, Singapore and [25]Shandong Provincial Institute of Dermatology and Venereology, Shandong Academy of Medical Science, Jinan, Shandong, People's Republic of China

*To whom correspondence should be addressed at: Jianjun Liu, Human Genetics, Genome Institute of Singapore, A*STAR, 60 Biopolis Street, Genome #02-01, Singapore 138672. Tel: +65-6808-8088; Fax: +65-6808-8304; Email: liuj3@gis.a-star.edu.sg. Eng-King Tan, Department of Neurology, Singapore General Hospital, Level 6, Academia, 20 College Road, Singapore 169856. Tel: +65-6326 5003; Fax: +65-6220-3321; Email: tan.eng.king@sgh.com.sg

Abstract

Genome-wide association studies (GWAS) on Parkinson's disease (PD) have mostly been done in Europeans and Japanese. No study has been done in Han Chinese, which make up nearly a fifth of the world population. We conducted the first Han Chinese GWAS analysing a total of 22,729 subjects (5,125 PD cases and 17,604 controls) from Singapore, Hong Kong, Malaysia, Korea, mainland China and Taiwan. We performed imputation, merging and logistic regression analyses of 2,402,394 SNPs passing quality control filters in 779 PD cases, 13,227 controls, adjusted for the first three principal components. 90 SNPs with association $P < 10^{-4}$ were validated in 9 additional sample collections and the results were combined using fixed-effects inverse-variance meta-analysis. We observed strong associations reaching genome-wide significance at *SNCA*, *LRRK2* and *MCCC1*, confirming their important roles in both European and Asian PD. We also identified significant ($P < 0.05$) associations at 5 loci (*DLG2*, *SIPA1L2*, *STK39*, *VPS13C* and *RIT2*), and observed the same direction of associations at 9 other loci including *BST1* and *PARK16*. Allelic heterogeneity was observed at *LRRK2* while European risk SNPs at 6 other loci including *MAPT* and *GBA-SYT11* were non-polymorphic or very rare in our cohort. Overall, we replicate associations at *SNCA*, *LRRK2*, *MCCC1* and 14 other European PD loci but did not identify Asian-specific loci with large effects (OR > 1.45) on PD risk. Our results also demonstrate some differences in the genetic contribution to PD between Europeans and Asians. Further pan-ethnic meta-analysis with European GWAS cohorts may unravel new PD loci.

Introduction

Parkinson's disease (PD; OMIM:168600) is among the most common age-related neurodegenerative diseases worldwide. It occurs at similar incidence in Europeans and Asians, and this is expected to increase over the coming decades as the size of the aged population grows. Large-scale meta-analyses of genome-wide association studies (GWAS) on PD in the European population analysing up to 13,708 cases and 95,282 controls have led to the identification of 32 loci significantly associated with disease risk (1). One of the first PD GWAS was conducted in Japan (2), which led to the identification of loci such as *PARK16* and *BST1* that were subsequently replicated in Europeans (1–9), and also demonstrated the absence of the *MAPT* risk variant in Asians. The Han Chinese population is the largest worldwide, with 1.37 billion people in mainland China alone. This ethnic group makes up about 19% of the world population and a significant fraction of PD patients worldwide, thus it will be important to elucidate the genetic risk factors underlying morbidity in this population. To date, no study has been done in Han Chinese samples. Here, we conducted the first Han Chinese GWAS in 779 Singapore Chinese cases and 13,227 controls with validation in samples from Hong Kong, Malaysia, Korea, mainland China and Taiwan, analysing a total of 22,729 subjects (5,125 cases and 17,604 controls) from across East Asia.

Results

Discovery-stage GWAS

At the discovery stage, we analysed a total of 2,402,394 SNPs in 779 cases and 13,227 controls after quality control filtering (Supplementary Material, Fig. S1). Genome-wide inflation of association evidence was low (lambda GC = 1.017) in the discovery analysis, suggesting minimal bias resulting from population stratification (Supplementary Material, Fig. S2). We observed strong association at the *SNCA* locus crossing the threshold for genome-wide significance (Supplementary Material, Fig. S1). There was no secondary signal observed within the *SNCA* locus, although we might not have had sufficient power in the current study to identify independent associations at variants with smaller effects or lower minor allele frequencies.

