Blood RNA biomarkers in prodromal PARK4 and rapid eye movement sleep behavior disorder show role of complexin 1 loss for risk of Parkinson’s disease

Suna Lahut1,2, Suzana Gispert1, Özgür Ömür1,2, Candan Depboylu3, Kay Seidel4, Jorge Antolio Domínguez-Bautista5, Nadine Brehm1, Hülya Tireli1, Karl Hackmann6, Caroline Pirkevi2, Barbara Leube7, Vincent Ries3, Kerstin Reim3, Nils Brose8, Wilfred F. den Dunnen9, Madrid Johnson10, Zsuzsanna Wolf11, Marc Schindewolf12, Wiebke Schrempf13, Kathrin Reetz14, Peter Young15, David Vadasz3, Achilleas S. Frangakis10, Evelin Schroeter15, Helmut Steinmetz1, Marina Jendrach1, Udo Rüb4, Ayşe Nazli Başak1, Wolfgang Oertel*3 and Georg Auburger*1,2

ABSTRACT
Parkinson’s disease (PD) is a frequent neurodegenerative process in old age. Accumulation and aggregation of the lipid-binding SNARE complex component α-synuclein (SNCA) underlies this vulnerability and defines stages of disease progression. Determinants of SNCA levels and mechanisms of SNCA neurotoxicity have been intensely investigated. In view of the physiological roles of SNCA in blood to modulate vesicle release, we studied blood samples from a new large pedigree with SNCA gene duplication (PARK4 mutation) to identify effects of SNCA gain of function as potential disease biomarkers. Downregulation of complexin 1 (CPLX1) mRNA was correlated with genotype, but the expression of other Parkinson’s disease genes was not. In global RNA-Seq profiling of blood from presymptomatic PARK4 individuals, bioinformatics detected significant upregulations for platelet activation, hemostasis, lipoproteins, endocytosis, lysosome, cytokine, Toll-like receptor signaling and extracellular pathways. In PARK4 platelets, stimulus-triggered degranulation was impaired. Strong SPP1, GZMH and PLTP mRNA upregulations were validated in PARK4. When analysing individuals with rapid eye movement sleep behavior disorder, the most specific known prodromal stage of general PD, only blood CPLX1 levels were altered. Validation experiments confirmed an inverse mutual regulation of SNCA and CPLX1 mRNA levels. In the 3′-UTR of the CPLX1 gene we identified a single nucleotide polymorphism that is significantly associated with PD risk. In summary, our data define CPLX1 as a PD risk factor and provide functional insights into the role and regulation of blood SNCA levels. The new blood biomarkers of PARK4 in this Turkish family might become useful for PD prediction.

KEY WORDS: α-synuclein, Complexin 1, PARK4, Rapid eye movement sleep behavior disorder, Parkinson’s disease, Biomarkers

INTRODUCTION
Parkinson’s disease (PD) is the second most frequent age-associated brain degeneration disorder, affecting ∼1% of the population >65 years of age. The PD-specific progressive movement deficit is mostly attributable to the severe affliction and cell death of midbrain nigrostriatal dopaminergic neurons (Braak et al., 2003). Surviving neurons in vulnerable regions exhibit aggregates predominantly consisting of the protein α-synuclein (SNCA), which are visualized as Lewy neurites and Lewy bodies, both in sporadic late-onset PD and in most familial early-onset PD variants (Goedert et al., 2013). Autosomal dominant PD with early clinical manifestation was observed in rare families, leading to the identification of SNCA gene missense mutations (PARK1 variant), such as A53T, and of SNCA gene duplication/triplication events (PARK4 variant) as the strongest causes of this pathology (Polymeropoulos et al., 1997; Singleton et al., 2003). Further recruitment of PD families led to the identification of several disease genes responsible for monogenic PD (Corti et al., 2011). In addition, genome-wide association studies (GWAS) of very large collectives of late-manifesting sporadic PD cases identified two regions on chromosome 4 (SNCA locus and CPLX1/GAK/TMEM175/DGKQ locus) that contain genetic variants predisposing to multifactorial PD (Lill et al., 2012; Nalls et al., 2014). Variations in the SNCA gene 3′-untranslated region (3′-UTR) and its promoter were strongly correlated with PD risk (Rhim et al., 2012).

SNCA is physiologically concentrated in axon terminals. It is associated with the lipid membranes of synaptic vesicles and interacts with synaptobrevin, a component of the SNARE complex,
modulating vesicle exocytosis and neurotransmission (Diao et al., 2013). Its toxic gain of function leads over time to impaired synaptic vesicle release and synaptic failure (Garcia-Reitbock et al., 2010; Nemani et al., 2010; Platt et al., 2012; Ianzek et al., 2013). Current investigations aim to elucidate SNCA-triggered pathology, concentrating on disease stages before the occurrence of irreversible cell loss, when neuroprotective therapies might still be efficacious. In the prodromal stage of PD, non-motor symptoms such as hyposmia, constipation, depression or rapid eye movement (REM) sleep behavior disorder (RBD) were documented, of which RBD is now recognized as the most specific and predictive prodromal phenotype. Individuals suffering from RBD carry a risk of >85% to manifest PD after 15-20 years, and the associated neurodegenerative process is a synucleinopathy in 95% of cases (Stiasny-Kolster et al., 2005; Albers et al., 2012; Boeve et al., 2013; Iranzo et al., 2013, 2014; Mahowald and Schenck, 2013).

SNCA is abundantly expressed in blood (Shin et al., 2000; Barbour et al., 2008). The accumulation of SNCA in short-lived blood cells was found to result in diverse subtle phenotypes. Enhanced apoptotic vulnerability of human PARK1 lymphocytes and SNCA-transfected myeloma and leukemia cell lines to oxidative stress (Kim et al., 2004; Battisti et al., 2008), impaired innate immune functions of mouse leukocytes with SNCA overexpression (Gardai et al., 2013) and dose-dependent inhibition of α-granule release in human platelets exposed to exogenous SNCA (Park et al., 2002) provide evidence that biomarkers of elevated SNCA abundance and of the risk of synucleinopathy can be identified in peripheral tissues.

