User Manual for BASSAA v0.4.0

Introduction

Welcome to the BASSAA (B27 Ankylosing Spondylitis and Acute Anterior Uveitis Atlas) platform! To uncover pathological signatures in scRNA-seq data, BASSAA employs the Sparse Decomposition of Arrays (SDA) algorithm, a soft-clustering method that identifies transcriptional signatures as components. These components provide weights for genes (loadings) and rankings for cells (scores), revealing patterns of biological and pathological relevance. This manual highlights key findings from SDA analysis while offering an interactive interface for exploring corroborated biology and disease pathology. Explore the platform through our interactive website and R package at https://eisascience.shinyapps.io/BASSAA/. This application integrates Sparse Decomposition of Arrays (SDA) and other powerful tools to provide insights into biological patterns, gene expression, and cell-type-specific information.

BASSAA offers: - Interactive visualization tools. - Flexible selection of data subsets. - Integration of curated datasets (e.g., Mahyari-Gill). - Predefined SDA components representing biologically meaningful signatures.

Use BASSAA to explore gene expression, enrich insights into cell populations, and visualize significant patterns in your data.

Getting Started

Home Page

Welcome to BASSAA! This platform serves as a gateway to understanding the immune mechanisms underlying Ankylosing Spondylitis (AS) and Acute Anterior Uveitis (AAU). These diseases are characterized by chronic inflammation and systemic involvement, affecting joints, eyes, and other tissues. Both conditions share a strong genetic association with HLA-B27, a molecule implicated in immune regulation and dysbiosis of gut microbiota. While AS primarily impacts the musculoskeletal system, AAU manifests as painful ocular inflammation. The co-occurrence of AS and AAU underscores the need for high-resolution immune profiling to unravel their shared and distinct pathophysiologies.

Using single-cell CITE sequencing, BASSAA identifies unique transcriptomic and immunophenotypic differences across AS, AAU, and HLA-B27 positive and negative healthy controls (HC). Highlights include:

- Dysregulated immune pathways driven by IL-17/23, TNF, and IFNg in AS and AAU.
- Distinct immune signatures such as B-cell maturation defects in AS and cytotoxic TNK cell enrichment in AAU.
- Common myeloid activation across AS and AAU, with impaired homeostatic function in monocytes.
- Validation of these findings using independent datasets.

Dive into the Main Tab to explore SDA-derived components that reveal these intricate immunophenotypes and enhance our understanding of AS and AAU.

Main Tab: Features and Usage

The Main Tab is the core workspace of BASSAA. It is divided into multiple panels, each serving a specific purpose in the analysis workflow.

Visualization Inputs

- **2D Reduction**:
    - Select dimensionality reduction methods (e.g., UMAP).
    - Example: "UMAPintg" is a previously computed integrated UMAP using the Mahyari-Gill data as a reference, leveraging rPCA, an anchor-based method, to align the cells between the two datasets.

- **Data Origin**:
  - Choose the data source, such as "Mahyari-Gill" (which includes HLA-B27 information and stratification for AAU, AAU+AS, AS, and HC patients) or "GSE194315" (a validation dataset lacking HLA-B27 expression and only including AS and HC patients).
- **Metadata Selection**:
  - Focus analysis on metadata categories, such as:
    * Major Cell Types
    * Immunophenotypes
    * Donors
    * HLA-B27

Cell Selection Inputs

- **Condition**:
  - Filter data based on experimental conditions, such as "HC only" or "AS_AAU HC".
- **Major Cell Type**:
  - Explore specific cell populations (e.g., TNKs, Myeloids, B cells).

SDA Inputs

- **SDA Components**:
  - Select a predefined component, each representing key immunological signatures derived from SDA analysis:
  - **B-Cell Defect Signature (BCDS)**: Distinguishes mature, class-switched B cells (positively scored) from naive or early-stage B cells (negatively scored). Higher scores indicate adaptive immune activity, while lower scores highlight early-stage B cell presence.
  - **Myeloid Regulation Imbalance (MRI)**: Reflects an imbalance in myeloid regulation. Positive scores highlight active immune/inflammatory response genes associated with dendritic cells (DCs) and CD16+ monocytes, while negative scores emphasize homeostatic roles of CD14+ monocytes, typically enriched in healthy controls.
  - **pDC-Monocyte Differential Signature (PMDS)**: Differentiates pDCs and DCs (positively scored) from monocytes (negatively scored). This imbalance highlights active inflammatory responses in AS and AAU, with monocyte enrichment and DC depletion.
  - **TNK Cytotoxic Module (TCM)**: Captures cytotoxic TNK cell activity. Positive scores indicate active cytotoxic responses (e.g., effector TNK cells), while negative scores represent naive and central memory T cells. Enrichment patterns distinguish AAU and HLA-B27pos HC from AS and HLA-B27neg HC.
- **Score Threshold**:
  - Set a numerical threshold to filter SDA scores and focus on significant patterns.
- **Gene Output**:
  - Specify the number of genes to retrieve (default: 30).
  - Use "CopyPosGenes" or "CopyNegGenes" buttons to copy gene lists.
  - Download the gene list using "DownloadGenes".

