Activity dependent transcription pathways and its implication in Alzheimer's Disease

Universitat Oberta de Catalunya

Xavier Fernandez Olalla

MU Bioinf. i Bioest.

Área de trabajo final

Tutora de TF:

Yolanda Guillén Montalbán

Profesor responsable de la asignatura:

David Merino Arranz

14/01/2023

LIBERTAS PERFVNDET

OMNIA LVCE

UNIVERSITAT DE BARCELONA

1. Table of contents

2. List of abbreviations

Aβ Amyloid beta

AD Alzheimer's disease

Cc Cell culture

DEG Differentially-expressed genes

DGE Differential gene expression

GO Gene ontology

Hs *Homo sapiens*

IEGs Immediate early genes

KCl Potassium chloride

LTP Long term potentiation

Mm *Mus musculus*

NT Neurotransmitters

PEA Pathway enrichment analysis

PMID PubMed Identifier

RNA Ribonucleic acid

TFs Transcription factors

TTX Tetrodotoxin

WT Wild type

3. Abstract

Neuronal activity dependent pathways and synaptic plasticity are essential to memory formation and learning processes, with a huge implication of immediate early genes. Alterations on those pathways and genes are related with a wide number of mental disorders and cognitive impairment, including Alzheimer's disease (AD). Since there is no effective treatment for AD, I propose to analyze activity dependent genes as possible target for therapeutical approach for AD. At the end of this work, from the analysis and comparison of RNA-sequencing of activated neuronal cultures, AD mice model brain, and human AD patients' samples, different activity dependent genes have been found to be altered in AD and could be a potential target for future therapeutic approach.

4. Introduction

a. Brain and memory processes

The brain is in charge of receiving environmental and internal body stimuli, adapting, and responding by promoting a response and establishing a memory that can be stored and retrieved later. All of those processes are carried out by the complex network of interactions between neurons and non-neuronal cells in neural circuits. The processes of learning and memory are based on changes in the synapses (interactions between neurons), known as synaptic plasticity.

Santiago Ramón y Cajal hypothesized that changes in strength and growth of connections between neurons supported cognitive processes in the brain (Ramón y Cajal, 1894). Half century later, Donald Hebb upheld his idea postulating that the synaptic efficiency between presynaptic and postsynaptic neurons should be strengthened upon activity (Hebb, 1950).

To understand the synaptic plasticity mechanism, first we need to know what is a synapse. The Oxford English Dictionary describes synapse as "*a structure that permits a neuron (or nerve cell) to pass an electrical or chemical signal to another neuron or to the target effector cell*". In the brain, the synapses are formed between 2 neurons. The information is transmitted unidirectionally from the pre-synaptic neuron's axon to the post-synaptic neuron's dendrite.

When the pre-synaptic neuron is activated (figure 1a), it generates an action potential that depolarizes the membrane of the presynaptic axon, what induces the Ca^{2+} voltage-gated channels activation and the entrance of Ca^{2+} (figure 1b), triggering the liberation of the neurotransmitters (NT) to the synaptic cleft (figure 1c). There, the liberated NT bind to their specific receptors in the postsynaptic membrane and induce downstream changes in the postsynaptic neuron (figure 1d). The changes that occur both in the post-synaptic neurons and in the presynaptic neuron that have an effect on the synapsis performance are the called synaptic plasticity.

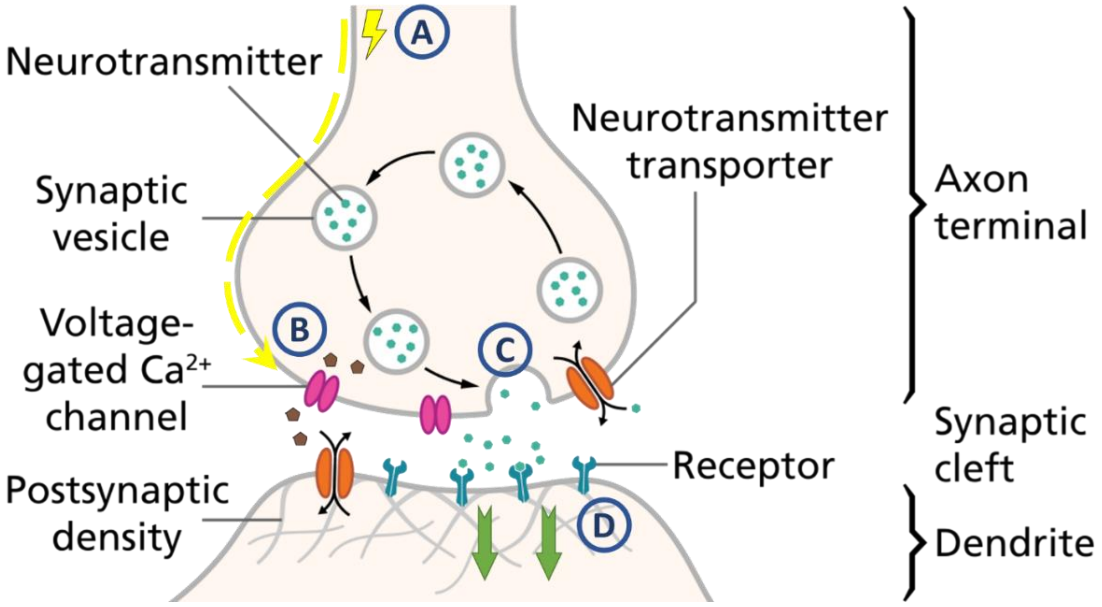

Figure 1. Neuron to neuron synapsis process scheme. Different parts of the synapse process are illustrated: A) Action potential from the soma to the axonal presynaptic membrane. B) Depolarization of presynaptic membrane and activation of calcium channel and consequent calcium entrance to the neuron. C) Liberation of neurotransmitters to the synaptic cleft due to the increasing intracellular calcium levels. D) Activation of postsynaptic receptors by

neurotransmitters and downstream effects on postsynaptic neuron. Image modified from "SynapseSchematic" by Thomas Splettstoesser is licensed under CC BY 4.0.

There are two different types of modifications that can occur in synaptic plasticity: short-term and long-term synaptic plasticity. In short-term synaptic plasticity, the liberation of NT induces local changes in both pre-synaptic and post-synaptic neurons that modify the next incoming signals. Those local changes are transients and quick (from milliseconds to few minutes) and do not need de novo protein synthesis or degradation (Zucker & Regehr, 2002). On the other hand, in long-term synaptic plasticity the changes promoted in the synapses are more complex and require transcriptional and protein synthesis changes but are more stable in time (from hours to years) and are the base of memory and learning processes.

One of the most important synaptic plasticity processes in learning and memory formation is long-term potentiation (LTP). In LTP process, in a similar way than in other synaptic plasticity changes, an input signal in the post-synaptic neuron promotes the overexpression, translation and export of AMPA receptors to the post-synaptic membrane in order to strengthen the connection between the pre-synaptic and post-synaptic neurons. In this process, the next signals in this synapse will have the same presynaptic input but the postsynaptic output will be significantly increased (Harris, 2020).

The current theory of LTP mechanism hold that LTP develops through at least three phases, including initial short-term potentiation (STP), early LTP (E-LTP), and late LTP (L-LTP) (Roberson et al., 1996). STP and E-LTP are transient and consist in rapid mechanisms such as an increase of receptor release from vesicles to synaptic membranes. In contrast, L-LTP requires additional changes like structural modifications to the formation of new synapses and remodeling and grow of preexisting synapses, what requires new gene expression and protein synthesis (Bailey et al., 2015). So, long-term synaptic plasticity and memory require de novo gene expression and protein synthesis that are regulated by the called **Activity dependent pathways**.

b. Activity dependent pathways

Transcription and translation inhibition impair learning and memory. This indicates that de novo gene expression and protein synthesis are essential for memory processes (Kandel, 2001). Extracellular stimuli can induce a rapid and robust change of the transcription of specific genes, known as immediate early genes (IEGs) and are cell specific.

Here, we are interested in those IEGs whose regulation depends on specific stimuli of synaptic activity. In the synapsis, the activity of NT promotes changes in the cytoplasmic Ca^{2+} levels that regulates certain signaling pathways involved in the expression of different neuronal activity-regulated transcription factors (TFs) such as CREB, SRF or MeCP2 among others (Flavell & Greenberg, 2008).

The proto-oncogene *c-fos* was the first IEG reported to be regulated by neuronal activity. Later, several genes have been reported to be regulated by neuronal activity. Many of the activity dependent genes codify for TFs that could intervene in many roles that are indispensable for learning, memory and cognitive processes, like neuronal survival, dendritic and axonal growth, synaptic development, and neuronal plasticity (Greer & Greenberg, 2008).

Disruption of activity dependent pathways have been linked with many cognitive disorders (Figure 2) including the first stages of Alzheimer's disease (AD) development (de Haan et al., 2012), what indicate the huge importance of those pathways for brain and cognitive processes.

