Supplementary Material: A step-by-step guide to I-ATAC, validating pipeline with two case studies

- 1 I-ATAC: Interactive pipeline for the management and pre-2 processing of ATAC-seq samples 3 4 Zeeshan Ahmed¹ and Duygu Ucar² 5 ¹University of Connecticut Health Center, 195 Farmington Ave, Farmington, CT, USA 6 7 ²The Jackson Laboratory for Genomic Medicine, 10 Discovery Dr., Farmington, CT, 8 USA 9 Corresponding authors: Zeeshan Ahmed¹ and Duvgu Ucar² 10 Email address: zahmed@uchc.edu and duygu.ucar@jax.org 11 12 Software Availability: I-ATAC is programmed in Java and built at both Mac-OS-X and Windows 13 14 platform. 15 • Its source code and executable are freely available at: https://github.com/UcarLab/I-ATAC 16 17 • Example dataset is available at: https://zenodo.org/record/46079#.WAe315MrK7Y 18 • Supporting software and dependencies are available at: 19 https://zenodo.org/record/162023#.WAe3dJMrK7Y 20 21 • For additional information, please refer to the project webpage: https://www.jax.org/research-and-faculty/tools/i-atac 22
- 23

Table of Contents

26	1	Motivation	3
27	2	I-ATAC	4
28	3	Design Description	6
29	3.		
30	3.	.2 Applications integration, data processing pipeline and project's directory structure	7
31	3.	.3 Comments workflow, operating systems and physical data storage in data cluster	9
32	4	GUI Description	10
33	5	Integrated Applications Details	13
34	5.		
35	5.	2 Trimmomatic	13
36	5.		
37	5.	.4 SAMtools	
38	5.		
39	5.		
40	5.		
41	5.	.8 MACS2	14
42	6	Installation and Configuration	15
43	7	Case Studies	
44	7.	1 Example Dataset	
45		7.1.1 Dataset Details	17
46		7.1.2 Input	
47		7.1.3 Output	
48	7.	.2 Case Study 2: Using GM12878 – CD4 T- Cells	
49		7.2.1 Dataset Details	
50		7.2.2 Input	
51		7.2.3 Output	25
52	8	Conclusions	29
53	9	Acknowledgments	30
54	10	Funding:	30
55	11	Conflict of Interests:	30
56	12	Additional Requirements	30
57	13	References	30
58	10		
20			

60 1 Motivation

- 61 The use of high-throughput sequencing technologies has brought an enormous increase in the amount
- 62 of heterogeneous genomic data production in the last decades. The importance of genomic dataset
- 63 processing in the genomic community is well known; as it plays important role in analysing the
- 64 dynamics and complexities of gene regulation with modelling and implementation of different
- 65 statistical methods utilizing data processing pipelines.
- 66 Traditional way of next generation sequencing (NGS) data pre-processing is complex and based on
- 67 running a series of command-line applications in Unix, Linux, MAC and DOS environments, which
- requires good knowledge of bioinformatics tools and good programming skills. There are over 200
- 69 tools available for the genome and exome sequencing data pre-processing and analysis (Pabinger *et*
- 70 *al.*, 2013) but most of them are non-interactive and command line based. Writing complex command
- 71 line scripts and pipelines, and running non-interactive mode applications might be convenient for the
- scientists with good bioinformatics background but it is very hard for the biologist with no
- 73 programming skills to conduct complex data analyses. The focus of our research is toward the
- 74 application of a novel epigenomic profiling assay for transposase-accessible chromatin with high
- 75 throughput sequencing (ATAC-seq) for integrative epigenomic analysis (Buenrostro *et al.*, 2013).
- ATAC-seq is a protocol to capture open chromatin sites (Buenrostro *et al.*, 2013; Buenrostro *et al.*, 2015a) by performing adaptor ligation and fragmentation of open chromatin regions (Tsompana and
- 77 2015a) by perform78 Buck, 2014).

79

 RAW Datasets

 Quality Check
 Adapter Check

 Quality Trimming

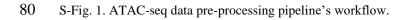
 Start
 Alignment

 Sorting
 Removing

 Duplicates

 Peak Calling

 Processed Data



- 81 ATAC-seq has been a popular chromatin profiling technology for clinical samples and has been used
- 82 for the assessment of chromatin accessibility in various cells and tissues in human and model
- 83 organisms e.g. (Moskowitz *et al.* 2017; Miskimen *et al.*, 2017; Bao *et al.* 2015; Buenrostro *et al.*,
- 84 2015b) etc. Due to its efficiency in requirement of biological sample and in library preparation time,
- many scientists are generating ATAC-seq libraries to decipher the chromatin landscape in a given
 cell type and condition of interest. To generate ATAC-seq libraries, a hyperactive molecule, Tn5 is
- cell type and condition of interest. To generate ATAC-seq libraries, a hyperactive molecule, ThS is
 used to cut the open chromatin and then short reads are sequenced typically from both ends (i.e.,
- paired end). The next step is the processing of ATAC-seq samples. A typical ATAC-seq data
- 89 processing pipeline's workflow is shown in S-Fig. 1, which starts with the quality check and adapter
- 90 trimming, then alignment, shifting, removing duplicates, sorting and peak calling to find potential
- 91 open chromatin sites, indicating active regulatory elements in each cell.
- 92 Processing and analysis of large number of ATAC-seq samples is a challenge for non-computational
- 93 scientists since usually multiple tests are required to find the optimal algorithms and parameter
- 94 settings. Interactive-ATAC (I-ATAC) (Ahmed and Ucar, 2017) is the first interactive, cross platform,
- 95 user-friendly desktop application, which supports reproducible and automatic pre-processing of
- 96 ATAC-seq (Buenrostro et al., 2013; Buenrostro et al., 2015) samples.

97 **2 I-ATAC**

98 The targeted end users of I-ATAC are mainly the biologists, who are familiar and comfortable with

99 interactive operating systems (e.g. Windows, Mac-OS-X) and applications (e.g. web based browsers

100 or client based viewers), yet have limited experience with programming, shell scripting, and with the

- 101 Unix environment. Moreover, I-ATAC could be a helpful tool for bioinformaticians, who are new to
- 102 the field of epigenomic data analysis and are not familiar with ATAC-seq data processing steps.
- 103 The GUI of the I-ATAC (S-Fig. 2 A and B) is designed for simplicity and ease by following human
- 104 computer interaction (HCI) guidelines (Ahmed *et al.*, 2014). The concept behind designing I-ATAC
- 105 GUI was to implement "One Click Operations" concept, similar to a Google search that requires
- 106 users to enter one natural language based query and click a search button. Similarly, along with the
- default or customized settings (S-Fig. 2 B), I-ATAC requires only path to the sample data files
- 108 (zipped or unzipped "FASTQ" files), project name and pressing button "Run ATAC-Seq" (S-Fig. 2
- 109 A) to perform following tasks:
- 110 Get user login credentials
- Connect to the data cluster or local computer
- Create output directory structure
- 113 Locate input data
- Copy & paste or create soft links of data to process
- 115 Load modules, compilers & interpreters
- Write command line instructions to integrate applications
- Compose shell scripts (pipeline)
- Create & queue jobs (Unix based Secure Shell Scripts) at cluster or execute instructions on local computer
- Place output files in created directory structure
- 121 Start data processing
- Disconnect to the connected data cluster

• • •		I-A	TAC ver 1.0.1 - Public	Release		
			Process Setting	JS		
Seq. Protocol 🛛 2x 💠 Project	MACProject	Input Direct	ory /ATAC_PROJECT	S/gz_fastq_files		
	II All	Auto Correct	Output Direct			
Run ATAC Seq.	V FastQC	 Auto Correct Trimmomatic 			✓ Insert Size	
	BAM Shifter	Sam tools	Samtools Index	Bed Tools	Macs2: Peak Calling	
Please write for assistance and troubleshooting	at: zeeshan.ahmed@jax.o	org OR please call at: 860-8	37-2063		I-ATAC ver. 1.0.1 is developed by Dr. Zees	han AHMED

124 S-Fig. 2 (A): Graphical User Interface of I-ATAC: Create and run data processing jobs.

0	I-ATAC ver 1.0.1 - Public Release	
	Process Settings	
Reference Genome	/data/seqdma/pipelines/QC_PIPELINE/INDEXES/HUMAN/BWA/hg19.fa	Host data-cluster
FastQC	/data/seqdma/software/FastQC/fastqc	User user
Trimmomatic	/data/seqdma/software/Trimmomatic-0.32/trimmomatic-0.32.jar	Password
Trim. Adapters	/data/seqdma/software/Trimmomatic-0.32/adapters/NexteraPE-PE.fa	
BWA	/data/seqdma/software/bwa-0.7.15/bwa	🗹 Multi Queued Jobs
Picard SortSam	/data/seqdma/software/picard/1.95/SortSam.jar	Put in Single Queue
Mark Duplicates	/data/seqdma/software/picard/1.95/MarkDuplicates.jar	Merge Replicates
nsert Size Metrics	/data/seqdma/software/picard/1.95/CollectInsertSizeMetrics.jar	Wall time 00:10:00
BAM Gap Align	/data/seqdma/software/ATAC_BAM_shifter_gappedAlign/ATAC_BAM_shifter_gappedAlign.pl	nodes 1
Samtools	samtools	ppn 1
Bedtools	bedtools	Email
MACS	macs2	name@email.com
Output Directory	/data/zahmed/ATAC_PROJECTS	🗹 Create & Queue Jobs
	Save Parameters into File Load Parameters from File	Direct Processing
	Defualt Parameters Reset Paths	Soft-Links 🗌 Cop
	Clear Parameters	✓ *.gz – zipped files

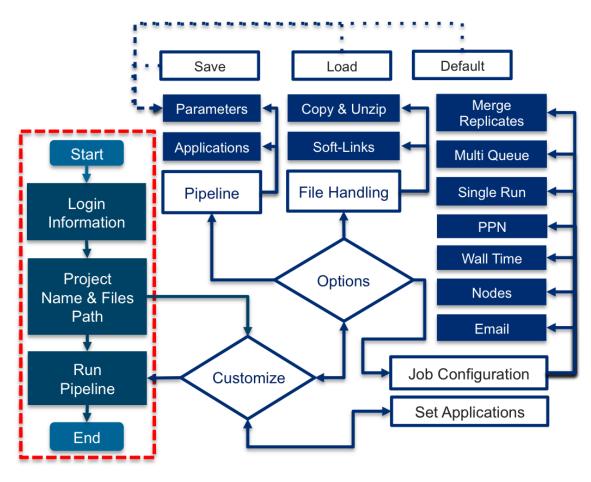
126 S-Fig. 2 (B): Graphical User Interface of I-ATAC: Set parameters and user credentials.

128 **3 Design Description**

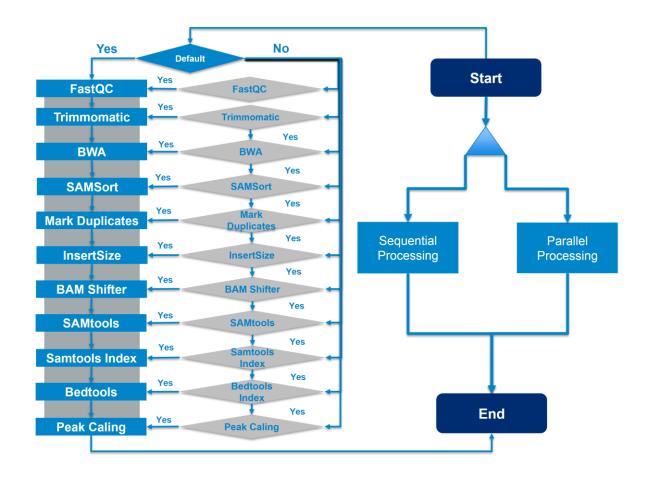
- 129 I-ATAC is a platform designed by following software engineering principles for the sustainable
- 130 bioinformatics software implementation (Ahmed *et al.*, 2014). Here, we present its operational
- 131 workflow, data structure and components' orientation.

132 3.1 Operational Workflow of I-ATAC

- 133 Following default workflow (S-Fig. 3), user can process ATAC-seq samples with the application of
- 134 complete pipeline, which involves the execution of all integrated applications (FASTQC,
- 135 Trimmomatic, BWA, Picard, ATAC_BAM_shiftrt_gappedAlign.pl, bedtools and macs2) but user is
- 136 not limited in the use of I-ATAC (S-Fig. 3). User can chose to run a single application as well as
- 137 customize applications' workflow, following pre and post-requisites e.g. in case user is only
- 138 interested in having FASTQC reports or trimming of low quality reads and adapters or user has
- already trimmed filtered FASTQ files but would like to map to reference genome only or may be
- 140 only interested in generating BED files from BAM and peak calling etc. I-ATAC supported such
- 141 customization and it can be very helpful, especially in trouble shooting situations, where due to any
- 142 reason either pipeline could not fully execute or if there is already data exists in a form which does
- 143 not require all steps of ATAC-seq pipeline. This customization can save time and computational
- 144 resources.



146 S-Fig. 3: I-ATAC: Operational workflow of I-ATAC



147

S-Fig. 4: I-ATAC: Customization of ATAC-seq data pre-processing pipeline with sequential (multiple jobs in one script)
 and parallel (multiple jobs in multiple scripts, one of each) processing.

150 User can remotely handle sample data files for processing by either keeping them in the same parent

directory and putting only pre-processed results in the main project and sub-project directories or by

first copying compressed files into the project directory, unzips them and then process them. User

153 can configure job (UNIX based Secure Shell Scripts) settings by processing one or multiple samples

154 at a time as one job or multiple jobs (one for each sample).

155 I-ATAC also enables users to customize parameters used for data pre-processing steps by letting the

user to choose between applications as well as by setting different parameters (S-Fig. 4), which

157 enables customizing this pipeline for the analyses of other data types, such as ChIP-seq data. As the

158 output, I-ATAC produces data quality reports that can be visualized within the platform. It also

159 outputs ATAC-seq reads that are filtered, trimmed and aligned as well as peak calls from these reads.

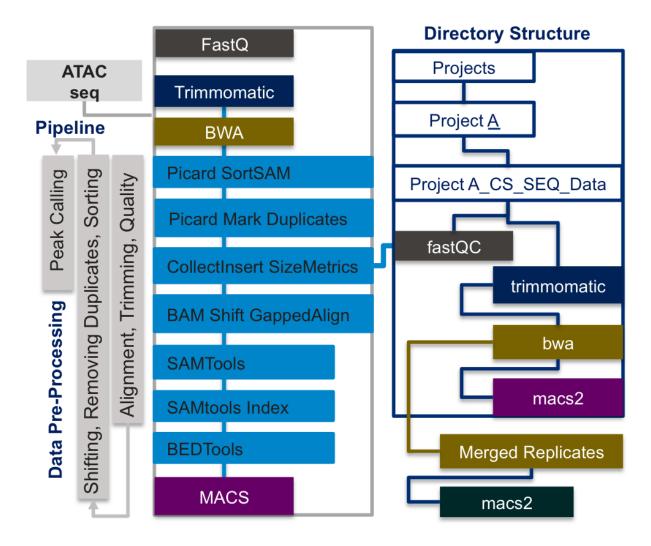
160 **3.2** Applications integration, data processing pipeline and project's directory structure

161 ATAC-seq data processing pipeline starts with the quality check, then paired end reads are trimmed,

aligned, filtered, and sorted in a "sam" file. The "sam" file is compressed and indexed to a bam file,

- 163 which is then used as input for peak calling. To manage pre-processed data, proposed directory
- structure is followed and automatically created in data cluster before data processing (S-Fig. 5).

- 166 ATACseq Projects 167 \rightarrow Project Name 168 $\rightarrow \rightarrow Sample_A_R1_Sample_A_R2$ 169 $\rightarrow \rightarrow \rightarrow fastQC$ 170 $\rightarrow \rightarrow \rightarrow$ trimmomatic 171 $\rightarrow \rightarrow \rightarrow \rightarrow bwa$ 172 $\rightarrow \rightarrow \rightarrow \rightarrow \rightarrow macs^2$ 173 \rightarrow mergedreplicated $\rightarrow \rightarrow macs^2$
- 174
- 175 All the quality reports ("zip" and "html" files) are placed in "fastQC" sub-directory. Compressed
- files contain different output files including text ("txt") and web page ("html"). Text file contains 176
- 177 information about basic statistics, file name, file type, encoding, total sequences, sequence flagged
- quality, sequence length, base number, mean, median, lower, quartile, upper, quartile, 10th 178
- 179 percentile, 90th percentile, quality, Count, per base sequence content, per sequence GC content, per
- 180 base N content, sequence length distribution, sequence duplication levels, overrepresented sequences,
- 181 adapter content and Kmer content. Whereas html file visualize quantitative results.



183 S-Fig. 5: I-ATAC: Applications and project directory structure.

184 All trimmed and filtered "FASTQ" files are placed in "trimmomatic" sub-directory, all the sorted,

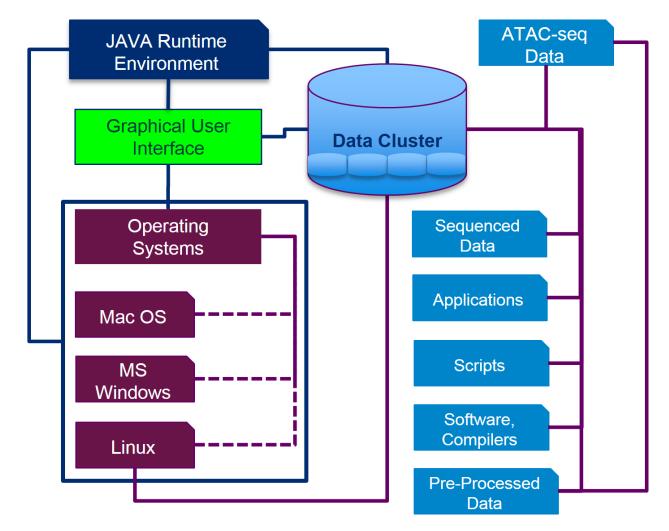
185 shifted "sam", indexed "bam" and "bed" files are placed in "bwa" sub-directory. All the observed

186 peak files are placed in the "macs2" sub-directory. The nested directory structure provides an

187 organized and modular storage for multi-level ATAC-seq data analysis pipeline. Produced results in

188 the form of sorted "sam" and "bam" files, as well as peaks can be visualized using available genome

189 data browsers (e.g. USCS, Chipster etc.) and viewers (e.g. IGV etc.).



