**Specific Aims**

Over 50% of human genes utilize alternative Transcription Start Sites (TSSs) resulting in mRNA isoforms with distinct 5′ Untranslated Regions (5′ UTRs).1 5′ UTRs harbor translational control elements that alter ribosomal recruitment to the mRNA.2,3 As such, changes in TSS usage (hereon referred to as TSS switching) are well poised to direct differences in translational output. However, our understanding of the connections between TSSs and translational output is incomplete due to insufficient mapping of TSS positions and relatively few well-characterized translational control elements in 5′ UTRs. My *long-term goal* is to uncover how TSS switching impacts translational control elements in 5′ UTRs and elucidate the mechanisms by which these elements modulate translation.

To explore TSS switching in a biological context where we know widespread changes happen both in transcription and translation, I will examine breast cancer metastasis that arises via the Epithelial to Mesenchymal Transition (EMT). EMT occurs when adhesive epithelial cells acquire mesenchymal-like morphology, motility, and invasiveness.4,5 Co-transcriptional processes including alternative splicing and polyadenylation are altered throughout EMT in breast cancer6–8 but global switching signatures are unknown. My *overall objective* is to generate high-resolution maps of genome wide TSS usage during EMT in breast cancer and assess the impact of the resulting 5′ UTRs on translation. **I hypothesize that TSS switching in breast cancer alters the inclusion of translational control elements in 5′ UTRs thus affecting the translational landscape and contributing to cancer progression and metastasis**. Investigating TSS profiles and downstream translational control will 1) reveal the 5′ UTR regulatory code and deconvolute its contribution to gene expression, 2) demonstrate the sufficiency of TSS switching in driving aggressive cancer phenotypes to uncover novel anticancer targets, and 3) allow for predictions of protein output from mRNA sequence which can have applications in gene therapy and mRNA vaccine development.

**Aim 1: Map TSS switching and identify 5′ UTR-embedded translational control elements in breast cancer.** I will induce EMT in breast cancer cell lines of varying aggressiveness. I will then map global TSS switching events in EMT intermediates of all cell lines. I expect to see different TSS profiles between the cell lines and across EMT intermediates particularly with genes that encode proliferation, migration, and invasion proteins. I will then assess the translational potential of 5′ UTR isoforms that switch across EMT intermediates. Isoforms with significant differences in translational potential will be further inspected to identify translational enhancer and repressor motifs in 5′ UTRs.

**Aim 2: Investigate the sufficiency of TSS switching in promoting aggressive breast cancer phenotypes.** I will use CRISPR-activation tools in breast cancer cell lines of low aggressiveness to artificially express 5 UTR isoforms that are preferentially translated in aggressive breast cancer cells (as previously characterized9). I will then assess the morphology, migration, and invasion properties of the CRISPR-transfected cells and compare those properties to highly aggressive breast cancer. I expect that upon transfection, cells will exhibit EMT-like morphology and higher migration and invasion properties. These results will demonstrate the sufficiency of TSS switching in reprogramming cancer cells to more aggressive phenotypes.

**Intellectual Merit and Broader Impacts.** This work will investigate translational control elements in 5′ UTRs in high throughput to develop mechanistic hypotheses about the role of TSS switching in the molecular and phenotypic reprogramming of cancer. Our results will facilitate predictions of translational output from 5′ UTR sequences. They will also enable therapeutic strategies such as selective targeting of pathogenic 5′ UTR isoforms and collective targeting of translational control motifs that predominantly alter gene expression in cancer.

**Significance**

Over 50% of human genes utilize alternative Transcription Start Sites (TSSs) resulting in mRNA isoforms with distinct 5′ UTR sequences.1 A change in TSS usage (hereon referred to as TSS switching) is well poised to direct differences in translational output (**Schematic 1**). However, the mechanisms by which TSS switching impacts gene expression are not well understood and the 5′ UTR regulatory code is largely unknown. In this proposal, I will 1) generate high resolution maps of genome wide TSS usage, 2) identify 5′ UTR isoforms with differential translational potential to pinpoint regulatory elements, and 3) examine the sufficiency of TSS switching in altering gene expression and inducing pathological phenotypes. The proposed experiments will enable advances in RNA-based therapeutics including anticancer treatments, gene therapy, and mRNA vaccines.

