2020

Insights into Gene Expression and the Influence of Regulatory Mechanisms within Meningioma - A Bioinformatic Approach

Carr, L.

The Plymouth Student Scientist
University of Plymouth

Insights into Gene Expression and the Influence of Regulatory Mechanisms within Meningioma - A Bioinformatic Approach

Louise Carr

Project Advisor: Professor Matthias Futschik, School of Biomedical Sciences, University of Plymouth, Drake Circus, Plymouth, PL4 8AA

Abstract

Meningioma is cancer of the meninges, the protective lining of the brain and spinal cord. Currently meningioma is classified using WHO Grades (I-III) which is based on histological characteristics of the tumours. This approach can give inaccurate indication of tumour aggression and therefore an inappropriate treatment course is chosen. The aim of this study was to identify differentially expressed genes that may act as biomarkers to indicate tumour aggressiveness. A further aim was to infer and investigate the role of specific gene expression regulatory mechanisms within meningioma. Bioinformatic approaches for transcriptomic analysis were used to study microarray data of 62 meningioma tumour patients. Comparison of gene expression was carried out between the 3 WHO grades and between groups with different clinical rates of recurrence. Differential gene expression analysis was completed using online tools GEO2R and Network Analyst, enrichment analysis was performed using WebGestalt and X2KWeb was used to investigate transcription factor influence. Identified potential biomarkers of aggression include an upregulation of PTTG1, SRSF6, and FOXM1 and downregulation of LEPR and SFRP1. Furthermore, through enrichment analysis cell cycle, metabolic pathways and spliceosomes were identified to be overrepresented in the upregulated genes of both grade III and aggressive comparison groups. In conclusion this study identified potential genetic characteristics and the associated biological pathways and processes that are dysregulated in disease state. It also provides potential biomarkers of meningioma aggression for further functional validation.

Introduction

Meningioma is a cancer of the meninges and accounts for up to 36% of all primary central nervous system (CNS) tumours and is the most common intracranial tumour (Louis et al, 2016). Meningiomas are typically diagnosed in patients between the ages of 40-60, with incidence increasing with age and are twice more common in women compared to men (Magill et al, 2018). Meningioma also occurs within children and adolescents; however this is a rare occurrence and accounts for 4.6% of patients with CNS tumours (Pećina-Šlaus et al, 2016).

Meningioma originates from progenitor cells that develop into arachnoidal cap cells. These cells constitute the thin layer called the meninges which covers the brain and spinal cord (Pećina-Šlaus et al, 2016). Meningioma tumour growth can be either intraspinal or intracranial, the latter being more common with the primary sites of meningioma being intraspinal in only ~12% of patients (Saraf et al, 2011). Intracranial tumours can occur at the gyri (ridges) or sulci (folds) of the brain with the specific location having direct impact on prognosis surgical resection success and symptom presentation of patients (Desai and Patel, 2016). A theory of a relationship between hormones and meningioma progression exists and is based on observations such as the gender bias in prevalence, identification of hormone receptors on tumours and the influence of pregnancy and hormone replacement therapy has on tumour growth (Wiemels et al, 2010 and Gurcay et al, 2018).

Classification of meningioma tumours by the World Health Organisation (WHO) consists of 3 groups: slow growing benign WHO grade I, atypical WHO grade II or malignant WHO grade III; these groups represent approximately 80%, 18% and 1-3% of all meningiomas respectively (Louis et al, 2016). Grade I meningiomas criteria include 9 histologically defined subgroups and must lack criteria of the other two grades; as a result, there are difficulties for diagnostics in borderline cases, especially between grade I and grade II cases (Harter et al, 2017), as outlined in Figure 1.
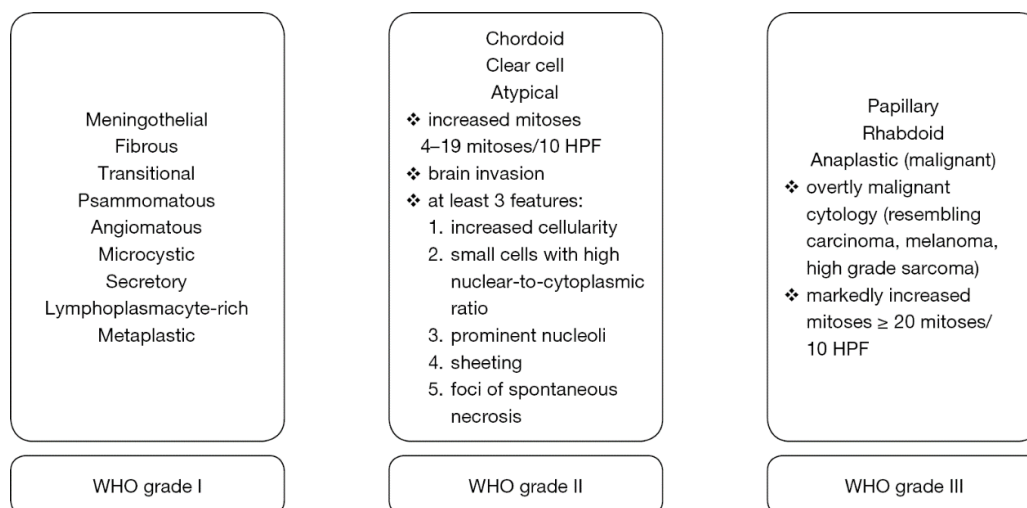
Figure 1: Histological criteria for WHO grading of meningioma. Adapted from (From Harter et al, 2017 under CC BY 3.0). Mitotic rate is based on the assessment of number of mitoses per 10 high power fields (HPF) and a non-standard sample area.

80% of meningioma cases are classed as slow growing benign tumours (WHO grade I), while the remaining 20% of tumours, WHO grade II and III, portray aggressive and metastasising behaviour with high propensity for recurrence thus have an increased morbidity and mortality (Pećina-Šlaus et al, 2016).

The majority of symptomatic meningiomas can be successfully treated by surgical resection with or without adjuvant radiotherapy; yet, within WHO grade I the preferred management option for most patients is observation using magnetic resonance imaging (MRI) (Wang and Osswald, 2018). This is due to the morbidity associated with surgery and its variable success that is dependent on tumour location which ultimately dictates the degree of resection. Prediction of tumour recurrence is commonly based on the WHO classification and the extent of resection. However, other factors also influence risk such as tumour size, location, age and gender, increased mitotic activity and genetic characteristics (Dunn et al, 2018). Currently 5-year recurrence rates following total resection is at 3%, 38%, and 78% for grades I, II and III respectively, demonstrating there is a poor relapse-free survival for higher grade meningioma (Dunn et al, 2018). While the WHO classification system is often adequate in predicting outcome, in some cases grading is not sufficient long term as, for example, up to 20% of grade I tumours recur within 20 years of total resection (Harter et al, 2017). Demonstrating the need for improved and more accurate tumour grading to reduce the frequency of inappropriate therapy decisions in meningioma to improve patient outcome.

The genetic landscape of meningioma has been well characterised, with monosomy 22 and inactivation mutations of NF2 gene being the most reported and the first genetic anomalies to be correlated with meningioma development (Pećina-Šlaus et al, 2016). The remaining ~40% of meningiomas that do not carry a mutated NF2 gene but include mutations in TRAF7, AKT1, KLF4, PTCH1, PIK3CA, SUFU and PRKAR1A gene (Dunn et al, 2018). The recent development of high throughput genomic sequencing coupled with bioinformatic analyses has initiated research into detecting differentially expressed genes (DEGs) and their corresponding biological role within the different WHO grades to link particular genetic characteristics to a presenting phenotype. For example, Schmidt et al (2016) completed a transcriptomic analysis on the 3 WHO grades and clinical subgroups of 144 patients that were split into a discovery set (n=62) and a validation set (n=82). They identified many DEGs associated with the different subgroups and confirmed findings using the larger validation set and immunohistochemistry staining. The study concluded that the aggressiveness of the disease was associated with upregulation of PTTG1 and downregulation of LEPR.

By taking a bioinformatic approach to study the transcriptome of meningioma, the entire repertoire of RNA transcripts from a population of cells or tissues can be analysed simultaneously (Futschik et al, 2018). Measuring the transcriptome of a cell can be indicative of the relative gene expression taking place and can point to the relevant upstream regulation of expression. By taking a broader approach to understand gene expression instead of using targeted assays, it allows the investigation of coordinated trends that would have otherwise been lost. This gives a greater overview of gene interactions and networks and thus offers a more accurate representation of complex disorders, such as cancer (Lowe et al, 2017).

In the last decade, many non-protein coding sequences, accounting for ~98% of the genome, have been found to function in gene expression regulation such as cis-

regulatory regions and microRNAs (miRNAs). Recently through transcriptomic analysis, miRNA-200a has been identified to be downregulated by ~25-fold in meningioma and the overexpression of miRNA-109a has been associated with higher recurrence rates (Suppiah et al, 2019). Parallel with the identification of the role of theses regulatory sequences is the identification of non-coding mutations which have been implicated in cancer, proving the importance transcriptomics to analyse these regulatory mechanisms (Patel and Wang, 2018). Transcription factors (TFs) are proteins that bind to cis-regulatory regions to regulate gene expression; hence dysregulation of TF expression or binding have been identified in cancer (Patel and Wang, 2018). For instance, RNA sequencing identified FOXM1 TF as a driver for proliferation and a marker of poor outcome in meningioma (Vasudevan et al, 2018).

This project aimed to investigate the genomic landscape of the WHO classification grades to develop biomarkers to improve the identification of meningioma grades as well as to infer their biological role. It was hypothesised that there would be significantly DEGs between the WHO grades and aggression tumour groups Furthermore, a closer observation of the driving forces of aggressive meningioma was carried out, such that various biomarkers from DEGs with potential use in the prognosis and risk stratification of patients could be identified. Additionally, different regulatory mechanisms and their potential influence on gene expression, and thus the subsequent phenotype presented in meningioma tumours, were examined.

Material and Methodology

For the bioinformatic analysis of transcriptomic data, a variety of web-based tools were used that give biological insights from complex and extensive raw microarray gene expression data. The material and methodology will describe, in chronological order, the data and tools used and the underlying statistical and bioinformatic algorithms and theories.