Our identification of a new large pedigree of autosomal dominant PD attributable to SNCA gene duplication with 12 presymptomatic PARK4 heterozygotes has provided a unique opportunity to explore blood biomarkers and permitted the definition of a molecular signature at the RNA level that predicts PARK4 PD. For validation, the results were assessed in individuals with a risk of developing PD because of manifestation of RBD as a highly specific prodromal sign. Our data on blood biomarkers as a diagnostic tool might contribute to the assessment of the risk of multifactorial PD in individuals without a positive family history. The most relevant biomarker is the SNARE component complexin 1, which acts as risk factor for PD by itself.

RESULTS

RNA levels in blood from presymptomatic PARK4 heterozygotes are reduced for CPLX1

An exceptionally large pedigree with autosomal dominant PD (Fig. 1A) was identified in Turkey. Blood DNA genotyping demonstrated a genomic tandem duplication at the SNCA locus as a known cause of PD (Singleton et al., 2003; Fig. 1B) and detected presymptomatic PARK4 heterozygotes, 12 of whom were available and included in this study. The two clinically affected family members plus the 12 presymptomatic heterozygotes (mean age 45.5 years, range 29-56 years, six males) and the 12 age-matched control relatives (mean age 44.6 years, range 31-57 years, six males) underwent overnight fasting and had whole peripheral blood protein and RNA samples collected and processed in parallel to assess the SNCA-dependent expression profiles of blood. As a consequence of the tandem duplication, the blood mRNA levels of the neighbouring genes GPRIN3, SNCA and MMRN1 were increased to ~1.5-fold in blood of PARK4 individuals versus control relatives (11 versus 9). Given that the MMRN1 gene dosage has also been implicated in cognitive decline (Nishioka et al., 2006; Fuchs et al., 2007; Mutez et al., 2011), it will be interesting to perform a neuropsychological characterization of the PARK4 heterozygotes in this family in the future, but currently the cognitive score in one patient is still unaffected (mini-mental state examination score, MMSE=30). In an initial candidate gene study, we studied the blood expression of all PARK genes, of the previously reported PD blood biomarker ST13 and of promising transcripts that encode putative SNCA-interactor proteins (Fig. 2A,B) and showed altered levels in our previous global transcriptome profiling of midbrain tissue in a carefully characterized synucleinopathy mouse model with dopaminergic deficits, impaired synaptic plasticity and mitochondrial dysfunction (Kurz et al., 2010; Platt et al., 2012; Tozzi et al., 2012; Gispert et al., 2015a,b; Subramaniam et al., 2014; Brehm et al., 2015b). SNCA gene duplication in the Turkish pedigree showed no correlation with the expression of other genes with known association with PD risk or of the previously reported PD blood biomarker ST13 (Scherzer et al., 2007). A nominally significant inverse correlation was observed for CPLX1 and a trend towards inverse correlation for YWHAE (Fig. 2B), two phenomena that we had previously identified as synucleinopathy markers in mouse mutant midbrain (Gispert et al., 2015b; Brehm et al., 2015a). CPLX1 levels were reduced to 0.67-fold, so the effect size was moderate; despite the nominal significance (P=0.04), a demonstration of real significance after Bonferroni or Benjamini–Hochberg correction would be possible only in much larger sample collections. Quantitative immunoblots of protein extracts from corresponding whole blood PARK4 samples showed a ~1.5-fold accumulation of SNCA monomer, but no SNCA aggregates were detectable, even after delipidation for epitope unmasking, probably because of the short lifespan of blood cells (Sharon et al., 2003; Fig. 1C). The protein levels of complexin 1 could not be detected reliably in blood. Together with previous mouse findings (Chandra et al., 2004; Gispert et al., 2015b), these data suggest that CPLX1 mRNA levels in blood mirror the gain of physiological function of SNCA rather than aggregation pathology. This provides proof of principle that blood RNA profiling can aid the development of risk prediction biomarkers for synucleinopathies.

Blood platelets from presymptomatic PARK4 heterozygotes suggest reduced stimulus-triggered degranulation in the absence of SNCA aggregates

Given that exogenous recombinant human SNCA and deletion of murine Snca are known to modulate the stimulus-triggered vesicular release from blood platelets (Park et al., 2002; Reheman et al., 2010), we investigated platelet function in two presymptomatic PARK4 heterozygotes (male aged 48 years and female aged 45 years) and two age- and sex-matched first degree control relatives (male aged 43 years and female aged 45 years) who could travel from Turkey to Germany. As revealed by flow cytometry, stimulation-induced degranulation increased the cell surface presence of the α-granule antigen CD62P (P-selectin) and the lysosomal CD63 antigen. In comparison to their control relatives and to the normal German population, these degranulation-dependent changes were diminished in presymptomatic PARK4 heterozygotes (Table 1). Electron microscopic analysis did not detect protein aggregates in the PARK4 platelets (Fig. 3). The presymptomatic PARK4 heterozygotes did not report any manifest blood coagulation disorder. These functional data might explain the expression dysregulation of the SNARE complex component CPLX1 as a compensatory effort within the platelet activation pathway.

Global RNA profiling of blood from presymptomatic PARK4 heterozygotes detects several strongly upregulated pathways

Global expression profiling by RNA-seq in blood samples from five presymptomatic PARK4 heterozygotes (male aged 53 years, female...
aged 59 years, male aged 50 years, female aged 47 years and male aged 50 years) versus five age- and sex-matched controls (male aged 45 years, female aged 54 years, male aged 57 years, female aged 42 years and male aged 42 years) was used to identify additional molecular effects of \( \text{SNCA} \) gene duplication. All data were deposited at the European Nucleotide Archive (ENA) public database (accession number PRJEB8960), and a table containing the ranked gene list of expression changes is provided in Table S1.

In order to identify pathway dysregulations, the data were assessed bioinformatically with Gene Set Enrichment Analysis (GSEA). Downregulated pathways were not apparent, but several strong upregulations (with \( P \)-values=0.0 and \( q \)-values=0.0) were documented and concerned cytokine signaling in the immune system (normalized enrichment score, NES 4.8), adaptive immune system (NES 4.5), hemostasis (NES 4.4), lysosome (NES 4.3), platelet activation signaling and aggregation (NES 4.1), innate immune system (NES 4.0), endocytosis (NES 3.9), Toll-like receptor signaling (NES 3.3), metabolism of lipids and lipoproteins (NES 3.3) and extracellular region (NES 2.9). These unbiased findings are in excellent agreement with the known features and functions of SNCA in blood, such as altered coagulation and altered immune competence, and might reflect a compensatory increase in the well-established lysosomal degradation of SNCA (Westbroek et al., 2011).

