ONLINE ONLY

Supplemental material

Gene expression analysis during progression of malignant meningioma compared to benign meningioma

Maier et al.

https://thejns.org/doi/abs/10.3171/2022.7.JNS22585

DISCLAIMER The *Journal of Neurosurgery* acknowledges that the following section is published verbatim as submitted by the authors and did not go through either the *Journal's* peer-review or editing process.

Supplementary Material

S1 – Customization of the NanoString Neuroinflammation panel

Table S1.

ARNT2	*KLF4*	*SFRP1*
CDKN2A	*MAPK4*	*SFRP4*
CTNNB1	*MELK*	*SMARCB1*
ELN	*MYBL2*	*SMARCE1*
EPHB3	*NF2*	*SMO*
FOXM1	*NOTCH1*	*SUFU*
FZD8	*PBRM1*	*TERT*
H3-3A	*POLR2B*	*TRAF7*
HOXB2	*RBP4*	*TTK*
IGF2	*RUNX1*	*VIT*

List of 30 extra genes added to the Nanostring nCounter Human Neuroinflammation Panel.

Nanostring nCounter Human Neuroinflammation Panel gene list and annotations are available here: www.nanostring.com/wp-content/uploads/2021/01/LBL-10496_Human_Neuroinflammation_Gene_List.xlsx

S2 – Differentially expressed genes calculations (formulas)

Probabilistic Index Model (Analysis #1, #2 and #3)

In the comparison of two groups of individuals (between subjects), we identify the significant differences in gene expression employing the probabilistic index regression model[1]. It estimates the probability that the gene expression in the tumor of a randomly selected patient in one group is higher than the gene expression in the tumor of a randomly selected patient in the other group. We account for sex and age at tumor resection. Hence, the probabilistic index is a generalization of the Mann-Whitney-Wilcoxon rank test which allows us to control for patient demographics[2].

Let Y_{gi} be the normalized expression of gene $g = 1,2,\ldots,G$ for patient $i = 1,2,..,m$ who belongs to group $A_i \in \{0,1\}$ and with vector of covariates X_i (e.g. sex and age at first WHO grade III tumor), then we define the probabilistic index model:

$$\text{logit}(\mathbb{P}(Y_{gi} \leq Y_{gj}|A_i, A_j, X_i, X_j) = \beta_{g1}(A_j - A_i) + \beta_{g2}(X_j - X_i)$$

and the probabilistic index for gene g is estimated as:

$$\text{PI}_g = \hat{\mathbb{P}}(Y_{gi} \leq Y_{gj}|A_i, A_j) = \frac{\exp^{\hat{\beta}_{g1}(A_j - A_i)}}{1 + \exp^{\hat{\beta}_{g1}(A_j - A_i)}}$$

When the gene g is not differentially expressed the probabilistic index PI_g has value around 50%; meanwhile, if there is strong evidence for a difference in gene expression the PI_g is either close to 100% or to 0% depending on in which group the gene expression is higher. Moreover, a Wald-type test for β_{g1} is performed to test if the gene is differentially expressed. The null hypothesis is $H_0: \beta_{g1} = 0$ which is equivalent into test that $\text{PI}_g = 0.5$.

Signed Rank Test (Analysis #4)

When comparing two recurrences within the same individual, the significant differences in gene expressions are determined by the Wilcoxon signed rank test[2]. This test accounts for the fact that two measurements are coming from the same individual. It considers sign and magnitude of the difference between gene expressions in two measurements. Let $Y_{k,gi}$ be the normalized gene expression of $g = 1,2,\ldots,G$ for recurrence $k \in \{0,1\}$ in individual $i = 1,2,\ldots,m$. The test calculates the difference $d_{gi} = (Y_{1,gi} - Y_{0,gi})$ for each gene g in individual i, than it orders the absolute values of the differences and assigns rank $R_{gi} = 1$ to the largest and $R_{gi} = m$ to the smallest absolute difference. Finally, the test statistic is:

$$W_g = \sum_{i=1}^{m} \text{sign}\,(Y_{1,gi} - Y_{0,gi}) \cdot R_{gi}$$

where $\text{sign}\,(d_{gi}) = 1$ if the difference $d_{gi} > 0$ and $\text{sign}\,(d_{gi}) = -1$ when the gene expression in recurrence 0 is larger than the gene expression in recurrence 1 $d_{gi} < 0$.

