Transcriptomic similarities and differences between mouse models and human Alzheimer's Disease

Marco Antônio De Bastiani[1]*, Bruna Bellaver[1]*, Giovanna Carello-Collar[1], Stefania Forner[2], Alessandra Cadete Martini[3], Tharick A. Pascoal[4], Eduardo R. Zimmer[1,5,6]

[1]Graduate Program in Biological Sciences: Biochemistry, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil;

[2]Institute for Memory Impairments and Neurological Disorders (UCI MIND), University of California, Irvine, USA;

[3]Department of Pathology & Laboratory Medicine, University of California, Irvine, USA;

[4]Department of Neurology and Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA;

[5]Department of Pharmacology, UFRGS, Porto Alegre, Brazil;

[6]Graduate Program in Biological Sciences: Pharmacology and Therapeutics, UFRGS, Porto Alegre, Brazil.

*Both authors contributed equally to the article.

Marco Antônio De Bastiani, PhD. Department of Biochemistry, UFRGS, Porto Alegre, Brazil. Rua Ramiro Barcelos 2600, Porto Alegre, Brazil - 90035-003, Phone: +55 51 3308-5556; FAX: +55 51 3308-5535; e-mail: tyrev@hotmail.com

Abstract

Alzheimer's disease (AD) is a multifactorial pathology with the vast majority of cases having a sporadic presentation. Yet, preclinical research studies relied for decades on transgenic models overexpressing human genes found in AD autosomal dominant patients. Recently, knock-in (KI) models, such as the novel hAβ-KI mouse, were introduced; however, its translational potential is still being addressed. To unveil core molecular programs in AD, we explored hippocampal transcriptomic profiles of transgenic (5xFAD and APP/PS1) and KI (hAβ-KI) mouse models with early- (EOAD) and late- (LOAD) onset AD patients. Our analysis revealed that all three mouse models presented more Gene Ontology biological processes (GOBP) terms and enriched signaling pathways in common with LOAD than with EOAD subjects. Furthermore, semantic similarity of enriched GOBP terms highlighted model-specific biological alterations that might be further explored. Finally, we identified 17 transcription factor potentially acting as master regulators of AD in all models. Our cross-species exploratory analyses evidenced that the hAβ-KI mice presented a remarkable specificity to LOAD, which might support its use to advance our understanding of this AD subtype.

Keywords: Alzheimer's disease mouse models; humanized Aβ knock-in; APP/PS1; 5xFAD; bioinformatics.

Introduction

Alzheimer's disease (AD) is commonly categorized into early- (EOAD; <65) and late-onset (LOAD; ≥65) based on the age of clinical onset, the latter being the most prevalent form of the disease (Reitz *et al*, 2020). This broad and simplistic definition includes both the mendelian and non-mendelian EOAD, which confers an additional degree of heterogeneity to this group. However, both forms of EOAD seem to present a more prominent brain atrophy, glucose hypometabolism and increased tau PET uptake when compared to LOAD patients (Cho *et al*, 2017; McDade *et al*, 2018; Möller *et al*, 2013; Rabinovici *et al*, 2010; Schöll *et al*, 2017).

Although the autosomal dominant inheritance (mendelian EOAD) due to mutations in APP, PSEN1 and PSEN2 only accounts for ~10% of EOAD cases (Reitz *et al.*, 2020), the genetically driven early-onset forms of the disease represented, for many years, the main target of animal modeling. Indeed, the most used animal models of AD overexpress one or more genes associated with human autosomal dominant AD, such as APP and PSEN1. However, this approach has undergone extensive scrutiny in the past years (Götz *et al*, 2018; Saito *et al*, 2014) due to the fact that these overexpression models do not resemble many important aspects of LOAD. Thus, a major challenge in the field has been the development of new animal models that better recapitulate LOAD. The emergence of new human Aβ knock-in (KI) models, which may better resemble LOAD due to the lack of supra-physiological expression of human transgenes, may avoid some of the pitfalls that overexpression models presented (Baglietto-Vargas *et al*, 2021; Saito *et al.*, 2014; Serneels *et al*, 2020).

In this regard, comparisons between overexpression and KI models with human pathology at the molecular level may help understanding to what extent they can resemble human disease. Additionally, the ability to recapitulate key molecular pathways activated or repressed in AD is crucial for model validity. In this study, we aimed to ascertain the genes and pathways overlapping between two gene overexpression (5xFAD and APP/PS1) and one KI (hAβ-KI) mouse models, with EOAD and LOAD. With this in mind, we compared the hippocampal transcriptomic profiling of these animal models and AD subtypes, as well as their

molecular specificity to the human disease. Finally, we employed regulatory network-based approach to infer and investigate common master regulators (MR) between animal models and AD.

Results

Mouse models exhibit more DEGs overlapping with LOAD

DEA between hippocampal region of hAβ-KI, 5xFAD and APP/PS1 mice and their WT controls identified 1537, 3231 and 1768 DEGs, respectively (unadjusted p-value < 0.05; **Figure 2A, B; Supplemental Figure S1, Supplemental Table 2**). We observed that the hAβ-KI animals presented more DEGs overlapping with the 5xFAD mice than with APP/PS1 model [389 (25.3%) *versus* 235 (15.3%) DEGs, respectively; Chi-square adjusted p-value < 0.001; **Figure 2C**]. In addition, the comparison with AD human data demonstrated that hAβ-KI mice exhibited more DEGs in common with LOAD than with EOAD individuals (381 *versus* 164, respectively; **Figure 2D, E**). Despite being mouse models carrying AD familial mutations, APP/PS1 and 5xFAD mice shared more DEGs with LOAD than with EOAD patients (**Figure 2D, E**). Moreover, considering the total of DEGs identified in the mouse models as reference (model-disease overlap), the intersection of DEGs between 5xFAD and EOAD was significantly higher (14%) than the overlap between both hAβ-KI (10.7%; Chi-square adjusted p-value = 0.0054) and APP/PS1 (10.9%; Chi-square adjusted p-value = 0.0061) with EOAD (**Figure 2D - right**). On the other hand, the overlap of LOAD DEGs with 5xFAD (28%) and with APP/PS1 (25.3%) mice were not significantly different from the overlap with hAβ-KI (24.8 %; **Figure 2E - right**; Chi-square adjusted p-value = 0.123 and Chi-square adjusted p-value = 0.063, respectively). When we compared only the adjusted p-value DEGs [Benjamini & Hochberg (BH) > 0.1] we observed a similar profile/proportion of genes overlap, especially regarding to the hAβ-KI model and LOAD (**Supplemental Figure S2**). Thus, we opted for using unadjusted p-values < 0.05 for subsequent functional enrichment analysis, which allows for an exploratory cross-species analysis of core molecular programs that can be further

validated. Interestingly, the hAβ-KI model shared 212 DEGs exclusively with LOAD, while only 98 with EOAD (**Figure 2A, B**). The PPI network revealed that PTGES3, GNB1, ARIH2, SMURF1, EIF3A are the gene hubs of the four clusters formed according to protein interaction (**Figure 2F**). However, only two genes revealed in the PPI network, GNB1 and NKTR, were found differentially expressed in the hAβ-KI model after BH adjustment. Considering the DEGs exclusively shared with the hAβ-KI model and EOAD subjects, only 11 remained in the PPI network, and only the PRPF40A gene remained significant in the hAβ-KI after multiple comparisons correction (**Figure 2G**). Next, we compared DEGs of each mouse model with a database from MS patients to verify the specificity of these models for AD pathology. Both 5xFAD and APP/PS1 mice presented greater overlap of DEGs with AD than with MS subjects (**Figure 2H, I; Supplemental Figure S3A, B**). Interestingly, no significant differences were observed in DEGs shared between hAβ-KI mice and EOAD (10.7%) or MS patients (9.6%; **Figure 2J; Supplemental Figure S3C**), while the overlap of hAβ-KI DEGs with LOAD was significantly higher than with both EOAD and MS (24.8%; Chi-square adjusted p-value < 0.001; **Figure 2J**; **Supplemental Figure S3C**).

