2021-7

Investigating the role of Schizophrenia-associated gene expression in the developing human brain using Machine Learning

Katie Kelly
Technological University Dublin, x00120104@mytudublin.ie

Recommended Citation

Kelly, K. (2021). Investigating the role of Schizophrenia-associated gene expression in the developing human brain using Machine Learning. Technological University Dublin. DOI: 10.21427/R4FM-9950

Investigating the role of Schizophrenia-associated gene expression in the developing human brain using Machine Learning

Katie Kelly

Technological University Dublin

M.Sc **2021**

Sample Spine of Thesis

M.Sc.	**Katie Kelly**	**2021**

Investigating the role of schizophrenia-associated gene expression in the developing human brain using Machine Learning

A Thesis Presented for the Award of Masters by Research

Katie Kelly (B.Sc.)

Technological University Dublin – Tallaght Campus

Department of Science

For Research Carried Out Under the Guidance of

Dr Eugene Hickey & Dr Therese Murphy

Submitted to Technological University Dublin

July 2021

DECLARATION

I certify that this thesis which I now submit for examination for the award of _____Masters_____________ is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for graduate study by research of the Technological University Dublin (TU Dublin) and has not been submitted in whole or in part for another award in any other third-level institution.

The work reported in this thesis conforms to the principles and requirements of the TU Dublin's guidelines for ethics in research.

(The following sentence may be deleted if access to the thesis is restricted according to Section 4.8 of the TU Dublin Research regulations)

TU Dublin has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis is duly acknowledged.

Signature ________Katie Kelly_________________ Date _____12/07/21_____

Candidate

Acknowledgements

I would like to thank my supervisor’s Dr Eugene Hickey & Dr Therese Murphy who have been supportive and encouraging throughout this project. I would like to thank Eugene for his positive outlook and assurance which gave me confidence in my research which inspired me to keep going.

I would also like to thank Lorcán Dooley for being a wonderful distraction and a huge supporter. I love you.

Dedicated to Jim and Siobhan Kelly who have always had faith that I could do anything I put my mind to. I love you both.

Abstract

Schizophrenia is a debilitating condition that affects 1% of the population, causes significant hardship and though there are treatments available they are characterised by several limitations. It is a complex mental disorder where some individuals show mild subclinical cognitive symptoms before psychosis onset in adolescence. The treatments available only target a portion of the symptoms and although extensive research has been conducted, a comprehensive understanding of the nature of schizophrenia remains elusive. Unlike other neurodevelopmental disorders, schizophrenia symptoms do not typically present themselves until adolescence. This study aimed to discover gene co-expression networks at multiple developmental stages to identify candidate therapeutic targets to better treat and manage schizophrenia.

Recent genome-wide association studies have identified 145 genetic loci associated with schizophrenia. Allen Brain Atlas's BrainSpan resource provides brain development data from neurotypical brains. Using this resource it was possible to study the gene expression of 316 schizophrenia-associated genes, identified previously in a large-scale GWAS, across each of the developmental stages available in the Allen Brain Atlas. K means Clustering and a systems biology approach (WGCNA) was applied to these schizophrenia-associated genes at each developmental stage where modules within networks were created by grouping co-expressed genes. To facilitate biological interpretation of these modules co-expressed genes were visualised using Cytoscape and gene ontology pathway enrichment analysis was applied.

We identified 21 hub genes using WGCNA. Of the 316 schizophrenia-associated genes, 27 modules were identified and 3 hub genes *GPR52, INA, SATB2* were common in multiple developmental stages. Our results suggest that *GPR52, INA, SATB2* represent candidate genes for future evaluation of their potential as therapeutic targets of schizophrenia. Additional hub genes included *TRANK1* and *ALMS1*, genes which were previously identified as expression quantitative trait loci. Taken together our results add further evidence that these genes could be good candidates for further research as they may regulate several schizophrenia-related genes in their respective modules. Finally, our enrichment analysis implicated a role for positive regulation of macrophage proliferation and cellular response to catecholamine stimulus, and cellular response to diacyl bacterial lipopeptide at each developmental stage. The immune system and catecholamines, including dopamine, have long been associated with schizophrenia and our results provide further support for these hypotheses.

List of Abbreviations

5HT - 5- hydroxytryptamine receptors
ABA - Allen Brain Atlas
AHBA - Allen Human brain Atlas
ADHD - Attention-Deficit Hyperactivity Disorder
ANPs - Antipsychotic Naïve Patients
ASD - autism spectrum disorder
BP - bipolar disorder
CNVs - Copy Number Variants
DMAs - Dopamine modulating antipsychotics
eQTL - Expression quantitative trait loci
FDR - False Discovery Rate
GABA - Gamma aminobutyric acid
GCN - Gene co-expression networks
GMV - Grey Matter volume
GO - Gene Ontology
GS - Gene significance
GWAS - Genome Wide Association Study
HCs - Healthy controls
HG - Hub Genes
IBD - Inflammatory bowel disease
ID - Intellectual disability
IL1β - Interleukin 1 beta
IL6 - Interleukin 6
INDELS - Insertions and deletions
ISH - In Situ Hybridisation
LD - Linkage disequilibrium
LoF - Loss of Function
LMD - Laser Microdissection
LSD - Lysergic acid diethylamide
ME - Module eigengene
MIA - Maternal Immune activation
MK-801 - Dizocilpine
MM - Module membership
MNI - Montreal Neurological Institute
MS - Module significance
MRI - Magnetic Resonance Imaging
mRNA - Messenger ribonucleic acid
NS - node significance
NMDAR - N-methyl-D-aspartate receptors
OTU - Operational Taxonomic Units
PCA - Principal component analysis
PCP - Phenylcyclidine
PEN - Polyethylene naphthalene
PFC - Prefrontal cortex
PNNs - perineuronal nets
PV - Parvalbumin
QC - Quality control

RNA	-	Ribonucleic acid
SNPs	-	Single Nucleotide Polymorphisms
TO	-	Topological Overlap
TOM	-	Topological Overlap Matrix
TNF-α	-	Tumour necrosis factor alpha
TNF-β	-	Tumour necrosis factor beta
UHR	-	Ultra high risk
VTA	-	Venteral Tegmental Area
WES	-	Whole Exome Sequencing
WGS	-	Whole Genome Sequencing
WGCNA	-	Weighted Correlation Network Analysis

Table of Contents

Chapter 1 – Introduction to schizophrenia

Schizophrenia is a debilitating psychiatric condition that manifests itself early in adolescence and can last a lifetime. It has a 1% global prevalence and comes with significant societal and economic costs as well as substantial mortality and morbidity (1). Its complex nature is believed to originate from a mixture of genetic and environmental factors including prenatal exposure to infection and lack of nutrients which cause disruptions during early brain development *in utero* (2,3). There are many hypotheses about the causes of schizophrenia, but the aetiology is unknown. This lack of understanding of schizophrenia is evident in its treatments which haven't significantly advanced since the introduction of the first-generation antipsychotic medication such as Chlorpromazine in the 1950s (4). This complex condition presents with three modes of clinical features namely: positive symptoms (psychoses manifesting as delusions and hallucinations, paranoia, hyperactivity and agitation), negative symptoms (social withdrawal, lack of motivation, asociality, avolition, affective flattening, consummatory and anticipatory anhedonia, and alogia), and cognitive symptoms (trouble with critical thinking, working memory and difficulty integrating feelings, thoughts and behaviour, attention and vigilance, verbal learning, reasoning and problem solving, and social cognition) as well as motor disturbances which regularly results in a poor quality of life (5–7). The presentation of symptoms is heterogeneous which makes schizophrenia both difficult to diagnose and treat (8). The negative and cognitive symptoms are chronic and are closely related to functional outcomes, and contribute greatly to illness burden, (6) the positive symptoms usually relapse and remit (8). Despite some progress in the understanding of several of the fundamental mechanisms involved in schizophrenia's aetiology, the current treatments available come with serious side effects, inconsistent efficacy, and lack of evidence that they substantially improves the outcomes (9,10). At present, schizophrenia’s treatments consist of antipsychotic drugs, social support, rehabilitation, and psychological therapies (8). Current antipsychotics are associated with serious limitations. Firstly, around 30% of sufferers are treatment-resistant, secondly, they mainly ameliorate positive symptoms only leaving cognitive and negative symptoms untreated and lastly, antipsychotics trigger both neurological and metabolic side effects (11). As a result, there is a clear need for more efficient and effective treatments as well as uncovering a model for prediction of efficacy as currently determining the most effective treatment of schizophrenia is a trial and error method (7,12). It is important to study patients at several clinical stages to give insight into the effects of schizophrenia itself, its progression and what alterations are caused by the pharmacological treatments (13).

1.1 Schizophrenia’s aetiology

At present, there are no clinical diagnostic tests available for schizophrenia so diagnosis relies on clinical observations and self-report (14). Schizophrenia remains incurable and the best outcome continues to be managing symptoms and preservation of independence and functionality (9). Until there is a more complete understanding of schizophrenias aetiology, there is little hope for improving diagnosis, predicting susceptibility, management, and treatments for those with schizophrenia.

1.2 Understanding of schizophrenia to date

Schizophrenia’s complicated and unknown underlying mechanisms has meant that there has been no fundamental innovation in schizophrenia treatments since the introduction of first-generation antipsychotics in the 1950s (15). As treatments mainly target the positive symptoms there is a clear need for a focus on cognitive and negative symptom domains. These types of studies could lead to new endophenotypic markers which could promote novel treatment discovery and could initiate concurrent medication strategies with current antipsychotics (15).

1.2.1 Immune System

Schizophrenias pathogenesis is elusive, and though animal models have been used to understand elements, the human central nervous system (CNS) and immune system are much more complex and intricate (16). Both systems share common features in developmental mechanisms, so therefore CNS and immune system dysregulation should be studied in humans (16,17). The immune hypothesis of schizophrenia has been around for a long time and is supported by epidemiological, genetic, imaging and biomarker studies (17). The accumulating evidence that anti-inflammatory and immunosuppressive medications are effective treatments and that autoimmune conditions and immune activation are risk factors for developing schizophrenia provides perhaps the most convincing evidence of the immunes system involvement (18,19).

Dysregulation of the innate and adaptive immune system has been identified by epidemiological, genetic, postmortem and therapeutic studies and are likely to contribute to some of the symptoms of schizophrenia (20). Though there have been a large number of studies with significant funding devoted to better understand schizophrenia outcomes remain poor and hope remains in advances in psychoneuroimmunology and other advanced technological research areas to provide more consistent and successful management of schizophrenia (12). Several

autoimmune conditions display neuropsychiatric symptoms suspected to be caused by brain reactive antibodies (21). Schizophrenia and autoimmune diseases are often comorbid likely because of some genetic overlap, (22) affecting common underlying pathways which entail inflammatory immune response antibodies which can attack brain tissue (23)(22). A national cohort found if a patient had a prior autoimmune disease they are 29% more likely to develop schizophrenia in adolescence (21).
Maternal Immune Activation (MIA) can disrupt normal fetal brain development and has been linked to schizophrenia for over a century, it is estimated that if MIA could be avoided that 30% of schizophrenia cases would be averted (20). Lower levels of acute-phase proteins in neonates which increases the susceptibility of infection have also been hypothesized to increase the risk of psychosis in adulthood (22). Patients experiencing acute episodes of schizophrenia often have increased levels of Interleukin-1- beta (IL-1 β), Interleukin-6 (IL-6), and transforming growth factor-beta (TGF- β) (23). In unmedicated patients, tumour necrosis factor-alpha (TNF α) protein levels and *IL-1β* messenger RNA (mRNA) is seen to be elevated (24).

Within the body exists a dynamic population of gut microbes which houses many bacteria approximately 10^{14} cells. The biological biodiversity is established in the first couple of months of existence, and has a continuous role throughout life, and is very susceptible to environmental factors (25). The gut microbiome can control how the brain behaves and functions via the microbiota-gut-brain (MGB) and it has been reported to be related to changes in cognition, anxiety, and memory, as well as development, maturation of immune, neural and endocrine systems in animal models (26). These physiological and behavioural processes are often impaired in people with schizophrenia. A high α-diversity score is usually a sign of good health (27). In a study performed by Zheng et al. medicated and unmedicated patients with schizophrenia, it has been observed that they have a decreased α-diversity in their microbiome when compared to healthy controls (HCs) (26). It was also found that Veillonellaceae and Lachnospiraceae found in the microbiome environment were associated with symptom severity in schizophrenia (26). β- diversity analysis of schizophrenia patients and HCs found clear differences in the compositions of each microbiome by looking at operational taxonomic units (OTU) levels (26). In one study when the linear discriminating analysis effect size was applied to 77 differential OTUs it was observed that 23 out of the 77 OTUs saw an increase in patients with schizophrenia patients when compared to controls. The OTUs belonged to the bacterial families Veillonellaceae, Coriobacteriaceae, Bacterioidaceae, and Prevotellaceae, the other 54 OTU levels were seen to be decreased in patients with schizophrenia (Lachnospiraceae, Norank, Ruminococcaceae and Enterobacteriaceae). (26)

1.2.2 Neurodevelopmental hypothesis

Epidemiological, basic, and clinical neuroscience research has presented evidence that schizophrenia is of neurodevelopmental origin (28). This hypothesis is now widely accepted but what differentiates schizophrenia from other neurodevelopmental conditions is its time of onset, in adolescence (29). Autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD) and intellectual disabilities (ID) characteristically present themselves much earlier in childhood (29). Schizophrenia shares many phenotypic and clinical similarities and is often comorbid with these neurodevelopmental disorders but because of its delayed presentation, they were not initially linked (29). Before they were connected it was then hypothesised that schizophrenia may be a neurodegenerative disorder, but when post-mortem studies failed to identify traumatic, neurotoxic, or neurodegenerative mechanisms in the brain this theory was disproven, and the neurodevelopmental hypothesis replaced it (29). In neonatal primates and rodents, prenatal cortical lesions were shown to lead to the emergence of abnormalities that mimicked schizophrenia in early adolescence, proving that early developmental abnormalities could have an impact on cortical function in later life, making the neurodevelopmental hypotheses for schizophrenia plausible (30). Instead of each neurodevelopmental disorder being viewed independently an alternative view was proposed, that these neurodevelopmental disorders lie on an etiological continuum with a diverse range of outcomes that follows from early brain development disturbances because of shared genetic variants and environmental factors (29). These neurodevelopmental disorders are diagnosed based on symptoms, the timing of onset, severity/persistence, and abnormal brain development (29). Before the first psychotic episode, schizophrenia presents itself very similarly to the other neurodevelopmental disorders, but only the negative and cognitive symptoms. There are several rare copy number variants (CNVs), genes affected by loss of function (LoF) mutations, genes enriched with *3* nonsynonymous mutation and alleles that have significant associations with schizophrenia, ASD, ADHD and ID which represent direct outcomes of the rare pathogenic mutations that they share (29). This would also suggest that the risk of developing positive symptoms is not mediated by cognitive impairment.

1.2.3 Dopamine hypothesis

Dopamine is a catecholamine neurotransmitter in the brain which regulates critical neurological processes such as cognition, motor control, reward and learning (31). In the 1950s Chlorpromazine an antipsychotic drug and affective antagonist for the D2, D3, and D5 receptors was released and the treatments dopamine receptor antagonists have remained the most prevalent therapeutic (11). Chlorpromazine controlled the positive symptoms of schizophrenia patients and the theory that dopamine alterations in the mesolimbic pathway caused positive symptoms was strengthened and confirmed (11). Other key evidence supporting the dopamine hypothesis was when amphetamines were administered which increase the extracellular concentrations of dopamine and psychotic symptoms like schizophrenias appeared (32). This evidence was reinforced when treatments that depleted the concentration of dopamine such as alpha-methyl-para-tyrosine and reserpine were shown to reduce psychotic symptoms (32). These antipsychotics target other dopamine receptors, serotonin, norepinephrine, acetylcholine and histamine as well (32). It is often seen has that in the associative striatum there is an increased dopamine synthesis capacity for people who have psychotic disorders including schizophrenia (33). The increased dopamine synthesis is detectable in ultra-high-risk (UHR) subjects and before early symptoms of people who eventually develop schizophrenia thus are not a consequence of antipsychotic exposure or psychotic episodes (33). Rodent models have been able to replicate this, these models have also shown that increased synthesis and release of striatal dopamine can be a result of acute stressors and inflammatory challenges *in utero* (34). These developmental disruptions cause the dopamine system to become hyper-responsive later in life, in the rodent equivalent of adolescence (34). Recent studies have pinned that part of the cause for schizophrenia to be a combination of an increased spontaneous dopamine release and decreased dopamine release for relevant stimuli (34). Studies using amphetamines have been important for proving this. At moderate doses, amphetamines act as a reward predicting cue by increasing the levels of striatal dopamine appropriately while at larger doses, the amphetamines blunt adaptive responses, which alters the behavioural response and increases the spontaneous transients (spikes in the levels of dopamine) (35). These spontaneous transients may explain the inappropriate phasic firing of dopamine neurons known to be part of schizophrenia (35). All psychostimulants including amphetamines have the effect of increasing spontaneous transients in the striatum which correlate and could explain some of the positive symptoms of schizophrenia. Some of the primary negative symptoms of schizophrenia could be explained by the decreased adaptive transients in the striatum. (35) It is thought that many antipsychotic drugs perform in the same manner and affect the adaptive and spontaneous transients similarly, where one cannot be fixed without aggravating the other (34).

Around 30% of patients with schizophrenia do not respond to antipsychotics with high D2 occupancy and do not respond to treatments that diminish the levels of presynaptic dopamine concentrations. (32,36) Demjaha et.al found that people who responded to typical antipsychotic treatment had higher dopamine synthesis capacity and that increased synaptic dopamine may be used to predict treatment responsiveness. (36). Treatment-resistant patients did not have this capacity, this demonstrated that there may be a subtype of schizophrenia which is non-dopaminergic. Accumulating evidence has shown that schizophrenia’s core pathophysiology may also involve dysfunction in glutamatergic, serotonergic and gamma-aminobutyric acid (GABA). (37)

1.2.4 Glutamate hypothesis

The dopamine hypothesis can account for a portion of the psychopathology of schizophrenia, in particular positive symptoms (38). Atypical antipsychotic drugs apart from Clozapine have little to no effect on negative and cognitive symptoms (11). Negative and cognitive symptoms are neglected by antipsychotics and persist causing chronic disability (4). In patients with chronic schizophrenia cortical atrophy correlates with the negative and cognitive symptoms but not with the severity of the psychosis, (39) showing that although some of the cognitive and negative symptoms may be caused by dysregulation in dopamine pathways, not all are.

Glutamatergic pathways are primarily the excitatory neurotransmitters in the brain and glutamatergic neurons utilise between 60-80% of the total brain metabolic activity (32). Glutamate pathways have been linked to the limbic system, cortex, thalamus and are mediated by N-methyl-D-aspartate receptors (NMDARs) (37). Glutamate was originally associated with schizophrenia because it was observed that there were decreased levels of glutamate in cerebrospinal fluid (CSF) of patients with schizophrenia (32). There is now mounting evidence that glutamatergic dysregulation in the prefrontal cortex causes dopamine hyperactivity in the ventral tegmental area (VTA) which causes auditory hallucinations and paranoid delusions (40). Studies using NMDAR antagonists (Ketamine and phencyclidine (PCP), dizocilpine (MK-801)) on HCs induce schizophrenia-like symptoms (negative and cognitive) and increased prefrontal glutamine levels, these can last up to two weeks (41). PCP and Ketamine are non-competitive antagonists that bind at the NMDA subtype of glutamate receptor (39,42). From observing the effects of the NMDAR antagonists on healthy individuals, it has been proposed that certain symptoms of schizophrenia may result from the hypofunction of NMDAR (43). It has also been observed that patients with schizophrenia undergoing long term treatment have increased levels of glutamine in the anterior dorsal cingulate cortex which was linked with the severity of psychotic symptoms (41), suggesting that despite the treatment with antipsychotic treatments there is a basal increase of presynaptic glutamate which is consistent with the NMDAR hypofunction pathophysiological model of schizophrenia (41). It has been seen in patients with schizophrenia that increased synaptic release of glutamine is associated with psychosis, while glutamate metabolism is related to cognitive impairments (41). In one metanalysis it was observed that glutamate in the frontal region was lower but glutamine is higher in people with schizophrenia when compared with controls, over time the levels of both reduce which could suggest a progressive load of synaptic activity (44). Patients with schizophrenia who don’t respond to typical antipsychotic treatment seem to have more marked glutamatergic abnormalities while treatment responders have dopaminergic abnormalities (36).

The role of glutamate in the pathophysiology of schizophrenia has been investigated in Genome-wide association studies (GWAS) they have highlighted several genes associated with glutamatergic neurotransmission or with downstream mediators (*GRM3, GRIN2A, and GRIA1*) (32).

1.2.5 Gamma-aminobutyric Acid (GABA)

GABA is a major inhibitory neurotransmitter located in the CNS (37). Results from animal models and postmortem studies suggest that part of schizophrenias pathophysiology is caused by both dysfunctions of GABAergic interneurons and

NMDARs(45,46). In human postmortem studies in individuals with schizophrenia, alterations were seen in GABA-related epigenetic, transcript, synaptic, and protein markers especially evident was the subpopulation of GABA neurons which encompass calcium-binding protein parvalbumin (PV) (47). GABA interneurons are an important part of the brains rhythm generating network, they are also important in controlling neural oscillations which are fundamental mechanisms for memory, perception and consciousness (5). The third layer of the prefrontal cortex houses a microcircuit where GABAergic PV cells and glutamatergic cells synchronize neural oscillations (11). The PFC PV neurons have lower levels of PV and proteins and a GABA synthesizing enzyme GAD67 (45). These structural and molecular alterations are hypothesised to alter GABA neurotransmission and weaken the PFC gamma oscillations in people with schizophrenia (47). GABA antagonists have been shown to effective in improving some of the core symptoms of schizophrenia in clinical studies (48). Benzodiazepine which works on the GABA-A receptor allosteric site is used often with antipsychotic medications to treat schizophrenia (5).

1.2.6 Serotonin Hypothesis (5-hydroxytryptamine, 5-HT)

Although the serotonin hypotheses are one of the oldest in regards to schizophrenias pathogenesis it remains highly topical because of the lack of reproducible results (49). Serotonin has been linked to schizophrenia’s pathophysiology since studies looking at the interaction between 5-HT and the hallucinogenic drug lysergic acid diethylamide (LSD) which resulted from antagonism of 5-HT in the CNS (37,50). Psychotic symptoms due to dementia and Parkinson’s are successfully treated with 5HT2A antagonists without D2 antagonism which halts excess serotonin being released which stops the downstream release of glutamate which can activate the mesolimbic dopamine pathway (40). Sizeable evidence from multiple methods suggests that a subpopulation of patients with schizophrenia display serotonergic function abnormalities (49). It is believed that 5-HT receptors (5-HT$_{1A}$, 5-HT$_{2A}$, 5-HT$_{2C}$) may underlie cognitive symptoms and motivational disabilities and this shown by atypical antipsychotics which act on 5-HT$_{1A}$ and 5-HT$_{2A}$ receptors and how they ameliorate negative symptoms and mood disturbances (51). In mutant mice who display decreased willingness to work for a reward, there were D$_2$ receptors overexpression and up-regulation of 5-HT$_{2C}$ receptors in the striatum (51). The extrapyramidal effects of antipsychotics can be ameliorated by serotonin antagonists, (37) though the pathogenesis of schizophrenia has not been explicitly linked with serotonergic dysfunction, 5HT-3 and 5HT-6 continue to be looked into as potential therapeutic targets.

1.3 Environmental risks for schizophrenia

Many epidemiological studies have investigated the impact of the environment on the development of schizophrenia. Several risk factors have been highlighted, such as being raised in an urban environment, early life adversities, and the use of cannabis (especially compounds with high tetrahydrocannabinol levels) early and frequently which has an impact on the developing social brain (2,3,8). Immigration (first and second-generation) has also been studied as well as an increase in the rate of incidence in individuals with young parents or with relatively old parents. (2,8,52,53). Prenatal exposure to infection, preterm births, social disadvantage, and lack of nutrients in the womb have all been linked to heightening the risk of developing schizophrenia (28,54). Prenatal stress increases the basal secretion of glucocorticoid hormones which can reprogram the hypothalamic-pituitary-adrenal axis (3). In rodent models, following prenatal stress malformations in the DNA methylation in GABAergic neurons were observed and connected to schizophrenia-like symptoms (3). Exposure to prenatal infection has been shown to induce epigenetic modifications which can cause the downregulation of genes essential for synaptic plasticity, transmission, working memory, and social cognition (55). Obstetric complications such as bleeding during pregnancy, pre-eclampsia and traumatic births can also increase the risk of developing schizophrenia (3). Severe famine at the time of conception or early in the pregnancy increased the risk of developing schizophrenia two-fold, while mothers with inadequate weight gain increased the risk of psychosis for their offspring by 9-fold (56). Prenatal immune system activation can affect brain development negatively and can slow or alter the neurodevelopmental trajectories which can cause behavioural and cognitive impairments later in life (3). The brain is especially vulnerable in the first and second trimester of pregnancy during critical brain development so maternal stress such as bereavements, unwanted pregnancies and other serious life events are positively linked with the development of schizophrenia and other serious mental disorders (3). Particular childhood and adolescent risk factors are capable of predicting the age of manifestation in patients with and without a relative who has schizophrenia (53). Studies have shown that some patients who develop schizophrenia during adolescence experienced delayed developmental milestones in their first year, had hearing impairments, emotional problems, low IQ in childhood, and interpersonal difficulties (2).

Some of the positive symptoms of schizophrenia have been linked to developmental trauma, cannabis use, living in an urbanised area and the minority group position in that area, for these studies cultural bias and selective migration were inspected and found to not impact these association (56). Living in an area densely populated with the same ethnic group and moving from an urban area to a rural environment decreases the risk of developing any kind of psychosis (2,56). The accumulating evidence that environmental exposures occurring preconception through to adolescence and adulthood play a role in the susceptibility of schizophrenia, as well as ample evidence that exposures to environmental factors *in utero* produce brain anomalies as well as phenotypes similar to schizophrenia (56). Though the associations with environmental factors are robust the observational epidemiology cannot distinguish true causation from the association as a result of pleiotropy or reverse causation (8).

The role of environment in schizophrenia has been hypothesised for decades but the lack of biological models and methodologic limitations has made it a difficult test to what extent they are involved (52). In recent years genetics has dominated the discourse of schizophrenia’s aetiology (9,52). Twin studies have shown a discordance rate for monozygotic twins who develop schizophrenia of 40-55%. Monozygotic twins have identical genomes this illustrates that the risk of developing schizophrenia is not solely genetic but plays an important role (56). The most plausible explanation for this discordance is exposure to environmental factors which are likely to occur as early as *in utero*, or gene-environment interactions during crucial brain development (2,56).

1.4 MRI Findings

The first MRI (magnetic resonance imaging) study focused on schizophrenia was conducted in 1984 and with advancements in technology over recent years many more followed. These studies have shown that there is not distinct diagnostic neuropathology for schizophrenia, but any of the subtle changes which are evident is apparent when a patient first becomes symptomatic (49,57). The lack of evidence of distinct neuropathology for schizophrenia could be explained by schizophrenias diverse presentation, range and severity of the symptoms and if the patient has been treated with antipsychotic medications before the time of MRI scan (48,57). Studies have found that several of the brain abnormalities which can occur are evident before any symptoms appear, hinting at schizophrenias neurodevelopmental nature and that these abnormalities may change over time (21,26). Understanding these changes in brain structure could prove most valuable in prognosis, treatment, and intervention. Across studies, reduced volume in the intracranial is seen especially with male patients and because 90% of the intracranial volume is usually reached at the age of five this suggests that there is an early developmental cause (59). Reduced total brain volume has also been observed consistently with a marked reduction in grey matter volume (GMV), while cerebrospinal fluid, third and lateral ventricles, and the left side of the planum temporale have increased volumes and are associated with more severe symptoms (60). The levels of GMV reduction is associated with elevated doses of antipsychotics and duration of illness (27). In antipsychotic naive patients (ANPs) most of the brain abnormalities observed are the same as those found in medicated patients but to a lesser extent, in ANPs GMV and total brain volume, the effect size was up to 30% less (24). Conversely thalamic and caudate nucleus volumes are more prominent in ANPs which strengthens the evidence that typical antipsychotic medication enlarges the volume of the basal ganglia (61). The levels of white matter reductions are similar in medicated patients and ANPs which suggests that the levels of white matter do not considerably change after onset (60). Some post-mortem studies have also uncovered neuroinflammation in the brains of schizophrenia patients which are unrelated to treatments (9,20).

1.5 Genetics of Schizophrenia

Schizophrenia is heritable meaning having a family member with the disorder heightens the odds of developing it during a lifetime (53,62). Offspring with one parent with schizophrenia have a risk rate of 7% while offspring, where neither parent has schizophrenia, is 0.86% (63). Twin studies have been pivotal in furthering our understanding of the role that genetics plays in schizophrenia's aetiology (62,63). A study performed using the Danish Wide Twin Register found that the propandwise concordance for schizophrenia is 7% for dizygotic twins (Fraternal) and 33% for monozygotic twins (Identical) in terms of disease liability (8). Although previous studies have found higher rates of concordance, with monozygotic twins achieving 30-40% and heritability estimates for schizophrenia >80% (6,21). Longitudinal twin studies have shown that children of the unaffected monozygotic twin have a similar risk as to the affected twins' children of developing schizophrenia or a schizophrenia-related disease in their lifetimes (9). Taken together, these findings highlighted a clear genetic susceptibility to schizophrenia.

As a result of multiple technological advances and extensive collaboration, there have been remarkable advances in the genetics of schizophrenia in the past decade (64). Genome-wide association studies (GWAS) have been valuable in uncovering many schizophrenia risk loci, including single nucleotide polymorphisms (SNPs), copy number variants (CNVs) and insertions or deletions in bases in the genome (INDELs) (48). These GWA studies aim to identify areas of the genome that increase an individual's risk for developing schizophrenia (65). Schizophrenia is a complex polygenic psychotic disorder meaning it is not caused by one genetic variation with a large effect but rather by a combination of multiple genetic variants that each subtly increases the risk of the disease developing (66,67).

Genome-wide association studies (GWAS) are a powerful tool used for studying the genetic architecture of diseases (68). It is an experimental design used to uncover associations between traits of interest and genetic variants, with the aim of better understanding the underlying biology which could lead to better treatments and prevention strategies (69). GWASs have also been successful in uncovering disease-associated biological processes and assisting in risk prediction. (69). GWAS exploits linkage disequilibrium (LD) to measure an association at one genetic variant as a proxy for other genetic variants, the statistical power of these studies depends on the sample size, the distribution of effect size of the casual genetic variants, their frequency in the general population and the LD between genotyped DNA variants and unknown causal variants (65,69). In 2009 the first robustly associated loci linked with schizophrenia were identified using a sample size of 3000, and in 2014 using a sample size of 35,000 cases the number of genetic variants increased to 128 common variant associations across 108 genetic loci (1). In 2018 a schizophrenia meta-analysis identified 179 independent significant SNPs which mapped to 145 loci (1,66).

Formatted: Not Highlight

The 2018 meta-analysis GWA study for schizophrenia had 11,260 people with schizophrenia (cases) and 25,542 healthy controls (HCs). Potential schizophrenia risk genes were generated by taking proximity of a gene to SNPs into account but also the kind of genetic variant, expression quantitative trait loci (eQTLs), chromosome conformation data and genomic finemapping (70). Despite the large sample size, there was not a huge overlap of results from the previous GWAS, which can be expected from studying a complex, polygenic disorder like schizophrenia. (1,66). This study found that associations converge in specific cell types like pyramidal cells, some interneurons and medium spiny neurons (67). It was estimated that from a third to a half of genetic liability derive from common alleles, and a large portion of rare variant architecture comes from mutation intolerant genes which have also consistently been observed in other neurodevelopmental disorder (1,66). It was also noted that in the case of schizophrenia there was an enrichment of common variants associated with loss of function (LoF) and mutation intolerant genes and that these genes accounted for 30% of SNPs based heritability (66). People with schizophrenia have decreased fecundity and early mortality, but the common risk alleles persist in the population which could be because of 1) balance selection that schizophrenia-related alleles have reproductive advantages so are preserved because of their association with positively selected alleles or 2) the effects of gene-environment interaction on these rare variants (8).

Despite this improved knowledge, the understanding of the underlying biological mechanism has not progressed far enough to develop new treatments or cultivate preventative strategies.

1.6 Aim of the project

Understanding how gene expression and regulation differ between individuals has advanced the understanding of healthy tissues and the origins of diseases and complex traits (71). To get a better understanding of the control of gene expression it’s important to understand the relationship between genotype and phenotype and RNA sequencing which is a more quantitatively accurate absolute transcript (72).

This study aims to examine gene expression of the schizophrenia-associated genetic loci, identified by Pardiñas et al., in the developing human brain using the BrainSpan Atlas of the ABA repository. Of the 145- independent schizophrenia-associated loci provided by Pardiñas et al. 316 genes were available on ABA’s resource, BrainSpan for further investigation. The expression profile of the 316 schizophrenia-associated genes was investigated using 1) K-means to identify underlying patterns in the genetic data and 2) network analysis using weight-gene co-expression network analysis (WGCNA) in the developmental stages available in ABA’s Brainspan resource. Both steps were performed across the five developmental stages of the ABA dataset. Gene modules identified using WGCNA were further characterized by identifying hub genes and performing enrichment analysis to identify schizophrenia-related biological pathways. Identifying relevant biological pathways can further our understanding of disease aetiology and present new targets for novel therapeutics which could provide better outcomes for schizophrenia patients.

Chapter 2 – Methods

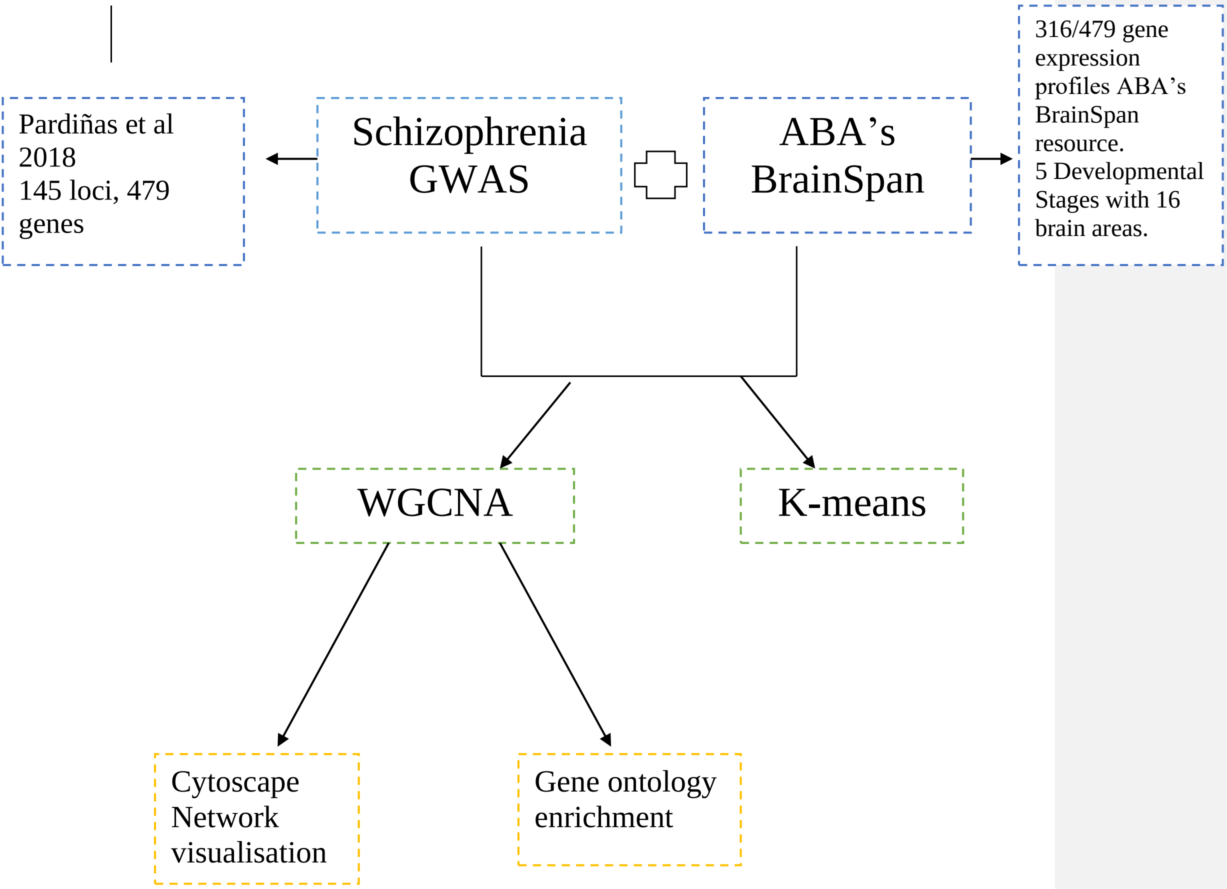


Figure 1.6.1 Diagram of methods performed in this study.

Chapter 2 - Methods

2.1 Collation of Schizophrenia-associated Genes

Our knowledge of schizophrenia's genetics has vastly improved in the past decade, however, identifying gene targets has proven difficult (1). Pardiñas et al. completed a GWAS meta-analysis in 2018 which had 11,260 cases obtained from a CLOZUK sample on genetic information from people with schizophrenia in the UK and 24,542 controls (66). The meta-analysis used cases from the CLOZUK GWAS and combined them with Psychiatric Genetic Consortium (PGC) datasets from the 2014 GWAS excluding any overlapping samples which brought the total number of cases to 40,657 cases and 64,643 controls, the meta-analysis highlighted 179 independent genome-wide SNPs which were significant, which mapped to 145 loci (66). 93 of the 145 loci had been previously identified by the PGC GWAS in 2014 and had shown an increased association in the 2018 meta-analysis (1,66). Summary statistics were added to the ClOZUK genes so a combined analysis could occur. The PGC data was re-examined with a fixed-effects procedure derived from standard errors and polygenic risk scores were calculated for the whole dataset (66). In this study, schizophrenia-associated genes (n=316) identified by the GWAS meta-analysis conducted in 2018 by Pardiñas et al. who identified 145 loci total with $P < 5 \times 10^8$. If the loci did not overlap with a gene, the closest gene within a 500kb radius was used. Of the 145 loci, 316 genes were available from ABA's BrainSpan resource (66).

2.2 Allen Brain Atlas

The brain is the most complex system in the human body with approximately 86 billion neurons and around a trillion synapses per cubic centimetre of the cortex (73). Its circuitry, cellular and structural diversity and the regulation of its transcriptome are far from being completely understood. One billion people are suffering worldwide with brain diseases and disorders but there is a lack of diseased tissue to study. There is a need for another approach to uncovering their aetiologies (74). The challenge of this type of research is the scarcity of high-quality post-mortem human brains. These brains are normally dissected at brain banks and distributed to various research groups and thus data derived from these analyses have diversified hypotheses that are non-parallel as well as different types of research methods, which has hindered the analysis of brain disorders (73). Although other species model systems have been useful, analyses of the human brain itself are essential to get a true understanding (73).

ABA is a public resource that gives access to gene expression, connectivity, and neuroanatomical data for mouse, primate, adult humans, and developing brains for humans and mice which integrates MRI, genomic and anatomic information, histology, diffusion tensor imaging and gene expression data derived from ISH and microarray methods (75). The original Allen Human Brain Atlas (AHBA) uses high-quality post-mortem brains from males and females between the ages of 18-68 with no known neuropathological or neuropsychiatric history and maps the genes expression to the stereotaxic space. (76) This valuable tool can help researchers trying to comprehend how spatial variation on the molecular scale associate with macroscopic neuroimaging phenotypes (76). While there are other human atlases, only Allen Human Brain Atlas (AHBA) possesses high-resolution coverage of the majority of the brain (76).

The brain tissue underwent several tests including serology, toxicology and tested the RNA quality to determine if it meets the inclusion factors (74). If the brain tissue samples passed, they were then sent to tissue repositories for initial tissue processing. After which the brain tissue was frozen, after and sent to the Allen Institute where thorough quality control (QC) tests were performed, the brain tissues that passed this threshold have histological data collected from them and the tissues were subdivided and categorised based on if they contain cortical or subcortical substructures. Additional tissue containing subcortical structures were collected and then placed on membrane slides so laser microdissection (LMD) could occur. Both cortical and subcortical tissue samples were collected for microarray analysis. The microarray analysis quantified the expression levels of thousands of genes at once by measuring the hybridisation of Cy3-labelled RNA (cRNA) to a probe on a microarray chip (Agilent 8 X 60K custom design arrays) (74). The probes were mapped to a specific location of DNA that contains single-stranded nucleic acid profiles which recombine with their complementary targets during hybridisation. The gene expression levels in the tissue samples were quantified by measuring the fluorescence at the sequence-specific locations which correlate to the levels of mRNA (76).

2.2.1 Brainspan Atlas of the Developing Human Brain

The human brain develops following a complex series of histogenic occurrences that depend on differential gene expression and its complex development is not fully understood (77). During the first 6 months of embryonic life, the brains general architecture is formed this is driven by strong genetic influences which are silenced in the third trimester allowing for environmental factors to influence the last trimester (78). Mice and non-human primates' models have been useful in developing some knowledge of the brain but the differences between species is a huge limitation. Firstly, because of the difference in size, in addition to this the evolutionary differences which are seen in the superficial layers of the neocortex and secondly the developmental differences in the evolution of GABAergic interneurons (77). The shortage of human prenatal tissue and the use of different species models which have their restrictions has hindered the development of an anatomically comprehensive atlas of the prenatal human brain which could be used for studying the roots of neurodevelopmental and psychiatric disorders (79). The ABA resource BrainSpan transcriptional atlas of developing human brain is a repository of RNA sequences expression profiles of 16 brain structures from 8 weeks post-conception (prenatal) to 40 years of age (80). The stages are outlined in Table 1. The prenatal stage is made up of four high quality mid gestational brains, two from fifteen to sixteen post-conceptual weeks, and two twenty-one post-conceptual weeks specimens. These tissues had no history of maternal drug or alcohol abuse or potential agents that could disturb their development or relations with HIV 1 or 2 or HepB or HepC (77). The specimens were donated from the birth defects research lab at Washington University and the Advanced Biosciences resource in California (77). The left hemispheres were coronally, serially cytosectioned onto polyethylene naphthalene (PEN) membrane slides for LMD and histologically stained for detailed structure identification, and three hundred regions per specimen were isolated (77). The right hemisphere of two of the specimens was handled similarly and was used further for In situ hybridisation and Nissl staining for structure identification (81).

The sample locations were mapped to MRI coordinates and then to the Montreal Neurological Institute (MNI) coordinate space (81). This data was anatomically delineated to create a digital reference atlas which allows for the visualisation of transcriptome data in its exact coordinates. The atlas resources also include MRI, diffusion-weighted MRI from three brains with the approximate same post-conceptual weeks as well as the white matter reconstruction for three additional brains (77).

Table 2.1: Age categories from the developmental stages for ABA's resource BrainSpan available for download in R.

Stage	Age category
1	Prenatal
2	Infant (0-2 years)
3	Child (3-11 years)
4	Adolescent (12-19 years)
5	Adult (>19 years)

2.2.2 ABA and its application in research

The scarcity of suitable brain tissue available for research led scientists to develop Allen Brains Atlas human brain resource. ABA's gene expression data being accessible at high neuroanatomical data makes it possible to identify intricate gene expression patterns for healthy human brains, these profiles for healthy brains can be used as a baseline to identify genes involved in neurological conditions by using machine learning techniques which could be connected with a neurological condition. This approach was successfully applied by Negi et al. where they applied machine learning methods such as hierarchical clustering and weighted co-expression on ABAs gene expression profiles across brain regions (82). From there they were able to build supervised classification models for Autism and Parkinson's with 84% and 81% accuracies respectively (82). Researchers can solely use ABAs resources alone or can apply external data from GWAS or MRI studies to aid their analysis. McCarthy et al. applied the latter technique when investigating Bipolar disorder (BD), they took 58 genes identified to be involved with BD from a previously published GWAS and looked at their expression pattern across 900 areas (83). They also compiled a meta-analysis of MRI studies looking for structural abnormalities across patients diagnosed with BD (83). They aimed to see if they could link unusual gene expression in the BD genes with the brain structural differences (83). Using ABA's Brainspan human brain transcriptome database Mahfouz et al. hypothesised that understanding the functional relationships between ASD candidate genes during normal development could provide insight into ASD's genetic heterogeneity (80). Over human development, the heterogeneous ASD candidate genes share transcriptional networks related to protein turnover, mitochondrial function and synapse elimination and formation (80).

Figure 2.2.2.1 Schizophrenia-associated genes identified by Pardiñas et al and their position on the chromosomes.

2.2.3 Collation of schizophrenia-associated genes from ABA

The ABA data was downloaded into R (version 4.0.0) using the R packages ABAData and ABAEnrichment (52,74). The complete ABA dataset has 17,245 genes expressed in 16 distinct areas over five developmental stages from prenatal to adulthood (See Table 2.1 for more detail). The genes found to be significantly related to schizophrenia identified by Pardiñas et al. determined by their p-values (66) were exported by CSV file into R. The 17,245 genes available in ABAData were filtered into five dataframes for each of the developmental stages for further analysis. The dataframes were shaped into a wide format using the `pivot_wider` function in R where the brain areas are the column names, and the row names were converted to the schizophrenia-associated gene names using `column_to_rowname` function in R and the gene expression for each gene was scaled. The schizophrenia-associated genes are available in Table 37 in the Appendix.

Figure 2.2.2.1 shows the schizophrenia-associated genes identified by Pardiñas et al. which were available on ABA's resource and where they lie on the chromosome.

2.3 Machine Learning and Clustering

Regression analysis, feature selection methods, and classification are elements of the term Machine Learning (84). Classification can be subdivided into supervised, semi-supervised, and unsupervised. Supervised classification deals with objects that are labelled beforehand and build a learning algorithm which is then used to predict the classification of unlabelled data. Semi-supervised uses labelled and unlabelled data to train an algorithm (85). Unsupervised classification defines classes without help from previously known labels (84), clustering is a form of unsupervised learning.

In genetics, large datasets of genes and their expression are given to a clustering algorithm to cluster genes whose expression are similar to each other. These algorithms can be used for prediction, classification, and identification in DNA sequences but can also be taught to distinguish between phenotypes and identify possible biomarkers (85).

2.3.1 Unsupervised Learning and clustering for gene expression data

Unsupervised learning is a machine learning technique that looks for natural structures in data and groups them without classifying them (85). Gene expression data is massively complex. Clustering is an unsupervised learning approach capable of discovering subgroups within a dataset, each of these subgroups or clusters have similar observations within them. This type of analysis has been a cornerstone for interpreting biological information from large gene datasets (86). Clustering can group genes based on their similar expression across brain areas and discover patterns in the data. Clustering can suggest regulatory relationships between genes and transcription factors and can further genome annotation by using the principle of guilt by association, as well as give a better understanding of how diseases manifest and can progress over time (87).

2.3.2 K-means clustering and NbClust

One of the most fundamental modes of understanding learning is to organise data into sensible groupings (88). K-means is a numerical, unsupervised, iterative, non-deterministic method that is classified as a partitional clustering algorithm (89). The k-means algorithm finds a split so the squared error between the points in a cluster and the empirical mean is minimized in each of the clusters. To perform k-means the number of clusters (k), distance metric and cluster initialisation must be pre-specified before the algorithm can be run (88). The goal of k-means is to produce groupings each with a high degree of similarity and a low degree of similarity with the other groupings (90). One of the issues with k-means is deciding the number of clusters (k) that are suitable for a dataset, there are many different indices to determine this but a package in R called NbClust integrates thirty different indices in one package to determine the optimal number of clusters in a dataset (91).

2.3.3 Kmeans analysis of schizophrenia-associates genes

In this study, unsupervised machine learning techniques were performed on the schizophrenia-associated gene set to identify underlying patterns. The NbClust package in R was used to determine the optimal number of clusters for each of the developmental stages. NbClust uses thirty different methods of determining cluster number and produces a bar chart to visualise which cluster number fits best (91). The optimal cluster number was put into the kmeans function in R for the centres. Twenty-five was selected to be the optimum number for initial configurations (nstart) and Euclidean distance was used. Each cluster is filled with genes with similar expression patterns. After the K-means analysis was performed each of the clusters was visualised using the `fvis_cluster` from the factoextra package in R (92). A table to show module assignment for each of the genes is available in table 17 in the appendix.

2.4 Co-expression Network Analysis

The information found in gene expression data can be used to link genes with unknown function to biological processes, identify candidate genes for disease, determine transcriptional regulatory systems, and identify novel targets for therapeutics (93). Co-expression network analysis recognises genes that show coordinated expression patterns, and the networks can be shown as gene-gene similarity matrixes in later analyses. Co-expression looks to identify relationships between pairs of genes by using mutual background information or correlation (93). These pair-wise correlations between them are then rolled out to the other genes in the dataset until a network is formed where multiple modules are fashioned and each node signifies a gene, and the edges represent the presence and strength of the relationship. Functional enrichment analysis can be

Formatted: Not Highlight

applied to the modules formed after applying the co-expression network analysis method, these modules can often represent biological processes (93).

2.5 Weighted correlation network analysis (WGCNA)

Genes do not work alone, and each gene can work with between four to eight genes which in turn could be involved in up to ten biological processes (94). Any dysfunction in these pathways can potentially lead to diseases. There are many ways to analyse complex, multi-dimensional genetic data, and one of the most popular methods are correlation networks. This technique is a useful way of discovering the underlying intrinsic organisation of the transcriptome. Constructing gene co-expression networks (GCN) for complex diseases is an important method of identifying genes involved in disease, highlighting highly connected genes within the networks and modules that can lead to novel therapeutics or biomarkers for diagnosis. WGCNA is an unsupervised learning systems biology network analysis method for associating correlation patterns among genes across gene expression microarray samples. The WGCNA package which is available for download in R can construct gene networks, identify modules, and can detect highly connected genes that are representative of the module using hierarchical clustering (95). When WGCNA is performed the algorithm evaluates the expression for each gene, pairing them based on topological overlap (TO) and then considering the degree of shared neighbours looking for consistent gene expression patterns and placing them into modules (82). Once the modules are defined the module eigengene (ME) which is the first principal component of the module is isolated and centralised. Highly connected nodes which are most like the ME and that are representative of the modules are specified and these are called hub genes (HG). The module membership (MM) calculates the degree of correlation between the genes within a module and the ME (96).

Using module significance (MS) methods can help detect important modules which contain high average node significance (NS) and the gene significance (GS) which is the correlation between a node and a phenotype of interest (95). The WGCNA algorithm can execute network construction, module detection, gene selection, data simulation, visualisation, and calculate topological properties (95). WGCNA has been applied successfully with cancers, mice and yeast genetics, and brain imaging data.

2.5.1 Network analysis using WGCNA

After the initial unsupervised learning analysis looking for underlying patterns, network analysis was performed using WGCNA in R. WGCNA was performed on the developmental stages because it is an effective way to characterise correlation patterns within the schizophrenia gene set, genes that correlate sometimes are related biologically. Networks were constructed using an adjacency matrix which looked at the co-expression similarity between a pair of genes and constructed a hierarchical graph. Pairwise correlations were used to identify modules where genes with similar gene expression are grouped into modules. To construct the weighted gene network, a soft threshold power analysis was first performed using the `pickSoftThreshold` function within the WGCNA package to calculate the adjacency by using gene co-expression, the power in the `pickSoftThreshold` function was calculated independently for each developmental stage and verbose was set at five (94,95). Once the power is chosen to calculate adjacencies, the adjacency is transformed into Topological Overlap Matrix (TOM) and used to calculate the dissimilarity. A clustering dendrogram is made from the genes using TOM-based dissimilarity and subsequently a minimum module size of 10 was chosen the genes were assigned to modules with genes of similar expression profiles. HG and ME were identified in each of the modules using the `moduleEigengenes` function and the `chooseTopHubInEachModule` function in WGCNA. Each gene in the module is annotated to its distance from the ME this is the MM.

2.5.2 Networks and their applications

Networks are abstract models made up of nodes, vertices, and a set of edges. The nodes are the entities and the edges are the information that connects them (97). There are different types of networks for different situations that can yield different outputs, directed networks are formed when nodes are asymmetrical and mean one can influence the second, but the second cannot influence the first. Undirected networks are when the relationship between the nodes is symmetrical and is most useful for exploratory analysis of genes (97). Understanding the intricate relationship between diseases or disorders and their underlying mechanism is a subject that continues to challenge the areas of medicine and biology. There is clear evidence that there are disease-disease associations where two or more conditions can have similar or identical underlying mechanisms and understanding one can further the understanding of the other (80). The advancements of high throughput technologies like DNA microarray and next-generation sequencing have given researchers large scale genomic datasets (98). Constructing new biological pathways is generally achieved by using the interactions found from previous studies with gene regulation information for specific diseases or tissues, using system-level biological data is predicted to improve current knowledge of underlying mechanisms and lead to improvements in diagnosis, prognosis and treatments (99).

2.5.3 Cytoscape

Cytoscape is a free software project which combines expression data with biomolecular networks and aids in visualising, querying, and linking the data to functional annotation databases (100). Functional proteomics and genomics techniques allow for measurements of expression profiles and interactions between cells and tissue to be collected which could potentially map cellular processes and their dynamics. From these expression profiles active biological processes can be identified using enriched gene annotation and by combining expression profiles and cellular network interactions changes in biological activity could be explained (101). Cytoscape allows for protein or gene properties to be associated with the nodes and edges by changing their appearances which allows for numerous types of data to be seen in a network context. It also includes a range of environments that can model gene transcription kinetics, biochemical reactions, and metabolic control, which advance biological research (100). To gain insight into the structure and organisation of a network Cytoscape's NetworkAnalyzer plugin was developed for visualisation and analysis (102). NetworkAnalyzer computes a set of topological parameters, including, the number of nodes, edges, network diameter, radius, density, centralization, heterogeneity, connected components, clustering coefficient and shortest path lengths (103).

2.5.4 Visualisation of modules using data from WGCNA in Cytoscape

Gene relationship data, module membership, weight and direction of the edge was saved to csv files Node files for each module in the developmental stage, this data was generated using WGCNA analysis in R. These CSV files were exported into Cytoscape (104) (version 3.7.1) using the `exportNetworkToCytoscape` function so each of the modules for the developmental stages could be visualised. Firstly, the edge data was exported, source and target node columns were selected, and the p-value for SNP inclusion determined by Pardiñas et al. (66) was marked as the source node attribute and weight which was filtered to 0.8 and above was selected as the edge attribute. To the node file produced by Cytoscape, the node table for each of the modules was loaded in for the module membership of each node. Cytoscape’s NetworkAnalyser was applied to each of the modules, each of the modules was treated as undirected. Using the visualise parameters function in NetworkAnalyzer the size of the Node (which referred to the genes) used the MM, edge width was mapped to the weight of the schizophrenia genes and the node colours were charted to the p-value determined by Pardiñas et al. (66).

2.6 Gene Ontologies (GO)

A gene ontology defines a gene's function and how the functions of other genes are related to each other (105). GO is described with respects to three features: molecular function (the activities performed at a molecular level by the gene products), cellular components (where the gene product performs a function relative to a cellular structure), and, biological processes (the biological programs which are completed by several molecular activities) (105,106). As the knowledge of gene ontology is expanding so too are the databases that house them.

anRichment a package available in R is used to calculate ontology enrichment within the modules provided when compared to known reference gene sets such as KEGG, GO, Reactome, etc. (107). By using the function `enrichmentAnalysis` in anRichment and providing a module in the classes input and a collection reference gene set, GO enrichment analysis was applied (107). GO enrichment using anRichment was performed on each of the modules to identify biological processes which are over-represented in each of the modules (107). To run the analysis using `anRichment` the `enichmentAnalysis` function was run using GOcollection which is built using `org.Hs.eg.db` R package (108) and specifying species as human, the threshold was set at 1e-4, the threshold type was Bonferroni, getOverlapEntrez = TRUE, getOverlapSymbols = TRUE and ignoring the grey module.

Formatted: Left, Right: 0 cm, Line spacing: single

Chapter 3 -Results

3.1 Data Pre-processing

The schizophrenia-associated genes (genes tagged) from the GWAS meta-analysis by Pardiñas et al. (66) are found in the supplementary materials section NIHMS958804-supplement-Supplementary_Table.xlsx on the sheet titled "Supplementary Data Table 4: Independent genome-wide significant association signals from the CLOZUK + PGC meta-analysis, clumped and amalgamated into loci". This data was downloaded into R (66). The `dataset_5_stages` function available within the ABAData package was loaded into R, each of the five age categories was separated into their data frame and then filtered so they only included the schizophrenia-associated genes identified by Pardiñas et al (66). This gave five dataframes for each developmental Stage with 316 schizophrenia-associated genes expression over 16 distinct brain regions. Each of the dataframes was placed in a wide format where the gene names were rows, and the brain regions were columns. This is illustrated below in Table 3.1.

Table 3.1: Wide-format of the schizophrenia-associated genes data frame. Brain areas available from ABAs Brainspan are the names of the columns in bold and the schizophrenia-associated genes are the row names. Each cell contains the scaled gene expression for each brain area

	10163	**10173**	**10185**	**10194**	**10209**	**10225**
ABCB1	0.645635	0.193074	0.103555	-1.09892	0.05936	-1.24916
ABCB9	-0.62942	0.871415	1.128378	-0.6889	-0.32854	0.351194
ABCD2	1.410626	0.488282	0.541653	0.88314	1.281896	0.40718
ACO2	-0.66424	-0.09813	-0.02023	-0.50082	-0.88345	-1.04544
ACP2	-0.5259	-0.74472	-0.28589	-0.63244	-0.70708	-0.90644
ACTR1A	0.84579	0.475071	0.904972	0.463709	1.186486	0.717958
ACTR5	-0.97613	0.244193	0.215631	-0.26563	-1.25836	-0.95691

3.2 Unsupervised learning using K-means analysis on the schizophrenia-associated genes

Unsupervised learning looks for patterns in data. During this project, a large amount of gene expression data from the ABA was used to associate schizophrenia-associated genes based on the similarity of their expression profiles across brain areas. Unsupervised learning allows users to group the schizophrenia-associated genes into clusters.

3.3 Determining Optimal cluster Number using NbClust

K-means analysis was used to cluster each developmental stage to show the schizophrenia-associated genes cluster. As K-means requires the number of clusters to be prespecified, the `NbClust` function in R was used to determine the optimum number of clusters in each developmental stage as is illustrated in Figure 3.3.1.

`NbClust` uses thirty optima (91) ways of determining cluster number and produces a bar plot to represent how many times each number of clusters appeared. The optimal cluster numbers for developmental stages 1-5 are as seen in Figure 3.3.1. The `kmeans` function available in R was used to run the analysis and the package `factoextra` was used to visualise the clusters (109).

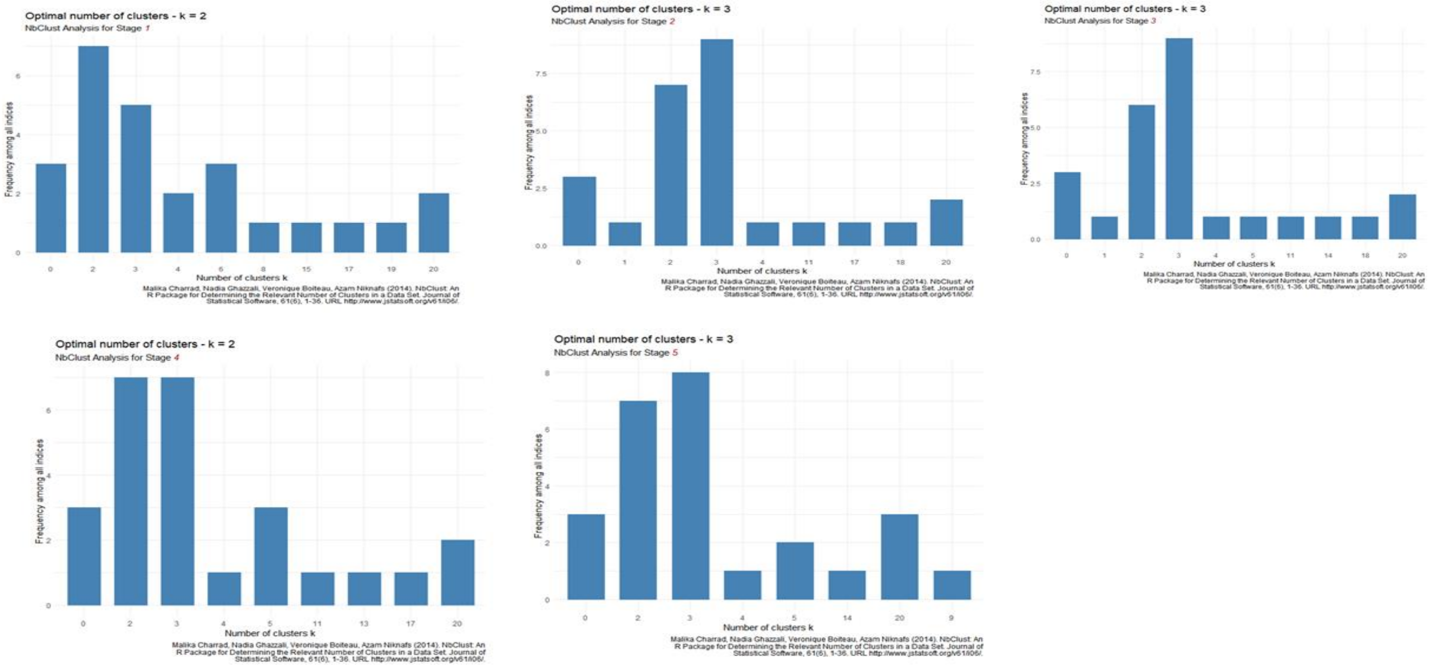


Figure 3.3.1 NbClust analysis performed on the schizophrenia genes identified by Pardiñas et al. to identify the optimal number of clusters (looking between 1-10 clusters) for K-means clustering on each of the five developmental stages available on ABA's BrainSpan resource

From the results in Figure 3.3.1, there is a clear indication of the optimal number for all clusters except Stage Four where 2 and 3 are both optimal.

Figure 3.3.2.-3.3.6 visualise the cluster assignments for each of the schizophrenia-associated genes in each developmental stage. Figure 3.3.7 exhibits the clusters of each developmental stage side by side for comparison.

The sum of squares of a cluster measures the total variance within a cluster. A smaller sum of squares means a more compact cluster which means there is internal cohesion in the cluster. A low total variance tells us that the genes in the Developmental stage are like each other, in good clustering, a total variance would achieve a high percentage where the difference between the groups would explain a majority of the total variance and the within-cluster variance would explain the small fraction left.

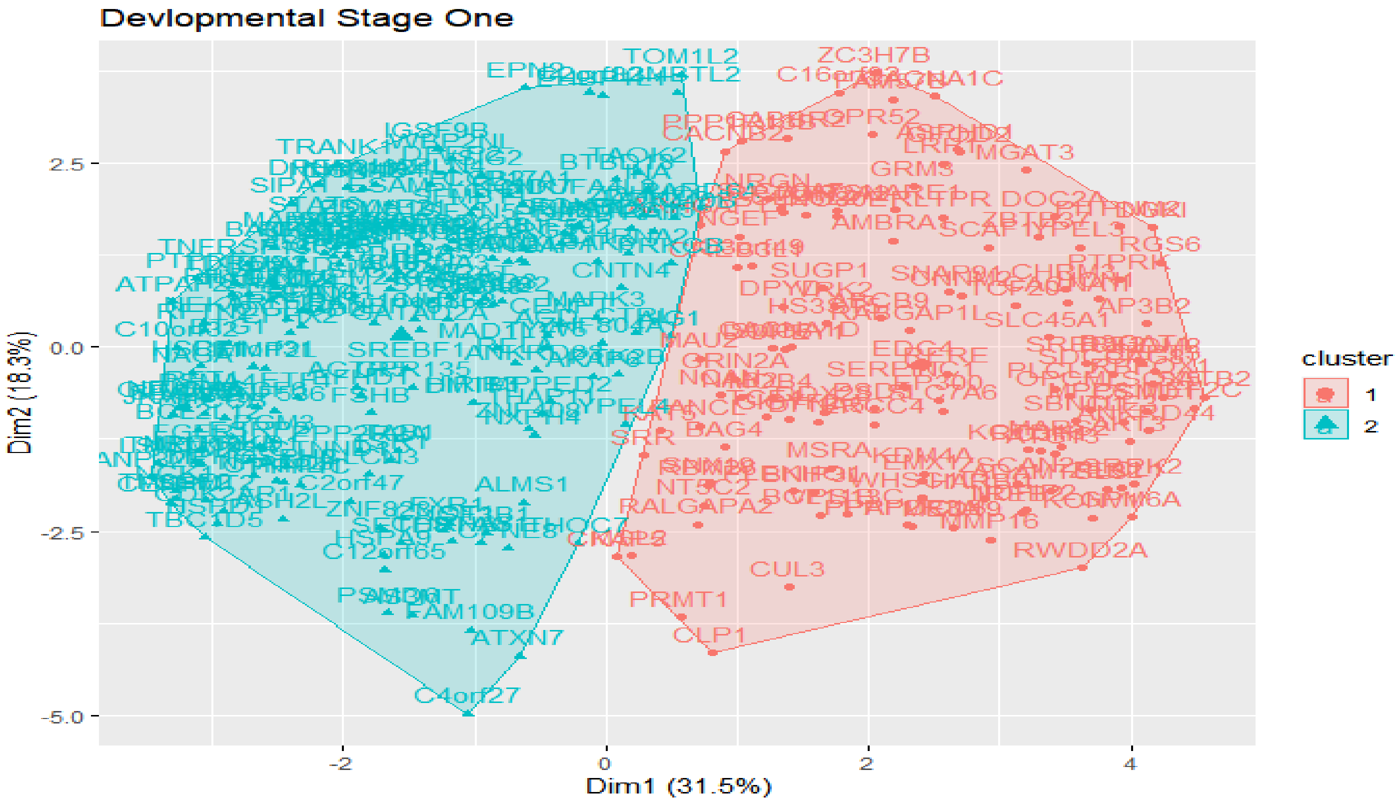


Figure 3.3.2: K-means analysis on the Developmental Stage One genes using the kmeans function in R with 2 centres selected as per the NbClust recommendation and nstart= 25

Table 3.2: Sum of squares for each module in Developmental Stage One determined using the `k-means` function in R.

Cluster	***1***	***2***
Sum of Squares per cluster	1714.970	2115.253

The total variance in the data (Between SS/ Total SS) = 24%

In Figure 3.3.2 cluster One is the most compact cluster, and this is confirmed by it having a smaller sum of squares seen in Table 3. Both cluster One and cluster Two have a very high within-cluster variance. A total variance value of 24% shows that the gene expression data in all of the genes are similar and most of the variance is explained by within-cluster variance. A higher total variance is more desirable with only a portion of the variance being explained within the clusters.

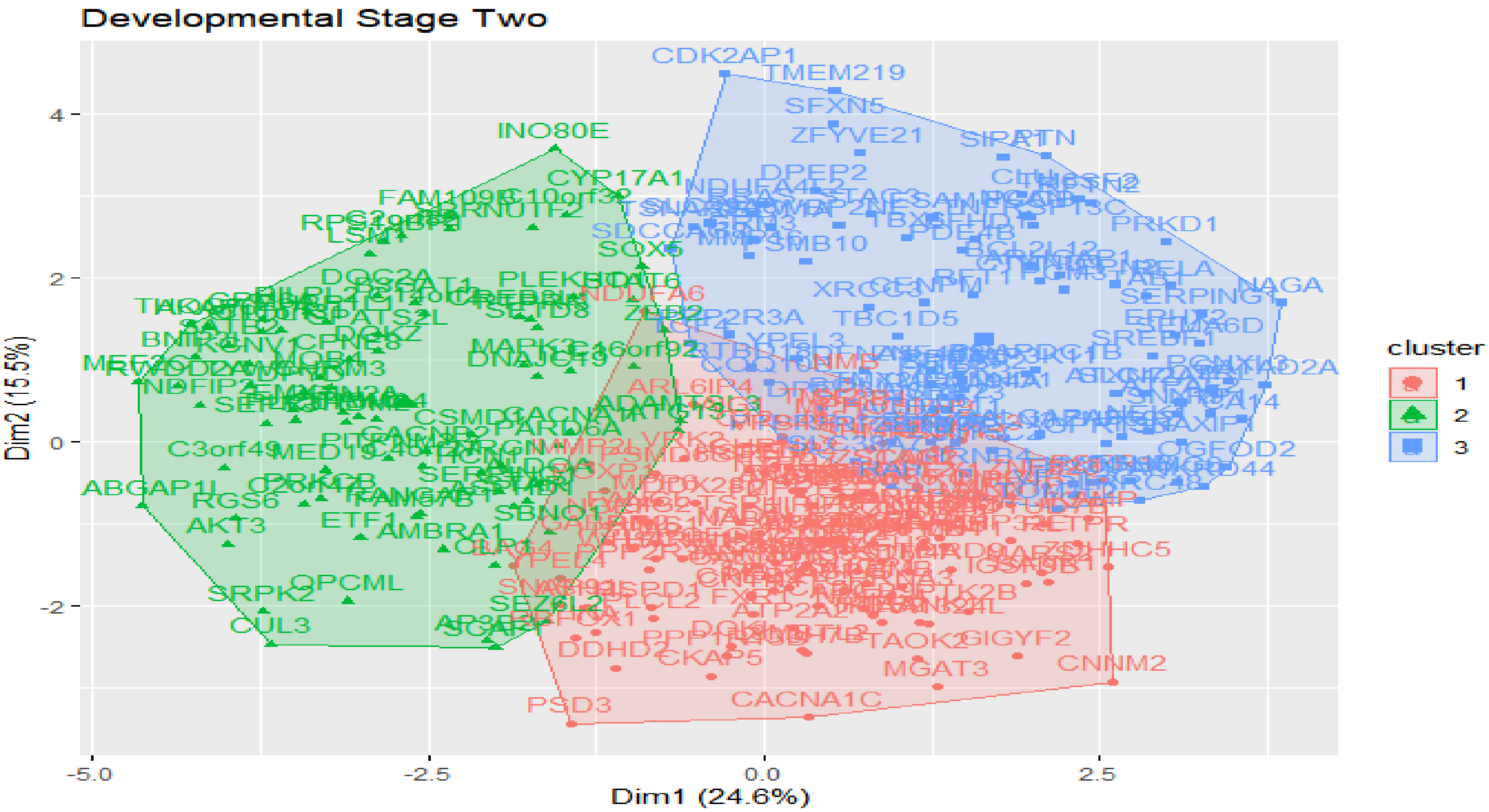


Figure 3.3.3: K-means analysis on the Developmental Stage Two genes using the kmeans function in R with 3 centres selected as per the NbClust recommendation and nstart= 25.

Table 3.3: Sum of squares for each module in Developmental Stage Two determined using the kmeans function in R

Cluster	***1***	***2***	***3***
Sum of squares of cluster	1339.484	1128.824	1247.380

Total variance in the data = 26.3%

In the clusters in Figure 3.3.3 each of the sum of squares is high, this tells us that there is a lot of in-cluster variation, the total variance of the data is also low at 26.3%. The high sum of squares within the clusters explains a lot of the total variance in the dataset.

