

Cell Reports

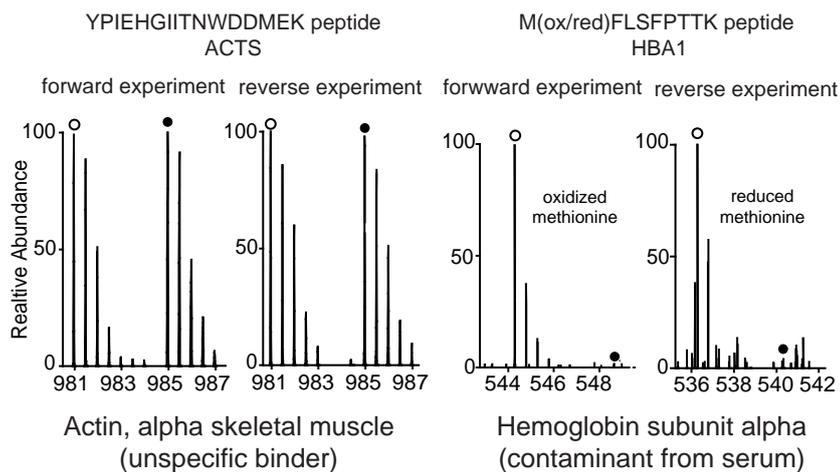
Supplemental Information

Quantitative Interaction Proteomics of Neurodegenerative Disease Proteins

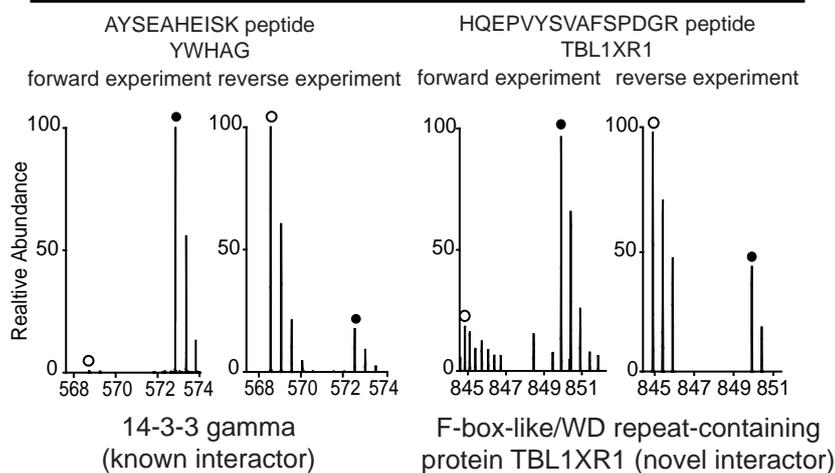
Fabian Hosp, Hannes Vossfeldt, Matthias Heinig, Djordje Vasiljevic, Anup Arumughan, Emanuel Wyler, the Genetic and Environmental Risk for Alzheimer's Disease (GERAD1) Consortium, Markus Landthaler, Norbert Hubner, Erich E. Wanker, Lars Lannfelt, Martin Ingelsson, Maciej Lalowski, Aaron Voigt, and Matthias Selbach