hAβ-KI mice and LOAD patients present higher functional changes similarities

Functional enrichment analysis of gene ontology for biological processes (GOBPs) revealed that ~ 92% of the GOBPs enriched in hAβ-KI mice overlap with enriched terms in LOAD patients (**Figure 3B, E - right**) in the model-disease approach, while the intersection with EOAD was markedly low (around 32%, **Figure 3A, D - right**). The remaining 8% of GOBP terms not shared between hAβ-KI mice and LOAD were related to RNA splicing and protein phosphorylation (**Supplemental Table 3**). Additionally, the hAβ-KI mice presented a higher number of enriched GOBP terms in common with the 5xFAD (79.2%) than with the APP/PS1 (56.1%) model (**Figure 3C**; Chi-square adjusted p-value < 0.001). Unexpectedly, both 5xFAD and APP/PS1 shared more than 75% of their enriched GOBPs with LOAD patients (**Figure 3B, E - right**), an overlap greater than the observed with EOAD (~ 45%) (**Figure 3A, D - right**). Considering the disease-model overlap, the 5xFAD mice was the only one

presenting significant higher intersection with both EOAD (55%) and LOAD (46.8%; **Figure 3D - left, E - left**). Specifically, it was observed about 50% of GOBPs overlap between EOAD and LOAD with 5xFAD, indicating a lack of specificity for this model. The comparison among the three models with MS GOBPs demonstrated, however, a great specificity of these mouse models for AD (**Supplemental Figure S3**). In fact, **Figure 3F-H** shows that all mouse models evaluated presented less than 5% of enriched GOBP terms in common with MS.

hAβ-KI mice present less GOBP terms intersection with EOAD patients in comparison to APP/PS1 and 5xFAD models

The union of enriched GOBP intersecting terms in the mouse models and human AD were generated to better visualize the common biological processes altered in each group. "Regulation of cytokine secretion" (13 nodes), "regulation of catabolic processes" (3 nodes), "NFκB signaling" (9 nodes) and "intracellular transport and secretion" (21 nodes) were among the GOBP terms enriched in the hippocampus of all three mouse models and EOAD patients (**Figure 4; light gray circles**). Of note, the majority of the GOBP terms in common with EOAD come from the 5xFAD and APP/PS1 models (102 nodes, **orange circles**). These terms are mainly related to cellular response to stressor agents, hormones and cytokines, immune response and calcium homeostasis and transport. Interestingly, GOBP terms related with oxidative phosphorylation found in EOAD only appeared enriched in the APP/PS1 mouse model (**purple circles**).

hAβ-KI mice and LOAD subjects exhibit more overlap among the enriched GOBP terms

Figure 5 shows that hAβ-KI mice GOBP terms are more represented among the GOBPs in common with LOAD (**yellow, purple and pink circles**) than with EOAD. "PI3K and NFκB signaling pathways" (20 nodes), "regulation of neuronal development" (19 nodes), "regulation of neurotransmitter and hormone secretion" (31 nodes), "immune response" (22 nodes), "antigen processing and presentation" (21 nodes) and "phospholipid and ribose phosphate metabolism" (10 nodes) are among the enriched biological processes shared with all

the three mouse models and LOAD (**pink circles**). Interestingly, biological processes as "glucose, cholesterol and purine metabolism", "regulation of MAPK cascade" and "calcium homeostasis" only appeared enriched in 5xFAD and APP/PS1 mice (**gray circles**).

The majority of the hAβ-KI enriched KEGGs is shared with LOAD

The most affected pathways related to changes in transcriptome profile were identified using enrichment analysis of KEGG canonical pathways. **Figure 6A, B** shows that among the 10 KEGGs found significantly enriched in the hAβ-KI model, six were also identified in LOAD subjects ["glutamatergic and GABAergic synapse" (adjusted p-value = 0.017), "calcium signaling" (adjusted p-value = 0.040), "Rap1 (adjusted p-value = 0.040) and Ras signaling (adjusted p-value = 0.040)" and "choline metabolism in cancer" (adjusted p-value = 0.041)], while only one was enriched in EOAD patients ["amyotrophic lateral sclerosis" (adjusted p-value = 0.040); **Supplemental Table 4**]. KEGG analysis also revealed that APP/PS1 was the mouse model that presented the highest percentage of pathway overlap with both EOAD (36.7%; **Figure 6C**) and LOAD (73.3%; **Figure 6D**). Interestingly, most of the intersection between APP/PS1 and human AD is represented by KEGGs related to other neurodegenerative diseases (e.g. Parkinson's disease, Huntington's disease) or bacterial infection (e.g. *Escherichia coli*, *Salmonella* spp.; **Supplemental Table 4**). Despite PI3K pathway-related GOBPs appeared enriched in LOAD, they were only significantly altered in the hAβ-KI model (adjusted p-value = 0.040) in the KEGG analysis. "Rap1 and Ras signaling pathways" and "adhesion and apoptosis" were found consistently altered in the APP/PS1 and hAβ-KI models (**Figure 6H, I; Supplemental Table 4**). In addition, the 5xFAD was the mouse model that presented more enriched KEGG terms (**Figure 6G; Supplemental Table 4**). Similar to the GOBP enrichment analysis, 5xFAD presented more pathways shared with LOAD (56.9%) than with EOAD (13.8%; **Figure 6C, D**), which were mainly related to synaptic neurotransmission and insulin regulation (**Supplemental Table 4**). Additionally, alterations in apoptosis- and endocannabinoid system-related pathways were observed in the 5xFAD model and in the human disease (**Supplemental Table 4**). Despite an overlap in KEGGs associated to

inflammatory diseases and bacterial infections between 5xFAD and AD, well-studied signaling pathways in the context of inflammation were only observed in the 5xFAD model (e.g. TNF, Toll-like, NFκB).

EOAD and LOAD patients share two enriched MR with AD mouse models

To identify elements located in higher positions of the biological system hierarchy, we performed an MRA to search for transcription factors potentially driving the biological alterations observed in AD. In addition, we asked if these elements were also driving the transcriptional profile changes in the mouse models. Our analysis revealed a total of 95 MR enriched with DEGs that were in at least one experimental group (**Figure 7A, B; Supplemental Table 5**). Interestingly, both 5xFAD and APP/PS1 mouse models presented more MR in common with LOAD than with EOAD, while hAβ-KI mice exhibited a similar overlap of MR with both subtypes of human AD (**Figure 7A, B**). Among the 17 MR identified in $\geq$ 4 groups, only PARK2 and SOX9 were enriched in all the three models and in the human disease (**Figure 7C**). Next, two-tail GSEA was performed to infer the activation state of each MR candidate. We observed that PARK2 and SOX9 are respectively repressed and activated across the disease/animal model phenotypes evaluated here (**Figure 7D**). In addition, FOXC2 and ZNF461 were identified exclusively in the hAβ-KI mice and LOAD individuals (**Supplemental Table 5)**. The three mouse models share eight enriched MR, most of which involved in the regulation of cell cycle and apoptosis (**Supplemental Table 5**). A different methodological approach to infer activation of transcription factors was also applied and similar results were observed for the 17 MR candidates (**Supplemental Figure S4**). Finally, a comparison with a previously published study that investigated MR associated with AD showed that 11 out of 17 MR identified here were also enriched in that study (**Figure 7E**).