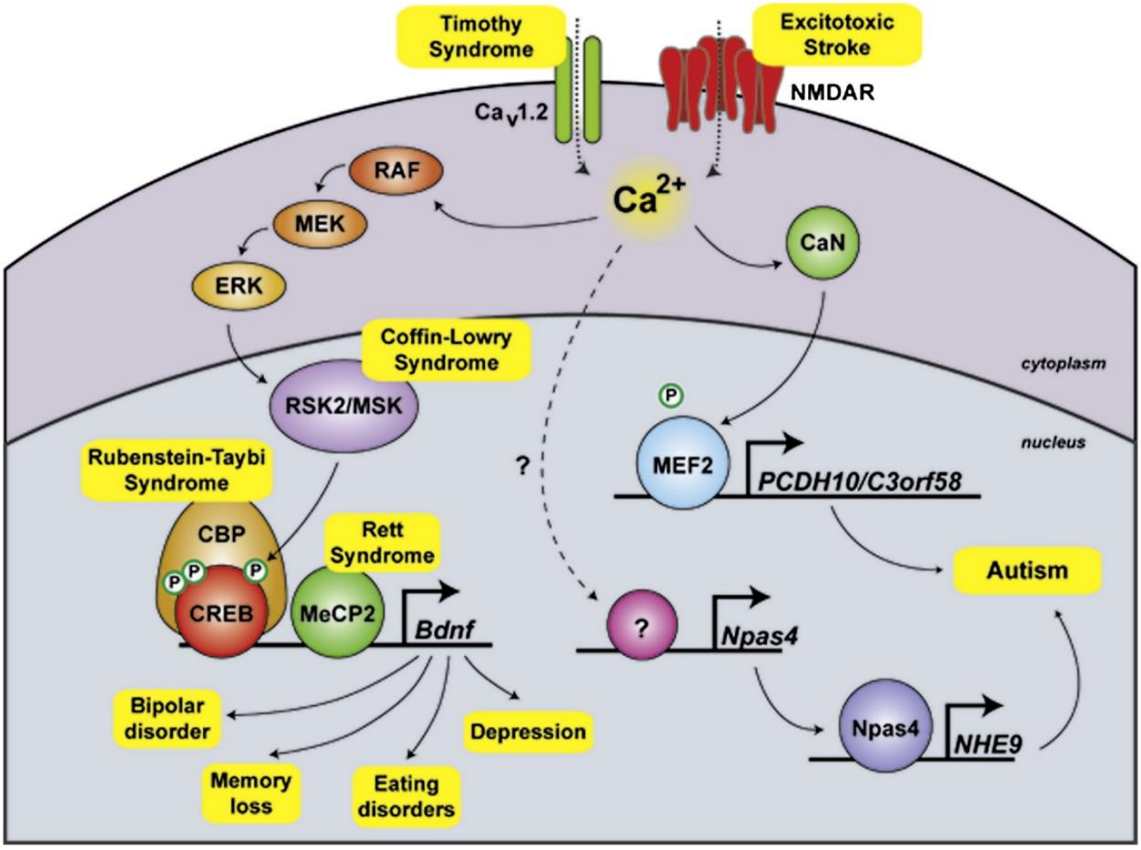

Figure 2. Mutation of Components of the Activity-Dependent Gene Expression Pathways Results in Human Cognitive Disorders. Figure from "From Synapse to Nucleus: Calcium-Dependent Gene Transcription in the Control of Synapse Development and Function. Greer P, Greenberg M. Neuron (2008) 59(6) 846-860"

c. Alzheimer's disease

Alzheimer's disease (AD), the most common cause of dementia worldwide, is characterized by memory loss and progressive cognition impairment that disable the autonomy of patients. Currently, 50 million people worldwide live with dementia and this number will rise to 152 million by 2050, what will provoke an enormous health, economic and social burden (World Health Organization, 2022).

AD development is spatiotemporally described. The first affected region is the entorhinal cortex and its alterations start years before symptoms onset. After that, the disease expands to the limbic cortex correlating with the appearance of first mild cognitive symptoms. Finally, the pathology appears over the isocortex: almost all brain structures are affected and damaged, the cognitive impairment aggravates and culminates with dementia (Palmer, 2011).

The different histopathological hallmarks of AD are neuroinflammation, synaptic pathways disruption (specially the cholinergic pathway), intracellular processes alteration, neuronal death, disruption of cerebral vascularity, and protein misfolding, especially amyloid-beta (Aβ) accumulation in plaques and hyperphosphorylated tau protein aggregation in neurofibrillary tangles (Figure 3).

Nowadays, there is no effective treatment to avoid, stop or ameliorate AD. Current pharmacological treatments are based on Acetylcholinesterase inhibitors for maintaining the cholinergic levels that are lost due to the cholinergic pathways alteration and ameliorate AD symptoms. However, they are not enough to reduce the disease progression.

In the latest decades, scientific community has focused on trying to target main candidate that underlies the disease, the amyloid plaques. Different pharmacological approaches against those proteins, like Aducanumab (Sevigny et al., 2016), have failed in stopping or ameliorating AD (Walsh et al., 2021), so different approaches are necessary to develop successful treatments against AD.

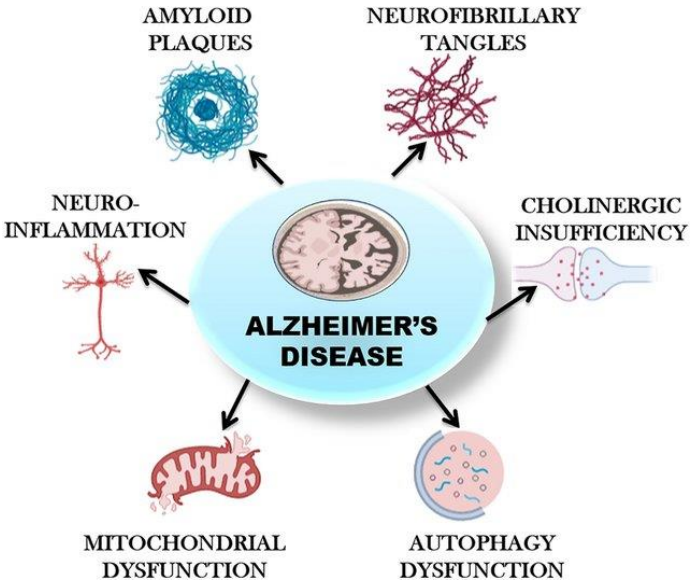

Figure 3. Histopathological hallmarks of Alzheimer's Disease. Figure from "Recent advances in molecular pathways and therapeutic implications targeting neuroinflammation for Alzheimer's disease. Dhapola et al., 2021

It has been described that neuronal-activity-dependent pathways and memory circuits are altered in asymptomatic and first stages of AD (de Haan et al., 2012; Sperling et al., 2010). At this point, several genomics, transcriptomics and proteomics analyses of AD and activity dependent pathways have been performed (Guerreiro et al., 2013; Patel et al., 2019), but the heterogeneity of the data and the results makes hard to define the transcriptional mechanisms associated with AD. The aim of this work is to analyze transcriptional data of AD samples from different biological sources (human and mice brain tissues, and cell lines) to find common activity dependent genes altered in AD.

5. Hypothesis

The hypothesis of this master thesis is that Alzheimer's disease is associated with certain gene expression alterations involved in activity dependent pathways.

6. Objectives

The objectives of this master thesis are the following:

1. To find out neuronal activity genes using stimulated neuronal cultures
2. To study the altered gene expression induced by Alzheimer's disease in patients' brain samples and in AD mice models brains.

3. To sum up the information of the objectives 1 and 2 to uncover the activity dependent genes altered in Alzheimer's disease

7. Methods

a. Analysis design

The aim of the work is to put together data from different public sources, which have been analyzed independently, and reanalyze them using the same computational pipeline to compare the results.

There are multiple methods to process transcriptional data, which makes it difficult to compare the results obtained from independent studies. Furthermore, there are different versions of the same genome, which can difficult the results comparison. To solve that, raw data of published experiments will be curated with the same conditions and mapped to the latest genomes versions. Differential gene expression analysis will be performed using the same statistical analysis algorithm. Then, a metanalysis of all the results will be made to get a general overview of activity dependent pathways, and their potential implication in AD.

As we are interested in gene expression and activity-dependent transcription pathways, RNA-sequencing data from AD human patients, AD mice models and activated neuron cultures will be used for the analysis.

Table 1. Samples data information. hs: Homo sapiens, mm: Mus musculus, cc: cell culture, PMID: PubMed Identifier, AD: Alzheimer's Disease, WT: Wild Type, TTX: Tetrodoxin, KCl: Potassium Chloride.

