# Lysosomal storage disorder gene variants in multiple system atrophy

Lasse Pihlstrøm[1,2,3], Lucia Schottlaender[1,2], Viorica Chelban[1,2], MSA Exome Consortium, Wassilios G. Meissner[4], Monica Federoff[5], Andy Singleton[5] and Henry Houlden[1,2*]

1. Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK

2. National Hospital for Neurology and Neurosurgery, Queen Square, London, UK

3. Oslo University Hospital, Oslo, Norway.

4. French Reference Center for MSA, Department of Neurology, University Hospital Bordeaux, 33000 Bordeaux and Institute of Neurodegenerative Diseases, CNRS UMR 5293, University Bordeaux, 33000 Bordeaux, France.

5. Laboratory of Neurogenetics, NIH/NIA, Bethesda, USA.

* h.houlden@ucl.ac.uk

Sir,

We have read with great interest the recent article by Robak and colleagues, reporting an excessive burden of rare, likely damaging variants in lysosomal storage disorder genes associated with Parkinson's disease (Robak *et al.*, 2017). The authors analysed whole exome sequencing data from 1156 Parkinson's disease patients and 1679 control subjects in the discovery phase, investigating cumulative rare variant burden across 54 lysosomal storage disorder genes using the sequence kernel association test - optimal (SKAT-O) (Lee *et al.*, 2012). A significant association was found for non-synonymous variants and likely damaging variants with a minor allele frequency < 3%, which was also replicated in a sample set of 6713 cases and 5964 controls genotyped on the NeuroX array. Furthermore, SKAT-O was nominally significant for non-synonymous variants with minor allele frequency < 1% in the exome data from 436 patients and 169 controls in the Parkinson's Progression Marker Initiative (PPMI), albeit not significant following adjustment for multiple testing. Thus, these data further suggest a link between lysosomal dysfunction and α-synuclein aggregation (see figure 1)

Multiple system atrophy (MSA) is known to demonstrate several major features overlapping with Parkinson's disease, including clinical Parkinsonism and histopathological manifestations of α-

synucleinopathy (Gilman *et al.*, 2008). However, in contrast to Parkinson's disease and other neurodegenerative disorders, efforts to identify genetic causes of MSA have thus far had limited success. A Japanese study has proposed mutations in the *COQ2* gene as pathogenic in rare cases of familial MSA (Multiple-System Atrophy Research Collaboration, 2013), yet these findings have not been supported by subsequent independent investigations (Scholz and Bras, 2015). An association between MSA risk and common genetic variation of the α-synuclein locus has been reported in several Caucasian cohorts (Al-Chalabi *et al.*, 2009; Scholz *et al.*, 2009), yet was not replicated in the only genome-wide association study published to date, including 908 MSA patients (Sailer *et al.*, 2016). Furthermore, this study found no signals at genome-wide significance.

Pathogenic *GBA* mutations cause Gaucher's disease in the homozygous state and are strong risk factors for Parkinson's disease in heterozygous carriers (Sidransky *et al.*, 2009). The powerful example of *GBA* provided the main rationale for the broader study of lysosomal storage disorder gene variants by Robak and colleagues. A few reports indicate that rare coding variants in *GBA* may be a shared risk factor for Parkinson's disease and MSA (Mitsui *et al.*, 2015; Sklerov *et al.*, 2017). Neuropathological evidence also support a possible role of lysosomal dysfunction in MSA pathogenesis (Makioka *et al.*, 2012). Available sample sets for genetic studies of MSA are generally far smaller than in Parkinson's disease, yet the nominally significant replication reported in the moderately sized PPMI dataset indicated that a similar analysis might be attempted in MSA. We therefore investigated the association of rare variants in lysosomal storage disorder genes in whole exome sequencing data from MSA and healthy control subjects, adopting the same methodological approach as Robak et al.

Whole exome sequencing was performed in 264 pathologically confirmed MSA cases from the MSA Brain Bank and DNA Collaboration and 462 neuropathologically normal post-mortem controls with age at death >65 years, all of Caucasian origin. We further included exome sequencing data from 111 clinically diagnosed MSA cases and 116 cardiovascular controls, generating a final combined dataset of 375 MSA patients and 587 controls. Brain tissue obtained from Queen Square Brain Bank was donated for research using ethically approved protocols and stored under a license from the Human Tissue Authority. DNA was extracted and investigated under approval of the joint ethics committee of UCL Institute of Neurology and the National Hospital for Neurology and Neurosurgery, London, UK (UCLH: 04/N034). All living participants

were enrolled based on written, informed consent in accordance with research protocols approved by relevant institutional review boards.

Exome data was filtered using PLINK 1.9 to remove samples with high missingness rate, excess heterozygosity, sex check failure, non-Caucasian ancestry or evidence of cryptic relatedness, and filter out variants showing high missingness rate or deviation from Hardy-Weinberg equilibrium. Similar to Robak et al., we categorized variants into three nested groups, including (i) non-synonymous variants, (ii) likely damaging variants with Combined Annotation Dependent Depletion score (CADD) ≥ 12.37 (Kircher *et al.*, 2014) and (iii) loss of function variants (stopgain, stoploss, frameshift or splicing mutations). SKAT-O was performed for these three variant categories at minor allele frequency thresholds of $< 3\%$ and $< 1\%$, aggregated across the list of 54 lysosomal disorder genes using the SKAT R package (Lee *et al.*, 2012). Covariates included sex and top two principal components, which were calculated based on a subset of common, linkage disequilibrium-pruned single-nucleotide polymorphisms (SNPs). Conservatively including only SNPs in the analysis, which are less prone to sequencing error than indels, we observed no signs of inflated p-values in QQ-plots of single gene SKAT-O results.

