Post-transcriptional control drives Aurora kinase A expression in human cancers

Roberta Cacioppo[1,3,*], Deniz Rad[1,4], Giulia Pagani[2], Paolo Gandellini[2] and Catherine Lindon[1,*].

[1] Department of Pharmacology, University of Cambridge, Cambridge, CB2 1PD, UK

[2] Department of Biosciences, University of Milan, Milan, 20133, Italy

[3] Current address: MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH, UK

[4] Current address: Trinity College, University of Cambridge, Cambridge, CB2 1TQ, UK

*corresponding author. Email: RC, rcacioppo@mrc-lmb.cam.ac.uk; CL, acl34@cam.ac.uk.

ABSTRACT

Aurora kinase A (AURKA) is a major regulator of the cell cycle. A prominent association exists between high expression of AURKA and cancer, and impairment of AURKA levels can trigger its oncogenic activity. In order to explore the contribution of post-transcriptional regulation to AURKA expression in different cancers, we carried out a meta-analysis of -omics data of 18 cancer types from The Cancer Genome Atlas (TCGA). Our study confirmed a general trend for increased AURKA mRNA in cancer compared to normal tissues and revealed that AURKA expression is highly dependent on post-transcriptional control in several cancers. Correlation and clustering analyses of AURKA mRNA and protein expression, and expression of AURKA-targeting *hsa-let-7a* miRNA, unveiled that *hsa-let-7a* is likely involved to varying extents in controlling AURKA expression in cancers. We then measured differences in the short/long ratio (SLR) of the two alternative cleavage and polyadenylation (APA) isoforms of AURKA mRNA across cancers compared to the respective healthy counterparts. We suggest that the interplay between APA and *hsa-let-7a* targeting of AURKA mRNA may influence AURKA expression in some cancers. *hsa-let-7a* and APA may also independently contribute to altered AURKA levels. Therefore, we argue that AURKA mRNA and protein expression are often discordant in cancer as a result of dynamic post-transcriptional regulation.

INTRODUCTION

AURKA is a key regulator of the cell cycle, controlling centrosome maturation and mitotic spindle assembly[1]. AURKA is overexpressed and represents a marker of poor prognosis in a broad range of human malignancies[2–4]. Because of the significant association between high AURKA expression and cancer progression, poor prognosis, and drug resistance, AURKA is a preferred target for anti-cancer strategies, especially the use of small molecule kinase inhibitors[4,5] and targeted proteolytic tools[6,7]. However, even the most promising AURKA inhibitors are still under clinical studies and only alisertib has concluded phase III clinical assessment[4]. Furthermore, the existence of kinase-independent AURKA activity suggests that deregulation of expression may be sufficient to promote some of its oncogenic functions. Accordingly, anti-cancer strategies aimed at reducing AURKA expression levels have been shown to be effective in suppressing the carcinogenicity of AURKA[8–13].

The literature mainly reports that AURKA overexpression in cancers is due to increased gene copy number, transcription, or protein stability[14]. Various single nucleotide polymorphisms (SNPs) of the *AURKA* gene have also been associated with cancer development and susceptibility[15]. Additionally, certain microRNAs (miRNAs) control AURKA expression in cancers where its overexpression is a driving factor or a marker of poor prognosis[16–20]. However, overexpression of AURKA protein in cancer does not always correlate with the above mechanisms, suggesting that other deregulated post-transcriptional events may occur[21–24].

One important step in the maturation of mRNAs is the cleavage and polyadenylation (C/P) process, which involves cutting the 3' end of precursor mRNAs (pre-mRNAs) and adding a poly(A) tail[25]. A polyadenylation signal (PAS), which is found 10–30 nucleotides upstream, as well as UGUA and U-rich motifs usually precede the cleavage site, whereas U- and GU-rich motifs usually follow; these elements altogether form the C/P site[26]. Most human pre-mRNAs have multiple C/P sites[27], which allows for

different transcript isoforms to be expressed for the same gene by alternative cleavage and polyadenylation (APA).

The function of microRNAs in controlling genes involved in the cell cycle and the significance of this control in cancer are widely recognised[28,29]. The *hsa-let-7* miRNA family consists of 11 closely related genes that map in chromosomal areas that are usually deleted in human tumours[30]. Because of their pathogenic downregulation in cancer, *hsa-let-7* miRNAs are classified as tumour suppressors[28,31]. A link between *hsa-let-7a* expression and clinical characteristics has been shown in triple negative breast cancer[32,33], and roles for *hsa-let-7a* in breast tumour development and metastasis have been hypothesised[34,35]. Multiple studies also reported a role for *hsa-let-7a* in controlling AURKA expression[19,24,31,36]. Indeed, we recently described a post-transcriptional pathway of AURKA regulation through APA of AURKA mRNA that underlies cell cycle-dependent translational efficiency[24]. We found that perturbation of AURKA APA allowed AURKA expression to evade regulation by the tumour suppressor miRNA *hsa-let-7a* and was sufficient for acquisition of cellular properties associated with oncogenic transformation[24]. However, the extent of deregulated post-transcriptional events influencing AURKA expression across cancers is currently unknown.

Several studies have determined the expression profile of *AURKA* and its prognostic significance in a wide range of cancers using whole-genome datasets from TCGA[4,37–41], although limited by two main caveats. First, these studies used incongruent datasets that included patients who had undergone prior treatments or who had been diagnosed with other types of cancers or metastases, as well as datasets that do not meet TCGA standards. In these studies cancer expression information was downloaded directly from user-friendly online resources, which perform downstream analysis of the entire group of datasets by TCGA, such as UALCAN[42] or GEPIA2[43]. Second, the existence of multiple AURKA mRNA isoforms[44] was never considered. Given the key role of post-transcriptional control in cancer[45], analyses of isoform-specific expression may reveal mechanisms converging on mRNA that underlie dysregulated *AURKA* expression and that the studies carried out to date have overlooked.

In this article, analysis of protein and mRNA expression of *AURKA* in all cancers was performed using genomic datasets available from TCGA and explored in previous published studies[4,37–40]. However, we improved the accuracy and reliability of the analysis by filtering datasets from patients with unusual medical histories or from those who have undergone specific prior treatments, as well as non-standard datasets as determined by TCGA. Additionally, we sought to understand the role of post-transcriptional regulation in *AURKA* expression in cancer. Based on our previous finding that *hsa-let-7a* miRNA and APA of AURKA mRNA can control AURKA protein levels[24], we explored the extent to which such association may provide a general mechanism for AURKA overexpression in cancer. Based on our findings, we hypothesise that AURKA protein expression in cancer could result from dynamic post-transcriptional regulation by *hsa-let-7a* and APA, in combination or independently.

MATERIALS AND METHODS

Pan-cancer analysis of *AURKA* and *hsa-let-7a* expression. Public datasets of harmonized gene, miRNA and protein expression quantification of 18 primary tumour types from TCGA[41] (**Table 1**) were downloaded directly from the TCGA repository. Only solid tumours and tumours for which at least three datasets of matched normal tissues existed were considered for analysis. Datasets were excluded if they: (i) were derived from patients under systemic, radiation, neoadjuvant, hormone, or other treatments prior to the malignancy, or concomitant to the malignancy but claimed unsuitable by TCGA; (ii) were derived from patients previously or concomitantly diagnosed with other types of cancers or with metastases; (iii) were derived from patients with recurrent tumours; (iv) did not meet TCGA study protocols or did not have clear tumour classification. R[46] was used to manage downloaded datasets and perform downstream analysis. Data points relative to AURKA mRNA and *hsa-let-7a* miRNA expression were extracted from transcriptome profiling datasets and plotted as fragments per kilobase of transcript per million mapped reads (FPKM) and as reads per million miRNA mapped (RPM), respectively, in $\log_2$ scale. Abundance of mature *hsa-let-7a* was calculated as the sum of reads from all miRNA IDs corresponding to the same unique MIMAT identifier (MIMAT0000062) from miRBase[47]. Data

points of AURKA protein expression were extracted from proteome profiling datasets and plotted as $\log_2$ arbitrary units (a.u.). Bioinformatics pipelines used by TCGA to generate mRNA, miRNA, and protein expression quantification datasets can be found at docs.gdc.cancer.gov/Data/Introduction.

Correlational analysis. RStudio[46] was used to match variables to their originating sample according to the Sample ID for each TCGA cancer project. The variables were first matched in pairs; paired data points belonging to individual samples were then plotted in XY graphs to perform the Spearman's rank test, and the strength of correlation was expressed as value of the Spearman's rank correlation coefficient (r). Degrees of correlation were considered as follows: r < |0.1| none; r = [|0.1|-|0.3|] weak; r = [|0.3|-|0.5|] moderate; r = [|0.5|-|0.7|] strong; r > |0.7| very strong. p values were corrected using the Bonferroni test.

