Supplementary data for “Prion-like α-synuclein pathology in the brain of infants with Krabbe disease” by Hatton et al

Supplementary Methods

Supplementary Methods 1: Post-mortem tissue preparation and fractionation

KD cases were included if they had clinical features consistent with infantile KD, alongside enzymatic or genetic evidence of β-galactocerebrosidase deficiency or *GALC* mutation, and typical neuropathological features of KD (globoid cells and demyelination). KD4 did not have genetic or enzymatic evidence of KD; however, they were included as they had typical clinical and pathological features of KD and a sibling exhibited reduced galactosylceramidase activity. Demographic information for all cases is included in Supplementary Table 1.

At autopsy of NBTR tissue, the brainstem was removed at the level of the red nucleus and the cerebrum hemisected through the corpus callosum. The right hemisphere was immersion-fixed in neutral-buffered formalin whilst the left hemisphere was sliced into approximately 1 cm coronal slices and snap frozen at -120°C prior to storage at -80°C. After five weeks fixation, the right hemisphere was sub-dissected for neuropathological evaluation and processed for paraffin wax embedding.

Approximately 400 mg of frozen grey matter from temporal cortex (BA20), in addition to 400 mg of underlying white matter, was dissected from tissue slices at -20°C. Small samples were retained for real-time quaking-induced conversion (RT-QuIC) analysis, whilst the remainder was suspended in 0.2 M TEAB buffer supplemented with Complete protease inhibitors and PhoStop tablets (Roche, Basel, Switzerland) and homogenised for 30 seconds in a rotator-stator homogeniser.

Approximately 200 µl was removed from each sample as the crude tissue lysate. Samples were centrifuged at 16,000 rcf for 30 minutes at 4°C with the supernatant recovered as the soluble tissue fraction. Samples were then washed by resuspension in 1:10 w/v Tris-HCl buffer (50 mM Tris-HCl, 150 mM sodium chloride, 2 mM EDTA) and centrifuged twice at 16,000 rcf for 30 minutes at 4°C. The pellet was resuspended 1:5 w/v in Tris-HCl buffer containing 5% SDS,

vortexed, and incubated at room temperature for 5 minutes before centrifugation at 16,000 rcf for 30 minutes at 10°C, with the supernatant recovered as the SDS-soluble fraction.

Temporal cortex (Brodmann area 20) was chosen for molecular analysis of α-synuclein as it contained sufficient tissue for downstream analyses and consistently contained globoid cells and neuritic and neuronal α-synuclein pathology in KD cases. Frozen tissue was only available for two KD cases (KD 1 and KD 3) and these were compared to four infant control cases and a DLB case which was included as a synucleinopathy disease control (Supplementary Table 1).

Supplementary Methods 2: Immunohistochemical and fluorescent staining

6 µm sections from KD frontal, temporal, parietal, and occipital cortex, in addition to samples from medulla, midbrain, and cerebellum, were initially stained with Syn1 and p62/SQSTM1 to identify the presence of Lewy bodies (Supplementary Methods Table 1). Frontal cortex, temporal cortex and medulla only was evaluated in KD4 due to limited tissue available. Additional α-synuclein antibodies (Syn-O2, 5G4 and pS129) were evaluated in cortex and compared to controls, based on observations from Syn1 staining (Supplementary Methods Table 1). Antibody visualisation used the Menarini MenaPath X-Cell Linked Plus HRP detection kit, as per manufacturer's instructions (Menarini Diagnostics, Berkshire, UK). Sections were evaluated on a Nikon Eclipse 90i microscope with DsFi1 camera coupled to a computer with NIS Elements software v3 (Nikon, Tokyo, Japan). Immunofluorescent staining was performed to identify co-localisation of proteins of interest and visualised with secondary antibodies conjugated to fluorophores (Supplementary Methods Table 1). Fluorescence was evaluated on a Leica SP8 confocal microscope coupled to a computer with LASX software (Leica, Wetzlar, Germany).

Supplementary Methods Table 1: Antibodies using for immunohistochemistry and immunofluorescence

IHC antibody	Isotype	Manufacturer	Dilution	Antigen retrieval
Anti-α-synuclein Syn1/Clone 42	Mouse IgG1	BD Biosciences #610786	250 µg/ml	Citrate pH 6, formic acid
Anti-oligomeric α-synuclein (Syn-O2)	Mouse IgG1	El-Agnaf laboratory	100 ng/ml	Citrate pH 6, formic acid
Anti-aggregated α-synuclein (5G4)	Mouse IgG1	Merck MABN389	200 ng/ml	Citrate pH 6, formic acid

Anti-α-synuclein pS129	Rabbit IgG	Abcam ab51253	6 µg/ml	Citrate pH 6, formic acid
Anti-SQSTM1/p62	Mouse IgG1	BD Biosciences #610832	8 µg/ml	Citrate pH 6

IF antibody	**Isotype**	**Manufacturer**	**Dilution**	**Secondary antibody**
Anti-α-synuclein Syn1/Clone 42	Mouse IgG1	BD Biosciences #610786	1 µg/ml	Goat anti-mouse IgG1-AF488 (Thermo #A-21121)
Anti-oligomeric α-synuclein (Syn-O2)	Mouse IgG1	El-Agnaf laboratory	500 ng/ml	Goat anti-mouse IgG1-AF488 (Thermo #A-21121)
Anti-aggregated α-synuclein (5G4)	Mouse IgG1	Merck MABN389	200 ng/ml	Goat anti-mouse IgG1-AF488 (Thermo #A-21121)
Anti-SQSTM1/p62	Rabbit IgG	Abcam ab207305	2 µg/ml	Goat anti-rabbit IgG-AF546 (Thermo #A-11035)
Anti-NeuN	Mouse IgG2b	Abcam ab104224	5 µg/ml	Goat anti-mouse IgG2b-647 (Thermo #A-21242)

Supplementary Methods 3: Frozen tissue preparation for SDS-PAGE/western blot

Temporal cortex (Brodmann area 20) was chosen for molecular analysis of α-synuclein as it contained sufficient tissue for downstream analyses and consistently contained globoid cells and neuritic and neuronal α-synuclein pathology in KD cases. Frozen tissue was only available for two KD cases (KD 1 and KD 3) and these were compared to four infant control cases and a DLB case which was included as a synucleinopathy disease control (Table 1). Tissue was fractionated by centrifugation into aqueous- and detergent- soluble fractions (Supplementary Methods 1). Staining of western blot membranes used antibodies detailed in Supplementary Methods Table 2.

