



# Expression Atlas update :from tissues to single cells

---

한국뇌연구원 신경유전학 Lab

# CONTENTS

## PRESENTATION



Introduction

Expression Atlas

Single cell Expression Atlas

Future Directions

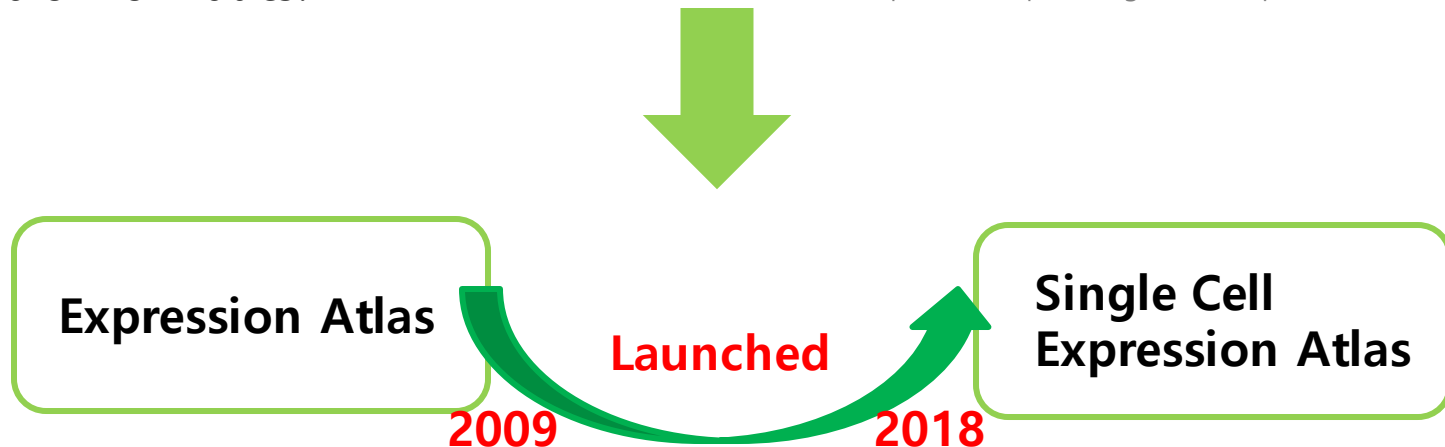


## Introduction



The **European Bioinformatics Institute (EMBL-EBI)** is an International Governmental Organization (IGO) which, as part of the European Molecular Biology Laboratory (EMBL) family, focuses on research and services in bioinformatics.

출처: [https://en.wikipedia.org/wiki/European\\_Bioinformatics\\_Institute](https://en.wikipedia.org/wiki/European_Bioinformatics_Institute)



Expression Atlas is EMBL-EBI's **resource for gene and protein expression** in various biological systems and contexts and provides **open access** to this data for the research community



# Expression Atlas



Expression Atlas is an added-value **online bioinformatics resource** that includes a database, user interface and web-service enabling easy access to information about gene expression across species, tissues, cells experimental conditions and diseases

## Datasets

- NCBI's Gene Expression Omnibus (GEO)
- European Nucleotide Archive
- GTEx
- Gramene
- Pancancer Analysis of Whole Genomes (PCAWG)

## Protein expression information datasets

- PRIDE
- Human Protein Atlas

## Gene Expression datasets

- Open Targets



# Single Cell Expression Atlas



Single Cell Expression Atlas includes 123 studies, including almost a **million assays from 12 different species**.

It enables to providing easy input to **software for further analysis** **gene** of expression matrices, as well as supporting the development of **computational methods** for downstream analyses

## Datasets

: All datasets are sourced from public archives

- Array Express and GEO
- Experimental Factor Ontology terms (EFO)
- Cell Ontology

## Data analysis

- Gene expression quantification

: The pipeline comprises quality filtering, quality trimming and sequencing artifact removal with FASTX-Toolkit, poly-A and uncalled base filtering with fastq-utils and contamination check. And so on.

