With profound appreciation, I extend my heartfelt thanks to Her Royal Highness Princess Maha Chakri Sirindhorn for her gracious support in granting Thai youth the chance to participate in the SIPGA internship activities in the year 2024.

Chonthicha Khotchakasorn

Applied Biological Sciences, Chulabhorn Graduate Institute

With the support of

Agency for Science, Technology and Research (A*STAR)

National Science and Technology Development Agency (NSTDA)

Princess Sirindhorn IT Foundation (PSIT)

Preface

This report has been written based on the experience gained from participating in the Singapore International Pre-Graduate Award (SIPGA) program in 2024 in Singapore, from August 26, 2024 to February 25, 2025. It includes general information about the SIPGA Program and the research conducted during the program.

Participating in the SIPGA Program has been one of the most valuable periods of my life, and I sincerely hope that my experiences will be of some benefit to the readers.

Chonthicha Khotchakasorn

Thai representative in the SIPGA Program for the year 2024

Acknowledgements

I wish to express my deepest gratitude for the boundless royal grace of Her Royal Highness Princess Maha Chakri Sirindhorn, who most graciously granted me the esteemed opportunity to participate in the 2024 Singapore International Pre-Graduate Award (SIPGA) program at the Agency for Science, Technology and Research (A*STAR), Singapore.

I am also profoundly thankful to the National Science and Technology Development Agency (NSTDA) for their exceptional support and guidance in preparing me for this opportunity, as well as for their continued assistance throughout the duration of the program.

My sincere appreciation is extended to Professor Chen Kok Hao, my research advisor, along with his research team, for their invaluable mentorship, academic guidance, and unwavering support throughout the course of my research.

Lastly, I am truly grateful to my family, friends, and all mentors who provided steadfast encouragement and support in preparing me for this meaningful and transformative experience.

Chonthicha Khotchakasorn

Thai representative in the SIPGA Program for the year 2024

Content

SIPGA

The Singapore International Pre-Graduate Award (SIPGA) offers international students a unique opportunity to immerse themselves in Singapore's dynamic scientific landscape and collaborate with world-renowned researchers at A*STAR's research institutes.

This program is open to undergraduate and master's students in fields such as Computing and Information Science, Biomedical Science, Physical Science, and Engineering and Technology, providing short-term research attachments ranging from two to six months.

A*STAR (Agency for Science, Technology and Research) is Singapore's leading government agency dedicated to advancing science and innovation. It plays a central role in driving research that supports economic growth, national development, and societal well-being.

Founded in 1991 (originally as NSTB), it was renamed ASTAR in 2002 to reflect its mission of making Singapore a star in science and tech. ASTAR oversees and funds cutting-edge R&D, bridging academia, research institutes, and industries.

A*STAR's research activities are housed in two major research hubs:

1. Biopolis (Life Sciences Hub)

Biopolis focuses on biomedical sciences and healthcare innovation. Key research institutes here include:

- **Genome Institute of Singapore (GIS)** – Genomics, DNA sequencing, precision medicine.
- **Singapore Immunology Network (SIgN)** – Immunology and infection.
- **Institute of Molecular and Cell Biology (IMCB)** – Molecular biology, stem cells, cancer research.
- **Institute of Bioengineering and Bioimaging (IBB)** – Bioengineering, medical devices, imaging tech.
- **Bioinformatics Institute (BII)** – Computational biology, AI in healthcare.

- **Experimental Drug Development Centre (EDDC)** – Drug discovery and translational medicine.

2. Fusionopolis (Physical Sciences & Engineering Hub)

Fusionopolis focuses on engineering, advanced materials, and digital tech. Key institutes include:

- **Institute of Materials Research and Engineering (IMRE)** – Smart materials, polymers, nanotech.
- **Institute for Infocomm Research (I^2R)** – AI, communications, cybersecurity, IoT.
- **Institute of High Performance Computing (IHPC)** – Computational science, digital twins, simulations.
- **Advanced Remanufacturing and Technology Centre (ARTC)** – Industry 4.0, robotics, advanced manufacturing.
- **Singapore Institute of Manufacturing Technology (SIMTech)** – Automation, additive manufacturing, industrial innovation.

Research Report

1. Introduction

Intron/exon project:

Nascent transcripts dynamics in human cancer cell lines MCF-7, A549 and normal lung tissue fibroblast cell line IMR-90 were studied as part of the first project.

