Rを用いた空間解析の基礎
と
空間データベースの紹介

京都大学医生物学研究所 准教授

VANDENBON Alexis

京都大学

本日の目標・想定する受講者

空間トランスクリプトーム解析の基礎
&
DeepSpaceDBを含む公共データベースの
特徴や利用方法

- 空間トランスクリプトーム・データの読み取り方や解析手法の概略を知りたい方
- 公共データベースの所在や特徴を知りたいといった方
- 公共の空間トランスクリプトームデータを比較・解析してみたい方
- Rの経験がある方が望ましいです

自己紹介

バンデンボン　アレクシス
VANDENBON Alexis　京都大学 医生物学研究所 准教授

略歴（1981年9月生まれ・ベルギー出身）

2005年	来日
2006年～2009年	東京大学 新領域創成科学研究科 博士課程 （指導教員：中井 謙太 教授）
2009年～2014年	大阪大学 免疫学フロンティア研究センター(iFReC) 特任研究員
2014年～2017年	大阪大学 免疫学フロンティア研究センター特任助教
2017年～2022年	京都大学 医生物学研究所 講師
2022年～現在	京都大学 医生物学研究所 准教授

Image from https://wallpaperaccess.com/

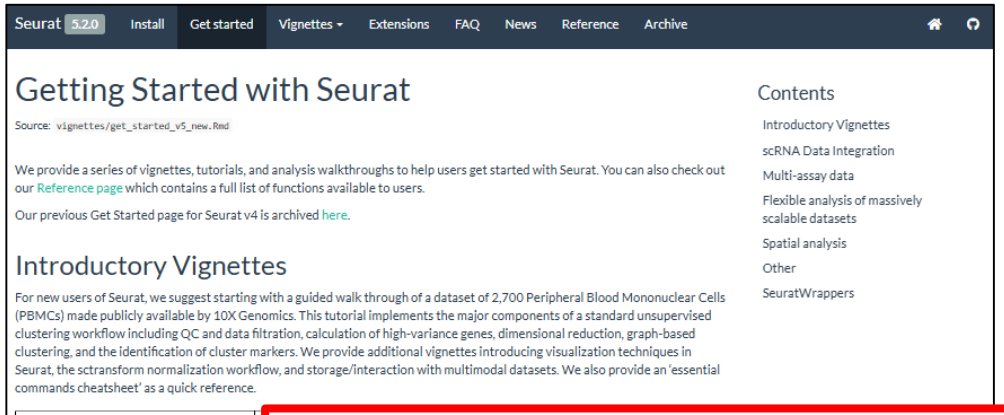
概要

- Rを使った空間トランスクリプトミクスデータ解析
  - Visium example
  - Other platforms
- 既存の空間トランスクリプトミクスデータベースの紹介
- DeepSpaceDBの紹介
  - Visium interface
  - Xenium interface
- まとめ

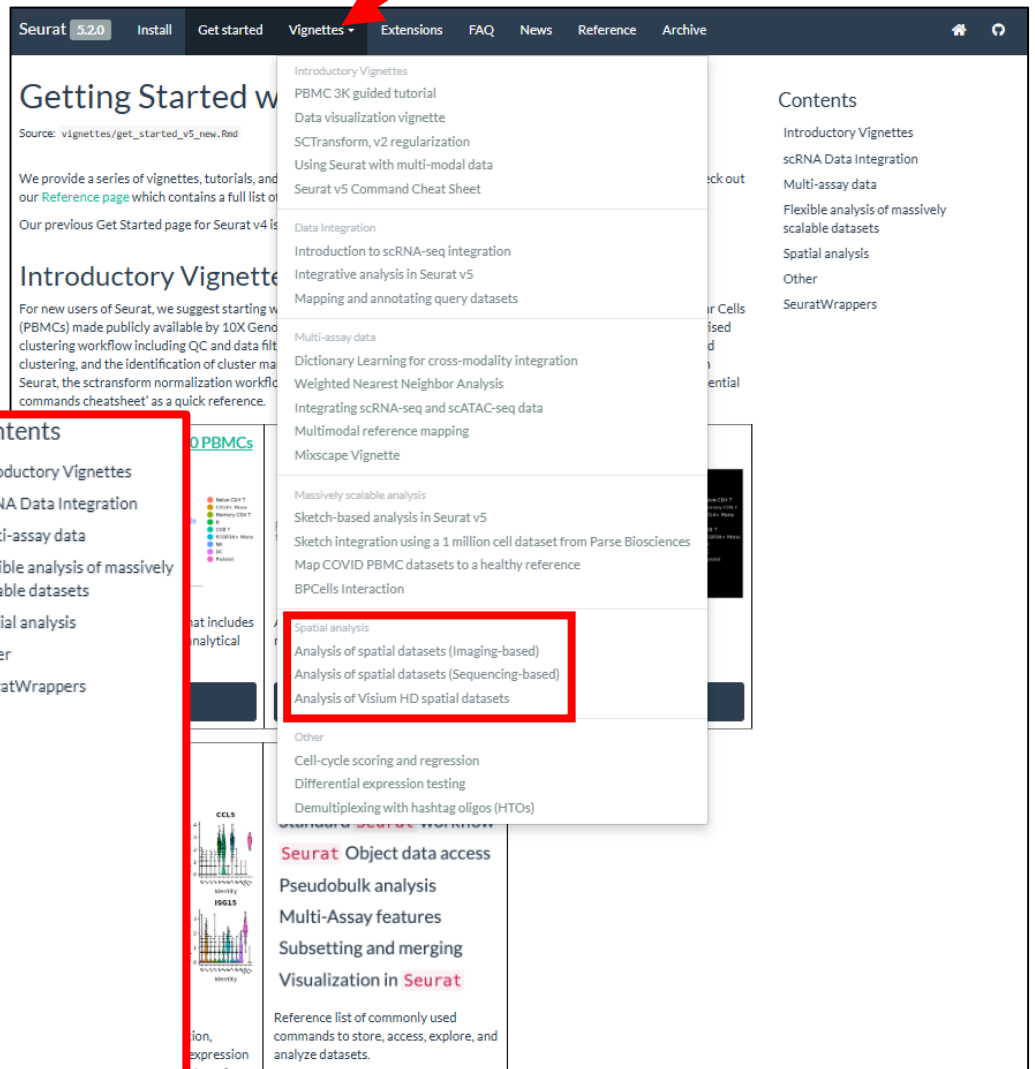
The R Seurat package

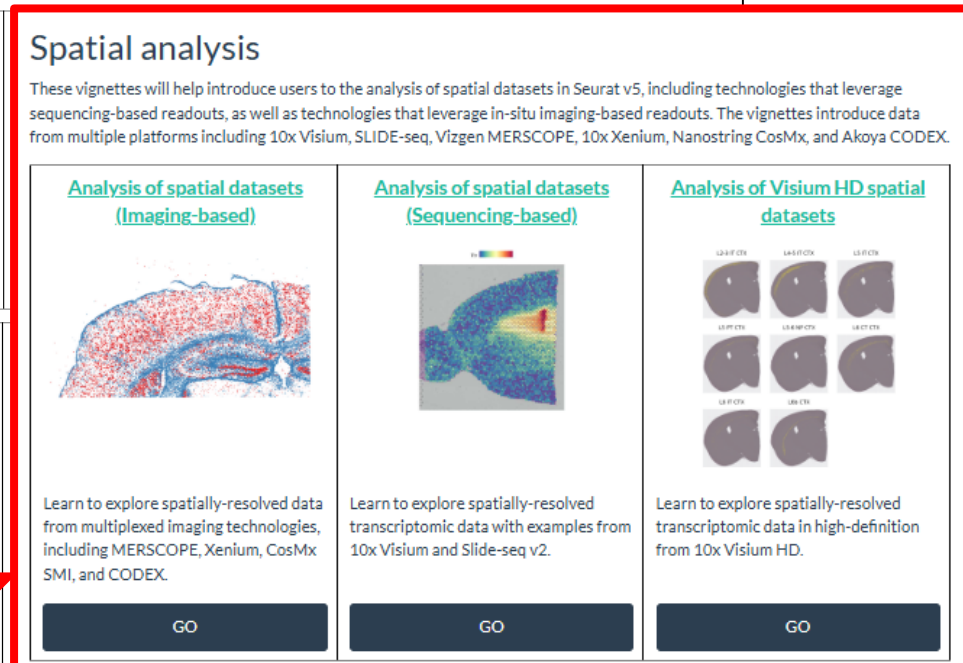
https://satijalab.org/seurat/

https://satijalab.org/seurat/articles/get_started_v5_new

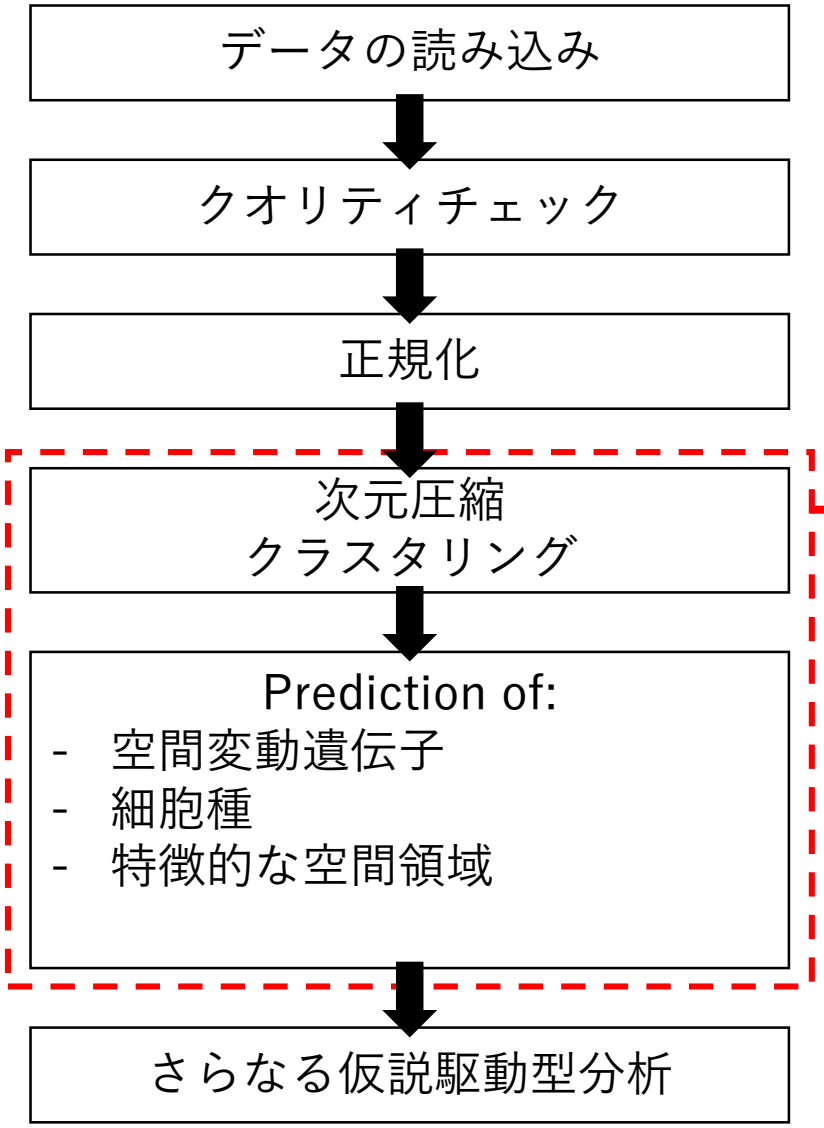


or





ワークフローの概要



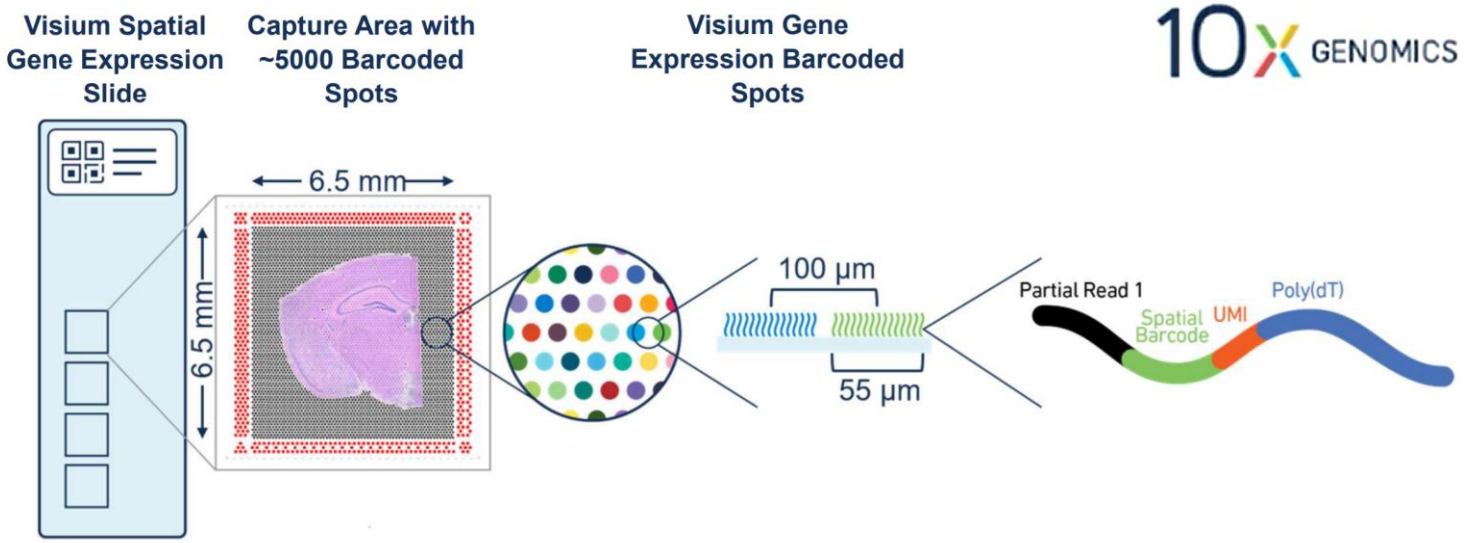
https://satijalab.org/seurat/articles/spatial_vignette

https://github.com/alexisvdb/ajacs_20251218

データ探索

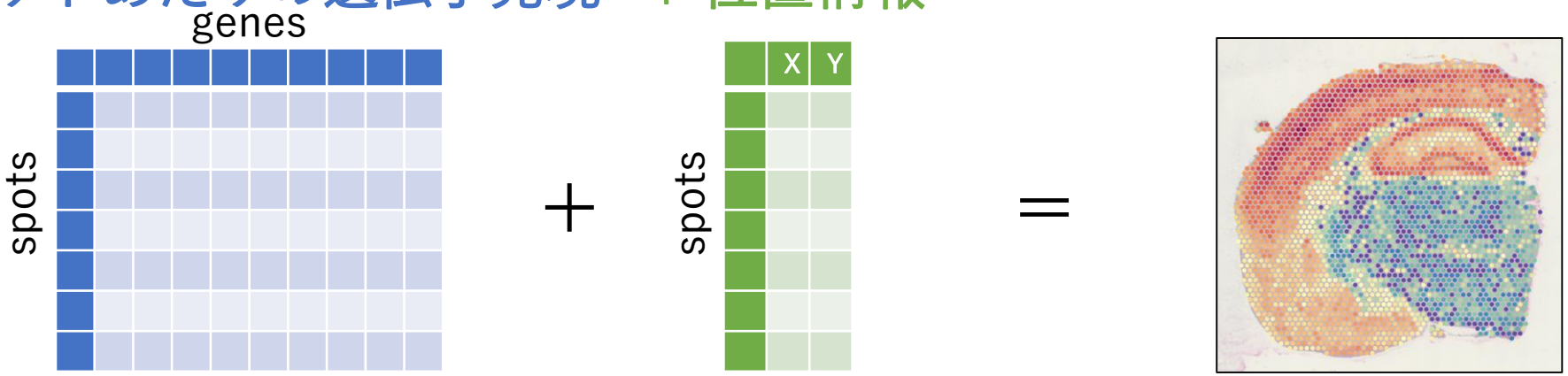
- 高次元データの解析で特に重要となる
- データの品質評価や異常値の特定
- データ中の主要な構造を見つける
- 仮説の生成と検証
- 単純に正解をだすような解法は無い

10X Genomics Visium platform



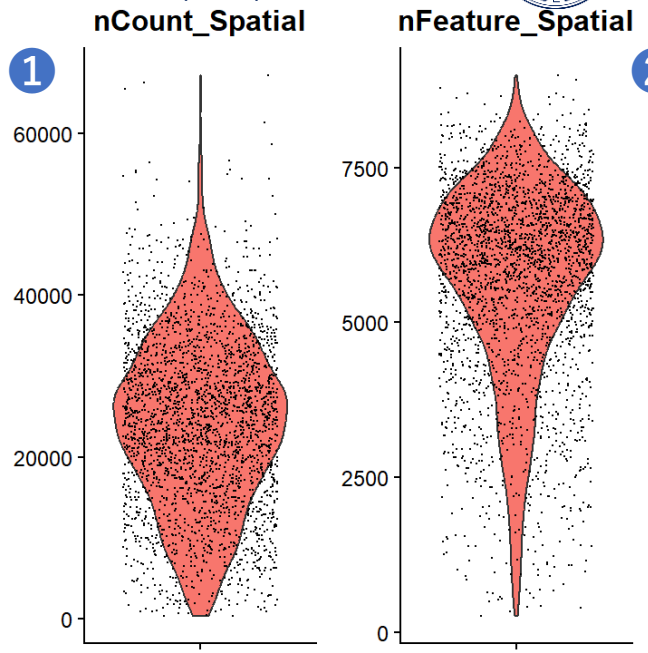
- 各スポットには約10個の細胞が含まれている（“single cell”ではない）
- Final data:

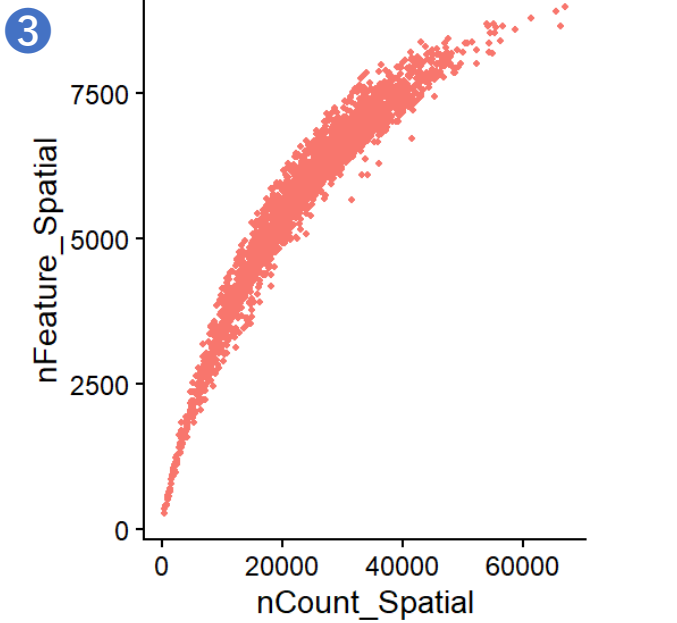
スポットあたりの遺伝子発現　＋位置情報



Seurat tutorial – クオリティチェック (1)


```
#### load packages ####
library(Seurat) # version 5.3.0
library(SeuratData)

#### Install and load example Visium dataset ####
InstallData("stxBrain")
brain <- LoadData("stxBrain", type = "anterior1")

#### Quality check ####
VlnPlot(brain, features = "nCount_Spatial")
VlnPlot(brain, features = "nFeature_Spatial")
FeatureScatter(brain, feature1 = "nCount_Spatial",
                      feature2 = "nFeature_Spatial")
```





クオリティの指標として使われるもの

1. スポット当たりのリード数
2. スポット当たりの遺伝子検出数

1細胞データとの共通点：

1 2. スポット間にかなりのばらつきがある
3. リード数と遺伝子検出数に相関がある

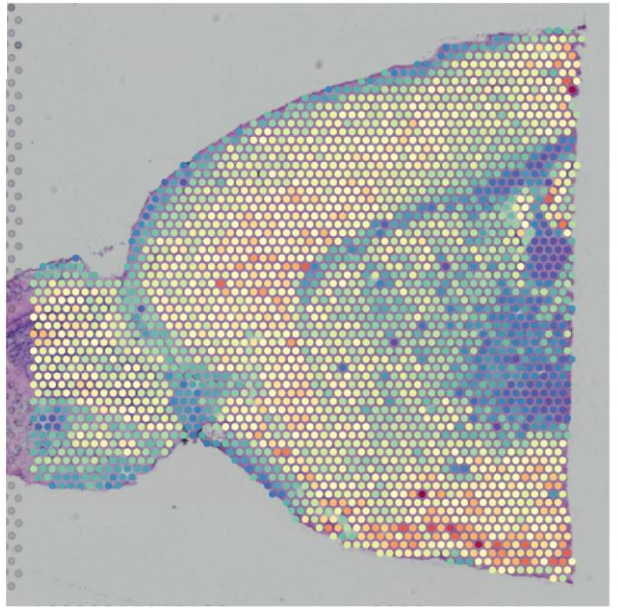
1細胞データとの相違点：「Doublet」という概念がない

Seurat tutorial – クオリティチェック (2)

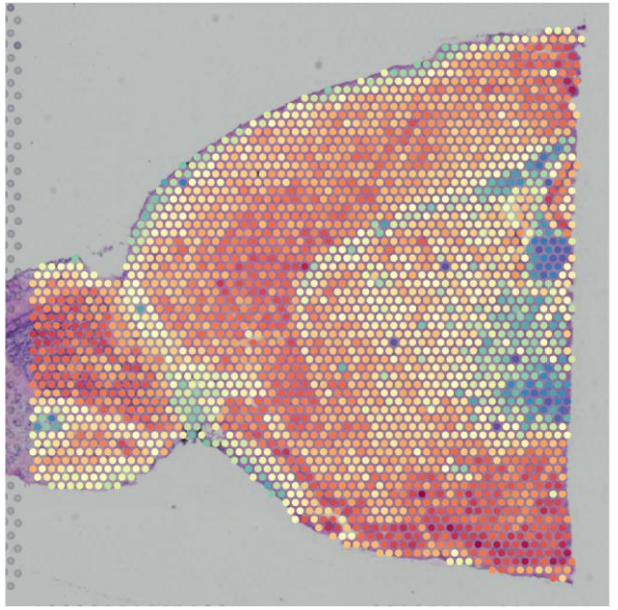

```
#### Quality within slices ####
SpatialFeaturePlot(brain, features = "nCount_Spatial") ← 1
SpatialFeaturePlot(brain, features = "nFeature_Spatial") ← 2
```

- クオリティの指標として使われるもの
  1. スポット当たりのリード数
  2. スポット当たりの遺伝子検出数
- シングルセル解析以上に注意深い解釈が必要となる
  - 組織内における細胞密度のばらつき
  - 細胞種の違い
  - 組織の透過性の違い
- 脳の組織構造を反映した空間的パターンを見ることが出来る

1



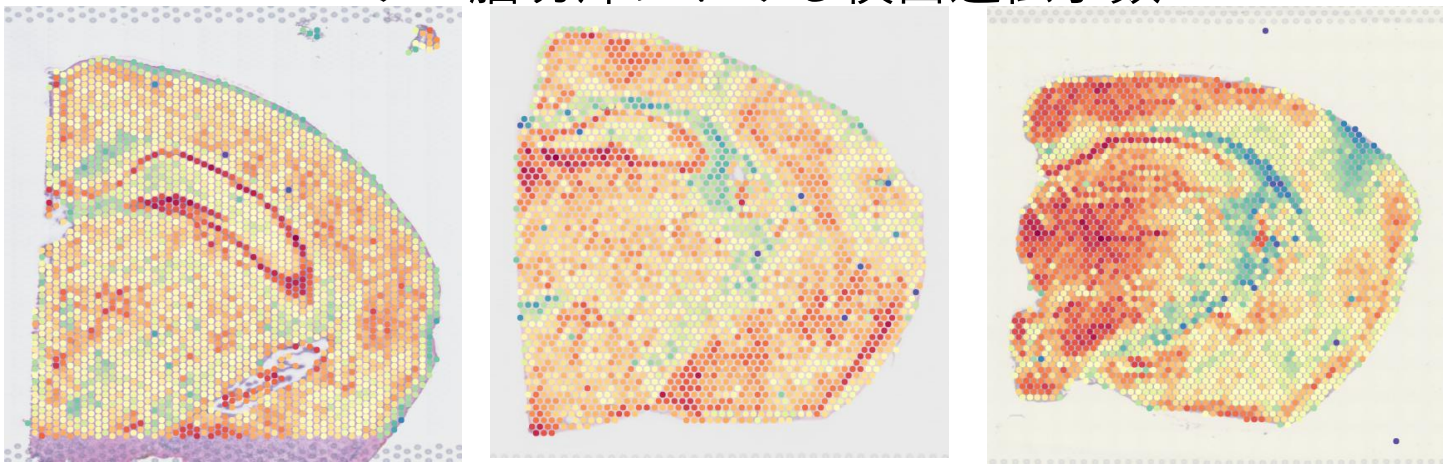
2



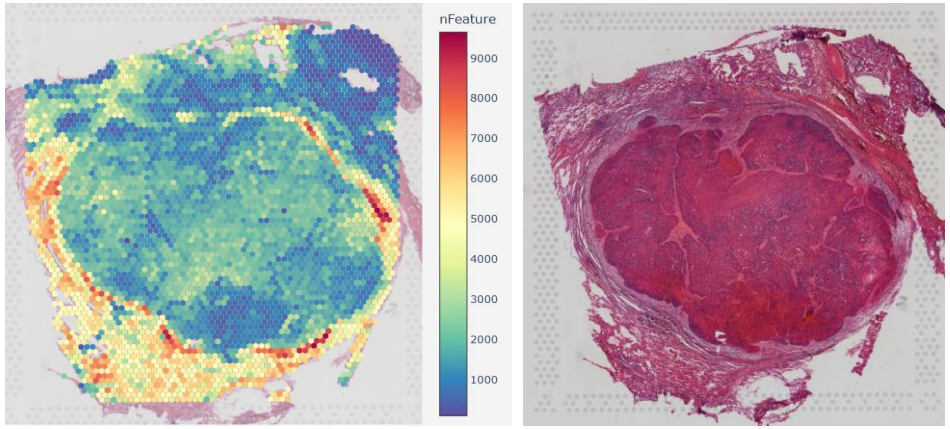
クオリティ値の空間パターン

- DeepSpaceDBからの例
- 一部のパターンでは生物学的な特徴を反映している
  - 脳の神経解剖学的構造
  - ガン組織の存在

マウス脳切片における検出遺伝子数

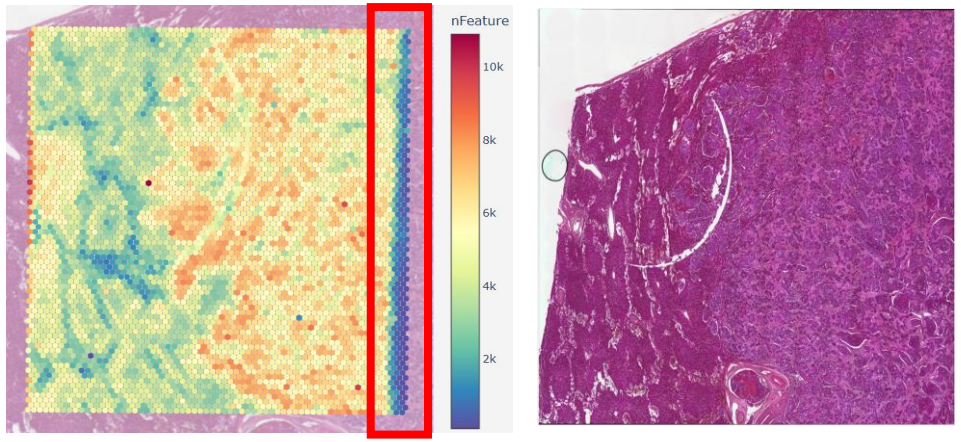


肝芽腫 (sample ID DSID001904)



技術的な問題でパターンが生じることもある

乳がん由来の肝臓転移
(sample ID DSID001370)



Seurat チュートリアル – 正規化

- 生データのUMIカウントは、スポット間のシーケンシング深度の違いに左右される
- スポット間の比較を妥当とするには正規化が必要である
- オプション:

  Option 2. ライブラリサイズ（シーケンシング深度）による正規化

  Option 1. SCTransform　負の二項分布に基づく複雑な正規化アプローチ

```
#### Normalization ####
# option 1: sctransform residuals for all genes
brain <- SCTransform(brain, assay = "Spatial")

# option 2: standard log normalization for comparison
# brain <- NormalizeData(brain, assay = "Spatial")
```

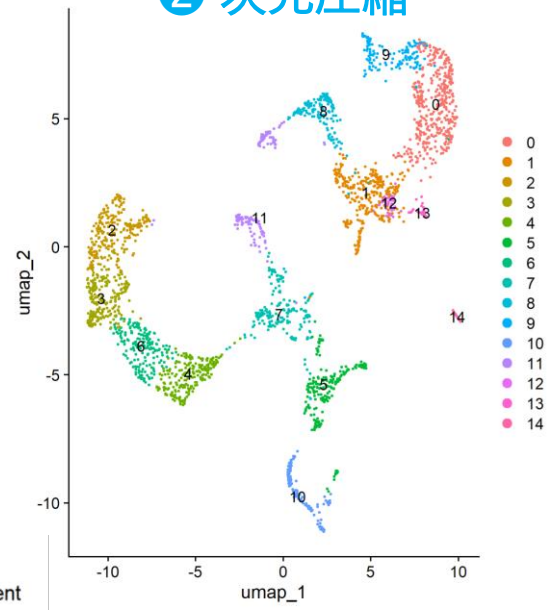
Seurat チュートリアル – 次元圧縮

探索的データ解析

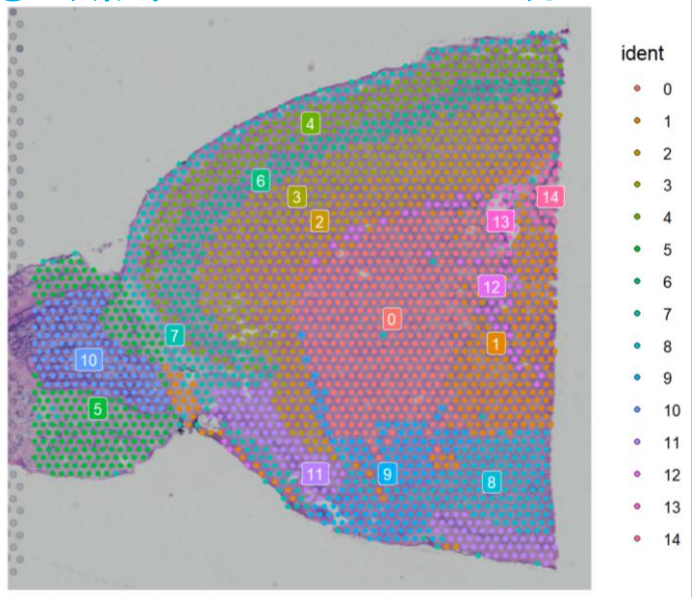
高次元データ → 人間が理解しやすい 単純化された要約

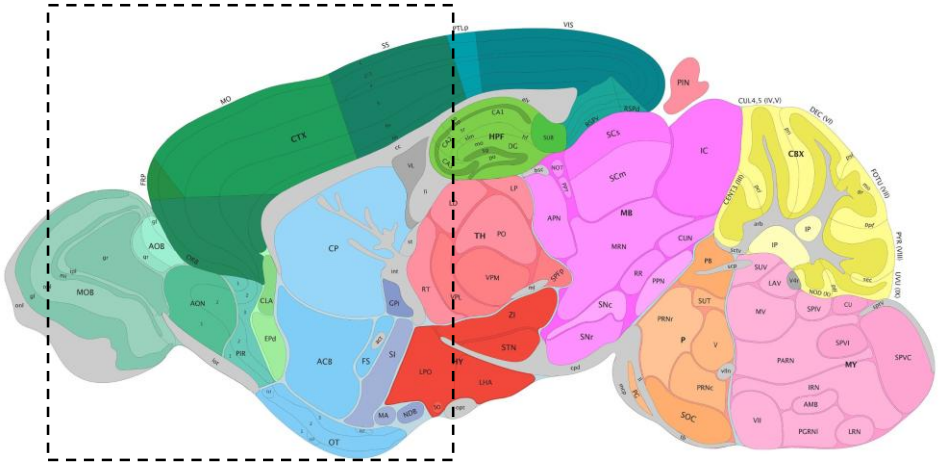
```
#### Dimensionality reduction & clustering ####
brain <- RunPCA(brain, assay = "SCT")
brain <- FindNeighbors(brain, reduction = "pca", dims = 1:30)
brain <- FindClusters(brain)
SpatialDimPlot(brain, label = TRUE, label.size = 3)   ← 1

brain <- RunUMAP(brain, reduction = "pca", dims = 1:30)
DimPlot(brain, reduction = "umap", label = TRUE)      ← 2
```

❷ 次元圧縮



❶ 類似性によるグループ分け

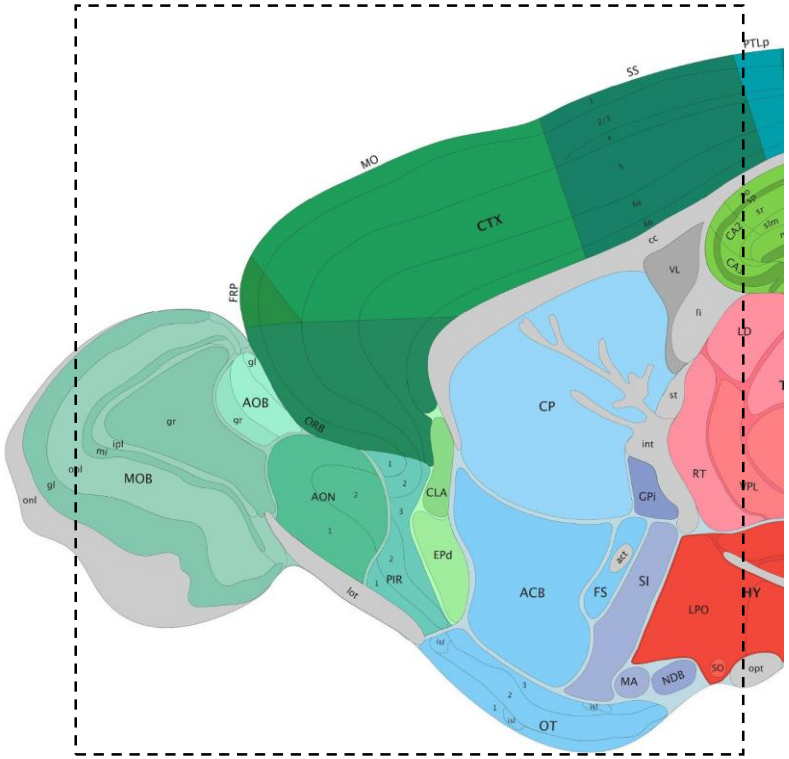




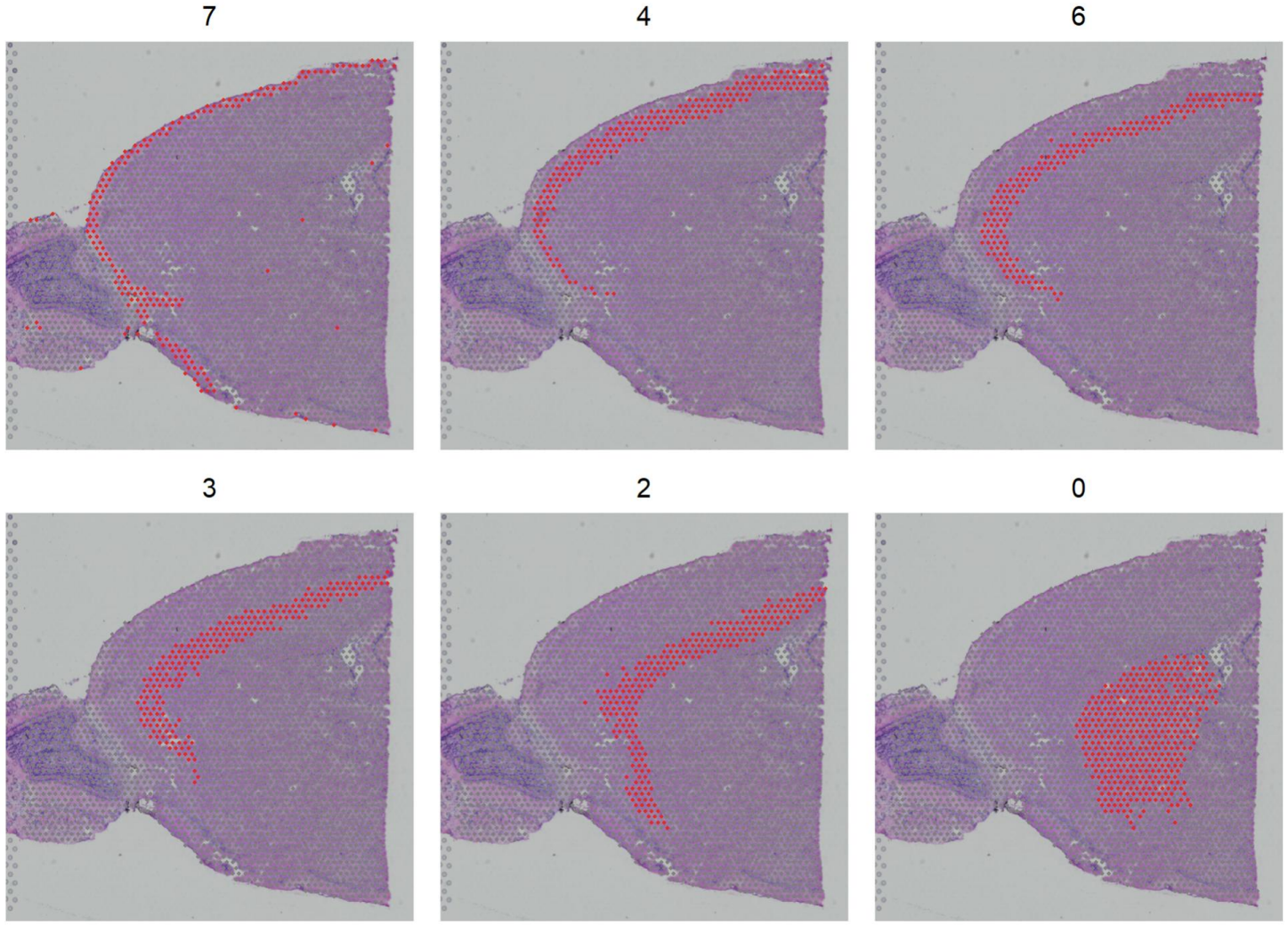
マウス脳 サジタル断面 (Image from Allen brain atlas)

Seurat tutorial – クラスタリング結果の検証


```
# highlighting clusters
SpatialDimPlot(brain,
   cells.highlight = CellsByIdentities(object =
                brain, idents = c(7, 4, 6, 3, 2, 0)),
   facet.highlight = TRUE, ncol = 3)
```



マウス脳サジタル断面
(Image adapted from Allen brain atlas)



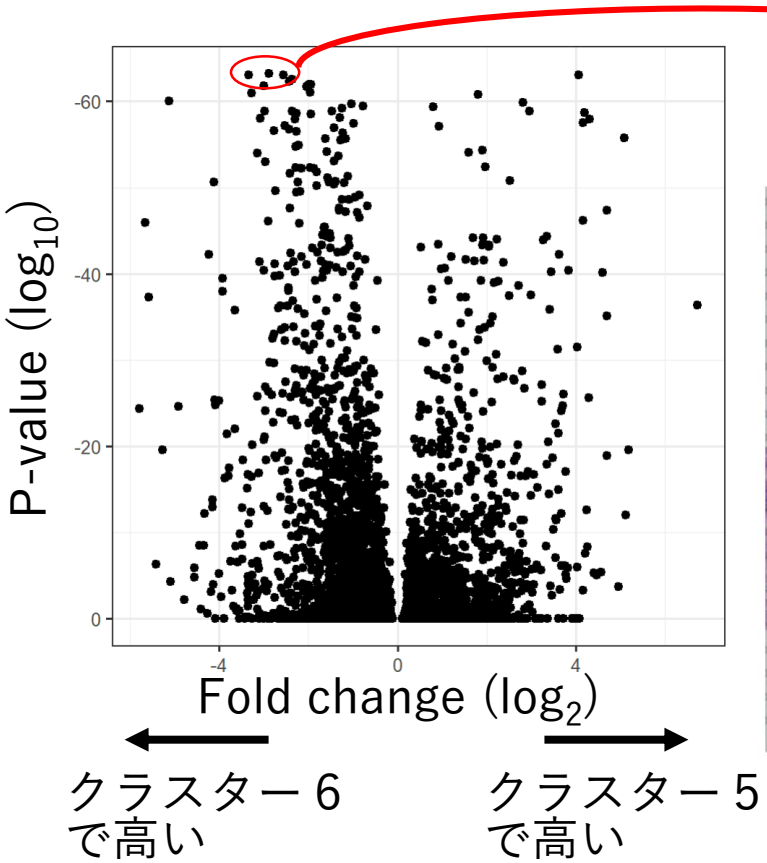
このクラスタリングは遺伝子発現パターンの類似性にのみ基づいており、空間的情報は一切用いていない

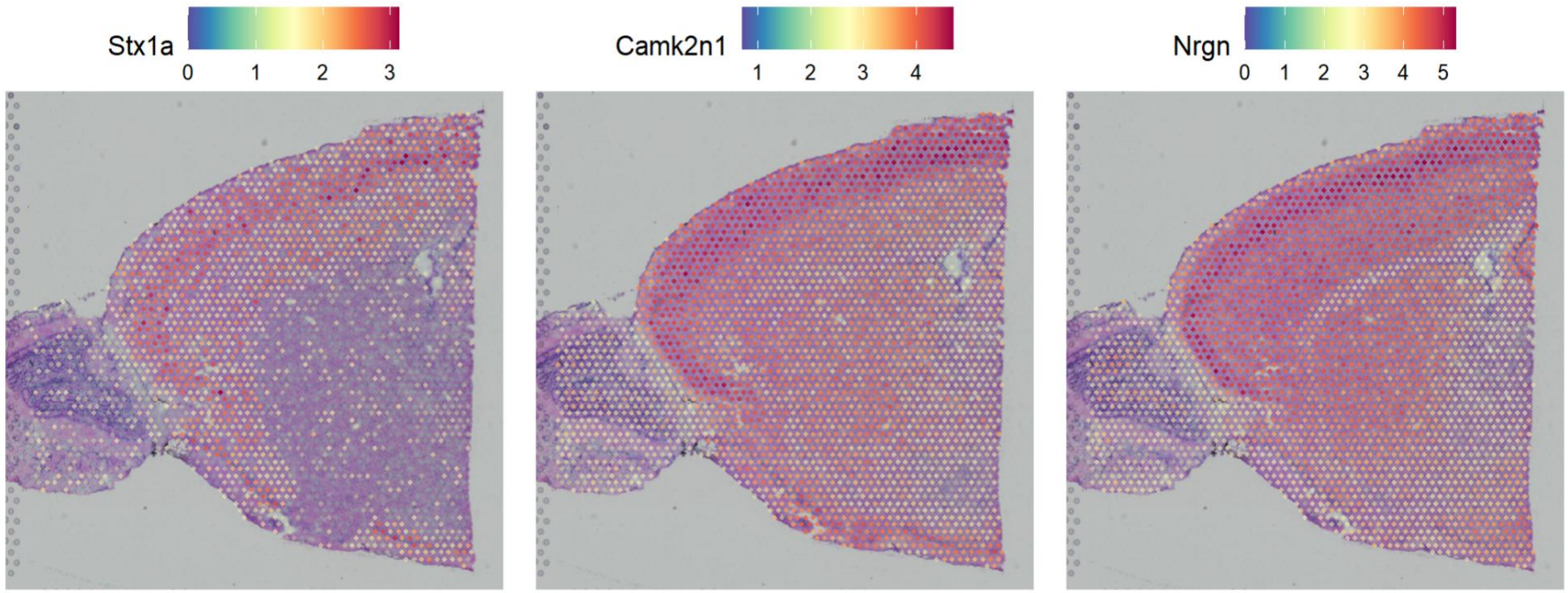
Seurat tutorial – 空間的変動遺伝子（SVG）(1)


```
#### Identifying spatially variable features ####
# comparisons between clusters (here: between 5 and 6)
de_markers <- FindMarkers(brain, ident.1 = 5, ident.2 = 6)
# plot the top 3 genes with the highest expression in cluster 6 vs 5
SpatialFeaturePlot(object = brain, features = rownames(de_markers)[1:3], alpha = c(0.1, 1), ncol = 3)
```