Clustering analysis. *k*-means clustering was used to partition two observations into clusters, in which each cancer belongs to the cluster with the nearest mean of the observations' values (cluster centroid) serving as a prototype of the cluster. The total sum of the squared distance between centroid and each member of the cluster (i.e., total SSE) was 0.2 and the SSE per cluster was ≤ 0.1.

Pan-cancer analysis of AURKA PAS usage. Mapped RNA-seq datasets in .bam format from the 18 selected TCGA solid primary tumours (**Table 1**) were obtained with permission of the NIH, and were analysed using the public software APAtrap[48] according to the developers' instructions. The analysis was only directed at *AURKA* and *CCND1* genes. GRCh38/hg38 was used as reference genome and the GENCODE release 36[49] was used for gene annotations. The analysis was run using APAtrap default parameters but with a minimum average coverage required for each 3'UTR of 10. Data relative to APA isoform usage of the genes of interest were plotted as short/long ratio (SLR) of individual samples, or as mean cancer to normal fold change of SLR values for individual normal-cancer sample pairs (i.e., SLR fold change).

RESULTS

Pan-cancer analysis of AURKA mRNA and protein expression

We first explored AURKA mRNA expression levels across cancers. Before performing a pan-cancer analysis of AURKA expression, we first explored whether the inclusion or exclusion of unusual samples affected fold change estimates between cancer and normal tissues. To do this, two pools of TCGA gene expression quantification datasets were created (**Table 1**). One contained the entire set of datasets from both cancer and normal samples as downloadable from TCGA (n = 8483). The second pool (n = 7758) excluded datasets from irregular cases, which were from both cancer and normal tissue samples. The difference between the fold change of median AURKA FPKM in cancer over normal samples when the entire pool is used (a) and when the filtered pool is used (b) was calculated as: difference (%) = [(a – b)/b] x 100. While this difference was < |5|% in 13 cancers, it was ~5–10% in four cancers (ESCA, LUSC, PRAD, LUAD), and ~24% in BLCA (**Table 1**). Differences were observed with both positive and negative trends, indicating that using nonconforming datasets can lead to both overestimate and underestimate of changes in AURKA mRNA expression between cancer and normal samples.

The pan-cancer analysis of AURKA mRNA expression revealed a general trend of increased mRNA levels in cancer tissues, although to varying extents (**Figure 1A**). Other studies reported different statistical significances due to the use of alternative statistical tests and a different number of control samples[4,37–40]. THCA represents the only example in which AURKA mRNA is lower in cancer compared to normal tissue, although its expression is overall very low, and this difference is therefore non-significant (**Table 1**). Most cancer samples showed an extensive range of data distribution, indicating variable levels of AURKA mRNA expression among cancer patients. The median AURKA mRNA expression across normal tissues was tissue-dependent, and the overall AURKA mRNA levels in cancer also varied across tissues suggesting tissue-specific dysregulation.

Next, how AURKA mRNA expression correlated with protein expression ('protein vs mRNA') was examined in cancer samples (**Figure 1B**, **Supplementary Figure 1**). Proteomic datasets were also filtered for irregularities before analysis. For all the cancers, except PRAD and THCA, there existed a positive correlational relationship (r > 0.1) of varying degrees between AURKA mRNA and protein expression levels (**Figure 1B**). The wide range of r coefficient values suggests that mechanisms exist that modulate the coupling of AURKA mRNA and protein expression (**Figure 1C**). High r values would indicate that AURKA protein levels are dictated by mechanisms that determine the number of its mRNA molecules, such as gene copy number variation (CNV), the rate of transcription or of mRNA degradation. Contrarily, low r values might indicate that AURKA protein abundance likely results from mechanisms instead influencing the rate of translation or of protein degradation. Therefore, mechanisms governing AURKA translation or degradation play a greater role in cancers with lower AURKA mRNA-protein correlation (**Figure 1C**).

Pan-cancer analysis of *hsa-let-7a* miRNA expression

Since aberrant AURKA protein expression in cancer may derive from alterations in post-transcriptional events, this was further investigated in light of the reported role of *hsa-let-7a* miRNA in controlling AURKA expression[19,24,31,36]. Initially, the selected 18 cancers were subjected to analysis of *hsa-let-7a* expression levels. Two pools of datasets were created, one containing the entire set of datasets from both cancer and normal samples as downloaded from TCGA (n = 8376), and a second one that excluded datasets pertaining to irregular cases (n = 7702) (**Table 2**). The percentage difference (%) between the fold change of median *hsa-let-7a* RPM in cancer over normal samples when the entire pool is used and when the filtered pool is used was calculated as per **Table 1**. While the difference was minimal for nine cancers (< |1|%), it was |1–5|% for six cancers (BLCA, BRCA, COAD, HNSC, LIHC, THCA), and ~|6–9|% for three cancers (ESCA, KIRP, LUAD). Similarly to the analysis of AURKA mRNA expression, differences could be either positive or negative (**Table 2**).

The levels of *hsa-let-7a* expression in the selected cancers compared to corresponding normal tissues are shown in **Figure 2A**. Despite *hsa-let-7a*'s status as a known tumour suppressor miRNA[28,31,36], and therefore the expectation that its expression is down-regulated in cancer, five cancers (BLCA, CESC, ESCA, PAAD, PRAD) showed increased *hsa-let-7a* expression compared to normal tissues (**Figure 2A**, **Table 2**). Two cancers (STAD, UCEC) did not show any particular change in *hsa-let-7a* expression compared to their respective controls. On the other hand, 11 cancers showed a decrease in *hsa-let-7a* expression compared to normal tissues, which was statistically significant for all with the exception of KIRP. By examining *hsa-let-7a* expression in the normal samples across tissue types, it is evident that this varied considerably and was therefore tissue-dependent[30]. The general level of *hsa-let-7a* expression in cancer was also fluctuating across tissue types, implying tissue-specific mechanisms of dysregulation.

Involvement of *hsa-let-7a* in AURKA expression across cancers

Subsequently, the correlation between AURKA mRNA and *hsa-let-7a* expression levels ('mRNA vs *hsa-let-7a*') in individual cancer samples was interrogated as a route to infer the role of *hsa-let-7a* in modulating AURKA expression (**Figure 2B, Supplementary Figure 2A**). A correlation was found to exist (r ≥ |0.1|) with statistical significance for six cancers (BRCA, LUAD, LUSC, PRAD, THCA, UCEC). Nonetheless, in accordance with the role of *hsa-let-7a* in suppressing AURKA expression[19,24,31,36], negative r coefficients were measured in all cases where r > |0.1|, except in PRAD, despite the general lack of statistical significance. A potentially existing correlation suggests that *hsa-let-7a* may act upon AURKA mRNA stability, whereas lack of correlation may not exclude that *hsa-let-7a* controls AURKA expression by tuning its translation, provided that a link exists between *hsa-let-7a* and AURKA protein expression.

The relationship between AURKA protein expression and *hsa-let-7a* expression ('protein vs *hsa-let-7a*') was then explored for the 18 cancers of interest (**Figure 2C, Supplementary Figure 2B**). An

existing correlation (r ≥ |0.1|) with statistical significance was found for two cancers (BRCA, UCEC). In addition, negative r coefficients were measured for all cancers where r > |0.1|, with the exception of HNSC and PAAD. The correlations between *hsa-let-7a* and AURKA protein levels were overall weaker than those measured between *hsa-let-7a* and AURKA mRNA expression but might reflect a connection between *hsa-let-7a* and protein levels, underpinned by a role of *hsa-let-7a* in either AURKA mRNA abundance or translation rate, according to whether a link exists between *hsa-let-7a* and AURKA mRNA or not. A clustering analysis was therefore carried out between two chosen observations: r coefficients of the 'protein vs mRNA' correlation; r coefficients of the 'protein vs *hsa-let-7a*' correlation. Five clusters were identified (**Figure 2D**), the members and features of which, as well as speculations on possible involvements of *hsa-let-7a* in AURKA expression, are summarized in **Figure 2E**. We point to a heterogeneous landscape of how *hsa-let-7a* might be implicated in the regulation of AURKA expression in cancer. Based on our analysis, the involvement of *hsa-let-7a* could be hypothesised to range from being null to negatively controlling the rate of degradation or of translation of AURKA mRNA.

