Differential Impacts of TOX3 Genetic Variants on Restless Legs Syndrome and Parkinson's Disease

Nora Patel,a,b, Jiajun Li,c,d, Emma R. Davis,b,e, Yifan Liu,f, Daniel P. Clark,b,e,g, Olivia H. Wright,h, Matthew J. Taylor,h, Chloe S. White,b,e, Feng Zhang,c,d, Isabella Martin,i, Liam Anderson,j, Ava R. Lewis,b,e,g, Ethan J. Carter,b,e,g,*

[a] Department of Experimental Medicine, McGill University, Montréal, Québec, Canada
[b] Montreal Neurological Institute, McGill University, Montréal, Québec, Canada
[c] Laboratoire de neurogénétique, Centre de recherche, Institut universitaire en santé mentale de Montréal, Montréal, Québec, Canada
[d] Département de psychiatrie, Université de Montréal, Montréal, QC, Canada
[e] Department of Neurology and Neurosurgery, McGill University, Montréal, Québec, Canada
[f] Department of Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada
[g] Department of Human Genetics, McGill University, Montréal, Québec, Canada
[h] Department of Neurology, The Johns Hopkins Bayview Medical Center, Baltimore, MD, USA
[i] Division of Neurosciences, CHU de Québec - Université Laval and the Department of Medicine, Faculty of Medicine, Laval University.
[j] National Reference Center for Narcolepsy, Sleep Unit, Department of Neurology, Gui-de-Chauliac Hospital, CHU Montpellier, University of Montpellier, Inserm U1061, Montpellier, France

Abstract

Introduction: Parkinson disease (PD) and restless legs syndrome (RLS) may be clinically and/or etiologically related, yet this association is under debate. Single nucleotide polymorphisms (SNPs) in the *TOX3* gene locus were implicated in both RLS and PD genome-wide association studies (GWAS), suggesting a potential pleiotropy.

Methods: Two case-control cohorts including 644 PD patients, 457 RLS patients and 945 controls were genotyped for one known RLS-related SNP (rs3104767) and one PD-related SNP (rs4784226) in the *TOX*3 locus. The associations between genotype and PD and RLS risk were tested using multivariate regression models.

Results: The allele frequencies of RLS-related SNP rs3104767 in RLS patients and controls were 0.35 and 0.43, respectively (OR 0.70, *p*=0.0007). Regression model suggested that this association

is derived by homozygous carriage of rs3104767 (adjusted p=0.008). A nominal association was observed for homozygous carriers of the rs3104767 SNP in PD (OR 1.62, 95% CI 1.05-2.54, p=0.034), i.e. with an opposite direction of effect on RLS and PD, but this was not significant after Bonferroni correction. However, data from published GWASs of RLS and PD, and from the PDgene database, further supported these inverse associations.

Conclusions: Our results confirm the association between the *TOX3* SNP rs3104767 and RLS, and suggest that *TOX3* variants are involved in both RLS and PD, but with different, or even opposite effects. Studies in larger populations of different ethnicities are required to further refine the *TOX3* locus is involved in RLS and PD.

Keywords: Restless legs syndrome; Parkinson's disease; Genetics; TOX3

Introduction

Parkinson disease (PD) and restless legs syndrome (RLS) are common neurological disorders affecting millions around the world. RLS prevalence varies in different populations from 5%-15% (Yeh et al. 2012) while PD is prevalent in 1-2% of individuals older than 65 (de Rijk et al. 2000). Due to their high prevalence, co-morbidity with both RLS and PD may be common. Furthermore, it was suggested that PD treatment may lead to RLS or RLS-like symptoms (Lee et al. 2009). Some epidemiological studies suggested that PD is more common in RLS patients and RLS is more common in PD patients, while other studies reported lack of association (Rijsman et al. 2014). Therefore, it was hypothesized that RLS and PD could be etiologically related. However, this association is controversial, as other epidemiological, pathological and imaging studies did not support this hypothesis (Garcia-Borreguero et al. 2003). Therefore, it is still unclear if these two disorders are etiologically related.

