Submitted to the Annals of Applied Statistics

SUPPLEMENTARY MATERIALS TO: RADIO-IBAG: RADIOMICS-BASED INTEGRATIVE BAYESIAN ANALYSIS OF MULTIPLATFORM GENOMIC DATA

By Youyi Zhang [1], Jeffrey S. Morris[1], Shivali Narang Aerry[2], Arvind U.K. Rao[3] and Veerabhadran Baladandayuthapani [*3]

[1] *The University of Texas MD Anderson Cancer Center*
[2] *Johns Hopkins University*
[3] *University of Michigan Ann Arbor*

S1. Bayesian Hierarchical Modeling We applied the normal gamma (NG) prior for our hierarchical linear regression model in both stage II and stage III with different hierarchical levels. The subsequent sections include the general full conditional posterior distribution of the parameters, initial values for MCMC and hyperparameter settings.

S1.1 Full conditional posterior distribution In both the *radiogenomic model* and the *radiogenomic clinical model*, we applied a linear regression model with the Bayesian normal gamma shrinkage method to achieve sparse signaling detection.

Consider the linear regression model: $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

In the *radiogenomic model*, $\boldsymbol{Y}$ denotes the specific RmF or RF, and X is the matrix of the genomic platform combinations. In the *radiogenomic clinical model*, $\boldsymbol{Y}$ denotes the clinical outcome, and X represents the matrix of the RmF or RF combinations modulated by different gene expression parts explained by different genomic platforms. For our general regression model, the full posterior distribution can be expressed as

$$\boldsymbol{\beta}|rest \sim Normal((X^TX + \sigma^2 D_\tau^{-1})^{-1}X^TY, (X^TX + \sigma^2 D_\tau^{-1})^{-1}\sigma^2)$$

$$\sigma^2|result \sim IG(a + n/2, b + (Y - X\boldsymbol{\beta})^T(Y - X\boldsymbol{\beta})/2)$$

$$\psi_{ji}|rest \sim GIG(a = \gamma_j^{-2}, b = \beta_{ji}^2, p = \lambda_j - \frac{1}{2})$$

$$\lambda_j|rest \sim (1/\lambda_j)^{\tilde{a}} exp\{-\tilde{b}\gamma_j^{-2}/(2\lambda_j) - c\lambda_j\} \times \prod_{i=1}^{p_j} \psi_{ji}^{\lambda_j} / \{(\Gamma(\lambda_j))^{p_j}(2\gamma_j^2)^{p_j\lambda_j}\}$$

*Correspondence, E-mail: veerab@umich.edu

$$\gamma_j^{-2}|rest \sim Gamma(\tilde{a} + p_j\lambda_j, (\tilde{b}/\lambda_j + \sum_{i=1}^{p_j}\psi_{ji})/2),$$

If applying to the *radiogenomic clinical model*, j denotes the RmF combination that are modulated by the gene expression that is explained by the j^{th} platform, and k represents the k^{th} RmF; if applying to the *radiogenomic model*, j is the genomic platform type index, i is the gene index.

Specifically, λ_j is sampled through the Metropolis-Hastings method, the proposed family is $exp(\sigma_\lambda^2 z)\lambda_j$, and z comes from the standard normal distribution. The acceptance rate is controlled between 20% and 30%.

S1.2 Initial values and hyperparameter settings For the general model, the initial values and hyperparameter settings are as follows.

- For the prior distribution of σ^2, we choose to apply the uninformative prior and set $a = b = 0.001$.
- The hyperparameter for λ_j is at c; we set $c = 1$ [Griffin et al. (2010)].
- For the prior distribution of γ_j^{-2}, $\tilde{a} = 2$ and if the design matrix X is non-singular, we set $\tilde{b}$ to be the mean of the least squares solution $\hat{\beta}_{ji}^2$. If X is singular, $\tilde{b}$ is set to be the average of the minimum-length least squares (MLLS) solution $\hat{\beta}_{ji}^2$ across all subjects.
- The initial $\boldsymbol{\beta}$ is estimated from the frequentist elastic net algorithm.
- The initial value of σ^2 is estimated by the mean squared error from the frequentist elastic net algorithm.
- The initial values for γ_j^{-2}, λ_j and ψ_{ji} are set to 1.

S2 Data preprocessing for GBM In this section, we mainly present data preprocessing techniques and procedures, especially for GBM MRI imaging data, including imaging preprocessing, feature extraction and description. Further, we illustrate our preliminary checking results for radiomic features (RFs) and describe how radiomic-meta-features (RmFs) are generated. For the genomic platform data sets, we perform missing value imputation; detailed information is described in the last subsection.

S2.1 Radiomic-feature preprocessing Our GBM MRI imaging data set was downloaded from The Cancer Imaging Archive with two imaging modalities: T1-post contrast and T2-weighted FLAIR.

In brief, the dicom MRI images were converted to the nifti format (.nii) using MRIConvert software ($http://lcni.uoregon.edu/\sim jolinda/MRIConvert/$). We then performed pre-processing steps on the MRI images by following a certain pipeline, and further obtained the 3D tumor volumes. The pipeline is described here.

a. Non-uniformity correction: We used a nonparametric intensity non-uniformity normalization (N3) correction module in MIPAV (v 6.0) [McAuliffe et al. (2001)] to correct the shading artifacts that resulted from partial volume averaging errors of the MRI instrument. The N3 algorithm iteratively estimates the true tissue intensity distribution, as the shading artifacts lead to reduced signal intensities in certain image regions.

b. Registration: We used medical image processing, analysis and visualization (MIPAV) software to register the T2-FLAIR N3 corrected images to the respective T1-POST N3 corrected images. We used the normalized mutual information criterion in the optimized automatic registration module in MIPAV ((http://mipav.cit.nih.gov/).

c. Segmentation: Semi-automated segmentation of the tumor regions in 3D was performed by our clinical experts using the Medical Image Interaction Toolkit 3M3 (http://www.mint-medical.de/). This software features slice-by-slice contour drawing, correction tools and 3D interpolation of the tumor region, which are utilized to perform the tumor segmentation.

d. Re-slicing: Lastly, we re-sliced the original as well as the pre-processed images to widths of 1 millimeter using the NIFIT toolbox in MATLAB ($http://research.baycrest.org/\sim jimmy/NIfTI$). Image Feature Computation: Textural and regional image features. Computer-based texture analysis of medical images such as MRI scans depicts some quantitative properties like the measurement of regional intensity variations. Using these properties in the textural analysis can further help in the prediction of several factors such as molecular subtypes of the disease and patient survival times.

In this study, the image-derived textural features are utilized in analyzing the relationship between imaging data, genomic multiplatform data sets, and clinical outcomes. For this analysis, we derived textural features from the largest axial 2D slice of the tumor area [Zhou et al. (2014)]. These textural features were obtained from the following two-step process.

1. Image filters: We used Laplacian of Gaussian (LoG) [Haralick (1979) and Gaussian filters [Haralick (1992)] to filter the MR images at five different scales, so as to obtain fine, medium and coarse transforms of the 2D tumor region. The LoG filter, commonly used for edge detection, is a measure of the second spatial derivative of an image. The Laplacian filter is applied to an image that has been smoothed using a Gaussian filter; hence, reducing the sensitivity to noise. We used five standard deviations (σ) to derive fine features at 0.2mm and 0.4mm, medium features at 1.5mm and 2.5mm, and coarse features at 5mm.

2. Texture features and summary measures: In terms of textural features, we calculated Haralick features for both T2-weighted FLAIR and T1-weighted post contrast images. The gray-level-co-occurrence matrices (GLCMs) were derived for both the original and pre-processed images. Further, from the GLCMs, we computed Haralick statistical features at 4 different distances (1mm, 2mm, 4mm and 8mm). Besides the textural features, we obtained some summary features such as the mean intensity of the images, entropy and uniformity measures. These features are also known as TxR features [Ganeshan et al. (2010)]. In addition to the textural features, we computed the area and mean intensity of the largest tumor slice. These are referred to as the regional features.

After the extraction of the features, we conducted further normalization by calculating two different ratios: type-1 and type-2. Ratio type-1 corresponds to the ratio between the features computed at different filters and the features computed from the original images (unfiltered features). Ratio type-2 corresponds to the ratio between features computed at the coarsest scale and those computed at the finer scale.

In our analysis, we considered three major types of RFs: Haralick features, histogram features and regional features. We have 13 total Haralick features, with the names and corresponding calculation formulas listed in **Supplementary Table S1**. Histogram features are calculated from the histogram of the image intensity distribution, with a total of 3 typical features, including the mean intensity, entropy and uniformity characteristics, with formula shown in **Supplementary Table S2**. As mentioned in the imaging preprocessing section, the two regional features that we utilized were the area and mean intensity of the largest tumor slice.

TABLE S1
Haralick features and formulas

Textural Features	**Formula**
Energy	$f_1 = \sum_i \sum_j \{p(i,j)\}^2$
Contrast	$f_2 = \sum_{n=0}^{N_g-1} n^2 \{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \vert\vert i-j\vert = n\}$
Correlation	$f_3 = \frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Sum of squares: variance	$f_4 = \sum_i \sum_j (i-\mu)^2 p(i,j)$
Inverse difference moment (local homogeneity)	$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
Sum average	$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$
Sum variance	$f_7 = \sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i)$
Sum entropy	$f_8 = -\sum_{i=2}^{2N_g} p_{x+y}(i) log\{p_{x+y}(i)\}$
Entropy	$f_9 = -\sum_i \sum_j p(i,j) log\{p(i,j)\}$
Difference variance	f_{10} =variance of p_{x-y}
Difference entropy	$f_{11} = -\sum_{i=0}^{N_g-1} p_{x-y}(i) log\{p_{x-y}(i)\}$
Cluster shade	$f_{12} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i+j-\mu_x-\mu_y\}^3 p(i,j)$
Cluster prominence	$f_{13} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i+j-\mu_x-\mu_y\}^4 p(i,j)$

* N_g denotes the number of distinct grey levels used; μ_x,μ_y,σ_x and σ_y are the means and standard deviations of the partial probability density p_x and p_y; $p_x(i) = \sum_{j=1}^{N_g} p(i,j)$, $p_y(j) = \sum_{i=1}^{N_g} p(i,j)$; $p_{x+y}(k) = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j), i+j = k$ and $k = 2, 3, ..., 2N_g$; $p_{x-y}(k) = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j), |i-j| = k$ and $k = 0, 1, ..., N_g - 1$.

TABLE S2
Histogram Features and formulas

Histogram Features	**Formula**
Mean	$m = \sum_{i=0}^{L-1} z_i p(z_i)$
Uniformity	$U = \sum_{i=0}^{L-1} p^2(z_i)$
Entropy	$e = -\sum_{i=0}^{L-1} p(z_i) log_2^{p(z_i)}$

* z_i denotes a random variable indicating intensity, $p(z_i)$ is the histogram of the the intensity values within ragion of interest (ROI), and L is the number of possible intensity levels.

S2.2 RF and RmF description In total, we have 972 RFs that we extracted from the original imaging data. Considering that some features were extracted from the same imaging modality, or were processed using the same algorithm, we categorized 972 RFs into 20 groups based on their properties, extracting procedure and imaging modalities. **Table S3** illustrates the name

of the RF groups and the corresponding brief description.

The 972 preprocessed RFs are highly correlated when checking for Pearson correlation. Figure S1 clearly shows that the RFs are highly correlated with the block structure due to two imaging modalities as well as two major normalizing approaches (ratio-1 and ratio-2) applied during feature extraction.

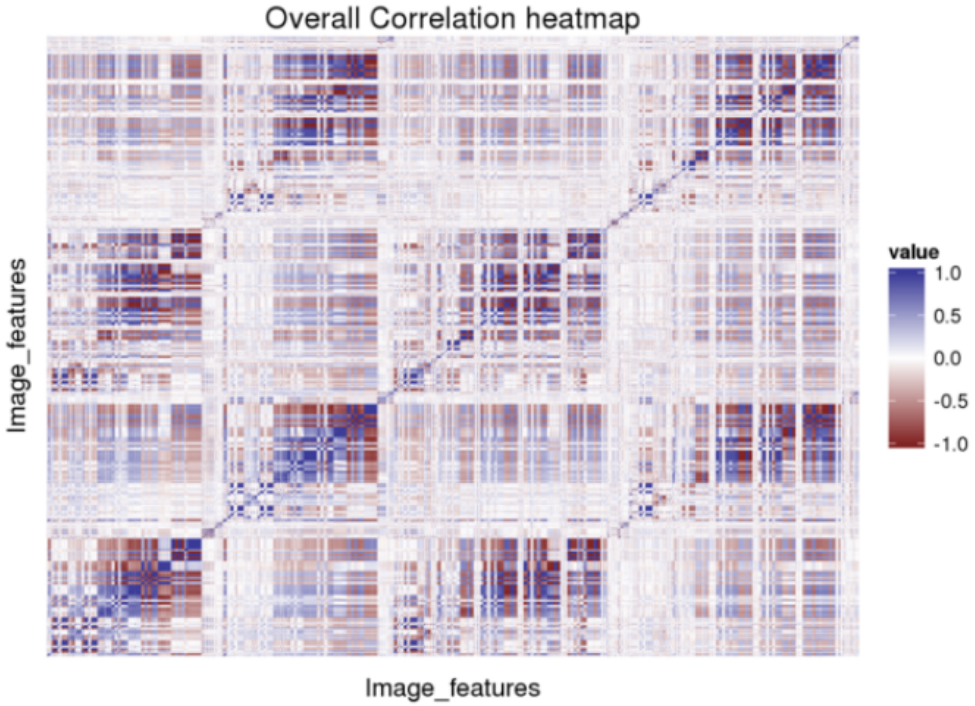


Fig S1: Heatmap of the Pearson correlation among radiomic features (972 features)

In our analysis, we chose to apply dimensional reduction approaches to our RFs. The typical technique is principal component analysis (PCA), however, the key limitation lies in that it does not lead to sparse loadings, making it harder to interpret the results. Hence, alternatively, the sparse PCA algorithm developed by [Zou, Hastie and Tibshirani (2006)] was applied with the formulation described below.

Suppose Z_i is the i^{th} principal component derived from ordinary PCA of

TABLE S3
RF groups and description

RF group name	**Description**
F_Region	Regional features including tumor area, maximum intensity, minimum intensity, mean intensity of T2-weighted FLAIR image
F_LoG_Tex_R1	Ratio1(filter/unfiltered) Haralick texture features derived from LoG filtered T2-weighted FLAIR image
F_Unft_Hist	Histogram features derived from unfiltered T2-weighted FLAIR image
F_LoG_Hist_R1	Ratio1(filter/unfiltered) histogram features derived from LoG filtered T2-weighted FLAIR image
T1_Region	region features including tumor area, maximum intensity, minimum intensity, mean intensity of T1-post contrast image
T1_LoG_Tex_R1	Ratio1(filter/unfiltered) Haralick texture features derived from LoG filtered T1-post contrast image
T1_Unft_Hist	Histogram features of unfiltered T1-post contrast image
T1_LoG_Hist_R1	Ratio1(filter/unfiltered) histogram features of LoG filtered T1-post contrast image
F_LoG_Tex_Fine	Haralick features derived from LoG filtered FLAIR image with fine scale
F_LoG_Tex_R2	Ratio2(coarse/fine) haralick features derived from LoG filtered T2-weighted FLAIR image
F_LoG_Hist_Fine	Histogram features of fine LoG filtered T2-weighted FLAIR image
F_LoG_Hist_R2	Ratio2(coarse/fine) histogram features of LoG filtered T2-weighted FLAIR image
F_Gauss_Hist_Fine	Histogram features of Gaussian filtered FLAIR image with fine scale
F_Gauss_Hist_R2	Ratio2(coarse/fine) histogram features derived from Gaussian filtered T2-weighted FLAIR image
T1_LoG_Tex_Fine	Haralick features derived from LoG filtered T1-post contrast image with fine scale
T1_LoG_Tex_R2	Ratio2(coarse/fine) haralick texture features derived from Gaussian filtered T2-weighted FLAIR image
T1_LoG_Hist_Fine	Histogram features of LoG filtered T1-post contrast image with fine scale
T1_LoG_Hist_R2	Ratio2(coarse/fine) histogram features derived from LoG filtered T1-post contrast image
T1_Gauss_Hist_Fine	Histogram features of Gaussian filtered T1-post contrast image with fine scale
T1_Gauss_Hist_R2	Ratio2(coarse/fine) histogram features derived from Gaussian filtered T1-post contrast image

matrix $\boldsymbol{X}$ with n samples and p predictors, where the loading matrix is denoted as V_i. By regressing PC on X with penalization, sparse loading can be achieved.

$$\hat{\beta} = \arg_{\beta} min||Z_i - \boldsymbol{X}\beta||^2 + \lambda||\beta||^2 + \lambda_1||\beta||_1$$

, where $||\beta||_1 = \sum_{j=1}^{p} |\beta_j|$. The updated sparse loading can be expressed as $\hat{V}_i = \frac{\hat{\beta}}{||\hat{\beta}||}$, and $\boldsymbol{X}\hat{V}_i$ is the i^{th} approximated principal component. For a more detailed theorem and proof, see the appendix for the publication by Zou and Hastie (2006).

The loadings as well as the leading principal components were derived from both ordinary PCA and sparse PCA, based on RF-prespecified groups, the squared loading proportion of the principal analysis is calculated respectively. This information is shown in Figure S2 below and in Figure 2 in the main text. When comparing these two heatmaps, we can explicitly see a great difference in the sparsity level. We utilized 22 leading principal components derived from sparse PCA as our imaging features in modeling stage II and stage III. We call them "radiomic-meta features" (RmFs) in the analysis.

S2.3 Dataset Sample Size Figure S3 shows the diagram of the sample size description of genomic data of different platform, radiomic data sample size and the clinical data sample size as well as their intersection sample sizes.

mRNA: continuous data
CN: log transformed continuous data
microRNA: continuous data
Radiomic data: continuous
Clinical: continuous data (we took the survival in month with log2 transformation as the outcomes)

S2.4 Missing value imputation for genomic platform data We have missing values in the copy number data set, with 4.3% of the data missing. To impute the missing values, we chose to use the following steps. First, we impute each NA with the average values of the other patients (mean imputation). Second, using the complete matrix with mean imputation to calculate the correlation matrix between markers, for each target marker with missing elements, we select 3 markers that are the most highly positively correlated with this marker. Third, we regress the target marker on the 3 selected markers and obtain the predicted values. Lastly, we replace

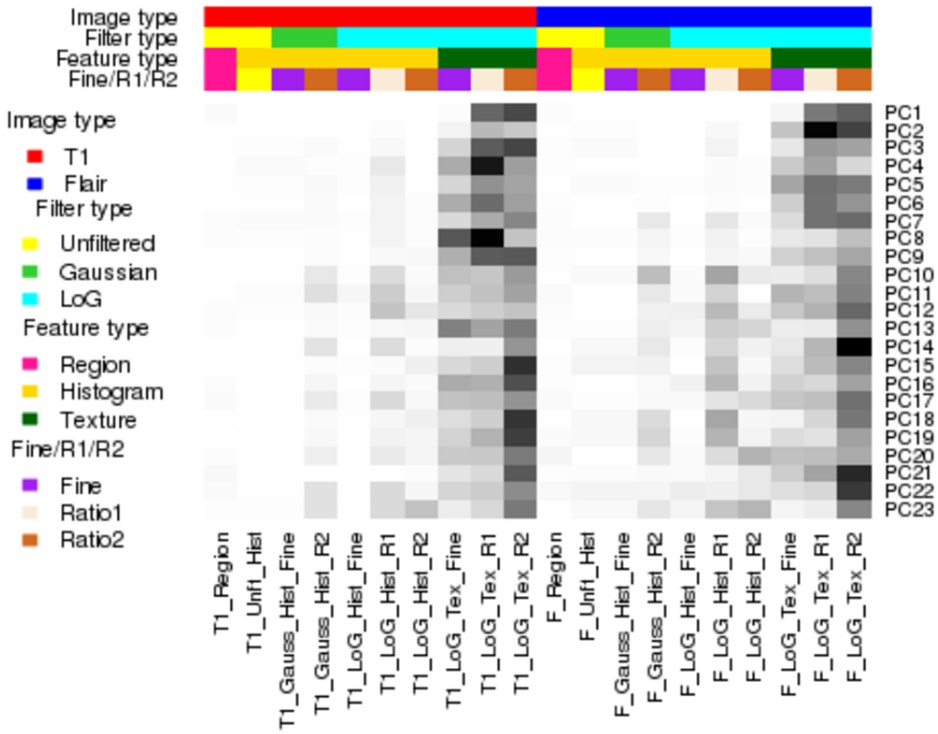


Fig S2: PCA Squared loadings proportion for each RF group. For each of the 23 PC scores, the sum of the squared loadings of each group is calculated after dividing by the total sum of the squared loadings that equals exactly 1. The heatmap shows this values in grey level, interpreted the RF group importance for each PC component. The grey level ranging from white to black matches the proportional values ranging from 0 to 1.

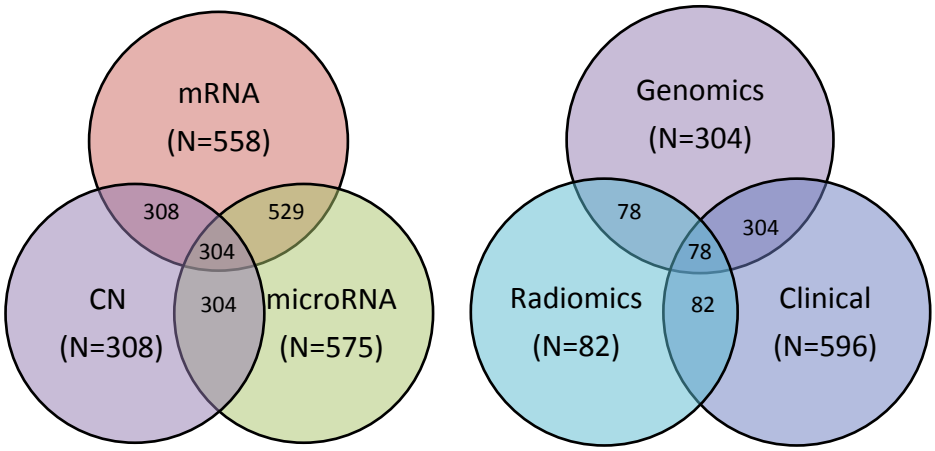


Fig S3: Diagram of sample size, there are 304 samples having all mRNA, CN and microRNA information, and 78 samples have clinical and Radiomic information.

the predicted values for the missing elements of the target marker in the original matrix.