Pan-cancer profiling of AURKA APA mRNA isoforms

One source of variability in the correlation between *hsa-let-7a* and AURKA expression in different cancers could arise from differences in AURKA mRNA processing that would determine sensitivity to *hsa-let-7a* targeting. For example, we previously showed that *hsa-let-7a* targets only the long APA isoform of AURKA mRNA[24] and a different short/long ratio (SLR) of AURKA APA isoforms is present between several types of normal and cancer cells[24,50–53]. Here, pre-processed RNA-seq datasets from the selected TCGA cancers were subjected to analysis of PAS usage using APAtrap[48]. The APAtrap analysis only detected two of the AURKA transcripts annotated in ENSEMBL, namely ENST00000371356.6 (AURKA-203) and ENST00000395914.5 (AURKA-207). These transcripts share the same 3' UTR region, therefore the APAtrap results are identical, and we only report those for AURKA-203. The depth of coverage of reads mapped to the reference genome in is shown in **Figure 3A** for representative BRCA primary tumour samples, which show different AURKA SLRs. Most cancers showed an increase in mean AURKA SLR (AURKA-203) in cancer samples compared to respective matched normal samples (SLR fold change > 1) (**Figure 3B**). In most cases the cancer-dependent increase in mean SLR was more than 1.5-fold. In particular, BRCA showed an increase of ~2.5 fold change in mean SLR, and this recapitulates results that others have obtained using different methods[24,50–52]. AURKA-207 followed the same trends as AURKA-203 (data not shown). As a positive control to our analysis, the APAtrap analysis was extended to the cyclin D1 (CCND1) mRNA, also known to have an increased SLR in cancer[54–57]. We found that the cancer/normal fold change of mean CCND1 SLR values was > 1 in most cases (**Figure 3C**), confirming the ability of APAtrap to report on APA events. In summary, our data indicate that APA represents a mechanism of AURKA post-transcriptional dysregulation. Given the higher translation rate of the short APA isoform[24], APA likely contributes to the increased expression of AURKA in cancers.

Interplay of APA and *hsa-let-7a* in regulating AURKA expression

We then examined the extent to which APA is linked to AURKA expression and how it may be associated with *hsa-let-7a* in mediating AURKA expression across cancers. First, a positive correlation (r > 0.1) between AURKA SLR and protein expression levels ('protein vs SLR') was found in six cancers (**Figure 4A, Supplementary Figure 3A**), suggesting that APA modulates AURKA expression in these cancers (UCEC, BRCA, KIRP, LUAD, READ, KIRC). Furthermore, the 'protein vs SLR' correlation was analysed in light of the cancer-dependent changes in AURKA SLR (**Figure 4B**). A positive correlation was found (r = 0.2) when we estimated how the correlation of AURKA protein expression with its SLR varied according to SLR change in different cancers, although this has no statistical significance (p = 0.714). Our findings may potentially be consistent with the idea that APA positively contributes to shaping AURKA protein expression in certain cancers.

We then probed the correlation of AURKA SLR to AURKA mRNA expression ('mRNA vs SLR') across the six cancers (**Figure 4C, Supplementary Figure 3B**). Despite the degree of uncertainty in our measurements, these observations may be consistent with an effect of APA on AURKA mRNA stability in READ, BRCA, and UCEC, and lack of correlation in LUAD, KIRC, and KIRP instead possibly points to an effect on AURKA translation in these cancers. Finally, analysis of the correlation between AURKA SLR and *hsa-let-7a* expression levels ('*hsa-let-7a* vs SLR') suggests varying extents to which *hsa-let-7a* might be associated with AURKA APA. Different trends could be identified when AURKA 'protein vs SLR' correlation was compared to the '*hsa-let-7a* vs SLR' correlation (**Figure 4D**). This raises the hypothesis that APA could influence AURKA expression via interplay with *hsa-let-7a* (UCEC, BRCA, KIRP, READ) or independently of *hsa-let-7a* (LUAD, KIRC).

Heterogeneity of AURKA splicing isoforms across cancers

Seventeen high-scoring AURKA transcript variants have been annotated to date, nine of which are listed in the NCBI database[58] and eight in the Ensembl database[59] (**Figure 5A**). The 'matched annotation from NCBI and EMBL-EBI' (MANE)[60] label was recently given to two AURKA transcripts that were revealed identical (NM_198437.3 and ENST00000395915.8/AURKA-208). The 16 distinct isoforms are 5' untranslated region (UTR) splicing variants, resulting from alternative splice sites and exon skipping. The coding sequence and the 3'UTR instead follow canonical splicing, except for ENST00000395907.5 (AURKA-204), whose coding sequence includes the last intron and has a truncated 3'UTR; this transcript variant is however characterized by a poorer annotation score.

The public web server GEPIA2[43] was consulted to retrieve information on the distribution of AURKA transcripts across TCGA cancers. **Figure 5B** shows the expression levels of each of the eight Ensembl transcripts in the selected TCGA cancers. It is evident that each isoform has an individual profile of expression across cancers. In particular, some transcripts are abundantly expressed in most cancers, whereas others are expressed only in specific cancers. Reciprocally, some cancers express most isoforms quite equally, while other cancers express only a specific group of transcripts. Since these are alternatively spliced isoforms in the 5'UTR with different combinations of exons[44] it can be concluded that splicing and unknown features of the 5'UTR play a role in determining the context-dependent expression profile of AURKA mRNA.

DISCUSSION

This article describes an exploratory analysis around the expression of *AURKA* on a pan-cancer scale. Following the analysis of publicly available data from TCGA[41], it was clear that *AURKA* overexpression at the level of the mRNA is prevalent in many cancers, a finding previously reported[4,37–40]. However, our analyses of TCGA datasets avoided TCGA data-processing servers (e.g., UALCAN or GEPIA2). In addition, we carefully selected datasets to exclude those with annotations containing 'disqualifying' information about patients or samples. TCGA itself suggests reviewing datasets according to their accompanying annotations prior to running an analysis, although this practice is not commonly observed. Importantly, since a great number of flagged datasets derive from patients having received radiation, hormone, or other treatments, removing nonconforming datasets counteracts the bias in gene expression introduced by treatment-derived selective pressure, which constitutes a basis of tumour heterogeneity[61].

We investigated to what extent post-transcriptional regulation is responsible for defining and dysregulating *AURKA* expression in human cancers. To this aim, the extent to which the abundance of AURKA protein follows changes in the abundance of AURKA mRNA was measured across TCGA cancer samples. Our results suggest that mechanisms affecting AURKA translation or protein degradation are involved in tuning AURKA protein expression in various cancers, with a role of increasing importance as the degree of mRNA-protein correlation decreases. This has already been proposed as a general mechanism of gene expression regulation[62–64], and is concordant with results reported by others specifically regarding AURKA expression[21,65–67].

Investigations of the possible involvement of *hsa-let-7a* miRNA in AURKA expression in cancer were also performed. First, analysis of expression levels of *hsa-let-7a* indicated that *hsa-let-7a* is downregulated in most cancers, in accordance with its tumour suppressor role[28,31,36], however with some exceptions. Such pan-cancer analysis of *hsa-let-7a* expression has no precedent in the literature to our knowledge. Our data support the idea that *hsa-let-7a* tunes the expression of AURKA in some cancers. Weak relationships were generally found across all the cancers, apparently inconsistent with a hypothesis that AURKA mRNA is a major target of *hsa-let-7a*[19,31] unlike other oncogenes like RAS and MYC[68]. We hypothesise that mechanisms influencing the ability of *hsa-let-7a* to target AURKA mRNA could limit the coupling between *hsa-let-7a* abundance and AURKA expression measurable in our pan-cancer analysis. For example, *hsa-let-7a*'s targeting ability can indeed be subject to contextual regulation, such as cell cycle phase-dependent, since AURKA mRNA is specifically repressed by *hsa-let-7a* in G_1 and S phases but not in G_2[24].

Other groups have used alternative software tools to measure genome-wide APA changes in TCGA samples, but in these studies AURKA SLR changes were not reported in the published lists of genes with cancer-dependent changes in SLR[69–74]. In contrast, other studies using alternative sample sources and methods of analysis found AURKA SLR increased in cancer samples, specifically in breast and lung cancers, where such increase even correlated with poor prognosis and survival[24,50,52]. Here, using the APAtrap software[48] to interrogate our curated TCGA samples, AURKA SLR was discovered to have increased in almost all cancers compared to their normal tissues. Results presented here using high quality RNA-seq data are therefore consistent with what has previously been observed in breast and lung cancers using microarray data[24,50,52], and expand the observation to other cancers.

