Using global team science to identify genetic parkinson's disease worldwide

Eva-Juliane Vollstedt, Katja Lohmann, Harutyun Madoev, Shalini Padnamabhan, Anna Naito, Sonja Petkovic, Jan O. Aasly, Charles H. Adler, Azlina Ahmad-Annuar, Bashayer R. Al-Mubarak, Nada Al Tassan, Alberto Albanese, Roy N. Alcalay, Victoria Álvarez , Hassan Anhar, Grazia Annesi, Silke Appel-Cresswell, Gonzalo Arboleda, David Arkadir, Thomas R. Barber, Soraya Bardien, Melinda Barkhuizen, Matthew J. Barrett, Bruno A. Benitez, Kailash P. Bhatia, Vincenzo Bonifati, Maria Bozi, Alexis Brice, Laura Brighina, Kathrin Brockmann, Francisco Cardoso, Andrea Carmine Belin, Jonathan Carr, Piu Chan, Bruce Chase, Alice Chen-Plotkin, Sun J. Chung, Jordi Clarimon, Lorraine N. Clark, Jean-Christophe Corvol, Patrick Cras, David Crosiers, Carlos Cruchaga, Nir Giladi, Joana Damásio, Parimal Das, Patricia de Carvalho Aguiar, Giuseppe De Michele, Jolanta Dorszewska, Sevda Erer, Matthew J. Farrer, Rosangela Ferese, Ondrej Fiala, Tatiana Foroud, Andrzej Friedman, Manabu Funayama, Gaetan Garraux, Emilia M. Gatto, Gençer Genç, Stefano Goldwurm, Juan Carlos Gómez-Esteban, Ji-feng Guo, Hasmet A. Hanagasi, Nobutaka Hattori, Robert Hauser, Peter Hedera, Faycal Hentati, Jens Michael Hertz, Janice L. Holton, Michele Hu, Takeshi Ikeuchi, Sergey Illarioshkin, Jon Infante, Joseph Jankovic, Beom Seok Jeon, Hiroshi Kataoka, Hideshi Kawakami, Yun Joong Kim, Péter Klivényi, Sullen Koks, Vladimir S. Kostic, Anna Krygowska-Wajs, Anthony E. Lang, Mark S. LeDoux, Shen-Yang Lim, Chin-Hsien Lin, Chin-Song Lu, Tim Lynch, Maciej Machaczka, Kari Majamaa, Katerina Markopoulou, Karen Marder, Mika H. Martikainen, Ignacio F. Mata, George D. Mellick, Pablo Mir, Renato Puppi Munhoz, Diana Olszewska, Coro Paisán-Ruiz, Haydeh Payami, Joel Perlmutter, Márcia M. G. Pimentel, Peter P. Pramstaller, Teeratorn Pulkes, Andreas Puschmann, Gerhard Ransmayr, Olaf Riess, Owen Ross, Federico Rodriguez-Porcel, Ekaterina Rogaeva, Javier Ruiz-Martinez, Esther M. Sammler, Marta San Luciano, Wataru Satake, Rachel Saunders-Pullman, Ali Sazci, Clemens R. Scherzer, Anette Schrag, Manu Sharma, Ellen Sidransky, Gabriella Silvestri, Andrew B. Singleton, Maria Skaalum Petersen, Mariana Spitz, Leonidas Stefanis, Carolyn M. Sue, Matthew Swan, Ricardo Taipa, Eng-King Tan, Bei-Sha Tang, Avner Thaler, Astrid Thomas, Tatsushi Toda, Mathias Toft, Eduardo Tolosa, Enza Maria Valente, Christine Van Broeckhoven, Laszlo Vecsei, Tom Warner, Caroline Williams-Gray, Juliane Winkelmann, Dirk Woitalla, Zbigniew K. Wszolek, Ruey-Meei Wu, Yih-Ru Wu, Tao Xie, Bao-rong Zhang, Alexander Zimprich, Rejko Krüger, Matthew Farrer, Inke R. König, Meike Kasten, Christine Klein

Key points

Question

Can global team science help to overcome difficulties posed by rare disease subtypes such as genetic Parkinson's disease?

