


Pipeline Library

Stanome provides a list of pre-configured pipelines evaluated against the model organisms in the Pipeline Library.

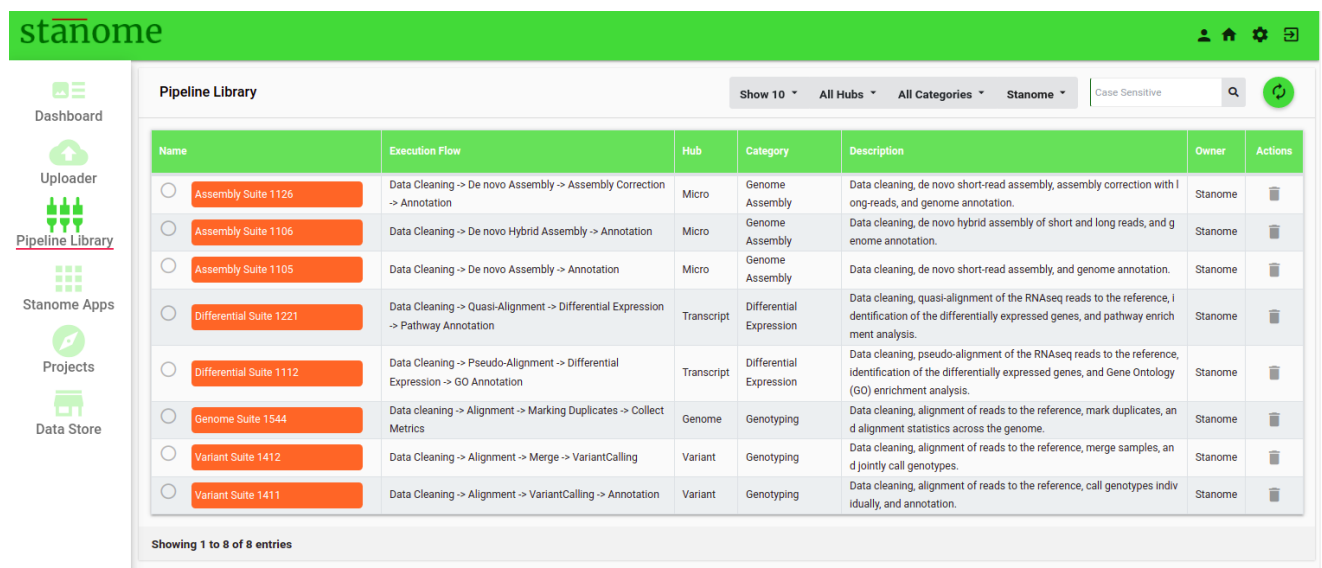
- [Pipeline Library](#)









Pipeline Library

A pipeline is defined as a set of computational tools, which run either sequentially or parallelly to achieve a specific data analysis objective. Bioinformatics tools (and their respective commands) are designated as steps within a pipeline. Stanome provides a list of pre-configured pipelines evaluated against the model organisms in the **Pipeline Library**.

Click the  icon in the Utilities menu to access the **Pipeline Library** window (Fig. 1).

The **Pipeline Library** window shows the list view of all the pre-configured pipelines.




Name	Execution Flow	Hub	Category	Description	Owner	Actions
<input type="radio"/> Assembly Suite 1126	Data Cleaning -> De novo Assembly -> Assembly Correction -> Annotation	Micro	Genome Assembly	Data cleaning, de novo short-read assembly, assembly correction with long-reads, and genome annotation.	Stanome	
<input type="radio"/> Assembly Suite 1106	Data Cleaning -> De novo Hybrid Assembly -> Annotation	Micro	Genome Assembly	Data cleaning, de novo hybrid assembly of short and long reads, and genome annotation.	Stanome	
<input type="radio"/> Assembly Suite 1105	Data Cleaning -> De novo Assembly -> Annotation	Micro	Genome Assembly	Data cleaning, de novo short-read assembly, and genome annotation.	Stanome	
<input type="radio"/> Differential Suite 1221	Data Cleaning -> Quasi-Alignment -> Differential Expression -> Pathway Annotation	Transcript	Differential Expression	Data cleaning, quasi-alignment of the RNAseq reads to the reference, identification of the differentially expressed genes, and pathway enrichment analysis.	Stanome	
<input type="radio"/> Differential Suite 1112	Data Cleaning -> Pseudo-Alignment -> Differential Expression -> GO Annotation	Transcript	Differential Expression	Data cleaning, pseudo-alignment of the RNAseq reads to the reference, identification of the differentially expressed genes, and Gene Ontology (GO) enrichment analysis.	Stanome	
<input type="radio"/> Genome Suite 1544	Data cleaning -> Alignment -> Marking Duplicates -> Collect Metrics	Genome	Genotyping	Data cleaning, alignment of reads to the reference, mark duplicates, and alignment statistics across the genome.	Stanome	
<input type="radio"/> Variant Suite 1412	Data Cleaning -> Alignment -> Merge -> Variant Calling	Variant	Genotyping	Data cleaning, alignment of reads to the reference, merge samples, and jointly call genotypes.	Stanome	
<input type="radio"/> Variant Suite 1411	Data Cleaning -> Alignment -> Variant Calling -> Annotation	Variant	Genotyping	Data cleaning, alignment of reads to the reference, call genotypes individually, and annotation.	Stanome	

Showing 1 to 8 of 8 entries

The list view shows six fields:

- **Name:** Pipeline name
- **Execution Flow:** Sequential order of the tools used in the pipeline
- **Hub:** Name of the relevant Hub
- **Category:** The functional tag(s) associated with the pipeline
- **Description:** Functional details of each tool
- **Owner:** Owner of the pipeline

Details of a pipeline can be viewed by clicking on the pipeline name. In addition to the list view, the pipeline window (Fig. 2) provides additional details about steps and tools. Go to pipeline creation to know more. Click the  icon to go back to the **Pipeline Library** window. Users can't delete pipelines from the **Pipeline Library** but can copy them into individual

projects and customize them according to their needs.

Dashboard

Uploader

Pipeline Library

Stanome Apps

Projects

Data Store

Pipeline Library : 6085189646 , Genome Suite 1544

Name *

Genome Suite 1544

Hub

Genome

Description *

Data cleaning, alignment of reads to the reference, mark duplicates, and alignment statistics across the genome.

Execution Flow *

Data cleaning -> Alignment -> Marking Duplicates -> Collect Metrics

Category

Genotyping

Date *

28/Jul/21 7:09 PM

Details

A basic pipeline for exploring the quality of the whole-genome sequencing data. Sequencing reads are aligned to the reference genome with BWA followed by marking the duplicates with the Picard tool. Several metrics such as coverage, GC bias, and insert sizes are calculated.

Input data:

Reference genome (Fasta)

Whole-genome sequencing data (Fastq)

▼ Pipeline Steps

Step	Name	Tool	Command	Predecessor	Merge	Input Source	Actions
1	DataCleaning	Trimmomatic	Trimmomatic	NA	No	Data Store	⌵ ⌶
2	Alignment	BWA	BWAAligner	1	No	1	⌵ ⌶
3	Sam2bam	Samtools	Sam2Bam	2	No	2	⌵ ⌶