Transcriptome profiling-based evidences reiterate the intrinsic role of MAPKAPK2 in facilitating molecular crosstalk during HNSCC pathogenesis

Sourabh Soni[1,3,#], Prince Anand[1,3], Mohit Kumar Swarnkar[2], Vikram Patial[1,3], Narendra V. Tirpude[1,3], Yogendra S. Padwad[1,3,*]

[*]Corresponding author: Yogendra S. Padwad (yogendra@ihbt.res.in, +919816747680)

Authors' information

[1]Pharmacology and Toxicology Laboratory, CSIR-Institute of Himalayan Bioresource Technology (CSIR-IHBT), Palampur-176061, India;

[2]Biotechnology Division, CSIR-Institute of Himalayan Bioresource Technology (CSIR-IHBT), Palampur-176061, India;

[3]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad-201002, India.

Sourabh Soni: sourabhsoniplp@gmail.com

Prince Anand: princeanandsaharsa4u@gmail.com

Mohit Kumar Swarnkar: mohitks@ihbt.res.in

Vikram Patial: vikrampatial@ihbt.res.in

Narendra V. Tirpude: narendra@ihbt.res.in

*Yogendra S. Padwad: yogendra@ihbt.res.in

[#]Current address: Department of Pediatric Newborn Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA. Email: ssoni4@bwh.harvard.edu

Abstract

Background: Transcriptomic profiling has been pivotal in better comprehending the convoluted biology of tumors including head and neck squamous cell carcinoma (HNSCC). Recently, mitogen-activated protein kinase-activated protein kinase-2 (MAPKAPK2/MK2) has been implicated in many human diseases including tumors. We recently reported that MK2 is a critical regulator of HNSCC and functions *via* modulating the transcript turnover of crucial genes involved in HNSCC pathogenesis. Consequently, to expand our insight into the biological relevance of MK2 and intricate cross-talks in tumor milieu, we conceptualized a comprehensive transcriptome analysis of HNSCC.

Results: We performed an extensive transcriptomic profiling to ascertain global patterns of gene expression in both *in vitro* and *in vivo* experimental models of HNSCC which exquisitely emulates the tumor microenvironment. Transcriptomic characterization substantiated an intrinsic role of MK2 and certain MK2-regulated genes in HNSCC pathogenesis, an outcome that reiterates our recent findings. Annotation and differential gene expression analysis revealed candidate genes whilst pathway enrichment analysis corroborated the biological significance of findings. Furthermore, advanced gene expression assays through nCounter system (primary validation) in conjunction with immunohistochemical analysis (secondary validation) validated the transcriptome profiling outcomes quite robustly.

Conclusions: Our results have underpinned the importance of seven differentially expressed MK2-regulated genes which are constitutively involved in HNSCC pathogenesis and could serve as potential candidates for future endeavors pertaining to therapeutic interventions and diagnosis pertaining to HNSCC. Collectively, our findings have paved the way towards the identification and development of new effective tumor markers and potential molecular targets for HNSCC management and improved clinical outcomes.

Keywords

Transcriptomics, Head and neck squamous cell carcinoma, MAPKAPK2, RNA-binding proteins, Pathogenesis, Differentially expressed genes, Gene regulation, MK2-regulated genes, nCounter Gene expression assays, Therapeutics

Background

Multifaceted regulatory networks tend to connect genes within a myriad of cellular processes. The involvement of a plethora of genes in fundamental biological processes *viz*. cell differentiation, growth and programmed cell death, and their involvement in many diseases is presently known (Spataro *et al.*, 2017). However, the apprehension of their roles at a global level is still incomplete. Gene transcription and regulatory networks in conjunction with new genome-wide approaches have garnered huge attention in the pretext of gene regulation. Nevertheless, post-transcriptional mechanisms like transcript stability and translation are also highly crucial and require intricate regulation *via* a multitude of intracellular signaling pathways. Particularly, the control of transcript stability through phosphorylation-mediated regulation of RNA-binding proteins (RBPs) by mitogen-activated protein kinases (MAPKs) has been a topic of huge interest (Lasa *et al.*, 2000; Venigalla *et al.*, 2012, Soni[1] *et al.*, 2019, Soni[2] *et al.*, 2019).

Head and neck squamous cell carcinomas (HNSCCs) having an incidence rate of ~600,000 cases yearly, is the seventh most common cancer worldwide and one of the most lethal with overall mortality rate of 40-50% (Jemal *et al.*, 2011; Ferlay *et al.*, 2015). HNSCCs are classified either histologically (Woolgar *et al.*, 2011) or *via* analysis of global transcription that employs etiology specific profiles (Martin *et al.*, 2014; Van Waes *et al.*, 2017). However, when these parameters were used for patient clustering, specific differences were observed in clinical behavior of patients as well as their response to therapy (Denaro *et*

al., 2014). The survival rates of HNSCC patients have not improved much, hence, HNSCC has been rightly termed as a malignant tumor having a low survival rate (Narayanan *et al.*, 2015). Consequently, an improved mechanistic insight into the molecular basis of HNSCC pathogenesis is urgently required to help in early diagnosis and development of effective therapeutics aimed at improved clinical outcomes (Jamali *et al.*, 2015).

The role of differentially expressed genes (DEGs) and endogenous RNA networks in HNSCC is not fully deciphered. Past reports on genome and transcriptome studies in various human tumors have revealed aberrant regulatory programs, driver mutations, and disease subtypes (Weinberg, 2014). The cancer genome study has acted as a valuable tool for classification, diagnosis, and prognosis in HNSCC. There have been many past reports pertaining to genomic alterations in HNSCC (Chau *et al.*, 2016; Bossi *et al.*, 2016). Recently, the cancer genome atlas (TCGA) led to a global analysis of the major molecular changes, comprehensive landscape of transcriptomic alterations and pathogenesis-linked signalling pathways in tumors thus contributing to the identification of novel prognostic biomarkers or specific anticancer molecular candidate groups (Cancer Genome Atlas Network, 2015; Campbell *et al.*, 2018). However, there is still need of extensive research insights to decipher the prognostic value attributed by these genomic alterations in tumors like HNSCC. A variety of biomarkers like MAPK phosphatase-1 (MKP-1), p16, p27, p53, tumor necrosis factor-α (TNF-α), and vascular endothelial growth factor (VEGF) have been shown to be linked with HNSCC (Thomas *et al.*, 2005; Soni[2] *et al.*, 2019), but they have not proved to be sufficient in accurately defining HNSCC pathogenesis. Single biomarkers have generally proved insufficient in the prediction of therapeutic response thereby necessitating research on combinatorial markers through high-resolution “omic” profiling (Bhatia *et al.*, 2012). Consequently, identification of new reliable molecular biomarkers associated with HNSCC

using omics-based analysis is the need of the hour in order to develop them as novel potential therapeutic targets (Fiore *et al.*, 2016).

Lately, the mRNA regulatory networks involved in tumor progression have garnered huge research interest. Recently, many reports have showcased the role of these intricate networks in tumorigenesis (Fang *et al.*, 2018). However, research insights in this area are quite limited, thereby pointing to a pertinent need of comprehensive analysis of mRNAs and regulatory networks. Next generation sequencing (NGS) has fast evolved as an important tool for epigenomic, genomic, and transcriptomic profiling (Cieślik *et al.*, 2018). Technological advances in mining and deciphering the vast cancer transcriptomic data has enabled us to better comprehend the complexity of various tumors and has streamlined efforts to discover novel biomarkers and therapeutic targets aimed at tumor management (Cieślik *et al.*, 2018).

In our recent study, we elucidated the role of mitogen-activated protein kinase-activated protein kinase-2 (MAPKAPK2 or MK2) in HNSCC pathogenesis using clinical tissue samples, cell lines and heterotopic xenograft mice model (Soni[2] *et al.*, 2019). We reported that MK2 is critically important in regulating HNSCC, and acts by modulating the transcript stability of crucial pathogenesis-related genes. We also established that knockdown of MK2 attenuates tumor progression in a xenograft mice model. Thereupon, in an attempt to delve deeper into the mechanistic insights of the role of MK2 and to decipher the molecular markers responsible for MK2-mediated changes in HNSCC pathogenesis, we conceptualized to perform a comprehensive transcriptome profiling of HNSCC cell line and xenograft-derived tumor samples evaluated in our recent study.

In the present study, we have evaluated the global mRNA expression profiles of our previously established HNSCC experimental model sets using transcriptome analysis on NovaSeq 6000 system (Illumina Inc., USA). The *in vitro* HNSCC cell line model including CAL27-MK2$_{WT}$ (wild-type) and CAL27-MK2$_{KD}$ (knockdown) cells cultured in both

normoxic as well as tumor microenvironment mimicking hypoxic conditions comprised the first set. The *in vivo* heterotopic HNSCC xenograft bearing tumors from CAL27-MK2$_{WT}$ and CAL27-MK2$_{KD}$ cells in immunocompromised mice as previously described (Soni[2] *et al.*, 2019) formed the second set. The findings of the transcriptomic profiling corroborated with our recently published results, hence, ascertaining a crucial role of MK2 in HNSCC pathogenesis *via* transcript stability regulation (Soni[2] *et al.*, 2019). Thus, by comprehensive transcriptome analysis we have elucidated certain specific MK2-mediated DEGs and regulatory networks that play an integral role in HNSCC pathogenesis in our experimental models. Furthermore, we performed gene expression assays on the nCounter system (NanoString Technologies, Inc., USA) in order to obtain a sensitive, highly multiplexed, and reliable detection of the specific mRNA targets based on our initial transcriptomic profiling. Finally, we cross validated the nCounter gene expression analysis results in *in vitro* setting using immunohistochemical analysis and found that the findings are in consonance and reiterate our recent discovery that MK2 is critically involved in regulating HNSCC progression and functions by modulating the transcript stability of crucial pathogenesis-related genes. The assays yielded highly precise and reproducible data which confirmed transcriptomic findings and yielded seven candidate MK2-regulated genes intrinsically involved in HNSCC pathogenesis which could potentially aid in discovery of new molecular markers for HNSCC management and diagnostic benefit.

Results

Qualitative assessment of the generated cDNA library followed by filtering and assembly of reads depicted optimum alignment

The RNA samples used for sequencing were isolated from appropriate cells/tissues using RNeasy Mini kit (Qiagen, Germany) and the quality assessment of the isolated total-RNA as well as the cDNA library formed was performed using Bioanalyzer (Agilent 2100, Agilent

Technologies, USA). The RNA integrity number (RIN) value of all the RNA samples was found to be >5 (exclusion criteria for this study). It is quite evident from the outcome of the quality check analysis that all the RNA samples utilized, as well as the cDNA library constructed were of apt quality for their use in downstream experiments (Figure 1a, S1 and S2).

Deep sequencing of RNA obtained from NovaSeq 6000 platform resulted in 349 million raw reads (~58.2 million raw reads per sample) of average insert size 210bp. Figure 1a and S3 summarizes the quality check (QC) results of the sequencing experiment. The raw FASTQ sequences were filtered using NGSQC tool kit to obtain high quality (HQ) reads based on the set parameters, generating 258 million filtered HQ reads (~43.1 million HQ reads per sample), amounting to 74.1% of total raw reads, hence, implying that the obtained data was of good quality. A total of 25.7 giga bases of data was generated which foretells the enormity and wide complexity of the human genome.

The HQ reads obtained were further considered for downstream analysis and mapped over the human reference genome. The alignment performed employing the Kallisto pipeline was optimum with approximately 88.7% of the HQ reads on an average getting mapped to the human reference genome (Figure 1a, 1b and S4).

Annotation and differential gene expression analysis revealed candidate genes

The present study was conducted using two distinct experimental datasets as mentioned in Figure 1a and Table 1. Using the criteria of fragments per kilobase of transcript per million mapped (FPKM)≥0.1, we identified on an average around 62791 expressing transcripts in all the samples, representing thousands of genes. A-D datasets depicted the *in vitro* HNSCC cell line samples while F vs E illustrated the *in vivo* xenograft dataset. As detailed in Figure 1a and Table 1, we performed analysis in $MK2_{KD}$ vs $MK2_{WT}$ scenario under both normoxic and hypoxic conditions to obtain a comprehensive picture of the changes in the global gene

expression pattern in the backdrop of tumor core emulating hypoxic niche. Firstly, we evaluated the differential gene expression in the various studied datasets and found a large pool of DEGs, precisely 1403 in B vs A, 924 in D vs C, 1360 in D vs B, 1456 in C vs A, and 984 in F vs E as assessed by the pre-set cut-off values. Figure 1c represents the total number of upregulated (fold change (FC)>2) and downregulated (FC<-2) genes among all the DEGs in the analyzed datasets.

Pathway enrichment analysis exhibited the biological significance of our findings

The multitude of DEGs in various experimental datasets in the transcriptomic profiling are implicated in hundreds of significant biologies/biological processes as summed up in Figure 1d. To gain further insight into the biological significance of changes in gene expression and to attain a global picture of the molecular pathways possibly contributing to HNSCC pathogenesis, pathway enrichment analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. It yielded an integrated network analysis and revealed the top five biological processes enriched in the experimental datasets. The top gene ontology pathway analysis for the DEGs is depicted in Figure 1e. It was notably found that a significant percentage (~5%) of the total DEGs in the various studied combinations belonged to the pathways involved in cancer progression (Figure 1e). The DEGs showing significant changes between various groups were then selected, followed by generation of heat maps to assess clustering of gene expression profiles among the experimental datasets (Figure 2a). Figure 2b further depicts the distribution of all the transcripts on the two dimensions of -log(P) and FC by way of volcano plots with differentially expressed transcripts highlighted in blue.

In A-D datasets, the DEGs belonged to a large number of biological processes thereby limiting the information that could be harnessed. Henceforth, in an effort to filter down data and fulfill the aim of extracting valuable leads, we selected the 77 elements/biological

processes that were common in A-D combinations using Venny 2.1.0 (http://bioinfogp.cnb.csic.es/ tools/) (Figure 2c). These processes belonged to numerous categories which are listed in Table S1. On the background of this knowledge, we further filtered down the common genes in these 77 biological processes (Figure 2d). Consequently, we obtained five genes common in A-D combinations namely EH domain binding protein 1 (EHBP1), SMC5-SMC6 complex localization factor 2 (SLF2), death associated protein 3 (DAP3), inositol hexakisphosphate kinase 2 (IP6K2) and runt related transcription factor 1 (RUNX1) as listed in Table S2. Notably, these genes play specific roles in HNSCC pathogenesis possibly *via* MK2-mediated regulation.

Gene regulatory networks and pathways depicted the importance of candidate genes in HNSCC pathogenesis

Gene regulatory networks for these five common genes in A-D datasets furnished a detailed overview of the various inter-connections and the biological processes affected by them. Collectively, the results showed that IP6K2 plays an intrinsic role in nucleotide binding, while DAP3 in apoptosis and poly(A) RNA binding as confirmed in past reports (Thomas *et al.*, 2014; Wazir *et al.*, 2015). Interestingly, these genes showed differential regulation in various experimental datasets (Figure S5-S8), hence, clearly signifying that MK2-knockdown as well as hypoxic tumor microenvironment affects the genes and pathways in HNSCC *via* differential regulation, pointing to a central role of MK2 in transcriptional regulation of HNSCC.

Furthermore, to retrieve the information about the MK2-regulated candidate genes intrinsic to HNSCC, we narrowed down our search to cancer related processes for further analysis. This filtered down our dataset to 16 cancer-specific biological processes that were common in A-D datasets (listed in Table S3). In these 16 processes only two genes, DAP3 and RUNX1, were found to be common as per the analysis performed using Venny 2.1.0

(Figure 2e and Table S4). Collectively, these findings have indicated that these two genes are intrinsically involved in HNSCC pathogenesis and warrants further investigation in this regard. Notably, the 16 processes were found to be clustered in 5 major biological pathways (Figure S9), clearly showcasing that MK2 portrays an intrinsic role in the hypoxic tumor microenvironment by regulating processes like apoptosis and transcription, hence, potentiating and strengthening our latest *in vitro* findings (Soni[2] *et al.*, 2019).

Similarly, on comparing the mice bearing CAL27-MK2$_{KD}$ tumors with those bearing CAL27-MK2$_{WT}$ tumors (F vs E xenograft dataset), we found that the DEGs were clustered in 14 biological processes of relevance to tumor pathogenesis as listed in Table S5. The genes involved in these biological processes were then filtered out and Figure S10 shows the gene regulatory network for the same. This analysis provided us with certain candidate genes like TRAF2 (apoptosis); EPB41L1 (cytoskeleton); FOXO3, H2AFY and YAP1 (transcriptional regulation); and DIDO1 (RNA binding) that are involved in key cellular processes in the xenograft model. These findings can be explored further to decipher the putative role of these potential candidate genes in HNSCC pathogenesis with an aim to define probable therapeutic targets for HNSCC management.

MK2 mediated master of regulatory network functions by modulating the transcript stability

The prime objective of this study was to gather comprehensive information about the various DEGs and pathways rendering essential roles in HNSCC pathogenesis in the presence and absence of MK2 in both normoxic as well as the tumor core mimicking hypoxic conditions. Thereupon, keeping MK2 at the nexus of further analysis, we shifted the focus on elucidating in detail, the regulation of MK2 pathway and its downstream targets in various experimental datasets. It was found that MK2 is involved in regulation of six major biological pathways as shown in Figure S11. Recent i*n vitro* findings by our team have asserted that MK2 controls

transcript stability of critical genes involved in HNSCC pathogenesis *via* RBP-mediated regulation (Soni[2] *et al.*, 2019). Hence, we analyzed the transcriptomic data in regard to deciphering the role of MK2 in the regulation of mRNA stability. Interestingly, MK2 was found to accomplish the task of regulating the transcript stability through RBP-mediated regulation with HuR (ELAVL1) and TTP (ZFP36) playing intrinsic roles (Figure 3a), thus clearly affirming the hypothesis and corroborating our previous findings (Soni[2] *et al.*, 2019). The levels of expression of these RBPs and hence regulation was different in various experimental datasets clearly suggesting that tumor environment, in association with presence/absence of MK2, plays an important role in HNSCC pathogenesis (Figure 3a).