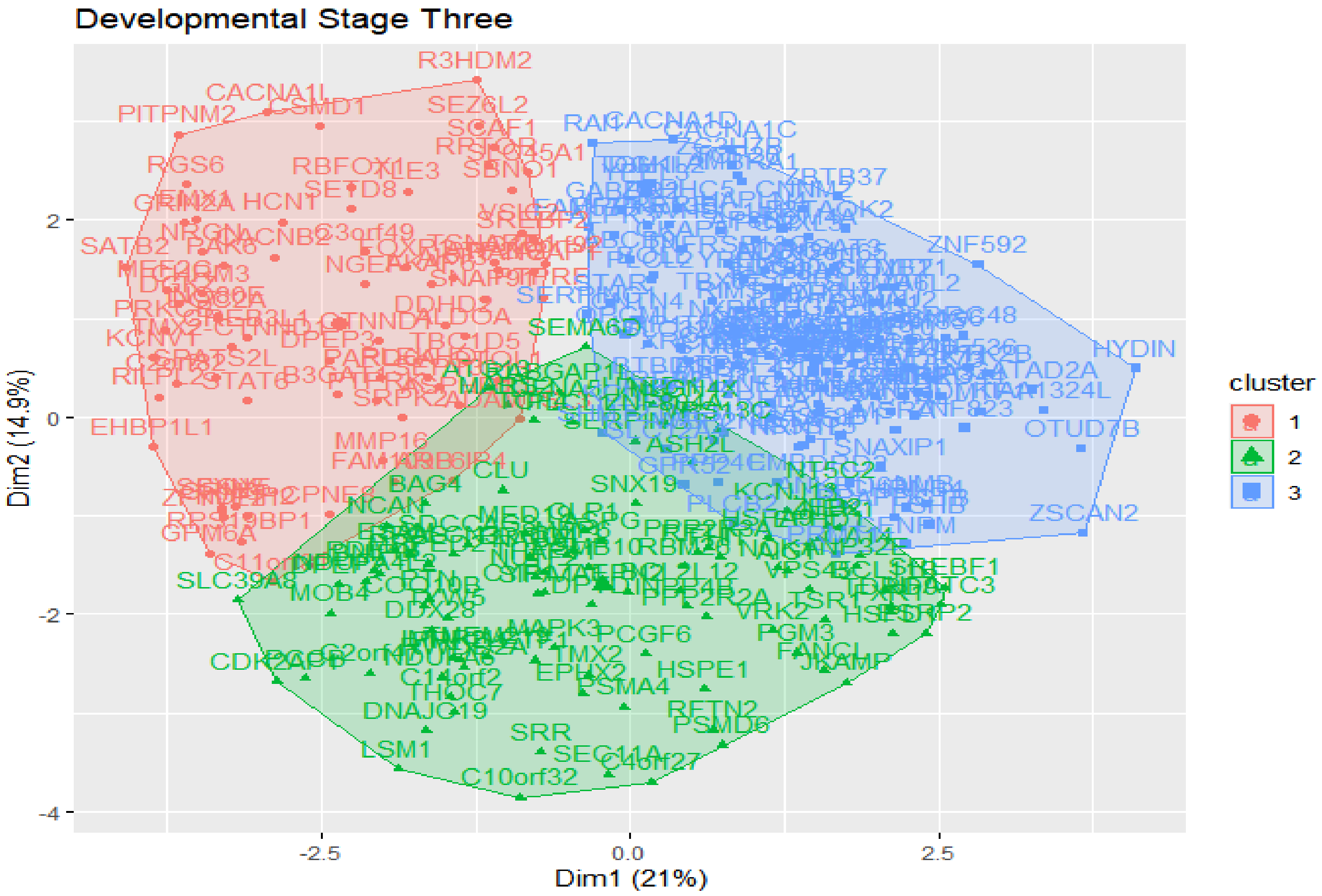


Figure 3.3.4: K-means analysis on the Developmental Stage Three genes using the kmeans function in R with 3 centres selected as per the NbClust recommendation and nstart= 25

Table 3.4: Sum of squares for each module in Developmental Stage Three determined using the kmeans function in R

Cluster	***1***	***2***	***3***
Sum of squares of cluster	925.8711	1544.4462	1389.3763

The total variance of data in module = 23.4%

The sum of squares of both clusters in Figure 3.3.4 tells us that again there is some variance in the clusters. The low total variance in the data also tells us that the clusters are similar. Ideally, the properties of good clustering would have clusters that are alike, and the other clusters would be very different giving a total variance of the data a percentage closer to 1.

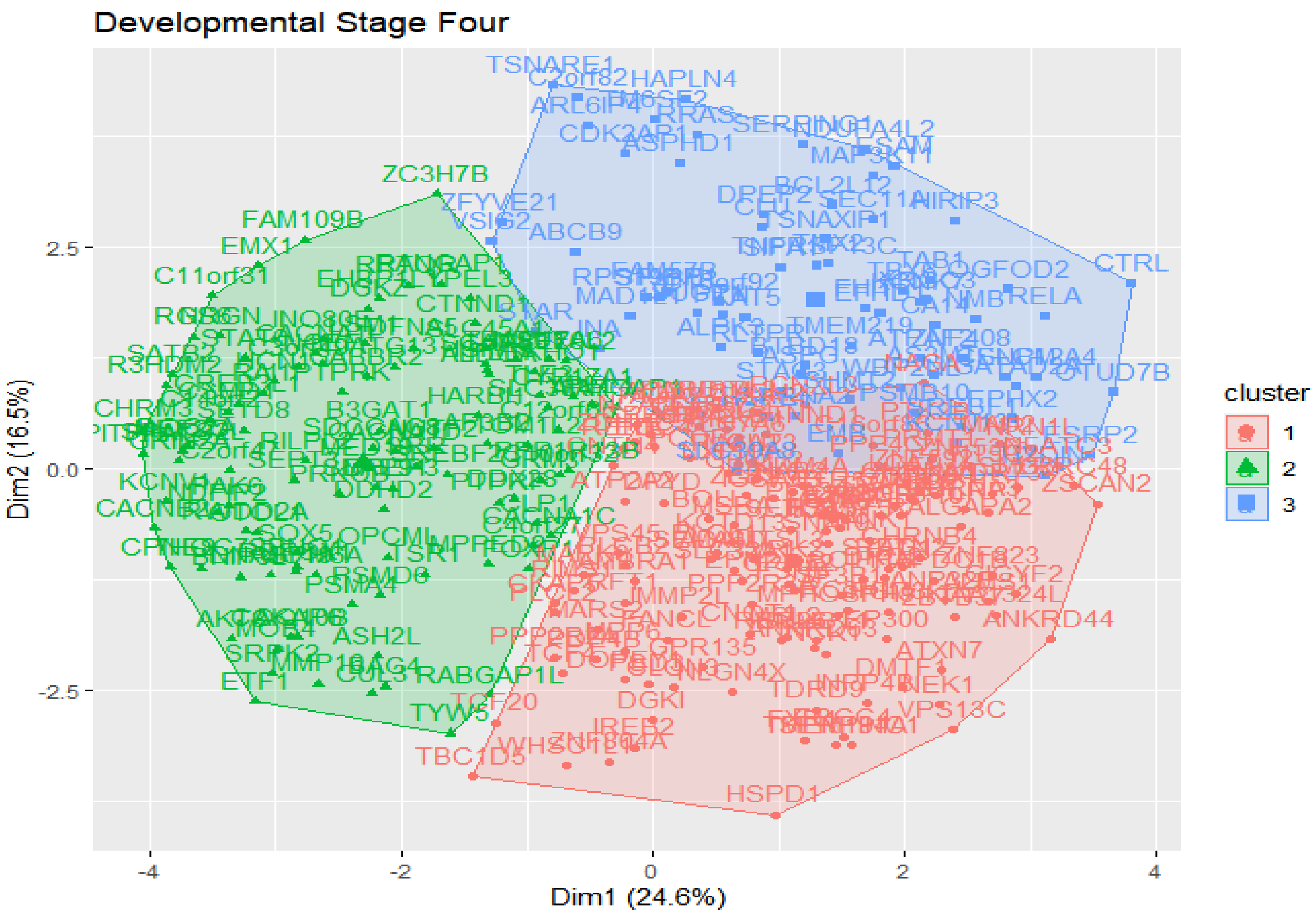


Figure 3.3.5: K-means analysis on the Developmental Stage Four genes using the kmeans function in R with 3 centres selected as per the NbClust recommendation and nstart= 25.

Table 3.5: Sum of squares for each module in Developmental Stage Four determined using the kmeans function in R

Cluster	***1***	***2***	***3***
Sum of squares of cluster	1430.5423	1308.7500	993.0729

The total variance of data in module = 25.9%

The in-cluster variation in Figure 3.3.5 is especially high for cluster One whereas cluster Two and Three there has a similar cluster sum of squares. A large portion of the total variance in the module would be explained by cluster One.

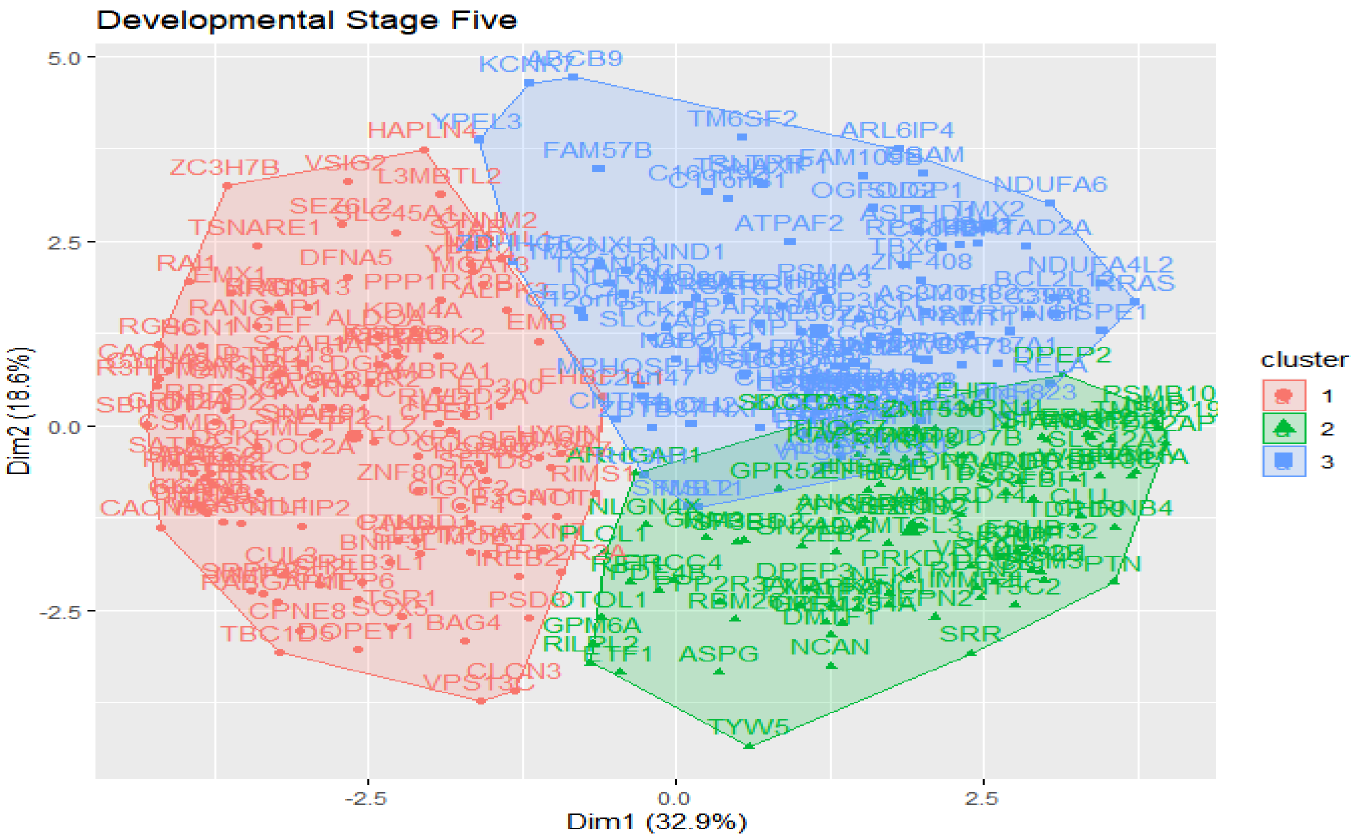


Figure 3.3.6 K-means analysis on the developmental stage Five genes using the kmeans function in R with 3 centres selected as per the NbClust recommendation and nstart= 25.

Table 3.6: Sum of squares for each module in developmental stage Five determined using the kmeans function in R

Cluster	***1***	***2***	***3***
Sum of squares of cluster	1268.6299	952.0207	1120.2446

The total variance of data in module = 33.7 %

In Figure 3.3.6 all of the clusters have high in-cluster variation and a cluster that is not compact, and it can be seen the total variance in the data is low at 33.7% meaning both clusters are similar so a large portion of the 33.7% variation will be explained by the within-cluster variance.

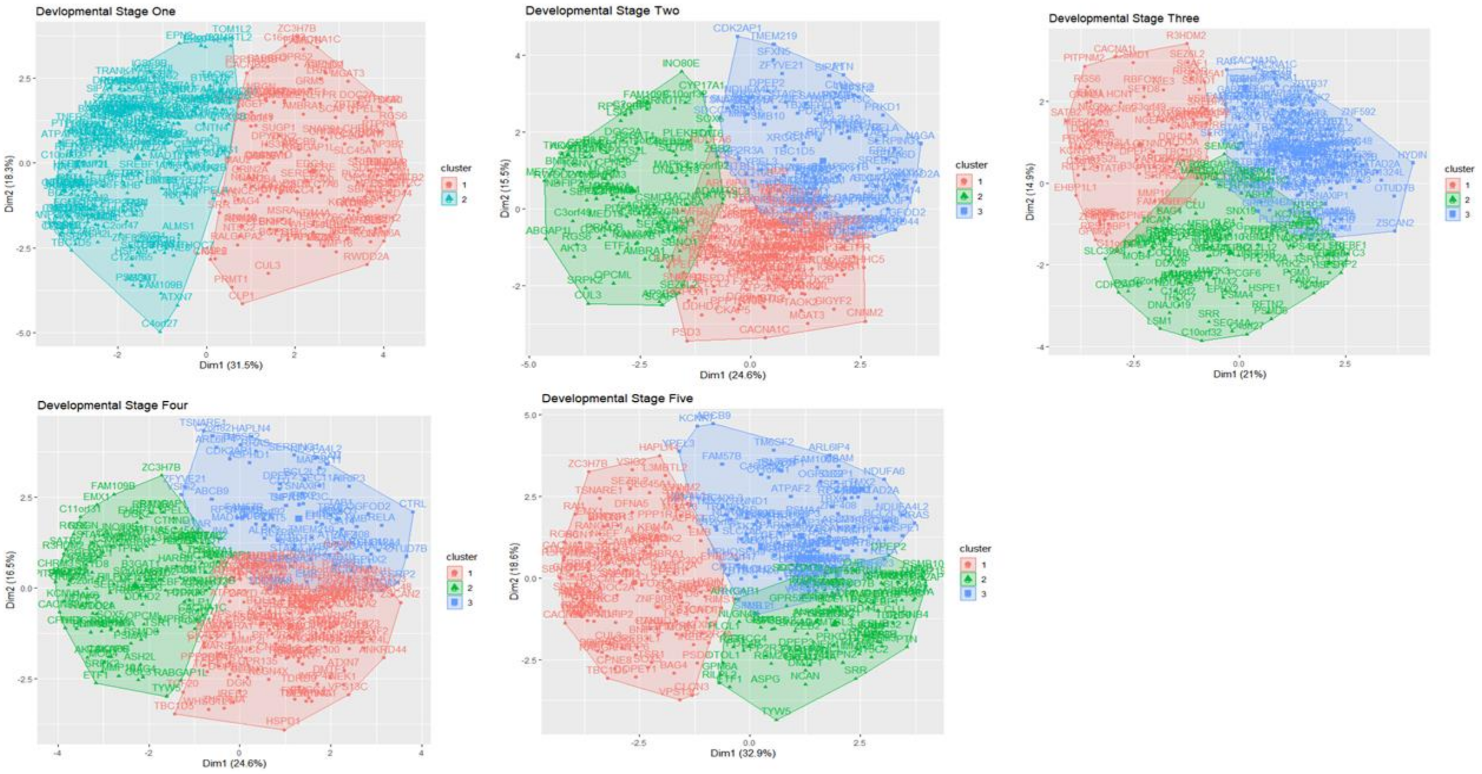


Figure 3.3.7: Kmeans analysis run on schizophrenia genes identified by Pardiñas et al. to determine intrinsic patterns within the genes at each of the developmental stages from ABAs Brainspan resource

3.4 WGCNA on schizophrenia-associated genes and Network Visualisation using Cytoscape

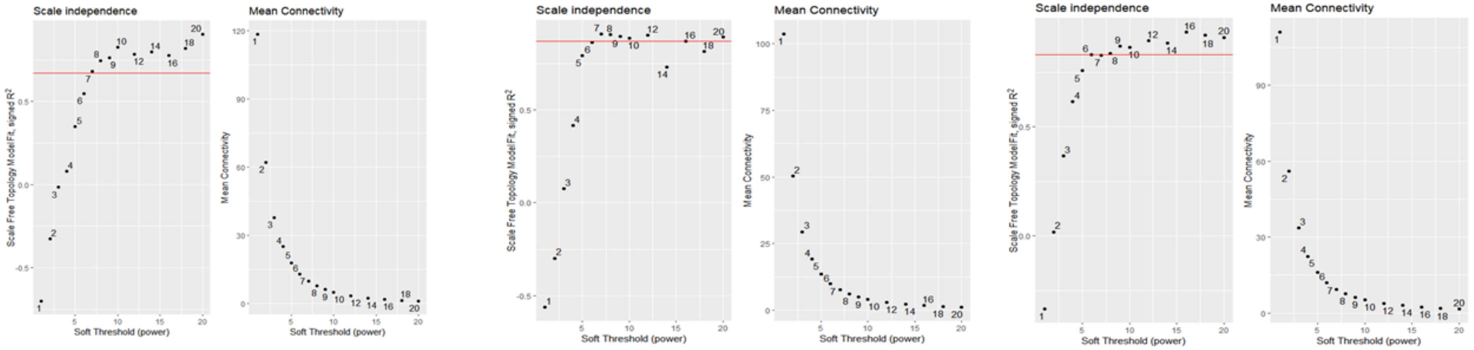
We next employed `WGCNA` to undertake a system-level approach to identify networks of co-expressed modules of schizophrenia-associated genes. `WGCNA` is performed to organise highly correlated genes into gene modules.

First, an analysis of the topology of the networks at various soft thresholding powers was performed separately for developmental stages (results are illustrated in Figure 3.4.1). The soft threshold is calculated to identify the power of the gene correlation should be raised. By raising the correlation to this power, it will reduce the noise of any correlations in the adjacency matrix**.**

Table 3.7: The Soft Thresholding power of each developmental stage calculated using WGCNA and shown in Figure 3.4.1

Developmental Stage	**Soft Thresholding power**
Stage One	7
Stage Two	6
Stage Three	7
Stage Four	9
Stage Five	9

`WGCNA` is used to organise highly correlated genes into modules. Below a gene co-expression network is constructed which is represented by an adjacency matrix which signifies similar co-expression between a gene pair. Hierarchical clustering is used to identify modules and uses topological overlap to measure dissimilarity. Once the schizophrenia-associated genes are separated into modules these modules were summarised by calculating a module eigengene and defining an intramodular hub gene. Then the modules were and visualised using Cytoscape’s `NetworkAnalyze`r. Next, GO enrichment was performed on each of the modules using `anRichment` an R package (107).



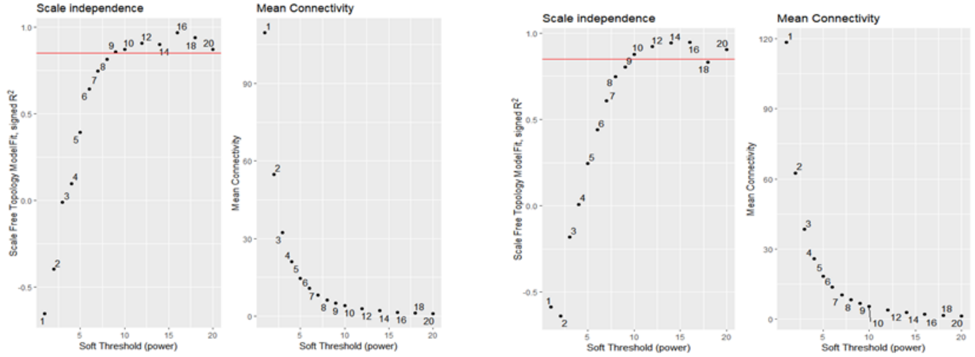
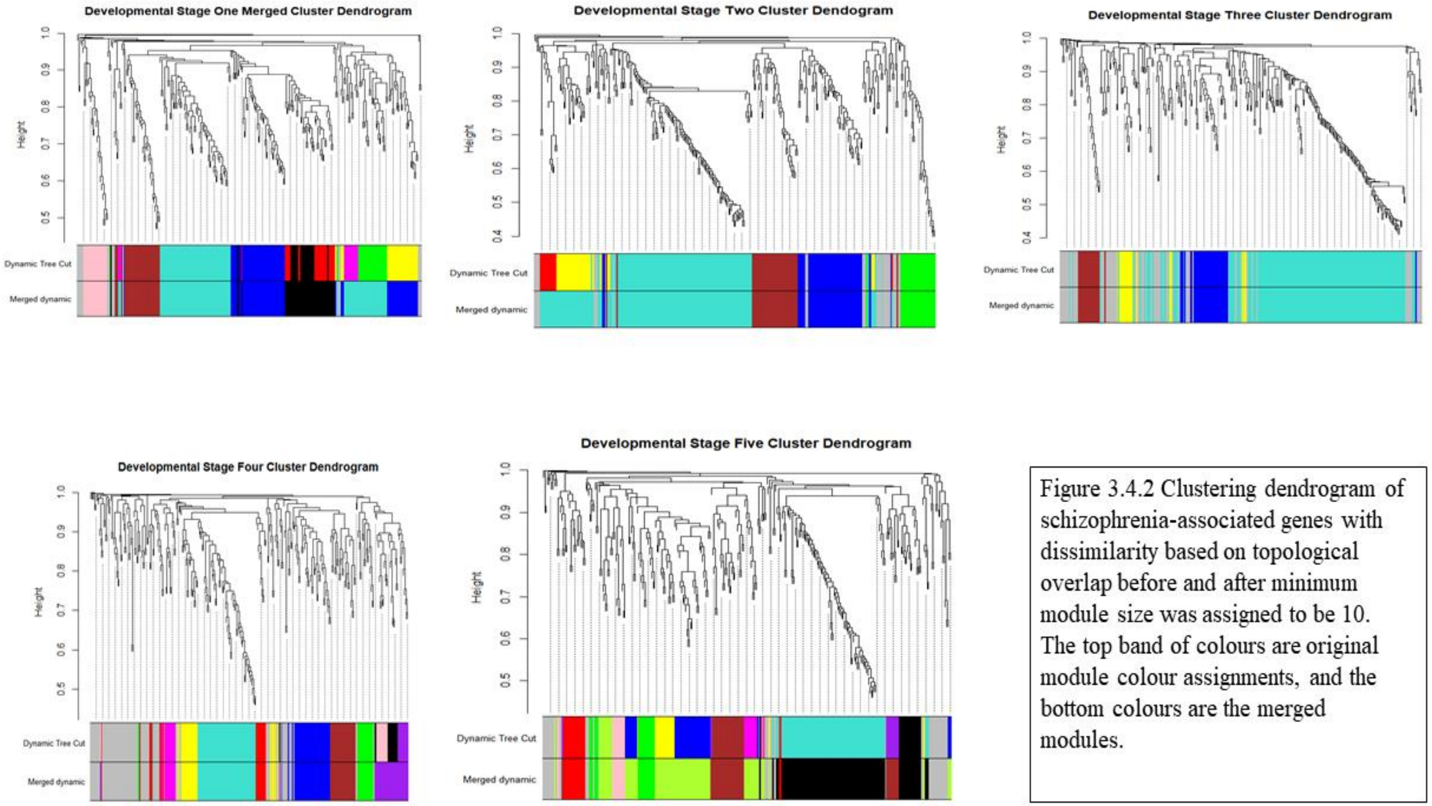


Figure 3.4.1: Analysis of the topology of the networks at various soft-thresholding powers. The scale independence shows the y-axis as a function of the soft-threshold power on the x-axis. The mean connectivity displays the mean as a function of the soft-thresholding power. Developmental stages One through Three are the top (left to right) and developmental stages four and five are on the bottom. The red line through the scale independence graph shows which soft threshold power which was chosen at each developmental stage and can be seen in Table 3.4.1.

After the differential topological matrix is calculated, a gene clustering dendrogram is plotted using the `hclust` function in R. Each leaf of the dendrogram is a gene and after the minimum module is set, the genes may remerge. The dendrogram clusters the branches into coloured modules but some of the modules may need to merge because their genes are highly co-expressed. This is done by calculating eigengenes of each module and re-clustering based on the module eigengene dissimilarity correlations using the `mergeCloseModules` function in WGCNA. Once the modules have merged the module eigengene is re-calculated. WGCNA identified 7 modules in Stage One, 3 modules in Stage Two, 5 modules in Stage Three, 6 modules in Stage Four and 9 modules in Stage Five.

The merged modules and initial modules are illustrated in Figure 3.4.2. After the modules are merged the eigengene is recalculated and in Figure 3.4.3 the adjacency of the eigengene compared to the other eigengenes in the developmental stage is shown.

Network heatmap plot for the developmental stages one through five, using the function `TOMplot` in `WGCNA`, were created to visualise the topological overlap matrix (illustrated in Figure 3.4.4). This TOM matrix uses the adjacency matrix to build another adjacency matrix which takes topological overlap (the number of shared neighbours of the nodes).



Developmental Stage One Merged Cluster Dendrogram

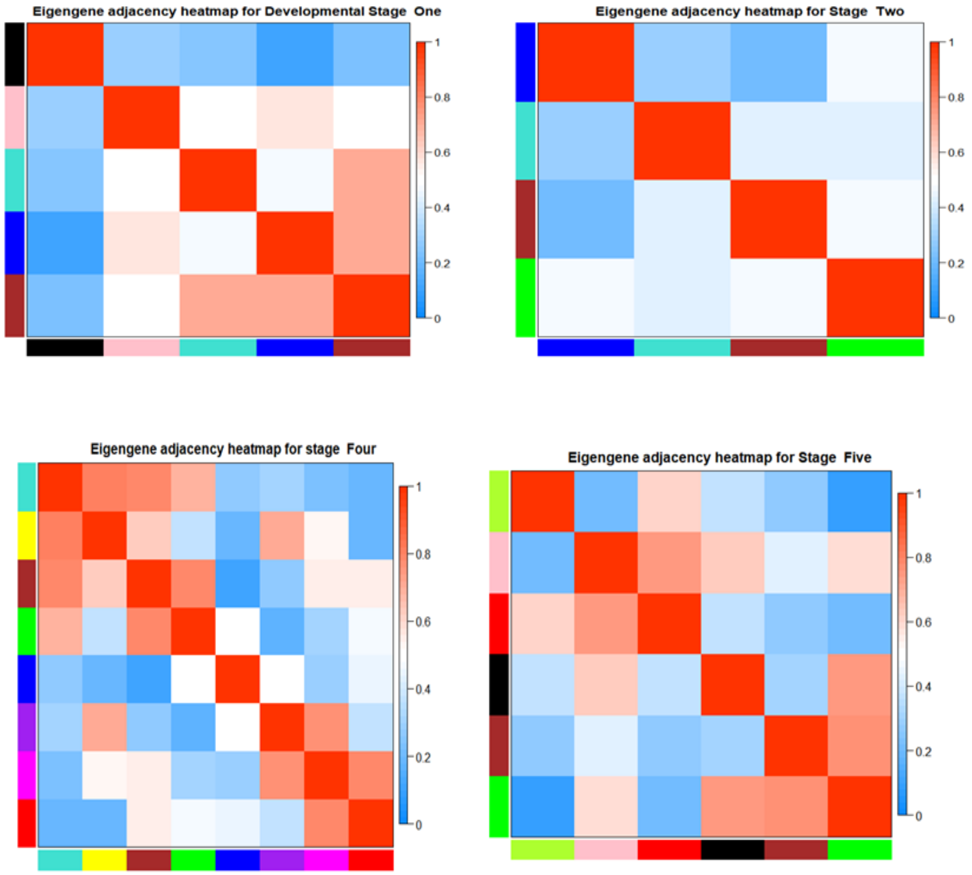
Developmental Stage Two Cluster Dendogram

Developmental Stage Three Cluster Dendrogram

Developmental Stage Four Cluster Dendrogram

Developmental Stage Five Cluster Dendrogram

Figure 3.4.2 Clustering dendrogram of schizophrenia-associated genes with dissimilarity based on topological overlap before and after minimum module size was assigned to be 10. The top band of colours are original module colour assignments, and the bottom colours are the merged modules.



Eigengene adjacency heatmap for Developmental Stage One

Eigengene adjacency heatmap for Stage Two

Eigengene adjacency heatmap for Stage Three

Eigengene adjacency heatmap for stage Four

Eigengene adjacency heatmap for Stage Five

Figure 3.4.3 Eigengene adjacency heatmap for each eigengene network in each developmental stage. Each row and column indicate an eigengene labelled by its module colour. Orange indicates high correlation and blue indicates low correlation.

3.5 Intramodular Hub Genes and Network Analysis

Hub genes (absolute module membership ⩾ 0.8) for each module within each stage were identified using the function `chooseTopHubInEachModule` in the `WGCNA` package. Identified HGs for each module, the function of each gene, as defined by NCBI and their association with disease phenotypes are outlined in Table 9.

Table 3.8: Gene Functions and the phenotypes they are involved in for each hub genes identified by the WGCNA function in R when performed on the schizophrenia-associated genes identified by Pardiñas et al. for the five developmental stages available on ABA's Brainspan.

Module Colour	**Hub Gene**	**Gene Name**	**Genomic Location**	**Function (NCBI gene and Gene ontology)**	**Association with other Conditions**	**References (PMID ID)**
Stage One						
Black	*SOX5*	SRY-Box Transcription Factor 5	NC_000012.12	This gene encodes a member of the SOX (SRY-related HMG-box) family of transcription factors involved in the regulation of embryonic development and the determination of cell fate. The encoded protein may act as a transcriptional regulator after forming a protein complex with other proteins. The encoded protein may play a role in chondrogenesis	Lamb-Shaffer Syndrome and Optic Nerve Hypoplasia, Bilateral.	31578471
Blue	*FHIT*	Fragile Histidine Triad Diadenosine Triphosphatase	NC_000003.12	The protein encoded by this gene is a P1-P3-bis(5'-adenosyl) triphosphate hydrolase involved in purine metabolism. This gene encompasses the common fragile site FRA3B on chromosome 3, where carcinogen-induced damage can lead to translocations and aberrant transcripts. Aberrant transcripts from this gene have been found in about half of all oesophageal, stomach, and colon carcinomas. The encoded protein is also a tumour suppressor, as loss of its activity results in replication stress and DNA damage.	Renal Cell Carcinoma, Nonpapillary and Sporadic Breast Cancer.	28404875
Brown	*SLC12A4*	Solute Carrier Family 12 Member 4	NC_000016.10	The encoded protein controls the movement of potassium and chloride ions across the plasma membrane.	Sickle Cell and Fisheye disease	31792382
Pink	*OTOL1*	Otolin 1	NC_000003.12	Secreted glycoprotein with a C-terminal complement Cq1-like globular domain that belongs to the C1q/tumour necrosis factor-related protein (CTRP) family. The encoded protein is expressed in the inner ear and forms a multimeric complex called the otoconia, together with cerebellin-1 and otoconin-90, as part of the otoconial membrane. It contains extensive posttranslational modifications including hydroxylated prolines and glycosylated lysine's	Benign Paroxysmal Positional Nystagmus and Vestibular Disease	29533337, 31120422

Module Colour	**Hub Gene**	**Gene Name**	**Genomic Location**	**Function (NCBI gene and Gene ontology)**	**Association with other Conditions**
Stage Two					
Blue	*KCNV1*- Potassium Voltage-Gated Channel Modifier Subfamily V Member 1	NC_000008.11	Potassium channel subunit that does not form functional channels by itself. Modulates KCNB1 and KCNB2 channel activity by shifting the threshold for inactivation to more negative values and by slowing the rate of inactivation. Can down-regulate the channel activity of KCNB1, KCNB2, KCNC4 and KCND1, possibly by trapping them in intracellular membranes.	Atrial Septal Defect 5 and Familial Adult Myoclonic Epilepsy	25969726
Brown	*NDRG4*-NDRG Family Member 4	NC_000016.10	This gene is a member of the N-myc downregulated gene family which belongs to the alpha/beta hydrolase superfamily. The protein encoded by this gene is a cytoplasmic protein that is required for cell cycle progression and survival in primary astrocytes and may be involved in the regulation of mitogenic signalling in vascular smooth muscles cells	Infantile Myofibromatosis and Pulmonary Atresia With Ventricular Septal Defect.	31832525, 19711485
Green	*GPR52*- G Protein-Coupled Receptor 52	NC_000001.11	Members of the G protein-coupled receptor (GPR) family play important roles in signal transduction from the external environment to the inside of the cell	Huntington's disease	33796846, 24587241
Turquoise	*STAG1*-Stromal Antigen 1	NC_000003.12	SCC3 family and is expressed in the nucleus. It encodes a component of cohesin, a multi-subunit protein complex that provides sister chromatid cohesion along the length of a chromosome from DNA replication through prophase and prometaphase, after which it is dissociated in preparation for segregation during anaphase.	Mental Retardation, Autosomal Dominant 47 and Cornelia De Lange Syndrome.	2467316, 28430577, 28119487, 32778134

Module Colour	**Hub Gene**	**Gene Name**	**Genomic Location**	**Function (NCBI gene and Gene ontology)**	**Association with other Conditions**	**References (PMID ID)**
Stage Three						
Blue	*SATB2*	SATB homeobox 2	NC_000002.12 (199269500.199471266, complement)	SATB2 encodes for a DNA binding protein that binds specifically at nuclear matrix attachment regions. These regions are involved in chromatin remodelling and transcription regulation.	Glass syndrome (with intellectual disability)	24301056
Brown	*INA*	Internexin Neuronal Intermediate Filament Protein Alpha	NC_000010.11	Neurofilaments are type IV intermediate filament heteropolymers composed of light, medium, and heavy chains. Neurofilaments comprise the axoskeleton and they functionally maintain the neuronal calibre. They may also play a role in intracellular transport to axons and dendrites. This gene is a member of the intermediate filament family and is involved in the morphogenesis of neurons	Gastroenteropancreatic Neuroendocrine Neoplasm and Medulloepithelioma.	29339073
Turquoise	*TRANK1*	Tetratricopeptide repeat and ankyrin repeat-containing 1	NC_000003.12 (36826817.36945662, complement)		Associated with BPD	24309898
Yellow	*ANKRD63*	Ankyrin repeat domain 63	NC_000015.10 (40278372.40282586, complement)			

Module Colour	**Hub Gene**	**Gene Name**	**Genomic Location**	**Function (NCBI gene and Gene ontology)**	**Association with other Conditions**	**References (PMID ID)**
Stage Four						
Blue	*MEF2C*	Myocyte Enhancer Factor 2C	NC_000005.10	Encodes a member of the MADS-box transcription enhancer factor 2 (MEF2) family of proteins, which play a role in myogenesis. The encoded protein, MEF2 polypeptide C, has both trans-activating and DNA binding activities. This protein may play a role in maintaining the differentiated state of muscle cells. Mutations and deletions at this locus have been associated with severe cognitive disability, stereotypic movements, epilepsy, and cerebral malformation	Mental Retardation, Autosomal Dominant 20 and autism spectrum disorder.	32418612, 27779093
Brown	*SMG6*	SMG6 Nonsense Mediated mRNA Decay Factor	NC_000017.11	This gene encodes a component of the telomerase ribonucleoprotein complex responsible for the replication and maintenance of chromosome ends. The encoded protein also plays a role in the nonsense-mediated mRNA decay (NMD) pathway, providing the endonuclease activity near the premature translation termination codon that is needed to initiate NMD	Pancreatic Adenosquamous Carcinoma and Lissencephaly.	25770585
Green	*TAOK2*	TAO Kinase 2	NC_000016.10	Involved in many different processes, including, cell signalling, microtubule organization and stability, and apoptosis.	Wilson-Turner X-Linked Mental Retardation Syndrome and Syndromic X-Linked Intellectual Disability	29467497
Magenta	*OPCML*	Opioid Binding Protein/Cell Adhesion Molecule Like	NC_000011.10	Bind's opioids in the presence of acidic lipids; probably involved in cell contact.	Ovarian Cancer and Hypogonadotropic Hypogonadism 14 With Or Without Anosmia.	29907679, 33777925, 31577955
Purple	*GPR52*	G Protein-Coupled Receptor 52	NC_000001.11	Members of the G protein-coupled receptor (GPR) family play important roles in signal transduction from the external environment to the inside of the cell	Psychiatric disorders	33796846, 24587241

Red	*WHSC1L1/NSD3*	Nuclear Receptor Binding SET Domain Protein 3	NC_000008.11	Histone methyltransferase. Preferentially dimethylates 'Lys-4' and 'Lys-27' of histone H3 forming H3K2me2 and H3K27me2. H3 'Lys-4' methylation represents a specific tag for epigenetic transcriptional activation, while 'Lys-27' is a mark for transcriptional repression	Wolf-Hirschhorn Syndrome and Nut Midline Carcinoma	31190890, 27285764, 25942451
Turquoise	*CA8*	Carbonic anhydrase 8	NC_000008.11 (60185412. 60281400, complement)	In the carbonic anhydrase family but carbonic anhydrase activity (i.e., the reversible hydration of carbon dioxide) The absence of CA8 gene transcription in the cerebellum of the lurcher mutant in mice with a neurologic defect suggests an important role for this acatalytic form.	Mutations in this gene are associated with cerebellar ataxia, mental retardation, and disequilibrium syndrome 3 (CMARQ3). Polymorphisms in this gene are associated with osteoporosis, and overexpression of this gene in osteosarcoma cells suggests an oncogenic role.	19461874
Yellow	*ALMS1*	ALMS1 Centrosome and Basal Body Associated Protein	NC_000002.12	Involved in PCM1-dependent intracellular transport. Required, directly or indirectly, for the localization of NCAPD2 to the proximal ends of centrioles. Required for proper formation and/or maintenance of primary cilia (PC), microtubule-based structures that protrude from the surface of epithelial cells.	Alstrom Syndrome and Premature Ovarian Failure 1.	30421101, 32808654

Developmental Stage	**Hub Gene**	**Gene Name**	**Genomic Location**	**Function (NCBI gene and Gene ontology)**	**Association with other Conditions**	**References (PMID ID)**
Stage Five						
Black	*C16orf86*	Chromosome 16 Open Reading Frame 86	NC_000016.10	Protein Coding Gene		33639916
Brown	*RFTN2*	Raftlin Family Member 2	NC_000002.12	Upon bacterial lipopolysaccharide stimulation, mediates clathrin-dependent internalization of TLR4 in dendritic cells, resulting in activation of TICAM1-mediated signalling and subsequent IFNB1 production. May regulate B-cell antigen receptor-mediated signalling.	Glass Syndrome.	
Green	*NFATC3*	Nuclear Factor of Activated T Cells 3	NC_000016.10	Acts as a regulator of transcriptional activation. Plays a role in the inducible expression of cytokine genes in T-cells, especially in the induction of the IL-2 (PubMed:18815128). Along with NFATC4, involved in embryonic heart development	Crouzon Syndrome with Acanthosis Nigricans and Leukostasis	31249342, 33520407
Greenyellow	*SATB2*	SATB homeobox 2	NC_000002.12 (199269500.199471266, complement)	SATB2 encodes for a DNA binding protein that binds specifically at nuclear matrix attachment regions. These regions are involved in chromatin remodelling and transcription regulation.	Glass syndrome (with intellectual disability)	24301056
Pink	*CHRNA5*	Cholinergic Receptor Nicotinic Alpha 5 Subunit	NC_000015.10	The protein encoded by this gene is a nicotinic acetylcholine receptor subunit and a member of a superfamily of ligand-gated ion channels that mediate fast signal transmission at synapses.	Smoking as a Quantitative Trait Locus 3and Tobacco Addiction	33752734, 30366711, 32817066, 33511332
Red	*INA*	Internexin Neuronal Intermediate Filament Protein Alpha	NC_000010.11	Neurofilaments are type IV intermediate filament heteropolymers composed of light, medium, and heavy chains. Neurofilaments comprise the axoskeleton and they functionally maintain the neuronal calibre. They may also play a role in intracellular transport to axons and dendrites. This gene is a member of the intermediate filament family and is involved in the morphogenesis of neurons	Gastroenteropancreatic Neuroendocrine Neoplasm and Medulloepithelioma.	29339073

3.6 Cytoscape and Network Visualisation

`NetworkAnalyzer`, a Cytoscape plugin, aids in visualisation but can also calculate a network's topological properties (103). Using data loaded into Cytoscape including source nodes, target nodes, the p-value for SNP inclusion identified by Pardiñas et al, weights and node attributes, the network could be visualised using `NetworkAnalyzer`. `NetworkAnalyzer` also calculates the properties of a network as shown in Figure 3.6.1, treating the network as undirected. The clustering coefficient is a ratio of closed triangles over the total open and closed triangles (110). Connected components measure the number of separated fragments in the overall network. Network diameter is the largest number of edges to transverse the network. Network centralization measures the centrality of each node in the network, a network that is highly centralized (= 1) contains a few nodes that dominate the network and without these nodes, the network would become fragmented and leave unconnected sub-networks (111). Network density is a measure of how densely a network is filled with edges, where 0 means there are no edged and 1 means the network is highly populated with edges. Network heterogeneity is a measure of the diversity of the number of connections a node has shown by the node degrees where a homogenous network is equal to 0 and 1 is heterogeneous (112). Networks containing biological data are usually very heterogeneous where most nodes have very few edges apart from HGs which are highly connected (113). The network visualisations for the modules can be seen in the appendix and the network topological parameters are below in Table 3.9 -3.13.

Table 3.9 Network topology parameters calculated by NetworkAnalyzer in Cytoscape for Developmental Stage One.

Module Colour	Cluster Coefficient	Diameter	Network Centralization	Number of Nodes	Network Density	Network Heterogeneity
Black	0.61	5	0.414	47	0.299	0.63
Blue	0.55	6	0.271	74	0.161	0.777
Brown	0.634	5	0.378	32	0.323	0.66
Pink	0.72	4	0.421	20	0.411	0.551
Turquoise	0.477	7	0.258	101	0.127	0.805

Table 3.10 Network topology parameters calculated by NetworkAnalyzer in Cytoscape for Developmental Stage Two.

Module Colour	Cluster Coefficient	Diameter	Network Centralization	Number of Nodes	Network Density	Network Heterogeneity
Blue	0.595	5	0.356	33	0.197	0.654
Brown	0.585	4	0.308	14	0.429	0.554
Turquoise	0.57	11	0.281	144	0.177	0.922
Yellow	0.365	6	0.21	16	0.217	0.62

Table 3.11 Network topology parameters calculated by NetworkAnalyzer in Cytoscape for Developmental Stage Three

Module Colour	Cluster Coefficient	Diameter	Network Centralization	Number of Nodes	Network Density	Network Heterogeneity
Blue	0.595	5	0.356	33	0.197	0.654
Brown	0.585	4	0.308	14	0.429	0.554
Turquoise	0.57	11	0.281	144	0.177	0.922
Yellow	0.365	6	0.21	16	0.217	0.62

Table 3.12 Network topology parameters calculated by NetworkAnalyzer in Cytoscape for Developmental Stage Four.

Module Colour	Cluster Coefficient	Diameter	Network Centralization	Number of Nodes	Network Density	Network Heterogeneity
Blue	0.636	4	0.348	38	0.265	0.558
Brown	0.578	5	0.274	25	0.29	0.417
Green	0.373	4	0.2125	17	0.25	0.522
Magenta	0.397	4	0.327	12	0.273	0.593
Purple	0.582	4	0.338	32	0.264	0.553
Red	0.37	4	0.359	14	0.308	0.598
Turquoise	0.708	4	0.316	58	0.379	0.537
Yellow	0.648	5	0.379	21	0.357	0.512

Formatted: Table

Table 3.13 Network topology parameters calculated by NetworkAnalyzer in Cytoscape for Developmental Stage Five.

Module Colour	Cluster Coefficient	Diameter	Network Centralization	Number of Nodes	Network Density	Network Heterogeneity
Black	0.643	5	0.289	104	0.202	0.781
Brown	0.642	5	0.244	36	0.256	0.522
Green	0.614	4	0.379	19	0.327	0.634
Greenyellow	0.544	6	0.262	77	0.179	0.678
Pink	0.538	4	0.309	12	0.379	0.425
Red	0.44	4	0.288	19	0.298	0.51

3.7 Gene Ontologies

Gene ontologies that are enriched for each module within each stage were identified using the function `enrichmentAnalysis` which uses the GO collection database in the anRichment Bioconductor package in R. Finding GOs which are enriched in each of the modules can deepen the understanding of schizophrenia and lead to the discovery of novel therapeutic targets.

Table 10 – Table 14 displays the most enriched gene ontologies per module per stage. In the appendix Tables 10 – 27 illustrate the full gene ontologies identified by anRichment. It is clear from the ontology results that positive regulation of macrophage proliferation is very important as it appears in the top 3 ontologies in a significant number of the modules per Developmental Stage. Cellular response to catecholamine stimulus and cellular response to diacyl bacterial lipopeptide appear in the top three ontologies at least once per stage also. These enriched ontologies are candidates for future investigation.

Table 3.9: Gene Ontologies of the top enriched gene ontologies in Stage One using anRichment

GOID	DEFINITION	ONTOLOGY	Module	GO Process/ Term	FDR
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Black	positive regulation of macrophage proliferation	2.92E-15
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Black	binding	1.88E-10
GO:0043231	The organised structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Black	intracellular membrane-bounded organelle	6.74E-10
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	1.75E-29
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Blue	intracellular organelle lumen	2.74E-14
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl) benzene-1,2-diol] and derivatives formed by substitution.	BP	Blue	cellular response to catecholamine stimulus	4.03E-13
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	2.10E-15
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Brown	binding	3.36E-10
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	4.32E-09
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Pink	protein binding	1.61E-08
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Pink	positive regulation of macrophage proliferation	2.30E-07
GO:0008152	The chemical reactions and pathways, including anabolism and catabolism, by which living organisms transform chemical substances. Metabolic processes typically transform small molecules, but also include macromolecular processes such as DNA repair and replication, and protein synthesis and degradation.	BP	Pink	metabolic process	6.41E-07
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	9.49E-49
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	1.07E-17
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain	MF	Turquoise	armadillo repeat domain binding	4.10E-17

	of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.				

Table 3.10: Gene Ontologies for the most enriched ontologies in Stage Two using anRichment

GOID	DEFINITION	ONTOLOGY	Module	GO Process/ Term	FDR
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	9.38E-19
GO:0043231	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Blue	intracellular membrane-bounded organelle	2.90E-12
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Blue	cytoplasm	1.74E-10
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	7.94E-15
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	2.07E-09
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Brown	cytoplasm	9.37E-09
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Green	molecular function	2.31E-12
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Green	positive regulation of macrophage proliferation	7.16E-10
GO:1902644	The chemical reactions and pathways involving tertiary alcohol	BP	Green	tertiary alcohol metabolic process	2.55E-09
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	1.10E-69
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.	MF	Turquoise	armadillo repeat domain binding	3.94E-28
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	4.57E-25

Table 3.11: Gene Ontologies for the most enriched ontologies in Stage Three using anRichment

GOID	DEFINITION	ONTOLOGY	Module	GO Process/ Term	FDR
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	2.37E-12
GO:0043231	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Blue	intracellular membrane-bounded organelle	1.70E-10
GO:0044238	The chemical reactions and pathways involving those compounds are formed as a part of the normal anabolic and catabolic processes. These processes take place in most, if not all, cells of the organism.	BP	Blue	primary metabolic process	1.03E-09
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	4.06E-08
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	8.00E-06
GO:0099582	Any neurotransmitter receptor activity that is involved in regulating the concentration of calcium in the presynaptic cytosol.	MF	Brown	neurotransmitter receptor activity involved in the regulation of presynaptic cytosolic calcium ion concentration	1.25E-05
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	5.64E-77
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl) benzene-1,2-diol] and derivatives formed by substitution.	BP	Turquoise	cellular response to catecholamine stimulus	4.05E-29
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	1.01E-23
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Yellow	binding	3.20E-07
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Yellow	positive regulation of macrophage proliferation	3.44E-07
GO:0044237	The chemical reactions and pathways by which individual cells transform chemical substances.	BP	Yellow	cellular metabolic process	1.60E-06

Table 3.12: Gene Ontologies for the most enriched ontologies in Stage Four using anRichment

GOID	DEFINITION	ONTOLOGY	Module	GO Process/ Term	FDR
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	1.70E-10
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Blue	cytoplasm	4.22E-09
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Blue	binding	4.46E-08
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	1.89E-08
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Brown	binding	4.85E-07
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	2.09E-06
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Green	positive regulation of macrophage proliferation	3.43E-06
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Green	intracellular organelle lumen	9.16E-05
GO:0035373	Interacting selectively and non-covalently with a chondroitin sulphate proteoglycan, any proteoglycan containing chondroitin sulphate as the glycosaminoglycan carbohydrate unit.	MF	Green	chondroitin sulphate proteoglycan binding	0.000174
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Magenta	cytoplasm	7.51E-05
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Magenta	molecular function	0.000264
GO:0019222	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways within a cell or an organism.	BP	Magenta	regulation of metabolic process	0.000537
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Purple	positive regulation of macrophage proliferation	4.37E-10
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Purple	intracellular organelle lumen	6.00E-08
GO:1901666	Any process that activates or increases the frequency, rate, or extent of NAD+ ADP-ribosyltransferase activity	BP	Purple	positive regulation of NAD+ ADP-ribosyltransferase activity	2.40E-07
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Red	positive regulation of macrophage proliferation	3.43E-06
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Red	protein binding	0.0004

GO:0043412	The covalent alteration of one or more monomeric units in a polypeptide, polynucleotide, polysaccharide, or other biological macromolecules, resulting in a change in its properties.	BP	Red	macromolecule modification	0.000479
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	3.42E-25
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Turquoise	intracellular organelle lumen	2.23E-12
GO:0005634	A membrane-bounded organelle of eukaryotic cells in which chromosomes are housed and replicated. In most cells, the nucleus contains all of the cell's chromosomes except the organellar chromosomes and is the site of RNA synthesis and processing. In some species of specialized cell types, RNA metabolism or DNA replication may be absent.	CC	Turquoise	nucleus	5.73E-11
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Yellow	molecular function	6.12E-08
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Yellow	positive regulation of macrophage proliferation	2.55E-07
GO:0043227	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Yellow	membrane-bounded organelle	9.65E-07

Table 3.13: Gene Ontologies for the most enriched ontologies in Stage Five using anRichment

GOID	DEFINITION	ONTOLOGY	Module	GO Process/ Term	FDR
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Black	positive regulation of macrophage proliferation	2.43E-38
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Black	cellular response to diacyl bacterial lipopeptide	3.66E-14
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl) benzene-1,2-diol] and derivatives formed by substitution.	BP	Black	cellular response to catecholamine stimulus	3.49E-13
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	1.14E-11
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Brown	binding	1.07E-09
GO:0044237	The chemical reactions and pathways by which individual cells transform chemical substances.	BP	Brown	cellular metabolic process	9.83E-09
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Green	positive regulation of macrophage proliferation	5.52E-06
GO:0098809	Catalysis of the reaction: nitrite + acceptor = product(s) of nitrate reduction + reduced acceptor.	MF	Green	nitrite reductase activity	1.04E-05
GO:0043226	Organized structure of distinctive morphology and function. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes and the cytoskeleton, and prokaryotic structures such as anammoxosomes and pirellulosomes. Excludes the plasma membrane.	CC	Green	organelle	1.33E-05
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Greenyellow	positive regulation of macrophage proliferation	1.22E-34
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Greenyellow	cellular response to diacyl bacterial lipopeptide	3.08E-18
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Greenyellow	collagen V binding	7.65E-15
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Pink	positive regulation of macrophage proliferation	0.000262
GO:0019866	The inner, i.e. lumen-facing, the lipid bilayer of an organelle envelope; usually highly selective to most ions and metabolites.	CC	Pink	organelle inner membrane	0.000645
GO:0031966	Either of the lipid bilayers surrounds the mitochondrion and form the mitochondrial envelope.	CC	Pink	mitochondrial membrane	0.001661
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Red	positive regulation of macrophage proliferation	1.84E-07

GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Red	protein binding	1.43E-05
GO:0070013	An organelle lumen is part of an intracellular organelle.	CC	Red	intracellular organelle lumen	3.82E-05

4.0 Discussion

Schizophrenia is a chronic and disabling disorder that affects 1% of the general population whose causes remain unclear even with much research into disease aetiology (4). Schizophrenia is a neurodevelopmental disorder, therefore, constructing and exploring networks of genes, previously identified as associated with schizophrenia, over key developmental stages could aid our understanding of schizophrenia's aetiology and identify novel therapeutic targets of disease.

Although schizophrenia is researched extensively from many angles its mechanisms remain elusive. It is difficult to predict who will develop schizophrenia because of the complex interactions between genetic and environmental factors. Schizophrenia is classed as a neurodevelopmental disorder but as it doesn't manifest until adolescence it's very important to study it at multiple clinical stages (13,53,162). Insults to brain development *in utero* can have an impact on the severity of schizophrenia's symptoms later in life. Changes in gene expression throughout schizophrenia have been observed, the study performed by Ota et al. looked at participants blood and compared expression levels of schizophrenia genes in clinically high-risk patients through to chronic schizophrenia. They observed changes in gene expression profiles at different clinical stages (13). Studies like these highlight the importance of looking at schizophrenia at multiple stages to captures its heterogeneity.

In this study, genes previously identified in a large-scale schizophrenia GWAS as being significantly associated with schizophrenia were used to filter ABA's BrainSpan resource to include gene expression data for these genes across 16 brain areas and five developmental stages. To find underlying patterns in the gene expression data K-means analysis was utilised. A systems biology approach (WGCNA) was used to describe the pairwise relationship between genes at each development stage and to create networks (i.e., modules) of schizophrenia-associated genes which were co-expressed in each

Developmental Stage. Next, GO enrichment analysis was applied to each of the modules using anRichment to aid biological interpretation of the identified networks in each developmental stage.

4.1 K-means analysis on the schizophrenia-associated genes

K-means clustering was performed on each Developmental Stage. Each schizophrenia-associated gene was assigned to the closest centre (k). Developmental Stage One to Five can be seen in Figure 3.3.2 to 3.3.7 and a comparison of the five stages in 3.3.7.

If a cluster is filled with genes that are very similar the within-cluster sum of squares of the cluster will be small and when the cluster is visualised it will appear small and compact. Each of the clusters in this analysis had a large within-cluster sum of squares, this means there is variance within the cluster. The total within variance measures the deviation from the mean. The total within variance in this dataset is low which means that the dataset is very similar. Usually, only a small amount of the total variance is explained by the within-cluster sum of squares but in this case, the total variance is low and the within the sum of squares is high. This tells us the schizophrenia-associated genes expression profile in the brain are similar to each other. From the k-means results, it is clear that these genes do not cluster well together, if you increase the number of clusters for each of the Developmental Stages the sum of squares with each cluster decreases and the total variance increases but this would mean ignoring the optimal cluster number for each stage previously calculated.

4.2 Weighted Gene Correlations Network Analysis on schizophrenia-associated genes

WGCNA is a systems biology method that uses gene transcripts to describe pairwise relationships between the genes. WGCNA was used to calculate modules of schizophrenia-associated genes which were co-expressed for each developmental stage. WGCNA identified three modules in developmental stage One and Five, four in developmental stages Two and two in developmental Stages Three and Four. The ME of each cluster was identified and the gene most like the ME is classed as the HG. Each HG can be seen in Table 9. WGCNA relies strongly on the assumption that gene co-expression networks follow a scale-free topology where highly connected genes are essential for a functioning system (163). It assumes that gene products associated with the same phenotype usually participate in the same module (163). When WGCNA was applied to the expression profiles of the 316 schizophrenia-associated genes in the brain, each gene was assigned to a module and each module was given a colour. A hierarchical clustering graph of initial and final (after minimum module size of 10) module assignments for each Developmental Stage can be seen in Figure 3.4.2. After module construction, each genes module membership and weight were calculated for downstream visualisation using Cytoscape.

4.3 Visualisation of modules using data from WGCNA.

First, genes were assigned to modules. Using each gene's MM the ME and thus HG was identified, and the weight calculated. This information was collated into csv files, which were uploaded to Cytoscape so the modules could be visualised. Cytoscape allows for the important attributes to be highlighted in the network graph. In each graph node size is shown by the module membership, edge width signifies weight and node colour displays p-value as calculated by Pardiñas et al. These p-values range from 2.12×10^{-44} – 4.88×10^{-08} the lower the value the greater the association. Each Cytoscape network can be seen in Figure 3.6.1 to 3.6.27. Only the edges width of 0.8 and over were included to capture the most important edges.

4.3.1 Developmental Stage One - Prenatal

BrainSpan's Developmental Stage One includes gene expression for healthy prenatal brains. Figure 3.6.1 shows the Black module which contains 47 genes. The HG for this module is *SOX5*. This gene encodes a transcription factor that mediates DNA binding and nuclear trafficking (114). *SOX5* modulates the timing of important processes during corticofugal neuron production so consequently neocortical neuron diversity and sub-type specific differentiation (164). *SOX5* has been linked to Lamb Shaffer syndrome which is a developmental disorder that involves ID, language, and motor deficits and serves various roles in multiple cancer types (165,166). Figure 3.6.2 shows the blue module and contains 74 genes, two of which have split into a small module. The HG for this module is *FHIT*, this genes inactivation, deletion and decreased expression is seen in most cancers (116). It can induce cell apoptosis, stunt the cell growth cycle which can impede tumour proliferation (167). Figure 3.6.3 displays the Brown module which contains 32 genes and whose HG is *SLC12A4* codes for KCC1. KCC1 is a protein that facilitates the symport of chlorine and potassium through cells surfaces and is involved in cancer growth, bone turnover and sickle cell formation (117). Na-K-Cl and K-Cl co-transporters regulate Cl levels, shifts in chloride electrochemical gradients can affect GABAergic transmission which is seen to be impaired in schizophrenia especially in ANPs (168,169). Figure 3.6.4 shows the pink module with 20 genes, the HG *OTOL1* which is essential for hearing and vestibular function (119,120). It is not known what function *OTOL1* has in the brain. Figure 3.6.5 shows the large turquoise module which contains 101 genes. The HG for the Turquoise module is *TMEM194A* also known as *NEMP1*. This gene is poorly understood but the NEMP1 protein has been linked to eye development from the gastrula through to the neurula stage (123).

Stage One is the prenatal stage, *SOX5* being associated with a neurodevelopmental disorder and *NEMP1* being linked to eye development are interesting. There are several physiological and structural impairments in the eye associated with schizophrenia. The retina and the brain develop from the same tissue the neuroectoderm and it has been proposed that retinal changes can be a marker for progressive brain tissue loss and function (170). This may mean that retinal structure changes could parallel the brain (171). Performing a longitudinal study observing retinal changes and linking them to schizophrenia symptoms could potentially map the progression of schizophrenia.

4.3.2 Developmental Stage Two – Infant (0-2 years)

BrainSpan's Developmental Stage Two includes gene expression for a healthy infant's brain from 0-2 years old. 3.6.6 shows the blue module which has 43 genes, the HG is *KCNV1*, this gene is involved in another neurodevelopmental disorder autism spectrum disorder (ASD) (124). It codes for a subunit that regulates potassium channels which regulates neurotransmission release, neuronal excitability, and epithelial electrolyte transport controlled by BK channels (125). The BK channels have been linked to schizophrenia (172). Figure 3.6.7 displays the brown module with 31 genes and whose HG is *NDRG4* which is involved in cell proliferation, differentiation, and development (127). *NDRG4* deficient mice are at a greater risk of cerebral ischemia and exhibit poor spatial learning (127,173). Knockout of *NDRG4* induces glioblastoma cell apoptosis which contributes to neurological damage (127). Figure 3.6.8 shows the green module and has 26 genes, two of the genes in the modules have split into their submodule. The HG for the module is *GPR52* which is a G protein-coupled receptor, it is thought that it could regulate dopaminergic and glutamatergic transmission which is responsible for cognitive function and is affected in schizophrenia (128). Figure 3.6.9 displays the turquoise module; this module has six sub-modules and contains 147 genes in total. The HG for the module is *STAG1* is needed for cohesion at telomeres and DNA replication (131). STAG1 deletions and point mutations can attribute to syndromic unspecific ID (132).

4.3.3 Developmental Stage Three – Child (3- 11 years)

Developmental Stage Three contains gene expression across the brains of children between the ages of 3 and 11 years. Figure 3.6.10 displays the blue module and contains one connected component and contains 33 genes. The HG for the module is *SATB2,* which is involved in transcription regulation and chromatin remodelling (133). *SATB2* is a transcription factor that regulates neocortical circuitry and organisation (133). *SATB2* has been shown to cause *SATB2*-associated syndrome and developmental delays (133,134). Figure 3.6.11 shows the brown module which contains 14 genes and whose HG is *INA.* It is found in developing neuroblasts and in cerebellar granule in the adult CNS (135,174). *INA* maintains the morphogenesis of neurons (135). Figure 3.6.12 shows the turquoise module and has 144 genes and whose HG is *TRANK1*. *TRANK1* encodes a protein in the brain with unknown function (175). Decreased expression of *TRANK1* affected the expression of several genes which are involved in neural development (175). Low levels of *TRANK1* mRNA expression is a BP risk factor (175,176). The yellow module is shown in Figure 3.6.13 and has16 genes in the module. The HG is *ANKRD63* but not much is known about the function of this gene.

There is not a clear link between the HGs in Developmental Stage Three. *SATB2*-syndrome is a neurodevelopmental disorder that leads to developmental delays, it also regulates chromatin remodelling. Defects in chromatin remodelling are often seen in neurodevelopmental disorders, these defects can compound over time and impair brain circuit establishment (177). These impairments in neurodevelopmental disorders can lead to decreased cognitive function. Although not much is known about the exact function of *TRANK1* it has been linked to BP, schizophrenia and BP share some symptomatology and significant genetic overlap with a genetic correlation from common risks estimated to be 0.6-0.7, and 114 loci contributing to both (178).

4.3.4 Developmental Stage Four – Adolescent (12-18 years)

Developmental Stage Four captures gene expression in the brain for adolescents. Typically schizophrenia manifests its symptoms during this time frame (179). Figure 3.6.14 illustrates the blue Module with 38 genes. The HG for this module is *MEF2C*, this gene regulates expression across development in processes such as synapse formation and development, neuronal differentiation (140). It has also been shown to be involved in numerous neurodevelopmental disorders including ASD (180). Figure 3.6.15 displays the brown module with 25 genes. This module's HG is *SMG6*, although there is not much known about this gene's exact function. Still, it has been linked to pancreatic adenosquamous carcinoma and lissencephaly (181). *SMG6* hyperfunction has been shown in epileptic seizures (182). Figure 3.6.16 displays the green module with 17 genes, this module has very few edges and is filled with genes with very low p-values showing a statistically strong link to schizophrenia. The HG for this module is *TAOK2* which regulates neurodevelopment, synapse formation and the development of synapses through modulation of the cytoskeleton which can be linked to schizophrenia (141). *TAOK2* has also been seen to regulate apoptosis by activating T cells and is a neurodevelopmental risk gene (141). Figure 3.6.17 shows the magenta module which contains only 12 genes. *OPCML* is the HG for this module and has been connected to the lung, brain, and cervical cancers (143,145). This gene regulates synaptogenesis and synaptic plasticity. Figure 3.6.18 shows the purple module with 32 genes. The HG is *GPR52* which is a G protein-coupled receptor, it may regulate dopaminergic and glutamatergic transmission which is responsible for cognitive function (129). Figure 3.6.19 displays the red module which is a small module with 14 genes. The HG of this module is *WHSC1L1/NSD3*. This gene has been linked to numerous cancers, cell cycle progression and promotes antiviral innate immunity (146,147). Figure 3.6.20 presents the turquoise module with 58 genes. The HG is *CA8* which is a member of the carbonic anhydrase family, it is expressed in Purkinje cells in the cerebellum. It is a IP_3R1 inhibitor that regulates calcium levels which help key cellular processes (150,183). Figure 3.6.21 shows the yellow module with 21 genes. The HG is *ALMS1* which is involved in the maintenance of centromere cohesion, transcription and actin organisation and endosomal trafficking. Mutations in ALMS1 can cause Alström syndrome which can sometimes include ID and can cause psychotic like symptoms in adults which is also when psychotic symptoms appear in schizophrenia (184).

MEF2C and *TAOK2* are both involved in synapse formation and development. Dysregulated synaptic development has been hypothesised as underlying altered neuronal function in schizophrenia (185). Synaptic pruning occurs during adolescence Developmental Stage Four (12 – 18 years old) 30% of synapses which are formed during adolescence in the dorsolateral prefrontal cortex are lost but in people with schizophrenia it is closer to 60% (186). The synaptic pruning roughly ends at the time of schizophrenia onset leading to the hypothesis that altered synaptic pruning may be a part of the pathophysiology of schizophrenia (185).

4.3.5 Developmental Stage Five – Adult (>19 years)

Figure 3.6.22 shows the Black module which has 104 genes that have formed their submodule. The HG for this module is *C16orf86* which is associated with insulin sensitivity in human skeletal muscles, insulin sensitivity in a subgroup of patients with schizophrenia who are anti-psychotic resistant (155). Figure 3.6.23 displays the brown module with 36 modules. The modules HG is *RFTN2* is implicated in glass syndrome which is characterised by developmental delay, speech development and ID. Glass syndrome has some symptom similarities to schizophrenia which would suggest some shared underlying pathways. Figure 3.6.24 shows the green module with 19 genes. Its HG is *NFATC3*, a member of the calcineurin nuclear factor of activation of T cells who play an essential role in the immune system (187). It is a transcription factor that is involved in the development and progression of tumours and is important for brain tissue homeostasis (158) and is a prompter of neural progenitor cell differentiation into neurons and astrocytes (188). Figure 3.6.25 shows the Greenyellow which houses 77 genes. The modules HG is *SATB2* which encodes a protein that is involved in transcription regulation and chromatin remodelling (133). *SATB2* is a transcription factor that regulates neocortical circuitry and organisation (133). *SATB2* has been shown to cause *SATB2*-associated syndrome and developmental delays (133,134). Figure 3.6.26 displays that the pink module houses only 12 genes. *CHRNA5* is the HG for this module which is a nicotinic acetylcholine receptor, three other members of the family are in this module, altered cholinergic neural transmission has been shown to increase susceptibility to cognitive deficits, and in the Chinese Han population, it has been linked to early-onset and more severe symptoms of schizophrenia (160,189). Figure 3.6.27 shows the red module and has 19 genes. *INA* is the HG in this module which is found in developing neuroblasts and in cerebellar granule in the adult CNS and *INA* maintains the morphogenesis of neurons (135). It is a prognostic marker for poor survival rates in colorectal cancer patients(190).

RFTN2 and *SATB2* are both involved in glass syndrome that is also known as SATB2-syndrome, which is characterised by ID, craniofacial abnormalities, dental abnormalities, and behavioural problems. *CHRNA*5 has been shown to increase cognitive deficits, these and the genes involved in the *SATB2* syndrome are linked to the cognitive and negative symptoms of schizophrenia.

4.3.6 Recurring Hub Genes Across Developmental Stages

Of the 316 schizophrenia-associated genes 27 modules were formed and three genes repeated over two modules, *GPR52*, *INA* and *SATB2*.

GPR52 is an orphan G- protein-coupled receptor (GPCR) which is selectively expressed in the striatum and regulates various brain functions including homeostasis, immune function, neurotransmission and metabolism (191,192). It is expressed in the striatum and nucleus accumbens which have been linked with psychiatric disorders. GPCR's are mediators of signal transduction in the CNS and have been actively investigated for their role in the development of mood disorders (193). GPCR's are the most common targets of antipsychotics and play crucial roles in controlling brain function by regulating numerous downstream signalling pathways (192). *GPR52* has been highlighted as a potential therapeutic target for schizophrenia, it is thought that *GPR52* signalling via 5'-cyclic adenosine monophosphate (cAMP) could oppose D2 signalling activity in the striatum while stimulating D1/NMDA function in the frontal cortex (194,195). The role of *GPR52* in the dopaminergic system may aggravate the symptoms of schizophrenia (193).

SATB2 is a transcription factor that regulates neocortical organisation and circuitry (133). *SATB2* is required for the projection of upper-layer neurons and it can regulate other genes by mediating chromatin loop formation and can modify higher-order chromatin structure (133). It can control the expression of genes that are involved in pluripotency and self-renewal (196). During CNS development the SATB2 protein is expressed in the superficial cortical layers and determines neuron projection. In adult CNS it is expressed in pyramidal neurons of all cortical layers and regulates long-term memory and synaptic plasticity which is linked to cognition (197). *SATB2* facilitates callosal projection by repressing the *BCL11B* gene whose protein is required for subcortical projection neuron identity and postnatal development of the hippocampus (133). This can lead to altered cognition which is seen in schizophrenia. Loss of *BCL11B* leads to weakened hippocampal memory and learning (133). De novo structural or single-nucleotide variants in *SATB2* is linked to *SATB2*-associated syndrome which is characterised by intellectual disability, developmental delays, abnormal craniofacial features and behavioural issues (133,197).

INA (α-Internexin) codes a class IV neuronal intermediate filament protein which maintains neuronal morphogenesis and provides strength to the cell (135). This protein is a structural component of the cytoskeleton and is involved in neurogenesis (136). INA regulates the expression of other neurofilaments during brain development (198). It is widely accepted that *INA* is involved in neuronal development, but its function remains unknown. *INA* is the main component of neuronal IF inclusion disease which causes rare frontal dementia and behavioural and personality changes and has been linked to brain tumours (135). The INA protein is believed to be involved in tumour initiation and progression and is one of the most overexpressed proteins in gliomas (174,190).

INA and *SATB2* are repeated in Developmental Stage Three and Developmental Stage Five and *GPR52* is repeated in Developmental Stage Two and Four. *GPR52* is thought to aggravate schizophrenia's positive symptoms, and *SATB2* and *INA* mutations can cause symptoms that look like schizophrenia's negative and cognitive symptoms. Our results suggest that *INA*, *SATB2* and *GPR52* represent candidate genes for future evaluation of their potential as a therapeutic target for schizophrenia.

Pardiñas et al. study applied Summary-data-based Mendelian Randomisation (SMR) analysis to the schizophrenia-associated gene with dorsal prefrontal cortex expression quantitative trait locus (eQTL) using CommonMind Consortium. This analysis aimed to uncover variants that could be causally linked to expression changes in specific genes. They applied a threshold of 0.05 which highlighted colocalised signals due to a single casual variant. From this, they discovered 22 candidate variants at 19 loci with an FDR P <0.05. *ALMS1* the HG in the green Module in Developmental Stage Two and *TRANK1* the HG in the brown Module of Developmental Stage Three identified in our study were also identified in the SMR analysis. Thus our results suggesting these genes are HGs adds further evidence that they would be good candidates for further research as they may regulate several schizophrenia-related genes in their respective modules.

4.4 Gene Ontologies

On each of the modules calculated by WGCNA, anRichment was run. anRichment is run to biologically interpret the modules. P-values were calculated and GO's which were lower than the threshold of 0.05 after Bonferroni correction was applied were retained. As anRichment produced many significant GO's, GO parent terms were excluded to focus on more specific pathways. The complete ontologies calculated from anRichment in R including the genes involved can be found in the appendix from Table 3 to Table 27.