Data Collection

Publicly available microarray data deposited on the NCBI Gene Expression Omnibus (GEO) database was used to select a relevant experimental publication (Edgar et al, 2002). Schmidt et al (2016) utilised the Illumina HumanHT-12 V4.0 expression beadchip microarray platform for measuring RNA.
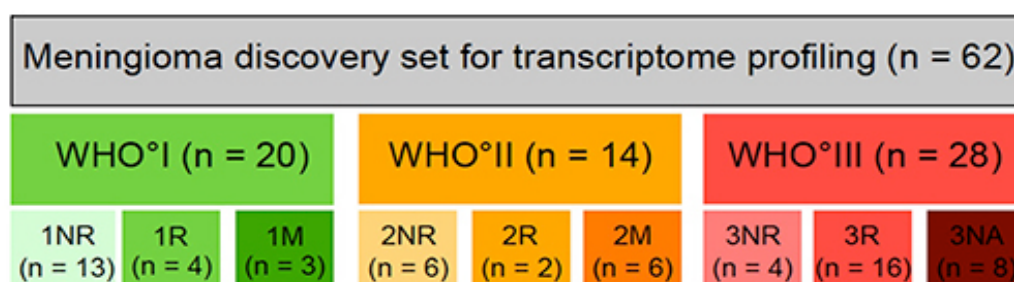

Figure 1: Sample distribution among WHO grade and clinical outcome subgroup, (From Schmidt et al, 2016 under CC BY 3.0). NR= tumours without any further recurrence in the 36 month period, R= subsequent recurrent tumour of the same WHO grade after complete resection, M= subsequent recurrent tumour of a higher WHO grade after complete resection, NA= no follow up data obtained or incomplete resection of tumour. n= the number of samples in a category

GEO2R

Normalised data deposited in the NCBI GEO database, GEO accession number GSE74385, was used in GEO2R differential gene expression analysis (DGEA) (Smyth, 2004). GEO2R online tools use a modified *t*-test implemented in the Bioconductor Limma (Linear Models for Microarray Analysis) package that overcomes the problem of small sample sizes in *t*-tests, thus allowing more complex experimental designs (Ritchie et al, 2015). Limma uses linear models to analyse microarray data for DGEA. Linear modelling is used to obtain the probability that two groups have the same mean, which indicates there are no DEGs and thereby suggests that there is insufficient evidence to reject the null hypothesis (Smyth, 2005). GEO2R for DGEA uses the GEO database to annotate Illumina probe IDs with non-redundant gene symbols.

The Benjamini-Hochberg method was used to correct for multiple testing and to obtain the false discovery rate (FDR) (Wu et al, 2016). A FDR of 0.05 represents 5% of declared DEGs being false positive, hence genes with an adjusted p-value >0.05 are deemed insignificant (Rogers and Weiss, 2009). Pairwise comparisons of the 3 WHO grades and aggression groups defined from clinical outcome were made. The Non-Aggressive group (n=23) includes the 23 non-recurrent cases over the 3 WHO grades. Aggressive group (n=39) includes recurrent (R) and malignant (M) of the 3 grades and the 3 not available (NA) of grade III as this grade has the highest recurrence rate of the WHO grades, Figure 2.

Genes outputted from GEO2R into an excel file were classed as significant DEGs based on FDR and were divided into upregulated and downregulated genes based on a positive or negative logFC value respectively. The LogFC is the log2-fold change between the two comparison groups that describes the factor with which the expression is increased or decreased (Pacholewska, 2017). The grade I versus grade III comparison (I v.s. III) DEGs were filtered further by reducing the significance threshold to adjusted p-value<0.01 due to the large number of IDs. The overall work flow of DEGs from GEO2R for grade I v.s. III is shown in Figure 3, with a similar workflow implemented for other comparisons. From the GEO2R results 'GO.Function' column was used in a preliminary investigation into potential TF. Filtered 'transcription factor' and then 'transcription factor activity' to identify any DEGs that function as TF.
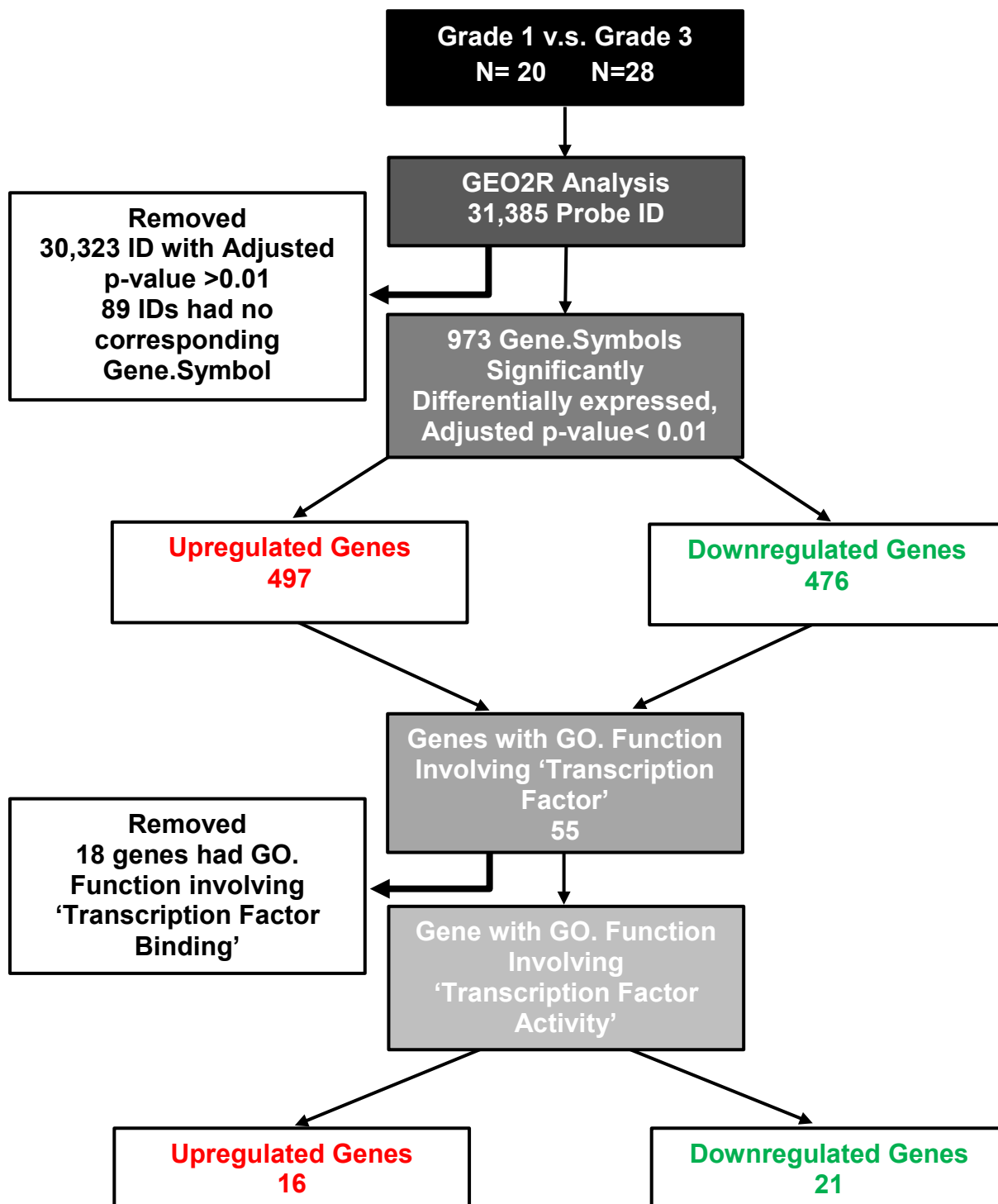
Figure 3: Workflow of GEO2R data. Red highlighted are upregulated genes and Green highlighted represent the downregulated genes. The filtering was completed within a Microsoft Excel file using a cut-off of adjusted p-value<0.01 to include only significantly differentially expressed genes and the LogFC to divide upregulated genes from downregulated genes. Other comparisons used a cut-off of adjusted p-value<0.05 instead. Further filtering was completed in a preliminary analysis to identify differentially expressed transcription factors using the GO.Function associated with each gene.

NetworkAnalyst
The expression matrix file from GEO was downloaded and the information tags were edited appropriately for the specific group pairwise comparisons and were uploaded to NetworkAnalyst 2019 version for DGEA and visualisation of results (Xia et al, 2015). Similarly to GEO2R, Limma is used for the DGEA with significance thresholds set as adjusted p-value <0.05 and Log2 fold change= 1. For the visualisation of DEGs between the comparison groups, heatmaps and 3D principle component analysis (PCA) clustering analysis was used. The KEGG pathway database was used to complete enrichment analysis on all identified DEGs, as well as within focused subsets of genes that were identified (via heatmaps) to have a relevant pattern of expression. PCA 3D analysis is a common unsupervised approach used to visualise high dimensional data. It is a mathematical algorithm where a multivariate dataset is linearly transformed whilst keeping maximum variation in the dataset. It reduces the dimension of the data, making is easier to visualise, and identifies the directions, known as principal components, of which a maximum of the variation of the data is explained. Visually, it might then be possible to identify whether particular groups overlap or form separate clusters (Metsalu and Vilo, 2015). Additionally, heatmaps represent individual expression values as colour gradients, such that each gene per sample is assigned a colour that relates to its expression level. This allows identification of visible clustering of genes with similar or different expression values (Zhao et al, 2014).

WebGestalt
WebGestalt (WEB-based Gene SeT AnaLysis Toolkit) 2017 version (Wang et al, 2017) was used to complete two types of functional enrichment analysis: over-representation analysis (ORA) and gene set enrichment analysis (GSEA).
In ORA, genes that were both within an annotated functional category, such as a pathway or biological process, and set of DEG are counted (Dong et al, 2016). The overlapping genes have their significance assessed by Fisher's exact test (Wang et al, 2017); therefore, ORA assesses whether genes from a particular pathway or biological process are overrepresented among the significantly DEGs (Dong et al, 2016). In contrast, GSEA ranks all genes based on the relative level of differential expression (Liu and Ruan, 2014). A maximum enrichment score (MES) is calculated from the list of ranked genes in a specific category - miRNA or TF motifs. An enriched p-value is calculated from the comparison of ranked MES to randomly generate MES distributions which assesses whether an annotated gene set has a tendency towards being upregulated or downregulated, or if they are randomly distributed along the ranked list (Tipney and Hunter, 2010).
The gene symbol list of DEGs was inputted into WebGestalt for ORA, where the upregulated and downregulated genes were analysed separately. The Illumina HumanHT-12 V.4.0 was selected as the reference set for analysis. Benjamini-Hochberg FDR at FDR<0.05 was used in the advanced parameters. For both the upregulated and downregulated gene lists ORA using Gene Ontology (GO) 2019 version (Ashburner et al, 2000), a functional database with computational models of biological systems, was used to investigate gene enrichment in the functional category 'Biological_Process'. Additionally, ORA of 'pathway' functional category using the KEGG pathway database was performed (Kanehisa et al, 2019). Kyoto Encyclopedia of Genes and Genomes (KEGG) is a manually constructed database which curates' models of pathways and biological processes.

GSEA was completed by inputting both the gene symbol list with the corresponding LogFC values of the DEGs. No reference gene list was needed for the analysis. Network 'miRNA_target' and 'Transcription_Factor_Target' functional categories were used which utilised the GSEA Molecular Signatures Database (MSigDB) 2017 version (Subramanian et al, 2005) to give insights into the regulatory mechanisms of gene expression of the two *trans*-regulators: miRNAs and TFs. MSigDB uses motif gene sets, which are a collection of genes that are grouped based on shared short sequence motifs. These motifs represent known or likely sequences within either their 3'-untranslated regions (3'-UTRs) or promoters that are potential binding targets of miRNA or transcription factors, respectively (Subramanian et al, 2005). The miRNA targets are based on miRbase (2005), a microRNA database that gathers miRNA sequences and annotations found within scientific publications. Targets of transcription factors are catalogued from Xie et al (2005) study.