190

191 S-Fig. 6: I-ATAC: Components workflow, operating systems and physical data storage in data cluster.

192 **3.3** Comments workflow, operating systems and physical data storage in data cluster

193 The components workflow (S-Fig. 6) of I-ATAC depends on the Java Run Time Environment

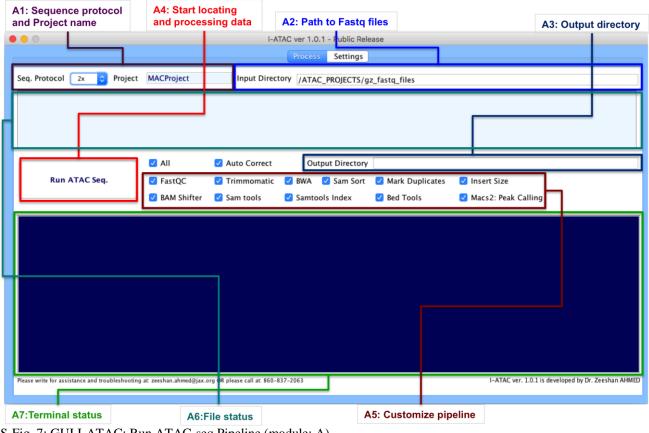
194 (<u>http://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html</u>), to be

- installed at in-use operation system, which can be Mac-OS-X, Microsoft Windows and Linux etc.
- 196 The sample, sequenced data files, applications (S-Table. 3), compilers and interpreters (S-Table. 4),
- 197 pre-processed data and scripts are need to be placed in data cluster.

198

200 4 **GUI Description**

- 201 As shown in (S-Fig. 2 A and B), the overall GUI of the I-ATAC is divided in to two modules:
- 202 Process and Settings.
- 203 The *Process* module is to generate and run pipeline. *Process* provides six major features: A1:
- 204 Sequence protocol and Project Name, A2: Output directory, A3: Path to FASTQ file, A4: Start
- 205 locating and processing data, A5: Customize pipeline, A6: File Status, and A7: Terminal status (S-
- 206 Fig. 7 and S-Table 1).



 $207 \\ 208$ S-Fig. 7: GUI I-ATAC: Run ATAC-seq Pipeline (module: A).

Number	Feature	Description
Al	Sequence protocol and Project name	 Requires user to select either "1x" or "2x" for single and paired end data. Requires user to enter a Project Name, without any spaces and special characters. Reason to avoid spaces and special characters is, that, I-ATAC will automatically create a new directory, where data will be referenced (soft links) or copied for processing and results will be placed.
A2	Path to the FASTQ files	• Requires path to the directory, where data (FASTQ files) are placed.
A3	Output directory	• Requires path to the directory, where processed data (outcome/output files) will be placed.
A4	Start locating and processing data	 Starts locating sample data files, copy from source to the main destination (project) director, unzip compressed sample data files,

		automatically generate script, queue script(s) as job(s), perform data processing and place output in proposed directory structure.		
A5	Customize Pipeline	 Default parameters include I/O redirected, sequential combination of integrated applications and parameters. Twelve different options are integrated: FASTQC Trimmomatic BWA Sam Sort Mark Duplicates Insert Size BAM Shifter SAM tools SAM tools index BED tools MACS2 All Marc Correct Option 12 is to select all options and perform data processing with default settings and option 13 is to correct the sequence I/O. 		
A6	File Status	 Provides information about located data samples in the data cluster, using provided input path. 		
A7	Terminal Status	Provides information about execution of job in data cluster.		

- 210 S-Table. 1: Features description of GUI-A: Run ATAC-seq Pipeline
- 211 The GUI-B module is mainly used to set the parameters of the applications and directory paths. As
- shown in the figure (S-Fig. 8), it provides only four features: *Applications Parameters, Directory*
- 213 Paths, Save and load Parameters, and Reset Paths (S-Fig. 8 and S-Table 2).

B1: Edit parame	eters and modules	B3: Login information
• • •	I-ATAC ver 1.0.1 - Public Release	
	Process Settings	
Reference Genome	/data/seqdma/pipelines/QC_PIPELINE/INDEXES/HUMAN/BWA/hg19.fa	Host data-cluster
FastQC	/data/seqdma/software/FastQC/fastqc	User <mark>user </mark>
Trimmomatic	/data/seqdma/software/Trimmomatic-0.32/trimmomatic-0.32.jar	Password
Trim. Adapters	/data/seqdma/software/Trimmomatic-0.32/adapters/NexteraPE-PE.fa	Multi Queued Jobs
BWA	/data/seqdma/software/bwa-0.7.15/bwa	Put in Single Queue
Picard SortSam	/data/seqdma/software/picard/1.95/SortSam.jar	Merge Replicates
	/data/seqdma/software/picard/1.95/MarkDuplicates.jar	
Insert Size Metrics	/data/seqdma/software/picard/1.95/CollectInsertSizeMetrics.Jar	Wall time 00:10:00
BAM Gap Align	/data/seqdma/software/ATAC_BAM_shifter_gappedAlign/ATAC_BAM_shifter_gappedAlign.pl	nodes 1
Samtools	samtools	ppn 1
Bedtools	bedtools	Email
MACS	macs2	name@email.com
Output Directory	/data/zahmed/ATAC_PROJECTS	✓ Create & Queue Jobs
	Save Parameters into File Load Parameters from File	Direct Processing
	Defualt Parameters Reset Paths	Soft-Links 🗌 Copy
	Clear Parameters	✓ *.gz – zipped files
B2: Load, set a	nd save parameters B4: Job settings	

214 215 S-Fig. 8: GUI I-ATAC: Set Script Parameters (module: B).

A step-by-step guide to I-ATA	C, validating pipeline	with two case studies
-------------------------------	------------------------	-----------------------

B1 Applications Parameters Paths and calling instructions to the following integrates and referenced genome and adapters:	rm of a text file, of a text file.
Image: Second	of a text file.
2. FASTQC 3. Trimmomatic 4. Adapters 5. BWA 6. SamSort 7. Mark Duplicates 8. Insert Size Metrics 9. BAM Gap Align 10. SAM tools 11. BED tools 12. Samtools Index 13. MACS2 14. Output Directory B2 Load, set and save parameters Five options (buttons) are provided: 1. Save Parameters: To save the set parameters in the fort to reuse and share settings. 2. Load Parameters: To load saved settings in the form to a scient and parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login • User requires entering name of the host (attached data c of the personal computer), user login name and passwow ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings	of a text file.
3. Trimmomatic 4. Adapters 5. BWA 6. SamSort 7. Mark Duplicates 8. Insert Size Metrics 9. BAM Gap Align 10. SAM tools 11. BED tools 12. Samtools Index 13. MACS2 14. Output Directory B2 Load, set and save parameters Five options (buttons) are provided: 1. Save Parameters: To save the set parameters in the for to reuse and share settings. 2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login B4 Job Settings Default job (set of instructions, written in the form of a script and cores sample ("FASTQ") and applications to perform data processing	of a text file.
4. Adapters5. BWA6. SamSort7. Mark Duplicates8. Insert Size Metrics9. BAM Gap Align10. SAM tools11. BED tools12. Samtools Index13. MACS214. Output DirectoryB2Load, set and save parametersFive options (buttons) are provided:1. Save Parameters: To save the set parameters in the for to reuse and share settings.2. Load parametersB3User LoginB4Job SettingsDefault job (set of instructions, written in the form of a script andB4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
5. BWA6. SamSort7. Mark Duplicates8. Insert Size Metrics9. BAM Gap Align10. SAM tools11. BED tools12. Samtools Index13. MACS214. Output DirectoryB2Load, set and save parametersFive options (buttons) are provided:1. Save Parameters: To save the set parameters in the for to reuse and share settings.2. Load Parameters: To clear all parameter fields.4. Default Parameters: To load default parameters.5. Reset Paths: To apply modifications.B3User Login• User requires entering name of the host (attached data c of the personal computer), user login name and passwor 	of a text file.
6. SamSort7. Mark Duplicates8. Insert Size Metrics9. BAM Gap Align10. SAM tools11. BED tools12. Samtools Index13. MACS214. Output DirectoryB2B2Load, set and save parametersFive options (buttons) are provided:1. Save Parameters: To save the set parameters in the for to reuse and share settings.2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To load default parameter fields.4. Default Parameters: To load default parameters.5. Reset Paths: To apply modifications.B3B4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
7. Mark Duplicates8. Insert Size Metrics9. BAM Gap Align10. SAM tools11. BED tools12. Samtools Index13. MACS214. Output DirectoryB2B2Load, set and save parametersFive options (buttons) are provided:1. Save Parameters: To save the set parameters in the for to reuse and share settings.2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To load default parameters.B3B3B4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
9. BAM Gap Align10. SAM tools11. BED tools12. Samtools Index13. MACS214. Output DirectoryB2Load, set and save parametersFive options (buttons) are provided:1. Save Parameters: To save the set parameters in the for to reuse and share settings.2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To clear all parameter fields.4. Default Parameters: To load default parameters.5. Reset Paths: To apply modifications.B3User Login•User requires entering name of the host (attached data co of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processingB4Job Settings	of a text file.
10. SAM tools11. BED tools12. Samtools Index13. MACS214. Output Directory15. B216. Load, set and save parameters17. Save Parameters: To save the set parameters in the for to reuse and share settings.18. Clear Parameters: To load saved settings in the form of 3. Clear Parameters: To load saved settings in the form of 3. Clear Parameters: To load default parameters.183193193193194194194195 <td>of a text file.</td>	of a text file.
11. BED tools12. Samtools Index13. MACS214. Output DirectoryB2Load, set and save parametersFive options (buttons) are provided:1. Save Parameters: To save the set parameters in the for to reuse and share settings.2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To load saved settings in the form of 3. Clear Parameters: To load default parameters.B3User Login••B4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
12. Samtools Index 13. MACS2 14. Output DirectoryB2Load, set and save parametersFive options (buttons) are provided: 1. Save Parameters: To save the set parameters in the for to reuse and share settings.B2Load, set and save parametersFive options (buttons) are provided: 1. Save Parameters: To save the set parameters in the for to reuse and share settings.B2Load, set and save parametersFive options (buttons) are provided: 1. Save Parameters: To load saved settings in the form of 3. Clear Parameters: To clear all parameter fields. 4. Default Parameters: To load default parameters. 5. Reset Paths: To apply modifications.B3User Login• User requires entering name of the host (attached data c of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processingB4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
13. MACS2 14. Output Directory B2 Load, set and save parameters Five options (buttons) are provided: 1. Save Parameters: To save the set parameters in the form to reuse and share settings. 2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To load default parameter fields. 4. Default Parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login • User requires entering name of the host (attached data cof the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings	of a text file.
B2 Load, set and save parameters Five options (buttons) are provided: B2 Load, set and save parameters Five options (buttons) are provided: B2 Load, set and save parameters Five options (buttons) are provided: B2 Load, set and save parameters Five options (buttons) are provided: B2 Load, set and save parameters To save the set parameters in the form to reuse and share settings. Clear Parameters: To load saved settings in the form of 3. Clear Parameters: To clear all parameter fields. 4. Default Parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login • B4 Job Settings Default job (set of instructions, written in the form of a script and	of a text file.
B2Load, set and save parametersFive options (buttons) are provided:1.Save Parameters: To save the set parameters in the for to reuse and share settings.2.Load Parameters: To load saved settings in the form of 3.3.Clear Parameters: To load default parameter fields.4.Default Parameters: To load default parameters.5.Reset Paths: To apply modifications.B3User Login4.User requires entering name of the host (attached data of of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processingB4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
parameters1.Save Parameters: To save the set parameters in the for to reuse and share settings.2.Load Parameters: To load saved settings in the form of 3.3.Clear Parameters: To clear all parameter fields.4.Default Parameters: To load default parameters.5.Reset Paths: To apply modifications.B3User Login•User requires entering name of the host (attached data of of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processingB4Job SettingsDefault job (set of instructions, written in the form of a script and	of a text file.
2. Load Parameters: To load saved settings in the form of 3. Clear Parameters: To clear all parameter fields. 4. Default Parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login • User requires entering name of the host (attached data of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings	luster or name
3. Clear Parameters: To clear all parameter fields. 4. Default Parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login • User requires entering name of the host (attached data confit the personal computer), user login name and passwort ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings	luster or name
4. Default Parameters: To load default parameters. 5. Reset Paths: To apply modifications. B3 User Login • User requires entering name of the host (attached data c of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings Default job (set of instructions, written in the form of a script and	
5. Reset Paths: To apply modifications. B3 User Login • User requires entering name of the host (attached data condition of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings Default job (set of instructions, written in the form of a script and	
B3 User Login • User requires entering name of the host (attached data construction of the personal computer), user login name and passwork ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings Default job (set of instructions, written in the form of a script and script)	
B4 Job Settings of the personal computer), user login name and password ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing	
ATAC successfully login into to host and access sample ("FASTQ") and applications to perform data processing B4 Job Settings Default job (set of instructions, written in the form of a script and	
B4 Job Settings Default job (set of instructions, written in the form of a script and	
	5.
program (executable software) to perform certain set of operation	
parameters. Furthermore I-ATAC provides eight different options	to customize
script generation and job submission: 1. Multi Queued Job : Processes multiple samples at a tin	he hv
generating and submitting parallel-multiple data process	
for each).	8 Jaco (a
2. Put in Single Queue: Processes one or multiple sample	
generating and submitting one data processing job (one	
3. Merge Replicates: Applicable only in case of processir	
samples at a time by submitting one data processing job	
enables selection of all generated "bam" files from all the processed samples directories (bwa) and performs peak	
4. Wall Time: Sets time to be allocated for the processing	
job. In case of multiple-parallel jobs, it will set provided	
jobs.	
5. Nodes: Sets the number of nodes (connection points) re	quested for job.
Default set node is 1.	• 1
6. Processor per node (ppn): Sets the number of cores (v processors) per node per. Default set ppn is 1.	intual
 7. Email: Sets to get notification (cancelled, completed) at 	bout the status
of submitted job.	e status
8. Create & Queue Jobs: In case host is data cluster, then	I-ATAC will
prepare and submit jobs.	
9. Direct Processing: In case host is personal computer, the	nen I-ATAC
will prepare and submit instructions.	C will and the
10. Creates soft links: Having checked this option, I-ATA soft links of FASTQ files in to output directory.	will create
11. Copy : Having checked this option, I-ATAC will create	CODV FASTO
files in to output directory.	
12. *.gz zijped files: Having checked this option, I-ATAC	will expect
input FASTQ files are zipped otherwise not.	-

- 217 The sole objective of developing I-ATAC is to help with the provision of interactive ATAC-seq data
- 218 processing pipeline that is why; we have not developed features for file handling between data cluster
- and operating systems. There are already some interactive tools available for such purposes e.g. File
- Zilla (<u>https://filezilla-project.org</u>), WinSCP (<u>http://winscp.net/eng/download.php</u>), Cyberduck
 (<u>https://cyberduck.io/?l=en</u>) etc.
- 222 **5** Integrated Applications Details
- ATAC-seq data processing pipeline consists of different third party applications (S-Table. 4); I/O
- 224 (input/output) redirected (one's output is treated as another's input, in terms of both data analysis and
- 225 processing) and integrated method (S-Fig. 6). Additionally, it requires all needed compilers and
- interpreters to be downloaded and installed as well (S-Table. 4).

227 **5.1 FASTQC:**

- 228 It is a command line based, non-interactive tool for the high throughput sequence data. It is
- 229 programmed in Java and requires <u>Java Runtime Environment</u> and <u>Picard</u> BAM/SAM libraries to be
- 230 installed in the data cluster. Its output is based on Basic Statistics, Per base sequence quality, Per tile
- 231 sequence quality, Per sequence quality scores, Per base sequence content, Per sequence GC content,
- Per base N content, Sequence Length Distribution, Sequence Duplication Levels, Overrepresented
- sequences, Adapter Content and Kmer Content. FASTQC used version details, including input,
- output and download details are given in S-Table. 3, Row No.: 1.

235 **5.2 Trimmomatic**

- 236 It is a command line based, non-interactive tool for the trimming of reads (Bolger *et al.*, 2014) using
- paired-end and single ended data produced by the Illumina next generation sequencing technology
- 238 (<u>http://www.illumina.com/</u>). It takes compressed or uncompressed FASTQ (phred-33 and phred-64
- 239 quality scores) file as input andmainly performs adapter filtering, sliding window trimming, base
- 240 cutting (start and end of reads, as well, at specific number) and removes below quality reads.
- Trimmomatic's used version details, including input, output and download details are given in S-
- 242 Table. 3, Row No.: 2.

243 **5.3 BWA**

- 244 Burrows-Wheeler Alignment tool (BWA) is a software application for aligning short nucleotide
- sequences to a reference genome (Li and Durbin, 2009). It implements BWA-backtrack for reading
- 246 sequence up to 100bp, and BWA-SW and BWA-MEM algorithms are for reading longer sequences
- between 70bp to 1Mbp. BWA's used version details, including input, output and download details
- 248 are given in S-Table. 3, Row No.: 3.

249 **5.4 SAMtools**

- 250 Sequence Alignment/Map (SAM) tools is a software package with various utilities, mainly used for
- sequence data formatting (Li, 2011; Li, *et al.*, 2009). It helps in performing complex operations at
- sequence data files, including variant calling, alignment, sorting, indexing, viewing, data extraction
- and format conversion. SAMtools applied package's version details, including input, output and
- download details are given in S-Table. 3, Row No.: 4.

