Multi-omics approach to infer cancer therapeutic targets on chromosome 20q across tumor types

Antoine M Snijders and **Jian-Hua Mao**

Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States

Abstract

The identification of good targets is a critical step for the development of targeted therapies for cancer treatment. Here, we used a multi-omics approach to delineate potential targets on chromosome 20q, which frequently shows a complex pattern of DNA copy number amplification in many human cancers suggesting the presence of multiple driver genes. By comparing the amounts of individual mRNAs in cancer from 11 different human tissues with those in their corresponding normal tissues, we identified 18 genes that were robustly elevated across human cancers. Moreover, we found that higher expression levels of a majority of these genes were associated with poor prognosis in many human cancer types. Using DNA copy number and expression data for all 18 genes obtained from The Cancer Genome Atlas project, we discovered that amplification is a major mechanism driving overexpression of these 18 genes in the majority of human cancers. Our integrated analysis suggests that 18 genes on chromosome 20q might serve as novel potential molecular targets for targeted cancer therapy.

Keywords

cancer; therapeutic targets; chromosome 20q; omics; bioinformatics

Introduction

Targeted therapy is a newer type of cancer treatment that uses drugs or other substances to more precisely identify and attack cancer cells, usually while doing little damage to normal cells. However, the development of targeted therapies requires the identification of good targets. A few approaches can be used to discover potential targets. The first is to identify

abnormalities in chromosomes (such as translocation) that are present in cancer cells but not in normal cells. Chromosome translocations can lead to fusion proteins, which are potential targets for targeted therapies. For example, Gleevec® targets the BCR-ABL fusion protein for treating some leukemia cells that contain BCR-ABL translocation[1,2]. The second approach is to identify the gene mutations that drive cancer progression. An example is the *BRAF* V600E mutation, which is present in a wide range of cancers including a high percentage of malignant melanomas[3]. Vemurafenib targets this mutant form of the BRAF protein and it has been approved for treating patients with inoperable or metastatic melanoma that contains this altered BRAF protein[4]. Finally, by comparing the amounts of individual proteins in cancer cells with those in normal cells, one can identify proteins that are more or less abundant in cancer cells. An example of such a differentially expressed target is the human epidermal growth factor receptor 2 protein (HER-2). HER-2 is expressed at high levels on the surface of some cancer cells. Several targeted therapies are directed against HER-2 including trastuzumab, which is approved to treat certain breast and stomach cancers that overexpress HER-2[5]. Microarray and next generation sequencing technologies have become invaluable tools used to catalog these genomic abnormalities occurring in human cancers, and they can be used to identify new potential therapeutic targets.

The availability of large cancer genomic data sets allows for unbiased approaches to identify genes that are important in tumor progression. Gene transcript-based signatures that predict prognosis have successfully been developed for many different tumor types. However, it remains a challenge to distinguish cancer driver genes from passenger genes; the latter referring to genes that are correlated (in expression) to driver genes and are likely prognostic biomarkers but are, nonetheless, not actively contributing to the carcinogenic process.

An essential early step in the pathogenesis of most cancers is losing one of the defense mechanisms that controls the integrity of the genome, making it possible for a cell to rapidly acquire genomic changes. In a majority of epithelial cancers, genomic instability occurs at the chromosomal level, affecting numerous genes and thereby causing tumor progression. Amplifications — defined as regions of focal high-level DNA copy number change — are likely to represent aberrations under continuous selection for tumor growth since amplified DNA is unstable[6]. Thus, gene amplifications focus on genes that exist in a region with candidate oncogenes contributing to cancer development. DNA amplifications on chromosome 20q are often observed in many human cancers, suggesting that genes which reside on chromosome 20q play a causal role in tumorigenesis. Moreover, 20q amplifications are often highly complex, indicating the presence of multiple genes is important in tumor development[7,8].

Here, we aggregated available cancer databases to identify cancer driver genes across tumor types by combining gene transcript and DNA copy number across chromosome 20q to select tumor-type specific signatures that predict patient prognosis. Our strategy identified critical genes and pathways in tumor development that are important for designing better treatment strategies.

Materials and methods

Gene transcript data of normal (non-tumor) and tumor tissues across 11 different tumor types were obtained from the National Center for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO). They include: brain, GSE4290 (tumor vs. healthy donor); breast, GSE10780 (tumor vs. adjacent normal) and GSE3744 (tumor vs. healthy donor); colon, GSE8671 (tumor vs. adjacent normal); gastric, GSE13911 (tumor vs. adjacent normal); head and neck, GSE6791 (tumor vs. healthy donor) and GSE12452 (tumor vs. healthy donor); liver, GSE6764 (tumor vs. healthy donor); lung, GSE31210 (tumor vs. adjacent normal) and GSE19188 (tumor vs. adjacent normal); ovarian, GSE14407 (tumor vs. healthy donor); cervix, GSE6791 (tumor vs. healthy donor); pancreas, GSE16515 (tumor vs. adjacent normal); and prostate cancer, GSE3325 (tumor vs. healthy donor). Fold change was calculated for each gene and its significance was tested using t-test ($p < 0.05$). Survival multivariate analysis and risk assessment for individual genes and gene signatures in human cancer data sets were performed using SurvExpress[9] in the following datasets: ovarian (GSE9891 and GSE32062); head and neck (TCGA and E-MTAB-1328); breast (TCGA and GSE20685); liver (TCGA and GSE17856); lung adeno-carcinoma (TCGA); lung squamous cell carcinoma (TCGA); pancreatic (GSE28735 and GSE21501); stomach (TCGA); colon (TCGA and GSE17536); brain low-grade glioma (TCGA); brain glioblastoma multiforme (GSE16011); and prostate (MSKCC). Genomic alterations and mRNA expression levels for The Cancer Genome Atlas (TCGA) studies were obtained from cBioPortal[10,11]. We used a rank-based non-parametric test (Kruskal-Wallis) to determine whether the gene expression levels were significantly different between the copy number groups ($p < 0.05$ was used as a threshold for significance). All information concerning these samples can be downloaded from cBioPortal (http://www.cbioportal.org/data_sets.jsp). Gene ontology enrichment analysis was performed using the web-based gene set analysis toolkit ($p < 0.05$ was used as a threshold for significance)[12].