Next, to attain a clear picture of what was happening at the transcriptional level, we narrowed down our analysis by filtering down to selected genes of the MK2 pathway that were previously studied in MK2-centric studies and were also analyzed in our recent *in vitro* study. The genes selected for this analysis were p38, MK2, AUF1, TTP, CUGBP1, CEBPδ, HuR, MKP-1, p27, TNF-α and VEGF, which have been known to be involved in a number of key cellular pathways. The analyzed genes were clustered into eight major biological processes as shown in Figure S12. Elucidation of the MAPK signaling cluster in detail indicated that in the background of MK2-knockdown in normoxia (B vs A dataset), VEGF and TNF-α tend to show downregulation (Figure 3b) which is in complete consonance with our previously established findings. Collectively, transcriptomic analysis results were quite promising and corroborated very well with previous findings, thereby, strongly validating our hypothesis that MK2 is the master regulator of transcript stability of genes critical to HNSCC pathogenesis.

3'-untranslated region-based filtering furnished information regarding important MK2-regulated downstream target genes

In the present study, our prime focus was role of MK2 and MK2-regulated genes in HNSCC pathogenesis. MK2 can potentially regulate the transcript stability of only those targets that possess the binding regions for RBPs in their 3'-untranslated region (3'-UTR). Practically, only these genes could be the probable candidates of MK2 downstream targeting *via* RBP-mediated regulation. Hence, we narrowed down the transcriptomic analysis to only those DEGs that harbored RBP binding regions in their 3'-UTRs. To accomplish this task, we fetched out the 3'-UTR regions of all DEGs using Ensembl (http://www.ensembl.org/) (Aken *et al.*, 2016). Next, the domain sequences of RBPs were assessed using catalog of inferred sequence binding preferences of RBPs (CISBP-RNA) database. Lastly, the transcripts that harbour RBP specific regions in their 3'-UTRs were filtered out using RBPmap v1.1 web tool (http://rbpmap.technion.ac.il/). Using this approach, we analyzed the top two upregulated and downregulated genes in previously selected 16 cancer specific pathways (Table S3) in the transcriptomic profiling of the *in vitro* HNSCC cell line model (A-D datasets) which resulted in 34 MK2-regulated genes listed in Table S6. Similarly, for the *in vivo* heterotopic HNSCC xenograft experimental dataset (F vs E dataset), we filtered down the topmost upregulated and downregulated genes in all the cancer specific pathways which provided us 48 MK2-regulated genes that are listed in Table S7. This specific filter-based transcriptomic analysis brought into limelight possible MK2-downstream target genes that could be integral in HNSCC pathogenesis. Furthermore, to cross-validate findings of the transcriptomic profiling, these filtered candidate genes along with the 5 common genes in the 77 common elements in A-D datasets listed in Table S2 were used for further *in vitro* experimentation. One gene (H2AFY) was common in the transcriptomic analysis for both A-D and F vs E datasets resulting in a total of 86 genes (34 MK2-regulated genes and 5 common genes for A-D datasets, 48 MK2 regulated genes in F vs E dataset, and 1 gene was common in both) that were selected for further validation.

Highly efficient and precise detection of gene expression *via* nCounter gene expression assays potentiated transcriptomic outcomes

Routinely, the findings of transcriptomic analysis are generally validated in *in vitro* setting *via* gene expression analysis employing qRT-PCR and using the same RNA sample in order to maintain homogeneity. In lieu of the high-throughput nature of our validation, qRT-PCR analysis could have been prone to numerous errors; hence, we employed a latest and highly precise gene expression assay-based nCounter system approach (NanoString Technologies, Inc.) to validate the transcriptome analysis. To accomplish this assay, 90 specific custom designed molecular probes corresponding the selected candidate genes were procured from NanoString Technologies, Inc. These probes aided in imaging and detection of multiple transcripts (90 in this case) in single reaction with high fidelity rate. The gene set comprised of the 86 selected genes from the transcriptomic profiling as well as four housekeeping genes (HKG). The HKG were selected based on extensive literature survey which ascertained that their levels remain unchanged in HNSCC (listed in Table S8 and S9, respectively).

The assay was performed using standard procedure as highlighted in Figure S13 and S14 and detailed in the methods. Briefly, the custom synthesized probes were hybridized to the target RNA samples followed by washing of the excess probes. Further, immobilization of the probe/RNA complexes on the nCounter cartridge was performed, samples were run on the nCounter instrument and finally data was collected. The statistical analysis ($p<0.05$) indicated that some of these genes were differentially expressed in the various experimental datasets. We next performed pathway enrichment analysis of the nCounter gene expression assay data to determine the biological significance of changes in gene expression. KEGG enrichment analysis revealed the top five biological processes (Figure 4a and 4b), while heat map analysis deciphered the clustering of gene expression profiles among various combinations (Figure 4c). The results from the nCounter assays correlated with the transcriptomic analysis (Table 2 and 3), hence substantiating our findings.

Furthermore, considering the individual CodeSet of 39 genes for the *in vitro* HNSCC cell line model (A-D datasets) and 48 genes for the *in vivo* heterotopic HNSCC xenograft experimental dataset (F vs E dataset), the heat-map and box-plot representation of the results as analyzed by nCounter gene expression assay have been shown in Figure 4d and 4e, respectively. These results indicated the expression profile and variation in FC among DEGs. Figure 4d clearly depicts the upregulated and downregulated DEGs in the nCounter assays and the results were in consonance with transcriptomic profiling with a high percentage of genes showing similar pattern of expression and even matching fold change values as showcased in Table 2 and 3. For B vs A dataset, a total of 39 genes were analyzed, out of which 24 matched with the transcriptomic analysis (61.6% matching score). The matched genes were then analyzed for FC where, BRD2 was found to be the only upregulated gene (FC>2) and CLK2 was the only downregulated gene (FC<2). Similarly, after analysis of all the datasets we found that 12 DEGs were common in our transcriptomic and nCounter analysis (the results have been summed up in Table 4). Notably, these genes are key players in important processes like cell cycle progression, apoptosis and transcription regulation, hence, potentiating their role as important MK2-regulated genes involved in HNSCC pathogenesis. Thus, these results strengthen our recent findings that MK2 is critically important in regulating HNSCC and functions by modulating the transcript stability of crucial pathogenesis-related genes. Further, detailed statistical analysis accentuated that expression of only 7 (MELK, ZNF662, BMP7, CREB3L1, IGFBP2, MUC4, PRKAR2B), out of the 12 candidate genes were significantly changed (FC>2 or <-2, $p<0.05$) (Table 2 and 3). MELK was the only gene belonging to the C vs A dataset (cell line comparison) while the other 6 genes were from the F vs E dataset (xenograft comparison). Hence, we next analyzed these 6 genes (ZNF662, BMP7, CREB3L1, IGFBP2, MUC4, PRKAR2B) *in vitro* by immunohistochemistry (IHC) to ascertain our transcriptomic and nCounter findings in an experimental HNSCC xenograft model (Figure 5a).

Immunohistochemical analysis indicated probable role of candidate genes in HNSCC pathogenesis

To add another of confidence to the dataset and probe the role of candidate genes in HNSCC pathogenesis immunohistochemical analysis was performed in tumor xenograft tissues. The six filtered candidate DEGs that were common in transcriptomic and nCounter data analysis in F vs E dataset were then analyzed in tumor sections from xenografted animals using IHC to evaluate the expression levels of these proteins (Figure 5). Results obtained from IHC largely strengthen our findings about the common DEGs in our transcriptomic and nCounter analysis. The *in situ* expression of IGFB2, MUC4, PRKARB2 was mostly present in the cellular cytoplasm and stroma of squamous cell carcinoma (Figure 5). The tumor sections with the expression have cellular pleumorphism, mitotic figures and formation of nests of tumor cells. It is clearly evident from the IHC outcome that expression levels of three MK2-regulated candidate genes (IGFB2, MUC4, PRKARB2) were upregulated in tumor xenografts created using CAL27-MK2$_{KD}$ cells as compared to CAL27-MK2$_{WT}$ cells (Figure 5). These genes are widely considered imperative to processes like cell cycle progression, apoptosis and transcription regulation. There was no significant change in the protein expression levels of the three other analyzed genes (data not shown). Findings of the Immunohistochemical analysis clearly implied the elevated expression level of IGFB2, MUC4, and PRKARB2 in tumor xenografts, it corroborates the obtained datasets from transcriptome analysis and nCounter gene expression validation. These finding clearly upholds the importance of these MK2 regulated candidate genes in HNSCC pathogenesis.

Discussion

To improve understanding of the convoluted biology and leverage the outcomes to optimize the management of HNSCC, there have been many efforts to characterize HNSCC at the transcriptome level. Methodological breakthroughs in the recent past have revolutionized the

area of transcriptome profiling by providing a link between the molecular mechanisms and cellular phenotypes (Cheng *et al.*, 2018; Cieślik *et al.*, 2018). In recent times, a comprehensive landscape of genomic and transcriptomic alterations in squamous tumors including HNSCC had emerged from the TCGA network (Lawrence *et al.*, 2015; Campbell *et al.*, 2018). Cellular models that can comprehensively characterize metastatic HNSCC are still lacking and hence translationally relevant transcriptome profiling underlying the basis of HNSCC metastasis has proved to be a powerful tool for preclinical research (Byron *et al.*, 2016; Nisa *et al.*, 2018). In the context of the present study, this link presented us with an opportunity to thoroughly dissect the huge complexity and large heterogeneity in head and neck tumors in an attempt to discern novel biomarkers and potential therapeutic targets.

Generally, there are two well characterized ways to analyze cancer transcriptomic data. First is the differential approach, in which tumor expression profiles relative to the normal patient-matched or even unmatched tissues are analyzed. Second is the relative approach, in which transcript levels across different tumors or other samples are compared. The simplest example of the differential approach is identification of the up/down-regulated DEGs in a particular sample. There are many established methods that help in detection of DEGs for both microarray-based approach as well as RNA-sequencing (Gentleman *et al.*, 2006; Love *et al.*, 2014). A typical transcriptome profiling result is generally a never-ending list comprising of thousands of DEGs, hence, it has always been very difficult to interpret this data without additional filtering *via* functional annotations. A large variety of methods are available for analysis of DEGs and for obtaining critical understanding of the pathways, gene regulatory and co-expression networks involved (Ackermann *et al.*, 2009; Mitrea *et al.*, 2013).

Keeping in mind the critical findings from the previous studies, we performed transcriptome profiling of both the *in vitro* cell line as well as *in vivo* xenograft tumor samples resulting in thousands of DEGs. These genes were segregated based on their

clustering in various biological processes (Table S1, S3 and S5). In line with our primary goal, we filtered the processes on the basis of relevance in cancer and ended up with five DEGs that were common in cell line data sets (A-D) namely EHBP, SLF2, DAP3, IP6K2 and RUNX1 as listed in Table S2. Eps15 homology domain binding protein is encoded by the EHBP1 gene and this protein has been shown to portray a role in actin reorganization and endocytic trafficking (Guilherme *et al.*, 2004). Polymorphism in this gene at a single nucleotide level has been reported to cause prostate cancer (Sun *et al.*, 2011). SLF2 is a DNA damage response pathway gene and functions by regulating genomic stability by post-replication repair of damaged DNA (Räschle *et al.*, 2015). DAP3 has been shown to be involved in mediating interferon (IFN)-γ induced cell death in addition to its role in organelle biogenesis as well as maintenance and mitochondrial translation (Kissil *et al.*, 1995; Tang *et al.*, 2009). DAP3 has been characterised by its pro-apoptotic function as a prognostic factor in gastric cancer (Jia *et al.*, 2014). The protein encoded by IP6K2 gene has been shown to affect growth suppression and apoptotic action of IFN-β in physiologic regulation of apoptosis in ovarian cancers (Morrison *et al.*, 2002). It has been reported that deletion of IP6K2 leads to HNSCC predisposition (Morrison *et al.*, 2009). Lastly, the protein encoded by RUNX1 has been shown to be involved in the process of normal haematopoiesis with chromosomal translocations related to this gene showing association with leukemias (Sood *et al.*, 2017). Further, it has been reported that RUNX1 depletion in human HNSCC cells caused growth arrest (Taniuchi *et al.*, 2012). Collectively, it is quite evident that all these MK2-regulated genes are playing vital role in tumor pathogenesis, hence showing consistency with our finding of their involvement in HNSCC pathogenesis. Therefore, they can be further explored as candidates for the development of novel biomarkers and utilized as potential therapeutic targets in HNSCC management.

Additionally, gene regulatory networks in the transcriptome profiling provided information of the various biological processes regulated by these candidate genes. It

supplied us with a wealth of information which can be further utilized to study the pathogenesis of HNSCC in detail especially in the background of MK2-knockdown and varied tumor microenvironment (normoxia/hypoxia). Figure 3a is a representation of gene regulatory network showing the role of MK2 in regulation of mRNA stability in various experimental dataset (F vs E dataset). The figure clearly demonstrates that MK2 regulates transcript stability *via* RBP-mediated regulation with HuR (ELAVL1) and TTP (ZFP36) playing integral part. Similarly, Figure 3b portraying the regulatory network that represents MAPK signaling cluster of the selected MK2 pathway genes (p38, MK2, AUF1, TTP, CUGBP1, CEBPδ, HuR, MKP-1, p27, TNF-α and VEGF) in the transcriptome profiling data of the *in vitro* HNSCC cell line dataset (B vs A, normoxic microenvironment) indicates VEGF and TNF-α down regulation. Interestingly, these transcriptomic profiling results are in complete consonance with our recently published findings and hence succeed to potentiate and validate the fact that MK2-knockdown destabilized VEGF and TNF-α in normoxia *via* RBP-mediated interaction.

Transcriptome analysis techniques are being commonly utilized in endeavors to decipher various molecular mechanisms of tumorigenesis and to fetch out novel prognostic and therapeutic markers (Bergers *et al.*, 2008; Trapnell *et al.*, 2010). In this study, we aimed to assess the MK2-regulated candidate genes playing prominent role in HNSCC pathogenesis. To accomplish this, we filtered out the DEGs based on the presence of adenine uridine rich elements (ARE)-regions in their 3'-UTRs where RBPs could bind and modulate their function *via* MK2-mediated regulation. This provided us with a list of 34 genes in A-D datasets while 48 genes in the xenograft dataset (listed in Table S6 and S7). The data was further validated using the nCounter gene expression assay system (NanoString Technologies, Inc.).

nCounter gene expression assays enable the digital quantification and single molecule imaging of multiple target RNA molecules using multicolor-coded molecular

barcodes (Figure S2 and S15). This system provides discrete and accurate counts of RNA transcripts at a high level of sensitivity and precision (Brumbaugh *et al.*, 2011). We chose four commonly used reference genes *viz.* ABCF1, GAPDH, POLR2A, and RPL19, in nCounter data analysis because of their baseline expression in HNSCC tumorigenesis as well as in MK2-knockdown conditions, as listed in Table S9 (Lallemant *et al.*, 2009; Palve *et al.*, 2018).

Gene expression assays are independent of any enzymatic reactions or amplification protocols and have no reliance on degree of fluorescence intensity to determine target abundance. As a result of these characteristics, and also the highly automated nature of bar-coded sample processing, these assays result in highly precise and reproducible outcomes. On an average, we obtained around 52% matching score of transcriptome profiling data with nCounter gene expression assay-based validation which is considered a good percentage match. Filtering of the DEGs in the matched data revealed a list of 12 genes (6 upregulated and 6 downregulated in various experimental datasets) that were common in our comprehensive nCounter system-based validation of transcriptomic profiling (Table 4). Intriguingly, these genes portray crucial roles in processes like apoptosis (CLK2, MELK, MUC4), cell cycle regulation (CLK2, MELK) and transcription regulation (BRD2, H2AFY, SAMD4B, ZNF662) suggesting that they could be potentially used as molecular candidates for further investigations in regard to design of molecular markers and therapeutics for HNSCC management. Hence, transcriptome analysis followed by nCounter assay-based validation of our data has provided us with valuable findings that can aid in extending the field of HNSCC research.

Six candidate genes (ZNF662, BMP7, CREB3L1, IGFBP2, MUC4, PRKAR2B) that showed statistically significant up/downregulation in F vs E dataset were further analyzed in tumor sections from xenografted animals using IHC to evaluate the expression levels of these proteins. Interestingly, our IHC findings confirmed the validity of our *in-silico* findings.

Consistent with *in silico* findings the expression levels of three genes (IGFB2, MUC4, PRKARB2) were found to be upregulated, suggesting the roles of these MK2-dependent candidate genes in pathogenesis and progression of HNSCC. Insulin like Growth Factor Binding Protein 2 (IGFBP2) has been shown as a growth promoter gene in several tumors and considered as a central hub of oncogenic signaling network governing transcriptional regulation and promoting epithelial to mesenchymal transition, invasion, angiogenesis and metastasis (Lindström 2019). Mucin 4 (MUC4) serves as a major constituent of mucus secreted by epithelial cells and found overexpressed in a variety of cancers like papillary thyroid carcinomas. It is known for promotion of tumor growth, proliferation and migration (Yu *et al.,* 2020). Recent insights have been made into the transcriptional regulation of protein kinase cAMP-dependent type II regulatory subunit beta (PRKAR2B) by miRNAs and X-Box binding protein 1 leading to a better understanding of PRKAR2B-driven prostate cancer progression (Xia *et al.,* 2020). Consistent to our findings these genes have been suggested to be prognostic indicators and therapeutic targets in various cancers including HNSCC (Park *et al.*, 2015; Zhao *et al.*, 2018; Lindström 2019; Xia *et al.,* 2020; Yu *et al.*, 2020). It is worth mentioning that all of these genes are MK2-regulated and play specific roles in pathogenesis. In conclusion, present study vindicates involvement of more MK2-dependent genes as critically important in regulating HNSCC progression by modulating the transcript stability of pathogenesis-related genes *via* RBP-mediated regulation. This study has made it possible to filter down from thousands of DEGs to a few potential candidate genes using a comprehensive transcriptomic and *in vitro* approach. In an attempt to delve deeper into the mechanistic insights and to decipher the molecular markers responsible for MK2-mediated changes in HNSCC pathogenesis, the role of these potential therapeutic targets warrants further detailed investigation.