256 **5.5 Picard**

- 257 It is Java based non-interactive tool, which requires Java Runtime Environment to execute. It is
- 258 mainly used for the sequence data manipulation in sam and bam files. Both sam and bam files
- contain same data structure and format, sam is human readable, whereas, bam is machine-readable
- 260 format (binary). It's used version's details, including input, output and download details are given in
- S-Table. 3. It performs sorting in order and can read information about library, platform, sample,
- sequence, predicted insert size etc. Picard's used version details, including input, output and
- 263 download details are given in S-Table. 3, Row No.: 5.

264 **5.6 BEDtools**

- 265 Browser Extensible Data (BED) tools (Quinlan and Hall, 2010) is a software application for
- 266 converting "bam" to "bed" files and compare large sets of genomic features. Moreover, it can be used
- 267 for converting BEDPE intervals to BAM and BAM to FASTQ, finding closest and potentially non-
- 268 overlapping interval, creating HTML pages to link UCSC locations, finding pairs that overlap other
- 269 pairs and intervals in various ways, randomly redistributed and adjust size of intervals and tag bam
- alignment etc. BEDtools used version details, including input, output and download details are given
- 271 in S-Table. 3, Row No.: 6.

272 5.7 ATAC_BAM_shifter_gappedAlign.pl

ATAC_BAM_shifter_gappedAlign.pl is an open source Perl script, which can be used to perform

- read shifting based on the read quality. It takes aligned "bam" file as an input and offsets by 4bp for
- the positive strand (sequence containing instructions for building a protein) and –5bp for the negative

strand (merely contains the complementary sequence and according to the base-pairing rules it is not

277 normally transcribed into RNA nor translated into protein). Users can use any other tools for shifting

- 278 the reads. ATAC_BAM_shifter_gappedAlign version details, including input, output and download
- details are given in S-Table. 3, Row No.: 7.

280 **5.8 MACS2**

281 Model-based Analysis of ChIP-Seq (MACS) (Zhang, *et al.*, 2008) is a tool for analyzing short reads

for the spatial resolution of the predicted sites, capturing local biases in the genome and generation of

283 peaks with detailed information about length, genome coordinates, summit, p-value, q-values, false-

discovery rate (FDR) and fold enrichment. MACS2's used version details, including input, output

and download details are given in S-Table. 3, Row No.: 8.

No.	Applications	Versions	Download Web links	Input File Formats	Outputs File Formats
1	FASTQC	0.11.2	http://www.bioinformatics.b abraham.ac.uk/projects/fast qc/	FASTQ	html, zip, txt.
2	Trimmomatic (Bolger <i>et al.</i> , 2014)	0.32	http://www.usadellab.org/c ms/?page=trimmomatic	FASTQ	FASTQ, trimU.fastq,
3	Burrows-Wheeler Alignment tool (BWA) (Li and Durbin, 2009)	0.7.10	http://bio- bwa.sourceforge.net	fastq_filtered	sam
4	Sequence Alignment/Map (SAM) tools (Li, 2011; Li, <i>et al.</i> , 2009)	0.1.19	http://samtools.sourceforge. net	sam	Sam. txt, pdf
5	Picard	1.95	http://broadinstitute.github.i	sam, bam	sam, bam

			o/picard/		
6	Browser Extensible Data (BED) tools (Quinlan and Hall, 2010)	2.22.0	http://bedtools.readthedocs. org/en/latest/content/overvie w.html	bam	bed
7	ATAC_BAM_shifter_gappe dAlign	1	https://github.com/acdaughe rty/scripts/blob/master/Most Used/ATAC_BAM_add1.pl	bam	bam
8	Model-based Analysis of ChIP-Seq (MACS) (Zhang, <i>et al.</i> , 2008)	2.1.0.20151 222	http://liulab.dfci.harvard.ed u/MACS/	bed	bed, bdg, broadPeak, gappedPeak

286 S-Table. 3: Integrated applications in I-ATAC data processing pipeline

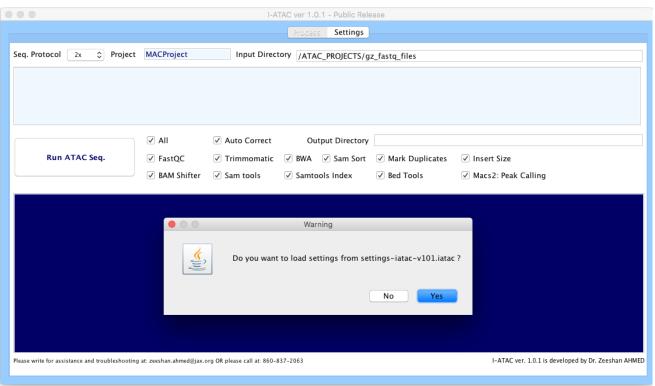
287

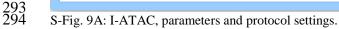
No.	Compiler / Interpreter	Versions	Download Web links
1	JAVA	8	https://www.java.com/en/download/
2	R	3.2.3	https://www.r-project.org
3	Perl	5.10.1	https://www.perl.org
4	Python	2.7.3	https://www.python.org

288 S-Table. 4: Needed compilers and interpreters

289 **6** Installation and Configuration

- 290 The software executable (JAR file) is open source and freely available and to execute I-ATAC, major
- 291 requirement is the installation of Java Runtime Environment
- 292 (http://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html).





	I-ATAC ver 1.0.1 - Public Release	
	Process Settings	
Reference Genome	/INDEXES/HUMAN/BWA/hg19.fa	Host data-cluster
FastQC	/software/FastQC/fastqc	User user
Trimmomatic	/software/Trimmomatic-0.32/trimmomatic-0.32.jar	Password
Trim. Adapters	/software/Trimmomatic-0.32/adapters/NexteraPE-PE.fa	
BWA	/software/bwa-0.7.15/bwa	Multi Queued Jobs
Picard SortSam	/software/picard/1.95/SortSam.jar	 Put in Single Queue Merge Replicates
Mark Duplicates	/software/picard/1.95/MarkDuplicates.jar	Merge Replicates
Insert Size Metrics	/software/picard/1.95/CollectInsertSizeMetrics.jar	Wall time 00:10:00
BAM Gap Align	/software/ATAC_BAM_shifter_gappedAlign/ATAC_BAM_shifter_gappedAlign.pl	nodes 1
Samtools	samtools	ppn 1
Bedtools	bedtools	Email
MACS	macs2	name@email.com
Output Directory	/ATAC_PROJECTS	✓ Create & Queue Jobs
	Save Parameters into File Load Parameters from File	Direct Processing
	Defualt Parameters Reset Paths	🗸 Soft-Links 🗌 Copy
	Clear Parameters	✓ *.gz – zipped files

295

297

296 S-Fig. 9B: I-ATAC, default setting.

seqdma/pipelines/QC_PIPELINE/INDEXES/HUMAN/BWA/hg19.fa eqdma/software/FastQC/fastqc eqdma/software/Trimmomatic-0.32/trimmomatic-0.32.jar seqdma/software/Trimmomatic-0.32/adapters/NexteraPE-PE.fa eqdma/software/bwa-0.7.15/bwa eqdma/software/picard/1.95/SortSam.jar eqdma/software/picard/1.95/MarkDuplicates.jar eqdma/software/picard/1.95/CollectInsertSizeMetrics.jar eqdma/software/ATAC_BAM_shifter_qappedAlign/ATAC_BAM_shifter_gappedAlign.pl	Host helix User zahmed Password ••••••••••••••••••••••••••••••••••••
eqdma/software/Trimmomatic-0.32/trimmomatic-0.32.jar seqdma/software/Trimmomatic-0.32/adapters/NexteraPE-PE.fa eqdma/software/bwa-0.7.15/bwa eqdma/software/picard/1.95/SortSam.jar eqdma/software/picard/1.95/MarkDuplicates.jar eqdma/software/picard/1.95/CollectInsertSizeMetrics.jar	User zahmed Password Multi Queued Jobs V Put in Single Queue V Merge Replicates
seqdma/software/Trimmomatic-0.32/adapters/NexteraPE-PE.fa eqdma/software/bwa-0.7.15/bwa eqdma/software/picard/1.95/SortSam.jar eqdma/software/picard/1.95/MarkDuplicates.jar eqdma/software/picard/1.95/CollectInsertSizeMetrics.jar	Password Multi Queued Jobs
eqdma/software/bwa-0.7.15/bwa eqdma/software/picard/1.95/SortSam.jar eqdma/software/picard/1.95/MarkDuplicates.jar eqdma/software/picard/1.95/CollectInsertSizeMetrics.jar	 Put in Single Queue Merge Replicates
eqdma/software/picard/1.95/SortSam.jar eqdma/software/picard/1.95/MarkDuplicates.jar eqdma/software/picard/1.95/CollectInsertSizeMetrics.jar	Merge Replicates
eqdma/software/picard/1.95/MarkDuplicates.jar eqdma/software/picard/1.95/CollectInsertSizeMetrics.jar	
	Wall time 10:20:30
andma/software/ATAC RAM shifter nannedAlinn/ATAC RAM shifter nannedAlinn nl	
equina/solware/ATAC_bain_sinter_gappedAligh/ATAC_bain_sinter_gappedAligh.pr	nodes 1
ls	ppn 1
ls	Email
	zeeshan.ahmed@jax.org
zahmed/ATAC_PROJECTS	🗹 Create & Queue Jobs
Parameters into File Load Parameters from File	Direct Processing
alt Parameters Reset Paths	🗹 Soft–Links 🗌 Copy
ar Parameters	✓ *.gz – zipped files
	ahmed/ATAC_PROJECTS Parameters into File Load Parameters from File alt Parameters Reset Paths

298 S-Fig. 9C: I-ATAC, user setting, loaded from "settings-iatac-v101.iatac".

- 299 After executing I-ATAC and before starting data processing, it is important to set valid applications
- 300 paths and calling protocols (section: Graphical User Interface of I-ATAC). Our default parameters
- 301 (S-Fig. 9A, 9B and 9C) are set according to our data cluster and installed versions of application (S-
- 302 Table. 3), and Compilers/Interpreters (S-Table. 4).
- 303 Using default configuration settings; I-ATAC will consider logged-in user with a default directory of
- 304 same name as of user in the data cluster (e.g. Zeehan \rightarrow "d:/data/Zeeshan/ATAC_PROJECTS/").
- 305 However, user can alter, reset and save default project directory settings.

306 7 Case Studies

- 307 In order to validate the performance of I-ATAC and to guide the users, we present two case studies.
- 308 First involves using the example data; where we have created small size example dataset (provided in 309 supplementary material and can be downloaded from the following web link:
- 310 <u>https://zenodo.org/record/46079#.VsJMg7S5LHM</u>) with artificial names (to explain the process,
- 311 execution steps in simpler way.). The reason for giving example study is to let the user, use the
- 312 application and observe results in possible shortest time. Moreover, it will also help in figuring out
- and resolving trouble shooting conditions (e.g. could be due to inappropriate installation of
- 314 downloaded application and compilers/interpreter or any other exceptional reason etc.). Second study
- is using publically available data (GM12878, CD4); where we have processed publically available
- 316 data, which a trained user can download and process using I-ATAC. In both case studies, I-ATAC is
- 317 run at the Mac-OS-X-Yosemite 10.10.5 platform.

318 **7.1 Example Dataset**

319 7.1.1 Dataset Details

- 320 Raw dataset and produced results mentioned in this example case study, which can be downloaded
- 321 from the provided project web link. Sequenced, paired sample data ("FASTQ" or "FASTQ.gz") files
- 322 are need to be collected and placed in the attached data cluster.

323 7.1.2 Input

- 324 The input to I-ATAC is the path to ATAC-seq sample data, which in our case is:
- 325 *"/data/zahmed/ATAC PROJECTS/gz fastq files"*
- As shown in S-Fig. 10, there are two samples available (paired data, four "FASTQ" zipped files) in the above-mentioned directory i.e. "gz fastq files", which are:
- 328Firt_SampleData_R1.fastq.gz329Firt_SampleData_R2.fastq.gz330Second_SampleData_R1.fastq.gz331Second_SampleData_R2.fastq.gz

```
    chined@helix:/data/zahmed/ATAC_PROJECTS/gz_fastq_files — ss...
    [zahmed@helix gz_fastq_files]$ pwd
/data/zahmed/ATAC_PROJECTS/gz_fastq_files
[zahmed@helix gz_fastq_files]$ ls -l
total 1558
-rw-r--r-- 1 zahmed jaxuser 126452 Jan 13 14:17 Firt_SampleData_R1.fastq.gz
-rw-r--r-- 1 zahmed jaxuser 131611 Jan 13 14:17 Firt_SampleData_R2.fastq.gz
-rw-r--r-- 1 zahmed jaxuser 126454 Jan 13 14:17 Second_SampleData_R1.fastq.gz
-rw-r--r-- 1 zahmed jaxuser 131613 Jan 13 14:17 Second_SampleData_R2.fastq.gz
    [zahmed@helix gz_fastq_files]$
```

332

- 333 S-Fig. 10: Screen shot (Linux Terminal, using Mac-OS-X) of compressed sample data files
- After setting parameters and input path to the I-ATAQ-seq, pressed button "Run ATAC-seq", an
- information message will appear (S-Fig. 11) to verify the input sample data source location, output
- 336 directory location and set job parameters.

		I-A	TAC ver 1.0.1	- Public Rele	ase		
			Process	Settings			
Seq. Protocol 🛛 2x 🗘 Project	Example_with_gz	_files Input Direct	ory /data/za	ahmed/ATA0	_PROJECTS/gz_fastq_	files	
Run ATAC Seq.	 ✓ AII ✓ FastQC ✓ BAM Shifter 	 ✓ Auto Correct ✓ Trimmomatic ✓ Sam tools 		Sam Sort	/data/zahmed/ATAC ✓ Mark Duplicates ✓ Bed Tools	_PROJECTS/Example_with_gz ✓ Insert Size ✓ Macs2: Peak Calling	_files
						Hi zahmed, its JAX-ATACSec-	GUI
				(it)	-> Project: Ex -> walltime: 1 -> nodes: 1 -> ppn: 1 -> Input Files	should be at: /data/zahmed/	/ATAC_PROJECTS/gz_fastq_files
Please write for assistance and troubleshooting	g at: zeeshan.ahmed@jax.o	rg OR please call at: 860-8	137-2063			II be at: /data/zahmed/ATAC_	PROJECTS/Example_with_gz_file

338 S-Fig. 11: I-ATAC, input sample data and set parameters' verification

		-A	TAC ver 1.0.1 - Public	Release		
			Process Setting	S		
Seq. Protocol 🛛 2x 🗘 Projec	ct Example_with_g	z_files Input Direc	tory /data/zahmed/A	TAC_PROJECTS/gz_fastq_	_files	
Total FastQ Files: 4						
Samples>File R1: Firt_Sample	eData_R1.fastqFile R	2: Firt_SampleData_	R2.fastq			
Samples>File R1: Second_Sa	mpleData_R1.fastqFi	le R2: Second_Samp	bleData_R2.fastq			
	✓ All	✓ Auto Correct	Output Direct	ory /data/zahmed/ATAC	_PROJECTS/E>	xample_with_gz_files
Run ATAC Seq.	✓ FastQC	✓ Trimmomatic	🗹 BWA 🗹 Sam Se	rt 🗹 Mark Duplicates	✓ Insert Si	ize
zahmed@helix ATAC_PROJECTS zahmed@helix Example_with_g]\$ cd Example_with_		✓ Samtools Index	✓ Bed Tools	✓ Macs2:	Peak Calling
[zahmed@helix ATAC_PROJECTS [zahmed@helix ATAC_PROJECTS [zahmed@helix Example_with_g total 0 [zahmed@helix Example_with_g logout Disconnected to helix drwxr-sr-x 22 zahmed jaxuse drwxr-sr-x 22 zahmed jaxuse Irwxrwxrwx 1 zahmed jaxuse	c) mkdir Example_w [] cd Example_with_ [z_files] S s - [z_files] S n - s / data, [z_files] S exit er 184 Dec 5 19: er 65 Dec 5 19: er 69 Dec 5 19:3 er 69 Dec 5 19:3 er 69 Dec 5 19:3	ith_gz_files gz_files /zahmed/ATAC_PR(34 .	OJECTS/gz_fastq_ files ,R2.fastq.gz ata_R2.fastq.gz ,R1.fastq.gz			

339

- 340 S-Fig. 12: I-ATAC, identified sample data
- 341 At successful verification, file status window (S-Fig. 12) provides the information about located
- 342 sample data files, which were copied, pasted and unzipped in the project directory
- 343 (Example_with_gz_files). At second successful verification, the ATAC-seq data processing pipeline
- 344 was automatically scripted (S-Fig. 13) and created job was queued to the data cluster (S-Fig. 14).