**Blood qPCR shows an increase in GZMH, SPP1 and PLTP in presymptomatic PARK4, but cannot distinguish prodromal multifactorial PD (RBD), whereas CPLX1 is useful as a biomarker in PARK4 and RBD**

**Studies on presymptomatic PARK4 heterozygotes**

Particularly strong effects within the lysosome, immune and lipid pathways were studied as promising biomarker candidates for PARK4. We decided to validate the strong GZMH transcript upregulation (to 204%, \( P=0.04 \) in Student’s paired \( t \)-test) by the independent qPCR technique in the same individuals plus additional members of the PARK4 pedigree, because the lysosomal enzymes cathepsin D and B, but not cathepsin G-like 2 (or granzyme H, GZMH) were previously implicated in degradation versus aggregation of SNCA (Crabtree et al., 2014; Tsujimura et al., 2014). The similar strong upregulation (to 244%, \( P=0.01 \)) of the immunity regulator osteopontin (\( \text{SPP1} \)) transcript was assessed, because osteopontin levels were previously identified as a biomarker of PD in blood serum, cerebrospinal fluid, microglia and affected neurons (Iczkiewicz et al., 2006; Maetzler et al., 2007; Westbroek et al., 2011).
Fig. 2. Correlation between whole blood mRNA levels (qPCR) of candidate genes and PARK4 genotype, comparing 14 adult PARK4 heterozygotes versus nine age-matched control relatives. (A) Scattergrams represent the expression levels in whole blood of individual transcripts determined by qPCR and normalized versus TBP loading control, with mean and s.e.m. (B) The bar graph summarizes these scattergrams by illustrating nominal Student’s unpaired t-test P-values. The selected candidate genes are either known to be mutated in monogenic PD (black) or were previously claimed to constitute a blood expression biomarker of PD (yellow) or represent components of the interactome of α-synuclein (SNCA), including β-synuclein (SNCB) (orange), 14-3-3 isoforms (YWHAG, YWHAB and YWHAE) (purple) and complexin 1 (CPLX1) (blue) within the presynaptic SNARE complex. *P<0.05.
Shi et al., 2015). The similar strong upregulation (to 189%, \(P=0.08\)) of the phospholipid transfer protein (\(PLTP\)) transcript was assessed, because PLTP modifies ataxia, Alzheimer’s disease and tau phosphorylation and serves as lipopolysaccharide interactor (Dong et al., 2009; Gautier and Lagrost, 2011; Albers et al., 2012). The qPCR analysis confirmed significant increases for both transcripts (to 179% with \(P=0.009\) for \(GZMH\); to 201% with \(P=0.016\) for \(SPP1\); Fig. 4A). In comparison, the recently published longitudinally dynamic biomarkers of PD in blood, \(HNF4A\) and \(PTBP1\), did not show significantly altered levels at this prodromal stage (\(n=9\) control versus \(n=12\) PARK4 individuals).

**Studies on individuals with RBD but asymptomatic for parkinsonian motor symptoms**

To investigate whether these candidate biomarkers might also be useful for risk prediction in individuals without a PD family history who manifest symptoms that might represent incipient PD (Postuma et al., 2012), we collected blood RNA from individuals with RBD after overnight fasting. The qPCR analysis of 46 individuals with RBD (mean age 65 years, range 34-83 years, with 36 males; since sampling, five individuals have converted to PD and one to Lewy body dementia, without DATscan confirmation) versus 19 matched controls (mean age 46 years, range 30-74 years, with 12 males) failed to detect a significant increase in \(GZMH\) \((P=0.29)\), \(SPP1\) \((P=0.35)\); but upon exclusion of one exceptionally high value among controls, an upregulation to 174% with \(P=0.04\) was observed) and \(PLTP\) \((P=0.66)\). By contrast, a significant downregulation was confirmed for the \(CPLX1\) transcript (to 74%, \(P=0.015\); upon exclusion of one exceptionally high value among controls, a statistical trend towards downregulation with \(P=0.06\) remained; Fig. 4B).

To define the stability of \(CPLX1\) mRNA levels in blood as a biomarker further, their dependence on sex and age was investigated. When the younger control individuals from Turkey were combined with the older control individuals from Central Europe, the expression did not show a significant correlation with age in linear regression analysis \((P=0.33)\). No deviation from linearity was detectable in the runs test \((P=0.53)\). There was no dependence on sex \((P=0.50)\). The PARK4 cohort showed a trend towards reduced \(CPLX1\) values at older age \((P=0.08)\) and a trend towards lower values in females \((P=0.08)\).

Overall, the blood qPCR observations validate the RNA-seq findings and suggest that \(GZMH, SPP1\) and \(PLTP\) upregulations are indeed strong effects of PD pathogenesis. However, they might not be specific enough for the risk diagnostics of sporadic multifactorial PD, whereas the comparatively subtle \(CPLX1\) downregulation is useful in large populations with parallel sample processing, even if its fold change is too small for individual diagnostics.

**Reduced \(CPLX1\) mRNA in human SNCA-transfected neuroblastoma**

To further evaluate the SNCA-triggered expression changes, we transiently overexpressed wild-type SNCA in human SH-SY5Y neuroblastoma cells, a widely used PD cell model with catecholaminergic properties. A downregulation of \(CPLX1\) mRNA in parallel to accumulation of complexin 1 protein occurred within 2 days after transfection (Fig. 5). This demonstrates that these effects are initiated early in neural cells. The observed upregulation of complexin 1 protein in radioimmunoprecipitation assay (RIPA) buffer fractions of these cells after 2 days contrasts with the downregulation of complexin 1 protein in 8 M urea fractions that was recently observed in the prefrontal cortex of autopsied individuals who had PD in old age (Dumitriu et al., 2016) and might be explained by altered \(CPLX1\) solubility, by progression from acute to chronic SNCA gain of function or by other differences between neuroblastoma cells versus cortical postmitotic neurons.

