Deciphering the phenotypic heterogeneity and drug response in cancer cells using genome-wide activity and interaction of chromatin domains.

Neetesh Pandey*[1], Madhu Sharma*[1], Arpit Mathur[1], Chukwuemeka George Anene-Nzel[2,6,7], Muhammad Hakimullah[2], Indra Prakash Jha[1], Omkar Chandra[1], Shreya Mishra[1], Ankur Sharma[2,4,5], Roger Foo[2], Amit Mandoli[3], Ramanuj DasGupta[2], Vibhor Kumar[#1]

1. Department of computational Biology, Indraprastha Institute of Information Technology, New Delhi.
2. Genome Institute of Singapore, A*STAR, Singapore
3. National Institute of Pharmaceutical Education and Research-Ahmedabad (NIPER-A), Palaj, Gandhinagar 382355, Gujarat, India.
4. Harry Perkins Institute of Medical Research, QEII Medical Centre, Nedlands, Western Australia, Australia
5. Curtin Medical School, Curtin University, Bentley, Western Australia, Australia
6. Cardiovascular Disease Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.
7. Montreal Heart Institute, Department of Medicine, University of Montreal, Quebec, Canada

*equal contribution
corresponding author, email: vibhor@iiitd.ac.in

Abstract

Chromatin is organised in the form of domains known to be insulated from neighbouring regions in the genome. The colocalisation of genes sharing similar functions in the topologically associated domains (TAD) and the orchestration of their activity in cancer-cells and their drug-response have not been comprehensively explored. Here we analyzed patterns in the activity of TADs in cancer-cells along with chromatin interaction profiles to understand their modes of drug-response. Using our approach on the transcriptome of 819 cancer cell lines revealed that for multiple drugs, IC50 values were substantially more correlated with the activity of a few TADs than any of the genes. Such a result signifies the potential utility of TAD activity in highlighting common mechanisms of drug response in spite of heterogeneity in the association of genes with drug-resistance. Further we performed analysis of 9014 cancer sample transcriptomes from 20 cancer types to find the

relationship between TAD-activity and survival to highlight context of cause and actionable mechanism for better therapy. Overall our analysis highlighted higher importance and influence of combinatorics of genes in many TADs in comparison to independent effect of their own genes , on the patient-survival and drug-response of cancer-cells.

Introduction

Chromatin organisation and orchestration of activity of different segments of genomes are important influencing factors in the transition of cellular states and response of cells. Multiple groups have investigated the link between chromatin organisation and different properties of cancer-cells [1] and variable mutation rates in their genomes [2]. Jones *et al.* have mentioned that more than 50% of human cancers develop mutations in enzymes involved in chromatin organisation [3]. Even though genetic mutation and escape from apoptosis are involved in the development of cancerous state, regulatory changes in genes through perturbation in the activity of different chromatin segments always become instrumental in such transition. It is evident from previous reports that cancer therapy tends to spare some cancer-cells, even when they lack resistance mutations. [4]. In the past, several studies have tried to find reasons for drug-resistance of cancer-cells. Using mutation and expression profiles, a group involved in Cancer Cell Line Encyclopedia (CCLE) attempted to estimate response for many different drugs [5]. Many other groups have also tried to provide non-genetic explanations using reprogramming and stemness which is linked with increased xenobiotic resistance, in part, due to the efflux pumps expression [6–11], greater DNA repair effectiveness, [6] resilience and stress response [12–14]. While others have attempted to explain it in terms of the heterogeneity of tumors and the capacity of a few cancer-cells to transition states. In every scenario, chromatin organisation plays a crucial role in the growth of cancer and its drug-resistant versions [15,16].

Chromatin is known to be organised in the form of different domains, often called topologically associated domains (TAD). TADs are naturally defined genomic regions that are segregated [17]. TADs have been found to be largely stable across cell-types [18,19] and during differentiation [20]. Besides mammalian genomes, other non-mammalian genomes are also organised in such domains. In fact, a few studies have demonstrated the conservation of a few TAD boundaries across evolution such that approximately 54% of TAD boundaries in the human genome are also

conserved in homologous regions in the genome of mice [18]. TADs are present in the genome, but gene co-localization also exhibits a pattern driven by cellular and evolutionary mechanisms. Such as, a gene-dense locus 3p21.31 on chromosome 3 is known to harbor the largest cluster of tumor suppressor genes. In the past, reports of numerous deletions in locus 3p21.31 on chromosome 3 in a number of tumors have been made (18). The co-localisation and synteny of genes involved in similar or related functions could be due to multiple reasons. One of the reasons is selection pressure to avoid the recombination of beneficial alleles in a locus. Salem *et al.* have shown that paralogs are more often in the same TAD than random pairs of genes [21]. TADs define gene-sets that are activated by eukaryotic cells in an orchestrated manner in different developmental stages and cellular states. The mutation rates in TADS and their boundaries have been studied in the context of the development of cancer and survival analysis [1]. A recent analysis by Akedmir *et al.* [22] revealed that changes in somatic mutational load in cancer genomes were co-localized with TAD boundaries. Several studies have discussed TAD-boundary disruption during the development of cancer. TAD boundaries insulate genes and enhancers and their activity from the effect of neighboring domains [23]. Hence abnormality in activity of genes inside a TAD could be a more direct indication of effective disruption of the TAD boundary. Therefore, we have used an approach to estimate the activity of gene-set consisting of genes inside a TAD to study their role in cancer and drug response.

We initially performed analysis using chromatin interaction profiles of head and neck cancer (HNSC) cell lines. Specifically, we used head and neck squamous cell carcinoma (HNSC) - patient-derived primary cultures (PDCs) for this aim established by Chia *et al.* [24] and used in other studies [25] to understand drug-response of cancer cells. We profiled the chromatin interaction landscape for four HNSC cell lines to comprehend the genomic domain activation pattern in drug resistance. We expanded our study by analyzing the transcriptome of 819 cell lines reported in the CCLE database and response for 544 drugs reported in the CTRPV2 portal. Further, we used our approach to draw inferences from transcriptome profiles of 9014 patients across 20 cancer types made available by the TCGA consortium.

Results

Oral cancer HiC-profiling, TAD patterns, and drug resistance connections

In order to study the chromatin organisation in HNSC cell lines, we performed HiC for four HNSC cell lines. The four HNSC cell lines for which we profiled chromatin interaction are primary oral cancer-cells of three patients' primary tumors (HN148P, HN120P, HN137P) and metastatic cancer-cells of patient HN120 (HN120M) (Supplementary Figure S1). Additionally, we carried out ChIP-seq for H3K27ac and H3K4me3 histone modifications in HN137P, HN120M, and HN148P. We followed the approach of estimating the activity of a TAD by calculating the enrichment of the set of genes contained by it. For single-cell expression profiles, we used UniPath to calculate TAD gene-set enrichment in every single-cell. To calculate TAD activity using the bulk expression profile, we used GSVA [26]. Here we call TAD gene-set enrichment score TAD activity. A brief outline of the work has also been shown in Figure 1A.

TAD activity alterations with cisplatin resistance development

We contrasted the TAD activity scores of oral cancer cell lines with that of their cisplatin-resistant variants. For this purpose, we used the single-cell expression profiles of HNSC cell line of HN137P (paired-end), HN137PCR, HN137M, and HN137MCR published by Sharma *et al.* [27] and have the same genotype [28]. We also used the single-cell expression profile of cell lines HN120M and HN120MCR as they also have the same genotype. Using the Wilcoxon rank sum test and fold change calculation, we found that 20 TADs were downregulated in cisplatin resistant versions of multiple cell lines (HN137PCR, HN120MCR, HN137MCR) (Figure 1B, Supplementary Table 1-6). Three essential TADs (chr11_70150000_71300000, chr11_70200000_71300000, chr11_70225000_71350000) related to keratin-associated proteins were found among the 20 common downregulated TADs. These 3 TADs are localised in chromosomal band 11q13, especially in 11q13.3 and 11q13.4 bands, and have some hierarchy and overlap. The TADs present in 11q13.3 and 11q13.4 band and downregulated in the cisplatin resistant version of all oral cancer cell lines consisting of 5 known keratin associated protein genes (KRTAP5-7, KRTAP5-10, KRTAP5-9, KRTAP5-8, and KRTAP5-11), one actin-binding protein (CTNN) and a few more genes with both known and unknown functions (Figure 1C-D). We also found that 14 TADs were upregulated in cisplatin-resistant versions of 3 HNSC cell lines (HN137PCR, HN120MCR, HN137MCR) using Wilcoxon rank sum test and fold change calculation (Supplementary Figure S2, Supplementary Table 6). We observed that some of the TADs associated significantly with cancer cell progression, and their genes are mainly accountable for the resistance to HNSC. Based

on the genes TMC6 and TMC8, which originate through keratinocytes and lymphocytes defects, respectively, in a TAD's (chr17_76100000_76430000) geneset demonstrates how ion channel transport across the cell membrane plays a vital role in fundamental tumor cells functions like cell volume regulation, cell cycle progression, proliferation, migration, and apoptosis. Human papillomavirus (HPV) infection is strongly related to the development and prognosis of head and neck squamous cell carcinoma (HNSC). The gene transmembrane channel-like 8 (TMC8) has a significant function in T cell modulation and is a major influence on HPV susceptibility [29–31]. When we visualized the chromatin interaction landscape of chr11, we found that the genomic region in 11q13.3 is highly connected to other regions (Figure 1E, regions R1, R2 R3, R4, R5, R6, R7, R8). Such results indicate that TADs in 11q13.3 and 11q13.4 could be part of a larger regulatory condensate. Hence we chose H3K27ac peaks (active enhancers and promoters) present in commonly upregulated TADs and regions connected to them. The motif enrichment analysis for selected peaks showed top motifs for TFs EWS: ERG, FLI1, and ETV2, which belong to the ETS family (Figure 1F). Involvement of ERG and FLI1 in multiple cancers has been reported previously [32]. The FLI1 is itself known to be involved in cell-cycle regulation [33]. We also carried out an examination of gene-ontology enrichment for H3K27ac peaks on chromosome 11 and regions interacting with 11q13.3 and 11q13.4 bands. In addition, cisplatin induced down-regulated TAD in 11q13.4, (with genes KRTAP5-7, KRTAP5-8, KRTAP5-9, KRTAP5-9, KRTAP5-10, KRTAP5-11) and the region R1 connected to it also has many genes which transcribe keratin associated protein (KRTAP5-2, KRTAP5-1, KRTAP5-4, KRTAP5-3, and KRTAP5-5). Thus, HNSC chromatin interaction profile revealed that KRTAP5 family genes are in spatial proximity and possibly in the same regulatory unit whose activity could be induced in response to cisplatin. Among other connected regions, R6 and R7 harbor genes associated with the stress response. Similarly, other connected regions harbor genes with enrichment of function related to stress response and anti-apoptosis (Figure 1G). Overall, TAD activity based comparison among HNSC cell lines and their cisplatin-resistant versions revealed a common hub of chromatin interactions that could be involved in response to stress and cell survival in presence of cisplatin.

Generalization of our understanding of the relationship between drug resistance and TAD by evaluation of CCLE data

After finding significant regulation of activity of a few TADs in HNSC cell lines in presence of cisplatin, we tried to find their association with drug-response in other cancer cell lines. Hence first, we compiled a common TAD list by making a union list of TADs detected in our HNSC cell lines, and TADs reported in the TADKB database. Our approach is based on the reports by multiple researchers on the conservation of TAD boundaries across species and cell-types, which we further verified in cancer cells also. Using the TADKB database, with given TAD boundaries, we found that TAD boundaries across multiple cancer cell-types are also conserved (Figure 2A-B). Such observation encouraged us to use a common TAD list to calculate their TAD activity across 819 CCLE cell lines. We calculated the correlation between the activity of TADs in the union list and the IC50 value of drugs reported in the CCLE database. Using HNSC cell lines for calculating correlation revealed an interesting trend. There was a high correlation between the activity of a few TADs with IC50 value of many drugs (Supplementary Figure S3). Such as, the activity of TAD in location hg19:chr13_58180000_59940000 had a high correlation (R=0.79) with IC50 value of the drug Importazole. The activity of another TAD located at genomic location hg19:chr1_12310000_13770000 was negatively correlated with IC50 (positively correlated with PIC50) value of multiple drugs like Bemcentinib, Dactolisib, and LRRK2-IN-1. In order to have a systematic analysis, we performed clustering of values of correlation between TAD activity and drug PIC50 values (see Figure 2C). We divided the TADs into 8 major clusters and observed a class of TADs whose activity had both a positive correlation with PIC50 values for a group of drugs and a negative correlation with PIC50 values for another group of drugs. As shown in Figure 2C, cluster 3 TAD activity had a negative correlation with PIC50 values of drugs in the D1 group. Whereas TADs in the same cluster-2 also had a positive correlation with PIC50 values of the D2 group of drugs. On the other hand, TADs in cluster-1 and cluster-5 had a negative correlation with PIC50 value of most drugs. In addition, drugs belonging to group D3 showed quite a distinct pattern in each cluster. Whereas drugs belonging to group D4 had a PIC50 value negatively correlated with the activity of TADs in cluster-5 and positively correlated with cluster-2 TAD (Figure 2C).

When we repeated the same analysis using cell lines for all cancer types in the CCLE database, the correlation values between TAD activity and drug PIC50 value decreased (Figure 2D). Such a decrease in correlation values indicates that the association of TAD activity with response of cancer cell towards drugs is influenced by the context of cancer type. We found varying

correlations for the same set of TAD and drugs across different cancer-types (Supplementary Figure S4). Nevertheless, there were still a few TADs whose activity had a notable correlation with drug response even when cell-lines of all types of cancers in the CCLE set were used. We again divided the TADs into 8 clusters based on correlation with IC50 values (Figure 2D). Such as, TADs activity in clusters 3,4,5,6 and 8 exhibited a strong positive correlation with the IC50 value (i.e., a negative correlation with the PIC50) of multiple drugs.

The functions most enriched for genes in TADs of cluster-2 are the formation of the cornified envelope, keratinization, and laminin interactions (Supplementary Figure S5A). Cornification, a type of controlled cell death distinct from apoptosis, is the result of epidermal keratinocytes' terminal differentiation. The mechanism of cornification remains poorly known, although the regulation of cell death in the epidermis has gained widespread interest due to its potential implications for skin diseases [34]. Besides, reductions in keratin expression have been linked to higher metastatic potential, according to numerous studies [35–39]. Additionally, keratins interact with nuclear lamins, which have been demonstrated to bind DNA [40–42]. It is intriguing to consider the possibility that keratin filaments take a role in signal transmission from the extracellular matrix to the nucleus. Downregulation of keratin may cause the extracellular matrix to stop signaling, which would then cause genes related to migration to become active. The desmosome's key structural components are keratin filaments [43,44]. Keratin downregulation can cause desmosomal attachments to break down, which can result in decreased adhesion and an increased inclination to migrate. Berens *et al.* showed that by interfering with KRTAP5-5 activity, the keratin cytoskeleton was disrupted, significantly changing the expression of cytoskeleton regulators and the distribution of cell surface markers. The study also demonstrated that KRTAP-5 reduction from cancer cells resulted in cell blebbing, keratins 14 and 18 loss, in addition to the overexpression of vimentin intermediate filaments [45]. Vimentin has been shown to be expressed at considerably higher amounts in cancer cells that have developed a resistance to treatment, such as tamoxifen-resistant breast cancer [46], GR malignant glioma [47], and prostate cancer resistant to paclitaxel [48]. Henceforth, in addition to its physiologic function, a keratin-associated protein can also act as a switch for malignancy development and therapeutic responsiveness.

