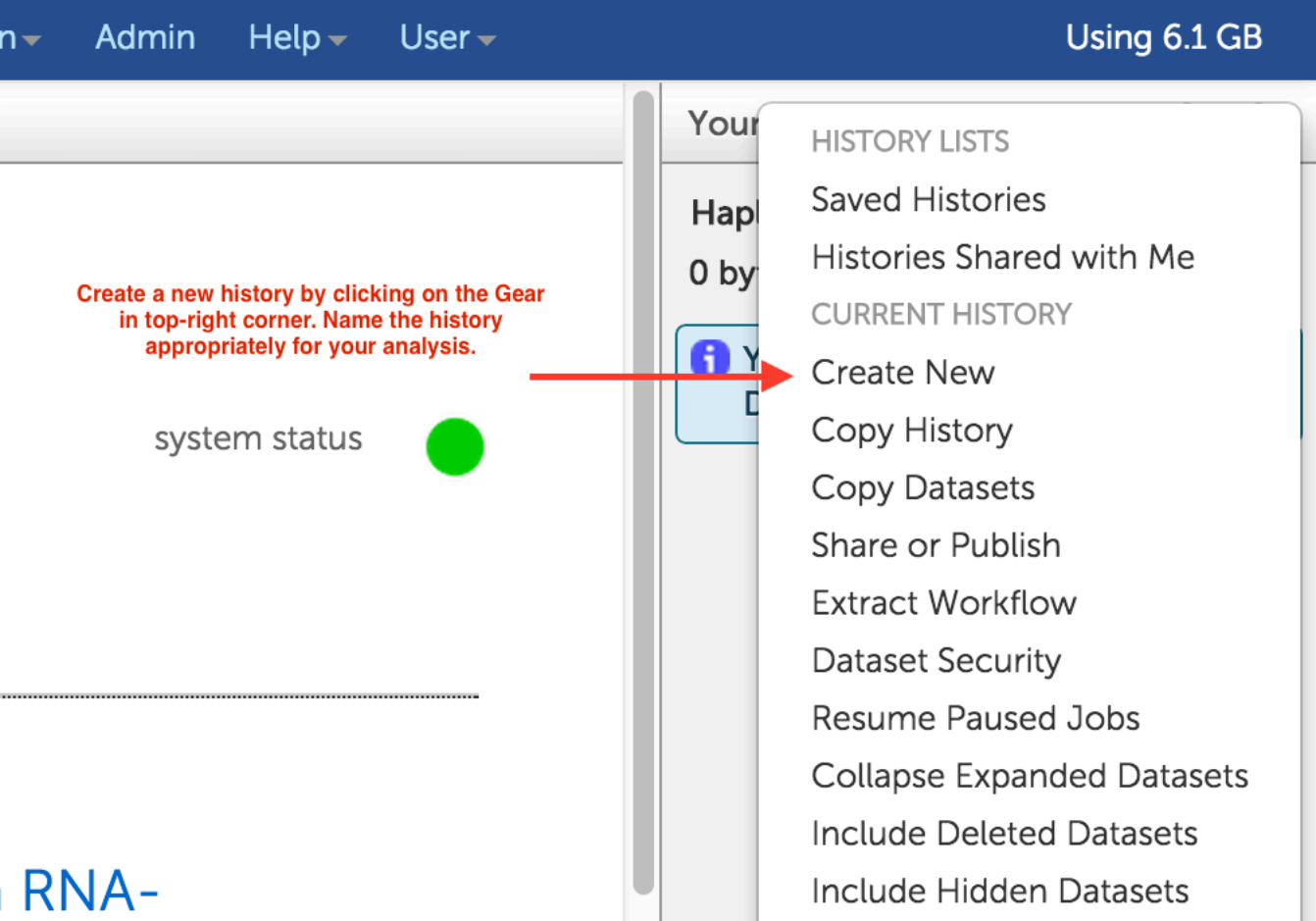
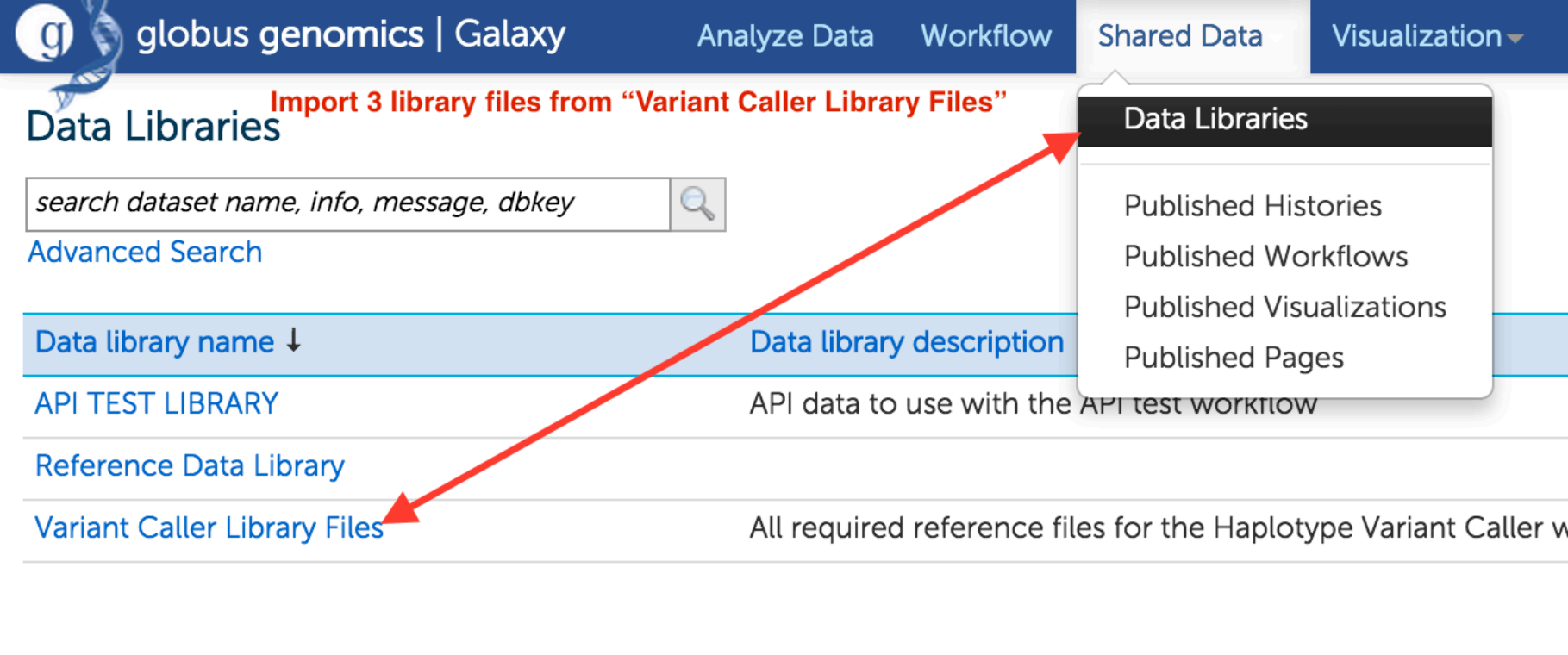