Conclusions

In conclusion, present study vindicates involvement of more MK2-dependent genes as critically important in regulating HNSCC by modulating the transcript stability of pathogenesis-related genes possibly *via* RBP-mediated regulation and these results are in perfect consonance and augmentation with our recent findings (Soni[2] *et al.*, 2019). Comprehensively, few crucial MK2-regulated putative candidate genes were identified in this study and their plausible mode of action in HNSCC pathogenesis was elucidated, which could be further explained as possible targets in the pathway for HNSCC management (Figure 6). This study has made it possible to filter down from thousands of DEGs to a few potential candidate genes using a comprehensive transcriptomic and *in vitro* approach. In an attempt to delve deeper into the mechanistic insights and to decipher the molecular markers responsible for MK2-mediated changes in HNSCC pathogenesis, the role of these potential therapeutic targets warrants further detailed investigation.

Methods

Culturing and maintenance of cell lines

Homo sapiens tongue squamous cell carcinoma cell line CAL27 (CRL-2095™, ATCC, USA) and *Homo sapiens* palatal mesenchyme cell line HEPM (CRL-1486™, ATCC) were grown in specific media supplemented with 10% fetal bovine serum and 1% antibiotic-antimycotic (Gibco, USA). The cells were cultured in normal conditions (37°C, 5% CO_2 incubator with 95% humidity) and were free from any kind of contamination. Further, MK2-specific short hairpin RNA-green fluorescent protein (shRNA-GFP) constructs were used to stably knockdown MK2 from cultured CAL27 cells, thus forming CAL27-$MK2_{KD}$ cells (as previously described in Soni[2] *et al.*, 2019). For hypoxia experimentation, the cultured cells plated in petriplates were incubated for 48 hours in 0.5% O_2 at 37°C in a hypoxia chamber (Bactrox, Shel-Lab, USA).

Xenograft mice model generation

In order to mimic the tumor microenvironment more closely to the human tumors, a biologically relevant heterotopic xenograft model of HNSCC was developed in non-obese diabetic/severe combined immune deficiency (NOD/SCID) mice. The immunocompromised mice were randomly assigned into control (CAL27-$MK2_{WT}$) and experimental (CAL27-$MK2_{KD}$) group based on specific cell-type injected (as previously described in Soni[2] *et al.*, 2019). Briefly, for xenograft generation, one million cultured cells suspended in 100µl of 1x phosphate buffered saline were injected subcutaneously in the right flank of mice. Seven weeks after graft inoculation, the mice were euthanized by CO_2 asphyxiation; tumors were aseptically excised, weighed and used for tissue embedding/RNA isolation.

RNA extraction and sample preparation

CAL27 cells (both WT and KD) cultured in petriplates and grown in both normoxic and hypoxic environment as well as tumors resected from xenografted mice were used for isolation of total cellular RNA using RNeasy Mini kit (Qiagen) following manufacturer recommended protocol (sample details are provided in Figure 1a and Table 1). Consequently, the quality as well as concentration of all the RNA samples was determined by employing NanoDrop 2000C spectrophotometer (Thermo Fisher Scientific, USA) and Bioanalyzer (Agilent 2100, Agilent Technologies). RNA samples having RIN>5 was used for cDNA library preparation. For each sample, RNA was isolated from at least three individual plates/tumors and was pooled for library construction and further analysis.

cDNA library preparation and sequencing

Total RNA (5µg) from each sample was used to isolate poly-A mRNA and to prepare cDNA library using TruSeq mRNA sample preparation kit v2 (Illumina Inc.). Each sample was tagged with a unique TruSeq index tag in order to prepare multiplexed libraries. Six paired-

end adapters with unique six base index sequences were used for the library preparation. The index tag permits accurate differentiation between each sample. The quantification of prepared libraries was performed on Qubit fluorometer using Qubit dsDNA BR assay kit (Life Technologies, USA) and size of the libraries was checked on Bioanalyzer DNA 1000 series II chip (Agilent Technologies). The flow chart of the sequential steps involved in TruSeq library preparation is given in Figure S15. The insert size and purity of libraries were then validated on Bioanalyzer DNA 1000 series II chip (Agilent Technologies). All the libraries (six in total, 4 from cell line model and 2 from animal model) had an average insert size of 210bp and were pooled by taking 10µl from each library. Final pool was loaded in one lane of a S2 flow cell using NovaSeq XP protocol. Cluster amplification and sequencing data generation were performed on NovaSeq 6000 system (Illumina Inc.) using 2x100 paired-end cycles (Figure 1a). Raw data quality control was performed using NGSQC tool kit v2.3 using default parameters (Patel *et al.*, 2012).

Reference based assembly and homology search

Raw FASTQ files with low-quality reads of sequencing data were filtered to obtain high quality reads. High quality filtered data was aligned with the reference genome (genome reference consortium human build 38 patch release 12, GRCh38.p12). Kallisto pipeline was used for alignment and identification of transcript coding regions followed by quantitation and annotation using default parameters (Huang *et al.*, 2009). Furthermore, we performed the removal of multi-mapped reads and finally converted them to read counts for annotated genes. Figure 1b is a flowchart representation of the various steps involved in the sequencing data analysis.

Annotation, differential gene expression and pathway analysis

Expression of the commonly found transcripts in all the samples was analyzed based on their FPKM values. Transcripts were given a score for their expression by Cufflinks-based maximum likelihood method and values of FPKM≥0.1 were considered significant for downstream analysis. Transcripts uniquely expressed in each sample were considered specific and were analyzed separately. False discovery rate (FDR) was employed to correct the statistical significance of the p-values for multiple tests. DEGs in multiple combinations were identified by DESeq analysis pipeline using a fold change threshold of absolute log2 fold change (FC)≥2 and a statistically significant Student's t-test p-value threshold adjusted for FDR<0.001. Hence, transcripts having FC<-2 were considered as downregulated while those with FC>2 were deemed upregulated. Statistically significant enriched functional classes with a p-value adjusted for FDR>0.05 derived using the hypergeometric distribution test corresponding to DEGs were determined using Student's t-test with Benjamini Hocheberg FDR test.

Unsupervised hierarchical clustering of DEGs was performed using Cluster 3.0 and visualized using Java TreeView (Huang *et al.*, 2009) (Figure 1b). Gene ontologies and pathways that harbour expressed transcripts were identified using DAVID functional annotation tool (http://david.abcc.ncifcrf.gov/home.jsp) (Dennis *et al.*, 2003; Sherman *et al.*, 2009). For DEGs, heat maps and volcano plots were generated using 'gplots' and 'heat map' packages. Gene ontology (GO) as well as KEGG pathway analysis was accomplished for the assembled transcripts with reference to UniProt database. The data of total DEGs (upregulated and downregulated) was explored further using Cytoscape v3.5.0 (http://www.cytoscape.org/) to better understand the gene regulatory networks and for mapping of the results (Montojo *et al.*, 2010). In figures depicting gene regulatory networks, gene nodes are sized according to their p-values and colored according to their FC, where red

shows upregulation, green shows downregulation and yellow shows baseline expression; processes are showed in rectangular boxes and colored in blue.

nCounter gene expression assay

In order to validate the leads obtained from transcriptomic profiling, we utilized custom designed molecular barcodes (NanoString Technologies, Inc.) for single molecule imaging, thereby making it possible to detect and count hundreds of different transcripts in one reaction (Figure 1a and S13). RNA quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies) (Figure 1a and S1). Gene expression was analyzed on the nCounter system (NanoString Technologies, Inc.) following manufacturer's recommendations. Briefly, the custom synthesized probes were hybridized overnight to target RNA followed by washing away of excess probes, immobilization of CodeSet/RNA complexes in the nCounter cartridge and finally data collection on the nCounter system. Gene expression levels were measured in triplicate for total RNA from the cell line and xenografted tumor samples, normalized to the four HKG and analyzed using the nSolver software (NanoString Technologies, Inc.). Each nCounter assay contained synthetic spike-in controls in the preparatory mix. These controls are necessary to allow correction of sample-to-sample variation arising due to common experimental errors such as differences in amount of input transcripts or reagents (Brumbaugh *et al.*, 2011). The counts were normalized with the positive controls and averaged for the samples of each mRNA type. Normalization involved spiked-in positive and negative control probes for background correction in addition to four reference or housekeeping genes. Data analysis was performed on nSolver 3.0 analysis and advanced analysis software (NanoString Technologies, Inc.).

Immunohistochemical analysis

The levels of expression and activation status of specific proteins in xenografted tumor sections were analyzed using IHC to validate the overall findings of the transcriptomic and

the nCounter gene expression analysis in an *in vivo* xenograft model. 5 μm thin sections were fixed on poly-l-lysine coated slides, followed by deparaffinization and rehydration. Antigen retrieval was performed using sodium citrate buffer (pH 6.0) followed by quenching of endogenous peroxidases using BLOXALL™ blocking solution (Vector Laboratories, Inc., USA). Further, incubation of the sections with 2.5% normal horse serum blocked the exposed sites. Sections were incubated with appropriately diluted specific primary antibody overnight at 4°C followed by horseradish peroxidase-conjugated secondary antibody for 1 hour at room temperature. Rinsed sections were then incubated with 3,3'-diaminobenzidine substrate and Mayer's hematoxylin served as a counterstain. Five field views were obtained from each slide having tissue sections for designated antibody and used for analysis. Observations and imaging of the sections were carried out by a pathologist in a blinded fashion using BX53 bright field microscope (Olympus Corporation, Japan). Protein expression was quantified using ImageJ software 1.8.0 (https://imagej.nih.gov/ij/).

Availability of Data

The generated datasets (raw reads) from NovaSeq 6000 have been deposited to the National Centre for Biotechnology Information-Sequence Read Archive (NCBI-SRA) repository (https://www.ncbi.nlm.nih.gov/sra). The SRA BioProject accession number for the submitted bio projects are PRJNA646850 and PRJNA646851.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors acknowledge CSIR for financial support (CSIR-IHBT; Project MLP0204).

Authors' contributions

YSP conceptualized the work and designed the manuscript. SS and PA performed experiments and data analysis. MKS helped in RNA sequencing run on NovSeq and data analysis, VP performed IHC imaging and analysis, NVT helped in the xenograft model generation. SS, PA, VP and YSP wrote and edited the manuscript. All the authors read and approved the final version of the manuscript

Acknowledgements

The authors would like to thank the Director, CSIR-IHBT, Palampur for his consistent support and encouragement. SS is immensely thankful to CSIR and PA to DST for providing Ph.D. fellowship and Academy of Scientific and Innovative Research (AcSIR), Ghaziabad for Ph.D. registration. The authors would like to thank Bionivid Technology Private Limited, Bengaluru for help in transcriptome data analysis. IHBT communication number of this manuscript is 4682.

References

[1] Ackermann M, Strimmer K. A general modular framework for geneset enrichment analysis. BMC bioinformatics. 2009 Dec 1;10(1):47.

[2] Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, García Girón C, Hourlier T, Howe K. The Ensembl gene annotation system. Database. 2016 Jan 1;2016.

[3] Bergers G, Hanahan D. Modes of resistance to anti-angiogenic therapy. Nature Reviews Cancer. 2008 Aug;8(8).

[4] Bhatia S, Frangioni JV, Hoffman RM, Iafrate AJ, Polyak K. The challenges posed by cancer heterogeneity. Nature biotechnology. 2012 Jul;30(7).

[5] Bossi P, Bergamini C, Siano M, Rocca MC, Sponghini AP, Favales F, Giannoccaro M, Marchesi E, Cortelazzi B, Perrone F, Pilotti S. Functional genomics uncover the biology behind the responsiveness of head and neck squamous cell cancer patients to cetuximab. Clinical Cancer Research. 2016 Aug 1;22(15).

[6] Brumbaugh CD, Kim HJ, Giovacchini M, Pourmand N. NanoStriDE: normalization and differential expression analysis of NanoString nCounter data. BMC bioinformatics. 2011 Dec;12(1).

[7] Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nature Reviews Genetics. 2016 May;17(5).

[8] Campbell JD, Yau C, Bowlby R, Liu Y, Brennan K, Fan H, Taylor AM, Wang C, Walter V, Akbani R, Byers LA. Genomic, pathway network, and immunologic features distinguishing squamous carcinomas. Cell reports. 2018 Apr 3;23(1).

[9] Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015 Jan;517(7536).

[10] Chau NG, Li YY, Jo VY, Rabinowits G, Lorch JH, Tishler RB, Margalit DN, Schoenfeld JD, Annino DJ, Goguen LA, Thomas T. Incorporation of next-generation sequencing into routine clinical care to direct treatment of head and neck squamous cell carcinoma. Clinical cancer research. 2016 Jun 15;22(12).

[11] Cheng H, Yang X, Si H, Saleh AD, Xiao W, Coupar J, Gollin SM, Ferris RL, Issaeva N, Yarbrough WG, Prince ME. Genomic and transcriptomic characterization links cell lines with aggressive head and neck cancers. Cell reports. 2018 Oct 30;25(5).

[12] Cieślik M, Chinnaiyan AM. Cancer transcriptome profiling at the juncture of clinical translation. Nature Reviews Genetics. 2018 Feb;19(2).

[13] Denaro N, Russi EG, Adamo V, Merlano MC. State-of-the-art and emerging treatment options in the management of head and neck cancer: news from 2013. Oncology. 2014;86(4).

[14] Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. Genome biology. 2003 Sep;4(9).

[15] Fang XN, Yin M, Li H, Liang C, Xu C, Yang GW, Zhang HX. Comprehensive analysis of competitive endogenous RNAs network associated with head and neck squamous cell carcinoma. Scientific reports. 2018 Jul 12;8(1).

[16] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer. 2015 Mar 1;136(5).

[17] Fiore M, Forli S, Manetti F. Targeting mitogen-activated protein kinase-activated protein kinase 2 (MAPKAPK2, MK2): medicinal chemistry efforts to lead small molecule inhibitors to clinical trials. Journal of medicinal chemistry. 2016 Apr 28;59(8).

[18] Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, editors. Bioinformatics and computational biology solutions using R and Bioconductor. Springer Science & Business Media; 2006 Jan 27.

[19] Guilherme A, Soriano NA, Bose S, Holik J, Bose A, Pomerleau DP, Furcinitti P, Leszyk J, Corvera S, Czech MP. EHD2 and the novel EH domain binding protein EHBP1 couple endocytosis to the actin cytoskeleton. Journal of Biological Chemistry. 2004 Mar 12;279(11).

[20] Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research. 2009 Jan 1;37(1).

[21] Jamali Z, Aminabadi NA, Attaran R, Pournagiazar F, Oskouei SG, Ahmadpour F. MicroRNAs as prognostic molecular signatures in human head and neck squamous cell carcinoma: a systematic review and meta-analysis. Oral oncology. 2015 Apr 1;51(4).

[22] Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA: a cancer journal for clinicians. 2011 Mar;61(2).

[23] Jia Y, Ye L, Ji K, Zhang L, Hargest R, Ji J, Jiang WG. Death-associated protein-3, DAP-3, correlates with preoperative chemotherapy effectiveness and prognosis of gastric cancer patients following perioperative chemotherapy and radical gastrectomy. British journal of cancer. 2014 Jan;110(2).

[24] Kissil JL, Deiss LP, Bayewitch M, Raveh T, Khaspekov G, Kimchi A. Isolation of DAP3, a novel mediator of interferon-γ-induced cell death. Journal of Biological Chemistry. 1995 Nov 17;270(46).

[25] Lallemant B, Evrard A, Combescure C, Chapuis H, Chambon G, Raynal C, Reynaud C, Sabra O, Joubert D, Hollande F, Lallemant JG. Reference gene selection for head and neck squamous cell carcinoma gene expression studies. BMC molecular biology. 2009 Dec;10(1).

[26] Lasa M, Mahtani KR, Finch A, Brewer G, Saklatvala J, Clark AR. Regulation of cyclooxygenase 2 mRNA stability by the mitogen-activated protein kinase p38 signaling cascade. Molecular and cellular biology. 2000 Jun 15;20(12).

[27] Lindström MS. Expanding the scope of candidate prognostic marker IGFBP2 in glioblastoma. Bioscience reports. 2019 Jul;39(7):BSR20190770.

[28] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014 Dec 1;15(12).

[29] Martin D, Abba MC, Molinolo AA, Vitale-Cross L, Wang Z, Zaida M, Delic NC, Samuels Y, Lyons JG, Gutkind JS. The head and neck cancer cell oncogenome: a platform for the development of precision molecular therapies. Oncotarget. 2014 Oct;5(19).

[30] Mitrea C, Taghavi Z, Bokanizad B, Hanoudi S, Tagett R, Donato M, Voichita C, Draghici S. Methods and approaches in the topology-based analysis of biological pathways. Frontiers in physiology. 2013 Oct 10;(4).

[31] Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, Morris Q, Bader GD. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics. 2010 Nov 15;26(22).

[32] Morrison BH, Bauer JA, Hu J, Grane RW, Ozdemir AM, Chawla-Sarkar M, Gong B, Almasan A, Kalvakolanu DV, Lindner DJ. Inositol hexakisphosphate kinase 2 sensitizes ovarian carcinoma cells to multiple cancer therapeutics. Oncogene. 2002 Mar;21(12).

[33] Morrison BH, Haney R, Lamarre E, Drazba J, Prestwich GD, Lindner DJ. Gene deletion of inositol hexakisphosphate kinase 2 predisposes to aerodigestive tract carcinoma. Oncogene. 2009 Jun;28(25).

[34] Narayanan SP, Singh S, Gupta A, Yadav S, Singh SR, Shukla S. Integrated genomic analyses identify KDM1A's role in cell proliferation via modulating E2F signaling activity and associate with poor clinical outcome in oral cancer. Cancer letters. 2015 Oct 28;367(2).

[35] Nisa L, Barras D, Medová M, Aebersold DM, Medo M, Poliaková M, Koch J, Bojaxhiu B, Eliçin O, Dettmer MS, Angelino P. Comprehensive genomic profiling of patient-matched head and neck cancer cells: a preclinical pipeline for metastatic and recurrent disease. Molecular cancer research. 2018 Dec 1;16(12).

[36] Palve V, Pareek M, Krishnan NM, Siddappa G, Suresh A, Kuriakose MA, Panda B. A minimal set of internal control genes for gene expression studies in head and neck squamous cell carcinoma. PeerJ. 2018 Aug 14;6.

[37] Park J, Lee J, Choi C. Evaluation of drug-targetable genes by defining modes of abnormality in gene expression. Scientific reports. 2015 Sep 4;(5).

[38] Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PloS one. 2012 Feb 1;7(2).

