


## Article

# *C9orf72*-G<sub>4</sub>C<sub>2</sub> Intermediate Repeats and Parkinson's Disease; A Data-Driven Hypothesis

Hila Kobo<sup>1,2,†</sup>, Orly Goldstein<sup>1,†</sup>, Mali Gana-Weisz<sup>1</sup>, Anat Bar-Shira<sup>1</sup>, Tanya Gurevich<sup>2,3</sup>, Avner Thaler<sup>2,3,4</sup>, Anat Mirelman<sup>2,3,4</sup>, Nir Giladi<sup>2,3,4</sup> and Avi Orr-Urtreger<sup>1,2,\*</sup> 

<sup>1</sup> The Genomic Research Laboratory for Neurodegeneration, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv 64239, Israel; Hilakobo@tauex.tau.ac.il (H.K.); orlyg@tlvmc.gov.il (O.G.); maligw@tlvmc.gov.il (M.G.-W.); anatbn@tlvmc.gov.il (A.B.-S.)

<sup>2</sup> Sackler Faculty of Medicine, Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel; tanyag@tlvmc.gov.il (T.G.); avnert@tlvmc.gov.il (A.T.); anatmi@tlvmc.gov.il (A.M.); nirg@tlvmc.gov.il (N.G.)

<sup>3</sup> Movement Disorders Unit, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv 64239, Israel

<sup>4</sup> Laboratory for Early Markers of Neurodegeneration, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv 64239, Israel

\* Correspondence: aviorr@tlvmc.gov.il; Tel.: +972-3-6974704

† The two authors contributed equally to the work.

**Abstract:** Pathogenic *C9orf72*-G<sub>4</sub>C<sub>2</sub> repeat expansions are associated with ALS/FTD, but not with Parkinson's disease (PD); yet the possible link between intermediate repeat lengths and PD remains inconclusive. We aim to study the potential involvement of these repeats in PD. The number of *C9orf72*-repeats were determined by flanking and repeat-primed PCR assays, and the risk-haplotype was determined by SNP-array. Their association with PD was assessed in a stratified manner: in PD-patients-carriers of mutations in *LRRK2*, *GBA*, or *SMPD1* genes ( $n = 388$ ), and in PD-non-carriers (NC,  $n = 718$ ). Allelic distribution was significantly different only in PD-NC compared to 600 controls when looking both at the allele with higher repeat's size ( $p = 0.034$ ) and at the combined number of repeats from both alleles ( $p = 0.023$ ). Intermediate repeats (20–60 repeats) were associated with PD in PD-NC patients ( $p = 0.041$ ; OR = 3.684 (CI 1.05–13.0)) but not in PD-carriers ( $p = 0.684$ ). The *C9orf72* risk-haplotype, determined in a subgroup of 588 PDs and 126 controls, was observed in higher frequency in PD-NC (dominant model, OR = 1.71, CI 1.04–2.81,  $p = 0.0356$ ). All 19 alleles within the risk-haplotype were associated with higher *C9orf72* RNA levels according to the GTEx database. Based on our data, we suggest a model in which intermediate repeats are a risk factor for PD in non-carriers, driven not only by the number of repeats but also by the variants' genotypes within the risk-haplotype. Further studies are needed to elucidate this possible role of *C9orf72* in PD pathogenesis.

**Keywords:** Parkinson's disease (PD); *C9orf72*; intermediate repeats; hexanucleotide expansions



**Citation:** Kobo, H.; Goldstein, O.; Gana-Weisz, M.; Bar-Shira, A.; Gurevich, T.; Thaler, A.; Mirelman, A.; Giladi, N.; Orr-Urtreger, A. *C9orf72*-G<sub>4</sub>C<sub>2</sub> Intermediate Repeats and Parkinson's Disease; A Data-Driven Hypothesis. *Genes* **2021**, *12*, 1210. <https://doi.org/10.3390/genes12081210>

Academic Editor: Suzanne Lesage

Received: 6 July 2021

Accepted: 3 August 2021

Published: 5 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Parkinson's disease (PD) is a common neurodegenerative disorder, affecting about 2% of the elderly population worldwide [1]. Its complex genetic background has been revealed in the past decades, implicating many genes associated with the disease. A wide variety of genetic changes and mechanisms are involved in PD, including rare and common variants, recessive, dominant, and oligogenic inheritance, and epigenetics [2–8]. However, the full range of genetic changes in PD is still evolving.

G<sub>4</sub>C<sub>2</sub> Hexanucleotide repeat expansions in *C9orf72* are strongly associated with amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) [9,10], mostly in European and North American populations [11,12]. Although 30 repeats and over are considered pathogenic for ALS and FTD, most patients that are *C9orf72*-associated-ALS or -FTD, carry an expanded allele with hundreds, or even thousands, of repeats [12–14]. Interestingly, parkinsonism was observed in more than 40% of FTD and FTD/ALS patients with

pathogenic *C9orf72* expansions [15]. This observation has led researchers to investigate the possible association of *C9orf72* expansions with PD.

While rare cases of PD patients with 30 to 60 repeat expansions or more in *C9orf72* were detected (<0.7%), there was no association with PD [16–20]. Few studies have suggested that intermediate-size repeat lengths in *C9orf72* may be a risk factor for PD; however, these studies suggested different repeat lengths for this association:  $\geq 7$  repeats in Han Chinese [21], and  $\geq 20$  repeats in Caucasians [17]. In a multi-center meta-analysis of mostly Caucasian and Asian populations, the pathogenicity expansions threshold was determined as >60 repeats, and the intermediate repeat size was set as 17–60 repeats [22], suggesting an effect on PD-risk for the cutoff of 17 repeats, and even as little as 10 repeats as a stand-alone or as a cutoff. In a recent comprehensive review by Bourinaris and Houlden [23], *C9orf72* intermediate repeat lengths were reported in several parkinsonism and movement disorders, including in Dopa-responsive PD, atypical parkinsonisms including PSP and MSA, essential-tremor plus parkinsonism, and spinocerebellar ataxia.

As previous studies have shown the pivotal role of genetically homogeneous populations, such as the Ashkenazi Jews (AJ), in understanding the genetic background of neurodegenerative diseases [24–27], we hereby determined *C9orf72* repeats' size in PD patients of Ashkenazi origin and examined their potential association with PD. We also studied the possible association of the shared risk-haplotype, which is observed in carriers of intermediate repeats, with PD-risk.

## 2. Methods

### 2.1. Population

This study included a cohort of consecutively recruited unrelated 1106 PD patients of full Ashkenazi Jewish origin (Table 1). Patients were recruited between the years 2005 and 2015. The diagnostic criteria, recruitment, and genotyping for *LRRK2*, *GBA*, and *SMPD1* mutations have been previously described [24,25,28]. Carrier patients (PD-carriers) were defined with one or more of the following mutations in *LRRK2* (p.G2019S), *SMPD1* (p. L302P), or any of the 10 *GBA* mutations (c.84insG, IVS2 + 1G > A, p.V394L, p.N370S, p.L444P, p.R496H, p.E326K, p.T369M, p.R44C, and 370Rec). Patients who did not carry any of these mutations were determined as non-carriers (PD-NC). The cohort of 600 ethnically matched control individuals used in this study has been previously described [26].