X2KWeb
2018 version eXpression2Kinase Web (X2KWeb) was used to computationally predict potential TF influence on observed DEGs (Clarke et al, 2018). Transcription factor enrichment analysis (TFEA) was completed on the DEG list inputted into the tool to identify TFs that were most likely to regulate the DEGs and generate the mRNA expression levels observed. TFEA was completed using the ChIP-X database, a gene-list database compiled from 84 published experiments that report on the binding of specific TFs to DNA proximal to target gene loci (at the genome-wide level) using ChIP-seq and ChIP-chip methods (Lachmann et al, 2010). TFEA ranks individual TFs based on how likely they are to cause the observed gene expression and thus identifies the over-representation of specific TF targets within the gene symbol list. X2KWeb uses Fisher's exact test to complete statistical enrichment and Bonferroni's adjustment to account for multiple testing (Clarke et al, 2018).

Morpheus
Morpheus (2019) version was used to visually present the differentially expressed TFs identified from GO.Function filtering of GEO2R results within a heatmap with gene clustering. The expression matrix file was edited to only include the Ilumina probe IDs that corresponded to identified TF gene symbols and then uploaded to Morpheus. Both the grade and the level of aggression was annotated for each sample column and rows were annotated with gene symbols. Hierarchical clustering of the TF genes was completed using the average linkage method with a minimum and maximum distance of -1 and 1, respectively.

Results
To study dysregulation in meningioma and predict the upstream regulator mechanisms, publicly available microarray datasets generated by Schmidt et al (2016) were reanalysed. The aim is to identify groups of genes associated with pathways and processes that are atypical in subgroups of meningioma by completing enrichment analysis on the DEGs. Furthermore, I undertook a prediction of the influence of TF and miRNA by the observed expression of genes that contain specific target motifs.
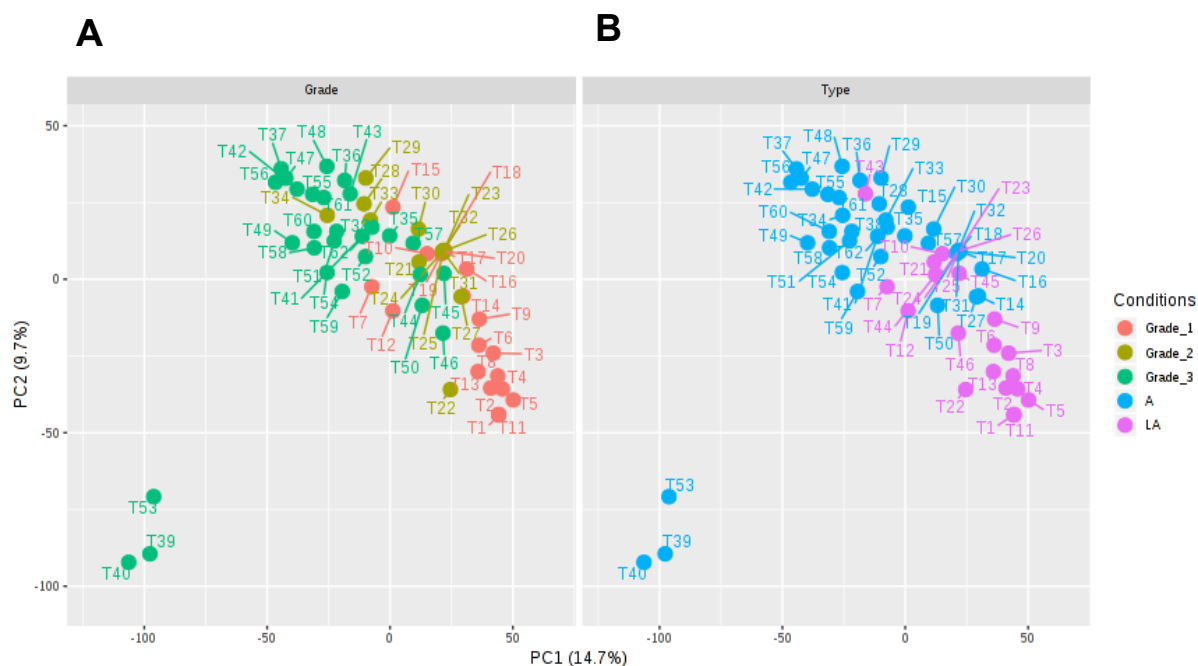
Quality Control
Data Distribution

For most microarray analyses, it is preferred that data from different samples share a similar distribution due to simplicity in downstream statistical analysis and interpretation. Figure 4 depicts a histogram of all the microarray data from the 62 tumour samples to evaluate the overall distribution of data prior to analysis with GEO2R. As seen below, there is little variance in the means and the standard deviations among the samples, hence it suggests that the data from different samples follow a similar general distribution.
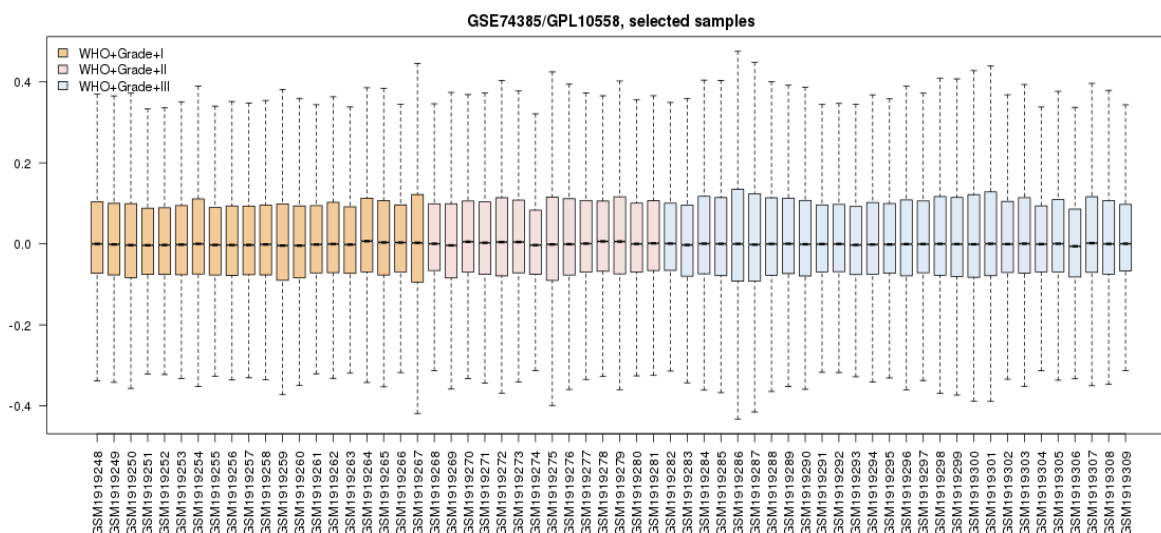
Figure 2: Histogram of the 62 microarray samples to check for normal distribution, histogram constructed using online tool GEO2R. 20 Orange are WHO Grade I, 14 Pink are WHO Grade II, and the 28 Blue are WHO Grade III.

Clustering of Data
Another quality control step taken was to perform cluster analysis on all the samples before filtering for DEGs, using PCA plot clustering to identify possible outliers where samples do not cluster with their WHO grade. Figure 5 A&B shows that T53, T40 and T39 (classified as 3R, 3NA and 3NA, respectively) were separated from the main cluster of samples and may indicate potential sample contamination (Zhao et al, 2014). However, heatmap cluster analysis based on DEGs found the 3 samples were all clustered in the largest aggressive cluster shown in Figure 6, therefore they were kept and used in downstream analysis.

9

Figure 3: A&B- PCA clustering plot for quality control of the two comparisons, 20 Red- Grade I, 14 Brown-Grade II, 28 Green- Grade III, 39 Blue- Aggressive, 23 Pink- Non-Aggressive. Constructed using online tool NetworkAnalyst.

In Figure 5, the samples of the same grade did appear to be located in close proximity to each other, with grade II samples situated in the middle of the main cluster and both grade I and grade III tumour samples being at opposite ends of the cluster. However, there is no clear distinction between the grades.

Figure 6 shows the clustering of samples based on gene expression of significantly differentially expressed cell cycle enriched gene of the non-aggressive and aggressive comparison. The largest cluster consists of aggressive samples, whereas the non-aggressive samples group into smaller clusters. This suggests a greater heterogeneity in gene expression levels in the non-aggressive samples. Figure 6 also illustrates the expression levels of genes that were found through ORA to be enriched in the cell cycle pathway. From the clustering within the aggressive samples there is upregulation in the genes involved in cell cycle compared to the non-aggressive group which are frequently downregulated.

Differentially Expressed Genes

By completing GEO2R differential gene expression analysis, DEGs at a significance level of adjusted p-value<0.05 were found for the comparisons of I v.s. III and NA v.s. A alone. Further filtering of I v.s. III comparison resulted in 973 genes (adjusted p-value<0.01) whereas NA v.s. A comparison resulted in 1426 genes (adjusted p-value<0.05). Figure 7 shows the distribution of the DEGs and the 498 significantly upregulated and 475 downregulated genes from the I v.s III comparison. Figure 8 shows 818 significantly upregulated genes and 763 significantly downregulated genes in the NA v.s. A comparison.
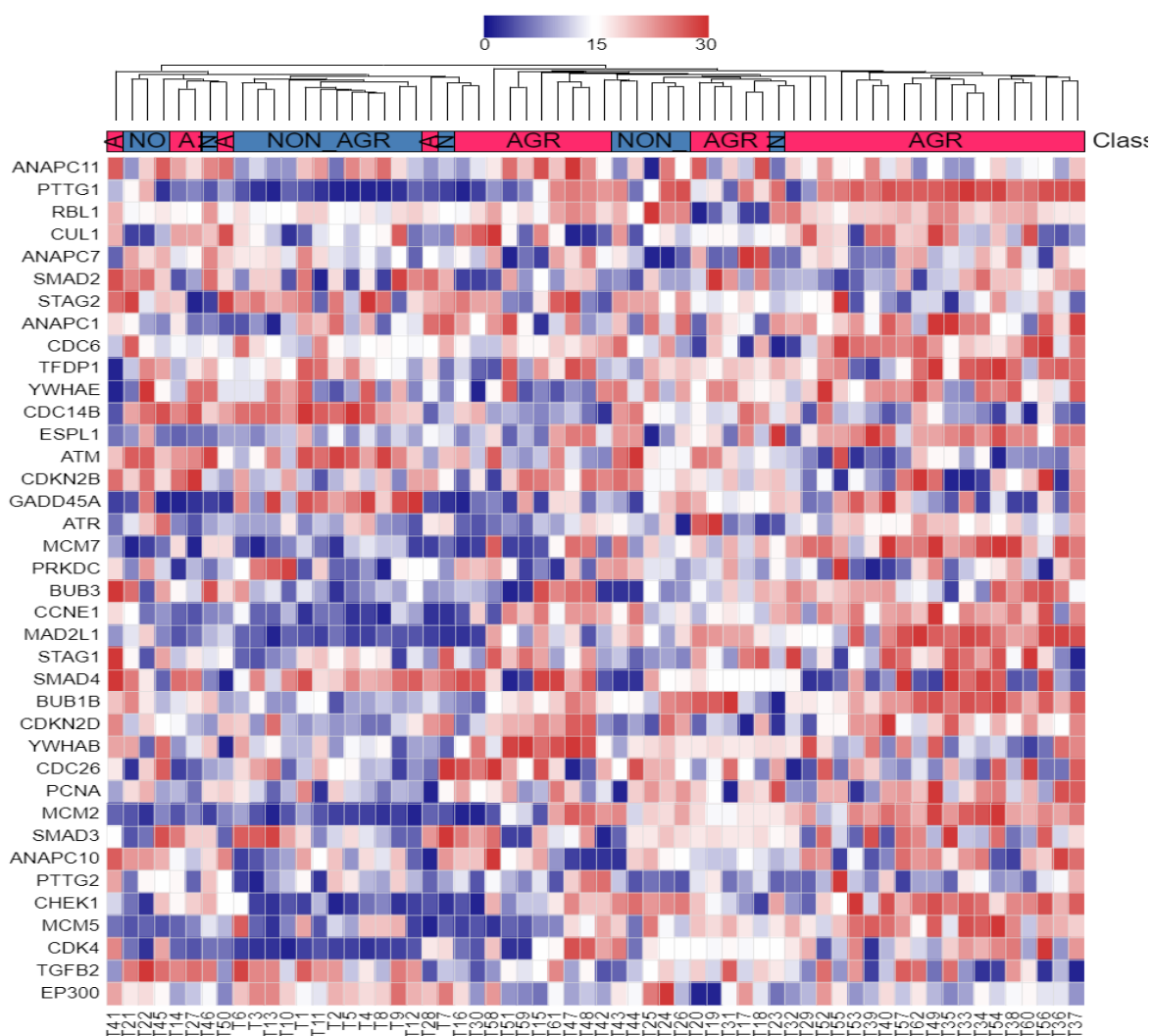
Figure 4: Clustered heatmap of genes that are differentially expressed between non-aggressive and aggressive tumours and that are associated with the cell cycle based on ORA enrichment analysis. Colour spectrum correlates to level of gene expression such that red denotes upregulation, blue denotes downregulation. The Network Analyst ORA produce multiple heatmaps of enriched biological processes, which shows the up and down expression of genes from the non-aggressive versus aggressive tumour comparison. Biological process 'Cell Cycle' had the most enriched genes (n=38, p-value <0.05) and therefore is included in this paper. The block of colours at the top indicate the clustered groups of the two different classes. Samples T53, T40 and T39 are highlighted by an asterisks (*), pink represents the aggressive class and blue represents the non-aggressive class of tumour samples.