A group of TAD (cluster-1) had their activity negatively correlated with IC50 values (positively with PIC50). Gene enrichment for cluster-1 genes highlighted some crucial biological processes involved in drug response, such as negative regulation of interleukin-6 production, Rac protein

signal transduction, NF-kappaB signaling, etc. (Supplementary Figure S5B). It has been demonstrated that people with different forms of cancer have higher levels of interleukin (IL)-6. Also, RAC1 hyperactivation is a frequent occurrence in human cancer and may be brought on by overexpression, aberrant upstream inputs, uncontrolled degradation, and/or altered intracellular localization [49–52]. Thus, overall, the correlation values of TAD activity with drug IC50 for multiple types of cancer cells, highlighted a few common mechanisms of response towards multiple drugs in spite of heterogeneity among different types of cancers.

However, when we used all types of cancer cell lines of the CCLE database, the association between the activity of a few TADs and the IC50 values for many drugs was considerably higher than the same estimate for any single-gene (Figure 2E). This result provides another evidence that our approach of using the TAD gene-set has the potential to highlight the common mechanism of drug response across several cancer types (Supplementary Figure S6 A-B). Thus even if genes involved in response to drugs could vary in different types of cancer the mechanism would have remained the same. Given that many paralogous genes with similar functions often co-localise, our observations could be easily explained. We further tried to model drug PIC50 values using TAD activity scores or gene expression profiles. We found that for many drugs, even if we used multiple genes to model the PIC50 values, the accuracy was much lower than the value achieved by using the TAD activity score (Figure 2F). We also tried to compare the correlation of a given TAD with its respective gene-set to confirm the hypothesis of co-functionality and found that neither of the genes of that particular TAD had a higher value (Figure 2G-H). The observation goes same for both in the case of drug sensitivity and drug resistance (Supplementary Figure S7). Such results indicate that TAD activity based analysis has the potential to highlight common mechanisms of drug response in spite of heterogeneity among cancer-cells. Examples include the drug Ciclopirox, LY-2183240, Lenvatinib, Axitinib, and Foretinib, where TAD activity-based prediction had significantly greater association than the gene-based model.

TAD activity pattern and its influence in TCGA PAN-cancer

There could be multiple reasons for variation in TAD activity among patients, such as i) change in the expression level of genes due to differences in cellular state of cancer cells or mutation at TAD boundary or genomic structural variation which could cause shuffling of genomic regions causing small TAD formation. ii) CNV of the genomic region consisting of the TAD. Nevertheless, each

of these events has an effect on the estimated activity of genes within known TADs. Hence we studied variations in TAD activity using the transcriptome of cancer samples provided by the TCGA consortium. For this purpose, we used a union list comprising TAD detected in HiC profiles in HNSC cell lines and TADs reported in the TADKB database [53]. Initially, we estimated fold change in median TAD activity in tumor and normal samples belonging to a tissue type. We found multiple TADs which had higher activity, specifically according to cancer type. However, there were indeed a few TADs that had higher activity compared to normal samples in multiple cancer types. As shown in Figure 3A, TADs belonging to cluster 12 were more active in multiple cancer types. Further, we tried to find enriched functions for genes in TADs belonging to different clusters. Such as, the genes belonging to the TADs in class 11 had enrichment for cellular glucuronidation (P-value = 0.0000014) containing UGT genes which have a role in cancer progression [54]. The growth of numerous cancers has been linked to the dysregulation of UGT expression and activity, supporting the idea that UGTs may serve as disease-accelerating agents or as mediators of oncogenic pathways in the absence of therapeutic intervention [55].

Since TAD activity seems to have an important role across all types of cancer. We performed a PAN-cancer analysis of the influence of TAD activity on survival. We found that fractions (or number) of TADs associated with survival in 2 and 3 cancer types were more than the TADs with survival-association in only one cancer type (criteria: P-value < 0.05) (Figure 3B). The set of TADs that appeared to be associated ($P < 0.05$) with survival in more than two cancer types also included TAD (chr11_70150000_71300000) from 11q13.4 region (Figure 3C). Further, we tested the hypothesis that such results could be artifacts of the effect of a few genes which could be independently more influential for survival than the activity of its host TAD. We found that for multiple TADs, their activity had a higher influence on survival than any of the genes inside them (Figure 3D-G). Such as, the TAD at 11q13.4 region had a significant association with survival in HNSC patients, but none of its genes had a comparable influence (Figure 3D). Similar results about more such TADs are shown in Figure 3D. Such results indicate that survival in cancer could be influenced by the combinatorial effect of genes (or their function) inside a TAD.

We further tried to find coherence between the association of TAD activity with response to drugs in cancer cell lines (CCLE database) and survival in a cohort of patients in TCGA. We focussed on results based on association with TCGA-HNSC patients and correlation with drugs in HNSC cell-lines in the CCLE database. There were many TADs that were found to be negatively

associated with drug PIC50 value and had a significant association with survival (P-value < 0.05). We sorted the TADs according to the number of drugs with which they had an absolute correlation values higher than 0.3 for HNSC cell lines in the CCLE database. We called TADs associated with survival (P-value < 0.05) and with the top 10 most frequently positive correlations (> 0.30) with PIC50 values of drugs as top10_positive. Similarly, we called TADs associated with survival and the top 10 most frequent negative correlation with PIC50 value of drugs as top10_negative. Similarly, we made groups of top 50 and top 100 positive and negative TADs. We observed the hazard-ratio distribution of TADs across different categories of top 10, 50, and 100. We found that in each category, the median hazard ratio of TADs with negative correlation with PIC50 values was above 1 and vice-versa for TADs (Figure 3H) with positive correlation. In order to further confirm the observation, we also looked at the frequency of TADs with hazard cut-off greater than 1.25 (Figure 3I) and less than 0.75 (Figure 3J). The result showed that most of the negatively correlated TADs (Negative_set) had a hazard ratio above 1.25 (low survival), whereas positively correlated TADs (Positive_set) were mostly having a hazard ratio below 0.75 (high survival). Thus TADs appearing as negatively correlated with sensitivity for multiple-drugs in cancer cell lines also had an association with low survival rate in cancer patients.

The association between TAD activity and survival across several cancer types revealed an interesting trend. For HNSC, the TAD in 11q13.4 region, whose activity was also significantly associated with survival harbor KRTAP5 family genes (KRTAP5-7, KRTAP5-10, KRTAP5-11) involved in keratinization (Figure 4A). Our findings showed that the keratinization pathway in HNSC was strongly linked to a higher survival rate. Among the other major TADs whose activity was also related to HNSC survival, a TAD (chr7_51500000_57100000) containing genes (especially EGFR) involved in EGFR associated pathways was also present (Figure 4B). A term "PTK6 promotes HIF1A stabilization" was also found enriched for genes in the TAD (chr7-51500000_57100000). Protein tyrosine kinase 6 (PTK6) is a cytoplasmic non-receptor tyrosine kinase that is abundantly expressed in a variety of tumor types, including breast- carcinomas, colorectal-carcinoma, prostate-carcinoma, lung-cancer, head and neck-carcinoma, and B and T lymphomas. Previous research demonstrated that PTK6 in breast and colorectal cancer cell lines was rapidly stabilized in hypoxic circumstances in a posttranslational HIF-1-independent mechanism. PTK6 was intended for the proteasome at normal oxygen levels, but this process was suppressed in hypoxia. Enhanced hypoxia-dependent migration and invasion followed PTK6

activation driven by hypoxia [56]. EGFR activity is itself associated with tumor growth[57]. However, as shown in Figure 3G, the association of EGFR expression with survival did not have a significant P-value. Thus the combinatorial effect of function of genes in TAD (chr7_51500000_57100000) appears to have a more profound effect on the survival of HNSC and CESC patients (Figure 4B, Supplementary Figure S8A, Supplementary Figure S8C) than any of the genes it hosts.

Another intriguing TAD (chr3_97760000_98290000) which appeared to be associated with the survival of HNSC patients, has olfactory receptors (OR), a class of genes that belong to a major sub-family of G protein-coupled receptors. The widespread expression of ORs in cancerous cells and tissues could be exploited for potential biomarkers for the disease (Figure 4C) [58]. Recent studies of patients survivors of head and neck cancer have connected altered taste and smell to olfactory receptor activity [59]. Yet another intriguing TAD (chr1_156110000_156900000) was likewise associated with HNSC survival. It harbors genes for which enriched pathway terms are "signalling to P38 via RIT and RIN" and "NGF-independent TRKA Activation" (Figure 4D).

We also performed an analysis of the association between TAD activity and survival for other types of cancer using TCGA transcriptome profiles. The results of CESC and LUSC cancer are provided in supplementary figures (Supplementary Figure S8 A-C). HiC visualisation of survival-associated TADs was also done to explore 3-dimensional associations of such TADs with other different regions in HNSC cell lines (Supplementary Figure S9). Overall, our analysis using TAD activity, revealed its potential to highlight the possible effect of co-localisation of genes in influencing cancer and possibly a few unknown mechanisms or conserved coordination of genes towards a major biological process which influence survival in cancer.

Furthermore, to assess the significance of individual genes of respective TAD, we took one of the top HNSC survival associated TAD i.e., chr11_70150000_71300000 (HNSC-TAD-1) containing some important genes such as CTTN, SHANK2, KRTAP5-7, KRTAP5-8, KRTAP5-9, KRTAP5-10, KRTAP5-11 etc. Subsequently, we performed survival analysis using the KRTAP gene family and KRTAP5 geneset in HNSC patients. It was found that the association between KRTAP5 gene-set and HNSC patient survival was not significant enough to point toward the absolute role of these genes (P-value =0.054). On the other hand, HNSC patient survival had a noticeable association

with gene-set of KRTAP family (P-value = 0.0043). However, it was not as significant as HNSC-TAD-1 comprising its full geneset (P-value = 0.0013) (Figure 5A-C). Such a result highlighted the fact that there could also be other genes in HNSC-TAD-1 that could be supporting the survival of HNSC patients in addition to KRTAP5 genes. It also hints about a possible evolutionary force for the co-localisation of genes in HNSC-TAD-1, supporting a few mechanisms supplementing each other for a response towards a common biological role. Hence the TAD activity-based approach highlighted the cumulative effect of genes in TAD with KRTAP5 family on survival, even if the HNSC cells in different patients could have a heterogeneous expression of KRTAP5 genes (Figure 5D). There could be effect of conditional dependencies in the associations found by our survival based analysis, therefore we also used Bayesian network modelling where every TAD was represented by a node (see Methods). Bayesian analysis of HNSC cancer patients also highlighted the importance of the same survival-associated TADs (including HNSC-TAD-1) as a crucial factor associated with the risk factor of survival (Figure 5E-F).

Discussion

Here, we have used a novel approach of calculating TAD activity using gene expression profiles. TAD activity is also indicative of the activity of other regulatory elements in a TAD such as enhancers. One of the reasons of our approach is that there are multiple mechanisms involved in formation of TAD boundaries. Barutcu et al. have shown that deleting CTCF-rich reason near TAD boundaries also do not disrupt the TADs [60]. The high number of conserved interactions within TADs also indicates that TADs are regulatory units. Hence our approach could be truely a direct way to estimate effect of disrupted TAD based regulatory units. Using our approach in the context of drug response of cancer-cells could also highlight the common mechanism of action. Such as, by examination of the changes in TAD activity brought on by cisplatin in HNSC cell lines, we recognised a chromosome 11 region in the 11q13.3–11q13.4 band. Our analysis has also added a dimension of chromatin structure to the role of TADs in 11q13.3-11q13.4 bands in providing robustness to cancer-cells towards multiple drugs. Our results indicate that TADS in 11q13.3 and 11q13.4 bands could be harboring enhancers or genes helpful in activating or regulating other connected regions which have stress-response genes. Most notably, a region interacting with TADs at 11q13.3 and 11q13.4 band also harbor many genes for keratin associated proteins. Keratin-associated proteins have not been commonly reported for association with

properties of cancer-cells except by a few anecdotal studies [45]. It could be possible that clusters of genes for keratin-associated proteins and other response genes co-localise for coexpression for a major biological process.

Our analysis using TCGA HNSC samples revealed many TADs associated with survival have genes for enriched pathways such as keratinization, EGFR-PTK6 based HIF1A stabilization, expression and translocation of olfactory receptors, PI3K-AKT activation. Clusters of multiple paralogous genes associated with HNSC survival were found, suggesting their importance in physiological and pathological roles via evolutionary conservation. Our analysis clearly indicates that TAD activity could be used as a biomarker for predicting drug response or survival estimation to cope with heterogeneity among tumors.

The current paradigm of chemotherapy for cancer involves multiple kinds of drugs and their combinations, as studied by different consortiums like CCLE and TCGA. Developing a framework for selecting effective drug combinations to prevent redundancies can be aided by the common mechanism revealed by TAD-based analysis. Our unbiased analysis of linking TAD based regulatory units to cancer survival highlighted a few mechanisms which could be targeted for chemotherapy. There is also another benefit of our approach of analysis; it could narrow down the search space for the identification of cis-regulatory elements like super-enhancers regulating multiple genes in a relevant TAD for the purpose of cancer therapeutics.

An interesting result that emerged in our analysis is that for many TADs, which are either associated with drug response or survival, the genes within their boundaries did not seem to be influential individually. Such results highlight the effectiveness of evolutionary conservation of synteny of genes in TADs to support collaborative and combinatorial action of genes towards some common function. Expansion of our analysis to multiple kinds of regulation analysis in normal cells could reveal many such conserved functional gene-sets naturally defined by TAD boundaries. Thus in the future, we will further expand our analysis to highlight TADs associated with different kinds of cellular states and their responses.

Methods

Experimental method

The Genome Institute of Singapore performed the culture of HNSC patient-derived cell lines (HN137P, HN120P, HN120M, and HN148P) as well as their ChiP-seq and Hi-C based profiling and sequencing.

Culture of HNSC PDC and their cisplatin-resistant model

The creation of the PDC model of HNSC is described by Chia *et al.* [24]. Cells were kept in a humidified atmosphere of 5% CO_2 at 37 °C. The identity of Cell line identity was authenticated and confirmed by Chia *et al.* through a comparison of STR profiles cell lines to their original tumor. The cisplatin version of the same HNSC cell lines was made by Sharma *et al.* by treating mice xenograft model with cisplatin for 3 weeks. The cells dissociated from the residual xenograft model were isolated into single-cell suspension followed by flow-cytometric analysis of cell-type and stem-state markers. Flow cytometry data were analyzed by FlowJo.

