Supplementary Methods

1. Sample recruitment

Parkinson's disease (PD) patients were recruited from the movement disorders clinics of the National Neuroscience Institute at the Singapore General Hospital and Tan Tock Seng Hospital Campuses. Patients were diagnosed with PD using the UK Brain Bank Criteria. Population controls were recruited from various centres across Singapore (Foo et al., 2014). All participants gave informed written consent and the study received approval from the institutional ethics committee (SingHealth and Nanyang Technological University, Singapore).

2. Exome array genotyping

We recruited a total of 834 cases and 9,094 controls of Chinese ethnicity from Singapore which were genotyped successfully on the Illumina HumanExome BeadChip Exome_Asian array. We excluded samples based on the following criteria: (i) first degree relative of other genotyped samples with the lower sample call rate, (ii) outliers on the first five principal components, (iii) have high heterozygosity and >5% missing genotypes and iv) the early-onset PD samples which have been exome-sequenced for the analysis of coding variants, as described below. Following quality control filtering, the remaining 710 PD cases (later onset > 55 years; age at sample collection: 69.4 ± 9.2 years, age of onset 65.3 ± 9.3 years; 56.1% males) and 9,046 population controls (age at sample collection: 50.5 ± 13.2 years, 55.9% males) were used in the downstream analysis. We analyzed 43 reported variants as well as 146 common non-synonymous coding variants (MAF>1%) in the 33 genes that were genotyped as part of the Illumina exome array. All the 710 PD cases were verified to be genetically matched to the controls by principal components analysis (Foo et al., 2014). Given the reported prevalence of PD in Singapore Chinese subjects,

we expect that not more than 1-2% of randomly selected population controls will develop PD in their lifetime, hence the low misclassification bias of < 2% is expected to have a much lower reduction of statistical power compared to the gain in statistical power from having a large population control for this analysis (WTCCC, 2007). Single variant analyses were done using PLINK.

3. Whole exome sequencing

We selected a total of 104 of the earliest-onset PD cases and 104 healthy elderly controls from Singapore for the analysis of coding variants by whole exome sequencing. 99 PD cases (age-of-onset ≤ 55 years, aged 54.7 ± 7.2 years, 67% male, onset 48.3 ± 5.8 years, 60 with onset ≤ 50 years, 23% with affected relatives) and 99 healthy elderly controls free of neurodegenerative diseases (aged 71.9 ± 10.8 years, 58% male) remained following the removal of first degree relatives, samples which were determined to be non-Chinese in ethnicity, and samples with gender discrepancy. We analyzed all coding variants of the 33 PD-causing genes in these 198 Chinese PD patients and controls. Given the relatively small number of early-onset PD patients analysed, we matched early-onset PD patients to elderly controls to avoid the inclusion of controls which were misclassified since the presence of early onset PD cases among these controls will have a large impact on statistical power. Rare non-synonymous coding variants (MAF ≤ 1% and present at < 1% frequency in 1000 Genomes, HapMap and ExAC populations) within these 33 PD genes were identified from whole exome sequencing data. SIFT algorithm (Sim et al., 2012) was used to predict effect of each missense variant. For gene-based tests, we compared the number of carriers of such variants between cases and controls using two-tailed Fisher's exact tests.