Table 1 and Table 2 present the genes achieved the highest significance. Interestingly, upregulation of *PTTG1* and downregulation of *MAN1C1, ANAPC16* and *LEPR* was found in both I v.s. III and NA v.s. A comparisons. For a complete DEG list excel file contact author. From Table 1, the upregulated genes, *TPX2* and *ASPM,* associated with microtubule functional and are involved in cell cycle; however, *PRC1* is upregulated whereas *ANAPC16* is downregulated. A similar divergence of cell cycle genes can be found in Table 2 with *CDCA5* upregulated and *ANAPC16* downregulated. A literature review revealed that many common genes associated previously with meningioma were not included in the DEG list with only 1 identified in

I v.s. III comparison, *SFRP1*, and 2 in the NA v.s. A comparison, *SFRP1* and *PIK3CA* indicating the complementary information contained in the microarray experiment. A note of caution: When completing DGEA using NetworkAnalyst, there were discrepancies between the lists of DEG generated. NetworkAnalyst tool identified the *PTEN* gene, known for its association with multiple cancers, to be significantly downregulated (adjusted p-value= 0.008041); this was not identified in GEO2R analysis.

Table 1: Grade I v.s Grade III differential gene expression analysis results, the 5 most significantly upregulated and downregulated genes are shown in red and green, respectively. The 3 genes in bold are genes identified through a literature search that have been previously reported in meningioma, the symbol * indicates key genes identified in original Schmidt et al (2016) study.

Gene Symbol	Gene Title	Adjusted P value	LogFC
OIP5	Opa interacting protein 5	0.00000436	0.522
TPX2	TPX2, microtubule nucleation factor	0.00000436	1.35
PRPF3	pre-mRNA processing factor 3	0.00000441	0.743
PRC1	protein regulator of cytokinesis 1	0.00000456	1.9
ASPM	abnormal spindle microtubule assembly	0.00000805	1.79
PNRC2	proline rich nuclear receptor coactivator 2	0.00000436	-0.712
ANAPC16	anaphase promoting complex subunit 16	0.00000441	-0.697
CRYL1	crystallin lambda 1	0.00000789	-0.778
ZC2HC1C	zinc finger C2HC-type containing 1C	0.00000883	-2.23
MAN1C1	mannosidase alpha class 1C member 1	0.00001038	-1.62
PTTG1*	**pituitary tumor-transforming 1**	**0.00007359**	**0.932**
LEPR*	**leptin receptor**	**0.000477**	**-2.14**
SFRP1	**secreted frizzled related protein 1**	**0.00689228**	**-2.3**

Table 2: Non-Aggressive v.s. Aggressive differential gene expression analysis results, the 5 most significantly upregulated and downregulated genes are shown in red and green, respectively. The 4 genes in bold are genes identified through a literature search that have been previously reported in meningioma, * key genes identified in original Schmidt et al (2016) study.

Gene Symbol	Gene Title	Adjusted P value	LogFC
RCN1	reticulocalbin 1	0.00009055	1.342013
CDCA5	cell division cycle associated 5	0.00011333	0.509429
IMP4	IMP4 homolog, U3 small nucleolar ribonucleoprotein	0.0001998	0.760898
PRELID1	PRELI domain containing 1	0.0001998	0.567713
ADRM1	adhesion regulating molecule 1	0.0001998	0.342146
FAM167B	family with sequence similarity 167 member B	0.00002581	-0.3384
MAN1C1	mannosidase alpha class 1C member 1	0.00009507	-1.34947
ANAPC16	anaphase promoting complex subunit 16	0.000113	-0.56758
YPEL2	yippee like 2	0.000142	-0.61066
RAVER2	ribonucleoprotein, PTB binding 2	0.0002	-0.70577
PTTG1*	**pituitary tumor-transforming 1**	**0.00443553**	**0.670423**
LEPR*	**leptin receptor**	**0.00042187**	**-2.02564**
PIK3CA	**phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha**	**0.043989**	**-0.14251**
SFRP1	**secreted frizzled related protein 1**	**0.003253**	**-2.0098**

Volcano Plot of Differentially Expressed Genes, Grade I v.s. Grade III
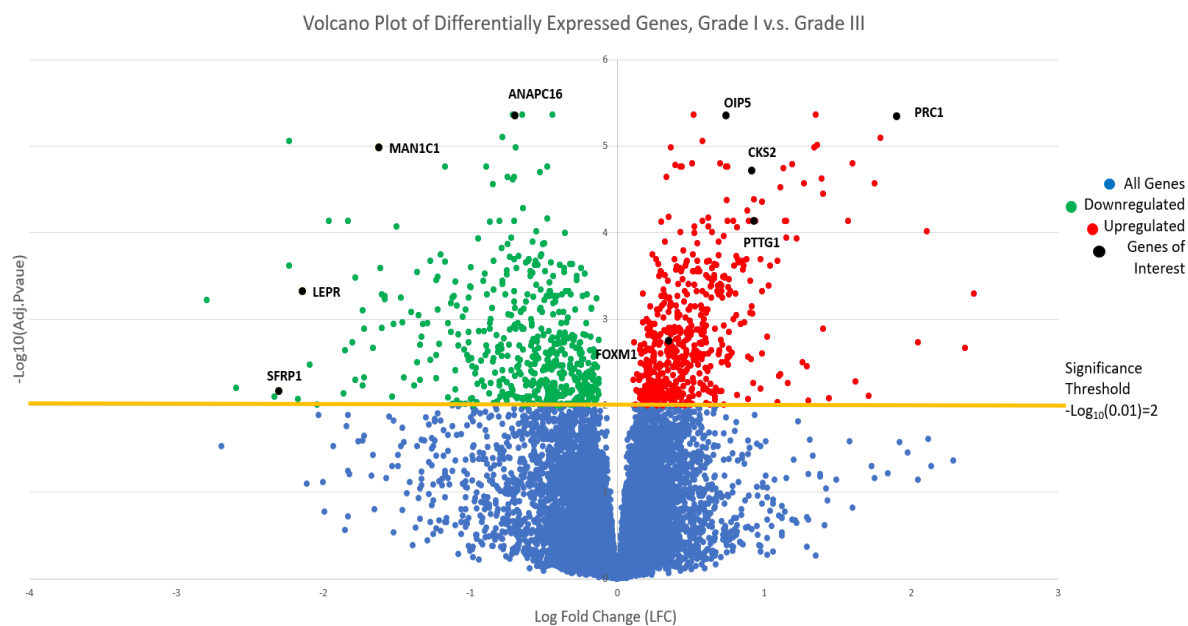
Figure 5: Volcano plot of differentially expressed genes from Grade I v.s. Grade III comparison, significantly (adjusted p-value<0.01) upregulated and downregulated genes are highlighted in red and green, respectively. Interesting genes, either are the top 5 up or downregulated genes or found in a literature search, which are highlighted in black and have been labelled with gene symbol.

Volcano Plot of Differentially Expressed Genes, Non Aggressive v.s.Aggressive
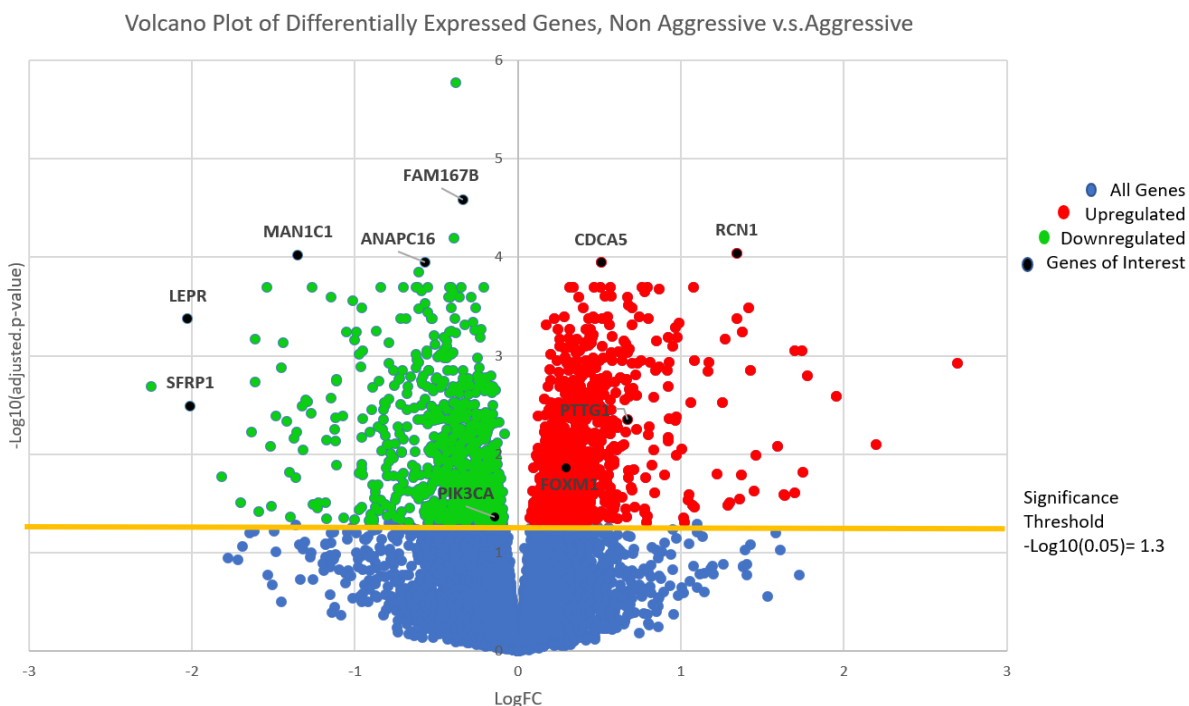
Figure 6: Volcano plot of differentially expressed genes from Non-Aggressive v.s. Aggressive comparison, significantly (adjusted p-value<0.05) upregulated and downregulated genes are highlighted in red and green, respectively. Interesting genes, either are the top 5 up or downregulated genes or found in a literature search, which are highlighted in black and have been labelled with gene symbol.

