

Comparison respect to NOKS, conditional on treatment

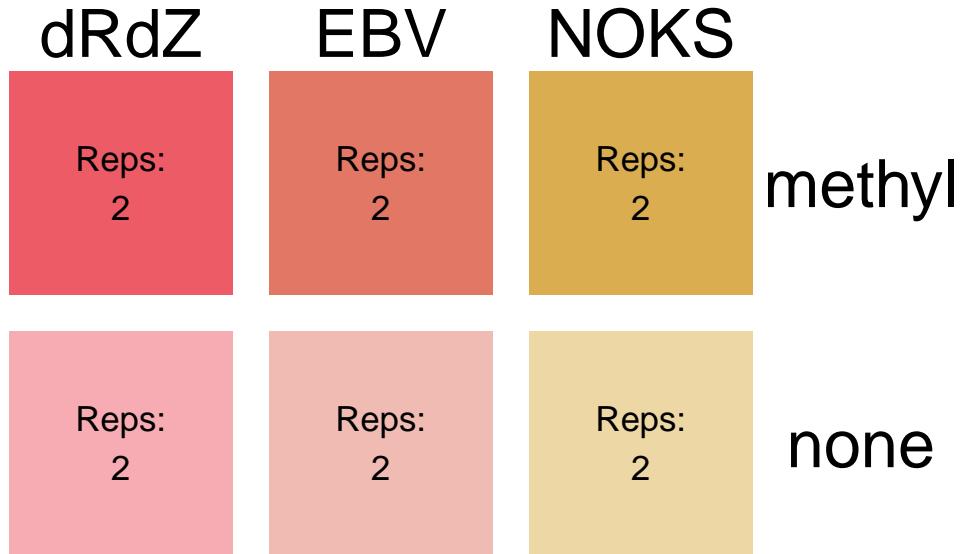
Contents

Intro	1
Differential expression analysis	1
Rank analysis	2
Joint analysis	3
Pathway analysis	5
All genes (HALLMARK pathways)	5
All genes (curated pathways)	20
Differentially expressed genes (defined by padj ≤ 0.05 and HALLMARK pathways)	21

Intro

Mark made this analysis under a different approach. For the analysis, he compared the log2FC of the EBV and dRdZ expression normalized by NOKS. I think that approach would be equivalent to make interaction contrasts with a fixed treatment. Hence, we are going to:

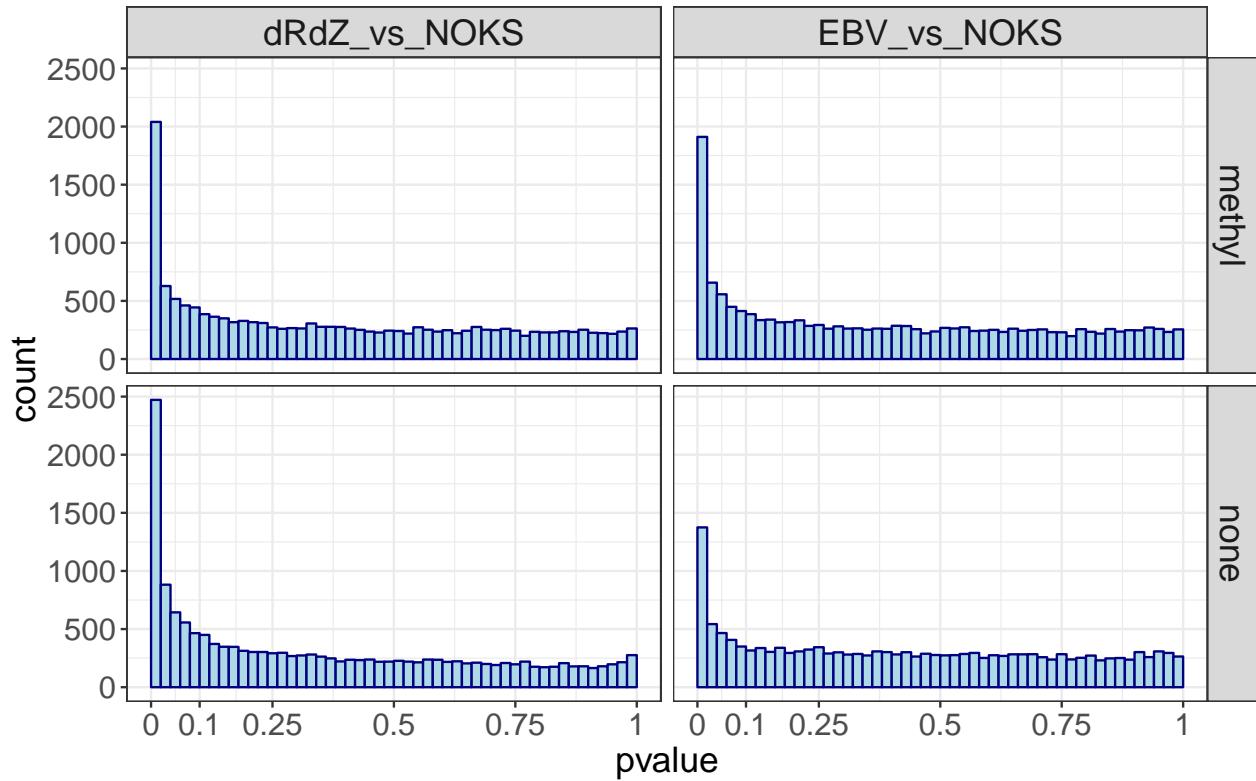
1. Perform contrasts of EBV and dRdZ against NOKS, with and without treatment.
2. Using all the genes, perform pathway analysis using GSEA
3. Compare the log2FC (calculated with the rlog matrices), of EBV and dRdZ normalized by NOKS.



Differential expression analysis

First, we used DESeq2 to estimate a signal-to-noise measure for the genes that are differentially expressed between EBV/dRdZ Vs. NOKS cells conditional on a treatment. It is expected that if the expression patterns are going to differ when conditioning by the methyl treatment or its absence. It is worth mentioning, that we removed from the sample the genes that have < 20 reads accumulated across all samples.

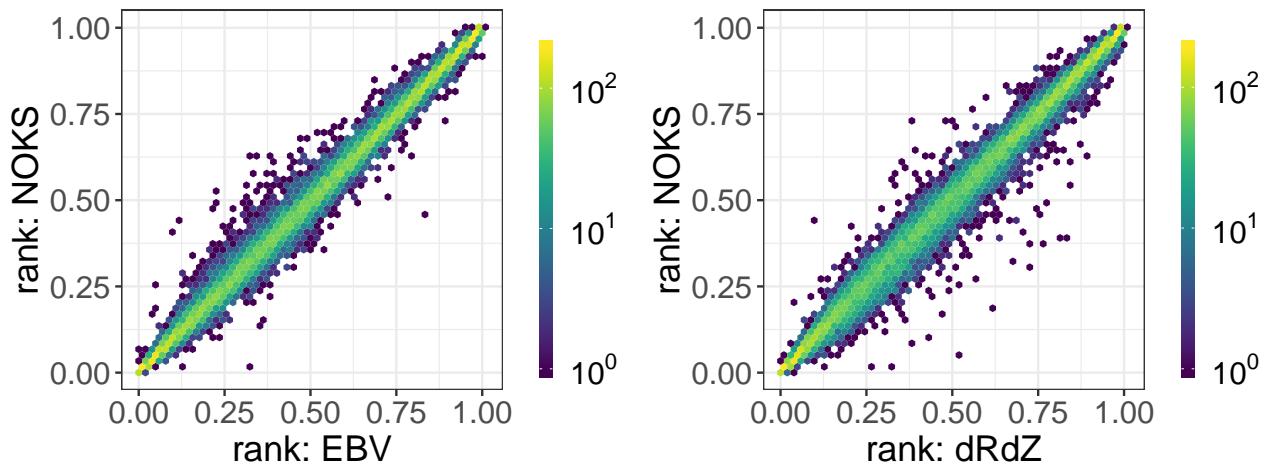
The figure below exhibits the p.value distribution for the 4 contrasts that we are testing. Even though, we are not going to use multiple test correction it is good to know that these tests satisfy the assumptions. I gave a quick look to Biostars and found two posts: 1 and 2, and it seems that there is not a pre-defined way when using a set of genes for the pathway analysis (either all genes or only the diff. expressed ones).



Rank analysis

With MC treatment

The figure below shows a decreasing ranking of the average rlog of the samples of a given cell and the `methyl` treatment.

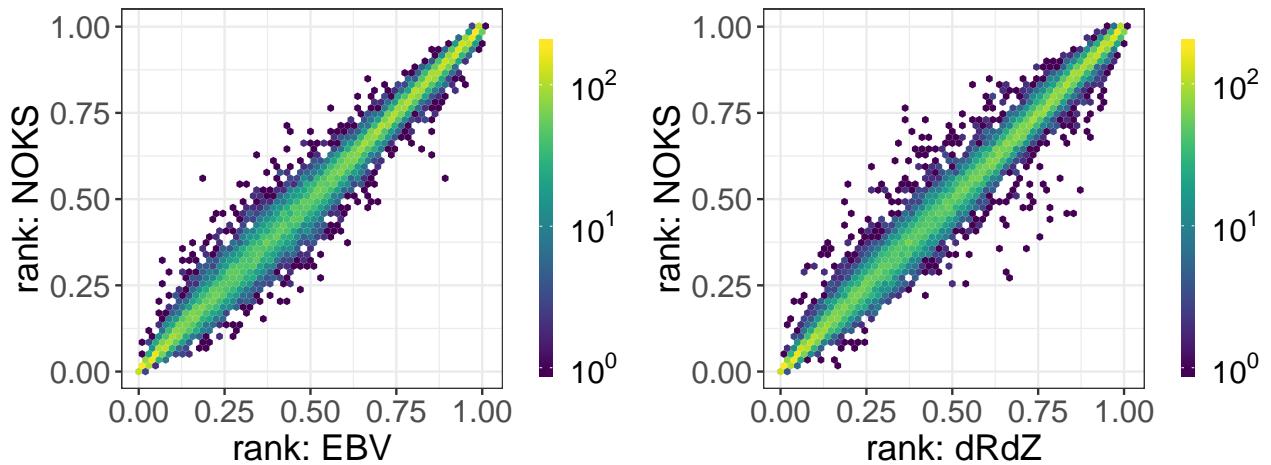


The closer the point is to the zero, is mean that the gene in question is more expressed. It appears that the

yellow/green diagonal band is a bit wider for the figure in the right. This would mean, that potentially there are more genes being differentially expressed in the dRdZ contrast than in the EBV one. To review if this was true, for both contrast we count the amount of genes that are differentially expressed in a given contrast:

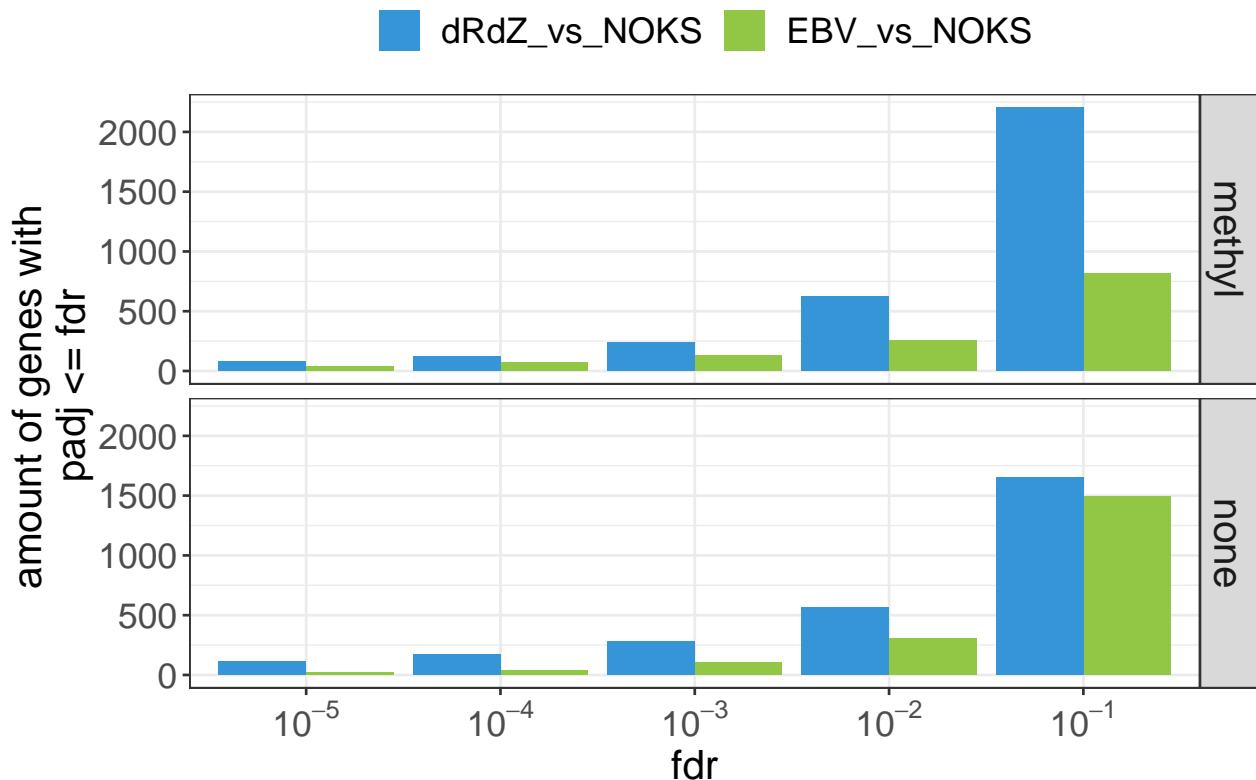
Without treatment

This analysis is analogous to the previous one. The only difference is that in this one, we are conditioning respect to the **none** treatment instead of the **methyl** one.

