medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Effect of prioritized *SORL1* missense variants supports clinical consideration for familial Alzheimer's Disease

Henne Holstege[1,2,3,4*], Matthijs W. J. de Waal[1,4*], Niccolo Tesi[1,2,4], Sven J. van der Lee[1,3,4], ADES-consortium, ADSP consortium, StEP-AD consortium, Knight-ADRC, UCSF/NYGC/UAB, Maartje Vogel[5], Resie van Spaendonk[5], Marc Hulsman[1,2,4], Olav M. Andersen[6*]

1. Section Genomics of Neurodegenerative Diseases and Aging, Department of Clinical Genetics, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands
2. Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands
3. Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands.
4. Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands
5. Dept. of Human Genetics, Genome Diagnostics Laboratory, Amsterdam UMC, Vrije Universiteit Amsterdam, the Netherlands
6. Dept. of Biomedicine, Aarhus University, Denmark

*Authors contributed equally to this work.

Corresponding Author:

Henne Holstege

h.holstege@amsterdam.nl

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background: Rare variants in the *SORL1* gene have been associated with increased risk of Alzheimer's disease (AD). While protein-truncating variants (PTVs) are observed almost exclusively in AD patients, most variants are rare missense variants that can be benign, risk-increasing, and recent reports have indicated that some variants are causative for disease. However, since *SORL1* is currently not considered an autosomal dominant Alzheimer Disease gene (ADAD), segregation analyses are not performed, which complicates the identification of additional clinically important missense variants.

Methods: We prioritized highly conserved and functionally relevant *SORL1* missense variants by considering the functional effects of homologous variants on proteins that share domains with *SORL1* (domain-mapping of disease mutations, DMDM) into. We used this variant prioritization approach to annotate *SORL1* variants identified in a previously assembled exome sequencing dataset encompassing 18,959 AD cases and 21,893 non-demented controls, and we tested the effect of high, moderate, low and no priority missense variants and specific variant subtypes on disease risk and age at onset.

Results: High priority missense variants (HPV) associated with a 6.4-fold increased risk of AD (95%CI: 4.3 – 9.7, $p=2.1x10^{-24}$), which concentrated on early onset AD (OR_{EOAD} 10.5, 95%CI: 6.8 - 16.3, $p=3.0x10^{-29}$) vs. late onset AD (OR_{LOAD}=4.5, 95%CI 2.85 - 6.94; $p=4.9x10^{-11}$). The median age at onset of HPV carriers was >8-years earlier than carriers of wild-type *SORL1*. Intriguingly, specific subtypes of HPVs, including those affecting residues in the YWTD-motif or the calcium cage, occurred only in AD cases and carriers of these variants had an *earlier* age at onset compared to carriers of PTVs, indicative of a dominant negative effect. Carriers of other HPVs had an age at onset that overlapped with carriers of PTVs, suggesting they lead to haploinsufficiency. Yet other variants had a slightly later age at onset than PTVs, suggesting that their effect on *SORL1* function was milder than losing a copy. Variants annotated as moderate, low and no priority did not have an effect on AD.

Conclusions: Next to carriers of *SORL1* PTVs, carriers of selected missense variants should be considered for segregation analyses, which will likely provide evidence for autosomal dominant inheritance for additional *SORL1* missense variants.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

The SORL1 protein, also known as SORLA, encoded by the *SORL1* gene, share large homologous regions with members of the low-density lipoprotein receptors family (LDLRs) which are part of larger trafficking complexes and integrated in the membranes of the cell, endocytosed vesicles and the tubes that connect them. LDLR is the cargo-binding entity of the LDLR-clathrin complex, which regulates the endocytosis of cargo. SORL1 is the cargo-binding entity of the SORL1-retromer complex which regulates cargo-transport from the endosome back to the trans-Golgi network ('retrograde' pathway) and the transport of endocytosed receptors back to the cell surface ('recycling' pathway). SORL1 cargo includes both Amyloid-β and the amyloid precursor protein (APP): APP binding by retromer-SORL1 prevents APP-trafficking to the early endosome, thereby warding off APP cleavage and the subsequent the formation and secretion of Amyloid-β (Small et al., 2015). Furthermore, retromer-SORL1 binds endocytosed GLUA1 at the endosome and recycles it back to dendritic spines, thereby supporting healthy glutamate signaling (Mishra et al., 2022).

In addition to the involvement of *SORL1* in hallmark processes of Alzheimer's disease (AD), evidence for the association of genetic variants in *SORL1* with the AD risk has accumulated since 2007 (Bettens et al., 2008; Pottier et al., 2012; Rogaeva et al., 2007; Vardarajan et al., 2015). SORL1 is a large, 2214 amino-acid protein, encoded by 48 exons of the *SORL1* gene. By virtue of its size, *SORL1* genetic sequence is vulnerable for acquiring mutations. To date, >500 coding variants have been reported across diverse populations, which range from having no to little effect to having deleterious effects on protein function. In total, potentially damaging *SORL1* variants affect as many as 2.75% of all unrelated early onset AD patients (EOAD, with age at onset <65 years) and 1.5% of unrelated late onset AD cases (LOAD, with age at onset >65 years) (Holstege et al., 2022). Of these, protein truncating variants (PTV) in the *SORL1* gene occur almost exclusively in AD cases (Holstege et al., 2017, 2022), and several clinical genetics labs now consider *SORL1* PTVs as clinically important. However, most *SORL1* variants observed in AD patients are *rare missense SORL1* variants, some of which are risk-increasing or possibly causative for disease, and many are benign (Campion et al., 2019; Thonberg et al., 2017; Van der Lee et al., 2018). We and others recently identified several *SORL1* missense variants, in which AD segregated with a *SORL1* variant with an autosomal dominant inheritance pattern. We identified the p.Y1816C 'Tyrosine Corner' variant in three unrelated pedigrees: the p.Y1816C substitution affects a strongly conserved tyrosine residue in the third SORL1 3Fn-domain, and this mutation leads to impaired SORL1 dimerization and retromer binding, and ultimately to autosomal dominant AD (ADAD) (A. M. G. Jensen et al., 2023). Second, is the p.D1545V 'Icelandic' missense

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

variant, which was identified in a large informative Icelandic family (Bjarnadottir, 2023) and which leads to ADAD as a result of protein misfolding and ER retention (Blacklow et al., 1996). Recently, the p.R953C 'Seattle variant' was reported to segregate with AD, and severe AD pathology was observed in affected family members (Fazeli et al., 2023). These reports provide the first indications that *SORL1* might be considered a fourth ADAD gene. However, *SORL1* is currently considered a risk gene, not a designated ADAD gene, such that variant carriers are often not clinically identified, and segregation analysis are not performed. Together with the variable age at onset of affected carriers, this complicates the identification of clinically important missense variants.

Here, we took advantage of the increasing prior knowledge of SORL1 function (see Box), and we learned from the effect of specific missense variants on the function of proteins that share domains with SORL1. This domain-mapping of disease mutations (DMDM) led to the prioritization of variants that affect highly conserved and functionally relevant *SORL1* variants. In the current manuscript, we applied this classification approach on *SORL1* variants identified in sequencing data of 18,959 AD cases and 21,893 non-demented controls (Holstege et al., 2022). We then associated the different missense mutation classes with risk and age at onset of AD.

BOX FIGURE WITH TEXT:

SORL1 is a large, 2214 amino acid protein with 11 overall domains, most of which consist of multiple repeated domain elements, each of which includes many strictly or moderately conserved residues important for protein domain folding and/or the binding of ligands (**Fig 1**). When the SORLA protein is transcribed as the ribosome, he protein starts with signal peptide (res 1-28) that is cleaved off upon translocation to the endoplasmatic reticulum. Then comes a pro-domain (res 29-81) that is speculated to prevent binding of certain ligands to the VPS10p-domain in the endoplasmatic reticulum (ER), where receptor and ligand are co-expressed. The pro-domain is cleaved off by Furin once SORL1 leaves the trans-Golgi-Network, where it can engage in ligand binding and trafficking. After the pro-domain comes the VPS10p domain (res 82-617), a ten-bladed β-propeller domain, a flat disc that is stabilized at its bottom face by the 10CC-domain (res 618-753), VPS10p binds ligand at its top face. The VPS10p domain has a large hydrophobic tunnel at its center, allowing interaction with small lipophilic ligands such as the Amyloid-β peptide. The domain contains two protrusions (loop structures, loop L1 and loop L2): the VPS10p domain can bind ligand at neutral pH and while L1 blocks part of the tunnel, the L2 protrusion pushes the ligand against the tunnel wall. After trafficking to a more acidic part of the cell (i.e. the lysosome), L1 and L2 change conformation and release the ligands from the VPS10p pore. C-terminal to the 10CC domain is a ligand-binding YWTD β-propeller (res 754-1013), which is stabilized at its bottom face by an EGF-domain (res 1014-

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

1074, fully encoded by exon 22), such that ligand-interactions with both the VPS10p and YWTD β-propellers occur at their top faces. The combined action of VPS10p β-propeller and the YWTD β-propeller might enable interactions with large ligands including co-receptors in multimeric complexes, or large soluble ligands requiring two adjacent β-propellers for efficient binding, akin to what was recently identified for LRP4/Agrin/MusK signalling complex (Xie et al., 2023). C-terminal to the EGF domain comes the CR-cluster (res 1075-1550) which is the interacting site of at least half of the SORL1-ligands, including APP. This cluster is like a flexible necklace composed of 11 unique ~40 amino-acid CR domains, each encoded by a single exon (exons 23-33), that each form the 'pearls' on the string. These can wrap around larger ligands and engage in minimal motif interactions with multiple sites of a ligand, leading to high-affinity ligand binding. Each CR domain includes 16 strictly conserved amino acids, including six disulfide bridge-forming cysteines, such that all CR domains have a similar compact folding. Each CR domain further contains four conserved residues that form an octahedral 'calcium cage' which stabilizes the domain, and in combination with two backbone carbonyls, coordinates a calcium ion, which is critical for calcium-dependent domain folding. The side chains of these two residues engage in minimal-motif ligand binding, which explains why ligand binding to CR-domains relies on Ca2+. Substituting these may impair the binding of specific ligands, but do not affect overall folding and stability of CR-domains. Perturbation of the calcium cage on the other hand, which occurs in carriers of the autosomal dominant p.D1545V 'Icelandic' mutation, leads to a misfolded SORL1 protein that is retained in the ER (Bjarnadottir, 2023). C-terminal of the CR-cluster is the 3Fn-cassette (res 1551-2121), containing 6 ellipsoid 3Fn domains, each containing several conserved and partly conserved residues, and involved in SORL1 dimerization (A. M. Jensen et al., 2023). Therefore, genetic variants affecting one of the conserved residues in 3Fn domain is likely to disturb SORL1 dimerization, such observed with the p.Y1816C 'tyrosine corner' mutation (A. M. G. Jensen et al., 2023). Lastly, SORL1 has a transmembrane and cytoplasmic tail domain (res 2122-2214) which can interact with the VPS26 subunit of the retromer complex (Fjorback et al., 2012). Recent evidence suggests that SORL1 matures (by *N*- and *O*-glycosylation) at the ER/Golgi in a monomer form, then travels to the endosome where it dimerizes at its 3Fn domain and its VPS10 domain (A. M. Jensen et al., 2023). The dimerized SORL1 uses its cytoplasmic tail domain to interact with the VPS26 subunit of the retromer complex, allowing SORL1 to engage in retromer-dependent cargo trafficking through the endolysosomal system.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

Samples: We extracted *SORL1* genetic variants from the assembled whole exome sequencing (WES) and whole genome sequencing (WGS) data as previously described (Holstege et al., 2022), which includes data contributed by the European ADES cohort and the ADSP, StEP-AD, Knight-ADRC and UCSF/NYGC/UAB cohorts, Procedures for AD diagnosis were similar across cohorts, and occurred according to the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association criteria (McKhann et al., 1984) or the National Institute on Aging-Alzheimer's Association criteria (Mckhann et al., 2011). Carriers of a pathogenic variant(s) in *PSEN1*, *PSEN2*, or *APP* or in any other gene associated with Mendelian dementia were excluded. Relatives up to 3rd degree of relatedness were excluded to avoid any family-based effects.

