

ORIS User Manual

Urminder Singh

Table of Contents

Getting Started.....	2
Prerequisite.....	2
Installation	2
Running ORIS.....	2
Origin Finder	3
Skew Methods	4
Skew Calculator.....	5
Auto-Correlation Method	5
Z-curve	5
Automatic Prediction	5
Pattern Search.....	6
Search Sequence	6
DnaA box and Yeast ACS	7
Weight Matrix Search	7
Extract Sequence	8
DNA Analysis	8
DNA Bending Analysis	8
Information	9
Shannon's Entropy	9
Renyi's Entropy	10
Binary Entropy	10
Redundancy Plot	10
Composition.....	10
Help.....	10
ORIS Settings.....	10
User Guide	11
License.....	12

Getting Started

ORIS is a freely available and open source Java software tool. Latest version of ORIS can be downloaded from <https://github.com/urmi-21/ORIS/releases>.

Prerequisite

- Supported Operating Systems: Windows, Linux and Mac
- Java 7 or higher

Installation

The .jar file downloaded from <https://github.com/urmi-21/ORIS/releases> is the ORIS program, which can be executed directly. No further installation is required.

Running ORIS

To execute ORIS program, double click on the .jar file and the ORIS main window should open. Advance users can open terminal and type the following command to execute the jar file:

Java -jar oris.jar

After successfully executing the above the step, The ORIS main window will be displayed (Figure 1).

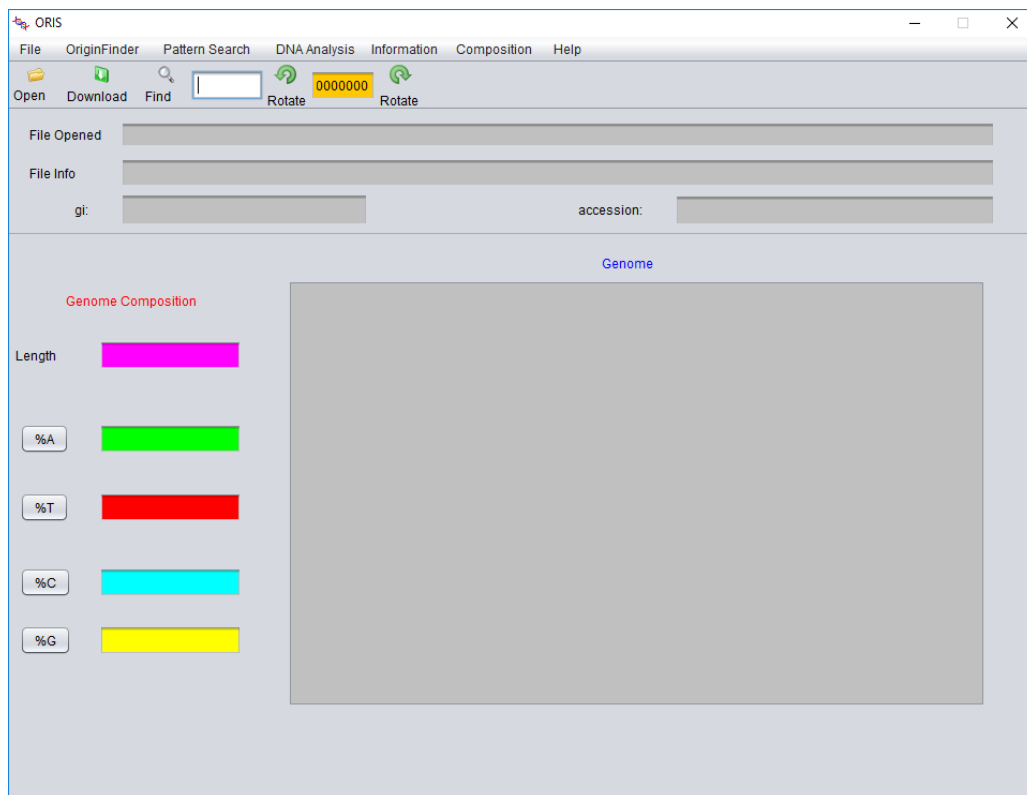


Figure 1 Main Window of ORIS

Then, User must read a valid genome file in fasta format stored as plain txt file. To do this click “Open” button and locate the fasta file. The file will be loaded and the sequence data will be displayed in the main window. One can also download the required whole genome sequence from NCBI’s FTP server by clicking the “Download” button. Once the file is downloaded, it can be read with the “Open” button. Now, the user can proceed to do interactive analysis with ORIS in order to find putative replication origin sites in the whole genome sequence (Figure 2).

