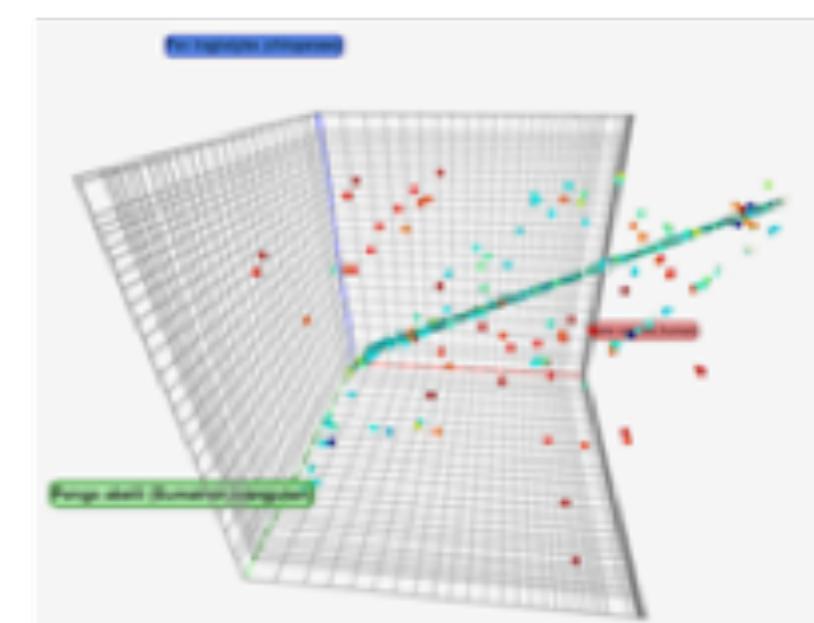
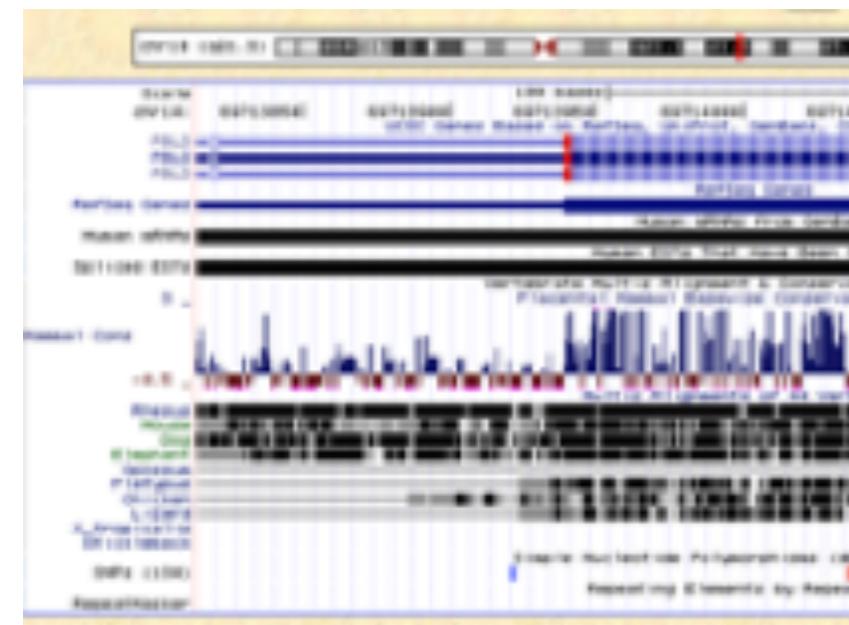


Computational Genomics

Introduction to **Text Manipulation(s)**



Introduction to Text Manipulation(s)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

search tools

Upload Data

HPRC

Get Data

Send Data

Collection Operations

Lift-Over

Text Manipulation

Convert Formats

Filter and Sort

Join, Subtract and Group

Fetch Alignments/Sequences

Operate on Genomic Intervals

Statistics

Graph/Display Data

Phenotype Association

BED

Annotation

Multiple Alignments

NCBI BLAST+

Mapping

SAM/BAM

Assembly

FASTQ Quality Control

FASTA/FASTQ

RNA-seq

CD-HIT

Datamash

EMBOSS

MUMmer4

Nanopore

Best Practices for Kaiser Galaxy

- Kaiser Galaxy (docs, slides) is configured for teaching purposes so all users have a file quota of 1TB. How to permanently delete nonessential files.
- Only certain tools that support multi-core processing have the Job Resources Parameters option which allow you to select cores, memory and time.
 - The default job resource parameters for all tools is 1 core with 7GB memory for 24 hours (24 SUs).
 - Configuring a job to use 48 cores for 1 hour requires 48 SUs. (48 cores for 168 hours = 7872 SUs).
 - Configuring a job to use 360GB memory for 1 hour requires 48 SUs. (360GB memory for 168 hours = 7872 SUs).
 - If a tool you used failed because it needs the Job Resource Parameters option added, contact the HPRC helpdesk.

History

search datasets

L05-A

2 shown

1.2 MB

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A



Take an interactive tour: Galaxy UI History Window Manager Deferred Datasets

Galaxy is an open platform for supporting data intensive research. Galaxy is developed by The Galaxy Team with the support of many contributors.

The Galaxy Project is supported in part by NHGRI, NSF, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Johns Hopkins University.

Introduction to Text Manipulation(s)

**Table containing
12 fields (columns) and 15 records (rows or lines)**

Field 01
Record 01

Field 01
Record 08

Field 12
Record 01
Head -2

Tail -4

1	2	3	4	5	6	7	8	9	10	11	12
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											

Introduction to Text Manipulation(s)

Select first lines from a dataset (head)

What it does: This tool outputs specified number of lines from the beginning of a dataset

Example: Selecting 2 lines from this:

chr7	56632	56652	D17003_CTCF_R6	310	+
chr7	56736	56756	D17003_CTCF_R7	354	+
chr7	56761	56781	D17003_CTCF_R4	220	+
chr7	56772	56792	D17003_CTCF_R7	372	+
chr7	56775	56795	D17003_CTCF_R4	207	+

will produce:

chr7	56632	56652	D17003_CTCF_R6	310	+
chr7	56736	56756	D17003_CTCF_R7	354	+

Introduction to Text Manipulation(s)

Select first lines from a dataset (head)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

select first

Select first lines from a dataset
Select random lines from a file
Select last lines from a dataset
Change Case of selected columns
Change Case of selected columns
CollectHsMetrics compute metrics about datasets generated through hybrid-selection (e.g. exome)
skipseq Reads and writes sequences, skipping first few
Extract dataset
Count occurrences of each record
Filter data on any column using simple expressions
Extract features from GFF data
bedtools FlankBed create new intervals from the flanks of existing intervals
CollectRnaSeqMetrics collect metrics about the alignment of RNA to various functional classes of loci in the genome
MAF to BED Converts a MAF formatted file to the BED format
Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome
Filter SAM on bitwise flag values
Upload File from your computer
Column Regex Find And Replace
NCBI BLAST+ blastn Search nucleotide database with nucleotide query sequence(s)
NCBI BLAST+ blastp Search protein database with protein query sequence(s)

WORKFLOWS
All workflows

Select first lines from a dataset (Galaxy Version 1.0.1)

Select first

10 lines from 1: Hs_Chr21_Table_A

Dataset has a header No
if checked the header is kept and is not counted as a kept line

Job Resource Parameters

Specify job resource parameters

Memory (GB) 7 Maximum Job Memory

Time (hours) 24 Maximum job time

Execute

What it does

This tool outputs specified number of lines from the beginning of a dataset

Example

Selecting 2 lines from this:

```
chr7 56632 56652 D17003_CTCF_R6 310 +
chr7 56736 56756 D17003_CTCF_R7 354 +
chr7 56761 56781 D17003_CTCF_R4 220 +
chr7 56772 56792 D17003_CTCF_R7 372 +
chr7 56775 56795 D17003_CTCF_R4 207 +
```

will produce:

```
chr7 56632 56652 D17003_CTCF_R6 310 +
chr7 56736 56756 D17003_CTCF_R7 354 +
```

Requirements: ?
- coreutils (Version 8.31)

History

search datasets

L05-A
2 shown
1.2 MB

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Select first lines from a dataset (head)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

select first x

Upload Data

Show Sections

Select first lines from a dataset

Select random lines from a file

Select last lines from a dataset

Change Case of selected columns

Change Case of selected columns

CollectHsMetrics compute metrics about datasets generated through hybrid-selection (e.g. exome)

skipseq Reads and writes sequences, skipping first few

Extract dataset

Count occurrences of each record

Filter data on any column using simple expressions

Extract features from GFF data

bedtools FlankBed create new intervals from the flanks of existing intervals

CollectRnaSeqMetrics collect metrics about the alignment of RNA to various functional classes of loci in the genome

MAF to BED Converts a MAF formatted file to the BED format

Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome

Filter SAM on bitwise flag values

Upload File from your computer

Column Regex Find And Replace

NCBI BLAST+ blastn Search nucleotide database with nucleotide query sequence(s)

NCBI BLAST+ blastp Search protein database with protein query sequence(s)

WORKFLOWS

All workflows

Chrom Start End Name Score Strand ThickStart ThickEnd ItemRGB BlockCount BlockSizes

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB	BlockCount	BlockSizes
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+					
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+					
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+					
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+					
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+					
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+					
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+					
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+					
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+					
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+					

History

search datasets x

L05-A

3 shown
1.2 MB

3: Select first on data 1 x

10 regions
format: bed, database: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom 2. Start 3. End 4

1. Chrom	2. Start	3. End	4
chr21	10538723	10538734	ENST00000622113
chr21	10541111	10541165	ENST00000622113
chr21	10542394	10542448	ENST00000622113
chr21	10552656	10552716	ENST00000622113
chr21	10559493	10559544	ENST00000622113

2: Hs_Chr21_Table_B x

1: Hs_Chr21_Table_A x

8,803 regions
format: bed, database: hg38

python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom 2. Start 3. End 4

1. Chrom	2. Start	3. End	4
chr21	10538723	10538734	ENST00000622113
chr21	10541111	10541165	ENST00000622113
chr21	10542394	10542448	ENST00000622113
chr21	10552656	10552716	ENST00000622113
chr21	10559493	10559544	ENST00000622113

Introduction to Text Manipulation(s)

Select last lines from a dataset (tail)

What it does: This tool outputs specified number of lines from the end of a dataset

Example: Selecting 2 lines from this:

chr7	57134	57154	D17003_CTCF_R7	356	-
chr7	57247	57267	D17003_CTCF_R4	207	+
chr7	57314	57334	D17003_CTCF_R5	269	+
chr7	57341	57361	D17003_CTCF_R7	375	+
chr7	57457	57477	D17003_CTCF_R3	188	+

will produce:

chr7	57341	57361	D17003_CTCF_R7	375	+
chr7	57457	57477	D17003_CTCF_R3	188	+

Introduction to Text Manipulation(s)

Select last lines from a dataset (tail)

HPRC Kaiser Galaxy Using 0%

Tools

- select last
- Upload Data
- Show Sections

Select last lines from a dataset

Select random lines from a file

Select first lines from a dataset

Change Case of selected columns

Change Case of selected columns

CollectHsMetrics compute metrics about datasets generated through hybrid-selection (e.g. exome)

NCBI BLAST+ blastn Search nucleotide database with nucleotide query sequence(s)

NCBI BLAST+ blastp Search protein database with protein query sequence(s)

NCBI BLAST+ tblastn Search translated nucleotide database with protein query sequence(s)

NCBI BLAST+ blastx Search protein database with translated nucleotide query sequence(s)