	Human AD samples 1 (hs1)	Human AD samples 2 (hs2)	Mouse AD model 1 (mm1)	Mouse AD model 2 (mm2)	Human neuronal culture 1 (cc1)	Human neuronal culture 2 (cc2)
Study PMID	36224601	32989324	32049030	34654824	32992958	33542524
Type of sample	Occipital lobe from human patients	Lateral temporal lobe from human patients	Mice Hippocampus	Mice Hippocampus	Human SH-SY5Y neuroblastoma culture neuronally diferentiated	Cultured hiPSC-derived GABAergic neurons
Number of samples	6 Control and 6 AD patients	6 Control and 6 AD patients	6 WT and 6 AD models	6 WT and 6 AD models	3 Controls and 3 stimulated	4 Controls and 4 stimulated
Type of disease	Late Onset AD	Late Onset AD	-	-	-	-
Model mutations	-	-	J20 (APPSw,Ind)	5xFAD (APPSw,Flor,Lnd, PSEN1^{M146L,L286V})	-	-
Treatment	-	-	-	-	KCl 100 mM 2h	TTX 1 uM O/N + KCl 53 mM 2h
Sequencer	Illumina HiSeq 4000	Illumina NextSeq 500	Illumina HiSeq 2500	Illumina NextSeq 500	Illumina NextSeq 500	Illumina NextSeq 500
Type of Raw data	Paired-end reads	Single-end reads	Paired-end reads	Paried-end reads	Single-end reads	Single-end reads

b. Data collection

The most similar studies with the most homogeneous information possible were selected, and the maximum number of samples were reanalyzed.

Due to the computational resources limit, 2 studies of human AD patients (Caldwell et al., 2022; Nativio et al., 2020), 2 studies of AD mice models (Castanho et al., 2020; Forner et al., 2021) and 2 studies of human neuronal cultures (Boulting et al., 2021; Kiltschewskij & Cairns, 2020) were selected. Data information is shown in table 1.

The characteristics of the samples "intra-group" are not exactly the same (Table 1), and it has to be taken into account. ENA browser was used to download the raw data of the different studies.

i. Quality control

FastQCR (Kassambara, 2019) R package was used to analyze de quality of the reads of the samples. All samples, except mm2, were uploaded by the authors after the QC (Supplementary information 1). Fastp (S. Chen et al., 2018) tool for Linux Bash was used to filter mm2 raw reads. During the filtering, adapters, polyG and polyA tails were trimmed and bases with less than 40% of bases with lower quality score than threshold of 20 were kept, using the following code:

```
fastp -i R1.fastq.gz -o R1.trimmed.fastq.gz -I R2.fastq.gz -O
R2.trimmed.fastq.gz -g -W 5 -q 20 -u 40 -x -3
```

c. Mapping and read counting

Curated reads were mapped to indexed genomes (hg38 and mm39), using the R package Rsubread (Liao et al., 2019). Genomes were full single-block indexed with a threshold of 100 repetitions of 16mers (sequences of 16 bp). Once the genomes were indexed, fastq files were aligned against their respective genomes (hg38, for human AD samples and neuronal cultures, and mm39 for AD mice model samples), and sorted. R code used is available in Supplementary information 2.

Sorted BAM files were used for gene counts analysis, obtaining a gene counts table for each sample and study.

d. Differential gene expression (DGE) analysis

Differential expression analysis was performed using Deseq2 (Love et al., 2014) R package that uses the non-normalized read counts due to it uses the average expression strength of each gene, across all samples, as its filter criterion, and it omits all genes with mean normalized counts below a filtering threshold from multiple testing adjustment (Love et al., 2014). Comparisons were performed per groups:

- Stimulated neurons vs Non stimulated (control) in neuronal cultures datasets.
- AD samples vs Control samples in human brain tissues datasets.
- AD Mutants vs WT control samples in mice hippocampal datasets.

Differentially-expressed genes (DEG) tables were obtained.

e. Biological analysis of DGEs

DEGs were analyzed using different R package: VennDiagram (H. Chen, 2022) and EnhancedVolcano (Blighe et al., 2021) to analyze DEGs by diagrams and volcano-plots, respectively. ClusterProfiler (Wu et al., 2021) was used to perform gene ontology (GO).

8. Results and discussion

a) Neuron activity by KCl-induced membrane depolarization enhances IEGs transcriptional up-regulation.

From the mapping and gene counting of the two neuronal culture studies, we obtained two different lists of DGE:

- 5101 DEGs from cc1 dataset. In this case, there is a homogeneous distribution where half of DEGs is up-regulated and the other half are downregulated (2701 up-regulated vs 2397 down-regulated). (Figure 4A).
- 1388 DEG from cc2 dataset. Here, similarly than in cc1 dataset, there is a homogeneous distribution in regulation (785 up-regulated vs 603 down-regulated), although some of the regulated genes have higher statistically strength in the different expression (Figure 4B).

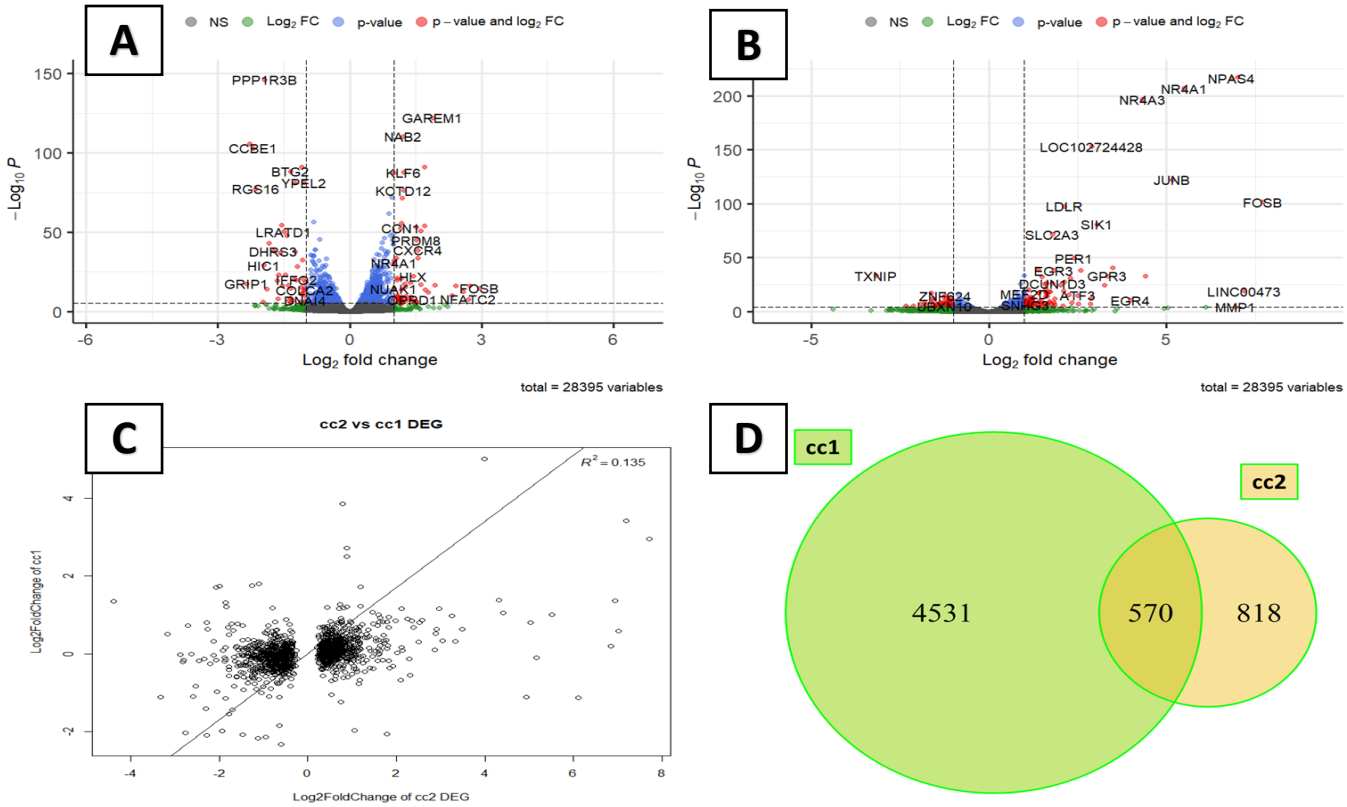

Figure 4. A) Volcanoplot of DEG Log2FoldChange of cc1 dataset. B) Volcanoplot of DEG Log2FoldChange of cc2 dataset. C) Log2Foldchange of cc2 DEG comparing with cc1 DEG. D)Ven diagram of DEG from neuronal cultures. In green (cc1) are represented the DEG of the neuronal culture study 1. In orange (cc2) the DEG of the neuronal culture 2.

When the Log2FoldChange of DEGs from cc2 dataset is compared with the Log2FoldChange of the same genes of the cc1 dataset, no correlation has been found (Figure 4C). This means that genes do not change its expression in the same way when neurons are stimulated in two independent experiments with different conditions.

When comparing the DEGs of both datasets, there are 570 DEGs in common (Figure 4D).

Once had the genes differentially expressed in those datasets, the most known IEGs were checked to assess if KCl depolarization of this neuronal cultures change the expression of already described IEGs.