Findings

A total of 8819 PD patients with mutations in PD genes were reported by 128 researchers from 117 centers in 41 countries. The number of reported monogenic PD patients exceeds the number of published cases by a factor of 2.8. Likewise, completeness of individual data is markedly higher than published information. Willingness to collaborate is high (98%).

Meaning

Novel ways of team science including large numbers of researchers worldwide with shared interests and similar areas of expertise are both necessary and successful when it comes to addressing the challenges posed by rare diseases, an approach that can be readily extended to similar clinical and research questions requiring a worldwide contribution of the scientific community.

Abstract

Importance: Rare diseases challenge the international research community with small sample sizes at individual sites calling for novel ways of worldwide team science.

Objective: To identify a worldwide cohort of patients and families with genetic Parkinson's disease (PD) as an example of a new, comprehensive approach to global collaboration, adapted from team science approaches, and to thereby address the challenge that rare diseases pose to clinicians, researchers and patients (and caregivers).

Design: In 2018, we conducted a worldwide, systematic online survey on the availability of demographic, clinical, genetic, and additional data of patients with genetic PD due to *SNCA, LRRK2, VPS35, Parkin, PINK1, DJ1,* and *GBA* mutations.

Setting: The study included researchers worldwide.

Participants: Researchers were identified extracting all corresponding authors of articles represented in the "Movement Disorders Society Genetic mutation database" (MDSGene) and through the "Genetic Epidemiology of Parkinson's disease" (GEoPD) consortium and invited to participate in the build-up of this worldwide cohort.

Main Outcome(s) and Measure(s): The main outcome was the number of patients with genetic PD globally available. Other outcomes include the availability of demographic, clinical, and genetic data, as well as biospecimens and the researchers' interest in further collaboration.

Results: In total, we identified 354 researchers at 117 international sites in 41 countries to participate in our study, 162 (50%) of whom completed the survey. They reported a total of 8819 PD patients of >9 ethnicities with mutations in SNCA (n=274), LRRK2 (n=3231), VPS35 (n=39), Parkin (n=1733), PINK1 (n=286) or DJ1 (n=32), and risk alleles in GBA (n=3224). Demographic data was reported to be available by most researchers (>93%), availability of clinical data ranged from 59% (nonmotor signs) to 91% (age at onset); biospecimens were available from 7% (cerebrospinal fluid) to 80% (DNA). 98% of the researchers indicated their interest in further collaboration.

Conclusions and Relevance: The number of patients with genetic PD included almost triples the number of cases reported in the literature, and data availability, by far exceeds currently published clinical information, indicating that accessing patients and data for rare diseases requires novel approaches and ways of communication. The overwhelmingly positive response and willingness to collaborate impressively highlights the relevance and power of team science.

Introduction

Rapidly advancing sequencing technologies offer new and cost-effective approaches to identify genetic causes of rare diseases and increasingly define genetic subtypes of common diseases. Parkinson's disease (PD) is a common disease but can be genetically stratified into subgroups of patients with rare genetic forms of PD due to mutations in *SNCA, LRRK2, VPS35, Parkin, PINK1, DJ1,* and *GBA*, the latter as the strongest known genetic risk factor. While studying small numbers of patients may suffice to elucidate the pathophysiology of rare disease, establishing accurate genotype-phenotype relationships and conducting significant clinical studies and trials requires large sample sizes and thus pose a major challenge on investigating relatively rare diseases. Additional layers of complexity are added as clinical and genetic data are often reported in an inconsistent and incomplete fashion, sometimes due to reporting and naming habits rather than the availability of this data to the reporting author. Clinical expression and penetrance of gene mutations may vary considerably across different populations and ethnicities (1). Illustrating the magnitude of the problem, up to 300 million patients are estimated to suffer from a rare disease worldwide (2) including ~300,000 patients with genetic forms of PD, representing 5% of an estimated 6 million patients with PD in 2018 (3). Clinical information is reported in the international medical literature, published in English, but only for a fraction of these cases and with a bias towards unusual cases, i.e. 1769 patients with monogenic PD (Pubmed and Movement Disorder Society Genetic mutation database; MDSGene, www.mdsgene.org). As gene-specific therapies are increasingly being developed for PD with several frontrunners now in clinical trials(4)(5), identifying patients and families in whom disease is driven by genetic factors is imperative.