Application of our comprehensive quality control procedures for a *SORL1*-specific analysis allowed the retention of more AD cases and controls, compared to the previously published genome-wide analysis (Holstege et al., 2022). We further maximized analysis power by including *SORL1* variants identified in individuals with non-European ancestry, the rationale for this is that rare *SORL1* variants have been reported to associated with AD risk across all populations studied thus far (Miyashita et al., 2013), such that it is unlikely that population-specificity will influence association statistics. Together, the current sample included *SORL1* sequences from 40,852 individuals: 18,959 AD cases and 21,893 non-demented controls, comprising 13.62% African, 0.07% East Asian, 0.64% South Asian, 6.16% Admixed Americans and 79.5% European (**Table 1**), for a PCA representing the sample by population background see **Figure S1**.

Variant Quality Control: The raw sequence data was processed with a uniform pipeline as described previously (Holstege et al., 2022). In brief, the data was processed relative to the GRCh37 reference genome, after which extensive quality control was applied which led to the exclusion of likely false positive variant-calls from analysis. Other variants were excluded due to differential missingness, positions for which coverage across cases and controls differed >5%. These included all variants in exon 1 (res 1-95), which codes for the signal peptide (res 1-28), the pro-domain (29-81), and the first 10 residues of the VPS10p domain. See **Table S1** for excluded variants.

Variant annotation: We annotated *SORL1* variants that occur in the canonical transcript (T260197). All variants were annotated with the 'non-neuro popmax' MAF using the GnomAD database version v.2.1.1. Variants that were absent from GnomAD database were annotated by their MAF in the total current sample. Variants with

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

MAF>0.05% (which relates to having at least 21 carriers in this sample) were considered non-rare. Variants with a MAF<0.05% were considered rare, and included in a domain-specific rare variant burden analysis (**Fig 2**).

We used the Variant Effect Predictor in Ensembl database (VEP, version v.94.542) to identify variants with a possible consequence on protein function. Missense variants were annotated with the REVEL score (Ioannidis et al., 2016), which ranges from 0 (no predicted effect on protein function), to 1 (high predicted effect). Variants comprising two consecutive missense variants that give rise to two consecutive amino acid substitutions could not be annotated by REVEL, and were conservatively annotated according to the substitution with the lowest REVEL score. Protein truncating variants (PTV) were identified using LOFTEE (version v.1.0.2) (Karczewski et al., 2020), which annotates nonsense, frameshift and splice variants that lead to protein truncation as 1, and non-PTVs as 0. Note that PTVs in the last exon 48 should be considered deleterious, as this encodes the cytoplasmic tail domain which includes the FANSHY motif (for retromer binding), DDLGEDDED motif (sequence for binding cytoplasmic AP1 and AP2) and the DVPMV motif (sequence for GGA1 and GGA2 binding) which are all necessary for cellular trafficking and activity. Since LOFTEE did not annotate exon 48 PTVs, we manually included them in the PTV list. Splice variants with LOFTEE score 0 were evaluated using Splice AI (Jaganathan et al., 2019) and those with a potential splice effect were evaluated manually by a trained clinical geneticist (MV), and those with expected effects on splicing were added to the list of PTVs.

Prioritization of rare missense variants

Rare missense variants (MAF <0.05%) were separated into high priority variants (HPVs), moderate priority variants (MPVs) low priority variants (LPVs) and no priority variants (NPVs), according to the scheme presented in **Table 2**, identified variants are listed in **Table S2**. HPVs were identified as according the DMDM analysis described previously (Andersen et al., 2023), independent of REVEL score, and are indicated in **Fig 3** and listed in **Table S4**. An exception is the p.VPS10p domain and 10CC domain combination, as the 5 proteins of the VPS10p-receptor family hold no/only few disease associated variants. Therefore, DMDM analysis was not possible for the variants in the VPS10p domain, such that, apart from the variants involving cysteines in VPS10p loops L1 and L2, we relied on applying a REVEL score threshold of >0.5, for which we previously found the strongest effect on AD risk (Holstege et al., 2022). MPVs are rare missense variants annotated as moderate priority by DMDM as indicated in **Fig 3** and listed in **Table S5**). LPVs are rare missense variants with a REVEL score >0.5 that are not in the VPS10p domain, and NPVs comprise all remaining rare variants that are not HPV, MPV or LPV.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Rare variant association with Alzheimer's Disease: Effects for non-rare variants (MAF>0.05%) on AD were evaluated per variant using a logistic regression model (**Table S3)**, while correcting for population effects using PCA components (PC1-PC6) (**Fig S1**). Variants with a MAF<0.05% were considered rare, and included in a domain-specific rare variant burden analysis (**Fig 2**). We associated carrying a variant with MAF <0.05% appertaining to a specific priority category with AD (burden test). We repeated analyses stratifying EOAD patients (AD-aao <65 years) and LOAD patients (AD-aao >65 years), relative to the same group of controls. Since the number of specific variants number of carriers was low, the significance of the association was determined using a Fisher exact test, and corrected for multiple testing (Bonferroni), p_{adj}<0.05 was considered significant. Calculations were performed using the epiR package (v.2.0.38).

Age at onset curves: Since controls were overall much younger than cases, we compared the effect of different *SORL1* variants on age at onset variants in a case-only age at onset analysis. We used Kaplan-Meier survival analysis (CI of 95%) to estimate age at onset curves (in R using Survival (v.3.3-1). For each variant priority category, we compared the age at onset with the age at onset of *SORL1*-WT carriers in our cohort. Log rank tests were performed to test for differences between age at onset curves. Additionally, we stratified according to *APOE* genotype.

Effect of APOE: We investigated a possible interaction-effect between *APOE* genotype (no, one or two APOE-E4 alleles) and *SORL1* priority category. To avoid confounding an interaction signal by samples in which APOE status was part of selection criteria, we performed an interaction analysis on ADES dataset only, for which there was no selection for *APOE*-genotype. We tested for both additive effect (*SORL1 + APOE*) and interactive effect (*SORL1 * APOE*) using a logistic regression models adjusted for PCA components (PC1-PC6). Interaction effects were tested using a Likelihood Ratio Test.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Results

We included 40,852 individuals in this sample: 18,959 cases (mean age of 72.4 ± 10.7, 59.49% females 50.4% APOE ε4 carriers) and 21,893 controls (mean age of 71.1 ± 16.7, 58.23% females, 17.2% APOE ε4 carriers) (**Table 1**). After quality control we observed 646 unique coding *SORL1* variants across these individuals. Of these, 52 were non-rare variants with a MAF>0.05% (**Table S3**), 74 PTVs and 520 rare missense variants were further stratified into 269 NPVs, 73 LPVs, 67 MPVs, and 111 HPVs (variants listed in **Table S2**), as indicated in **Fig 2**.

PTV, Protein Truncating Variants

The 74 rare PTV variants included nonsense, frameshift and splice variants, which were observed in 89 cases and 6 controls (aged 49-, 53-, 64-, 75-, 80-, 85-year-old controls at last screening), associated with an overall 17.2-fold increased risk of AD (95%CI 7.5 – 39.3; p =1.2x10^{-21}). Specifically, PTVs associated with a 35.3-fold increased risk of EOAD (95%CI 15.24 – 81.8; p =2.2x10^{-31}) and an 8.6-fold increased risk of LOAD (95%CI 3.6 – 20.6; p=3.6x10^{-7}) (**Table 3**). A survival analysis indicated that the median age at onset of a *SORL1*-PTV carrier was 10-years earlier AD than the age at onset of *SORL1*-WT carriers (95%CI -12 – -8; 2.66x10^{-11}) (**Fig 4A, Table 3**).

HPV, High priority missense variants

We identified 111 missense HPVs carried by 149 AD cases and 27 controls, and these associated with an overall 6.4-fold increased risk of AD (95%CI: 4.3 – 9.7, p=2.1x10^{-24}). Specifically, HPVs associated with a 10.5-fold increased risk of EOAD (95%CI: 6.8 - 16.3, p=3.0x10^{-29}), and to a 4.5-fold increased risk of LOAD (95%CI 2.85 - 6.94; p=4.9x10^{-11}), indicating that EOAD patients are enriched with these variants (**Table 3**). Of the 111 missense HPVs, 82 (74%) were singletons (72/82 in AD cases), 18 variants were carried by two individuals (30/36 were AD cases) 11 variants were carried by three or more individuals, and one variant (Y391C) was carried by 12 individuals in the sample (all AD cases). A survival analysis indicated that carrying a missense HPV expedites AD-aao by on average 8.2 years (95%CI 10 – 6; 1.01x10^{-8}), relative to *SORL1* WT carriers (**Figure 4A, Table 3**).

VPS10p domain (res 82-617) and 10CC domain (res 618-753): The 27 missense HPVs in the VPS10p domain associate with overall 8.8-fold increased risk of AD (95%CI 3.5- 22.34; p=3.2 x10^{-07}), and specifically with a 15.7-fold increased risk of EOAD (95%CI 5.94- 41.48; p=1.79 x10^{-09}) and with a 5.48-fold increased risk of LOAD (95%CI: 2.01 – 14.95; p=0.0077) (**Table 3**). We identified 6 missense HPVs in the p.VPS10p domain that involve cysteines in one of the two loops L1 and L2. Of all receptors with a VPS10p domain, only SORL1 has these two loops, which are involved in ligand-binding (Kitago et al., 2015). Intriguingly, 12 unrelated AD cases and no

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

controls gained a cysteine in L1 (Y391C): carriers came from diverse cohorts and countries, with an median age at onset of 67.5 years, which was an average 5.5 year later compared to PTVs carriers, but 4.5 years earlier than *SORL1* WT carriers, (**Fig 4B, Table S6**). Two AD cases lost a cysteine in L2 (C467Y, C473S) and two AD cases gained a cysteine in L2 (R480C). Last, two controls had gained a cysteine in L1 (G398C) or L2 (S474C). We further identified 16 missense HPVs in the p.VPS10p domain with a REVEL score ≥0.50, carried by 15 cases and 3 controls, which in aggregate had an average age at onset 2.5 (95%CI: 4 - 13) year later than PTVs carriers and 12.5 years earlier than *SORL1* WT carriers (**Fig 4B, Table 3**). Of particular interest are the four variants that affect the Asp-box, that stabilizes the β-propeller by forming interactions between propeller blades, with the L1/L2 loops and with the nearby 10CC-domains (Andersen et al., 2023): three cases with S564G, T570I or S138F, and two cases with D236G. The 10CC domain, C-terminal to the p.VPS10p domain, stabilizes VPS10p β-propeller, and losing one of the 10 highly conserved cysteines or gaining a cysteine will likely impair domain folding: we observed one AD case who carried C716W, and two AD cases who carried Y722C. In aggregate, we observed 19 variants affecting the 10CC domain, with a median age at onset of 67 years, 5 years earlier than *SORL1* WT carriers (**Fig 4B, Table 3**).