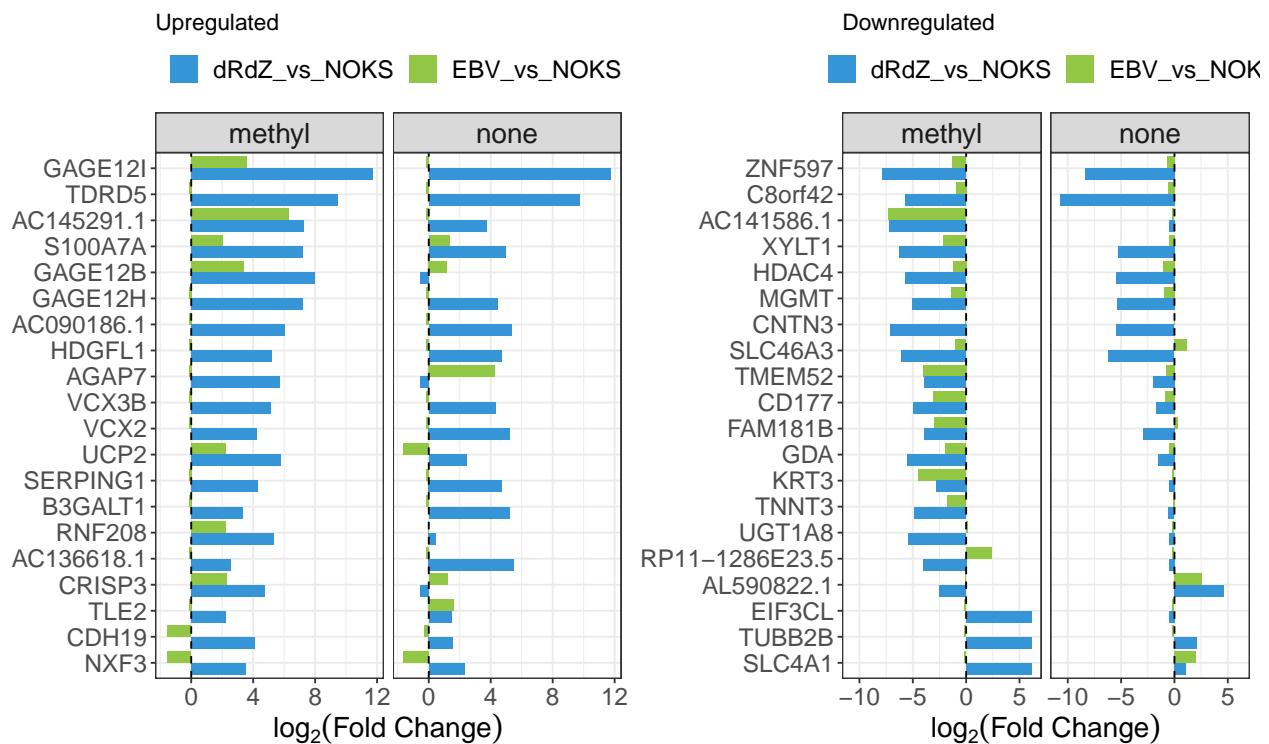


Joint analysis

The figure below shows that when counting the number of genes that are differentially expressed, the increment between the dRdZ and EBV contrasts is higher, when the MC treatment is applied.

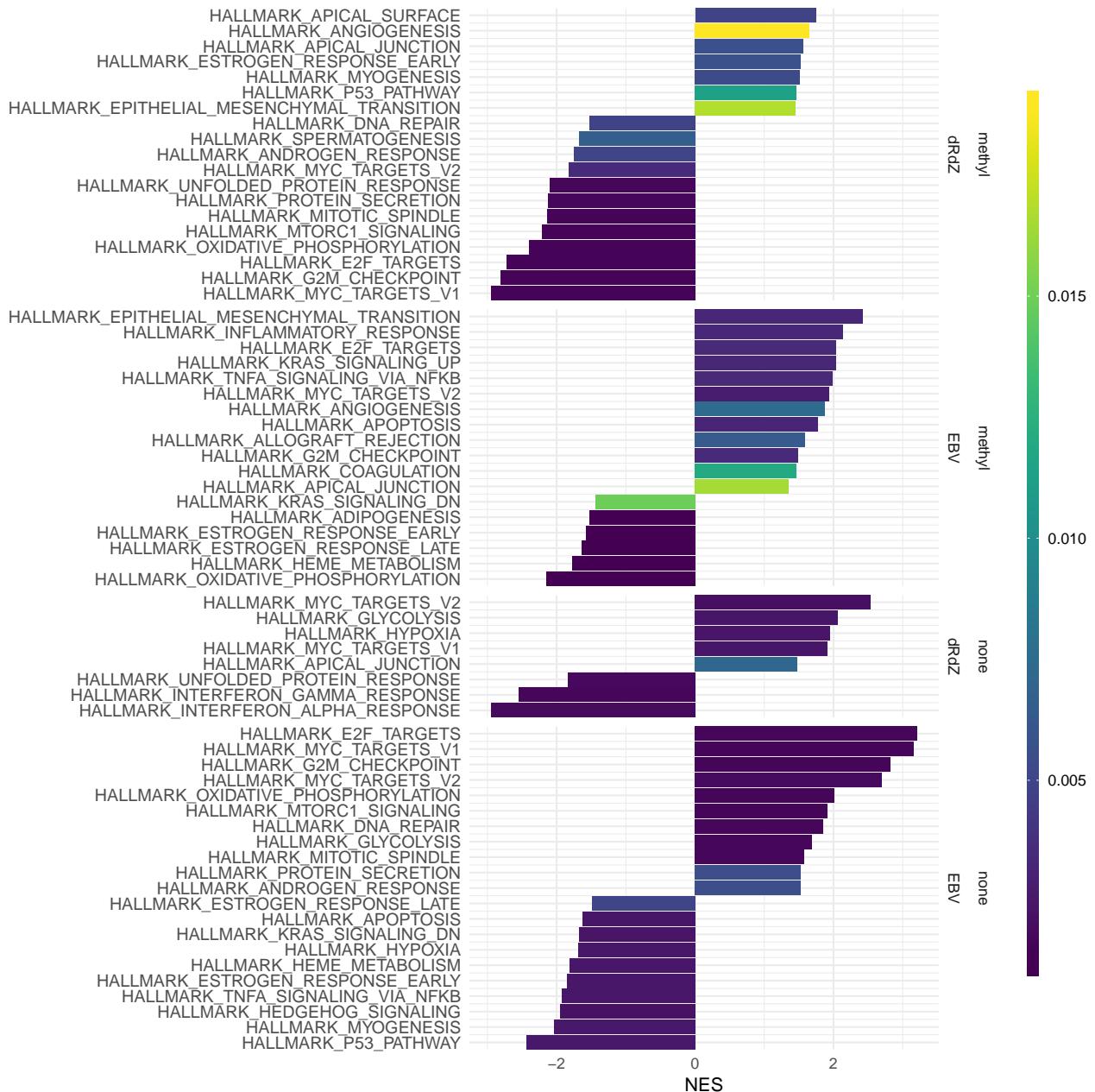


Since more gene are expressed in the mutant type EBV-dRdZ than in the EBV cell lines when a treatment is applied, then we examine the degree to which the most upregulated (downregulated) genes are expressed by comparing the top (bottom) 20 genes with the highest (lowest) log₂ fold change (dRdZ/EBV divided by NOKS):

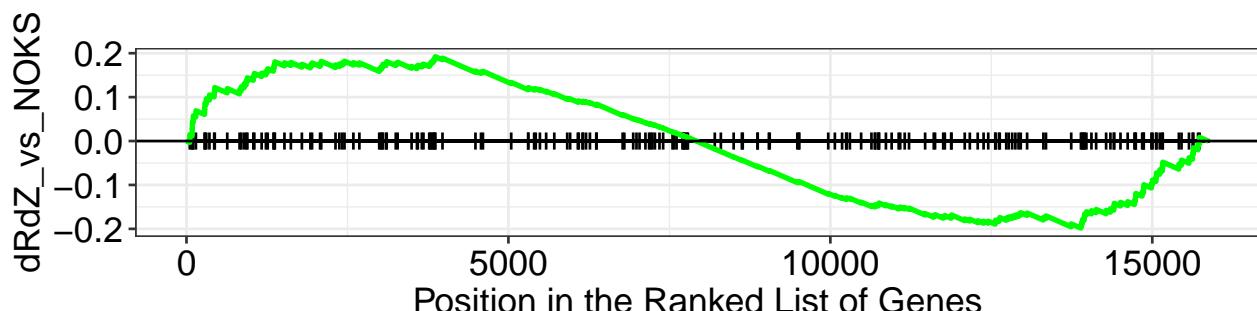
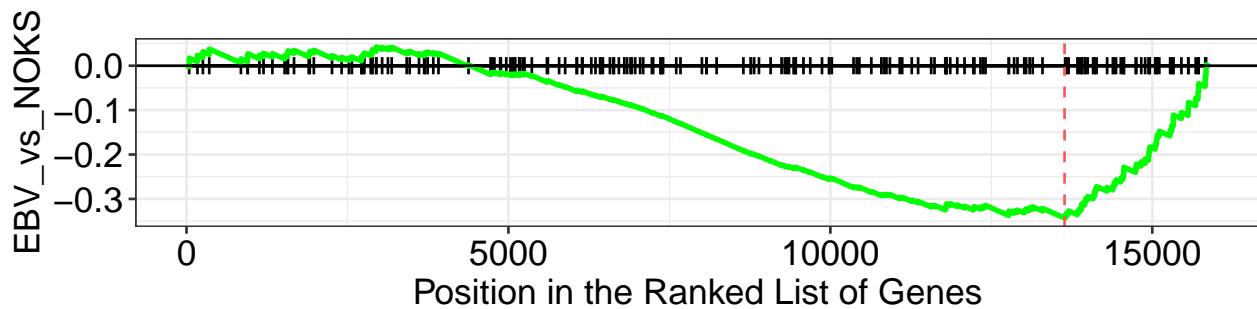


Pathway analysis

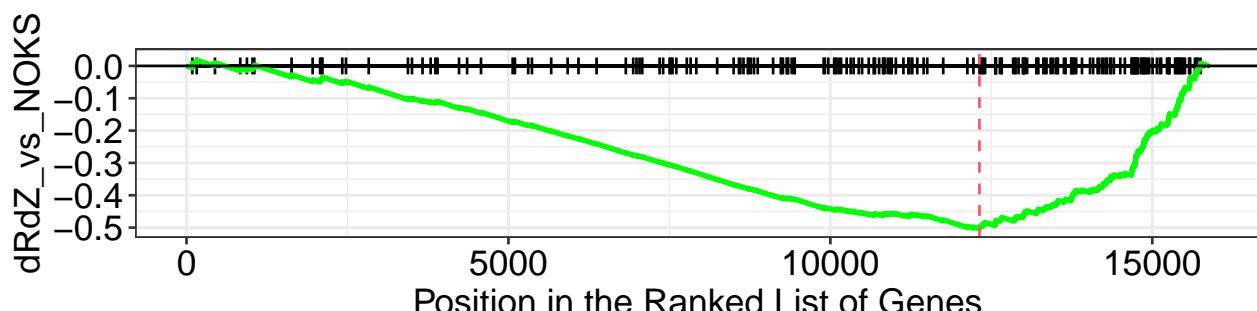
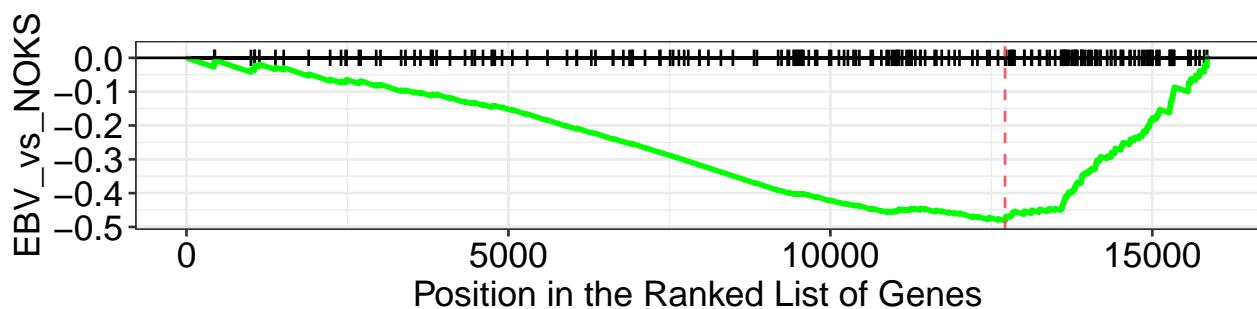
All genes (HALLMARK pathways)



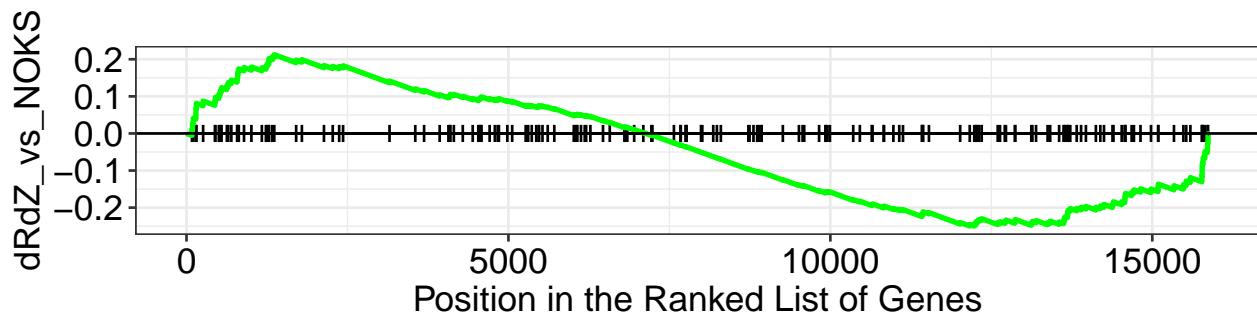
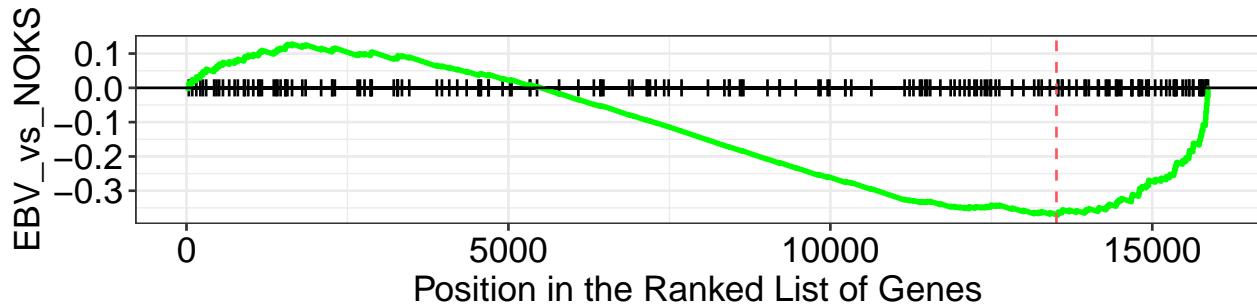
HALLMARKADIPOGENESIS