A screenshot of a computer

Description automatically generated

**Schematic 1.** *TSS switching can modulate gene expression via 5′ UTRs.*TSS switching produces mRNA transcripts with 5′ UTR isoforms that may harbor translational enhancers (thus increasing ribosome recruitment; left) or translational repressors (thus decreasing ribosome recruitment; right). Examples of translational control elements in 5′ UTRs are shown in the black box at the bottom.

Low-throughput experiments have already elucidated several 5′ UTR-embedded translational control elements10,11 and TSS switching in particular genes has been associated with aggressive phenotypes in breast cancer9,12. However, less is known about how such mechanisms are employed genome wide and whether they drive disease progression. As such, there remains a need to uncover the 5′ UTR regulatory code and mechanistically analyze its impact on gene expression.13 The work proposed here will address this gap by employing high-throughput methods within a metastatic cancer model to uncover disease-causing TSS switching events and identify novel avenues for therapeutics. Specifically, this work will be grounded in breast cancer models given the diversity of this disease and the ability to utilize cell lines of varying aggressiveness (from the less aggressive luminal subtypes to the more aggressive triple negative/basal cells).14 Additionally, I will be inducing Epithelial to Mesenchymal Transition (EMT). EMT occurs when adhesive epithelial cells acquire mesenchymal-like morphology, motility, and invasiveness.4,5 Co-transcriptional processes including alternative splicing and polyadenylation are altered throughout EMT in cancer6–8, but global TSS switching signatures are unknown. Mapping genome wide TSS switching in breast cancer throughout EMT will uncover novel translational control elements and elucidate their mechanistic role in promoting cancer metastasis.

Particularly, aim 1 will elucidate TSS switching in breast cancer and develop a mechanistic understanding of how this switching contributes to more aggressive cancer phenotypes whereas aim 2 will demonstrate the sufficiency of TSS switching, irrespective of other co- or post-transcriptional processes, in altering protein output and exerting observable phenotypic effects on cancer cells in terms of morphology, migration, and invasion properties. Importantly, the workflow optimized in this study can be expanded to a variety of cancer cell lines to identify conserved TSS switching that contributes to aggressive cancer progression.

Broader implications of these results include the development of therapeutics that target disease specific TSS switching or translational control elements that are primarily utilized in cancer. In addition, knowledge of these 5’ UTR regulatory elements can facilitate the prediction of protein output from the mRNA sequence. This knowledge has applications in modulating physiological levels of exogenous protein expression in gene therapy and designing high efficacy mRNA vaccines.

**Innovation**

This work is conceptually innovative because it will be the first to generate TSS maps in several breast cancer cell lines and across EMT treatment allowing for direct comparisons of mapped TSSs and thus providing a temporal aspect into which switching events drive more aggressive phenotypes. Additionally, the results of this work will deconvolute the role of TSS switching in modulating gene expression and mechanistically characterize the 5′ UTR regulatory code.

This work is technically innovative because it utilizes novel high-throughput techniques including 1) ReCappable-seq which enriches for mRNAs with intact 5′-ends and enables accurate TSS identification of all cellular mRNAs (unlike other TSS detection methods which can only capture capped mRNA transcripts) and 2) DART which both quantifies translational potential and allows for causative dissection of regulatory elements. Since its conception, DART has uncovered a minimum of 163 novel regulatory elements in each experiment.

Overall, the work outlined here will forward our understanding of the mechanisms by which TSS selection modulates genome wide gene expression in disease enabling the design of novel therapeutics.