Hi-C profiling

In situ Hi-C was performed as previously described [61] (PMID 30717603). Five million HN137 and HN120 cells were cross-linked with 1.5% formaldehyde for 10 minutes at room temperature, after which glycine was added to stop the reaction. Cross-linked cells were lysed in 500 µL cold lysis buffer (10 mmol/L Tris-HCl, pH 8.0, 10 mmol/L NaCl, 0.2% (vol/vol) Igepal CA630, mixed with protease inhibitors) on ice for 30 minutes. The pellet was washed twice with cold 1× NEBuffer 2, and chromatin was solubilized in 50 µL 0.5% SDS and incubated at 62°C for 10 minutes. SDS was quenched using Triton X-100 (Sigma 93443), and chromatin was digested overnight by adding 25 µL 10X NEBuffer 2 and 100 U Mbo I (New England Biolabs, catalog No. R0147). Digestion efficiency was confirmed by isolating DNA from a 25-µL aliquot of the suspension and performing agarose gel electrophoresis. Digested ends were filled and labeled with biotin by adding 37.5 µL 0.4 mmol/L Biotin-14-dATP (Thermo Fisher Scientific, catalog No. 19524016), 1.5 µL 10 mmol/L dCTP, 1.5 µL of 10 mmol/L dGTP, 1.5 µL of 10 mmol/L dTTP, 8 µL of 5 U/µL DNA polymerase I, and large Klenow fragment (New England Biolabs, catalog No. M0210); and incubated at 37°C for 90 minutes. Blunt-end ligation was performed by adding 463 µL water, 120 µL 10X NEB T4 DNA ligase buffer, 200 µL 5% Triton X-100, 12 µL 10 mg/mL BSA, and 5 µL T4 DNA Ligase (New England Biolabs, catalog No. M0202). The solution was incubated at 16°C for 4 hours. After ligation, cross-link was reversed overnight, and DNA was purified with the phenol-chloroform method. Next, Hi-C libraries were treated with T4 DNA

Polymerase (New England Biolabs, catalog No. M0203) to remove biotin from unligated ends. To do this, 5 µg Hi-C library was combined with 2.5 µL of 1 mmol/L dATP, 2.5 µL of 1 mmol/L dGTP, 10 µL of 10X NEBuffer 2.1, and 5 µL of T4 DNA Polymerase (New England Biolabs, catalog No. M0203), and water was added to a final volume of 100 µL. The reactions were incubated at 20°C for 4 hours and stopped by incubating at 75°C for 20 minutes. Samples were sheared to a mean fragment length of 400 base pairs (bp) using Covaris M220. Pull-down of biotinylated DNA was performed with Dynabeads My One T1 Streptavidin beads (Thermo Fisher Scientific, catalog No.65601), and library preparation was performed using NEBNext Ultra II DNA Library Prep Kit for Illumina. Samples were quantified with KAPA, and sequencing was performed on a Hiseq 4000.

ChIP-seq profiling

Cells were cross-linked for 30 minutes (10% FBS/DMEM containing 1.5% PFA) followed by the addition of glycine (1/20 volume of 2.5M) and incubation on ice (5 min.). It was re-dissolved in a buffer followed by sonication. The buffer used for re-dissolving contained 10 mM Tris-HCl, pH8, 1 mM EDTA, 0.5 mM EGTA, and 1X protease inhibitor cocktail (Roche). For pulling the chromatin H3K4me3 or H3K427ac antibody-coupled Dynabeads and 150 µl of ChIP buffer (433 mM NaCl, 0.43% sodium deoxycholate, 4.3% Triton X-100) were used. For the input control library, 1/10 of the chromatin mixture was kept separately. RIPA buffer was used to wash the beads which were later incubated in 100 100 µl of elution buffer (50 mM Tris-HCl, pH8, 1% SDS, 10 mM EDTA) at 65 °C for 20 min with brief vortexing. De-crosslinking of immunoprecipitated chromatin and input samples was done by incubating at 65 °C overnight and subsequently incubating with 0.2 µg/µl of RNase A at 37 °C for 1 h and with 0.2 µg/µl of proteinase K at 55 °C for 2 h. DNA was isolated from the chromatin using phenol/chloroform/isoamyl alcohol (49:49:2) and ethanol and precipitated in presence of 40 µg of glycogen. Pellet was then suspended in 10mM Tris-JCL, pH8 after washing with 70% ethanol. Library preparation and multiplexing for ChIP-seq libraries were performed using the NEBNext ChIP-seq library Prep kit (NEB). The sequencing of Chip-seq and input control libraries was performed using Illumina HiSeq-2500.

Computational analysis

Analysis of HiC profile

To construct genome-wide maps with millions of interactions between genomic locus pairs, Hi-C creates hundreds of millions of read-pairs [62]. In our analysis HiCUP tool was used to process the Hi-C data. Sequence data from experiments using the Hi-C and Capture Hi-C (CHi-C) technologies are usually processed using the HiCUP pipeline to examine the three-dimensional structure of the genome. Data are mapped to a particular reference genome via the same pipeline, which also removes artifacts that would otherwise impede future study [63,64].

Analysis using HiCUP pipeline

In HiCUP pipeline involves numerous programs like hicup_truncater, hicup_mapper, hicup_filter, hicup_deduplicator, and hicup_digester- which prepares the genome digest file used by hicup_filter. The paired FASTQ files were used to generate Hi-C di-tag paired reads, which were then mapped to the hg19 reference genome. This HiCUP script controlled the pipeline carrying output from one section of the program to the next and executed each script sequentially. The configuration file for the HiCUP program also contains the parameters for the complete pipeline. The configuration file specified the names of the FASTQ file pairs that needed to be processed [63].

While the great majority of reads produced by the HiCUP Mapper script are undoubtedly authentic Hi-C products, a substantial minority are likely not and should be eliminated. For the purpose of finding legitimate Hi-C pairs, the HiCUP filter script analyses paired reads along with the file produced by the HiCUP Digester. This might not necessarily be the case, though, as the sequenced region may contain the Hi-C ligation junction. As a result, potentially useful data will be lost when such reads are eliminated from the Hi-C pipeline during the mapping process. By locating ligation junctions among reads and removing sequences downstream of the restriction enzyme recognition site, the HiCUP truncater script helps to fix this [63].

Pipeline implementation

For computational analysis, we used the DpnII digest profile of the hg19 reference genome created using hicup_digester. This file represented all possible DpnII fragments in the genome and was used to identify Hi-C artifacts. The alignment process was performed using HiCUP v0.5.10 with Bowtie2 (hg19), and the filtering step's minimum and maximum di-tag range were set to 150 and 1000, respectively. The default settings for the other parameters were used.

Generation of HiC contact matrix

We first converted the binary alignment mapped file into arrowhead input format using the samtools view command to create the HiC contact matrix [65]. In order to extract the necessary fields from the bam file, we also utilized the awk command. The normalized Hi-C contact matrix (.hic file) was created using the Juicer Tools Pre command with the help of the arrowhead input file. The dump command of Juicer Tools was used to export the observed contact count for each region of the Hi-C map with balanced Knight-Ruiz (KR) normalization. With this, we were able to obtain the text Hi-C matrix (.txt) at 25 kb resolution for all the chromosomes.

Detection of topologically associated domains (TADs)

We used the DomainCaller approach from Dixon *et al.* to identify topologically associated domains at 25 kb on all chromosomes [18,66]. We started with a square contact matrix of the 25 kb HiC contact map, and subsequently directionality indices were created using DI_from_matrix.pl with a bin size of 25,000 along with a window size of 125,000. This vector of directionality indices was provided as input for the HMM_calls.m script. Further, we obtained TAD calls by processing HMM via file_ends_cleaner.pl, converter_7col.pl, hmm_probablity_correcter.pl, hmm-state_caller.pl, and, hmm-state_domains.pl. The HMM probability correction script was run with the following parameters as input: min = 2, prob = 0.99, and bin size = 25,000 [66]. Secondly, we also got TADs from the TADKB database, which gives TADs at resolutions of 50 kb and 10 kb from a variety of cell types, namely human HMEC, NHEK, GM12878, IMR90, K562, HUVEC, and KBM7 [19,53]. TAD locations for each cell type were obtained with the help of the Directionality Index (DI) [53].

Analysis of ChIP-seq profiles

Analysis of ChIP-seq profiles was done using DFilter [67] to call peaks as well as make custom tracks. The parameters used for calling peaks were according to the instructions provided on the DFilter help page (https://reggenlab.github.io/DFilter/).

TAD gene-set procurement and calculation of TAD activity score

The TADs gene-set was constructed using genes found in certain genomic locations on the hg19 genome version after intersecting with the corresponding promoter regions. The widespread availability of big data sets from multiple assays with numerous complex phenotypes inspired us to study the functional mechanism of TADs in the context of transcriptomics. To facilitate the interpretation and organization of TAD-genesets in large microarray experiments, gene set enrichment (GSE) was done. Using a framework called gene set enrichment analysis (GSE), it is possible to summarise a signature or a pattern from gene expression profiles. Compared to GSVA single gene analysis, this approach has the advantages of noise & dimension reduction, as well as better biological interpretability [68]. GSVA tool was used to evaluate the gene set enrichment variation within a sample population [26]. GSVA is an unsupervised and non-parametric technique that eliminates the traditional strategy of directly modeling phenotypes within the enrichment score.

We defined TADs activity scores as enrichment scores of its gene-set so that we could better understand biological events. The Gene Set Variation Analysis (GSVA) R-package was used to calculate TADs activity scores on log2(10*normalized expression matrix + 1) with its own constructed TADs gene set file. For TCGA (https://portal.gdc.cancer.gov/) [69] patient data, RNASeq gene expression profiles of 20 cancer types out of 24 cancer types (PANCANCER) were transformed into TADs activity scores. The data from the Cancer Cell Line Encyclopedia (CCLE) with 819 samples from different cancer cell lines were used to calculate TADs activity scores. 819 CCLE-cancer cell lines were chosen based on their coherence with 544 drugs listed in the Cancer Therapeutics Response Portal (CTRPV2). We used the single-cell expression profiles of HNSC cancer cell lines published by Sharma et al. [27] to calculate TADs activity scores using UniPath R-package [70]. The gene enrichment scores thus obtained were analysed for differential activity among resistant and non-resistant cell lines using Wilcoxon rank-sum test (Supplementary Table 1-4).

Calculation of survival P-value

To calculate the significance of association of TAD activity and gene expression with survival for TCGA patients, R packages survival and survminer were used [71]. Input files were GSVA scores and clinical data containing days_to_death, gender, etc where clinical data was fetched from the Xena browser [72]. 'Status' column 0 or 1 in clinical data stands for alive and dead, respectively.

For calculating the survival P-value for each TAD, the median of each row of GSVA was calculated, and another "median-group" column was created, signifying 1 for GSVA greater than the median and 0 for GSVA less than the median. the clinical information (days-to-death, status (0/1), and patient ID) was merged with GSVA scores and median-group column. Hence for each TAD, we created a file from which data was fit in the 'survfit' function from R package survival and survminer. Survfit output includes survival P-value for each of the concerned TAD. In addition, we also calculated the hazard ratio for each TAD using the coxph function. Different cut-off combinations were implemented to identify TADs significance in survival (high-low), such as 51-49, 60-40, and 70-30 percentiles to define the group of patients based on the activity of a TAD, in addition to the MaxStat package [73]. Subsequently, TADs which were common among such combinations, were selected for further analysis (Supplementary Figure S10).

In order to study the association of patient survival with drug resistance, we created two sets of TADs i.e., Positive_set and Negative_set, with correlation above 0.3 and below -0.3 respectively, and examined their survival hazard ratio in multiple categories. Furthermore, we also explored the distribution of TADs across different cut-offs of hazard ratio.

Bayesian network-based analysis of survival in head and neck squamous cell carcinoma (HNSC)

Patient, wise matrix from TCGA-PANCANCER, with risk factor (0 or 1) and TAD GSVA score was made for each cancer type. Here the risk node deteremines whether the patient is at high risk (risk = 1, days to the death of the patient < m, where m is the median of days to death of all patients) or at low risk (risk = 0, days to death of the patient > m). Top 200 TADs with significant survival P-value along with the risk factor column were used with the bnlearn package [74] as input to learn a bayesian network. We used the Bayesian Information Criterion (BIC) as a score function along with the Hill Climbing (HC) structure learning method from bnlearn. Thereafter, we bootstrapped the learning and majority voting across 101 BNs to strengthen the graphical model. The data was discretized based on quantiles to reduce the computational time required to learn the network.

Figures

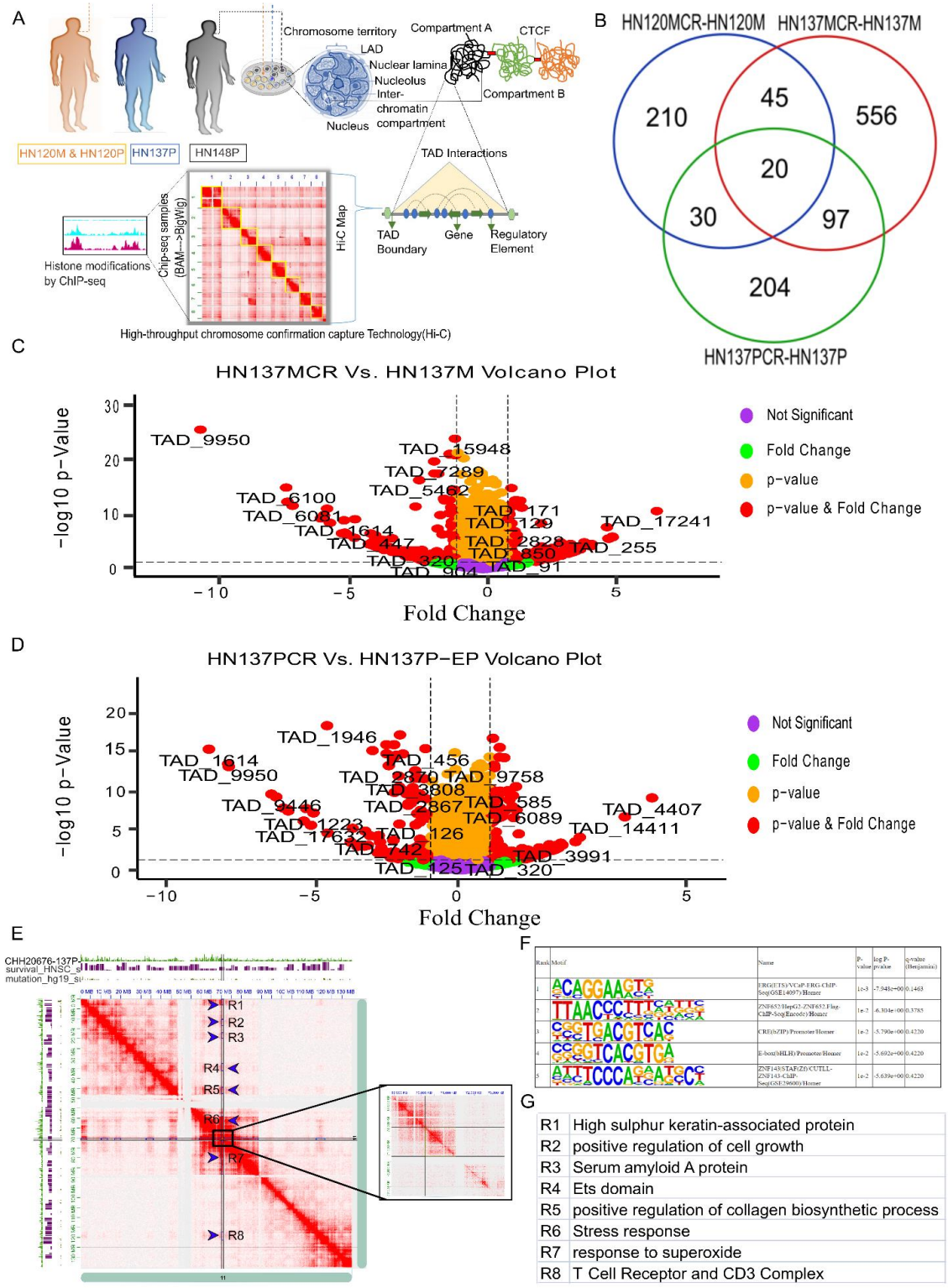


Figure 1: Analysis of chromatin interaction profiles and domain activity of head and neck cancer patient derived cell lines. A) A description of the profiling and analysis of the chromatin

interaction profiles of four HNSC cell lines. Two of the cell lines are from the same patient HN120, while the other two are models of primary cancer-cells from patients HN137 and HN148. The HiC based chromatin interaction profile from HNSC cell lines was used to extract TAD boundaries and further downstream analysis. **B)** The TAD activity was calculated using expression profile of single cells for 3 HNSC cell lines and their cisplatin-resistance version. The Venn diagram shows the number of TADs downregulated in cisplatin-resistant versions of HNSC cell lines. For upregulated genes see supplementary Figure S2. **C)** The fold change and significance (P-value) of the TAD activity between 137P and its cisplatin-resistant version 137PCR are depicted on the volcano plot. **D)** An volcano plot of fold change and significance (P-value) for TAD activity differences between the 137M cell line as well as its cisplatin version. **E)** The interaction matrix visualisation for chromosome 11. The TADs (in the 11q13.3 band) downregulated in all three HNSC cell lines is the present region inside the rectangle indicated in the middle. Other regions with distal interaction with the TAD in 11q13.3 band are shown with arrows as R1, R2, R3 .. R6). **F)** The enriched motifs in H2K27ac peaks in down-regulated TAD in 11q13.3-11q13.4 band and R1, R2 …. R6 region. **G)** The enriched gene-ontology terms of genes present in different regions interacting with 11q13.3-11q13.4 bands.