Discussion

Great efforts were made in the last decade to generate mouse models that better recapitulate the sporadic characteristics of LOAD; however, their translational value is still

being addressed. Here, we evaluated hippocampal similarities and differences of three mouse models (APP/PS1, 5xFAD and hAβ-KI) at the transcriptional level. An exploratory cross-species comparative transcriptomics between these models with EOAD and LOAD subjects was also conducted. Overall, we observed that the hAβ-KI mice presented more similarities with 5xFAD compared to the APP/PS1 model. Interestingly, all mouse models showed more similarities to LOAD than to EOAD. Finally, the comparison with human AD pointed to a striking specific transcriptomic similarity of hAβ-KI with LOAD.

The specificity of the hAβ-KI model for LOAD was first observed in the exploratory DEG analysis, as the hAβ-KI mice presented twice more DEGs exclusively overlapping with LOAD than with EOAD subjects. Further PPI network analysis of these DEGs identified gene products that interact with each other to accomplish certain cellular functions. For example, we found clusters involved in ribosomal RNA processing, inflammatory response and E3-ubiquitin ligase-related immune response, all phenomena well described in AD (Ding *et al*, 2005; Gjoneska *et al*, 2015; Hegde *et al*, 2019). Corroborating, Baglietto-Vargas and colleagues demonstrated that the hAβ-KI mouse presented a decreased production of the anti-inflammatory cytokines IL-2 and IL-10, compared to age-matched WT animals (Baglietto-Vargas *et al.*, 2021). Importantly, a cluster of genes encoding G protein and G protein-coupled receptors (GPCRs) was also evidenced in our study. G proteins act as modulators or transducers in various transmembrane signaling systems and GPCRs are implicated in the pathogenesis of AD, in multiple stages of APP processing, and can be directly or indirectly impacted by Aβ (Thathiah & De Strooper, 2011). Indeed, primary rat neuronal cultures and brain slices treated with a metabotropic, but not ionotropic, glutamate receptor (mGluR) agonist stimulated APP secretion, suggesting that G protein signaling cascade in response to mGluRs activation is linked to APP processing by α-secretase (Lee *et al*, 1995). Interestingly, this process seems to be dependent of the PI3K signaling activation (Ledonne *et al*, 2015), a pathway significantly altered in the hAβ-KI mice. More informative than evaluate single genes or elements, network-based approaches are powerful strategies to identify elements, pathways or functional modules

altered in several diseases (Basu *et al*, 2021; Santiago *et al*, 2017); the clusters identified here might shed light in the understanding of pathogenic mechanisms of LOAD that can be recapitulated by the novel hAβ-KI mouse model, facilitating the search for therapeutic targets.

The replication of transcriptomic changes at the pathway level is more consistent than single gene analysis, and it has the potential to offer insights into the biological processes disturbed in AD. In this sense, we observed that the hAβ-KI mouse presented almost a complete overlap of enriched hippocampal GOBP terms with LOAD, while only about one third was shared with EOAD. This apparent specificity for LOAD seems to be a unique and important feature of this novel KI model, as the 5xFAD and APP/PS1 mice did not discern between AD subtypes. Despite of that, the scanty overlap of GOBP terms among AD mouse models and MS confirmed that these models present transcriptomic features specific to AD rather than general alterations shared among other neurodegenerative diseases. In addition, this resemblance with AD appears to be a characteristic specifically of rodent models carrying human mutated genes directly related to AD, such as APP and PSEN1, or humanized Aβ. Indeed, Burns and colleagues observed that Tg4510 mice, that resemble tau pathology but also harbor mutations found in familial frontotemporal dementia, presented highest enrichment of genes in common with human amyotrophic lateral sclerosis (ALS) and Huntington's disease rather than with AD (Burns *et al*, 2015).

Compromised calcium (Ca^{2+}) handling by neurons precedes the formation of Aβ plaques and neurofibrillary tangles and has been implicated in virtually all the major molecular alterations underlying the pathogenesis of AD (Tong *et al*, 2018). The association between PS1 mutations and altered Ca^{2+} signaling in neurons has been known for many years (Tong *et al*, 2016). However, the lack of rodent models that resemble more than the AD hallmarks (i.e. Aβ load and cognitive impairment) has been a major limitation in the field (2017). In addition, early imbalance between excitatory/inhibitory (E/I) neurotransmission, with a loss of neuronal network stability, is a well-attested phenomenon in AD (Frere & Slutsky, 2018; Palop &

Mucke, 2016). The average incidence of seizure in AD patients is around 15%, which is 7-fold higher when compared to individuals without dementia (Horváth *et al*, 2016; Imfeld *et al*, 2013). Aβ is able to impair the long-term potentiation of synaptic strength, promote depression of synaptic activity and alter brain network (Palop & Mucke, 2016), and affect the E/I balance by impairing the inhibitory activity of the parvalbumin-expressing interneurons (Verret *et al*, 2012). Interestingly, the modulation of interneuron function might improve brain rhythms and cognitive functions in AD (Tong *et al*, 2014; Verret *et al.*, 2012). Our study identified "calcium signaling" and "glutamatergic and GABAergic signaling" to be exclusively altered in the hAβ-KI mice and LOAD, pointing this KI mouse model as an important tool to better understand calcium signaling and the E/I imbalance in AD.

Transcription factors play key roles orchestrating phenotypic determination through the regulation of numerous transcriptional targets that coordinate complex cellular processes (Carro *et al*, 2010; Vargas *et al*, 2018). Based on reverse engineering co-expression regulatory network reconstruction, we identified two transcription factors exclusively altered in the hAβ-KI model and LOAD individuals' hippocampi: ZNF461 and FOXC2. ZNF461 role in brain function is still poorly explored; however, this transcription factor was identified among genes that represent a polygenetic risk for psychiatric disorders and its alteration might contribute to cortical atrophy and changes in functional connectivity (Lee *et al*, 2017). Accordingly, changes in cortical thickness show strong correlation with the clinical stages of AD (Ossenkoppele *et al*, 2019). In the same way, a decreased functional connectivity is not only positively correlated with disease severity in autosomal dominant AD and LOAD patients (Thomas *et al*, 2014), but also predicts cognitive decline in asymptomatic preclinical AD individuals (Buckley *et al*, 2017). On the other hand, FOXC2 function in the brain is implicated in cell proliferation and invasion in glioblastoma (Li *et al*, 2013), in angiogenic processes during fetal brain development (Siegenthaler *et al*, 2013), and is directly regulated by cyclin-dependent kinase 5 (Cdk5) phosphorylation to control peripheral lymphatic vase development

(Liebl *et al*, 2015). Interestingly, it has been demonstrated that the deregulation of Cdk5 contributes to AD pathology preceding tau hyperphosphorylation and loss of synaptic proteins (Kurbatskaya *et al*, 2016). On the other hand, Saito and colleagues reported that the App^{NL-F} KI model did not generate p25, necessary for aberrant Cdk5 activation, suggesting that the reports about its relevance in AD might be an artifact due to the overexpression of APP in other mouse models (Saito *et al*, 2016). However, there are several studies using organotypic hippocampal slices and primary neural cells exposed to Aβ show increased p25 generation independently of APP overexpression (Ma *et al*, 2013; Seo *et al*, 2014; Zheng *et al*, 2005), suggesting that experimental validation of FOXC2 and Cdk5 in hAβ-KI mice is needed to understand the potential link among FOXC2, Cdk5 and AD pathology.

