

Identifying Cell-Cell Communication in Alzheimer's Disease Pathology

Promoter:

Prof. De Strooper, B.

Dr. Fiers, M.

VIB-KU LEUVEN

CENTER FOR BRAIN & DISEASE RESEARCH

Dissertation presented in
fulfilment of the requirements
for the degree of Master of Science:

Bioinformatics

Nenghan Lin

This dissertation is part of the examination and has not been corrected after defense for eventual errors. Use as a reference is permitted subject to written approval of the promotor stated on the front page.

Acknowledgments

I would like to start by expressing my appreciation to Prof. Bart De Strooper and Dr. Mark Fiers, who gave me an opportunity to work in such a fascinating field and whose knowledge and advice have been crucial to my development as a scientist. Additionally, I want to thank everyone in the lab for their hard work in this research. Additionally, I want to express my gratitude to Ashley Lu for all her guidance and support given to me during these trying times. Finally, I would like to thank my parents for their support and understanding.

Abstract

The dominant idea used to explain the pathophysiology of AD in recent years is the amyloid hypothesis. Based on this hypothesis, some studies have been done to monitor transcriptional changes occurring near the Amyloid- β (A β) plaque niche to better comprehend AD pathophysiology. Multiple cell types, such as astrocytes and microglia, were observed to enrich around the A β plaque. Therefore, more research, such as integrating single-cell transcriptomics, is required to identify the cell-cell communications (CCCs) between these plaque associated cells (PACs) and how they contribute to the cascade in AD. CellphoneDB and NicheNet are two publicly available repositories to study CCCs. This thesis will utilize CellphoneDB and NicheNet to find some potential ligand-receptor interactions in CCCs between PACs. These two methods have different algorithms, which complement one another's shortcomings, and at the same time, have room for improvement. Through these two methods, 18 potential ligand-receptor interactions were found. More research should be done on these ligand-receptor interactions because they are probably involved in the CCCs between PACs.

List of Abbreviations and Symbols

AD	Alzheimer's Disease
APP	Amyloid precursor protein
AUPR	Area under the precision/recall curve
AUROC	Area under the receiver operating characteristic curve
A β	Amyloid- β
BBB	Blood-brain barrier
CCC	Cell-cell communication
CNS	Central nervous system
DE	Differential expression
GWAS	Genome-wide association study
ISS	<i>In Situ</i> Sequencing
LFC	Log fold change
NBB	Netherlands Brain Bank
OLIG	Oligodendrocyte gene
PAC	Pathology Associated Cell
PIG	Plaque-induced gene
PPR	Personalized PageRank
pTau	Phosphorylated Ta
scRNA-seq	Single-cell RNA sequencing
snRNA-seq	Single-nuclei RNA sequencing
ST	Spatial Transcriptomics
TREM2	Triggering receptor expressed on myeloid cells 2

List of Figures

Figure 1: Interactions between proteins or non-proteins such as steroid hormones, small molecules, or neurotransmitters are stored in the CellphoneDB database.....	17
Figure 2: The null distribution of the mean expression of a given interaction between two cell types.....	18
Figure 3: The generation paths of two weighted networks.....	20
Figure 4: The generation process of the prior model of the ligand-target matrix (Browaeys et al., 2020).....	21
Figure 5: The total number of ligand-receptor interactions in three phenotypes (AD, RES and ND).....	31
Figure 6: The scatter plots of mean expression of all possible interactions in AD against the mean expression of interactions in RES when receiver cell type was astro_5.....	32
Figure 7: The number of celltype-specific interactions with significant p-values in two phenotypes (AD and RES).....	33
Figure 8: The Venn diagrams of celltype-specific interaction among PACs.....	34
Figure 9: The distribution of celltype-specific interactions in sender-receiver cell pairs.....	35
Figure 10: Dot plot of celltype-specific interactions when the receiver was astro_0.....	38
Figure 11: The simple ligands and simple receptors found in celltype-specific interactions when the receiver was astro_0.....	38
Figure 12: Line plot of background genes number in five cell types (micro_5, astro_0, oligo_0, oligo_3 and oligo_4).....	40
Figure 13: Volcano plots of DE genes in ADvsND and RESvsND for each cell type in PACs.....	42
Figure 14: Comparison of liagnd_activity_ADvsND and ligand_activity_RESvsND when the receiver was one cell type in PACs and the rest were senders.....	45
Figure 15: The Venn plot of active ligands found in each receiver.....	47
Figure 16: From left to right, the plots represent the ligand activity, expression and log fold change(LFC) of each active ligand for receiver astro_0	49
Figure 17: From left to right, the plots represent the ligand-receptor weights, expression and log fold change(LFC) of each receptor in receiver astro_0.....	49
Figure 18: Venn plots of ligands and receptors in the databases of CellphoneDB and NicheNet (ligands and receptors are simple proteins in the CellphoneDB database).....	50
Figure 19: Venn plot of the number of ligands found in NicheNet and CellphoneDB.....	51

List of Tables

Table 1: The number of celltype-specific interactions found in each receiver cell type.....	38
Table 2: The number of DEGs found for receiver cell type in ADvsND and RESvsND.	42
Table 3: The number of ligands, background genes and DEGs for each receiver cell type and the active ligands (top 20 ligands in ADvsND or RESvsND).	46
Table 4: The final ligand-receptor pairs found in both NicheNet analysis and CellphoneDB analysis.	52

Contents

1 Context and Aims	9
2 Literature Review	10
2.1 Alzheimer's Disease	10
2.1.1 Introduction	10
2.1.2 The amyloid hypothesis of AD.....	11
2.1.3 The cellular responses of AD	13
2.1.4 The multicellular gene co-expression network.....	16
2.2 Single-cell Transcriptomics.....	16
2.2.1 CellphoneDB	17
2.2.2 NicheNet.....	19
3 Materials and Methods	23
3.1 Sample collection and data acquisition	23
3.1.1 Sequencing data collection	23
3.1.2 Meta-data collection	23
3.2 Data preprocessing	24
3.2.1 Pre-analysis.....	24
3.2.2 The Seurat object.....	25
3.2.3 The AnnData object.....	25
3.4 Differential expression analysis.....	26
3.4 CellphoneDB analysis	27
3.5 NicheNet analysis.....	29
4 Results	30
4.1 Celltype-specific ligand-receptor interactions identified by CellphoneDB.....	30
4.1.1 All possible interactions found in AD, RES and ND by CellphoneDB	30
4.1.2 Visualization of mean expressions and p-values of all possible interactions	31
4.1.3 548 celltype-specific interactions were found as candidates for CCCs.....	32
4.1.4 Distribution of celltype-specific interactions in sender-receiver cell pairs	34
4.1.4 All celltype-specific interactions were divided into five receiver groups	35
4.2 Active ligands identified by NicheNet	39
4.2.1 Control the quantity of background genes to ensure the accuracy of ligand predictions	39
4.2.2 Determine DEGs by DE analysis to ensure the accuracy of ligand predictions.....	40

4.2.3 Visualization of ligand activity in ADvsND and RESvsND	43
4.2.4 Top 20 ligands ranked by activity scores	45
4.2.2 Additional information of active ligands.....	47
4.3 Intersection between findings of CellphoneDB and NicheNet.....	50
4.3.1 Data sources of CellphoneDB and NicheNet	50
4.3.2 Potential cell-cell communications between PACs	50
5 Discussion.....	55
5.1 CellphoneDB	55
5.1.1 Mean expression and p-value for each interaction	55
5.1.2 Celltype-specific interactions	56
5.2 NicheNet.....	57
5.2.1 The quantity control of background genes	57
5.2.2 DE analysis.....	57
5.2.3 Ligand prediction analysis.....	58
5.2.4 Active ligands.....	59
5.3 Potential CCCs in PACs	59
5.4 Evaluation of CellphoneDB.....	60
5.5 Evaluation of NicheNet	61
6. Conclusion.....	62
5 References	63
7 Appendix	69

1 Context and Aims

Alzheimer's disease (AD) is a damaging neurodegenerative disease that begins with memory loss and progresses to cognitive loss. For more than 20 years, the amyloid hypothesis has been the primary theory explaining the pathophysiology of AD (Ricciarelli & Fedele, 2017). The amyloid hypothesis proposes a complex, multicellular, inflammatory cascade initiated by A β pathology (De Strooper & Karraan, 2016). To better comprehend the pathophysiology of AD as well as discover new therapeutic strategies, it is crucial to study the accompanying cellular responses and changes at the molecular level in the cascade. Chen et al. (2020) used Special Transcriptomics (ST) and In Situ Sequencing (ISS) to track transcriptional alteration happening around the A β plaque niche in a mouse model of AD. Multiple cell types, such as astrocytes and microglia, were shown to associate with one another in the A β niche, and multicellular gene co-expression networks were discovered. However, human pathologies cannot be recalculated by using a mouse model. Therefore, this thesis integrated single-cell transcriptomics in the human post-mortem model in three different phenotypes (AD, RES and ND) to understand the cell-cell communications (CCCs) between these cells and how they contribute to the cascade in Alzheimer's disease. 'AD' represents the plaque is present and the Alzheimer's symptoms occur; 'RES' represent the plaque is present but the Alzheimer's symptoms are 'resilient'; 'ND' represent control samples without plaque and Alzheimer's symptoms.

CellphoneDB and NicheNet are two publicly available repositories to investigate CCCs in single-cell transcriptomics studies. In this thesis, we aim to utilize CellphoneDB and NicheNet together to find some interesting ligand-receptor pairs in CCCs between cells around the A β plaques. CellphoneDB focuses on the expression of ligands and receptors in sender-cells and receiver-cells, which indicates interactions with ligand-receptor pair that have relatively high gene expression and cell-type specificity, but interactions found from CellphoneDB lack downstream gene confirmation of receptors. In contrast, NicheNet uses a prior model of the ligand-target matrix to find ligands that are responsible for the differential response of downstream target genes. Therefore, this thesis investigates CCCs by combining these two approaches, which fill in each other's deficiencies. Additionally, this thesis does not entirely follow the tutorials for these two methods, and some algorithms in this thesis have been modified in accordance with the author's recommendations in order to better fit our own data.

2 Literature Review

2.1 Alzheimer's Disease

2.1.1 Introduction

Alzheimer's disease (AD) is a common type of dementia that causes memory loss at the beginning and leads to loss of cognitive abilities in daily life. Aging is the risk factor to get Alzheimer's, and there is no known cure for this disease (Matthews et al., 2018).

Alzheimer's disease is a kind of neurodegenerative disease named after Alois Alzheimer, who is a German psychiatrist and neurologist. In 1906, a woman with an atypical mental illness died after suffering memory loss, disorientation, language problems, and unpredictable behaviors. After her death, Dr. Alzheimer discovered misshapen clumps of proteins and twisted bundles of fibers in her brain, and his findings were documented in his report as distinctive plaques and neurofibrillary, or tau, tangles. A few years later, in 1910, another German doctor named Emil Kraepelin codified Alzheimer's discoveries and assigned the syndrome the name "Alzheimer's disease" (Cipriani et al., 2011).

In the early 1980s, thousands of scientists began investigating the molecular studies of Alzheimer's disease as a result of the disease's crazily rising prevalence. One intriguing idea was proposed by George Glenner in 1984, which is that the specific amyloidogenic protein deposition in AD ($A\beta$) could be causative. Many have questioned this audacious hypothesis, but it has gained more and more support as data from several preclinical and clinical trials continue to mount (Selkoe et al., 2016). The amyloid cascade hypothesis has taken the lead as the principal theoretical basis for illuminating the pathogenesis and etiology of Alzheimer's disease since 1992. According to this hypothesis, the deposition of Amyloid-beta is the initial event that causes neuronal degradation (Reitz et al., 2012). This accumulation then kicks off a series of events that finally leads to brain cell death and the development of Alzheimer's disease. Many academic and pharmacological studies over the past 20 years have been inspired and driven by this hypothesis, which is driving the generation of potential therapies (Karran et al., 2011).

In addition to amyloid- β accumulation in amyloid plaques and subsequent biochemical processes, the delicate balance of connections between neurons, astrocytes, microglia, and vascular cells necessary for normal brain function is disturbed during the disease (Henstridge et al., 2019). The complicated cellular phase of the disease, which has been developing over decades, initially exhibits innocuous symptoms before turning chronic. Therefore, both the amyloid cascade concept and the intricate cellular interactions may offer significant insight into creating potent treatments for Alzheimer's disease (De Strooper & Karran, 2016).

2.1.2 The amyloid hypothesis of AD

In Alzheimer's disease, what triggers brain cells to degenerate and die has always remained a mystery. Although the ultimate pathogenesis of AD is still undetermined, various hypotheses have been propounded as potential candidates. Among these theories, the amyloid hypothesis has been the predominant theory of AD pathogenesis for more than 20 years (Ricciarelli & Fedele, 2017). In 1992, Hardy and Higgins proposed the amyloid- β peptide deposition in the brain parenchyma as the initial event in the pathogenesis of AD (HARDY & HIGGINS, 1992).

2.1.2.1 A cascade model initiated by amyloid beta accumulation

Amyloid beta ($A\beta$) refers to peptides of 36-43 amino acids, which constitute the majority of the amyloid plaques discovered in the brains of AD sufferers. The peptides come from the amyloid precursor protein (APP), which beta secretase and gamma secretase can break down to produce $A\beta$. With the development of AD, these peptides cluster collectively to form a type of poisonous oligomer which is also called plaques. The human body is unable to decompose or eradicate such plaques, so they slowly deposit in the brain. Amyloidosis, which is resulted from the deposition of amyloid, is suspected to play a role in several neurodegenerative illnesses (Wang et al., 2021)

The accumulation of $A\beta$ oligomers serves as the catalyst in the linear cascade model of the amyloid hypothesis of AD. Gradually, this first event will trigger a cascade of processes such as Tau pathology, synapse loss, neuronal death, oxidative stress, neuroinflammation, and finally, loss of memory and cognitive ability (De Strooper & Karran, 2016).

2.1.2.2 The controversies of the amyloid hypothesis

A number of studies support the amyloid cascade hypothesis. For instance, mutated genes in AD patients are consistently related to the biological process of A β formation or accumulation (Julia et al., 2017). Moreover, numerous supporting research created mutant mice that generate amyloid plaques and exhibit symptoms of human AD, which include memory loss in mazes (Nilsson et al., 2014).

However, this hypothesis remains highly debated. For instance, many people with trisomy 21 (Down's Syndrome) were discovered to have AD by the age of 40. Such Down's individuals would have a triple copy of APP, which results in excess A β production. The amyloid hypothesis states that these people with A β accumulation should develop AD. Paradoxically, although plaques accumulate and A β expression is enhanced, not all Down's syndrome sufferers develop AD (Zigman et al., 2008).

2.1.2.3 The extension of the amyloid hypothesis

Although not everyone endorses the amyloid hypothesis, there is no denying A β has dominated investigations on AD for decades. The amyloid hypothesis suggests that A β is an ideal therapeutic target, yet numerous therapeutic approaches focused on reducing A β have so far been shown to be ineffective in clinical studies. The failure of therapies to reduce A β results in more people doubting the Amyloid Hypothesis (Teich & Arancio, 2012).

However, AD is probably a complicated disease controlled by many variables. De Strooper and Karran highlighted that the discovery of poisonous A β oligomers leads to a variety of molecular pathways, which complicates and raises questions about literary studies, but they proposed another interpretation for the controversy surrounding the pathogenesis of AD, which is that molecular pathways have diverse effects at various points in the development of this disease (De Strooper & Karran, 2016). It is becoming abundantly evident that the accumulation of A β uses a variety of enzymes and signaling pathways that are involved in a wide range of cellular interactions. Thus, the therapeutic failure of reducing A β may indicate that controlling A β directly is an unrealistic way, given the intricate function of A β in neuronal physiology, rather than that the amyloid hypothesis is erroneous (Teich & Arancio, 2012). Therefore, in addition to the amyloid hypothesis, it is essential to take into account the role of different cells as well as their interactions when investigating the

pathogenesis of AD.

2.1.3 The cellular responses of AD

The amyloid hypothesis mainly focuses on the toxicity of A β to neurons and synapses, but neurons do not function alone. The cellular interactions between neurons, astrocytes, microglia, and vascular cells build a balance in the healthy human brain (Villegas-Llerena et al., 2015). It is crucial to investigate how these connections alter as AD develops in order to better comprehend the pathophysiology of the disease and identify new treatments. In addition to the cascade events in the amyloid hypothesis, the brain experiences the innate immune response and the inflammatory response (De Strooper & Karan, 2016).

2.1.3.1 Astrocytes in the neuroinflammation

Astrocytes or astroglia are a type of glial cells in the brain and spinal cord as microglia. They have the largest number in the brain (Freeman & Rowitch, 2013). By delivering cholesterol to neurons and other glial cells, astrocytes can control cell signaling in the central nervous system (CNS) (Wang et al., 2021). In healthy human brains, astrocytes have a characteristic spongiform morphology and play an essential role in the CNS, and they have various functions like maintaining synaptic activity, eliminating waste, including A β , and repairing neuronal damages (Ransom & Ransom, 2011, Jessen et al., 2015, Frost & Li, 2017).

Astrocytes are the central players in maintaining synapses and are involved in various signaling systems like glutamate. They assist in the removal of A β by the calcium concentration in response to neurotransmitter release (De Strooper & Karan, 2016). Neuropsychiatric diseases like AD may be driven by changes in astrocyte morphology and function, and the A β accumulation may trigger astrocytes to stop functioning normally, which would exacerbate the disease. A β interacts with the receptors like P2Y1 (Delekate et al., 2014), α 7-nAChRs (Xiu et al., 2005), and mGluR5 (Grolla et al., 2013) on the surface of astrocytes, causing calcium entry, the abnormal calcium concentration then strongly hinders neurotransmitter release from astroglia, which has a significant negative impact on synaptic transmission and neuronal homeostasis and ultimately leads to neurodegeneration (González-Reyes, 2017). In addition to the signaling assistance function of astrocytes, they are involved in the system of A β removal as well. Numerous proteases, including NEP, are produced by

astrocytes and are capable of cleaving the A β during AD (Mulder et al., 2012). Some proteases have the ability to hydrolyze A β oligomers, and some astrocyte-secreted proteases can degrade plaques in their monomeric and fibrillar extracellular forms (Ries & Sastre, 2016). In the AD mouse model, Yan et al. (2006) discovered that the number of proteases expressed by astrocytes dramatically rises around A β plaques.

Inflammation is a biological response of the human body to potentially hazardous stimuli, such as toxic A β oligomers. It is a protective mechanism by removing potentially dangerous components and repairing tissue. Compared to other parts of the human body, the CNS in the brain and spinal cord has a distinct inflammatory response (Ferrero-Miliani, 2007). The difference is caused by the blood-brain barrier (BBB), which is a highly selective semipermeable border to block leukocytes (Ransohoff et al., 2015). When A β plaques develop in the brain, neuroinflammation activates to protect the brain cells due to the interactions between glial cells. Although inflammation serves as a form of defense mechanism, it may ultimately harm the brain since it disrupts the cellular environment. Such aberrant neuroinflammation is typically linked to neurodegenerative disorders, including AD (González-Reyes, 2017). When there is damage in the brain, astrocytes are stimulated to surround the damaged brain region. In AD, activated astrocytes, also known as reactive astrocytes, are clustered around the plaques to assault A β plaques, and these cells become neurotoxic via interacting with some proinflammatory cytokines like IFN- γ (Hashioka et al., 2021). The concentration of proinflammatory cytokines is significantly increased after A β accumulation, and Liddelow et al. (2017) discovered that microglia, another type of glial cells around plaques, play a key role in activating astrocytes.

2.1.3.2 Microglia response in inflammation

About 7% of the brain's cells are microglia, a type of glial cell that serves as a macrophage cell in the CNS. Microglia play a vital role in the health of the entire brain; they build up a network of glial cells and are continually removing plaques from the central nervous system (Gehrman et al., 1995). Thus, in the early stages of AD, the proliferation and activation of microglia are frequently localized around the A β to phagocytose plaques (Villegas-Llerena et al., 2015).

In the recent ten years, genome-wide association studies (GWASs) have discovered more than 20 genetic loci that implicate the high risk the AD, such as the APOE gene, which explains a significant

part of the heritable risk for sporadic AD (Lambert et al., 2013). Along with these genetic variants, the triggering receptor expressed on myeloid cells 2 (TREM2) gene mutation nearly triples the risk of AD (Guerreiro et al., 2013; Jonsson et al., 2013). Through its interaction with the activating adaptor protein DAP12, TREM2 serves as a cell surface receptor on microglia, and it activates the signaling to boost cellular chemotaxis, phagocytosis of A β , and proliferation (Wang et al., 2015; Mazaheri et al., 2017). TREM2 is necessary for microglial phagocytosis of a variety of substrates, including A β and other lipoproteins. When A β aggregates with lipoproteins like LDL, they are swallowed up by microglia considerably more effectively (Wang et al., 2015). Importantly, there was reduced indication of A β internalization in vivo and in vitro in microglia with missing TREM2 (Yeh et al., 2016). The mutations in TREM2 disrupted the interaction of TREM2 with its lipoprotein ligands like LDL. Some transgenic Trem2-deletion mice models have been examined to establish the function of Trem2 in vivo. A remarkable drop in the population of microglia was observed around the A β plaques, along with a significant reduction in inflammatory cytokines. TREM2 thus contributes significantly to the inflammatory response of the microglia (Ulrich et al., 2014, Wang et al., 2015). In addition to phagocytosis, microglia surround A β plaques with a protective barrier in an attempt to compress plaques and halt the rapid accumulation of A β . When the toxic A β oligomers proceed to deposit in the brain, microglia have a preventative function to surround and control the development of A β plaques and condense plaques into a less toxic formation (Condello et al., 2015).

2.1.3.3 Is inflammation beneficial or harmful?

Inflammation is the fundamental factor in AD development, and it shows various effects at different phases of AD. Inflammation might be a protective response to clear the toxic A β plaques, but abnormal inflammation could be damaging to brain cells. Astroglia and microglia are the main cells that participate in the neuroinflammation in the CNS. The accumulation of A β as well as the interactions between cells around A β plaques, are able to exacerbate inflammation (González-Reyes, 2017). Therefore, investigating cellular reactions and interactions at different phases of AD is crucial. In addition, aging is a risk factor as well for abnormal inflammation, Cribbs et al. (2012) discovered that the expression of genes related to inflammation increases significantly with increasing age.

2.1.4 The multicellular gene co-expression network

To better understand the complex, multi-cellular, inflammatory cascade initiated by A β pathology in the amyloid cascade hypothesis, the changes that occur at the molecular level merit more investigation. Chen et al. (2020) utilized special transcriptomics (ST) and in situ sequencing (ISS) to monitor transcriptional changes happening around the A β plaque niche in an AD mouse model. Multiple cell types co-localizing in the A β niche were observed, and two multicellular gene co-expression networks were generated.

Chen et al. (2020) built a network consisting of 57 plaque-induced genes (PIGs), which represents intercellular crosstalk between cells around the plaques, particularly between astrocytes and microglia. For example, PIGs include a secreted protein called C1q that is essential to initiate the cascade or trigger inflammation of astrocytes. The regulator APOE of C1q is present in PIGs as well, which controls the activation of C1q (Yin et al., 2019). The activation of the astroglia may result from interactions between these proteins, which may also start the cascade.

Another gene co-expression network created by Chen et al. (2020) was primarily made up of myelin-related oligodendrocyte genes (OLIGs). The majority of genes in this network are involved in myelin processing, which has an impact on neuronal signaling. The transcriptomic profile of these genes has been shown to be dysregulated by the A β accumulation. Therefore, the crosstalk between cells like astrocytes and microglia was found in the multicellular inflammatory environment around A β plaques. Further investigations, such as integrating single-cell transcriptomics to comprehend the crosstalk between these cells and how they can contribute to the cascade in AD, which would be useful for the potential of therapies.

2.2 Single-cell Transcriptomics

Single-cell transcriptomics studies are appropriate for this investigation of cell-cell communication since they focus on the gene expression of individual cells. By employing various techniques to analyze the expression pattern of various genes, it is feasible to find signaling pathways between cells surrounding A β plaques as well as some interesting ligand-receptor pairs that could be exploited in biological and medical research. Several methods like CellphoneDB and NicheNet could be applied

in this study to find these intriguing ligand-receptor pairings involved in cell-cell communication.

2.2.1 CellphoneDB

CellphoneDB is a publicly available repository to study the cellular interactions mediated by ligand-receptor complexes. It is a repository of ligands, receptors, and their interactions in humans. An unusual feature of this repository is that heteromeric complexes can be represented accurately in CellphoneDB by taking into account the subunit architecture of both ligands and receptors, which makes it possible to analyze cellular communications between multi-subunit protein complexes (Efremova et al., 2020).

2.1.1.1 The mean expression of interactions

CellphoneDB has a comprehensive database, which consists of proteins, steroid hormones, neurotransmitters, and other organic compounds like histamine. The majority of these compounds are annotated as ligands, receptors, or adhesion molecules, while the rest of them are assembled into heterodimeric complexes as subunits instead of working as single partners in the interaction (Figure 1).

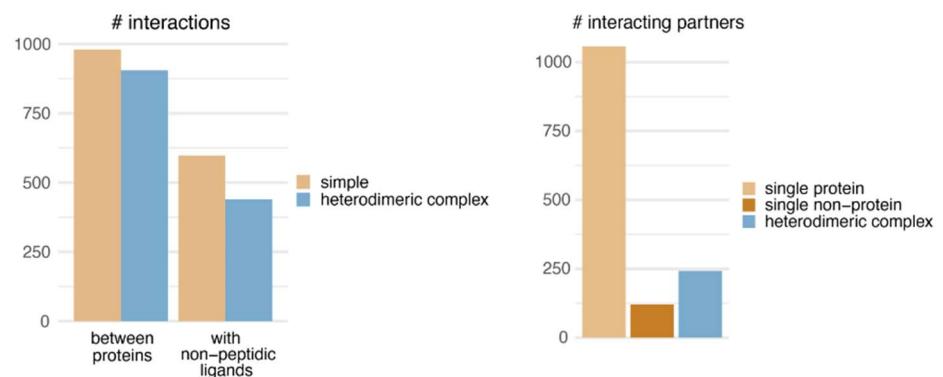


Figure 1: Interactions between proteins or non-proteins such as steroid hormones, small molecules, or neurotransmitters are stored in the CellphoneDB database. CellphoneDB has a total of 2923 interactions in its latest database (v4.1.0), including roughly 1000 interactions involving non-peptidic ligands (Left). The majority of the interaction partners in all interactions included in CellphonDB are single proteins, with the remainder being single non-proteins and heterodimeric complexes (Right). (Efremova et al., 2021).

In CellphoneDB, the measurement used to determine if the cell-to-cell communication is active is based on the gene expression of all elements involved in the interaction, including single proteins, single non-proteins, and heterodimeric complexes. By averaging the expression of each gene

participant in the corresponding producing cells, the mean expression of a simple interaction for each potential pair of interacting cells is generated. When the analysis includes heteromeric complexes, the member of the complex with the lowest expression is used to calculate the mean expression.

2.1.1.2 The null distribution and p-value

CellphoneDB adopts a non-parametric methodology to assign a p-value to the mean expression values of ligand-receptor pairs, aiming to offer statistical inference. In essence, it calculates the p-value by comparing the interaction's mean expression value with the null distribution specific to each sender-receiver pair. In short, CellphoneDB employs a random permutation of all cell type labels 1,000 times to calculate the mean of the average expression level of the interaction in each type of sender-receiver pair (Figure 2). Finally, A null distribution is constructed for each interaction in each pairwise comparison between two cell types, and the p-value for the likelihood of cell-type specificity of a given receptor–ligand complex can be calculated by the null hypothesis that the mean expression of this interaction in this sender-receiver pair is equal to the mean of the null distribution (Figure 2). If the multi-subunit heteromeric complexes are engaged in the interaction, random shuffling is performed on the member of the complex with the lowest expression.

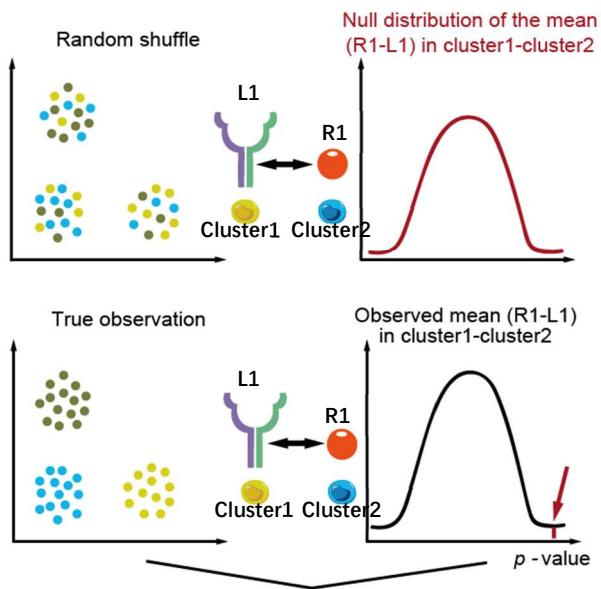


Figure 2: The null distribution of the mean expression of a given interaction between two cell types. In the random shuffle, cell-type labels are permuted 1000 times, and each time the mean expression of R1-L1 (receptor-ligand) in cluster1-cluster2 is calculated. In true observation, the p-value of the observed mean expression of R1-L1 in the null distribution would be calculated (Efremova et al., 2021).

2.1.1.3 The significant interactions

CellphoneDB contains all cell-cell interactions that could possibly occur in the dataset, and it prioritizes interactions that are highly enriched between cell types based on the number of significant pairs. Besides, CellphoneDB is able to reduce the number of candidate interactions if a list of differentially expressed genes is provided, which will remove candidate interactions without these provided genes. However, cells from the single-cell transcriptomics data could come from different developmental phases in various vitro systems, so CellphoneDB might give the wrong answer without additional information because cell types that do not co-appear in time and space cannot interact. To solve this problem, CellphoneDB allows users to divide cells into different microenvironments as the prior information.

2.2.2 NicheNet

NicheNet is a computational framework used for investigating intercellular interactions, based on the gene expression profiles of cells in humans or mice (Browaeys et al., 2020). NicheNet is able to exploit transcriptomics data as the input to build a model to predict the ligand-receptor signaling that is the potential to influence gene expression. Compared to other computational methods like CellphoneDB, NicheNet illustrates the influence of ligand-receptor interactions on target gene expression by involving intracellular signaling (Browaeys et al., 2020; Efremova et al., 2020).

2.2.2.1 Integrated networks

In NicheNet, interactions were collected from a variety of complimentary ligand-receptor, signaling, and gene-regulatory data sources to build several integrated networks (Figure 3). Every data source was transformed into a source-specific directed graph, in which each node represents a protein or a regulated gene, and each edge represents an interaction found in this data source. After that, each graph had a corresponding adjacency matrix generated, and each matrix had a weight given to it to build a weighted sum of adjacency matrices:

$$W = \sum_{i=1}^n (\omega_i A_i) \quad (1)$$

n: the number of open data sources.