Overall, the diverse results on AURKA cancer-specific SLR across studies are not yet conducive to a unifying hypothesis for the APA-mediated regulation of AURKA. Interestingly, transcript AURKA-203 contains exon III and transcript AURKA-207 contains exon II, both of which are thought to be implicated in breast cancer[75,76], which is also the type of cancer showing the second highest change in AURKA SLR as recorded by the APAtrap analysis. Furthermore, while APAtrap exposes the abundance of one 3'UTR isoform compared to another of the same gene, abundance levels of each isoform can be otherwise influenced by mechanisms linked to alternative splicing. It is highly possible that there exists a combination of AURKA APA and alternative splicing isoforms in different contexts.

In conclusion, the data presented here cumulatively suggest that post-transcriptional mechanisms can constitute the primary form of regulation of the expression of AURKA in some cancers. As a consequence, impairments in mechanisms affecting mRNA dynamics offer a substantial basis for alterations in AURKA protein expression in cancer. This is a previously underrated notion, as the basis for AURKA oncogenic activation by means of overexpression has generally been attributed to enhanced gene copy number, transcription, or protein stability. We therefore add a new facet to the complex oncogenic expression program of the *AURKA* cancer gene, highlighting molecular mechanisms that could represent actionable targets of both DNA- and RNA-based therapeutics.

DATA AVAILABILITY
Mapped RNA-sequencing data underlying this article were provided by the TCGA Research Network by permission. Any other data underlying this article are freely available in the TCGA repository at https://portal.gdc.cancer.gov.

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

REFERENCES

1. Willems, E. *et al.* The functional diversity of Aurora kinases: a comprehensive review. *Cell Div.* **13**, 7 (2018).

2. Xu, J. *et al.* Aurora-A identifies early recurrence and poor prognosis and promises a potential therapeutic target in triple negative breast cancer. *PLoS One* **8**, e56919 (2013).

3. Zhang, J., Li, B., Yang, Q., Zhang, P. & Wang, H. Prognostic value of Aurora kinase A (AURKA) expression among solid tumor patients: a systematic review and meta-analysis. *Jpn. J. Clin. Oncol.* **45**, 629–636 (2015).

4. Du, R., Huang, C., Liu, K., Li, X. & Dong, Z. Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy. *Mol. Cancer* **20**, 15 (2021).

5. Bavetsias, V. & Linardopoulos, S. Aurora Kinase Inhibitors: Current Status and Outlook. *Front. Oncol.* **5**, 278 (2015).

6. Wang, R. *et al.* Selective targeting of non-centrosomal AURKA functions through use of a targeted protein degradation tool. *Commun. Biol.* **4**, 640 (2021).

7. Bozilovic, J. *et al.* Novel, highly potent PROTACs targeting AURORA-A kinase. *Curr. Res. Chem. Biol.* **2**, 100032 (2022).

8. Hata, T. *et al.* RNA interference targeting aurora kinase a suppresses tumor growth and enhances the taxane chemosensitivity in human pancreatic cancer cells. *Cancer Res.* **65**, 2899–2905 (2005).

9. Lu, Y. *et al.* Knocking down the expression of Aurora-A gene inhibits cell proliferation and induces G2/M phase arrest in human small cell lung cancer cells. *Oncol. Rep.* **32**, 243–249 (2014).

10. dos Santos, E. O., Carneiro-Lobo, T. C., Aoki, M. N., Levantini, E. & Bassères, D. S. Aurora kinase targeting in lung cancer reduces KRAS-induced transformation. *Mol. Cancer* **15**, 12 (2016).

11. Zhong, N. *et al.* Silencing Aurora-A with siRNA inhibits cell proliferation in human lung adenocarcinoma cells. *Int. J. Oncol.* **49**, 1028–1038 (2016).

12. Wang, C., Yan, Q., Hu, M., Qin, D. & Feng, Z. Effect of AURKA Gene Expression Knockdown on Angiogenesis and Tumorigenesis of Human Ovarian Cancer Cell Lines. *Target. Oncol.* **11**, 771–781 (2016).

13. Yang, Y. *et al.* Silencing of AURKA augments the antitumor efficacy of the AURKA inhibitor MLN8237 on neuroblastoma cells. *Cancer Cell Int.* **20**, 9 (2020).

14. Marumoto, T., Zhang, D. & Saya, H. Aurora-A - a guardian of poles. *Nat. Rev. Cancer* **5**, 42–50 (2005).

15. Wu, P.-J. *et al.* The impact of Aurora kinase A genetic polymorphisms on cervical cancer progression and clinicopathologic characteristics. *Int. J. Med. Sci.* **18**, 2457–2465 (2021).

16. Ma, Z.-L. *et al.* Tanshinones suppress AURKA through up-regulation of miR-32 expression in non-small cell lung cancer. *Oncotarget* **6**, 20111–20120 (2015).

17. Yuan, Y. L., Yu, H., Mu, S.-M., Dong, Y. D. & Li, D. Y. MiR-26a-5p Inhibits Cell Proliferation and Enhances Doxorubicin Sensitivity in HCC Cells via Targeting AURKA. *Technol. Cancer Res. Treat.* **18**, 1533033819851833 (2019).

18. Fadaka, A. O., Sibuyi, N. R. S., Madiehe, A. M. & Meyer, M. MicroRNA-based regulation of Aurora A kinase in breast cancer. *Oncotarget* **11**, 4306–4324 (2020).

19. Liu, X. *et al.* Tanshinone Inhibits NSCLC by Downregulating AURKA Through Let-7a-5p. *Front. Genet.* **11**, 838 (2020).

20. Zhang, H., Bao, J., Zhao, S., Huo, Z. & Li, B. MicroRNA-490-3p suppresses hepatocellular carcinoma cell proliferation and migration by targeting the aurora kinase A gene (AURKA). *Arch. Med. Sci.* **16**, 395–406 (2020).

21. Lai, C.-H. *et al.* Translational up-regulation of Aurora-A in EGFR-overexpressed cancer. *J. Cell. Mol. Med.* **14**, 1520–1531 (2010).

22. Dobson, T., Chen, J. & Krushel, L. A. Dysregulating IRES-dependent translation contributes to overexpression of oncogenic Aurora A Kinase. *Mol. Cancer Res.* **11**, 887–900 (2013).

23. Lai, C.-H. *et al.* Translational upregulation of Aurora-A by hnRNP Q1 contributes to cell proliferation and tumorigenesis in colorectal cancer. *Cell Death Dis.* **8**, e2555 (2017).

24. Cacioppo, R., Akman, H. B., Tuncer, T., Erson-Bensan, A. E. & Lindon, C. Differential translation of mRNA isoforms underlies oncogenic activation of cell cycle kinase Aurora A. *Elife* **12**, RP87253 (2023).

25. Gruber, A. J. & Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* **20**, 599–614 (2019).

26.   Gruber, A. J. *et al.* A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. *Genome Res.* **26**, 1145–1159 (2016).

27.   Derti, A. *et al.* A quantitative atlas of polyadenylation in five mammals. *Genome Res.* **22**, 1173–1183 (2012).

28.   Bueno, M. J. & Malumbres, M. MicroRNAs and the cell cycle. *Biochim. Biophys. Acta* **1812**, 592–601 (2011).

29.   Ghafouri-Fard, S., Shoorei, H., Anamag, F. T. & Taheri, M. The Role of Non-Coding RNAs in Controlling Cell Cycle Related Proteins in Cancer Cells. *Frontiers in Oncology* vol. 10 (2020).

30.   Lee, H., Han, S., Kwon, C. S. & Lee, D. Biogenesis and regulation of the let-7 miRNAs and their functional implications. *Protein Cell* **7**, 100–113 (2016).

31.   Johnson, C. D. *et al.* The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer Res.* **67**, 7713–7722 (2007).

32.   Avery-Kiejda, K. A., Braye, S. G., Mathe, A., Forbes, J. F. & Scott, R. J. Decreased expression of key tumour suppressor microRNAs is associated with lymph node metastases in triple negative breast cancer. *BMC Cancer* **14**, 51 (2014).

33.   Turashvili, G. *et al.* Novel prognostic and predictive microRNA targets for triple-negative breast cancer. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* fj201800120R (2018) doi:10.1096/fj.201800120R.

34.   Thammaiah, C. K. & Jayaram, S. Role of let-7 family microRNA in breast cancer. *Non-coding RNA Res.* **1**, 77–82 (2016).

35.   Shi, Y. *et al.* Let-7a-5p inhibits triple-negative breast tumor growth and metastasis through GLUT12-mediated warburg effect. *Cancer Lett.* **495**, 53–65 (2020).