**Does complexin 1 protein aggregate in the midbrain in PD?**

Protein levels of SNCA and complexin 1 are increased in midbrain tissue from individuals with idiopathic PD and from mouse models of synucleinopathy (Basso et al., 2004; Gispert et al., 2015b). Such accumulations might be attributable to elevated gene dosage with increased transcript levels or to impaired degradation during an aggregation process. The latter mechanism might be responsible for
the decreased SNCA transcript levels that were described in some individuals with PD (Dächsel et al., 2007). Several other components of the SNARE complex that mediate membrane fusion and vesicle exocytosis, namely SNAP25, syntaxin 1 and synaptobrevin 2 (VAMP2), are contained in Lewy bodies and are depleted from surrounding neural tissue in PD (Garcia-Reitbock et al., 2010; Mukaetova-Ladinska et al., 2013). To test whether complexin 1 is aggregating within Lewy bodies, as was previously observed for SNCA and 14-3-3, we performed immunohistochemical analyses of midbrains from individuals with PD. A low-titer antibody (Acris) detecting complexin 1 with high specificity stained aggregates in neurites and neuronal perikarya of fixed tissue after antigen retrieval (Fig. 6A-D). Double immunofluorescence revealed co-localization of α-synuclein and complexin 1 in fixed tissue only after maximal exposure and contrast adjustment of the channel depicting complexin 1 (Fig. 6E-H).

These observations might support the notion that complexin 1 is sequestered into Lewy bodies, but of course the evidence must be considered with caution given that numerous proteins might be entrapped in aggregates without a specific role in pathogenesis. The recent report of a 20% reduction of CPLX1 protein in 8 M urea fractions of prefrontal cortex from autopsied individuals with PD (Dumitriu et al., 2016) is also compatible with a sequestration of complexin 1 into Lewy bodies. Given that we documented a 33% reduction of CPLX1 expression in blood of presymptomatic PARK4 individuals, the dysregulation appears to be relatively stable over decades and might not be useful as a biomarker of disease progression from the prodromal to the final stages of PD. Therefore, extensive investigations of CPLX1 expression in individuals with PD of varying severity were not attempted.

CPLX1 gene variant enhances PD risk

It remains unclear whether CPLX1 modulation is neuroprotective or instead contributes to pathogenesis. Interestingly, the CPLX1 gene is encoded within a locus of confirmed association with PD risk according to GWAS meta-analyses (Lill et al., 2012). The corresponding PD risk haplotype, with an odds ratio of 1.45 at chromosome 4p16.3, spanned the genes PCGF3/LOC100129917/CPLX1/GAK/TMEM175/DGKQ. The GWAS approach cannot reliably dissect such loci further, although it can detect relative disease risks as low as 1.1-fold. In the latest meta-analysis of GWAS studies in PD (Nalls et al., 2014), two single nucleotide polymorphisms (SNPs) in the intron 1 of CPLX1 directly next to the GAK gene reached genome-wide significance (rs76444973 and rs34006598), but it is difficult to predict how these variants would influence the function of CPLX1 protein. We therefore used a complementary approach and searched for a low-frequency CPLX1 gene variant, which shows its risk association already in small case-control collectives and explains a substantial part of the disease association contained in the locus. For this purpose, 360 random individuals with idiopathic general (non-monogenic) PD and 358 controls were studied regarding several candidate SNPs outside the linkage disequilibrium block around GAK, within the CPLX1 3′-UTR. The selected SNPs show minor allele frequencies between 0.1 and 0.5 in Caucasian populations and were available as TaqMan genotyping assays. Significant association with PD risk was observed for the G allele of the SNP rs10794536 (Table 2), with an odds ratio of 1.33. This 3′-UTR variant might influence the stability of CPLX1 mRNA. These data suggest that complexin 1 not only serves as a downstream marker of the physiological α-synuclein function, but also modulates PD risk.

Cplx1 gene ablation upregulates α-synuclein levels

In order to determine whether complexin 1 loss of function enhances or alleviates the susceptibility to PD, the relevant brain region was studied in a Cplx1 null mouse. This mouse line was previously shown to have an early and strong cerebellar phenotype (Reim et al., 2001; Glynn et al., 2005). In addition, these mice display dystonia, shuffled walking and reduced novelty seeking,
several signs that are characteristic effects of nigrostriatal dysfunction; intriguingly, these mice also display resting tremor, a diagnostic hallmark in humans for the onset of intermediate stage PD (Glynn et al., 2005). However, only cerebellar and not cerebral tissues have been investigated to date. Brains of the Cplx1 null mice revealed significantly increased levels (1.3- to 1.5-fold) of Snca mRNA and SNCA protein by 3 months of age (Fig. 7). The data provide further evidence that complexin 1 levels are not only a downstream marker of SNCA function, but are also involved in SNCA abundance, apparently by reciprocal feedback regulation of expression.

DISCUSSION

In summary, we report a new, exceptionally large PARK4 pedigree, provide proof of principle that the risk for a future manifestation of PD is reflected in the global transcriptome of blood, identify diverse pathways and, in particular, the levels of SPP1, GZMH, PLTP and CPLX1 as biomarkers of pathogenesis, show subtle CPLX1 downregulations to distinguish prodromal PD cohorts (presymptomatic PARK4 heterozygotes and RBD individuals) from controls, and suggest that complexin 1 loss of function acts as a PD risk factor.

Twenty-five years ago, the characterization of the large Italian Contursi PARK1 kindred triggered research into PD genetics (Golbe et al., 1990; Polymeropoulos et al., 1996, 1997). Now, the cooperativity of the Turkish PARK4 pedigree is a unique opportunity for further studies into human synucleinopathy, such as brain imaging, and the existence of 12 presymptomatic heterozygotes of the SNCA gene duplication in this family might be crucial for the identification of a molecular signature of disease risk. Our deposition

Fig. 5. Accumulation of complexin 1 protein and downregulation of CPLX1 mRNA levels after overexpression of WT SNCA in SH-SY5Y cells.

