T.R.
GEBZE TECHNICAL UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY

ETS GENE REGULATION IN NEURONAL DIFFERENTIATION: TRANSCRIPTOMIC ANALYSIS AND EVOLUTIONARY ASPECTS

YIGIT KORAY BABAL
A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
DEPARTMENT OF BIOTECHNOLOGY

GEBZE
2023

T.R.

GEBZE TECHNICAL UNIVERSITY

INSTITUTE OF BIOTECHNOLOGY

ETS GENE REGULATION IN NEURONAL DIFFERENTIATION: TRANSCRIPTOMIC ANALYSIS AND EVOLUTIONARY ASPECTS

YIGIT KORAY BABAL

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOTECHNOLOGY

THESIS SUPERVISOR

PROF. DR. ISIL KURNAZ

GEBZE

2023

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
BİYOTEKNOLOJİ ENSTİTÜSÜ

TRANSKRİPTOMİK VE EVRİMSEL YAKLAŞIMLA NÖRONAL FAKLILAŞMADA ETS GEN DÜZENLENMESİ

YİĞİT KORAY BABAL

DOKTORA TEZİ
BİYOTEKNOLOJİ ANABİLİM DALI

DANIŞMANI
PROF. DR. IŞIL KURNAZ

GEBZE
2023

GEBZE TEKNİK ÜNİVERSİTESİ	**DOKTORA JÜRİ ONAY FORMU**

GTÜ Biyoteknoloji Enstitüsü Yönetim Kurulu'nun 16/01/2023 tarih ve 2023/01 sayılı kararıyla oluşturulan jüri tarafından 19/01/2023 tarihinde tez savunma sınavı yapılan Yiğit Koray Babal'ın tez çalışması Biyoteknoloji Anabilim Dalında DOKTORA tezi olarak kabul edilmiştir.

JÜRİ

ÜYE
(TEZ DANIŞMANI) : PROF. DR. IŞIL KURNAZ

ÜYE : PROF. DR. TUNAHAN ÇAKIR

ÜYE : DR. ÖĞR. ÜYESİ MEHMET ÖZBİL

ÜYE : PROF. DR. KAZIM YALÇIN ARĞA

ÜYE : DR. ÖĞR. ÜYESİ BAŞAK KANDEMİR

ONAY

Gebze Teknik Üniversitesi Biyoteknoloji Enstitüsü Yönetim Kurulu'nun
……/……./...…. tarih ve …..…/…..... sayılı kararı.

SUMMARY

One of the key transcription factor superfamilies that regulates development, cell proliferation, and cancer progression is the ETS domain transcription factor family. ETS member proteins have been shown to have a variety of activities in the nervous system, but more research is needed to understand how these proteins affect neuronal cell differentiation and the evolutionary conservation of their target genes. In this thesis study, I focused on comparative transcriptomics to examine the regulatory function of ETS transcription factors in both glioma progression and neural development with their functional evolution. To investigate the regulatory function of ETS transcription factors in neural development across species, transcriptional patterns of ETS members and their target genes were compared using human, gorilla, mouse, fruit fly, and worm transcriptomics datasets. ETS member proteins control genes involved in neuronal differentiation, nervous system development, axon and synaptic modulation in many animals, according to bioinformatics methods used to evaluate the evolutionary conservation of ETS transcription factors during neuronal differentiation. As a result, we have identified both particular and widespread regulatory networks for these ETS proteins. Using a gene regulatory network inference technique, this study compares the transcriptome effects of ETS transcription factors on neuronal development. ETS members are in fact evolutionarily conserved in the control of neuronal development, according to an analysis of gene regulatory networks. However, it was shown that the ETS, PEA3, and ELF subfamilies have comparatively higher levels of activity as transcription factors involved in the transcriptional control of neural development.

Key Words: ETS Family, Gene Regulatory Networks, Transcriptomics.

ÖZET

Gelişimi, hücre proliferasyonunu ve kanser progresyonunu düzenleyen anahtar transkripsiyon faktörü süper ailelerinden biri ETS domain transkripsiyon faktörü ailesidir. ETS üyesi proteinlerin sinir sisteminde çeşitli aktivitelere sahip olduğu gösterilmiştir, ancak bu proteinlerin nöronal hücre farklılaşmasını ve hedef genlerinin evrimsel korunumunu nasıl etkilediğini anlamak için daha fazla araştırmaya ihtiyaç vardır. Bu tez çalışmasında, ETS transkripsiyon faktörlerinin hem glioma ilerlemesi hem de nöral gelişimdeki düzenleyici işlevini fonksiyonel evrimleriyle birlikte incelemek için karşılaştırmalı transkriptomiklere odaklandım. Türler arasında nöral gelişimde ETS transkripsiyon faktörlerinin düzenleyici işlevini araştırmak için, ETS üyelerinin ve hedef genlerinin transkripsiyonel modelleri insan, goril, fare, meyve sineği ve solucan transkriptomik veri kümeleri kullanılarak karşılaştırıldı. Nöronal farklılaşma sırasında ETS transkripsiyon faktörlerinin evrimsel korunumunu değerlendirmek için kullanılan biyoinformatik yöntemlere göre, ETS üyesi proteinler birçok hayvanda nöronal farklılaşma, sinir sistemi gelişimi, akson ve sinaptik modülasyonda yer alan genleri kontrol etmektedir. Sonuç olarak, bu ETS proteinleri için hem özel hem de yaygın düzenleyici ağlar belirledik. Bu çalışma, bir gen düzenleyici ağ çıkarım tekniği kullanarak, ETS transkripsiyon faktörlerinin nöronal gelişim üzerindeki transkriptom etkilerini karşılaştırmaktadır. Gen düzenleyici ağların analizine göre, ETS üyeleri aslında nöronal gelişimin kontrolünde evrimsel olarak korunmuştur. Bununla birlikte, ETS, PEA3 ve ELF alt ailelerinin, nöral gelişimin transkripsiyonel kontrolünde yer alan transkripsiyon faktörleri olarak nispeten daha yüksek aktivite seviyelerine sahip olduğu gösterilmiştir.

Anahtar Kelimeler: ETS Ailsi, Gen Regülasyon Ağları, Transkriptomiks.

ACKNOWLEDGEMENTS

Firstly, I would like to thank with a great sincerity to Prof. Dr. Isil AKSAN KURNAZ for her supervision, guidance and advice during the period of my research. I feel so lucky to be her student.

I would like to thank all the AxanLab members and allumi, my dear lab mates: Basak KANDEMİR, Ekin SONMEZ, Melis SAVASAN SOGUT, Oya ARI UYAR, Merve USTUN, Dogukan METINER, Büsra NUR CICEK, Betul ERTURAL, Esranur YIGIT, Pınar MUTLU, Atılay ILGUN, Sedef YUSUFOGLU for all their supports, friendship and all lovely memories.

I would like to thank Tugberk TALAK, Kadir KOCABAS, Ozgenur AKKUS, Ecehan ABDİK, Regan ODONGO, Emel KOKREK, Batuhan CAKIR, Gülben AVŞAR and Emre Taylan DUMAN for their friendship and moral supports.

I specially thank to my lovely mother Pınar BABAL for always being with me, she is most important thing in my life. I also endless thank to my brother Efe BABAL, my grandmother Melek BABAL, my aunt Sukran KASARCI, my uncle Erdal KASARCI, my cousins Taner KASARCI, Erdem BABAL, Didem BABAL and Haluk KASARCI for all their supports, patience and their endless love.

The most importantly, I also very special thanks to my dear wife Simge SENGUL for always being with me, her support, patience and her precious love. I wouldn't have come this far without her support.

TABLE of CONTENTS

	Page
SUMMARY	v
ÖZET	vi
ACKNOWLEDGEMENTS	vii
LIST of ABBREVIATIONS and ACRONYMS	x
LIST of FIGURES	xi
LIST of TABLES	xiii
1. INTRODUCTION	1
1.1. ETS Transcription Factor Family	1
1.2. Evolutionary and developmental role of ETS Family	2
1.3. Glioma progression and ETS proteins	4
1.4. Purpose of Thesis	6
2. MATERIALS and METHODS	7
2.1. Materials	7
2.1.1. Datasets	7
2.1.2 Cell lines	8
2.1.3. Kits and cell culture solutions	9
2.2. Methods	10
2.2.1. Raw data preprocessing and differential gene expression analysis of RNA-seq datasets	10
2.2.2. Preprocessing and differential gene expression analysis of micro array datasets	11
2.2.3. Functional enrichment analysis	11
2.2.4. Gene regulatory network constructions	11
2.2.5. Differentiation of SH-SY5Y cells with retinoic acid and qPCR validation for genes associated with neuronal differentiation	14
3. RESULTS	17
3.1. Glioma grade progression and transcriptomic gene regulation related with the ETS transcription factors	17

3.1.1. Identification of differentially expressed genes 17
3.1.2. Transcriptional gene regulatory networks of glioma grades 20
3.2. Transcriptomic comparison of neuronal differentiation across species under the regulation of ETS TFs 26
3.2.1. ETS subfamilies in different species 26
3.2.2. Transcriptomic comparison of neuronal differentiation and gene regulatory network of ETS TFs 28
3.3. Transcriptomic comparison of primary neurons and RA-differentiated SH-SY5Y cells 46
4. DISCUSSION and CONCLUSION 52
REFERENCES 56
BIOGRAPHY 64
APPENDICES 65

LIST of ABBREVIATIONS and ACRONYMS

Abbreviations and Acronyms		Explanations
ETS	:	E26 Transformation Specific
PNT	:	Pointed
FGF	:	Fibroblast Growth Factor
MAPK	:	Mitogen Activated Protein Kinase
SRF	:	Serum Response Factor
PEA3	:	Polyoma Enhancer Activator 3
SRE	:	Serum Response Element
SH-SY5Y	:	Human Neuroblastoma Cell Line
FBS	:	Fetal Bovine Serum
DMEM	:	Dulbecco's Modified Eagle Medium
PCA	:	Principal Component Analysis
DEGs	:	Differentially expressed genes
NSC	:	Neural stem cell
NPC	:	Neural progenitor cell
GRN	:	Gene regulatory network

LIST of FIGURES

Figure No:		Page
1.1:	Representation ETS family members and relative location of ETS and PNT domain.	1
2.1:	Methodological flowchart of glioma progression project.	13
2.2:	Methodological flowchart of neuronal differentiation project.	14
3.1:	Results of differential gene expression analyses.	18
3.2:	Functional enrichment analysis results of DEGs in glioma grades.	20
3.3:	ETS-mediated gene regulation network size and regulon activity.	21
3.4:	Individual regulon targets that have positive and negative regulation were identified by two-tailed GSEA analysis.	22
3.5:	Glioma grade-specific gene regulatory network controlled by ETS transcription factors.	23
3.6:	GO enrichment analysis of certain regulatory targets that are both positively and negatively regulated	25
3.7:	KEGG enrichment analysis of certain regulatory targets that are both positively and negatively regulated.	26
3.8:	Distribution of ETS subfamilies and members located in species in this study.	28
3.9:	Explanatory investigation of transcriptomic changes between stem or progenitor cells vs corresponding mature neurons from transcriptomic datasets	30
3.10:	GO over-representation analysis of the biological processes for significantly changed genes resulted from differentially gene expression analyses.	31
3.11:	Regulon activity of ETS members. Heatmaps represent the expression changes of significantly changed ETS members with normalized regulon activity scores.	34
3.12:	Gene regulatory network of human neuronal differentiation under the regulation of ETS transcription factors.	36
3.13:	GO enrichment analysis of human individual regulons and their positively and negatively regulated targets.	37

3.14:	Gene regulatory network of mouse neuronal differentiation under the regulation of ETS transcription factors.	38
3.15:	GO enrichment analysis of mouse individual regulons and their positively and negatively regulated targets.	39
3.16:	Gene regulatory network of Drosophila neuronal differentiation under the regulation of ETS transcription factors.	40
3.17:	GO enrichment analysis of Drosophila individual regulons and their positively and negatively regulated targets.	41
3.18:	Gene regulatory network of C. elegans neuronal differentiation under the regulation of ETS transcription factors.	42
3.19:	GO enrichment analysis of C. elegans individual regulons and their positively and negatively regulated targets.	43
3.20:	Gene regulatory network of gorilla neuronal differentiation under the regulation of ETS transcription factors.	44
3.21:	GO enrichment analysis of gorilla individual regulons and their positively and negatively regulated targets.	45
3.22:	Combined functional transcriptional network of ETS members.	46
3.23:	PCA plots of the SH-SY5Y RA dataset and Primary neuron dataset.	47
3.24:	Volcano plot for Primary neuron and SH-SY5Y data sets after differential gene expression analysis.	47
3.25:	Comparison of genes significantly changed as a result of differential gene expression analysis with venn diagram.	48
3.26:	GO analysis result of genes with reduced gene expression	48
3.27:	GO analysis result of genes with increased gene expression.	49
3.28:	Network representation of common genes in the two datasets and genes associated with neuronal differentiation by GO analysis.	50
3.29:	Morphological structures of SH-SY5Y cells during RA differentiation.	50
3.30:	qPCR validation of selected genes.	51

LIST of TABLES

Table No:		**Page**
2.1:	Descriptions of transcriptomic datasets.	8
2.2:	List of kits for qPCR validation assays.	9
2.3:	List of cell culture solutions.	9
2.4:	Primer lists	16
4.1:	Published and submitted research papers.	55

1. INTRODUCTION

1.1. ETS Transcription Factor Family

The ETS domain transcription factor superfamily, which consists of 13 subfamilies and 29 different members, originated from the viral oncogene v-ets-1 of the E-twenty-six (E26) leukemogenic virus and its protooncogene counterpart c-ets-1 [1]. Although the location of these domains on the protein might vary between subfamilies, all members of the ETS superfamily have the conserved DNA binding domain called the ETS domain, which binds a consensus GGAA/T pattern on its target promoters. There is also an extra helix-loop-helix domain found in some subfamilies known as the "pointed" domain, including the ETS and ERG subfamilies [2]. Despite these intra-family variances, it is thought that all ETS proteins are conserved among metazoan, evolving through gene duplications, and providing a substantial amount of functional redundancy during the development (Figure 1.1) [1],[2],[3].

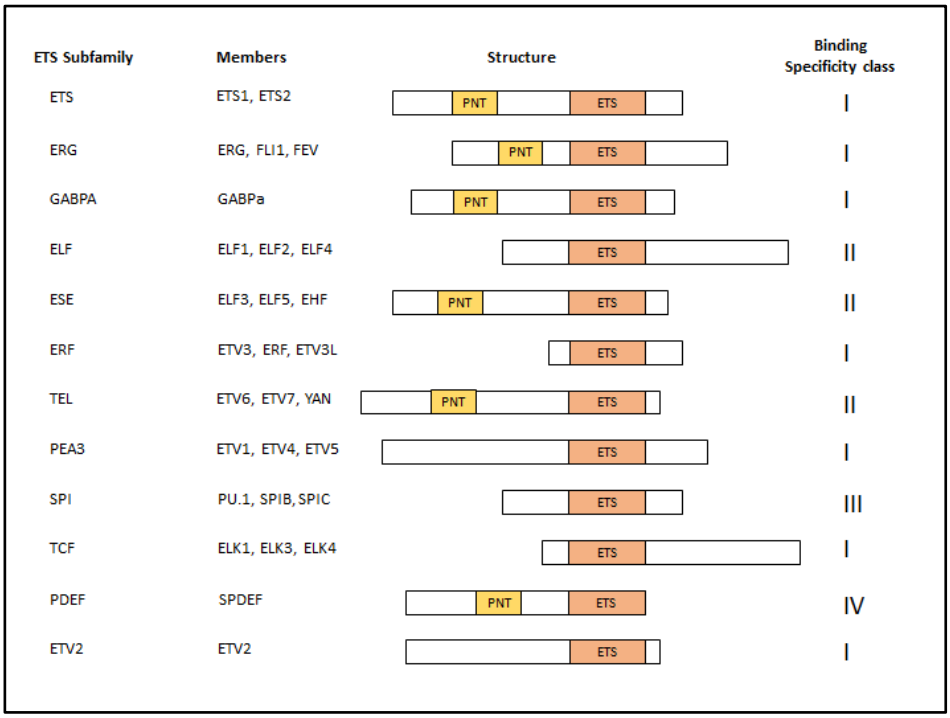


Figure 1.1: Representation ETS family members and relative location of ETS and PNT domain.

Several biological processes, including tissue morphogenesis, cell growth, death, and embryonic development, are influenced by ETS proteins. [2],[4]. It's interesting to note that the tissue expression patterns of ETS proteins display significant similarities, with several members of this superfamily exhibiting widespread expression [5]. Not unexpectedly, ETS members also generally show overlapping and possibly redundant DNA binding, as demonstrated by genomic sequence tenancy and other investigations. [6]. Regarding their functions in tumorigenesis, they have also been subjected to a lot of research. For instance, in Ewing's sarcoma, the ETS protein Fli-1 is fused to an RNA-binding protein EWS as a result of a chromosomal translocation, which increases the capacity for transactivation and results in malignancy. Similarly, chronic myelomonocytic leukemias show TEL/ETV6 translocation to the PDGF receptor b gene. Numerous ETS proteins have been identified to be overexpressed or to have mutations in various cancers [2],[7],[8].

1.2. Evolutionary and developmental role of ETS Family

The ETS family of transcription factors has a large protein distribution in the animal kingdom. Therefore, the evolutionary history of ETS proteins is as broad as their differential distribution of gene expression in tissues. The ETS, PEA3, GABP, and ER71 subfamilies seem to have split off more later in the phylogeny through gene duplications peculiar to the vertebrate lineage than the SPI subfamily in terms of evolutionary connections [1]. For instance, the TCF subfamily member Elk-1 was discovered to be conserved across the metazoan kingdom, from the earliest deuterostomes to vertebrates. Cnidarians and hemichordates have a single defined Elk1 gene, whereas vertebrates like the axolotl and humans have three Elk (TCF) subfamily members [9]. The annelid *Hediste diversicolor*'s Ets genes, Nd ets and Nd erg, have been found to be functional homologs of Ets-1 and Fli-1 in human [10].

Ets proteins have a variety of roles in embryonic development, including the hematological system, neurological system, and cardiovascular system development [11],[12]. From sponges and mollusks to humans, a considerable number of ETS proteins, including TEL, ETS1, Cf-ets, and Friend leukemia integration 1 transcription factor (Fli1), are crucial for the development of hematopoietic lineages and the immune system [13]. It has been demonstrated that the ETS proteins Ets1/2 and

Elk1/3/4 control FGF signaling in the anterior neural plate during the development of the nervous system in marine invertebrates like *Ciona intestinalis* [14]. It has been demonstrated that the UNC-6/netrin signaling interacts with the *Caenorhabditis elegans* ETS protein AST-1 (for axonal steering), which controls neuronal development and axon guidance [15]. Zebrafish with simultaneous Pea3, Etv5, and Erm knockdown displayed abnormalities resembling Fgf deficit in embryos throughout cardiac development and patterning [16]. C-ets1 was discovered to be expressed during rat development in the neural pituitary at E14 as well as in the future hypothalamus, despite the fact that its function in both areas was thought to be blood vessel construction [17]. Er81 knockout mice had postnatal lethality, loss of monosynaptic connections, and developmental issues with proprioceptive afferents [11].

Along with other organs, the eye and brain expressed Xl-Ets-2 during *Xenopus* embryogenesis, and XER81 was discovered in the cortical plate and later in discrete frontal cortex, midbrain, and hindbrain subdomains. In contrast, Ck-Ets-1 was expressed in neural crest cells in chicken embryogenesis, and the Ck-Pea3 subfamily was found in motor neurons in the embryonic spinal cord [4],[18]. The ER81 protein was discovered to be expressed in the developing hindbrain, notably in the inferior olive nucleus, in both chicks and mice [19]. It has been discovered that longitudinal and midline glial cell differentiation, as well as photoreceptor differentiation in the eye, are crucial for the development of the central nervous system (CNS) in *Drosophila* pointedP1 and pointedP2, as well as their vertebrate homologs c-ets1 and c-ets2 [20].

Through controlling the expression of acetylcholine receptors on postsynaptic membrane, conditional deletion of the ETS gene GABPα in skeletal muscle has revealed its importance at the neuromuscular junction [21]. It was discovered that the Pea3 subfamily's mesenchymal Etv1, Etv4, and Etv5 were crucial for cochlear length but were not involved in the patterning of the sensory progenitors [22]. Etv5 was also demonstrated to be a MAPK target during the development of the ventral telencephalon, notably in the progenitors of the ventricular zone [23]. In the early stages of zebrafish brain development, it was discovered that Fgf3 and Fgf8 signaling controlled the expression of Erm and Pea3 [24].

Interesting evidence suggests that many ETS proteins are involved in serotoninergic differentiation. In zebrafish, it was discovered that Erm/Etv5b controls the activation of hypothalamic serotoninergic neuron development in response to Fgf

signaling [25]. In serotonergic neurons of the raphe nuclei in the midbrain and pons of humans, the ETS protein FEV was measured [26]. On the other side, serotoninergic development was flawed in mouse Pet-1 knockout mice [27].

1.3. Glioma progression and ETS proteins

Due to their localization and sometimes locally invasive development, brain tumors are marked by severe morbidity and death. The majority of neoplastic brain lesions are metastases from non-central nervous system malignancies, which are 5–10 times more prevalent than original brain tumors. The most prevalent primary brain tumors are gliomas. The bulk of primary brain tumor fatalities are caused by gliomas, which make up almost 30% of all primary brain tumors and 80% of all malignant ones [28].

ETS proteins have critical roles in both the formation and progression of cancer as well as normal developmental processes and growth. They do this through regulating epithelial-stromal interactions, the endothelial transition, as well as cell growth, migration, or angiogenesis. [29]. Numerous ETS TFs, particularly PEA3, ETS1, and ETS2, have enhanced activity in malignancies. Among other aspects of tumorigenesis, these ETS TFs are involved in the development of tumors, the mesenchymal transformation, metastasis, and angiogenesis. [30],[31]. When genetic modifications in PPI, PTM, and/or protein stabilization arise, it frequently leads to an increase in the transcriptional potency of ETS TFs in a variety of malignancies, as is the case with c-ETS1 in leukemia. Other times, ETS TFs undergo chromosomal relocations that result in TMPRSS-ERG fusion proteins in much over half of prostate cancers. [30].

Breast cancer was discovered to have excessive levels of Ets-1, which was linked to a bad prognosis [31]. Ets-1 controls energy metabolism in several cancer types such as breast and ovary cancer in addition to being a crucial invader regulator [31],[32]. Through the control of hTERT expression, it has also been demonstrated to contribute to the preservation of telomeres [33]. The ternary complex factor (TCF) subfamily member Elk-1, which is phosphorylated at serine 417, has been linked to the degree of differentiation of colonic adenocarcinomas [34]. Furthermore, ELK-1 was discovered to be related with disease recurrence in later stages of prostate cancer [35]. In several

cancer types such as brain, sarcomas and skin, ELK1 expression has been shown to be greater than ELK3 expression [36]. The MZF/Elk-1 transcription factor complex boosted PKC expression in hepatocellular carcinoma in conjunction with cell migration and tumorigenicity [37]. A variety of malignancies, including lung tumors with MET amplification, were also affected by PEA3, ERM and ER81 and they were discovered to be involved in migration [38]. PEA3 has been demonstrated to support malignancy in colorectal cancer [39]. It has been demonstrated that the overabundance of ETV4 and ETV5 in ovarian cancer is caused by the loss of PEA3 subfamily repressors [40].

The most prevalent and fatal primary brain tumors are malignant gliomas. Characterization of glioma grades is based on the cell proliferation activity, angiogenesis and biomarkers, which are also used as diagnostic criteria. [41],[42]. The most aggressive and deadly glioma is glioblastoma multiforme (GBM), which is categorized as a WHO grade IV disease. A fast-proliferating tumor core and highly invasive cells surround the tumor core in high-grade gliomas [43].

In high-grade gliomas, ETV2, a key player in the vascular reprogramming of CD133+ cells, was shown to be overexpressed. This transdifferentiation is expected to make CD133+ GBM stem cells resistant to anti-angiogenic treatment [44]. It was discovered that the gene ERG is a biomarker for vascular cells in CNS malignancies. This finding may be utilized to analyze the vascularization of gliomas [45]. PU.1, identified with a crucial function in blood cell development, had also been found to be substantially expressed in glioma patients, suggesting a role in glioma progression. [46]. Fli1 has also been shown to be altered expression in glioma as same as PU.1 [47]. ETS proteins have the ability to control other facets of carcinogenesis in addition to their function in tumor vascularization. Etv5 was shown to be a regulator in Nf1 mutant mice and studies confirmed the elevated expression of Etv5 and its target genes in glioma [48].