• •	
#1/bin/bash	
#PBS -N Example_with_gz_files #PBS -l nodes=1:ppn=1,walltime=10:20:30 #PBS -M zeeshan.ahmed@jax.org	
cd /data/shed/TL_PMOLET/Lample dif.g_list, is dif.g	<pre>tatzahed/ATAC_PRODECTS/Example_with_gr_files_Escond_sampledus_ALSecond_sampledus_T2: nodule load prime. Indule load prime. The second sampledus_ALSecond_sampledus_ALSecond_sampledus_ALSecond_sampledus_ALSecond_sampledus_ALSecond_sampledus_ALS: nodule load prime. The second sampledus_ALSecond_sampledus_ALS: nodule load prime. The second sampledus_ALS and the second sampledus_ALS. Second_sampledus_ALS: nodule load prime. The second sampledus_ALS. Second_sampledus_ALS: nodule load prime. The second sampledus_ALS: nodule load prime. The second sampledus_</pre>

345

1,0-1 All

346 S-S-Fig. 13: I-ATAC automatically generated script

	t Example_with_g	z_files Input Direc	tory /data/	zahmed/ATA0	_PROJECTS/gz_fastq_	files
Total FastQ Files: 4						
amples>File R1: Firt_Sample	eData_R1.fastqFile R	2: Firt_SampleData_	R2.fastq			
Samples>File R1: Second_Sa	mpleData_R1.fastqFi	le R2: Second_Samp	oleData_R2.f	astq		
	II 🔽	✓ Auto Correct	Out	put Directory	/data/zahmed/ATAC	_PROJECTS/Example_with_gz_files
Run ATAC Seq.	✓ FastQC	✓ Trimmomatic	✓ BWA	✓ Sam Sort	✓ Mark Duplicates	✓ Insert Size
	✓ BAM Shifter	✓ Sam tools	✓ Samtoo	ols Index	✓ Bed Tools	✓ Macs2: Peak Calling
2280021.helix-master fa_m 2280022.helix-master STDI		0 C high_r 00:00:08 R batch 00:00:00 R batch				
2280021.helix-master fa_m 2280022.helix-master STDI 2280023.helix-master STDI 2280030.helix-master rapso 2280037.helix-master ser 2280038.helix-master ser 2280039.helix-master ser 2280039.helix-master ser 2280040.helix-master ser 2280041.helix-master s.f 2280042.helix-master ser	wgs_p1.sh slek N leed1 N hongb sarch.sh hongb s_bcells.sh emarquez sus_cd14.sh emarque 8.sorted.sh emarquez all2016.sh emarquez ensus_th.sh emarque	00:00:08 R batch 00:00:00 R batch 0 Q batch 2 01:00:40 R ba 2 00:58:55 R b 2 00:59:28 R b 2 00:59:21 R ba	mem atch batch batch atch atch atch			0

347

348 S-Fig. 14: I-ATAC job queued.

349 **7.1.3 Output**

- After the successful execution of the ATAC-seq data processing pipeline, the system's generated output can be located in the mentioned output directory (S-Fig. 15). The project directory contains
- 352 automatically generated and run script:

353	Example_with_gz_files_Example_with_gz_files.sh
354	copied, pasted and unzipped "FASTQ" files:
355 356 357 358	Firt_SampleData_R1.fastq Firt_SampleData_R2.fastq Second_SampleData_R1.fastq Second_SampleData_R2.fastq
359	and sub-directories:
360 361 362	Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2, Example_with_gz_files_Second_SampleData_R1_Second_SampleData_R2, MergedSamples
363 364 365 366	The output files were placed in proposed system's automatically created sub-directory structure (Section: Applications integration, data processing pipeline and project's directory structure), as shown in S-Fig. 14. We also input two samples and asked system to produce merged replicates as well. So, we observed results for both samples as well as merged replicates.

	● ● ① 2ahmed — zahmed@helix:/data/zahmed/ATAC_PROJECTS/Example_with_gz_files — ssh — 183×14
368	[zahmed@helix Example_with_gz_files]\$ ls -l total 4153 -rwr-r-1 zahmed jaxuser 29832 Jan 13 17:19 Example_with_gz_files_Example_with_gz_files_sh drwxr-sr-x 4 zahmed jaxuser 5 Jan 13 7:19 Example_with_gz_files_firt_SampleData_R1_Firt_SampleData_R2 -rwr-r- 1 zahmed jaxuser 192 Jan 13 7:19 Example_with_gz_files_econd_SampleData_R1_Firt_SampleData_R2 -rwr-r- 1 zahmed jaxuser 5 Jan 13 7:19 Example_with_gz_files_econd_SampleData_R1_Firt_SampleData_R2 -rwr-r- 1 zahmed jaxuser 7 J4798 Jan 13 7:18 Firt_SampleData_R1.fastq -rw-r-r-r- 1 zahmed jaxuser 7 J4798 Jan 13 17:28 Excond_SampleData_R1.fastq -rw-r-r-r- 1 zahmed jaxuser 7 J4798 Jan 13 17:28 Firt_SampleData_R1.fastq -rw-r-r-r- 1 zahmed jaxuser 7 J4798 Jan 13 17:28 Firt_SampleData_R1.fastq -rw-r-r-r- 1 zahmed jaxuser 7 J4798 Jan 13 17:28 Second_SampleData_R1.fastq -rw-r-r-r- 1 zahmed jaxuser 7 J4798 Jan 13 17:28 Second_SampleData_R2.fastq -rw-r-r-r-r-1 zahmed jaxuser 7 J4798 Jan 13 17:28 Second_SampleData_R1.fastq -rw-r-r-r-r-1 zahmed jaxuser 7 J4798 Jan 13 17:28 Second_SampleData_R2.fastq -rw-r-r-r-r-1 zahmed jaxuser 7 J4798 Jan 13 17:28 Second_SampleData_R2.fastq [zahmed@helix Example_with_gz_files]5
369 370	S-Fig. 15: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output project directory and files
371 372	The produced results from First_SampleData are shown in S-Fig. 16, including quality reports in FASTQC directory, which are:
373 374 375 376 377	Firt_SampleData_R1_fastqc.html Firt_SampleData_R1_fastqc.zip Firt_SampleData_R2_fastqc.html Firt_SampleData_R2_fastqc.zip trimmed "FASTQ" files in trimmomatic directory, which are:
378 379 380 381	Firt_SampleData_R1.fastq_filtered Firt_SampleData_R1.trimU.fastq Firt_SampleData_R2.fastq_filtered Firt_SampleData_R2.trimU.fastq
382 383	all sorted, shifted and indexed "sam" and "bam", "bed" and related files are placed in "bwa" directory, which are:
384 385 386 387 388 389 390 391 392 393	Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2.sam Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_sorted.sam Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted.bam Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_bam Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam.bai Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam.bai Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted.bed Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_metrics.txt Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_metrics.pdf Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_insertSize.txt
394	and all produced results at peak calling were placed in "macs2" directory, which are:
395 396 397 398 399 400 401 402 403	Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_control_lambdabdg Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_peaks.broadPeak Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_peaks.gappedPeak Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_peaks.sls Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_peaks.xls Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_peaks.xls Example_with_gz_files_Firt_SampleData_R1_Firt_SampleData_R2_rmdup_shifted_sorted.bam_sorted_treat_pileup.bd g
404	

- 406 S-Fig. 16: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output files for
- 407 Firt_SampleData

405

408 Likewise, First_SampleData, the produced results from Second_SampleData are shown in S-Fig. 17.

- 410 S-Fig. 17: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output files for
- 411 Second_SampleData
- 412 The produced results from merged samples are shown in S-Fig. 18.

<pre>[zahmedghelix Example_with_gz_files]s of MergedSamples/ [zahmedghelix KergedSamples]s ts -1 total 5866 -rw-rr 1 zahmed jaxuser 252113 Jan 13 17:20 Example_with_gz_files_mergedSample.bam -rw-rr 1 zahmed jaxuser 1793149 Jan 13 17:20 Example_with_gz_files_mergedSample.sam -rw-rr 1 zahmed jaxuser 181296 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai -rw-rr 1 zahmed jaxuser 181296 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai -rw-rr 1 zahmed jaxuser 181296 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai -rw-rr 1 zahmed jaxuser 176901 Jan 13 17:20 macs2 [zahmedghelix macs2]s ts -1 total 968 -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloreAdPeak -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloreAdPeak -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloreAdPeak -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloreAdPeaks.slo</pre>		
<pre>total 5866 - nwr-r 1 zahned jaxuser 252113 Jan 13 17:20 Example_with_gz_files_mergedSample.bam -nwr-r 1 zahned jaxuser 1973140 Jan 13 17:20 Example_with_gz_files_mergedSample.sam -nwr-r 1 zahned jaxuser 252113 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam -nwr-r 1 zahned jaxuser 1421265 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam -nwr-r 1 zahned jaxuser 374 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed dixxr=sr-x 2 zahned jaxuser 374 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed dixxr=sr-x 2 zahned jaxuser 374 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_bed dixxr=sr-x 2 zahned jaxuser 374 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -nwr-r 1 zahned jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -nwr-r 1 zahned jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.spapedPeak -nwr-r 1 zahned jaxuser 1655 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -nwr-r 1 zahned jaxuser 1055 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -nwr-r 1 zahned jaxuser 1055 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -nwr-r</pre>		
<pre>-rwrw1 zahmed jaxuser 725113 Jan 13 17:20 Example_with_gz_files_mergedSample.bam -rwrw 1 zahmed jaxuser 1973140 Jan 13 17:20 Example_with_gz_files_mergedSample.bam -rwrw 1 zahmed jaxuser 7451240 Jan 13 17:20 Example_with_gz_files_mergedSample_corted.bam.bai -rwrw 1 zahmed jaxuser 7452680 Jan 13 17:20 Example_with_gz_files_mergedSample_corted.bam.bai /rwrw 1 zahmed jaxuser 552680 Jan 13 17:20 Example_with_gz_files_mergedSample_corted.bam.bai /zahmed@helix MergedSamples]\$ cd macs2/ /zahmed@helix MergedSamples]\$ cd macs2/ /zahmed@helix Macs25 ls -1 total 968 -rwrw 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rwrw 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_ceaks.spapedPack -rwrw 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloraddPack -rwrw 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloraddPack -rwrw 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sloraddPack</pre>		j)\$ (5 - L
<pre>-rwr 1 zahmed jaxuser 1973149 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam -rw-rr 1 zahmed jaxuser 252113 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai -rw-rr 1 zahmed jaxuser 1481296 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai -rw-rr 1 zahmed jaxuser 562680 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed drwxr-sr-x 2 zahmed jaxuser 374 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed Zahmed@helix MergedSamples]\$ cd macs2/ Zahmed@helix MargedSamples]\$ cd macs2/ Zahmed@helix MargedSamples]\$ cd macs2/ Zahmed@helix macs2]\$ ls -l total 968 -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.spapedPeak -rw-rr 1 zahmed jaxuser 1655 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr 1 zahmed jaxuser 1855 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr 1 Jahmed jaxuser 1855 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr 1 Jahmed jaxuser 200331 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr 1 Jahmed jaxuser 200331 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr 1 Jahmed jaxuser 200331 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr 1 Jahmed jaxuser 200332 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls -rw-rr</pre>		- 253112 Jan 12 17/20 Eventle with an files manadfample han
<pre>-rwr 1 zahed jaxuser 725113 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam. -rwr 1 zahed jaxuser 1481260 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai /rw-rr 1 zahed jaxuser 552600 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed drxr-sr-x 2 zahed jaxuser 374 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed /zaheddehelix MergedSamples]s cd macs2 /zaheddehelix macs2]s ls -l total 960 -rw-rr 1 zahed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr 1 zahed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_ceaks.gappedPeak -rw-rr 1 zahed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.storadPeak -rw-rr 1 zahed jaxuser 1365 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.storedPeaks.storedPeaks.storedPeaks.storedSample_sorted_peaks.storedSample_sorted_peaks.storedSample_sorted_peaks.storedSample.sorted_peaks.storedSample.sorted_peaks.storedSample_sorted_peaks.storedS</pre>		
<pre>-rwr 1 zahmed jaxuser 1481296 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bam.bai -rwr 2 zahmed jaxuser 562680 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed drwx-rs-x 2 zahmed jaxuser 374 Jan 13 17:20 macs2 [zahmed@helix macs2]\$ is -l total 968 -rw-rr- 1 zahmed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.sls</pre>		
<pre>-rwrw 1 zahmed jaxuser 552680 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed drxw-rs-r 2 zahmed jaxuser 374 Jan 13 17:20 macs2 [zahmedghelix MergedSamples]s cd macs2/ [zahmedghelix macs2]s l = -1 total 968 -rwrw 1 zahmed jaxuser 176991 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rwrw 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rwrw 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rwrw 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.slpcadPeak -rwrw 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.slpcadPeak -rwrw 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.slpcadPeak -rwrw 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg</pre>		
drxx-sr-x 2 zahmed jaxuser 374 Jan 13 17:20 macs2 [zahmedghelix MergedSamples]\$ cd macs2/ total 968 -nw-rr 1 zahmed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -nw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -nw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.appedPeak -nw-rr 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.als -nw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg		
Izahmedghelix MergedSamples]s ed macs2/ Izahmedghelix macs2]s ls -l Izahmedghelix macs2]s ls -l -rw-rr- 1 zahmed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr- 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 10 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.gappedPeaks -rw-rr 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.gappedPeaks	-rw-rr 1 zahmed jaxuse	r 562680 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted.bed
Izahmedghelix mac52]\$ ls -l total 960 -mwr-rwr-1 zahmed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -mwr-rwr-1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -mwr-rwr-1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.appedPeak -mwr-rwr-1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.als -mwr-rwr-1 zahmed jaxuser 1053 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.als	drwxr-sr-x 2 zahmed jaxuse	r 374 Jan 13 17:20 macs2
total 968 -rw-rr-1 zahmed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr-1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr-1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls -rw-rr-1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls	[zahmed@helix MergedSample	s]\$ cd macs2/
total 968 -rw-rr-1 zahmed jaxuser 176901 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr-1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr-1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls -rw-rr-1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treaks.xls -rw-rr-1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg	zahmed@helix macs21\$ ls -	
-rw-rr1 zahmed jaxuser 176001 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_control_lambda.bdg -rw-rr 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.agpedPeak -rw-rr 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls -rw-rr 1 zahmed jaxuser 105 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg		
-rw-rr- 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.broadPeak -rw-r-r-r- 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.gappedPeak -rw-r-r 1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls		r 176901 Jan 13 17:20 Example with oz files mergedSample sorted control lambda.bdg
-rw-rr- 1 zahmed jaxuser 0 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.gappedPeak -rw-r-r 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls -rw-r-r 1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg		
-rw-rr 1 zahmed jaxuser 1165 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_peaks.xls -rw-rr 1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg		
-rw-rr- 1 zahmed jaxuser 200932 Jan 13 17:20 Example_with_gz_files_mergedSample_sorted_treat_pileup.bdg		
	-rw-rr 1 zanmed jaxuse [zahmed@helix macs2]\$	Zobasz Jan is 1/:20 Example_with_gz_files_mergeosample_soried_treat_pileup.Dog

414 S-Fig. 18: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output files for Merged Samples

415 **7.2** Case Study 2: Using GM12878 – CD4 T- Cells

416 7.2.1 Dataset Details

- 417 Information about Raw dataset (GM12878 CD4 T- Cell, Day 1, Rep1 SRR891275 and Rep2
- 418 SRR891276) is available at web link
- 419 (https://catalog.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM12878&product=CC and
- 420 produced results, which are mentioned in this case study can be downloaded from the following
- 421 project web link.

422 **7.2.2 Input**

- 423 The input to I-ATAC is the path to ATAC-seq sample data (S-Fig. 19), which in our case is:
- 424

413

"/data/zahmed/ATAC_seq_data/CD4"

425

/uuu/2unmeu/ATAC_seq_uuu/CD4

```
    Comparison C
```

426

- S-Fig. 19: Screen shot (Linux Terminal, using Mac-OS-X) of compressed sample data files of GM12878 CD4 T- Cell,
 Day 1, Rep1 SRR891275 and Rep2 SRR891276.
- 429 Likewise, earlier discussed case study, at successful identification and verification of sample data
- 430 files (S-Fig. 20), data processing job was created (S-Fig. 21) and successfully queued (S-Fig. 22).

