Visium Spatial Gene Expression for FFPE: RNA-templated Ligation Probe Design

Introduction

The sequestration and degradation of RNA that occurs in formaldehyde fixed paraffin embedded (FFPE) tissues presents a major challenge to transcriptomic analyses. The Visium Spatial Gene Expression for FFPE assay overcomes traditional FFPE limitations by avoiding reliance on the polyA tail of transcripts or the need for extended reverse transcription and template switching. Instead, Visium Spatial Gene Expression for FFPE uses a highly specific RNA-templated ligation probe based strategy to assay protein-coding RNA expression across the whole transcriptome and map the expression of each transcript to its spatial location in the tissue. This Technical Note describes highly specific 10x Genomics probes for maximum capture efficiency from FFPE samples, for use in the Visium Spatial Gene Expression for FFPE assay.

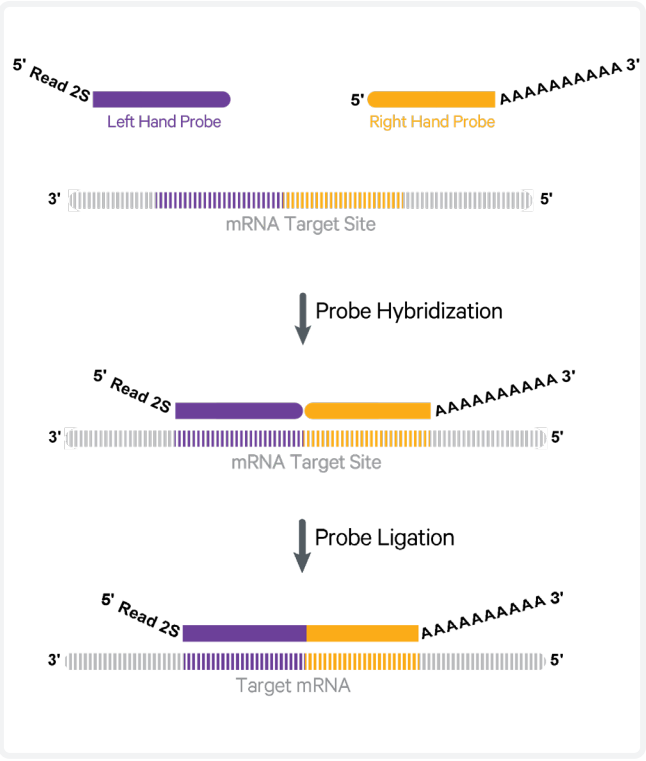

Figure 1. Probe hybridization and ligation

Assay Overview

To capture mRNAs of interest, human or mouse whole transcriptome probe panels contain pairs of specific probes for each targeted gene. The left hand probe introduces a partial Read 2S sequence, while the right hand probe introduces a synthetic polyA tail. These probe pairs are hybridized to the target mRNA and ligated together (Figure 1). The target RNA is then digested, allowing the single stranded ligation product to be released and captured by the Visium Spatial Gene Expression slide via the poly(dT) capture sequence. For more information, consult the Visium Spatial Gene Expression Reagent Kits for FFPE User Guide (CG000407).

Probe Design and Specificity

Probe Design

An initial list of genes and transcripts were selected as the GENCODE basic transcripts for each protein coding gene in APPRIS for mouse or human reference transcriptomes (GRCh38: GENCODE v32/Ensemble 98, mm10: GENCODE vM23/Ensembl 98) (Figure 2). Probes were not designed against mitochondrial, ribosomal, and a large fraction of variable/joining TCR immunoglobulin genes.

Candidate RNA-templated ligation probes were generated by tiling transcripts. All candidate probes were evaluated and scored based on multiple factors such as complete transcript coverage, position within annotated transcripts, GC content, presence of homopolymers, probe specificity, overlap with repetitive or low complexity sequences, and overlap with SNPs. These factors were used to select one probe pair per gene for inclusion in the whole transcriptome probe set (Figure 2). If the probe set did not capture all annotated transcripts of a gene, additional probes were designed when possible to improve transcript coverage. These probe panels are not designed to detect isoforms or SNPs. A full list of the probes and their targets can be found in Space Ranger 1.3.0, which can be found on the downloads page of the 10x Genomics Support website.