This project first began with the testing of a mini library for introns and exons of four genes, including ESR1, MYC, GREB1 and TFF1 gene, using 361 probes to collect information on spatial distribution and copy numbers of RNAs.

Nuclear receptors in cells are activated when stimulated by steroid hormones such as estrogen, which control coregulators for gene transcription. The overall goal of this project was to model nascent transcript activity for specific genes of interest, particularly ones with changes in expressions under the influence of estrogenic effects (1).

L37 library designed for studying genes relevant to pancreatic cancer from Norbert Ha was tested on the cell lines for optimizing library amplification protocol, encoding probe concentrations and readout probe concentrations for better imaging and data quality.

A549 and IMR-90 cell lines were then tested for probes designed in the L75 human library by Xinrui Zhou. The library consists of 5176 probes for 85 genes at interest, 12 introns and 10 exons as control, as well as 1 gene with high intron burst and high FPKM.

Split-FISH encoding probes were then tested as a method to improve hybridization efficiency. The encoding probes were digested into two halves with USER enzyme, replacing traditional restriction enzyme, to create a pair of probes and prevent hanging nucleotides after the enzymatic digestion. By applying this method, the encoding probes would then be hybridized and bridged for readout to bind on and display signals only if both sides of the pair of encoding probes successfully hybridize to the RNA, enabling higher binding specificity.

2. Results and Work in Progress

Intron/exon project:

Intron/exon project: The four-gene mini library demonstrated distinguishable results between introns and exons of genes specified under basal conditions. Imaging results (fig.1) showed aligned results compared to the study conducted by (1), with clear intron signals in MYC gene control and exon signals across most panels. The probe design pipeline and experimental procedures were verified through trials of MERFISH and suggested intron probes sequence hybridization could be identified for analysis.

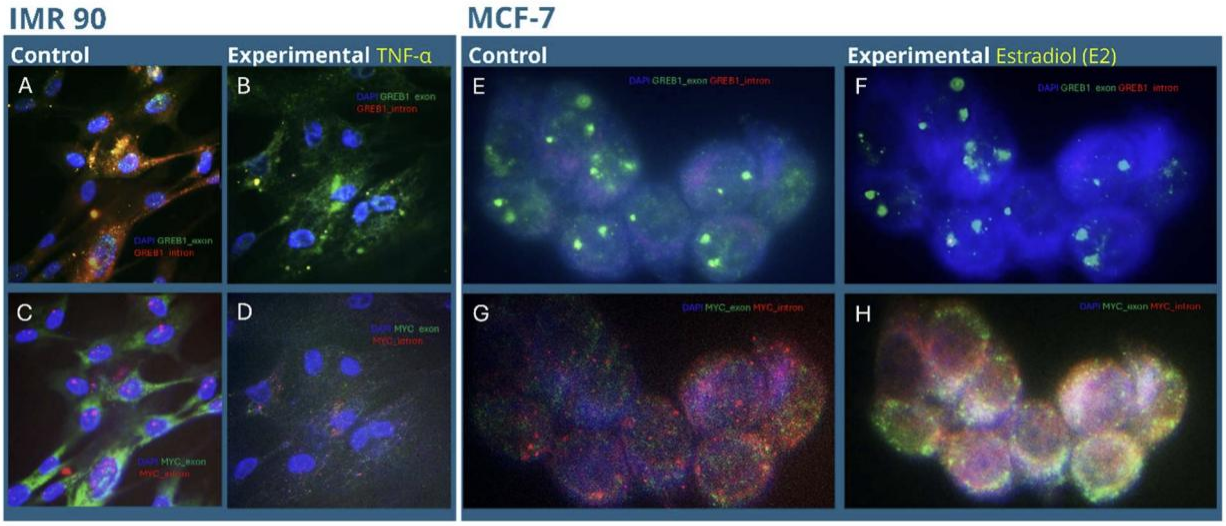

Figure 1. Representative images of IMR-90 and MCF-7 cells hybridized with probes detecting GREB1 and MYC gene. Cells in (B, D, F, H) are stimulated by estrogen for 48 hours. For (A-H), blue represents the nucleus stained by DAPI, green represents exons, mature RNA, and red represents intron, nascent RNA. (A) GREB-1 gene exons and introns on IMR-90 sample with no stimulation. (B) GREB-1 gene exons and introns on IMR-90 sample with stimulation using TNF-α. (C) MYC gene exons and introns on IMR-90 sample with no stimulation. (D) MYC gene exons and introns on IMR-90 sample with stimulation using TNF-α. (E) GREB-1 gene exons and introns on MCF-7 sample with no stimulation. (F) GREB-1 gene exons and introns on MCF-7 sample with stimulation using estradiol. (G) MYC gene exons and introns on MCF-7 sample with no stimulation. (H) MYC gene exons and introns on MCF-7 sample with stimulation using estradiol.