Over Representation Analysis

ORA using GeneOntology of DEGs from I v.s. III comparison identified significantly overrepresented biological processes within the upregulated genes but not in the downregulated genes. Also, no downregulated genes were enriched in the KEGG pathway ORA.
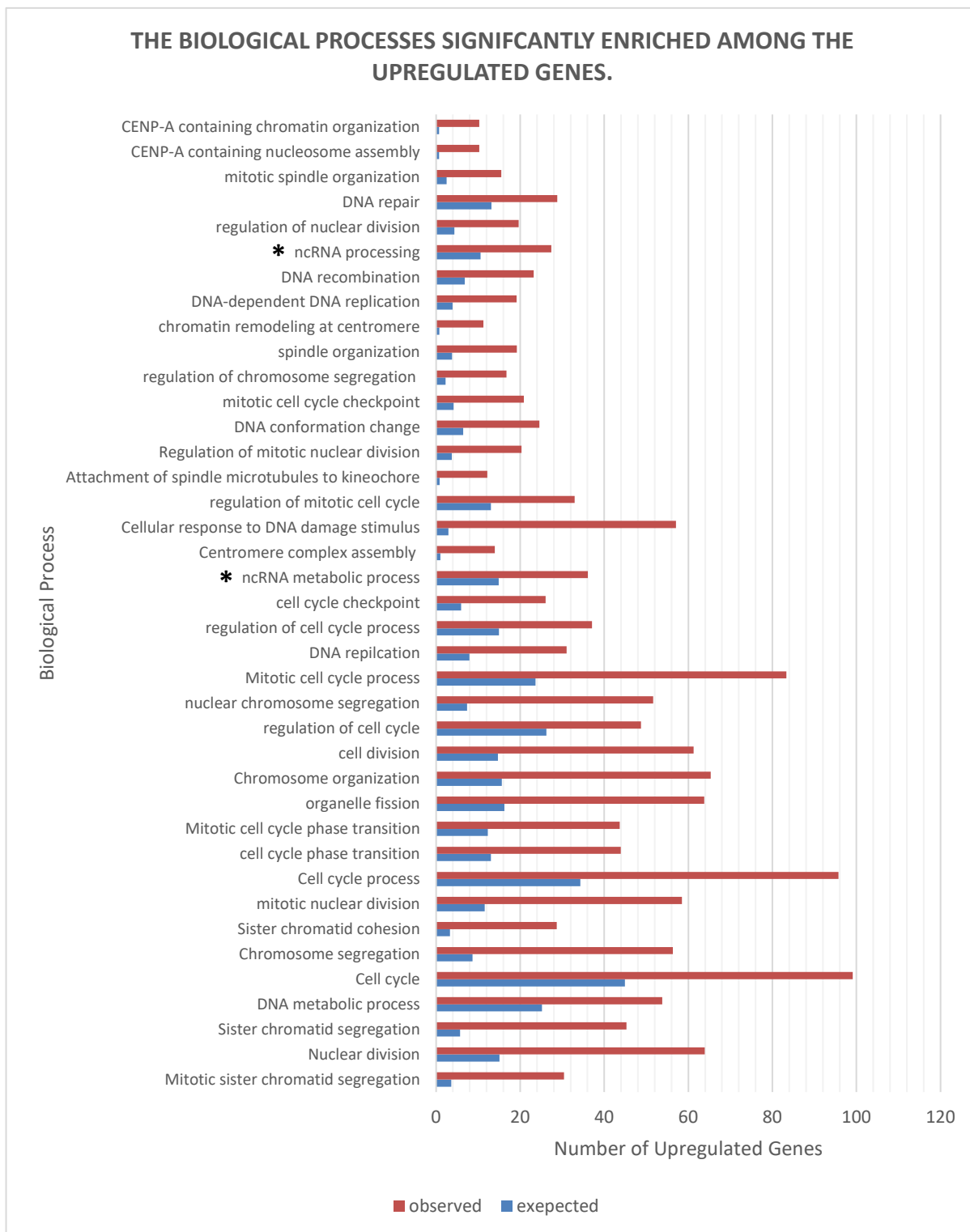

Figure 7: Bar Chart illustrating the ORA GeneOntology results of grade I v.s. grade III, the 39 enriched biological pathways are shown with the number of gene observed in the upregulated gene list and the number of genes that would be expected based on size of

gene list is given in orange and blue, respectively. * The 2 biological processes that are overrepresented but are not a subcategory of cell cycle or mitosis.
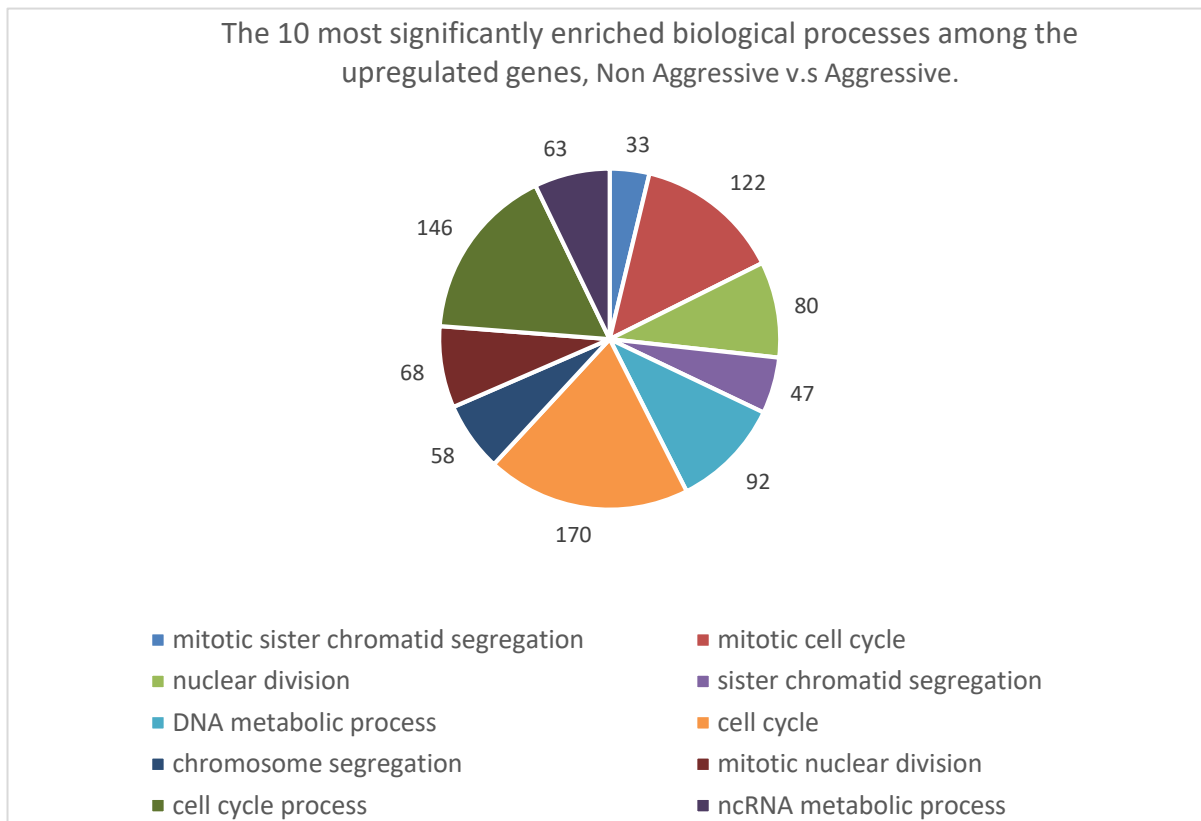
9 biological processes were overrepresented among the upregulated genes. However only 2 biological processes, 'ncRNA metabolic process' and 'ncRNA processing', were not involved in any aspect of mitosis and cell cycle as indicated by asterisk (*) in Figure 9. This trend was also found in the KEGG pathway analysis as of the 8 pathways found to be over represented within the upregulated genes, the second most significantly enriched was the cell cycle pathway, see Figure 10. Figure 9 shows that 'cellular response to DNA damage stimulus' had the largest odd ratio as it had the greatest difference in the number of expected (3) and observed genes (57). This was also found in ORA of pathways as Fanconi anaemia pathway which functions in DNA damage repair, was reported to be over represented and had 18 genes associated in the pathway, Figure 10. In contrast metabolic pathways were identified as the most significantly overrepresented pathway as there were 59 annotated associated genes however this was not identified in the ORA biological process analysis.
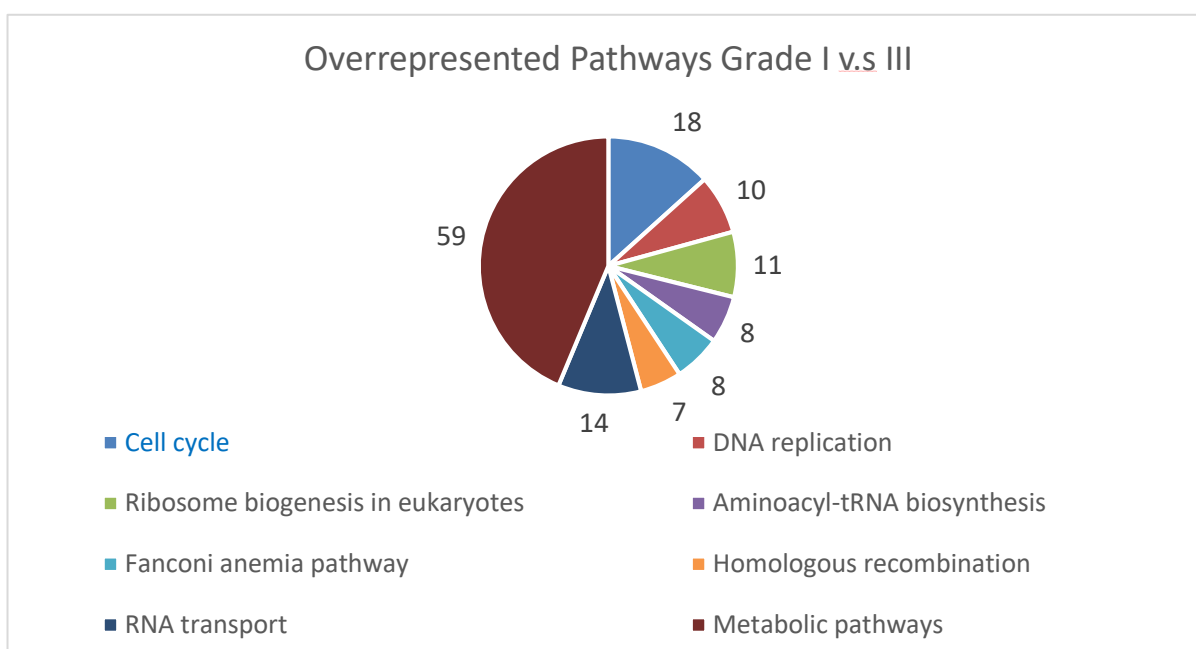

Figure 8: Pie Chart of ORA KEGG pathway results of the upregulated genes from Grade I v.s. Grade III comparison. The surrounding numbers display the number of observed genes associated with the specific pathway.