ω_i : the weight given to a source-specific network i, which indicates how significant of a contribution the network i has on the performance of the final model.

A_i : the corresponding adjacency matrix for a source-specific network i , in which each entry represents the interaction between nodes.

In the end, a hub correction factor was utilized to reduce the weight of edges directed to hubs proportional to their indegree in order to reduce the potential detrimental impact of over-dominant hubs on the final model:

$$W_{corrected} = WD^{-h} \quad (2)$$

D : the diagonal matrix that indicates each node's indegree.

h : the hub correction factor to correct the weight matrix.

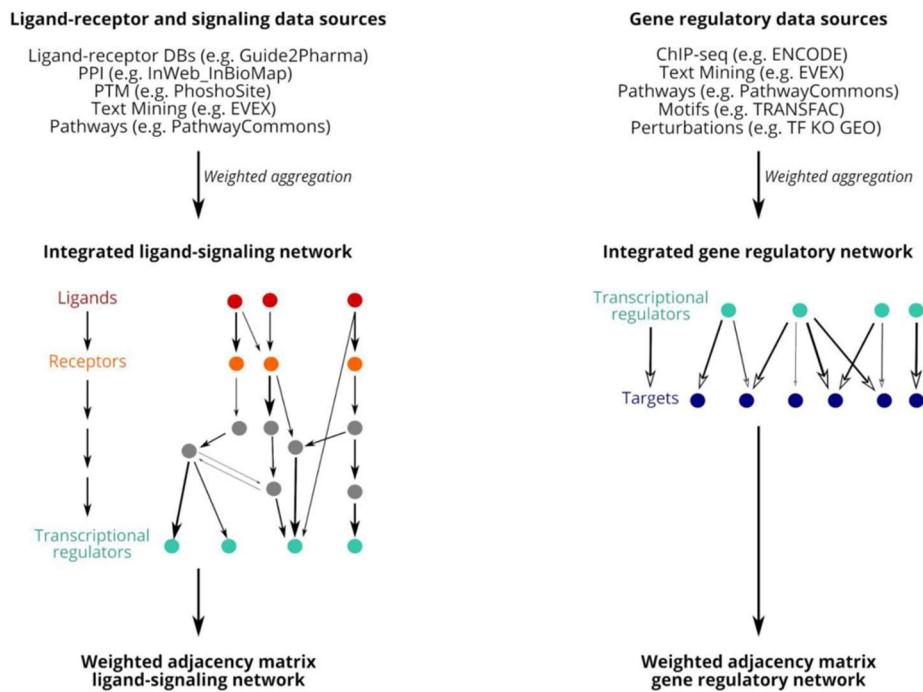


Figure 3: The generation paths of two weighted networks. The path on the left shows the process to generate the weighted adjacency matrix of the ligand-signaling network, including protein-protein interactions consisting of the signaling path from ligands to the transcriptional regulators. The path on the right shows the process of generating the weighted adjacency matrix of the gene regulatory network, which only includes interactions between regulators and target genes (Browaeys et al., 2020).

2.2.2.2 The prior model of ligand-target matrix

The prior model was built based on integrated networks generated in the NicheNet, which indicates the regulatory potential scores for each ligand-target pair (Figure 4). For every ligand, a signaling importance score was calculated for every gene by applying Personalized PageRank (PPR) on the weighted ligand-signaling network. The PPR value of each ligand-gene pair represents the probability of signaling to start from a particular ligand to a transcriptional regulator, so a ligand-regulator signaling importance scores matrix was created.

By multiplying the ligand-regulator signaling matrix with the weighted gene regulatory network, the ligand-target matrix was obtained, in which each entry represents a regulatory potential score corresponding to the confidence that a particular ligand can regulate the expression of a particular target gene:

$$l_{ij} = \sum_{k=1}^m (PPR_{jk} \times GRN_{kj}) \quad (3)$$

PPR_{jk} : the significance of transcriptional regulator k in ligand i signaling pathway as assessed by PPR.

GRN_{kj} : the weighted gene regulatory network that indicates the evidence that the expression of target gene j is regulated by the transcriptional regulator k.

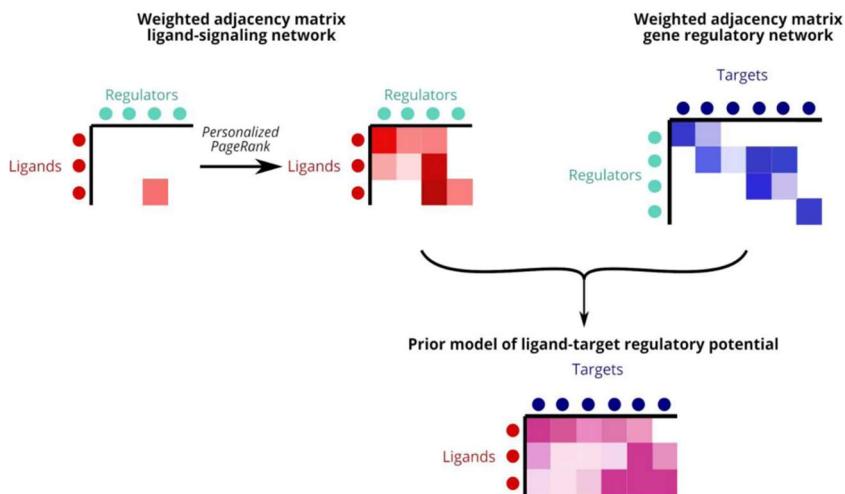


Figure 4: The generation process of the prior model of the ligand-target matrix (Browaeys et al., 2020).

2.2.2.3 Ligand activity prediction

The prior model of the ligand-target matrix can be used to predict each possible ligand expressed in the sender cells. First, NicheNet binarizes all targets based on differential expression analysis of interest. This involves assigning one to the significantly differentially expressed genes as one, while assigning zero to all other genes. Then the prior model performs a binary classification on this binary data to predict which ligand is active or not, using the regulator potential scores of each ligand. NicheNet performs several measurements to make the ligand activity prediction, including AUROC (area under the receiver operating characteristic curve), AUPR (area under the precision/recall curve), and Pearson correlation coefficient. After comparison with the validation datasets, the Pearson correlation coefficient is used as the ligand activity to rank ligands and indicate which ligand is more likely to be responsible for the transcriptional response. If a ligand has a relatively high ligand activity, it is possible that there exists cell-cell communication between the receiver cells and the sender cells

that have the high expression of this ligand.

After obtaining top-ranked ligands, the regulatory potential scores between these ligands and target genes of interest can then be generated as weights and displayed in the heatmap. The genes of interest denote the genes that are part of the target gene set and the most strongly predicted targets of at least one of the top-ranked ligands, which needs the prior model of ligand-target loaded.

2.2.2.4 Further analysis

One of the further analyses is identifying which receptors of the receiver cell population might be able to bind to the prioritized ligands from the sender cell population. In this analysis, the best upstream receptors are selected from the ligand-receptor network instead of the best upstream ligands, and the weights of the ligand-receptor interactions are determined as in the NichiNet model. The hierarchical clustering is performed to order the potential interactions between ligands and receptors, and the results can be visualized as a heatmap.

Another frequently used approach is visualizing the expression of ligands and target genes in a composite heatmap. NicheNet ranks ligand activities exclusively by their ability to regulate downstream target genes, irrespective of the expression levels of the ligands themselves. Therefore, it is vital to have post hoc visualization which combines the expression of ligand, target genes, the ligand activity, and the regulatory potential of ligand-target side-by-side in heatmap format, to better understand the validity of the prediction.

3 Materials and Methods

3.1 Sample collection and data acquisition

A total of 25 people were chosen for this study from the Netherlands Brain Bank (NBB). Individuals with genotyping, bulk-tissue transcriptome profiling, clinical and pathological evaluations were included in the sample collection. They were divided into two groups: the "no-pathology" group, which consisted of 9 people with no or little pathology, and the "AD-pathology" group, which included 16 people with varying degrees of amyloid-beta and neurofibrillary tangle deposition. All individuals in the no-pathology group had been clinically diagnosed as cognitively normal, while the AD-pathology group included eight individuals with clinically normal cognitive function and eight individuals with clinical cognitive impairment. The distribution of individuals in both groups was balanced across sexes (no-pathology: 4 males, 5 females; AD-pathology: 10 males, 6 females), and the groups were matched in terms of age (average age of the no-pathology group: 83; the average age of the AD-pathology group: 86). Each post-mortem subject had biopsies performed on the superior frontal gyrus of their brain. The samples were stored at -70°C and transported to the De Strooper Lab for further analysis.

3.1.1 Sequencing data collection

To perform Visium Spatial Gene Expression sequencing, frozen brain blocks were cryosectioned at a thickness of 10 µm using a cryostat. One section per donor was mounted onto a Visium Spatial Gene Expression Slide (10x Genomics) for further analysis. 15 sections with a thickness of 50 µm were collected from the same frozen brain blocks for single-nuclei RNA sequencing (snRNA-seq). These sections were stored at -70°C until the snRNA-seq procedure.

3.1.2 Meta-data collection

Several incubation procedures were carried out to identify amyloid-beta and phosphorylated Tau (pTau) on the tissue slices utilized for Visium analysis. Three distinct antibodies, including anti-amyloid-beta 17-24 (4G8), donkey anti-mouse IgG (H+L) Alexa555, and a combination of an Alexa647-conjugated anti-RBFOX3/NeuN antibody (1B7), were treated with the tissues

independently. Finally, DAPI and streptavidin were incubated with the tissues. Amyloid plaques and pTau pathology were identified by applying a pixel classification threshold and an additional size threshold of $100 \mu\text{m}^2$ for amyloid plaques and $1 \mu\text{m}^2$ for pTau pathology, respectively. Based on morphological criteria, various plaque types were classified. After classification, staining intensity analysis was used to annotate diffuse plaques.

3.2 Data preprocessing

3.2.1 Pre-analysis

The snRNA-seq data and Visium data underwent preprocessing, clustering, and integration prior to the commencement of this master thesis project.

Single-cell RNA-seq processing and cell type annotation

The raw FASTQ files were processed using Cell Ranger V6.0.1 (10x Genomics) with default settings. Souporcell was utilized to determine the sample IDs for each cell, and the Cell Ranger output BAM files were used to calculate allele frequency tables for each sequencing library. Subsequently, the count matrices were imported into Scanpy and merged. Cells expressing fewer than 200 genes and genes detected in fewer than 5 cells were filtered out, and Scrublet was employed to remove potential doublets. 7 major cell types were identified by using the expression patterns of cell type marker genes from the study by Leng et al. (2021), and the identified major cell clusters were further subclustered, curated, and manually annotated into a total of 54 subtypes.

Visium spatial transcriptomics data analysis

The raw FASTQ files containing Visium reads were processed using Space Ranger 1.2 (10x Genomics) with the 10X distributed GRCh38 reference (version 2020-A) to maintain consistency with the single-nuclei (SN) data. To spatially map brain cell states identified through the SN analysis onto the Visium data, we employed the cell2location method.

Differential cell abundance analysis

In order to investigate how cell abundances change in response to plaque and pTAU pathology, as well as to compare these changes between the AD and RES cohorts, we performed a differential abundance (DA) analysis using the edgeR package. Similar analyses were also applied to the different

plaque types, which were categorized based on their morphology during the image analysis. These plaque types include core plaques, neuritic plaques, diffused plaques and unclassified plaques. This allows us to assess the differential changes in cell abundances in response to plaque and pTAU pathology and compare these changes between AD and RES cohorts (Supplement Figure 21). In DA analysis, five different cell types, such as Astro_0, micro_5, oligo_0, oligo_3 and oligo_4, were clustered into one group, which named plaque associated cells (PACs). These PACs were used to do single-cell transcriptomics analysis in this thesis.

3.2.2 The Seurat object

After the pre-analysis, the Seurat object with clean data was obtained. There is only an ‘RNA’ assay in the object, which contains the log-transformed gene expression information (RNA-seq data) of every single cell. The ‘RNA’ assay is represented as a gene expression matrix, where each row represents a gene and each column represents a cell. There are 34766 features (genes) for 112698 cells in the gene matrix. The metadata of the Seurat object contains the meta-information about each cell, such as the number of RNA features detected (nFeature_RNA) and the cell type. One additional information, which is ‘first_subtypes’, was added for every single cell in the pre-analysis, indicating the subtypes of each cell type. For example, there exist micro_0, micro_1, micro_2, micro_3, micro_4, micro_5, and micro_6 for the cell type micro. These subtypes were used to replace cell type in this thesis. Another important piece of information for each cell in the metadata is ‘Phenotype’, where ‘AD’ represents the plaque is present and the Alzheimer’s symptoms occur; ‘RES’ represent the plaque is present but the Alzheimer’s symptoms are ‘resilient’; ‘ND’ represent control samples without plaque and Alzheimer’s symptoms.

3.2.3 The AnnData object

Three AnnData objects were created using the counts matrix generated from the Seurat object. Each AnnData object represents one phenotype, such as ‘AD’, ‘RES’ and ‘ND’. The metadata used for each AnnData object was created, where each row represents a single cell and the first column indicates the subtypes of the cell.

3.4 Differential expression analysis

Differential expression (DE) analysis was done by running *FindMarkers()* in the package Seurat. The default statistical hypothesis used in *FindMarkers()* is the Wilcoxon Test, which is a non-parametric test to find differentially expressed genes in two conditions (Wilcoxon, 1945). The Wilcoxon Test has a null hypothesis for each gene to calculate the p-value, which is that the gene expressions have the same population in two conditions (two groups). Firstly, the gene expressions in two conditions are combined and ranked from the smallest to the largest. The rank sum is calculated for each group (T1 and T2):

$$T_1 + T_2 = \frac{N * (N + 1)}{2} \quad (4)$$

With N is the total sample number. Under the null hypothesis, T1 is assumed to follow its distribution:

$$\mu = \frac{n_1 * (N + 1)}{2} \quad (5)$$

$$\sigma = \frac{n_1 * n_2 (N + 1)}{12} \quad (6)$$

With μ is the mean and σ is the standard variance. T1 is then compared to each gene's distribution under the null hypothesis to determine the p-value for that gene.

3.4 CellphoneDB analysis

CellphoneDB has three methods to analyze scRNA-seq data, including *cpdb_analysis_method()*, *cpdb_statistical_analysis_method()* and *cpdb_degs_analysis_method()*. The function *cpdb_analysis_method()* only calculates the mean expression of interactions, whereas the function *cpdb_statistical_analysis_method()* includes the mean expression calculation and the analysis of significance to find cell-specificity of each interaction. The function *cpdb_degs_analysis_method()* has a filter to retrieve interactions where at least one partner (ligand or receptor) is differentially expressed, which is not helpful in this thesis since those ligands and receptors that are not differentially expressed could also be involved in CCCs. Therefore, *cpdb_statistical_analysis_method()* was used in this thesis to find interesting interactions.

cpdb_statistical_analysis_method() is a statistical inference of interaction specificity. *cpdb_statistical_analysis_method()* was applied to the AnnData object containing the scRNA-seq data to predict enriched ligand-receptor interactions between two cell types based on the expression of a receptor by one cell type and a ligand by another cell type:

$$\text{mean expression} = \frac{\text{expression}_{\text{ligand}} + \text{expression}_{\text{receptor}}}{2} \quad (7)$$

If the multi-subunit heteromeric complex is involved in the interaction as ligand or receptor, the expression of ligand or receptor is equal to the minimum expression of subunits:

$$\text{expression}_{\text{ligand/receptor}} = \min(\text{subunit}_1, \text{subunit}_2, \text{subunit}_3 \dots) \quad (8)$$

After generating the mean expression matrix, *cpdb_statistical_analysis_method()* computed the *p-value* for the likelihood of cell-type specificity of each interaction. Before *p-value* calculation, ligands and receptors that were expressed in fewer than a threshold percentage of the cells in the given cluster were removed (threshold = 0.1), and all of the complex's subunits had to be expressed above this threshold for multi-unit heteromeric complexes to pass this filter. The cluster label of each cell was then randomly permuted 1000 times, and a null distribution was created by calculating the mean expression of each interaction in each pairwise comparison between two cell types. The *p-value* for the likelihood of cell-type specificity of the interaction in one cell-pair was determined by counting

the number of mean expressions that are higher than the actual mean expression over all 1000 trials:

$$p = \frac{\text{the number of mean expression higher than the actual mean expression}}{1000} \quad (9)$$

The *p-values* have a range from 0 to 1 in the algorithm of CellphoneDB, which could lead to misunderstanding because the null distribution only allowed *p-values* to be close to 0 or 1. Therefore, we added 0.0001 to all 0 and subtracted 0.0001 from all non-zero values in the *p-value* matrix, bringing the range of *p-values* to 0.0001 to 0.9999.

False discovery rate (FDR) correction was applied by the function *multitest()* from package *statsmodels()* in the end, which provided the *adjusted p-value* for each *p-value*. The method used was '*fdr_bh*', which represents the Benjamini-Hochberg Procedure in the FDR correction. The FDR was controlled at a level alpha (alpha = 0.05) and each p-value was given a rank from smallest to largest by the Benjamini-Hochberg procedure:

$$\text{adjusted } p \text{ value} = p \text{ value} * \frac{\text{rank}}{\text{total length}} \quad (10)$$

3.5 NicheNet analysis

NicheNet provides a function `predict_ligand_activities()` to predict whether a ligand is active or not and rank all ligands based on the ligand activity, which is the Pearson correlation coefficient between the ligand-target weights and the responses of differentially expressed targets. For each ligand expressed in given sender cell types, the ligand-target weights were selected from the prior model offered by NicheNet, and the responses of differentially expressed targets were binary data generated from the Seurat object containing the scRNA-seq data (response = 1 if the gene is regarded as DE gene; response = 0 if the gene is not regarded as DE gene).

Before the function `predict_ligand_activities()` was performed, background genes and DEGs in the receiver needed to be determined. Controlling the quantity of both gene-sets is crucial since the number of background genes and DEGs has a significant impact on ligand prediction. The parameter `pct`, which controls the quantity of background genes, indicated that all background genes were expressed in more receiver cells than a threshold percentage. `p_val_adj` and `avg_log2FC`, two measurements of whether a gene is differentially expressed in two conditions, were used to identify the DEGs in the output of DE analysis. Based on the suggestion from the NicheNet package author, the number of DEGs should be in the range of 20-1000 and the number of background genes should be larger than 10000.

The function `predict_ligand_activities()` utilized `cor.test()` to determine whether the correlation between the ligand-target weights and the responses is significant by computing the pearson correlation coefficient (`cor_p`):

$$cor_p = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{\sum(x - m_x)^2 \sum(y - m_y)^2}} \quad (11)$$

Where x are the ligand-target weights, y are the responses of differentially expressed targets, m_x is the mean of x , m_y is the mean of y . The final output of `predict_ligand_activities()` ranked all ligands based on `cor_p`, which is used as the ligand activity in NicheNet.

4 Results

4.1 Celltype-specific ligand-receptor interactions identified by CellphoneDB

Only PACs were selected to be studied in this thesis because these cells were found to be more likely to have CCCs in pre-analysis. In CellphoneDB analysis, celltype-specific interactions in PACs were found through *cpdb_statistical_analysis_method()* this thesis. The mean expressions and p-values of each possible interaction happened were calculated, and some celltype-specific ligand-receptor interactions were found through specific filters and summarized in the end of CellphoneDB analysis.

4.1.1 All possible interactions found in AD, RES and ND by CellphoneDB

The function *cpdb_statistical_analysis_method()* was performed three times to obtain results in three different phenotypes (AD, RES and ND), using the AnnData objects containing the scRNA-seq data in AD, RES and ND. Each time *cpdb_statistical_analysis_method()* generated a mean expression matrix and a p-value matrix for every possible interaction in each cell-type pair. The total number of interactions found in AD was 2851 (1914 interactions included multi-subunit complex and 937 interactions only included simple proteins); The total number of interactions found in RES was 2836 (1912 interactions included multi-subunit complex and 924 interactions only included simple proteins); The total number of interactions found in ND was 2857(1919 interactions included multi-subunit complex and 938 interactions only included simple proteins) (Figure 5). The number of interactions in the three phenotypes were similar, and only a few interactions could not be found in all three phenotypes.

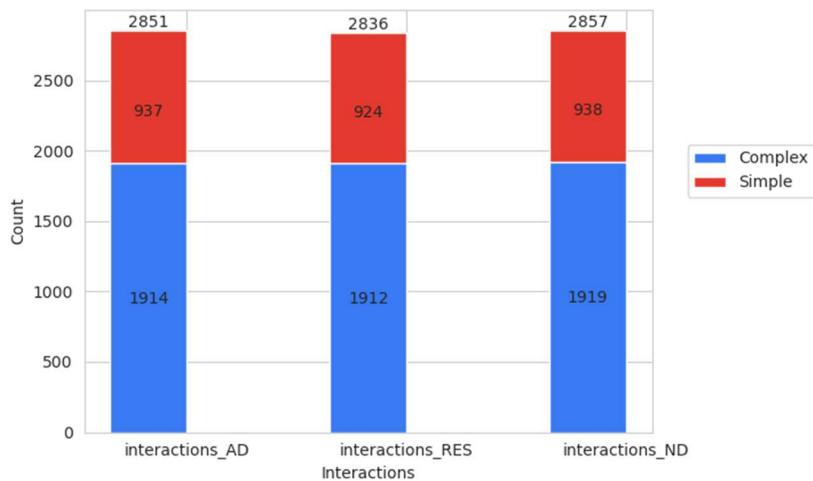


Figure 5: The total number of ligand-receptor interactions in three phenotypes (AD, RES and ND). Out of a total of 2923 interactions collected in the CellphoneDB database, we have summarized the total number of ligand-receptor interactions expressed in our data. The red bar indicates the number of interactions, including the multi-subunit heteromeric complex, whereas the blue indicates the number of interactions only, including simple proteins.

Before selecting cell-type pairs in PACs, the mean expression matrix and p-value matrix included all possible cell-type pairs in the AnnData object that has 57 cell subtypes, so there were 2916 possible cell-type pairs in these two matrices. There was not one interaction found in all cell-type pairs, and the majority of the interactions had mean expressions of 0 and p-values of 0.9999 in the entry of matrices. For each p-value, the adjusted p-value, which was used to do further analysis, was calculated according to the Benjamini-Hochberg Procedure.

4.1.2 Visualization of mean expressions and p-values of all possible interactions

All possible interactions were visualized in the scatter plot to investigate the changes between two phenotypes AD and RES. Four scatter plots of receiver micro_5 were shown as an example in Figure 6, and scatter plots of other PACs were displayed in Supplement Figures 1-4. In Figure 6, it was noticed that the mean expressions of interactions did not change a lot between AD and RES, and some interactions were only found in AD or RES. Besides, the mean expressions did not show a correlation with cell-specificity (The adjusted p-values). Few interactions were only celltype-specific in AD or RES, and more interactions showed cell-specificity in both AD and RES.

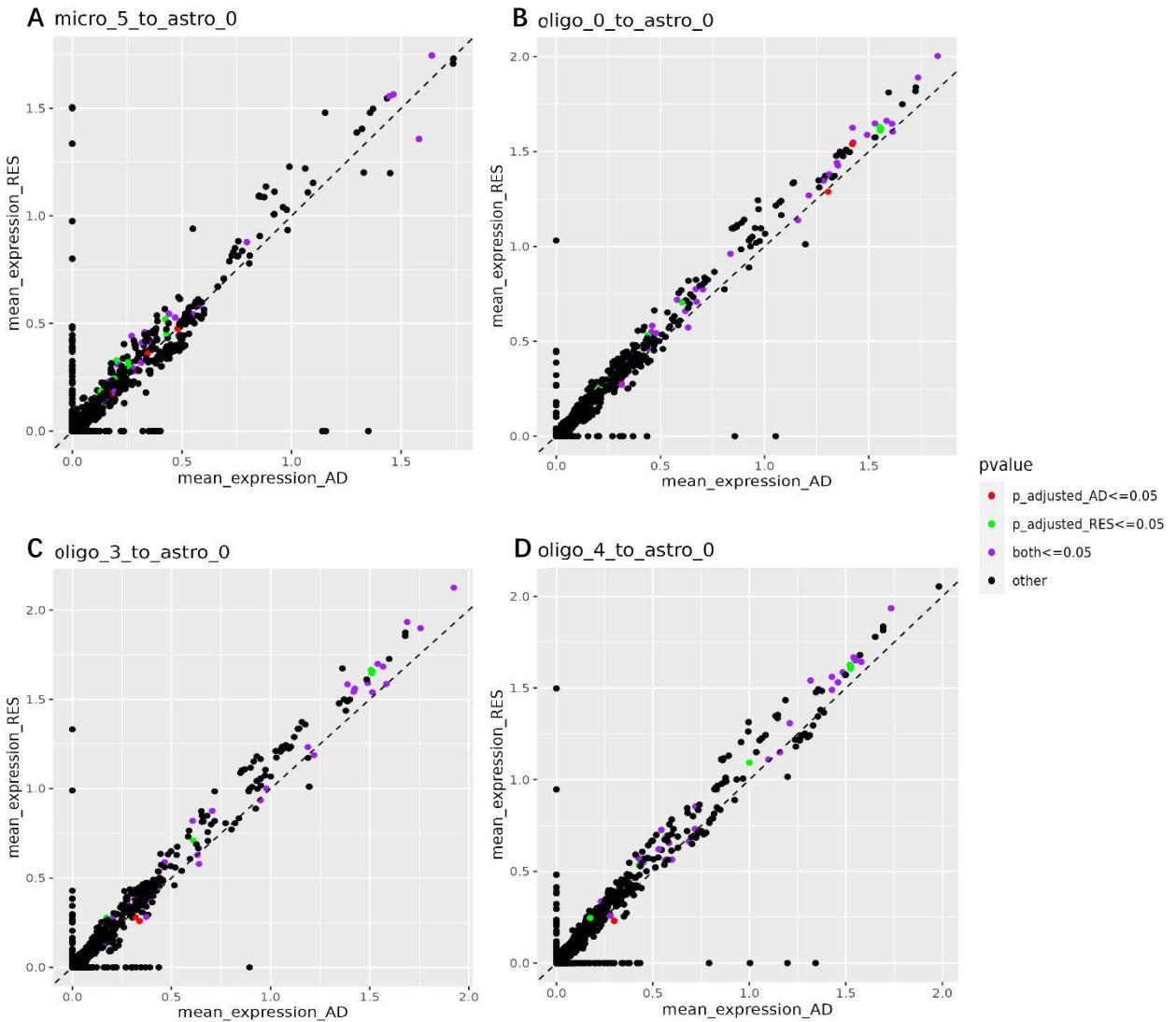


Figure 6: The scatter plots of mean expression of all possible interactions in AD against the mean expression of interactions in RES when receiver cell type was astro_5. The dashed line (slope = 1, intercept = 0) was used to indicate the changes in interactions in AD and RES. The red dots represented interactions that only had $p_{adjusted_AD} \leq 0.05$, the green dots represented interactions that only had $p_{adjusted_RES} \leq 0.05$, the purple dots represented interactions with $p_{adjusted_AD} \leq 0.05$ and $p_{adjusted_RES} \leq 0.05$, and the black dots represented the remaining interactions. (A) interaction found when micro_5 is sender cell-type. (B) interaction found when oligo_0 is sender cell-type. (C) interaction found when oligo_3 is sender cell-type. (D) interaction found when oligo_4 as sender cell-type

4.1.3 548 celltype-specific interactions were found as candidates for CCCs

A filter was then built to drop interactions that were not cell-type specific, which was achieved by setting a threshold on the adjusted p-value. Only interactions showed celltype-specificity in AD and no celltype-specificity in ND (the adjusted p-value ≤ 0.05 & the adjusted p-value ≥ 0.1) or interactions showed celltype-specificity in RES and no celltype-specificity in ND (the adjusted p-value ≤ 0.05 in RES & the adjusted p-value ≥ 0.1) would pass the filter. The number of interactions

decreased to 548 after the filter, which dropped around 80% of non-celltype-specific interactions. For PACs, there were a total of 500 celltype-specific interactions with an adjusted p-value of 0.05 in RES (189 celltype-specific interactions included multi-subunit complex and 311 celltype-specific interactions only included simple proteins), while there were 353 celltype-specific interactions with an adjusted p-value of 0.05 in AD (101 celltype-specific interactions included multi-subunit complex and 252 celltype-specific interactions only included simple proteins) (Figure 7). In the 548 celltype-specific interactions between PACs, 305 interactions had significant adjusted p-values in both AD and RES, as opposed to 48 interactions and 195 interactions, respectively, that only had significant adjusted p-values in AD and RES (Figure 8).

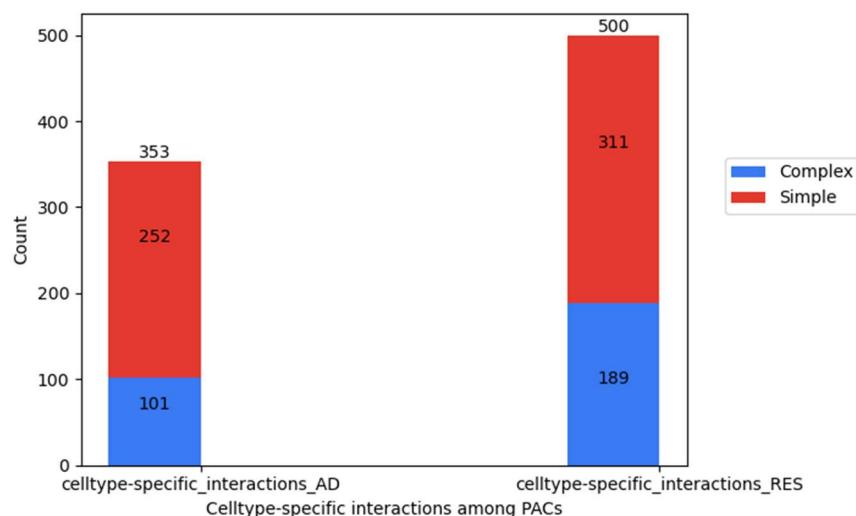


Figure 7: The number of celltype-specific interactions with significant p-values in two phenotypes (AD and RES). The red bar indicates the number of interactions including the multi-subunit heteromeric complex, whereas the blue indicates the number of interactions only including simple proteins.