**Approach**

**Aim 1: Map TSS switching and identify 5′ UTR-embedded translational control elements in breast cancer***.*

*Rationale and overview.* Our understanding of the connections between TSSs and translational output remains incomplete, due in part to insufficient mapping of TSS positions and relatively few well-characterized translational control elements. To address this need, in this aim, I will generate high resolution maps of genome wide TSS usage and uncover global translational control elements to identify *functional* TSS switching. Given that TSS switching has been associated with aggressive phenotypes in breast cancer9,12, I will utilize breast cancer cell lines of varying aggressiveness as a model for investigating whether TSS switching *drives* aggressive phenotypes. To investigate metastasis, I will also optimize EMT induction in those cell lines and determine EMT intermediate states. To detect TSS switching events that include or exclude translational control elements, I will identify 5′ UTR isoforms with significant differences in translational potential.

*Approach.* To examine breast cancer cell lines with varying aggressiveness, I will use epithelial luminal A breast cancer cells (MCF7) and epithelial triple negative cells (MDA-MB-468) with mammary epithelial cells (HMEC) as a healthy counterpart. Luminal A tumors are of relatively lower aggressiveness and they are more responsive to treatment while triple negative A are highly aggressive breast tumors often associated with a poor pronosis.14 Both cell lines are epithelial and neither naturally undergoes EMT. I will optimize an EMT induction treatment for all cell lines using TGF-β15 or Epidermal Growth Factor (EGF)16 as has been done previously. For each cell line, I will maintain two controls, an untreated group and a group treated with the solution used to resuspend TGF-β or EGF (for example, citric acid). Importantly, upon EMT induction, intermediate states exhibiting both epithelial and mesenchymal properties arise.17 Previous work determines that intermediate states are more aggressive than a complete transition into the mesenchymal phenotype.18 I will identify early, middle, and late EMT intermediates by monitoring gene expression, phenotypic properties, and morphology. Specifically, I will use RT-qPCR and Western Blots to measure the EMT marker vimentin and the ratio of the epithelial marker E-cadherin to the mesenchymal marker N-cadherin.15,19 Measurements of Gapdh in each sample will be used to normalize the data in both experiments. For RT-qPCR, I will use commercially available random primers to reverse transcribe RNA extracted from cells at various EMT intermediate states. I will then use primers which are specific to each marker for the amplification of the desired cDNA fragments and the commercially available EvaGreen for their quantification. For Western Blots, I will use fluorescent secondary antibodies to enable quantification of the signal. For phenotypic properties, I will identify the migratory and invasion properties of cells using Transwell assays to monitor cancer aggressiveness.20 Briefly, I will use commercially available chambers which are either coated with a polycarbonate membrane (for testing migration) or proteins of the extracellular matrix (for testing invasion). Cells seeded on top of the chamber will be in serum-fee media. Media in the well underneath the chamber will contain 10% FBS to facilitate chemotaxis. Morphology will be inspected visually using a tissue culture microscope. I anticipate that epithelial cells will have a cobble-like shape (capable of forming closely packed colonies) that changes into a spindle-like morphology followed by the appearance of individual circular cells that do not form closely packed colonies.16 When no further changes are observed in marker expression, phenotypes, and morphology, treatment will be stopped. If the expected EMT changes are not observed, I will optimize EMT induction via various steps including increasing the concentration of the inducer, testing different inducers (*i.e.,*EGF instead of TGF-β), or altering media composition (for example removing FBS from the media 24 hours before initiating treatment) as has been done before21. Once robust EMT timepoints (designated as early, middle, and late) are characterized, three biological replicates will be generated for each intermediate state in all cell lines.

TSS profiling will be conducted on RNA extracted from EMT intermediates. I will extract RNA from cells at those timepoints using a TRIzol-based extraction. To map global TSS switching events, I will use ReCappable-seq.1 In ReCappable-seq, sequencing libraries are constructed by enzymatically uncapping all transcripts (using yDcpS) then recapping (using Vaccinia Capping Enzyme) with a biotin-labelled GTP analog that can be pulled down with a streptavidin matrix. Two groups will be sequenced for each replicate of each cell line. One group will be treated with a phosphatase prior to yDcpS while the other will be untreated. The phosphatase-treated group will only contain capped mRNAs (products of RNA polymerase II). The untreated group will contain all RNA species (including products of RNA polymerase I and III). Therefore, in addition to identifying TSSs for mRNAs to reveal 5′ UTR isoforms, this technique will also enable us to detect non-coding RNAs expressed exclusively in metastatic breast cancer, expanding our avenues for identifying targetable anticancer therapeutics. EMT intermediates across all cell lines will have two technical replicates for each biological replicate. TSS peaks will be identified as Tags Per Million (TPM) as previously done.1 To identify TSS switching events, I will search for a 10% or higher change in TPM for genes that exhibit more than one TSS across EMT intermediates for each cell line.