Across the developmental stages, anRichment has highlighted the immune system and inflammation, specifically concerning macrophage proliferation (GO:00120041). Inflammation in the CNS is facilitated by astrocytes, microglial cells, proinflammatory cytokines, invading immune cells which includes macrophages, monocytes, and T or B lymphocytes. For appropriate function, a well-regulated inflammatory response is essential, but uncontrolled inflammation caused by infectious agents, genetics or physical trauma can be detrimental (199). The macrophages of the brain are called microglia and they play a crucial role in the innate immunity of the CNS and represent up to 10% of total brain cells, but their cell density depends on the area of the brain (200,201). Microglial cells are derived from the yolk sac progenitors during embryogenesis and migrate throughout the CNS, they are maintained through adulthood by self-renewal and rapid cell turnover (201,202). Microglia are involved in the synaptic organisation, phagocytosis of apoptotic cells during development, maintenance of neuronal excitability, trophic neuronal support in the developing brain and brain protection and repair (202). In post-mortem cortical tissue of patients with schizophrenia, the synapse density is reduced. This excessive pruning reflects abnormalities in synaptic structures and microglia like cells

(203). A lot of this synaptic pruning occurs during adolescence, this is when schizophrenia symptoms occur (203). *In utero* MIA is a risk factor for neurodevelopmental disorders like schizophrenia. O'Loughlin et al. administered lipopolysaccharide to mice on embryonic day 12 to induce MIA. This induced a pro-inflammatory cytokine profile which continued in the amygdala to early adulthood. These alterations in the foetal brain elicited by MIA can lead to alterations to microglia (54). Pre and perinatal activation of the immune system can increase the immune system's sensitivity throughout life (18). Diverse immune alteration has been observed in those with schizophrenia and autoimmune disorders and severe infections are linked to schizophrenia risk (21,204).

Cellular response to diacyl lipopeptide stimulus (GO:0071726) showed up in the top three gene ontologies in at least one module per Developmental Stage apart from Developmental Stage Four.

Cellular response to catecholamine stimulus was highlighted by anRichment in at least one module per Developmental Stage (GO:0071870). Catecholamines, which are neurotransmitters in the CNS and peripheral nervous system, include dopamine, norepinephrine and adrenaline (127). Catecholamine signalling underlies the mesocorticolimbic system and affects executive function and cognition (206). Catecholamine signalling pathways are pharmacological therapy targets for patients with neuropsychiatric disorders because of their relationship with affective, executive and cognitive functions (206). Catecholamines have versatile functions as slow-acting neurotransmitters in synaptic neurotransmission and controlling the effects of fast-acting neurotransmitters (205). Dopamine has been linked with schizophrenia for many years and dopamine receptor antagonists continue to be the leading therapy for schizophrenia (5). Dopamine neurotransmission is altered in many neural pathways in schizophrenia (207), these alterations include hyperactive dopaminergic transmission in the striatum, hippocampus and mesolimbic areas and hypoactive transmission in the PFC of patients which schizophrenia (207). Dopamine displays regulatory effects on an immune response which depends on dopamine concentration, sub-type of receptors, time of exposure, type of immune cell and immune cell activation (31). This has been shown to affect cognitive functions (207). Immune cells in particular T-cells, microglial and peripheral monocytes collaborate with the CNS and have cognitive and behavioural function which are seen to be altered in schizophrenia (31,207). Dopamine has been seen to influence the activity of these immune cells since they express dopamine receptors (207). Changes in dopamine concentration and/or receptors in T cells are thought to be the cause of abnormal immune functions in people with schizophrenia and Parkinson's (208). Low levels of dopamine neurotransmission, as well as serotonin and glutamatergic neurotransmission, are seen in people with schizophrenia. These are connected to low levels of neuroinflammation, which has been hypothesised to be the reason for CNS volume loss and low levels of microglial activation in schizophrenia patients in neuroimaging studies (18). Renalase is thought to metabolise dopamine. In a study conducted by Catak et al. which used thirty-three schizophrenia patients it was found that the levels of renalase in these patients was significantly lower than the control group. This could be a potential biomarker for schizophrenia (209).

Regulation of neurotransmitter secretion (GO:0046928) was enriched in Developmental Stage One and Three and Neurotransmitter receptor activity (GO:0030594) in Developmental Stage Four. Multiple neurotransmitters have been implicated in being involved in schizophrenia including Dopamine, Serotonin, GABA and NMDA (37).

In Developmental Stage One, oestrogen 16-alpha-hydroxylase activity (GO:0101020) was highlighted as being enriched in the turquoise module. Oestrogen is produced in the brain in areas including the hippocampus, cortex, amygdala, hypothalamus and cerebellum (210). Because oestrogen can have a powerful effect on numerous areas of the brain it can affect mood, cognition and behaviour (211). It plays a critical role in influencing dopamine, serotonin, glutamate and GABA neurotransmitters which are key in schizophrenia (210).

Although oestrogen is involved in brain development the exact role of oestrogen 16-alpha-hydroxylase activity in schizophrenia is less obvious.

The immune system, inflammation, catecholamines and neurotransmitter dysfunctions have been identified as being important for each developmental stage. One of the important findings from the study performed by Pardinas et al. was the significance of the SNP which was found in the major histocompatibility complex (xMHC) region which is linked to the adaptive immune system. Although SNPs associated with the MHC region was included in the gene list which was provided by Pardinas's study, ABA's BrainSpan resource did not include gene expression data for it and thus were not included in our downstream analysis. Our findings support the role of the immune system in Schizophrenia over each of the Developmental Stages.

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto, Border: : (No border), Pattern: Clear

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto, Border: : (No border), Pattern: Clear

4.5 Limitations of the study

There were several limitations of this study. Firstly, ABA's BrainSpan resource developed in 2014 uses healthy brains for each developmental stage. ABA's BrainSpan contains gene expression data across 16 brain regions from 8 post-conception to 40 years, covering the complete development process (79). The atlas contains next-generation RNA sequencing data which has collected 579 tissue samples from thirty neurologically unremarkable brains over five developmental stages (80). A limitation to the BrainSpan resource is that the brain is divided into just 16 regions and that the brain sample size is small. In addition, ~80% of participants had transcriptomes missing from at least one brain region which means you're not getting complete pictures of the brain (79). ABA uses only neurotypical brains, we applied similar methods which had proven to be successful in studies performed by McCarthy et al. and Negi et al. (82,83). An additional gene expression database is the genotype tissue expression (GTEx) project. This database was created to enable the study of human gene regulation and variation of gene expression in multiple tissues (212). The GTEx consortium collected 14,787 transcriptomes from 948 patients including 13 brain regions (79). The largest limitation of using GTEx is that it does not have developing brain data which is why ABA was used in this study.

Brain tissue is comprised of diverse basic cell types, if the expression at tissue levels changes it may be due to alterations in the proportion of basic cell types (213). Single-cell sequencing (scRNASeq) allows for the dissection of gene expression at single-cell resolution, using scRNASeq can lead to findings in cell expression alterations and dynamics (214). This revolutionary tool if applied to schizophrenia could uncover the uniqueness of each brain cell at microscopic resolution (215). PsychENCODE is a resource that was built to elucidate the underlying molecular mechanisms of psychiatric disorders. The resource includes the integration of data from ENCODE, GTEx, Roadmap and relevant single-cell studies. PsychENCODE's main success has been the multi-omics approach of non-coding elements and transcriptome in neurotypical developing brains and adults with psychiatric disorders (213). Unfortunately, PsychENCODE does not have a developmental database and thus was not used in this current study.

The schizophrenia-associated gene set which was used in this study was taken from a study performed by Pardiñas et al. (66). Pardiñas et al. used these results as a training set to create risk profile scores to identify SNPs at high confidence (66). This study is the largest schizophrenia GWAS to date with 40,675 cases. The study size could be larger and could produce more reproducible SNP's. The SNPs identified were mapped to the closest loci and if there were no overlapping genes it was mapped to the closest gene within a 500 KB distance (66). As with all GWAS, the best way to discover more and yield more accurate results is to increase the sample size of the study. The Pardiñas study mapped to 479 genes, of these genes only 316 genes were available on ABA's BrainSpan transcriptome atlas. Missing a large portion of the Pardiñas gene set is a limitation of the current study and some SNPs did not map to any genes available in ABA.

K-means as a method has several limitations, firstly because the user has to decide on the number of clusters k before beginning the analysis. Secondly, k-means has a bias of creating modules of a similar size which may not accurately represent the group (90). Lastly, k-means centroids are immensely affected by outliers and can give the outliers disproportionate importance (89,90). K-means was used in this analysis to look for underlying patterns within the genetic data per stage. When the Pardiñas gene set was put through k-means the intercluster variation explained most of the total variance in the data which is not typical for good clustering.

An additional limitation of this study is that there was no clinical data that could be applied to the WGCNA data. Calculating gene trait significance also allows for the gene expression to be linked to biologically relevant traits. Studies which had clinical data were able to determine the GS threshold by combining it with module membership. Those with the highest GS can be considered HG for a module. In this study, MM was calculated by determining the node distance from the central ME. The MM calculated measured by Pearson correlation between ME and gene expression should only contain genes with the highest correlation but one study concluded that 25% be a better fit in other modules (216). Usually, WGCNA analysis uses much more data, more genes could mean more than one HG for a module and can give a better insight into the critical underlying pathways within a module.

4.6 Future Directions

To get a more comprehensive understanding of the brain changes which result in severe psychiatric diseases it is important to have a clear understanding of mechanisms that occur in normal brain development. When there is a comprehensive understanding of normal brain development it will be easier to observe the changes that occur in neuropsychiatric conditions like schizophrenia. At present, it is difficult to obtain full affected brains at any stage and as there are no biomarkers it is difficult to predict who will develop schizophrenia until it eventually manifests. This would mean prenatal, child and early adolescence brains which are obtained could not be confirmed that they would develop schizophrenia. In schizophrenia, it is important to understand the brains mechanisms and what goes awry in the early stages. To create a developmental resource researcher would have to use brains from people who are high risk (perhaps based on the polygenic risk scores for schizophrenia) or those who have close relatives with schizophrenia.

To combat these limitations a longitudinal study where participants who are more high risk to develop schizophrenia than the general population could be studied from birth to adulthood. This type of analysis would use blood samples and MRIs but a limitation for this kind of study is the inability to study arguably the most important stage during development *in utero*, where crucial decisions involving cell fate and distinctive development of the brain occur. Thus, the incorporation of genetic, epigenome and gene expression data and machine learning approaches will allow for a better understanding of abnormal brain development.

Gene Ontology databases are ever-growing with continual studies uncovering molecular functions, cellular locations, and biological processes. With more studies utilising WGS there will be a more absolute understanding of the genome and what it does and continuously expand the scope of GO. Using GO enrichment analysis like in this study could identify biological differences between controls and schizophrenia patients which could provide pharmaceutical companies with novel therapeutic targets.

Functional work on the HG’s of interest at each of the stages could be performed, looking at risk variants. These genes could be studied in knock out mice. It would be particularly interesting to study *ALMS1* and *TRANK1* which were highlighted as being interesting in this study and Pardiñas et al. Clustered regularly interspaced short palindromic repeat (CRISPR-Cas9) is a genome-editing tool that can induce double-strand breaks at target regions and can alleviate a gene's function (217,218). As *ALMS1* and *TRANK1* were highlighted as being important in both this study and Pardiñas study these would be great candidate genes for further downstream analysing including CRISPR-Cas9.

In the future, a method similar to this could be applied to other neuropsychiatric disorders including ASD, MD and BPD. Studying conditions like these over time could assist in more accurate prediction of whom will develop, how severe the symptoms could be, and which treatments could be effective as well as identify candidate genes for downstream analysis. This type of analysis could be easily adapted when newer GWAS produce more comprehensive gene lists.

4.7 Conclusion

Schizophrenia is still a long way from being fully understood. From the schizophrenia-associated genes which were highlighted by Pardiñas and available in ABA’s BrainSpan resource, 316 genes were available. Once GO enrichment was applied to the modules produced by WGCNA, it was clear that macrophage proliferation and catecholamine dysfunction were important mechanisms underlying schizophrenia in each developmental stage. The importance of the immune system and catecholamine’s most notably dopamine has been highlighted in multiple previous studies. There is an interaction between the immune system and dopamine and evidence for this is co-morbidities of schizophrenia and autoimmune diseases. Researching them together may give additional insight into their interdependency.

To more accurately pinpoint central ontologies for each stage a study like this must be done comparing controls and patients with schizophrenia. Studying them and looking at immune and catecholamine processes specifically throughout development could produce novel therapeutics with better efficacy and less severe side effects.

5.0 Bibliography

1. Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, et al. Biological Insights From 108 Schizophrenia-Associated Genetic Loci. Nature. 2014;511(7510):421–7. Available from: http://www.nature.com/authors/editorial_policies/license.html#terms
2. Stilo SA, Murray RM. Non-Genetic Factors in Schizophrenia. Curr Psychiatry Rep. 2019;21(100).
3. Pugliese V, Bruni A, Carbone EA, Calabrò G, Cerminara G, Sampogna G, et al. Maternal stress, prenatal medical illnesses and obstetric complications: Risk factors for schizophrenia spectrum disorder, bipolar disorder and major depressive disorder. Psychiatry Res. 2019;271:23–30.
4. Haller CS, Padmanabhan JL, Lizano P, Torous J, Keshavan M. Recent advances in understanding schizophrenia. F1000Prime Rep. 2014;6(57):1–11.
5. Amanda Krogmann, Luisa Peters, Laura von Hardenberg, Katja Bödeker, Viktor B. Nöhles and CUC. Keeping up with the therapeutic advances in schizophrenia: a review of novel and emerging pharmacological entities Amanda. CNS Spectr. 2019;24:41–68.
6. Carbon M, Correll CU. Thinking and acting beyond the positive: The role of the cognitive and negative symptoms in schizophrenia. CNS Spectr. 2014;19(S1):35–53.
7. Gobira PH, Ropke J, Aguiar DC, Crippa JAS, Moreira FA. Animal models for predicting the efficacy and side effects of antipsychotic drugs. Rev Bras Psiquiatr. 2013;35:132–9.
8. Owen MJ, Sawa A, Mortensen PB. Schizophrenia. Lancet. 2016;388(10039):86–97.
9. Avramopoulos D. Recent Advances in the Genetics of Schizophrenia. Mol Neuropsychiatry [Internet]. 2018 [cited 2019 Sep 3];4:35–51. Available from: www.karger.com/mnp
10. Insel TR. Rethinking schizophrenia. Nature. 2010. p. 187–93.
11. Stępnicki P, Kondej M, Kaczor AA. Current concepts and treatments of schizophrenia. Molecules. 2018;23(2087).
12. Strous RD, Shoenfeld Y. Schizophrenia, autoimmunity and immune system dysregulation: A comprehensive model updated and revisited. J Autoimmun. 2006;27:71–80.
13. Ota VK, Moretti PN, Santoro ML, Talarico F, Spindola LM, Xavier G, et al. Gene expression over the course of schizophrenia: from clinical high-risk for psychosis to chronic stages. npj Schizophr. 2019;5(5).
14. Gejman P V., Sanders AR, Duan J. The role of genetics in the etiology of schizophrenia. Psychiatr Clin North Am. 2010;33(1):35–66.
15. Lally J, MacCabe JH. Antipsychotic medication in schizophrenia: A review. Br Med Bull. 2015;114:169–79.
16. Muller N, J. Schwarz M. The Role of Immune System in Schizophrenia. Curr Immunol Rev. 2010;6(3):213–20.
17. Van Kesteren CFMG, Gremmels H, De Witte LD, Hol EM, Van Gool AR, Falkai PG, et al. Immune involvement in the pathogenesis of schizophrenia: A meta-analysis on postmortem brain studies. Transl Psychiatry. 2017;7(3). Available from: http://dx.doi.org/10.1038/tp.2017.4
18. Müller N. Inflammation in schizophrenia: Pathogenetic aspects and therapeutic considerations. Schizophr Bull. 2018;44(5):973–82.
19. Müller N, Weidinger E, Leitner B, Schwarz MJ. The role of inflammation in schizophrenia. Frontiers in Neuroscience. 2015.
20. Horváth S, Mirnics K. Immune system disturbances in schizophrenia. Biol Psychiatry. 2014;75(4):316–23. Available from: http://dx.doi.org/10.1016/j.biopsych.2013.06.010
21. Benros ME, Nielsen PR, Nordentoft M, Eaton WW, Dalton SO, Mortensen PB. Autoimmune diseases and severe infections as risk factors for schizophrenia: A 30-year population-based register study. Am J Psychiatry. 2011;168(12):1303–10.
22. Khandaker GM, Cousins L, Deakin J, Lennox BR, Yolken R, Jones PB. Inflammation and immunity in schizophrenia: Implications for pathophysiology and treatment. The Lancet Psychiatry. 2015;2(3):258–70.

23. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B, Brian J. Miller, Peter Buckley, Wesley Seabolt, Andrew Mellor and BK. Meta-Analysis of Cytokine Alterations in Schizophrenia: Clinical Status and Antipsychotic Effects Brian. Biol Psychiatry. 2011;70(7):663–71.
24. Song XQ, Lv LX, Li WQ, Hao YH, Zhao JP. The Interaction of Nuclear Factor-Kappa B and Cytokines Is Associated with Schizophrenia. Biol Psychiatry. 2009;65:481–8.
25. Dinan TG, Borre YE, Cryan JF. Genomics of schizophrenia: Time to consider the gut microbiome? Mol Psychiatry. 2014;19:1252–7.
26. Zheng P, Zeng B, Liu M, Chen J, Pan J, Han Y, et al. The gut microbiome from patients with schizophrenia modulates the glutamate-glutamine-GABA cycle and schizophrenia-relevant behaviors in mice. Sci Adv. 2019;5.
27. Kelly JR, Minuto C, Cryan JF, Clarke G, Dinan TG. The role of the gut microbiome in the development of schizophrenia. Schizophr Res. 2020;
28. Magdalon J, Mansur F, Teles e Silva AL, de Goes VA, Reiner O, Sertié AL. Complement System in Brain Architecture and Neurodevelopmental Disorders. Front Neurosci. 2020;14(23).
29. Owen MJ, O'donovan MC. Schizophrenia and the neurodevelopmental continuum:evidence from genomics. World Psychiatry. 2017;16:227–35.
30. Gupta S, Kulhara P. What is schizophrenia: A neurodevelopmental or neurodegenerative disorder or a combination of both A critical analysis. Indian J Psychiatry. 2010;52(1):21–7.
31. S.M Matt P. G. Where is Dopamine and How Do Immune Cells See It?: Dopamine-Mediated Immune Cell Function in Health and Disease. J Neuroimmune Pharmacol. 2020;15(1):114–64.
32. Howes O, McCutcheon R, Stone J. Glutamate and dopamine in schizophrenia: An update for the 21st century. J Psychopharmacol. 2015;29(2):97–115.
33. Kesby JP, Eyles DW, McGrath JJ, Scott JG. Dopamine, psychosis and schizophrenia: The widening gap between basic and clinical neuroscience. Transl Psychiatry. 2018;8(30).
34. Murray RM, Bhavsar V, Tripoli G, Howes O. 30 Years on: How the Neurodevelopmental Hypothesis of Schizophrenia Morphed into the Developmental Risk Factor Model of Psychosis. Schizophr Bull. 2017;43(6):1190–6.
35. Frank TVM and MJ. An Integrative Perspective on the Role of Dopamine in Schizophrenia. Biol Psychiatry. 2017;176(1):52–66.
36. Demjaha A, Egerton A, Murray RM, Kapur S, Howes OD, Stone JM, et al. Antipsychotic treatment resistance in schizophrenia associated with elevated glutamate levels but normal dopamine function. Biol Psychiatry. 2014;75(5).
37. Yang AC, Tsai SJ. New targets for schizophrenia treatment beyond the dopamine hypothesis. Int J Mol Sci. 2017;18(1689).
38. Brisch R, Saniotis A, Wolf R, Bielau H, Bernstein HG, Steiner J, et al. The role of dopamine in schizophrenia from a neurobiological and evolutionary perspective: Old fashioned, but still in vogue. Front Psychiatry. 2014;5(47).
39. Uno Y, Coyle JT. Glutamate hypothesis in schizophrenia. Psychiatry Clin Neurosci. 2019;73:204–15.
40. Stahl SM. Beyond the dopamine hypothesis of schizophrenia to three neural networks of psychosis: Dopamine, serotonin, and glutamate. CNS Spectr. 2018;23(3):187–91.
41. Bustillo JR, Chen H, Jones T, Lemke N, Abbott C, Qualls C, et al. Increased glutamine in patients undergoing long-term treatment for schizophrenia: A proton magnetic resonance spectroscopy study at 3 T. JAMA Psychiatry. 2014;71(3):265–72.
42. Aghajanian GK, Marek GJ. Serotonin model of schizophrenia: emerging role of glutamate mechanisms. In: Brain Research Reviews. 2000. p. 302–12.
43. Sánchez-Blázquez P, Rodríguez-Muñoz M, Garzón J. The cannabinoid receptor 1 associates with NMDA receptors to produce glutamatergic hypofunction: Implications in psychosis and schizophrenia. Front Pharmacol. 2014;4(169).
44. Marsman A, Van Den Heuvel MP, Klomp DWJ, Kahn RS, Luijten PR, Hulshoff Pol HE. Glutamate in schizophrenia: A focused review and meta-analysis of 1H-MRS studies. Schizophr Bull. 2013;39(1):120–9.
45. Egerton A, Modinos G, Ferrera D, McGuire P. Neuroimaging studies of GABA in schizophrenia: A

systematic review with meta-analysis. Transl Psychiatry. 2017;7:1–10.
46. Cohen SM, Tsien RW, Goff DC, Halassa MM. The impact of NMDA receptor hypofunction on GABAergic neurons in the pathophysiology of schizophrenia. Schizophr Res. 2015;167(1–3):98–107.
47. Glausier JR, Lewis DA. GABA and schizophrenia: Where we stand and where we need to go. Schizophr Res. 2017;181:2–3.
48. Birnbaum R, Weinberger DR. Genetic insights into the neurodevelopmental origins of schizophrenia. Nat Rev Neurosci. 2017;18:727–40.
49. Quednow BB, Geyer MA, Halberstadt AL. Serotonin and Schizophrenia. In: Handbook of Behavioral Neuroscience. 2010. p. 585–620.
50. Abi-Dargham A. Alterations of Serotonin Transmission in Schizophrenia. Int Rev Neurobiol. 2007;78:133–64.
51. Sumiyoshi T, Kunugi H, Nakagome K. Serotonin and dopamine receptors in motivational and cognitive disturbances of schizophrenia. Front Neurosci. 2014;8(395).
52. Van Os J, Kenis G, Rutten BPF. The environment and schizophrenia. Nature. 2010;468(7321):203–12. Available from: http://dx.doi.org/10.1038/nature09563
53. Janoutová J, Janáčková P, Šerý O, Zeman T, Ambroz P, Kovalová M, et al. Epidemiology and risk factors of schizophrenia. Neuroendocrine Letters. 2016;37(1):26994378–370116. Available from: www.nel.edu
54. O'Loughlin E, Pakan JMP, Yilmazer-Hanke D, McDermott KW. Acute in utero exposure to lipopolysaccharide induces inflammation in the pre- and postnatal brain and alters the glial cytoarchitecture in the developing amygdala. J Neuroinflammation. 2017;14(1):1–12.
55. Hollins SL, Cairns MJ. MicroRNA: Small RNA mediators of the brains genomic response to environmental stress. Prog Neurobiology. 2016;143:61–81. Available from: http://dx.doi.org/10.1016/j.pneurobio.2016.06.005
56. Brown AS. The environment and susceptibility to schizophrenia. Prog Neurobiol. 2011;93(1):23–58.
57. Shenton ME, Dickey CC, Frumin M, Mccarley RW. A review of MRI findings in schizophrenia. Schizophr Res . 2001;49(1):1–52. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2812015/
58. Dietsche B, Kircher T, Falkenberg I. Structural brain changes in schizophrenia at different stages of the illness: A selective review of longitudinal magnetic resonance imaging studies. Aust N Z J Psychiatry. 2017;1–9.
59. Haijma S V., Van Haren N, Cahn W, Koolschijn PCMP, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: A meta-analysis in over 18 000 subjects. Schizophr Bull. 2013;39(5):1129–38.
60. Fusar-Poli P, Smieskova R, Kempton MJ, Ho BC, Andreasen NC, Borgwardt S. Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. Neurosci Biobehav Rev. 2013;37(8):1680–91.
61. Scheepers FE, De Wied CCG, Pol HEH, Van De Flier W, Van Der Linden JA, Kahn RS. The effect of clozapine on caudate nucleus volume in schizophrenic patients previously treated with typical antipsychotics. Neuropsychopharmacology. 2001;24(1):47–54.
62. Hilker R, Helenius D, Fagerlund B, Skytthe A, Christensen K, Werge TM, et al. Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. Biol Psychiatry. 2018;83(6):492–8. Available from: http://dx.doi.org/10.1016/j.biopsych.2017.08.017
63. Schwab SG, Wildenauer DB. Genetics of psychiatric disorders in the GWAS era: An update on schizophrenia. Eur Arch Psychiatry Clin Neurosci. 2013;263:147–54.
64. Henriksen MG, Nordgaard J, Jansson LB. Genetics of schizophrenia: Overview of methods, findings and limitations. Front Hum Neurosci. 2017;11(322).
65. Duncan EL, Brown MA. Genome-Wide Association Studies. In: Genetics of Bone Biology and Skeletal Disease: Second Edition. 2017.
66. Antonio F. Pardiñas1, Peter Holmans1, Andrew J. Pocklington1, Valentina Escott- Price1, Stephan Ripke2, 3, Noa Carrera1, Sophie E. Legge1, Sophie Bishop1, Darren Cameron1, Marian L. Hamshere1, Jun Han1, Leon Hubbard1, Amy Lynham1, Kiran Mantripragada1 ER. Common schizophrenia alleles are enriched in mutation intolerant genes and in region under strong

background selection. Nature Genetics. 2018;50(3):381–9. Available from: file:///C:/Users/Carla Carolina/Desktop/Artigos para acrescentar na qualificação/The impact of birth weight on cardiovascular disease risk in the.pdf

67. Skene NG, Bryois J, Bakken TE, Breen G, James J, Gaspar HA, et al. Genetic identification of brain cell types underlying schizophrenia. 2019;50(6):825–33.
68. Bush WS, Moore JH. Chapter 11: Genome-Wide Association Studies. PLoS Comput Biol. 2012;8(12):1–10.
69. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. American Journal of Human Genetics. 2017. p. 5–22.
70. van Mierlo HC, Schot A, Boks MPM, de Witte LD. The association between schizophrenia and the immune system: Review of the evidence from unbiased ‘omic-studies.’ Schizophrenia Research. 2020;217:114–23. Available from: https://doi.org/10.1016/j.schres.2019.05.028
71. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204–13.
72. Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM, et al. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. Med Sci Monit Basic Res. 2014;20:138–41.
73. Ball S, Gilbert TL, Overly CC. The Human Brain Online: An Open Resource for Advancing Brain Research. PLoS Biol. 2012;10(12).
74. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012 Sep 20;489(7416):391–9. Available from: http://www.nature.com/articles/nature11405
75. Sunkin SM, Ng L, Lau C, Dolbeare T, Gilbert TL, Thompson CL, et al. Allen Brain Atlas: An integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Res. 2013;41(database issue).
76. Arnatkevičiūtė A, Fulcher BD, Fornito A. A practical guide to linking brain-wide gene expression and neuroimaging data. Neuroimage. 2019;189:353–67.
77. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, et al. Transcriptional landscape of the prenatal human brain. Nature. 2014;508:199–206.
78. Vasung L, Abaci Turk E, Ferradal SL, Sutin J, Stout JN, Ahtam B, et al. Exploring early human brain development with structural and physiological neuroimaging. Neuroimage. 2019;187:226–54.
79. Pei G, Wang YY, Simon LM, Dai Y, Zhao Z, Jia P. Gene expression imputation and cell-type deconvolution in human brain with spatiotemporal precision and its implications for brain-related disorders. Genome Res. 2021;1(1):146–58.
80. Mahfouz A, Ziats MN, Rennert OM, Lelieveldt BPF, Reinders MJT. Shared Pathways Among Autism Candidate Genes Determined by Co-expression Network Analysis of the Developing Human Brain Transcriptome. J Mol Neurosci. 2015;57:580–94.
81. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012;489:391–9.
82. Negi SK, Guda C. Global gene expression profiling of healthy human brain and its application in studying neurological disorders OPEN. Nat Reports. 2017 Dec 18 ;7(897):1–12. Available from: www.nature.com/scientificreports/
83. Mccarthy MJ, Liang S, Spadoni AD, Kelsoe JR, Simmons AN. Whole Brain Expression of Bipolar Disorder Associated Genes: Structural and Genetic Analyses. PLoS One . 2014 ;9(6):1–13. Available from: www.plosone.org
84. Rodriguez MZ, Comin CH, Casanova D, Bruno OM, Amancio DR, Costa L da F, et al. Clustering algorithms: A comparative approach. PLoS One. 2019;14(1).
85. Jayanthi K, Mahesh C. A Study on machine learning methods and applications in genetics and genomics. Int J Eng Technol. 2018;7(1.7):201–4.
86. Saelens W, Cannoodt R, Saeys Y. A comprehensive evaluation of module detection methods for gene expression data. Nat Commun. 2018;9(1090):1–12.
87. Gibbons FD, Roth FP. Judging the quality of gene expression-based clustering methods using gene annotation. Genome Res. 2002;12:1574–81.

88. Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognit Lett. 2010;31:651–66.
89. Shi N, Liu X, Guan Y. Research on k-means clustering algorithm: An improved k-means clustering algorithm. In: 3rd International Symposium on Intelligent Information Technology and Security Informatics, IITSI 2010. IEEE; 2010. p. 63–7.
90. Morissette L, Chartier S. The k-means clustering technique: General considerations and implementation in Mathematica. Tutor Quant Methods Psychol. 2013;9(1):15–24.
91. Charrad M, Ghazzali N, Boiteau V, Niknafs A. Nbclust: An R package for determining the relevant number of clusters in a data set. J Stat Softw. 2014;61(6):1–36.
92. Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses [Computer sofware, package]. R package version. 2017.
93. van Dam S, Võsa U, van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. Brief Bioinform. 2018;19(4):575–92.
94. Li J, Zhou D, Qiu W, Shi Y, Yang JJ, Chen S, et al. Application of Weighted Gene Co-expression Network Analysis for Data from Paired Design. Sci Rep. 2018;8(622).
95. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(559).
96. Liu Y, Gu HY, Zhu J, Niu YM, Zhang C, Guo GL. Identification of Hub Genes and Key Pathways Associated With Bipolar Disorder Based on Weighted Gene Co-expression Network Analysis. Front Physiol. 2019;10:1–9.
97. Costantini G, Epskamp S, Borsboom D, Perugini M, Mõttus R, Waldorp LJ, et al. State of the aRt personality research: A tutorial on network analysis of personality data in R. J Res Pers. 2015;54:13–29.
98. Yu D, Kim M, Xiao G, Hwang TH. Review of Biological Network Data and Its Applications. Genomics Inform. 2013;11(4):200–10.
99. Sun K, Gonçalves JP, Larminie C, Pržulj N. Predicting disease associations via biological network analysis. BMC Bioinformatics. 2014;15(304):1–13.
100. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
101. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using cytoscape. Nat Protoc. 2007;2(10):2336–82.
102. Gang Su1, John H. Morris2, Barry Demchak3 and GDB. BIOLOGICAL NETWORK EXPLORATION WITH CYTOSCAPE 3. Curr Protoc Bioinforma. 2014;74(8).
103. Assenov Y, Ramírez F, Schelhorn SESE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. Bioinformatics. 2008;24(2):282–4.
104. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. Methods Mol Biol. 2011;
105. Carbon S, Dietze H, Lewis SE, Mungall CJ, Munoz-Torres MC, Basu S, et al. Expansion of the gene ontology knowledgebase and resources: The gene ontology consortium. Nucleic Acids Res. 2017;45(Database issue):331-D338.
106. Blake JA, Christie KR, Dolan ME, Drabkin HJ, Hill DP, Ni L, et al. Gene ontology consortium: Going forward. Nucleic Acids Res. 2015;43(D1):1049–56.
107. Langfelder P. anRichment: Collections and annotation data for use with anRichmentMethods. 2020.
108. Carlson M. org.Hs.eg.db: Genome wide annotation for Human. 2019.
109. Kassambara A, Mundt F, Kassambara, A.; Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. URL http//www.sthda.com/english/rpkgs/factoextra BugReports. 2017;
110. Barabási AL, Oltvai ZN. Network biology: Understanding the cell's functional organization. Nat Rev Genet. 2004;5:101–13.
111. Pavlopoulos GA, Secrier M, Moschopoulos CN, Soldatos TG, Kossida S, Aerts J, et al. Using graph theory to analyze biological networks. BioData Min. 2011;4(10).
112. Jacob R, Harikrishnan KP, Misra R, Ambika G. Measure for degree heterogeneity in complex networks and its application to recurrence network analysis. R Soc Open Sci. 2017;4(1).
113. Dong J, Horvath S. Understanding network concepts in modules. BMC Syst Biol. 2007;1(24).

114. Zawerton, Ash; Mignot, Cyril; Sigafoos, Ashley; Blackburn, Patrick R.; Haseeb, Abdul; McWalter, Kirsty; Ichikawa, Shoji; Nava, Caroline; Keren, Boris; Charles, Perrine; Marey, Isabelle; Tabet, Anne-Claude; Levy, Jonathan; Perrin, Laurence; Hartmann, Andre C. Widening of the genetic and clinical spectrum of Lamb-Shaffer syndrome, a neurodevelopmental disorder due to SOX5 haploinsufficiency. Genet Med. 2019;0(0).
115. Manfred Schartl1, 2, 3, Susanne Schories1, Yuko Wakamatsu1, Yusuke Nagao4, Hisashi Hashimoto4, Chloé Bertin5, Brigitte Mourot5, Cornelia Schmidt1, Dagmar Wilhelm6, Lazaro Centanin7 YG and AH. Sox5 is involved in germ-cell regulation and sex determination in medaka following co-option of nested transposable elements. BMC Biol. 2018;16(16).
116. Tae-Gul Lee1, 2 , Eun-Hui Jeong1, Seo Yun Kim1, Hye-Ryoun Kim1 HK, Kim C-H. Fhit, a tumor suppressor protein, induces autophagy via 14-3-3τ in non-small cell lung cancer cells. Oncotarget. 2017;8(19):31923–37.
117. Garneau AP, Slimani S, Tremblay LE, Fiola MJ, Marcoux AA, Isenring P. K^+-Cl^- cotransporter 1 (KCC1): a housekeeping membrane protein that plays key supplemental roles in hematopoietic and cancer cells. J Hematol Oncol [Internet]. 2019 Dec 11;12(1):74. Available from: https://jhoonline.biomedcentral.com/articles/10.1186/s13045-019-0766-x
118. Fiona C. Brown, 1 Ashlee J. Conway, 1 Loretta Cerruti, 1 Janelle E. Collinge, 2 Catriona McLean, 3 James S. Wiley, 4 Ben T. Kile, 2 Stephen M. Jane, 1, 3, 5,* and David J. Curtis1iona C. Brown, 1 Ashlee J. Conway, 1 Loretta Cerruti, 1 Janelle E. Collinge, and DJC. Activation of the erythroid K-Cl cotransporter Kcc1 enhances sickle cell disease pathology in a humanized mouse model. Blood. 2015;126(26):2863–70.
119. Kourosh Parham, M.D., Ph.D.1, George A. Kuchel, M.D.2, Janet E. McElhaney, M.D.4, and Laura Haynes PD. A relationship between blood levels of otolin-1 and vitamin D. Otol Neurotol. 2018;39(4):269–73.
120. Murat Doğan , Mustafa Şahin YK. Otolin-1, as a Potential Marker for Inner Ear Trauma after Mastoidectomy. J Int Adv Otol. 2019;15(2):200–3.
121. Yanyan Liu1 2 · Chuang Tong2 · Jun Cao2 · Maoming Xiong1. NEMP1 Promotes Tamoxifen Resistance in Breast Cancer Cells. Biochem Genet. 2019;57(6):813–26.
122. Yonit Tsatskis1*, Robyn Rosenfeld1, 2*, Joel D. Pearson1*, Curtis Boswell2, 3*, Yi Qu1, Kyunga Kim1, 4, Lacramioara Fabian5, Ariz Mohammad6, Xian Wang7, 8, Michael I. Robson9, Karen Krchma10, Jun Wu10, João Gonçalves1, Didier Hodzic11, Shu Wu1, Daniel Pot HM. The NEMP family supports metazoan fertility and nuclear envelope stiffness. Sci Adv Genet. 2020;6.
123. Hiroshi Mamada, Noriyuki Takahashi MT. Involvement of an inner nuclear membrane protein, Nemp1, in Xenopus neural development through an interaction with the chromatin protein BAF. Dev Biol. 2009;327:497–507.
124. Marta Codina-Solà, Benjamín Rodríguez-Santiago, Aïda Homs, Javier Santoyo, Maria Rigau, Gemma Aznar-Laín, Miguel del Campo, Blanca Gener, Elisabeth Gabau, María Pilar Botella, Armand Gutiérrez-Arumí1, Guillermo Antiñolo, Luis Alberto Pérez-Jurado and IC. Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. Mol Autism. 2015;6(21).
125. Mitsuru Ebihara, Hisako Ohba, Mika Kikuchi TY. Structural characterization and promoter analysis of human potassium channel Kv8.1 (KCNV1) gene. Gene. 2004;325:89–96.
126. Takashi Hashimoto 1, Shohei Maekawa SM. IgLON cell adhesion mol- ecules regulate synaptogenesis in hippocampal neurons. Cell Biochem Functopm. 2009;7:496–8.
127. Lan Wen a, Lei Liu b, Linyan Tong a, Jinfang Li a, Keming Zhang a, Qinbin Zhang a CL a. NDRG4 prevents cerebral ischemia/reperfusion injury by inhibiting neuronal apoptosis. Genes Dis. 2019;6:448–54.
128. Mitsuo Wada, Kayo Yukawa, Hiroyuki Ogasawara, ..., Takeshi Ohta, Eunyoung Lee TM. GPR52 accelerates fatty acid biosynthesis in a ligand-dependent manner in hepatocytes and in response to excessive fat intake in mice. iScience. 2021;24(102260).
129. Hidetoshi Komatsu1*., Minoru Maruyama2., Shuuhei Yao3, Tokuyuki Shinohara4, Kensuke Sakuma2, Sachiko Imaichi4, Tomoko Chikatsu2, Kanako Kuniyeda5, Foo Kok Siu6, Lam Sock Peng6, Katherine Zhuo7, Lay Sock Mun7, Tan Min Han7, Yoshio Matsumoto3, Tadatoshi Has MM.

Anatomical Transcriptome of G Protein-Coupled Receptors Leads to the Identification of a Novel Therapeutic Candidate GPR52 for Psychiatric Disorders. PLoS One. 2014;9(2).
130. Lorena Benedetti1, 2 , Matteo Cereda1, 2 , LeeAnn Monteverde1, 2 , Nikita Desai1 2 and, Ciccarelli FD. Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1. Oncotarget. 2017;8(23):37619–32.
131. Petra van derPetra van der Lelij1,* , Joseph A Newman2,*, Simone Lieb3,*, Julian Jude1,*, Vittorio Katis2, Thomas Hoffmann1, Matthias Hinterndorfer1, Gerd Bader3, Norbert Kraut3, Mark A Pearson3, Jan-Michael Peters1, 4, Johannes Zuber1, 4, Opher Gileadi2, MP. STAG1 vulnerabilities for exploiting cohesin synthetic lethality in STAG2-deficient cancers. Life Sci Alliance. 2020;3(7).
132. Daphné Lehalle, 1, 2 Anne-Laure Mosca-Boidron, 3, 4 Amber Begtrup, 5 Odile Boute-Benejean, 6 Perrine Charles, 7 Megan T Cho, 5 Amanda Clarkson, 8 Orrin Devinsky, 9 Yannis Duffourd, 4 Laurence Duplomb-Jego, 3, 4 Bénédicte Gérard, 10 Aurélia Jacquette, 7 Pa 4 Laurence Faivre1. STAG1 mutations cause a novel cohesinopathy characterised by unspecific syndromic intellectual disability. BMJ. 2017;54(7):479–88.
133. Whitton L, Apostolova G, Rieder D, Dechant G, Rea S, Donohoe G, et al. Genes regulated by SATB2 during neurodevelopment contribute to schizophrenia and educational attainment. PLoS Genet. 2018;14(7).
134. Qian Y, Liu J, Yang Y, Chen M, Jin C, Chen P, et al. Paternal low-level mosaicism-caused SATB2-associated syndrome. Front Genet. 2019;
135. Schult, D., Ho¨lsken MB, Schlaffer, S.-M. Siegel, S. Buslei, I. Kreitschmann-Andermahr , R. Fahlbusch R. Expression pattern of neuronal intermediate filament a-internexin in anterior pituitary gland and related tumors. Pituitary. 2015;18:465–473.
136. Yingjie Li, Liangliang Bai, Huichuan Yu, Du Cai, Xiaolin Wang, Baoyuan Huang, Shaoyong Peng, Meijin Huang, Guangwen Cao, Andrew M. Kaz, William M. Grady, Jianping Wang, and Yingjie Li, Liangliang Bai1, Huichuan Yu, Du and YLL. Epigenetic Inactivation of α-Internexin Accelerates Microtubule Polymerization in Colorectal Cancer. Cancer Res. 2020;80(23):5203–15.
137. Whelan R, Clair DS, Mustard CJ, Hallford P, Wei J. Study of novel autoantibodies in schizophrenia. Schizophr Bull. 2018;44(6):1341–9.
138. Xiaoyu Xu1, 2,† Zhao Zeng1,† Li Huo1,† Hong Liu1, 3 Yan Yu1 Ling Zhang1 Jiannong Cen1 Huiying Qiu1, 3 Xiaowen Tang1, 3 Chengcheng Fu1, 3 Yue Han1, 2, 3 Miao Miao1, 3 Zhengming Jin1, 3 Changgeng Ruan1, 2, 3 Depei Wu1, 2, 3 Suning Chen1, 2, 3 Qinrong Wang1 3, Lingzhi Yan1 3. High expression of myocyte enhancer factor 2C predicts poor prognosis for adult acute myeloid leukaemia with normal karyotype. BJH Corresp. 2019;189(1):23–7.
139. Adam J Harrington 1, Catherine M Bridges 2, Stefano Berto 3, Kayla Blankenship 1, Jennifer Y Cho 2, Ahlem Assali 1, Benjamin M Siemsen 4, Hannah W Moore 5, Evgeny Tsvetkov 1, Acadia Thielking 1, Genevieve Konopka 3, David B Everman 5, Michael D Scofield 4 CWC 6. MEF2C Hypofunction in Neuronal and Neuroimmune Populations Produces MEF2C Haploinsufficiency Syndrome-like Behaviors in Mic. Biol Psychiatry. 2020;15(88):488–99.
140. Adam J Harrington1 2†, , Aram Raissi2† , Kacey Rajkovich3 SB, Jaswinder Kumar2 4, Molinaro3 G, Raduazzo2 J, Guo2 Y, et al. MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to neurodevelopmental disorders. Elife. 2016;5.
141. Melanie Richter1, Nadeem Murtaza2, 3 Robin Scharrenberg1 Sean H. White2, 3 Ole Johanns1 Susan Walker4, 5 , Ryan K. C. Yuen 4, 5 Birgit Schwanke1 Bianca Bedürftig1 Melad Henis1, 6 Sarah Scharf7 Vanessa Kraus7 , Ronja Dörk7 Jakob Hellmann7 Zsuzsa Lindenmaie 3 Froylan Calderon de Anda. Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. Mol Psychiatry. 2019;24:1329–50.
142. Chih-Yeu Fang 1, Tsung-Ching Lai 2, 3, Michael Hsiao 1, 4 and Y-CC. The Diverse Roles of TAO Kinases in Health and Diseases. J Mol Sci. 2020;27(7463).
143. Yang Shao1, 2, Jing Kong1† HX, , Xiaoli Wu1, YuePeng Cao1 , Weijian Li 1 , Jing Han4 , Dake Li 1 KX and JW. OPCML Methylation and the Risk of Ovarian Cancer: A Meta and Bioinformatics Analysis. Front Cell Dev Biol. 2021;9(570908).
144. Zhengrong Zhang, Maoqing Ye, Qiongwei Li, Rena Li, Jun Li DZ. The Schizophrenia Susceptibility Gene OPCML Regulates Spine Maturation and Cognitive Behaviors through Eph-Cofilin Signaling.

Cell Rep. 2019;29:49–61.
145. Jane Antony1, 2, 3, Elisa Zanini1 , Zoe Kelly1, Tuan Zea Tan2, Evdoxia Karali1, Mohammad Alomary1, Youngrock Jung1, Katherine Nixon1, Paula Cunnea1, Christina Fotopoulou1, Andrew Paterson1, Sushmita Roy-Nawathe1, Gordon B Mills4, Ruby Yun-Ju Huang2, 5, Je & CR. The tumour suppressor OPCML promotes AXL inactivation by the phosphatase PTPRG in ovarian cancer. EMBO J. 2081;19(e45670).
146. Vougiouklakis1 T, Ryuji Hamamoto1 Y, Vassiliki N&, Saloura* 1. The NSD family of protein methyltransferases in human cancer. Epigenomics. 2015;7(5).
147. Li LYLYQLC. Downregulation of NSD3 (WHSC1L1) inhibits cell proliferation and migration via ERK1/2 deactivation and decreasing CAPG expression in colorectal cancer cells. Onco Targets Ther. 2019;12:3933–43.
148. Chunmei Wang, 1, 2 Qinlan Wang, 3 Xiaoqing Xu, 1 Bin Xie, 3 Yong Zhao, 4 Nan Li, 2 and Xuetao Cao1, 2 3. The methyltransferase NSD3 promotes antiviral innate immunity via direct lysine methylation of IRF3. J Exp Med. 2017;214(12):3597–610.
149. Vassiliki Saloura1 , Theodore Vougiouklakis1, Makda Zewde1, Kiyotani1 Kazuma, Jae-Hyun Park1, Guimin Gao2, Theodore Karrison2, Mark Lingen3 , Yusuke Nakamura1, 4 RH. WHSC1L1 drives cell cycle progression through transcriptional regulation of CDC6 and CDK2 in squamous cell carcinoma of the head and neck. Oncotarget. 2016;7(27):42527–38.
150. Paternoster L, Soblet J, Aeby A, De Tiège X, Goldman S, Yue WW, et al. Novel homozygous variant of carbonic anhydrase 8 gene expanding the phenotype of cerebellar ataxia, mental retardation, and disequilibrium syndrome subtype 3. Am J Med Genet Part A. 2020;97(3):516–20.
151. Ma HL, Yu SJ, Chen J, Ding XF, Chen G, Liang Y, et al. CA8 promotes RCC proliferation and migration though its expression level is lower in tumor compared to adjacent normal tissue. Biomed Pharmacother. 2020;121(109578).
152. Zhuang GZ, Keeler B, Grant J, Bianchi L, Fu ES, Zhang YP, et al. Carbonic anhydrase-8 regulates inflammatory pain by inhibiting the ITPR1-cytosolic free calcium pathway. PLoS One. 2015;10(3).
153. Tom Hearn. ALMS1 and Alström syndrome: a recessive form of metabolic, neurosensory and cardiac deficits. J Mol Med. 2019;97:1–17.
154. Shao-Yan Zhang, Chao Xuan, Yi Wang, Shao-Qiang Zhang, Hui Li, Guo-Wei He and Q-WT. Association between ALMS 1 variants and early-onset coronary artery disease: a case–control study in Chinese population. Biosci Rep. 2020;40(BSR20193637).
155. Hemang M. Parikh1 2*, , Targ Elgzyri2, Amra Alibegovic3, Natalie Hiscock4, Ola Ekström2, Karl-Fredrik Eriksson2, Allan Vaag3, Leif C. Groop2, 5, Kristoffer Ström2 6† and Ola Hansson. Relationship between insulin sensitivity and gene expression in human skeletal muscle. BMC Endocr Disord. 2021;21(32).
156. Jakub Tomasik, PhD1; Santiago G. Lago, PhD1; Javier Vázquez-Bourgon, MD P. Association of Insulin Resistance With Schizophrenia Polygenic Risk Score and Response to Antipsychotic Treatment. JAMA Psychiatry. 2019;76(8).
157. Zaoa X, *, Jin Chengb *, Congle Shena , Guiwen Guana, Xiaoyu Fenga, Jun Zoua, Jing Zhanga TL, Huiling Zhenga, Ting Zhanga, Jie Wanga, Jia Liua, Deyao Lia, Fengmin Lu, Fuping You and XC. NFATc3 inhibits hepatocarcinogenesis and HBV replication via positively regulating RIG-I-mediated interferon transcription. Oncoimmunology. 2021;10(1).
158. Katia Urso, Andrés Fernández, Patricia Velasco, Javier Cotrina, Belén de Andrés, Pilar Sánchez-Gómez , Aurelio Hernández-Laín, Sonsoles Hortelano JMR& EC. NFATc3 controls tumour growth by regulating proliferation and migration of human astroglioma cells. Nat Sci Reports. 2019;9(9361).
159. Sridevi Venkatesan EKL. Chrna5 is essential for a rapid and protected response to optogenetic release of endogenous acetylcholine in prefrontal cortex. bioXriv. 2020;40(38):7255–68.
160. Thaka’a K. Al-Omoush1, Karem H. Alzoubi, Omar F. Khabour, Fawzi M. Alsheyab, Ahmed Abu-siniyeh, Nour A. Al-Sawalha, Fadia A. Mayyas, Caroline O. Cobb TE. The CHRNA5 Polymorphism (rs16969968) and its Association with Waterpipe Smoking Addiction among Jordanians. Arab J Basix Appl Sci. 2020;27(1):450–5.
161. Myriam Srour, Noriaki Shimokawa, Fadi F. Hamdan, Christina Nassif, Chantal Poulin, Lihadh Al Gazali, Jill A. Rosenfeld, Noriyuki Koibuchi, Guy A. Rouleau, Aisha Al Shamsi and JLM.

Dysfunction of the Cerebral Glucose Transporter SLC45A1 in Individuals with Intellectual Disability and Epilepsy. CellPress. 2017;100(5):824–30.

162. Thyme SB, Pieper LM, Li EH, Pandey S, Wang Y, Morris NS, et al. Phenotypic Landscape of Schizophrenia-Associated Genes Defines Candidates and Their Shared Functions. Cell [Internet]. 2019;177(2):478–91. Available from: https://doi.org/10.1016/j.cell.2019.01.048

163. Sant'Anna Feltrin A, Tahira AC, Simões SN, Brentani H, Martins DC. Assessment of complementarity of WGCNA and NERI results for identification of modules associated to schizophrenia spectrum disorders. PLoS One [Internet]. 2019 [cited 2019 Jul 3];14(1):1–22. Available from: https://doi.org/10.1371/journal.pone.0210431

164. Lai T, Jabaudon D, Molyneaux BJ, Azim E, Arlotta P, Menezes JRL, et al. SOX5 Controls the Sequential Generation of Distinct Corticofugal Neuron Subtypes. Neuron. 2008;57:232–47.

165. Sun C, Ban Y, Wang KAI, Sun Y, Zhao Z. SOX5 promotes breast cancer proliferation and invasion by transactivation of EZH2. Oncol Lett. 2019;17(3):1–9.

166. You J, Zhao Q, Fan X, Wang J. SOX5 promotes cell invasion and metastasis via activation of twist-mediated epithelial– mesenchymal transition in gastric cancer. Onco Targets Ther. 2019;12:2465–76.

167. Fu Y, Shan X, Song W, Xu K, Jiao C, Zhang Q. Correlations of breast cancer FHIT gene with the incidence and prognosis of breast cancer. JBUON. 2019;24(9):40–7.

168. Sullivan CR, Funk AJ, Shan D, Haroutunian V, McCullumsmith RE. Decreased chloride channel expression in the dorsolateral prefrontal cortex in schizophrenia. PLoS One. 2015;10(3).

169. Lewis DA, Narendran R. In Vivo Measurement of GABA Transmission in Healthy Subjects and Schizophrenia Patients W. Am J Psychiatry. 2016;172(11):1148–59.

170. Silverstein SM, Rosen R. Cognition Schizophrenia and the eye. Schizophr Res. 2015;2(Cognition 2):46–55.

171. Almonte MT, Capellàn P, Yap TE, Cordeiro MF. Retinal correlates of psychiatric disorders. Ther Adv Chronic Dis. 2020;11:1–21.

172. Zhang L, Li X, Zhou R, Xing G. Possible role of potassium channel , big K in etiology of Schizophrenia. Med Hypotheses. 2006;67:41–3.

173. Yamamoto H, Kokame K, Okuda T, Nakajo Y, Yanamoto H, Miyata T. NDRG4 Protein-deficient Mice Exhibit Spatial Learning Deficits and Vulnerabilities to Cerebral Ischemia * □. J Biol Chem. 2011;286(29):26158–65.

174. F. Ducray, MD E. Crinie`re, MSc A. Idbaih, MD, PhD K. Mokhtari, MD Y. Marie, MSc S. Paris, MSc S. Navarro, MD F. Laigle-Donadey, MD C. Dehais, MD J. Thillet, PhD K. Hoang-Xuan, MD, PhD, J.-Y. Delattre, MD M. Sanson, MD P. alpha-Internexin expression identifies 1p19q codeleted gliomas. Neurology. 2009;72(2):156–61.

175. Jiang X, Detera-Wadleigh SD, Akula N, Mallon BS, Hou L, Xiao T, et al. Sodium valproate rescues expression of TRANK1 in iPSC-derived neural cells that carry a genetic variant associated with serious mental illness. Mol Psychiatry [Internet]. 2019;24(4):613–24. Available from: http://dx.doi.org/10.1038/s41380-018-0207-1

176. Li W, Cai X, Li HJ, Song M, Zhang CY, Yang Y, et al. Independent replications and integrative analyses confirm TRANK1 as a susceptibility gene for bipolar disorder. Neuropsychopharmacology. 2020;0:1–10.

177. Mossink B, Negwer M, Schubert D, Nadif Kasri N. The emerging role of chromatin remodelers in neurodevelopmental disorders: a developmental perspective [Internet]. Vol. 78, Cellular and Molecular Life Sciences. Springer International Publishing; 2021. 2517–2563 p. Available from: https://doi.org/10.1007/s00018-020-03714-5

178. Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW, et al. Genomic dissection of bipolar disorder and schizophrenia including 28 subphenotypes. bioRxiv. 2017;173(7):1705–15.

179. Laskaris LE, Di Biase MA, Everall I, Chana G, Christopoulos A, Skafidas E, et al. Microglial activation and progressive brain changes in schizophrenia. Br J Pharmacol. 2016;173(4):666–80.

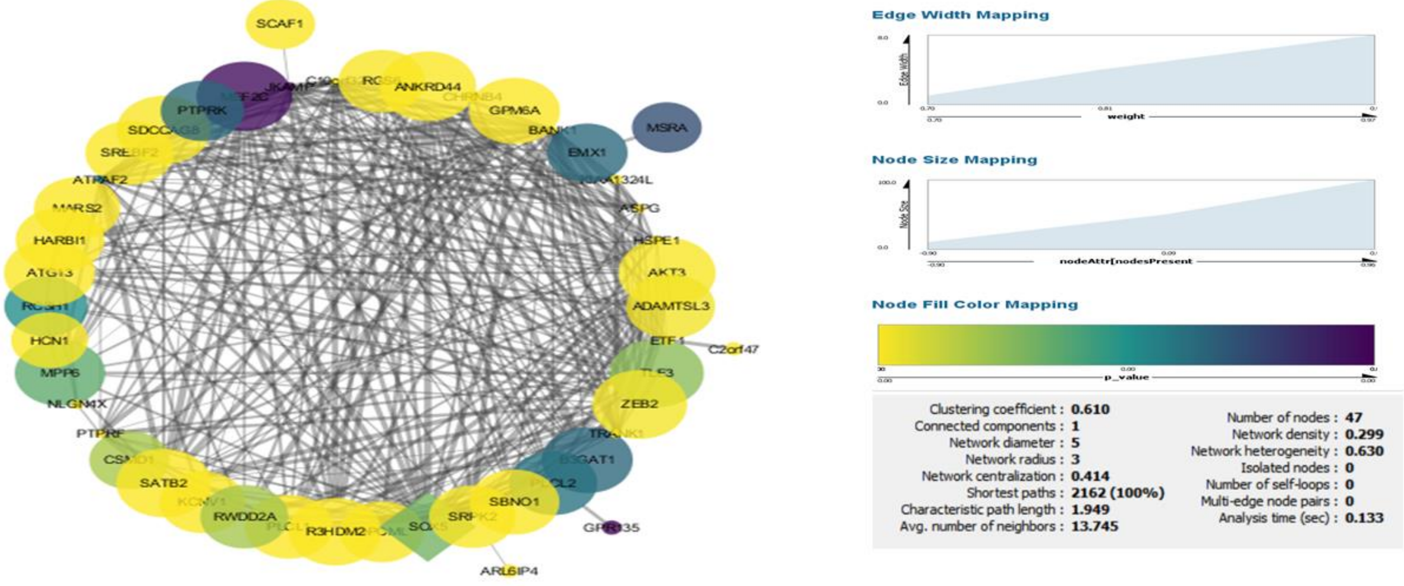
180. Li M, Santpere G, Kawasawa YI, Evgrafov O V., Gulden FO, Pochareddy S, et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. Science (80-). 2018;362(6420).

181. Tangliang Li1, Yue Shi2, Pei Wang1, Luis Miguel Guachalla3, Baofa Sun2, Tjard Joerss1, Yu-Sheng

Chen2, Marco Groth1, Anja Krueger1, Matthias Platzer1, Yun-Gui Yang2, Karl Lenhard Rudolph1, 3 & Zhao-Qi Wang1 4, Abstract. Smg6Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. Embo J. 2015;34(12).
182. Tao H, Zhou X, Xie Q, Ma Z, Sun F, Cui L, et al. SRR intronic variation inhibits expression of its neighbouring SMG6 gene and protects against temporal lobe epilepsy. J Cell Mol Med. 2018;22(3):1883–93.
183. Chi TH, Hsieh BY, Liang PS, Han TH, Hsieh M. Expression and Functional Study of Single Mutations of Carbonic Anhydrase 8 in Neuronal Cells. Cell Mol Neurobiol [Internet]. 2020;(0123456789). Available from: https://doi.org/10.1007/s10571-020-00907-1
184. Citton V, Favaro A, Bettini V, Gabrieli J, Milan G, Greggio NA, et al. Brain involvement in Alström syndrome. Ophanet J Rare Dis. 2013;8(24).
185. Wang X, Christian KM, Song H, Ming G. Synaptic dysfunction in complex psychiatric disorders : from genetics to mechanisms. Genome Med. 2018;10(9):9–11.
186. Bennett MR. Synapse formation and regression in the cortex during adolescence and in schizophrenia. Med J Aust. 2009;190(4):4–6.
187. Minami T. Calcineurin-NFAT activation and DSCR-1 auto-inhibitory loop : how is homoeostasis regulated ? J Biochem. 2014;155(4):217–26.
188. Kipanyula MJ, Kimaro WH, Etet PFS. The Emerging Roles of the Calcineurin-Nuclear Factor of Activated T-Lymphocytes Pathway in Nervous System Functions and Diseases. J Aging Res. 2016;5081021.
189. Wei Han, Tianxiao Zhang, Tong Ni, Li Zhu, Dan Liu, Gang Chen, Huali Lin, Teng Chen FG. Relationship of common variants in CHRNA5 with early-onset schizophrenia and executive function. Schizophr Res. 2019;206:407–12.
190. Yuhong Wang, Yuanjia Chen, Xiaoxing Li, Wanming Hu, Yu Zhang, Luohai Chen MC and JC. Loss of expression and prognosis value of alpha-internexin in gastroenteropancreatic neuroendocrine neoplasm. BMC Cancer. 2018;18(691).
191. Xi Lin, Mingyue Li, Niandong Wang, Yiran Wu, Zhipu Luo, Shimeng Guo, Gye-Won Han, Shaobai Li, Yang Yue1, Xiaohu Wei Xin Xie, Yong Chen, Suwen Zhao, Jian Wu ML& FX. Structural basis of ligand recognition and self-activation of orphan GPR52. Nature. 2020;579(7797):152–7.
192. Hidetoshi Komatsu MF and YH. Potential Utility of Biased GPCR Signaling for Treatment of Psychiatric Disorders. Int J Mol Sci. 2019;20(3207).
193. Orlandi C, Watkins LR. Orphan g protein coupled receptors in affective disorders. Genes (Basel). 2020;11(6):1–32.
194. Pingyuan Wang, Daniel E. Felsing, Haiying Chen, Sonja J. Stutz, Ryan E. Murphy, Kathryn A. Cunningham JAA and J. Discovery of Potent and Brain-Penetrant GPR52 Agonist that Suppresses Psychostimulant Behavior. J Med Chem. 2020;22:13951–13972.
195. Nishiyama K, Suzuki H, Harasawa T, Suzuki N, Kurimoto E, Kawai T, et al. FTBMT, a novel and selective GPR52 agonist, demonstrates antipsychotic-like and procognitive effects in rodents, revealing a potential therapeutic agent for schizophrenia. J Pharmacol Exp Ther. 2017;363(2):253–64.
196. Roy SK, Shrivastava A, Srivastav S, Shankar S, Srivastava RK. SATB2 is a novel biomarker and therapeutic target for cancer. J Cell Mol Med. 2020;24(19):11064–9.
197. Gary Donohoe DWMICLW, Georg Dechant GA. Genes encoding SATB2-interacting proteins in adult cerebral cortex contribute to human cognitive ability. PLoS Genet. 2019;15(2).
198. Lu X, Chen X, Huang L, Zhu C, Gu Y, Ye S. Anti- a -Internexin Autoantibody from Neuropsychiatric Lupus Induce Cognitive Damage via Inhibiting Axonal Elongation and Promote Neuron Apoptosis. PLoS One. 2010;5(6).
199. Müller N, Weidinger E, Leitner B, Schwarz MJ. The role of inflammation in schizophrenia. Front Neurosci. 2015;9(372).
200. Monji A, Kato TA, Mizoguchi Y, Horikawa H, Seki Y, Kasai M, et al. Neuroinflammation in schizophrenia especially focused on the role of microglia. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2013;42:115–21. Available from: http://dx.doi.org/10.1016/j.pnpbp.2011.12.002

201. Pepe G, De Maglie M, Minoli L, Villa A, Maggi A, Vegeto E. Selective proliferative response of microglia to alternative polarization signals. J Neuroinflammation. 2017;14(236).
202. Bachiller S, Jiménez-Ferrer I, Paulus A, Yang Y, Swanberg M, Deierborg T, et al. Microglia in neurological diseases: A road map to brain-disease dependent-inflammatory response. Front Cell Neurosci. 2018;12(488).
203. Sellgren CM, Gracias J, Watmuff B, Biag JD, Thanos JM, Whittredge PB, et al. Increased synapse elimination by microglia in schizophrenia patient-derived models of synaptic pruning. Nat Neurosci [Internet]. 2019;22(3):374–85. Available from: http://dx.doi.org/10.1038/s41593-018-0334-7
204. Miller BJ, Goldsmith DR. Towards an Immunophenotype of Schizophrenia: Progress, Potential Mechanisms, and Future Directions. Neuropsychopharmacology. 2017;42:299–317.
205. Nagatsu T. The catecholamine system in health and disease - Relation to tyrosine 3-monooxygenase and other catecholamine-synthesizing enzymes. Proc Japan Acad Ser B Phys Biol Sci. 2006;82:388–415.
206. Fujita H, Carlson ES. Catecholamine signaling that modulates cerebellar operations in cognition. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 2021.
207. Vidal PM, Pacheco R. The Cross-Talk Between the Dopaminergic and the Immune System Involved in Schizophrenia. Front Pharmacol. 2020;11(394).
208. Sarkar C, Basu B, Chakroborty D, Dasgupta PS, Basu S. The immunoregulatory role of dopamine: An update. Brain Behav Immun [Internet]. 2010;24(4):525–8. Available from: http://dx.doi.org/10.1016/j.bbi.2009.10.015
209. Catak Z, Kocdemir E, Ugur K, Yardim M, Sahin I, Kaya H, et al. A novel biomarker renalase and its relationship with its substrates in schizophrenia. J Med Biochem. 2019;38(3):299–305.
210. Crider A, Pillai A. Estrogen signaling as a therapeutic target in neurodevelopmental disorders. J Pharmacol Exp Ther. 2017;360(1):48–58.
211. Gogos A, Sbisa AM, Sun J, Gibbons A, Udawela M, Dean B. A Role for Estrogen in Schizophrenia: Clinical and Preclinical Findings. Int J Endocrinol. 2015;2015(615356).
212. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45(6):580–5.
213. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, et al. Comprehensive functional genomic resource and integrative model for the human brain. Science (80-). 2018;362(6420).
214. Chen G, Ning B, Shi T. Single-cell RNA-seq technologies and related computational data analysis. Front Genet. 2019;10(317).
215. Hedlund E, Deng Q. Single-cell RNA sequencing : Technical advancements and biological applications. Mol Aspects Med [Internet]. 2018;59:36–46. Available from: https://doi.org/10.1016/j.mam.2017.07.003
216. Botía JA, Vandrovcova J, Forabosco P, Guelfi S, D'Sa K, Hardy J, et al. An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. BMC Syst Biol. 2017;
217. Barrangou R, Doudna JA. Applications of CRISPR technologies in research and beyond. Nat Biotechnol. 2016;34(9):933–41.
218. Chen CH, Xiao T, Xu H, Jiang P, Meyer CA, Li W, et al. Improved design and analysis of CRISPR knockout screens. Bioinformatics. 2018;34(23):4095–101.