FindMarkersでは空間情報（位置や隣接関係など）は一切利用しない

クラスター５と６の間で発現に差がある遺伝子





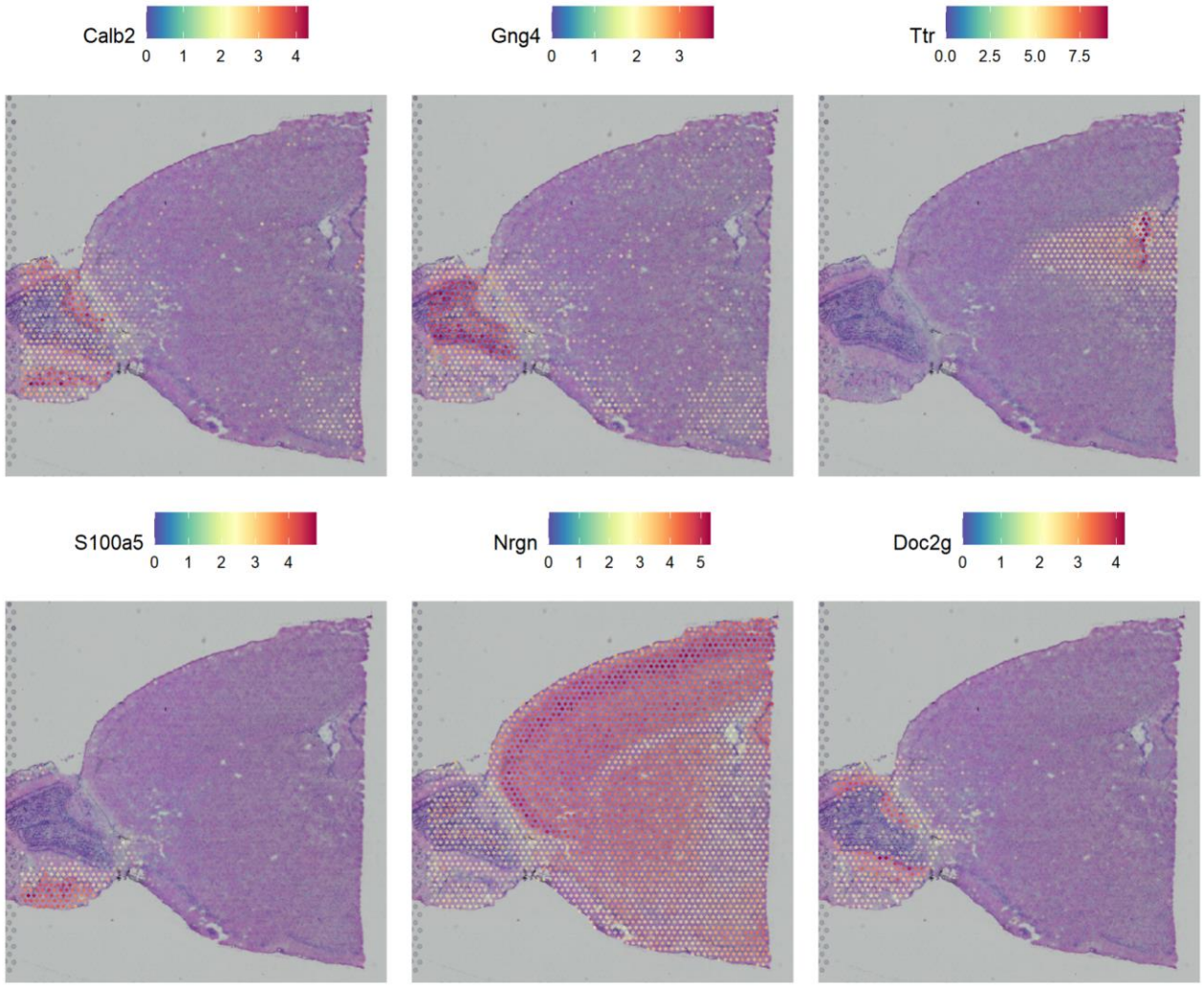
Seurat tutorial – 空間的変動遺伝子（SVG）(2)

空間情報を取り込んだSVG 予測法

- Moran's I (used in this tutorial)
- SPARK-X (Zhu *et al.*, Genome Biol, 2021)
- binSpect (Giotto package; Dries *et al.*, Genome Biol, 2021)
- singleCellHaystack (Vandenbon and Diez, Nat Commun, 2020)

```
# Use spatial info to find SVGs (slow)
# Optional speed-up: install.packages('Rfast2')
brain <- FindSpatiallyVariableFeatures(
   brain, assay = "SCT",
   features = VariableFeatures(brain)[1:1000],
   selection.method = "moransi")

# plot the top-scoring genes
top.features <- head(SpatiallyVariableFeatures(brain,
                        selection.method = "moransi"), 6)

SpatialFeaturePlot(brain, features = top.features,
                           ncol = 3, alpha = c(0.1, 1))
```

Moran's I 法によるトップ変動遺伝子



概要

- Rを使った空間トランスクリプトミクスデータ解析
  - Visium example
  - Other platforms
- 既存の空間トランスクリプトミクスデータベースの紹介
- DeepSpaceDBの紹介
  - Visium interface
  - Xenium interface
- まとめ

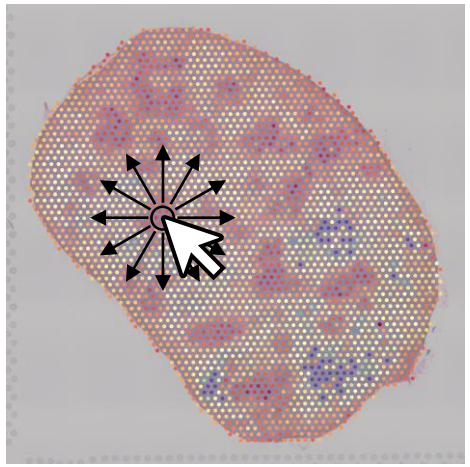
他のプラットフォームについて

- The Seuratのウェブサイト上のチュートリアルはシーケンスベースのプラットフォームを対象としている
  - Visium
  - Slide-seq
- しかし、イメージベースのプラットフォームも同様のワークフローで解析可能である
  - Vizgen MERSCOPE
  - Xenium
  - Nanostring CosMx Spatial Molecular Imager
  - Akoya CODEX

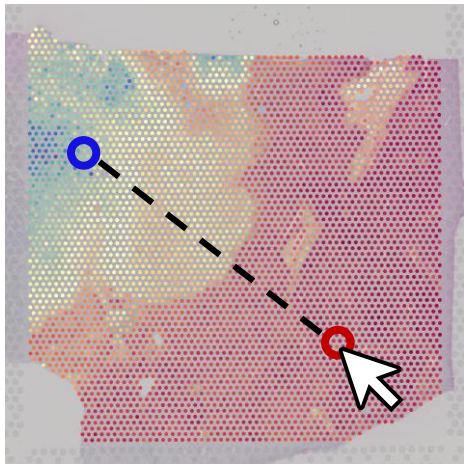
Seurat tutorial – 結論

- Seurat は空間データ解析ツールとして大変優れている
- 解析の大部分はscRNA-seq 解析と同様の流れである
- しかし空間的な要素を完全には活用していない
- インタラクティブな探索には限界がある

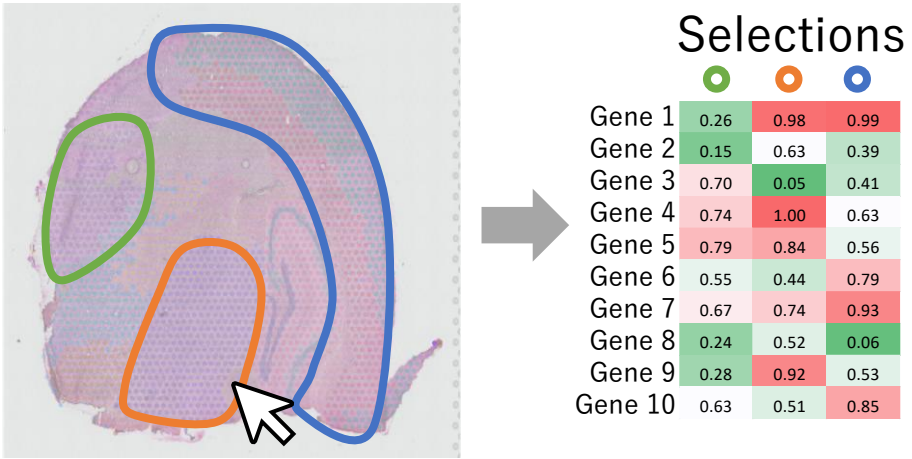
遺伝子発現は、ある点からの距離に応じてどのように変化するのか？



この軸に沿って起こっている遺伝子発現変化は何か?



マウスで選んだ領域どうしを比較したい。どの遺伝子の発現に差があるのか？



概要

- Rを使った空間トランスクリプトミクスデータ解析
  - Visium example
  - Other platforms
- **既存の空間トランスクリプトミクスデータベースの紹介**
- DeepSpaceDBの紹介
  - Visium interface
  - Xenium interface
- まとめ

Spatial transcriptomics databases

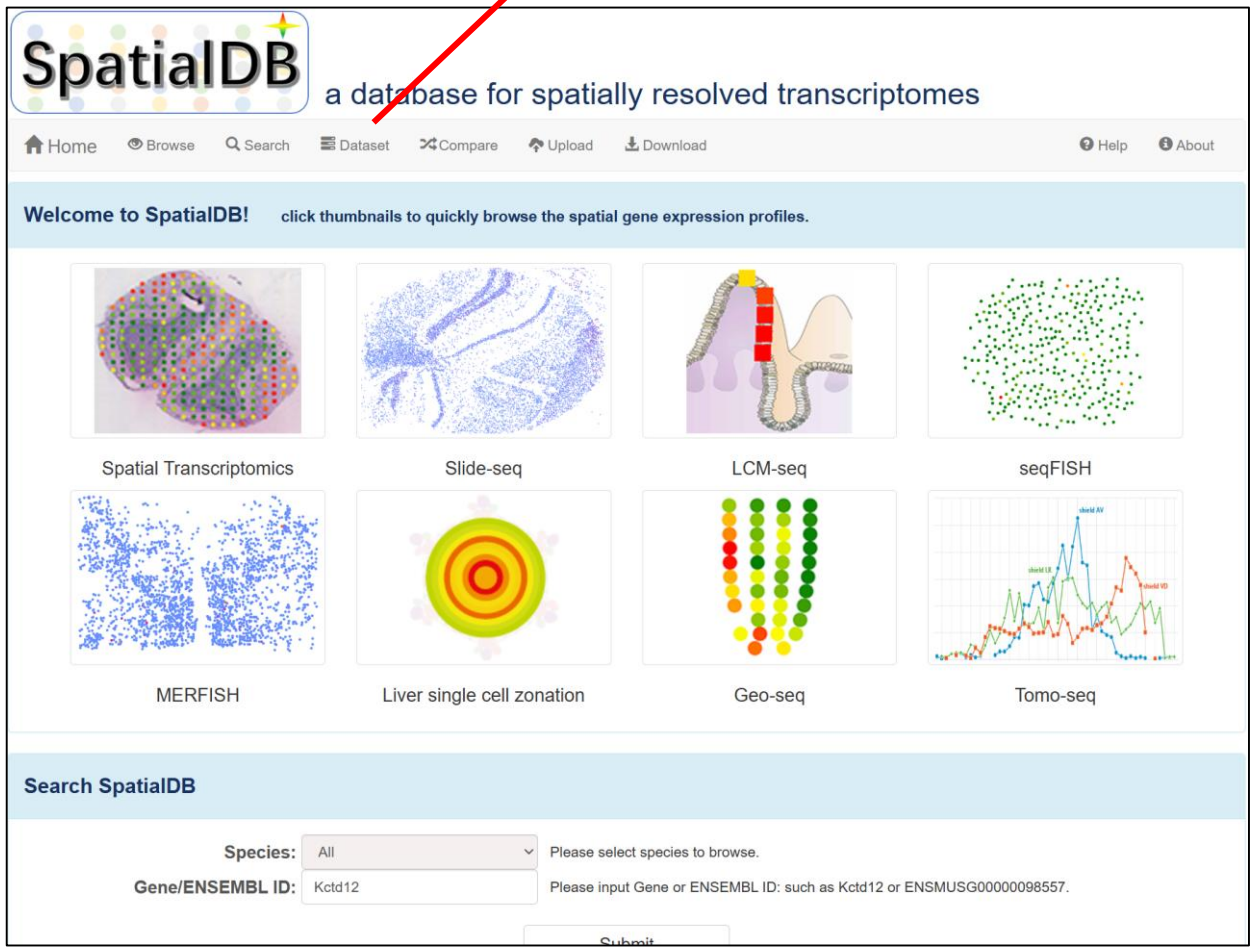
- 多くのデータベースが公開されている

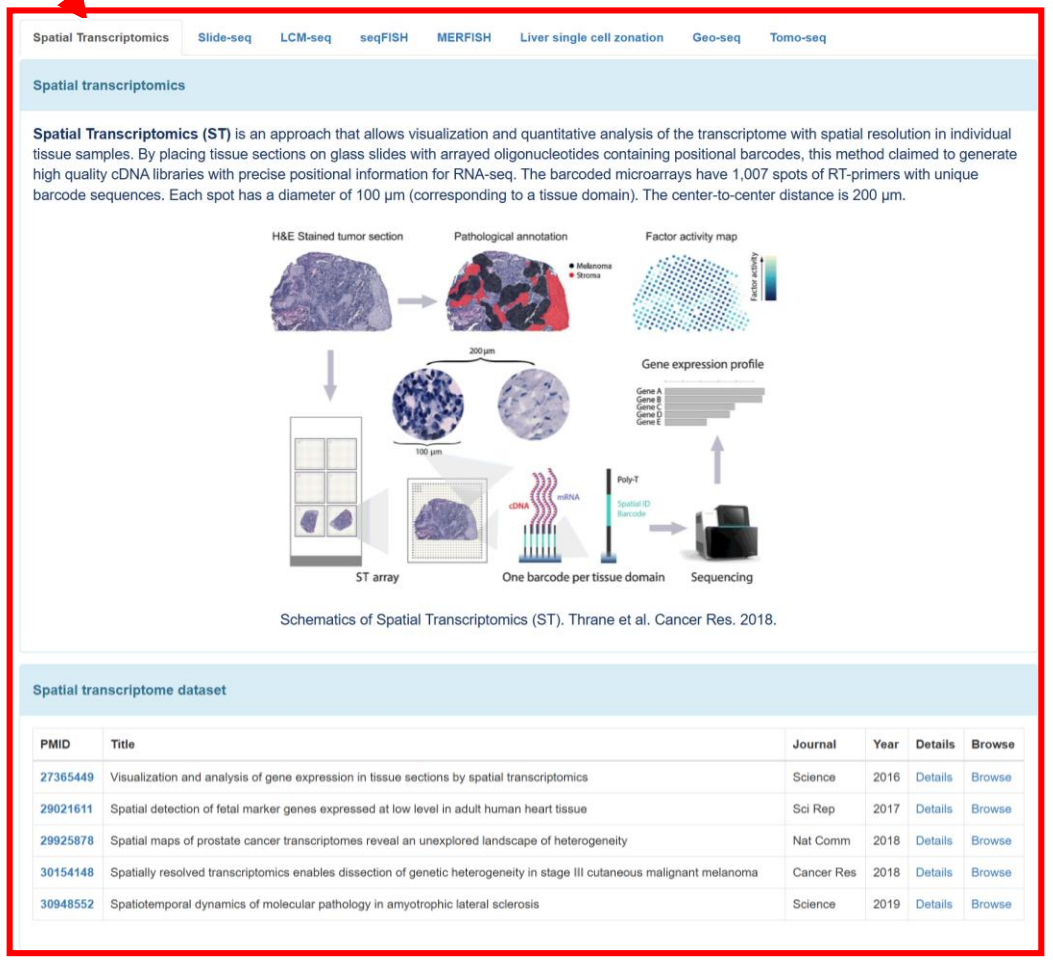
	Database	URL	Publication	PMID	Comment
★	SpatialDB	link	Nucleic Acids Res, 2020	31713629	No longer maintained?
★	SODB	link	Nat Methods, 2023	36797409	Sometimes long load times
	SCAR	link	Nucleic Acids Res, 2024	37739405	Focus on cancer
	SORC	link	Nucleic Acids Res, 2024	37811897	Focus on cancer
★	STOmicsDB	link	Nucleic Acids Res, 2024	37953328	>15,000 samples
★	SOAR	link	Sci Adv, 2025	40498826	Sometimes long load times
★	DeepSpaceDB	link	Nucleic Acids Res, 2025	41160880	Our database!

★:本日の紹介予定

SpatialDB

異なる実験手法・技術からなるサンプル



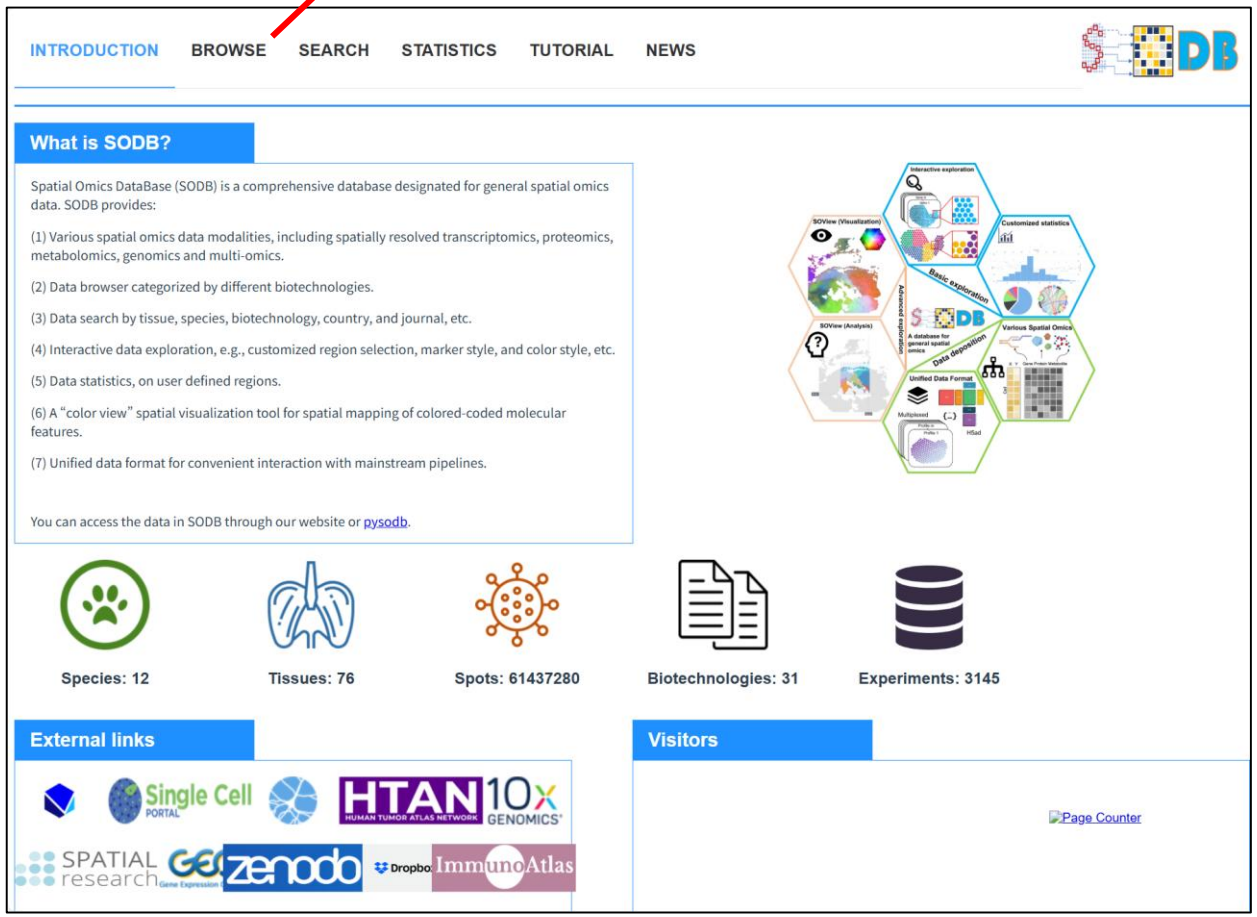


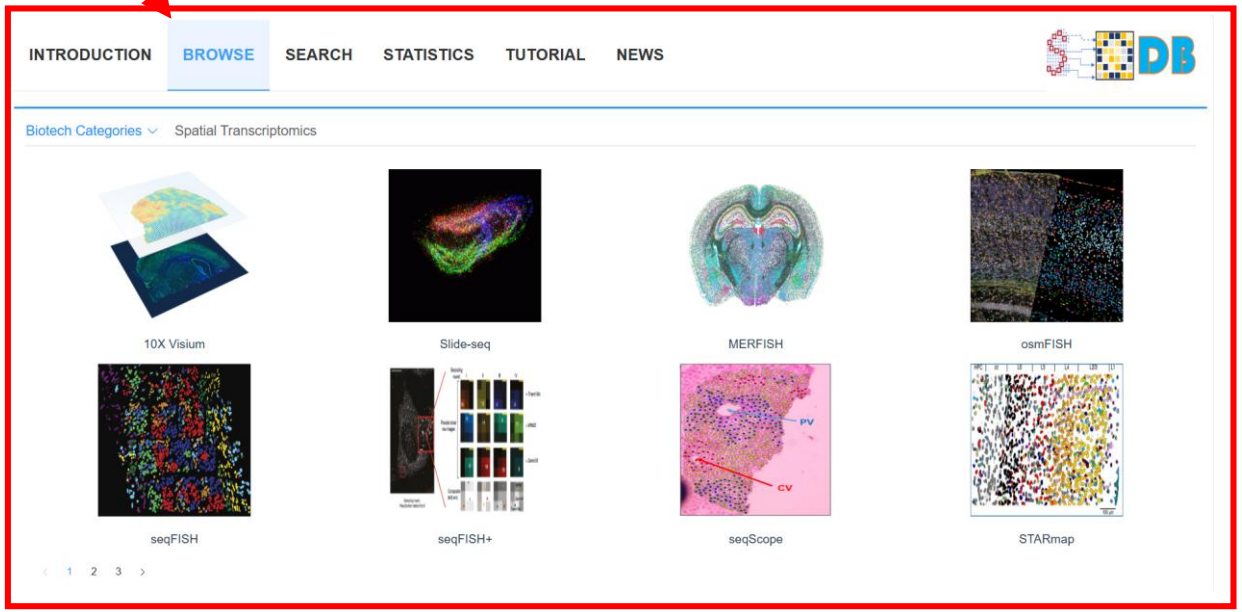
- 24 studies
- 8 Biotechnologies
- 305 samples

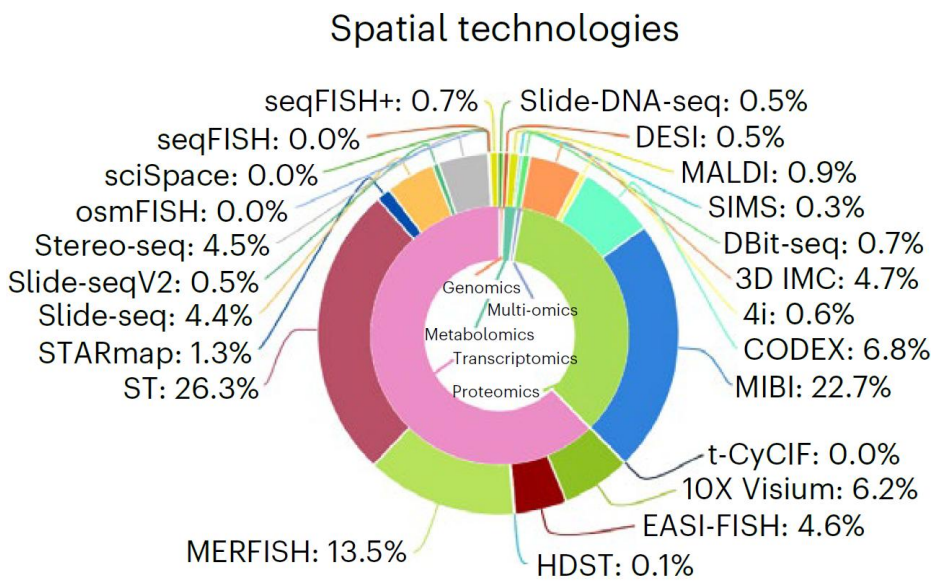
- 組織切片上に発現遺伝子をプロットできるが
- **公開以来一度も更新されていない**
- **いくつかのツールが動かない**

SODB

異なる実験手法・技術の数々



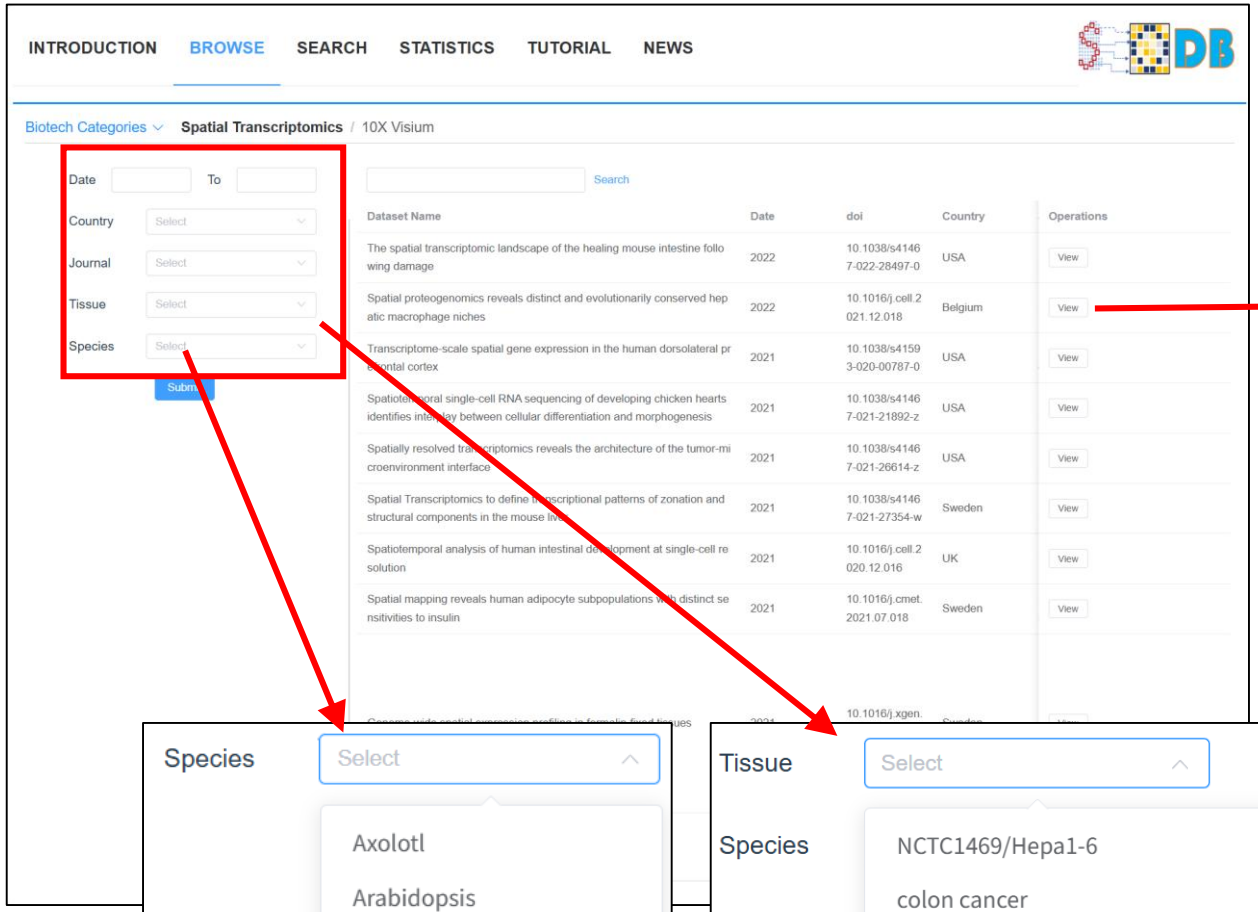


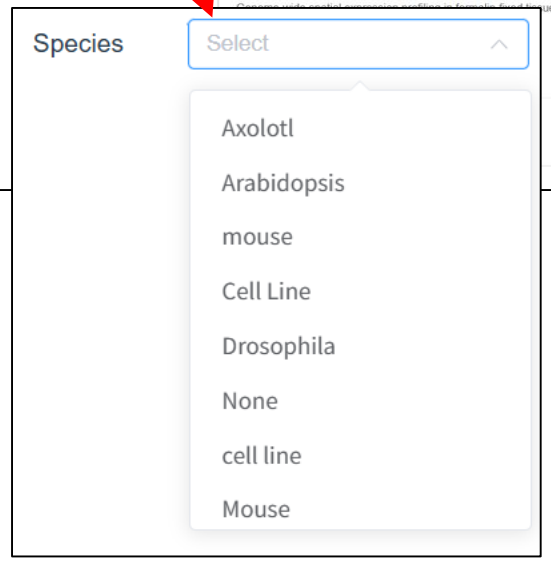


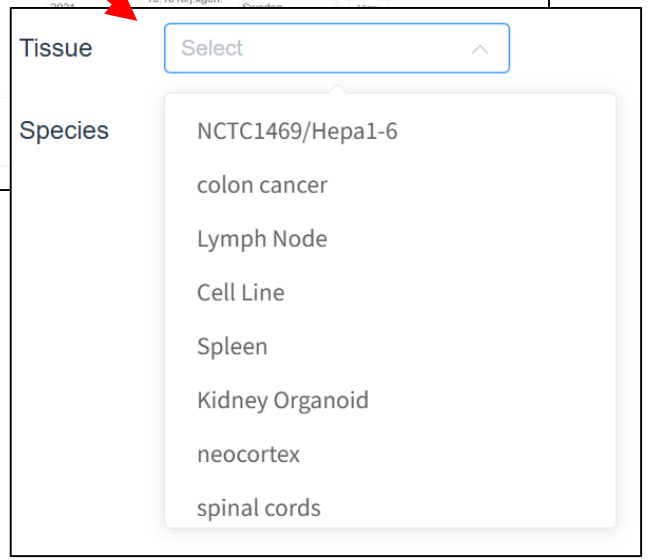
From Yuan et al., Nat Methods, 2023

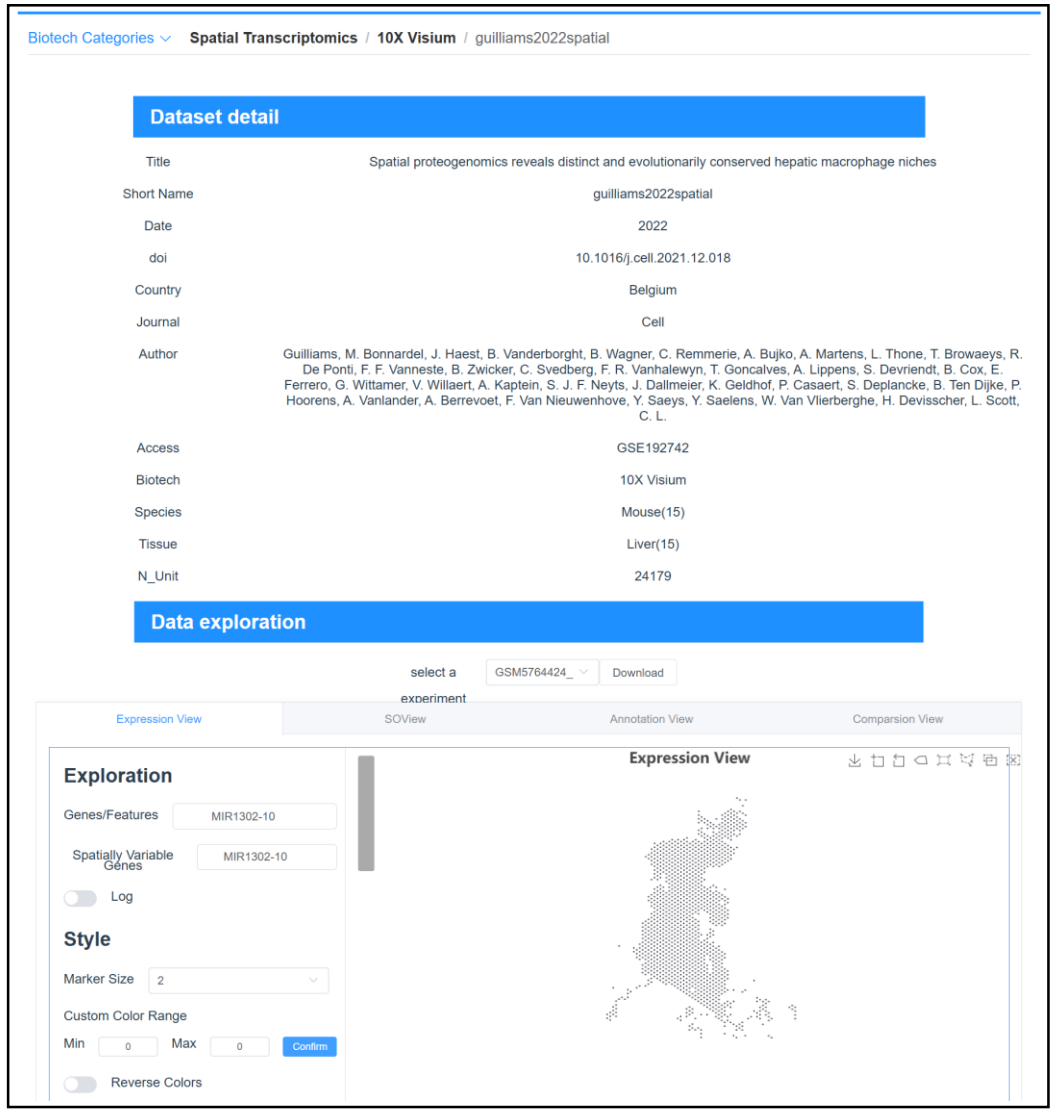
- 31 Biotechnologies
- 3145 samples
- しばしば反応が遅くなる

SODB – browsing Visium data







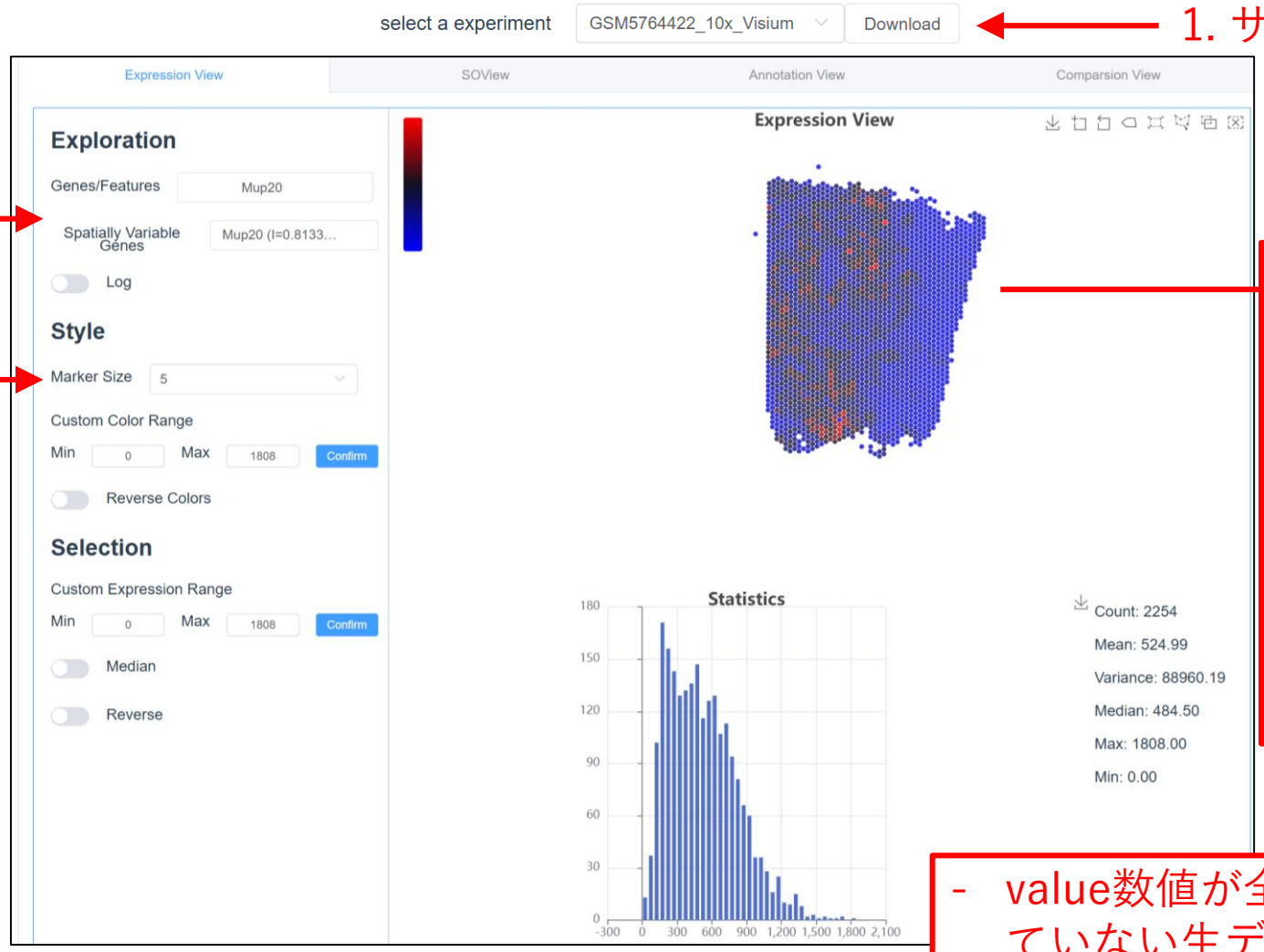


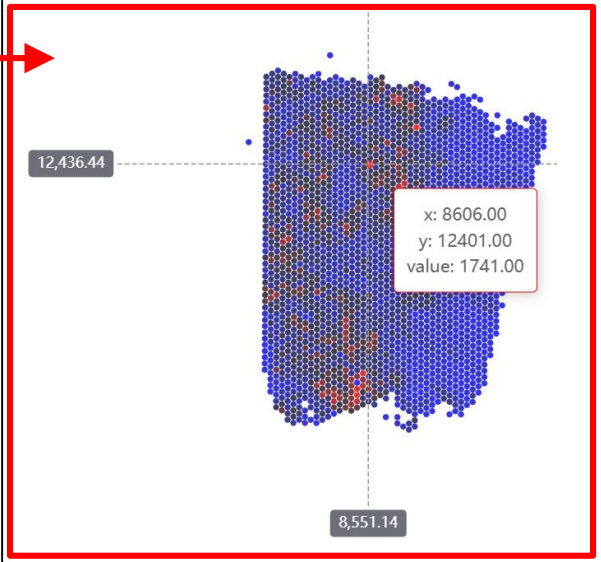
SODB – Expression View

1. サンプルを選ぶ GSM5764422

2. 遺伝子を選ぶ

3. (optional)
マーカーサイズ
を変更する

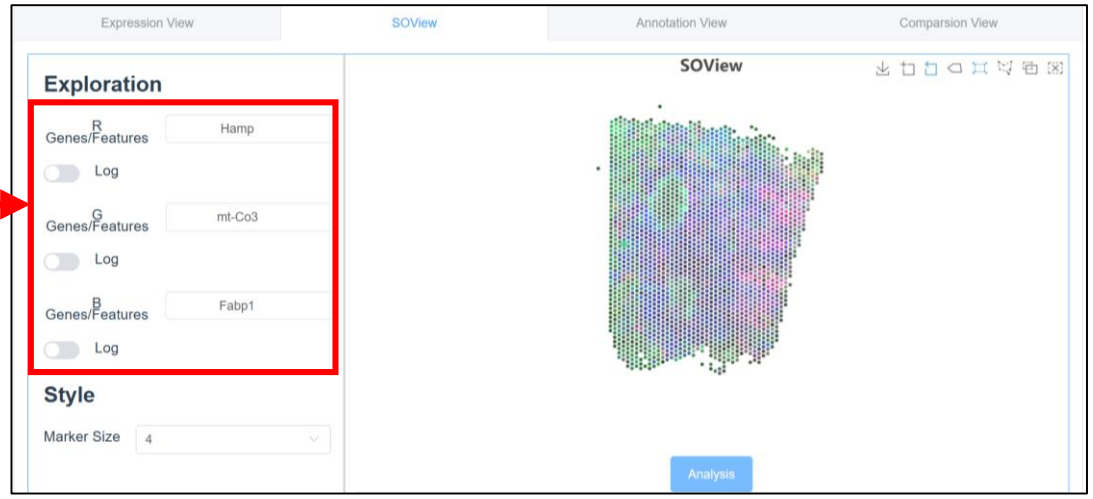




- value数値が全て整数なので、正規化していない生データのままの可能性がある
- 切片画像データが付いていない

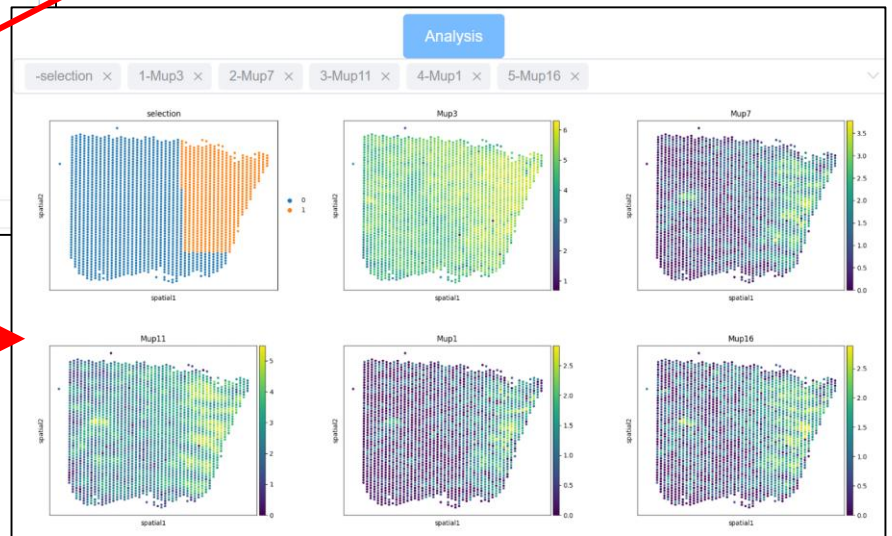
SODB – SOView

遺伝子発現分布を
３つまで表示する
ことができる（赤青
緑）

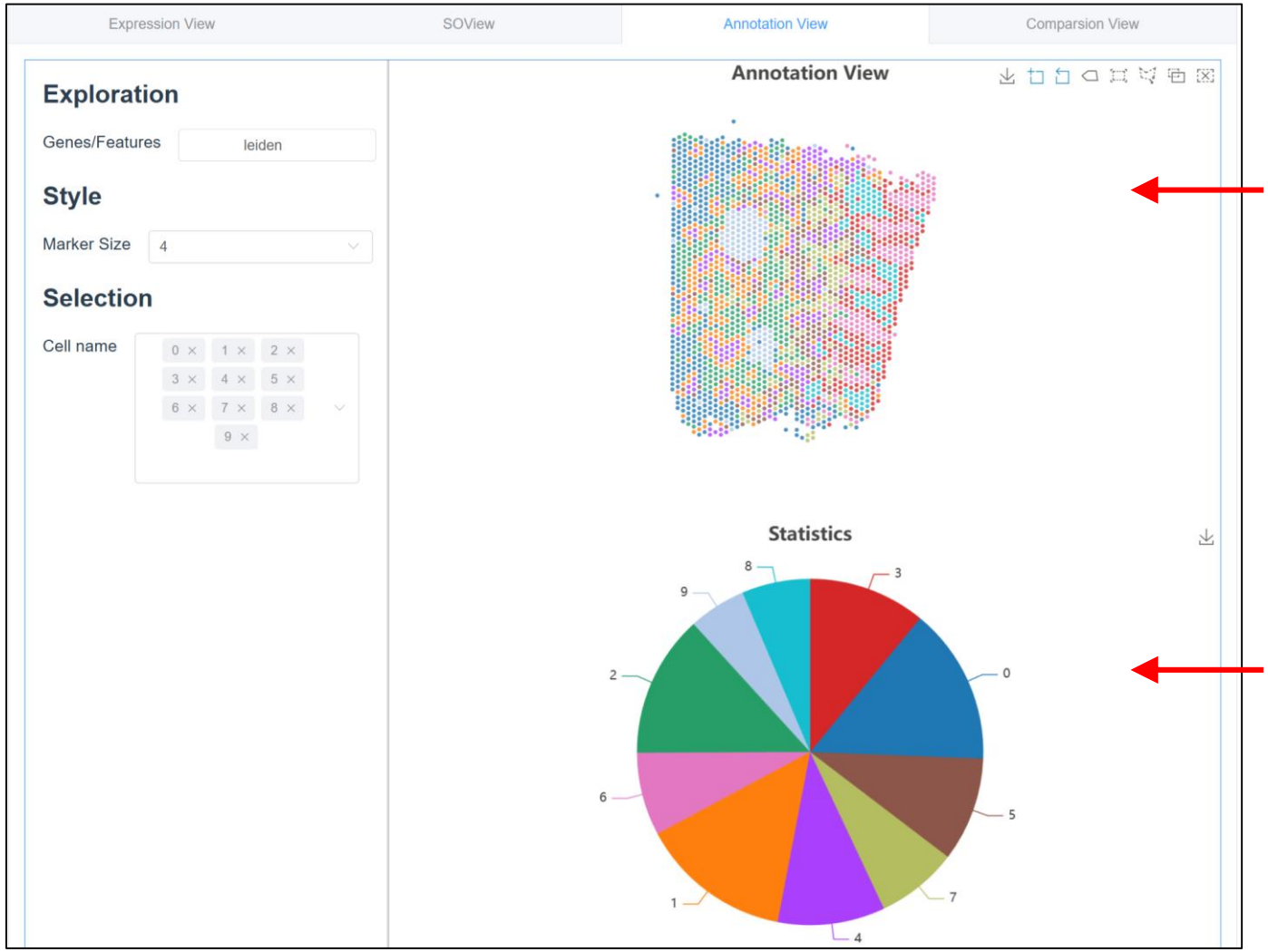


興味ある領域を選んで “Analysis”
をクリックすると

選択した領域内で発現が高い遺伝子が
表示される.