YWTD domain (res 754-1013) and EGF domain (res 1014-1074): We identified 6 variants that affect the highly conserved YWTD-motif, which maintains the structural and functional integrity of the β-propeller. These were carried by 8 non-related AD cases, with an average age at onset 8 years earlier than *SORL1* WT carriers (**Fig 4B, Table 3)**. We further identified 5 AD cases and one control with a R953H, in which a positively charged arginine at domain position 38 in the 5th YWTD domain: DMDM analysis revealed that losing the positively charged Arginine at this position is deleterious (Andersen et al., 2023). Ages at onset ranged from 46 to 78. We did not identify variants at this position in other YWTD subdomains. The EGF domain, C-terminal to the YWTD domain, includes eight cysteines that most likely form four intradomain disulphide bridges to stabilize the whole EGF:YWTD β-propeller unit, such that variants involving a cysteine were prioritized as HPV. We identified one AD patient who carried a variant leading to Y1064C substitution and one control who carried a variant leading to a C1026R substitution. Taken together, carrying a variant affecting the YWTD domain leads to a 4.6-fold increased risk of AD (**Table 3)**.

CR domain (res 1075-1550): calcium cage and cysteines: DMDM analysis provides strong evidence that variants disrupting the calcium cage in the CR domain leads to a dysfunctional SORL1 protein. We identified 12 unique variants that lead to substitutions of the calcium cage residues which were observed only in AD cases (n=13), the average age at onset of AD patients who carry a variant affecting a calcium cage was 12 years earlier

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

than SORL1 WT carriers, and an average 2 years *earlier* compared to PTVs carriers. (**Fig 4B, Table 3)**. One calcium cage variant, D1108N, was observed in three unrelated AD cases. DMDM analysis further provides strong evidence that an odd number of cysteines (ONC), i.e. a disruption of the conserved pattern of 6 cysteines results in dysfunctional CR domain folding. We identified 24 unique cysteine-affecting variants (15 cysteine-gained and 9 cysteine-lost variants) carried by 36 AD cases and 8 controls such that carrying such an HPV, associates with a 5.2-fold increased risk of AD (95%CI: 2.2-11.2; $p=5.9x10^{-5}$) (**Table 3**). Most variants were caried by one or two individuals, but five unrelated cases and two controls carried R1490C, four unrelated cases carried R1080C variant, and three unrelated cases and one control carried R1124C. The average age at onset of these AD patients with an ONC variant was 68 which was 4 years earlier than *SORL1* WT carriers, but on average a 6-year *later* than PTV carriers (**Fig 4B, Table 3**).

3Fn domain (res 1551-2121) The 3Fn domain is necessary for SORL1-dimerization, complex formation with retromer, and possibly cargo binding at the endosome (A. M. Jensen et al., 2023). We recently established that a substitution at position 83 (the p.Y1816C Tyrosine corner mutation) critically disturbs the domain-stabilizing "Tyrosine corner" interaction with a Leucine at domain position 77, and a proline at position 79, which leads to ADAD in assembled pedigrees (A. M. G. Jensen et al., 2023). In this sample, we observed the p.Y1816C variant in 6 AD unrelated AD cases with an average age at onset of 60.2 years . One AD case carried a variant affecting the proline at position 79 (p.P1619Q) with age at onset 59 years. And one 45-year-old control carried a variant affecting the leucine at position 77 (p. L1617V); given the importance of the tyrosine corner in SORLA function, it is not unlikely that this individual will develop AD at a later age.

One AD case carried a variant affecting the glycine at position 96 (p.G1732A) for which potential consequence on domain folding or cargo interactions remains to be established. This HPV was previously reported to segregate with AD in a Swedish family (Thonberg et al., 2017), supporting variant damagingness (Andersen et al., 2023). One AD patient, with age-at-onset 46, carried a variant affecting the glycine at position 36 (p.G1681D), which may affect domain stability or ligand binding. Lastly, four cases and one control affect the residues that contribute to the hydrophobic core, i.e. the tryptophan at position 25 and the tyrosine at position 41 which acts as the 'glue', that holds the two β-sheets of the 3Fn-domain sandwich fold together. Furthermore, substitution of the moderately prioritized prolines at positions 6, and 7 that occur in some 3Fn domains were observed in only AD cases.

Transmembrane and tail domain (res 2161-2214): There were no variants prioritized in these domains

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

MPV, Moderate Priority Variants

We identified 67 unique MPVs, carried by 83 AD cases and 61 controls. In aggregate, the risk of carrying an MPV concentrates on the LOAD patients, in which we find the strongest effect (OR=1.7, 95%CI 1.2 - 2.4, $p=7.5x10^{-2}$) (**Table 3**). Our age at onset analysis supports this as compared to *SORL1 WT* carriers, carrying a *SORL1* MPV does not lead to an expedited age at the EOAD stage, but did expedite age at onset at the LOAD stage (**Fig 4A, Table 3**). Note that HPVs with REVEL>50 associated with a 2.23-fold increased risk of (95%CI 1.5-3.4), $p=1.3x10^{-4}$, indicating that selection of functionally relevant MPVs may more closely capture the risk they hold. We identified MPVs due to their position in a functional domain, but for which associated damagingness is currently unclear and requires more evidence (Andersen et al., 2023). We identified 6 MPVs that affected residues that contribute to the hydrophobic core of the YWTD domain (res 754-1013) at domain positions 6, 8, 15 and 42, carried by 6 cases and three controls. Furthermore, 2 AD cases and 3 controls who carried an MPV leading to N924S substitution, that affects domain position 4 (part of the SBiN-motif) in the 5th blade of YWTD. In the CR-cluster (aa 1075-1550), we identified ten unique variants affecting the partly conserved glycine at position 38 which occurs in eight of the eleven CR-domains in SORL1, which were carried by 26 cases and 15 controls that affected in CR domains 5, 6, 9, 10, or 11, and in aggregate associated with a two-fold increased risk (OR = 2.0; 95%CI 1.06 – 3.78; p = 0.040). Of these, variant G1536S, was carried by 10 cases and 5 controls, contributing substantially to the signal. We further identified one AD patient who carried S1148R, which affects a domain-stabilizing serine at domain position 46, as part of an 'Asx-turn'. Other MPVs in the CR domain were the fingerprint at residue 39, which we observed in 5 AD cases and 3 controls. We identified 14 MPVs in the 3Fn domains (aa 1551-2121): which were carried by 8 cases and 10 controls. Lastly, in the tail domain (aa 2161-2214), we observed 4 variants affecting the conserved FANSHY motif, which is essential for SORL1 binding to the retromer core complex: three cases carried respectively A2173T, S2175R (11:121498424:C>A), and H2176R. A fourth variant, S2175R (11:121498424:C>G) was carried by 10 cases and 11 controls. Lastly, two AD cases carried the D2207G variant in the GGA-binding motif of the SORL1 tail, involved in binding the adapter protein AP1. Overall, more research is necessary to understand the effect of specific tail-motif substitutions in the FANSHY or the GGA-binding motifs on AD risk.

Rare variants with low and no priority: LPVs and NPVs.

Of all the rare variants, with MAF <0.05%, we identified 73 LPVs, which were carried by 87 AD cases with average age at onset of 70 years (95%CI: 67-75) and 85 non-demented controls. Carrying an LPV does not

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

associate with an increased risk of AD (OR=1.2, 95%CI (0.9 – 1.6; p=1) (**Table 3**). Similarly, we identified 269 NPVs, which were carried by 303 AD cases with average age at onset 73 years, (95%CI: 71-74) and 312 controls. Carrying an NPV also does not associate with increased risk of AD (OR=1.1; 95%CI 1.0 - 1.3, p=1) (**Table 3**). Furthermore, a survival analysis indicated that carrying an LPV or an NPV does not lead to a significantly expedited age at onset of AD relative to carriers of WT *SORL1* (**Fig 4A, Table 3**).

Non-rare *SORL1* variants with MAF>0.05% have only small effect on AD risk:

A total of 8,578 in our sample (21%) carried at least one of the 52 variants that were considered non-rare, 18 of which were common enough for imputation in the latest AD GWAS including 111,326 cases and 677,663 controls (Bellenguez et al., 2022), allowing the identification of variant specific effects (**Table S3**). The most common *SORL1* variant is the A528T substitution (rs2298813, CADD score 25.5, REVEL score 0.112, carried by 3.6% of the individuals in the sample), which associates with a small, 1.11-fold increased AD risk in the GWAS ($p=5.79x10^{-8}$) (**Table S3**). The E270K substitution also maps in the VPS10p domain (rs117260922, CADD score of 31, REVEL score of 0.31, carried by 1.9% of all individuals in the sample), but we find no evidence for an effect on AD (GWAS OR=1.02; $p=5.89x10^{-1}$) (**Table S3**). The age at onset of carriers from A528T and E270K variant carriers fully overlaps with carriers of WT *SORL1*. The D2065V substitution in the 3Fn domain (rs140327834, CADD score of 28.4 and a REVEL score of 0.568, carried by 0.46% of the sample), is associated with a 1.36-fold increased risk of AD in the GWAS ($p=1.61x10^{-6}$) (**Table S3**). Our analysis provides no evidence that any of the other non-rare variants in our sample associates with altered risk of AD (**Fig 4C, Table 3**)(**Table S3**).

Effect of *APOE*-e4 allele:

In the dataset (excluding the ADSP cohort, see methods) AD risk increases 3.1-fold for each added *APOE*-e4 allele (95% CI 2.9 – 3.3, p=0). The number of *APOE*-e4 alleles also affected age at onset: AD cases with no E4 allele have a median age at onset of 75 years, those with a single E4 allele have a median age at onset 70 years and homozygous E4 carriers have an age-at-onset of 64 years (**Fig 5**). In the *APOE* negative individuals in our dataset, carrying an PTV or HPV expedited age at onset by respectively 6 (95% CI 10 – 2) and 10 (95% CI 13 – 4) years. In AD cases who carry a single *APOE*-e4 allele, carrying an additional PTV or HPV expedited the median age at onset by respectively 9 (95% CI 10 – 6) and 7 (95% CI 9 – 4)years. In AD cases who carry two *APOE*-e4 alleles, carrying an additional PTV or HPV expedited age at onset by respectively 4 (95% CI 6 – NA) and 7 (95% CI 9 –

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

NA) years. For the MPVs, LPVs and NPVs there was no change in age at onset relative to *SORL1* WT, consistent with the negligible effect on AD risk (**Fig S2)**. While we observe a major additive effect of carrying a *SORL1* HPV and PTV on top *APOE*-e4, we also tested for interaction effect in the PTV and HPVs carriers (p=0.04 and p=0.06 respectively). However, we note that controls who carried a PTV or HPV and for whom *APOE* genotype was available were few (n=18), and predominantly negative for the E4 allele (n=15, 83%) and younger individuals (52% age < 65). This leads us to caution that this case/control analysis design lacks power, and may lead to incorrect inferences regarding a possible interaction effect, as we cannot take possible future conversion into account. Nevertheless, the additive and possibly synergistic effect of *APOE*-e4 allele explains, at least in part, the diverging age at onset of carriers of the same variant. Indeed, examining the age at onset among carriers of the same specific PTVs, HPVs and MPVs, while considering their *APOE* genotype, reveals substantial (**Fig S3**). The age at onset of the twelve Y391C cases ranged between 60 – 86 years, of five case-carriers of the 744R/* PTV ages ranged between 56-91 years, of the four case-carriers of the 866R/* PTV ages at onset ranged between 60-73 years, of the five case-carriers of the R953H HPV between 46-78 years, and of the six case-carriers of the Y1816C HPV onset ranged between 56 and 74 years. Indeed, close inspection indicates that the older cases often carry a protective *APOE*-e2 allele, while the earlier onset cases were more likely to have at least one *APOE*-e4 allele (**Fig S4**).