[39] Räschle M, Smeenk G, Hansen RK, Temu T, Oka Y, Hein MY, Nagaraj N, Long DT, Walter JC, Hofmann K, Storchova Z. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. Science. 2015 May 1;348(6234).

[40] Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009 Jan;4(1).

[41] Soni[1] S, Anand P, Padwad YS. MAPKAPK2: the master regulator of RNA-binding proteins modulates transcript stability and tumor progression. Journal of Experimental & Clinical Cancer Research. 2019 Dec 1;38(1).

[42] Soni[2] S, Saroch MK, Chander B, Tirpude NV, Padwad YS. MAPKAPK2 plays a crucial role in the progression of head and neck squamous cell carcinoma by regulating transcript stability. Journal of Experimental & Clinical Cancer Research. 2019 Dec;38(1).

[43] Sood R, Kamikubo Y, Liu P. Role of RUNX1 in hematological malignancies. Blood. 2017 Apr 13;129(15).

[44] Spataro N, Rodríguez JA, Navarro A, Bosch E. Properties of human disease genes and the role of genes linked to Mendelian disorders in complex disease aetiology. Human molecular genetics. 2017 Feb 1;26(3).

[45] Sun J, Kader AK, Hsu FC, Kim ST, Zhu Y, Turner AR, Jin T, Zhang Z, Adolfsson J, Wiklund F, Zheng SL. Inherited genetic markers discovered to date are able to identify a significant number of men at considerably elevated risk for prostate cancer. The Prostate. 2011 Mar 1;71(4).

[46] Tang T, Zheng B, Chen SH, Murphy AN, Kudlicka K, Zhou H, Farquhar MG. hNOA1 interacts with complex I and DAP3 and regulates mitochondrial respiration and apoptosis. Journal of Biological Chemistry. 2009 Feb 20;284(8).

[47] Taniuchi I, Osato M, Ito Y. Runx1: no longer just for leukemia. The EMBO journal. 2012 Nov 5;31(21).

[48] Thomas GR, Nadiminti H, Regalado J. Molecular predictors of clinical outcome in patients with head and neck squamous cell carcinoma. International journal of experimental pathology. 2005 Dec;86(6).

[49] Thomas MP, Potter BV. The enzymes of human diphosphoinositol polyphosphate metabolism. The FEBS journal. 2014 Jan 1;281(1).

[50] Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature biotechnology. 2010 May;28(5).

[51] Van Waes C, Musbahi O. Genomics and advances towards precision medicine for head and neck squamous cell carcinoma. Laryngoscope investigative otolaryngology. 2017 Oct;2(5).

[52] Venigalla RK, Turner M. RNA-binding proteins as a point of convergence of the PI3K and p38 MAPK pathways. Frontiers in immunology. 2012 Dec 26;3.

[53] Wazir U, Orakzai MM, Khanzada ZS, Jiang WG, Sharma AK, Kasem A, Mokbel K. The role of death-associated protein 3 in apoptosis, anoikis and human cancer. Cancer Cell International. 2015 Dec;15(1).

[54] Weinberg RA. Coming full circle—from endless complexity to simplicity and back again. Cell. 2014 Mar 27;157(1).

[55] Woolgar JA, Triantafyllou A. Squamous cell carcinoma and precursor lesions: clinical pathology. Periodontology 2000. 2011 Oct;57(1).

[56] Xia L, Han Q, Chi C, Zhu Y, Pan J, Dong B, Huang Y, Xia W, Xue W, Sha J. Transcriptional regulation of PRKAR2B by miR-200b-3p/200c-3p and XBP1 in human prostate cancer. Biomedicine & Pharmacotherapy. 2020 Apr 1;124.

[57] Yu J, Xu L, Yan J, Yu J, Wu X, Dai J, Guo J, Kong Y. MUC4 isoforms expression profiling and prognosis value in Chinese melanoma patients. Clinical and Experimental Medicine. 2020 Mar 14.

[58] Zhao C, Zou H, Zhang J, Wang J, Liu H. An integrated methylation and gene expression microarray analysis reveals significant prognostic biomarkers in oral squamous cell carcinoma. Oncology reports. 2018 Nov 1;40(5).

List of Tables

Table 1: Sample description and comparison details.

Tabular representation of the codes and details of all the samples used for transcriptome analysis and details of the various experimental datasets (comparisons) used in the study.

S. No.	Analysis Code	Sample Type	Sample Description
1.	A	Cultured Cell Line (*in vitro*)	Normal CAL27 cells (Normoxia)
2.	B		MK2-knockdown CAL27 cells (Normoxia)
3.	C		Normal CAL27 cells (Hypoxia)
4.	D		MK2-knockdown CAL27 cells (Hypoxia)
5.	E	Dissected Xenografts (*in vivo*)	Normal CAL27 cells grafted
6.	F		MK2-knockdown CAL27 cells grafted

Comparison/ Dataset Code	Comparison/Dataset Detail
B vs A **or** KD(N) vs N(N)	MK2-knockdown CAL27 cells (Normoxia) vs Normal CAL27 cells (Normoxia)
D vs C **or** KD(H) vs N(H)	MK2-knockdown CAL27 cells (Hypoxia) vs Normal CAL27 cells (Hypoxia)
D vs B **or** KD(H) vs KD(N)	MK2-knockdown CAL27 cells (Hypoxia) vs MK2-knockdown CAL27 cells (Normoxia)
C vs A **or** N(H) vs N(N)	Normal CAL27 cells (Hypoxia) vs Normal CAL27 cells (Normoxia)
F vs E **or** KD(X) vs N(X)	MK2-knockdown CAL27 cells grafted vs Normal CAL27 cells grafted

Table 2: Summary of the comparison

Tabular representation of the summary of the comparison between the data obtained from transcriptome profiling and nCounter gene expression assay analysis for all the experimental datasets.

Combination	No. of genes analyzed	No. and % of genes matching with transcriptome analysis	Upregulated genes	Downregulated genes
B vs A	39	**24** (61.6 %)	**1** (BRD2)	**1** (CLK2)
D vs C	39	**17** (43.6%)	-	-
D vs B	39	**19** (48.7%)	-	**1** (SAMD4B)
C vs A	39	**20** (51.3%)	-	**2** (H2AFY, MELK)
F vs E	48	**26** (54.2%)	**5** (BMP7, CREB3L1, IGFBP2, MUC4, PRKAR2B)	**2** (CDSN, ZNF662)

Table 3: Tabular representation of the list of 39 genes in *in vitro* HNSCC cell line model (C vs A dataset) showing the match of the average log2 fold change values in the last two columns in both transcriptome profiling and the NanoString gene expression assay. The average log2 fold change values are colored according to the change in gene expression (red indicates upregulation while green indicates downregulation). The gene highlighted in yellow is the matched candidate DEG that shows statistically significant change in expression.

S.No.	Gene Symbol	Entrez_GeneID	Gene Name	C_VS_A Transcriptome	C_VS_A Nanostring
1	ADNP2	22850	ADNP homeobox 2(ADNP2)	0.26	-1.1
2	BAZ2B	29994	bromodomain adjacent to zinc finger domain 2B(BAZ2B)	-0.28	-1.12
3	BRD2	6046	bromodomain containing 2(BRD2)	-	1.71
4	BRD4	23476	bromodomain containing 4(BRD4)	0.39	1.58
5	CALD1	800	caldesmon 1(CALD1)	-0.24	-1.31
6	CAMKK2	10645	calcium/calmodulin dependent protein kinase kinase 2(CAMKK2)	0.78	-1.74
7	CAMTA2	23125	calmodulin binding transcription activator 2(CAMTA2)	0.24	1.07
8	CLK2	1196	CDC like kinase 2(CLK2)	-2.71	-1.19
9	CTBP2	1488	C-terminal binding protein 2(CTBP2)	-0.14	-1.15
10	DAP3	7818	death associated protein 3(DAP3)	-1.49	-1.92
11	DICER1	23405	dicer 1, ribonuclease III(DICER1)	2.57	-2.21
12	EHBP1	23301	EH domain binding protein 1(EHBP1)	-3.28	-1.34
13	ERF	2077	ETS2 repressor factor(ERF)	-0.04	1.21
14	FER	2241	FER tyrosine kinase(FER)	-0.75	1.09
15	FOXJ3	22887	forkhead box J3(FOXJ3)	1.18	1.2
16	H2AFY	9555	H2A histone family member Y(H2AFY)	-4.65	-1.95
17	IP6K2	51447	inositol hexakisphosphate kinase 2(IP6K2)	0.55	1.71
18	IRAK1	3654	interleukin 1 receptor associated kinase 1(IRAK1)	-8.54	-1.93
19	KDM5C	8242	8242 [Description: lysine demethylase 5C] [Gene Type: protein-coding]	0.25	1.4
20	LATS1	9113	large tumor suppressor kinase 1(LATS1)	-0.66	-1.27
21	MAP4K4	9448	mitogen-activated protein kinase kinase kinase kinase 4(MAP4K4)	4.16	-1.36
22	MELK	9833	maternal embryonic leucine zipper kinase(MELK)	-3.68	-4.59
23	MINK1	50488	misshapen like kinase 1(MINK1)	-0.24	1.37
24	NCOA6	23054	nuclear receptor coactivator 6(NCOA6)	-	1.11
25	NCOR1	9611	nuclear receptor corepressor 1(NCOR1)	-0.30	-1.05
26	NEK9	91754	NIMA related kinase 9(NEK9)	5.34	-1.1
27	NR3C1	2908	nuclear receptor subfamily 3 group C member 1(NR3C1)	-0.15	1.28
28	PAK4	10298	p21 (RAC1) activated kinase 4(PAK4)	-	1.03
29	PASK	23178	PAS domain containing serine/threonine kinase(PASK)	-	-1.99
30	PBRM1	55193	55193 [Description: polybromo 1] [Gene Type: protein-coding]	-	-1.37
31	PPP1R12A	4659	protein phosphatase 1 regulatory subunit 12A(PPP1R12A)	1.00	1.1
32	RUNX1	861	runt related transcription factor 1(RUNX1)	0.47	-1.03
33	SAMD4B	55095	sterile alpha motif domain containing 4B(SAMD4B)	-2.54	-1.52
34	SLF2	55719	SMC5-SMC6 complex localization factor 2(SLF2)	-0.81	-1.52
35	SNAPC4	6621	small nuclear RNA activating complex polypeptide 4(SNAPC4)	7.41	1.09
36	SP3	6670	Sp3 transcription factor(SP3)	0.14	-1.02
37	TAF1	6872	TATA-box binding protein associated factor 1(TAF1)	-0.76	-1.82
38	UTRN	7402	utrophin(UTRN)	-9.26	-1
39	ZNF189	7743	zinc finger protein 189(ZNF189)	-1.01	1.3

Table 4: Tabular representation of the list of 48 genes in the *in vivo* heterotopic HNSCC xenograft experimental dataset (F vs E dataset) showing the match of the average log2 fold change values in the last two columns in both transcriptome profiling and the NanoString gene expression assay. The average log2 fold change values are colored according to the change in gene expression (red indicates upregulation while green indicates downregulation). The genes highlighted in yellow are the matched candidate DEGs that show statistically significant change in expression.

S.No.	Gene Symbol	Entrez_GeneID	Gene Name	F_VS_ETranscriptome	F vs E Nanostring
1	APBB2	323	amyloid beta precursor protein binding family B member 2(APBB2)	-3.74	-1.16
2	ATP13A2	23400	ATPase 13A2(ATP13A2)	3.75	1.45
3	BMP7	655	bone morphogenetic protein 7(BMP7)	5.95	47.09
4	CDC25B	994	cell division cycle 25B(CDC25B)	0.73	1.18
5	CDSN	1041	corneodesmosin(CDSN)	-2.35	1.72
6	CPEB2	132864	cytoplasmic polyadenylation element binding protein 2(CPEB2)	-3.07	-1.14
7	CREB3L1	90993	cAMP responsive element binding protein 3 like 1(CREB3L1)	3.46	15.34
8	DDR1	780	discoidin domain receptor tyrosine kinase 1(DDR1)	6.31	-1.18
9	DPYSL3	1809	dihydropyrimidinase like 3(DPYSL3)	-	16.6
10	DST	667	dystonin(DST)	-6.97	1.09
11	EIF4E	1977	eukaryotic translation initiation factor 4E(EIF4E)	3.08	1.1
12	EZH1	2145	enhancer of zeste 1 polycomb repressive complex 2 subunit(EZH1)	-	1.37
13	FOXO3	2309	forkhead box O3(FOXO3)	-7.23	-1.5
14	FREM1	158326	FRAS1 related extracellular matrix 1(FREM1)	-5.95	1.51
15	GUK1	2987	guanylate kinase 1(GUK1)	0.11	1.2
16	H2AFY	9555	H2A histone family member Y(H2AFY)	3.77	1.34
17	HNRNPD	3184	heterogeneous nuclear ribonucleoprotein D(HNRNPD)	-1.10	1.29
18	IGFBP2	3485	insulin like growth factor binding protein 2(IGFBP2)	3.86	12.21
19	ITPR1	3708	inositol 1,4,5-trisphosphate receptor type 1(ITPR1)	7.67	-1.11
20	JAK1	3716	Janus kinase 1(JAK1)	-0.08	-1
21	KMT2C	58508	lysine methyltransferase 2C(KMT2C)	8.64	1.1
22	LIMK1	3984	LIM domain kinase 1(LIMK1)	-4.37	1.26
23	LIMS1	3987	LIM zinc finger domain containing 1(LIMS1)	6.28	1.27
24	MKL2	57496	MKL1/myocardin like 2(MKL2)	2.40	1.14
25	MUC4	4585	mucin 4, cell surface associated(MUC4)	2.82	7.74
26	NDRG2	57447	57447 Description: NDRG family member 2] [Gene Type: protein-coding]	-	2.72
27	PCBP4	57060	poly(rC) binding protein 4(PCBP4)	-	1.07
28	PFKM	5213	5213 [Description: phosphofructokinase, muscle] [Gene Type: protein-coding]	1.00	2.62
29	PFKP	5214	phosphofructokinase, platelet(PFKP)	-5.36	-1.09
30	PI4KB	5298	phosphatidylinositol 4-kinase beta(PI4KB)	4.87	1.05
31	PPP1R12C	54776	protein phosphatase 1 regulatory subunit 12C(PPP1R12C)	-4.14	1.12
32	PPP2R1B	5519	protein phosphatase 2 scaffold subunit Abeta(PPP2R1B)	-5.42	1.03
33	PRKAR2B	5577	protein kinase cAMP-dependent type II regulatory subunit beta(PRKAR2B)	2.06	4.28
34	PTPN22	26191	protein tyrosine phosphatase, non-receptor type 22(PTPN22)	0.54	1.09
35	SEMA7A	8482	semaphorin 7A (John Milton Hagen blood group)(SEMA7A)	-	-1.74
36	SLC35B2	347734	solute carrier family 35 member B2(SLC35B2)	1.57	-1.01
37	SMAD3	4088	SMAD family member 3(SMAD3)	0.00	-1.29
38	SS18	6760	SS18, nBAF chromatin remodeling complex subunit(SS18)	-4.34	1.25
39	STK25	10494	serine/threonine kinase 25(STK25)	-	1.15
40	TGFBRAP1	9392	transforming growth factor beta receptor associated protein 1(TGFBRAP1)	3.62	1.22
41	TRAK1	22906	trafficking kinesin protein 1(TRAK1)	-	-1.26
42	UBE3A	7337	7337 [Description: ubiquitin protein ligase E3A] [Gene Type: protein-coding]	-0.06	1.14
43	ZBED1	9189	zinc finger BED-type containing 1(ZBED1)	-3.63	1.68
44	ZC3H13	23091	zinc finger CCCH-type containing 13(ZC3H13)	0.27	-1.03
45	ZNF131	7690	zinc finger protein 131(ZNF131)	-	1.26
46	ZNF48	197407	zinc finger protein 48(ZNF48)	-0.98	-3.04
47	ZNF544	27300	zinc finger protein 544(ZNF544)	-	1.16
48	ZNF662	389114	zinc finger protein 662(ZNF662)	-3.48	-3.44

Figure Legends

Fig. 1 Schematic pictorial illustration of the experimental and analysis workflow used for transcriptomic profiling.

a) Detailed experimental work plan and sample preparation scheme indicating the two distinct sample types used for the experimental analysis (cell lines and xenografted tumor).

b) Workflow of sequencing analysis showing sequential steps and various tools and pipelines employed for the transcriptome profiling.

c), d) A comparative bar diagram representation of the number of: (c) Differentially expressed genes; (d) Significant biologies/Biological processes, present in various experimental datasets in the transcriptome profiling study**.**

e) Pie chart representation of the top five gene ontologies and pathways summary on the basis of all the differentially expressed genes in the various transcriptome profiling experimental datasets showing approximately 5% of the total DEGs belonging to the pathways involved in cancer progression.

Fig. 2 Transcriptomic profiling, analysis and filtering of data.

a) Non-hierarchical heat map representation depicting the expression profile and variation in average log2 fold change among differentially expressed genes in various experimental datasets in the transcriptome profiling study. Color bar represents the expression values with green representing the lowest (downregulation) and red representing the highest (upregulation) expression levels. The various experimental datasets used for expression profiling are labelled on the top.

b) Volcano plot representation of the complete transcript list according to their average log2 fold change and p-values for various experimental datasets in the transcriptome profiling study with differential transcripts highlighted in blue. The plot displays differentially expressed genes along the dimensions of biological (average log2 fold change-FC) and statistical significance (p). Genes with absolute log2 fold change>2 and a p-value<0.05 were considered as differentially expressed genes.

c) Venn diagram representation created using Venny 2.1.0 showing 77 common elements in the transcriptome profiling of the *in vitro* HNSCC cell line model (A-D datasets).

d) Venn diagram representation created using Venny 2.1.0 showing the five common genes in 77 common elements in the transcriptome profiling of the *in vitro* HNSCC cell line model (A-D datasets).

e) Venn diagram representation created using Venny 2.1.0 showing the two common genes in 16 common elements in the transcriptome profiling of the *in vitro* HNSCC cell line model (A-D datasets).

