# User Manual for TISCalling Prediction in JBrowse

## **Basic usage of JBrowse**

## 1. Linear genome view

To open a Linear Genome View (LGV), go to the menu bar and select:  $Add \rightarrow Linear$ Genome View. Choose an available genome to display TIS prediction results. Currently, supported genomes include *Arabidopsis thaliana* and *Solanum lycopersicum*.

## 2. Available Tracks of TIS-prediction scores and Ribo-seq

To open the **Available Tracks** menu, click the **open track selector button** at the bottom of the **Linear Genome View (LGV)** panel. The menu will appear on the right-hand side, listing available tracks such as:

- Reference sequence
- Gene annotation
- TIS prediction score
- Ribo-seq data collected from different resources

To display a track, check the arrow ( $\blacktriangleright$ ) to show the available datasets and click the box next to it to display the tracks in the **LGV panel**.

## 3. Scrolling

- Use the mouse wheel or click and drag to scroll horizontally.
- Alternatively, use the **left and right pan buttons** in the **LGV header** to scroll in the respective directions.

## 4. Zooming

• Use the **zoom buttons** and **slider bar** in the **LGV header** to zoom in and out.

## 5. Reordering Tracks

• Click and drag the **drag handle** (indicated by six vertical dots) on the left side of the track labels to reorder tracks.



Figure 1 Basic usage of Jbrowse. 1. Select Add  $\rightarrow$  Linear Genome View to open a new LGV. 2. Click the track selector button to display available tracks. 3. Use the Pan buttons to scroll left or right. 4. Use the Zoom buttons or slider to adjust the zoom level. 5. Tracks can be reordered by clicking and dragging the drag handle (indicated by six vertical dots).

## **Using the Location Search Box**

The Location Search Box is located at the top of the Linear Genome View (LGV).

You can search for a location in multiple ways by typing in the search box:

• By region and coordinates: e.g., chr1:1..100, chr1:1-100, or chr1 1 100

≡ > Araport11			
Chr1 >>>>5M >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	20M 2	5M > > > > 3d
	AT5G16820 Q 🛛	2.8Кbp Q	Q
Chr1 12,974,000 12,974,400 12,9	AT5G16820	5,600 12,976,000	12,976,400
$ \begin{array}{c} \vdots \times \text{Araport11_201606} \\ \vdots \\ \vdots \times \text{At 5UTRATG LR} \\ \vdots \\ \vdots \times \text{At CDSATG LR} \\ \vdots \\ \end{array} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \end{bmatrix} $	AT5G16820.1 AT5G16820.2 HSF3:five_prime_UTR:1 (AT5G16820:five_prime_UTR:1) HSF3:five_prime_UTR:2 (AT5G16820:five_prime_UTR:2) HSF3:three_prime_UTR:1 (AT5G16820:three_prime_UTR:1)	AT1TE42305	ATITE42315 plus minus plus minus

• By gene name or feature keyword: e.g., AT5G16820

Figure 2: Users can search for gene names or other features using the Location Search Box.

## **Available Tracks**

### **Reference Sequence Track**

The **Reference Sequence** track displays the plus strand or the minus (reverse complement) strand nucleotide sequences and the corresponding amino acid sequences (in six reading frames) of the genome, with its visibility depending on the zoom level:

- At the **kilobase-pair range**, the track highlights **start codons** (red) and **stop codons** (green) across different reading frames.
- At the **100-base-pair range**, the nucleotide sequence becomes visible.

#### Annotation Track

The **Annotation** track shows the locations of genomic features, including **exons, coding sequences (CDS),** and associated **gene IDs**. Arrows indicate the transcription direction of each gene.

For better visualization:

- Exons are typically represented as blue blocks.
- CDS regions are shown as yellow blocks, visually distinguishing coding sequences.
- Introns are often depicted as connecting lines between exons.

#### **TIS Prediction Score Tracks**

The **TIS Prediction** <u>Score</u> tracks include the scores computed by four TIS-type models: **5'UTR ATG, 5'UTR non-ATG, CDS ATG, and CDS non-ATG**.

- Colors indicate the strand direction that the predicted site is read in:
  - **Red** represents predictions on the **plus strand**.
  - Blue represents predictions on the minus strand.
- **Prediction scores** range from **0 to 1**.

### **Ribo-Seq Tracks**

When available, the **Ribo-Seq** track displays **ribosome binding signals**, providing insights into translation activity.

- The individual sample names re displayed.
- Values are normalized to CPM (Counts Per Million) to ensure comparability across samples.
- **Colors** indicate the predicted site location:
  - **Red** represents predictions on the **plus strand**.



• Blue represents predictions on the minus strand.

Figure 3. LGV displaying multiple tracks. The tracks include Reference Sequence, Annotation, and TIS Prediction. Additionally, Ribo-Seq tracks are displayed if available.