S3 Nonlinearity Checking for Genomic Model We applied generalized additive model (GAM) in *Genomic Model* for each gene given that GAM, compared with General Linear Model (GLM), can achieve higher flexibility in modeling the genomic mechanisms. To check the existence of nonlinearity, we show the comparison of GLM and GAM in terms of ANOVA p-values for model comparison across all 49 genes. Moreover, we also show the fitted smooth curves and the corresponding confidence interval lines for several genes and the platforms in below.

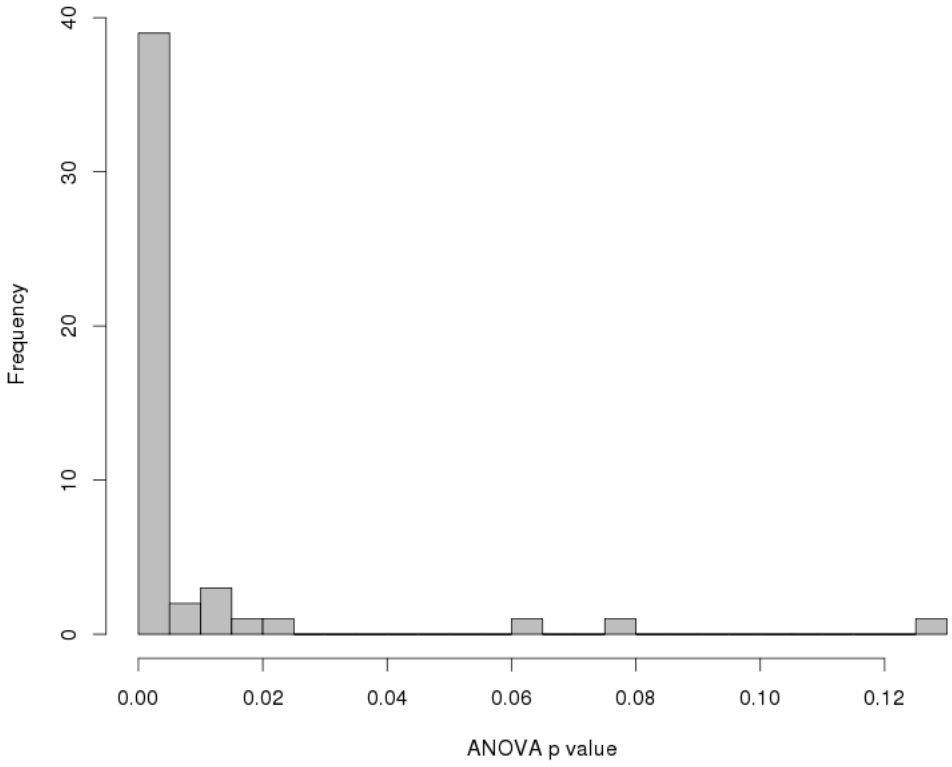


Fig S4: Histogram of ANOVA p-value when doing model comparison (GLM vs. GAM and GLM is nested into GAM), small p-value indicates the two models have significant difference and GAM is preferred over GLM.

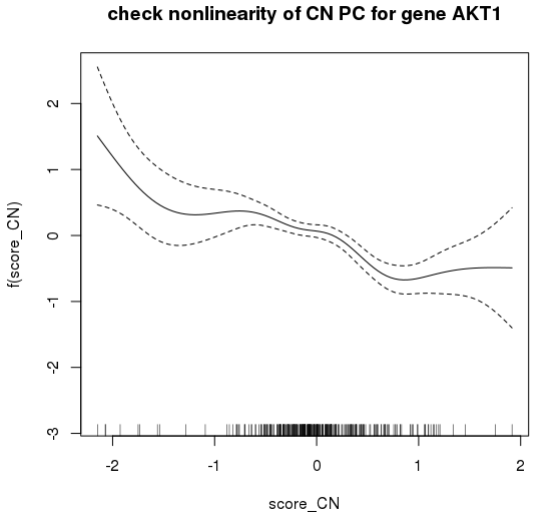


Fig S5: Fitted smoothing curve for the 1st leading PC score of copy number alteration of gene AKT1, the figure shows the existence of nonlinearity

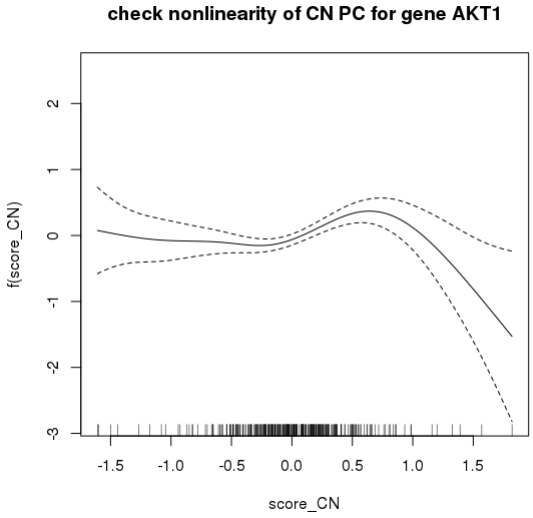


Fig S6: Fitted smoothing curve for the 2st leading PC score of copy number alteration of gene AKT1, the figure shows the existence of nonlinearity

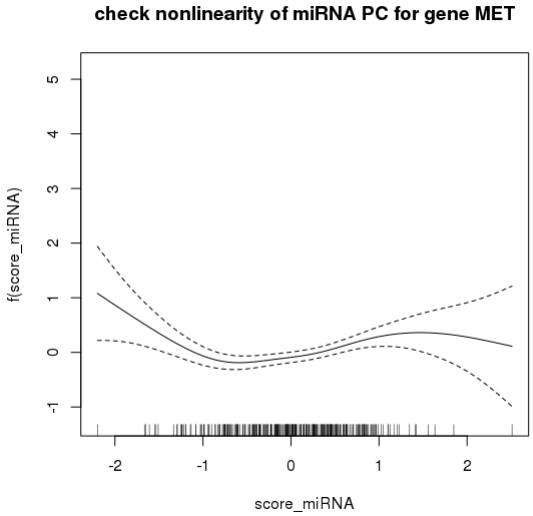


Fig S7: Fitted smoothing curve for the 1st leading PC score of microRNA mapped with gene MET, the figure shows the existence of nonlinearity

S4 Additional Results

S4.1 Magnitude in Stage II In the *radiogenomic model*, we dig into the relationship between the multiplatform genomic data and RmFs; the magnitudes (posterior mean of the coefficients) are shown in Figure S3.

S4.2 Convergence Checking We applied Bayesian Normal Gamma shrinkage model taking account of the multi-scale datasets in both stage II and stage III, thus, we check the converence here for the parameters and hyperparameters in stage II and stage III.

In stage III, we have totally 185 parameters including β_{ji}, ψ_{ji} and $j = 1,2,3,4; i = 1,2,3,...,22$ σ^2, λ_j $(j = 1,2,3,4)$ and $1/\gamma_j^2$ $(j = 1,2,3,4)$. We ran MCMC for 30000 iterations and summarize the results using 20000 samples with the burn-in samples removed. We evaluate the convergence by checking traceplots of 3 main parameter vectors: βs, ψs and σ^2, as well as the ratio of $\lambda_j/(1/\gamma_j^2)$ which identifiably leads to the estimation of ψs. We show the corresponding traceplots for all the parameters in stage III and stage II in separate supplementary file "Traceplot_stage2_stage3.pdf".

In addition, we check the convergence of the parameters in Stage III us-

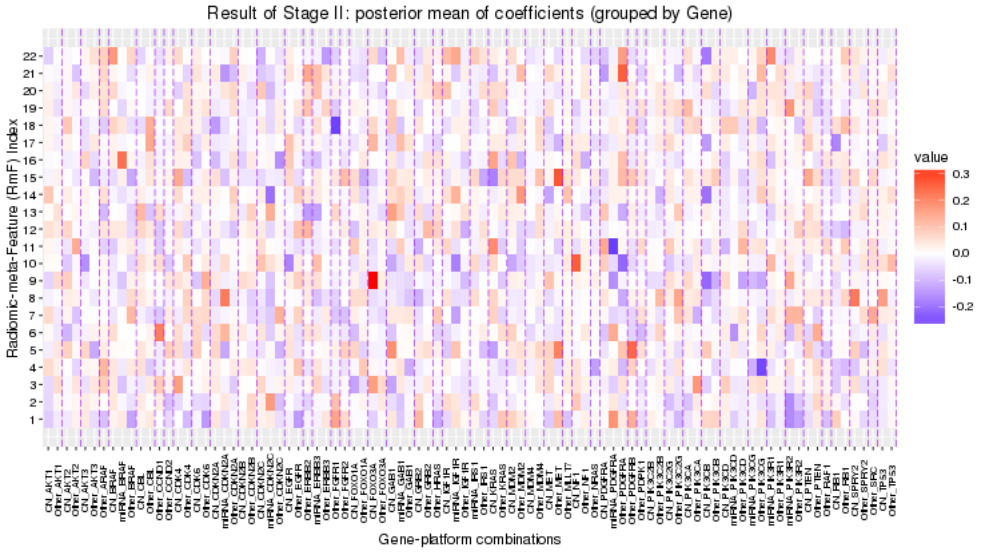


Fig S8: Posterior mean of magnitude from stage II (radiogenomic model). For each RmF, the posterior mean of β_{jg} is the magnitude of the g^{th}'s mRNA part explained by the j^{th} genomic platform. After filtering, 92 gene-platform combinations are sorted and grouped by gene, and the positive and negative effects are respectively illustrated in red and purple.

ing Geweke test Geweke et al. (1991) where we test for the mean difference between first 10% proportion and the last 50% of the samples for all the parameters. P values are illustrated via histogram plots as in the following figures. The histogram is not skewed to the right which indicates proper mixing for MCMC iterations.

In all, our result summary is based on the chains which are long enough to guarantee the convergence.

S4.3 Results of Different Thresholds Checking Here in this section, we check the results when setting different thresholds for both stage II and stage III, respectively δ_1 and δ_2. For δ_1 we check 2 more results when $\delta_1 = 0.05$ and $\delta_1 = 0.1$ and compare the result with our result in the main text with $\delta_1 = 0.075$. Similary, we check the results for stage III when $\delta_2 = 0.02$ and $\delta_2 = 0.08$ and compare the result with $\delta_2 = 0.05$. These are all shown in figures from **Fig S18** to **Fig S30**.

The resutls show that for both stage II and stage III, the number of the

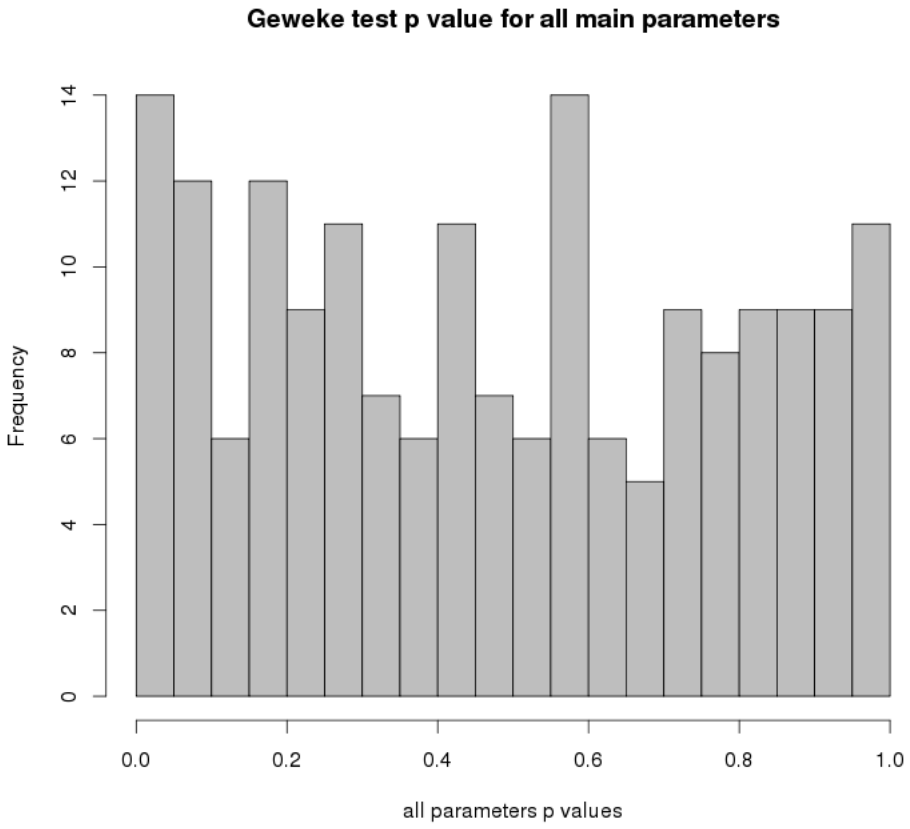


Fig S9: Geweke test for all parameters, 7.7% have p value smaller than 0.05.

signals that are flagged changes with the threshold settings. However, the order of the posterior probabilities are similar and specifically, for stage III, the results do not change much as the threshold decreases or increases within proper range. Thus, we choose the specific threshold based on how it can balance the sparsity and biological interpretation.

S5 Stage III Result Using Bayessian Lasso In our paper, we applied Normal Gamma prior in both stage II and stage III. The reason we chose to apply normal gamma prior here is mainly considering about the flexibility of the shrinkage where Normal Gamma contains two parameters of the prior settings in coefficients. We also applied Bayesian Lasso Park and Casella (2008) in stage III, we also allow different hyperparameters for different Radiomic groups (Radiomic components driven by different genomic

factors). The prior settings are:

For the *radiogenomic clinical model*,

$$\boldsymbol{Y} = \boldsymbol{\mathcal{I}}\boldsymbol{\alpha} + \boldsymbol{\epsilon}$$

$$\boldsymbol{Y} \sim Normal(\boldsymbol{\mathcal{I}}\boldsymbol{\alpha}, \sigma^2 \boldsymbol{I}_{\boldsymbol{N_{\mathcal{GIC}}}})$$

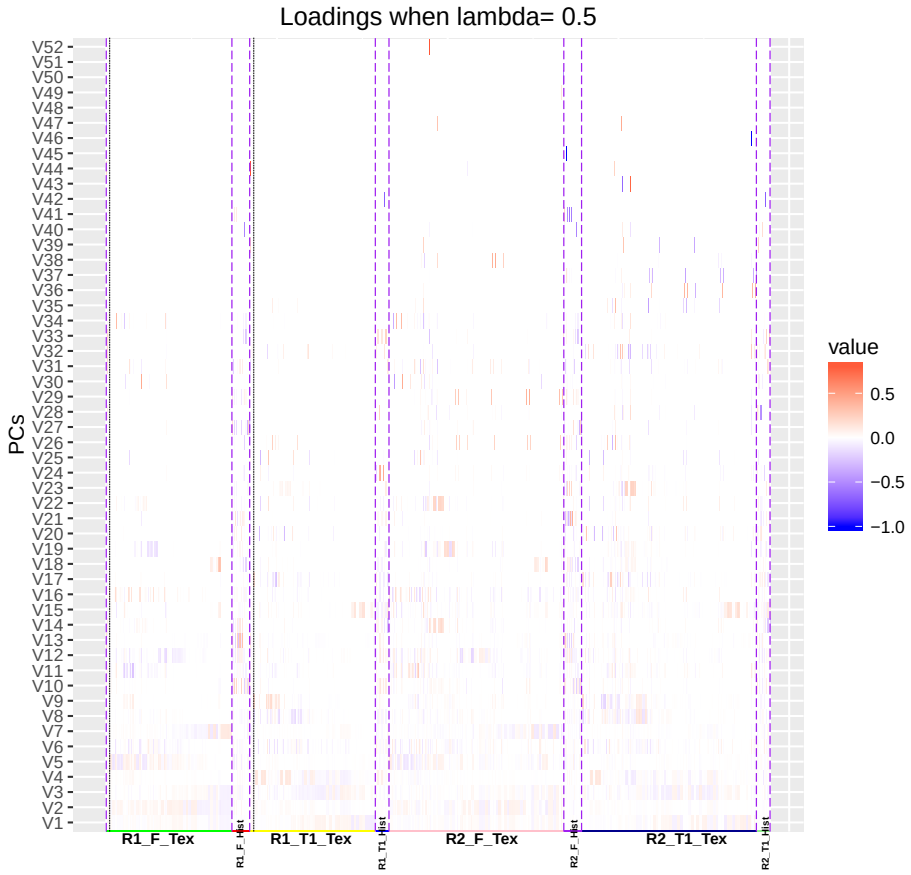
$$\boldsymbol{\alpha} \sim Normal(\mathbf{0}, D_\psi)$$

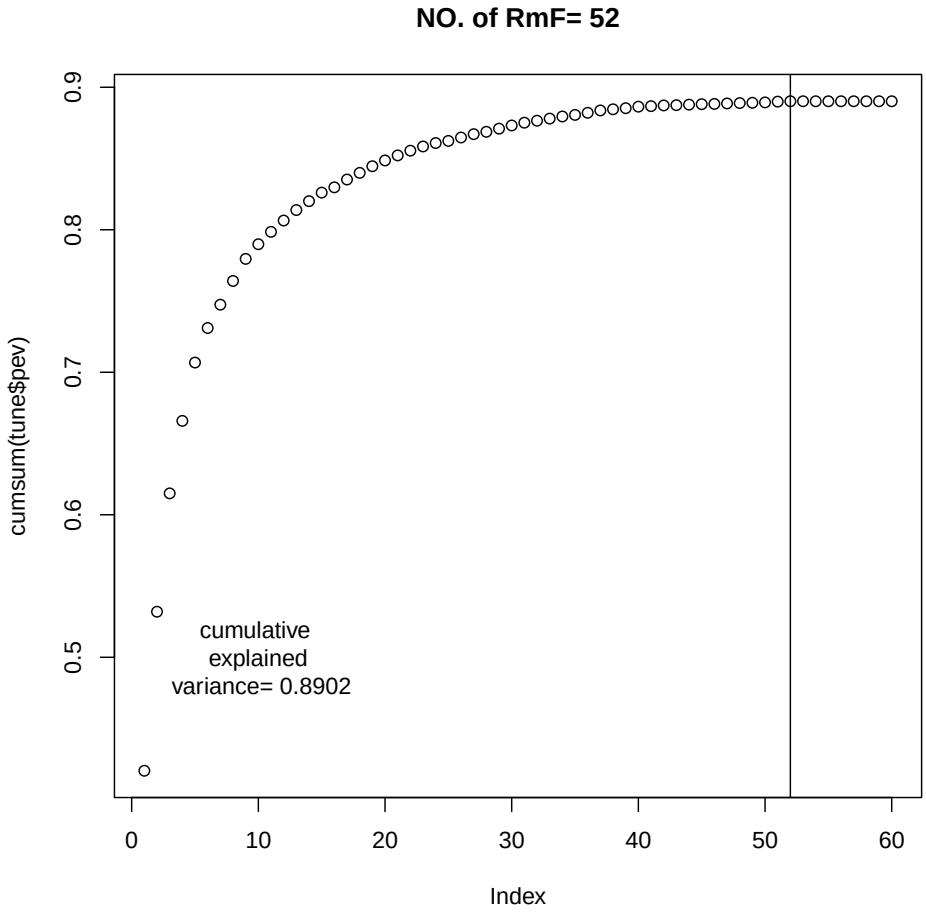
$$D_\psi = diag(\psi_{1,1}, \psi_{1,2}, ..., \psi_{1,K}, \psi_{2,1}, \psi_{2,2}, ..., \psi_{2,K}, ..., \psi_{J,1}, \psi_{J,2}, ..., \psi_{J,K}),$$

where J denotes the total number of different RF combination types ($j = 1, 2, 3, ..., J$, our $J = 4$), k denotes the RF index ($k = 1, 2, 3, ..., K$). Further, we assign our prior and hyper-prior distributions as $\psi_{j,k} \sim Gamma(1, 1/(2\gamma_j^2))$, $\sigma^2 \sim InverseGamma(u_1, u_2)$ and $1/(2\gamma_j^2) \sim Gamma(\tilde{e}, \tilde{f}/(2\lambda_j))$. Finally, RF combination selection is based on the posterior probability of the MCMC samples.

The result shows that Bayesian Lasso, compared with Normal Gamma prior, leads to less shrinkage. The result is illustrated in Fig S10. However, since Bayesian Lasso has one hyperparameter, which lacks enough flexibility, Normal Gamma is more preferred as the prior distribution in our analysis.

S6 Result Using Different Regularization Parameter in sPCA In our paper, the regularization parameter in implementing Sparse PCA Zou, Hastie and Tibshirani (2006) that we choose is $\lambda = 2.5$, which gave us proper balance between the variance that the leading PCs can explain, the sparsity of the loadings which leads to better interpretation, and the number of the leading RmF. However, we explored large range of the λs ($\lambda = 0.5, 1, 1.5, ..., 5$) and the corresponding loadings and the loadings are shown in below. As we can see from the loadings of the neighboring λ, $\lambda = 2$ and $\lambda = 3$, the main Radiomic features that contribute to the RmFs are similar. We also plot the results of stage III for $\lambda = 2$ and $\lambda = 3$ and we can see that, for example, RmF 7 driven by CN component is selected to be important when $\lambda = 2$, and when checking the loadings, we noticed that F_LoG_Tex_R1 and F_LoG_Tex_R2 count for the main part of the RmF 7, which is similar with our results. We chose $\lambda = 2.5$ with the total cumulative variance equal to 0.807. Thus, Sparse PCA is stable with the choice of the regularization parameters within a specific range, and the interpretation is similar as well.