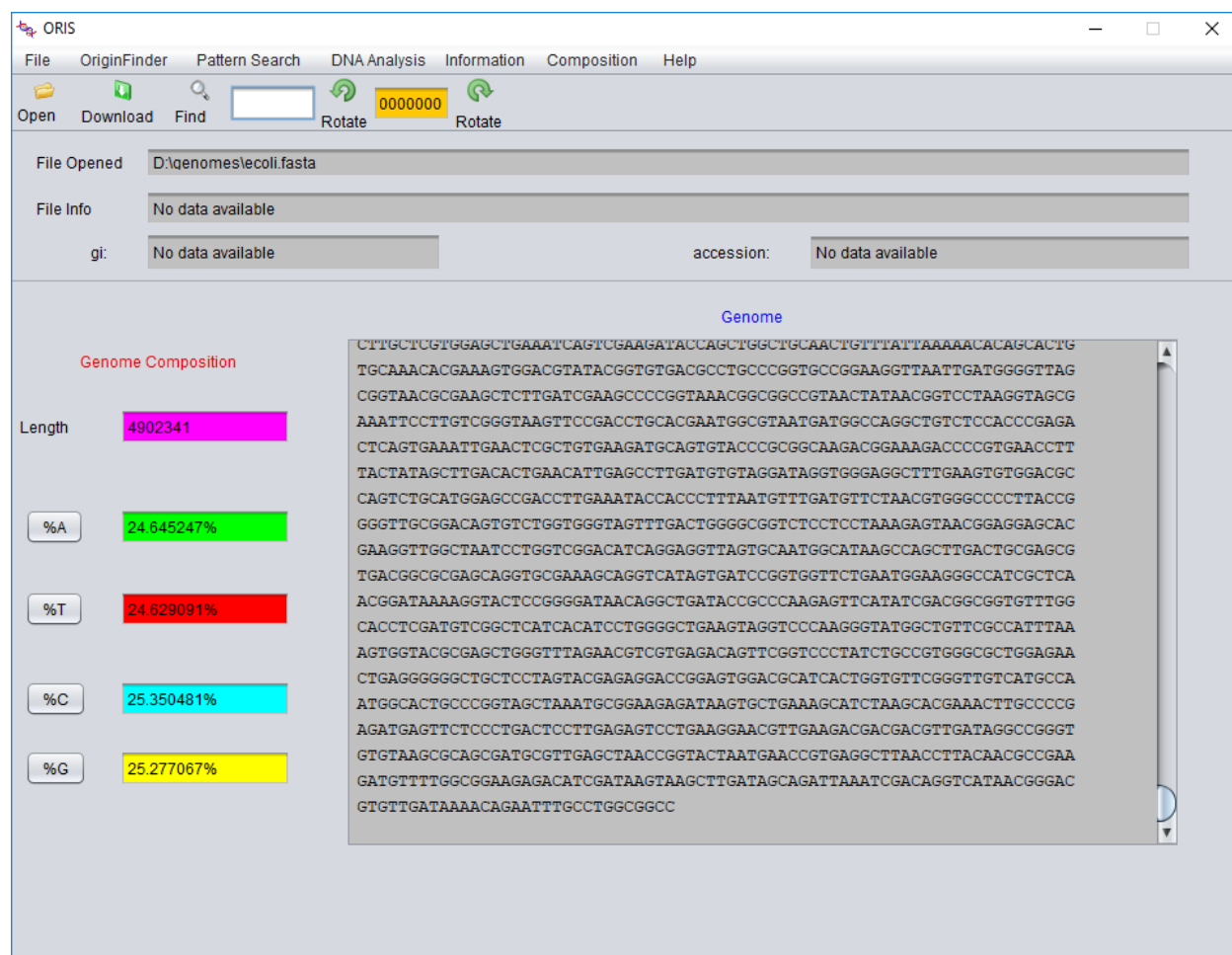


Figure 2 ORIS after data is loaded

The rest of this document describes how to access various methods implemented in ORIS, which allows to search for putative origin of replication sites in an interactive manner through ORIS’ GUI. To read about the methods in detail and the interpretation of the plots, please see the *Supplementary Document* at <https://github.com/urmi-21/ORIS/>.

Origin Finder

The origin finder menu has methods to find origin of replication like regions in the genome using the sequence composition. To access origin finder menu, go to Menubar → OriginFinder (Figure 3).