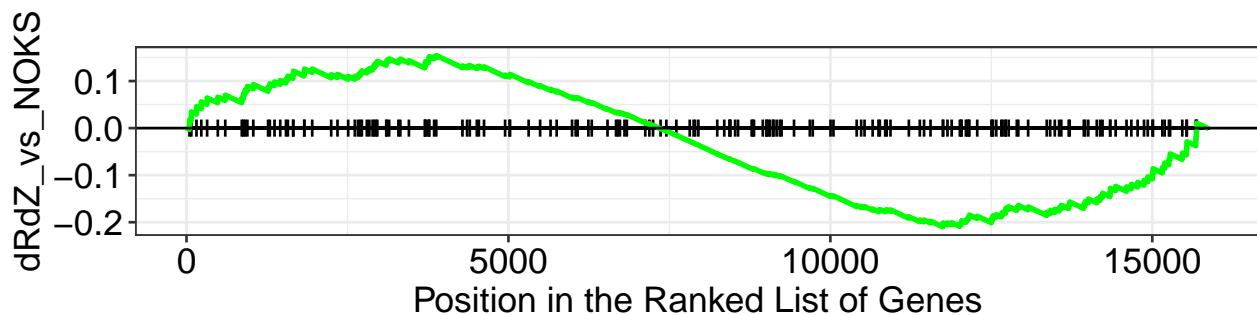
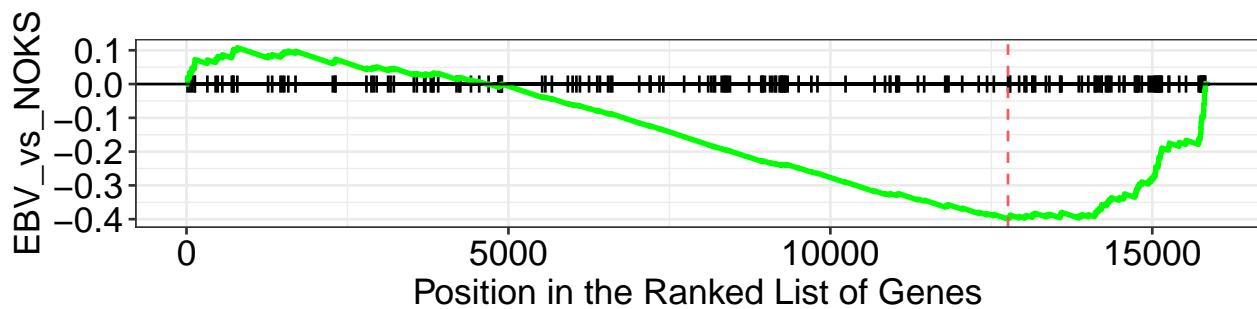
HALLMARKOXIDATIVE_PHOSPHORYLATION



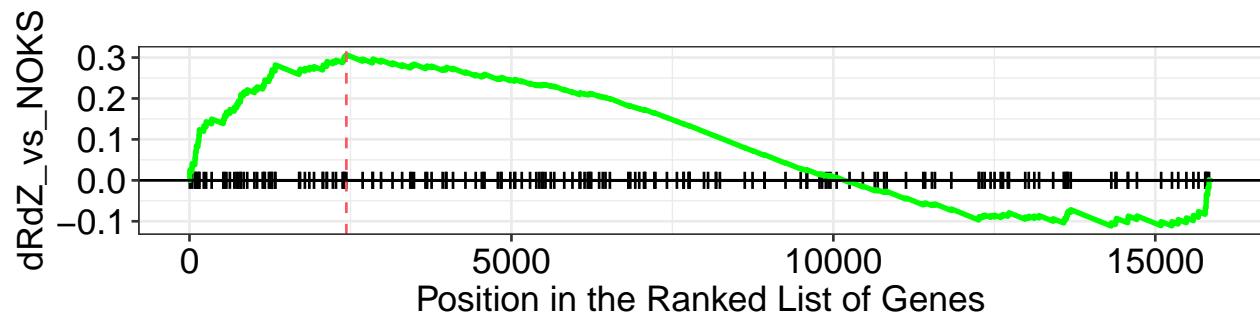
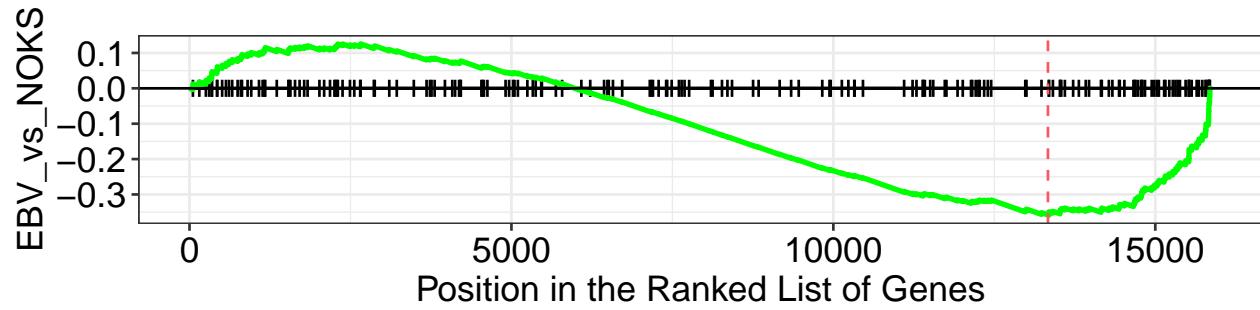
HALLMARK_ESTROGEN_RESPONSE_LATE



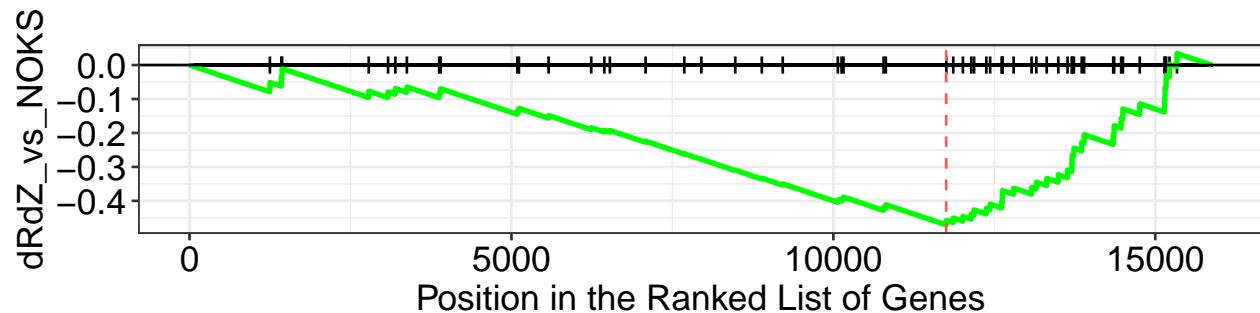
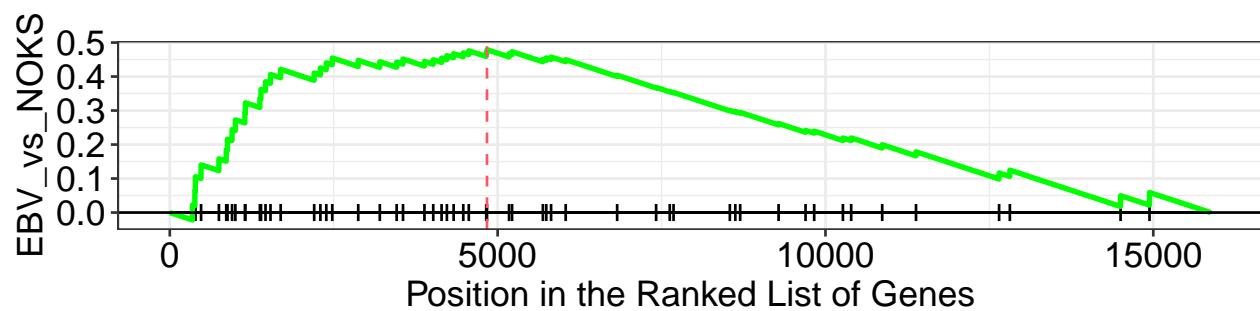
HALLMARK_HEME_METABOLISM



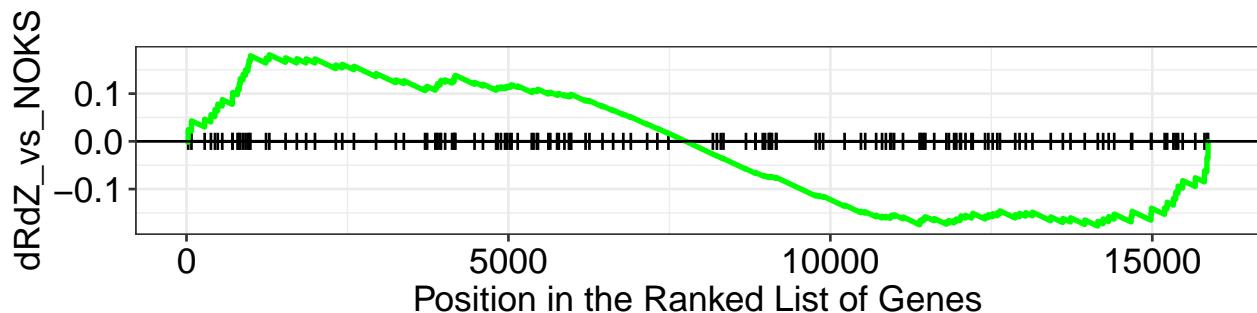
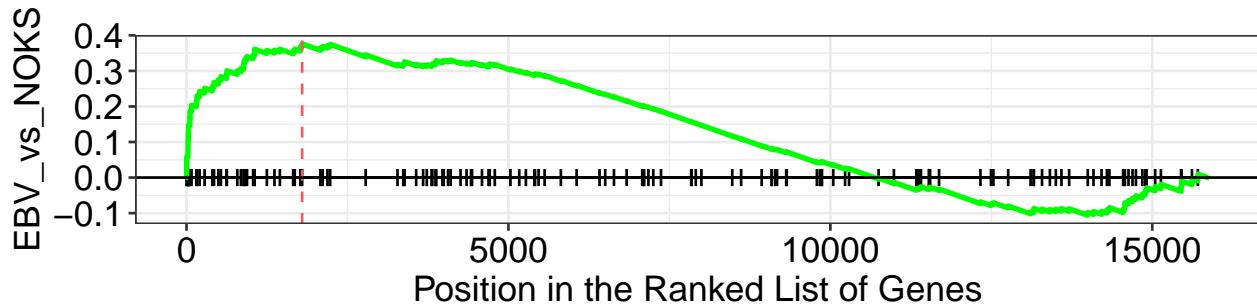
HALLMARK_ESTROGEN_RESPONSE_EARLY



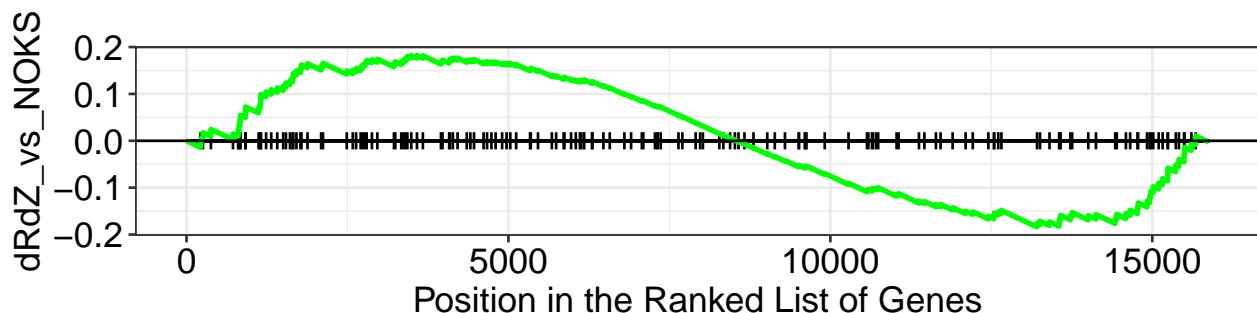
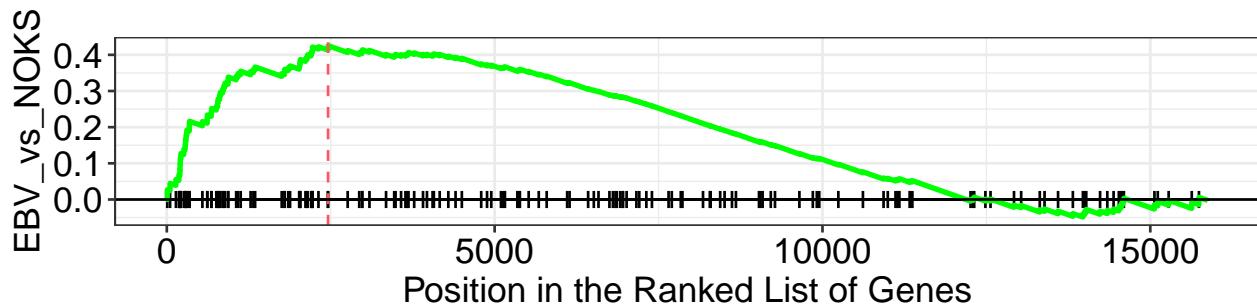
HALLMARK_MYC_TARGETS_V2



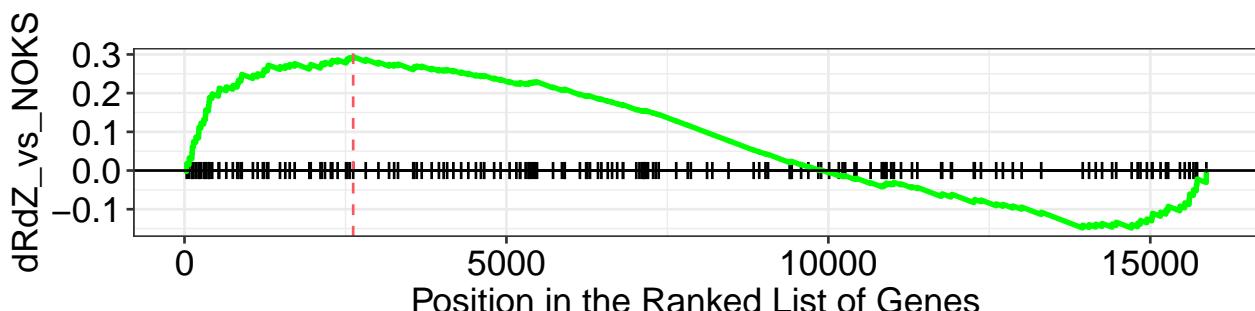
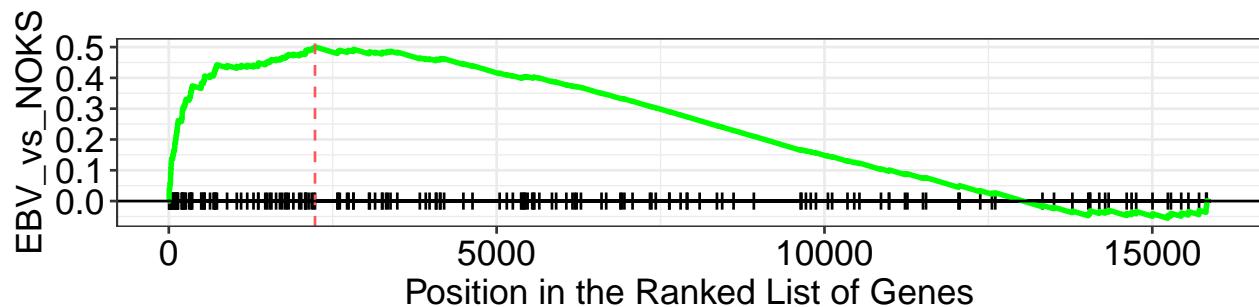
HALLMARK_APOPTOSIS