Further, in the MRA, we identified SOX9 as a transcription factor consistently activated in the mouse models and human AD. SOX9 is a key factor in the nervous system development, especially for astrocyte and oligodendrocyte cell fate specification (Molofsky *et al*, 2013; Nagao *et al*, 2016; Stolt *et al*, 2003). Recently, Sun and colleagues identified SOX9 as an astrocyte-specific nuclear marker in the adult human and mouse brains, presenting a remarkable expression in murine hippocampus and cortex, when compared to the cerebellum (Sun *et al*, 2017). In addition, the upregulation of SOX9 in mouse models of ALS, spinal cord injury and stroke was already reported in the literature (McKillop *et al*, 2013; McKillop *et al*, 2016; Sun *et al.*, 2017). However, there are no studies to date linking this transcription factor with AD, and our results highlight a promising new target for investigation in AD pathophysiology. In addition, PARK2 gene was also identified in the MRA as a transcription factor consistently repressed among all groups of this study. PARK2 encodes an E3 ubiquitin ligase and its involvement in autosomal recessive parkinsonism is well established. Although less explored, its role in AD has already been demonstrated by computational (Kumar & Kumar, 2019; Vargas *et al.*, 2018) and experimental (Goiran *et al*, 2018; Martín-Maestro *et al*, 2019; Martín-Maestro *et al*, 2016) approaches. Specifically, mitophagy failure, promoted by a repression in PARK2 ability to stimulate PS1, was evidenced in cellular and animal models

(Goiran *et al.,* 2018; Martín-Maestro *et al.,* 2019). On the other hand, the overexpression of PARK2 promoted diminished ubiquitinated proteins accumulation, improved its targeting to mitochondria and potentiated autophagic vesicle synthesis (Martín-Maestro *et al.,* 2016). Additionally, a meta-analysis from genome-wide association studies identified PARK2 as a potential risk gene for AD. Our findings highlight the value of hAβ-KI, 5xFAD and APP/PS1 mouse models to better understand these particularly underexplored aspects in AD.

Decades of use of animal models in AD research underline that each one is able to mimic a slightly different aspect of the disease. Therefore, one could argue that animal models of AD should be elected according to the biological aspect aimed for investigation, rather than be seen as a generic model. The clusterization by semantic similarity of enriched GOBP terms performed here allowed to recognize biological alterations specific of each animal mouse model. For example, several GOBP terms associated to oxidative phosphorylation, purine metabolism and the MAPK pathway were altered in 5xFAD and APP/PS1 mice but not in the hAβ-KI model. On the other hand, inflammatory-related processes seem to be a feature of AD pathology present in all three mouse models. Thus, the comparison of altered biological processes among mouse models and human AD presented here might help to guide future methodological decisions regarding the most appropriate animal model to answer a specific research question. Finally, the transcription factors potentially acting as master regulators of AD observed in this study need to be further explored and validated experimentally, having the potential to greatly aid future research and mechanistic understanding of the disease.

Methods

Mouse Models Data Acquisition

RNA sequencing (RNA-seq) data from 5xFAD [4, 8 and 12 months-old, n = 23 hemizygous; 26 wild-type (WT)] and hAβ-KI (22 months-old, n = 7 homozygous; 8 WT) AD mouse models (**Supplemental Table 1**) were obtained from AMP-AD Knowledge Portal (https://adknowledgeportal.synapse.org/) using synapser (version 0.7.64) and synapserutils

(version 0.1.6) packages. Specifically, gene expression information was collected from https://www.synapse.org/#!Synapse:syn16798173 and https://www.synapse.org/#!Synapse:syn18634479 for 5xFAD and hAβ-KI models, respectively. APP/PS1 (8 and 12 months-old, n = 8 APP/PS1; 8 WT) mouse model RNA-seq data (**Supplemental Table 1**) was combined from two Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) datasets [GSE149661 (Unger *et al*, 2020) and GSE145907] and downloaded through NCBI Sequence Read Archive using SRAToolKit (https://github.com/ncbi/sra-tools). After quality control evaluation, the following samples were removed from further analyses: sample "67-2" from hAβ-KI; samples "466", "305", "456", "497" from 5xFAD.

Known phenotypic features of the three mouse models evaluated in this study are depicted in **Figure 1A**.

Human Data Acquisition

Human AD hippocampal processed microarray data of five studies were obtained from GEO repository [GSE28146 (Blalock *et al*, 2011), GSE29378 (Miller *et al*, 2013), GSE36980 (Hokama *et al*, 2014), GSE48350 (Berchtold *et al*, 2008) and GSE84422 (Wang *et al*, 2016)], downloaded using GEOquery package (v2.56.0) (Davis & Meltzer, 2007) and combined under common gene symbol annotations. Afterwards, batch correction was implemented using the sva package (v3.36.0) (Leek *et al*, 2012) and data was split into EOAD (age at death < 65, n = 4 EOAD; 10 control) and LOAD (age at death ≥ 65, n = 59 LOAD; 63 control) for further analyses (**Supplemental Table 1**). AD diagnosis, control definition and exclusion criteria of each GSE study are depicted in **Figure 1B**. Multiple sclerosis (n = 5 MS; 5 control) hippocampal RNA-seq data (**Supplemental Table 1**) was also obtained from GEO under the identifier GSE123496 (Voskuhl *et al*, 2019) and downloaded through NCBI Sequence Read Archive. Sample demographics can be found in **Figure 1C**.

RNA-seq Processing

Raw data for each RNA-seq dataset was downloaded and transcript alignment was performed using Salmon (v1.3.0) (Patro *et al*, 2017). Transcripts were mapped to genome using indexes derived from *Mus musculus* GRCm38 Ensembl build (ftp://ftp.ensembl.org/pub/release-96/fasta/mus_musculus) and *Homo sapiens* GRCH38 Ensembl build (ftp://ftp.ensembl.org/pub/release-96/fasta/homo_sapiens) for the mouse models and human data, respectively. Aligned reads were summarized using tximport (v1.12.3) (Soneson *et al*, 2015) and genes with minimum mean of counts per million cut-off < 2 were filtered out.

Differential Expression Analyses

Differential expression was computed on processed microarray data using the limma package (Ritchie *et al*, 2015) *lmFit* function to fit multiple linear models by generalized least squares. In addition, *eBayes* function was used to compute moderated t-statistics, moderated F-statistic and log-odds of differential expression by empirical Bayes moderation of the standard errors towards a common value. For RNA-seq datasets, processed expression data from each study was submitted to DESeq2 (v1.28.1) (Love *et al*, 2014) method using previously created tximport Summarized Experiment. Differential expression analysis (DEA) based on the Negative Binomial distribution was computed with the *DESeq* function followed by log fold change shrinkage with the *lfcShrink* function (shrinkage estimator type = "ashr"). Genes with unadjusted p-value < 0.05 were considered as differentially expressed genes (DEGs). For information about BH adjusted p-values, see **Supplemental Table 2**.

Venn diagrams for DEGs were constructed using VennDiagram package (v1.6.20) (Chen & Boutros, 2011). The proportion differences between overlapped genes between models and diseases were computed by Pearson's Chi−squared test with Yates' continuity correction for count data, followed by post-hoc pairwise Bonferroni adjustment.

Protein-Protein Interaction Network Reconstruction

We used the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) Consortium to build protein-protein interaction (PPI) networks. STRING is a biological database and web resource of known and predicted PPI which contains information from numerous sources, including experimental data, computational prediction methods and public text collections. The construction of the PPI networks was implemented in R using the STRINGdb (v2.0.2), RedeR (v1.36.0) and igraph (v1.2.6) packages (Castro *et al*, 2012; Mora & Donaldson, 2011; Szklarczyk *et al*, 2019). For the final networks, we retained only the edges with a combined interaction score > 0.7 from all sources and highly connected nodes for the final networks.

Functional Enrichment Analyses

DEGs (unadjusted p-value < 0.05) from human or mouse model studies were submitted to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis using the clusterProfiler package (v3.16.1) *enrichKEGG* and *enrichGO* functions. The GO terms were clustered by semantic similarity using the *mgoSim* function from GOSemSim (v2.14.2) package (Yu *et al*, 2010) (arguments measure = "Wang" and combine = NULL). The resulting similarity matrices were represented as GO networks using the RedeR (v1.36.0) package (Castro *et al.*, 2012) for interactive visualization and manipulation of nested networks. Clusters of GO terms obtained from GOSemSim algorithm were manually named for their biological interpretation. Venn diagrams for enriched GO/KEGG terms were constructed using VennDiagram (v1.6.20) package (Chen & Boutros, 2011). Nested networks were constructed by maintaining only the intersecting GOBP terms among the mouse models and each human pathology (either EOAD or LOAD). Finally, Jaccard coefficient > 0.7 was used to filter out edges with low gene intersection between terms.