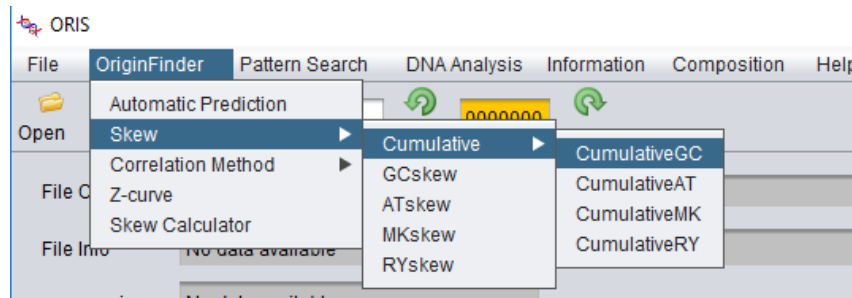


Figure 3 Accessing Skew Methods in the OriginFinder menu

Skew Methods

Skew methods menu contains methods like GC-Skew, AT-Skew, MK-Skew etc. To access this menu, go to Menubar → OriginFinder → Skew (Figure 3).

After clicking an option (e.g., GCskew) a new window will be displayed. This window will allow the user to enter the window size and increment parameters. *These parameters are required for any sliding-window analysis method and thus will be often displayed when using ORIS as most of the methods are based on a sliding-window analysis scheme (Figure 4).*

 The image shows a dialog box titled 'Choose Window-size and Increment'. It contains the following fields and options:

- 'Window Size' text box with the value '50000'.
- 'Increment' text box with the value '10000'.
- A checkbox labeled 'Save plot data to text file' which is currently unchecked.
- 'Enter File name' text box.
- A horizontal separator line.
- A checked checkbox labeled 'Use Whole Genome'.
- 'Start pos' and 'End pos' text boxes, both currently empty.
- A note below the boxes: '*Position index starts from 1'.
- A 'Go' button at the bottom.

Figure 4 Window to input the window size and increment parameters.

Skew Calculator

To allow users to enter their own formula, we have implemented skew calculator option in ORIS. To access this menu, go to Menubar → OriginFinder → Skew Calculator. A new frame will be displayed, which provides an interface to enter a custom mathematical formula (Figure 5).

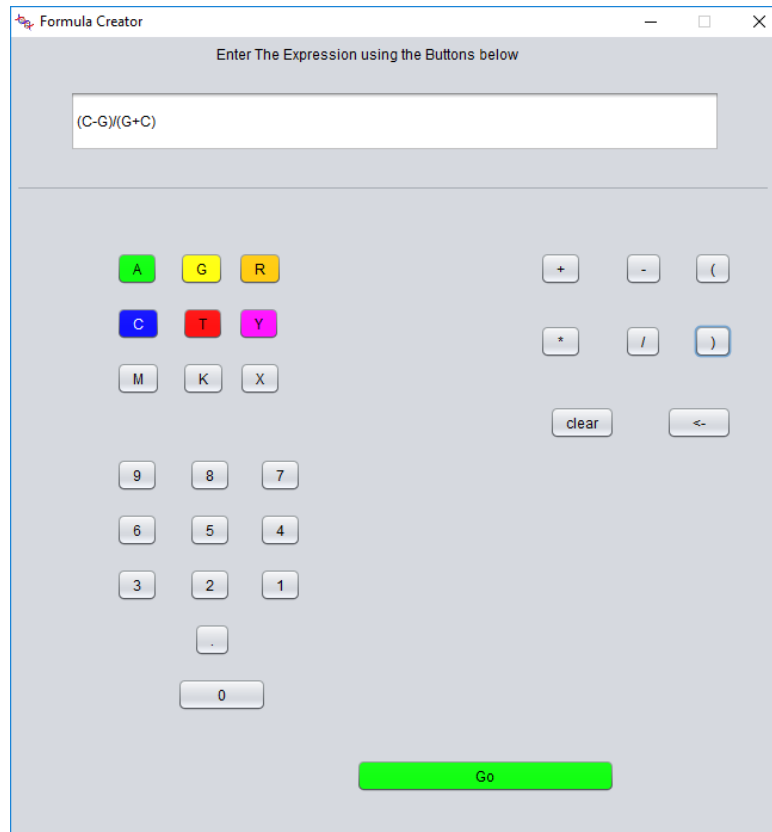


Figure 5 Window to input custom skew formula. This example shows the GC-skew entered as a mathematical formula.