Supplementary Methods Table 2: Antibodies used for western blot analysis

Antibody	**Manufacturer**	**Dilution**	**Secondary antibody**	**Dilution**
Anti-α-synuclein Syn1/Clone 42	BD Biosciences #610786	250 ng/ml	Goat anti-mouse	1:5,000
Anti-α-synuclein pS129	Abcam ab51253	3 μg/ml	Goat anti-rabbit	1:2,500

Supplementary Methods 4: Purification of α-synuclein and preparation of tissue for RT-QuIC

α-synuclein was produced using the pRK172 plasmid containing cDNA for the human *SNCA* gene expressed in E. Coli BL21 (DE3) and purified using size exclusion and Mono Q anion exchange chromatography. α-Synuclein was diluted in 20 mM Tris/HCl pH 7.4, 100 mM NaCl, and protein concentrations were determined using the bicinchoninic acid assay (BCA; Pierce, Waltham, MA, USA) and aliquots (300 µL of 1 mg/mL) were prepared and stored at −80 °C. Prior to use, the proteins were filtered (100 kDa spin filter) and the protein concentration was again determined by BCA assay.

Brain tissue lysates from KD1 and KD3 and from four infant control cases were homogenised 10% w/v on ice in TBS (20 mM Tris-HCl pH 7.4, 150 mM NaCl) and 5 mM EDTA with protease and phosphatase inhibitors (Thermo Fisher, Paisley, UK). Samples were centrifuged at 3000 × g at 4 °C for 30 min and the supernatant recovered as the TBS (aqueous)-soluble fraction. The pellet was then resuspended in CelLytic buffer (Sigma, St Louis, MO, USA), homogenized on ice, and centrifuged at 3000× g at 4 °C for 30 min with the supernatant recovered as the detergent-soluble fraction. The total protein concentration was measured in both fractions by BCA assay (Pierce, Thermo Fisher, Paisley, UK) and aliquots were prepared and stored at −80 °C.

Supplementary Methods 5: *GALC* gene and genetic risk of DLB

This analysis utilized data described in two previous studies: an exome sequenced cohort of 1,118 DLB patients [1] and a cohort of whole-genome genotyped samples described in [2] (Supplementary Methods Table 3). The data described here has been lifted over to the most recent genome build (GRCh38/hg38). The genotyping data were imputed on the TOPMed platform [3].

Supplementary Methods Table 3: Cohort descriptions. Whole exome sequencing (WES) and whole genome genotyped (WGG) data.

Cohort Name	Publication	Data format	Cases	Controls	Average Age
DLB exomes	[1]	WES	1,118	0	78.8
DLB WGG	[2]	WGG	1,296	4,379	79.64

Quality control

Individual datasets were subject to variant QC conducted primarily in bcftools version 1.12 [4]. All datasets were aligned to the latest reference genome panel (GRCh38/hg38). Variant annotations were generated using SNPeff [5]. Variant and sample QC involved removing samples with low variant quality (DP < 8, GQ < 20, and missingness > 0.15) and discordant allelic depth and genotypes. Additional sequencing and QC information on each dataset can be found in corresponding publications. Further genotyping samples described in the CNV analysis, but not included in the cohort from Guerreiro et al (2018), were genotyped on Illumina's HumanOmniExpress chip.

Single variant association

The association to disease status of both rare and common variants was assessed using logistic and Firth regression. Logistic regression was performed on common (MAF > 0.05) variants using PLINK 2.0 [6] using the first 10 principal components (PCs) as covariates. Rare variant

(MAF < 0.05) association was conducted using Firth regression, given its superior ability to control for type I errors [7, 8]. This test was performed using the implementation available in rvtests [9], also using the first 10 PCs as covariates.

Since no controls were available to complement the WES data, Gnomad allele counts from the non-Finnish European population, specifically those taken from the non-neurologically affected subset (https://gnomad.broadinstitute.org/about), were used in a fisher exact test to determine significance of variants. This test was implemented in R using the fisher.test() function.

Gene burden analysis

Gene burden tests were conducted using SkatO as available through rvtests [9]. The first 10 PCs were used as covariates. The burdens of missense, high impact, and missense plus loss of function variants were conducted separately. Variant functional impact was determined using SNPeff [5].

CNV analysis

The entirety of the genotyping samples (n=1,296) and a subset of the exome sequenced samples with genotyping information available (n=174) were screened for CNVs. CNVs were called on GenomeStudio with cnvPartition v.3.2.0 with standard parameter and the GenTrain score > 0.7 filter ON. Results from cnvPartition were exported and CNVs on and around the gene *GALC* were analysed and confirmed visually on GenomeStudio.

References

1. Orme, T., et al., *Analysis of neurodegenerative disease-causing genes in dementia with Lewy bodies.* Acta Neuropathol Commun, 2020. **8**(1): p. 5.

2. Guerreiro, R., et al., *Investigating the genetic architecture of dementia with Lewy bodies: a two-stage genome-wide association study.* Lancet Neurol, 2018. **17**(1): p. 64-74.

3. Taliun, D., et al., *Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program.* Nature, 2021. **590**(7845): p. 290-299.

4. Danecek, P., et al., *Twelve years of SAMtools and BCFtools.* Gigascience, 2021. **10**(2).

5. Cingolani, P., et al., *A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.* Fly (Austin), 2012. **6**(2): p. 80-92.

6. Chang, C.C., et al., *Second-generation PLINK: rising to the challenge of larger and richer datasets.* Gigascience, 2015. **4**: p. 7.

7. Ma, C., et al., *Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants.* Genet Epidemiol, 2013. **37**(6): p. 539-50.

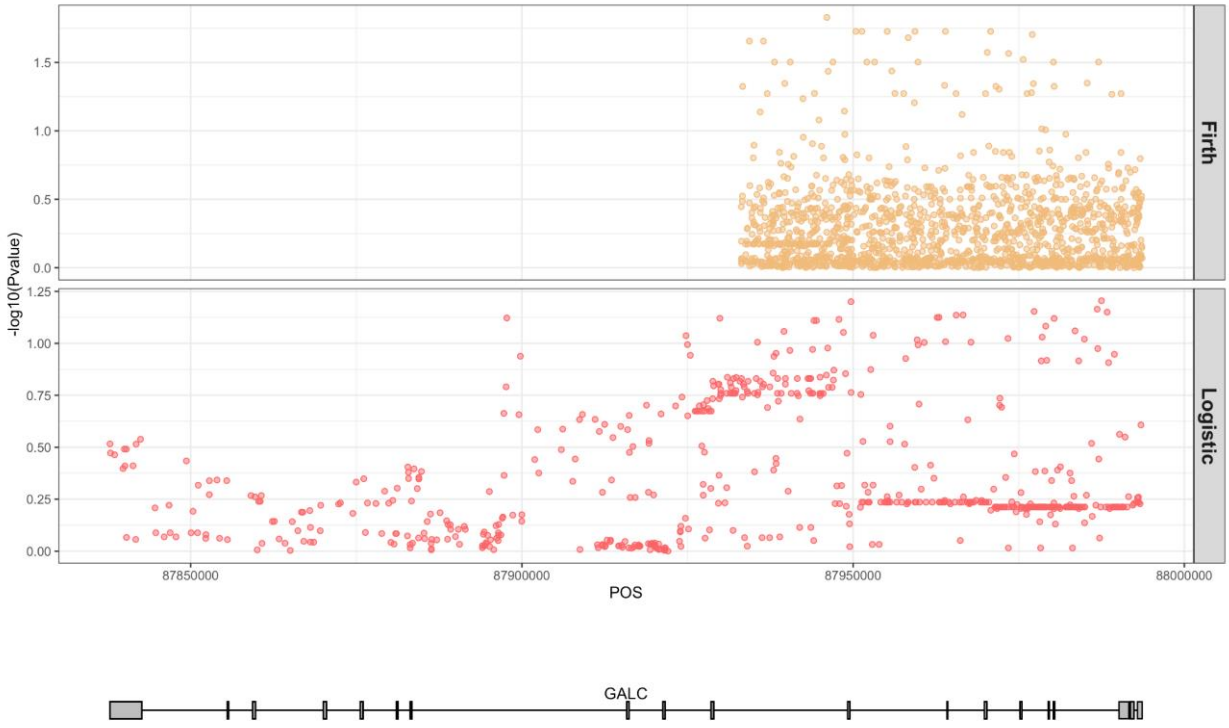
8. Wang, X., *Firth logistic regression for rare variant association tests.* Front Genet, 2014. **5**: p. 187.