SODB – Annotation View

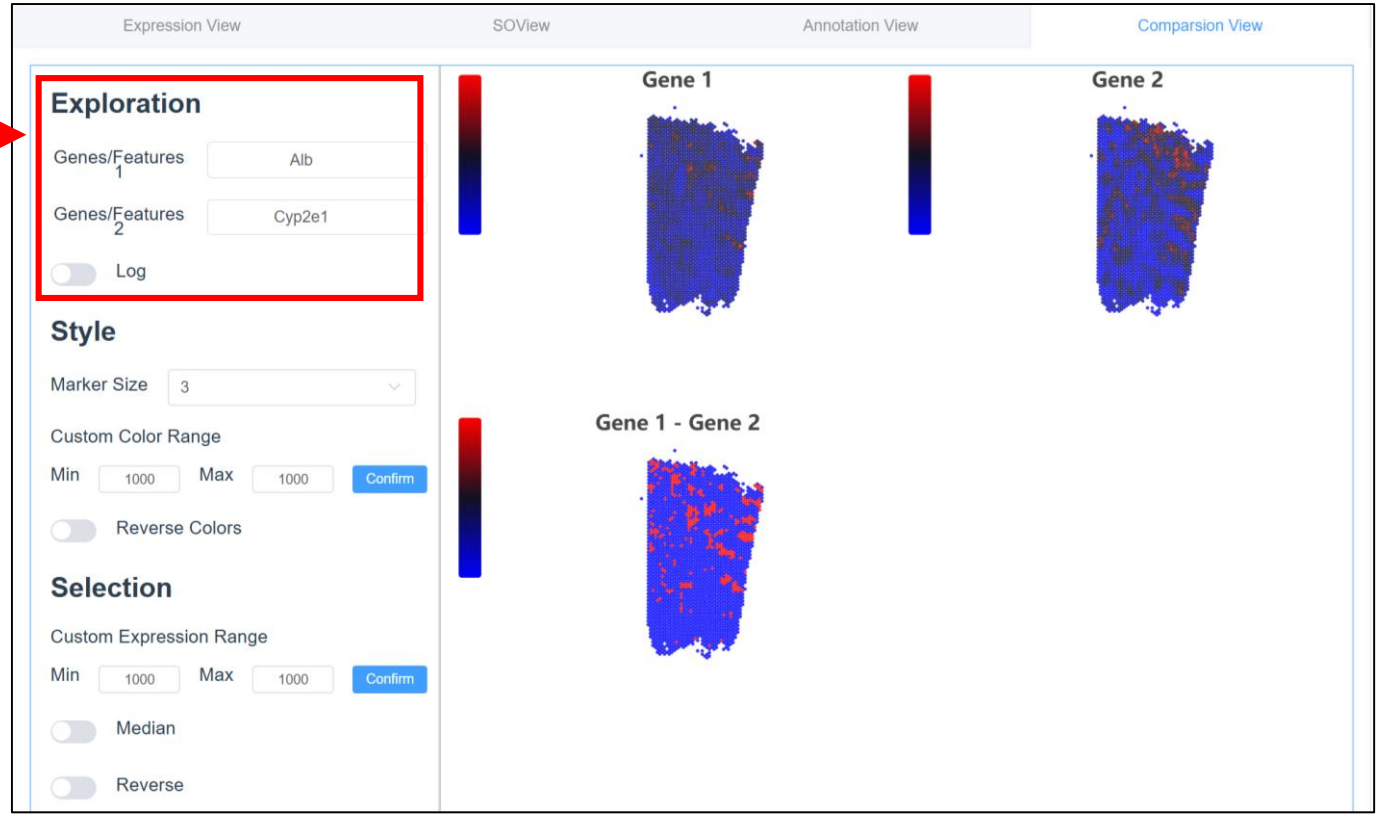


クラスタリング結果

クラスター当たりのスポット数

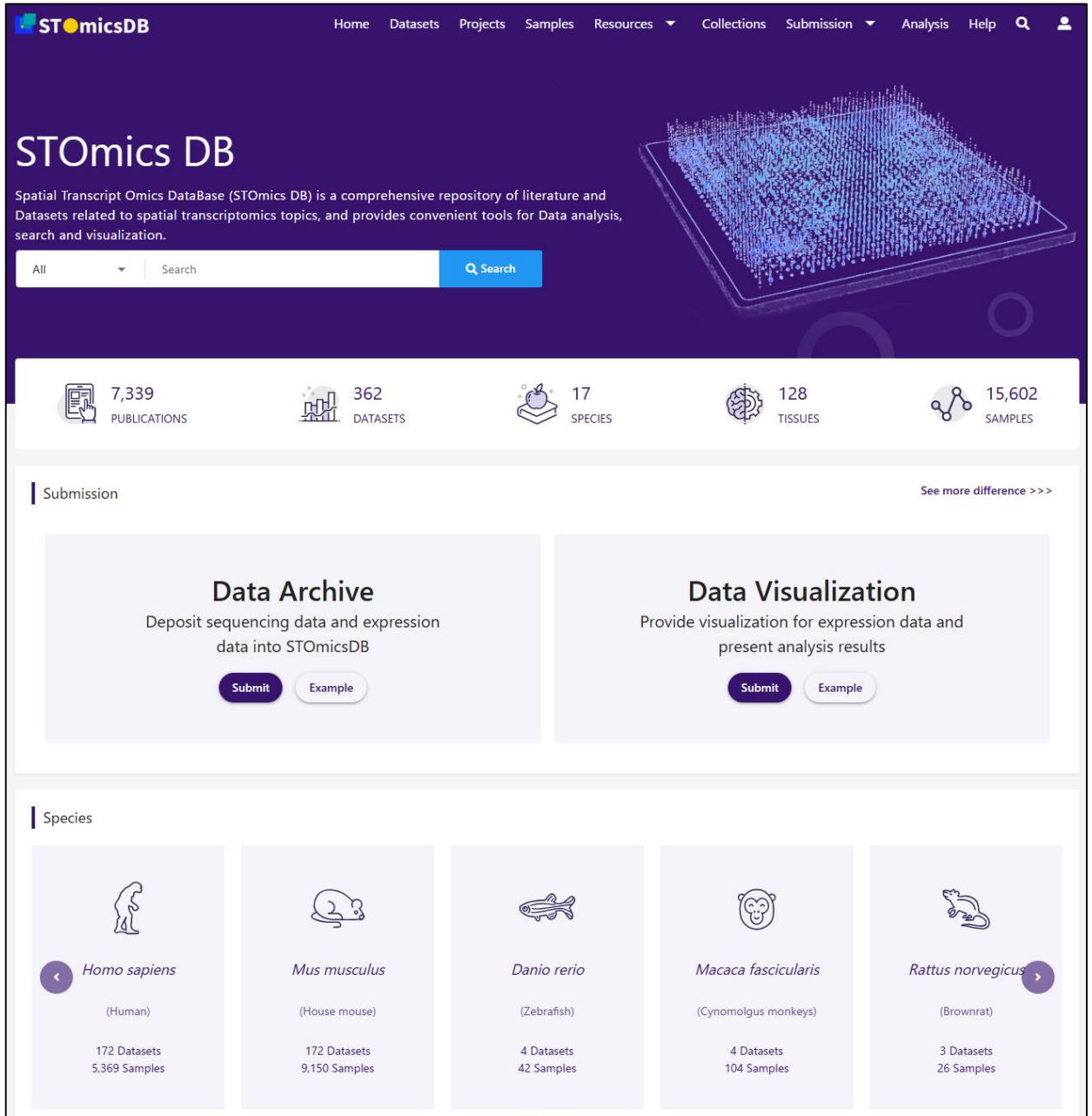
SODB – Comparison View

遺伝子を２つ選択

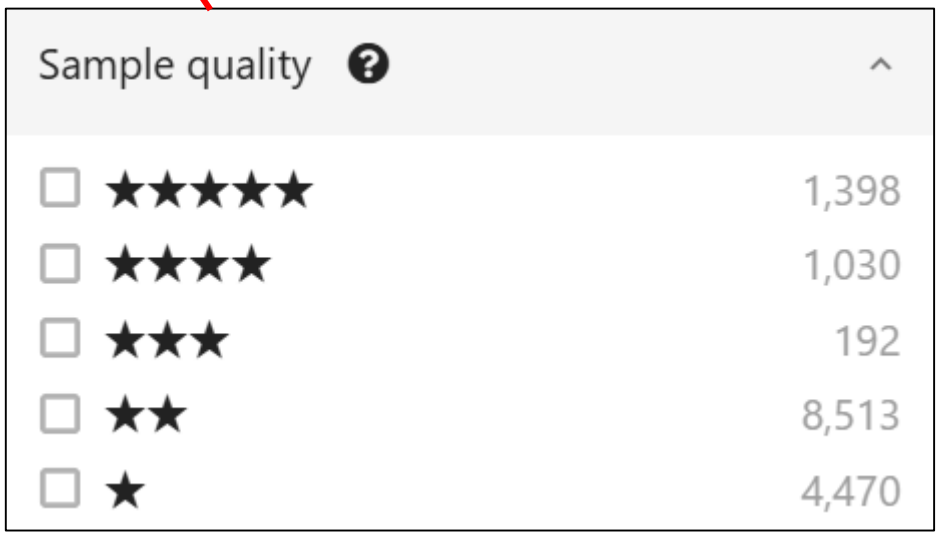


- 遺伝子１の発現分布
- 遺伝子２の発現分布
- 両者間の差

STOmicsDB

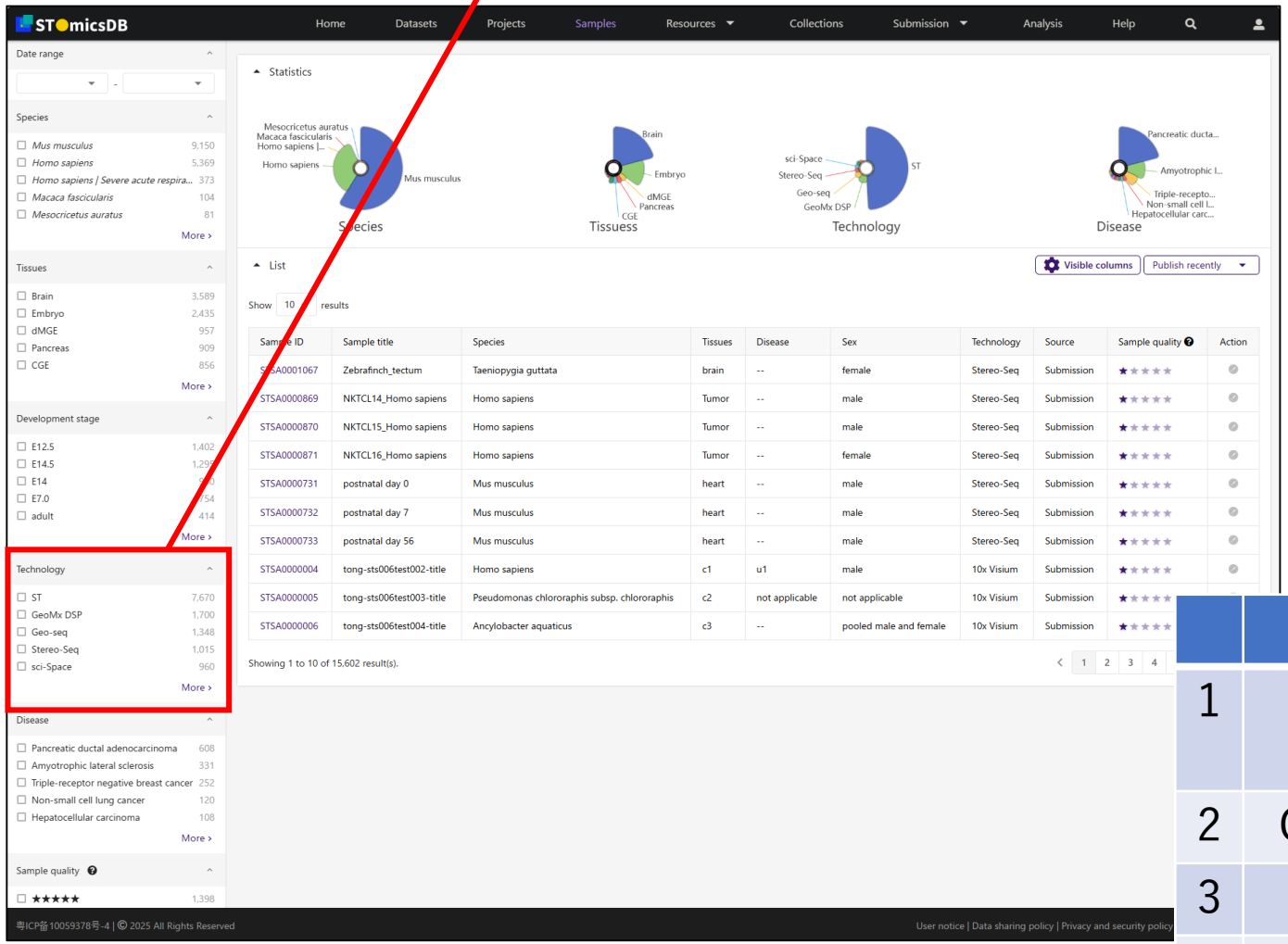


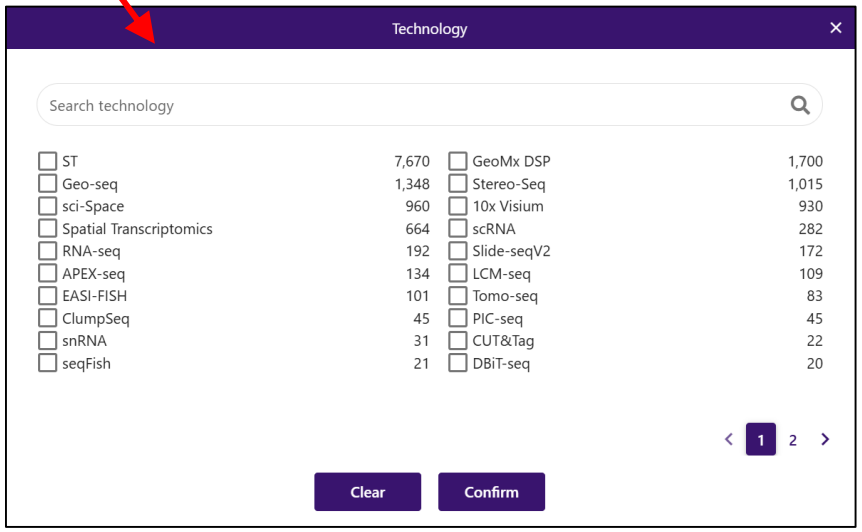
- 362 studies
- 8 technologies
- 15,602 samples (!) 但し品質の良いサンプルは少ない



5★: spatial coordinates & images
4★: spatial coordinates only
3★: images only
2★: lacks coordinates & images
1★: not processed

STOmicsDB – “Samples” page





様々な技術

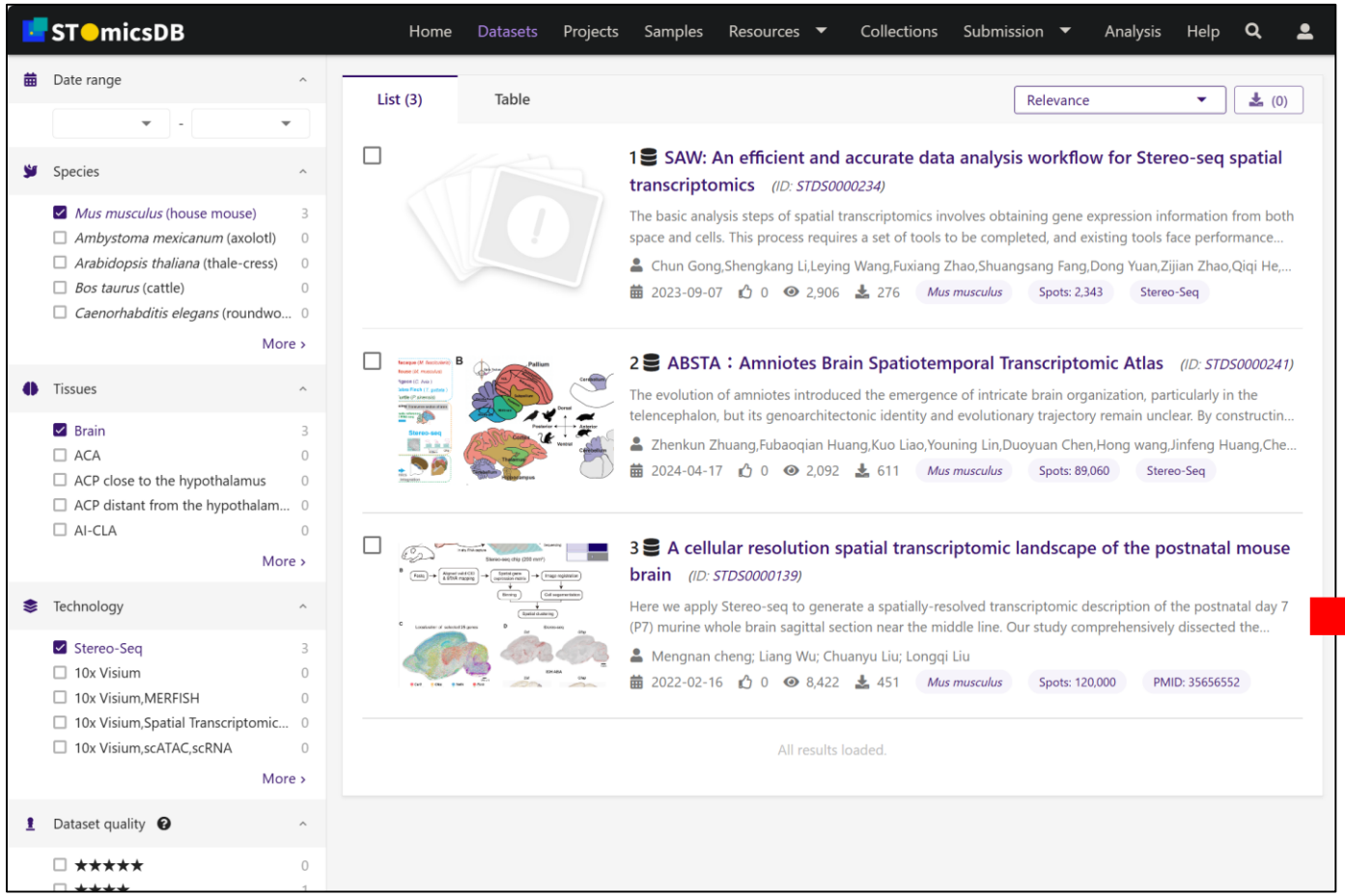
	Technology	Count	Comment
1	ST	7,670	≈ predecessor of Visium
2	GeoMX DSP	1,700	By Nanostring
3	Geo-seq	1,348	
4	Stereo-seq	1,015	By BGI
5	Sci-Space	960	
6	10x Visium	930	By 10X Genomics

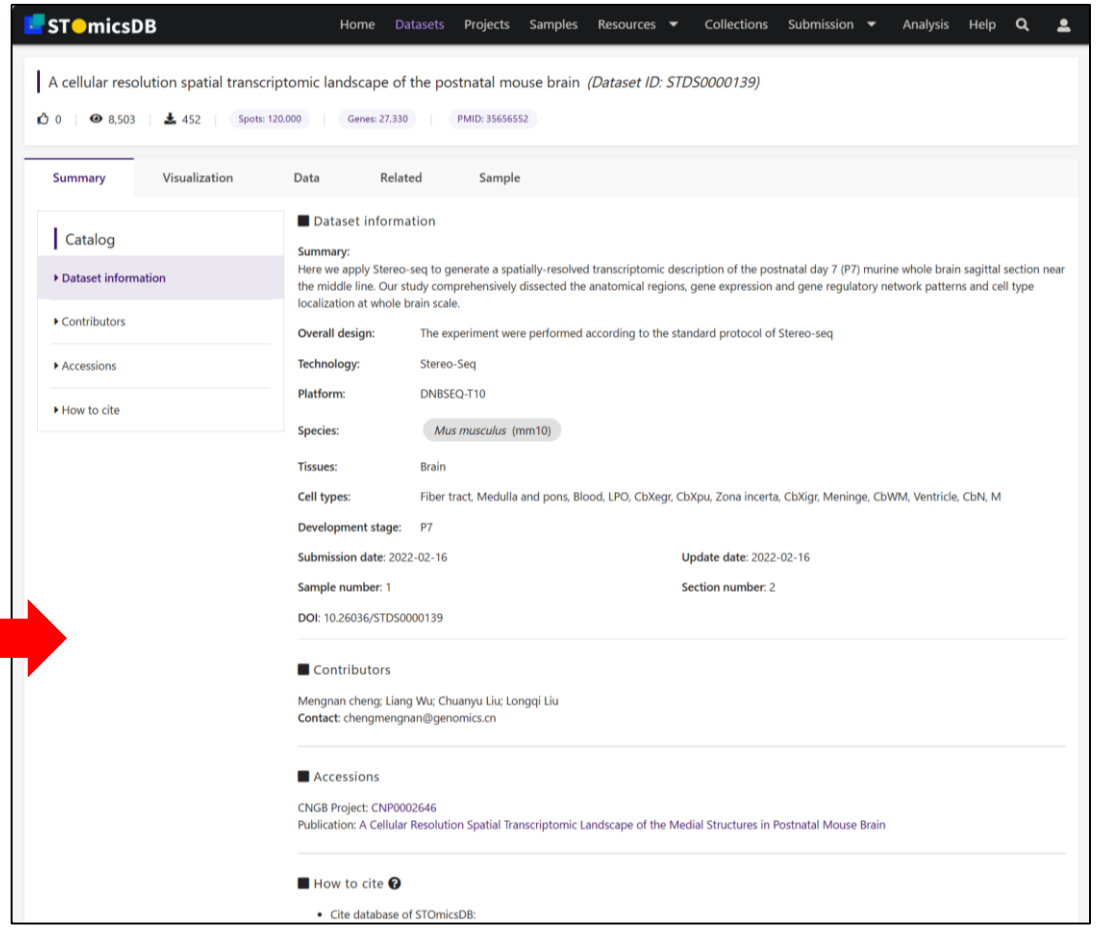
異なる技術由来のサンプルを満遍なく集めているが、利用できる解析ツールはサンプルごとに限定されている。

STOmicsDB – “Datasets” page

フィルターをセットする: Species = “mouse”,
Tissues = “Brain”, Technology = “Stereo-seq”

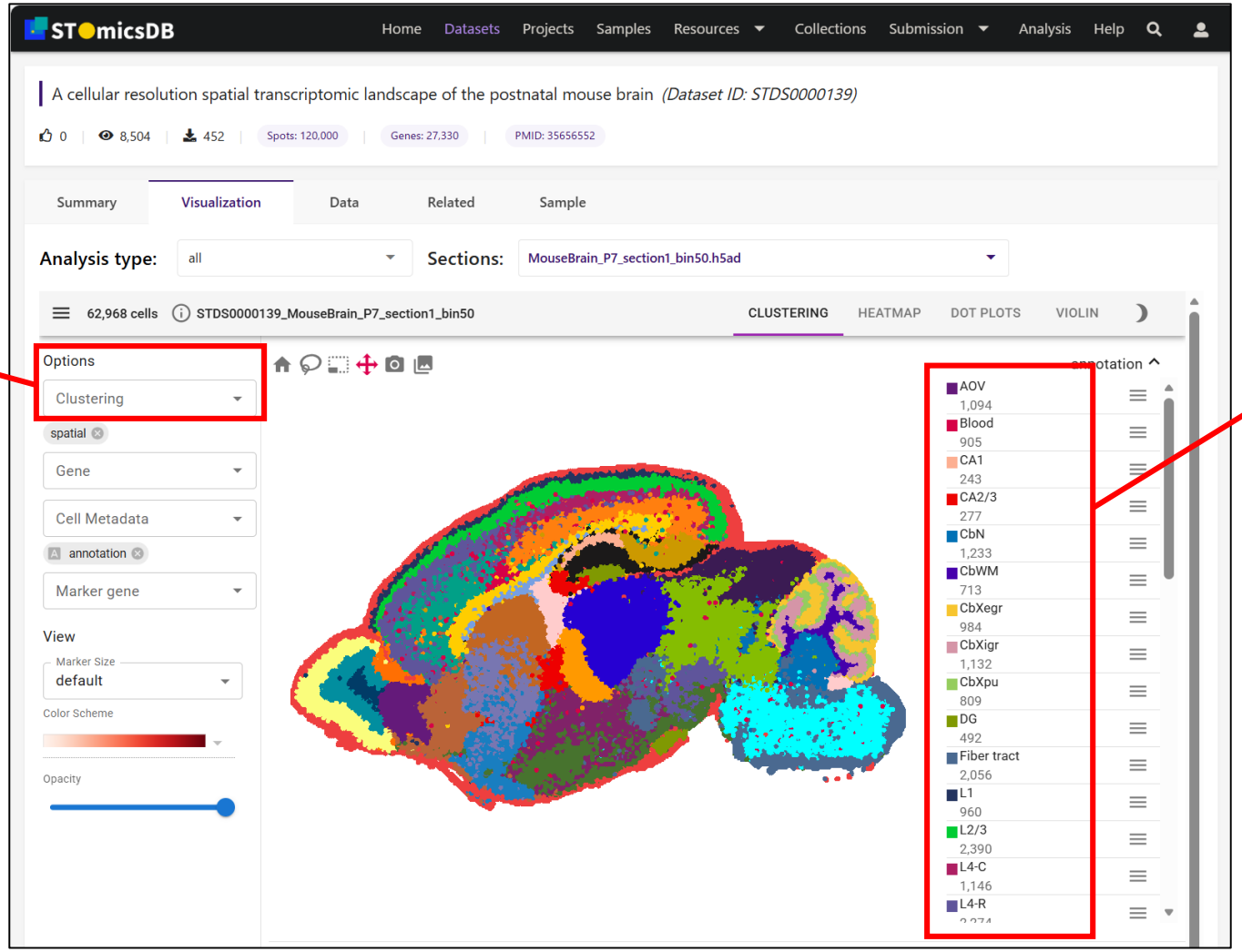
左のリストから一つ選んだ結果





STOmicsDB – Visualization (1)

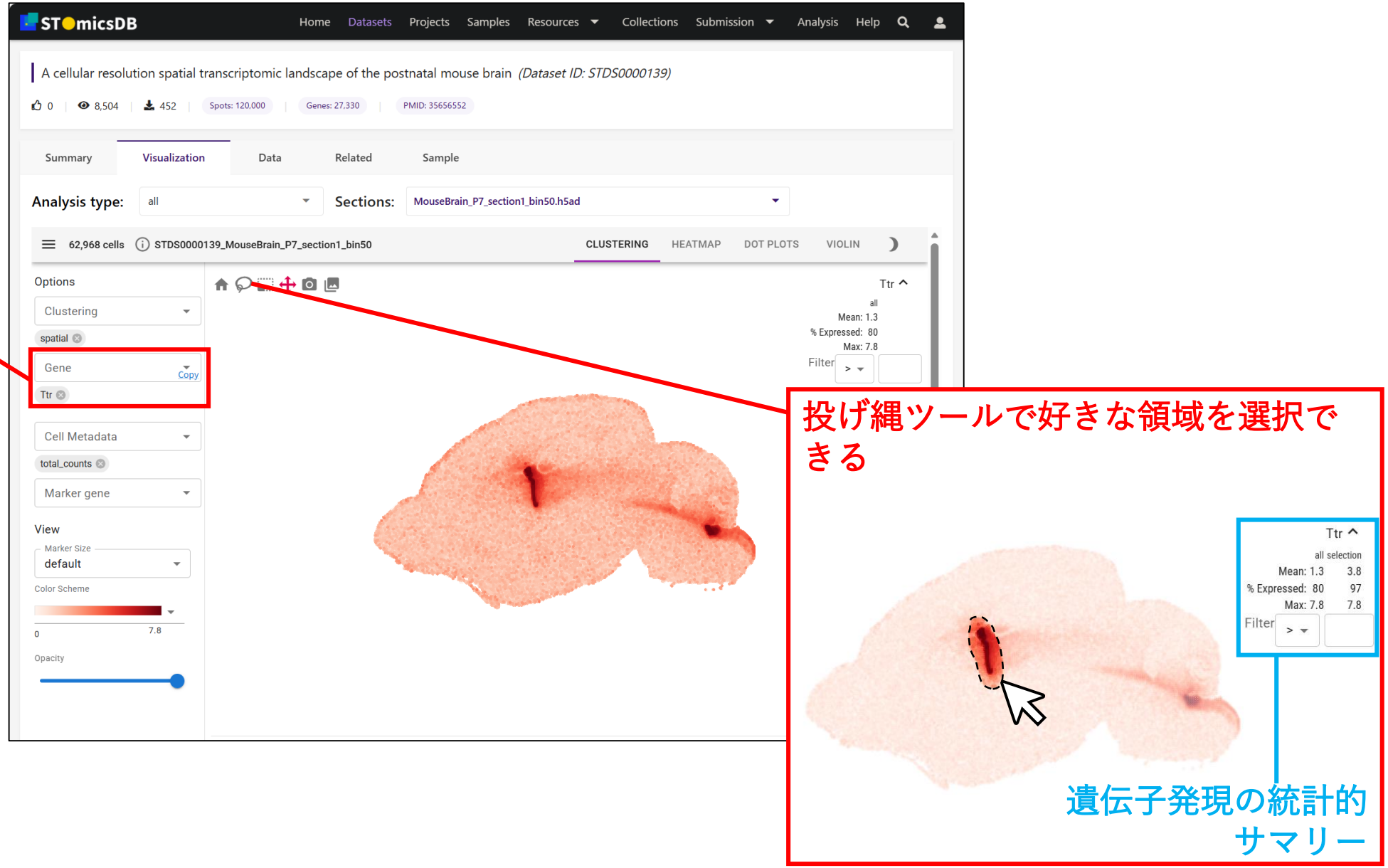
右パネルの画像表示種類を選択できるように見えるが、Stereo-seqサンプルの場合は1種類しかない



Here:
細胞種や組織のアノテーション

STOmicsDB – Visualization (2)

遺伝子発現パターンを見る



STOmicsDB – Data

STOmicsDB Home Datasets Projects Samples Resources Collections Submission Analysis Help

A cellular resolution spatial transcriptomic landscape of the postnatal mouse brain *(Dataset ID: STDS0000139)*

0 | 8,507 | 452 | Spots: 120,000 | Genes: 27,330 | PMID: 35656552

Summary Visualization **Data** Related Sample

Data type

stomics

Show 10 results

Analysis type

Download table (0)

	File name	File type	Analysis type	Sample name	Size	Identifiers
☐	MouseBrain_P7_section1_bin50.h5ad	h5ad	processed	P7_section1_Bin50	3.0G	
☐	MouseBrain_P7_section1_singlecell.h5ad	h5ad	processed	P7_section1_singlecell	1.5G	

Showing 1 to 2 of 2 result(s).

1

*h5ad: AnnData format of original data, without any additional modifications by STOmicsDB.
*processde.h5ad: AnnData format, with analysis results generated by STOmicsDB.

h5ad 形式でダウンロード可能

Summary Visualization Data **Related** Sample

Related type	Related ID	Related link	Related title
CNSA project	CNP0002646	https://db.cngb.org/search/project/CNP0002646/	--

Summary Visualization Data Related **Sample**

Show 10 results

Sample ID	Sample title	Species	Tissues	Disease	Sex	Technology	Source	Sample quality	action
STSP0005129	P7_section1_singlecell	Mus musculus	Brain	--	--	Stereo-Seq	CNGB	★★★★★	
STSP0005187	P7_section1_Bin50	Mus musculus	Brain	--	--	Stereo-Seq	CNGB	★★★★★	

Showing 1 to 2 of 2 result(s).

1

生データへのリンク

CNGBdb GENOS Scientific databases Submission Analysis Download About Login

Home / Data resources / Project / CNP0002646

Spatial transcriptomics of brain for the postantal day 7 mouse

Source: CNGBdb Project (ID CNP0002646)

474 270

Description: Brain develops in an intricately orchestrated sequence of stages, which starts with neurulation from the ectoderm of the embryo and takes a long time to postnatal to mature. A complete understanding of this process requires a systematic characterization of cell states over the entire spatiotemporal range of brain development. Recently technology of spatial transcriptome has brought new insight to the field of brain science. This new era capitalizes on high-resolution molecular tools that enable large-scale mapping of markers, which, in an unbiased and unsupervised fashion, capture the complexity of the nervous system. Here we use stereo-seq performed the spatial characteristic of the whole brain of the developing mouse at a key timepoint- postnatal day 7 (P7).

Data type: Raw sequence reads

Sample scope: Monoisolate

Submitter: 成梦南(Grace Cheng); 深圳华大生命科学研究院

Release date: 2022-04-29

Last updated: 2022-01-27

DOI: 10.26036/CNP0002646

Statistics: 2 samples; 2 experiments; 12 runs

Data size: 2.2TB

Download data | Download metadata | Get FTP links of all files

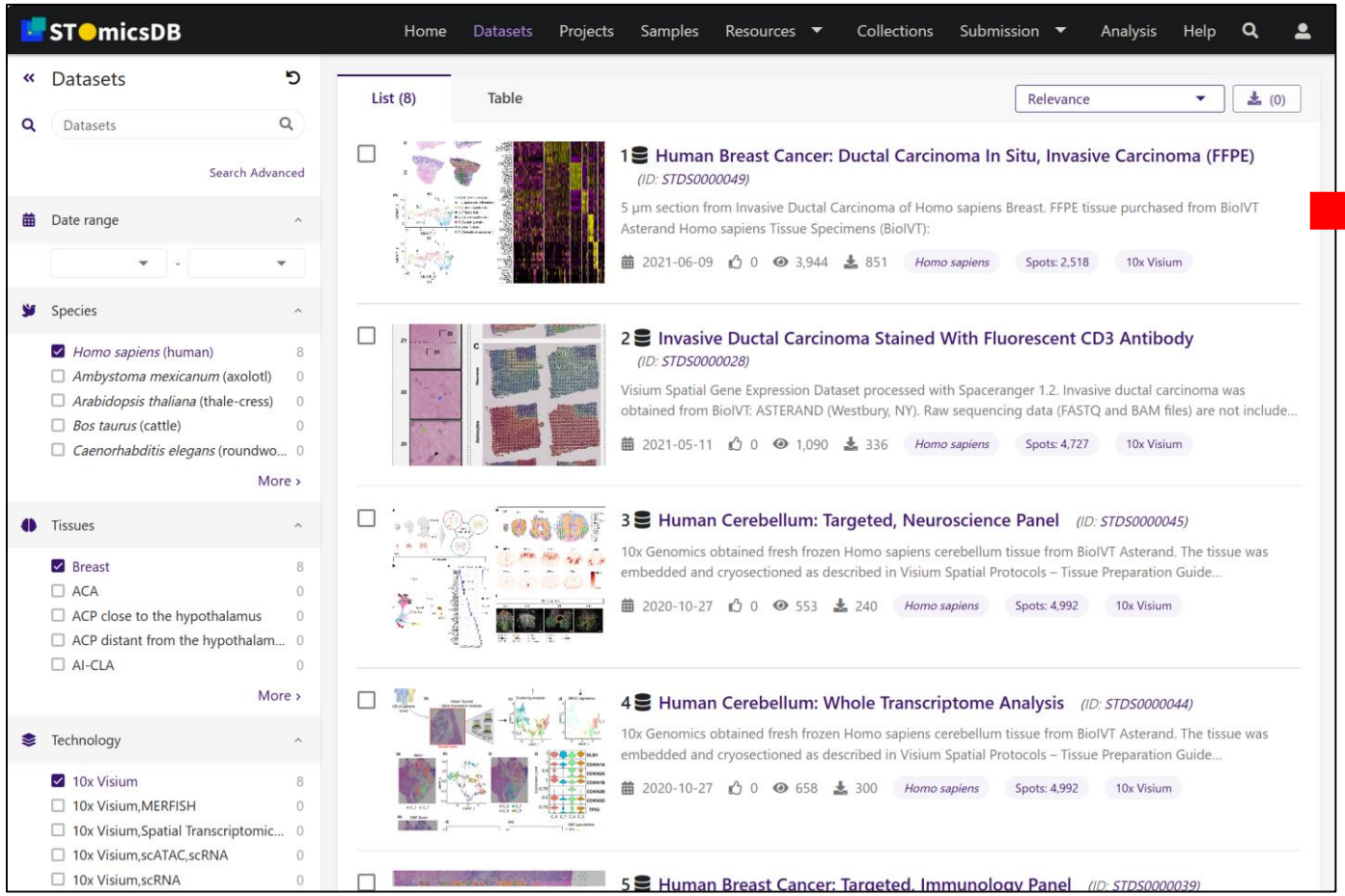
Runs

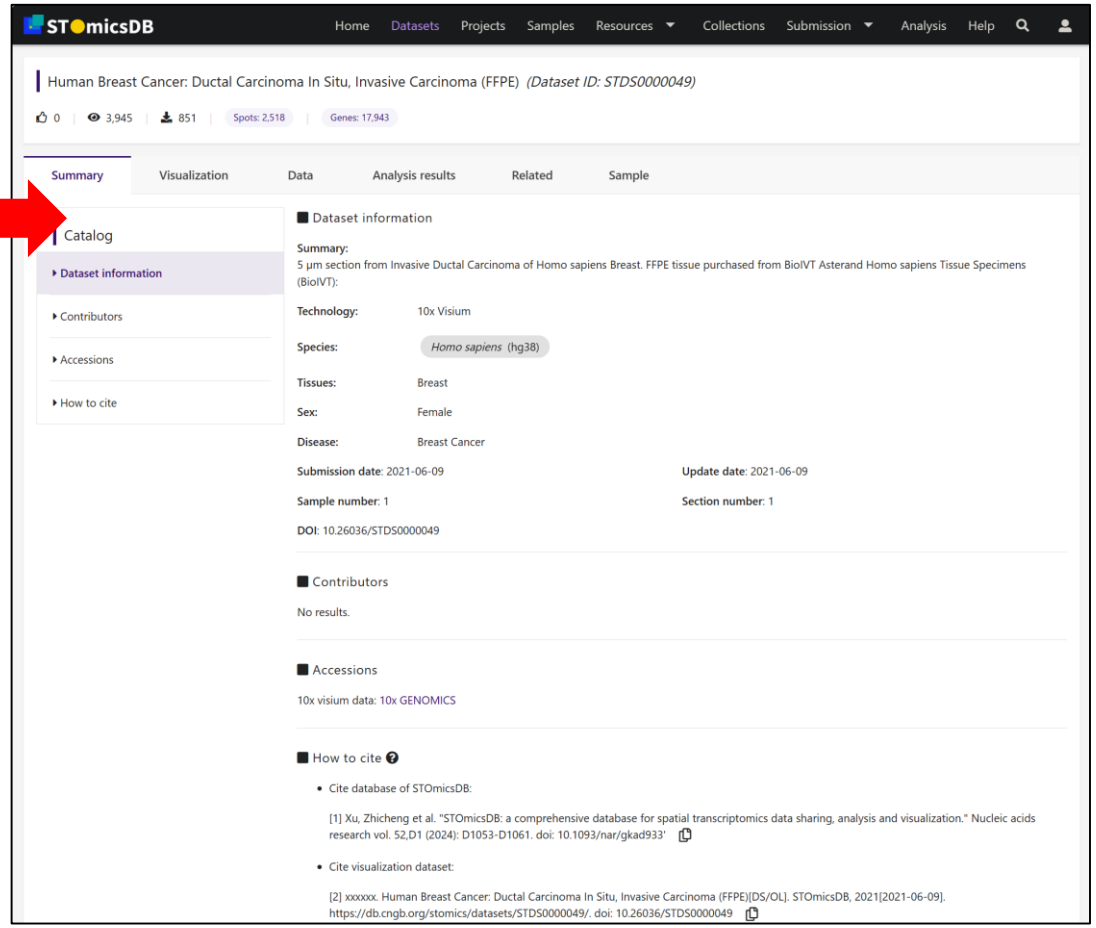
Run ID	Platform	Library layout	Organism	Experiment ID	Sample ID	Operation
CNR0495629	DNBSEQ-T7	paired	Mus musculus	CNX0412126	CNS0513001	Download data (183.0GB)
CNR0495630	DNBSEQ-T7	paired	Mus musculus	CNX0412126	CNS0513001	Download data (192.8GB)
CNR0495631	DNBSEQ-T7	paired	Mus musculus	CNX0412126	CNS0513001	Download data (185.66GB)
CNR0495632	DNBSEQ-T7	paired	Mus musculus	CNX0412126	CNS0513001	Download data (220.12GB)
CNR0495633	DNBSEQ-T7	paired	Mus musculus	CNX0412126	CNS0513001	Download data (189.58GB)

STOmicsDB – Visium example

フィルターをセットする: Species = “human”,
Tissues = “Breast”, Technology = “10x Visium”