**Table 1.** Characteristics of the 1106 Parkinson's disease patients of Ashkenazi Jewish origin.

	Non-Carrier PD Patients <sup>a</sup>	PD Patients Carriers of <i>LRRK2</i> , <i>GBA</i> , or <i>SMPD1</i> Mutations
N	718	388 <sup>c</sup>
Women, N (%)	266 (37.0)	171 (44.1)
AAE, mean (SD), y	68.5 (10.2) <sup>d</sup>	65.4 (10.1)
AAO, mean (SD), y	61.4 (11.5)	58.4 (10.6)
Family history of PD <sup>b</sup> , N (%)	144 (20)	120 (30.9)

Abbreviations: PD, Parkinson disease; N, number of individuals; AAE, age at enrollment; SD, standard deviation; y, years; AAO, age at disease onset. <sup>a</sup> Patients without the *LRRK2*, *GBA*, or *SMPD1* mutations (specified in the Methods section). <sup>b</sup> Patients with 1st or 2nd degree family members with PD. <sup>c</sup> Carrier patients included: 133 individuals with *LRRK2* mutation, 223 individuals with *GBA* mutations, 8 individuals with *SMPD1* mutation, 23 individuals with mutations in both *LRRK2* and *GBA*, and 1 patient with both *GBA* and *SMPD1* mutations. <sup>d</sup> Data regarding AAE was not available for one individual.

### 2.2. Determining the G<sub>4</sub>C<sub>2</sub> Hexanucleotide Repeat Length in the *C9orf72* Gene

To determine the number of repeats in *C9orf72*, flanking and repeat-primed PCR assays were performed as previously described [9,10], with some modifications [26]. This method detects all repeats expansion but can determine the number of repeats up to 55 repeats. Therefore, in all individuals that carried an allele with 30 repeats or more, an additional method was used to determine accurately the repeat number up to 145 repeats (Asuragen assay kit AmpliX<sup>®</sup> PCR/CE *C9orf72* Kit; Asuragen Genetics; Austin, TX, USA).

### 2.3. Assembly of the Risk-Haplotype within the *C9orf72* Locus

To determine the presence of the risk-haplotype in *C9orf72*-locus, we used the genotype data (from Vacic et al. [29], Affymetrix SNP6.0) of 127 AJ controls (all part of our cohort of 600 controls) and 597 AJ-PD patients (594 are part of our cohort of 1106 PD patients). When determination of the presence of the risk-haplotype was impossible due to missing genotypes, these individuals were excluded (Control = 1, PD = 5). One PD was excluded due to low genotype rate, one was not tested for repeat size, and two PDs who carried >145 repeats were also excluded. In total, 126 controls and 588 PD patients were analyzed.

### 2.4. Statistical Analyses of *C9orf72* G<sub>4</sub>C<sub>2</sub> Hexanucleotide Repeats

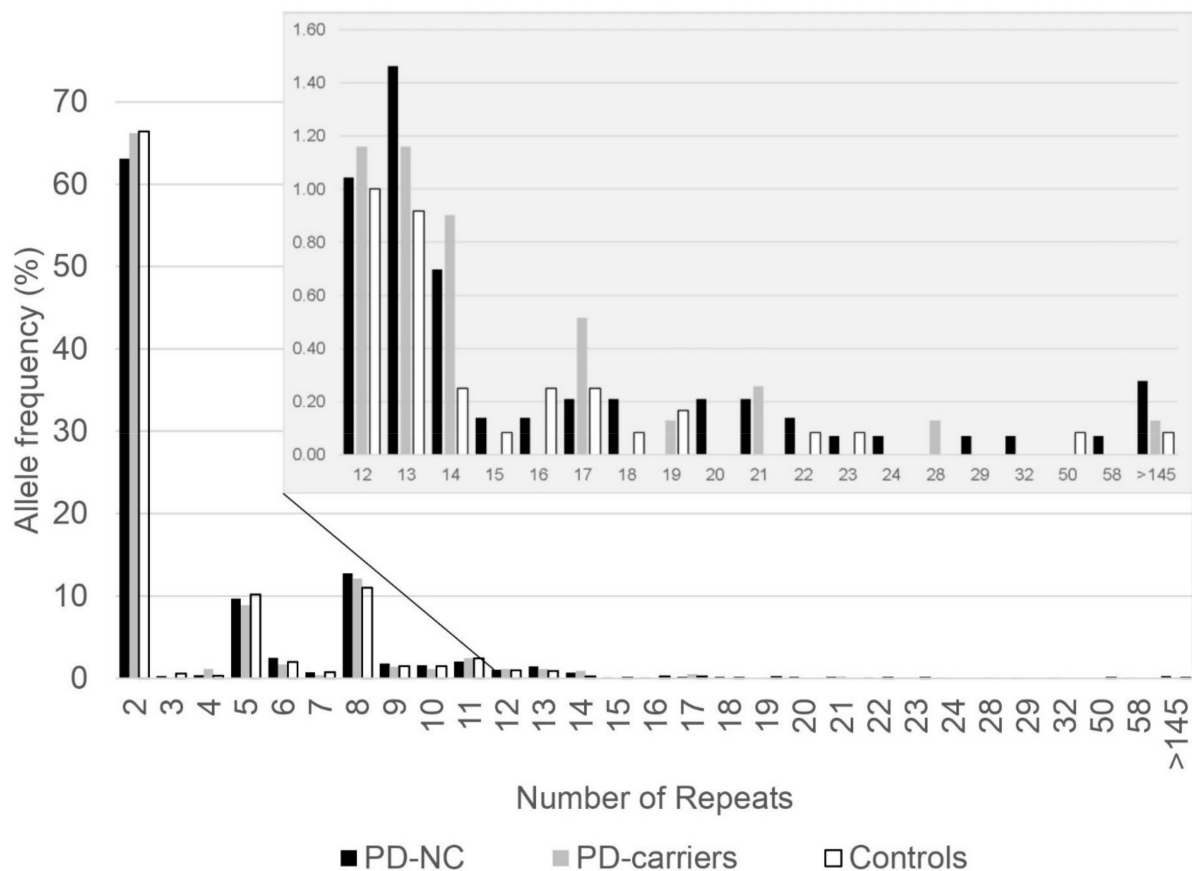
All statistical analyses were performed using IBM SPSS statistics software v25 (IBM Corporation, New York, USA). Differences in continuous variables were tested using Mann–Whitney U test or *t*-test (2-tailed). To test the difference in *C9orf72* repeat lengths between patients and controls, both alleles repeat sizes (per individual) were included.

Categorical variables were compared using 2-sided  $\chi^2$ -test, or Fisher's exact test when numbers were less than 5. Odds ratio (OR), with 95% confidence interval (CI), was applied to assess the association of *C9orf72* G<sub>4</sub>C<sub>2</sub> repeat lengths with PD. This association was examined using the longest repeat size (per individual) as the independent variable. Association of the risk-haplotype with PD was examined using a dominant model. Logistic regression analysis was performed when using repeat units as a quantitative trait (the largest allele or the sum of both alleles).

## 3. Results

### 3.1. Allele Frequencies of *C9orf72* G<sub>4</sub>C<sub>2</sub> Hexanucleotide Repeats in Ashkenazi PD Patients

Allele frequencies of *C9orf72* repeats were determined in our cohort of 1106 Ashkenazi PD patients (Table 1), that was divided into two groups based on their genotypic status, either carriers of PD-associated mutations (see Methods section), or non-carrier patients (PD-NC). We ran a stratified analysis based on the carrier status in *LRRK2*, *GBA*, and *SMPD1*, as a high percentage of our PD cohort carry risk alleles in these 3 genes (35.1%, 388/1106), and these carrier-patients may mask the effect of the hexanucleotide repeat length on PD-risk in non-carrier patients. The most frequent alleles found were 2, 8 and 5 repeat units (66.2%, 12.1%, 8.9% in carrier, and 63.1%, 12.7%, 9.7% in non-carrier patients; Figure 1 and Supplement Table S1). These alleles were also shown as the most common alleles in our previously published data of Ashkenazi controls (66.4%, 11.0%, and 10.2%) [26]. No significant difference in allele distribution was observed between patients with mutations (in *LRRK2*, *GBA*, or *SMPD1* genes) and controls (Mann–Whitney *U* test  $p = 0.756$ ;  $4.23 \pm 6.13$ ,  $N = 776$  and  $4.09 \pm 5.36$ ,  $N = 1200$ , respectively). However, a significant difference in allele distribution was detected in PD-NC (Mann–Whitney *U* test  $p = 0.034$ ;  $4.71 \pm 8.40$ ,  $N = 1436$ ). This was also significant for the total number of repeats (combining the numbers of repeats from both alleles; excluding individuals with expanded alleles of >145 repeats) in PD-NC (Mann–Whitney *U* test  $p = 0.023$ ;  $8.63 \pm 5.57$ ,  $N = 714$ , and  $7.94 \pm 5.01$ ,  $N = 599$ ) and was not significant in PD-carriers (Mann–Whitney *U* test  $p = 0.565$ ;  $8.10 \pm 4.90$ ,  $N = 387$ ).