NCBI BLAST+ tblastx Search translated nucleotide database with translated nucleotide query sequence(s)

NCBI BLAST+ rpstblastn Search protein domain database (PSSMs) with translated nucleotide query sequence(s)

NCBI BLAST+ rpsblast Search protein domain database (PSSMs) with protein query sequence(s)

My HPRC SU balance

NCBI BLAST+ convert2blastmask Convert masking information in lower-case masked FASTA input to file formats suitable for makeblastdb

NCBI BLAST+ blastdbcmd entry(s) Extract sequence(s) from BLAST database

Workflow Visualize Shared Data Admin Help User

Select last lines from a dataset (Galaxy Version 1.0.1)

Select last

10 lines from 1: Hs_Chromosome_A

Dataset has a header: No

Job Resource Parameters

Specify job resource parameters

Memory (GB): 7

Time (hours): 24

Execute

What it does

This tool outputs specified number of lines from the end of a dataset

Example

- Input File:

chr7	57134	57154	D17003_CTCF_R7	356	-
chr7	57247	57267	D17003_CTCF_R4	207	+
chr7	57314	57334	D17003_CTCF_R5	269	+
chr7	57341	57361	D17003_CTCF_R7	375	+
chr7	57457	57477	D17003_CTCF_R3	188	+
- Show last two lines of above file. The result is:

chr7	57341	57361	D17003_CTCF_R7	375	+
chr7	57457	57477	D17003_CTCF_R3	188	+

Requirements: coreutils (Version 8.31)

History

search datasets

L05-A

3 shown
1.2 MB

3: Select first on data 1

2: Hs_Chromosome_A

1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Select last lines from a dataset (tail)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User Using 0%

Tools

select last x

Upload Data

Show Sections

Select last lines from a dataset

Select random lines from a file

Select first lines from a dataset

Change Case of selected columns

Change Case of selected columns

CollectHsMetrics compute metrics about datasets generated through hybrid-selection (e.g. exome)

NCBI BLAST+ blastn Search nucleotide database with nucleotide query sequence(s)

NCBI BLAST+ blastp Search protein database with protein query sequence(s)

NCBI BLAST+ tblastn Search translated nucleotide database with protein query sequence(s)

NCBI BLAST+ blastx Search protein database with translated nucleotide query sequence(s)

NCBI BLAST+ tblastx Search translated nucleotide database with translated nucleotide query sequence(s)

NCBI BLAST+ rpstblastn Search protein domain database (PSSMs) with translated nucleotide query sequence(s)

NCBI BLAST+ rpsblast Search protein domain database (PSSMs) with protein query sequence(s)

My HPRC SU balance

NCBI BLAST+ convert2blastmask Convert masking information in lower-case masked FASTA input to file formats suitable for makeblastdb

NCBI BLAST+ blastdbcmd entry(s) Extract sequence(s) from BLAST database

Chrom Start End Name Score Strand ThickStart ThickEnd ItemRGB BlockCount BlockSizes History

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB	BlockCount	BlockSizes	History
chr21	46644305	46644488	ENST00000440086.5_cds_3_0_chr21_46644306_f	0	+						
chr21	46648457	46648619	ENST00000440086.5_cds_4_0_chr21_46648458_f	0	+						
chr21	46649574	46649739	ENST00000440086.5_cds_5_0_chr21_46649575_f	0	+						
chr21	46661799	46661936	ENST00000440086.5_cds_6_0_chr21_46661800_f	0	+						
chr21	46663382	46663554	ENST00000440086.5_cds_7_0_chr21_46663383_f	0	+						
chr21	46664294	46664327	ENST00000440086.5_cds_8_0_chr21_46664295_f	0	+						
chr21	46644342	46644488	ENST00000455177.3_cds_0_0_chr21_46644343_f	0	+						
chr21	46648457	46648619	ENST00000455177.3_cds_1_0_chr21_46648458_f	0	+						
chr21	46649574	46649739	ENST00000455177.3_cds_2_0_chr21_46649575_f	0	+						
chr21	46652844	46652952	ENST00000455177.3_cds_3_0_chr21_46652845_f	0	+						

search datasets x

L05-A
4 shown
1.2 MB checkbox comment chat

4: Select last on data 1 eye pencil x
10 regions
format: bed, database: hg38
display in IGB View display with IGV local Human hg38 display at UCSC main test

1. Chrom 2. Start 3. End 4

1. Chrom	2. Start	3. End	4
chr21	46644305	46644488	ENST00000440086
chr21	46648457	46648619	ENST00000440086
chr21	46649574	46649739	ENST00000440086
chr21	46661799	46661936	ENST00000440086
chr21	46663382	46663554	ENST00000440086

3: Select first on data 1 eye pencil x

2: Hs_Chromosome_B eye pencil x

1: Hs_Chromosome_A eye pencil x

8,803 regions
format: bed, database: hg38
python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View display with IGV local Human hg38 display at UCSC main test

1. Chrom 2. Start 3. End 4

1. Chrom	2. Start	3. End	4
chr21	10538723	10538734	ENST00000622113..
chr21	10541111	10541165	ENST00000622113..
chr21	10542394	10542448	ENST00000622113..
chr21	10552656	10552716	ENST00000622113..
chr21	10559493	10559544	ENST00000622113..

Introduction to Text Manipulation(s)

Remove beginning of a file

What it does: This tool removes a specified number of lines from the beginning of a dataset

Example:

Input File:

chr7	56632	56652	D17003_CTCF_R6	310	+
chr7	56736	56756	D17003_CTCF_R7	354	+
chr7	56761	56781	D17003_CTCF_R4	220	+
chr7	56772	56792	D17003_CTCF_R7	372	+
chr7	56775	56795	D17003_CTCF_R4	207	+

After removing the first 3 lines the dataset will look like this:

chr7	56772	56792	D17003_CTCF_R7	372	+
chr7	56775	56795	D17003_CTCF_R4	207	+

Introduction to Text Manipulation(s)

Remove beginning of a file

HPRC Kaiser Galaxy Workflow Visualize Shared Data Admin Help User Using 0%

Tools

remove

Remove beginning of a file
Remove, rearrange and/or rename columns in txt-converted FCS files
bedtools SubtractBed remove intervals based on overlaps
Remove sequencing artifacts
Cutadapt Remove adapter sequences from FASTQ/FASTA
unique_line remove duplicate lines
cutseq Removes a specified section from a sequence
degapseq Removes gap characters from sequences
noreturn Removes carriage return from ASCII files
Filter MAF blocks by Species
Filter empty datasets
Filter collection
Filter failed datasets
Trim leading or trailing characters
Filter MAF by specified attributes
RevertSam revert SAM/BAM datasets to a previous state
Subtract the intervals of two datasets
Join MAF blocks by Species
CD-HIT PROTEIN Cluster a protein dataset into representative sequences
CD-HIT-EST Cluster a nucleotide dataset into representative sequences
WORKFLOWS
All workflows

Remove beginning of a file (Galaxy Version 1.0.0)

Remove first
10 lines from 1: Hs_Chromosome_A

Job Resource Parameters
Specify job resource parameters
Memory (GB) 7
Maximum Job Memory
Time (hours) 24
Maximum job time

Execute

What it does
This tool removes a specified number of lines from the beginning of a dataset.

Example
Input File:

```
chr7 56632 56652 D17003_CTCF_R6 310 +
chr7 56736 56756 D17003_CTCF_R7 354 +
chr7 56761 56781 D17003_CTCF_R4 220 +
chr7 56772 56792 D17003_CTCF_R7 372 +
chr7 56775 56795 D17003_CTCF_R4 207 +
```

After removing the first 3 lines the dataset will look like this:

```
chr7 56772 56792 D17003_CTCF_R7 372 +
chr7 56775 56795 D17003_CTCF_R4 207 +
```

History

search datasets

L05-A
4 shown
1.2 MB

4: Select last on data 1
3: Select first on data 1
2: Hs_Chromosome_B
1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Remove beginning of a file

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

- remove x
- Upload Data**
- Show Sections**
- Remove beginning of a file**
- Remove, rearrange and/or rename columns in txt-converted FCS files**
- bedtools SubtractBed** remove intervals based on overlaps
- Remove sequencing artifacts**
- Cutadapt** Remove adapter sequences from FASTQ/FASTA
- unique_line** remove duplicate lines
- cutseq** Removes a specified section from a sequence
- degapseq** Removes gap characters from sequences
- noreturn** Removes carriage return from ASCII files
- Filter MAF blocks by Species**
- Filter empty datasets**
- Filter collection**
- Filter failed datasets**
- Trim leading or trailing characters**
- Filter MAF by specified attributes**
- RevertSam** revert SAM/BAM datasets to a previous state
- Subtract** the intervals of two datasets
- Join MAF blocks by Species**
- CD-HIT PROTEIN** Cluster a protein dataset into representative sequences
- CD-HIT-EST** Cluster a nucleotide dataset into representative sequences
- WORKFLOWS**
- All workflows

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB	BlockCount	Blocks
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f	0	+					
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f	0	+					
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f	0	+					
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f	0	+					
chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f	0	+					
chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f	0	+					
chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f	0	+					
chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f	0	+					
chr21	10603561	10603632	ENST00000622113.4_cds_21_0_chr21_10603562_f	0	+					
chr21	10605416	10605552	ENST00000622113.4_cds_22_0_chr21_10605417_f	0	+					
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+					
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+					
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+					
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+					
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+					
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+					
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+					
chr21	10538723	10538734	ENST0000427445.6_cds_3_0_chr21_10538724_f	0	+					
chr21	10541111	10541165	ENST0000427445.6_cds_4_0_chr21_10541112_f	0	+					
chr21	10542394	10542448	ENST0000427445.6_cds_5_0_chr21_10542395_f	0	+					
chr21	10559493	10559544	ENST0000427445.6_cds_6_0_chr21_10559494_f	0	+					
chr21	10561029	10561191	ENST0000427445.6_cds_7_0_chr21_10561030_f	0	+					
chr21	10567669	10567789	ENST0000427445.6_cds_8_0_chr21_10567670_f	0	+					
chr21	10569436	10569536	ENST0000427445.6_cds_9_0_chr21_10569437_f	0	+					
chr21	10569682	10569746	ENST0000427445.6_cds_10_0_chr21_10569683_f	0	+					
chr21	10570484	10570549	ENST0000427445.6_cds_11_0_chr21_10570485_f	0	+					
chr21	10577459	10577520	ENST0000427445.6_cds_12_0_chr21_10577460_f	0	+					
chr21	10578336	10578418	ENST0000427445.6_cds_13_0_chr21_10578337_f	0	+					
chr21	10578516	10578605	ENST0000427445.6_cds_14_0_chr21_10578517_f	0	+					
chr21	10590461	10590523	ENST0000427445.6_cds_15_0_chr21_10590462_f	0	+					
chr21	10592292	10592373	ENST0000427445.6_cds_16_0_chr21_10592293_f	0	+					
chr21	10595981	10596087	ENST0000427445.6_cds_17_0_chr21_10595982_f	0	+					
chr21	10598014	10598094	ENST0000427445.6_cds_18_0_chr21_10598015_f	0	+					
chr21	10602057	10602150	ENST0000427445.6_cds_19_0_chr21_10602058_f	0	+					
chr21	10603561	10603632	ENST0000427445.6_cds_20_0_chr21_10603562_f	0	+					
chr21	10605416	10605552	ENST0000427445.6_cds_21_0_chr21_10605417_f	0	+					
chr21	10561159	10561191	ENST0000612746.1_cds_4_0_chr21_10561160_f	0	+					
chr21	10567669	10567789	ENST0000612746.1_cds_5_0_chr21_10567670_f	0	+					
chr21	10569436	10569536	ENST0000612746.1_cds_6_0_chr21_10569437_f	0	+					
chr21	10569682	10569746	ENST0000612746.1_cds_7_0_chr21_10569683_f	0	+					
chr21	10570484	10570549	ENST0000612746.1_cds_8_0_chr21_10570485_f	0	+					
chr21	10577459	10577520	ENST0000612746.1_cds_9_0_chr21_10577460_f	0	+					
chr21	10578336	10578418	ENST0000612746.1_cds_10_0_chr21_10578337_f	0	+					
chr21	10578516	10578605	ENST0000612746.1_cds_11_0_chr21_10578517_f	0	+					