Diffuse and non-diffuse gliomas are the two broad categories for gliomas. The WHO categorization of CNS malignancies is used to identify and classify diffuse gliomas, such as oligodendrogliomas and astrocytomas, which resemble glial precursors [49]. The consistency in diagnosis of II and III grades of gliomas might be difficult due to their somewhat variable character. Differential diagnosis is performed by a number of genetic markers, such as IDH mutations or TERT alterations, which provide ETS binding sites [33],[41]. ERG, PEA3, ERM and ER81 over expressed

patients were identified in association with poor prognosis in clinical prostate tumor samples, showing that ETS gene expression could be used as a biomarker [50].

1.4. Purpose of Thesis

In this thesis study, I have mainly focused on ETS superfamily proteins during nervous system development and the functional evolution across species. For the first time, I present a comprehensive analysis of the regulation of the ETS transcription factor throughout the development of the nervous system in the metazoan lineage. I have shown that ETS members affect relatively similar genes that participate in the development of the nervous system, including genes involved in synaptic and axonal regulation and neuronal differentiation. Additionally, in order the investigate the transcriptional regulation of ETS proteins, I have also studied glioma progression as an aspect of constantly proliferative cell condition. In this part of the thesis work, I mainly focused on the different ETS genes that are upregulated in glioma grades found in transcriptomic data from patients. The general purpose of this thesis is to provide a comprehensive view of gene regulation under the ETS transcription factor family in terms of nervous system development and glioma progression with an evolutionary perspective using systems biology approach.

2. MATERIALS and METHODS

2.1. Materials

2.1.1. Datasets

All transcriptomic datasets were obtained from Gene Expression Omnibus database [51]. For glioma grade progression project, patient-derived micro-array datasets with control tissue samples were used including different stages of glioma (grade 2-4) [52]. For transcriptomic comparison of neuronal differentiation, transcriptomic datasets from 5 different organisms (human, gorilla, mouse, fruit fly and *C. elegans*) including neural stem cells (NSCs), neural progenitor cells (NPCs) and their mature neuron counterparts were used. Additionally, transcriptomic dataset from SH-SY5Y cells with retinoic acid (RA) differentiation procedure was used to validate the SH-SY5Y cell line as an *in vitro* model for neuronal differentiation [53]. Detailed information about all transcriptomic datasets in this study is listed in Table 2.1.

An 18-day neuronal differentiation strategy was used to develop H1-derived NPCs to mature neurons in the human dataset (GEO accession GSE56785). Data from day 0 (NPC) and day 18 (neuron) were used to distinguish between NPC and mature neurons. All additional samples (days 1, 2, 4, 5, and 11) were used to build a gene regulatory network. The adult mouse brain dataset (GEO accession GSE94991) comprised several phases of neurogenic differentiation from neural stem cells to neurons (NSCs, early and late neuroblasts, neurons). NSC and mature neuron samples were used for differential gene expression analysis, just like the human dataset, and all samples were used for gene regulatory network construction. The larval neuroblasts and their derived daughter neurons from the *Drosophila* dataset (GEO accession GSE38764) were used. The pan-neuronal differentiation protocol's NPC and neuron samples were used in the gorilla dataset (GEO accession GSE120271). Finally, the pan-neuronal cells and their embryonic counterparts from the *C. elegans* dataset (GEO accession GSE8004) were used.

Table 2.1: Descriptions of transcriptomic datasets.

GEO accession	Organism	Title	Reference
GSE4290	*Homo sapiens*	Expression data of glioma samples from Henry Ford Hospital	[52]
GSE56785	*Homo sapiens*	RNA-Seq characterization of human H1-derived NPC differentiation time course	[54]
GSE71817	*Homo sapiens*	Neuro-like differentiation of SH-SY5Y human cell line	[53]
GSE94991	*Mus musculus*	Dynamic control of translation during adult neurogenesis	[55]
GSE38764	*Drosophila melanogaster*	Transcriptome analysis of *Drosophila* neural stem cells reveals a transcriptional network for self-renewal.	[56]
GSE120271	*Gorilla gorilla*	RNA-seq and ChIP-seq for human-specific regulation of neural maturation identified by cross-primate transcriptomics	[57]
GSE8004	*Caenorhabditis elegans*	Cell-specific microarray profiling of the *C. elegans* nervous system.	[58]

2.1.2 Cell lines

In cell culture studies, SH-SY5Y human neuroblastoma cells (ATCC number: CRL-2266) were used in all experimental procedures.

2.1.3. Kits and cell culture solutions

Table 2.2: List of kits for qPCR validation assays.

Kit	Company	Catalog/Lot Number
innuPREP RNA Mini Kit 2.0	AnalytikJena	035-19
iScript cDNA Synthesis Kit	Bio Rad	1708891
SsoAdvanced Universal SYBR Green Supermix	Bio Rad	172-5274

Table 2.3: List of cell culture solutions.

Solutions	Company	Catalog/Lot Number
Dulbecco's Modified Eagle Medium 1X (DMEM) with L-glutamine, pyruvate and 4,5g/L Glucose	Gibco	41966-029
Fetal Bovine Serum (FBS)	Gibco	10270
Penicillin,Streptomycin (10000U-ml Penicillin, 10mg/ml Streptomycin)	PAN Biotech	P06-07100
Dulbecco's Phosphate Buffered Saline 1X (DPBS)	Gibco	14190-94
0.05% Tyrpsin-EDTA Solution 1X	Gibco	25300-054

2.2. Methods

2.2.1. Raw data preprocessing and differential gene expression analysis of RNA-seq datasets

Using the SRA Run Selector, each RNA-seq raw sequence files was uploaded to the Galaxy online platform for preprocessing [59]. For quality control of the raw and trimmed sequence reads, FastQC was used with default settings [60]. Then, using Trim Galore! raw reads were trimmed to remove sequence adapters and poor-quality reads (26 Phred quality score cutoff) [61]. Following trimming, reads were aligned with the reference genome (for *H. sapiens*, hg19, built-in Galaxy; for *Mus musculus*, mm10; for *D. melanogaster*, dm3; for *G. gorilla*, gorGor6, from UCSC genomes) to produce a mapped read BAM file using HISAT2 [62]. Finally, the featureCounts tool was used to produce a count matrix [63].

DESeq2 were used for differential gene expression analyses [64]. Prior to doing a differential gene expression analysis, all sum of raw read counts for each genes with lower-than-10 summed row count values across samples were removed. The built-in vst function of DESeq2 was used to apply a variance stabilizing transformation on the count matrix in order to eliminate the variance's dependency on the mean, to handle the large variance of the logarithm of count data when the mean is low. The plotPCA function was used to perform principal component analysis (PCA), after which PC1 and PC2 were chosen as the primary principal components for display. Although normalized data was used for PCA and all other down-stream analysis, raw count matrix for mature neurons and their associated stem cell or progenitor cell in each dataset were used for differential gene expression analysis to compare the transcriptomic profiles. Differentially expressed genes (DEGs) were identified using the DESeq function and displayed using the EnhancedVolcano package with an adjusted p-value 0.05 or 0.01 using Benjamini-Hochberg correction [65].

2.2.2. Preprocessing and differential gene expression analysis of micro array datasets

The dataset's processed expression matrix for datasets which include *C. elegans*, glioma patients and differentiated SH-SY5Y with RA were retrieved from Bioconductor using the R package GEOquery [66]. To examine the association between the samples, the expression matrices from each dataset were Log2 normalized and employed in PCA analysis. Prior to doing the differential analysis of gene expression, PCA plot was used to determine outliers. pan-neuronal cells and their embryonic reference samples (*C. elegans* dataset) and each glioma grades and control brain tissue samples (glioma dataset) were compared using the R package limma from Bioconductor to analyze differential gene expression [67]. Volcano plots were created using the R package EnhancedVolcano from Bioconductor to display the DEGs that were found using Benjamini-corrected Hochberg's adjusted p-value 0.05 or 0.01 significant threshold.

2.2.3. Functional enrichment analysis

Gene Ontology (GO) and KEGG over-representation analyses were carried out with an adjusted p-value 0.05 Benjamini-Hochberg correction cutoff by clusterProfiler using up-and down-regulated genes produced by differentially expressed genes as well as target genes produced by gene regulatory network algorithm [68]. Significantly enriched terms were displayed as bar graphs after enrichment analyses.

2.2.4. Gene regulatory network constructions

RTN R package was used to build gene regulatory networks [69]. Using a gene expression matrix and a predetermined TF list, this method calculates the TF-gene interaction network. RTN method uses transcriptomic data from RNA-seq or microarray to examine the mutual information between TFs and possible targets. Additionally, it offers capabilities that use bootstrap analysis and steps in data processing to filter out interactions that are weak. In order to assess both the positive and negative gene target sets of TFs regarded as regulons, gene set enrichment analysis

(GSEA) was also incorporated into RTN algorithms. Edge-weighted TF-gene interactions are present in constructed gene regulatory networks, and all of the networks derived from various species were employed for comparative study. In glioma grade project, gene regulatory network was used identify ETS genes and corresponding target genes associated with glioma grade progression. The RTN algorithm was utilized in the algorithmic setup with 1000 nPermutations, p-value 0.05 parameters, normalized gene expression matrices, and significantly altered ETS members as input. Cytoscape was used to display the gene regulatory networks that were retrieved from RTN [70]. The significant targets of ETS members in each dataset's gene regulatory network were also used in GO over-representation analyses. In the glioma patient's dataset, the generated network was further refined using DEGs from our earlier microarray experiments [8],[71].

Using gene regulatory networks, evolutionary distance, and the findings of regulon-specific GO enrichment analysis, a combined gene regulatory network was built in order to compare the properties of gene regulatory networks across species. The human, gorilla, mouse, fruit fly, and worm evolutionary dendrograms were taken from the TimeTree database [72]. The dendrogram was then integrated with all gene regulatory networks in graph style. GO terms were categorized to show the integrated network targets of ETS members as identified by the gene regulatory network analysis.

This method involved combining all GO results and creating a word cloud out of the most prevalent keyword phrases. All target genes were grouped into the appropriate GO keyword cluster using this word cloud, including development, axon, synaptic, differentiation, neuron, metabolic, filament, cellular, and cell cycle. Cytoscape was used to display the merged and classified network in its final form. Methodological flowcharts are shown in Figure 2.1-2.2

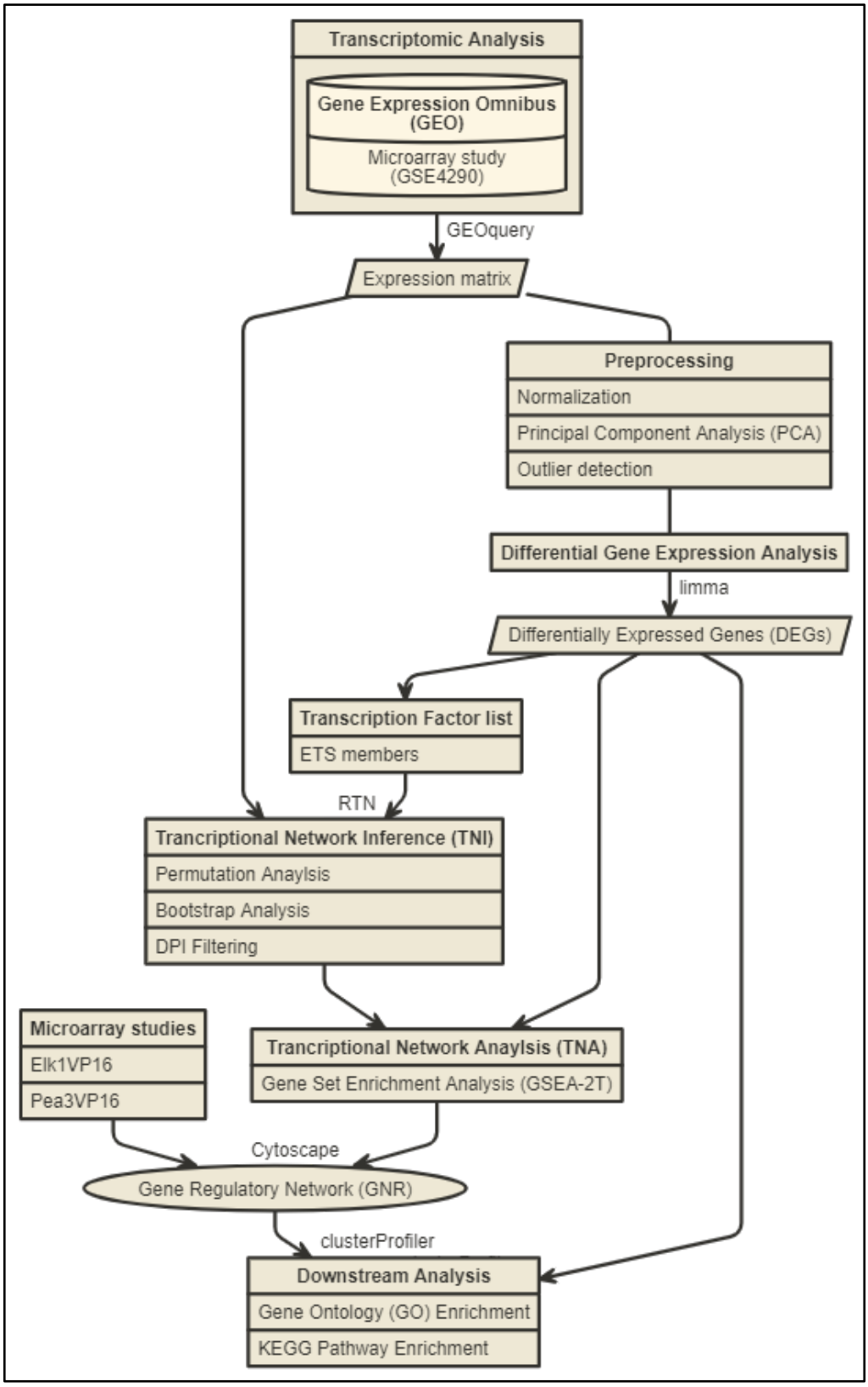


Figure 2.1: Methodological flowchart of glioma progression project.

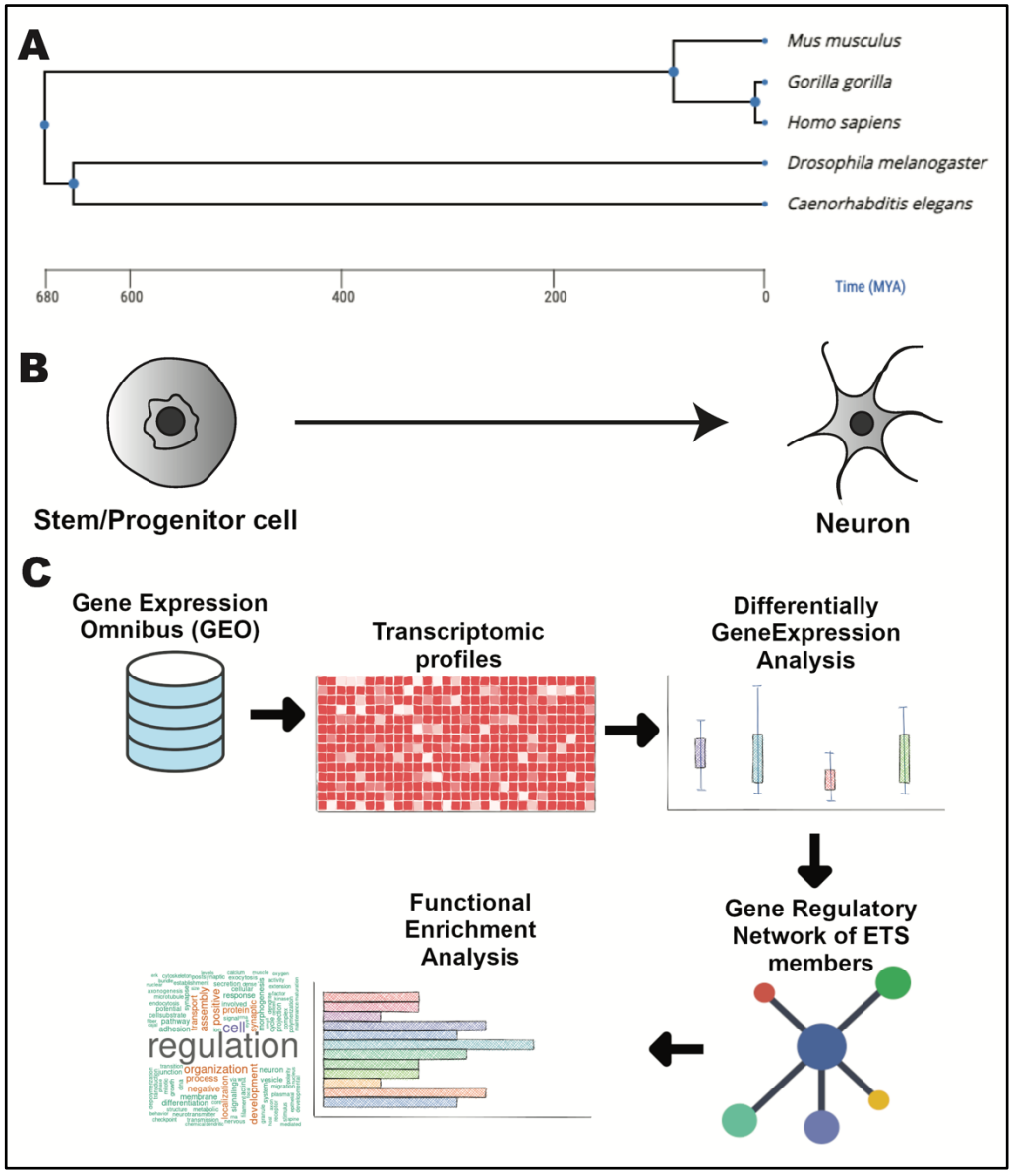


Figure 2.2: Methodological flowchart of neuronal differentiation project.

2.2.5. Differentiation of SH-SY5Y cells with retinoic acid and qPCR validation for genes associated with neuronal differentiation

SH-SH5Y cells were grown in standard culture medium DMEM with 4,5g/L glucose, L-glutamine, pyruvate was used with addition of 10 % FBS and 1 % penicillin/streptomycin antibiotics. When cells reach 80 % confluency in flask or plate, the cells are sub-cultured and were seeded onto 100 mm cell culture plate. For differentiation of SH-SY5Y cell growth medium DMEM with 10% FBS (4,5g/L glucose, L-glutamine, pyruvate, 10 % FBS and 1 % penicillin/streptomycin

antibiotics) containing SH-SY5Y cells to be differentiated was replaced with medium containing 1% FBS (4,5g/L glucose, L-glutamine, pyruvate, 1 % FBS and 1 % penicillin/streptomycin antibiotics) after 36 hours seeding the cells. Cells were incubated in medium containing 1% FBS for 24 hours, after which this medium was replaced with medium containing 1% FBS + 10 µM Retinoic acid. After the addition of retinoic acid, the medium was replaced with 1% FBS + Retinoic acid every 2 days. Cells were morphologically observed to differentiate for 2 days after the addition of retinoic acid.

After the RA differentiation protocol, cells within flask or plate are washed with PBS to eliminate dead, unbound cells and their debris. After that cells were collected into tube by cell scraper, then RNA extraction (innuPREP RNA Mini Kit 2.0) and cDNA synthesis (iScript cDNA Synthesis Kit) were performed for SYBER Green qPCR (SsoAdvanced Universal SYBR Green Supermix) protocol. For RNA isolation, first, the medium of cultured cells was discarded and 200 µl of PBS was added to the plate to collect the cell into the tube with a scraper. The tubes were centrifuged at 2000g for 5 minutes and the supernatant is discarded. Then 600 µl of lysis buffer containing 6µl 2-mercaptoethanol was added for each sample and the tubes were vortexed well. The mixture was then passed through the syringe 10-13 times for homogenization. After this step, 600 µl ethanol was added to each sample and the tubes were vortexed. Then 700µl of sample in the tubes were transferred to the column and centrifuged at 12000g for 15 seconds. This step was repeated until no sample remains. Then 700 µl of First Wash Buffer was added to the column and the columns are centrifuged at 12000g for 15 seconds and the supernatant is discarded. Then 500 µl of Second Wash Buffer was added to the column and the columns were centrifuged at 12000g for 15 seconds. Then the supernatant was discarded and repeated the Second Wash Buffer step. To remove ethanol, the columns were centrifuged at 12000g for 1-2 minutes without adding any solution (dry centrifugation) and the columns were transferred to clean collection tubes. For elution, 30 µl RNase-free water is added to the column. Columns were incubated for 1 min and columns are centrifuged at maximum speed for 2 min and RNA samples are stored at -80 °C.

cDNA synthesis was performed according to the kit protocol (kit): 20 µl total volume for 1x reaction, 4 µl of reaction mixture, 1 µl of reaction enzyme, 500 ng/µl of RNA sample and the required amount of nuclease-free water. Reaction protocol was

performed in thermal cycle according to following incubation time and temperature, respectively:

- 5 minutes at 25 °C
- 30 minutes at 42 °C
- 5 minutes at 85 °C

The resulting cDNAs were stored at -20 °C for short-term storage and at -80 °C for long-term storage. From cell culture to production of condition specific cDNAs, all experiment were carried out with assistance of my colleagues Ekin SÖNMEZ, Pınar MUTLU and Atılay İLGÜN.

The corresponding primers for qPCR analysis of the selected genes associated with neuronal phenotype and resulted in Figure 3.28 were designed with the help of NCBI Primer-BLAST (Table 2.4). For qRT-PCR reaction, 0.5 ng of synthesized cDNAs, 300 nM of gene-specific primers and 5ul SYBR Green supermix were used. Primers specific for β-actin and *GAPDH* genes will be used as reference genes along with primers designed for target genes. The qRT-PCR reaction will be performed on an Applied Biosystems StepOnePlus device in accordance with the SYBR Green protocol and the gene expression values were determined relatively between differentiated and undifferentiated sample groups by calculating the Ct values and $2^{-\Delta\Delta Ct}$ values.

Table 2.4: Primer lists

Gene	F-Primer (5'-3')	R-Primer (5'-3')
hACTB	ACGAAACTACCTTCAACTCC	GATCTTGATCTTCATTGTGCTGG
hGAPDH	CATCTTCCAGGAGCGAGATCC	AAATGAGCCCCAGCCTTCTCC
hSEMA3D	ACTCGATCCCTTGGGCCTAC	TTTGCTCCATTGAGCCAGTAGT
hSEMA5A	GGAGGAGAGCCTGAGCATGA	CACGGAGACCACACACCAAA
hETV5	TTTGATCTTGGTTGGAGGTGGGG	CTGATGATGAACAGTTTGTCCCAG
ROBO2	CCACATTTCTCAGGAGGCCA	ACAAACTGTGGGGGAGCAAC
PLXNB1	GAGTACCGTCCCCTGACCTT	TCAACATCAAGGGTGCGAGG
PLXNA4	CGGATCATCACAGAGAATGGGGA	GATCTTGTCCACTCGCAGCAG

3. RESULTS

3.1. Glioma grade progression and transcriptomic gene regulation related with the ETS transcription factors

3.1.1. Identification of differentially expressed genes

Transcriptomic datasets including to various glioma stages of patients with corresponding epilepsy patients as non-tumor sample were gathered and examined, before assessing the expression of genes in the ETS superfamily. These glioma samples' transcriptome analyses produced a collection of differentially expressed genes (DEGs), and principal component analysis was used to confirm that these DEGs were effective in differentiating tumor cells from healthy cells (Figure 3.1). The PCA analysis revealed that while non-tumor and glioma samples could be easily distinguished from one another, discrimination between glioma grades was less prominent; grade 2 and 4 gliomas were comparatively distinct, but grade 3 tumors were not easily distinguished from the other stages (Figure 3.1.A). As a result, any outlier data were removed before the analysis of differential gene expression. In the tumor samples, differential gene expression analysis revealed that 8402 of 19,225 genes had altered substantially (adjusted p-value 0.05 and $-0.6 > \log2FC > 0.6$ criteria) (grades 2–4). Differential gene expression analysis was applied to identify upregulated and downregulated genes in each of all classes of gliomas, and the DEGs were visualized as volcano plot (Figure 3.1.B–D).