Thus far, genome-wide studies (GWAS) identified 19 genetic loci for RLS (Schormair et al. 2017) and 41 loci for PD (Chang et al. 2017). In the RLS GWAS, pleiotropy with PD was examined, and it was concluded that there is no pleiotropy between RLS and PD, probably ruling out shared etiology. Furthermore, previous studies that examined the role of some of the RLS genetic markers in PD identified no association (Gan-Or et al. 2015). However, different single nucleotide polymorphisms (SNPs) in the *TOX3* locus were implicated in RLS (rs3104767), or just below the significance threshold in a PD GWAS database (rs4784226, $p=3.37 \times 10^{-6}$, www.pdgene.org). Interestingly, a new large-scale PD GWAS implicated another SNP (rs4784227, almost in full linkage disequilibrium, LD, with rs4784226) in the *TOX3* locus, which was associated with a slightly increased risk for PD, and significant at the genome-wide level (OR 1.09, $p=9.75 \times 10^{-11}$) (Chang et al. 2017). To further examine the potential role of these SNPs in

RLS and PD risk, we performed an association study of rs3104767 and rs4784226 in our cohorts of RLS and PD.

Methods

Study Population

Case-control population of Parkinson disease

The PD population included unrelated, consecutively recruited patients (n=644) and controls (n=613), of European ancestry, of French or French-Canadian origin. French-Canadian patients and controls were recruited from clinics across Québec, Canada, including the Quebec Parkinson's Network (http://rpq-qpn.ca/). French patients and controls were recruited at Montpellier, France. PD was diagnosed using the UK Brain Bank Criteria. The average age at enrollment of PD patients was 65.6±9.8 years (data were missing for 22 patients), with 64.4% men (data were missing for 3 patients). The average age of the control population was 52.5±13.1 years (data were available for 538 individuals), with 58.2% men. As there are differences between age and sex ($p<0.05$ for both), adjustment for age and sex was performed (see statistical analysis).

Case-control population of restless legs syndrome

The RLS study included unrelated and consecutively recruited RLS patients (n=424) and controls (n=325) of European ancestry, recruited as previously described (Gan-Or et al. 2015). The average age at enrollment of RLS patients was 56.4±14.1 years (data were missing for two patients), with 39.2% men (data were missing for 1 patient). Patients were diagnosed based on the international RLS study group (IRLSSG) criteria (Allen et al. 2014). The average age at enrollment of the control population was 52.8±16.8 years (data were missing for 21 individuals), with 38.4% men

(data were missing for 15 individuals). All participants in both PD and RLS cohorts signed informed consent at enrollment, and the study protocols w ere approved by the institutional ethics review boards.

Selection of SNPs and genotyping

Two single nucleotide polymorphism (SNPs) in the *TOX3* locus on chromosome 16q12.1 were selected for genotyping in the two cohorts. One SNP which was associated with risk for RLS, rs3104767 (Winkelmann et al. 2011), and a second SNP, rs4784226, associated with risk for PD (www.pdgene.org) (Chang et al. 2017). Of note, the rs4784227 SNP from the recent PD GWAS (Chang et al. 2017) was not selected for analysis as this work was done prior to the publication of the GWAS and was based on the data in PDgene. Importantly, these two SNPs, rs4784226 and rs4784227, are in high LD (r^2=0.9, D'=0.95), thus genotyping only one as a proxy is sufficient. DNA was extracted from blood using a standard salting out protocol. The two SNPs were genotyped by TaqMan SNP genotyping assays following the manufacturer's instructions. The genotypes were called using the QuantStudio^{TM} 7 Flex Real-Time PCR System and Software (v 1.0) (Applied Biosystem).