Celltype-specific interactions between PACs

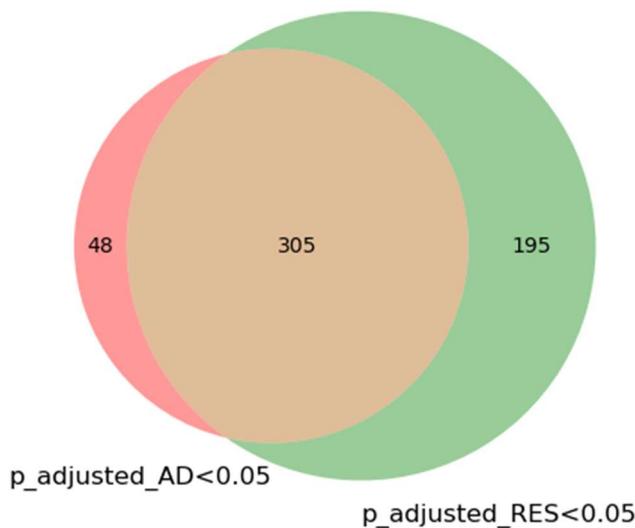


Figure 8: The Venn diagrams of celltype-specific interaction among PACs. The red part on the left indicated celltype-specific interactions only had significant adjusted p-values in AD, the green part on the right indicated celltype-specific interactions only had significant adjusted p-values in RES, and the brown part on the middle indicated celltype-specific interactions had significant adjusted p-values in both AD and RES.

4.1.4 Distribution of celltype-specific interactions in sender-receiver cell pairs

There were 548 celltype-specific interactions in all, of which 195 interactions only show celltype-specificity in RES and 48 interactions only show celltype-specificity in AD. Three heatmaps were constructed to visualize the distribution of these interactions between PACs (Figure 9). AD-celltype-specific interactions were centralized in three cell-type pairs: astro_0-micro_5, astro_0-oligo_4 and oligo_0-micro_5 (Figure 9A); RES-celltype-specific interactions were centralized in one cell-type pair: oligo_3-astro_0, and astro_0 was discovered to communicate with other PACs as sender cells substantially more frequently (Figure 9B); all interactions found were regarded as plaque-celltype-specific interactions, which were centralized in three cell-type pairs: astro_0-micro_5, oligo_3-astro_0 and micro_5-astro_0, and astro_0 was discovered to be commonly used as sender cell or receiver cell in cell-to-cell communications between PACs (Figure 9C).

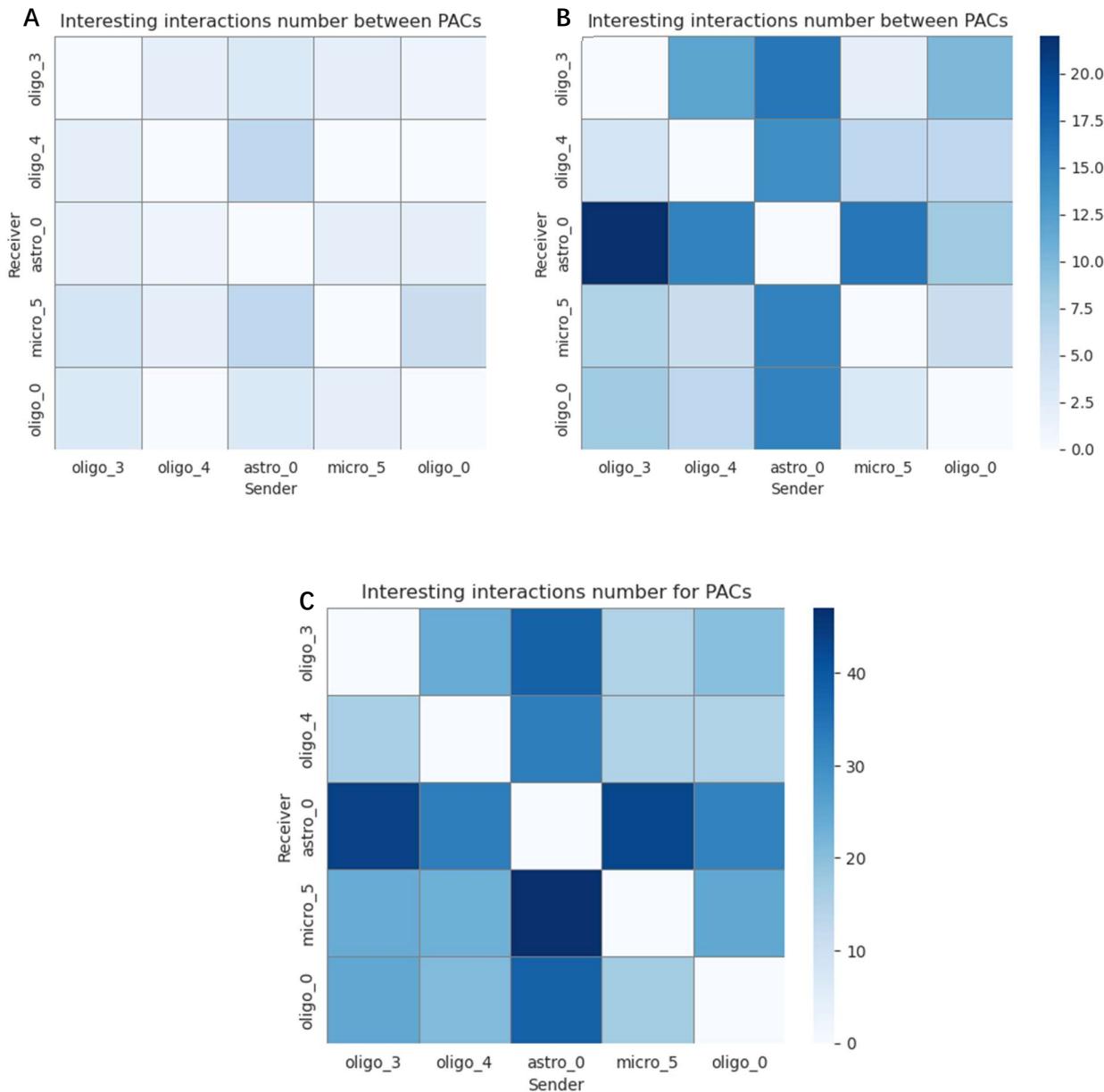


Figure 9: The distribution of celltype-specific interactions in sender-receiver cell pairs. (A)The distribution of AD-celltype-specific interactions between PACs. X-axis represents 5 PACs as sender cells and y-axis represents 5 PACs as receiver cells. The total number of interactions shown in plot-A is 48. (B) The distribution of RES-celltype-specific interactions between PACs and the total interactions number is 195. (C) The distribution of plaque-celltype-specific interactions between PACs and the total interactions number is 548.

4.1.4 All celltype-specific interactions were divided into five receiver groups

Five groups were established from the 548 celltype-specific interactions between PACs: Receiver_astro_0, Receiver_micro_5, Receiver_oligo_0, Receiver_oligo_3 and Receiver_oligo_4. These groups represented which cell type was used as the receiver in the interactions, and the other

PACs were used as the sender. Celltype-specific interactions in group Receiver_astro_0 were visualized in Figure 10, and dot plots of other PACs were displayed in Supplement Figure 5-8. The celltype-specific interactions found in each receiver cell type were counted (Table 1), and the simple ligands and simple receptors for receiver astro_0 were displayed in Figure 11. For the other simple ligands and simple receptors for other receiver cell types, they were displayed in Supplement Figures 17-20.

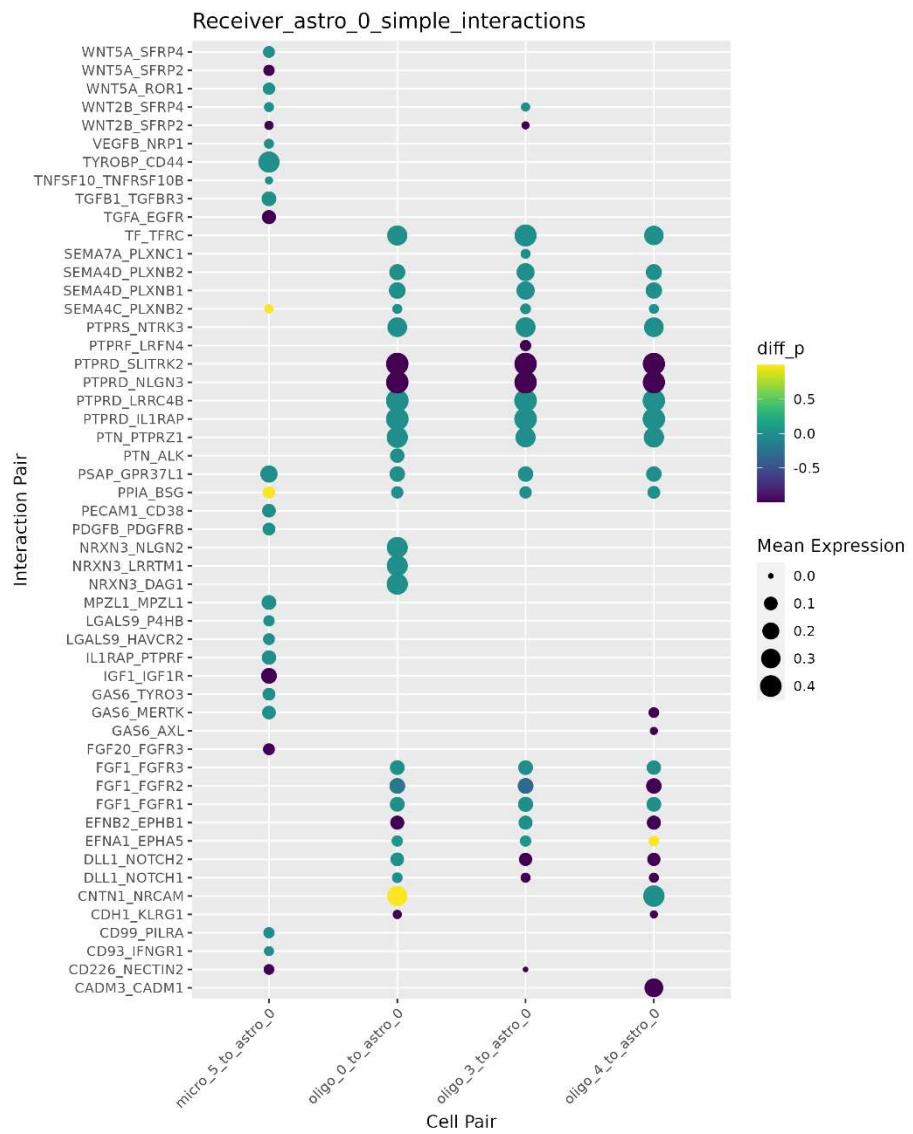
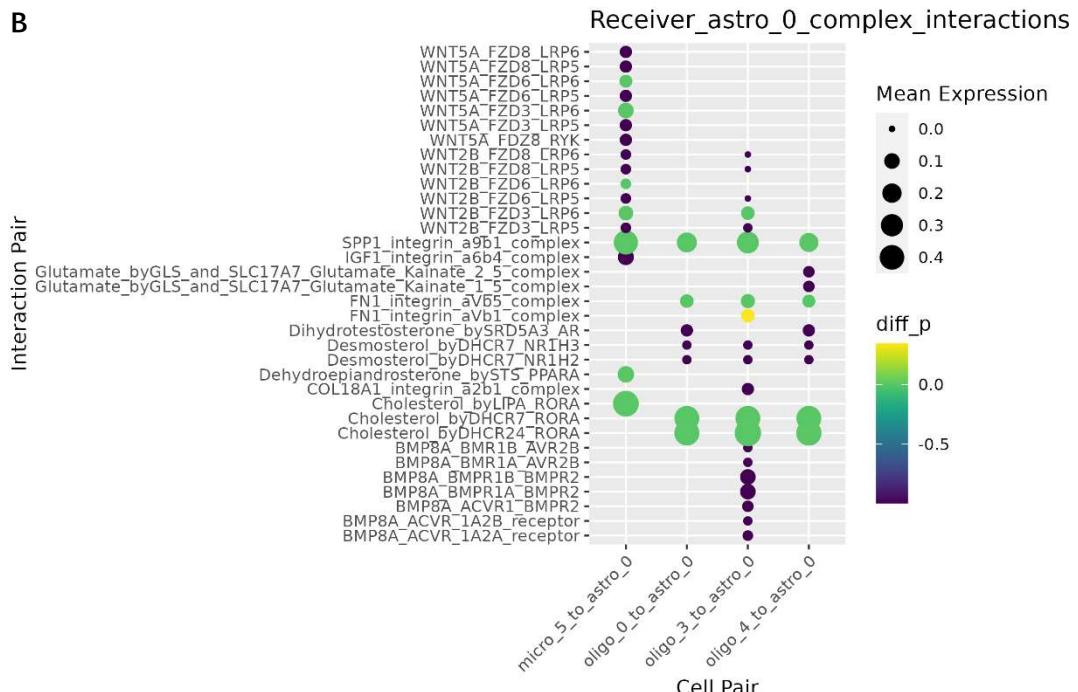
A**B**

Figure 10: Dot plot of celltype-specific interactions when the receiver was astro_0. The other PACs were senders, which were micro_5, oligo_0, oligo_3 and oligo_4. $\text{diff_p} = -(\text{p_adjusted_AD} - \text{p_adjusted_RES})$. Mean expression = $(\text{mean expression in AD} + \text{mean expression in RES} + \text{mean expression in ND})/3$. All of these interactions were celltype-specific in AD or RES, but not in ND. If the color of the dot was yellow, the interaction was celltype-specific in AD, if the color was purple, the interaction was celltype-specific in RES. If the color was green, the interaction was celltype-specific in both AD and RES (A) Dot plot of celltype-specific interactions only containing simple ligands and simple receptors. (B) Dot plot of celltype-specific interactions containing complex.

Table 1: The number of celltype-specific interactions found in each receiver cell type

Receiver	micro_5	astro_0	oligo_0	oligo_3	oligo_4
# Celltype-specific interactions	119	152	101	97	79

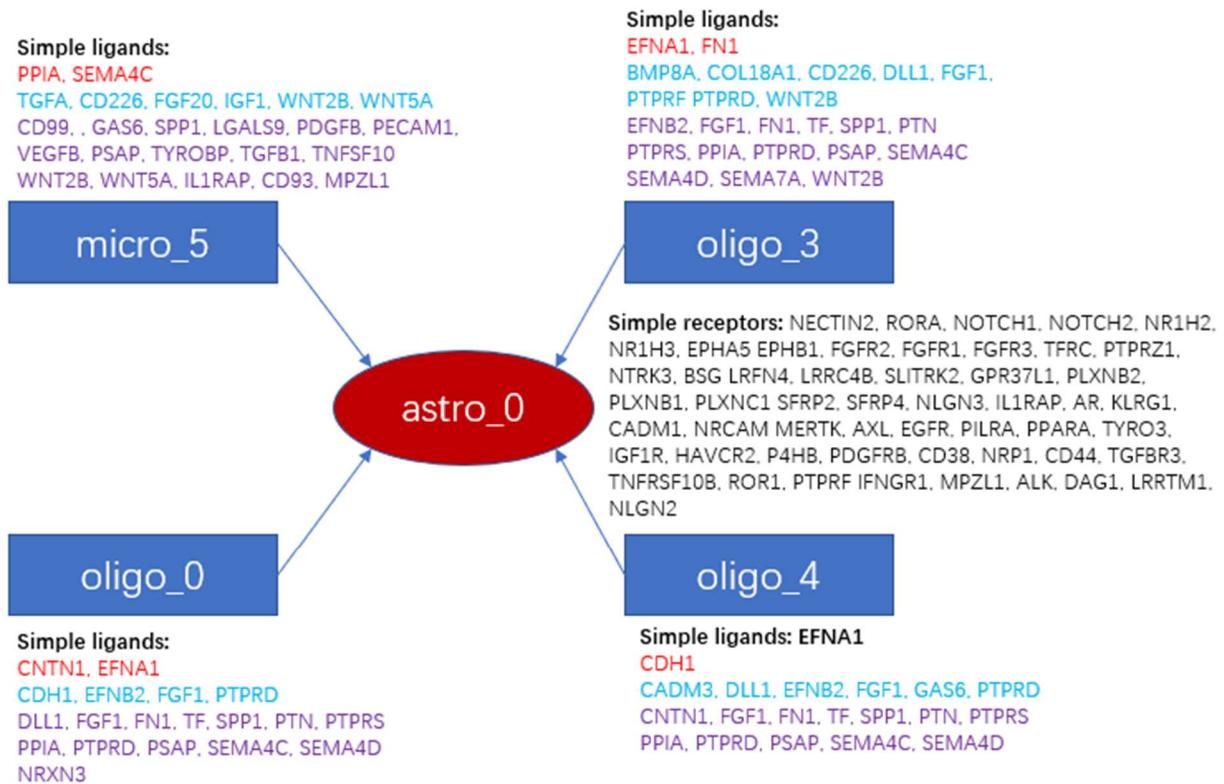


Figure 11: The simple ligands and simple receptors found in celltype-specific interactions when the receiver was astro_0. The red circle represents receivers, the blue squares represent senders. The ligands found in the interactions that were celltype-specific in AD were marked as red; The ligands found in the interactions that were celltype-specific in RES were marked as blue; The ligands found in the interactions that were celltype-specific in both AD and RES were marked as purple.

4.2 Active ligands identified by NicheNet

In contrast to detecting potential ligand-receptor pairs in the CellphoneDB analysis by using the gene expression, active ligands were identified in the NicheNet analysis by using the transcriptional response. One ligand was likely to be active if the majority of its target genes in the signaling downstream were differentially expressed between two conditions (ADvsND or RESvsND). Before running ligand prediction analysis, the quantity control of background genes and DE analysis had to be done. The ligand activities were then visualized and high-activity ligands were selected. Some additional information like the expression and log fold change of ligands were done because more information is needed to find truly active ligands if NicheNet was performed alone.

4.2.1 Control the quantity of background genes to ensure the accuracy of ligand predictions

Too many background genes would make every ligand's activity small, whereas too few background genes would produce extremely high but unreal ligand activity. According to the suggestion from the package author, the numbers of background genes were needed to exceed 10000. Therefore, the quantity control of background genes in the receiver had to be accomplished to ensure the accuracy of ligand prediction before *predict_ligand_activities()* was performed. The parameter *pct* was used to control the number of background genes in each of the five cell types in PACs. As *pct* fell in Figure 12, the quantities of background genes increased significantly for each cell type, with astro_0 showing a much greater abundance than the other cell types. All cell types had background genes over 5000 when *pct* = 0.1; all cell types had background genes larger than 7500 when *pct* = 0.05; and all cell types had background genes larger than 12000 when *pct* = 0.01 (Figure 12). Thus, *pct* = 0.1 was utilized in this thesis to ensure the numbers of background genes exceeded 10000 in every receiver cell type.

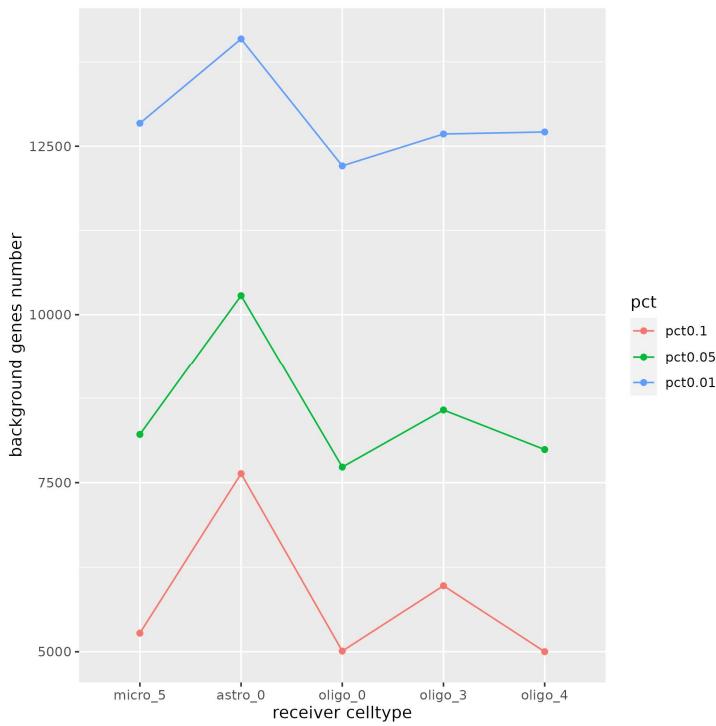
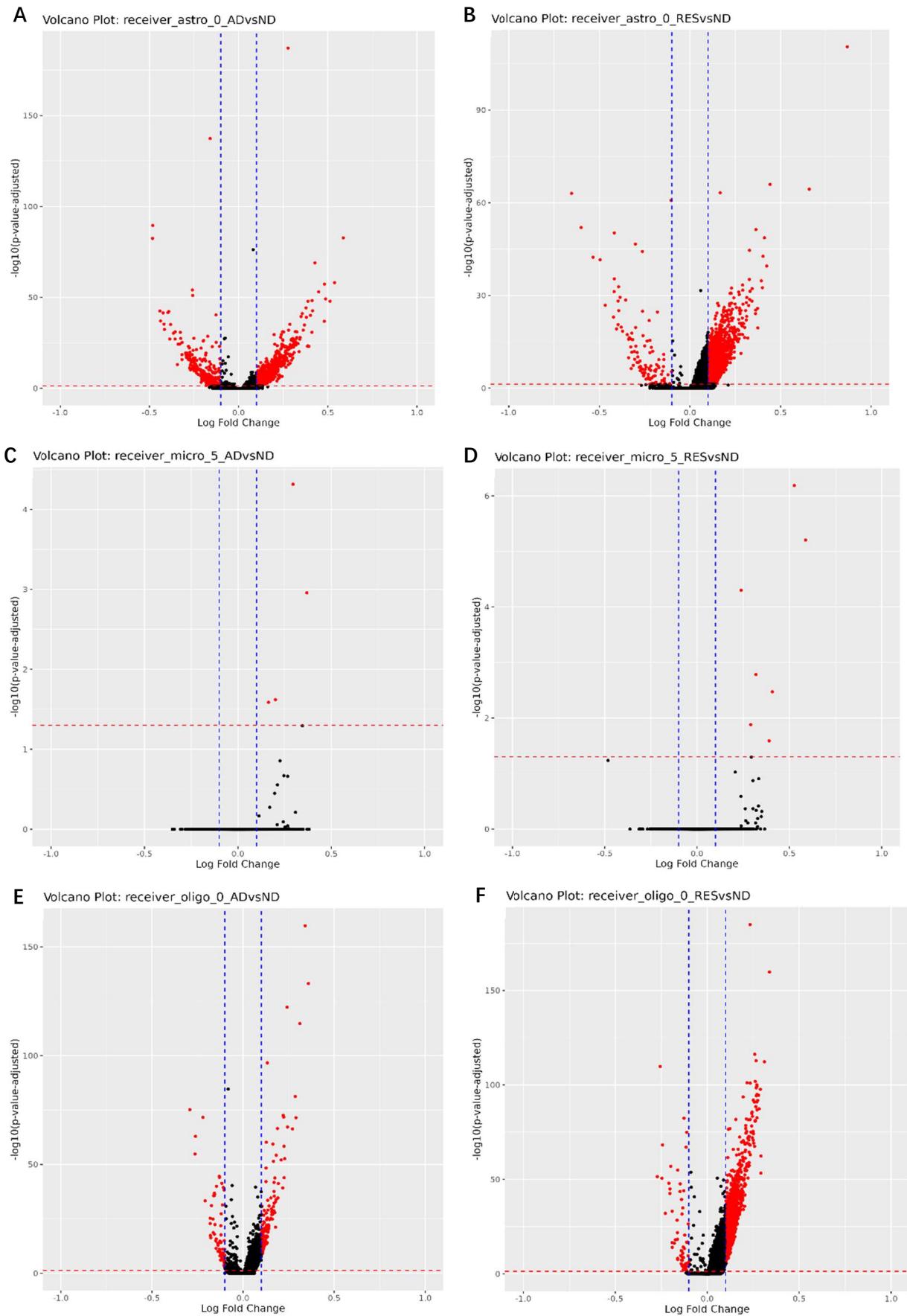


Figure 12: Line plot of background genes number in five cell types (micro_5, astro_0, oligo_0, oligo_3 and oligo_4). The red line represents pct = 0.1, the green represents pct = 0.05, and the blue line represents pct = 0.01

4.2.2 Determine DEGs by DE analysis to ensure the accuracy of ligand predictions

It is essential to determine what are DEGs in the receiver when running NicheNet, which will have a huge impact on the ligand prediction. Too many DEGs might result in a lot of active but unreal ligands, while too few DEGs might lead to false or missing active ligand. According to the suggestion from the package author, the number of DEGs is better to be controlled in the range of 20-1000. Therefore, the *FindMarkers()* was done before the ligand prediction analysis for every cell type in PACs in two combinations of phenotypes: ADvsND and RESvsND . Volcano plots of ADvsND and RESvsND were generated to visualize the DE genes in five cell types. In the DE output of the Wilcoxon test, only genes with $p_val_adj \leq 0.05$ & $abs(avg_log2FC) \geq 0.1$ were annotated as DEGs with significantly differential expression. The number of DEGs found for each receiver cell type in ADvsND and RESvsND was recorded in Table 2. Unfortunately, micro_5 in ADvsND and RESvsND and oligo_4 in ADvsND did not have enough DEGs (< 20), while astro_0 in RESvsND had too many DEGs (>1000). Additionally, astro_0 showed an obviously large amount of DEGs in both ADvsND and RESvsND, whereas micro_5 showed an obviously small amount of DEGs in both ADvsND and RESvsND (Table 2). One interesting finding is that all cell types in PACs showed a

larger number of DEGs in RESvsND than that in ADvsND.



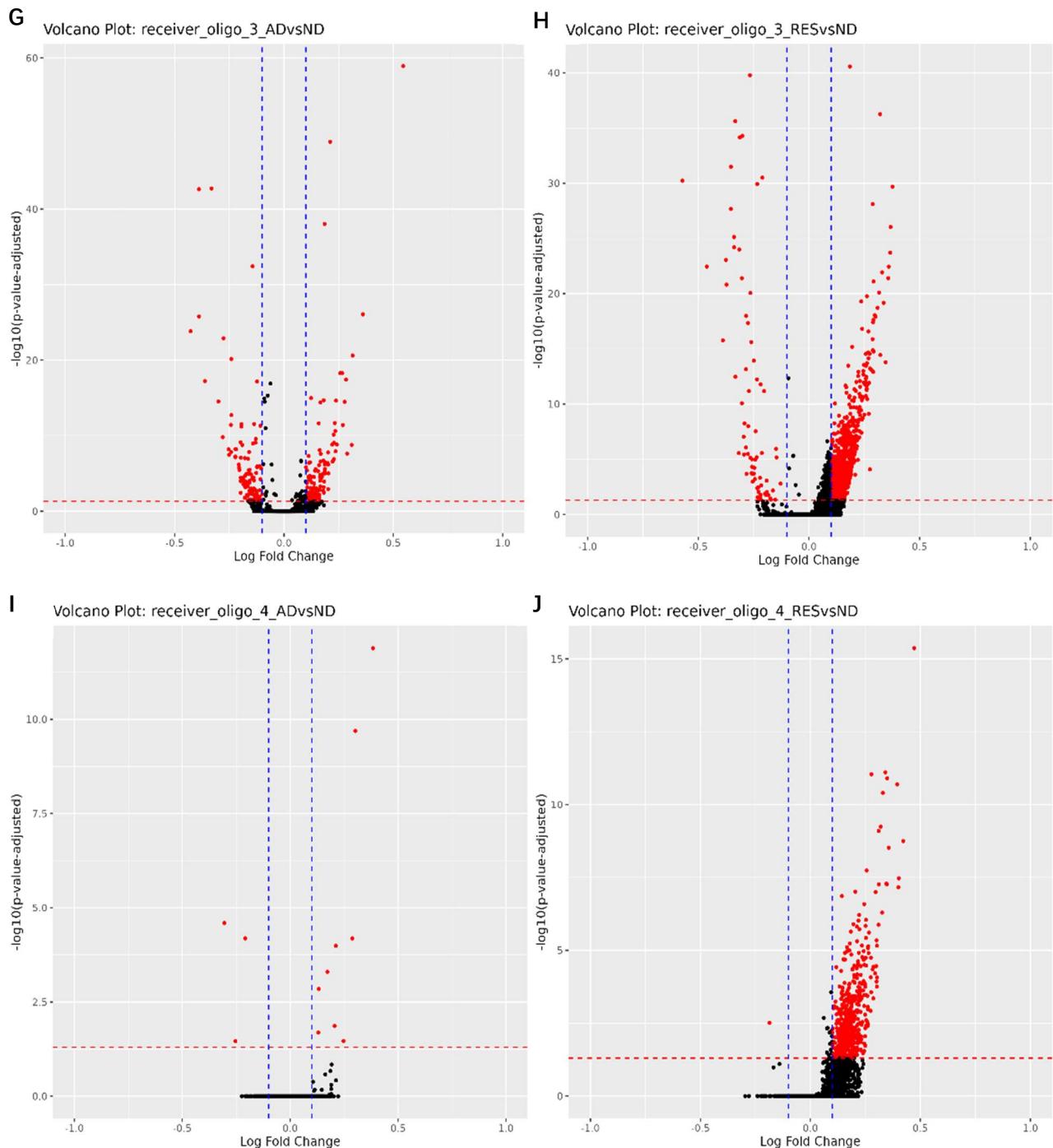


Figure 13: Volcano plots of DE genes in ADvsND and RESvsND for each cell type in PACs. Genes with $p_val_adj \leq 0.05$ & $\text{abs}(\text{avg_log2FC}) \geq 0.1$ were marked as red. (A) and (B) represent astro_0; (C) and (D) represent micro_5; (E) and (F) represent oligo_0; (G) and (H) represent oligo_3; (I) and (J) represent oligo_4.

Table 2: The number of DEGs found for receiver cell type in ADvsND and RESvsND.

Receiver	astro_0	micro_5	oligo_0	oligo_3	oligo_4
# DEGs	ADvsND	911	4	180	201
	RESvsND	1578	5	903	468

4.2.3 Visualization of ligand activity in ADvsND and RESvsND

After the quantity control of background genes and DEGs, the function *predict_ligand_activities()* was performed for PACs. One cell type in PACs was always chosen as the receiver while the other cell types were chosen as senders when *predict_ligand_activities()* was running. There were two sets of DEGs in the receiver because of three phenotypes of data (AD, RES and ND), so scatter plots were drawn to compare the ligand activities of ADvsND and ligand activities of RESvsND for each receiver (Figure 14).

