Analysis of Genetic Mutations in the BioFIND Cohort

BioFind Cohort Overview

The Fox Investigation for New Discovery of Biomarkers (BioFIND) is a cross-sectional study sponsored by MJFF with support from the National Institute of Neurological Disorders and Stroke (NINDS). Clinical data and biological specimens were collected at 8 sites across the United States. Biospecimen collection was performed at baseline and samples were drawn within 1-3 hours of most recent PD medication administration. Biospecimens collected when subjects were off their PD medications are available from a single follow-up visit 14 days later. This cohort contains 122 well-defined, moderately advanced PD patients and 101 healthy controls. Available de-identified data include clinical motor data, clinical non-motor (cognitive, neurobehavioral, neuropsychological, autonomic, sleep), as well as biologic data including spinal fluid alpha-synuclein, amyloid-beta, tau, and phosphorylated tau levels. Available biospecimens include plasma, DNA and RNA from blood, whole blood pellet, cerebrospinal fluid, urine, and saliva. More information on the BioFIND cohort can be found at https://www.michaeljfox.org/biospecimens.

Genotyping Methods

Study 1 (Project ID 108 & 124)

PI: Dena Hernandez (NIA)

Method: Illumina Human Omni Express Exome+ v1.3 NeuroX array

The OmniExpress NeuroX array is an Illumina Infinium iSelect HD Custom Genotyping array containing 961,000 markers including 267,607 Illumina standard content exonic variants, 635,000 tagged SNPs from HapMap and an additional 24,706 custom variants designed for neurological disease studies. Of the custom variants, approximately 12,000 are designed to study Parkinson's disease and are applicable to both large population studies of risk factors and to investigations of familial disease and known mutations. Genotyping was performed per manufacturers protocols. The Genotyping Analysis Module within Genome Studio version 1.9.4 was used to analyze data. The threshold call rate for sample inclusion was 95%. The genetic variants in table 1 were directly typed using the NeuroX array. This list includes the data dictionary description for each variant; note this field includes the ancestral and minor alleles as defined by dbSNP as of April 10th 2014 (build creation 123, build update 138.

DATA DICTIONARY ENTRY	GENE	OTHER
rs114138760 C/G (FWD) G:Ancestral C:Minor	*GBA/SYT11*	rs114138760_C
rs76763715 C/T (FWD) T:Ancestral C:Minor	*GBA*	rs76763715_GBA_p.N370S_C
rs71628662 C/T (FWD) T:Ancestral C:Minor	*GBA/SYT11*	rs71628662_C
rs823118 C/T (FWD) C:Ancestral T:Minor	*RAB7L1*	rs823118_C
rs10797576 C/T (FWD) C:Ancestral T:Minor	*SIP A1L2*	rs10797576_T
rs6430538 C/T (FWD) T:Ancestral C:Minor	*ACMSD/TMEM163*	rs6430538_T
rs1955337 G/T (FWD) G:Ancestral T:Minor	*STK39*	rs1955337_T
rs12637471 A/G (FWD) G:Ancestral A:Minor	*MCCC1*	rs12637471_A
rs34884217 (G/T) REV T:Ancestral C:Minor	GAK	rs34884217_G
rs34311866 A/G (REV) A:Ancestral C:Minor	*GAK*	rs34311866_G
rs11724635 A/C (FWD) A:Ancestral A:Minor	*BST1*	rs11724635_C
rs6812193 C/T (FWD) C:Ancestral T:Minor	*F AM47E/SCARB2*	rs6812193_T
rs356181 C/T (REV) T:Ancestral A:Minor	*SNCA*	rs356181_C
rs3910105 C/T (REV) T:Ancestral G:Minor	*SNCA*	rs3910105_C
rs8192591 A/G (REV) G:Ancestral T:Minor	*HLA*	rs8192591_T
rs9275326 (was rs115462410) C/T (FWD) C:Ancestral T:Minor	*HLA*	rs115462410_T
rs199347 C/T (REV) C:Ancestral G:Minor	*GPNMB*	rs199347_C
rs591323 A/G (FWD) G:Ancestral A:Minor	*FGF20*	rs591323_A
rs118117788 C/T (FWD) C:Ancestral T:Minor	*INPP5F*	rs118117788_T
rs329648 C/T (FWD) T:Ancestral T:Minor	*MIR4697*	rs329648_T
rs76904798 C/T (FWD) T:Ancestral T:Minor	*LRRK2*	rs76904798_T
rs34995376 A/G (FWD) G:Ancestral A:Minor	*LRRK2*	rs34995376_LRRK2_p.R1441H_A
rs35801418 A/G (FWD) A:Ancestral G:Minor	*LRRK2*	rs35801418_LRRK2_p.Y1699C_G
rs34637584 A/G (FWD) G:Ancestral A:Minor	*LRRK2*	rs34637584_LRRK2_p.G2019S_A
rs35870237 C/T (FWD) T:Ancestral C:Minor	*LRRK2*	rs35870237_LRRK2_p.I2020T_C
rs11060180 A/G (FWD) A:Ancestral G:Minor	*CCDC62*	rs11060180_G
rs11158026 C/T (FWD) T:Ancestral T:Minor	*GCH1*	rs11158026_T
rs2414739 A/G (FWD) G:Ancestral G:Minor	*VPS13C*	rs2414739_G
rs14235 A/G (FWD) G:Ancestral A:Minor	*BCKDK/ STX1B*	rs14235_A
rs11868035 A/G (FWD) G:Ancestral A:Minor	*SREBF/RAI1*	rs11868035_A
rs17649553 C/T (FWD) T:Ancestral T:Minor	*MAPT*	rs17649553_T
rs12456492 A/G (FWD) G:Ancestral G:Minor	*RIT2*	rs12456492_G
rs55785911 A/G (FWD) G:Ancestral A:Minor	*DDRGK1*	rs55785911_A

