



## scVelo.ComputeVelocity Documentation

**Description:** Perform single-cell RNA-velocity analysis using scVelo

**Author(s):** Theis Lab, Institute of Computational Biology, Helmholtz Zentrum München.  
Wrapped as a module by Anthony S. Castanza and Barbara A. Hill Mesirov Lab,  
UCSD School of Medicine.

**Contact:** [genepattern.org/help](http://genepattern.org/help)

**Summary:** “RNA velocity enables the recovery of directed dynamic information by leveraging splicing kinetics. scVelo generalizes the concept of RNA velocity (La Manno et al., Nature, 2018). by relaxing previously made assumptions with a stochastic and a dynamical model that solves the full transcriptional dynamics. It thereby adapts RNA velocity to widely varying specifications such as non-stationary populations.” [\(source\)](#) for more about scVelo see [the scVelo about page](#).

This module implements the basic, dynamical modeling, and differential kinetics scVelo analysis pipelines downstream of Scanpy preprocessing.

**Source Publication:** Bergen, V., Lange, M., Peidli, S. et al. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol 38, 1408–1414 (2020). <https://doi.org/10.1038/s41587-020-0591-3>

See: <https://scvelo.readthedocs.io/references/> for full citation information.

### Parameters:

Name	Description
h5ad*	Anndata specification h5ad file, this should contain data preprocessed with Scanpy that also contains the velocity required “spliced” and “unspliced” data layers.
Marker Genes	A simple text file containing a list of genes that will be plotted on the dataset results. Should be given as one gene per line.
Velocity Mode*	Mode for performing the velocity estimation. Supports 'deterministic', 'stochastic' (default), or 'dynamical' methods for computing splicing kinetics. ( <a href="#">See the full scVelo documentation for details</a> )
min_shared_counts*	Filter genes by minimum shared counts across cells
n_top_genes*	Top N variable genes to use for velocity calculation
recalculate_highly_variable_genes*	If highly_variable_genes have already been calculated externally, force their re-computation with Scanpy using the selected method. If set to 'false' and the highly_variable_genes key isn't found in the data, the calculation will be automatically performed using the Seurat_v3 method.  Selecting options other than 'false' will force re-calculation of PCA, the KNN graph, and UMAP embeddings if they were previously calculated by Scanpy.
Enforce Normalization*	Normally, the top level counts matrix used by most tools such as scanpy and seurat will already be normalized. scVelo requires that the separate spliced and unspliced counts layers used only by velocity calculations also be normalized. Setting Enforce.Normalization to "all" will cause scVelo to force normalization of all three layers of the data. Setting Enforce.Normalization to "velocity" (default) will only

# GenePattern

	normalize the velocity specific spliced and unspliced layers. Setting Enforce.Normalization to "none" will use the dataset as is for all layers. Note: selecting "all" will force re-calculation of PCA, the KNN graph, and UMAP embeddings.
N_PCs*	Number of principal components used in calculation of the gene moments of spliced/unspliced abundances.
N_Neighbors*	Number of nearest neighbors used in calculation of the gene moments of spliced/unspliced abundances.
Differential Kinetics Analysis*	Perform differential kinetics analysis on the basis of dataset clustering using dynamical velocity estimations.
Embedding*	Select the type of embedding used in ScanpyUtilities. Supports UMAP and tSNE. If the specified embedding is not present in the dataset, or any selected options require the rerunning of lower-dimensional embedding this option will be used.
Clustering*	Autodetect existing clusters in the dataset, or run clustering using choice of leiden or louvian algorithms. If no clustering is detected with "autodetect_existing", "leiden" will be used.
Cluster Resolution*	Resolution to use if running leiden or louvain clustering on the dataset.
Regression Keys*	Any keys to regress out of the dataset (i.e. batch) using scanpy.pp.regress_out(). Should be specified as a non-quoted comma separated list.  Regressing out a key will force re-calculation of PCA, the KNN graph, and UMAP embeddings.
Plot Batches*	If a "batch" key is given in the data, "true" will output individual velocity plots for each batch in the dataset (if present).
Save Plot Type*	Save velocity plots as png (default) or svg (or pdf; not recommended due to pdf element count limitations).
output file base*	Output basename for velocity calculation results. By default the outputs will be named based on the input h5ad file.

\* = required

**Output File(s):** An .h5ad file containing the full results of the scVelo pipeline, along with major plots from the selected analysis (.png format by default), and tab delimited text files containing numerical values underlying some of the computations. Plots will be named based on the specific analysis options selected, and separate plots will be produced for the differential kinetics analysis mode if enabled.

**Module Language:** Python 3

**Source Repository:** <https://github.com/genepattern/scVelo.ComputeVelocity/tree/v0.10>

**Docker image:** genepattern/compute\_scvelo:beta1

Version	Date	Comment
0.10	2022-XX-XX	Initial beta release.