9. Zhan, X., et al., *RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data.* Bioinformatics, 2016. **32**(9): p. 1423-6.

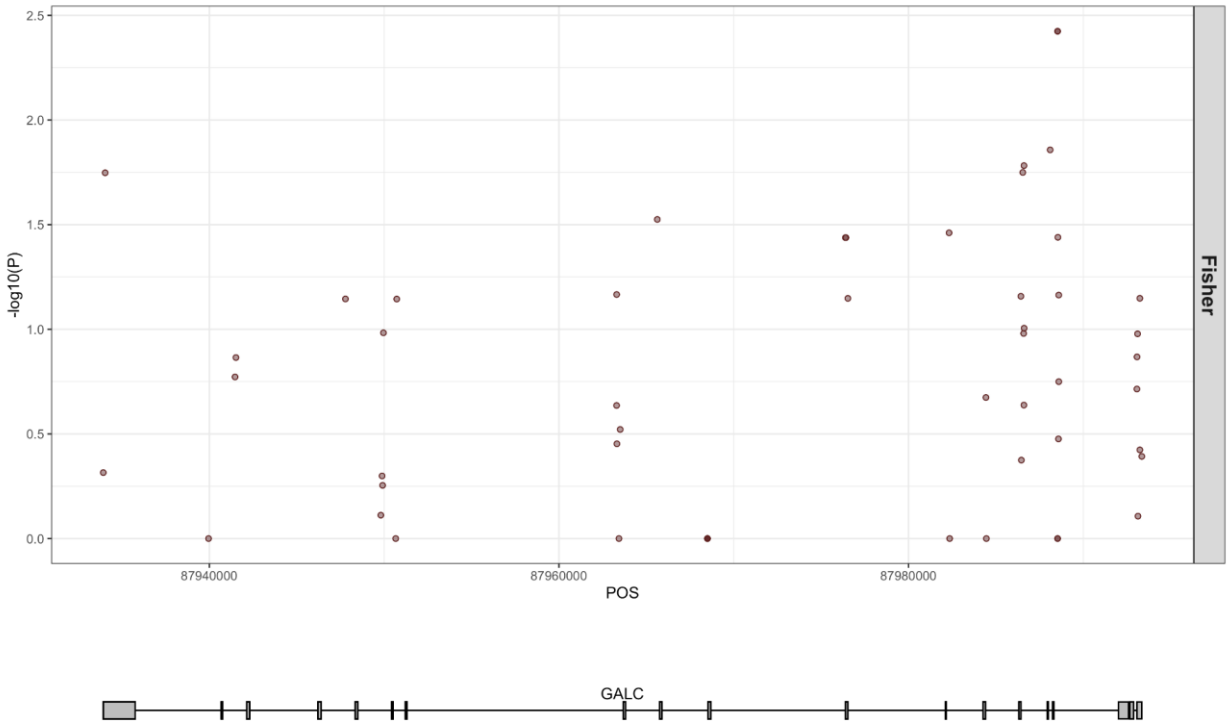
Supplementary Results

Supplementary Results 1: Single variant association

We detected no association between common variants within the *GALC* gene region and disease status (Fig 1). We also detected no association between rare variants within the *GALC* gene region (Figure 1 & 2). The most significant result was a variant, rs561184126 (14-87988525-T-TA), from the fisher test in the exomes with a p-value of 0.003 and an OR of 53.83.



Supplementary Results Figure 1. Logistic (bottom) and Firth (top) regression results for the whole genome



Supplementary Results Figure 2. Fisher exact test results comparing allele counts in DLB exomes to those in non-neurologically affected non-Finnish Europeans from the Gnomad database

Supplementary Results 2: Gene burden analysis

Gene burden analysis demonstrated no significant associations in SkatO (Supplementary Data Table 1).

Supplementary Table 1: SkatO gene burden results for exome and genotyping data

	Variant	**NumVar**	**NumPolyVar**	**Q**	**rho**	**pvalue**
DLB Exomes vs ADSP Exome Controls	High Impact	14	3	NA	NA	NA
	Missense + LOF	112	33	198372	0	0.186102
	Missense	98	30	110412	1	0.062663
DLB WGG	High Impact	3	2	22.0663	1	0.668939
	Missense + LOF	48	33	6242.71	1	0.474872
	Missense	45	31	16177.9	1	0.144063

Supplementary Results 3: Prioritised rare variants

No high-impact variants were found in the genotyping data. There were 2 heterozygous missense variants found only in DLB cases (1 case each). The first of these variants, rs777305549, had a CADD score of 22.5 and a frequency of 1.55E-05 in the Gnomad Non-Finnish European population. The second of these variants, rs371523347, was reported on ClinVar with uncertain significance for galactosylceramide beta-galactosidase deficiency. This variant had a frequency of 9.30E-05 in the NFE Gnomad population, and a CADD score of 32.

