geneviz-java

1. Brief

    1.1 Geneviz-Java is a Java-based application for visualizing genomic data. It provides tools for loading, parsing, and displaying genomic sequences and annotations from various file formats such as FASTA and GFF.

2. Features

    2.1 - Load and parse FASTA files to visualize DNA sequences.

    2.2 - Load and parse GFF files to visualize genomic features and annotations.

    2.3 - Basic Statistics on DNA Sequences and Features.

    2.4 - FASTA Visualization Panel.
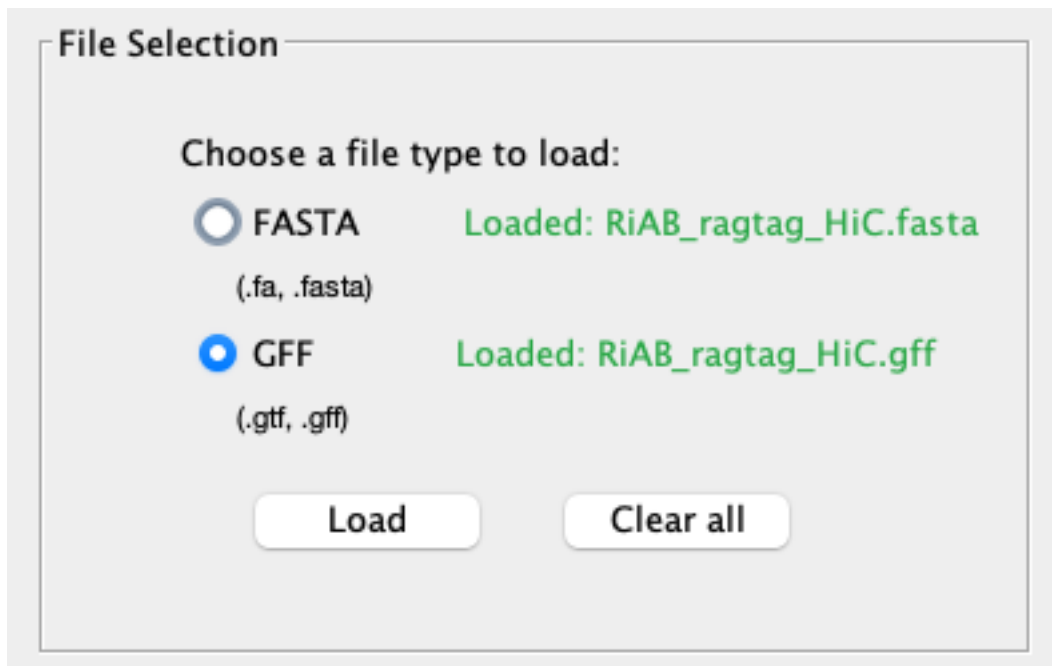
    2.5 - GFF Visualization Table.

3. Requirements

    3.1 Java 21 or higher

4. Installation

    4.1 Ensure you have Java 21 or higher installed on your system.

    4.2 Download the JAR file from /geneviz-java/target/genevis-java*.jar.

    4.3 Double-click the JAR file to run the application.

5. File Selection

    5.1 The file expects the user to load a FASTA file and a GFF3 file (*Note: currently only GFF3 file support is available*).