- Downstream analysis

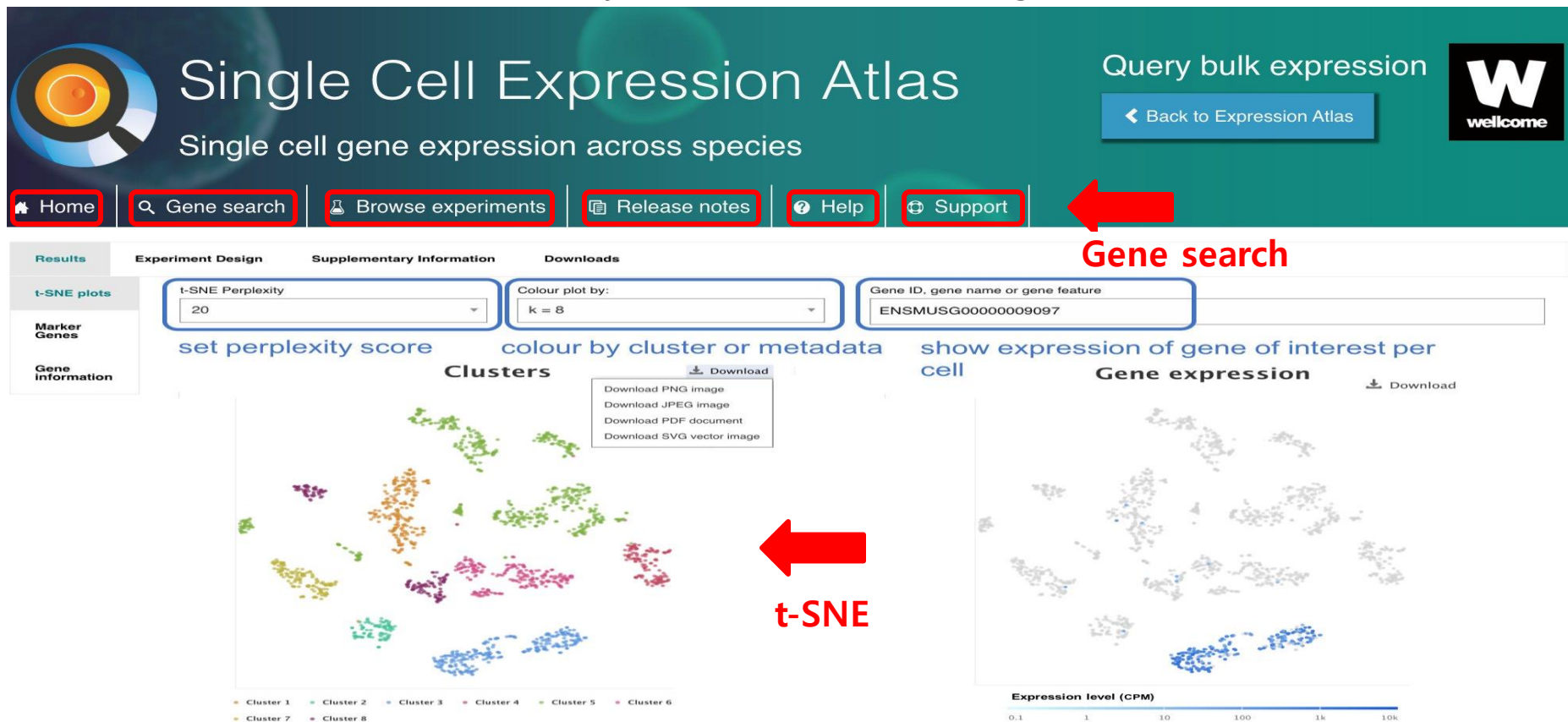
: Final output files are provided in standard file formats where available, including a 10x-like matrix market format for expression matrices.



# Single Cell Expression Atlas

## User interface

- The website **can be searched by genes** within or across species to reveal experiments, tissues and cell types where this gene is expressed or under which conditions it is a marker gene.
- Cells **can be visualized** using a pre-calculated **t-SNE plot** and **can be colored** by different metadata or by cell clusters, based on gene expression





## Future Directions

1. The Expression Atlas **intergrated data from multiple datasets**, in a single interface and processed in a uniform way
2. We will continue to **develop Single Cell Expression Atlas** and will focus on **combining the precision of gene expression** at the single cell level to the breadth of gene expression across a variety of conditions and tissues within the main components of Expression Atlas.
3. We will explore possibilities to **provide a wider range of data visualization** rather than t-SNE plots, including UMAP or PCA to name a few.
4. To address this challenge we will be **working towards seamless integration** of all Expression Atlas components into a single interface.