History

search datasets

L05-A

5 shown
1.8 MB

5: Remove beginning o n data 1

8,793 regions
format: bed, database: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom	2. Start	3. End	4
chr21	10577459	10577520	ENST0000062211
chr21	10578336	10578418	ENST0000062211
chr21	10578516	10578605	ENST0000062211
chr21	10590461	10590523	ENST0000062211
chr21	10592292	10592373	ENST0000062211

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

8,803 regions
format: bed, database: hg38

python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom	2. Start	3. End	4
chr21	10538723	10538734	ENST0000062211
chr21	10541111	10541165	ENST0000062211
chr21	10542394	10542448	ENST0000062211
chr21	10552656	10552716	ENST0000062211
chr21	10559493	10559544	ENST0000062211

Introduction to Text Manipulation(s)

Concatenate datasets tail-to-head (cat)

What it does: Concatenates datasets

Example

Concatenating Dataset:

chrX	151087187	151087355	A	0	-
chrX	151572400	151572481	B	0	+

with Dataset1:

chr1	151242630	151242955	X	0	+
chr1	151271715	151271999	Y	0	+
chr1	151278832	151279227	Z	0	-

and with Dataset2:

chr2	100000030	200000955	P	0	+
chr2	100000015	200000999	Q	0	+

will result in the following:

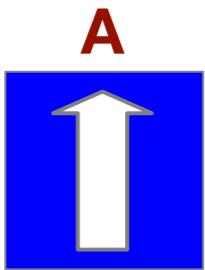
chrX	151087187	151087355	A	0	-
chrX	151572400	151572481	B	0	+
chr1	151242630	151242955	X	0	+
chr1	151271715	151271999	Y	0	+
chr1	151278832	151279227	Z	0	-
chr2	100000030	200000955	P	0	+
chr2	100000015	200000999	Q	0	+

Introduction to Text Manipulation(s)

Concatenate datasets tail-to-head (cat)

Given three Datasets A, B, and C

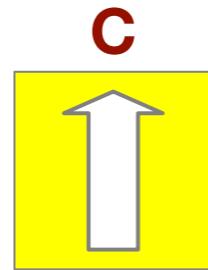
Dataset



Dataset



Dataset

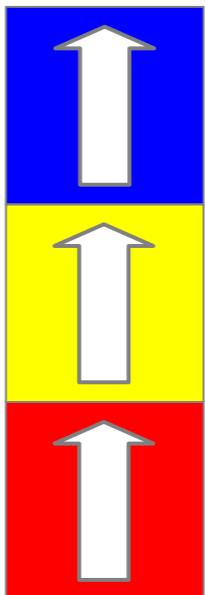


If Concatenated as ... Result in ...

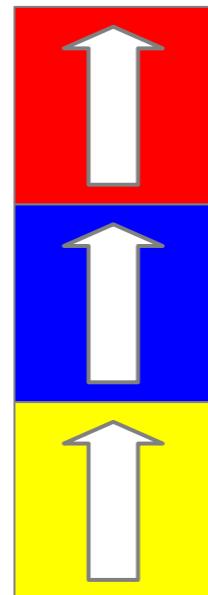
$A + B + C$



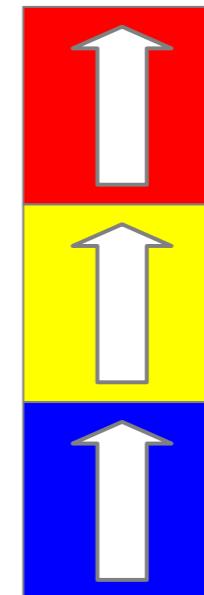
$A + C + B$



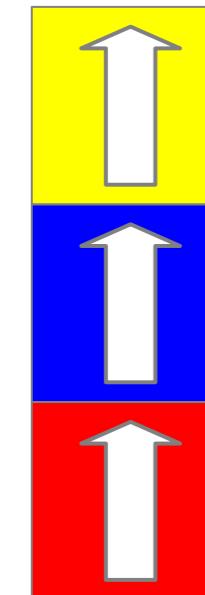
$B + A + C$



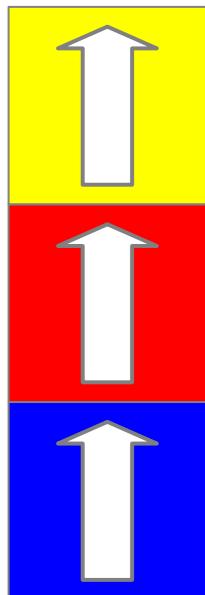
$B + C + A$



$C + A + B$



$C + B + A$



Using CAT

Introduction to Text Manipulation(s)

Concatenate datasets tail-to-head (cat)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User Using 0%

Tools

concatenate

Concatenate two BED files
Concatenate two BED files
Concatenate datasets tail-to-head
Group data by a column and perform aggregate operation on other columns.
MAF to FASTA Converts a MAF formatted file to FASTA format
Cutadapt Remove adapter sequences from FASTQ/FASTA

WORKFLOWS

All workflows

Concatenate datasets tail-to-head (Galaxy Version 1.0.0)

Concatenate Dataset

1: Hs_Chromosome_A

Dataset

1: Dataset

Select

2: Hs_Chromosome_B

+ Insert Dataset

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

WARNING: Be careful not to concatenate datasets of different kinds (e.g., sequences with intervals). This tool does not check if the datasets being concatenated are in the same format.

What it does

Concatenates datasets

Example

Concatenating Dataset:

```
chrX 151087187 151087355 A 0 -
chrX 151572400 151572481 B 0 +
```

with Dataset1:

```
chr1 151242630 151242955 X 0 +
chr1 151271715 151271999 Y 0 +
chr1 151278832 151279227 Z 0 -
```

and with Dataset2:

```
chr2 100000030 200000955 P 0 +
chr2 100000015 200000999 Q 0 +
```

History

search datasets

L05-A

5 shown
1.8 MB

5: Remove beginning on data 1
4: Select last on data 1
3: Select first on data 1
2: Hs_Chromosome_B
1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Concatenate datasets tail-to-head (cat)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User Using 0%

Tools

- concatenate x
- Upload Data
- Show Sections
- Concatenate two BED files
- Concatenate two BED files
- Concatenate datasets tail-to-head
- Group data by a column and perform aggregate operation on other columns.
- MAF to FASTA Converts a MAF formatted file to FASTA format
- Cutadapt Remove adapter sequences from FASTQ/FASTA
- WORKFLOWS
- All workflows

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB	BlockCount	Blocks
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+					
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+					
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+					
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+					
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+					
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+					
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+					
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+					
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+					
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+					
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f	0	+					
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f	0	+					
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f	0	+					
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f	0	+					
chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f	0	+					
chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f	0	+					
chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f	0	+					
chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f	0	+					
chr21	10603561	10603632	ENST00000622113.4_cds_21_0_chr21_10603562_f	0	+					
chr21	10605416	10605552	ENST00000622113.4_cds_22_0_chr21_10605417_f	0	+					
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+					
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+					
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+					
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+					
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+					
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+					
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+					
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+					
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+					
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+					
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+					
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+					
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+					
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+					
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+					
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+					
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+					
chr21	10578336	10578418	ENST00000427445.6_cds_13_0_chr21_10578337_f	0	+					
chr21	10578516	10578605	ENST00000427445.6_cds_14_0_chr21_10578517_f	0	+					
chr21	10590461	10590523	ENST00000427445.6_cds_15_0_chr21_10590462_f	0	+					
chr21	10592292	10592373	ENST00000427445.6_cds_16_0_chr21_10592293_f	0	+					
chr21	10595981	10596087	ENST00000427445.6_cds_17_0_chr21_10595982_f	0	+					
chr21	10598014	10598094	ENST00000427445.6_cds_18_0_chr21_10598015_f	0	+					
chr21	10602057	10602150	ENST00000427445.6_cds_19_0_chr21_10602058_f	0	+					

History

search datasets ? x

L05-A
6 shown
3 MB ? x m

6: Concatenate datasets on data 2 and data 1 ? x

17,606 regions
format: **bed**, database: **hg38**
? x m
display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom	2. Start	3. End	4
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f
chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f
chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f
chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f
chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f

5: Remove beginning of data 1 ? x

4: Select last on data 1 ? x

3: Select first on data 1 ? x

2: Hs_Chromosome_B ? x

1: Hs_Chromosome_A ? x

8,803 regions
format: **bed**, database: **hg38**
python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory
? x m
display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom	2. Start	3. End	4
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f
chr21	105922		

Introduction to Text Manipulation(s)

Cut columns from a table (cut)

What it does: This tool selects (cuts out) specified columns from the dataset.

Columns are specified as c1, c2, and so on.

Column count begins with 1

Columns can be specified in any order (e.g., c2,c1,c6)

If you specify more columns than actually present – empty spaces will be filled with dots

Input Example: Input dataset (six columns: c1, c2, c3, c4, c5, and c6):

```
chr1 10  1000  gene1 0 +
chr2 100 1500  gene2 0 +
```

cut on columns "c1,c4,c6" will return:

```
chr1 gene1 +
chr2 gene2 +
```

cut on columns "c6,c5,c4,c1" will return:

```
+ 0 gene1 chr1
+ 0 gene2 chr2
```

cut on columns "c1-c3" will return:

```
chr1 10  1000
chr2 100 1500
```

cut on columns "c8,c7,c4" will return:

```
. . gene1
. . gene2
```

Introduction to Text Manipulation(s)

Cut columns from a table (cut)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User Using 0%

Tools

cut

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

All workflows

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns
c1,c2,c3,c4

Delimited by
Tab

From
1: Hs_Chromosome_Table_A

Job Resource Parameters

Specify job resource parameters

Memory (GB)
7

Maximum Job Memory

Time (hours)
24

Maximum job time

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:
apple,is,good
windows,is,bad

will give:
apple good
windows bad

What it does

This tool selects (cuts out) specified columns from the dataset.

- Columns are specified as **c1, c2**, and so on. Column count begins with **1**
- Columns can be specified in any order (e.g., **c2,c1,c6**)
- If you specify more columns than actually present - empty spaces will be filled with dots

Example

Input dataset (six columns: c1, c2, c3, c4, c5, and c6):

```
chr1 10 1000 gene1 0 +  
chr2 100 1500 gene2 0 +
```

History

search datasets

L05-A

6 shown
3 MB

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_Table_B

1: Hs_Chromosome_Table_A

8,803 regions
format: bed, database: hg38

python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1.Chrom	2.Start	3.End	4
chr21	10538723	10538734	ENST00000622113..
chr21	10541111	10541165	ENST00000622113..
chr21	10542394	10542448	ENST00000622113..
chr21	10552656	10552716	ENST00000622113..
chr21	10559493	10559544	ENST00000622113..