Results

Meta-analysis identified an 18-gene signature frequently upregulated across human tumor types

We conducted a meta-analysis of genes on chromosome 20q that are consistently upregulated across different human tumor types. We collected gene transcript data of normal and tumor tissues across 11 different tumor types including brain, breast, colon, gastric, head and neck, liver, lung, ovarian, cervix, pancreas, and prostate cancers. We calculated the differential expression of all 301 genes present on chromosome 20q for which gene transcript data was available. We then filtered for genes that were upregulated in tumors by at least 1.5 fold ($p < 0.05$) in seven or more tumor types. This resulted in a gene signature of 18 genes for further downstream analysis (**Figure 1A**). To identify tumor types that were the strongest contributors to the 18-gene signature, we generated a heatmap of the p-values pertaining to the expression difference between normal and tumor tissues across all tumor types (**Figure 1B**). We found that breast, lung, gastric, brain, and colon cancers have the strongest expression difference between tumor and normal tissues, while prostate and liver cancers have weaker contributions.

The 18-gene signature is associated with disease-free survival across tumor types

To investigate whether individual genes in our signature were associated with disease-free survival, a log-rank test was performed and the split-point of the patient cohort was chosen, where the p-value between two patient cohorts was the lowest. An example is shown in **Figure 2A** for the *RAE1* gene, where increased expression of *RAE1* was associated with decreased disease-free survival in breast, brain, head and neck, lung, ovarian, and pancreatic cancers ($p < 0.05$). **Figure 2B** shows a p-values' heatmap of the prognostic property of each gene within our signature across 10 tumor types (18 total data sets). Increased expression of each gene was associated with disease-free survival in at least three tumor types. The increased expression of *TPX2*, *UBE2C*, *CSE1L*, *AURKA*, and *RAE1* were associated with decreased disease-free survival in at least nine data sets. In some instances, we found that increased gene expression was associated with better prognosis (**Figure 2B**). For example, the increased expressions of *MMP9* and *SULF2* (two neighboring genes) were associated with increased survival in five data sets.

We then examined whether gene expression of the complete 18-gene signature was associated with cancer patient survival across 12 tumor types. Risk groups were assigned by dividing the patient cohort (ranked by prognostic index) into two equal parts. We observed significant association of our signature with disease-free survival in all 18 independent data sets ($0.012 < p < 1.608 \times 10^{-9}$; **Figure 3**). These data indicated that our signature is broadly predictive for disease-free survival independent of tumor type.

Increase in DNA copy number as a mechanism for increased expression

A variety of mechanisms can result in a change in gene expression, which modifies tumor growth and progression and these include point mutations, epigenetic alterations, genomic rearrangements, and DNA copy number aberrations. For example, it has been shown that gene amplifications are under continuous selection pressures and when the selection pressure is removed, amplifications are not maintained and eventually disappear. Thus, amplifications focus on those genes that are important for tumor development. To identify whether gene amplification was associated with increased gene expression within our 18-gene signature, we used a rank-based non-parametric test to determine the significance of differences in gene expression between tumors for each gene, with and without an increase in gene copy number. For example, the *RAE1* expression was found to be significantly associated with DNA copy number in 20 tumor types (**Figure 4A**). The p-values' heatmap of the rank-based test is shown in **Figure 4B** for all 18 genes across 21 tumor types. Strong associations between the DNA copy number and gene expression were observed in the majority of tumor types, whereas sparse associations were observed in acute myeloid leukemia (AML), pheochromocytoma and paraganglioma (PCPG), and thyroid cancer (THCA) (**Figure 4B**). Elevated DNA copy numbers of *MMP9* and *SULF2* were associated with increased gene expressions in only two and seven tumor types, respectively. Unsurprisingly, in AML, the *RAE1* expression was not significantly associated with the copy number since the *RAE1* copy number was only observed in one tumor (although the *RAE1* expression in that tumor was increased compared to tumors without the *RAE1* copy number gain). This analysis showed that the DNA copy number is a major mechanism for

our 18-gene signature on chromosome 20q, by which cells increase gene expression levels as they progress toward malignancy.