To assess which of these TSS switching events is *functional,* I will use Direct Analysis of Ribosomal Targeting (DART).22 In DART, a user-defined 5′ UTR library containing ~12,000 sequences is *in vitro* transcribed via a T7 promoter (appended to the sequences) and then enzymatically capped using Vaccinia Capping Enzyme. Capped mRNAs are then introduced to a cellular extract of interest (here, I will be using commercially available HeLa extracts) and translation is allowed to proceed. Importantly, a translation elongation inhibitor, cycloheximide, is used to halt ribosomes upon their assembly. The extract is then loaded onto a sucrose gradient and ribosome-bound mRNAs are separated from free mRNAs. The mRNA is extracted from each fraction, reverse transcribed into cDNA, amplified, and sent for sequencing. A ribosomal recruitment score (RRS) is calculated for each isoform based on its abundance (reads per million) in the input versus 80S-bound mRNA fraction. Significance in RRSs between isoforms will be determined with a Bonferroni-corrected two-tailed t test. To infer translational enhancers (or repressors), unique sequences in longer 5′ UTRs with relatively higher (or lower) RRSs will be analyzed using STREME23**.**

*Expected results*. I expect to see differences in TSS abundance for genes with alternative TSSs across EMT intermediates in each cell line. Particularly, I anticipate that TSSs observed in middle and late EMT intermediates might play interesting roles in promoting metastasis which can be further analyzed via DART. Moreover, I expect that TSS switching which is shared among the two cancer cell lines but not the healthy mammary cell line will elucidate conserved switching in breast cancer (EMT-induced switching that drive aggressiveness in cancer cells). On the other hand, TSS usage unique to the more aggressive MDA-MB-468 cell line might elucidate EMT-related switching that push the cells to aggressiveness unattained by the luminal A cells. Similarly, TSS switching observed in the less aggressive MCF-7 cell line at later EMT intermediates, but which is present in MDA-MB-468 cells pre-EMT induction might reveal metastasis-related switches that occur in aggressive breast tumors. By using DART, I expect to reveal 5′ UTR isoforms with significant differences in translational potential allowing for the identification of translational control elements that play causal roles in altering translational output.

*Alternative Approaches/Unexpected Results*. Highly stable RNAs could mask the detection of TSS switching events. For example, if an RNA expressed pre-EMT induction is highly stable, it may be detected with high abundance post-EMT induction despite a switch in the TSS choice that reduces continued expression of this RNA. If not many significant TSS switching events are observed between the EMT intermediates across the cell lines, I can gain a temporal understanding of TSS usage by incorporating TimeLapse-seq24 into the ReCappable-seq protocol. In TimeLapse-seq, 4-thiouridine is added into cell media at a desired time point (*i.e.,* at EMT induction) and then converted into a cytosine analog post RNA extraction to create mappable T to C mutations in the sequenced cDNA. This strategy will allow us to detect any newly synthesized isoforms and identify stable mRNA expressed pre-EMT induction for accurate quantification of TSS usage at each intermediate stage.

If no major differences in RRSs are detected by DART, it would be interesting to prepare and investigate various human extracts. The initial experimental design proposes using commercially available HeLa extracts. However, to match the model in which the TSS switching events will be detected, we can prepare extracts from breast cancer cell lines including MCF-7 and MDA-MB-468. It would also be informative to design a study that tests the same library in various human cellular extracts prepared from HeLa (standard), MCF-7, and MDA-MB-468 cells. Significant differences in RRSs detected in a breast cancer cellular extract but not other cancer extracts may indicate translational control motifs that interact with proteins exclusively expressed in breast cancer.