ORA of the DEGs from the NA v.s. A comparison identified no overrepresented downregulated genes within Gene Ontology biological processes or KEGG pathway at a significance level of adjusted p-value<0.05. ORA of the upregulated genes identified 302 enriched categories, of which the 10 most significantly enriched categories were analysed further, see Figure 11. 9 out of 10 enriched processes were sub-groups of cell cycle, for example 'DNA metabolic process' and 'nuclear division', which are needed for proliferation of cells, with 'ncRNA metabolic processes' being the exception. This trend is highlighted in Figure 12 as the hierarchical structure shows a clustering of linked biological processes involved in cell cycle. From the ORA using the KEGG database there were 11 pathways that were significantly overrepresented. Similarly, to the Gene Ontology enrichment, cell cycle was the most significant result with 20 associated upregulated genes.

Interestingly spliceosomes classed as non-coding RNA (ncRNA), were also identified from KEGG analysis, with 17 associated upregulated genes. From Figure 11 and Figure 12 of the ORA biological process analysis, ncRNA metabolic processes was found to be significantly enriched, therefore furthering the significance of spliceosomes as an enriched pathway in the upregulated genes of aggressive compared to non-aggressive tumours.

Figure 9: ORA Gene Ontology results of the upregulated genes from Non-Aggressive v.s. Aggressive comparison. Only the 10 most significantly overrepresented are reported as the analysis identified 302 enriched biological processes. The surrounding numbers display the number of observed genes associated with the specific biological process.
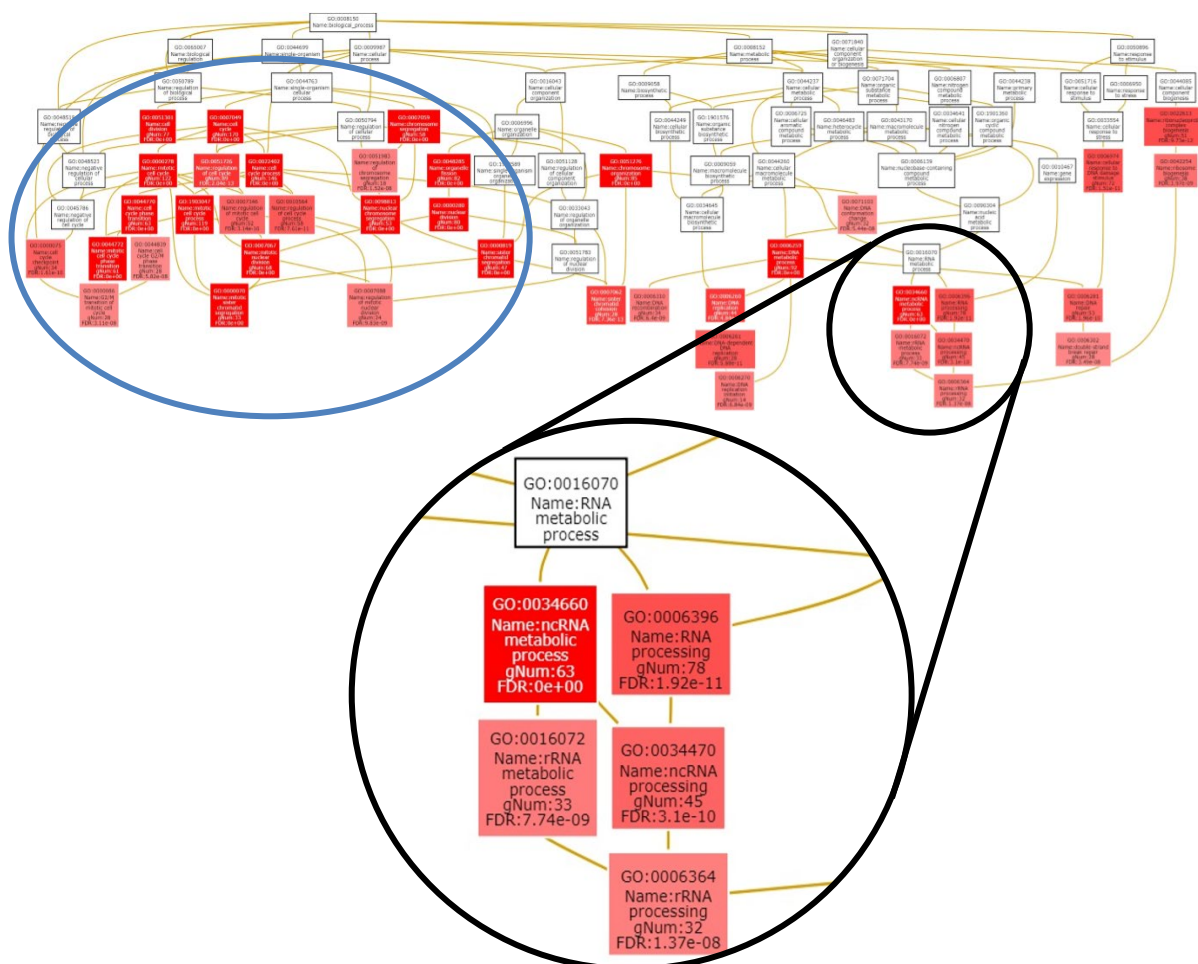
Figure 10: ORA of biological process results as shown in an inverted tree structure to represent Gene Ontology were biological processes are placed in a hierarchical structure. For ease of investigation I have circled and enhanced the biological processes of interest. The different shades of red correlate to significance of enrichment, ncRNA metabolic process is in a darker shade of red indicating is more significantly enriched than the surrounding and related biological processes. The blue circle indicates the main cluster of biological processes under the broad category of cell cycle process.

Influence of Regulatory Mechanisms

Transcription factor influence was investigated by collating results from GO.Function filtering, GSEA enrichment using 'Transcription_Factor_Target' functional category and X2KWeb which the upregulated and downregulated genes were analysed separately. MiRNA influence was described from GSEA enrichment using 'MiRNA_Target' functional category results. See Table 3 for an overview of results and Supplementary Material 1 & 2 for complete results.

Table 3: Results of transcription factor and miRNA influence analysis. 3 different approaches on the differentially expressed gene list are shown with the number of transcription factors (TF) and miRNA identified using each approach within the two comparison groups.

	Grade I v.s. Grade III	Non-aggressive v.s aggressive
GO.Function filtering (TF)	37	83
GSEA (TF)	20	20
GSEA (miRNA)	20	20
X2KWeb upregulated (TF)	103	104
X2KWeb downregulated (TF)	104	103

When examining all methods applied for the investigation of regulatory mechanisms a trend in the results is the identification of a family of transcription factors, Forkhead Box (FOX). There were 10 and 12 instances of FOX involvement recorded in the I v.s. III comparison and NA v.s A comparison, respectively. Using the X2KWeb analysis, the family of TFs were found to have targets in the upregulated and downregulated gene list, indicating both excitatory and inhibitory regulatory roles. A further trend within the TF results was the identification of Pax4 TF in both comparison groups. Figure 13 shows two clusters; the first cluster consists of downregulated and upregulated TF genes in grade I and grade III, respectively, whilst the converse observation was found in the second cluster. There are two sub clusters that show the most significant DEGs as shown by the deep colour gradient, 15 and 23 TF genes are shown in sub cluster 1 and 2, respectively. Similar to Figure 6, Figure 13 shows different TF expression between the 3 grades but also between the clinical aggression groups. *FOXM1* TF is shown to be upregulated in the higher grades but also, based on the variation of expression in the heatmap, it is upregulated in the aggressive group with the same expression pattern identified for the *PTTG1* gene. From the miRNA results there were 7 miRNAs (highlighted green in Supplementary Material 1 and 2) identified in both comparison groups however no other significant trends were identified. The influence these TF and miRNA preliminary results have on gene expression level will be evaluated in the discussion below along with the investigation of potential biomarkers of aggression from the DEG analysis results.
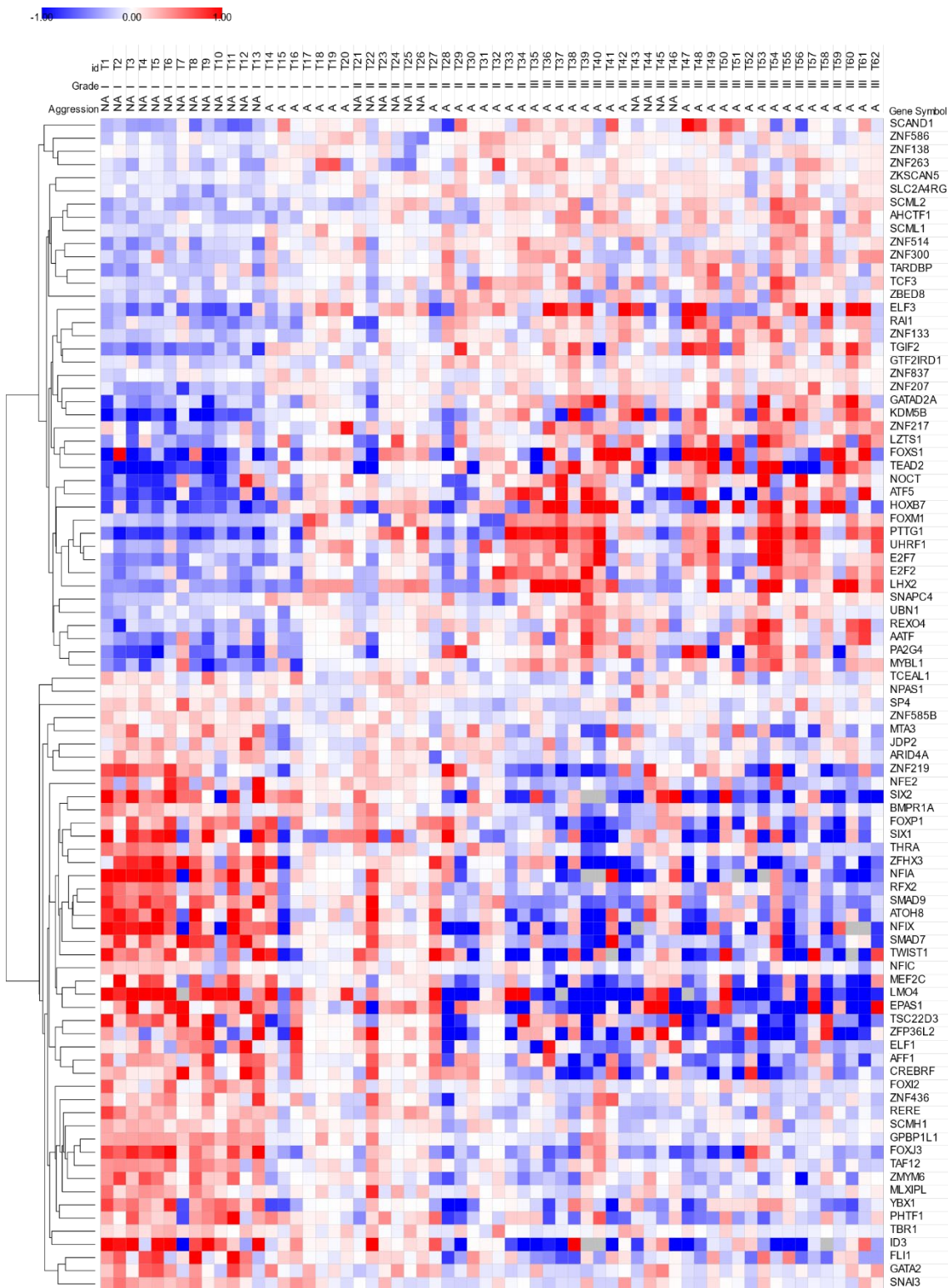
Figure 11: Heatmap Cluster Analysis of DEGs encoding TFs. Colour spectrum correlates to level of gene expression. Red is upregulated and blue is downregulated. Heatmap created using Morpheus online tool and edited expression matrix file. Cluster analysis shows two main clusters that show opposite TF gene expression and sub clusters based on significantly DEGs that show differences between WHO grades and clinical groups.

