ONLINE RESOURCES

Rare variants in *LRRK1* and Parkinson's disease

Eva C. Schulte, MD[a,b]; Daniel C. Ellwanger, MSc[c]; Sybille Dihanich, PhD[d]; Claudia Manzoni, PhD[d]; Katrin Stangl, MD[e]; Barbara Schormair, PhD[b,f]; Elisabeth Graf, MSc[b]; Sebastian Eck, MSc[b]; Brit Mollenhauer, MD[g,h]; Dietrich Haubenberger, MD[i]; Walter Pirker, MD[i]; Alexander Zimprich, MD[i]; Thomas Brücke, MD[j]; Peter Lichtner, PhD[b,f]; Annette Peters, PhD[k]; Christian Gieger, PhD[l]; Claudia Trenkwalder, MD[g,m]; Hans-Werner Mewes, PhD[c]; Thomas Meitinger, MD[b,f]; Patrick A. Lewis, PhD[d,n]; Hans H. Klünemann, MD[e]; Juliane Winkelmann, MD[a,b,f,]*

[a] Klinik und Poliklinik für Neurologie, Klinikum rechts der Isar, Technische Universität München, Munich, Germany
[b] Institut für Humangenetik, Helmholtz Zentrum München, Munich, Germany
[c] Chair of Genome-Oriented Bioinformatics, Technische Universität München, Freising, Germany
[d] Department of Molecular Neuroscience, UCL Institute of Neurology, University College London, London, UK
[e] Klinik und Poliklinik für Psychiatrie, Psychosomatik und Psychotherapie, Universität Regensburg, Regensburg, Germany
[f] Institut für Humangenetik, Technische Universität München, Munich, Germany
[g] Paracelsus Elena Klinik, Kassel, Germany
[h] Neurochirurgische Klinik, Georg August Universität Göttingen, Göttingen, Germany
[i] Department of Neurology, Medical University of Vienna, Vienna, Austria
[j] Neurologische Klinik, Wilhelminenspital, Vienna, Austria
[k] Institut für Epidemiologie II, Helmholtz Zentrum München, Munich, Germany
[l] Institut für Genetische Epidemiologie, Helmholtz Zentrum München, Munich, Germany
[m] Neurologische Klinik, Georg August Universität Göttingen, Göttingen, Germany
[n] School of Pharmacy, University of Reading, Whiteknights, Reading, UK

*corresponding author:
Prof. Dr. Juliane Winkelmann, Klinik und Poliklinik für Neurologie, Klinikum rechts der Isar, Technische Universität München, Ismaningerstr. 22, 81675 München, Germany
Phone: +49 89 4140 4688, Fax: +49 89 4140 7681, Email: winkelmann@lrz.tum.de

ONLINE RESOURCES METHODS

Description of case/control sample

All cases used in variant screening and genotyping were recruited at Paracelsus-Elena Klinik, a hospital specializing in Parkinson's disease (PD), in Kassel, Germany, as well as at the departments of neurology at Wilhelminenspital and Allgemeines Krankenhaus in Vienna, Austria. PD diagnosis was made in accordance with the UK Brain Bank Criteria. Controls belong to a large general population cohort (KORA) based in the region around Augsburg in Southern Germany and have been described previously.[1] KORA-AGE represents a subset of the KORA cohort collected in 2009 as a gender- and age-stratified subsample of the KORA S1-S4 studies comprising participants born before 1944. All individuals taking dopaminergic drugs were excluded from the control sample.

Bioinformatic prioritization of variants

Multiple Sequence Alignment

A multiple sequence alignment was computed using ClustalW based on *LRRK1/LRRK2* pairs from the following organisms: *Homo sapiens* (NP_078928.3, NP_940980.3), *Mus musculus* (NP_666303.3, NP_080006.3), *Rattus norvegicus* (NP_001178553.1, NP_001178718.1), *Bos taurus* (NP_001192703.1, NP_001193015.1), *Canis familiaris* (XP_545823.2, XP_543734.2), *Danio rerio* (XP_002667476.2, NP_001188385.1), *Callithrix jacchus* (XP_002749111.1, XP_002752413.1), *Macaca mulatta* (XP_001084079.1, XP_002798616.1), *Ailuropoda melanoleuca* (XP_002922722.1, XP_002925880.1), *Equus caballus* (XP_001489911.1, XP_001914702.1), *Monodelphis domestica* (XP_001373133.2, XP_001367394.1), *Pan troglodytes* (XP_510623.3, XP_001168494.1), *Meleagris gallopavo* (XP_003209642.1, XP_003201970.1), *Pongo abelii* (XP_002825936.1, XP_002823165.1), *Xenopus (Silurana)*