Fig. 3 Transcriptomic profiling of MK2 pathway and its regulation.

a) Gene regulatory network showing the role of MK2 in regulation of mRNA stability in the transcriptome profiling of the *in vitro* HNSCC cell line model (A-D datasets) and the *in vivo* heterotopic HNSCC xenograft experimental dataset (F vs E dataset). The figure clearly demonstrates that MK2 regulates transcript stability *via* RBP-mediated regulation with HuR (ELAVL1) and TTP (ZFP36) playing integral part. The gene nodes are sized according to their p-values and colored according to their average log2 fold change, where red shows upregulation while green shows downregulation and yellow indicates baseline expression; processes are showed in rectangular boxes and colored in blue.

b) Gene regulatory network showing MAPK signaling cluster of the selected MK2 pathway genes (p38, MK2, AUF1, TTP, CUGBP1, CEBPδ, HuR, MKP-1, p27, TNF-α and VEGF) in the transcriptome profiling data of the *in vitro* HNSCC cell line dataset (B vs A, normoxic microenvironment) indicating VEGF and TNF-α down regulation. The gene nodes are sized according to their p-values and colored according to their average log2 fold change, where red shows upregulation while green shows downregulation and yellow indicates baseline expression; processes are showed in rectangular boxes and colored in blue.

Fig. 4 nCounter gene expression assay analysis.

a, b) Pie chart representation of the top five gene ontologies and pathways summary on the basis of all the differentially expressed genes in the various experimental datasets used for nCounter gene expression assay. (a) *in vitro* HNSCC cell line model (A-D datasets) and (b) *in vivo* heterotopic HNSCC xenograft experimental dataset (F vs E dataset) shows the total DEGs belonging to the various biological processes.

c) Representative non-hierarchical heat map representation depicting the expression profile and variation in average log2 fold change among differentially expressed genes in various experimental datasets analyzed by nCounter gene expression assay considering the complete CodeSet of 86 genes. Color bar represents the expression values with green representing the highest (upregulation) and red representing the lowest (downregulation) expression levels. The various experimental datasets used for expression profiling are labelled on the top: *in vitro* HNSCC cell line model (A-D datasets) where Lane 1-KDH is Dataset D (MK2-knockdown CAL27 cells in Hypoxia; Lane 2-NH is Dataset C (Normal CAL27 cells in Hypoxia); Lane 5-KDN is Dataset B (MK2-knockdown CAL27 cells in Normoxia); Lane 6-

NN is Dataset A (Normal CAL27 cells in Normoxia); and *in vivo* heterotopic HNSCC xenograft experimental dataset where Lane 3-KDX is MK2-knockdown CAL27 cells grafted (F dataset) and Lane 4-NX is Normal CAL27 cells grafted (E dataset).

d) Representative non-hierarchical heat map representation depicting the expression profile and variation in average log2 fold change among differentially expressed genes in various experimental datasets analyzed by nCounter gene expression assay considering the individual CodeSet of 39 genes for the *in vitro* HNSCC cell line model (A-D datasets) and 48 genes for the *in vivo* heterotopic HNSCC xenograft experimental dataset. Color bar represents the expression values with green representing the highest (upregulation) and red representing the lowest (downregulation) expression levels. The various experimental datasets used for expression profiling are labelled on the top and the two distinct clusters of upregulated and downregulated genes are visible in each combination.

e) Box-plot representation depicting the differences between the expression profile of differentially expressed genes in various experimental datasets analyzed by nCounter gene expression assay considering the individual CodeSet of 39 genes for the *in vitro* HNSCC cell line model (A-D datasets) and 48 genes for the *in vivo* heterotopic HNSCC xenograft experimental dataset. The counts were normalized with the positive controls and averaged for the samples of each mRNA type ($p<0.05$).

Fig. 5 Representation of the analysis flow for all datasets, and Secondary validation through immunohistochemical analysis.

a) Figure depicts pictorial representation of the key outcomes obtained with the workplan. The primary analysis section represents initial filtering of the transcriptome data for identification of crucial biological processes and genes for all the experimental datasets. Additionally, expression levels and gene networks were checked for two categories of genes, first for the genes that are transcriptionally regulated by MK2 and second, for the genes having their mRNA stability controlled by MK2 through RBP-mediated regulation. Thereafter, MK2-regulated transcripts that harbor RBP specific regions in their 3'-UTRs were identified in all datasets, providing 34 genes in A to D comparisons and 48 in F vs E comparisons which are further analyzed by nCounter gene expression analysis. Finally, 7 common DEGs in transcriptomic and nCounter analysis were identified and 6 of them from F vs E dataset were subjected to IHC validation where IGFBP2, MUC4 and PRKAR2B were found to be upregulated and hence in consonance with our findings.

b) Representation of the protein expression levels of the three candidate MK2-regulated genes in HNSCC xenografted mice tumor sections by immunohistochemical analysis. Color

bar represents the expression values in terms of mean score, blue bar represents the expression level in CAL27-MK2$_{WT}$ (Dataset E) and the red bar signifies protein expression in CAL27-MK2$_{KD}$ sections (Dataset F). Data clearly implies the upregulation in expression levels of IGFB2, MUC4, and PRKARB2 proteins in CAL27-MK2$_{KD}$ tissue sections as compared to CAL27-MK2$_{WT}$. Here, parametric Welch t-test was used for evaluating the statistical significance using GraphPad Prism 7.0 software with ** denoting $p<0.01$; n=5 field views for IHC analysis.

Fig. 6 Pictorial representation of the analyzed MAPK pathway.

Graphical illustration depicts analyzed MAPK pathway in this study including activation of MAPKAPK2 and the plausible mode of action in HNSCC pathogenesis. Figure elucidates the final MK2-regulated putative candidate genes obtained in this study which could be further explained as possible targets in the pathway for HNSCC management.

a)

Normal CAL27 cells (CAL27-MK2$_{WT}$)
MK2-knockdown CAL27 cells (CAL27-MK2$_{KD}$)

Hypoxia
Hypoxia chamber
Normoxia
Subcutaneous injection of cells
Xenografted tumor

Normal CAL27 cells (CAL27-MK2$_{WT}$)
MK2-knockdown CAL27 cells (CAL27-MK2$_{KD}$)

Cell culture and xenograft murine model

Cell lines RNA (Dataset A, B, C, D)

Xenograft tumor RNA (Dataset E, F)

RNA quality assessment

RNA isolation and quality check

cDNA library

cDNA library quality assessment

cDNA library preparation and quality assessment

Next generation sequencing (NovaSeq 6000)

Data quality check and analysis

nCounter gene expression assay (Primary validation)

Immunohistochemical analysis (Secondary validation)

b)

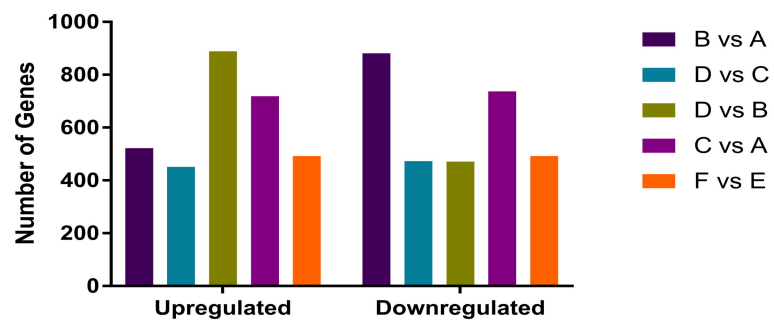
HQ Reads

Alignment (KALLISTO) ← Reference Genome : NCBI mRNA Human Data

Differential Expression Analysis (DESEQ) ← Log2 ratio>= ±2

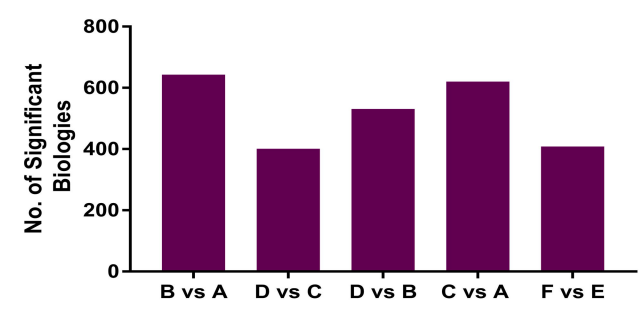
Annotation, Functional Enrichment Analysis and Interpretation ← **Annotation:** DAVID **Heat Maps:** Cluster 3.0 and Java Tree View **Networks:** Cytoscape

c)



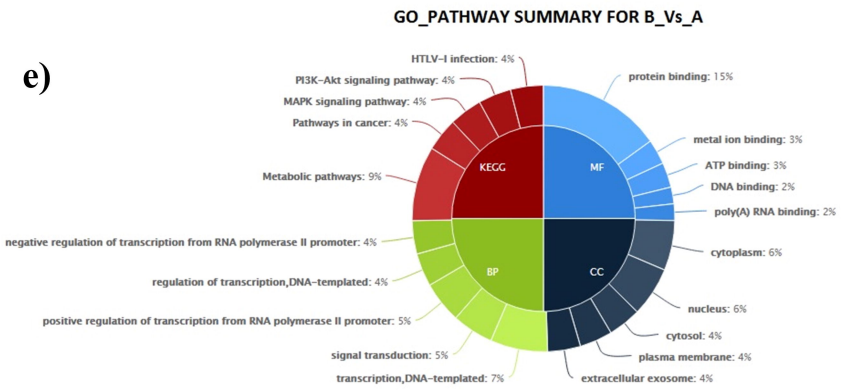
	B vs A	D vs C	D vs B	C vs A	F vs E
Up	522	451	889	719	492
Down	881	473	471	737	492

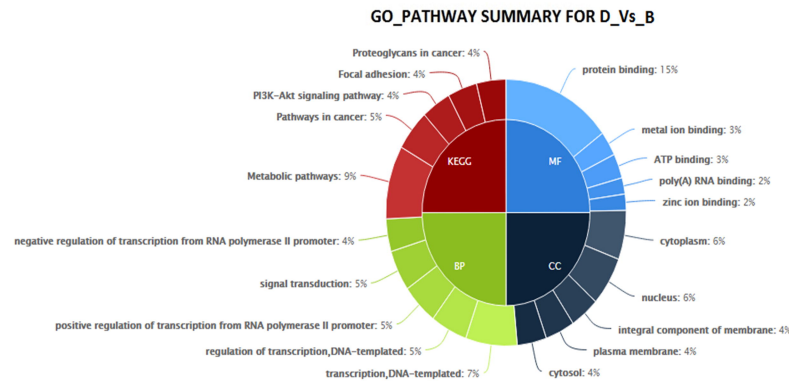
d)

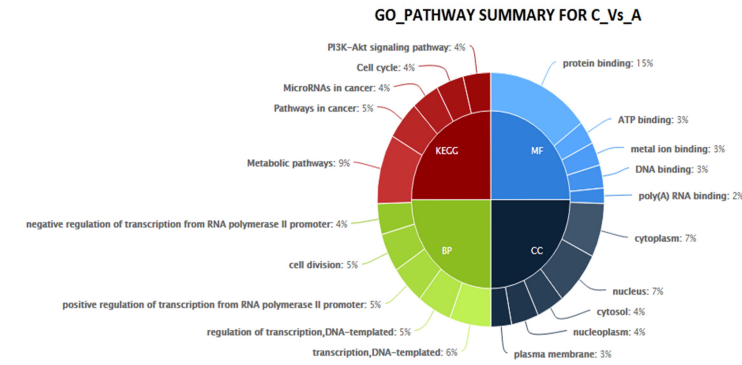


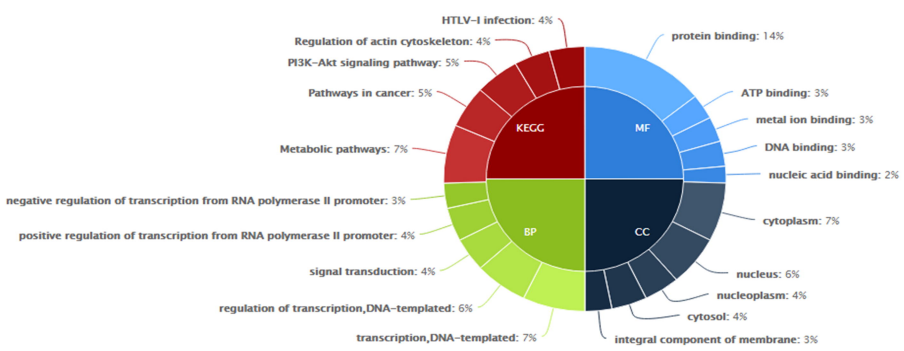
	B vs A	D vs C	D vs B	C vs A	F vs E
	643	401	531	620	408

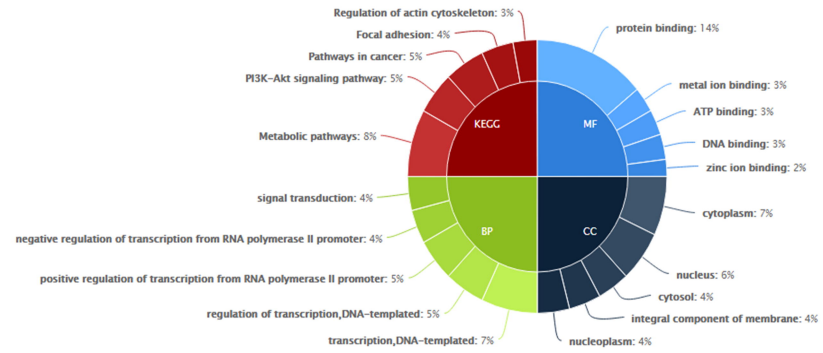
e)