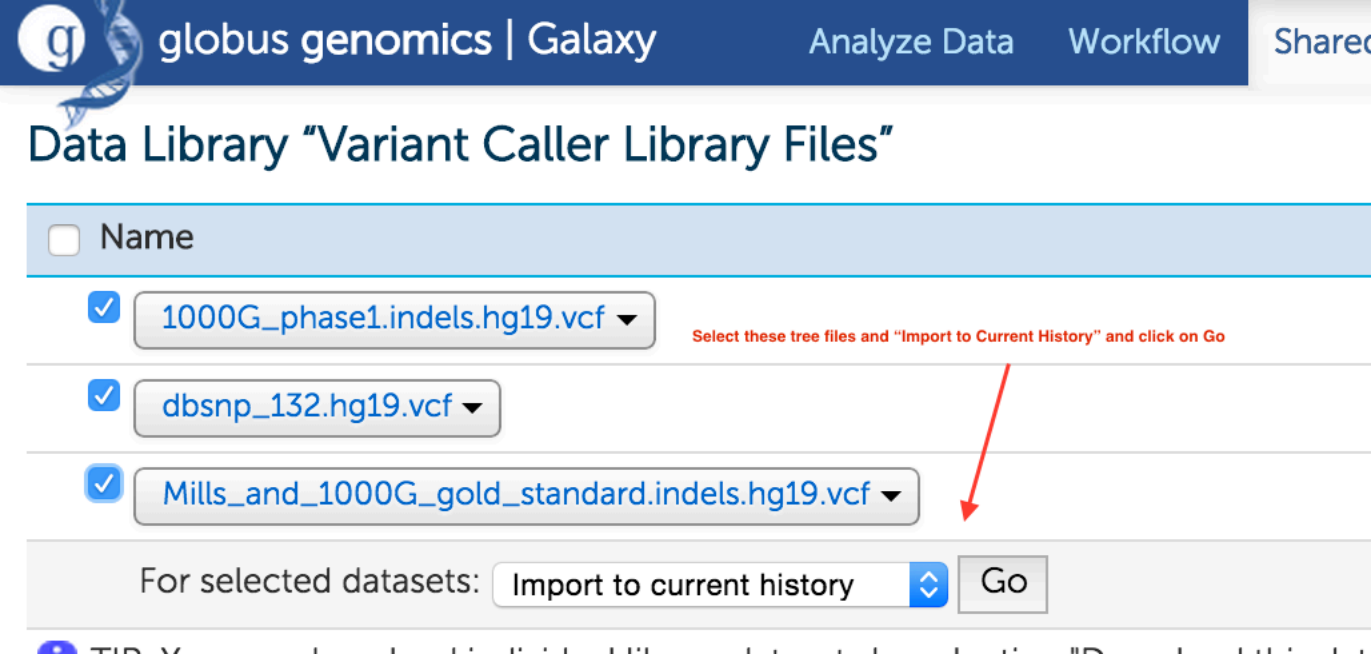