Results are shown in Table 1. We found no significant evidence of excess burden of lysosomal storage disorder gene variants in MSA. P-values < 0.1 for both the largest variant set and the most damaging variant set (loss of function) might indicate that significant results could conceivably be obtained with larger sample sizes, yet this remains purely speculative. Furthermore, none of the included lysosomal storage disorder genes were individually significant in SKAT-O burden analyses.

In conclusion, despite an intriguing hypothesis and suggestively significant results in the similarly sized PPMI analysis by Robak and colleagues in PD, our investigation of rare variant burden across 54 lysosomal storage disorder genes in 375 MSA patients and 587 controls failed to show an association. Given the difficulty with genetic power in rare diseases, further international collaborations aiming to increase sample sets of several thousand patients will be crucial in future efforts to elucidate the genetic contribution to MSA. Larger genome-wide association studies are clearly needed to investigate common susceptibility variants. In our opinion, the approach taken by Robak and colleagues also exemplifies well how hypothesis-driven exploration of whole exome sequencing data can be used to identify rare variants and larger gene pathways contributing to neurodegenerative disease.

**Acknowledgements:** The authors are grateful to all study participants and to the Queen Square Brain Bank, UCL and the London Neurodegenerative diseases Brain Bank, Institute of Psychiatry, King's College London, the MRC Sudden Death Brain and Tissue Bank Edinburgh (UK), the IDIBAPS Brain Bank Barcelona (Spain), the Human Brain and Spinal Fluid Resource Center Los Angeles (USA), and the Harvard Brain Tissue Resource Center Belmont MA, the GIE-Neuro CEB brain bank Paris (France), and the Netherlands brain bank for the samples. We are also grateful to the MSA Trust, The MSA coalition, The Wellcome Trust and the MRC for funding.

**Table 1 Number of variants and SKAT-O association results**

| Variant category | Minor allele frequency<0.01 | | Minor allele frequency < 0.03 | |
|---|---|---|---|---|
| | SNPs | P-value | SNPs | P-value |
| Non-synonoymous | 432 | 0.14 | 446 | 0.084 |
| CADD > 12.36 | 294 | 0.17 | 300 | 0.24 |
| Loss of function | 10 | 0.088 | 10 | 0.088 |

Sex and top two principal components were used as covariates in SKAT-O analyses.
CADD ≥ 12.37 corresponds to the predicted 2% most damaging of all possible single-nucleotide changes in the genome. SKAT-O = sequence kernel association test - optimal; SNP = single-nucleotide polymorphism; CADD = Combined Annotation Dependent Depletion algorithm

**References:**

Al-Chalabi A, Durr A, Wood NW, Parkinson MH, Camuzat A, Hulot JS, *et al.* Genetic variants of the alpha-synuclein gene SNCA are associated with multiple system atrophy. PLoS One 2009; 4: e7114.

Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, *et al.* Second consensus statement on the diagnosis of multiple system atrophy. Neurology 2008; 71: 670-6.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014; 46: 310-5.

Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics 2012; 13: 762-75.

Makioka K, Yamazaki T, Takatama M, Nakazato Y, Okamoto K. Activation and alteration of lysosomes in multiple system atrophy. Neuroreport 2012; 23: 270-6.

Mitsui J, Matsukawa T, Sasaki H, Yabe I, Matsushima M, Durr A, *et al.* Variants associated with Gaucher disease in multiple system atrophy. Ann Clin Transl Neurol 2015; 2: 417-26.

Multiple-System Atrophy Research Collaboration. Mutations in COQ2 in familial and sporadic multiple-system atrophy. N Engl J Med 2013; 369: 233-44.

Robak LA, Jansen IE, van Rooij J, Uitterlinden AG, Kraaij R, Jankovic J, *et al.* Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. Brain 2017; 140: 3191-203.

Sailer A, Scholz SW, Nalls MA, Schulte C, Federoff M, Price TR, *et al.* A genome-wide association study in multiple system atrophy. Neurology 2016; 87: 1591-8.

Scholz SW, Bras J. Genetics Underlying Atypical Parkinsonism and Related Neurodegenerative Disorders. Int J Mol Sci 2015; 16: 24629-55.

Scholz SW, Houlden H, Schulte C, Sharma M, Li A, Berg D, *et al.* SNCA variants are associated with increased risk for multiple system atrophy. Ann Neurol 2009; 65: 610-4.

Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, *et al.* Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med 2009; 361: 1651-61.

Sklerov M, Kang UJ, Liong C, Clark L, Marder K, Pauciulo M, *et al.* Frequency of GBA variants in autopsy-proven multiple system atrophy. Mov Disord Clin Pract 2017; 4: 574-81.