IEGs list (de Haan et al., 2012; Flavell & Greenberg, 2008; Patel et al., 2019; Tong et al., 2001; Yan et al., 2021):

- FOS
- JUN
- EGR-1
- FOSB
- JUNB
- NR4A1
- ZENK
- ARC
- ATF3
- C-MYC
- EGR2
- FRA-1
- FRA-2
- JUND
- MYC
- NPAS4
- PKG-1
- PKG-2
- EGR3
- EGR4
- FOSL1
- FOSL2

Of those 22 IEGs:

- 10 IEGs were in cc1 DEGs
- 11 IEGs were in cc2 DEGs
- 8 IEGs in common DEGs

Table 2 and 3. IEGs differentially expressed IEGs in neuronal cultures, the log2FoldChange of their expression, and the p-adjusted valor of the change in expression.

cc1 IEGs

SYMBOL	log2FoldChange	padj
FOSB	2,96	2,12E-15
MYC	1,52	5,92E-54
EGR2	1,47	5,78E-06
NR4A1	1,01	2,67E-31
FOSL2	0,98	8,44E-46
EGR3	0,95	3,55E-05
JUND	0,68	8,37E-05
FOS	0,58	2,25E-05
ARC	0,52	0,015566571
JUN	0,46	1,66E-08

cc2 IEGs

SYMBOL	log2FoldChange	padj
FOSB	7,71	3,67E-102
FOS	7,01	4,11E-218
NPAS4	6,84	4,70E-217
NR4A1	5,50	7,68E-208
JUNB	5,16	7,14E-123
EGR4	3,98	3,68E-11
FOSL2	3,25	4,26E-25
ARC	2,89	5,22E-13
ATF3	2,49	8,96E-16
EGR3	1,83	6,59E-39
JUND	0,84	4,63E-08

All of those IEGs in cc1 and cc2, were up-regulated when neurons were depolarized with KCl, as shown by the positive values of log2FoldChange in tables 2 and 3.

b) Neuron depolarization by KCl enhances transcriptional changes related with transcriptional regulation.

The gene ontology analysis of the DEGs of neuronal cultures, reveal that in both neuronal cultures, the stimulation of neurons with KCl regulates processes related with DNA transcription regulation (Figure 5):

- cc1: transcription coregulator activity, DNA-binding transcription factor binding, protein transferase activity, catalytic activity acting on DNA, transcription coactivator activity, RNA polymerase II-specific DNA-binding transcription factor binding, among others (Figure 5A).
- cc2: DNA-binding transcription repressor activity, RNA polymerase II-specific, DNA-binding transcription repressor activity, DNA-binding transcription activator activity, DNA-binding transcription activator activity, RNA polymerase II-specific and transcription coregulator activity (Figure 5B).
- 507 common DEGs between cc1 and cc2: DNA-binding transcription activator activity, DNA-binding transcription activator activity, RNA polymerase II-specific, transcription coregulator activity, DNA-binding transcription factor binding … (Figure 5C)

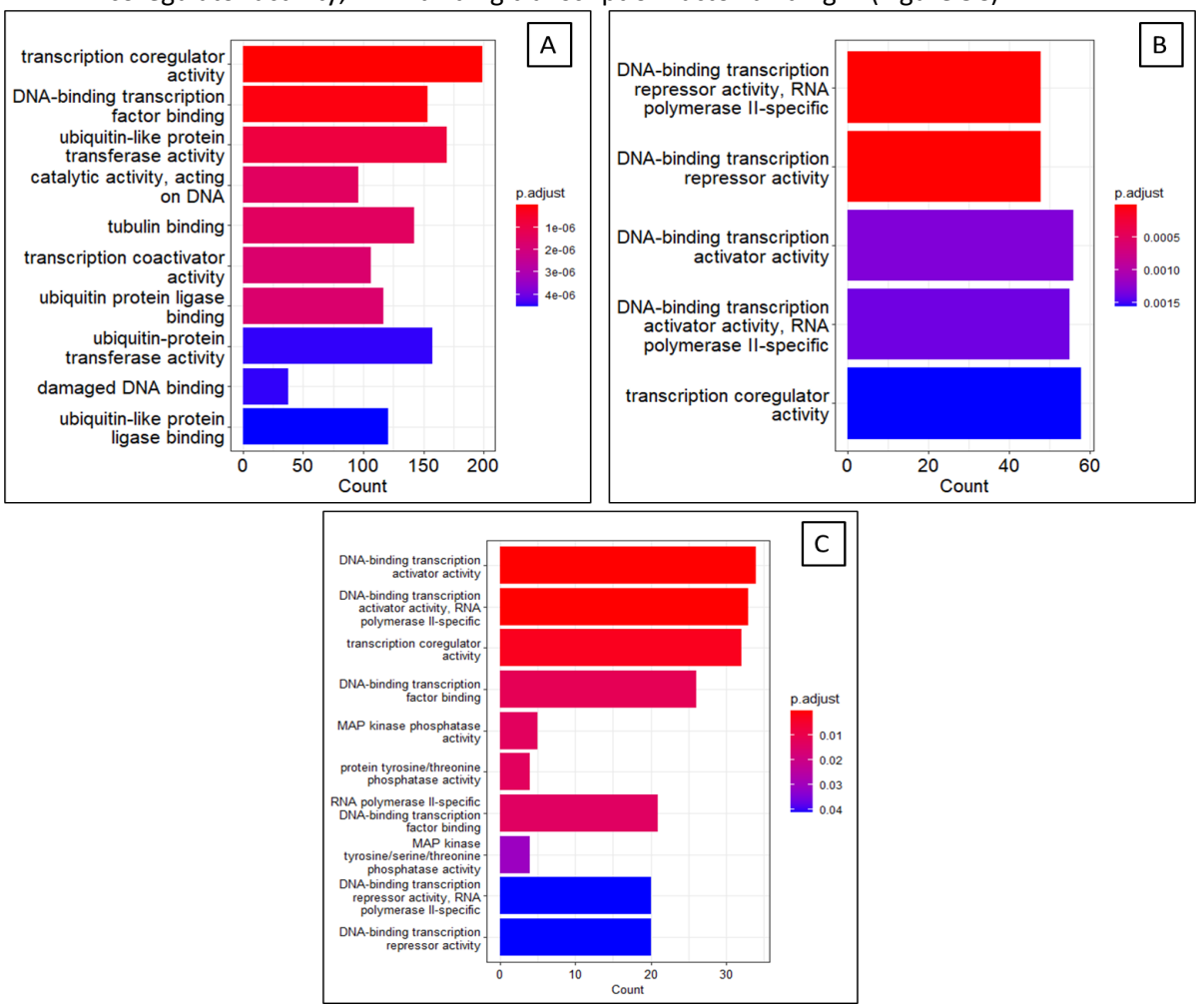

Figure 5.Gene Ontology (GO) analysis of the DEGs of cc1 dataset (A), cc2 dataset (B) and the common DEGs between cc1 and cc2 (C)

In both neuronal culture datasets, the neuronal membrane depolarization induces changes in the transcription regulation machinery, as described by the authors (Boulting et al., 2021; Kiltschewskij & Cairns, 2020). It also matches with previous literature (Flavell & Greenberg, 2008;

Greer & Greenberg, 2008), what suggests that the experimental model and data analysis are correct and reliable.

Interestingly, as we can see in Figure 4C and 4D, despite both datasets correspond to stimulated neuronal cultures, there are remarkable biological differences reflected in the number and the direction of the expression changes. To assess that, it's important to focus on the characteristics of the cultures.

Neurons from cc1 (Table 1) come from a differentiated human immortalized neuron-like cell line obtained from a neuroblastoma, SH-SY5Y cell line (Neuronal differentiation information in supplementary information 3). In contrast, cc2 are GABAergic neurons derived from hiPSC (Cell line information in supplementary information 3). So, the neuronal cultures come from different cell lines. Whereas the GABAergic neurons from cc2 are a specific cell type specialized in inhibitory signals, cc1 neurons are non-specific neuron type. It is well known that GABAergic neurons have a very specific transcriptomic profile that let them being separated and clustered easily from other type of neurons (Yao et al., 2021). That could explain de differences in the DEGs and GO results between the two neuronal cultures.

c) Low DEG in human AD samples

In the first human dataset (hs1), RNA sequencing data was obtained from occipital lobe tissue and DGE analysis was performed comparing control samples versus AD samples. But no genes were found to be differentially expressed. There are many reasons that could explain that.

- For this analysis, only 6 samples per group were analyzed. In the original work (Caldwell et al., 2022) 20 samples per group were analyzed. The wide range of variation in human samples makes it difficult to gather strong statistical data that can produce significant results. In this case, 6 samples per group could not be enough to obtain statistical power.
- In this case, bulk tissue is used for RNA-sequencing so, transcripts of every cell are analyzed, not only neurons. That could mask the alteration in expression of some genes that are altered only in a specific type of cell (Li & Wang, 2021).

In the second human dataset (hs2) 193 DEGs were found (Figure 6A). The low DEG rate obtained in this case can be attributed to the same reasons as in hs1. In this case, only 1 IEG has been detected as DEG: EGR4.

EGR4 is a gene that codes for a protein called EGR4 (early growth response 4). This protein is expressed in the nervous system and plays a role in the development and function of neurons. It is involved in the regulation of gene expression and has been shown to be important for proper brain development. Dysregulation of EGR4 has been implicated in various neurological disorders, such as schizophrenia and autism (Pérez-Cadahía et al., 2011). And it has not been related with Alzheimer's disease before.