Finally, it would be interesting to assess changes in the CDS and 3′ UTR that accompany a TSS switching event for their effect on translational potential of the mRNA. This may be inspected with polysome profiling and long-read sequencing.25 Additionally, we may incorporate long-read sequencing with ReCappable-seq.26

**Aim 2: Investigate the sufficiency of TSS switching in promoting aggressive breast cancer phenotypes.**

*Rational and overview.* To establish TSS switching events as viable targets for anticancer therapeutics, it is important to investigate their sufficiency in promoting cancer aggressiveness. To investigate whether TSS switching is sufficient to drive aggressive breast cancer phenotypes, I will leverage CRISPR activation systems to artificially express EMT-specific isoforms in uninduced MCF7 cells and monitor changes in morphology, motility, and invasion.

*Approach.* Since the shorter 5′ UTR isoforms of the plasticity inducing proteins SNAIL and NANOG were reported to be preferentially translated in stressed breast cancer cells that mimic an EMT signature9, I hypothesize that the shorter isoforms act as *drivers* of EMT. I will test this hypothesis using 5′ RACE to identify all SNAIL and NANOG isoforms in uninduced MCF7 cells and EMT intermediates. I will then utilize luciferase-based assays to establish differences in translational potential. Briefly, I will clone the 5′ UTR of each isoform for both genes upstream of a luciferase gene. I will then *in vitro* transcribe mRNAs containing the 5′ UTR of interest and luciferase coding region. Capping will be done enzymatically using Vaccinia Capping Enzyme. I will then add the mRNAs to a cellular extract to allow translation to occur. After 30 minutes at 37 °C, the addition of the luciferase substrate, luciferin, will result in the production of light that corresponds to how much luciferase was translated and which can be detected via a plate reader. I will use an Fluc mRNA (which encodes luciferase) for normalization with three technical replicates for each tested mRNA.

To test whether TSS switching is sufficient to drive more aggressive phenotypes, I will artificially express the short isoforms of SNAIL and NANOG using a puromycin-tagged VPR-dCas9 plasmid in epithelial MCF7 cells. The puromycin tag will allow me to select for cells which have up taken the VPR-dCas9 plasmid. I will then use Transwell assays to compare the migration and invasion properties of CRISPR-transfected MCF7 cells to middle and late MCF7 EMT-intermediates. It will be critical to drive predominant expression of the short isoforms with minimal changes in overall mRNA levels. I will validate increased short isoform expression and consistent total mRNA levels using 5′ RACE and Northern blots. 5′ RACE allows for the identification of all 5′ UTR isoforms present while Northern blots allow for the quantification of all mRNA isoforms. Northern blots will also eliminate concerns of alternative splicing or other changes to the coding region of the mRNA. Additionally, RT-qPCR can be utilized to quantify overall mRNA levels and ensure their consistency pre- and post-transfection. These results will demonstrate the sufficiency of TSS switching in reprogramming cancer cells to more aggressive phenotypes.

*Expected results.* I expect that the shorter isoforms of SNAIL and NANOG will exhibit higher translational potential. I also expect that transfected MCF7 cells will exhibit EMT-like morphology and higher migration and invasion properties relative to untransfected cells. Comparing those cells to MCF7 middle or late EMT intermediates will provide an understanding of the extent to which TSS switching can drive aggressive phenotypes.

*Alternative Approaches/Unexpected Results.* If no significant differences in translational potential are observed for the isoforms of the two genes, I will utilize available RNA-seq data and computational TSS identification tools (such as DeeReCT-TSS27) to identify TSS switching events in breast cancer tissues. I will then test the translational potential for isoforms of the identified genes. If CRISPR-transfected MCF7 cells do not show increased cancer phenotypes, I will investigate other candidates as identified by DeeReCT-TSS.