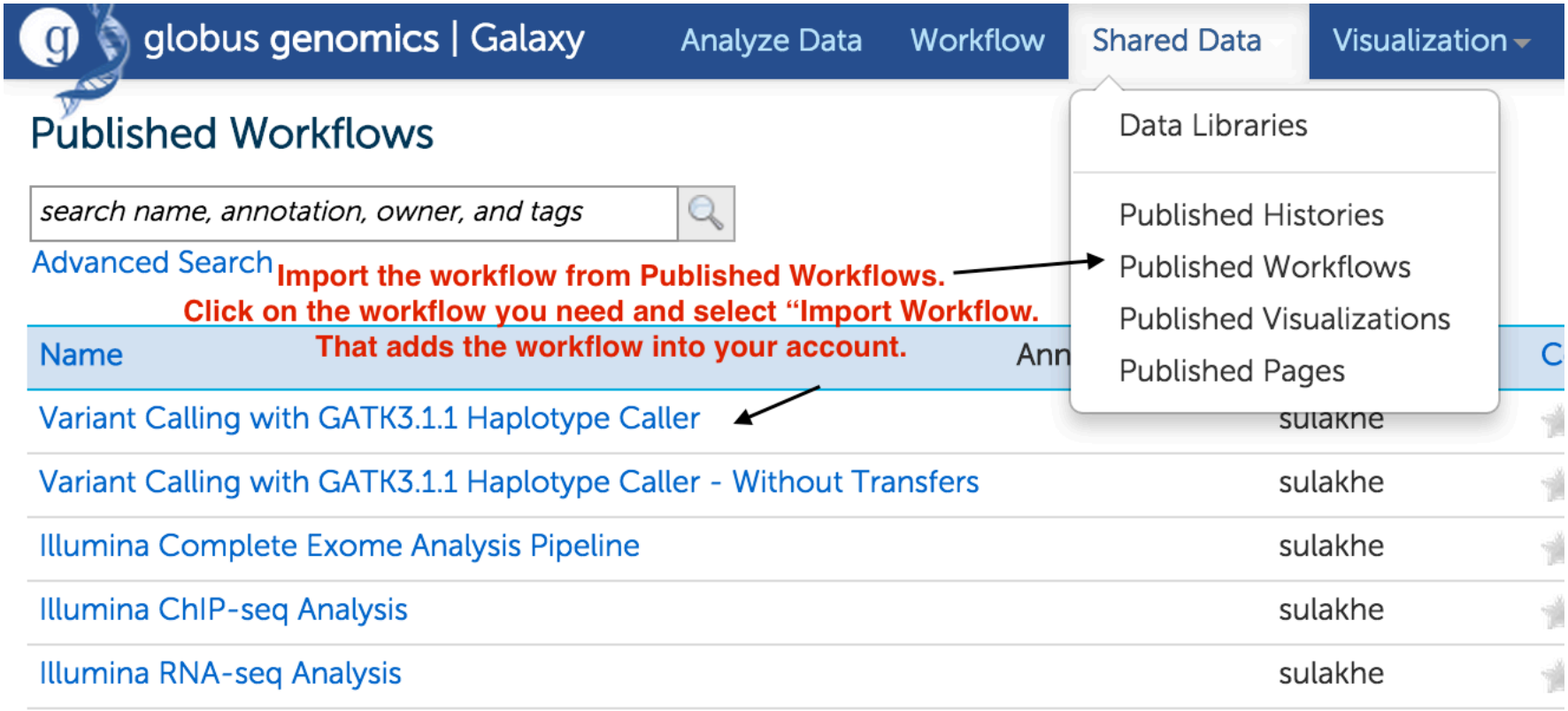
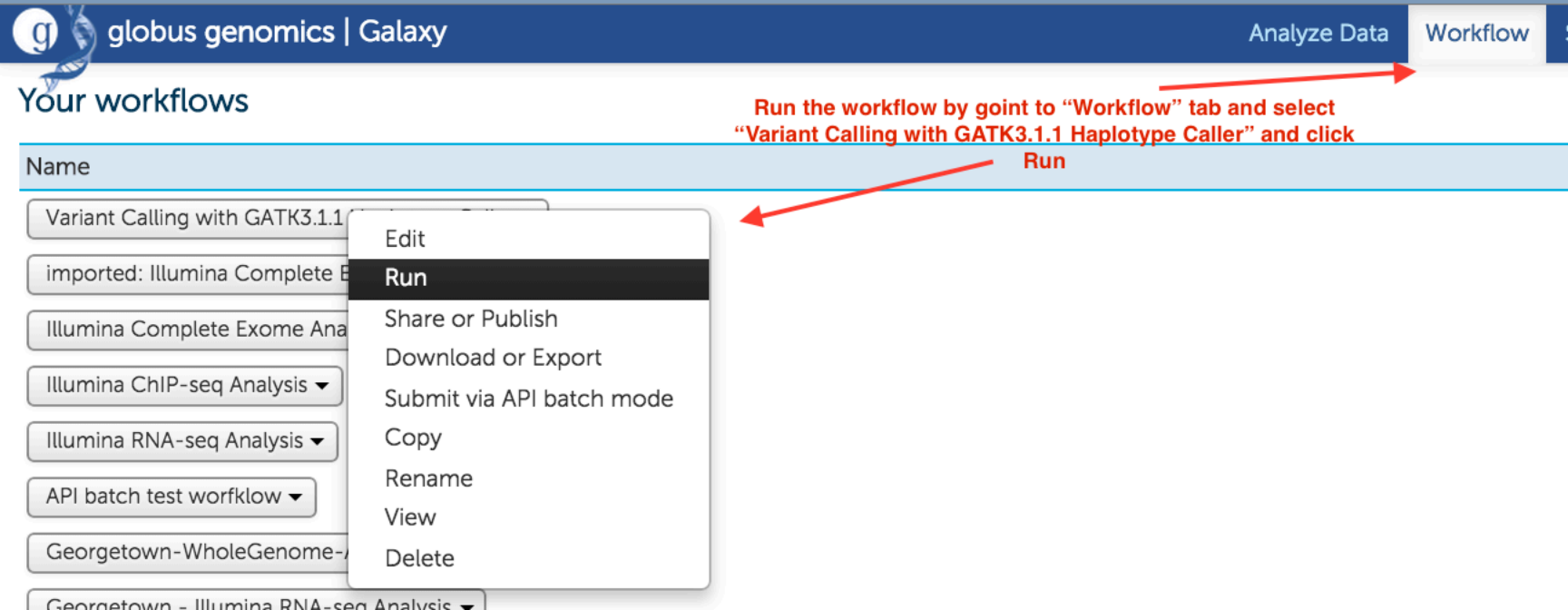