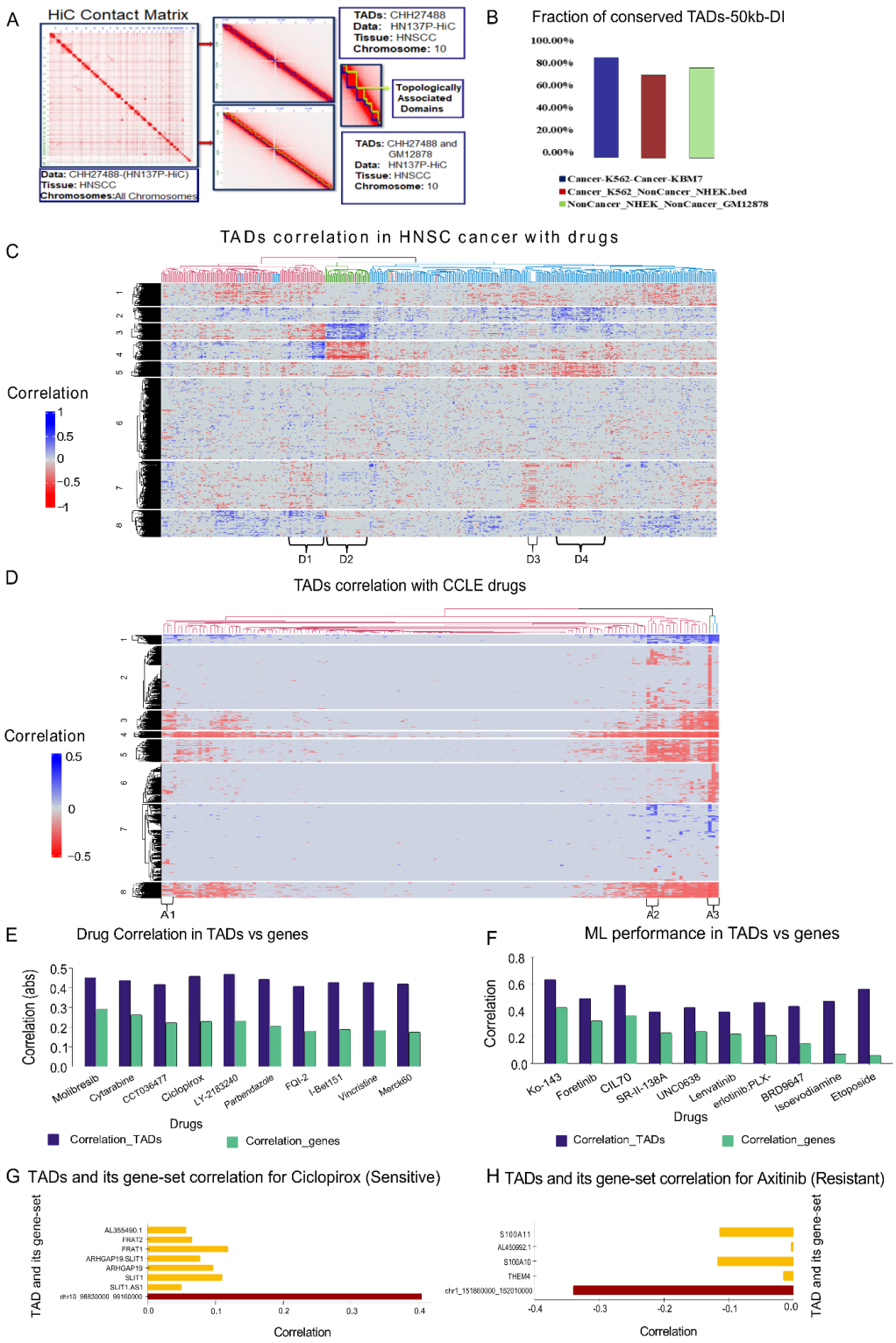


Figure 2: Conservation of TAD boundaries and pattern of correlation between TAD activity and drug response. A) Visualization of TAD boundaries and chromatin interaction from two

different cell lines (GM12878 and HN137P). A Zoomed image of chromatin interaction on chromosome 10 is shown for two cell lines with their TAD boundaries (green and blue). The TAD boundaries are almost similar in the two cell lines. **B)** Overlap of TAD boundary regions (within 1 KB from the boundary) among different cell-types. The Y axis shows the percentage of boundaries overlapping among two cell-lines. **C)** Heatmap illustration for correlation among TAD activity and PIC50 value for drugs in HNSC cell lines from CCLE database. Only those TADs (in rows) are shown whose activity has an absolute correlation value of more than 0.3 with PIC50 value of at least one drug. Similar to this, only drugs that have an absolute correlation value of more than 0.3 with the activity of at least one TAD are displayed. Four sub-clusters of drugs are also highlighted here. **D)** Heatmap of correlation between TAD activity and PIC50 value of drugs for cell lines from all cancer types present in CCLE database. Only those TADs (rows) or drugs (columns) are shown which have atleast one absolute correlation value of more than 0.3. A few sub-clusters of drugs are also shown. **E)** Bar plot of best positive correlation achieved for few drugs between their PIC50 values with TAD activity (blue) or single gene expression (green). **F)** The actual and predicted PIC50 values using a machine learning model are shown in a bar chart where the features were either TAD activity or gene expression. For many drugs, even when numerous genes are used, the prediction of PIC50 values (the effectiveness or accuracy) was noticeably lower than the value attained by TAD activity score. **G)** Comparison of CCLE drug-sensitivity (Ciclopirox) correlation with activity of a TAD (red) (chr10_98830000_99160000) across all cancer types (CCLE database) and its respective genes (yellow). In the shown figure, none of the genes individually have as high a correlation with PIC50 value of Ciclopirox as their host TAD activity. **H)** Comparison of CCLE drug-resistant (Axitinib) correlation values of activity of a TAD (red) (chr1_151880000_152010000) and expression of genes in its respective gene-set (yellow).

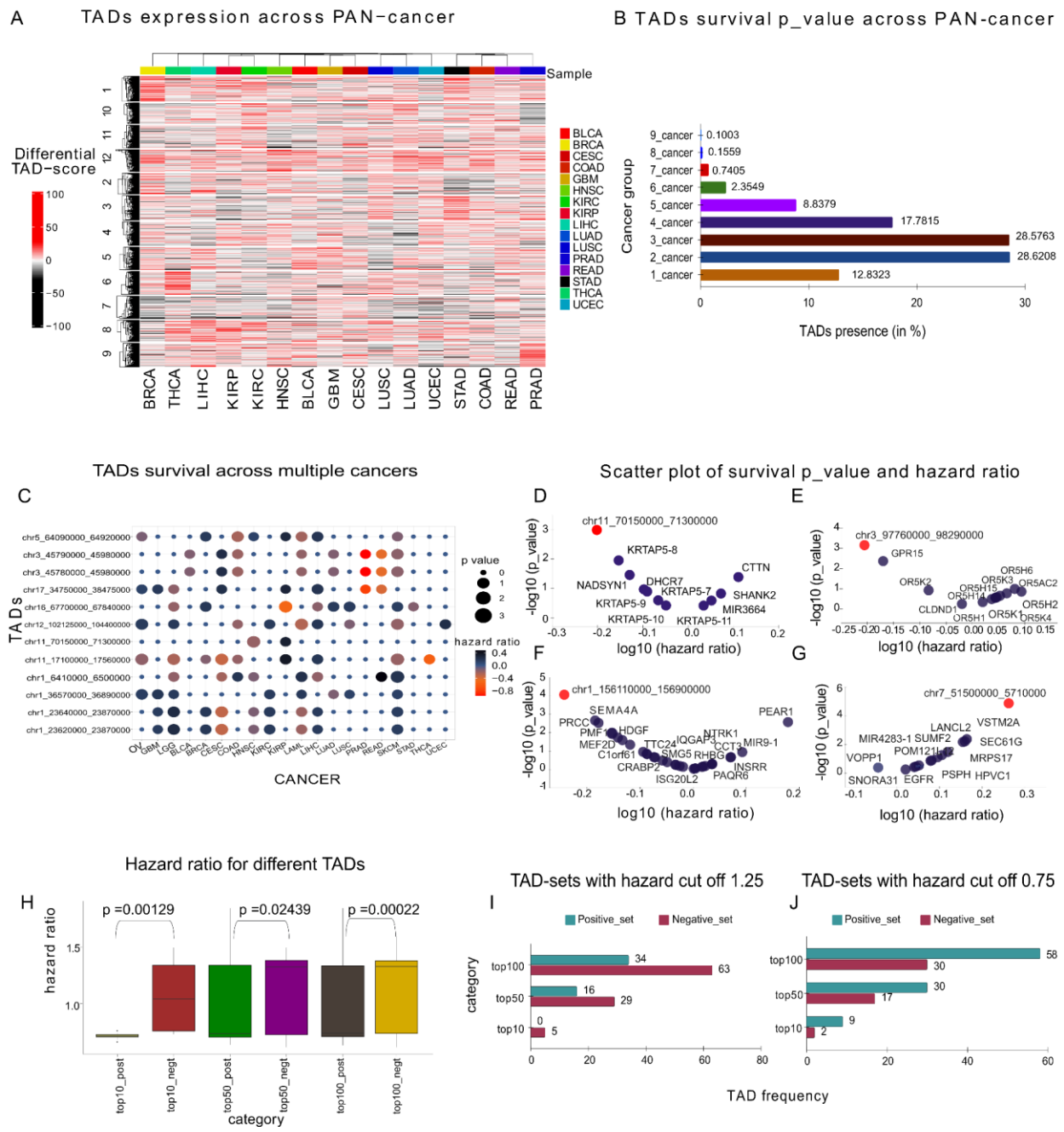


Figure 3: Activity of TADs across multiple cancer TCGA samples and their association with survival. A) Heatmap of fold changes of median TAD activity (tumor Vs normal sample) across 16 cancer types, calculated using TCGA data. **B)** TADs frequency is shown in PAN-cancer. TAD frequency was determined by calculating the survival's P-value. **C)** Based on their P-values and hazard ratios for survival, some TADs patterns across PAN-cancer are graphically shown by dot plots. **D)** TAD (chr11_70150000_71300000)-gene survival P-value comparisons are illustrated

using scatter plots. **E)** TAD (chr3_97760000_98290000)-gene survival P-value comparisons are illustrated using scatter plots. **F)** TAD (chr1_156110000_156900000)-gene survival P-value comparisons are illustrated using scatter plots. **G)** TAD (chr7_51500000_57100000)-gene survival P-value comparisons are illustrated using scatter plots. **H)** hazard-ratio distribution of drug-correlated TADs across the top 10, 50, and 100 categories to analyse the role of TADs in drug resistance. It was observed that the hazard ratio of negatively correlated TADs was above 1 in each of the categories, while the inverse was true for positively associated TADs. **I)** Frequency of TADs with hazard cutoffs greater than 1.25. **J)** Frequency of TADs with hazard cutoffs less than 0.75. The results demonstrated that the majority of negatively correlated TADs (Negative_set) had hazard ratios above 1.25 (poor survival), whereas the majority of positively correlated TADs (Positive_set) had hazard ratios below 0.75. (high survival).