Loadings when lambda= 0.5
PCs
V52
V51
V50
V49
V48
V47
V46
V45
V44
V43
V42
V41
V40
V39
V38
V37
V36
V35
V34
V33
V32
V31
V30
V29
V28
V27
V26
V25
V24
V23
V22
V21
V20
V19
V18
V17
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
R1_F_Tex
R1_F_Hist
R1_T1_Tex
R1_T1_Hist
R2_F_Tex
R2_F_Hist
R2_T1_Tex
R2_T1_Hist
value
0.5
0.0
−0.5
−1.0


NO. of RmF= 52
cumsum(tune$pev)
0.5
0.6
0.7
0.8
0.9
cumulative
explained
variance= 0.8902
0
10
20
30
40
50
60
Index

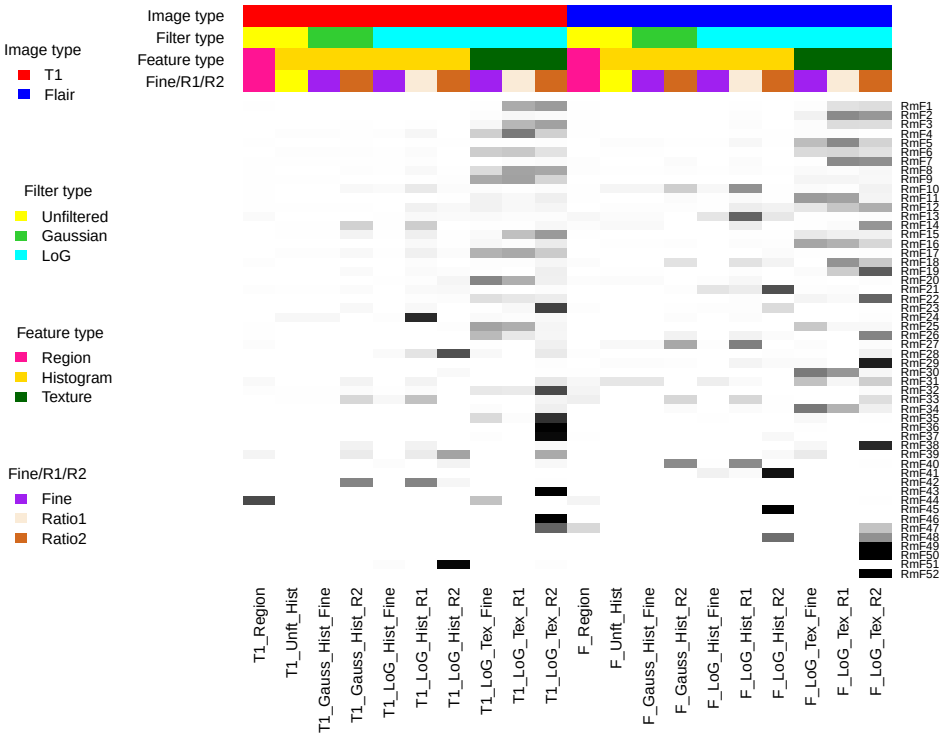
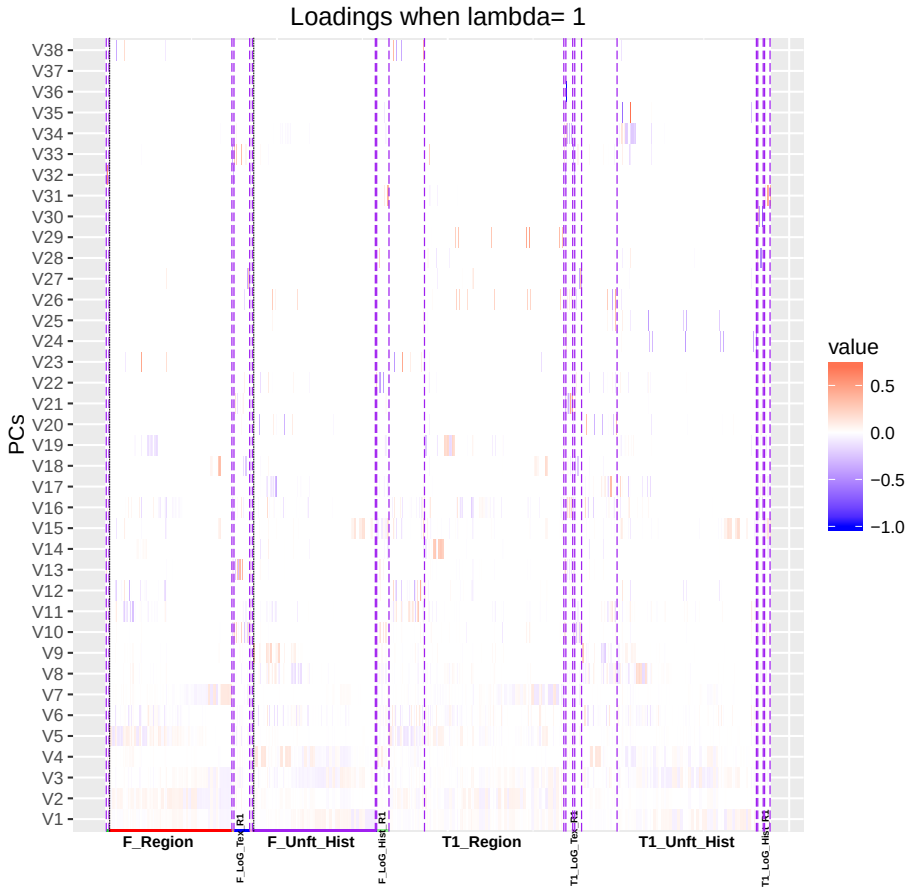
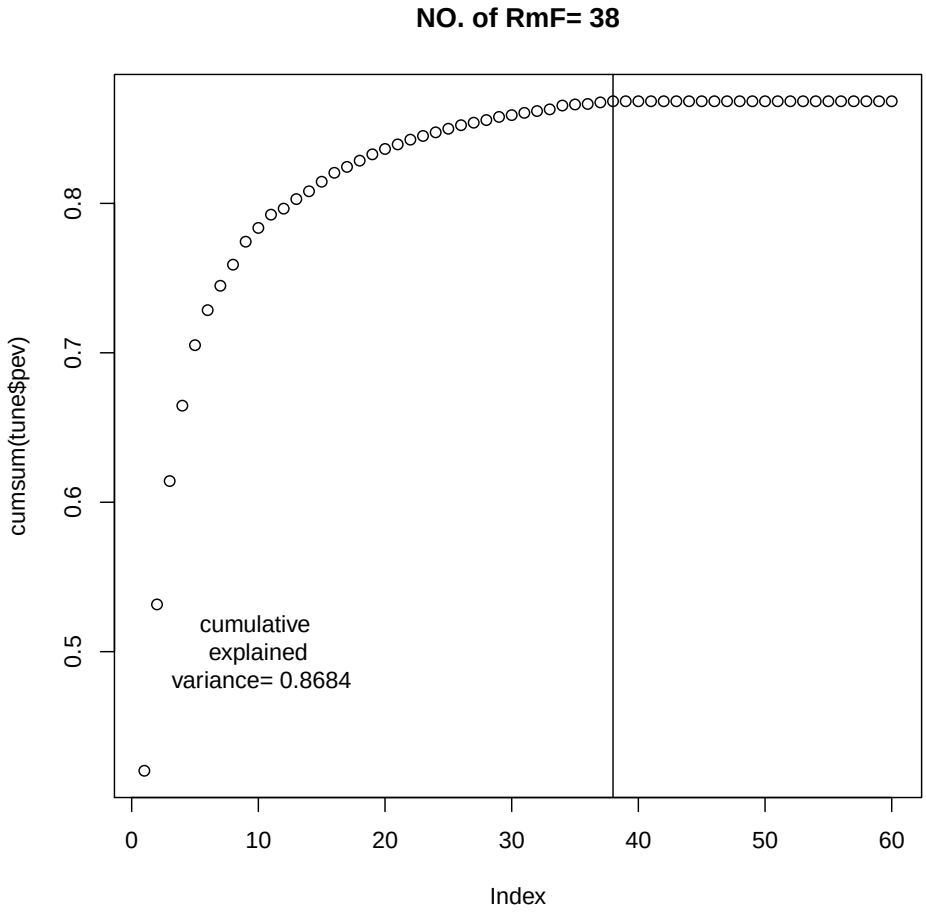


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
RmF17
RmF18
RmF19
RmF20
RmF21
RmF22
RmF23
RmF24
RmF25
RmF26
RmF27
RmF28
RmF29
RmF30
RmF31
RmF32
RmF33
RmF34
RmF35
RmF36
RmF37
RmF38
RmF39
RmF40
RmF41
RmF42
RmF43
RmF44
RmF45
RmF46
RmF47
RmF48
RmF49
RmF50
RmF51
RmF52
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 1
PCs
V38
V37
V36
V35
V34
V33
V32
V31
V30
V29
V28
V27
V26
V25
V24
V23
V22
V21
V20
V19
V18
V17
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Text_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Text_R1
T1_Unft_Hist
T1_LoG_Hist_R1
value
0.5
0.0
−0.5
−1.0


NO. of RmF= 38
cumsum(tune$pev)
0.5
0.6
0.7
0.8
cumulative
explained
variance= 0.8684
0
10
20
30
40
50
60
Index

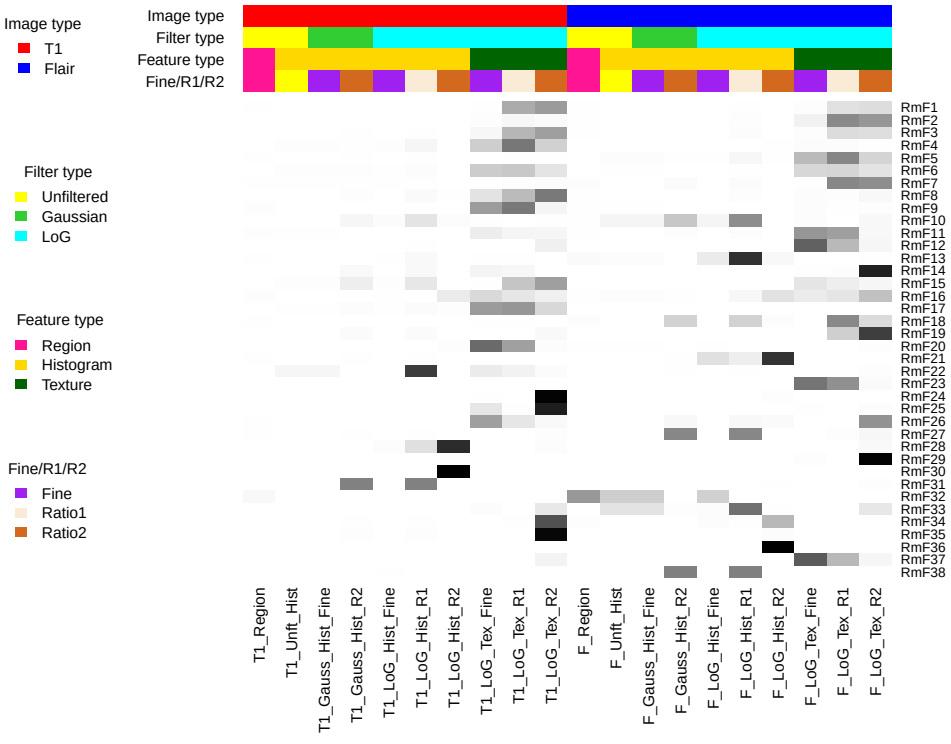
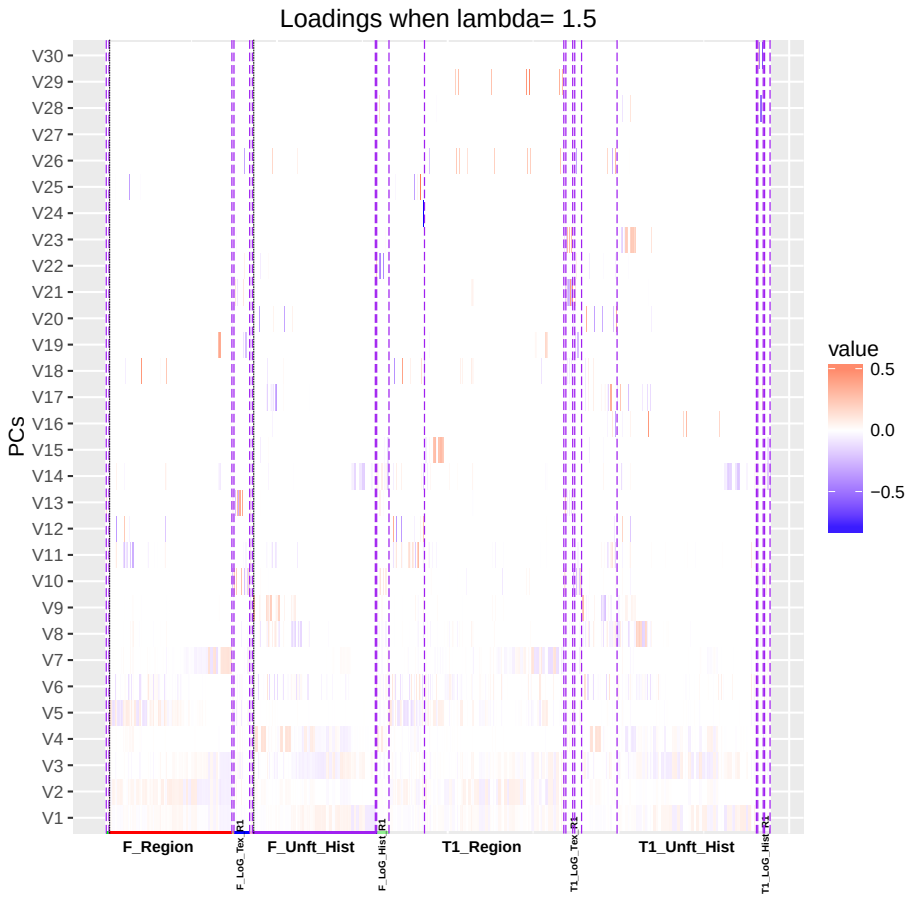
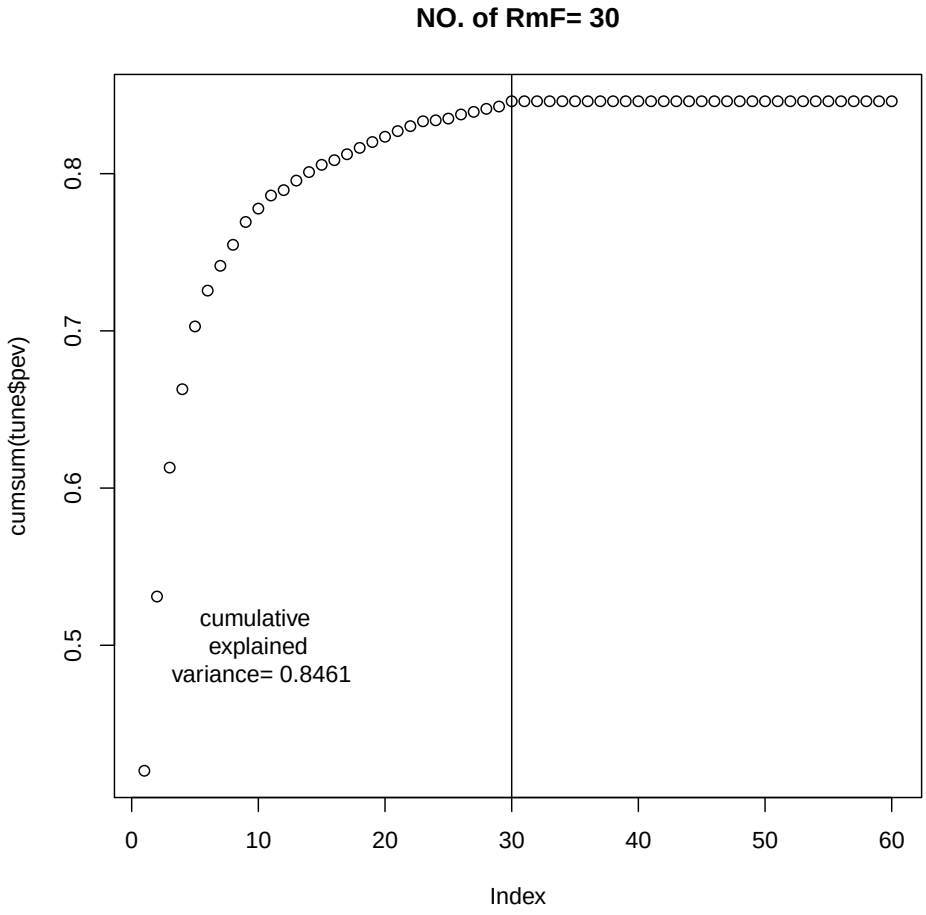


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
RmF17
RmF18
RmF19
RmF20
RmF21
RmF22
RmF23
RmF24
RmF25
RmF26
RmF27
RmF28
RmF29
RmF30
RmF31
RmF32
RmF33
RmF34
RmF35
RmF36
RmF37
RmF38
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 1.5
PCs
V30
V29
V28
V27
V26
V25
V24
V23
V22
V21
V20
V19
V18
V17
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Tex_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Tex_R1
T1_Unft_Hist
T1_LoG_Hist_R1
value
0.5
0.0
−0.5


NO. of RmF= 30
cumsum(tune$pev)
0.5
0.6
0.7
0.8
cumulative
explained
variance= 0.8461
0
10
20
30
40
50
60
Index

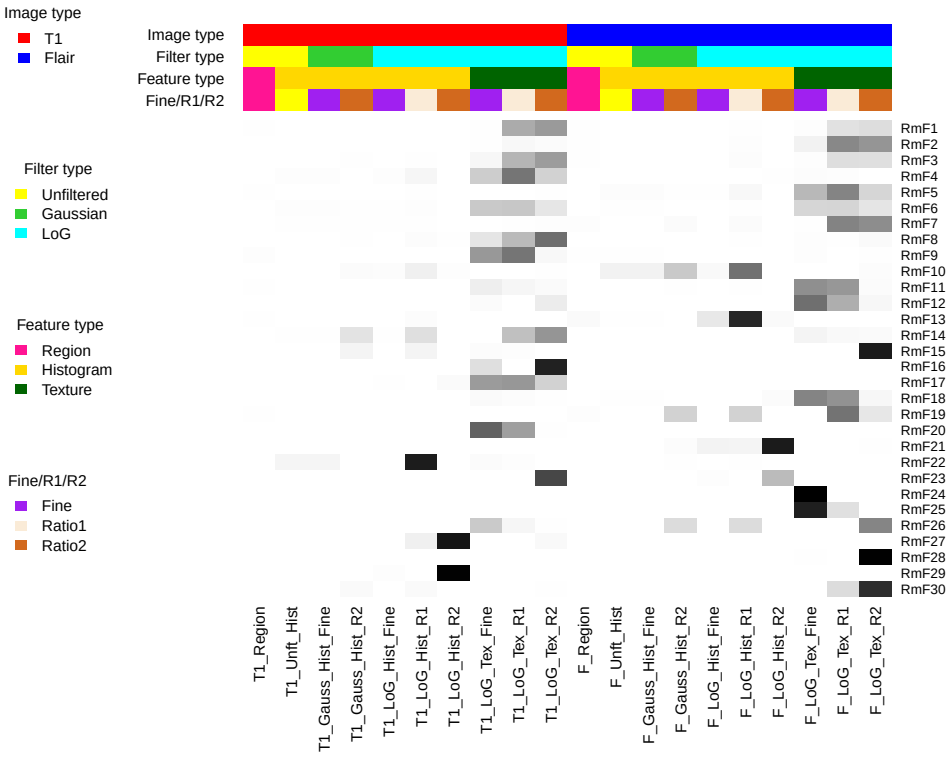
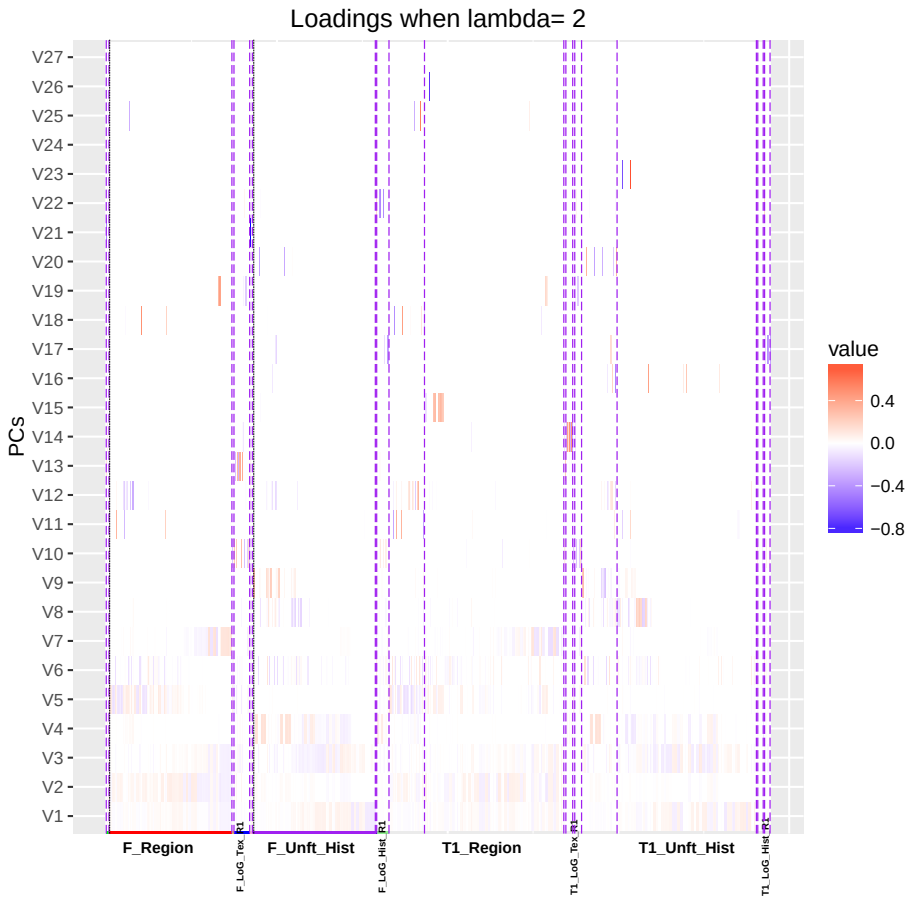
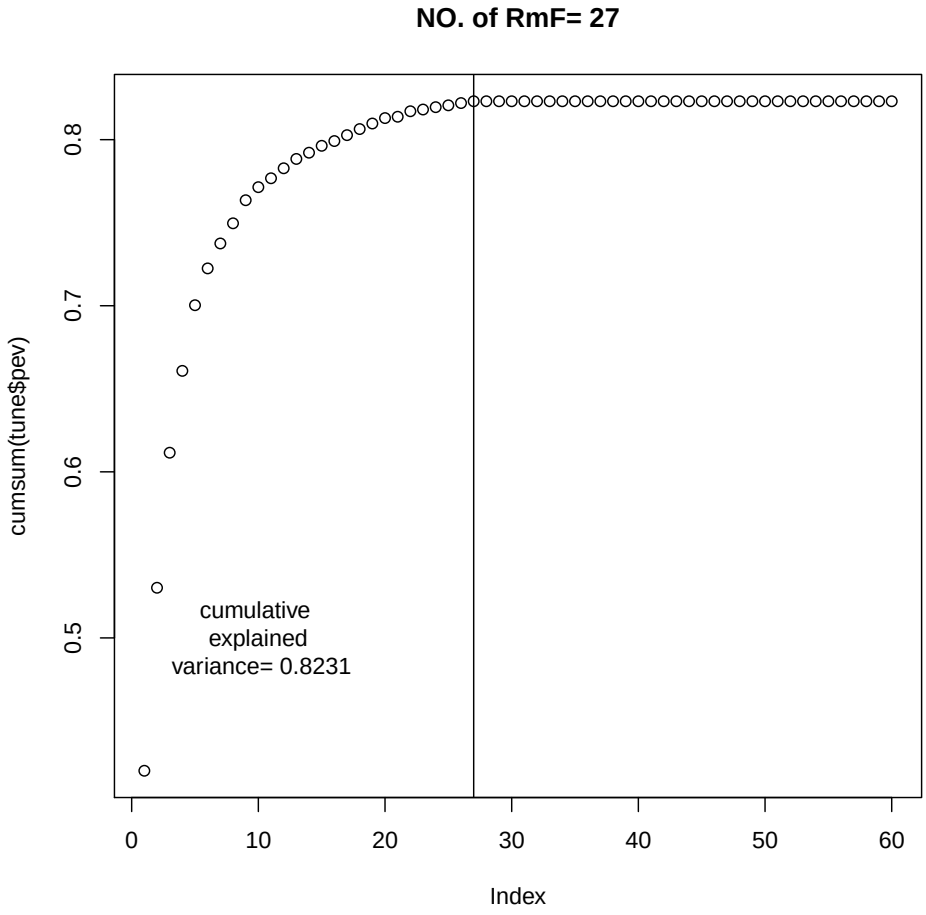