Imaging results by hybridizing the pancreatic cancer library L37 (fig. 2) displayed high background and signs of non-specific binding signals. Split-FISH analysis of images for hybridizing the L37 library on cell lines repeatedly reported a low correlation between counts of RNA signals detected and Fragments Per Kilobase of transcript per Million mapped reads (FPKM).

RNA spot counts for some genes were less than 10, while many were not found. The counts for blank probes which are designed to be signal-free displayed higher counts than expected, suggesting that images taken or the hybridization performed exhibited strong background noise and low hybridization efficiency for many genes.

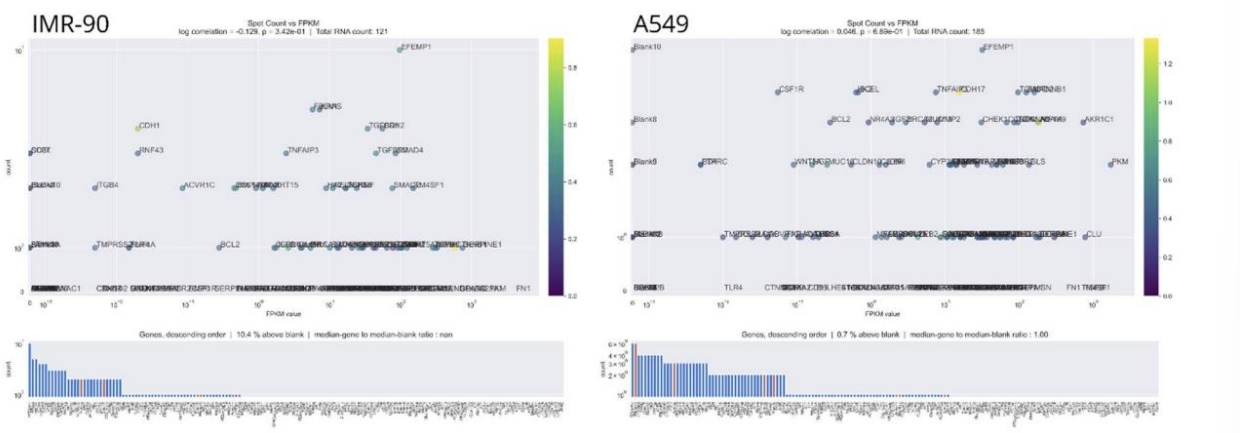

Figure 2. RNA Spot counts vs FPKM for IMR-90 and A549 cells hybridized with probes of L37 pancreatic cancer library. IMR-90-L37 hybridization indicated a negative correlation of -0.129. A549-L37 hybridization indicated a correlation of 0.046.

Similarly, imaging results from the hybridization of newly designed human library L75 on IMR-90 cells (fig. 3a) reflected low RNA spot counts result and low correlation of RNA counts with FPKM values. Output images recorded from MERFISH rounds displayed strong background signals despite adjusting to optimal image settings and details (fig. 3b and 3c). Correlation increased when intron and exon control probes were excluded from the analysis. Blank probes also exhibited unexpectedly high signals, which indicates strong background signals. The overall RNA spot counts however were higher than the L37 pancreatic cancer library, with certain genes expressing more than 100 spots count.