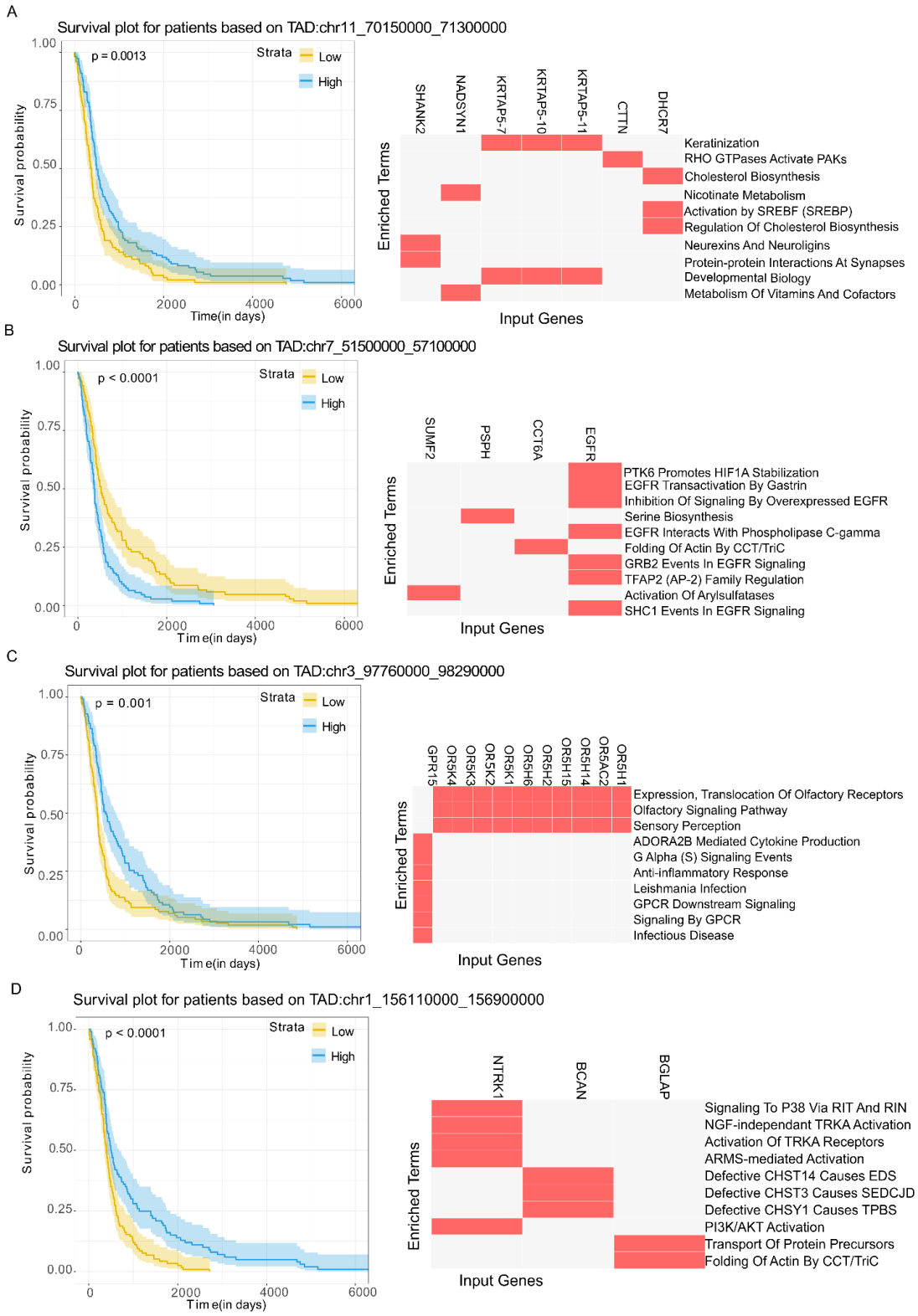


Figure 4: **A)** For patients with head and neck cancer (HNSC), a Kaplan-Meier (KM) plot shows the survival rate. KM plots were created for two groups of HNSC patients where TAD activity

(hg19:chr11-70150000_71300000) was high and those who had low TAD activity. On the right panel, the top enriched ontology terms for genes in the TAD, are displayed alongside the associated genes. **B)** A KM-plot for the survival rate of HNSC patients based on TAD activity (hg19:chr7_51500000_57100000) is displayed, along with the top enriched terms for the relevant genes. **C)** The KM-plot for the survival rate for HNSC patients based on TAD (hg19:chr3-97760000_98290000) activity is displayed. Also displayed are the enriched terms for the genes contained in that TAD. **D)** The KM-plot for the survival rate for HNSC patients based on TAD (hg19:chr1-156110000_156900000) activity is displayed. Also displayed are the enriched terms for the genes contained in that TAD.

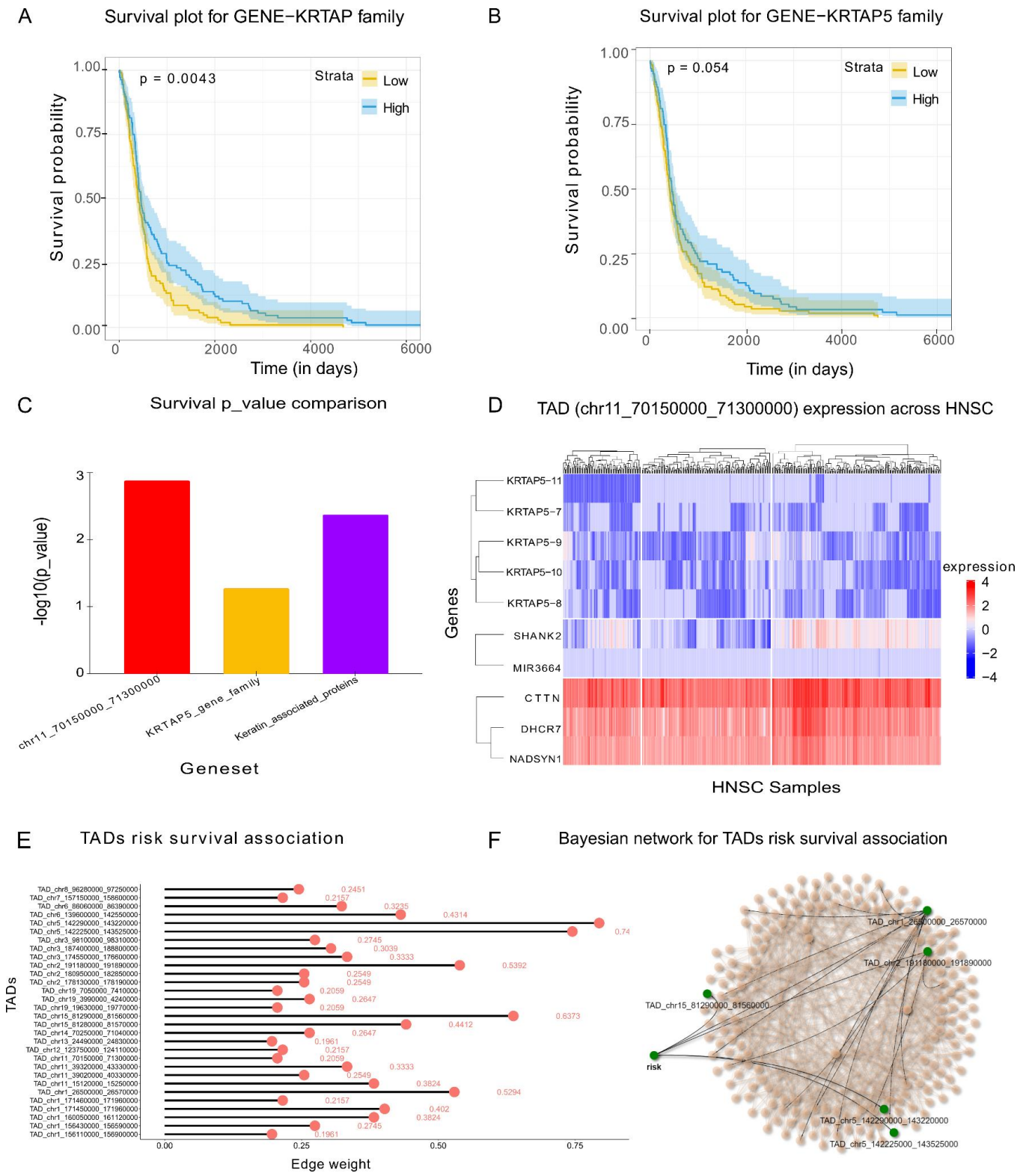


Figure 5: A) Kaplan-Meier (KM) plot shows the survival rate in TCGA-HNSC patients for activity of the gene-set of KRTAP family genes. **B)** The KM plot for activity of KRTAP5 family gene-set. **C)** Bar-plot illustrating the survival P-values associated with the activity of the gene-set,

a TAD (hg19:chr11-70150000_71300000), KRTAP-gene-set, and KRTAP5-geneset family. **D)** Heatmap of expression of genes in TAD (hg19:chr11-70150000_71300000) in TCGA-HNSC samples. **E)** Bar-plot explaining edge strength of risk node and its neighbor nodes based on Bayesian network modeled using TAD activity score. The Bayesian network was modeled to check the direct influence of the TAD activity on survival. **F)** A visualisation of the Bayesian network learned for relationship between activity of TADs and survival.

References

1. Li L, Barth NKH, Pilarsky C, Taher L. Cancer Is Associated with Alterations in the Three-Dimensional Organization of the Genome. Cancers . 2019;11. doi:10.3390/cancers11121886

2. Polak P, Karlić R, Koren A, Thurman R, Sandstrom R, Lawrence M, et al. Cell-of-origin chromatin organization shapes the mutational landscape of cancer. Nature. 2015;518: 360–364.

3. Jones PA, Issa J-PJ, Baylin S. Targeting the cancer epigenome for therapy. Nat Rev Genet. 2016;17: 630–641.

4. Pisco AO, Huang S. Non-genetic cancer cell plasticity and therapy-induced stemness in tumour relapse: “What does not kill me strengthens me.” Br J Cancer. 2015;112: 1725–1732.

5. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modeling of anticancer drug sensitivity. Nature. 2012;483: 603.

6. Kenyon J, Gerson SL. The role of DNA damage repair in aging of adult stem cells. Nucleic Acids Res. 2007;35: 7557–7565.

7. DeNicola GM, Karreth FA, Humpton TJ, Gopinathan A, Wei C, Frese K, et al. Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature. 2011;475: 106–109.

8. Nakasone ES, Askautrud HA, Kees T, Park J-H, Plaks V, Ewald AJ, et al. Imaging tumor-stroma interactions during chemotherapy reveals contributions of the microenvironment to resistance. Cancer Cell. 2012;21: 488–503.

9. Rosenzweig SA. Acquired resistance to drugs targeting receptor tyrosine kinases. Biochem Pharmacol. 2012;83: 1041–1048.

10. Challen GA, Little MH. A side order of stem cells: the SP phenotype. Stem Cells. 2006;24: 3–12.

11. Zhou S, Schuetz JD, Bunting KD, Colapietro AM, Sampath J, Morris JJ, et al. The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat Med. 2001;7: 1028–1034.

12. Dean M, Fojo T, Bates S. Tumour stem cells and drug resistance. Nat Rev Cancer. 2005;5: 275–284.

13. Donnenberg VS, Donnenberg AD. Multiple drug resistance in cancer revisited: the cancer stem cell hypothesis. J Clin Pharmacol. 2005;45: 872–877.

14. Medema JP. Cancer stem cells: the challenges ahead. Nat Cell Biol. 2013;15: 338–344.

15. Morgan MA, Shilatifard A. Chromatin signatures of cancer. Genes Dev. 2015;29: 238–249.

16. Kleppe A, Albregtsen F, Vlatkovic L, Pradhan M, Nielsen B, Hveem TS, et al. Chromatin organisation and cancer prognosis: a pan-cancer study. Lancet Oncol. 2018;19: 356–369.

17. Tena JJ, Santos-Pereira JM. Topologically Associating Domains and Regulatory Landscapes in Development, Evolution and Disease. Front Cell Dev Biol. 2021;9: 702787.

18. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485: 376–380.

19. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159: 1665–1680.

20. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015;518: 331–336.

21. Ibn-Salem J, Muro EM, Andrade-Navarro MA. Co-regulation of paralog genes in the three-dimensional chromatin architecture. Nucleic Acids Res. 2017;45: 81–91.

22. Akdemir KC, Le VT, Kim JM, Killcoyne S, King DA, Lin Y-P, et al. Somatic mutation distributions in cancer genomes vary with three-dimensional chromatin structure. Nat Genet. 2020;52: 1178–1188.

23. Gong Y, Lazaris C, Sakellaropoulos T, Lozano A, Kambadur P, Ntziachristos P, et al. Stratification of TAD boundaries reveals preferential insulation of super-enhancers by strong boundaries. Nat Commun. 2018;9: 542.

24. Chia S, Low JL, Zhang X, Kwang XL, Chong FT, Sharma A, et al. Phenotype-driven precision oncology as a guide for clinical decisions one patient at a time. Nat Commun. 2017;8. doi:10.1038/s41467-017-00451-5

25. Suphavilai C, Chia S, Sharma A, Tu L, Da Silva RP, Mongia A, et al. Predicting heterogeneity in clone-specific therapeutic vulnerabilities using single-cell transcriptomic signatures. Genome Med. 2021;13. doi:10.1186/s13073-021-01000-y

26. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14: 7.

27. Sharma A, Cao EY, Kumar V, Zhang X, Leong HS, Wong AML, et al. Longitudinal single-cell RNA sequencing of patient-derived primary cells reveals drug-induced infidelity in stem cell hierarchy. Nat Commun. 2018;9: 4931.

28. Ramazzotti D, Angaroni F, Maspero D, Ascolani G, Castiglioni I, Piazza R, et al. Variant calling from scRNA-seq data allows the assessment of cellular identity in patient-derived cell lines. Nature communications. 2022. p. 2718.

29. Lang F, Stournaras C. Ion channels in cancer: future perspectives and clinical potential. Philos Trans R Soc Lond B Biol Sci. 2014;369: 20130108.

30. Lin B, Wang S, Yao Y, Shen Y, Yang H. Comprehensive co-expression analysis reveals TMC8 as a prognostic immune-associated gene in head and neck squamous cancer. Oncol Lett. 2021;22: 498.

31. Wu C-J, Li X, Sommers CL, Kurima K, Huh S, Bugos G, et al. Expression of a TMC6-TMC8-CIB1 heterotrimeric complex in lymphocytes is regulated by each of the components. J Biol Chem. 2020;295: 16086–16099.

32. Martens JHA, Mandoli A, Simmer F, Wierenga B-J, Saeed S, Singh AA, et al. ERG and FLI1 binding sites demarcate targets for aberrant epigenetic regulation by AML1-ETO in acute myeloid leukemia. Blood. 2012;120: 4038–4048.

33. Fuchs O. Importance and Presentation of Transcription Factor Fli-1 in Hematopoiesis and in Hematological and Other Malignancies. ARC Journal of Hematology. 2017;2: 23–37.

34. Lachner J, Mlitz V, Tschachler E, Eckhart L. Epidermal cornification is preceded by the expression of a keratinocyte-specific set of pyroptosis-related genes. Sci Rep. 2017;7: 17446.

35. Lindberg K, Rheinwald JG. Suprabasal 40 kd keratin (K19) expression as an immunohistologic marker of premalignancy in oral epithelium. Am J Pathol. 1989;134: 89–98.

36. Miettinen M, Franssila K. Immunohistochemical spectrum of malignant melanoma. The common presence of keratins. Lab Invest. 1989;61: 623–628.

37. Thompson EW, Paik S, Brünner N, Sommers CL, Zugmaier G, Clarke R, et al. Association of increased basement membrane invasiveness with absence of estrogen receptor and expression of vimentin in human breast cancer cell lines. J Cell Physiol. 1992;150: 534–544.

38. Hendrix MJ, Seftor EA, Chu YW, Seftor RE, Nagle RB, McDaniel KM, et al. Coexpression of vimentin and keratins by human melanoma tumor cells: correlation with invasive and metastatic potential. J Natl Cancer Inst. 1992;84: 165–174.

39. Crowe DL, Milo GE, Shuler CF. Keratin 19 downregulation by oral squamous cell carcinoma lines increases invasive potential. J Dent Res. 1999;78: 1256–1263.

40. Stenvall C-GA, Nyström JH, Butler-Hallissey C, Jansson T, Heikkilä TRH, Adam SA, et al. Cytoplasmic keratins couple with and maintain nuclear envelope integrity in colonic epithelial cells. Mol Biol Cell. 2022;33: ar121.

41. Stenvall C-GA, Nyström JH, Butler-Hallissey C, Adam SA, Foisner R, Ridge KM, et al. Keratins couple with the nuclear lamina and regulate proliferation in colonic epithelial cells. bioRxiv. 2020. p. 2020.06.22.164467. doi:10.1101/2020.06.22.164467

42. Taniura H, Glass C, Gerace L. A chromatin binding site in the tail domain of nuclear lamins that interacts with core histones. J Cell Biol. 1995;131: 33–44.

43. Vielmuth F, Wanuske M-T, Radeva MY, Hiermaier M, Kugelmann D, Walter E, et al. Keratins Regulate the Adhesive Properties of Desmosomal Cadherins through Signaling. J Invest Dermatol. 2018;138: 121–131.