tropicalis (XP_002939309.1, XP_002932250.1), *Anolis carolinensis* (XP_003225956.1, XP_003221560.1), *Nomascus leucogenys* (XP_003277581.1, XP_003252348.1), and *Sus scrofa* (XP_003121698.1, NP_001106908.1). We determined *LRRK2* mutations related to PD (rs33939927: *LRRK2* p.Arg1441Gly, rs35801418: *LRRK2* p.Tyr1699Cys, rs34637584: *LRRK2* p.Gly2019Ser, rs35870237: *LRRK2* p.Ile2020Thr) on the *LRRK1* peptide sequence and introduced the *LRRK2* nucleotide mutation to the corresponding *LRRK1* coding triplet (rs33939927: *LRRK1* p.Lys746Glu, rs35801418: *LRRK1* p.Phe1022Cys, rs34637584: *LRRK1* p.Gly1411Arg, rs35870237: *LRRK1* p.Ile1412Thr).

Multi-model Ensemble

We implemented a multi-model ensemble of prediction algorithms (PolyPhen[2], PolyPhen-2[3], Phd-SNP[4], SIFT[5], SNPs3D[6], MutationTaster[7] and Pmut[8], each contributing equally). Since each model provides different scoring schemes, their solution space $s \in S$ was transformed by a function *p* computing the probability score of a variant to affect the function of the protein:

(1) $$p(s, r, class) = \begin{cases} [0.0, 0.5[& if\ class = 0 \\ ]0.5, 1.0] & if\ class = 1 \\ 0.5 & else \end{cases}$$

If a reliability value $r \in R$ of the classification was denoted by the algorithm, it was considered in the transformation, otherwise *r* was set to 1. A low confidence converges the probability score to 0.5, e.g. a non-reliable prediction was scored as $p(s, 0, class) = 0.5$. The score distributions of each class were determined by means of an exhaustive set of predictions provided by the algorithms' databases. The probability scores of each algorithm $i = \{1, \ldots, n\}$ were combined into a single score:

(2) $$P_{score} = \frac{\sum_{i=1}^{n} p_i(s, r, class)}{n}$$

Structural Analysis of Mutation

LRRK1 amino acid sequences containing annotated protein domains were folded by the multiple-threading approach I-TASSER[9] and the predicted tertiary structures with the highest confidence scores were selected. Mutations were mapped to the peptide structure with the SWISSPDB[10] viewer. Energy minimizations were performed by the NOMAD-Ref algorithm[11] with the conjugate gradient method for the wildtype and the variant structures. We computed the root mean square deviation (RMSD) between the wildtype peptide θ_1 and the variant peptide θ_2 :

(3) $$RMSD\,(\theta_1, \theta_2) = \sqrt{\frac{\sum_{i=1}^{n} (x_{1,i} - x_{2,i})^2}{n}}$$

Based on all variants considered, the *RMSD* was normalized for each functional domain *m* and a deviation score was calculated:

(4) $$Dscore = \frac{RMSD(\theta_m, \theta_2)}{\max(RMSD(\theta_m, \theta_m))}$$

Scaling of Prediction Values

To include tertiary structure information, we combined the *PScore* and the *Dscore* by means of a weighted harmonic mean to a mutation score:

(5) $$Mscore = (1 + \beta^2) \times \frac{Pscore \times Dscore}{(Pscore \times \beta^2) + Dscore}$$

Since *ab initio* tertiary structure determination is rather inaccurate, we selected $\beta = 0.25$, thus giving a higher weight to the prediction ensemble.