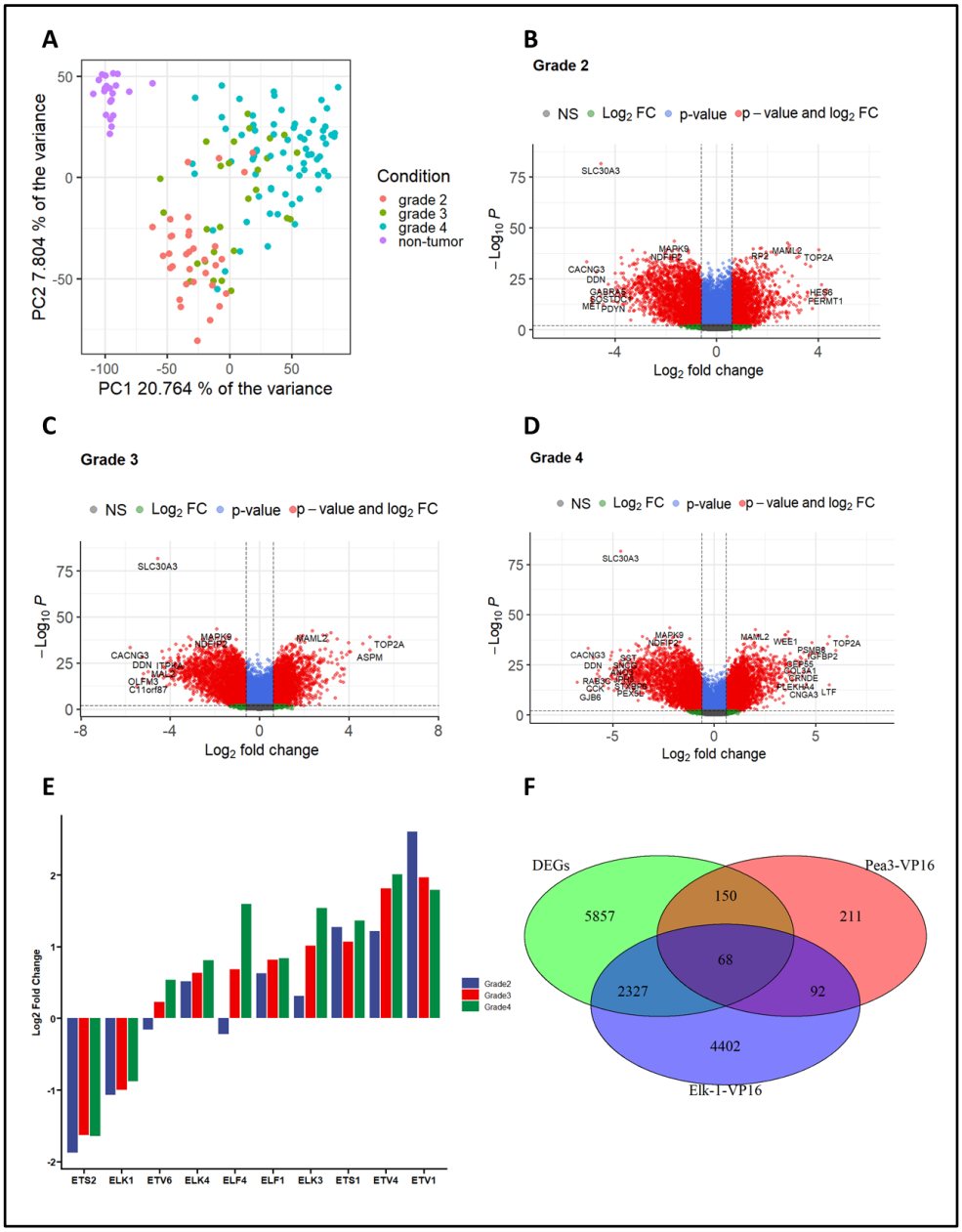


Figure 3.1: Results of differential gene expression analyses. (A). The normalized expression matrix's PCA graphic shows the principal components. Each dot denotes a distinct sample. (B–D). The associated differentially expressed genes (DEGs) were identified by applying the limma software with a 0.6 absolute log2 fold change, an adjusted p-value 0.01 with FDR cutoff, and comparisons of non-tumor vs. particular glioma grades. (E). Comparing samples with and without tumors using differential gene expression analysis to determine the relative FC of substantially altered ETS TFs for gliomas. (F). Venn diagrams were used to show the intersection of DEGs between our previously reported micro arrays Elk1VP16 and PEA3VP16 and limma findings.

The second question was, if ETS genes are DEGs, how do various glioma stages affect their expression during glioma progression? (Figure 3.1.E). For that reason, I have focused on ETS subfamilies, specifically the ETS [73],[74],[75], TCF

[8],[76],[77], ELF [78],[79], PEA3 [48],[80], and TEL [81],[82] subfamilies, which have all been previously documented to be implicated in gliomas. ELK-1 and ETS2 expressions were found to be downregulated in all grades, however ELK4, ELF1, ELK3, ETS1, ETV4, and ETV1 expressions were found to be upregulated in all grades, according to the data. Strangely, grade 2 had a downregulation of ELF4 and ETV6 expressions whereas grades 3 and 4 saw an upregulation. While ETV1 expression was found to be maximum in grade 2 gliomas and gradually decrease with glioma stage, the expression of ETV4, ELK3, and to a lesser degree ELK4 was shown to eventually increase with glioma grade. The PU.1, which has been demonstrated to be associated with the development of gliomas. and the expression levels were discovered to increase as glioma grades (Figure 3.1.E) [46].

In our earlier microarray investigations, a variety of transcriptional targets for these ETS proteins were discovered when constitutively active PEA3-VP16 or Elk-1-VP16 were overexpressed in SH-SY5Y neuroblastoma cells. I have compared DEGs from glioma grades with the transcriptomic dataset of our previous studies (PEA3-VP16 and Elk-1-VP16 overexpressing cells) in order to identify overlapping genes (Figure 3.1.F).

Enrichment analysis was carried out to elucidate the functional characteristics of the discovered DEGs in the glioma grades, and significantly enriched GO and KEGG terms were acquired (Figure 3.2). Up- and downregulated genes of the various glioma grades were originally examined for the GO analysis for biological processes. The downregulated gene clusters of grades demonstrated neuronal behavior, including synaptic transmission, as would be assumed, but the upregulated gene groups of grades were seen with cell cycle associated phenotype in all glioma grades (Figure 3.2.A). While the p53, TGF-β, and Notch signaling pathways were enriched in the KEGG analysis for the upregulated genes, substantial downregulated gene clusters fell into synaptic activity pathways, including such glutamatergic, GABAergic, serotonergic, and dopaminergic pathways (Figure 3.2.B). For instance, it was discovered that as gliomas progressed, the TGF-β signaling pathway was altered. This was confirmed by an increase in the concentration of BMP molecules, such as BMP3 and BMP4, as well as its targets, Smad1/5/8 and Id. Additionally, there was a dramatic increase in the expression of the cell cycle proteins p15 and cMyc. However, genes involved in the Notch signaling system, including Delta, Notch, and Fringe, were shown to be overexpressed in gliomas. The expression of genes including Ras, MEK1, ERK, JNK,

and Elk-1 was discovered to be downregulated, and the MAPK signaling pathway was revealed to be enriched in these clusters. All these enrichment analysis findings supported the idea that cells in all classes of glioma down-regulated signaling pathways directly related to neural function, but up-regulated signaling pathways generally associated with cell growth and survival. It's important to notice that the grade 4 glioma samples showed strong PI3K pathway overexpression rather than significant TGF-B pathway upregulation (Figure 3.2.B).

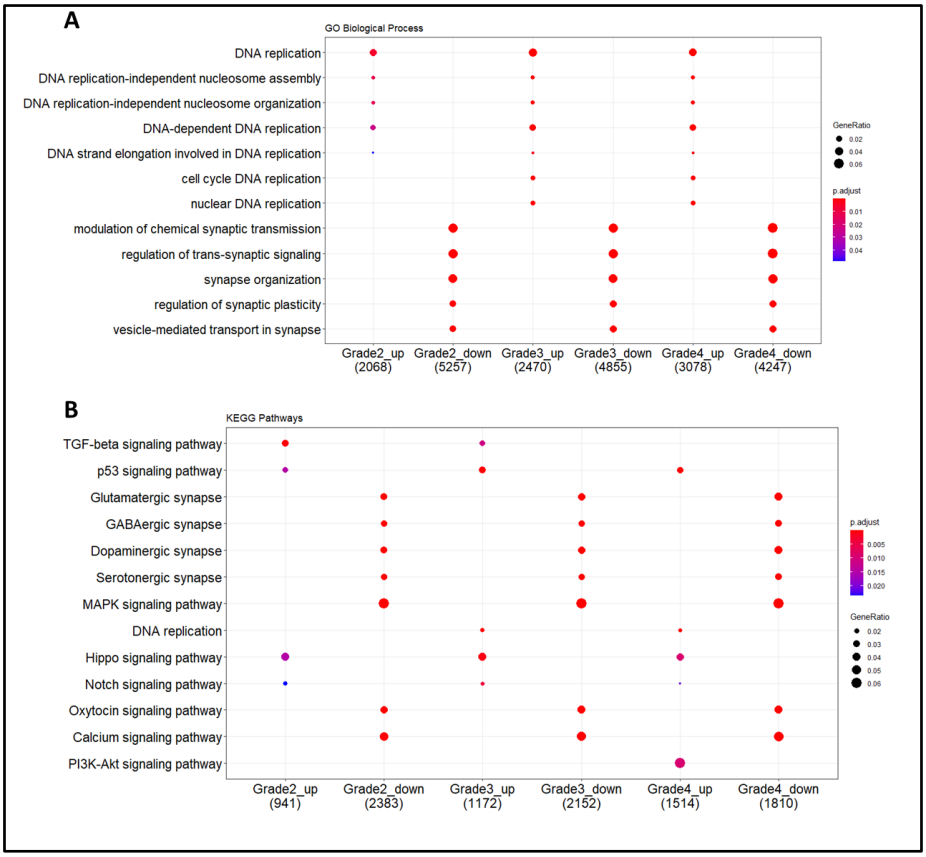


Figure 3.2: Functional enrichment analysis results of DEGs in glioma grades. adjusted p-value < 0.05 was used to identify significant (A) Gene Ontology and (B) KEGG terms.

3.1.2. Transcriptional gene regulatory networks of glioma grades

To explore ETS transcriptional regulatory networks specific to each glioma progression, I combined DEGs with the count matrix, ETS TFs as regulons. 10,762 target genes, 23,181 interactions overall, and 11 ETS member regulons were included in the first transcriptional network inference (TNI) network. When regulon activity

scores were analyzed, it was shown that while ELK-1 and ETS2 had strong regulon activity in the absence of a tumor (magenta, Figure 3.3.A), ETV1 displayed a high activity score on mostly grade 2 gliomas (green, Figure 3.3.A). While grade 3 glioma regulon activity was scattered between grade 2 and grade 4, the other ETS proteins had substantial regulon activity in the grade 4 cluster. The transcriptional network analysis (TNA) approach was utilized to include DEGs into the original network, which was followed by the construction of a targeted network of glioma grade. Nine important ETS regulons were discovered to be enriched with various target genes using the TNA algorithm (Figure 3.3.B).

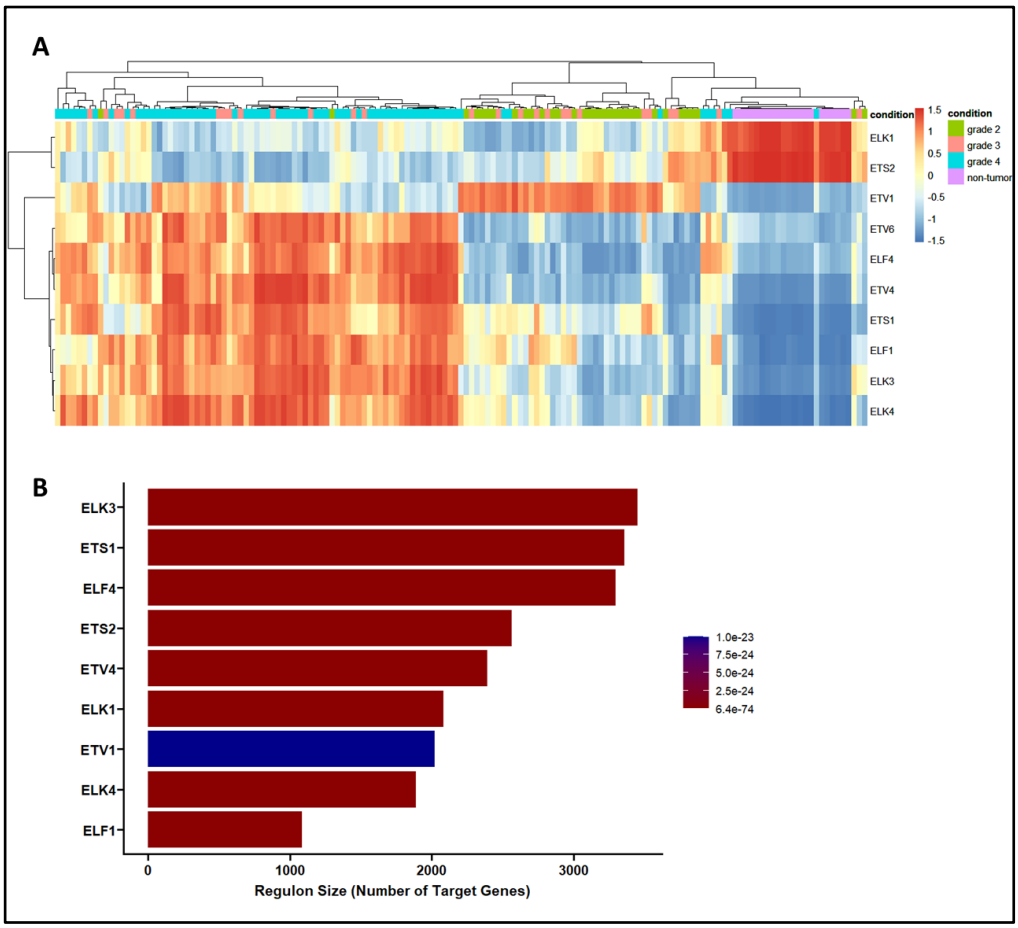


Figure 3.3: ETS-mediated gene regulation network size and regulon activity. (A). Regulon activity scores for ETS transcription factors across the glioma samples. (B). Regulon size of ETS TFs resulting after TNA algorithm.

Two-tailed GSEA was used to identify the directionality of regulations, and target gene distribution co-expression patterns for each unique ETS regulon were established for both positive and negative regulation (Figure 3.4). The peak of each

figure represents the enrichment score for the gene specified (dotted lines), and the colored bar displays the positively and negatively linked genes. In this study, genes were prioritized for their fold changes on the x-axis and enrichment scores on the y-axis. These findings imply that enriched ETS regulons have both distinct and common gene targets in gliomas, as shown by a distinct separation of negatively and positively linked targets in regulons like ETS2 and ELK1 (unique) and overlap between negatively and positively linked targets in negative and positive regulons like those of ELF4 and ELK3 (common).

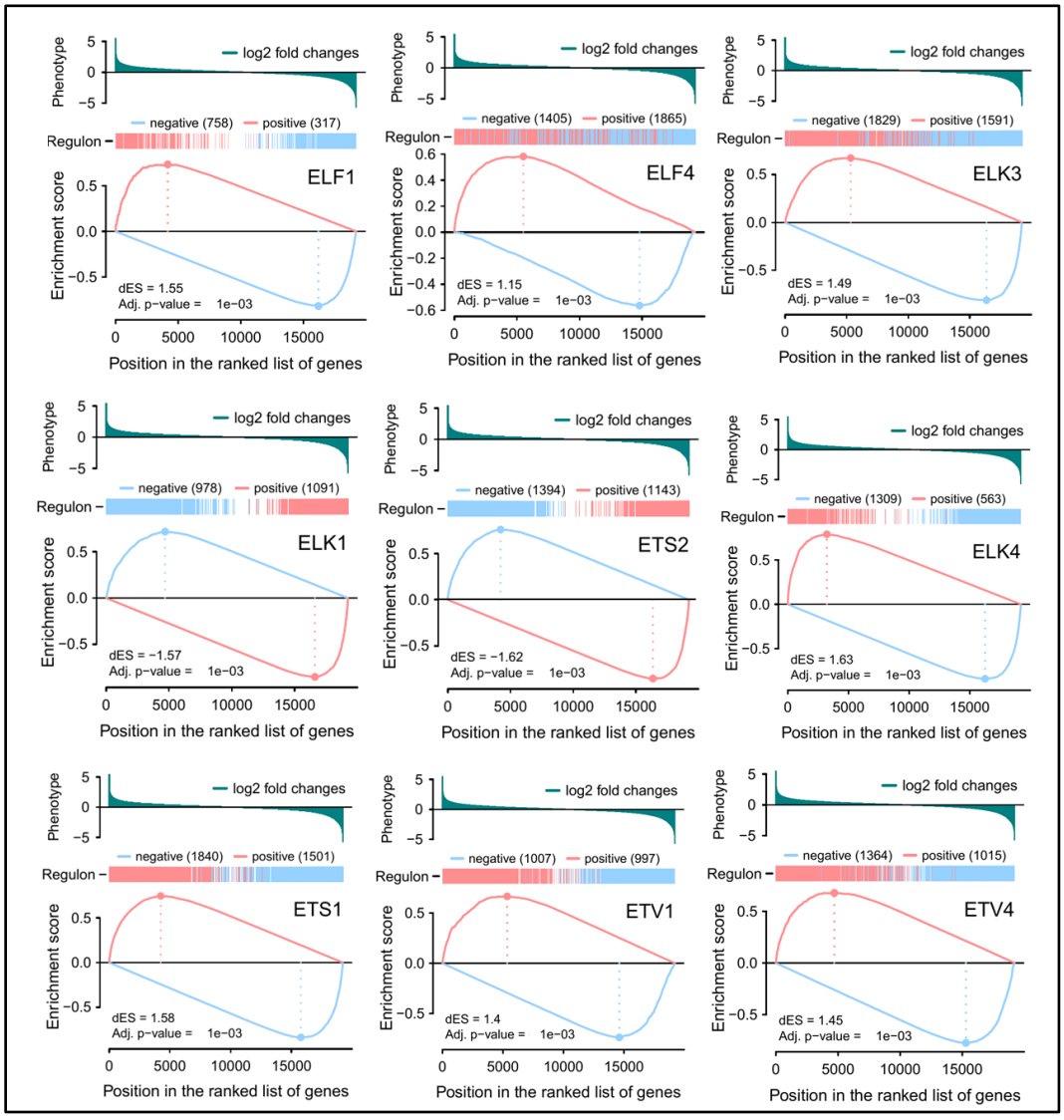


Figure 3.4: Individual regulon targets that have positive and negative regulation were identified by two-tailed GSEA analysis.

To produce a distinctive regulatory network of ETS TFs, the resulting network was filtered with DEGs from Elk1 and PEA3 overexpression transcriptomic data from our laboratory. Filtering led to the construction of a final regulatory network with 3366 target genes and 6610 connections with ETS regulons. Cytoscape was used to visualize the gene regulatory network (Figure 3.5). The common targets are clustered in the centre of this network representation, which also depicts FC of DEGs and their interactions with ETS regulons (Figure 3.5).

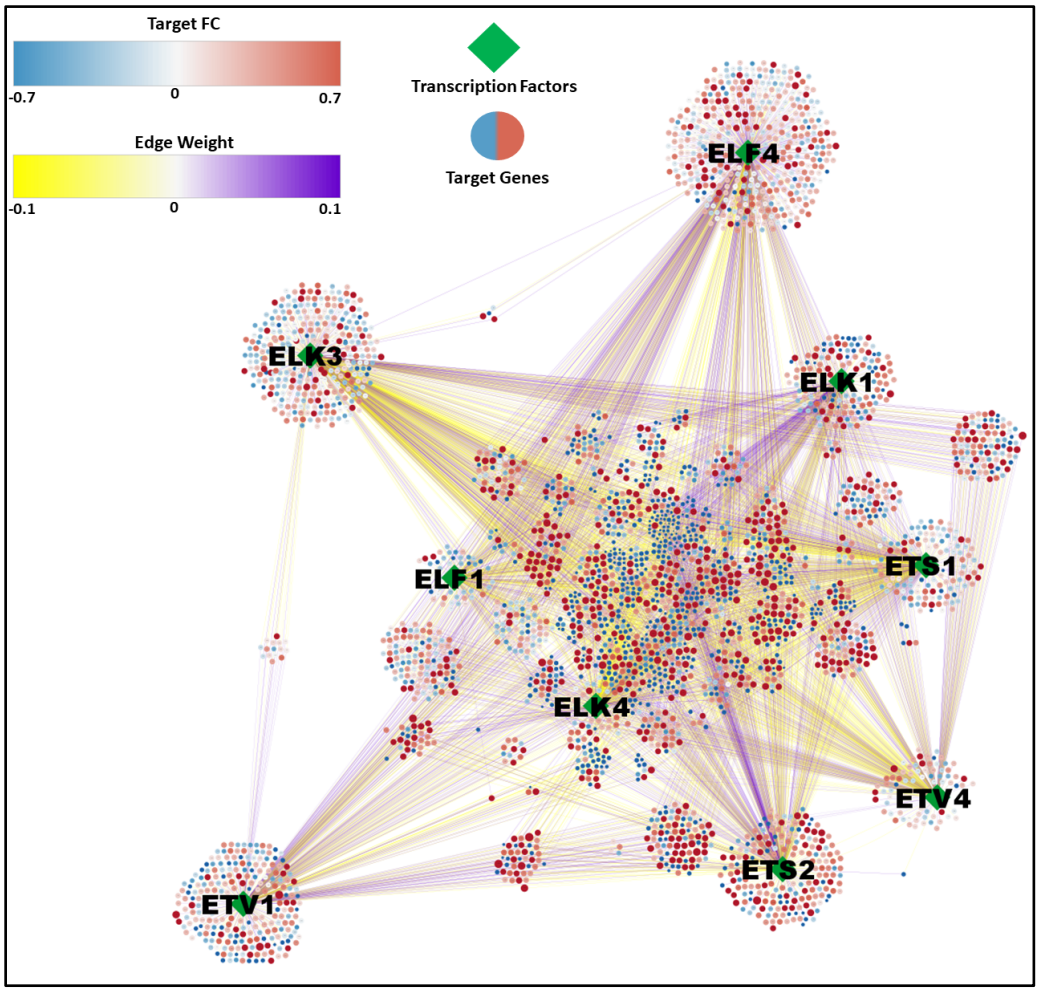


Figure 3.5: Glioma grade-specific gene regulatory network controlled by ETS transcription factors. Circle nodes stand in for the target genes, whereas diamond nodes indicate ETS TFs as a regulator. Due to differential gene expression analyses, the color of the circle represents the mean fold change of glioma grades in comparison to non-tumor samples. The GSEA analysis's edge colors display the enrichment scores for each target gene and its related regulators.

After that GO and KEGG enrichment analysis was carried out using positive and negative regulatory targets of ETS TFs on the network. It was shown that biological processes such as cell-cell contact, protein localization and synapse formation were enriched in the positive and negative regulatory targets of ETS TFs, some of which are shared by several family members and others of which are exclusive to one or a few (Figure 3.6). ETS2 and ELK1 were also shown to have targets inside the same biological pathways, while ELF4 and ELF1 regulons appear to correspond to related biological processes (Figure 3.6.A). While ETS1, ELK4, ELK and to some extent, ETV4 show to regulate analogous biological pathways, the ETV1 regulon shows to have a unique collection of positive regulatory targets (Figure 3.6.A). While ELF1, ELK4, ETS1, ELK3 and ETV4 appear to control comparable biological processes, the ELF4 regulon and the ETS2 regulon each consist of different targets, this categorization is not the same for negatively regulated targets (Figure 3.6.B). The nucleosome and chromatin disassembly-related processes were prominent in the positive regulatory targets of the ETV1 regulon, however no substantial negatively regulated targets for the ETV1 regulon were found. The ELK3 regulon had negatively regulated targets in synaptic transmission, synaptic vesicle signaling and synaptic plasticity pathways, as well as favorably regulated targets in protein maturation, ECM organization pathways.