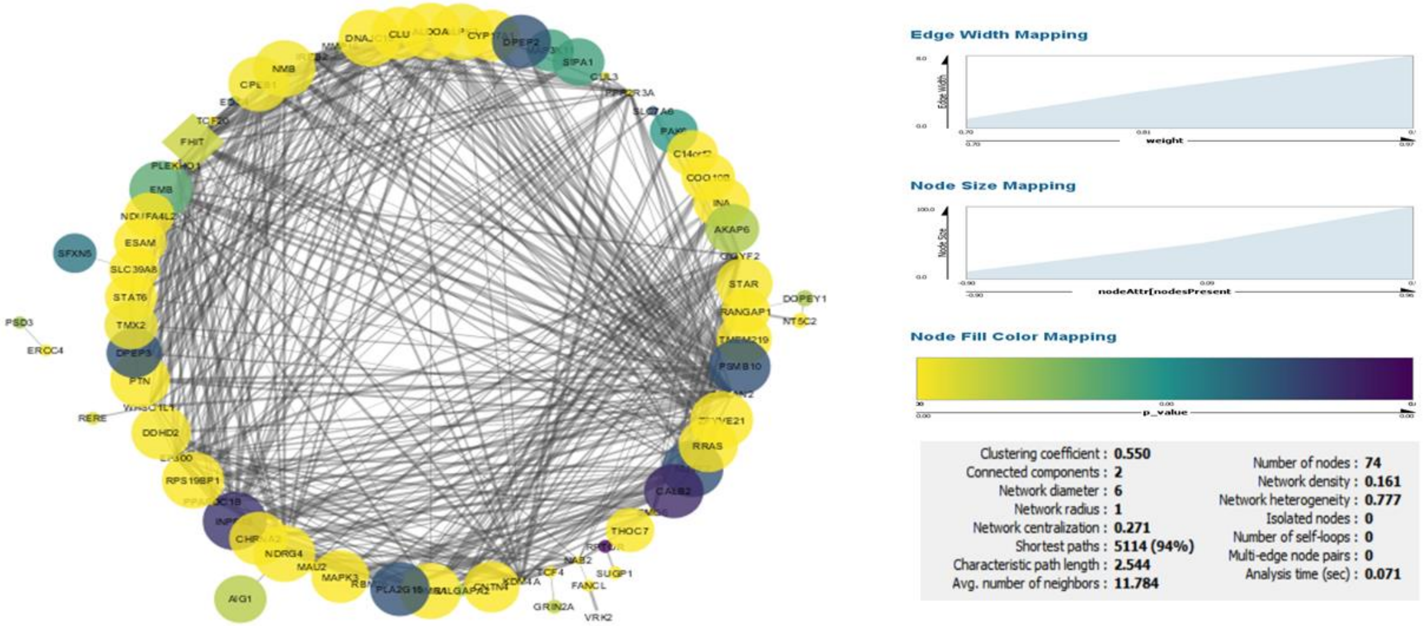
6.0 Appendix



Clustering coefficient : **0.610**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.414**
Shortest paths : **2162 (100%)**
Characteristic path length : **1.949**
Avg. number of neighbors : **13.745**

Number of nodes : **47**
Network density : **0.299**
Network heterogeneity : **0.630**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.133**

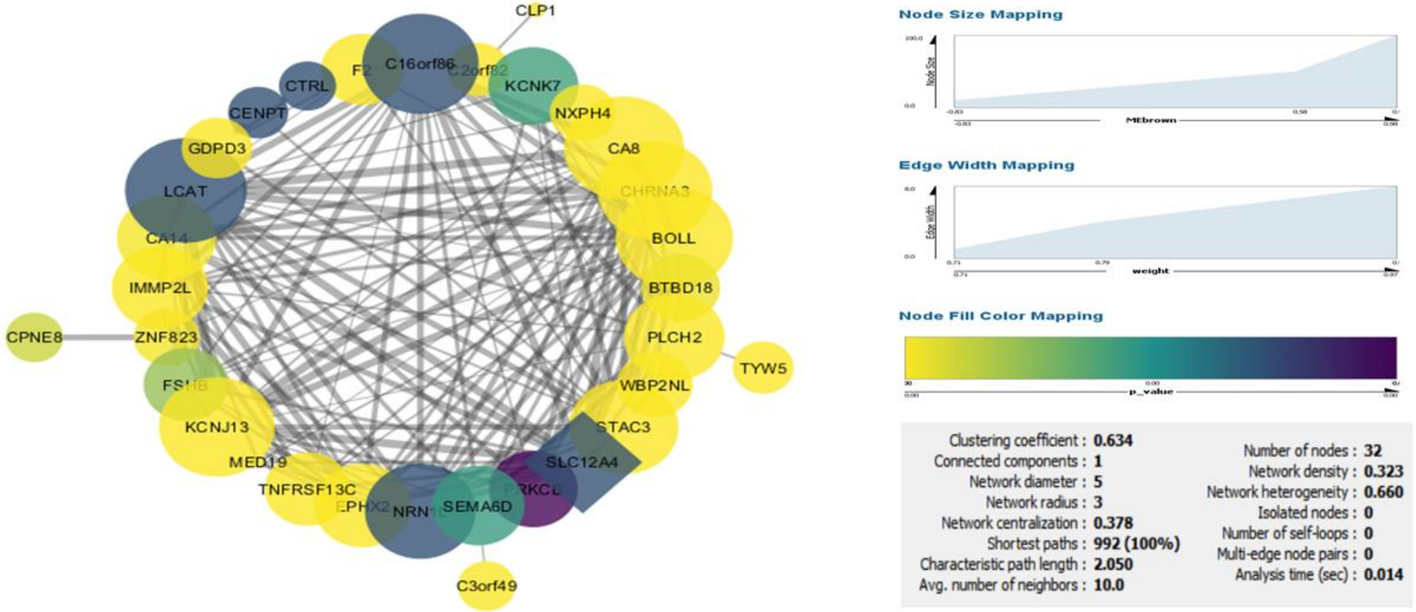
Figure 6.1.1 Black Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.550**
Connected components : **2**
Network diameter : **6**
Network radius : **1**
Network centralization : **0.271**
Shortest paths : **5114 (94%)**
Characteristic path length : **2.544**
Avg. number of neighbors : **11.784**

Number of nodes : **74**
Network density : **0.161**
Network heterogeneity : **0.777**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.071**

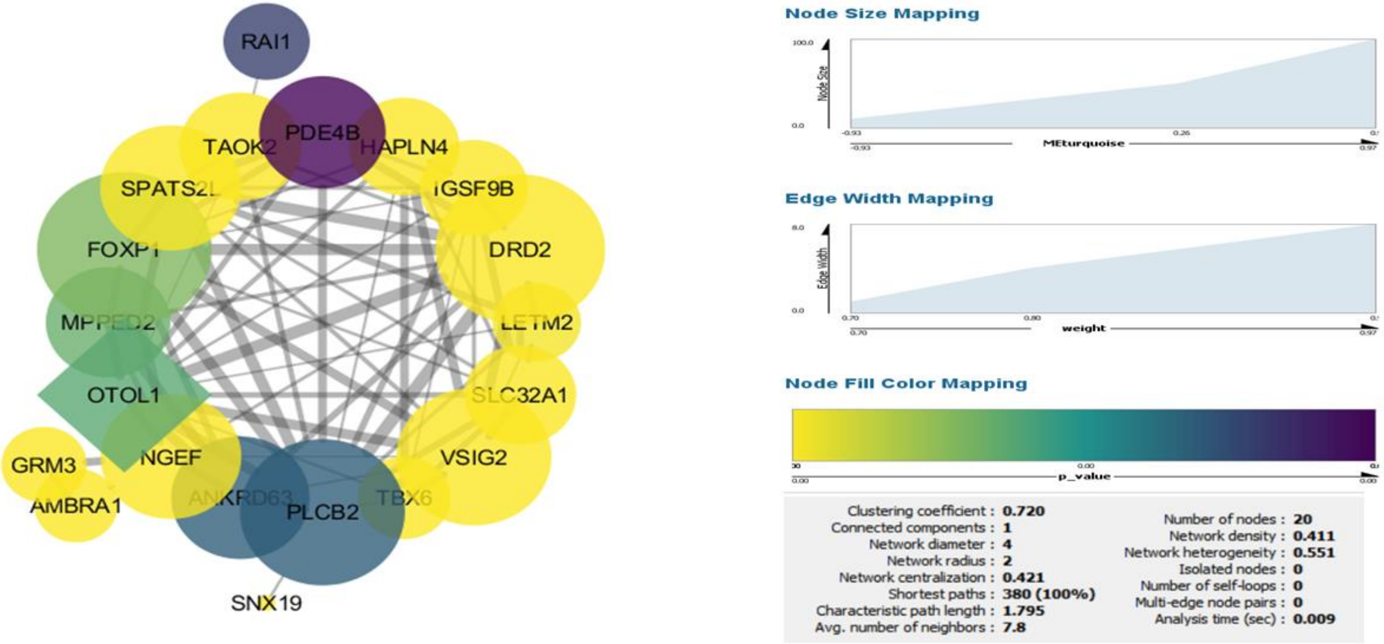
Figure 6.1.2 Blue Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.634**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.378**
Shortest paths : **992 (100%)**
Characteristic path length : **2.050**
Avg. number of neighbors : **10.0**

Number of nodes : **32**
Network density : **0.323**
Network heterogeneity : **0.660**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.014**

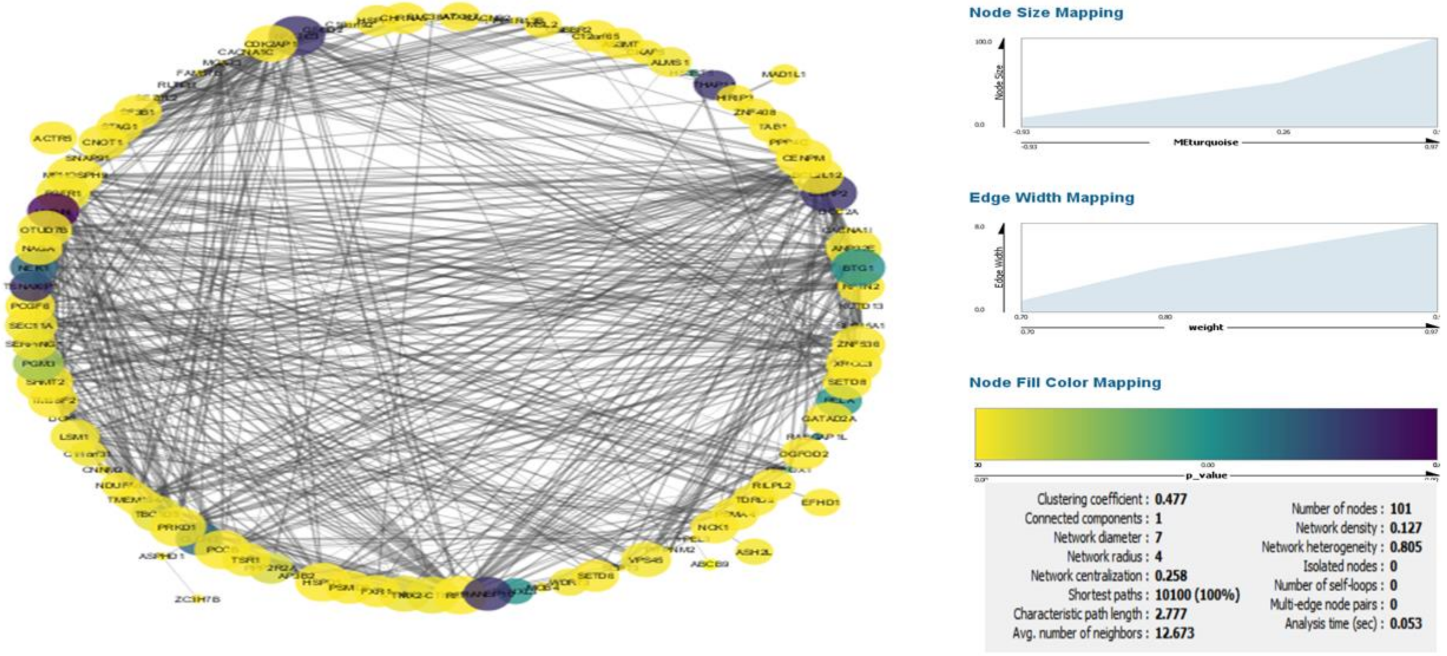
Figure 6.1.3 Brown Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.720**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.421**
Shortest paths : **380 (100%)**
Characteristic path length : **1.795**
Avg. number of neighbors : **7.8**

Number of nodes : **20**
Network density : **0.411**
Network heterogeneity : **0.551**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.009**

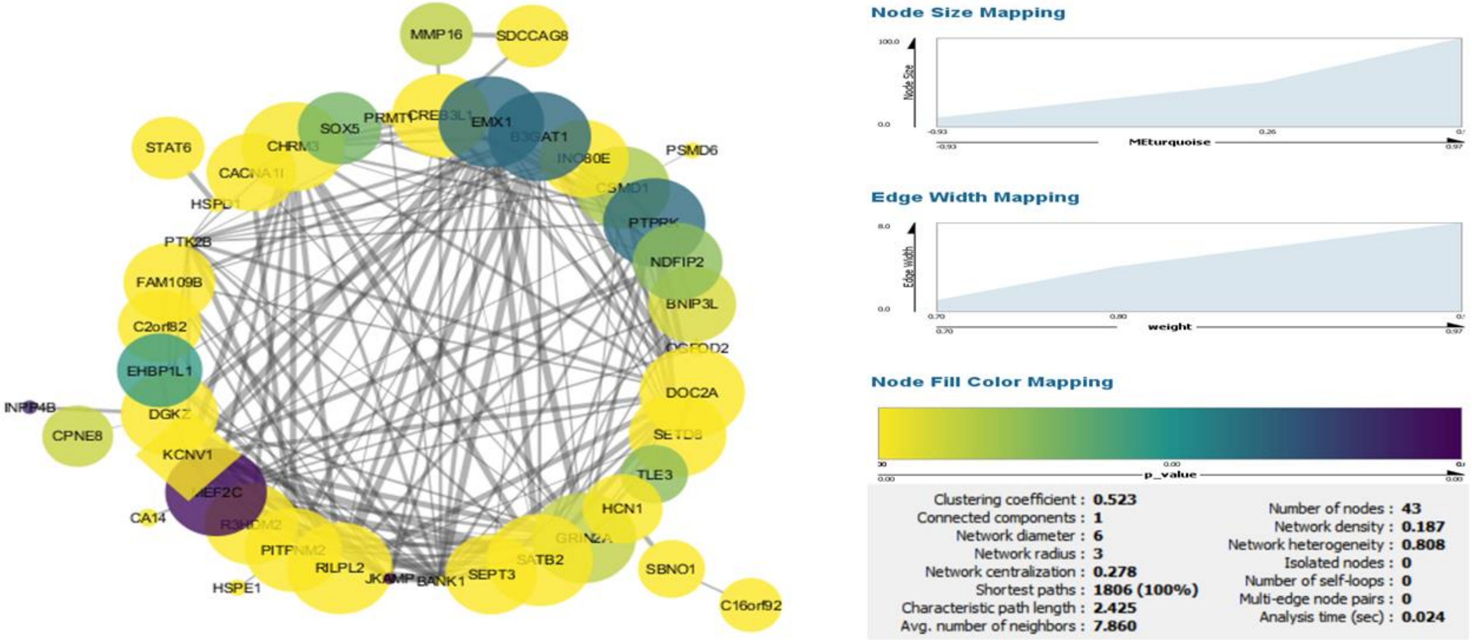
Figure 6.1.4 Pink Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : 0.477
Connected components : 1
Network diameter : 7
Network radius : 4
Network centralization : 0.258
Shortest paths : 10100 (100%)
Characteristic path length : 2.777
Avg. number of neighbors : 12.673

Number of nodes : 101
Network density : 0.127
Network heterogeneity : 0.805
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.053

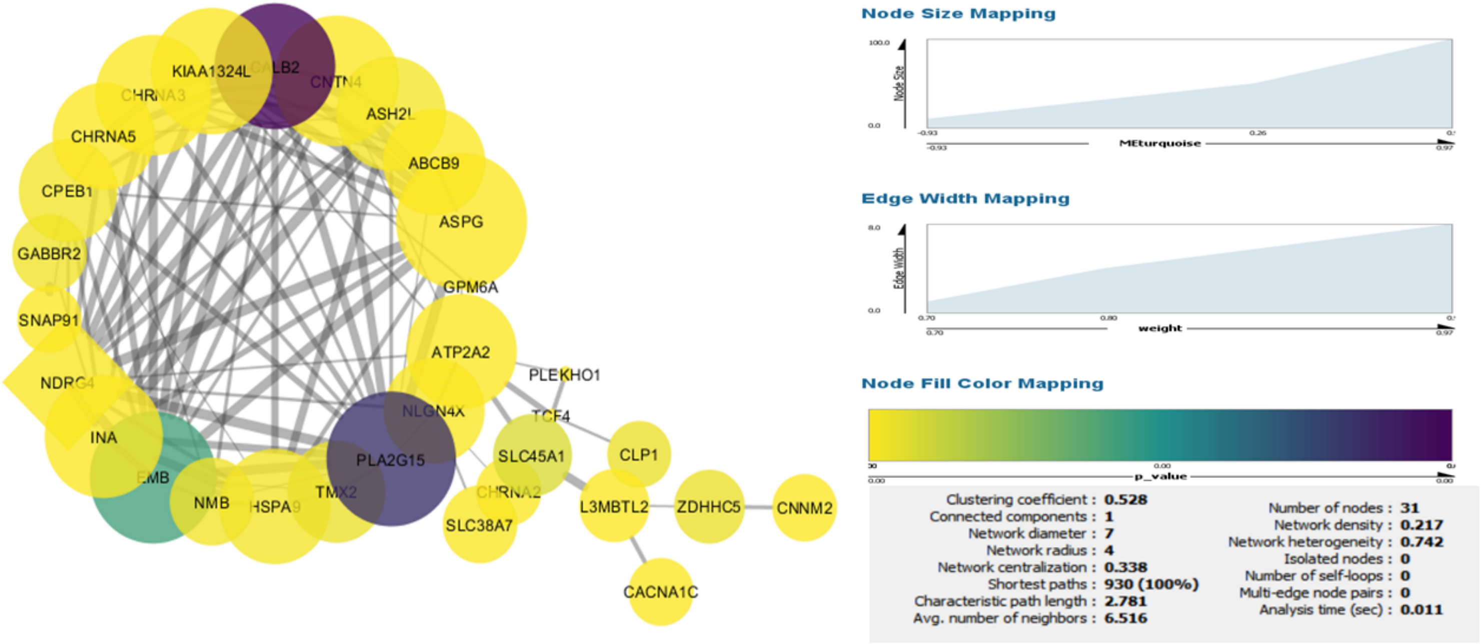
Figure 6.1.5 Turquoise Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : 0.523
Connected components : 1
Network diameter : 6
Network radius : 3
Network centralization : 0.278
Shortest paths : 1806 (100%)
Characteristic path length : 2.425
Avg. number of neighbors : 7.860

Number of nodes : 43
Network density : 0.187
Network heterogeneity : 0.808
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.024

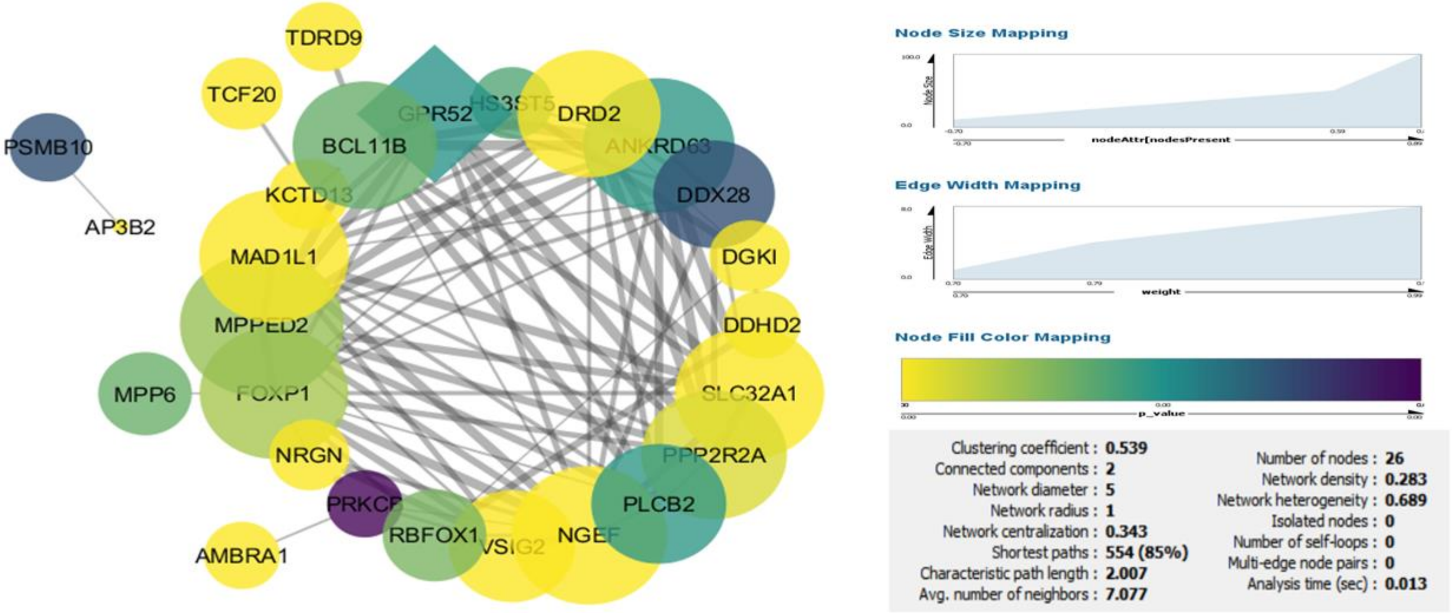
Figure 6.1.6 Blue Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.528**
Connected components : **1**
Network diameter : **7**
Network radius : **4**
Network centralization : **0.338**
Shortest paths : **930 (100%)**
Characteristic path length : **2.781**
Avg. number of neighbors : **6.516**

Number of nodes : **31**
Network density : **0.217**
Network heterogeneity : **0.742**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.011**

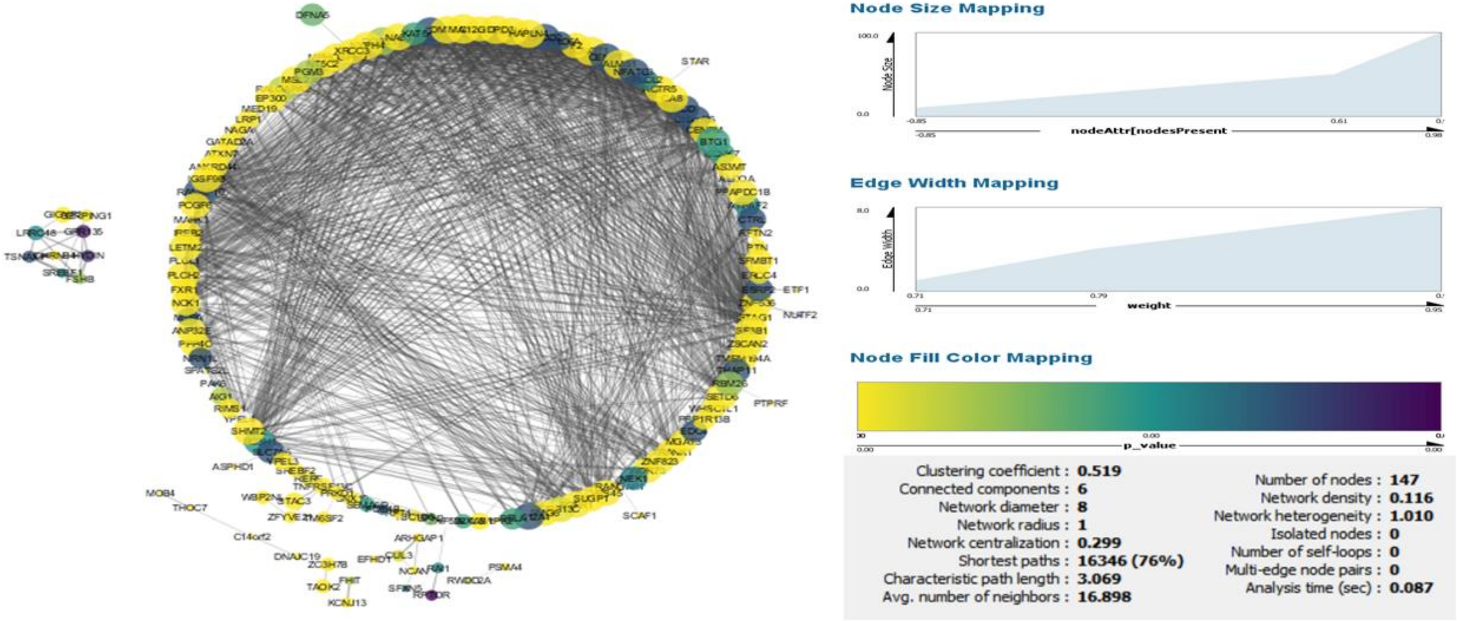
Figure 6.1.7 Brown Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.539**
Connected components : **2**
Network diameter : **5**
Network radius : **1**
Network centralization : **0.343**
Shortest paths : **554 (85%)**
Characteristic path length : **2.007**
Avg. number of neighbors : **7.077**

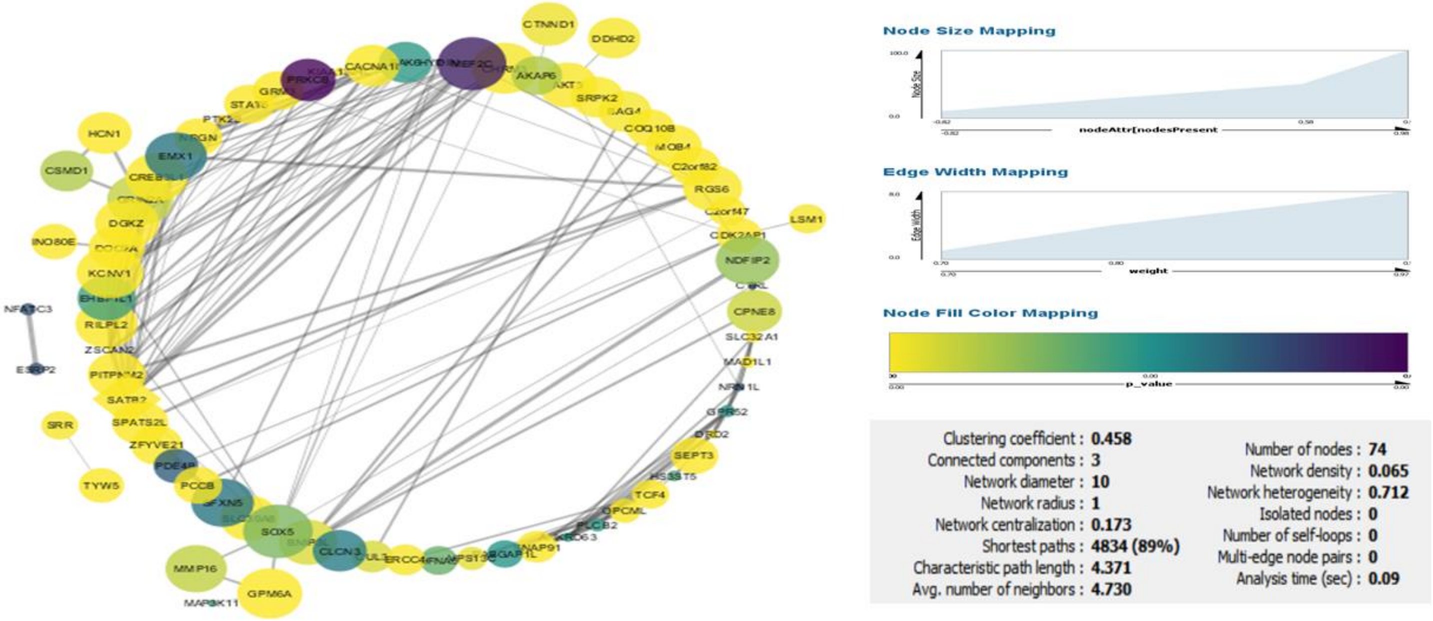
Number of nodes : **26**
Network density : **0.283**
Network heterogeneity : **0.689**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.013**

Figure 6.1.8 Green Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



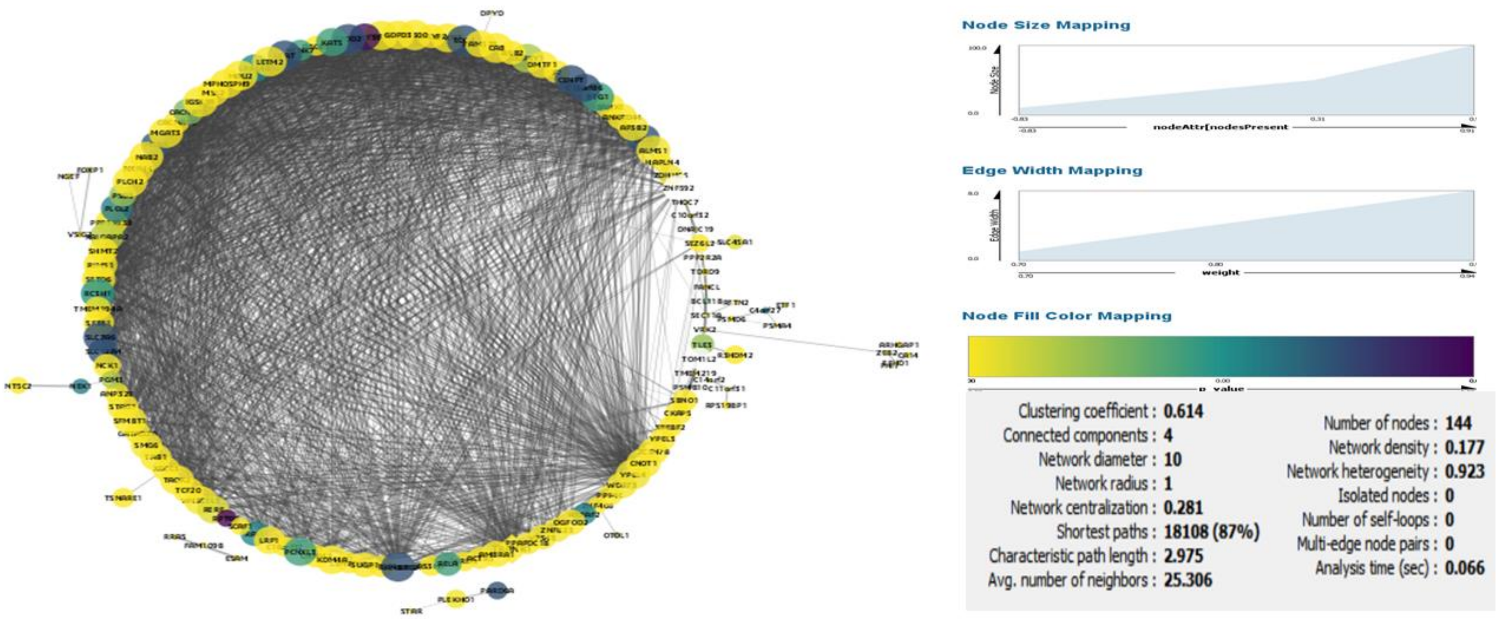
Clustering coefficient : **0.519**
Connected components : **6**
Network diameter : **8**
Network radius : **1**
Network centralization : **0.299**
Shortest paths : **16346 (76%)**
Characteristic path length : **3.069**
Avg. number of neighbors : **16.898**
Number of nodes : **147**
Network density : **0.116**
Network heterogeneity : **1.010**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.087**

Figure 6.1.9 Turquoise Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



```
     Clustering coefficient : 0.458           Number of nodes : 74
     Connected components : 3                 Network density : 0.065
          Network diameter : 10          Network heterogeneity : 0.712
            Network radius : 1                  Isolated nodes : 0
   Network centralization : 0.173          Number of self-loops : 0
           Shortest paths : 4834 (89%)    Multi-edge node pairs : 0
Characteristic path length : 4.371          Analysis time (sec) : 0.09
  Avg. number of neighbors : 4.730
```

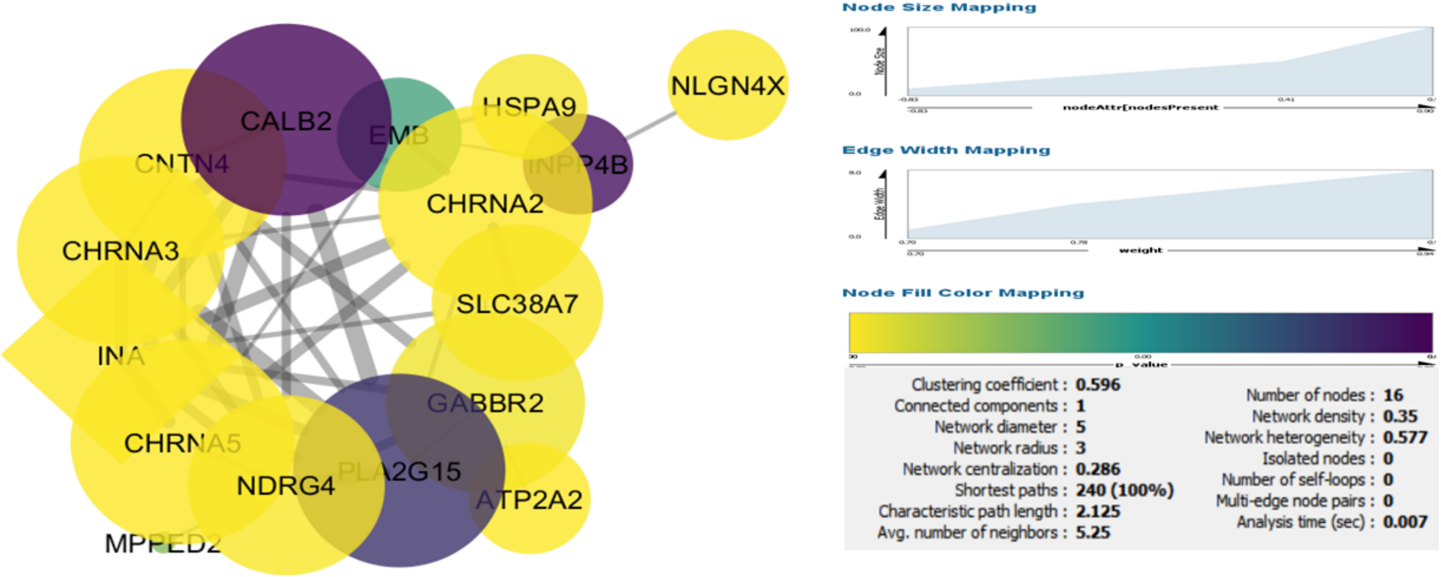
Figure 3.6.10 Brown Module for developmental Stage Three where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.614**
Connected components : **4**
Network diameter : **10**
Network radius : **1**
Network centralization : **0.281**
Shortest paths : **18108 (87%)**
Characteristic path length : **2.975**
Avg. number of neighbors : **25.306**

Number of nodes : **144**
Network density : **0.177**
Network heterogeneity : **0.923**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.066**

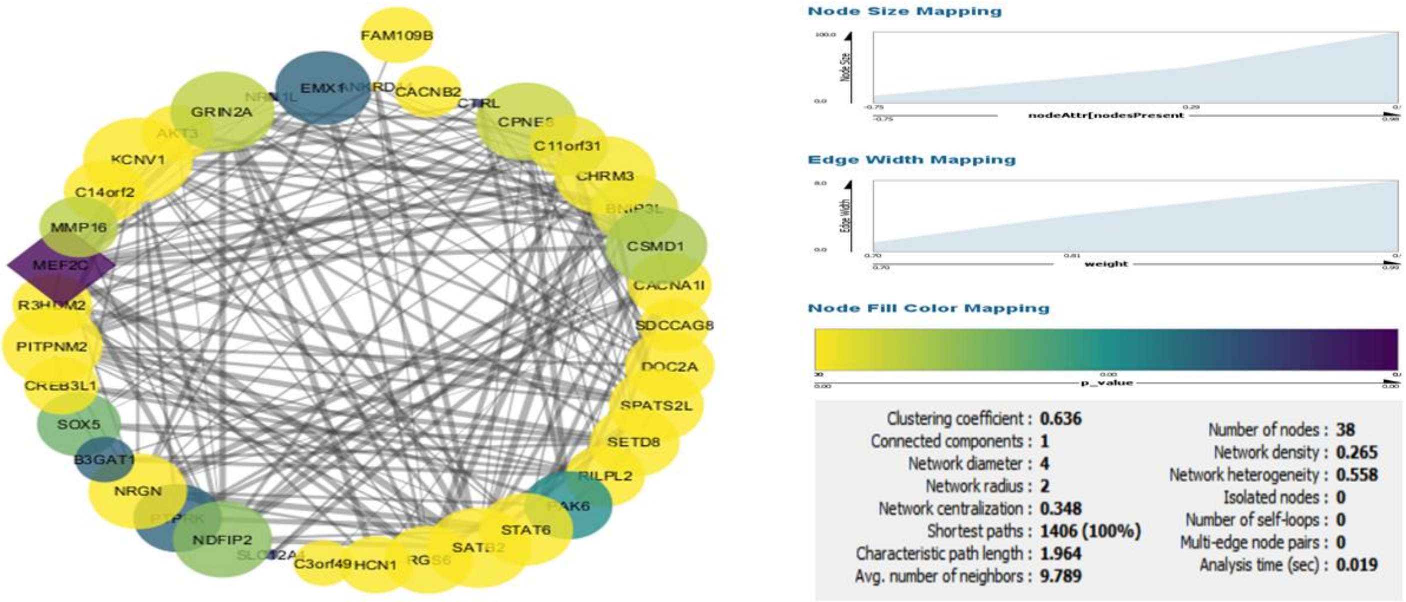
Figure 6.1.11 Turquoise Module for developmental Stage Three where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : 0.596
Connected components : 1
Network diameter : 5
Network radius : 3
Network centralization : 0.286
Shortest paths : 240 (100%)
Characteristic path length : 2.125
Avg. number of neighbors : 5.25

Number of nodes : 16
Network density : 0.35
Network heterogeneity : 0.577
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.007

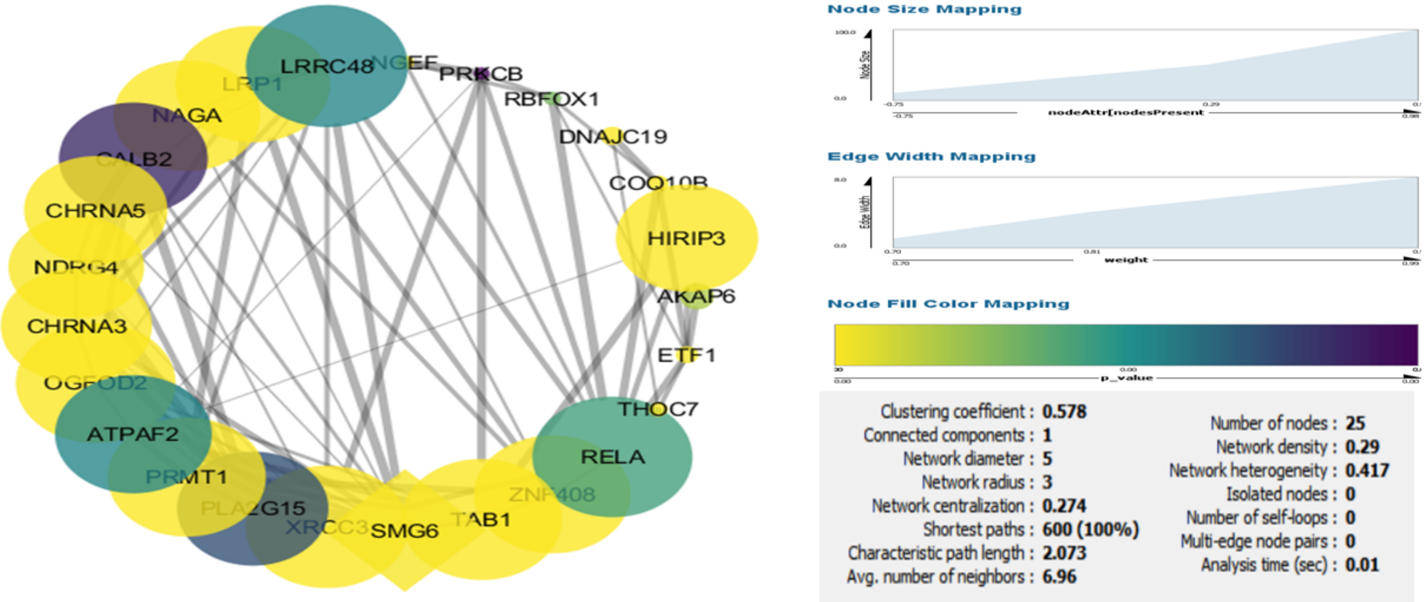
Figure 6.1.12 Yellow Module for developmental Stage Three where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.636**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.348**
Shortest paths : **1406 (100%)**
Characteristic path length : **1.964**
Avg. number of neighbors : **9.789**

Number of nodes : **38**
Network density : **0.265**
Network heterogeneity : **0.558**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.019**

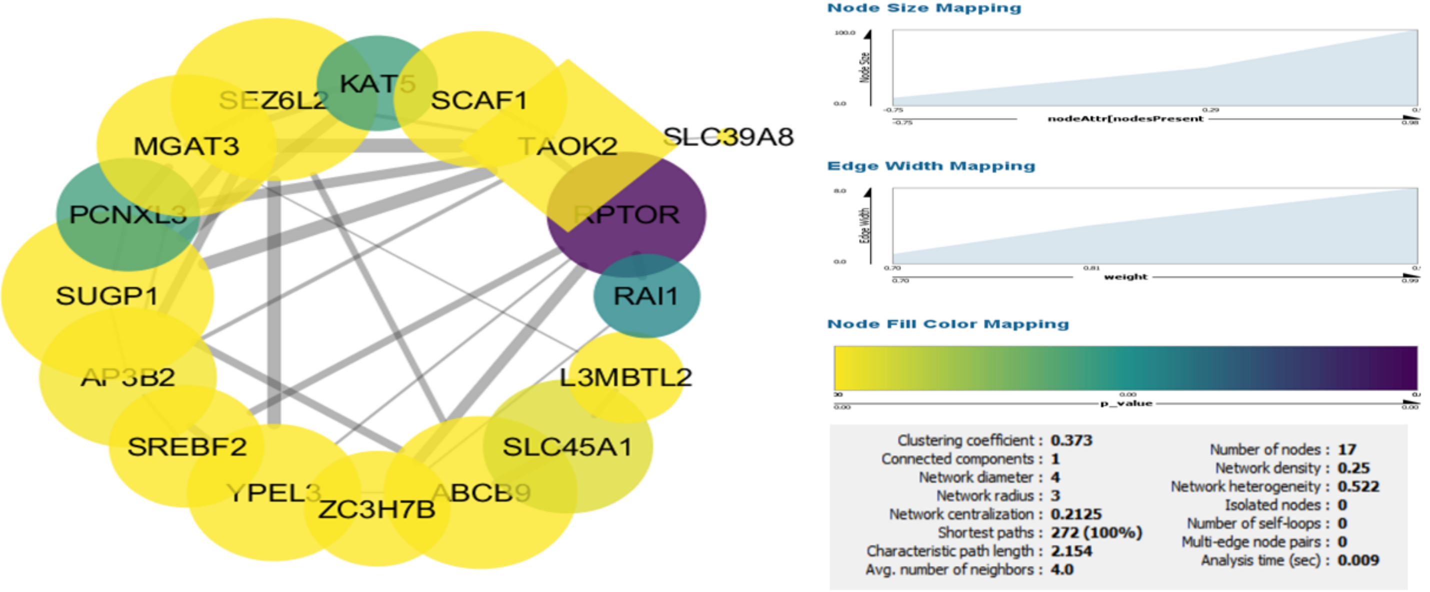
Figure 6.1.13 Blue Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.578**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.274**
Shortest paths : **600 (100%)**
Characteristic path length : **2.073**
Avg. number of neighbors : **6.96**

Number of nodes : **25**
Network density : **0.29**
Network heterogeneity : **0.417**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.01**

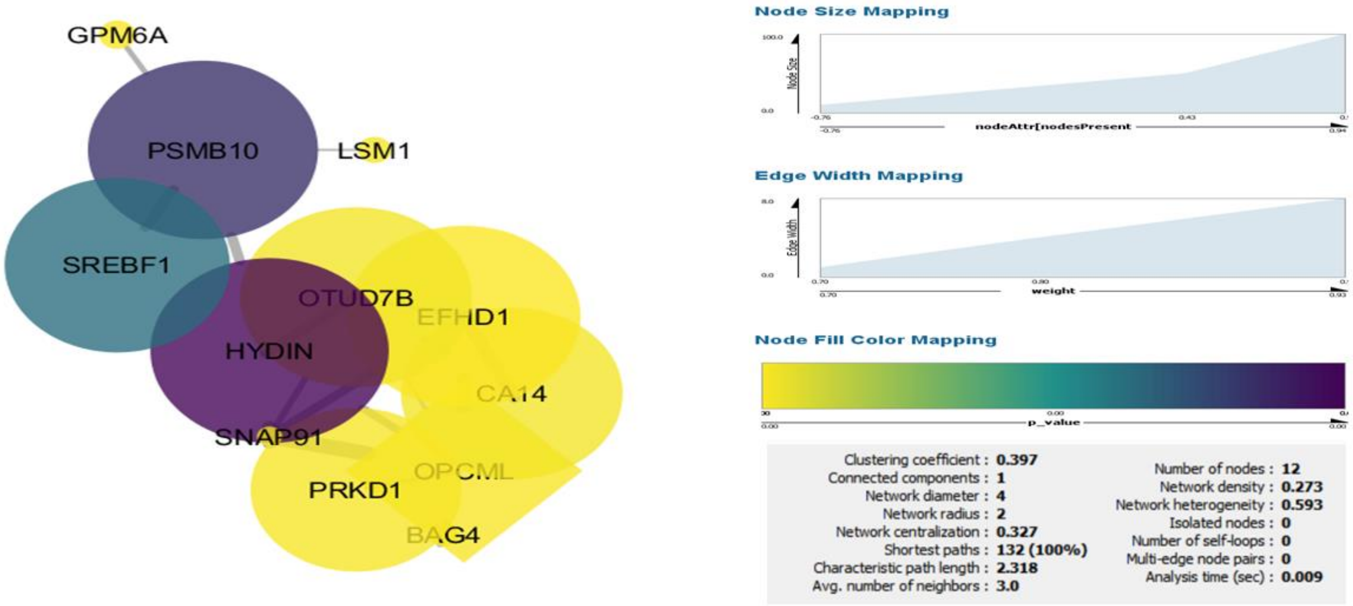
Figure 6.1.14 Brown Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.373**
Connected components : **1**
Network diameter : **4**
Network radius : **3**
Network centralization : **0.2125**
Shortest paths : **272 (100%)**
Characteristic path length : **2.154**
Avg. number of neighbors : **4.0**

Number of nodes : **17**
Network density : **0.25**
Network heterogeneity : **0.522**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.009**

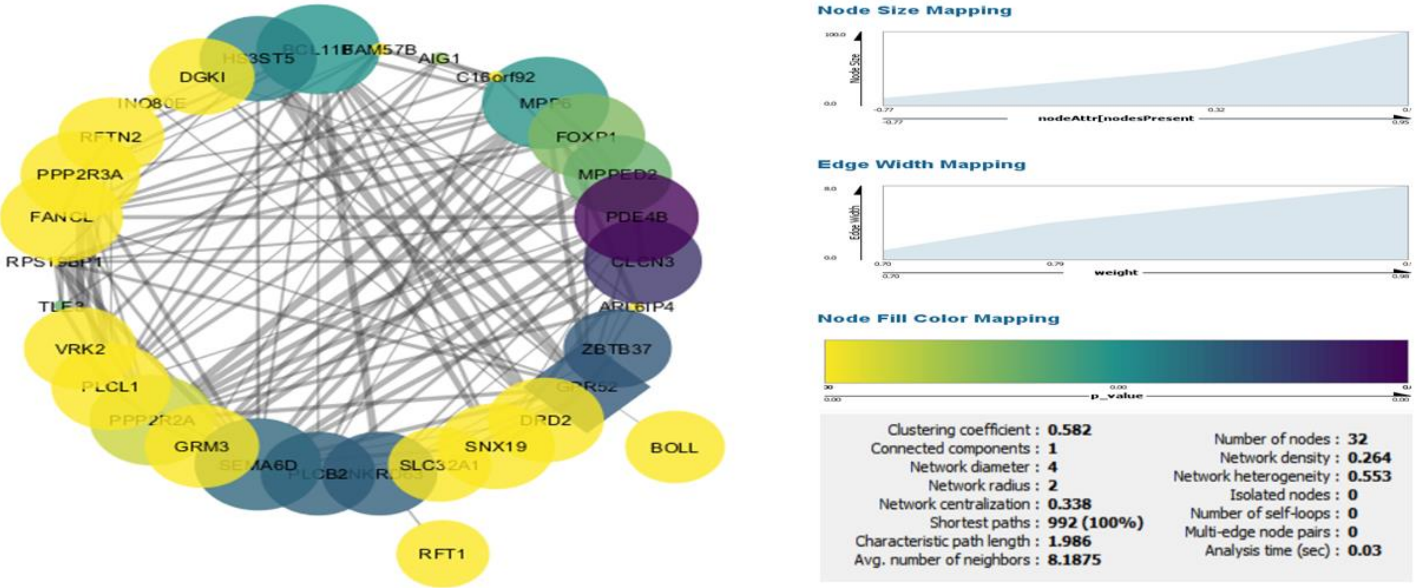
Figure 6.1.15 Green Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.397**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.327**
Shortest paths : **132 (100%)**
Characteristic path length : **2.318**
Avg. number of neighbors : **3.0**

Number of nodes : **12**
Network density : **0.273**
Network heterogeneity : **0.593**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.009**

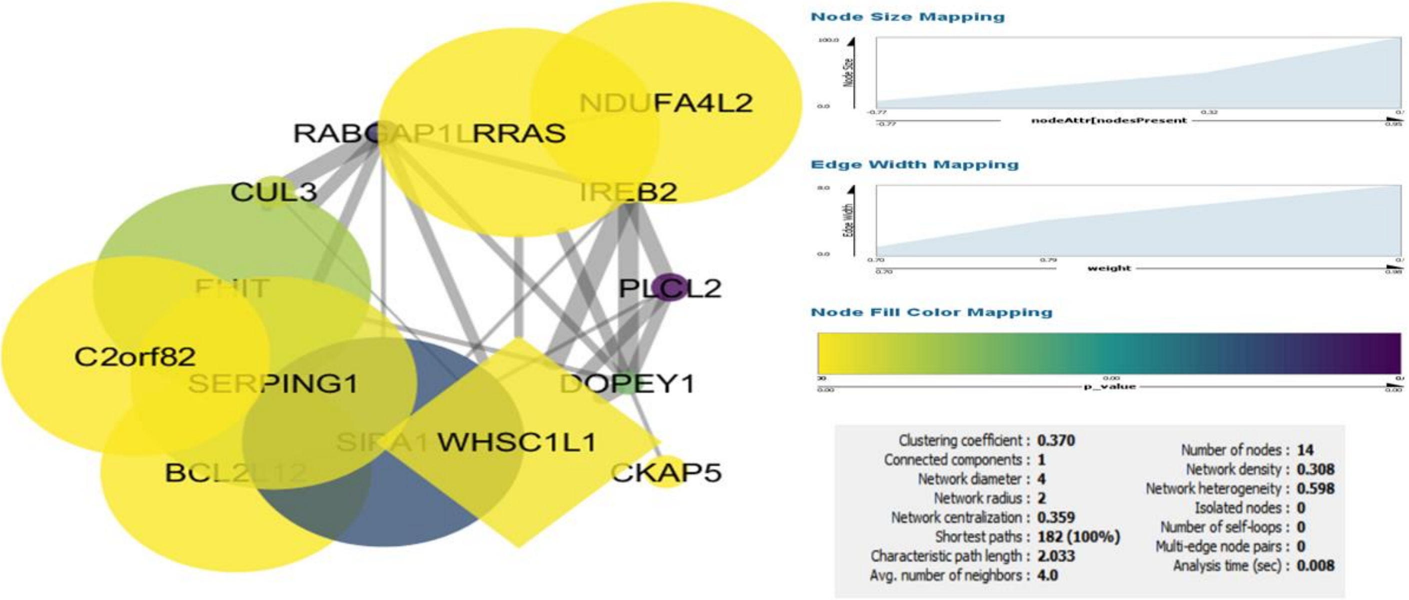
Figure 6.1.16 Magenta Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.582**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.338**
Shortest paths : **992 (100%)**
Characteristic path length : **1.986**
Avg. number of neighbors : **8.1875**

Number of nodes : **32**
Network density : **0.264**
Network heterogeneity : **0.553**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.03**

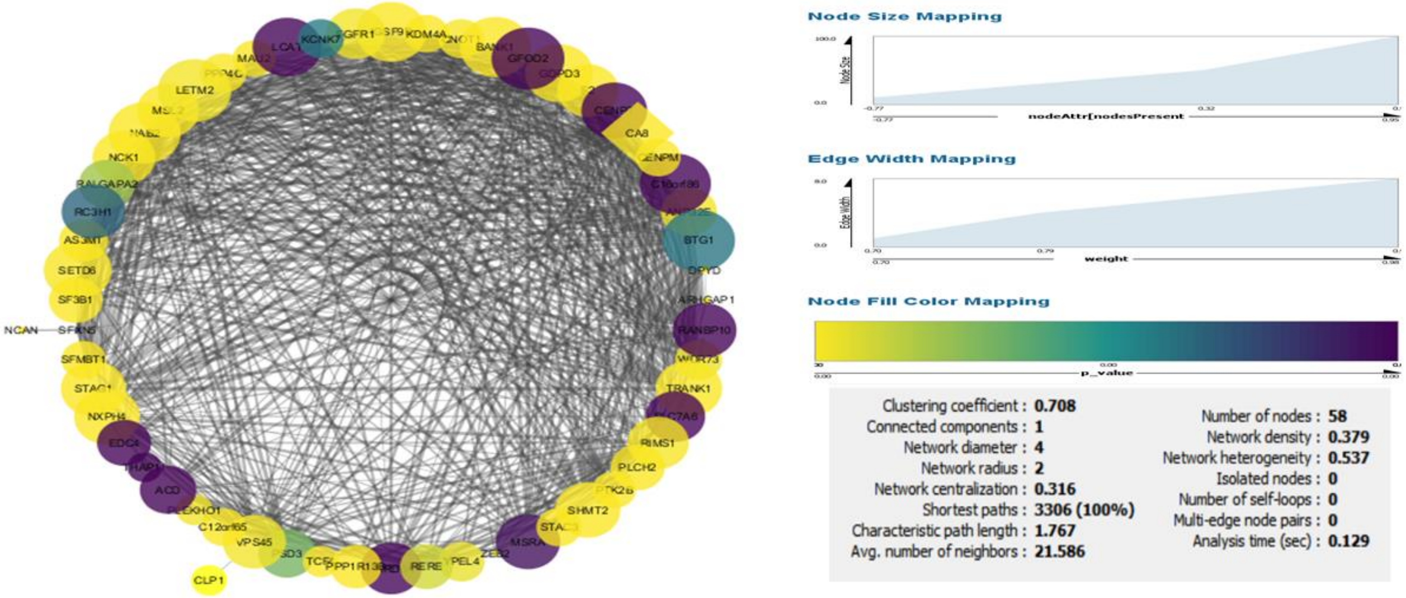
Figure 6.1.17 Purple Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.370**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.359**
Shortest paths : **182 (100%)**
Characteristic path length : **2.033**
Avg. number of neighbors : **4.0**

Number of nodes : **14**
Network density : **0.308**
Network heterogeneity : **0.598**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.008**

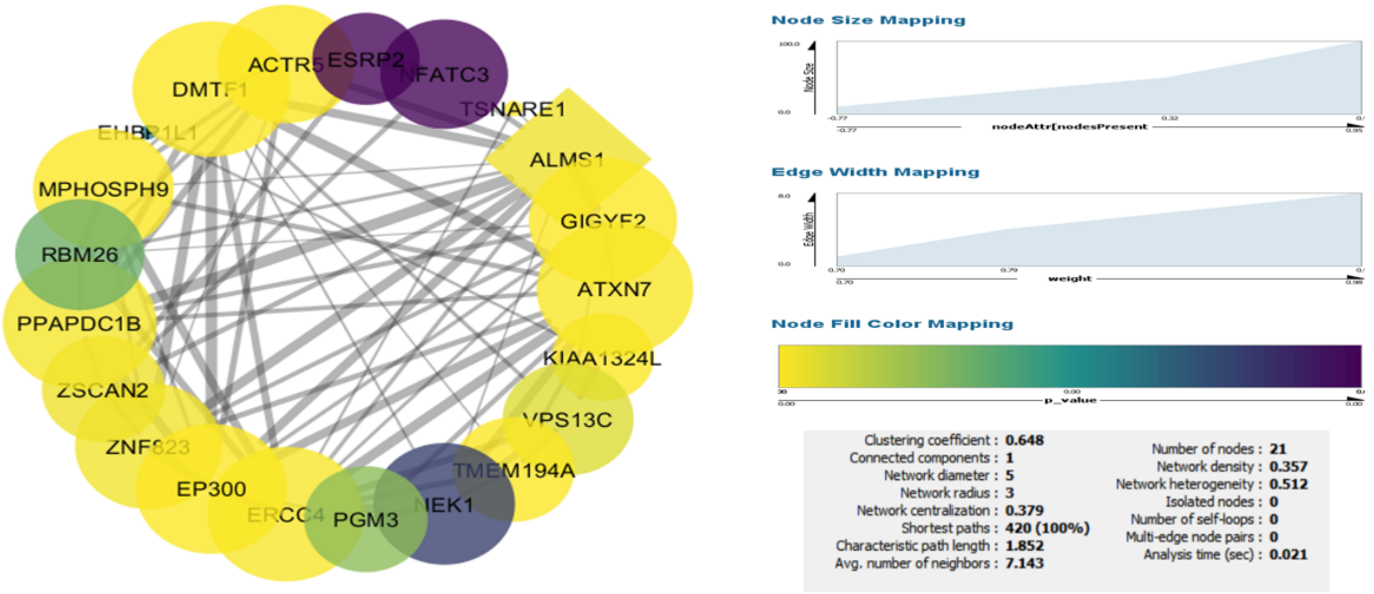
Figure 6.1.18 Red Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.708**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.316**
Shortest paths : **3306 (100%)**
Characteristic path length : **1.767**
Avg. number of neighbors : **21.586**

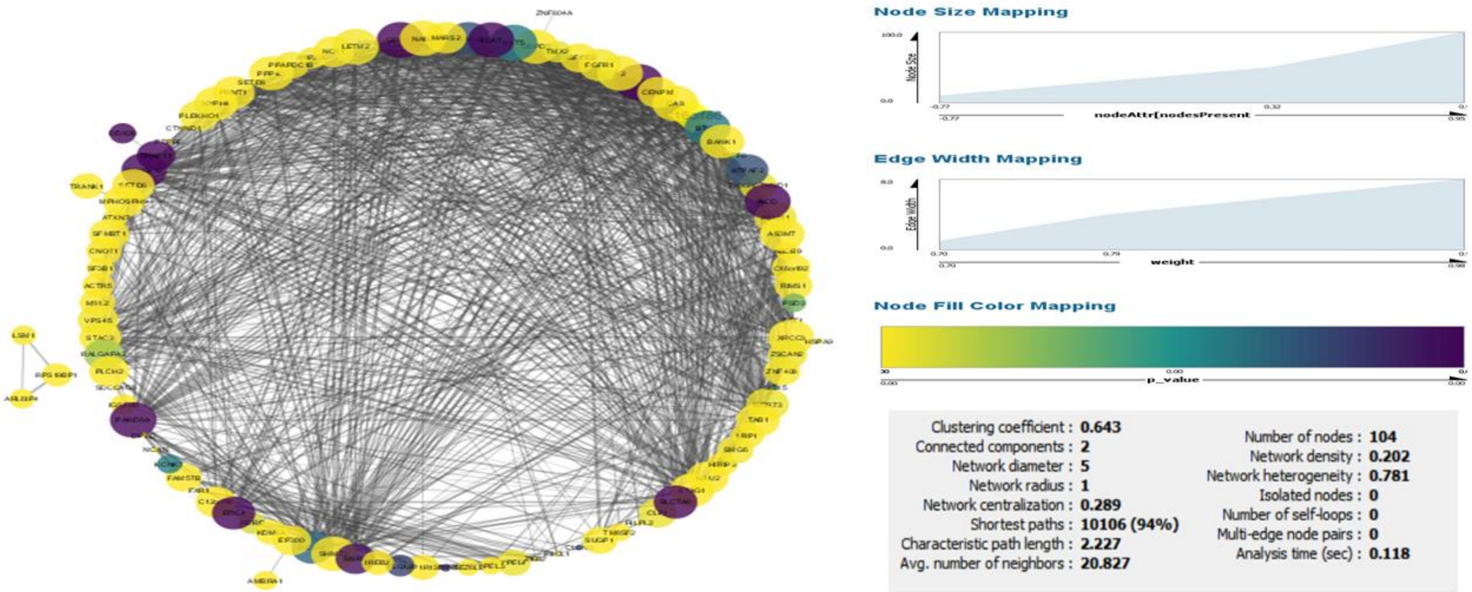
Number of nodes : **58**
Network density : **0.379**
Network heterogeneity : **0.537**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.129**

Figure 6.1.19 Turquoise Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



```
Clustering coefficient : 0.648
Connected components : 1
Network diameter : 5
Network radius : 3
Network centralization : 0.379
Shortest paths : 420 (100%)
Characteristic path length : 1.852
Avg. number of neighbors : 7.143

Number of nodes : 21
Network density : 0.357
Network heterogeneity : 0.512
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.021
```

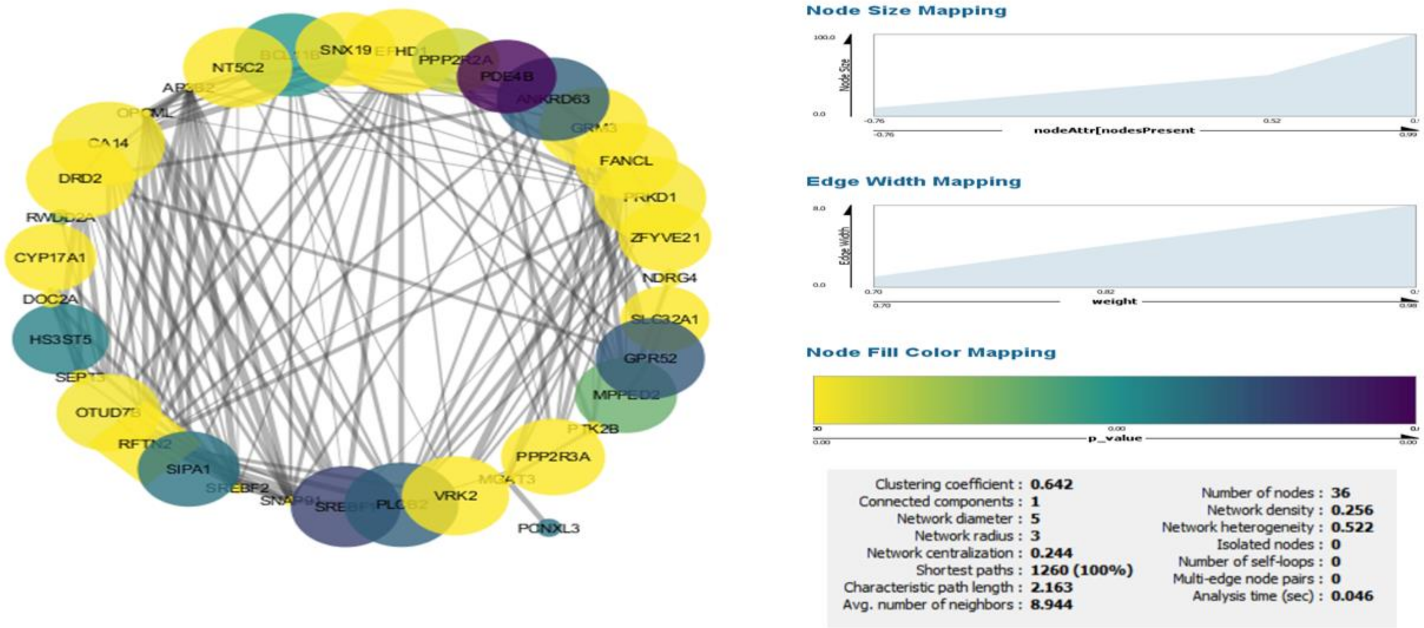
Figure 6.1.20 Yellow Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.643**
Connected components : **2**
Network diameter : **5**
Network radius : **1**
Network centralization : **0.289**
Shortest paths : **10106 (94%)**
Characteristic path length : **2.227**
Avg. number of neighbors : **20.827**

Number of nodes : **104**
Network density : **0.202**
Network heterogeneity : **0.781**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.118**

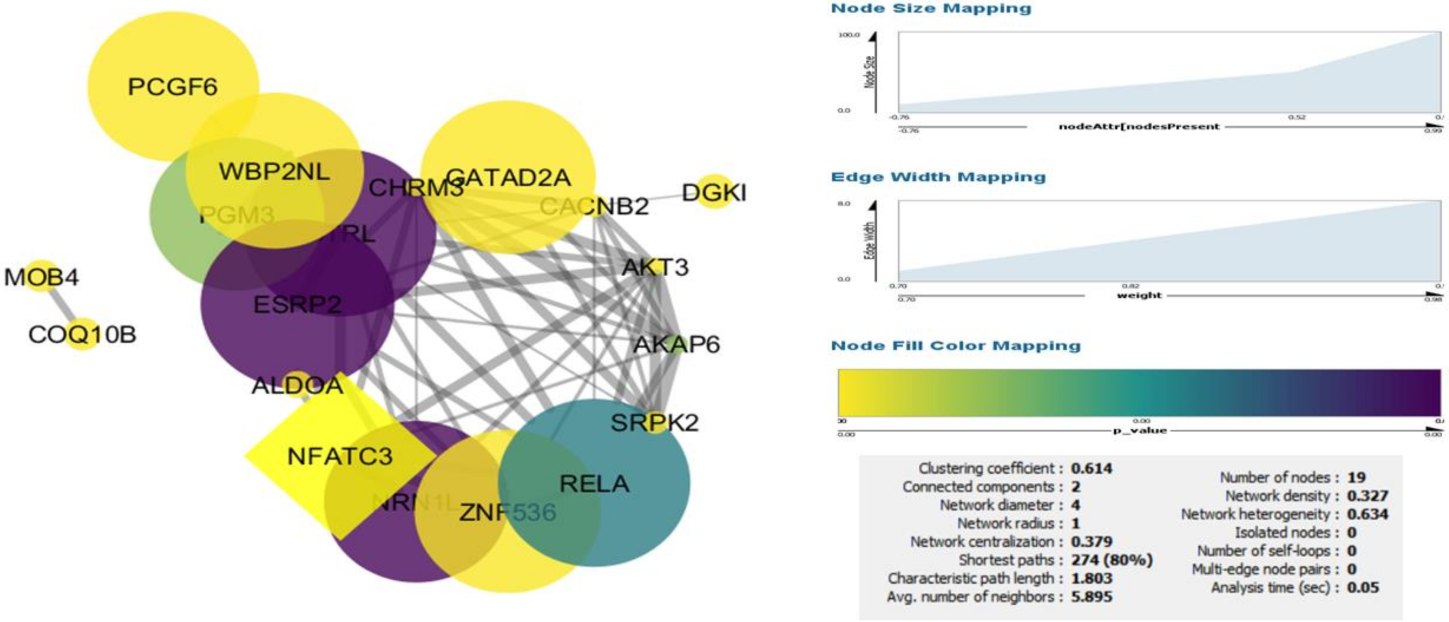
Figure 6.1.21 Black Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.642**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.244**
Shortest paths : **1260 (100%)**
Characteristic path length : **2.163**
Avg. number of neighbors : **8.944**

Number of nodes : **36**
Network density : **0.256**
Network heterogeneity : **0.522**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.046**

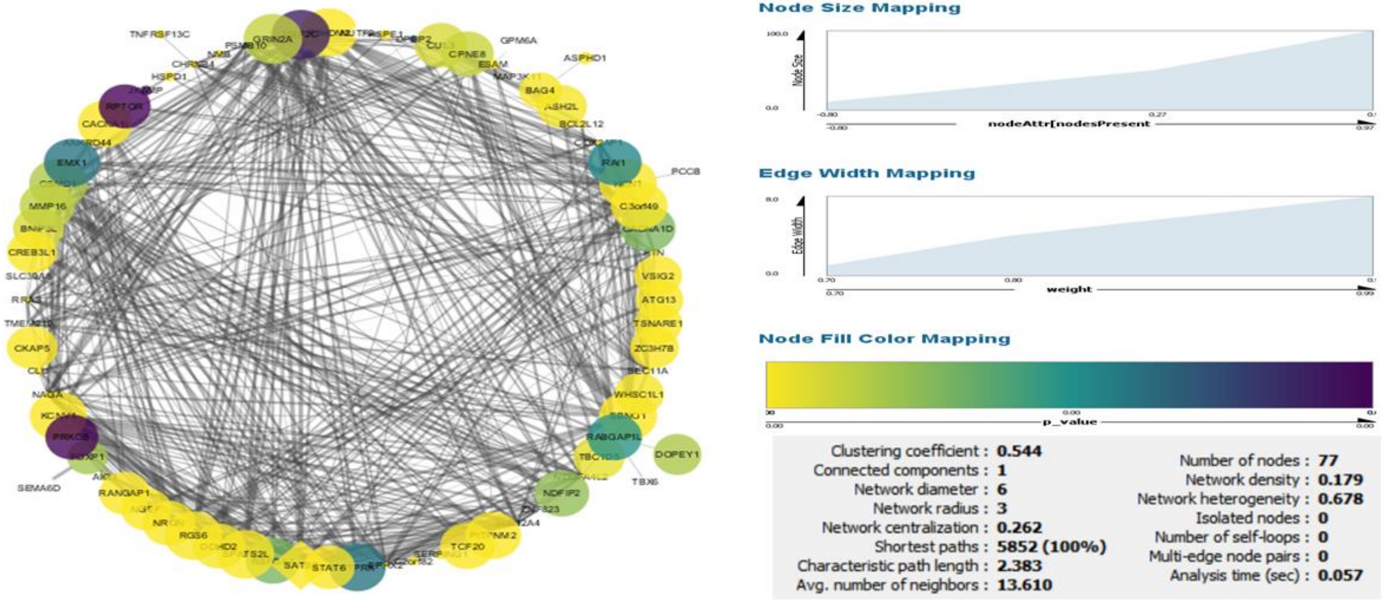
Figure 6.1.22 Brown Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : 0.614
Connected components : 2
Network diameter : 4
Network radius : 1
Network centralization : 0.379
Shortest paths : 274 (80%)
Characteristic path length : 1.803
Avg. number of neighbors : 5.895

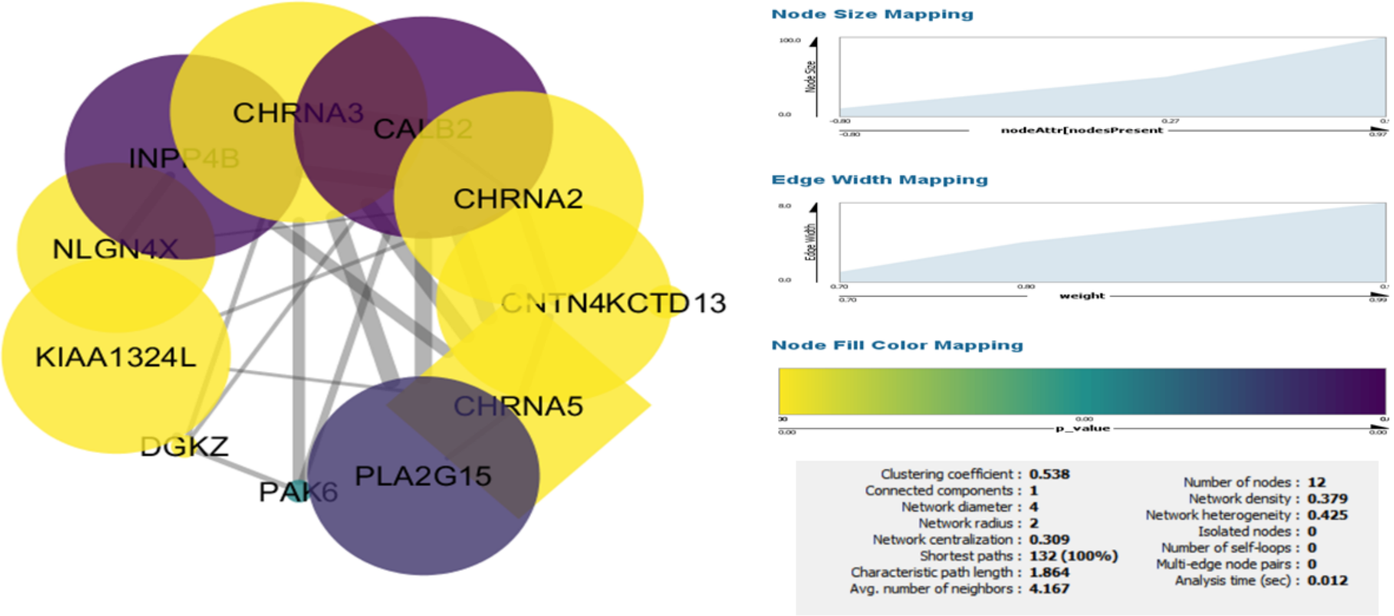
Number of nodes : 19
Network density : 0.327
Network heterogeneity : 0.634
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.05

Figure 6.1.23 Green Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



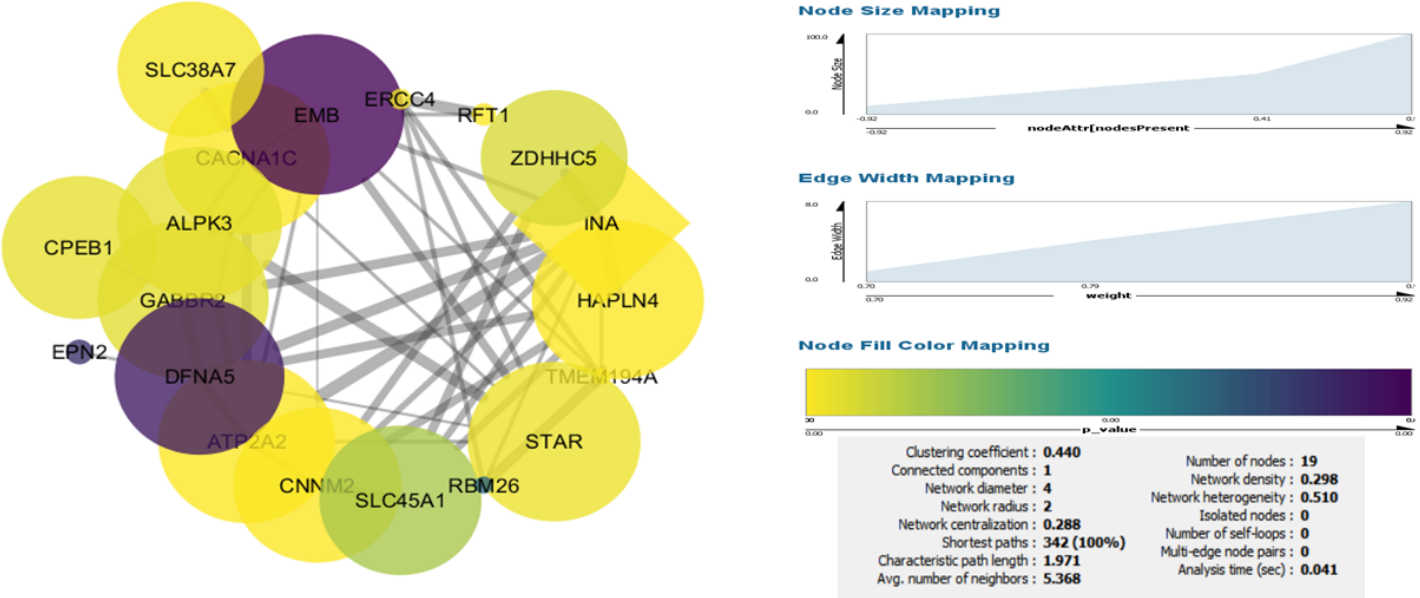
Clustering coefficient : **0.544**
Connected components : **1**
Network diameter : **6**
Network radius : **3**
Network centralization : **0.262**
Shortest paths : **5852 (100%)**
Characteristic path length : **2.383**
Avg. number of neighbors : **13.610**
Number of nodes : **77**
Network density : **0.179**
Network heterogeneity : **0.678**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.057**

Figure 6.1.24 Greenyellow Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



```
Clustering coefficient :  0.538
Connected components :  1
Network diameter :  4
Network radius :  2
Network centralization :  0.309
Shortest paths :  132 (100%)
Characteristic path length :  1.864
Avg. number of neighbors :  4.167

Number of nodes :  12
Network density :  0.379
Network heterogeneity :  0.425
Isolated nodes :  0
Number of self-loops :  0
Multi-edge node pairs :  0
Analysis time (sec) :  0.012
```

Figure 6.1.25 Pink Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : 0.440
Connected components : 1
Network diameter : 4
Network radius : 2
Network centralization : 0.288
Shortest paths : 342 (100%)
Characteristic path length : 1.971
Avg. number of neighbors : 5.368

Number of nodes : 19
Network density : 0.298
Network heterogeneity : 0.510
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.041

Figure 6.1.26 Red Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.