Cellular analyses

Reported (p.Lys651Ala and p.Lys1270Trp) [12] and newly identified (p.Arg631Trp, p.Arg1261His, p.Arg1271Glu and p.Tyr1410Asp) variants were inserted into the open reading frame of *LRRK1* in a 2xMyc tag vector by site directed mutagenesis using the Quickchange II XL kit (Agilent) according to the manufacturer's instructions and verified by sequencing.

SHSY-5Y cells (ATTC # CRC-2266) human neuroblastoma cells were cultured in Dulbeccos Modified Eagles Medium supplemented with 10% fetal bovine serum (both Life technologies) at 37°C and 5% CO_2. Cells were transfected with *LRRK1* using Fugene 6 reagent according to manufacturer's instructions. Protein expression was assessed by immunoblot analysis. Cells in $10cm^2$ dishes were harvested in 1ml of ice-cold RIPA buffer (Cell Signaling) supplemented with complete protease inhibitors (Roche) and lysed at 4°C for 30 minutes. Lysates were clarified by centrifugation at 10000*g* for 10 minutes at 4°C, the protein concentration in the supernatents quantified by BCA assay (Pierce) and samples diluted to equivalent concentration. 10µg of lysate for each construct was loaded onto 4-12% Bis-Tris acrylamide gels (Life technologies) and electrophoresised at 160V for 80 minutes. Protein was transferred to PVDF membrane (Millipore) by western blot, and resulting membranes blocked with 5% milk in TBST. To detect LRRK1, anti-myc mouse monoclonal antibody (Sigma) was used at a 1:2000 dilution followed by anti-mouse HRP conjugated secondary antibody at a 1:5000 dilution. Anti β-Actin mouse monoclonal antibody was used at a 1:5000 dilution, followed by probing with HRP conjugated secondary at a 1:10000 dilution. Bands were detected by incubation with ECL reagent (Pierce) and exposure to SuperRX film (Fujifilm), developed on a Konica SRX101A processor.

Cell death assays were carried out by MTT assay. Cell culture medium was supplemented with (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Sigma Aldrich) to a final concentration of 500 µg/ml for 3 hours. Cell medium was then discarded and the formazan crystals accumulated within the energetically active cells were dissolved in pure DMSO. The plate was then analyzed in a multi-well plate reader accessing the absorbance of every single well at the wavelength of 570 nm. The results were reported as percentage of cell viability after treatment in comparison with untreated, control cells. Graphs and statistical analyses were performed by Prism software.

For immunocytochemistry, cells were seeded on coverslips in 24 wells plates at the concentration of $2x10^5$ cell/ml (0.5ml each well). 24 hours following transfection cells were washed twice in DPBS and fixed a room temperature for 15 minutes in a solution of 4% paraformaldehyde in DPBS. Cell were washed three times in DPBS, blocked and permeabilized at room temperature for 30 minutes by using a solution of 15% normal goat serum (S1000, Vector) and 0.1% Triton X-100 in DPBS. After washing, cells were incubated overnight at 4°C with the primary antibody. Anti-mouse, secondary antibody (A21124, Alexa Fluor, emission at 568 nm) was used to reveal the primary antibody staining and nuclei were labelled with a 0.05% solution of Hoechst in DPBS before the sealing the coverslips with Fluoromount G mounting medium (Southern Biotech). Images were acquired with a Leica DM5500 B fluorescence microscope.

ONLINE RESOURCES REFERENCES

1. Wichmann HE, Gieger C, Illig T (2004) KORAgen—Resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen. 67 Suppl 1:26-30.

2. Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. Nucleic Acid Res. 30:3894-3900.

3. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods. 7:248-249.

4. Capriotti E Calabrese R Casadio R (2006) Predicting the insurgence of human genetic diseases associtated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 22:2729-2734.

5. Ng PC, Henikoff S (2006) Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet. 7:61-80.

6. Yue P, Melamud E, Moult J (2006) SNPs3D: candidate gene and SNP selection for association studies. BMC Bioinformatics. 7:166.

7. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods. 7:575-576.

8. Ferrer-Costa C, Gelpi JL, Zamakola L, Parraga I, de la Cruz X, Orozco M (2005) PMUT: a web-based tool for annotation of pathological mutations on proteins. Bioinformatics. 21:3176-3178.

9. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc, 5:725-738.