We selected 62 SNPs with $P < 10^{-4}$ representing 46 loci for further validation. Taking into consideration that some true signals may be masked by inclusion of other disease cases amongst our controls or by the differential genotyping quality of SNPs across the different arrays, we included another 11 SNPs representing 9 loci with $P < 10^{-4}$ if we analyse only healthy controls (8,996 controls, excluding other disease cases), and another 23 SNPs representing 21 loci crossing the same threshold if we analyse only samples genotyped on similar arrays (OmniZhonghua and OmniExpress; 3,496 controls, 489,262 overlapping genotyped SNPs). For loci tagged only by imputed SNPs, we selected only those with more than 1 imputed SNPs crossing the $P < 10^{-4}$ threshold for validation. In total, 96 SNPs representing 76 independent loci were taken forward for validation in eight independent sample collections across East Asia (first validation) (Supplementary Material, Table S2). These made up a total of 3,769 cases and 3,762 controls that were genotyped on the Sequenom or Taqman platforms. Of the 76 loci, only 4 (*MCCC1*, *LRRK2*, *SNCA* and *DLG2*) have been previously reported (Fig. 1).

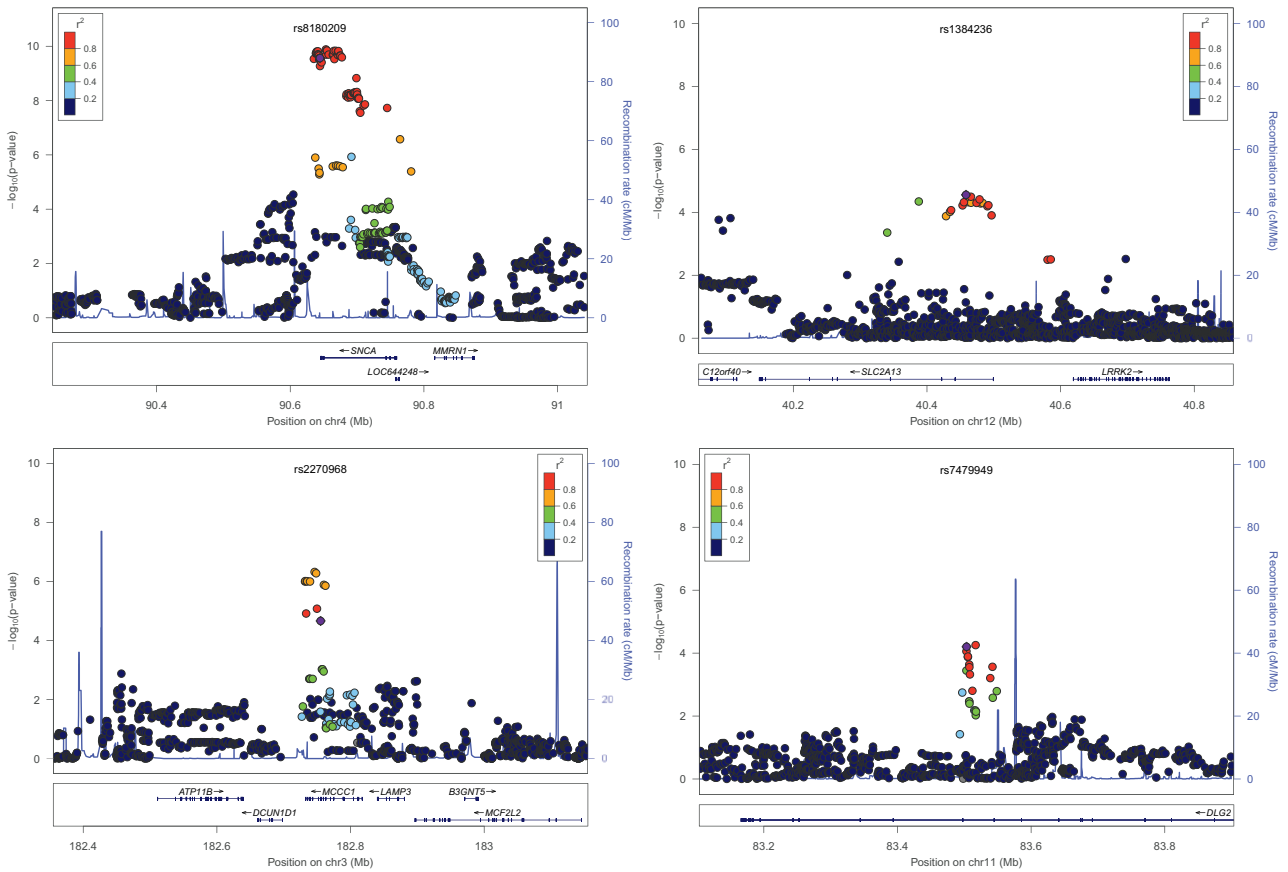

Figure 1. Recombination plot (10) showing association *P*-values of SNPs at 4 known loci taken forward for validation, with colours representing the LD with respect to the validated SNP (purple diamond and labelled on the plot).

Validation of top-associated SNPs in 9 independent sample collections

90 SNPs (73 loci) passed design and genotyping successfully in at least 6 out of 8 validation cohorts (Supplementary Material, Table S3). We tested each SNP for association in each collection using logistic regression analysis, and combined the results using a fixed-effects meta-analysis across six to nine sample collections. From this combined meta-analysis across the discovery and validation samples, only SNPs in three of the previously reported loci (*MCCC1*, *LRRK2*, *SNCA*) reached genome-wide significance (Fig. 2). We brought four other SNPs (including *DLG2*) with $5 \times 10^{-8} < P < 10^{-4}$ in the meta-analysis for further validation in another independent collection of 577 cases and 615 controls from Taiwan (second validation). We observed a consistent association at *DLG2* rs7479949, another previously reported locus which was close to genome-wide significant ($P = 5.90 \times 10^{-8}$; Fig. 3). None of the other three loci were validated (with $P < 0.05$) nor reached genome-wide significance in the combined analysis of 5,125 cases and 17,604 controls (See Supplementary Material, Table S4 for the other three loci).