Statistical analysis

Goodness-of-fit test was applied to look for deviation from Hardy-Weinberg equilibrium (HWE). $\chi2$ and Student's *t*-tests were used to determine differences in sex and age, respectively. To examine the association of *TOX3* locus variants, binary logistic regression was applied with status (patient/control) as a dependent variable, SNP genotypes as independent variables and age and sex as covariates. Recessive and dominant regression models were also performed. All statistical analysis was done using Statistical Package for Social Science (SPSS) software V.23 (IBM Inc)

and PLINK. Since two SNPs were tested in three models, Bonferroni correction for multiple comparisons was applied ($p<0.0083$ was considered statistically significant).

Results

The distribution of the two tested SNPs in the control populations did not deviate from HWE. The two SNPs were successfully genotyped in all samples and were in partial linkage disequilibrium (e.g. based on PD controls, D'=0.663, r^2=0.21, p for LD <0.001, nearly identical in the RLS controls). Table 1 details the distribution of genotypes of both SNPs in PD and RLS patients and controls. The allele frequencies of RLS-related SNP rs3104767 in RLS patients and controls were 0.35 and 0.43, respectively (OR=0.70, 95%CI=0.56-0.86, p=0.0007), in concordance with the previously reported GWAS (Winkelmann et al. 2011). The allele frequencies of this SNP in PD patients and controls, however, were 0.46 and 0.44, with opposite, non-significant directionality (OR=1.12, 95%CI=0.95-1.31, p=0.17). The allele frequencies of the PD-associated SNP rs4784226 were neither associated with RLS (OR=0.79, 95%CI=0.62-1.01, p=0.07, Bonferroni corrected p=0.42) nor with PD (OR=1.01, 95%CI=0.85-1.20, p=0.92).

We further performed logistic regression adjusted for age and sex (Table 1). Interestingly, a nominal association, with an inverse effect to that of RLS was observed for homozygous carriers of the rs3104767 SNP in PD (OR=1.62, 95%CI=1.05-2.54, p=0.034, Bonferroni corrected p=0.204). Furthermore, the regression model of the rs3104767 SNP in RLS suggested that the observed association with RLS is derived by homozygous carriage of the rs3104767 SNP (p=0.008). In age- and sex-adjusted recessive model, as expected, biallelic carriage of the rs3104767 SNP was nominally associated with a reduced risk for RLS (OR=0.59, 95%CI=0.38-

0.90, *p*=0.01), while in the dominant model this effect was reduced (OR=0.72, 95%CI=0.52-0.996, *p*=0.047).

Discussion

Our results demonstrate that rs3104767 is associated with risk for RLS, as was previously reported, with similar allele frequencies and effect size (Winkelmann et al. 2011), probably driven by homozygous carriage of the SNP. Interestingly, homozygous carriage of this SNP was also nominally associated with increased risk for PD (OR 1.62). When considering our results together with the RLS and PD GWASs (Chang et al. 2017, Schormair et al. 2017) and the data in PDgene (www.pdgene.org), it seems that there is an opposite effect of *TOX3* haplotypes in PD and RLS. While in RLS the minor alleles are associated with decreased risk, in PD they are associated with increased risk (Table 2). In our cohorts, homozygous carriage of rs3104767 was associated with an almost two-fold reduced risk for RLS, and 60% increase in risk for PD (yet with only nominally significant *p* value, probably due to reduced power). In the PDgene database, all three SNPs, rs3104767, rs4784226 and rs4784227 were nominally associated with an increased risk for PD ($p<5.04x10^{-5}$ for each of the SNPs), and in the recent PD GWAS, they all reached genome-wide significance (Chang et al. 2017). Therefore, *TOX3* may play a role in both RLS and PD, but possibly with an opposite effect.

