User Guide

Arima-HiC⁺ Kit

Arima-HiC Bioinformatics User Guide for Visualization of HiC Data with HiGlass in a Mac or Linux Based OS

Material Part Number: A101020
Document Part Number: A160601 v00
Release Date: September 2021

This product is intended for research use only. This product is not intended for diagnostic purposes.

This document and its contents are proprietary to Arima Genomics, Inc ("Arima Genomics"). Use of this document is intended solely for Arima Genomics customers for use with the Arima-HiC+ Kit, P/N A101020, and for no other purpose. This document and its contents shall not be used, distributed or reproduced in whole or in part and/or otherwise communicated or disclosed without the prior written consent of Arima Genomics.

This user manual must be read in advance of using the product and strictly followed by qualified and properly trained personnel to ensure proper use of the Arima-HiC+ kit. Failure to do so may result in damage to the product, injury to persons, and/or damage to other property. Arima Genomics does not assume any liability resulting from improper use of its products or others referenced herein.

U.S. Patent No. US 9,434,985 and 9,708,648 pertains to the use of this product.

TRADEMARKS
Illumina®, MiSeq®, MiniSeq®, NextSeq®, HiSeq®, and NovaSeq™ are trademarks of Illumina, Inc.

Document	Date	Description of Change
Material Part Number: A101020 **Document Part Number:** A160601 v00	September 2021	Initial Release

Table of Contents

1 Introduction

HiGlass (https://docs.higlass.io/index.html) is an extensible and customizable viewer for multi-scale genomic data sets and is an alternative to Juicebox (https://github.com/aidenlab/Juicebox). HiGlass runs on web browsers in a Docker container.

2 Installing HiGlass

To install HiGlass, you will need the software packages below:

- Docker (https://docs.docker.com/engine/install/): Docker is program that lets you run "containers" hosting software and its dependencies.
- Python 3 (https://www.python.org/downloads/): Python 3 is required to install HiGlass.
- higlass-manage (https://github.com/higlass/higlass-manage): This package is a wrapper for the Docker commands used to run a local instance. Run the commands below to install:

```
pip install higlass-python
pip install higlass-manage
```

- SAMtools (https://github.com/samtools/samtools): This is needed for processing BAM files.
- bgzip (https://github.com/samtools/tabix): This is needed for compressing the output files.
- pairix (https://github.com/4dn-dcic/pairix): This is needed for indexing and querying on a block-compressed text file containing HiC interactions. Run the command below to install:

```
git clone https://github.com/4dn-dcic/pairix
```

You will need to add SAMtools, bgzip, pairix, and higlass-manage installation paths to your $PATH variable so that you can run these programs from any directory.

3 Computational Resources

We recommend using at least 4 GB of memory to process the data for visualization. The most time-consuming step will be the bam2pairs script which may run up to 12 hours for samples containing a few hundred million of read pairs.

4 Example Usage

Before you start, download Arima Genomics Public CHiC data from using the command below:

```
wget ftp://ftp-
arimagenomics.sdsc.edu/pub/ARIMA_TEST_DATASET/BAM/test_300M.bam
```

In order to visualize the HiC interactions using HiGlass, the first step is to convert the BAM file to a format called "pairs" containing HiC interactions using the command below.

```
~/tools/pairix/util/bam2pairs/bam2pairs -c hg38.chrom.sizes
test_300M.bam test_300M
```

Next convert the *.bsorted.pairs.gz file to a *.cool file in binary format, which stores large matrices of HiC interactions.

```
~/tools/anaconda3/bin/cooler cload pairs -c1 2 -p1 3 -c2 4 -p2 5
hg38.chrom.sizes:10000 test_300M.bsorted.pairs.gz test_300M.10kb.cool
```

Finally, convert the *.cool format to *.mcool format which contains multiple levels of resolutions of your choice. You can use "-n" option to utilize multiple processors to speed up the computation. When you specify different resolutions, please always include 1 million and 2 million base pair resolutions in order for the browser to fully display the entire genome (see **Fig. 1**) and arrange all other resolutions to be differ by no more than a factor of 5. The "--balance" option will apply iterative correction (i.e. matrix balancing normalization) to the HiC data.

```
~/tools/anaconda3/bin/cooler zoomify --balance -r
20000,50000,100000,200000,500000,1000000,2000000 test_300M.10kb.cool
```

Once you have the "*.mcool" binary file, you can view your heatmap image using your web browser. In your Linux terminal, please run:

```
higlass-manage start
```

Then the following command will open your default web browser and re-direct you to a webpage looks like this: http://localhost:8989/app/?config=XXXXXX

```
higlass-manage view test_300M.10kb.mcool
```