The proportion differences between overlapped terms between models and diseases were computed by Pearson's Chi−squared test with Yates' continuity correction for count data, followed by post-hoc pairwise Bonferroni adjustment.

Reverse Engineering of Transcriptional Network

The transcriptional network (TN) centered on transcription factors (TF) and their predicted target genes were inferred using a large cohort of neurologically and neuropathologically normal individuals (n = 122) obtained from GEO under the identifier GSE60862 (Trabzuni *et al*, 2013). Herein, the terms "regulatory unit" or "regulon" are used to describe the groups of inferred genes and their associated TFs. RTN (v2.12.1) package was used to reconstruct and analyze TNs based on the mutual information (MI) using the Algorithm for the Reconstruction of Accurate Cellular Networks (ARACNe) method (Fletcher *et al*, 2013; Margolin *et al*, 2006a; Margolin *et al*, 2006b). In summary, the regulatory structure of the network is derived by mapping significant associations between known TFs and all potential targets. To create a consensus bootstrap network, the interactions below a minimum MI threshold are eliminated by a permutation step and unstable interactions are additionally removed by bootstrap. Finally, data processing inequality algorithm is applied with null tolerance to eliminate interactions that are likely to be mediated by a third TF. The reference hippocampus TN was built using the package's default number of 1000 permutations and 100 bootstraps (p-value < 0.001).

Master Regulators Inference and Two-Tailed Gene Set Enrichment Analysis

Master regulator (MR) analysis (MRA) was employed for the MR inference (Carro *et al.,* 2010). MRA computes the statistical overrepresentation of DEGs (p-value < 0.05) obtained from DEA in the regulatory units of the reference TN. The regulons were considered altered in the disease if they presented (1) statistical enrichment of DEGs, (2) regulon size > 50 and (3) $\geq$ 80% of the queried case-control studies. Two-tailed Gene Set Enrichment Analysis (GSEA) was also performed using the RTN package (version 2.4.6, p-value < 0.05 and 1000 permutations). Briefly, Pearson's correlation was used to split the regulatory units into positively (A) and negatively (B) associated targets. Afterwards, the phenotype association of each subgroup was tested using the GSEA (Subramanian *et al*, 2005) statistics, resulting in

independent enrichment scores (ES) for each subgroup. Finally, we tested the differential enrichment (ESA – ESB) considering the following desirable criteria for clear association: (1) a maximum deviation from zero near opposite extremes and (2) a good separation of the two distributions. Thus, a high negative differential score implies that the regulon is repressed in the disorder phenotype, while a high positive one indicates that the regulon is induced.

Virtual Inference of Protein Activity by Enriched Regulon Analysis

The virtual inference of protein activity by enriched regulon analysis (VIPER) is another regulatory-network based approach to infer protein activity from gene expression profiles. Similar to MRA, VIPER systematically analyze the expression of the regulatory units previously identified by ARACNe algorithm. However, VIPER uses a fully probabilistic, yet efficient enrichment analysis framework based on analytic rank-based enrichment analysis. The analysis was implemented using the viper (v1.22.0) package in R (Alvarez *et al*, 2016).

Acknowledgments

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.synapse.org/). The IU/JAX/UCI MODEL-AD Center was established with funding from The National Institute on Aging (U54 AG054345-01 and AG054349). Aging studies are also supported by the Nathan Shock Center of Excellence in the Basic Biology of Aging (NIH P30 AG0380770). ERZ receives financial support from CNPq [435642/2018-9] and [312410/2018- 2], Instituto Serrapilheira [Serra-1912-31365], Brazilian National Institute of Science and Technology in Excitotoxicity and Neuroprotection [465671/2014-4], FAPERGS/MS/CNPq/SESRS–PPSUS [30786.434.24734.23112017] and ARD/FAPERGS [54392.632.30451.05032021]. MAB receives financial support from CNPq PDJ [150293/2019-4]. BB receives financial support from CAPES [88887.336490/2019-00].

Author Contributions

MAB selected the databases and performed transcriptomics analysis. MAB, BB and GCC analyzed and interpreted the data and conceived figures and tables. MAB and BB

wrote the manuscript. ERZ conceived and supervised the project. ERZ, GCC, TAP, SF and ACM edited and revised the manuscript for intellectual content.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

(2017) Calcium Hypothesis of Alzheimer's disease and brain aging: A framework for integrating new evidence into a comprehensive theory of pathogenesis. *Alzheimers Dement* 13: 178-182.e117

Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Ye BH, Califano A (2016) Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet* 48: 838-847

Baglietto-Vargas D, Forner S, Cai L, Martini AC, Trujillo-Estrada L, Swarup V, Nguyen MMT, Do Huynh K, Javonillo DI, Tran KM *et al* (2021) Generation of a humanized Aβ expressing mouse demonstrating aspects of Alzheimer's disease-like pathology. *Nat Commun* 12: 2421

Basu A, Ash PE, Wolozin B, Emili A (2021) Protein Interaction Network Biology in Neuroscience. *Proteomics* 21: e1900311

Berchtold NC, Cribbs DH, Coleman PD, Rogers J, Head E, Kim R, Beach T, Miller C, Troncoso J, Trojanowski JQ *et al* (2008) Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proc Natl Acad Sci U S A* 105: 15605-15610

Blalock EM, Buechel HM, Popovic J, Geddes JW, Landfield PW (2011) Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. *J Chem Neuroanat* 42: 118-126

Braak H, Braak E (1991) Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* 82: 239-259

Buckley RF, Schultz AP, Hedden T, Papp KV, Hanseeuw BJ, Marshall G, Sepulcre J, Smith EE, Rentz DM, Johnson KA *et al* (2017) Functional network integrity presages cognitive decline in preclinical Alzheimer disease. *Neurology* 89: 29-37

Burns TC, Li MD, Mehta S, Awad AJ, Morgan AA (2015) Mouse models rarely mimic the transcriptome of human neurodegenerative diseases: A systematic bioinformatics-based critique of preclinical models. *Eur J Pharmacol* 759: 101-117

Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY, Sulman EP, Anne SL, Doetsch F, Colman H *et al* (2010) The transcriptional network for mesenchymal transformation of brain tumours. *Nature* 463: 318-325

Castro MA, Wang X, Fletcher MN, Meyer KB, Markowetz F (2012) RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. *Genome Biol* 13: R29

Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12: 35

Cho H, Choi JY, Lee SH, Lee JH, Choi YC, Ryu YH, Lee MS, Lyoo CH (2017) Excessive tau accumulation in the parieto-occipital cortex characterizes early-onset Alzheimer's disease. *Neurobiol Aging* 53: 103-111

Davis S, Meltzer PS (2007) GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 23: 1846-1847

Ding Q, Markesbery WR, Chen Q, Li F, Keller JN (2005) Ribosome dysfunction is an early event in Alzheimer's disease. *J Neurosci* 25: 9171-9175

Fletcher MN, Castro MA, Wang X, de Santiago I, O'Reilly M, Chin SF, Rueda OM, Caldas C, Ponder BA, Markowetz F *et al* (2013) Master regulators of FGFR2 signalling and breast cancer risk. *Nat Commun* 4: 2464

Frere S, Slutsky I (2018) Alzheimer's Disease: From Firing Instability to Homeostasis Network Collapse. *Neuron* 97: 32-58

Gahete MD, Rubio A, Durán-Prado M, Avila J, Luque RM, Castaño JP (2010) Expression of Somatostatin, cortistatin, and their receptors, as well as dopamine receptors, but not of neprilysin, are reduced in the temporal lobe of Alzheimer's disease patients. *J Alzheimers Dis* 20: 465-475