**Timeline And Future Directions**

I expect to identify EMT intermediates in all cell lines by the middle of the first year upon starting the proposed work. I anticipate having obtained ReCappable-seq data by the end of the first year. DART and VPR-dCas9 transfections will be performed in the second year.

|  |  |  |  |
| --- | --- | --- | --- |
| Year 1: first half | Year 1: second half | Year 2: first half | Year 2: second half |
| Aim 1: Establish EMT intermediates in at least two cell lines |  |  |  |
|  | Aim 1: Optimize ReCappable-seq libraries and perform sequencing |  |  |
|  | Aim 1: Analyze ReCappable-seq data | |  |
|  |  | Aim 1: DART |  |
|  |  | Aim 2: VPR-dCas9 | |
|  |  |  | Aim 1: EMT, ReCappable-seq, and DART for remaining cell lines of interest |
|  |  |  | Aim 1: follow up on motifs identified via DART |

Future directions will focus on 1) following up with identified translational control elements from DART (by performing low throughput experiments that mutate identified motifs to validate their effect on translation), 2) investigating additional cancer cell lines (such as colon cancer) to identify conserved cancer-related TSS switching events, and 3) utilizing CRISPR-cas9 methods to artificially repress cancer-specific isoforms to assess TSS switching events as targets for anticancer therapeutics.

**References**

(1) Yan, B.; Tzertzinis, G.; Schildkraut, I.; Ettwiller, L. Comprehensive Determination of Transcription Start Sites Derived from All RNA Polymerases Using ReCappable-Seq. *Genome Res.* **2022**, *32* (1), 162–174. https://doi.org/10.1101/gr.275784.121.

(2) Lim, Y.; Arora, S.; Schuster, S. L.; Corey, L.; Fitzgibbon, M.; Wladyka, C. L.; Wu, X.; Coleman, I. M.; Delrow, J. J.; Corey, E.; True, L. D.; Nelson, P. S.; Ha, G.; Hsieh, A. C. Multiplexed Functional Genomic Analysis of 5’ Untranslated Region Mutations across the Spectrum of Prostate Cancer. *Nat. Commun.* **2021**, *12* (1), 4217. https://doi.org/10.1038/s41467-021-24445-6.

(3) Leppek, K.; Das, R.; Barna, M. Functional 5′ UTR MRNA Structures in Eukaryotic Translation Regulation and How to Find Them. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (3), 158–174. https://doi.org/10.1038/nrm.2017.103.

(4) Chen, T.; You, Y.; Jiang, H.; Wang, Z. Z. Epithelial-Mesenchymal Transition (EMT): A Biological Process in the Development, Stem Cell Differentiation, and Tumorigenesis. *J. Cell. Physiol.* **2017**, *232* (12), 3261–3272. https://doi.org/10.1002/jcp.25797.

(5) Pradella, D.; Naro, C.; Sette, C.; Ghigna, C. EMT and Stemness: Flexible Processes Tuned by Alternative Splicing in Development and Cancer Progression. *Mol. Cancer* **2017**, *16* (1), 8. https://doi.org/10.1186/s12943-016-0579-2.

(6) Qiu, Y.; Lyu, J.; Dunlap, M.; Harvey, S. E.; Cheng, C. A Combinatorially Regulated RNA Splicing Signature Predicts Breast Cancer EMT States and Patient Survival. *RNA* **2020**, *26* (9), 1257–1267. https://doi.org/10.1261/rna.074187.119.

(7) Shapiro, I. M.; Cheng, A. W.; Flytzanis, N. C.; Balsamo, M.; Condeelis, J. S.; Oktay, M. H.; Burge, C. B.; Gertler, F. B. An EMT–Driven Alternative Splicing Program Occurs in Human Breast Cancer and Modulates Cellular Phenotype. *PLoS Genet.* **2011**, *7* (8), e1002218. https://doi.org/10.1371/journal.pgen.1002218.

(8) Neumann, D. P.; Pillman, K. A.; Dredge, B. K.; Bert, A. G.; Bracken, C. P.; Hollier, B. G.; Selth, L. A.; Beilharz, T. H.; Goodall, G. J.; Gregory, P. A. *The Landscape of Alternative Polyadenylation during EMT and Its Regulation by the RNA-Binding Protein Quaking*; preprint; Molecular Biology, 2022. https://doi.org/10.1101/2022.12.01.518773.