Analysis of PD risk loci

We then investigated 32 SNPs that were previously reported to have reached genome-wide significance in the larger European PD GWAS (1) in our Han Chinese GWAS discovery dataset (Supplementary Material, Table S5). Besides rs12637471 in *MCCC1* ($r^2 > 0.8$ with rs2270968), 13 other reported SNPs showed consistent directions of effect with that reported in European studies (Supplementary Material, Table S5), out of which 4 (in *RIT2, VPS13C, STK39 and SIP1AL2*) were statistically significant ($P < 0.05$) (Supplementary Material, Fig. S3A), although none of the SNPs within these loci crossed our threshold ($P < 10^{-4}$) in selecting SNPs for validation (Supplementary Material, Fig. S3B). This included *BST1* and *PARK16*, loci that were first reported in Japanese (2) and subsequently validated in Europeans (1). Another six reported SNPs were non-polymorphic in our samples and in East Asians samples from the 1000 genomes project (Supplementary Material, Table S5) and none of the other SNPs across these loci (including *MAPT and GBA;* Supplementary Material, Fig. S3C) showed a significant association in our samples, suggesting that they may be European-specific loci. There was no evidence of consistent association at 5 other SNPs, although larger numbers of samples will be needed to assess the small effects that were previously reported at these loci (Supplementary Material, Fig. S3D). There were also another 7 previously reported SNPs that could not be confidently imputed (with >98% of samples with imputation genotype quality above 0.9, info score >0.8). Within the loci tagged by these SNPs, we observed evidence of association at neighbouring SNPs for the previously reported *SNCA* SNP rs356182 ($r^2 > 0.5$ with reported SNP in 1000 genomes Asians) (10,11) and the reported *DLG2* SNP rs3793947 ($r^2 > 0.8$ with reported SNP in 1000 genomes Asians) (10,11), but there was little evidence of association at neighbouring SNPs within the other 5 loci (Supplementary Material, Fig. S3E).