All analyses

In all four analyses, the same test is employed for 787 genes. We account for multiple testing by adjusting the p-values with the Benjamini and Hochberg correction method[3]. Thus, we control the false discovery rate (FDR) at most 5%, which is the proportion of genes that are falsely declared differentially expressed.

S3 – Sub analyses based on pathway annotations in analysis #1 – Microglia and cytokines

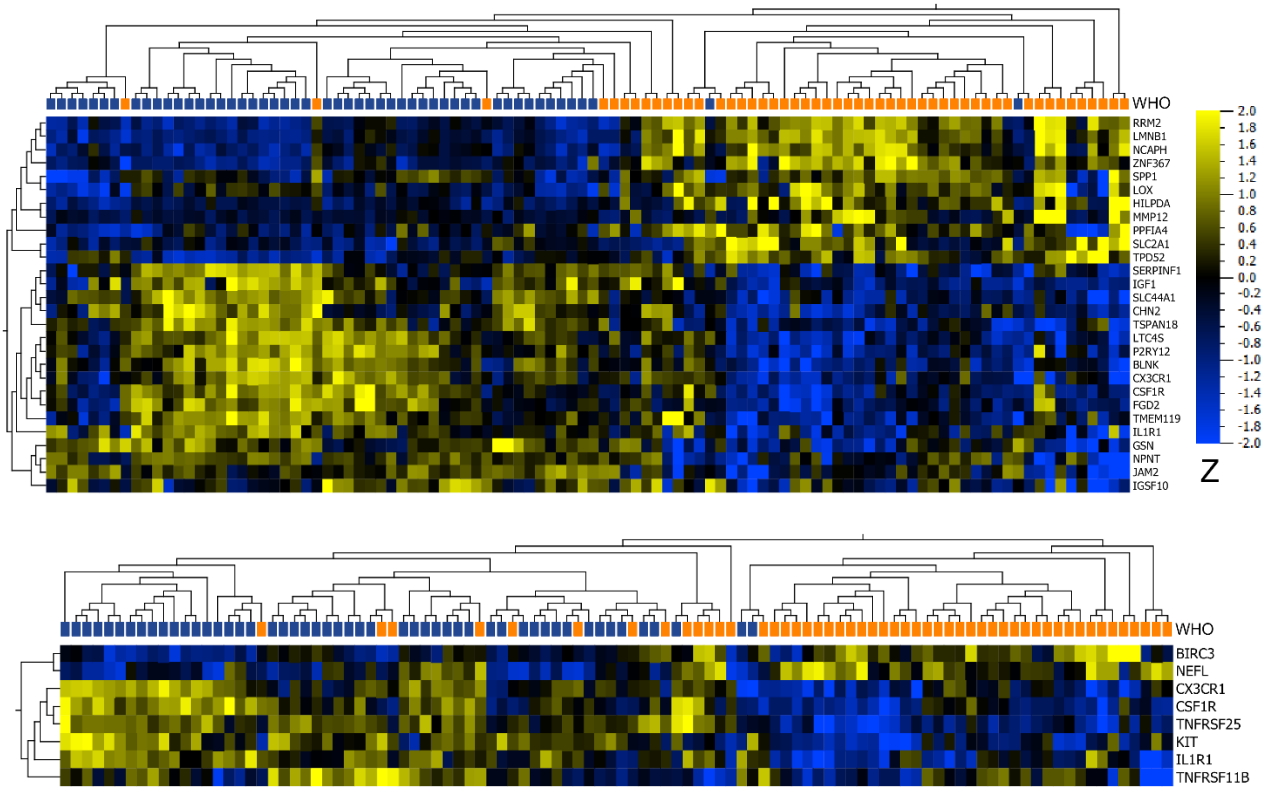


Fig. S3. Heatmaps and UHCL based on t-tests in analysis #1 (comparing 51 grade III meningioma to 51 grade I meningiomas from benign controls). Differentially expressed genes (DEG) were filtered based on neuroinflammatory pathways as annotated by NanoString. Upper panel shows DEG associated with microglia regulation. Of the 187 genes annotated to this pathway, 28 had significantly differential gene expression in grade III compared to grade I from benign controls. Lower panel shows cytokine associated genes in WHO grade III vs. grade I controls, 8 DEG (t-tests) out of 117 annotated. Grade III meningiomas are marked orange and grade I are marked blue in both subanalyses.