431

I

		I-ATAC ver 1.0.1 - Public R		
		Process Settings	_	
Seq. Protocol 2x 🗘 Projec	ct GM12878_CD4_Day	1 Input Directory /data/zahmed/AT	AC_PROJECTS/CD4	
		Auto Correct Output Director	v /data/zahmed/ATAC	_PROJECTS/GM12878_CD4_Day1
Run ATAC Seq.		Trimmomatic 🗹 BWA 🗹 Sam Sor		✓ Insert Size
	✓ BAM Shifter ✓	Sam tools 🗹 Samtools Index	✓ Bed Tools	✓ Macs2: Peak Calling
	• • •	Hi zahmed, its JAX-ATACSe	c-GUI	
		Here is your job description: -> Project: GM12878_CD4_Day1		
	2 -	-> walltime: 300:30:30		
		-> nodes: 1 -> ppn: 1		
		-> Input Files should be at: /data/zahm -> Results will be at: /data/zahmed/AT		
		Do you want me submit it to the helix ?		
		,	No	Yes
Please write for assistance and troubleshoot -Fig. 20: I-ATAC, inpu		^{pr} please call at: 860-837-2063 et parameters and verification	on	I-ATAC ver. 1.0.1 is developed t
		et parameters and verification // <i>data/zahmed/ATAC_PRO</i>	JECTS/CD4	I-ATAC ver. 1.0.1 is developed t
		et parameters and verification	JECTS/CD4	I-ATAC ver. 1.0.1 is developed b
-Fig. 20: I-ATAC, inpu	ıt sample data, se	et parameters and verification // <i>data/zahmed/ATAC_PRO</i>	JECTS/CD4 Day1	I-ATAC ver. 1.0.1 is developed h
-Fig. 20: I-ATAC, inpu //in/tash 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1	ut sample data, se	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1	JECTS/CD4 Day1 D4_Day1 - esh - 261x68	
-Fig. 20: I-ATAC, inpu //in/tash 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1	ut sample data, se	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1	JECTS/CD4 Day1 D4_Day1 - esh - 261x68	
-Fig. 20: I-ATAC, inpu //in/tash 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1 85 -4 oft1225_CG4_Day1	ut sample data, se	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1	JECTS/CD4 Day1 D4_Day1 - esh - 261x68	
-Fig. 20: I-ATAC, input	tt sample data, se tt sample data, se t //data/zabwed/AMAC_MADICTS/0413270_C ddle laad bestools ; nedule load MASS ddl r ranseati//bus/mas2 ; of /data/ framewati/bus/bus/2 ; of /data/ framewati/bus/bus/bus/2 ; of /data/ framewati/bus/bus/bus/2 ; of /data/ framewati/bus/bus/bus/bus/bus/bus/bus/bus/bus/bus	et parameters and verificatio /data/zahmed/ATAC_PRO GM12878_CD4_L wmd - zahmedBreitz/data/zahmed/ATAC_PROJECTE/GM12876 C - zahmedBreitz/data/zahmed/ATAC_PROJECTE/GM12876 C - zahmedBreitz/data/zahmed/ATAC_PROJECTE/GM12876 C - January - Janua	JECTS/CD4 Day1 D4_Day1 - ssh - 261+58 AL_CD4.ATACseq_Day1_8e01_578 ay1/0413878_C04_Day1_6401_578 ay1/0413878_C04_Day1_6401_578 ay1/0413878_C04_Day1_6401_578 ay1/0413878_C04_Day1_6401_578 ay1/04138801278_st_241_5785 ay1/04138801278_st_241_5785 ay1/04138801278_st_241_5785 ay1/04138801278_st_241_5785 ay1/04138801278_st_241_5785 ay1/04138801258012580000000000000000000000000000	X1_ALL : nodule load python : nodule load R : nodule 1 yl_Repl_SR891273_R1_ALL_(D+_ATACken_DayRepl_SR8912 D+_ATACken_DayRayALL_YALL_YALL_YALL_YALL_YALL_YALL_YAL
-Fig. 20: I-ATAC, input	tt sample data, se ti sample data, se ti sample data, se ti sample data, se ti sample data ti sample dat	et parameters and verificatio /data/zahmed/ATAC_PRO GM12878_CD4_1 wmed - zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_C /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed@nelic/data/zahmed/ATAC_PROLECTS/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac_prolects/0012878_CA /zahmed/atac/prolects/00128788_CA /zahmed/atac/prolects/00128788_CA /zahmed/atac/prolects/00128788_CA /zahmed/atac/prolects/00128788_CA /zahmed/atac/prolects/	JECTS/CD4 Day1 D4.Day1 - 10h - 201+58 ALL-CD4- ATACseq_Bay1_Rep1_S98891275 ay2cdC1297_CD4_Bay2_Bay2_Bay2_Bay2_Bay2_Bay2_Bay2_Bay2	BJ.ALL : models load python : models load B : models ly ALL : models load python : models load B : models ly ALL : COM - ATACCHE, Day J. Bell, DBBBD12 : Del. ATACHE, Day J. Bell, DBBBD12 : Del. ATACHE, DBJ.C.C. / ATACHE, DBJ.C. / ATACHE, DDJ.C.
-Fig. 20: I-ATAC, input -Fig. 20: I-ATAC, input ************************************	tt sample data, se ti sample data, se r/des/demo///def_PROJECT/001973, d wite lad beforit; neuvic tas Mac model lad beforit; neuvic tas Mac model rimmedic/bu/mac2; c //def_Artace_Dat_Res_SMB0273, d rimmedic/bu/mac2; c rimmedic/bu/mac2; c rimmedic/bu/m	et parameters and verificatio /data/zahmed/ATAC_PRO GM12878_CD4_1 wmed - zahmed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 //ansed@helic/data/zahmed/ATAC_PROJECTE/001287.04 /ansed	JECTS/CD4 Day1 D4.Day1 - tsh - 201+58 ALLCO4- ATACseq_Bay1_8ep1_598891275 ay2601297LC9-20158 ALLCO4- ATACseq_Bay1_8ep1_598891275 ay2601297LC9-20158892125,125 ALLCO4- ATACseq_Bay1_8ep1_588 ALLCO4- ATACseq_Bay1_8ep1_58	BJ.ALL : models load python : models load B : models ly ALL : models load python : models load B : models ly ALB : models load B : models l
-Fig. 20: I-ATAC, input -Fig. 20: I-ATAC, input ////////////////////////////////////	/deta/zahmed/ATMC_M0021CT%_00112870_C /deta/zahmed/ATMC /deta/zahmed/ATMC <td>et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/ata/sahmed/ATAC_PROJECTS/01287.cd /data/zahmed/ata/sahmed/</td> <td>JECTS/CD4 Day1 D4_Day1 - ssh - 281×68 ALL_C04 - ATACSER_BAY1.Rep1_S98891275 sp1/of1287_C04_Bay1_C04_ATACSER_BAY ACKSER_D04_Rep1_S9891275_R1_ALL_C ACKSER_D04_Rep1_S9891275_R1_ALL_C Immontic-1.32.1 pr Ff_/da1/athech ACKSER_D04_Rep1_S9891275_R1_ALL_C D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_R04_R048_D04_R048_D04_R048_D04_R048_R048_R048_R048_R048_R048_R048_R0</td> <td>JEJ_ALL ; mobile load python ; mobile load A ; mobile] Mobile J. Mobile J. Mobile</td>	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/ata/sahmed/ATAC_PROJECTS/01287.cd /data/zahmed/ata/sahmed/	JECTS/CD4 Day1 D4_Day1 - ssh - 281×68 ALL_C04 - ATACSER_BAY1.Rep1_S98891275 sp1/of1287_C04_Bay1_C04_ATACSER_BAY ACKSER_D04_Rep1_S9891275_R1_ALL_C ACKSER_D04_Rep1_S9891275_R1_ALL_C Immontic-1.32.1 pr Ff_/da1/athech ACKSER_D04_Rep1_S9891275_R1_ALL_C D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_R04_R048_D04_R048_D04_R048_D04_R048_R048_R048_R048_R048_R048_R048_R0	JEJ_ALL ; mobile load python ; mobile load A ; mobile] Mobile J. Mobile
-Fig. 20: I-ATAC, input -Fig. 20: I-ATAC, input ////////////////////////////////////	/deta/zahmed/ATMC_M0021CT%_00112870_C /deta/zahmed/ATMC /deta/zahmed/ATMC <td>et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/ata/sahmed/ATAC_PROJECTS/01287.cd /data/zahmed/ata/sahmed/</td> <td>JECTS/CD4 Day1 D4_Day1 - ssh - 281×68 ALL_C04 - ATACSER_BAY1.Rep1_S98891275 sp1/of1287_C04_Bay1_C04_ATACSER_BAY ACKSER_D04_Rep1_S9891275_R1_ALL_C ACKSER_D04_Rep1_S9891275_R1_ALL_C Immontic-1.32.1 pr Ff_/da1/athech ACKSER_D04_Rep1_S9891275_R1_ALL_C D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_R04_R048_D04_R048_D04_R048_D04_R048_R048_R048_R048_R048_R048_R048_R0</td> <td>JEJ_ALL ; mobile load python ; mobile load A ; mobile] Mobile J. Mobile J. Mobile</td>	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/milc/data/zahmed/ATAC_PROJECTS/011287.cd /data/zahmed/ata/sahmed/ATAC_PROJECTS/01287.cd /data/zahmed/ata/sahmed/	JECTS/CD4 Day1 D4_Day1 - ssh - 281×68 ALL_C04 - ATACSER_BAY1.Rep1_S98891275 sp1/of1287_C04_Bay1_C04_ATACSER_BAY ACKSER_D04_Rep1_S9891275_R1_ALL_C ACKSER_D04_Rep1_S9891275_R1_ALL_C Immontic-1.32.1 pr Ff_/da1/athech ACKSER_D04_Rep1_S9891275_R1_ALL_C D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_Rep1_S9891275_R2_ALL_sorted:as D4_D04_R04_R048_D04_R048_D04_R048_D04_R048_R048_R048_R048_R048_R048_R048_R0	JEJ_ALL ; mobile load python ; mobile load A ; mobile] Mobile J. Mobile
//interfactors/	At sample data, se at sample data at a at sample data at a at sample data at a at sample data at a at a sample data at a	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmed - zahmed@melic/data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/aTaca/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ataca/zahmed/ataca/zahmed/ataca/zahmed/ataca/zahmed/	JECCTS/CD4 Day1 D4.Day1 - ssh - 201x58 D4.Day1 - 201	 J.B.Z.ALL ; module load python ; module load A ; module 1 Y.L.Reyl J.MMB1277, N.S.ALL (Chi - ATACAER, Dayl, Mod. J.S.BMB12 (Chi - A) (Chi - A)
Articles 20: I-ATACe, input Fig. 20: I-ATAC, input Fig. 20: I-ATACe, input </td <td>//dota/cabmed/ATMC_MP024CT5/0012878.c //dota/cabmed/ATMC_MP024CT5/0012878.c //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc /dota/cabmed/atmc /dota/cabmed/atmc</td> <td>et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@melic/data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/atacs_projects/samed/atacs/projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs/ /</td> <td>JECCTS/CD4 Day1 D4.Day1 - ssh - 201x68 D4.Day1 - 201x6 D4.Day1 - 201x7 D4.Day1 - 201x7 D4.Day2 - 201x7 D4</td> <td> J.B. ALL ; module load python ; module load R ; module 1 (1, 1, 2, 4, 1, 2, 6, 1, 1, 1, 2, 6, 1, 2, 1, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,</td>	//dota/cabmed/ATMC_MP024CT5/0012878.c //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc //dota/cabmed/atmc /dota/cabmed/atmc	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@melic/data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/atacs_projects/samed/atacs/projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs/ /	JECCTS/CD4 Day1 D4.Day1 - ssh - 201x68 D4.Day1 - 201x6 D4.Day1 - 201x7 D4.Day1 - 201x7 D4.Day2 - 201x7 D4	 J.B. ALL ; module load python ; module load R ; module 1 (1, 1, 2, 4, 1, 2, 6, 1, 1, 1, 2, 6, 1, 2, 1, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
Frige. 20: I-ATACC, input Frige. 20: I-ATAC, I-ATAC, I-ATAC, I-A, I-A, I-A, I-A, I-A, I-A, I-A, I-A	<pre>// dots/samed/ATMC_PROJECTS/OR12878_C // dots/samed/atmc/projects/samed/samed/ // dots/samed/atmc/projects/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/samed/ // dots/samed/s</pre>	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmd - zahmed@melic/data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/atacs_projects/samed/atacs/projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs_projects/samed/atacs/ /data/zahmed/atacs/ /	JECCTS/CD4 Day1 D4.Day1 - ssh - 201x68 D4.Day1 - 201x6 D4.Day1 - 201x7 D4.Day1 - 201x7 D4.Day2 - 201x7 D4	 J.R. ALL ; module load python ; module load R ; module 1 (1, 1, 2, 4, 1, 2, 6, 1, 1, 1, 2, 6, 1, 2, 1, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
-Fig. 20: I-ATACC, input	<pre>// dots/samed/ATMC_PROJECTS/OR12878_C // dots/samed/atmc/projects/samed/samed/ // dots/samed/atmc/projects/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/ // dots/samed/samed/samed/samed/samed/samed/samed/ // dots/samed/s</pre>	et parameters and verification /data/zahmed/ATAC_PRO GM12878_CD4_1 mmed - zahmed@melic/data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ATAC_PROJECTS/01287_Cd /data/zahmed/ataches/zahmed/ataches/zahm	JECCTS/CD4 Day1 D4.Day1 - ssh - 201x68 D4.Day1 - 201x6 D4.Day1 - 201x7 D4.Day1 - 201x7 D4.Day2 - 201x7 D4	 J.R. ALL ; module load python ; module load R ; module 1 (1, 1, 2, 4, 1, 2, 6, 1, 1, 1, 2, 6, 1, 2, 1, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1,8-1 All

438 S-Fig. 21: I-ATAC generated script

439 **7.2.3 Output**

- 440 As in the earlier discussed case study, all the produced results were placed in the proposed and auto
- 441 generated directory structure (S-Fig. 22, 23and 24).

[zahmed@helix GM12878 CD4 Da	/11\$ \s -\
total 12408081	
-rwxr-xr-x 1 zahmed jaxuser	2351167270 Jan 15 14:52 CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL.fastq
rwxr-xr-x 1 zahmed jaxuser	2351167270 Jan 15 14:52 CD4+ ATACseq Day1 Rep1 SRR891275 R2 ALL.fastq
rwxr-xr-x 1 zahmed jaxuser	3279168066 Jan 15 14:52 CD4+ ATACseq Day1 Rep2 SRR891276 R1 ALL.fastq
rwxr-xr-x 1 zahmed jaxuser	3279168066 Jan 15 14:52 CD4+ ATACseq Day1 Rep2 SRR891276 R2 ALL.fastq
rwxr-sr-x 4 zahmed jaxuser	53 Jan 15 14:55 GM12878 CD4 Day1 CD4+ ATACseq Day1 Rep1 SRR891275 R1 ALL CD4+ ATACseq Day1 Rep1 SRR891275 R2 ALL
rwxr-sr-x 4 zahmed jaxuser	53 Jan 15 15:49 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL
rw 1 zahmed jaxuser	4966515 Jan 15 17:03 GM12878_CD4_Day1.e985775
rw-rr 1 zahmed jaxuser	10964 Jan 15 14:55 GM12878_CD4_Day1_GM12878_CD4_Day1.sh
rw 1 zahmed jaxuser	272 Jan 15 15:50 GM12878_CD4_Day1.0985775
irwxr-sr-x 3 zahmed jaxuser	303 Jan 15 17:00 MergedSamples
	y1]\$ cd GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL_CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL
	y1_CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL_CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL]\$ cd fastQC/
[zahmed@helix fastQC]\$ ls -l	
total 2878	
	383986 Jan 15 14:56 CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL_fastqc.html
	518046 Jan 15 14:56 CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL_fastqc.zip
	376161 Jan 15 14:56 CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL_fastqc.html
	504211 Jan 15 14:56 CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL_fastqc.zip
zahmed@helix fastQC]\$ cd	
	y1_CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL_CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL]\$ cd trimmomatic/
zahmed@helix trimmomatic]\$.s -l
otal 5019864	
irwxr-sr-x 3 zahmed jaxuser	1353 Jan 15 15:47 bwa
	2222603428 Jan 15 14:58 CD4+_ATACseq_Day1_Rep1_SRR091275_R1_ALL.fastq_filtered
	61389838 Jan 15 14:58 CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL.trimU.fastq
	2225050298 Jan 15 14:58 CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL.fastq_filtered
-rw-rr i zanmed jaxuser [zahmed@helix trimmomatic]\$	49900792 Jan 15 14:58 CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL.trimU.fastq
	.a dwa/
[zahmed@helix bwa]\$ ls –l total 10945305	
-rw-rr 1 zahmed jaxuser	8621 Jan 15 15:42 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL_CD4+_ATACseq_Day1_Rep1_SRR891275_R2_ALL_rmdup_insertSize.txt
rw-rr 1 zahmed jaxuser	oozi Jan 15.42 GM12878_CO4_DAY_CO4+ATACseq_DAY1Repi_SRR891275_R1ALL_CO4+ATACSeq_DAY1_Repi_SRR891275_R2_ALL_IMUdp_metrics.cdf
rw-rr 1 zahmed jaxuser	1475 Jan 15.476 UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676 2538 Jan 15 15:41 UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ676_UNIZ6
	236 Jai 13.5.4 OHZAZO/CUC-JUCY_LOV-INCSELD91_CPE1_SNR091275.11_ALL_CUCY_INCSELD91_CPE1_SNR091275_R2_LL_TWUDP.ICO
	1900200325 Jan 19 19 19 19 19 19 19 19 19 19 19 19 19
	370832294 Jan 15 15:47 GH12878 CO4 Dayi LCH-A HACSeq Dayi Rep1_SHNB91275 R1 ALL CH+ ATACSeq Dayi Rep1_SHNB91275 R2 ALL HUDu Shifted sorted bam
rw-rr 1 zahmed jaxuser	5847992 Jan 15 15:47 GM1287 C04 Day1 C04 ATACKED Day1 RRF91275 R1 ALL C04 ATACKED Day1 Rep1 SRR891275 R2 ALL rndup shifted sorted.bai
	483741438 Jan 15 15:47 (M12872 C04 Day1 CD4 ATACSeg Day1 Rep SRR891275 R1 ALL CD4+ ATACSeg Day1 Rep1 SRR891275 R2 ALL rndup shifted sorted.bed
	3396028093 Jan 15 15:35 GM12878 (D4 Day1 CD4+ ATACSeg Day1 Rep1 SRR891275 R1 ALL CD4+ ATACSeg Day1 Rep1 SRR891275 R2 ALL sam
	3396028038 Jan 15 15:39 GM12878 Cp4 Davi C04+ ATACseo Davi Rep1 SRR891275 R1 ALL C04+ ATACseo Davi Rep1 SRR891275 R2 ALL sorted.sam
rwxr-sr-x 2 zahmed jaxuser	809 Jan 15 15:49 macs2
zahmed@helix bwal\$ cd macs2	
zahmed@helix macs21\$ ls -l	
otal 555680	
	168991285 Jan 15 15:48 GM12878 CD4 Dav1 CD4+ ATACseg Dav1 Rep1 SRR891275 R1 ALL CD4+ ATACseg Dav1 Rep1 SRR891275 R2 ALL rmdup shifted sorted.bam sorted control lambda.bdg
	1149502 Jan 15 15:49 GM12878 CD4 Dav1 CD4+ ATACseg Dav1 Rep1 SRR891275 R1 ALL CD4+ ATACseg Dav1 Rep1 SRR891275 R2 ALL imdup shifted sorted bam sorted peaks.broadPeak
	13531059 Jan 15 15:49 GM12878 CD4 Dav1 CD4+ ATACseg Dav1 Rep1 SRR091275 R1 ALL CD4+ ATACseg Dav1 Rep1 SRR091275 R2 ALL mmdup shifted sorted bam sorted peaks.gappedPeak
	11743053 Jan 15 15:48 GM12878 CD4 Davi CD4+ ATACseg Davi Rep1 SRR091275 R1 ALL CD4+ ATACseg Davi Rep1 SRR091275 R2 ALL modup shifted sorted bam sorted peaks.xls
	298027077 Jan 15 15:48 GM12878 CD4 Dav1 CD4+ ATACseg Dav1 Rep1 SRR891275 R1 ALL CD4+ ATACseg Dav1 Rep1 SRR891275 R2 ALL rmdup shifted sorted.bam sorted treat pileup.bdg
zahmed@helix macs2]\$	