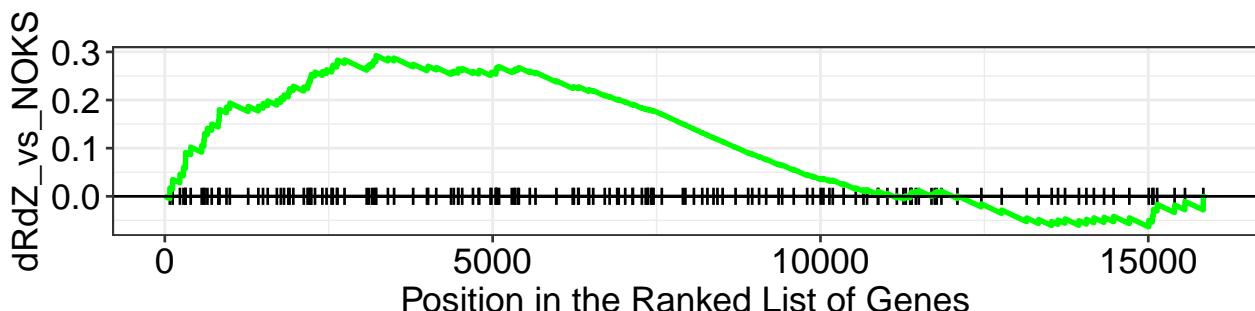
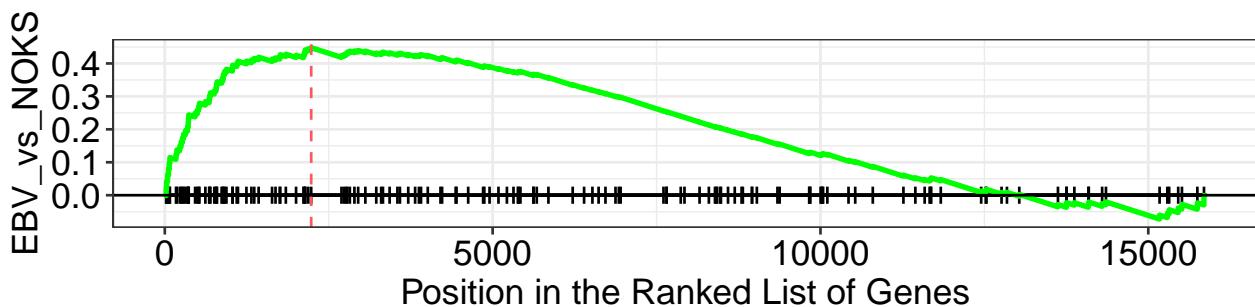
HALLMARK_KRAS_SIGNALING_UP



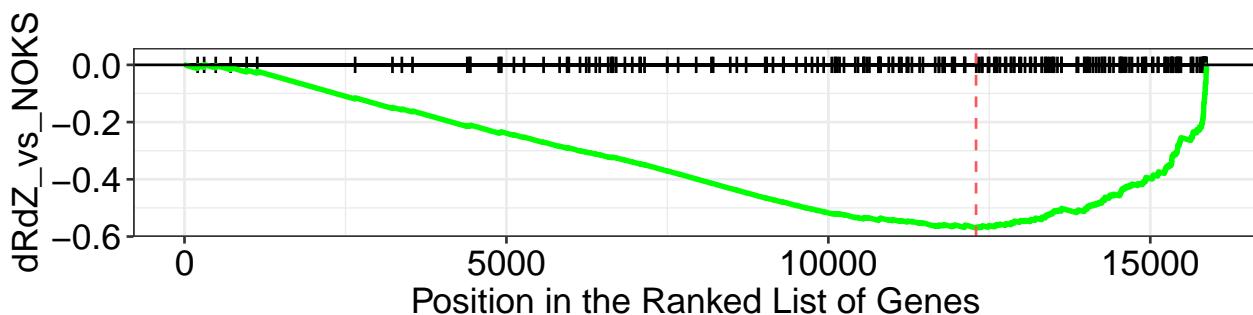
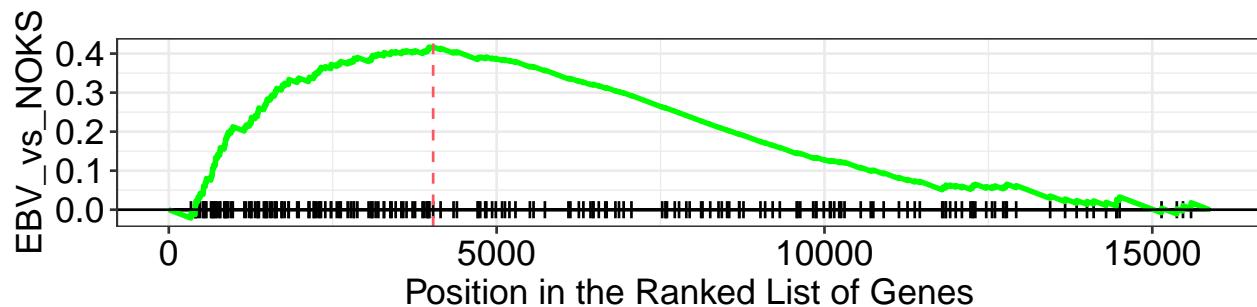
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION



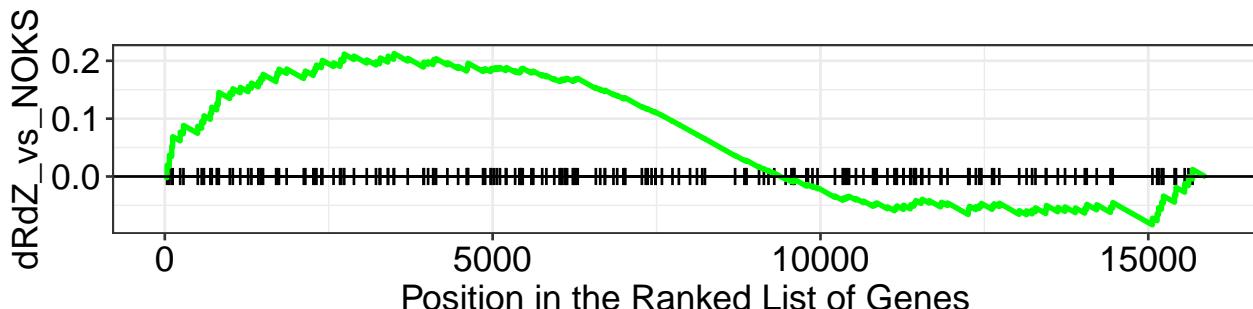
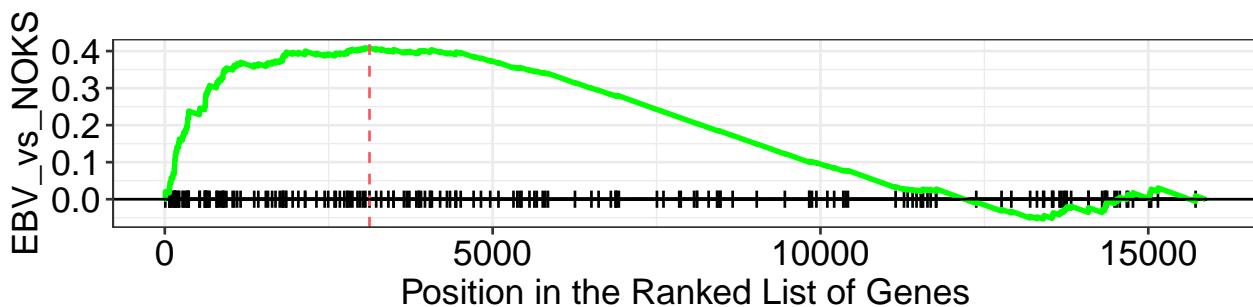
HALLMARK_INFLAMMATORY_RESPONSE



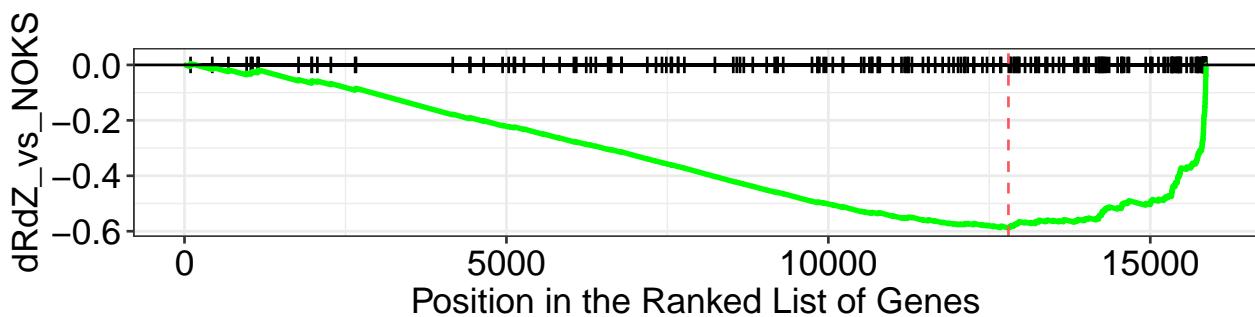
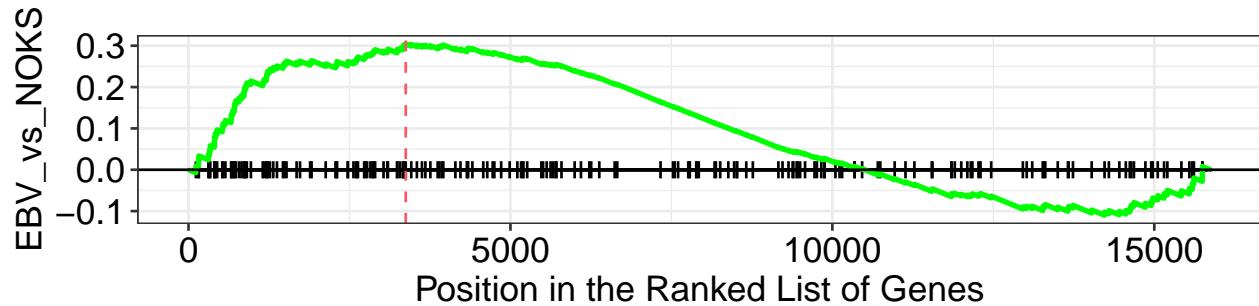
HALLMARK_E2F_TARGETS



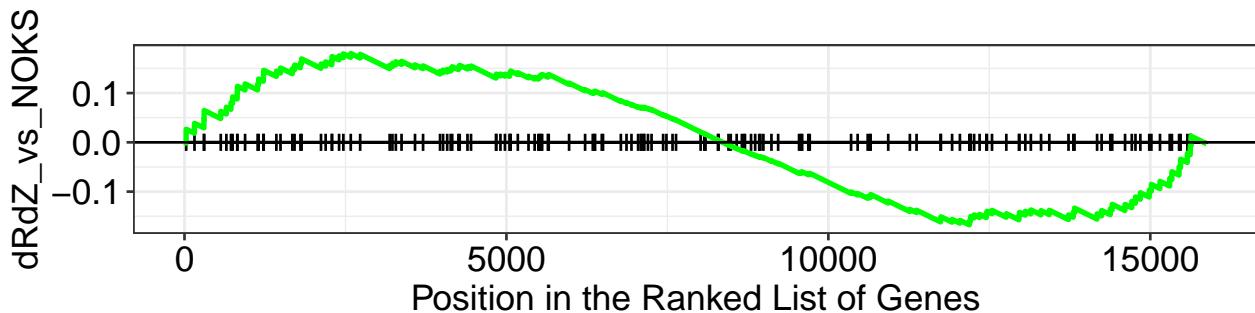
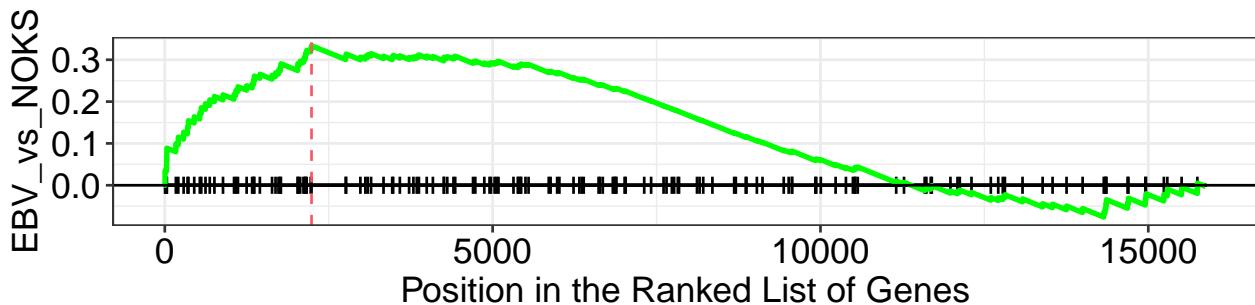
HALLMARK_TNFA_SIGNALING_VIA_NFKB



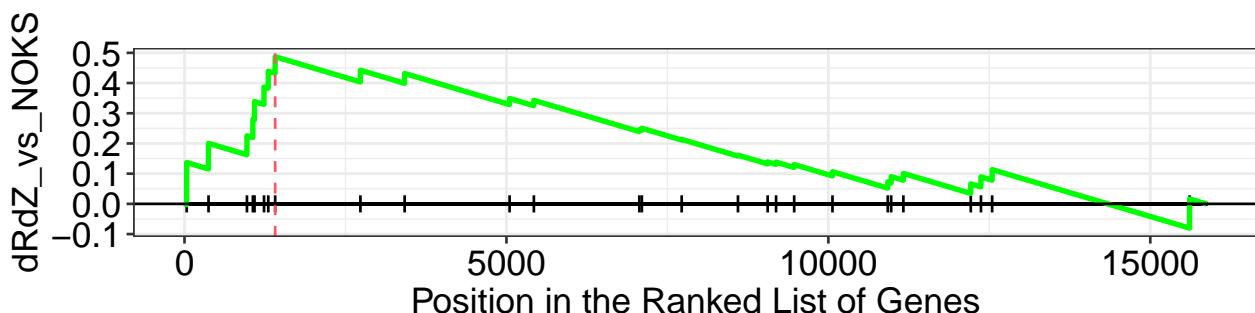
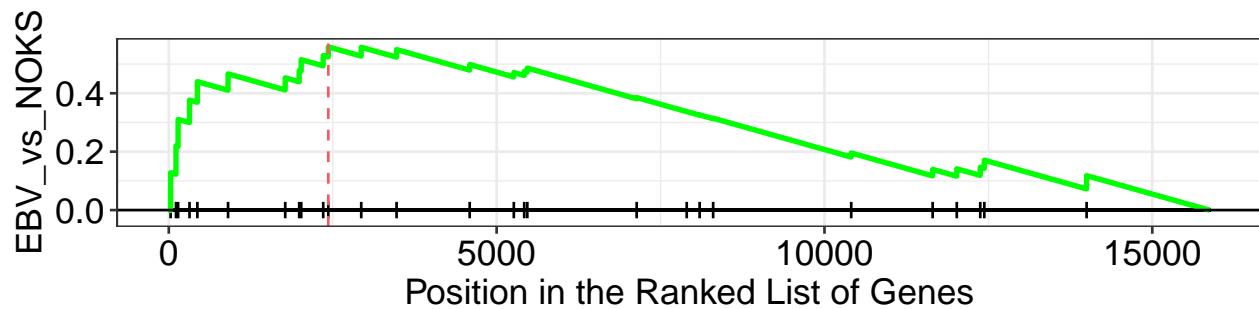
HALLMARK_G2M_CHECKPOINT