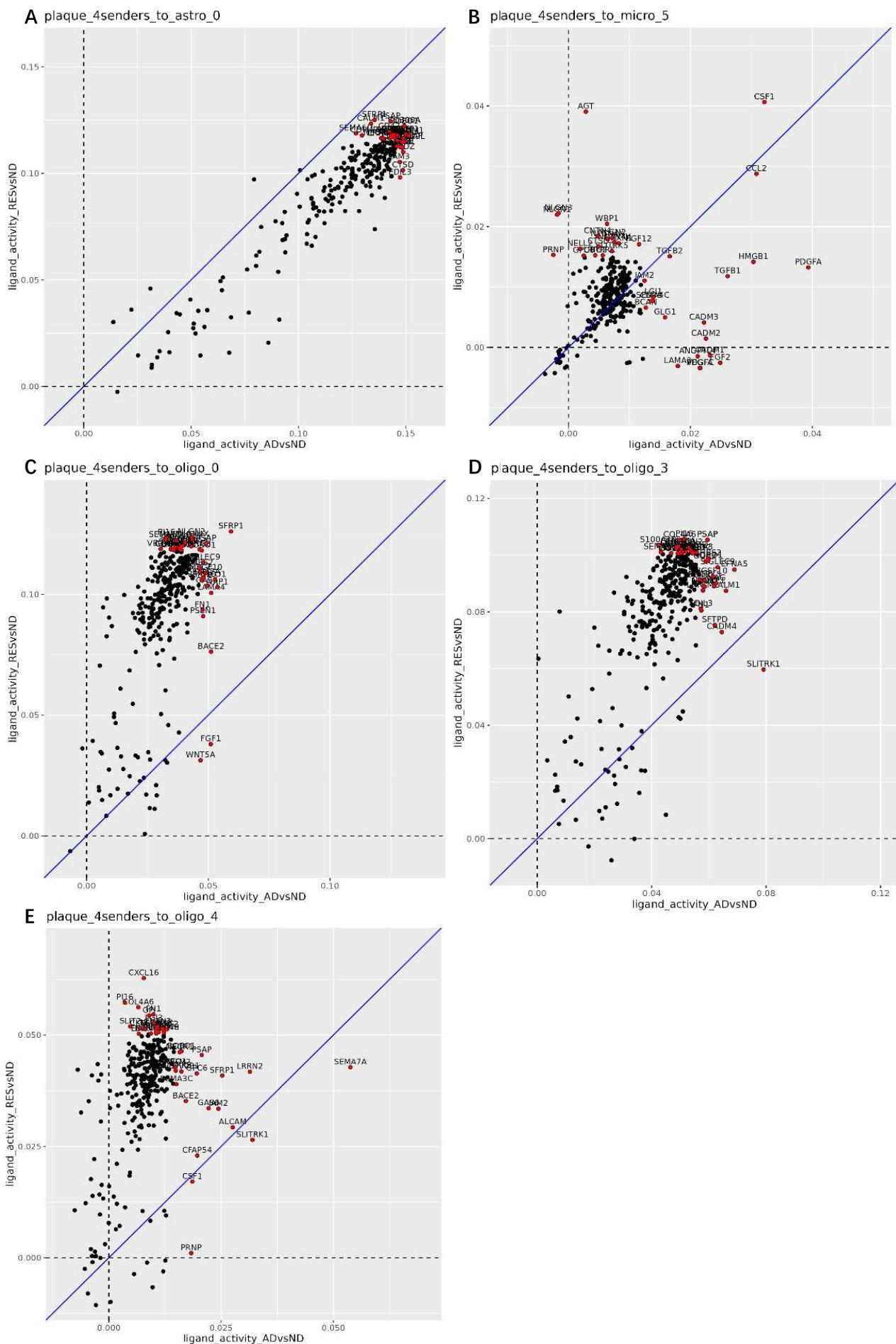


Figure 14: Comparison of liagnd_activity_ADvsND and ligand_activity_RESvsND when the receiver was one cell type in PACs and the rest were senders. The blue line (slope = 1, intercept = 0) was used to discover the changes in two activity sets. The top 20 ligands in liagnd_activity_ADvsND and ligand_activity_RESvsND were marked as red and labeled. (A) receiver was astro_0; (B) receiver was micro_5; (C) receiver was oligo_4; (D) receiver was oligo_3; (E) receiver was oligo_0.

4.2.4 Top 20 ligands ranked by activity scores

Ligand prediction analysis was performed for each receiver cell type in ADvsND and RESvsND, 307 ligands were found for receiver oligo_0; 250 ligands were found for receiver astro_0; 288 ligands were found for receiver micro_5; 308 ligands were found for receiver oligo_3; 314 ligands were found for receiver oligo_4 (Table 3). The top 20 ligands that had positive and relatively high activities were selected for the further visualization (Table 3). Throughout this thesis, these selected ligands based on ligand activity scores were referred to as active ligands. For each receiver cell type, two ligand prediction analyses (ADvsND and RESvsND) were performed, and active ligands from two analyses were listed in the Table 3.

Table 3: The number of ligands, background genes and DEGs for each receiver cell type and the active ligands (top 20 ligands in ADvsND or RESvsND).

Receiver	micro_5	astro_0	oligo_0	oligo_3	oligo_4
# Ligands	288	250	307	308	314
# Background genes	12840	14091	12208	12681	12711
# DEGs	ADvsND: 4 RESvsND: 5	ADvsND: 911 RESvsND: 1578	ADvsND: 180 RESvsND: 903	ADvsND: 201 RESvsND: 902	ADvsND: 12 RESvsND: 468
Active ligands	PDGFA, CSF1, CCL2, HMGB1, TGFB1, FGF2, CADM1, CADM2, CADM3, PDGFC, VEGFA, ANGPTL4, LAMA2, TGFB2, GLG1, LGI1, SEMA4C, CD38, BCAN, JAM2, FGF1, AGT, NLGN3, NLGN1, WBP1, CNTN4, NLGN2, NLGN4X, NLGN4Y, NRXN1, FGF12, CTSD, NELL2, SLITRK5, PRNP, LRIG2, PTPRK, GPC6	NCAM1, GSTP1, ALOX5AP, CD300A, ROBO1, MPDZ, SEMA6A, CTSD, KIAA0319L, EFNB2, ARF1, A2M, CLU, EFNA1, EDIL3, JAM3, CD58, CLDN11, EMC1, SIGLEC9, SFRP1, PSAP, CALM1, CD47, SEMA6D, PTPRK, SLITRK2, PLEKHO2, CDH4, LRFN5, TGOLN2, SIRPB2, NECTIN1, LRRC4B	SFRP1, NRP1, NEO1, LAMA4, BACE2, FGF1, GAS6, PSAP, MEGF10, SEMA7A, SIGLEC9, MAG, PSEN1, FN1, ROBO1, VSIG10, WNT5A, THBS2, SDC2, NLGN2, PI16, NLGN4X, SEMA6D, LRRC4B, NLGN4Y, COL21A1, SEMA6A, THBS4, COL5A3, COL4A5, GSTP1, COL4A6, LRFN5, TCTN1	SLITRK1, EFNA5, CALM1, CADM4, SIGLEC9, MAG, SFTP1, NRP1, IGSF10, THBS2, ROBO1, PSAP, SDC2, LRRN2, SLITRK5, NEO1, LRRTM2, FN1, EDIL3, NRG2, PI16, COL4A6, S100A10, SEMA6A, NLGN2, LRRC4B, CLDN11, NLGN4Y, SEMA3D, THBS4, SEMA6D, COL5A3, FGF11, PLEKHO2, CNTN3, LRFN5, NLGN4X, FGF14, COL21A1	SEMA7A, SLITRK1, LRRN2, ALCAM, SFRP1, JAM2, GAS6, PSAP, CFAP54, GPC6, CSF1, PRNP, BACE2, ROBO1, HSP90B1, NEGR1, SEMA3C, SLITRK5, NEO1, LRRTM2, CXCL16, PI16, COL4A6, FN1, GPI, LGI3, SLIT2, CNTN3, NLGN4X, NID1, CKLF, TG, FLRT2, BOC, TSPAN5, TGOLN2, LRRC4B, NLGN2, NLGN4Y, EDA

Venn plot was made to observe the overlaps among these active ligands for different receiver cell types (Figure 15). There were a lot of overlaps between different active ligand sets, especially among oligo_0, oligo_3 and oligo_4, which was one of the motivations that we combined oligo_0, oligo_3 and oligo_4 into one cluster ‘oligos’.

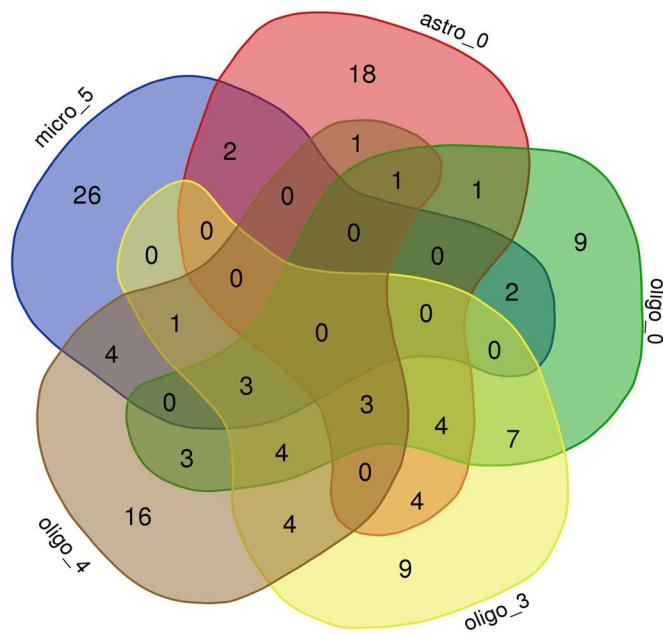


Figure 15: The Venn plot of active ligands found in each receiver. Active ligands for each receiver cell type were listed in Table 3. Each color represents active ligands for one receiver: blue = micro_5, red = astro_0, green = oligo_0, yellow = oligo_3, and gray = oligo_4.

4.2.2 Additional information of active ligands

A lot of information about ligands and receptors is lost in NicheNet's ligand prediction study because it focuses solely on which ligands should be responsible for the response of DGEs. Therefore, it is better to illustrate the expression and log fold change of ligands and receptors in possible cell signaling pathways after identifying active ligands in NicheNet. It is clear to notice that although some ligands had high ligand activities, their expressions and receptors' expressions were low in senders and receivers (Figure 16 & Figure 17). The additional information of active ligands of other receivers were displayed in Supplement Figures 9-16.

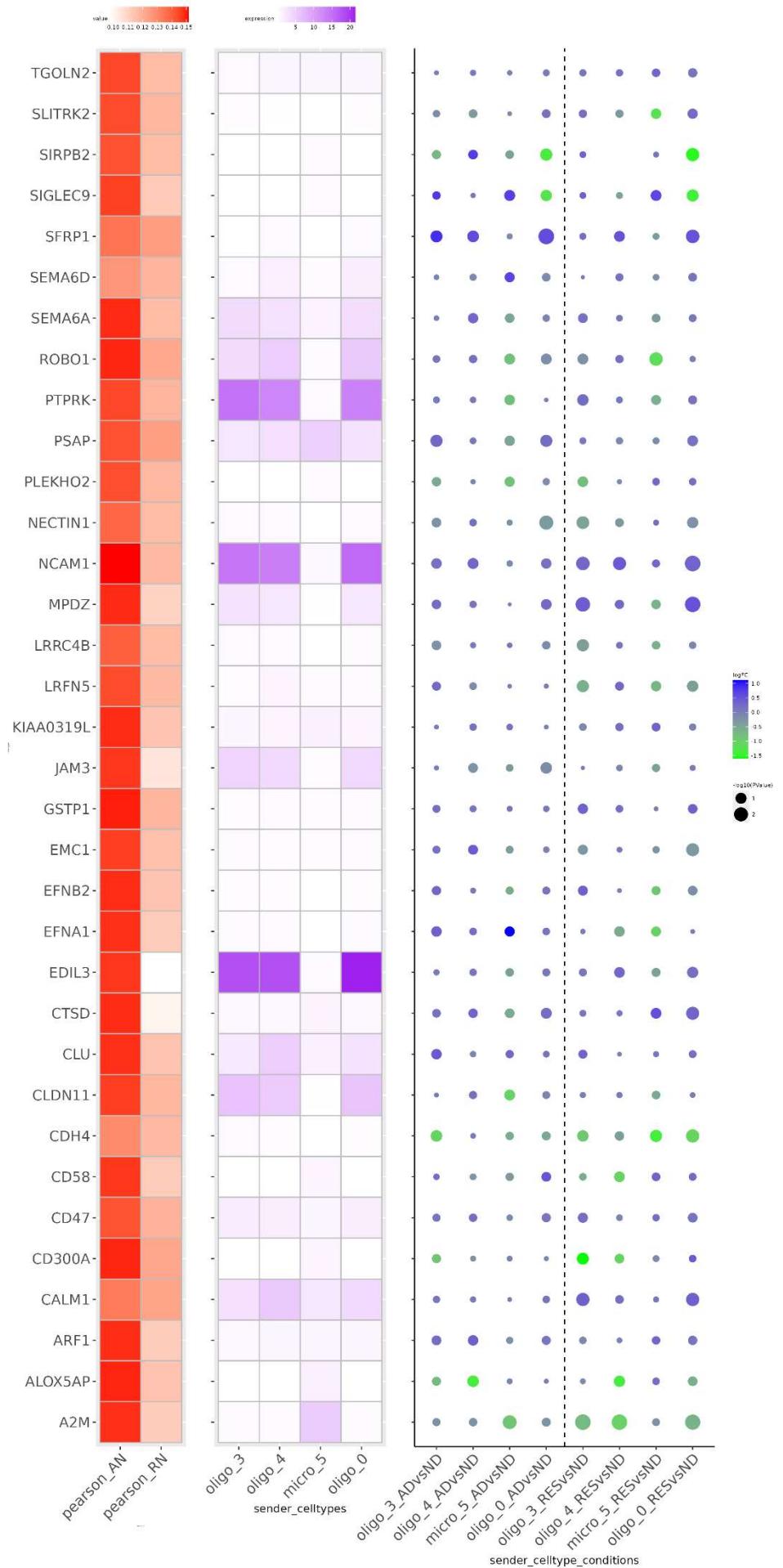


Figure 16: From left to right, the plots represent the ligand activity, expression and log fold change(LFC) of each active ligand for receiver astro_0. In the dot plot to visualize the log fold change of ligands, the color of the dots indicates the LFC value, and the size of the dots represents the $-\log_{10}(p\text{-value})$ of the DE ligands.

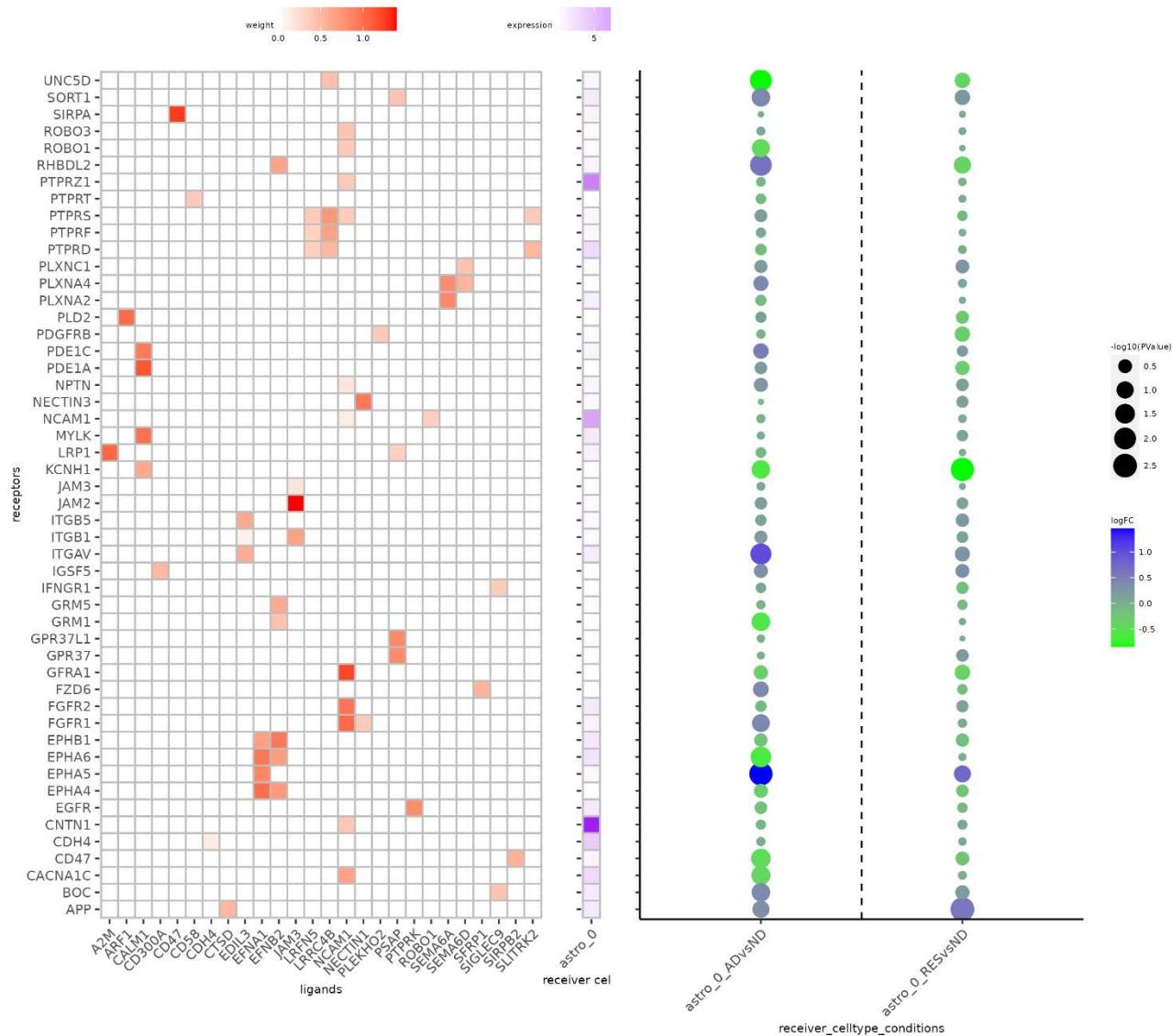


Figure 17: From left to right, the plots represent the ligand-receptor weights, expression and log fold change(LFC) of each receptor in receiver astro_0. The ligand-receptor weights matrix was generated from the database of NicheNet. In the dot plot to visualize the log fold change of receptors, the color of the dots indicates the LFC value, and the size of the dots represents the $-\log_{10}(p\text{-value})$ of the DE receptors.

4.3 Intersection between findings of CellphoneDB and NicheNet

4.3.1 Data sources of CellphoneDB and NicheNet

The size and data sources of the databases employed in CellphoneDB and NicheNet are quite different, which has a significant impact on the outcome. Simple proteins and multi-subunit complexes are included in CellphoneDB's interaction input, which has been collected from a range of data sources including PubMed, HMRbase, Uniprot, and Reactome. NicheNet has a ligand-receptor network, which collects ligands and receptors from data sources like KEGG, Ramilowski et al. (2015) and IUPHAR Guide to Pharmacology (via Harmonizome). Simple proteins were selected from the interaction input of CellphoneDB, which were used to compare with the ligands and receptors selected from the ligand-receptor network of NicheNet. There are 1226 ligands and 1067 receptors in the database of NicheNet, and 598 ligands and 512 receptors (both are simple proteins) in the database of CellphoneDB. In ligands, there was a 312-protein overlap, and in receptors, there was a 204-protein overlap (Figure 18).

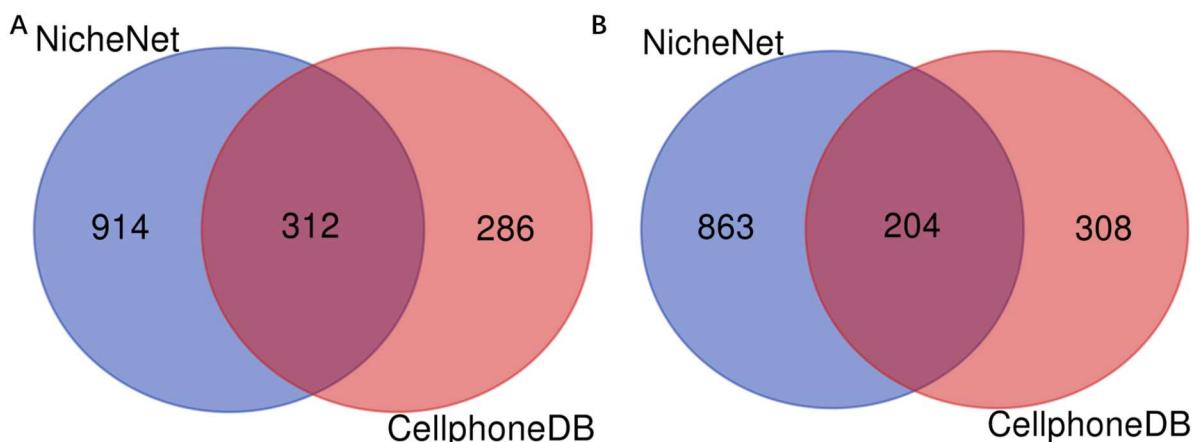


Figure 18: Venn plots of ligands and receptors in the databases of CellphoneDB and NicheNet (ligands and receptors are simple proteins in the CellphoneDB database). The blue represents proteins in NicheNet and the red represents proteins in CellphoneDB. (A)Ligands, (B)Receptors.

4.3.2 Potential cell-cell communications between PACs

Following the CellphoneDB analysis, 548 celltype-specific interactions were discovered, 93 of which involved ligands with simple structures. In the NicheNet analysis, A total of 122 active ligands were identified. There was an overlap between findings from CellphoneDB and NicheNet, and these

ligands were CSF1, TGFB1, FGF2, CADM1, CADM3, VEGFA, TGFB2, SEMA4C, JAM2, FGF1, NRXN1, EFNB2, CLU, EFNA1, JAM3, PSAP, GAS6, SEMA7A, FN1, WNT5A, PI16, VEGFB, COL4A5, EFNA5 and SEMA3C (Figure 19). A more informative filter was made to remove ligands that were not sent to the same receiver cell type in CellphoneDB, and interactions that were not collected in the ligand-receptor network in NicheNet (Table 4).

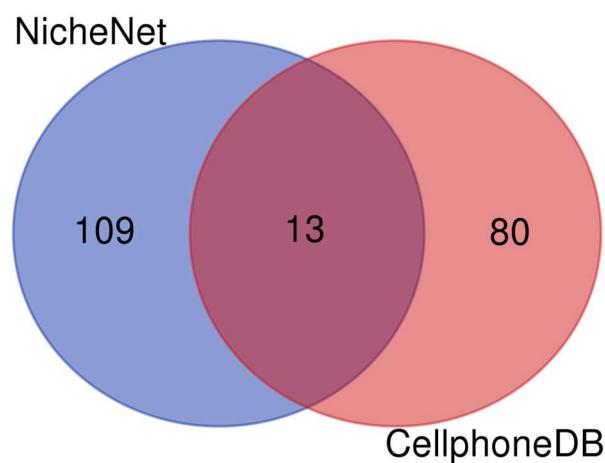


Figure 19: Venn plot of the number of ligands found in NicheNet and CellphoneDB. The blue represents NicheNet, and the red represents CellphoneDB.

Table 4: The final ligand-receptor pairs found in both NicheNet analysis and CellphoneDB analysis.

ligand	receptor	sender	receiver
CSF1	CSF1R	oligo_3	micro_5
		astro_0	micro_5
TGFB1	TGFBR1	astro_0	micro_5
TGFB1	TGFBR2	astro_0	micro_5
VEGFA	NRP1	astro_0	micro_5
VEGFA	NRP2	astro_0	micro_5
TGFB2	TGFBR1	astro_0	micro_5
TGFB2	TGFBR2	astro_0	micro_5
SEMA4C	PLXNB2	oligo_3	micro_5
		oligo_4	micro_5
		oligo_0	micro_5
JAM2	JAM3	astro_0	oligo_4
		micro_5	oligo_4
FGF2	ITGAV	astro_0	micro_5
FGF1	ITGAV	oligo_3	micro_5
		oligo_4	micro_5
FGF1	FGFR2	oligo_0	micro_5
		astro_0	oligo_0
		oligo_3	oligo_0
		oligo_4	oligo_0
NRXN1	NLGN1	astro_0	oligo_0
		oligo_3	astro_0
		oligo_4	astro_0
EFNB2	EPHB1	oligo_0	astro_0
		oligo_3	astro_0

		oligo_4	astro_0
EFNA1	EPHA5	oligo_0	astro_0
		oligo_3	astro_0
		oligo_4	astro_0
		oligo_0	astro_0
		micro_5	astro_0
		micro_5	oligo_3
		oligo_0	oligo_3
PSAP	GPR37L1	oligo_3	oligo_3
		micro_5	oligo_3
		oligo_4	oligo_3
		oligo_3	oligo_3
		astro_0	oligo_4
		oligo_0	oligo_4
		micro_5	oligo_4
		oligo_3	oligo_4
		astro_0	oligo_0
		oligo_4	oligo_0
PSAP	GPR37	micro_5	oligo_0
		oligo_3	oligo_0
		oligo_3	oligo_0
PSAP	GPR37L1	micro_5	oligo_0
		oligo_3	oligo_0
		astro_0	oligo_0
GAS6	TYRO3	micro_5	oligo_4
FN1	ITGA2	oligo_4	oligo_3
		micro_5	oligo_3

		astro_0	oligo_3
		oligo_4	oligo_3
		oligo_0	oligo_3
		astro_0	oligo_4
		oligo_4	oligo_0
		astro_0	oligo_0
FN1	ITGAV	oligo_3	oligo_4
		oligo_3	oligo_0
WNT5A	ROR1	micro_5	oligo_0
PI16	TNFRSF21	oligo_4	oligo_0
EFNA5	EPHB2	astro_0	oligo_3

Since oligo_0, oligo_3, and oligo_4 shared a lot of similarities in interactions, they were grouped into oligos, and a new table was created to display the cell-to-cell connections between astro_0, micro_5, and oligos (Table 5). Additionally, the articles that support the function of interactions in AD were included, along with the phenotype-specificities including AD-specificity, RES-specificity, and plaque-specificity for each interaction. It is clear that not every phenotype-specificity was consistently represented in the output of CellphoneDB and NicheNet.

Table 5: Table of potential CCCs between PACs and the papers supporting the influence of interactions in Alzheimer's Disease. Plaque, AD and RES indicate whether the interaction is plaque-specific, AD-specific or RES-specific in NicheNet and CellphoneDB, which was manually annotated by comparing ligand activity from NicheNet in ADvsND and RESvsND and celltype-specificity from CellphoneDB in AD, RES and ND. In NicheNet, AD-specificity indicates the ligand was only active in ADvsND, RES-specificity indicates the ligand was only active in RESvsND, and plaque-specificity indicates the ligand was active in both ADvsND and RESvsND. In CellphoneDB, AD-specificity indicates the interaction was only celltype-specific in AD, RES-specificity indicates the interaction was only celltype-specific in RES, and plaque-specificity indicates the interaction was celltype-specific in both AD and RES.

sender	receiver	ligand	receptor	paper related to AD	NicheNet	CellphoneDB
astro_0	micro_5	CSF1	CSF1R	Cheng et al., 2021	plaque	plaque
		TGFB1	TGFBR1	Lively et al., 2018	plaque	plaque
		TGFB1	TGFBR2	Moursel et al., 2018	plaque	plaque
		VEGFA	NRP1	Seto ET AL., 2023	AD	plaque
		VEGFA	NRP2	Seto ET AL., 2023	AD	plaque
		TGFB2	TGFBR1	Roychowdhury et al., 2021	plaque	plaque
		TGFB2	TGFBR2	Moursel et al., 2018	plaque	plaque
		FGF2	ITGAV	Hutter-Schmid et al., 2018	AD	AD
	oligos	JAM2	JAM3	Cen et al., 2020	plaque	plaque
		FGF1	FGFR2	Gama et al., 2005	AD	plaque
		NRXN1	NLGN1	Loers et al., 2023	RES	RES
		PSAP	GPR37L1	Kaufmann et al., 2022	RES	plaque
		FN1	ITGA2	Gwon et al., 2018	RES	plaque
		NRXN1	LRRTM3	Swaminathan et al., 2011	RES	plaque
		EFNA5	EPHB2	Liu et al., 2017	RES	plaque
		GAS6	TYRO3	Tondo et al., 2019	RES	plaque
		astro_0	EFNA1	EPHA5	AD	AD
micro_5	oligos	JAM2	JAM3	Cen et al., 2020	plaque	RES
		EFNA1	EPHA5	Paik et al., 2020	AD	AD
		PSAP	GPR37L1	Kaufmann et al., 2022	RES	plaque
		PSAP	GPR37	Meyer et al., 2014	RES	plaque
		GAS6	TYRO3	Tondo et al., 2019	RES	plaque
		WNT5A	ROR1	Becker et al., 2018	plaque	plaque
		oligos	astro_0	PSAP	AD	plaque
		EFNB2	EPHB1	Cissé et al., 2015	AD	plaque

	EFNA1	EPHA5	Paik et al., 2020	AD	AD
micro_5	CSF1	CSF1R	Cheng et al., 2021	plaque	plaque
	SEMA4C	PLXNB2	Carulli ., 2021	plaque	plaque
	FGF1	ITGAV	Bountali et al., 2019	RES	AD
	FGF1	FGFR2	Gama et al., 2005	RES	AD
oligos	FGF1	FGFR2	Gama et al., 2005	AD	plaque
	EFNA1	EPHA5	Paik et al., 2020	AD	AD
	PSAP	GPR37L1	Kaufmann et al., 2022	plaque	plaque
	PSAP	GPR37	Meyer et al., 2014	plaque	plaque
	FN1	ITGA2	Gwon et al., 2018	RES	RES
	FN1	ITGAV	Hall-Roberts et al., 2020	RES	AD
	PI16	TNFRSF21	Zhang et al., 2021	RES	RES

5 Discussion

5.1 CellphoneDB

5.1.1 Mean expression and p-value for each interaction

In order to gain a deeper understanding of the unique expression of ligand-receptor pairs in samples of AD, RES, and ND, we conducted *cpdb_statistical_analysis()* separately for each of these phenotypes. Figure 5 shows the overall number of interactions across three phenotypes as well as the number of interactions involving protein complexes and simple proteins. Three CellphoneDB outputs had comparable interaction number structures, Three CellphoneDB outputs exhibited similar levels of LR interactions and composition, making it feasible to compare phenotypes downstream. Moreover, the CellphoneDB database that is a curated database contains a total of 2923 interactions, and nearly all of these interactions were expressed in the cells analyzed in this thesis. This indicates that the database is highly suitable for studying ligand-receptor interactions among human brain samples.

Two files generated from *cpdb_statistical_analysis_method()* were essential in finding celltype-specific interactions, containing the mean expressions matrix and the p-values matrix of each interaction. CellphoneDB did not provide the FDR-controlling procedure to control false positives when conducting multiple comparisons, so the FDR control based on the Benjamini-Hochberg Procedure was done to obtain the adjusted p-values of each interaction, which was utilized to replace the p-value in this thesis. The scatter plot comparing mean expression of interactions between AD and RES in Figure 6 highlights a few interactions that were exclusively present in AD or RES which are the ones that contribute to the minor variation in the number of interactions across different phenotypes. The mean expression level is stable for those interactions found in both AD and RES. However, the mean expression and cell-specificity did not show a strong correlation in Figure 6. Interactions could have celltype-specificity with either high mean expression or low mean expression.