Quality control of sample handling was determined by comparing the subject's sex reported by Coriell Institute for Medical Research with the genotypic sex estimated from X chromosome heterogeneity. X chromosome heterogeneity calculations were based on common SNPs from the International HapMap Project that had genotypes with missingness <5% and hardy-Weinberg equilibrium p values >1E-5. Samples considered heterozygosity outliers (>±3sd from the sampled mean) or with discrepancies between reported sex and genotypic estimated sex were excluded.

Study 2 (Project ID 114)

PI: Andrea Dardis & Stefania Zampieri (University Hospital Santa Maria della Misericordia)

Method: Parallel End Pair Sequencing via Illumina MiSeq

The whole genomic sequence of the *GBA1* gene (GenBank J03059.1) was analyzed by massive parallel pair end sequencing. For each sample, the whole genomic *GBA1* sequence was PCR--amplified in two overlapped fragments of 2870bp and 4492bp, using primers designed to selectively amplify the gene and not the homologous pseudogene. Pair end libraries were generated using a Nextera XT DNA sample preparation kit (Illumina) and sequenced on an Illumina MiSeq platform. Pair end reads were mapped to the reference human genome (hg19, Chr 1 155211205-155204618; Amplicon 1: 155211205-155208234; Amplicon 2: 155209110-155204618). Only samples with a sequencing quality control score ≥30 and with a minimal read depth of 200X were taking into consideration for variant analysis. All SNV and indel information are output in variant call format. Annotation of SNV was performed with wANNOVAR [http://wannovar.usc.edu/]. Exonic variants were confirmed by Sanger sequencing and the presence of two variants in the same allele was confirmed by cloning and sequencing the PCR product (containing both variants). The possible presence of the RecΔ55allele (non-detectable by NGS) has been ruled out by Sanger sequencing of exon 9.

Mutation Carriers

Below you will find a summary of the mutations detected in the BioFIND cohort along with patient IDs corresponding to the samples.

Mutation	Patient ID
GBA	
GBA N370S	BF-1008
GBA N370S	BF-1156
GBA L483P	BF-1016
GBA R296Q	BF-1021
GBA S235P/A495P	BF-1080
GBA A495P	BF-1128
GBA R324H	BF-1199
GBA L29Afs*18	BF-1236
LRRK2	
LRRK2 G2019S	BF-1108
LRRK2 G2019S	BF-1198
LRRK2 R1441C	BF-1222
DJ-1	
c.-24+2T>A	BF-1044
Parkin	
Parkin R275W	BF-1257

More Information

Full datasets and methods descriptions for the BioFIND cohort are available at https://biofind.loni.usc.edu/.

Last Updated June 2023