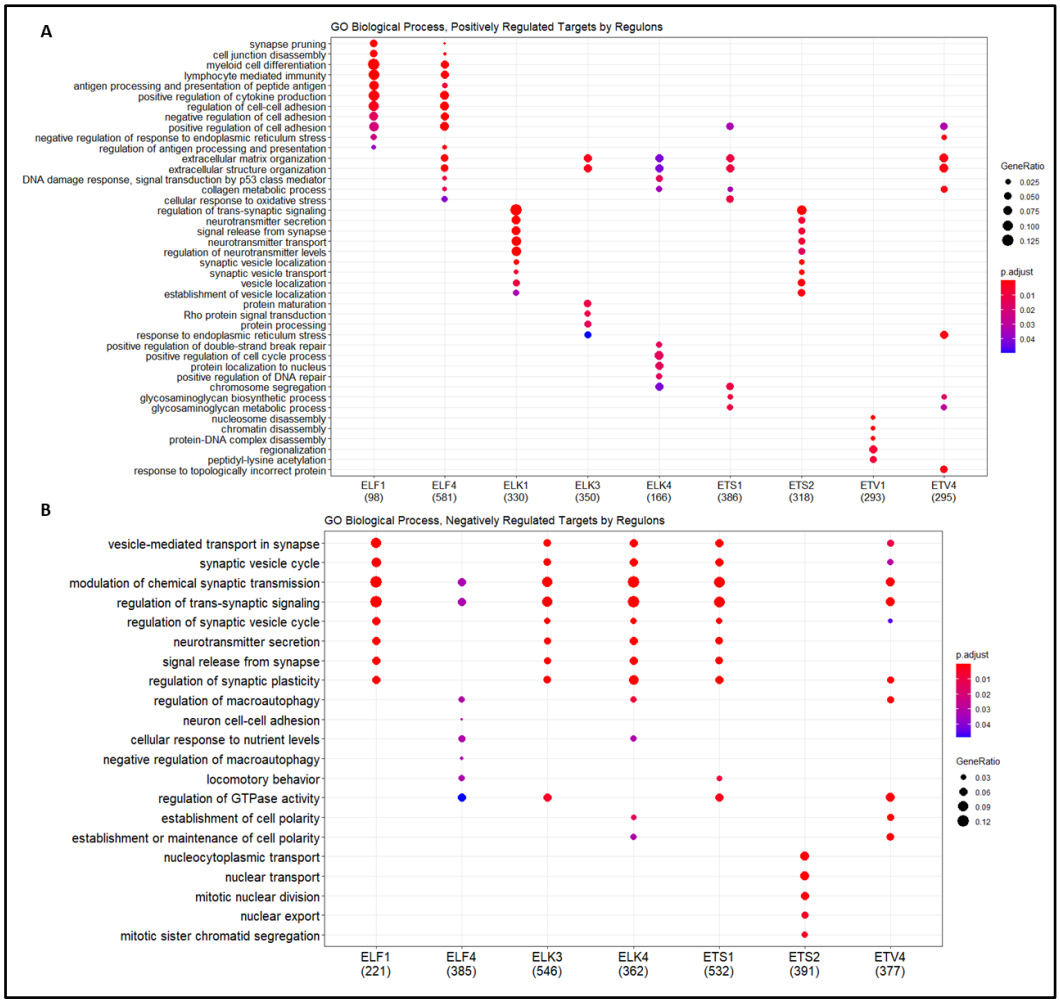


Figure 3.6: GO enrichment analysis of certain regulatory targets that are both positively and negatively regulated. With an adjusted p-value of 0.05, clusterProfiler carried out the GO analysis. The gene ratio shows the percentage of all the genes included in the enrichment analysis that are enriched with the associated GO terms.

KEGG pathway enrichment analyses revealed that, each ETS regulon controlled a different signaling pathway. For example, the ETV4 regulon's positively regulated cluster was enriched for the MAPK and PI3K-Akt signaling pathways, while its negatively regulated cluster was enriched for endocytosis and the synaptic vesicle cycle. The positive regulatory targets of ELK1 were enriched for calcium signaling pathway (Figure 3.7). Except for ELK1, most ETS regulons shared the synaptic vesicle cycle and endocytosis as common signaling pathways.

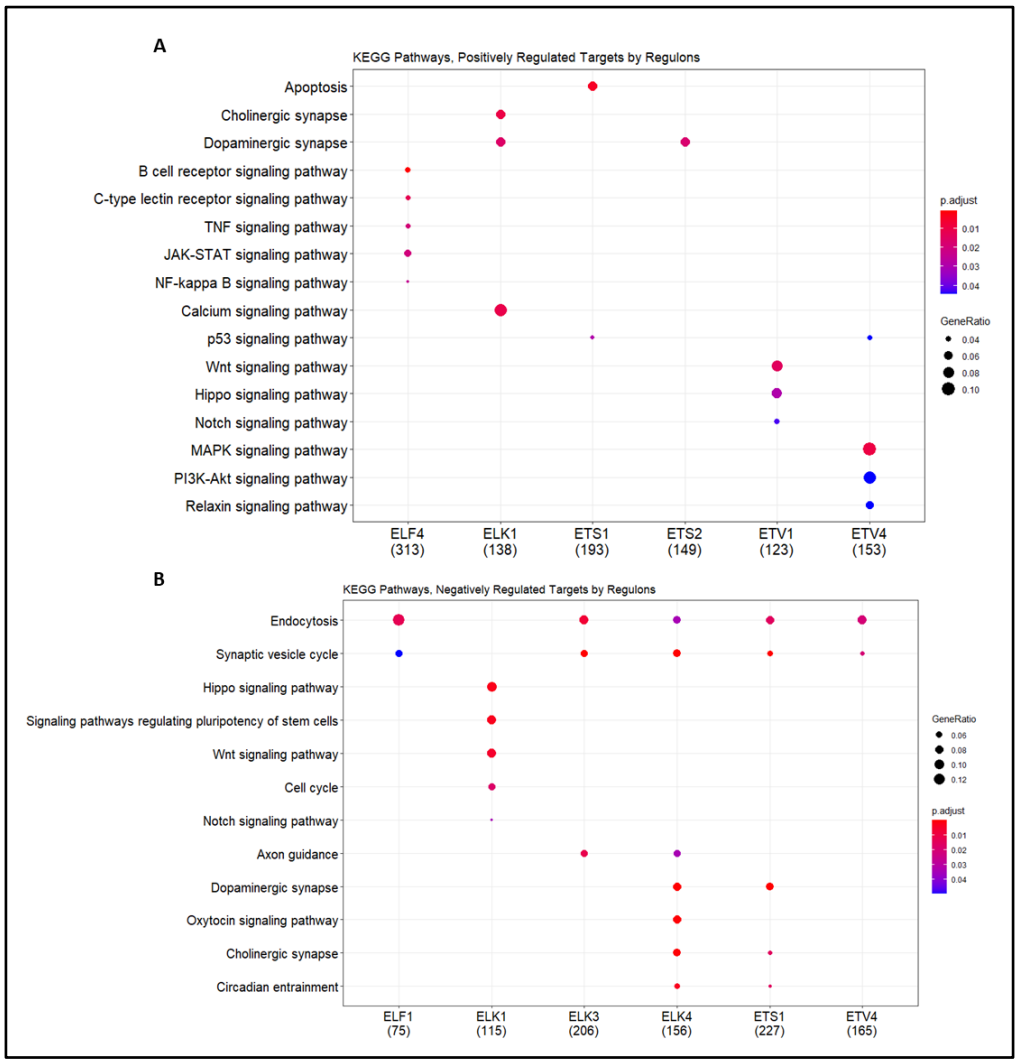


Figure 3.7: KEGG enrichment analysis of certain regulatory targets that are both positively and negatively regulated. With an adjusted p-value of 0.05, clusterProfiler carried out the KEGG analysis. The gene ratio shows the percentage of all the genes included in the enrichment analysis that are enriched with the associated KEGG terms.

3.2. Transcriptomic comparison of neuronal differentiation across species under the regulation of ETS TFs

3.2.1. ETS subfamilies in different species

To investigate the functional evolution of ETS domain proteins across species in relation to nervous system development, I first identified neuronal differentiation RNA-seq datasets from *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*,

Gorilla gorilla, and *Caenorhabditis elegans* to be used in this investigation. I have created a map of the ETS subfamilies and subfamily members in these species, together with domain structures of each subfamily, prior to the transcriptome study of gene regulatory networks (Figure 3.8). The ETS subfamily contains the human and gorilla ETS1 and ETS2 proteins, the mouse Ets1 and Ets2 proteins, which have both the DNA-binding ETS domain and the recognizable pointed domain on the protein, and the *Drosophila* Pointed 1 and 2 (PntP1, PntP2) proteins, which have only the ETS domain in PntP1 but both the ETS and pointed domains in PntP2. (Figure 3.8, blue lines).

Human, gorilla, and mouse ERG, FLI1, FEV, and *Drosophila* ERG subfamily member Ets65A, which only has the ETS domain and *Drosophila* Ets21C, which only has the ETS domain, and *C. elegans* ets-5, which only has the ETS domain, all belong to the ERG subfamily (Figure 3.8, petrol green lines). ETS and pointed domains are present in members of the ELG subfamily such as GABPα and ELG in humans and gorillas, Gabpα in mice, and Ets97D in *Drosophila*, although *C. elegans* ets-9 only possesses the ETS domain (Figure 3.8, mauve line colors).

ELF subfamily is defined by ELF1, ELF2 and ELF4 in humans and gorillas, Elf2 and Elf4 in mice, E74A and E74B in *Drosophila*, and elf-1 in *C. elegans*, and solely contains the ETS domain (Figure 3.8, cardinal red lines). While *C. elegans* ets-7 only has the ETS domain, the TEL subfamily is recognized by ETV6, ETV7, and YAN in humans and gorillas, Etv6 in mice, and Aop/YAN in *Drosophila* (Figure 3.8, lime green lines). ETV1, ETV2, ETV4, and ETV5 are members of the PEA3 subfamily, and Etv1, Etv4, and Etv5 are members of the PEA3 subfamily in mice, *Drosophila*'s Ets96B, and C. elegans' Ets-8. The PEA3 subfamily solely includes the ETS domain (Fig.1, orange lines). Only PDEF in humans, gorillas, mice, and ets-4 in *C. elegans* are members of the PDEF subfamily and have both the ETS and pointed domains, whereas *Drosophila* Ets98B only has the ETS domain (Figure 3.8, green lines). SPI, SPIB, and SPIC are members of the SPI subfamily, which includes just the ETS domain in humans, gorillas, and mice. *Drosophila* and *C. elegans* have not yet been shown to have any members of the SPI subfamily (Figure 3.8, mustard yellow lines).

The only members of the TCF subfamily that contain the ETS domain are ELK1, ELK3, and ELK4 in humans, gorillas, and mice. *Drosophila* and *C. elegans* have yet to be identified as members of the TCF subfamily (Figure 3.8, purple lines). While no *Drosophila* ortholog has been found, members of the ESE subfamily include ESE1,

ESE2, and ESE3 in humans and gorillas, just Ese1 in mice, and svh-5 in *C. elegans* (Fig.1, yellow lines). Lastly the ERF subfamily only contains the ETS domain and includes the ERF and ETV3 in humans, gorillas, and mice (Figure 3.8, red lines).

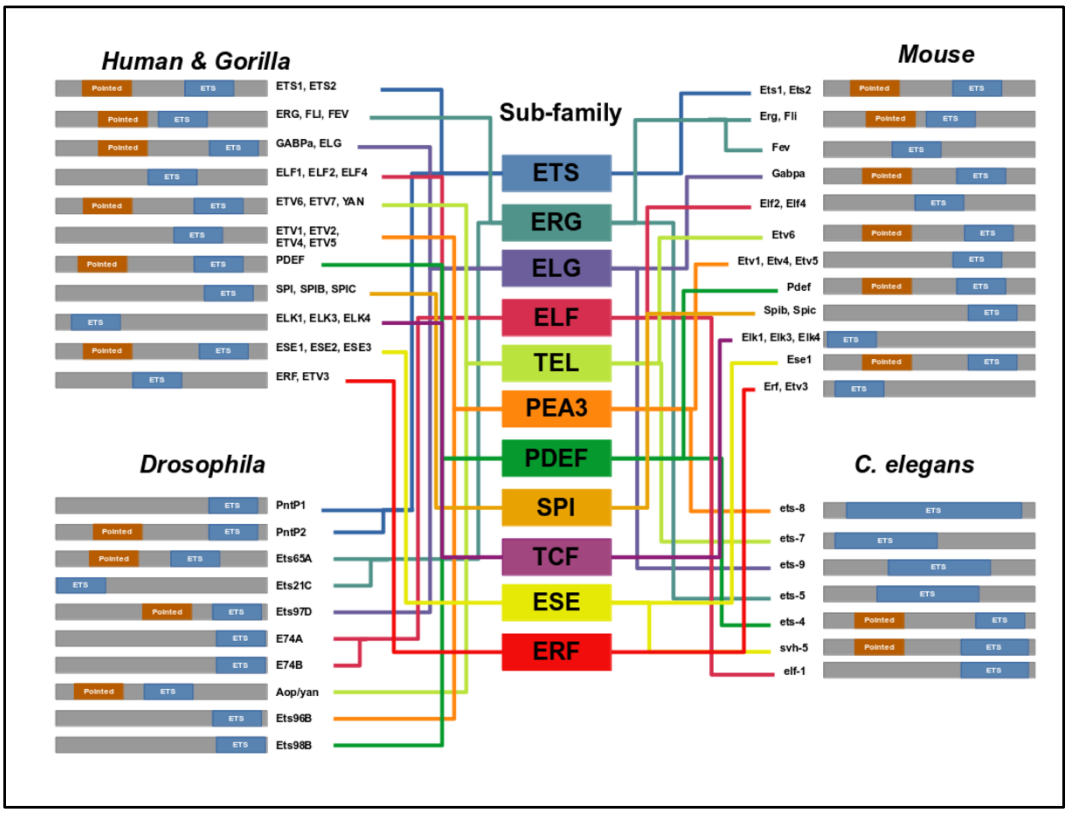


Figure 3.8: Distribution of ETS subfamilies and members located in species in this study. Each colored lines represent the connection between subfamilies and members.

3.2.2. Transcriptomic comparison of neuronal differentiation and gene regulatory network of ETS TFs

To better understand the functional evolution of ETS proteins in relation to nervous system development, transcriptome datasets were looked from other species that provided information on neuronal differentiation or embryonic nervous system development, and discovered datasets from five distinct organisms: *Homo sapiens*, *Gorilla gorilla*, *Mus musculus*, *Drosophila melanogaster*, and *Caenorhabditis elegans*. In terms of evolutionary connections, *D. melanogaster* and *C. elegans*, as well as *H. sapiens* and *G. gorilla*, are the most recent, having emerged in the neo-proterozoic period. *M. musculus*, on the other hand, first appeared in the mesozoic era.

I first built subfamily and domain association maps amongst the creatures to be examined, and then we investigated neuronal differentiation transcriptome information from these organisms to build gene regulatory networks relevant to nervous system development. To that purpose, I have gathered microarray datasets for *C. elegans* neuronal development and RNA-seq datasets of neural stem or progenitor cells vs adult neurons for human, gorilla, mouse, and *Drosophila*. Following the examination of these transcriptome profiles for DEGs, gene regulatory networks of ETS superfamily members were created, and functional enrichment analysis was performed.

The volcano plots revealed that all transcriptome studies included significantly changed genes that were differentially expressed (Figure 3.9). When these DEGs were analyzed in terms of biological processes, it was found that, in general, downregulated genes in all organisms were related to cell division and growth processes, whereas upregulated genes in all organisms were related to neuronal differentiation and function processes, though there were some minor species-dependent differences (Figure 3.10), suggesting that, despite the possibility that nervous systems may have varied, they were likely related to the same biological processes.

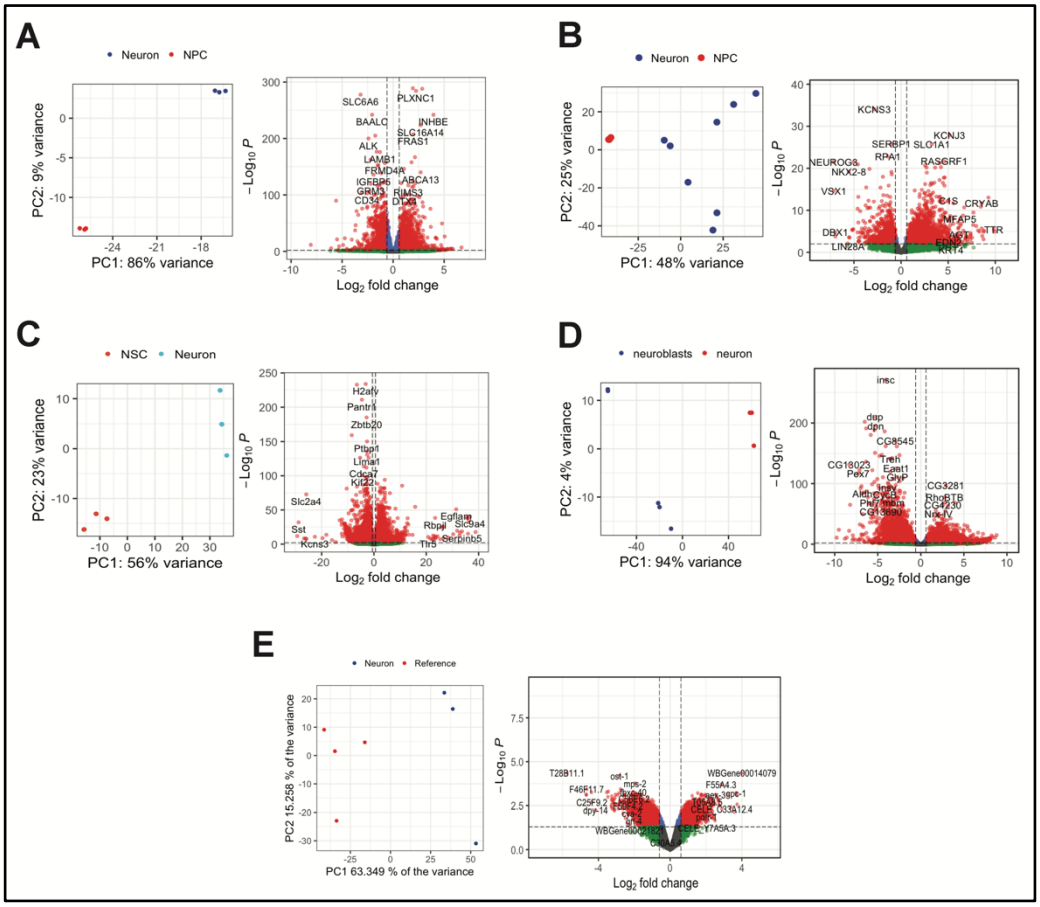


Figure 3.9: Explanatory investigation of transcriptomic changes between stem or progenitor cells vs corresponding mature neurons from transcriptomic datasets of (A) human, (B) gorilla, (C) mouse, (D) drosophila, (E) *C. elegans*. PCA plots show the sample distribution and volcano plots show the relative changes of DEGs.

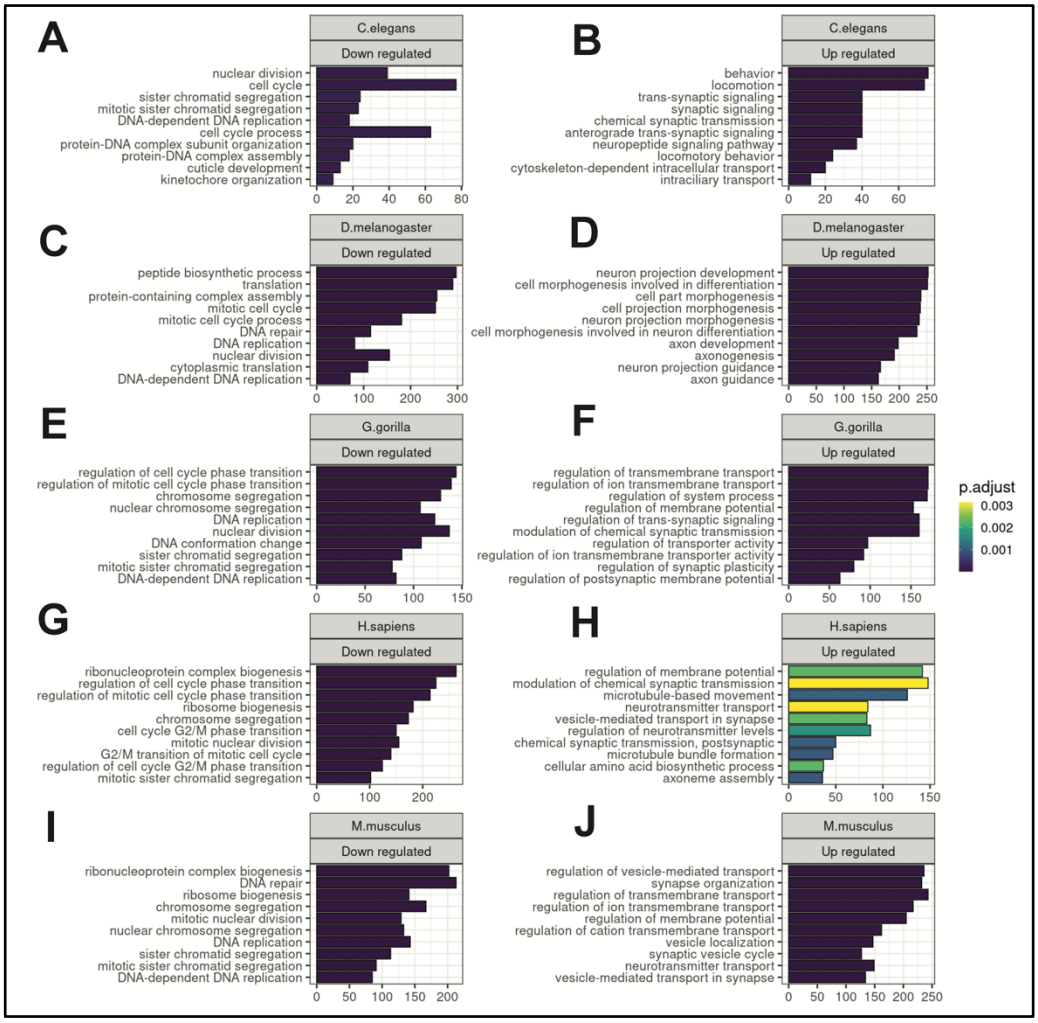


Figure 3.10: GO over-representation analysis of the biological processes for significantly changed genes resulted from differentially gene expression analyses. All analyses were performed with up and down-regulated DEG lists of the corresponding organism as an input: (A) downregulated genes in *C. elegans*, (B) upregulated genes in *C. elegans*, (C) downregulated genes for *D. melanogaster*, (D) upregulated genes for *D. melanogaster*, (E) downregulated genes for *G. gorilla*, (F) upregulated genes for *G. gorilla,* (G) downregulated genes for *H. sapiens*, (H) upregulated genes for *H. sapiens*, (I) downregulated genes for *M. musculus*, (J) upregulated genes for *M. musculus*. Significant GO terms were determined with an adjusted p-value < 0.05 Benjamini-Hochberg correction (see color code scale).

As a result, I have been looking at the ETS regulation networks in each of these transcriptomic datasets. Although it was not feasible to compare the changes in gene expression of every member of the ETS family across species because not every member of the family had a substantial change in each transcriptome, several important points were nevertheless discovered.

Only in the *H. sapiens* dataset, where expression levels of several ETS members were examined on Day 0, Day 1, Day 2, Day 4, Day 5, Day 11 and Day 18, where a

very dynamic expression pattern was seen, were time-dependent gene expression changes during neural differentiation conceivable (Figure 3.11.A). ELF2 was a member of the ELF subfamily and displayed essentially minimal expression on Day 0, but have very fluctuating gene expression pattern compared to other ETS members. SPIB, on the other hand, showed downregulation between Days 0 and 1, an increase between Days 2, 4, and 5, a decrease on Day 11, and an increase once again on Day 18. A number of different subfamily members, including the PEA3 subfamily members ETV4 and ETV5, the TCF subfamily member ELK3, the ERF subfamily member ERF, the TEL subfamily member ETV6, the ETS subfamily member ETS1, and the FLI subfamily member FLI1, displayed a profile that was similar to SPIB. The main difference was that these subfamily members had high expression on Day 0 instead of a similar downregulation on Day 1, upregulation to different levels on (Figure 3.11.A). Interestingly, members of the PEA3 subfamily ETV1 and ETS subfamily ETS2 had distinct expression patterns from their aforementioned family members: Both ETV1 and ETS2 showed little to no expression on Days 2–5, but upregulation on Day 11. ETV1 was substantially expressed on Days 0 and 1, but ETS2 was highly expressed only on Day 0 and subsequently downregulated on Day 1. (Figure 3.11.A).

TCF subfamily members ELK1 and ELK3 display opposite patterns of expression in the gorilla dataset, with ELK1 expressed at lower levels in NPCs and greater levels in mature neuron datasets, and ELK3 expressed at higher levels in NPCs but at a downregulated level in neuron datasets (Figure 3.11.B). ETV4 and FLI1 are substantially expressed in NPCs but have decreasing yet varied levels of expression in other neuron samples, whereas PDEF subfamily member SPDEF in gorilla follows a pattern similar to ELK1 (Figure 3.11.B).

The fact that mouse datasets include NSCs, early and late neuroblasts, and mature neurons makes them more similar to human datasets. The dynamic profile of ETS members' expression was similar to that of the human and gorilla datasets: Etv4 and Spdef, members of the Pea3 subfamily, had decreased expression in NSCs and early neuroblasts but increased expression in late neuroblasts and mature neurons (Figure 3.11.C). Etv1, a different member of the Pea3 subfamily that in human datasets was strongly expressed on Day 0 but subsequently dropped on Day 18 but exhibited low expression in mice NSCs, was increased in early and late neuroblasts, and then decreased again in mature neurons (Figure 3.11.C). Elf4 and Elf2 were both

extensively expressed in NSCs, and in Elf2's case, early neuroblasts, but their expression levels were significantly reduced in late neuroblasts and mature neurons (Figure 3.11.C). Elk-1 expression levels were drastically downregulated in mouse NSCs, elevated in early and late neuroblasts, and lowered once more in mature neurons, in contrast to TCF subfamily members in the human and gorilla transcriptomes (Figure 3.11.C).