442

S-Fig. 22: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output main project directory and files,
 and for sample CD4+_ATACseq_Day1_Rep1_SRR891275

<pre>medmblix GH127F_COL_DyvL_CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL]s cd fastQC/ medmblix fastQC is 1s -1 l 2934 r 1 zahmed jaxuser 389214 Jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_fastqc.html r 1 zahmed jaxuser 324573 Jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_fastqc.html r 1 zahmed jaxuser 324573 Jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_fastqc.html r 1 zahmed jaxuser 324573 Jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL]s cd rimmomatic/ medMeltuk fastQCI is cd 1 G93240 r 1 zahmed jaxuser 3892642 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL]s cd rimmomatic/ medMeltuk fastQCI is cd 1 G93240 r 1 zahmed jaxuser 3892680854 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL]s cd rimmomatic/ medMeltuk trimmomaticJ is cd 1 G93240 r 1 zahmed jaxuser 3892680854 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL.freq r 1 zahmed jaxuser 3892680834 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r 1 zahmed jaxuser 9568 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fostq r 1 zahmed jaxuser 9568 Jan 15 16:44 CM12878_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL.redup_metrics.pdf r 1 zahmed jaxuser 9568 Jan 15 16:44 CM12878_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL.redup_metrics.pdf r 1 zahmed jaxuser 2223 Jan 15 16:43 CM12878_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL.redup_metrics.pdf r 1 zahmed jaxuser 2223 Jan 15 16:43 CM12878_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.pdf r</pre>	
<pre>medBetLx fistClS is -1 1 2934 Fr 1 zahmed jaxuser 389214 Jan 15 15:50 (D4+_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_fastqc.rin Fr 1 zahmed jaxuser 327273 Jan 15 15:50 (D4+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastqc.rin Fr 1 zahmed jaxuser 327272 Jan 15 15:50 (D4+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastqc.rin Fr 1 zahmed jaxuser 327272 Jan 15 15:50 (D4+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastqc.rin medBetLx GT2278_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastqc.rin medBetLx GT2278_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastq_filtered Fr 1 zahmed jaxuser 327253 Jan 15 16:51 bus Fr 1 zahmed jaxuser 327253 Jan 15 15:51 C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastq_filtered Fr 1 zahmed jaxuser 327253 Jan 15 16:51 bus Fr 1 zahmed jaxuser 327253 Jan 15 15:51 C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastq_filtered Fr 1 zahmed jaxuser 327253 Jan 15 15:51 C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastq_filtered Fr 1 zahmed jaxuser 3271551 C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastq_filtered Fr 1 zahmed jaxuser 6012766 Jan 15 15:51 C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_fastq_filtered Fr 1 zahmed jaxuser 15283 Jan 15 16:44 GM1287_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR891276_R2_ALL_rinuU_insertSize.txt Fr 1 zahmed jaxuser 12283 Jan 15 16:44 GM1287_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR891276_R2_ALL_rinuU_insertSize.txt Fr</pre>	🖲 🔵 🌑 🏠 zahmed — zahmed@helix:/data/zahmed/ATAC_PROJECTS/GM12878_CD4_Day1/GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_DAY1_Rep2_SRR891276_R1_ALL_CD4+_ATACSeq_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSeq_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSeq_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SRR891276_R1_ALL_CD4+_ATACSEQ_DAY1_REP2_SR891ARL_CD4+_ATACSEQ_R1_R1_R1_R1_R1_R1_R1_R1_R1_R1_R1_R1_R1_
<pre>1 2914 1 2914 1</pre>	[zahmed@helix GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL]\$ cd fastQC/
<pre>r 1 zahmed jacuser 392714 Jan 15 15:50 CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL_fostqc.ring r 1 zahmed jacuser 326927 Jan 15 15:50 CD+_ATACseq_Dayl_Rep2_SR801726_R2_ALL_fostqc.ring r 1 zahmed jacuser 327827 Jan 15 15:50 CD+_ATACseq_Dayl_Rep2_SR801726_R2_ALL_fostqc.ring meddPelik f1287_CDL_Dayl_CD+_ATACseq_Dayl_Rep2_SR801726_R2_ALL_fostqc.ring meddPelik f1287_CDL_Dayl_CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL_fostqc.ring meddPelik f1287_CDL_Dayl_CD+_ATACseq_Dayl_Rep2_SR801726_R2_ALL_fostqc.ring meddPelik f1287_CDL_Dayl_CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL_fostqc.ring meddPelik f1287_CDL_Dayl_CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL.fostqc.filtered r 1 zahmed jacuser 3025609865 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL.trinut.fostq r 1 zahmed jacuser 3025609865 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL.trinut.fostq r 1 zahmed jacuser 3025609865 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL.trinut.fostq r 1 zahmed jacuser 3025609865 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SR801726_R1_ALL.trinut.fostq r 1 zahmed jacuser 3025609865 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SR801276_R1_ALL.trinut.fostq r 1 zahmed jacuser 3052609865 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SR801276_R1_ALL.trinut.fostq r</pre>	[zahmed@helix fastQC]\$ ls -l
<pre>r 1 zahmed jaxuer 52457 Jan 15 15:50 CD+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_fastqc.thi r 1 zahmed jaxuer 52782 Jan 15 15:50 CD+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_fastqc.thi medMetLix forLi378_CD4_Jayl_CD4+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_fastqc.thi medMetLix forLi378_CD4_Jayl_CD4+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SR801276_R2_ALL]5 cd trimmonatic/ 1 590240 </pre>	total 2934
<pre>r 1 zahmed jaxuer 524579 jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_[fastqc.htm r 1 zahmed jaxuer 526720 Jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_fastqc.htm r 1 zahmed jaxuer 527220 Jan 15 15:50 CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_fastqc.htm medMeilix follor r 1 zahmed jaxuer 52680956 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_fastqc.htm S092240 1 zahmed jaxuer 9268260956 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_fastqc.htm S092240 1 zahmed jaxuer 9256256 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL.fastq_filtered r 1 zahmed jaxuer 9256256 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL.fastq_filtered r 1 zahmed jaxuer 9256256 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL.trimU.fastq medMeilix follor r 1 zahmed jaxuer 9256256 Jan 15 15:51 CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL.trimU.fastq medMeilix fullor r 1 zahmed jaxuer 9560 Jan 15 15:40 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL.CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL.trimU.fastq medMeilix fullor r 1 zahmed jaxuer 9560 Jan 15 16:44 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 12203 Jan 15 16:44 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 2723 Jan 15 16:43 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 4205830664 Jan 15 16:43 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 4205830664 Jan 15 16:43 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 4205830664 Jan 15 16:43 (M12078_CO4_Dayl_CO+_ATACseq_Dayl_Rep2_SR801276_R1_ALL_CO+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 4205831</pre>	rw-rr 1 zahmed jaxuser 389214 Jan 15 15:50 CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_fastqc.html
<pre>r 1 zahmed jauuer 527829 Jan 15 15:59 (D++_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_C0+_ATACseq_Dayl_Rep2_SR8891276_R2_ALL]s dd trimmomatic/ meddwelix is/10278_CO4_Dayl_C0+_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_C0+_ATACseq_Dayl_Rep2_SR8891276_R2_ALL]s dd trimmomatic/ meddwelix is/10278_CO4_Dayl_C0+_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_fastq_filtered r 1 zahmed jauuer 302162516 Jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_fastq_filtered r 1 zahmed jauuer 302162516 Jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_trinut_fastq r 1 zahmed jauuer 302162516 Jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SR8891276_R2_ALL.fratu_fastq r 1 zahmed jauuer 692182680 Jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SR8891276_R2_ALL.trinut_fastq meddwelix inimanaticjs cd bawa/ meddwelix bulls is -1 1 zahmed jauuer 795680963 Jan 15 16:44 CM12878_C04_Dayl_C04_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_trinut_fastq meddwelix bulls is -1 1 zahmed jauuer 7925630646 Jan 15 16:44 CM12878_C04_Dayl_C04_ATACseq_Dayl_Rep2_SR8891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR8991276_R2_ALL_redup_metrics.pdf r 1 zahmed jauuer 79258 Jan 15 16:44 CM12878_C04_Dayl_C04_ATACseq_Dayl_Rep2_SR8991276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR8991276_R2_ALL_redup_metrics.stxt r 1 zahmed jauuer 79259 Jan 15 16:43 CM12878_C04_Dayl_C04_ATACseq_Dayl_Rep2_SR8991276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR8991276_R2_ALL_redup_metrics.pdf r 1 zahmed jauuer 792593 Jan 15 16:43 CM12878_C04_Dayl_C04_ATACseq_Dayl_Rep2_SR8991276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR8991276_R2_ALL_redup_metrics.stxt r 1 zahmed jauuer 792593 Jan 15 16:43 CM12878_C04_Dayl_C04_ATACseq_Dayl_Rep2_SR8991276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SR8991276_R2_ALL_redup_shifted_bam r</pre>	-rw-rr 1 zahmed jaxuser 524579 Jan 15 15:50 CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_fastqc.zip
<pre>medHelix fistOC[s c1 medHelix fistOC[s c1 m</pre>	-rw-rr 1 zahmed jaxuser 390892 Jan 15 15:50 CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_fastqc.html
medghelix immonatic1 is -1 i G992240 -fr-rx 3 zahned jaxuser 1333 Jan 15 16:51 kma -r 1 zahned jaxuser 1353 Jan 15 16:51 kma -r 1 zahned jaxuser 135893934 Jan 15 16:45 (M12276_CMA_Dyrl_CMA_ATACseq_Dayl_Rep2_SR891276_R1_ALL.triul.fastq medghelix timomatic1s cd bwa/ medghelix kmanatils cd bwa/ medghelix bwa/ medghelix bwa/ -r 1 zahned jaxuser 19268 Jan 15 16:43 (M12276_CMA_Dyrl_CM4_ATACseq_Dyrl_Rep2_SR891276_R1_ALL_CM4_ATACseq_Dyrl_Rep2_SR891276_R2_ALL_redup_metrics_petrics_txt -r 1 zahned jaxuser 19278 Jan 15 16:43 (M12276_CMA_Dyrl_CM4_ATACseq_Dyrl_Rep2_SR891276_R1_ALL_CM4_ATACseq_Dyrl_Rep2_SR891276_R2_ALL_redup_shifted_sorted.bam -r 1 zahned jaxuser 149278 Jan 15 16:50 (M12276_CMA_Dyrl_CM4_ATACseq_Dyrl_Rep2_SR891276_R1_ALL_CM4_ATACseq_Dyrl_Rep2_SR891276_R2_ALL_redup_shifted_sorted.bam bai -r 1 zahned jaxuser 149278 Jan 15 16:51 (M12276_CMA_Dyrl_CM4_ATACseq_Dyrl_Rep2_SR891276_R1_ALL_CM4_ATACseq_Dyrl_Rep2_SR891276_R2_ALL_redup_shifted_sorted.bam bai -r 1 zahned jaxuser 149278 Jan 15 16:51 (M12276_CMA_Dyrl_CM4_ATACseq_Dyrl_Rep2_SR891276_R1_ALL_CM4_ATACseq_Dyrl_Rep2_SR891276_R2_ALL_redup_shifted_so	-rw-rr 1 zahmed jaxuser 527829 Jan 15 15:50 CD4+_ATACseg_Day1_Rep2_SRR891276_R2_ALL_fastgc.zip
<pre>med@helik it rimmomaticls is1 i commod jaxuser 3253 Jan 15 16:51 bwa r-r-r 1 zahmed jaxuser 325639654 Jan 15 16:51 bwa r-r-r 1 zahmed jaxuser 325639654 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r-r 1 zahmed jaxuser 325639654 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.frain_fittered r-r 1 zahmed jaxuser 325639654 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.frain_fittered r-r 1 zahmed jaxuser 325639654 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.frain_fittered redPetik basis Js -1 I 1 zahmed jaxuser 325637664 Jan 15 15:51 CD+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.cD+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.pdf r-r 1 zahmed jaxuser 325638646 Jan 15 16:43 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.pdf r-r 1 zahmed jaxuser 22530 Jan 15 16:44 0M12278_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.pdf r-r 1 zahmed jaxuser 22530 Jan 15 16:43 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.pdf r-r 1 zahmed jaxuser 2455638664 Jan 15 16:43 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.txt r-r 1 zahmed jaxuser 2455638664 Jan 15 16:43 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.txt r-r 1 zahmed jaxuser 35989580 Jan 15 16:51 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metrics.txt r-r 1 zahmed jaxuser 425283253 Jan 15 16:51 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metric4_bam_bai r-r 1 zahmed jaxuser 425283257 Jan 15 16:51 0M12276_CD4_Dayl_CD4_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_redup_metric4_bam_bai r-r 1 zahmed jaxuser 425283257 Jan 15 16:53 0M12876_CD4_Dayl_CD4_ATACs</pre>	Izahmed@helix fastOC]\$ cd
<pre>1 699244 Fr-r 1 zahmed jaxuser 3025689865 Jan 15 15:51 D04-XTACseq_Dayl_Rep2_SR891276_R1_ALL.fstq_filtered Fr 1 zahmed jaxuser 3025689865 Jan 15 15:51 C04+XTACseq_Dayl_Rep2_SR891276_R1_ALL.triuU.fastq Fr 1 zahmed jaxuser 3025690863 Jan 15 15:51 C04+XTACseq_Dayl_Rep2_SR891276_R1_ALL.triuU.fastq Fr 1 zahmed jaxuser 30152686 Jan 15 15:51 C04+XTACseq_Dayl_Rep2_SR891276_R1_ALL.triuU.fastq medMetLix trimomaticly cd bwa/ medMetLix trimomaticly cd bwa/ medMet</pre>	[zahmed@helix GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL]\$ cd trimmomatic/
<pre>r-r: 1 zahmed javuser 1353 Jan 15 16:51 bwa r-r: 1 zahmed javuser 20263516 Jan 15 15:51 C04_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r-r: 1 zahmed javuser 20263516 Jan 15 15:51 C04_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r-r: 1 zahmed javuser 69182680 Jan 15 15:51 C04_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r-r: 1 zahmed javuser 69182680 Jan 15 15:51 C04_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r-r: 1 zahmed javuser 69182680 Jan 15 15:51 C04_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fastq_filtered r-r: 1 zahmed javuser 69182680 Jan 15 15:51 C04_ATACseq_Dayl_Rep2_SRR891276_R1_ALL.fot+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL.rmdup_insertSize.txt r-r: 1 zahmed javuser 20560 Jan 15 16:44 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_rmdup_metrics.pdf r-r: 1 zahmed javuser 225538664 Jan 15 16:43 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_rmdup_metrics.txt r-r: 1 zahmed javuser 2455838664 Jan 15 16:43 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_rmdup_shifted_bam r-r: 1 zahmed javuser 2455838664 Jan 15 16:43 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_rmdup_shifted_bam r-r: 1 zahmed javuser 2455838664 Jan 15 16:43 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_rmdup_shifted_bam r-r: 1 zahmed javuser 49208411 Jan 15 16:59 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_rmdup_shifted_sorted.bam r-r: 1 zahmed javuser 4722832873 Jan 15 16:59 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_indup_shifted_sorted.bam.soit r-r: 1 zahmed javuser 4722832873 Jan 15 16:53 GM12878_C04_Dayl_C04+_ATACseq_Dayl_Rep2_SRR891276_R1_ALL_C04+_ATACseq_Dayl_Rep2_SRR891276_R2_ALL_indup_shifted_sorted.bam.soit r-r: 1 zahmed javuse</pre>	zahmed@helix trimmomatic]\$ ls -l
<pre>r: 1 zahmed javuer 3025609365 jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SRR801276_R1_ALL.trinut.fattq r: 1 zahmed javuer 3055000034 jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SRR801276_R2_ALL.trinut.fattq r: 1 zahmed javuer 305500034 jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SRR801276_R2_ALL.trinut.fattq r: 1 zahmed javuer 305500034 jan 15 15:51 C0+_ATACseq_Dayl_Rep2_SRR801276_R2_ALL.trinut.fattq redeNetLix Value5 15 -1 1 14846505 r: 1 zahmed javuer 305500064 jan 15 16:44 CM12076_C04_Dayl_C0+_ATACseq_Dayl_Rep2_SRR801276_R1_ALL_C0+_ATACseq_Dayl_Rep2_SR801276_R2_ALL_redup_insertSize.txt r</pre>	total 6993240
<pre>r 1 zahmed jaxuer 92162316 Jan 15 15:10 C0+_ATACseq_Day1_Rep2_SRR891276_R1_ALL.triauL.fastq r 1 zahmed jaxuer 92080934 Jan 15 15:51 C0+_ATACseq_Day1_Rep2_SRR891276_R2_ALL.triauL.fastq medMetLix insumatic1s_cd bwa/ medMetLix bual5 ts -1 r 1 zahmed jaxuer 9560 Jan 15 16:44 CM12878_C04_Day1_C04_ATACseq_Day1_Rep2_SRR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SRR891276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 9560 Jan 15 16:44 CM12878_C04_Day1_C04_ATACseq_Day1_Rep2_SRR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 2723 Jan 15 16:44 CM12878_C04_Day1_C04_ATACseq_Day1_Rep2_SRR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_redup_metrics.pdf r 1 zahmed jaxuer 255638664 Jan 15 16:43 CM12878_C04_Day1_C04_ATACseq_Day1_Rep2_SRR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_redup_metrics.txt r 1 zahmed jaxuer 4295638664 Jan 15 16:43 CM12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_redup_shifted_bam r 1 zahmed jaxuer 491084511 Jan 15 16:49 CM12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_redup_shifted_somt r 1 zahmed jaxuer 491084511 Jan 15 16:30 CM12878_C04_Day1_C04_ATACSeq_Day1_Rep2_SR891276_R1_ALL_C04_ATACSeq_Day1_Rep2_SR891276_R2_ALL_redup_shifted_somt r</pre>	drwxr-sr-x 3 zahmed jaxuser 1353 Jan 15 16:51 bwa
<pre>r 1 zahmed jaxuer 3095000034 Jan 15 15:51 CD4ATAC:seq_Day1_Rep2_SRR091276_R2_ALL.trinuLiftered r 1 zahmed jaxuer 50122608 Jan 15 15:51 CD4ATAC:seq_Day1_Rep2_SRR091276_R2_ALL.trinuLiftered medMetLik valuer 1203 Jan 15 16:44 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_insertSize.txt r 1 zahmed jaxuer 50560 Jan 15 16:44 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_insertSize.txt r 1 zahmed jaxuer 1203 Jan 15 16:44 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_insertSize.txt r 1 zahmed jaxuer 42014051 Jan 15 16:49 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_insertSize.txt r 1 zahmed jaxuer 42014051 Jan 15 16:49 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_insertSize.txt r 1 zahmed jaxuer 42014051 Jan 15 16:90 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_instretd_sorted.bam r 1 zahmed jaxuer 50900561 Jan 15 16:19 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_instretd_sorted.bam.bai r 1 zahmed jaxuer 50900561 Jan 15 16:151 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_instretd_sorted.bam.bai r 1 zahmed jaxuer 421282857 Jan 15 16:36 GM12078_CD4_Day1_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATAC:seq_Day1_Rep2_SRR091276_R2_ALL_rmdup_instretd_sorted.bam.bai r</pre>	-rw-rr 1 zahmed jaxuser 3092689056 Jan 15 15:51 CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL.fastq_filtered
<pre>r 1 zahmed jaxuser 60182680 Jan 15 15:51 (D4+_ATACseq_Day1_Rep2_SR8091276_R2_ALL.trimU.fastq medMelix timumantic1s cd bow/ medMelix imumantic2s cd bow/ medMelix imumantic3s cd bow/</pre>	-rw-rr 1 zahmed jaxuser 92162516 Jan 15 15:51 CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL.trimU.fastq
medMetik tvimomaticjs cd bwa/ medMetik tvimomaticjs cd bwa/ medMetik tvimomaticjs cd bwa/ medMetik tvisols 1s -1 1 14884585 r-r 1 zahmed jaxuser 9560 Jan 15 16:44 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_intertis.pdf r-r 1 zahmed jaxuser 12283 Jan 15 16:44 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_intertis.pdf r-r 1 zahmed jaxuser 2273 Jan 15 16:44 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_intertis.pdf r-r 1 zahmed jaxuser 15090500 Jan 15 16:49 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_intertis.stt r-r 1 zahmed jaxuser 50909508 Jan 15 16:59 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_interd_sorted.bam r-r 1 zahmed jaxuser 4721827823 Jan 15 16:51 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_interd_sorted.bam.bai r-r 1 zahmed jaxuser 4721827823 Jan 15 16:53 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_interd_sorted.bam.bai r-r 1 zahmed jaxuser 4721827827 Jan 15 16:36 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_sorted.bam r-r 1 zahmed jaxuser 4721827827 Jan 15 16:36 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_sorted.bam r-r 1 zahmed jaxuser 4721827827 Jan 15 16:36 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_sorted.bam r-r 1 zahmed jaxuser 472182787 Jan 15 16:35 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_cD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_sorted.bam r-r 1 zahmed jaxuser 478243247 R09 Jan 15 16:53 GM12078_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_CD4+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_mdup_shifted_sorted.bam_sorted_co	-rw-rr 1 zahmed jaxuser 3095898034 Jan 15 15:51 CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL.fastq_filtered