Admin
Help
User
Using 6.1 GB
Create a new history by clicking on the Gear in top-right corner. Name the history appropriately for your analysis.
system status
HISTORY LISTS
Saved Histories
Histories Shared with Me
CURRENT HISTORY
Create New
Copy History
Copy Datasets
Share or Publish
Extract Workflow
Dataset Security
Resume Paused Jobs
Collapse Expanded Datasets
Include Deleted Datasets
Include Hidden Datasets
RNA-


globus genomics | Galaxy
Analyze Data
Workflow
Shared Data
Visualization
Import 3 library files from "Variant Caller Library Files"
Data Libraries
search dataset name, info, message, dbkey
Advanced Search
Data Libraries
Published Histories
Published Workflows
Published Visualizations
Published Pages
Data library name ↓
Data library description
API TEST LIBRARY
API data to use with the API test workflow
Reference Data Library
Variant Caller Library Files
All required reference files for the Haplotype Variant Caller w


globus genomics | Galaxy
Analyze Data
Workflow
Shared
Data Library "Variant Caller Library Files"
Name
1000G_phase1.indels.hg19.vcf
dbsnp_132.hg19.vcf
Mills_and_1000G_gold_standard.indels.hg19.vcf
Select these tree files and "Import to Current History" and click on Go
For selected datasets:
Import to current history
Go