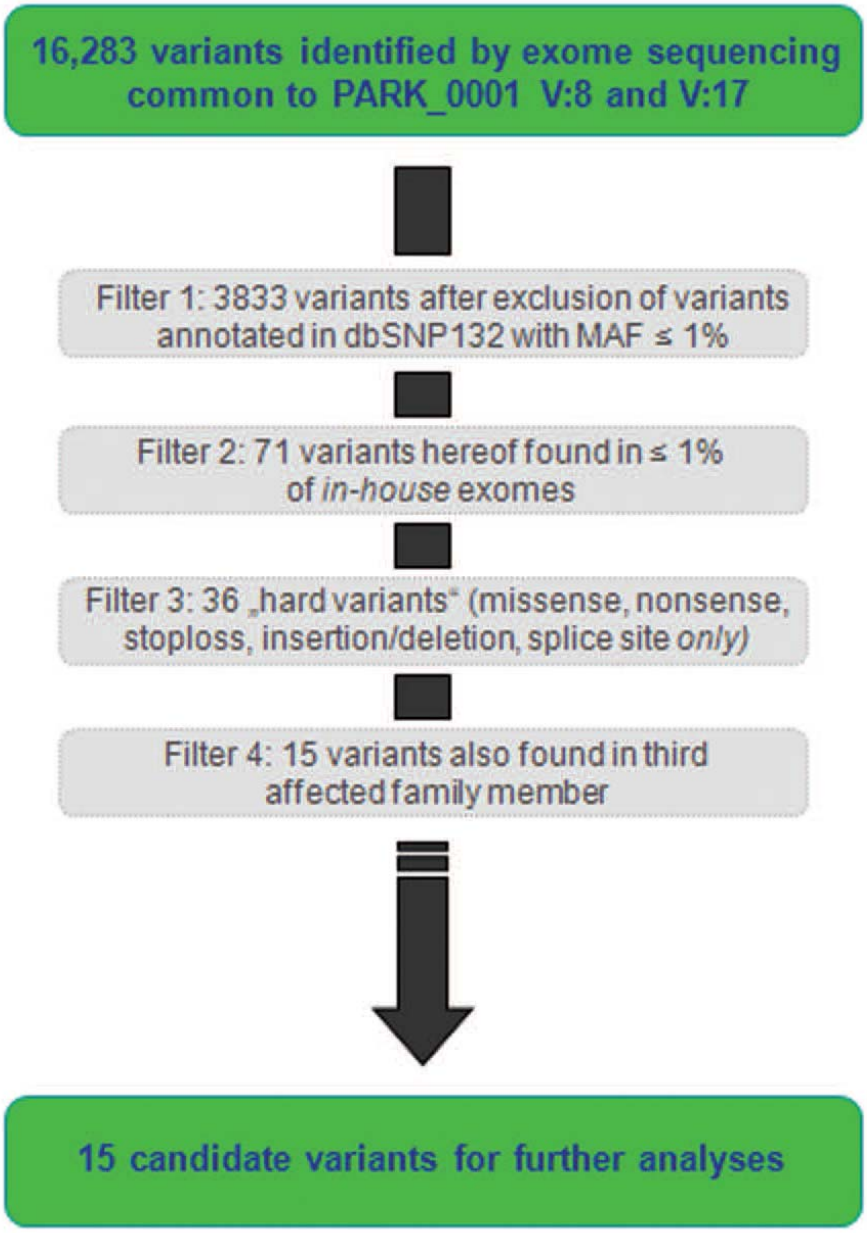
10. Guex N, Peitsch MC (1997) SWISS-MODEL and Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 18:2714-2723.

11. Lindahl E, Azuara C, Koehl P, Delarue M (2006) NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. Nucleid Acid Res. 34:W52-56.

12. Korr D, Toschi L, Donner P, Pohlenz HD, Kreft B, Weiss B (2006) LRRK1 protein kinase activity is stimulated upon binding of GTP to its Roc domain. Cell Signal. 18:910-920.

ONLINE RESOURCES FIGURES

Online Resources Figure 1



Filtering scheme for variants identified by exome sequencing in the two affected family members examined.