<pre>medBetLix buols 1s -1 1 4804585 r 1 zahmed jaxuser 9560 Jan 15 16:44 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_insertSize.txt r 1 zahmed jaxuser 2223 Jan 15 16:44 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_intert.cs.pdf r 1 zahmed jaxuser 422114851 Jan 15 16:43 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_intert.cs.pdf r 1 zahmed jaxuser 422114851 Jan 15 16:43 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_intert.cs.txt r 1 zahmed jaxuser 422114851 Jan 15 16:43 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_interd.sorted.bam r 1 zahmed jaxuser 5980980 Jan 15 16:51 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_interd.sorted.bam.bai r 1 zahmed jaxuser 5980980 Jan 15 16:51 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_interd.sorted.bam.bai r 1 zahmed jaxuser 7203232 Jan 15 16:53 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_indup_interd_sorted.bam.sorted.bed r 1 zahmed jaxuser 7203232 Jan 15 16:53 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_indup_interd_sorted.bam.sorted.bed r 1 zahmed jaxuser 7203232 Jan 15 16:53 GM1287_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_indup_interd_sorted.bam.sorted_control_inabda.bdg r</pre>	rw-rr 1 zahmed jaxuser 69102608 Jan 15 15:51 CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL.trimU.fastq
1 1 200000 1 2000000000000000000000000000000000000	zahmed@helix trimmomatic]\$ cd bwa/
<pre>r 1 ahmed javuser 9560 Jan 15 16:44 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_insertSize.txt r 1 ahmed javuser 2123 Jan 15 16:44 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_insertSize.txt r 1 ahmed javuser 422114851 Jan 15 16:43 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_insertSize.txt r 1 ahmed javuser 422114851 Jan 15 16:43 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_insertSize.txt r 1 ahmed javuser 4202114851 Jan 15 16:49 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_insertGay1ed_Sorted.bam r 1 ahmed javuser 5090980 Jan 15 16:51 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_instited_sorted.bam.bai r 1 ahmed javuser 72132337 Jan 15 16:51 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_instited_sorted.bam.bai r 1 ahmed javuser 721323327 Jan 15 16:51 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_rmdup_instited_sorted.bam.bai r 1 ahmed javuser 72132337 Jan 15 16:53 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R2_ALL_ICM+ATACseq_Day1_Rep2_SR8891276 R2_ALL_iSM r 1 ahmed javuser 72132337 Jan 15 16:53 GH12872_COM_Day1_COM_ATACseq_Day1_Rep2_SR8891276 R3_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R3_ALL_ISM r 1 ahmed javuser 809 Jan 15 16:53 GH12872_COM_Day1_CM+ATACseq_Day1_Rep2_SR8891276 R3_ALL_COM_ATACseq_Day1_Rep2_SR8891276 R3_ALL_ISM r 1 ahmed javuser 809 Jan 15 16:53 GH12872_COM_DAy1_COM_ATACseq_Day1_Rep2_SR8891276_R3_ALL_ISM r</pre>	zahmed@helix bwa]\$ ls -l
<pre>r 1 ahmed jaxuser 12203 Jun 15 16:44 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891275_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rndup_metrics.txf r 1 ahmed jaxuser 2253 Jun 15 16:43 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891275_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 4211481 Jun 15 16:43 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 42121481 Jun 15 16:54 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 42121481 Jun 15 16:56 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 421232332 Jun 15 16:56 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rndup_shifted_sorted.bam_soit r 1 ahmed jaxuser 421233332 Jun 15 16:56 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_son r 1 ahmed jaxuser 421233332 Jun 15 16:56 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_son r 1 ahmed jaxuser 42233333 Jun 15 16:53 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_son r 1 ahmed jaxuser 4203333 Jun 15 16:53 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM+_ATACseq_Day1_Rep2_SR891276_R2_ALL_son r 1 ahmed jaxuser 4921333 Jun 15 16:53 GH1287_COM_Day1_CdH_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_son r</pre>	otal 14884585
<pre>r 1 ahmed javuser 2723 Jan 15 16:43 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891275 R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_sht r 1 ahmed javuser 452536664 Jan 15 16:43 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891275 R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_sht1ed.bam r 1 ahmed javuser 4525143 Jan 15 16:43 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891275 R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_sht1ed.bam r 1 ahmed javuser 45273723 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_sht1ed.bam.bai r 1 ahmed javuser 7597823 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_sht1ed_sorted.bam.bai r 1 ahmed javuser 721832873 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276 R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_sht1ed_sorted.bam.bai r 1 ahmed javuser 721832873 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 ahmed javuser 721832873 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 ahmed javuser 721832873 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276_R1_ALL_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 ahmed javuser 728313491 Jan 15 16:53 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR891276_R2_ALL_GM2_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r</pre>	
<pre>r 1 ahmed jaxuser 2455538664 Jun 15 16:43 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR89127G_R1_ALL_COM_ATACseq_Day1_Rep2_SR89127G_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 491095411 Jun 15 16:49 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR89127G_R1_ALL_COM_ATACseq_Day1_Rep2_SR89127G_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 502072832 Jun 15 16:51 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR89127G_R1_ALL_COM_ATACseq_Day1_Rep2_SR89127G_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 502072832 Jun 15 16:51 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR89127G_R1_ALL_COM_ATACseq_Day1_Rep2_SR89127G_R2_ALL_rndup_shifted.bam r 1 ahmed jaxuser 502072832 Jun 15 16:51 GH1287g_COM_Day1_COM_ATACseq_Day1_Rep2_SR89127G_R1_ALL_COM_ATACseq_Day1_Rep2_SR89127G_R2_ALL_rndup_shifted.bam r</pre>	
<pre>r 1 zahmed javuser 492114851 Jan 15 16:49 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891275 R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam r 1 zahmed javuser 5998980 Jan 15 16:59 GH12878_CO4_Day1_CO4_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.soi r 1 zahmed javuser 62973022 Jan 15 16:51 GH12878_CO4_Day1_CO4_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.soi r 1 zahmed javuser 4221832827 Jan 15 16:53 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 zahmed javuser 4221832827 Jan 15 16:53 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 zahmed javuser 422183287 Jan 15 16:54 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 zahmed javuser 429134931 Jan 15 16:53 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 zahmed javuser 429134931 Jan 15 16:53 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r</pre>	<pre>rrw=rr=rr= 1 zahmed jaxuser 2723 Jan 15 16:43 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rndup_metrics.txt</pre>
<pre>r 1 zahmed javuser 491985411 Jan 15 16:56 (H12878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam r 1 zahmed javuser 62597282 Jan 15 16:51 GH12878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sonted.bam r 1 zahmed javuser 472183285 Jan 15 16:36 GH12878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sonted.som r 1 zahmed javuser 472183285 Jan 15 16:36 GH12878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sonted.som r 1 zahmed javuser 472183285 Jan 15 16:36 GH12878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_c04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sonted.som red(betLix bua)5 cd macs2/ med(betLix bua)5 cd macs2/</pre>	-rw-rr 1 zahmed jaxuser 2455630664 Jan 15 16:43 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup.sam
<pre>r 1 zahmed javuser 55080968 Jan 15 16:51 GH12878_CO4_Day1_C04_ATACseq_Day1_Rep2_SR891275 R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.bai rr 1 zahmed javuser 4721323821 Jan 15 16:51 GH12878_CO4_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam rr- 1 zahmed javuser 4721323821 Jan 15 16:34 GH12878_CO4_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r-sr-x 2 zahmed javuser 4721323821 Jan 15 16:34 GH12878_CO4_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_Sam r-sr-x 2 zahmed javuser 4721332832 Jan 15 16:46 GH12878_CO4_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r-sr-x 2 zahmed javuser 472133281 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r-sr-x 1 zahmed javuser 472133281 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r-r-r 1 zahmed javuser 472133281 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.sorted_control_lambda.bdg r-rr 1 zahmed javuser 899139 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.sorted_peaks.appedFeak rr 1 zahmed javuser 899139 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_D391_Rep2_SR891276_R1_ALL_C04+_ATACseq_D391_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.sorted_peaks.appedFeak rr 1 zahmed javuser 8297313 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_D39176_R1_R1_C04+_ATACseq_D391_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.sorted_peaks.appedFeak rr 1 zahmed javuser 82973131 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseqD391_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.sorted_peaks.appedFeak rr 1 zahmed javuser 829731311 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseqD391_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam.sorted_peaks.appedFea</pre>	
<pre>r 1 ahmed javuser 625972823 Jan 15 16:15 (0112878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam_sorted.bed r 1 ahmed javuser 422183285 Jan 15 16:36 (0112878_C04_bay1_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sont red(betLx bua)5 cd macs2/ med(betLx bua)5 cd macs2/ me</pre>	
r 1 zahmed javuser 472183282 Jan 15 16:36 (H12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 zahmed javuser 472183287 Jan 15 16:46 (H12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sam r 1 zahmed javuser 282034031 Jan 15 16:53 (H12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_sorted.sam r 1 zahmed javuser 28031031 Jan 15 16:53 (H12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891376_R2_ALL_rndup_shifted_sorted.bam_sorted_control_lambda.bdg r 1 zahmed javuser 28031031 Jan 15 16:53 (H12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891376_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891376_R2_ALL_rndup_shifted_sorted.bam_sorted_peaks.broadPeak r 1 zahmed javuser 28031031 Jan 15 16:53 (H12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891376_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891376_R2_ALL_rndup_shifted_sorted.bam_sorted_peaks.broadPeak r 1 zahmed javuser 28031031 Jan 15 16:53 (H12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891376_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891376_R2_ALL_rndup_shifted_sorted.bam_sorted_peaks.broadPeak r 1 zahmed javuser 28031031 Jan 15 16:53 (H12876_C04_Day1_C04+_ATACseq_D3y1_Rep2_SR891376_R1_ALL_C04+_ATACseq_D3y1_Rep2_SR891376_R2_ALL_rndup_shifted_sorted.bam_sorted_peaks.broadPeak r 1 zahmed javuser 28031311 Jan 15 16:53 (H12876_C04_Day1_C04+_D3KCseq_D3y1_Rep2_SR891376_R1_ALL_C04+_ATACseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R1_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R14_ALL_C04+_R14Cseq_D3y1_Rep2_SR891376_R14_ALL_C04+_R140Cseq_D3y1_Rep2_SR891376_R14_ALL_C04+_R140Cseq_D3y1_Rep2_SR891376_R14_ALL_C04+_R140Cseq_D3y1_Rep2_SR891376_R14_ALL_C04+_R140Cseq_D3y1_Rep2_SR	
<pre>r 1 zahmed jaxuser 4221832857 Jan 15 16:48 GH12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_sorted.sam recdPetLix bua]5 cd macs2/ medPetLix bua]5 cd macs2/ medPetLix bua]5 cd macs2/ medPetLix bua]5 cd macs2/ medPetLix bua]5 cd macs2/ ner/</pre>	
r-sr-x 2 zahmed javuser 809 Jan 15 16:53 macs2 meddetlik Janks 6 d macs2 meddetlik Janks 6 d macs2 meddetlik macs215 1s - 1 1 yaB152 T 1 zahmed javuser 800138 Jan 15 16:53 CH12078 (DAL Day1_CAL_ATACseq_Day1_Rep2_SR801376_R1_ALL_CAL_ATACseq_Day1_Rep2_SR801376_R2_ALL_rmdup_shifted_sorted.bam_sorted_control_lambda.bdg T 1 zahmed javuser 800138 Jan 15 16:53 CH12078_CDA_Day1_CAL_ATACseq_Day1_Rep2_SR801376_R1_ALL_CAL_ATACseq_Day1_Rep2_SR801376_R2_ALL_rmdup_shifted_sorted.bam_sorted_pakks.proadPeak T 1 zahmed javuser 800138 Jan 15 16:53 CH12078_CDA_Day1_CAL_ATACseq_Day1_Rep2_SR801376_R1_ALL_CAL_ATACseq_Day1_Rep2_SR8013776_R2_ALL_rmdup_shifted_sorted.bam_sorted_pakks.proadPeak T 1 zahmed javuser 8201373 Jan 15 16:53 CH12078_CDA_Day1_COL_ATACseq_Day1_Rep2_SR801376_R1_ALL_COL_ATACseq_Day1_Rep2_SR8013776_R2_ALL_rmdup_shifted_sorted.pakks.proadPedeak T 1 zahmed javuser 8201373 Jan 15 16:53 CH12078_COL_Day1_COL_ATACseq_Day1_Rep2_SR801376_R1_ALL_COL_ATACseq_Day1_Rep2_SR8013776_R2_ALL_rmdup_shifted_sorted.pakks.proadPedeak	
medBelik bagis dd macs2/ medBelik bagis2]s is -i l 785152 r 1 rahmed jaxuser 220334931 Jan 15 16:53 GM12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR8891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR8891276_R2_ALL_mdup_shifted_sorted.bam_sorted_control_lambda.bdg rr 1 rahmed jaxuser 8091139 Jan 15 16:53 GM12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR8891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR8891276_R2_ALL_mdup_shifted_sorted.bam_sorted_control_lambda.bdg rr 1 rahmed jaxuser 8098027 Jan 15 16:53 GM12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR8891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR8891276_R2_ALL_mdup_shifted_sorted.peaks.papedPeak rr 1 rahmed jaxuser 8227853 Jan 15 16:53 GM12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR8891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR8891276_R2_ALL_mdup_shifted_sorted.peaks.appedPeak rr 1 rahmed jaxuser 8227853 Jan 15 16:53 GM12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR8891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR8891276_R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.xls	
medBelLix mac2]s ls -L 1 789152 r-r 1 zahmed jaxuser 228334931 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam_sorted_control_lambda.bdg r-r 1 zahmed jaxuser 48081139 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam_sorted_peaks.broadPeak r-r 1 zahmed jaxuser 4808027 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.peaks.appedPeak r-r 1 zahmed jaxuser 48278153 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.peaks.appedPeak r-r 1 zahmed jaxuser 48278153 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.peaks.appedPeak r-r 1 zahmed jaxuser 48278153 Jan 15 16:53 GH12876_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276_R2_ALL_rmdup_shifted_sorted.bam_sorted_treats_tileup.bdg	
l 795152 r-r 1 rahmed jaxuser 220334931 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_C04+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_shifted_sorted.bam_sorted_control_lambda.bdg r-r 1 rahmed jaxuser 8091139 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_C04+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_shifted_sorted.bam_sorted_peaks.broadPeak r-r 1 rahmed jaxuser 8098027 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_C04+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_shifted_sorted.bam_sorted_peaks.broadPeak r-r 1 rahmed jaxuser 8278153 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_C04+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_shifted_sorted.bam_sorted_peaks.slp r-r 1 rahmed jaxuser 8278153 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SRR091276_R1_ALL_C04+_ATACseq_Day1_Rep2_SRR091276_R2_ALL_rmdup_shifted_sorted.bam_sorted_trest_pileup.bdg	
r1 zahmed javuser 228330911 Jan 15 16:53 GH1287E_C04_Day1_C04_ATACseq_Day1_Rep2_SR8891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_control_lambda.bdg rr1 zahmed javuser 8091139 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.broadPeak rr1 zahmed javuser 9488827 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.broadPeak rr1 zahmed javuser 8278153 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.broadPedPeak rr1 zahmed javuser 8278153 Jan 15 16:53 GH12878_C04_Day1_C04+_ATACseq_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.sxls rr1 zahmed javuser 8278153 Jan 15 16:53 GH12878_C04_Day1_C04+_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.sxls rr1 zahmed javuser 8278153 Jan 15 16:53 GH12876_C04_Day1_C04+_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.sxls rr1 zahmed javuser 8278153 Jan 15 16:53 GH12876_C04_Day1_C04+_Day1_Rep2_SR891276 R1_ALL_C04+_ATACseq_Day1_Rep2_SR891276 R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.sxls	
r 1 zahmed javuser 8091139 Jan 15 16:33 (412876_004_0ay)_C04_ATACseq_0ay]_Rep_SRR891276_R1_ALL_C04+_ATACseq_0ay]_Rep_SRR891276_R12_ALL_mdu_shited_sorted.bam_sorted_peaks.broadPeak r 1 zahmed javuser 820852 Jan 15 16:33 (412876_004_0ay)_C04+_ATACseq_0ay]_Rep_SRR891276_R1_ALL_C04+_ATACseq_0ay]_Rep_SRR891276_R2_ALL_mdu_shited_sorted.bam_sorted_peaks.appedPeak r 1 zahmed javuser 8220853 Jan 15 16:33 (412876_004_0ay)_C04+_ATACseq_0ay]_Rep_SRR891276_R1_ALL_C04+_ATACseq_0ay]_Rep_SRR891276_R2_ALL_mdu_shited_sorted.bam_sorted_peaks.appedPeak r 1 zahmed javuser 8220853 Jan 15 16:33 (412876_004_0ay)_C04+_ATACseq_0ay]_Rep_SRR891276_R1_ALL_C04+_ATACseq_0ay]_Rep_SRR891276_R2_ALL_mdu_shited_sorted.bam_sorted_treats_tieup.bdg	otal 705152
r-r 1 zahmed jaxuser 9488027 Jan 15 16:53 GH12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR8801276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR801276_R2_ALLmdup_shifted_sorted.bam_sorted_peaks.gappedPeak r-r 1 zahmed jaxuser 8278153 Jan 15 16:53 GH12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR801276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR801276_R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.xls r-r 1 zahmed jaxuser 393530111 Jan 15 16:53 GH12878_C04_Day1_C04_ATACseq_Day1_Rep2_SR801276_R1_ALL_C04+_ATACseq_Day1_Rep2_SR801276_R2_ALL_mdup_shifted_sorted.bam_sorted_peaks.xls	
rr1 rahmed javuser 9278153 Jan 15 16:53 (H12876_C04_Day1_C04_ATACseq_Day1_Rep2_SRR801276_R1_ALL_C04+_ATACseq_Day1_Rep2_SRR801276_R12_ALL_rndu_shifted_sorted.bam_sorted_reaks.xtis	
rr 1 zahmed jaxuser 393539111 Jan 15 16:53 GM12878_CD4_Day1_CD4+_ATACseq_Day1_Rep2_SRR891276_R1_ALL_CD4+_ATACseq_Day1_Rep2_SRR891276_R2_ALL_rmdup_shifted_sorted.bam_sorted_treat_pileup.bdg	
ned@helix macs2]\$	[zahmed@helix macs2]\$