(9) Jewer, M.; Lee, L.; Leibovitch, M.; Zhang, G.; Liu, J.; Findlay, S. D.; Vincent, K. M.; Tandoc, K.; Dieters-Castator, D.; Quail, D. F.; Dutta, I.; Coatham, M.; Xu, Z.; Puri, A.; Guan, B.-J.; Hatzoglou, M.; Brumwell, A.; Uniacke, J.; Patsis, C.; Koromilas, A.; Schueler, J.; Siegers, G. M.; Topisirovic, I.; Postovit, L.-M. Translational Control of Breast Cancer Plasticity. *Nat. Commun.* **2020**, *11* (1), 2498. https://doi.org/10.1038/s41467-020-16352-z.

(10) Chen, J.; Tresenrider, A.; Chia, M.; McSwiggen, D. T.; Spedale, G.; Jorgensen, V.; Liao, H.; van Werven, F. J.; Ünal, E. Kinetochore Inactivation by Expression of a Repressive MRNA. *eLife* **2017**, *6*, e27417. https://doi.org/10.7554/eLife.27417.

(11) McGeachy, A. M.; Ingolia, N. T. Starting Too Soon: Upstream Reading Frames Repress Downstream Translation. *EMBO J.* **2016**, *35* (7), 699–700. https://doi.org/10.15252/embj.201693946.

(12) Nakanishi, T.; Bailey-Dell, K. J.; Hassel, B. A.; Shiozawa, K.; Sullivan, D. M.; Turner, J.; Ross, D. D. Novel 5′ Untranslated Region Variants of *BCRP* MRNA Are Differentially Expressed in Drug-Selected Cancer Cells and in Normal Human Tissues: Implications for Drug Resistance, Tissue-Specific Expression, and Alternative Promoter Usage. *Cancer Res.* **2006**, *66* (10), 5007–5011. https://doi.org/10.1158/0008-5472.CAN-05-4572.

(13) Rendleman, J.; Haizel, S.; Wu, S.; Liu, J.; Ge, X.; Zou, H.; Mohammad, M. P.; Pressler, M.; Maity, S.; Hronová, V.; Gao, Z.; Herrmannová, A.; Lei, A.; Allgoewer, K.; Sultanov, D.; Hinckley, W. E.; Szkop, K. J.; Topisirovic, I.; Larsson, O.; Hatzoglou, M.; Valášek, L. S.; Vogel, C. *Regulatory Start-Stop Elements in 5’ Untranslated Regions Pervasively Modulate Translation*; preprint; Systems Biology, 2021. https://doi.org/10.1101/2021.07.26.453809.

(14) Dai, X.; Cheng, H.; Bai, Z.; Li, J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *J. Cancer* **2017**, *8* (16), 3131–3141. https://doi.org/10.7150/jca.18457.

(15) Liu, Q.-Q.; Huo, H.-Y.; Ao, S.; Liu, T.; Yang, L.; Fei, Z.-Y.; Zhang, Z.-Q.; Ding, L.; Cui, Q.-H.; Lin, J.; Yu, M.; Xiong, W. TGF‑β1‑induced Epithelial‑mesenchymal Transition Increases Fatty Acid Oxidation and OXPHOS Activity via the P‑AMPK Pathway in Breast Cancer Cells. *Oncol. Rep.* **2020**, *44* (3), 1206–1215. https://doi.org/10.3892/or.2020.7661.

(16) Devaraj, V.; Bose, B. Morphological State Transition Dynamics in EGF-Induced Epithelial to Mesenchymal Transition. *J. Clin. Med.* **2019**, *8* (7), 911. https://doi.org/10.3390/jcm8070911.

(17) Goetz, H.; Melendez-Alvarez, J. R.; Chen, L.; Tian, X.-J. A Plausible Accelerating Function of Intermediate States in Cancer Metastasis. *PLOS Comput. Biol.* **2020**, *16* (3), e1007682. https://doi.org/10.1371/journal.pcbi.1007682.