Hosp et al. *Figure S1 (Related to Figure 1A)*

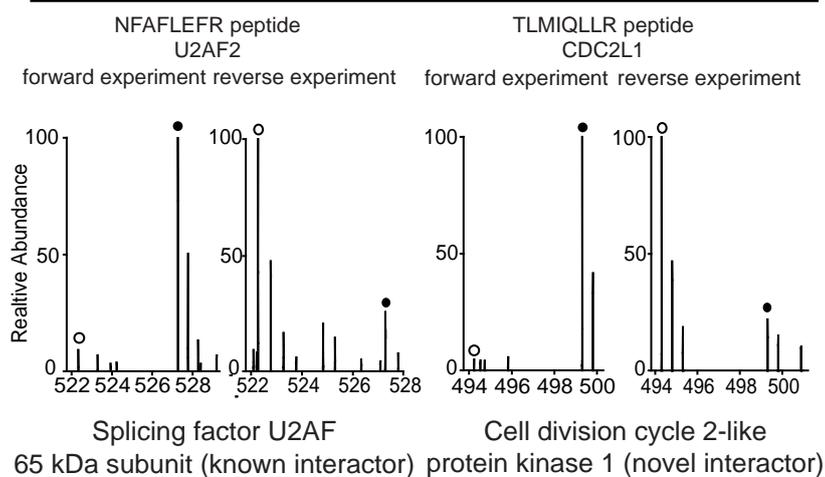
A myc-tag control vs. Ataxin1-Q30



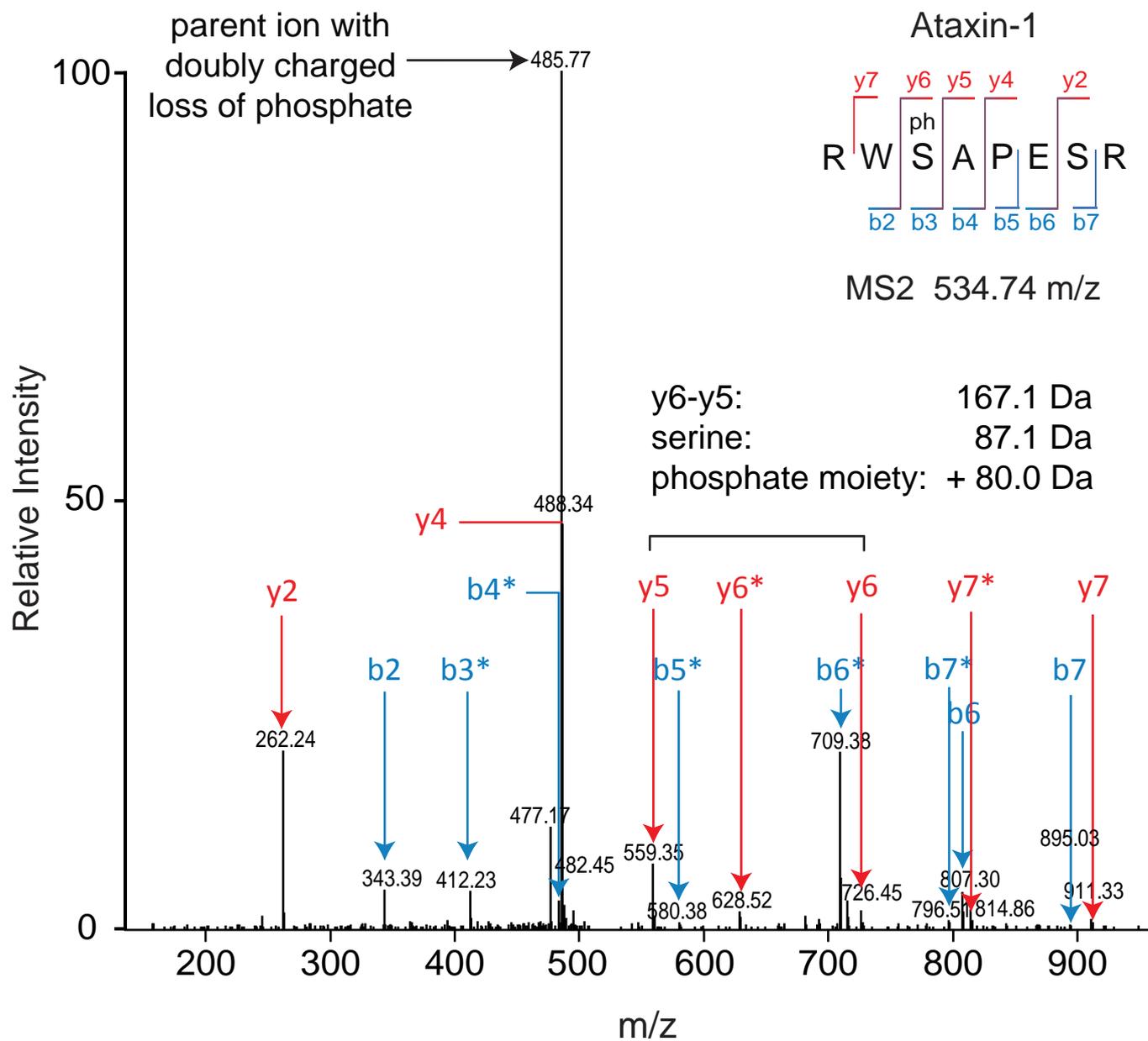
B myc-tag control vs. Ataxin1-Q30



C myc-tag control vs. Ataxin1-Q82



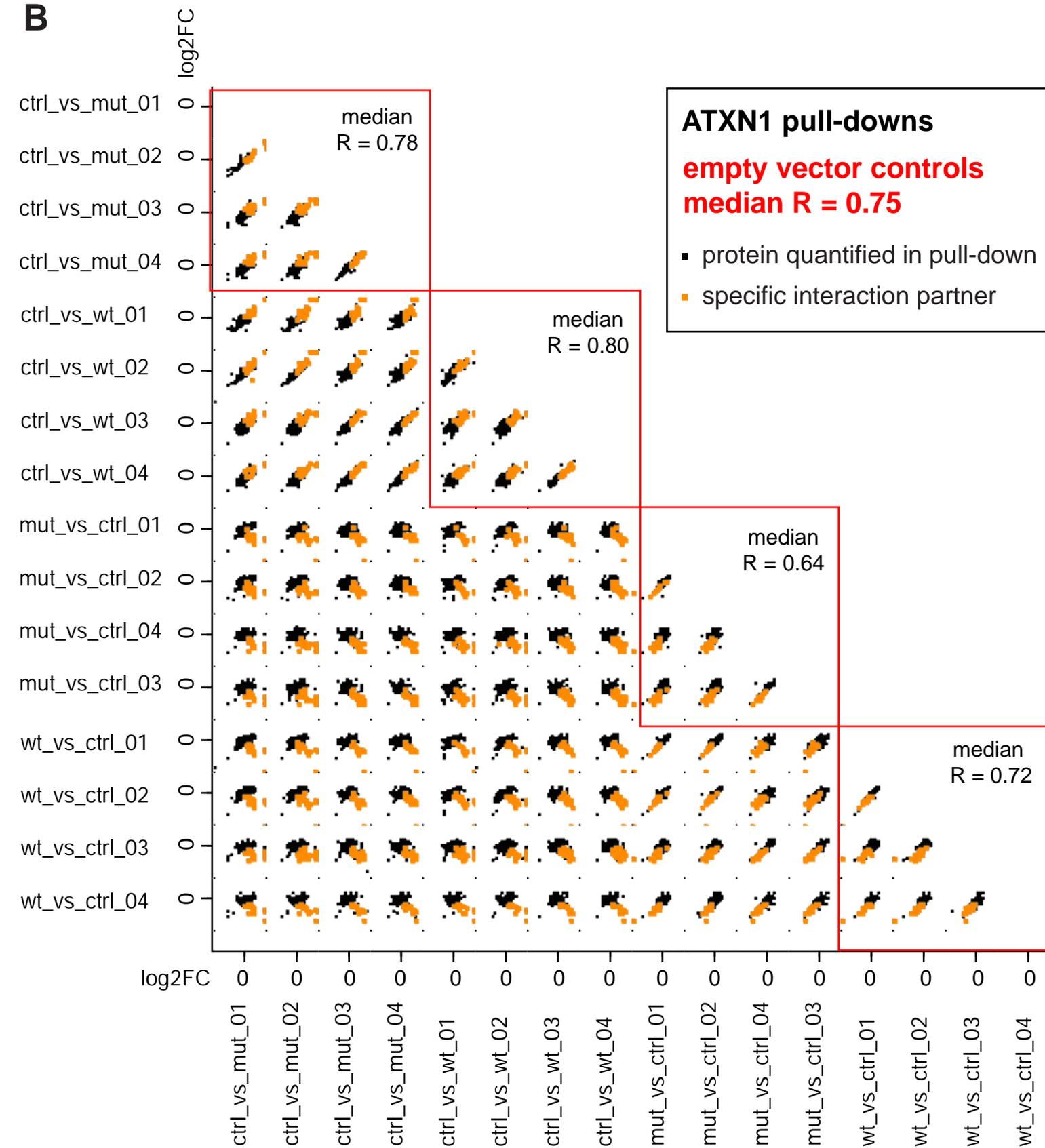
Hosp et al. *Figure S2 (Related to Figure 1B+C)*



A

#	Neurodegenerative disease	Gene	Name	Variant	median R
1	Alzheimer's disease (AD)	APP	Amyloid beta A4 protein	wild-type	0.47
2				K670N/M671L	
3		PSEN1	Presenilin-1	wild-type	0.94
4				A431E	
5	Huntington's disease (HD)	HTT	Huntingtin (N-term 506 aa)	wild-type Q23	0.96
6				mutant Q145	
7			Huntingtin (Exon1)	wild-type Q23	0.92
8				mutant Q79	
9	Parkinson's disease (PD)	PARK2	Parkin	wild-type	0.94
10				Q329Stop	
11	Spinocerebellar ataxia type 1 (SCA1)	ATXN1	Ataxin-1	wild-type Q30	0.75
12				mutant Q82	

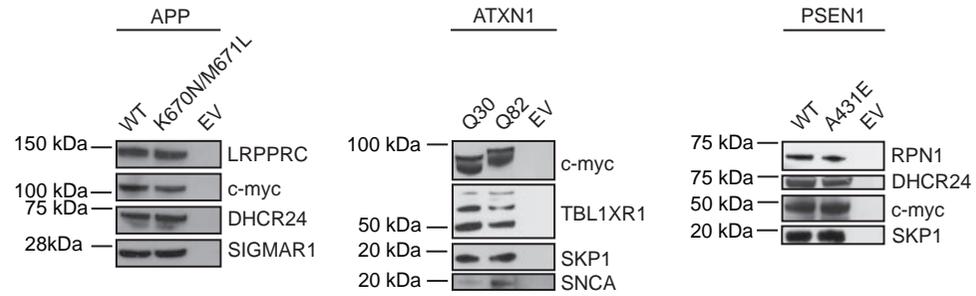
B



Hosp et al. Figure S4 (Related to Figure 1D)