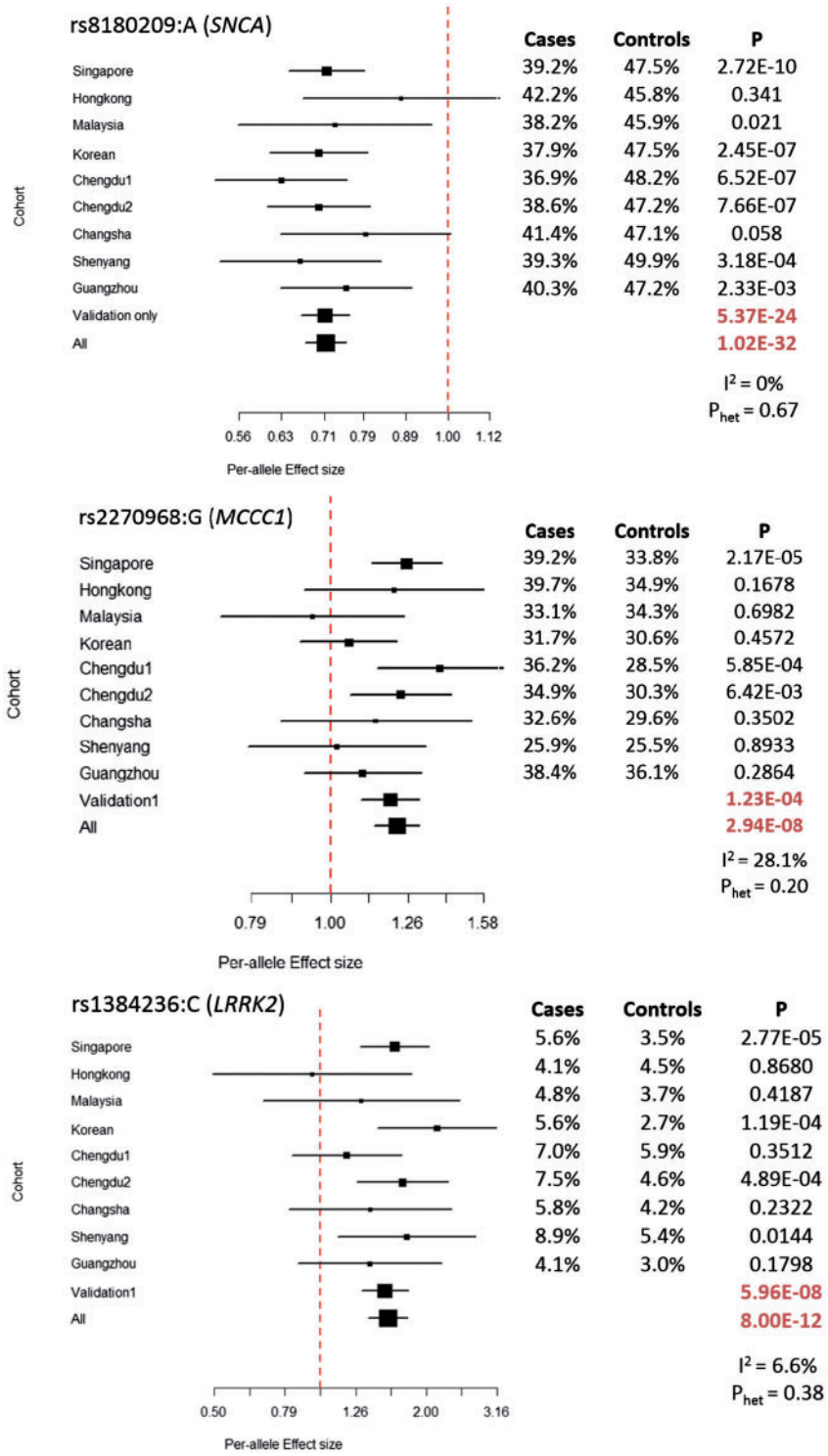

Figure 2. Forest plot showing odds ratios, allele frequencies and *P*-values in each sample collection at validated SNPs within known loci *SNCA*, *MCCC1* and *LRRK2*.

Pathway analysis

To identify pathways and gene sets with a high proportion of genes significantly associated with PD, we ran a gene set enrichment analysis (GSEA) using the i-GSEA4GWAS (improved GSEA analysis for GWAS) web server (12). Among the top-ranked significant gene sets ($P < 0.001$), we identified potentially interesting pathways related to neurodegeneration (Alzheimer's disease), long-term potentiation, apoptosis and some immune-related pathways (Supplementary Material, Table S6). Most of the contributing association signals are individually weak ($P \sim 10^{-2}$ to 10^{-3}), and future studies will confirm if these are true association signals.

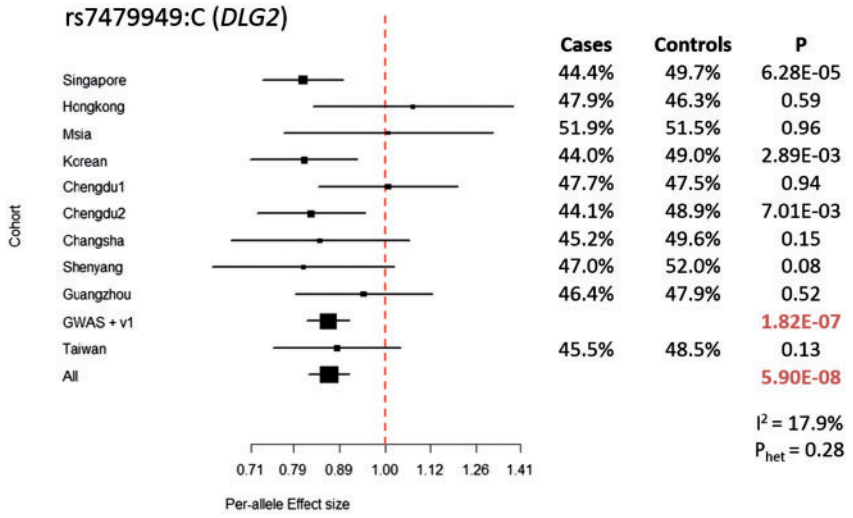

Figure 3. Forest plot showing odds ratios, allele frequencies and *P*-values in each sample collection at validated SNPs within known locus *DLG2*.

Discussion

We have conducted the largest East Asian GWAS on PD to date, analysing 22,729 subjects (5,125 PD cases and 17,604 controls) from six different countries across East Asia, namely Singapore, Hong Kong, Malaysia, Korea, mainland China and Taiwan. It is also the first comprehensive GWAS to be done on Han Chinese subjects at the discovery stage. We demonstrated strong association signals in *MCCC1*, *LRRK2* and *SNCA*, confirming these as the most important PD loci in both Asians and Europeans, and observed consistent association in 14 other previously-reported loci. Like Satake *et al.* (2) we did not observe any evidence for association across the *MAPT* locus (Supplementary Material, Fig. S3C) and the reported risk SNPs were non-polymorphic in our samples (Supplementary Material, Table S5), even though this locus showed one of the strongest associations in the European population (1). We did not identify any Asian-specific loci with large effects, e.g., odds ratios >1.3 with minor allele frequency >30% or odds ratios >1.45 for minor allele frequency >10% (Supplementary Material, Fig. S4) (13).