Similar to another member of the family, Gabpα in mice, *Drosophila* Ets97D, a member of the ELG subfamily, was strongly expressed in neuroblasts and downregulated in mature neurons (Figure 3.11.D). The TCF, TEL, PDEF, PEA3 and ETS subfamily members Ets65A, aop, Ets98B, Ets96b, and pnt, on the other hand, were all elevated during *Drosophila*'s neural development (Figure 3.11.D). In the *C. elegans* sample, only ets-5 and ets-6 were significantly changed, and both were downregulated in neurons (Figure 3.11.E).

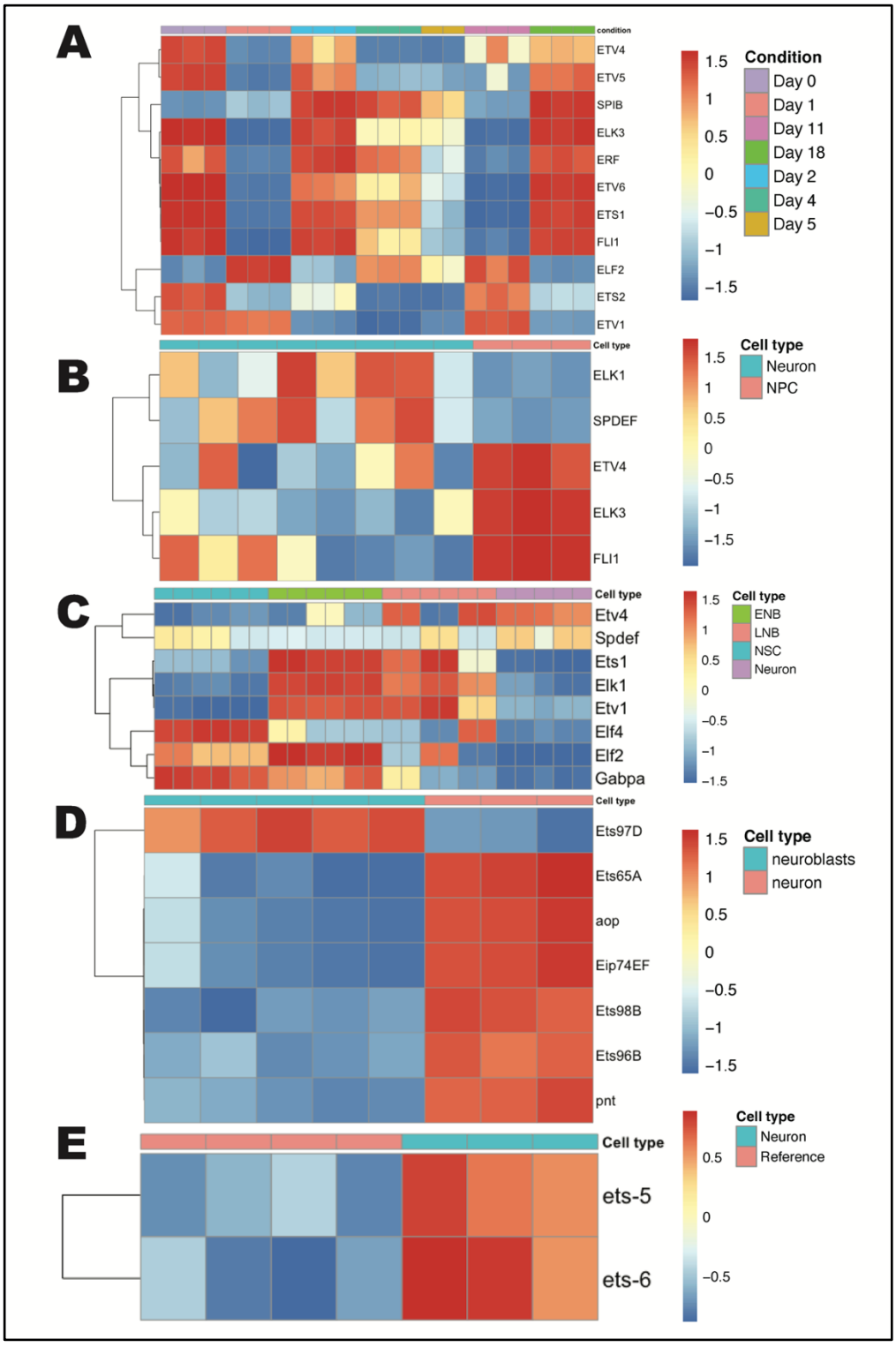


Figure 3.11: Regulon activity of ETS members. Heatmaps represent the expression changes of significantly changed ETS members with normalized regulon activity scores. (A) H. sapiens profiles during Day 0, 1, 2, 4, 5, 11 and 18 of neuronal differentiation; (B) G. gorilla profiles of NPCs vs mature neurons; (C) M. musculus profiles of NSC, ENB, LNB and mature neurons; (D) D. melanogaster profiles of neuroblasts vs neurons; (E) C. elegans profiles of mature neurons vs reference cells. NPC: neuronal precursor cell, NSC: neural stem cell, ENB: early neuroblast, LNB: late neuroblast.

Then, for each species dataset, I performed GO functional enrichment analysis and gene regulatory network analysis for significantly changed ETS proteins (Figure 3.12-21). ETV4, ETV5, FLI1, ETS1, ELK3, SPIB, and ETV6 appear to have more shared targets in the human regulatory network than ETV1, which mostly regulates its own particular network (Figure 3.12-13). Elf2 and Ets1 have the most shared targets in the mouse dataset, but Elf4 has a separate independent regulatory network (Figure 3.14-15). In the *Drosophila* dataset, targets for apo, Eip74EF, Ets97D, pnt, and Ets65A are shared, but Ets96B and Ets98B have separate regulatory networks (Figure 3.16-17). Only 2 significantly changed genes, ets-5 and ets6, were discovered in *C. elegans*, and they were discovered to have their own distinct regulatory networks (Figure 3.18-19). The gorilla dataset's gene regulatory network study similarly produced a very straightforward network, with ELK1 and SPDEF both having their own distinct and independent regulatory networks (Figure 3.20-21).

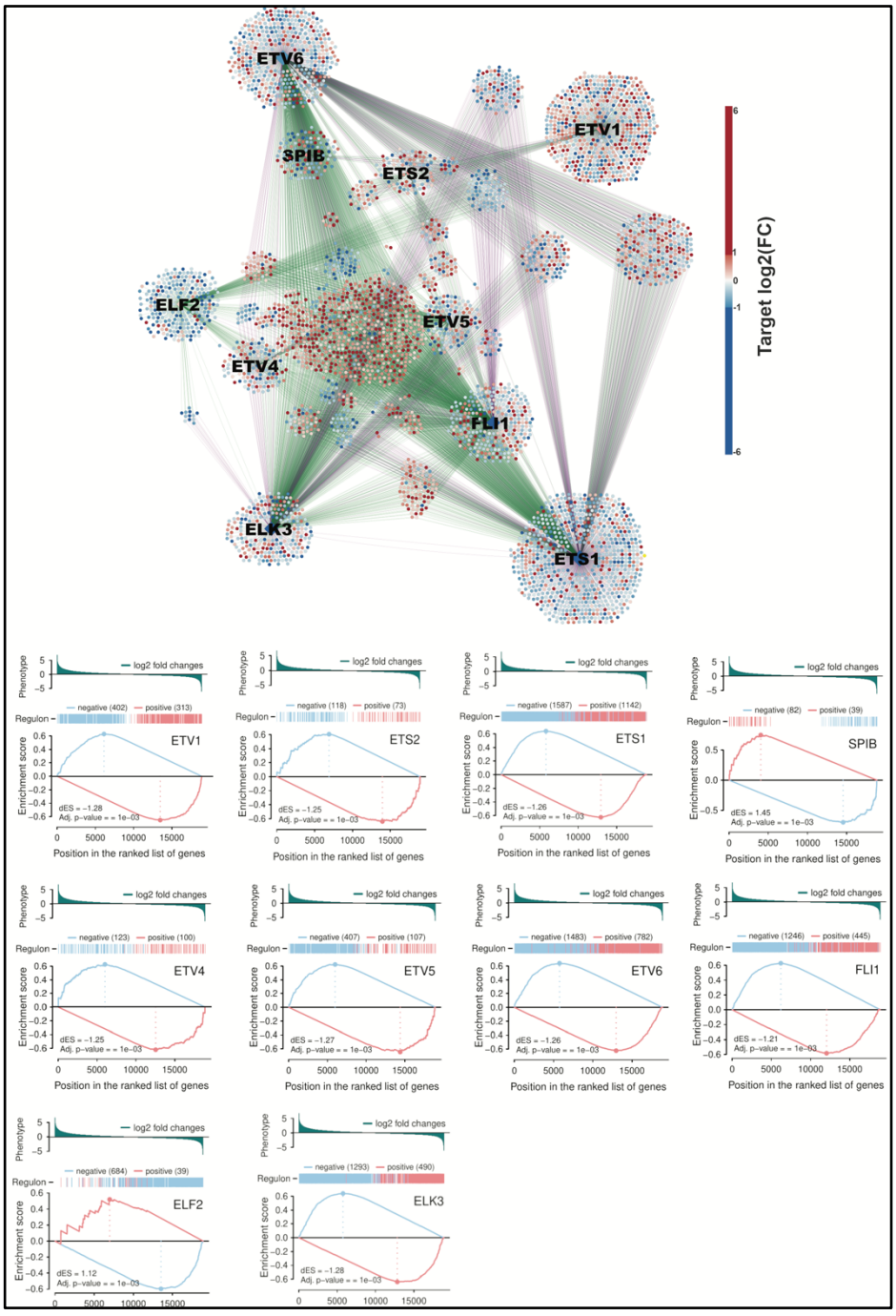


Figure 3.12: Human neural development gene regulatory network governed by ETS transcription factors. Circle nodes stand in for the target genes, whereas diamond nodes indicate ETS members as a regulator. The circle's color represents the differential gene expression analysis's log 2-fold change in target genes between neuron and NPC. According to the GSEA analysis, edge colors represent the enrichment score of each target gene with its related regulators. While green borders represent negative regulation, purple edges represent positive regulation. The two-tailed GSEA analysis related to positively and negatively regulated targets of certain regulons is shown in the bottom panel. Gene expression analysis ranks target genes, and enrichment analysis scores them by calculating the edge weight of the gene regulatory network.

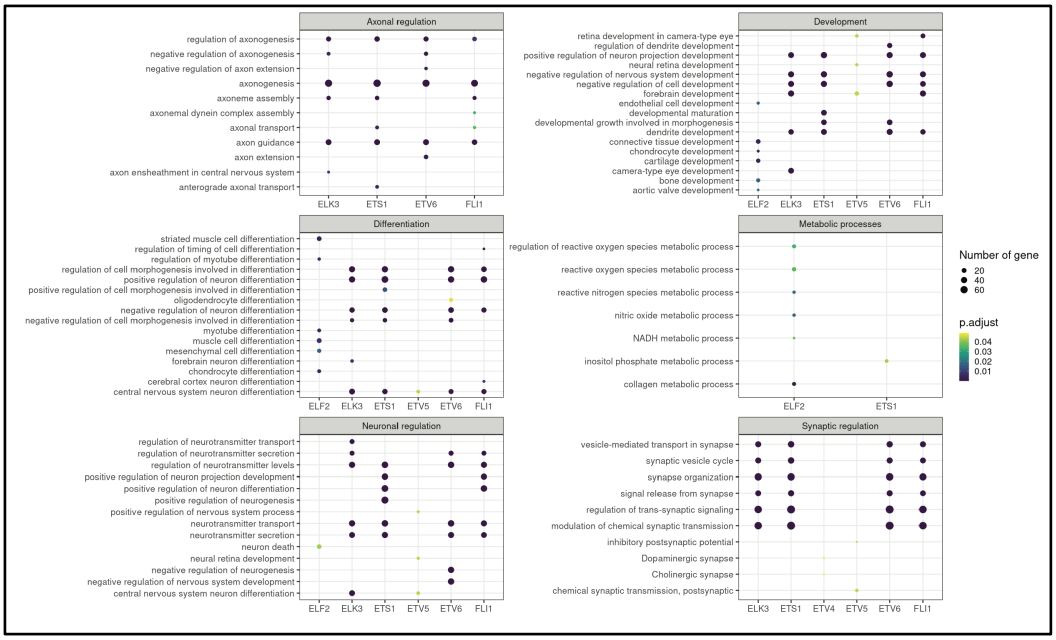


Figure 3.13: Gene Ontology encrichment analysis of human ETS regulons. Enriched terms were identified using adjusted p-value < 0.05 with Benjamini-Hochberg correction.

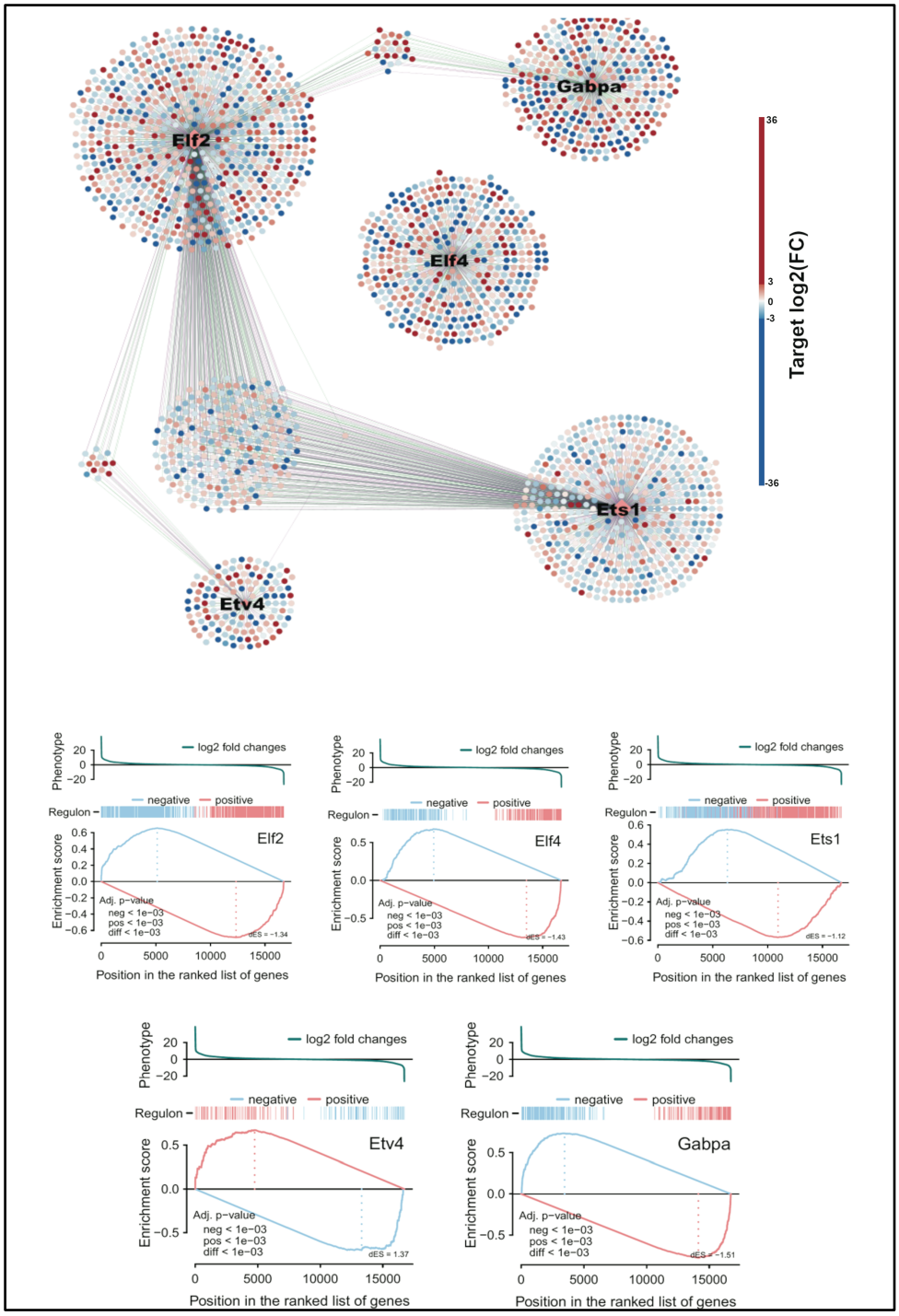


Figure 3.14: Mouse neural development gene regulatory network governed by ETS transcription factors. Circle nodes stand in for the target genes, whereas diamond nodes indicate ETS members as a regulator. The circle's color represents the differential gene expression analysis's log 2-fold change in target genes between neuron and NPC. According to the GSEA analysis, edge colors represent the enrichment score of each target gene with its related regulators. While green borders represent negative regulation, purple edges represent positive regulation. The two-tailed GSEA analysis related to positively and negatively regulated targets of certain regulons is shown in the bottom panel. Gene expression analysis ranks target genes, and enrichment analysis scores them by calculating the edge weight of the gene regulatory network.

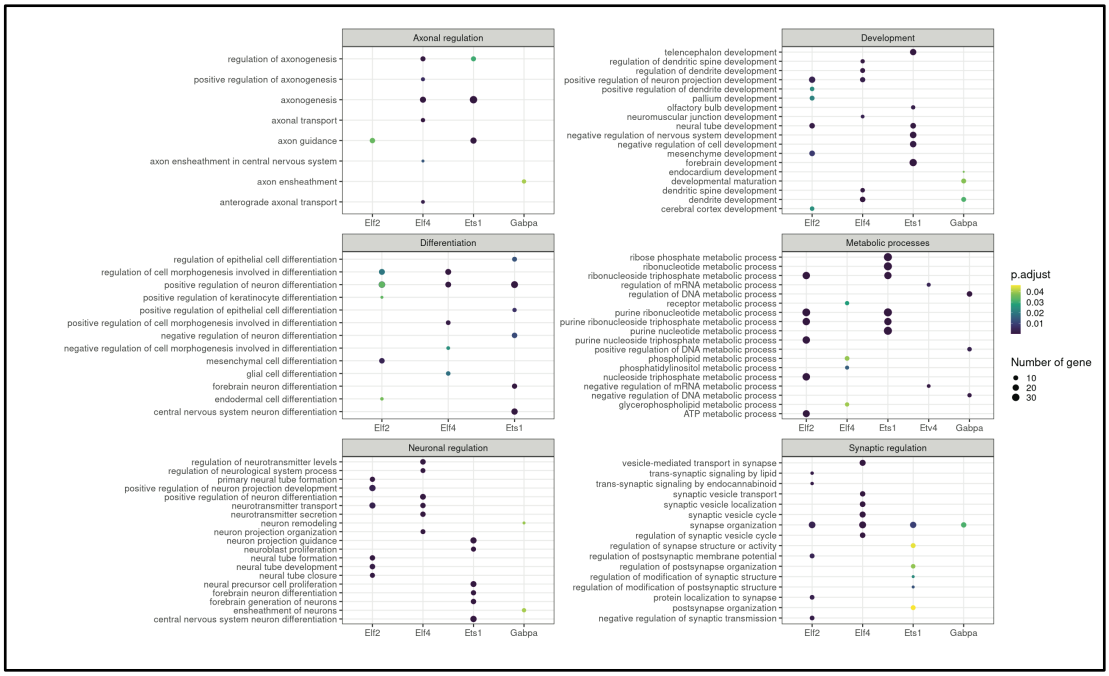


Figure 3.15: Gene Ontology encrichment analysis of mouse ETS regulons. Enriched terms were identified using adjusted p-value < 0.05 with Benjamini-Hochberg correction.

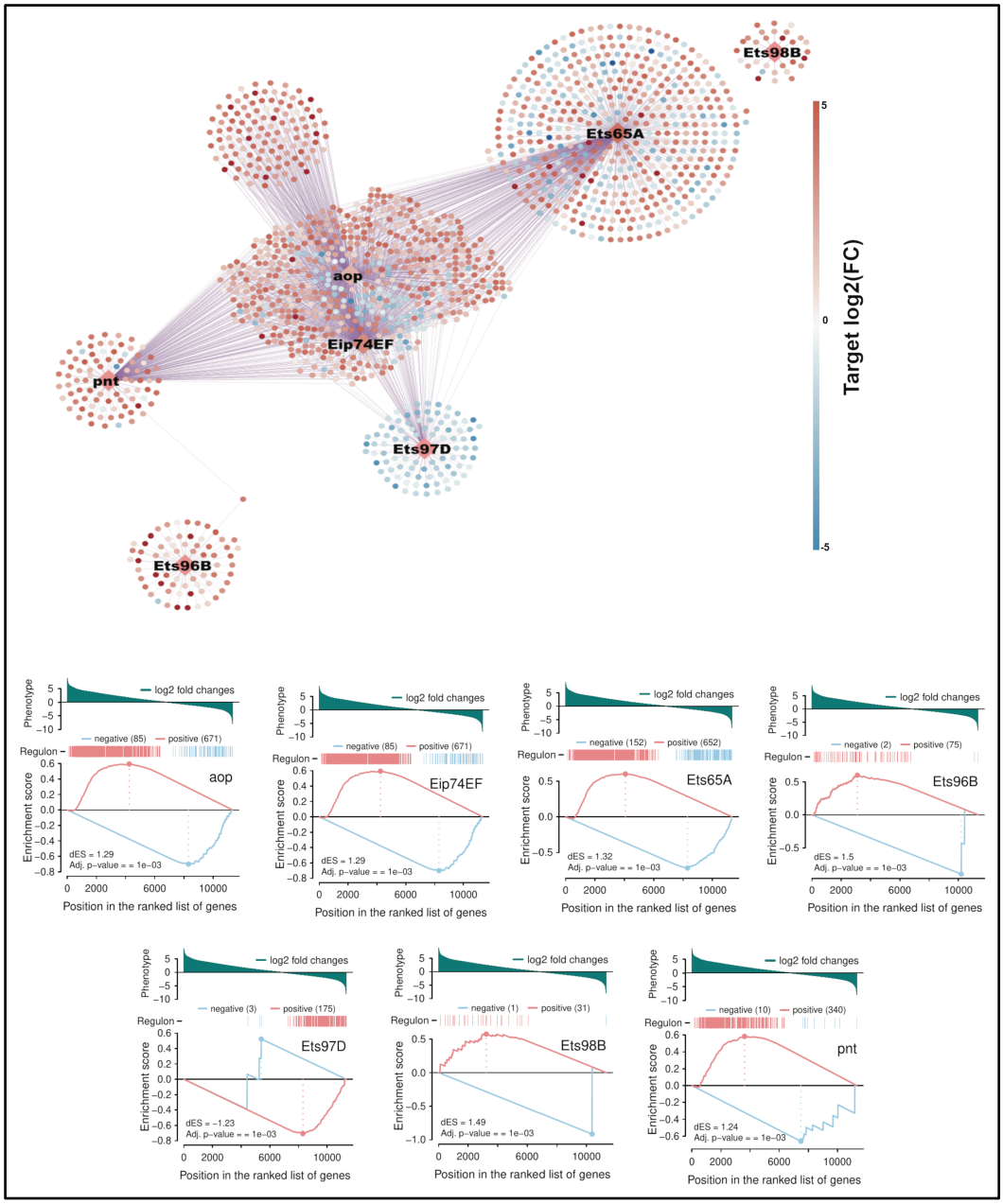


Figure 3.16: *Drosophila* neural development gene regulatory network governed by ETS transcription factors. Circle nodes stand in for the target genes, whereas diamond nodes indicate ETS members as a regulator. The circle's color represents the differential gene expression analysis's log 2-fold change in target genes between neuron and NPC. According to the GSEA analysis, edge colors represent the enrichment score of each target gene with its related regulators. While green borders represent negative regulation, purple edges represent positive regulation. The two-tailed GSEA analysis related to positively and negatively regulated targets of certain regulons is shown in the bottom panel. Gene expression analysis ranks target genes, and enrichment analysis scores them by calculating the edge weight of the gene regulatory network.