Since the 1990s, there has been a growing interest and investment in large-scale, team-based research initiatives to address complex and multifaceted problems that require collaboration across different disciplines (6). Likewise, there is a growing necessity for – ideally global-scale - team science approaches of clinicians and researchers with *similar* interests joining forces to tackle a global task. Employing novel ways of team science, electronic databasing, and global communication, we performed a worldwide survey of genetic PD with an emphasis on availability of demographic, clinical and biomarker data and materials, to foster global collaboration.

Methods

Selection of participants

In order to identify possible participants for our survey, we compiled the names of corresponding authors from articles included in the Movement Disorders Society Genetic mutation database (MDSGene, see www.mdsgene.org). Six causes of monogenic PD were included: *SNCA*, *LRRK2*, *VPS35, Parkin*, *PINK1*, and *DJ1*. When the project was conceived *GBA* had yet to be included in the MDSGene database. Hence, we screened literature according to MDSGene criteria to identify corresponding authors of eligible articles (articles published in English with clinical information available). In addition, the "Genetic Epidemiology of Parkinson's disease" consortium (https://geopd.lcsb.uni.lu) contributed members not already identified as a corresponding author of publications represented in MDSGene. During the survey phase of the project, additional contacts were included upon recommendation of participants (see Figure S1).

Data collection

We developed an online survey on demographic, clinical, genetic, and additional data (for details, see Table S1) and invited the previously identified researchers to report availability of information in their samples of genetic PD patients focusing on *SNCA, LRRK2, VPS35, Parkin, PINK1, DJ1,* and *GBA*. To avoid multiple reporting of the same cases, we asked participants to indicate sharing of samples and encouraged participants from the same center to nominate one person to report all cases or to divide up the cases between reporting researchers. Two rounds of email reminders were sent out after two weeks each to enhance the response rate. The duration of the survey was 6 weeks. Participants received no financial compensation for their contribution.

Results

A total of 8819 PD patients with mutations in PD genes were reported by 128 participants of the survey from 117 centers in 41 countries (Figure 1). Samples from participating researchers include diverse ethnicities: Caucasian (n=85 researchers), Asian (n=26), Ashkenazi Jewish (n=17), African (n=10), Hispanic (n=9), Arab (n=5), Indian (n=5), North African (n=5), Mennonites (n=3), and Others (n=11). Patients were reported to carry mutations in *SNCA* (n=274), *LRRK2* (n=3231), *VPS35* (n=39), *Parkin* (n=1733), *PINK1* (n=286), *DJ1* (n=32), or GBA risk alleles (n=3224). The response rate was 50% (n=162) and 79% (n=128) of the respondents completed the survey (see Figure S1). Of these, 98% (n=125) indicated their interest in further collaborations, and 45% (n=57) participants sent personal emails expressing their interest to contribute further to this project and/or suggesting additional collaborative projects (Figure 2). Demographic data was available for >93% of participating researchers, while the availability of clinical data ranged from basic information (91% for age at onset) to more specific items (59% for nonmotor signs, see Table S1). Additional information on environment and lifestyle was less commonly noted but still

available for about a third of the cases (39% and 32%, respectively). Most researchers reported to have DNA (80%), whereas other blood-derived biospecimens (RNA, Serum, Plasma, whole blood) were reported to be available by <25% of the researchers. Samples that are more difficult to obtain (cerebrospinal fluid, brain tissue, fibroblasts, induced pluripotent stem cells) were even more rare (<11%).

Discussion

The number of reported monogenic PD cases in the present study exceeds published cases by a factor of 2.8 (n=5845 versus n=2083 cases, both excluding GBA), indicating a large number of patients with genetic PD is not reflected in the literature. This worldwide "publication gap" further suggests that there is a large number of well-characterized genetically defined patients readily available but not yet included in research projects or clinical trials. Of particular note, our study suggested an unprecedented level of data completeness, a paramount prerequisite when aiming to perform meaningful genotype-phenotype correlations and when selecting patients for clinical studies and trials.