Introduction to Text Manipulation(s)

Cut columns from a table (cut)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

cut

Upload Data

Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

All workflows

	chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f
7: Cut on data 1	chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f
6: Concatenate datasets on data 2 and data 1	chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f
5: Remove beginning on data 1	chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f
4: Select last on data 1	chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f
3: Select first on data 1	chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f
2: Hs_Chr21_Table_B	chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f
1: Hs_Chr21_Table_A	chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f
python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory	chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f
display in IGB View	chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f
display with IGV local Human hg38	chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f
display at UCSC main test	chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f
5 6	chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f
00000622113.4_cds_3_0_chr21_10538724_f 0 +	chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f
00000622113.4_cds_4_0_chr21_10541112_f 0 +	chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f
00000622113.4_cds_5_0_chr21_10542395_f 0 +	chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f
00000622113.4_cds_6_0_chr21_10552657_f 0 +	chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f
00000622113.4_cds_7_0_chr21_10559494_f 0 +	chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f
00000622113.4_cds_8_0_chr21_10561030_f 0 +	chr21	10603561	10603632	ENST00000622113.4_cds_21_0_chr21_10603562_f
00000622113.4_cds_9_0_chr21_10567670_f 0 +	chr21	10605416	10605552	ENST00000622113.4_cds_22_0_chr21_10605417_f
00000622113.4_cds_10_0_chr21_10569683_f 0 +	chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f
00000622113.4_cds_11_0_chr21_10570485_f 0 +	chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f
00000622113.4_cds_12_0_chr21_10577460_f 0 +	chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f
00000622113.4_cds_13_0_chr21_10578337_f 0 +	chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f
00000622113.4_cds_14_0_chr21_10578517_f 0 +	chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f
00000622113.4_cds_15_0_chr21_10590462_f 0 +	chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f
00000622113.4_cds_16_0_chr21_10592293_f 0 +	chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f
00000622113.4_cds_17_0_chr21_10595982_f 0 +	chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f
00000622113.4_cds_18_0_chr21_10598015_f 0 +	chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f
00000622113.4_cds_19_0_chr21_10602058_f 0 +	chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f
00000622113.4_cds_20_0_chr21_10602058_f 0 +	chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f
00000622113.4_cds_21_0_chr21_10561030_f 0 +	chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f
00000622113.4_cds_22_0_chr21_10567670_f 0 +	chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f
00000622113.4_cds_23_0_chr21_10569437_f 0 +	chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f
00000622113.4_cds_24_0_chr21_10569683_f 0 +	chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f
00000622113.4_cds_25_0_chr21_10570485_f 0 +	chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f
00000622113.4_cds_26_0_chr21_10577460_f 0 +	chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f
00000622113.4_cds_27_0_chr21_10578337_f 0 +	chr21	10578336	10578418	ENST00000427445.6_cds_13_0_chr21_10578337_f
00000622113.4_cds_28_0_chr21_10578517_f 0 +	chr21	10578516	10578605	ENST00000427445.6_cds_14_0_chr21_10578517_f
00000622113.4_cds_29_0_chr21_10590462_f 0 +	chr21	10590461	10590523	ENST00000427445.6_cds_15_0_chr21_10590462_f
00000622113.4_cds_30_0_chr21_10592293_f 0 +	chr21	10592292	10592373	ENST00000427445.6_cds_16_0_chr21_10592293_f
00000622113.4_cds_31_0_chr21_10595982_f 0 +	chr21	10595981	10596087	ENST00000427445.6_cds_17_0_chr21_10595982_f
00000622113.4_cds_32_0_chr21_10598015_f 0 +	chr21	10598014	10598094	ENST00000427445.6_cds_18_0_chr21_10598015_f
00000622113.4_cds_33_0_chr21_10602058_f 0 +	chr21	10602057	10602150	ENST00000427445.6_cds_19_0_chr21_10602058_f
00000622113.4_cds_34_0_chr21_10603562_f 0 +	chr21	10603561	10603632	ENST00000427445.6_cds_20_0_chr21_10603562_f

History

search datasets

L05-A

7 shown

3.56 MB

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

8,803 regions

format: bed, database: hg38

python: error while loading shared libraries: libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View

display with IGV local Human hg38

display at UCSC main test

00000622113.4_cds_3_0_chr21_10538724_f 0 +

00000622113.4_cds_4_0_chr21_10541112_f 0 +

00000622113.4_cds_5_0_chr21_10542395_f 0 +

00000622113.4_cds_6_0_chr21_10552657_f 0 +

00000622113.4_cds_7_0_chr21_10559494_f 0 +

Introduction to Text Manipulation(s)

Cut columns from a table (`cut`)

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

cut x

Upload Data

Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/Fasta

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

All workflows

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns

c5,c6

Delimited by

Tab

From

1: Hs_Chromosome_A

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

i The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:

apple,is,good
windows,is,bad

will give:

apple good
windows bad

What it does

This tool selects (cuts out) specified columns from the dataset.

- Columns are specified as **c1, c2**, and so on. Column count begins with **1**
- Columns can be specified in any order (e.g., **c2,c1,c6**)
- If you specify more columns than actually present - empty spaces will be filled with dots

Example

Input dataset (six columns: c1, c2, c3, c4, c5, and c6):

chr1 10 1000 gene1 0 +
chr2 100 1500 gene2 0 +

History

search datasets x

L05-A

7 shown
3.56 MB

7: Cut on data 1

8,803 regions
format: interval, database: hg38

display with IGV local Human hg38
display at UCSC main test

1.Chrom 2.Start 3.End 4

chr21 10538723 10538734 ENST00000622113...
chr21 10541111 10541165 ENST00000622113...
chr21 10542394 10542448 ENST00000622113...
chr21 10552656 10552716 ENST00000622113...
chr21 10559493 10559544 ENST00000622113...

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_B

1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Cut columns from a table (cut)

The screenshot shows the HPRC Kaiser Galaxy web interface. The top navigation bar includes links for Workflow, Visualize, Shared Data, Admin, Help, User, and various system icons. A progress bar at the top right indicates "Using 0%".

Tools

- cut (selected)
- Upload Data
- Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

- All workflows

History

- L05-A (8 shown, 3.6 MB)
- 8: Cut on data 1 (8,803 lines, tabular format, hg38 database)
- 7: Cut on data 1 (8,803 regions, interval format, hg38 database)
- 6: Concatenate dataset (s on data 2 and data 1)
- 5: Remove beginning o n data 1
- 4: Select last on data 1
- 3: Select first on data 1
- 2: Hs_Chr21_Table_B
- 1: Hs_Chr21_Table_A (8,803 regions, bed format, hg38 database)

Introduction to Text Manipulation(s)

Paste two files side by side

What it does: This tool merges two datasets side by side

If the first (left) dataset contains column assignments such as chromosome, start, end and strand, these will be preserved

However, if you would like to change column assignments, click the pencil icon in the history item

Example:

First dataset:	Second dataset:
a 1	20
a 2	30
a 3	40

Pasting them together will produce:

Final dataset:
a 1 20
a 2 30
a 3 40

Introduction to Text Manipulation(s)

Paste two files side by side

HPRC Kaiser Galaxy Using 0%

Tools

paste

Upload Data

Show Sections

Paste two files side by side

pasteseq Insert one sequence into another

Extract Pairwise MAF blocks given a set of genomic intervals

Show-Coords Parse delta file and report coordinates and other information

MergeBamAlignment merge alignment data with additional info stored in an unmapped BAM dataset

AddOrReplaceReadGroups add or replaces read group information

Extract MAF blocks given a set of genomic intervals

FastqToSam convert Fastq data into unaligned BAM

Map with BWA - map short reads (< 100 bp) against reference genome

Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome

Bowtie2 - map reads against reference genome

WORKFLOWS

All workflows

Workflow Visualize Shared Data Admin Help User

Paste two files side by side (Galaxy Version 1.0.0)

Paste

7: Cut on data 1

and

8: Cut on data 1

Delimit by

Tab

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

Paste preserves column assignments of the first dataset.

What it does

This tool merges two datasets side by side. If the first (left) dataset contains column assignments such as chromosome, start, end and strand, these will be preserved. However, if you would like to change column assignments, click the pencil icon in the history item.

Example

First dataset:

```
a 1
a 2
a 3
```

Second dataset:

```
20
30
40
```

Pasting them together will produce:

```
a 1 20
a 2 30
a 3 40
```

History

search datasets

L05-A

8 shown

3.6 MB

8: Cut on data 1

8,803 lines

format: tabular, database: hg38

7: Cut on data 1

8,803 regions

format: interval, database: hg38

1: Chrom 2:Start 3:End 4

chr21	10538723	10538734	ENST00000622113..
chr21	10541111	10541165	ENST00000622113..
chr21	10542394	10542448	ENST00000622113..
chr21	10552656	10552716	ENST00000622113..
chr21	10559493	10559544	ENST00000622113..

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Paste two files side by side

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

paste x

Upload Data

Show Sections

Paste two files side by side

pasteseq Insert one sequence into another

Extract Pairwise MAF blocks given a set of genomic intervals

Show-Coords Parse delta file and report coordinates and other information

MergeBamAlignment merge alignment data with additional info stored in an unmapped BAM dataset

AddOrReplaceReadGroups add or replaces read group information

Extract MAF blocks given a set of genomic intervals

FastqToSam convert Fastq data into unaligned BAM

Map with BWA - map short reads (< 100 bp) against reference genome

Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome

Bowtie2 - map reads against reference genome

WORKFLOWS

All workflows

Chromosome	Start	End	Description	Score	Orientation
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f	0	+
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f	0	+
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f	0	+
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f	0	+
chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f	0	+
chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f	0	+
chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f	0	+
chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f	0	+
chr21	10603561	10603632	ENST00000622113.4_cds_21_0_chr21_10603562_f	0	+
chr21	10605416	10605552	ENST00000622113.4_cds_22_0_chr21_10605417_f	0	+
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543382_f	0	+
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552716_f	0	+
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+
chr21	10567669	10567789	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+
chr21	10578336	10578418	ENST00000427445.6_cds_13_0_chr21_10578337_f	0	+
chr21	10578516	10578605	ENST00000427445.6_cds_14_0_chr21_10578517_f	0	+
chr21	10590461	10590523	ENST00000427445.6_cds_15_0_chr21_10590462_f	0	+
chr21	10592292	10592373	ENST00000427445.6_cds_16_0_chr21_10592293_f	0	+
chr21	10595981	10596087	ENST00000427445.6_cds_17_0_chr21_10595982_f	0	+
chr21	10598014	10598094	ENST00000427445.6_cds_18_0_chr21_10598015_f	0	+
chr21	10602057	10602150	ENST00000427445.6_cds_19_0_chr21_10602058_f	0	+
chr21	10603561	10603632	ENST00000427445.6_cds_20_0_chr21_10603562_f	0	+

DANGEROUS OPERATION!!