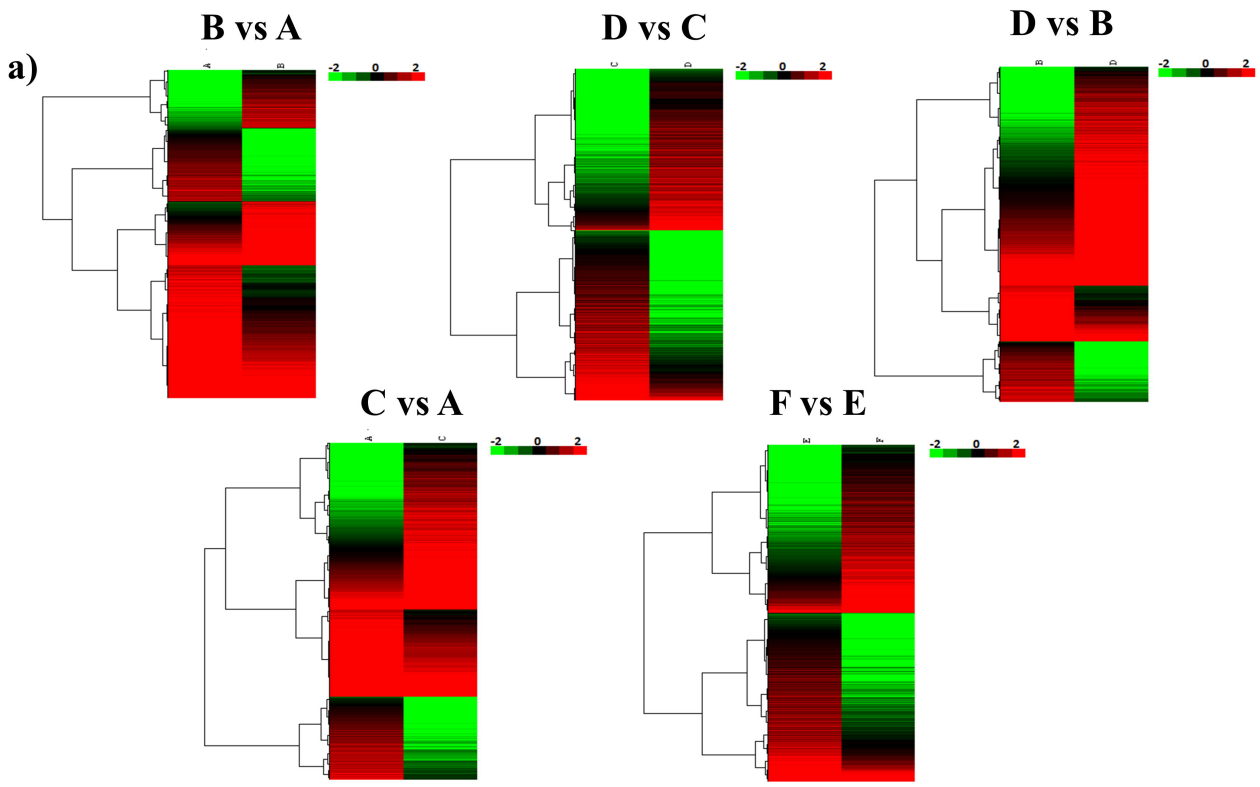


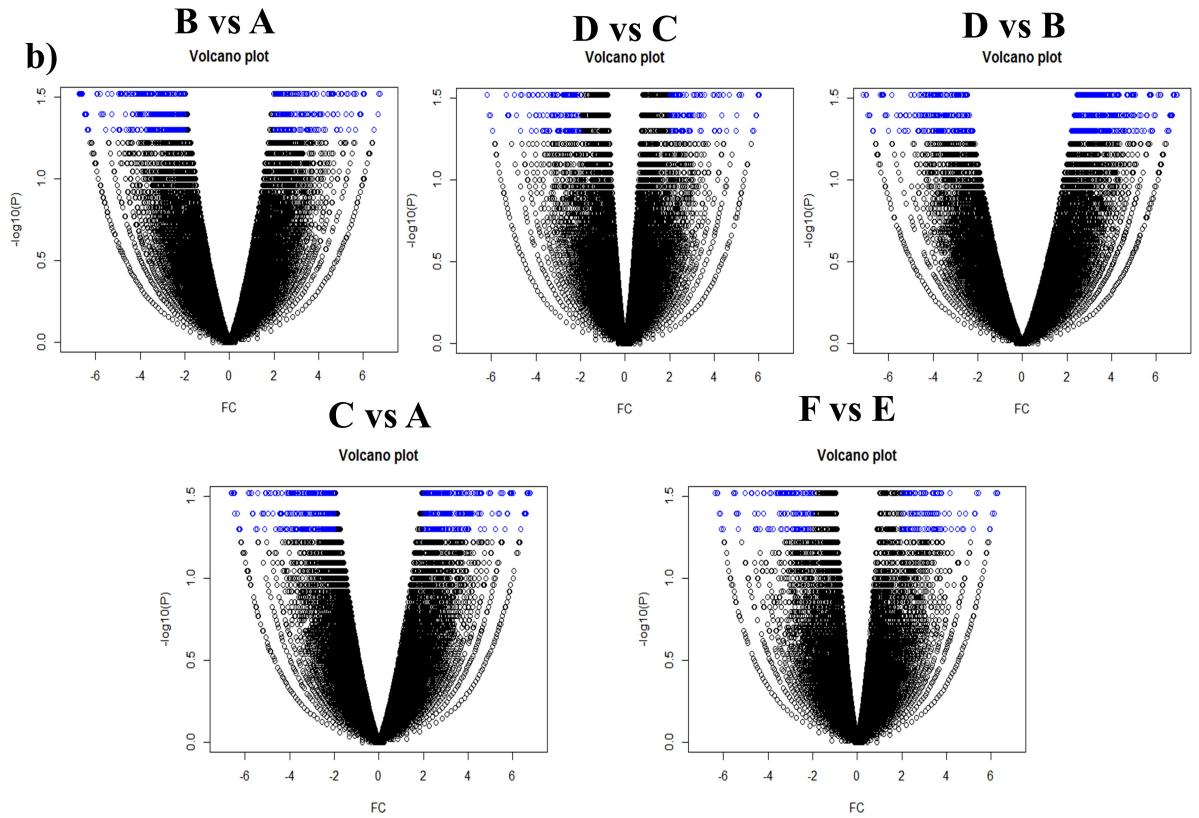


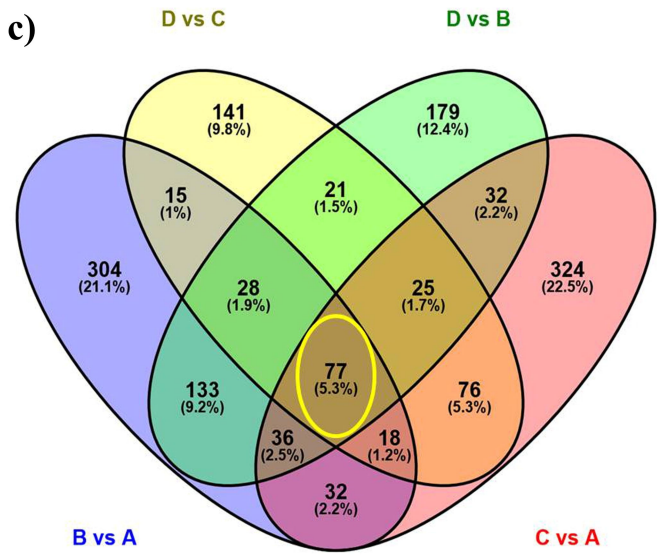


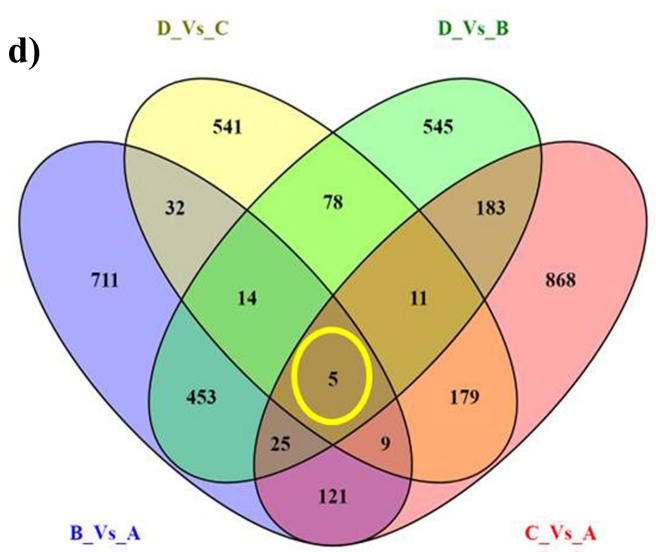


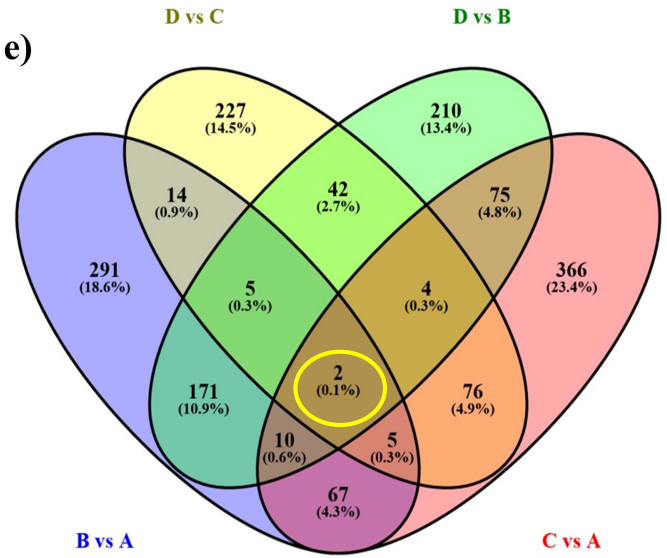



a)
B vs A
D vs C
D vs B
C vs A
F vs E

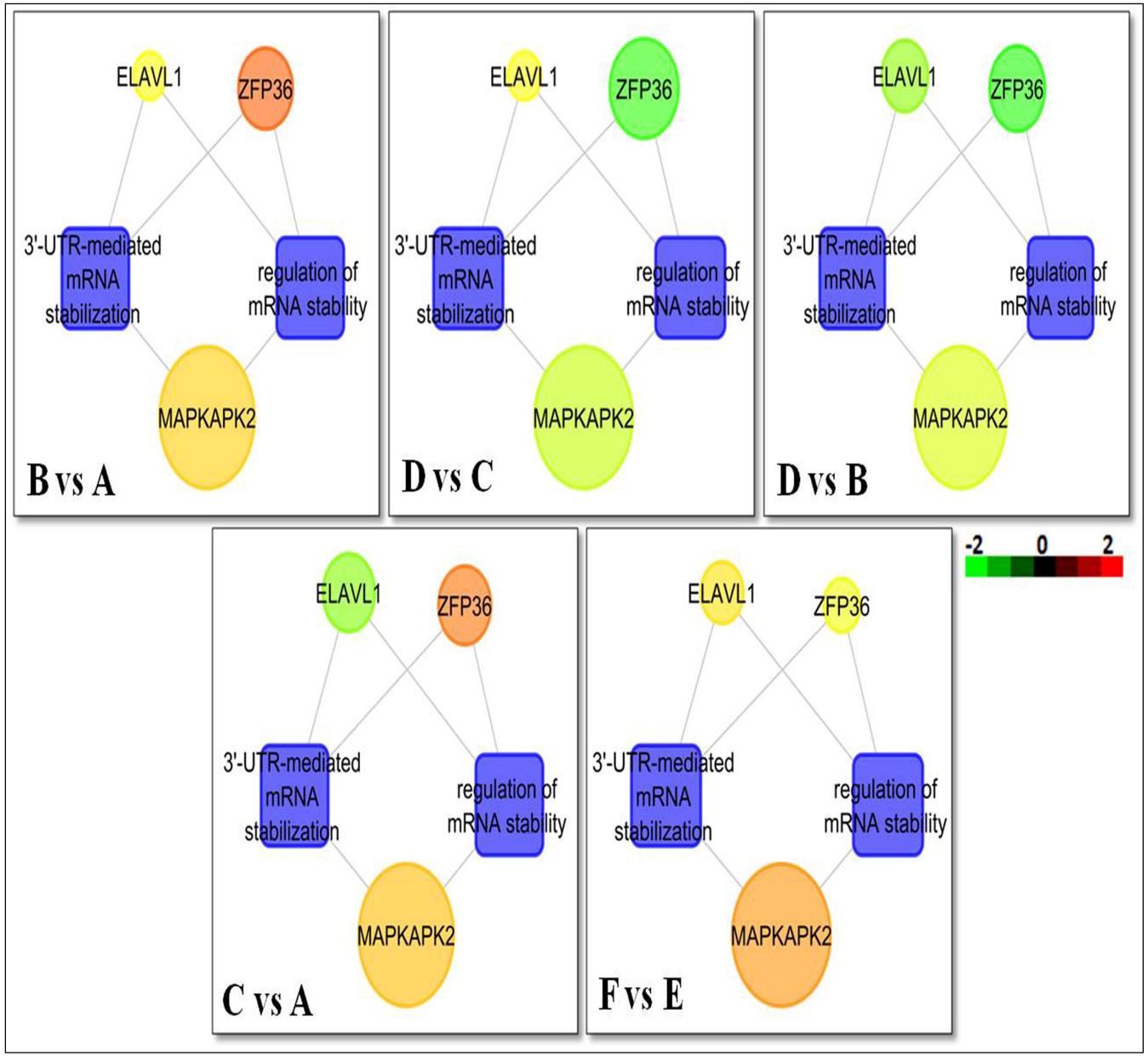

b)
B vs A
D vs C
D vs B
C vs A
F vs E
Volcano plot
-log10(P)
FC


c)
D vs C
D vs B
B vs A
C vs A
141 (9.8%)
179 (12.4%)
15 (1%)
21 (1.5%)
32 (2.2%)
304 (21.1%)
28 (1.9%)
25 (1.7%)
324 (22.5%)
77 (5.3%)
133 (9.2%)
76 (5.3%)
36 (2.5%)
18 (1.2%)
32 (2.2%)

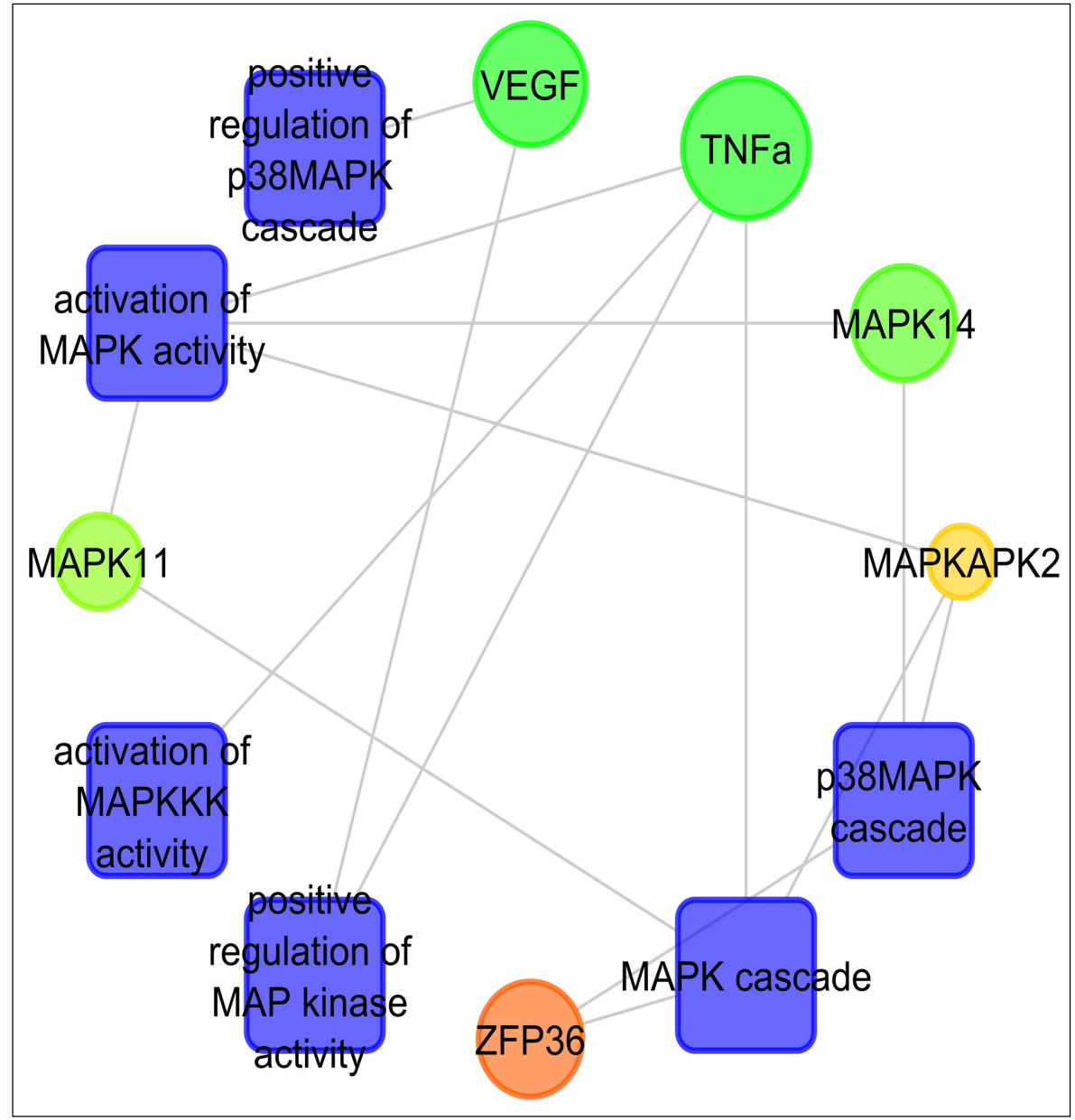

d)
D_Vs_C
D_Vs_B
B_Vs_A
C_Vs_A
541
545
32
78
183
711
14
11
868
5
453
179
25
9
121

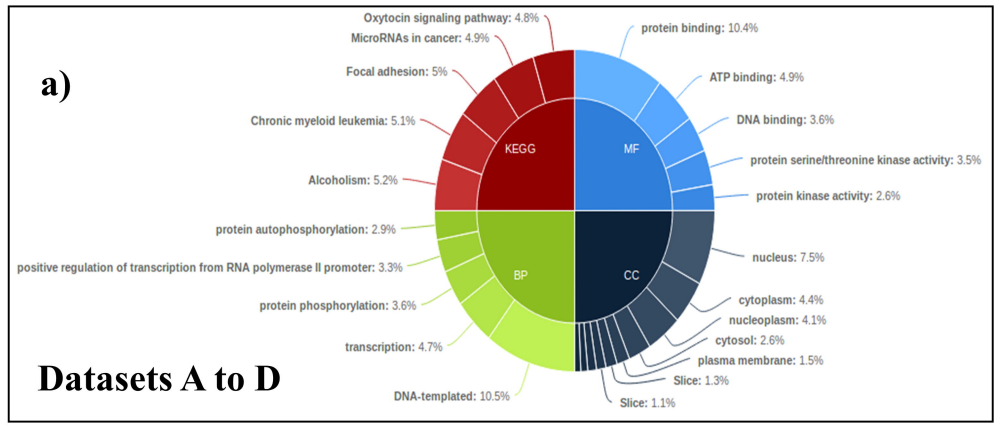

e)
D vs C
D vs B
B vs A
C vs A
227 (14.5%)
210 (13.4%)
14 (0.9%)
42 (2.7%)
75 (4.8%)
291 (18.6%)
5 (0.3%)
4 (0.3%)
366 (23.4%)
2 (0.1%)
171 (10.9%)
76 (4.9%)
10 (0.6%)
5 (0.3%)
67 (4.3%)

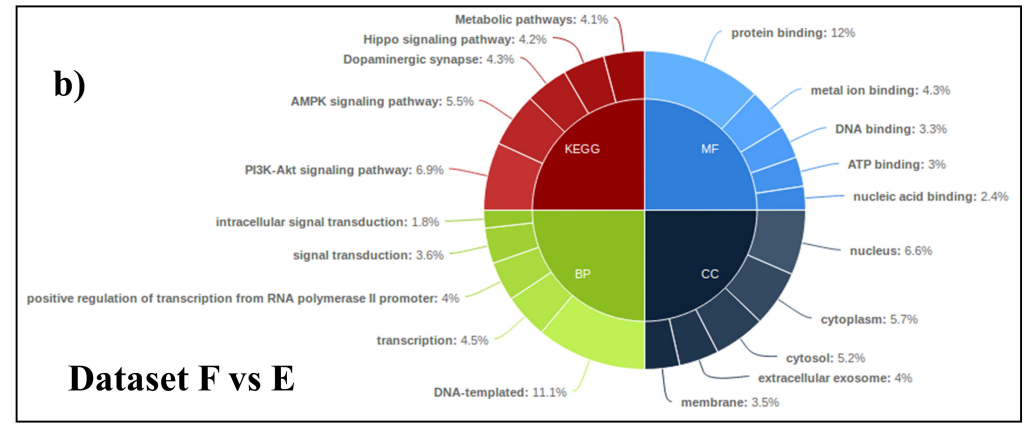
a)

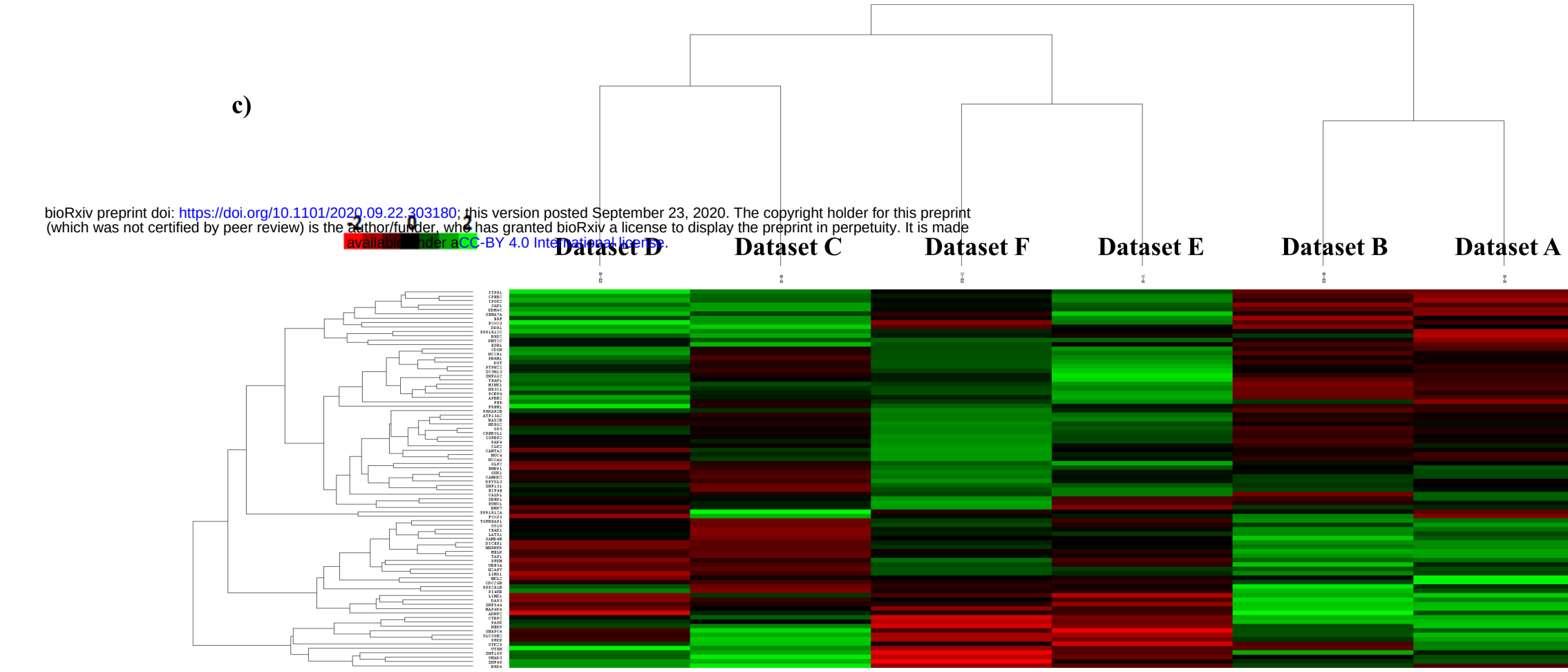

ELAVL1
ZFP36
3'-UTR-mediated mRNA stabilization
regulation of mRNA stability
MAPKAPK2
B vs A
D vs C
D vs B
C vs A
F vs E
-2
0
2