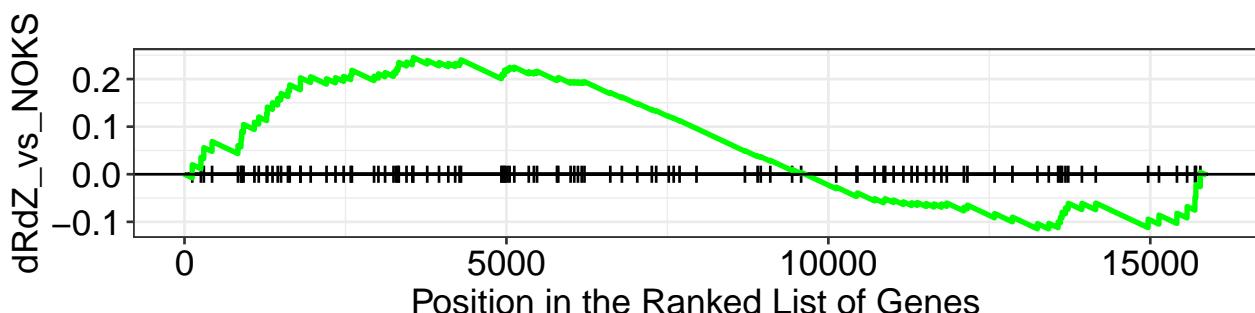
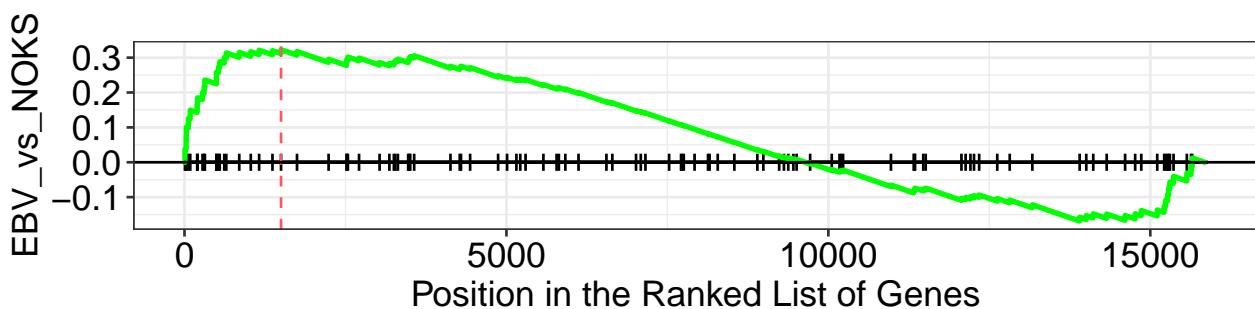
HALLMARK_ALLOGRAFT_REJECTION



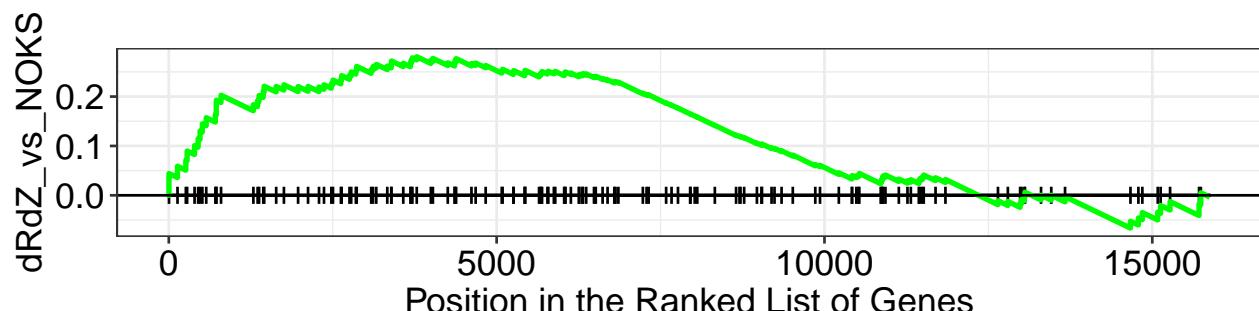
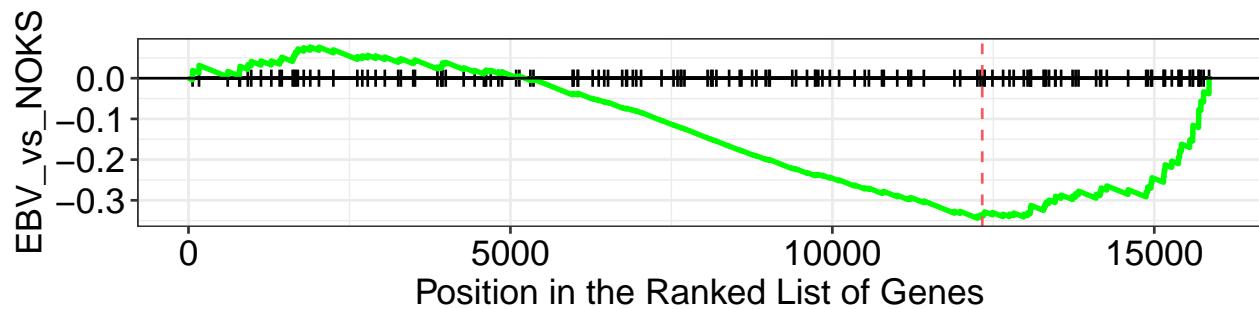
HALLMARK_ANGIOGENESIS



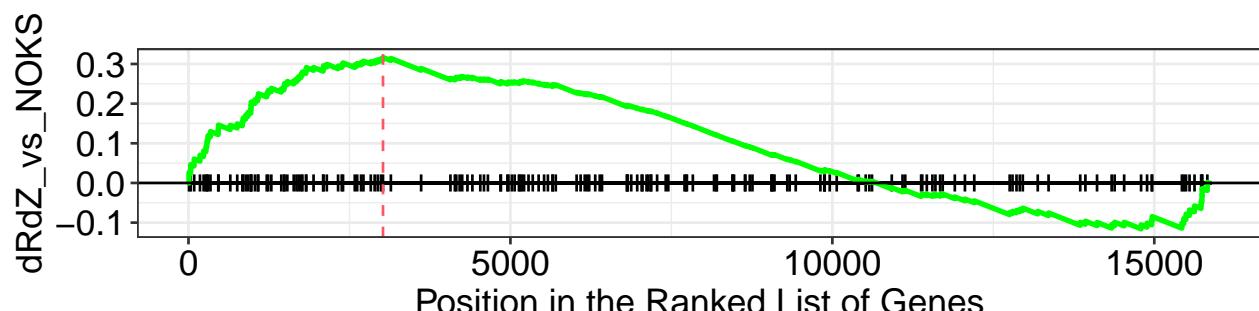
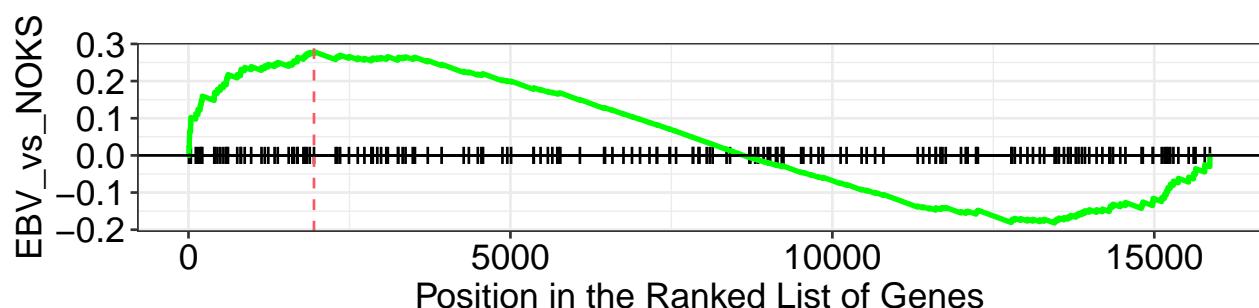
HALLMARK_COAGULATION



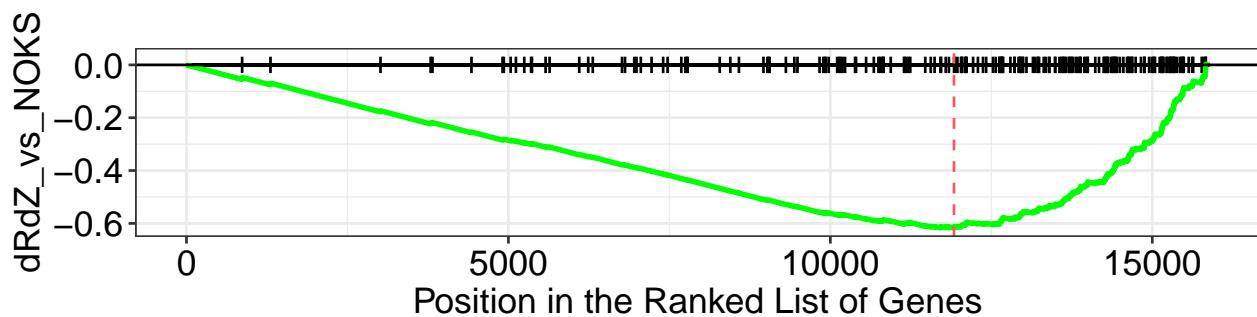
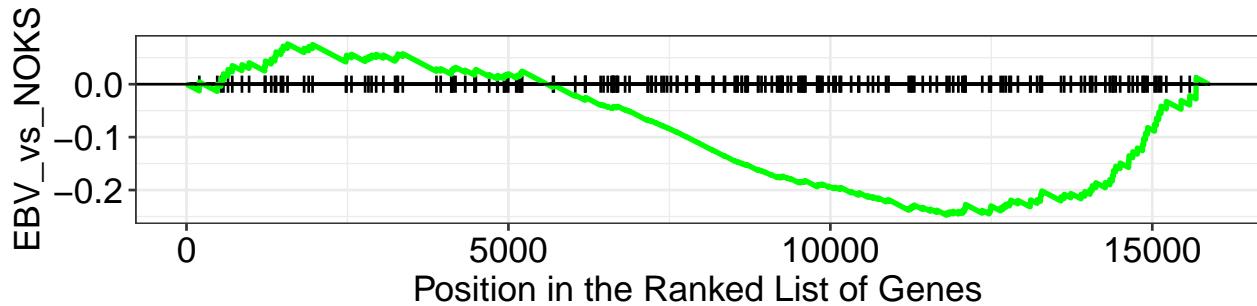
HALLMARK_KRAS_SIGNALING_DN



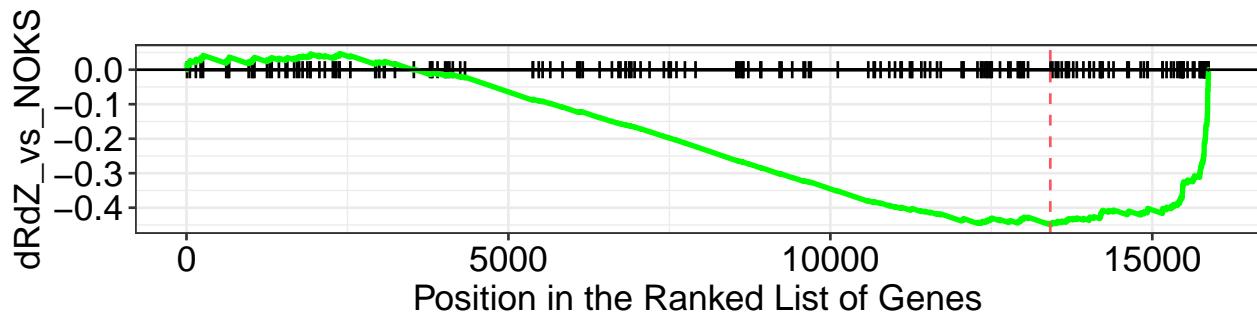
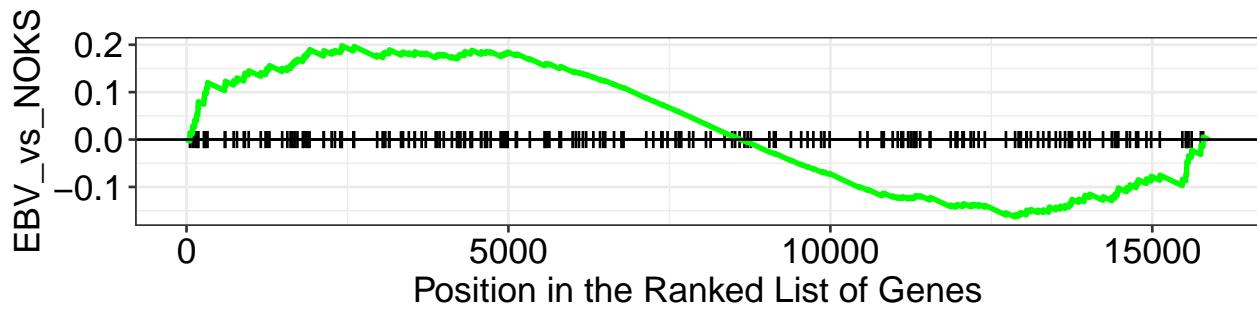
HALLMARK_APICAL_JUNCTION



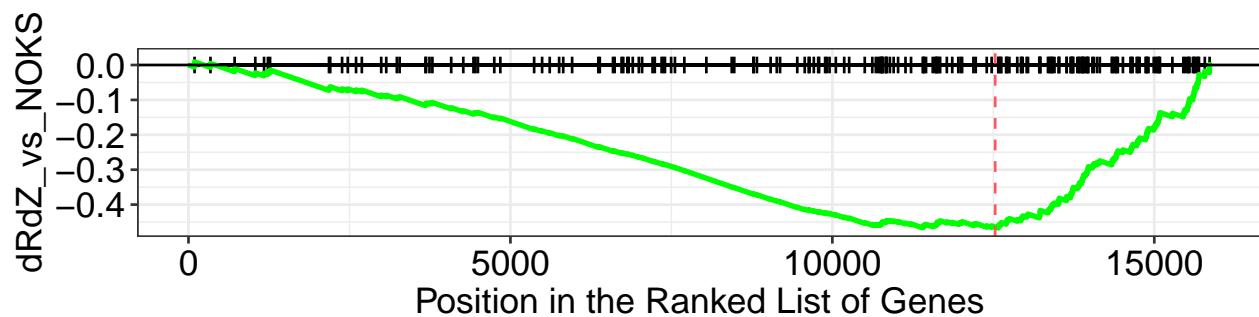
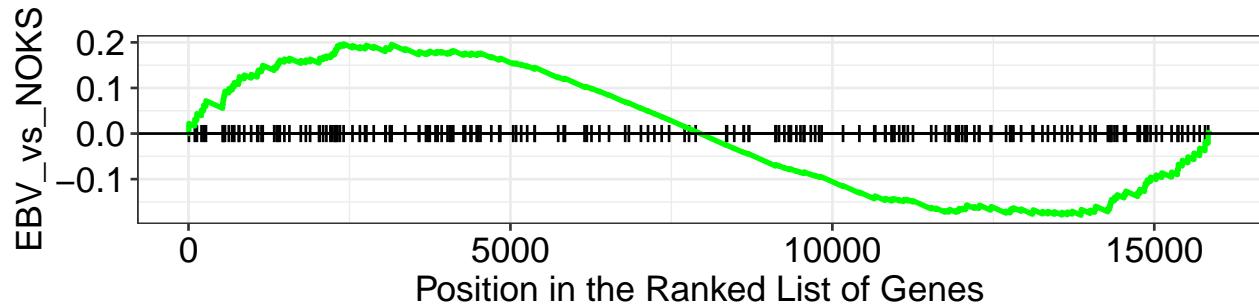
HALLMARK_MYC_TARGETS_V1