History

search datasets

L05-A

9 shown
4.2 MB

9: Paste on data 8 and data 7

8,803 regions
format: interval, database: hg38

display with IGV local Human hg38
display at UCSC main test

1. Chrom 2. Start 3. End 4

chr21 10538723 10538734 ENST00000622113

chr21 10541111 10541165 ENST00000622113

chr21 10542394 10542448 ENST00000622113

chr21 10552656 10552716 ENST00000622113

chr21 10559493 10559544 ENST00000622113

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Join two files

What it does: This tool joins two tabular files based on a common key column

Example:

First File	
Fruit	Color
Apple	red
Banana	yellow
Orange	orange
Melon	green

Second File	
Fruit	Price
Orange	7
Avocado	8
Apple	4
Banana	3

Joining both files, using key column 1 and a header line, will return:

Joined File		
Fruit	Color	Price
Apple	red	4
Avocado	.	8
Banana	yellow	3
Melon	green	.
Orange	orange	7

Introduction to Text Manipulation(s)

Join two files - Preparing The Data

HPC Kaiser Galaxy Using 0%

Tools

- cut
- Upload Data
- Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

- All workflows

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns

c1,c2,c3,c4

Delimited by

Tab

From

9: Paste on data 8 and data 7

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:

```
apple,is,good  
windows,is,bad
```

will give:

```
apple good  
windows bad
```

What it does

This tool selects (cuts out) specified columns from the dataset.

- Columns are specified as **c1**, **c2**, and so on. Column count begins with **1**
- Columns can be specified in any order (e.g., **c2,c1,c6**)
- If you specify more columns than actually present - empty spaces will be filled with dots

Example

Input dataset (six columns: c1, c2, c3, c4, c5, and c6):

```
chr1 10 1000 gene1 0 +
```

History

search datasets

L05-A

9 shown

4.2 MB

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_B

1: Hs_Chromosome_A

8,803 regions

format: bed, database: hg38

python: error while loading shared libraries:
libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1.Chrom	2.Start	3.End	4
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494

Introduction to Text Manipulation(s)

Join two files - Preparing The Data

HPC Kaiser Galaxy Workflow Visualize Shared Data ▾ Admin Help ▾ User ▾ 🔍 🔔 🔞 🔍

Using 0%

Tools

cut x

Upload Data

Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

All workflows

chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724_f
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112_f
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395_f
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657_f
chr21 10559493 10559544 ENST00000622113.4_cds_7_0_chr21_10559494_f
chr21 10561029 10561191 ENST00000622113.4_cds_8_0_chr21_10561030_f
chr21 10567669 10567789 ENST00000622113.4_cds_9_0_chr21_10567670_f
chr21 10569436 10569536 ENST00000622113.4_cds_10_0_chr21_10569437_f
chr21 10569682 10569746 ENST00000622113.4_cds_11_0_chr21_10569683_f
chr21 10570484 10570549 ENST00000622113.4_cds_12_0_chr21_10570485_f
chr21 10577459 10577520 ENST00000622113.4_cds_13_0_chr21_10577460_f
chr21 10578336 10578418 ENST00000622113.4_cds_14_0_chr21_10578337_f
chr21 10578516 10578605 ENST00000622113.4_cds_15_0_chr21_10578517_f
chr21 10590461 10590523 ENST00000622113.4_cds_16_0_chr21_10590462_f
chr21 10592292 10592373 ENST00000622113.4_cds_17_0_chr21_10592293_f
chr21 10595981 10596087 ENST00000622113.4_cds_18_0_chr21_10595982_f
chr21 10598014 10598094 ENST00000622113.4_cds_19_0_chr21_10598015_f
chr21 10602057 10602150 ENST00000622113.4_cds_20_0_chr21_10602058_f
chr21 10603561 10603632 ENST00000622113.4_cds_21_0_chr21_10603562_f
chr21 10605416 10605552 ENST00000622113.4_cds_22_0_chr21_10605417_f
chr21 10538723 10538734 ENST00000612957.4_cds_3_0_chr21_10538724_f
chr21 10541111 10541165 ENST00000612957.4_cds_4_0_chr21_10541112_f
chr21 10543328 10543382 ENST00000612957.4_cds_5_0_chr21_10543329_f
chr21 10552656 10552716 ENST00000612957.4_cds_6_0_chr21_10552657_f
chr21 10559493 10559544 ENST00000612957.4_cds_7_0_chr21_10559494_f
chr21 10561029 10561191 ENST00000612957.4_cds_8_0_chr21_10561030_f
chr21 10567669 10567789 ENST00000612957.4_cds_9_0_chr21_10567670_f
chr21 10538723 10538734 ENST00000427445.6_cds_3_0_chr21_10538724_f
chr21 10541111 10541165 ENST00000427445.6_cds_4_0_chr21_10541112_f
chr21 10542394 10542448 ENST00000427445.6_cds_5_0_chr21_10542395_f
chr21 10559493 10559544 ENST00000427445.6_cds_6_0_chr21_10559494_f
chr21 10561029 10561191 ENST00000427445.6_cds_7_0_chr21_10561030_f
chr21 10567669 10567789 ENST00000427445.6_cds_8_0_chr21_10567670_f
chr21 10569436 10569536 ENST00000427445.6_cds_9_0_chr21_10569437_f
chr21 10569682 10569746 ENST00000427445.6_cds_10_0_chr21_10569683_f
chr21 10570484 10570549 ENST00000427445.6_cds_11_0_chr21_10570485_f
chr21 10577459 10577520 ENST00000427445.6_cds_12_0_chr21_10577460_f
chr21 10578336 10578418 ENST00000427445.6_cds_13_0_chr21_10578337_f
chr21 10578516 10578605 ENST00000427445.6_cds_14_0_chr21_10578517_f
chr21 10590461 10590523 ENST00000427445.6_cds_15_0_chr21_10590462_f
chr21 10592292 10592373 ENST00000427445.6_cds_16_0_chr21_10592293_f
chr21 10595981 10596087 ENST00000427445.6_cds_17_0_chr21_10595982_f
chr21 10598014 10598094 ENST00000427445.6_cds_18_0_chr21_10598015_f
chr21 10602057 10602150 ENST00000427445.6_cds_19_0_chr21_10602058_f
chr21 10603561 10603632 ENST00000427445.6_cds_20_0_chr21_10603562_f

History

search datasets

L05-A

10 shown
4.76 MB

10: Cut on data 9

8,803 regions
format: interval, database: hg38

display with IGV local Human hg38
display at UCSC main test

1. Chrom 2. Start 3. End 4

chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724_f
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112_f
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395_f
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657_f
chr21 10559493 10559544 ENST00000622113.4_cds_7_0_chr21_10559494_f

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_B

1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Join two files - Preparing The Data

HPC Kaiser Galaxy Using 0%

Tools

- cut
-
-

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

- All workflows

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns
c4,c5,c6

Delimited by
Tab

From
2: Hs_Chromosome_B

Job Resource Parameters
Specify job resource parameters

Memory (GB)
7

Maximum Job Memory

Time (hours)
24

Maximum job time

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:

```
apple,is,good  
windows,is,bad
```

will give:

```
apple good  
windows bad
```

What it does

This tool selects (cuts out) specified columns from the dataset.

- Columns are specified as **c1**, **c2**, and so on. Column count begins with **1**
- Columns can be specified in any order (e.g., **c2,c1,c6**)
- If you specify more columns than actually present - empty spaces will be filled with dots

Example

Input dataset (six columns: c1, c2, c3, c4, c5, and c6):

```
chr1 10 1000 gene1 0 +
```

History

search datasets

L05-A
10 shown
4.76 MB

10: Cut on data 9
9: Paste on data 8 and data 7
8: Cut on data 1
7: Cut on data 1
6: Concatenate datasets on data 2 and data 1
5: Remove beginning on data 1
4: Select last on data 1
3: Select first on data 1
2: Hs_Chromosome_B
8,803 regions
format: bed, database: hg38
python: error while loading shared libraries:
libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View
display with IGV local Human hg38
display at UCSC main test

2: Start 3: End 4 5 6
10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724_f 0 +
10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112_f 0 +
10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395_f 0 +
10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657_f 0 +
10559493 10559544 ENST00000622113.4_cds_7_0_chr21_10559494_f 0 +

1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Join two files - Preparing The Data

HPC Kaiser Galaxy Workflow Visualize Shared Data Admin Help User 🔍 🔔 🔞 🔒 Using 0%

Tools

cut x

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

All workflows

ENST00000622113.4_cds_3_0_chr21_10538724_f 0 +
ENST00000622113.4_cds_4_0_chr21_10541112_f 0 +
ENST00000622113.4_cds_5_0_chr21_10542395_f 0 +
ENST00000622113.4_cds_6_0_chr21_10552657_f 0 +
ENST00000622113.4_cds_7_0_chr21_10559494_f 0 +
ENST00000622113.4_cds_8_0_chr21_10561030_f 0 +
ENST00000622113.4_cds_9_0_chr21_10567670_f 0 +
ENST00000622113.4_cds_10_0_chr21_10569437_f 0 +
ENST00000622113.4_cds_11_0_chr21_10569683_f 0 +
ENST00000622113.4_cds_12_0_chr21_10570485_f 0 +
ENST00000622113.4_cds_13_0_chr21_10577460_f 0 +
ENST00000622113.4_cds_14_0_chr21_10578337_f 0 +
ENST00000622113.4_cds_15_0_chr21_10578517_f 0 +
ENST00000622113.4_cds_16_0_chr21_10590462_f 0 +
ENST00000622113.4_cds_17_0_chr21_10592293_f 0 +
ENST00000622113.4_cds_18_0_chr21_10595982_f 0 +
ENST00000622113.4_cds_19_0_chr21_10598015_f 0 +
ENST00000622113.4_cds_20_0_chr21_10602058_f 0 +
ENST00000622113.4_cds_21_0_chr21_10603562_f 0 +
ENST00000622113.4_cds_22_0_chr21_10605417_f 0 +
ENST00000612957.4_cds_3_0_chr21_10538724_f 0 +
ENST00000612957.4_cds_4_0_chr21_10541112_f 0 +
ENST00000612957.4_cds_5_0_chr21_10543329_f 0 +
ENST00000612957.4_cds_6_0_chr21_10552657_f 0 +
ENST00000612957.4_cds_7_0_chr21_10559494_f 0 +
ENST00000612957.4_cds_8_0_chr21_10561030_f 0 +
ENST00000612957.4_cds_9_0_chr21_10567670_f 0 +
ENST00000427445.6_cds_3_0_chr21_10538724_f 0 +
ENST00000427445.6_cds_4_0_chr21_10541112_f 0 +
ENST00000427445.6_cds_5_0_chr21_10542395_f 0 +
ENST00000427445.6_cds_6_0_chr21_10559494_f 0 +
ENST00000427445.6_cds_7_0_chr21_10561030_f 0 +
ENST00000427445.6_cds_8_0_chr21_10567670_f 0 +
ENST00000427445.6_cds_9_0_chr21_10569437_f 0 +
ENST00000427445.6_cds_10_0_chr21_10569683_f 0 +
ENST00000427445.6_cds_11_0_chr21_10570485_f 0 +
ENST00000427445.6_cds_12_0_chr21_10577460_f 0 +
ENST00000427445.6_cds_13_0_chr21_10578337_f 0 +
ENST00000427445.6_cds_14_0_chr21_10578517_f 0 +
ENST00000427445.6_cds_15_0_chr21_10590462_f 0 +
ENST00000427445.6_cds_16_0_chr21_10592293_f 0 +
ENST00000427445.6_cds_17_0_chr21_10595982_f 0 +
ENST00000427445.6_cds_18_0_chr21_10598015_f 0 +
ENST00000427445.6_cds_19_0_chr21_10602058_f 0 +
ENST00000427445.6_cds_20_0_chr21_10603562_f 0 +

History

search datasets

L05-A
11 shown
5.16 MB

11: Cut on data 2
10: Cut on data 9
9: Paste on data 8 and data 7
8: Cut on data 1
7: Cut on data 1
6: Concatenate datasets on data 2 and data 1
5: Remove beginning on data 1
4: Select last on data 1
3: Select first on data 1
2: Hs_Chr21_Table_B
1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Join two files - Performing the JOIN Operation

HPC Kaiser Galaxy Using 0%

Workflow Visualize Shared Data Admin Help User

Tools

join

Upload Data Show Sections

Join the intervals of two datasets side-by-side

Join two Datasets side by side on a specified field

Combine FASTA and QUAL into FASTQ

FASTQ interlacer on paired end reads

Join MAF blocks by Species

Group data by a column and perform aggregate operation on other columns.