Gene Inputs

- **Gene Search**:
  - Enter gene names to visualize expression (e.g., "S100A11").
- **Expression Threshold**:
  - Filter cells by gene expression levels.
  - Example: Set a threshold of 0.1 to focus on cells with significant expression.

Visual Outputs

Chromosomal Location

- **Plot**: Displays genomic coordinates with gene loadings.
- **Download**: Save the visualization using the "Download" button.

Cell Score (Ridge)

- **Plot**: Visualizes the distribution of SDA scores across metadata categories.
- **Use Case**: Compare SDA signatures across conditions or cell types.

Enrichment Heatmap

- **Plot**: Shows the enrichment of cells with SDA scores above the threshold, calculated using a chi-squared test.
- **Scaling**: The residuals from the chi-squared test are scaled using the inverse hyperbolic sine (asinh) transformation to normalize and enhance interpretability.
- **Importance**: This approach highlights significant deviations from expected values, allowing for a clearer understanding of cell enrichment patterns across conditions. The asinh transformation ensures that both small and large residuals are effectively visualized without distortion, facilitating comparisons between enriched and depleted cell populations.

GO Analysis (Positive/Negative)

- **Positive GO Terms**: Highlights biological processes positively correlated with SDA components.
- **Negative GO Terms**: Identifies processes negatively correlated with SDA components.
- **Bi-Directional Insights**: For each SDA component, we sort the top positive and negative genes based on their loadings. Each direction is biologically informative and can reveal distinct, unrelated processes. Positive and negative loadings may not always correlate but instead highlight complementary or contrasting biological pathways.

2D Visualizations

- **Cell Score (2D)**: Overlays SDA scores onto the precomputed rPCA integrated UMAP, providing insights into the spatial distribution of component scores.
- **Gene Expression (2D)**: Maps specific gene expression levels onto the same precomputed UMAP, enabling visualization of expression patterns across the dataset.

Metadata Clustering

- **Metadata (2D)**: Visualizes the precomputed rPCA integrated UMAP with the selected metadata categories overlaid (e.g., cell types, conditions).

Outputs

Gene Tables

- **Positive/Negative Genes**:
  - Tables of top-loaded genes for SDA components.
  - Download options available for both tables.

Tips for Exploration

- Experiment with different combinations of cell types, conditions, and SDA components to uncover hidden patterns.
- Use default settings as starting points but adjust thresholds and filters to suit your research questions.
- Visualize and download plots for presentations or further analyses.

Citing BASSAA

To acknowledge BASSAA in your research, please use the following citation:

Mahyari, E., Davin, S., Ogle, K., Fale-Olsen, E., Shaut, C., Martin, T.M., Ahuja, J., Suhler, E., Deodhar, A., Rosenbaum, J., & Gill, T. (In Review). "Divergent B cell and cytotoxic TNK cell activation signatures in HLA-B27-associated ankylosing spondylitis and acute anterior uveitis."

For more information, contact the corresponding author:

Tejpal Gill, Ph.D.
Oregon Health and Science University

Troubleshooting

- **Errors Loading Data**: Ensure the selected SDA run is valid.
- **Slow Performance**: The application is optimized for the Mahyari-Gill dataset with curated top 4K genes for faster load times.
- **Missing Outputs**: Verify threshold settings and ensure they are within valid ranges.

For additional support, contact the development team through the "Gill Lab" or "@eisamahyari" links in the sidebar.

Acknowledgements

Author Contributions

Conceptualization: T.G., J.T.R., and J.A.

Subject recruitment: A.D., E.S., T.M.M., K.O., E.F.O.

Data collection: S.D., C.S., K.O., E.F.O.

Analysis: E.M., T.G., and J.A.