S4 – IPA analysis – Canonical pathways, networks and analysis match

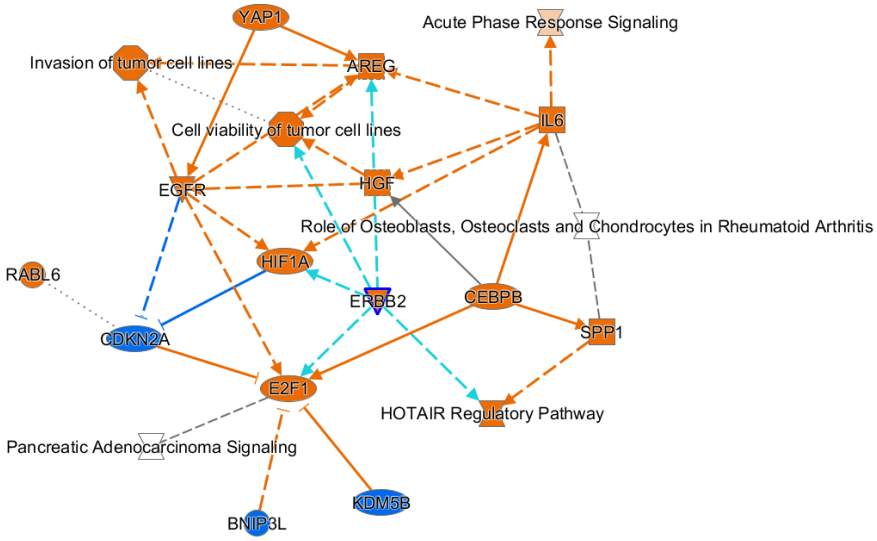
The IPA software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis)[4] was used for bioinformatics analysis, in which we investigated canonical pathways, networks and analysis match (how patterns in our dataset relate to other data sets). For network generation, IPA uses a network algorithm to map multiple molecules into network and assign scores for each network (the score is -log to the p-value of Fisher's exact test at the right tail, p score = (-log10(p-value)). For analysis match, the pattern matching is based on a z-score which indicates how well the activated or inhibited entities match in another signature. The score is normalized to range from 100% to -100% (overall similarity score) with corresponds to perfect match and a perfect 'anti-match'; threshold for significant overlap was set to -50% and 50%.

The 78 DEG (analysis #1, t-tests) were uploaded to the IPA (fig. S4 shows a graphical summary). Table S4 shows all canonical pathways with -log(p) > 3. The network with the highest p-score (-log10(p-value) = 31) is shown in figure S4. The IPA analysis match analysis yielded no datasets with an overall similarity >50% or -<50%. The datasets with the highest similarity overall z-score were a bladder carcinoma data set from OncoGeo (testing p53-like vs. basal, overall similarity z-score of 44.3%) and a TCGA low grade glioma dataset (testing *CCNE1* somatic mutation vs. wildtype, overall similarity z-score of 43.4%).

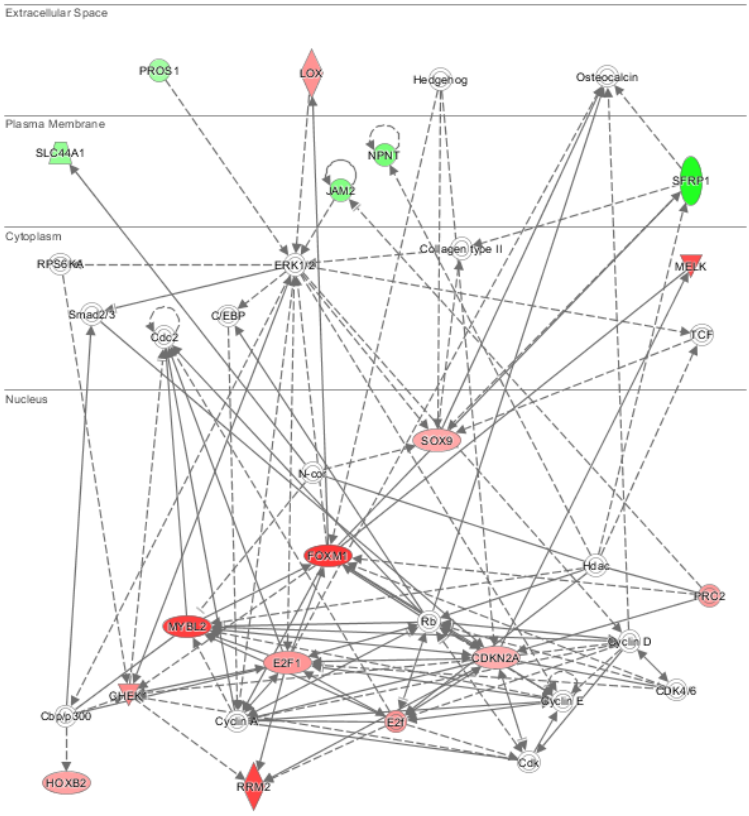
Table S4.