445

- 446 S-Fig. 23: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output files for sample
- 447 CD4+_ATACseq_Day1_Rep2_SRR891276

• •	😭 zahmed —	zahmed@helix:/data/zahmed/ATAC_PROJECTS/GM12878_CD4_Day1/MergedSamples/macs2 — ssh — 14	2×20
[zahmed@helix	GM12878_CD4_Da	y1]\$ cd MergedSamples/	
[zahmed@helix	MergedSamples]	\$ ls -l	
total 7175522			
-rw-rr 1 z	ahmed jaxuser	861237223 Jan 15 16:55 GM12878_CD4_Day1_mergedSample.bam	
-rw-rr 1 z	ahmed jaxuser	3676571000 Jan 15 17:00 GM12878_CD4_Day1_mergedSample.sam	
-rw-rr 1 z	ahmed jaxuser	861237223 Jan 15 16:59 GM12878_CD4_Day1_mergedSample_sorted.bam	
-rw-rr 1 z	ahmed jaxuser	6057328 Jan 15 16:59 GM12878_CD4_Day1_mergedSample_sorted.bam.bai	
-rw-rr 1 z	ahmed jaxuser	1109713470 Jan 15 17:00 GM12878_CD4_Day1_mergedSample_sorted.bed	
drwxr-sr-x 2 z	ahmed jaxuser	349 Jan 15 17:03 macs2	
[zahmed@helix	MergedSamples]	\$ cd macs2/	
[zahmed@helix	macs2]\$ ls -l		
total 693456			
-rw-rr 1 z	ahmed jaxuser	206569472 Jan 15 17:02 GM12878_CD4_Day1_mergedSample_sorted_control_lambda.bdg	
-rw-rr 1 z	ahmed jaxuser	5015956 Jan 15 17:03 GM12878_CD4_Day1_mergedSample_sorted_peaks.broadPeak	
-rw-rr 1 z	ahmed jaxuser	6688771 Jan 15 17:03 GM12878_CD4_Day1_mergedSample_sorted_peaks.gappedPeak	
-rw-rr 1 z	ahmed jaxuser	5231644 Jan 15 17:03 GM12878_CD4_Day1_mergedSample_sorted_peaks.xls	
-rw-rr 1 z	ahmed jaxuser	405274624 Jan 15 17:02 GM12878_CD4_Day1_mergedSample_sorted_treat_pileup.bdg	
[zahmed@helix	macs2]\$		

449

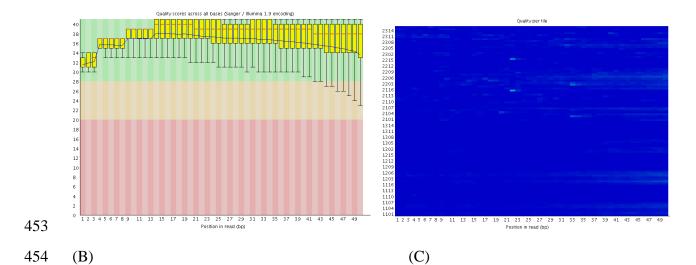
450 S-Fig. 24: Screen shot (Linux Terminal, using Mac-OS-X) of produced I-ATAC output files for Merged Samples

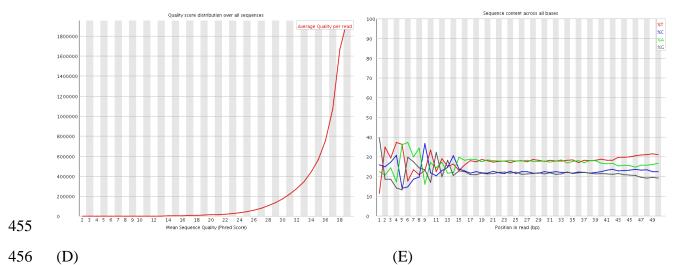
Basic Statistics

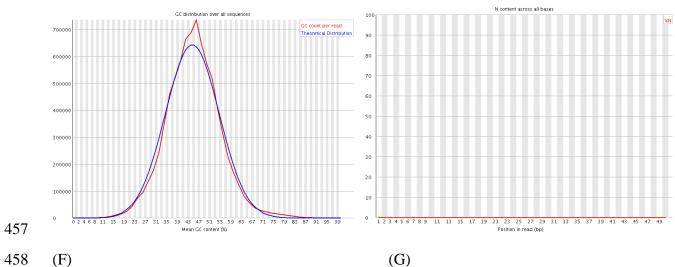
Measure	Value
Filename	CD4+_ATACseq_Day1_Rep1_SRR891275_R1_ALL.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	8086940
Sequences flagged as poor quality	0
Sequence length	50
%GC	44



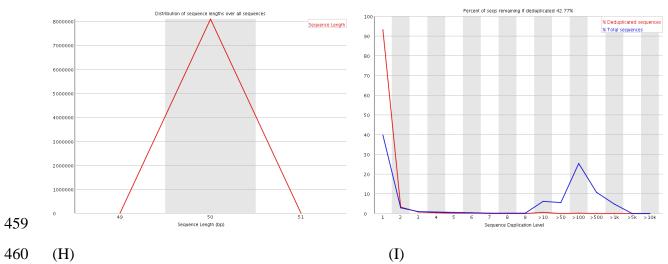
452 (A)



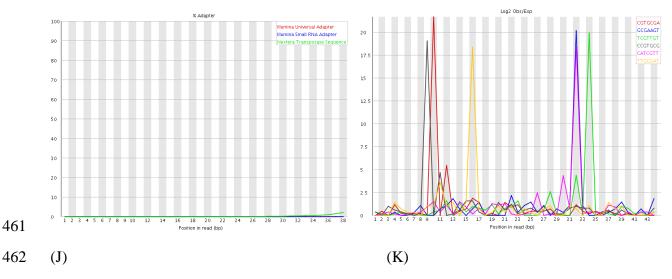








(F)



463 S-Fig. 25. Statistics and quality results produced by the FASTQC for the CD4+_ATACseq_Day1_Rep1_SRR891275.

464

465 Quality stats produced by the FASTQC (at CD4+_ATACseq_Day1_Rep1_SRR891275) are shown in

466 S-Fig. 26, which are: Basic Statistics (S-Fig. 25. A), Per base sequence quality (S-Fig. 25. B), Per tile

467 sequence quality (S-Fig. 25. C), Per sequence quality scores (S-Fig. 25. D), Per base sequence

468 content (S-Fig. 25. E), Per sequence GC content (S-Fig. 25. F), Per base N content (S-Fig. 25. G),

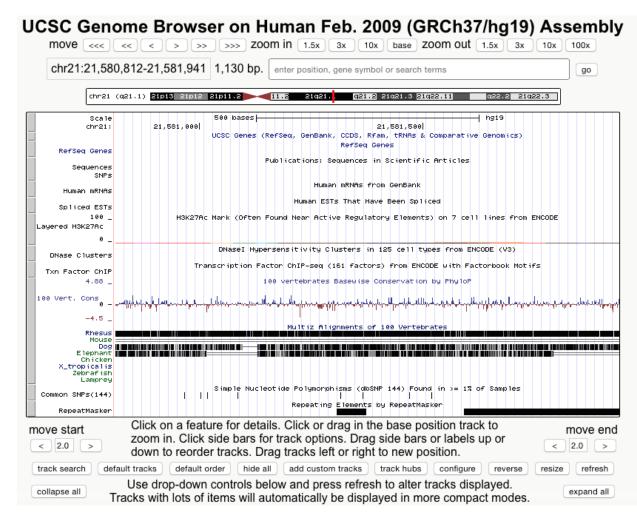
- 469 Sequence Length Distribution (S-Fig. 25. H), Sequence Duplication Levels (S-Fig. 25. I), Adapter
- 470 Content (S-Fig. 25. J) and Kmer Content (S-Fig. 25. K).
- 471



472

473 S-Fig. 26. Example Visualization of "bam" files (CD4+_ATACseq_Day1_Rep1_SRR891275,

474 CD4+_ATACseq_Day1_Rep2_SRR891276 and GM12878_CD4_Day1_mergedSample) using IGV.



- 476 S-Fig. 27 Example Visualization using "bdg" file (GM12878_CD4_Day1_mergedSample_sorted_control_lambda.bdg)
 477 using USCS Genome Browser.
- 478 The detailed output of samples a CD4+_ATACseq_Day1_Rep1_SRR891275 and
- 479 CD4+_ATACseq_Day1_Rep2_SRR891276 are attached in the supplementary material. Example
- 480 visualization of produced results is created by visualizing sorted "bam" files
- 481 (CD4+_ATACseq_Day1_Rep1_SRR891275, CD4+_ATACseq_Day1_Rep2_SRR891276 and
- 482 GM12878_CD4_Day1_mergedSample) using IGV (S-Fig. 26) and peak file
- 483 (CD4+_ATACseq_Day1_Rep1_SRR891275) using USCS Genome browser (S-Fig. 27).

484 8 Conclusions

- 485 To the best of our knowledge, I-ATAC platform is the first desktop tool that is specialized to
- 486 processing and analysis of ATAC-seq data. I-ATAC provides a flexible algorithm and parameter
- 487 setting GUI for non-computational scientists and a time-efficient parallel data analysis environment
- 488 for computational scientists. Future work includes incorporating visualization and differential
- 489 analysis modules in I-ATAC platform.
- 490

475

492 **9** Acknowledgments

The Jackson Laboratory (JAX) supports and owns this project. Special thanks to Ucar and
Banchereau labs as well as the Genome Technologies and Computational Sciences cores at JAX,
who provided insight and expertise that greatly, assisted the research and development of this
platform.

497 **10 Funding:**

498 This work was supported by The Jackson Laboratory, USA.

499 **11 Conflict of Interests:**

500 The authors declare that they have no competing interests.

501 12 Additional Requirements

- 502 For additional information, please refer to the project webpage: <u>https://www.jax.org/research-and-</u> 503 <u>faculty/tools/i-atac</u>
- 504 Source code, JAR files for MAC OS X and Windows, and complete source code package for Eclipse 505 IDE is available at <u>https://github.com/UcarLab/I-ATAC</u>
- 506 Example dataset is available at: <u>https://zenodo.org/record/46079#.WAe315MrK7Y</u>
- 507 Supporting software and dependencies are available at: 508 https://zenodo.org/record/162023#.WAe3dJMrK7Y

509 13 References

- 510 Ahmed Z, Ucar D. (2017) A standalone software platform for the interactive management and pre-
- 511 processing of ATAC-seq samples. *PeerJ Preprints.*, **5**:e2942v1 512 https://doi.org/10.7287/peerj.preprints.2942v1
- 513 Ahmed Z, Zeeshan S, Dandekar T. (2014) Developing sustainable software solutions for 514 bioinformatics by the "Butterfly" paradigm. *F1000Res.*, **3**, 71.
- 515 Bauch A. (2011) openBIS: a flexible framework for managing and analyzing complex data in 516 biology research. *BMC Bioinf.*, **12**, 468.
- 517 Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. et al. (2013) Transposition of native
- 518 chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and
- 519 nucleosome position. *Nat Methods.*, **10**, 1213-1218.
- Bolger AM, Lohse M, Usadel B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.*, **30**, 2114-20.
- 522 Dander A, Pabinger S, Sperk M, Fischer M, Stocker G, Trajanoski Z. (2014) SeqBench: integrated 523 solution for the management and analysis of exome sequencing data. *BMC Res Notes.*, **7**, 43.
- 524 Giardine B. (2005) Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* **15**, 1451-1455.

- 526 Horak CE, Snyder M. (2002) ChIP-chip: A genomic approach for identifying transcription factor
- 527 binding sites. *Methods Enzymol.*, **350**, 469–483
- 528 Li H, Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform.
- 529 *Bioinformatics.*, **25**, 1754-60.
- 530 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000
- 531 Genome Project Data Processing Subgroup. (2009) The Sequence alignment/map (SAM) format and 532 SAMtools. *Bioinformatics.*, **25**, 2078-2079.
- Mariette J, Escudié F, Allias N, Salin G, Noirot C, Thomas S, Klopp C. (2012) NG6: Integrated next
 generation sequencing storage and processing environment. *BMC Genomics*. 13, 462.
- 535 McLellan A S, Dubin RA, Jing Q, Broin PÓ, Moskowitz D, Suzuki M, Calder RB, Hargitai J,
- 536 Golden A, Greally JM. (2012) The Wasp System: an open source environment for managing and 537 analyzing genomic data. *Genomics.*, **100**, 345-51.
- Mount DM. (2004). Bioinformatics: Sequence and Genome Analysis (2nd ed.). Cold Spring Harbor
 Laboratory Press: Cold Spring Harbor, NY. ISBN 0-87969-608-7.
- 540 Orvis J, Crabtree J, Galens K, Gussman A, Inman JM, Lee E, Nampally S, Riley D, Sundaram JP,
- 541 Felix V, Whitty B, Mahurkar A, Wortman J, White O, Angiuoli SV. (2010) Ergatis: a web interface
- and scalable software system for bioinformatics workflows. *Bioinformatics*, **26**, 1488-92.
- Quinlan, A R, Hall IM. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.*, 26, 841-842.
- 545 Scholtalbers J. (2013) Galaxy LIMS for next-generation sequencing. *Bioinformatics*. **29**,1233-1234.
- Tsompana M, Buck MJ. (2014) Chromatin accessibility: a window into the genome. *Epigenetics & Chromatin.*, 7, 33.
- 548 Venco F, Vaskin Y, Ceol A, Muller H. (2014) SMITH: a LIMS for handling next-generation
- 549 sequencing workflows. *BMC Bioinf.*, **15**, (Suppl 14):S3.
- 550 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS.
- 551 (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
- 552