**Figure 1.** Graphic representation of *C9orf72* G<sub>4</sub>C<sub>2</sub> hexanucleotide repeat allele frequencies in Ashkenazi Parkinson's disease patients and controls. The repeats' allele frequencies for each group are presented: PD patients without *LRRK2*, *GBA*, and *SMPD1* mutations (PD-non-carriers, black); PD patients carrying mutation in *LRRK2*, *GBA*, or *SMPD1* (PD-carriers, gray); and controls as previously published (Reference [26], white). Upper panel is a zoomed-in graph of 12 repeats and higher.

### 3.2. The Association of *C9orf72* G<sub>4</sub>C<sub>2</sub> Hexanucleotide Intermediate Repeat Lengths with PD in Ashkenazi Patients

We examined the association of the longest repeat allele in each individual with PD (in PD-carriers and PD-NC). First, we examined whether large expansion lengths (>145 repeats) are present within our cohort of PD patients and controls. We found one PD-carrier patient (carrying the *LRRK2* p.G2019S mutation, 1/388 = 0.3%), four PD-NC (4/718 = 0.6%), and one control (1/600 = 0.2%) that carried an expanded allele. No significant association was detected in any of the patients' groups compared to controls (Fisher's Exact Test  $p = 1.000$  and  $p = 0.384$ , respectively). Of interest, one of these PD patients was later diagnosed with ALS, and two had dementia. For further analysis, we excluded these five carriers with *C9orf72* G<sub>4</sub>C<sub>2</sub> repeats expansion (all with > 145 repeats). None of our patients or controls carried alleles between 60 and 145 repeats.

Next, we examined if intermediate *C9orf72* hexanucleotide repeats (20–60 repeats) were associated with PD in our cohort. Among the 1101 PD patients, 3 PD-carriers (0.8%, 3/387) and 13 PD-NC (1.8%, 13/714) carried an allele with intermediate repeat lengths, compared to 3 among the controls (0.5%, 3/599). An association with increased risk for PD was observed in PD-NC patients compared to controls (Fisher's exact test,  $p = 0.041$ ; OR = 3.684, CI = 1.045–12.990; Table 2, Figure 1), but no association was detected when comparing PD-carrier patients to controls ( $p = 0.684$ , Table 2).

**Table 2.** The association 2–60 *C9orf72* hexanucleotide repeats with Parkinson’s disease.

Cohort	Non-Carrier PD Patients <sup>a</sup>	PD Patients Carriers of <i>LRRK2</i> , <i>GBA</i> , or <i>SMPD1</i> Mutations	Controls <sup>b</sup>
2–19 repeats, N (%)	701 (98.2)	384 (99.2)	596 (99.5)
20–60 repeats, N (%)	13 (1.8)	3 (0.8)	3 (0.5)
Odds Ratio (95% CI)	3.684 (1.045–12.990)	1.552 (0.312–7.729)	
<i>p</i> -value <sup>c</sup>	0.041	0.684	

Abbreviations: PD, Parkinson’s disease; N, Number of individuals; CI, confidence interval. The longest allele in each individual was recorded. <sup>a</sup> Patients without the *LRRK2*, *GBA*, or *SMPD1* mutations (specified in the Methods section). <sup>b</sup> The values in the control cohort were previously published [26]. <sup>c</sup> 2 × 2 Fisher’s exact test (2-sided).

Measuring the effect of the risk associated with an increasing number of repeats ( $\beta$ , regression analysis for each one repeat unit) shows a significant effect in PD-NC when looking either at the largest allele or at the sum of alleles ( $\beta = 0.032$ ,  $p = 0.015$ , OR = 1.032, CI = 1.006–1.059;  $\beta = 0.025$ ,  $p = 0.021$ , OR = 1.025, CI = 1.004–1.047; respectively). No effect was observed in PD-carriers ( $\beta = 0.009$ ,  $p = 0.542$ , OR = 1.01, CI = 0.979–1.041;  $\beta = 0.007$ ,  $p = 0.617$ , OR = 1.007, CI = 0.981–1.033; respectively). No effect on age of motor symptoms onset was observed in PD-NC ( $p$  value = 0.978).