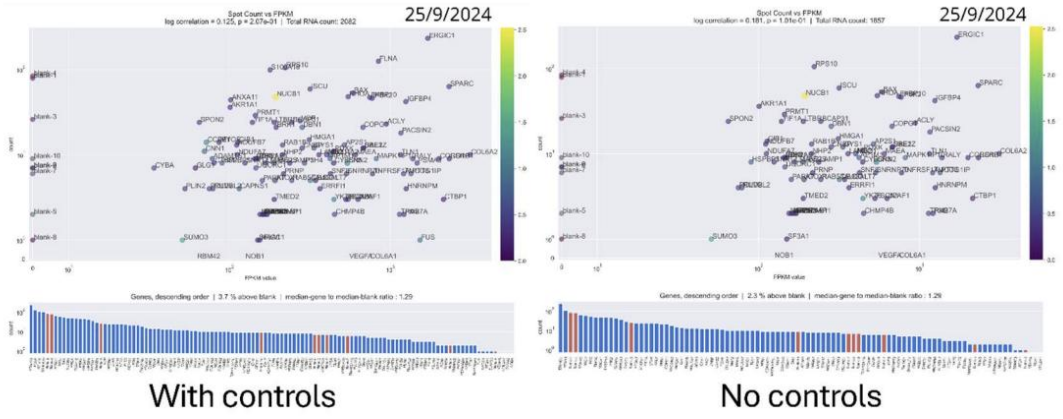

Figure 3a. Spot Counts vs FPKM with IMR-90 cells hybridized with L75 human library. The hybridization was conducted using 10% wash buffer, with encoding probe hybridization incubated at 37°C. Data plotted on the graph with control probes indicated a correlation of 0.125. Data plotted on the graph without control probes indicated a correlation of 0.181.

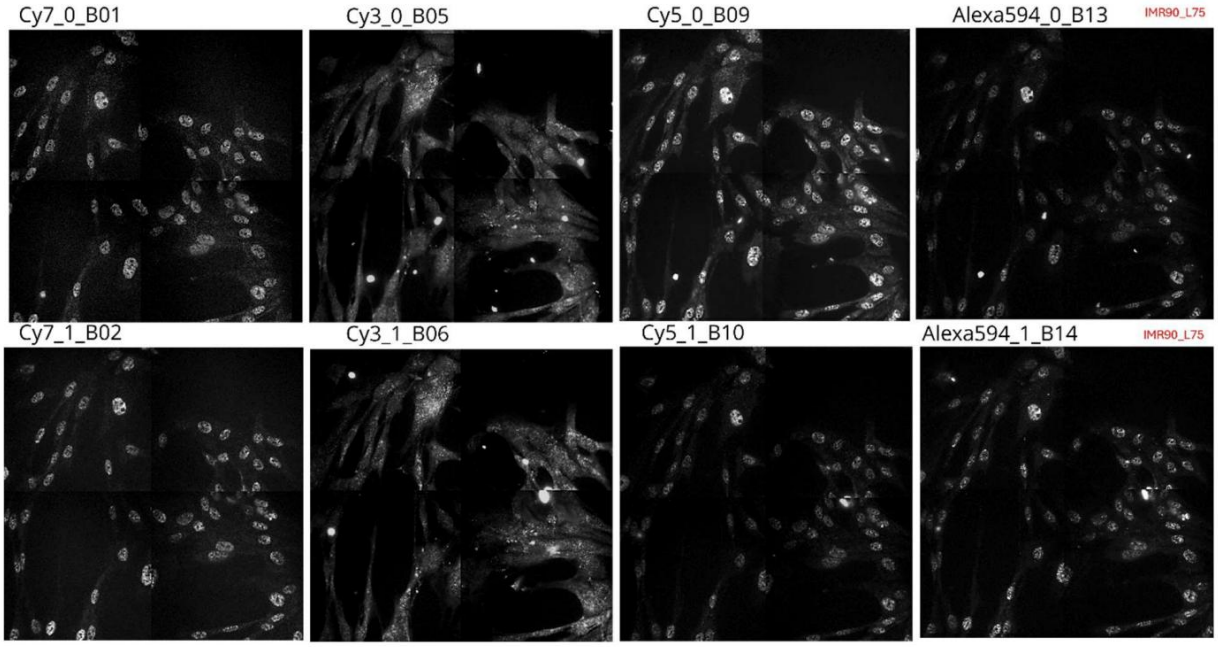

Figure 3b. Output images of L75 human library hybridization on IMR-90 in corresponding readouts (B01, B05, B09, B13, B02, B06, B10, B14).