Comparison of prioritization scheme vs using only REVEL scores.

We conducted a comparative analysis to evaluate the effectiveness of our variant prioritization approach as opposed to relying solely on the REVEL score (**Fig S5**). Using our prioritization approach we identified 111 unique HPVs, which associated with a 6.42-fold increased risk of AD (i.e EAOD and LOAD). The aggregate of variants with a REVEL threshold of >0.7 associated with 3-fold increased risk of AD, and implementation of higher REVEL thresholds associated with risks never exceeded OR of 4. These results indicate that our approach outperforms the utilization of the REVEL prediction tool alone.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

In this work, we applied the results of an DMDM analysis as performed by Andersen et al (Andersen et al., 2023) to rare genetic *SORL1* missense variants observed in our assembled sample of 18,959 AD cases and 21,893 controls. Missense variants identified as 'high priority' HPV associated with an overall >6 fold increased AD risk (>10-fold increased risk of EOAD) and leads to a median 8.2 year earlier onset of AD in comparison to *SORL1* WT carriers. In comparison PTVs, which are considered clinically relevant, associated with an overall 12-fold increased AD risk (36-fold increased risk of EOAD) and leads to a median 10-year earlier onset of AD. In aggregate, HPV variants associate with a stronger increased risk compared to the aggregate risk of variants with the highest revel scores, indicating that this variant prioritization approach outperforms the application of the REVEL score.

Clearly, the effect of *SORL1* genetic variants on AD concentrates on the PTVs and HPVs. Variants that were annotated as moderate priority variants (MPVs) conferred a 1.6-fold increased risk of AD, and had negligible effects on age at onset. MPVs affecting the hydrophobic core of the YWTD domain or the conserved glycine at position 38 of the CR domain, seemed to occur more often in AD cases than in controls, but there is currently not enough functional evidence to include these variants in the HPV category. For common variants and for rare variants annotated as low and no priority (LPVs and NPVs), we observed no or only very limited associations with risk of AD, and no effect on age at onset. In contrast, HPV variants and PTVs lead to as similar average 8-10 year expedited age at onset relative to carriers of WT *SORL1*. Age at onset of specific carriers varies, and we observed that this is, at least in part, dependent on the additive, and possibly interactive effect of carrying no, one or two *APOE*-e4 alleles, which is in line with previous observations by us and others (Louwersheimer et al., 2017; Schramm et al., 2022). In our dataset, *APOE* genotype explained a >10-year difference in the median age at onset, recapitulating the effect of *APOE* effects observed in population studies (Desikan et al., 2017; van der Lee et al., 2018). Furthermore, it is according to expectations, that in addition to the *APOE* genotype, other risk alleles will further influence AD-age at onset (Bellenguez et al., 2022; Ryman et al., 2014).

Interestingly, we observed that specific HPVs had an *earlier* age-at-onset than the carriers of PTV variants, suggesting that such variants might have a dominant negative effect relative to the effect of haploinsufficiency (Verheijen et al., 2016). This was true for carriers of variants affecting the calcium cage in the CR domain, and carriers of variants affecting the YWTD-motif, and these were observed only in AD cases. One explanation for a

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

dominant negative effect may be that WT SORL1 dimerizes (or possibly polymerizes) at the luminal side of the endosomal tubule membrane (A. M. Jensen et al., 2023). However, a disruption of the calcium cage impairs protein folding (Blacklow et al., 1996; Fass et al., 1997), which precludes SORL1 maturation, such that the receptor is retained in the ER (Andersen et al., 2023). However, local misfolding of the CR- or YWTD-domain leaves the 3Fn-domain intact, such that the mutant receptor can still dimerize (and possibly polymerize) with both mutant and WT receptor, while retained at the ER. This way, less than half of SORL1 protein (as compared to WT SORL1 carriers) can leave the ER to perform its cargo trafficking functions. This may explain the autosomal dominant effect observed for specific variants. This is in agreement with the functional evidence associated with calcium cage variant p.D1545V observed in an Icelandic AD family, which binds WT receptor while strongly retained at the ER (Bjarnadottir, 2023). Furthermore, our data provide a preliminary indication that carriers of variants affecting the YWTD-motif and variants strongly affecting the VPS10p domain have an earlier age at onset than carriers of PTV variants. It is likely that such variants also affect the folding of the receptor, leading to a similar autosomal dominant effect due to dimerization of the mutant and wild-type receptor at the ER. Another example is the p.R953H variant that affects a conserved Arginine at position 38 of the YWTD domain, which predominantly observed in EOAD patients. Recent evidence on the p.R953C variant (the 'Seattle' mutation) indicates that substitution of an Arginine at this location leads to mis-localization within cells, leading to decreased maturation and shedding of the sSORL1 protein, suggesting that both p.R953H and p.R953C might also have a dominant negative effect on protein function (Fazeli et al., 2023).

On average however, HPVs have a similar effect on age at onset variants as PTVs, suggesting that the effects of most HPVs may be similar to the haploinsufficiency associated with PTV. An example of such a variant is the p.Y1816C that affects the 'tyrosine corner' substitution, a residue that contributes strongly to the stability of 3Fn-domains. In aggregated pedigrees, we recently showed that this variant leads to AD with an autosomal dominant inheritance pattern (A. M. G. Jensen et al., 2023). Functional experiments indicated that p.Y1816C mutant is efficiently matured, and trafficked from the ER to the endosome, but there it fails to form the dimer-dependent complex with retromer (A. M. G. Jensen et al., 2023), such that the SORL1 mutant cannot contribute to retromer sorting. However, the wild type allele still can, suggesting haploinsufficiency. This is supported by functional studies: the properly matured WT SORLA protein reaches the cell surface where it is cleaved and shedded as soluble SORLA (sSORLA) in the interstitial space. In a separate CSF analysis we observed that sSORLA levels for this mutant were ~50%, mimicking the effect we observe for PTVs on sSORLA shedding (Holstege et al., 2023; A. M. G. Jensen et al., 2023).

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

On the other hand, carriers of some HPVs have a slightly *later* onset than PTV variant carriers, suggesting that their effects are less damaging than the effect of losing one *SORL1* copy. Specifically, a Y391C substitution affecting Loop L1 in the VPS10p domain was observed in 12 unrelated cases and no controls. Functional evidence is needed, but we speculate that this variant might affect the binding of Amyloid-β (and other ligands) to the VPS10p-domain, to decreased lysosomal delivery of Amyloid-β and to an increase of secreted Amyloid-β (Caglayan et al., 2014); since other SORL1 functions may still be in-tact, this may explain a less deleterious effect compared to PTV. Furthermore, relative to carriers of PTV variants, we also observe later ages at onset for carriers of variants that affect the 10CC domain, or that that lead to losing or gaining a cysteine in the CR domain, suggesting that these mutants still have some residual activity. Possibly, these variants may lead to an unstable receptor, which may be in part removed by the ER-associated degradation (ERAD) pathway while some may escape the ER control check and be exported to subsequent cellular compartments. However, we can currently not present any evidence that supports this.

It is important to note that the associations with AD risk and observed effects on age-at-onset are specific for the AD patients in this dataset, which is relatively enriched with EOAD cases, such that derived incidence curves do not represent the incidence in the overall population. Nevertheless, it is valid to compare risk and age at onset effects for the different SORL1 classes of variants *within* our sample. We further acknowledge that determining odds ratio's for EOAD and LOAD separately only partly accounts for the influence of age on the effect of *SORL1* variants on AD risk. Furthermore, 33% of the controls was younger than 50, and 50% was younger than 65, at the time of sample inclusion, and it is likely that a non-negligible fraction of controls may develop AD at a later age, such that effect sizes presented here are conservative. Ideally, risk and age-at-onset analysis are assessed in population-based follow-up studies.

During our analyses, we came across several idiosyncrasies in the *SORL1* gene that might be considered when analyzing *SORL1* variants. For example, exons 23-33 each translate one of the 11 CR domains. It is currently unclear whether exon-skipping splice variants may translate to the in-frame removal of a complete CR domain, yielding a (partly) functional SORL1 protein with one missing CR domain, or whether this leads to the generation of alternative and non-productive transcripts (Le Guennec et al., 2018). An exception is the 7th CR domain encoded by exon 29, since joining exons 28 and 30 produces a nonsense-codon. Furthermore, we noticed that variants in exon 1, which was excluded from analysis due to differential missingness, included 5 PTVs which occurred 5 AD cases and 4 controls, suggesting a possible enrichment of PTVs in exon 1. Since PTVs in exon 1 are similarly

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

likely to lead to *SORL1* haploinsufficiency, the only reasonable explanation for a possible enrichment is the use of an alternative transcriptional start site, as a back-up mechanism for *SORL1* transcription. However, we currently have no supporting evidence to substantiate this. Furthermore, we acknowledge that in our risk assessment we included only the variants observed in our exome sequencing dataset, but this list of variants is far from exhaustive. We have seen several publications reporting the identification of *SORL1* variants that were not observed in our sample namely: R953C (Fazeli et al., 2023), R1303C (Thonberg et al., 2017), R1084C (El Bitar et al., 2019), C1192Y (Cao et al., 2021), C1344R (Thonberg et al., 2017), C1453S and C1249S (Verheijen et al., 2016). Furthermore, there may be non-coding and structural variants in and around the *SORL1* gene that may affect protein function, and we did not take these into account in the current analysis. Lastly, while most HPV or PTV variants occurred only once or twice in the sample, several variants occurred more frequently, these included among others Y391C (12 cases), R1490C (5 cases/2 controls), Y1816C (6 cases), R953H (6 cases/1 control), 744R/* (5 cases) and 866R/* (4 cases). This prompts an investigation of whether these unrelated individuals share a founder mutation, allowing a segregation analysis across assembled pedigrees of variant carriers. Alternatively, these variants may have occurred de novo in each pedigree, which would provide first preliminary evidence that mutational hotspots exist in *SORL1* (Nesta et al., 2021).

Conclusions: While several laboratories are currently considering *SORL1* PTVs as clinically important, we show that next to PTVs, selected missense variants also deserve clinical attention. We show that, depending on the affected residue and domain, variant-effects range from dominant negative, haploinsufficient to risk-increasing. We propose that HPVs reported in the manuscript should be reported back to clinicians, so they may consider performing segregation analyses. We hypothesize this will provide additional evidence of autosomal dominant inheritance patterns for many additional *SORL1* variants.

Acknowledgments

The authors are grateful to all study participants, their family members, the participating medical staff, general practitioners, pharmacists and all laboratory personnel involved in patient diagnosis, blood collection, blood biobanking, DNA preparation and sequencing. The data used in this work was collected using the funding obtained by the following study cohorts: ADES-FR, AgeCoDe-UKBonn; Barcelona SPIN; AC-EMC; ERF and Rotterdam; ADC-Amsterdam; 100-plus study; EMIF-90+; Control Brain Consortium; PERADES; StEP-AD; Knight-ADRC;

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

UCSF/NYGC/UAB; UCL-DRC EOAD; ADSP. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu/). Full consortium acknowledgements and funding sources are listed in the Supplementary Note.