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
RmF17
RmF18
RmF19
RmF20
RmF21
RmF22
RmF23
RmF24
RmF25
RmF26
RmF27
RmF28
RmF29
RmF30
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 2
PCs
V27
V26
V25
V24
V23
V22
V21
V20
V19
V18
V17
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Tex_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Tex_R1
T1_Unft_Hist
T1_LoG_Hist_R1
value
0.4
0.0
−0.4
−0.8


NO. of RmF= 27
cumsum(tune$pev)
0.5
0.6
0.7
0.8
cumulative
explained
variance= 0.8231
0
10
20
30
40
50
60
Index

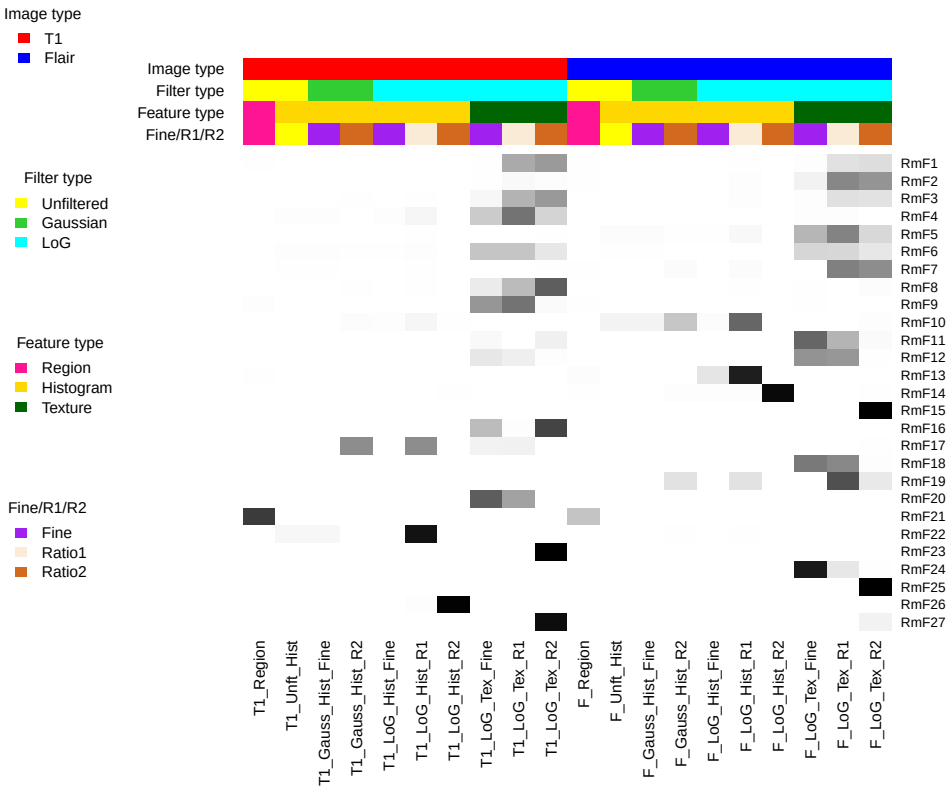
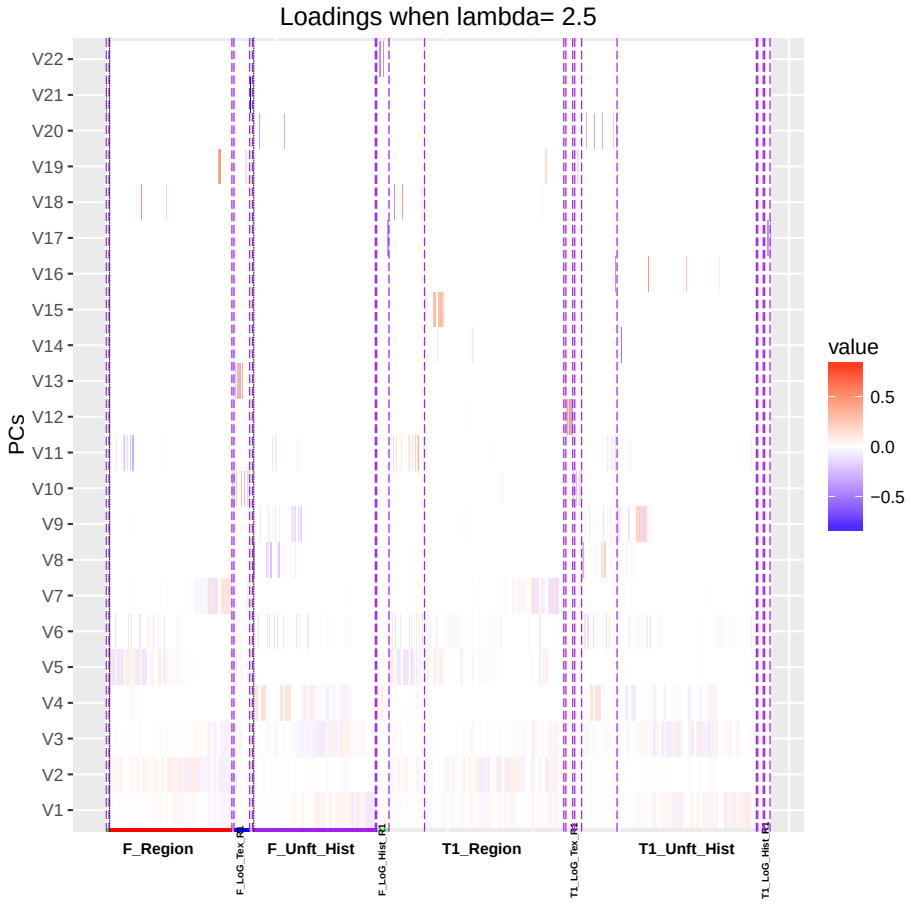
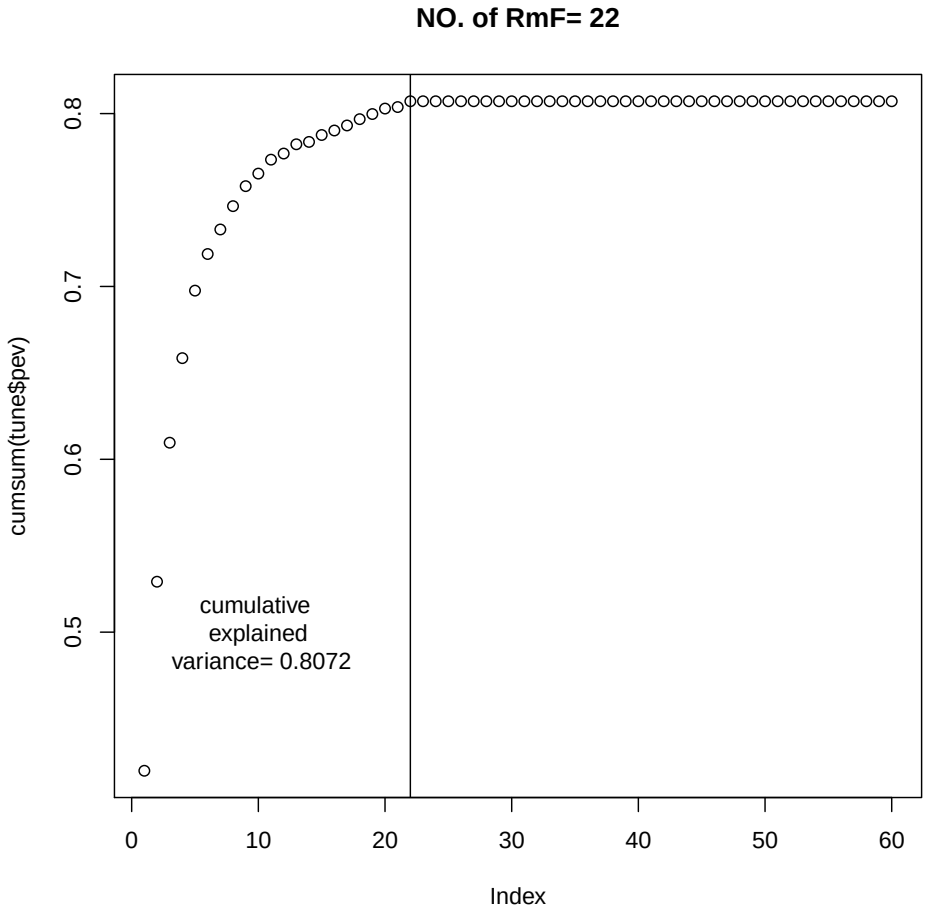


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
RmF17
RmF18
RmF19
RmF20
RmF21
RmF22
RmF23
RmF24
RmF25
RmF26
RmF27
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 2.5
PCs
V22
V21
V20
V19
V18
V17
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_Unft_Hist
T1_Region
T1_Unft_Hist
value
0.5
0.0
−0.5


NO. of RmF= 22
cumsum(tune$pev)
0.5
0.6
0.7
0.8
cumulative
explained
variance= 0.8072
0
10
20
30
40
50
60
Index

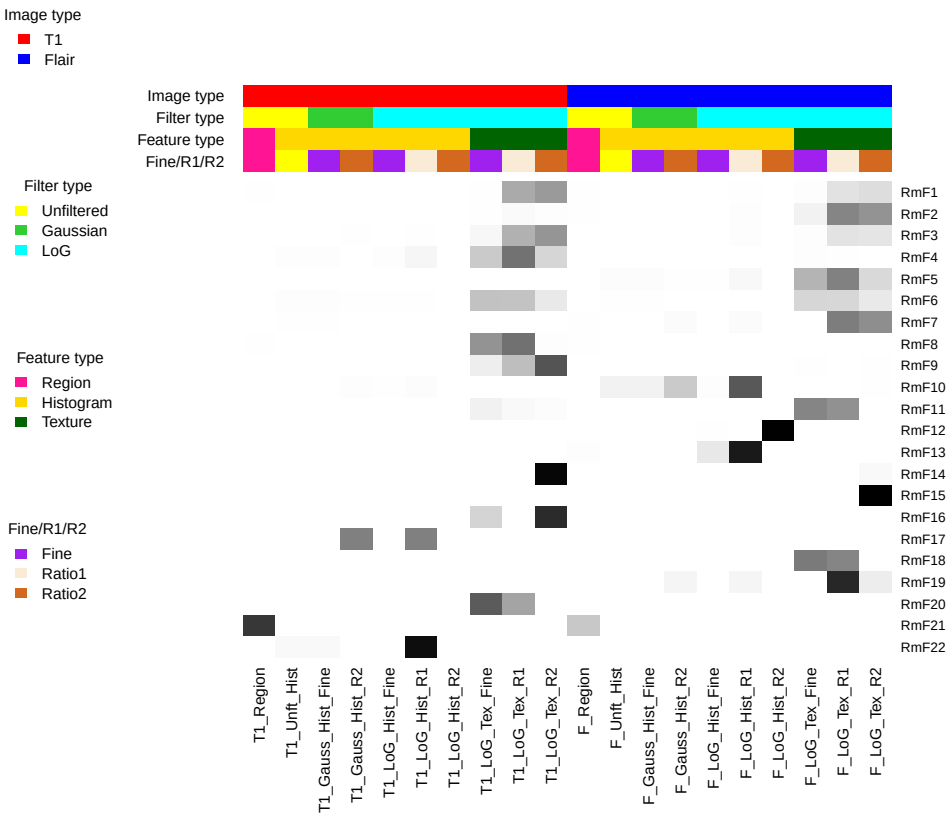
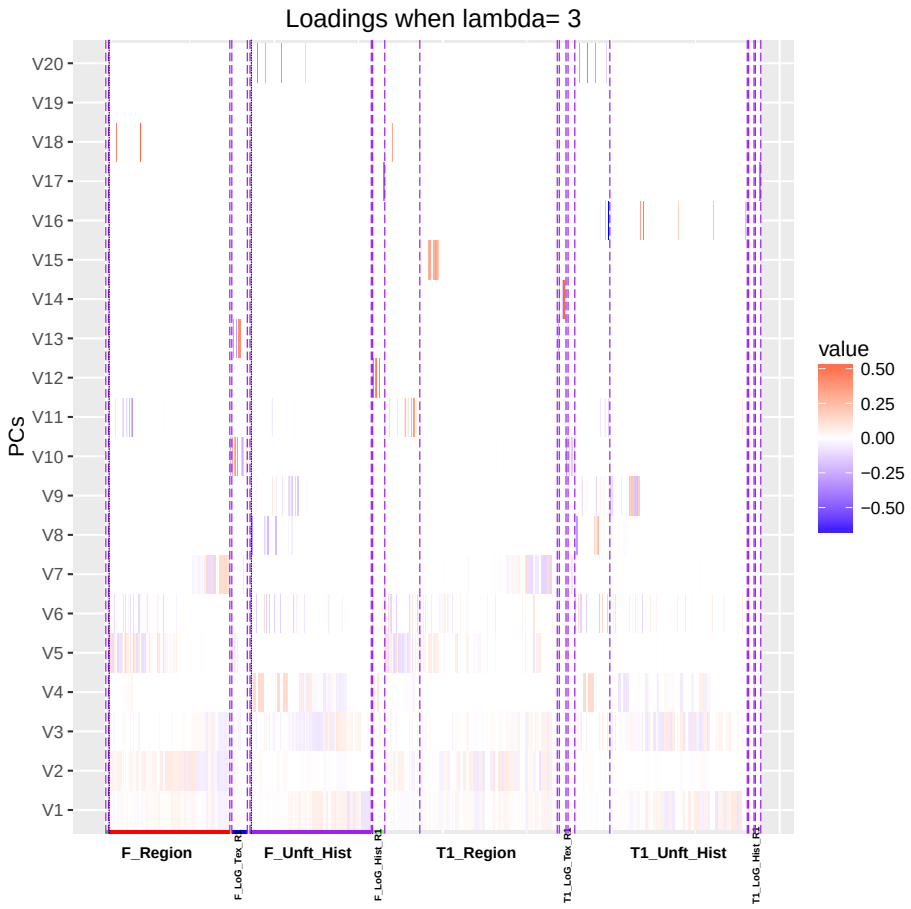
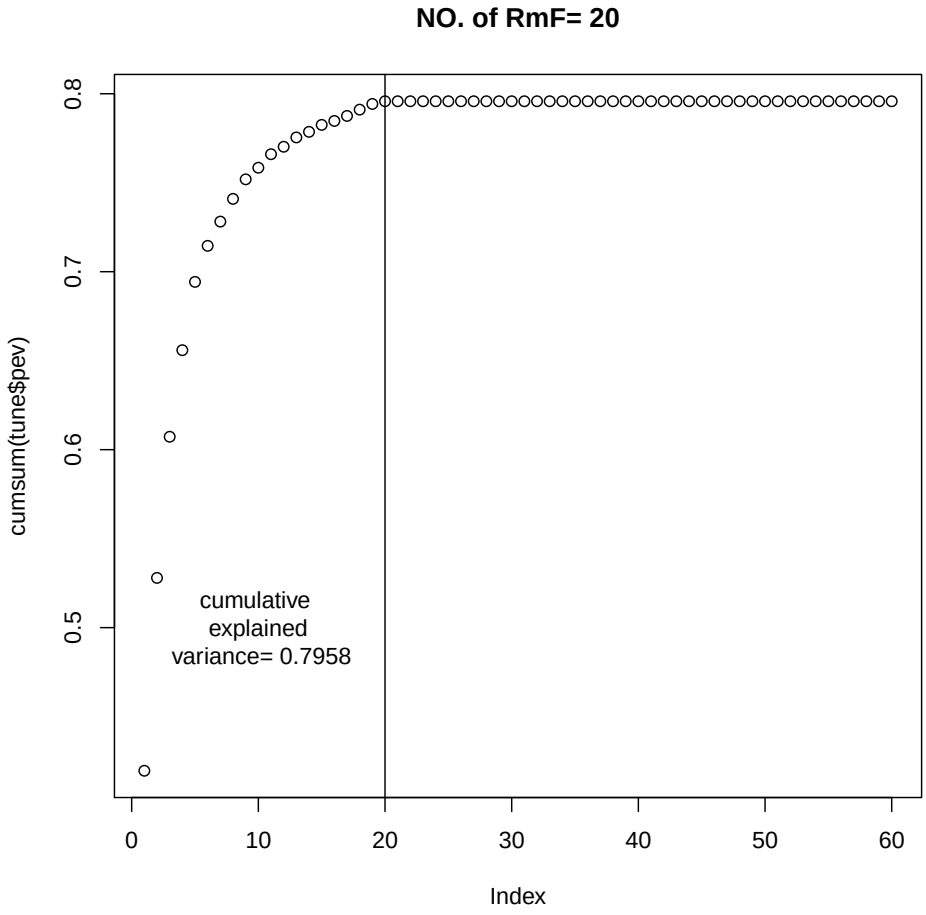


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
RmF17
RmF18
RmF19
RmF20
RmF21
RmF22
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 3
PCs
V20
V19
V18
V17
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Tex_Rgn
F_Unft_Hist
F_LoG_Hist_Rgn
T1_Region
T1_LoG_Tex_Rgn
T1_Unft_Hist
T1_LoG_Hist_Rgn
value
0.50
0.25
0.00
−0.25
−0.50


NO. of RmF= 20
cumsum(tune$pev)
0.5
0.6
0.7
0.8
cumulative
explained
variance= 0.7958
0
10
20
30
40
50
60
Index

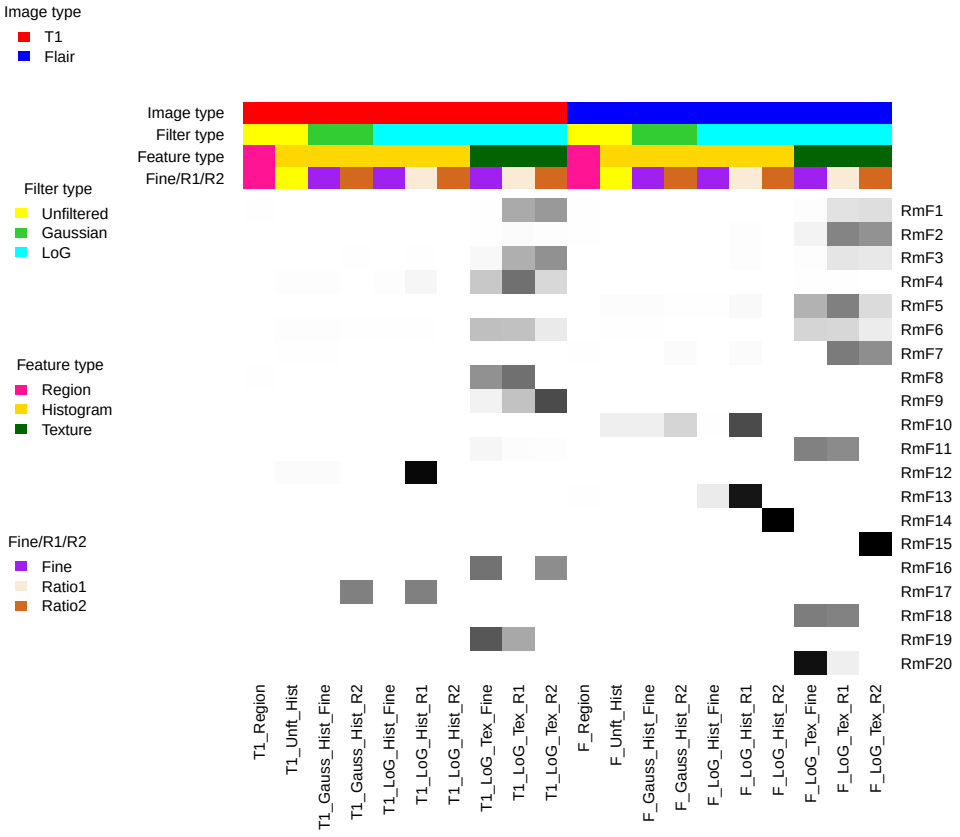
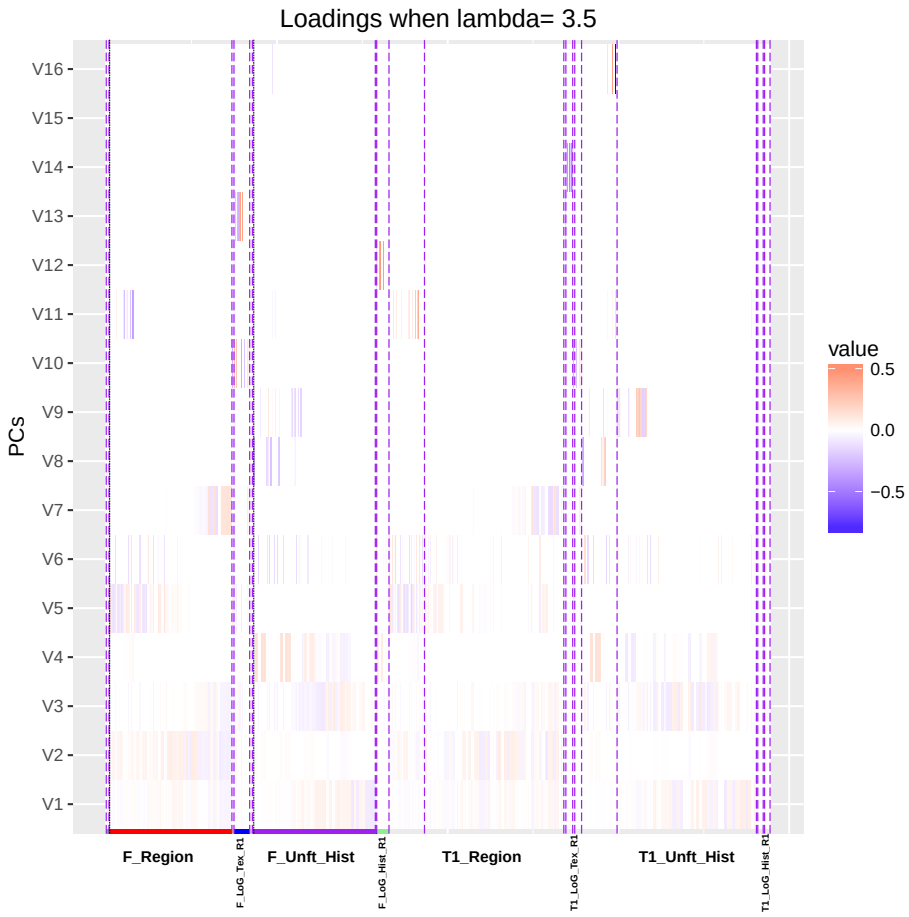
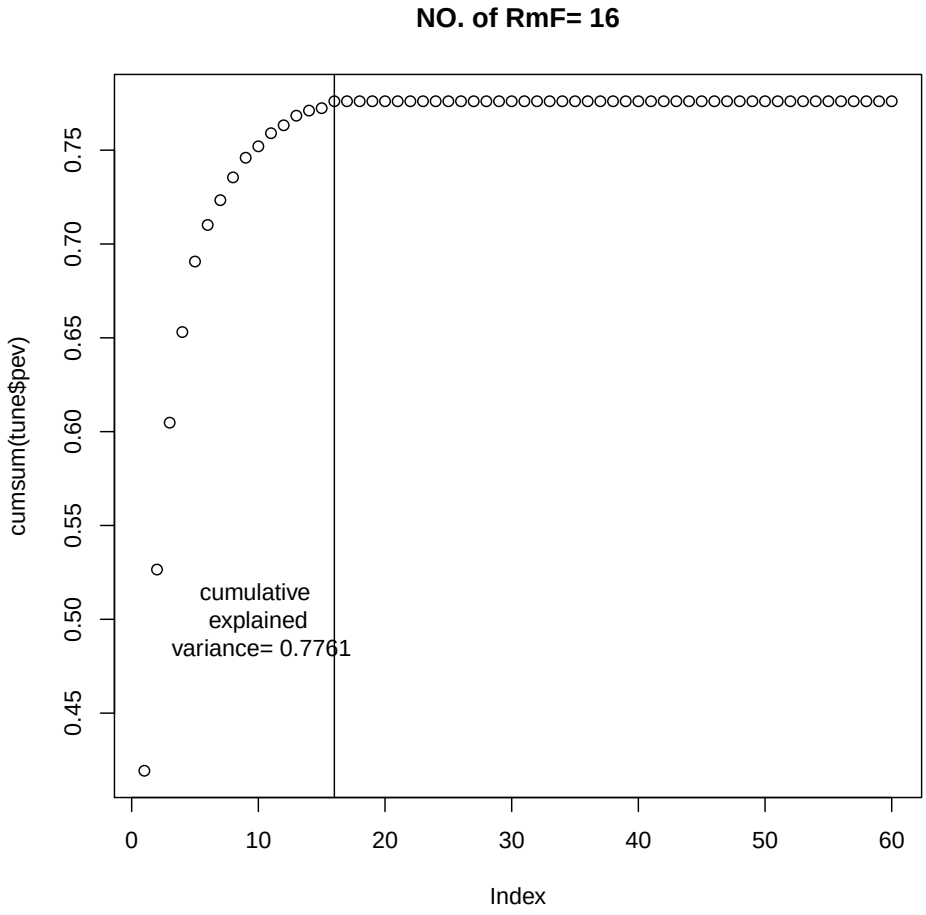