36.   Barh, D., Malhotra, R., Ravi, B. & Sindhurani, P. MicroRNA let-7: an emerging next-generation cancer therapeutic. *Curr. Oncol.* **17**, 70–80 (2010).

37.   Shan, B. *et al.* AURKA Increase the Chemosensitivity of Colon Cancer Cells to Oxaliplatin by Inhibiting the TP53-Mediated DNA Damage Response Genes. *Biomed Res. Int.* **2020**, 8916729 (2020).

38.   Miralaei, N., Majd, A., Ghaedi, K., Peymani, M. & Safaei, M. Integrated pan-cancer of AURKA expression and drug sensitivity analysis reveals increased expression of AURKA is responsible for drug resistance. *Cancer Med.* **10**, 6428–6441 (2021).

39.   Almilaibary, A. Targeting aurora kinase a (AURKA) in cancer: molecular docking and dynamic simulations of potential AURKA inhibitors. *Med. Oncol.* **39**, 246 (2022).

40.   Lu, K., Yuan, X., Zhao, L., Wang, B. & Zhang, Y. Comprehensive pan-cancer analysis and the regulatory mechanism of AURKA, a gene associated with prognosis of ferroptosis of adrenal cortical carcinoma in the tumor micro-environment. *Front. Genet.* **13**, 996180 (2022).

41.   Tomczak, K., Czerwińska, P. & Wiznerowicz, M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol.* **19**, A68-77 (2015).

42.   Chandrashekar, D. S. *et al.* UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* **19**, 649–658 (2017).

43.   Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560 (2019).

44.   Cacioppo, R. & Lindon, C. Regulating the regulator: a survey of mechanisms from transcription to translation controlling expression of mammalian cell cycle kinase Aurora A. *Open Biol.* **12**, 220134 (2022).

45.   Obeng, E. A., Stewart, C. & Abdel-Wahab, O. Altered RNA Processing in Cancer Pathogenesis and Therapy. *Cancer Discov.* **9**, 1493–1510 (2019).

46.   Team, R. C. R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput. Vienna, Austria* (2022).

47.   Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).

48.   Ye, C., Long, Y., Ji, G., Li, Q. Q. & Wu, X. APAtrap: identification and quantification of alternative polyadenylation sites from RNA-seq data. *Bioinformatics* **34**, 1841–1849 (2018).

49.   Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).

50.   Lembo, A., Di Cunto, F. & Provero, P. Shortening of 3'UTRs correlates with poor prognosis in breast and lung cancer. *PLoS One* **7**, e31129 (2012).

51.   Akman, H. B., Oyken, M., Tuncer, T., Can, T. & Erson-Bensan, A. E. 3'UTR shortening and EGF signaling: implications for breast cancer. *Hum. Mol. Genet.* **24**, 6910–6920 (2015).

52.   Wang, L. *et al.* Dissecting the heterogeneity of the alternative polyadenylation profiles in triple-negative breast cancers. *Theranostics* **10**, 10531–10547 (2020).

53.   Chen, Y. *et al.* A systematic benchmark of Nanopore long read RNA sequencing for transcript level analysis in human cell lines. *bioRxiv* 2021.04.21.440736 (2021) doi:10.1101/2021.04.21.440736.

54.   Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).

55.   Wang, Q. *et al.* Cell Cycle Regulation by Alternative Polyadenylation of CCND1. *Sci. Rep.* **8**, 6824 (2018).

56.   Masamha, C. P. & Wagner, E. J. The contribution of alternative polyadenylation to the cancer phenotype. *Carcinogenesis* **39**, 2–10 (2018).

57.   Chan, J. J., Tabatabaeian, H. & Tay, Y. 3'UTR heterogeneity and cancer progression. *Trends Cell Biol.* **33**, 568–582 (2023).

58.   Sayers, E. W. *et al.* Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26 (2022).

59.   Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res.* **50**, D988–D995 (2022).

60.   Morales, J. *et al.* A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. *Nature* **604**, 310–315 (2022).

61.   Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).

62.   Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).

63.   Koussounadis, A., Langdon, S. P., Um, I. H., Harrison, D. J. & Smith, V. A. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci. Rep.* **5**, 10775 (2015).

64.   Franks, A., Airoldi, E. & Slavov, N. Post-transcriptional regulation across human tissues. *PLOS Comput. Biol.* **13**, e1005535 (2017).

65.   Zhou, H. *et al.* Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. *Nat. Genet.* **20**, 189–193 (1998).

66.   Gritsko, T. M. *et al.* Activation and overexpression of centrosome kinase BTAK/Aurora-A in human ovarian cancer. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **9**, 1420–1426 (2003).

67.   Jeng, Y.-M., Peng, S.-Y., Lin, C.-Y. & Hsu, H.-C. Overexpression and amplification of Aurora-A in hepatocellular carcinoma. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **10**, 2065–2071 (2004).

68.   Wang, X. *et al.* Regulation of let-7 and its target oncogenes. *Oncol. Lett.* **3**, 955–960 (2012).

69.   Xia, Z. *et al.* Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. *Nat. Commun.* **5**, 5274 (2014).

70.   Yang, X. *et al.* Genome-wide profiling reveals cancer-related genes with switched alternative polyadenylation sites in colorectal cancer. *Onco. Targets. Ther.* **11**, 5349–5357 (2018).

71.   Xiang, Y. *et al.* Comprehensive Characterization of Alternative Polyadenylation in Human Cancer. *J. Natl. Cancer Inst.* **110**, 379–389 (2018).

72.   Xue, Z. *et al.* Recurrent tumor-specific regulation of alternative polyadenylation of cancer-related genes. *BMC Genomics* **19**, 536 (2018).

73.   Zingone, A. *et al.* A comprehensive map of alternative polyadenylation in African American and European American lung cancer patients. *Nat. Commun.* **12**, 5605 (2021).

74.   Zhang, Y., Xu, Y. & Wang, Y. Alternative polyadenylation associated with prognosis and therapy in colorectal cancer. *Sci. Rep.* **12**, 7036 (2022).

75.   Li, C. *et al.* Cell type and culture condition-dependent alternative splicing in human breast cancer cells revealed by splicing-sensitive microarrays. *Cancer Res.* **66**, 1990–1999 (2006).

76.   Dago, D. N. *et al.* Estrogen receptor beta impacts hormone-induced alternative mRNA splicing in breast cancer cells. *BMC Genomics* **16**, 367 (2015).

Table 1: Comparison of the two pools of datasets for analysis of AURKA mRNA expression. For each pool (entire or filtered), the table displays the available number of RNA-seq datasets (n°), the median AURKA mRNA FPKM, and the ratio between tumour (T) to normal tissue (NT) of median FPKM, per cancer project and tissue type. The table also shows the % difference between the T / NT ratio when the filtered set is used and that when the entire set is used.