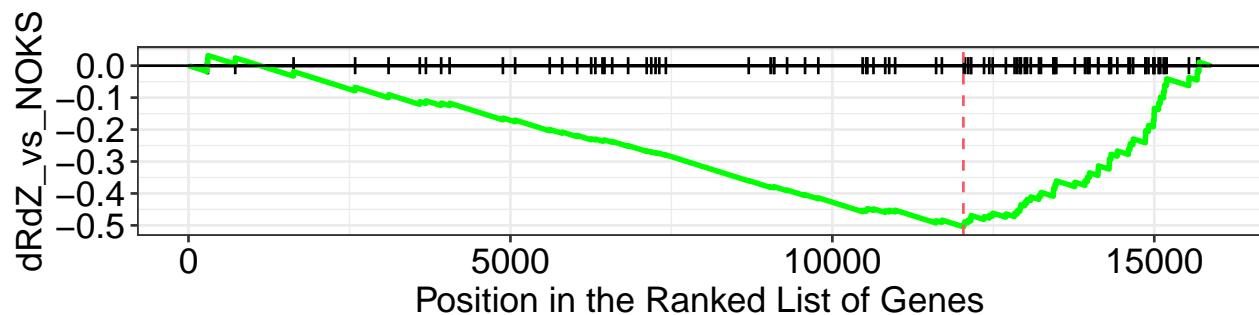
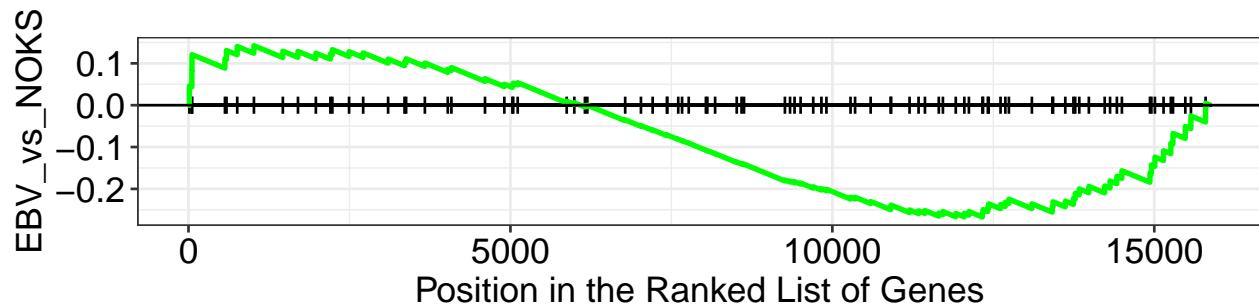
HALLMARK_MITOTIC_SPINDLE



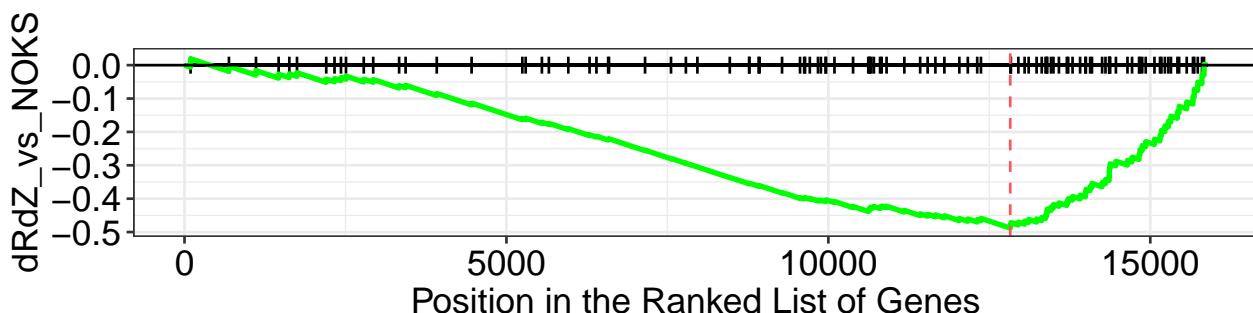
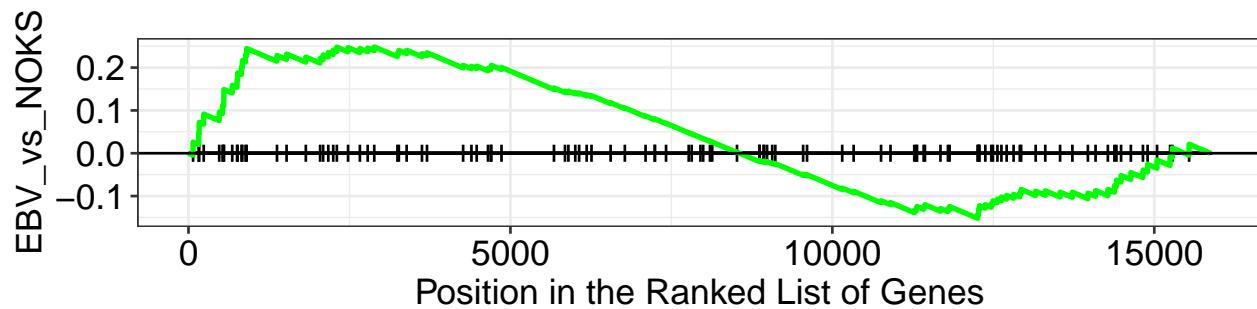
HALLMARK_MTORC1_SIGNALING



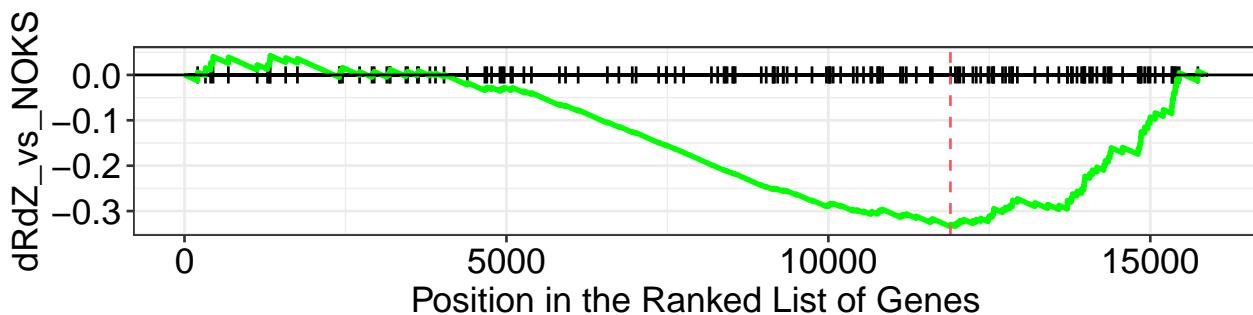
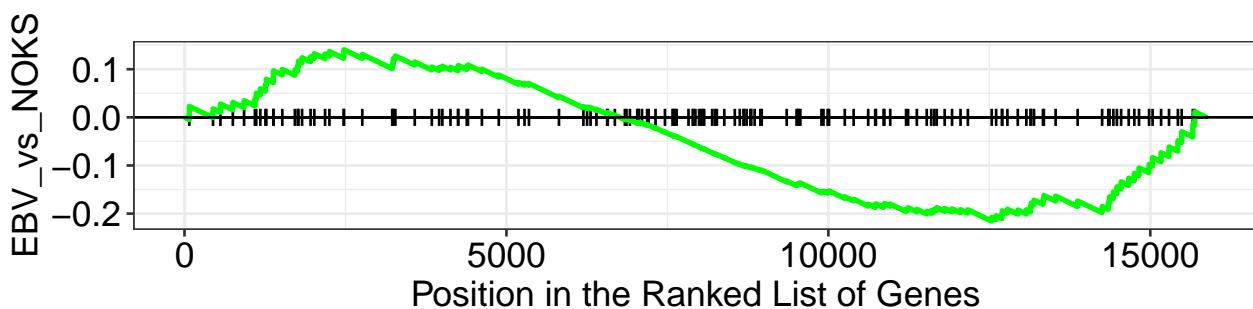
HALLMARK_PROTEIN_SECRETION



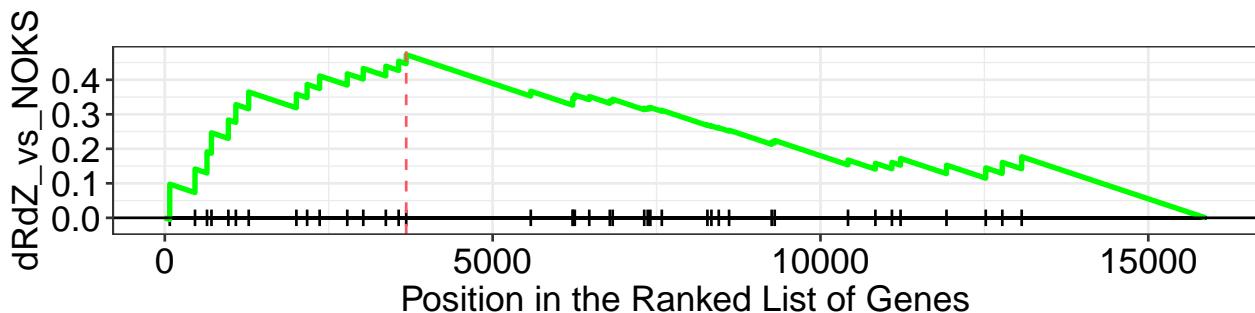
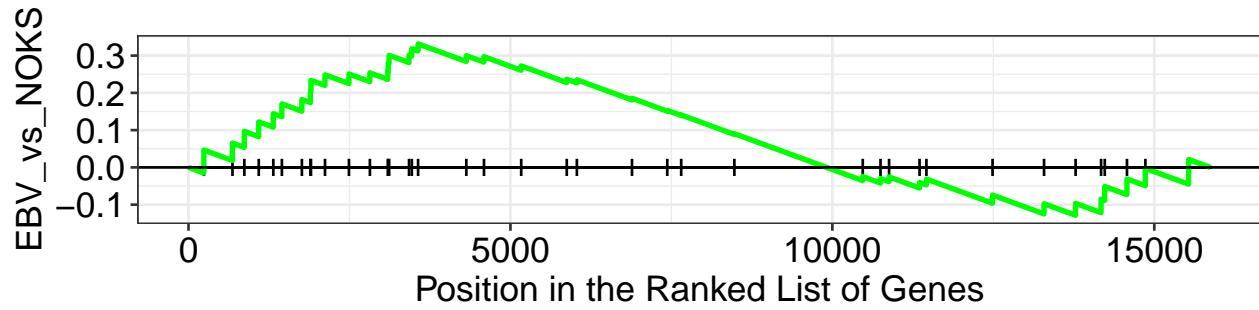
HALLMARK_UNFOLDED_PROTEIN_RESPONSE



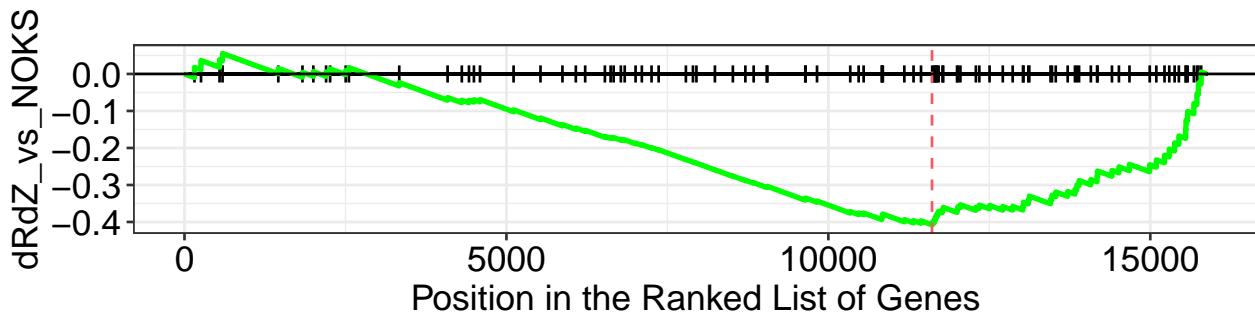
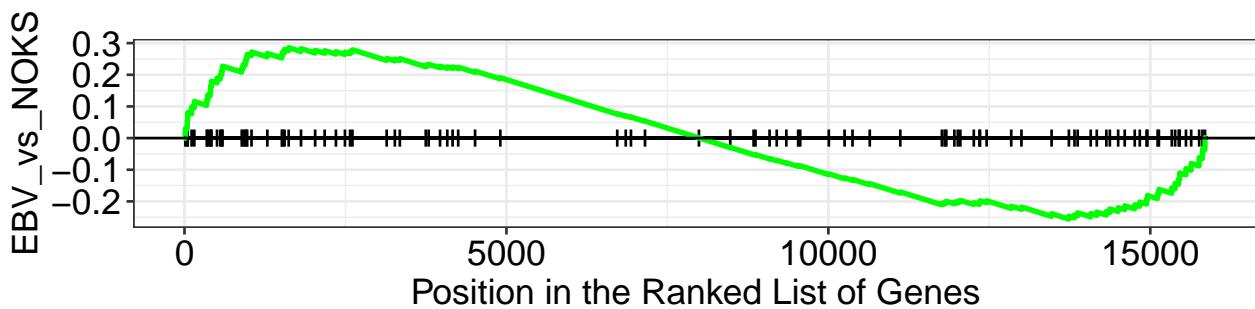
HALLMARK_DNA_REPAIR



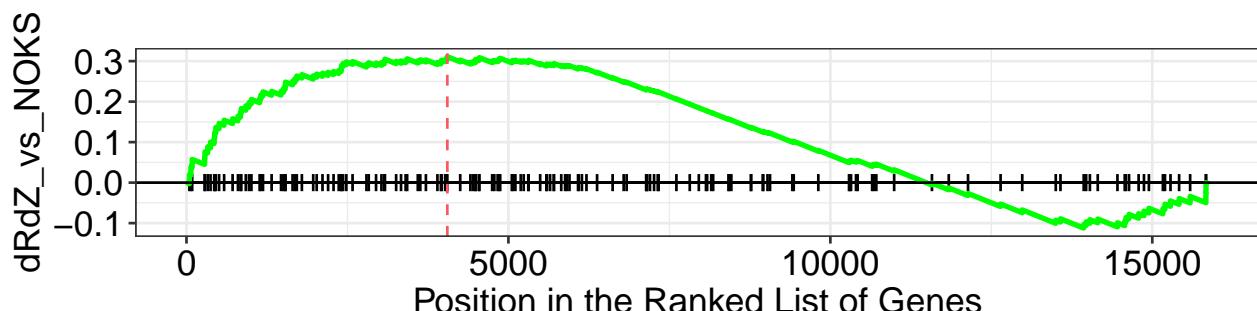
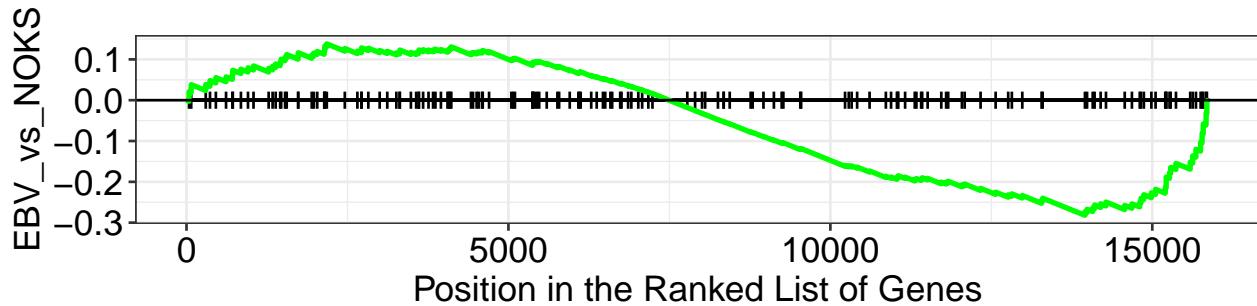
HALLMARK_APICAL_SURFACE