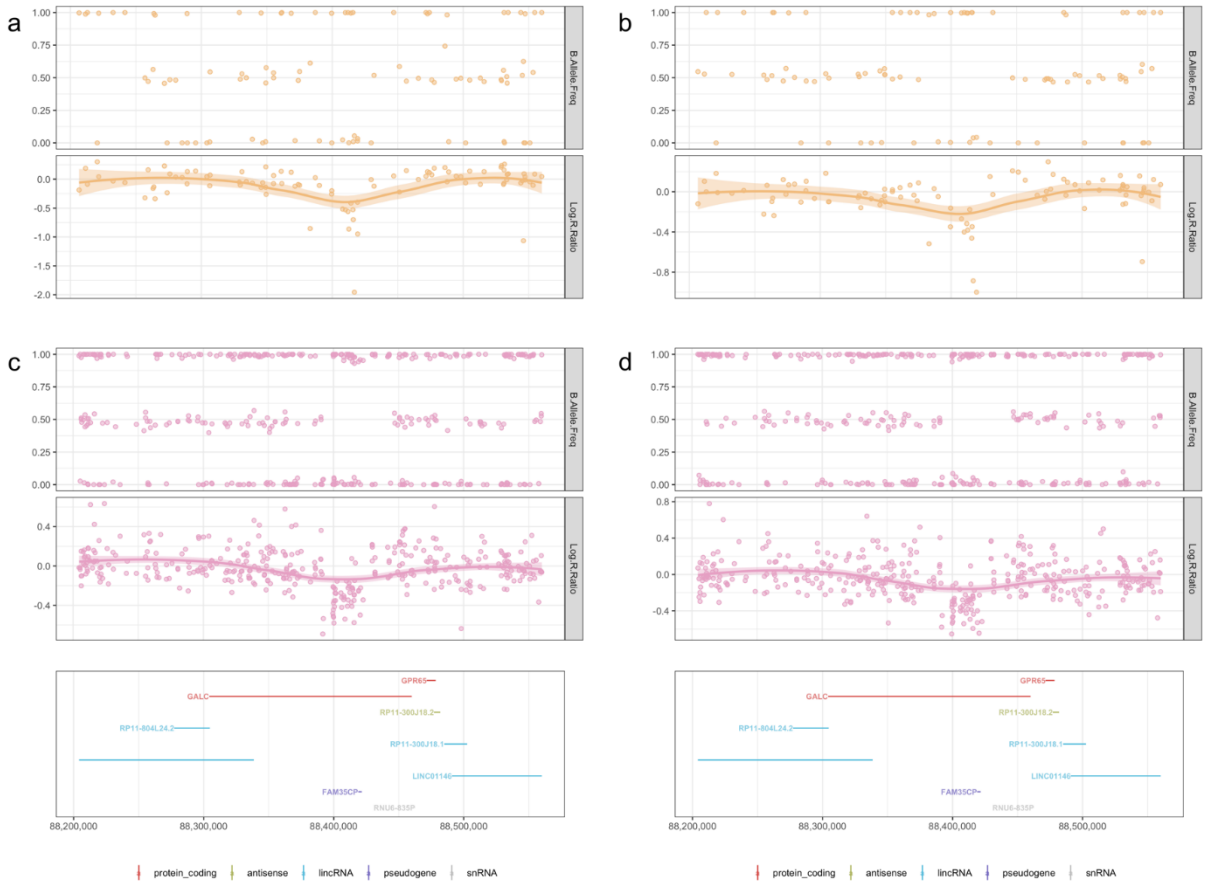
The exome data similarly did not contain any case-specific homozygous variants. The most notable variants from the exomes were four missense, heterozygous variants with CADD scores > 20. However, none of these variants were described on ClinVar (Supplementary Data Table 2).

Supplementary Results Table 2: The most interesting prioritized variants from exome and genotyping data. Positions given in hg38.

SNP ID	Position	REF/ALT	Amino Acid Change	Gnomad MAF NFE	CADD	ClinVarReport	Data source
rs777305549	87950702	T/A	p.Asn403Ile	1.55E-05	22.5	NA	WGG
rs371523347	87988513	C/T	p.Arg69Gln	9.30E-05	32	uncertain significance for galactosylceramide beta-galactosidase deficiency	WGG
chr14:87984483:A:G	87984483	A/G	p.Tyr165His	NA	31	NA	WES
rs772928724	87993063	A/C	p.Cys34Trp	7.75E-05	27.6	NA	WES
chr14:87993083:C:G	87993083	C/G	p.Ala28Pro	NA	22.6	NA	WES
rs372285275	87993100	C/T	p.Gly22Asp	1.55E-05	22.8	NA	WES

Supplementary Results 4: CNV analysis

Of the 1,470 samples which CNVs were called in, four samples were found to have heterozygous deletions within the *GALC* gene (Supplementary Results Figure 3). Breakpoints and sizes of these CNVs are described in Table 7. This frequency of minor alleles (0.001) is in line with what would be expected for European populations (MAF gnomAD European for DEL_14_144665 = 0.001), suggesting that heterozygosity of this deletion (that is known to be pathogenic in homozygosity for Krabbe disease) is not associated with the risk of DLB.



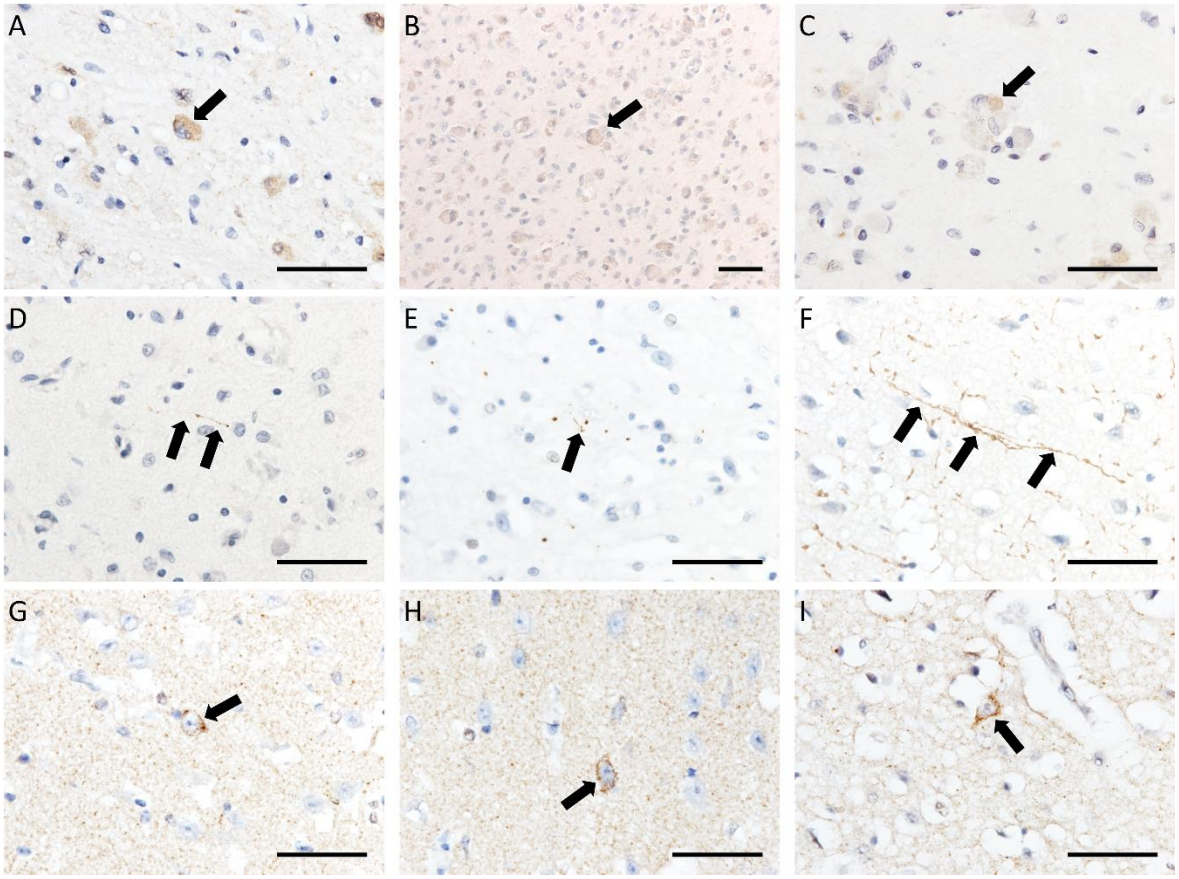
Supplementary Results Figure 3: CNVs called within GALC gene region. The colour of the SNPs represents genotyping chip (Yellow (a.b) = OmniX, Pink (c,d) = HumanOmniExpress). Gene positions were derived from Ensembl, and both genes and CNVs are plotted in build GRCh37.