Our study has some limitations. We do not have data on the presence of RLS in our PD cohort or on the presence of PD in the RLS cohort, which could have improved our analysis. However, given the concordance of our results with the much larger GWASs, this limitation probably did not affect the results. In addition, as we discussed above, our cohorts, despite being relatively large for a single center, were underpowered to detect the association with PD after

correction for multiple comparisons. Of note, since the association with PD could be considered as a replication of the GWAS results, nominally significant *p* value may also be considered as a true association. Lastly, we could not perform a direct comparison between PD and RLS patients, due to the ethnic background of the studied populations. Although both PD and RLS cohorts included only patients and controls of European ancestry, our PD cohort was enriched in French genetic background, while our RLS cohort was of mixed European ancestry (French-Canadian and US patients of different European origins). Therefore, future studies directly comparing PD and RLS of the same ethnicity are warranted.

In the *TOX3* locus, except for the *TOX3* gene, there is a non-coding RNA gene (*CASC16*), and it cannot be ruled out based on the current data that it is involved in PD. *TOX3* encodes TOX high mobility group (HMG) box family member 3, which acts as a nuclear transcription regulation factor. Based on data from the Genotype-Tissue Expression portal (GTEx, www.gtexportal.org), TOX3 is expressed in parts of the brain, including in the substantia nigra and the cortex, but less in other basal ganglia, and there is no data on the association between the tested *TOX3* SNPs and TOX3 expression. TOX3 may serve as a neuronal survival factor, through transcription induction which is probably calcium mediated (Yuan et al. 2009), yet its potential functional roles in RLS and PD are still to be uncovered.

TOX3 is also involved in breast cancer, and it was suggested that it may serve as a regulator of estrogen receptor-mediated gene expression (Seksenyan et al. 2015). This may raise a hypothesis for its differential involvement in RLS and PD. In both RLS and PD, sex is associated with risk for the disorder, but with an inverse association. In PD men are more susceptible (de Rijk et al. 2000) and in RLS women are more susceptible (Yeh et al. 2012). If this differential sex distribution in PD and RLS is related to estrogen regulation, perhaps the inverse associations of

TOX3 variants with RLS and PD are involved in the sex differences in these two disorders, through different effects on estrogen receptor-mediated gene expression. Currently, there are no supporting evidence for this hypothesis, and further study is needed to examine this hypothesis and the differences in *TOX3*-associated risk in RLS and PD.

Acknowledgements

References

Allen, R. P., D. L. Picchietti, D. Garcia-Borreguero et al. (2014). Restless legs syndrome/Willis-Ekbom disease diagnostic criteria: updated International Restless Legs Syndrome Study Group (IRLSSG) consensus criteria--history, rationale, description, and significance. Sleep Med 15(8): 860-873.

Chang, D., M. A. Nalls, I. B. Hallgrimsdottir et al. (2017). A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nat Genet 49(10):1511-1516.

de Rijk, M. C., L. J. Launer, K. Berger et al. (2000). Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Neurology 54(11 Suppl 5): S21-23.

Gan-Or, Z., R. N. Alcalay, A. Bar-Shira et al. (2015). Genetic markers of Restless Legs Syndrome in Parkinson disease. Parkinsonism Relat Disord 21(6): 582-585.

Gan-Or, Z., S. Zhou, A. Ambalavanan et al. (2015). Analysis of functional GLO1 variants in the BTBD9 locus and restless legs syndrome. Sleep Med 16(9): 1151-1155.

Garcia-Borreguero, D., P. Odin and C. Serrano (2003). Restless legs syndrome and PD: a review of the evidence for a possible association. Neurology 61(6 Suppl 3): S49-55.

Lee, J. E., H. W. Shin, K. S. Kim and Y. H. Sohn (2009). Factors contributing to the development of restless legs syndrome in patients with Parkinson disease. Mov Disord 24(4): 579-582.

Rijsman, R. M., L. F. Schoolderman, R. S. Rundervoort and M. Louter (2014). Restless legs syndrome in Parkinson's disease. Parkinsonism Relat Disord 20 Suppl 1: S5-9.

Schormair, B., C. Zhao, S. Bell et al. (2017). Identification of novel risk loci for restless legs syndrome in genome-wide association studies in individuals of European ancestry: a meta-analysis. Lancet Neurol 16(11): 898-907.