Discussion

The identification of good targets for the development of targeted therapies remains difficult. Cancer genome analysis has shown great promise in identifying key aberrations in tumor growth and survival pathways that could serve as targets for therapeutic intervention. Our integrated multi-omics analysis of genes on chromosome 20q identified 18 genes that might serve as novel potential molecular targets for targeted therapy. Gene ontology analysis revealed significant enrichment of cell cycle and mitosis-related biological processes in our 18-gene signature (**Figure 5**), suggesting that a cluster of functionally related genes localize to chromosome 20q. In addition, a number of these genes have previously been identified as oncogenes, including: *AURKA*, *UBE2C*, *TPX2*, *FAM83D*, *ZNF217*, *SALL4*, *MMP9*, and therapeutic targets: *AURKA* and *UBE2C*.

Our analysis identified several well-known oncogenes that are involved in mitosis, including: *AURKA*, *DSN1*, *UBE2C*, *FAM83D*, *TPX2*, and *MYBL2*. AURKA, a mitotic spindle-associated kinase that regulates key events during mitosis is frequently over-expressed in cancer and is associated with aneuploidy in cancer cells, suppression of apoptosis and enhancing proliferation, and cell survival[13,14]. Our gene signature also identified the microtubule-associated protein TPX2, which binds AURKA and targets it to the mitotic spindles[15]. A number of inhibitors targeting AURKA have been developed and are currently evaluated in clinical trials[16,17]. In addition, our analysis also identified RAE1, an essential mitotic checkpoint regulator involved in spindle organization[18,19].

A meta-analysis of *ZNF217* and *SALL4* (both identified in our study) concluded that their over-expressions were associated with poor survival across different tumor types[20,21]. Furthermore, *in vitro* studies showed that the overexpression of ZNF217 in mammary epithelial cells drives many of the hallmarks associated with cancer progression, whereas knockdown studies of SALL4 were shown to inhibit proliferation[22,23].

SULF2 is a sulfatase that can edit the sulfation pattern of heparin sulfate proteoglycans on the outside of the cell, thus affecting key signaling pathways[24]. Furthermore, SULF2 was recently shown to be a diagnostic and prognostic marker in lung cancer[25]. Similarly, the overexpression of *CSE1L* on chromosome 20q has been suggested to predict distant metastasis in breast cancer[26]. Mechanistically, CSE1L regulates the association of alpha and beta-tubulin and its increased expression increases the extension of MCF-7 breast cancer cell protrusions, thereby promoting migration[27].

It has been reported that the FAM83D expression is elevated in various cancers[28,29]. Moreover, higher levels of FAM83D expression positively correlate with a poor prognosis in many cancer types[28]. Two recent studies have identified FAM83D as a prognostic marker for hepatocellular carcinoma[30,31]. Furthermore, higher levels of FAM83D have been significantly associated with shorter overall and metastatic relapse-free survival, particularly in patients with estrogen receptor positive (ER+) and luminal subtype tumors[28]. Forced

expression of *FAM83D* in non-malignant cells in culture promoted the proliferation and invasion of breast cancer cells, and down-regulated the expression of tumor suppressor gene, *FBXW7*[32].

The observation that gene expression of the complete 18-gene signature was more significantly associated with cancer patient survival compared to individual genes suggests that multiple genes independently contribute to the overall survival. These data are consistent with the complex pattern of DNA copy number amplifications often observed on chromosome 20q and the presence of multiple driver genes.

Conclusion

In conclusion, our integrative multi-omics analysis of genes on chromosome 20q is paving the way to the development of additional therapeutic targets for cancers with 20q amplifications. This analysis pipeline could furthermore be potentially applied to other tumor amplicons.

Acknowledgements

References

1. Goldman JM, Melo JV. Chronic myeloid leukemia – Advances in biology and new approaches to treatment. N Engl J Med. 2003; 349:1451–1464. doi: 10.1056/NEJMra020777. [PubMed: 14534339]

2. Fausel C. Targeted chronic myeloid leukemia therapy: Seeking a cure. Am J Health Syst Pharm. 2007; 64(24):S9–S15. doi: 10.2146/ajhp070482. [PubMed: 18056932]

3. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, et al. Mutations of the *BRAF* gene in human cancer. Nature. 2002; 417:949–954. l doi: 10.1038/nature00766. [PubMed: 12068308]

4. Larkin J, Ascierto PA, Dréno B, Atkinson V, Liszkay G, et al. Combined vemurafenib and cobimetinib in *BRAF*-mutated melanoma. N Engl J Med. 2014; 371:1867–1876. doi: 10.1056/NEJMoa1408868. [PubMed: 25265494]

5. Scaltriti M, Nuciforo P, Bradbury I, Sperinde J, Agbor-Tarh D, et al. High HER2 expression correlates with response to the combination of lapatinib and trastuzumab. Clin Cancer Res. 2015; 21:569–576. doi: 10.1158/1078-0432.CCR-14-1824. [PubMed: 25467182]

6. Snijders AM, Fridlyand J, Mans DA, Segraves R, Jain AN, et al. Shaping of tumor and drug-resistant genomes by instability and selection. Oncogene. 2003; 22:4370–4379. doi: 10.1038/sj.onc. 1206482. [PubMed: 12853973]

7. Wilting SM, Snijders PJF, Meijer GA, Ylstra B, van den Ijssel PRLA, et al. Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. J Pathol. 2006; 209(2):220–230. doi: 10.1002/path.1966. [PubMed: 16538612]

8. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, et al. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer. 2006; 6:96. doi: 10.1186/1471-2407-6-96. [PubMed: 16620391]