Table 6.1: Schizophrenia-associated gene set from the 145 loci identified by Pardiñas et al. which are available in ABA's BrainSpan resource.

ABCB9	BTBD18	CNOT1	ERCC4	IGSF9B	MPPED2	PITPNM2	RELA	SLC39A8	TNFRSF13C
ACD	BTG1	CNTN4	ESAM	IMMP2L	MSL2	PLA2G15	RERE	SLC45A1	TOM1L2
ACTR5	C10orf32	COQ10B	ESRP2	INA	MSRA	PLCB2	RFT1	SLC7A6	TRANK1
ADAMTSL3	C11orf31	CPEB1	ETF1	INO80E	NAB2	PLCH2	RFTN2	SMG6	TSNARE1
AIG1	C12orf65	CPNE8	F2	INPP4B	NAGA	PLCL1	RGS6	SNAP91	TSNAXIP1
AKAP6	C14orf2	CREB3L1	FAM109B	IREB2	NCAN	PLCL2	RILPL2	SNX19	TSR1
AKT3	C16orf86	CSMD1	FAM57B	JKAMP	NCK1	PLEKHO1	RIMS1	SOX5	TYW5
ALDOA	C16orf92	CTNND1	FANCL	KAT5	NDFIP2	PPAPDC1B	RLTPR	SPATS2L	VPS13C
ALMS1	C2orf47	CTRL	FGFR1	KCNJ13	NDRG4	PPP1R13B	RPS19BP1	SREBF1	VPS45
ALPK3	C2orf82	CUL3	FHIT	KCNK7	NDUFA4L2	PPP2R2A	RPTOR	SREBF2	VRK2
AMBRA1	C3orf49	CYP17A1	FOXP1	KCNV1	NDUFA6	PPP2R3A	RRAS	SRPK2	VSIG2
ANKRD44	C4orf27	DDHD2	FSHB	KCTD13	NEK1	PPP4C	RWDD2A	SRR	WBP2NL
ANKRD63	CA14	DDX28	FXR1	KDM4A	NFATC3	PRKCB	SATB2	STAC3	WDR73
ANP32E	CA8	DFNA5	GABBR2	KIAA1324L	NGEF	PRKD1	SBNO1	STAG1	WHSC1L1
AP3B2	CACNA1C	DGKI	GATAD2A	L3MBTL2	NLGN4X	PRMT1	SCAF1	STAR	XRCC3
ARHGAP1	CACNA1D	DGKZ	GDPD3	LCAT	NMB	PSD3	SDCCAG8	STAT6	YPEL3
ARL6IP4	CACNA1I	DMTF1	GFOD2	LETM2	NRGN	PSMA4	SEC11A	SUGP1	YPEL4
AS3MT	CACNB2	DNAJC19	GIGYF2	LRP1	NRN1L	PSMB10	SEMA6D	TAB1	ZBTB37
ASH2L	CALB2	DOC2A	GPM6A	LRRC48	NT5C2	PSMD6	Sep-03	TAOK2	ZC3H7B
ASPG	CDK2AP1	DOPEY1	GPR135	LSM1	NUTF2	PTK2B	SERPINC1	TBC1D5	ZDHHC5
ASPHD1	CENPM	DPEP2	GPR52	MAD1L1	NXPH4	PTN	SERPING1	TBX6	ZEB2
ATG13	CENPT	DPEP3	GRIN2A	MAP3K11	OGFOD2	PTPRF	SETD6	TCF20	ZFYVE21
ATP2A2	CHRM3	DPYD	GRM3	MAPK3	OPCML	PTPRK	SETD8	TCF4	ZNF408
ATPAF2	CHRNA2	DRD2	HAPLN4	MARS2	OTOL1	R3HDM2	SEZ6L2	TDRD9	ZNF536
ATXN7	CHRNA3	EDC4	HARBI1	MAU2	OTUD7B	RABGAP1L	SF3B1	THAP11	ZNF592
B3GAT1	CHRNA5	EFHD1	HCN1	MED19	PAK6	RAI1	SFMBT1	THOC7	ZNF804A
BAG4	CHRNB4	EHBP1L1	HIRIP3	MEF2C	PARD6A	RALGAPA2	SFXN5	TLE3	ZNF823

BANK1	CKAP5	EMB	HS3ST5	MGAT3	PCCB	RANBP10	SHMT2	TM6SF2	ZSCAN2
BCL11B	CLCN3	EMX1	HSPA9	MMP16	PCGF6	RANGAP1	SIPA1	TMEM194A	
BCL2L12	CLP1	EP300	HSPD1	MOB4	PCNXL3	RBFOX1	SLC12A4	TMEM219	
BNIP3L	CLU	EPHX2	HSPE1	MPHOSPH9	PDE4B	RBM26	SLC32A1	TMX2	
BOLL	CNNM2	EPN2	HYDIN	MPP6	PGM3	RC3H1	SLC38A7	TMX2-CTNND1	

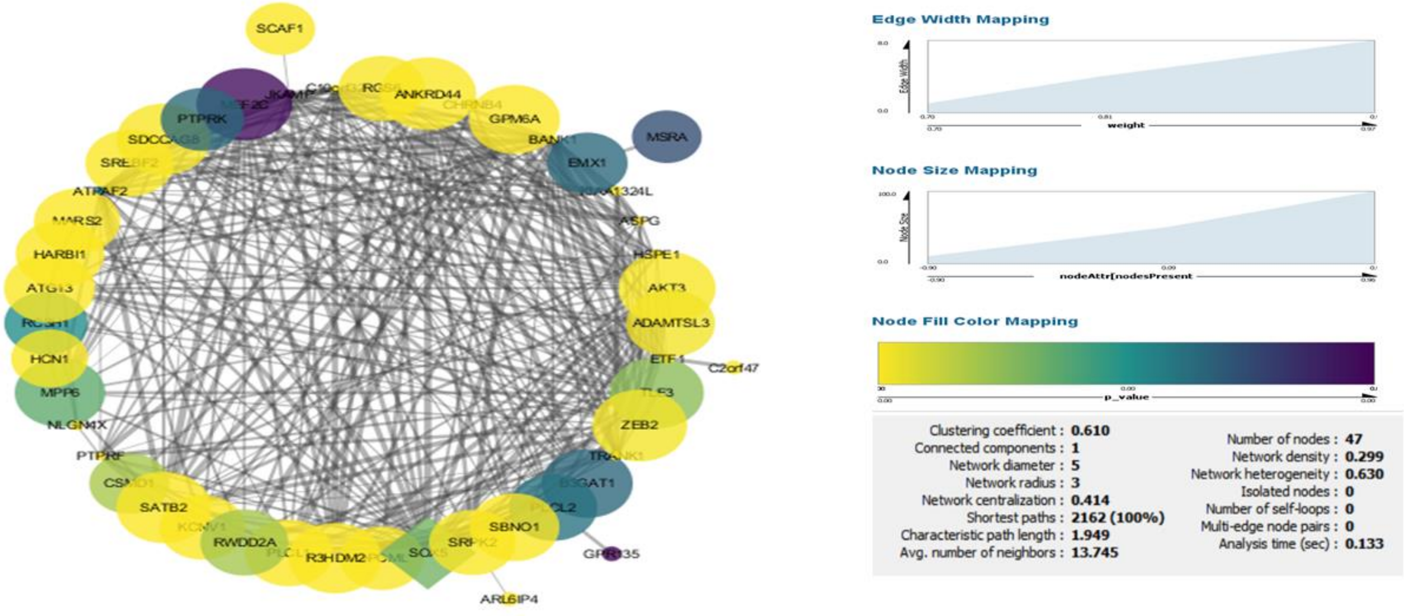
Table 6.2: Cluster assignments for each schizophrenia-associated gene over the five stages using the kmeans function available in R.

Gene Name	Developmental Stage One	Developmental Stage Two	Developmental Stage Three	Developmental Stage Four	Developmental Stage Five
ABCB1	3	2	2	2	2
ABCB9	1	2	1	2	1
ABCD2	2	1	2	2	2
ACO2	3	2	2	2	2
ACP2	3	2	1	2	1
ACTR1A	2	1	1	2	1
ACTR5	3	2	1	2	2
ADAMTSL3	2	2	2	2	2
ADAMTSL4	3	2	1	2	2
AIG1	3	2	1	2	2
AKAP6	3	1	2	1	1
AKT3	2	1	2	1	1
ALAS1	2	1	1	2	2
ALDOA	3	1	2	2	1
ALMS1	3	2	1	2	2
ANAPC7	3	2	1	2	2
ANKRD44	2	2	1	2	2
ANKRD45	2	1	1	2	2
ANKRD63	3	2	1	2	2
APOPT1	3	1	2	1	2
ARL5B	2	1	2	1	1
ARTN	3	2	1	2	2
AS3MT	3	2	1	2	2
ATF4	1	1	1	1	2
ATP13A1	1	2	1	2	1
ATPAF2	3	2	1	2	2
ATXN7	2	2	1	2	1

B9D1	3	1	2	2	2
BAG4	2	1	2	1	1
BANK1	3	2	1	2	2
BCL11B	2	2	1	2	2
BCL2L12	3	2	1	2	2
BNIP3L	2	1	2	1	1
BRD8	3	2	1	2	1
BTBD18	1	2	1	2	1
C2orf47	3	1	2	1	2
C2orf82	1	1	2	1	2
CA8	3	2	1	2	2
CACNA1C	1	2	1	2	1
CACNA1D	2	2	1	1	1
CACNA1I	1	1	2	1	1
CACNB2	1	1	2	1	1
CALB2	3	2	1	2	2
CENPM	3	2	1	2	2
CENPT	1	2	1	2	2
CEP170	2	1	1	2	1
CHRNA2	3	2	1	2	2
CHRNA3	3	2	1	2	2
CLCN3	3	2	2	2	1
CLDN23	2	2	2	2	2
CNOT1	3	2	1	2	1
CNTN4	3	2	1	2	2
CSMD1	2	1	2	1	1
CUL3	2	1	2	1	1
DFNA5	2	2	2	1	1
DGKI	1	2	1	2	1
DNAJC19	3	1	2	1	2
DOPEY1	2	2	1	2	1

DPYD	1	1	2	2	2
DRD2	3	2	1	2	2
EMB	3	2	1	2	1
EMX1	1	1	2	1	1
ESAM	3	2	2	2	2
FANCL	1	1	1	2	2
FHIT	3	2	1	2	2
FOXP1	3	1	2	1	1
FTSJ2	3	1	2	1	1
GABBR2	1	1	1	1	1
GPM6A	2	1	2	1	2
GRIA1	1	2	1	2	2
GRIN2A	1	1	2	1	1
GRM3	1	2	2	2	2
HCN1	1	1	2	1	1
IGSF9B	3	2	1	2	1
IL20RB	3	2	1	2	2
IMMP2L	3	1	2	2	2
INHBC	3	2	1	2	2
INPP4B	3	2	1	2	2
ME1	3	2	2	1	2
MGAT3	1	2	1	2	1
MMP16	2	2	2	1	1
NDFIP2	2	1	2	1	1
NLGN4X	3	2	1	2	2
OPCML	2	1	1	1	1
PDE4B	3	2	2	2	2
PSD3	1	1	1	2	1
PTPRK	1	1	2	1	1
RBFOX1	2	1	2	1	1
RERE	2	2	1	2	1

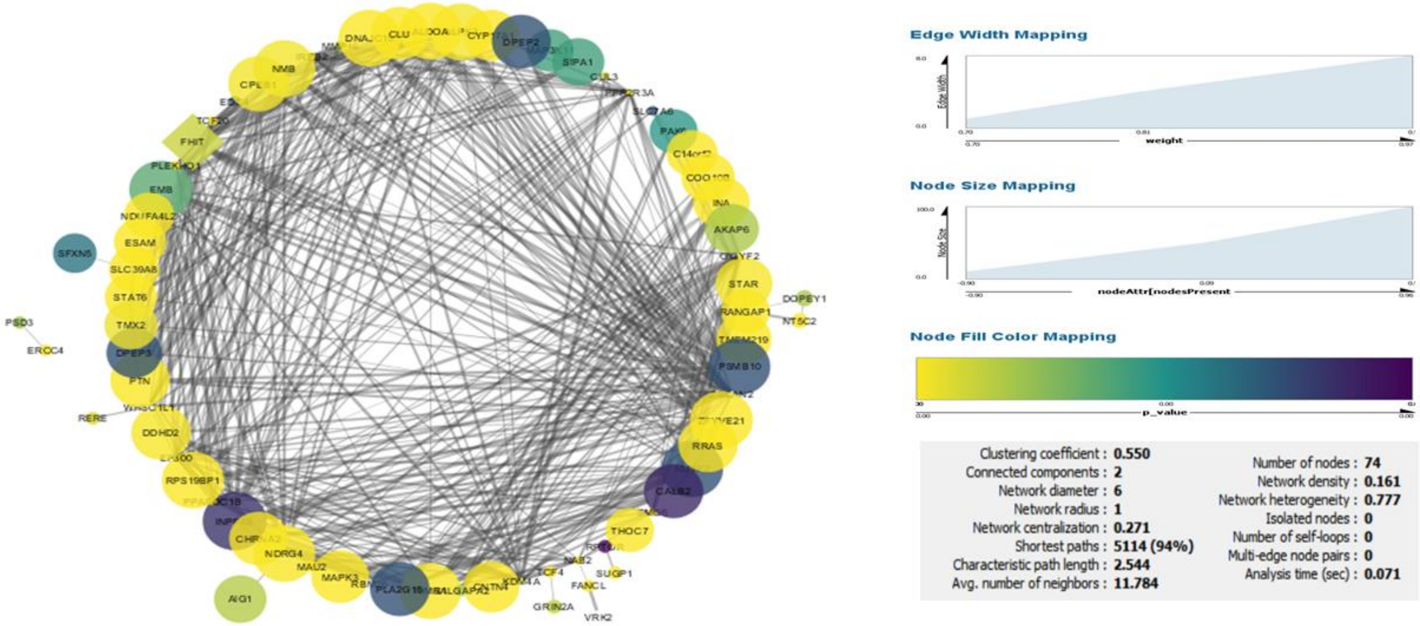
RGS6	1	1	2	1	1
RIMS1	3	2	1	2	2
RPTOR	1	2	2	1	1
SATB2	2	1	2	1	1
SEMA6D	3	2	2	2	1
SNX19	2	2	2	2	2
SPATS2L	3	1	2	1	1
TBC1D5	3	2	2	2	1
TCF4	1	2	2	2	1
TRANK1	3	2	1	2	1
TSNARE1	1	2	2	1	1
ZEB2	2	2	1	2	2
ZNF440	2	2	1	2	2
ZNF536	3	2	1	2	2
ZNF804A	3	2	1	2	1



Clustering coefficient : **0.610**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.414**
Shortest paths : **2162 (100%)**
Characteristic path length : **1.949**
Avg. number of neighbors : **13.745**

Number of nodes : **47**
Network density : **0.299**
Network heterogeneity : **0.630**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.133**

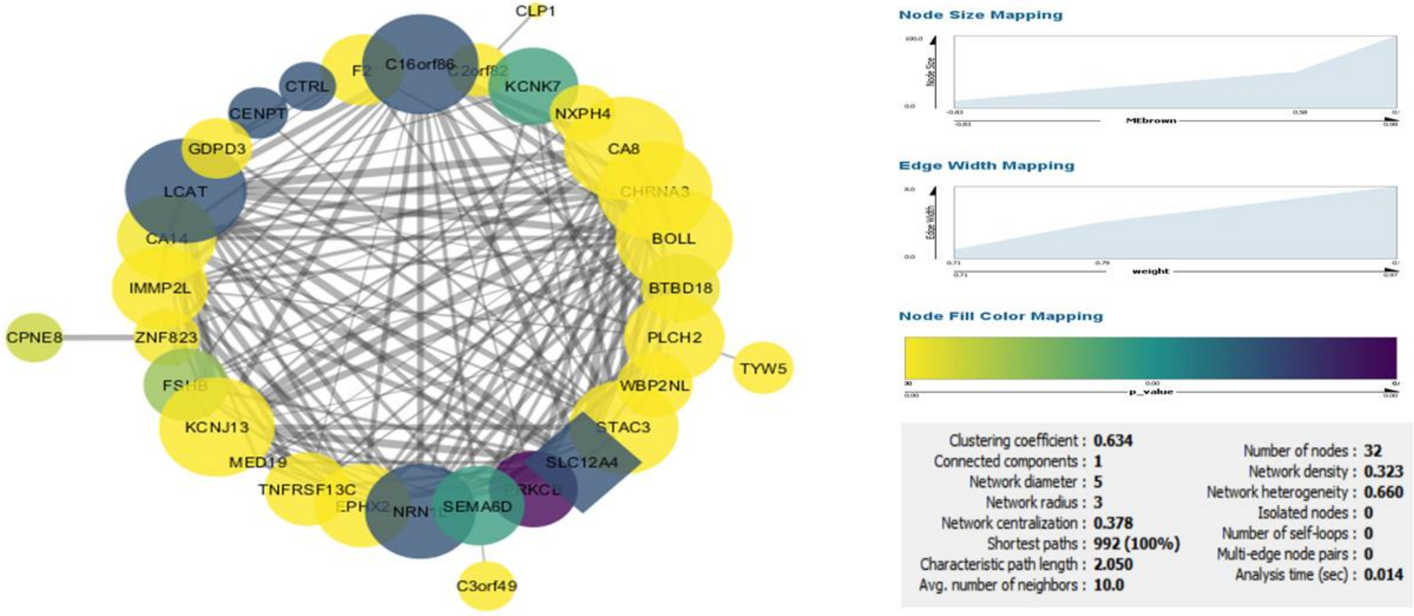
Figure 6.1.1 Black Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.550**
Connected components : **2**
Network diameter : **6**
Network radius : **1**
Network centralization : **0.271**
Shortest paths : **5114 (94%)**
Characteristic path length : **2.544**
Avg. number of neighbors : **11.784**

Number of nodes : **74**
Network density : **0.161**
Network heterogeneity : **0.777**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.071**

Figure 6.1.2 Blue Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.634**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.378**
Shortest paths : **992 (100%)**
Characteristic path length : **2.050**
Avg. number of neighbors : **10.0**

Number of nodes : **32**
Network density : **0.323**
Network heterogeneity : **0.660**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.014**

Figure 6.1.3 Brown Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.

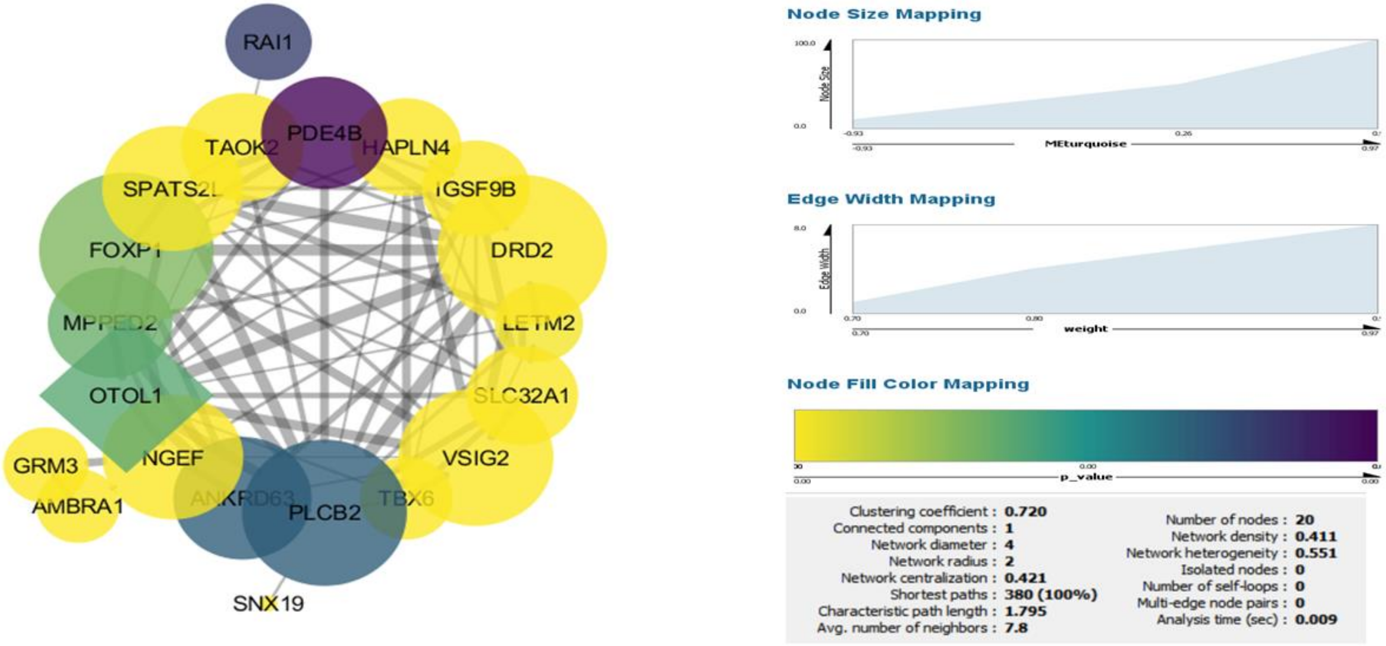
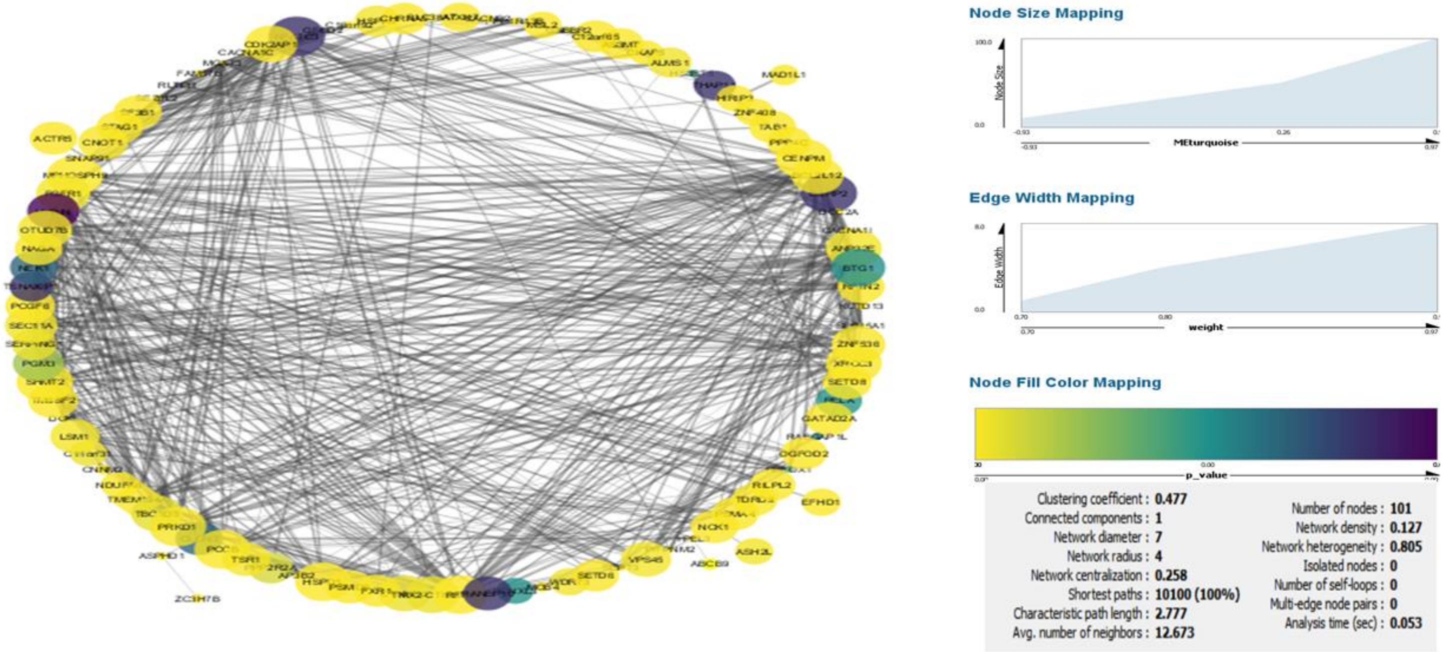


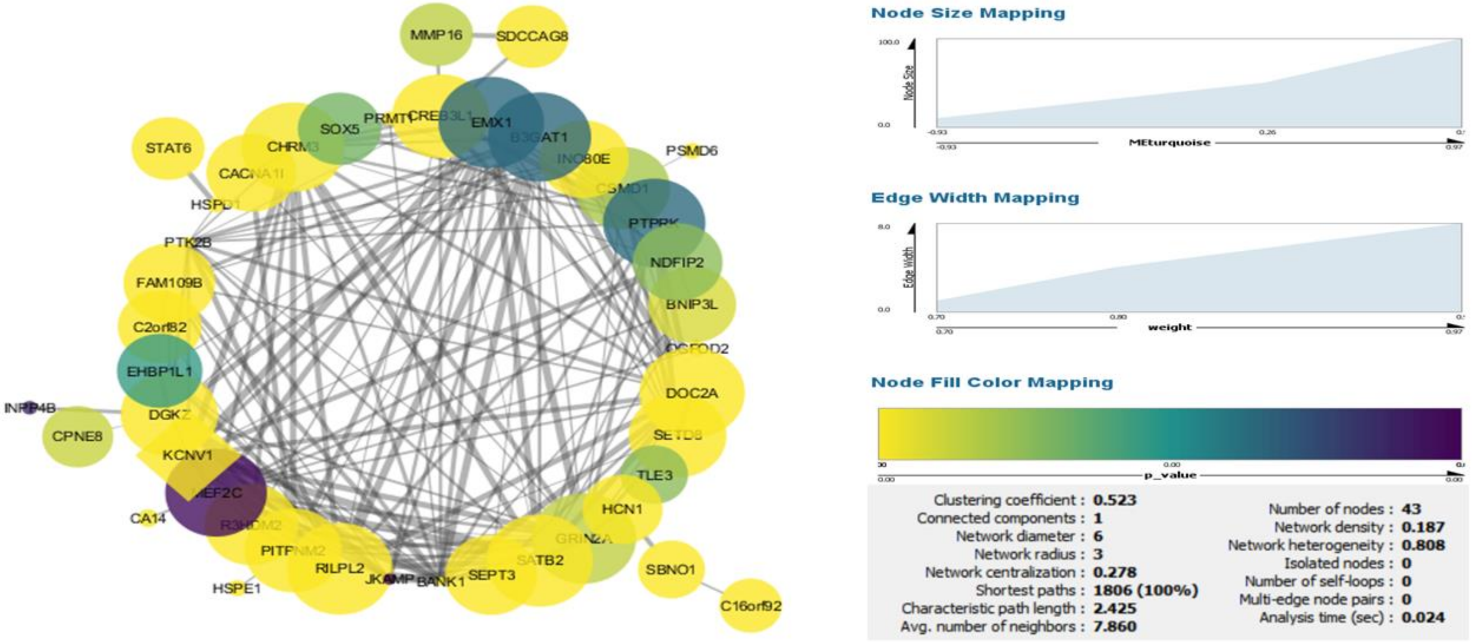
Figure 6.1.4 Pink Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.477**
Connected components : **1**
Network diameter : **7**
Network radius : **4**
Network centralization : **0.258**
Shortest paths : **10100 (100%)**
Characteristic path length : **2.777**
Avg. number of neighbors : **12.673**

Number of nodes : **101**
Network density : **0.127**
Network heterogeneity : **0.805**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.053**

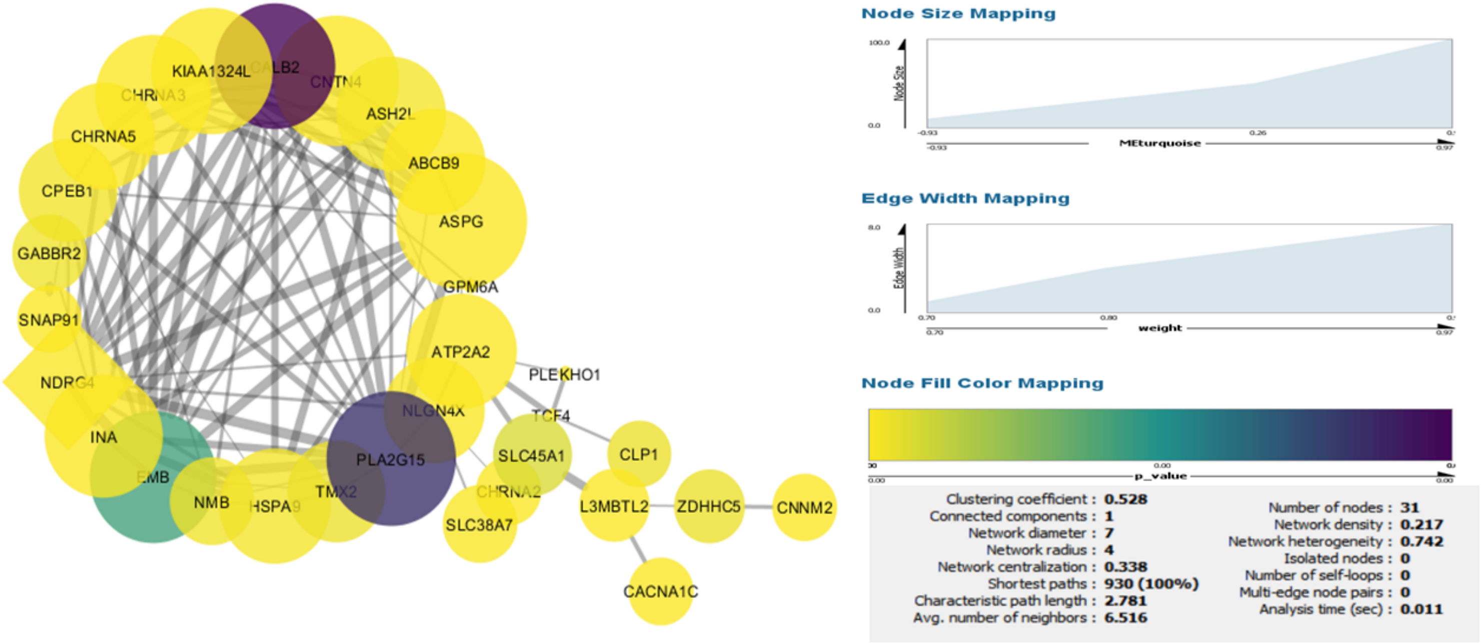
Figure 6.1.5 Turquoise Module for developmental Stage One where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.523**
Connected components : **1**
Network diameter : **6**
Network radius : **3**
Network centralization : **0.278**
Shortest paths : **1806 (100%)**
Characteristic path length : **2.425**
Avg. number of neighbors : **7.860**

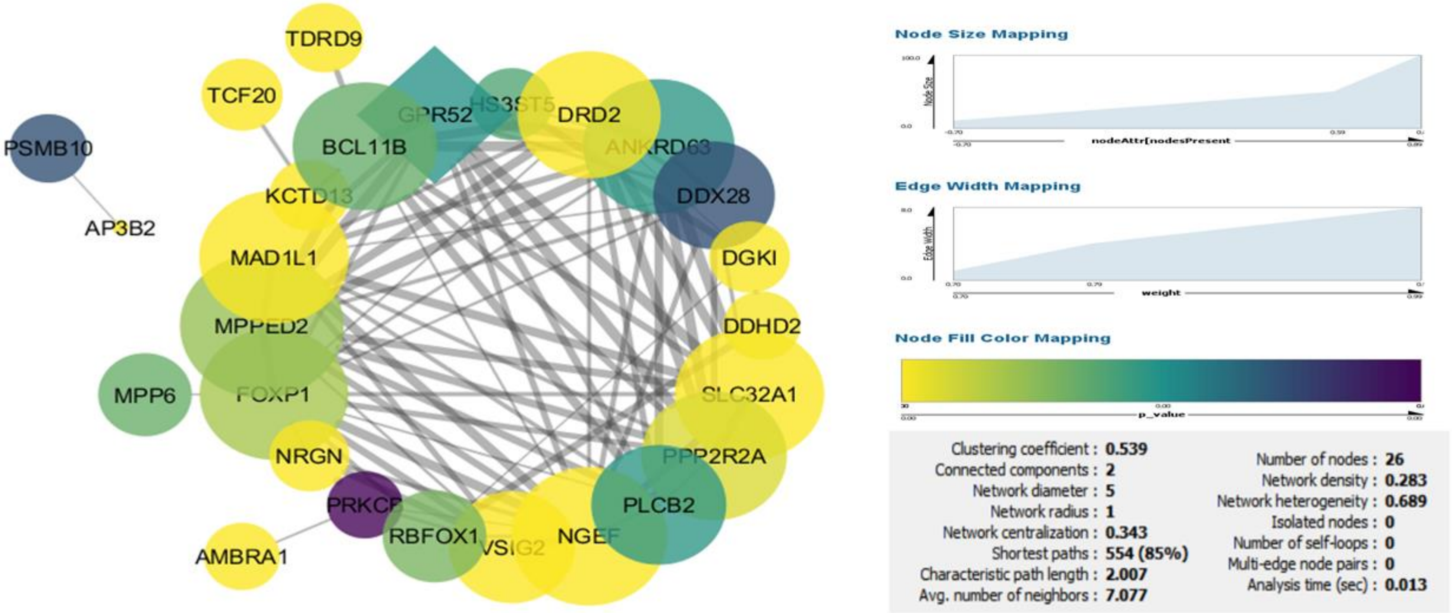
Number of nodes : **43**
Network density : **0.187**
Network heterogeneity : **0.808**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.024**

Figure 6.1.6 Blue Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



```
Clustering coefficient : 0.528
Connected components : 1
Network diameter : 7
Network radius : 4
Network centralization : 0.338
Shortest paths : 930 (100%)
Characteristic path length : 2.781
Avg. number of neighbors : 6.516

Number of nodes : 31
Network density : 0.217
Network heterogeneity : 0.742
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.011
```

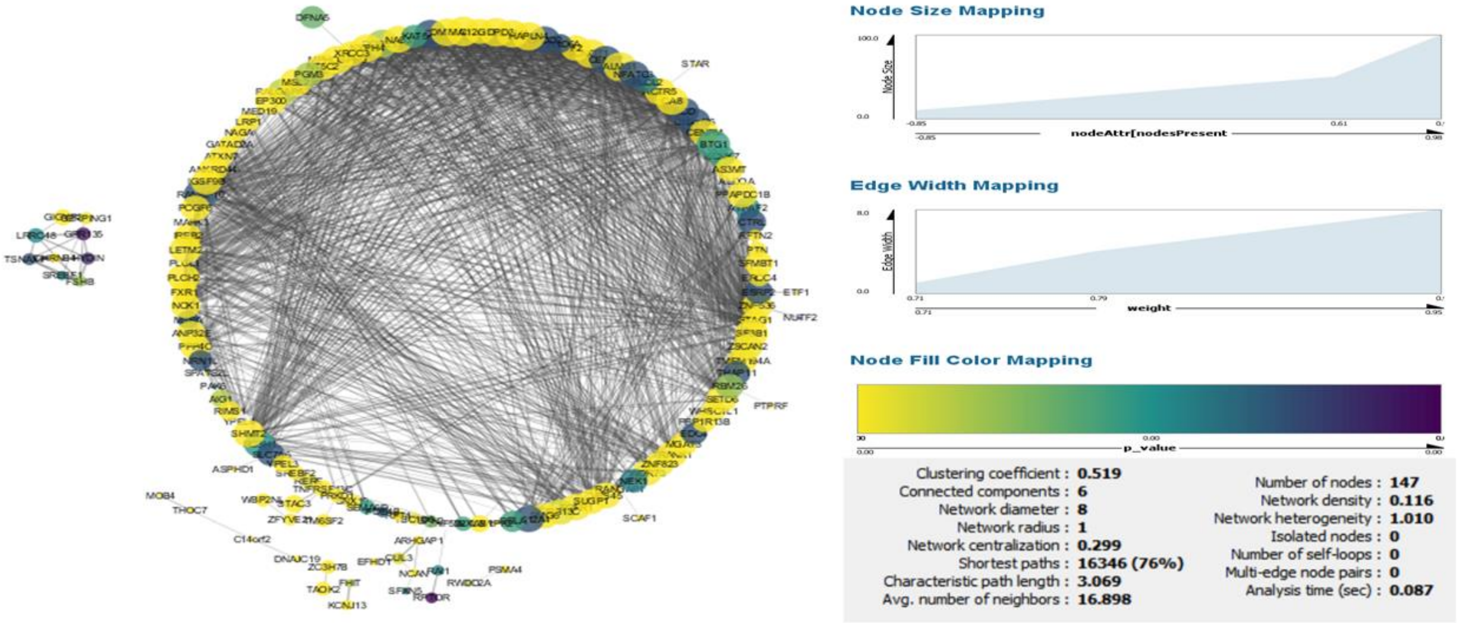
Figure 6.1.7 Brown Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.539**
Connected components : **2**
Network diameter : **5**
Network radius : **1**
Network centralization : **0.343**
Shortest paths : **554 (85%)**
Characteristic path length : **2.007**
Avg. number of neighbors : **7.077**

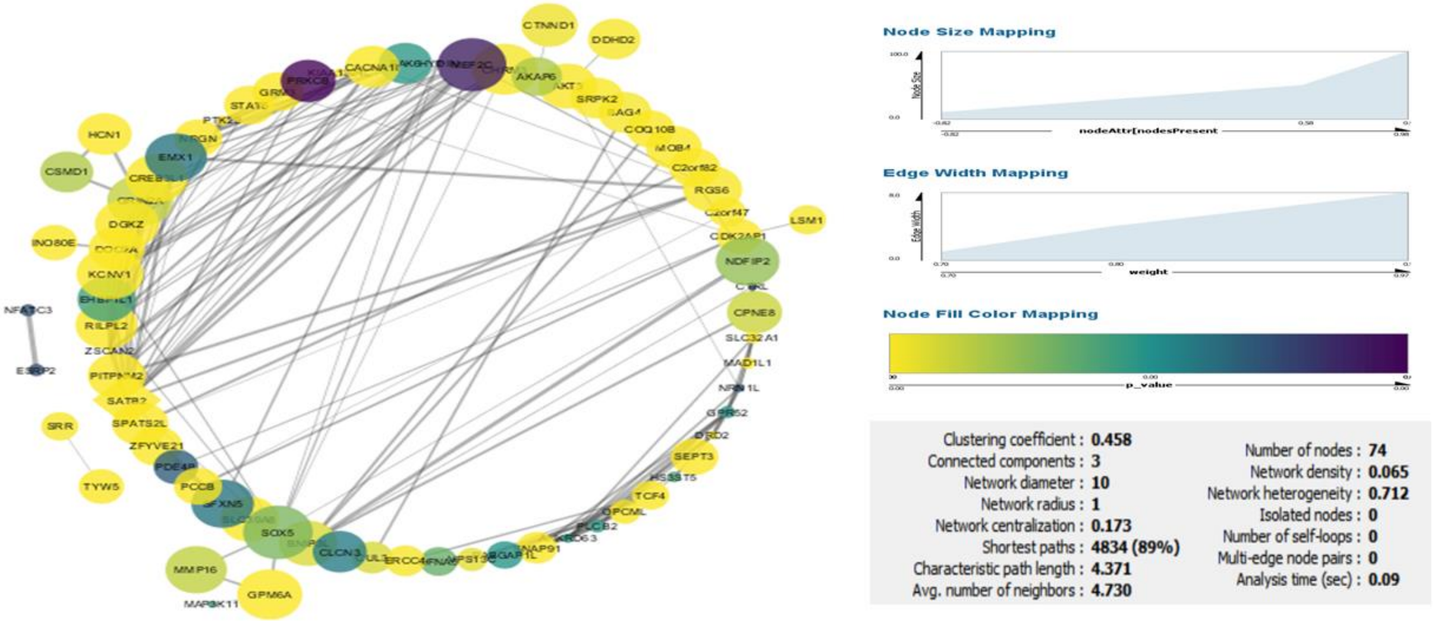
Number of nodes : **26**
Network density : **0.283**
Network heterogeneity : **0.689**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.013**

Figure 6.1.8 Green Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



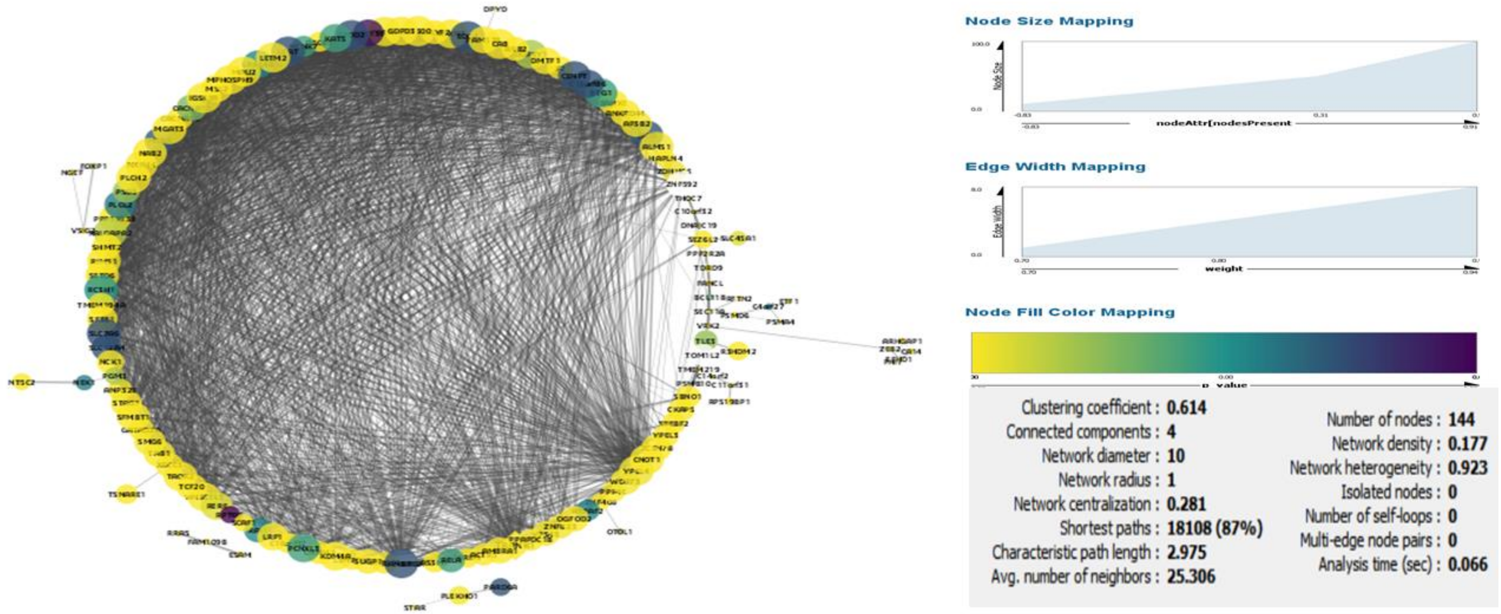
```
Clustering coefficient : 0.519
Connected components : 6
Network diameter : 8
Network radius : 1
Network centralization : 0.299
Shortest paths : 16346 (76%)
Characteristic path length : 3.069
Avg. number of neighbors : 16.898

Number of nodes : 147
Network density : 0.116
Network heterogeneity : 1.010
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.087
```

Figure 6.1.9 Turquoise Module for developmental Stage Two where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.458**
Connected components : **3**
Network diameter : **10**
Network radius : **1**
Network centralization : **0.173**
Shortest paths : **4834 (89%)**
Characteristic path length : **4.371**
Avg. number of neighbors : **4.730**
Number of nodes : **74**
Network density : **0.065**
Network heterogeneity : **0.712**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.09**

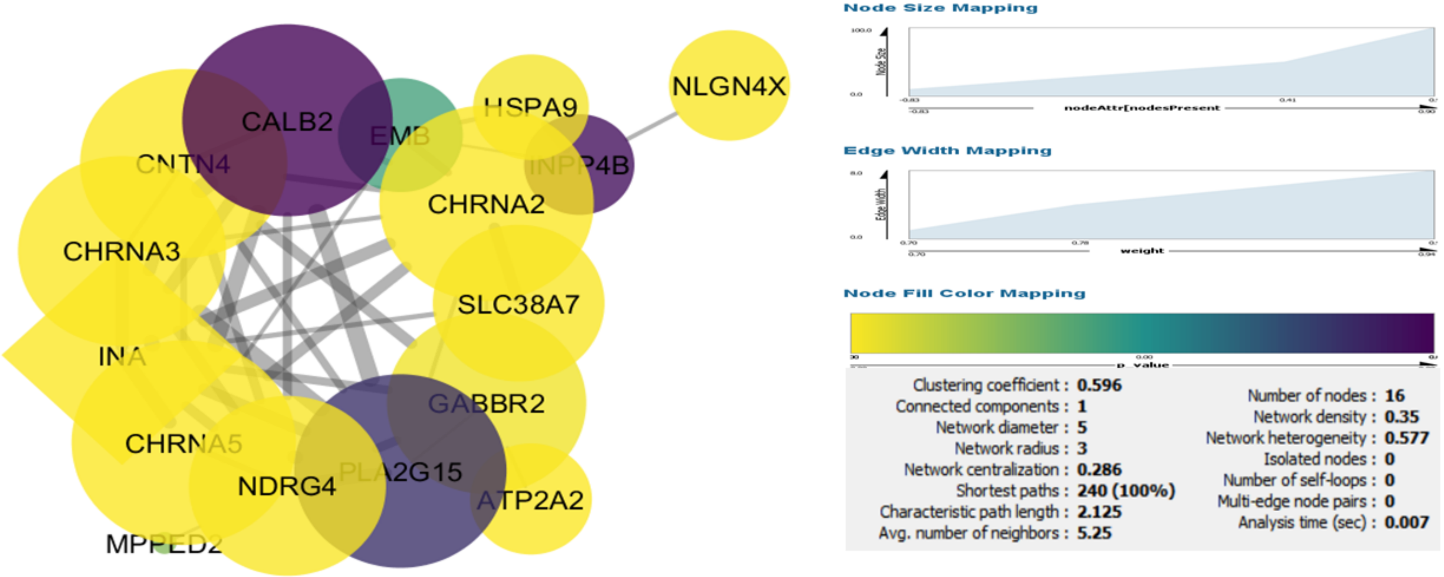
Figure 3.6.10 Brown Module for developmental Stage Three where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.614**
Connected components : **4**
Network diameter : **10**
Network radius : **1**
Network centralization : **0.281**
Shortest paths : **18108 (87%)**
Characteristic path length : **2.975**
Avg. number of neighbors : **25.306**

Number of nodes : **144**
Network density : **0.177**
Network heterogeneity : **0.923**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.066**

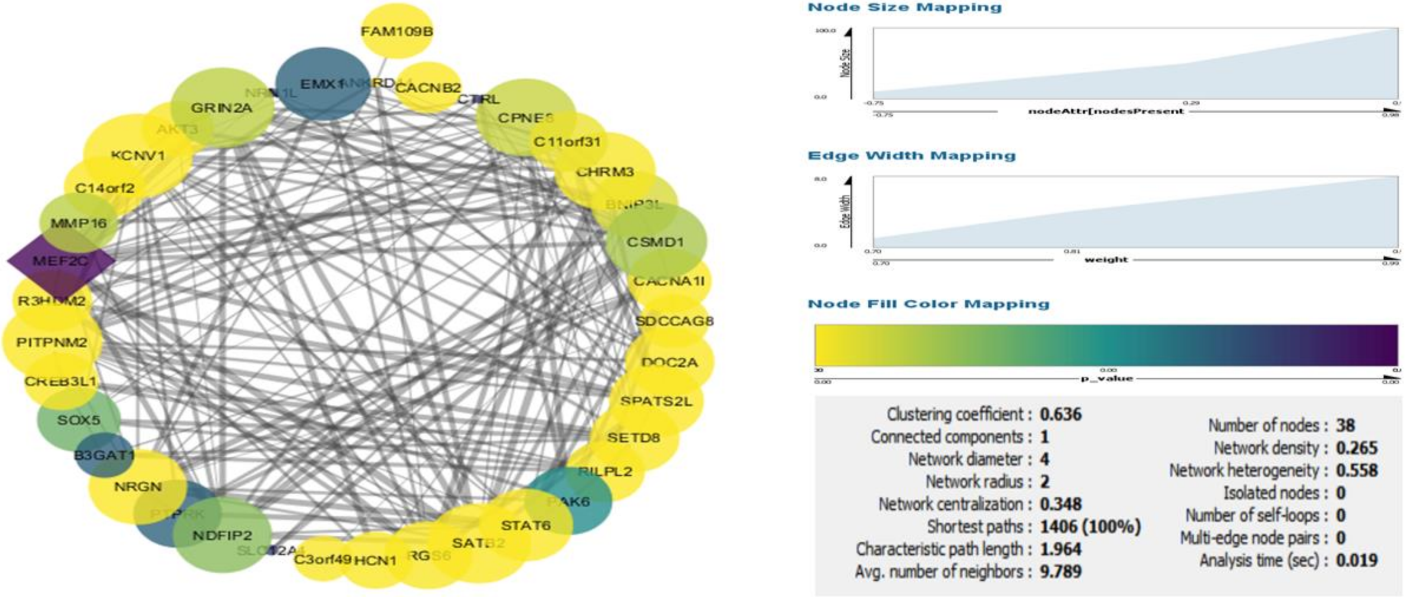
Figure 6.1.11 Turquoise Module for developmental Stage Three where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.596**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.286**
Shortest paths : **240 (100%)**
Characteristic path length : **2.125**
Avg. number of neighbors : **5.25**

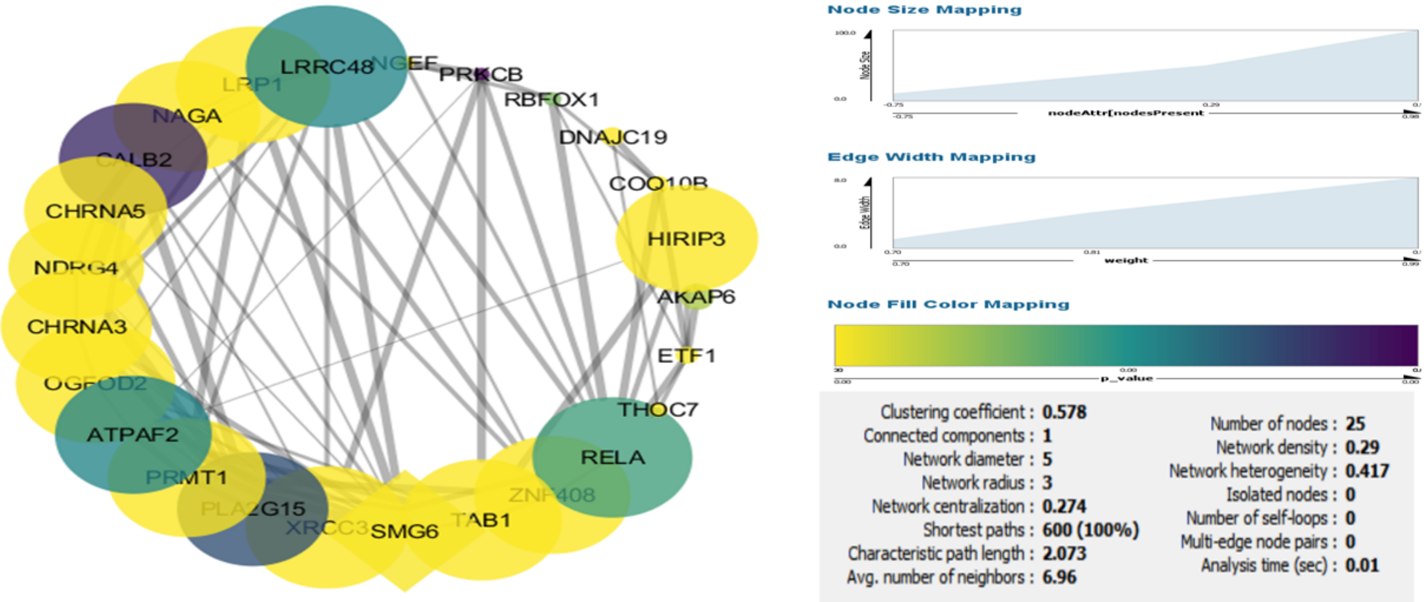
Number of nodes : **16**
Network density : **0.35**
Network heterogeneity : **0.577**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.007**

Figure 6.1.12 Yellow Module for developmental Stage Three where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient :	**0.636**
Connected components :	**1**
Network diameter :	**4**
Network radius :	**2**
Network centralization :	**0.348**
Shortest paths :	**1406 (100%)**
Characteristic path length :	**1.964**
Avg. number of neighbors :	**9.789**
Number of nodes :	**38**
Network density :	**0.265**
Network heterogeneity :	**0.558**
Isolated nodes :	**0**
Number of self-loops :	**0**
Multi-edge node pairs :	**0**
Analysis time (sec) :	**0.019**

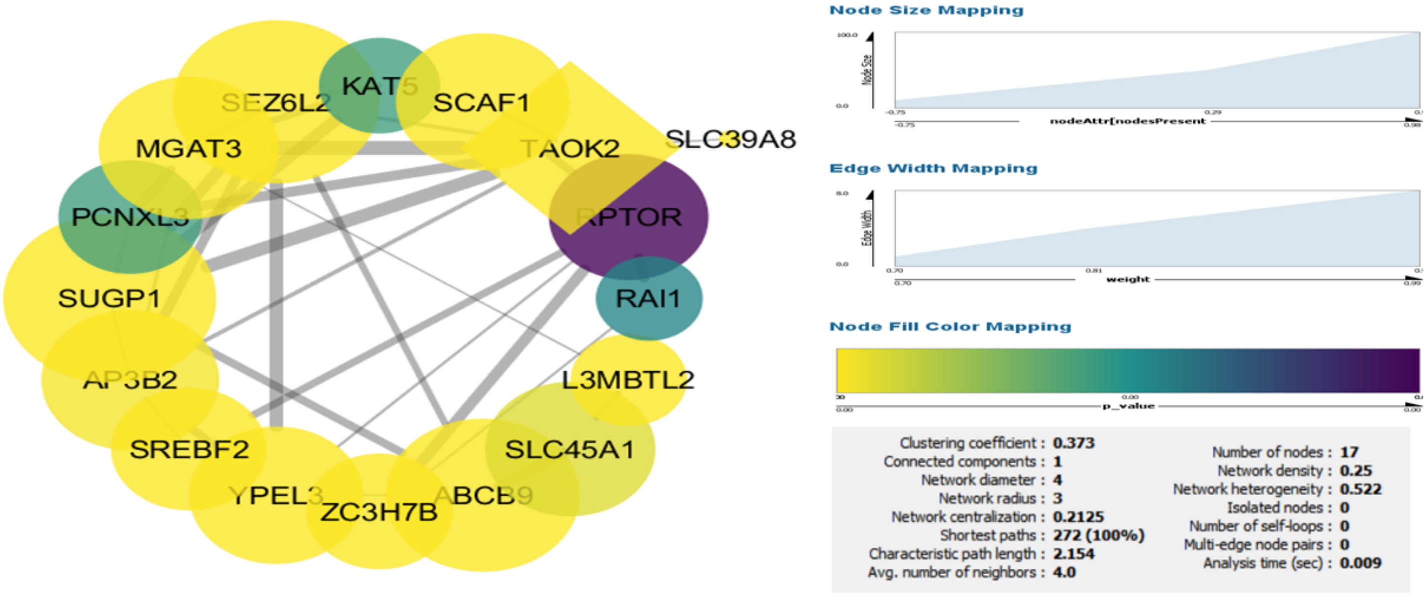
Figure 6.1.13 Blue Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.578**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.274**
Shortest paths : **600 (100%)**
Characteristic path length : **2.073**
Avg. number of neighbors : **6.96**

Number of nodes : **25**
Network density : **0.29**
Network heterogeneity : **0.417**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.01**

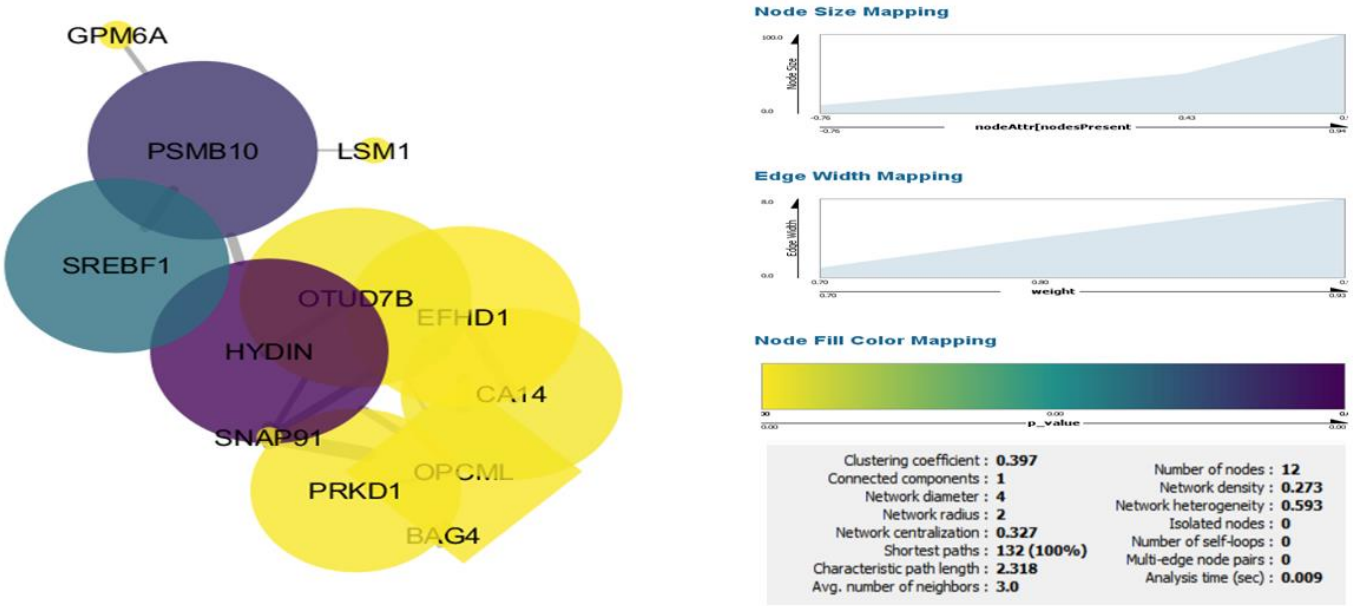
Figure 6.1.14 Brown Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : 0.373
Connected components : 1
Network diameter : 4
Network radius : 3
Network centralization : 0.2125
Shortest paths : 272 (100%)
Characteristic path length : 2.154
Avg. number of neighbors : 4.0

Number of nodes : 17
Network density : 0.25
Network heterogeneity : 0.522
Isolated nodes : 0
Number of self-loops : 0
Multi-edge node pairs : 0
Analysis time (sec) : 0.009

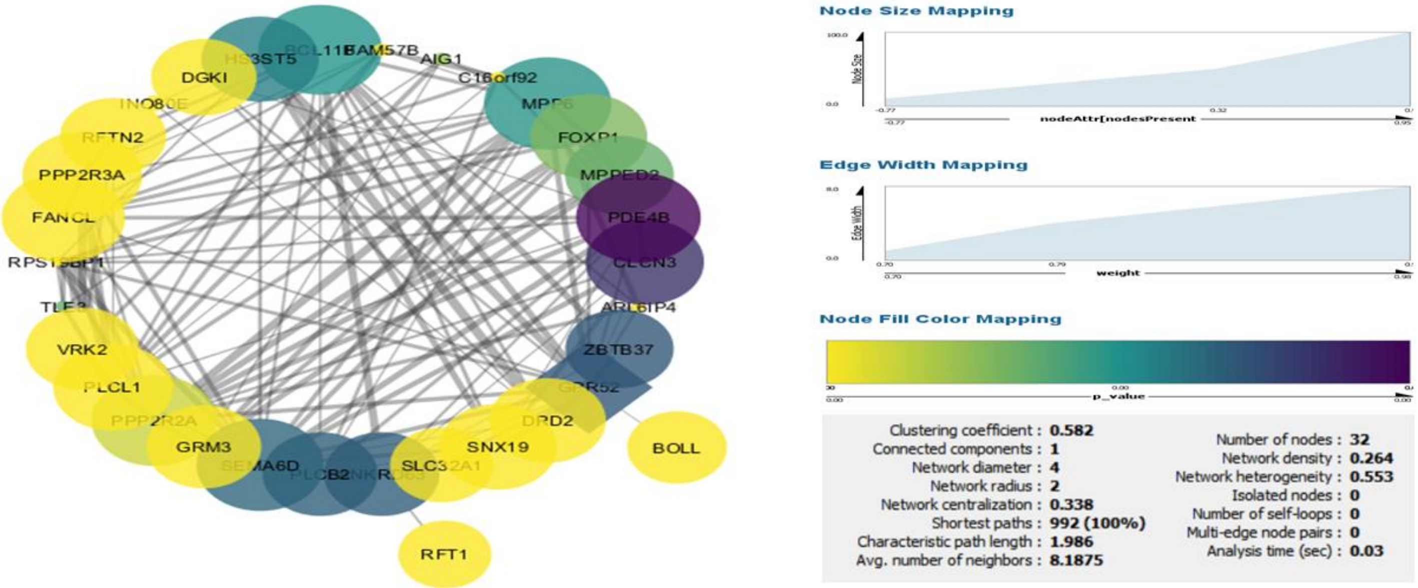
Figure 6.1.15 Green Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.397**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.327**
Shortest paths : **132 (100%)**
Characteristic path length : **2.318**
Avg. number of neighbors : **3.0**

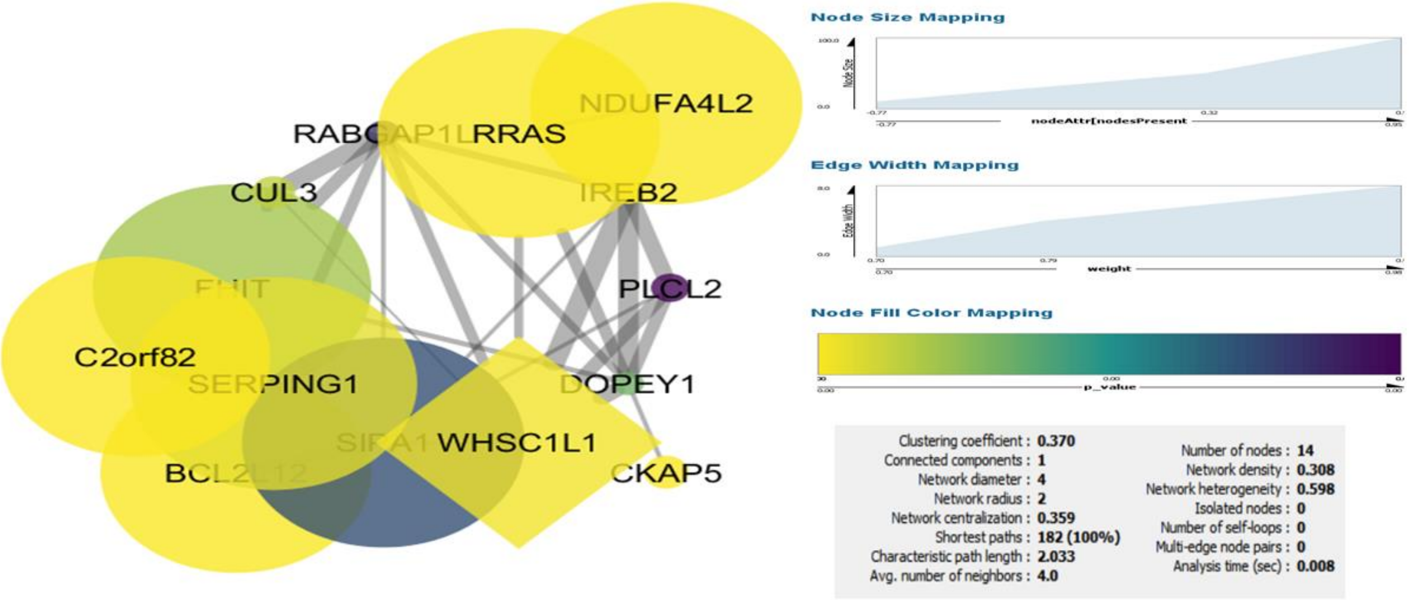
Number of nodes : **12**
Network density : **0.273**
Network heterogeneity : **0.593**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.009**

Figure 6.1.16 Magenta Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.582**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.338**
Shortest paths : **992 (100%)**
Characteristic path length : **1.986**
Avg. number of neighbors : **8.1875**
Number of nodes : **32**
Network density : **0.264**
Network heterogeneity : **0.553**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.03**

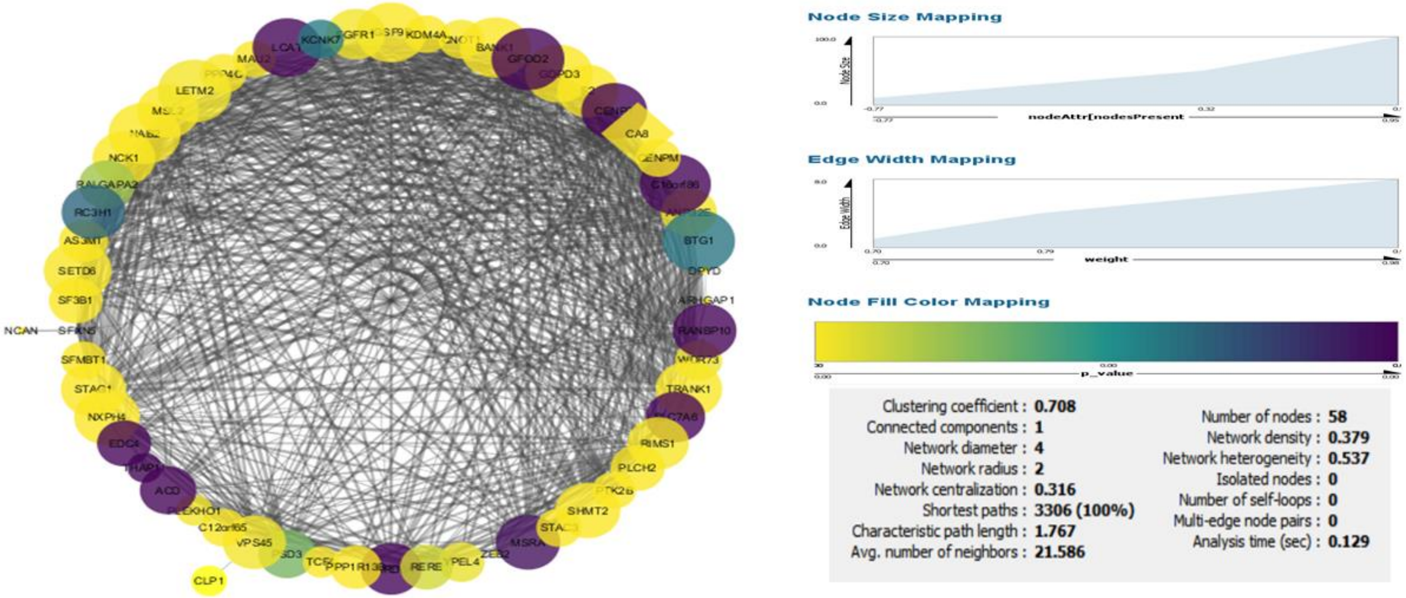
Figure 6.1.17 Purple Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.370**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.359**
Shortest paths : **182 (100%)**
Characteristic path length : **2.033**
Avg. number of neighbors : **4.0**

Number of nodes : **14**
Network density : **0.308**
Network heterogeneity : **0.598**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.008**

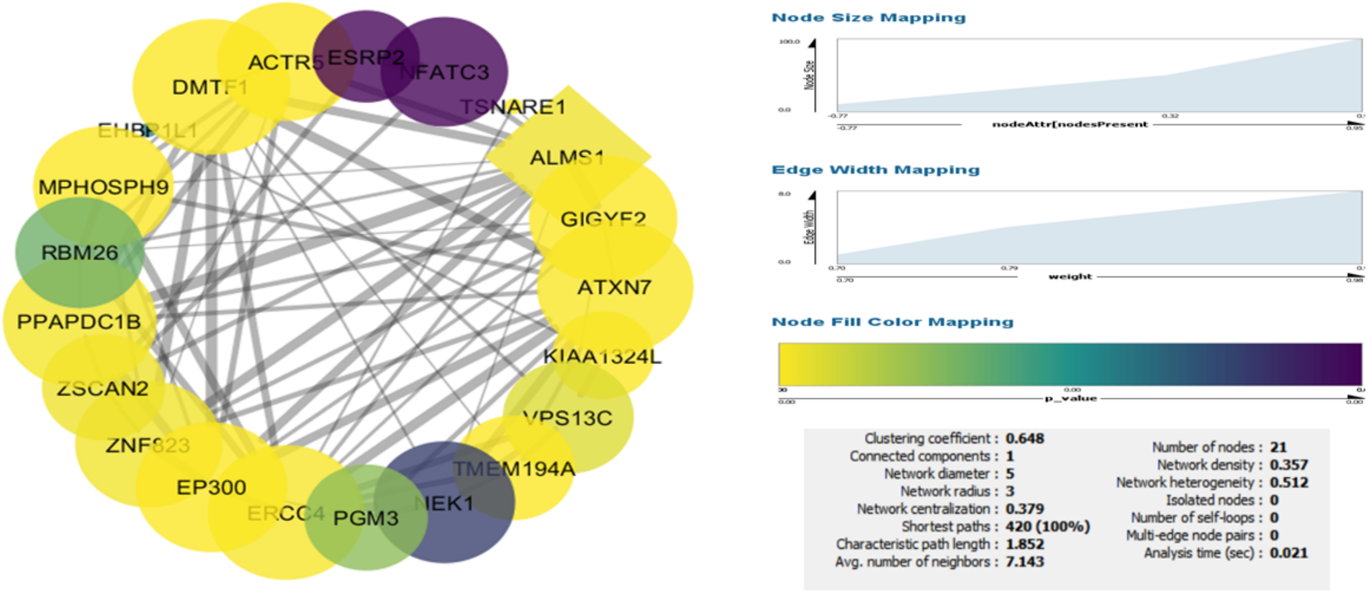
Figure 6.1.18 Red Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.708**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.316**
Shortest paths : **3306 (100%)**
Characteristic path length : **1.767**
Avg. number of neighbors : **21.586**