Compare two Datasets to find common or distinct rows

WORKFLOWS

All workflows

Join two Datasets side by side on a specified field (Galaxy Version 2.1.3)

Join

using column

with

and column

Keep lines of first input that do not join with second input

Keep lines of first input that are incomplete

Fill empty columns

Keep the header lines

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

DO NOT Use This Tool!

Use This Tool!!!

History

search datasets

L05-A

11 shown

5.16 MB

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

This tool will force the output datatype to tabular. To change metadata assignments click on the "edit attributes" link of the history item generated by this tool.

TIP: If your data is not TAB delimited, use Text Manipulation->Convert

Syntax

This tool joins lines of two datasets on a common field. An empty string ("") is not a valid identifier. You may choose to include lines of your first input that do not join with your second input.

Introduction to Text Manipulation(s)

Join two files - Performing the JOIN Operation

HPC Kaiser Galaxy Using 0%

Tools

- join
- Upload Data
- Show Sections

Join the intervals of two datasets side-by-side

Join two Datasets side by side on a specified field

Combine FASTA and QUAL into FASTQ

FASTQ interlacer on paired end reads

Join MAF blocks by Species

Group data by a column and perform aggregate operation on other columns.

Compare two Datasets to find common or distinct rows

WORKFLOWS

- All workflows

Workflow Visualize Shared Data Admin Help User

Join two Datasets side by side on a specified field (Galaxy Version 2.1.3)

Join

10: Cut on data 9

using column

Column: 4

with

11: Cut on data 2

and column

Column: 1

Keep lines of first input that do not join with second input

No

Keep lines of first input that are incomplete

No

Fill empty columns

No

Keep the header lines

No

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

⚠ This tool will force the ouput datatype to tabular. To change metadata assignments click on the "edit attributes" link of the history item generated by this tool.

ℹ TIP: If your data is not TAB delimited, use *Text Manipulation->Convert*

Syntax

This tool joins lines of two datasets on a common field. An empty string ("") is not a valid identifier. You may choose to include lines of your first input that do not join with your second input.

History

search datasets

L05-A

11 shown
5.16 MB

11: Cut on data 2

8,803 lines
format: tabular, database: hg38

10: Cut on data 9

8,803 regions
format: interval, database: hg38

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

Introduction to Text Manipulation(s)

Join two files - Performing the JOIN Operation

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User History Using 0%

Tools

- join x
- Upload Data**
- Show Sections**
- Join the intervals of two datasets side-by-side
- Join two Datasets side by side on a specified field
- Combine FASTA and QUAL into FASTQ
- FASTQ interlacer on paired end reads
- Join MAF blocks by Species
- Group data by a column and perform aggregate operation on other columns.
- Compare two Datasets to find common or distinct rows
- WORKFLOWS**
- All workflows

chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f	ENST00000622113.4_cds_13_0_chr21_10577460_f	0	+
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f	ENST00000622113.4_cds_14_0_chr21_10578337_f	0	+
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f	ENST00000622113.4_cds_15_0_chr21_10578517_f	0	+
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f	ENST00000622113.4_cds_16_0_chr21_10590462_f	0	+
chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f	ENST00000622113.4_cds_17_0_chr21_10592293_f	0	+
chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f	ENST00000622113.4_cds_18_0_chr21_10595982_f	0	+
chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f	ENST00000622113.4_cds_19_0_chr21_10598015_f	0	+
chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f	ENST00000622113.4_cds_20_0_chr21_10602058_f	0	+
chr21	10603561	10603632	ENST00000622113.4_cds_21_0_chr21_10603562_f	ENST00000622113.4_cds_21_0_chr21_10603562_f	0	+
chr21	10605416	10605552	ENST00000622113.4_cds_22_0_chr21_10605417_f	ENST00000622113.4_cds_22_0_chr21_10605417_f	0	+
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+
chr21	10578336	10578418	ENST00000427445.6_cds_13_0_chr21_10578337_f	ENST00000427445.6_cds_13_0_chr21_10578337_f	0	+
chr21	10578516	10578605	ENST00000427445.6_cds_14_0_chr21_10578517_f	ENST00000427445.6_cds_14_0_chr21_10578517_f	0	+
chr21	10590461	10590523	ENST00000427445.6_cds_15_0_chr21_10590462_f	ENST00000427445.6_cds_15_0_chr21_10590462_f	0	+
chr21	10592292	10592373	ENST00000427445.6_cds_16_0_chr21_10592293_f	ENST00000427445.6_cds_16_0_chr21_10592293_f	0	+
chr21	10595981	10596087	ENST00000427445.6_cds_17_0_chr21_10595982_f	ENST00000427445.6_cds_17_0_chr21_10595982_f	0	+
chr21	10598014	10598094	ENST00000427445.6_cds_18_0_chr21_10598015_f	ENST00000427445.6_cds_18_0_chr21_10598015_f	0	+
chr21	10602057	10602150	ENST00000427445.6_cds_19_0_chr21_10602058_f	ENST00000427445.6_cds_19_0_chr21_10602058_f	0	+
chr21	10603561	10603632	ENST00000427445.6_cds_20_0_chr21_10603562_f	ENST00000427445.6_cds_20_0_chr21_10603562_f	0	+

History

search datasets

L05-A
12 shown
6.12 MB

12: Join two Datasets on data 11 and data 10
11: Cut on data 2
10: Cut on data 9
9: Paste on data 8 and data 7
8: Cut on data 1
7: Cut on data 1
6: Concatenate datasets on data 2 and data 1
5: Remove beginning on data 1
4: Select last on data 1
3: Select first on data 1
2: Hs_Chr21_Table_B
1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Join two files - Fixing the File

HPC Kaiser Galaxy Using 0%

Tools

- cut
- Upload Data
- Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

- All workflows

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns

c1,c2,c3,c4,c6,c7

Delimited by

Tab

From

12: Join two Datasets on data 11 and data 10

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

WARNING: This tool breaks column assignments. To re-establish column assignments run the tools and click on the pencil icon in the latest history item.

The output of this tool is always in tabular format (e.g., if your original delimiters are commas, they will be replaced with tabs). For example:

Cutting columns 1 and 3 from:

```
apple,is,good
windows,is,bad
```

will give:

```
apple    good
windows bad
```

What it does

This tool selects (cuts out) specified columns from the dataset.

- Columns are specified as **c1**, **c2**, and so on. Column count begins with **1**
- Columns can be specified in any order (e.g., **c2,c1,c6**)
- If you specify more columns than actually present - empty spaces will be filled with dots

Example

Input dataset (six columns: c1, c2, c3, c4, c5, and c6):

```
chr1 10 1000 gene1 0 +
```

History

search datasets

L05-A

12 shown
6.12 MB

12: Join two Datasets on data 11 and data 10

8,803 lines
format: tabular, database: hg38

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_1_Table_B

1: Hs_Chromosome_1_Table_A

Introduction to Text Manipulation(s)

Join two files - Fixing the File

HPC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

cut

Upload Data

Show Sections

Cut columns from a table

Cutadapt Remove adapter sequences from FASTQ/FASTA

Filter by quality

Trimmomatic flexible read trimming tool for Illumina NGS data

Trim sequences

cutseq Removes a specified section from a sequence

WORKFLOWS

All workflows

chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724_f 0 +
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112_f 0 +
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395_f 0 +
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657_f 0 +
chr21 10559493 10559544 ENST00000622113.4_cds_7_0_chr21_10559494_f 0 +
chr21 10561029 10561191 ENST00000622113.4_cds_8_0_chr21_10561030_f 0 +
chr21 10567669 10567789 ENST00000622113.4_cds_9_0_chr21_10567670_f 0 +
chr21 10569436 10569536 ENST00000622113.4_cds_10_0_chr21_10569437_f 0 +
chr21 10569682 10569746 ENST00000622113.4_cds_11_0_chr21_10569683_f 0 +
chr21 10570484 10570549 ENST00000622113.4_cds_12_0_chr21_10570485_f 0 +
chr21 10577459 10577520 ENST00000622113.4_cds_13_0_chr21_10577460_f 0 +
chr21 10578336 10578418 ENST00000622113.4_cds_14_0_chr21_10578337_f 0 +
chr21 10578516 10578605 ENST00000622113.4_cds_15_0_chr21_10578517_f 0 +
chr21 10590461 10590523 ENST00000622113.4_cds_16_0_chr21_10590462_f 0 +
chr21 10592292 10592373 ENST00000622113.4_cds_17_0_chr21_10592293_f 0 +
chr21 10595981 10596087 ENST00000622113.4_cds_18_0_chr21_10595982_f 0 +
chr21 10598014 10598094 ENST00000622113.4_cds_19_0_chr21_10598015_f 0 +
chr21 10602057 10602150 ENST00000622113.4_cds_20_0_chr21_10602058_f 0 +
chr21 10603561 10603632 ENST00000622113.4_cds_21_0_chr21_10603562_f 0 +
chr21 10605416 10605552 ENST00000622113.4_cds_22_0_chr21_10605417_f 0 +
chr21 10538723 10538734 ENST00000612957.4_cds_3_0_chr21_10538724_f 0 +
chr21 10541111 10541165 ENST00000612957.4_cds_4_0_chr21_10541112_f 0 +
chr21 10543328 10543382 ENST00000612957.4_cds_5_0_chr21_10543329_f 0 +
chr21 10552656 10552716 ENST00000612957.4_cds_6_0_chr21_10552657_f 0 +
chr21 10559493 10559544 ENST00000612957.4_cds_7_0_chr21_10559494_f 0 +
chr21 10561029 10561191 ENST00000612957.4_cds_8_0_chr21_10561030_f 0 +
chr21 10567669 10567789 ENST00000612957.4_cds_9_0_chr21_10567670_f 0 +
chr21 10538723 10538734 ENST00000427445.6_cds_3_0_chr21_10538724_f 0 +
chr21 10541111 10541165 ENST00000427445.6_cds_4_0_chr21_10541112_f 0 +
chr21 10542394 10542448 ENST00000427445.6_cds_5_0_chr21_10542395_f 0 +
chr21 10559493 10559544 ENST00000427445.6_cds_6_0_chr21_10559494_f 0 +
chr21 10561029 10561191 ENST00000427445.6_cds_7_0_chr21_10561030_f 0 +
chr21 10567669 10567789 ENST00000427445.6_cds_8_0_chr21_10567670_f 0 +
chr21 10569436 10569536 ENST00000427445.6_cds_9_0_chr21_10569437_f 0 +
chr21 10569682 10569746 ENST00000427445.6_cds_10_0_chr21_10569683_f 0 +
chr21 10570484 10570549 ENST00000427445.6_cds_11_0_chr21_10570485_f 0 +
chr21 10577459 10577520 ENST00000427445.6_cds_12_0_chr21_10577460_f 0 +
chr21 10578336 10578418 ENST00000427445.6_cds_13_0_chr21_10578337_f 0 +
chr21 10578516 10578605 ENST00000427445.6_cds_14_0_chr21_10578517_f 0 +
chr21 10590461 10590523 ENST00000427445.6_cds_15_0_chr21_10590462_f 0 +
chr21 10592292 10592373 ENST00000427445.6_cds_16_0_chr21_10592293_f 0 +
chr21 10595981 10596087 ENST00000427445.6_cds_17_0_chr21_10595982_f 0 +
chr21 10598014 10598094 ENST00000427445.6_cds_18_0_chr21_10598015_f 0 +
chr21 10602057 10602150 ENST00000427445.6_cds_19_0_chr21_10602058_f 0 +
chr21 10603561 10603632 ENST00000427445.6_cds_20_0_chr21_10603562_f 0 +