Supplementary Results Table 3: Breakpoints, lead SNPs, and sizes for CNVs. Plot column indicates label in Figure 10 representing CNV. Positions on chromosome 14 are listed in GRCh37. Sizes are listed in kilobases.

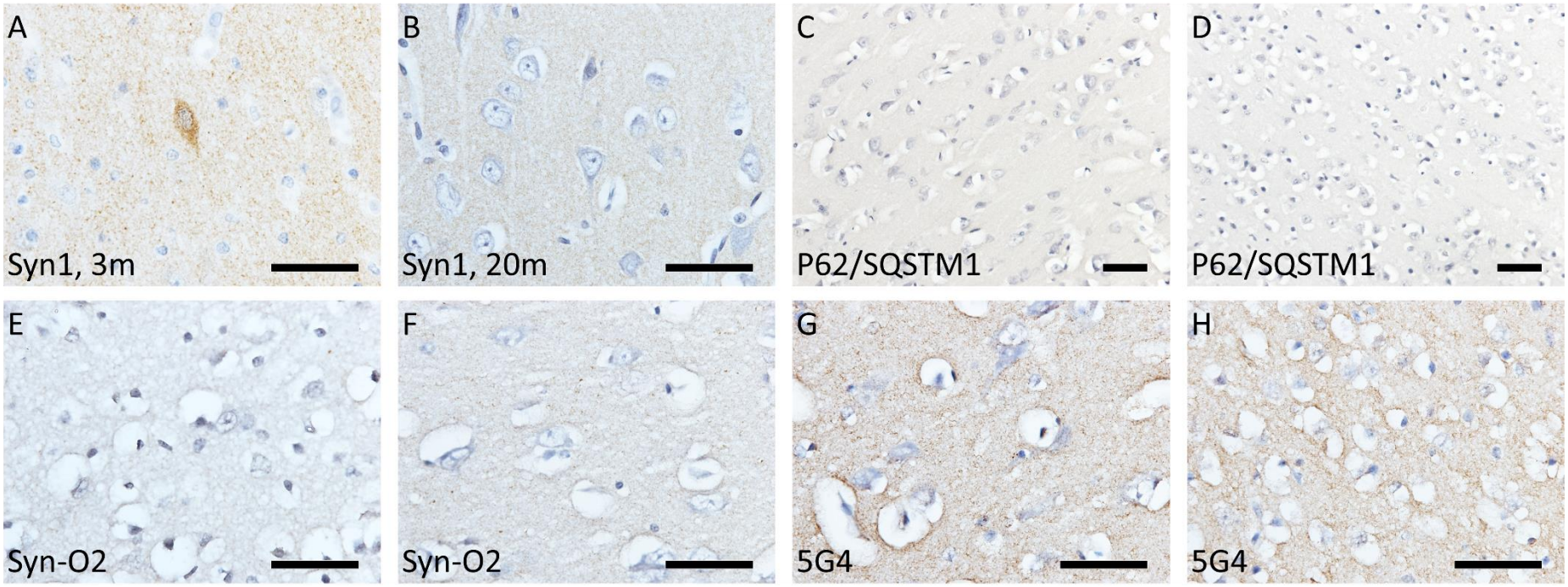
Plot Label	Start SNP	End SNP	Start Position	End Position	Size (Kb)
a	rs440765	rs17687109	88390135	88429690	39.555
b	rs382299	rs12888666	88383098	88431898	48.8

c	rs10142867	rs72629370	88391765	88422740	30,975
d	rs10142867	rs17123945	88391765	88420788	29.023