### 3.3. The *C9orf72* Risk-Haplotype Is Associated with Higher RNA Expression Levels and with PD

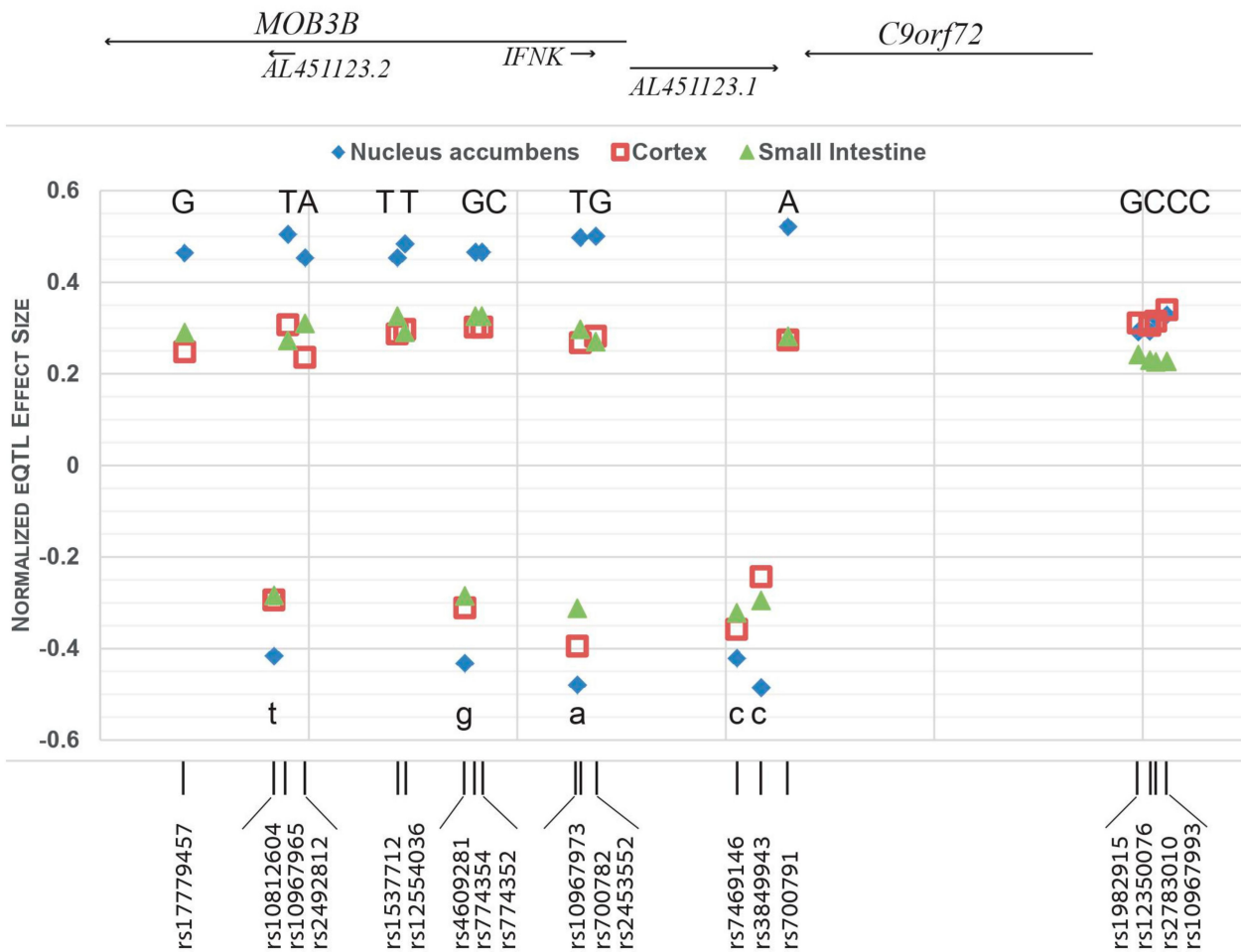
We have previously shown that the *C9orf72* expansions in AJ shared a risk-haplotype that expands 107Kb (Goldstein et al. [26], hg19: chr9:27484575-27591569), encompasses the complete *C9orf72* gene, and includes 44 informative single nucleotide variants (SNVs), with significant association with higher number of repeats (over 8 repeats). We further examined here the effects of these 44 SNVs (within the risk haplotype) on the RNA expression levels (eQTL) and splicing (sQTL) of their adjacent genes, as reported by the Genotype-Tissue Expression (GTEx) project [30]. Of them, 11 SNVs had no QTLs, 4 had low effect size (absolute NES < 0.2), and one had no significant effect ( $m$ -value < 0.9). Nine other variants were also excluded due to high allele frequency in AJs (> 0.5 in gnomAD v3.1 database of non-neuro cases). The other 19 SNVs within this 107Kb risk-haplotype were all associated with higher *C9orf72* RNA expression levels compared to the expression levels of the non-risk alleles (Table 3, Figure 2). Cerebellum and Nucleus-accumbens were the tissues with the highest normalized effect size (NES). Moreover, as GTEx evaluates the effect of each SNV on the neighboring genes within a 2 Mb interval (1 Mb upstream and 1 Mb downstream), it is important to note that all 19 SNVs affected exclusively *C9orf72*-RNA levels and not any other genes within that 2 Mb window. These SNVs also affected splice variants (sQTL) in an exclusive manner, only for *C9orf72*, mainly in cerebellum (Table 3).

We, therefore, tested if carrying the risk-haplotype is associated with PD. Genotyping data from 127 AJ controls and 597 AJ-PD patients were assembled (see Methods), and the presence of the 19-SNVs-risk-haplotype was determined. Overall enrichment of the risk-haplotype was observed in PDs compared to controls: 167 out of 588 PDs carried one or two copies of the risk-haplotype (28.4%) compared to 24 out of 126 controls (19.0%). When stratifying based on mutation carrier status (PD-carriers and PD-NC), a significant association was detected in PD-NC: 28.6% of them carried one or two copies of the risk-haplotype (OR = 1.71, CI = 1.04–2.81,  $p = 0.0356$ ), and tendency was shown in PD-carriers (28.0%, OR = 1.65, CI = 0.97–2.82,  $p = 0.0656$ ).

**Table 3.** The effect of the 19 variations within the risk-haplotype on RNA expression levels and splice variants as reported by GTEx database.

