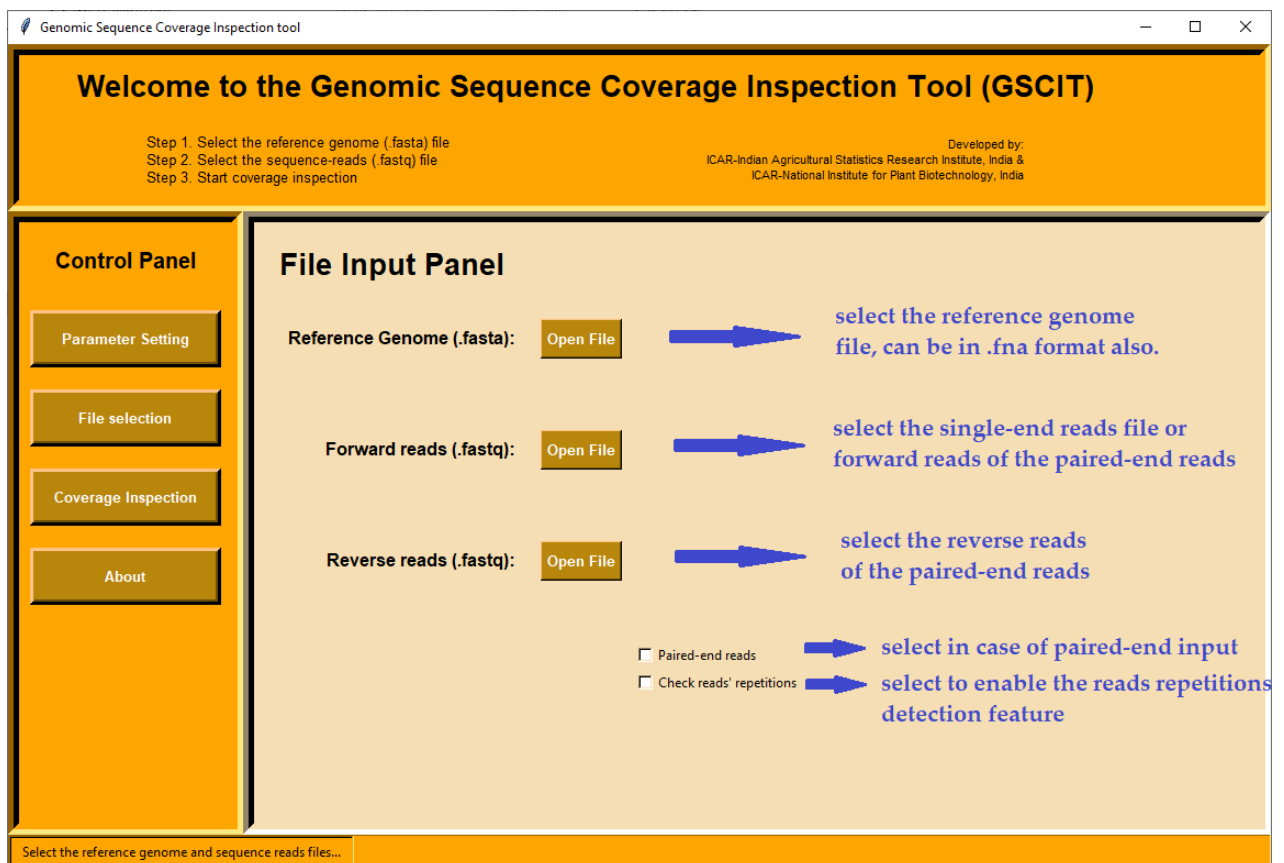


# GSCIT user manual

The GSCIT software is designed to estimate the coverage depth and breadth and for quality-profiling of the sequenced reads quickly. GSCIT is a platform-independent software implementing a novel Hash Table-based mapping algorithm in addition to various other prominent reads-profiling functionalities. The tool's coverage-indicating metrics provide novel insights regarding the input data with a user-friendly interface. In addition, the software does not need user-parameter tuning, as the artificial intelligence-based modules calculate multiple parameters on their own on run-time, which keeps the software from crashing while processing large genomes and limits the processing time. Once clicked the software may take 20-30 seconds to show up, depending upon the computer's configuration. For each execution run, the software creates a new folder in the directory the software is running. The graphs, plots, and output logs are stored in the same folder. For any query contact at [samarth.godara@gmail.com](mailto:samarth.godara@gmail.com).

Samarth Godara,  
Scientist, Division of Computer Applications,  
ICAR-Indian Agricultural Statistics Research Institute (IASRI),  
Library Avenue, Pusa, New Delhi-110012

1. GSCIT Panels
  - a. File Selection panel



b. Parameter Setting panel

Genomic Sequence Coverage Inspection tool

## Welcome to the Genomic Sequence Coverage Inspection Tool (GSCIT)

Step 1. Select the reference genome (.fasta) file  
Step 2. Select the sequence-reads (.fastq) file  
Step 3. Start coverage inspection

Developed by:  
ICAR-Indian Agricultural Statistics Research Institute, India &  
ICAR-National Institute for Plant Biotechnology, India

**Control Panel**











Parameter Setting

File selection

Coverage Inspection

About

**Parameter Panel**

	Range	Default values	
Ref Genome Sample size :	<input type="text"/>	0.001 - 1.0	to be calculated  what fraction of the reference genome is to be used?
Sequence Read Sample size :	<input type="text"/>	0.001 - 1.0	1.0  what fraction of the NGS reads is to be used?
Reads Limit :	<input type="text"/>	1KB - 1KGB	6221673  What is the maximum number of reads GSCIT will process?
Quality Threshold :	<input type="text"/>	10 - 100	30  Reads above this PHRED score are to be used for mapping
Read Length Threshold :	<input type="text"/>	20 - 1000	30  reads shorter than this are not used for mapping
Key Length :	<input type="text"/>	5 - 20	10  minimizing this will increase accuracy but it will take longer time
Mismatch Threshold :	<input type="text"/>	0 - 20	2  number of mutations acceptable per read for mapping
Read length (Upper Limit):	<input type="text"/>	100 - 1000	1000  reads should not be longer than this (used for buffering)
Starting adapter size:	<input type="text"/>	0 - 30	0  number of base pairs to remove from the beginning of reads
Ending adapter size:	<input type="text"/>	0 - 30	0  number of base pairs to remove from the ending of reads

Set the inspection parameters... (Empty cells represent default values)

c. Coverage Inspection panel

Genomic Sequence Coverage Inspection tool

## Welcome to the Genomic Sequence Coverage Inspection Tool (GSCIT)

Step 1. Select the reference genome (.fasta) file  
Step 2. Select the sequence-reads (.fastq) file  
Step 3. Start coverage inspection

Developed by:  
ICAR-Indian Agricultural Statistics Research Institute, India &  
ICAR-National Institute for Plant Biotechnology, India

**Control Panel**

Parameter Setting






File selection

Coverage Inspection

About

Start inspection

Stop inspection

button to start the inspection

button to stop the inspection

Output panel

Click on 'Start Inspection' to start the coverage inspection...