After finish viewing your heatmap image using your web browser, please run:
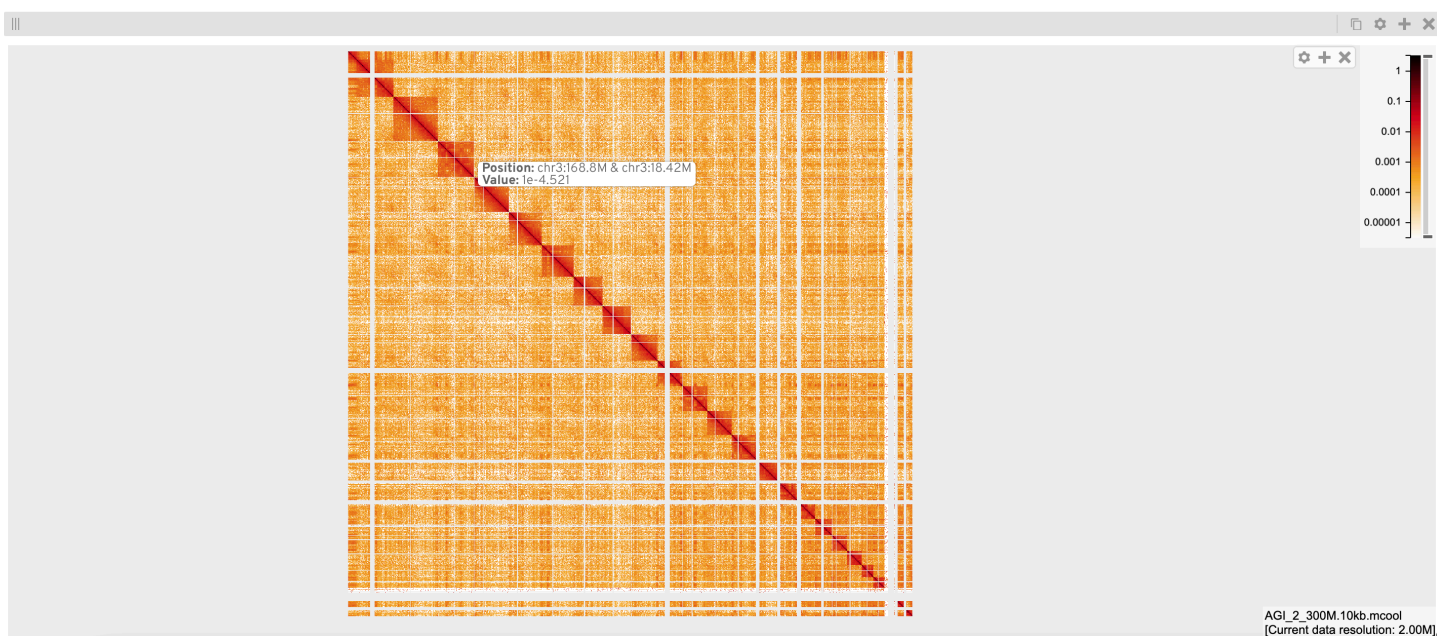
```
higlass-manage stop
```

Fig. 1. Visualization of the Arima CHiC data using HiGlass.

5 How to Cite

Kerpedjiev et al. (2018), HiGlass: Web-based visual comparison and exploration of genome interaction maps. Genome Biology, 19:125.

Warranty and Contact Info

WARRANTY DISCLAIMERS

THE EXPRESS WARRANTIES AND THE REMEDIES SET FORTH ABOVE ARE IN LIEU OF, AND ARIMA GENOMICS AND ITS LICENSORS, SUPPLIERS AND REPRESENTATIVES HEREBY DISCLAIM, ALL OTHER REMEDIES AND WARRANTIES, EXPRESS, STATUTORY, IMPLIED, OR OTHERWISE, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY, SATISFACTORY QUALITY, NONINFRINGEMENT OR FITNESS FOR A PARTICULAR PURPOSE, OR REGARDING RESULTS OBTAINED THROUGH THE USE OF ANY PRODUCT OR SERVICE (INCLUDING, WITHOUT LIMITATION, ANY CLAIM OF INACCURATE, INVALID OR INCOMPLETE RESULTS), IN EACH CASE HOWEVER ARISING, INCLUDING WITHOUT LIMITATION FROM A COURSE OF PERFORMANCE, DEALING OR USAGE OF TRADE, OR OTHERWISE. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, ARIMA AND ITS LICENSORS, SUPPLIERS AND REPRESENTATIVES SHALL NOT BE LIABLE FOR LOSS OF USE, PROFITS, REVENUE, GOODWILL, BUSINESS OR OTHER FINANCIAL LOSS OR BUSINESS INTERUPTION, OR COSTS OF SUBSTITUTE GOODS OR SERVICES, OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, EXEMPLARY OR INDIRECT DAMAGES FOR BREACH OF WARRANTY.

WARRANTY

All warranties are personal to the Purchaser and may not be transferred or assigned to a third-party, including an affiliate of the Purchaser. The warranty described below excludes any stand-alone third-party goods that may be acquired or used with the Product. Arima Genomics only warrants that the kit reagents will be made and tested in accordance with Arima Genomics manufacturing and quality control processes. Arima Genomics makes no warranty that the reagents provided in this kit will work as intended by the Purchaser or for the Purchaser's intended uses. ARIMA GENOMICS MAKES NO OTHER WARRANTY, EXPRESSED OR IMPLIED. THERE IS NO WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. The warranty provided herein and the data and descriptions of Arima Genomics products appearing in Arima Genomics product literature and website may not be altered except by express written agreement signed by an officer of Arima Genomics. Representations, oral or written, which are inconsistent with this warranty or such publications are not authorized and if given, should not be relied upon.

The foregoing warranties do not apply to the extent a non-conformance is due to (i) abuse, misuse, neglect, negligence, accident, improper storage, or use contrary to the Documentation or Specifications, (ii) use that is an Excluded Use, (iii) improper handling, (iv) unauthorized alterations, (v) natural disasters, or (vi) use with a third-party's good that is not specified in the product documentation. In the event of a breach of the foregoing warranty, customer shall promptly contact Arima Genomics customer support to report the non-conformance and shall cooperate with Arima Genomics in confirming or diagnosing the non-conformance. Additionally, Arima Genomics may request return shipment of the non-conforming product at Arima Genomics cost. Arima Genomics sole obligation shall be to replace the applicable product or part thereof, provided the customer notifies Arima Genomics within 90 days of any such breach. If after exercising reasonable efforts, Arima Genomics is unable to replace the product, then Arima Genomics shall refund to the Purchaser all monies paid for such applicable product.

CONTACT US

Technical Support: techsupport@arimagenomics.com
Order Support: ordersupport@arimagenomics.com