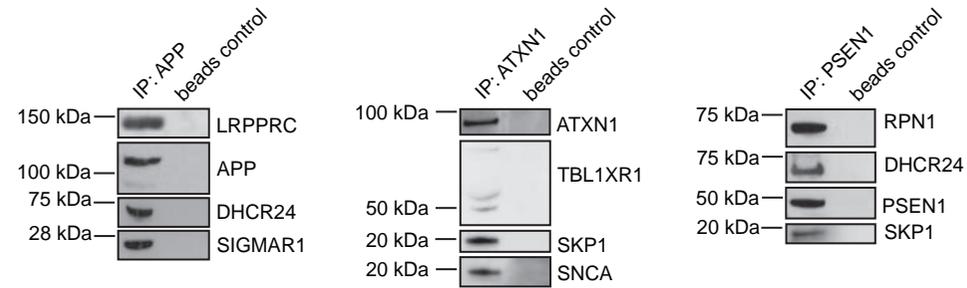
A

HEK293T (overexpressed 1x c-myc::bait)
IP: c-myc

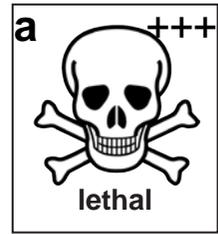


B

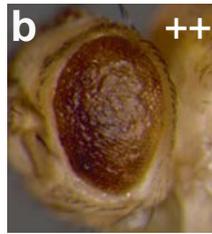
SH-SY5Y (endogenous bait)



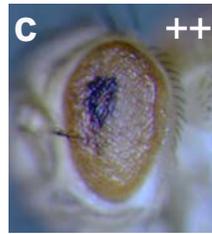
Hosp et al. *Figure S5 (Related to Figure 2)*



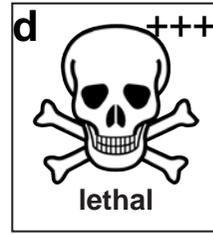
ATXN1-Q82
RNAi_CG4263
CDC2L1



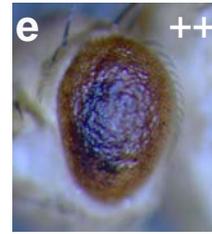
ATXN1-Q82
RNAi_CG6198
CHORDC1



ATXN1-Q82
RNAi_CG2128
HDAC3



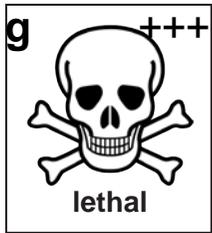
ATXN1-Q82
RNAi_CG4147
HSPA5



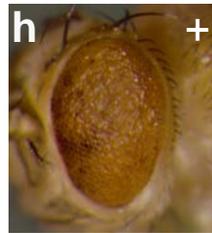
ATXN1-Q82
RNAi_CG8937
HSPA8



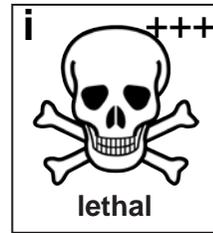
ATXN1-Q82
RNAi_CG1245
MED27



ATXN1-Q82
RNAi_CG4097
PSMB1



ATXN1-Q82
RNAi_CG8881
SKP1



ATXN1-Q82
RNAi_CG5677
SPCS3

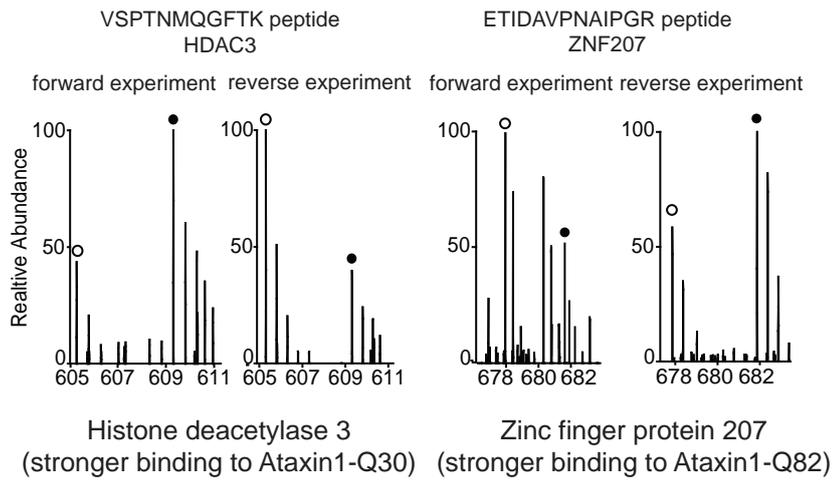


ATXN1-Q82
RNAi_CG17870
YWHAZ

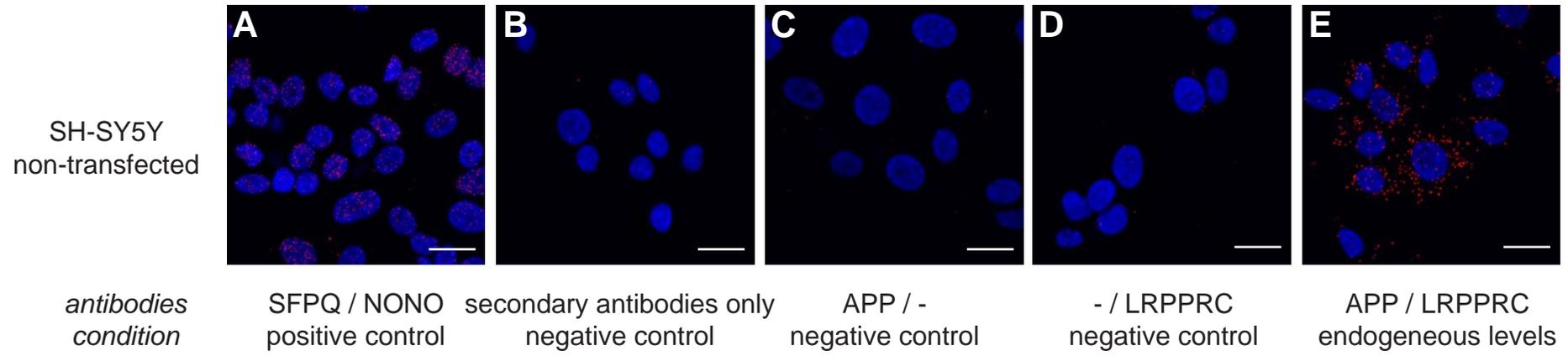
Hosp et al. *Figure S6 (Related to Figure 4B)*

A

Ataxin1-Q82 vs. Ataxin1-Q30



Hosp et al. *Figure S7 (Related to Figure 5C)*



Supplemental Experimental Procedures

Plasmids

For mammalian protein expression, wild-type and disease-associated variants of selected bait proteins were cloned into pcDNA5/FRT/TO_1xN-term_c-Myc overexpression vectors using the GATEWAY system (Invitrogen).