The work in this manuscript was carried out on the Cartesius supercomputer, which is embedded in the Dutch national e-infrastructure with the support of SURF Cooperative. Computing hours were granted to H. H. by the Dutch Research Council ('100plus': project# vuh15226, 15318, 17232, and 2020.030; 'Role of VNTRs in AD'; project# 2022.31, 'Alzheimer's Genetics Hub' project# 2022.38). H.H. and O.M.A. are a part of the EU Joint Programme-Neurodegenerative Disease Research (JPND) Working Group SORLA-FIX under the 2019 ''Personalized Medicine'' call (JPND2019-466-197, ZonMW 733051110, Danish Innovation Foundation and the Velux Foundation Denmark). H.H., S.L., are recipients of ABOARD, a public-private partnership receiving funding from ZonMW (#73305095007) and Health~Holland, Topsector Life Sciences & Health (PPP-allowance; #LSHM20106). S.L. is recipient of ZonMW funding (#733050512). H.H. was supported by the Hans und Ilse Breuer Stiftung (2020), Dioraphte 16020404 (2014) and the HorstingStuit Foundation (2018). O.M.A. is recipient of funding related to the 2022 Research Prize from the Alzheimers Research Foundation Denmark. Additional funding was provided by: Direktør Emil C. Hertz og Hustru Inger Hertz' Fond (to M.L.); HH has a collaboration contract with Muna Therapeutics, PacBio, Neurimmune and Alchemab. She serves in the scientific advisory boards of Muna Therapeutics and is an external advisor for Retromer Therapeutics.

Author contributions: Conceived the study: HH and OMA; Wrote the manuscript: HH, MdW, NT, SvdL, MH, OMA; Cohort participant collection and sequencing: ADES, ADSP, StEP-AD, Knight ADRC, UCSF/NYGC/UAB. Genetic Analyses: HH, MdW, NT, SvdL, MV, RvS, MH, OMA

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Tables

	Cases	Controls
Total N	18,959	21,893
Females (%)	59.49%	58.23%
Mean Age (SD)	72.4 ± 10.7 years	71.1 ± 16.7 years
Early onset AD (%)	32.46%	N/A
APOE ε4 (%)	Hom: 9.3% Het: 41.1% Non: 49% Missing: 0.6%	Hom: 1.1% Het: 16.1% Non: 53.5% Missing: 29.3%
Non-European ancestry	14.9%	25.45%

Table 1. Cohort characteristics

Prioritization category	Selection criteria
PTV: protein truncating variants	All truncating variants: i.e. nonsense, frameshift and splice variants*
HPV: High priority missense variants	Variants that affect *high-priority* residues (**Table S4)** Variants in the p.VPS10p and 10CC domains *with* REVEL score ≥0.50
MPV: Moderate priority variants	Variants that affect *moderate-priority* residues (**Table S5)**
LPV: Low priority variants	Variants *not prioritized with* REVEL score ≥0.50
NPV: No priority variants	All remaining variants

Table 2. Prioritization scheme of rare variants. Rare variants with MAF<0.05% were considered for prioritization. HPV and MPVs affect residues corresponding to the black and grey residues in **Figure 3**. [1]**MAF**: Variant Minor Allele Frequency in the non-neuro pop-max dataset (Gnomad v.2.1.1). When unavailable, we used the variant MAF in the sample. Non-neuro sample in GnomAD is the sample without individuals with neurological diseases. The pop-max is the population with the highest frequency (pop-max). [2]**REVEL** (Ioannidis et al): Variant effect prediction algorithm: Rare Exome Variant Ensemble Learner. Scores range from 0 to 1 and variants with higher scores are predicted to be more deleterious.

Coding rare variants	Variant Subcategory	carriers	# unique variants	Variant effect on AD risk						Effect on age at onset		
	domain residues	**all/EOAD/LOAD/ controls**	**all/EOAD/LOAD/controls**	**All carriers OR**	**All carriers p value**	**EOAD OR**	**EOAD p value**	**LOAD OR**	**LOAD p value**	**age at onset (95%CI)**	**Δ age at onset vs WT (95% CI)**	**p value***
SORL1 WT												
WT		31,331/4,644/9,680/17,007	**N/A**	0.89 (0.85 - 0.93)	1.12^{E}-05	0.88 (0.83 - 0.94)	**7.0E-3**	0.89 (0.85 - 0.94)	**2.6E-04**	72 (72 - 73)	NA	N/A
Non-rare variants												
—Non-rare	E270K*	1,409/212/479/718	1/1/1/1	1.02 (0.92 - 1.13)	6.88E-01	0.94 (0.81 - 1.1)	4.3E-01	1.1 (0.95 - 1.2)	2.9E-01	72 (71 - 73)	0 (-1 — 0)	1
	A528T*	2,719/495/961/1,263	1/1/1/1	1.2 (1.08 - 1.26)	6.15E-05	1.1 (1.03 - 1.27)	1.2E-02	1.2 (1.08 - 1.28)	1.6 E-04	71 (70 - 72)	-1 (-2 — -1)	1.7E-01
	D2065V*	353/55/106/192	1/1/1/1	0.9 (0.7 - 1.07)	1.79E-01	0.87 (0.64 - 1.18)	3.7E-01	0.9 (0.68 - 1.09)	2.2E-01	70 (68 - 72)	-2 (-4 — -1)	7.1E-01
	All other*	4,035/500/1,314/2,221	48/34/38/45	1.0 (0.95 - 1.09)	6.0E-01	1.0 (0.91 – 1.12)	8.2E-01	1.0 (0.95 - 1.10)	6.4E-01	74 (73 - 74)	-2 (-1 — -1)	1
Protein Truncating Variants												
PTV	**Total**	95/59/30/6	74/52/25/6	17.2 (7.5 - 39.3)	1.2E-21	35.3 (15.2 - 81.8)	5.9E-31	8.6 (3.6 - 20.6)	3.6E-07	62 (60 - 65)	-10 (-12 — -8)	2.7E-11
Rare Missense variants												
HPV: high priority	Total	176/79/70/27	111/62/53/22	6.4 (4.3 - 9.7)	2.1E-24	10.5 (6.8 - 16.3)	3.0E-29	4.5 (2.9 - 6.9)	4.9E-11	63.8 (62 - 67)	-8.2 (-10 — -6)	1.0E-08
—VPS10p	Total	43/22/16/5	27/17/10/5	8.8 (3.46 - 22.34)	3.2E-07	15.7 (5.9 - 41.5)	2.1E-09	5.5 (2.0 - 15)	9.1E-03	59.5 (56 - 78)	-12.5 (-16 —-5)	1
	—Cysteins gained & cysteins lost L1/L2	18/8/8/2	6/4/2/2	9.24 (2.1 - 40.2)	6.5E-03	14.2 (3.2 - 67.1)	4.3E-03	6.8 (1.5 - 32.2)	1.8E-01			
	—Asp-box, REVEL>0.5	6/3/3/0	5/3/3/0	NA	2.7E-01	NA	3.0E-01	NA	1			
	—Remaining p.VPS10p variants, REVEL>0.5	19/11/5/3	16/10/5/3	6.2 (1.8 - 21.2)	2.5E-02	13.1 (3.6 - 46.8)	2.7E-04	2.9 (0.7 - 11.9)	1			
—10CC	Total	30/11/13/6	19/8/10/3	4.6 (1.9 - 11.3)	8.4E-03	6.5 (2.4 - 17.7)	4.9E-03	3.7 (1.4 - 9.8)	1.9E-01	67 (63 - 73)	-5 (-9 — 0)	1
—YWTD	Total	25/13/7/5	12/9/6/5	4.6 (1.7 - 12.3)	2.5E-02	9.3 (3.3 - 26)	2.0E-04	2.4 (0.8 - 7.6)	1			
	—YWTD-motif (17, 18, 19, 20)	8/6/2/0	6/5/2/0	NA	5.9E-02	NA	3.0E-03	NA	1	64 (48 - NA)	-8 (-24 — NA)	2.6E-03
	—Highly conserved residues (29, 35)	8/2/3/3	3/2/2/3	1.9 (0.5 - 8.1)	1	2.37 4(0.4 - 14.2)	1	1.7 (0.4 - 8.5)	1			
	—partly conserved residues (9,38)	8/4/2/2	2/1/2/2	3.5 (0.7 - 17.2)	1	7.1 (1.3 - 38.9)	6.5E-01	1.71(0.2 - 12.1)	1			
—EGF	Cysteines gained or lost	2/1/0/1	2/1/0/1	1.15 (0.1 - 18.5)	1	3.6 (0.2 - 56.9)	1	NA	1			
—CR	Total	59/24/27/8	38/22/20/6	7.4 (3.5 - 15.6)	5.4E-09	10.7 (4.8 - 23.9)	6.5E-09	5.8 (2.6 - 12.7)	3.9E-05			
	—Calcium Cages (D, D, D, E) 37, 41, 47, 48	13/9/4/0	12/9/4/0	NA	1.2E-03	NA	3.2E-05	NA	4.9E-01	60 (56 - NA)	-12 (-16 — NA)	7.7E-04
	—Cysteines gained or lost	44/13/23/8	24/11/16/6	5.2 (2.2 - 11.2)	5.9E-05	5.8 (2.4 - 14.0)	2.4E-03	4.9 (2.2 - 11.0)	1.0E-03	68 (63 - 74.3)	-4 (-9 — 1.3)	1
	—Asx-turn (D) (44)	2/2/0/0	2/2/0/0	NA	1	NA	1	NA	NA			
—3FN	Total	15/8/5/2	10/5/4/2	7.50 (1.7 - 33.3)	7.6E-02	14.2 (3.0 - 67.1)	4.3E-03	4.3 (0.8 - 22.0)	1			
	—Partly conserved glycines (36, 96)	2/1/1/0	2/1/1/0	NA	1	NA	1	NA	1			
	—Partly conserved prolines (6, 7, 79)	4/2/2/0	4/2/2/0	NA	1	NA	1	NA	1			
	—highly conserved residues (25,41,77, 83)	9/5/2/2	4/2/1/2	4.0 (0.8 - 19.5)	1	8.9 (1.7 - 45.9)	1.9E-01	1.7 (0.2 - 12.1)	1			
MPV: Moderate priority	all	144/22/61/61	67/18/39/33	1.6 (1.1 - 2.2)	1.7E-01	1.3 (0.8 - 2.1)	1	1.7 (1.2 - 2.4)	7.5E-02	72 (71 - 74)	0 (-1 —1)	1
LPV: low priority	all	172/35/52/85	73/23/32/46	1.2 (0.9 - 1.6)	1	1.5 (1.0 - 2.2)	1	1.1 (0.7 -1.5)	1	70 (67 - 75)	2 (-5 - 2)	1
NPV: no priority	all	615/102/201/312	269/76/137/158	1.12 (1.0 - 1.3)	1	1.2 (0.9 - 1.5)	1	1.1 (0.9 - 1.3)	1	73 (71 - 74)	1 (1 — 1)	1

Table 3. Effect of variant subtypes on AD risk, and effect on age at onset: OR: odds ratio, calculated using a Fisher's Exact test. *Odds rations for all non-rare variants were calculated using a logistic regression model on the minor allele dosages. Note that all effect sizes were calculated relative to the same control group, which was relatively young, such that effect sizes may be conservative.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figures