Gjoneska E, Pfenning AR, Mathys H, Quon G, Kundaje A, Tsai LH, Kellis M (2015) Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* 518: 365-369

Goiran T, Duplan E, Chami M, Bourgeois A, El Manaa W, Rouland L, Dunys J, Lauritzen I, You H, Stambolic V *et al* (2018) β-Amyloid Precursor Protein Intracellular Domain Controls Mitochondrial Function by Modulating Phosphatase and Tensin Homolog-Induced Kinase 1 Transcription in Cells and in Alzheimer Mice Models. *Biol Psychiatry* 83: 416-427

Götz J, Bodea LG, Goedert M (2018) Rodent models for Alzheimer disease. *Nat Rev Neurosci* 19: 583-598

Hegde AN, Smith SG, Duke LM, Pourquoi A, Vaz S (2019) Perturbations of Ubiquitin-Proteasome-Mediated Proteolysis in Aging and Alzheimer's Disease. *Front Aging Neurosci* 11: 324

Hokama M, Oka S, Leon J, Ninomiya T, Honda H, Sasaki K, Iwaki T, Ohara T, Sasaki T, LaFerla FM *et al* (2014) Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama study. *Cereb Cortex* 24: 2476-2488

Horváth A, Szűcs A, Barcs G, Noebels JL, Kamondi A (2016) Epileptic Seizures in Alzheimer Disease: A Review. *Alzheimer Dis Assoc Disord* 30: 186-192

Hyman BT (1997) The neuropathological diagnosis of Alzheimer's disease: clinical-pathological studies. *Neurobiol Aging* 18: S27-32

Imfeld P, Bodmer M, Schuerch M, Jick SS, Meier CR (2013) Seizures in patients with Alzheimer's disease or vascular dementia: a population-based nested case-control analysis. *Epilepsia* 54: 700-707

Kumar D, Kumar P (2019) Aβ, Tau, and α-Synuclein aggregation and integrated role of PARK2 in the regulation and clearance of toxic peptides. *Neuropeptides* 78: 101971

Kurbatskaya K, Phillips EC, Croft CL, Dentoni G, Hughes MM, Wade MA, Al-Sarraj S, Troakes C, O'Neill MJ, Perez-Nievas BG *et al* (2016) Upregulation of calpain activity precedes tau phosphorylation and loss of synaptic proteins in Alzheimer's disease brain. *Acta Neuropathol Commun* 4: 34

Ledonne A, Nobili A, Latagliata EC, Cavallucci V, Guatteo E, Puglisi-Allegra S, D'Amelio M, Mercuri NB (2015) Neuregulin 1 signalling modulates mGluR1 function in mesencephalic dopaminergic neurons. *Mol Psychiatry* 20: 959-973

Lee A, Shen M, Qiu A (2017) Psychiatric polygenic risk associates with cortical morphology and functional organization in aging. *Transl Psychiatry* 7: 1276

Lee RK, Wurtman RJ, Cox AJ, Nitsch RM (1995) Amyloid precursor protein processing is stimulated by metabotropic glutamate receptors. *Proc Natl Acad Sci U S A* 92: 8083-8087

Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28: 882-883

Li W, Fu X, Liu R, Wu C, Bai J, Xu Y, Zhao Y, Xu Y (2013) FOXC2 often overexpressed in glioblastoma enhances proliferation and invasion in glioblastoma cells. *Oncol Res* 21: 111-120

Liebl J, Zhang S, Moser M, Agalarov Y, Demir CS, Hager B, Bibb JA, Adams RH, Kiefer F, Miura N *et al* (2015) Cdk5 controls lymphatic vessel development and function by phosphorylation of Foxc2. *Nat Commun* 6: 7274

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550

Ma Y, Bao J, Zhao X, Shen H, Lv J, Ma S, Zhang X, Li Z, Wang S, Wang Q *et al* (2013) Activated cyclin-dependent kinase 5 promotes microglial phagocytosis of fibrillar β-amyloid by up-regulating lipoprotein lipase expression. *Mol Cell Proteomics* 12: 2833-2844

Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, Califano A (2006a) ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 7 Suppl 1: S7

Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A (2006b) Reverse engineering cellular networks. *Nat Protoc* 1: 662-671

Martín-Maestro P, Gargini R, García E, Simón D, Avila J, García-Escudero V (2019) Mitophagy Failure in APP and Tau Overexpression Model of Alzheimer's Disease. *J Alzheimers Dis* 70: 525-540

Martín-Maestro P, Gargini R, Perry G, Avila J, García-Escudero V (2016) PARK2 enhancement is able to compensate mitophagy alterations found in sporadic Alzheimer's disease. *Hum Mol Genet* 25: 792-806

Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, Menon M, He L, Abdurrob F, Jiang X *et al* (2019) Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* 570: 332-337

McDade E, Wang G, Gordon BA, Hassenstab J, Benzinger TLS, Buckles V, Fagan AM, Holtzman DM, Cairns NJ, Goate AM *et al* (2018) Longitudinal cognitive and biomarker changes in dominantly inherited Alzheimer disease. *Neurology* 91: e1295-e1306

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (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

McKillop WM, Dragan M, Schedl A, Brown A (2013) Conditional Sox9 ablation reduces chondroitin sulfate proteoglycan levels and improves motor function following spinal cord injury. *Glia* 61: 164-177

McKillop WM, York EM, Rubinger L, Liu T, Ossowski NM, Xu K, Hryciw T, Brown A (2016) Conditional Sox9 ablation improves locomotor recovery after spinal cord injury by increasing reactive sprouting. *Exp Neurol* 283: 1-15

Miller JA, Woltjer RL, Goodenbour JM, Horvath S, Geschwind DH (2013) Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med* 5: 48

Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, Vogel FS, Hughes JP, van Belle G, 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

Möller C, Vrenken H, Jiskoot L, Versteeg A, Barkhof F, Scheltens P, van der Flier WM (2013) Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. *Neurobiol Aging* 34: 2014-2022

Molofsky AV, Glasgow SM, Chaboub LS, Tsai HH, Murnen AT, Kelley KW, Fancy SP, Yuen TJ, Madireddy L, Baranzini S *et al* (2013) Expression profiling of Aldh1l1-precursors in the developing spinal cord reveals glial lineage-specific genes and direct Sox9-Nfe2l1 interactions. *Glia* 61: 1518-1532

Mora A, Donaldson IM (2011) iRefR: an R package to manipulate the iRefIndex consolidated protein interaction database. *BMC Bioinformatics* 12: 455

Nagao M, Ogata T, Sawada Y, Gotoh Y (2016) Zbtb20 promotes astrocytogenesis during neocortical development. *Nat Commun* 7: 11102

Nilsson P, Sörgjerd K, Kakiya N, Sasaguri H, Watamura N, Shimozawa M, Tsubuki S, Zhou Z, Loera-Valencia R, Takamura R *et al* (2020) Somatostatin receptor subtypes 1 and 4 redundantly regulate neprilysin, the major amyloid-beta degrading enzyme in brain. *bioRxiv*: 2020.2005.2009.085795

Ossenkoppele R, Smith R, Ohlsson T, Strandberg O, Mattsson N, Insel PS, Palmqvist S, Hansson O (2019) Associations between tau, Aβ, and cortical thickness with cognition in Alzheimer disease. *Neurology* 92: e601-e612

Palop JJ, Mucke L (2016) Network abnormalities and interneuron dysfunction in Alzheimer disease. *Nat Rev Neurosci* 17: 777-792

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14: 417-419

Rabinovici GD, Furst AJ, Alkalay A, Racine CA, O'Neil JP, Janabi M, Baker SL, Agarwal N, Bonasera SJ, Mormino EC *et al* (2010) Increased metabolic vulnerability in early-onset Alzheimer's disease is not related to amyloid burden. *Brain* 133: 512-528