9. Aguirre-Gamboa R, Gomez-Rueda H, Martínez-Ledesma E, Martínez-Torteya A, Chacolla-Huaringa R, et al. SurvExpress: An online biomarker validation tool and database for cancer gene expression data using survival analysis. PLoS ONE. 2013; 8:e74250. doi: 10.1371/journal.pone. 0074250. [PubMed: 24066126]

10. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6(269):pl1. doi: 10.1126/scisignal.2004088. [PubMed: 23550210]

11. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2:401–404. doi: 10.1158/2159-8290.CD-12-0095. [PubMed: 22588877]

12. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): Update 2013. Nucleic Acids Res. 2013; 41(W1):W77–W83. doi: 10.1093/nar/gkt439. [PubMed: 23703215]

13. Katayama H, Brinkley WR, Sen S. The Aurora kinases: Role in cell transformation and tumorigenesis. Cancer Metastasis Rev. 2003; 22(4):451–464. doi: 10.1023/A:1023789416385. [PubMed: 12884918]

14. Warner SL, Bearss DJ, Han H, Von Hoff DD. Targeting Aurora-2 kinase in cancer. Mol Cancer Ther. 2003; 2(6):589–595. [PubMed: 12813139]

15. Kufer TA, Silljé HH, Körner R, Gruss OJ, Meraldi P, et al. Human TPX2 is required for targeting Aurora-A kinase to the spindle. J Cell Biol. 2002; 158(4):617–623. doi: 10.1083/jcb.200204155. [PubMed: 12177045]

16. Malumbres M, Pérez de Castro I. Aurora kinase A inhibitors: Promising agents in antitumoral therapy. Expert Opin Ther Targets. 2014; 18(12):1377–1393. doi: 10.1517/14728222.2014.956085. [PubMed: 25200357]

17. Katayama H, Sen S. Aurora kinase inhibitors as anticancer molecules. Biochim Biophys Acta. 2010; 1799(10–12):829–839. doi: 10.1016/j.bbagrm.2010.09.004. [PubMed: 20863917]

18. Blower MD, Nachury M, Heald R, Weis K. A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. Cell. 2005; 121(2):223–234. doi: 10.1016/j.cell.2005.02.016. [PubMed: 15851029]

19. Babu JR, Jeganathan KB, Baker DJ, Wu X, Kang-Decker N, et al. Rae1 is an essential mitotic checkpoint regulator that cooperates with Bub3 to prevent chromosome missegregation. J Cell Biol. 2003; 160(3):341–353. doi: 10.1083/jcb.200211048. [PubMed: 12551952]

20. Cheng J, Gao J, Shuai X, Tao K. Oncogenic protein SALL4 and ZNF217 as prognostic indicators in solid cancers: A meta-analysis of individual studies. Oncotarget. 2016; 7(17):24314–24325. doi: 10.18632/oncotarget.8237. [PubMed: 27007163]

21. Cohen PA, Donini CF, Nguyen NT, Lincet H, Vendrell JA. The dark side of ZNF217, a key regulator of tumorigenesis with powerful biomarker value. Oncotarget. 2015; 6(39):41566–41581. doi: 10.18632/oncotarget.5893. [PubMed: 26431164]

22. Chen YY, Li ZZ, Ye YY, Xu F, Niu RJ, et al. Knockdown of SALL4 inhibits the proliferation and reverses the resistance of MCF-7/ADR cells to doxorubicin hydrochloride. BMC Mol Biol. 2016; 17:6. doi: 10.1186/s12867-016-0055-y. [PubMed: 26935744]

23. Littlepage LE, Adler AS, Kouros-Mehr H, Huang G, Chou J, et al. The transcription factor ZNF217 is a prognostic biomarker and therapeutic target during breast cancer progression. Cancer Discov. 2012; 2:638–651. doi: 10.1158/2159-8290.CD-12-0093. [PubMed: 22728437]

24. Rosen SD, Lemjabbar-Alaoui H. Sulf-2: An extracellular modulator of cell signaling and a cancer target candidate. Expert Opin Ther Targets. 2010; 14(9):935–949. doi: 10.1517/14728222.2010.504718. [PubMed: 20629619]

25. Lui NS, Yang YW, van Zante A, Buchanan P, Jablons DM, et al. SULF2 expression is a potential diagnostic and prognostic marker in lung cancer. PLoS ONE. 2016; 11:e0148911. doi: 10.1371/journal.pone.0148911. [PubMed: 26882224]

26. Yuksel UM, Turker I, Dilek G, Dogan L, Gulcelik MA, et al. Does CSE1L overexpression affect distant metastasis development in breast cancer? Oncol Res Treat. 2015; 38(9):431–434. doi: 10.1159/000438501. [PubMed: 26278417]

27. Tai CJ, Shen SC, Lee WR, Liao CF, Deng WP, et al. Increased cellular apoptosis susceptibility (CSE1L/CAS) protein expression promotes protrusion extension and enhances migration of MCF-7 breast cancer cells. Exp Cell Res. 2010; 316(17):2969–2981. doi: 10.1016/j.yexcr.2010.07.019. [PubMed: 20688056]

28. Walian PJ, Hang B, Mao JH. Prognostic significance of *FAM83D* gene expression across human cancer types. Oncotarget. 2016; 7(3):3332–3340. doi: 10.18632/oncotarget.6620. [PubMed: 26678035]