44. Hatzfeld M, Keil R, Magin TM. Desmosomes and Intermediate Filaments: Their Consequences for Tissue Mechanics. Cold Spring Harb Perspect Biol. 2017;9. doi:10.1101/cshperspect.a029157

45. Berens EB, Sharif GM, Schmidt MO, Yan G, Shuptrine CW, Weiner LM, et al. Keratin-associated protein 5-5 controls cytoskeletal function and cancer cell vascular invasion. Oncogene. 2017;36: 593–605.

46. Liu H, Zhang H-W, Sun X-F, Guo X-H, He Y-N, Cui S, et al. Tamoxifen-resistant breast cancer cells possess cancer stem-like cell properties. Chin Med J . 2013;126: 3030–3034.

47. Yan Y-R, Xie Q, Li F, Zhang Y, Ma J-W, Xie S-M, et al. Epithelial-to-mesenchymal transition is involved in BCNU resistance in human glioma cells. Neuropathology. 2014;34: 128–134.

48. Hu Y, Zang J, Qin X, Yan D, Cao H, Zhou L, et al. Epithelial-to-mesenchymal transition correlates with gefitinib resistance in NSCLC cells and the liver X receptor ligand GW3965 reverses gefitinib resistance through inhibition of vimentin. Onco Targets Ther. 2017;10: 2341–2348.

49. Olson MF, Ashworth A, Hall A. An essential role for Rho, Rac, and Cdc42 GTPases in cell cycle progression through G1. Science. 1995;269: 1270–1272.

50. Wang Z, Pedersen E, Basse A, Lefever T, Peyrollier K, Kapoor S, et al. Rac1 is crucial for Ras-dependent skin tumor formation by controlling Pak1-Mek-Erk hyperactivation and hyperproliferation in vivo. Oncogene. 2010;29: 3362–3373.

51. Wu C-YC, Carpenter ES, Takeuchi KK, Halbrook CJ, Peverley LV, Bien H, et al. PI3K regulation of RAC1 is required for KRAS-induced pancreatic tumorigenesis in mice. Gastroenterology. 2014;147: 1405–16.e7.

52. Mizukawa B, Wei J, Shrestha M, Wunderlich M, Chou F-S, Griesinger A, et al. Inhibition of Rac GTPase signaling and downstream prosurvival Bcl-2 proteins as combination targeted therapy in MLL-AF9 leukemia. Blood. 2011;118: 5235–5245.

53. Liu T, Porter J, Zhao C, Zhu H, Wang N, Sun Z, et al. TADKB: Family classification and a knowledge base of topologically associating domains. BMC Genomics. 2019;20: 1–17.

54. Allain EP, Rouleau M, Lévesque E, Guillemette C. Emerging roles for UDP-glucuronosyltransferases in drug resistance and cancer progression. Br J Cancer. 2020;122: 1277–1287.

55. Stingl JC, Bartels H, Viviani R, Lehmann ML, Brockmöller J. Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. Pharmacol Ther. 2014;141: 92–116.

56. Pires IM, Blokland NJ, Broos AW, Poujade F-A, Senra JM, Eccles SA, et al. HIF-1α-independent hypoxia-induced rapid PTK6 stabilization is associated with increased motility and invasion. Cancer Biol Ther. 2014;15: 1350–1357.

57. Sasaki T, Hiroki K, Yamashita Y. The role of epidermal growth factor receptor in cancer metastasis and microenvironment. Biomed Res Int. 2013;2013: 546318.

58. Chung C, Cho HJ, Lee C, Koo J. Odorant receptors in cancer. BMB Rep. 2022;55: 72–80.

59. Alfaro R, Crowder S, Sarma KP, Arthur AE, Pepino MY. Taste and Smell Function in Head and Neck Cancer Survivors. Chem Senses. 2021;46. doi:10.1093/chemse/bjab026

60. Barutcu AR, Maass PG, Lewandowski JP, Weiner CL, Rinn JL. A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat Commun. 2018;9: 1444.

61. Lee DP, Tan WLW, Anene-Nzelu CG, Lee CJM, Li PY, Luu TDA, et al. Robust CTCF-Based Chromatin Architecture Underpins Epigenetic Changes in the Heart Failure Stress-Gene Response. Circulation. 2019;139: 1937–1956.

62. Forcato M, Nicoletti C, Pal K, Livi CM, Ferrari F, Bicciato S. Comparison of computational methods for Hi-C data analysis. Nat Methods. 2017;14: 679–685.

63. Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, et al. HiCUP: pipeline for mapping and processing Hi-C data. F1000Res. 2015;4: 1310.

64. Aljogol D, Thompson IR, Osborne CS, Mifsud B. Comparison of Capture Hi-C Analytical Pipelines. Front Genet. 2022;0. doi:10.3389/fgene.2022.786501

65. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of

SAMtools and BCFtools. Gigascience. 2021;10. doi:10.1093/gigascience/giab008

66. Dali R, Blanchette M. A critical assessment of topologically associating domain prediction tools. Nucleic Acids Res. 2017;45: 2994–3005.

67. Kumar V, Muratani M, Rayan NA, Kraus P, Lufkin T, Ng HH, et al. Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol. 2013;31: 615–622.

68. Hung J-H, Yang T-H, Hu Z, Weng Z, DeLisi C. Gene set enrichment analysis: performance evaluation and usage guidelines. Brief Bioinform. 2012;13: 281–291.

69. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, et al. Toward a Shared Vision for Cancer Genomic Data. N Engl J Med. 2016;375: 1109–1112.

70. Chawla S, Samydurai S, Kong SL, Wu Z, Wang Z, Tam WL, et al. UniPath: a uniform approach for pathway and gene-set based analysis of heterogeneity in single-cell epigenome and transcriptome profiles. Nucleic Acids Res. 2021;49: 1801.

71. Tm T, Grambsch PM. Modeling survival data: extending the Cox model. New York Springer; 2000.

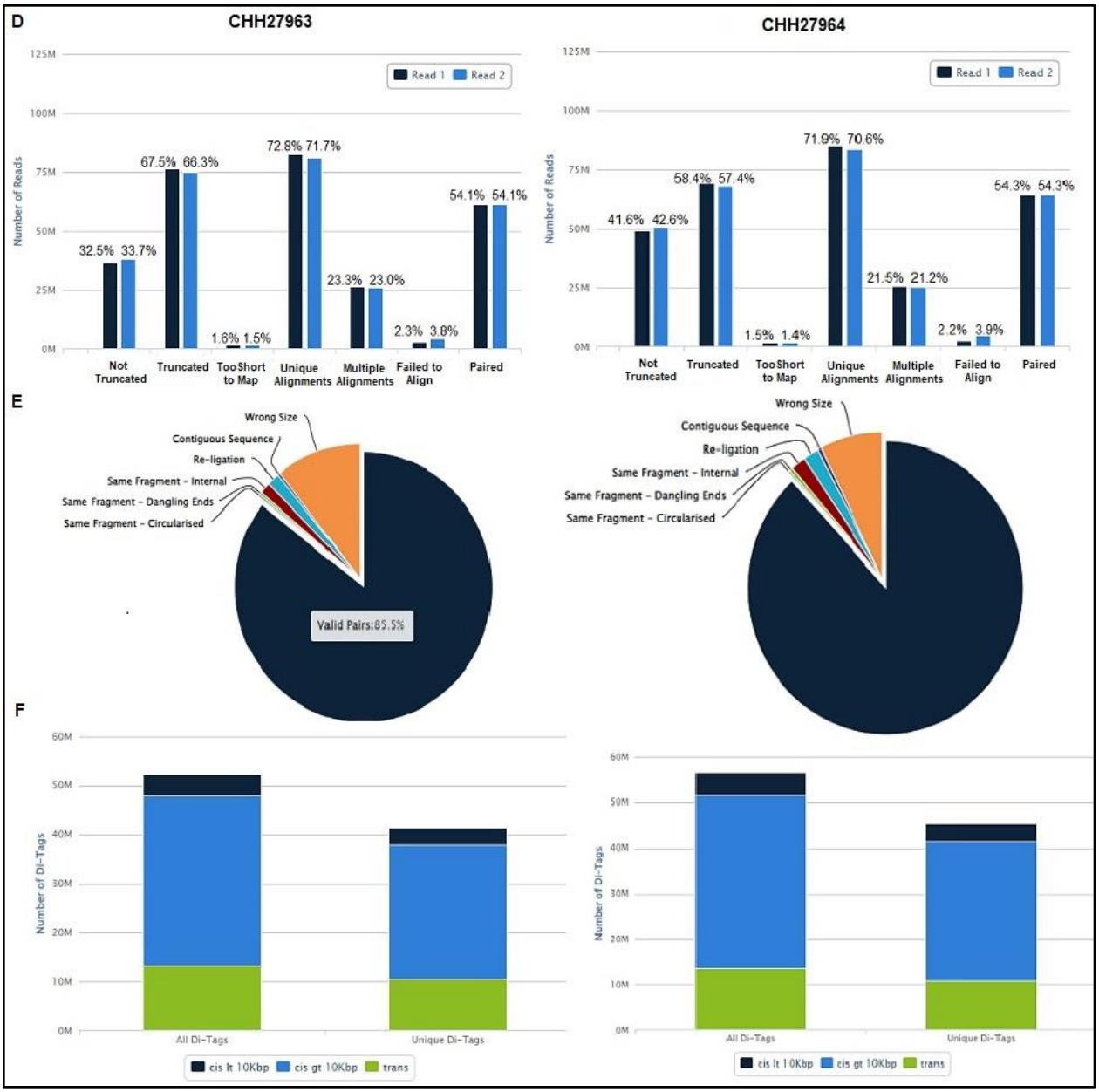
72. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38: 675–678.

73. Lausen B, Schumacher M. Maximally Selected Rank Statistics. Biometrics. 1992. p. 73. doi:10.2307/2532740

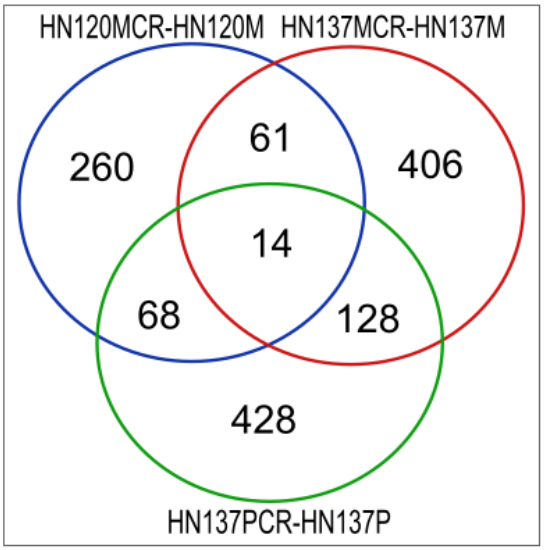
74. Scutari M, Denis J-B. Bayesian Networks: With Examples in R. CRC Press; 2014.

Supplementary Figures

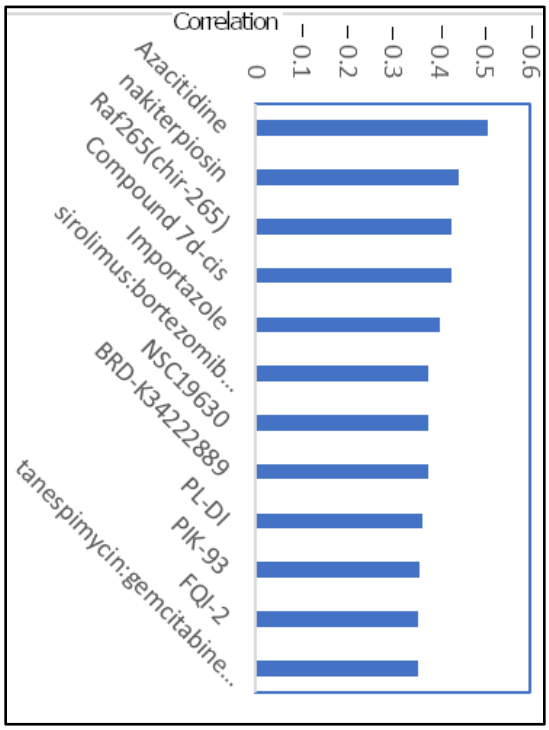
A
CHH27488
CHH27962
Read 1
Read 2
Number of Reads
250M
200M
150M
100M
50M
0M
175M
125M
75M
25M
25.1% 25.6%
74.9% 74.4%
3.1% 3.3%
68.2% 67.1%
24.5% 24.3%
4.2% 5.3%
45.3% 45.3%
33.5% 34.5%
66.5% 66.5%
1.8% 1.7%
73.0% 71.7%
22.8% 22.6%
2.4% 4.1%
54.1% 54.1%
Not Truncated
Truncated
Too Short to Map
Unique Alignments
Multiple Alignments
Failed to Align
Paired
B
Wrong Size
Contiguous Sequence
Re-ligation
Same Fragment - Internal
Same Fragment - Dangling Ends
Same Fragment - Circularised
Valid Pairs:78.8%
Valid Pairs:83.0%
C
Number of Di-Tags
100M
80M
60M
40M
20M
0M
All Di-Tags
Unique Di-Tags
cis lt 10Kbp
cis gt 10Kbp
trans



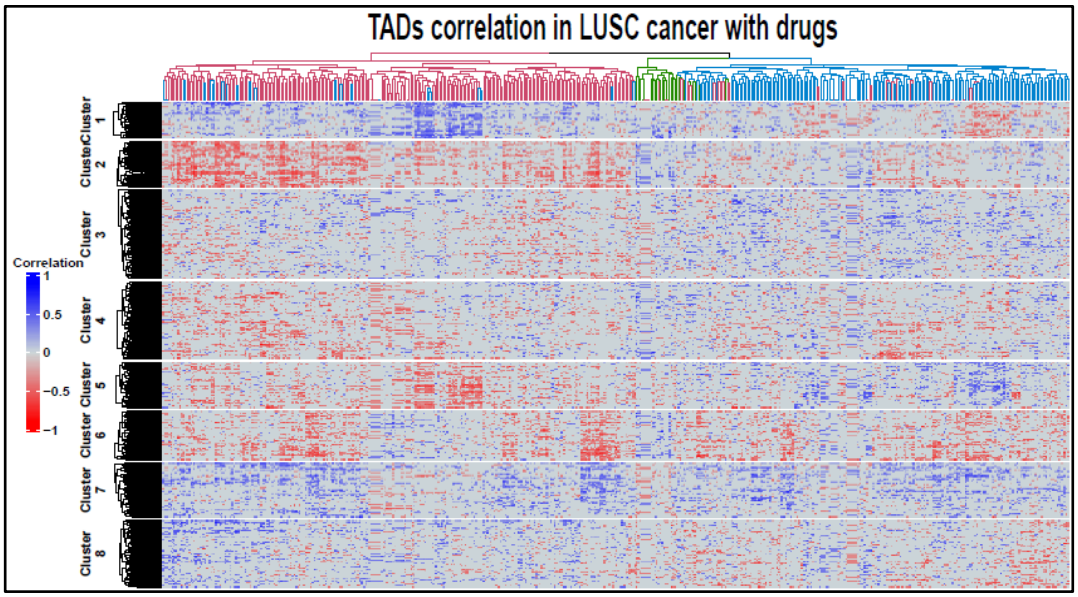
Supplementary Figure S1: Quality control reports of HiC-profiles: **A)** Truncating and mapping of HiC reads of sample HN137P and HN120M from HNSC cancer sample. **B)** Filtering of uninformative reads and explaining valid pairs in sample HN137P and HN120M from HNSC cancer sample. **C)** De-duplication bar-chart in respect of cis-trans read pairs in sample HN137P and HN120M from HNSC cancer sample. **D)** Truncating and mapping of HiC reads of sample HN120P and HN148P from HNSC cancer sample. **E)** Filtering of uninformative reads and explaining valid pairs in sample HN120P and HN148P from HNSC cancer sample. **F)** De-duplication bar-chart in respect of cis-trans read pairs in sample HN120P and HN148P from HNSC cancer sample.