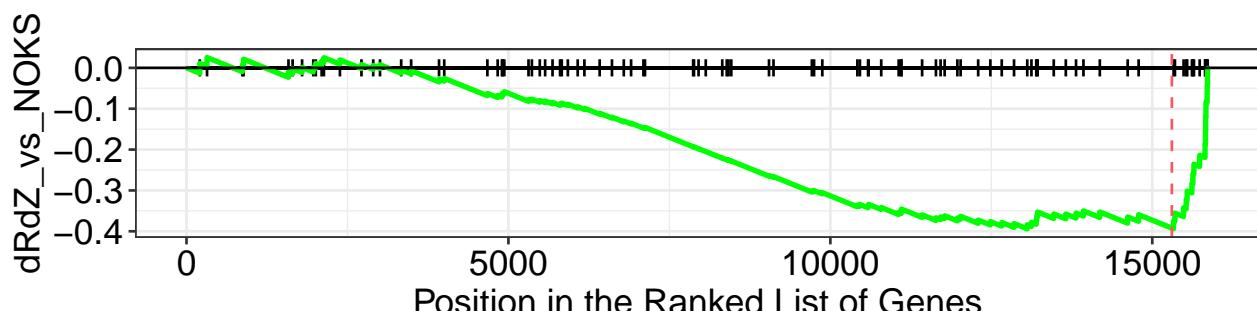
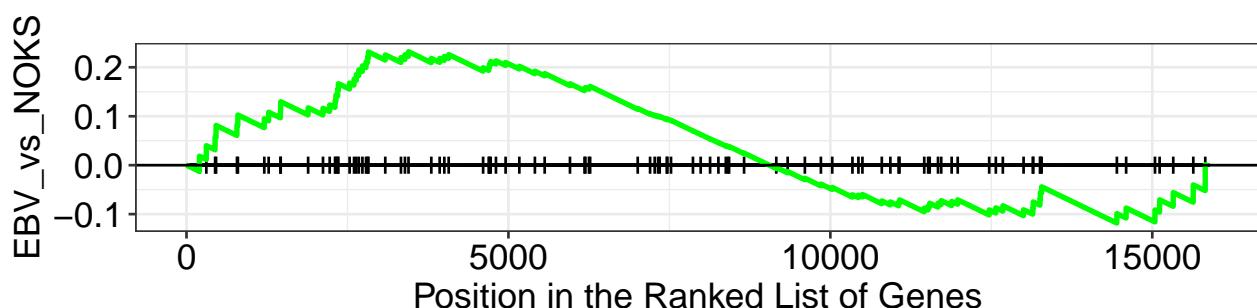
HALLMARK_ANDROGEN_RESPONSE



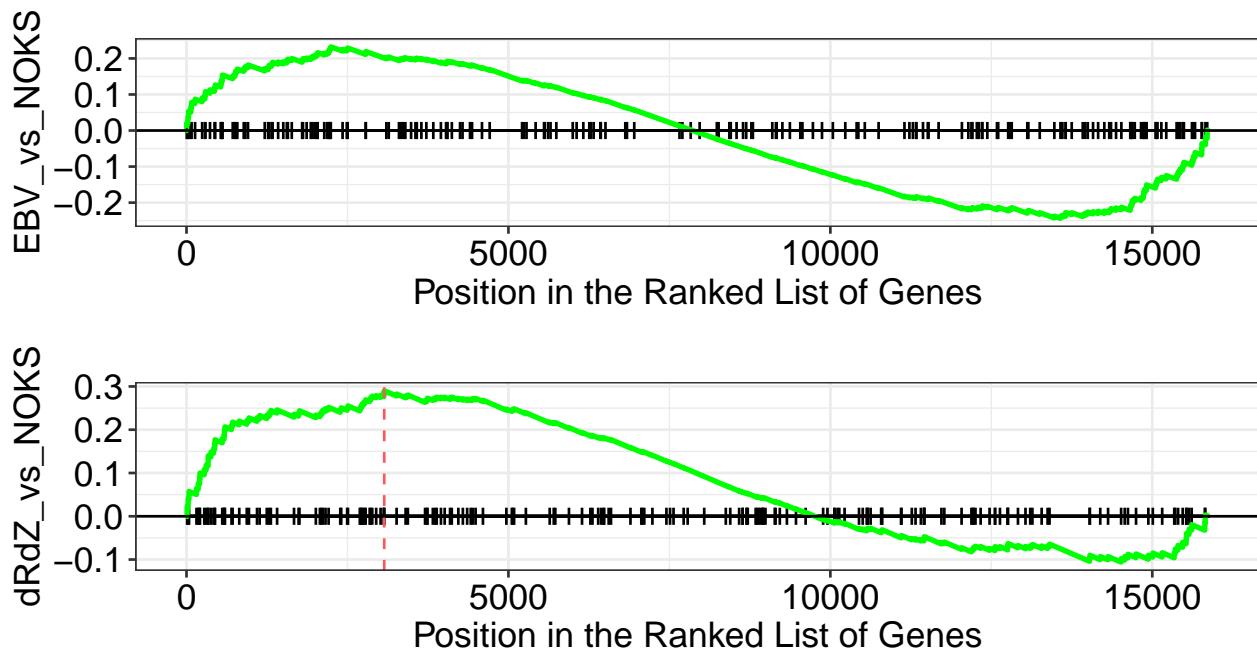
HALLMARK_MYOGENESIS



HALLMARK_SPERMATOGENESIS



HALLMARK_P53_PATHWAY

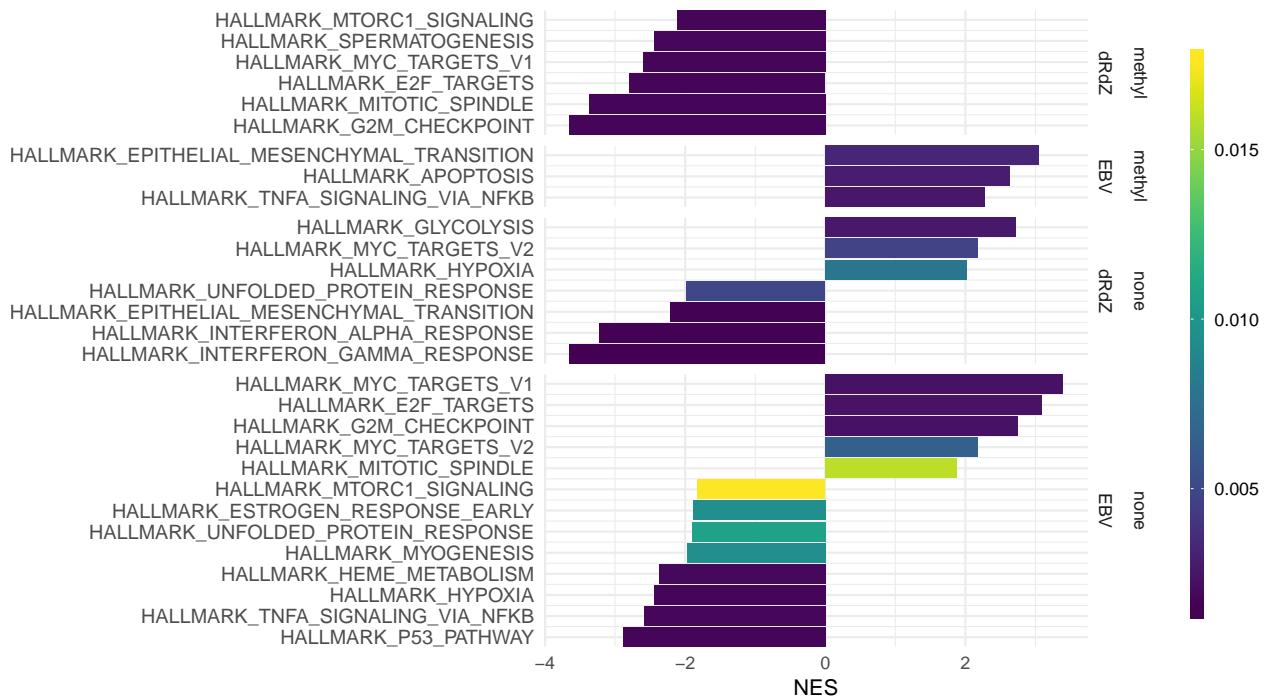


All genes (curated pathways)

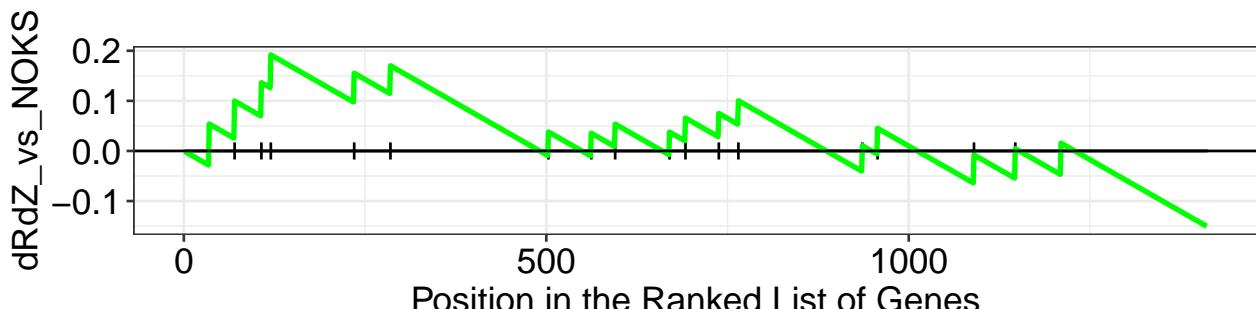
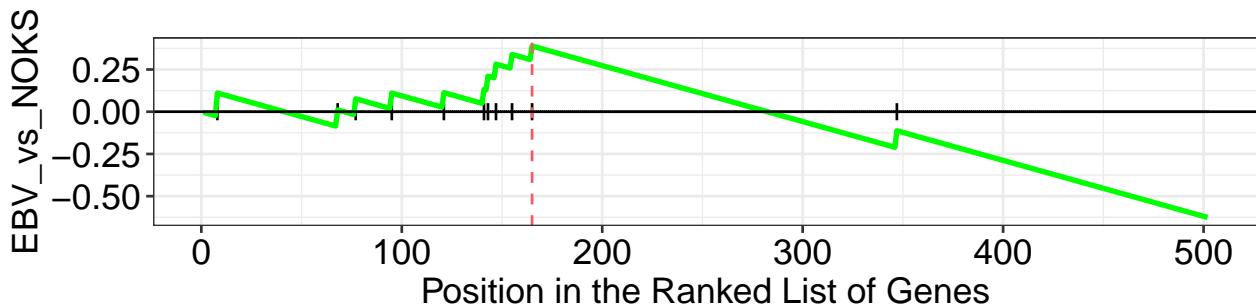


- Note: There are only those 3 pathways using all genes. When tried to make the same plot using differentially expressed genes didn't get anything.

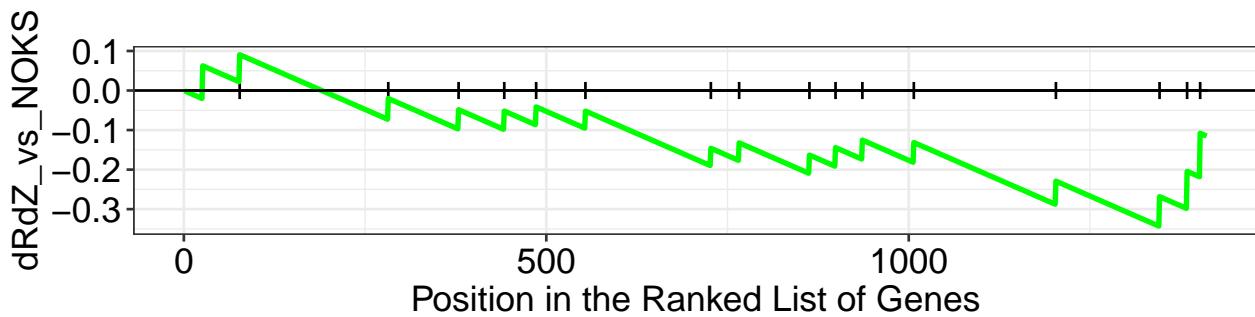
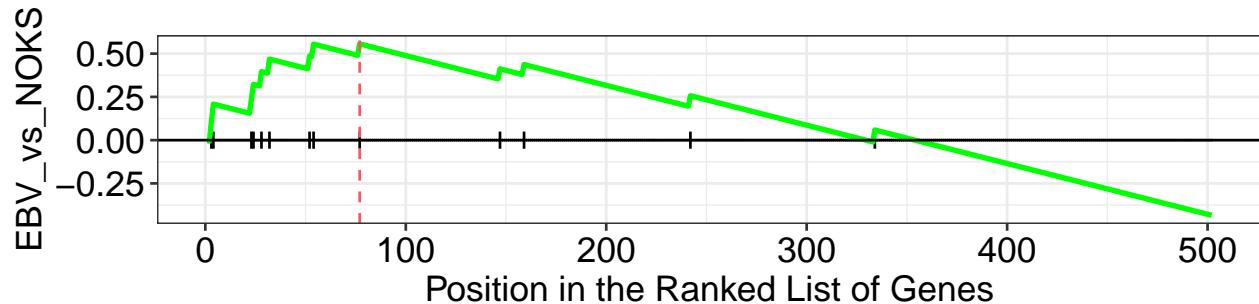
Differentially expressed genes (defined by $p_{adj} \leq 0.05$ and HALLMARK pathways)