History

search datasets

L05-A

13 shown
6.72 MB

13: Cut on data 12

12: Join two Datasets on data 11 and data 10

View data

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

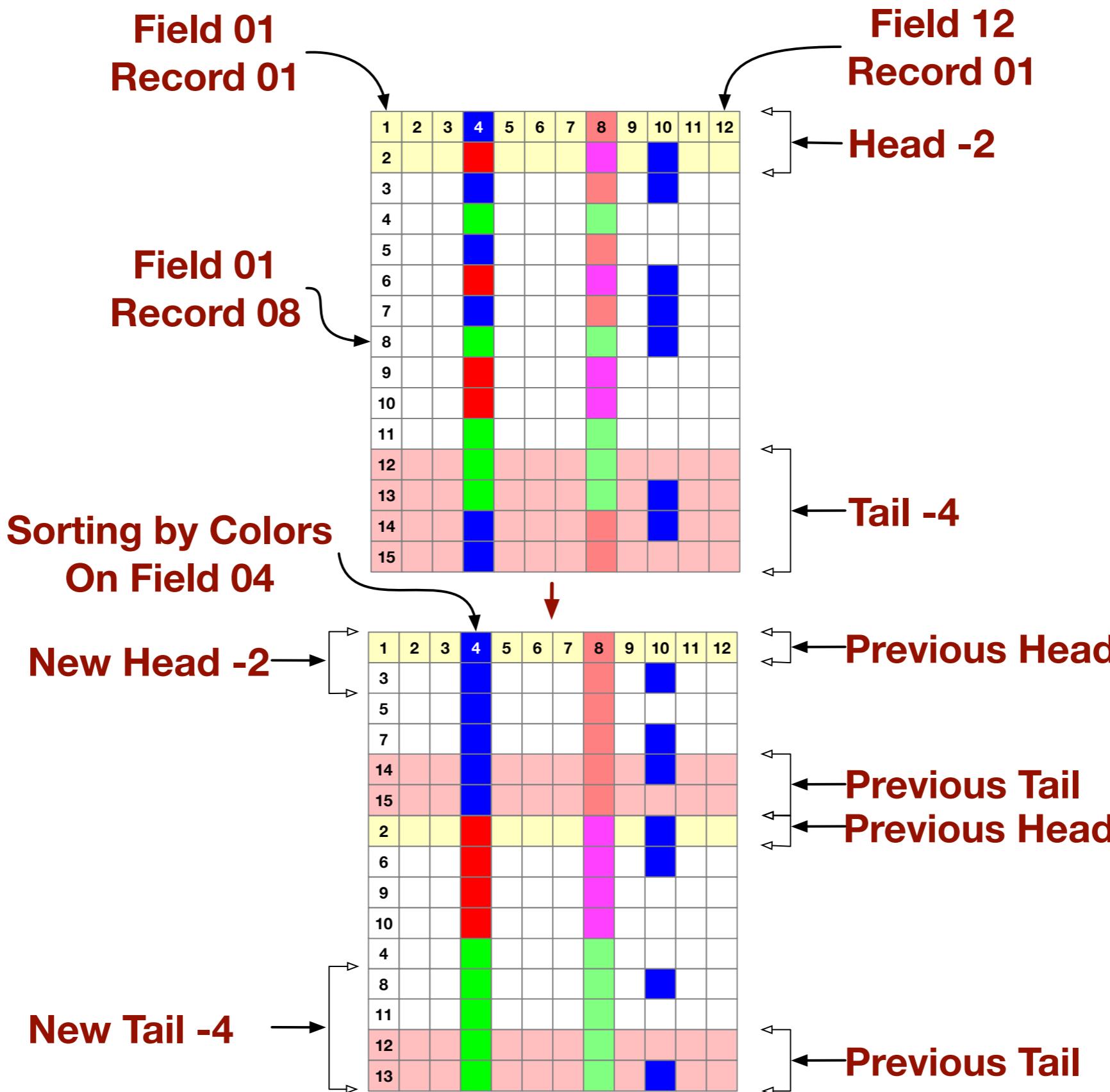
3: Select first on data 1

2: Hs_Chromosome_1_Table_B

1: Hs_Chromosome_1_Table_A

Introduction to Text Manipulation(s)

Table containing
12 fields (columns) and 15 records (rows or lines)



Introduction to Text Manipulation(s)

**Table containing
12 fields (columns) and 15 records (rows or lines)**

**Sorting by Colors
On Field 04**



1	2	3	4	5	6	7	8	9	10	11	12
3											
5											
7											
14											
15											
2											
6											
9											
10											
4											
8											
11											
12											
13											

Introduction to Text Manipulation(s)

**Table containing
12 fields (columns) and 15 records (rows or lines)**

Removing Duplicate Fields



1	2	3	4	5	6	7	8	9	10	11	12
3											
5											
2											
9											
4											
8											

Introduction to Text Manipulation(s)

Sort data in ascending or descending order

This tool sorts an input file.

Sorting Styles:

- **Fast Numeric:** sort by numeric values. Handles integer values (e.g. 43, 134) and decimal-point values (e.g. 3.14). Does not handle scientific notation (e.g. -2.32e2)
- **General Numeric:** sort by numeric values. Handles all numeric notations (including scientific notation). Slower than fast numeric, so use only when necessary.
- **Natural Sort:** Sort in 'natural' order (natural to humans, not to computers).
- **Alphabetical sort:** Sort in strict alphabetical order.
- **Human-readable numbers:** Sort human readable numbers (e.g. 1G > 2M > 3K > 400)
- **Random order:** return lines in random order

Introduction to Text Manipulation(s)

Sort data in ascending or descending order

Example – Header line

Input file (note first line is a header line, should not be sorted):

Fruit.	Color	Price
Banana	Yellow	4.1
Avocado	Green	8.0
Apple	Red	3.0
Melon	Green	6.1

Sorting by numeric order on column 3, with header, will return:

Fruit	Color	Price
Apple	Red	3.0
Banana	Yellow	4.1
Melon	Green	6.1
Avocado	Green	8.0

Introduction to Text Manipulation(s)

Sort data in ascending or descending order

Example – Natural vs. Alphabetical sorting

Given the following list:

```
chr4  
chr1  
3  
chr1  
chr1  
0
```

Alphabetical sort would produce the following sorted list:

```
chr1  
chr1  
0  
chr1  
3  
chr2
```

Natural Sort would produce the following sorted list:

```
chr1  
chr2  
chr4  
chr1  
0  
chr1
```

Introduction to Text Manipulation(s)

Sort data in ascending or descending order

HPRC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User BI

Using 0%

Tools

search tools

HPRC

- Get Data
- Send Data
- Collection Operations
- Lift-Over
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- Fetch Alignments/Sequences
- Operate on Genomic Intervals
- Statistics
- Graph/Display Data
- Phenotype Association
- BED
- Annotation
- Multiple Alignments
- NCBI BLAST+
- Mapping
- SAM/BAM
- Assembly
- FASTQ Quality Control
- FASTA/FASTQ
- RNA-seq
- CD-HIT
- Datamash
- EMBOSS
- MUMmer4
- Nanopore

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB	Bl
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+				
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+				
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+				
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+				
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+				
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+				
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+				
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+				
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+				
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+				
chr21	10577459	10577520	ENST00000622113.4_cds_13_0_chr21_10577460_f	0	+				
chr21	10578336	10578418	ENST00000622113.4_cds_14_0_chr21_10578337_f	0	+				
chr21	10578516	10578605	ENST00000622113.4_cds_15_0_chr21_10578517_f	0	+				
chr21	10590461	10590523	ENST00000622113.4_cds_16_0_chr21_10590462_f	0	+				
chr21	10592292	10592373	ENST00000622113.4_cds_17_0_chr21_10592293_f	0	+				
chr21	10595981	10596087	ENST00000622113.4_cds_18_0_chr21_10595982_f	0	+				
chr21	10598014	10598094	ENST00000622113.4_cds_19_0_chr21_10598015_f	0	+				
chr21	10602057	10602150	ENST00000622113.4_cds_20_0_chr21_10602058_f	0	+				
chr21	10603561	10603632	ENST00000622113.4_cds_21_0_chr21_10603562_f	0	+				
chr21	10605416	10605552	ENST00000622113.4_cds_22_0_chr21_10605417_f	0	+				
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+				
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+				
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+				
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+				
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+				
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+				
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+				
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+				
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+				
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+				
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+				
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+				
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+				
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+				
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+				
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+				
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+				
chr21	10578336	10578418	ENST00000427445.6_cds_13_0_chr21_10578337_f	0	+				
chr21	10578516	10578605	ENST00000427445.6_cds_14_0_chr21_10578517_f	0	+				
chr21	10590461	10590523	ENST00000427445.6_cds_15_0_chr21_10590462_f	0	+				
chr21	10592292	10592373	ENST00000427445.6_cds_16_0_chr21_10592293_f	0	+				
chr21	10595981	10596087	ENST00000427445.6_cds_17_0_chr21_10595982_f	0	+				
chr21	10598014	10598094	ENST00000427445.6_cds_18_0_chr21_10598015_f	0	+				
chr21	10602057	10602150	ENST00000427445.6_cds_19_0_chr21_10602058_f	0	+				

History

search datasets

L05-A

13 shown
6.72 MB

7: Concatenate data 1 and data 2

6: Concatenate datasets on data 2 and data 1

17,606 regions
format: bed, database: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1.Chrom 2.Start 3.End 4

chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724_f
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112_f
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395_f
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657_f
chr21 10559493 10559544 ENST00000622113.4_cds_7_0_chr21_10559494_f

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

8,803 regions
format: bed, database: hg38

python: error while loading shared libraries:
libpython3.7m.so.1.0: cannot open shared object file: No such file or directory

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1.Chrom 2.Start 3.End 4

chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724_f
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112_f
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395_f
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657_f

Introduction to Text Manipulation(s)

Sort data in ascending or descending order

HPRC Kaiser Galaxy Using 0%

Tools

sort

Sort data in ascending or descending order

SortSam sort SAM/BAM dataset

Sort collection

bedtools SortBED order the intervals

Samtools sort order of storing aligned sequences

FixMateInformation ensure that all mate-pair information is in sync between each read and its mate pair

bedtools Genome Coverage compute the coverage over an entire genome

bedtools MapBed apply a function to a column for each overlapping interval

MergeSamFiles merges multiple SAM/BAM datasets into one

SAM-to-BAM convert SAM to BAM

ReplaceSamHeader replace header in a SAM/BAM dataset

BED-to-bigBed converter

CollectGcBiasMetrics charts the GC bias metrics

MeanQualityByCycle chart distribution of base qualities

RevertSam revert SAM/BAM datasets to a previous state

QualityScoreDistribution chart quality score distribution

CollectBaseDistributionByCycle charts the nucleotide distribution per cycle in a SAM or BAM dataset

Collect Alignment Summary Metrics writes a file containing summary alignment metrics

bedtools MultiCovBed counts coverage from multiple BAMs at specific intervals

Workflow Visualize Shared Data Admin Help User

Sort data in ascending or descending order (Galaxy Version 1.2.0)

Sort Dataset

13: Cut on data 12

on column

Column: 2

with flavor

Numerical sort

everything in

Ascending order

Column selection

+ Insert Column selection

Number of header lines to skip

0

characters are already considered as comments and kept

Job Resource Parameters

Specify job resource parameters

Memory (GB)

7

Maximum Job Memory

Time (hours)

24

Maximum job time

Execute

TIP: If your data is not TAB delimited, use *Text Manipulation->Convert*

Syntax

This tool sorts the dataset on any number of columns in either ascending or descending order.