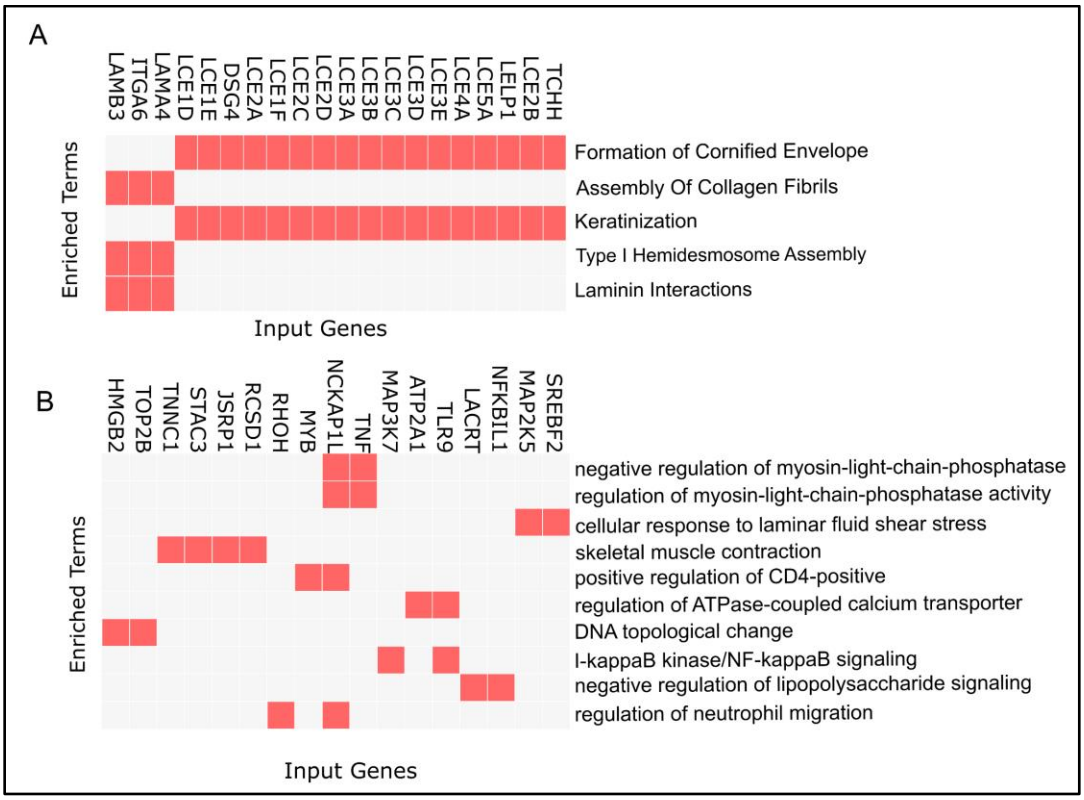
Supplementary Figure S2: The TAD activity was calculated using expression profile of single cells for 3 HNSC cell lines and their cisplatin-resistance version. The venn diagram shows number of TADs upregulated in cisplatin-resistant versions of HNSC cell lines.



Supplementary Figure S3: Correlation of activity of TAD chr11_70150000_71300000 with PIC50 values of a few drugs in HNSC cell lines (CCLE database). This TAD seems to be associated with resistance towards the shown drugs as the correlation is negative.

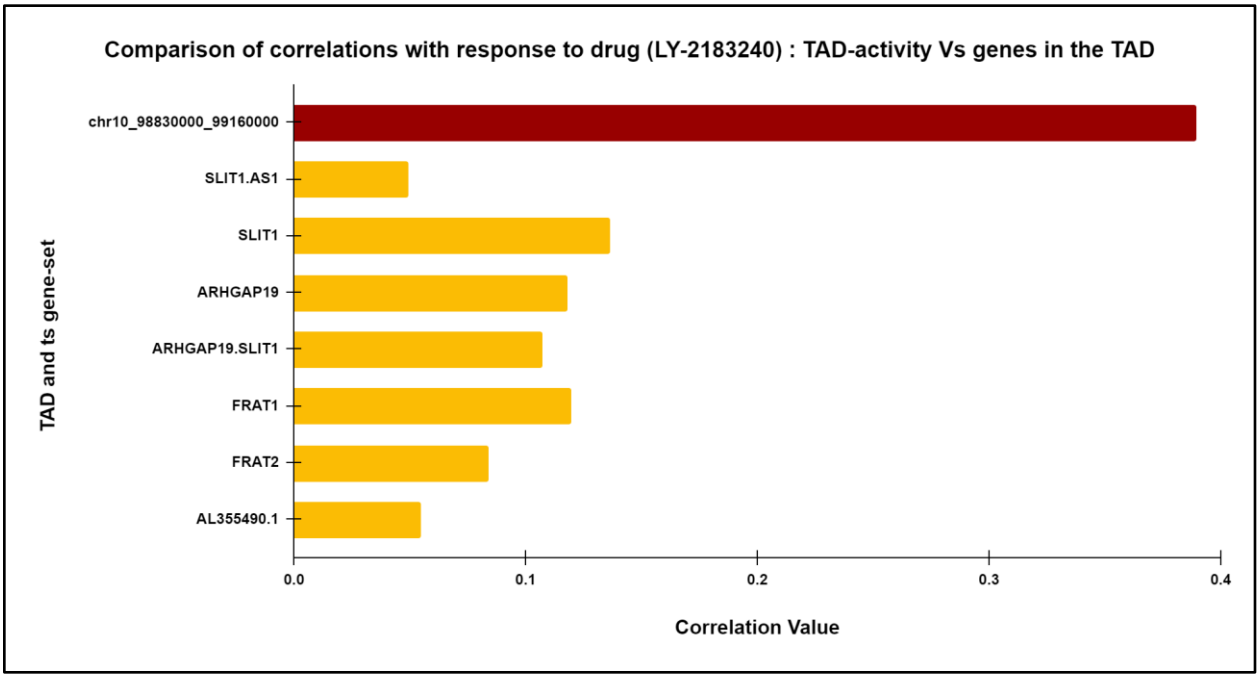


Supplementary Figure S4: Heatmap plot based on correlation: Heatmap illustration for correlation among TAD activity and PIC50 value for drugs in LUSC cell lines from CCLE database.

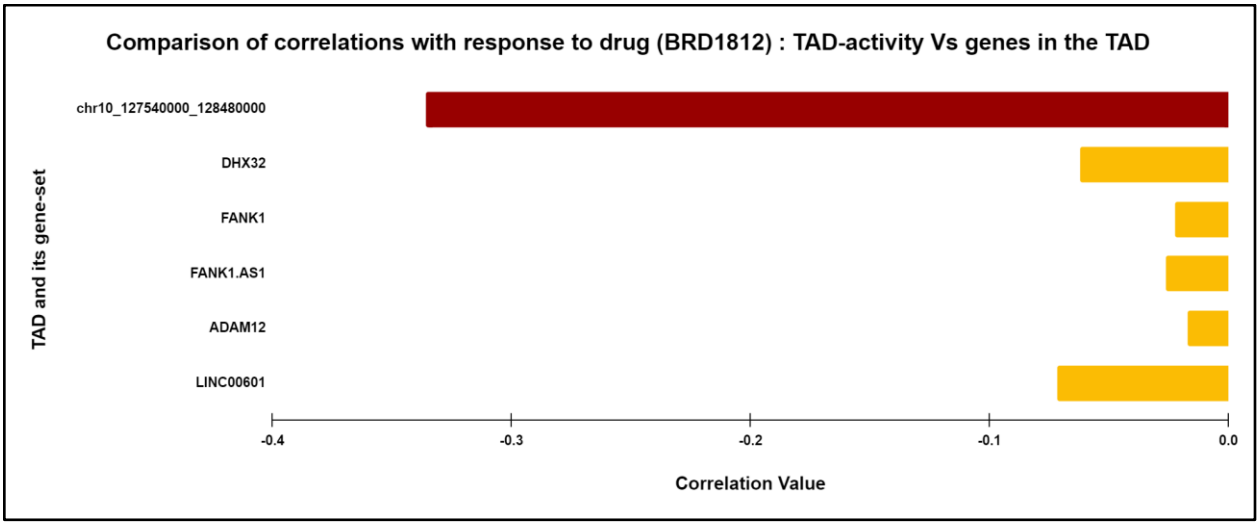


Supplementary Figure S5: A) Gene enrichment for TADs in cluster 2. **B)** Gene enrichment for TADs in cluster 1.

A

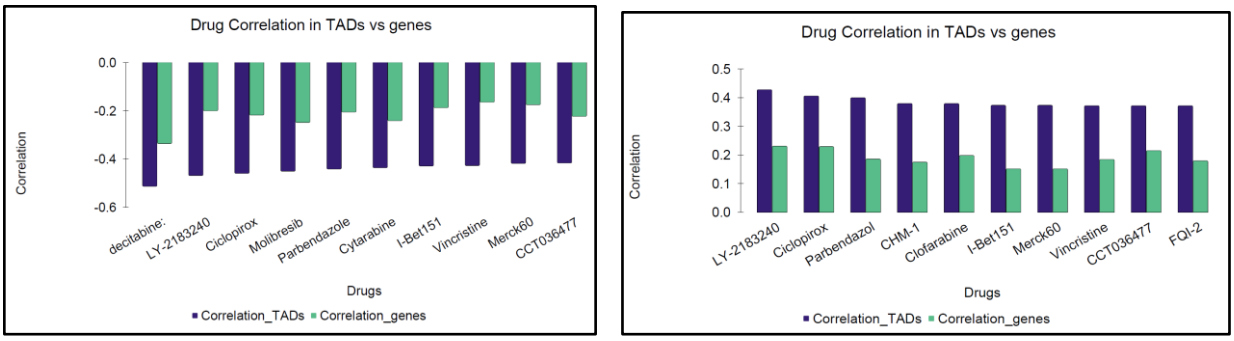


B

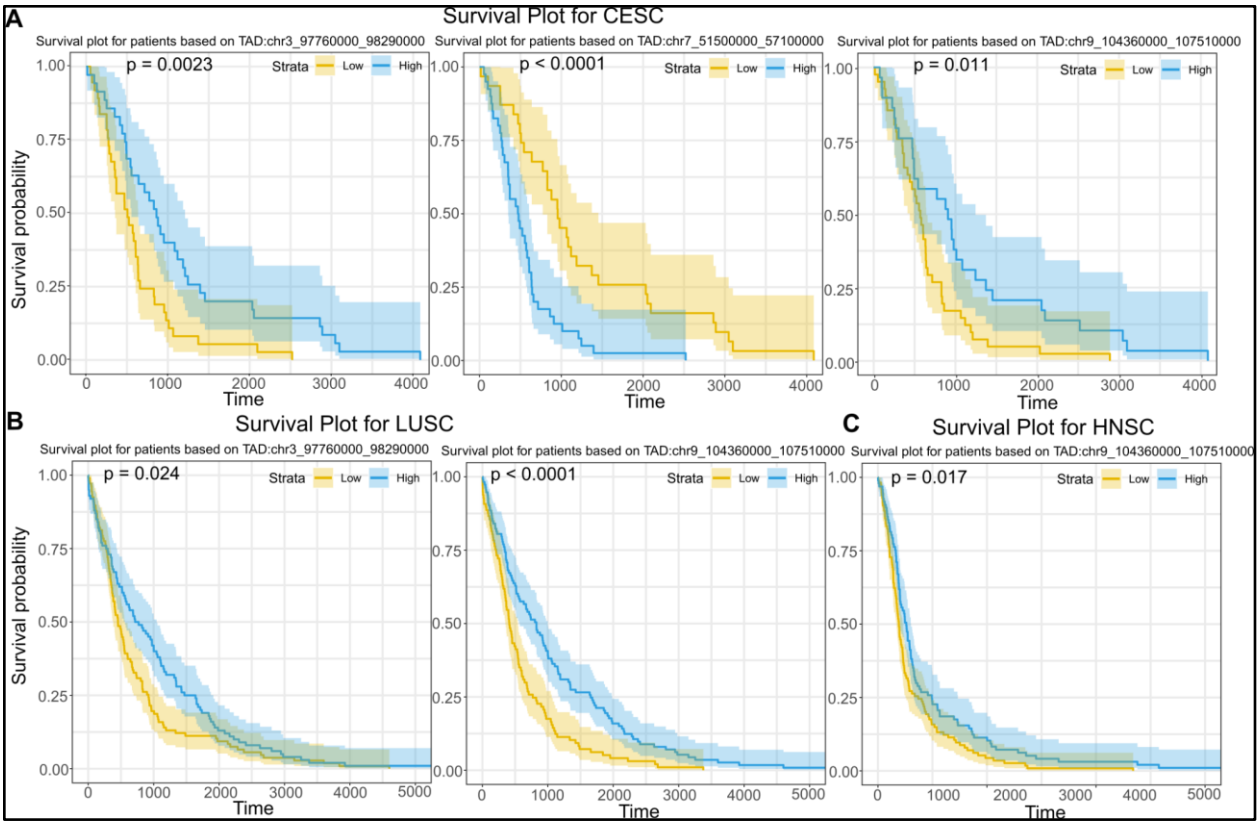


Supplementary Figure S6: A) Correlation comparison between TAD and its gene-sets for drug LY-2183240. **B)** Correlation comparison between TAD and its gene-sets for drug BRD1812.