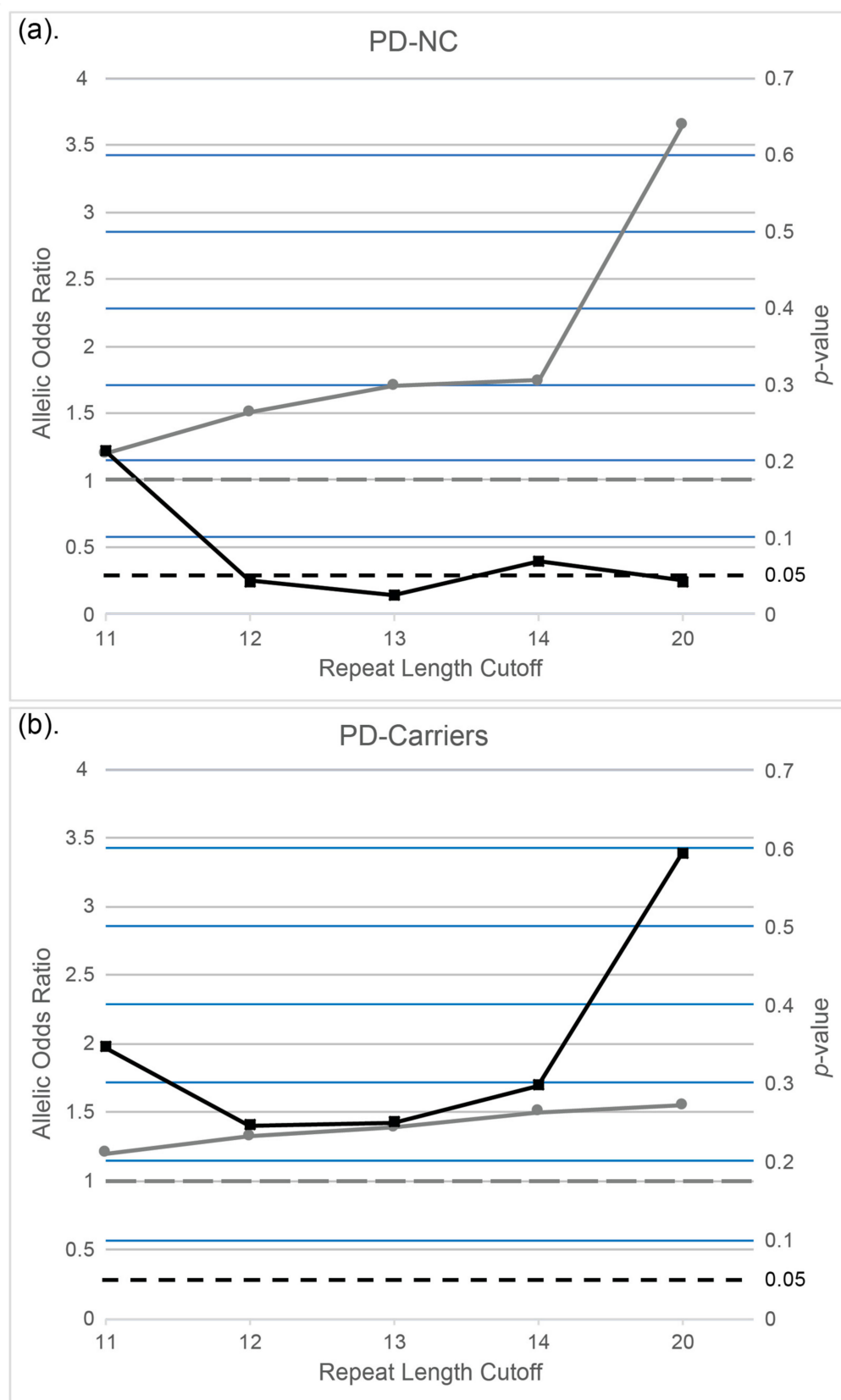
Location, chr9 (hg38)	rs ID	Gene	Risk Allele	Ref > Alt <sup>a</sup>	Highest eQTL for <i>C9orf72</i> , NES (Tissue)	eQTL for <i>C9orf72</i> in Cerebellum, NES	Highest sQTL for <i>C9orf72</i> , NES (Tissue)	GnomAD v3.1- AF for the Risk Allele (in Non-Neuro Cases)		
								AJ	European (Non-Finnish)	All Populations
27488094	rs17779457	<i>MOB3B</i>	<b>G</b>	<b>T &gt; G</b>	<b>0.464 (N.A.)</b>	<b>0.341</b>	−0.88 (Cer)	0.2097	0.2428	0.2490
27496663	rs10812604	<i>MOB3B</i>	<b>T</b>	<b>T &gt; G</b>	−0.416 (N.A.)	−0.343	0.70 (Cer)	0.2252	0.2861	0.3555
27497990	rs10967965	<i>MOB3B</i>	<b>T</b>	<b>A &gt; T</b>	0.505 (N.A.)	0.360	−0.96 (C.H.)	0.1742	0.2185	0.1547
27499629	rs2492812	<i>MOB3B</i>	<b>A</b>	<b>C &gt; A</b>	0.454 (N.A.)	0.334	−0.88 (C.H.)	0.2091	0.2419	0.2486
27508491	rs1537712	<i>MOB3B</i>	<b>T</b>	<b>C &gt; T</b>	0.454 (N.A.)	0.332	−0.75 (Cer)	0.2323	0.2753	0.2915
27509213	rs12554036	<i>MOB3B</i>	<b>T</b>	<b>G &gt; T</b>	0.484 (N.A.)	0.345	−1.0 (C.H.)	0.2054	0.2335	0.1796
27514964	rs4609281	<i>MOB3B</i>	<b>G</b>	<b>G &gt; T</b>	−0.432 (N.A.)	−0.364	0.71 (Cer)	0.2417	0.2865	0.3509
27515969	rs774354	<i>MOB3B</i>	<b>G</b>	<b>A &gt; G</b>	0.466 (N.A.)	0.345	−0.80 (C.H.)	0.2354	0.2754	0.2909
27516592	rs774352	<i>MOB3B</i>	<b>C</b>	<b>T &gt; C</b>	0.466 (N.A.)	0.345	−0.80 (C.H.)	0.2354	0.2755	0.2910
27525753	rs10967973	<i>IFNK</i>	<b>A</b>	<b>A &gt; G</b>	−0.602 (Cer)	−0.602	0.66 (Cer)	0.3570	0.4640	0.5599
27526049	rs700782	<i>IFNK</i>	<b>T</b>	<b>C &gt; T</b>	0.498 (N.A.)	0.370	−0.98 (C.H.)	0.2109	0.2440	0.2461
27527514	rs2453552	<i>MOB3B</i>	<b>G</b>	<b>T &gt; G</b>	0.501 (N.A.)	0.36	−1.0 (C.H.)	0.2110	0.2450	0.2650
27541043	rs7469146	<i>C9orf72</i>	<b>C</b>	<b>C &gt; T</b>	−0.542 (Cer)	−0.542	0.65 (Cer)	0.3736	0.5119	0.5916
27543384	rs3849943	<i>C9orf72</i>	<b>C</b>	<b>C &gt; T</b>	−0.485 (N.A.)	−0.336	0.96 (Cer, C.H.)	0.2005	0.2362	0.2178
27545962	rs700791	<i>C9orf72</i>	<b>A</b>	<b>C &gt; A</b>	0.521 (N.A.)	0.388	−1.1 (C.H.)	0.1958	0.2293	0.1979
27579562	rs1982915	Intergenic	<b>G</b>	<b>A &gt; G</b>	0.407 (Cer)	0.407	−0.49 (Cer)	0.4212	0.4982	0.4745
27580676	rs12350076	Intergenic	<b>C</b>	<b>A &gt; C</b>	0.408 (Cer)	0.408	−0.51 (Cer)	0.4262	0.5104	0.4964
27581241	rs2783010	Intergenic	<b>C</b>	<b>T &gt; C</b>	0.418 (Cer)	0.418	−0.56 (Cer)	0.4273	0.4645	0.4808
27582313	rs10967993	Intergenic	<b>C</b>	<b>T &gt; C</b>	0.424 (Cer)	0.424	−0.54 (F.C.)	0.4025	0.5044	0.5032

<sup>a</sup> In bold is the allele that is associated with higher *C9orf72* RNA expression levels, as reported by GTEx database. Ref = reference allele; Alt = Alternate allele; eQTL = expression quantitative trait loci; sQTL = spliced quantitative trait loci; NES = normalized effect size; AF = Allele frequency; AJ = Ashkenazi Jews; N.A. = Nucleus accumbens; Cer = Cerebellum; C.H. = Cerebellar hemisphere; F.C. = Frontal Cortex.



**Figure 2.** Higher levels of *C9orf72* RNA expression for all 19 SNVs within the risk-haplotype: Normalized eQTL effect size in Nucleus accumbens (diamond), Cortex (square), and small intestine (triangle). Upper panel and upper-case letters are the alternate allele observed in the risk haplotype, associated with higher levels of *C9orf72* expression; lower panel and lower-case letters are the reference allele observed in the risk haplotype, associated with higher levels of *C9orf72* expression.

We also attempted to define the correlation between the existence of risk-haplotype and number of repeats, by looking at all alleles ( $n = 1428$  alleles): 100% negative correlation existed between the risk-haplotype and the 2-repeats' allele (with zero percent risk-haplotype), and 100% positive correlation was observed in carriers of 14–60 repeats (100% carried the risk-haplotype, Supplement Figure S1). Thus, we used 14 repeats as the best assessor for carrying the risk-haplotype and re-calculated the association of 14 repeats and higher with PD in our cohorts. Among all alleles in PD-NC, 2.3% had 14–60 repeats, compared to only 1.3% in controls, showing a trend toward significance (Figure 3a, OR = 1.75, CI = 0.96–3.19,  $p = 0.069$ , uncorrected), with no significance observed in PD-carriers (Figure 3b, OR = 1.46, CI = 0.72–2.97,  $p = 0.296$ ). Further analysis showed a significant association in PD-NC when the cutoff of 13 repeats or 12 repeats was selected (OR = 1.70, CI = 1.07–2.72,  $p = 0.0257$  and OR = 1.51, CI = 1.01–2.25,  $p = 0.044$ , respectively, Figure 3a, uncorrected), while there was no significant association in PD-carriers (OR = 1.46, CI = 0.72–2.97,  $p = 0.296$ ; OR = 1.39, CI = 0.79–2.42,  $p = 0.249$  and OR = 1.32, CI = 0.83–2.12,  $p = 0.245$ , respectively, Figure 3b). No significance was shown at 11 repeats cutoff, both in PD-NC and PD-carriers (Figure 3, OR = 1.22, CI = 0.89–1.69,  $p = 0.214$  and OR = 1.20, CI = 0.82–1.74,  $p = 0.345$ , respectively).



**Figure 3.** The effect of each cutoff of *C9orf72*-repeats-size on Parkinsons' disease risk in non-carriers of mutations ((a), PD-NC) and in carriers of mutation ((b), PD-Carriers; see Methods). Grey lines (circles) depict the Odds ratio and black lines (squares) depict the *p*-value. Dashed grey lines represent Odd Ratio of 1.0, and dashed black lines represent *p*-value of 0.05.