Probe Specificity

All candidate probe sequences generated during probe design were aligned to the reference transcriptome using BLAST to determine all possible sites for off-target hybridization.

To enable accurate estimation of the off-target activity at each possible hybridization site, a set of experiments were designed using probes with intentionally introduced nucleotide substitutions disrupting probe hybridization and ligation efficiencies. A set of probes targeting well-expressed genes with no BLAST off-target alignments identified was selected for the experiments.

In the first experiment, nucleotides were substituted within the hybridizing sequence (but outside the ligation site) of one of the probe pairs, reducing the hybridization strength of the probe to its targeted transcript(s). By comparing the UMI counts of unmodified probes to probes with intentionally introduced mismatches, a probe hybridization strength threshold could be identified. Probes with hybridization strength under this threshold are not expected to hybridize and produce off-target UMI counts (Figure 3A).

In the second experiment, mismatches were randomly introduced at various positions near the ligation junction, thereby disrupting the ligation efficiency of the probe pair. By comparing UMI

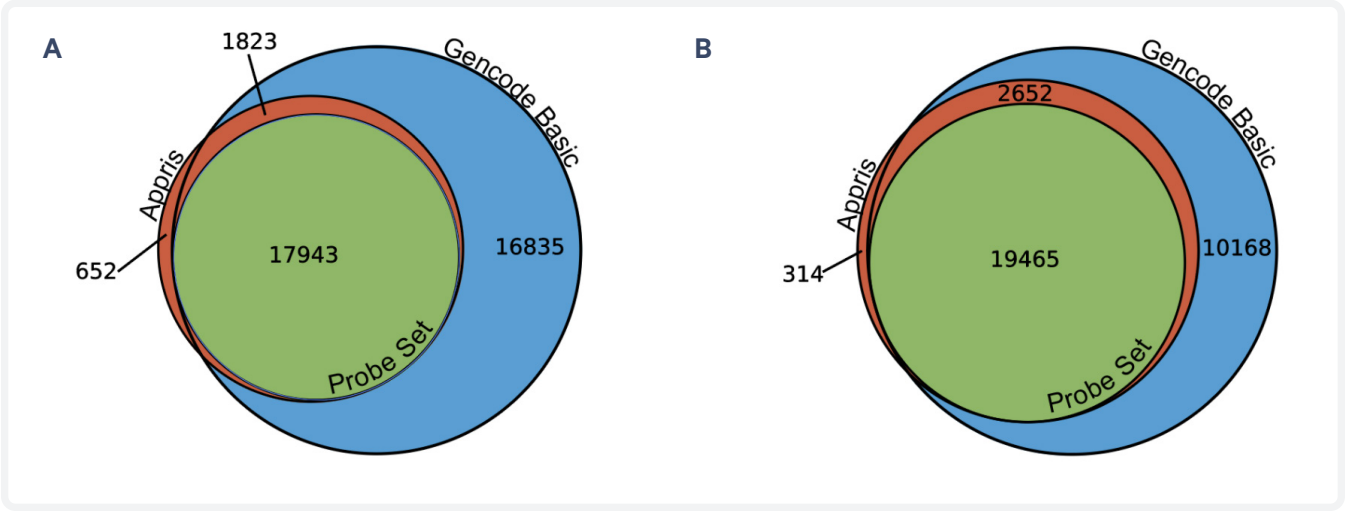

Figure 2. Targeted Genes in the RNA-templated Ligation probe set for human (A) and mouse (B). Probes were designed by tiling GENCODE basic transcripts of protein coding genes in APPRIS. The full list of probes and targets are available in the probe set csv downloadable from the 10x support website.

counts to unmodified probes, mismatched positions that hinder probe ligation could be identified (Figure 3B). By assaying each of the identified off-target alignments using the hybridization strength threshold and the ligase disturbing mismatch pattern, a filtered set of active off-target alignment could be identified and used to select probes with little expected off-target activity (Figure 2).

For some groups of highly homologous genes, all candidate probes have sufficient cross-gene sequence similarity to produce off-target ligation products. Probes for these genes are included in the probe sets and the potential non-specific targets for these probes are specified in the respective offtarget csv file. Note, any gene that has at least one probe with predicted off-target activity will be excluded from filtered outputs of the Space Ranger data analysis pipeline. Instructions for including these probes in the analysis can be found in the Probe Filter section of the Space Ranger Algorithms Overview on the 10x Genomics Support website.