Discussion

Schmidt et al (2016) conducted the original microarray study of which the data was publicly available. This cohort was mainly chosen because it included a large cohort of grade 3 meningioma patients. Although they not representative of true prevalence, the balanced numbers allowed statistical comparisons to be made. Furthermore, the choice of storage of the tumour biopsy samples used in this study was advantageous as they were stored at -80°C instead of formalin-fixed paraffin-embedded (FFPE) processed samples, which can lead to fragmentation of DNA and RNA (Choi et al, 2017). A limitation of Schmidt et al (2016) study was that they only described the gene transcripts in the different grades. I have complemented this with a system bioinformatics approach for the prediction of causative mechanisms behind the described gene transcripts by investigating the role of regulatory mechanisms, a novel aspect of analysis of this data set.

The discrepancies in identified DEG sets from the two independent online tools, GEO2R and NetworkAnalyst, illustrates the influence of bioinformatic platform used in analysis and is due to the differences within the Illumina probe annotation files. This give additional confidence in the DEGs identified by two independent computational platforms and is an advantage of my method design as shows that neither of the tools are comprehensive in identifying DEGs and should therefore be used to complement each other. Benjamini-Hochberg adjustment was used to control FDR over the Bonferroni adjustment for the significant threshold of DEGs and to account for multiple testing (Chen et al, 2017), because Bonferroni adjustment tends to be too stringent in the case of transcriptomic analysis. However, the FDR method assumes independence of tests, which in the context of gene expression leads to the assumption that DEGs identified work independently (Stevens et al, 2017). Therefore, enrichment analysis was performed to rectify this, and has advantages over a ranked list of genes from DGEA as it is more representative of the complexities of interacting gene groups as typically one gene has no independent influence on a phenotype (Tipney and Hunter, 2010). A limitation of enrichment analysis was that I did not account for the non-independent functional biological process categories and therefore all genes enriched within a specific category will also be enriched in the more general category. This artificial increase in number of linked significant categories maybe an explanation for the number of enriched categories associated in cell cycle processes found in both comparisons. This could have been controlled for by considering a specific size of categories or by using the 'Biological_process_non_redundant' functional category in WebGestalt. Overall the PCA did not show clear separation in the global transcriptome between the grades or clinical aggression groups and therefore, to separate the different grades and predict aggression, the identification of particular DEGs is required. A plethora of conclusions can be drawn from the unbiased examination of generated results. I will, however, focus on a few conclusions that offer the potential as biomarkers or have relevance in meningioma pathogenesis as they make biological sense or have been previously reported to have influence in meningioma or cancer in general. Further experimental research will be needed to confirm the influence of identified potential biomarkers, for example knockout experiments to investigate the role of the genes and immunohistochemical staining to confirm observed expression changes translate onto protein level (Schmidt et al, 2016).

A common trend from enrichment analysis over both comparison groups and functional category was the over representation of upregulated genes involved in cell

cycle processes. This indicates that cell cycle process is upregulated in grade III and aggressive tumours compared to grade I and non-aggressive, respectively. Limitless replicative potential is a hallmark of cancer and through atypical proliferation, which required both increased cell cycle and loss of regulatory check points, generates a mass of cell that constitutes a tumour (Hanahan and Weinberg, 2000). This increased cell cycle leads to progressive erosion of telomeres, with shortened telomeres being linked to high tumour grades (Chan et al, 2001), and culminates in chromosomal fusions and instability. Therefore, increased cell cycle could explain the chromosomal alteration in meningioma such as del(1p36) that are associated with higher grades and indicate poor outcome (Pećina-Šlaus et al, 2016).

From Schmidt et al (2016) two novel prognostic biomarkers independent of WHO grade were identified, *LEPR* and *PTTG1*, and malignancy-associated protein expression changes were confirmed via protein staining. From DEGs I also identified a downregulation of *LEPR* and upregulation of *PTTG1* in aggressive and grade III groups. Downregulation of *LEPR*, encoding for leptin receptor, has been previously identified as a biomarker of recurrence in meningioma in Menghi et al (2011) microarray gene expression analysis study. Although the influence of *LEPR* in meningioma is unknown, knockout experiments of *LEPR* in rats results in leptin resistance, a characteristic of obesity, and whole-exome sequencing identified multiple *LEPR* mutations within obesity patients (Picó et al, 2002 and Gill et al, 2014). Notably, obesity has been found to be significantly increased within meningioma patients and therefore suggest a relationship. *PTTG1*, pituitary tumour-transforming gene-1, is a novel gene in meningioma but is a known oncogene in other cancer with its expression levels correlating with a degree of malignancy (Schmidt et al, 2016). Yoon et al (2012) found *PTTG1* was highly expressed in breast cancer patients and enhanced migratory and invasive behaviour of the tumour by inducing epithelial to mesenchymal transition (EMT). *PTTG1* increases Snail, Slug, Twist and Zeb1 TF which have activator and repressor function on genes that results in morphological changes associated with EMT. Within X2KWeb results both Zeb1 and Twist1 TF were identified to have targets in both up and down regulated DEGs and therefore indicate the involvement of *PTTG1* in the more aggressive meningiomas.

Spliceosome are a novel aspect in meningioma pathogenesis and were discovered to be an enriched pathway in the upregulated genes of aggressive v.s. non-aggressive samples with 17 genes identified. Serine and arginine rich splicing factor 6, *SRSF6*, has been associated in progression of breast cancer with epidemiological and experiment evidence highlighting the involvement of estrogen and estrogen receptor (ER) in the splicing factors function. In normal cells, estrogen reduces levels of *SRSF6* which causes an exon skipping RNA splicing event of $CRH\text{-}R_1$ gene, a hormone receptor, and the alternative splicing affects the receptor functionality leading to the suppression of estrogen-induced cell proliferation (Silipo et al, 2015). The influence of hormones is shared in both meningioma and breast cancer progression and from the results I infer that in aggressive meningioma the upregulated *SRSF6* gene appropriately splices the $CRH\text{-}R_1$ gene and thus the receptor functions and leads to cell proliferation and tumour growth which occurs in more aggressive tumours. Potentially the ER located on some meningioma tumours act to sequester the function of estrogen to inhibit *SRSF6* expression and therefore allow the downstream signalling resulting in cell proliferation and more aggressive presentation (Lal et al, 2013). The identification of *SRSF6* role in cell proliferation and its regulation is a potential mechanism behind the correlation of ER positive

meningioma patients with poor prognosis compared to ER negative patients (Hua et al, 2018).

Many differentially expressed TF or TF that had targets within the DEG set can be further classed as pioneer factors that influence gene expression by targeting nucleosome DNA. They cause chromatin remodelling, for example chromatin opening, so many previously inaccessible genes can be expressed and thus greatly alter the cell-fate (Iwafuchi-Doi, 2018). FOX, forkhead box, TF family are prime examples and were found frequently in my results. Vasudevan et al (2018) transcriptomic analysis of RNA-Seq identified *FOXM1* as a key TF in meningioma proliferation and associated with poor prognosis. *FOXM1* functions to dysregulate the Wnt signalling pathway to drive proliferation and tumour growth. This Wnt dysregulation has been documented in meningioma previously and is a trend identified in my DEG results. *FOXM1* interacts directly with beta-catenin to transduce Wnt signals but also suppresses antagonists to have sustained Wnt signalling and proliferation. *SFRP1* supresses the Wnt pathway however this was identified in results as significantly downregulated. This was also identified in Vasudenven et al (2018) and was suggested to be epigenetically silenced through the hypermethylation of *SFRP1* promotor. Both the results of my transcriptomic analysis and other transcriptomic studies found *FOXM1* as a potential biomarker of meningioma aggression and indicated its role in the epigenetic silencing of *SFRP1* gene, a further biomarker in meningioma.

Of the miRNA identified no significant trends were found with only one miRNA, mir-145, having been previously identified to have influence in meningioma. However, the predicative function of mir-145 inferred from observed DEGs does not correlate with previous findings that mir-145 levels are decreased in higher grades (Galani and Lampri, 2017). Although no conclusive evidence of the influence of miRNA in meningioma aggression was found in my analysis this avenue of research has great potential and further research is needed. Many non-coding regions are considered pharmaceutically 'druggable' and therefore have application in reducing meningioma burden (Amodio et al, 2017).

Conclusion

In conclusion by taking a bioinformatic approach for the description of DEGs and the prediction of regulatory mechanisms I have identified multiple potential biomarkers of aggression in meningioma. These will aid classification of tumour grade and therefore be a more accurate tool for prognosis prediction. Furthermore, the biological role of novel genes and pathways has been inferred and adds to the understanding of the pathogenesis of meningioma. Finally, by investigating the upstream regulatory mechanism of abnormal gene expression, therapies can be designed to target these to correct the dysregulated gene expression and to then reduce the burden of meningioma and improve patients' health.

References

Amodio. N, Raimondi. L, Juli. G, Stamato. M, Caracciolo, D, Tagliaferri. P, Tassone. P (2018) 'MALAT: a druggable long non-coding RNA for targeted anti-cancer approaches', *Journal of Hematology and Oncology*, 11(**63**), pp. 1

Ashburner. M, Ball. C, Blake. J et al (2000) 'Gene Ontology: tool for the unification of biology', *Nature Genetics*, 25(**1**), pp. 25-29.