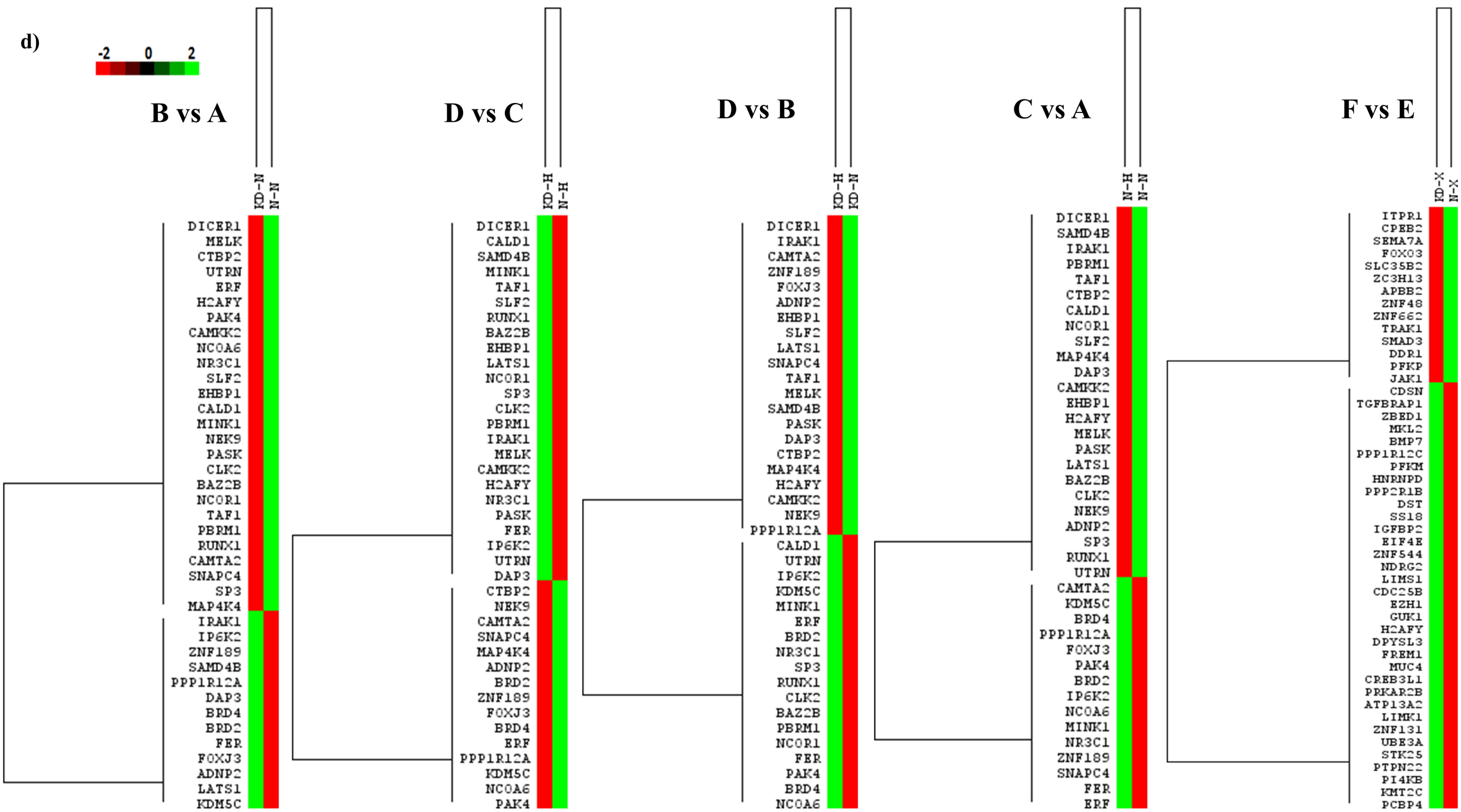
b)


positive regulation of p38MAPK cascade
VEGF
TNFa
activation of MAPK activity
MAPK14
MAPK11
MAPKAPK2
activation of MAPKKK activity
p38MAPK cascade
positive regulation of MAP kinase activity
MAPK cascade
ZFP36