Reitz C, Rogaeva E, Beecham GW (2020) Late-onset vs nonmendelian early-onset Alzheimer disease: A distinction without a difference? *Neurol Genet* 6: e512

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47

Saito T, Matsuba Y, Mihira N, Takano J, Nilsson P, Itohara S, Iwata N, Saido TC (2014) Single App knock-in mouse models of Alzheimer's disease. *Nat Neurosci* 17: 661-663

Saito T, Matsuba Y, Yamazaki N, Hashimoto S, Saido TC (2016) Calpain Activation in Alzheimer's Model Mice Is an Artifact of APP and Presenilin Overexpression. *J Neurosci* 36: 9933-9936

Santiago JA, Bottero V, Potashkin JA (2017) Dissecting the Molecular Mechanisms of Neurodegenerative Diseases through Network Biology. *Front Aging Neurosci* 9: 166

Schöll M, Ossenkoppele R, Strandberg O, Palmqvist S, Jögi J, Ohlsson T, Smith R, Hansson O (2017) Distinct 18F-AV-1451 tau PET retention patterns in early- and late-onset Alzheimer's disease. *Brain* 140: 2286-2294

Seo J, Giusti-Rodríguez P, Zhou Y, Rudenko A, Cho S, Ota KT, Park C, Patzke H, Madabhushi R, Pan L *et al* (2014) Activity-dependent p25 generation regulates synaptic plasticity and Aβ-induced cognitive impairment. *Cell* 157: 486-498

Serneels L, T'Syen D, Perez-Benito L, Theys T, Holt MG, De Strooper B (2020) Modeling the β-secretase cleavage site and humanizing amyloid-beta precursor protein in rat and mouse to study Alzheimer's disease. *Mol Neurodegener* 15: 60

Siegenthaler JA, Choe Y, Patterson KP, Hsieh I, Li D, Jaminet SC, Daneman R, Kume T, Huang EJ, Pleasure SJ (2013) Foxc1 is required by pericytes during fetal brain angiogenesis. *Biol Open* 2: 647-659

Soneson C, Love MI, Robinson MD (2015) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* 4: 1521

Stolt CC, Lommes P, Sock E, Chaboissier MC, Schedl A, Wegner M (2003) The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes Dev* 17: 1677-1689

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-15550

Sun W, Cornwell A, Li J, Peng S, Osorio MJ, Aalling N, Wang S, Benraiss A, Lou N, Goldman SA *et al* (2017) SOX9 Is an Astrocyte-Specific Nuclear Marker in the Adult Brain Outside the Neurogenic Regions. *J Neurosci* 37: 4493-4507

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P *et al* (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47: D607-d613

Thathiah A, De Strooper B (2011) The role of G protein-coupled receptors in the pathology of Alzheimer's disease. *Nat Rev Neurosci* 12: 73-87

Thomas JB, Brier MR, Bateman RJ, Snyder AZ, Benzinger TL, Xiong C, Raichle M, Holtzman DM, Sperling RA, Mayeux R *et al* (2014) Functional connectivity in autosomal dominant and late-onset Alzheimer disease. *JAMA Neurol* 71: 1111-1122

Tong BC, Lee CS, Cheng WH, Lai KO, Foskett JK, Cheung KH (2016) Familial Alzheimer's disease-associated presenilin 1 mutants promote γ-secretase cleavage of STIM1 to impair store-operated Ca2+ entry. *Sci Signal* 9: ra89

Tong BC, Wu AJ, Li M, Cheung KH (2018) Calcium signaling in Alzheimer's disease & therapies. *Biochim Biophys Acta Mol Cell Res* 1865: 1745-1760

Tong LM, Djukic B, Arnold C, Gillespie AK, Yoon SY, Wang MM, Zhang O, Knoferle J, Rubenstein JL, Alvarez-Buylla A *et al* (2014) Inhibitory interneuron progenitor transplantation restores normal learning and memory in ApoE4 knock-in mice without or with Aβ accumulation. *J Neurosci* 34: 9506-9515

Trabzuni D, Ramasamy A, Imran S, Walker R, Smith C, Weale ME, Hardy J, Ryten M (2013) Widespread sex differences in gene expression and splicing in the adult human brain. *Nat Commun* 4: 2771

Unger MS, Li E, Scharnagl L, Poupardin R, Altendorfer B, Mrowetz H, Hutter-Paier B, Weiger TM, Heneka MT, Attems J *et al* (2020) CD8(+) T-cells infiltrate Alzheimer's disease brains and

regulate neuronal- and synapse-related gene expression in APP-PS1 transgenic mice. *Brain Behav Immun* 89: 67-86

Vargas DM, De Bastiani MA, Zimmer ER, Klamt F (2018) Alzheimer's disease master regulators analysis: search for potential molecular targets and drug repositioning candidates. *Alzheimers Res Ther* 10: 59

Verret L, Mann EO, Hang GB, Barth AM, Cobos I, Ho K, Devidze N, Masliah E, Kreitzer AC, Mody I *et al* (2012) Inhibitory interneuron deficit links altered network activity and cognitive dysfunction in Alzheimer model. *Cell* 149: 708-721

Voskuhl RR, Itoh N, Tassoni A, Matsukawa MA, Ren E, Tse V, Jang E, Suen TT, Itoh Y (2019) Gene expression in oligodendrocytes during remyelination reveals cholesterol homeostasis as a therapeutic target in multiple sclerosis. *Proc Natl Acad Sci U S A* 116: 10130-10139

Wang H, Muiznieks LD, Ghosh P, Williams D, Solarski M, Fang A, Ruiz-Riquelme A, Pomès R, Watts JC, Chakrabartty A *et al* (2017) Somatostatin binds to the human amyloid β peptide and favors the formation of distinct oligomers. *Elife* 6

Wang M, Roussos P, McKenzie A, Zhou X, Kajiwara Y, Brennand KJ, De Luca GC, Crary JF, Casaccia P, Buxbaum JD *et al* (2016) Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med* 8: 104

Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S (2010) GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* 26: 976-978

Zheng YL, Kesavapany S, Gravell M, Hamilton RS, Schubert M, Amin N, Albers W, Grant P, Pant HC (2005) A Cdk5 inhibitory peptide reduces tau hyperphosphorylation and apoptosis in neurons. *Embo j* 24: 209-220

Figure legends

Figure 1 | Phenotypic features of AD mouse models, AD diagnosis, control group definition, and demographics of human samples. APP/PS1 (top), 5xFAD (middle) and hAβ-KI mice (bottom) timelines showing specific phenotypes of each mouse model across different ages (in months (**A**). The numbers in bold correspond to the time points evaluated in this study. Data for APP/PS1 and 5xFAD mice was available at AlzForum. hAβ-KI information was taken from Baglietto-Vargas et al (Baglietto-Vargas *et al.*, 2021). The subjects were diagnosed with AD according to established clinical criteria and/or neuropathological *post-mortem* evaluation in each selected GSE study **(B)**. Age, sample size and sex information are specified for EOAD, LOAD and MS subjects and their age-matched controls **(C)**. NINCDS-ADRDA 1984 (McKhann *et al*, 1984); CERAD (Mirra *et al*, 1991); NIA-Reagan Institute criteria (Hyman, 1997); Braak staging (Braak & Braak, 1991); NFT = neurofibrillary tangles; MMSE = Mini-Mental State Examination; DLB = Dementia with Lewy-Bodies; PD = Parkinson's disease; MS = Multiple Sclerosis. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease.