Auto-Correlation Method

To access the auto-correlation based methods for prediction of origin regions, go to Menubar → OriginFinder → Correlation method and choose the required method.

Z-curve

To make a z-curve plot, go to Menubar → OriginFinder → Z-curve

Automatic Prediction

ORIS also has an automatic prediction module. This module automatically detects the sites of origin of replication by examining the various skew based and auto-correlation based methods. Motif and gene search is also performed for more reliable predictions. The output is a list of putative origin of replication sites with evidence.

To perform automatic prediction with ORIS, go to Menubar → OriginFinder → Automatic Prediction. A new frame will be displayed which will allow user to enter the parameters for the search (Figure 6).

Prediction of OriC

☒ Choose dnaA box E. Coli:TTATCCACA

☐ Choose ORB motif (Archea) Methanococcacea:TTWKATTCAyrATWTWTW

☐ Enter Other

Max. Mismatches Allowed 1

Minimum no. of dnaA matches required 2

Open .ptt file

Open Glimmer output file

Check Methods Below

☒ GC skew ☒ MK skew ☒ RY skew ☒ Correlation ☒ Entropy

Depth of Search

Go

Figure 6 Automatic prediction parameters in ORIS. User can choose which motifs to search and minimum number of matches required. The annotation could be read as .ptt file or output from the Glimmer program. The various methods used for prediction are listed as checkboxes. Depth of search parameter indicates how thorough the search should be performed. Sliding the slider to right means deeper search and results in higher number of putative origins of replication sequences.

Pattern Search

The pattern search menu has methods to perform basic motif search in the genome and generate sequence logos. To access origin finder menu, go to Menubar → Pattern Search.

Search Sequence

Any sequence can be searched with the Search sequence option. Go to Menubar → Pattern Search → Search Sequence (Figure 7).

Figure 7 Searching a motif in ORIS. User can enter a motif to search using A, G, C, T, R, Y, W, S, M, K or X characters. Various options for visualizing the results are also available e.g., generating a sequence logo, displaying the hits along the genome coordinates and saving results to text file.

DnaA box and Yeast ACS

In ORIS we have provided a list of several motifs which are found near with origin of replication sites. These are DnaA boxes and Yeast ACS sequences. To directly search one of these motifs, go to Menubar → Pattern Search → DnaA Box OR Menubar → Pattern Search → Yeast ACS

Weight Matrix Search

Users can read a non-gapped seq-alignment file to create a weight matrix or can create a weight matrix on their own. Then one can enter a sequence in which search is performed and results are stored in a .txt file specified by the user.

To perform weight matrix search, go to Menubar → Pattern Search → Weight Matrix Search (Figure 8).

Enter Wt-Matrix

Enter no. of columns :

Create

Or

Choose MSA file

Figure 8 Performing weight matrix search with ORIS. User can read the weight matrix file or create his/her own weight matrix.

Extract Sequence

To extract DNA sequence from the genome using given coordinates, go to Menubar → Pattern Search → Extract Sequence (Figure 9).

Extract Sequence

Enter start and end positions to get the sequence (positions start from 1)

Start : End :

Select an option below if required

☐ Complement

☐ Reverse

Enter file name to save as .txt

GO

Figure 9 Extracting DNA sequence from the genome from given coordinates. User can enter the coordinates and select options to reverse or complement or both. A filename should be provided to save the results.

DNA Analysis

DNA Bending Analysis

To access DNA Bending Analysis method, go to Menubar → DNA Analysis → DNA Bending Analysis (Figure 10).

Select Parameters for DNA Bending Analysis

Enter the origin site interval(default)

To

Or

☐ Enter the center of origin site

Enter Window size:

Enter Increment

☐ Save plot data as txt file

Enter File Name

Go

Figure 10 DNA bending analysis with ORIS. User can enter the coordinates of the replication site or enter the coordinate which lies in the center of the origin of replication site.

Information

The information menu has methods to examine the sequence composition using methods based on information theory. To access the information menu, go to Menubar → Information (Figure 11).