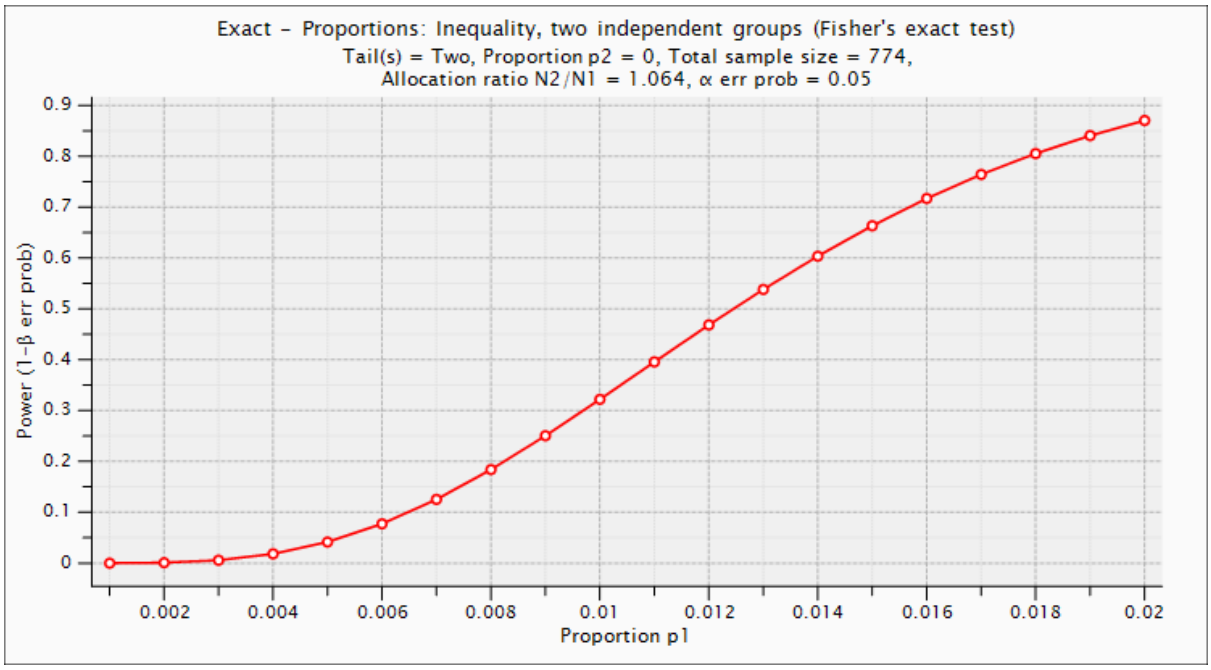
4. Targeted sequencing

We further validated candidate variants in *CD36* by targeted sequencing of 576 ethnic Chinese and Korean samples collected from Hong Kong, Malaysia, South Korea, which included 276 early-onset PD cases (96 Chinese: age at sample collection: 53.9 ± 8.1 years, age of onset 42.7 ± 6.6 years; 65.3% males, 180 Koreans: age at sample collection: 52.1 ± 8.4 years, age of onset 41.4 ± 4.7 years; 47.8% males) and 300 population controls (120 Chinese: age at sample collection: 69.5 ± 3.3 years, 65.8% males, 180 Koreans: age at sample collection: 75.6 ± 3.5 years, 64.4% males).

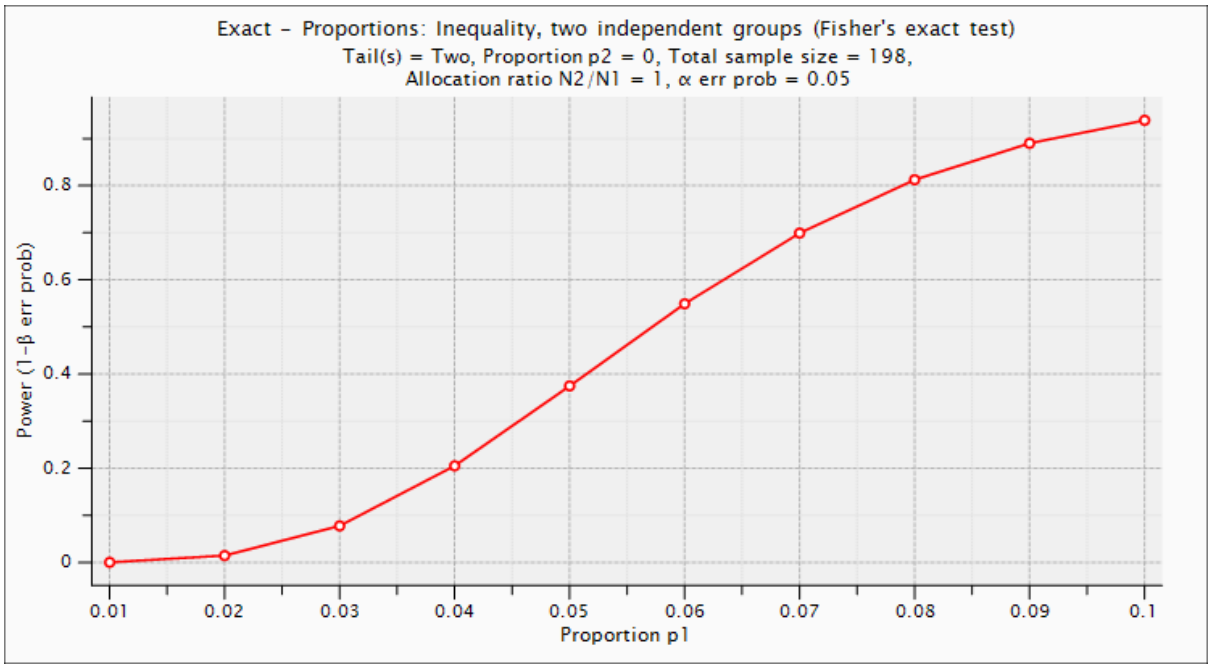
5. Power calculations

To assess the power we had to detect the reported novel PD variants in our Chinese PD cases, we performed power calculations using Fisher's exact test with post-hoc power analysis for $\alpha = 0.05$ (G*Power, Version 3.1.9.2 (Faul et al., 2009; Faul et al., 2007)). Minor allele frequencies used for power calculations were based on number of instances each mutation was reported in cases/controls by Siitonen et. al. and Jansen et. al..