    5.2 The user may also load these files from the File menu bar.

6. Basic Statistics

6.1 The basic statistics tab displays two options: FASTA and Features.
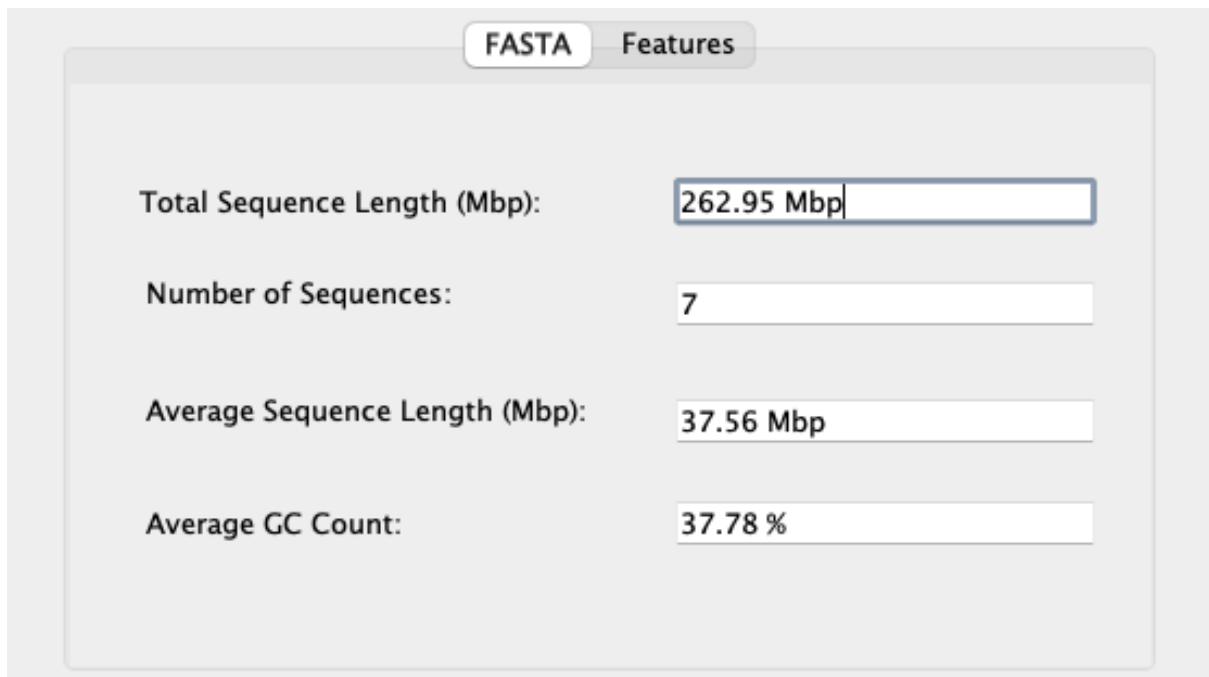
6.2 FASTA:

6.2.1 Displays the total length of the all the sequences in the FASTA file in megabases — *Note: Support for filtering on individual sequences will be made available in a future release.*

6.2.2 Displays the number of sequences in the FASTA file.

6.2.3 Displays the average sequence length in megabases of the FASTA file.

6.2.4 Displays the average GC count expressed as a percentage of the sequences in the FASTA file.

FASTA	**Features**
Total Sequence Length (Mbp):	262.95 Mbp
Number of Sequences:	7
Average Sequence Length (Mbp):	37.56 Mbp
Average GC Count:	37.78 %

6.3 **Features:**

6.3.1 Default gene selection is "Select All". This allows the user to interact with features in the file considering all genes. The user has the ability to select a specific gene — of type in a specific gene from the drop down box to explore further this gene. Click the update button to generate statistics on the user choice.

6.3.2 Displays the {average} number of genes per feature selected.

6.3.3 Displays the longest feature from the selected gene.

6.3.4 Displays the shortest feature from the selected gene.

6.3.5 Displays the average feature length from the selected gene.

6.3.6 Per gene features include [mRNA, exon, intron, CDS].

FASTA **Features**

Select Gene:
jg2077

Number of exons in jg2077:
2.00

Select Feature:
exon

Longest exon in jg2077:
590

Update

Shortest exon in jg2077:
175

Average exon Length in jg2077:
382.50

7. FASTA Visualization

    7.1 The FASTA Visualization tab allows the user to see the FASTA sequences at a selected gene location. In addition, the user can highlight features of interest with the feature dropdown box. Click the update button to calculate a new selection.

    7.2 The gene visualization shows a visual representation of the gene model. In the current release the gene model will show left-to-right orientation with a label indicating the direction of the gene. *Note: future release will automatically re-orientate the gene model based on strand direction.*