#### 4. Discussion

More than 40 diseases, most of which affect the nervous system, were identified with a genetic basis of expansions of simple short DNA sequence (reviewed by Reference [31]). Among these diseases are myotonic dystrophy, Huntington's disease, spinocerebellar ataxia (SCA), spinal and bulbar muscular atrophy (SBMA), and Fragile-X syndrome. A common finding in these disorders is the correlation of the number of repeats with clinical phenotypes and penetrance. Interestingly, emerging studies suggest that expanded repeats and intermediate repeats can cause or act as risk factors for different neurological diseases, depending on the number of repeats. This was suggested for the *FMR1* gene, when over 200 CCG repeats cause mental retardation (Fragile-X syndrome), while the premutations of 45–200 repeats are a risk factor for Fragile-X-tremor/ataxia syndrome (FXTAS) in males and premature ovarian failure 1 (FXPOI) in females. Other examples are *ATXN2*, in which more than 34 repeats cause SCA2, while 29–33 are risk factor for ALS [32,33], and *ATXN1*, in which more than 38 repeat cause SCA1, while  $\geq 33$  are risk factor for ALS, mostly in *C9orf72* expansion carriers [34]. The latter is an example of the complexity of these mechanisms, as some of these risk alleles are more significant in specific subgroups of patients. To see if a similar phenomenon exists also for *C9orf72* intermediate repeats, we analyzed Parkinson's disease patients of Ashkenazi origin, in a stratified manner, in carriers of mutations in *LRRK2*, *GBA*, and/or *SMPD1* genes, and in patients that do not carry these mutations. We showed that the expanded alleles (>60 repeats) had no association with PD, as shown by other groups and in a meta-analysis [16–22]; however, intermediate-size repeats of 20–60 are significantly enriched in PD-NC, increasing the risk for PD, while, in PD-carriers, there was no effect. The high significant odds ratio of 3.68 in non-carriers may be due to the exclusion of those patients who carry known risk alleles in *LRRK2*, *GBA*, or *SMPD1*, as we believe that in these patients the risk for PD is likely influenced by these mutations and not by the *C9orf72* intron 1 hexanucleotide repeat numbers. Of note is that Xi et al. also reported that intermediate repeats of 20–29 were found only in PD-NC, and not in PD-patients that carry the *LRRK2* p.G2019S mutation [19]. Based on these observations, we propose that *C9orf72* G<sub>4</sub>C<sub>2</sub> hexanucleotide repeats in intron 1 act as a risk factor for PD when number of repeats are intermediates. Although the exact definition of intermediate- and pathogenic-length size is still in debate, and may differ in different populations, we believe that the presence of alleles lower than the 60 repeats, in the intermediate range, should not be dismissed as a potential risk for PD.

How may intermediate numbers of repeats affect the risk to develop PD?

In ALS, although the role of *C9orf72* expansions is not yet fully established, three main mechanisms are proposed to contribute to its pathogenicity: *C9orf72* loss-of-function, generation of toxic RNA aggregates, and short peptide accumulation [12,14]. As G<sub>4</sub>C<sub>2</sub> repeats in the *C9orf72* gene are located within intron 1 and are near the promoter region, they may lead to changes in promoter regulation, depending on the number of repeats. Indeed, it was shown that ALS/FTD-expanded alleles are highly methylated and lead to lower levels of *C9orf72* mRNA and protein [35–37].

Do these same mechanisms contribute to Parkinson's disease risk? We demonstrated here that the risk-haplotype, which is shared by the intermediate *C9orf72* G<sub>4</sub>C<sub>2</sub> repeats, includes many SNVs that have the same effect of increasing RNA expression levels of *C9orf72*, as reported in GTE<sub>x</sub>. We, therefore, suggest that the mechanism involved in the effect of intermediate-size repeats on PD-increased-risk, could be the higher expression of *C9orf72*. This mechanism was recently suggested for a different neurodegenerative disease, Corticobasal Degeneration (CBD), a rare neurodegenerative disease that shares some similar clinical features with PD [38]. Researchers showed a significant enrichment of intermediate repeats in autopsy-proven CBD, as well as increased *C9orf72*-RNA expression levels in human brain tissues and in CRISPR/cas9 knock-in iPSC cells, but no association with pathologic RNA foci or dipeptides aggregates.

One important question is whether intermediate-repeats-sizes or the increase of *C9orf72* expression, which is associated with the risk-haplotype, may drive the risk for PD.

As these two mostly go together, this question should be answered in an experimental set-up that separates the two events. Cali et al. tried to answer this question by knocking-in 28 repeats into iPSCs that normally have either 2 or 6 repeats (suggestive of cells that do not carry the risk haplotype) and demonstrated that these knocked-in cells show higher expression of *C9orf72* [38]. As GTEx data suggest higher expression of *C9orf72* for all 19 variants within the risk-haplotype, it is tempting to suggest that PD-risk may be determined by the level of *C9orf72* expression, mediated by both the risk-haplotype and the number of repeats, a hypothesis that needs further evaluation. This hypothesis raises other questions: does the overall level of *C9orf72* expression depend on the effect of the total number of repeats in both alleles, and whether the genomic region of the *C9orf72*-risk-haplotype might expand to a larger interval than the minimal linkage disequilibrium of 107 Kb, as suggested by gnomAD database? In addition, the GTEx data show that the risk-haplotype effect on *C9orf72* RNA expression levels is not uniform in all tissues. The effect size is large and significant mostly in brain tissues, as well as in the small intestine, but much smaller in whole blood and lymphocytes.

What could be the effect of intermediate-size repeats on cellular expression? Cali et al. performed a comparative genes expression analysis between cells with intermediate repeats and cells with low number of repeats [38], demonstrating upregulation of genes that are enriched for vesicle trafficking and protein degradation pathways, including golgi vesicle transport, response to ER-stress, and autophagy, pathways which are involved in PD.

In summary, our stratified analysis suggests that intermediate-size hexanucleotide repeats in *C9orf72* are a risk factor for PD in individuals who do not carry common AJ founder mutations in *LRRK2*, *GBA*, or *SMPD1*. These results should be interpreted with caution as no correction for multiple comparisons was performed, and similar analyses should be performed on a larger cohort of PD patients. However, we propose a model in which the risk for PD may be driven not only by the number of repeats, but also by the genotypes of SNVs within the risk-haplotype, affecting *C9orf72* RNA expression levels. Further studies are needed to elucidate the possible role of *C9orf72* in PD pathogenesis.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/genes12081210/s1>, Table S1. Allele frequencies of G<sub>4</sub>C<sub>2</sub> hexanucleotide repeat lengths in *C9orf72* in Parkinson's disease patients of Ashkenazi Jewish origin and controls. Figure S1. Number of alleles observed for each *C9orf72*-G<sub>4</sub>C<sub>2</sub>-repeat-size with risk or non-risk haplotypes in the Ashkenazi Jews (PDs and controls).

**Author Contributions:** Conceptualization, H.K., O.G., A.O.-U.; methodology, H.K., O.G., A.O.-U.; validation, H.K., O.G.; formal analysis, H.K., O.G.; investigation, H.K., O.G.; resources, M.G.-W., A.B.-S., T.G., A.T., A.M., N.G.; data curation, H.K., O.G., A.O.-U.; writing—original draft preparation, H.K., O.G., A.O.-U.; writing—review and editing, O.G., M.G.-W., A.B.-S., T.G., A.T., A.M., N.G., A.O.-U.; visualization, H.K., O.G.; supervision, O.G., A.O.-U.; project administration, M.G.-W., A.O.-U.; funding acquisition, A.O.-U. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Kahn Foundation and Michael J Fox Foundation.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (The Institutional Helsinki Committee of the Tel Aviv Sourasky Medical Center, Tel Aviv, Israel; protocol code 05-069, date of approval 8 September 2005), and by the Israeli National Supreme Helsinki Committee for Genetic Studies (protocol code 920050035, date of approval 8 September 2005).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available in the Supplementary Materials Table S1.