Cell culture and SILAC

HEK293T (System Biosciences) and SH-SY5Y (kind gift from the Erich Wanker lab, MDC Berlin) were cultivated at 37°C with 5% CO₂ and split every second or third day. SILAC media were essentially prepared as described (Paul et al., 2011). Briefly, we used Dulbecco's Modified Eagle Medium (DMEM) Glutamax lacking arginine and lysine (custom preparation from Gibco) supplemented with 10% (HEK293T) or 20% (SH-SY5Y) dialyzed fetal bovine serum (dFBS, Gibco). To prepare 'heavy' (H) SILAC media we added 28 mg/l ¹³C₆¹⁵N₄ L-arginine and 49 mg/l ¹³C₆¹⁵N₂ L-lysine. Labeled amino acids were purchased from Sigma Isotec. 'Light' (L) SILAC medium was prepared by adding the corresponding non-labeled amino acids (Sigma). Transient transfections were performed using linear polyethylenimine (Sigma with 15 µg plasmid DNA and 30 µg PEI per 2*10⁷ cells. Twenty-four hours after transfection, cells were washed once in ice-cold PBS and lysed by scratching in RIPA buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1% Na-Deoxycholate, 0.1% SDS) supplemented with complete protease inhibitor (Roche) and 1% Benzoylase (Merck). After incubation on ice for 30 min, cell debris was cleared by centrifugation at 13,000 g for 10 min.

Drosophila models and genetic screen

Flies were raised and maintained at 25°C on standard food. Crossbreeding of adult *Drosophila melanogaster* were conducted at 25°C. GMR-GAL4 control and UAS-ATXN1-82 disease strains were recombined in order to generate the screening stock for the polyQ modifier screen (GMR>ATXN1-Q82). GMR-ATXN1-Q82 virgins were crossed to males carrying UAS-RNAi constructs, purchased from the Vienna Drosophila RNAi Centre (VDRC) considering known or predicted human homologues to the fly genes (Homologene build64). F1 females (GMR>ATXN1-Q82) in combination with the respective UAS-RNAi expression) were selected for REP evaluation 1-5 days post eclosion. For assessment of REP modulation, at least five female flies were analyzed for changes in the severity of eye degeneration. RNAi lines exhibiting such effects in the GMR-GAL4 control flies were excluded from subsequent experiments due to impact unconnected to expression of elongated polyQ. Candidate lines were tested for polyQ specificity by rescreening with Tau[R406W] screening stock. *Drosophila* compound eyes were pictured using an Olympus zoom stereo microscope at 6.3x magnification and Cell A software.

Pull-down experiments

Immunoprecipitations of c-myc-tagged bait proteins were performed using the μ MACS c-myc Isolation Kit (Miltenyi Biotec) according to the manufacturer's instructions. Cell lysates with overexpressed c-myc::bait fusion proteins from 2×10^7 HEK293T cells were incubated with 75 μ l of anti-c-myc μ MACS beads for 60 min on ice. Next, the beads were transferred to a μ MACS M column placed in the magnetic field of the OctoMACS separator. The beads on the column were washed three times with RIPA buffer and once with wash buffer 2 that is part of the Miltenyi μ MACS kit. Proteins bound to the beads were eluted by applying three times 100 μ l elution buffer (100 mM glycine pH 2.5). The collected protein eluates were combined in a fresh 2 ml Eppendorf tube and set to ethanol precipitation and subsequent in-solution digestion followed by LC-MS/MS analysis.

Co-immunoprecipitations of endogenous proteins were performed by crosslinking specific anti-bait antibodies to NHS-activated Sepharose. Next, the antibody-coupled NHS-Sepharose was quenched by 0.2 M ethanolamine pH 8.0, washed with alternate pH washing buffers (0.1 M acetate, 0.5 M NaCl pH 5; 0.1 M acetate, 0.5 M NaCl pH 5), equilibrated in lysis buffer and incubated with SH-SY5Y or whole mouse brain lysate on a spinning wheel overnight at 4°C. Unspecific binders were removed by washing (20 mM Tris pH 7.6, 300 mM NaCl, 5 mM $MgCl_2$) and bound proteins were eluted by boiling in SDS loading buffer and subsequently analyzed by Western blotting. Antibodies used: APP (B-4), ATXN1 (L-19), c-myc (9E10), DHCR24 (H-300), LRPPRC (H-300), PSEN1 (N-19), RPN1 (E-7), SKP1 (H-6), SIGMAR1 (L-20), TBL1XR1 (L-08).

Mitochondrial activity tests

Mitochondria were isolated from 10^4 - 10^5 HEK293T cells using the Mitochondria Isolation kit (Miltenyi Biotec) according to the manufacturer's instructions. Mitochondrial activity was then assessed by determining the mitochondrial aconitase activity (Aconitase Activity Assay Kit, Merck) according to the manufacturer's instructions. Values were normalized to the mitochondrial protein concentration.

Quantitative RT-PCR

Expression changes at the transcript level were analyzed by qRT-PCR (quantitative reverse transcription polymerase chain reaction). Total RNA from cells was isolated using the miRNeasy kit (Qiagen) according to the manufacturer's protocol. RNA concentration was measured using a NanoDrop system (Thermo Scientific, Wilmington, DE, USA). Reverse transcription was performed using 2.2 μ g total RNA. After DNaseI (Invitrogen) digestion for 15 minutes at 25°C, the sample was split in 2x20 μ l reactions and one of them was reverse transcribed using SuperScript III (Invitrogen) and an oligo-dT18 primer for 1 hour at 50°C. For quantitative PCR, the 2x SYBR green PCR Mastermix (Applied Biosciences, Carlsbad, CA, USA) was used in a 20 μ l reaction with 400 nM of each primer and 0.4 μ l of the RT reaction. The amplification was measured and quantified using the StepOne system from applied biosciences using relative quantification for all primer pairs. Used primers: COX1 (for: CAGCAGTCCTACTTCTCCTATCTCT; rev: GGGTCGAAGAAGGTGGTGT), LRPPRC (for:

GAAGATGCCTTGAACCTTCAAAGA; rev: GCCTACATACTTGCCGGTGT), Vinculin (for: CTCGTCCGGGTTGGAAAAGAG; rev: AGTAAGGGTCTGACTGAAGCAT).

Proximity ligation assay (PLA)

PLA was carried out with HEK293T or SH-SY5Y cells according to the manufacturer's instructions (Olink Bioscience, Uppsala, Sweden). Data acquisition was done on a Leica TCS SP5 with 40x magnification and 3x zoom. Antibodies/siRNA used: APP (B-4), LRPPRC (H-300), NONO (H-85); SFPQ (39-1), siLRPPRC (sc-44734).