ONLINE RESOURCES TABLES

Online Resources Table 1

Clinical Phenotype of Genotyped Affected Individuals

Individual ID	Age at Onset	Disease Duration	IS	B	R	RT	PI	L-Dopa/ DA	Additional Features
V:8	56	17	RT	+	+	+	++	++	MCI, episodes of depression, hyperreflexia
V:9	58	12	B	+	++	+	+	+	Dementia, episodes of depression, Babinski's sign bilaterally and general hyperreflexia
V:17	56	9	B	+	+	++	+	+	MCI, episodes of depression, Babinski's sign on the left and general hyperreflexia

IS = initial symptom, B = bradykinesia, R = rigor, RT = resting tremor, PI = postural instability, DA = dopamine agonist, MCI = mild cognitive impairment

Online Resources Table 2

Non-Synonymous and Indel Variants Identified in Variant Screening of *EEF1D* and *LRRK1*.

gene	genomic position (hg19)	dbSNP132	variation		frequency		domain
			nucleotide	amino acid	cases (n=862)	controls (n=940)	
EEF1D	chr8:144671439-144671422	novel	c.813_830del18bp	c.813_830del18bp		1	n/a
EEF1D	chr8:144671384	novel	c.868 G>A	p.Gly290Arg	1		n/a
EEF1D	chr8:144671279	novel	c.973 G>A	p.Ala325Thr	2		n/a
EEF1D	chr8:144671194	novel	c.1058 G>A	p.Arg353Gln		1	n/a
EEF1D	chr8:144662764	novel	c.1622 G>A	p.Arg541Ile		2	n/a
EEF1D	chr8:144662740	novel	c.1646 G>A	p.Ala549Val*	1		n/a
EEF1D	chr8:144662286	novel	c.1801 G>A	p.Pro601Ser	1		n/a
LRRK1	chr15:101562626	novel	c.1891 C>T	p.Arg631Trp	1		ROC
LRRK1	chr15:101565017	novel	c.2072 G>A	p.Val693Met		1	ROC
LRRK1	chr15:101565029	novel	c.2089 G>A	p.Val697Ile	8	9	ROC
LRRK1	chr15:101566195	novel	c.2258 T>C	p.Leu753Pro		1	ROC
LRRK1	chr15:101567500	novel	c.2440 G>A	p.Gly814Arg		1	ROC
LRRK1	chr15:101567909	novel	c.2593 G>A	p.Asp865Asn	3		ROC
LRRK1	chr15:101567912	rs56003881	c.2596 G>A	p.Asp866Asn	5	4	COR
LRRK1	chr15:101567959	novel	c.2643 G>T	p.Gln881His		1	COR
LRRK1	chr15:101569374	rs41531245	c.2900 C>T	p.Thr967Met	2		COR
LRRK1	chr15:101569388	novel	c.2914 T>C	p.Phe972Leu		1	COR
LRRK1	chr15:101586235	novel	c.3013 G>A	p.Gly1005Ser	1		COR
LRRK1	chr15:101586332-101586344	novel	c.3110_3122del13bp	c.3110_3122del13bp	1		COR
LRRK1	chr15:101588745	novel	c.3182 C>T	p.Thr1061Ile		1	COR
LRRK1	chr15:101589988	novel	c.3439 G>A	p.Ala1147Thr		1	COR
LRRK1	chr15:101593161	novel	c.3724 G>A	p.Glu1242Gln		1	kinase
LRRK1	chr15:101593187	novel	c.3730 G>C	p.Glu1244Gln	1		kinase
LRRK1	chr15:101593213	novel	c.3776 G>A	p.Arg1259Gln		2	kinase
LRRK1	chr15:101593219	novel	c.3782G>A	p.Arg1261Gln*	4	8	kinase
LRRK1	chr15:101593249	novel	c.3812 G>A	p.Arg1271His	1		kinase
LRRK1	chr15:101593457	novel	c.3886 G>A	p.Asp1296Asn	1		kinase
LRRK1	chr15:101593508	novel	c.3937 A>G	p.Ala1313Thr	1		kinase
LRRK1	chr15:101595324	novel	c.4228 T>G	p.Tyr1410Asp	1		kinase

n/a=not available; for *EEF1D*, no known protein domains are annotated in UniProt (accessed February 3, 2012). An asterix denotes the original variant identified in exome sequencing.