To our knowledge, the MDSGene database currently represents the only systematic, comprehensive and fully curated collection of genetic and clinical information on genetic PD. It is also based on a global effort currently including >70 movement disorder specialists and geneticists working on data extraction from worldwide literature. However, the data gaps are considerable. Especially information on nonmotor signs is scarce in the literature, even for cognition, information is available only for about a third (660/2064) of the cases with monogenic PD (www.mdsgene.org). In contrast, our present approach indicates availability of nonmotor signs in general for about three quarters (6592/8787) of the reported cases. Given the increasing difficulty to publish clinical-genetic case reports or series, along with broad(er) access to diagnostic genetic testing, we predict this trend of, first, non-reporting of

genetic PD patients and, second, of not publishing data at the individual patient level to quickly grow.

Based on the enthusiastic responses and willingness to collaborate, we are confident that we have successfully established an international collaboration that will enable (i) clinicogenetic studies to address comprehensive genotype-related characterization, (ii) performing modifier screens with regard to expressivity and penetrance of genetic PD, and (iii) conducting well-informed, large-scale clinical trials for these rare subtypes of PD, each of which is likely amenable to individualized, gene-specific treatment approaches. We further conclude that novel team science including large numbers of researchers worldwide with shared interests and similar areas of expertise are both necessary and successful when it comes to addressing the challenges posed by rare diseases, an approach that can be readily extended to similar clinical and research questions requiring a worldwide contribution of the scientific community.

References

1. Kasten M, Hartmann C, Hampf J, Schaake S, Westenberger A, Vollstedt E-J, et al. Genotype-phenotype relations for the Parkinson's Disease genes Parkin , PINK1 , DJ1: MDSGene Systematic Review. Mov Disord. 2018 Apr 11;

2. Gammie T, Lu CY, Babar ZU-D. Access to Orphan Drugs: A Comprehensive Review of Legislations, Regulations and Policies in 35 Countries. Garattini S, editor. PLoS One. 2015 Oct 9;10(10):e0140002.

3. Dorsey ER, Constantinescu R, Thompson JP, Biglan KM, Holloway RG, Kieburtz K, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. Neurology. 2007;68(5):384–6.

4.   A Global Study to Assess the Drug Dynamics, Efficacy, and Safety of GZ/SAR402671 in

        Parkinson's Disease Patients Carrying a Glucocerebrosidase (GBA) Gene Mutation

        (MOVES-PD) [Internet]. Available from:

        https://clinicaltrials.gov/ct2/show/NCT02906020

5.   Olanow CW, Kordower JH. Targeting α-Synuclein as a therapy for Parkinson's disease:

        The battle begins. Mov Disord. 2017 Feb;32(2):203–7.