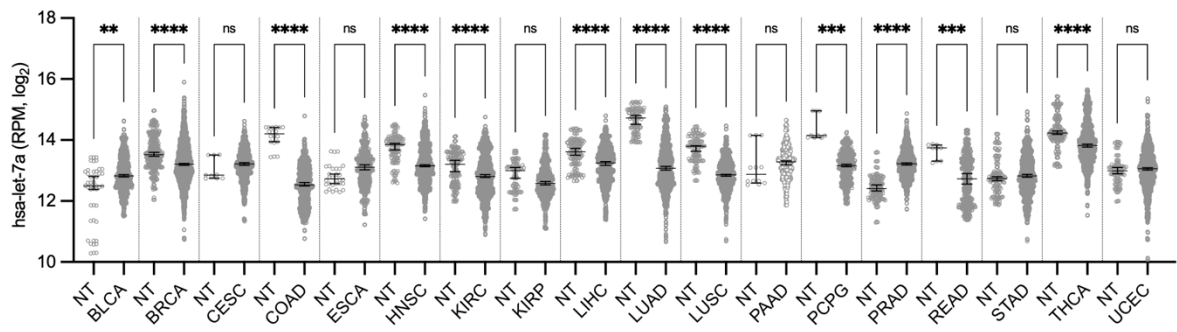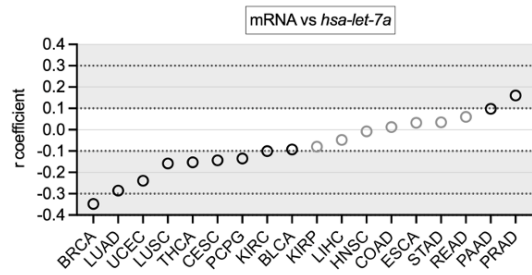|	Entire set			Filtered set			difference (%)
	n°	Median FPKM	Median FPKM(T) / Median FPKM(NT)	n°	Median FPKM	Median FPKM(T) / Median FPKM(NT)	
NT BLCA	19 412	1.29 12.04	9.33	15 362	1.03 11.89	11.54	23.68
NT BRCA	113 1118	0.93 7.33	7.88	110 1090	0.93 7.36	7.91	0.41
NT CESC	3 306	0.51 15.37	30.13	3 303	0.51 15.42	30.23	0.33
NT COAD	41 483	6.20 18.48	2.98	30 408	6.11 18.43	3.02	1.20
NT ESCA	13 185	1.92 12.36	6.43	12 181	1.74 12.36	7.10	10.34
NT HNSC	44 522	2.78 11.33	4.08	39 483	2.68 11.22	4.19	2.72
NT KIRC	72 542	1.04 1.71	1.64	64 427	1.04 1.70	1.63	-0.58
NT KIRP	32 291	1.00 1.99	1.99	28 261	0.98 2.02	2.06	3.58
NT LIHC	50 374	0.54 5.30	9.81	43 355	0.56 5.35	9.55	-2.67
NT LUAD	59 541	1.21 6.89	5.69	46 462	1.16 6.94	5.98	5.07
NT LUSC	51 502	1.17 12.33	10.54	43 458	1.11 12.46	11.23	6.52
NT PAAD	4 179	1.53 3.68	2.41	4 157	1.52 3.74	2.46	2.30
NT PCPG	3 184	0.78 1.36	1.74	3 181	0.78 1.35	1.73	-0.74
NT PRAD	52 502	0.63 1.11	1.76	44 475	0.60 1.12	1.87	5.95
NT READ	10 167	6.15 19.97	3.25	9 146	6.18 20.02	3.24	-0.24
NT STAD	36 412	2.82 13.53	4.80	36 376	2.82 13.46	4.77	-0.52
NT THCA	59 513	0.86 0.77	0.89	57 482	0.86 0.77	0.89	0.16
NT UCEC	35 554	0.67 7.98	11.91	34 521	0.69 7.90	11.45	-3.87
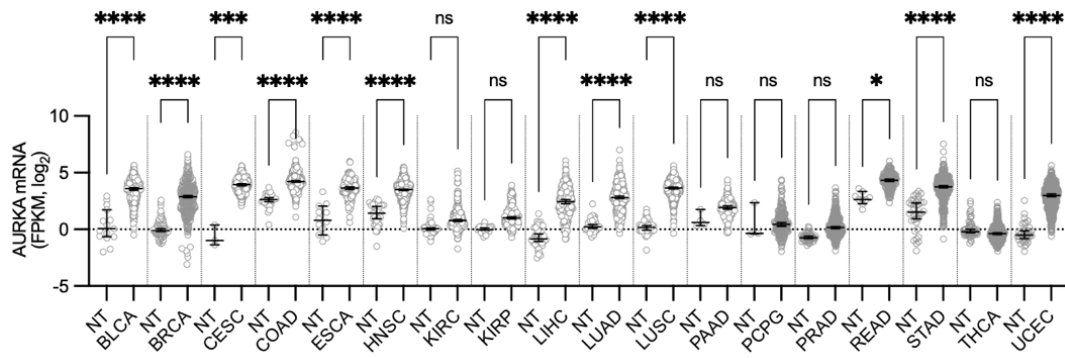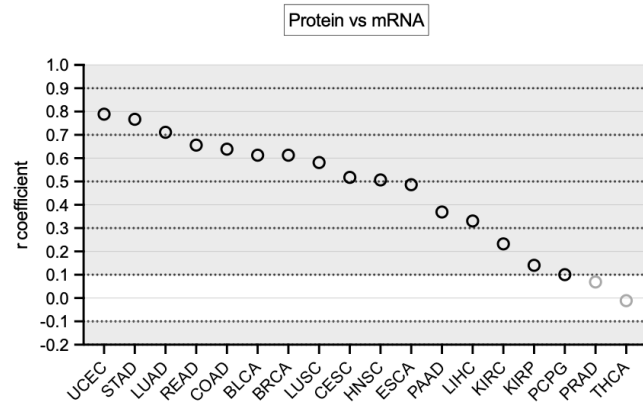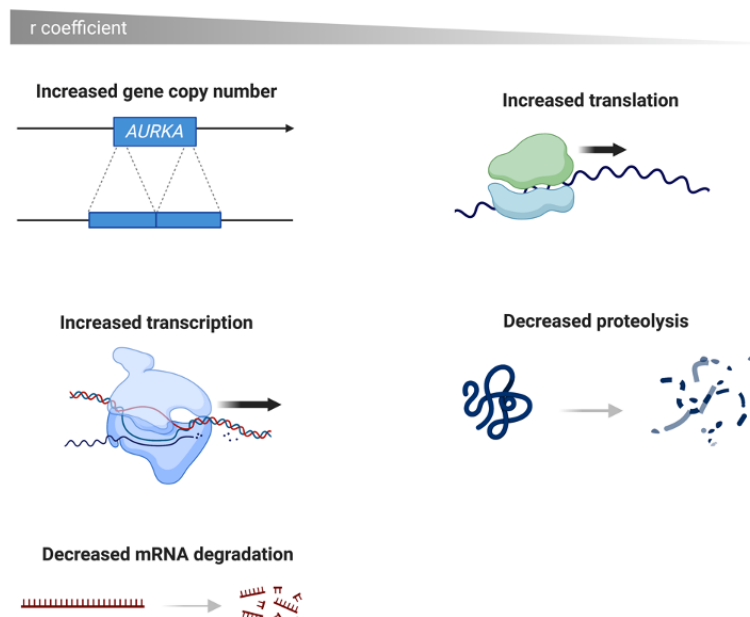
Figure 1: Pan-cancer analysis of AURKA mRNA and protein expression. (A) Median and 95% CI of AURKA mRNA FPKM values in 18 TCGA cancers compared to their respective normal tissue (NT). Number of datasets per condition shown in **Table 1**, filtered set. Kruskal-Wallis with Dunnett's multiple comparisons test. ns, not significant; *, p<0.02; ****, p<0.0001. **(B)** Distribution of r coefficients of the AURKA mRNA-protein correlation across cancers. **(C)** Graphic representation of molecular processes likely underlying the AURKA mRNA-protein correlation in cancer samples. Figure created with BioRender.com.

Table 2: Comparison of the two pools of datasets for analysis of *hsa-let-7a* expression. For each pool (entire or filtered), the table displays the available number of miRNA-seq datasets (n°), the median *hsa-let-7a* RPM, and the ratio between tumour (T) to normal tissue (NT) of median RPM, per cancer project and tissue type. The table also shows the % difference between the T / NT ratio when the filtered set is used and that when the entire set is used.

		Entire set			Filtered set		
	n°	Median RPM	Median RPM(T) / Median RPM(NT)	n°	Median RPM	Median RPM(T) / Median RPM(NT)	difference (%)
NT	57	5640		45	5780		
BLCA	1254	7371	1.31	1098	7250	1.25	-4.02
NT	312	11684		306	11760		
BRCA	3309	9528	0.82	3192	9474	0.81	-1.21
NT	9	7385		9	7385		
CESC	927	9557	1.30	897	9505	1.29	-0.54
NT	24	18908		18	18841		
COAD	1371	5803	0.31	1158	5919	0.31	2.36
NT	39	7258		30	6788		
ESCA	561	8864	1.22	552	8848	1.30	6.73
NT	132	14321		114	14751		
HNSC	1575	9241	0.65	1464	9128	0.62	-4.10
NT	213	9460		189	9460		
KIRC	1635	7149	0.76	1368	7153	0.76	0.06
NT	102	7486		87	8161		
KIRP	876	6194	0.83	786	6132	0.75	-9.19
NT	150	12050		129	12519		
LIHC	1125	9609	0.80	1068	9710	0.78	-2.73
NT	138	25917		99	27139		
LUAD	1563	8974	0.35	1344	8638	0.32	-8.08
NT	135	13831		117	14195		
LUSC	1434	7197	0.52	1254	7434	0.52	0.64
NT	12	7594		12	7594		
PAAD	537	9967	1.31	456	9974	1.31	0.07
NT	9	18148		9	18148		
PCPG	552	9179	0.51	549	9185	0.51	0.07
NT	156	5440		132	5440		
PRAD	1497	9507	1.75	1416	9551	1.76	0.46
NT	9	13717		9	13717		
READ	486	6748	0.49	435	6801	0.50	0.79
NT	135	6823		135	6823		
STAD	1338	7302	1.07	1275	7258	1.06	-0.60
NT	177	19193		171	19193		
THCA	1542	14643	0.76	1449	14483	0.75	-1.09
NT	99	8123		99	8123		
UCEC	1638	8549	1.05	1635	8554	1.05	0.06