(18) Kisoda, S.; Mouri, Y.; Kitamura, N.; Yamamoto, T.; Miyoshi, K.; Kudo, Y. The Role of Partial-EMT in the Progression of Head and Neck Squamous Cell Carcinoma. *J. Oral Biosci.* **2022**, *64* (2), 176–182. https://doi.org/10.1016/j.job.2022.02.004.

(19) Tian, M.; Schiemann, W. P. TGF-β Stimulation of EMT Programs Elicits Non-Genomic ER-α Activity and Anti-Estrogen Resistance in Breast Cancer Cells. *J. Cancer Metastasis Treat.* **2017**, *3*, 150–160. https://doi.org/10.20517/2394-4722.2017.38.

(20) Guan, Y.-X.; Zhang, M.; Chen, X.-Z.; Zhang, Q.; Liu, S.-Z.; Zhang, Y.-L. Lnc RNA SNHG20 Participated in Proliferation, Invasion, and Migration of Breast Cancer Cells via MiR-495. *J. Cell. Biochem.* **2018**, *119* (10), 7971–7981. https://doi.org/10.1002/jcb.26588.

(21) Yin, Y.; Feng, L.; Wang, L.; Ding, L. The Role of Curcumae Rhizoma-Sparganii Rhizoma Medicated Serum in Epithelial-Mesenchymal Transition in the Triple Negative Breast Cancer. *Biomed. Pharmacother.* **2018**, *99*, 340–345. https://doi.org/10.1016/j.biopha.2017.11.139.

(22) Niederer, R. O.; Rojas-Duran, M. F.; Zinshteyn, B.; Gilbert, W. V. Direct Analysis of Ribosome Targeting Illuminates Thousand-Fold Regulation of Translation Initiation. *Cell Syst.* **2022**, *13* (3), 256-264.e3. https://doi.org/10.1016/j.cels.2021.12.002.

(23) Bailey, T. L. STREME: Accurate and Versatile Sequence Motif Discovery. *Bioinformatics* **2021**, *37* (18), 2834–2840. https://doi.org/10.1093/bioinformatics/btab203.

(24) Schofield, J. A.; Duffy, E. E.; Kiefer, L.; Sullivan, M. C.; Simon, M. D. TimeLapse-Seq: Adding a Temporal Dimension to RNA Sequencing through Nucleoside Recoding. *Nat. Methods* **2018**, *15* (3), 221–225. https://doi.org/10.1038/nmeth.4582.

(25) Nguyen, L. A. C.; Inada, T.; Galipon, J. Nanopore Direct RNA Sequencing of Monosome- and Polysome-Bound RNA. In *Nanopore Sequencing*; Arakawa, K., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2023; Vol. 2632, pp 281–297. https://doi.org/10.1007/978-1-0716-2996-3\_20.

(26) Ugolini, C.; Mulroney, L.; Leger, A.; Castelli, M.; Criscuolo, E.; Williamson, M. K.; Davidson, A. D.; Almuqrin, A.; Giambruno, R.; Jain, M.; Frigè, G.; Olsen, H.; Tzertzinis, G.; Schildkraut, I.; Wulf, M. G.; Corrêa, I. R.; Ettwiller, L.; Clementi, N.; Clementi, M.; Mancini, N.; Birney, E.; Akeson, M.; Nicassio, F.; Matthews, D. A.; Leonardi, T. Nanopore ReCappable Sequencing Maps SARS-CoV-2 5′ Capping Sites and Provides New Insights into the Structure of SgRNAs. *Nucleic Acids Res.* **2022**, *50* (6), 3475–3489. https://doi.org/10.1093/nar/gkac144.

(27) Zhou, J.; Zhang, B.; Li, H.; Zhou, L.; Li, Z.; Long, Y.; Han, W.; Wang, M.; Cui, H.; Li, J.; Chen, W.; Gao, X. Annotating TSSs in Multiple Cell Types Based on DNA Sequence and RNA-Seq Data via DeeReCT-TSS. *Genomics Proteomics Bioinformatics* **2022**, *20* (5), 959–973. https://doi.org/10.1016/j.gpb.2022.11.010.