**Acknowledgments:** This work was performed in partial fulfillment of the requirements for a Ph.D. degree by Hila Kobo at the Sackler Faculty of Medicine, Tel Aviv University, Israel. We thank Haike Reznik-Wolf for assistance with Asuragene genotyping.

**Conflicts of Interest:** Authors H.K., O.G., M.G.-W. and A.B.-S. declare no conflicts of interest. Author T.G. has served as an advisor for Cytora, Synnerva, Teva, Medison, and Allergan, received honoraria from AbbVie and Neuroderm, research support from Parkinson’s Foundation, University Tel-Aviv, Phonetica Ltd., and Israel innovation authority and travel support from AbbVie, Medison, Medtronic, and Allergan. Author AT has received honoraria from AbbVie Inc. and research grants from Michael J Fox Foundation. Author AM has received consulting fees from NeuroDerm and research grants from Michal J Fox Foundation, the Israeli Science Foundation and the US Department of Defense. Author NG serves as consultant to Intec Pharma, NeuroDerm, Denali, Abbvie, Sanofi-Genzyme, Biogen, Vibrant, BOL, LTI, Idorsia and Neuron23, Pharma2B, receives payment for lectures at Abbvie, Sanofi-Genzyme and Movement Disorder Society, received research support from the Michael J Fox Foundation, the National Parkinson Foundation, and the Israel Science Foundation, as well as from Teva NNE program, Biogen, The Aufzien Center, and the Sieratzki family Foundation at Tel-Aviv University, serves on the advisory board of LTI, NeuroDerm, Sionara, Sanofi-Genzyme, Biogen, Denali, Intec Pharma, Idorsia, and own stocks in Lysosomal Therapeutic Ltd., Vibrant, BOL, and serves as a member of the Editorial Board for the Journal of Parkinson’s Disease. Author AOU has received research support from Michael J Fox Foundation and Chaya Charitable Fund and has received honoraria from Sanofi Genzyme. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

### Abbreviations

AAE: age at enrollment; AAO, age-at-onset, ALS, amyotrophic lateral sclerosis; C9orf72, chromosome 9 open reading frame 72; CI, confidence interval; FTD, frontotemporal dementia; GBA, glucosidase  $\beta$  acid; LRRK2, leucine-rich repeat kinase 2; SMPD1, sphingomyelin phosphodiesterase 1; N, number of individuals; OR, odds ratio; PD, Parkinson’s disease; SD, standard deviation; y, years.

### References

1. De Lau, L.M.; Breteler, M.M. Epidemiology of Parkinson’s disease. *Lancet Neurol.* **2006**, *5*, 525–535. [[CrossRef](#)]
2. Blauwendraat, C.; Nalls, M.A.; Singleton, A.B. The genetic architecture of Parkinson’s disease. *Lancet Neurol.* **2020**, *19*, 170–178. [[CrossRef](#)]
3. Corti, O.; Lesage, S.; Brice, A. What genetics tells us about the causes and mechanisms of Parkinson’s disease. *Physiol. Rev.* **2011**, *91*, 1161–1218. [[CrossRef](#)]
4. Farrer, M.J. Genetics of Parkinson disease: Paradigm shifts and future prospects. *Nat. Rev. Genet.* **2006**, *7*, 306–318. [[CrossRef](#)] [[PubMed](#)]
5. Gasser, T.; Hardy, J.; Mizuno, Y. Milestones in PD genetics. *Mov. Disord.* **2011**, *26*, 1042–1048. [[CrossRef](#)]
6. Lesage, S.; Brice, A. Parkinson’s disease: From monogenic forms to genetic susceptibility factors. *Hum. Mol. Genet.* **2009**, *18*, R48–R59. [[CrossRef](#)]
7. Lubbe, S.J.; Escott-Price, V.; Gibbs, J.R.; Nalls, M.A.; Bras, J.; Price, T.R.; Nicolas, A.; Jansen, I.E.; Mok, K.Y.; Pittman, A.M.; et al. Additional rare variant analysis in Parkinson’s disease cases with and without known pathogenic mutations: Evidence for oligogenic inheritance. *Hum. Mol. Genet.* **2016**, *25*, 5483–5489. [[CrossRef](#)]
8. Van Heesbeen, H.J.; Smidt, M.P. Entanglement of Genetics and Epigenetics in Parkinson’s Disease. *Front. Neurosci.* **2019**, *13*, 277. [[CrossRef](#)] [[PubMed](#)]
9. DeJesus-Hernandez, M.; Mackenzie, I.R.; Boeve, B.F.; Boxer, A.L.; Baker, M.; Rutherford, N.J.; Nicholson, A.M.; Finch, N.A.; Flynn, H.; Adamson, J.; et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **2011**, *72*, 245–256. [[CrossRef](#)] [[PubMed](#)]
10. Renton, A.E.; Majounie, E.; Waite, A.; Simon-Sanchez, J.; Rollinson, S.; Gibbs, J.R.; Schymick, J.C.; Laaksovirta, H.; van Swieten, J.C.; Myllykangas, L.; et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **2011**, *72*, 257–268. [[CrossRef](#)] [[PubMed](#)]
11. Liu, Y.; Yu, J.T.; Zong, Y.; Zhou, J.; Tan, L. C9ORF72 mutations in neurodegenerative diseases. *Mol. Neurobiol.* **2014**, *49*, 386–398. [[CrossRef](#)] [[PubMed](#)]
12. Shu, L.; Sun, Q.; Zhang, Y.; Xu, Q.; Guo, J.; Yan, X.; Tang, B. The Association between C9orf72 Repeats and Risk of Alzheimer’s Disease and Amyotrophic Lateral Sclerosis: A Meta-Analysis. *Parkinsons Dis.* **2016**, *2016*, 5731734. [[CrossRef](#)] [[PubMed](#)]
13. Dols-Icardo, O.; Garcia-Redondo, A.; Rojas-Garcia, R.; Sanchez-Valle, R.; Noguera, A.; Gomez-Tortosa, E.; Pastor, P.; Hernandez, I.; Esteban-Perez, J.; Suarez-Calvet, M.; et al. Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. *Hum. Mol. Genet.* **2014**, *23*, 749–754. [[CrossRef](#)]
14. Haeusler, A.R.; Donnelly, C.J.; Rothstein, J.D. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nat. Rev. Neurosci.* **2016**, *17*, 383–395. [[CrossRef](#)] [[PubMed](#)]