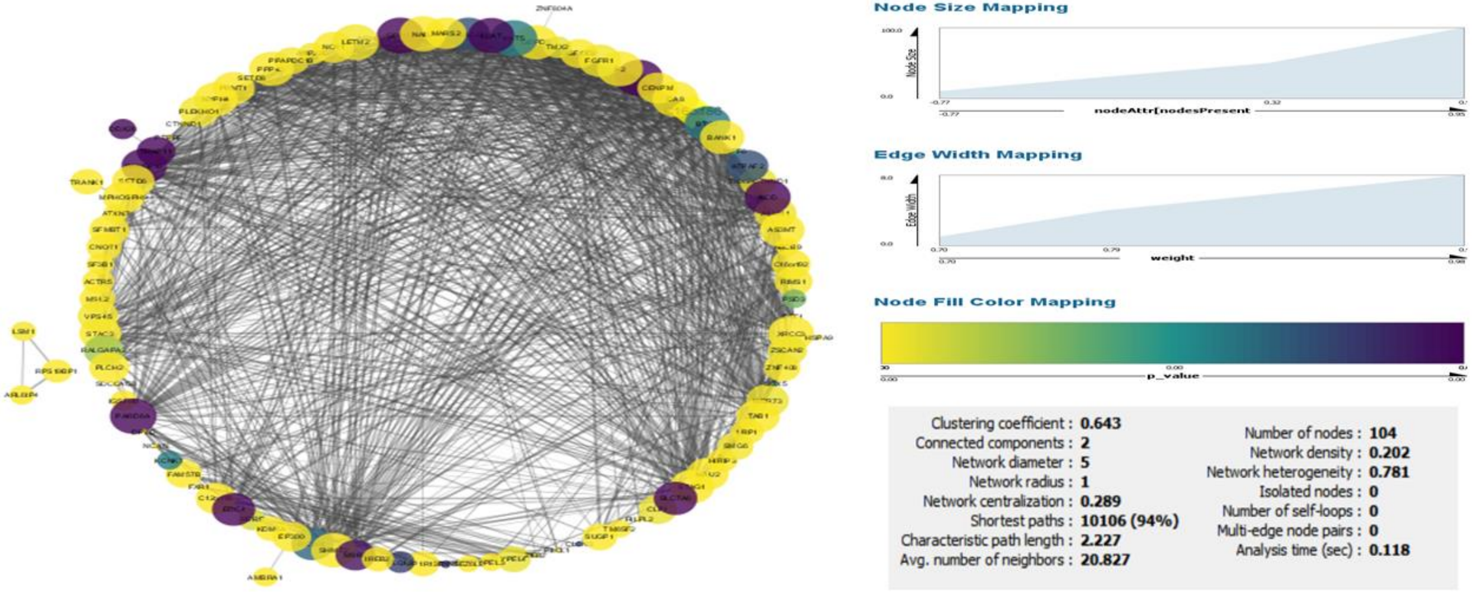
Number of nodes : **58**
Network density : **0.379**
Network heterogeneity : **0.537**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.129**

Figure 6.1.19 Turquoise Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



```
      Clustering coefficient : 0.648            Number of nodes : 21
    Connected components : 1                  Network density : 0.357
        Network diameter : 5           Network heterogeneity : 0.512
          Network radius : 3                   Isolated nodes : 0
   Network centralization : 0.379          Number of self-loops : 0
          Shortest paths : 420 (100%)    Multi-edge node pairs : 0
Characteristic path length : 1.852          Analysis time (sec) : 0.021
 Avg. number of neighbors : 7.143
```

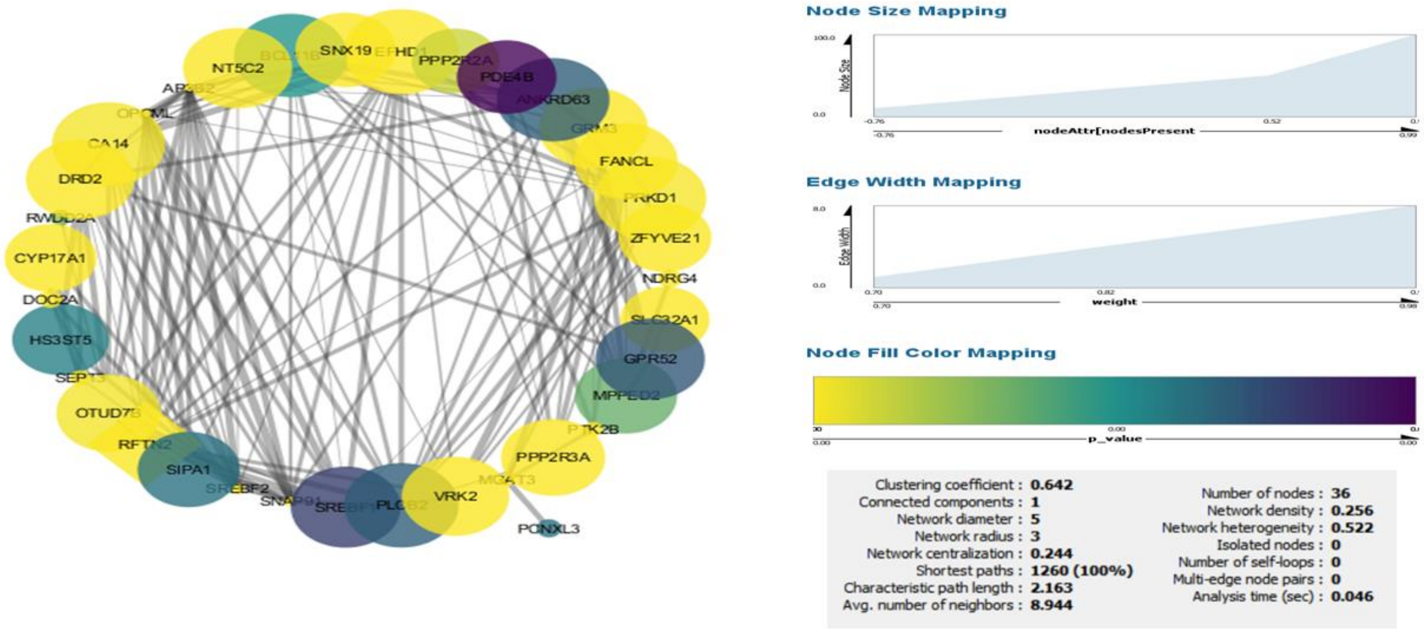
Figure 6.1.20 Yellow Module for developmental Stage Four where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.643**
Connected components : **2**
Network diameter : **5**
Network radius : **1**
Network centralization : **0.289**
Shortest paths : **10106 (94%)**
Characteristic path length : **2.227**
Avg. number of neighbors : **20.827**

Number of nodes : **104**
Network density : **0.202**
Network heterogeneity : **0.781**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.118**

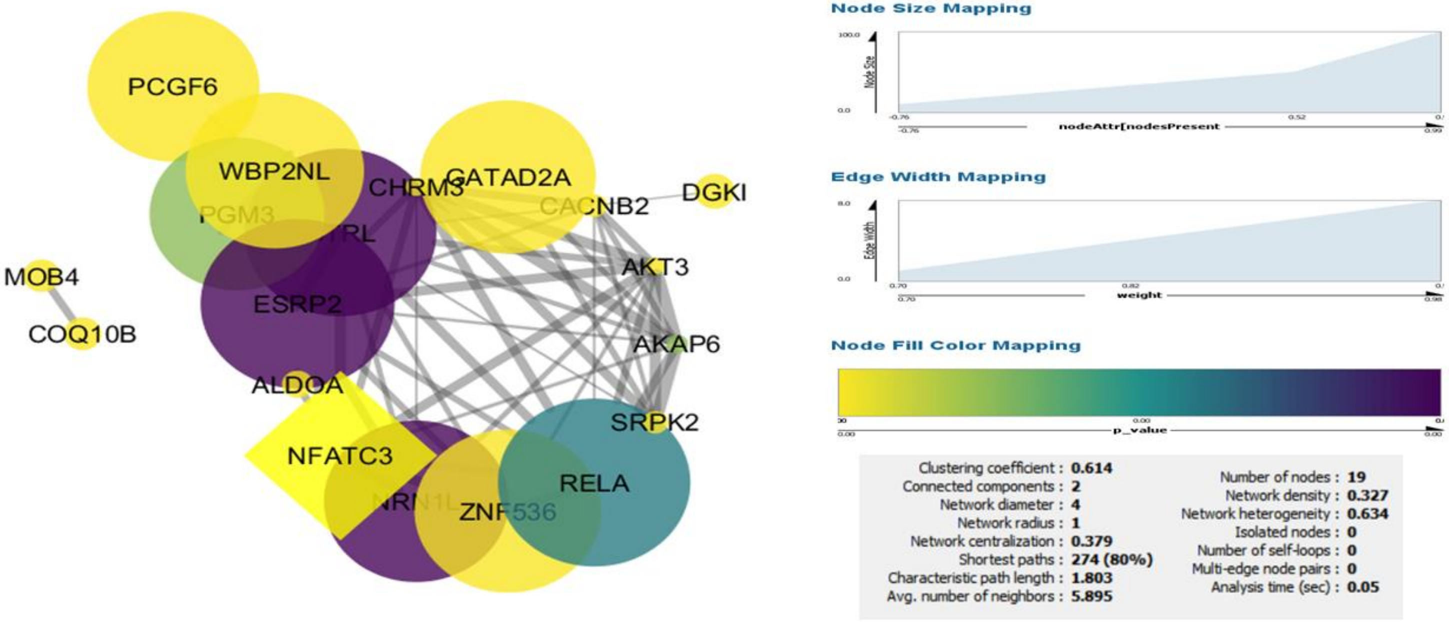
Figure 6.1.21 Black Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.642**
Connected components : **1**
Network diameter : **5**
Network radius : **3**
Network centralization : **0.244**
Shortest paths : **1260 (100%)**
Characteristic path length : **2.163**
Avg. number of neighbors : **8.944**

Number of nodes : **36**
Network density : **0.256**
Network heterogeneity : **0.522**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.046**

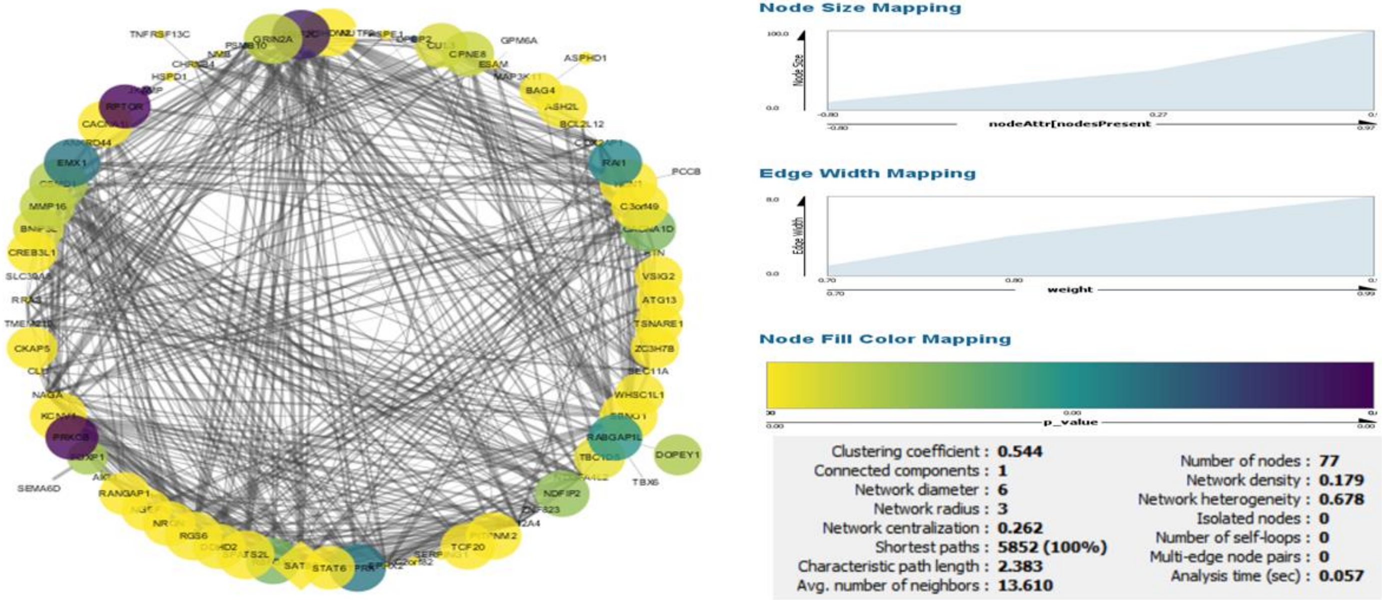
Figure 6.1.22 Brown Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.614**
Connected components : **2**
Network diameter : **4**
Network radius : **1**
Network centralization : **0.379**
Shortest paths : **274 (80%)**
Characteristic path length : **1.803**
Avg. number of neighbors : **5.895**

Number of nodes : **19**
Network density : **0.327**
Network heterogeneity : **0.634**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.05**

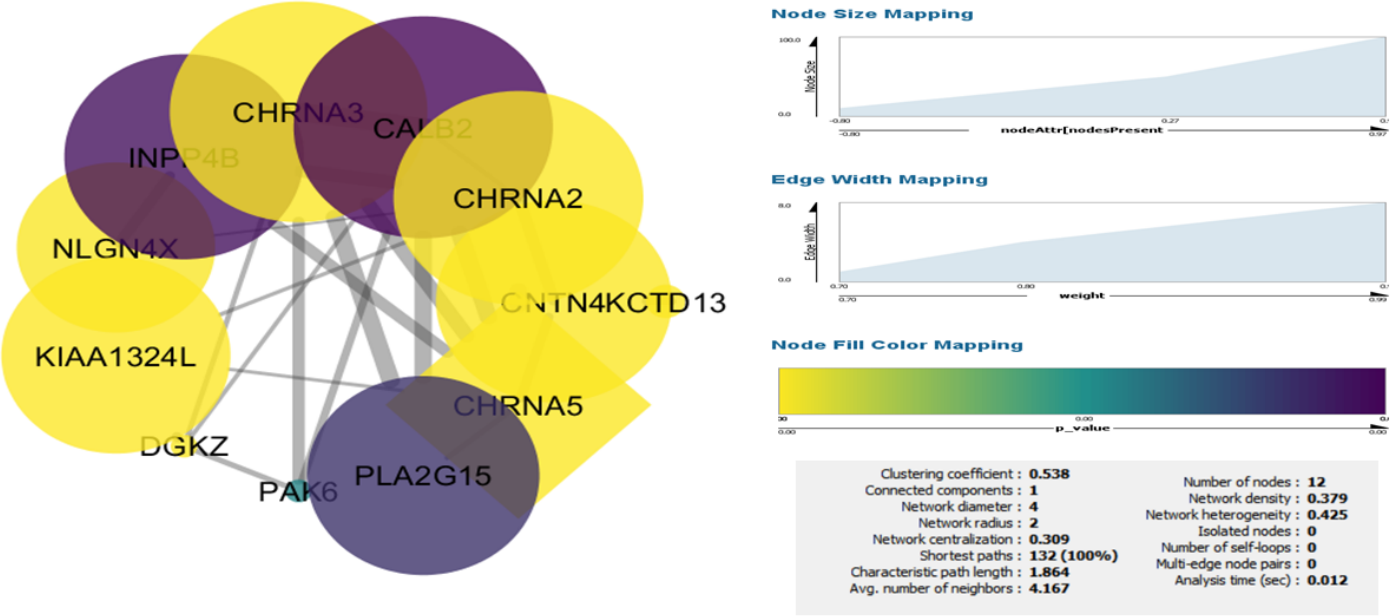
Figure 6.1.23 Green Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.544**
Connected components : **1**
Network diameter : **6**
Network radius : **3**
Network centralization : **0.262**
Shortest paths : **5852 (100%)**
Characteristic path length : **2.383**
Avg. number of neighbors : **13.610**

Number of nodes : **77**
Network density : **0.179**
Network heterogeneity : **0.678**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.057**

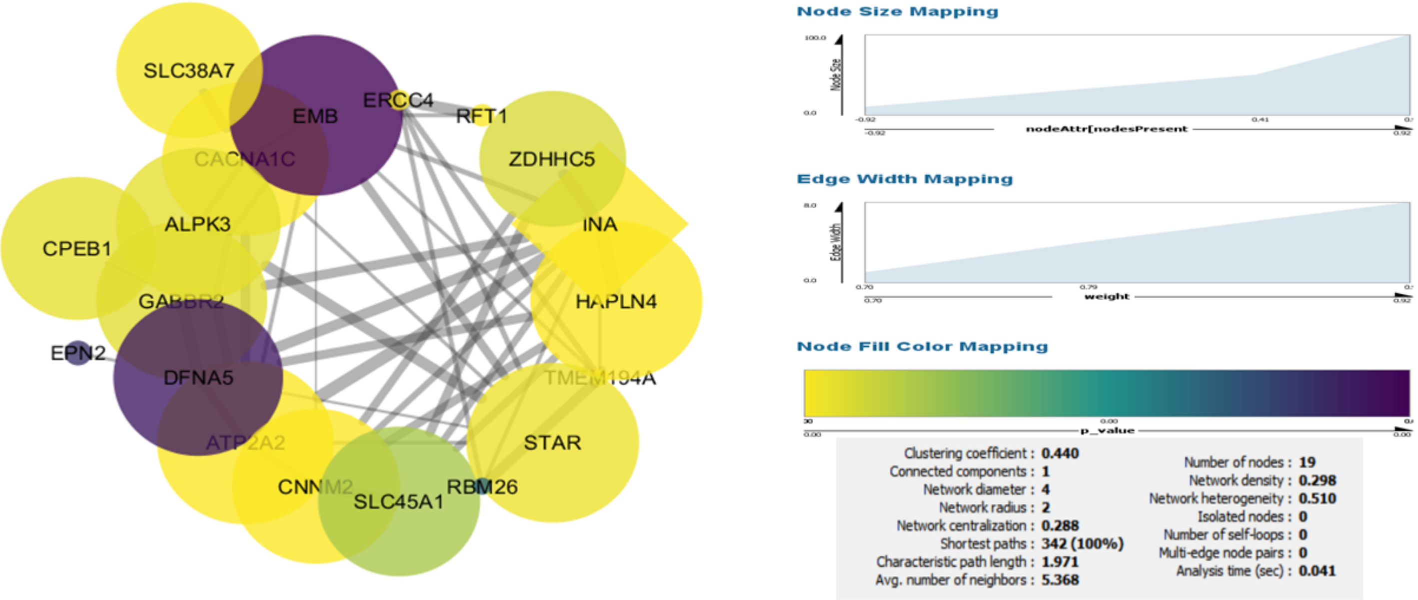
Figure 6.1.24 Greenyellow Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.538**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.309**
Shortest paths : **132 (100%)**
Characteristic path length : **1.864**
Avg. number of neighbors : **4.167**

Number of nodes : **12**
Network density : **0.379**
Network heterogeneity : **0.425**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.012**

Figure 6.1.25 Pink Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.



Clustering coefficient : **0.440**
Connected components : **1**
Network diameter : **4**
Network radius : **2**
Network centralization : **0.288**
Shortest paths : **342 (100%)**
Characteristic path length : **1.971**
Avg. number of neighbors : **5.368**

Number of nodes : **19**
Network density : **0.298**
Network heterogeneity : **0.510**
Isolated nodes : **0**
Number of self-loops : **0**
Multi-edge node pairs : **0**
Analysis time (sec) : **0.041**

Figure 6.1.26 Red Module for developmental Stage Five where the size of the node is measured using module membership, p-values are based on p-values for SNP inclusion which was determined by Pardiñas et al. and shown by node fill colour and, edge width is measured on weight. The diamond shaped node is the HG identified by WGCNA. The weight of the edges was filtered to only include 0.8 and above.

Table 6.3: Gene Ontologies identified for the Black module in developmental stage One using the anRichment function as part of WGCNA in R using the default settings.

GOID	DEFINITION	ONTOLOGY	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Black	positive regulation of macrophage proliferation	2.92E-15	ARHGAP1, CYP17A1, FSHB, GRM3, HSPE1, KCNJ13, MMP16, NRGN, STAT6, FXR1, AP3B2, CUL3, INPP4B, DGKI, GPR52, KDM4A, PSMD6, RABGAP1L, MPHOSPH9, TAB1, SDCCAG8, SMG6, SATB2, PSD3, ABCB9, KCNV1, FOXP1, B3GAT1, TMX2, MSL2, NDUFA4L2, AS3MT, SUGP1, DPEP2, CPEB1, MAIP1, EFHD1, YPEL3, IMMP2L, C12orf65, OTOL1, LETM2, PHETA2, YPEL4, STAC3
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Black	binding	1.88E-10	ARHGAP1, MPPED2, CYP17A1, FSHB, HSPE1, MMP16, NRGN, STAT6, FXR1, CUL3, INPP4B, DGKI, KDM4A, PSMD6, RABGAP1L, TAB1, SDCCAG8, SMG6, SATB2, PSD3, ABCB9, FOXP1, B3GAT1, BANK1, MSL2, SUGP1, DPEP2, CPEB1, MAIP1, EFHD1, YPEL3, C12orf65, OTOL1, LETM2, PHETA2, YPEL4, STAC3
GO:0043231	Organised structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Black	intracellular membrane-bounded organelle	6.74E-10	CYP17A1, HSPE1, MMP16, NRGN, STAT6, FXR1, AP3B2, CUL3, DGKI, KDM4A, PSMD6, RABGAP1L, MPHOSPH9, TAB1, SMG6, SATB2, ABCB9, FOXP1, B3GAT1, MSL2, NDUFA4L2, SUGP1, CPEB1, MAIP1, EFHD1, YPEL3, IMMP2L, C12orf65, LETM2, PHETA2, YPEL4, STAC3
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Black	cytoplasm	5.00E-09	ARHGAP1, CYP17A1, FSHB, HSPE1, MMP16, NRGN, STAT6, FXR1, AP3B2, CUL3, INPP4B, DGKI, KDM4A, PSMD6, RABGAP1L, MPHOSPH9, TAB1, SDCCAG8, SMG6, ABCB9, B3GAT1, NDUFA4L2, AS3MT, CPEB1, MAIP1, EFHD1, IMMP2L, C12orf65, LETM2, PHETA2, STAC3
GO:0016020	A lipid bilayer along with all the proteins and protein complexes embedded in it an attached to it.	CC	Black	membrane	2.01E-08	ARHGAP1, CYP17A1, GRM3, HSPE1, KCNJ13, MMP16, NRGN, STAT6, FXR1, AP3B2, CUL3, DGKI, GPR52, MPHOSPH9, TAB1, PSD3, ABCB9, KCNV1, B3GAT1, TMX2, NDUFA4L2, DPEP2, CPEB1, MAIP1, EFHD1, IMMP2L,

						LETM2, STAC3
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Black	intracellular organelle lumen	9.92E-08	ARHGAP1, CYP17A1, FSHB, GRM3, HSPE1, KCNJ13, MMP16, NRGN, STAT6, FXR1, CUL3, INPP4B, DGKI, GPR52, KDM4A, PSMD6, RABGAP1L, TAB1, SDCCAG8, SMG6, SATB2, PSD3, KCNV1, FOXP1, TMX2, BANK1, CPEB1, MAIP1, EFHD1, YPEL3, STAC3

Table 6.4: Gene Ontologies for the Blue module in developmental stage One using the anRichment function of WGCNA on the schizophrenia-associated genes.

GOID	DEFINITION	ONTOLOGY	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	1.75E-29	More than 50 overlapping genes
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Blue	intracellular organelle lumen	2.74E-14	More than 50 overlapping genes
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl) benzene-1,2-diol] and derivatives formed by substitution.	BP	Blue	cellular response to catecholamine stimulus	4.03E-13	CLCN3, DRD2, EP300, EPHX2, ETF1, PTK2B, GPM6A, LRP1, NEK1, PPP2R2A, PRKCB, RELA, SIPA1, TCF4, HIRIP3, ASH2L, INA, AKAP6, BAG4, SNAP91, AKT3, CLP1, SF3B1, VPS13C, SRR, BCL11B, ACD, SETD6, PCGF6, WDR73, ATPAF2, DNAJC19, RILPL2, KCTD13, INO80E, HCN1, BTBD18
GO:0005737	All the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Blue	cytoplasm	1.15E-11	SERPING1, CALB2, CLCN3, DRD2, EP300, EPHX2, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PCCB, PGM3, PLCL1, PPP2R2A, PRKCB, PSMB10, RANGAP1, RELA, SIPA1, STAR, VRK2, DGKZ, INA, AKAP6, BAG4, ATP5MPL, SNAP91, AKT3, RAI1, CLP1, DOP1A, VPS13C, SRR, BOLL, SETD6, DRC3, TLCD3B, WDR73, ATPAF2, SFXN5, BORCS7, DNAJC19, RILPL2, HS3ST5
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) because of a diacylated bacterial lipopeptide stimulus.	BP	Blue	cellular response to diacyl bacterial lipopeptide	1.70E-11	SERPING1, DRD2, EP300, EPHX2, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PCCB, PGM3, PLCL1, PPP2R2A, PRKCB, PSMB10, RANGAP1, RELA, RRAS, STAR, TCF4, VRK2, DGKZ, ASH2L, BAG4, AKT3, RAI1, CLP1, ZC3H7B, SF3B1, AIG1, ADAMTSL3, ALPK3, SCAF1, SRR, BCL11B, ACD, BOLL, ZNF408, SETD6, TLCD3B, PCGF6, MED19, HS3ST5, KCTD13, INO80E, BTBD18
GO:0016020	A lipid bilayer along with all the proteins and protein complexes embedded in it an attached to it.	CC	Blue	membrane	2.79E-11	CALB2, CLCN3, DRD2, PTK2B, GPM6A, LRP1, MAP3K11, PLCL1, PRKCB, RANGAP1, RRAS, SIPA1, STAR, VRK2, DGKZ, INA, AKAP6, BAG4,

						ATP5MPL, SNAP91, AKT3, DOP1A, SEZ6L2, SLC45A1, AIG1, CNNM2, VPS13C, SLC38A7, SRR, TLCD3B, WDR73, SFXN5, BORCS7, TMEM219, RFTN2, DNAJC19, CPNE8, C16orf92, RILPL2, HS3ST5, HCN1
GO:0097178	The aggregation, arrangement and bonding together of a set of components to form a ruffle, a projection at the leading edge of a crawling cell; the protrusions are supported by a microfilament meshwork. The formation of ruffles (also called membrane ruffling) is thought to be controlled by a group of enzymes known as Rho GTPases, specifically RhoA, Rac1 and cdc42.	BP	Blue	ruffle assembly	1.53E-10	CLCN3, DRD2, EP300, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PCCB, PRKCB, RANGAP1, RELA, RRAS, STAR, TCF4, VRK2, DGKZ, BAG4, AKT3, CLP1, ZC3H7B, SF3B1, CNNM2, ALPK3, SCAF1, SRR, BCL11B, ACD, BOLL, ZNF408, PCGF6, HS3ST5, HCN1
GO:0048519	Any process that stops, prevents, or reduces the frequency, rate or extent of a biological process. Biological processes are regulated by many means; examples include the control of gene expression, protein modification or interaction with a protein or substrate molecule.	BP	Blue	negative regulation of biological process	1.73E-10	SERPING1, DRD2, EP300, ETF1, PTK2B, LRP1, PLCL1, PPP2R2A, PRKCB, PSMB10, RANGAP1, RELA, RRAS, SIPA1, STAR, AKAP6, BAG4, AKT3, RAI1, CLP1, ZC3H7B, VPS13C, BCL11B, ACD, BOLL, SETD6, TLCD3B, PCGF6, WDR73, HS3ST5, KCTD13, BTBD18
GO:0005634	A membrane-bounded organelle of eukaryotic cells in which chromosomes are housed and replicated. In most cells, the nucleus contains all of the cell's chromosomes except the organellar chromosomes and is the site of RNA synthesis and processing. In some species, or in specialized cell types, RNA metabolism or DNA replication may be absent.	CC	Blue	nucleus	5.27E-10	CALB2, EP300, PTK2B, LRP1, NEK1, PPP2R2A, PRKCB, PSMB10, RANGAP1, RELA, SIPA1, TCF4, VRK2, HIRIP3, DGKZ, ASH2L, INA, AKAP6, BAG4, AKT3, RAI1, CLP1, ZC3H7B, SF3B1, ALPK3, SCAF1, BCL11B, ACD, ZNF408, SETD6, PCGF6, ATPAF2, MED19, KCTD13, INO80E, BTBD18
GO:0006996	A process that is carried out at the cellular level which results in the	BP	Blue	organelle organization	6.74E-10	CLCN3, EP300, EPHX2, PTK2B, LRP1, NEK1, PPP2R2A, PRKCB, RELA, SIPA1, HIRIP3, ASH2L,

	assembly, arrangement of constituent parts, or disassembly of an organelle within a cell. An organelle is an organized structure of distinctive morphology and function. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes and the cytoskeleton. Excludes the plasma membrane.					INA, BAG4, SNAP91, AKT3, VPS13C, ACD, SETD6, PCGF6, WDR73, DNAJC19, RILPL2, KCTD13, INO80E, BTBD18
GO:0060255	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways involving macromolecules, any molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.	BP	Blue	regulation of macromolecule metabolic process	1.47E-09	SERPING1, DRD2, EP300, EPHX2, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PLCL1, PPP2R2A, PRKCB, PSMB10, RELA, RRAS, STAR, TCF4, VRK2, ASH2L, BAG4, RAI1, CLP1, ZC3H7B, SF3B1, BCL11B, ACD, BOLL, ZNF408, SETD6, PCGF6, MED19, KCTD13, BTBD18
GO:1901561	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a benomyl stimulus.	BP	Blue	response to benomyl	3.08E-09	CLCN3, EP300, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PCCB, PRKCB, RANGAP1, RELA, RRAS, TCF4, VRK2, DGKZ, BAG4, AKT3, CLP1, ZC3H7B, SF3B1, CNNM2, ALPK3, SCAF1, SRR, BCL11B, ACD, BOLL, ZNF408, PCGF6, HS3ST5, HCN1
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Blue	regulation of fertilization	1.07E-08	SERPING1, DRD2, EP300, EPHX2, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PLCL1, PPP2R2A, PRKCB, PSMB10, RELA, RRAS, STAR, TCF4, VRK2, DGKZ, ASH2L, BAG4, RAI1, BCL11B, ACD, BOLL, ZNF408, SETD6, PCGF6, MED19, BTBD18
GO:0043168	Interacting selectively and non-covalently with anions, charged atoms or groups of atoms with a net negative charge.	MF	Blue	anion binding	1.45E-08	CLCN3, PTK2B, LRP1, MAP3K11, NEK1, PCCB, PLCL1, PRKCB, RELA, RRAS, VRK2, DGKZ, SNAP91, AKT3, CLP1, CNNM2, ALPK3, SRR, CPNE8, HS3ST5, HCN1

GO:0006464	The covalent alteration of one or more amino acids occurring in proteins, peptides and nascent polypeptides (co-translational, post-translational modifications) occurring at the level of an individual cell. Includes the modification of charged tRNAs that are destined to occur in a protein (pre-translation modification).	BP	Blue	cellular protein modification process	1.59E-08	DRD2, EP300, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PGM3, PLCL1, PPP2R2A, PRKCB, PSMB10, RANGAP1, RELA, RRAS, VRK2, ASH2L, BAG4, AKT3, ALPK3, SETD6, PCGF6, HS3ST5, KCTD13, INO80E
GO:1901666	Any process that activates or increases the frequency, rate or extent of NAD+ ADP-ribosyltransferase activity.	BP	Blue	positive regulation of NAD+ ADP-ribosyltransferase activity	2.84E-08	SERPING1, DRD2, EP300, ETF1, PTK2B, LRP1, MAP3K11, NEK1, PCCB, PGM3, PLCL1, PPP2R2A, PRKCB, PSMB10, RANGAP1, RELA, RRAS, VRK2, ASH2L, BAG4, AKT3, ADAMTSL3, ALPK3, SRR, BOLL, SETD6, TLCD3B, PCGF6, HS3ST5, KCTD13, INO80E
GO:0032502	A biological process whose specific outcome is the progression of an integrated living unit: an anatomical structure (which may be a subcellular structure, cell, tissue, or organ), or organism over time from an initial condition to a later condition.	BP	Blue	developmental process	5.31E-08	SERPING1, CLCN3, DRD2, EP300, PTK2B, GPM6A, LRP1, PGM3, PRKCB, PSMB10, RELA, RRAS, STAR, TCF4, ASH2L, INA, AKAP6, AKT3, RAI1, CLP1, ALPK3, SRR, BCL11B, BOLL, SETD6, TLCD3B, DNAJC19, RILPL2, HCN1, BTBD18
GO:1901679	The directed movement of nucleotide across a membrane.	BP	Blue	nucleotide transmembrane transport	6.96E-08	DRD2, EP300, EPHX2, ETF1, PTK2B, PGM3, PLCL1, PRKCB, PSMB10, RELA, STAR, TCF4, DGKZ, ASH2L, RAI1, CLP1, SCAF1, SRR, BCL11B, ACD, BOLL, ZNF408, SETD6, TLCD3B, PCGF6, MED19, HS3ST5, KCTD13, BTBD18
GO:0044249	The chemical reactions and pathways resulting in the formation of substances, carried out by individual cells.	BP	Blue	cellular biosynthetic process	2.45E-07	DRD2, EP300, EPHX2, ETF1, PTK2B, PGM3, PRKCB, PSMB10, RELA, STAR, TCF4, DGKZ, ASH2L, RAI1, CLP1, SCAF1, SRR, BCL11B, ACD, BOLL, ZNF408, SETD6, TLCD3B, PCGF6, MED19, HS3ST5, KCTD13, BTBD18
GO:0006950	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene	BP	Blue	response to stress	2.85E-07	SERPING1, DRD2, EP300, EPHX2, PTK2B, LRP1, MAP3K11, NEK1, PRKCB, PSMB10, RANGAP1, RELA, SIPA1, STAR, VRK2, DGKZ, ASH2L, BAG4, AKT3, VPS13C, ACD, SETD6, INO80E

	expression, etc.) as a result of a disturbance in organismal or cellular homeostasis, usually, but not necessarily, exogenous (e.g., temperature, humidity, ionizing radiation).					
GO:0031325	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways by which individual cells transform chemical substances.	BP	Blue	positive regulation of cellular metabolic process	2.96E-07	DRD2, EP300, PTK2B, LRP1, MAP3K11, NEK1, PRKCB, RELA, STAR, TCF4, DGKZ, ASH2L, BAG4, RAI1, VPS13C, BCL11B, ACD, BOLL, MED19, KCTD13, BTBD18
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.	MF	Blue	armadillo repeat domain binding	2.96E-07	SERPING1, CALB2, CLCN3, DRD2, EP300, EPHX2, PTK2B, GPM6A, LRP1, PLCL1, PRKCB, PSMB10, RANGAP1, STAR, DGKZ, AKAP6, BAG4, AKT3, CNNM2, VPS13C, ACD, BOLL, HCN1
GO:0032559	Interacting selectively and non-covalently with an adenyl ribonucleotide, any compound consisting of adenosine esterified with (ortho)phosphate or an oligophosphate at any hydroxyl group on the ribose moiety.	MF	Blue	adenyl ribonucleotide binding	3.98E-07	CLCN3, PTK2B, MAP3K11, NEK1, PCCB, PRKCB, VRK2, DGKZ, AKT3, CLP1, CNNM2, ALPK3, SRR, HS3ST5, HCN1

Table 6.5: Gene Ontology Brown Module for developmental stage One using the anRichment function as part of WGCNA in R using the default settings

GOID	DEFINITION	ONTOLOGY	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	2.10E-15	SERPINC1, ATP2A2, CACNB2, CHRM3, CHRNA2, CTRL, IREB2, NAB2, NAGA, NDUFA6, SREBF2, SRPK2, DOC2A, CACNA1I, GABBR2, ZNF536, R3HDM2, IGSF9B, DDHD2, MOB4, PARD6A, SFMBT1, ZSCAN2, SBNO1, ZNF823, RPTOR, RBM26, COQ10B, GFOD2, L3MBTL2, CREB3L1, WBP2NL, KIAA1324L, EHBP1L1, KMT5A, SNX19
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Brown	binding	3.36E-10	SERPINC1, ATP2A2, CACNB2, CHRM3, CHRNA2, IREB2, NAB2, NAGA, SREBF2, SRPK2, DOC2A, CACNA1I, GABBR2, ZNF536, R3HDM2, IGSF9B, DDHD2, MOB4, PARD6A, SFMBT1, ZSCAN2, ZNF823, RPTOR, RBM26, COQ10B, L3MBTL2, CREB3L1, WBP2NL, KMT5A, SNX19
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	4.32E-09	SERPINC1, ATP2A2, CACNB2, CHRM3, CHRNA2, IREB2, NAB2, SREBF2, SRPK2, DOC2A, CACNA1I, GABBR2, ZNF536, IGSF9B, DDHD2, PARD6A, SFMBT1, ZSCAN2, SBNO1, ZNF823, RPTOR, RBM26, L3MBTL2, CREB3L1, WBP2NL, KMT5A, SNX19
GO:0043231	Organised structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Brown	intracellular membrane-bounded organelle	2.02E-08	SERPINC1, ATP2A2, IREB2, NAB2, NAGA, NDUFA6, SREBF2, SRPK2, DOC2A, ZNF536, R3HDM2, DDHD2, MOB4, PARD6A, SFMBT1, ZSCAN2, SBNO1, ZNF823, RPTOR, RBM26, COQ10B, L3MBTL2, CREB3L1, WBP2NL, KMT5A

Table 6.6: Gene Ontology for Pink Module in Developmental Stage One using the anRichment function as part of WGCNA in R using the default settings

GOID	DEFINITION	ONTOLOGY	Module	GO Process	FDR	Genes
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Pink	protein binding	1.61E-08	ERCC4, HSPA9, PPP2R3A, PPP4C, SHMT2, XRCC3, TBC1D5, NUTF2, STAG1, KAT5, LSM1, HPF1, PAK6, OTUD7B, SEMA6D, THOC7, TYW5, TOM1L2, TSNARE1
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Pink	positive regulation of macrophage proliferation	2.30E-07	ERCC4, HSPA9, PPP4C, SHMT2, XRCC3, PLCH2, TBC1D5, NUTF2, STAG1, KAT5, LSM1, HPF1, TSR1, PAK6, OTUD7B, SEMA6D, THOC7, RFT1, TYW5, TOM1L2, TSNARE1
GO:0008152	The chemical reactions and pathways, including anabolism and catabolism, by which living organisms transform chemical substances. Metabolic processes typically transform small molecules, but also include macromolecular processes such as DNA repair and replication, and protein synthesis and degradation.	BP	Pink	metabolic process	6.41E-07	ERCC4, HSPA9, PPP2R3A, PPP4C, SHMT2, XRCC3, PLCH2, TBC1D5, NUTF2, STAG1, KAT5, LSM1, HPF1, TSR1, PAK6, OTUD7B, THOC7, TYW5
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Pink	intracellular organelle lumen	8.50E-07	ERCC4, HSPA9, PPP2R3A, PPP4C, SHMT2, XRCC3, PLCH2, TBC1D5, NUTF2, STAG1, KAT5, LSM1, HPF1, PAK6, OTUD7B, SEMA6D, RFT1, TOM1L2

Table 6.7: Gene Ontologies for Turquoise Module in Developmental Stage One using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	9.49E-49	More than 50 overlapping genes
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	1.07E-17	More than 50 overlapping genes
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.	MF	Turquoise	armadillo repeat domain binding	4.10E-17	BNIP3L, CACNA1C, CACNA1D, CHRNA3, CHRNA5, CHRNB4, CLU, CTNND1, DPYD, EMX1, F2, FGFR1, GRIN2A, PRMT1, MEF2C, NMB, PDE4B, PRKD1, PSMA4, PTN, SLC12A4, SREBF1, TLE3, ALMS1, TAOK2, KCNK7, VPS45, RIMS1, PLCL2, PPP1R13B, NGEF, GIGYF2, PLEKHO1, TM6SF2, NLGN4X, SLC39A8, CSMD1, ZNF804A, TNFRSF13C, CARMIL2, CNTN4
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl)benzene-1,2-diol] and derivatives formed by substitution.	BP	Turquoise	cellular response to catecholamine stimulus	1.97E-14	RERE, BNIP3L, BTG1, CHRNA3, CHRNB4, CLU, NCAN, CTNND1, EMX1, F2, FGFR1, PRMT1, MEF2C, PRKD1, PTN, PTPRK, ATXN7, SLC12A4, SREBF1, TLE3, ALMS1, TAOK2, ATG13, CKAP5, ZEB2, EPN2, RIMS1, CNOT1, PPP1R13B, MAU2, NGEF, PLEKHO1, HYDIN, NSD3, AMBRA1, NLGN4X, NDRG4, CENPM, ACTR5, CENPT, ANP32E, ESAM, ZNF804A, TDRD9, EMB, CARMIL2, CNTN4
GO:0005829	The part of the cytoplasm that does not contain organelles, but which does contain other particulate matter, such as protein complexes.	CC	Turquoise	cytosol	8.03E-14	BNIP3L, CLU, CTNND1, GSDME, DPYD, FGFR1, FHIT, PRMT1, MSRA, NFATC3, PDE4B, PRKD1, PSMA4, ATXN7, SREBF1, ALMS1, TAOK2, RGS6, ATG13, CKAP5, ZEB2, DMTF1, EPN2, NT5C2, RIMS1, CNOT1, PPP1R13B, EDC4, NGEF, SPATS2L,

						GIGYF2, AMBRA1, RALGAPA2, PITPNM2, RANBP10, NDRG4, CENPM, CENPT, RPS19BP1, HARBI1, ASPG
GO:0048523	Any process that stops, prevents, or reduces the frequency, rate or extent of a cellular process, any of those that are carried out at the cellular level, but are not necessarily restricted to a single cell. For example, cell communication occurs among more than one cell, but occurs at the cellular level.	BP	Turquoise	negative regulation of cellular process	2.03E-13	RERE, BNIP3L, BTG1, CACNA1C, CLU, CTNND1, GSDME, F2, FGFR1, FHIT, PRMT1, MEF2C, MGAT3, NFATC3, NMB, PRKD1, PSMA4, PTN, PTPRK, SOX5, SREBF1, TLE3, TAOK2, RGS6, ZEB2, EPN2, CNOT1, PLCL2, PPP1R13B, NGEF, GIGYF2, GATAD2A, AMBRA1, THAP11, NLGN4X, NDRG4, BCL2L12, TDRD9, CARMIL2, CNTN4
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Turquoise	regulation of fertilization	8.45E-13	RERE, BTG1, CHRNA3, CLU, GSDME, EMX1, F2, FGFR1, FHIT, GRIN2A, PRMT1, MEF2C, MGAT3, NFATC3, PRKD1, PSMA4, PTN, PTPRK, SOX5, SREBF1, TLE3, TAOK2, ZNF592, ATG13, ZEB2, DMTF1, CNOT1, PLCL2, GIGYF2, TM6SF2, NDFIP2, RBFOX1, GATAD2A, NSD3, AMBRA1, THAP11, NDRG4, ACTR5, ESRP2, CENPT, BCL2L12, ZBTB37, TNFRSF13C
GO:0071986	A eukaryotically conserved protein complex; in humans, it is comprised of LAMTOR1, LAMTOR2, LAMTOR3, LAMTOR4, and LAMTOR5. The complex is anchored to lipid rafts in late endosome membranes via LAMTOR1, constitutes a guanine nucleotide exchange factor (GEF) for the Rag GTPases.	CC	Turquoise	Ragulator complex	4.49E-12	CACNA1C, CACNA1D, CHRNA3, CHRNA5, CHRNB4, CLU, CTNND1, GSDME, F2, FGFR1, FHIT, GRIN2A, OPCML, PDE4B, PRKD1, PTN, PTPRK, SLC12A4, RGS6, CKAP5, KCNK7, RIMS1, PPP1R13B, VSIG2, CA14, ZDHHC5, PLEKHO1, RALGAPA2, NLGN4X, SLC39A8, GPR135, NDRG4, PLPP5, ESAM, ZNF804A, TNFRSF13C, EMB, CARMIL2, CNTN4, HARBI1, SNORC
GO:1901666	NA	BP	Turquoise	positive regulation of NAD+ ADP-ribosyltransferase activity	7.48E-12	BNIP3L, BTG1, CHRNA3, CLU, NCAN, GSDME, DPYD, F2, FGFR1, FHIT, GRIN2A, PRMT1, MEF2C, MGAT3, MSRA, PDE4B, PRKD1, PSMA4, PTN, PTPRK, ATXN7, SREBF1, TAOK2, ATG13, ZEB2, NT5C2, CNOT1, PLCL2, SEC11A, PLA2G15, ZDHHC5, GIGYF2, JKAMP, NDFIP2, NSD3,

						FANCL, NDRG4, GDPD3, ACTR5, BCL2L12, MARS2, TNFRSF13C, ASPHD1, ASPG
GO:0120069	Any process that increases the frequency, rate or extent of any stomach fundus smooth muscle contraction.	BP	Turquoise	positive regulation of stomach fundus smooth muscle contraction	3.89E-11	RERE, CHRNA3, EMX1, FGFR1, PRMT1, MEF2C, PRKD1, PTN, PTPRK, ALMS1, TAOK2, CKAP5, ZEB2, RIMS1, NGEF, PLEKHO1, HYDIN, NDRG4, ZNF804A, EMB, CARMIL2, CNTN4
GO:0005886	The membrane surrounding a cell that separates the cell from its external environment. It consists of a phospholipid bilayer and associated proteins.	CC	Turquoise	plasma membrane	5.94E-11	CACNA1C, CACNA1D, CHRNA3, CHRNA5, CHRNB4, CTNND1, GSDME, F2, FGFR1, FHIT, GRIN2A, OPCML, PDE4B, PRKD1, PTN, PTPRK, SLC12A4, RGS6, CKAP5, KCNK7, RIMS1, PPP1R13B, VSIG2, CA14, ZDHHC5, PLEKHO1, RALGAPA2, NLGN4X, SLC39A8, GPR135, NDRG4, PLPP5, ESAM, ZNF804A, TNFRSF13C, EMB, CARMIL2, CNTN4, HARBI1
GO:0016021	The component of a membrane consisting of the gene products and protein complexes having at least some part of their peptide sequence embedded in the hydrophobic region of the membrane.	CC	Turquoise	integral component of membrane	2.67E-10	BNIP3L, CACNA1C, CACNA1D, CHRNA3, CHRNA5, CHRNB4, FGFR1, GRIN2A, MGAT3, PDE4B, PTPRK, SLC12A4, SREBF1, TAOK2, KCNK7, VPS45, NEMP1, SEC11A, VSIG2, CA14, ZDHHC5, GIGYF2, JKAMP, TM6SF2, NDFIP2, NLGN4X, SLC39A8, CSMD1, GPR135, GDPD3, PLPP5, ESAM, TNFRSF13C, EMB, ASPHD1, SNORC, PCNX3
GO:0043232	Organised structure of distinctive morphology and function, not bounded by a lipid bilayer membrane and occurring within the cell. Includes ribosomes, the cytoskeleton and chromosomes.	CC	Turquoise	intracellular non-membrane-bounded organelle	3.17E-10	CACNA1C, CACNA1D, CLU, EMX1, FGFR1, FHIT, MEF2C, MSRA, PDE4B, PSMA4, ATXN7, ALMS1, TAOK2, CKAP5, ZEB2, RIMS1, CNOT1, MAU2, EDC4, SPATS2L, GIGYF2, ARL6IP4, HYDIN, GATAD2A, NSD3, AMBRA1, CENPM, ACTR5, CENPT, ANP32E, RPS19BP1, ZNF804A, TDRD9, CARMIL2, HARBI1
GO:0120060	Any process that modulates the frequency, rate or extent of any gastric emptying process, the process in which the liquid and liquid-suspended solid contents of the	BP	Turquoise	regulation of gastric emptying	4.35E-10	CACNA1C, CHRNA3, CHRNA5, CHRNB4, CLU, CTNND1, GRIN2A, NMB, PDE4B, PTPRK, ALMS1, TAOK2, NGEF, ZDHHC5, GIGYF2, PLEKHO1, HYDIN, AMBRA1,

	stomach exit through the pylorus into the duodenum.					NLGN4X, NDRG4, ZNF804A, EMB, CARMIL2, CNTN4
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Turquoise	collagen V binding	4.49E-10	RERE, BNIP3L, BTG1, CLU, NCAN, EMX1, F2, FGFR1, FHIT, PRMT1, MEF2C, MSRA, NFATC3, PSMA4, ATXN7, SREBF1, TLE3, TAOK2, CKAP5, ZEB2, DMTF1, PPP1R13B, MAU2, EDC4, PLA2G15, SPATS2L, ARL6IP4, GATAD2A, NSD3, FANCL, THAP11, CENPM, ACTR5, ESRP2, CENPT, ANP32E, RPS19BP1, MARS2
GO:0044267	The chemical reactions and pathways involving a specific protein, rather than of proteins in general, occurring at the level of an individual cell. Includes cellular protein modification.	BP	Turquoise	cellular protein metabolic process	8.60E-10	BNIP3L, BTG1, CHRNA3, CLU, GSDME, F2, FGFR1, FHIT, GRIN2A, PRMT1, MEF2C, MGAT3, MSRA, PRKD1, PSMA4, PTN, PTPRK, ATXN7, SREBF1, TAOK2, ATG13, ZEB2, CNOT1, PLCL2, SEC11A, ZDHHC5, GIGYF2, JKAMP, NDFIP2, NSD3, FANCL, NDRG4, ACTR5, BCL2L12, MARS2, ASPHD1
GO:0010468	Any process that modulates the frequency, rate or extent of gene expression. Gene expression is the process in which a gene's coding sequence is converted into a mature gene product or products (proteins or RNA). This includes the production of an RNA transcript as well as any processing to produce a mature RNA product or an mRNA or circRNA (for protein-coding genes) and the translation of that mRNA or circRNA into protein. Protein maturation is included when required to form an active form of a product from an inactive precursor form.	BP	Turquoise	regulation of gene expression	9.68E-10	RERE, BTG1, CLU, EMX1, F2, FGFR1, PRMT1, MEF2C, NFATC3, PDE4B, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, RIMS1, CNOT1, EDC4, GIGYF2, NDFIP2, RBFOX1, GATAD2A, NSD3, THAP11, ACTR5, ESRP2, CENPT, BCL2L12, ZBTB37, ZNF804A, TNFRSF13C, TDRD9
GO:0031175	The process whose specific outcome is the progression of a neuron projection over time, from its formation to the mature structure. A neuron projection is any process extending from a neural cell, such	BP	Turquoise	neuron projection development	2.13E-09	RERE, CHRNA3, EMX1, FGFR1, PRMT1, MEF2C, PRKD1, PTN, PTPRK, TAOK2, ZEB2, RIMS1, NGEF, NDRG4, ZNF804A, EMB, CNTN4

	as axons or dendrites (collectively called neurites).					
GO:1901555	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a paclitaxel stimulus.	BP	Turquoise	response to paclitaxel	5.31E-09	RERE, BTG1, CLU, DPYD, EMX1, FGFR1, FHIT, GRIN2A, PRMT1, MEF2C, NFATC3, PDE4B, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, NT5C2, CNOT1, EDC4, GIGYF2, ARL6IP4, RBFOX1, GATAD2A, NSD3, FANCL, THAP11, ACTR5, ESRP2, CENPT, BCL2L12, ZBTB37, MARS2, TDRD9, HARBI1
GO:0009653	The process in which anatomical structures are generated and organized. Morphogenesis pertains to the creation of form.	BP	Turquoise	anatomical structure morphogenesis	1.15E-08	RERE, BTG1, CACNA1C, CHRNA3, CLU, EMX1, F2, FGFR1, MEF2C, PRKD1, PSMA4, PTN, TLE3, TAOK2, ZEB2, EPN2, RIMS1, NGEF, PLEKHO1, NDRG4, ESRP2, TNFRSF13C, EMB, CARMIL2, CNTN4
GO:1901679	The directed movement of nucleotide across a membrane.	BP	Turquoise	nucleotide transmembrane transport	1.22E-08	RERE, BTG1, CLU, NCAN, DPYD, EMX1, FGFR1, PRMT1, MEF2C, MGAT3, NFATC3, PDE4B, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, NT5C2, CNOT1, PLCL2, ZDHHC5, GIGYF2, GATAD2A, NSD3, THAP11, PITPNM2, ACTR5, CENPT, BCL2L12, ZBTB37, MARS2, TNFRSF13C, ASPG
GO:0043005	A prolongation or process extending from a nerve cell, e.g. an axon or dendrite.	CC	Turquoise	neuron projection	1.83E-08	CACNA1C, CHRNA3, CHRNA5, CHRNB4, CLU, CTNND1, GRIN2A, NMB, PDE4B, PTPRK, TAOK2, NGEF, ZDHHC5, GIGYF2, NLGN4X, ZNF804A, EMB, CNTN4
GO:0035556	The process in which a signal is passed on to downstream components within the cell, which become activated themselves to further propagate the signal and finally trigger a change in the function or state of the cell.	BP	Turquoise	intracellular signal transduction	2.69E-08	CA8, CACNA1C, CLU, GSDME, F2, FGFR1, FHIT, GRIN2A, PRMT1, MEF2C, NFATC3, PRKD1, PSMA4, TAOK2, RGS6, ZEB2, CNOT1, PLCL2, PPP1R13B, NGEF, NDFIP2, RALGAPA2, PITPNM2, NDRG4, BCL2L12
GO:0007417	The process whose specific outcome is the progression of the central nervous system over time, from its formation to the mature structure. The central nervous system is	BP	Turquoise	central nervous system development	2.88E-08	RERE, CLU, NCAN, CTNND1, EMX1, F2, FGFR1, GRIN2A, PTN, ZEB2, GIGYF2, HYDIN, NLGN4X, NDRG4, CNTN4, HAPLN4

	the core nervous system that serves an integrating and coordinating function. In vertebrates it consists of the brain and spinal cord. In those invertebrates with a central nervous system it typically consists of a brain, cerebral ganglia and a nerve cord.					
GO:0044249	The chemical reactions and pathways resulting in the formation of substances, carried out by individual cells.	BP	Turquoise	cellular biosynthetic process	3.33E-08	RERE, BTG1, CLU, NCAN, DPYD, EMX1, FGFR1, PRMT1, MEF2C, MGAT3, NFATC3, PDE4B, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, NT5C2, CNOT1, ZDHHC5, GIGYF2, GATAD2A, NSD3, THAP11, PITPNM2, ACTR5, CENPT, BCL2L12, ZBTB37, MARS2, TNFRSF13C, ASPG
GO:0050793	Any process that modulates the frequency, rate or extent of development, the biological process whose specific outcome is the progression of a multicellular organism over time from an initial condition (e.g. a zygote, or a young adult) to a later condition (e.g. a multicellular animal or an aged adult).	BP	Turquoise	regulation of developmental process	4.70E-08	BTG1, CHRNA3, EMX1, F2, FGFR1, PRMT1, MEF2C, NFATC3, PRKD1, PSMA4, PTN, SOX5, TAOK2, ZEB2, EPN2, RIMS1, CNOT1, NGEF, PLEKHO1, NDRG4, BCL2L12, ZNF804A, TNFRSF13C, CNTN4
GO:0051252	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways involving RNA.	BP	Turquoise	regulation of RNA metabolic process	5.30E-08	RERE, BTG1, CLU, EMX1, FGFR1, PRMT1, MEF2C, NFATC3, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, CNOT1, GIGYF2, RBFOX1, GATAD2A, NSD3, THAP11, ACTR5, ESRP2, CENPT, BCL2L12, ZBTB37
GO:0005654	That part of the nuclear content other than the chromosomes or the nucleolus.	CC	Turquoise	nucleoplasm	6.26E-08	RERE, BNIP3L, BTG1, PRMT1, MEF2C, MSRA, NFATC3, PSMA4, ATXN7, SREBF1, TLE3, DMTF1, PPP1R13B, MAU2, EDC4, PLA2G15, ARL6IP4, GATAD2A, NSD3, FANCL, THAP11, CENPM, ACTR5, ESRP2, CENPT, ANP32E, RPS19BP1
GO:0010604	Any process that increases the frequency, rate or extent of the chemical reactions and	BP	Turquoise	positive regulation of macromolecule	6.68E-08	RERE, CHRNA3, CLU, GSDME, F2, FGFR1, GRIN2A, PRMT1, MEF2C, NFATC3, PDE4B,

	pathways involving macromolecules, any molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.			metabolic process		PRKD1, SOX5, SREBF1, TAOK2, ATG13, ZEB2, DMTF1, RIMS1, CNOT1, GIGYF2, NDFIP2, NSD3, NDRG4, BCL2L12, ZNF804A, TNFRSF13C
GO:0033554	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating the organism is under stress. The stress is usually, but not necessarily, exogenous (e.g. temperature, humidity, ionizing radiation).	BP	Turquoise	cellular response to stress	7.77E-08	BNIP3L, CLU, PRMT1, MEF2C, MGAT3, MSRA, PRKD1, PSMA4, PTN, PTPRK, SREBF1, TAOK2, ATG13, ZEB2, CNOT1, GIGYF2, JKAMP, FANCL, AMBRA1, ACTR5, BCL2L12
GO:0090317	Any process that decreases the frequency, rate or extent of the directed movement of proteins within cells.	BP	Turquoise	negative regulation of intracellular protein transport	8.52E-08	RERE, BTG1, CLU, EMX1, FGFR1, PRMT1, MEF2C, NFATC3, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, CNOT1, EDC4, GIGYF2, ARL6IP4, RBFOX1, GATAD2A, NSD3, FANCL, THAP11, ACTR5, ESRP2, CENPT, BCL2L12, ZBTB37, MARS2, TDRD9, HARBI1
GO:0031325	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways by which individual cells transform chemical substances.	BP	Turquoise	positive regulation of cellular metabolic process	9.84E-08	RERE, BNIP3L, CHRNA3, CLU, GSDME, F2, FGFR1, GRIN2A, PRMT1, MEF2C, NFATC3, PRKD1, SOX5, SREBF1, TAOK2, ATG13, ZEB2, DMTF1, CNOT1, GIGYF2, NDFIP2, NSD3, AMBRA1, NDRG4, BCL2L12, TNFRSF13C
GO:0051128	Any process that modulates the frequency, rate or extent of a process involved in the formation, arrangement of constituent parts, or disassembly of cell structures, including the plasma membrane and any external encapsulating structures such as the cell wall and cell envelope.	BP	Turquoise	regulation of cellular component organization	1.16E-07	BNIP3L, BTG1, CHRNA3, CLU, F2, FGFR1, MEF2C, PRKD1, PTN, SREBF1, ALMS1, TAOK2, ATG13, ZEB2, RIMS1, CNOT1, PPP1R13B, NGEF, NSD3, NDRG4, ZNF804A, CARMIL2
GO:0006811	The directed movement of charged atoms or small charged molecules into, out of or within a cell, or between cells, by means	BP	Turquoise	ion transport	1.49E-07	CACNA1C, CACNA1D, CHRNA3, CHRNA5, CHRNB4, F2, FGFR1, GRIN2A, MEF2C, NMB, PDE4B, SLC12A4, KCNK7, RIMS1,

	of some agent such as a transporter or pore.					CA14, NDFIP2, PITPNM2, SLC39A8, EMB
GO:0009966	Any process that modulates the frequency, rate or extent of signal transduction.	BP	Turquoise	regulation of signal transduction	1.94E-07	CLU, CTNND1, GSDME, F2, FGFR1, GRIN2A, PRMT1, MEF2C, PRKD1, PSMA4, TLE3, TAOK2, RGS6, ZEB2, EPN2, RIMS1, CNOT1, PLCL2, PPP1R13B, NGEF, NDFIP2, RALGAPA2, NLGN4X, NDRG4, BCL2L12
GO:0006928	The directed, self-propelled movement of a cell or subcellular component without the involvement of an external agent such as a transporter or a pore.	BP	Turquoise	movement of cell or subcellular component	2.51E-07	RERE, BTG1, CACNA1C, CACNA1D, F2, FGFR1, MEF2C, MGAT3, PDE4B, PRKD1, PTN, PTPRK, TAOK2, ZEB2, PLEKHO1, HYDIN, NDRG4, ESAM, EMB, CARMIL2, CNTN4
GO:0050804	Any process that modulates the frequency or amplitude of synaptic transmission, the process of communication from a neuron to a target (neuron, muscle, or secretory cell) across a synapse. Amplitude, in this case, refers to the change in postsynaptic membrane potential due to a single instance of synaptic transmission.	BP	Turquoise	modulation of chemical synaptic transmission	2.71E-07	CACNA1D, CHRNA3, CHRNA5, CHRNB4, GRIN2A, MEF2C, PTN, RIMS1, PLCL2, NLGN4X, CNTN4
GO:0099192	A synapse formed by a cerebellar Golgi cell synapsing on to a cerebellar granule cell.	CC	Turquoise	cerebellar Golgi cell to granule cell synapse	2.76E-07	CACNA1D, CHRNA3, CHRNA5, CHRNB4, GRIN2A, MEF2C, PTN, RIMS1, PLCL2, NLGN4X, CNTN4
GO:0043412	The covalent alteration of one or more monomeric units in a polypeptide, polynucleotide, polysaccharide, or other biological macromolecule, resulting in a change in its properties.	BP	Turquoise	macromolecule modification	3.04E-07	BTG1, CHRNA3, CLU, GSDME, F2, FGFR1, PRMT1, MEF2C, MGAT3, MSRA, PRKD1, PSMA4, PTN, PTPRK, ATXN7, SREBF1, TAOK2, ATG13, ZEB2, PLCL2, ZDHHC5, NDFIP2, GATAD2A, NSD3, FANCL, NDRG4, ACTR5, TDRD9, ASPHD1
GO:0070025	Interacting selectively and non-covalently with carbon monoxide (CO).	MF	Turquoise	carbon monoxide binding	3.11E-07	CACNA1C, CACNA1D, CHRNA3, CLU, F2, FGFR1, GRIN2A, MEF2C, PDE4B, PRKD1, PTN, TAOK2, RGS6, ATG13, ZEB2, RIMS1, PLCL2, MAU2, NGEF, NDFIP2, AMBRA1, RALGAPA2, ANP32E, BCL2L12
GO:0060079	A process that leads to a temporary increase in postsynaptic potential due to the flow of positively charged ions into the	BP	Turquoise	excitatory postsynaptic potential	3.77E-07	CHRNA3, CHRNA5, CHRNB4, GRIN2A, MEF2C, RIMS1, NLGN4X

	postsynaptic cell. The flow of ions that causes an EPSP is an excitatory postsynaptic current (EPSC) and makes it easier for the neuron to fire an action potential.					
GO:0009059	The chemical reactions and pathways resulting in the formation of a macromolecule, any molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.	BP	Turquoise	macromolecule biosynthetic process	4.21E-07	RERE, BTG1, CLU, NCAN, EMX1, FGFR1, PRMT1, MEF2C, MGAT3, NFATC3, PRKD1, PSMA4, PTPRK, SOX5, SREBF1, TLE3, ZNF592, ZEB2, DMTF1, CNOT1, ZDHHC5, GIGYF2, GATAD2A, NSD3, THAP11, ACTR5, CENPT, BCL2L12, ZBTB37, MARS2, TNFRSF13C

Table 6.8: Gene Ontologies for Blue Module in Developmental Stage Two using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	9.38E-19	CLU, NCAN, CTNND1, CYP17A1, EPHX2, FGFR1, GRM3, HSPD1, HSPE1, MMP16, NAGA, NDUFA6, OPCML, PCCB, STAT6, TBX6, TLE3, INPP4B, DGKI, GABBR2, KDM4A, MPHOSPH9, VPS45, SATB2, PSD3, ZDHHC5, FOXP1, B3GAT1, LSM1, RBFOX1, RALGAPA2, SUGP1, NDRG4, ZFYVE21, GDPD3, COQ10B, DRC3, YPEL3, L3MBTL2, PLPP5, C12orf65, RPS19BP1, TNFRSF13C, RFTN2, LETM2, C16orf92, WBP2NL, RILPL2, YPEL4, MED19
GO:0043231	Organised structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Blue	intracellular membrane-bounded organelle	2.90E-12	CLU, NCAN, CTNND1, CYP17A1, EPHX2, FGFR1, HSPD1, HSPE1, MMP16, NAGA, NDUFA6, PCCB, STAT6, TBX6, TLE3, DGKI, KDM4A, MPHOSPH9, VPS45, SATB2, FOXP1, B3GAT1, LSM1, RBFOX1, RALGAPA2, SUGP1, NDRG4, COQ10B, YPEL3, L3MBTL2, C12orf65, RPS19BP1, LETM2, WBP2NL, YPEL4, MED19
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Blue	cytoplasm	1.74E-10	CLU, NCAN, CTNND1, CYP17A1, EPHX2, FGFR1, HSPD1, HSPE1, MMP16, NAGA, NDUFA6, PCCB, STAT6, INPP4B, DGKI, GABBR2, KDM4A, MPHOSPH9, VPS45, B3GAT1, LSM1, RBFOX1, RALGAPA2, NDRG4, ZFYVE21, GDPD3, COQ10B, DRC3, PLPP5, C12orf65, RPS19BP1, LETM2, WBP2NL, RILPL2
GO:0044237	The chemical reactions and pathways by which individual cells transform chemical substances.	BP	Blue	cellular metabolic process	6.66E-09	CLU, NCAN, CYP17A1, EPHX2, FGFR1, HSPD1, HSPE1, NAGA, NDUFA6, PCCB, STAT6, TBX6, TLE3, INPP4B, DGKI, KDM4A, SATB2, ZDHHC5, FOXP1, B3GAT1, LSM1, RBFOX1, SUGP1, NDRG4, GDPD3, COQ10B, L3MBTL2, PLPP5, C12orf65, TNFRSF13C, WBP2NL, MED19
GO:0044238	The chemical reactions and pathways involving those compounds which are formed as a part of the normal anabolic and catabolic processes. These	BP	Blue	primary metabolic process	2.27E-08	CLU, NCAN, CYP17A1, EPHX2, FGFR1, HSPD1, HSPE1, MMP16, NAGA, PCCB, STAT6, TBX6, TLE3, INPP4B, DGKI, KDM4A, SATB2, ZDHHC5, FOXP1, B3GAT1, LSM1, RBFOX1, SUGP1, NDRG4, GDPD3, L3MBTL2, PLPP5, C12orf65, TNFRSF13C, WBP2NL, MED19

	processes take place in most, if not all, cells of the organism.					
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Blue	cellular response to diacyl bacterial lipopeptide	3.35E-08	CLU, NCAN, CYP17A1, EPHX2, FGFR1, HSPD1, HSPE1, MMP16, NAGA, PCCB, STAT6, TBX6, TLE3, INPP4B, DGKI, KDM4A, SATB2, ZDHHC5, FOXP1, B3GAT1, LSM1, RBFOX1, SUGP1, NDRG4, GDPD3, COQ10B, L3MBTL2, PLPP5, C12orf65, TNFRSF13C, WBP2NL, MED19
GO:0016020	A lipid bilayer along with all the proteins and protein complexes embedded in it an attached to it.	CC	Blue	membrane	1.24E-07	CLU, CTNND1, CYP17A1, FGFR1, GRM3, HSPD1, HSPE1, MMP16, NDUFA6, OPCML, STAT6, DGKI, GABBR2, MPHOSPH9, VPS45, PSD3, ZDHHC5, B3GAT1, RALGAPA2, NDRG4, GDPD3, COQ10B, PLPP5, TNFRSF13C, RFTN2, LETM2, C16orf92, RILPL2
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Blue	protein binding	1.24E-07	MPPED2, CLU, CTNND1, EPHX2, FGFR1, HSPD1, HSPE1, NAGA, PCCB, STAT6, TBX6, TLE3, INPP4B, DGKI, GABBR2, KDM4A, VPS45, SATB2, PSD3, FOXP1, LSM1, RBFOX1, RALGAPA2, SUGP1, NDRG4, ZFYVE21, L3MBTL2, RPS19BP1, WBP2NL, RILPL2, MED19
GO:0043233	The internal volume enclosed by the membranes of a particular organelle; includes the volume enclosed by a single organelle membrane, e.g. endoplasmic reticulum lumen, or the volume enclosed by the innermost of the two lipid bilayers of an organelle envelope, e.g. nuclear lumen.	CC	Blue	organelle lumen	2.00E-07	CLU, NCAN, EPHX2, FGFR1, HSPD1, HSPE1, MMP16, PCCB, STAT6, TBX6, TLE3, DGKI, KDM4A, SATB2, FOXP1, SUGP1, YPEL3, L3MBTL2, C12orf65, RPS19BP1, YPEL4, MED19
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Blue	collagen V binding	2.00E-07	CLU, NCAN, EPHX2, FGFR1, HSPD1, HSPE1, MMP16, PCCB, STAT6, TBX6, TLE3, DGKI, KDM4A, SATB2, FOXP1, SUGP1, YPEL3, L3MBTL2, C12orf65, RPS19BP1, YPEL4, MED19

Table 6.9: Gene Ontologies for Brown Module in Developmental Stage Two using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	7.94E-15	BTG1, CACNA1C, CHRNA5, ETF1, F2, PTK2B, KCNJ13, NEK1, NRGN, MAPK3, TCF4, FXR1, ATP5MPL, KCNK7, NXPH4, IGSF9B, SMG6, PPP1R13B, MAU2, NGEF, AMBRA1, TSNAXIP1, NDUFA4L2, ADAMTSL3, AS3MT, SRR, DPEP3, BOLL, ZNF408, SETD6, CENPT, EFHD1, WDR73, CREB3L1, TMEM219, SNORC, PCNX3
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	2.07E-09	BTG1, CACNA1C, CHRNA5, ETF1, F2, PTK2B, KCNJ13, NEK1, NRGN, MAPK3, TCF4, FXR1, KCNK7, NXPH4, IGSF9B, SMG6, PPP1R13B, MAU2, NGEF, AMBRA1, BOLL, ZNF408, SETD6, CENPT, EFHD1, WDR73, CREB3L1, TMEM219
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Brown	cytoplasm	9.37E-09	BTG1, CACNA1C, ETF1, F2, PTK2B, NEK1, NRGN, MAPK3, FXR1, ATP5MPL, SMG6, PPP1R13B, NGEF, AMBRA1, TSNAXIP1, NDUFA4L2, AS3MT, SRR, DPEP3, BOLL, SETD6, CENPT, EFHD1, WDR73, CREB3L1, SNORC
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Brown	protein binding	2.41E-08	BTG1, CACNA1C, CHRNA5, ETF1, F2, PTK2B, NEK1, NRGN, MAPK3, TCF4, FXR1, NXPH4, IGSF9B, SMG6, PPP1R13B, MAU2, NGEF, AMBRA1, ADAMTSL3, SRR, BOLL, ZNF408, SETD6, CENPT, CREB3L1, TMEM219
GO:0016043	A process that results in the assembly, arrangement of constituent parts, or disassembly of a cellular component.	BP	Brown	cellular component organization	4.64E-08	BTG1, ETF1, F2, PTK2B, NEK1, MAPK3, TCF4, IGSF9B, SMG6, PPP1R13B, MAU2, NGEF, AMBRA1, SRR, DPEP3, SETD6, CENPT, EFHD1, WDR73, CREB3L1
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-	BP	Brown	cellular response to catecholamine stimulus	7.59E-08	BTG1, ETF1, F2, PTK2B, NEK1, MAPK3, TCF4, IGSF9B, SMG6, PPP1R13B, MAU2, NGEF, AMBRA1, SRR, DPEP3, SETD6, CENPT, EFHD1, WDR73, CREB3L1

	aminoethyl)benzene-1,2-diol] and derivatives formed by substitution.					
GO:0016020	A lipid bilayer along with all the proteins and protein complexes embedded in it an attached to it.	CC	Brown	membrane	1.22E-07	CACNA1C, CHRNA5, F2, PTK2B, KCNJ13, NRGN, MAPK3, FXR1, ATP5MPL, KCNK7, IGSF9B, PPP1R13B, NGEF, AMBRA1, NDUFA4L2, SRR, DPEP3, EFHD1, WDR73, CREB3L1, TMEM219, SNORC, PCNX3