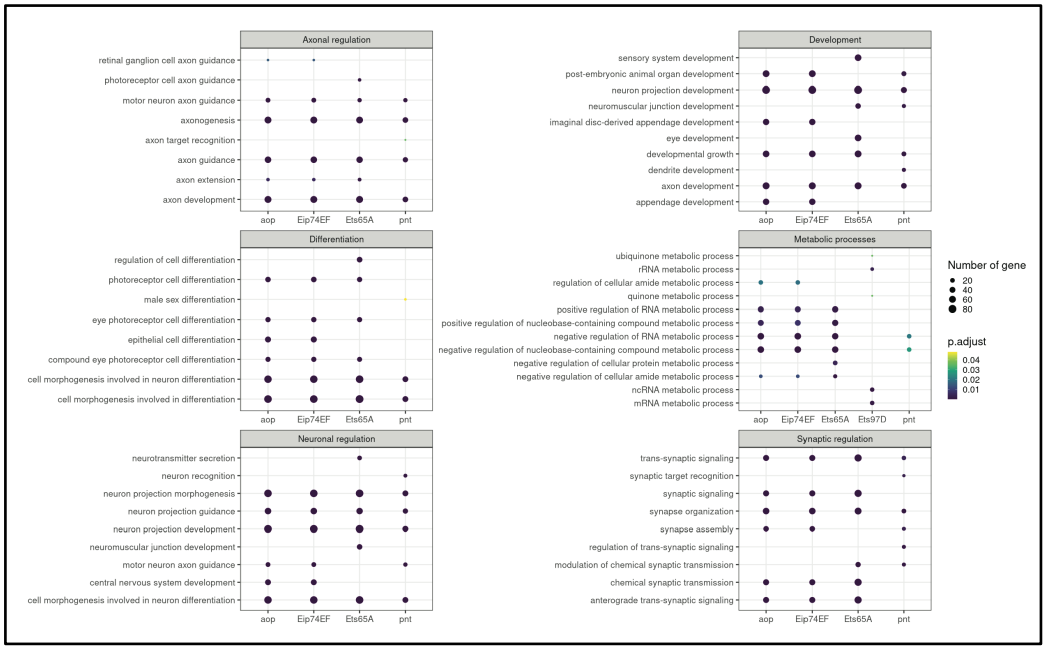


Figure 3.17: Gene Ontology encrichment analysis of *drosophila* ETS regulons. Enriched terms were identified using adjusted p-value < 0.05 with Benjamini-Hochberg correction.

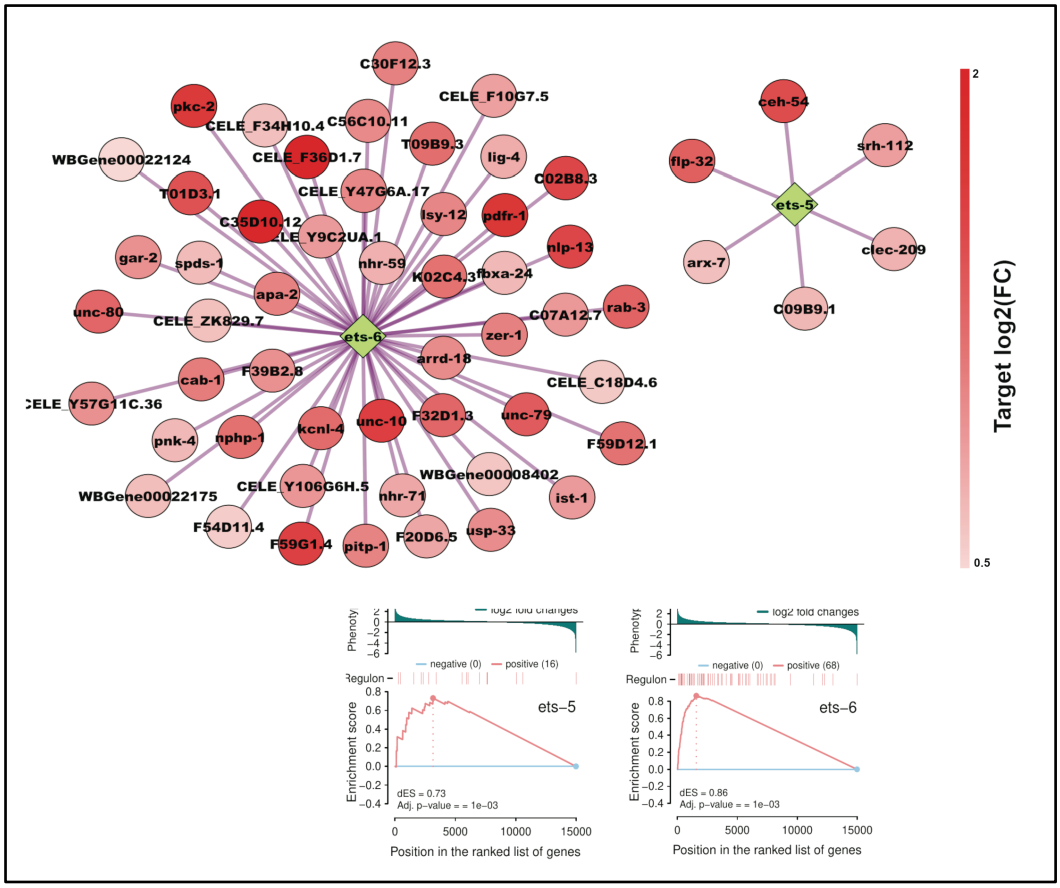


Figure 3.18: *C. elegans* neural development gene regulatory network governed by ETS transcription factors. Circle nodes stand in for the target genes, whereas diamond nodes indicate ETS members as a regulator. The circle's color represents the differential gene expression analysis's log 2-fold change in target genes between neuron and NPC. According to the GSEA analysis, edge colors represent the enrichment score of each target gene with its related regulators. While green borders represent negative regulation, purple edges represent positive regulation. The two-tailed GSEA analysis related to positively and negatively regulated targets of certain regulons is shown in the bottom panel. Gene expression analysis ranks target genes, and enrichment analysis scores them by calculating the edge weight of the gene regulatory network.

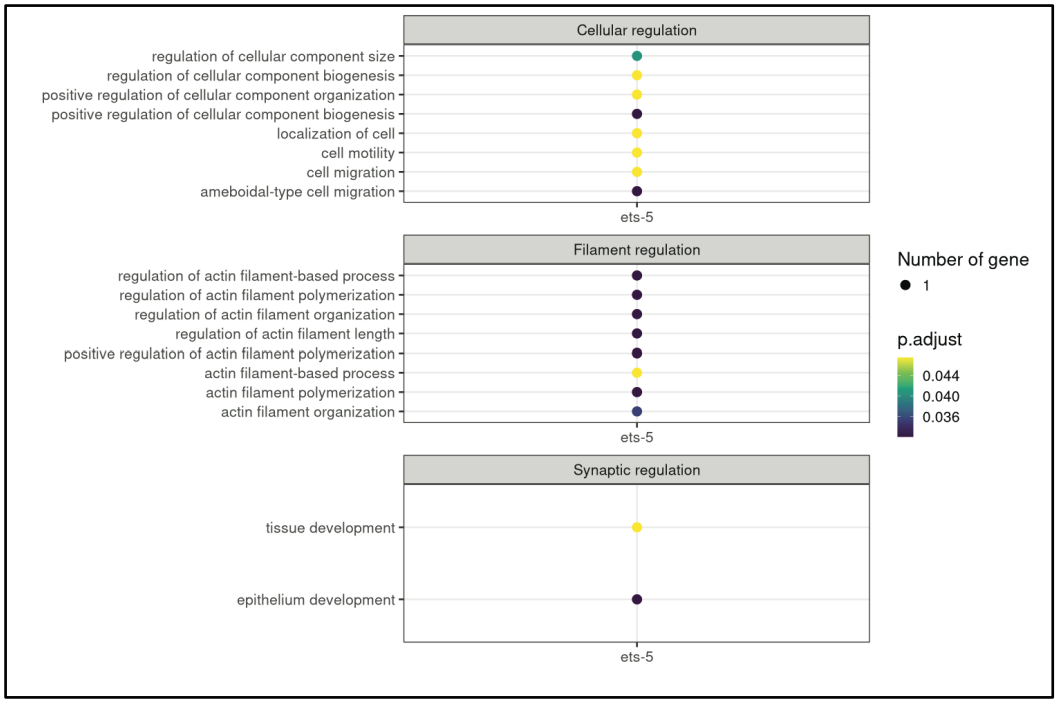


Figure 3.19: Gene Ontology encrichment analysis of *C. elegans* ETS regulons. Enriched terms were identified using adjusted p-value < 0.05 with Benjamini-Hochberg correction.

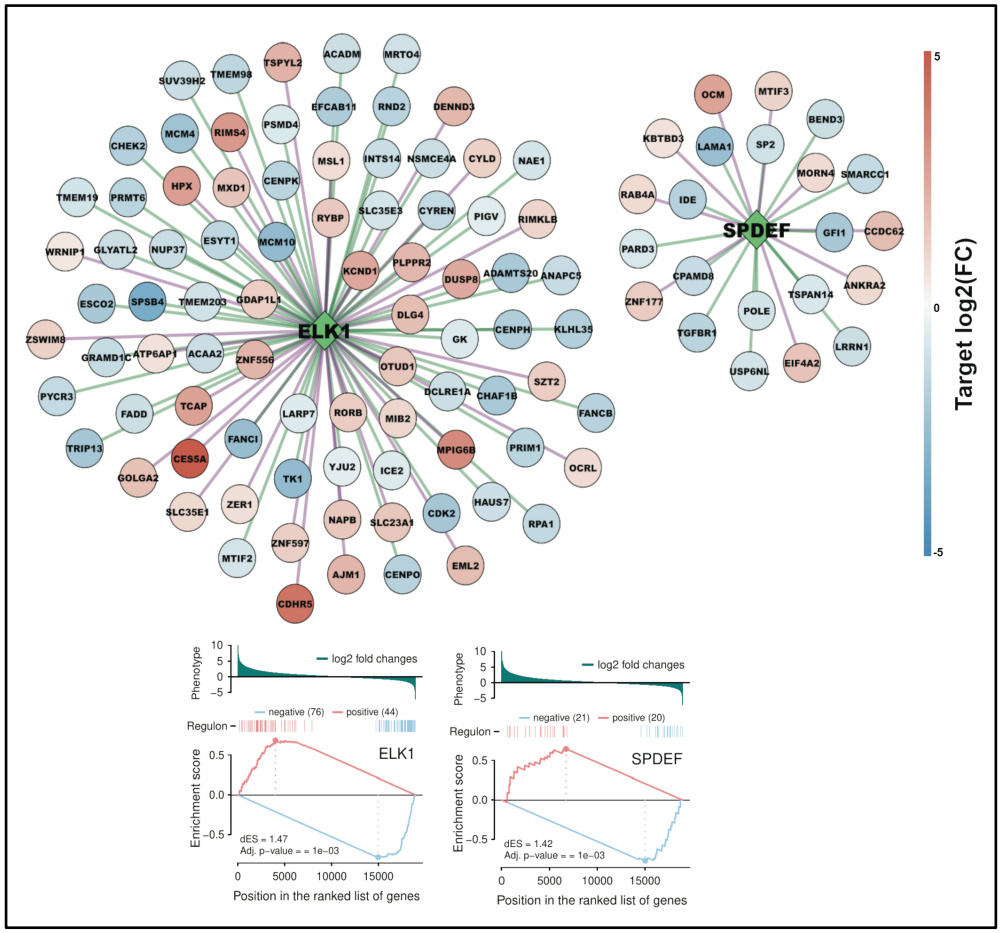


Figure 3.20: Gorilla neural development gene regulatory network governed by ETS transcription factors. Circle nodes stand in for the target genes, whereas diamond nodes indicate ETS members as a regulator. The circle's color represents the differential gene expression analysis's log 2-fold change in target genes between neuron and NPC. According to the GSEA analysis, edge colors represent the enrichment score of each target gene with its related regulators. While green borders represent negative regulation, purple edges represent positive regulation. The two-tailed GSEA analysis related to positively and negatively regulated targets of certain regulons is shown in the bottom panel. Gene expression analysis ranks target genes, and enrichment analysis scores them by calculating the edge weight of the gene regulatory network.

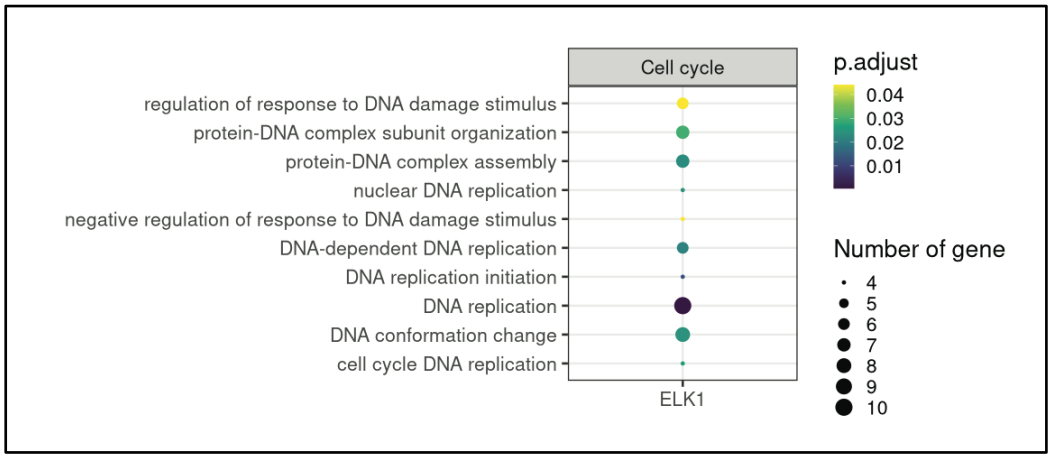


Figure 3.21: Gene Ontology encrichment analysis of gorilla ETS regulons. Enriched terms were identified using adjusted p-value < 0.05 with Benjamini-Hochberg correction.

When these functional enrichments and the gene regulatory networks of the ETS members were visualized as a combined network including the evolutionary tree of each species, it was observed that while many of the ETS members are involved in processes related to nervous system development, such as neuronal regulation, synaptic regulation, differentiation and development, some ETS proteins have unique regulatory networks in specific biological processes, such as *C. elegans* ets-5 in cellular regulation and filament regulation and gorilla ELK1 in the cell cycle (Figure 3.22). In the context of processes relevant to nervous system development, no obvious subfamily-dependent functional division of labor was seen among ETS members, confirming considerable redundancy among family members. Most of the ETS members in *Drosophila* shared targets for axonal regulation, neuronal regulation, development, and synaptic regulation; however, pnt had a smaller network for differentiation, synaptic regulation, and metabolic processes, and Ets97D had entirely different regulatory targets for metabolic processes (Figure 3.22). ETS members also shown substantial levels of redundancy in the human dataset, with several ETS proteins sharing regulatory targets for differentiation, axonal control, and synaptic regulation (Fig.6, Suppl Fig 2). The mouse dataset, on the other hand, showed that each ETS member appeared to have their own regulatory network, with some similarities, such as Elf4 and Ets1 in axonogenesis, Ets1 and Elf2 in metabolic processes, and Elf2 and Elf4 in neuronal control (Figure 3.22).

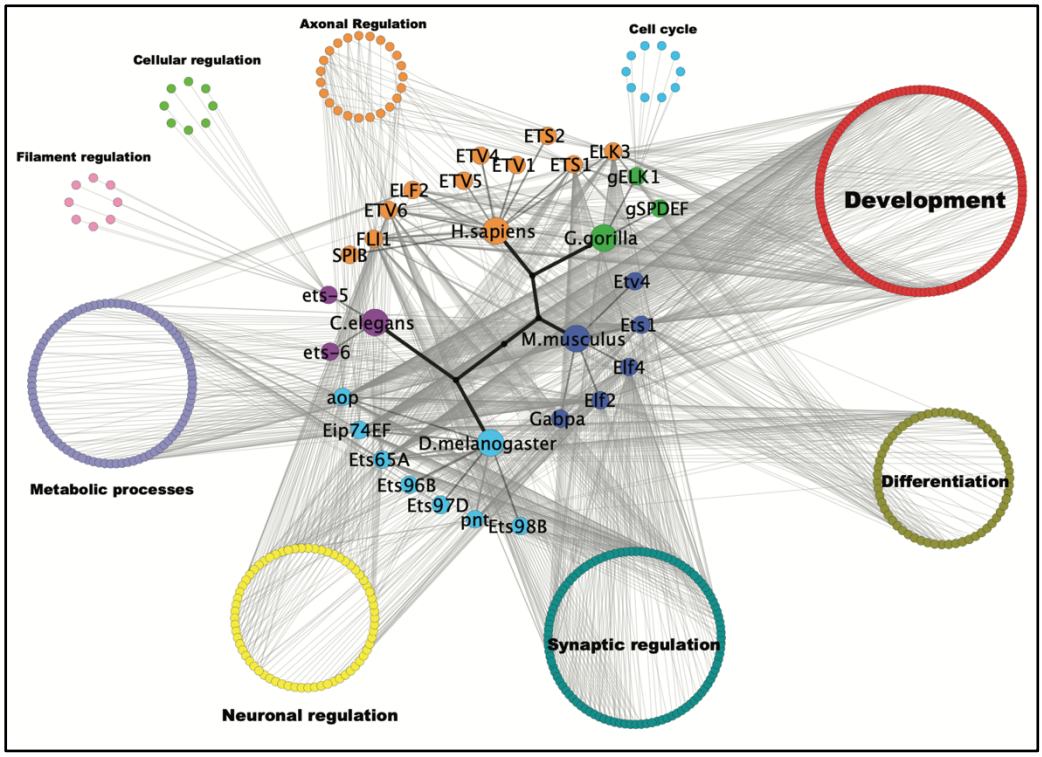


Figure 3.22: Combined functional transcriptional network of ETS members. Phylogenetic distances are represented in the middle of the network as an evolutionary tree, each ETS member is connected with their organism. Annotated nodes indicate the individual enriched GO terms resulting from functional enrichment analysis of ETS regulon targets.

3.3. Transcriptomic comparison of primary neurons and RA-differentiated SH-SY5Y cells

Individual data sets were compared to control samples for differential gene expression analysis. In the RA-SH-SY5Y dataset, 18-day RA differentiated SH-SY5Y cells were compared to undifferentiated control samples. After performing all raw data analysis and normalization steps, principal components of individual data sets were investigated to check sample distributions and detect outliers. According to the PCA plots, the percentage of PCs and sample distributions were observed and decided to be suitable for differential gene expression analysis (Figure 3.23).

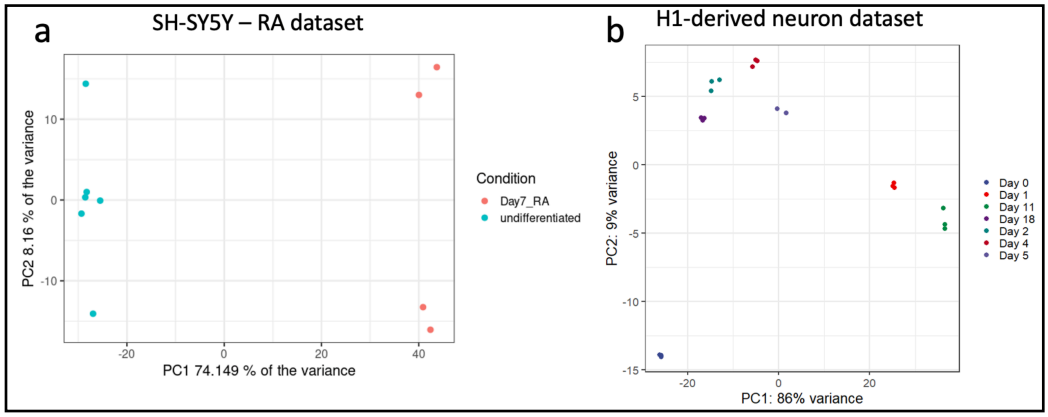


Figure 3.23: (A) PCA plots of the SH-SY5Y RA dataset and (B) Primary neuron dataset. Each dot represents a sample in the dataset. Colors represent groups of samples in the dataset.

In the primal neuron dataset, day 18 of the neuronal differentiation protocol was compared to a day 0 control. Significantly altered genes were identified after performing differential-gene expression analysis. In both datasets, genes such as PLXNC1 and SNCAIP show significantly increased gene expression for mature neuronal phenotypes (Figure 3.24). After differential gene expression analysis, DEGs were compared with each other for the two datasets and 5699 common DEGs were identified, as can be seen in the Venn diagram (Figure 3.25).

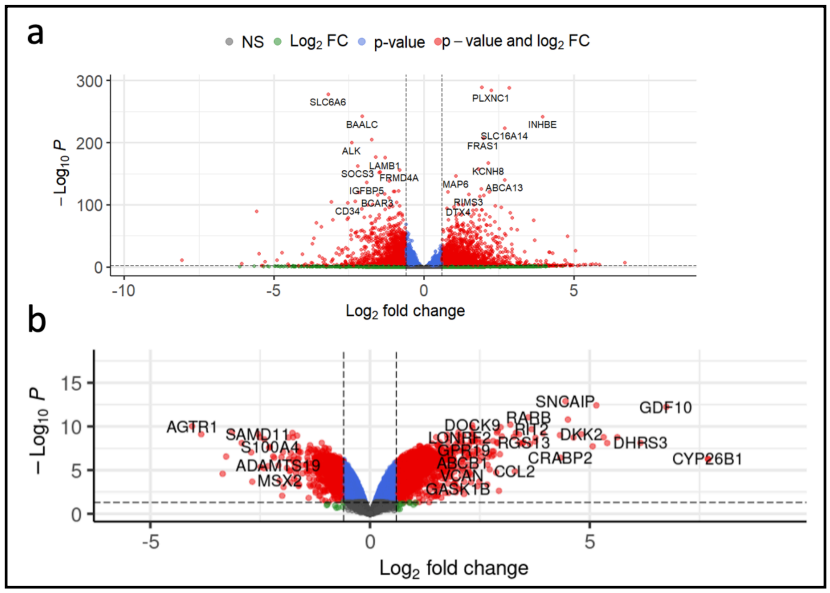


Figure 3.24: Volcano plot for (A) Primary neuron and (B) SH-SY5Y data sets after differential gene expression analysis. The y-axis shows the degree of statistical significance, and the x-axis shows the fold change of genes compared to the control group.

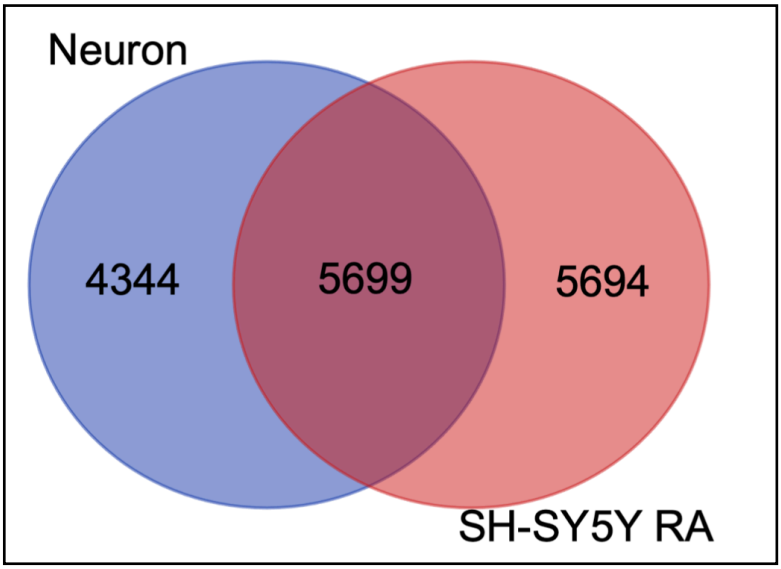


Figure 3.25: Comparison of genes significantly changed as a result of differential gene expression analysis with venn diagram.

GO enrichment analysis was performed on common DEGs to investigate common phenotypic changes for both differentiation procedures. First, DEGs were grouped into increasing and decreasing gene expression for separate datasets and GO enrichment analysis was performed. For decreasing genes, mostly terms related to the nucleus or cell division were obtained (Figure 3.26). On the other hand, increasing genes were mostly identified with neuronal phenotypes (Figure 3.27).

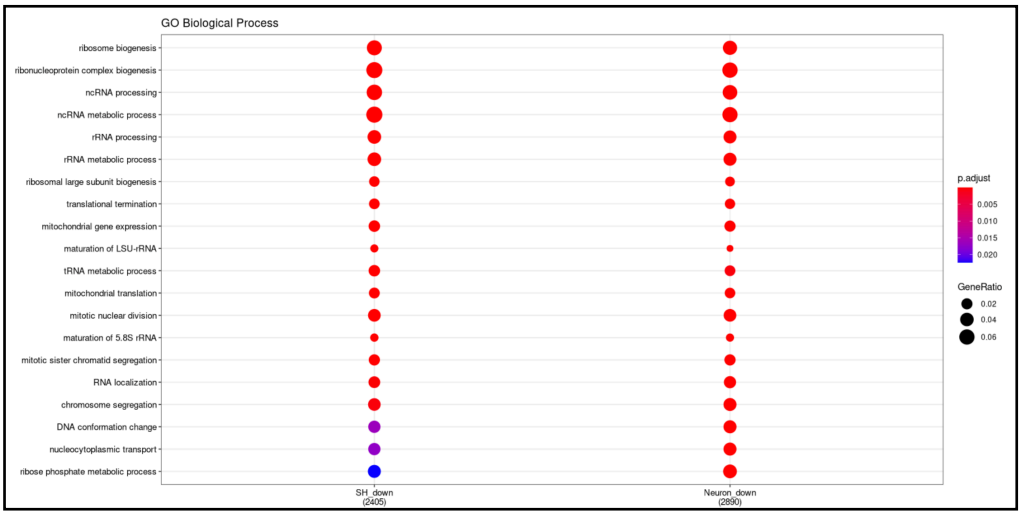


Figure 3.26: GO analysis result of genes with reduced gene expression. The size of each dot indicates how many genes are included for the defined GO term and the colors indicate the degree of statistical significance.