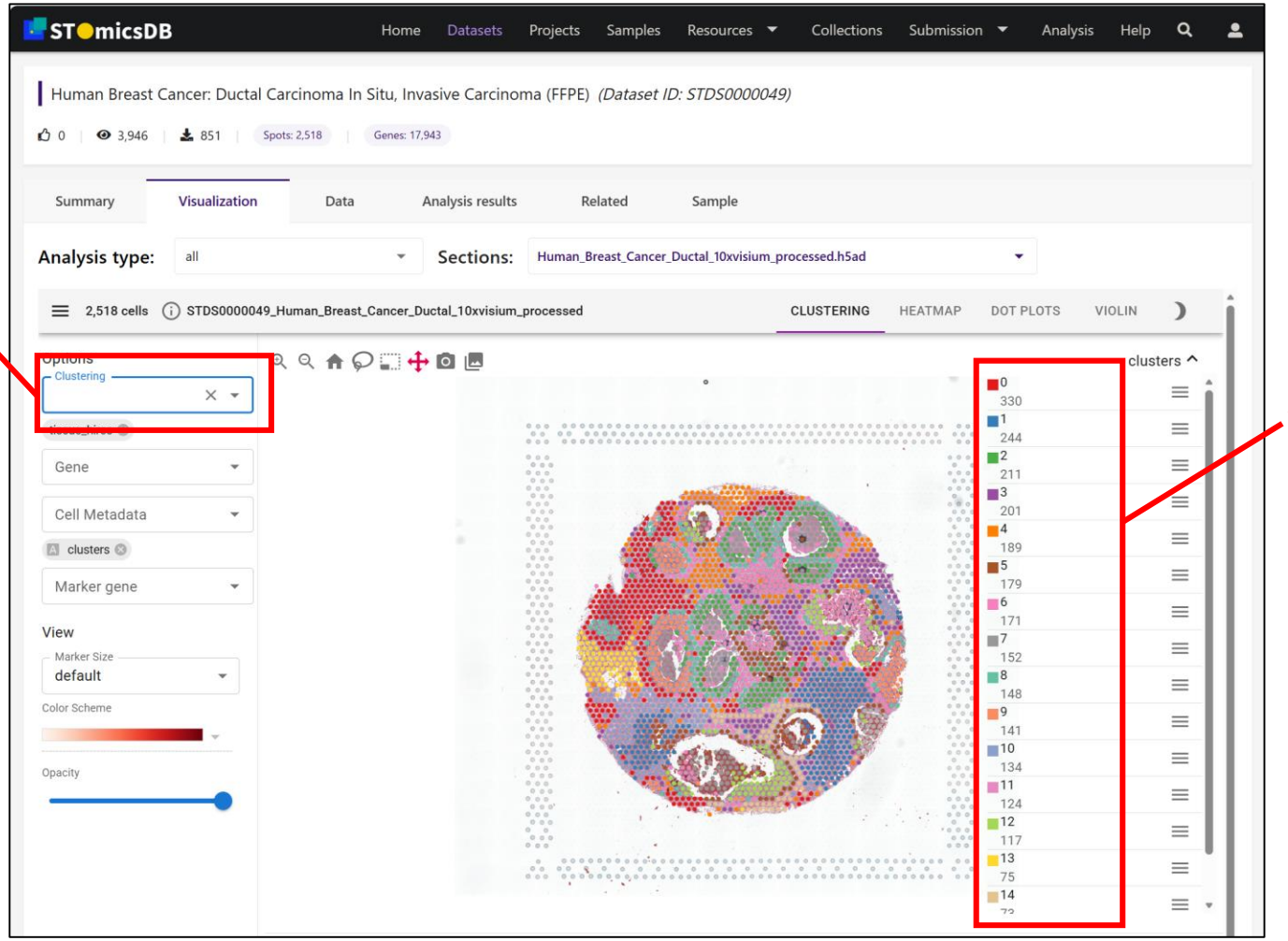
左のリストから一つ選ぶ





STOmicsDB – Visualization (Visium) (1)

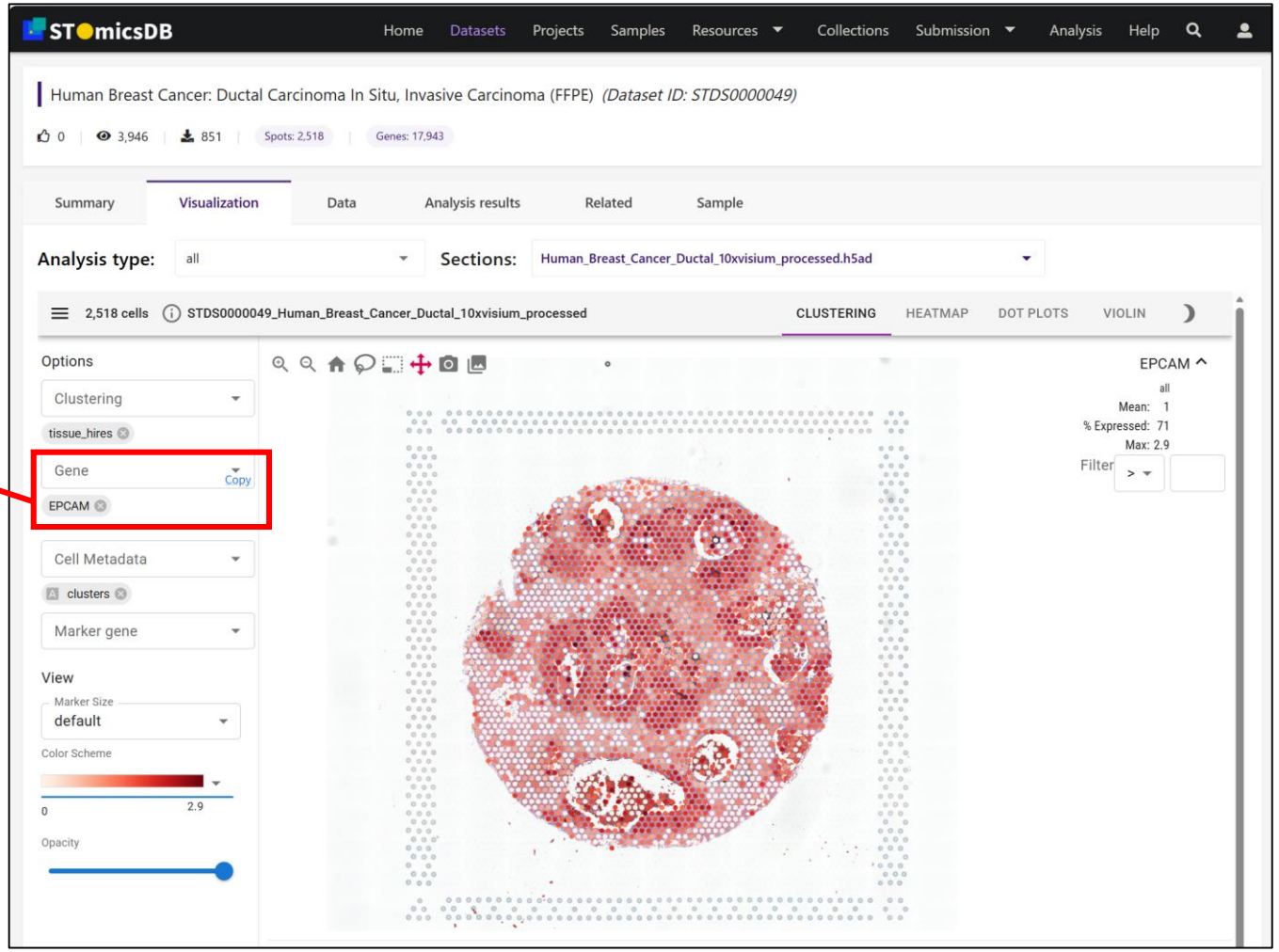
右側の画像の表示方法の選択
- spatial
- tissue-hires
- umap



クラスター

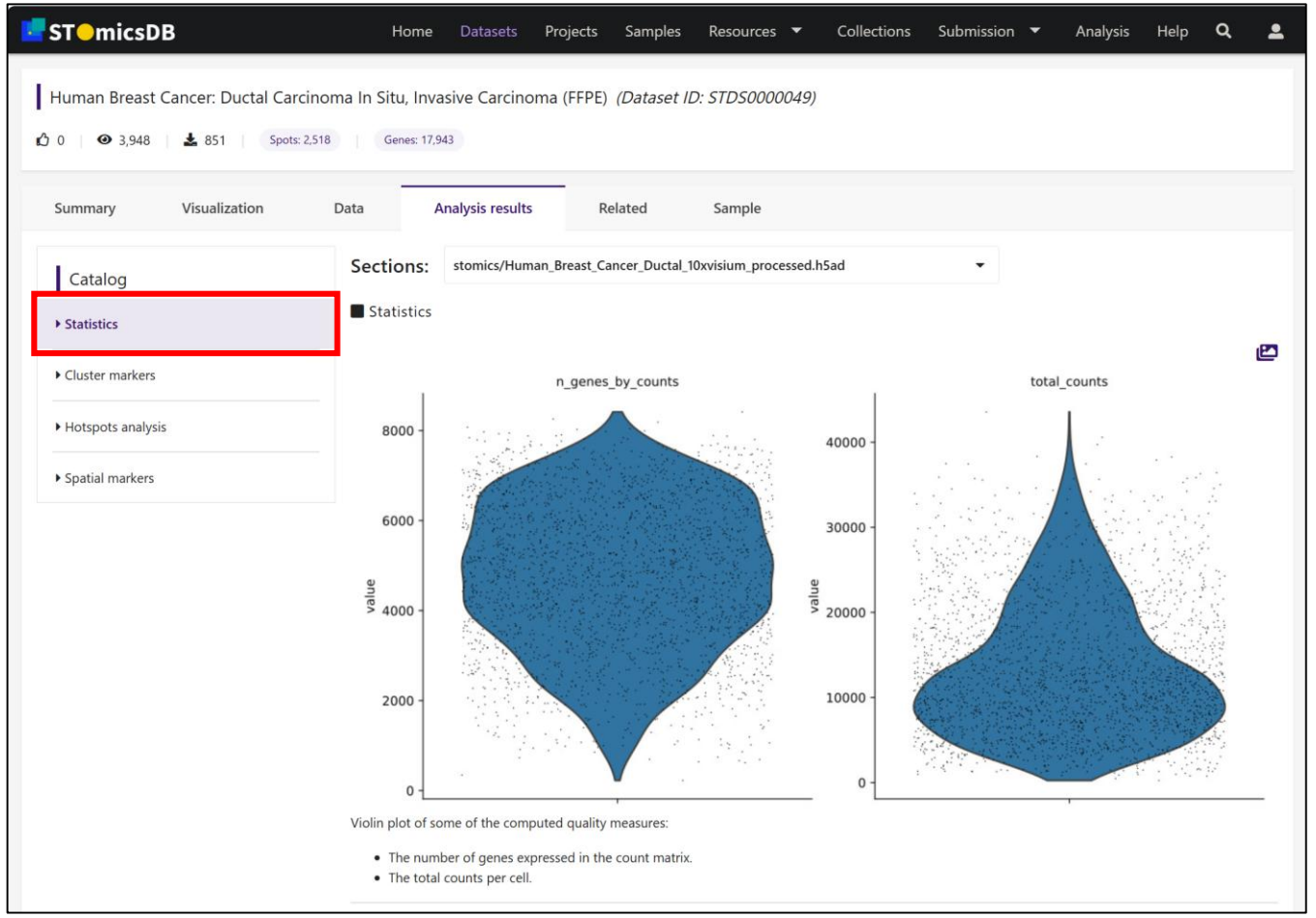
STOmicsDB – Visualization (Visium) (2)

遺伝子を選び、遺伝子発現パターンを表示させる



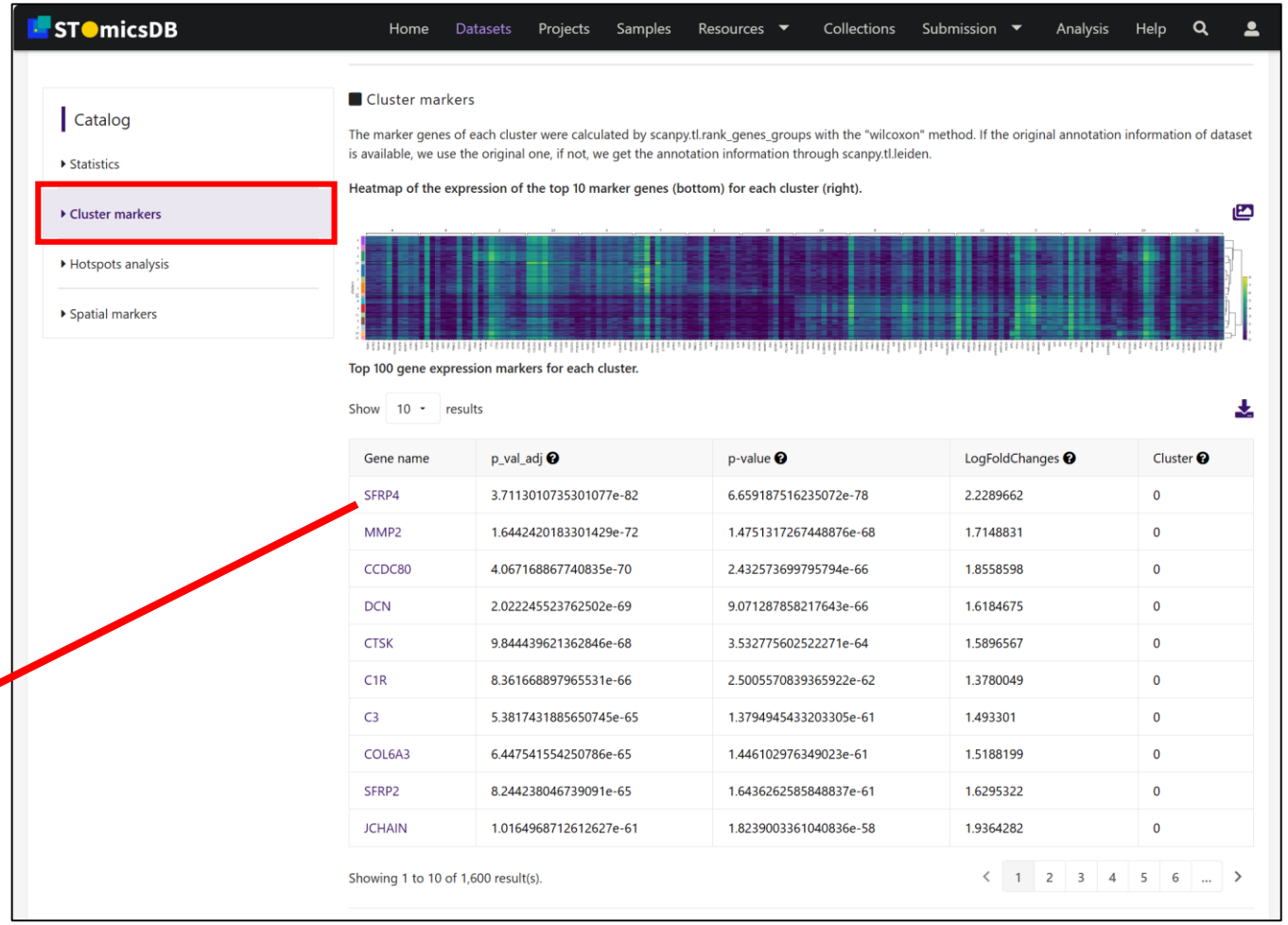
STOmicsDB – Analysis results (Visium) (1)

Note: プラットフォームによって利用できるツールが異なる

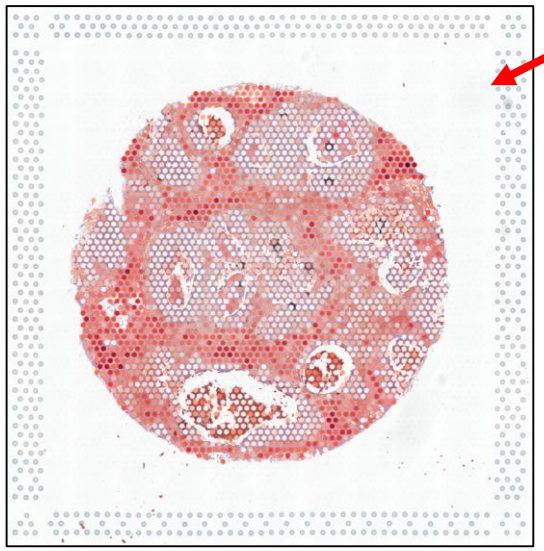


STOmicsDB – Analysis results (Visium) (2)

Note: プラットフォームによって利用できるツールが異なる

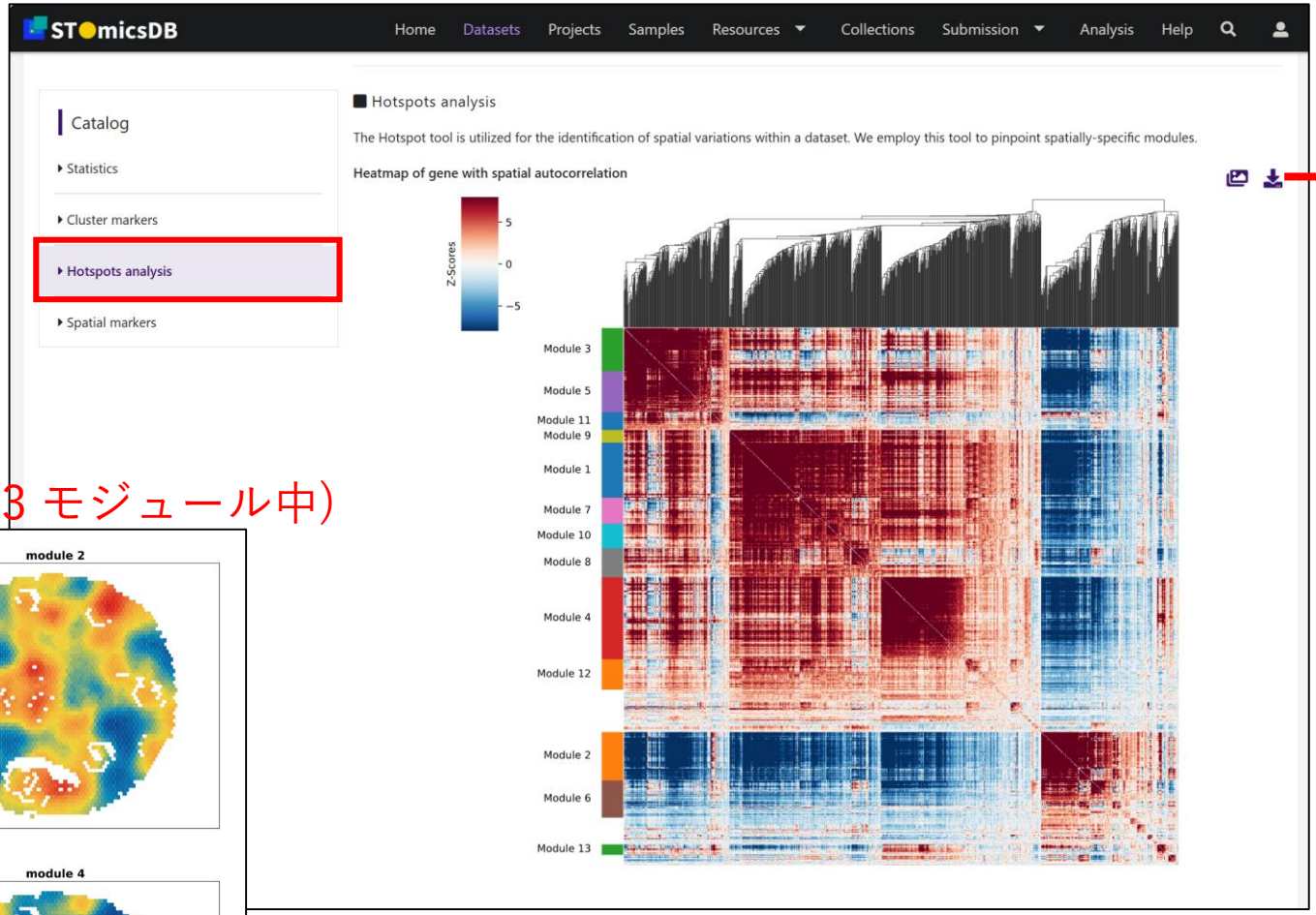


Geneをクリックすると、発現パターンのページにリンクする.



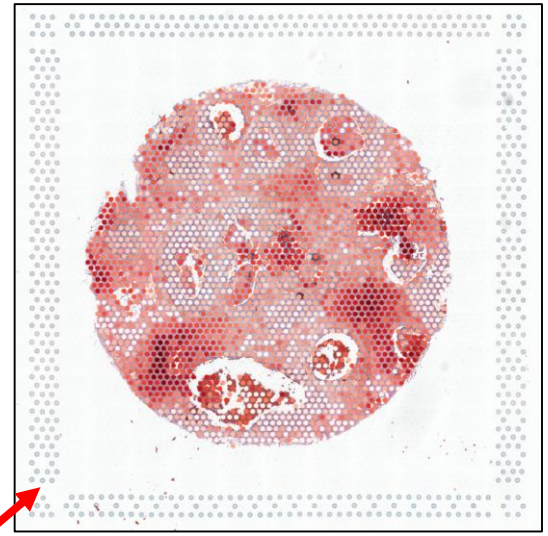
STOmicsDB – Analysis results (Visium) (3)

Note: プラットフォームによって利用できるツールが異なる

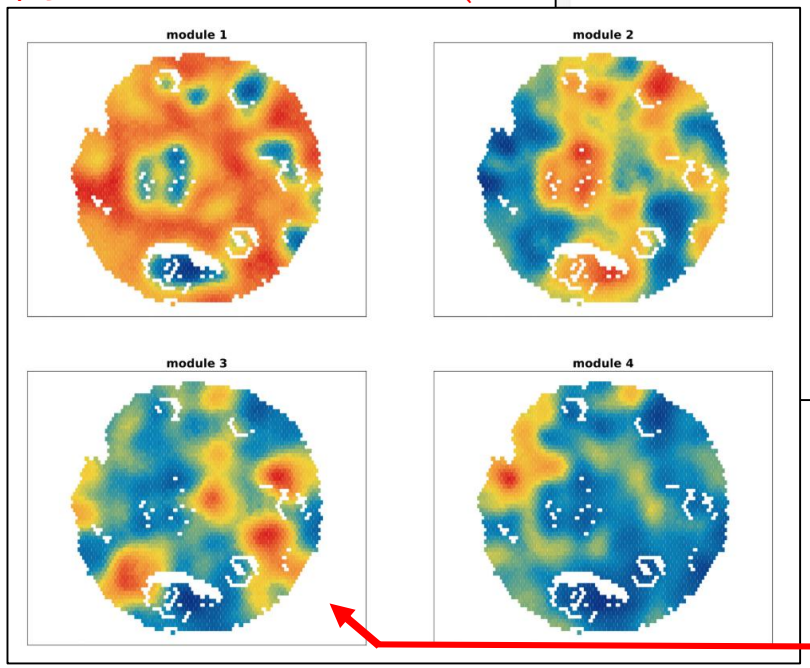


ダウンロード（約1000個の遺伝子リスト）
例としてmodule3に属するAPOC1を見てみると

APOC1発現パターン



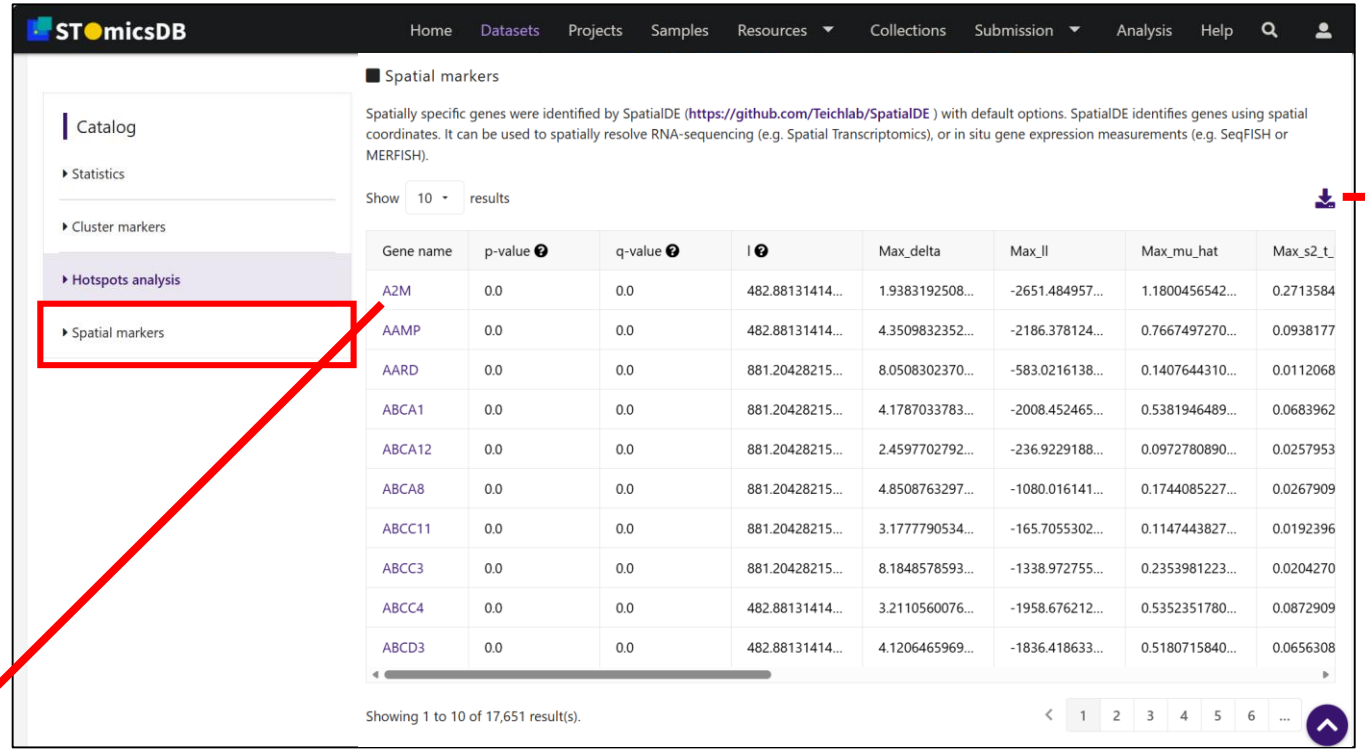
例: モジュール1 ~ 4 (全13 モジュール中)



発現パターンはモジュール3と似ている事がわかる

STOmicsDB – Analysis results (Visium) (4)

Note: プラットフォームによって利用できるツールが異なる



ダウンロード（17651個の遺伝子リスト）

遺伝子をクリックすると発現パターンページにジャンプする

SOAR

- 3,461 サンプル
- 13 種
- 19 テクノロジー

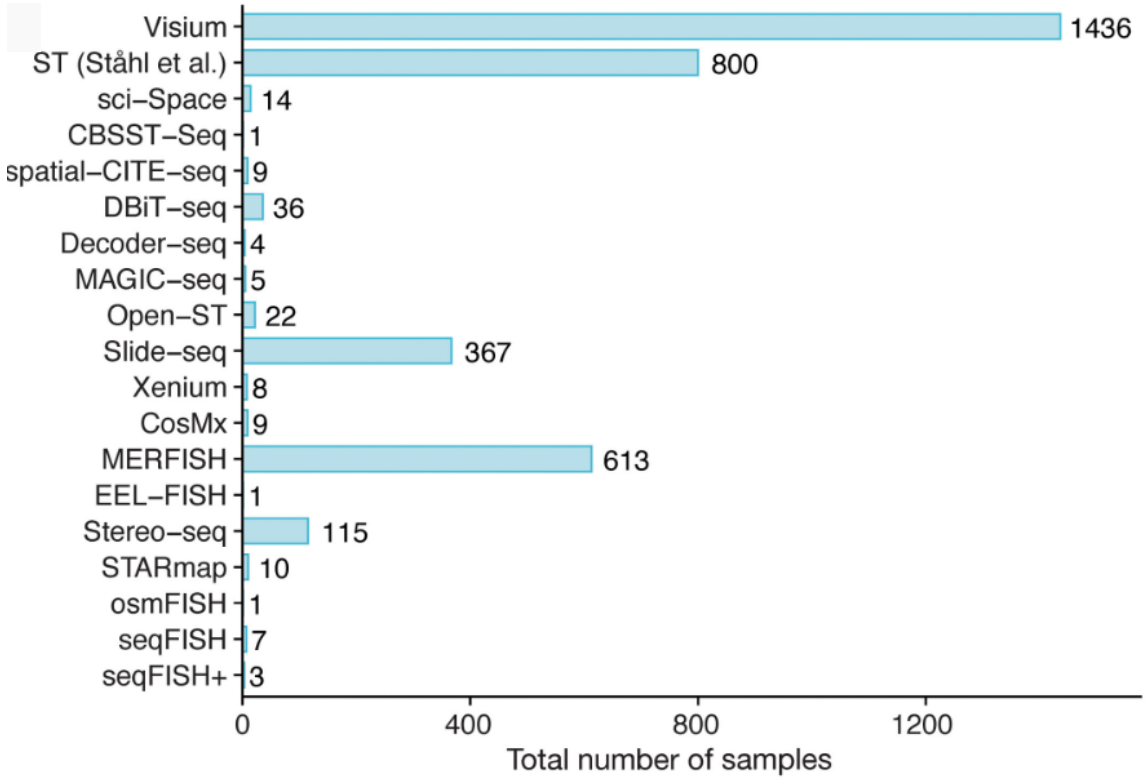


Figure from Li *et al.*, Scientific Advances, 2025

SOAR – Data Browser (Visium) (1)

SOAR | Data Browser | Gene & Cell Analysis | Drug Discovery | Download | Help

Data Browser

- Click on a Sample number to visualize spatial gene expression and view its spatial variability analysis results
- The datasets are categorized into different research topics (the Topic column), and sample-wise conditions are recorded in the Condition column

Show 100 entries Search:

Dataset	Sample	Technology	Species	Organ	Tissue	Condition	Topic	#Spots	#Genes	Publication.year	Source
All	All	All	All	All	All	All	All	All	All	All	All
10x_demo_GE_breast_cancer_sec1	1	Visium	Human	Breast	Breast Cancer	Malignant	Pathological	3797	21129	2020	10x website
10x_demo_FFPE_brain_IF	1	Visium	Mouse	Brain		Normal	Normal	2407	15180	2020	10x website
10x_demo_FFPE_brain	1	Visium	Mouse	Brain		Normal	Normal	2258	15202	2020	10x website
10x_demo_FFPE_breast_cancer	1	Visium	Human	Breast	Breast Cancer	Malignant	Pathological	2516	15946	2020	10x website
10x_demo_FFPE_kidney	1	Visium	Mouse	Kidney		Normal	Normal	3106	15588	2020	10x website
10x_demo_FFPE_prostate_adj_norm	1	Visium	Human	Prostate	Normal	Adjacent Normal of Cancer	Normal	3459	15497	2020	10x website
10x_demo_FFPE_prostate_cancer	1	Visium	Human	Prostate	Prostate Cancer	Malignant	Pathological	4369	15574	2020	10x website
10x_demo_FFPE_prostate	1	Visium	Human	Prostate		Normal	Normal	2542	14840	2020	10x website
10x_demo_GE_breast_cancer_sec2	1	Visium	Human	Breast	Breast Cancer	Malignant	Pathological	3983	21187	2020	10x website
10x_demo_GE_coronal	1	Visium	Mouse	Brain		Normal	Normal	2697	18768	2020	10x website

Showing 1 to 10 of 3,461 entries Previous 1 2 3 4 5 ... 347 Next

Spatial Gene Expression | Spatial Clustering | Clustering (non-spatial) | Cellular Deconvolution/Annotation

Spatial gene expression in 10x_demo_FFPE_brain_IF.1

Data Download

JPG PDF

Gene

A2m

*Type to search for a gene

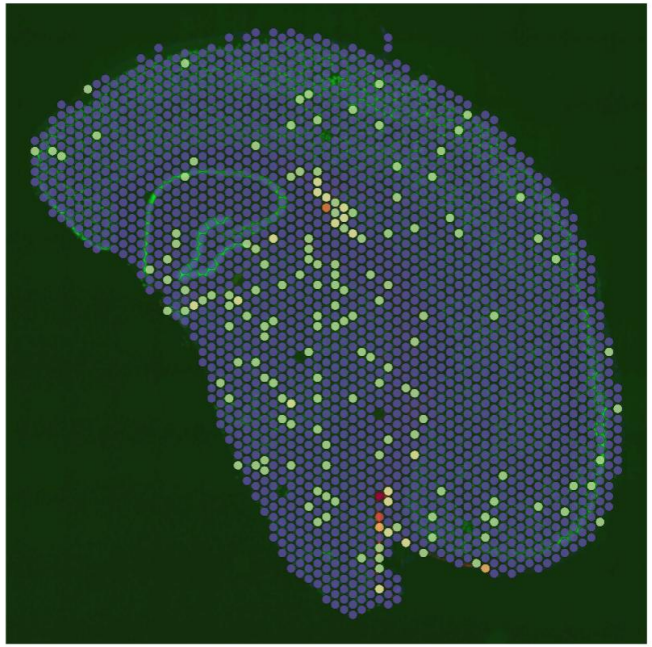
Normalized expression

0.0 0.5 1.0 1.5 2.0

Adjust the opacity of spots

0 0.8 1

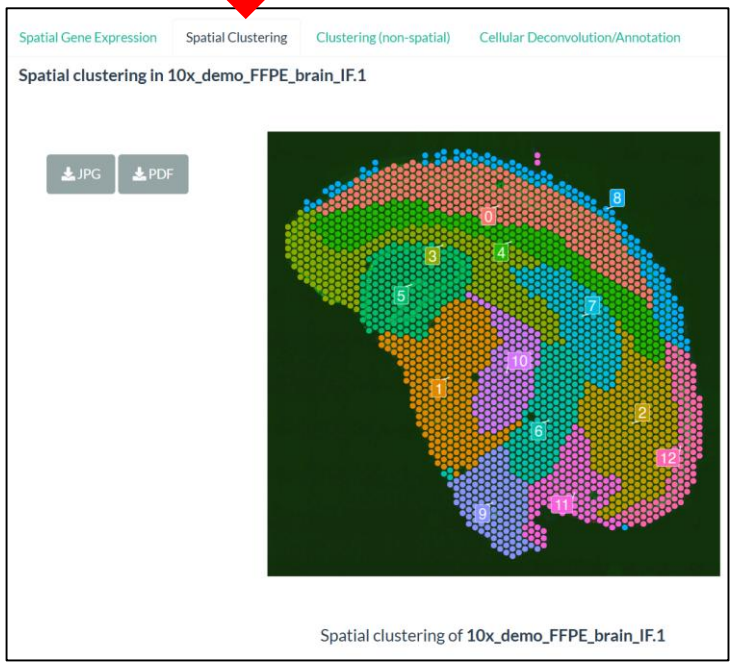
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1



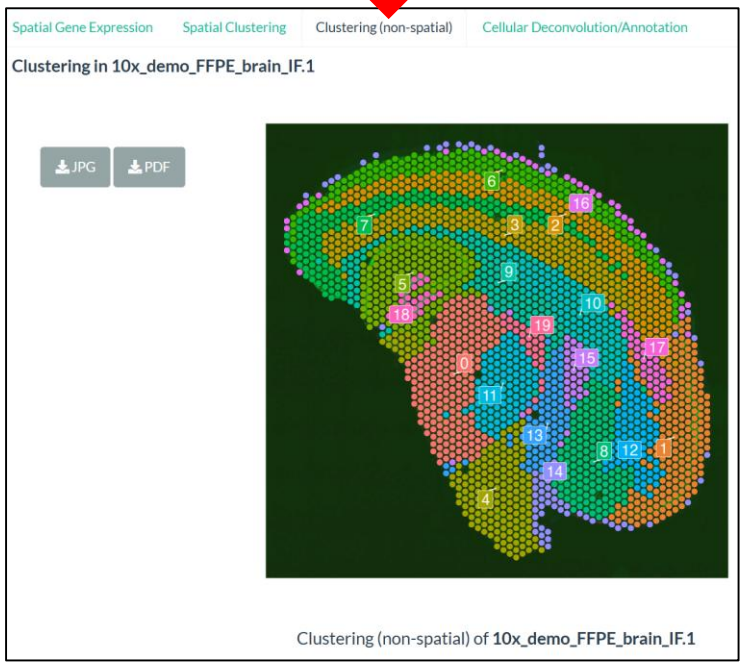
Normalized spatial gene expression of **A2m** in **10x_demo_FFPE_brain_IF.1**

SOAR – Data Browser (Visium) (2)

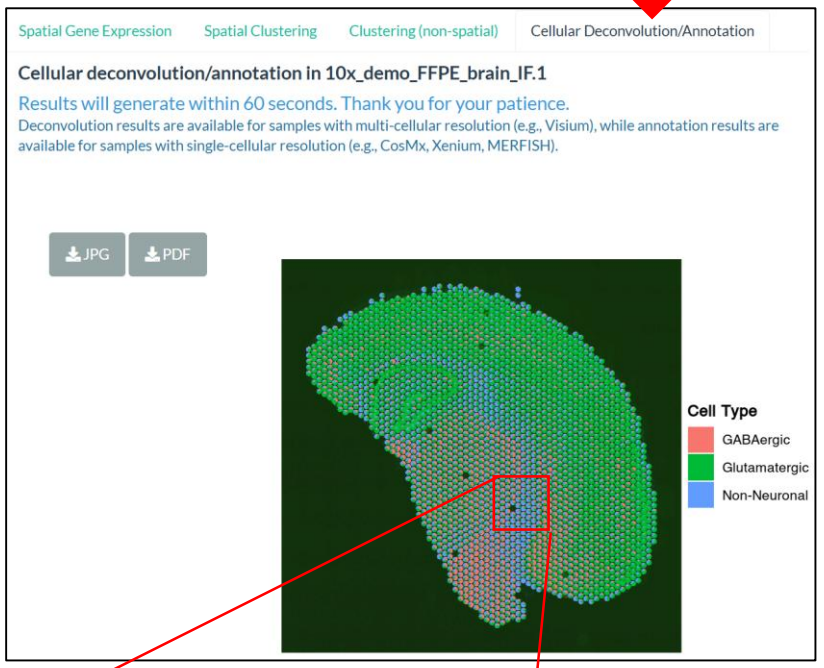
STAGATEを用いた
spatial domains

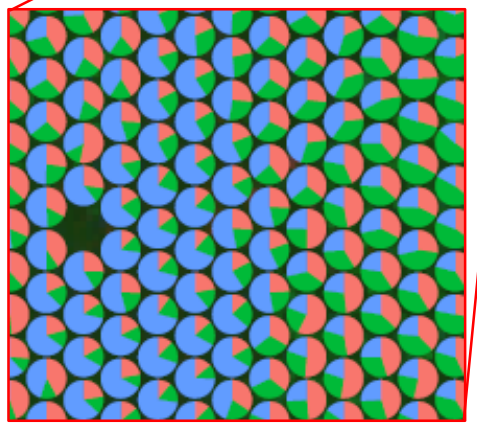


遺伝子発現のみに基づく
クラスタリング



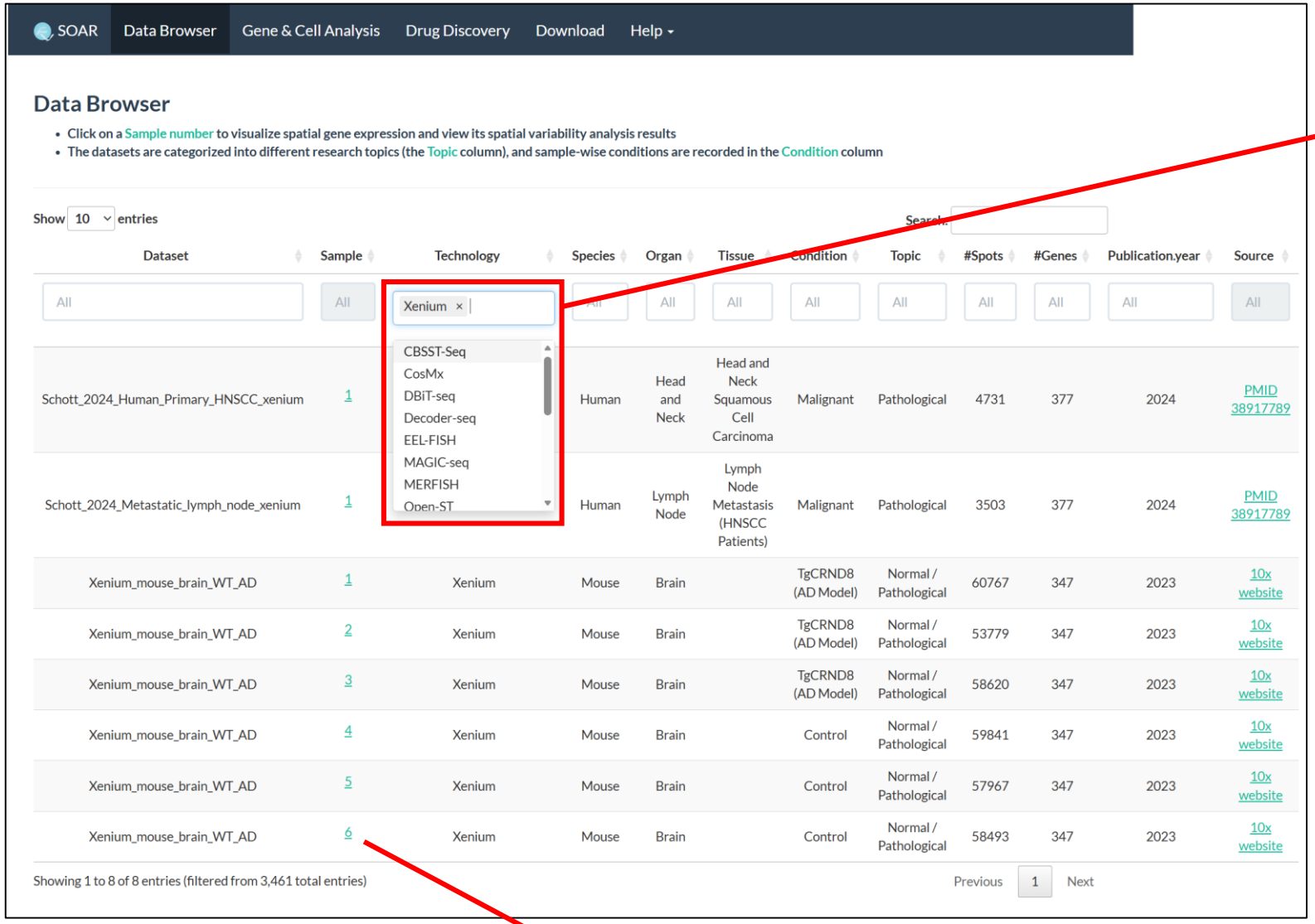
細胞種予測





予測された細胞種の割合がスポット
にパイチャートで表示される

SOAR – Data Browser (Xenium) (1)



Dataset	Sample	Technology	Species	Organ	Tissue	Condition	Topic	#Spots	#Genes	Publication.year	Source
Schott_2024_Human_Primary_HNSCC_xenium	1	Xenium	Human	Head and Neck	Head and Neck Squamous Cell Carcinoma	Malignant	Pathological	4731	377	2024	PMID 38917789
Schott_2024_Metastatic_lymph_node_xenium	1	Xenium	Human	Lymph Node	Lymph Node Metastasis (HNSCC Patients)	Malignant	Pathological	3503	377	2024	PMID 38917789
Xenium_mouse_brain_WT_AD	1	Xenium	Mouse	Brain		TgCRND8 (AD Model)	Normal / Pathological	60767	347	2023	10x website
Xenium_mouse_brain_WT_AD	2	Xenium	Mouse	Brain		TgCRND8 (AD Model)	Normal / Pathological	53779	347	2023	10x website
Xenium_mouse_brain_WT_AD	3	Xenium	Mouse	Brain		TgCRND8 (AD Model)	Normal / Pathological	58620	347	2023	10x website
Xenium_mouse_brain_WT_AD	4	Xenium	Mouse	Brain		Control	Normal / Pathological	59841	347	2023	10x website
Xenium_mouse_brain_WT_AD	5	Xenium	Mouse	Brain		Control	Normal / Pathological	57967	347	2023	10x website
Xenium_mouse_brain_WT_AD	6	Xenium	Mouse	Brain		Control	Normal / Pathological	58493	347	2023	10x website