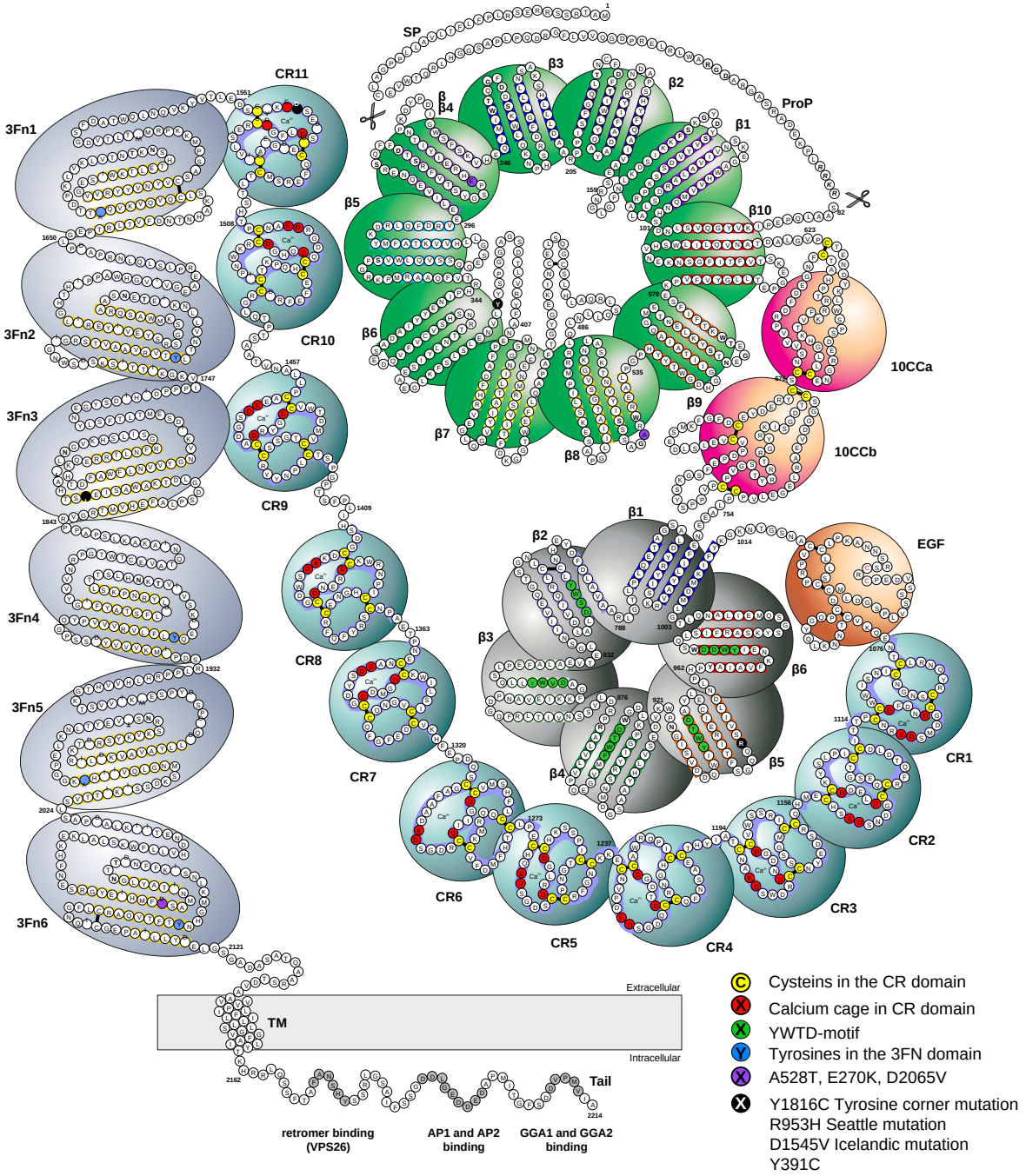

Figure 1: Schematic of the SORL1 protein, highlighting the variants with the strongest effects on AD risk. Note how conserved residues exist in all repeats in a domain.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

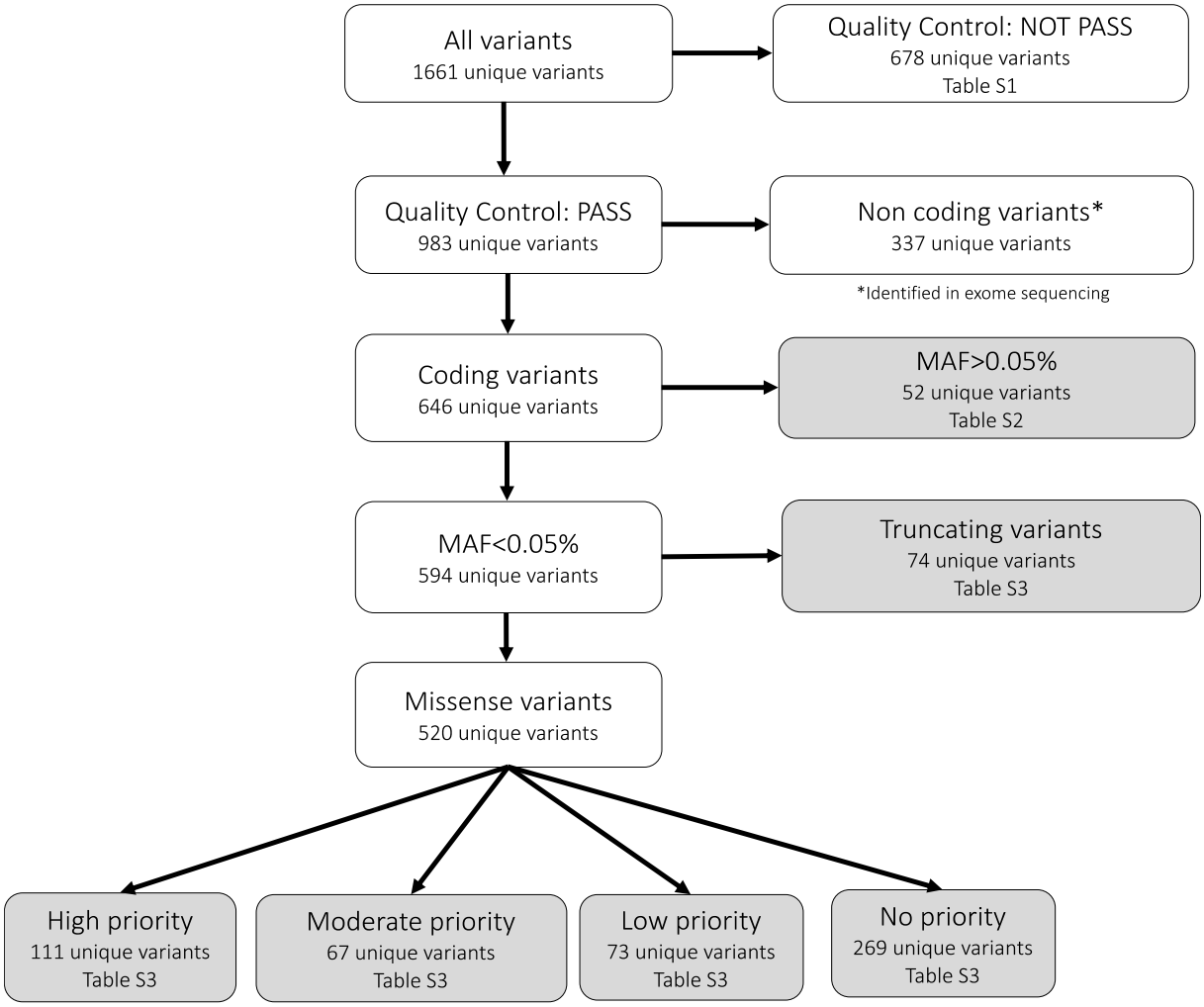

Figure 2: Flowchart of selection procedures for *SORL1* variant subtypes. *Non-coding variants were identified in the padding of exome sequencing or in exome excerpts from whole genome sequences.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

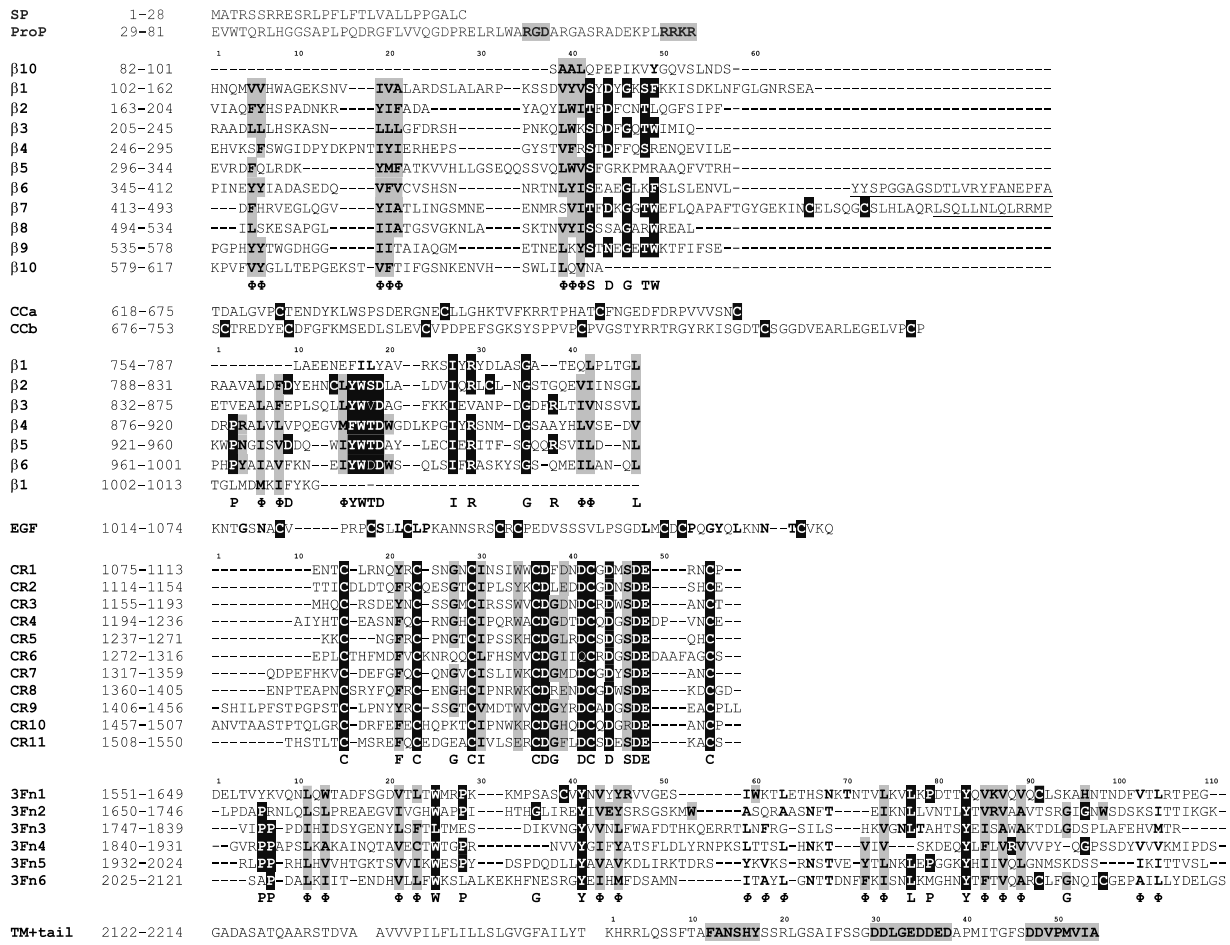

Figure 3: Protein sequence alignments per SORL1 subdomain. The protein sequence was aligned for each repeat in each SORL1 subdomain, revealing conserved residues. Residues likely to harbor deleterious mutations based on either domain sequence conservation or because the DMDM analysis were prioritized. **Black:** High priority variants (HPVs) affect residues annotated in black; **Grey:** Moderate priority variants (MPVs) affect residues indicated in grey.