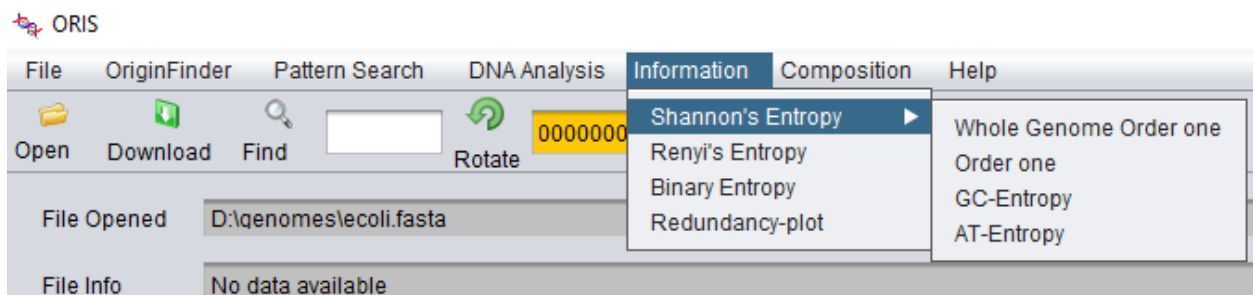


Figure 11 Selecting the "Information" menu in ORIS

Shannon's Entropy

To compute the Shannon's entropy over the genome, go to Menubar → Information → Shannon's Entropy and select appropriate method.

Renyi's Entropy

To compute the Shannon's entropy over the genome, go to Menubar → Information → Renyi's Entropy.

Binary Entropy

To compute the Shannon's entropy over the genome, go to Menubar → Information → Binary Entropy.

Redundancy Plot

To make the redundancy plot, go to Menubar → Information → Redundancy plot

Composition

The composition menu has methods to plot the sequence composition using a sliding window approach. To access the information menu, go to Menubar → Composition and select the required option (Figure 12).

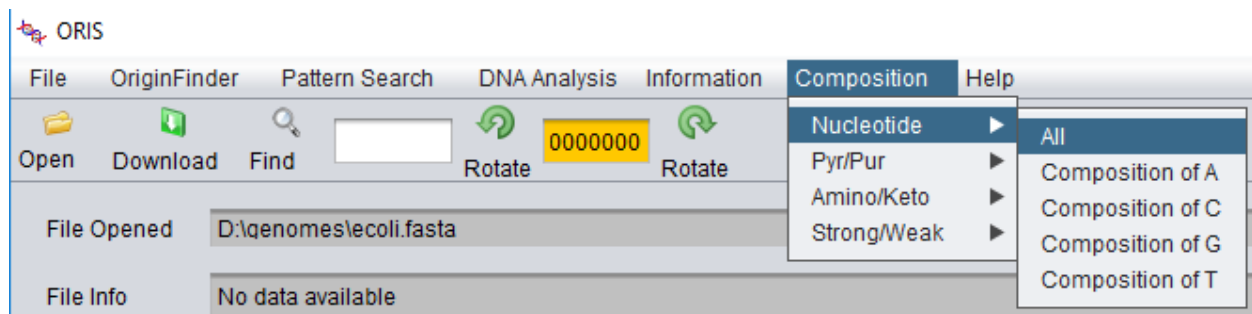


Figure 12 Selecting the "Composition" menu in ORIS

Help

The Help menu provides options to help user with the analysis. To access the information menu, go to Menubar → Help.

ORIS Settings

To adapt ORIS to one's needs, one can tweak the settings from the settings menu. Go to Menubar → Help → Settings. A settings form will be displayed (Figure 13).

Settings

Change Settings

Restart Application for Settings to take Effect

Enter Default Window size: 50000

Enter Default Increment size: 10000

Enter Default Open Directory: D:\genomes/ Choose

Enter Default Save Directory: Choose

Select App Priority(No restart reqd) : High

*High priority will utilize max CPU cores and perform Faster

Save

Figure 13 ORIS settings form. User can enter default window size and increment parameters and set default open and save directories. Users can also set the priority which indicates amount of CPU used by multithreaded computations in ORIS.

User Guide

There is a small user guide which could be accessed from within the ORIS program. To access the use guide, go to Menubar → Help → User Guide (Figure 14).

Oris User Guide

Index

1. Windowing system
2. Skew Methods
3. Correlation
4. Pattern Search
5. Weight Matrix
6. Extract Sequence
7. Bending Analysis
8. Entropy
9. Composition
10. Settings

Most methods in ORIS use a sliding window method to compute results. To do this a window of some size is chosen and an increment is defined. Then computations are done on the sequence under the window. The window is then slid with an increment value and procedure is repeated.

Sliding window is used as sequences are very large and doing computation on whole sequence will be very slow

Figure 14 ORIS user guide. To navigate to different topics, user can click on the index on the left panel and results will be displayed on the right panel.

License

ORIS is available under the GNU General Public License. See <https://github.com/urmi-21/ORIS/blob/master/LICENSE> for more information.