Ingenuity Canonical Pathways	-log(p-value)	zScore	Molecules
HIF1α Signaling	3,2	1,34	*IGF1,IGF2,MMP12,SERPINE1,SLC2A1*
Hepatic Fibrosis Signaling Pathway	3,33	0,38	*EZH2,IL1R1,MAPK10,PRKAR2B,SERPINE1,SPP1,TNFRSF11B*
Senescence Pathway	3,34	0,82	*CDKN2A,CHEK1,E2F1,EZH2,MAPK4,SERPINE1*
Kinetochore Metaphase Signaling Pathway	3,39	1,00	*BIRC5,KIF2C,PTTG1,TTK*
Tumor Microenvironment Pathway	3,49	1,34	*IGF1,IGF2,MMP12,SLC2A1,SPP1*
p53 Signaling	3,5	-1,00	*BIRC5,CDKN2A,CHEK1,E2F1*
HOTAIR Regulatory Pathway	3,68	2,24	*EZH2,FOXM1,JAM2,MMP12,SPP1*
Autophagy	4,12	-0,82	*E2F1,IGF1,MAPK10,PRKAR2B,SESN1,TNFRSF11B*
LXR/RXR Activation	4,26	-0,45	*C4A/C4B,IL1R1,RBP4,SERPINF1,TNFRSF11B*
Neuroinflammation Signaling Pathway	5,02	-0,38	*BIRC3,BIRC5,CSF1R,CX3CR1,IL1R1,MAPK10,MAPK4,SLC6A1*
Acute Phase Response Signaling	6,74	0,45	*C4A/C4B,CP,IL1R1,RBP4,SERPINA3,SERPINE1,SERPINF1,TNFRSF11B*

Table S4. Canonical pathways from the Ingenuity Pathway Analysis with corresponding p-values and z-scores. The only canonical pathway with a significant p-value and concomitant z-score >2 is the HOTAIR regulatory pathway.



(a)



(b)

Fig. S4. (a) Graphical summary from the IPA analysis based on the 78 DEG shows entities with p-values <0.05 and z-scores > 2 (orange, activated nodes) and <-2 (blue, inhibited nodes) for diseases, functions and upstream regulators in WHO grade III vs. grade I meningiomas. **(b)** Top network based on the 78 DEG from analysis #1 with a p-score of 31 (p-score = -log10(p-value)). Upregulated genes in this network (grade III vs. I comparison) are marked with red and downregulated genes in the network are marked with green.