“Xenium”を選ぶ

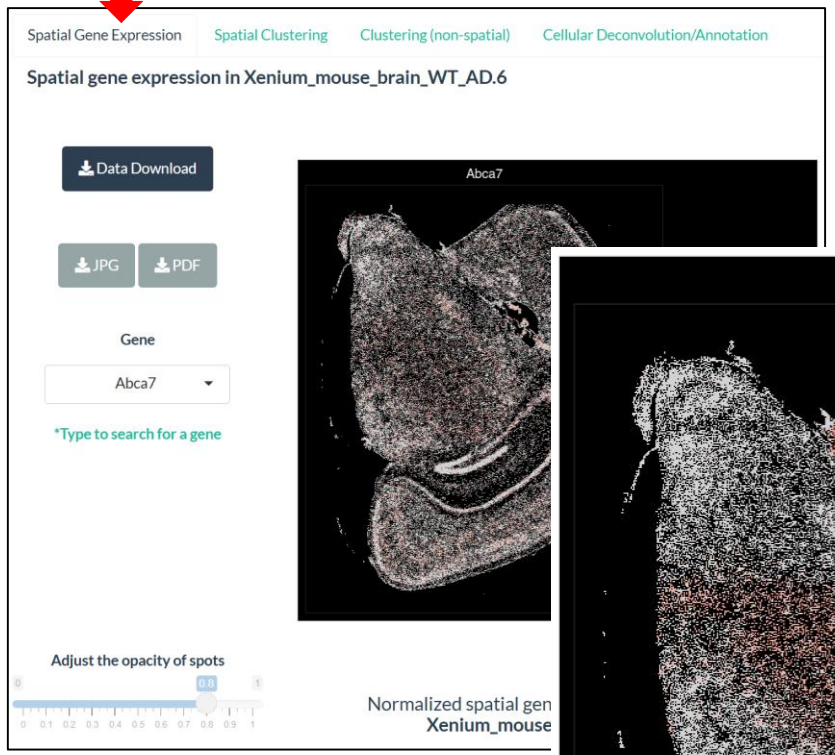
SOAR にはXeniumが
8サンプル集められている

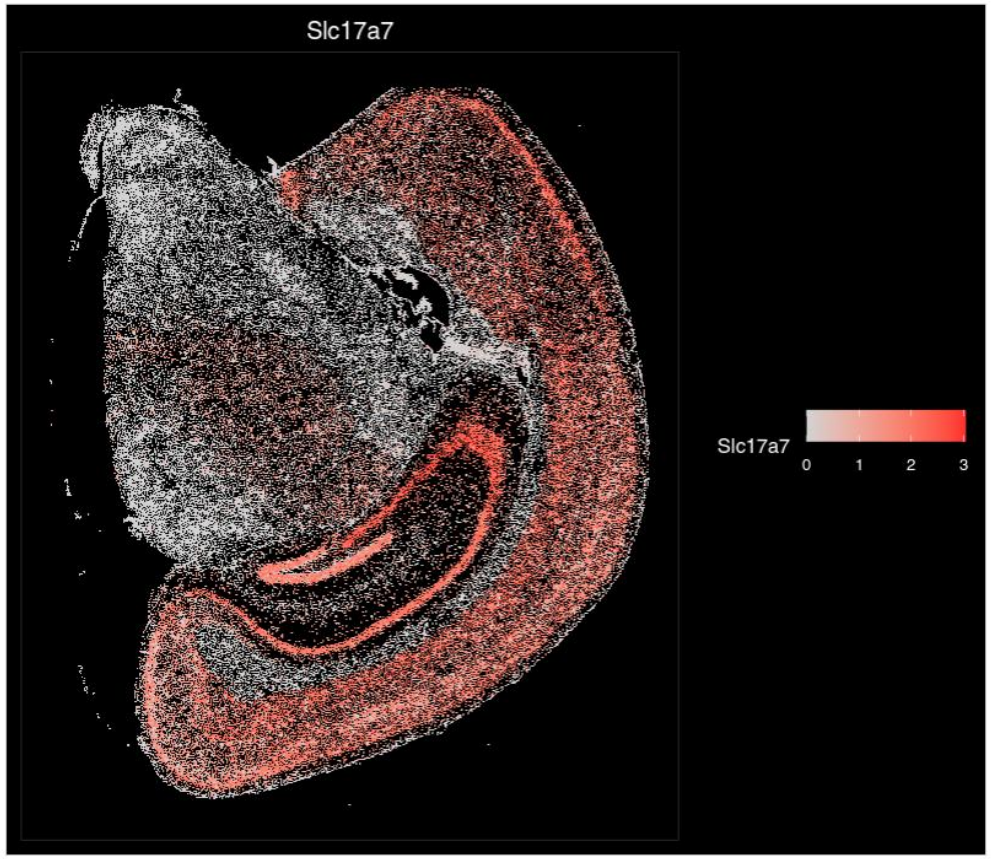
このマウスの脳サンプルをクリックすると

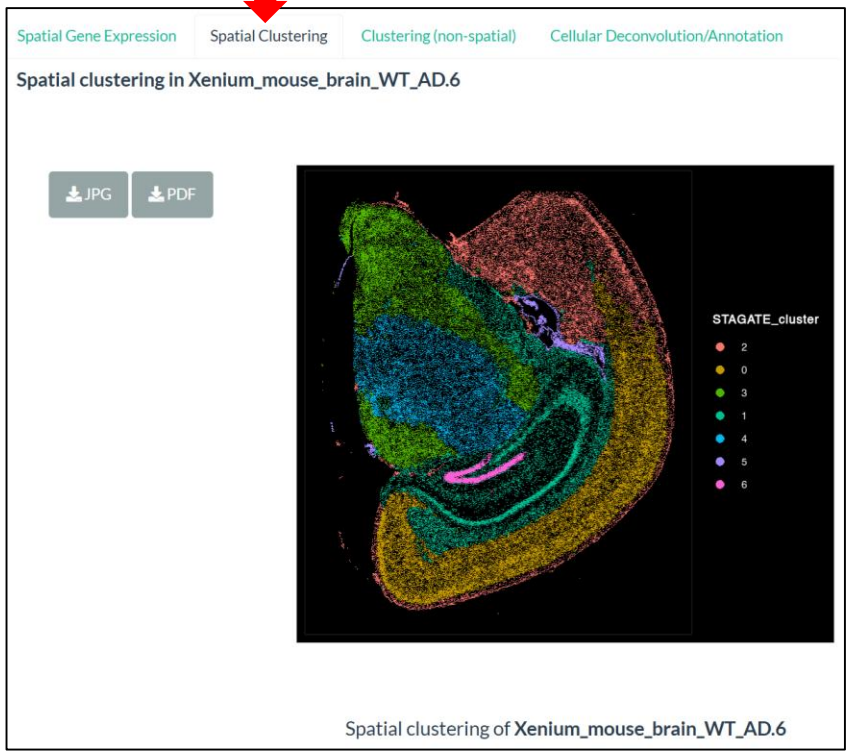
SOAR – Data Browser (Xenium) (2)

遺伝子発現

STAGATE を用いた
空間的クラスタリング







既存のデータベースに共通する特徴

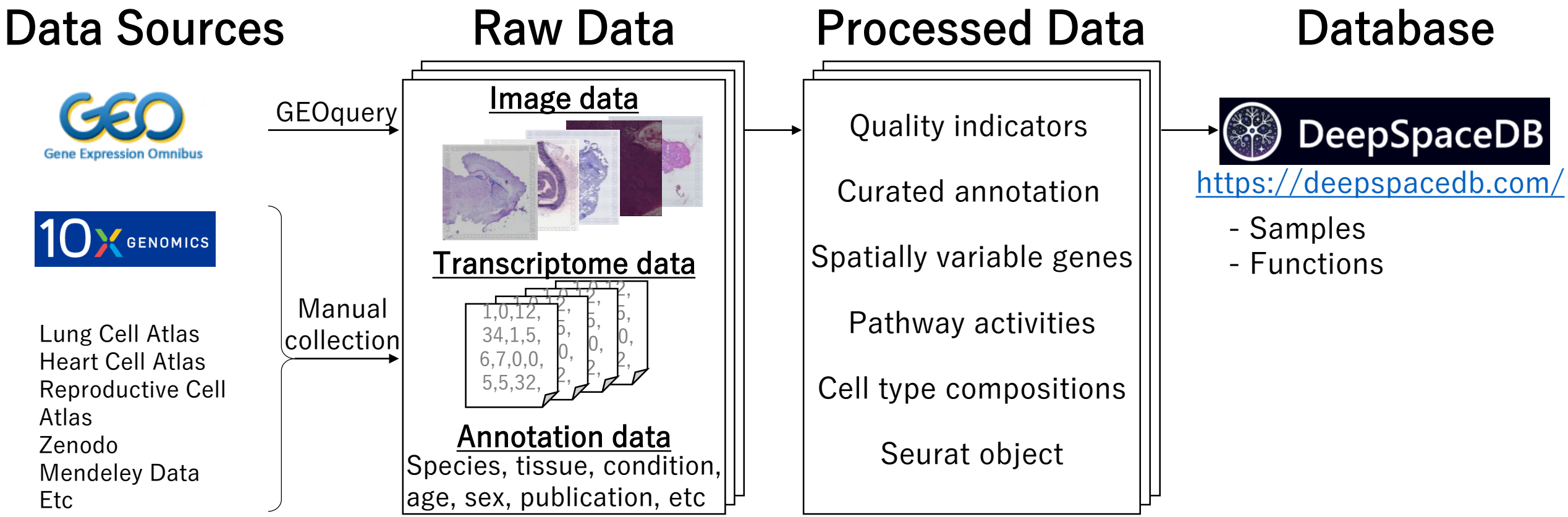
- 優れた点
  - 多彩なプラットフォーム
  - サンプル数が多い

- だが問題が多い
  - 低品質なデータが多く含まれている
  - 解析機能が限られている
  - 直観的な操作ではない上に反応がしばしば遅い

概要

- Rを使った空間トランスクリプトミクスデータ解析
  - Visium example
  - Xenium example
- 既存の空間トランスクリプトミクスデータベースの紹介
- DeepSpaceDBの紹介
  - Visium interface
  - Xenium interface
- まとめ

ワークフロー


Data Sources
Raw Data
Processed Data
Database
Gene Expression Omnibus
GEOquery
10x GENOMICS
Manual collection
Lung Cell Atlas
Heart Cell Atlas
Reproductive Cell Atlas
Zenodo
Mendeley Data
Etc
Image data
Transcriptome data
1,0,12,
34,1,5,
6,7,0,0,
5,5,32,
Annotation data
Species, tissue, condition, age, sex, publication, etc
Quality indicators
Curated annotation
Spatially variable genes
Pathway activities
Cell type compositions
Seurat object
DeepSpaceDB
https://deepspacedb.com/
- Samples
- Functions

アノテーションデータの収集例

NCBI

Gene Expression Omnibus

HOME | SEARCH | SITE MAP | GEO Publications | FAQ | MIAME | Email GEO

NCBI > GEO > **Accession Display** | Not logged in | Login

Scope: Self | Format: HTML | Amount: Quick | GEO accession: GSE225766 | GO

Series GSE225766 Query DataSets for GSE225766

Status	Public on Jun 16, 2023
Title	A cellular and molecular spatial atlas of dystrophic muscle
Organism	Mus musculus
Experiment type	Expression profiling by high throughput sequencing Other
Summary	Asynchronous skeletal muscle degeneration/regeneration is a hallmark feature of Duchenne muscular dystrophy (DMD); however, traditional -omics technologies that lack spatial context make it difficult to study the biological mechanisms of how asynchronous regeneration contributes to disease progression. Here, using the severely dystrophic D2-mdx mouse model, we generated a high-resolution cellular and molecular spatial atlas of dystrophic muscle by integrating spatial transcriptomics and single-cell RNAseq datasets.
Overall design	Gene expression profiling analysis of spatial transcriptome data taken from wildtype (DBA2/J) and dystrophic (D2-mdx) mouse gastrocnemius/plantaris muscle
Contributor(s)	Stec M, Su Q, Adler C, Zhang L, Ni M, Wei Y, Atwal G, Halasz G
Citation(s)	Stec MJ, Su Q, Adler C, Zhang L et al. A cellular and molecular spatial atlas of dystrophic muscle. *Proc Natl Acad Sci U S A* 2023 Jul 18;120(29):e2221249120. PMID: 37410813
Submission date	Feb 21, 2023
Last update date	Sep 18, 2023
Contact name	Michael Stec
Organization name	Regeneron Pharmaceuticals
Street address	777 Old Saw Mill River Rd
City	Tarrytown
State/province	NY
ZIP/Postal code	10591
Country	USA
Platforms (1)	GPL19057 Illumina NextSeq 500 (Mus musculus)
Samples (11)	GSM7055901 WT1
Less...	GSM7055902 WT2
	GSM7055903 MDX1
	GSM7055904 MDX2
	GSM7055905 WT3
	GSM7055906 MDX3
	GSM7055907 MDX4

Sample GSM7055907 Query DataSets for GSM7055907

Status	Public on Jun 16, 2023
Title	MDX4
Sample type	SRA
Source name	Gastrocnemius
Organism	Mus musculus
Characteristics	tissue: Gastrocnemius genotype: MDX
Extracted molecule	polyA RNA
Extraction protocol	10x Genomics Visium Spatial Gene Expression for Fresh Frozen Tissue v1.0
Library strategy	RNA-Seq
Library source	transcriptomic
Library selection	cDNA
Instrument model	Illumina NextSeq 500
Data processing	Spaceranger 1.3.0 Assembly: mm10
Submission date	Feb 21, 2023
Last update date	Sep 18, 2023
Contact name	Michael Stec
Organization name	Regeneron Pharmaceuticals
Street address	777 Old Saw Mill River Rd
City	Tarrytown
State/province	NY
ZIP/Postal code	10591
Country	USA
Platform ID	GPL19057
Series (1)	GSE225766 A cellular and molecular spatial atlas of dystrophic muscle

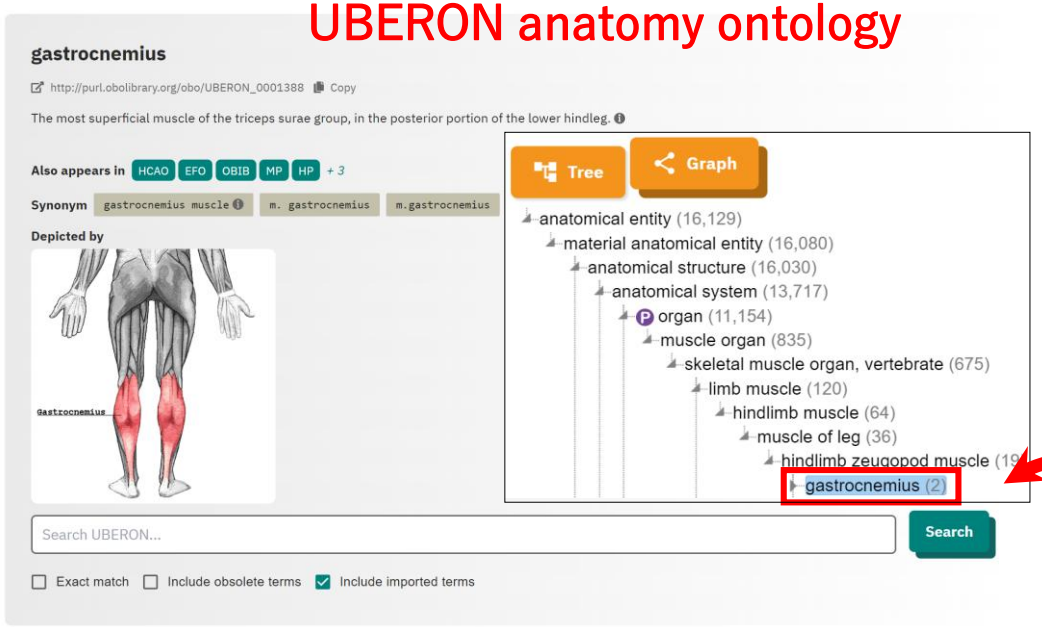
Relations

BioSample SAMN33402466

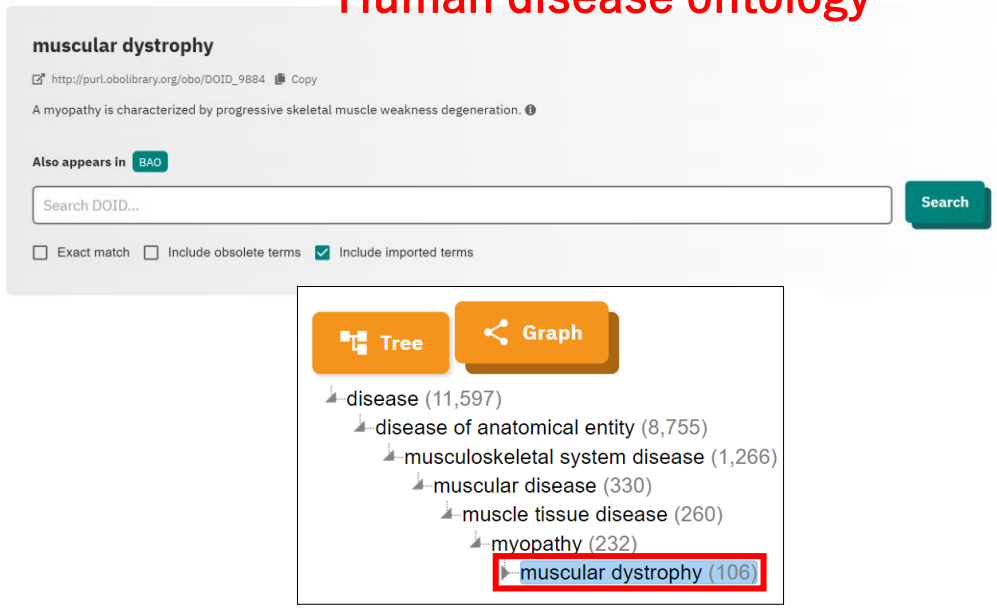
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE225766

アノテーション及びオントロジー

UBERON anatomy ontology



Human disease ontology



データビューの一例

(sample DSID000466)

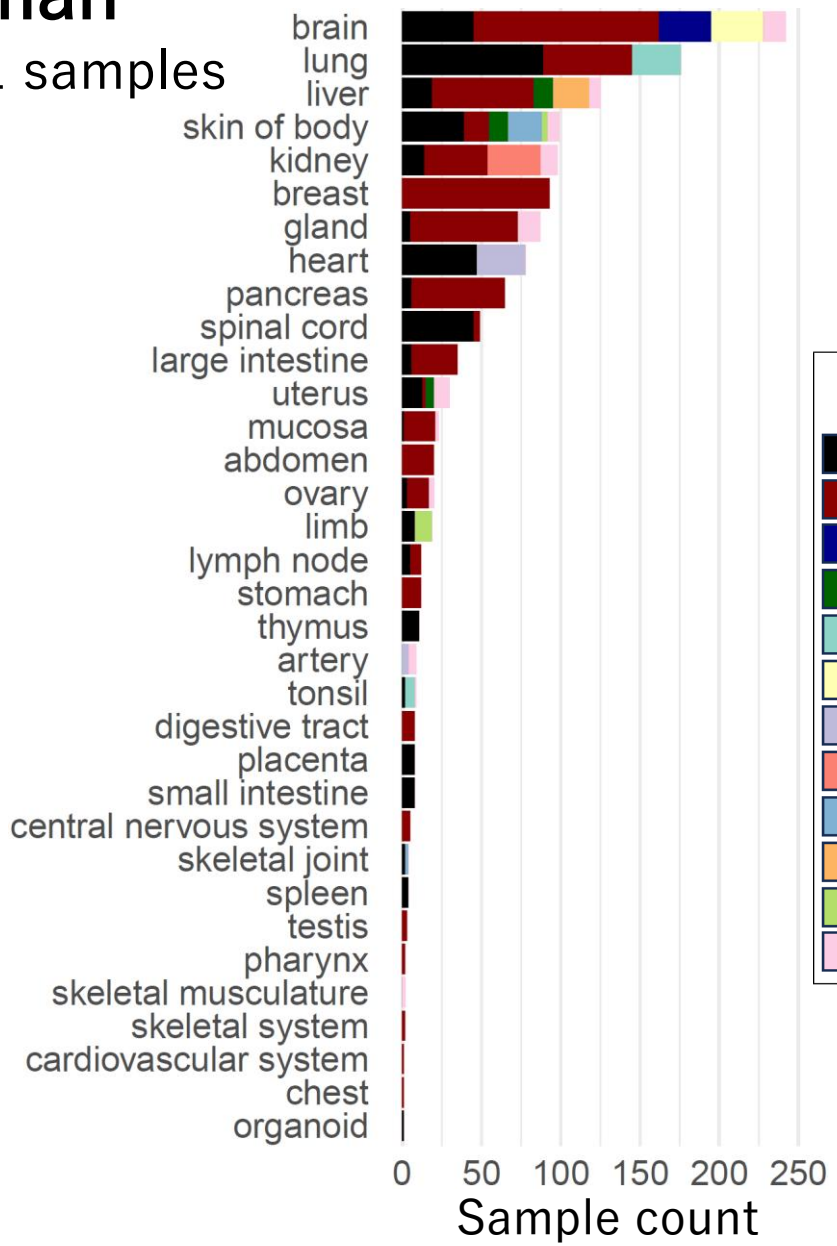
DeepSpace ID	DSID000466
Sample ID	GSM7055907
Series ID	GSE225766
Platform	Visium V1
Organism	mouse
Organ	skeletal musculature
Organ Detailed	gastrocnemius
Condition	muscle tissue disease
Condition Detailed	muscular dystrophy
Sex	male
Age or Stage	6 week
Post Mortem	--
Ethnicity or Strain	D2-mdx
Description	severely dystr
Source	GEO
Source URL	GSE225766
PMID URL:	37410813
Date Published:	2023-06-16

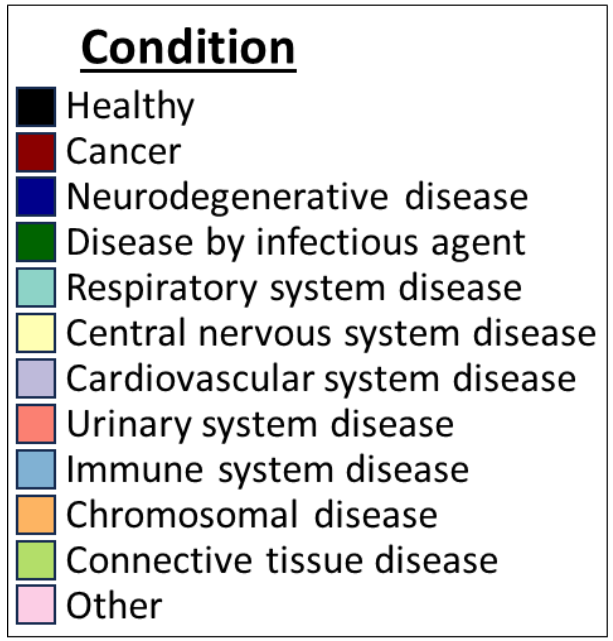
- アノテーション作業は非常に時間がかかる
- 高度な専門性が必要となる
- LLMsを使い、部分的な自動化を目指したい

総サンプル数　ヒト及びマウス (version 1.1)

Human

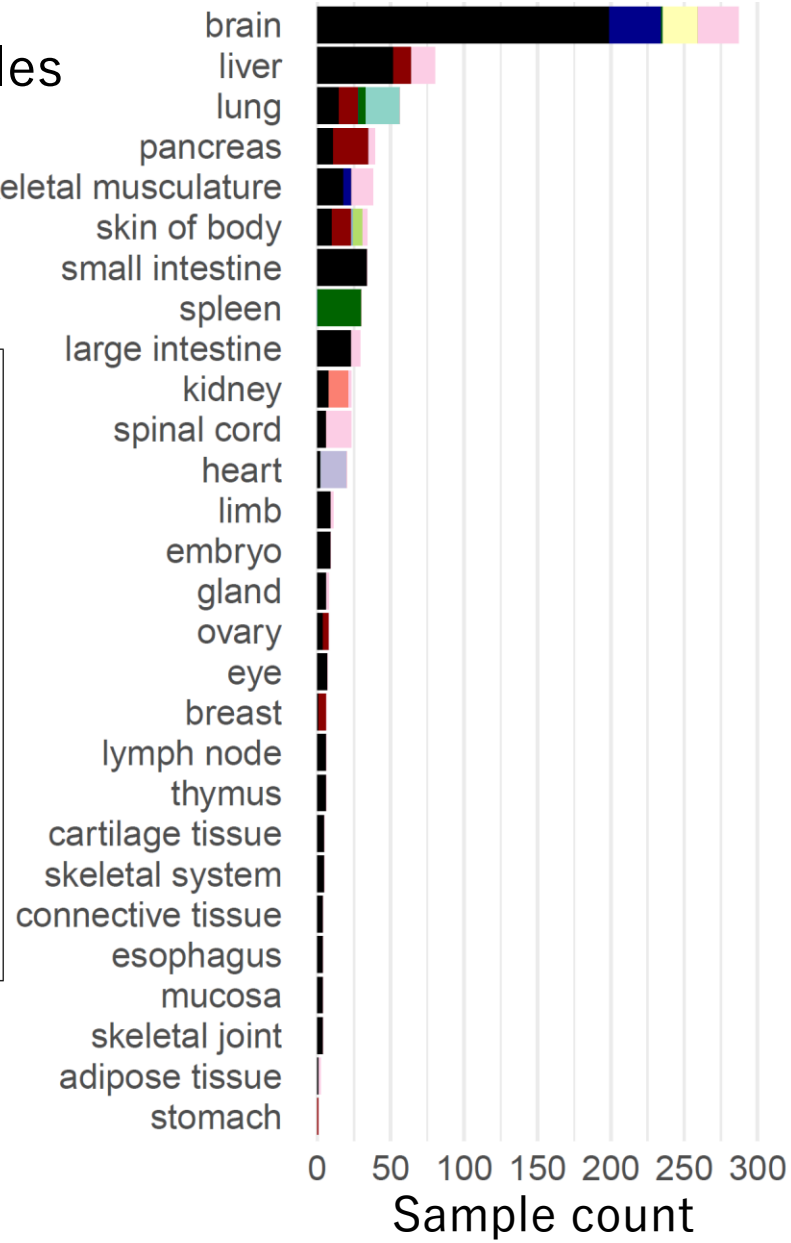
1,361 samples





Mouse

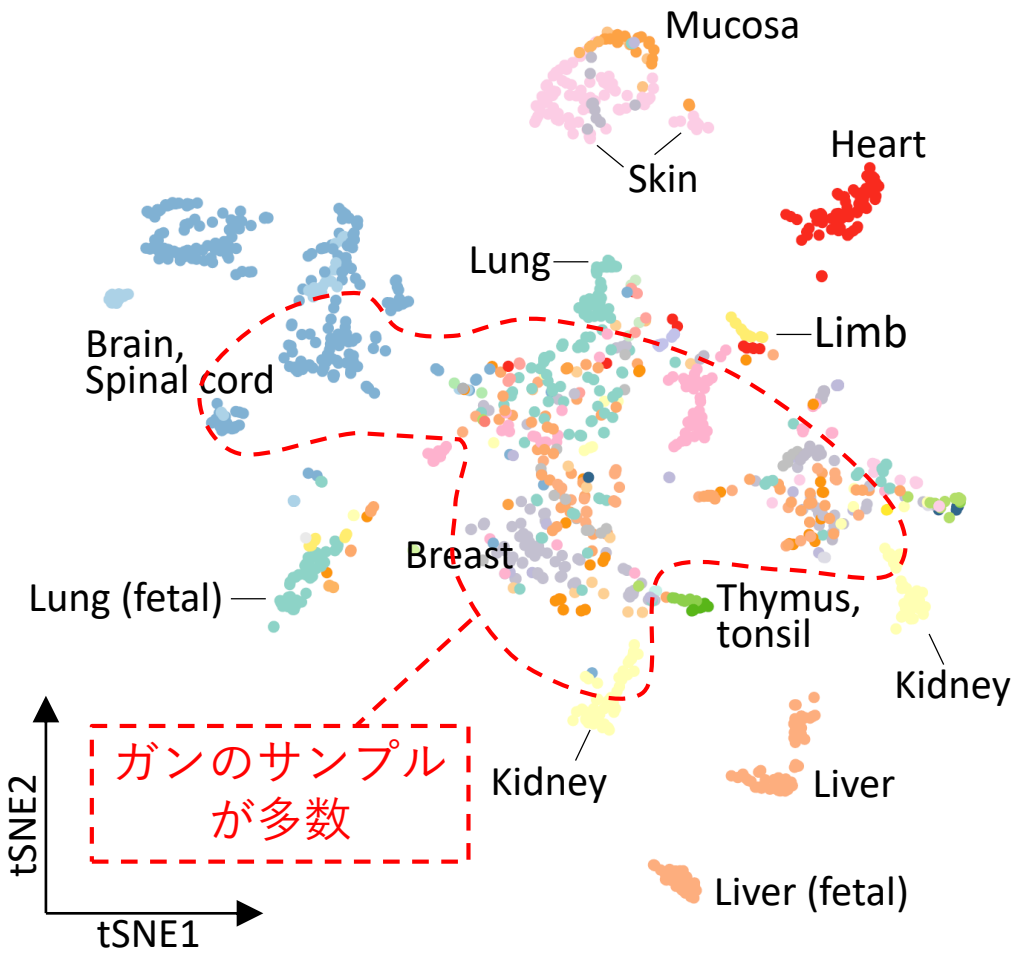
783 samples



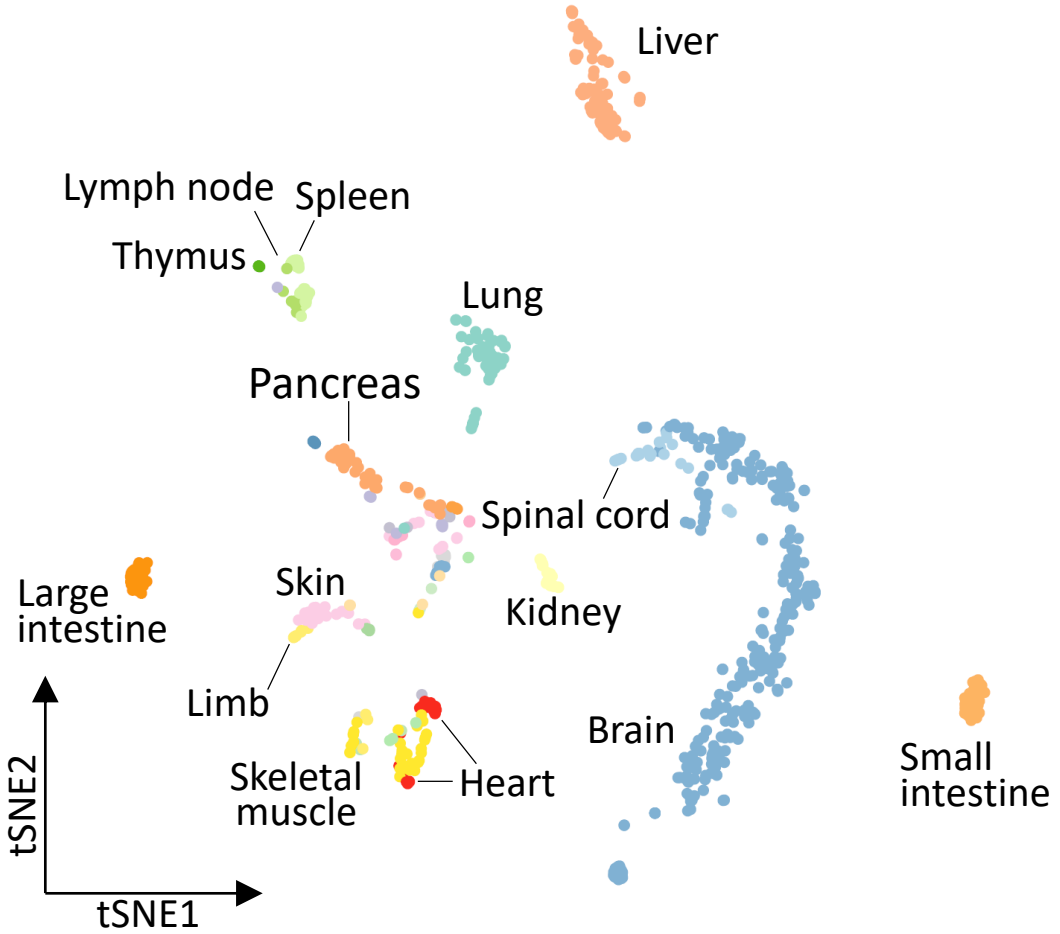
データの生物学的一貫性

- 全てのサンプルを“pseudo-bulk”として埋め込み2次元に圧縮
- 関連する組織は似た遺伝子発現パターンを示す

Human tissue slices

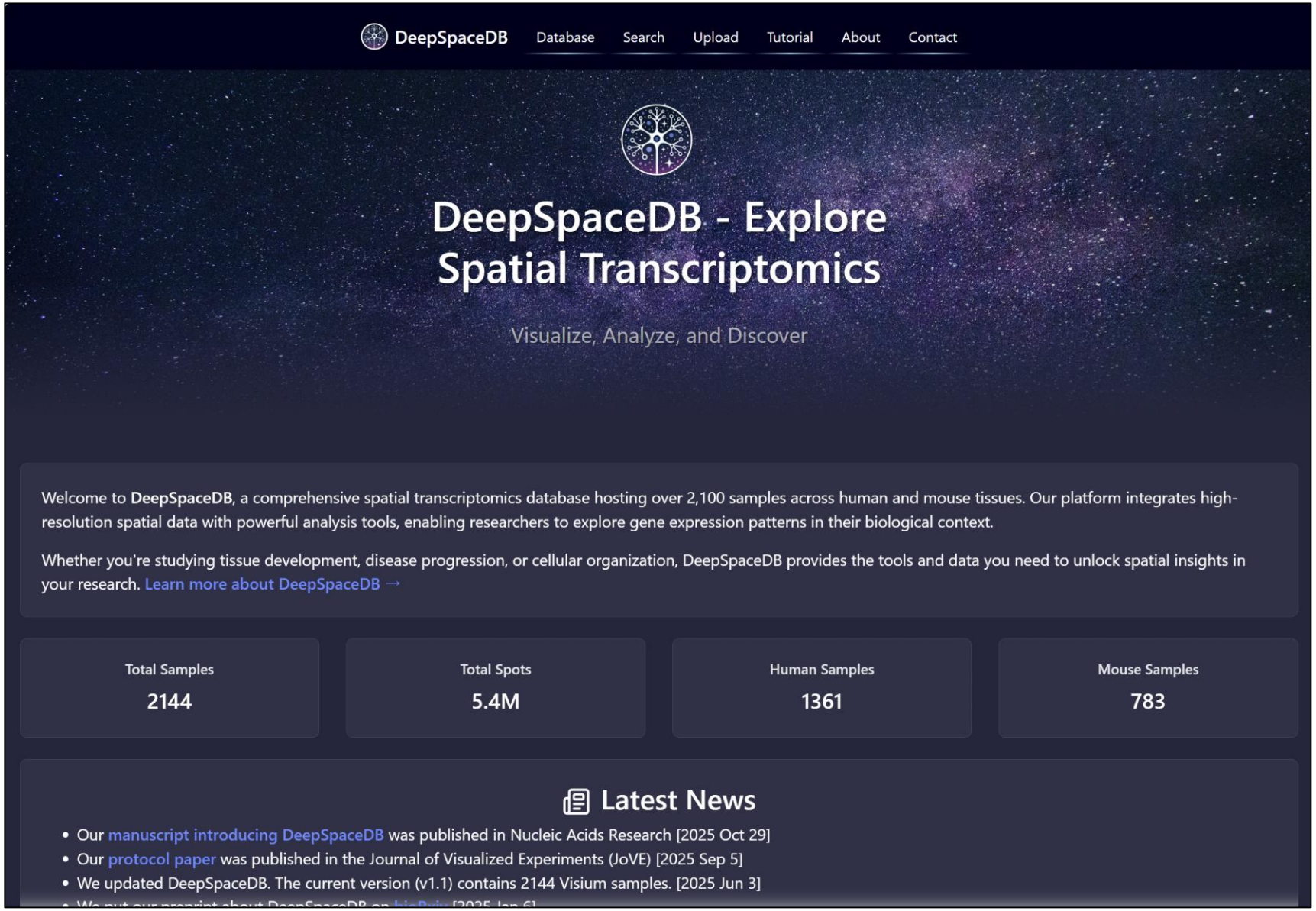


Mouse tissue slices

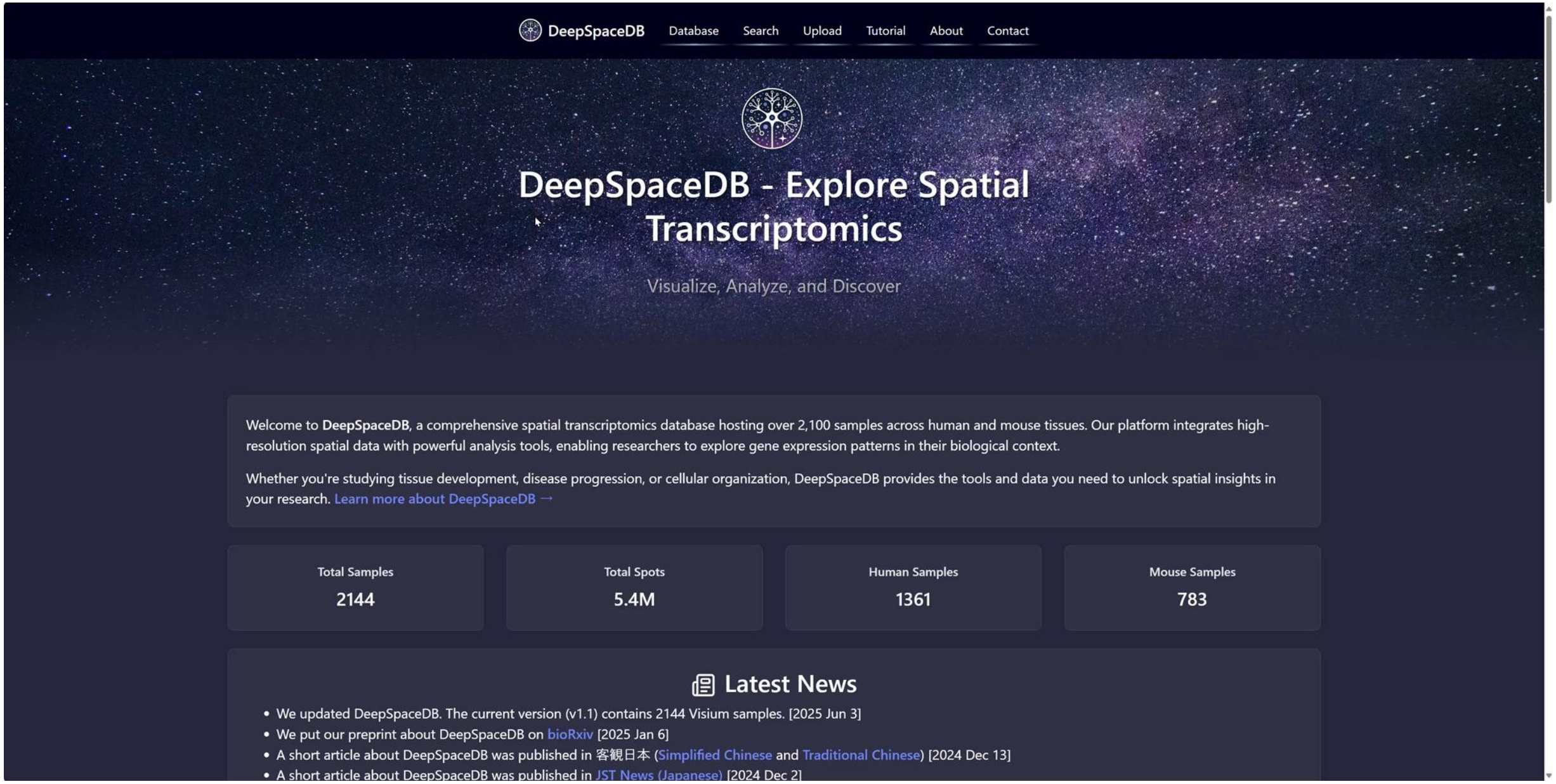


DeepSpaceDB – Top page

1. **Database:**
   サンプル検索
2. **Search:**
   遺伝子やパスウェイによる検索
3. **Upload:**
   自分のVisium dataをアップロード
4. **Tutorial:**
   ビデオとFAQs
5. **About:**
   概要研究資金引用
6. **Contact:**
   研究グループおよび研究協力者



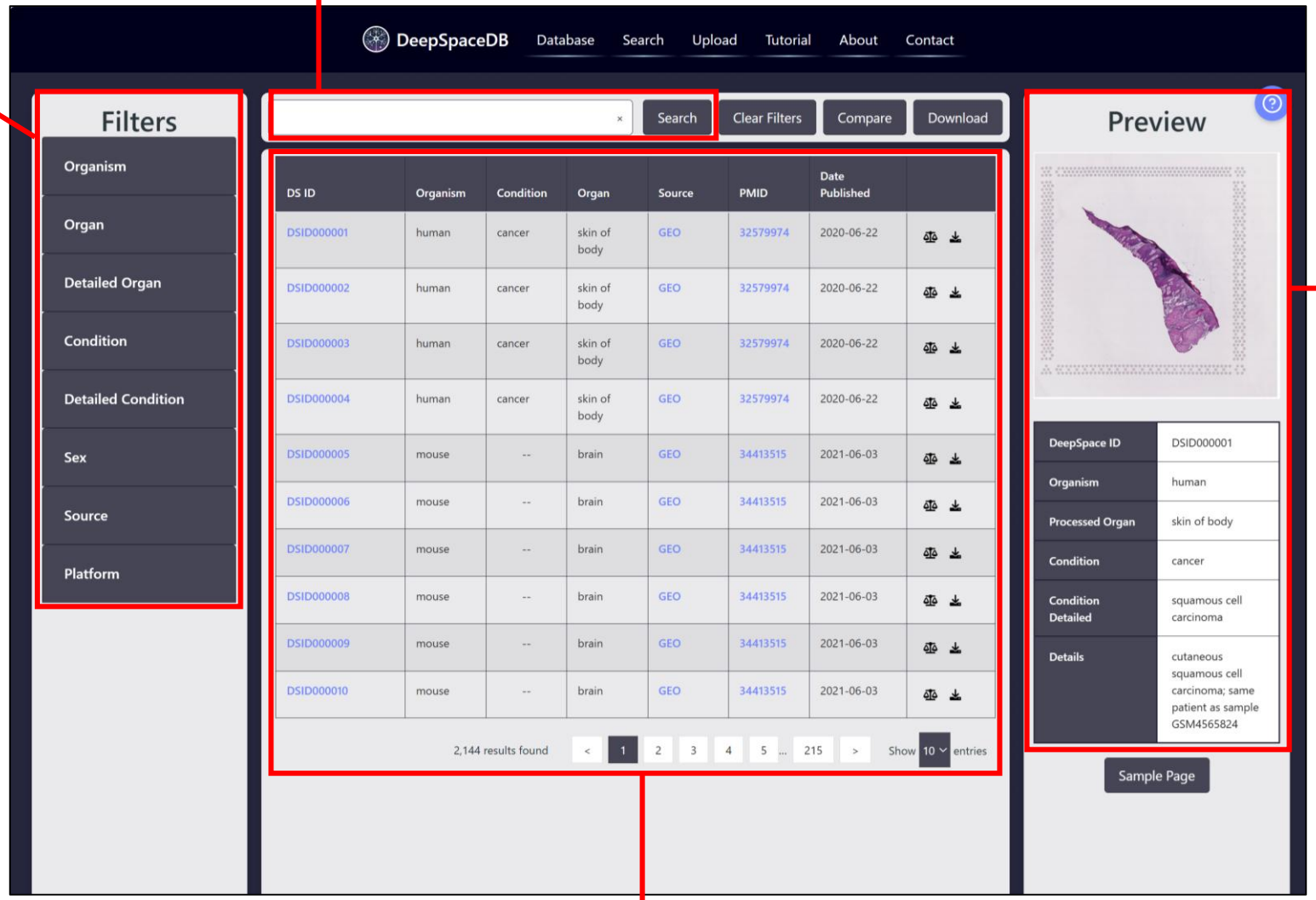
DeepSpaceDB – Demonstration (動画)