FASTA Visualization

TTCATTTGGTATCTGTTCCTTCTTTCCTGAACACGAGAGAGGCAAAGAGAGAGACGGGGACGGCGCAGTAGCTGGAACGGCGAA
GGCGGAGTGCGAGGACAGAGGCGGTGCTTAAGTGCGTCGGCGTCGCCGCCGGCAGAGAGCAAGTGCTGTGCTTAAGGAGGAAAA
CCGGCTGGTGTTGCATCGGAGACGGCGGTGCAGGTGGCAACAATGGATTGGCGAAGATAGCGGCAGCCGAGAAGCTCGACGCCG
GCCGGCCGGGCTTAAGGAGGAAGACCGGCGGCGGTGAGGTTGCTGCGAGAATGAACGGCGCGATGAATTGGACTGGGCACGTTA
GCGACGGCGACGATTGGGATCTGAAATGGGTCGCCGGGAGCAGAGGACTGAGCAGCTTGAAAACGGCGCCGGCAAGGATTGCGA
GAGGGATCGATTGAGGCAGTCCGGCCGGACTGGTTCTCGTGCAGCTTCCAAACGGCGCCGGCTTGGGCTGAAATTGGAATCCGAG
CGAGAGGTGCGGCATGCGTGGCCGTTTGTGGGAAGCGGTGGCTTCGTAATTGATGCCGTGAACAGAGAAGGGCTCGTGGTGATG
ATGCTGCAGAGGCGTGGCTTTAAGCTTCCAAACCTGCAAAATAAAACAAACGGACATAATTAGATTGATTGATGATTGTGAAACT
TGCAGGTGGTTCTCTTAAATAGCGAGAGGAAAGGGGAGCTCGGTGGAAGGGGTAGGTGATGGTGGTGTTGGTTGGTGAAGCTAGA
AGCCAAGGAGGAGGTGCCGGCAGTGGAGGAAAACCAGGGAAAAAATGGACTGTAGTCAAAGAGAAGTGGGAATGGTGGGCAGC
CGGCAT

Gene Visualization

<--- Reverse Strand

Select Gene:
jg2077

Select Feature:
exon

Update

8. GFF Visualization

8.1 The GFF Visualization table shows a tabular view of the GFF file. *Note: this table can take a while to load depending on the size of your input file. Max recommended size for this version is 66.7MB — support for faster parsing into data streams will be made available in future versions.*

GFF Visualization

SeqName	Source	Feature	Start	Stop	Score	Strand	Frame	Attribute
HiC_scaffol...	AUGUSTUS	gene	34910959	34913014	0.0	–		. ID=jg101...
HiC_scaffol...	AUGUSTUS	mRNA	34910959	34913014	0.0	–		. ID=jg101...
HiC_scaffol...	AUGUSTUS	stop_codon	34910959	34910961	0.0	–	0	ID=jg101...
HiC_scaffol...	AUGUSTUS	CDS	34910959	34911088	0.64	–	1	ID=jg101...
HiC_scaffol...	AUGUSTUS	exon	34910959	34911088	0.0	–		. ID=jg101...
HiC_scaffol...	AUGUSTUS	intron	34911089	34912547	0.0	–		. ID=jg101...
HiC_scaffol...	AUGUSTUS	CDS	34912548	34913014	0.86	–	0	ID=jg101...
HiC_scaffol...	AUGUSTUS	exon	34912548	34913014	0.0	–		. ID=jg101...
HiC_scaffol...	AUGUSTUS	start_codon	34913012	34913014	0.0	–	0	ID=jg101...
HiC_scaffol...	AUGUSTUS	gene	15890196	15891041	0.0	–		. ID=jg2077;
HiC_scaffol...	AUGUSTUS	mRNA	15890196	15891041	0.0	–		. ID=jg207...
HiC_scaffol...	AUGUSTUS	stop_codon	15890196	15890198	0.0	–	0	ID=jg207...
HiC_scaffol...	AUGUSTUS	CDS	15890196	15890785	0.82	–	2	ID=jg207...
HiC_scaffol...	AUGUSTUS	exon	15890196	15890785	0.0	–		. ID=jg207...
HiC_scaffol...	AUGUSTUS	intron	15890786	15890866	0.0	–		. ID=jg207...
HiC_scaffol...	AUGUSTUS	CDS	15890867	15891041	0.82	–	0	ID=jg207...
HiC_scaffol...	AUGUSTUS	exon	15890867	15891041	0.0	–		. ID=jg207...
HiC_scaffol...	AUGUSTUS	start_codon	15891039	15891041	0.0	–	0	ID=jg207...
HiC_scaffol...	AUGUSTUS	gene	25517536	25519088	0.0	+		. ID=jg322...
HiC_scaffol...	AUGUSTUS	mRNA	25517536	25519088	0.0	+		. ID=jg322...

9. Misc

9.1 File navigation bar allows the user to load a files outside of the file selection panel. Attempting to load multiple files is not recommended.

9.2 File navigation: clear all files will do exactly what it says on the tin. It will also clear and redirect the user to the file selection panel.

9.3 Users can only navigate to the secondary tabs once data is loaded.

9.4 *Note: The UX design leaves a little to be desired. The tabbed interface keeps the app simple and structured. The app would benefit from more flexible scaling approach. The current version does not scale if the user expands the interface.*