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
RmF17
RmF18
RmF19
RmF20
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 3.5
PCs
V16
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Tex_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Tex_R1
T1_Unft_Hist
T1_LoG_Hist_R1
value
0.5
0.0
−0.5


NO. of RmF= 16
cumsum(tune$pev)
0.45
0.50
0.55
0.60
0.65
0.70
0.75
cumulative
explained
variance= 0.7761
0
10
20
30
40
50
60
Index

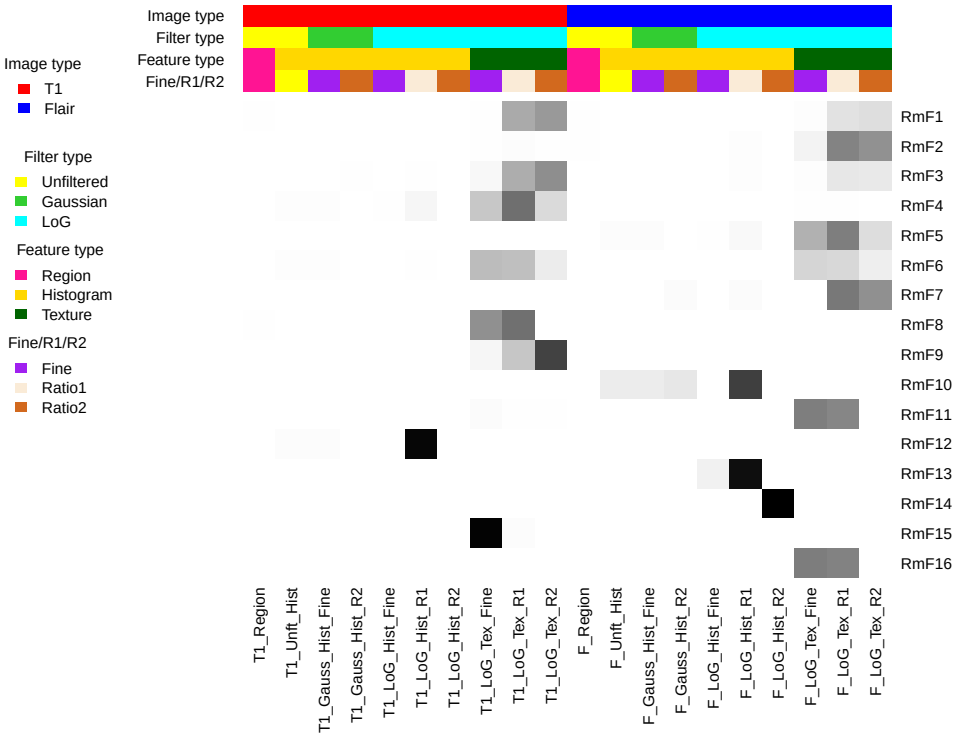
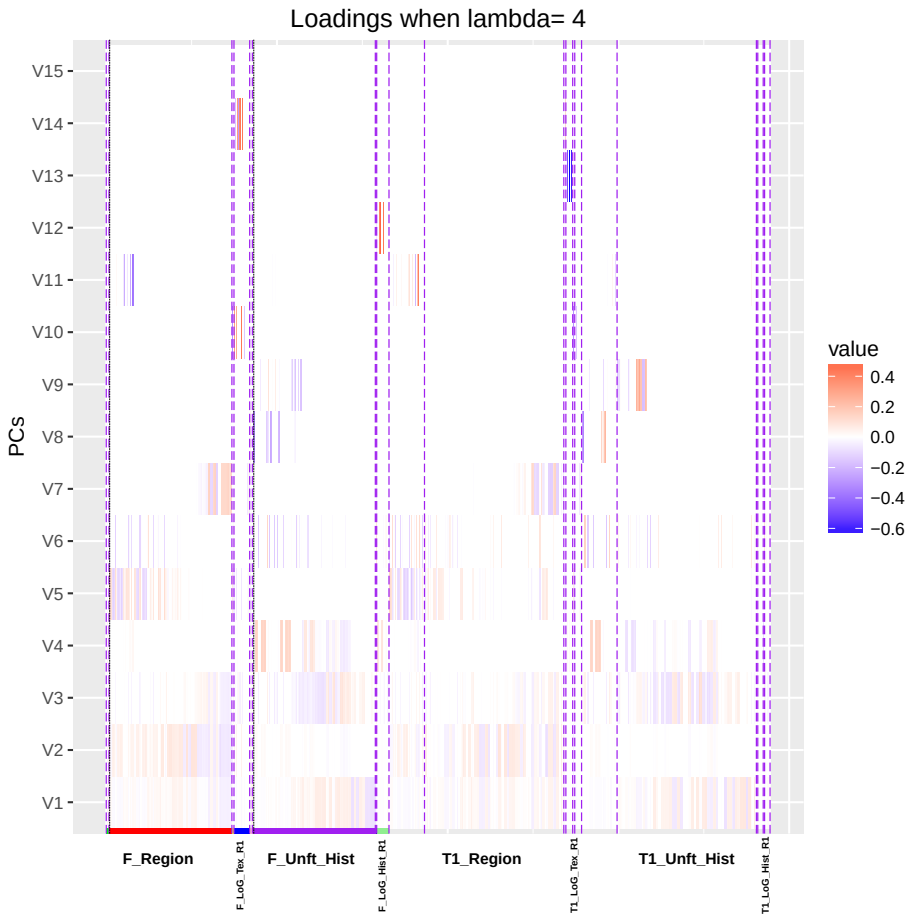
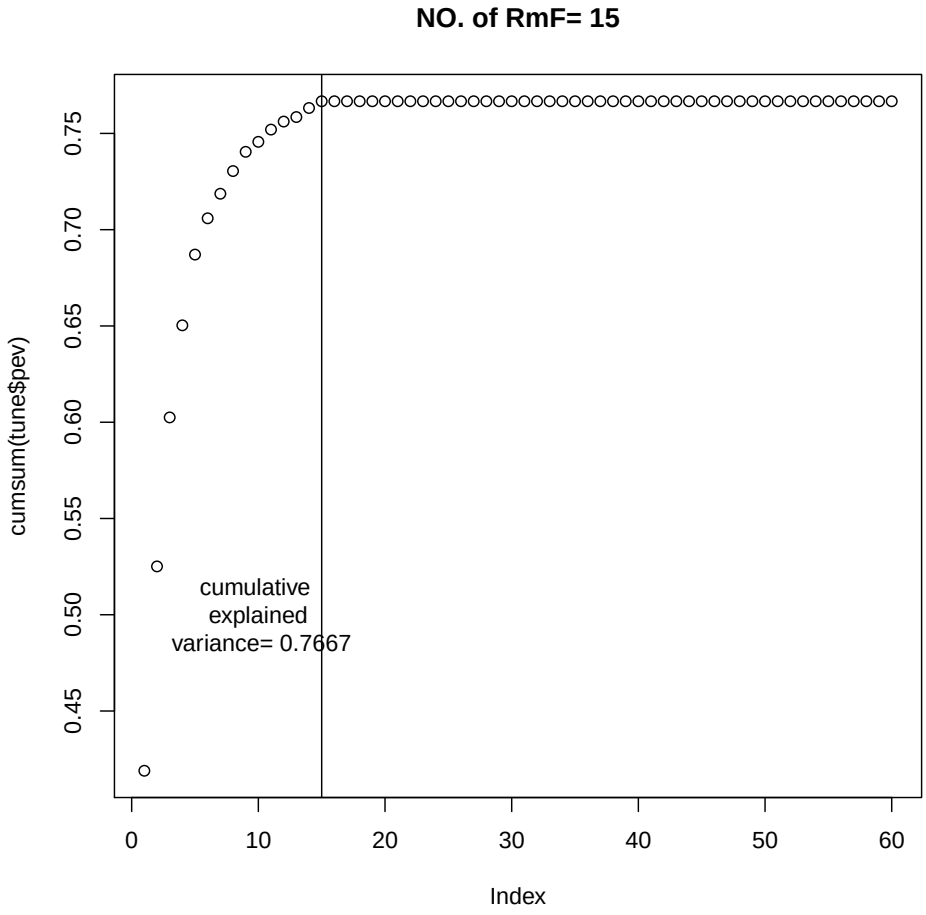


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
RmF16
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 4
PCs
V15
V14
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Tex_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Tex_R1
T1_Unft_Hist
T1_LoG_Hist_R1
value
0.4
0.2
0.0
−0.2
−0.4
−0.6


NO. of RmF= 15
cumsum(tune$pev)
0.45
0.50
0.55
0.60
0.65
0.70
0.75
cumulative
explained
variance= 0.7667
0
10
20
30
40
50
60
Index

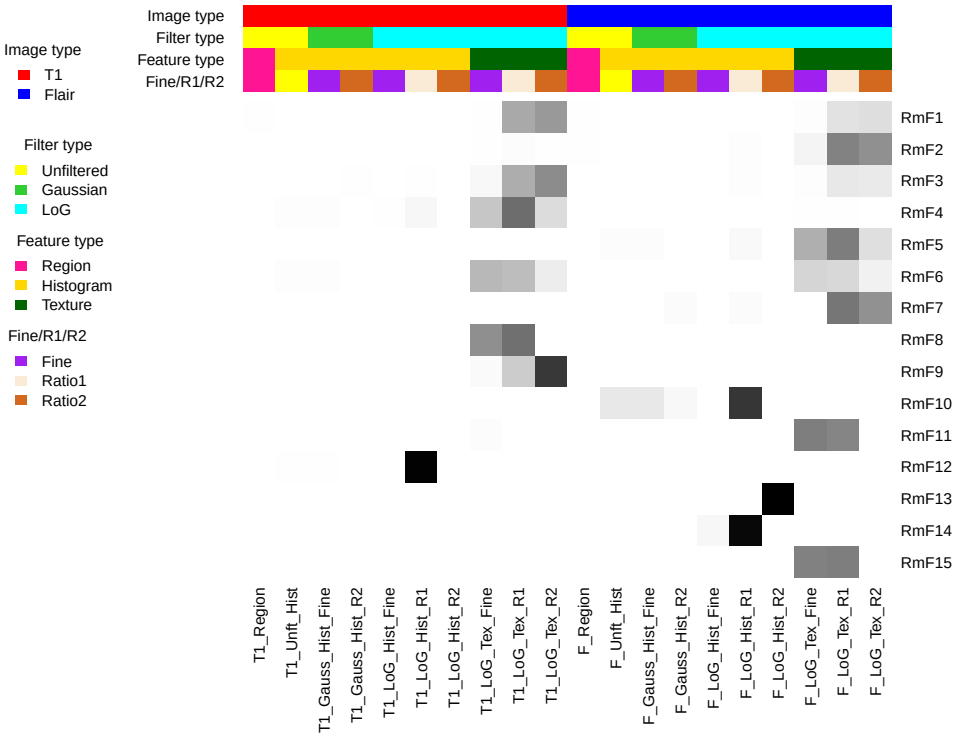
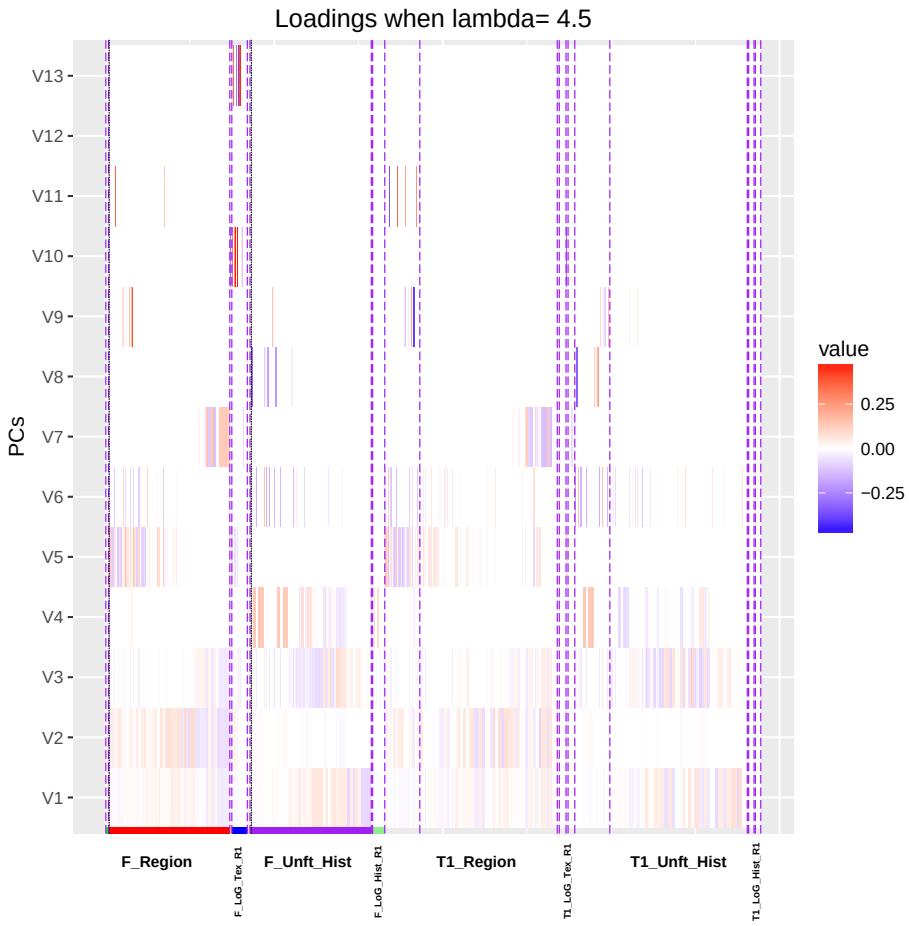
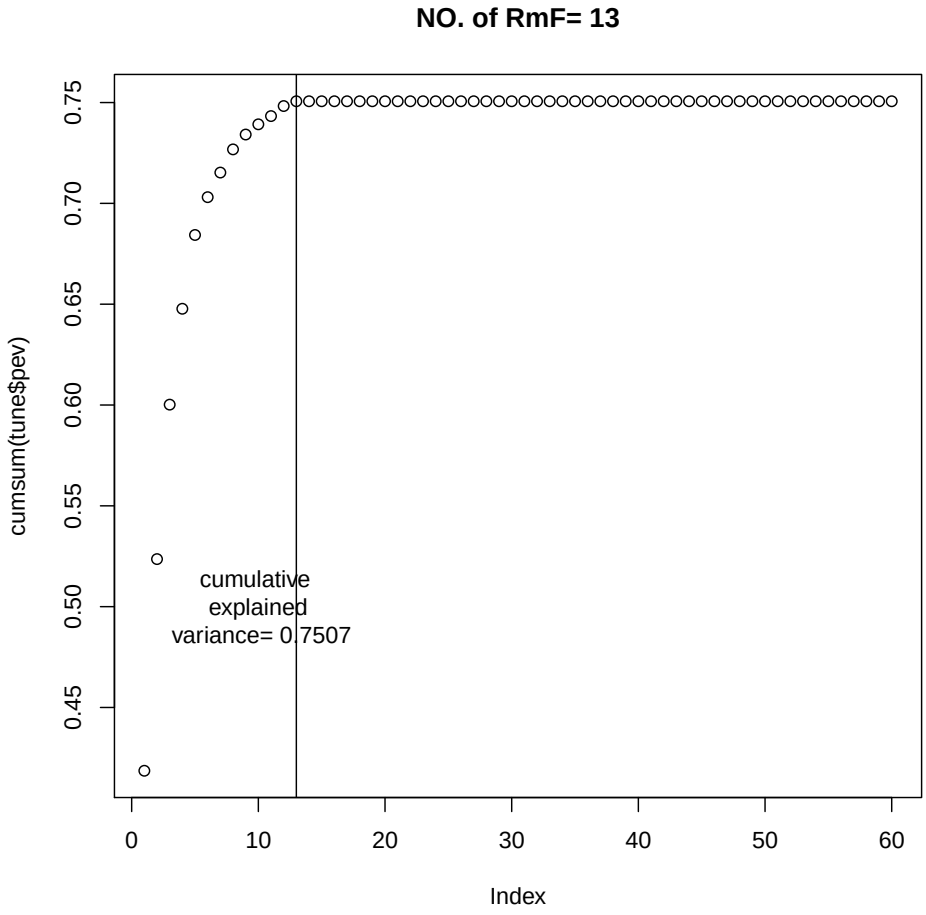


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
RmF14
RmF15
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 4.5
PCs
V13
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
value
0.25
0.00
−0.25
F_Region
F_LoG_Tex_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Tex_R1
T1_Unft_Hist
T1_LoG_Hist_R1


NO. of RmF= 13
cumsum(tune$pev)
0.45
0.50
0.55
0.60
0.65
0.70
0.75
cumulative
explained
variance= 0.7507
0
10
20
30
40
50
60
Index