Seksenyan, A., A. Kadavallore, A. E. Walts et al. (2015). TOX3 is expressed in mammary ER(+) epithelial cells and regulates ER target genes in luminal breast cancer. BMC Cancer 15: 22.

Winkelmann, J., D. Czamara, B. Schormair et al. (2011). Genome-wide association study identifies novel restless legs syndrome susceptibility loci on 2p14 and 16q12.1. PLoS Genet 7(7): e1002171.

Yeh, P., A. S. Walters and J. W. Tsuang (2012). Restless legs syndrome: a comprehensive overview on its epidemiology, risk factors, and treatment. Sleep Breath 16(4): 987-1007.

Yuan, S. H., Z. Qiu and A. Ghosh (2009). TOX3 regulates calcium-dependent transcription in neurons. Proc Natl Acad Sci U S A 106(8): 2909-2914.

Table 1. Logistic regression model adjusted for age and sex of the association between *TOX3* SNPs, rs3104767 and rs4784226, RLS and PD.

Population	SNP	Genotype	Genotype frequency RLS, % (n)	Genotype frequency Control, % (n)	OR (95% CI) [a]	*p* value [a]
RLS	rs3104767	GG	42.2% (179)	32.3% (105)	Ref	Ref
		GT	46.0% (195)	48.6% (158)	0.79 (0.56-1.11)	0.17
		TT	**11.8% (50)**	**19.1% (62)**	**0.52 (0.33-0.85)**	**0.008**
	rs4784226	CC	63.4% (269)	56.3% (183)	Ref	Ref
		CT	32.3% (137)	38.5% (125)	0.83 (0.60-1.16)	0.27
		TT	4.2% (18)	5.2% (17)	0.90 (0.43-1.86)	0.77
Population	SNP	Genotype	Genotype frequency PD, % (n)	Genotype frequency Control, % (n)	OR (95% CI) [a]	P value [a]
PD	rs3104767	GG	28.0% (180)	32.3% (198)	Ref	Ref
		GT	51.4% (331)	48.1% (295)	1.19 (0.85-1.66)	0.32
		TT	**20.7% (133)**	**19.6%((120)**	**1.62 (1.05-2.54)**	**0.034 [b]**
	rs4784226	CC	53.4% (344)	54.8% (336)	Ref	Ref
		CT	39.3% (253)	36.9% (226)	1.04 (0.77-1.42)	0.79
		TT	7.3% (47)	8.3% (51)	0.85 (0.47-1.54)	0.60

RLS, restless legs syndrome; PD, Parkinson's disease; SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval; n, number.
[a] Adjusted for age and sex
[b] Only trend after Bonferroni correction for multiple comparisons, *p*=0.068.

Table 2. Inverse associations of *TOX3* SNPs with RLS and PD in the current and previous studies

	rs3104767				rs4784226				rs4784227			
	Current study	RLS GWAS	PD GWAS	PDgene	Current study	RLS GWAS	PD GWAS	PDgene	Current study	RLS GWAS	PD GWAS	PDgene
Association with RLS	↓	↓	NT	NT	NS	↓	NT	NT	NT	↓	NT	NT
Association with PD	↑ᵃ	NT	↑	↑	NS	NT	↑	↑	NT	NT	↑	↑

The arrows point the direction of association of the minor allele. Arrow pointing down indicates reduced risk associated with the minor allele, arrow pointing up indicates increases risk associated with the minor allele. Large arrows indicate that the association is based on published data, and the small arrows indicate inferred effect based on linkage disequilibrium with the other SNPs, since all three SNPs are in LD (rs4784226 and rs4784227 in almost full LD, and rs3104767 is in partial LD with the two other SNPs).

RLS, restless legs syndrome; PD, Parkinson's disease; GWAS, genome-wide association study; PDgene, the PDgene database at www.pdgene.org; NT, not tested; NS, not significant.

[a] Nominal association