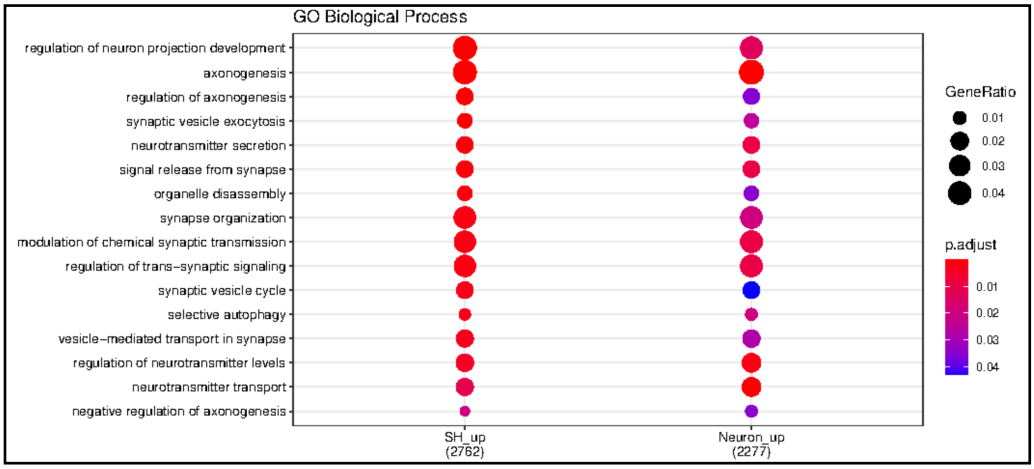


Figure 3.27: GO analysis result of genes with increased gene expression. The size of each dot indicates how many genes are included for the defined GO term and the colors indicate the degree of statistical significance.

To validate the gene expression of these DEGs in SH-SY5Y cells, genes belonging to two key GO terms associated with nervous system development were identified and visualized as a network format with the corresponding DEGs. Using this network, candidate validation genes were selected for qPCR (Figure 3.28). Genes associated with neuronal maturation at the basic level were selected based on the literature and ETS family members in the network were also selected for qPCR validation.

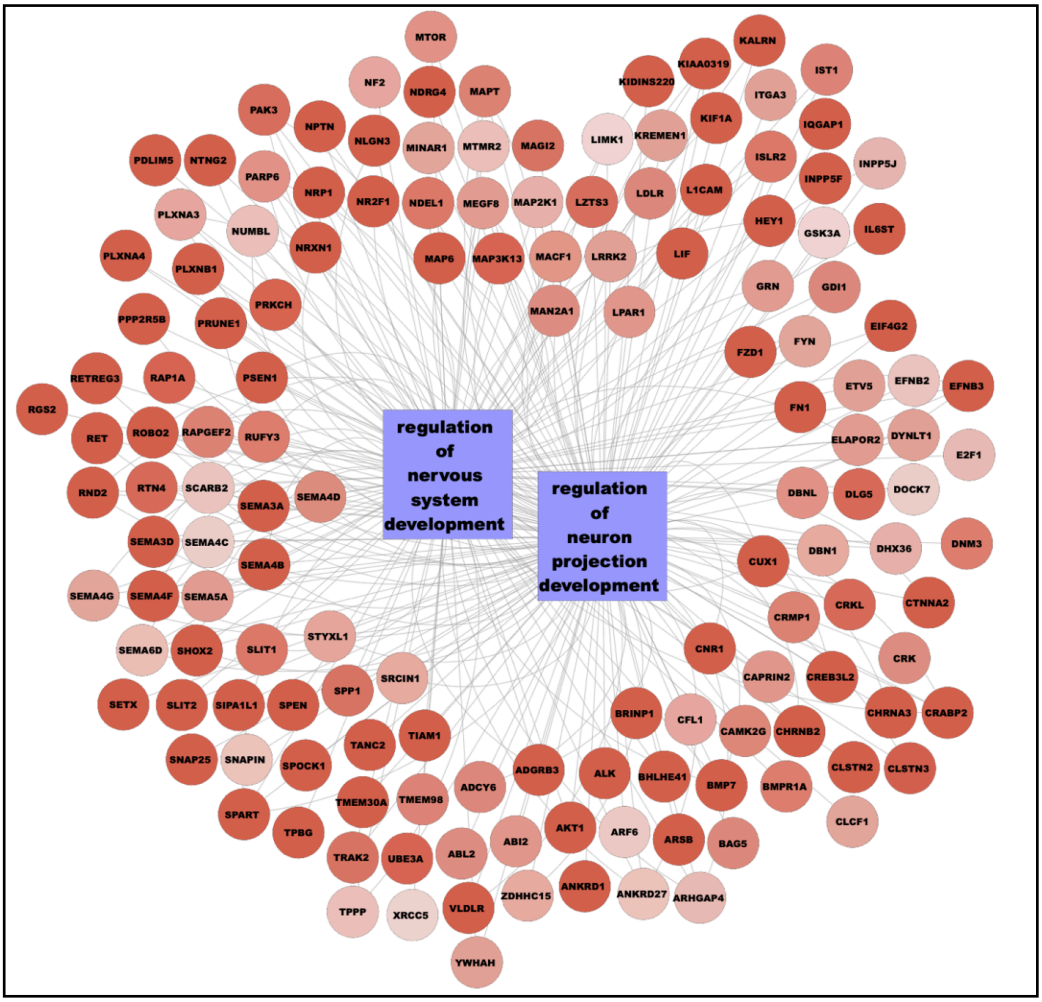


Figure 3.28: Network representation of common genes in the two datasets and genes associated with neuronal differentiation by GO analysis.

Before selecting validation gene targets, SH-SY5Y cells were differentiated with RA differentiation protocol and cell morphology was checked. Much more neuron-like morphology, such as prominent axons and thinner soma, is observed in 1% FBS and 1% FBS + RA groups (Figure 3.29).

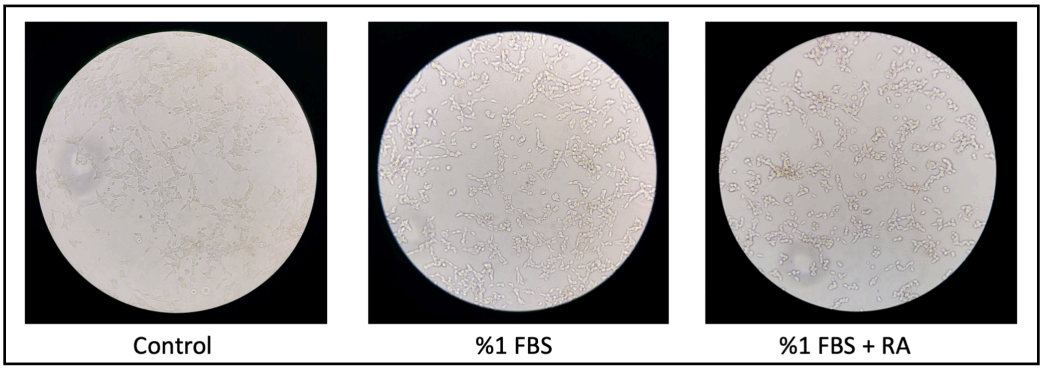


Figure 3.29: Morphological structures of SH-SY5Y cells during RA differentiation.

To validate the results from the transcriptomic comparison in SH-SY5Y cells, 5 genes associated with axonal guidance and development during neuronal differentiation were selected: PLXNA4, PLXNB1, ROBO2, SEMA3D and SEMA5A. In addition, an ETS member ETV5 was also observed in common DEGs and included in the qPCR validation list. As a result of qPCR analyses, an increase in the expression of all selected genes was observed in the 1% FBS + RA group compared to the control. However, only 3 of them were statistically significant (Figure 3.30).

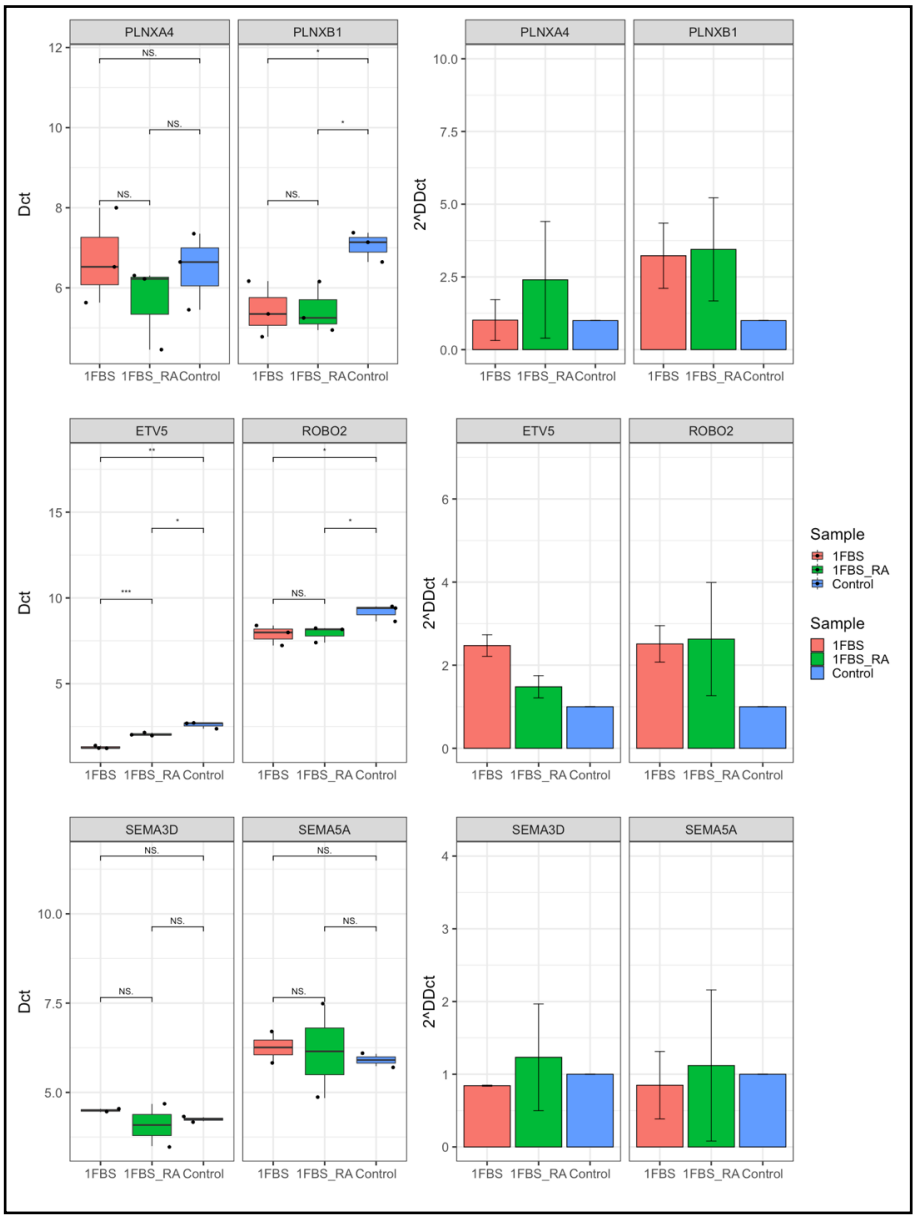


Figure 3.30: qPCR validation of selected genes. The top panel represents a box plot of Dct values with statistical analysis. The bottom panel represents the bar graph of relative gene expression calculated by housekeeping genes and control sample normalization.

4. DISCUSSION and CONCLUSION

In this thesis study I have primarily focused on the ETS superfamily of transcription factors in nervous system development across species however, it should be highlighted that this superfamily also plays a vital role in carcinogenesis, tumor progression, and metastasis via chromosomal translocations, the generation of chimeric proteins, overexpression that results in transformation and malignancy, or modulating epithelial-stromal interactions that enhance metastatic qualities [29]. Etv5 and its network expression have been identified as essential components of the neoplastic state in low-grade glioma [48]. Unfortunately, in the microarray datasets including glioma grades in this study, ETV5 did not seem to be one of the DEGs in tumor vs. non-tumor samples. However, I have discovered that ETV1, a different member of the PEA3 subfamily, is highly expressed in early stages gliomas and barely expressed in later stages (Figure 3.3). In prostate cancer, the ELK-1 protein is reported to be an important requisite for the androgen receptor (AR), and it has been discovered that its expression is linked to a clinically higher level and a predictive indication of disease repetition [35]. In the literature ELK1 is highly associated with cancer progression. In my analysis I have identified the ELK1 expression as downregulated. However, when I investigated this controversy, it is caused by epilepsy samples as control in the dataset. In epilepsy condition, ELK1 was expression very high level. Therefore, ELK1 expression were identified to be downregulated when a comparison was made between epilepsy and glioma grades (Appendix A). In grade 2 gliomas ETV6 and ELF4 were shown to be downregulated but increased in grades 3 and 4 gliomas (Figure 3.3). ETS proteins don't have much tissue specialization, and several family members really have widespread expression [5]. Therefore, it is not unexpected that an examination of ETS transcription factors' gene regulatory networks shows significant target overlap, demonstrating that functional redundancy occurs, at least to some level. However, when the positive and negative regulatory targets of ETS regulons were compared, it was shown that the negatively regulated targets, which mainly concern synaptic vesicle trafficking and synaptic transmission, heavily overlapped. These results also show that ETS proteins influence genes that are also involved in developmental processes.

Therefore, this part of the thesis study is a proof-of-concept study that demonstrates that, in addition to currently available molecular markers ETS genes can

be used as diagnostic approach for glioma grade progression. However, more physiologically relevant non-tumor controls are necessary in the future to fine-tune the results. Additionally, target genes, pathways and processes may be found using gene regulatory network analysis for ETS TFs, which can aid in understanding the molecular mechanisms across family members. I suggest that this sort of network analysis be expanded to distinguish between distinct carcinogenesis phases in additional tumor types as well as other tissue developmental stages.

Throughout metazoa, the ETS superfamily has divided into several subfamilies; among a subset of these ETS proteins, the fastest evolutionary rate was predicted for Er71, Er81, Sap-1, and Net, whereas Ets2, Fli, Erg, and Spi1 were reported to be comparatively more conserved than the previous set [1]. Understanding functional redundancy and the evolution of the subfamilies in development has been of significant interest because ETS proteins are engaged in numerous developmental processes. In many human disorders, including different cancers and diabetes, they are also remarkably uncontrolled. For instance, PEA3 subfamily members Pea3, Erm, and Er81 are expressed in the nervous system from neural tube closure onward, particularly in the brain neuroepithelium and the neural crest derivatives as well as in the spinal cord, in addition to developing organs like the heart, kidney, and lungs in the embryo [12]. Elk-1 has been demonstrated to control PA-induced inflammation in photoreceptors in an animal model of Type 2 diabetes [83]. It's interesting to note that the ELK1 transcription factor has been identified as a target for depression, and that high levels of ELK1 expression have been linked to treatment resistance in postmortem samples from depression-induced suicides [84].

Here, in this thesis project, gene regulatory networks were represented for ETS proteins in five species during the development of the nervous system using transcriptome data. I demonstrate that ETS members exhibit significant functional redundancy, and no obvious subfamily-based labor of division is easily discernible. Individual subfamily members may have variable expression levels and regulatory networks, although PEA3, FLI, TCF, and TEL family members in particular seem to be actively regulated throughout neural development. ETV4 and ETV5 of the PEA3 subfamily share important regulatory targets with humans, however ETV1 appears to mostly have its own regulatory network (Figure 3.12). In the human dataset, the TCF subfamily member ELK3 shows varying degrees of expression during the course of the 18-day neuronal differentiation process; however, in the gorilla dataset, ELK3 is

shown to be strongly expressed in NPCs and downregulated in neurons. However, in mice, Elk-1 expression is initially low in NSCs, is elevated in neuroblasts, and then decreases again in neurons. In contrast, in *Drosophila*, TCF member Ets65A is low in neuroblasts and high in neurons. These findings show that each ETS member and its gene regulatory network have been adapted for a particular species.

It should be emphasized that there are important distinctions between the nervous systems of vertebrates and invertebrates. While invertebrates like flatworms and insects have ganglionic nervous systems that are ventrally structured, vertebrates' cerebrospinal neural systems are dorsally oriented. Lower vertebrate species like fish and birds have smooth cerebral hemispheres, but mammals' cerebrums have a lot of grooves, which considerably increases the surface area. It is believed that the vertebrate nervous system's evolutionary success is due to the growth of the cerebral cortex in mammals, which required substantial myelination of the axons for faster neurotransmission [85]. As a result, vertebrates' glial cells had to become more diverse. Therefore, there has been a lot of interest in the evolution of the transcriptional program for these many neurological systems. Numerous important genes, including FOXP2, have been discovered in particular regionalization of the human brain as a result of the majority of research on developmental regulatory genes concentrating on these genes [86],[87].

Because of the wide variation in domain structural design within ETS subfamilies, distinct ETS proteins interact with other proteins quite differently [88]. Although we have seen various, often conflicting roles across subfamily members of the ETS, we have not clearly defined a species-specific functional division of labor among them. Understanding and identifying the protein-protein interaction networks, collaboration with other transcription factors or signaling proteins, and how these interactions function in differentiation will also aid in understanding the gene regulatory networks of various ETS family members in the development of the nervous system. Clarifying how ETS proteins contribute to the development of cancer requires a full understanding of their developmental roles and activities.

Additionally, using transcriptomic comparison and experimental approaches, we have shown that RA differentiation of SH-SY5Y cells can be a useful model for primary neuron differentiation. The important proof of this study is that RA-differentiated SH-SY5Y cells have a very similar transcriptomic profile compared to primary neurons (Figure 3.26-27 and Appendix B-C). Literature studies using SH-

SY5Y cells also show that SH-SY5Y cells subjected to differentiation procedures reflect the neuronal phenotype and are used as a model *in vitro* cell system [89],[90]. As a result, in this project, it was determined that SH-SY5Y cells are a suitable cell line to be used as a neuronal differentiation model in further studies, and that ETV5, one of the ETS family members, is involved in this differentiation procedure.

The functional evolution of the ETS transcription factor superfamily in terms of neuronal differentiation utilizing the gene regulatory network inference technique is the subject of this study, which is the first comparative transcriptomic investigation in the literature. Overall, gene regulatory network comparisons showed that ETS members are evolutionarily conserved regarding the control of neuronal development, despite the absence of any appreciable subfamily specialization. The ETS, PEA3, and ELF subfamilies, however, were found to be considerably more important for nervous system development. To completely understand the evolution of subfamily roles within the ETS superfamily, more analogous investigations for the development of various tissues and systems should be carried out. As a future aspect, transcriptomic comparison of ETS regulation during neuronal differentiation should be investigated in single cell transcriptomic level in order to elucidate cell type specific developmental trajectory. Additionally, multi omics approaches should be implemented into this research question such as ATAC-seq and ChIP-seq for in depth investigation of ETS regulation with motif binding. Overall outputs of this thesis study were listed in Table 4.1.

Table 4.1: Published and submitted research papers.

Output	**Reference**
Gene Regulatory Network of ETS Domain Transcription Factors in Different Stages of Glioma	[7]
Multiple Sclerosis Biomarker Candidates Revealed by Cell-Type-Specific Interactome Analysis	[91]
Nervous System-Related Gene Regulatory Networks and Functional Evolution of ETS proteins Across Species	Babal et al., 2023 under review

REFERENCES

[1] Laudet V., Hanni C., Stehelin D., Duterque-Coquillaud M., (1999), "Molecular phylogeny of the ETS gene family", Oncogene, 18(6), 1351-1359.

[2] Oikawa T., Yamada T., (2003), "Molecular biology of the Ets family of transcription factors", Gene, 303, 11-34.

[3] Degnan B. M., Degnan S. M., Naganuma T., Morse D. E., (1993), "The Ets Multigene Family Is Conserved Throughout the Metazoa", Nucleic Acids Research, 21(15), 3479-3484.

[4] Remy P., Baltzinger M., (2000), "The Ets-transcription factor family in embryonic development: lessons from the amphibian and bird", Oncogene, 19(55), 6417-6431.

[5] Hollenhorst P. C., McIntosh L. P., Graves B. J., (2011), "Genomic and biochemical insights into the specificity of ETS transcription factors", Annu Rev Biochem, 80, 437-471.

[6] Boros J., O'Donnell A., Donaldson I. J., Kasza A., Zeef L., Sharrocks A. D., (2009), "Overlapping promoter targeting by Elk-1 and other divergent ETS-domain transcription factor family members", Nucleic Acids Res, 37(22), 7368-7380.

[7] Babal Y. K., Kandemir B., Kurnaz I. A., (2021), "Gene Regulatory Network of ETS Domain Transcription Factors in Different Stages of Glioma", J Pers Med, 11(2).

[8] Sogut M. S., Venugopal C., Kandemir B., Dag U., Mahendram S., Singh S., et al., (2021), "ETS-Domain Transcription Factor Elk-1 Regulates Stemness Genes in Brain Tumors and CD133+ BrainTumor-Initiating Cells", J Pers Med, 11(2).

[9] Saxton J., Ferjentsik Z., Ducker C., Johnson A. D., Shaw P. E., (2016), "Stepwise evolution of Elk-1 in early deuterostomes", FEBS J, 283(6), 1025-1038.

[10] Bocquet-Muchembled B., Leroux R., Chotteau-Lelievre A., Vergoten G., Fontaine F., (2002), "Expression and evolution studies of ets genes in a primitive coelomate, the polychaete annelid, Hediste (Nereis) diversicolor", Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 132(4), 685-697.

[11] Bartel F. O., Higuchi T., Spyropoulos D. D., (2000), "Mouse models in the study of the Ets family of transcription factors", Oncogene, 19(55), 6443-6454.

[12] Maroulakou I. G., Bowe D. B., (2000), "Expression and function of Ets transcription factors in mammalian development: a regulatory network", Oncogene, 19(55), 6432-6442.

[13] Ma H., Wang J., Wang B., Zhao Y., Yang C., (2009), “Characterization of an ETS transcription factor in the sea scallop Chlamys farreri”, Dev Comp Immunol, 33(9), 953-958.

[14] Gainous T. B., Wagner E., Levine M., (2015), “Diverse ETS transcription factors mediate FGF signaling in the Ciona anterior neural plate”, Developmental Biology, 399(2), 218-225.

[15] Schmid C., Schwarz V., Hutter H., (2006), “AST-1, a novel ETS-box transcription factor, controls axon guidance and pharynx development in C. elegans”, Dev Biol, 293(2), 403-413.

[16] Znosko W. A., Yu S., Thomas K., Molina G. A., Li C., Tsang W., et al., (2010), “Overlapping functions of Pea3 ETS transcription factors in FGF signaling during zebrafish development”, Dev Biol, 342(1), 11-25.

[17] Laurent-Huck F. M., Egles C., Kienlen P., Stoeckel M. E., Felix J. M., (1996), “Expression of the c-ets1 gene in the hypothalamus and pituitary during rat development”, Brain Res Dev Brain Res, 97(1), 107-117.

[18] Kawachi K., Masuyama N., Nishida E., (2003), “Essential role of the transcription factor ETS-2 in Xenopus early development”, Journal of Biological Chemistry, 278(7), 5473-5477.

[19] Zhu Y., Guthrie S., (2002), “Expression of the ETS transcription factor ER81 in the developing chick and mouse hindbrain”, Dev Dyn, 225(3), 365-368.

[20] Albagli O., Klaes A., Ferreira E., Leprince D., Klambt C., (1996), “Function of ets genes is conserved between vertebrates and Drosophila”, Mechanisms of Development, 59(1), 29-40.

[21] O'Leary D. A., Noakes P. G., Lavidis N. A., Kola I., Hertzog P. J., Ristevski S., (2007), “Targeting of the ETS factor GABPalpha disrupts neuromuscular junction synaptic function”, Mol Cell Biol, 27(9), 3470-3480.

[22] Ebeid M., Huh S. H., (2020), “Mesenchymal ETV transcription factors regulate cochlear length”, Hearing Research, 396.