Probe Alignment and Counting in Space Ranger 1.3.0

In Space Ranger 1.3.0, probe ligation events are counted using a probe alignment algorithm. The probe alignment algorithm assigns probe IDs to each probe half identified within each sequencing read. Only reads with both probe halves mapped to one of the probe sequences outlined in the reference probe set csv file are used for UMI counting and downstream analysis. For genes targeted by more than one probe (e.g. to improve transcript coverage), all individual probe counts are summed and the probe count sum is presented in the spaceranger count output. For more information on probe counting and filtering see the Probe Alignment for FFPE and Probe Filtering for FFPE sections of the Gene Expressions Algorithms Overview"page on the 10x Genomics Support website.

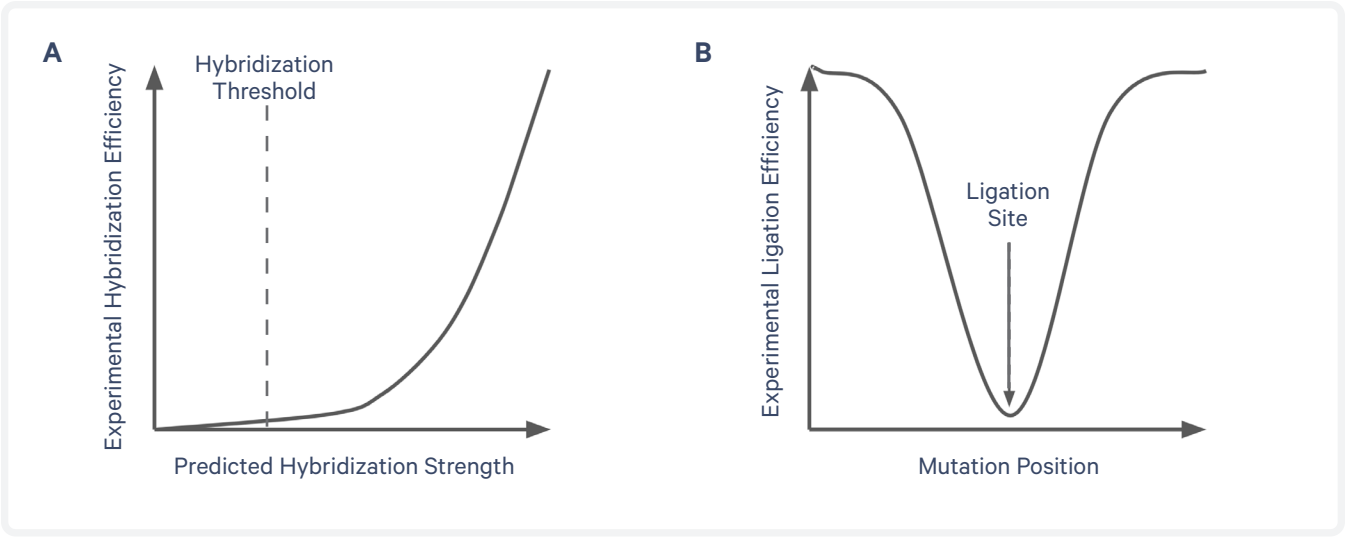

Figure 3. Estimation of off-target activity via intentional mismatch experiments. Hybridization strength predicted in silico accurately predicts experimental hybridization efficiency (A). Probe sequences that have no off-target matches are selected due to their predicted efficient hybridization. A mismatch adjacent to the ligation junction substantially reduces ligation efficiency (B).

Conclusions

Through careful selection and validation, the RNA-templated ligation approach was selected as a robust and highly specific method for capturing RNA of interest from FFPE samples.

Combined with immunofluorescence, Visium Spatial Gene Expression for FFPE enables simultaneous visualization of protein and gene expression.

References

1. Visium Spatial Gene Expression Reagents Kits for FFPE User Guide (CG000407)

Document Revision Summary

Document Number	CG000437
Title	Visium Spatial Gene Expression for FFPE: RNA-templated Ligation Probe Design
Revision	Rev A
Revision Date	October 2021

Contact:
support@10xgenomics.com
10x Genomics
6230 Stoneridge Mall Road
Pleasanton, CA 94588 USA