5.1.2 Celltype-specific interactions

There was a filter made to exclude interactions that lacked cell-specificity in AD or RES, resulting in 80% of interactions being dropped successfully. Since the database of NicheNet only contains simple proteins, 548 celltype-specific interactions, which only contained simple proteins as ligands or receptors, were selected to compare the results from NicheNet in the end of thesis (Figure 7). Most of these interactions showed cell-specificity in both AD and RES, while few of these interactions were only cell-specific in AD (Figure 8). Considering that cell-type specificity primarily relies on the variance in mean expression between the cell type of interest and the remaining PACs, these interactions suggest the activation of potentially more LR pairs in the resilient sample through protective mechanisms, or the presence of greater changes in the cellular composition within the PACs of the resilient sample. CCCs from astro_0 to micro_5 and CCCs from astro_0 to oligo_4 were much stronger (more celltype-specific interactions) than CCCs between other PACs in AD, while CCCs from oligo_0 to astro_0 were the strongest in RES (Figure 9A and 9B). astro_0 was used as an example in Figure 10 to visualize the celltype-specific interactions since it had the largest number of celltype-specific interactions and the most receptors (Table 1 & Figure 11). Although these findings were generated through in silico approaches, it is crucial to validate the relevance of these ligand-receptor (LR) pairs in the context of AD progression. NicheNet serves as an alternative that could be utilized to assess and compare the LR pairs identified by CellphoneDB. However, experimental validation techniques like *in situ* hybridization are necessary to confirm the findings and provide concrete evidence.

At the end of this thesis, all of these celltype-specific interactions were used to combine the NicheNet results, but if CellphoneDB was the only method available for identifying CCCs between cells, it was also essential to display the mean expressions and p-values for each cell pair. In Figure 10, interactions with higher mean expression (larger size) and celltype-specificity only in AD or RES (yellow color represent AD and purple color represent RES) have more research value than other interactions. In the final stages of this thesis, a combination of these cell-type-specific interactions with the active ligands predicted by NicheNet would be used to generate a consensus list. However, if CellphoneDB were the sole method used to identify CCCs between cells, post hoc filtering based on mean expression and p-value becomes crucial (Figure 10 & Supplement Figure 5-8). In Figure 10,

interactions that are both cell-type-specific and exhibit high expression (indicated by larger size) in either AD (represented by yellow color) or RES (represented by purple color) are given priority for further investigation.

5.2 NicheNet

5.2.1 The quantity control of background genes

Prior to using the NicheNet ligand prediction methodology, the quantity control of background genes was crucial because it directly affected the Pearson correlation coefficient (ligand activity). Typically, NicheNet manual recommends using genes expressed in a minimum of 10% of cells as background genes. However, the 10% threshold will reduce the number of background genes to 5000-8000 in our dataset. Upon consulting the original author, it was advised to have at least 10,000 background genes to ensure the performance of ligand prediction analysis. Following several iterations using different thresholds (10%, 5%, 1%), a threshold of 1% was chosen to ensure sufficient representation of the background genes (Figure 12).

5.2.2 DE analysis

Despite the fact that NicheNet does not offer comprehensive details regarding the DE analysis, the DE analysis is crucial for the ligand prediction analysis because the list of DEGs can affect all activities of ligands. In this thesis, DE analysis based on Wilcoxon Test determined the genes were differentially expressed in ADvsND or RESvsND, which was then converted to transcriptional responses that identified the active ligands. In DE analysis, $p_val_adj \leq 0.05$ & $abs(avg_log2FC) \geq 0.1$ were used to mark DEGs, and the numbers of DEGs varied considerably among various cell types (Figure 13). There were 911 DEGs and 1578 DEGs found for astro_0 in ADvsND and RESvsND, whereas only 4 DEGs and 5 DEGs were found for micro_5 in ADvsND and RESvsND. Astro_0 had a lot of DEGs and micro_5 had only few DEGs, so astro_0 had a high probability of being the receiver and micro_5 probably worked as the sender in CCCs. Neither too few or too many DEGs works well in ligand prediction, and the package author suggested a range of 20–1000 for the number of DEGs. The influence of DEGs on the performance of ligand prediction

would be demonstrated in the next step. The evaluation of NicheNet was performed during the step of ligand prediction analysis, using astro_0 and micro_5. Additionally, it was observed that there are more DGEs in RESvsND than in ADvsND. This finding suggests the presence of a potential protective mechanism in RES samples, potentially mediated through CCCs.

5.2.3 Ligand prediction analysis

The function `predict_ligand_activities()` was performed ten times for five receiver cell types. For each receiver cell type, ligand activities in ADvsND and RESvsND were obtained. Visualization of ligand activities was done to evaluate the performance of NicheNet under various numbers of DEGs, and the ligands with high activities were marked. In Figure 14, it was clear that more active ligands with high activity were found for a given number of background genes. For example, oligo_0 had 180 DEGs and 903 DEGs in ADvsND and in RESvs ND, so the ligand activities of the majority of ligands in RESvsND were higher than that in ADvsND (Figure 14C). However, the correlation between the trend of ligand activities and the number of DEGs would be curtailed when the number of DEGs exceeded 1000, for example, astro_0 had 911 DEGs and 1578 DEGs found for astro_0 in ADvsND and RESvsND, but the ligand activities of the majority of ligands did not significantly change in two cases (Figure 14A). In addition, it was helpful to distinguish a few ligands if the number of DEGs was low, for example, micro_5 only had 4 DEGs and 5 DEGs were found for micro_5 in ADvsND and RESvsND, so the distribution of ligands in the scatter plots appeared significantly different from others. The ligand activities of most ligands for micro_5 were close to zero, and the ligand activities of those distinguished ligands were only around 0.03, making it difficult to tell whether or not they were active (Figure 14B).

Since NichNet is very sensitive to the DEGs input list, by testing a range of DE cutoffs ($p_val_adj \leq 0.05 \ \& \ abs(avg_log2FC) \geq 0.2, 0.1, 0.05$), we have concluded that using a DEGs list with no less than 20 and no more than 1000 will give the optimal prediction. In particular, Figure 14 was shown that when the number of input genes is less than 20, it may lead to false predictions of active ligands. Conversely, if there are more than 1000 input genes, it becomes difficult to distinguish the true active ligands from the large pool of backgrounds.

5.2.4 Active ligands

There were hundreds of ligands for each type of receiver cell, but not all of them were recognized by NicheNet as being active. The top 20 ligands in ligand activities were referred to as active ligands in each ligand prediction analysis (Table 2). Ligands that were only classified as the top 20 in ADvsND or RESvsND probably contribute to the phenotype difference between AD and RES, while ligands that were classified as the top 20 in both ADvsND and RESvsND probably contribute to the plaque development in Alzheimer's disease. Some additional information was displayed for these active ligands, it is clear that some active ligands and their receptors have low expression and small LFC. It is one shortcoming of NicheNet because NicheNet does not take gene expression into account. If NicheNet were the sole method used to identify CCCs between cells, these figures could be used to do a post filtering on the ligands.

5.3 Potential CCCs in PACs

There were 548 celltype-specific interactions and 122 active ligands discovered by a study of CellphoneDB and NicheNet. These two methods utilized entirely different algorithms to identify CCCs among cells. One aim of this thesis is trying to determine whether any ligand-receptor interactions are predicted by both methods. 13 ligands were first discovered by simply looking the overlap between active ligands from NicheNet analysis and ligands (simple protein structure) involved in celltype-specific interactions from CellphoneDB analysis (Figure 19), but NicheNet's algorithm requires a double-check to see if they were produced by the same sender cell type and their receptors were produced by the same receiver cell type, as shown by CellphoneDB results. Therefore, Table 3 was made, which contains the final potential ligand-receptor interactions, the sender cell type and the receiver cell type. There were 18 ligand-receptor interactions in Table 3, some of which could happen in several sender-receiver combinations. Since oligo_0, oligo_3 and oligo_4 share a lot of similarities during the study like most interactions and ligands found were the same, these three clusters were merged into one cluster oligos. By looking through relevant studies, Table 4 was created to confirm the effect of these ligand-receptor interactions on Alzheimer's disease. All interactions found may contribute to plaque formation in Alzheimer's disease; but some of them can also be responsible for the phenotypic difference between AD and RES, based on the cell-type specificity

found in CellphoneDB and ligand activity found in NicheNet. For example, an interaction that was only celltype-specific in RES and the ligand involved in this interaction was only active in RESvsND could be responsible for the phenotype RES. Not all interactions show consistent phenotype-specificity in both methods. A limited number of interactions had the same phenotype-specificity, such as FGF2-ITGAV, NRXN1-NLGN1, EFNA1-EPHA5, FN1-ITGA2 and PI16-TNFRSF21 (Table 4). These five interactions in specific sender-receiver pairs are candidates that may contribute to the phenotypic difference between AD and RES.

5.4 Evaluation of CellphoneDB

CellphoneDB has a database including multi-unit complexes that function as ligands or receptors, which makes it possible to find some interesting ligand-receptor interactions involving multiple genes. The key of CellphoneDB to determine potential CCCs between cells is the expression of ligands and receptors. In theory, for a ligand-receptor interaction to be deemed valid, it should be expressed initially in the sender-receiver cells. Furthermore, cell-type specificity plays a significant role in identifying potential CCCs in CellphoneDB. If an interaction is found to be specific to a particular sender-receiver combination, meaning that the ligand and receptor are expressed more in this combination compared to other possible cell combinations, there is a strong likelihood that this interaction indeed takes place in that specific sender-receiver combination. Although CellphoneDB's output is simple to understand and visualize, it misses the transcriptional change that occurs in the signaling pathway downstream following the interaction binding. This interaction is not active and this cell-to-cell communication is not happening if the receiver failed to notice any changes as a result of the interaction or if the ligand and receptor did not even bind to one another. This is one weakness of CellphoneDB, which NicheNet addresses by taking into account the DE of the target genes. Besides, another big drawback is that the database of Cellphone is not big enough to detect more possible interactions in the CCCs. If we only focus on the simple proteins in the database, there are only 598 ligands and 512 receptors in the database (Figure 18), which is much fewer than proteins collected in the NicheNet database and leads to a lot of missing interactions.

5.5 Evaluation of NicheNet

The ligand-target weights matrix is used by NicheNet's prior model for the ligand prediction analysis to identify the ligands that are active in the samples. NicheNet attempts to identify the ligand most likely to be responsible for the transcriptional change in the receiver by focusing only on the DEGs in the receiver. Every ligand is given weights to all genes in the prior model (weights matrix), and the weights represent the strength of the association between the ligand and the target gene. NicheNet might decide that a ligand is active and effective in CCCs if the target genes with high weights for that ligand are differentially expressed in the receiver. Unlike CellphoneDB, NicheNet determines CCCs solely based on the differential expression of genes within the downstream signaling pathway. However, the NicheNet approach may encounter several challenges. For instance, when targets exhibit high differential expression but have multiple upstream regulators, their corresponding ligands may receive higher activity scores, even if the actual expression of those ligands/receptors is low. Furthermore, the goal of ligand prediction analysis is to find the Pearson correlation coefficient between ligand and the transcriptional response, so it is essential to determine genes that are differentially expressed. The quantity control of background genes and DEGs has to be done before the prediction, but some biological meaning may be lost in the process. For instance, in order to have more than 10000 background genes, genes expressed in at least 1% of cells were chosen in this thesis, yet it is more logical to set the threshold at 10%. Therefore, the requirements of NicheNet are more stringent than CellphoneDB.

NicheNet and CellphoneDB both have room for improvement. For CellphoneDB, the result would be more accurate if the expression or log fold change of target genes could be taken into account. For NicheNet, it is preferable to include the quantity control of background genes and DEGs in the algorithm, and the ligand activity would be more logical if the expression of the ligand and its receptor could be used in the calculation. In addition, the quality of databases in these two methods are essential to detect potential interactions because only LR interactions recorded in the database can be detected. The databases should be large enough to detect LR interactions of interest, for example, in this thesis brain-specific LR interactions will be essential for the study.

6. Conclusion

In this thesis, 18 potential ligand-receptor interactions were identified using CellphoneDB and NicheNet. All of them might contribute to plaque formation in Alzheimer's disease based on the algorithms in these two methods, and some of them, such as FGF2-ITGAV, NRXN1-NLGN1, EFNA1-EPHA5, FN1-ITGA2 and PI16-TNFRSF21, are more likely to be the key to the phenotypic change between AD and RES in Alzheimer's disease because they are only celltype-specific in AD and active in ADvsND or only celltype-specific in RES and active in RESvsND. Some papers indicate the function of these 18 interactions on Alzheimer's disease, but more specific experiments are required to confirm these predicted relationships. During the process of identifying the cell-cell communications between plaque-activated cells, the applicability and performance of CellphoneDB and NicheNet were evaluated. CellphoneDB directly utilizes the gene expression to calculate the mean expression and cell-specificity of each interaction, and selects interactions with high mean expression and cell-specificity as candidate cell-cell communication pathways. CellphoneDB is easy to understand and user-friendly. However, CellphoneDB neglects the information from the target genes downstream of the signaling pathway after the interaction binding, which might lead to false positive errors. In contrast to CellphoneDB, NicheNet seeks to study cell-cell communication by transcriptional responses in the receiver. NicheNet does not take gene expression into account, and focuses on finding the ligands that result in DEGs. It is challenging to control the accuracy of ligand prediction analysis in NicheNet because the user has to control the quantity of background genes and DEGs. There is no one requirement for quantity control of background genes and DEGs in NicheNet, but the quantity control can directly affect the ligand activity calculation. Furthermore, it is not a good idea to completely exclude the gene expression from the analysis, which might also cause some false positive errors in the results.

5 References

- Becker, J., & Wilting, J. (2018). WNT signaling, the development of the sympathoadrenal–paraganglionic system and neuroblastoma. *Cellular and Molecular Life Sciences : CMLS*, 75(6), 1057–1070. <https://doi.org/10.1007/s00018-017-2685-8>
- Bountali, A., Tonge, D. P., & Mourtada-Maarabouni, M. (2019). RNA sequencing reveals a key role for the long non-coding RNA MIAT in regulating neuroblastoma and glioblastoma cell fate. *International Journal of Biological Macromolecules*, 130, 878–891. <https://doi.org/10.1016/j.ijbiomac.2019.03.005>
- Browaeys, R., Saelens, W., & Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. *Nature Methods*, 17(2), 159–162. <https://doi.org/10.1038/s41592-019-0667-5>
- Cang, Z., Zhao, Y., Almet, A. A., Stabell, A., Ramos, R., Plikus, M. V., Atwood, S. X., & Nie, Q. (2023). Screening cell-cell communication in spatial transcriptomics via collective optimal transport. *Nature Methods*. <https://doi.org/10.1038/s41592-022-01728-4>
- Carulli, D., de Winter, F., & Verhaagen, J. (2021). Semaphorins in Adult Nervous System Plasticity and Disease. *Frontiers in Synaptic Neuroscience*, 13, 672891–672891. <https://doi.org/10.3389/fnsyn.2021.672891>
- Cen, Z., Chen, Y., Chen, S., Wang, H., Yang, D., Zhang, H., Wu, H., Wang, L., Tang, S., Ye, J., Shen, J., Wang, H., Fu, F., Chen, X., Xie, F., Liu, P., Xu, X., Cao, J., Cai, P., ... Luo, W. (2020). Biallelic loss-of-function mutations in JAM2 cause primary familial brain calcification. *Brain (London, England : 1878)*, 143(2), 491–502. <https://doi.org/10.1093/brain/awz392>
- Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Sala Frigerio, C., Corthout, N., Qian, X., Laláková, J., Kühnemund, M., Voytyuk, I., Wolfs, L., Mancuso, R., Salta, E., Balusu, S., Snellinx, A., Munck, S., Jurek, A., Fernandez Navarro, J., Saido, T. C., ... De Strooper, B. (2020). Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's . *Cell*, 182(4), 976–991.e19. <https://doi.org/10.1016/j.cell.2020.06.038>
- Cheng, B., Li, X., Dai, K., Duan, S., Rong, Z., Chen, Y., Lü, L., Liu, Z., Huang, X., Xu, H., Zhang, Y.-W., & Zheng, H. (2021). Triggering Receptor Expressed on Myeloid Cells-2 (TREM2) Interacts With Colony-Stimulating Factor 1 Receptor (CSF1R) but Is Not Necessary for CSF1/CSF1R-Mediated Microglial Survival. *Frontiers in Immunology*, 12, 633796–633796. <https://doi.org/10.3389/fimmu.2021.633796>
- Cipriani, G., Dolciotti, C., Picchi, L., & Bonuccelli, U. (2011). Alzheimer and his : a brief history. *Neurological Sciences*, 32(2), 275–279. <https://doi.org/10.1007/s10072-010-0454-7>
- Cissé, M., & Checler, F. (2015). Eph receptors: New players in Alzheimer's disease pathogenesis. *Neurobiology of Disease*, 73, 137–149. <https://doi.org/10.1016/j.nbd.2014.08.028>
- Condello, C., Yuan, P., Schain, A., & Grutzendler, J. (2015). Microglia constitute a barrier that prevents neurotoxic protofibrillar A β 42 hotspots around plaques. *Nature Communications*, 6(1), 6176–6176. <https://doi.org/10.1038/ncomms7176>
- Cribbs, D. H., Berchtold, N. C., Perreau, V., Coleman, P. D., Rogers, J., Tenner, A. J., & Cotman, C. W. (2012). Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and

neurodegeneration: A microarray study. *Journal of Neuroinflammation*, 9(1), 179–179. <https://doi.org/10.1186/1742-2094-9-179>

De Strooper, B., & Karan, E. (2016). The Cellular Phase of Alzheimer's . *Cell*, 164(4), 603–615. <https://doi.org/10.1016/j.cell.2015.12.056>

Delekate, A., Füchtemeier, M., Schumacher, T., Ulbrich, C., Foddis, M., & Petzold, G. C. (2014). Metabotropic P2Y1 receptor signalling mediates astrocytic hyperactivity in vivo in an Alzheimer's mouse model. *Nature Communications*, 5(1), 5422–5422. <https://doi.org/10.1038/ncomms6422>

Efremova, M., Vento-Tormo, M., Teichmann, S. A., & Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nature Protocols*, 15(4), 1484–1506. <https://doi.org/10.1038/s41596-020-0292-x>

Efremova, M., Vento-Tormo, R., Teichmann, S. A., & Vento-Tormo, M. (2021). CellphoneDB (v2.0.0) [Software]. GitHub. <https://github.com/Teichlab/cellphonedb>

Ferrero-Miliani, L., Nielsen, O. H., Andersen, P. S., & Girardin, S. E. (2007). Chronic inflammation: importance of NOD2 and NALP3 in interleukin-1beta generation. *Clinical and Experimental Immunology*, 47(2), 227–235.

Freeman, M. R., & Rowitch, D. H. (2013). Evolving Concepts of Gliogenesis: A Look Way Back and Ahead to the Next 25 Years. *Neuron (Cambridge, Mass.)*, 80(3), 613–623. <https://doi.org/10.1016/j.neuron.2013.10.034>

Frost, G. R., & Li, Y.-M. (2017). The role of astrocytes in amyloid production and Alzheimer's . *Open Biology*, 7(12), 170228–. <https://doi.org/10.1098/rsob.170228>

Gama, C. I., & Hsieh-Wilson, L. C. (2005). Chemical approaches to deciphering the glycosaminoglycan code. *Current Opinion in Chemical Biology*, 9(6), 609–619. <https://doi.org/10.1016/j.cbpa.2005.10.003>

Gehrmann, J., Matsumoto, Y., & Kreutzberg, G. W. (1995). Microglia: Intrinsic immune effector cell of the brain. *BRAIN RESEARCH REVIEWS*, 20(3), 269–287. [https://doi.org/10.1016/0165-0173\(94\)00015-H](https://doi.org/10.1016/0165-0173(94)00015-H)

González-Reyes, R. E., Nava-Mesa, M. O., Vargas-Sánchez, K., Ariza-Salamanca, D., & Mora-Muñoz, L. (2017). Involvement of astrocytes in Alzheimer's from a neuroinflammatory and oxidative stress perspective. *Frontiers in Molecular Neuroscience*, 10, 427–427. <https://doi.org/10.3389/fnmol.2017.00427>

Grolla, A. A., Fakhfouri, G., Balzaretti, G., Marcello, E., Gardoni, F., Canonico, P. L., DiLuca, M., Genazzani, A. A., & Lim, D. (2013). A β leads to Ca $^{2+}$ signaling alterations and transcriptional changes in glial cells. *Neurobiology of Aging*, 34(2), 511–522. <https://doi.org/10.1016/j.neurobiolaging.2012.05.005>

Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogaeva, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J. S. , 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. *The New England Journal of Medicine*, 368(2), 117–127. <https://doi.org/10.1056/NEJMoa1211851>

Gwon, S.-Y., Rhee, K.-J., & Sung, H. J. (2018). Gene and protein expression profiles in a mouse model of collagen-induced arthritis. *International Journal of Medical Sciences*, 15(1), 77–85. <https://doi.org/10.7150/ijms.22345>

Hall-Roberts, H., Agarwal, D., Obst, J., Smith, T. B., Monzón-Sandoval, J., Di Daniel, E., Webber, C., James, W. S., Mead, E., Davis, J. B., & Cowley, S. A. (2020). TREM2 Alzheimer's variant R47H causes similar transcriptional dysregulation to knockout, yet only subtle functional phenotypes in human iPSC-derived macrophages. *Alzheimer's Research & Therapy*, 12(1), 151–151. <https://doi.org/10.1186/s13195-020-00709-z>

HARDY, J., & HIGGINS, G. (1992). Alzheimer's : The Amyloid Cascade Hypothesis. *Science (American Association for the Advancement of Science)*, 256(5054), 184–185. <https://doi.org/10.1126/science.1566067>

Hashioka, S., Wu, Z., & Klegeris, A. (2021). Glia-Driven Neuroinflammation and Systemic Inflammation in Alzheimer's . *Current Neuropharmacology*, 19(7), 908–924. <https://doi.org/10.2174/1570159X1866620111104509>

Henstridge, C. M., Hyman, B. T., & Spires-Jones, T. L. (2019). Beyond the neuron–cellular interactions early in Alzheimer pathogenesis. *Nature Reviews. Neuroscience*, 20(2), 94–108. <https://doi.org/10.1038/s41583-018-0113-1>

Hutter-Schmid, B., & Humpel, C. (2018). Primary mouse brain pericytes isolated from transgenic Alzheimer mice spontaneously differentiate into a CD11b+ microglial-like cell type in vitro. *Experimental Gerontology*, 112, 30–37. <https://doi.org/10.1016/j.exger.2018.08.003>

Jessen, N. A., Munk, A. S. F., Lundgaard, I., & Nedergaard, M. (2015). The Glymphatic System: A Beginner's Guide. *Neurochemical Research*, 40(12), 2583–2599. <https://doi.org/10.1007/s11064-015-1581-6>

Jonsson, T., Stefansson, H., Steinberg, S., Jónsdóttir, I., Jonsson, P. V., Snaedal, J., Bjornsson, S., Huttenlocher, J., Levey, A. I., Lah, J. J., Rujescu, D., Hampel, H., Giegling, I., Andreassen, O. A., Engedal, K., Ulstein, I., Djurovic, S., Ibrahim-Verbaas, C., Hofman, A., ... Stefansson, K. (2013). Variant of TREM2 Associated with the Risk of Alzheimer's. *The New England Journal of Medicine*, 368(2), 107–116. <https://doi.org/10.1056/NEJMoa1211103>

Julia, T. C. W., & Goate, A. M. (2017). Genetics of β-amyloid precursor protein in alzheimer's. *Cold Spring Harbor Perspectives in Medicine*, 7(6), a024539–. <https://doi.org/10.1101/cshperspect.a024539>

Karran, E., Mercken, M., & Strooper, B. D. (2011). The amyloid cascade hypothesis for Alzheimer's : an appraisal for the development of therapeutics. *Nature Reviews. Drug Discovery*, 10(9), 698–712. <https://doi.org/10.1038/nrd3505>

Kaufmann, M., Schaupp, A.-L., Sun, R., Coscia, F., Dendrou, C. A., Cortes, A., Kaur, G., Evans, H. G., Mollbrink, A., Navarro, J. F., Sonner, J. K., Mayer, C., DeLuca, G. C., Lundeberg, J., Matthews, P. M., Attfield, K. E., Friese, M. A., Mann, M., & Fugger, L. (2022). Identification of early neurodegenerative pathways in progressive multiple sclerosis. *Nature Neuroscience*, 25(7), 944–955. <https://doi.org/10.1038/s41593-022-01097-3>

Lambert, J. (2013). F1–01–01: Meta-analysis in more than 74,000 individuals identifies 11 new susceptibility loci for Alzheimer's . *Alzheimer's & Dementia*, 9(4S_Part_3), P123–P123. <https://doi.org/10.1016/j.jalz.2013.04.040>

Lee, M.-C., Ting, K. K., Adams, S., Brew, B. J., Chung, R., & Guillemin, G. J. (2010). Characterisation of the expression of NMDA receptors in human astrocytes. *PloS One*, 5(11), e14123–e14123. <https://doi.org/10.1371/journal.pone.0014123>

Leng, K., Li, E., Eser, R., Piergies, A., Sit, R., Tan, M., Neff, N., Li, S. H., Rodriguez, R. D., Suemoto, C. K., Leite, R. E. P., Ehrenberg, A. J., Pasqualucci, C. A., Seeley, W. W., Spina, S., Heinsen, H., Grinberg, L. T., & Kampmann, M. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nature Neuroscience*, 24(2), 276–287. <https://doi.org/10.1038/s41593-020-00764-7>

Liddelow, S. A., Guttenplan, K. A., Clarke, L. E., Bennett, F. C., Bohlen, C. J., Schirmer, L., Bennett, M. L., Münch, A. E., Chung, W.-S., Peterson, T. C., Wilton, D. K., Frouin, A., Napier, B. A., Panicker, N., Kumar, M., Buckwalter, M. S., Rowitch, D. H., Dawson, V. L., Dawson, T. M., ... Barres, B. A. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. *Nature (London)*, 541(7638), 481–487. <https://doi.org/10.1038/nature21029>

Liu, Y.-H., Tsai, J.-W., Chen, J.-L., Yang, W.-S., Chang, P.-C., Cheng, P.-L., Turner, D. L., Yanagawa, Y., Wang, T.-W., & Yu, J.-Y. (2017). Ascl1 promotes tangential migration and confines migratory routes by induction of Ephb2 in the telencephalon. *Scientific Reports*, 7(1), 42895–42895. <https://doi.org/10.1038/srep42895>

Lively, S., Lam, D., Wong, R., & Schlichter, L. C. (2018). Comparing effects of transforming growth factor β 1 on microglia from rat and mouse: Transcriptional profiles and potassium channels. *Frontiers in Cellular Neuroscience*, 12, 115–115. <https://doi.org/10.3389/fncel.2018.00115>

Loers, G., Kleene, R., Bork, U., & Schachner, M. (2023). The Interactions of the 70 kDa Fragment of Cell Adhesion Molecule L1 with Topoisomerase 1, Peroxisome Proliferator-Activated Receptor gamma and NADH Dehydrogenase (Ubiquinone) Flavoprotein 2 Are Involved in Gene Expression and Neuronal L1-Dependent Functions. *International Journal of Molecular Sciences*, 24(3), 2097–. <https://doi.org/10.3390/ijms24032097>

Matthews, K. A., Xu, W., Gaglioti, A. H., Holt, J. B., Croft, J. B., Mack, D., & McGuire, L. C. (2018). Racial and ethnic estimates of Alzheimer's and related dementias in the United States (2015–2060) in adults aged \geq 65 years. *Alzheimer's & Dementia*. <https://doi.org/10.1016/j.jalz.2018.06.3063>

Mazaheri, F., Snaidero, N., Kleinberger, G., Madore, C., Daria, A., Werner, G., Krasemann, S., Capell, A., Trümbach, D., Wurst, W., Brunner, B., Bultmann, S., Tahirovic, S., Kerschensteiner, M., Misgeld, T., Butovsky, O., & Haass, C. (2017). TREM2 deficiency impairs chemotaxis and microglial responses to neuronal injury. *EMBO Reports*, 18(7), 1186–1198. <https://doi.org/10.15252/embr.201743922>

Meyer, R. C., Giddens, M. M., Coleman, B. M., & Hall, R. A. (2014). The protective role of prosaposin and its receptors in the nervous system. *Brain Research*, 1585, 1–12. <https://doi.org/10.1016/j.brainres.2014.08.022>

Moursel, L. G., van Roon-Mom, W. M. C., Kielbasa, S. M., Mei, H., Buermans, H. P. J., van der Graaf, L. M., Hettne, K. M., de Meijer, E. J., van Duinen, S. G., Laros, J. F. J., van Buchem, M. A., 't Hoen, P. A. C., van der Maarel, S. M., & van der Weerd, L. (2018). Brain transcriptomic analysis of hereditary cerebral hemorrhage with amyloidosis-Dutch type. *Frontiers in Aging Neuroscience*, 10, 102–102. <https://doi.org/10.3389/fnagi.2018.00102>

Mulder, S. D., Veerhuis, R., Blankenstein, M. A., & Nielsen, H. M. (2012). The effect of amyloid associated proteins on the expression of genes involved in amyloid- β clearance by adult human astrocytes. *Experimental Neurology*, 233(1), 373–379. <https://doi.org/10.1016/j.expneurol.2011.11.001>

Nilsson, P., Saito, T., & Saido, T. C. (2014). New Mouse Model of Alzheimer's. *ACS Chemical Neuroscience*, 5(7), 499–502. <https://doi.org/10.1021/cn500105p>

Paik, D. T. D. T., Tian, L. L., Williams, I. M. I. M., Rhee, S. S., Zhang, H. H., Liu, C. C., Mishra, R. R., Wu, S. M. S. M., Red-Horse, K. K., & Wu, J. C. J. C. (2020). Single-cell RNA-seq Unveils Unique Transcriptomic Signatures of Organ-Specific Endothelial Cells. *Circulation (New York, N.Y.)*, 142(19), 1848–1862. <https://doi.org/10.1161/CIRCULATIONAHA.119.041433>

Ramilowski, J. A., Goldberg, T., Harshbarger, J., Kloppman, E., Lizio, M., Satagopam, V. P., Itoh, M., Kawaji, H., Carninci,

P., Rost, B., & Forrest, A. R. R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. *Nature Communications*, 6(1), 7866–7866. <https://doi.org/10.1038/ncomms8866>

Ransohoff, R. M., Schafer, D., Vincent, A., Blachère, N. E., & Bar-Or, A. (2015). Neuroinflammation: Ways in Which the Immune System Affects the Brain. *Neurotherapeutics*, 12(4), 896–909. <https://doi.org/10.1007/s13311-015-0385-3>

Ransom, B. R., & Ransom, C. B. (2011). Astrocytes: Multitalented Stars of the Central Nervous System. In *Astrocytes*(Vol. 814, pp. 3–7). Humana Press. https://doi.org/10.1007/978-1-61779-452-0_1

Reitz, C. (2012). Alzheimer's and the Amyloid Cascade Hypothesis: A Critical Review. *International Journal of Alzheimer's*, 2012, 369808–369811. <https://doi.org/10.1155/2012/369808>

Ricciarelli, R., & Fedele, E. (2017). The Amyloid Cascade Hypothesis in Alzheimer's : It's Time to Change Our Mind. *Current Neuropharmacology*, 15(6), 926–935. <https://doi.org/10.2174/1570159X15666170116143743>

Ries, M., & Sastre, M. (2016). Mechanisms of A β clearance and degradation by glial cells. *Frontiers in Aging Neuroscience*, 8, 160–160. <https://doi.org/10.3389/fnagi.2016.00160>

Roychowdhury, T., Lu, H., Hornsby, W. E., Crone, B., Wang, G. T., Guo, D., Sendamarai, A. K., Devineni, P., Lin, M., Zhou, W., Graham, S. E., Wolford, B. N., Surakka, I., Wang, Z., Chang, L., Zhang, J., Mathis, M., Brummett, C. M., Melendez, T. L., ... Willer, C. J. (2021). Regulatory variants in TCF7L2 are associated with thoracic aortic aneurysm. *American Journal of Human Genetics*, 108(9), 1578–1589. <https://doi.org/10.1016/j.ajhg.2021.06.016>

Selkoe, D. J., & Hardy, J. (2016). The amyloid hypothesis of Alzheimer's at 25 years. *EMBO Molecular Medicine*, 8(6), 595–608. <https://doi.org/10.15252/emmm.201606210>

Seto, M., Dumitrescu, L., Mahoney, E. R., Sclafani, A. M., De Jager, P. L., Menon, V., Koran, M. E. I., Robinson, R. A., Ruderfer, D. M., Cox, N. J., Seyfried, N. T., Jefferson, A. L., Schneider, J. A., Bennett, D. A., Petyuk, V. A., & Hohman, T. J. (2023). Multi-omic characterization of brain changes in the vascular endothelial growth factor family during aging and Alzheimer's disease. *Neurobiology of Aging*, 126, 25–33. <https://doi.org/10.1016/j.neurobiolaging.2023.01.010>

Swaminathan, S., Kim, S., Shen, L., Risacher, S. L., Foroud, T., Pankratz, N., Potkin, S. G., Huentelman, M. J., Craig, D. W., Weiner, M. W., Saykin, A. J., & (ADNI), T. A. D. N. I. (2011). Genomic Copy Number Analysis in Alzheimer's Disease and Mild Cognitive Impairment: An ADNI Study. *International Journal of Alzheimer's Disease*, 2011, 729478–10. <https://doi.org/10.4061/2011/729478>

Teich, A. F., & Arancio, O. (2012). Is the Amyloid Hypothesis of Alzheimer's therapeutically relevant? *Biochemical Journal*, 446(2), 165–177. <https://doi.org/10.1042/BJ20120653>

Tondo, G., Perani, D., & Comi, C. (2019). TAM Receptor Pathways at the Crossroads of Neuroinflammation and Neurodegeneration. *Disease Markers*, 2019, 2387614–13. <https://doi.org/10.1155/2019/2387614>

Ulrich, J. D., Finn, M. B., Wang, Y., Shen, A., Mahan, T. E., Jiang, H., Stewart, F. R., Piccio, L., Colonna, M., & Holtzman, D. M. (2014). Altered microglial response to A β plaques in APPPS1-21 mice heterozygous for TREM2. *Molecular Neurodegeneration*, 9(1), 20–20. <https://doi.org/10.1186/1750-1326-9-20>

Villegas-Llerena, C., Phillips, A., Garcia-Reitboeck, P., Hardy, J., & Pocock, J. M. (2015). Microglial genes regulating neuroinflammation in the progression of Alzheimer's . *Current Opinion in Neurobiology*, 36, 74–81.