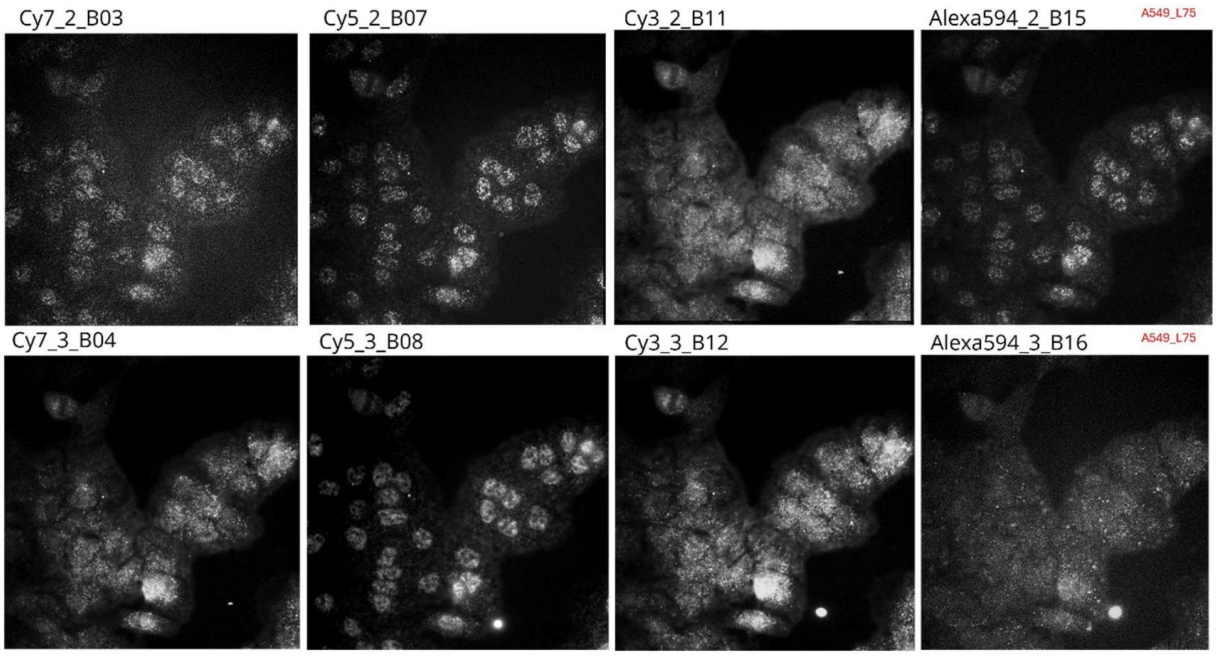

Fig 3c. Output images of L75 human library hybridization on A549 in corresponding readouts (B03, B07, B11, B15, B04, B08, B12, B16).

After further optimizing hybridization experimental conditions and analysis thresholds, absolute counts of RNA as well as the correlation between RNA spot counts and FPKM of L75 human library increased. Nevertheless, blank probes signals were still at unexpectedly high counts, implying the presence of unusual background signals. The RNA count increased from 2082 to 4239 after using a higher percentage formamide wash buffer and higher incubation 9 temperatures for encoding probe hybridization. The correlation between RNA spot counts and FPKM improved to 0.267 with higher magnitude, small spot thresholds, low cuts, and high cuts (Fig. 4).