S5 – Probabilistic Index Model results for analysis #3 and #4

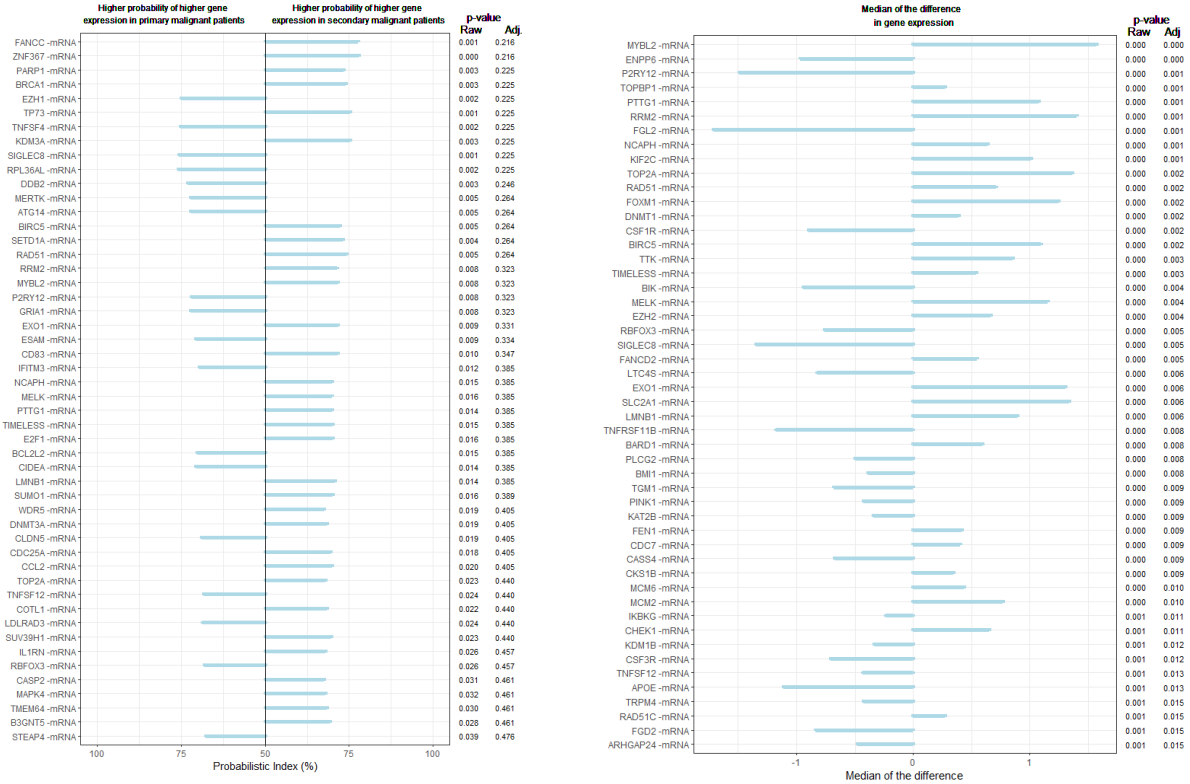


Fig. S5. Probabilistic indices, p-values, and adjusted p-values for analysis #3 (left) and #4 (right). (**a)** In analysis #3 the first WHO grade III meningioma in patients with secondary malignant meningioma (n=24) was compared to meningiomas from patients with primary malignant meningioma (n=27). The genes are ranked by adjusted p-value and top 50 are shown; analysis #3 yielded no significant DEG (adj. p-value <0.05) (**b)** In analysis #4 we made a comparison within patients with secondary malignant meningioma. The last premalignant meningioma was compared to the first grade III meningioma and the analysis yielded 119 DEG. The genes are sorted by the adjusted p-value and top 50 are shown.

S6 – Heatmap of gene expression comparison between WHO grade III meningiomas from patients with primary and secondary malignant meningioma (analysis #3)

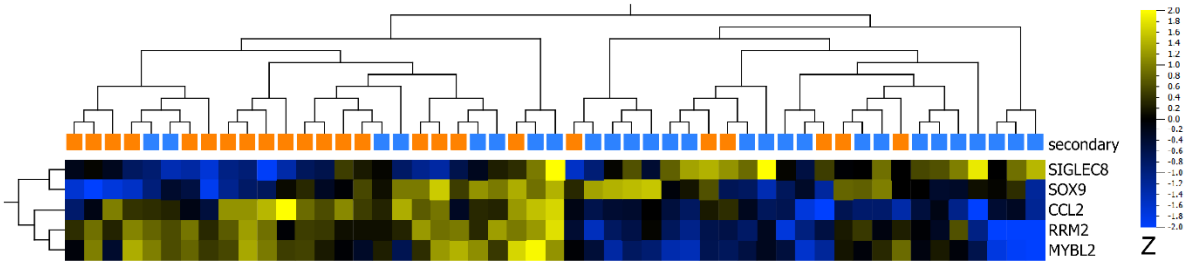


Fig. S6. Heatmap and unsupervised clustering based on t-tests comparing gene expression between 24 secondary WHO grade III meningioma, marked orange, and 27 primary grade III meningioma, marked blue (analysis #3). Upregulation in secondary vs. primary is colored yellow and downregulation blue. Eight differentially expressed genes where found with q = 0.51.