<https://doi.org/10.1016/j.conb.2015.10.004>

Wang, H., Kulas, J. A., Wang, C., Holtzman, D. M., Ferris, H. A., & Hansen, S. B. (2021). Regulation of beta-amyloid production in neurons by astrocyte-derived cholesterol. *Proceedings of the National Academy of Sciences - PNAS*, 118(33), 1–. <https://doi.org/10.1073/pnas.2102191118>

Wang, H., Kulas, J. A., Wang, C., Holtzman, D. M., Ferris, H. A., & Hansen, S. B. (2021). Regulation of beta-amyloid production in neurons by astrocyte-derived cholesterol. *Proceedings of the National Academy of Sciences - PNAS*, 118(33), 1–. <https://doi.org/10.1073/pnas.2102191118>

Wang, Y., Ulland, T. K., Ulrich, J. D., Song, W., Tzaferis, J. A., Hole, J. T., Yuan, P., Mahan, T. E., Shi, Y., Gilfillan, S., Cella, M., Grutzendler, J., DeMattos, R. B., Cirrito, J. R., Holtzman, D. M., & Colonna, M. (2016). TREM2-mediated early microglial response limits diffusion and toxicity of amyloid plaques. *The Journal of Experimental Medicine*, 213(5), 667–675. <https://doi.org/10.1084/jem.20151948>

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6), 80–83. <https://doi.org/10.2307/3001968>

Xiu, J., Nordberg, A., Zhang, J.-T., & Guan, Z.-Z. (2005). Expression of nicotinic receptors on primary cultures of rat astrocytes and up-regulation of the alpha₇, alpha₄ and beta₂ subunits in response to nanomolar concentrations of the beta-amyloid peptide(1-42). *Neurochemistry International*, 47(4), 281–290. <https://doi.org/10.1016/j.neuint.2005.04.023>

Yeh, F. L., Wang, Y., Tom, I., Gonzalez, L. C., & Sheng, M. (2016). TREM2 Binds to Apolipoproteins, Including APOE and CLU/APOJ, and Thereby Facilitates Uptake of Amyloid-Beta by Microglia. *Neuron (Cambridge, Mass.)*, 91(2), 328–340. <https://doi.org/10.1016/j.neuron.2016.06.015>

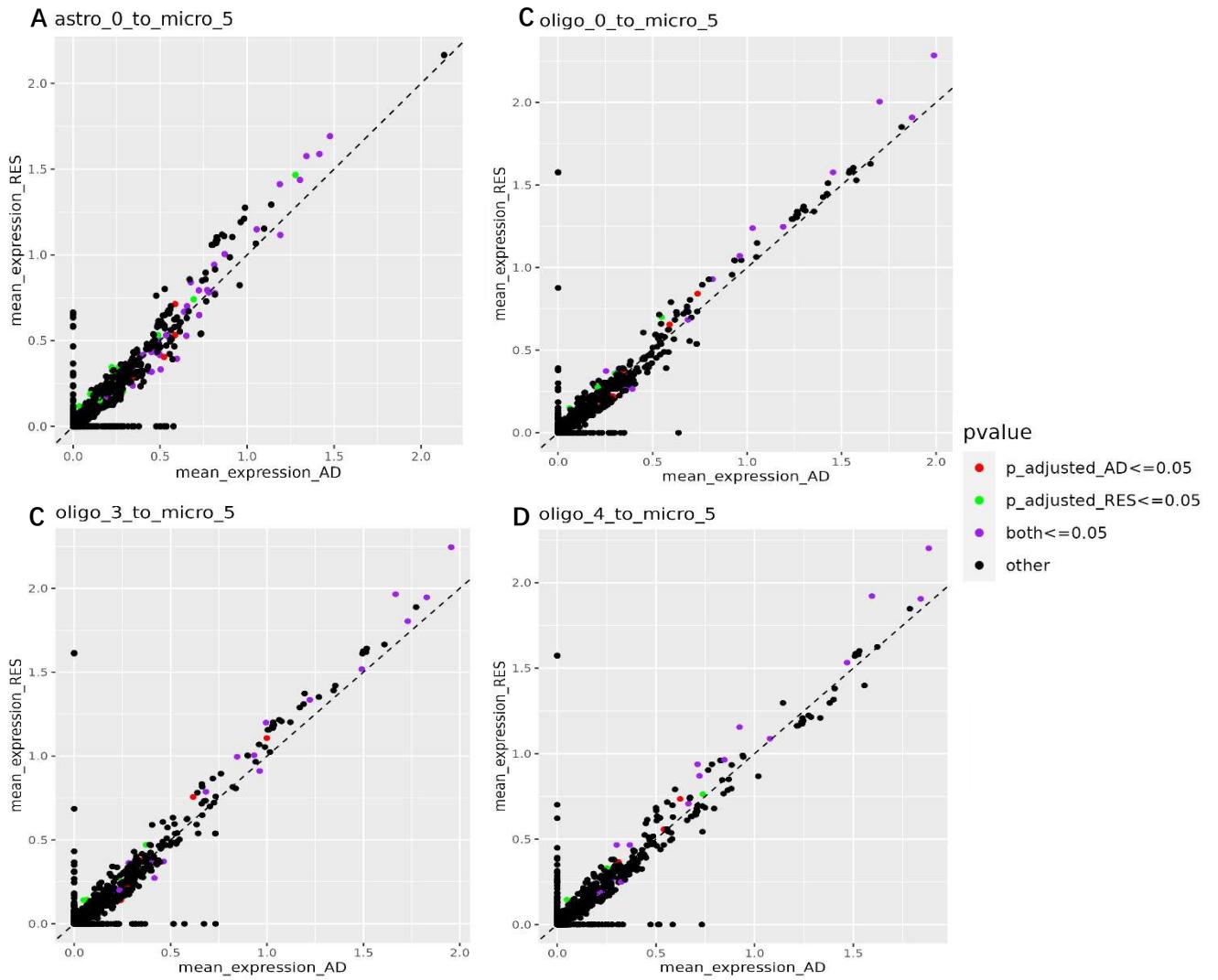
Yin, C., Ackermann, S., Ma, Z., Mohanta, S. K., Zhang, C., Li, Y., Nietzsche, S., Westermann, M., Peng, L., Hu, D., Bontha, S. V., Srikanthapu, P., Beer, M., Megens, R. T. A., Steffens, S., Hildner, M., Halder, L. D., Eckstein, H.-H., Pelisek, J., ... Habenicht, A. J. R. (2019). ApoE attenuates unresolvable inflammation by complex formation with activated C1q. *Nature Medicine*, 25(3), 496–506. <https://doi.org/10.1038/s41591-018-0336-8>

Yin, K.-J., Cirrito, J. R., Yan, P., Hu, X., Xiao, Q., Pan, X., Bateman, R., Song, H., Hsu, F.-F., Turk, J., Xu, J., Hsu, C. Y., Mills, J. C., Holtzman, D. M., & Lee, J.-M. (2006). Matrix Metalloproteinases Expressed by Astrocytes Mediate Extracellular Amyloid-beta Peptide Catabolism. *The Journal of Neuroscience*, 26(43), 10939–10948. <https://doi.org/10.1523/JNEUROSCI.2085-06.2006>

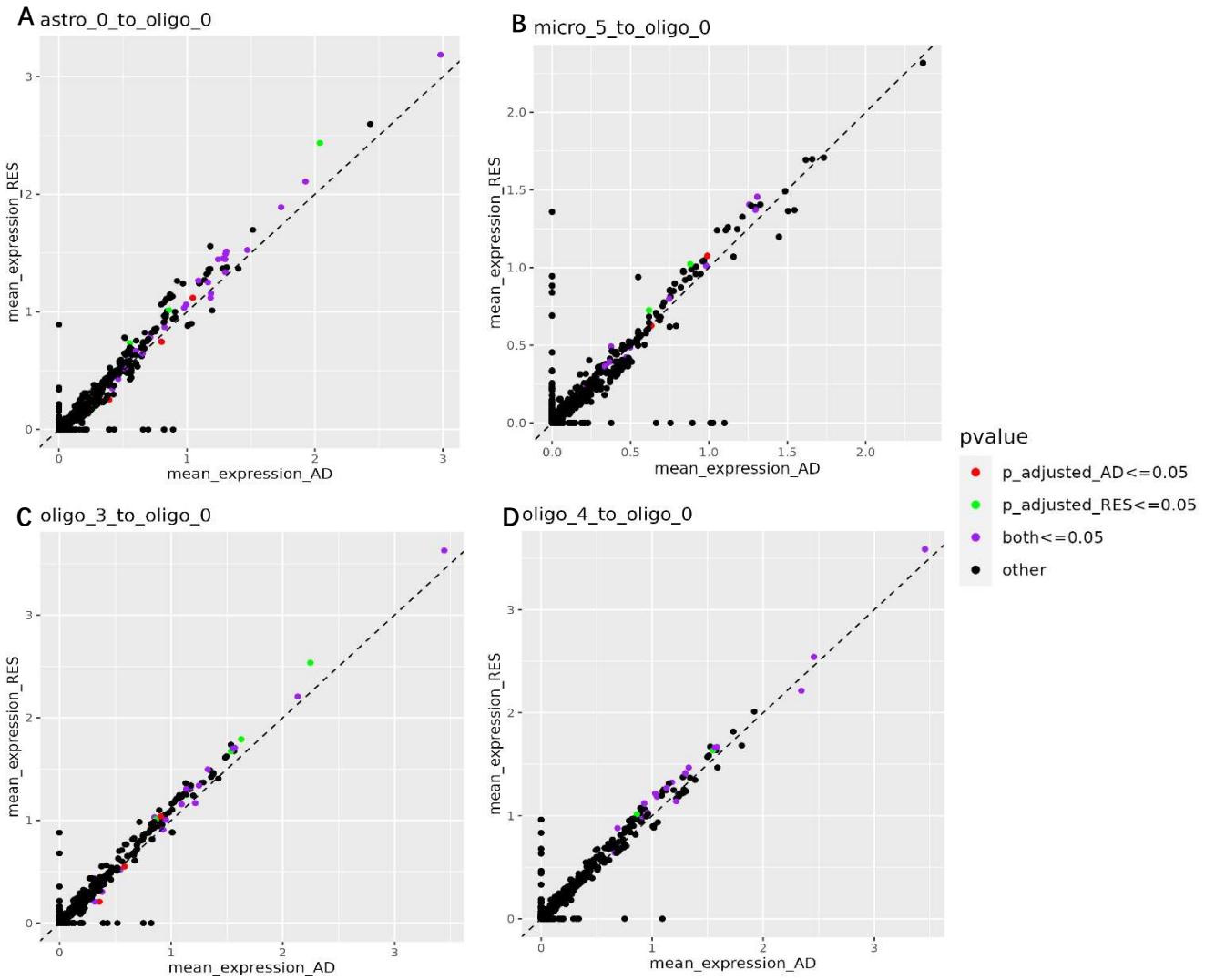
Zhang, T., Yu, J., Wang, G., & Zhang, R. (2021). Amyloid precursor protein binds with TNFRSF21 to induce neural inflammation in Alzheimer's Disease. *European Journal of Pharmaceutical Sciences*, 157, 105598–105598. <https://doi.org/10.1016/j.ejps.2020.105598>

Zigman, W. B., Devenny, D. A., Krinsky-McHale, S. J., Jenkins, E. C., Urv, T. K., Wegiel, J., Schupf, N., & Silverman, W. (2008). ALZHEIMER'S IN ADULTS WITH DOWN SYNDROME. *International Review of Research in Mental Retardation*, 36, 103–145. [https://doi.org/10.1016/S0074-7750\(08\)00004-9](https://doi.org/10.1016/S0074-7750(08)00004-9)

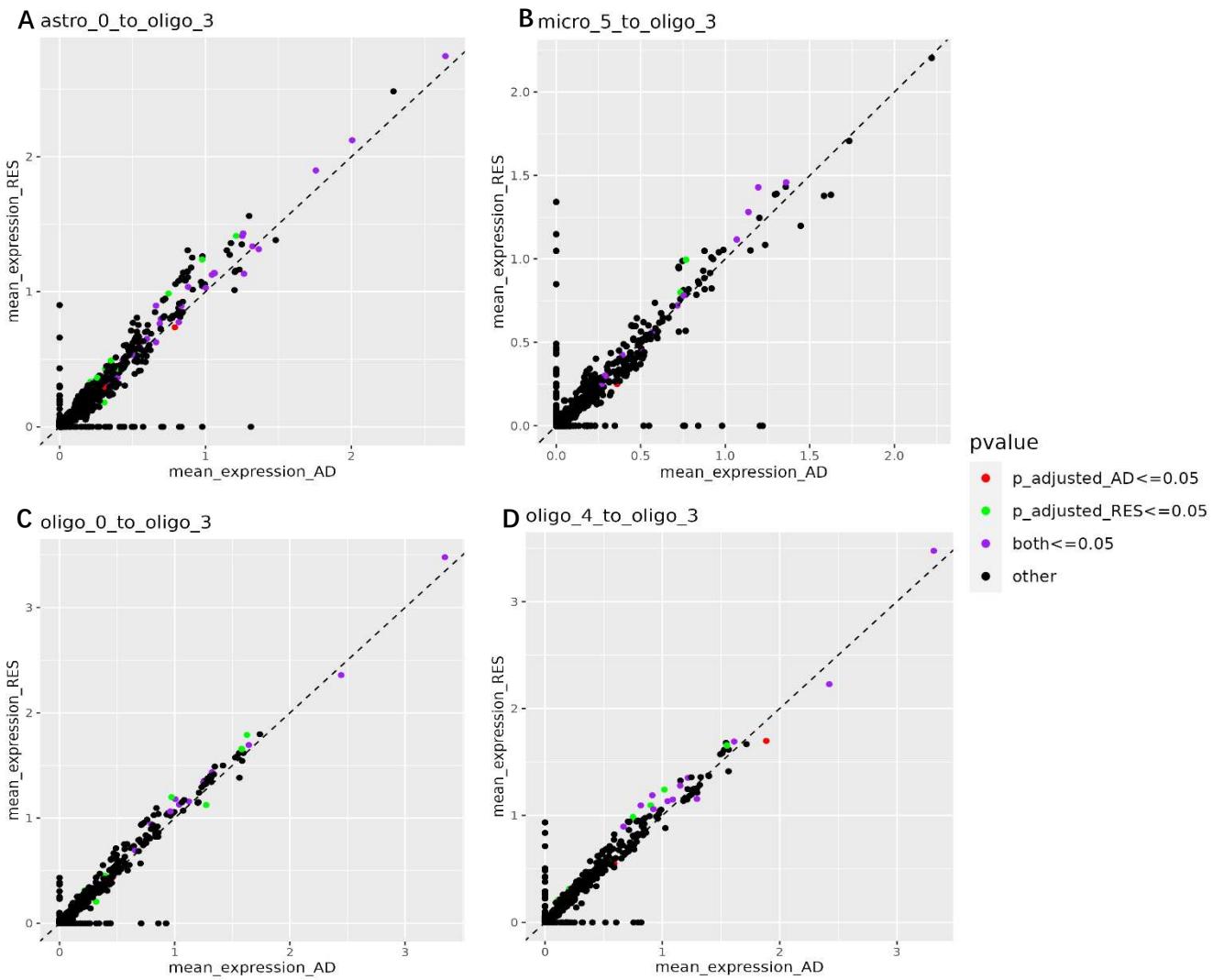
7 Appendix



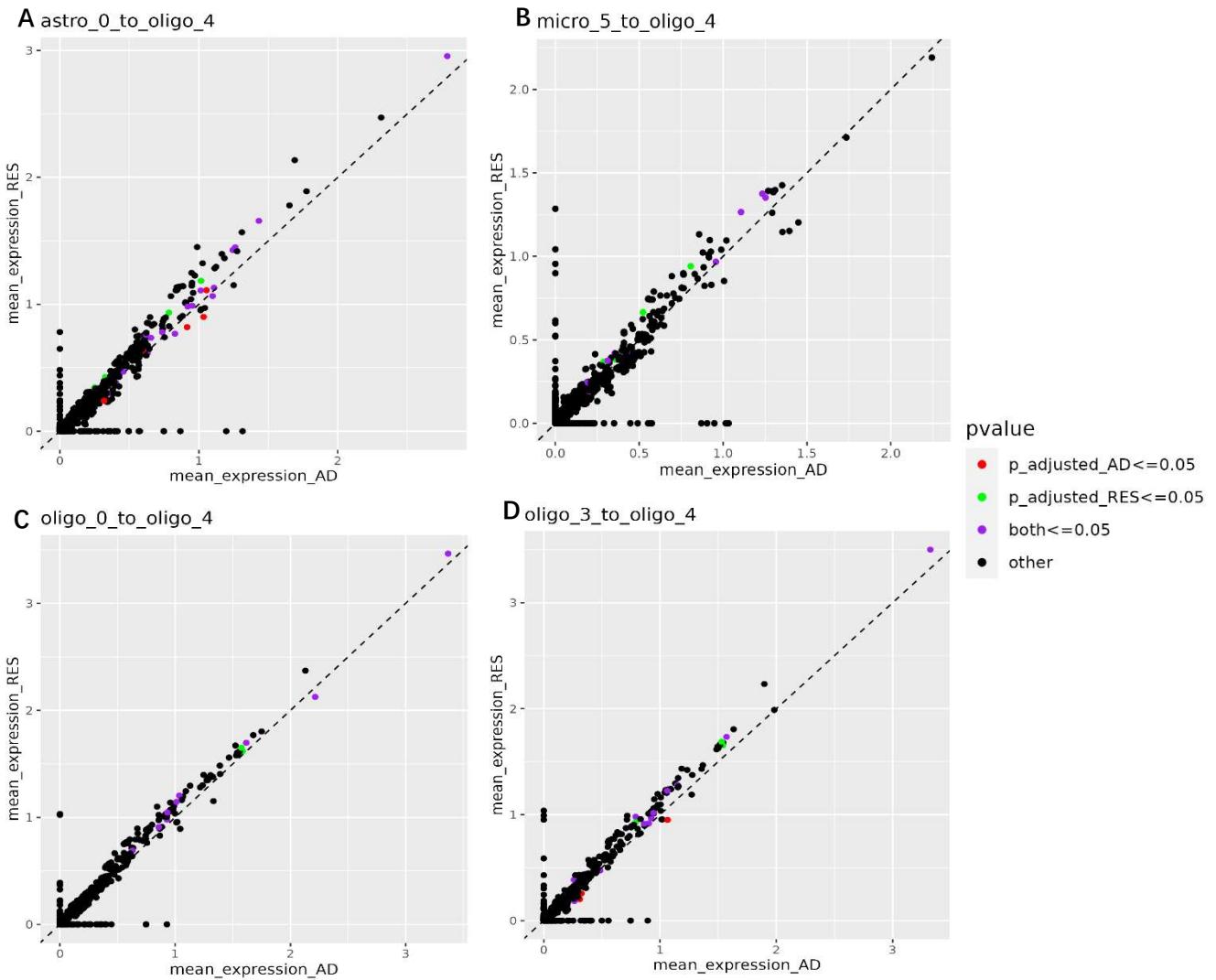
Supplement Figure 1: The scatter plots of mean expression of interactions in AD against mean expression of interactions in RES when receiver cell-type was micro_5. The dashed line (slope = 1, intercept = 0) was used to indicate the changes of interactions in AD and RES. The red dots represented interactions that only had $p_{adjusted_AD} \leq 0.05$, the green dots represented interactions that only had $p_{adjusted_RES} \leq 0.05$, the purple dots represented interactions with $p_{adjusted_AD} \leq 0.05$ and $p_{adjusted_RES} \leq 0.05$, and the black dots represented the remaining interactions. (A) interaction found when astro_0 as sender cell-type. (B) interaction found when oligo_0 as sender cell-type. (C) interaction found when oligo_3 as sender cell-type. (D) interaction found when oligo_4 as sender cell-type



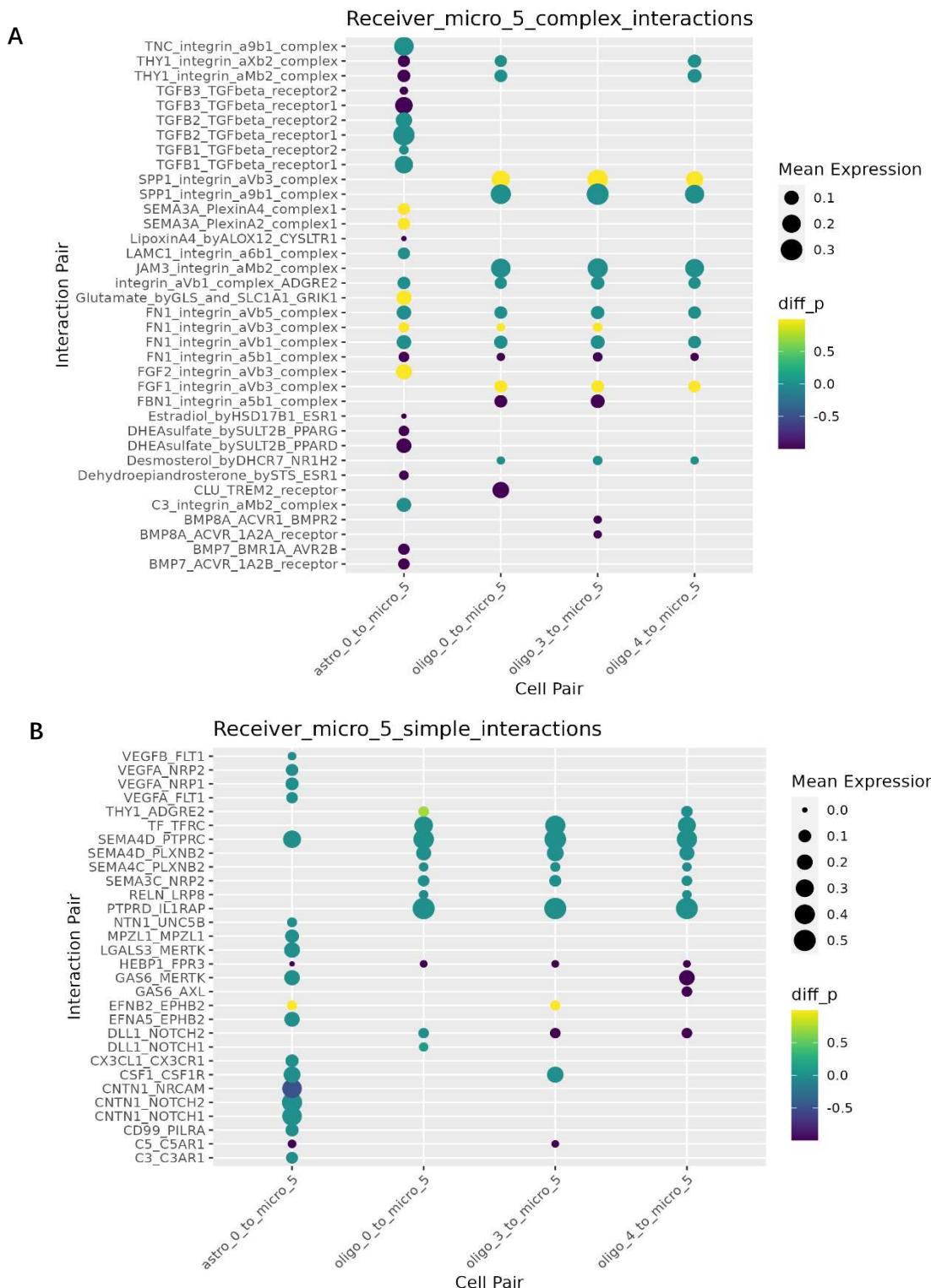
Supplement Figure 2: The scatter plots of mean expression of interactions in AD against mean expression of interactions in RES when receiver cell-type was oligo_0. The dashed line (slope = 1, intercept = 0) was used to indicate the changes of interactions in AD and RES. The red dots represented interactions that only had $p_{adjusted_AD} \leq 0.05$, the green dots represented interactions that only had $p_{adjusted_RES} \leq 0.05$, the purple dots represented interactions with $p_{adjusted_AD} \leq 0.05$ and $p_{adjusted_RES} \leq 0.05$, and the black dots represented the remaining interactions. (A) interaction found when astro_0 as sender cell-type. (B) interaction found when micro_5 as sender cell-type. (C) interaction found when oligo_3 as sender cell-type. (D) interaction found when oligo_4 as sender cell-type



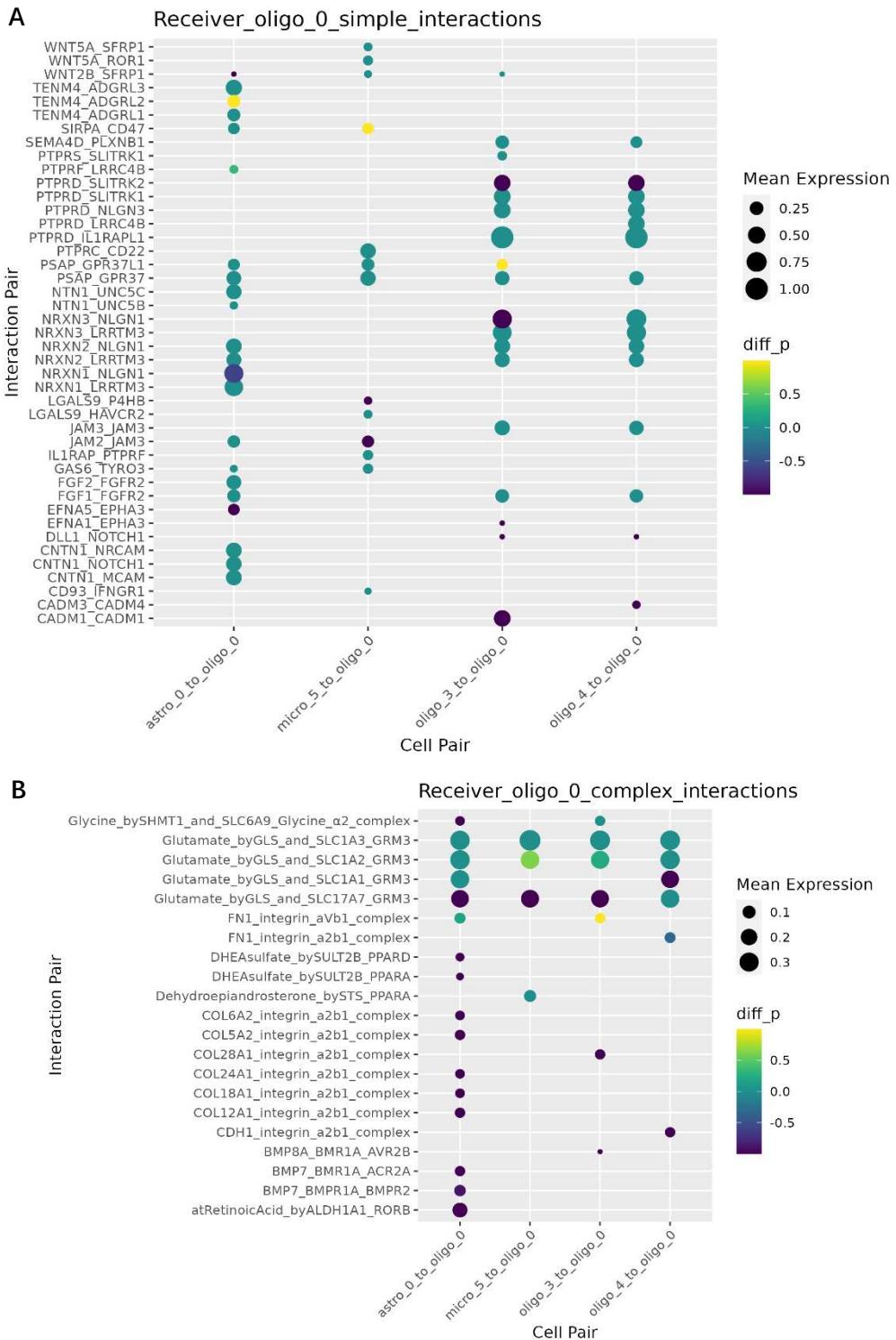
Supplement Figure 3: The scatter plots of mean expression of interactions in AD against mean expression of interactions in RES when receiver cell-type was oligo_3. The dashed line (slope = 1, intercept = 0) was used to indicate the changes of interactions in AD and RES. The red dots represented interactions that only had $p_{adjusted_AD} \leq 0.05$, the green dots represented interactions that only had $p_{adjusted_RES} \leq 0.05$, the purple dots represented interactions with $p_{adjusted_AD} \leq 0.05$ and $p_{adjusted_RES} \leq 0.05$, and the black dots represented the remaining interactions. (A) interaction found when astro_0 as sender cell-type. (B) interaction found when micro_5 as sender cell-type. (C) interaction found when oligo_0 as sender cell-type. (D) interaction found when oligo_4 as sender cell-type