Immunohistochemistry

The Swedish AD brain tissue was provided by Uppsala University brain bank. The specimens were obtained at routine autopsies of AD patients and age-matched controls with informed written consent from their families. The study had been approved by the Regional Ethical Committee in Uppsala, Sweden (2005-103; 2005-2006-29 and 2009-089; 2009-04-22), Ethical Committee of Helsinki University and conducted according to the ethical guidelines. One female AD patient from a Swedish family with confirmed APP 670/671 mutation (Mullan et al., 1992) and one aged-matched male control (died from a cardiac failure, two smaller cystic infarcts discovered in thalamus upon autopsy) were included in the study. The Swedish patient died at 61 years after three years of disease duration (Braak stage V-VI), while a healthy control at 63 years of age. The clinical symptoms of Swedish AD included patient memory failure, disorientation and dyspraxia. The brains were routinely fixed in buffered 4% formaldehyde, within twelve hours *post mortem* and widely sampled for embedding in paraffin. For comparative assessment of LRPPRC and APP expression levels in the brain, we utilized fresh frozen frontal and occipital lobe brain cortex samples from the same individuals. For immunohistochemistry on formalin-fixed paraffin embedded brain tissues, the well- characterized monoclonal and polyclonal antibodies against APP, Abeta and LRPPRC were used (see table below). After pre-treatment relevant for each antigen, the brain cortex tissue sections were incubated with the primary antibodies overnight at +4°C, followed by incubation with relevant secondary antibodies and visualization using the avidin-biotin-peroxidase method with 3,3'-diaminobenzidine and H₂O₂ (VECTASTAIN®ABC Reagent, Vector Laboratories, Peterborough, UK). The tissue sections were routinely counterstained with hematoxylin as described (Philipson et al., 2012). Negative controls included omission of primary and secondary antibodies.

Antibody	Epitope/Target	Application/pretreatment	Manufacturer
mAb 4G8	Aβ17-24	IHC (FA), 1:500-1:1,000	Covance, #SIG-39220
pAb AβPP	residues surrounding Thr668 of human AβPP ₆₉₅	IHC (Citrate), 1:200 WB, 1:1,000 in 5% BSA-TBS-T, ON.	Cell Signalling Technology, #2452S
mAb LRPPRC	C-terminus of	IHC (Citrate), 1:200	LRP130 (G-10), Santa Cruz

	LRPPRC (residues 974-1273)	WB, 1:1,000 in 3% BSA-TBS-T, ON	Biotechnology, Inc., #sc-166177
pAb LRPPRC	C-terminus of LRPPRC (residues 974-1273)	IHC (Citrate or Citrate/FA), 1:200 WB, 1:2,000 1:1,000 in 3% BSA-TBS-T, ON	LRP130, Santa Cruz Biotechnology, Inc., #sc-66844
mAb β -actin	cytoplasmic β -actin	WB, 1:1,000 1:1,000 in 3% BSA-TBS-T, ON	AC-15, Santa Cruz Biotechnology, Inc., #sc-69879

Abbreviations: mAb = mouse monoclonal antibody; pAb = rabbit polyclonal antibody; IHC- immunohistochemistry; WB- Western blotting; FA- formic acid (70%, 20 minutes), Citrate- 30 minutes of hydrated autoclaving in 10 mM *citric acid*, 0.05% Tween 20, *pH* 6.0, Citrate/FA- Citric acid followed by formic acid pretreatment; BSA- bovine serum albumin; TBS-T- Tris buffered saline/0.1% Tween 20. ON- overnight.

LC-MS/MS analysis

Protein precipitation and in-solution digest

Eluted proteins from the immunoprecipitation were ethanol precipitated overnight at 4°C and centrifuged the next day with 20,000 g at 4°C for 20 min. The supernatant was carefully removed and the protein pellet was air dried. Precipitated proteins were processed essentially as described (Shevchenko et al., 2006). Lysyl endopeptidase (LysC) (Wako, Osaka, Japan) and sequence-grade modified Trypsin (Promega, Madison, WI, USA) was used for in-solution digestion. Stop and go extraction (STAGE) tips containing C₁₈ empore disks (3M, Minneapolis, MN, USA) were used to purify and store peptide extracts (Rappsilber et al., 2003).

LC-MS/MS settings

For a single measurement, 5 μ l of peptide mixture was injected by an autosampler (LC PAL; CTC Analytics, Zwingen, Switzerland) onto a 15 cm silica microcolumn (inner diameter: 75 μ m) packed in-house with ReproSil-Pur C18-AQ 3- μ m resin (Dr. Maisch GmbH, Ammerbuch, Germany). Peptides were eluted by an Eksigent NanoLC-1D Plus liquid chromatography system (Eksigent, Dublin, CA, USA) on a 8-60% acetonitrile gradient (200 min) with 0.5% acetic acid at a nanoflow rate of 200 nl/min. Eluted peptides were sprayed via an electrospray ion source (Proxeon, Odense, Denmark) into an LTQ-Orbitrap or LTQ-Orbitrap XL hybrid mass spectrometer (Thermo Scientific, Bremen, Germany). Data dependent mode was used for the LTQ-Orbitrap instrument. For one scan cycle, one precursor ion scan was performed in the Orbitrap (*m/z* range = 300-1,700; R = 60,000; target value = 10⁶), followed by five most intense ions selected for fragmentation by collision induced dissociation (target value: 3,000; monoisotopic precursor selection enabled; wideband activation enabled) and MS/MS scans in

the LTQ (linear trap quadrupole) part of the machine. Ions were rejected if their charge state was unassigned or if they were singly charged. The dynamic exclusion duration for precursor ions selected for MS/MS scans was set to 60 s.

Whole proteome experiments were measured on LTQ-Orbitrap Velos systems. For this purpose, peptide mixtures were separated by reversed phase chromatography using the EASY-nLC system (Thermo Scientific) on in-house manufactured 20 cm fritless silica microcolumns with an inner diameter of 75 μm . Columns were packed with ReproSil-Pur C18-AQ 3 μm resin (Dr. Maisch GmbH). Peptides were separated on a 5-60% acetonitrile gradient (124 min) with 0.5% formic acid at a nanoflow rate of 200 nL/min. Eluting peptides were directly ionized by electrospray ionization and transferred into a Q Exactive mass spectrometer (Thermo Scientific). Mass spectrometry was performed in the data dependent positive mode with one full scan (m/z range = 300-1,700; $R = 60,000$; target value = 10^6). The 20 most intense ions with a charge state greater than one were selected for fragmentation by collision induced dissociation (target value = 5,000 monoisotopic; precursor selection enabled; wideband activation enabled) and MS/MS scans in the LTQ. Dynamic exclusion for selected precursor ions was set to 60 s.