6.   Stokols D, Hall KL, Taylor BK, Moser RP. The Science of Team Science. Am J Prev Med.

        2008 Aug;35(2):S77–89.

Figure 1: Worldwide centers reporting genetic PD patients (incl. GBA as a risk variant)
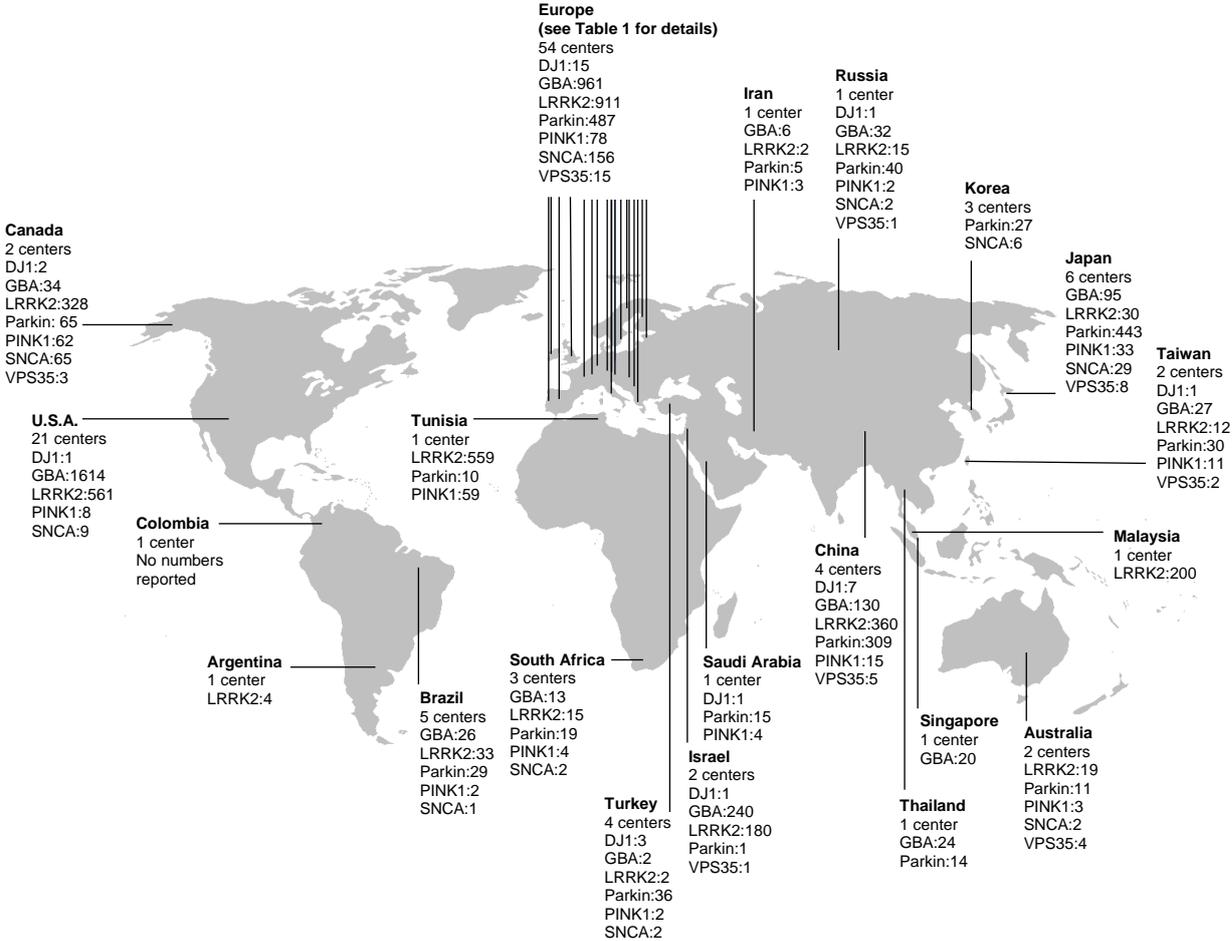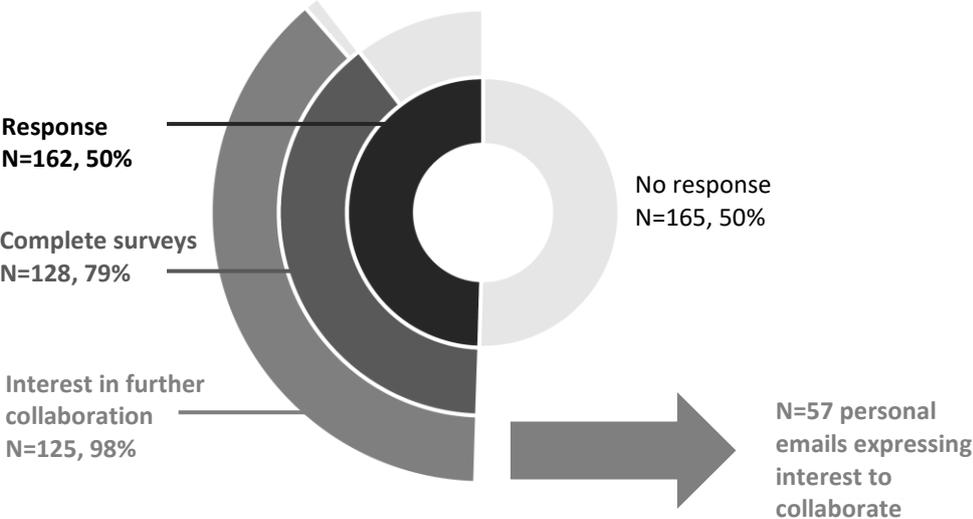
Europe
(see Table 1 for details)
54 centers
DJ1:15
GBA:961
LRRK2:911
Parkin:487
PINK1:78
SNCA:156
VPS35:15