A **B**

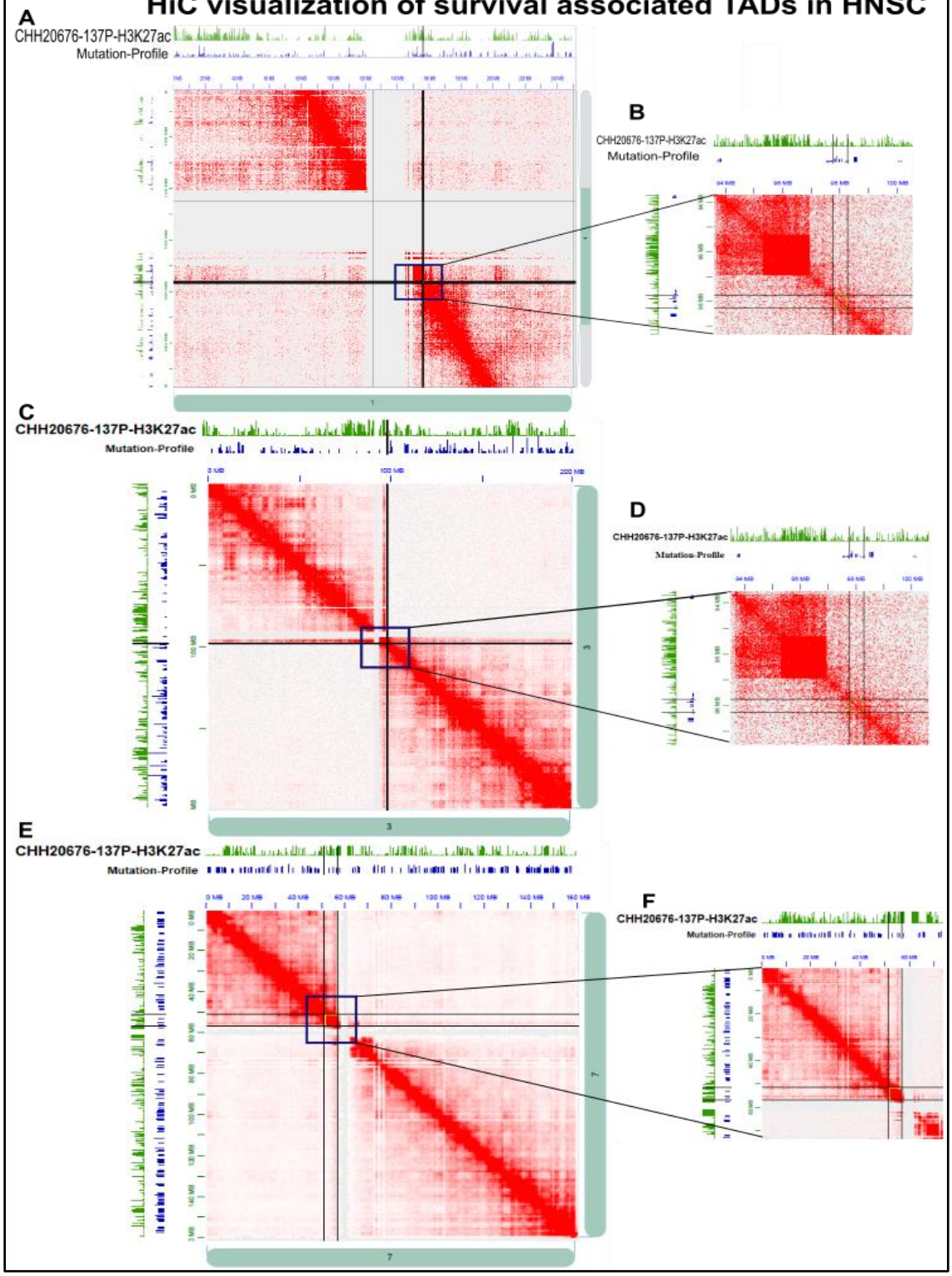


Supplementary Figure S7: Barplot demonstrating the association between the PIC50 values of a few drugs and TAD activity (blue) or a single gene's expression (green). **A)** Drug associations based on negative correlation. **B)** drug relationship measured by positive correlation.

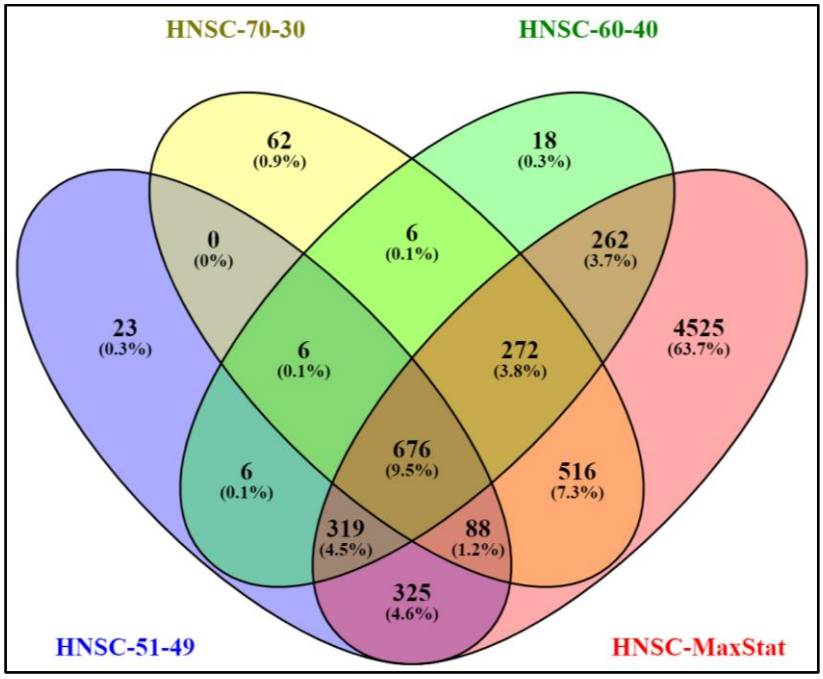


Supplementary Figure S8: TCGA survival results for CESC, LUSC, and HNSC cancers: **A)** Kaplan-Meier (KM) plot shows the survival rate for patients with cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC). **B)** Kaplan-Meier (KM) plot shows the survival rate for patients with lung squamous cell carcinoma (LUSC). **C)** Kaplan-Meier (KM) plot shows the survival rate for patients with Head and neck squamous cell carcinomas (HNSC).

HiC visualization of survival associated TADs in HNSC



Supplementary Figure S9: HiC profile (HN137P) visualisation of chromatin interaction matrix around TADs with significant association with survival in HNSC cancer cell line: **A)** HiC visualization of survival associated TAD-hg19:chr1-156110000_156900000 (Survival P-value: 0.0000946, Hazard ratio: 0.573137332). **B)** HiC visualization of survival associated TAD-hg19:chr1-156110000_156900000 in an expanded view. **C)** HiC visualization of survival associated TAD-hg19:chr3-97760000_98290000 (Survival P-value: 0.001000374, Hazard ratio: 0.632238033). **D)** HiC visualization of survival associated TAD-hg19:chr3-97760000_98290000 in expanded view. **E)** HiC visualization of survival associated TAD-hg19:chr7-51500000_57100000 (Survival P-value: 0.0000165, Hazard ratio: 1.841025503). **F)** HiC visualization of survival associated TAD-hg19:chr7-51500000_57100000) in expanded view.



Supplementary Figure S10: Common TADs (676) shown in a Venn diagram based on significant P-values for survival between different pairs of patients in HNSC, such as 70-30, 60-40, 51-49, and the Maxstat cuttoff approach.

Supplementary Tables

Supplementary Table 1: Significant downregulated TADs based on fold change and significance (P-value) of the TAD activity between cisplatin-resistant version of metastatic cancer-cells of patient 137MCR and metastatic cancer-cells of patient 137M.

TAD_numbers	TADs	Wilcoxon P-value	Fold change
TAD_9950	chr2_110300000_110850000	5.40E-27	-10.98462816
TAD_6100	chr18_28400000_29075000	4.57E-16	-7.642808579
TAD_6105	chr18_28560000_29100000	2.14E-13	-7.593933704
TAD_6081	chr18_25180000_29010000	1.22E-12	-7.392324844
TAD_9881	chr2_101870000_102500000	3.02E-10	-6.292640157
TAD_13475	chr4_70100000_71450000	7.34E-11	-6.170452969
TAD_13478	chr4_70150000_71420000	7.34E-11	-6.170452969
TAD_13480	chr4_70550000_71440000	7.34E-11	-6.170452969
TAD_7068	chr1_106325000_108925000	4.37E-12	-6.048265782
TAD_9986	chr2_113810000_114260000	2.02E-09	-5.938297313
TAD_9988	chr2_113830000_114260000	2.02E-09	-5.938297313
TAD_7820	chr1_20540000_20960000	1.92E-07	-5.467876641
TAD_1614	chr11_70330000_71470000	6.87E-10	-5.388454969
TAD_1615	chr11_70800000_71270000	6.87E-10	-5.388454969
TAD_1616	chr11_70800000_71290000	6.87E-10	-5.388454969
TAD_1617	chr11_70800000_71300000	6.87E-10	-5.388454969
TAD_2743	chr12_96300000_96930000	1.42E-07	-5.357908172
TAD_16019	chr7_127250000_128050000	3.48E-07	-5.272377141
TAD_8703	chr1_86850000_87560000	1.33E-06	-4.979127891

TAD_15243	chr6_166670000_167180000	4.25E-10	-4.960799813

Supplementary Table 2: Significant upregulated TADs based on fold change and significance (P-value) of the TAD activity between cisplatin-resistant version of metastatic cancer-cells of patient 137MCR and metastatic cancer-cells of patient 137M.

TADs_number	HN137MCR_HN137M_TADs	Wilcoxon P-value	Fold change
TAD_17241	chr8_23350000_24175000	1.22E-11	6.799716985
TAD_255	chr10_13390000_14530000	1.04E-06	5.082987
TAD_16793	chr7_92860000_93910000	7.94E-07	5.070768282
TAD_16797	chr7_92900000_93725000	7.94E-07	5.070768282
TAD_16500	chr7_47300000_48020000	2.08E-06	4.930253016
TAD_4110	chr15_55480000_55760000	1.79E-06	4.918034297
TAD_3378	chr14_23790000_23940000	1.32E-08	4.863050063
TAD_4202	chr15_65020000_65320000	1.98E-05	4.264332844
TAD_10729	chr2_228190000_228750000	5.54E-05	4.227676688
TAD_13185	chr4_187540000_190920000	9.55E-06	3.977192953
TAD_715	Chr10_88980000_89530000	0.000222295	3.81224025
TAD_228	chr10_127590000_128760000	0.000448072	3.647287547
TAD_14209	chr5_172675000_174200000	5.85E-05	3.610631391
TAD_3867	chr15_22660000_23260000	0.000711653	3.543428438
TAD_14981	chr6_132030000_132820000	1.96E-05	3.439569328
TAD_3870	chr15_22750000_23720000	0.001284547	3.366257016
TAD_4861	chr16_57580000_58050000	0.001255714	3.250179188

TAD_4862	chr16_57770000_58040000	0.001255714	3.250179188
TAD_14707	chr5_90780000_94160000	0.000262175	3.189085594
TAD_14718	chr5_92910000_94260000	0.000262175	3.189085594

Supplementary Table 3: Significant downregulated TADs based on fold change and significance (P-value) of the TAD activity between cisplatin-resistant version of primary cancer-cells of patient 137PCR and primary cancer-cells of patient 137P.

TAD_numbers	TADs	Wilcoxon P-value	Fold change
TAD_1614	chr11_70330000_71470000	3.05E-16	-8.4296
TAD_1615	chr11_70800000_71270000	3.05E-16	-8.4296
TAD_1616	chr11_70800000_71290000	3.05E-16	-8.4296
TAD_1617	chr11_70800000_71300000	3.05E-16	-8.4296
TAD_9950	chr2_110300000_110850000	2.34E-14	-7.78306
TAD_17636	chr8_82190000_82600000	5.28E-14	-7.76254
TAD_9446	chr21_42520000_42840000	1.79E-10	-6.33195
TAD_15606	chr6_52660000_52930000	4.21E-10	-6.17185
TAD_7322	chr1_152460000_153290000	1.14E-08	-5.92966
TAD_7078	chr1_108320000_109230000	1.68E-08	-5.79009
TAD_7321	chr1_152450000_153280000	2.60E-08	-5.76957
TAD_16660	chr7_73800000_75025000	5.18E-07	-5.24823
TAD_1612	chr11_70280000_71500000	1.34E-08	-5.16203

TAD_17632	chr8_81910000_82600000	1.78E-06	-5.00604
TAD_1223	chr11_2550000_2950000	4.68E-08	-4.92188
TAD_1946	chr12_112950000_113620000	2.94E-19	-4.46821
TAD_9952	chr2_110775000_111175000	1.78E-05	-4.44776
TAD_4830	chr16_55350000_56400000	0.000281	-3.66781
TAD_4834	chr16_55490000_56400000	0.000281	-3.66781
TAD_8703	chr1_86850000_87560000	4.39E-06	-3.60829

Supplementary Table 4: Significant upregulated TADs based on fold change and significance (P-value) of the TAD activity between cisplatin-resistant version of primary cancer-cells of patient 137PCR and primary cancer-cells of patient 137P.

TADs_number	HN137PCR_HN137P_EP_TADs	Wilcoxon P-value	Fold change
TAD_4407	chr15_89010000_89170000	5.49E-10	6.430469
TAD_14411	chr5_39570000_40830000	1.50E-07	5.517108
TAD_16945	chr8_11620000_11900000	6.28E-05	4.008524
TAD_17355	chr8_37620000_38030000	0.000234	3.864849
TAD_3991	chr15_40330000_40570000	0.000603	3.42151
TAD_17308	chr8_29180000_30030000	0.000603	3.42151
TAD_17309	chr8_29180000_30040000	0.000603	3.42151
TAD_3993	chr15_40570000_40730000	0.000356	3.370198
TAD_8890	chr20_23330000_23620000	0.00144	3.296308
TAD_11594	chr3_127290000_127770000	0.000729	3.210103

TAD_11596	chr3_127300000_127770000	0.000729	3.210103
TAD_3157	chr13_79960000_80940000	0.002364	3.183421
TAD_3158	chr13_79970000_80940000	0.002364	3.183421
TAD_383	chr10_3150000_4700000	0.002944	3.105426
TAD_7798	chr1_204130000_204480000	0.000572	2.988434
TAD_5744	chr17_62460000_62650000	0.001393	2.928911
TAD_14209	chr5_172675000_174200000	0.002886	2.900176
TAD_9182	chr20_55640000_56070000	0.006492	2.822181
TAD_9848	chr22_50150000_50500000	0.005879	2.770869
TAD_4302	chr15_75310000_75630000	0.004461	2.719557

Supplementary Table 5: The common TADs which were downregulated in cisplatin-resistant versions of the 3 HNSC cell lines (between HN120MCR-HN120M, HN137MCR-HN137M and HN137PCR-HN137P).

S. No.	Common TADs
1.	chr8_85530000_86570000
2.	chr8_81075000_82775000
3.	chr11_70200000_71300000
4.	chr8_80525000_89250000
5.	chr11_70225000_71350000
6.	chr8_83325000_86475000
7.	chr22_32450000_33475000
8.	chr8_80550000_82950000

9.	chr11_70800000_71290000
10.	chr1_208075000_210050000
11.	chr11_70150000_71300000
12.	chr11_70330000_71470000
13.	chr1_208150000_210100000
14.	chr15_100325000_101650000
15.	chr1_208125000_210075000
16.	chr11_70800000_71270000
17.	chr22_32425000_33575000
18.	chr11_70280000_71500000
19.	chr11_70800000_71300000
20.	chr1_208050000_210000000

Supplementary Table 6: The commonly upregulated TADs between HN120MCR-HN120M, HN137MCR-HN137M, and HN137PCR-HN137P (in cisplatin-resistant versions of HNSC cell lines).

S. No.	Common TADs
1.	chr2_228190000_228750000
2.	chr13_32890000_33780000
3.	chr13_32390000_34320000
4.	chr8_23300000_24250000
5.	chr19_40030000_41150000
6.	chr5_131400000_131710000

7.	chr5_36730000_38240000
8.	chr13_32420000_33240000
9.	chr17_76100000_76430000
10.	chr15_89000000_89570000
11.	chr17_4330000_4810000
12.	chr19_49130000_49490000
13.	chr5_36700000_38210000
14.	chr15_89000000_89400000