15. Boeve, B.F.; Boylan, K.B.; Graff-Radford, N.R.; DeJesus-Hernandez, M.; Knopman, D.S.; Pedraza, O.; Vemuri, P.; Jones, D.; Lowe, V.; Murray, M.E.; et al. Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. *Brain J. Neurol.* **2012**, *135*, 765–783. [[CrossRef](#)]
16. Cooper-Knock, J.; Frolov, A.; Highley, J.R.; Charlesworth, G.; Kirby, J.; Milano, A.; Hartley, J.; Ince, P.G.; McDermott, C.J.; Lashley, T.; et al. C9ORF72 expansions, parkinsonism, and Parkinson disease: A clinicopathologic study. *Neurology* **2013**, *81*, 808–811. [[CrossRef](#)]
17. Nuytemans, K.; Bademci, G.; Kohli, M.M.; Beecham, G.W.; Wang, L.; Young, J.I.; Nahab, F.; Martin, E.R.; Gilbert, J.R.; Benatar, M.; et al. C9ORF72 intermediate repeat copies are a significant risk factor for Parkinson disease. *Ann. Hum. Genet.* **2013**, *77*, 351–363. [[CrossRef](#)]
18. Majounie, E.; Abramzon, Y.; Renton, A.E.; Keller, M.F.; Traynor, B.J.; Singleton, A.B. Large C9orf72 repeat expansions are not a common cause of Parkinson’s disease. *Neurobiol. Aging* **2012**, *33*, 2527.e1–2527.e2. [[CrossRef](#)]
19. Xi, Z.; Zinman, L.; Grinberg, Y.; Moreno, D.; Sato, C.; Bilbao, J.M.; Ghani, M.; Hernandez, I.; Ruiz, A.; Boada, M.; et al. Investigation of c9orf72 in 4 neurodegenerative disorders. *Arch. Neurol.* **2012**, *69*, 1583–1590. [[CrossRef](#)]
20. Lesage, S.; Le Ber, I.; Condroyer, C.; Broussolle, E.; Gabelle, A.; Thobois, S.; Pasquier, F.; Mondon, K.; Dion, P.A.; Rochefort, D.; et al. C9orf72 repeat expansions are a rare genetic cause of parkinsonism. *Brain J. Neurol.* **2013**, *136*, 385–391. [[CrossRef](#)]
21. Jiao, B.; Guo, J.F.; Wang, Y.Q.; Yan, X.X.; Zhou, L.; Liu, X.Y.; Zhang, F.F.; Zhou, Y.F.; Xia, K.; Tang, B.S.; et al. C9orf72 mutation is rare in Alzheimer’s disease, Parkinson’s disease, and essential tremor in China. *Front. Cell. Neurosci.* **2013**, *7*, 164. [[CrossRef](#)]
22. Theuns, J.; Verstraeten, A.; Sleegers, K.; Wauters, E.; Gijssels, I.; Smolders, S.; Corsiers, D.; Corsmit, E.; Elinck, E.; Sharma, M.; et al. Global investigation and meta-analysis of the C9orf72 (G4C2)<sub>n</sub> repeat in Parkinson disease. *Neurology* **2014**, *83*, 1906–1913. [[CrossRef](#)]
23. Bourinaris, T.; Houlden, H. C9orf72 and its Relevance in Parkinsonism and Movement Disorders: A Comprehensive Review of the Literature. *Mov. Disord. Clin. Pract.* **2018**, *5*, 575–585. [[CrossRef](#)]
24. Orr-Urtreger, A.; Shifrin, C.; Rozovski, U.; Rosner, S.; Bercovich, D.; Gurevich, T.; Yagev-More, H.; Bar-Shira, A.; Giladi, N. The LRRK2 G2019S mutation in Ashkenazi Jews with Parkinson disease: Is there a gender effect? *Neurology* **2007**, *69*, 1595–1602. [[CrossRef](#)]
25. Gan-Or, Z.; Giladi, N.; Rozovski, U.; Shifrin, C.; Rosner, S.; Gurevich, T.; Bar-Shira, A.; Orr-Urtreger, A. Genotype-phenotype correlations between GBA mutations and Parkinson disease risk and onset. *Neurology* **2008**, *70*, 2277–2283. [[CrossRef](#)] [[PubMed](#)]
26. Goldstein, O.; Gana-Weisz, M.; Nefussy, B.; Vainer, B.; Nayshool, O.; Bar-Shira, A.; Traynor, B.J.; Drory, V.E.; Orr-Urtreger, A. High frequency of C9orf72 hexanucleotide repeat expansion in amyotrophic lateral sclerosis patients from two founder populations sharing the same risk haplotype. *Neurobiol. Aging* **2017**. [[CrossRef](#)]
27. Goldstein, O.; Nayshool, O.; Nefussy, B.; Traynor, B.J.; Renton, A.E.; Gana-Weisz, M.; Drory, V.E.; Orr-Urtreger, A. OPTN 691\_692insAG is a founder mutation causing recessive ALS and increased risk in heterozygotes. *Neurology* **2016**, *86*, 446–453. [[CrossRef](#)] [[PubMed](#)]
28. Goldstein, O.; Gana-Weisz, M.; Cohen-Avinoam, D.; Shiner, T.; Thaler, A.; Cedarbaum, J.M.; John, S.; Lalioti, M.; Gurevich, T.; Bar-Shira, A.; et al. Revisiting the non-Gaucher-GBA-E326K carrier state: Is it sufficient to increase Parkinson’s disease risk? *Mol. Genet. Metab.* **2019**, *128*, 470–475. [[CrossRef](#)] [[PubMed](#)]
29. Vacic, V.; Ozelius, L.J.; Clark, L.N.; Bar-Shira, A.; Gana-Weisz, M.; Gurevich, T.; Gusev, A.; Kedmi, M.; Kenny, E.E.; Liu, X.; et al. Genome-wide mapping of IBD segments in an Ashkenazi PD cohort identifies associated haplotypes. *Hum. Mol. Genet.* **2014**, *23*, 4693–4702. [[CrossRef](#)] [[PubMed](#)]
30. Consortium, G.T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **2020**, *369*, 1318–1330. [[CrossRef](#)]
31. Paulson, H. Repeat expansion diseases. *Handb. Clin. Neurol.* **2018**, *147*, 105–123. [[CrossRef](#)] [[PubMed](#)]
32. Elden, A.C.; Kim, H.J.; Hart, M.P.; Chen-Plotkin, A.S.; Johnson, B.S.; Fang, X.; Armarkola, M.; Geser, F.; Greene, R.; Lu, M.M.; et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **2010**, *466*, 1069–1075. [[CrossRef](#)]
33. Lee, T.; Li, Y.R.; Ingre, C.; Weber, M.; Grehl, T.; Gredal, O.; de Carvalho, M.; Meyer, T.; Tysnes, O.B.; Auburger, G.; et al. Ataxin-2 intermediate-length polyglutamine expansions in European ALS patients. *Hum. Mol. Genet.* **2011**, *20*, 1697–1700. [[CrossRef](#)] [[PubMed](#)]
34. Lattante, S.; Pomponi, M.G.; Conte, A.; Marangi, G.; Bisogni, G.; Patanella, A.K.; Meleo, E.; Lunetta, C.; Riva, N.; Mosca, L.; et al. ATXN1 intermediate-length polyglutamine expansions are associated with amyotrophic lateral sclerosis. *Neurobiol. Aging* **2017**. [[CrossRef](#)]
35. Xi, Z.; Zhang, M.; Bruni, A.C.; Maletta, R.G.; Colao, R.; Fratta, P.; Polke, J.M.; Sweeney, M.G.; Mudanohwo, E.; Nacmias, B.; et al. The C9orf72 repeat expansion itself is methylated in ALS and FTL D patients. *Acta Neuropathol.* **2015**, *129*, 715–727. [[CrossRef](#)] [[PubMed](#)]
36. Belzil, V.V.; Bauer, P.O.; Prudencio, M.; Gendron, T.F.; Stetler, C.T.; Yan, I.K.; Pregent, L.; Daugherty, L.; Baker, M.C.; Rademakers, R.; et al. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* **2013**, *126*, 895–905. [[CrossRef](#)]

- 
37. Waite, A.J.; Baumer, D.; East, S.; Neal, J.; Morris, H.R.; Ansorge, O.; Blake, D.J. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol. Aging* **2014**, *35*, 1779.e5–1779.e13. [[CrossRef](#)] [[PubMed](#)]
  38. Cali, C.P.; Patino, M.; Tai, Y.K.; Ho, W.Y.; McLean, C.A.; Morris, C.M.; Seeley, W.W.; Miller, B.L.; Gaig, C.; Vonsattel, J.P.G.; et al. C9orf72 intermediate repeats are associated with corticobasal degeneration, increased C9orf72 expression and disruption of autophagy. *Acta Neuropathol.* **2019**, *138*, 795–811. [[CrossRef](#)] [[PubMed](#)]