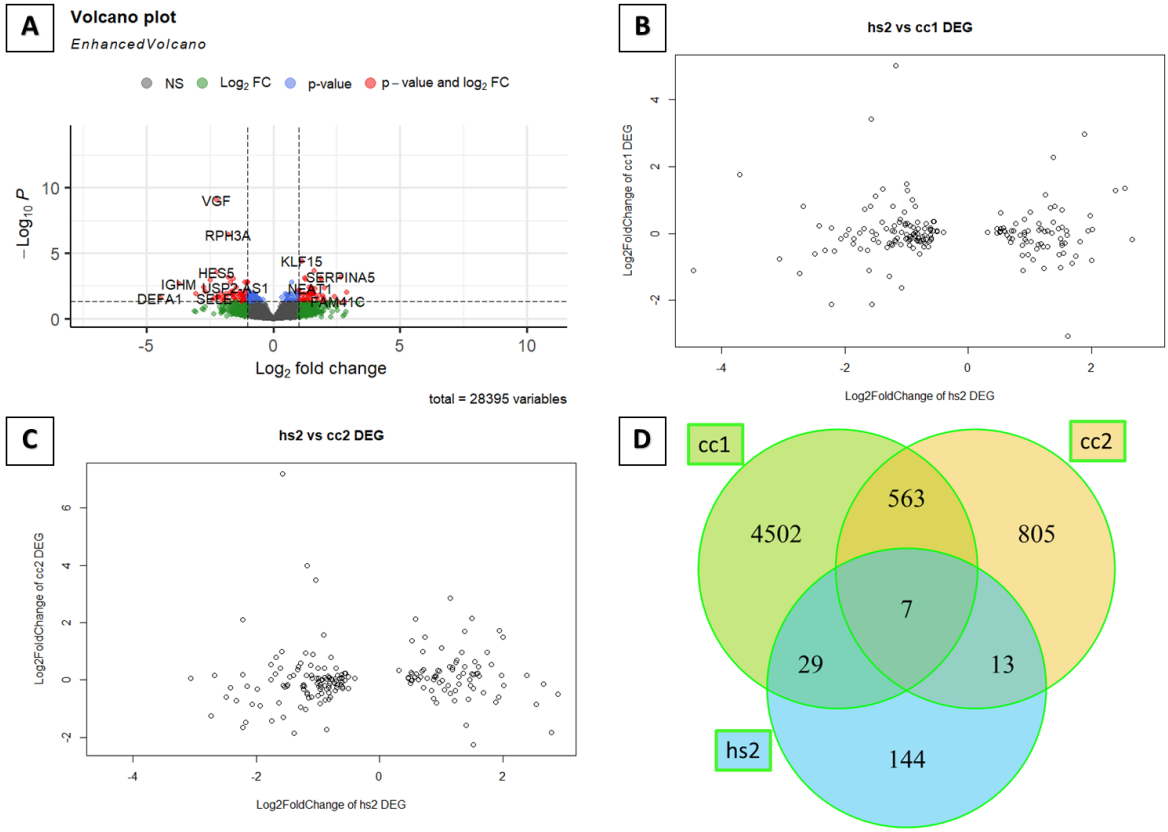

Figure 6 A) Volcanoplot of DEG Log2FoldChange of hs2 dataset. B) Log2Foldchange of hs2 DEG comparing with cc1 DEG. C) Log2Foldchange of hs2 DEG comparing with cc2 DEG. D) Ven diagram of DEG from hs2 vs neuronal cultures. In green (cc1) are represented the DEG of the neuronal culture study 1. In orange (cc2) the DEG of the neuronal culture 2. In blue (hs2) the DEG of the human AD samples 2.

d) 7 common DEGs between human AD samples and neuronal culture samples.

The comparison of transcriptional changes occurred in cell lines upon neuron stimulation and in AD samples indicated that DEGs found in cc1 (Figure 6B) or cc2 (6C) do not have the same direction. Nevertheless, hs2 analysis show 29 DEGs in common with cc1, 13 DEGs in common with cc2 and 7 DEGs present in the 3 datasets (Figure 6D).The seven common DEGs in both neuronal cultures and human brain AD samples are shown in table 4.

Table 4. Seven common genes differentially expressed between human AD samples and neuronal culture samples. In the first column the name of the gene, second column the log2FoldChange in hs2 dataset. Column 3 is the log2FoldChange in dataset cc1. Column 4 is the Log2FoldChange in dataset cc2. Last column describes briefly the function of the gene. Colored arrows indicate the change in transcription of DEG in comparison with the control group in each case (green= up-regulated, red=down-regulated).

SYMBOL	hs1 log2FoldChange	cc1 log2FoldChange	cc2 log2FoldChange	Gene/protein function
SLC6A17	-0,790 ⌄⌄	-0,611 ⌄⌄	0,798 ⌃⌃	Transporter for presynaptic uptake of NT, important for brain development
SLC38A2	1,009 ⌃⌃	0,637 ⌃⌃	1,121 ⌃⌃	Transporter of aminoacids at blood-brain-barrier
MAFF	1,385 ⌃⌃	0,343 ⌃⌃	1,703 ⌃⌃	bZIP transcription factor
EGR1	-1,032 ⌄⌄	0,632 ⌃⌃	3,491 ⌃⌃	Transcription regulator. IEG.
LINC-PINT	0,529 ⌃⌃	0,623 ⌃⌃	1,369 ⌃⌃	lncRNA for negative regulation of transcription
DUSP4	-0,901 ⌄⌄	1,001 ⌃⌃	1,584 ⌃⌃	Negativerly regulates MAP kinase superfamily (an active dependent pathway)
GADD45G	1,249 ⌃⌃	1,157 ⌃⌃	0,703 ⌃⌃	Stress response medator

Some activity dependent genes are dysregulated in AD (Table 4) as previously described (de Haan et al., 2012; Sperling et al., 2010) and they play role in different but very important biological functions essentials for correct neuron development as transcription regulation (Pérez-Cadahía et al., 2011) and ionic transporters (Greer & Greenberg, 2008).

Unfortunately, no Gene ontology could have been performed due the low number of genes in each group.

e) Heterogeneous number of DEGs in the mouse AD model hippocampal samples.

From the analysis of the mice hippocampal samples (mm1 and mm2 datasets), only 8 DEGs were reported in mm1, whereas 1071 DEGs were obtained in mm2.

To assess the differences, it is important to focus on the mice model used in each study. As explained in Table 1, the two studies were based in different mice models: J20 and 5xFAD (Oakley et al., 2006).
As we can see in Figure 7, the neuropathology differs substantially depending on the mutations that characterize each model. J20 mice have neuronal loss, changes in synaptic plasticity and cognitive impairment at the firsts months of age, and Aβ plaques and gliosis later. On the other hand, 5xFAD mice had changes in synaptic plasticity, gliosis and appearance of Aβ plaques at early months of life, but do not present cognitive impairment, synaptic loss and neuronal loss until the old age (Jawhar et al., 2012; Richard et al., 2015).

Taking into account this information, and considering that 5xFAD mice were exposed to a battery of behavioral test before being euthanized (Forner et al., 2021), it makes sense to find these transcriptional differences. One issue with animal models is that the statistical variation and the collection of bulk tissue for RNA-sequencing (previously mentioned in relation to human samples) can mask the biological differences of the studied disease.

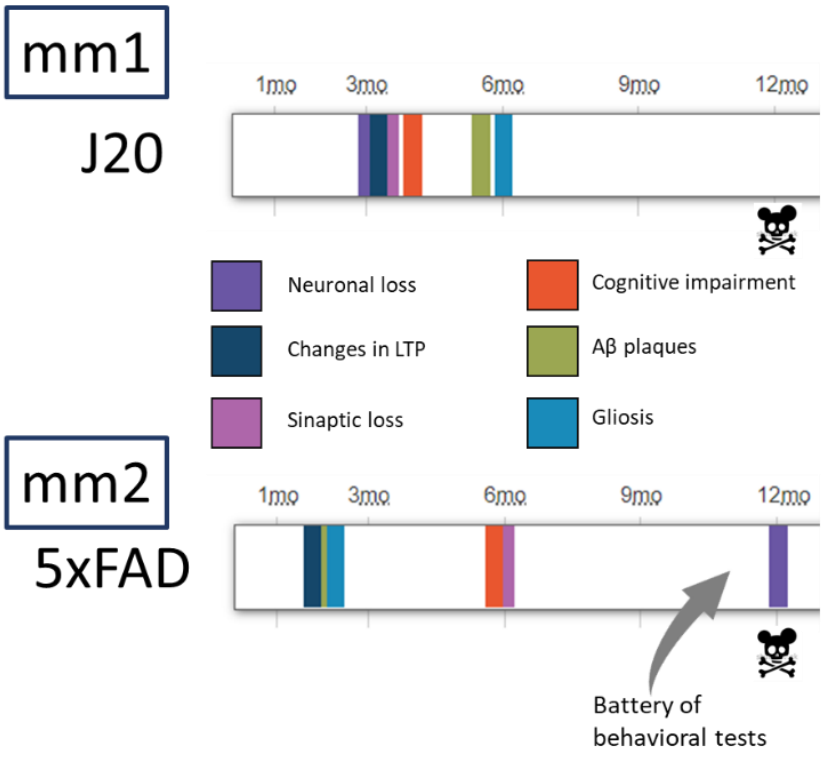

Figure 7. Comparison between the 2 different AD mice model used in each study.