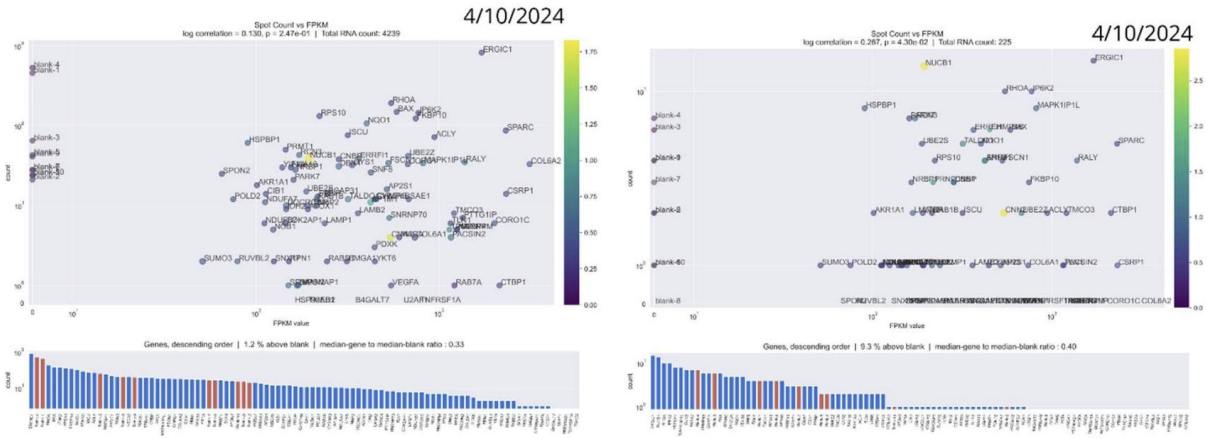

Figure 4. Spot Counts vs FPKM with IMR-90 cells hybridized with L75 human library, using 30% wash buffer, with encoding probe hybridization incubated at 47°C. Data plotted on left graph with a correlation of 0.130 and absolute count of RNA at 4239 was analyzed by split-FISH code with magnitude threshold: 0.4, small spot threshold: 0.8, low cut: 200, high cut: 800 . Data plotted on left graph with a correlation of 0.267 and absolute count of RNA at 225 was analyzed by split-FISH code with magnitude threshold: 0.8, small spot threshold: 1.2, low cut: 400, high cut: 1000.

Split-FISH with split probe method developed by (2) was successfully modified by using USER enzyme to create two split probes required (fig. 5) to hybridize to the RNA and for bridging probe to bind on and connect with readout probes for imaging signals, instead of conventional restriction enzymes that leaves hanging sequences after digestion.

Polyacrylamide gel electrophoresis image indicated the success of digesting probes into two by using the USER enzyme. The digested probes were found at 35 bp, matching the expected length of nucleotide with the control. Some undigested probes after in-vitro transcription were found at 55 bp while some primers leftovers were found at 20 bp.

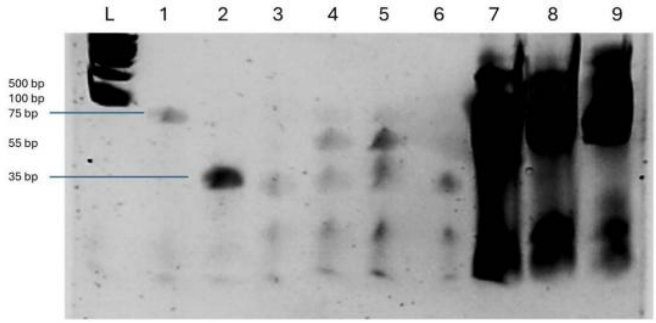

Figure 5. Polyacrylamide gel electrophoresis image after USER enzyme digestion of encoding probes. 100 bp ladder was loaded on the leftmost lane. Lane 1 was the benchmarking control encoding probe that was not digested with a size of 75 bp. Lane 2 was the benchmarking control encoding probe of 35nt after the USER digest. Lanes 3-6 represented bands at 55, 35, and 20 bp. Lane 7 showed band results from products of in-vitro transcription. Lanes 8-9 showed band results of product from reverse transcription.

3. Discussions and Future Directions

Intron/exon project:

Experimental protocol for hybridizing introns may require further optimization to improve intron signal strength and visibility, as well as reduce background noise via experimental, designing and analytical adjustments.

Applications of split-FISH probes would theoretically reduce the emission of signals from the background by increasing the binding specificity of encoding probes, by only allowing readout probes with fluorophores to bind to bridging probes that bind to pairs of encoding probes for the gene.

Furthermore, the TCEP disulfide method could be deployed to improve the washing of readout probes after each round of hybridization. The disulfide bond for linking fluorophore with the readout probe sequence would be removed using TCEP (tris(2-carboxyethyl)phosphine) as a reducing agent. Removing fluorophores with TCEP could lower background brightness and increase the visibility of true signals from probes on RNAs for the next round of hybridization. Protocols for cell preparation, library preparation, probe design, or imaging procedures such as time and temperature for incubation of hybridization as well as wash buffer concentrations could be further optimized for better hybridization conditions for each cell type in order to obtain better imaging results for downstream analysis of the spatial transcriptomics data.

4. References

1. Stossi F, Dandekar RD, Mancini MG, Gu G, Fuqua SAW, Nardone A, et al. Estrogen-induced transcription at individual alleles is independent of receptor level and active conformation but can be modulated by coactivators activity. Nucleic Acids Research. 2020;48(4):1800-10.

2. Goh JJL, Chou N, Seow WY, Ha N, Cheng CPP, Chang YC, et al. Highly specific multiplexed RNA imaging in tissues with split-FISH. Nature Methods. 2020;17(7):689-93.