globus genomics | Galaxy
Analyze Data
Workflow
Shared Data
Visualization
Data Libraries
Published Histories
Published Workflows
Published Visualizations
Published Pages
Published Workflows
search name, annotation, owner, and tags
Advanced Search
Import the workflow from Published Workflows.
Click on the workflow you need and select "Import Workflow.
That adds the workflow into your account.
Name
Ann
C
Variant Calling with GATK3.1.1 Haplotype Caller
sulakhe
Variant Calling with GATK3.1.1 Haplotype Caller - Without Transfers
sulakhe
Illumina Complete Exome Analysis Pipeline
sulakhe
Illumina ChIP-seq Analysis
sulakhe
Illumina RNA-seq Analysis
sulakhe


globus genomics | Galaxy
Analyze Data
Workflow
Your workflows
Run the workflow by goint to “Workflow” tab and select “Variant Calling with GATK3.1.1 Haplotype Caller” and click Run
Name
Variant Calling with GATK3.1.1
imported: Illumina Complete E
Illumina Complete Exome Ana
Illumina ChIP-seq Analysis
Illumina RNA-seq Analysis
API batch test worfklow
Georgetown-WholeGenome-
Georgetown - Illumina RNA-seq Analysis
Edit
Run
Share or Publish
Download or Export
Submit via API batch mode
Copy
Rename
View
Delete

Tools

search tools

MATERIAL SCIENCE

DATA TRANSFER
Globus Data Transfer
Get Data
Send Data

NGS APPLICATIONS
NGS: QC and manipulation
NGS: Mapping
NGS: Mapping QC
NGS: RNA Analysis
NGS: Peak Calling
NGS: SAM Tools
NGS: BAM Tools
NGS: Picard
NGS: Indel Analysis
NGS: GATK Tools
NGS: GATK2 Tools
NGS: GATK3 Tools
NGS: Variant Detection
NGS: Interval Tools
NGS: VCF Tools
NGS: CGA Tools
NGS: Simulation
SNP/WGA: Data; Filters
SNP/WGA: QC; LD; Plots
SNP/WGA: Statistical Models
Phenotype Association

DATA MANIPULATION
ENCODE Tools
Lift-Over
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Fetch Alignments
Get Genomic Scores
Operate on Genomic Intervals

Running workflow "Variant Calling with GATK3.1.1 Haplotype Caller"

Expand All | Collapse

Step 1: Get Data via Globus Online (version 1.0.0)

Provide the Globus Endpoint name and Path to the input file on the endpoint.