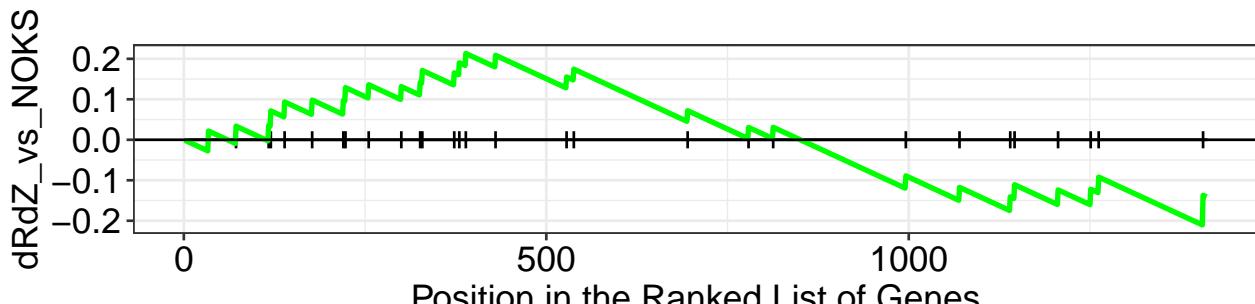
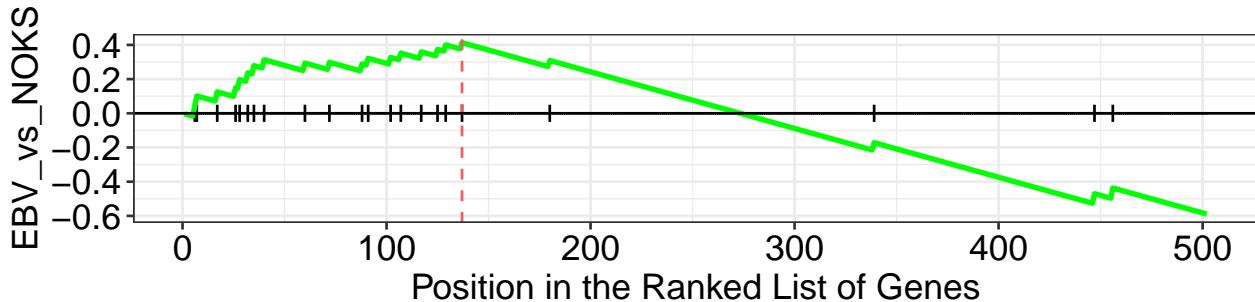
HALLMARK_TNFA_SIGNALING_VIA_NFKB



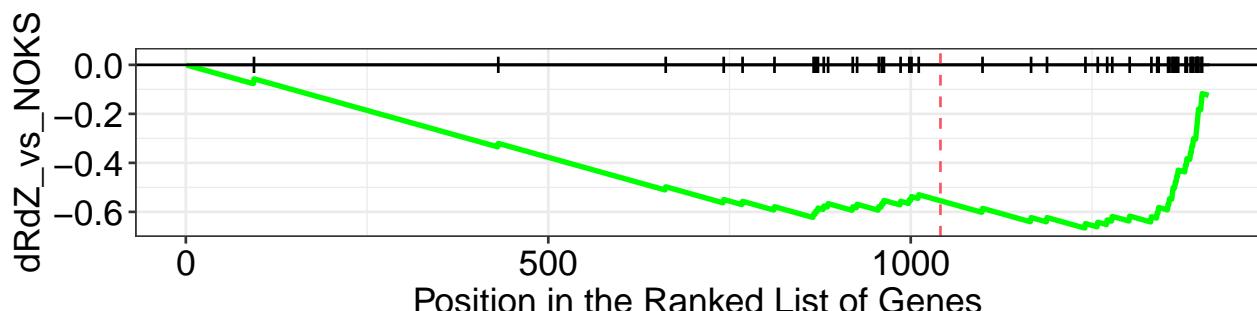
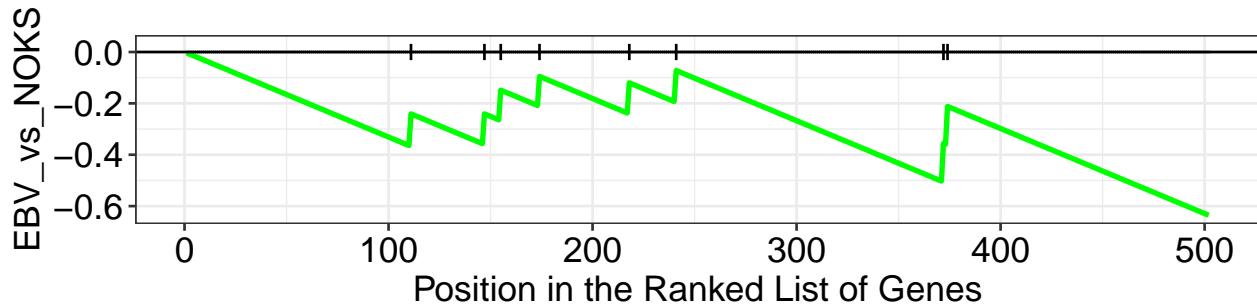
HALLMARK_APOPTOSIS



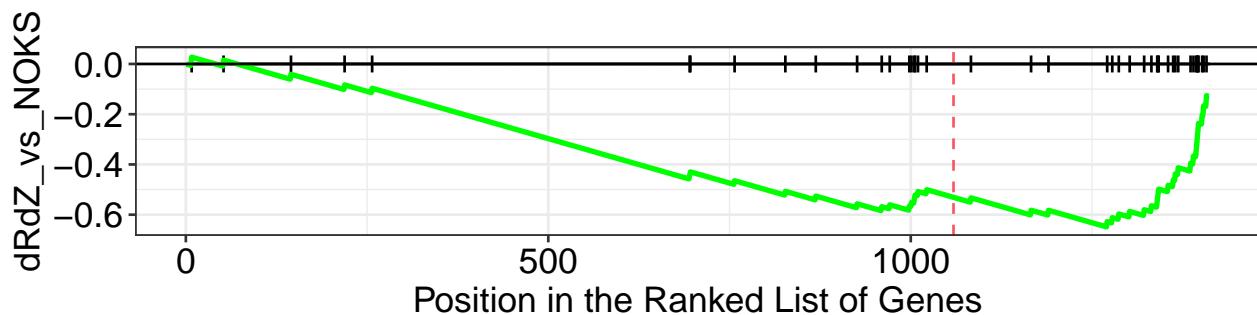
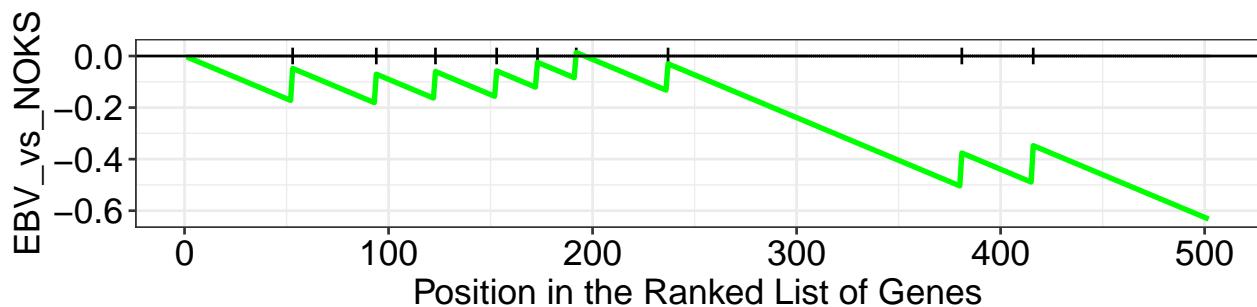
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION



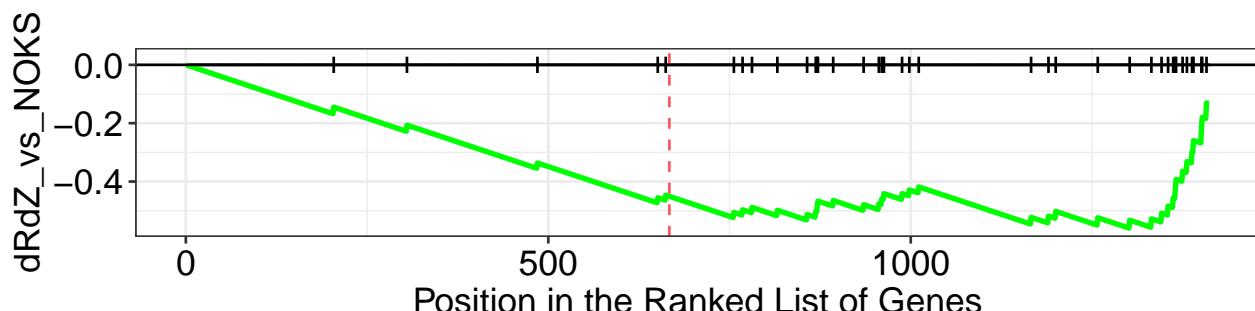
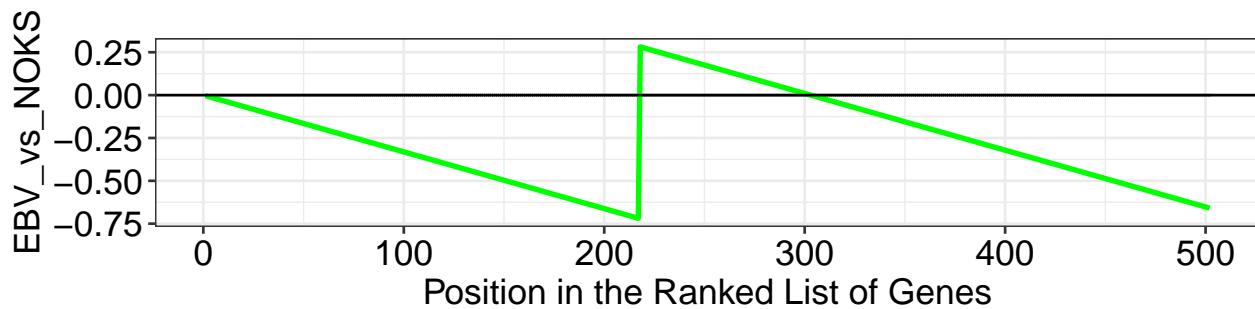
HALLMARK_G2M_CHECKPOINT



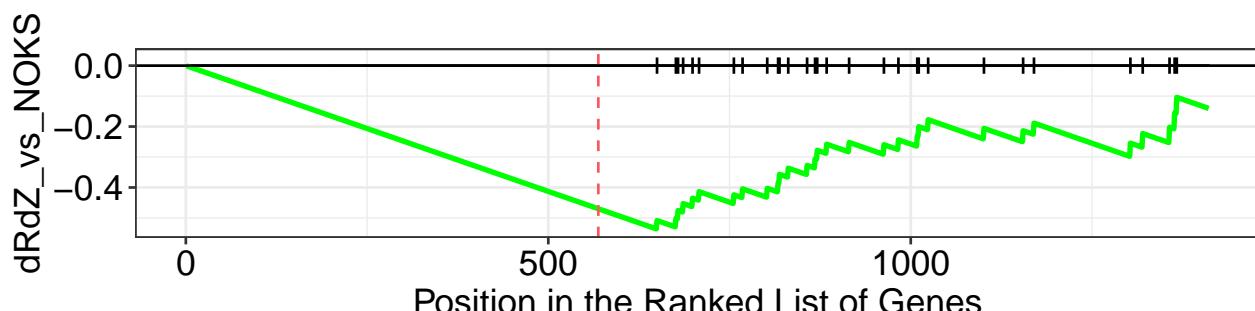
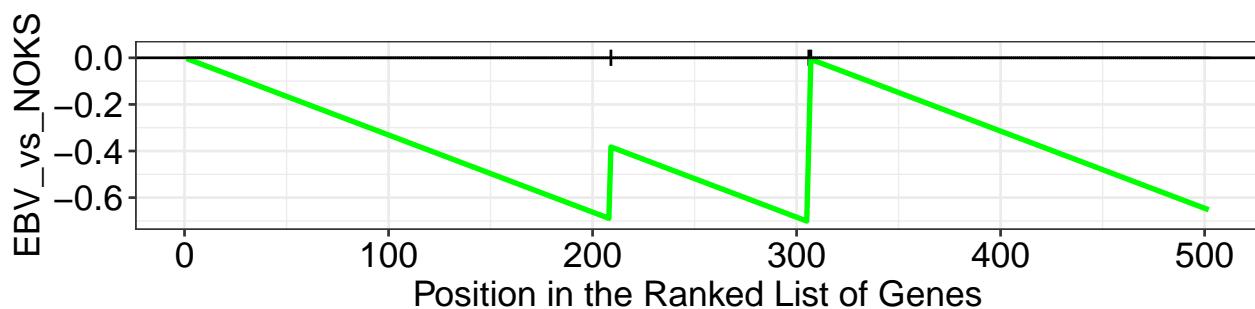
HALLMARK_MITOTIC_SPINDLE



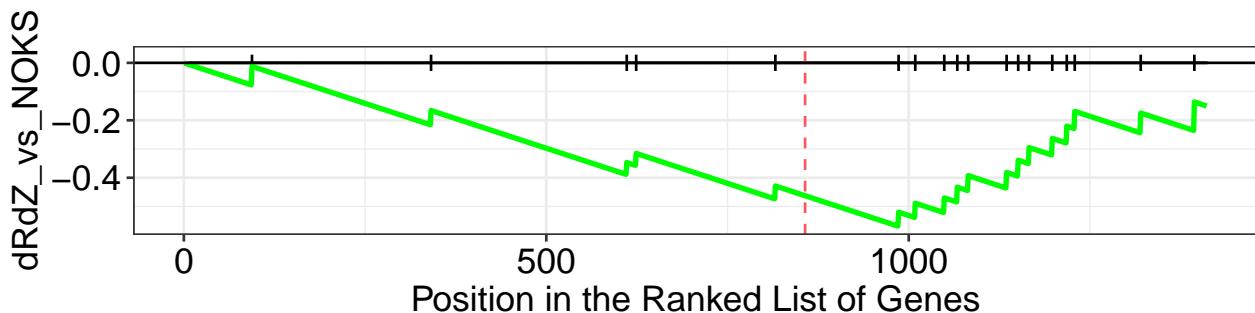
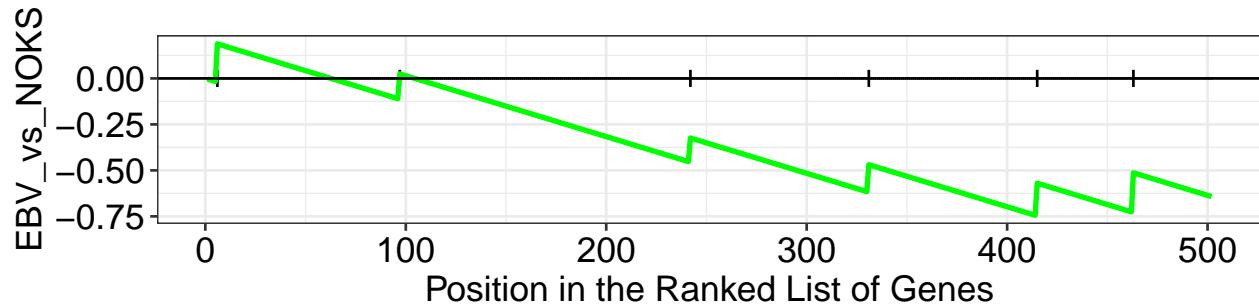
HALLMARK_E2F_TARGETS



HALLMARK_MYC_TARGETS_V1



HALLMARK_MTORC1_SIGNALING



HALLMARK_SPERMATOGENESIS