We observed evidence for allelic heterogeneity at *LRRK2*. We did not observe association at the *LRRK2* SNP rs76904798, which is independent ($r^2 < 0.2$) of the SNPs identified by us (including rs1384236) and by Satake et al (rs1994090) (1). We have previously performed direct genotyping of the Asian-specific *LRRK2* coding variant G2385R in a subset of the Singapore Chinese (discovery; 749 cases, 3925 controls) and Korean (first validation; 1078 cases, 665 controls) samples of this study (14). Consistent with what we have shown for rs1994090 (14), our top SNP rs1384236 is in moderate LD with G2385R in Chinese ($r^2 = 0.32$) and Koreans ($r^2 = 0.59$). Interestingly, rs1994090 and rs1384236 are in near-perfect LD in Koreans ($r^2 = 0.99$) and are also in high LD in the Chinese samples ($r^2 = 0.79$). Further studies comparing Chinese, Koreans and Japanese will help to dissect these subtle LD differences within the *LRRK2* locus among East Asian populations.

Overall, our results demonstrate the presence of some genetic heterogeneity in PD risk between Europeans and East Asians. Larger studies focused on this population may help to reveal novel loci that are Asian-specific or Han Chinese-specific. Furthermore, future trans-ethnic meta-analysis with the European samples will help to identify new loci and to fine-map shared loci, and may eventually contribute to our understanding of population-specific differences in the genetic risk underlying PD.

Materials and Methods

GWAS genotyping

We genotyped 834 PD cases and 50 controls on the Illumina Human OmniZhonghua8 v1A BeadChip for a total of 900,015 SNPs. For additional controls, we included samples from 13 other studies (Supplementary Material, Table S1) that were genotyped on different Illumina arrays (Human OmniZhonghua: 1,705, Human OmniExpress: 1,819, Human610-Quad: 6,113, Human 660W-Quad: 485, HumanHap 550-Quad: 1,029 and Human1M-Duo: 2,764). In total, we analysed 14,799 Han Chinese samples from Singapore that passed genotyping quality control (call rates >95%), comprising 834 PD cases and 13,965 controls. Patients were diagnosed with PD using the UK Brain Bank Criteria. This study was approved by the ethics committees or institutional review boards of the respective institutions.

Imputation and statistical analysis of GWAS samples

The software IMPUTE version 2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; date last accessed November 11, 2016) (15,16) was used for imputation of untyped SNPs in each dataset following pre-phasing using SHAPEIT (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html; date last accessed November 11, 2016) (17,18), and using the multi-ethnic 1000 genomes reference panel (March 2012) consisting of 36,648,992 autosomal SNP genotypes in 1,092 individuals from Africa, Asia, Europe, and the Americas (https://mathgen.stats.ox.ac.uk/impute/data_download_1000G_phase1_integrated.html; date last accessed November 11, 2016) (11) which has been shown to out-perform imputations with reference panels of matched ancestry. The imputation was run separately for each of the 14 study collections, after removing samples with call rates <95%, SNPs with call rates < 95%, minor allele frequencies (MAF) <1% and Hardy-Weinberg equilibrium (HWE) $P < 10^{-3}$ in

controls and/or $P < 10^{-6}$ in all samples as well as all non-autosomal SNPs (X, Y and mitochondrial chromosomes). The number of samples and SNPs used for imputation of each dataset is given in Supplementary Material, Table S1.