29. Varisli L. Meta-analysis of the expression of the mitosis-related gene *FAM83D*. Oncol Lett. 2012; 4(6):1335–1340. doi: 10.3892/ol.2012.925. [PubMed: 23205133]

30. Liao W, Liu W, Liu X, Yuan Q, Ou Y, et al. Upregulation of FAM83D affects the proliferation and invasion of hepatocellular carcinoma. Oncotarget. 2015; 6(27):24132–24147. doi: 10.18632/oncotarget.4432. [PubMed: 26125229]

31. Wang D, Han S, Peng R, Wang X, Yang XX, et al. FAM83D activates the MEK/ERK signaling pathway and promotes cell proliferation in hepatocellular carcinoma. Biochem Biophys Res Commun. 2015; 458(2):313–320. doi: 10.1016/j.bbrc.2015.01.108. [PubMed: 25646692]

32. Wang Z, Liu Y, Zhang P, Zhang W, Wang W, et al. FAM83D promotes cell proliferation and motility by downregulating tumor suppressor gene *FBXW7*. Oncotarget. 2013; 4(12):2476–2486. doi: 10.18632/oncotarget.1581. [PubMed: 24344117]
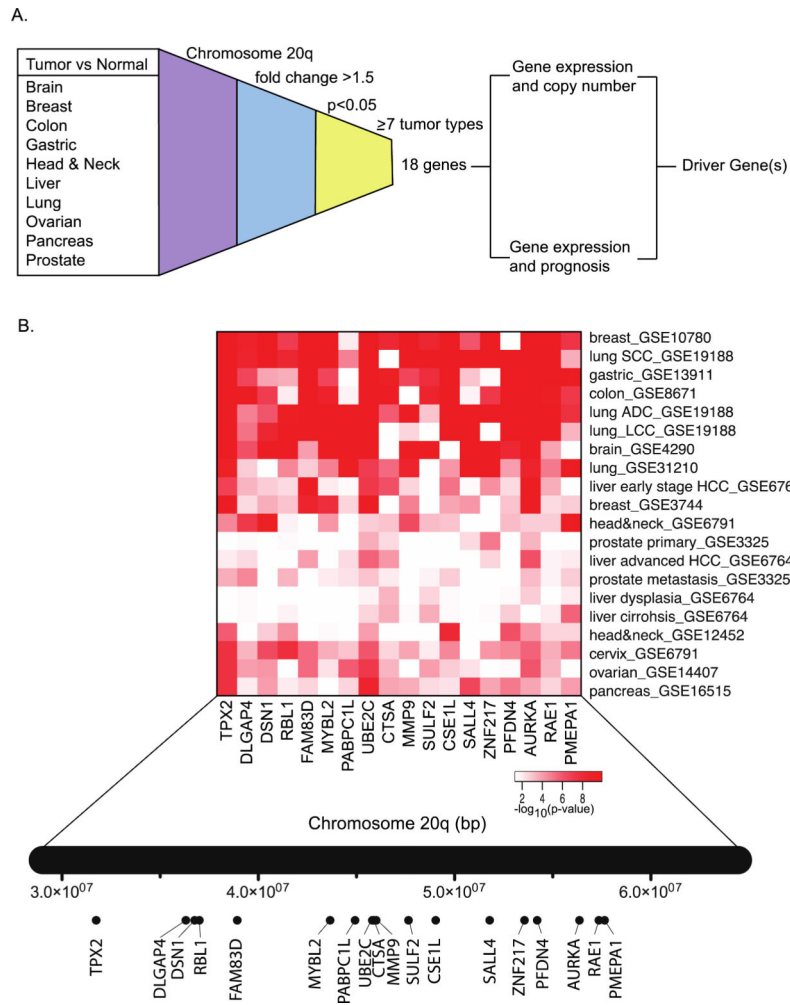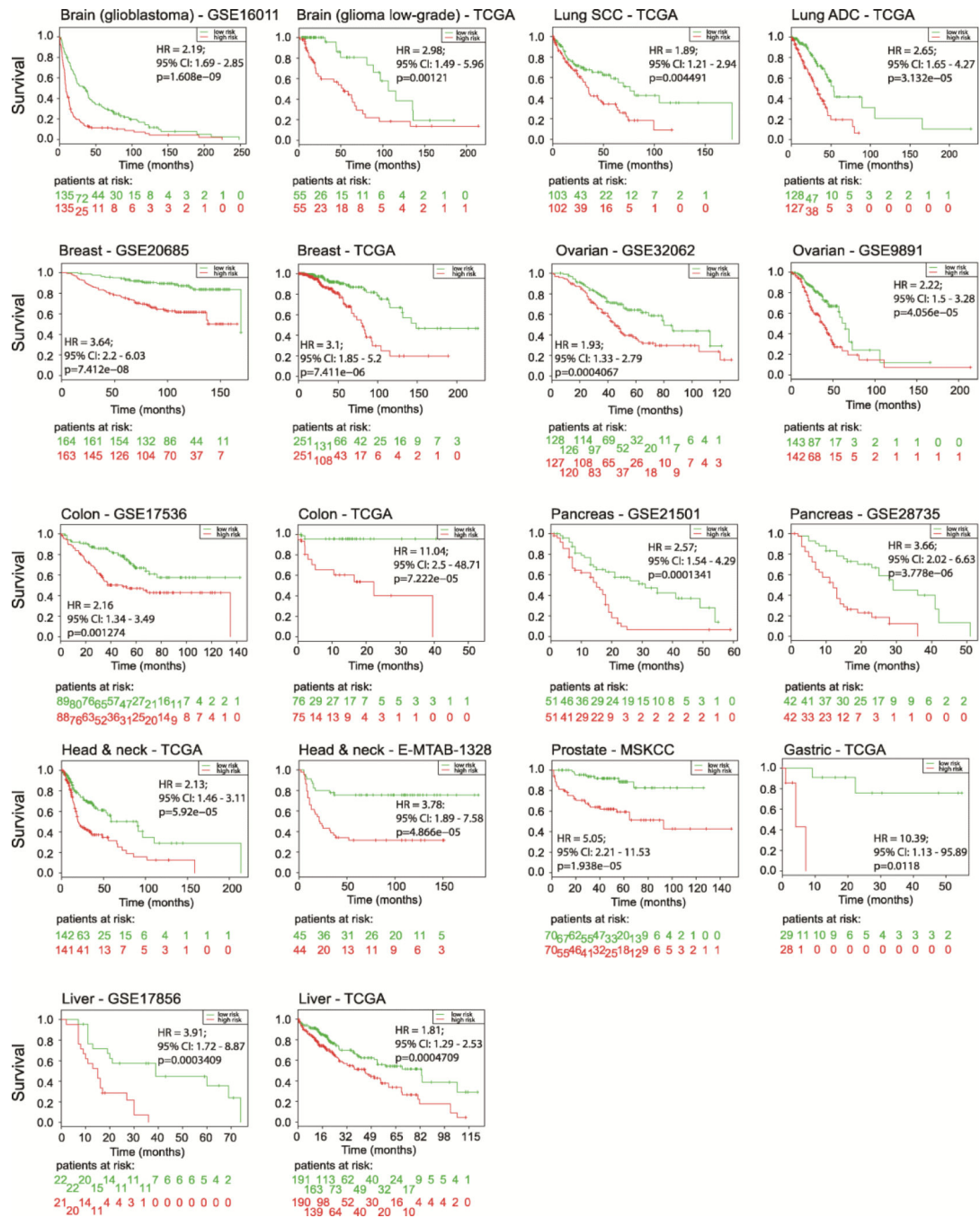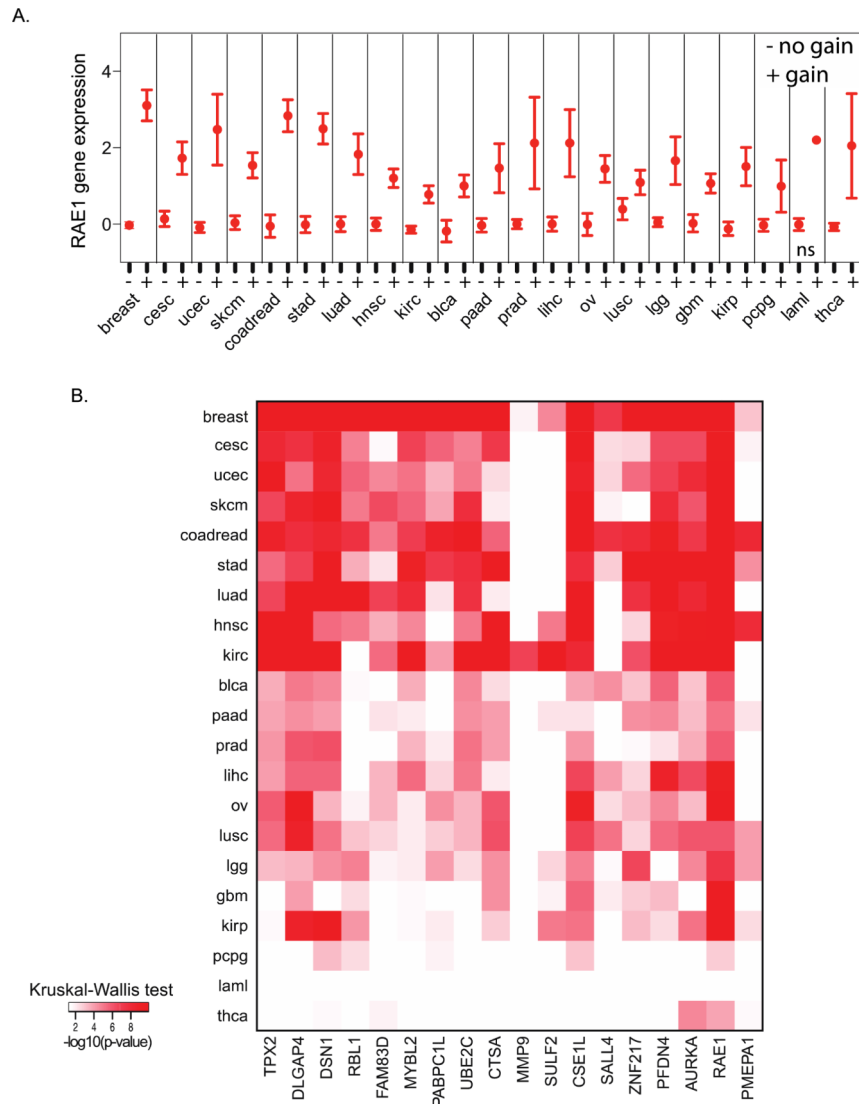
Figure 1.
An 18-gene signature on chromosome 20q is consistently upregulated in human cancers. (A) Gene expression