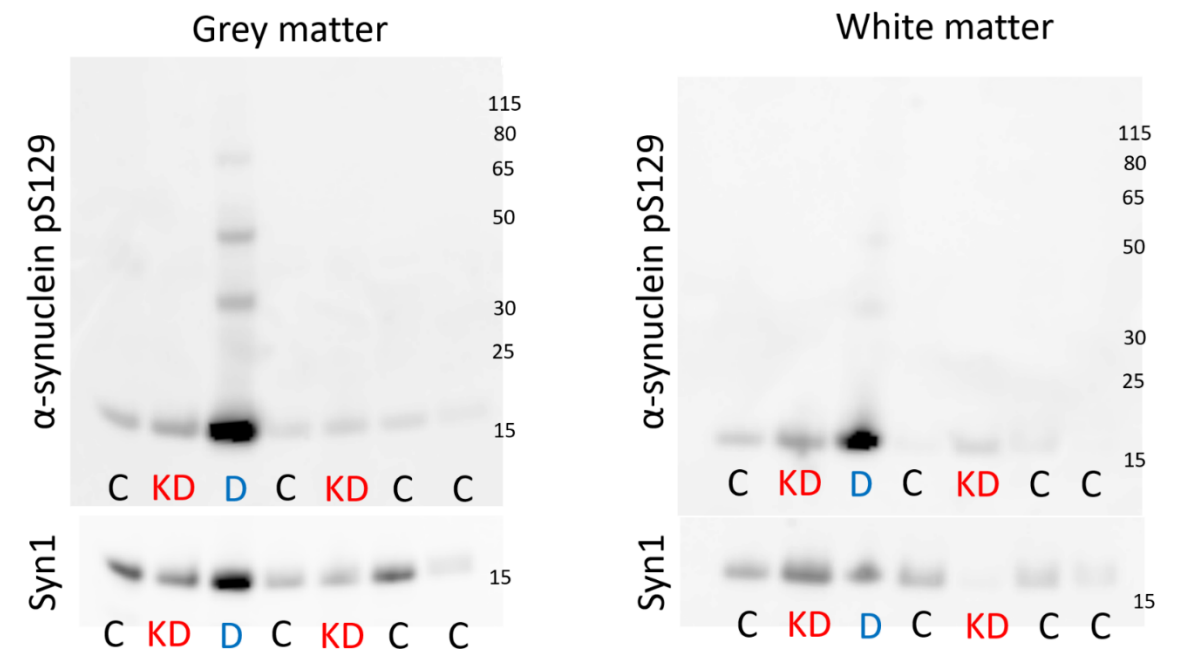
Supplementary Figures



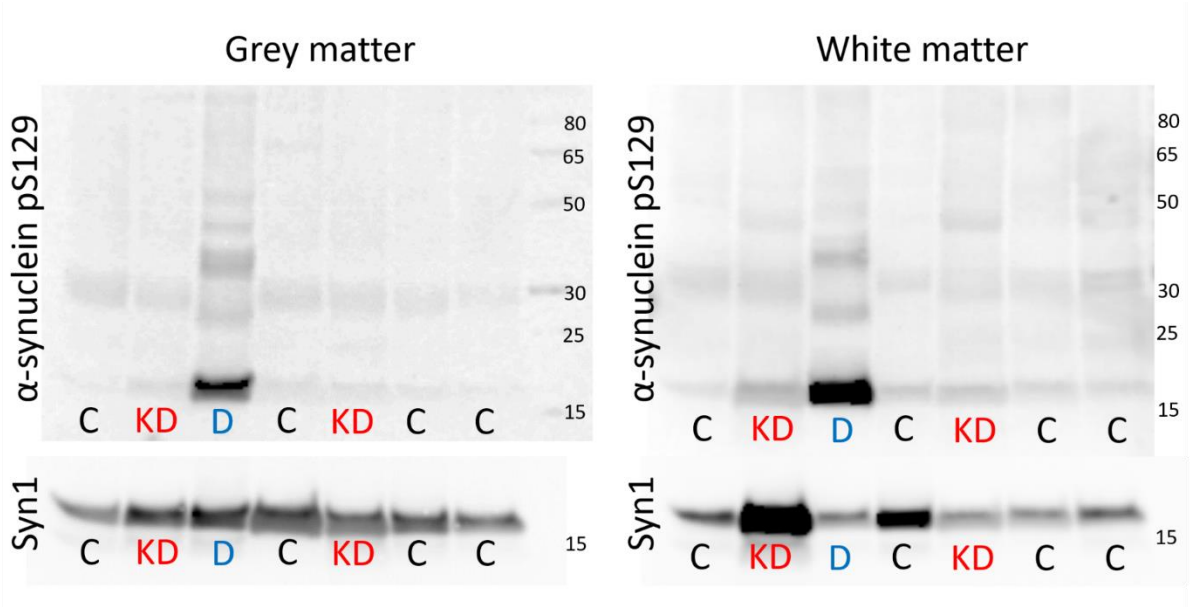
Supplementary Figure 1: Photomicrographs from KD cases demonstrating Syn1 immunoreactivity in globoid cells (A-C), p62+ neurites (D-F), and Syn-O2 in pyramidal neurons (G-I). Cases are KD1 (E & I), KD2 (B), KD3 (C, F & G). KD4 (A, D & H), scale bars are 50 μm.



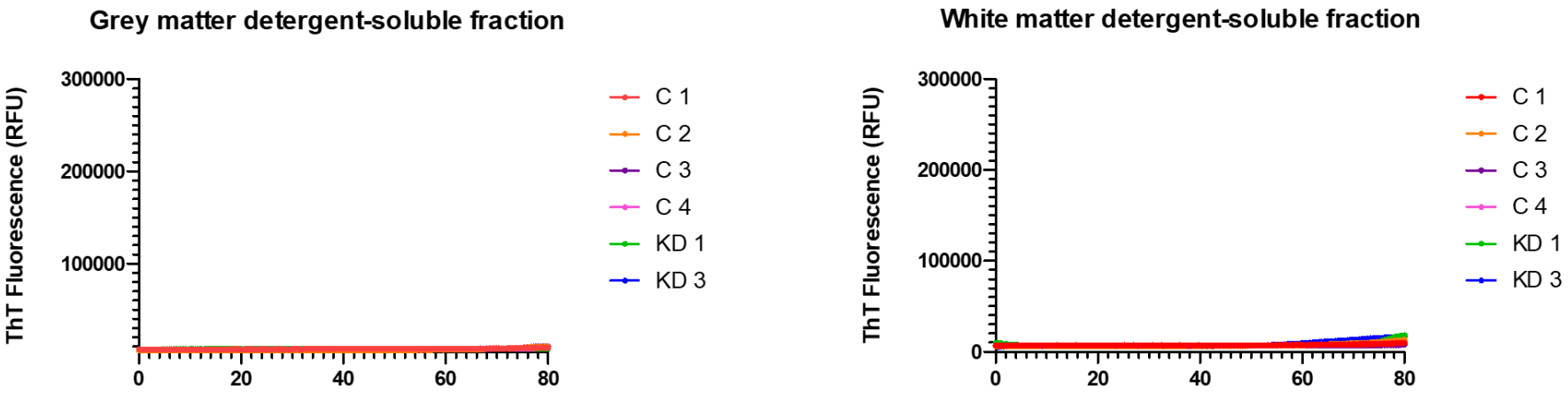
Supplementary Figure 2: Photomicrographs from infant control cases demonstrating α-synuclein in a sub-population of cortical pyramidal neurons in a three month old infant (A), but no immunoreactivity in a 20 month old infant (B). An absence of p62/SQSTM1 (C-D), Syn-O2 (E-F) and 5G4 (G-H) was also demonstrated in control cases. Scale bars = 50 μm.



Supplementary Figure 3: Western blot of crude temporal cortex lysates immunostained with pS129 (top) and Syn1 (bottom), KD cases (red) are indistinguishable from controls (black). In contrast, the DLB case (blue) demonstrates higher levels of pS129 in both grey and white matter, though high molecular weight bands are only obvious in grey matter. In all cases, Syn1 is variable across cases and particularly marked in DLB grey matter.