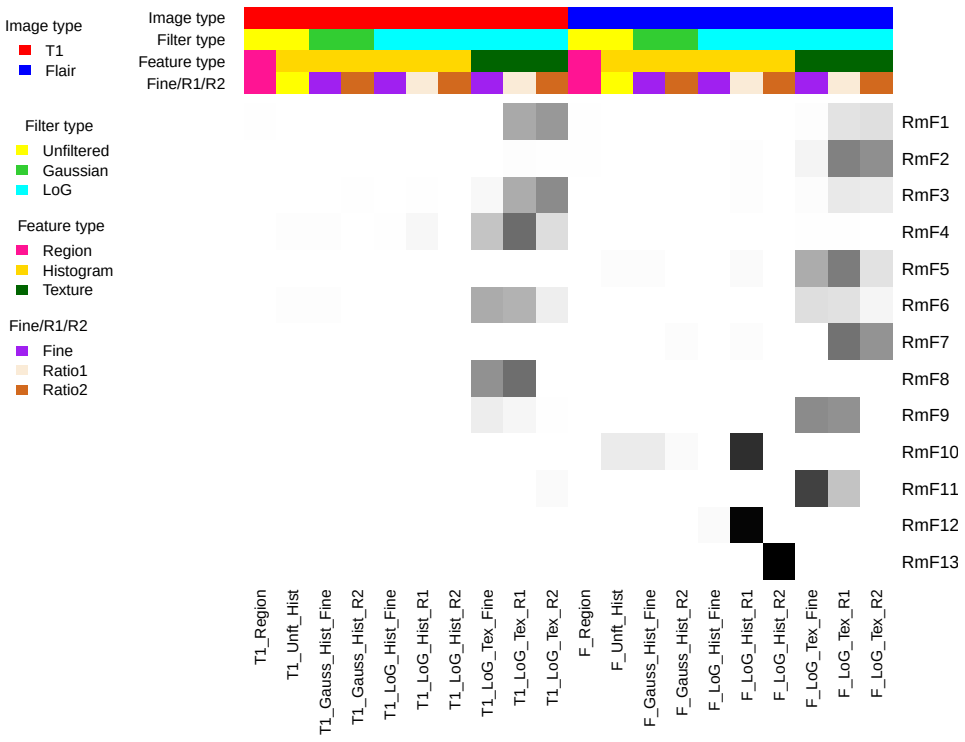
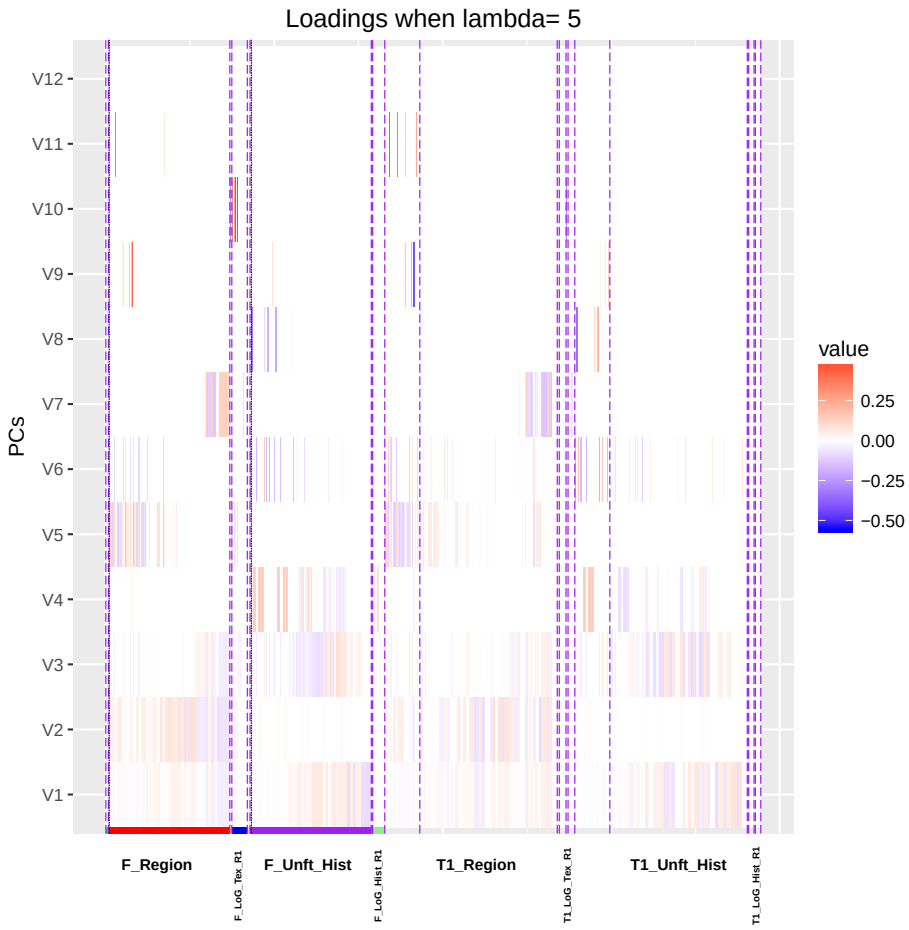
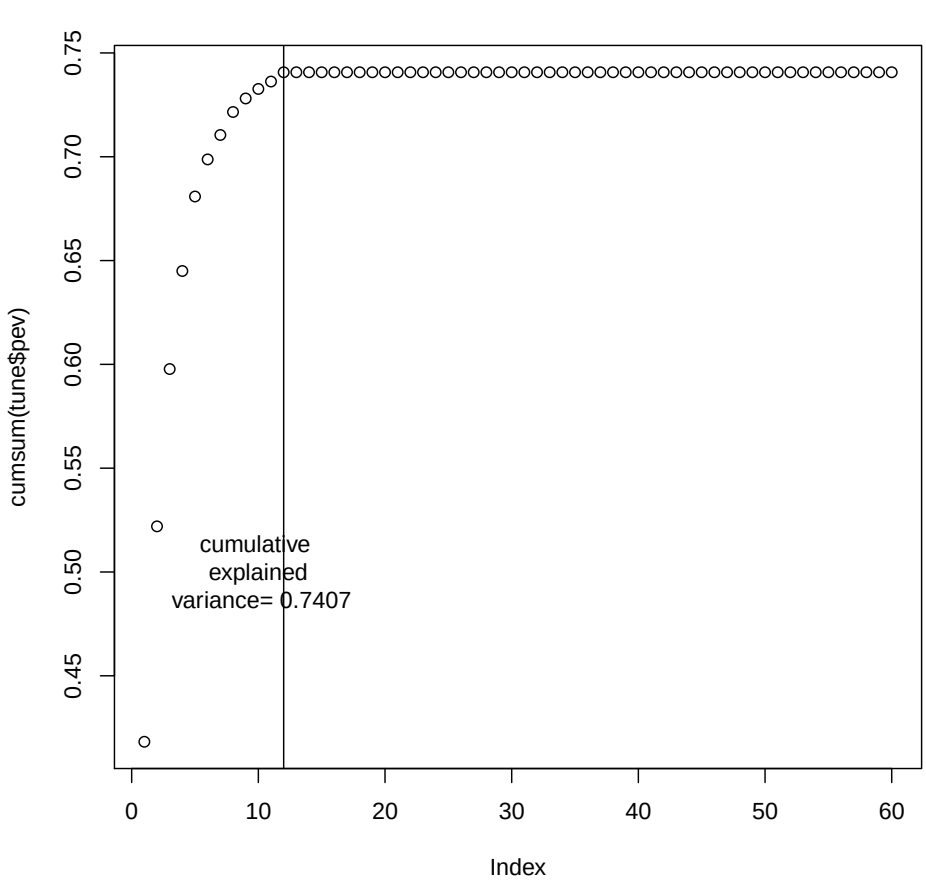


Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
RmF13
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2


Loadings when lambda= 5
PCs
V12
V11
V10
V9
V8
V7
V6
V5
V4
V3
V2
V1
F_Region
F_LoG_Tex_R1
F_Unft_Hist
F_LoG_Hist_R1
T1_Region
T1_LoG_Tex_R1
T1_Unft_Hist
T1_LoG_Hist_R1
value
0.25
0.00
−0.25
−0.50


NO. of RmF= 12
cumsum(tune$pev)
0.45
0.50
0.55
0.60
0.65
0.70
0.75
cumulative
explained
variance= 0.7407
0
10
20
30
40
50
60
Index

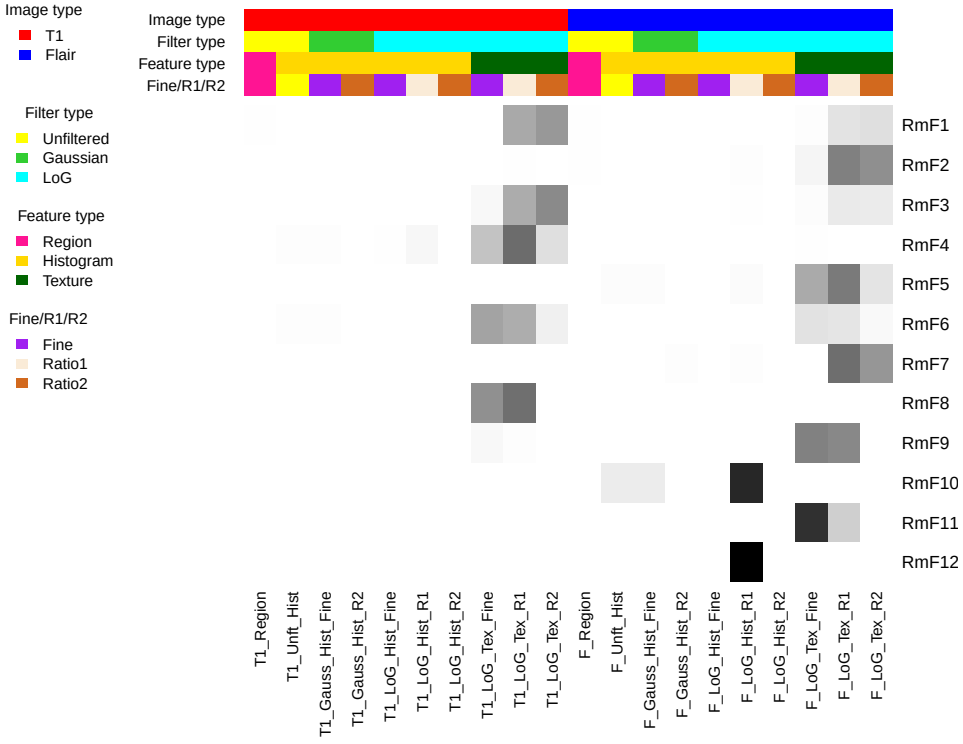
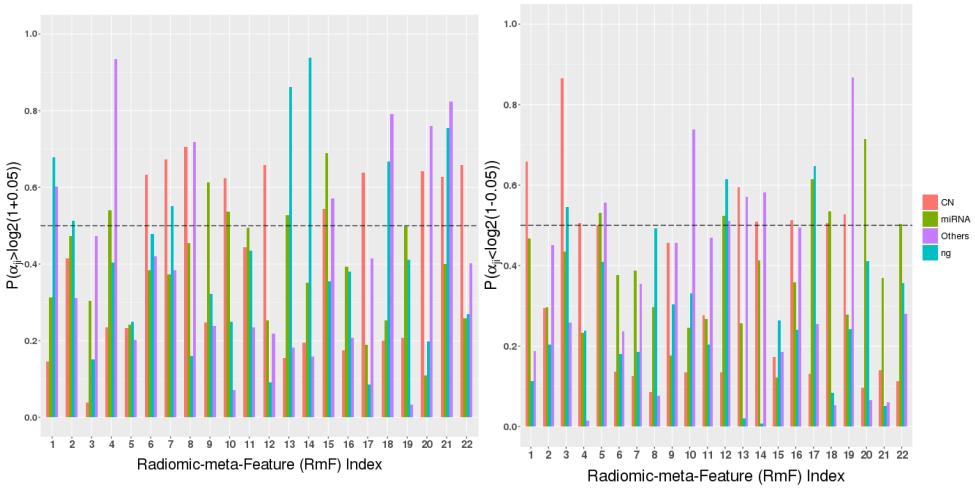


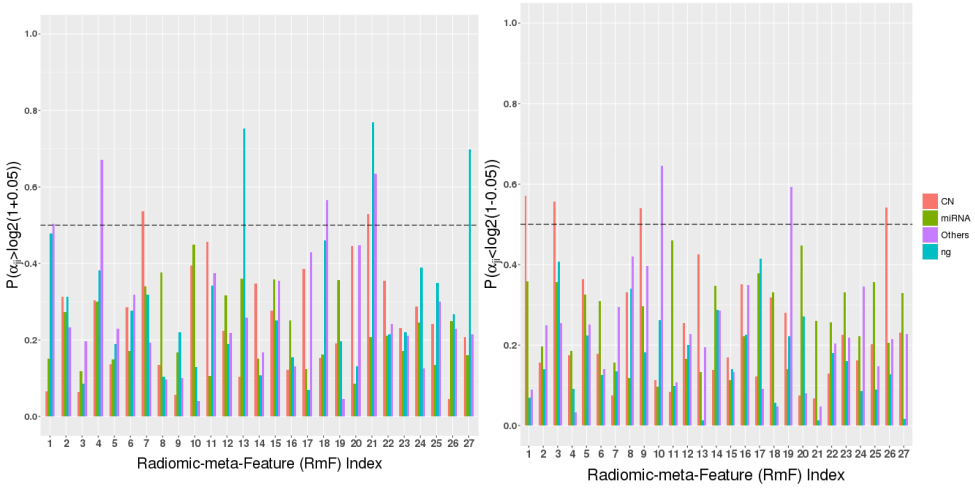
Image type
T1
Flair
Filter type
Unfiltered
Gaussian
LoG
Feature type
Region
Histogram
Texture
Fine/R1/R2
Fine
Ratio1
Ratio2
Image type
Filter type
Feature type
Fine/R1/R2
RmF1
RmF2
RmF3
RmF4
RmF5
RmF6
RmF7
RmF8
RmF9
RmF10
RmF11
RmF12
T1_Region
T1_Unft_Hist
T1_Gauss_Hist_Fine
T1_Gauss_Hist_R2
T1_LoG_Hist_Fine
T1_LoG_Hist_R1
T1_LoG_Hist_R2
T1_LoG_Tex_Fine
T1_LoG_Tex_R1
T1_LoG_Tex_R2
F_Region
F_Unft_Hist
F_Gauss_Hist_Fine
F_Gauss_Hist_R2
F_LoG_Hist_Fine
F_LoG_Hist_R1
F_LoG_Hist_R2
F_LoG_Tex_Fine
F_LoG_Tex_R1
F_LoG_Tex_R2



(a) positively significant RmF combinations

(b) negatively significant RmF combinations

Fig S10: Results of stage III (*radiogenomic clinical model*): Detecting postively and negatively significant RmF combinations. Each RmF is segmented into 4 parts, of which 3 parts are modulated by different genomic platform combinations denoted as $\boldsymbol{\mathcal{I}_{CN}}, \boldsymbol{\mathcal{I}_{miR}}$, and $\boldsymbol{\mathcal{I}_{O}}$. The 4^{th} part is modulated by unknown/unmeasured factors represented as $\boldsymbol{\mathcal{I}_{\bar{g}}}$ ("ng" in the legend). The barplot shows the posterior probabilities that the coefficient for each part $\alpha_{jk} > \delta_{+}^{*}$, where α_{jk} denotes the k^{th} RmF modulated by the j^{th} genomic platform. For each RmF, the probabilities of these 4 components, CN, miRNA, others, and ng, are respectively shown in red, green, purple and blue. Each probability in Figure (a) shows that 1 unit increment in the RmF component leads to at least 5% increase in survival time. Each probability in Figure (b) shows that 1 unit increment in the RmF component leads to at least 5% decrease in survival time. We consider the markers to be significant if this posterior probability is larger than 0.5.



(a) positively significant RmF combinations

(b) negatively significant RmF combinations

Fig S11: Results of Stage III when $\lambda = 2$

S7 Sensitivity Analysis Based on the prior settings described in the paper, for stage III, $\boldsymbol{\mathcal{I}} = \{\boldsymbol{\mathcal{I}_{CN}}, \boldsymbol{\mathcal{I}_{miR}}, \boldsymbol{\mathcal{I}_{O}}, \boldsymbol{\mathcal{I}_{\bar{g}}}\}$, and the effect parameter $\boldsymbol{\alpha} = \{\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}, \boldsymbol{\alpha_4}\}$, then the model and prior construction can be expressed as

$$\boldsymbol{Y} = \boldsymbol{\mathcal{I}\alpha} + \boldsymbol{\epsilon}$$

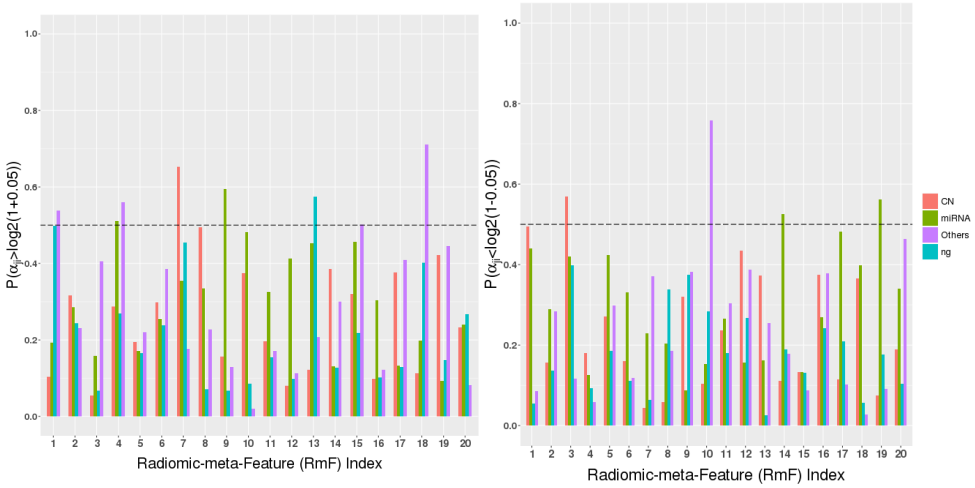
$$\boldsymbol{Y} \sim Normal(\boldsymbol{\mathcal{I}\alpha}, \sigma^2 \boldsymbol{I_{N_{\mathcal{GIC}}}})$$

$$\boldsymbol{\alpha} \sim Normal(\boldsymbol{0}, D_\psi)$$

$$D_\psi = diag(\psi_{1,1}, \psi_{1,2}, ..., \psi_{1,K}, \psi_{2,1}, \psi_{2,2}, ..., \psi_{2,K}, ..., \psi_{J,1}, \psi_{J,2}, ..., \psi_{J,K}),$$

where J denotes the total number of different RF combination types ($j = 1, 2, 3, ..., J$, our $J = 4$), k denotes the RF index ($k = 1, 2, 3, ..., K$). Further, we assign our prior and hyper-prior distributions as $\psi_{j,k} \sim Gamma(\lambda_j, 1/(2\gamma_j^2))$, $\sigma^2 \sim InverseGamma(u_1, u_2)$, $\lambda_j \sim exp(d)$, and $1/(2\gamma_j^2) \sim Gamma(\tilde{e}, \tilde{f}/(2\lambda_j))$.

We have hyperparameters d, $\tilde{e}$ and $\tilde{f}$. For hyperparameter $\tilde{f}$, it is suggested that $\tilde{f}$ comes from minimum-length least squares (MLLS) of the coefficients. Thus, we do the sensitivity analysis by adjusting d, $\tilde{e}$. In our analysis, we set up $d = 1$ and $\tilde{e} = 2$, in the sensitivity analysis, we set up $d = 0.5, 2$ and $\tilde{e} = 1, 4$ respectively.



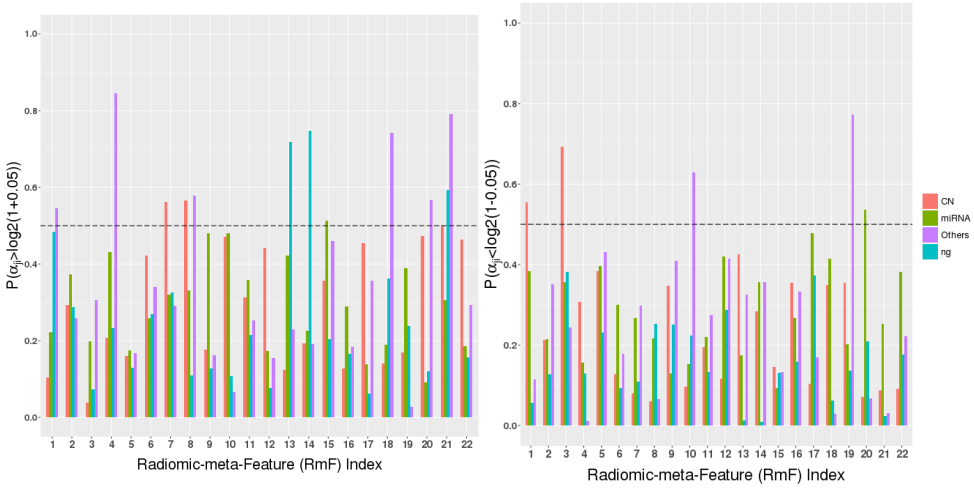
(a) positively significant RmF combinations

(b) negatively significant RmF combinations

Fig S12: Results of Stage III when $\lambda = 3$

Results in below show that the selected RmF components are almost the same across different hyperparameter settings. Even for some of the features which are not selected given lower value of d, they are close to the margin. Thus, our model is consistent with the hyperparameter settings. Since we applied the same prior for stage II, similar results will be drawn.

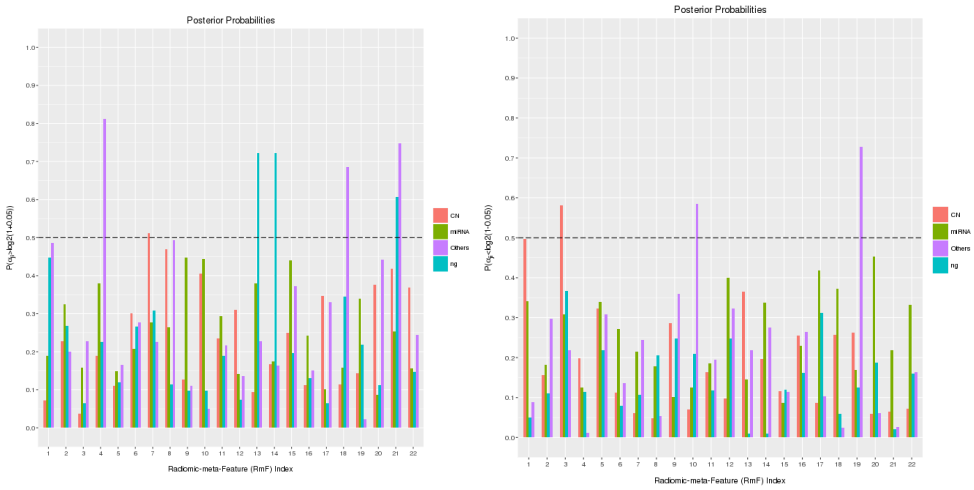
S8 Computing Time Checking We check the scale of the computing time for multiple numbers of genes that we analyze and the numbers of the RmFs. Following figures show that our computing time increases linearly with the number of the RmF, and with nonlinear trend with the number of genes. All the computing is implemented using computing machine with two 8-core Xeon E5-2690 (2.90GHz) processors and with 128G of memory. In stage II, we apply parallel computing with multiple cores.