After imputation and merging, we performed identity-by-descent analysis using overlapping genotyped SNPs in PLINK (19) and first-degree relative pairs were identified; the relative with a lower sample call rate was excluded. We then ran a further stringent quality control filtering on the SNP level, setting SNPs with genotype probabilities less than 0.9 to missing and removing SNPs with a call rate < 98%, and further excluding those with MAF <1%, info score <0.8, HWE P in controls $<10^{-3}$, HWE P in all samples $< 10^{-6}$ and heterogeneity in controls $P < 10^{-6}$. We also ran a principal component analysis (20) on 42,125 independent genotyped SNPs ($r^2 < 0.2$, pruned using PLINK (19) –indep-pairwise 50 5 0.2) overlapping across all the different arrays after exclusion of SNPs in the five conserved long-range LD regions in Chinese as previously described (21). We then excluded outliers on the first five principal components. Eventually, 2,402,394 SNPs remained for analysis in 779 cases and 13,227 controls. We ran logistic regression analyses adjusting for the first three principal components (Supplementary Material, Fig. S1) using PLINK (19). The full set of summary association statistics can be requested for download at https://www.nni.com.sg/research/research-platforms/Genomics/Pages/Home.aspx; date last accessed November 11, 2016.

Pathway analysis was conducted using the i-GSEA4GWAS web server (http://gsea4gwas.psych.ac.cn/; date last accessed November 11, 2016) (12) with default parameters and our genome-wide SNP logistic regression P-values as input. Power calculations were performed using the Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/; date last accessed November 11, 2016) (13) and regional association plots were generated using LocusZoom (http://locuszoom.sph.umich.edu/locuszoom/; date last accessed November 11, 2016) (10).

Validation samples

Patients and ethnically-matched controls from nine independent centres across East Asia were recruited by the respective hospitals and study groups (Supplementary Material, Table S2). Patients were diagnosed with PD using the UK Brain Bank Criteria. Blood samples were collected from each participant and DNA extraction was performed.

Validation genotyping

Genotyping analysis of the 96 SNPs selected for validation was performed using the MassArray system from Sequenom Locus-specific PCR and detection primers were designed using the MassArray Assay Design 3.0 software (Sequenom). For the second validation of the four selected SNPs, pre-designed TaqMan assays were obtained from Applied Biosystems/Life Technologies. TaqMan reactions were carried out in 5-μl volumes containing 10–20 ng DNA according to the manufacturer's protocols. Fluorescence data were obtained in the ABI PRISM 7900HT, and SDS 2.4 software (Applied Biosystems) was used to call genotypes.

Meta-analysis of GWAS and validation samples

After removing samples with an SNP call rate of <90% from further analysis, all SNPs passed quality control filters (call rates >95%, HWE $P > 0.001$). For all SNP calls, we visually inspected the clustering patterns of genotypes from TaqMan and Sequenom assays and confirmed that the genotypes were of good quality. In each of the validation samples, we performed logistic regression analyses and combined the results using a fixed-effects inverse variance meta-analysis in PLINK (19).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank Yen-Yek Ng, Ebonne Ng (National Neuroscience Institute) and the Genotyping platform at the Genome Institute of Singapore for their assistance with this project.

Conflict of Interest statement. None declared.

Funding

This work is supported by the National Medical Research Council under the Singapore Translational Research Investigator Award (STaR) award (to EK Tan) and Translational and Clinical Research Flagship Programme in Parkinson's disease (NMRC/TCR/013-NNI/2014) and Schizophrenia (NMRC/TCR/003/2008), Agency for Science, Technology and Research, Duke-NUS Graduate Medical School, and Singapore Millennium Foundation. JN Foo is a Singapore National Research Foundation Fellow (NRF-NRFF2016-03). SY Lim and AH Tan received funding support from the Malaysian Ministry of Higher Education grant for High-Impact Research (HIR) UM.C/625/1/HIR/MOHE/CHAN/11/03. YR Wu and CM Chen received funding support from grants 102-2314-B-182A-113-MY2 from the Taiwan Ministry of Science and Technology. K Song is supported by a Mid-career Researcher Program grant through the National Research Foundation of Korea (2014R1A2A1A09005824), funded by the Ministry of Science, Information & Communication Technology and Future Planning, the Republic of Korea. The funding sources had no role in the design and conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript.