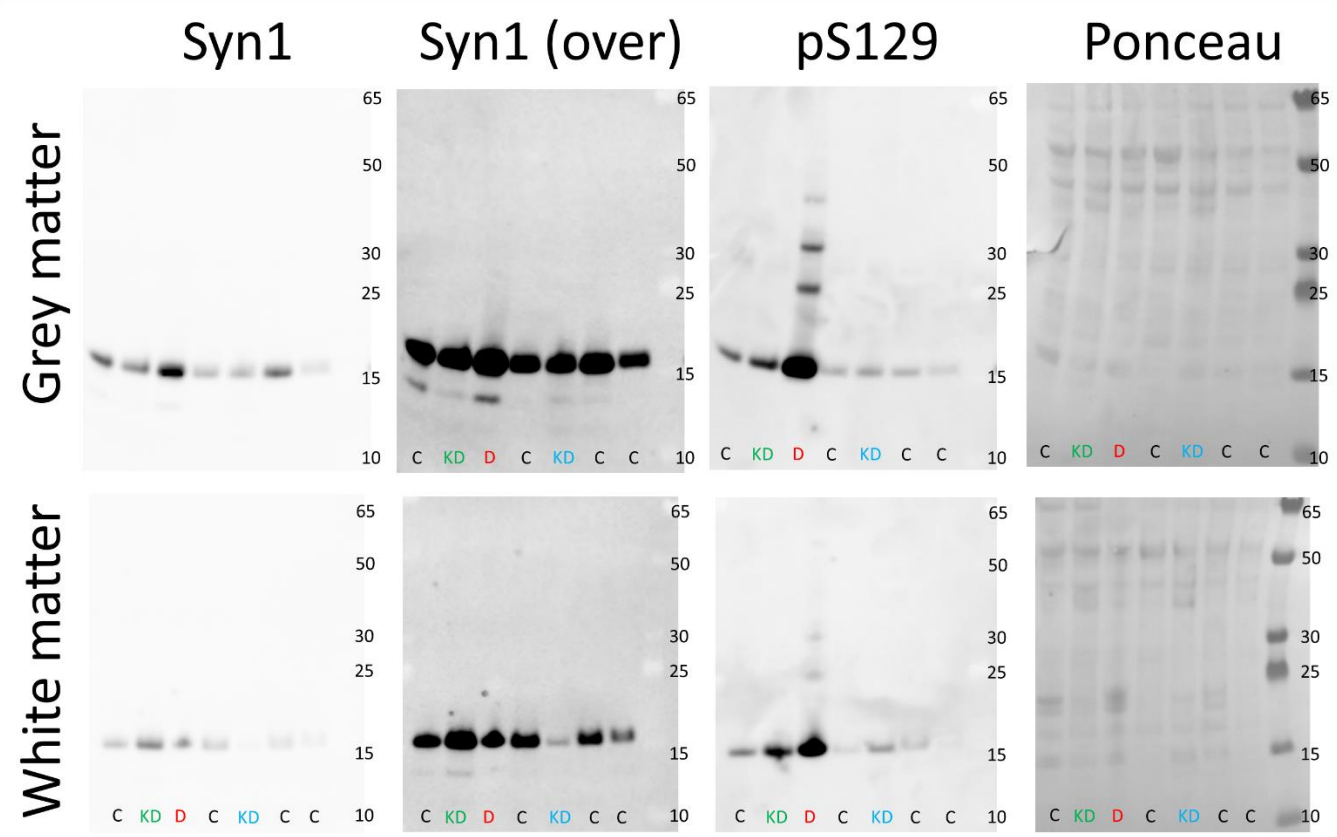
Supplementary Figure 4: Western blot of Triton X-100-soluble tissue lysates immunostained with pS129 (top) and Syn1 (bottom). KD cases (red) have higher molecular weight bands of pS129 at approximately 45 kDa in white matter (blue arrow) but are no different from controls in grey matter. The DLB case (blue) has numerous high molecular weight pS129 bands in both grey and white matter. Syn1 is not different in DLB compared to control in grey or white matter but KD2 has strikingly high levels of Syn1 in white matter. It is notable that KD2 had very high densities of Syn1-immunoreactive globoid cells in this area.



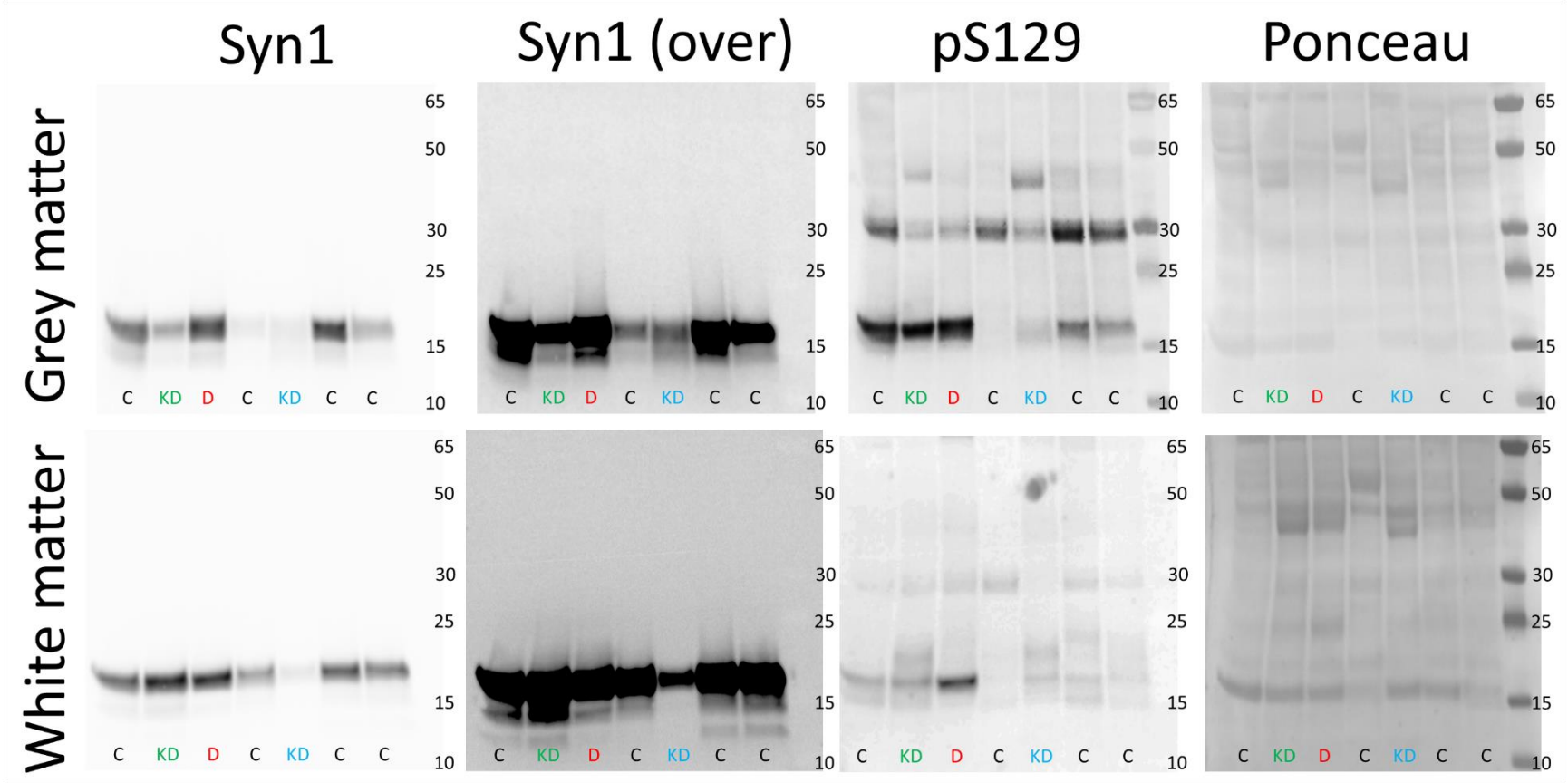
Supplementary Figure 5: RT-QuIC data demonstrating that the detergent soluble fractions of both grey and white matter did not demonstrate a positive reaction with the α-synuclein RT-QuIC assay.