Figure 2: Pan-cancer analysis of *hsa-let-7a* expression and involvement in AURKA expression. (A) Median and 95% CI of *hsa-let-7a* RPM values in 18 TCGA cancers compared to their respective normal tissue (NT). Number of datasets per condition shown in **Table 2**, filtered set. Kruskal-Wallis with Dunnett's multiple comparisons test. ns, not significant; **, p<0.01; ***, p<0.001; ****, p<0.0001. **(B)** Distribution of r coefficients of the AURKA mRNA-*hsa-let-7a* correlation across cancers. **(C)** Distribution of r coefficients of the AURKA protein-*hsa-let-7a* correlation across cancers. **(D)** Clusters found according to common trends in the two types of correlation. The r coefficient of the indicated correlation is plotted on each axis. **(E)** Table showing clusters composition and speculative involvement of *hsa-let-7a* in controlling AURKA expression.
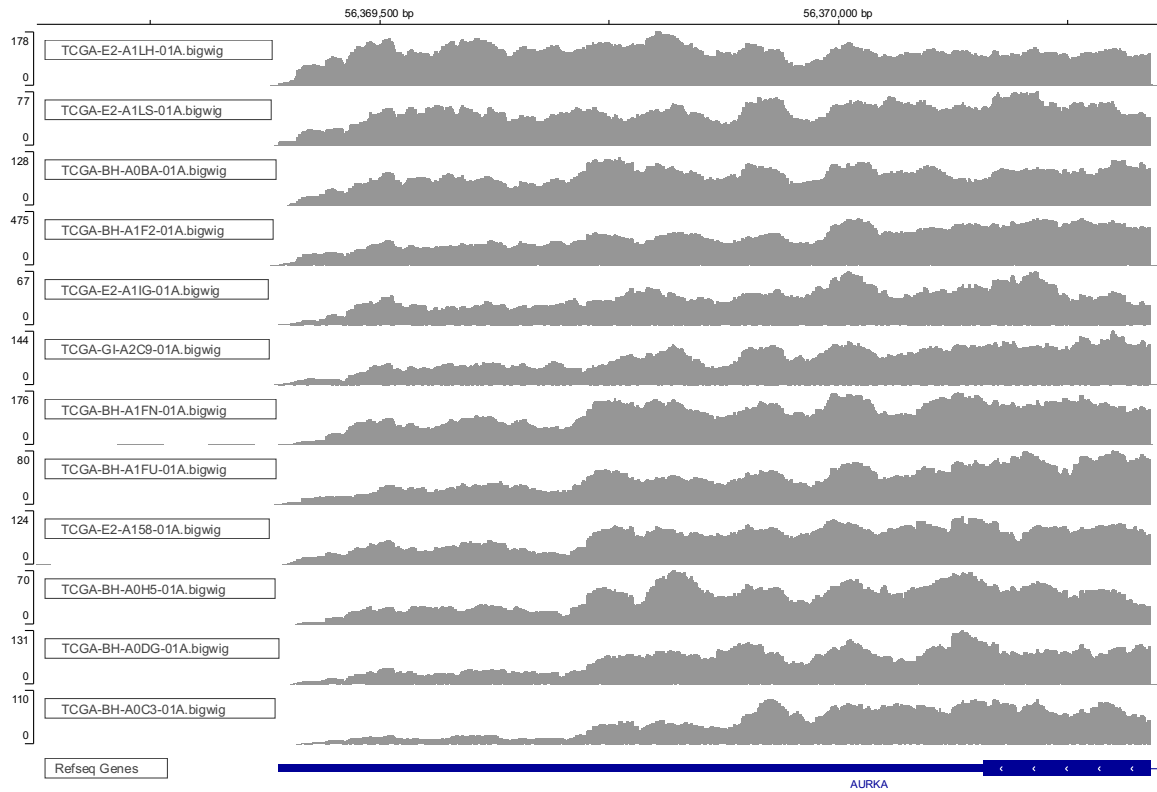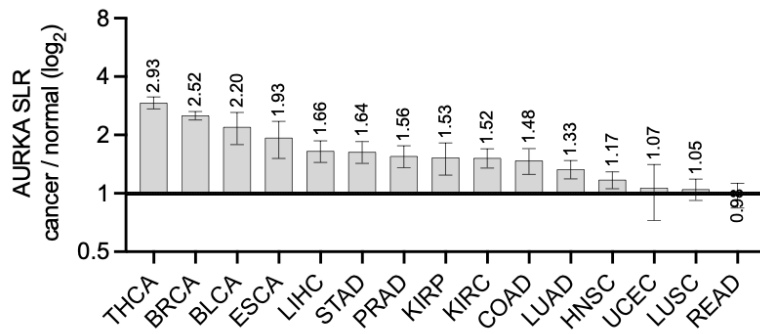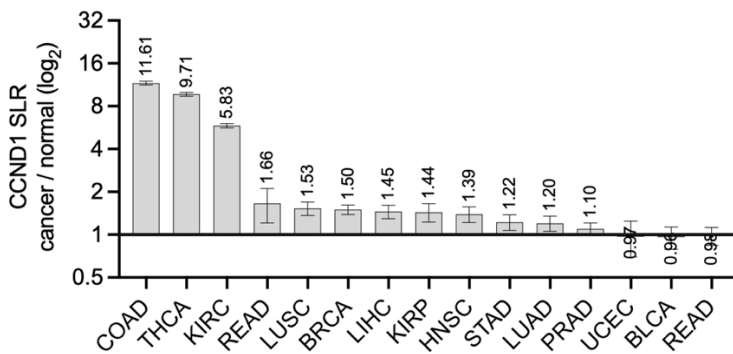
Figure 3: Pan-cancer profiling of AURKA APA mRNA isoforms. (A) Integrative Genomics Viewer (IGV) view of the depth of coverage of reads mapping to the 3' end of the AURKA-203 transcript (the last exon and 3' UTR are shown at the bottom, in blue). A subset of BRCA primary tumour samples is shown, selected to represent a range of SLRs, from low (top) to high (bottom). Reference genome GRCh38/hg38. Note, the y-axis scale (read depth) differs across panels. **(B)**, **(C)** Mean and standard error of the mean of SLR fold change values for cancer-normal sample pairs of AURKA transcript ENST00000371356.6 (B) and CCND1 transcript ENST00000227507.3 (C). Value of the mean indicated. SLR values were calculated using APAtrap and the analysis was performed once. n, number of cancer-normal sample pairs. Y axis in $\log_2$ scale.