Human SH-SYSY neuroblastoma cells were transiently transfected with plasmid pcDNA3.1(+) as an empty vector control or WT SNCA. (A) Immunoblot confirmation of successful SNCA overexpression and of the effect on complexin 1 protein levels on the second day after transfection, using GAPDH and β-actin as the loading controls (n=3 versus n=3 in independent experiments). (B) Densitometric quantification of complexin 1 (antibody from Synaptic systems) versus ACTB ratios, normalized against control, in a scattergram analysis with H-test. (C) Analysis of corresponding CPLX1 mRNA levels at 2 days after transfection. Asterisks indicate statistical significance versus control group (n=6 versus n=3 versus n=12). *P<0.05, **P<0.01. ACTB, β-actin; α-Syn, α-synuclein; WT, wild type.

Fig. 6. PD midbrain autopsies contain cytoplasmic and neuritic aggregates. Tissue sections (5 μm thick) from individuals with PD were stained with anti-complexin 1 (Acris) or double stained with anti-complexin 1 and anti- SNCA antibodies. (A-C) Neurons in the substantia nigra exhibiting Lewy body-like structures mildly stained with anti-complexin 1 (arrowheads). (D) Complexin 1 immunopositive Lewy neurite-like structure (asterisks) in the medulla at the level of the motor vagus nucleus. (E-H) Double immunostaining depicting possible co-localization of α-synuclein (SNCA; green) and complexin 1 (red), only after maximal exposure and contrast adjustment in the red channel. Sections were counterstained with Hoechst dye.
of PARK4 blood RNA-seq data at the ENA database aims to drive this effort, ensuring that PD risk, which is triggered by the genetically defined synucleinopathy of PARK4 in the Turkish pedigree and mirrored by specific pathways and molecular biomarkers, can now be compared with diverse ongoing data collections in individuals with manifest PD, usually without a family history and with multifactorial pathogenesis (Scherzer et al., 2007; Marek et al., 2011; Santiago and Potashkin, 2015).

It is unlikely that diagnostics of PD risk can be based on a single technique or a single molecular biomarker, so we evaluated the global transcriptome and the pathway dysregulations by GSEA bioinformatics. These approaches recently permitted us to identify three gene expression changes in dependence of SNCA function in three gene expression changes in dependence of SNCA function in brain tissue, concerning the chemotaxis factor Lect1 (Cnmd) in one digenic and two monogenic PD mouse models (Gispert et al., 2015a) and the midbrain-selective transcription factor Foxp1 and the SNARE complex component Cplx1 in a synucleinopathy model and SNCA-deficient mouse mutants (Gispert et al., 2015b). Although we were unable to detect substantial FOXP1 and LECT1 expression in human fresh whole blood PAXgene samples, despite strong signals having been reported in microarray studies of purified peripheral mononuclear blood cells, we found mRNA complexin 1 levels to be downregulated in whole blood RNA-seq GSEA; GBA is known to represent of the immunity pathway upregulation, and on PLTP levels as representative of the lipid metabolism pathway. The data clearly confirmed their role as biomarkers of PARK4, at least in this Turkish pedigree, so it might now be interesting to test whether progressive stages of synucleinopathy can be correlated with expression levels of these factors or of other pathway components, whether these factors represent molecular targets of neuroprotective therapies or are risk factors themselves. Interestingly, a moderate mRNA upregulation for the lysosomal enzyme glucocerebrosidase (GBA; log2 fold change=0.31) was notable in the PARK4 blood RNA-seq GSEA; GBA is known to contribute to SNCA degradation, and its mutations act as modifiers of PD risk (Yap et al., 2011; Sidransky and Lopez, 2012; Fishbein et al., 2014). But there are many triggers of lysosomal induction and immunity responses beyond SNCA accumulation and PD, and this probably explains why GZMH and SPP1 upregulations are not specific for individuals with RBD but are also detectable among some control individuals.

Table 2. Significant association of the CPLX1 3’-UTR SNP rs10794536 G allele with PD risk

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Genotypic frequencies (%)</th>
<th>Allelic frequencies (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>GT</td>
</tr>
<tr>
<td>PD</td>
<td>360</td>
<td>12 (3.33)</td>
<td>116 (32.22)</td>
</tr>
<tr>
<td>Control</td>
<td>358</td>
<td>5 (1.40)</td>
<td>100 (27.93)</td>
</tr>
</tbody>
</table>

$\chi^2=4.97$ $P=0.083$

$\chi^2=4.16$ $P=0.041$

Approximately half of the PD and control samples were recruited at Düsseldorf and Marburg Universities, with both case-control collectives showing similar trends in the $\chi^2$ statistics for allele and genotype associations.
Despite the small effect size of CPLX1 mRNA downregulation, this observation reproducibly distinguished cohorts of presymptomatic PARK4 heterozygotes in blood, of individuals with RBD in blood, and neuroblastoma cells after SNCA transfection. The corresponding protein levels are difficult to quantify in blood because of the abundance of hemoglobin, albumin and fibrinogen, and because small fold changes are difficult to detect with antibody-dependent techniques and non-linear enzyme kinetcs. CPLX1 protein accumulated in neuroblastoma cells (proteins extracted with RIPA buffer) in our experiments and is known to accumulate together with SNCA in midbrain tissue (proteins extracted with 7 M urea, 2 M thiourea and 4% CHAPS), whereas it is decreased in the prefrontal cortex of individuals with sporadic PD (proteins extracted with 8 M urea; Basso et al., 2004; Dumitriu et al., 2016). Thus, the abundance of CPLX1 protein in fractions of different solubility appears to vary between different brain areas or in dependence on the progression of SNCA aggregation. We speculate that the accumulation of insoluble SNCA through protein interaction sequesters CPLX1 protein, leading to its accumulation in solubility, later in insolubility, whereas a homeostatic negative autoregulation reduces the CPLX1 mRNA levels. The decrease of CPLX1 mRNA has clear advantages as risk biomarker in peripheral tissue, whereas complex 1 protein redistribution might become a biomarker of disease progression in nervous tissue. The formation of Lewy bodies has been shown to deplete SNARE complex components from neurons, although it is technically difficult to demonstrate their presence within the protein inclusion bodies (Garcia-Reitbock et al., 2010; Mukaetova-Ladinska et al., 2013). The unequivocal co-localization of SNCA and CPLX1 in aggregates is challenging. Co-immunoprecipitation studies failed to detect a direct association, but the high detergent conditions needed to enrich SNCA oligomers and fibrils would not leave protein associations intact.