S7 – Heatmap of gene expression comparison between WHO grade III meningiomas with rhabdoid/papillary and anaplastic morphology

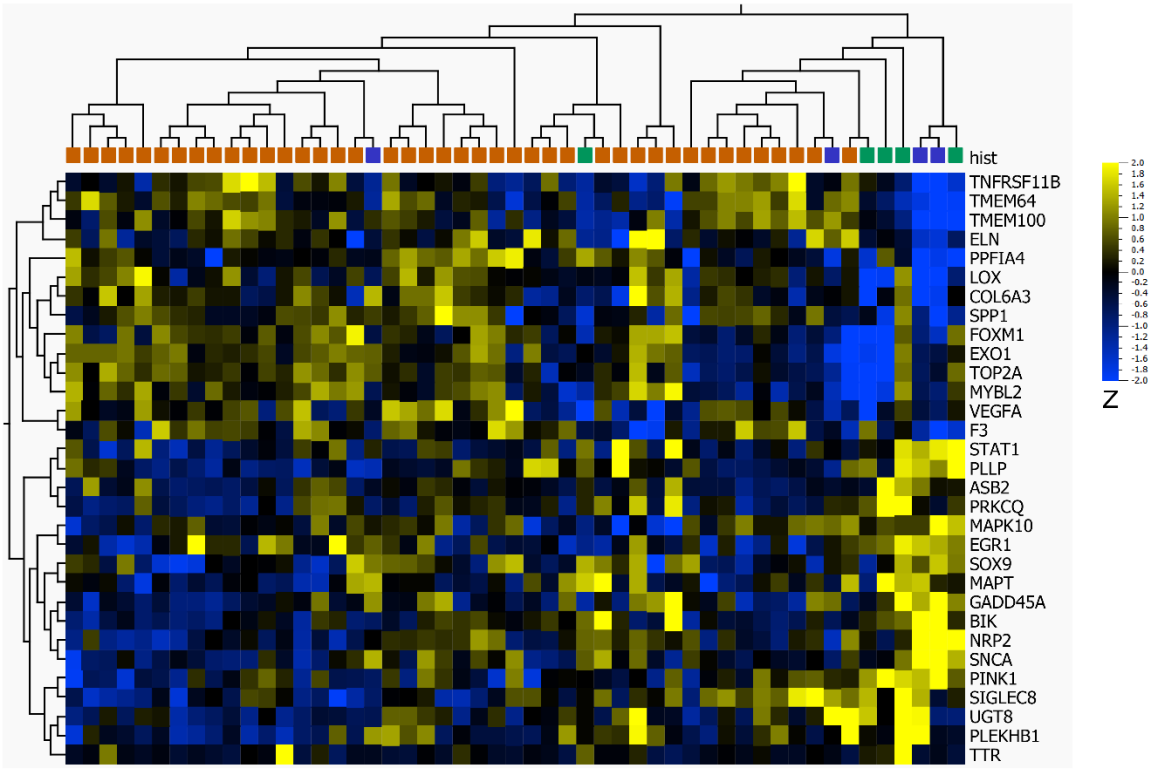


Fig. S7. Heatmap and unsupervised clustering based on t-tests comparing gene expression between 9 WHO grade III meningioma with rhabdoid or papillary morphology (4 papillary, marked blue, and 5 rhabdoid, marked green), and 42 anaplastic meningiomas. Upregulation in anaplastic vs. non-anaplastic is colored yellow and downregulation blue. 31 differentially expressed genes where found with q = 0.31. The cluster seen on the right with rhabdoid/papillary meningiomas contained some meningiomas with rhabdoid/papillary morphology and concomitant high mitotic index (>20) or anaplasia, thus no clear clustering based on concomitant malignant features.