levels were compared between tumor tissues and the corresponding normal tissues. Genes on chromosome 20q that were upregulated (1.5-fold change and $p < 0.05$) in at least seven tumor types were considered for downstream analysis to identify potential molecular targets for cancer therapy. (B) Heatmap of the p-values of the expression difference between tumor and the corresponding normal tissues. The relative location of genes on chromosome 20q is shown underneath.
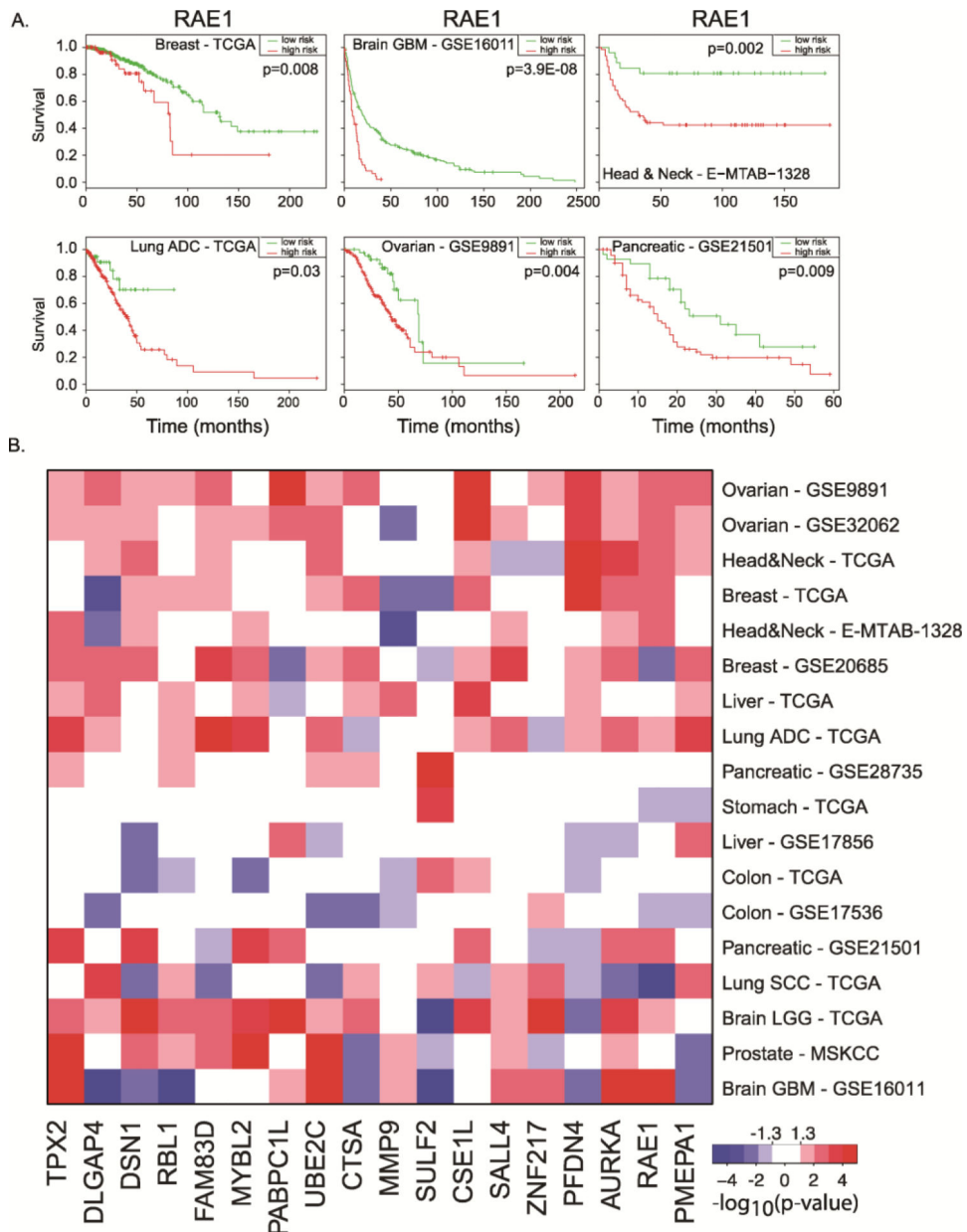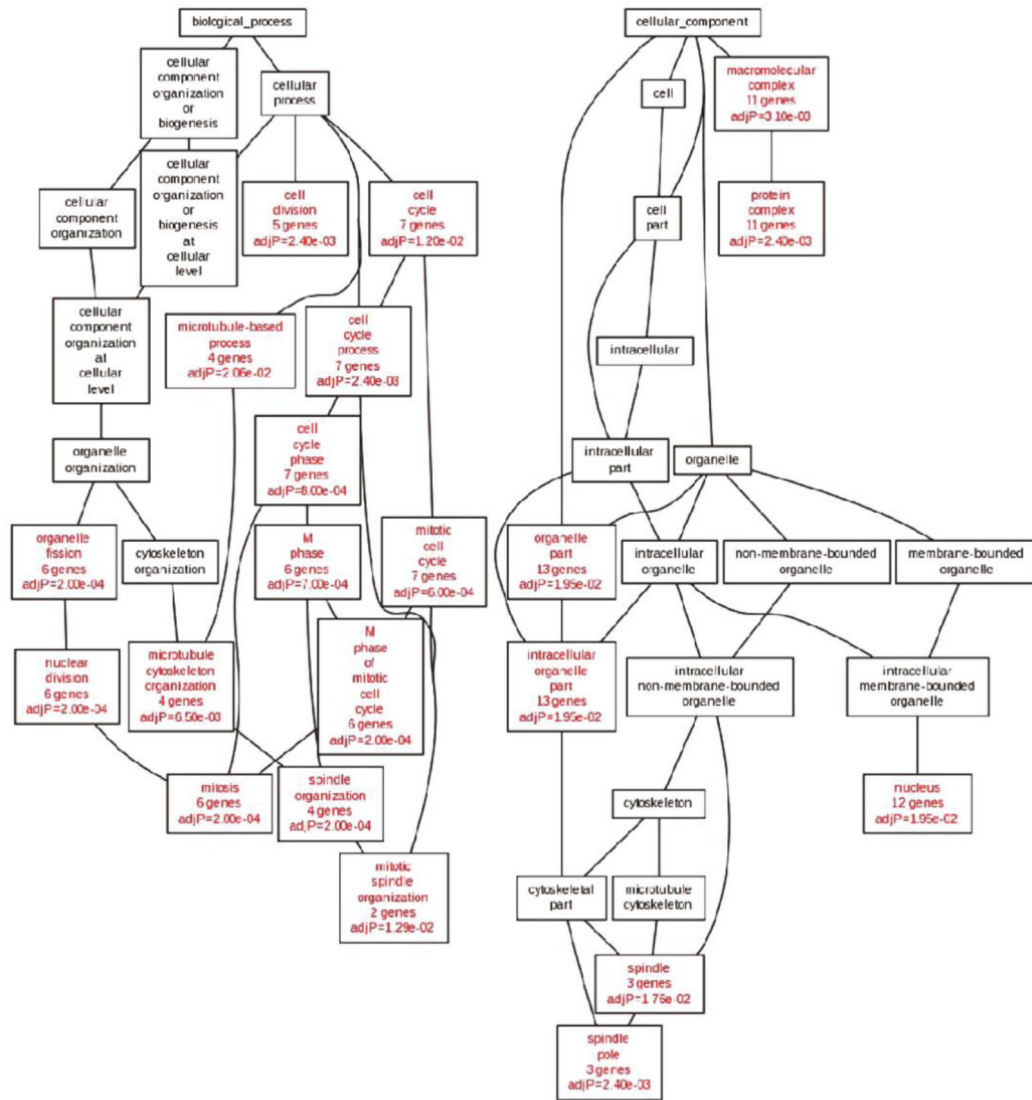
Figure 2.