The 8 DEGs obtained from mm1 dataset are:

APP CCN2 NXPH3 NECTIN3 PRSS12 PRG4 CCDC80 COL8A2

APP is the most differentially expressed gene being consistent with the mutations of the J20 mice model. J20 has two APP mutated forms from familiar Alzheimer's disease patients.
The rest of genes have different roles in cell structures, adhesion and neuronal growing factor. None of them are considered IEGs or activity dependent.

1071 DEGs were obtained from the mm2 dataset analysis. Among these genes, two are particularly interesting because they are immediate early genes (IEGs): ATF3 and MYC. Both of them are transcription factors regulated by the cAMP pathway, one of the main activity-dependent pathways in neurons (Ferrer et al., 2003).

f) AD mice model with the 5xFAD genetic mutations exhibits a transcriptional profile intermediate between neuronal cultures and Alzheimer's disease patients.

5xFAD AD model seems to have asymmetric transcriptomic profile in comparison with the previous analyzed groups (Figure 8A). In this case, most DEGs are up-regulated. Compared with the first neuronal culture and the human AD samples DEG lists (Figure 8B and 8C), mm2 DEGs have a better correlation with human AD samples (R^2=0.010) than with neuronal cultures (R^2=0.003). Moreover, when common DEGs are compared proportionally, human samples share a higher proportion of DEGs (hs1=8/149=0.053) than neuronal cultures (cc1= 191/4874 =0.039) (Figure 8D).

These results suggest that the transcriptomic profile of 5xFAD mice at 12 months is more similar to AD human patients than an activity dependent model (despite of the animals have been trained before euthanasia, what would be a model of neuronal activity). Our results match with previous published results (Landel et al., 2014; Oakley et al., 2006).

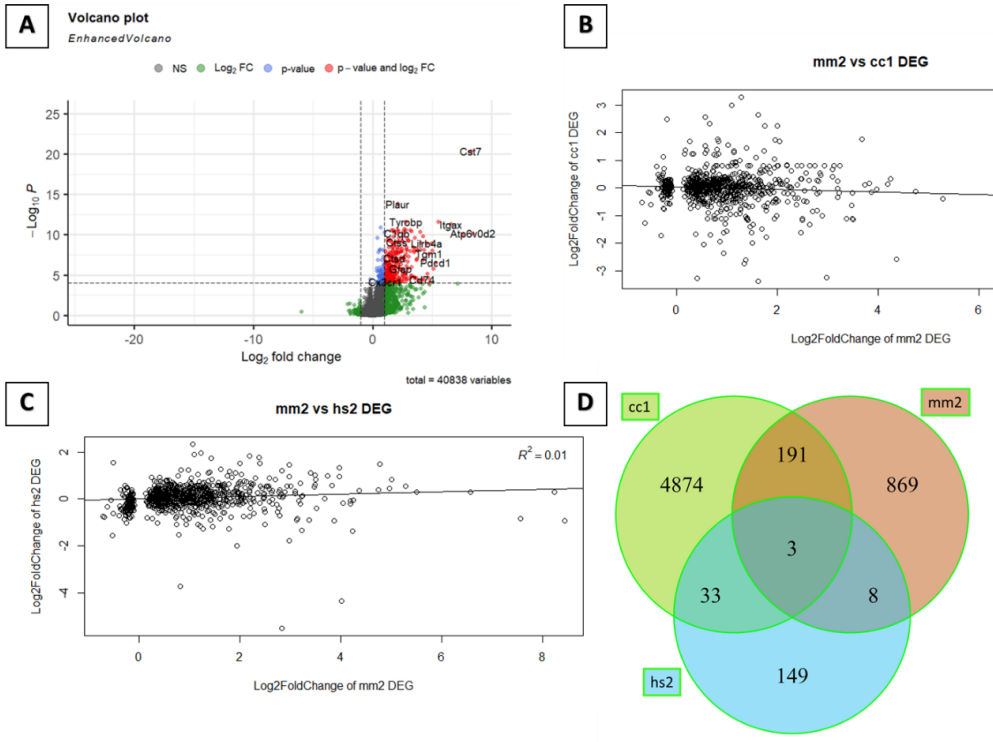

Figure 8. A) Volcanoplot of DEG Log2FoldChange of mm2 dataset. B) Log2Foldchange of mm2 DEG comparing with cc1 DEG. C) Log2Foldchange of mm2 DEG comparing with hs2 DEG. D) Ven diagram of DEG from mm2 vs cc1 and hs2.

In green (cc1) are represented the DEG of the neuronal culture study 1. In brown (mm2) the DEG of the mice model hippocampal samples 2. In blue (hs2) the DEG of the human AD samples 2

g) **mm2 DEGs are involved in neuroinflammation responses**.

Gene ontology analyses of the DEGs of 5xFAD mice reveal that the main cellular pathways altered by the mutations and/or behavioral training are related with immune responses (Figure 9).

In normal conditions, glial cells such as microglia and astrocytes detect and eliminate abnormal proteins and neuronal debris. If this is successful, the inflammatory response changes to an anti-inflammatory profile. However, in Alzheimer's disease, the production of certain proteins and an increase in neuronal death trigger a chronic, proinflammatory response called neuroinflammation that can be harmful to the brain and contribute to the development of AD (Minter et al., 2016; Rubio-Perez & Morillas-Ruiz, 2012).

Neuroinflammation has been widely associated with this AD mice model and others (Baik et al., 2019; Landel et al., 2014). Alzheimer's disease and most neurodegenerative diseases are associated with neuroinflammation processes that aggravate the pathological development (Palmer, 2011).

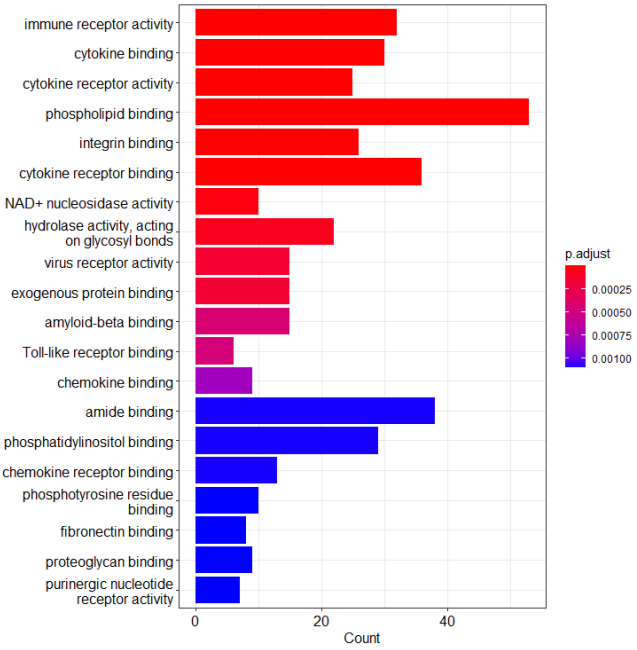

Figure 9. Gene ontology results from mm2 DEGs. Many of the functions are related with immune system and immune responses.

h) Common DEGs between AD mice models, neuronal cultures and AD human samples are involved in immune responses and neurotransmitter metabolism

When comparing the DEGs in the mm2 dataset with those in the human samples and neuronal cultures (Figure 8D), different metabolic pathways were observed. When comparing the DEGs from the 5xFAD mouse model with those from human samples of Alzheimer's patients, the ontology indicated different metabolic processes related to neurotransmitters (Figure 10A).

Surprisingly, when comparing the common DEGs between mm2 and neuronal cultures (cc1) only "cytokine receptor binding" molecular function was involved (Figure 10B).

Three common DEGs (FAM167B, RAB29, and LYNX1) were identified as being related to neurotransmitter activity (Figure 10C). NT alterations have been described to have an important role in cognitive impairment in AD and other dementias (Greer & Greenberg, 2008). As activity dependent genes, NT alterations have been found in different AD stadiums, at the point that some groups describe that can be used as biomarkers to predict and prevent AD progression (Beal & Growdon, 1986).