(a) positively significant RmF combinations

(b) negatively significant RmF combinations

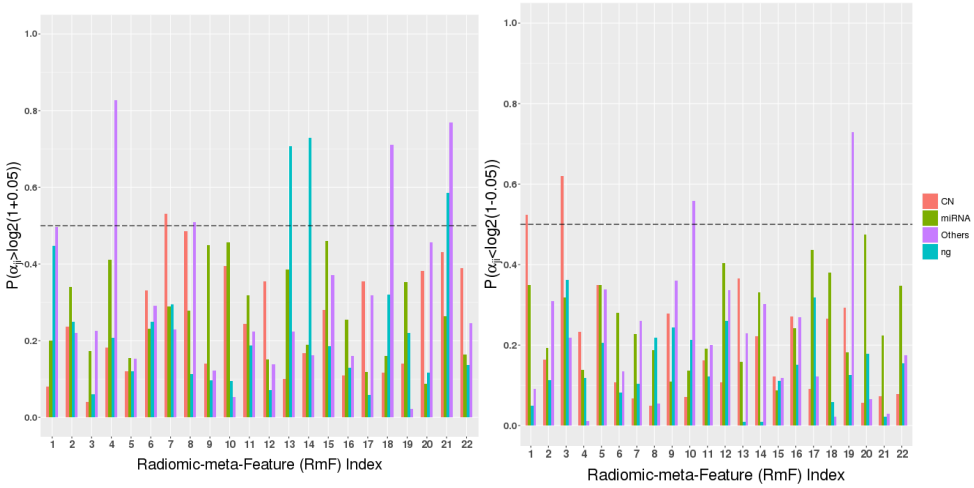
Fig S13: Results of Stage III when $d = 0.5$



(a) positively significant RmF combinations

(b) negatively significant RmF combinations

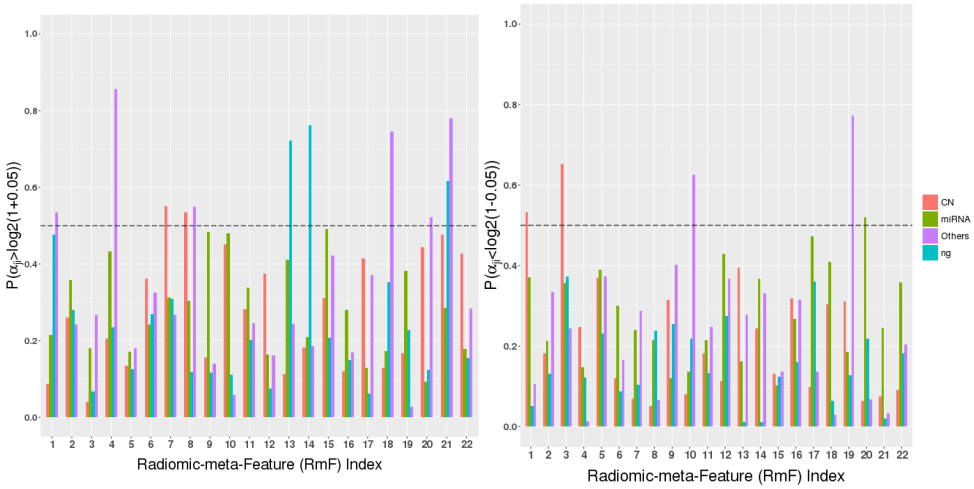
Fig S14: Results of Stage III when $d = 2$



(a) positively significant RmF combinations

(b) negatively significant RmF combinations

Fig S15: Results of Stage III when $\tilde{e} = 1$



(a) positively significant RmF combinations

(b) negatively significant RmF combinations

Fig S16: Results of Stage III when $\tilde{e} = 4$

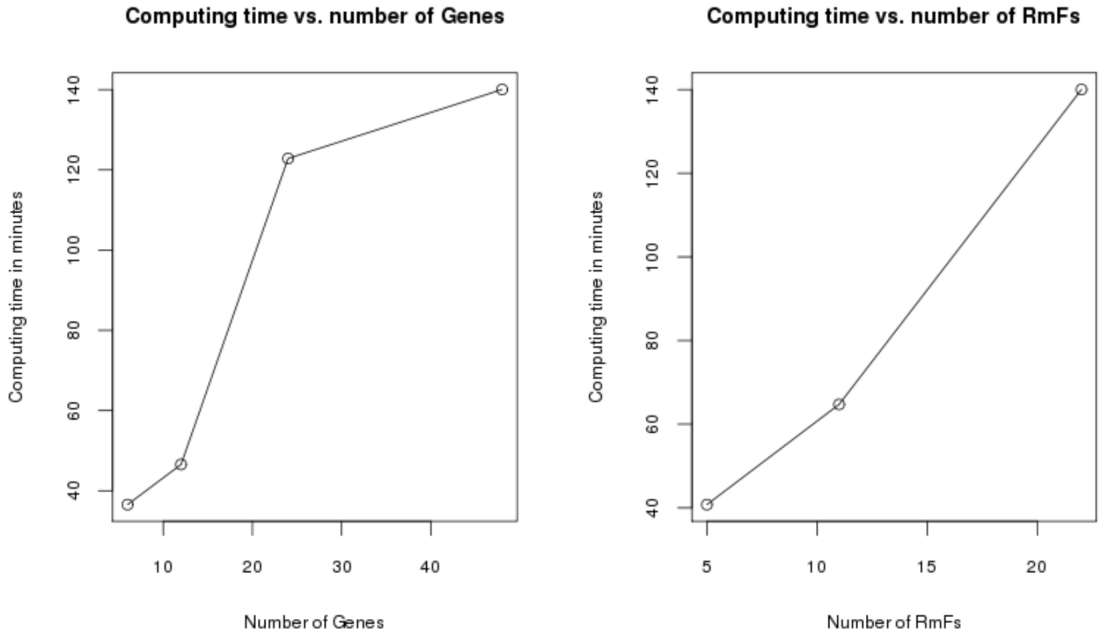


Fig S17: The computing time (in minutes) changes with the number of the genes and RmFs

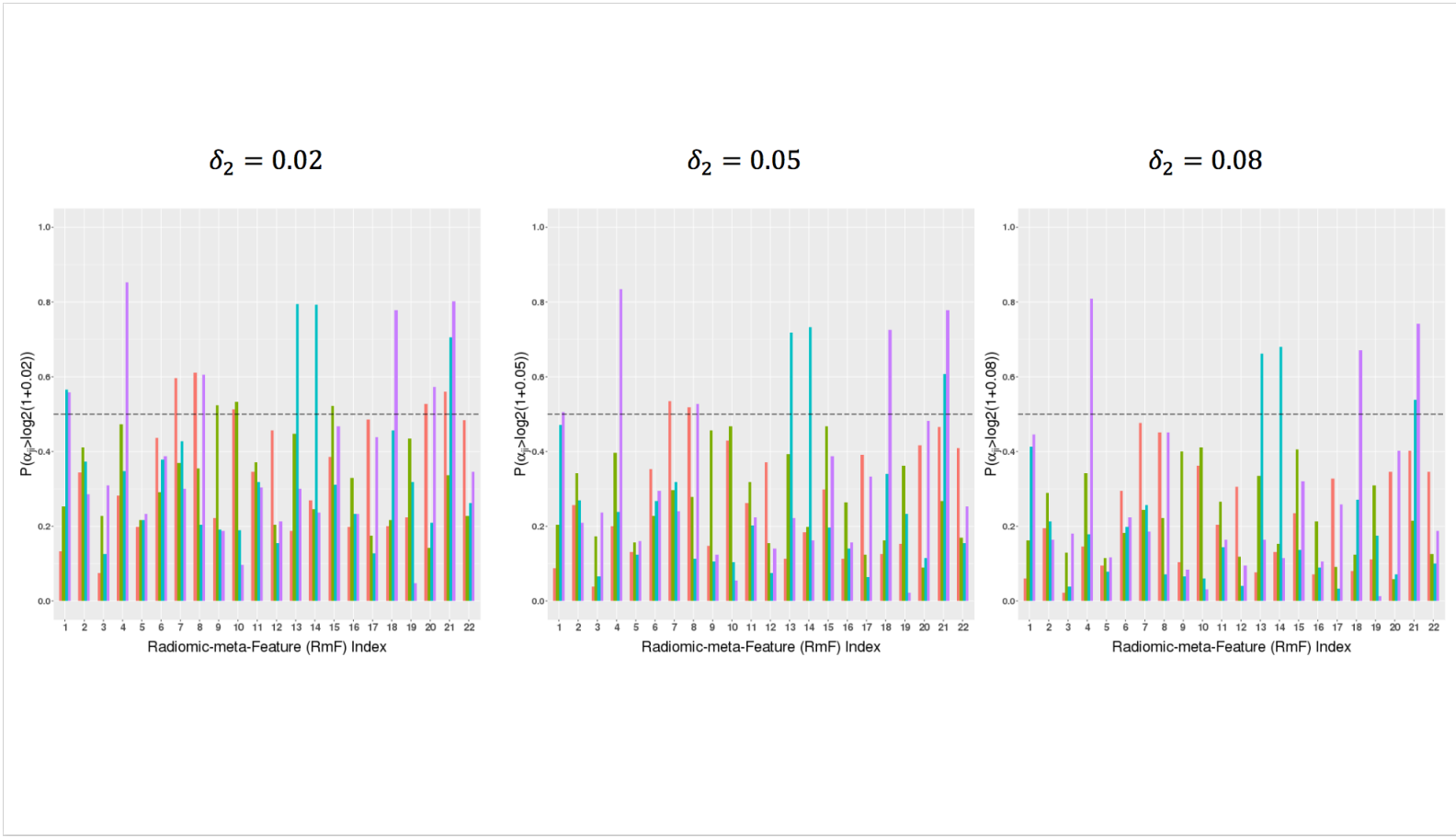


Fig S18: Stage III results: positively significant RmFs changing with the threshold

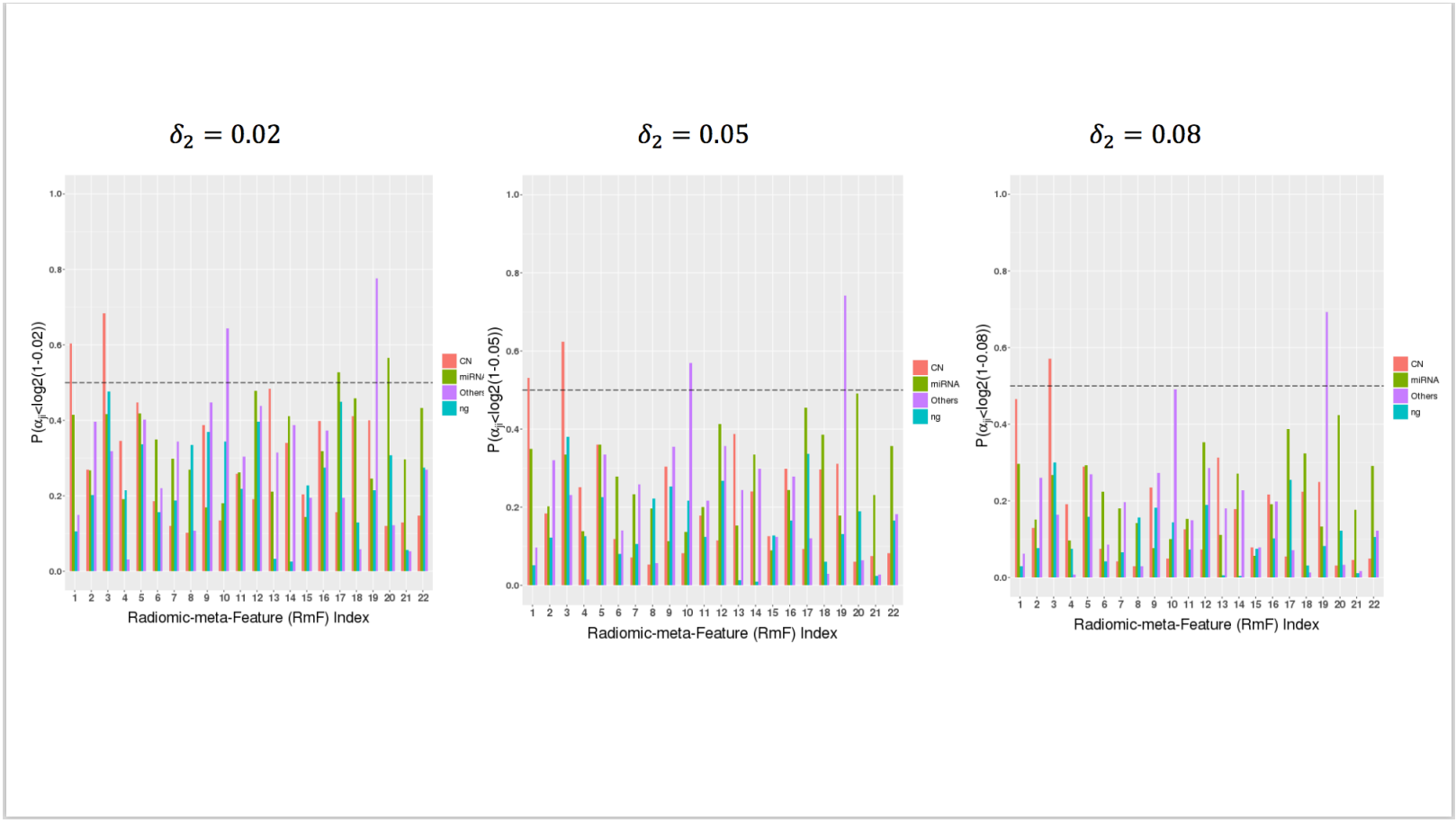


Fig S19: Stage III results: negatively significant RmFs changing with the threshold