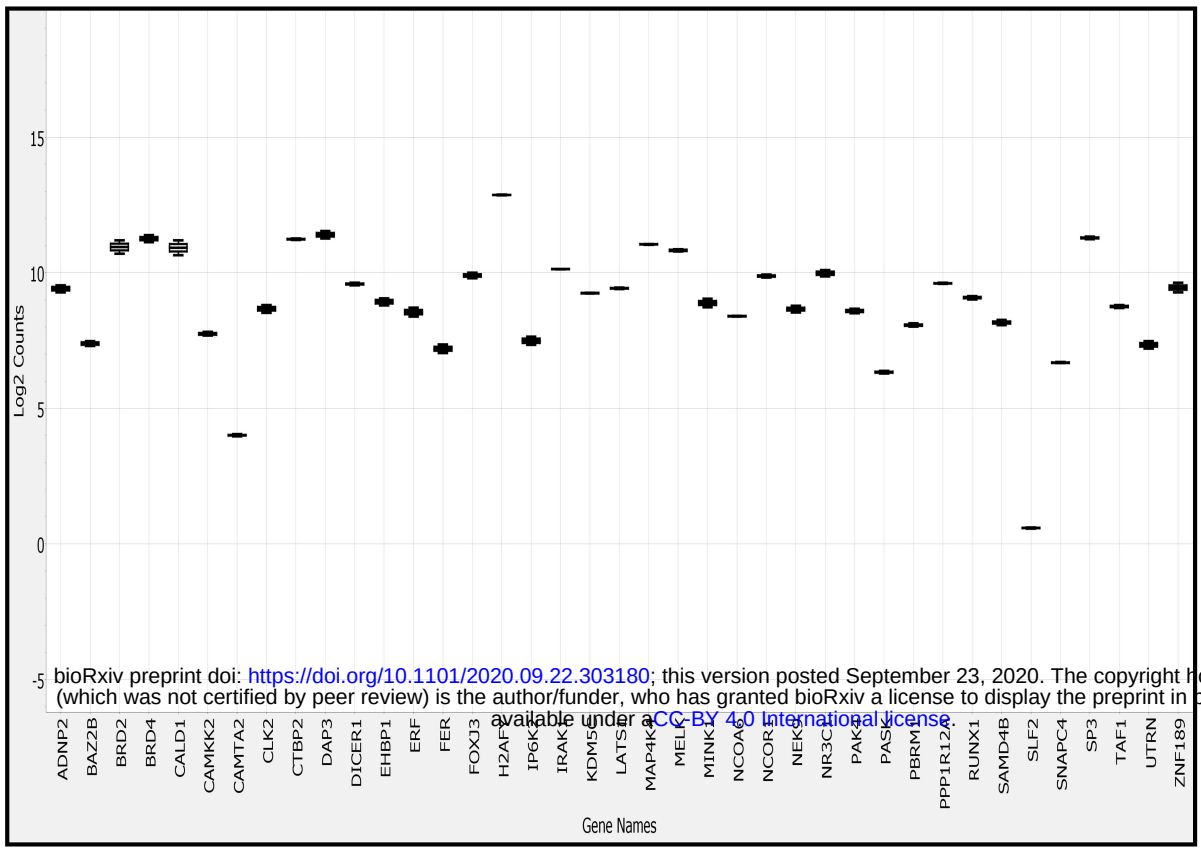


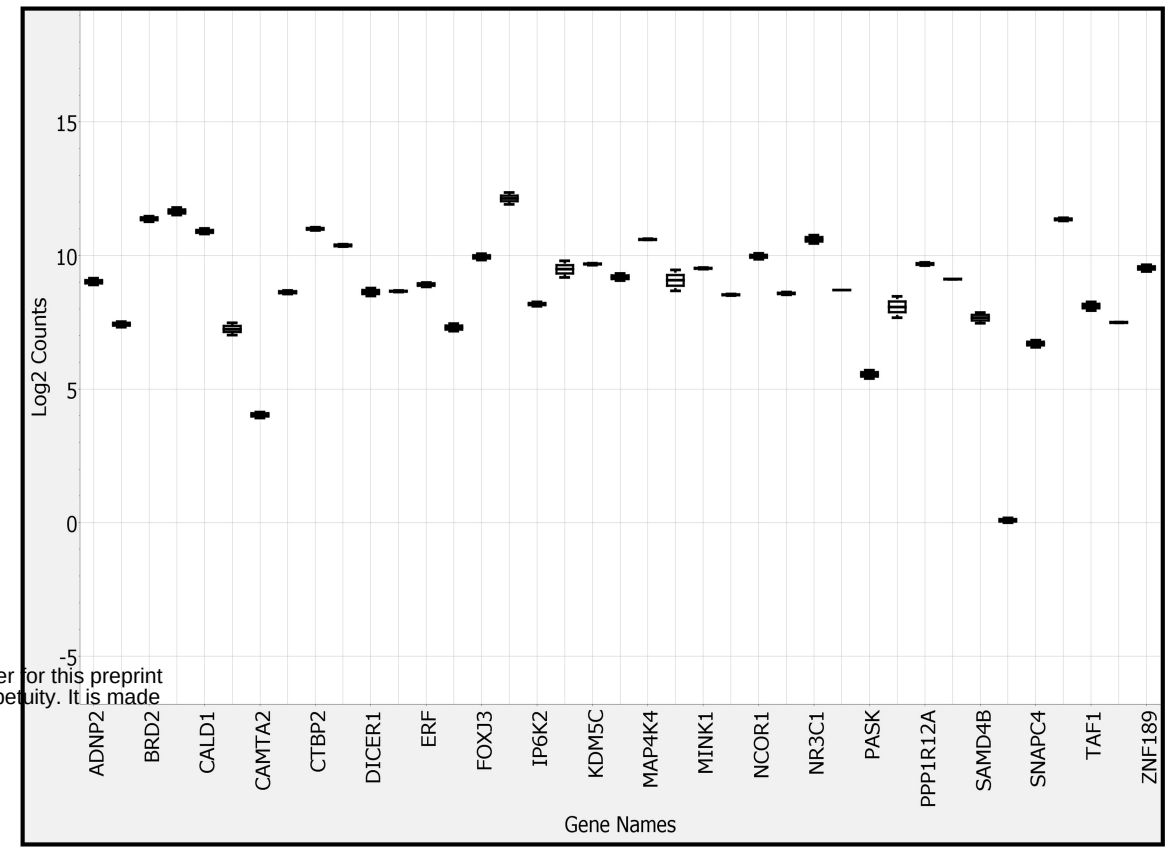


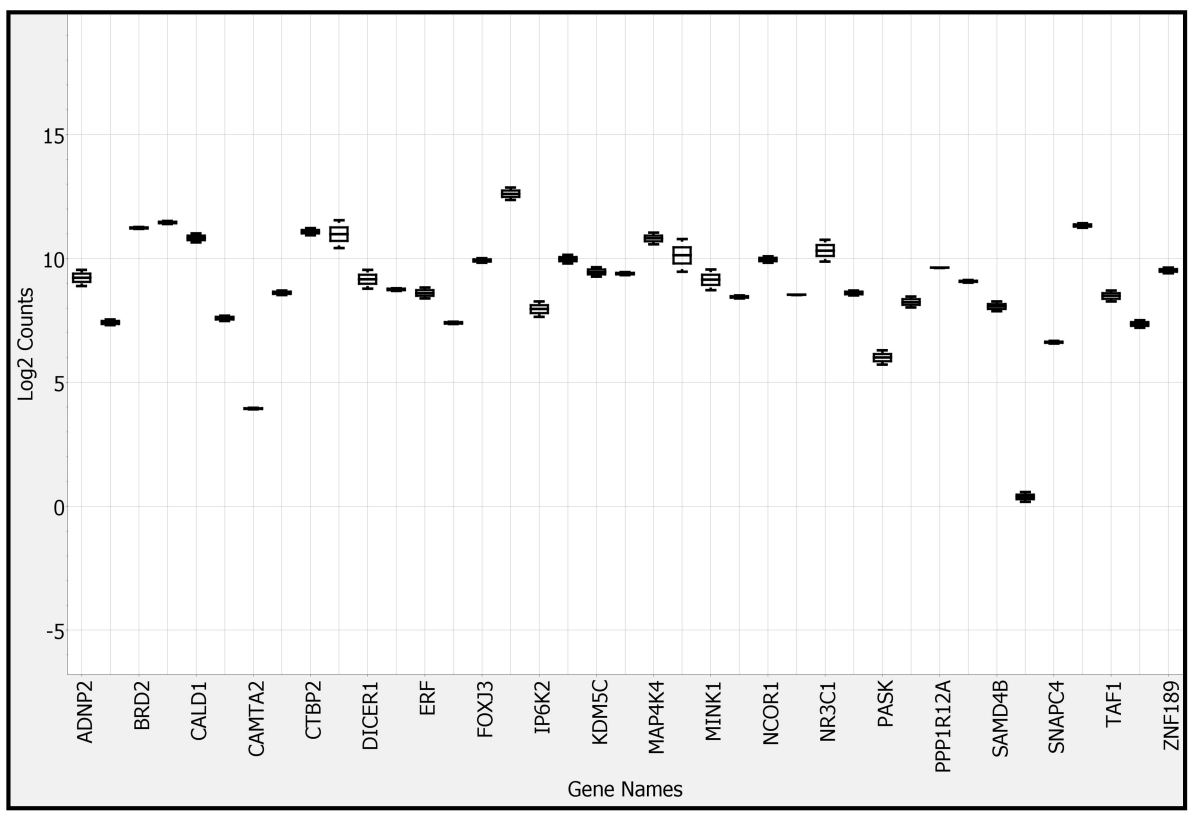


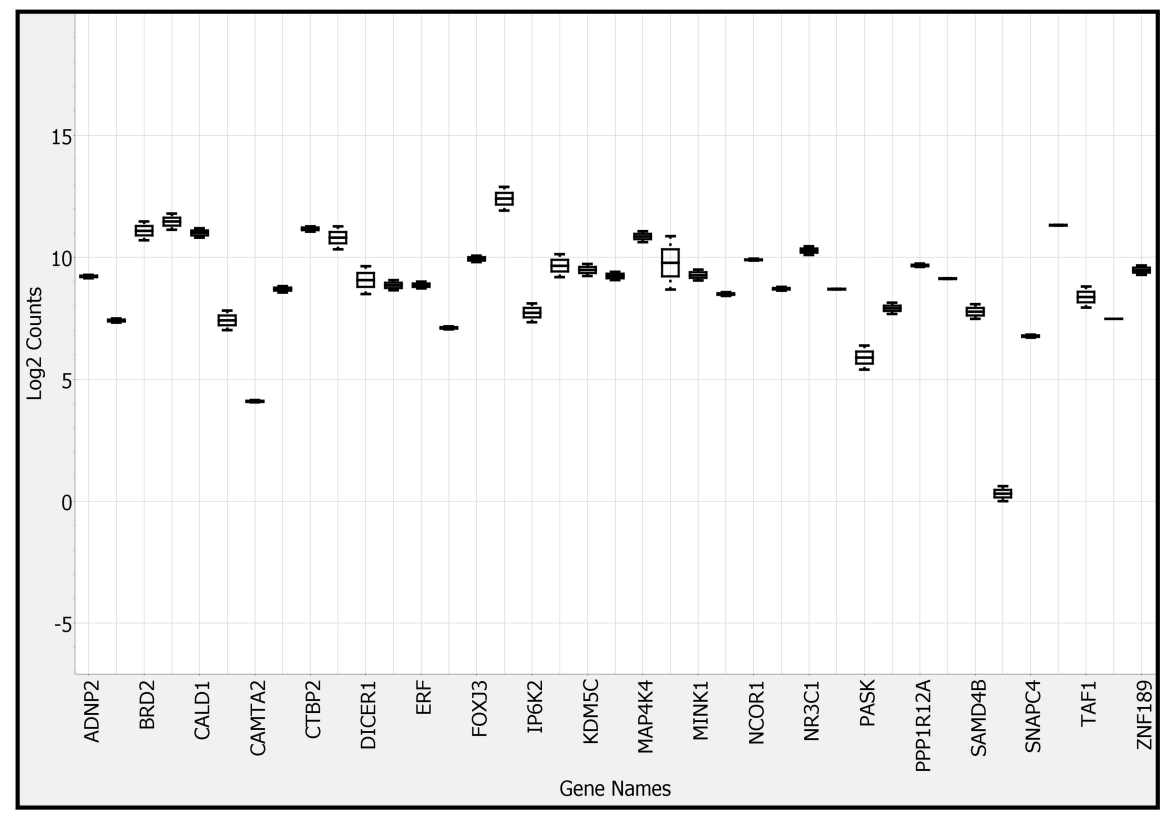


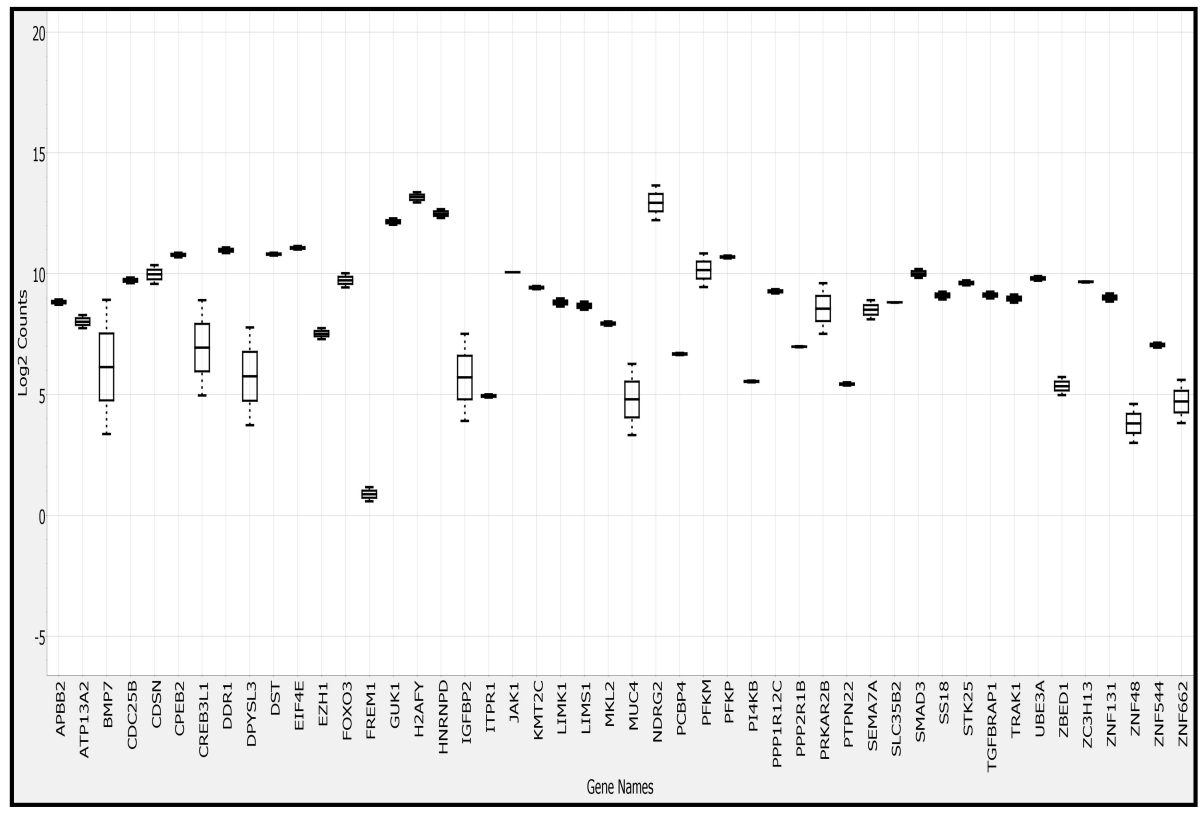
e)











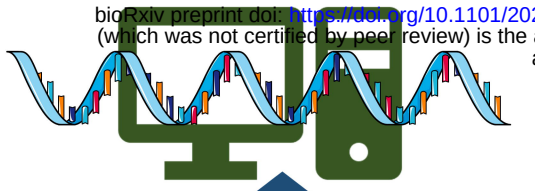
a)

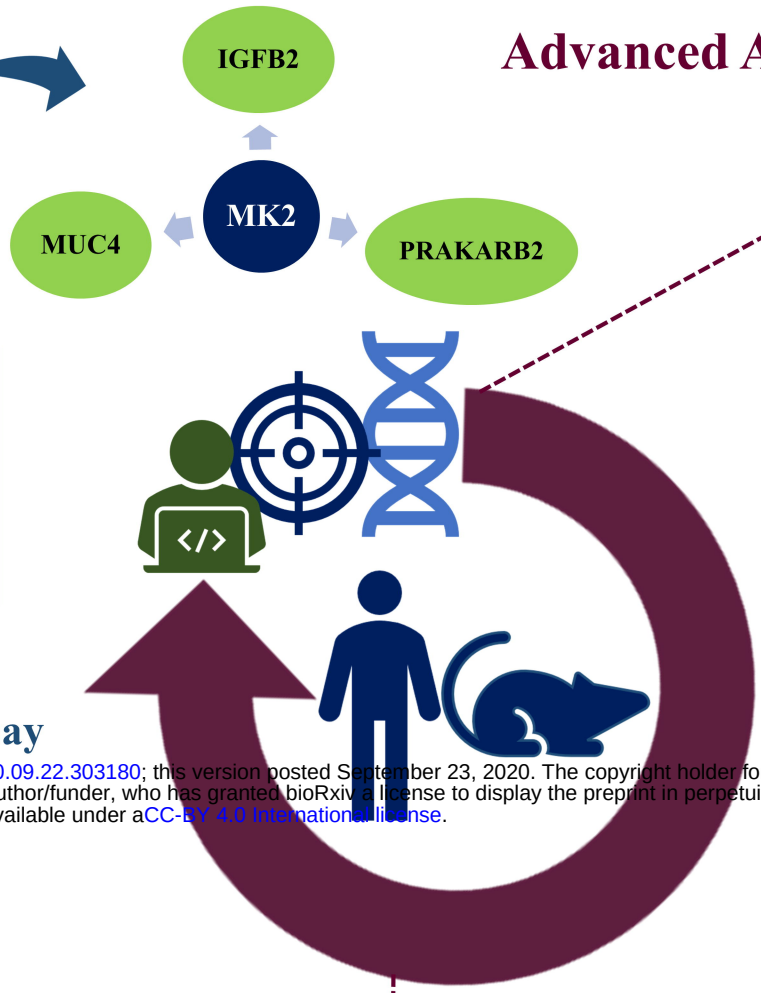
IGFBP2, MUC4, PRKAR2B

IHC analysis of 6 genes in F vs E dataset

7 common DEGs in transcriptomic and nCounter analysis (FC>2 or <-2, p<0.05)
MELK, ZNF662, BMP7, CREB3L1, IGFBP2, MUC4, PRKAR2B

n-Counter Gene Expression Assay



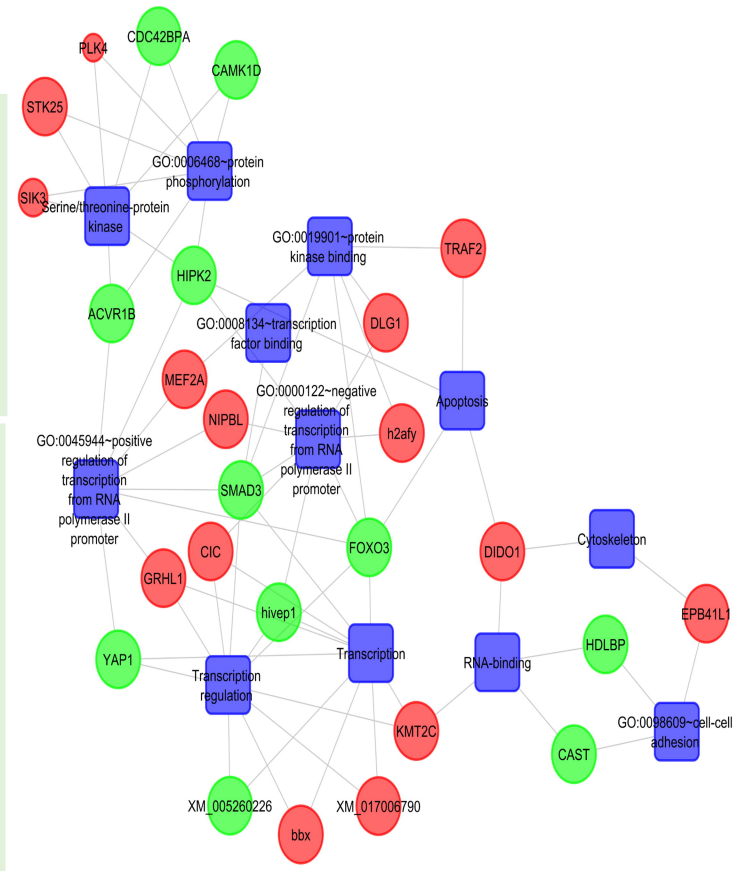


Advanced Analysis

Cell Line (A to D dataset)
16 cancer specific biological processes
2 common genes: DAP3, RUNX1

Tumor Xenograft (F vs E dataset)
14 biological processes relevant to tumor pathogenesis
Key Genes: TRAF2, EPB41L1, FOXO3, H2AFY , YAP1, DIDO1

Primary Analysis

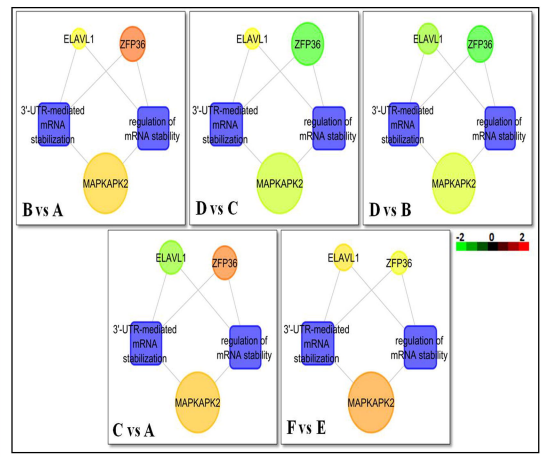


A to D datasets
Top 2 up/downregulated genes in 16 cancer specific biological processes = 34 MK2-regulated genes

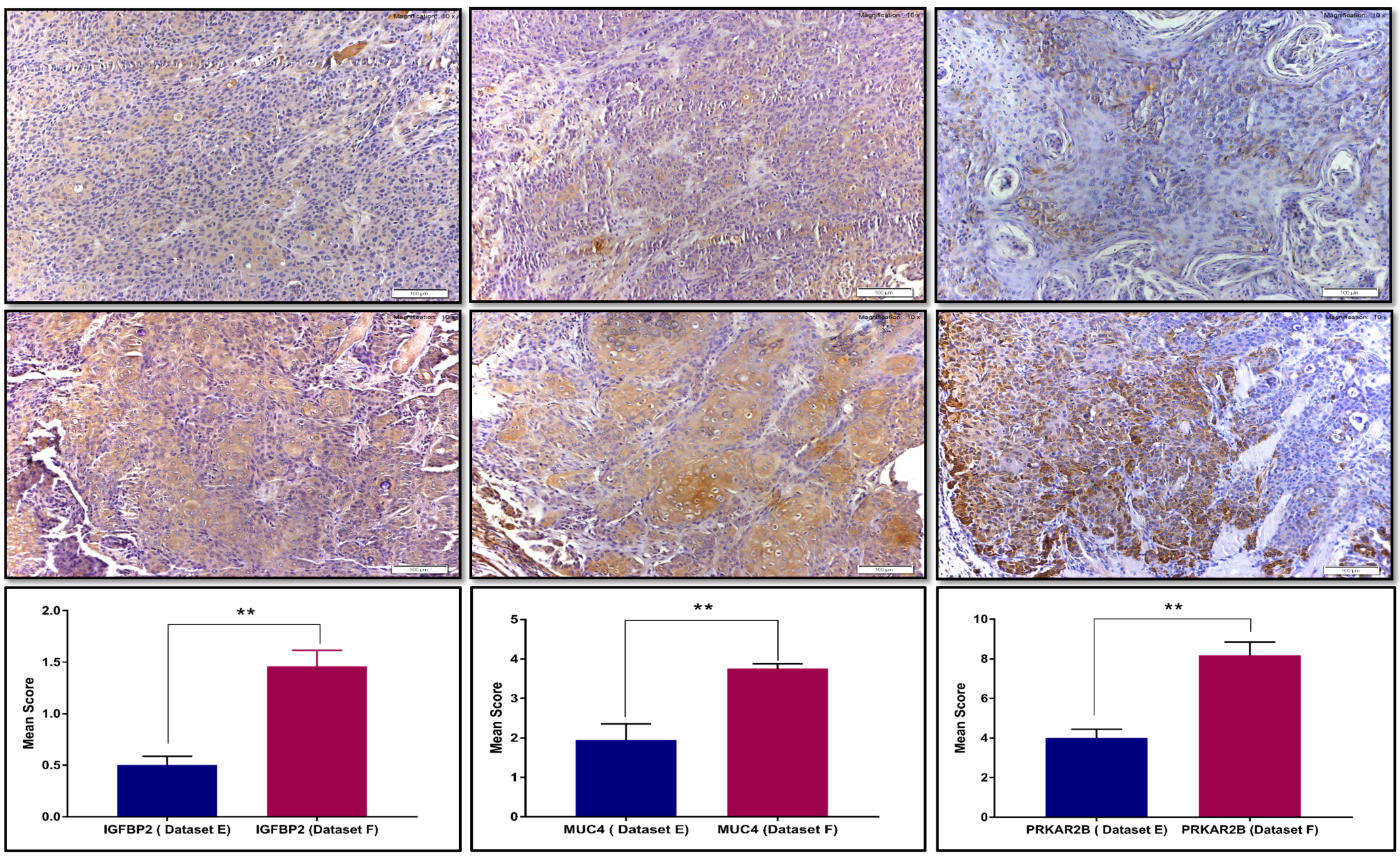
F vs E dataset
Top differentially expressed genes = 48 MK2-regulated genes

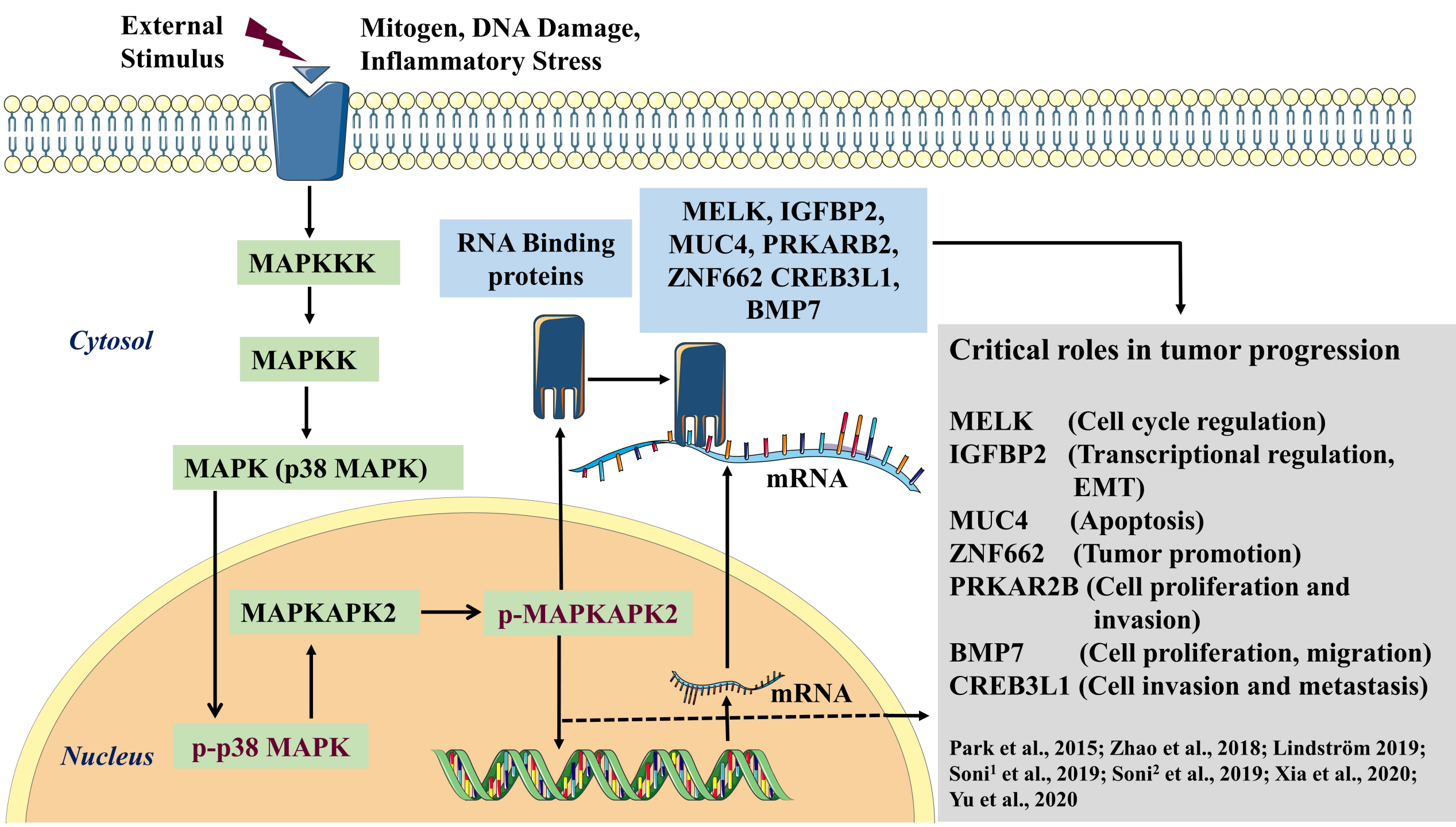
MK2 regulates the transcript stability of genes
MK2-regulated transcripts that harbor RBP specific regions in their 3'-UTRs

Cell Line and Xenograft
Levels of selected genes transcriptionally regulated by MK2 involved in key cellular pathways:
AUF1, TTP, CUGBP1, CEBPδ, HuR, MKP-1, p27, TNF-α, VEGF
MK2 regulates mRNA stability through RBP-mediated regulation:
HuR (ELAVL1) and TTP (ZFP36)



b)





Critical roles in tumor progression

MELK	(Cell cycle regulation)
IGFBP2	(Transcriptional regulation, EMT)
MUC4	(Apoptosis)
ZNF662	(Tumor promotion)
PRKAR2B	(Cell proliferation and invasion)
BMP7	(Cell proliferation, migration)
CREB3L1	(Cell invasion and metastasis)

Park et al., 2015; Zhao et al., 2018; Lindström 2019; Soni[1] et al., 2019; Soni[2] et al., 2019; Xia et al., 2020; Yu et al., 2020