Iran
1 center
GBA:6
LRRK2:2
Parkin:5
PINK1:3

Russia
1 center
DJ1:1
GBA:32
LRRK2:15
Parkin:40
PINK1:2
SNCA:2
VPS35:1

Korea
3 centers
Parkin:27
SNCA:6

Japan
6 centers
GBA:95
LRRK2:30
Parkin:443
PINK1:33
SNCA:29
VPS35:8

Canada
2 centers
DJ1:2
GBA:34
LRRK2:328
Parkin: 65
PINK1:62
SNCA:65
VPS35:3

Taiwan
2 centers
DJ1:1
GBA:27
LRRK2:12
Parkin:30
PINK1:11
VPS35:2

U.S.A.
21 centers
DJ1:1
GBA:1614
LRRK2:561
PINK1:8
SNCA:9

Tunisia
1 center
LRRK2:559
Parkin:10
PINK1:59

Colombia
1 center
No numbers
reported

China
4 centers
DJ1:7
GBA:130
LRRK2:360
Parkin:309
PINK1:15
VPS35:5

Malaysia
1 center
LRRK2:200

Argentina
1 center
LRRK2:4

Brazil
5 centers
GBA:26
LRRK2:33
Parkin:29
PINK1:2
SNCA:1

South Africa
3 centers
GBA:13
LRRK2:15
Parkin:19
PINK1:4
SNCA:2

Saudi Arabia
1 center
DJ1:1
Parkin:15
PINK1:4

Israel
2 centers
DJ1:1
GBA:240
LRRK2:180
Parkin:1
VPS35:1

Singapore
1 center
GBA:20

Australia
2 centers
LRRK2:19
Parkin:11
PINK1:3
SNCA:2
VPS35:4

Turkey
4 centers
DJ1:3
GBA:2
LRRK2:2
Parkin:36
PINK1:2
SNCA:2

Thailand
1 center
GBA:24
Parkin:14

Table 1: Number of centers and reported genetic PD patients in Europe.

	centers	DJ1	GBA	LRRK2	Parkin	PINK1	SNCA	VPS35
Austria	2	0	30	1	6	1	0	3
Belgium	3	0	12	43	1	0	1	0
Czech Republic	1	0	0	0	0	0	0	0
Denmark	2	0	0	0	2	0	0	0
Estonia	1	0	15	25	5	0	1	5
Faroer Islands	1	0	0	1	0	0	0	0
Finland	3	0	15	2	2	0	2	0
France	1	2	124	189	145	19	27	5
Germany	4	1	119	20	23	12	2	2
Greece	2	0	38	1	3	0	77	0
Hungary	1	0	3	1	0	0	0	0
Ireland	1	1	21	2	25	1	0	0
Italy	10	4	227	147	117	28	22	0
Netherlands	1	1	15	1	3	0	1	0
Norway	2	0	35	40	5	4	0	0
Poland	3	0	0	1	42	2	2	0
Portugal	2	1	0	0	4	1	0	0
Serbia	1	1	78	6	17	1	0	0
Spain	7	2	77	386	72	4	8	0
Sweden	2	2	37	14	2	0	6	0
United Kingdom	4	0	115	31	13	5	7	0

Figure 2: Response (for details, please see Figure S1).

**Response
N=162, 50%**

No response
N=165, 50%

**Complete surveys
N=128, 79%**

**Interest in further
collaboration
N=125, 98%**

**N=57 personal
emails expressing
interest to
collaborate**

Supplement

Figure S1: Response analysis. *including publications on GBA.