Table 6.10: Gene Ontology for Green Module Stage Two using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Green	molecular function	2.31E-12	BNIP3L, CACNB2, CLCN3, ERCC4, PRMT1, HSPA9, NCK1, PPP2R3A, PPP4C, PTPRK, PSMD6, NUTF2, KAT5, CLP1, NEMP1, SF3B1, SEC11A, ZSCAN2, TSR1, ACTR5, THOC7, IMMP2L, RFT1, TYW5, TSNARE1, ASPHD1, HAPLN4
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Green	positive regulation of macrophage proliferation	7.16E-10	BNIP3L, CACNB2, CLCN3, ERCC4, PRMT1, HSPA9, NCK1, PPP4C, PTPRK, PSMD6, NUTF2, KAT5, CLP1, NEMP1, SF3B1, SEC11A, ZSCAN2, TSR1, ACTR5, THOC7, IMMP2L, RFT1, TYW5, TSNARE1, ASPHD1, HAPLN4
GO:1902644	NA	BP	Green	tertiary alcohol metabolic process	2.55E-09	ERCC4, PRMT1, NCK1, PPP2R3A, PPP4C, PSMD6, KAT5, CLP1, SF3B1, SEC11A, ACTR5
GO:0032991	A stable assembly of two or more macromolecules, i.e. proteins, nucleic acids, carbohydrates or lipids, in which at least one component is a protein and the constituent parts function together.	CC	Green	protein-containing complex	7.70E-09	CACNB2, ERCC4, PRMT1, NCK1, PPP2R3A, PPP4C, PSMD6, NUTF2, KAT5, CLP1, SF3B1, SEC11A, TSR1, ACTR5, THOC7, IMMP2L, TSNARE1
GO:0043170	The chemical reactions and pathways involving macromolecules, any molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.	BP	Green	macromolecule metabolic process	4.82E-08	BNIP3L, ERCC4, PRMT1, NCK1, PPP2R3A, PPP4C, PTPRK, PSMD6, NUTF2, KAT5, CLP1, SF3B1, SEC11A, ZSCAN2, TSR1, ACTR5, THOC7, IMMP2L, TYW5, ASPHD1
GO:0044237	The chemical reactions and pathways by	BP	Green	cellular	1.56E-	BNIP3L, ERCC4, PRMT1, HSPA9,

	which individual cells transform chemical substances.			metabolic process	07	NCK1, PPP2R3A, PPP4C, PTPRK, PSMD6, KAT5, CLP1, SF3B1, SEC11A, ZSCAN2, TSR1, ACTR5, THOC7, IMMP2L, TYW5, ASPHD1
GO:0033554	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating the organism is under stress. The stress is usually, but not necessarily, exogenous (e.g. temperature, humidity, ionizing radiation).	BP	Green	cellular response to stress	1.69E-07	BNIP3L, ERCC4, PRMT1, HSPA9, NCK1, PPP4C, PTPRK, PSMD6, KAT5, ACTR5, IMMP2L
GO:0043227	Organised structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Green	membrane-bounded organelle	1.88E-07	BNIP3L, CLCN3, ERCC4, PRMT1, HSPA9, NCK1, PPP4C, PTPRK, PSMD6, NUTF2, KAT5, CLP1, NEMP1, SF3B1, SEC11A, ZSCAN2, TSR1, ACTR5, THOC7, IMMP2L, RFT1
GO:0005622	The living contents of a cell; the matter contained within (but not including) the plasma membrane, usually taken to exclude large vacuoles and masses of secretory or ingested material. In eukaryotes it includes the nucleus and cytoplasm.	CC	Green	intracellular	2.20E-07	BNIP3L, CLCN3, ERCC4, PRMT1, HSPA9, NCK1, PPP2R3A, PPP4C, PTPRK, PSMD6, NUTF2, KAT5, CLP1, NEMP1, SF3B1, ZSCAN2, TSR1, ACTR5, THOC7, IMMP2L, TYW5, TSNARE1

Table 6.11: Gene Ontology for Turquoise Module Stage Two using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	1.10E-69	More than 50 overlapping genes
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.	MF	Turquoise	armadillo repeat domain binding	3.94E-28	More than 50 overlapping genes
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	4.57E-25	More than 50 overlapping genes
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl) benzene-1,2-diol] and derivatives formed by substitution.	BP	Turquoise	cellular response to catecholamine stimulus	8.10E-24	More than 50 overlapping genes
GO:0005829	The part of the cytoplasm that	CC	Turquoise	cytosol	1.76E-20	More than 50 overlapping genes

	does not contain organelles, but which does contain other particulate matter, such as protein complexes.					
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Turquoise	regulation of fertilization	5.50E-19	More than 50 overlapping genes
GO:1901679	The directed movement of nucleotide across a membrane.	BP	Turquoise	nucleotide transmembrane transport	1.48E-16	More than 50 overlapping genes
GO:1901666	Any process that activates or increases the frequency, rate or extent of NAD+ ADP-ribosyltransferase activity.	BP	Turquoise	positive regulation of NAD+ ADP-ribosyltransferase activity	1.48E-16	More than 50 overlapping genes
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Turquoise	collagen V binding	3.74E-15	More than 50 overlapping genes
GO:0120060	Any process that modulates the frequency, rate or extent of any gastric emptying process, the process in which the liquid and liquid-suspended solid contents of the stomach exit through the pylorus into the duodenum.	BP	Turquoise	regulation of gastric emptying	5.54E-15	ATP2A2, CALB2, CHRM3, CHRNA3, CHRNB4, DRD2, GPM6A, GRIN2A, LRP1, NMB, PDE4B, PRKCB, PTPRF, RANGAP1, STAR, ALMS1, AP3B2, DOC2A, CUL3, TAOK2, SDCCAG8, GIGYF2, PARD6A, PLEKHO1, HYDIN, NLGN4X, RPTOR, CPEB1, BCL11B, ZNF804A, EMB, SLC32A1, CARMIL2, CNTN4, HCN1
GO:0006996	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly	BP	Turquoise	organelle organization	6.90E-15	ALDOA, RERE, ATP2A2, EP300, LRP1, MEF2C, PPP2R2A, PRKCB, PRKD1, RELA, ATXN7, SIPA1, SREBF1, SREBF2, SRPK2, ALMS1, MAD1L1, CUL3, HIRIP3, ASH2L,

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

	of an organelle within a cell. An organelle is an organized structure of distinctive morphology and function. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes and the cytoskeleton. Excludes the plasma membrane.					INA, TAOK2, BAG4, CKAP5, ZEB2, AKT3, STAG1, SDCCAG8, CNOT1, DDHD2, PARD6A, SFMBT1, HYDIN, VPS13C, NSD3, MSL2, PAK6, CENPM, MAIP1, ANP32E, PCGF6, DNAJC19, CARMIL2, TOM1L2, SNX19, BTBD18
GO:1901555	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a paclitaxel stimulus.	BP	Turquoise	response to paclitaxel	1.74E-14	More than 50 overlapping genes
GO:0099551	Cell-cell signalling between presynapse and postsynapse, via the vesicular release and reception of neuropeptide molecules, that modulates the synaptic transmission properties of the synapse.	BP	Turquoise	trans-synaptic signalling by neuropeptide, modulating synaptic transmission	8.65E-14	ATP2A2, CACNA1D, CALB2, CHRM3, CHRNA3, CHRNB4, DRD2, GRIN2A, MEF2C, PLCL1, PRKCB, PTN, RELA, SLC12A4, STAR, DOC2A, RIMS1, PLCL2, NLGN4X, SLC32A1, CNTN4, STAC3
GO:0046872	Interacting selectively and non-covalently with any metal ion.	MF	Turquoise	metal ion binding	9.32E-14	RERE, ATP2A2, CA8, CACNA1D, CALB2, DPYD, EP300, GRIN2A, IREB2, LRP1, PDE4B, PLCB2, PRKCB, PRKD1, SHMT2, SRPK2, DOC2A, ASH2L, ZNF592, PLCH2, ZNF536, ZEB2, RAI1, NT5C2, RIMS1, DDHD2, ZC3H7B, CA14, MOB4, GATAD2A, NSD3, MSL2, ZNF823, OTUD7B, THAP11, PITPNM2, RBM26, DPEP2, CPEB1, BCL11B, PCGF6, ZBTB37, ZNF804A, OTOL1, STAC3, HARBI1
GO:0097178	The aggregation, arrangement and	BP		ruffle assembly	1.29E-13	More than 50 overlapping genes

	bonding together of a set of components to form a ruffle, a projection at the leading edge of a crawling cell; the protrusions are supported by a microfilament meshwork. The formation of ruffles (also called membrane ruffling) is thought to be controlled by a group of enzymes known as Rho GTPases, specifically RhoA, Rac1 and cdc42.		Turquoise			
GO:0098935	The directed movement of organelles or molecules along microtubules in dendrites.	BP	Turquoise	dendritic transport	4.88E-13	ATP2A2, CACNA1D, CALB2, CHRM3, CHRNA3, CHRNB4, DRD2, GRIN2A, MEF2C, PLCL1, PRKCB, PTN, SLC12A4, STAR, DOC2A, RIMS1, PLCL2, NLGN4X, SLC32A1, CNTN4, STAC3
GO:0099552	Cell-cell signalling between presynapse and postsynapse, via the release and reception of lipid molecules, that modulates the synaptic transmission properties of the synapse.	BP	Turquoise	trans-synaptic signalling by lipid, modulating synaptic transmission	6.04E-13	ATP2A2, CACNA1D, CALB2, CHRM3, CHRNA3, CHRNB4, DRD2, GRIN2A, MEF2C, PLCL1, PRKCB, PTN, SLC12A4, STAR, DOC2A, RIMS1, PLCL2, NLGN4X, SLC32A1, CNTN4, STAC3
GO:0070025	Interacting selectively and non-covalently with carbon monoxide (CO).	MF	Turquoise	carbon monoxide binding	7.03E-13	ARHGAP1, SERPINC1, SERPING1, CACNA1D, CHRNA3, DRD2, EP300, FSHB, GRIN2A, LRP1, MEF2C, NAB2, PDE4B, PLCB2, PLCL1, PPP2R2A, PRKCB, PRKD1, PTN, PTPRF, RANGAP1, RELA, SIPA1, TAOK2, BAG4, RGS6, TBC1D5, ZEB2, RABGAP1L, TAB1, RIMS1, PLCL2, PAK6, RPTOR, RBM26, ANP32E, DNAJC19, STAC3
GO:2000225	Any process that stops, prevents, or reduces the frequency, rate or	BP		negative regulation of testosterone biosynthetic	1.12E-12	RERE, DRD2, EMX1, EP300, FSHB, IREB2, MEF2C, NAB2, NFATC3, PRKCB, PRKD1,

	extent of testosterone biosynthetic process.		Turquoise	process		PSMA4, PSMB10, RELA, SHMT2, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, TAB1, RAI1, CNOT1, GIGYF2, SFMBT1, GATAD2A, NSD3, BANK1, ZNF823, PAK6, OTUD7B, THAP11, RPTOR, CPEB1, BCL11B, PCGF6, ZBTB37, BTBD18
GO:1901561	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a benomyl stimulus.	BP	Turquoise	response to benomyl	1.16E-12	More than 50 overlapping genes
GO:0005886	The membrane surrounding a cell that separates the cell from its external environment. It consists of a phospholipid bilayer and associated proteins.	CC	Turquoise	plasma membrane	2.53E-12	More than 50 overlapping genes
GO:0005654	That part of the nuclear content other than the chromosomes or the nucleolus.	CC	Turquoise	nucleoplasm	3.26E-12	RERE, ATP2A2, EP300, MEF2C, MSRA, NFATC3, PPP2R2A, PRKCB, PSMA4, PSMB10, RANGAP1, RELA, ATXN7, SREBF1, SREBF2, SRPK2, CDK2AP1, CUL3, ASH2L, INA, DMTF1, STAG1, TAB1, RAI1, EDC4, PLA2G15, ARL6IP4, SFMBT1, GATAD2A, NSD3, MSL2, THAP11, RPTOR, CPEB1, CENPM, ESRP2, ANP32E, PCGF6, ATPAF2, STAC3
GO:0035556	The process in which a signal is passed on to downstream components within the cell, which become activated themselves to further propagate the signal and finally trigger a change in the	BP		intracellular signal transduction	5.12E-12	ARHGAP1, ATP2A2, CA8, GSDME, DRD2, EP300, FHIT, GRIN2A, LRP1, MEF2C, NFATC3, PLCB2, PLCL1, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SIPA1, SRPK2, CUL3, TAOK2, BAG4, RGS6, PLCH2, ZEB2, AKT3, TAB1, CNOT1, PLCL2,

	function or state of the cell.		Turquoise			BANK1, PAK6, OTUD7B, RPTOR, PITPNM2, STAC3
GO:0071986	A eukaryotically conserved protein complex; in humans, it is comprised of LAMTOR1, LAMTOR2, LAMTOR3, LAMTOR4, and LAMTOR5. The complex is anchored to lipid rafts in late endosome membranes via LAMTOR1, constitutes a guanine nucleotide exchange factor (GEF) for the Rag GTPases.	CC	Turquoise	Ragulator complex	5.36E-12	More than 50 overlapping genes
GO:1901558	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a metformin stimulus.	BP	Turquoise	response to metformin	6.72E-12	ALDOA, RERE, DRD2, EMX1, EP300, FSHB, IREB2, MEF2C, NAB2, NFATC3, PDE4B, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, STAR, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, TAB1, RAI1, NT5C2, CNOT1, SFMBT1, GATAD2A, NSD3, ZNF823, PAK6, OTUD7B, THAP11, RPTOR, SCAF1, BCL11B, PCGF6, ZBTB37, BTBD18
GO:0043232	Organised structure of distinctive morphology and function, not bounded by a lipid bilayer membrane and occurring within the cell. Includes ribosomes, the cytoskeleton and chromosomes.	CC	Turquoise	intracellular non-membrane-bounded organelle	8.10E-12	ALDOA, CACNA1D, CALB2, EMX1, EP300, FHIT, MEF2C, MSRA, PDE4B, PSMA4, RANGAP1, RELA, ATXN7, SHMT2, SRPK2, ALMS1, MAD1L1, DOC2A, CUL3, HIRIP3, ASH2L, INA, TAOK2, CKAP5, ZEB2, STAG1, SDCCAG8, RIMS1, CNOT1, DDHD2, EDC4, SPATS2L, GIGYF2, PARD6A, ARL6IP4, HYDIN, GATAD2A, NSD3, PAK6, RPTOR, CPEB1,

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

						CENPM, ANP32E, ZNF804A, CARMIL2, HARBI1
GO:0120069	Any process that increases the frequency, rate or extent of any stomach fundus smooth muscle contraction.	BP	Turquoise	positive regulation of stomach fundus smooth muscle contraction	2.94E-11	RERE, CHRNA3, DRD2, EMX1, EP300, GPM6A, LRP1, MEF2C, PRKD1, PTN, PTPRF, ALMS1, TAOK2, BAG4, CKAP5, ZEB2, SDCCAG8, RIMS1, PLEKHO1, HYDIN, BCL11B, SEMA6D, ZNF804A, EMB, CARMIL2, CNTN4
GO:0090317	Any process that decreases the frequency, rate or extent of the directed movement of proteins within cells.	BP	Turquoise	negative regulation of intracellular protein transport	3.44E-11	RERE, DRD2, EMX1, EP300, FSHB, MEF2C, NAB2, NFATC3, PPP2R2A, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, SRPK2, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, TAB1, RAI1, CNOT1, ZC3H7B, EDC4, GIGYF2, ARL6IP4, SFMBT1, GATAD2A, NSD3, ZNF823, PAK6, OTUD7B, THAP11, RPTOR, SCAF1, RBM26, CPEB1, BCL11B, ESRP2, PCGF6, ZBTB37, HARBI1, BTBD18
GO:0097707	A programmed cell death characterized morphologically by the presence of smaller than normal mitochondria with condensed mitochondrial membrane densities, reduction or vanishing of mitochondria crista, and outer mitochondrial membrane rupture. Activation of mitochondrial voltage-dependent anion channels and mitogen-activated protein kinases,	BP	Turquoise	ferroptosis	4.67E-11	RERE, DRD2, EMX1, EP300, FSHB, MEF2C, NAB2, NFATC3, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, TAB1, RAI1, CNOT1, SFMBT1, GATAD2A, NSD3, ZNF823, PAK6, OTUD7B, THAP11, RPTOR, SCAF1, BCL11B, PCGF6, ZBTB37, BTBD18

	upregulation of endoplasmic reticulum stress, and inhibition of cystine/glutamate antiporter are involved in the induction of ferroptosis. This process is characterized by the accumulation of lipid peroxidation products and lethal reactive oxygen species (ROS) derived from iron metabolism. Glutathione peroxidase 4 (GPX4), heat shock protein beta-1, and nuclear factor erythroid 2-related factor 2 function as negative regulators of ferroptosis by limiting ROS production and reducing cellular iron uptake, respectively. In contrast, NADPH oxidase and p53 act as positive regulators of ferroptosis by promotion of ROS production and inhibition of expression of SLC7A11 (a specific light-chain subunit of the cystine/glutamate antiporter), respectively. Misregulated ferroptosis has been implicated in multiple physiological and pathological processes.					
GO:0099192	A synapse formed by a cerebellar Golgi cell synapsing on to a cerebellar granule cell.	CC	Turquoise	cerebellar Golgi cell to granule cell synapse	4.68E-11	ATP2A2, CACNA1D, CALB2, CHRNA3, CHRNB4, DRD2, GRIN2A, MEF2C, PLCL1, PRKCB, PTN, STAR, RIMS1, PLCL2, NLGN4X, CNTN4
GO:1903593	Any process that modulates the frequency, rate or extent of histamine secretion by mast cell.	BP		regulation of histamine secretion by mast cell	6.09E-11	RERE, DRD2, EMX1, EP300, FSHB, MEF2C, NAB2, NFATC3, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, TAB1, RAI1,

			Turquoise			CNOT1, SFMBT1, GATAD2A, NSD3, ZNF823, PAK6, OTUD7B, THAP11, RPTOR, BCL11B, PCGF6, ZBTB37, BTBD18
GO:2001234	Any process that stops, prevents, or reduces the frequency, rate or extent of apoptotic signalling pathway.	BP	Turquoise	negative regulation of apoptotic signalling pathway	6.32E-11	RERE, DRD2, EMX1, EP300, FSHB, MEF2C, NAB2, NFATC3, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, TAB1, RAI1, CNOT1, SFMBT1, GATAD2A, NSD3, ZNF823, PAK6, OTUD7B, THAP11, RPTOR, BCL11B, PCGF6, ZBTB37, BTBD18
GO:0030054	A cellular component that forms a specialized region of connection between two or more cells or between a cell and the extracellular matrix. At a cell junction, anchoring proteins extend through the plasma membrane to link cytoskeletal proteins in one cell to cytoskeletal proteins in neighbouring cells or to proteins in the extracellular matrix.	CC	Turquoise	cell junction	8.62E-11	ATP2A2, CACNA1D, CALB2, CHRM3, CHRNA3, CHRNB4, DRD2, GPM6A, GRIN2A, LRP1, MEF2C, PDE4B, PPP2R2A, PRKCB, PTN, RELA, DOC2A, INA, SDCCAG8, RIMS1, PARD6A, PAK6, NLGN4X, CPEB1, ESAM, ZNF804A, EMB, SLC32A1
GO:0033554	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating the organism is	BP		cellular response to stress	1.64E-10	ATP2A2, EP300, LRP1, MEF2C, MGAT3, MSRA, PRKD1, PSMA4, PSMB10, PTN, PTPRF, RELA, SIPA1, SREBF1, SREBF2, CUL3, ASH2L, TAOK2, BAG4, ZEB2, AKT3, TAB1, CNOT1, GIGYF2, JKAMP, VPS13C, PAK6, RPTOR, CPEB1

	under stress. The stress is usually, but not necessarily, exogenous (e.g. temperature, humidity, ionizing radiation).		Turquoise			
GO:0006357	Any process that modulates the frequency, rate or extent of transcription mediated by RNA polymerase II.	BP	Turquoise	regulation of transcription by RNA polymerase II	2.45E-10	RERE, DRD2, EMX1, EP300, FSHB, MEF2C, NFATC3, PRKCB, PRKD1, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, RAI1, CNOT1, GATAD2A, ZNF823, OTUD7B, THAP11, BCL11B, PCGF6, ZBTB37, BTBD18
GO:0051172	Any process that stops, prevents, or reduces the frequency, rate or extent of the chemical reactions and pathways involving nitrogen or nitrogenous compounds.	BP	Turquoise	negative regulation of nitrogen compound metabolic process	3.20E-10	SERPINC1, RERE, SERPING1, DRD2, EP300, FHIT, GRIN2A, IREB2, MEF2C, MGAT3, NAB2, NFATC3, PTN, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF536, ZEB2, CNOT1, GIGYF2, PARD6A, SFMBT1, GATAD2A, BANK1, OTUD7B, THAP11, RPTOR, CPEB1, PCGF6
GO:0030425	A neuron projection that has a short, tapering, morphology. Dendrites receive and integrate signals from other neurons or from sensory stimuli, and conduct nerve impulses towards the axon or the cell body. In most neurons, the impulse is conveyed from dendrites to axon via the cell body, but in some types of unipolar neuron, the impulse does not travel via the cell body.	CC	Turquoise	dendrite	3.31E-10	CALB2, CHRM3, CHRNA3, DRD2, GPM6A, GRIN2A, LRP1, PDE4B, RANGAP1, TAOK2, GIGYF2, NLGN4X, RPTOR, CPEB1, ZNF804A, SLC32A1, HCN1
GO:0097473	Any apoptotic process in a retinal	BP		retinal rod cell apoptotic	3.48E-10	CALB2, CHRM3, CHRNA3, DRD2, GPM6A,

	rod cell, one of the two photoreceptor cell types of the vertebrate retina.		Turquoise	process		GRIN2A, LRP1, PDE4B, RANGAP1, TAOK2, GIGYF2, NLGN4X, RPTOR, CPEB1, ZNF804A, SLC32A1, HCN1
GO:0071516	The initial formation of a stable single-strand DNA lesion that triggers programmed gene conversion at the mating-type locus, thereby restricting mating-type interconversion to one of the two sister chromatids during DNA replication.	BP	Turquoise	establishment of imprinting at mating-type locus	6.63E-10	CHRM3, CHRNA3, DRD2, EP300, FSHB, LRP1, MEF2C, PDE4B, PRKCB, PTN, RANGAP1, RELA, SHMT2, SOX5, SREBF1, STAR, BAG4, TAB1, CNOT1, PARD6A, RPTOR, CPEB1, ESRP2, HCN1
GO:0019899	Interacting selectively and non-covalently with any enzyme.	MF	Turquoise	enzyme binding	8.13E-10	ARHGAP1, SERPINC1, ATP2A2, FHIT, LRP1, MEF2C, PPP2R2A, PRKCB, PTN, RANGAP1, RELA, SLC12A4, SREBF1, CDK2AP1, CUL3, TAOK2, BAG4, TBC1D5, RABGAP1L, TAB1, RIMS1, PARD6A, JKAMP, BANK1, PAK6, RPTOR, PITPNM2, SCAF1, TOM1L2
GO:0070901	The posttranscriptional addition of methyl groups to specific residues in a mitochondrial tRNA molecule.	BP	Turquoise	mitochondrial tRNA methylation	1.75E-09	ATP2A2, CHRM3, CHRNA3, GSDME, DRD2, EP300, FSHB, LRP1, MEF2C, MGAT3, MSRA, PDE4B, PRKCB, PRKD1, PSMA4, PSMB10, PTN, RANGAP1, RELA, SHMT2, SIPA1, SOX5, SREBF1, STAR, CUL3, INA, BAG4, TAB1, CNOT1, PARD6A, RPTOR, CPEB1, ESRP2, CPNE8, HCN1
GO:2000134	Any cell cycle regulatory process that prevents the commitment of a cell from G1 to S phase of the mitotic cell cycle.	BP		negative regulation of G1/S transition of mitotic cell cycle	2.39E-09	CHRNA3, DRD2, EMX1, EP300, FSHB, LRP1, MEF2C, NFATC3, PRKCB, PRKD1, PSMA4, PSMB10, PTN, PTPRF, RELA, SOX5, STAR, ASH2L, ZNF536, ZEB2, AKT3, RIMS1, PARD6A, SFMBT1, BCL11B, SEMA6D, ZNF804A, CNTN4

			Turquoise			
GO:0016477	The controlled self-propelled movement of a cell from one site to a destination guided by molecular cues. Cell migration is a central process in the development and maintenance of multicellular organisms.	BP	Turquoise	cell migration	4.23E-09	RERE, DRD2, FSHB, GPM6A, LRP1, MEF2C, MGAT3, PDE4B, PRKD1, PTN, PTPRF, CUL3, SLC7A6, TAOK2, BAG4, ZEB2, AKT3, SDCCAG8, PLEKHO1, PAK6, BCL11B, SEMA6D, ESAM, CARMIL2
GO:0016021	The component of a membrane consisting of the gene products and protein complexes having at least some part of their peptide sequence embedded in the hydrophobic region of the membrane.	CC	Turquoise	integral component of membrane	4.24E-09	ATP2A2, CACNA1D, CHRM3, CHRNA3, CHRNB4, DRD2, GPM6A, GRIN2A, LRP1, MGAT3, PDE4B, PTPRF, SLC12A4, SREBF1, SREBF2, CACNA1I, SLC7A6, GPR52, TAOK2, ABCB9, VSIG2, CA14, GIGYF2, SEZ6L2, SLC45A1, TMX2, AIG1, JKAMP, TM6SF2, CNNM2, NLGN4X, SLC39A8, CSMD1, GPR135, SEMA6D, TLCD3B, ESAM, SFXN5, DNAJC19, EMB, SLC32A1, KIAA1324L, STAC3, HCN1
GO:0050767	Any process that modulates the frequency, rate or extent of neurogenesis, the generation of cells in the nervous system.	BP	Turquoise	regulation of neurogenesis	5.08E-09	CHRNA3, DRD2, EMX1, EP300, LRP1, MEF2C, PRKD1, PTN, PTPRF, RELA, STAR, ZNF536, ZEB2, RIMS1, BCL11B, SEMA6D, ZNF804A, CNTN4
GO:0071727	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a triacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to triacyl bacterial lipopeptide	7.84E-09	ARHGAP1, CACNA1D, CHRNA3, DRD2, LRP1, PRKCB, RANGAP1, SREBF1, SREBF2, AP3B2, SLC7A6, TAOK2, BAG4, TBC1D5, CKAP5, RABGAP1L, RIMS1, DOP1A, ABCB9, PARD6A, VPS13C, BANK1, RBM26, MAIP1, SFXN5, DNAJC19, SLC32A1, TOM1L2, SNX19
GO:0031175	The process whose specific outcome is the progression of a	BP		neuron projection development	9.95E-09	RERE, CHRNA3, DRD2, EMX1, EP300, GPM6A, LRP1, MEF2C, PRKD1, PTN,

	neuron projection over time, from its formation to the mature structure. A neuron projection is any process extending from a neural cell, such as axons or dendrites (collectively called neurites).		Turquoise			PTPRF, TAOK2, ZEB2, RIMS1, BCL11B, SEMA6D, ZNF804A, EMB, CNTN4
GO:1901890	Any process that activates or increases the frequency, rate or extent of cell junction assembly.	BP	Turquoise	positive regulation of cell junction assembly	1.58E-08	CHRM3, CHRNA3, DRD2, GRIN2A, LRP1, MEF2C, PDE4B, PPP2R2A, PRKCB, PTN, RANGAP1, RELA, SHMT2, SREBF1, SREBF2, STAR, JKAMP, RPTOR, CPEB1, HCN1
GO:0051254	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways involving RNA.	BP	Turquoise	positive regulation of RNA metabolic process	1.58E-08	RERE, DRD2, EP300, FSHB, MEF2C, NFATC3, PRKCB, PRKD1, RELA, SOX5, SREBF1, SREBF2, ASH2L, ZEB2, DMTF1, STAG1, RAI1, CNOT1, GIGYF2, NSD3, RPTOR, CPEB1, BCL11B, BTBD18
GO:0101020	Catalysis of the reaction: oestrogen + donor-H2 + O2 = 16-alpha-hydroxyestrogen + H2O.	MF	Turquoise	oestrogen 16-alpha-hydroxylase activity	1.71E-08	ATP2A2, CACNA1D, CHRNA3, CHRNB4, DRD2, MEF2C, PRKCB, DOC2A, RIMS1, SLC32A1
GO:0070852	A neuron projection that is found in unipolar neurons and corresponds to the region between the cell body and the point at which the single projection branches.	CC	Turquoise	cell body fibre	2.40E-08	ATP2A2, CACNA1D, DRD2, GPM6A, GRIN2A, PDE4B, PLCB2, PRKCB, CACNA1I, PLCH2, CNNM2, SLC39A8, MAIP1, STAC3
GO:0098631	The binding by a cell-adhesion protein on the cell surface to an extracellular matrix component, to mediate adhesion of the cell to the external substrate or to another cell.	MF		cell adhesion mediator activity	2.43E-08	ARHGAP1, ATP2A2, CHRNB4, DRD2, LRP1, MGAT3, PRKD1, SLC12A4, SREBF1, SREBF2, AP3B2, DOC2A, CUL3, TBC1D5, TAB1, DOP1A, ABCB9, MOB4, TM6SF2, VPS13C, RPTOR, GPR135, TLCD3B, BORCS7, SLC32A1, SNX19

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

			Turquoise			
GO:0071329	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a sucrose stimulus.	BP	Turquoise	cellular response to sucrose stimulus	2.52E-08	CHRM3, CHRNA3, GSDME, DRD2, EP300, FSHB, LRP1, MEF2C, PDE4B, PRKCB, PRKD1, PSMA4, PSMB10, PTN, RANGAP1, RELA, SHMT2, SOX5, SREBF1, STAR, CUL3, INA, BAG4, TAB1, CNOT1, PARD6A, RPTOR, CPEB1, ESRP2, HCN1
GO:0072528	The chemical reactions and pathways resulting in the formation of a pyrimidine-containing compound, i.e. any compound that contains pyrimidine or a formal derivative thereof.	BP	Turquoise	pyrimidine-containing compound biosynthetic process	2.86E-08	ATP2A2, CACNA1D, DRD2, GPM6A, GRIN2A, PDE4B, PLCB2, PRKCB, CACNA1I, PLCH2, CNNM2, SLC39A8, MAIP1, STAC3
GO:0044248	The chemical reactions and pathways resulting in the breakdown of substances, carried out by individual cells.	BP	Turquoise	cellular catabolic process	2.95E-08	DPYD, DRD2, EP300, FHIT, LRP1, MGAT3, PDE4B, PPP2R2A, PRKD1, PSMA4, PSMB10, SHMT2, SREBF1, SREBF2, CUL3, TBC1D5, NT5C2, CNOT1, DDHD2, EDC4, PLA2G15, GIGYF2, AIG1, JKAMP, VPS13C, OTUD7B, RPTOR
GO:0098808	Interacting selectively and non-covalently with a 7-methylguanosine (m7G) group or derivative located at the 5' end of an mRNA molecule.	MF	Turquoise	mRNA cap binding	3.23E-08	CACNA1D, CALB2, CHRM3, DRD2, GPM6A, GRIN2A, PDE4B, PRKCB, PTN, DOC2A, RIMS1, NLGN4X, ZNF804A, SLC32A1
GO:0006468	The process of introducing a phosphate group on to a protein.	BP		protein phosphorylation	5.64E-08	CHRNA3, GSDME, DRD2, LRP1, MEF2C, NAB2, PLCL1, PRKCB, PRKD1, PSMA4, PSMB10, SRPK2, CDK2AP1, CUL3, TAOK2, BAG4, ZEB2, AKT3, TAB1, PLCL2, PARD6A, BANK1, PAK6, RPTOR, ALPK3

			Turquoise			
GO:0007420	The process whose specific outcome is the progression of the brain over time, from its formation to the mature structure. Brain development begins with patterning events in the neural tube and ends with the mature structure that is the centre of thought and emotion. The brain is responsible for the coordination and control of bodily activities and the interpretation of information from the senses (sight, hearing, smell, etc.).	BP	Turquoise	brain development	6.79E-08	RERE, DRD2, EMX1, GRIN2A, LRP1, PTN, STAR, INA, ZEB2, AKT3, HYDIN, NLGN4X, BCL11B, SEMA6D, SLC32A1, CNTN4
GO:1901678	The directed movement of an iron coordination entity into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore.	BP	Turquoise	iron coordination entity transport	6.97E-08	ALDOA, CTRL, DPYD, FHIT, GRIN2A, LRP1, MGAT3, PDE4B, PLCB2, PPP2R2A, PSMA4, PSMB10, RELA, SHMT2, CUL3, PLCH2, NT5C2, CNOT1, DDHD2, EDC4, PLA2G15, GIGYF2, AIG1, JKAMP, OTUD7B, ASPG
GO:0031399	Any process that modulates the frequency, rate or extent of the covalent alteration of one or more amino acid residues within a protein.	BP	Turquoise	regulation of protein modification process	8.87E-08	CHRNA3, GSDME, DRD2, EP300, LRP1, NAB2, PLCL1, PPP2R2A, PRKD1, PTN, RELA, SREBF1, CDK2AP1, CUL3, TAOK2, BAG4, ZEB2, TAB1, PLCL2, PARD6A, NSD3, BANK1, PAK6, RPTOR
GO:0140244	Any process that regulates translation occurring at the presynapse.	BP	Turquoise	regulation of translation at presynapse	1.07E-07	RERE, EMX1, EP300, MEF2C, NAB2, NFATC3, PRKCB, RELA, SOX5, SREBF1, SREBF2, ZNF592, ZNF536, ZEB2, DMTF1, STAG1, RAI1, SFMBT1, GATAD2A, NSD3, ZNF823, THAP11, BCL11B, PCGF6, ZBTB37
GO:0010558	Any process that decreases the	BP		negative regulation of	1.21E-07	RERE, EP300, IREB2, MEF2C, NAB2,

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

	rate, frequency or extent of the chemical reactions and pathways resulting in the formation of a macromolecule, any molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.		Turquoise	macromolecule biosynthetic process		NFATC3, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF536, ZEB2, CNOT1, GIGYF2, SFMBT1, GATAD2A, BANK1, OTUD7B, THAP11, CPEB1, PCGF6
GO:0043168	Interacting selectively and non-covalently with anions, charged atoms or groups of atoms with a net negative charge.	MF	Turquoise	anion binding	1.21E-07	SERPINC1, ATP2A2, GSDME, DPYD, LRP1, PDE4B, PLCL1, PRKCB, PRKD1, PTN, PTPRF, RELA, SHMT2, SRPK2, DOC2A, TAOK2, AKT3, PLCL2, ABCB9, PLA2G15, CNNM2, PAK6, NLGN4X, ALPK3, PITPNM2, CPNE8, CARMIL2, HCN1, SNX19
GO:0055065	Any process involved in the maintenance of an internal steady state of metal ions within an organism or cell.	BP	Turquoise	metal ion homeostasis	1.27E-07	ATP2A2, CACNA1D, CALB2, DRD2, GRIN2A, IREB2, LRP1, NMB, PLCB2, PRKCB, SLC12A4, PLCH2, CNNM2, SLC39A8, MAIP1
GO:0004620	Catalysis of the hydrolysis of a glycerophospholipid.	MF	Turquoise	phospholipase activity	1.37E-07	CHRM3, PLCB2, PLCL1, PLCH2, PLCL2, DDHD2, PLA2G15, ASPG
GO:0010647	Any process that increases the frequency, rate or extent of cell communication. Cell communication is the process that mediates interactions between a cell and its surroundings. Encompasses interactions such as signalling or attachment between one cell and another cell, between a cell and an extracellular matrix, or between a cell and any other	BP		positive regulation of cell communication	1.41E-07	ARHGAP1, CACNA1D, CALB2, CHRNB4, GSDME, DRD2, EP300, GRIN2A, LRP1, NMB, PRKCB, PRKD1, PSMA4, PSMB10, PTN, RELA, TAOK2, BAG4, ZEB2, AKT3, TAB1, RIMS1, BANK1, RPTOR

	aspect of its environment.		Turquoise			
GO:0120071	Any process that modulates the frequency, rate or extent of any pyloric antrum smooth muscle contraction.	BP	Turquoise	regulation of pyloric antrum smooth muscle contraction	1.44E-07	RERE, CHRNA3, DRD2, EMX1, GPM6A, LRP1, PTN, TAOK2, ZEB2, RIMS1, PLEKHO1, BCL11B, SEMA6D, EMB, CNTN4
GO:0023056	Any process that activates, maintains or increases the frequency, rate or extent of a signalling process.	BP	Turquoise	positive regulation of signalling	1.51E-07	ARHGAP1, CACNA1D, CALB2, CHRNB4, GSDME, DRD2, EP300, GRIN2A, LRP1, NMB, PRKCB, PRKD1, PSMA4, PSMB10, PTN, RELA, TAOK2, BAG4, ZEB2, AKT3, TAB1, RIMS1, BANK1, RPTOR
GO:0048858	The process in which the anatomical structures of a cell projection are generated and organized.	BP	Turquoise	cell projection morphogenesis	1.55E-07	RERE, CHRNA3, DRD2, EMX1, GPM6A, LRP1, PTN, TAOK2, ZEB2, RIMS1, PLEKHO1, BCL11B, SEMA6D, EMB, CNTN4
GO:0046928	Any process that modulates the frequency, rate or extent of the regulated release of a neurotransmitter from a cell.	BP	Turquoise	regulation of neurotransmitter secretion	1.56E-07	ATP2A2, CACNA1D, CHRNA3, CHRNB4, DRD2, MEF2C, PRKCB, RIMS1
GO:0010628	Any process that increases the frequency, rate or extent of gene expression. Gene expression is the process in which a gene's coding sequence is converted into a mature gene product or products (proteins or RNA). This includes the production of an RNA transcript as well as any processing to produce a mature RNA product or an mRNA or circRNA (for protein-coding genes) and the translation of that mRNA or circRNA into protein. Protein maturation is included	BP		positive regulation of gene expression	1.73E-07	DRD2, EP300, FSHB, LRP1, MEF2C, NFATC3, PDE4B, PRKD1, RELA, SOX5, SREBF1, SREBF2, SRPK2, STAR, ASH2L, ZEB2, DMTF1, STAG1, RAI1, RIMS1, NSD3, RPTOR, CPEB1, BCL11B, ZNF804A, BTBD18

	when required to form an active form of a product from an inactive precursor form.		Turquoise			
GO:2000226	Any process that modulates the frequency, rate or extent of pancreatic A cell differentiation.	BP	Turquoise	regulation of pancreatic A cell differentiation	2.32E-07	EP300, IREB2, MEF2C, NAB2, NFATC3, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF536, ZEB2, CNOT1, GIGYF2, SFMBT1, GATAD2A, BANK1, OTUD7B, THAP11, CPEB1, PCGF6
GO:0097733	A specialised 9+0 non-motile cilium found in photoreceptor cells. A ciliary transition zone called 'photoreceptor connecting cilium' links the photoreceptor outer segment to the inner segment.	CC	Turquoise	photoreceptor cell cilium	2.32E-07	ALDOA, ARHGAP1, SERPING1, CHRNB4, DRD2, GRIN2A, LRP1, PDE4B, SIPA1, SREBF1, SREBF2, AP3B2, DOC2A, TAOK2, TBC1D5, RABGAP1L, TAB1, DOP1A, DDHD2, GIGYF2, GPR135, ANP32E, SLC32A1, CARMIL2, EHBP1L1, SNX19
GO:0007166	A series of molecular signals initiated by activation of a receptor on the surface of a cell. The pathway begins with binding of an extracellular ligand to a cell surface receptor, or for receptors that signal in the absence of a ligand, by ligand-withdrawal or the activity of a constitutively active receptor. The pathway ends with regulation of a downstream cellular process, e.g. transcription.	BP	Turquoise	cell surface receptor signalling pathway	2.76E-07	CHRNA3, CHRNB4, DRD2, EP300, FSHB, GRIN2A, LRP1, MEF2C, NFATC3, PDE4B, PLCB2, PRKCB, PRKD1, PSMA4, PSMB10, PTN, PTPRF, RELA, SREBF1, CUL3, BAG4, ZEB2, TAB1, RIMS1, PLCL2, GIGYF2, PARD6A, NLGN4X, ESRP2, SEMA6D
GO:1901894	Any process that modulates the frequency, rate or extent of an ATPase-coupled calcium transmembrane transporter activity	BP	Turquoise	regulation of calcium-transporting ATPase activity	3.17E-07	CHRM3, CHRNA3, DRD2, LRP1, MEF2C, PDE4B, PRKCB, RANGAP1, RELA, SHMT2, SREBF1, STAR, RPTOR, CPEB1, HCN1

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

Table 6.12: Gene Ontologies for Blue Module Stage Three using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	2.37E-12	BNIP3L, CLU, GRM3, HSPD1, HSPE1, MAP3K11, MMP16, NAGA, PCCB, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, GABBR2, VPS45, SATB2, FOXP1, TM6SF2, RBFOX1, ZSCAN2, TSR1, NDUFA4L2, PAK6, RALGAPA2, BCL11B, NDRG4, ZFYVE21, GDPD3, DRC3, YPEL3, C16orf92, MED19
GO:0043231	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Blue	intracellular membrane-bounded organelle	1.70E-10	BNIP3L, CLU, HSPD1, HSPE1, MMP16, NAGA, PCCB, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, VPS45, SATB2, FOXP1, TM6SF2, RBFOX1, ZSCAN2, TSR1, NDUFA4L2, PAK6, RALGAPA2, BCL11B, NDRG4, YPEL3, MED19
GO:0044238	The chemical reactions and pathways involving those compounds which are formed as a part of the normal anabolic and catabolic processes. These processes take place in most, if not all, cells of the organism.	BP	Blue	primary metabolic process	1.03E-09	BNIP3L, CLU, HSPD1, HSPE1, MAP3K11, MMP16, NAGA, PCCB, PPP2R3A, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, SATB2, FOXP1, TM6SF2, RBFOX1, ZSCAN2, TSR1, PAK6, BCL11B, NDRG4, GDPD3, MED19
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Blue	intracellular organelle lumen	3.05E-09	BNIP3L, CLU, GRM3, HSPD1, HSPE1, MAP3K11, MMP16, PPP2R3A, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, GABBR2, VPS45, SATB2, FOXP1, TM6SF2, RBFOX1, ZSCAN2, PAK6, RALGAPA2, BCL11B, NDRG4, YPEL3, MED19
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Blue	cellular response to diacyl bacterial lipopeptide	6.83E-09	BNIP3L, CLU, HSPD1, HSPE1, MAP3K11, MMP16, NAGA, PCCB, PPP2R3A, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, SATB2, FOXP1, TM6SF2, RBFOX1, ZSCAN2, TSR1, PAK6, BCL11B, NDRG4, GDPD3, MED19

GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Blue	cytoplasm	1.69E-08	BNIP3L, CLU, HSPD1, HSPE1, MAP3K11, MMP16, NAGA, PCCB, PRKCB, STAR, STAT6, FXR1, DGKI, GABBR2, VPS45, TM6SF2, RBFOX1, TSR1, NDUFA4L2, PAK6, RALGAPA2, NDRG4, ZFYVE21, GDPD3, DRC3
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Blue	protein binding	4.22E-08	BNIP3L, CLU, HSPD1, HSPE1, MAP3K11, NAGA, PCCB, PPP2R3A, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, GABBR2, VPS45, SATB2, FOXP1, RBFOX1, PAK6, RALGAPA2, BCL11B, NDRG4, ZFYVE21, MED19
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Blue	regulation of fertilization	6.19E-08	CLU, HSPD1, HSPE1, MAP3K11, PPP2R3A, PRKCB, STAR, STAT6, TBX6, FXR1, SATB2, FOXP1, TM6SF2, RBFOX1, ZSCAN2, PAK6, BCL11B, NDRG4, MED19
GO:0031325	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways by which individual cells transform chemical substances.	BP	Blue	positive regulation of cellular metabolic process	8.68E-08	BNIP3L, CLU, HSPD1, HSPE1, MAP3K11, PRKCB, STAR, STAT6, TBX6, FXR1, SATB2, PAK6, BCL11B, NDRG4, MED19
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Blue	collagen V binding	2.27E-07	BNIP3L, CLU, HSPD1, HSPE1, MMP16, PCCB, PRKCB, STAR, STAT6, TBX6, FXR1, DGKI, SATB2, FOXP1, TSR1, PAK6, YPEL3, MED19
GO:0006807	The chemical reactions and pathways involving organic or inorganic compounds that contain nitrogen.	BP	Blue	nitrogen compound metabolic process	2.46E-07	BNIP3L, CLU, HSPD1, HSPE1, MAP3K11, MMP16, NAGA, PCCB, PPP2R3A, PRKCB, STAT6, TBX6, FXR1, SATB2, FOXP1, RBFOX1, ZSCAN2, TSR1, PAK6, BCL11B, NDRG4, GDPD3, MED19

Table 6.13: Gene Ontologies for Brown Module Stage Three using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	4.06E-08	CHRNA2, CHRNA5, ETF1, PTK2B, TCF4, ATP5MPL, PSMD6, KCNK7, IGSF9B, SMG6, MAU2, NGEF, AMBRA1, AS3MT, BOLL, ZNF408, SETD6, THOC7, TMEM219, ASPHD1
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	8.00E-06	CHRNA2, CHRNA5, ETF1, PTK2B, TCF4, PSMD6, KCNK7, IGSF9B, SMG6, MAU2, NGEF, AMBRA1, BOLL, ZNF408, SETD6, TMEM219
GO:0099582	Any neurotransmitter receptor activity that is involved in regulating the concentration of calcium in the presynaptic cytosol.	MF	Brown	neurotransmitter receptor activity involved in regulation of presynaptic cytosolic calcium ion concentration	1.25E-05	CHRNA2, CHRNA5, PTK2B, IGSF9B
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Brown	protein binding	2.88E-05	CHRNA5, ETF1, PTK2B, TCF4, PSMD6, IGSF9B, SMG6, MAU2, NGEF, AMBRA1, BOLL, ZNF408, SETD6, THOC7, TMEM219
GO:0060078	Any process that modulates the potential difference across a post-synaptic membrane.	BP	Brown	regulation of postsynaptic membrane potential	3.17E-05	CHRNA2, CHRNA5, PTK2B, IGSF9B

Table 6.14: Gene Ontology for Turquoise Stage Three using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	5.64E-77	More than 50 overlapping genes
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl)benzene-1,2-diol] and derivatives formed by substitution.	BP	Turquoise	cellular response to catecholamine stimulus	4.05E-29	More than 50 overlapping genes
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	1.01E-23	More than 50 overlapping genes
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.	MF	Turquoise	armadillo repeat domain binding	2.25E-18	More than 50 overlapping genes
GO:0005829	The part of the cytoplasm that does not contain organelles, but which does contain other particulate matter, such as protein complexes.	CC	Turquoise	cytosol	7.69E-18	More than 50 overlapping genes

GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Turquoise	collagen V binding	1.91E-15	More than 50 overlapping genes
GO:0033554	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating the organism is under stress. The stress is usually, but not necessarily, exogenous (e.g. temperature, humidity, ionizing radiation).	BP	Turquoise	cellular response to stress	4.35E-14	ATP2A2, EP300, LRP1, MEF2C, MSRA, NEK1, MAPK3, PSMA4, PSMB10, PTN, PTPRF, RELA, SIPA1, SREBF1, SREBF2, XRCC3, CUL3, ASH2L, TAOK2, BAG4, ATG13, AKT3, TAB1, CNOT1, B3GAT1, VPS13C, FANCL, RPTOR, CPEB1, ACD, ACTR5, IMMP2L, CREB3L1, INO80E
GO:0043232	Organised structure of distinctive morphology and function, not bounded by a lipid bilayer membrane and occurring within the cell. Includes ribosomes, the cytoskeleton and chromosomes.	CC	Turquoise	intracellular non-membrane-bounded organelle	5.38E-14	CACNA1C, EMX1, EP300, FGFR1, MEF2C, MSRA, NEK1, MAPK3, PSMA4, RELA, SHMT2, SRPK2, XRCC3, ALMS1, MAD1L1, DOC2A, CUL3, HIRIP3, ASH2L, INA, TAOK2, MPHOSPH9, STAG1, SDCCAG8, RIMS1, CNOT1, DDHD2, SPATS2L, LSM1, ARL6IP4, HYDIN, GATAD2A, NSD3, DDX28, RPTOR, CPEB1, ACD, CENPM, ACTR5, CREB3L1, C12orf65, RPS19BP1, ZNF804A, TDRD9, CARMIL2, WBP2NL, RILPL2, YPEL4, HARBI1, INO80E
GO:0097178	The aggregation, arrangement and bonding together of a set of components to form a ruffle, a projection at the leading edge of a crawling cell; the protrusions are supported by a microfilament meshwork. The formation of ruffles (also called membrane ruffling) is thought to be controlled by a group of enzymes known as Rho GTPases, specifically RhoA, Rac1 and cdc42.	BP	Turquoise	ruffle assembly	7.53E-14	More than 50 overlapping genes
GO:1901666	Any process that activates or increases the frequency, rate or extent of NAD+	BP	Turquoise	positive regulation of NAD+ ADP-ribosyltransferase	1.35E-13	More than 50 overlapping genes

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

	ADP-ribosyltransferase activity.			activity		
GO:1901679	The directed movement of nucleotide across a membrane.	BP	Turquoise	nucleotide transmembrane transport	1.45E-13	More than 50 overlapping genes
GO:1901561	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a benomyl stimulus.	BP	Turquoise	response to benomyl	1.71E-13	More than 50 overlapping genes
GO:1901555	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a paclitaxel stimulus.	BP	Turquoise	response to paclitaxel	1.71E-13	More than 50 overlapping genes
GO:0005654	That part of the nuclear content other than the chromosomes or the nucleolus.	CC	Turquoise	nucleoplasm	2.85E-13	RERE, ATP2A2, EP300, MEF2C, MSRA, NFATC3, PPP2R2A, MAPK3, PSMA4, PSMB10, RELA, SREBF1, SREBF2, SRPK2, XRCC3, CUL3, ASH2L, INA, DMTF1, NUTF2, STAG1, TAB1, RAI1, CLP1, PPP1R13B, PLA2G15, ARL6IP4, GATAD2A, NSD3, FANCL, MSL2, THAP11, RPTOR, SUGP1, CPEB1, ACD, CENPM, ACTR5, RPS19BP1, ATPAF2, STAC3, INO80E
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Turquoise	regulation of fertilization	1.39E-12	More than 50 overlapping genes
GO:0010604	Any process that increases the frequency, rate or extent of the chemical reactions and pathways involving macromolecules, any	BP	Turquoise	positive regulation of macromolecule metabolic process	1.59E-12	RERE, GSDME, DRD2, EP300, F2, FGFR1, FSHB, GRIN2A, LRP1, MEF2C, NAB2, NEK1, NFATC3, MAPK3, RELA, SOX5, SREBF1, SREBF2, SRPK2, CUL3, ASH2L,

	molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.					TAOK2, BAG4, ATG13, TBC1D5, DMTF1, STAG1, TAB1, RAI1, RIMS1, CNOT1, NDFIP2, NSD3, RPTOR, CPEB1, ACD, CREB3L1, ZNF804A, TNFRSF13C, WBP2NL, BTBD18
GO:0003824	Catalysis of a biochemical reaction at physiological temperatures. In biologically catalysed reactions, the reactants are known as substrates, and the catalysts are naturally occurring macromolecular substances known as enzymes. Enzymes possess specific binding sites for substrates, and are usually composed wholly or largely of protein, but RNA that has catalytic activity (ribozyme) is often also regarded as enzymatic.	MF	Turquoise	catalytic activity	1.69E-12	More than 50 overlapping genes
GO:0120060	Any process that modulates the frequency, rate or extent of any gastric emptying process, the process in which the liquid and liquid-suspended solid contents of the stomach exit through the pylorus into the duodenum.	BP	Turquoise	regulation of gastric emptying	2.44E-12	ATP2A2, CACNA1C, CHRNB4, DRD2, GPM6A, GRIN2A, LRP1, NMB, NRGN, MAPK3, PTPRF, ALMS1, AP3B2, DOC2A, CUL3, TAOK2, SDCCAG8, PSD3, ZDHHC5, LSM1, PLEKHO1, HYDIN, RPTOR, CPEB1, ZNF804A, EMB, SLC32A1, CARMIL2, CNTN4, WBP2NL, RILPL2, HCN1
GO:0048513	Development of a tissue or tissues that work together to perform a specific function or functions. Development pertains to the process whose specific outcome is the progression of a structure over time, from its formation to the mature structure. Organs are commonly observed as visibly distinct structures but may also exist as loosely associated clusters of cells that work	BP	Turquoise	animal organ development	2.56E-12	SERPINC1, RERE, CACNA1C, GSDME, DRD2, EMX1, EP300, FGFR1, FSHB, GPM6A, GRIN2A, IREB2, LRP1, MEF2C, NAB2, NRGN, PGM3, MAPK3, PSMA4, PSMB10, PTN, RELA, SOX5, SREBF1, MAD1L1, CUL3, ASH2L, INA, AKAP6, AKT3, TAB1, CLP1, HYDIN, ALPK3, IMMP2L, DNAJC19, SLC32A1, CNTN4, STAC3, HCN1, SNORC

	together to perform a specific function or functions.					
GO:0090317	Any process that decreases the frequency, rate or extent of the directed movement of proteins within cells.	BP	Turquoise	negative regulation of intracellular protein transport	5.06E-12	More than 50 overlapping genes
GO:0031325	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways by which individual cells transform chemical substances.	BP	Turquoise	positive regulation of cellular metabolic process	5.77E-12	RERE, GSDME, DRD2, EP300, F2, FGFR1, FSHB, GRIN2A, LRP1, MEF2C, NAB2, NEK1, NFATC3, MAPK3, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, TAOK2, BAG4, ATG13, TBC1D5, DMTF1, STAG1, TAB1, RAI1, CNOT1, NDFIP2, VPS13C, NSD3, RPTOR, CPEB1, ACD, CREB3L1, TNFRSF13C, WBP2NL, BTBD18
GO:0051173	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways involving nitrogen or nitrogenous compounds.	BP	Turquoise	positive regulation of nitrogen compound metabolic process	3.82E-11	RERE, GSDME, DRD2, EP300, F2, FGFR1, FSHB, GRIN2A, LRP1, MEF2C, NAB2, NEK1, NFATC3, MAPK3, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, TAOK2, BAG4, ATG13, DMTF1, STAG1, TAB1, RAI1, CNOT1, NDFIP2, NSD3, RPTOR, CPEB1, ACD, CREB3L1, TNFRSF13C, WBP2NL, BTBD18
GO:0071986	A eukaryotically conserved protein complex; in humans, it is comprised of LAMTOR1, LAMTOR2, LAMTOR3, LAMTOR4, and LAMTOR5. The complex is anchored to lipid rafts in late endosome membranes via LAMTOR1, constitutes a guanine nucleotide exchange factor (GEF) for the Rag GTPases.	CC	Turquoise	Ragulator complex	4.75E-11	SERPINC1, ATP2A2, CACNA1C, CHRNB4, CLCN3, GSDME, DRD2, F2, FGFR1, GPM6A, GRIN2A, LRP1, NRGN, MAPK3, PTN, PTPRF, CACNA1I, SLC7A6, GPR52, AKAP6, BAG4, RGS6, PLCH2, TBC1D5, SNAP91, RIMS1, PSD3, PPP1R13B, VSIG2, CA14, ZDHHC5, SEZ6L2, PLEKHO1, MPP6, SLC39A8, DPEP3, CPEB1, GPR135, ESAM, ZNF804A, TNFRSF13C, EMB, SLC32A1, CPNE8, CARMIL2, CNTN4, STAC3, HARBI1, HCN1, SNORC
GO:0035556	The process in which a signal is passed on to downstream components within the cell, which become activated themselves to further propagate the	BP	Turquoise	intracellular signal transduction	5.12E-11	ATP2A2, CA8, CACNA1C, GSDME, DRD2, EP300, F2, FGFR1, GRIN2A, LRP1, MEF2C, NEK1, NFATC3, MAPK3, PSMA4, PSMB10, RELA, SIPA1, SRPK2, CUL3,

	signal and finally trigger a change in the function or state of the cell.					TAOK2, AKAP6, BAG4, RGS6, PLCH2, AKT3, TAB1, CNOT1, PSD3, PPP1R13B, NDFIP2, OTUD7B, RPTOR, CREB3L1, STAC3
GO:0046872	Interacting selectively and non-covalently with any metal ion.	MF	Turquoise	metal ion binding	7.72E-11	RERE, ATP2A2, CA8, CACNA1C, DPYD, EP300, F2, GRIN2A, IREB2, LRP1, NEK1, PGM3, SHMT2, SRPK2, DOC2A, ASH2L, ZNF592, PLCH2, ZNF536, RAI1, NT5C2, RIMS1, DDHD2, ZC3H7B, CA14, MOB4, B3GAT1, GATAD2A, NSD3, FANCL, MSL2, ZNF823, OTUD7B, THAP11, DPEP2, DPEP3, CPEB1, ZBTB37, ZNF804A, YPEL4, STAC3, HARBI1
GO:0005886	The membrane surrounding a cell that separates the cell from its external environment. It consists of a phospholipid bilayer and associated proteins.	CC	Turquoise	plasma membrane	8.47E-11	SERPINC1, ATP2A2, CACNA1C, CHRNB4, CLCN3, GSDME, DRD2, F2, FGFR1, GPM6A, GRIN2A, LRP1, NRGN, MAPK3, PTN, PTPRF, CACNA1I, SLC7A6, GPR52, AKAP6, BAG4, RGS6, PLCH2, TBC1D5, SNAP91, RIMS1, PSD3, PPP1R13B, VSIG2, CA14, ZDHHC5, SEZ6L2, PLEKHO1, MPP6, SLC39A8, DPEP3, CPEB1, GPR135, ESAM, ZNF804A, TNFRSF13C, EMB, SLC32A1, CPNE8, CARMIL2, CNTN4, STAC3, HARBI1, HCN1
GO:1901558	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a metformin stimulus.	BP	Turquoise	response to metformin	2.46E-10	RERE, DRD2, EMX1, EP300, FGFR1, FSHB, IREB2, MEF2C, NAB2, NFATC3, PGM3, MAPK3, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, DMTF1, STAG1, TAB1, RAI1, CLP1, NT5C2, CNOT1, GATAD2A, NSD3, ZNF823, OTUD7B, THAP11, RPTOR, SCAF1, ACD, ACTR5, ZBTB37, CREB3L1, WBP2NL, BTBD18
GO:0006464	The covalent alteration of one or more amino acids occurring in proteins, peptides and nascent polypeptides (co-translational, post-translational	BP	Turquoise	cellular protein modification process	2.46E-10	SERPINC1, GSDME, DRD2, EP300, F2, FGFR1, LRP1, MEF2C, MSRA, NAB2, NEK1, PGM3, PPP2R2A, MAPK3, PSMA4, PSMB10, PTN, PTPRF, RELA, SHMT2,

	modifications) occurring at the level of an individual cell. Includes the modification of charged tRNAs that are destined to occur in a protein (pre-translation modification).					SREBF1, SRPK2, CUL3, ASH2L, TAOK2, BAG4, ATG13, AKT3, TAB1, ZDHHC5, B3GAT1, NDFIP2, NSD3, FANCL, MSL2, OTUD7B, RPTOR, ALPK3, ACTR5, HS3ST5, INO80E
GO:0050793	Any process that modulates the frequency, rate or extent of development, the biological process whose specific outcome is the progression of a multicellular organism over time from an initial condition (e.g. a zygote, or a young adult) to a later condition (e.g. a multicellular animal or an aged adult).	BP	Turquoise	regulation of developmental process	2.46E-10	DRD2, EMX1, EP300, F2, FGFR1, FSHB, LRP1, MEF2C, NAB2, NFATC3, PSMA4, PSMB10, PTN, PTPRF, RELA, SOX5, ASH2L, TAOK2, AKAP6, ZNF536, AKT3, RAI1, EPN2, RIMS1, CNOT1, DDHD2, LSM1, PLEKHO1, TLCD3B, CREB3L1, ZNF804A, TNFRSF13C, CNTN4
GO:0098631	The binding by a cell-adhesion protein on the cell surface to an extracellular matrix component, to mediate adhesion of the cell to the external substrate or to another cell.	MF	Turquoise	cell adhesion mediator activity	2.85E-10	ATP2A2, CHRNB4, CLCN3, DRD2, LRP1, NRGN, SREBF1, SREBF2, AP3B2, DOC2A, CUL3, AKAP6, TBC1D5, MPHOSPH9, NUTF2, TAB1, EPN2, DOP1A, ABCB9, MOB4, B3GAT1, NDFIP2, VPS13C, RPTOR, GPR135, TLCD3B, BORCS7, SLC32A1, HS3ST5
GO:0097707	A programmed cell death characterized morphologically by the presence of smaller than normal mitochondria with condensed mitochondrial membrane densities, reduction or vanishing of mitochondria crista, and outer mitochondrial membrane rupture. Activation of mitochondrial voltage-dependent anion channels and mitogen-activated protein kinases, upregulation of endoplasmic reticulum stress, and inhibition of cystine/glutamate antiporter are involved in the induction of ferroptosis. This process is	BP	Turquoise	ferroptosis	4.03E-10	RERE, DRD2, EMX1, EP300, FGFR1, FSHB, MEF2C, NAB2, NFATC3, MAPK3, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, DMTF1, STAG1, TAB1, RAI1, CLP1, CNOT1, GATAD2A, NSD3, ZNF823, OTUD7B, THAP11, RPTOR, SCAF1, ACTR5, ZBTB37, CREB3L1, WBP2NL, BTBD18

	characterized by the accumulation of lipid peroxidation products and lethal reactive oxygen species (ROS) derived from iron metabolism. Glutathione peroxidase 4 (GPX4), heat shock protein beta-1, and nuclear factor erythroid 2-related factor 2 function as negative regulators of ferroptosis by limiting ROS production and reducing cellular iron uptake, respectively. In contrast, NADPH oxidase and p53 act as positive regulators of ferroptosis by promotion of ROS production and inhibition of expression of SLC7A11 (a specific light-chain subunit of the cystine/glutamate antiporter), respectively. Misregulated ferroptosis has been implicated in multiple physiological and pathological processes.					
GO:0003676	Interacting selectively and non-covalently with any nucleic acid.	MF	Turquoise	nucleic acid binding	5.31E-10	RERE, EMX1, EP300, IREB2, LRP1, MEF2C, NFATC3, RELA, SOX5, SREBF1, SREBF2, SRPK2, XRCC3, BAG4, ZNF592, ZNF536, DMTF1, R3HDM2, RIMS1, CNOT1, ZC3H7B, SPATS2L, LSM1, ARL6IP4, GATAD2A, ZNF823, DDX28, OTUD7B, THAP11, SUGP1, SCAF1, CPEB1, ACD, ZBTB37, CREB3L1, C12orf65, RPS19BP1, ZNF804A, TDRD9, WBP2NL, SELENOH
GO:2000225	Any process that stops, prevents, or reduces the frequency, rate or extent of testosterone biosynthetic process.	BP	Turquoise	negative regulation of testosterone biosynthetic process	7.81E-10	RERE, DRD2, EMX1, EP300, FGFR1, FSHB, IREB2, MEF2C, NAB2, NFATC3, MAPK3, PSMA4, PSMB10, RELA, SHMT2, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, DMTF1, STAG1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, ZNF823, OTUD7B, THAP11, RPTOR, CPEB1, ACD,

						ACTR5, ZBTB37, CREB3L1, BTBD18
GO:0051130	Any process that activates or increases the frequency, rate or extent of a process involved in the formation, arrangement of constituent parts, or disassembly of cell structures, including the plasma membrane and any external encapsulating structures such as the cell wall and cell envelope.	BP	Turquoise	positive regulation of cellular component organization	1.03E-09	DRD2, EP300, FGFR1, GPM6A, LRP1, MAPK3, PTN, SREBF1, SREBF2, CUL3, BAG4, ATG13, TBC1D5, RIMS1, CNOT1, DDHD2, PPP1R13B, VPS13C, NSD3, ACD, ZNF804A, CARMIL2
GO:0120069	Any process that increases the frequency, rate or extent of any stomach fundus smooth muscle contraction.	BP	Turquoise	positive regulation of stomach fundus smooth muscle contraction	2.08E-09	RERE, DRD2, EMX1, EP300, FGFR1, GPM6A, LRP1, MEF2C, NEK1, MAPK3, PTN, PTPRF, ALMS1, TAOK2, BAG4, SDCCAG8, RIMS1, PLEKHO1, HYDIN, ZNF804A, EMB, CARMIL2, CNTN4, RILPL2
GO:1903593	Any process that modulates the frequency, rate or extent of histamine secretion by mast cell.	BP	Turquoise	regulation of histamine secretion by mast cell	2.25E-09	RERE, DRD2, EMX1, EP300, FGFR1, FSHB, MEF2C, NAB2, NFATC3, MAPK3, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, DMTF1, STAG1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, ZNF823, OTUD7B, THAP11, RPTOR, ACTR5, ZBTB37, CREB3L1, WBP2NL, BTBD18
GO:0016021	The component of a membrane consisting of the gene products and protein complexes having at least some part of their peptide sequence embedded in the hydrophobic region of the membrane.	CC	Turquoise	integral component of membrane	2.25E-09	ATP2A2, CACNA1C, CHRNB4, CLCN3, DRD2, FGFR1, GPM6A, GRIN2A, LRP1, PTPRF, SREBF1, SREBF2, CACNA1I, SLC7A6, GPR52, TAOK2, AKAP6, ABCB9, VSIG2, CA14, ZDHHC5, SEZ6L2, B3GAT1, SLC45A1, TMX2, AIG1, NDFIP2, SLC39A8, GPR135, TLCD3B, IMMP2L, ESAM, CREB3L1, SFXN5, TNFRSF13C, DNAJC19, EMB, SLC32A1, TSNARE1, KIAA1324L, HS3ST5, STAC3, HCN1, SNORC, PCNX3
GO:2001234	Any process that stops, prevents or reduces the frequency, rate or extent of	BP		negative regulation of apoptotic signalling pathway	2.32E-09	RERE, DRD2, EMX1, EP300, FGFR1, FSHB, MEF2C, NAB2, NFATC3, MAPK3,

	apoptotic signalling pathway.		Turquoise			PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, DMTF1, STAG1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, ZNF823, OTUD7B, THAP11, RPTOR, ACTR5, ZBTB37, CREB3L1, WBP2NL, BTBD18
GO:0071727	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) because of a triacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to triacyl bacterial lipopeptide	2.75E-09	CACNA1C, DRD2, LRP1, MAPK3, SREBF1, SREBF2, AP3B2, SLC7A6, TAOK2, AKAP6, BAG4, ATG13, TBC1D5, SNAP91, NUTF2, RIMS1, DOP1A, PPP1R13B, ABCB9, NDFIP2, VPS13C, ACD, MAIP1, IMMP2L, SFXN5, DNAJC19, SLC32A1, TOM1L2, RILPL2, TSNARE1
GO:0048699	The process in which nerve cells are generated. This includes the production of neuroblasts and their differentiation into neurons.	BP	Turquoise	generation of neurons	2.92E-09	RERE, GSDME, DRD2, EMX1, EP300, F2, FGFR1, GPM6A, LRP1, MEF2C, MAPK3, PTN, PTPRF, RELA, SOX5, TAOK2, ZNF536, SDCCAG8, RIMS1, LSM1, ZNF804A, EMB, CNTN4, HCN1
GO:0033043	Any process that modulates the frequency, rate or extent of a process involved in the formation, arrangement of constituent parts, or disassembly of an organelle.	BP	Turquoise	regulation of organelle organization	3.22E-09	LRP1, MEF2C, MAPK3, SREBF1, SREBF2, XRCC3, ALMS1, MAD1L1, CUL3, TAOK2, BAG4, ATG13, STAG1, SDCCAG8, CNOT1, DDHD2, PPP1R13B, VPS13C, NSD3, ACD, CARMIL2, TOM1L2
GO:0007049	The progression of biochemical and morphological phases and events that occur in a cell during successive cell replication or nuclear replication events. Canonically, the cell cycle comprises the replication and segregation of genetic material followed by the division of the cell, but in endocycles or syncytial cells nuclear replication or nuclear division may not be followed by cell division.	BP	Turquoise	cell cycle	4.58E-09	DRD2, EP300, FGFR1, NEK1, PPP2R2A, MAPK3, PSMA4, PSMB10, SIPA1, SRPK2, XRCC3, ALMS1, MAD1L1, CUL3, TAOK2, DMTF1, STAG1, SDCCAG8, CNOT1, PPP1R13B, RPTOR, DPEP3, TDRD9, TOM1L2, WBP2NL, BTBD18
GO:0030054	A cellular component that forms a specialized region of connection between two or more cells or between	CC		cell junction	4.95E-09	ATP2A2, CACNA1C, CHRNB4, CLCN3, DRD2, GPM6A, GRIN2A, LRP1, MEF2C, NRGN, PPP2R2A, MAPK3, PTN, RELA,

	a cell and the extracellular matrix. At a cell junction, anchoring proteins extend through the plasma membrane to link cytoskeletal proteins in one cell to cytoskeletal proteins in neighbouring cells or to proteins in the extracellular matrix.		Turquoise			DOC2A, INA, AKAP6, SNAP91, SDCCAG8, RIMS1, PSD3, CPEB1, ESAM, ZNF804A, EMB, SLC32A1
GO:0048468	The process whose specific outcome is the progression of the cell over time, from its formation to the mature structure. Cell development does not include the steps involved in committing a cell to a specific fate.	BP	Turquoise	cell development	5.13E-09	RERE, ATP2A2, DRD2, EMX1, EP300, F2, FGFR1, GPM6A, LRP1, MEF2C, MAPK3, PTN, PTPRF, RELA, CUL3, TAOK2, AKAP6, ZNF536, RIMS1, LSM1, HYDIN, ALPK3, ZNF804A, EMB, CNTN4, RILPL2, STAC3, HCN1
GO:0045202	The junction between a nerve fibre of one neuron and another neuron, muscle fibre or glial cell. As the nerve fibre approaches the synapse it enlarges into a specialized structure, the presynaptic nerve ending, which contains mitochondria and synaptic vesicles. At the tip of the nerve ending is the presynaptic membrane; facing it and separated from it by a minute cleft (the synaptic cleft) is a specialized area of membrane on the receiving cell, known as the postsynaptic membrane. In response to the arrival of nerve impulses, the presynaptic nerve ending secretes molecules of neurotransmitters into the synaptic cleft. These diffuse across the cleft and transmit the signal to the postsynaptic membrane.	CC	Turquoise	synapse	6.10E-09	ATP2A2, CACNA1C, CHRNB4, CLCN3, DRD2, GPM6A, GRIN2A, MEF2C, NRGN, PPP2R2A, PTN, RELA, DOC2A, INA, SNAP91, RIMS1, PSD3, CPEB1, ZNF804A, EMB, SLC32A1
GO:0034220	A process in which an ion is transported across a membrane.	BP		ion transmembrane transport	7.66E-09	ATP2A2, CACNA1C, CHRNB4, CLCN3, DRD2, F2, GPM6A, GRIN2A, MEF2C, CACNA1I, SLC7A6, AKAP6, PLCH2,