Data analysis

MaxQuant settings

The generated raw files from the immunoprecipitation experiments, containing the acquired MS and MS/MS spectra, were processed using the MaxQuant platform version 1.0.13.13 (Cox and Mann, 2008; Cox et al., 2009). SILAC pairs were assembled from detected isotope patterns, recalibrated and quantified in the Quant module (heavy labels: Lys-8; maximum of 3 labeled amino acids per peptide; maximum peptide charge of 6; top 6 MS/MS peaks per 100 Da; polymer detection enabled). The derived peak lists were searched using the MASCOT search engine (version 2.2, MatrixScience, Boston, MA, USA) against an in-house concatenated target-decoy database (Nesvizhskii and Aebersold, 2005) combining forward and reversed protein sequences from the *Homo sapiens* IPI protein database (release 3.64) and 31 common contaminants. LysC specificity, cleaving peptide bonds at the C-terminal side of lysine residues, was required. Carbamidomethylation of cysteine was used as fixed modification; oxidation of methionine and acetylation of the protein N-terminus were set as variable modifications. A maximum of 2 missed cleavages were allowed, and the mass tolerance for fragment ions was set at 0.5 Da. In the Identify module, the MASCOT-generated results were further filtered. For protein identification, a minimum peptide length of 6 amino acids was required. At least one peptide was required to be a unique peptide in a protein group, and a minimum peptide count required for a protein group was set to one. Maximum false discovery rate was set at 1% for both peptide and protein identifications, estimated based on the reverse hits matched in the target-decoy database. Protein quantification was based on unique peptides and non-unique peptides assigned to the protein group with highest number of peptides (razor peptides). At least one peptide SILAC ratio count was required for quantification of a protein group. Reproducibility analysis experiments were based on MaxQuant version 1.5.1.6.

Raw files derived from whole proteome experiments were processed using MaxQuant version 1.2.2.5. Here, MS/MS spectra were searched by the internal Andromeda (Cox et al., 2011b) search engine against the decoy *Homo sapiens* IPI protein database (release 3.84) containing forward and reverse sequences. Additionally the database included 248 common contaminants. MaxQuant analysis included an initial search with a precursor mass tolerance of 20 ppm the results of which were used for mass recalibration (Cox et al., 2011a). In the main Andromeda search precursor mass and fragment mass had an initial mass tolerance of 6 ppm and 20 ppm, respectively. The search included variable modifications of methionine oxidation and N-terminal acetylation, and fixed modification of carbamidomethyl cysteine. Minimal peptide length was set to six amino acids and a maximum of two miscleavages was allowed. The false discovery rate (FDR) was set to 0.01 for peptide and protein identifications. In the case of identified peptides that are all shared between two proteins, these are combined and reported as one protein group.

Protein quantification was based on unique peptides and non-unique peptides assigned to the protein group with highest number of peptides (razor peptides). At least one peptide SILAC ratio count was required for quantification of a protein group. Protein tables were filtered to eliminate identifications from the reverse database and common contaminants. Proteins were considered as interactors when they were enriched at least twofold (i.e. mean \log_2 fold change ≥ 1) and showed inverted ratio in the reverse experiment (i.e. $\log_2FC > 0$ in forward and $\log_2FC < 0$ in reverse experiment) compared to the empty vector controls. Preferential interaction partners of the wt or the disease associated variants had to pass three criteria. First, they had to be identified as specific interaction partners in the experiments with empty vector control (see above). Second, they were required to show a mean \log_2FC of at least 0.5 in the direct comparisons of the wt and disease-associated variant. Third, they were required to show inverted ratios in the label swap experiments ($\log_2FC > 0$ in forward and $\log_2FC < 0$ in reverse experiment or vice versa). As an exception, proteins that showed infinite SILAC ratios (isotope cluster detection in only one SILAC channel) were also considered as interaction partners. Out of 453 identified interactors, 20 proteins showed infinite ratios in particular experiments. For whole proteome experiments ratios were normalized to the median of all SILAC ratios.

Cluster analysis of gene ontology (GO) terms

In order to test whether identified PPIs are enriched for certain Gene Ontology (GO) terms, an enrichment analysis was performed using the online DAVID bioinformatics resource (<http://david.abcc.ncifcrf.gov/>) (Huang da et al., 2009). Calculation of over-represented GO terms was done by comparing the entire list of identified PPIs per bait proteins to the background, i.e. all human proteins. Significant count threshold was set at 1 and the EASE score (modified Fisher's exact test probability) cutoff was set to be 1. Terms with a p-value < 0.01 were selected, log- and z-transformed, hierarchically clustered and plotted as heatmap using in-house Perl and R scripts.

Statistical analysis

Statistical data analysis was done using the R project for Statistical Computing (Version 2.12.1, R Foundation for Statistical Computing, Vienna, Austria) or GraphPad Prism 5.0 (GraphPad Software, San Diego, CA, USA).

Network visualization

Interaction data was loaded to Biolayout 3D Express (version 2.0, open source software, The Roslin Institute, University of Edinburgh, UK) and then edited using standard settings (Theocharidis et al., 2009).

GWAS

Cohort data was acquired from two large genome wide association studies (GWAS) with 753/736 cases/controls (GenADA; (Filippini et al., 2009; Li et al., 2008)) and 3,941/7,848 cases/controls (GERAD; (Harold et al., 2009)). Genomic coordinates of all genes were obtained from BioMart (Ensembl release 54). SNP coordinates (NCBI36 assembly) and association summary statistics of the GenADA study were obtained from dbGAP (accession number phs000219v1). The same data structure for the GERAD study was kindly provided by the Genetic and Environmental Risk for Alzheimer's disease (GERAD1) Consortium (Harold et al., 2009). The GERAD1 sample included 3333 cases and 1225 elderly screened controls genotyped at the Sanger Institute on the Illumina 610-quad chip, referred to collectively hereafter as the 610 group. These samples were recruited by the Medical Research Council (MRC) Genetic Resource for AD (Cardiff University; Kings College London; Cambridge University; Trinity College Dublin), the Alzheimer's Research UK (ARUK) Collaboration (University of Nottingham; University of Manchester; University of Southampton; University of Bristol; Queen's University Belfast; the Oxford Project to Investigate Memory and Ageing (OPTIMA), Oxford University); Washington University, St Louis, United States; MRC PRION Unit, University College London; London and the South East Region AD project (LASER-AD), University College London; Competence Network of Dementia (CND) and Department of Psychiatry, University of Bonn, Germany and the National Institute of Mental Health (NIMH) AD Genetics Initiative. These data were combined with data from 608 AD cases and 853 elderly screened controls ascertained by the Mayo Clinic, Jacksonville, Florida; Mayo Clinic, Rochester, Minnesota; and the Mayo Brain Bank, which were genotyped using the Illumina HumanHap300 BeadChip. These samples were used in a previous GWAS of AD (Carrasquillo et al., 2009). All AD cases met criteria for either probable (NINCDS-ADRD (McKhann et al., 1984), DSM-IV) or definite (CERAD) (Mirra et al., 1991) AD. All elderly controls were screened for dementia using the MMSE or ADAS-cog, were determined to be free from dementia at neuropathological examination or had a Braak score of 2.5 or lower. A total of 5770 population controls were included in GERAD GWAS including the 1958 British Birth Cohort (1958BC) (<http://www.b58cgene.sgul.ac.uk>), NINDS funded neurogenetics collection at Coriell Cell Repositories (Coriell) (see <http://ccr.coriell.org/>), the KORA F4 Study (Wickham et al., 2002), Heinz Nixdorf Recall Study (Schmermund et al., 2002) and ALS Controls. The ALS Controls

were genotyped using the Illumina HumanHap300 BeadChip. All other population controls were genotyped using the Illumina HumanHap550 Beadchip.

