

GEGCURSOP025: Spatial Transcriptomics Curation Guide

Metadata

SOP identifier

GEGCURSOP025

Current version number

v1

SOP name

Spatial Transcriptomics Curation Guide

Author(s)

Nancy George

Contents

1. Overview
2. Implementation
3. Procedure
4. Interpretation
5. Process Details

Overview

This SOP is to guide users about the additional fields required for curating spatial transcriptomics datasets in Annotare and ArrayExpress.

A good introduction to Single Cell spatial transcriptomic datasets can be found here: <https://blog.bioturing.com/2020/11/09/explore-10x-visium-spatial-transcriptomics-data-at-ease-with-bioturing-browser/>

10x Visium documentation here: <https://support.10xgenomics.com/spatial-gene-expression> and library construction guide here: [CG000239_VisiumSpatialGeneExpression_UserGuide_RevE_.pdf](#)

Links to spatial barcode slide description here: <https://kb.10xgenomics.com/hc/en-us/articles/360041426992-Where-can-I-find-the-Space-Ranger-barcode-whitelist-and-their-coordinates-on-the-slide->

Links to spatial transcriptomics datasets in ArrayExpress: <https://www.ebi.ac.uk/arrayexpress/search.html?query=%22spatial+transcriptomics%22+>

Original MAGE-TAB specification can be found for download via Slack files here: http://fged.org/site_media/pdf/MAGE-TABv1.1_2011_07_28.pdf

To do:

Include more information about other spatial transcriptomics information e.g.

Nanostring Digital Spatial Profiling (DSP):

https://www.nanostring.com/wp-content/uploads/2020/12/BR_MK0981_GeoMx_Brochure_r19_FINAL_Single_WEB.pdf

2. Implementation

IDF fields to include

Comment[AdditionalFile:spatial barcode]

Experiment Type: spatial transcriptomics OR spatial transcriptomics by high-throughput sequencing

SDRF fields to include

1 sample = 1 capture area

Source Name Characteristic.... **Material Type** **Image File** **Extract Name**

Comment[library construction]: 10x Visium

Comment[spatial barcode read]: read1

Comment[spatial barcode size]: 16

Comment[spatial barcode offset]: 0

3.Procedure

4. Interpretation

SDRF

Samples

For **10x Visium** each **Capture Area = 1 sample** as represented in the Source Node. This is because 1 library comes from 1 capture area

Sample Barcode

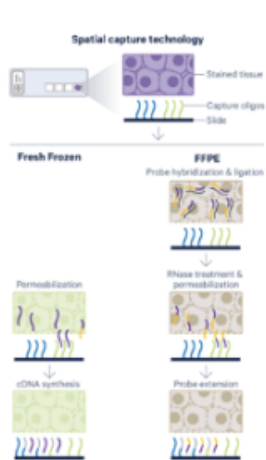
Each **Capture Area** has a **Sample Barcode**

Issues arise when multiple images are introduced onto a single capture area (e.g. multiple tissue sections per image). Unfortunately these then need to be described as mixed samples that describe the capture area.

Spatial Barcode

Example of 10x Visium Gene Expression library including the **Spatial Barcode** that replaces cell barcode in single cell SDRF. See below images that describe the process:

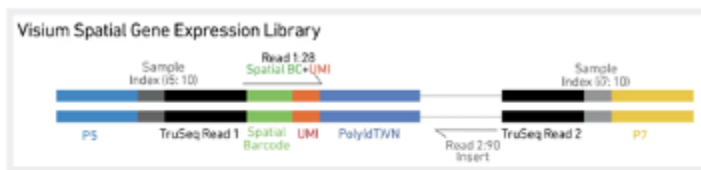
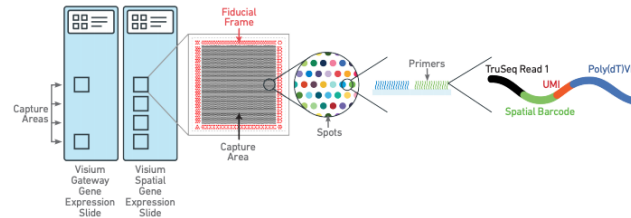
1. images are affixed to a Visium Slide - up to **4 capture areas** are present per slide
2. images are permeabilised and RNA extracted and captured by **capture oligos** affixed to the slide - these are the **spatial barcodes**. **Spatial barcodes** are specific to each **capture area**
 - a. **spatial barcodes** correspond to an **area on the capture area** and not a cell called a '**spot**'
3. Library consists of:
 - Comment[spatial barcode read]: read1
 - Comment[spatial barcode size]: 16bp
 - Comment[spatial barcode offset]: 0bp



Visium Spatial and Gateway Gene Expression Slides

The Visium Spatial Gene Expression Slide includes 4 Capture Areas (6.5 x 6.5 mm), each defined by a fiducial frame (fiducial frame + Capture Area is 8 x 8 mm). The Visium Gateway Gene Expression Slide includes 2 Capture Areas that are identical in size. The Capture Area has ~5,000 gene expression spots, each spot with primers that include:

- Illumina TruSeq Read 1 (partial read 1 sequencing primer)
- 16 nt Spatial Barcode (all primers in a specific spot share the same Spatial Barcode)
- 12 nt unique molecular identifier (UMI)
- 30 nt poly(dT) sequence (captures poly-adenylated mRNA for cDNA synthesis).



	Read 1	i7 Index	i5 Index	Read 2
Purpose	Spatial Barcode + UMI	Sample Index	Sample Index	Transcript
Length	(16 + 12) = 28	10	10	90

- Paired-end, dual-index libraries
- Recommend a minimum of 50k read pairs per spot covered with tissue

Image File

Within the original MAGE-TAB specification there are '**Nodes**' these act as directionality for the sample ID.

As you can see here in the table there is an **Image File** node. This originally referred to the associated microarray hybridisation image but can be applied to the **Visium** (or any spatial transcriptomic) image file.

The Image File comment acts as a Node in its own right so should be placed in the **SDRF** after **Source Name** and before **Extract Name** to reflect it's location in the experiment process. Image File is acquired before nuclear RNA extraction.

e.g

Source Name Characteristic.... **Material Type** **Image File** **Extract Name**

2.3 Sample and Data Relationship Format (SDRF)

2.3.1 SDRF Specification

Node/Edge	Associated attributes	Cardinality
Source Name	Characteristics, Provider, Material Type, Description, Comment	0..1
Sample Name	Characteristics, Material Type, Description, Comment	0..*
Extract Name	Characteristics, Material Type, Description, Comment	0..*
Labeled Extract Name	Characteristics, Material Type, Description, Label, Comment	0..1
Assay Name	Technology Type, Array Design File / REF, Comment	0..1
Scan Name	Comment	0..*
Normalization Name	Comment	0..*
Array Data File	Comment	0..*
Derived Array Data File	Comment	0..*
Array Data Matrix File	Comment	0..*
Derived Array Data Matrix File	Comment	0..*
Image File	Comment	0..*
Protocol REF	Term Source REF, Parameter, Performer, Date, Comment	0..*

IDF

Additional Files

10x Visium also contains spatial barcode location files. These map the spatial barcodes to the grid on the sample slide allowing for spatial reconstruction of the image once the transcriptomics analysis is done.

Depending on the Visium version, these need to be attached as the `Comment[AdditionalFile:spatial barcode]` in the **IDF**

These files are **versioned** by 10x and can be found here:

[visium-v2.txt](#) [visium-v2_coordinates.txt](#) [visium-v1.txt](#) [visium-v1_coordinates.txt](#)