			Turquoise			ABCB9, SLC39A8, MAIP1, SFXN5, EMB, SLC32A1, STAC3, HCN1
GO:0006357	Any process that modulates the frequency, rate or extent of transcription mediated by RNA polymerase II.	BP	Turquoise	regulation of transcription by RNA polymerase II	1.01E-08	RERE, DRD2, EMX1, EP300, FGFR1, FSHB, MEF2C, NFATC3, MAPK3, PSMA4, PSMB10, RELA, SOX5, SREBF1, SREBF2, CUL3, ASH2L, ZNF592, ZNF536, DMTF1, STAG1, RAI1, CNOT1, GATAD2A, ZNF823, OTUD7B, THAP11, ZBTB37, CREB3L1, BTBD18
GO:0043168	Interacting selectively and non-covalently with anions, charged atoms or groups of atoms with a net negative charge.	MF	Turquoise	anion binding	1.03E-08	SERPINC1, ATP2A2, CLCN3, GSDME, DPYD, F2, FGFR1, LRP1, NEK1, NRGN, MAPK3, PTN, PTPRF, RELA, SHMT2, SRPK2, XRCC3, DOC2A, TAOK2, SNAP91, AKT3, CLP1, ABCB9, PLA2G15, DDX28, ALPK3, TDRD9, CPNE8, CARMIL2, HS3ST5, HCN1
GO:0060322	The biological process whose specific outcome is the progression of a head from an initial condition to its mature state. The head is the anterior-most division of the body.	BP	Turquoise	head development	2.76E-08	RERE, DRD2, EMX1, EP300, FGFR1, GRIN2A, LRP1, NRGN, MAPK3, PTN, INA, AKT3, CLP1, HYDIN, IMMP2L, SLC32A1, CNTN4
GO:0019899	Interacting selectively and non-covalently with any enzyme.	MF	Turquoise	enzyme binding	3.58E-08	SERPINC1, ATP2A2, LRP1, MEF2C, PPP2R2A, MAPK3, PTN, RELA, SREBF1, CUL3, TAOK2, AKAP6, BAG4, ATG13, TBC1D5, SNAP91, NUTF2, TAB1, RIMS1, PSD3, FANCL, RPTOR, SCAF1, ACD, RPS19BP1, TOM1L2, RILPL2
GO:0006928	The directed, self-propelled movement of a cell or subcellular component without the involvement of an external agent such as a transporter or a pore.	BP	Turquoise	movement of cell or subcellular component	3.63E-08	RERE, ATP2A2, CACNA1C, DRD2, F2, FGFR1, FSHB, GPM6A, LRP1, MEF2C, MAPK3, PTN, PTPRF, AP3B2, CUL3, CACNA1I, SLC7A6, TAOK2, BAG4, AKT3, SDCCAG8, PLEKHO1, HYDIN, ESAM, EMB, CARMIL2, CNTN4
GO:0010647	Any process that increases the frequency, rate or extent of cell communication. Cell communication is the process that mediates	BP		positive regulation of cell communication	4.57E-08	CHRNB4, GSDME, DRD2, EP300, F2, FGFR1, GRIN2A, LRP1, NMB, NRGN, MAPK3, PSMA4, PSMB10, PTN, RELA, TAOK2, AKAP6, BAG4, AKT3, TAB1,

	interactions between a cell and its surroundings. Encompasses interactions such as signalling or attachment between one cell and another cell, between a cell and an extracellular matrix, or between a cell and any other aspect of its environment.		Turquoise			EPN2, RIMS1, PPP1R13B, NDFIP2, RPTOR
GO:0070762	A subcomplex of the nuclear pore complex (NPC) that spans the nuclear membrane and anchors the NPC to the nuclear envelope. In S. cerevisiae, the transmembrane ring is composed of Pom152p, Pom34p, and Ndc1p. In vertebrates, it is composed of Gp210, Ndc1, and Pom121. Components are arranged in 8-fold symmetrical 'spokes' around the central transport channel. A single 'spoke', can be isolated and is sometime referred to as the Ndc1 complex.	CC	Turquoise	nuclear pore transmembrane ring	4.65E-08	F2, GRIN2A, LRP1, PTN, SREBF1, SREBF2, AP3B2, TAOK2, AKAP6, BAG4, ATG13, TBC1D5, NUTF2, RIMS1, PPP1R13B, VPS13C, OTUD7B, ACD, MAIP1, IMMP2L, DNAJC19, TOM1L2, RILPL2, TSNARE1, STAC3
GO:0042592	Any biological process involved in the maintenance of an internal steady state.	BP	Turquoise	homeostatic process	4.67E-08	ATP2A2, CACNA1C, CLCN3, DRD2, EMX1, F2, FSHB, GRIN2A, IREB2, LCAT, LRP1, MEF2C, NMB, MAPK3, SREBF2, XRCC3, ALMS1, AKAP6, PLCH2, AKT3, TMX2, SLC39A8, ACD, MAIP1, TNFRSF13C
GO:0023056	Any process that activates, maintains or increases the frequency, rate or extent of a signalling process.	BP	Turquoise	positive regulation of signalling	4.85E-08	CHRNB4, GSDME, DRD2, EP300, F2, FGFR1, GRIN2A, LRP1, NMB, NRGN, MAPK3, PSMA4, PSMB10, PTN, RELA, TAOK2, AKAP6, BAG4, AKT3, TAB1, EPN2, RIMS1, PPP1R13B, NDFIP2, RPTOR
GO:0070025	Interacting selectively and non-covalently with carbon monoxide (CO).	MF		carbon monoxide binding	6.02E-08	SERPINC1, CACNA1C, DRD2, EP300, F2, FGFR1, FSHB, GRIN2A, LRP1, MEF2C, NAB2, NEK1, PPP2R2A, MAPK3, PTN, PTPRF, RELA, SIPA1, TAOK2, AKAP6,

			Turquoise			BAG4, RGS6, ATG13, TBC1D5, TAB1, RIMS1, NDFIP2, RPTOR, ACD, DNAJC19, STAC3
GO:0040011	Self-propelled movement of a cell or organism from one location to another.	BP	Turquoise	locomotion	6.02E-08	RERE, DRD2, EP300, F2, FGFR1, FSHB, GPM6A, GRIN2A, LRP1, MEF2C, MAPK3, PTN, PTPRF, CUL3, CACNA1I, SLC7A6, TAOK2, BAG4, AKT3, SDCCAG8, PLEKHO1, ESAM, EMB, CARMIL2, CNTN4
GO:0097733	A specialised 9+0 non-motile cilium found in photoreceptor cells. A ciliary transition zone called 'photoreceptor connecting cilium' links the photoreceptor outer segment to the inner segment.	CC	Turquoise	photoreceptor cell cilium	8.68E-08	CHRNB4, CLCN3, DRD2, FGFR1, GRIN2A, LRP1, NRGN, MAPK3, SIPA1, SREBF1, SREBF2, AP3B2, DOC2A, TAOK2, TBC1D5, SNAP91, TAB1, EPN2, DOP1A, DDHD2, NDFIP2, DPEP3, GPR135, SLC32A1, CARMIL2, PHETA2, EHBP1L1
GO:0006796	The chemical reactions and pathways involving the phosphate group, the anion or salt of any phosphoric acid.	BP	Turquoise	phosphate-containing compound metabolic process	1.35E-07	GSDME, DRD2, F2, FGFR1, LCAT, LRP1, MEF2C, NAB2, NEK1, PGM3, PPP2R2A, MAPK3, PSMA4, PSMB10, PTN, PTPRF, SHMT2, SRPK2, CUL3, INPP4B, TAOK2, BAG4, PLCH2, ATG13, AKT3, TAB1, CLP1, NT5C2, DDHD2, PLA2G15, RPTOR, ALPK3
GO:0070901	The posttranscriptional addition of methyl groups to specific residues in a mitochondrial tRNA molecule.	BP	Turquoise	mitochondrial tRNA methylation	2.13E-07	ATP2A2, GSDME, DRD2, EP300, FGFR1, FSHB, LRP1, MEF2C, MSRA, MAPK3, PSMA4, PSMB10, PTN, RELA, SHMT2, SIPA1, SOX5, SREBF1, CUL3, INA, AKAP6, BAG4, TAB1, EPN2, CNOT1, B3GAT1, RPTOR, CPEB1, CREB3L1, TNFRSF13C, CPNE8, HCN1
GO:0030135	Small membrane-bounded organelle formed by pinching off of a coated region of membrane. Some coats are made of clathrin, whereas others are made from other proteins.	CC	Turquoise	coated vesicle	2.27E-07	LRP1, NRGN, SREBF1, SREBF2, AP3B2, TBC1D5, SNAP91, EPN2, DDHD2, SLC32A1, PHETA2
GO:0009056	The chemical reactions and pathways resulting in the breakdown of substances, including the breakdown	BP		catabolic process	2.37E-07	CTRL, DPYD, DRD2, EP300, GRIN2A, LRP1, PPP2R2A, MAPK3, PSMA4, PSMB10, RELA, SHMT2, SREBF1,

	of carbon compounds with the liberation of energy for use by the cell or organism.		Turquoise			SREBF2, CUL3, PLCH2, ATG13, TBC1D5, NT5C2, CNOT1, DDHD2, PLA2G15, LSM1, AIG1, NDFIP2, VPS13C, OTUD7B, RPTOR
GO:0000902	The developmental process in which the size or shape of a cell is generated and organized.	BP	Turquoise	cell morphogenesis	2.46E-07	RERE, DRD2, EMX1, EP300, F2, GPM6A, LRP1, MEF2C, MAPK3, PSMB10, PTN, CUL3, TAOK2, RIMS1, PLEKHO1, EMB, CNTN4, RILPL2

Table 6.15: Gene Ontology for Yellow Module Stage Three using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	GO Process	FDR	Genes
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Yellow	binding	3.20E-07	BTG1, CHRM3, CYP17A1, PRMT1, HSPA9, NCK1, PLCL1, PPP4C, TLE3, RABGAP1L, KAT5, SF3B1, GIGYF2, JKAMP, CNNM2, TYW5, LETM2
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Yellow	positive regulation of macrophage proliferation	3.44E-07	BTG1, CHRM3, CYP17A1, PRMT1, HSPA9, NCK1, PLCL1, PPP4C, TLE3, RABGAP1L, KAT5, SF3B1, SEC11A, GIGYF2, JKAMP, CNNM2, TYW5, LETM2
GO:0044237	The chemical reactions and pathways by which individual cells transform chemical substances.	BP	Yellow	cellular metabolic process	1.60E-06	BTG1, CHRM3, CYP17A1, PRMT1, HSPA9, NCK1, PLCL1, PPP4C, TLE3, KAT5, SF3B1, SEC11A, GIGYF2, JKAMP, TYW5
GO:0043227	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Yellow	membrane-bounded organelle	1.00E-05	BTG1, CYP17A1, PRMT1, HSPA9, NCK1, PPP4C, TLE3, RABGAP1L, KAT5, SF3B1, SEC11A, GIGYF2, JKAMP, CNNM2, LETM2

Table 6.16: Gene Ontology for Blue Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Blue	positive regulation of macrophage proliferation	1.70E-10	CHRM3, CYP17A1, GRM3, HSPD1, HSPE1, MMP16, NDUFA6, PCCB, STAR, STAT6, DGKI, VPS45, SATB2, VSIG2, MOB4, FOXP1, LSM1, RBFOX1, TSR1, PAK6, RALGAPA2, NDRG4, ZFYVE21, GDPD3, MAIP1, COQ10B, DRC3, YPEL3, C12orf65, RPS19BP1, RILPL2, YPEL4, EHBP1L1, ASPG
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Blue	cytoplasm	4.22E-09	CYP17A1, HSPD1, HSPE1, MMP16, NDUFA6, PCCB, STAR, STAT6, DGKI, VPS45, MOB4, LSM1, RBFOX1, TSR1, PAK6, RALGAPA2, NDRG4, ZFYVE21, GDPD3, MAIP1, COQ10B, DRC3, C12orf65, RPS19BP1, RILPL2, EHBP1L1, ASPG
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Blue	binding	4.46E-08	MPPED2, CHRM3, CYP17A1, HSPD1, HSPE1, MMP16, PCCB, STAR, STAT6, DGKI, VPS45, SATB2, MOB4, FOXP1, LSM1, RBFOX1, TSR1, PAK6, RALGAPA2, NDRG4, ZFYVE21, GDPD3, MAIP1, COQ10B, YPEL3, C12orf65, RPS19BP1, RILPL2, YPEL4
GO:0043231	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane and occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Blue	intracellular membrane-bounded organelle	1.57E-07	CYP17A1, HSPD1, HSPE1, MMP16, NDUFA6, PCCB, STAR, STAT6, DGKI, VPS45, SATB2, MOB4, FOXP1, LSM1, RBFOX1, TSR1, PAK6, RALGAPA2, NDRG4, MAIP1, COQ10B, YPEL3, C12orf65, RPS19BP1, YPEL4

Table 6.17: Gene Ontology for Brown Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	1.89E-08	BNIP3L, CA8, ERCC4, ETF1, FGFR1, SIPA1, SREBF1, ATP5MPL, GABBR2, TAB1, IGSF9B, ABCB9, HYDIN, CNNM2, GATAD2A, ZSCAN2, AMBRA1, NLGN4X, RANBP10, SETD6, CENPT, TLCD3B, ZNF804A, SFXN5, STAC3
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Brown	binding	4.85E-07	BNIP3L, CA8, ERCC4, ETF1, FGFR1, SIPA1, SREBF1, GABBR2, TAB1, IGSF9B, ABCB9, CNNM2, GATAD2A, ZSCAN2, AMBRA1, NLGN4X, RANBP10, SETD6, CENPT, ANKRD44, ZNF804A, STAC3
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Brown	intracellular organelle lumen	2.09E-06	BNIP3L, CA8, ERCC4, ETF1, FGFR1, SIPA1, SREBF1, GABBR2, TAB1, IGSF9B, CNNM2, GATAD2A, ZSCAN2, AMBRA1, NLGN4X, SETD6, CENPT, TLCD3B, ZNF804A, STAC3

Table 6.18: Gene Ontology for Green Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Green	positive regulation of macrophage proliferation	3.43E-06	CACNA1C, FSHB, MAP3K11, PTN, PTPRF, GPR52, TBC1D5, RAI1, NXPH4, EPN2, DOP1A, CA14, THAP11, ACTR5, CARMIL2, TOM1L2, PCNX3
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Green	intracellular organelle lumen	9.16E-05	CACNA1C, FSHB, MAP3K11, PTN, PTPRF, GPR52, TBC1D5, RAI1, NXPH4, EPN2, THAP11, ACTR5, CARMIL2, TOM1L2
GO:0035373	Interacting selectively and non-covalently with a chondroitin sulphate proteoglycan, any proteoglycan containing chondroitin sulphate as the glycosaminoglycan carbohydrate unit.	MF	Green	chondroitin sulphate proteoglycan binding	0.000174	PTN, PTPRF
GO:0050896	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus. The process begins with detection of the stimulus and ends with a change in state or activity or the cell or organism.	BP	Green	response to stimulus	0.000336	CACNA1C, FSHB, MAP3K11, PTN, PTPRF, GPR52, TBC1D5, NXPH4, EPN2, ACTR5, CARMIL2, TOM1L2

Table 6.19: Gene Ontology for Magenta Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Magenta	cytoplasm	7.51E-05	BTG1, PLCB2, ALMS1, FXR1, CDK2AP1, TAOK2, CLP1, TM6SF2, PLPP5, LETM2, WBP2NL
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Magenta	molecular function	0.000264	BTG1, PLCB2, ALMS1, FXR1, CDK2AP1, TAOK2, CLP1, TM6SF2, BCL2L12, PLPP5, LETM2, WBP2NL
GO:0019222	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways within a cell or an organism.	BP	Magenta	regulation of metabolic process	0.000537	BTG1, ALMS1, FXR1, CDK2AP1, TAOK2, CLP1, TM6SF2, BCL2L12, WBP2NL
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Magenta	positive regulation of macrophage proliferation	0.000585	BTG1, PLCB2, ALMS1, FXR1, CDK2AP1, TAOK2, CLP1, TM6SF2, BCL2L12, PLPP5, LETM2, WBP2NL
GO:0044238	The chemical reactions and pathways involving those compounds which are formed as a part of the normal anabolic and catabolic processes. These processes take place in most, if not all, cells of the organism.	BP	Magenta	primary metabolic process	0.000978	BTG1, PLCB2, FXR1, CDK2AP1, TAOK2, CLP1, TM6SF2, BCL2L12, PLPP5, WBP2NL
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Magenta	cellular response to diacyl bacterial lipopeptide	0.001976	BTG1, PLCB2, FXR1, CDK2AP1, TAOK2, CLP1, TM6SF2, BCL2L12, PLPP5, WBP2NL

Table 6.20: Gene Ontology for Purple Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Purple	positive regulation of macrophage proliferation	4.37E-10	CHRNA2, CHRNA3, CLU, FHIT, GRIN2A, HSPA9, NCK1, PPP4C, SLC12A4, VRK2, AP3B2, AKAP6, CKAP5, PSMD6, KAT5, PLCL2, SEC11A, B3GAT1, FANCL, SBNO1, OTUD7B, RBM26, SEMA6D, THOC7, WDR73, RFT1, TNFRSF13C, BORCS7, TYW5, ASPHD1, HAPLN4
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Purple	intracellular organelle lumen	6.00E-08	CHRNA2, CHRNA3, CLU, FHIT, GRIN2A, HSPA9, NCK1, PPP2R3A, PPP4C, SLC12A4, VRK2, AKAP6, CKAP5, PSMD6, KAT5, PLCL2, BANK1, FANCL, SBNO1, OTUD7B, RBM26, SEMA6D, WDR73, RFT1, TNFRSF13C
GO:1901666	Any process that activates or increases the frequency, rate or extent of NAD+ ADP-ribosyltransferase activity.	BP	Purple	positive regulation of NAD+ ADP-ribosyltransferase activity	2.40E-07	CHRNA3, CLU, FHIT, GRIN2A, NCK1, PPP2R3A, PPP4C, VRK2, PSMD6, KAT5, PLCL2, SEC11A, B3GAT1, BANK1, FANCL, OTUD7B, TNFRSF13C, TYW5, ASPHD1
GO:0044267	The chemical reactions and pathways involving a specific protein, rather than of proteins in general, occurring at the level of an individual cell. Includes cellular protein modification.	BP	Purple	cellular protein metabolic process	4.22E-07	CHRNA3, CLU, FHIT, GRIN2A, NCK1, PPP2R3A, PPP4C, VRK2, PSMD6, KAT5, PLCL2, SEC11A, B3GAT1, BANK1, FANCL, OTUD7B, ASPHD1
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Purple	protein binding	7.41E-07	CHRNA3, CLU, FHIT, GRIN2A, HSPA9, NCK1, PPP2R3A, PPP4C, SLC12A4, VRK2, AKAP6, CKAP5, PSMD6, KAT5, PLCL2, BANK1, FANCL, OTUD7B, RBM26, SEMA6D, THOC7, BORCS7, TYW5

Table 6.21: Gene Ontology for Red Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Red	positive regulation of macrophage proliferation	3.43E-06	CACNA1D, DRD2, GPM6A, NAGA, PGM3, PLCL1, PSMB10, HIRIP3, ATG13, ZDHHC5, NDFIP2, SLC38A7, PITPNM2, CPEB1, GFOD2, TDRD9, HS3ST5
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Red	protein binding	0.0004	CACNA1D, DRD2, GPM6A, NAGA, PLCL1, PSMB10, HIRIP3, ATG13, NDFIP2, SLC38A7, PITPNM2, CPEB1, HS3ST5
GO:0043412	The covalent alteration of one or more monomeric units in a polypeptide, polynucleotide, polysaccharide, or other biological macromolecule, resulting in a change in its properties.	BP	Red	macromolecule modification	0.000479	DRD2, PGM3, PLCL1, PSMB10, ATG13, ZDHHC5, NDFIP2, TDRD9, HS3ST5
GO:0099080	A cellular component that consists of an indeterminate number of proteins or macromolecular complexes, organized into a regular, higher-order structure such as a polymer, sheet, network, or a fibre.	CC	Red	supramolecular complex	0.00068	CACNA1D, DRD2, GPM6A

Table 6.22: Gene Ontology for Turquoise Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Turquoise	positive regulation of macrophage proliferation	3.42E-25	More than 50 overlapping genes
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Turquoise	intracellular organelle lumen	2.23E-12	SERPINC1, RERE, ATP2A2, GSDME, EMX1, IREB2, LRP1, NAB2, NEK1, NFATC3, NMB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, MAD1L1, CUL3, INPP4B, CACNA1I, RGS6, ZNF592, PLCH2, ZNF536, DMTF1, AKT3, SDCCAG8, CNOT1, DDHD2, VPS13C, NSD3, ZNF823, SLC39A8, BCL11B, ZBTB37, CREB3L1, SLC32A1, CNTN4, MED19, HCN1, BTBD18
GO:0005634	A membrane-bounded organelle of eukaryotic cells in which chromosomes are housed and replicated. In most cells, the nucleus contains all of the cell's chromosomes except the organellar chromosomes and is the site of RNA synthesis and processing. In some species, or in specialized cell types, RNA metabolism or DNA replication may be absent.	CC	Turquoise	nucleus	5.73E-11	RERE, ATP2A2, EMX1, LRP1, NAB2, NEK1, NFATC3, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, MAD1L1, CUL3, RGS6, ZNF592, ZNF536, DMTF1, AKT3, R3HDM2, CNOT1, PLA2G15, SPATS2L, NSD3, ZNF823, BCL11B, CENPM, ZBTB37, CREB3L1, MED19, HARBI1, BTBD18
GO:2000225	Any process that stops, prevents, or reduces the frequency, rate or extent of testosterone biosynthetic process.	BP	Turquoise	negative regulation of testosterone biosynthetic process	3.29E-09	RERE, EMX1, IREB2, NAB2, NFATC3, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Turquoise	cellular response to diacyl bacterial lipopeptide	3.32E-09	SERPINC1, RERE, GSDME, EMX1, IREB2, LRP1, NAB2, NEK1, NFATC3, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, CUL3, INPP4B, SLC7A6, ZNF592, PLCH2, ZNF536, DMTF1, AKT3, CNOT1, DDHD2, PLA2G15, AIG1, NSD3, ZNF823, ADAMTSL3, DPEP2, BCL11B, ZBTB37, CREB3L1, MED19, HARBI1, BTBD18
GO:0016020	A lipid bilayer along with all the proteins and	CC	Turquoise	membrane	5.76E-	SERPINC1, ATP2A2, GSDME, LRP1, MAPK3,

	protein complexes embedded in it an attached to it.				09	SHMT2, SREBF2, CUL3, CACNA1I, SLC7A6, RGS6, PLCH2, AKT3, MPHOSPH9, CNOT1, DDHD2, PLA2G15, SEZ6L2, SLC45A1, AIG1, MPP6, VPS13C, SLC39A8, DPEP2, ESAM, CREB3L1, EMB, SLC32A1, CPNE8, CNTN4, KIAA1324L, HARBI1, HCN1
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Turquoise	regulation of fertilization	7.34E-09	SERPINC1, RERE, GSDME, EMX1, IREB2, LRP1, NAB2, NEK1, NFATC3, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:0006357	Any process that modulates the frequency, rate or extent of transcription mediated by RNA polymerase II.	BP	Turquoise	regulation of transcription by RNA polymerase II	1.03E-08	RERE, EMX1, NFATC3, MAPK3, PSMA4, RELA, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Turquoise	protein binding	1.04E-08	SERPINC1, RERE, ATP2A2, EMX1, IREB2, LRP1, NAB2, NEK1, NFATC3, NMB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, MAD1L1, CUL3, INPP4B, CACNA1I, RGS6, ZNF592, AKT3, SDCCAG8, CNOT1, AIG1, MPP6, NSD3, ADAMTSL3, BCL11B, ESAM, CREB3L1, EMB, CNTN4, MED19, HCN1
GO:1903593	Any process that modulates the frequency, rate or extent of histamine secretion by mast cell.	BP	Turquoise	regulation of histamine secretion by mast cell	1.95E-08	RERE, EMX1, NAB2, NFATC3, MAPK3, PSMA4, RELA, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:2001234	Any process that stops, prevents or reduces the frequency, rate or extent of apoptotic signalling pathway.	BP	Turquoise	negative regulation of apoptotic signalling pathway	1.96E-08	RERE, EMX1, NAB2, NFATC3, MAPK3, PSMA4, RELA, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:0016043	A process that results in the assembly, arrangement of constituent parts, or disassembly of a cellular component.	BP	Turquoise	cellular component organization	2.04E-08	RERE, ATP2A2, EMX1, LRP1, NEK1, MAPK3, RELA, SHMT2, SREBF2, MAD1L1, CUL3, SLC7A6, AKT3, SDCCAG8, CNOT1, DDHD2, MPP6, VPS13C, NSD3, BCL11B, CENPM,

						ESAM, CREB3L1, EMB, CNTN4, HCN1, BTBD18
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl)benzene-1,2-diol] and derivatives formed by substitution.	BP	Turquoise	cellular response to catecholamine stimulus	3.49E-08	RERE, ATP2A2, EMX1, LRP1, NEK1, MAPK3, RELA, SHMT2, SREBF2, MAD1L1, CUL3, SLC7A6, AKT3, SDCCAG8, CNOT1, DDHD2, MPP6, VPS13C, NSD3, BCL11B, CENPM, ESAM, CREB3L1, EMB, CNTN4, HCN1, BTBD18
GO:0097707	A programmed cell death characterized morphologically by the presence of smaller than normal mitochondria with condensed mitochondrial membrane densities, reduction or vanishing of mitochondria crista, and outer mitochondrial membrane rupture. Activation of mitochondrial voltage-dependent anion channels and mitogen-activated protein kinases, upregulation of endoplasmic reticulum stress, and inhibition of cystine/glutamate antiporter are involved in the induction of ferroptosis. This process is characterized by the accumulation of lipid peroxidation products and lethal reactive oxygen species (ROS) derived from iron metabolism. Glutathione peroxidase 4 (GPX4), heat shock protein beta-1, and nuclear factor erythroid 2-related factor 2 function as negative regulators of ferroptosis by limiting ROS production and reducing cellular iron uptake, respectively. In contrast, NADPH oxidase and p53 act as positive regulators of ferroptosis by promotion of ROS production and inhibition of expression of SLC7A11 (a specific light-chain subunit of the cystine/glutamate antiporter), respectively. Misregulated ferroptosis has been implicated in multiple physiological and pathological	BP	Turquoise	ferroptosis	3.52E-08	RERE, EMX1, NAB2, NFATC3, MAPK3, PSMA4, RELA, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18

	processes.					
GO:1901679	The directed movement of nucleotide across a membrane.	BP	Turquoise	nucleotide transmembrane transport	4.74E-08	RERE, EMX1, IREB2, NAB2, NFATC3, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF2, CUL3, INPP4B, ZNF592, PLCH2, ZNF536, DMTF1, CNOT1, DDHD2, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:1901558	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a metformin stimulus.	BP	Turquoise	response to metformin	8.94E-08	RERE, EMX1, IREB2, NAB2, NFATC3, MAPK3, PSMA4, RELA, SOX5, SREBF2, CUL3, ZNF592, ZNF536, DMTF1, CNOT1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19, BTBD18
GO:0140244	Any process that regulates translation occurring at the presynapse.	BP	Turquoise	regulation of translation at presynapse	1.50E-07	RERE, EMX1, NAB2, NFATC3, RELA, SOX5, SREBF2, ZNF592, ZNF536, DMTF1, NSD3, ZNF823, BCL11B, ZBTB37, CREB3L1, MED19
GO:0031325	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways by which individual cells transform chemical substances.	BP	Turquoise	positive regulation of cellular metabolic process	2.81E-07	RERE, GSDME, LRP1, NAB2, NEK1, NFATC3, MAPK3, RELA, SOX5, SREBF2, CUL3, DMTF1, CNOT1, VPS13C, NSD3, BCL11B, CREB3L1, MED19, BTBD18

Table 6.23: Gene Ontology for Yellow Module Stage 4 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Yellow	molecular function	6.12E-08	CALB2, DPYD, EP300, LCAT, MGAT3, MSRA, PPP2R2A, PRKCB, PRKD1, TBX6, DOC2A, NT5C2, RIMS1, ZC3H7B, PLEKHO1, ARL6IP4, NDUFA4L2, RPTOR, ALPK3, DPEP3, DNAJC19, KMT5A
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Yellow	positive regulation of macrophage proliferation	2.55E-07	CALB2, DPYD, EP300, LCAT, MGAT3, MSRA, PPP2R2A, PRKCB, PRKD1, TBX6, DOC2A, NT5C2, RIMS1, ZC3H7B, PLEKHO1, ARL6IP4, NDUFA4L2, RPTOR, ALPK3, DPEP3, DNAJC19, KMT5A
GO:0043227	Organized structure of distinctive morphology and function, bounded by a single or double lipid bilayer membrane. Includes the nucleus, mitochondria, plastids, vacuoles, and vesicles. Excludes the plasma membrane.	CC	Yellow	membrane-bounded organelle	9.65E-07	CALB2, EP300, LCAT, MGAT3, MSRA, PPP2R2A, PRKCB, PRKD1, TBX6, DOC2A, ZC3H7B, PLEKHO1, ARL6IP4, NDUFA4L2, RPTOR, ALPK3, DPEP3, DNAJC19, KMT5A
GO:0043229	Organized structure of distinctive morphology and function, occurring within the cell. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes, and the cytoskeleton. Excludes the plasma membrane.	CC	Yellow	intracellular organelle	1.06E-06	CALB2, EP300, MGAT3, MSRA, PPP2R2A, PRKCB, PRKD1, TBX6, DOC2A, RIMS1, ZC3H7B, PLEKHO1, ARL6IP4, NDUFA4L2, RPTOR, ALPK3, DPEP3, DNAJC19, KMT5A
GO:0008152	The chemical reactions and pathways, including anabolism and catabolism, by which living organisms transform chemical substances. Metabolic processes typically transform small molecules, but also include macromolecular processes such as DNA repair and replication, and protein synthesis and degradation.	BP	Yellow	metabolic process	5.41E-06	DPYD, EP300, LCAT, MGAT3, MSRA, PPP2R2A, PRKCB, PRKD1, TBX6, NT5C2, RIMS1, ZC3H7B, ARL6IP4, NDUFA4L2, RPTOR, ALPK3, DPEP3, KMT5A

Table 6.24: Gene Ontology for Black Module Stage 5 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Black	positive regulation of macrophage proliferation	2.43E-38	More than 50 overlapping genes
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Black	cellular response to diacyl bacterial lipopeptide	3.66E-14	More than 50 overlapping genes
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-(2-aminoethyl)benzene-1,2-diol] and derivatives formed by substitution.	BP	Black	cellular response to catecholamine stimulus	3.49E-13	RERE, ATP2A2, CACNB2, CHRNA3, EMX1, EP300, FGFR1, GPM6A, LRP1, PRKCB, MAPK3, PTPRF, RELA, SHMT2, SIPA1, SREBF1, SREBF2, MAD1L1, CUL3, SLC7A6, AKAP6, CKAP5, AKT3, SDCCAG8, EPN2, CNOT1, DDHD2, MAU2, MPP6, HYDIN, VPS13C, NSD3, BCL11B, CENPM, CENPT, IMMP2L, ESAM, CREB3L1, EMB, CARMIL2, CNTN4, TSNARE1, HCN1, BTBD18
GO:1901679	The directed movement of nucleotide across a membrane.	BP	Black	nucleotide transmembrane transport	1.02E-12	RERE, DPYD, EMX1, EP300, FGFR1, IREB2, NAB2, NFATC3, PDE4B, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF1, SREBF2, CUL3, INPP4B, ZNF592, PLCH2, ZNF536, PSMD6, DMTF1, TAB1, RAI1, NT5C2, CNOT1, DDHD2, GATAD2A, NSD3, BANK1, ZNF823, CPEB1, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, MARS2, TNFRSF13C, BTBD18
GO:2000225	Any process that stops, prevents, or reduces the frequency, rate or extent of testosterone biosynthetic process.	BP	Black	negative regulation of testosterone biosynthetic process	1.86E-12	RERE, EMX1, EP300, FGFR1, IREB2, NAB2, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, BANK1,

						ZNF823, CPEB1, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, BTBD18
GO:0070016	Interacting selectively and non-covalently with the armadillo repeat domain of a protein, an approximately 40 amino acid long tandemly repeated sequence motif first identified in the Drosophila segment polarity protein armadillo. Arm-repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation.	MF	Black	armadillo repeat domain binding	4.04E-12	SERPINC1, ATP2A2, CACNA1C, CACNB2, CHRNA3, DPYD, EMX1, EP300, FGFR1, GPM6A, IREB2, LRP1, NMB, PDE4B, PRKCB, MAPK3, PSMA4, SHMT2, SREBF1, SREBF2, CUL3, CACNA1I, AKAP6, PLCH2, PSMD6, AKT3, TMX2, VPS13C, SLC39A8, TNFRSF13C, SLC32A1, CARMIL2, CNTN4, HCN1
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Black	regulation of fertilization	1.59E-11	SERPINC1, RERE, CHRNA3, GSDME, EMX1, EP300, FGFR1, IREB2, LRP1, NAB2, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, CNOT1, RBFOX1, GATAD2A, NSD3, BANK1, ZNF823, CPEB1, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, TNFRSF13C, BTBD18
GO:1901561	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a benomyl stimulus.	BP	Black	response to benomyl	2.72E-11	RERE, ATP2A2, MPPED2, DPYD, EMX1, EP300, FGFR1, IREB2, LRP1, NFATC3, PDE4B, PRKCB, MAPK3, RELA, SHMT2, SOX5, SREBF1, SREBF2, ZNF592, ZNF536, DMTF1, AKT3, R3HDM2, NT5C2, CNOT1, MAU2, ABCB9, SPATS2L, ARL6IP4, RBFOX1, GATAD2A, ZNF823, ALPK3, CPEB1, BCL11B, CENPT, ZBTB37, CREB3L1, MARS2, HCN1
GO:0006357	Any process that modulates the frequency, rate or extent of transcription mediated by RNA polymerase II.	BP	Black	regulation of transcription by RNA polymerase II	3.13E-11	RERE, EMX1, EP300, FGFR1, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, RAI1, CNOT1, GATAD2A, ZNF823, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, BTBD18

GO:0048513	Development of a tissue or tissues that work together to perform a specific function or functions. Development pertains to the process whose specific outcome is the progression of a structure over time, from its formation to the mature structure. Organs are commonly observed as visibly distinct structures but may also exist as loosely associated clusters of cells that work together to perform a specific function or functions.	BP	Black	animal organ development	3.43E-11	SERPINC1, RERE, CACNA1C, GSDME, EMX1, EP300, FGFR1, GPM6A, IREB2, LRP1, NAB2, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, MAD1L1, CUL3, AKAP6, PSMD6, AKT3, TAB1, HYDIN, ALPK3, BCL11B, IMMP2L, SLC32A1, CNTN4, STAC3, HCN1
GO:0097178	The aggregation, arrangement and bonding together of a set of components to form a ruffle, a projection at the leading edge of a crawling cell; the protrusions are supported by a microfilament meshwork. The formation of ruffles (also called membrane ruffling) is thought to be controlled by a group of enzymes known as Rho GTPases, specifically RhoA, Rac1 and cdc42.	BP	Black	ruffle assembly	4.10E-11	RERE, ATP2A2, MPPED2, DPYD, EMX1, EP300, FGFR1, IREB2, LRP1, NFATC3, PDE4B, PRKCB, MAPK3, RELA, SHMT2, SOX5, SREBF1, SREBF2, ZNF592, ZNF536, DMTF1, AKT3, R3HDM2, NT5C2, CNOT1, MAU2, ABCB9, SPATS2L, ARL6IP4, RBFOX1, GATAD2A, ZNF823, ALPK3, CPEB1, BCL11B, CENPT, ZBTB37, CREB3L1, MARS2, HCN1
GO:1901555	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a paclitaxel stimulus.	BP	Black	response to paclitaxel	4.48E-11	RERE, DPYD, EMX1, EP300, FGFR1, IREB2, NAB2, NFATC3, PDE4B, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, NT5C2, CNOT1, ARL6IP4, RBFOX1, GATAD2A, NSD3, FANCL, ZNF823, CPEB1, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, MARS2, HARBI1, BTBD18
GO:0005829	The part of the cytoplasm that does not contain organelles, but which does contain other particulate matter, such as protein complexes.	CC	Black	cytosol	6.13E-11	GSDME, DPYD, EP300, FGFR1, IREB2, NFATC3, PDE4B, PRKCB, MAPK3, PSMA4, RELA, SIPA1, SREBF1, SREBF2, MAD1L1, CUL3, INPP4B, RGS6, CKAP5, PSMD6,

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

						DMTF1, TAB1, SDCCAG8, EPN2, NT5C2, CNOT1, DDHD2, SPATS2L, VPS13C, RALGAPA2, RANBP10, CPEB1, CENPM, CENPT, STAC3, HARBI1
GO:0009653	The process in which anatomical structures are generated and organized. Morphogenesis pertains to the creation of form.	BP	Black	anatomical structure morphogenesis	9.42E-11	RERE, CACNA1C, CHRNA3, EMX1, EP300, FGFR1, GPM6A, LRP1, NAB2, PRKCB, MAPK3, PSMA4, RELA, CUL3, PSMD6, AKT3, TAB1, SDCCAG8, EPN2, DDHD2, BCL11B, CREB3L1, TNFRSF13C, EMB, CARMIL2, CNTN4, HCN1
GO:1903593	Any process that modulates the frequency, rate, or extent of histamine secretion by mast cell.	BP	Black	regulation of histamine secretion by mast cell	1.07E-10	RERE, EMX1, EP300, FGFR1, NAB2, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, ZNF823, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, BTBD18
GO:2001234	Any process that stops, prevents, or reduces the frequency, rate or extent of apoptotic signalling pathway.	BP	Black	negative regulation of apoptotic signalling pathway	1.10E-10	RERE, EMX1, EP300, FGFR1, NAB2, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, ZNF823, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, BTBD18
GO:1901558	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a metformin stimulus.	BP	Black	response to metformin	1.33E-10	RERE, EMX1, EP300, FGFR1, IREB2, NAB2, NFATC3, PDE4B, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, NT5C2, CNOT1, GATAD2A, NSD3, ZNF823, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, BTBD18
GO:0097707	A programmed cell death characterized morphologically by the presence of smaller than normal mitochondria with condensed mitochondrial membrane densities, reduction or vanishing of	BP	Black	ferroptosis	2.84E-10	RERE, EMX1, EP300, FGFR1, NAB2, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, CNOT1, GATAD2A, NSD3, ZNF823, BCL11B, CENPT, BCL2L12, ZBTB37,

	mitochondria crista, and outer mitochondrial membrane rupture. Activation of mitochondrial voltage-dependent anion channels and mitogen-activated protein kinases, upregulation of endoplasmic reticulum stress, and inhibition of cystine/glutamate antiporter are involved in the induction of ferroptosis. This process is characterized by the accumulation of lipid peroxidation products and lethal reactive oxygen species (ROS) derived from iron metabolism. Glutathione peroxidase 4 (GPX4), heat shock protein beta-1, and nuclear factor erythroid 2-related factor 2 function as negative regulators of ferroptosis by limiting ROS production and reducing cellular iron uptake, respectively. In contrast, NADPH oxidase and p53 act as positive regulators of ferroptosis by promotion of ROS production and inhibition of expression of SLC7A11 (a specific light-chain subunit of the cystine/glutamate antiporter), respectively. Misregulated ferroptosis has been implicated in multiple physiological and pathological processes.					CREB3L1, BTBD18
GO:0090317	Any process that decreases the frequency, rate or extent of the directed movement of proteins within cells.	BP	Black	negative regulation of intracellular protein transport	9.81E-10	RERE, EMX1, EP300, FGFR1, NAB2, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SOX5, SREBF1, SREBF2, CUL3, ZNF592, ZNF536, PSMD6, DMTF1, TAB1, RAI1, CNOT1, ARL6IP4, RBFOX1, GATAD2A,

						NSD3, FANCL, ZNF823, CPEB1, BCL11B, CENPT, BCL2L12, ZBTB37, CREB3L1, MARS2, HARBI1, BTBD18
GO:0035556	The process in which a signal is passed on to downstream components within the cell, which become activated themselves to further propagate the signal and finally trigger a change in the function or state of the cell.	BP	Black	intracellular signal transduction	1.56E-09	ATP2A2, CA8, CACNA1C, GSDME, EP300, FGFR1, LRP1, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SIPA1, CUL3, AKAP6, RGS6, PLCH2, PSMD6, AKT3, TAB1, CNOT1, BANK1, RALGAPA2, BCL2L12, CREB3L1, STAC3
GO:0005886	The membrane surrounding a cell that separates the cell from its external environment. It consists of a phospholipid bilayer and associated proteins.	CC	Black	plasma membrane	4.80E-09	SERPINC1, ATP2A2, CACNA1C, CACNB2, CHRNA3, GSDME, FGFR1, GPM6A, LRP1, PDE4B, PRKCB, MAPK3, PTPRF, CACNA1I, SLC7A6, AKAP6, GABBR2, RGS6, PLCH2, CKAP5, SEZ6L2, MPP6, RALGAPA2, SLC39A8, CPEB1, ESAM, TNFRSF13C, EMB, SLC32A1, CPNE8, CARMIL2, CNTN4, STAC3, HARBI1, HCN1
GO:0033554	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating the organism is under stress. The stress is usually, but not necessarily, exogenous (e.g. temperature, humidity, ionizing radiation).	BP	Black	cellular response to stress	4.92E-09	ATP2A2, EP300, LRP1, MAPK3, PSMA4, PTPRF, RELA, SIPA1, SREBF1, SREBF2, CUL3, PSMD6, AKT3, TAB1, CNOT1, JKAMP, VPS13C, FANCL, CPEB1, BCL2L12, IMMP2L, CREB3L1
GO:0120060	Any process that modulates the frequency, rate or extent of any gastric emptying process, the process in which the liquid and liquid-suspended solid contents of the stomach exit through the pylorus into the duodenum.	BP	Black	regulation of gastric emptying	6.68E-09	ATP2A2, CACNA1C, CHRNA3, GPM6A, LRP1, NMB, PDE4B, PRKCB, MAPK3, PTPRF, AP3B2, CUL3, GABBR2, SDCCAG8, HYDIN, CPEB1, BCL11B, EMB, SLC32A1, CARMIL2, CNTN4, HCN1
GO:0071986	A eukaryotically conserved protein complex; in humans, it is comprised of LAMTOR1, LAMTOR2,	CC	Black	Ragulator complex	8.46E-09	SERPINC1, ATP2A2, CACNA1C, CACNB2, CHRNA3, GSDME, FGFR1, GPM6A, LRP1, PDE4B, PRKCB, MAPK3, PTPRF, CACNA1I,

	LAMTOR3, LAMTOR4, and LAMTOR5. The complex is anchored to lipid rafts in late endosome membranes via LAMTOR1, constitutes a guanine nucleotide exchange factor (GEF) for the Rag GTPases.					SLC7A6, AKAP6, GABBR2, RGS6, PLCH2, CKAP5, SEZ6L2, MPP6, RALGAPA2, SLC39A8, CPEB1, ESAM, TNFRSF13C, EMB, SLC32A1, CPNE8, CARMIL2, CNTN4, STAC3, HARBI1, HCN1
GO:0006928	The directed, self-propelled movement of a cell or subcellular component without the involvement of an external agent such as a transporter or a pore.	BP	Black	movement of cell or subcellular component	1.74E-08	RERE, ATP2A2, CACNA1C, CACNB2, FGFR1, GPM6A, LRP1, PDE4B, MAPK3, PTPRF, AP3B2, CUL3, CACNA1I, SLC7A6, AKT3, SDCCAG8, HYDIN, BCL11B, ESAM, EMB, CARMIL2, CNTN4
GO:0140244	Any process that regulates translation occurring at the presynapse.	BP	Black	regulation of translation at presynapse	2.89E-08	RERE, EMX1, EP300, NAB2, NFATC3, PRKCB, RELA, SOX5, SREBF1, SREBF2, ZNF592, ZNF536, DMTF1, RAI1, GATAD2A, NSD3, ZNF823, BCL11B, CENPT, ZBTB37, CREB3L1
GO:1901666	Any process that activates or increases the frequency, rate or extent of NAD+ ADP-ribosyltransferase activity.	BP	Black	positive regulation of NAD+ ADP-ribosyltransferase activity	3.59E-08	SERPINC1, CHRNA3, GSDME, DPYD, EP300, FGFR1, IREB2, LRP1, NAB2, PDE4B, PRKCB, MAPK3, PSMA4, PTPRF, RELA, SHMT2, SREBF1, CUL3, SLC7A6, PSMD6, AKT3, TAB1, NT5C2, CNOT1, PLA2G15, JKAMP, NSD3, BANK1, FANCL, ADAMTSL3, ALPK3, DPEP2, CPEB1, BCL2L12, IMMP2L, MARS2, TNFRSF13C
GO:0070901	The posttranscriptional addition of methyl groups to specific residues in a mitochondrial tRNA molecule.	BP	Black	mitochondrial tRNA methylation	4.16E-08	ATP2A2, CHRNA3, GSDME, EP300, FGFR1, LRP1, PDE4B, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SIPA1, SOX5, SREBF1, CUL3, AKAP6, PSMD6, TAB1, EPN2, CNOT1, CPEB1, CREB3L1, TNFRSF13C, CPNE8, HCN1
GO:0031981	The volume enclosed by the nuclear inner membrane.	CC	Black	nuclear lumen	6.21E-08	RERE, ATP2A2, EMX1, EP300, FGFR1, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SREBF1, SREBF2, CUL3, CKAP5, PSMD6, DMTF1, TAB1, RAI1, MAU2, PLA2G15, SPATS2L, ARL6IP4, GATAD2A, NSD3, FANCL, CPEB1, CENPM, CENPT, CREB3L1,

						STAC3
GO:0048699	The process in which nerve cells are generated. This includes the production of neuroblasts and their differentiation into neurons.	BP	Black	generation of neurons	8.25E-08	RERE, CHRNA3, GSDME, EMX1, EP300, FGFR1, GPM6A, LRP1, MAPK3, PTPRF, RELA, SOX5, ZNF536, SDCCAG8, BCL11B, EMB, CNTN4, HCN1
GO:0048468	The process whose specific outcome is the progression of the cell over time, from its formation to the mature structure. Cell development does not include the steps involved in committing a cell to a specific fate.	BP	Black	cell development	8.25E-08	RERE, ATP2A2, CHRNA3, EMX1, EP300, FGFR1, GPM6A, LRP1, MAPK3, PTPRF, RELA, CUL3, AKAP6, ZNF536, HYDIN, ALPK3, BCL11B, EMB, CNTN4, STAC3, HCN1
GO:0003824	Catalysis of a biochemical reaction at physiological temperatures. In biologically catalysed reactions, the reactants are known as substrates, and the catalysts are naturally occurring macromolecular substances known as enzymes. Enzymes possess specific binding sites for substrates, and are usually composed wholly or largely of protein, but RNA that has catalytic activity (ribozyme) is often also regarded as enzymatic.	MF	Black	catalytic activity	1.04E-07	MPPED2, CA8, DPYD, EP300, FGFR1, IREB2, PDE4B, PRKCB, MAPK3, PSMA4, PTPRF, SHMT2, CUL3, INPP4B, RGS6, PLCH2, PSMD6, AKT3, TAB1, NT5C2, CNOT1, DDHD2, ABCB9, PLA2G15, AIG1, NSD3, FANCL, ADAMTSL3, ALPK3, DPEP2, IMMP2L, MARS2, HARBI1
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Black	collagen V binding	1.39E-07	SERPINC1, RERE, ATP2A2, EMX1, EP300, FGFR1, NFATC3, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SREBF1, SREBF2, CUL3, CKAP5, PSMD6, DMTF1, TAB1, RAI1, MAU2, PLA2G15, SPATS2L, ARL6IP4, GATAD2A, NSD3, FANCL, CPEB1, CENPM, CENPT, CREB3L1, MARS2, STAC3
GO:0034220	A process in which an ion is transported across a membrane.	BP	Black	ion transmembrane transport	1.41E-07	ATP2A2, CACNA1C, CACNB2, CHRNA3, GPM6A, PDE4B, CACNA1I, SLC7A6, AKAP6, PLCH2, ABCB9, SLC39A8, EMB, SLC32A1, STAC3, HCN1
GO:0071329	Any process that results in a change			cellular response to sucrose	1.50E-	CHRNA3, GSDME, EP300, FGFR1, LRP1,

	in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a sucrose stimulus.	BP	Black	stimulus	07	PDE4B, PRKCB, MAPK3, PSMA4, RELA, SHMT2, SOX5, SREBF1, CUL3, AKAP6, PSMD6, TAB1, EPN2, CNOT1, CPEB1, CREB3L1, TNFRSF13C, HCN1
GO:0006996	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of an organelle within a cell. An organelle is an organized structure of distinctive morphology and function. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes and the cytoskeleton. Excludes the plasma membrane.	BP	Black	organelle organization	1.64E-07	RERE, ATP2A2, EP300, LRP1, PRKCB, MAPK3, RELA, SIPA1, SREBF1, SREBF2, MAD1L1, CUL3, CKAP5, AKT3, SDCCAG8, CNOT1, DDHD2, MAU2, HYDIN, VPS13C, NSD3, CENPM, CENPT, IMMP2L, CARMIL2, TSNARE1, BTBD18
GO:0046872	Interacting selectively and non-covalently with any metal ion.	MF	Black	metal ion binding	1.69E-07	RERE, ATP2A2, MPPED2, CA8, CACNA1C, DPYD, EP300, IREB2, LRP1, PDE4B, PRKCB, SHMT2, ZNF592, PLCH2, ZNF536, RAI1, NT5C2, DDHD2, GATAD2A, NSD3, FANCL, ZNF823, DPEP2, CPEB1, BCL11B, ZBTB37, STAC3, HARBI1
GO:0071516	The initial formation of a stable single-strand DNA lesion that triggers programmed gene conversion at the mating-type locus, thereby restricting mating-type interconversion to one of the two sister chromatids during DNA replication.	BP	Black	establishment of imprinting at mating-type locus	2.85E-07	CHRNA3, EP300, FGFR1, LRP1, PDE4B, PRKCB, MAPK3, RELA, SHMT2, SOX5, SREBF1, AKAP6, TAB1, CNOT1, CPEB1, CREB3L1, HCN1
GO:0003700	A protein or a member of a complex that interacts selectively and non-covalently with a specific DNA sequence (sometimes referred to as a motif) within the regulatory region of a gene to modulate transcription. Regulatory regions include promoters (proximal and distal) and	MF	Black	DNA-binding transcription factor activity	4.06E-07	RERE, EMX1, EP300, NFATC3, RELA, SOX5, SREBF1, SREBF2, ZNF592, ZNF536, DMTF1, RAI1, GATAD2A, ZNF823, BCL11B, CENPT, ZBTB37, CREB3L1

	enhancers. Genes are transcriptional units and include bacterial operons.					
GO:0016021	The component of a membrane consisting of the gene products and protein complexes having at least some part of their peptide sequence embedded in the hydrophobic region of the membrane.	CC	Black	integral component of membrane	4.11E-07	ATP2A2, CACNA1C, CACNB2, CHRNA3, FGFR1, GPM6A, LRP1, PDE4B, PTPRF, SREBF1, SREBF2, CACNA1I, SLC7A6, AKAP6, GABBR2, ABCB9, SEZ6L2, TMX2, AIG1, JKAMP, SLC39A8, IMMP2L, ESAM, CREB3L1, TNFRSF13C, EMB, SLC32A1, TSNARE1, KIAA1324L, STAC3, HCN1
GO:0051254	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways involving RNA.	BP	Black	positive regulation of RNA metabolic process	4.21E-07	RERE, EP300, NFATC3, PRKCB, MAPK3, RELA, SOX5, SREBF1, SREBF2, DMTF1, RAI1, CNOT1, NSD3, CPEB1, BCL11B, BCL2L12, CREB3L1, BTBD18
GO:0070852	A neuron projection that is found in unipolar neurons and corresponds to the region between the cell body and the point at which the single projection branches.	CC	Black	cell body fibre	4.52E-07	ATP2A2, CACNA1C, CACNB2, GPM6A, PDE4B, PRKCB, CACNA1I, AKAP6, PLCH2, SLC39A8, STAC3

Table 6.25: Gene Ontology for Brown Module Stage 5 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Brown	positive regulation of macrophage proliferation	1.14E-11	BTG1, CALB2, GRIN2A, HSPA9, NCK1, PLCB2, PPP4C, PTN, TLE3, VRK2, ALMS1, ASH2L, GPR52, TAOK2, KAT5, SEC11A, CA14, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, SEMA6D, THOC7, PLPP5, WDR73, TYW5, RFTN2, LETM2, WBP2NL, MED19, HS3ST5, HAPLN4
GO:0005488	The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.	MF	Brown	binding	1.07E-09	BTG1, CALB2, GRIN2A, HSPA9, NCK1, PLCB2, PPP2R3A, PPP4C, PTN, TLE3, VRK2, ALMS1, ASH2L, TAOK2, KAT5, CA14, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, SEMA6D, THOC7, TYW5, LETM2, WBP2NL, MED19, HS3ST5, HAPLN4
GO:0044237	The chemical reactions and pathways by which individual cells transform chemical substances.	BP	Brown	cellular metabolic process	9.83E-09	BTG1, GRIN2A, HSPA9, NCK1, PLCB2, PPP2R3A, PPP4C, PTN, TLE3, VRK2, ASH2L, TAOK2, KAT5, SEC11A, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, THOC7, PLPP5, TYW5, WBP2NL, MED19, HS3ST5
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Brown	cytoplasm	1.07E-08	BTG1, CALB2, GRIN2A, HSPA9, NCK1, PLCB2, PPP4C, PTN, VRK2, ALMS1, TAOK2, KAT5, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, SEMA6D, THOC7, PLPP5, WDR73, TYW5, LETM2, WBP2NL, HS3ST5
GO:0044238	The chemical reactions and pathways involving those compounds which are formed as a part of the normal anabolic and catabolic processes. These processes take place in most, if not all, cells of the organism.	BP	Brown	primary metabolic process	5.50E-08	BTG1, GRIN2A, NCK1, PLCB2, PPP2R3A, PPP4C, PTN, TLE3, VRK2, ASH2L, TAOK2, KAT5, SEC11A, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, THOC7, PLPP5, TYW5, WBP2NL, MED19, HS3ST5
GO:0043226	Organized structure of distinctive morphology and function. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes and the cytoskeleton, and prokaryotic structures	CC	Brown	organelle	1.51E-07	BTG1, CALB2, GRIN2A, HSPA9, NCK1, PPP4C, PTN, TLE3, VRK2, ALMS1, ASH2L, TAOK2, KAT5, SEC11A, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, SEMA6D, THOC7, WDR73, LETM2, WBP2NL, MED19, HS3ST5

	such as anammoxosomes and pirellulosomes. Excludes the plasma membrane.					
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Brown	cellular response to diacyl bacterial lipopeptide	2.99E-07	BTG1, GRIN2A, NCK1, PLCB2, PPP2R3A, PPP4C, PTN, TLE3, VRK2, ASH2L, TAOK2, KAT5, SEC11A, B3GAT1, OTUD7B, GDPD3, ACTR5, SETD6, THOC7, PLPP5, TYW5, WBP2NL, MED19, HS3ST5

Table 6.26: Gene Ontology for Green Module Stage 5 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Green	positive regulation of macrophage proliferation	5.52E-06	CHRM3, CHRNB4, CYP17A1, GRM3, MEF2C, MGAT3, PRKD1, ATXN7, INA, RABGAP1L, MOB4, PARD6A, PLEKHO1, CNNM2, ANP32E, ZNF804A, ASPHD1, ASPG
GO:0098809	Catalysis of the reaction: nitrite + acceptor = product(s) of nitrate reduction + reduced acceptor.	MF	Green	nitrite reductase activity	1.04E-05	CHRM3, CHRNB4, GRM3, MEF2C, INA, ZNF804A
GO:0043226	Organized structure of distinctive morphology and function. Includes the nucleus, mitochondria, plastids, vacuoles, vesicles, ribosomes and the cytoskeleton, and prokaryotic structures such as anammoxosomes and pirellulosomes. Excludes the plasma membrane.	CC	Green	organelle	1.33E-05	CHRM3, CHRNB4, CYP17A1, GRM3, MEF2C, MGAT3, PRKD1, ATXN7, INA, RABGAP1L, MOB4, PARD6A, PLEKHO1, CNNM2, ANP32E, ZNF804A
GO:0005737	All of the contents of a cell excluding the plasma membrane and nucleus but including other subcellular structures.	CC	Green	cytoplasm	8.25E-05	CHRNB4, CYP17A1, MEF2C, MGAT3, PRKD1, ATXN7, INA, RABGAP1L, MOB4, PARD6A, PLEKHO1, ANP32E, ZNF804A, ASPG
GO:0016020	A lipid bilayer along with all the proteins and protein complexes embedded in it an attached to it.	CC	Green	membrane	0.000105	CHRM3, CHRNB4, CYP17A1, GRM3, MGAT3, PRKD1, INA, MOB4, PARD6A, PLEKHO1, CNNM2, ZNF804A, ASPHD1

Table 6.27: Gene Ontology for Greenyellow Module Stage 5 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Greenyellow	positive regulation of macrophage proliferation	1.22E-34	More than 50 overlapping genes
GO:0071726	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a diacylated bacterial lipopeptide stimulus.	BP	Greenyellow	cellular response to diacyl bacterial lipopeptide	3.08E-18	More than 50 overlapping genes
GO:0070052	Interacting selectively and non-covalently with a type V collagen trimer.	MF	Greenyellow	collagen V binding	7.65E-15	BNIP3L, SERPING1, NCAN, ERCC4, HSPD1, HSPE1, MMP16, PCCB, PSMB10, STAT6, XRCC3, FXR1, CDK2AP1, HIRIP3, DGKZ, DGKI, KDM4A, ZEB2, NUTF2, CLP1, SATB2, SF3B1, FOXP1, MSL2, TSR1, PAK6, THAP11, RPTOR, SUGP1, ACD, MAIP1, ESRP2, YPEL3, L3MBTL2, C12orf65, YPEL4, INO80E, KMT5A
GO:0016020	A lipid bilayer along with all the proteins and protein complexes embedded in it an attached to it.	CC	Greenyellow	membrane	9.83E-12	ARHGAP1, BNIP3L, CACNA1D, CLCN3, HSPD1, HSPE1, KCNJ13, MAP3K11, MMP16, NDUFA6, OPCML, PLCL1, RRAS, STAT6, FXR1, DGKZ, DGKI, BAG4, ATG13, TBC1D5, SNAP91, NUTF2, VPS45, DOP1A, PSD3, ZDHHC5, GIGYF2, SLC45A1, NDFIP2, SLC38A7, NLGN4X, RPTOR, PITPNM2, SRR, DPEP3, NDRG4, MAIP1, COQ10B, RFT1, BORCS7, RILPL2, EHBP1L1, SNORC
GO:0071870	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a catecholamine stimulus. A catecholamine is any of a group of biogenic amines that includes 4-(2-aminoethyl) pyrocatechol [4-	BP	Greenyellow	cellular response to catecholamine stimulus	1.02E-10	BNIP3L, CLCN3, NCAN, ERCC4, HSPD1, MMP16, NDUFA6, XRCC3, HIRIP3, BAG4, KDM4A, ATG13, TBC1D5, ZEB2, SNAP91, CLP1, SATB2, SF3B1, MSL2, TSR1, PAK6, NLGN4X, RPTOR, SRR, DPEP3, NDRG4, ACD, MAIP1, L3MBTL2, C12orf65, RFT1, TDRD9, RILPL2, INO80E, KMT5A

	(2-aminoethyl)benzene-1,2-diol] and derivatives formed by substitution.					
GO:0031323	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways by which individual cells transform chemical substances.	BP	Greenyellow	regulation of cellular metabolic process	1.36E-10	BNIP3L, SERPING1, ERCC4, HSPD1, HSPE1, MAP3K11, PLCL1, PSMB10, RRAS, STAT6, FXR1, CDK2AP1, DGKZ, BAG4, KDM4A, ATG13, TBC1D5, ZEB2, PLCL2, SATB2, GIGYF2, FOXP1, NDFIP2, HPF1, ZSCAN2, PAK6, THAP11, RPTOR, RBM26, NDRG4, ACD, ESRP2, L3MBTL2, KMT5A
GO:0060255	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways involving macromolecules, any molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass.	BP	Greenyellow	regulation of macromolecule metabolic process	2.83E-10	SERPING1, ERCC4, HSPD1, HSPE1, MAP3K11, PLCL1, PSMB10, RRAS, STAT6, FXR1, CDK2AP1, BAG4, KDM4A, ATG13, TBC1D5, ZEB2, NUTF2, CLP1, PLCL2, SATB2, SF3B1, GIGYF2, FOXP1, NDFIP2, HPF1, ZSCAN2, PAK6, THAP11, RPTOR, NDRG4, ACD, ESRP2, L3MBTL2, TDRD9, KMT5A
GO:1901666	Any process that activates or increases the frequency, rate or extent of NAD+ ADP-ribosyltransferase activity.	BP	Greenyellow	positive regulation of NAD+ ADP-ribosyltransferase activity	1.20E-09	BNIP3L, SERPING1, NCAN, CTRL, HSPD1, HSPE1, MAP3K11, MMP16, NAGA, PCCB, PGM3, PLCL1, PSMB10, RRAS, FXR1, CDK2AP1, BAG4, KDM4A, ATG13, ZEB2, PLCL2, ZDHHC5, GIGYF2, NDFIP2, HPF1, MSL2, PAK6, RPTOR, SRR, DPEP3, NDRG4, C12orf65, INO80E, KMT5A
GO:0080154	Any process that modulates the rate, frequency or extent of fertilization. Fertilization is the union of gametes of opposite sexes during the process of sexual reproduction to form a zygote. It involves the fusion of the gametic nuclei (karyogamy) and cytoplasm (plasmogamy).	BP	Greenyellow	regulation of fertilization	9.82E-09	SERPING1, ERCC4, HSPD1, HSPE1, MAP3K11, PLCL1, PSMB10, RRAS, STAT6, FXR1, CDK2AP1, DGKZ, BAG4, KDM4A, ATG13, ZEB2, PLCL2, SATB2, GIGYF2, FOXP1, NDFIP2, HPF1, ZSCAN2, PAK6, THAP11, RPTOR, NDRG4, ACD, ESRP2, L3MBTL2, KMT5A
GO:0005654	That part of the nuclear content other than the chromosomes or the nucleolus.	CC	Greenyellow	nucleoplasm	1.24E-08	BNIP3L, ERCC4, PSMB10, STAT6, XRCC3, CDK2AP1, DGKZ, DGKI, KDM4A, NUTF2, CLP1, SATB2, SF3B1, FOXP1, MSL2, TSR1, THAP11, RPTOR, SUGP1, ACD, ESRP2, L3MBTL2, INO80E, KMT5A
GO:0005829	The part of the cytoplasm that does not	CC	Greenyellow	cytosol	2.46E-08	ARHGAP1, BNIP3L, HSPD1, PCCB, PGM3,

	contain organelles, but which does contain other particulate matter, such as protein complexes.					PSMB10, STAT6, XRCC3, FXR1, CDK2AP1, DGKI, BAG4, KDM4A, ATG13, TBC1D5, ZEB2, NUTF2, CLP1, DOP1A, GIGYF2, TSR1, RPTOR, PITPNM2, SRR, NDRG4, BORCS7, RILPL2, KMT5A
GO:0051171	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways involving nitrogen or nitrogenous compounds.	BP	Greenyellow	regulation of nitrogen compound metabolic process	2.49E-08	SERPING1, ERCC4, HSPD1, HSPE1, MAP3K11, PLCL1, PSMB10, RRAS, STAT6, FXR1, CDK2AP1, BAG4, KDM4A, ATG13, ZEB2, PLCL2, SATB2, GIGYF2, FOXP1, NDFIP2, HPF1, ZSCAN2, PAK6, THAP11, RPTOR, NDRG4, ACD, ESRP2, L3MBTL2, KMT5A
GO:0006810	The directed movement of substances (such as macromolecules, small molecules, ions) or cellular components (such as complexes and organelles) into, out of or within a cell, or between cells, or within a multicellular organism by means of some agent such as a transporter, pore or motor protein.	BP	Greenyellow	transport	4.37E-08	ARHGAP1, BNIP3L, SERPING1, CACNA1D, CLCN3, HSPD1, KCNJ13, PSMB10, DGKI, BAG4, ATG13, TBC1D5, SNAP91, NUTF2, VPS45, DOP1A, FOXP1, SLC45A1, NDFIP2, SLC38A7, NLGN4X, PITPNM2, RBM26, NDRG4, ACD, MAIP1, RFT1, RILPL2
GO:0048523	Any process that stops, prevents, or reduces the frequency, rate or extent of a cellular process, any of those that are carried out at the cellular level, but are not necessarily restricted to a single cell. For example, cell communication occurs among more than one cell, but occurs at the cellular level.	BP	Greenyellow	negative regulation of cellular process	7.30E-08	ARHGAP1, BNIP3L, SERPING1, ERCC4, HSPD1, PSMB10, RRAS, STAT6, XRCC3, FXR1, DGKI, BAG4, KDM4A, ZEB2, PLCL2, SATB2, GIGYF2, FOXP1, THAP11, NLGN4X, RPTOR, RBM26, NDRG4, ACD, L3MBTL2, TDRD9, KMT5A
GO:1901561	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a benomyl stimulus.	BP	Greenyellow	response to benomyl	7.32E-08	CLCN3, ERCC4, HSPD1, HSPE1, MAP3K11, PCCB, RRAS, STAT6, XRCC3, FXR1, DGKZ, DGKI, BAG4, ZEB2, CLP1, SATB2, SF3B1, GIGYF2, FOXP1, ZSCAN2, TSR1, PAK6, THAP11, SUGP1, SRR, RBM26, ACD, ESRP2, C12orf65, TDRD9
GO:0071725	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a triacylated bacterial lipopeptide stimulus.	BP	Greenyellow	response to triacyl bacterial lipopeptide	8.59E-08	ARHGAP1, BNIP3L, CACNA1D, HSPD1, BAG4, ATG13, TBC1D5, SNAP91, NUTF2, VPS45, DOP1A, FOXP1, SLC45A1, NDFIP2, SLC38A7, PITPNM2, RBM26, ACD, MAIP1, RFT1, RILPL2

Formatted: Font: (Default) Times New Roman, Font color: Black, Pattern: Clear

GO:0097178	The aggregation, arrangement and bonding together of a set of components to form a ruffle, a projection at the leading edge of a crawling cell; the protrusions are supported by a microfilament meshwork. The formation of ruffles (also called membrane ruffling) is thought to be controlled by a group of enzymes known as Rho GTPases, specifically RhoA, Rac1 and cdc42.	BP	Greenyellow	ruffle assembly	1.01E-07	CLCN3, ERCC4, HSPD1, HSPE1, MAP3K11, PCCB, RRAS, STAT6, XRCC3, FXR1, DGKZ, DGKI, BAG4, ZEB2, CLP1, SATB2, SF3B1, GIGYF2, FOXP1, ZSCAN2, TSR1, PAK6, THAP11, SUGP1, SRR, RBM26, ACD, ESRP2, C12orf65, TDRD9
GO:0033554	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating the organism is under stress. The stress is usually, but not necessarily, exogenous (e.g. temperature, humidity, ionizing radiation).	BP	Greenyellow	cellular response to stress	1.37E-07	BNIP3L, ERCC4, MAP3K11, PSMB10, STAT6, XRCC3, BAG4, ATG13, ZEB2, GIGYF2, FOXP1, HPF1, PAK6, RPTOR, ACD, YPEL3, INO80E, KMT5A
GO:0070025	Interacting selectively and non-covalently with carbon monoxide (CO).	MF	Greenyellow	carbon monoxide binding	1.57E-07	ARHGAP1, SERPING1, CACNA1D, ERCC4, HSPD1, HSPE1, MAP3K11, MMP16, PLCL1, DGKZ, DGKI, BAG4, ATG13, TBC1D5, ZEB2, PLCL2, NDFIP2, PAK6, RPTOR, RBM26, ACD
GO:0051649	Any process, occurring in a cell, that localizes a substance or cellular component. This may occur via movement, tethering or selective degradation.	BP	Greenyellow	establishment of localization in cell	2.30E-07	ARHGAP1, BNIP3L, SERPING1, CACNA1D, CLCN3, HSPD1, DGKI, BAG4, ATG13, TBC1D5, SNAP91, NUTF2, VPS45, DOP1A, NLGN4X, RBM26, NDRG4, ACD, MAIP1, RILPL2
GO:0044267	The chemical reactions and pathways involving a specific protein, rather than of proteins in general, occurring at the level of an individual cell. Includes cellular protein modification.	BP	Greenyellow	cellular protein metabolic process	2.60E-07	BNIP3L, SERPING1, HSPD1, HSPE1, MAP3K11, PGM3, PLCL1, PSMB10, RRAS, FXR1, CDK2AP1, BAG4, KDM4A, ATG13, ZEB2, PLCL2, ZDHHC5, GIGYF2, NDFIP2, HPF1, MSL2, PAK6, RPTOR, NDRG4, C12orf65, INO80E, KMT5A
GO:0043168	Interacting selectively and non-covalently with anions, charged atoms or groups of atoms with a net negative charge.	MF	Greenyellow	anion binding	2.82E-07	CLCN3, NCAN, HSPD1, HSPE1, MAP3K11, PCCB, PLCL1, RRAS, XRCC3, DGKZ, DGKI, SNAP91, CLP1, PLCL2, TSR1, PAK6, NLGN4X, PITPNM2, SRR, TDRD9

Table 6.28: Gene Ontology for Pink Module Stage 5 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate or extent of macrophage proliferation.	BP	Pink	positive regulation of macrophage proliferation	0.000262	EPHX2, ETF1, STAR, TCF4, ATP5MPL, IGSF9B, SMG6, NEMP1, AMBRA1, NDUFA4L2, ZNF408, C16orf92
GO:0019866	The inner, i.e. lumen-facing, lipid bilayer of an organelle envelope; usually highly selective to most ions and metabolites.	CC	Pink	organelle inner membrane	0.000645	STAR, ATP5MPL, NEMP1, NDUFA4L2
GO:0031966	Either of the lipid bilayers that surround the mitochondrion and form the mitochondrial envelope.	CC	Pink	mitochondrial membrane	0.001661	STAR, ATP5MPL, AMBRA1, NDUFA4L2
GO:0003674	A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.	MF	Pink	molecular function	0.002147	EPHX2, ETF1, STAR, TCF4, ATP5MPL, IGSF9B, SMG6, NEMP1, AMBRA1, NDUFA4L2, ZNF408

Table 6.29: Gene Ontology for Red Module Stage 5 using the anRichment function as part of WGCNA in R using the default settings

GOID	Definition	Ontology	Module	Go Process	FDR	Genes
GO:0120041	Any process that activates or increases the frequency, rate, or extent of macrophage proliferation.	BP	Red	positive regulation of macrophage proliferation	1.84E-07	ALDOA, CHRNA5, F2, PTK2B, FHIT, MSRA, PPP2R2A, RANGAP1, KCNK7, STAG1, ZC3H7B, PPP1R13B, NGEF, TSNAXIP1, BOLL, PCGF6, ATPAF2, TMEM219, PCNX3
GO:0005515	Interacting selectively and non-covalently with any protein or protein complex (a complex of two or more proteins that may include other nonprotein molecules).	MF	Red	protein binding	1.43E-05	ALDOA, CHRNA5, F2, PTK2B, FHIT, PPP2R2A, RANGAP1, STAG1, ZC3H7B, PPP1R13B, NGEF, BOLL, PCGF6, ATPAF2, TMEM219
GO:0070013	An organelle lumen that is part of an intracellular organelle.	CC	Red	intracellular organelle lumen	3.82E-05	ALDOA, CHRNA5, F2, PTK2B, FHIT, PPP2R2A, RANGAP1, KCNK7, STAG1, ZC3H7B, PPP1R13B, NGEF, BOLL, PCGF6, TMEM219
GO:0005829	The part of the cytoplasm that does not contain organelles, but which does contain other particulate matter, such as protein complexes.	CC	Red	cytosol	0.000218	ALDOA, PTK2B, FHIT, MSRA, PPP2R2A, RANGAP1, STAG1, PPP1R13B, NGEF, ATPAF2