Chen. H, Cho. C, Liang. C et al (2001) 'Differential telomerase expression and telomere length in primary intracranial tumors', *Chang Gung Medical Journal*, 24(**6**), pp352-360.

Chen. S, Feng. Z and Yi. X (2017) 'A general introduction to adjustment for multiple comparisons', *Journal of Thoracic Disease*, 9(**6**), pp. 1725-1729.

Choi. Y, Kim. A, Kim. J, Lee. J, Lee. S and Kim. C (2017) 'Optimization of RNA Extraction from Formalin-Fixed Paraffin-Embedded Blocks for Targeted Next-Generation Sequencing', *Journal of Breast Cancer*, 20(**4**), pp. 393-399.

Clarke. D, Kuleshov. M, Schilder. B, Torre. D, Duffy. M, Keenan. A, Lachmann. A, Feldmann. A, Gundersen. G, Silverstein. M, Wang. Z and Ma'ayan. A (2018) 'eXpression2Kinase (X2K) Web: linking expression signatures to upstream cell signalling networks', *Nucleic Acids Research*, 46(**1**), pp. 171-179.

Desai. P and Patel. D (2016) 'A study of meningioma in relation to age, sex, site, symptoms, and computerized tomography scan features', *International Journal of Medical Science and Public Health*, 5(**2**), pp. 331-334.

Dong. X, Hao. Y, Wang. X and Tian. W (2016) 'LEGO: a novel method for gene set over-representation analysis by incorporating network-based gene weights', *Scientific Reports*, 6(**1**), pp. 1-17.

Dunn. J, Ferluga. S, Sharma. V, Futschik. M, Hilton. D, Adams. C, Lasonder. E and Hanemann. C (2018) 'Proteomic analysis discovers the differential expression of novel proteins and phosphoproteins in meningioma including NEK9, HK2 and SET and deregulation of RNA metabolism', *EBioMedicine*, 1(**1**) pp. 1.

Edgar R, Domrachev M, Lash AE (2002) 'Gene Expression Omnibus: NCBI gene expression and hybridization array data repository', *Nucleic Acids Research,* 30(**1**), pp. 207-210

Futschik. M, Morkel. M, Schäfer. R and Sers. C (2018) 'Chapter 7 The Human Transcriptome Implications for Understanding, Diagnosing, and Treating Human Disease' in Coleman. W and Tsongalis. G, *Molecular Pathology*, Cambridge: Academic Press, pp. 135-164.

Galani. V, Lampri. E, Varouksi. A, Alexiou. G, Mitselou. A and Kyritsis. A (2017) 'Genetic and epigenetic alterations in meningioma', *Clinical Neurology and Neurosurgery,* 158(**1**), pp. 119-125.

Gill. R, Cheung. Y, Shen. Y et al (2014) 'Whole-exome sequencing identified novel LEPR mutation in individuals with severe early onset obesity', *Obesity*, 22(**2**), pp. 576-584.

Gurcay. A, Bozkurt. I, Senturk. S, Kazanci. A, Gurcan. O, Turkoglu. O and Beskonakli. E (2018) 'Diagnosis, Treatment, and Management Strategy of Meningioma during Pregnancy', *Asian Journal of Neurosurgery*, 13(**1**), pp. 86-89.

Hanahan. D and Weinberg. R (2000) 'The Hallmarks of Cancer', *Cell*, 100(**1**), pp. 57-70.

Harter. P, Braun. Y and Plate. K (2017) 'Classification of meningiomas-advances and controversies', *Chinese Clinical Oncology*, 6(**1**), pp. 1.

Hua. L, Zhu. H, Li. J et al (2018) 'Prognostic value of estrogen receptor in WHO Grade III meningioma: a long-term follow-up study from a single institution', *Journal of Neurosurgery*, 128(**6**), pp. 1698-1706.

Iwafuchi-Doi. M (2018) 'The mechanistic basis for chromtin regulation by pioneer transcription factors', *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 11(**1**), pp. 1.

Kanehisa. Sato. Y, Furumichi. M, Morishima. K and Tanabe. M (2019) 'New approach for understanding genome variations in KEGG' *Nucleic Acids Research*, 47(**1**), pp. 590-595.

Lachmann. A, Xu. H, Krishnan. J, Berger. S, Mazloom. A and Ma'ayan. A (2010) 'ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments', *Bioinformatics*, 26(**19**), pp. 2438-2444.

Lal. S, Allan. A, Markovic. D et al (2013) 'Estrogen Alters the Splicing of Type 1 Corticotropin-Releasing Hormone Receptor in Breast Cancer Cells', *Science Signaling*, 6(**282**), pp. 53.

Liu. L and Ruan. J (2014) 'Network-based Pathway Enrichment Analysis', *Biomedical Engineering*, 2013(**1**), pp. 218-221.

Louis. D, Perry. A, Reifenberger. G et al (2016) 'The 2016 World Health Organization Classification of Tumours of the Central Nervous System: a summary', *Acta Neuropathologica*, 131(**6**), pp. 803-820.

Lowe. R, Shirley. N, Bleackley. M, Dolan. S and Shafee. T (2017) 'Transcriptomics technologies', *PLOS Computational Biology*, 1(**1**), pp. 1.

Magill. S, Young. J, Chae. R, Aghi. M, Theodosopoulos. P and McDermott. M (2018) 'Relationship between tumor location, size, and WHO grade in meningioma', *Neurosurgical Focus*, 44(**4**), pp. 1-6.

Menghi. F, Orzan. F, Eoli. M et al (2011) 'DNA Microarray Analysis Identifies *CKS2* and *LEPR* as Potential Markers of Meningioma Recurrence', *Oncologist*, 16(**10**), pp. 1440-1450.

Metsalu. T and Vilo. J (2015) 'ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap', *Nucleic Acids Research*, 43(**1**), pp. 566-570.

miRBase (2005) *miRBase: the microRNA database*, Available at: http://www.mirbase.org/ (Accessed: 19 February 2019).

Morpheus (2019) *MORPHEUS*, Available at: https://software.broadinstitute.org/morpheus (Accessed: 6 March 2019).

Pacholewska. A (2017) '"Loget"- a Uniform Differential Expression Unit to Replace "logFC" and "log2FC"', *Matters*, 1(**1**), pp. 1-3.

Patel. M and Wang. J (2018) 'The Identification and Interpretation of *cis*-Regulatory Noncoding Mutations in Cancer', *High-Throughput*, 8(**1**), pp. 1-22.

Pećina-Šlaus. N, Kafka. A and Lechpammer. M (2016) 'Molecular Genetics of Intracranial Meningiomas with Emphasis on Canonical Wnt Signalling', *Cancers*, 8(**7**), pp. 67.

Pico. C, Sánchez. J, Oliver. P and Palou. A (2002) 'Leptin production by the stomach is up-regulated in obsess (fa/fa) Zucker rats', *Obesity research*, 10(**9**), pp. 932-938.

Ritchie. M, Phipson. B, Wu. D, Hu. Y, Law. C, Shi. W and Smyth. G (2015) '*limma* powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*, 43(**7**), pp. 47.

Rogers. A and Weiss. S (2009) 'Chapter 20- Epidemiologic and Population Genetic Studies', In: Robertson. D and Williams. G, *Clinical Translational Science*, Cambridge: Academic Press, pp. 289-299.

Saraf. S, McCarthy. B, and Villano. J (2011) 'Update on Meningiomas', *The Oncologist*, 16(**11**), pp. 1604-1613.

Schmidt. M, Mock. A, Jungk. C et al (2016) 'Transcriptiomic analysis of aggressive meningioma identifies PTTG1 and LEPR as prognostic biomarkers independent of WHO grade', *Oncotarget*, 7(**12**), pp. 14551-14568.

Silipo. M, Gautrey. H and Tyson-Capper. A (2015) 'Deregulation of splicing factors and breast cancer development', *Journal of Molecular Cell Biology*, 7(**5**), pp. 388-401.

Smyth. G (2004) 'Linear models and empirical Bayes methods for assessing differential expression in microarray experiments', *Statistical Applications in Genetics and Molecular Biology*, 3(**3**), pp. 1.

Smyth. G (2005) 'limma: Linear Models for Microarray Data', In: Gentleman. R, Carey. J, Huber. W, Irizarry. A., Dudoit. S, *Statistics for Biology and Health*, New York: Springer

Stevens. J, Masud. A, and Suyundikov. A (2017) 'A comparison of multiple testing adjustment methods with block-correlation positively-dependent tests', *PLoS One*, 12(**4**), pp. 1-12.

Subramanian. A, Tamayo. P, Mootha. V et al (2005) 'Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles', *Proceeding of the National Academy of Sciences of the United States of America*, 102(**43**), pp. 15545-15550.

Suppiah. S, Nassiri. F, Bi. W et al (2019) 'Molecular and translational advanced in meningiomas', *Neuro-Oncology*, 21(**1**), pp. 4-17

Tipney. H and Hunter. L (2010) 'An introduction to effective use of enrichment analysis software', *Human Genomics*, 4(**3**), pp. 202-206.

Vasudevan. H, Braustein. S, Phillips. J et al (2018) 'Comprehensive Molecular Profiling Identified FOXM1as a Key Transcription Factor for Meningioma Proliferation', *Cell Reports*, 22(**13**), pp. 3672-3683.

Wang. J, Vasaikar. S, Shi. Z, Greer. M and Zhang. B (2017) 'WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit', *Nucleic Acids Research*, 45(**1**), pp. 130-137

Wang. N and Osswalk. M (2018) 'Meningiomas: Overview and New Directions in Therapy', *Seminars Neurology*, 38(**1**), pp. 112-120.

Wiemels. J, Wrensch. M and Claus. E (2010) 'Epidemiology and etiology of meningioma', *Journal of Neuro-Oncology*, 99(**3**), pp. 307-314.

Wu. L, Dong. B, Zhang. F, Li. Y and Liu. L (2016) 'Prediction of the engendering mechanism and specific genes of primary melanoma by bioinformatics analysis', *Dermatologica Sinica*, 34(**1**), pp. 14-19

Xia. J, Gill. E and Hancock. R (2015) 'NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data', *Nature Protocols*, 10(**2015**), pp. 823-844.

Xie. X, Lu. J, Kulbokas. E, Golub. T, Mootha. V, Lindblad-Toh. K, Lander. E and Kellis. M (2005) 'Systematic discovery of regulatory motifs in human promoters and 3'UTRs by comparison of several mammals', *Nature*, 434(**1**), pp. 338-345.

Yoon. C, Kim. M, Lee. H et al (2012) '*PTTG1* Oncogene Promotes Tumor Malignancy via Epithelial to Mesenchymal Transition and Expansion of Cancer Stem Cell Population', *Journal of Biological Chemistry*, 287(**23**), pp. 19516-19527.

Zhao. S, Guo. Y, Sheng. Q and Shyr. Y (2014) 'Advanced Heat Map and Clustering Analysis Using Heatmap3', *BioMed Research International*, 2014(**986048**), pp. 1-6.

Appendices are available as 'supplementary files' (please see download area).