Individual genes of the 18-gene signature are associated with human cancer disease-free survival. (A) Overexpression of *RAE1* is associated with disease-free survival in different human cancer types. (B) Heatmap of the *p*-values for disease-free survival of individual genes in the 18-gene signature across 18 independent tumor data sets. Genes significantly (*p* < 0.05) associated with poor disease-free survival when upregulated are shown in red, while genes significantly (*p* < 0.05) associated with poor disease-free survival when downregulated are shown in blue.

Figure 3.

The 18-gene signature is associated with human cancer disease-free survival. For each tumor type, survival risk curves are shown; low and high risks are drawn in green and red, respectively. An estimate of the hazard ratio (HR) between groups is shown using the risk group prediction as the covariate in a Cox model (95% confidence interval (CI) is shown). The p-value represents the equality of survival curves based on a log-rank test. Bottom rows represent the number of patients not presenting the event at the specified time.

A.

B.

Figure 4.
DNA copy number of the 18-gene signature is strongly correlated with gene expression. (A)
The relationship between DNA copy number (gain versus no gain) and gene expression for
RAE1 across 21 tumor types. (B) Heatmap of the *p*-values of the relationship between DNA
copy number and gene expression (Kruskal-Wallis test). Tumor type abbreviations – breast:
breast invasive carcinoma; cesc: cervical squamous cell carcinoma and endocervical
adenocarcinoma; ucec: uterine corpus endometrial carcinoma; skcm: skin cutaneous
melanoma; coadread: colorectal adenocarcinoma; stad: stomach adenocarcinoma; luad: lung
adenocarcinoma; hnsc: head and neck squamous cell carcinoma; kirc: kidney renal clear cell
carcinoma; blca: urothelial bladder carcinoma; paad: pancreatic adenocarcinoma; prad:
prostate adenocarcinoma; lihc: liver hepatocellular carcinoma; ov: ovarian serous
cystadenocarcinoma; lusc: lung squamous cell carcinoma; lgg: low grade glioma; gbm:
gliobastoma multiforme; kirp: kidney renal papillary cell carcinoma; pcpg:
pheochromocytoma and paraganglioma; laml: acute myeloid leukemia; thca: thyroid
carcinoma.

Figure 5.
The 18-gene signature on chromosome 20q is significantly enriched for genes associated with cell cycle and mitosis. Directed acyclic graph structure of gene ontology results (biological processes: left; cellular components: right) for the 18-gene signature. Nodes with red labels indicate significantly enriched categories ($p < 0.05$), whereas nodes with black labels represent non-enriched parents.