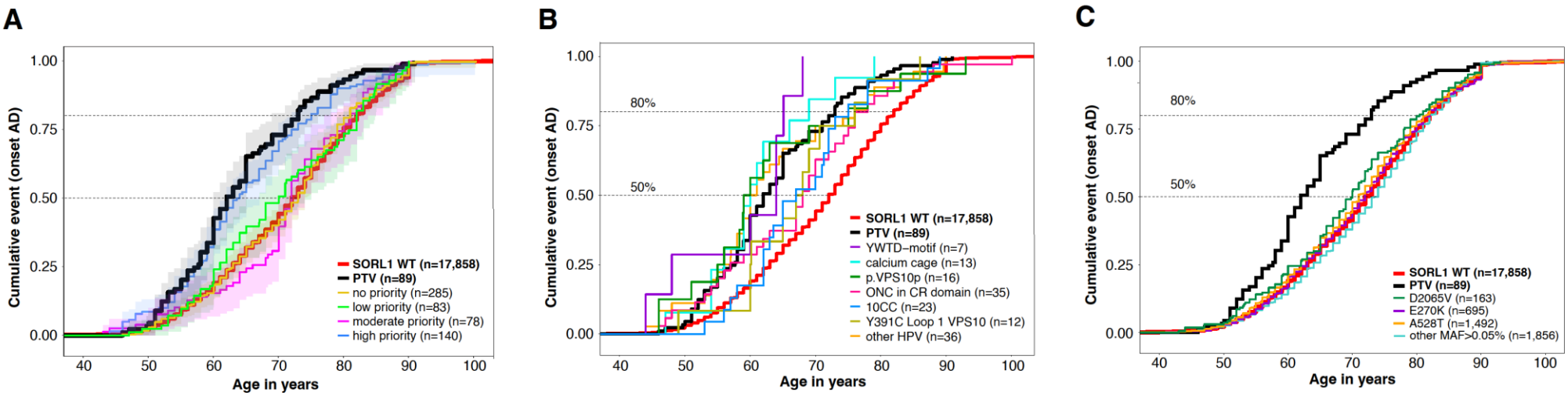

Figure 4. Age at onset analysis. AD patients were categorized according to the variant subtype they carry. **A)** Age at onset analysis for carriers of missense variants annotated as high, moderate, low and no priority, compared to carriers of PTVs and SORL1 WT carriers. **B)** Age at onset analysis of carriers of specific high priority missense variants, relative to carriers of PTVs and WTs. **C)** Age at onset analysis of common variants. Horizontal lines indicate the age at which 50% of all carriers with the same variant category have AD, and the age at which 80% of variant-carriers developed AD. Differences in ages at onset between carriers of variants appertaining to specific variant-groups are shown in **Table 3**.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

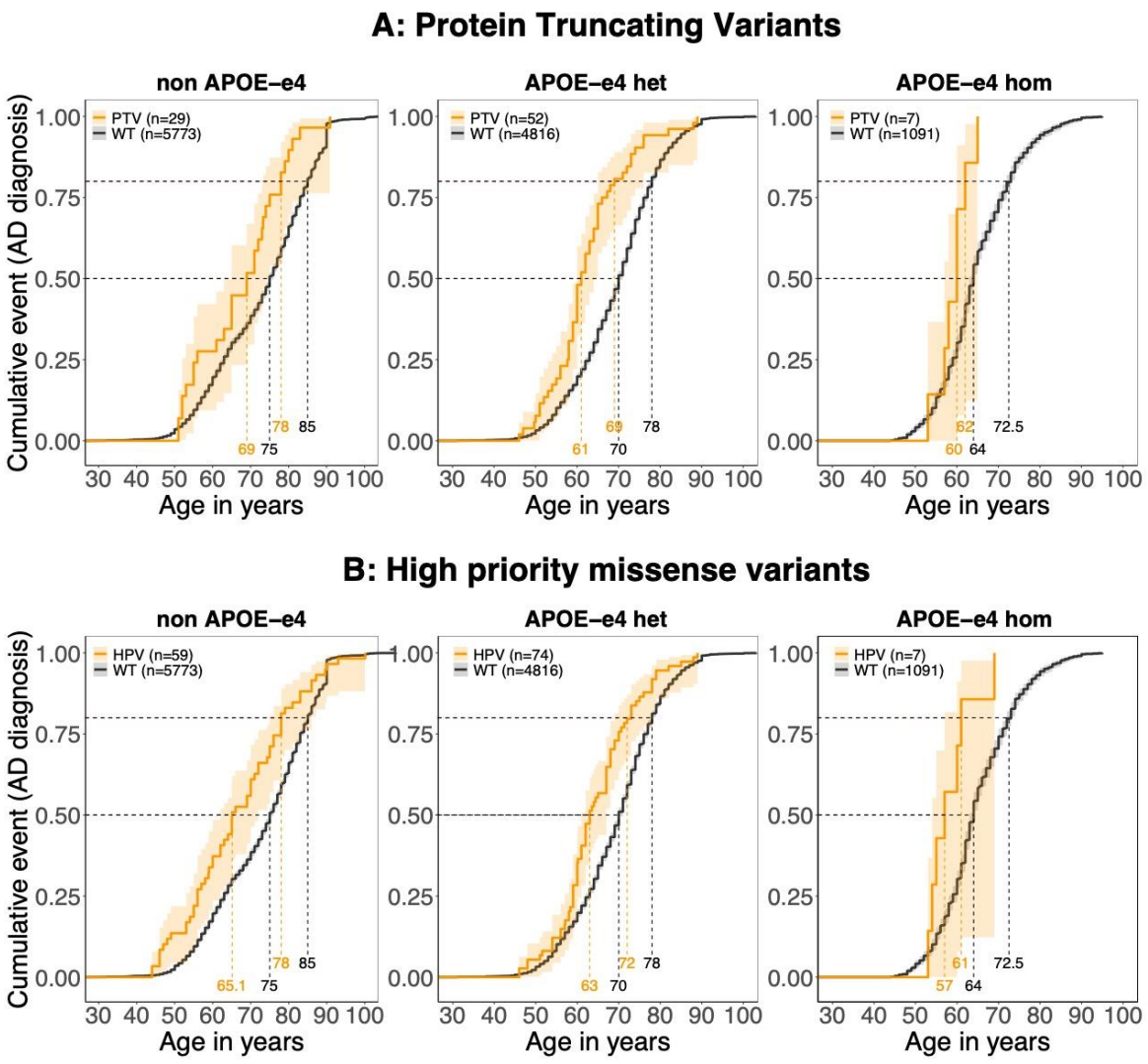

Figure 5: Rare *SORL1* variants in context of APOE genotype. A. AD patients who carried PTV in *SORL1* have an earlier age at onset compared to SORL1 wildtype carriers with the same APOE genotype. The orange ages indicate at what age respectively 50% and 80% of the variant carriers had Alzheimer's Disease, the black ages indicate at the age at which 50% or 80% of the SORL1 WT carriers developed AD. **B.** High priority variants in context of APOE genotype**.** carriers of a high priority variant have an earlier age at onset compared to wildtype carriers with similar APOE genotype. The effects of MPVs, LPVs and NPVs in context of APOE is shown in **Fig S2**.

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Reference List:

Andersen, O. M., Monti, G., Jensen, A. M. G., de Waal, M., Hulsman, M., Olsen, J. G., & Holstege, H. (2023). Relying on the relationship with known disease-causing variants in homologous proteins to predict pathogenicity of SORL1 variants in Alzheimer's disease. *BioRxiv*, 1–26.

Bellenguez, C., Küçükali, F., Jansen, I. E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A. C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., Holmans, P. A., Boland, A., Damotte, V., van der Lee, S. J., Costa, M. R., Kuulasmaa, T., Yang, Q., de Rojas, I., Bis, J. C., … Sánchez-Arjona, M. B. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nature Genetics*, *54*(4), 412–436. doi: 10.1038/s41588-022-01024-z

Bettens, K., Brouwers, N., Engelborghs, S., De Deyn, P. P., Van Broeckhoven, C., & Sleegers, K. (2008). SORL1 is genetically associated with increased risk for late-onset Alzheimer disease in the Belgian population. *Human Mutation*, *29*(5), 769–770. doi: https://doi.org/10.1002/humu.20725

Bjarnadottir, K. (2023). Recent SORL1 missense variant causing a SORLA maturation defect is associated with Alzheimer's disease. *In Preparation*.

Blacklow, S. C., & Kim, P. S. (1996). Protein folding and calcium binding defects arising from familial hypercholesterolemia mutations of the LDL receptor. *Nature Structural Biology*, *3*(9), 758–762.

Caglayan, S., Takagi-Niidome, S., Liao, F., Carlo, A.-S., Schmidt, V., Burgert, T., Kitago, Y., Füchtbauer, E.-M., Füchtbauer, A., Holtzman, D. M., Takagi, J., & Willnow, T. E. (2014). Lysosomal Sorting of Amyloid-β by the SORLA Receptor Is Impaired by a Familial Alzheimer's Disease Mutation. *Science Translational Medicine*, *6*(223), 223ra20-223ra20. doi: 10.1126/scitranslmed.3007747

Campion, D., Charbonnier, C., & Nicolas, G. (2019). SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data. *Acta Neuropathologica*, *138*(2), 173–186. doi: 10.1007/s00401-019-01991-4

Cao, L., Zhu, F., & Qiu, G. (2021). Early-onset Alzheimer's disease may be associated with sortilin-related receptor 1 gene mutation: A family report and review. *Radiology Case Reports*, *16*(1), 30–34. doi: https://doi.org/10.1016/j.radcr.2020.10.030

Desikan, R. S., Fan, C. C., Wang, Y., Schork, A. J., Cabral, H. J., Cupples, L. A., Thompson, W. K., Besser, L., Kukull, W. A., Holland, D., Chen, C.-H., Brewer, J. B., Karow, D. S., Kauppi, K., Witoelar, A., Karch, C. M., Bonham, L. W., Yokoyama, J. S., Rosen, H. J., … Dale, A. M. (2017). Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLOS*

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Medicine, *14*(3), e1002258. Retrieved from https://doi.org/10.1371/journal.pmed.1002258

El Bitar, F., Qadi, N., Al Rajeh, S., Majrashi, A., Abdulaziz, S., Majrashi, N., Al Inizi, M., Taher, A., & Al Tassan, N. (2019). Genetic Study of Alzheimer's Disease in Saudi Population. *Journal of Alzheimer's Disease*, *67*, 231–242. doi: 10.3233/JAD-180415

Fass, D., Blacklow, S., Kim, P. S., & Berger, J. M. (1997). Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. *Nature*, *388*(6643), 691–693. doi: 10.1038/41798

Fazeli, E., Child, D. D., Bucks, S. A., Stovarsky, M., Edwards, G., Yu, C.-E., Latimer, C., Kitago, Y., Bird, T., Andersen, O. M., Jayadev, S., & Young, J. E. (2023). A familial missense variant in the AD gene SORL1 impairs its maturation and endosomal sorting. *BioRxiv*, 2023.07.01.547348. doi: 10.1101/2023.07.01.547348

Fjorback, A. W., Seaman, M., Gustafsen, C., Mehmedbasic, A., Gokool, S., Wu, C., Militz, D., Schmidt, V., Madsen, P., Nyengaard, J. R., Willnow, T. E., Christensen, E. I., Mobley, W. B., Nykjær, A., & Andersen, O. M. (2012). Retromer Binds the FANSHY Sorting Motif in SorLA to Regulate Amyloid Precursor Protein Sorting and Processing. *The Journal of Neuroscience*, *32*(4), 1467 LP – 1480. doi: 10.1523/JNEUROSCI.2272-11.2012

Holstege, H., Hulsman, M., Charbonnier, C., Grenier-Boley, B., Quenez, O., Grozeva, D., van Rooij, J. G. J., Sims, R., Ahmad, S., Amin, N., Norsworthy, P. J., Dols-Icardo, O., Hummerich, H., Kawalia, A., Amouyel, P., Beecham, G. W., Berr, C., Bis, J. C., Boland, A., … Lambert, J.-C. (2022). Exome sequencing identifies rare damaging variants in ATP8B4 and ABCA1 as risk factors for Alzheimer's disease. *Nature Genetics*, *54*(12), 1786–1794. doi: 10.1038/s41588-022-01208-7

Holstege, H., Lunding, M., Tijms, B. M., van der Lee, S. J., Hulsman, M., Weiss, M. M., van Spaendonk, R., Visser, P. J., Jensen, A. M. G., Vægter, C. B., van der Flier, W. M., Teunissen, C. E., Monti, G., & Andersen, O. M. (2023). Decreased soluble SORLA levels as CSF biomarker of pathogenic SORL1 genetic variants in Alzheimer's disease patients. *In Preparation*.