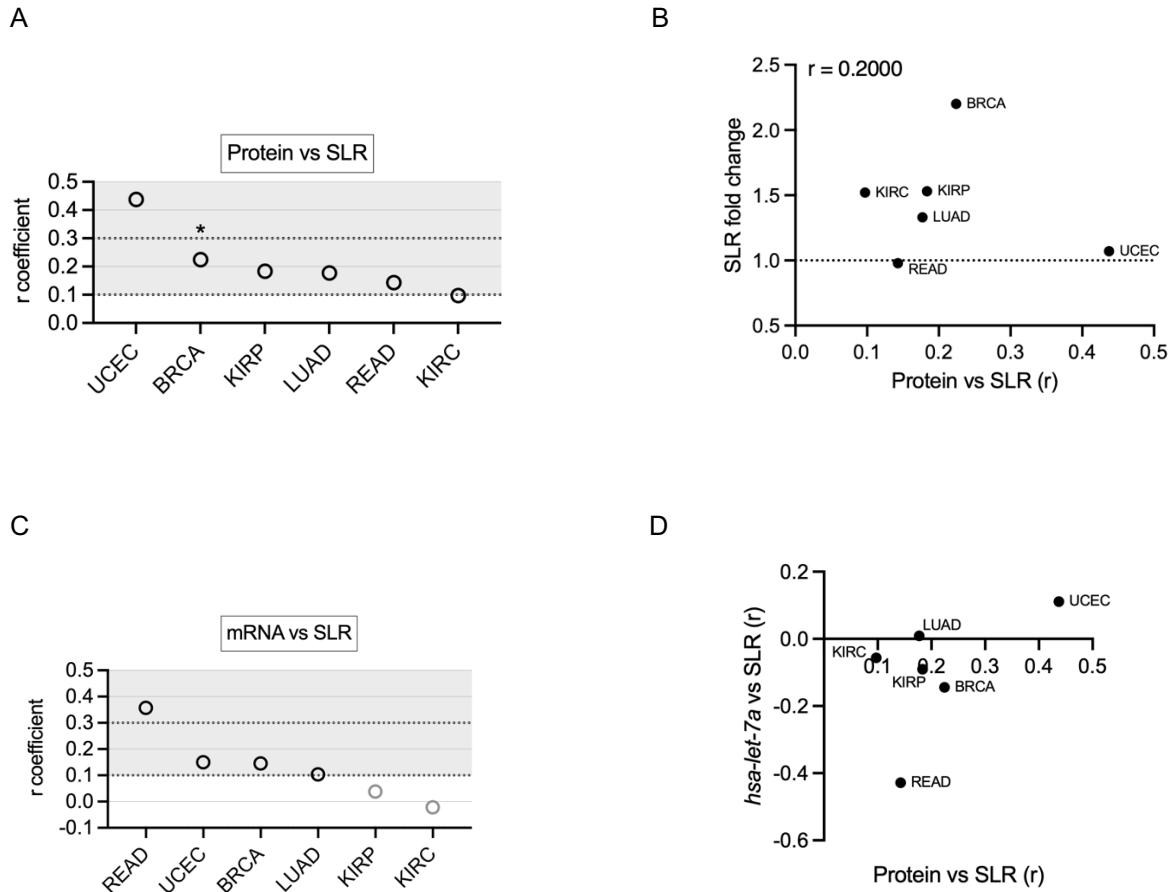
Figure 4: Interplay of APA and *hsa-let-7a* in regulating AURKA expression. (A) Distribution of r coefficients of the AURKA protein-SLR correlation across cancers. p<0.05 only for BRCA. **(B)** Scatter plot showing cancers according to the cancer/normal AURKA SLR fold change and to the value of the 'protein vs SLR' correlation coefficient r. **(C)** Distribution of r coefficients of the AURKA mRNA-SLR correlation across cancers. p>0.05 for all r coefficients. **(D)** Scatter plot showing cancers according to the r coefficient of the 'protein vs SLR' and of the '*hsa-let-7a* vs SLR' correlations.
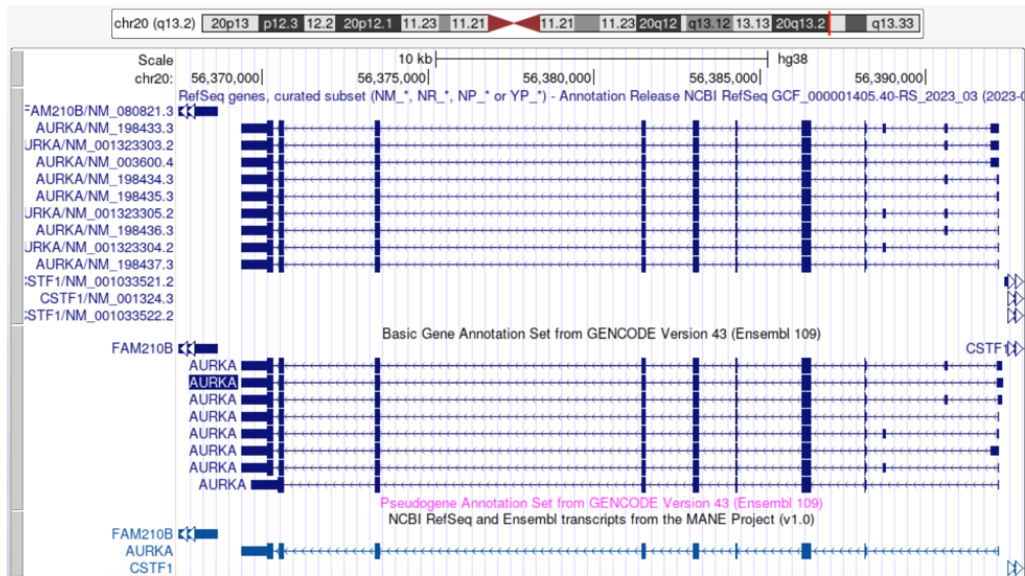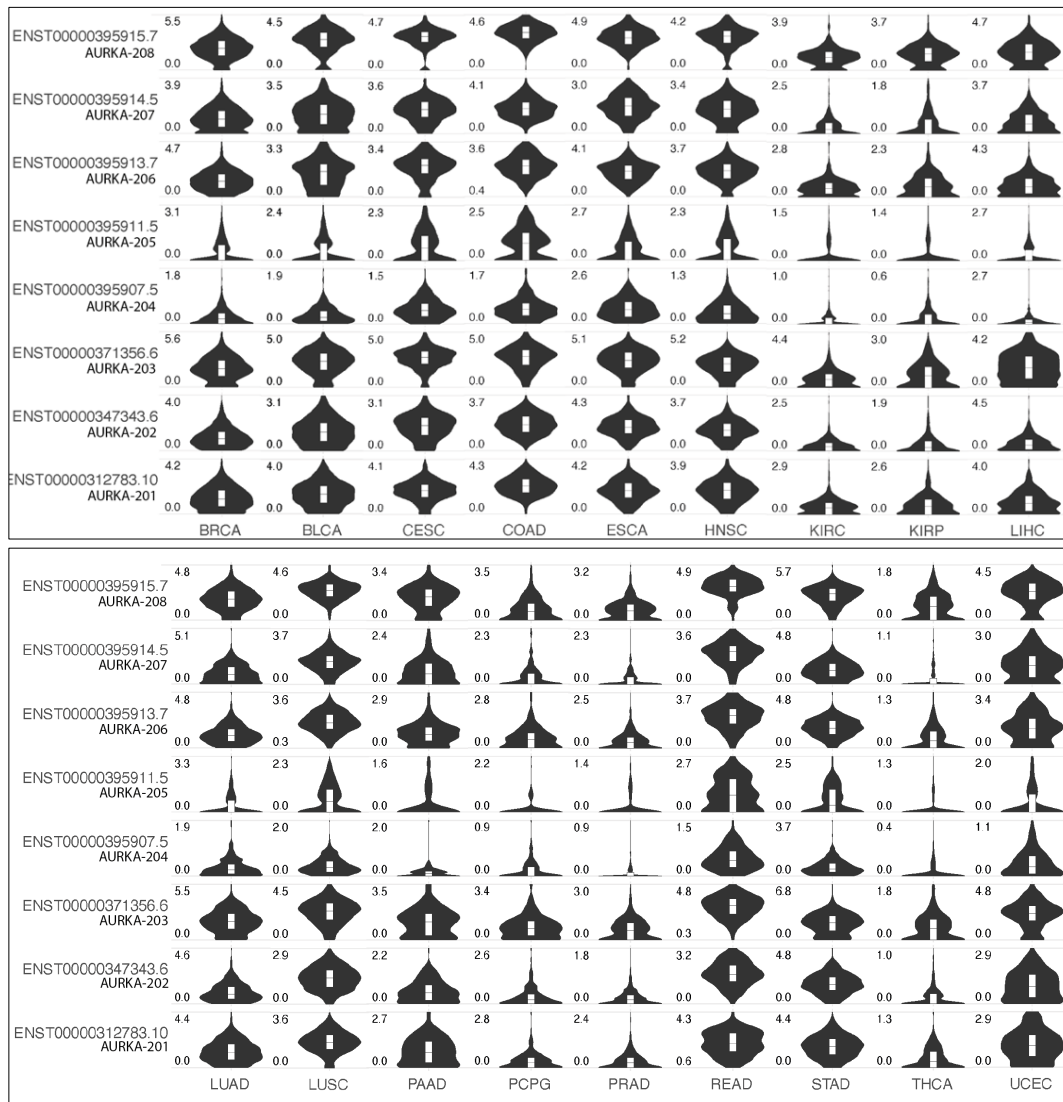
Figure 5: Heterogeneity of AURKA splicing isoforms across cancers. (A) UCSC Genome Browser view of AURKA transcript variants annotated in the NCBI (top) and Ensembl (bottom) databases. AURKA gene resides on the negative strand. The MANE isoform identical in both databases is shown in light blue at the bottom. **(B)** Violin plots showing the expression level [log2(TPM+1)] of the individual Ensembl AURKA transcripts in the TCGA cancers on the y axis. Cancer types in columns, AURKA transcripts in rows. TPM, transcript per million. Figure downloaded from the GEPIA2 platform. **(A)**, **(B)** ENST00000395907.5 (AURKA-204) transcript is displayed although it has lower annotation quality.