For each bait protein SNP sets were defined which comprise all SNPs that are located within a distance of 100 kb of any gene whose protein interacts with the corresponding bait protein. One-sided Wilcoxon-Mann-Whitney test was applied in order to detect differences of the means of the distribution of GWAS P-values within a SNP set compared to the global distribution of P-values (Heinig et al., 2010). Combined P-values were computed using Fisher's method.

The GWAS data from the "Multi-Site Collaborative Study for Genotype-Phenotype Associations in Alzheimer's Disease (GenADA)" were provided by the GlaxoSmithKline, R&D Limited. The datasets used for analyses described in this manuscript were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number [phs000219v1]. Furthermore, this study incorporated summary results from the GERAD1 genome-wide association study. GERAD1 Acknowledgements: Cardiff University was supported by the Wellcome Trust, Medical Research Council (MRC) and/or Alzheimer's Research UK (ARUK) and the Welsh Assembly Government. Cambridge University and Kings College London acknowledge support from the MRC. ARUK supported sample collections at the South West Dementia Bank and the Universities of Nottingham, Manchester and Belfast. The Belfast group acknowledges support from the Alzheimer's Society, Ulster Garden Villages, N.Ireland R&D Office and the Royal College of Physicians/Dunhill Medical Trust. The MRC and Mercer's Institute for Research on Ageing supported the Trinity College group. The South West Dementia Brain Bank acknowledges support from Bristol Research into Alzheimer's and Care of the Elderly. The Charles Wolfson Charitable Trust supported the OPTIMA group. Washington University was funded by NIH grants, Barnes Jewish Foundation and the Charles and Joanne Knight Alzheimer's Research Initiative. Patient recruitment for the MRC Prion Unit/UCL Department of Neurodegenerative Disease collection was supported by the UCLH/UCL Biomedical Centre and NIHR Queen Square Dementia Biomedical Research Unit. LASER-AD was funded by Lundbeck SA. The Bonn group was supported by the German Federal Ministry of Education and Research (BMBF), Competence Network Dementia and Competence Network Degenerative Dementia, and by the Alfried Krupp von Bohlen und Halbach-Stiftung. The GERAD1 Consortium also used samples ascertained by the NIMH AD Genetics Initiative. The KORA F4 studies were financed by Helmholtz Zentrum München; German Research Center for Environmental Health; BMBF; German National Genome Research Network and the Munich Center of Health Sciences. The Heinz Nixdorf Recall cohort was funded by the Heinz Nixdorf Foundation (Dr. jur. G.Schmidt, Chairman) and BMBF. Coriell Cell Repositories is supported by NINDS and the Intramural Research Program of the National Institute on Aging. We acknowledge use of genotype data from the 1958 Birth Cohort collection, funded by the MRC and the Wellcome Trust which was genotyped by the Wellcome Trust Case Control Consortium and the Type-1 Diabetes Genetics Consortium, sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development and Juvenile Diabetes Research Foundation International.

Genetic and Environmental Risk for Alzheimer's disease Consortium (GERAD1) Collaborators

Denise Harold¹, Richard Abraham¹, Paul Hollingworth¹, Rebecca Sims¹, Amy Gerrish¹, Jade Chapman¹, Giancarlo Russo¹, Marian Hamshere¹, Jaspreet Singh Pahwa¹, Valentina Escott-Price¹, Kimberley Dowzell¹, Amy Williams¹, Nicola Jones¹, Charlene Thomas¹, Alexandra Stretton¹, Angharad Morgan¹, Simon Lovestone², John Powell², Petroula Proitsi², Michelle K Lupton², Carol Brayne³, David C. Rubinsztein⁴, Michael Gill⁵, Brian Lawlor⁵, Aoibhinn Lynch⁵, Kevin Morgan⁶, Kristelle Brown⁶, Peter Passmore⁷, David Craig⁷, Bernadette McGuinness⁷, Stephen Todd⁷, Janet Johnston⁷, Clive Holmes⁸, David Mann⁹, A. David Smith¹⁰, Seth Love¹¹, Patrick G. Kehoe¹¹, John Hardy¹², Simon Mead¹³, Nick Fox¹⁴, Martin Rossor¹⁴, John Collinge¹³, Wolfgang Maier¹⁵, Frank Jessen¹⁵, Reiner Heun¹⁵, Britta Schürmann¹⁵, Alfredo Ramirez¹⁵, Tim Becker⁴⁴, Christine Herold⁴⁴, André Lacour⁴⁴, Dmitriy Drichel⁴⁴, Hendrik van den Bussche¹⁶, Isabella Heuser¹⁷, Johannes Kornhuber¹⁸, Jens Wiltfang¹⁹, Martin Dichgans^{20,21}, Lutz Frölich²², Harald Hampel^{23,24}, Michael Hüll²⁵, Dan Rujescu²⁴, Alison Goate²⁶, John S.K. Kauwe²⁷, Carlos Cruchaga²⁶, Petra Nowotny²⁶, John C. Morris²⁶, Kevin Mayo²⁶, Gill Livingston³¹, Nicholas J. Bass³¹, Hugh Gurling³¹, Andrew McQuillin³¹, Rhian Gwilliam³², Panagiotis Deloukas³², Ammar Al-Chalabi³³, Christopher E. Shaw³³, Andrew B. Singleton³⁴, Rita Guerreiro^{34, 43}, Thomas W. Mühlisen^{35,36}, Markus M. Nöthen^{35,36}, Susanne Moebus³⁷, Karl-Heinz Jöckel³⁷, Norman Klopp³⁸, H-Erich Wichmann^{38,39,40}, Minerva M. Carrasquillo⁴¹, V. Shane Pankratz⁴², Steven G. Younkin⁴¹, Peter Holmans¹, Michael O'Donovan¹, Michael J. Owen^{1†}, Julie Williams¹.