S8 – Gene Expression Trajectories across recurrences

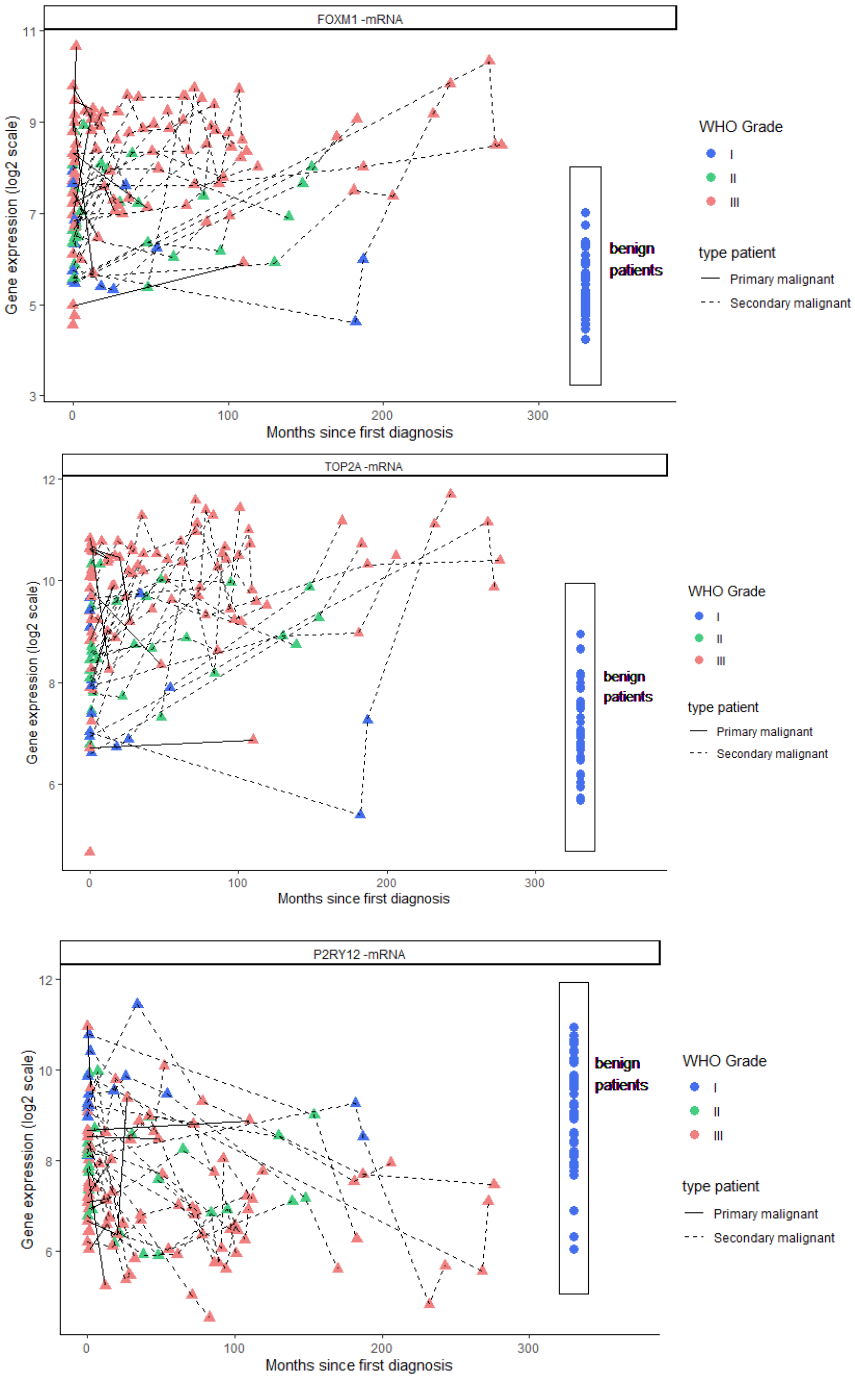


Fig. S8. Gene expression trajectories across recurrences for *FOXM1, TOP2A and P2RY12.* The genes were chosen as illustrative examples of two trajectories with increasing expression over time *(FOXM1* and *TOP2A*) and one trajectory diminishing over time (*P2RY12*). Log2 values of the 3 selected differentially expressed genes were plotted with months from first diagnosis and surgery date of the respective meningioma in malignant cases on the x-axis. The log2 gene expressions for the benign controls (a single WHO grade I meningioma per patient) are provided on the right. For patients with primary malignant meningioma the trajectory is showed with full lines and for patients with secondary malignant meningioma, it is shown with dotted lines. Color indicates WHO grade.

Supplementary References

1. Thas O, Neve J De, Clement L, Ottoy JP. Probabilistic index models. *J R Stat Soc Ser B Stat Methodol*. 2012;74(4):623-671. doi:10.1111/j.1467-9868.2011.01020.x

2. Miller E. THE SIGNED-RANK (WILCOXON) TEST. *Lancet*. 1969;293(7590):371. doi:10.1016/S0140-6736(69)91330-0

3. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

4. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*. 2014;30(4):523-530. doi:10.1093/bioinformatics/btt703