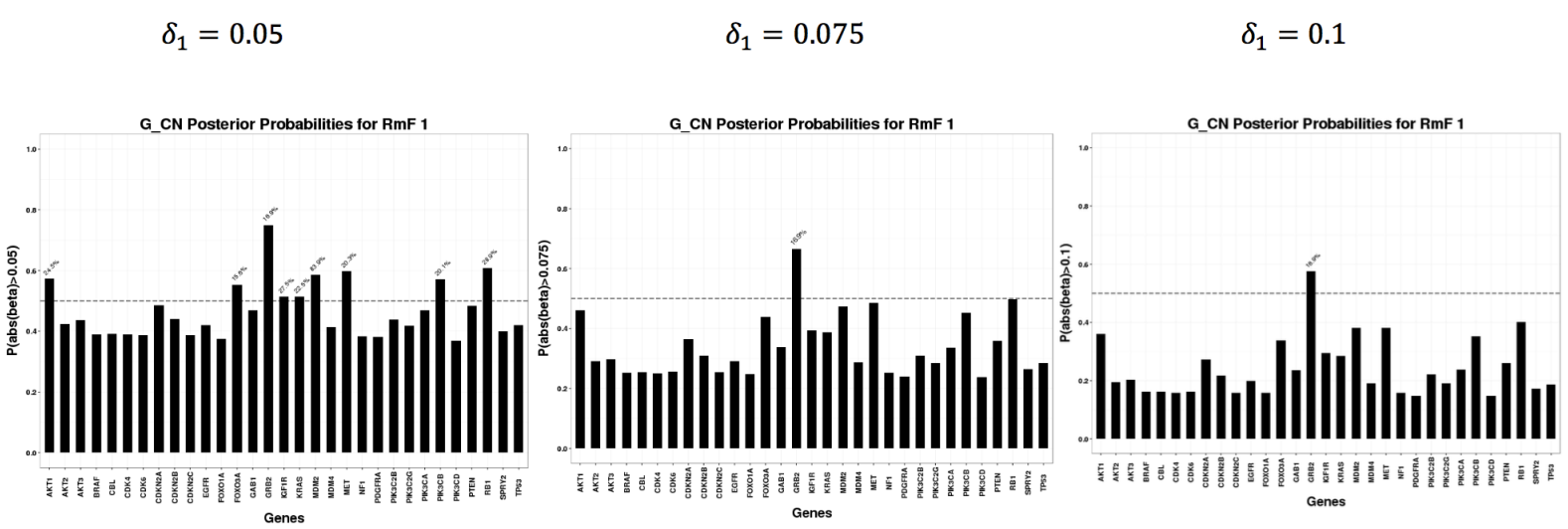


Fig S20: Stage II results changing with threshold

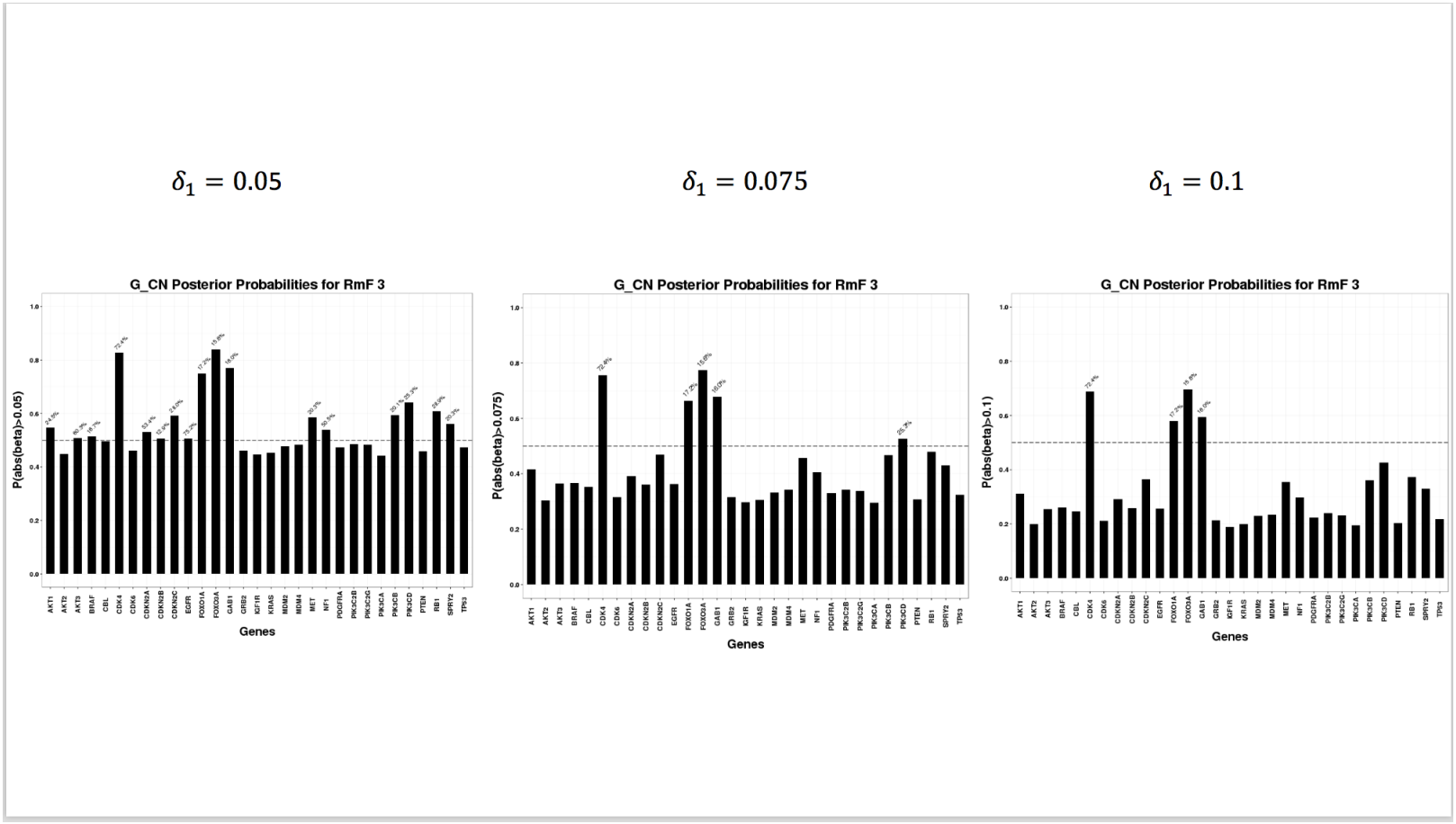


Fig S21: Stage II results changing with threshold

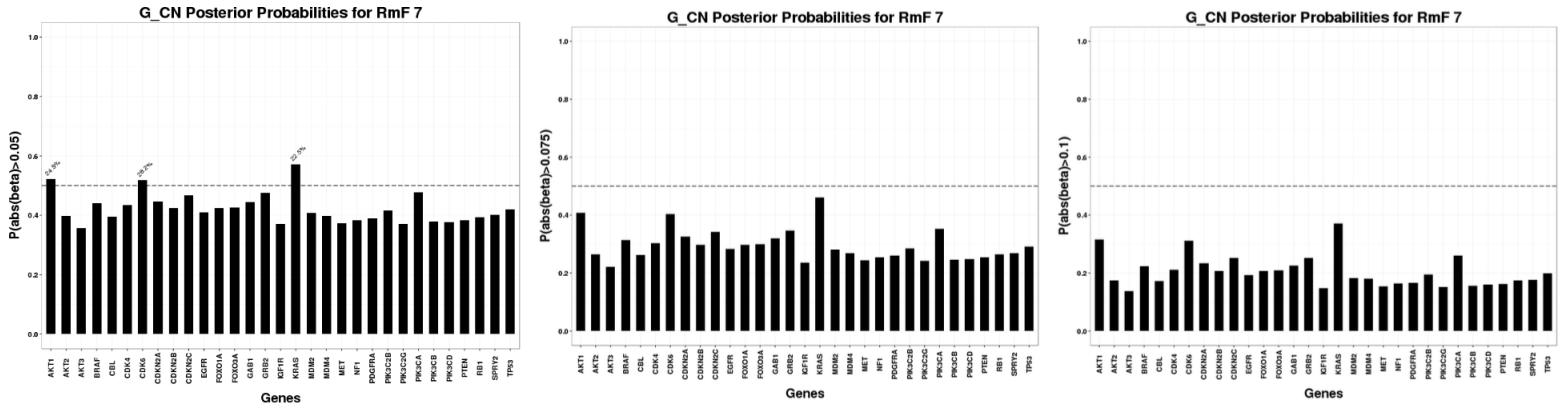


Fig S22: Stage II results changing with threshold

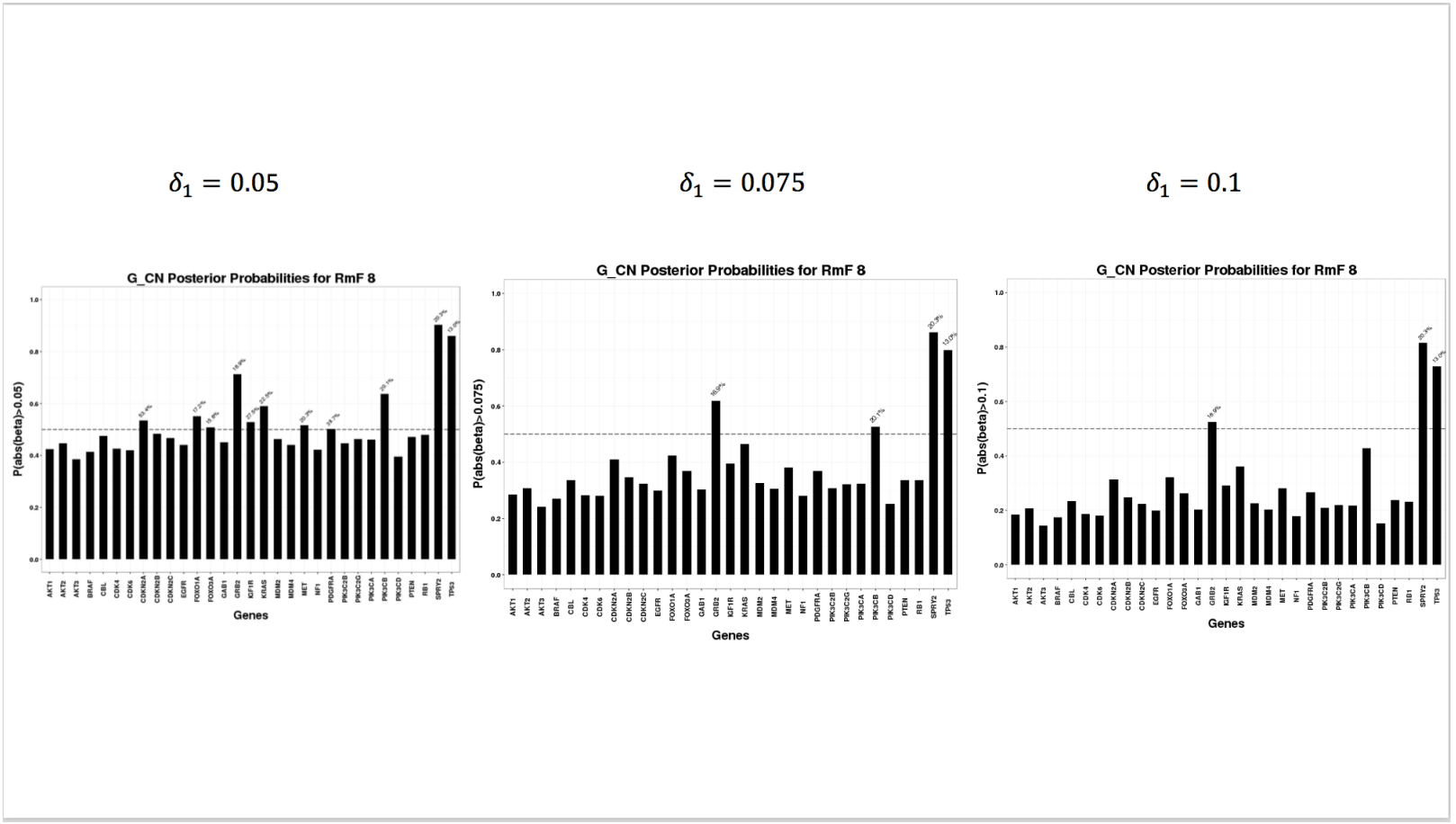


Fig S23: Stage II results changing with threshold

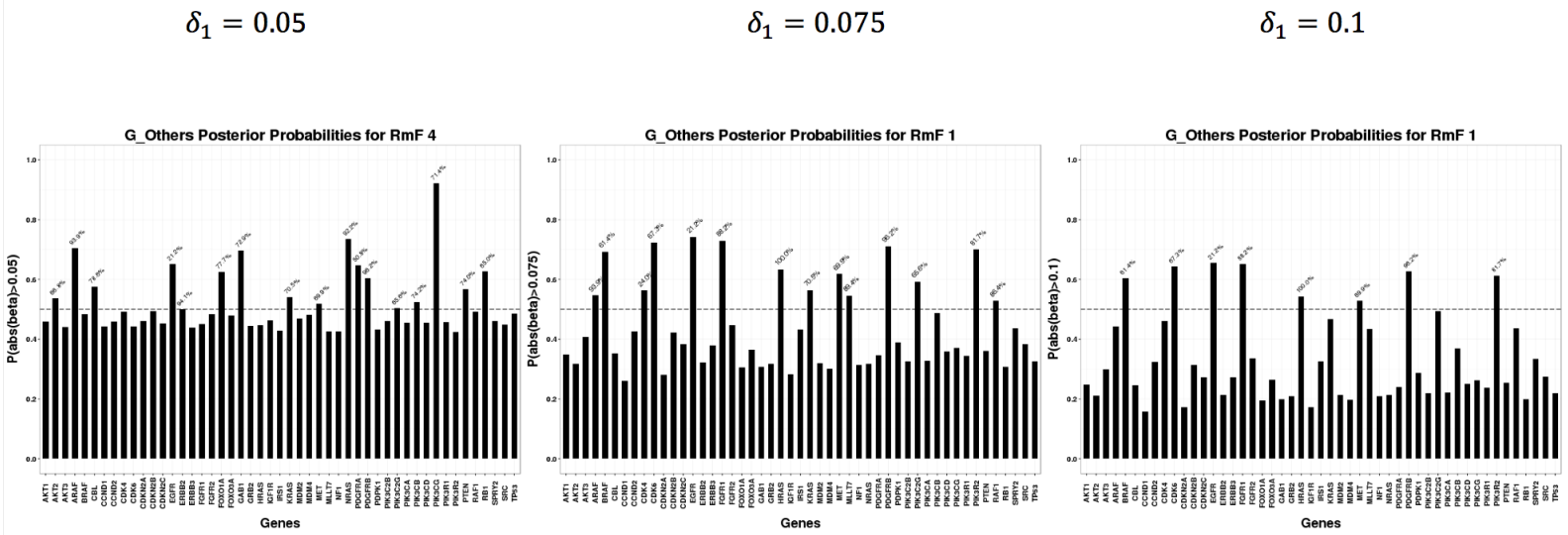


Fig S24: Stage II results changing with threshold

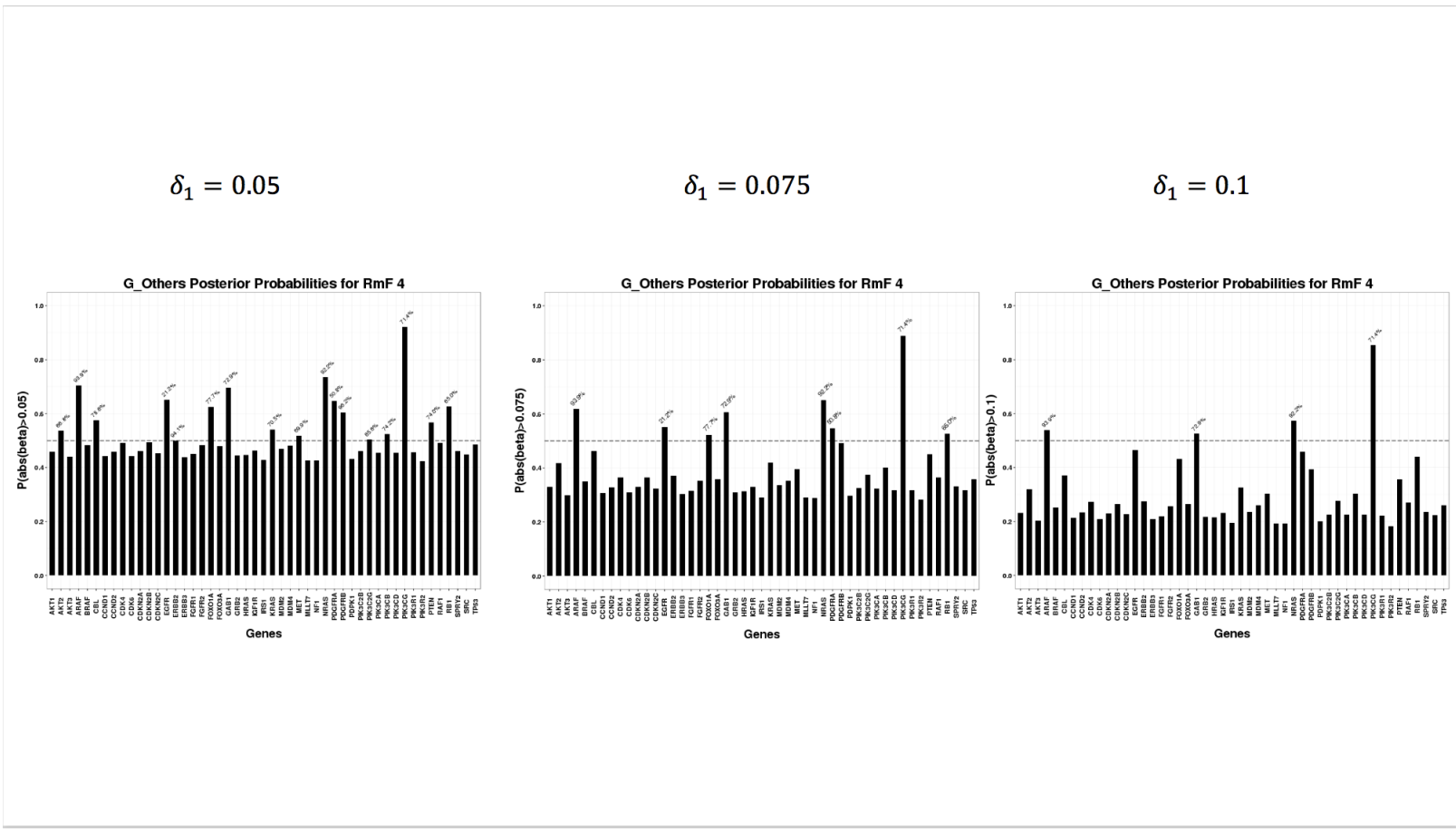


Fig S25: Stage II results changing with threshold

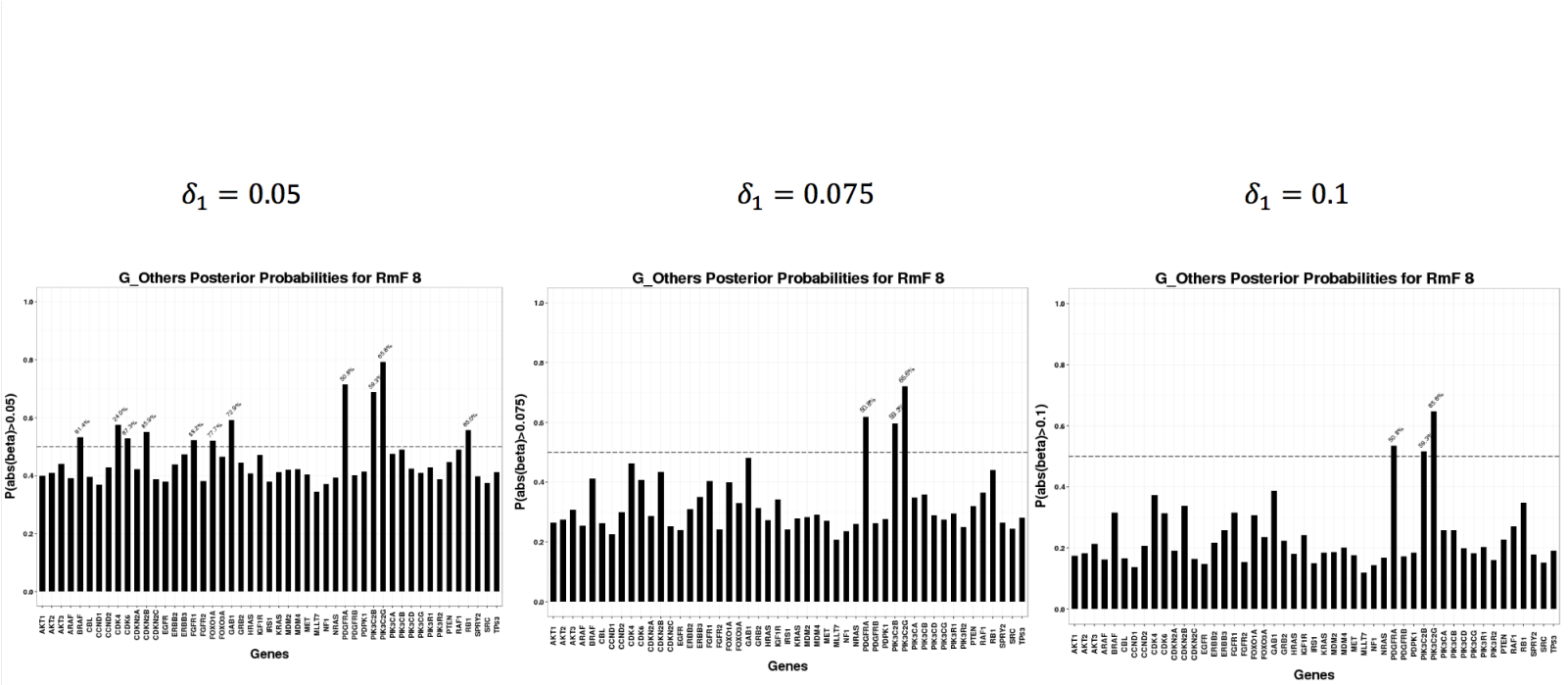


Fig S26: Stage II results changing with threshold

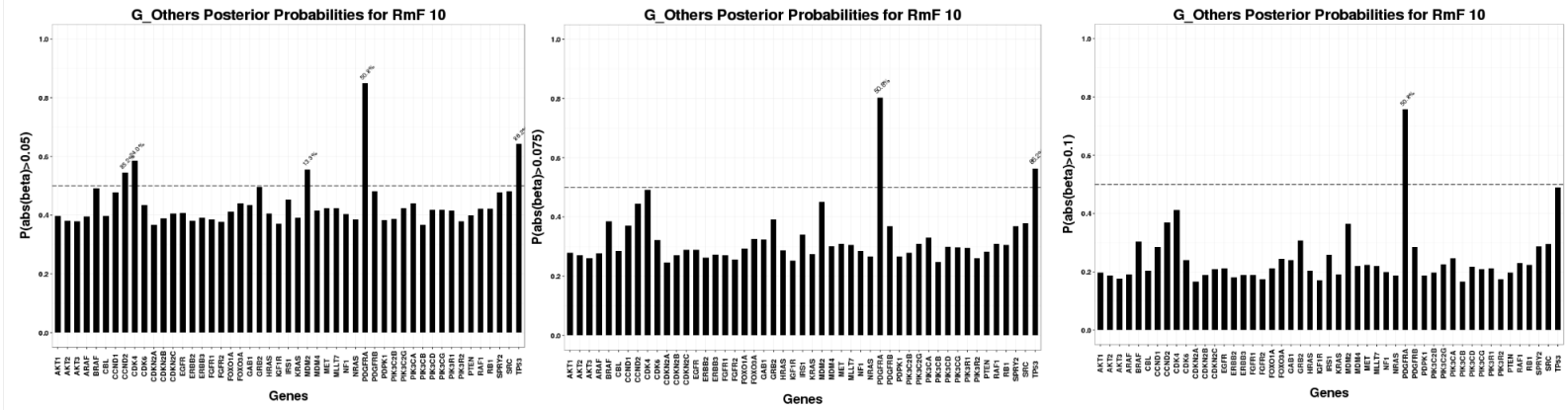


Fig S27: Stage II results changing with threshold

$\delta_1 = 0.05$ $\delta_1 = 0.075$ $\delta_1 = 0.1$

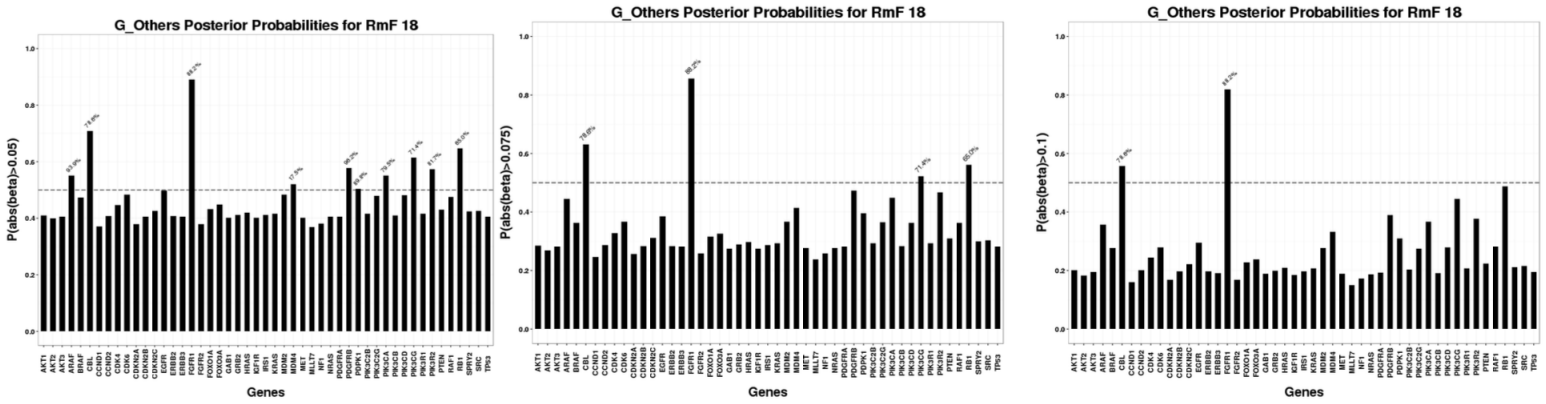


Fig S28: Stage II results changing with threshold

$\delta_1 = 0.05$ $\delta_1 = 0.075$ $\delta_1 = 0.1$

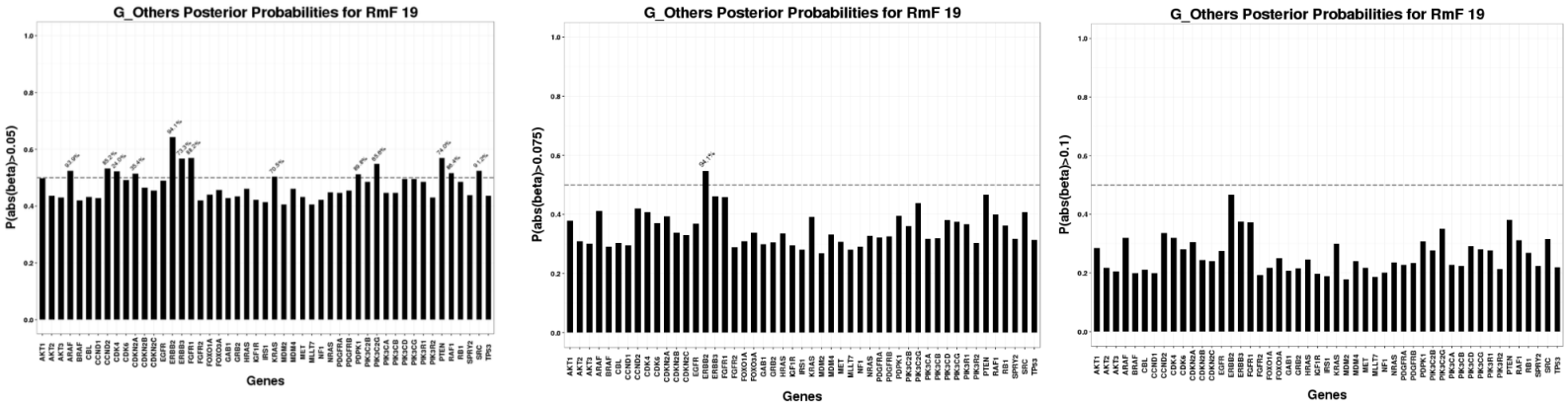


Fig S29: Stage II results changing with threshold

$\delta_1 = 0.05$ $\delta_1 = 0.075$ $\delta_1 = 0.1$

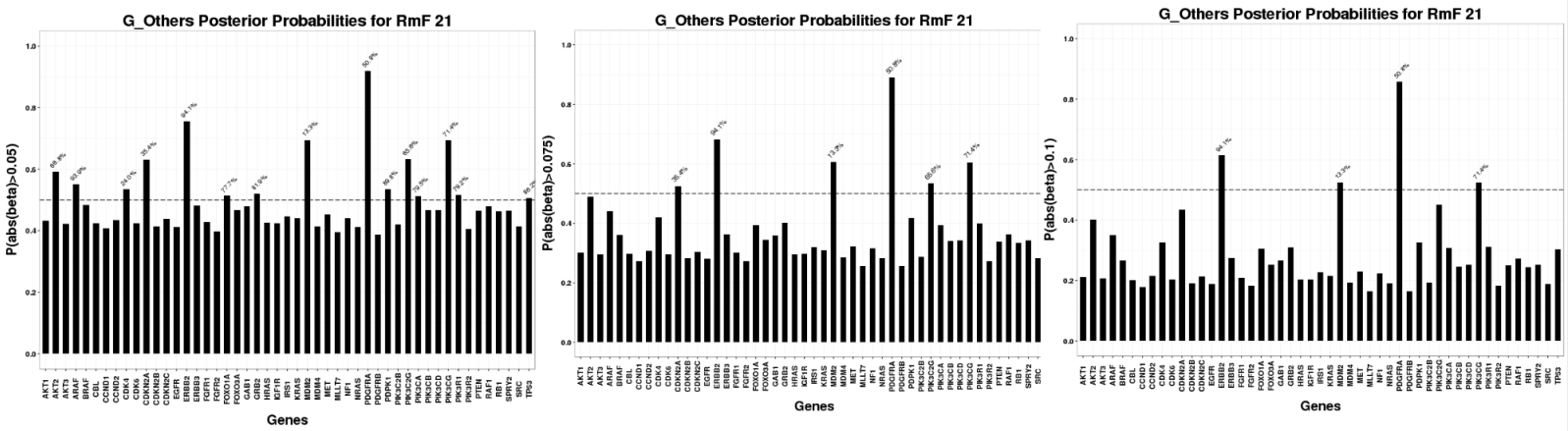


Fig S30: Stage II results changing with threshold

References

Ganeshan, B., Miles, K. A., Young, R. C., Chatwin, C. R., Gurling, H. M. and Critchley, H. D. (2010). Three-dimensional textural analysis of brain images reveals distributed grey-matter abnormalities in schizophrenia. *European radiology* **20** 941–948.

Geweke, J. et al. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments* **196**. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA.

Griffin, J. E., Brown, P. J. et al. (2010). Inference with normal-gamma prior distributions in regression problems. *Bayesian Analysis* **5** 171–188.

Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE* **67** 786–804.

Haralick, R. M. (1992). LG shapiro. Computer and Robot Vision, Vol. 1.

McAuliffe, M. J., Lalonde, F. M., McGarry, D., Gandler, W., Csaky, K. and Trus, B. L. (2001). Medical image processing, analysis and visualization in clinical research. In *Computer-Based Medical Systems, 2001. CBMS 2001. Proceedings. 14th IEEE Symposium on* 381–386. IEEE.

Park, T. and Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association* **103** 681–686.

Zhou, M., Hall, L., Goldgof, D., Russo, R., Balagurunathan, Y., Gillies, R. and Gatenby, R. (2014). Radiologically defined ecological dynamics and clinical outcomes in glioblastoma multiforme: preliminary results. *Translational oncology* **7** 5–13.

Zou, H., Hastie, T. and Tibshirani, R. (2006). Sparse principal component analysis. *Journal of computational and graphical statistics* **15** 265–286.

Youyi Zhang
Jeffrey S. Morris
Department of Biostatistics
The University of Texas MD Anderson Cancer Centers
1400 Pressler St,
Houston, TX77030
E-mail: YZhang39@mdanderson.org
E-mail: jefmorris@mdanderson.org

Shivali Narang Aerry
Department of Engineering for Professionals
Whiting School of Engineering
Johns Hopkins University
3400 N Charles St,
Baltimore, MD 21218
E-mail: shivali.narang13@gmail.com

Arvind U.K. Rao
Department of Computational Medicine & Bioinformatics
University of Michigan Ann Arbor
100 Washtenaw Avenue,
Ann Arbor, MI 48109-2218
E-mail: ukarvind@umich.edu

Veerabhadran Baladandayuthapani
Department of Biostatistics
University of Michigan Ann Arbor
100 Washtenaw Avenue,
Ann Arbor, MI 48109-2218
E-mail: veerab@umich.edu