Holstege, H., van der Lee, S. J., Hulsman, M., Wong, T. H., van Rooij, J. G. J., Weiss, M., Louwersheimer, E., Wolters, F. J., Amin, N., Uitterlinden, A. G., Hofman, A., Ikram, M. A., van Swieten, J. C., Meijers-Heijboer, H., van der Flier, W. M., Reinders, M. J. T., van Duijn, C. M., & Scheltens, P. (2017). Characterization of pathogenic SORL1 genetic variants for association with Alzheimer's disease: a clinical interpretation strategy. *European Journal of Human Genetics*, *25*(8), 973–981. doi: 10.1038/ejhg.2017.87

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q.,

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., … Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *The American Journal of Human Genetics*, *99*(4), 877–885. doi: https://doi.org/10.1016/j.ajhg.2016.08.016

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., Kosmicki, J. A., Arbelaez, J., Cui, W., Schwartz, G. B., Chow, E. D., Kanterakis, E., Gao, H., Kia, A., Batzoglou, S., Sanders, S. J., & Farh, K. K.-H. (2019). Predicting Splicing from Primary Sequence with Deep Learning. *Cell*, *176*(3), 535-548.e24. doi: https://doi.org/10.1016/j.cell.2018.12.015

Jensen, A. M. G., Raska, J., Fojtik, P., Monti, G., Lunding, M., Vochyanova, S., Pospisilova, V., van der Lee, S. J., van Dongen, J., Bossaerts, L., Van Broeckhoven, C., Dols, O., Lleo, A., Benussi, L., Ghidoni, R., Hulsman, M., Sleegers, K., Bohaciakova, D., Holstege, H., & Andersen, O. M. (2023). The SORL1 p.Y1816C variant causes impaired endosomal dimerization and autosomal dominant Alzheimer's disease. *MedRxiv*. doi: https://doi.org/10.1101/2023.07.09.23292253

Jensen, A. M., Kitago, Y., Fazeli, E., Vægter, C. B., Small, S., Petsko, G. A., & Andersen, O. M. (2023). Dimerization of the Alzheimer ' s disease pathogenic receptor SORLA regulates its association with retromer. *PNAS*, *120*(4), 1–11.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., … Consortium, G. A. D. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. doi: 10.1038/s41586-020-2308-7

Kitago, Y., Nagae, M., Nakata, Z., Yagi-Utsumi, M., Takagi-Niidome, S., Mihara, E., Nogi, T., Kato, K., & Takagi, J. (2015). Structural basis for amyloidogenic peptide recognition by sorLA. *Nature Structural & Molecular Biology*, *22*(3), 199–206. doi: 10.1038/nsmb.2954

Le Guennec, K., Tubeuf, H., Hannequin, D., Wallon, D., Quenez, O., Rousseau, S., Richard, A.-C., Deleuze, J.-F., Boland, A., Frebourg, T., Gaildrat, P., Campion, D., Martins, A., & Nicolas, G. (2018). Biallelic Loss of Function of SORL1 in an Early Onset Alzheimer's Disease Patient. *Journal of Alzheimer's Disease*, *62*, 821–831. doi: 10.3233/JAD-170981

Louwersheimer, E., Cohn-Hokke, P. E., Pijnenburg, Y. A. L., & Weiss, M. M. (2017). Rare Genetic Variant in SORL1 May Increase Penetrance of Alzheimer ' s Disease in a Family with Several Generations of APOE

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

- E4 Homozygosity. *Journal of Alzheimer's Disease*, *56*, 63–74. doi: 10.3233/JAD-160091

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease. *Neurology*, *34*(7), 939 LP – 939. doi: 10.1212/WNL.34.7.939

Mckhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer ' s disease : Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 263–269. doi: 10.1016/j.jalz.2011.03.005

Mishra, S., Knupp, A., Szabo, M. P., Williams, C. A., & Kinoshita, C. (2022). The Alzheimer's gene SORL1 is a regulator of endosomal traffic and recycling in human neurons. *Cellular and Molecular Life Sciences*, *79*, 1–22. doi: 10.1007/s00018-022-04182-9

Miyashita, A., Koike, A., Jun, G., Wang, L.-S., Takahashi, S., Matsubara, E., Kawarabayashi, T., Shoji, M., Tomita, N., Arai, H., Asada, T., Harigaya, Y., Ikeda, M., Amari, M., Hanyu, H., Higuchi, S., Ikeuchi, T., Nishizawa, M., Suga, M., … Kuwano, R. (2013). SORL1 Is Genetically Associated with Late-Onset Alzheimer's Disease in Japanese, Koreans and Caucasians. *PLOS ONE*, *8*(4), e58618. Retrieved from https://doi.org/10.1371/journal.pone.0058618

Nesta, A. V, Tafur, D., & Beck, C. R. (2021). Hotspots of Human Mutation. *Trends in Genetics*, *37*(8), 717–729. doi: https://doi.org/10.1016/j.tig.2020.10.003

Pottier, C., Hannequin, D., Coutant, S., Rovelet-Lecrux, A., Wallon, D., Rousseau, S., Legallic, S., Paquet, C., Bombois, S., Pariente, J., Thomas-Anterion, C., Michon, A., Croisile, B., Etcharry-Bouyx, F., Berr, C., Dartigues, J.-F., Amouyel, P., Dauchel, H., Boutoleau-Bretonnière, C., … Collaborators, P. G. (2012). High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Molecular Psychiatry*, *17*(9), 875–879. doi: 10.1038/mp.2012.15

Rogaeva, E., Meng, Y., Lee, J. H., Gu, Y., Kawarai, T., Zou, F., Katayama, T., Baldwin, C. T., Cheng, R., Hasegawa, H., Chen, F., Shibata, N., Lunetta, K. L., Pardossi-Piquard, R., Bohm, C., Wakutani, Y., Cupples, L. A., Cuenco, K. T., Green, R. C., … St George-Hyslop, P. (2007). The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nature Genetics*, *39*(2), 168–177. doi: 10.1038/ng1943

Ryman, D. C., Acosta-Baena, N., Aisen, P. S., Bird, T., Danek, A., Fox, N. C., Goate, A., Frommelt, P., Ghetti,

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

B., Langbaum, J. B. S., Lopera, F., Martins, R., Masters, C. L., Mayeux, R. P., McDade, E., Moreno, S., Reiman, E. M., Ringman, J. M., Salloway, S., … Network, A. the D. I. A. (2014). Symptom onset in autosomal dominant Alzheimer disease. *Neurology*, *83*(3), 253 LP – 260. doi: 10.1212/WNL.0000000000000596

Schramm, C., Charbonnier, C., Zaréa, A., Lacour, M., Wallon, D., Andriuta, D., Anthony, P., Auriacombe, S., Balageas, A.-C., Ballan, G., Barbay, M., Beaufils, E., Béjot, Y., Belliard, S., Benaiteau, M., Bennys, K., Blanc, F., Bombois, S., Boutoleau Bretonnière, C., … (ADSP), A. D. S. P. (2022). Penetrance estimation of Alzheimer disease in SORL1 loss-of-function variant carriers using a family-based strategy and stratification by APOE genotypes. *Genome Medicine*, *14*(1), 69. doi: 10.1186/s13073-022-01070-6

Small, S. A., & Petsko, G. A. (2015). Retromer in Alzheimer disease, Parkinson disease and other neurological disorders. *Nature Reviews Neuroscience*, *16*, 126–132.

Thonberg, H., Chiang, H.-H., Lilius, L., Forsell, C., Lindström, A.-K., Johansson, C., Björkström, J., Thordardottir, S., Sleegers, K., Van Broeckhoven, C., Rönnbäck, A., & Graff, C. (2017). Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. *Acta Neuropathologica Communications*, *5*(1), 43. doi: 10.1186/s40478-017-0441-9

Van der Lee, S. J., Minderhoud, C., van Spaendonk, R., Meijers-Heijboer, H., Van der Flier, W. M., Scheltens, P., & Holstege, H. (2018). P4-038: IS SORL1 AN AUTOSOMAL DOMINANT ALZHEIMER GENE? *Alzheimer's & Dementia*, *14*(7S_Part_27), P1447–P1447. doi: https://doi.org/10.1016/j.jalz.2018.06.2440

van der Lee, S. J., Wolters, F. J., Ikram, M. K., Hofman, A., Ikram, M. A., Amin, N., & van Duijn, C. M. (2018). The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. *The Lancet Neurology*, *17*(5), 434–444. doi: https://doi.org/10.1016/S1474-4422(18)30053-X

Vardarajan, B. N., Zhang, Y., Lee, J. H., Cheng, R., Bohm, C., Ghani, M., Reitz, C., Reyes-Dumeyer, D., Shen, Y., Rogaeva, E., St George-Hyslop, P., & Mayeux, R. (2015). Coding mutations in SORL1 and Alzheimer disease. *Annals of Neurology*, *77*(2), 215–227. doi: https://doi.org/10.1002/ana.24305

Verheijen, J., Van den Bossche, T., van der Zee, J., Engelborghs, S., Sanchez-Valle, R., Lladó, A., Graff, C., Thonberg, H., Pastor, P., Ortega-Cubero, S., Pastor, M. A., Benussi, L., Ghidoni, R., Binetti, G., Clarimon, J., Lleó, A., Fortea, J., de Mendonça, A., Martins, M., … Sleegers, K. (2016). A comprehensive study of the genetic impact of rare variants in SORL1 in European early-onset Alzheimer's disease. *Acta Neuropathologica*, *132*(2), 213–224. doi: 10.1007/s00401-016-1566-9

medRxiv preprint doi: https://doi.org/10.1101/2023.07.13.23292622; this version posted July 16, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Xie, T., Xu, G., Liu, Y., Quade, B., Lin, W., & Bai, X. (2023). Structural insights into the assembly of the agrin/LRP4/MuSK signaling complex. *Proceedings of the National Academy of Sciences*, *120*(23), e2300453120. doi: 10.1073/pnas.2300453120