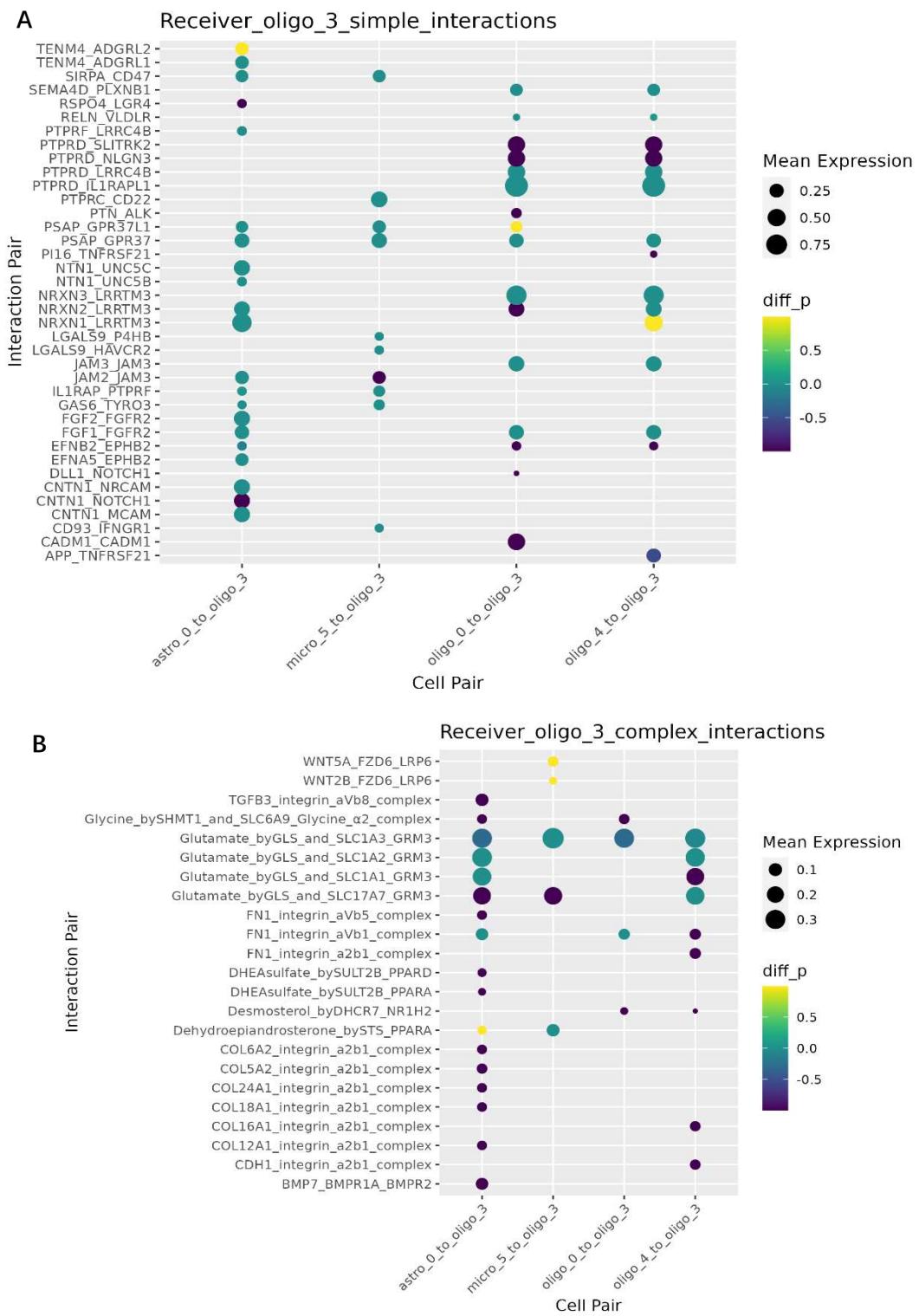
Supplement Figure 4: The scatter plots of mean expression of interactions in AD against mean expression of interactions in RES when receiver cell-type was oligo_4. The dashed line (slope = 1, intercept = 0) was used to indicate the changes of interactions in AD and RES. The red dots represented interactions that only had $p_{adjusted_AD} \leq 0.05$, the green dots represented interactions that only had $p_{adjusted_RES} \leq 0.05$, the purple dots represented interactions with $p_{adjusted_AD} \leq 0.05$ and $p_{adjusted_RES} \leq 0.05$, and the black dots represented the remaining interactions. (A) interaction found when astro_0 as sender cell-type. (B) interaction found when micro_5 as sender cell-type. (C) interaction found when oligo_0 as sender cell-type. (D) interaction found when oligo_3 as sender cell-type



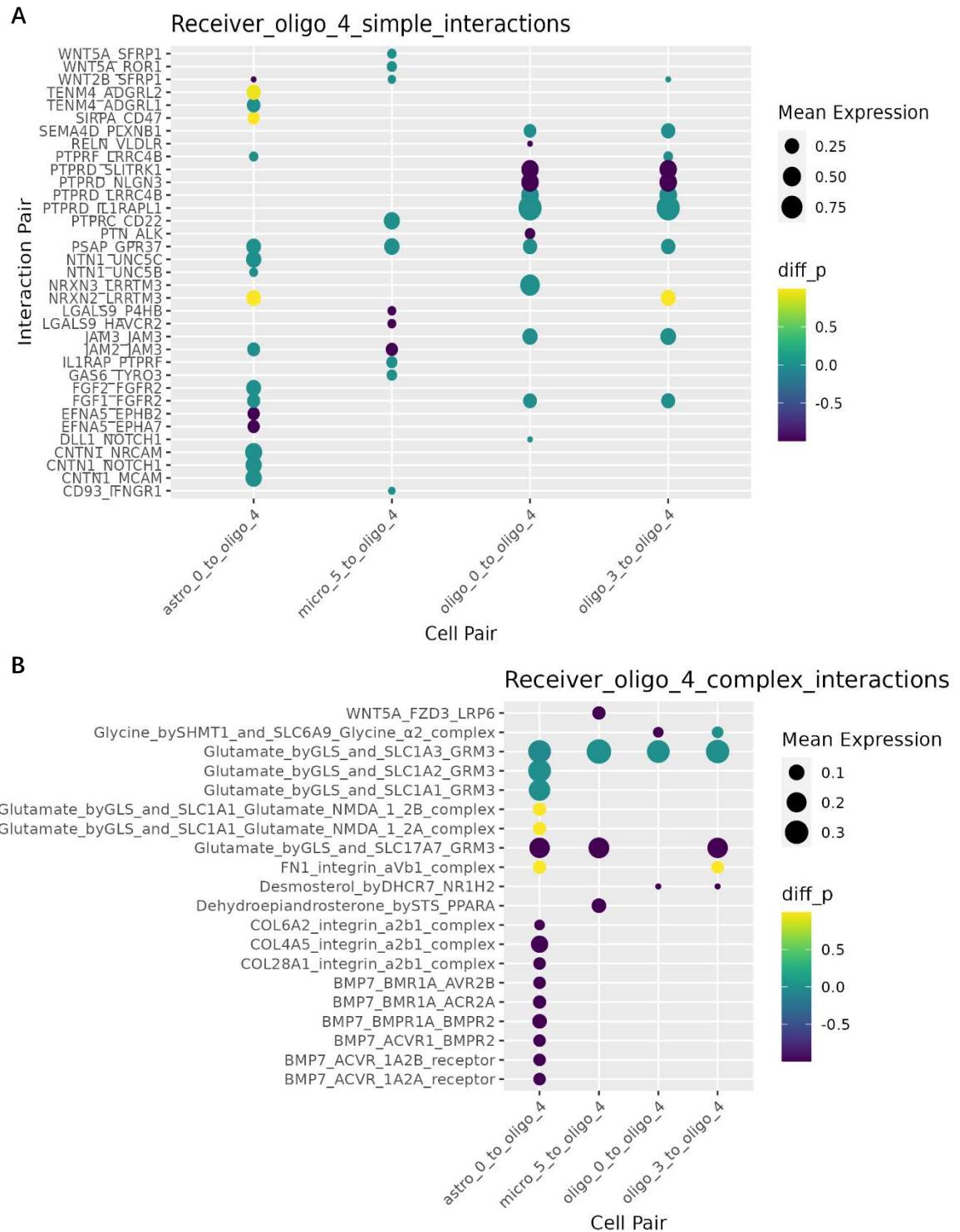
Supplement Figure 5: Dot plot of celltype-specific interactions when the receiver was micro_5. The other PACs were senders, which were astro_0, oligo_0, oligo_3 and oligo_4. $\text{diff}_p = -(\text{p}_{\text{adjusted_AD}} - \text{p}_{\text{adjusted_RES}})$. Mean expression = (mean expression in AD + mean expression in RES + mean expression in ND)/3. All of these interactions were celltype-specific in AD or RES, but not in ND. If the color of the dot was yellow, the interaction was celltype-specific in AD, if the color was purple, the interaction was celltype-specific in RES. If the color was green, the interaction was celltype-specific in both AD and RES (A) Dot plot of celltype-specific interactions only containing simple ligands and simple receptors. (B) Dot plot of celltype-specific interactions containing complex.



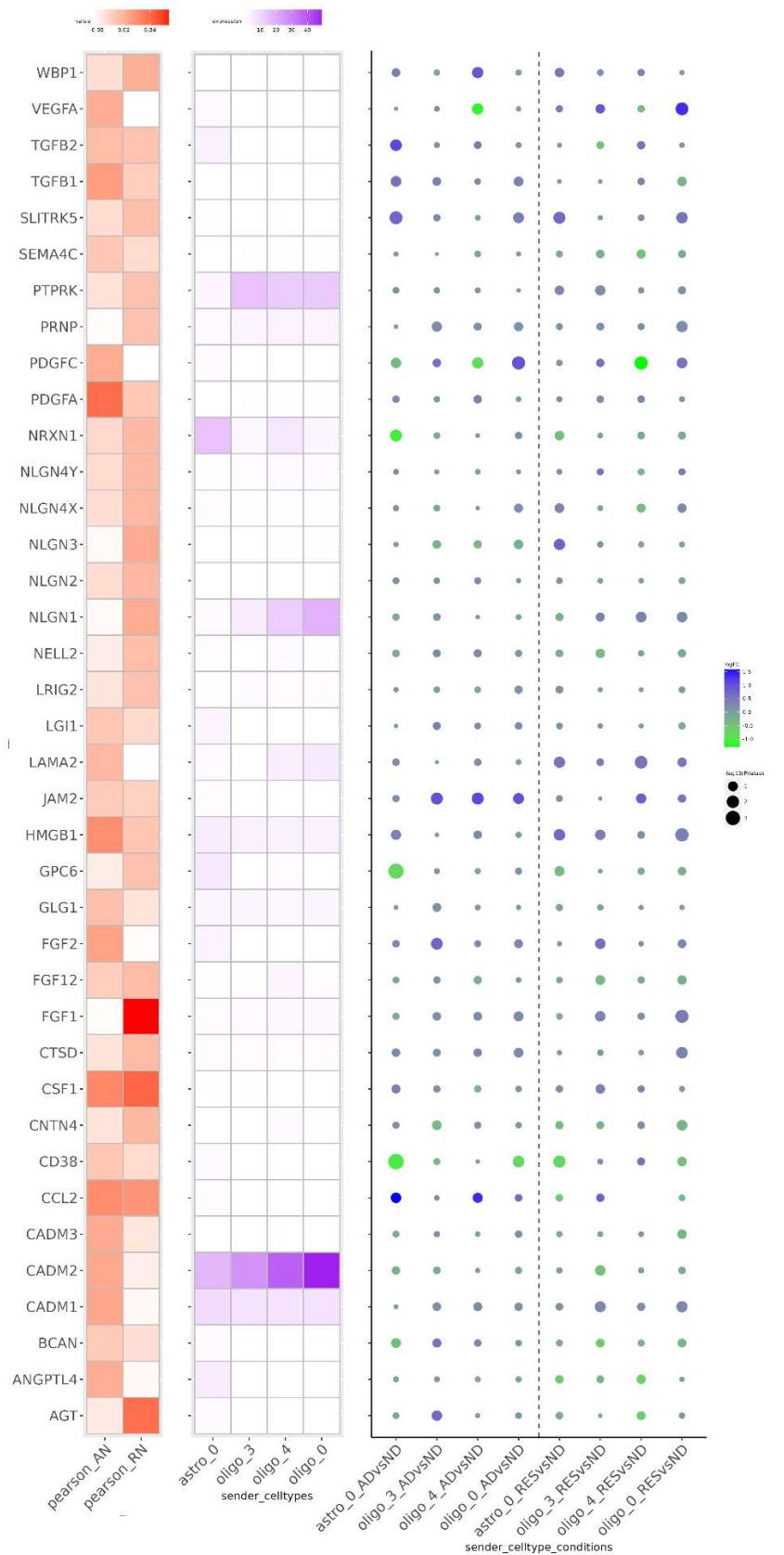
Supplement Figure 6: Dot plot of celltype-specific interactions when the receiver was oligo_0. The other PACs were senders, which were astro_0, micro_5, oligo_3 and oligo_4. $\text{diff_p} = -(\text{p_adjusted_AD} - \text{p_adjusted_RES})$. Mean expression = $(\text{mean expression in AD} + \text{mean expression in RES} + \text{mean expression in ND})/3$. All of these interactions were celltype-specific in AD or RES, but not in ND. If the color of the dot was yellow, the interaction was celltype-specific in AD, if the color was purple, the interaction was celltype-specific in RES. If the color was green, the interaction was celltype-specific in both AD and RES (A) Dot plot of celltype-specific interactions only containing simple ligands and simple receptors. (B) Dot plot of celltype-specific interactions containing complex.



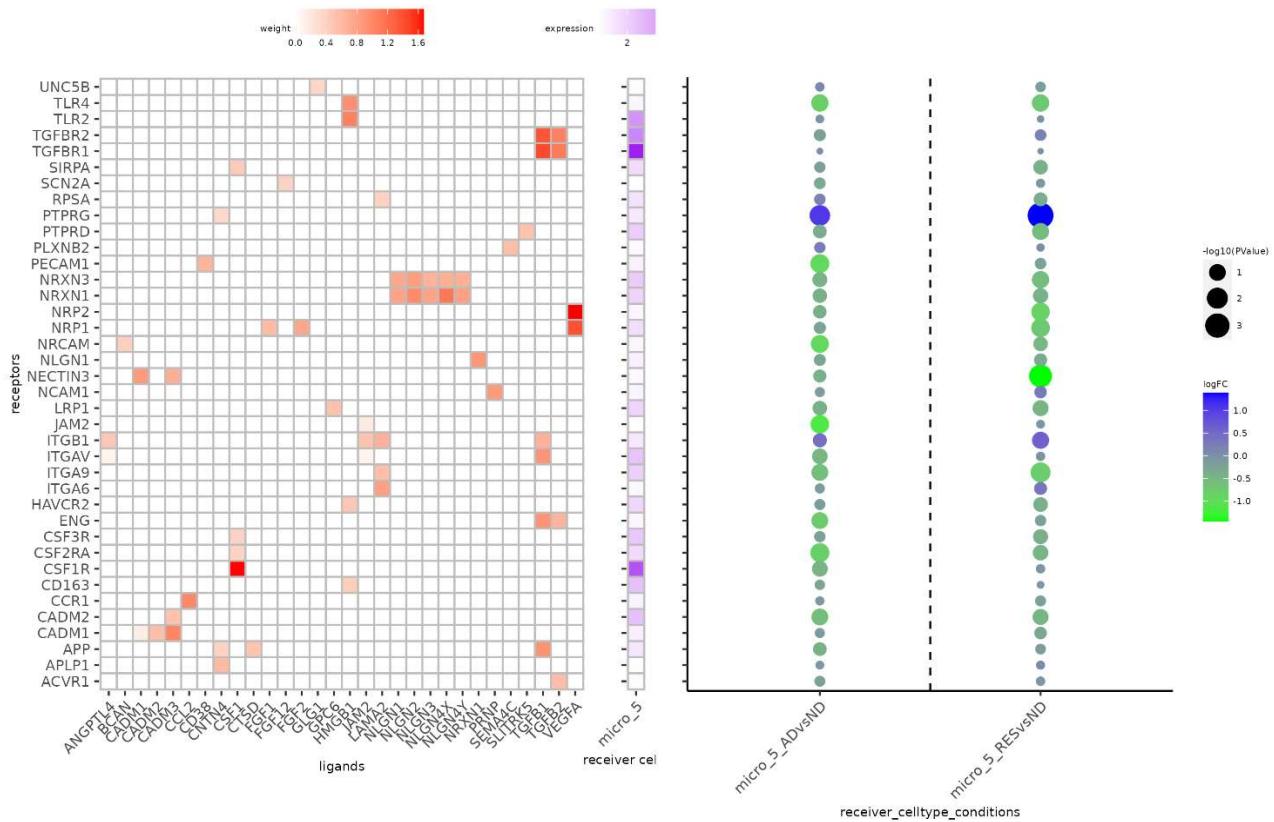
Supplement Figure 7: Dot plot of celltype-specific interactions when the receiver was oligo_3. The other PACs were senders, which were astro_0, oligo_0, micro_5 and oligo_4. $\text{diff_p} = -(\text{p_adjusted_AD} - \text{p_adjusted_RES})$. Mean expression = $(\text{mean expression in AD} + \text{mean expression in RES} + \text{mean expression in ND})/3$. All of these interactions were celltype-specific in AD or RES, but not in ND. If the color of the dot was yellow, the interaction was celltype-specific in AD, if the color was purple, the interaction was celltype-specific in RES. If the color was green, the interaction was celltype-specific in both AD and RES (A) Dot plot of celltype-specific interactions only containing simple ligands and simple receptors. (B) Dot plot of celltype-specific interactions containing complex.



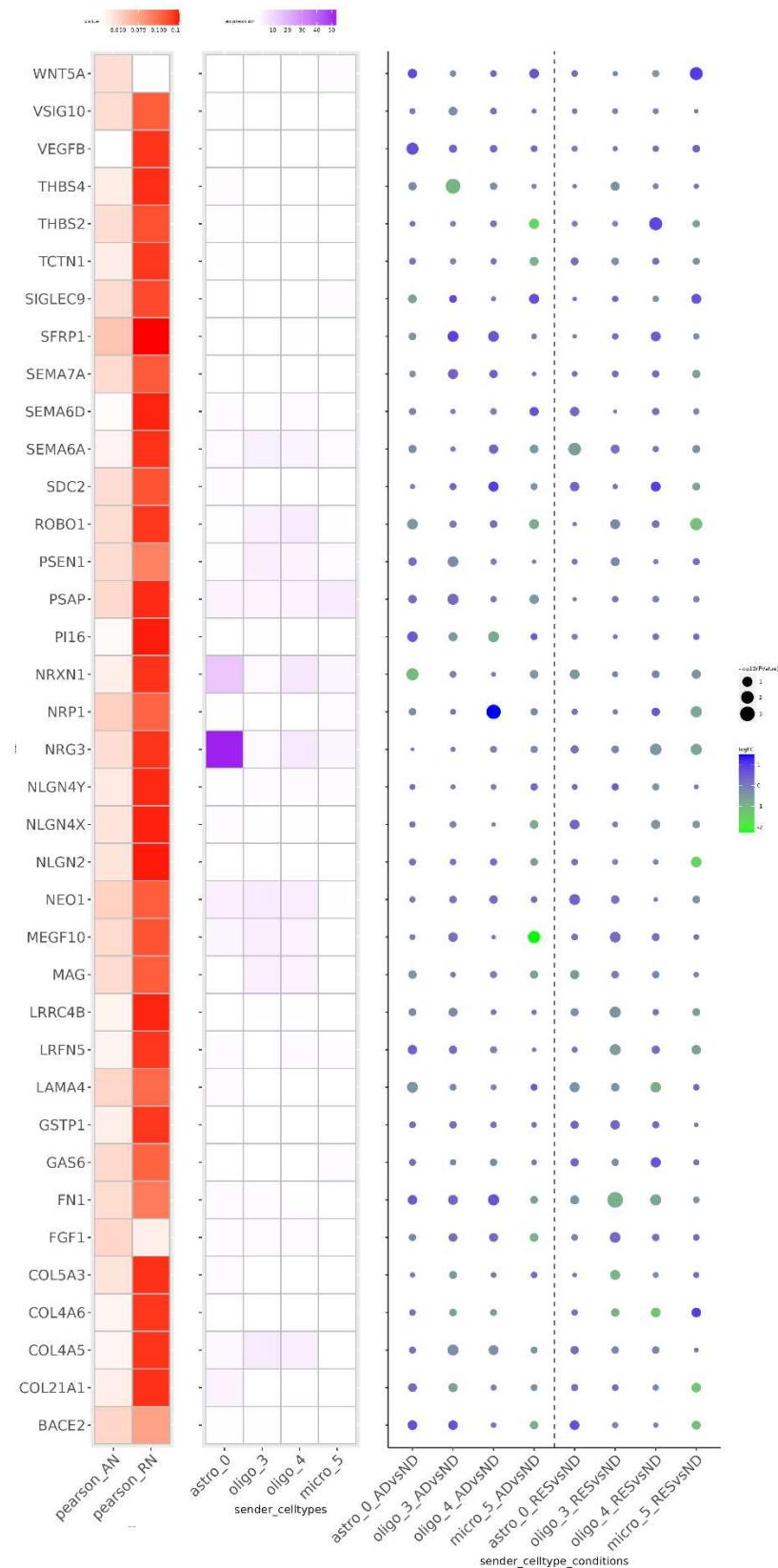
Supplement Figure 8: Dot plot of celltype-specific interactions when the receiver was oligo_4. The other PACs were senders, which were astro_0, oligo_0, oligo_3 and micro_5. $\text{diff_p} = -(\text{p_adjusted_AD} - \text{p_adjusted_RES})$. Mean expression = (mean expression in AD + mean expression in RES + mean expression in ND)/3. All of these interactions were celltype-specific in AD or RES, but not in ND. If the color of the dot was yellow, the interaction was celltype-specific in AD, if the color was purple, the interaction was celltype-specific in RES. If the color was green, the interaction was celltype-specific in both AD and RES (A) Dot plot of celltype-specific interactions only containing simple ligands and simple receptors. (B) Dot plot of celltype-specific interactions containing complex.



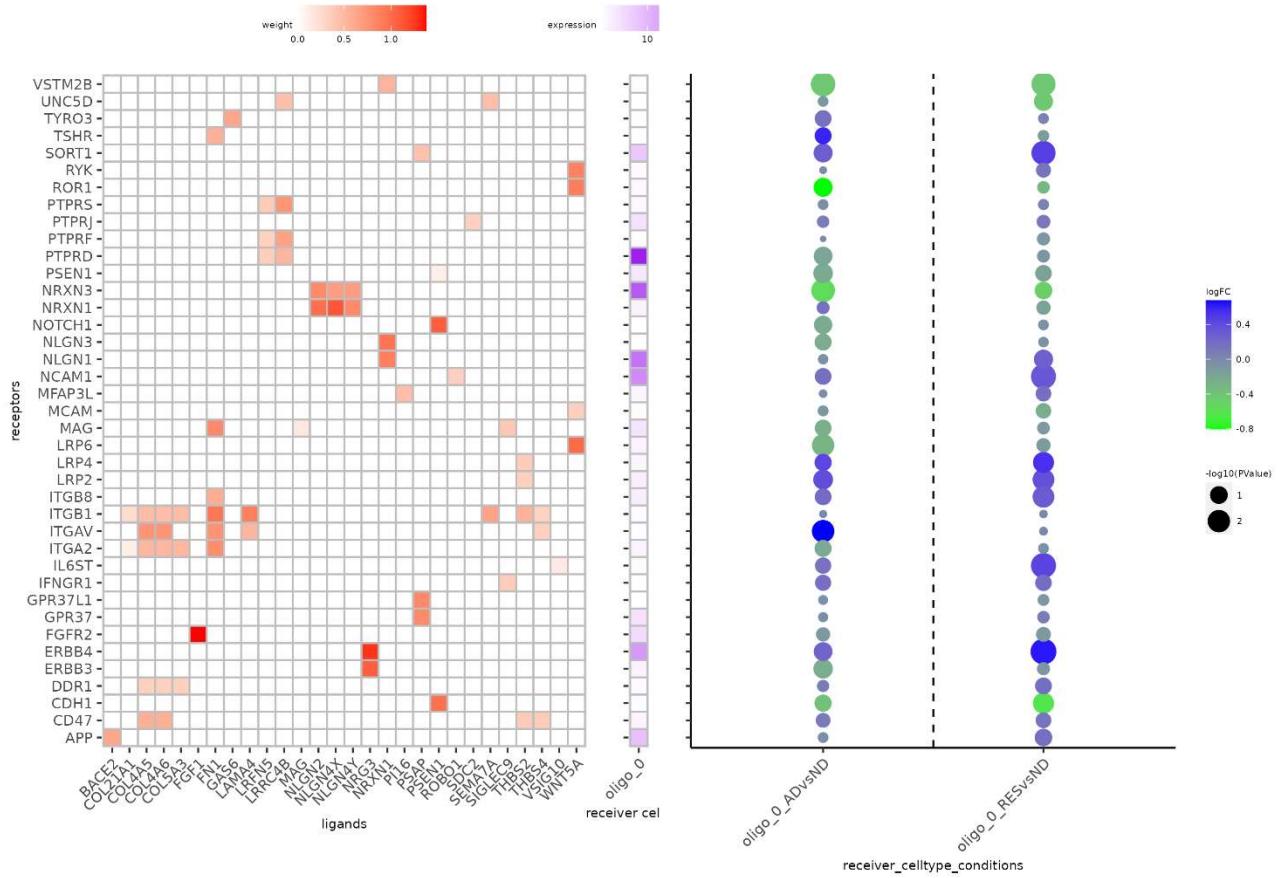
Supplement Figure 9: From left to right, the plots represent the ligand activity, expression and log fold change(lfc) of each potential ligand for receiver micro_5. In the dot plot to visualize log fold change of ligands, the color of dots indicates the lfc value, the size of dots represents the -log10(p-value) of the DE ligands.



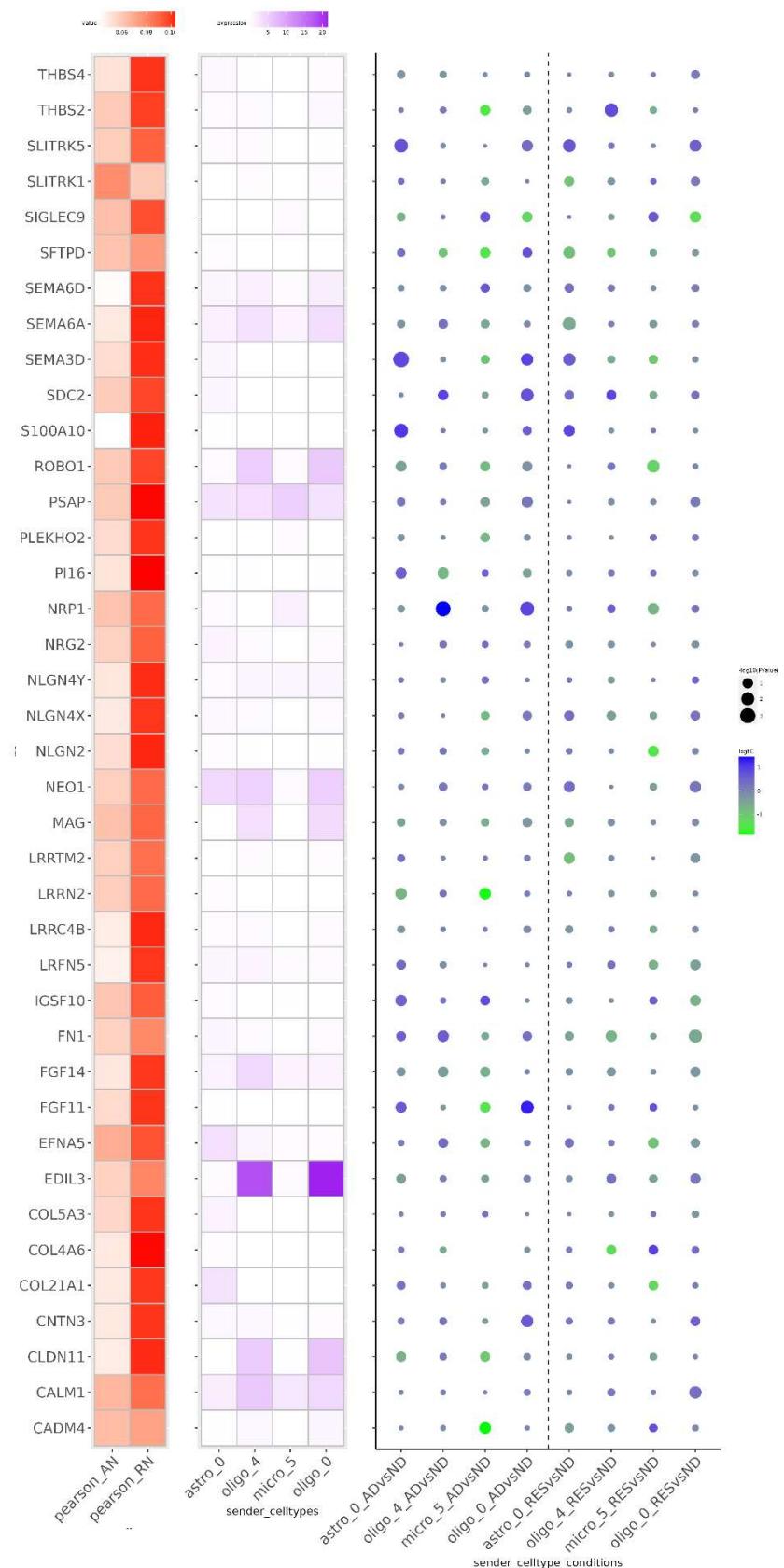
Supplement Figure 10: From left to right, the plots represent the ligand-receptor weights, expression and log fold change(lfc) of each receptor in receiver micro_5. The ligand-receptor weights matrix was generated from the database of NicheNet. In the dot plot to visualize log fold change of receptors, the color of dots indicates the lfc value, the size of dots represents the $-\log_{10}(p\text{-value})$ of the DE receptors.