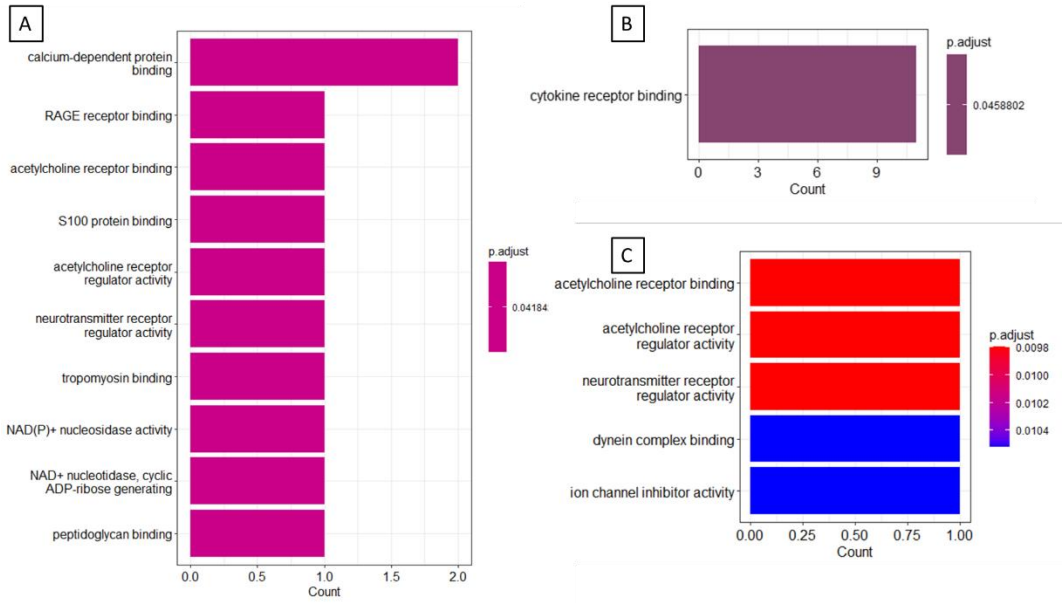

Figure 10. Gene ontology analysis of common DEGs between mice AD model and Neuronal cultures and human AD samples. A) GO analysis of the 11 DEGs shared by mm2 and hs2. B) GO analysis of the 194 DEGs shared by mm2 and cc1. C) GO analysis of the 3 DEGs shared by mm2, hs2 and cc1.

It was noted that there were no alterations in activity-dependent pathways, which may suggest that these alterations occur early in the disease and that both the mice and human samples were in advanced stages of Alzheimer's disease. As indicated in Figure 7, changes in LTP and synapses occur at firsts months of mice age, what could explain the absence of IEGs differentially expressed when comparing human and neuronal culture samples.

9. Conclusions

First of all, it is important to remark that in most datasets the raw reads uploaded by the authors were already filtered. This can lead to a bias in terms of initial qc parameters used.

We have seen in the stimulated neuronal cultures that activity induction by depolarization with KCl can be a good model of neuronal activity and to study IEGs. Despite the differences found between the two datasets of neuronal cultures, the number of IEGs in each dataset, and the common IEGs, show that this experimental approach can be valid. The differences between cultures can be attributed to the different origin of neurons.

Human samples results must be interpreted carefully. The high variability among individuals, and the use of bulk tissue instead of isolated neurons, can be masking the changes caused by the disease. For accurate results more samples should be collected and analyzed. Also, newer technologies such as single cell RNASeq must be used to describe the transcriptional changes in specific cell types. EGR4, an immediate early gene- transcript factor, has been found to be dysregulated, and could play an important role in Alzheimer's disease development. Apart from that, common DEGs between human samples and neuronal cultures show important changes in transcription regulation and synaptic processes that are activity dependent and could be important for first stages of Alzheimer’s disease development.

From the analysis of AD mice models we can conclude that it is very important to select a specific animal model. In this case, the two models used show prominent differences that have led to substantial dissimilar results.

When compared with previous results, the 5xFAD mice model highlights the neuroinflammation processes that are closely linked with AD and other neuropathologies.

To sum up, the bioinformatic analysis of all the collected samples in this work shows that Alzheimer's disease processes are related with activity dependent pathways. These can be studied using neuronal cultures, and mice models are important to study specific features of diseases like neuroinflammation.

10. Bibliography

Baik, S. H., Kang, S., Lee, W., Choi, H., Chung, S., Kim, J.-I., & Mook-Jung, I. (2019). A Breakdown in Metabolic Reprogramming Causes Microglia Dysfunction in Alzheimer's Disease. *Cell Metabolism*, *30*(3), 493-507.e6. https://doi.org/10.1016/j.cmet.2019.06.005

Bailey, C. H., Kandel, E. R., & Harris, K. M. (2015). Structural Components of Synaptic Plasticity and Memory Consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(7), a021758. https://doi.org/10.1101/cshperspect.a021758

Beal, M. F., & Growdon, J. H. (1986). CSF neurotransmitter markers in Alzheimer's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *10*(3–5), 259–270. https://doi.org/10.1016/0278-5846(86)90006-0

Blighe, K., Rana, S., & Lewis, M. (2021). *EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.* (R package version 1.12.0.).

Boulting, G. L., Durresi, E., Ataman, B., Sherman, M. A., Mei, K., Harmin, D. A., Carter, A. C., Hochbaum, D. R., Granger, A. J., Engreitz, J. M., Hrvatin, S., Blanchard, M. R., Yang, M. G., Griffith, E. C., & Greenberg, M. E. (2021). Activity-dependent regulome of human GABAergic neurons reveals new patterns of gene regulation and neurological disease heritability. *Nature Neuroscience*, *24*(3), 437–448. https://doi.org/10.1038/s41593-020-00786-1

Caldwell, A. B., Anantharaman, B. G., Ramachandran, S., Nguyen, P., Liu, Q., Trinh, I., Galasko, D. R., Desplats, P. A., Wagner, S. L., & Subramaniam, S. (2022). Transcriptomic profiling of sporadic Alzheimer's disease patients. *Molecular Brain*, *15*(1), 83. https://doi.org/10.1186/s13041-022-00963-2

Castanho, I., Murray, T. K., Hannon, E., Jeffries, A., Walker, E., Laing, E., Baulf, H., Harvey, J., Bradshaw, L., Randall, A., Moore, K., O'Neill, P., Lunnon, K., Collier, D. A., Ahmed, Z., O'Neill, M. J., & Mill, J. (2020). Transcriptional Signatures of Tau and Amyloid Neuropathology. *Cell Reports*, *30*(6), 2040-2054.e5. https://doi.org/10.1016/j.celrep.2020.01.063

Chen, H. (2022). *VennDiagram: Generate High-Resolution Venn and Euler Plots.* (R package version 1.7.3.).

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

de Haan, W., Mott, K., van Straaten, E. C. W., Scheltens, P., & Stam, C. J. (2012). Activity Dependent Degeneration Explains Hub Vulnerability in Alzheimer's Disease. *PLoS Computational Biology*, *8*(8), e1002582. https://doi.org/10.1371/journal.pcbi.1002582

Ferrer, I., Friguls, B., Dalfó, E., & Planas, A. M. (2003). Early modifications in the expression of mitogen-activated protein kinase (MAPK/ERK), stress-activated kinases SAPK/JNK and p38, and their phosphorylated substrates following focal cerebral ischemia. *Acta Neuropathologica*, *105*(5), 425–437. https://doi.org/10.1007/s00401-002-0661-2

Flavell, S. W., & Greenberg, M. E. (2008). Signaling Mechanisms Linking Neuronal Activity to Gene Expression and Plasticity of the Nervous System. *Annual Review of Neuroscience*, *31*(1), 563–590. https://doi.org/10.1146/annurev.neuro.31.060407.125631

Forner, S., Kawauchi, S., Balderrama-Gutierrez, G., Kramár, E. A., Matheos, D. P., Phan, J., Javonillo, D. I., Tran, K. M., Hingco, E., da Cunha, C., Rezaie, N., Alcantara, J. A., Baglietto-Vargas, D., Jansen, C., Neumann, J., Wood, M. A., MacGregor, G. R., Mortazavi, A., Tenner, A. J., … Green, K. N. (2021). Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. *Scientific Data*, *8*(1), 270. https://doi.org/10.1038/s41597-021-01054-y

Greer, P. L., & Greenberg, M. E. (2008). From Synapse to Nucleus: Calcium-Dependent Gene Transcription in the Control of Synapse Development and Function. *Neuron*, *59*(6), 846–860. https://doi.org/10.1016/j.neuron.2008.09.002

Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogaeva, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J. S. K., Younkin, S., Hazrati, L., Collinge, J., Pocock, J., Lashley, T., Williams, J., Lambert, J.-C., Amouyel, P., Goate, A., Rademakers, R., … Hardy, J. (2013). *TREM2* Variants in Alzheimer's Disease. *New England Journal of Medicine*, *368*(2), 117–127. https://doi.org/10.1056/NEJMoa1211851

Harris, K. M. (2020). Structural LTP: from synaptogenesis to regulated synapse enlargement and clustering. *Current Opinion in Neurobiology*, *63*, 189–197. https://doi.org/10.1016/j.conb.2020.04.009

Hebb, D. (1950). Hebb, D. O. The organization of behavior: A neuropsychological theory. New York: John Wiley and Sons, Inc., 1949. 335 p. $4.00. *Science Education*, *34*(5), 336–337. https://doi.org/10.1002/sce.37303405110

Jawhar, S., Trawicka, A., Jenneckens, C., Bayer, T. A., & Wirths, O. (2012). Motor deficits, neuron loss, and reduced anxiety coinciding with axonal degeneration and intraneuronal Aβ aggregation in the 5XFAD mouse model of Alzheimer's disease. *Neurobiology of Aging*, *33*(1), 196.e29-196.e40. https://doi.org/10.1016/j.neurobiolaging.2010.05.027

Kandel, E. R. (2001). The Molecular Biology of Memory Storage: A Dialogue Between Genes and Synapses. *Science*, *294*(5544), 1030–1038. https://doi.org/10.1126/science.1067020

Kassambara, A. (2019). *fastqcr: Quality Control of Sequencing Data. R package version 0.1.2.* (R package version 0.1.2.).