DeepSpaceDB – Database tab (1)

キーワード検索 (and)

各種フィルター



現在選んでいる
サンプルのサマリー

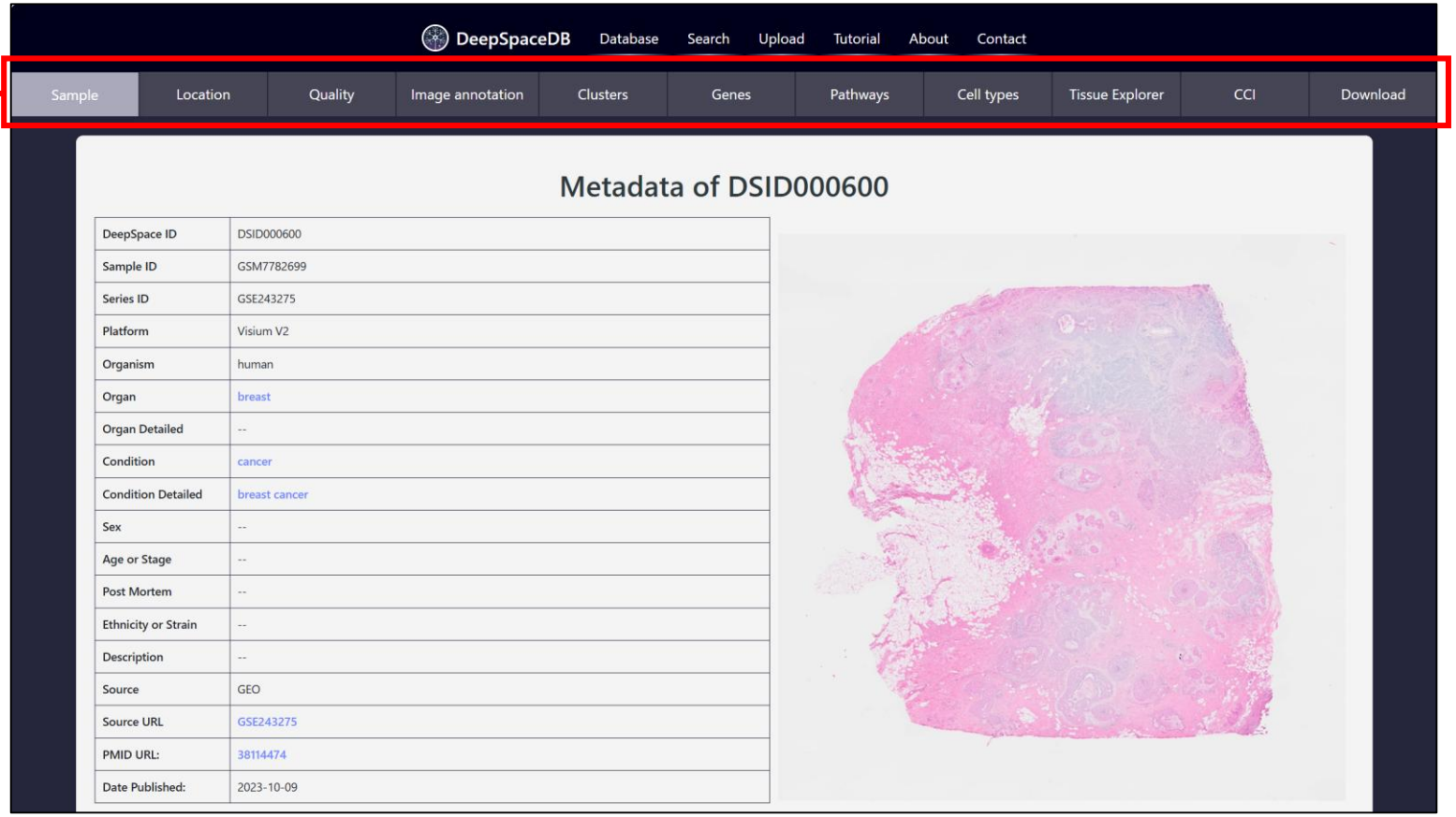
サンプルテーブル

DeepSpaceDB – Database tab (2)

フィルターにてヒト乳がんを選択した例

DeepSpaceDB – Sample page (1)

様々な情報を見ることが出来る

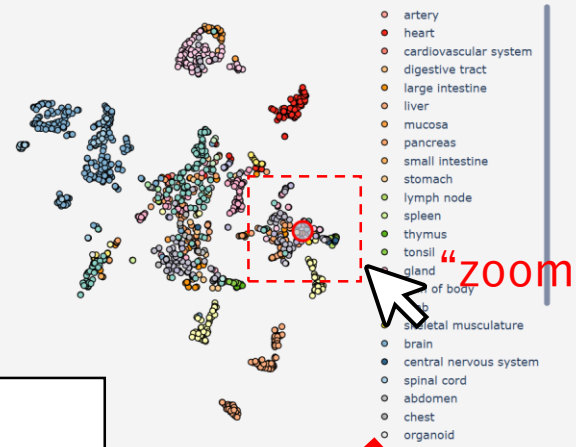


DeepSpaceDB – Sample page (2)

Sample | Location | Quality | Image annotation | Clusters | Genes | Pathways | Cell types | Tissue Explorer | CCI | Download

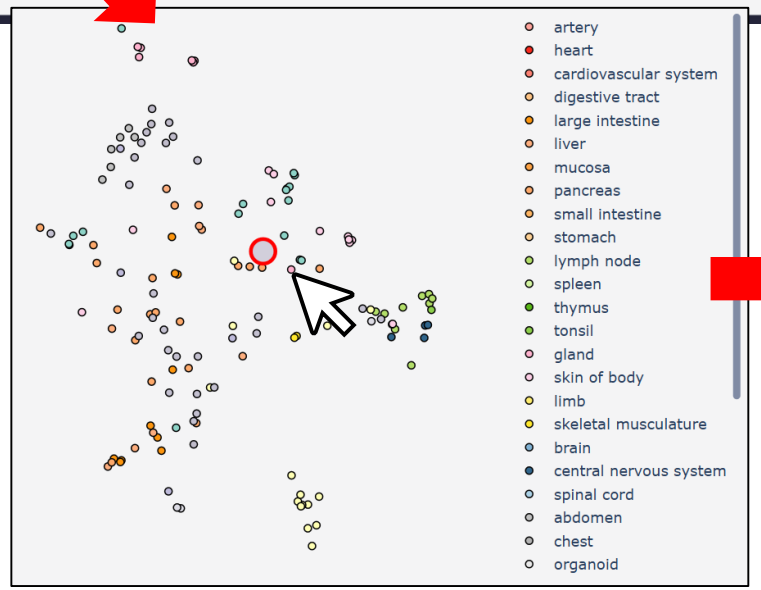
Location of the sample in the database

Embedding of all samples in DeepSpaceDB. The current sample is indicated as the point with a larger size. Select a spot to see it's preview information:



Locationタブでは、ユーザーは類似サンプルを簡単に探すことができる

“zoom” ツールで拡大



クリックでサンプルのプレビューが出る

Information about selected sample:

DeepSpace ID: DSID001012

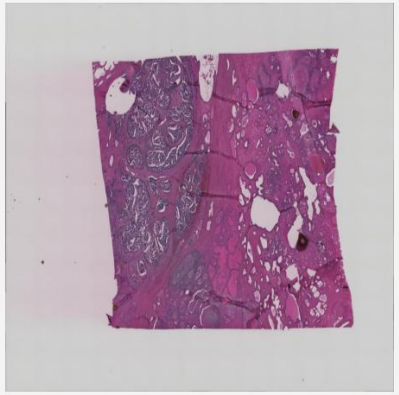
Organism: human

Organ: gland

Condition: cancer

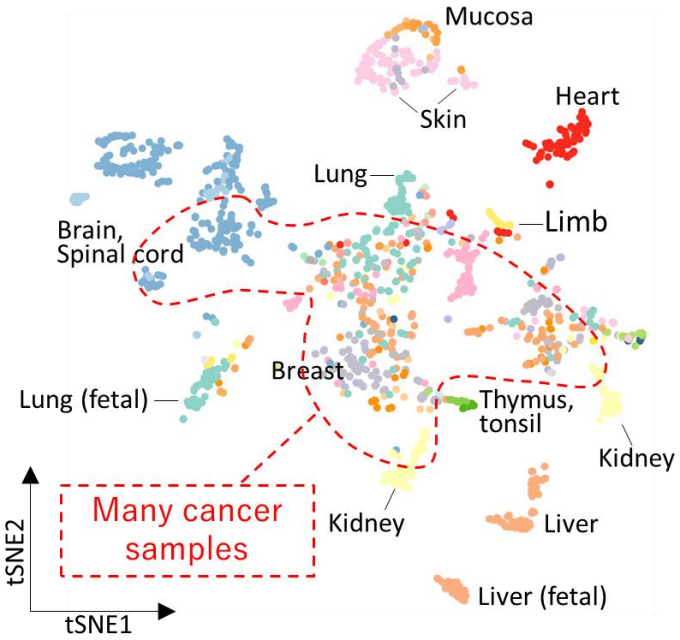
Detailed Condition: prostate cancer

Details: prostatic cancer; NEPC pT4N1M0 and ARPC pT2aN0M0, Gleason score of 4+4

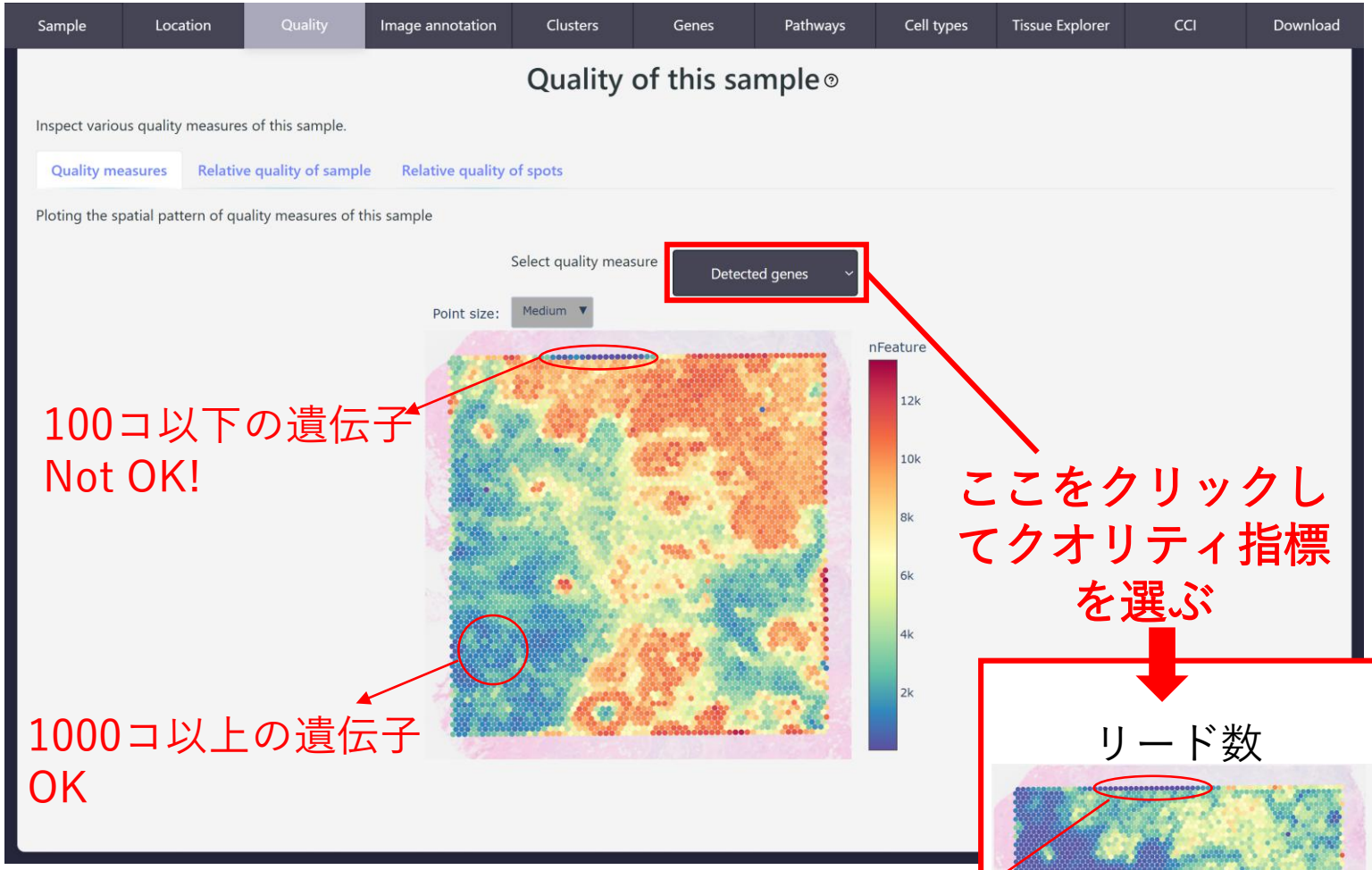


Close | Sample Page

Human tissue slices

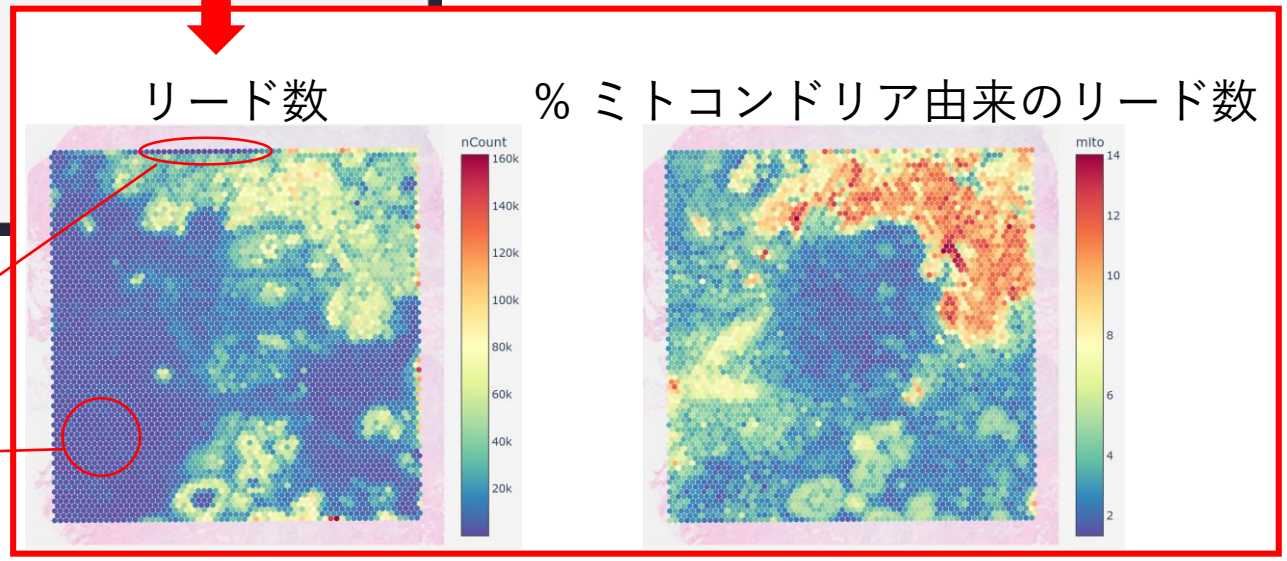


DeepSpaceDB – Sample page (3)



Qualityタブでは、サンプルのクオリティ（検出された遺伝子数やリード数など）を確認できる。
クオリティは技術的な問題による事も多いが、生物学的因子（細胞密度や細胞種など）によっても大きく左右される。

ここをクリックしてクオリティ指標を選ぶ



DeepSpaceDB – Sample page (4)

Image annotationタブ

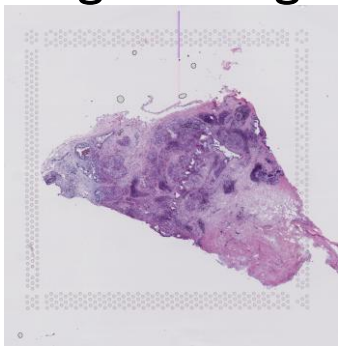
乳がん69サンプルは堀本義哉博士（東京医科大学乳腺科学）によるアノテーションがついている。

Visium H&E 画像は通常の診断用画像に比べて低画質であることが多いため、アノテーションは研究目的に限定されており、慎重に取り扱う必要がある

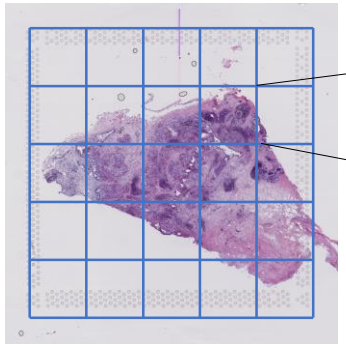
LLM（大規模言語モデル）ベースのアノテーション

LLMベースのアノテーションは注意が必要

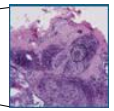
Original image



5 x 5 tiles



GPT-4oへのプロンプト

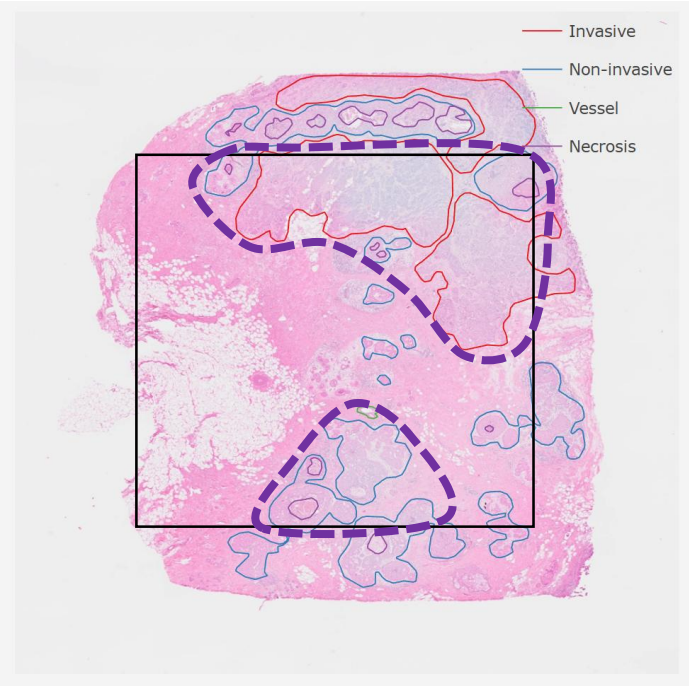


This is a part of an H&E image of a […] Describe any tissue features or signs of pathology in at most 100 characters. If there is not enough tissue to say anything, write 'empty'.

➡ 返答を収集

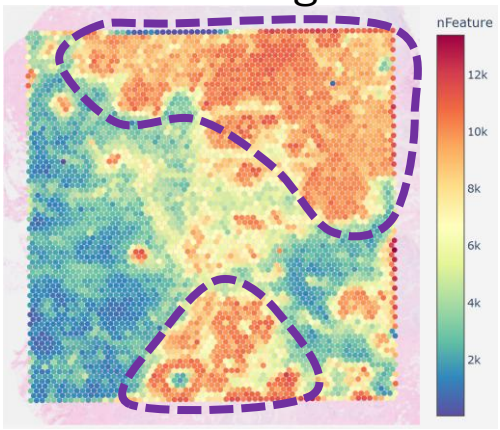
DeepSpaceDB – Sample page (5)

Image annotation

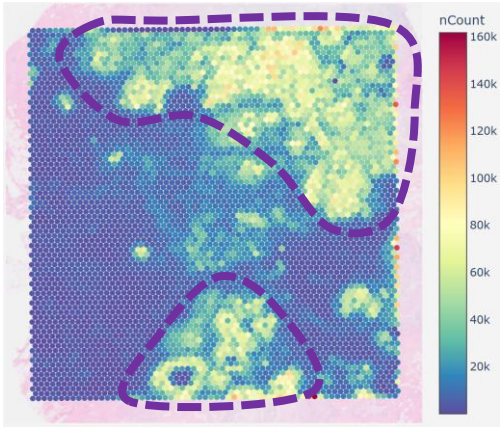


クオリティ指標

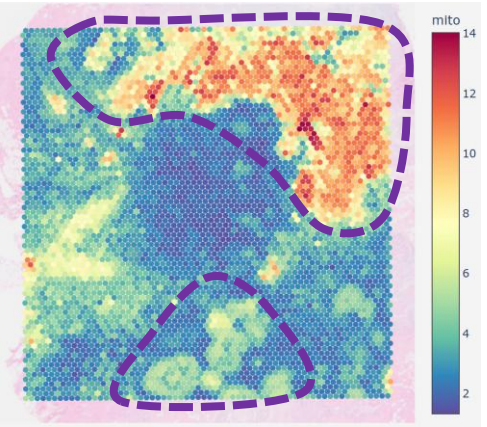
Detected genes



Read count



% mitochondrial reads



浸潤性、非浸潤性に関わらずガン組織は発現遺伝子数やリード数が多く、ミトコンドリア由来リード数も多いことが分かる。

DeepSpaceDB – Sample page (6)

Sample | Location | Quality | Image annotation | Clusters | Genes | Pathways | Cell types | Tissue Explorer | CCI | Download

Spot clustering

Visualization of spot clustering within a 2D embedding and inside the tissue. Colors represent clusters of this sample only, clusters found in the entire database, or BASS/BayesSpace domain clusters.

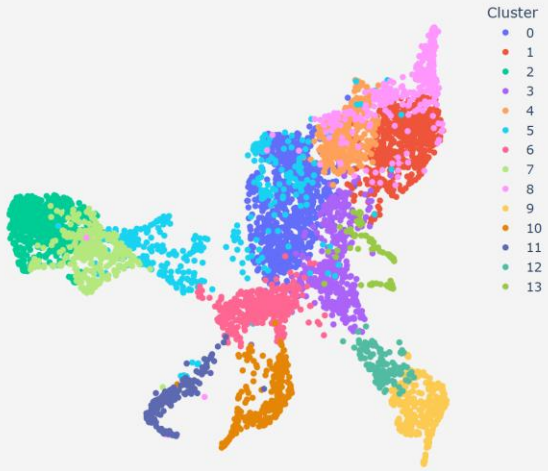
Choose clustering method:

◉ Standard Clustering ○ BASS Spatial Domains ○ BayesSpace Spatial Domains

Choose which clusters to display:

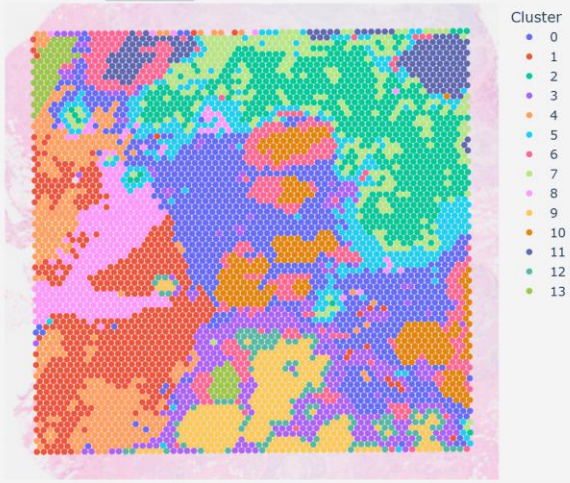
◉ Sample clusters ○ Database-wide clusters

Clusters inside 2D embedding



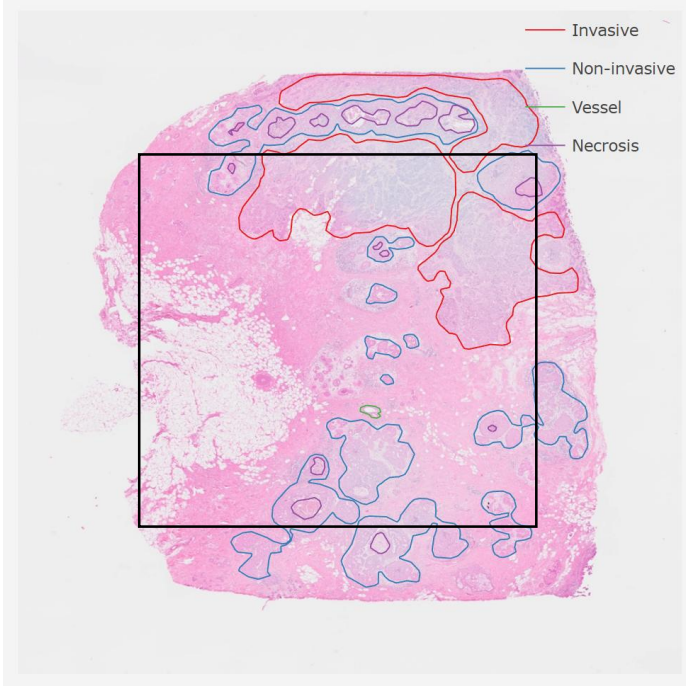
Clusters inside tissue section

Point size: Medium



Clustersタブ

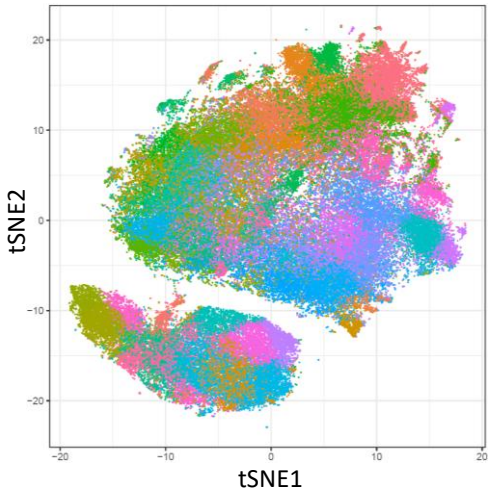
Image annotation



- サンプルのスポットは遺伝子発現パターンの類似性によりクラスタリングされる
- クラスターは生物学的特性をしばしば反映している（細胞種やガン組織など）

データベースワイドのクラスターアノテーション

ヒトの全サンプルデータ



これらのクラスターは
何を意味しているのか?

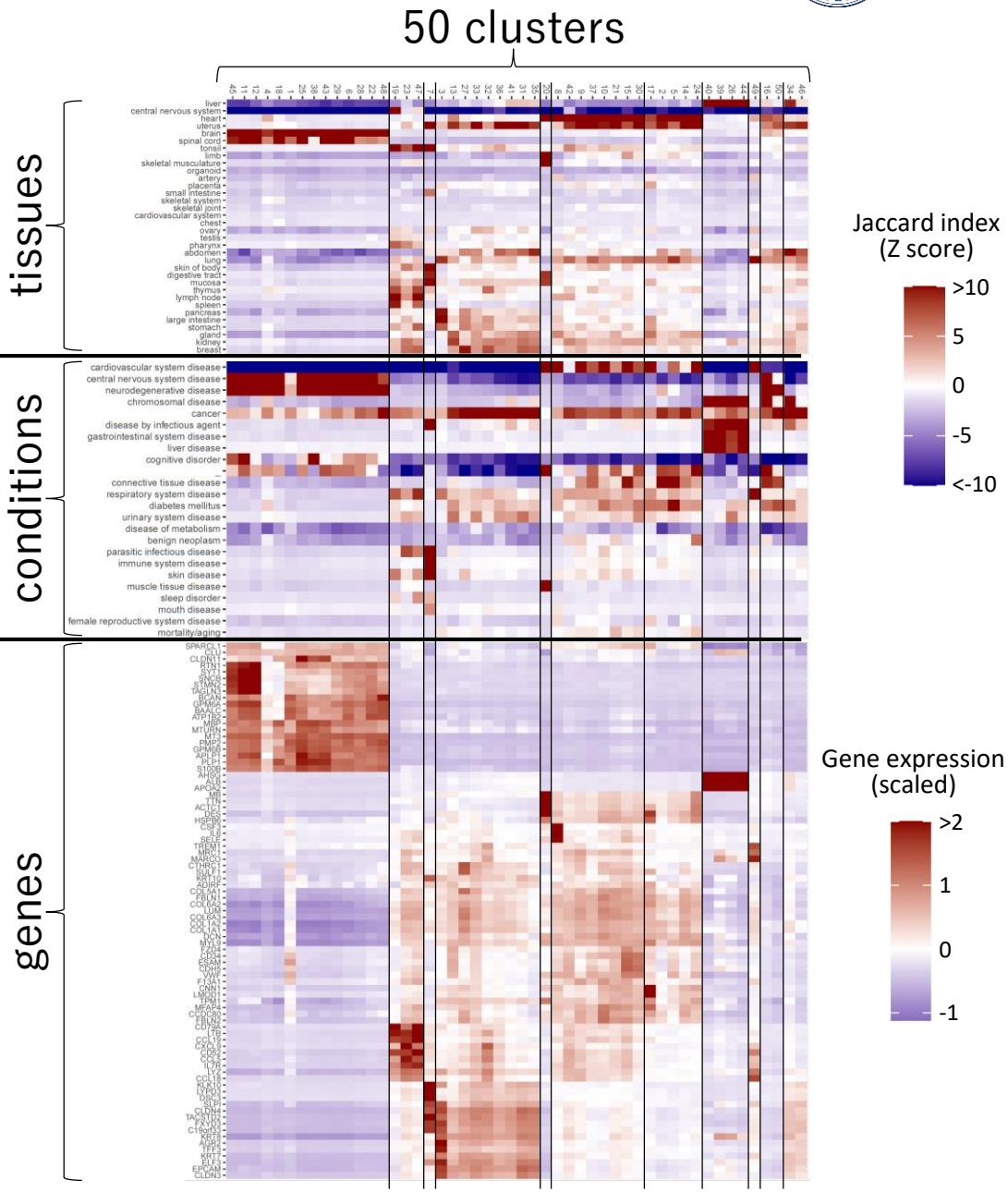
スポットクラスターの
アノテーション

(神経、ガン、上皮細胞など)

組織ごとのクラスターの
オーバーラップは？

色々な条件ごとのクラスター
のオーバーラップは？

それぞれのクラスターの
マーカー遺伝子は?



DeepSpaceDB – Sample page (7)

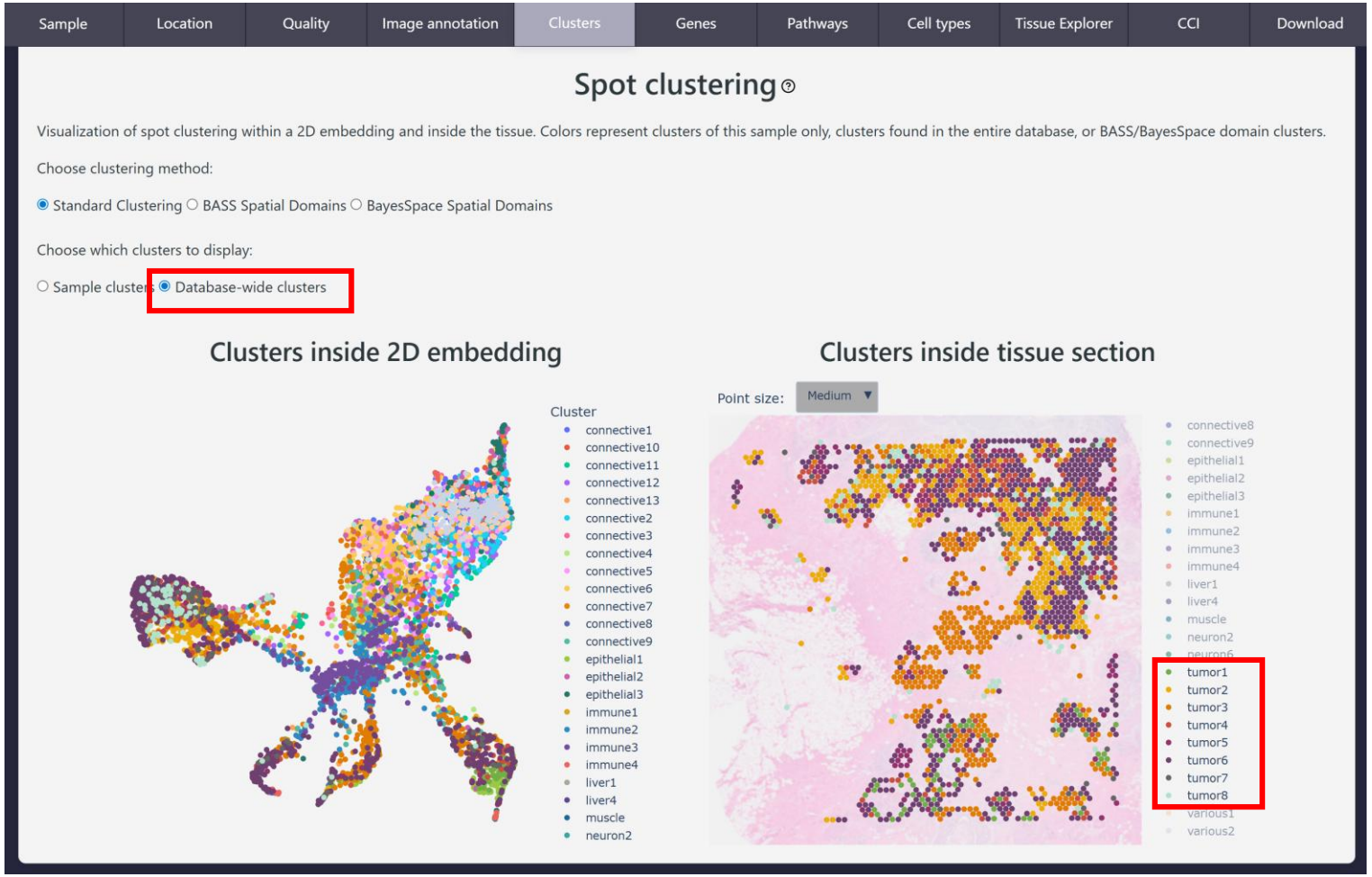
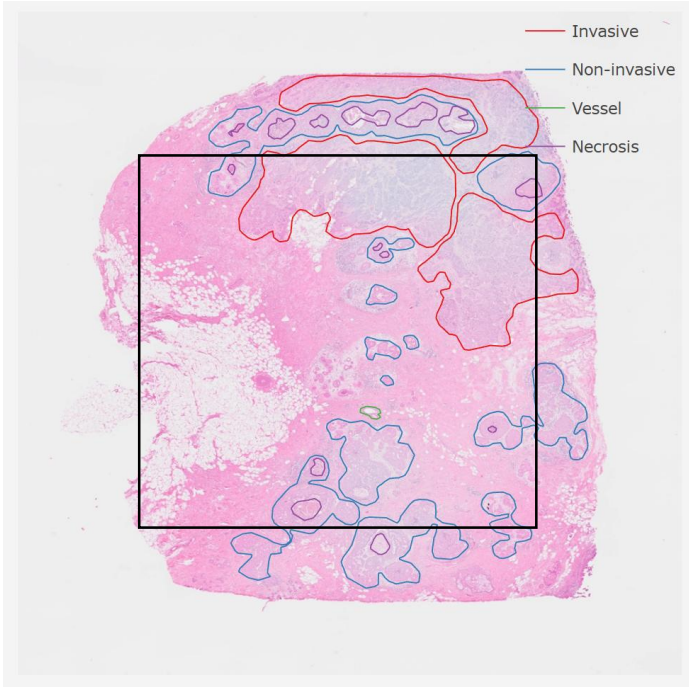
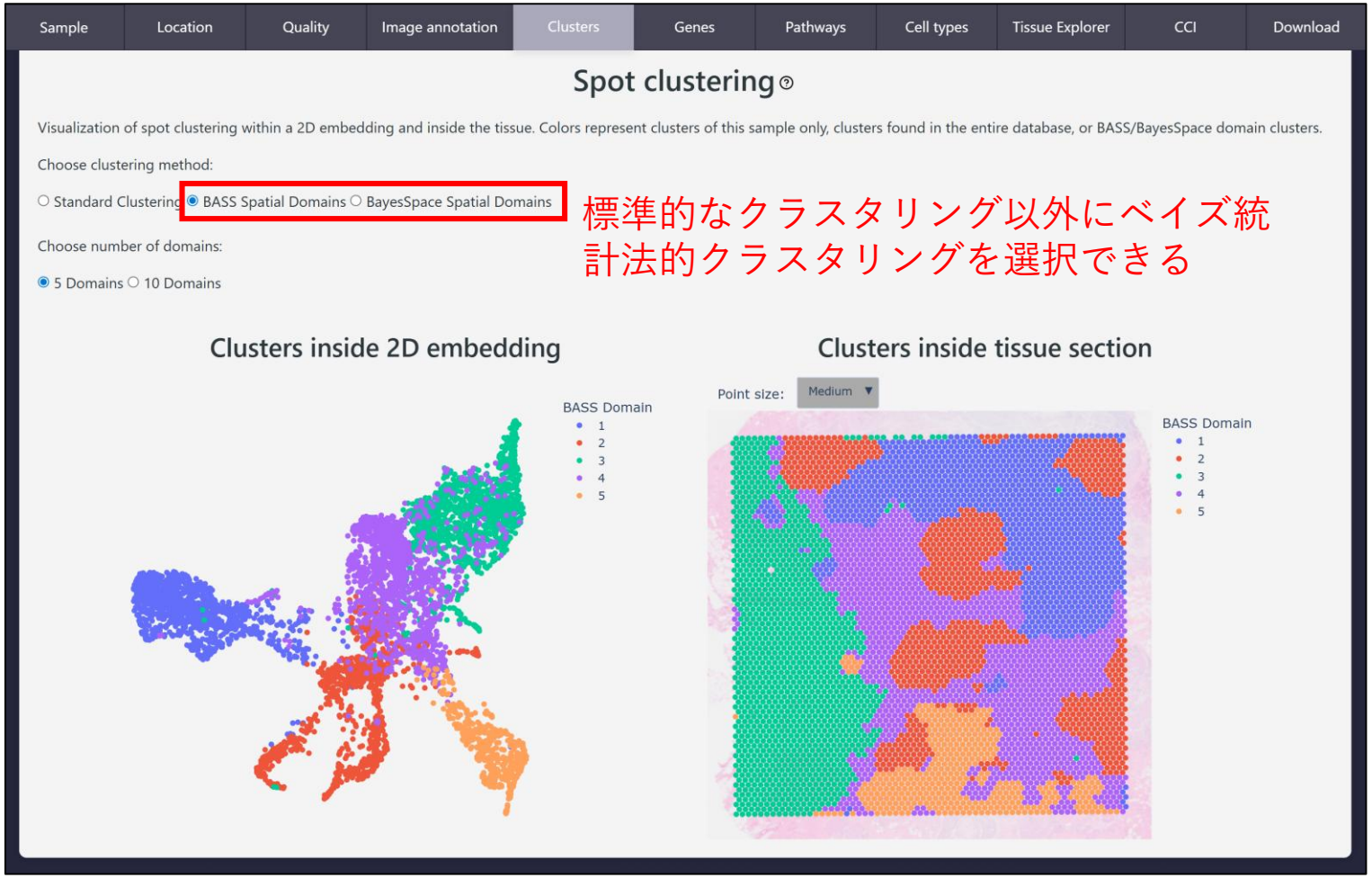


Image annotation

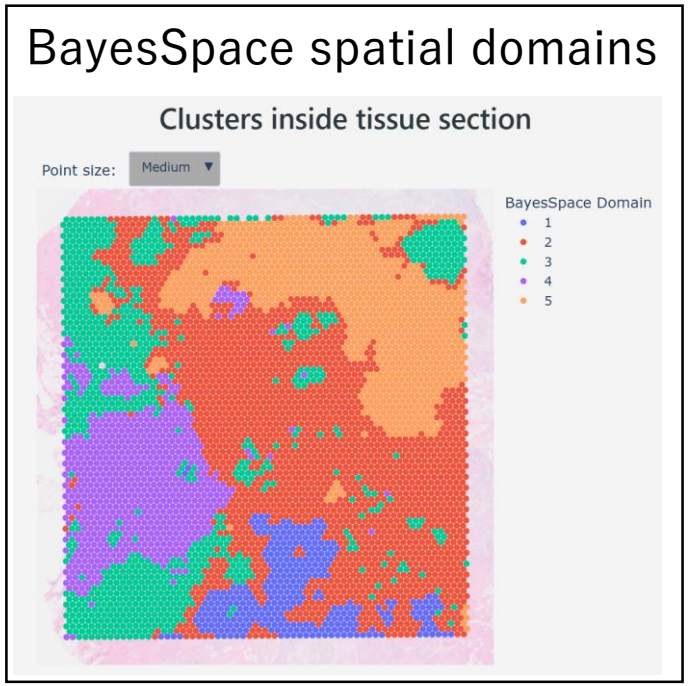


- データベースワイドのクラスターは Visium サンプルの組織構造を解釈する際に有益である。
- ここではtumorクラスター(1から8)がサンプルのガン組織上に局在する

DeepSpaceDB – Sample page (8)

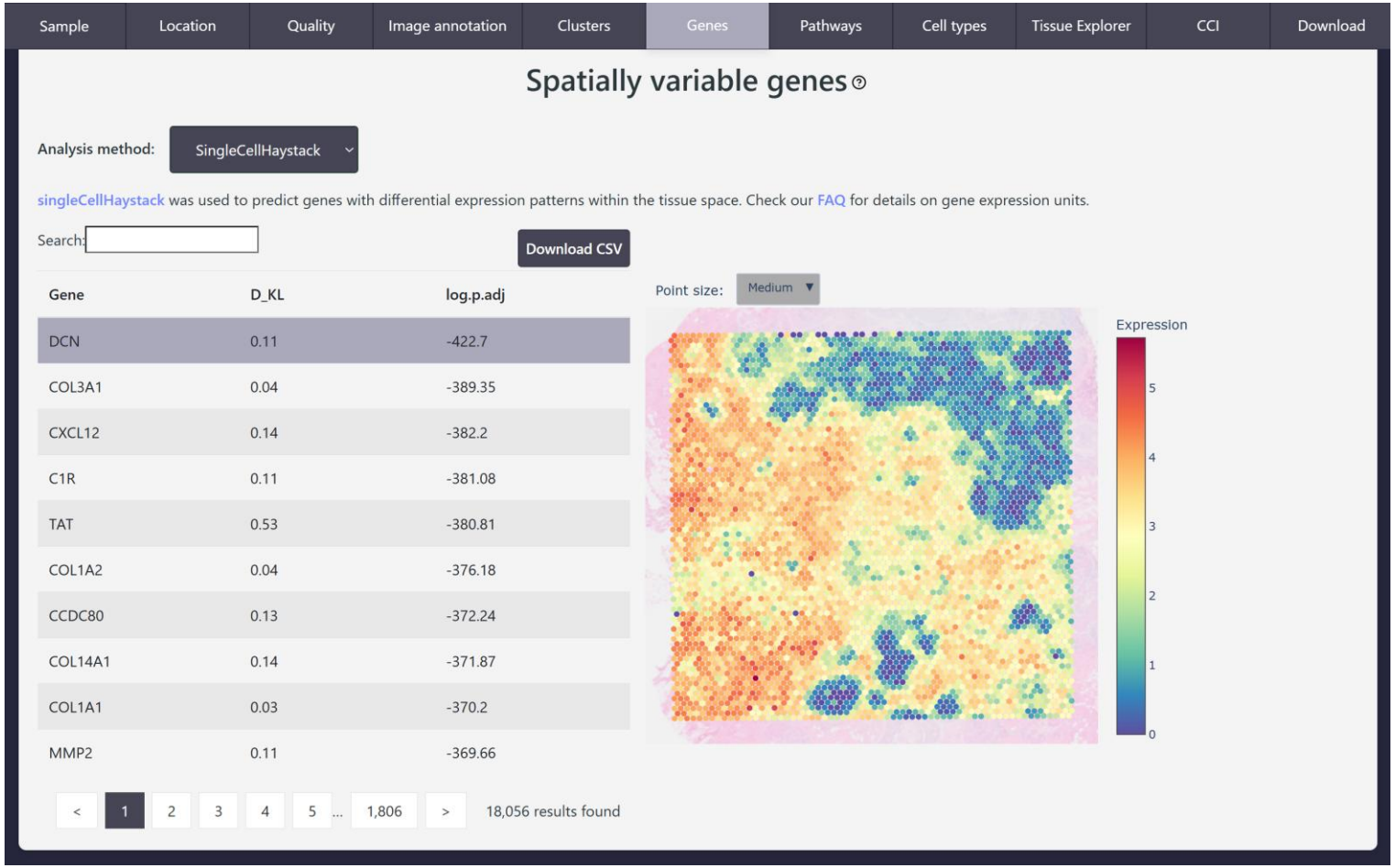


標準的なクラスタリング以外にベイズ統計法的クラスタリングを選択できる

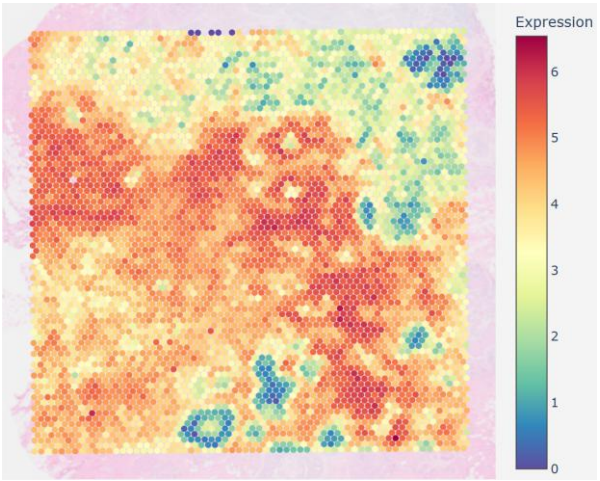


- Spatial domains は遺伝子発現情報に加えて空間情報を用いてクラスタリングを行う.
- DeepSpaceDB はBASS や BayesSpaceを用いたクラスタリングを提供している
- 解析から得られたドメインは生物学的特性をしばしば表している
- **当ウェビナー後半　酒井俊輔博士の講演**