Figure 2 | Shared DEGs among EOAD, LOAD and mouse models of AD. Venn diagram showing DEGs overlap between EOAD **(A)** and LOAD **(B)** with 5xFAD, APP/PS1 and hAβ-KI mice. Mosaic plot of hAβ-KI, 5xFAD and APP/PS1 overlap **(C)**. Mosaic plot of EOAD-model (left) and model-EOAD (right) DEGs overlap with 5xFAD, APP/PS1 and hAβ-KI mice **(D)**. Mosaic plot of LOAD-model (left) and model-LOAD (right) DEGs overlap with 5xFAD, APP/PS1 and hAβ-KI mice **(E)**. PPI network of DEGs exclusively shared between hAβ-KI mice and LOAD **(F)** or EOAD **(G)** subjects. Mosaic plot of EOAD, LOAD or MS DEGs overlap with 5xFAD mice **(H)**. Mosaic plot of EOAD, LOAD or MS DEGs overlap with APP/PS1 mice **(I)**. Mosaic plot of EOAD, LOAD or MS DEGs overlap with hAβ-KI mice **(J)**. The size of the red and yellow boxes reflects the proportion of overlapping and non-overlapping DEGs, respectively. Pearson's Chi−squared test with Yates' continuity correction was applied for the

mosaic plot analysis. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in. Genes with unadjusted p-value < 0.05 were considered as DEGs.

Figure 3 | **GOBP overlap among EOAD, LOAD and mouse models of AD.** Venn diagram of GOBPs intersections in EOAD **(A)** and LOAD **(B)** with 5xFAD, APP/PS1 and hAβ-KI mice. Mosaic plot of hAβ-KI-5xFAD and -APP/PS1 GOBP overlap **(C)**. Mosaic plot of EOAD-model (left) and model-EOAD (right) GOBP overlap with 5xFAD, APP/PS1 and hAβ-KI mice **(D)**. Mosaic plot of LOAD-model (left) and model-LOAD (right) GOBP overlap with 5xFAD, APP/ PS1 and hAβ-KI mice **(E)**. Mosaic plot of EOAD, LOAD or MS GOBP overlap with 5xFAD mice **(F)**. Mosaic plot of EOAD, LOAD or MS GOBP overlap with APP/PS1 mice **(G)**. Mosaic plot of EOAD, LOAD or MS GOBP overlap with hAβ-KI mice **(H)**. The size of the red and yellow boxes reflects the proportion of overlapping and non-overlapping GOBPs, respectively. Pearson's Chi−squared test with Yates' continuity correction was applied for the mosaic plot analysis. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in.

Figure 4 | **Nested networks of enriched GOBP intersections between EOAD and AD mouse models.** Blue, purple and orange circles represent the overlap between EOAD and protein overexpression mouse models. Intersection among EOAD and hAβ-KI are represented by circles in shades of gray. Each box represents a cluster of GOBP terms grouped by semantic similarity and named manually according to its main biological role. Circle sizes represent the number of genes enriched in the GO term. EOAD = early-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in.

Figure 5 | **Nested networks of enriched GOBP intersections between LOAD and AD mouse models.** Light purple, dark purple and green circles represent all the GOBP terms overlap between LOAD and hAβ-KI. Intersection among LOAD and overexpression protein mouse models are represented by circles in shades of gray. Each box represents a cluster of GOBP terms grouped by semantic similarity and named manually according to its main biological role.

Circle sizes represent the number of genes enriched in the GO term. LOAD = late-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in.

Figure 6 | **Functional enrichment analysis of KEGG terms in mouse models of AD and human disease.** Venn diagrams showing the KEGGs overlap among EOAD **(A)** and LOAD **(B)** subjects with 5xFAD, APP/PS1 and hAβ-KI mice. Mosaic plot of 5xFAD, APP/PS1 and hAβ-KI KEGGs overlap with EOAD **(C)** or LOAD **(D)**. Dual-plots showing the up- (red) and down- (blue) regulated KEGGs for EOAD **(E)**, LOAD **(F)**, 5xFAD **(G)**, APP/PS1 **(H)** and hAβ-KI **(I)**. The size of the red and yellow boxes reflects the proportion of overlapping and non-overlapping KEGGs, respectively. Pearson's Chi−squared test with Yates' continuity correction was applied for the mosaic plot analysis. The pathways in the dual-plots are represented in the Y axis; the left and the right part of the X axis corresponds to the number of genes enriched in the KEGG terms and their q-value (grey bars). EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in.

Figure 7 | **Master regulator analysis of mouse models of AD and human disease.** Venn diagrams of EOAD **(A)** or LOAD **(B)** with 5xFAD, APP/PS1 and hAβ-KI MRs overlap. Tile-plot of regulons significantly enriched with DEGs in EOAD, LOAD, 5xFAD, APP/PS1 and hAβ-KI (only regulatory units significant in ≥ 4 contexts are represented) **(C)**. Tile-plot showing the activation state of the transcription factors acting as MR in EOAD, LOAD, 5xFAD, APP/PS1 and hAβ-KI **(D)**. Venn diagram showing the overlap between previously MR candidates found by Vargas et al. and the current study **(E)**. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; hAβ-KI = human amyloid-β knock-in; ns = non-significant.

Supplemental Figure Legends

Figure S1 | **DEGs in mouse models of AD and human disease.** Volcano-plots and pie charts showing the up- and down-regulated genes in 5xFAD **(A)**, APP/PS1 **(B)**, hAβ-KI **(C)**, EOAD **(D)**, LOAD **(E)** and MS subjects **(F)**. The red points in the volcano-plots correspond to a DEG when compared to its control. In the pie charts, the pink and blue colors correspond respectively to the up- and down-regulated DEGs. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in. Genes with unadjusted p-value < 0.05 were considered as DEGs.

Figure S2 | Intersection of DEGs between animal models and AD subtypes. Mosaic plot of EOAD-model **(A)** and model-EOAD **(B)** DEGs overlap with 5xFAD, APP/PS1 and hAβ-KI mice. Genes with *BH adjusted p-value < 0.1* were considered as DEGs in the mosaic plots.

Figure S3 | **Shared DEGs and enriched GOBP among mouse models of AD and human disease.** Venn diagrams showing the DEGs (left) and GOBP terms (right) overlap among 5xFAD **(A, D)**, APP/PS1 **(B, E)** and hAβ-KI **(C, F)** mice with EOAD, LOAD and MS subjects. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; MS = multiple sclerosis; hAβ-KI = human amyloid-β knock-in. Genes with unadjusted p-value < 0.05 were selected as DEGs.

Figure S4 | **VIPER analysis of EOAD, LOAD and AD mouse models.** Plot of inferred regulon activity showing the up- (redish) and down- (blueish) regulated transcription factors in each group: EOAD **(A)**, LOAD **(B)**, 5xFAD **(C)**, APP/PS1 **(D)** and hAβ-KI **(E)**. EOAD = early-onset Alzheimer's disease; LOAD = late-onset Alzheimer's disease; hAβ-KI = human amyloid-β knock-in. *p-value < 0.05

A)

Plaques | Tangles | Neuronal loss | Gliosis
Synaptic loss | Changes in LTP/LTD | Cognitive impairment

B)

AD diagnosis and control definition			
GSE	AD diagnosis	Control definition	Exclusion criteria
28146	NINCDS-ADRDA + CERAD + NIA-Reagan or NFT count + Braak staging + amyloid plaque count	MMSE > 25	Not specified
29378	NINCDS-ADRDA + Braak staging + amyloid plaque count	Normal cognitive and functional examination	Not specified
36980	CERAD + Braak staging	No dementia	Not specified
84422	CERAD + Braak staging	Not specified	Not specified
48350	CERAD + NIA-Reagan + Braak staging	Cognitively and neurologically normal	DLB, PD, Binswanger's disease, MS, intracerebral hemorrhage

C)

Demographics of human samples						
	Control	EOAD	Control	LOAD	Control	MS
Age (median)	46	61	83	83	Unknown	Unknown
Sample size	10	4	63	59	5	5
Sex (% of females)	50	75	38	47	100	100

DIFFERENTIALLY EXPRESSED GENES

GENE ONTOLOGY BIOLOGICAL PROCESS