Kiltschewskij, D. J., & Cairns, M. J. (2020). Transcriptome-Wide Analysis of Interplay between mRNA Stability, Translation and Small RNAs in Response to Neuronal Membrane Depolarization. *International Journal of Molecular Sciences*, *21*(19). https://doi.org/10.3390/ijms21197086

Landel, V., Baranger, K., Virard, I., Loriod, B., Khrestchatisky, M., Rivera, S., Benech, P., & Féron, F. (2014). Temporal gene profiling of the 5XFAD transgenic mouse model highlights the importance of microglial activation in Alzheimer's disease. *Molecular Neurodegeneration*, *9*(1), 33. https://doi.org/10.1186/1750-1326-9-33

Li, X., & Wang, C.-Y. (2021). From bulk, single-cell to spatial RNA sequencing. *International Journal of Oral Science*, *13*(1), 36. https://doi.org/10.1038/s41368-021-00146-0

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, *47*(8), e47–e47. https://doi.org/10.1093/nar/gkz114

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Minter, M. R., Taylor, J. M., & Crack, P. J. (2016). The contribution of neuroinflammation to amyloid toxicity in Alzheimer's disease. *Journal of Neurochemistry*, *136*(3), 457–474. https://doi.org/10.1111/jnc.13411

Nativio, R., Lan, Y., Donahue, G., Sidoli, S., Berson, A., Srinivasan, A. R., Shcherbakova, O., Amlie-Wolf, A., Nie, J., Cui, X., He, C., Wang, L.-S., Garcia, B. A., Trojanowski, J. Q., Bonini, N. M., & Berger, S. L. (2020). An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease. *Nature Genetics*, *52*(10), 1024–1035. https://doi.org/10.1038/s41588-020-0696-0

Oakley, H., Cole, S. L., Logan, S., Maus, E., Shao, P., Craft, J., Guillozet-Bongaarts, A., Ohno, M., Disterhoft, J., van Eldik, L., Berry, R., & Vassar, R. (2006). Intraneuronal beta-Amyloid Aggregates, Neurodegeneration, and Neuron Loss in Transgenic Mice with Five Familial Alzheimer's Disease Mutations: Potential Factors in Amyloid Plaque Formation. *Journal of Neuroscience*, *26*(40), 10129–10140. https://doi.org/10.1523/JNEUROSCI.1202-06.2006

Palmer, A. M. (2011). Neuroprotective therapeutics for Alzheimer's disease: progress and prospects. *Trends in Pharmacological Sciences*, *32*(3), 141–147. https://doi.org/10.1016/j.tips.2010.12.007

Patel, H., Dobson, R. J. B., & Newhouse, S. J. (2019). A Meta-Analysis of Alzheimer's Disease Brain Transcriptomic Data. *Journal of Alzheimer's Disease*, *68*(4), 1635–1656. https://doi.org/10.3233/JAD-181085

Pérez-Cadahía, B., Drobic, B., & Davie, J. R. (2011). Activation and function of immediate-early genes in the nervous systemThis paper is one of a selection of papers in a Special Issue entitled 31st Annual International Asilomar Chromatin and Chromosomes Conference, and has undergone the Journal's usual peer review process. *Biochemistry and Cell Biology*, *89*(1), 61–73. https://doi.org/10.1139/O10-138

Richard, B. C., Kurdakova, A., Baches, S., Bayer, T. A., Weggen, S., & Wirths, O. (2015). Gene Dosage Dependent Aggravation of the Neurological Phenotype in the 5XFAD Mouse Model of Alzheimer's Disease. *Journal of Alzheimer's Disease*, *45*(4), 1223–1236. https://doi.org/10.3233/JAD-143120

Roberson, E. D., English, J. D., & Sweatt, J. D. (1996). A biochemist's view of long-term potentiation. *Learning & Memory*, *3*(1), 1–24. https://doi.org/10.1101/lm.3.1.1

Rubio-Perez, J. M., & Morillas-Ruiz, J. M. (2012). A Review: Inflammatory Process in Alzheimer's Disease, Role of Cytokines. *The Scientific World Journal*, *2012*, 1–15. https://doi.org/10.1100/2012/756357

Sevigny, J., Chiao, P., Bussière, T., Weinreb, P. H., Williams, L., Maier, M., Dunstan, R., Salloway, S., Chen, T., Ling, Y., O'Gorman, J., Qian, F., Arastu, M., Li, M., Chollate, S., Brennan, M. S., Quintero-Monzon, O., Scannevin, R. H., Arnold, H. M., … Sandrock, A. (2016). The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature*, *537*(7618), 50–56. https://doi.org/10.1038/nature19323

Sperling, R. A., Dickerson, B. C., Pihlajamaki, M., Vannini, P., LaViolette, P. S., Vitolo, O. v., Hedden, T., Becker, J. A., Rentz, D. M., Selkoe, D. J., & Johnson, K. A. (2010). Functional Alterations in Memory Networks in Early Alzheimer's Disease. *NeuroMolecular Medicine*, *12*(1), 27–43. https://doi.org/10.1007/s12017-009-8109-7

Tong, L., Thornton, P. L., Balazs, R., & Cotman, C. W. (2001). β-Amyloid-(1–42) Impairs Activity-dependent cAMP-response Element-binding Protein Signaling in Neurons at Concentrations in Which Cell Survival Is Not Compromised. *Journal of Biological Chemistry*, *276*(20), 17301–17306. https://doi.org/10.1074/jbc.M010450200

Walsh, S., Merrick, R., Milne, R., & Brayne, C. (2021). Aducanumab for Alzheimer's disease? *BMJ*, n1682. https://doi.org/10.1136/bmj.n1682

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*(3), 100141. https://doi.org/10.1016/j.xinn.2021.100141

Yan, P., Xue, Z., Li, D., Ni, S., Wang, C., Jin, X., Zhou, D., Li, X., Zhao, X., Chen, X., Cui, W., Xu, D., Zhou, W., & Zhang, J. (2021). Dysregulated CRTC1-BDNF signaling pathway in the hippocampus contributes to Aβ oligomer-induced long-term synaptic plasticity and memory impairment. *Experimental Neurology*, *345*, 113812. https://doi.org/10.1016/j.expneurol.2021.113812

Yao, Z., van Velthoven, C. T. J., Nguyen, T. N., Goldy, J., Sedeno-Cortes, A. E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., Ding, S.-L., Fong, O., Garren, E., Glandon, A., Gouwens, N. W., Gray, J., Graybuck, L. T., Hawrylycz, M. J., Hirschstein, D., … Zeng, H. (2021). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell*, *184*(12), 3222-3241.e26. https://doi.org/10.1016/j.cell.2021.04.021

Zucker, R. S., & Regehr, W. G. (2002). Short-Term Synaptic Plasticity. *Annual Review of Physiology*, *64*(1), 355–405. https://doi.org/10.1146/annurev.physiol.64.092501.114547

11. Supplementary information

1. Filtering information of raw data.

Dataset	Preprocessing	Software used
hs1	-Removed sequencing adaptors - quality threshold: Phred Q > 20.	TrimGalore!
hs2	Not specified	
mm1	-Ribosomal sequences removed -quality threshold: Phred Q > 20. -minimum sequence length: 35	F*astq-mcf*
mm2	None	
cc1	-Removed sequencing adaptors - quality threshold: Phred Q > 28.	Cutadapt'
cc2	Not specified	MAPtoFeatures (in-house software)

2. R code

All the R code used to develop this work, is abaliable in Github:

https://github.com/Xaviferol/XFO_Final_master_thesis.git

3. Neuronal cultures obtention:
    a. SH-SY5Y cultures were neuronally differentiated using all-trans retinoic acid (ATRA, Sigma-Aldrich, St. Louis, MO, USA).
    Day -1: cells were seeded at a density of 25,000 cells/cm2 in T175 culture flasks.
    Day 0: standard culture medium was replaced with ATRA-supplemented medium (10µM final).
    ATRA-medium was replaced on days 2, 4 and 6, after which the methods were continued as described on day 7.
    Successful differentiation was confirmed by examining neurite outgrowth and establishing the upregulation of known neuronal marker genes relative to naïve cells via mRNA sequencing
    (Kiltschewskij & Cairns, 2020)

    b. hiPSC-derived GNs were obtained commercially (Cellular Dynamics International iCell Neurons, NRC-100-010-001)
    Human induced pluripotent stem cell (hiPSC)-derived cultures highly enriched for developing forebrain GABAergic neurons (hGNs)
    (Boulting et al., 2021)