Thus, instead of pursuing the physical interaction between SNCA and CPLX1 proteins, we decided to assess their genetic interaction further. Two lines of evidence, (1) the identification of the CPLX1 gene variant rs10794536 that is associated with PD risk in a relatively small case-control study, and (2) the demonstration that genetic ablation of Cplx1 in mouse brain leads to an elevation of SNCA levels that is comparable to PARK4, both argue that complex 1 also has a role as risk factor in PD, not only a role as biomarker. The role of complex 1 appears to extend beyond monogenic synucleinopathy cases into sporadic PD, in view of (1) its accumulation in mass-spectrometry proteomics of PD midbrain autopsies, (2) the GWAS meta-analyses of risk loci for sporadic PD, and (3) the discovery of CPLX1 blood level downregulation in RBD cases.

It is noteworthy in this context that a recent report showed a knock-down of another SNARE protein, unphosphorylatable S187A-substituted SNAP25, to trigger increased abundance and mislocalization of SNCA, together with SNARE assembly problems and presynaptic disturbance, but without neurodegeneration until the mice were 1 year old. These authors concluded that dysfunctions at the SNARE complex in presynapses might induce elevated SNCA abundance and thus lead to an insidious neurodegenerative process (Nakata et al., 2012; Yasuda et al., 2013). This scenario is in good agreement with the recently identified genetic risk factors for idiopathic PD in individuals without monogenic traits. The reproducible GWAS loci of PD risk (Lill et al., 2012; Nalls et al., 2014) are surprisingly enriched in genes relevant for vesicle dynamics, such as the SYT11 (encoding synaptotagmin 11) and RAB25 genes at the chromosome 1q21.1-22 PD risk locus, the RAB29 gene at the chromosome 1q32 PARK16 locus, the NSF gene (encoding N-ethylmaleimide sensitive factor) within the risk haplotype at the chromosome 17q21 PD risk locus named after MAPT, and the SYT4 gene (encoding synaptotagmin 4) at the chromosome 18q12.3 PD risk locus named after RIT2. A remaining major GWAS PD risk locus contains the LRRK2 gene, which is responsible for autosomal dominant PARK8 and encodes a protein that modulates synaptic vesicle cycling and that interacts with syntaxin 1B/NSF/clathrin/AP2 (Piccoli et al., 2011). Thus, the SNCA and CPLX1 3′-UTR variants might be representative for various SNARE/RAB vesicle cycle alterations that modulate PD risk.

Overall, our initial characterization of an exceptionally large PARK4 pedigree from Turkey provided insights into the prodromal stage of monogenic synucleinopathy, defining a molecular signature in the blood transcriptome that might be relevant for predictive diagnostics in groups of at-risk individuals. Although several differences in expression between groups with presymptomatic PD versus controls were identified with statistical significance, individual differences at present would be insufficient for predictive diagnostics. Most importantly, the data identify CPLX1 as a biomarker and modifier gene of PD risk.

**MATERIALS AND METHODS**

**Neurological examination**

Neurological examination of members of the PARK4 family, which originated in the North Eastern region of Turkey, to document clinical manifestation or presymptomatic PARK4 heterozygote status adhered to the UK PD Society Brain Bank criteria (Hughes et al., 1992).

**Genotyping**

Genomic DNA from blood samples in EDTA tubes stored at +4°C was extracted with MagNA Pure Systems (Roche), after informed written consent and genetic counseling, with approval from the Ethics Commission of Bogaziçi University Istanbul. The molecular genetic diagnosis of PARK4 in an index individual with PD from this family (female aged 51 years with UPDRS (unified Parkinson’s disease rating scale) score 7 after a disease duration of 7 years, showing rigidity, resting tremor and bradykinesia, having 60 months of levodopa treatment with some benefit, but dyskinesias) was originally made by semi-quantitative multiplex PCR and multiplex ligation-dependent probe amplification (P051 kit) of SNCA exons 3 and 4 in comparison to Parkin and TTR exons 4, then extended to examine 22 pedigree members. Demonstration of SNCA gene duplication was complemented by the detection of the associated 4q21-3-4q22.3 haplotype of 17 microsatellite markers (D4S2691, D4S2409, D4S2462, D4S2622, D4S1542, D4S2929, D4S2371, D4S2461, D4S2304, D4S3459, D4S3457, D4S1544, D4S410, D4S1089, D4S414, D4S2404 and D4S2364) across 11 Mb around the SNCA locus in 12 pedigree members to demonstrate co-segregation. Further genotyping experiments were performed for the PARK4 family members using DNA probes specific to the SNCA gene and the quantitative real-time PCR method. Expression levels of the β-actin, GAPDH and TBP genes were separately used as endogenous controls for normalization. The gene expression fold changes were analyzed according to the 2−ΔΔCT method (Livak and Schmittgen, 2001).

**Molecular karyotyping**

Two PARK4 blood samples were assessed for copy number variation using a SurePrint G3 Human CGH (comparative genomic hybridization) microarray 2×400 K design 021850 (Agilent Technologies). Agilent’s labeling kit was used, and all procedures were carried out according to the manufacturer’s instructions. Scanning was performed on an Agilent microarray scanner and raw data were processed by Feature Extraction 9.5. Deleted and amplified regions were determined on Agilent microarray 2×400 K design 021850 (Agilent Technologies). Agilent's Workbench Standard Edition 5.0.14. A minimum of four consecutive pedigrees was analyzed in an index individual with PD from this family (female aged 51 years with UPDRS (unified Parkinson’s disease rating scale) score 7 after a disease duration of 7 years, showing rigidity, resting tremor and bradykinesia, having 60 months of levodopa treatment with some benefit, but dyskinesias) was originally made by semi-quantitative multiplex PCR and multiplex ligation-dependent probe amplification (P051 kit) of SNCA exons 3 and 4 in comparison to Parkin and TTR exons 4, then extended to examine 22 pedigree members. Demonstration of SNCA gene duplication was complemented by the detection of the associated 4q21-3-4q22.3 haplotype of 17 microsatellite markers (D4S2691, D4S2409, D4S2462, D4S2622, D4S1542, D4S2929, D4S2371, D4S2461, D4S2304, D4S3459, D4S3457, D4S1544, D4S410, D4S1089, D4S414, D4S2404 and D4S2364) across 11 Mb around the SNCA locus in 12 pedigree members to demonstrate co-segregation. Further genotyping experiments were performed for the PARK4 family members using DNA probes specific to the SNCA gene and the quantitative real-time PCR method. Expression levels of the β-actin, GAPDH and TBP genes were separately used as endogenous controls for normalization. The gene expression fold changes were analyzed according to the 2−ΔΔCT method (Livak and Schmittgen, 2001).