References

1. Nalls, M.A., Pankratz, N., Lill, C.M., Do, C.B., Hernandez, D.G., Saad, M., DeStefano, A.L., Kara, E., Bras, J., Sharma, M., *et al.* (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet*, **46**, 989–993.
2. Satake, W., Nakabayashi, Y., Mizuta, I., Hirota, Y., Ito, C., Kubo, M., Kawaguchi, T., Tsunoda, T., Watanabe, M., Takeda, A., *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.
3. Simon-Sanchez, J., Schulte, C., Bras, J.M., Sharma, M., Gibbs, J.R., Berg, D., Paisan-Ruiz, C., Lichtner, P., Scholz, S.W., Hernandez, D.G., *et al.* (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.*, **41**, 1308–1312.
4. Nalls, M.A., Plagnol, V., Hernandez, D.G., Sharma, M., Sheerin, U.M., Saad, M., Simon-Sanchez, J., Schulte, C., Lesage, S., Sveinbjornsdottir, S., *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.
5. International Parkinson's Disease Genomics Consortium (IPDGC); Wellcome Trust Case Control Consortium 2 (WTCCC2). (2011) A two-stage meta-analysis identifies several new loci for Parkinson's disease. *PLoS. Genet.*, **7**, e1002142.
6. Lill, C.M., Roehr, J.T., McQueen, M.B., Kavvoura, F.K., Bagade, S., Schjeide, B.M., Schjeide, L.M., Meissner, E., Zauft, U., Allen, N.C., *et al.* (2012) Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS. Genet.*, **8**,. e1002548.
7. Pankratz, N., Beecham, G.W., DeStefano, A.L., Dawson, T.M., Doheny, K.F., Factor, S.A., Hamza, T.H., Hung, A.Y., Hyman, B.T., Ivinson, A.J., *et al.* (2012) Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. *Ann. Neurol.*, **71**, 370–384.
8. Do, C.B., Tung, J.Y., Dorfman, E., Kiefer, A.K., Drabant, E.M., Francke, U., Mountain, J.L., Goldman, S.M., Tanner, C.M., Langston, J.W., *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.
9. Hamza, T.H., Zabetian, C.P., Tenesa, A., Laederach, A., Montimurro, J., Yearout, D., Kay, D.M., Doheny, K.F., Paschall, J., Pugh, E., *et al.* (2011) Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat. Genet.*, **42**, 781–785.
10. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R. and Willer, C.J. (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26**, 2336–2337.
11. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A. and Abecasis, G.R. 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature*, **526**, 68–74.
12. Zhang, K., Cui, S., Chang, S., Zhang, L. and Wang, J. (2010) i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. *Nucleic Acid Res.*, **38**, W90–W95.
13. Purcell, S., Cherny, S.S. and Sham, P.C. (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, **19**, 149–150.
14. Foo, J.N., Chung, S.J., Tan, L.C., Liany, H., Ryu, H.S., Hong, M., Koh, T.H., Irwan, I.D., Au, W.L., Prakash, K.M., *et al.* (2016) Linking a genome-wide association study signal to a LRRK2 coding variant in Parkinson's disease. *Mov. Disord.*, **31**, 484–487.
15. Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS. Genet.*, **5**, e1000529.
16. Howie, B., Marchini, J. and Stephens, M. (2011) Genotype imputation with thousands of genomes. *G3 (Bethesda)*, **1**, 457–470.
17. Delaneau, O., Marchini, J. and Zagury, J.F. (2012) A linear complexity phasing method for thousands of genomes. *Nat. Methods*, **9**, 179–181.
18. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. and Abecasis, G.R. (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.*, **44**, 955–959.
19. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. and Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
20. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
21. Chen, J., Zheng, H., Bei, J.X., Sun, L., Jia, W.H., Li, T., Zhang, F., Seielstad, M., Zeng, Y.X., Zhang, X. and Liu, J. (2009) Genetic structure of the Han Chinese population revealed by genome-wide SNP variation. *Am. J. Hum. Genet.*, **85**, 775–785.