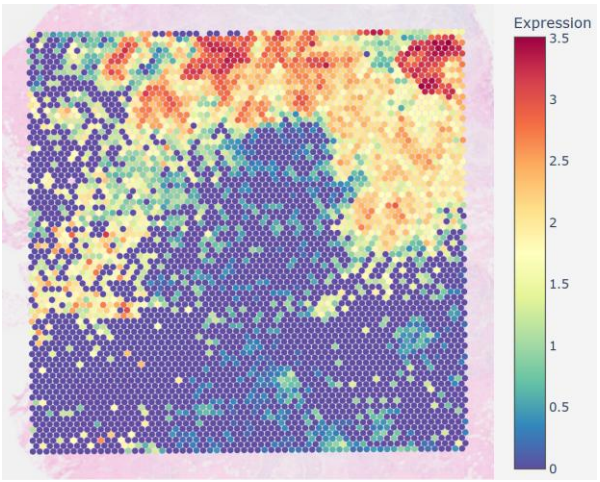
DeepSpaceDB – Sample page (9)



COL3A1



TAT

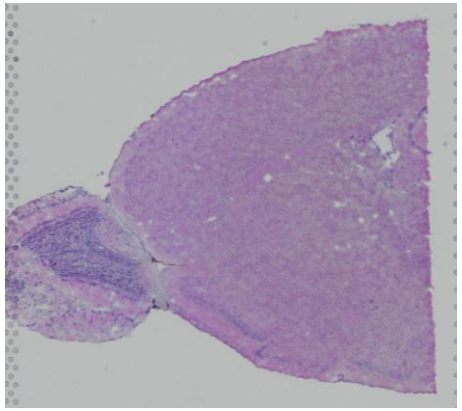


- Genesタブでは空間変動遺伝子 (SVG) を予測できる
  SVGとは明確な空間パターンをもつ遺伝子である
- DeepSpaceDB ではSVG予測プログラムのうち
  singleCellHaystack SPARK-X BinSpectが利用可能

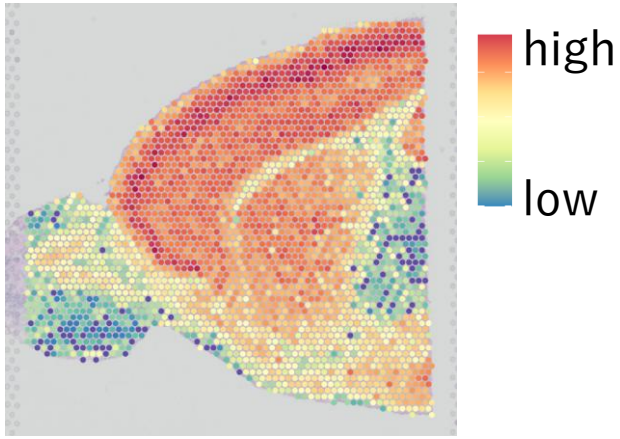
singleCellHaystack 方法論

singleCellHaystack は**細胞や遺伝子発現の分布**の偏りに基づき空間的発現変動遺伝子を予測する

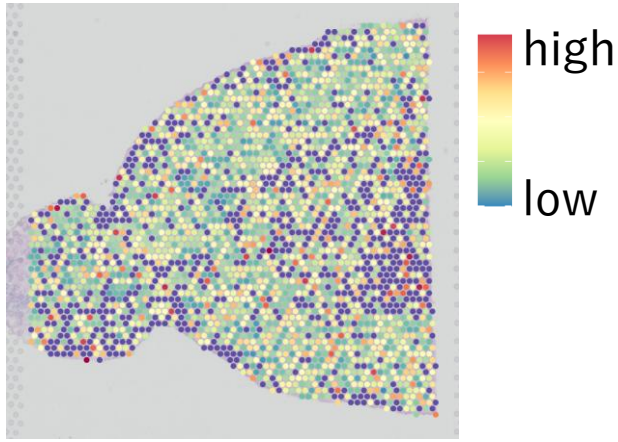
与えられた空間



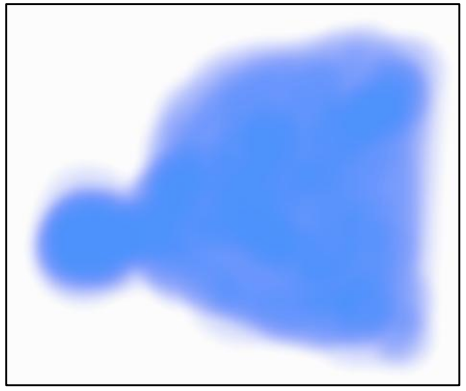
SVG



SVGではない

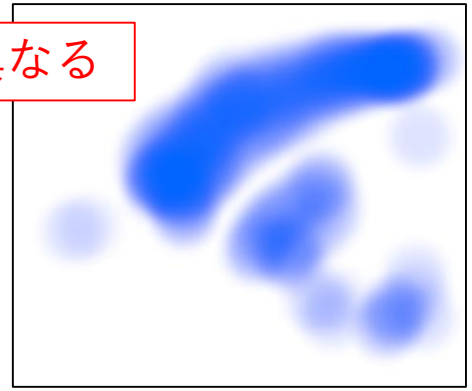


偏りのない基準分布

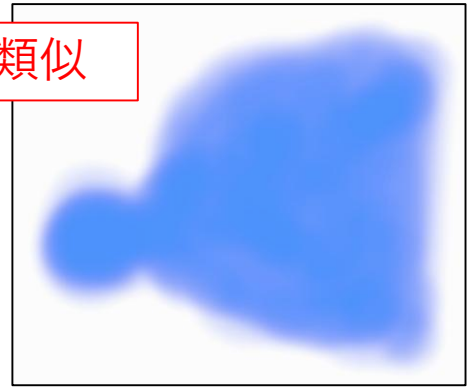


個々の遺伝子発現の分布

基準と異なる



基準に類似



Vandenbon and Diez, Nat Commun, 2020　and　Vandenbon and Diez, Sci Rep, 2023

DeepSpaceDB – Sample page (10)

Sample | Location | Quality | Image annotation | Clusters | Genes | Pathways | Cell types | Tissue Explorer | CCI | Download

Spatially variable pathways

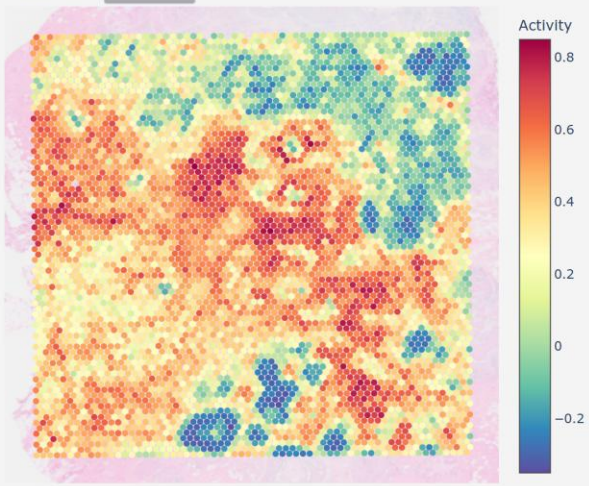
We estimated the activity of biological processes in each spot based on the expression of genes known to be involved in each process. singleCellHaystack was used to predict pathways with significant non-random patterns within the tissue. Check our FAQ for details on pathway activity units.

Search: Download CSV

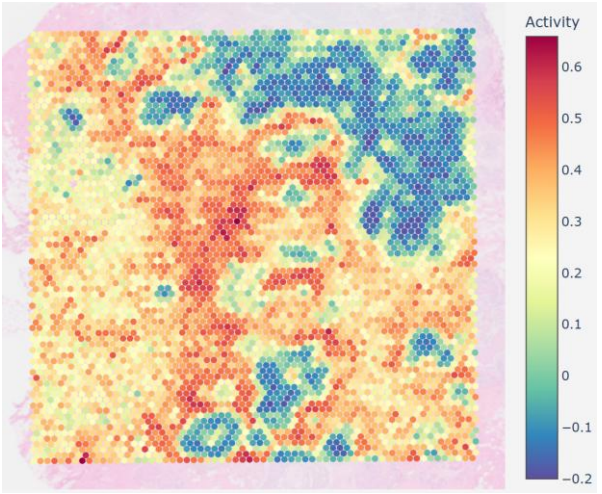
Pathway	D_KL	log.p.adj
COLLAGEN_FIBRIL_ORGANIZATION	0.04	-416.28
EXTRACELLULAR_MATRIX_ASSEMBLY	0.03	-394.88
COLLAGEN_METABOLIC_PROCESS	0.02	-382.52
COMPLEMENT_ACTIVATION	0.03	-374.87
SKELETAL_SYSTEM_MORPHOGENESIS	6.77e-03	-370.75
ATP_METABOLIC_PROCESS	3.63e-03	-369.11
COLLAGEN_CATABOLIC_PROCESS	0.03	-367.6
REGULATION_OF_EXTRACELLULAR_MATRIX_ORGANIZATION	0.03	-367.48
CELL_JUNCTION_DISASSEMBLY	0.04	-367.28
CARTILAGE_DEVELOPMENT	8.38e-03	-364.16

< 1 2 3 4 5 ... 316 > 3,152 results found

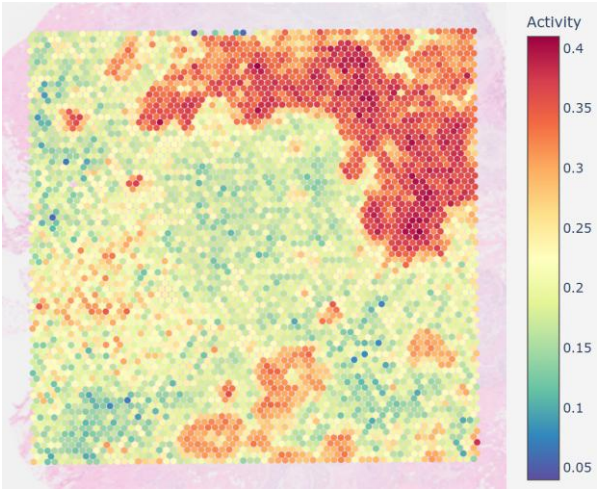
Point size: Medium



補体活性化

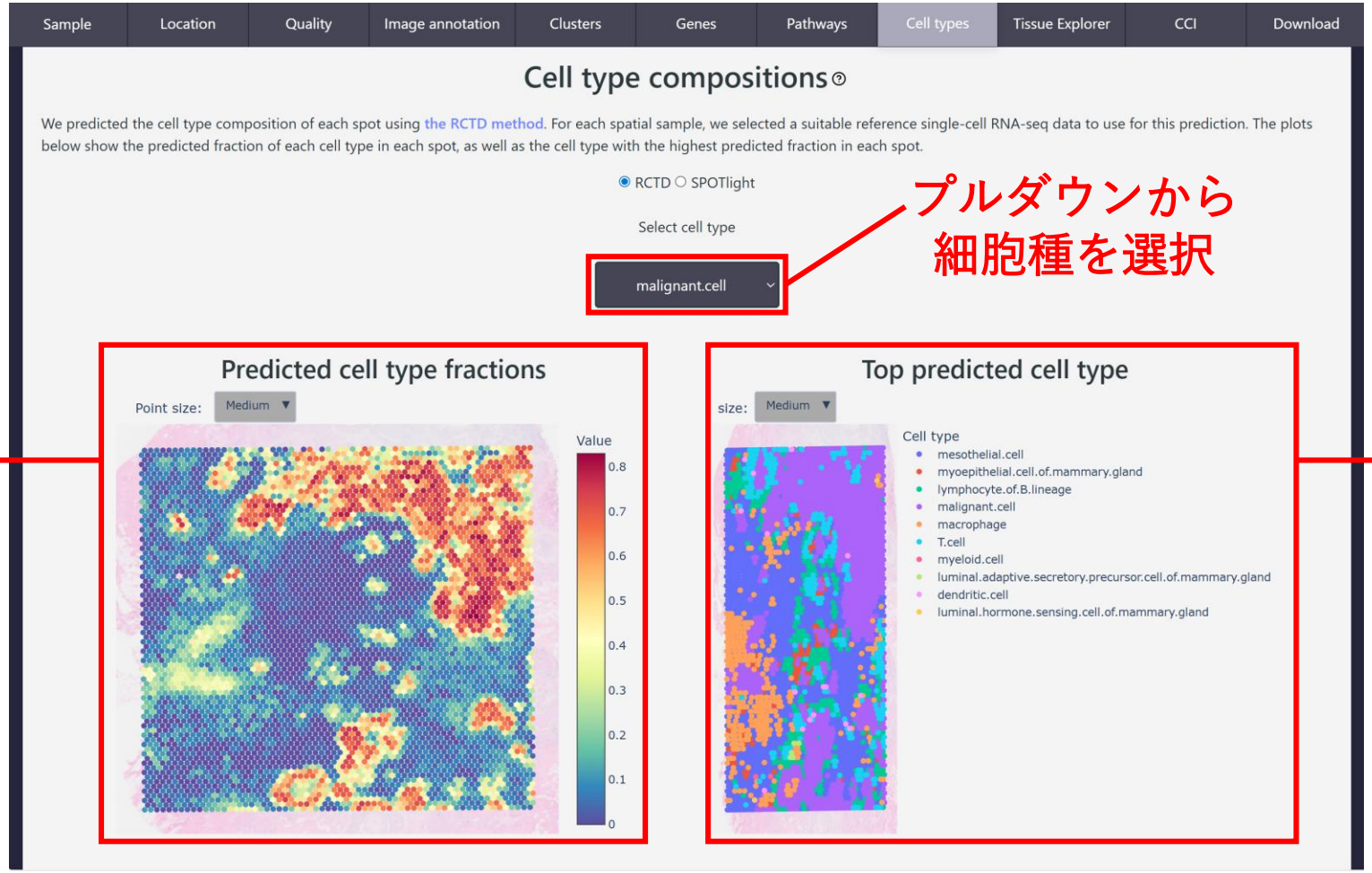


ATP 代謝プロセス



- 空間的変動パスウェイとはその活性が空間的に異なる分布を示すパスウェイである.
- 活性はそのパスウェイに関わる遺伝子群の発現に基づいている
- 予測は singleCellHaystackに基づいている

DeepSpaceDB – Sample page (11)

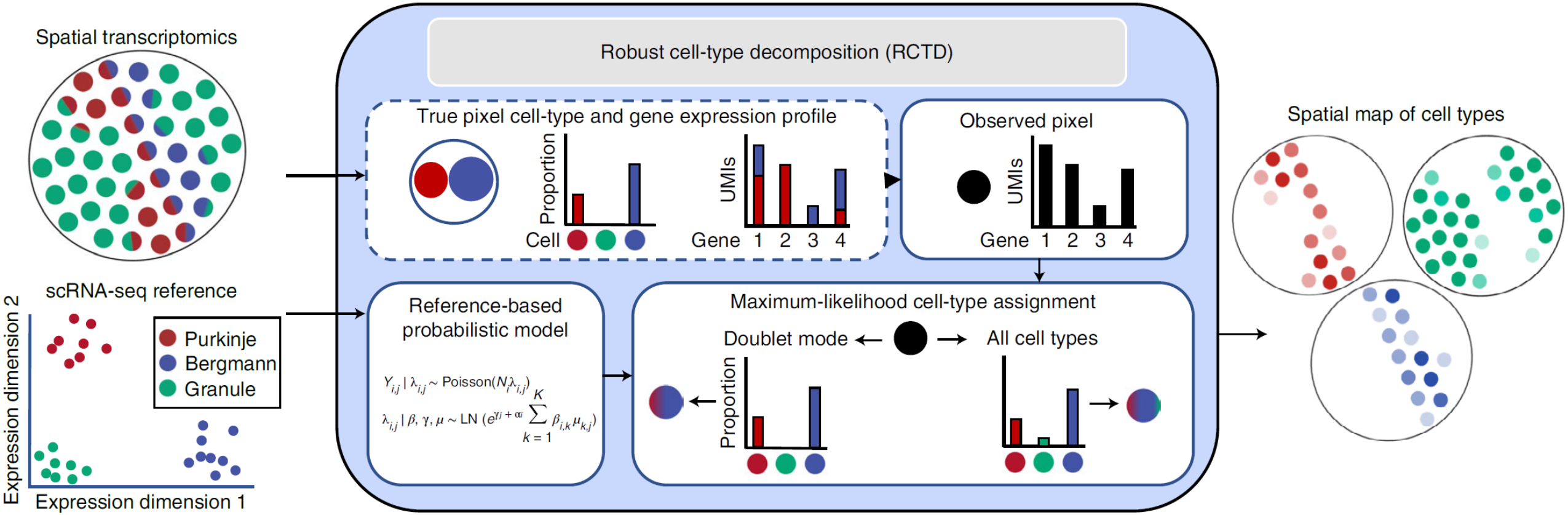


スポットごとの
malignant cells(悪性細胞)である確率が表現されている

スポットごとの
予測細胞種
(hard to see)

- **Cell typesタブでは**RCTD または SPOTlightを用いて細胞種を予測
- これらの予測はシングルセルRNA-seqデータとの比較に基づいている

RCTD 方法論



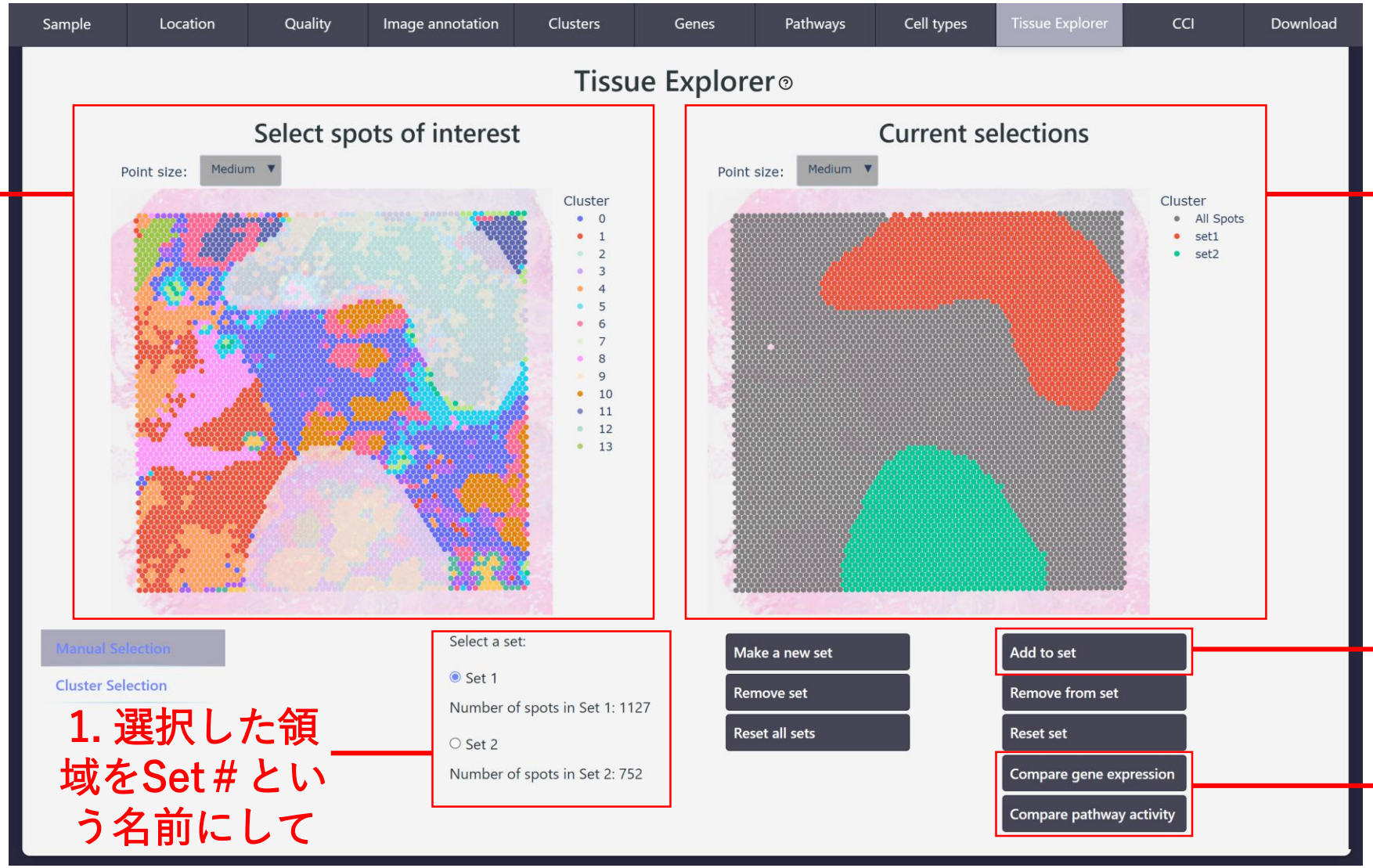
- RCTD は参照となる細胞種データを基に各スポットを細胞種の比率へと変換する
- Visium サンプルごとに最適の参照データセットを用意する事が重要

Figure from Cable *et al.*, Nat Biotech, 2021

DeepSpaceDB – Sample page (12)

“Tissue Explorer”タブ ではユーザーインタラクティブな比較が可能である

興味のある領域を自由に選択できる



現在の選択領域

1. 選択した領域をSet＃という名前にして

2. 比較リストに加え

3. セット間の差異を調べる

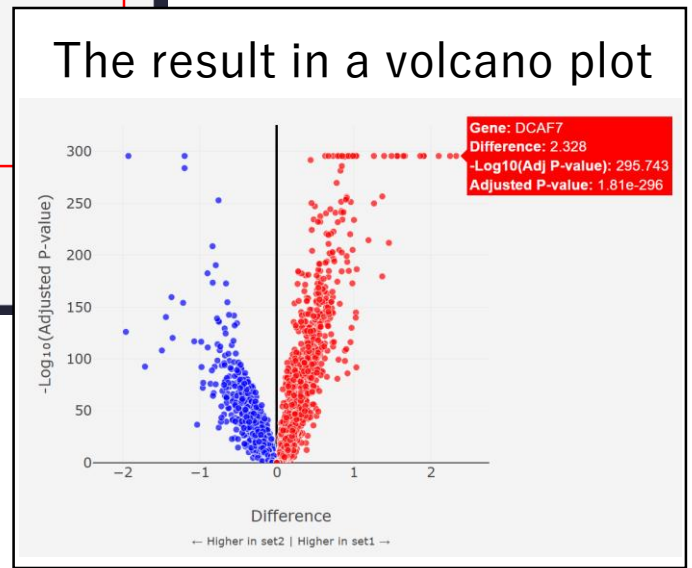
DeepSpaceDB – Sample page (13)

比較結果は“Tissue Explorer”タブの下方に表示される

遺伝子発現テーブル:
- 選択領域それぞれの平均値
- 発現の差
- P-value (t-test)
- Adjusted p-value

もし３つ以上の領域を選択した場合は、プルダウンでそのセットを表示するか選ぶことができる

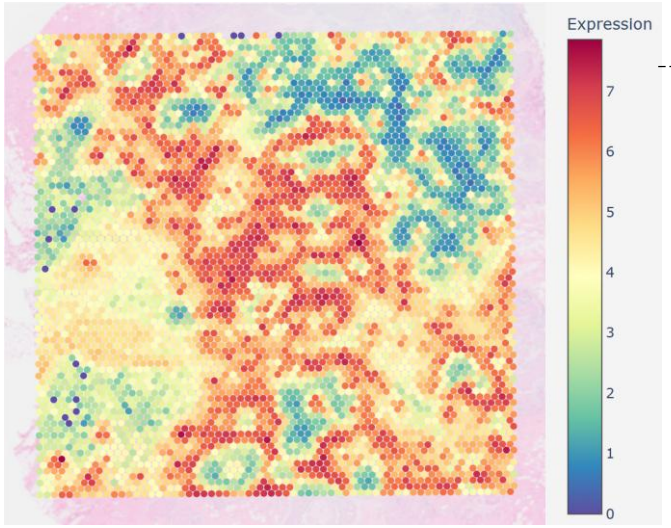
Visualization:
- scatter plot
- volcano plot



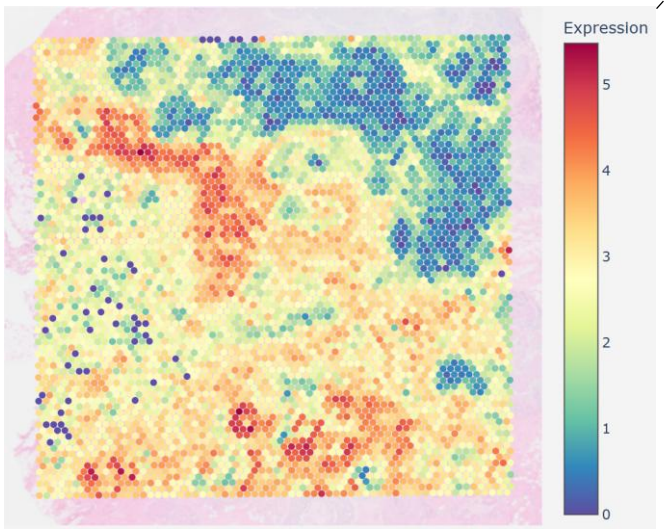
DeepSpaceDB – Sample page (14)

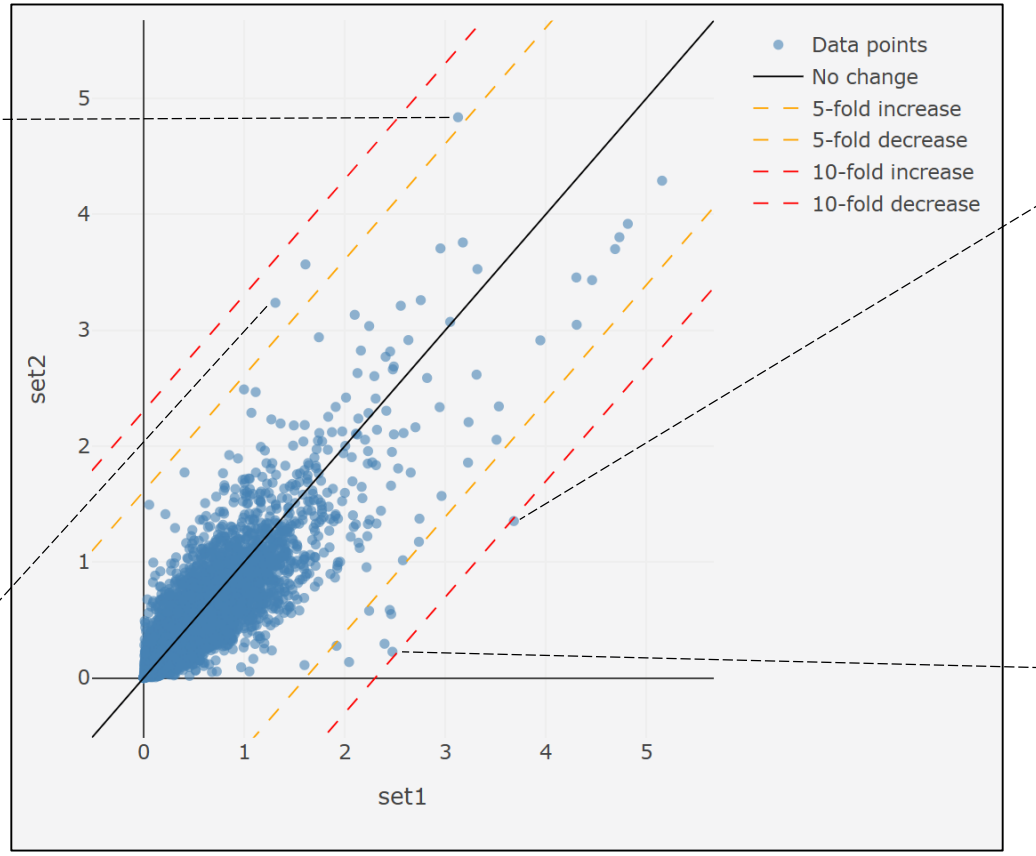
比較で得た結果は“Genes” タブで確認することができる

IGKC

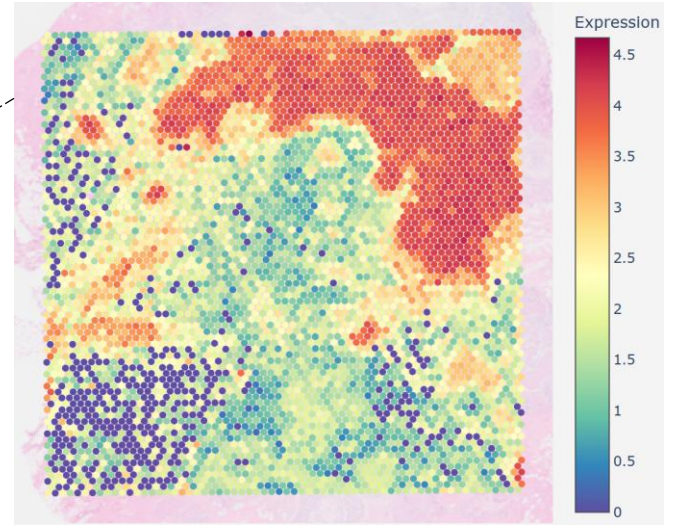


MGP

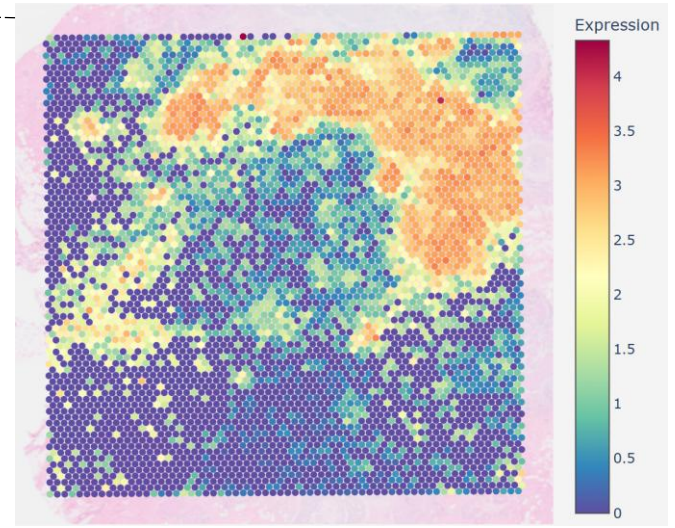




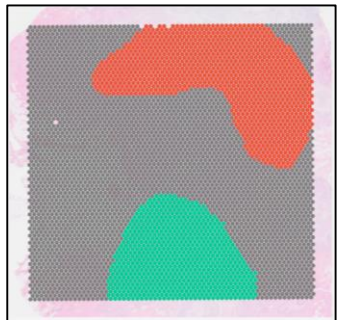
DCAF7



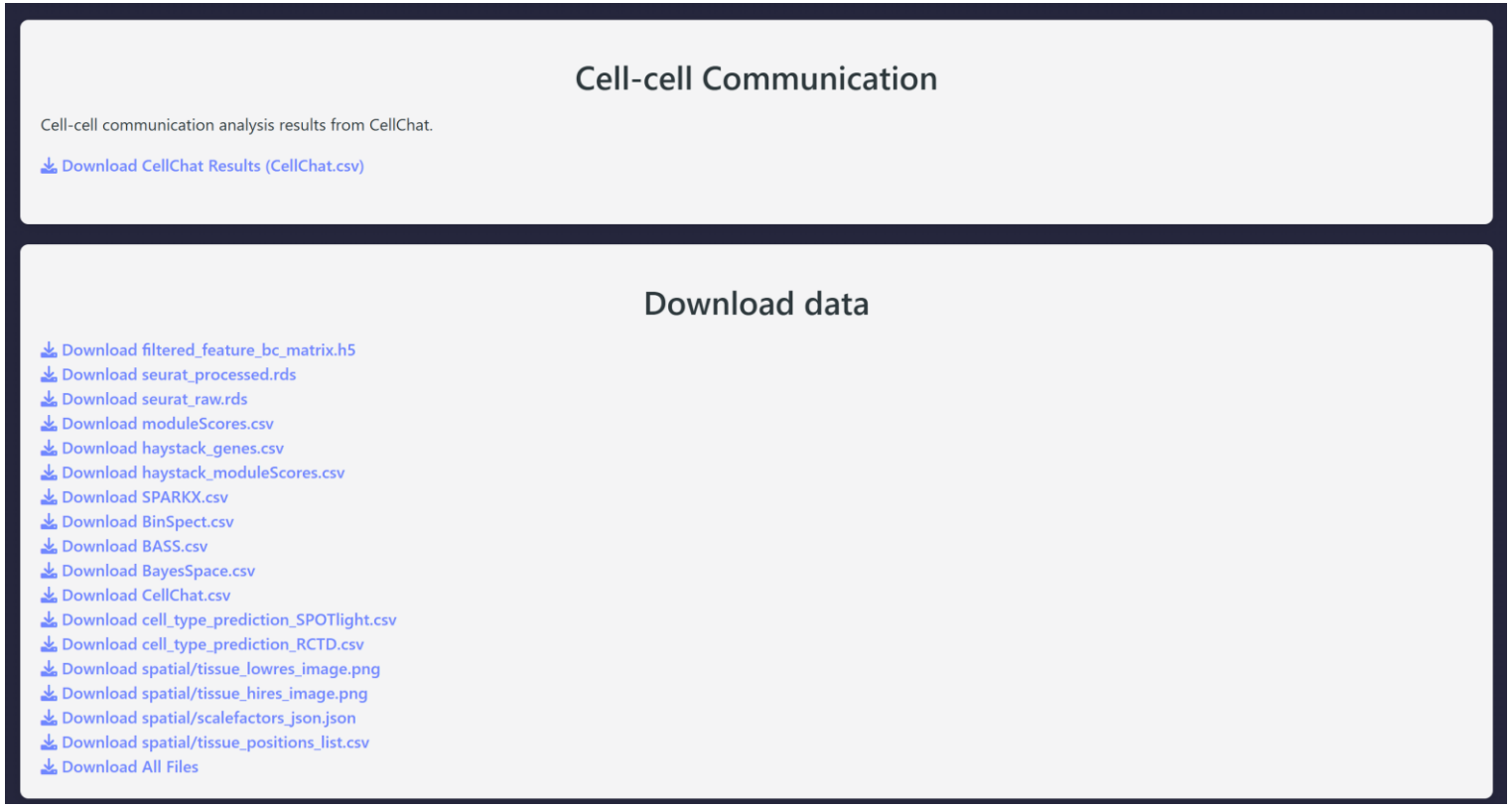
CDH2



マウス操作で選んだ領域→



DeepSpaceDB – Sample page (15)

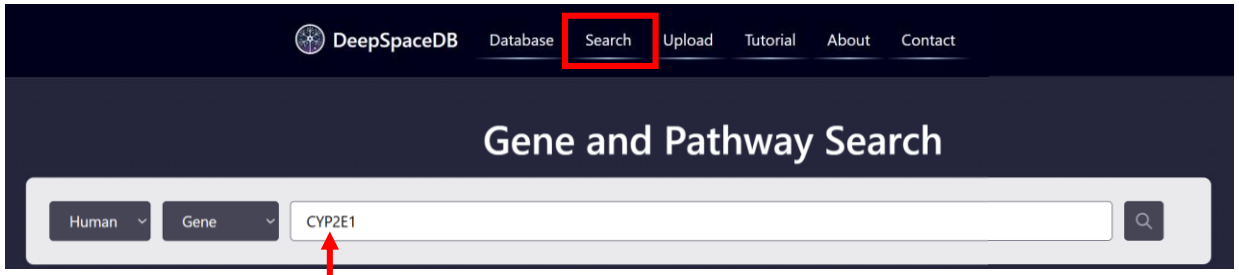


CellChatによる細胞間相互作用予測
(Jin *et al.*, Nat Protocols, 2024)

データや解析結果をダウンロードできる

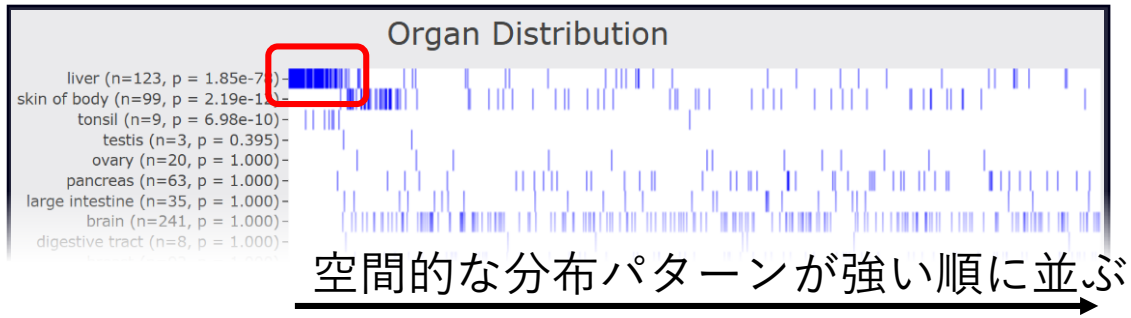
一日あたりダウンロードできるファイル数は制限あり

Search tool: データベース全体から検索

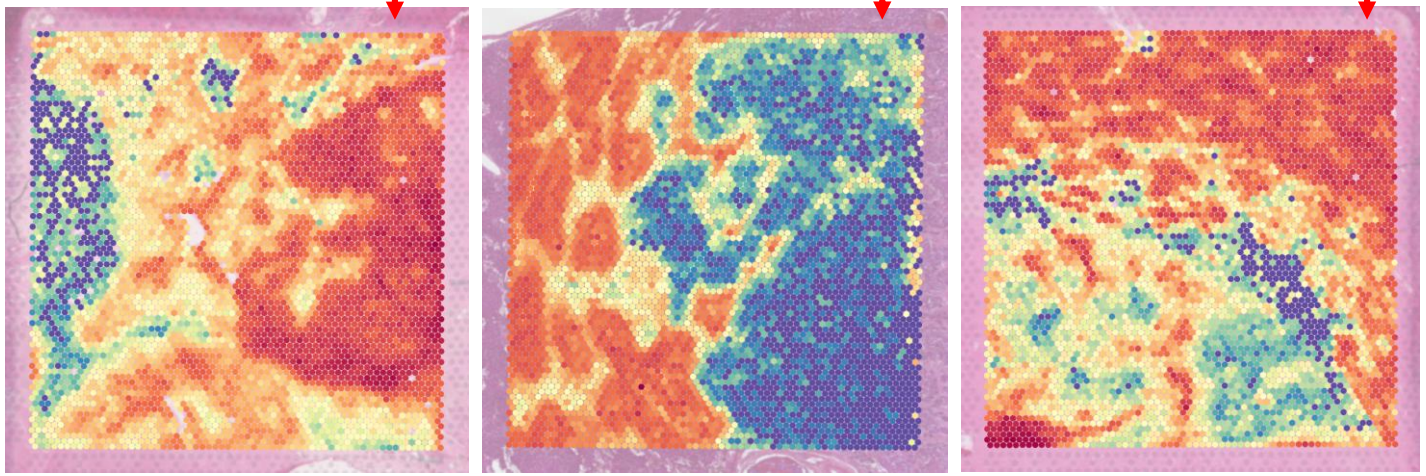


検索したい遺伝子やパスウェイを入力

- 遺伝子の機能について深い洞察へのヒントとなる
- 明確な分布パターンを持つサンプルを容易に検索できる



空間的な分布パターンが強い順に並ぶ



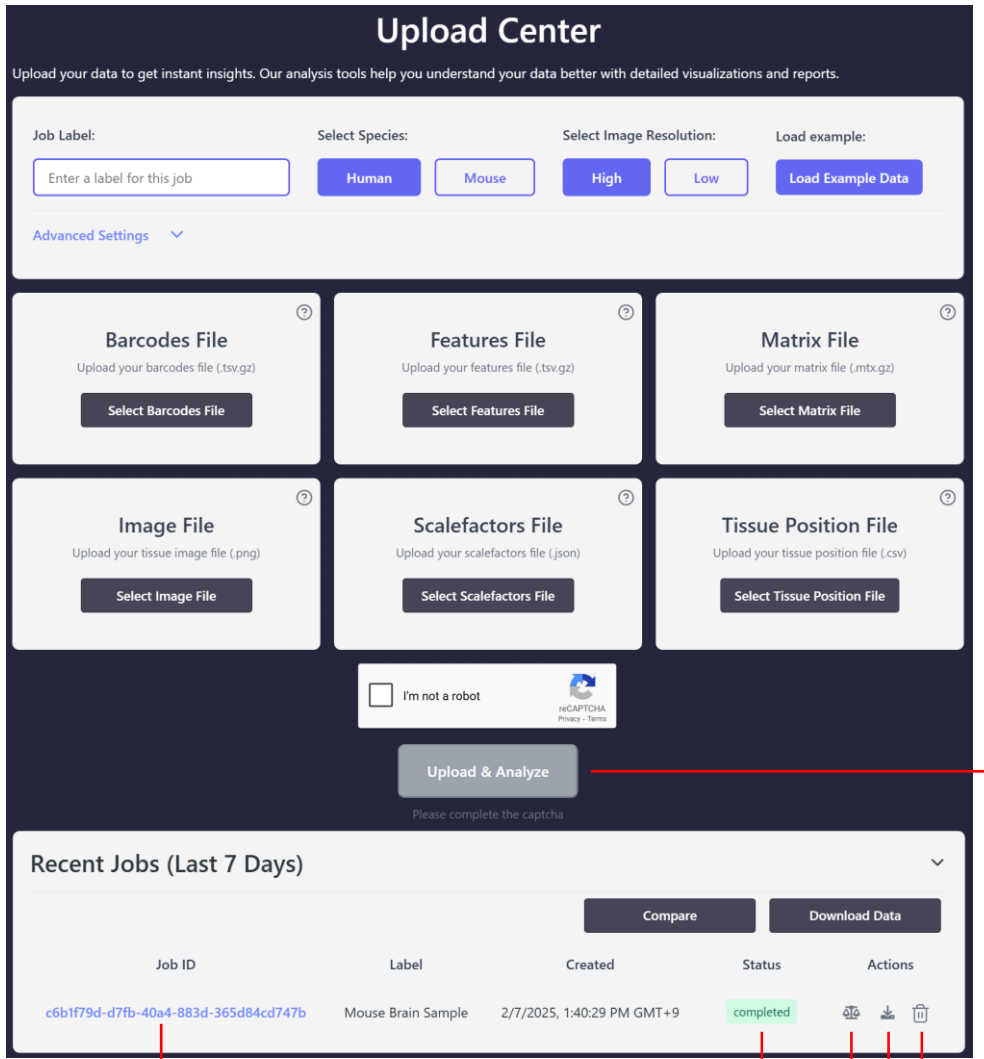
DS ID	Organ	Condition	Source	PMID	log.p.adj
DSID001919	liver	cancer	GEO	39900120	-536.15
DSID001370	liver	cancer	GEO	--	-487.46
DSID001917	liver	cancer	GEO	39900120	-415.15
DSID000838	liver	cancer	GEO	36652202	-365.63
DSID001916	liver	cancer	GEO	39900120	-355.89
DSID001913	liver	cancer	GEO	39900120	-346.59
DSID001912	liver	cancer	GEO	39900120	-336.28
DSID001914	liver	cancer	GEO	39900120	-322.7
DSID001918	liver	cancer	GEO	39900120	-306.73
DSID001915	liver	cancer	GEO	39900120	-298.14