**Disease Models & Mechanisms**
Whole blood transcriptomics and protein analysis

After overnight fasting of the individuals, whole blood samples from the PARK4 family were taken (after informed written consent and approval from the Ethics Commission of Boğaziçi University Istanbul) into PAXgene (Qiagen) and EDTA tubes. Also after overnight fasting, whole blood samples were collected into PAXgene tubes from a RBD cohort comprising 46 affected individuals and 19 healthy relatives (individuals with narcolepsy were excluded), after written informed consent and approval with consent from the Ethics Commission at Marburg University. Samples from affected individuals and controls were processed in parallel, and after 2 h incubation of the PAXgene tubes at room temperature, they were stored at −80°C until the RNA was extracted after a further overnight room temperature incubation, using the corresponding PAXgene kit. The isolated RNA was cleaned up by DNase-I amplification grade (Invitrogen) and converted to cDNA by SuperScript III Reverse Transcriptase (Invitrogen). The quality and quantity of the cDNA were determined by spectrophotometry and RNA integrity numbers (Agilent Bioanalyzer). Quantitative RT-PCR was performed with TaqMan assays (Applied Biosystems) Hs01018439_s1 (GPRIN3), Hs00201182_m1 (MMRN1), Hs00411197_m1 (LRK2), Hs01084510_m1 (GIGYF2), Hs00223032_m1 (ATPIA3), Hs01062447_m1 (ATXN3), Hs00697109_m1 (DTI), Hs00260868_m1 (PINK1), Hs00832556_s1 (STL3), Hs01038318_m1 (PARK2), Hs00201825_m1 (FBX07), Hs00188233_m1 (UCHL1), Hs01616483_m1 (G4A), Hs00348383_m1 (HTRA2), Hs00185926_m1 (PLA2G6), Hs00608151_m1 (SNCA), Hs00705917_s1 (YWHAQ), Hs01133056_m1 (SNCA), Hs00793604_m1 (YWHAH), HS00355749_g1 (YWHAE), Hs00362510_m1 (CPLV), Hs00277212_m1 (GZMH), Hs00959010_m1 (SPP1), Hs00272126_m1 (PTP), Hs00230853_m1 (HNF4A) and Hs00914687_g1 (PTBP1). After studies of various loading controls (e.g. GAPDH, ACTN, RPL13A, TBP, SDH and YWHA) versus SNCa expression in whole blood, Hs09999910_m1 (TBp) was observed to show the smallest variance and was therefore included as the housekeeping control for each expression test. A total of 25 ng cDNA was used in each reaction. The PCR conditions were 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C and 1 min at 60°C. The gene expression changes were analyzed according to the 2^−ΔΔCT method (Livak and Schmittgen, 2001).

Whole blood protein analysis

For protein extraction from the EDTA tubes, 300 μl blood were lysed with an equal amount of 1% SDS-RIPA buffer [50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1% Igepal CA-630 (Sigma), 0.5% sodium deoxycholate, 0.1% SDS, 1 mM PMSF and one tablet Complete Protease Inhibitor Cocktail (Roche)] and sonicated for 10 s. The blood lysates were transferred to PVDF (polyvinylidene fluoride) membranes. The membrane was blocked in 5% milk powder and incubated with primary antibodies against α-synuclein (1:1000, BD 610786), β-synuclein (1:1000, Abcam ab1665 and Upstate 36-009), complex 1 (1:500, Acris APS1050PU-N) and β-actin (1:1000, Sigma, A5441), visualizing the signals by the ECL method. Densitometric assessment was carried out using ImageJ software.

Blood platelet flow cytometry

After applying a light tourniquet, which was immediately released after venipuncture, blood was drawn using a 21-gauge butterfly needle. Blood was collected in 4.5 ml coagulation tubes containing 0.106 mol/l sodium citrate (Sarstedt AG) for flow cytometry evaluations. The following directly conjugated monoclonal mouse anti-human antibodies were obtained from BD Biosciences: phycoerythrin (PE)-conjugated anti-CD62p, PE-conjugated anti-CD63 and (Per-Cy5)-conjugated anti-CD61. Platelet immunostaining was performed as follows: 20 μl diluted citrated whole blood [10:1, Dulbecco’s modified phosphate saline without Ca²⁺ and Mg²⁺ (DBPS), Gibco BRL] was incubated with saturating concentrations of monoclonal antibodies for 15 min at room temperature. Thereafter, samples were diluted with 1 ml DPBS and analyzed immediately by FACs Calibur flow cytometer (BD Biosciences). Platelets were identified through CD61 positivity and their characteristic light scattering. Each analysis was performed on 5000 platelets. Quantification of surface expression of given antigens was obtained using CellQuestPro software (BD Biosciences). The mean fluorescence intensity was used to define molecule expressions on the platelet surface. Results were compared with 20 healthy Central European volunteers (control group). To characterize the α-granule and lysosome release response upon platelet activation, 20 μl of 1:10 diluted citrated whole blood was stimulated with thrombin receptor agonist peptide (TRAP-6, final concentration 10 μmol/l, Roche) and DPBS (control) for 10 min at room temperature. Platelets were immunostained as described above.