Affiliations

1. Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics, Neurosciences and Mental Health Research Institute, Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff, UK.
2. King's College London, Institute of Psychiatry, Department of Neuroscience, De Crespigny Park, Denmark Hill, London
3. Institute of Public Health, University of Cambridge, Cambridge, UK.
4. Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK.
5. Mercer's Institute for Research on Aging, St. James Hospital and Trinity College, Dublin, Ireland.
6. Institute of Genetics, Queen's Medical Centre, University of Nottingham, NG7 2UH, UK.
7. Ageing Group, Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, UK.
8. Division of Clinical Neurosciences, School of Medicine, University of Southampton, Southampton, UK.
9. Clinical Neuroscience Research Group, Greater Manchester Neurosciences Centre, University of Manchester, Salford, UK.
10. Oxford Project to Investigate Memory and Ageing (OPTIMA), University of Oxford, Level 4, John Radcliffe Hospital, Oxford OX3 9DU, UK.
11. University of Bristol Institute of Clinical Neurosciences, School of Clinical Sciences, Frenchay Hospital, Bristol, UK
12. Department of Molecular Neuroscience and Reta Lilla Weston Laboratories, Institute of Neurology, London, UK.
13. MRC Prion Unit, Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK.
14. Dementia Research Centre, Department of Neurodegenerative Diseases, University College London, Institute of Neurology, London, UK.

15. Department of Psychiatry, University of Bonn, Sigmund-Freud-Straße 25, 53105 Bonn, Germany.
16. Institute of Primary Medical Care, University Medical Center Hamburg-Eppendorf, Germany.
17. Department of Psychiatry, Charité Berlin, Germany.
18. Department of Psychiatry, Friedrich-Alexander-University Erlangen-Nürnberg, Germany.
19. LVR-Hospital Essen, Department of Psychiatry and Psychotherapy, University Duisburg-Essen, Germany.
20. Institute for Stroke and Dementia Reserach, Klinikum der Universität München, Marchioninstr. 15, 81377, Munich, Germany.
21. Department of Neurology, Klinikum der Universität München, Marchioninstr. 15, 81377, Munich, Germany.
22. Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany.
23. Discipline of Psychiatry, School of Medicine and Trinity College Institute of Neuroscience, Laboratory of Neuroimaging & Biomarker Research, Trinity College, University of Dublin, Ireland.
24. Institute for Memory and Alzheimer's Disease & INSERM, Sorbonne Universites, Pierre and Marie Curie University, Paris ; Institute for Brain and Spinal Cord Disorders (ICM), Department of Neurology, Hospital of Pitié-Salpêtrière.
25. Centre for Geriatric Medicine and Section of Gerontopsychiatry and Neuropsychology, University of Freiburg, Germany.
26. Departments of Psychiatry, Neurology and Genetics, Washington University School of Medicine, St Louis, MO 63110, US.
27. Department of Biology, Brigham Young University, Provo, UT, 84602, USA.
28. Neurodegenerative Brain Diseases group, Department of Molecular Genetics, VIB, Antwerpen, Belgium.
29. Institute Born-Bunge and University of Antwerp; Antwerpen, Belgium.
30. Memory Clinic and Department of Neurology, ZNA Middelheim, Antwerpen, Belgium.
31. Department of Mental Health Sciences, University College London, UK.
32. The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.
33. MRC Centre for Neurodegeneration Research, Department of Clinical Neuroscience, King's College London, Institute of Psychiatry, London, SE5 8AF, UK.
34. Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, 20892, USA.
35. Department of Genomics, Life & Brain Center, University of Bonn, Sigmund-Freud-Str. 25, D-53127 Bonn, Germany.
36. Institute of Human Genetics, University of Bonn, Wilhelmstr. 31, D-53111 Bonn, Germany.
37. Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, University Duisburg-Essen, Hufelandstr. 55, D-45147 Essen, Germany.
38. Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany.
39. Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.
40. Klinikum Grosshadern, Munich, Germany.
41. Department of Neuroscience, Mayo Clinic College of Medicine, Jacksonville, Florida 32224, USA.
42. Division of Biomedical Statistics and Informatics, Mayo Clinic and Mayo Foundation, Rochester, Minnesota 55905, USA.
43. Department of Molecular Neuroscience, Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK

44. Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Bonn
Supplemental References

- Carrasquillo, M.M., Zou, F., Pankratz, V.S., Wilcox, S.L., Ma, L., Walker, L.P., Younkin, S.G., Younkin, C.S., Younkin, L.H., Bisceglia, G.D., *et al.* (2009). Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nature genetics* **41**, 192-198.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26**, 1367-1372.
- Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V., and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nature protocols* **4**, 698-705.
- Cox, J., Michalski, A., and Mann, M. (2011a). Software lock mass by two-dimensional minimization of peptide mass errors. *Journal of the American Society for Mass Spectrometry* **22**, 1373-1380.
- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011b). Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* **10**, 1794-1805.
- Filippini, N., Rao, A., Wetten, S., Gibson, R.A., Borrie, M., Guzman, D., Kertesz, A., Loy-English, I., Williams, J., Nichols, T., *et al.* (2009). Anatomically-distinct genetic associations of APOE epsilon4 allele load with regional cortical atrophy in Alzheimer's disease. *NeuroImage* **44**, 724-728.
- Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M.L., Pahwa, J.S., Moskva, V., Dowzell, K., Williams, A., *et al.* (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nature genetics* **41**, 1088-1093.
- Heinig, M., Petretto, E., Wallace, C., Bottolo, L., Rotival, M., Lu, H., Li, Y., Sarwar, R., Langley, S.R., Bauerfeind, A., *et al.* (2010). A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* **467**, 460-464.
- Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57.
- Li, H., Wetten, S., Li, L., St Jean, P.L., Upmanyu, R., Surh, L., Hosford, D., Barnes, M.R., Briley, J.D., Borrie, M., *et al.* (2008). Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. *Archives of neurology* **65**, 45-53.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939-944.
- Mirra, S.S., Heyman, A., McKeel, D., Sumi, S.M., Crain, B.J., Brownlee, L.M., Vogel, F.S., Hughes, J.P., van Belle, G., and Berg, L. (1991). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* **41**, 479-486.
- Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B., and Lannfelt, L. (1992). A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nature genetics* **1**, 345-347.
- Nesvizhskii, A.I., and Aebersold, R. (2005). Interpretation of shotgun proteomic data: the protein inference problem. *Molecular & cellular proteomics : MCP* **4**, 1419-1440.
- Philipson, O., Lord, A., Lalowski, M., Soliymani, R., Baumann, M., Thyberg, J., Bogdanovic, N., Olofsson, T., Tjernberg, L.O., Ingelsson, M., *et al.* (2012). The Arctic amyloid-beta precursor protein (AbetaPP) mutation results in distinct plaques and accumulation of N- and C-truncated Abeta. *Neurobiology of aging* **33**, 1010 e1011-1013.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical chemistry* 75, 663-670.

Schmermund, A., Mohlenkamp, S., Stang, A., Gronemeyer, D., Seibel, R., Hirche, H., Mann, K., Siffert, W., Lauterbach, K., Siegrist, J., *et al.* (2002). Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. *American heart journal* 144, 212-218.

Theocharidis, A., van Dongen, S., Enright, A.J., and Freeman, T.C. (2009). Network visualization and analysis of gene expression data using BioLayout Express(3D). *Nature protocols* 4, 1535-1550.

Wickham, H., Walsh, C., Asherson, P., Gill, M., Owen, M.J., McGuffin, P., Murray, R., and Sham, P. (2002). Familiality of clinical characteristics in schizophrenia. *Journal of psychiatric research* 36, 325-329.