Supplementary Data

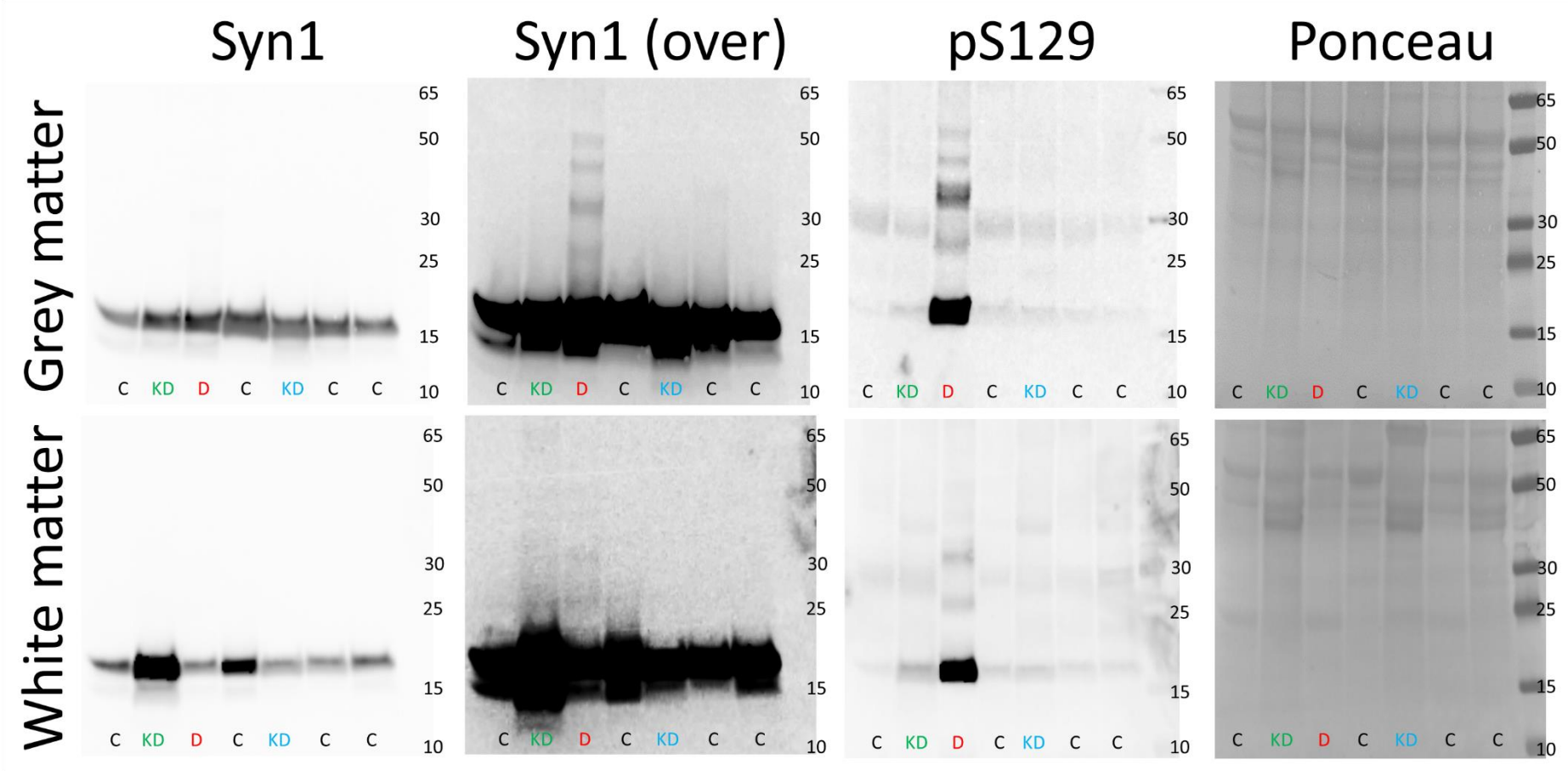
Supplementary Data 1: Complete western blot membranes



Supplementary Figure 1: Western blot membranes from crude tissue lysates, including under-exposed Syn1 membranes (Syn1) and over-exposed Syn1 membranes (Syn1 over), pS129 and the total protein stain Ponceau S, in both grey and white matter.



Supplementary Figure 2: Western blot membranes from aqueous-soluble tissue lysates, including under-exposed Syn1 membranes (Syn1) and over-exposed Syn1 membranes (Syn1 over), pS129 and the total protein stain Ponceau S, in both grey and white matter.



Supplementary Figure 3: Western blot membranes from aqueous-soluble tissue lysates, including under-exposed Syn1 membranes (Syn1) and over-exposed Syn1 membranes (Syn1 over), pS129 and the total protein stain Ponceau S, in both grey and white matter.

Supplementary Data 2: Acknowledgements for International DLB Genetics Consortium members

For the neuropathologically confirmed samples from Australia, tissues were received from the Sydney Brain Bank which is supported by Neuroscience Research Australia and the University of New South Wales. We would like to thank the South West Dementia Brain Bank (SWDBB) for providing brain tissue for this study. The SWDBB is supported by BRACE (Bristol Research into Alzheimer's and Care of the Elderly), Brains for Dementia Research and the Medical Research Council. The brain samples and/or bio samples were obtained from The Netherlands Brain Bank, Netherlands Institute for Neuroscience, Amsterdam (open access: www.brainbank.nl). All Material has been collected from donors for or from whom a written informed consent for a brain autopsy and the use of the material and clinical information for research purposes had been obtained by the NBB. This study was also partially funded by the Wellcome Trust, Medical Research Council and Canadian Institutes of Health Research (Dr. St. George-Hyslop). Work from Dr. Compta was supported by the CERCA Programme / Generalitat de Catalunya, Barcelona, Catalonia, Spain. This study was also partially funded by the Wellcome Trust, Medical Research Council, Canadian Institutes of Health Research, Ontario Research Fund. The Nottingham Genetics Group is supported by ARUK and The Big Lottery Fund. The effort from Columbia University was supported by the Taub Institute, the Panasci Fund, the Parkinson's Disease Foundation, and NIH grants NS060113 (L. Clark), P50AG008702 (P.I. Scott Small), P50NS038370 (P.I. R. Burke), and UL1TR000040 (P.I. H. Ginsberg). O.A.R. is supported by the Michael J. Fox Foundation, NINDS R01# NS078086. The Mayo Clinic Jacksonville is a Morris K. Udall Parkinson's Disease Research Center of Excellence (NINDS P50 #NS072187) and is supported by The Little Family Foundation and by the Mangurian Foundation Program for Lewy Body Dementia research and the Alzheimer Disease Research Center (P50 AG016547). The work from the Mayo Clinic Rochester is supported by the National Institute on Aging (P50 AG016574 and U01 AG006786). This work has received support from The Queen Square Brain Bank at the UCL Institute of Neurology; where TL is funded by an ARUK senior fellowship. Some of the tissue samples studied were provided by the MRC London Neurodegenerative Diseases Brain Bank and the Brains for Dementia Research project (funded by Alzheimer's Society and ARUK). This research was supported in part by both the NIHR UCLH Biomedical Research Centre and the Queen Square Dementia Biomedical Research Unit. This work was supported in part by the Intramural

Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services; project AG000951-12. The University of Pennsylvania case collection is funded by the Penn Alzheimer's Disease Core Center (AG10124) and the Penn Morris K. Udall Parkinson's Disease Research Center (NS053488). The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. Tissue samples from UCSD are supported by NIH grant AG05131. The authors thank the brain bank GIE NeuroCEB, the French program "Investissements d'avenir" (ANR-10-IAIHU-06). PJT and LM are supported by the Helsinki University Central Hospital, the Folkhälsan Research Foundation and the Finnish Academy.