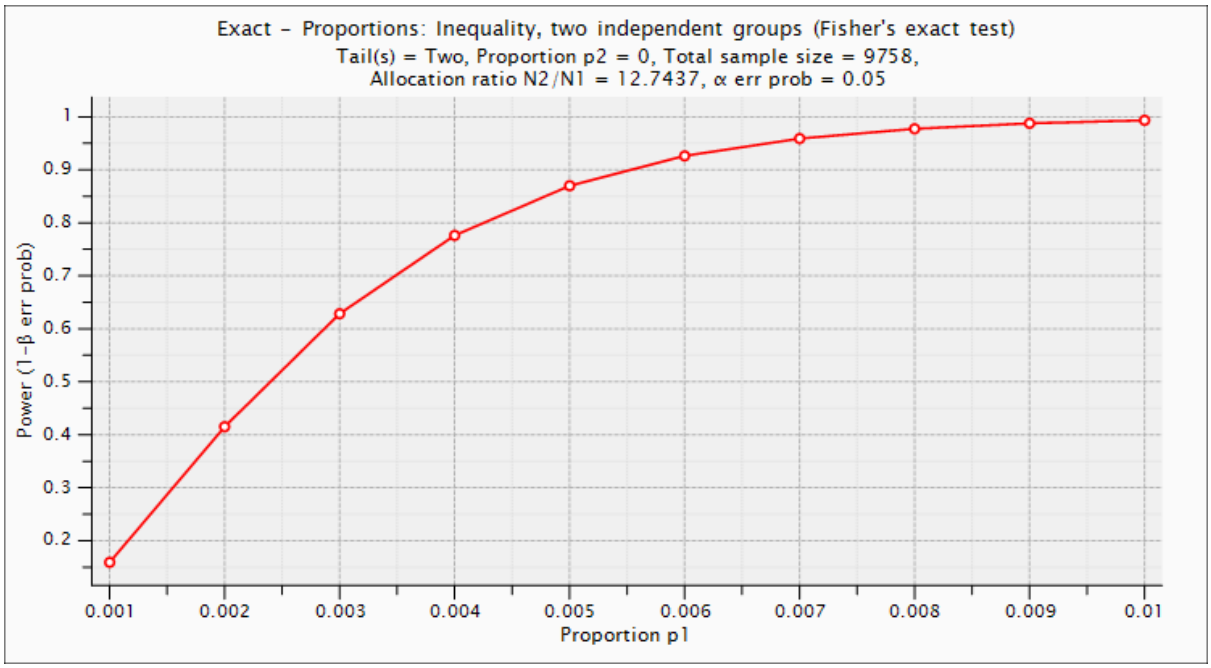
Supplementary Figures



Supplementary Figure 1: Plot of estimated statistical power for detection of rare variants in *CD36* present in proportion p1 of 375 early-onset patients, assuming near-complete penetrance (p2 ~0 in 399 controls) from Singapore, Hong Kong, Malaysia and Korea. We estimate we have 80% power (at α=0.05) to identify highly-penetrant *CD36* rare variants in at least 1.8% of our early-onset cases.



Supplementary Figure 2: Plot of estimated statistical power for detection of rare variants present in proportion p1 of 99 early-onset patients, assuming near-complete penetrance (p2 ~0 in 99 controls).



Supplementary Figure 3: Plot of estimated statistical power for detection of rare variants present in proportion p1 of 710 later-onset patients, assuming near-complete penetrance (p2 ~0 in 9046 controls).

References

Faul, F., Erdfelder, E., Buchner, A., Lang, A.G., 2009. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behavior research methods 41(4), 1149-1160.

Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods 39(2), 175-191.

Foo, J.N., Tan, L.C., Liany, H., Koh, T.H., Irwan, I.D., Ng, Y.Y., Ahmad-Annuar, A., Au, W.L., Aung, T., Chan, A.Y., Chong, S.A., Chung, S.J., Jung, Y., Khor, C.C., Kim, J., Lee, J., Lim, S.Y., Mok, V., Prakash, K.M., Song, K., Tai, E.S., Vithana, E.N., Wong, T.Y., Tan, E.K., Liu, J., 2014. Analysis of non-synonymous-coding variants of Parkinson's disease-related pathogenic and susceptibility genes in East Asian populations. Hum Mol Genet 23(14), 3891-3897.

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., Ng, P.C., 2012. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research 40(W1), W452-W457.

WTCCC, 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447(7145), 661-678.