Publications represented in MDSGene*
N=470

Corresponding authors,
no duplicates
N=265

Corresponding authors with
multiple publications (range 2-
19) N=71

Email delivered
N=198

Email not deliverable
N=67

Resent to
alternative email
address N=46

No alternative
email address
available N=21

Additional contacts
N=98

Email delivered
N=238

Email not deliverable
N=6

Response
N=32

Response
N=130

Completed surveys
N=23

Completed surveys
N=105

No participation
N=10

Retired
N=2

**Total completed surveys
N=128**

After first email
N=46

Survey completed by
another group/consortium
member N=11

Change in field of
interest
N=3

After first reminder
N=38

Could not complete survey
before deadline N=4

No own cohort
N=5

After second reminder
N=21

13

Figure S2: The online survey.

THE MICHAEL J. FOX FOUNDATION
FOR PARKINSON'S RESEARCH

UNIVERSITÄT ZU LÜBECK

MDSCene

Identifying global cohorts with monogenic Parkinson's disease

Cohort lead	
Cohort name	
Contact person	
Contact e-mail	
University/Hospital	
City	

Ethnic origin of cohort
☐ African ☐ Arab ☐ Ashkenazi Jews
☐ Asian ☐ Caucasian ☐ Hispanic
☐ Indian ☐ Mennonites ☐ North Africans
☐ Other

Country of origin — please select ▲▼

PD Genes

Please only report PD patients with proven mutation here.

☐ DJ1   n= [ ]
☐ GBA   n= [ ]
☐ LRRK2 n= [ ]
☐ Parkin n= [ ]
☐ PINK1 n= [ ]
☐ SNCA   n= [ ]
☐ VPS35 n= [ ]
☐ Others n= [ ]

Cohort information

	available	not available	can be obtained
Age	○	○	○
Sex	○	○	○
Ethnicity	○	○	○
Pedigree/Family history	○	○	○
Age at onset	○	○	○
UPDRS	○	○	○
Hoehn & Yahr score	○	○	○
Dopaminergic medications	○	○	○
Treatment response	○	○	○
Non-motor signs	○	○	○
Environmental exposures	○	○	○
Life style variables	○	○	○

Other data

	available	not available	can be obtained
Omics data	○	○	○
Imaging	○	○	○

Biospecimens

	available	not available	can be obtained
DNA	○	○	○
RNA	○	○	○
Serum	○	○	○
Plasma	○	○	○
Whole blood	○	○	○
CSF	○	○	○
Fibroblasts	○	○	○
iPSCs	○	○	○
Brain tissue	○	○	○

Comments

	yes	no
May we contact you for future information on your cohorts and possible collaboration?	○	○
Are you interested in joining the MJFF Parkin/LRRK2 or GBA consortium?	○	○

Done

15

Table S1: Availability of information.

	number of researchers who reported available information (percentage of all participating researchers)
Age	123 (96%)
Sex	123 (96%)
Ethnicity	119 (92%)
Pedigree	110 (86%)
Age at onset	117 (91%)
Unified Parkinson Disease Rating Scale (UPDRS)	81 (63%)
Hoehn & Yahr Scale	88 (69%)
Dopaminergic medications	93 (73%)
Nonmotor signs	76 (59%)
Environmental exposures	50 (39%)
Life style variables	41 (32%)
Treatment response	88 (69%)
Omics data	21 (16%)
-Genomics	16 (12%)
-Transcriptomics	3 (0.02%)
-Proteomics	2 (0.02%)
-Metabolomics	2 (0.02%)
Imaging	47 (37%)
-MRI	41 (32%)
-SPECT/PET	24 (19%)
-TCS	8 (0.06%)
DNA	103 (80%)
RNA	25 (20%)
Serum	31 (24%)
Plasma	28 (22%)
Whole blood	27 (21%)
CSF	9 (0.07%)
Fibroblasts	13 (10%)
iPSCs	14 (11%)
Brain tissue	14 (11%)