[23] Talley M. J., Nardini D., Qin S., Prada C. E., Ehrman L. A., Waclaw R. R., (2021), “A role for sustained MAPK activity in the mouse ventral telencephalon”, Dev Biol, 476, 137-147.

[24] Raible F., Brand M., (2001), “Tight transcriptional control of the ETS domain factors Erm and Pea3 by Fgf signaling during early zebrafish development”, Mech Dev, 107(1-2), 105-117.

[25] Bosco A., Bureau C., Affaticati P., Gaspar P., Bally-Cuif L., Lillesaar C., (2013), “Development of hypothalamic serotoninergic neurons requires Fgf signalling via the ETS-domain transcription factor Etv5b”, Development, 140(2), 372-384.

[26] Iyo A. H., Porter B., Deneris E. S., Austin M. C., (2005), “Regional distribution and cellular localization of the ETS-domain transcription factor, FEV, mRNA in the human postmortem brain”, Synapse, 57(4), 223-228.

[27] Hendricks T., Francis N., Fyodorov D., Deneris E. S., (1999), “The ETS domain factor Pet-1 is an early and precise marker of central serotonin neurons and interacts with a conserved element in serotonergic genes”, Journal of Neuroscience, 19(23), 10348-10356.

[28] Weller M., Wick W., Aldape K., Brada M., Berger M., Pfister S. M., et al., (2015), “Glioma”, Nat Rev Dis Primers, 1, 15017.

[29] Hsu T., Trojanowska M., Watson D. K., (2004), “Ets proteins in biological control and cancer”, Journal of Cellular Biochemistry, 91(5), 896-903.

[30] Kar A., Gutierrez-Hartmann A., (2013), “Molecular mechanisms of ETS transcription factor-mediated tumorigenesis”, Crit Rev Biochem Mol Biol, 48(6), 522-543.

[31] Furlan A., Vercamer C., Heliot L., Wernert N., Desbiens X., Pourtier A., (2019), “Ets-1 drives breast cancer cell angiogenic potential and interactions between breast cancer and endothelial cells”, Int J Oncol, 54(1), 29-40.

[32] Verschoor M. L., Verschoor C. P., Singh G., (2013), “Ets-1 global gene expression profile reveals associations with metabolism and oxidative stress in ovarian and breast cancers”, Cancer Metab, 1(1), 17.

[33] Fry E. A., Inoue K., (2018), “Aberrant expression of ETS1 and ETS2 proteins in cancer”, Cancer Rep Rev, 2(3).

[34] Morris J. F., Sul J. Y., Kim M. S., Klein-Szanto A. J., Schochet T., Rustgi A., et al., (2013), “Elk-1 phosphorylated at threonine-417 is present in diverse cancers and correlates with differentiation grade of colonic adenocarcinoma”, Hum Pathol, 44(5), 766-776.

[35] Pardy L., Rosati R., Soave C., Huang Y., Kim S., Ratnam M., (2020), “The ternary complex factor protein ELK1 is an independent prognosticator of disease recurrence in prostate cancer”, Prostate, 80(2), 198-208.

[36] Ahmad A. H., A., (2019), “Expression of oncogenes ELK1 and ELK3 in cancer.”, Annal Colorec Cancer Res, 1, 1–6.

[37] Yue C. H., Huang C. Y., Tsai J. H., Hsu C. W., Hsieh Y. H., Lin H., et al., (2015), “MZF-1/Elk-1 Complex Binds to Protein Kinase Calpha Promoter and Is Involved in Hepatocellular Carcinoma”, PLoS One, 10(5), e0127420.

[38] Kherrouche Z., Monte D., Werkmeister E., Stoven L., De Launoit Y., Cortot A. B., et al., (2015), “PEA3 transcription factors are downstream effectors of Met

signaling involved in migration and invasiveness of Met-addicted tumor cells", Mol Oncol, 9(9), 1852-1867.

[39] Mesci A., Taeb S., Huang X., Jairath R., Sivaloganathan D., Liu S. K., (2014), "Pea3 expression promotes the invasive and metastatic potential of colorectal carcinoma", World J Gastroenterol, 20(46), 17376-17387.

[40] Qi T., Qu Q., Li G., Wang J., Zhu H., Yang Z., et al., (2020), "Function and regulation of the PEA3 subfamily of ETS transcription factors in cancer", Am J Cancer Res, 10(10), 3083-3105.

[41] Olar A., Sulman E. P., (2015), "Molecular Markers in Low-Grade Glioma-Toward Tumor Reclassification", Semin Radiat Oncol, 25(3), 155-163.

[42] Wood M. D., Halfpenny A. M., Moore S. R., (2019), "Applications of molecular neuro-oncology - a review of diffuse glioma integrated diagnosis and emerging molecular entities", Diagn Pathol, 14(1), 29.

[43] Mehta S., Lo Cascio C., (2018), "Developmentally regulated signaling pathways in glioma invasion", Cell Mol Life Sci, 75(3), 385-402.

[44] Zhao C., Gomez G. A., Zhao Y., Yang Y., Cao D., Lu J., et al., (2018), "ETV2 mediates endothelial transdifferentiation of glioblastoma", Signal Transduct Target Ther, 3, 4.

[45] Haber M. A., Iranmahboob A., Thomas C., Liu M., Najjar A., Zagzag D., (2015), "ERG is a novel and reliable marker for endothelial cells in central nervous system tumors", Clin Neuropathol, 34(3), 117-127.

[46] Xu Y., Gu S., Bi Y., Qi X., Yan Y., Lou M., (2018), "Transcription factor PU.1 is involved in the progression of glioma", Oncol Lett, 15(3), 3753-3759.

[47] Vyazunova I., Maklakova V. I., Berman S., De I., Steffen M. D., Hong W., et al., (2014), "Sleeping Beauty mouse models identify candidate genes involved in gliomagenesis", PLoS One, 9(11), e113489.

[48] Pan Y., Duron C., Bush E. C., Ma Y., Sims P. A., Gutmann D. H., et al., (2018), "Graph complexity analysis identifies an ETV5 tumor-specific network in human and murine low-grade glioma", PLoS One, 13(5), e0190001.

[49] Barthel F. P., Wesseling P., Verhaak R. G. W., (2018), "Reconstructing the molecular life history of gliomas", Acta Neuropathol, 135(5), 649-670.

[50] Torres A., Alshalalfa M., Tomlins S. A., Erho N., Gibb E. A., Chelliserry J., et al., (2017), "Comprehensive Determination of Prostate Tumor ETS Gene Status in Clinical Samples Using the CLIA Decipher Assay", J Mol Diagn, 19(3), 475-484.

[51] Barrett T., Wilhite S. E., Ledoux P., Evangelista C., Kim I. F., Tomashevsky M., et al., (2013), "NCBI GEO: archive for functional genomics data sets--update", Nucleic Acids Res, 41(Database issue), D991-995.

[52] Sun L., Hui A. M., Su Q., Vortmeyer A., Kotliarov Y., Pastorino S., et al., (2006), "Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain", Cancer Cell, 9(4), 287-300.

[53] Lopes F. M., da Motta L. L., De Bastiani M. A., Pfaffenseller B., Aguiar B. W., de Souza L. F., et al., (2017), "RA Differentiation Enhances Dopaminergic Features, Changes Redox Parameters, and Increases Dopamine Transporter Dependency in 6-Hydroxydopamine-Induced Neurotoxicity in SH-SY5Y Cells", Neurotox Res, 31(4), 545-559.

[54] Sauvageau M., Goff L. A., Lodato S., Bonev B., Groff A. F., Gerhardinger C., et al., (2013), "Multiple knockout mouse models reveal lincRNAs are required for life and brain development", Elife, 2, e01749.

[55] Baser A., Skabkin M., Kleber S., Dang Y., Gulculer Balta G. S., Kalamakis G., et al., (2019), "Onset of differentiation is post-transcriptionally controlled in adult neural stem cells", Nature, 566(7742), 100-104.

[56] Berger C., Harzer H., Burkard T. R., Steinmann J., van der Horst S., Laurenson A. S., et al., (2012), "FACS purification and transcriptome analysis of drosophila neural stem cells reveals a role for Klumpfuss in self-renewal", Cell Rep, 2(2), 407-418.

[57] Linker S. B., Narvaiza I., Hsu J. Y., Wang M., Qiu F., Mendes A. P. D., et al., (2022), "Human-specific regulation of neural maturation identified by cross-primate transcriptomics", Curr Biol, 32(22), 4797-4807 e4795.

[58] Von Stetina S. E., Watson J. D., Fox R. M., Olszewski K. L., Spencer W. C., Roy P. J., et al., (2007), "Cell-specific microarray profiling experiments reveal a comprehensive picture of gene expression in the C. elegans nervous system", Genome Biol, 8(7), R135.

[59] Afgan E., Baker D., Batut B., van den Beek M., Bouvier D., Cech M., et al., (2018), "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update", Nucleic Acids Res, 46(W1), W537-W544.

[60] Web 1, (2020), www.bioinformatics.babraham.ac.uk/projects/fastqc/, (Date Accessed: 05.06.2020)

[61] Web 2, (2020), www.bioinformatics.babraham.ac.uk/projects/trim_galore/, (Date Accessed: 01.06.2020)

[62] Kim D., Langmead B., Salzberg S. L., (2015), "HISAT: a fast spliced aligner with low memory requirements", Nat Methods, 12(4), 357-360.

[63] Liao Y., Smyth G. K., Shi W., (2014), "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features", Bioinformatics, 30(7), 923-930.

[64] Love M. I., Huber W., Anders S., (2014), "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2", Genome Biol, 15(12), 550.

[65] Web 3, (2021), https://github.com/kevinblighe/EnhancedVolcano, (Date Accessed: 08.06.2021)

[66] Davis S., Meltzer P. S., (2007), "GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor", Bioinformatics, 23(14), 1846-1847.

[67] Ritchie M. E., Phipson B., Wu D., Hu Y., Law C. W., Shi W., et al., (2015), "limma powers differential expression analyses for RNA-sequencing and microarray studies", Nucleic Acids Res, 43(7), e47.

[68] Wu T., Hu E., Xu S., Chen M., Guo P., Dai Z., et al., (2021), "clusterProfiler 4.0: A universal enrichment tool for interpreting omics data", Innovation (N Y), 2(3), 100141.

[69] Castro M. A., de Santiago I., Campbell T. M., Vaughn C., Hickey T. E., Ross E., et al., (2016), "Regulators of genetic risk of breast cancer identified by integrative network analysis", Nat Genet, 48(1), 12-21.

[70] Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., Ramage D., et al., (2003), "Cytoscape: a software environment for integrated models of biomolecular interaction networks", Genome Res, 13(11), 2498-2504.

[71] Kandemir B., Dag U., Bakir Gungor B., Durasi I. M., Erdogan B., Sahin E., et al., (2017), "In silico analyses and global transcriptional profiling reveal novel putative targets for Pea3 transcription factor related to its function in neurons", PLoS One, 12(2), e0170585.

[72] Kumar S., Stecher G., Suleski M., Hedges S. B., (2017), "TimeTree: A Resource for Timelines, Timetrees, and Divergence Times", Mol Biol Evol, 34(7), 1812-1819.

[73] Xu H., Zhao G., Zhang Y., Jiang H., Wang W., Zhao D., et al., (2019), "Long non-coding RNA PAXIP1-AS1 facilitates cell invasion and angiogenesis of glioma by recruiting transcription factor ETS1 to upregulate KIF14 expression", J Exp Clin Cancer Res, 38(1), 486.

[74] Cam M., Charan M., Welker A. M., Dravid P., Studebaker A. W., Leonard J. R., et al., (2020), "DeltaNp73/ETS2 complex drives glioblastoma pathogenesis-targeting downstream mediators by rebastinib prolongs survival in preclinical models of glioblastoma", Neuro Oncol, 22(3), 345-356.

[75] Giese A., (2003), “Glioma invasion--pattern of dissemination by mechanisms of invasion and surgical intervention, pattern of gene expression and its regulatory control by tumorsuppressor p53 and proto-oncogene ETS-1”, Acta Neurochir Suppl, 88, 153-162.

[76] Mut M., Lule S., Demir O., Kurnaz I. A., Vural I., (2012), “Both mitogen-activated protein kinase (MAPK)/extracellular-signal-regulated kinases (ERK) 1/2 and phosphatidylinositide-3-OH kinase (PI3K)/Akt pathways regulate activation of E-twenty-six (ETS)-like transcription factor 1 (Elk-1) in U138 glioblastoma cells”, Int J Biochem Cell Biol, 44(2), 302-310.

[77] Wang Z., Yuan H., Sun C., Xu L., Chen Y., Zhu Q., et al., (2015), “GATA2 promotes glioma progression through EGFR/ERK/Elk-1 pathway”, Med Oncol, 32(4), 87.

[78] Cheng M., Zeng Y., Zhang T., Xu M., Li Z., Wu Y., (2021), “Transcription Factor ELF1 Activates MEIS1 Transcription and Then Regulates the GFI1/FBW7 Axis to Promote the Development of Glioma”, Mol Ther Nucleic Acids, 23, 418-430.

[79] Wang M., Yang C., Liu X., Zheng J., Xue Y., Ruan X., et al., (2020), “An upstream open reading frame regulates vasculogenic mimicry of glioma via ZNRD1-AS1/miR-499a-5p/ELF1/EMI1 pathway”, J Cell Mol Med, 24(11), 6120-6136.

[80] Padul V., Epari S., Moiyadi A., Shetty P., Shirsat N. V., (2015), “ETV/Pea3 family transcription factor-encoding genes are overexpressed in CIC-mutant oligodendrogliomas”, Genes Chromosomes Cancer, 54(12), 725-733.

[81] Gatalica Z., Xiu J., Swensen J., Vranic S., (2019), “Molecular characterization of cancers with NTRK gene fusions”, Mod Pathol, 32(1), 147-153.

[82] Busse T. M., Roth J. J., Wilmoth D., Wainwright L., Tooke L., Biegel J. A., (2017), “Copy number alterations determined by single nucleotide polymorphism array testing in the clinical laboratory are indicative of gene fusions in pediatric cancer patients”, Genes Chromosomes Cancer, 56(10), 730-749.

[83] Yu F., Ko M. L., Ko G. Y., (2021), “MicroRNA-150 and its target ETS-domain transcription factor 1 contribute to inflammation in diabetic photoreceptors”, J Cell Mol Med, 25(22), 10724-10735.

[84] Apazoglou K., Farley S., Gorgievski V., Belzeaux R., Lopez J. P., Grenier J., et al., (2018), “Antidepressive effects of targeting ELK-1 signal transduction”, Nature Medicine, 24(5), 591-+.

[85] Luisetto M. I., G.; Iatyschev, O.; Akram, M., (2019), “The evolution of the nervous system: invertebrates vs vertebrates as a useful instrument and model to research new pharmacological strategies in some human neurodegenerative conditions.”, Am J Biomed Sci & Res, 5(5).

[86] Arendt D., (2005), “Genes and homology in nervous system evolution: Comparing gene functions, expression patterns, and cell type molecular fingerprints”, Theory in Biosciences, 124(2), 185-197.

[87] Sousa A. M. M., Meyer K. A., Santpere G., Gulden F. O., Sestan N., (2017), “Evolution of the Human Nervous System Function, Structure, and Development”, Cell, 170(2), 226-247.

[88] Li R., Pei H., Watson D. K., (2000), “Regulation of Ets function by protein - protein interactions”, Oncogene, 19(55), 6514-6523.

[89] Nguyen K., Sanchez C. L., Brammer-Robbins E., Pena-Delgado C., Kroyter N., El Ahmadie N., et al., (2022), “Neurotoxicity assessment of QoI strobilurin fungicides azoxystrobin and trifloxystrobin in human SH-SY5Y neuroblastoma cells: Insights from lipidomics and mitochondrial bioenergetics”, Neurotoxicology, 91, 290-304.

[90] Zhao H., Xie J., Wu S., Sanchez O. F., Zhang X., Freeman J. L., et al., (2022), “Pre-differentiation exposure of PFOA induced persistent changes in DNA methylation and mitochondrial morphology in human dopaminergic-like neurons”, Environ Pollut, 308, 119684.

[91] Yurduseven K., Babal Y. K., Celik E., Kerman B. E., Kurnaz I. A., (2022), “Multiple Sclerosis Biomarker Candidates Revealed by Cell-Type-Specific Interactome Analysis”, OMICS, 26(5), 305-317.

BIOGRAPHY

Yigit Koray Babal achieved his bachelor's degree from Department of Biology, Marmara University in 2016. He has continued master's degree program in the Department of Molecular Biology and Genetics at Molecular Neurobiology Laboratory, Graduate School of Natural and Applied Sciences, Gebze Technical University, and graduated in 2018. Then he has continued Doctor of Philosophy degree in Institute of Biotechnology in Gebze Technical University since fall 2018. He also participated Brain Awareness Week activities, which is a funded project by FENS, for two years. He would like to study on computational neurobiology and bioinformatics.

APPENDICES

Appendix A: Expression level ETS members in epilepsy

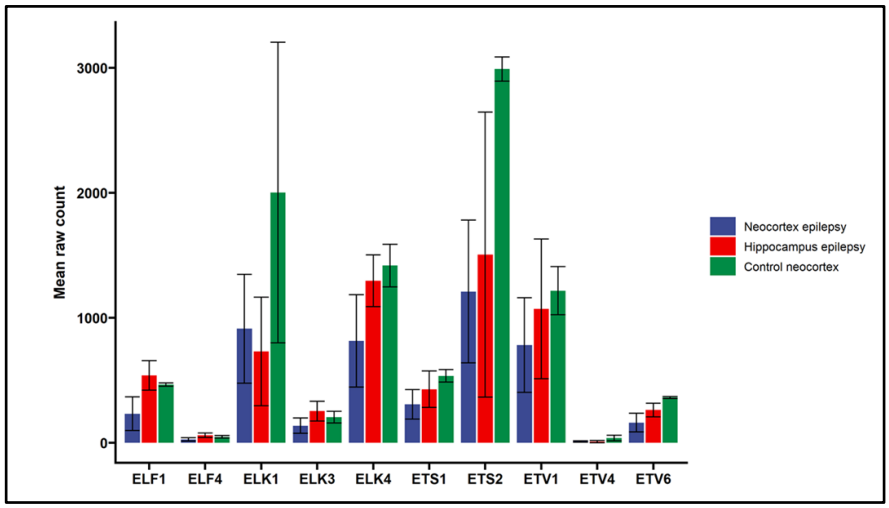


Figure A1.1: Expression level of ETS members on neocortex and hippocampus in epilepsy patients and healthy control from GSE134697.

Appendix B: Average expression comparison of SH-SY5Y cells and dopaminergic neuron

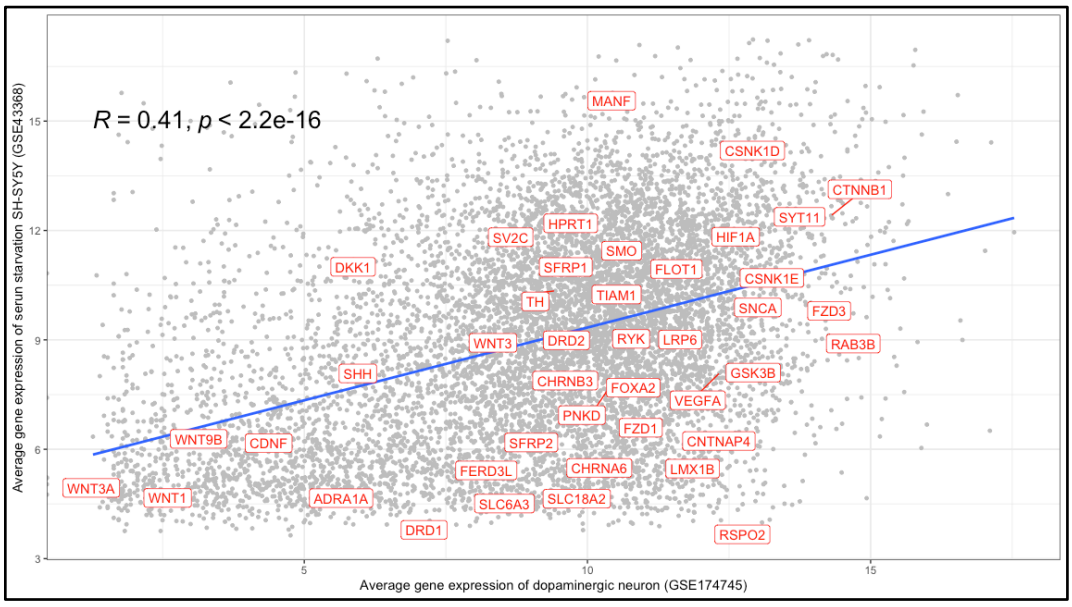


Figure B1.1: Correlation of average gene expression of serum starvation SH-SY5Y cells and dopaminergic neuron.

Appendix C: Transcriptomic comparison of RA differentiated SH-SY5Y cells and dopaminergic neurons.

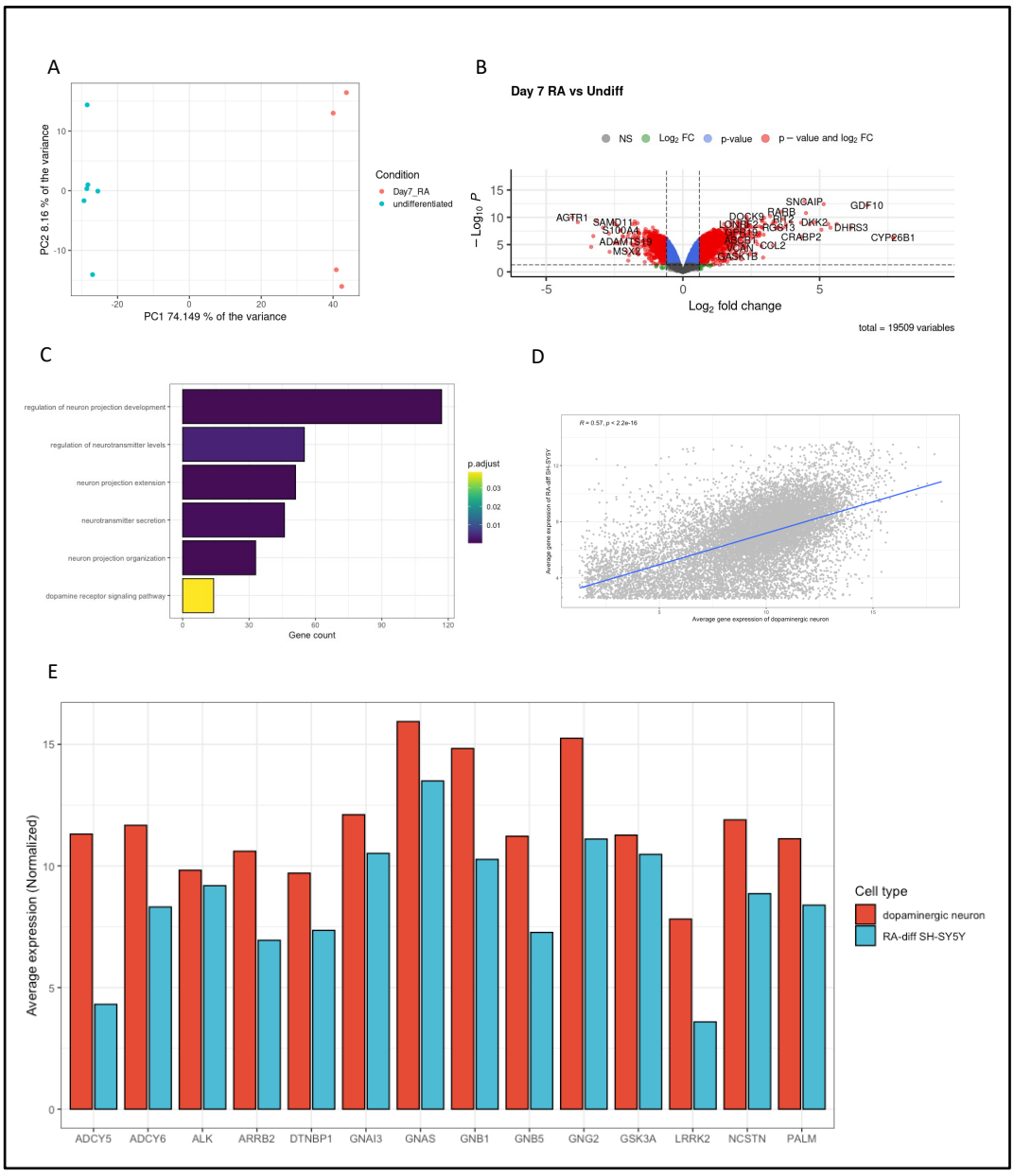


Figure C1.1: Transcriptomic comparison of RA differentiated SH-SY5Y cells and dopaminergic neurons.