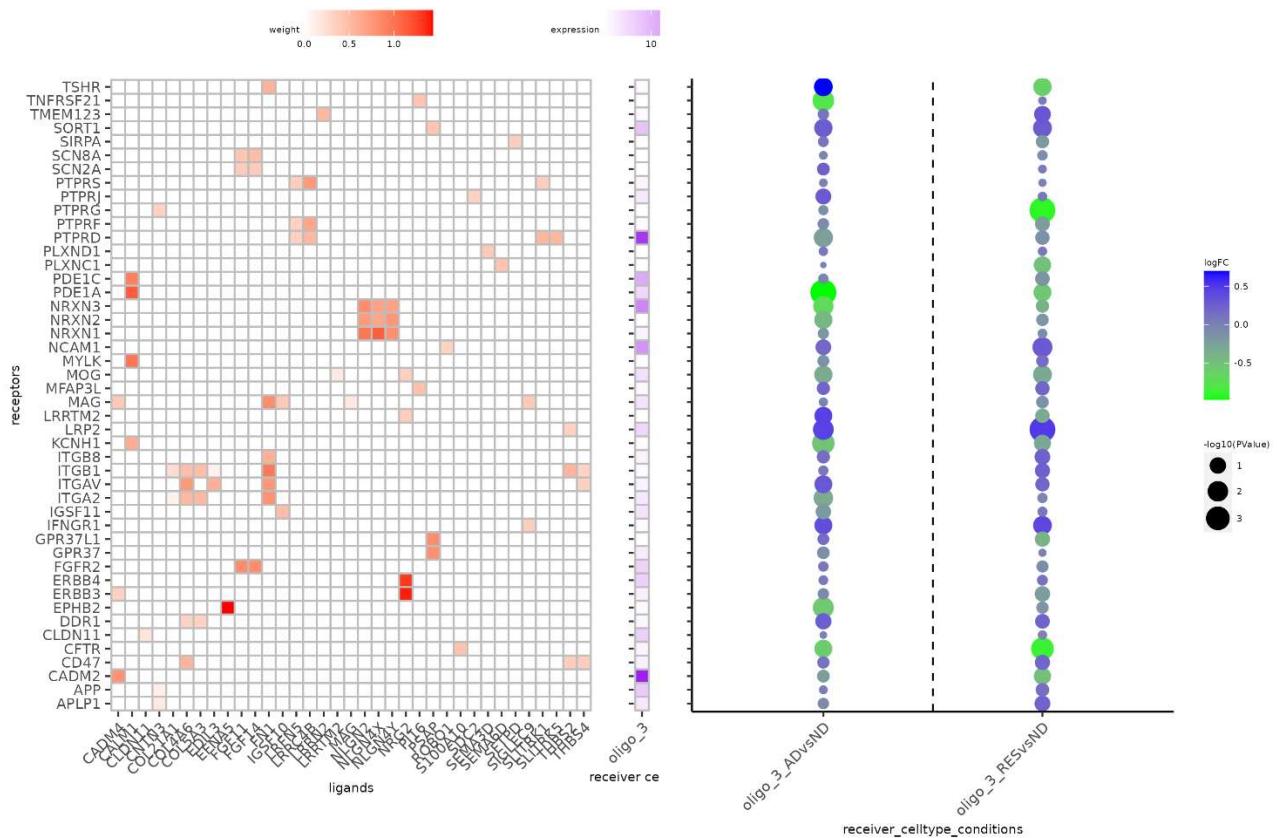
Supplement Figure 11: From left to right, the plots represent the ligand activity, expression and log fold change(lfc) of each potential ligand for receiver oligo_0. In the dot plot to visualize log fold change of ligands, the color of dots indicates the lfc value, the size of dots represents the -log10(p-value) of the DE ligands.



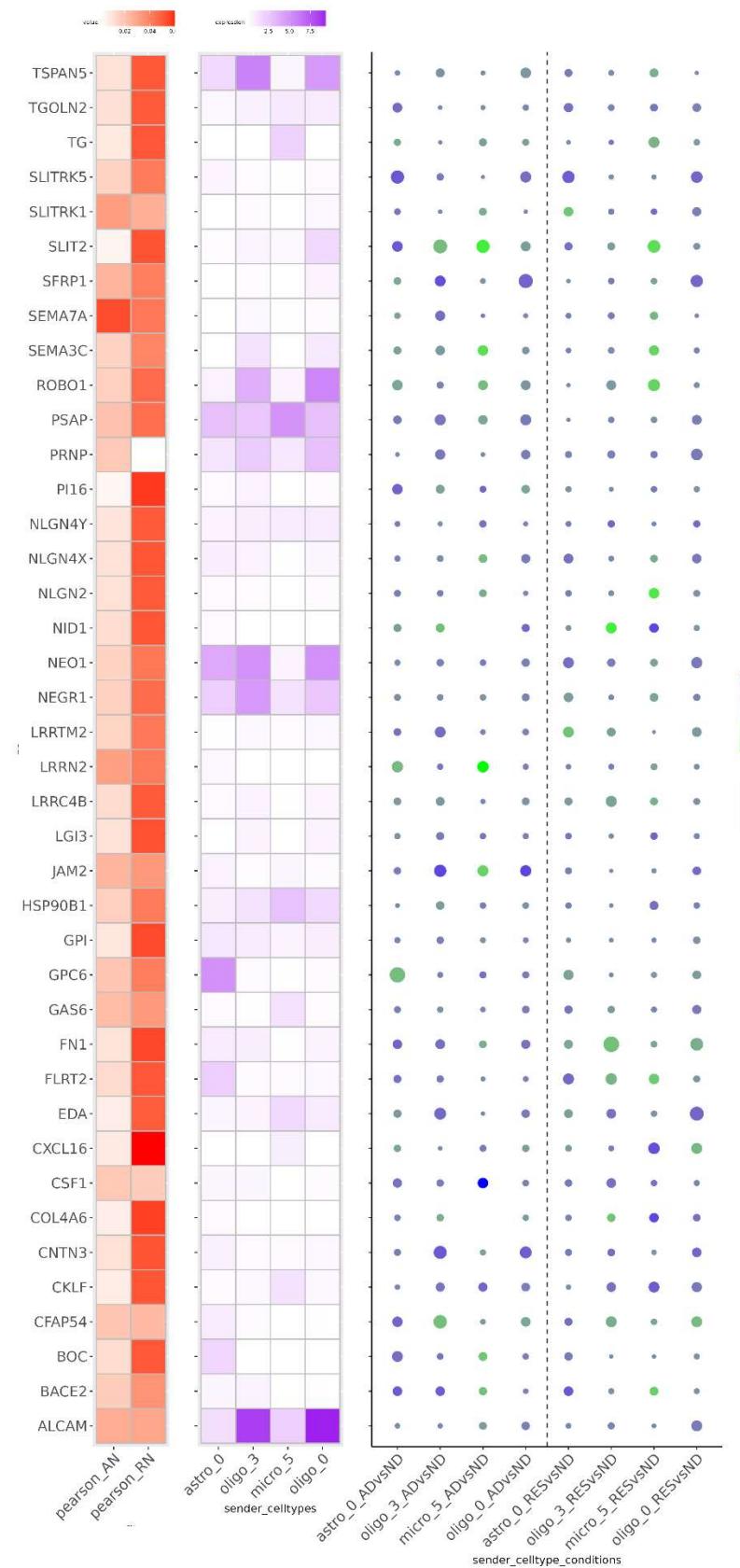
Supplement Figure 12: From left to right, the plots represent the ligand-receptor weights, expression and log fold change(lfc) of each receptor in receiver oligo_0. The ligand-receptor weights matrix was generated from the database of NicheNet. In the dot plot to visualize log fold change of receptors, the color of dots indicates the lfc value, the size of dots represents the $-\log_{10}(p\text{-value})$ of the DE receptors.



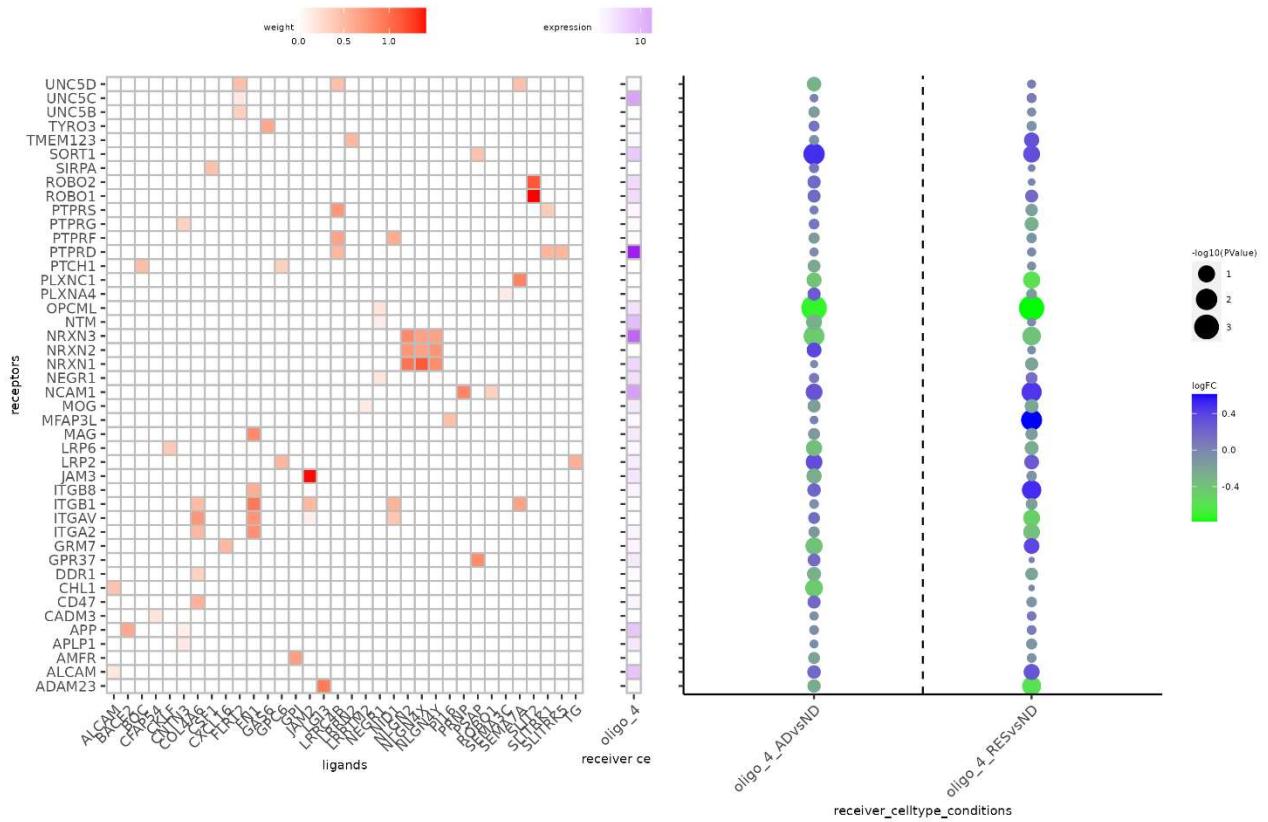
Supplement Figure 13: From left to right, the plots represent the ligand activity, expression and log fold change(lfc) of each potential ligand for receiver oligo_3. In the dot plot to visualize log fold change of ligands, the color of dots indicates the lfc value, the size of dots represents the -log10(p-value) of the DE ligands.



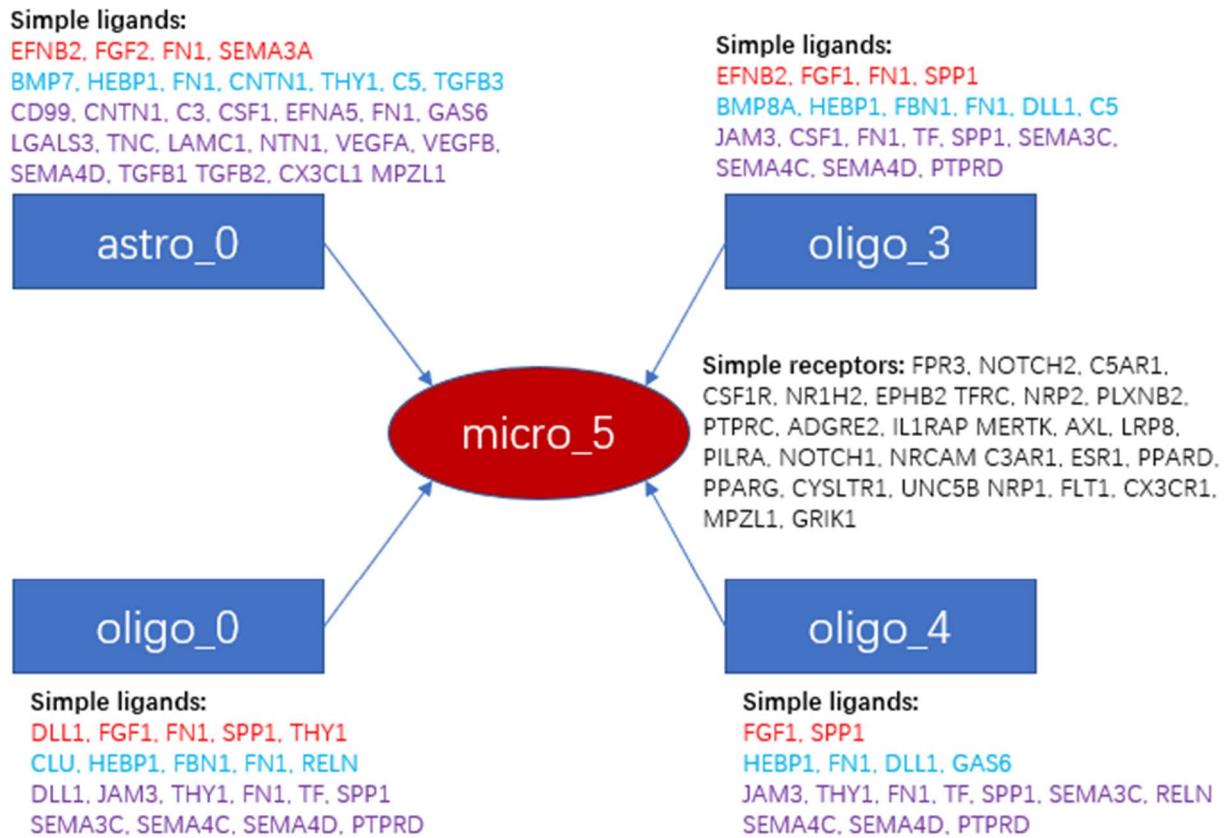
Supplement Figure 14: From left to right, the plots represent the ligand-receptor weights, expression and log fold change(lfc) of each receptor in receiver oligo_3. The ligand-receptor weights matrix was generated from the database of NicheNet. In the dot plot to visualize log fold change of receptors, the color of dots indicates the lfc value, the size of dots represents the $-\log_{10}(p\text{-value})$ of the DE receptors.



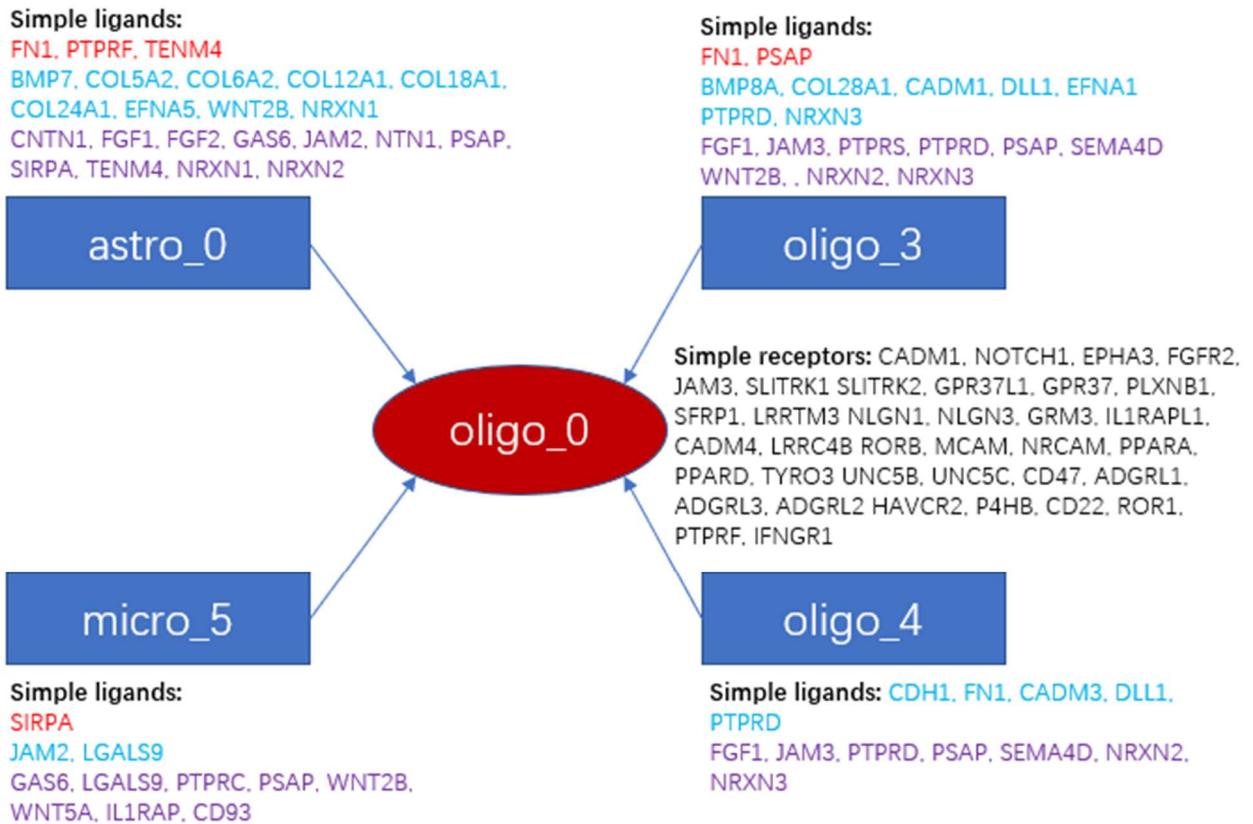
Supplement Figure 15: From left to right, the plots represent the ligand activity, expression and log fold change(lfc) of each potential ligand for receiver oligo_4. In the dot plot to visualize log fold change of ligands, the color of dots indicates the lfc value, the size of dots represents the -log10(p-value) of the DE ligands.



Supplement Figure 16: From left to right, the plots represent the ligand-receptor weights, expression and log fold change(lfc) of each receptor in receiver oligo_4. The ligand-receptor weights matrix was generated from the database of NicheNet. In the dot plot to visualize log fold change of receptors, the color of dots indicates the lfc value, the size of dots represents the $-\log_{10}(p\text{-value})$ of the DE receptors.



Supplement Figure 17: The simple ligands and simple receptors found in celltype-specific interactions when the receiver was **micro_5.** The red circle represents receivers, the blue squares represent senders. The ligands found in the interactions that were celltype-specific in AD were marked as red; The ligands found in the interactions that were celltype-specific in RES were marked as blue; The ligands found in the interactions that were celltype-specific in both AD and RES were marked as purple.



Supplement Figure 18: The simple ligands and simple receptors found in celltype-specific interactions when the receiver was **oligo_0.** The red circle represents receivers, the blue squares represent senders. The ligands found in the interactions that were celltype-specific in AD were marked as red; The ligands found in the interactions that were celltype-specific in RES were marked as blue; The ligands found in the interactions that were celltype-specific in both AD and RES were marked as purple.

Simple ligands:

IL1RAP, TENM4

BMP7, COL5A2, COL6A2, COL12A1, COL18A1, COL24A1, CNTN1, EFNB2, FN1, RSPO4, TGFB3
CNTN1, EFNA5, FGF1, FGF2, FN1, GAS6, JAM2, NTN1, PTPRF, PSAP, SIRPA, TENM4, NRXN1
NRXN2

astro_0

Simple ligands:

WNT2B, WNT5A

JAM2

GAS6, LGALS9, PTPRC, PSAP, SIRPA, IL1RAP
CD93

micro_5

oligo_3

oligo_0

oligo_4

Simple receptors: TNFRSF21, NR1H2, EPHB2, FGFR2, JAM3, LRRC4B, SLTRK2, VLDLR, GPR37, PLXNB1, LRRTM3, NLGN3, GRM3, IL1RAPL1, NOTCH1, MCAM, NRCAM, PPARA, PPARD, TYRO3, UNC5B, UNC5C, LGR4, GPR37L1, CD47, PTPRF, ADGRL1, ADGRL2, HAVCR2, P4HB, CD22, IFNGR1, CADM1, ALK

Simple ligands:

PSAP

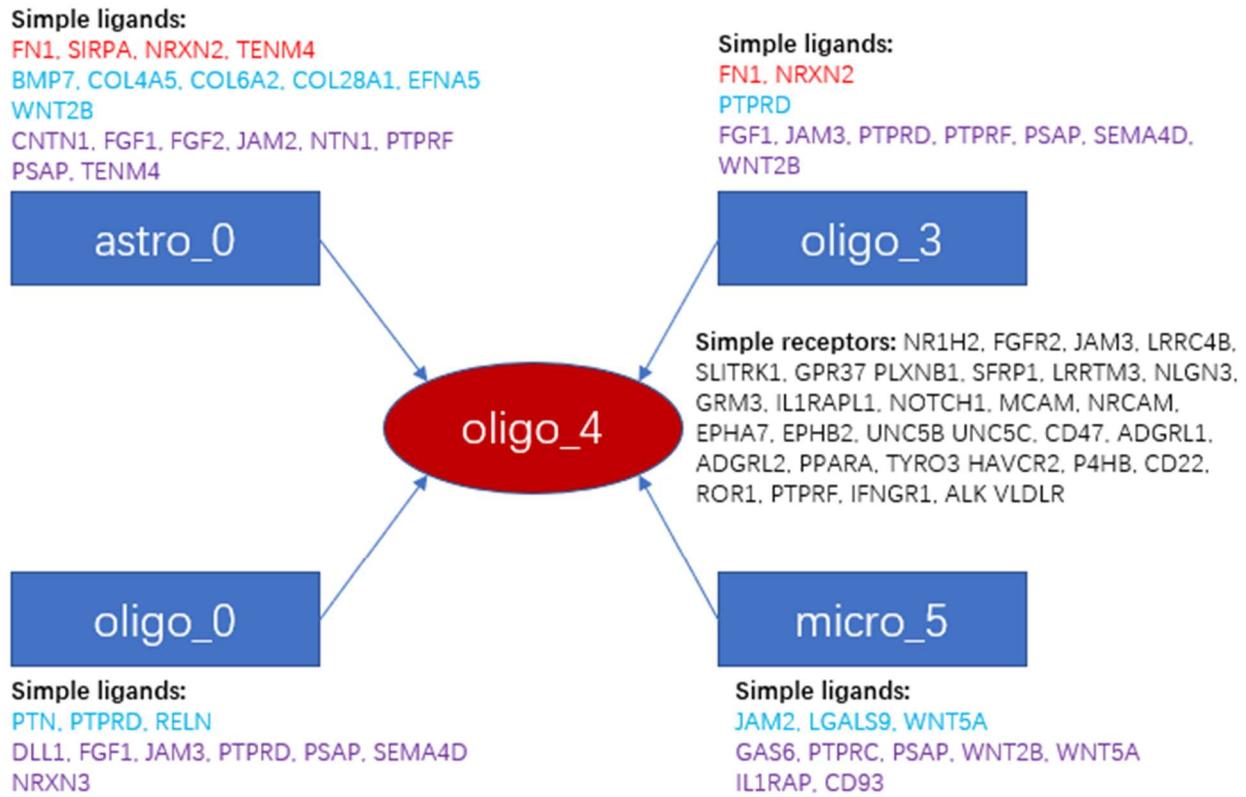
CADM1, DLL1, EFNB2, PTN, PTPRD, NRXN2
FGF1, FN1, JAM3, PTPRD, RELN, PSAP
SEMA4D, NRXN3

Simple ligands:

RELN, NRXN1

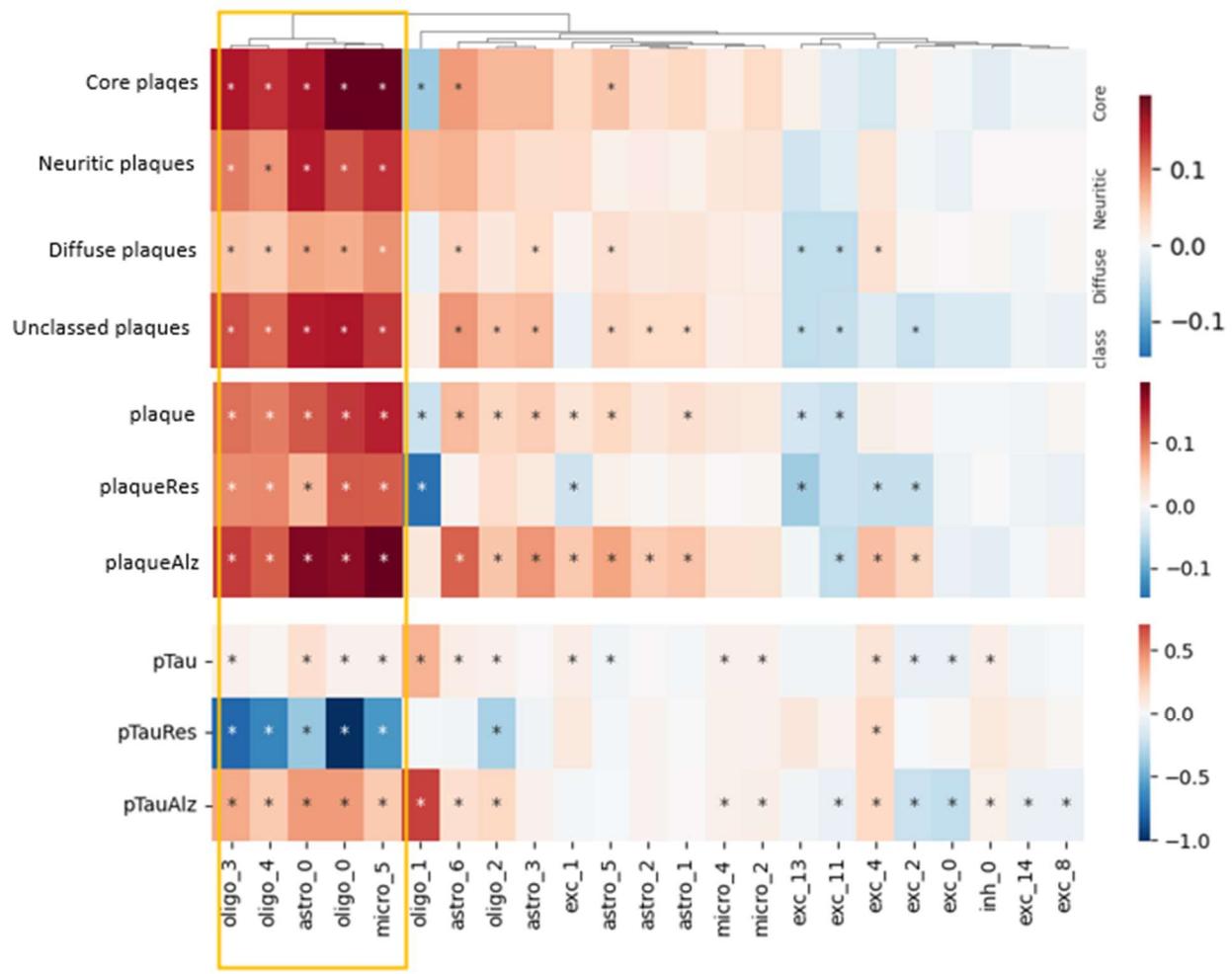
APP, PI16, CDH1, COL16A1, FN1, EFNB2, PTPRD
FGF1, JAM3, PTPRD, PSAP, SEMA4D, NRXN2
NRXN3

Supplement Figure 19: The simple ligands and simple receptors found in celltype-specific interactions when the receiver was oligo_3. The red circle represents receivers, the blue squares represent senders. The ligands found in the interactions that were celltype-specific in AD were marked as red; The ligands found in the interactions that were celltype-specific in RES were marked as blue; The ligands found in the interactions that were celltype-specific in both AD and RES were marked as purple.



Supplement Figure 20: The simple ligands and simple receptors found in celltype-specific interactions when the receiver was **oligo_4.** The red circle represents receivers, the blue squares represent senders. The ligands found in the interactions that were celltype-specific in AD were marked as red; The ligands found in the interactions that were celltype-specific in RES were marked as blue; The ligands found in the interactions that were celltype-specific in both AD and RES were marked as purple.

DA LFCs



Supplement Figure 21: Differential abundance (DA) analysis performed to investigate how cell abundances change in response to plaque pathology, pTAU pathology and different plaque types. Plaque type includes core plaques, neuritic plaques, diffused plaques and unclassified plaques. Plaque pathology includes plaque, plaqueRes (plaque is present without symptoms) and plaqueAlz (plaque is present with symptoms). pTAU pathology includes pTAU, pTAURes (pTAU is present without symptoms) and pTAUAlz (pTAU is present with symptoms). The orange square indicates five cell types are clustered into one group, named plaque associated cells (PACs)