1,337 results found < 1 2 3 4 5 ... 134 > Show 10 entries

Upload tool

データベースに保存されないのでご安心下さい

解析結果 →



アップロード
オプション

アップロードする
ファイルを選ぶ

アップロード及び解析

最近アップロードされ
たサンプル履歴

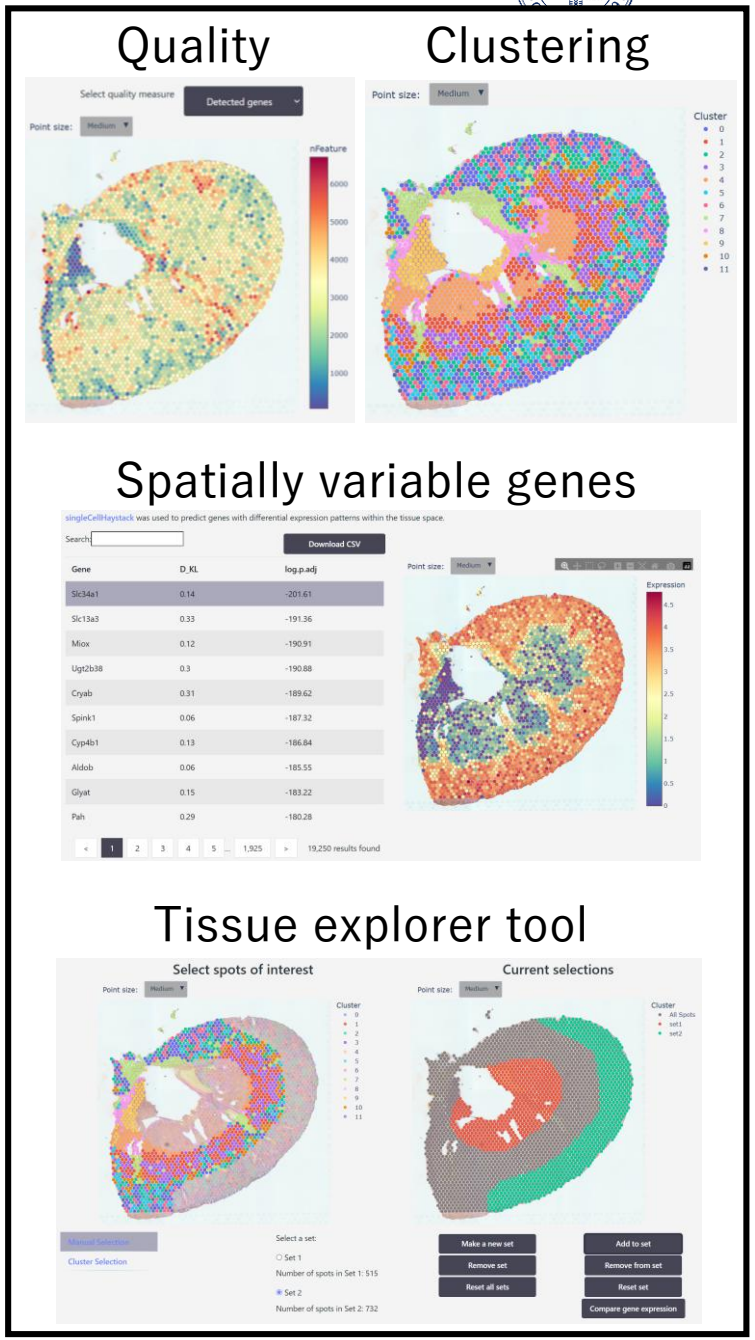
解析結果を表示

プロセシングステータス

サーバーからの消去

結果のダウンロード

他サンプルとの比較



概要

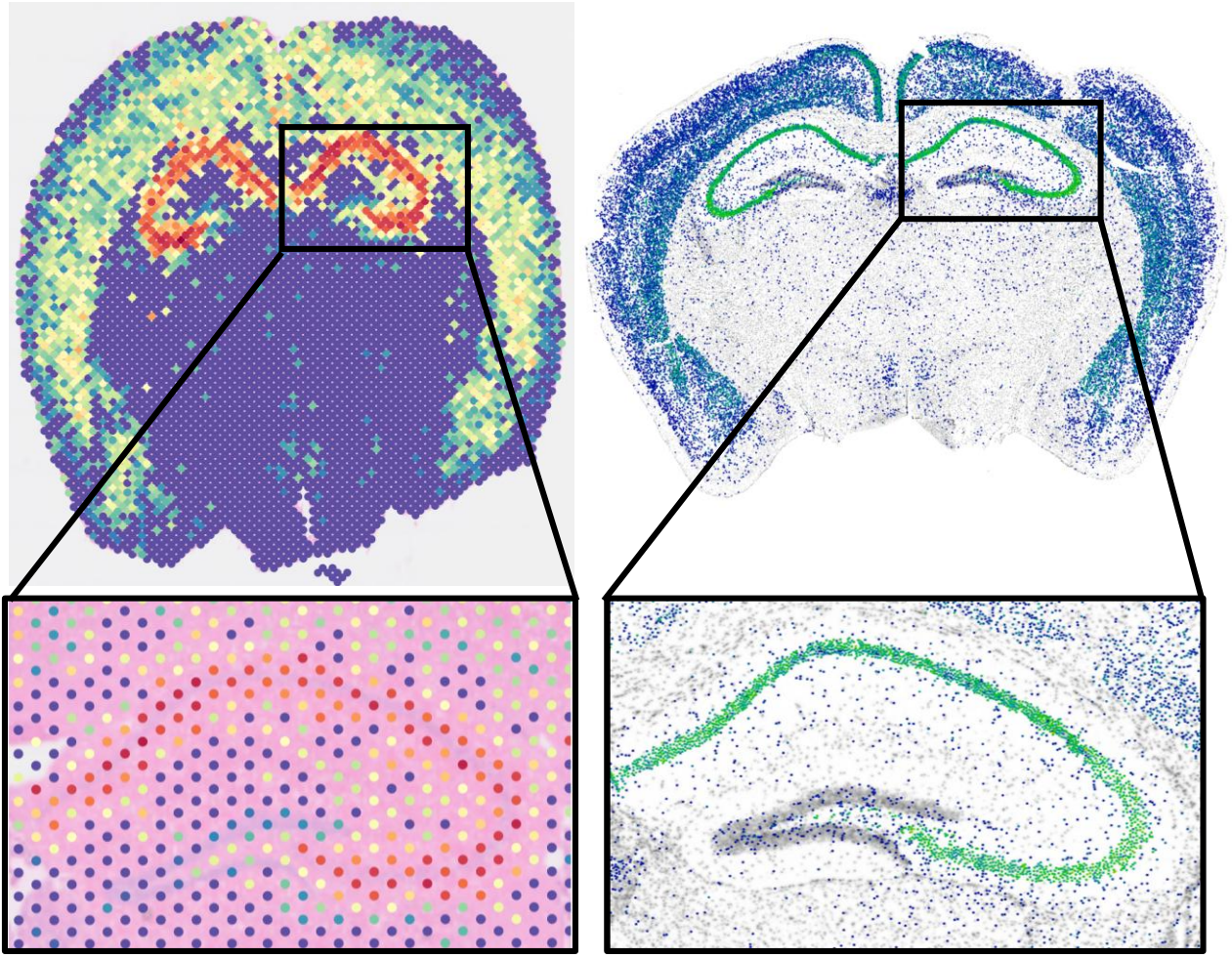
- Rを使った空間トランスクリプトミクスデータ解析
  - Visium example
  - Xenium example
- 既存の空間トランスクリプトミクスデータベースの紹介
- DeepSpaceDBの紹介
  - Visium interface
  - Xenium interface
- まとめ

VisiumとXeniumの比較例

Neurod6

Visium

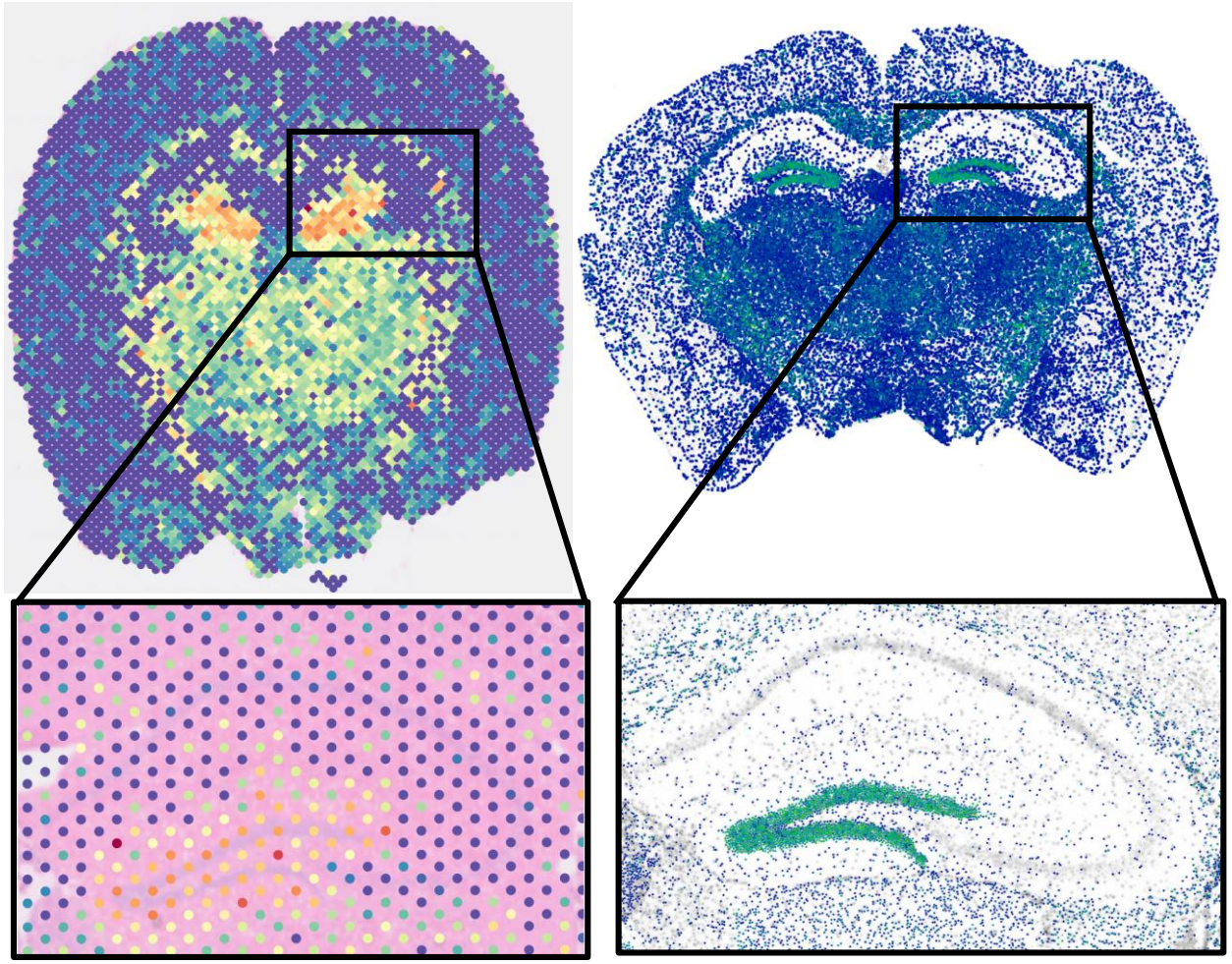
Xenium



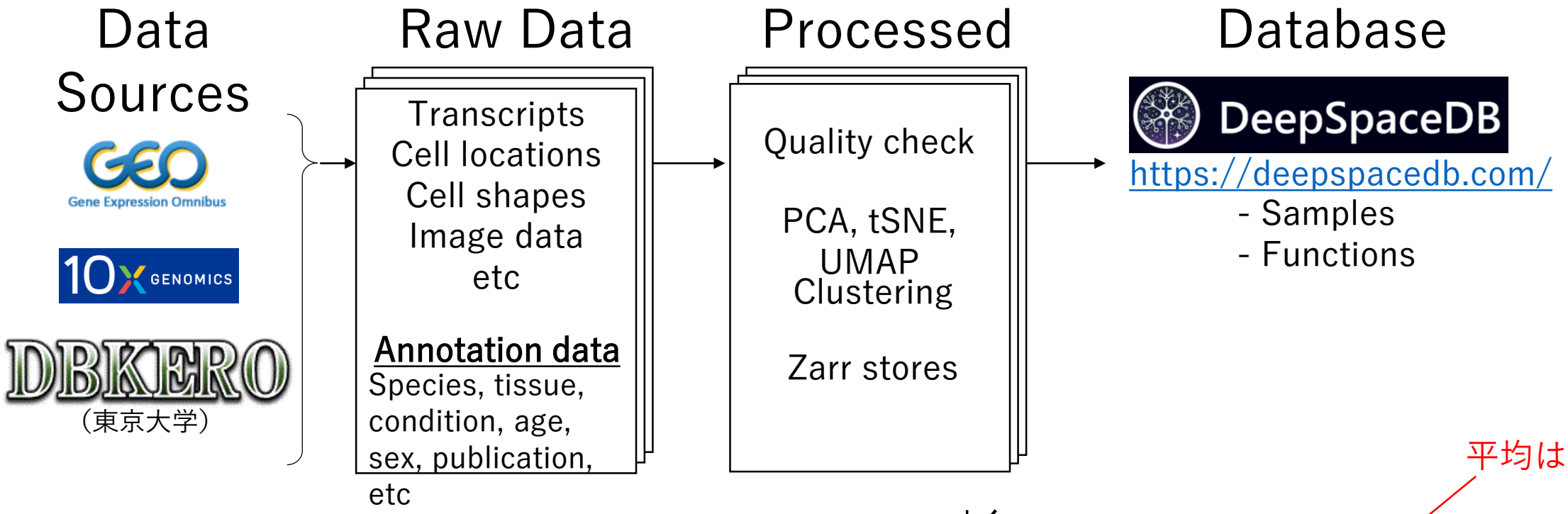
Prox1

Visium

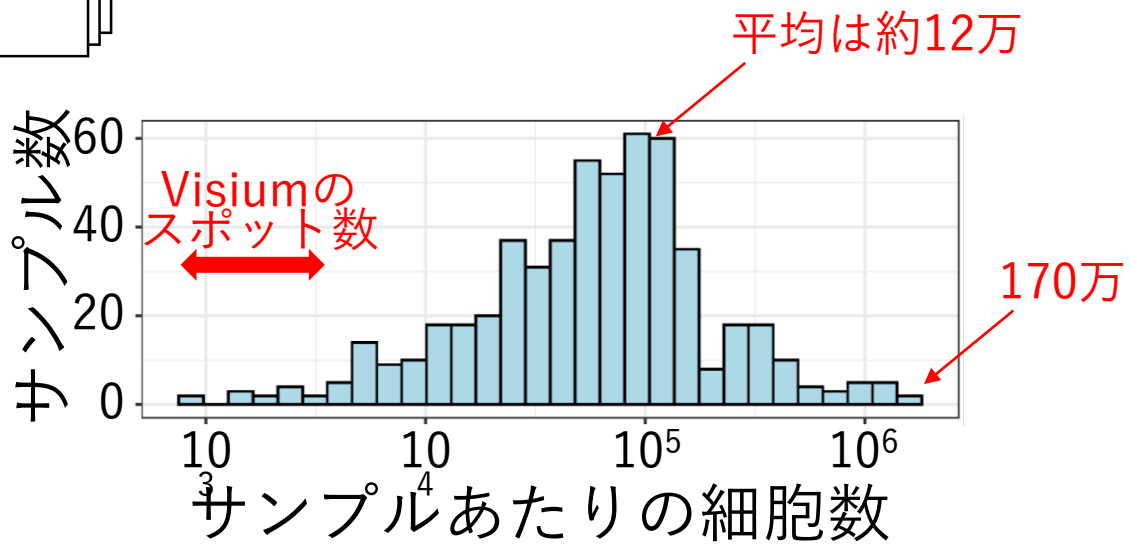
Xenium



Xeniumのワークフロー

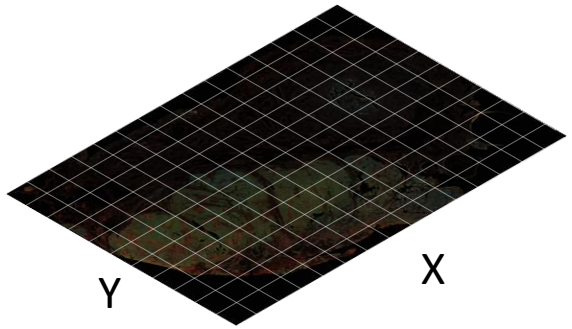


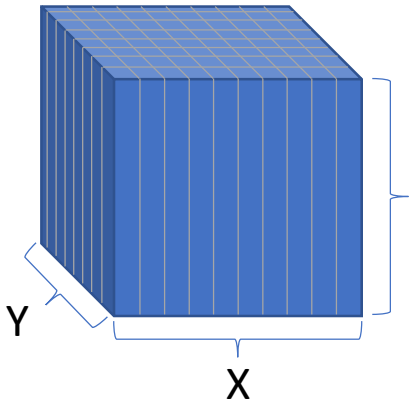
- Xeniumサンプル約600件を処理済み
- **高速アクセス実現のための工夫が必要**



目的ごとに最適化されたフォーマットを採用

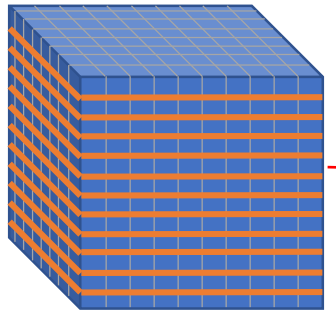
Xenium sample



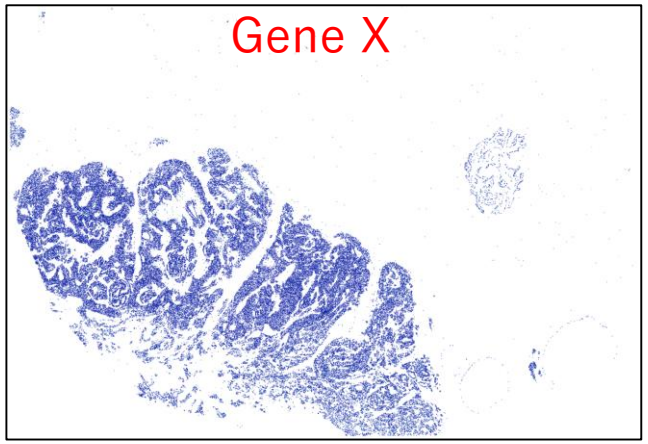


1つの遺伝子の可視化 -> その遺伝子の情報のみ読み込む

Zarr store #1



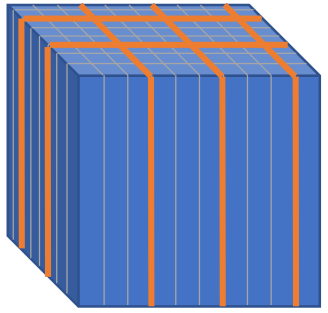
Gene X

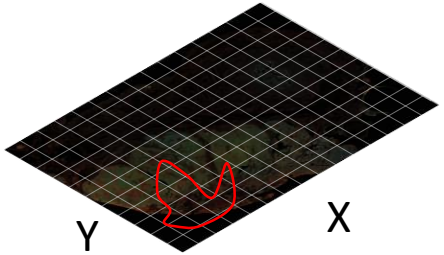


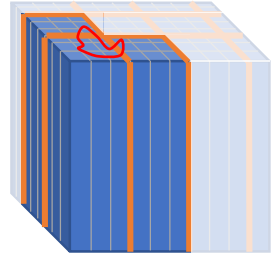
ある領域の全遺伝子の平均値

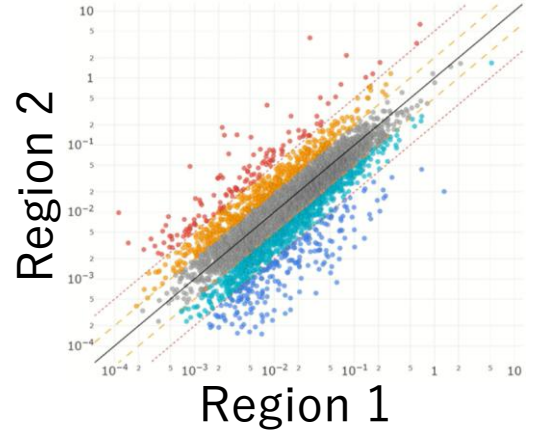
-> その領域の情報のみ読み込む

Zarr store #2

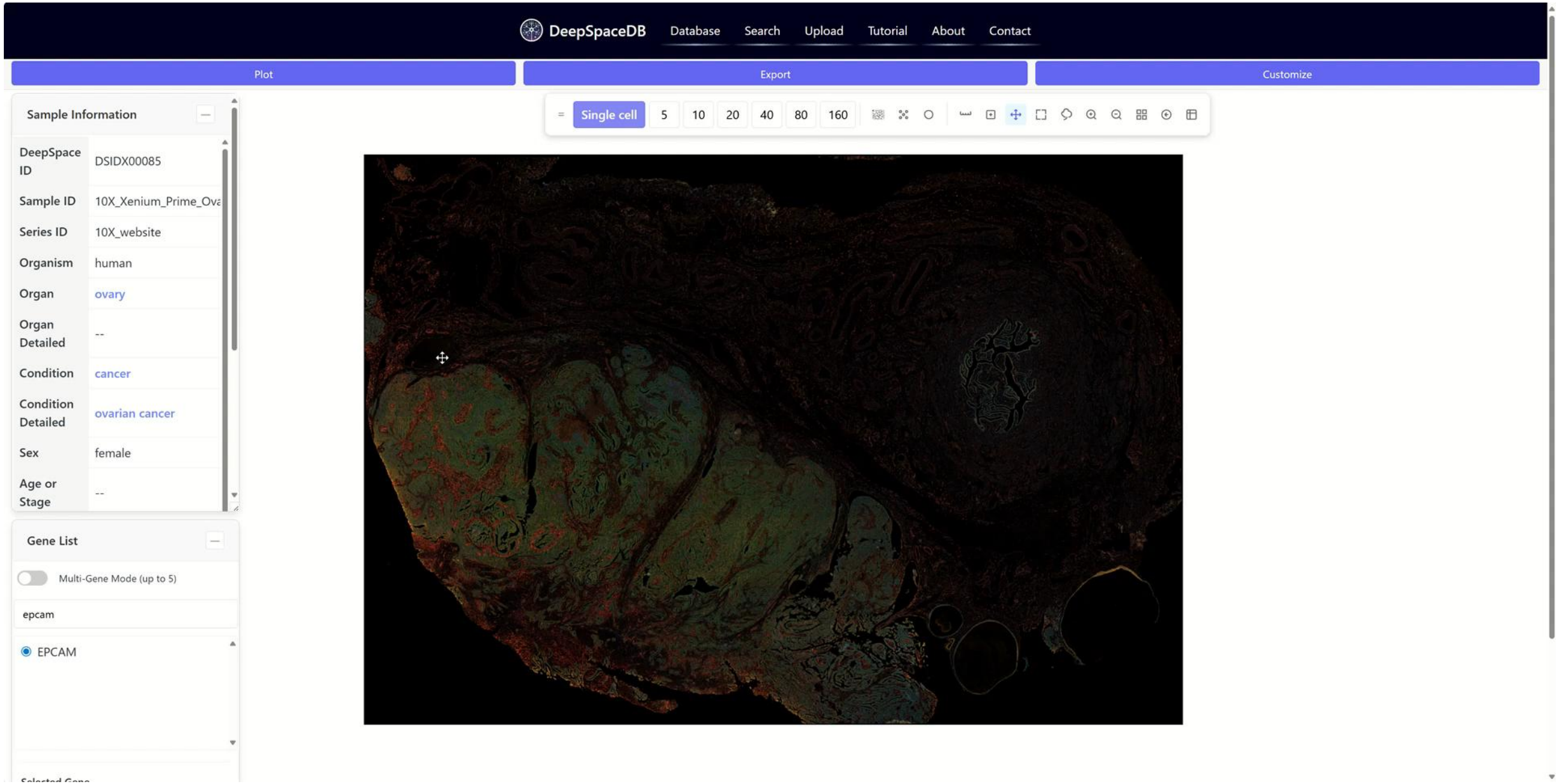




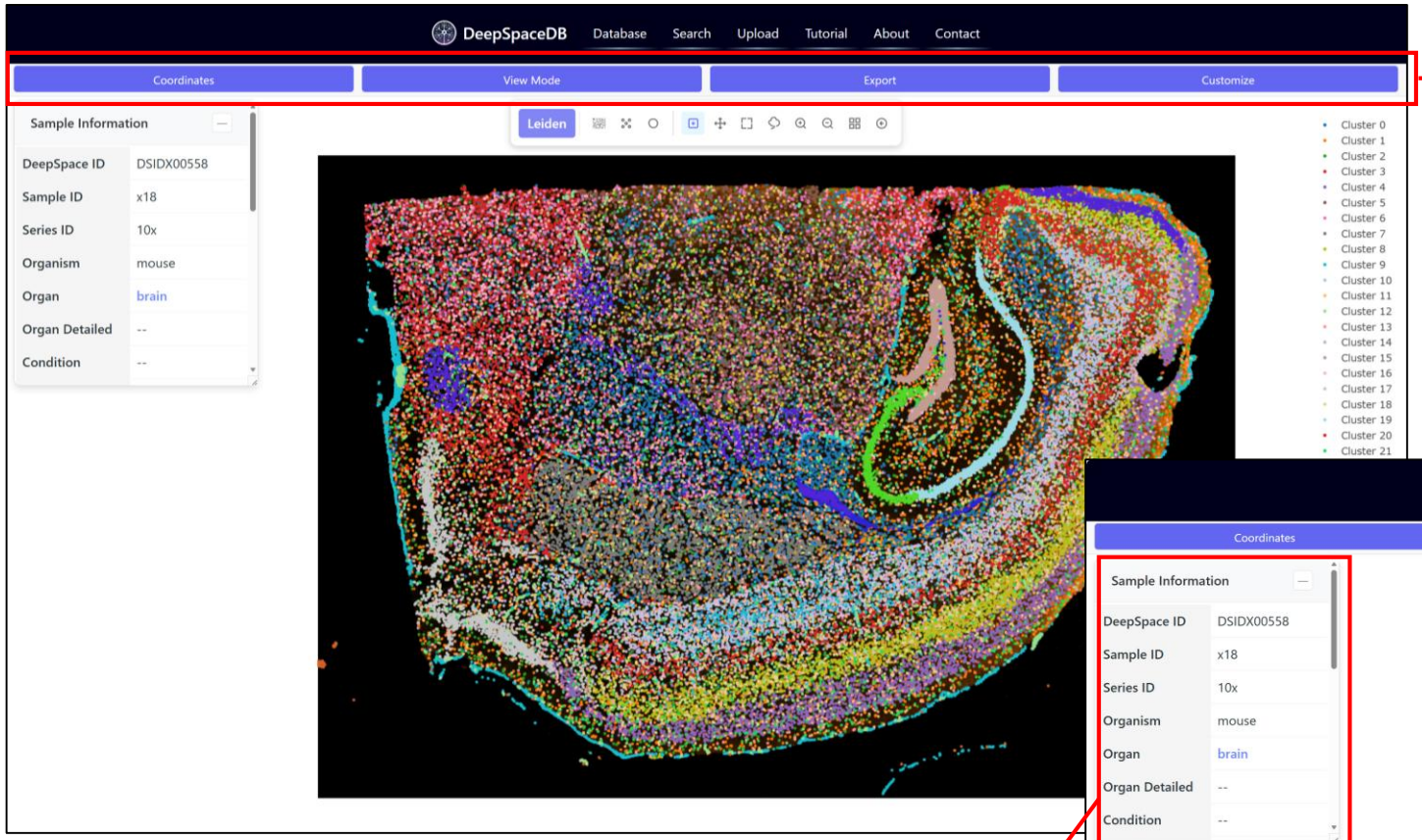




Xenium interface – demonstration (動画)

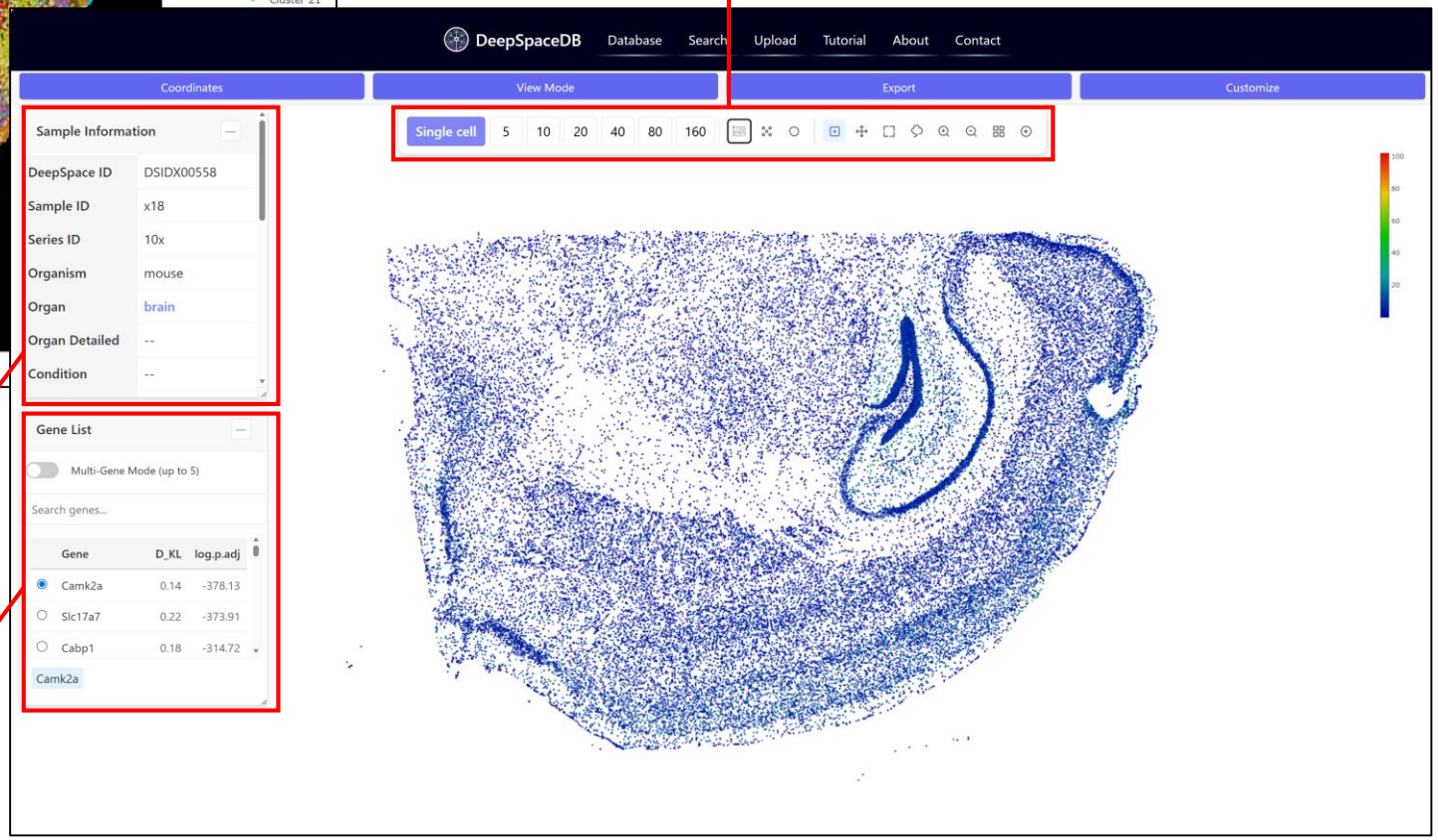


Xenium interface (1)



メニュー:
- クラスタリング
- 品質指標
- 遺伝子発現パターン

多様な操作ツール
（zoom in/out、画像表示切替など）

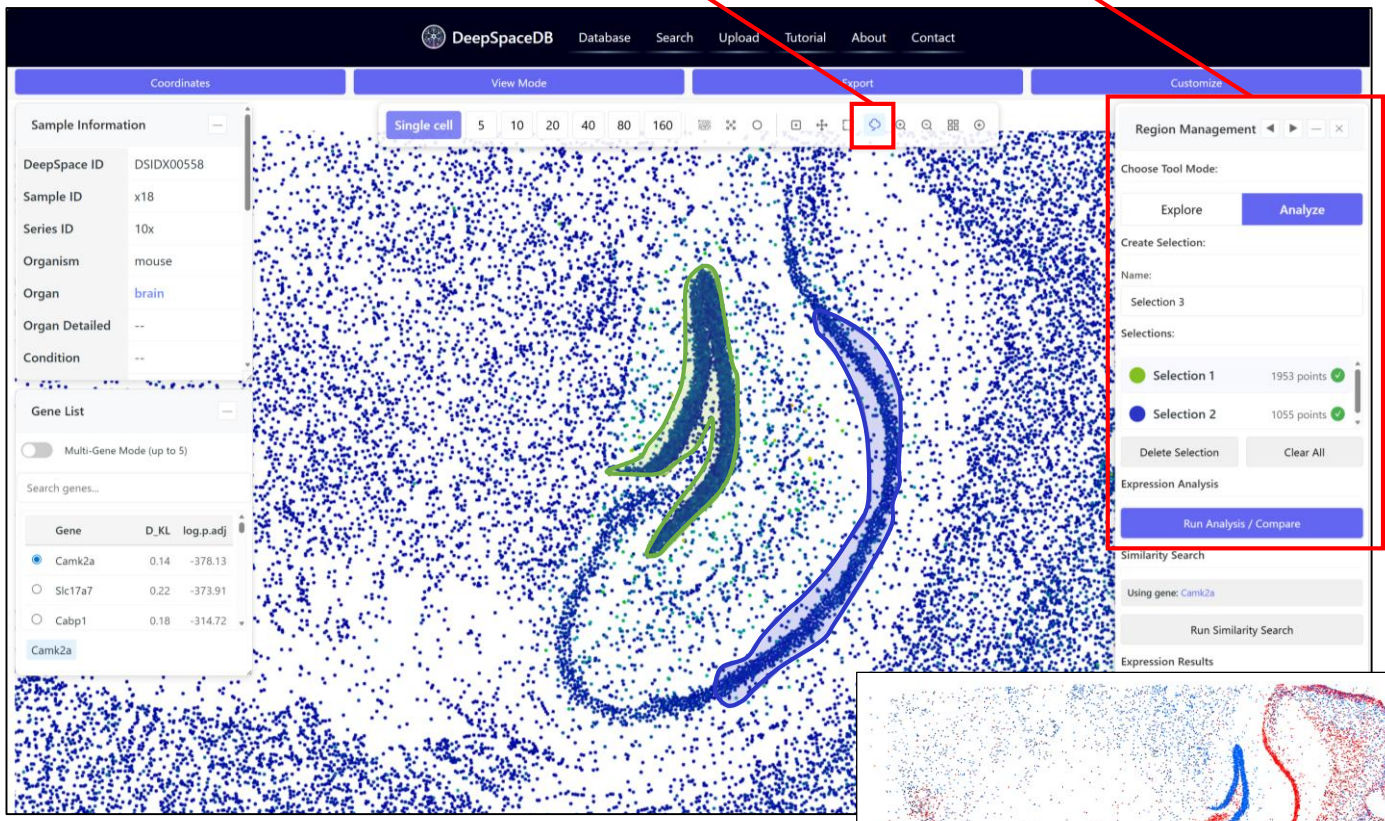


アノテーションデータ

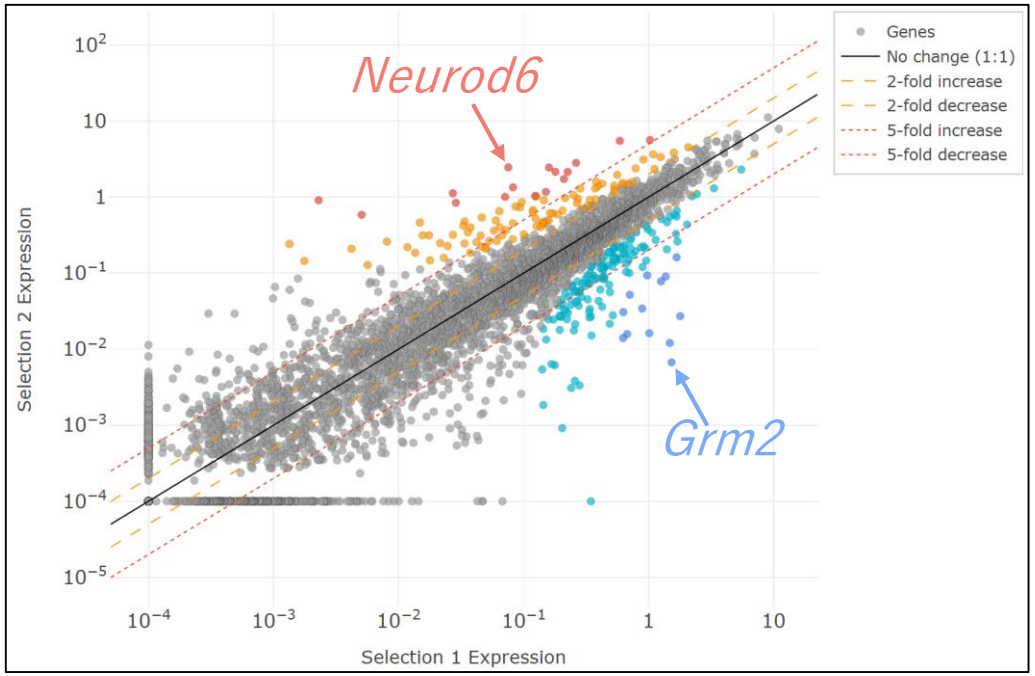
singleCellHaystackを使って空間的変動遺伝子を前もって予測計算している

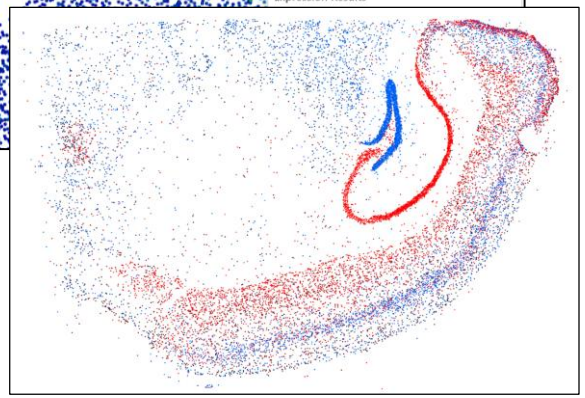
Xenium interface (2)

ユーザーは興味のある領域を自由に選択できる



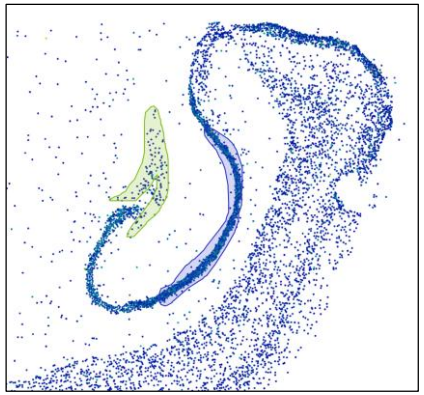
選択した領域間の比較



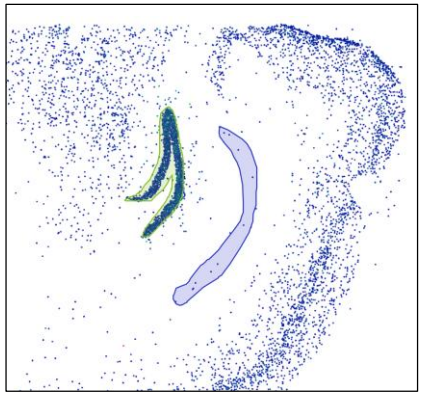


複数の遺伝子発現プロット

Neurod6 *Grm2*



Neurod6



Grm2

概要

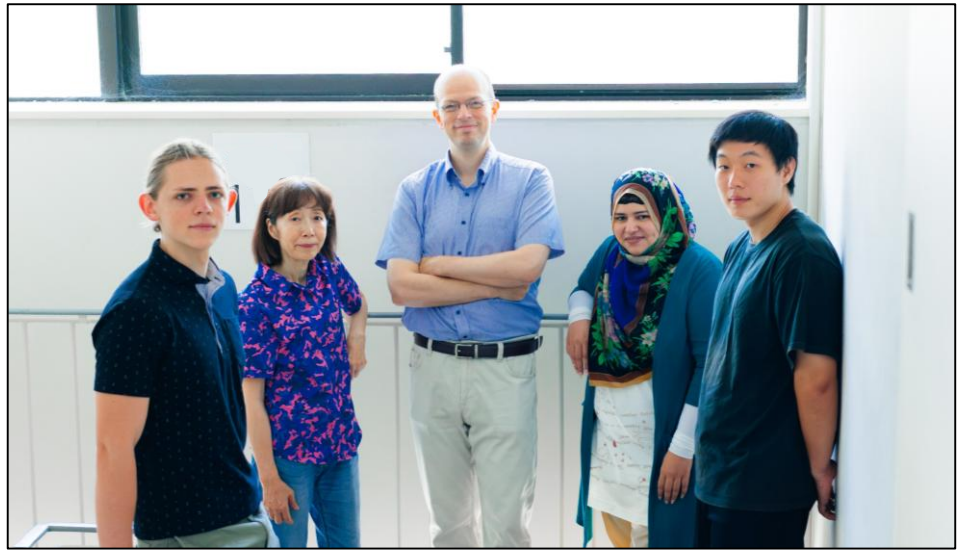
- Rを使った空間トランスクリプトミクスデータ解析
  - Visium example
  - Xenium example
- 既存の空間トランスクリプトミクスデータベースの紹介
- DeepSpaceDBの紹介
  - Visium interface
  - Xenium interface
- まとめ

謝辞

一緒に研究する仲間を募集しています。
学生の皆さん、大歓迎です。
alexisvdb@infront.kyoto-u.ac.jp

Our team

- Vladyslav HONCHARUK
- Afeefa ZAINAB
- TAKEMOTO Keiko
- Hongyi ZHAO



Collaborators & Advisors

- KAWAOKA Shinpei
  (Tohoku Univ & Kyoto Univ)
- SUZUKI Yutaka (Univ of Tokyo)
- HORIMOTO Yoshiya
  (Tokyo Medical Univ Hospital)
- Diego DIEZ (Osaka Univ)

Funding

This work is being supported by
JST NBDC Grant Number JPMJND2303
AMED Grant Number JP24gm2010003

DeepSpaceDBのまとめ

- 空間トランスクリプトームのデータベース構築
- 様々な組織や病歴を含んだ2,000以上の Visium サンプル
- インタラクティブかつ探索的な検索が可能
- 今後の計画:
  - Xenium用のインターフェースの公開
  - 解析ツールの追加
  - 他のプラットフォームも検討

Paper in *Nucleic Acids Research* Database Issue

deepspacedb nar or link