None

None

None

None

None

Source Endpoint

sulakhe#testS3dobynsbucket

Source Path

/newfolder/Exome-Sample_Forward_1.fastqsanger

None

None

Deadline (In minutes)

120

Provide your S3 endpoint and path to the input file in Step 1 and Step 2. (example shown)

Actions:

Hide this dataset.
Rename output 'out_file1' to 'Forward FASTQ'.
Set the datatype of output 'out_file1' to 'fastqsanger'

Step 2: Get Data via Globus Online (version 1.0.0)

Provide the Globus Endpoint name and Path to the input file on the endpoint.

None

None

None

None

None

Source Endpoint

sulakhe#testS3dobynsbucket

Source Path

/newfolder/Exome-Sample_Reverse_2.fastqsanger

None

None

Your History

Haplotype-with-Transfer
4.7 GB

3: Mills_and_1000G_gold_standard.indels.hg19.vcf

2: dbsnp_132.hg19.vcf

1: 1000G_phase1.indels.hg19.vcf

Tools

search tools

MATERIAL SCIENCE

DATA TRANSFER

Globus Data Transfer
Get Data
Send Data

NGS APPLICATIONS

NGS: QC and manipulation
NGS: Mapping
NGS: Mapping QC
NGS: RNA Analysis
NGS: Peak Calling
NGS: SAM Tools
NGS: BAM Tools
NGS: Picard
NGS: Indel Analysis
NGS: GATK Tools
NGS: GATK2 Tools
NGS: GATK3 Tools

Set the datatype of output out_file1 to fastqsanger

Step 3: Input dataset

Mills Indels

3: Mills_and_1000G_g..ls.hg19.vcf

type to filter

Step 4: Input dataset

1000G indels

1: 1000G_phase1.indels.hg19.vcf

type to filter

Select appropriate Indels and dbSnp files for Step 3, 4, 5 from the history

Step 5: Input dataset

dbSNP

2: dbsnp_132.hg19.vcf

type to filter

Step 6: Map with BWA-MEM (version 0.7.10)

Your History

Haplotype-with-Transfer
4.7 GB

3: Mills_and_1000G_gold_standard.indels.hg19.vcf

2: dbsnp_132.hg19.vcf

1: 1000G_phase1.indels.hg19.vcf

Step 8: Add or Replace Groups (version 1.56.0)

SAM/BAM dataset to add or replace read groups in

Output dataset 'output' from step 7

Read group ID (ID tag)

Test

Read group sample name (SM tag)

Test

Read group library (LB tag)

Test

Read group platform (PL tag)

Illumina

Read group platform unit

100

Specify additional (optional) arguments

Use pre-set defaults

Output bam instead of sam

True

Provide Add or Replace Groups parameters. If you don't know the correct params, use "Test" for now as show here.

Step 9: Mark Duplicate reads (version 1.56.0)

Tools

search tools

MATERIAL SCIENCE

DATA TRANSFER
Globus Data Transfer
Get Data
Send Data

NGS APPLICATIONS
NGS: QC and manipulation
NGS: Mapping
NGS: Mapping QC
NGS: RNA Analysis
NGS: Peak Calling
NGS: SAM Tools
NGS: BAM Tools
NGS: Picard
NGS: Indel Analysis
NGS: GATK Tools
NGS: GATK2 Tools
NGS: GATK3 Tools
NGS: Variant Detection
NGS: Interval Tools
NGS: VCF Tools
NGS: CGA Tools
NGS: Simulation
SNP/WGA: Data; Filters
SNP/WGA: QC; LD; Plots
SNP/WGA: Statistical Models
Phenotype Association

DATA MANIPULATION
ENCODE Tools
Lift-Over
Text Manipulation

Step 14: Haplotype Caller (version 3.1.1)

Covariates table recalibration file
Selection is Optional

Choose the source for the reference list
Locally cached

BAM file
Output dataset 'output_bam' from step 13

Using reference genome
hg19 (value not yet validated)

Basic or Advanced GATK options
Basic

Basic or Advanced Analysis options
Advanced

Mode for emitting experimental reference confidence scores
GVCF

activeRegionIn
Selection is Optional

activeRegionOut
False

Annotation Types
None

Additional annotations

Annotation Interfaces/Groups

Annotations to exclude
None

comp
Selection is Optional

contamination_fraction_to_filter
0.05

dbsnp
2: dbsnp_132.hg19.vcf

debug

Select dbSnp file in Step 14 (Haplotype Caller)