Platelet electron microscopy

Approximately 16 ml of blood was collected in 8.2 ml coagulation tubes containing 0.106 mol/l sodium citrate (Sarstedt AG). Fresh platelet-rich plasma (PRP) was prepared by centrifugation at 140 g for 10 min at room temperature. After being stored for 30 min, PRP samples were stimulated with TRAP-6, final concentration 30 μmol/l, for 30 s. Fixation of suspended platelets was accomplished by combining the sample with an equal volume of 0.1% glutaraldehyde in White’s saline [a 10% solution of a 1:1 mixture of the following: (1) 2.4 mmol/l NaCl, 0.1 mmol/l KCl, 46 mmol/l MgSO4 and 64 mmol/l Ca(NO3)2.4H2O; and (2) 0.136 mol/l NaHCO3, 8.4 mmol/l NaH2PO and 0.1 g/l of phenol red, pH 7.4]. After 15 min the samples were centrifuged to pellets, washed and resuspended in 3% glutaraldehyde in the same buffer, re-suspended and maintained at 4°C for 30 min, then sedimented to pellets. The supernatant was removed and replaced with 1% osmic acid in distilled water containing 1.5% potassium ferrocyanide for 1 h at 4°C. All samples were dehydrated in a graded series of alcohols and embedded in Epon. Thin sections of 100 nm were cut from the plastic blocks on a Leica ultramicrotome and examined after staining with uranyl acetate and lead citrate to enhance contrast. Samples were examined on a Tecnai F30 electron microscope at 300 kV equipped with a US 4000 camera.

RNA-seq

Global blood RNA sequencing for the five PARK4 versus five age- and sex-matched healthy individuals was outsourced to Alacris Theranostics GmbH, Berlin, Germany. Strand-specific RNA sequencing libraries were prepared from 500 ng of total RNA according to the Illumina TruSeq stranded mRNA protocol. mRNA was selected using oligo(dT) Dynabeads. Sequencing was performed on the Illumina HiSeq2500 platform, using paired-end 2×50 bp sequencing mode. For each library, 40-50 million passed filter reads were obtained. Raw data were processed according to the manufacturer’s instructions. Illumina’s bcl2fastq software v1.8.2 was used for base calling and demultiplexing. The sequencing reads were mapped to the human genome genome reference GRCh37/hg19 using bowa (version 0.5.9, http://bio-bwa.sourceforge.net/). Exonic reads in correct genomic orientation, defined by Ensembl Release 73 (http://www.ensembl.org/index.html), were counted using custom Python scripts. Expression values were calculated using RPKM (reads per kilobase per million mapped reads) normalization with custom R scripts (Mortazavi et al., 2008). Given that we observed huge amounts of hemoglobin in the samples, on hemoglobin genes were removed before the expression analysis. Differential expression between disease and control samples was calculated with the Bioconductor package edgeR (Robinson et al., 2010; http://www.bioconductor.org/packages/release/bioc/html/edgeR.html), applying an exact test for overdispersed data. As an additional criterion to judge variance and restrict the number of biomarker candidates, a custom R script was used to calculate Student’s paired t-tests between disease and control expression values for each gene.

Bioinformatic systems biology

RNA-seq data from five presymptomatic PARK4 heterozygotes and five matched controls were analyzed to compute average fold changes for each gene and rank them according the t-test statistics. For duplicate entries, the maximal value was used. The ranked list of gene symbols thus generated

RESEARCH ARTICLE

was subjected to nonspecific filtering and assessed by GSEA using the Java-based version GSEA-P (Subramanian et al., 2005, 2007). Permutations were performed on gene sets owing to the low number of biological replicates. We used the c2 (online pathway databases, Pubmed publications, expert of domain knowledge), c5 (Gene Ontology categories) and cc (Gene Ontology cellular component) gene sets from the MSigDB database (v.4.0, May 2013, http://www.broadinstitute.org/gsea/msigdb/index.jsp) to analyze the data sets.

**Statistical analyses**

Statistical analyses presented in bar graphs were performed using Student’s unpaired t-tests, those of CPLX1 level association with age or sex were performed with linear regression, runs test and two-way ANOVA, all being conducted and plotted with Prism 5.04 software (GraphPad, La Jolla, CA, USA). The allele association studies were carried out after excluding deviations from Hardy–Weinberg equilibrium by χ² statistics for genotype and allele distribution.

This article is part of a special subject collection ‘Neurodegeneration: from Models to Mechanisms to Therapies’, which was launched in a dedicated issue guest edited by Aaron Giller and James Shorter. See related articles in this collection at http://dmm.biologists.org/collection/neurodegenerative-disorders.

**Acknowledgements**

We thank B. Meseck-Seichow, M. Babi, M. Bouzrou, K. Weckmann, C. Hoett and E. Gerlach for technical assistance, the staff of the ZFE at the University Hospital Frankfurt for support, A. Ohle at the Animal Facility at the MPI for Experimental Medicine for mouse husbandry, and Tim Becker at the IMSDD of Bonn University. Manu Sharma at Tübingen University and Moritz Schütte at Alacris Theranostics in Berlin for statistical advice. We are grateful to Herbert Zimmermann for critical reading, and to the Turkish PARK4 family and the donors of PD and control DNA and brain autopsies for their cooperation.

**Competition finances**

The authors declare no competing or financial interests.

**Author contributions**

S.L.: organization and execution of research and statistical analysis; writing of the first draft. S.G.: design of research and statistical analysis; review and critique. O.O.: organization and execution of research and statistical analysis. H.T.: clinical organization and execution of research and statistical analysis. S.G.: review and critique. A.N.B.: organization and design of research; review and critique. W.O.: organization and design of research; review and critique. G.A.: organization and design of research; review and critique. The rest of the authors have contributed to execution of research and statistical analysis.

**Funding**

The study was financed by funds of the University Hospitals Frankfurt and Marburg, the Bundesministerium für Bildung und Forschung (BMBF) (NGFNplus Parkinson Net, Project 8.2.7), Deutsche Forschungsgemeinschaft (DFG) (GI 3423/1-1), Suna and Ian Kroc Foundation; Boğaziçi University Research Funds and International Parkinson Fonds, The Netherlands. W. H. Oertel is a Hertie Senior Research Professor (2014-2019), supported by the Charitable Hertie Foundation, Frankfurt/ Main, Germany.

**Data deposition**

All RNA-seq data are deposited in the European Nucleotide Archive (ENA) public database under accession numbers PRJEB8960 and ERP010003.

**Supplementary information**

Supplementary information available online at http://dmm.biologists.org/lookup/doi/10.1242/dmm.028035.supplemental

**References**