- Numerical sort** orders numbers by their magnitude, ignores all characters besides numbers, and evaluates a string of numbers to the value they signify.
- General numeric sort** orders numbers by their general numerical value. Unlike the numerical sort option, it can handle numbers in scientific notation too.
- Alphabetical sort** is a phonebook type sort based on the conventional order of letters in an alphabet. Each nth letter is compared with the nth letter of other words in the list, starting at the first letter of each word and advancing to the second, third, fourth, and so on, until the order is established. Therefore, in an alphabetical sort, 2 comes after 100 (1 < 2).

Examples

History

search datasets

L05-A

13 shown
6.72 MB

13: Cut on data 12

8,803 lines
format: tabular, database: hg38

1 2 3 4

chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_10542395
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_10552657
chr21 10559493 10559544 ENST00000622113.4_cds_7_0_chr21_10559494

12: Join two Datasets on data 11 and data 10

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Sort data in ascending or descending order

HPC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

sort x

Upload Data

Show Sections

Sort data in ascending or descending order

SortSam sort SAM/BAM dataset

Sort collection

bedtools SortBED order the intervals

Samtools sort order of storing aligned sequences

FixMateInformation ensure that all mate-pair information is in sync between each read and it's mate pair

bedtools Genome Coverage

compute the coverage over an entire genome

bedtools MapBed apply a function to a column for each overlapping interval

MergeSamFiles merges multiple SAM/BAM datasets into one

SAM-to-BAM convert SAM to BAM

ReplaceSamHeader replace header in a SAM/BAM dataset

BED-to-bigBed converter

CollectGcBiasMetrics charts the GC bias metrics

MeanQualityByCycle chart distribution of base qualities

RevertSam revert SAM/BAM datasets to a previous state

QualityScoreDistribution chart quality score distribution

CollectBaseDistributionByCycle charts the nucleotide distribution per cycle in a SAM or BAM dataset

Collect Alignment Summary Metrics writes a file containing summary alignment metrics

bedtools MultiCovBed counts coverage from multiple BAMs at specific intervals

chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+
chr21	10538723	10538734	ENST00000618007.5_cds_3_0_chr21_10538724_f	0	+
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+
chr21	10541111	10541165	ENST00000618007.5_cds_4_0_chr21_10541112_f	0	+
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+
chr21	10542394	10542448	ENST00000618007.5_cds_5_0_chr21_10542395_f	0	+
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+
chr21	10543328	10543382	ENST00000618007.5_cds_6_0_chr21_10543329_f	0	+
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+
chr21	10552656	10552716	ENST00000618007.5_cds_7_0_chr21_10552657_f	0	+
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+
chr21	10559493	10559544	ENST00000618007.5_cds_8_0_chr21_10559494_f	0	+
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+
chr21	10561029	10561191	ENST00000618007.5_cds_9_0_chr21_10561030_f	0	+
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+
chr21	10561159	10561191	ENST00000612746.1_cds_4_0_chr21_10561160_f	0	+
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000612746.1_cds_5_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000618007.5_cds_10_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+
chr21	10569436	10569536	ENST00000612746.1_cds_6_0_chr21_10569437_f	0	+
chr21	10569436	10569536	ENST00000618007.5_cds_11_0_chr21_10569437_f	0	+
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+
chr21	10569682	10569746	ENST00000612746.1_cds_7_0_chr21_10569683_f	0	+
chr21	10569682	10569746	ENST00000618007.5_cds_12_0_chr21_10569683_f	0	+
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+
chr21	10570484	10570549	ENST00000612746.1_cds_8_0_chr21_10570485_f	0	+
chr21	10570484	10570549	ENST00000618007.5_cds_13_0_chr21_10570485_f	0	+
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+
chr21	10577459	10577520	ENST00000612746.1_cds_9_0_chr21_10577460_f	0	+
chr21	10577459	10577520	ENST00000618007.5_cds_14_0_chr21_10577460_f	0	+

History

search datasets

L05-A
14 shown
7.32 MB

14: Sort on data 13
13: Cut on data 12
12: Join two Datasets on data 11 and data 10
11: Cut on data 2
10: Cut on data 9
9: Paste on data 8 and data 7
8: Cut on data 1
7: Cut on data 1
6: Concatenate datasets on data 2 and data 1
5: Remove beginning on data 1
4: Select last on data 1
3: Select first on data 1
2: Hs_Chr21_Table_B
1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Unique occurrences of each record

What it does:

This tool returns all unique lines using the 'sort -u' command.

It can be used with unsorted files

The input file needs to be tab separated. Please convert your file if necessary

Input File:

```
chr1  10  100  gene1
chr1  105 200  gene2
chr1  10  100  gene1
chr2  10  100  gene4
chr2 1000 1900  gene5
chr3  15  1656 gene6
chr2  10  100  gene4
```

Unique lines will result in:

```
chr1  10  100  gene1
chr1  105 200  gene2
chr2  10  100  gene4
chr2 1000 1900  gene5
chr3  15  1656 gene6
```

Introduction to Text Manipulation(s)

Unique occurrences of each record

HPC Kaiser Galaxy Using 0%

Tools

- unique
-
-

unique_line remove duplicate lines
Count occurrences of each record
Group data by a column and perform aggregate operation on other columns.
Mummer Align two or more sequences
Delta-Filter Filters alignment (delta) file from nucmer
bedtools FisherBed calculate Fisher statistic between two feature files
Nucmer Align two or more sequences
bedtools ClusterBed cluster overlapping/nearby intervals
AddOrReplaceReadGroups add or replaces read group information
FastqToSam convert Fastq data into unaligned BAM
Map with BWA - map short reads (< 100 bp) against reference genome
Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome
NCBI BLAST+ blastdbcmd entry(s)
Extract sequence(s) from BLAST database
Flye de novo assembler for single molecule sequencing reads
Merge collections
velvetg Velvet sequence assembler for very short reads
velveth Prepare a dataset for the Velvet velvetg Assembler
FASTA-to-Tabular converter
VelvetOptimiser Automatically optimize Velvet assemblies
Bowtie2 - map reads against reference genome

unique_line remove duplicate lines (Galaxy Version 1.0.0)

Original file 14: Sort on data 13

Job Resource Parameters

Specify job resource parameters

Memory (GB) Maximum Job Memory

Time (hours) Maximum job time

History

search datasets

L05-A
14 shown
7.32 MB

14: Sort on data 13
8,803 lines
format: tabular, database: hg38

1 2 3 4
chr21 10538723 10538734 ENST00000427445.6_cds_3_0_chr21_10538724
chr21 10538723 10538734 ENST00000612957.4_cds_3_0_chr21_10538724
chr21 10538723 10538734 ENST00000618007.5_cds_3_0_chr21_10538724
chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_10538724
chr21 10541111 10541165 ENST00000427445.6_cds_4_0_chr21_10541112

13: Cut on data 12

12: Join two Datasets on data 11 and data 10

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_B

1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Unique occurrences of each record

HPC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User Using 0%

Tools

- unique x
- Upload Data**
- Show Sections**
- unique_line** remove duplicate lines
- Count** occurrences of each record
- Group** data by a column and perform aggregate operation on other columns.
- Mummer** Align two or more sequences
- Delta-Filter** Filters alignment (delta) file from nucmer
- bedtools FisherBed** calculate Fisher statistic between two feature files
- Nucmer** Align two or more sequences
- bedtools ClusterBed** cluster overlapping/nearby intervals
- AddOrReplaceReadGroups** add or replaces read group information
- FastqToSam** convert Fastq data into unaligned BAM
- Map with BWA** - map short reads (< 100 bp) against reference genome
- Map with BWA-MEM** - map medium and long reads (> 100 bp) against reference genome
- NCBI BLAST+ blastdbcmd entry(s)**
Extract sequence(s) from BLAST database
- Flye** de novo assembler for single molecule sequencing reads
- Merge collections**
- velvetg** Velvet sequence assembler for very short reads
- velveth** Prepare a dataset for the Velvet velvetg Assembler
- FASTA-to-Tabular** converter
- VelvetOptimiser** Automatically optimize Velvet assemblies
- Bowtie2** - map reads against reference genome

chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+
chr21	10538723	10538734	ENST00000618007.5_cds_3_0_chr21_10538724_f	0	+
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+
chr21	10541111	10541165	ENST00000618007.5_cds_4_0_chr21_10541112_f	0	+
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+
chr21	10542394	10542448	ENST00000618007.5_cds_5_0_chr21_10542395_f	0	+
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+
chr21	10543328	10543382	ENST00000618007.5_cds_6_0_chr21_10543329_f	0	+
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+
chr21	10552656	10552716	ENST00000618007.5_cds_7_0_chr21_10552657_f	0	+
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+
chr21	10559493	10559544	ENST00000618007.5_cds_8_0_chr21_10559494_f	0	+
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+
chr21	10561029	10561191	ENST00000618007.5_cds_9_0_chr21_10561030_f	0	+
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+
chr21	10561159	10561191	ENST00000612746.1_cds_4_0_chr21_10561160_f	0	+
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000612746.1_cds_5_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000618007.5_cds_10_0_chr21_10567670_f	0	+
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+
chr21	10569436	10569536	ENST00000612746.1_cds_6_0_chr21_10569437_f	0	+
chr21	10569436	10569536	ENST00000618007.5_cds_11_0_chr21_10569437_f	0	+
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+
chr21	10569682	10569746	ENST00000612746.1_cds_7_0_chr21_10569683_f	0	+
chr21	10569682	10569746	ENST00000618007.5_cds_12_0_chr21_10569683_f	0	+
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+
chr21	10570484	10570549	ENST00000612746.1_cds_8_0_chr21_10570485_f	0	+
chr21	10570484	10570549	ENST00000618007.5_cds_13_0_chr21_10570485_f	0	+
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+
chr21	10577459	10577520	ENST00000612746.1_cds_9_0_chr21_10577460_f	0	+
chr21	10577459	10577520	ENST00000618007.5_cds_14_0_chr21_10577460_f	0	+

History

search datasets

L05-A
15 shown
7.92 MB

15: unique_line on data 14
8,803 lines
format: tabular, database: hg38

14: Sort on data 13

13: Cut on data 12

12: Join two Datasets on data 11 and data 10

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chr21_Table_B

1: Hs_Chr21_Table_A

Introduction to Text Manipulation(s)

Unique occurrences of each record

HPC Kaiser Galaxy Using 0%

Tools

unique

unique_line remove duplicate lines
Count occurrences of each record
Group data by a column and perform aggregate operation on other columns.
Mummer Align two or more sequences
Delta-Filter Filters alignment (delta) file from nucmer
bedtools FisherBed calculate Fisher statistic between two feature files
Nucmer Align two or more sequences
bedtools ClusterBed cluster overlapping/nearby intervals
AddOrReplaceReadGroups add or replaces read group information
FastqToSam convert Fastq data into unaligned BAM
Map with BWA - map short reads (< 100 bp) against reference genome
Map with BWA-MEM - map medium and long reads (> 100 bp) against reference genome
NCBI BLAST+ blastdbcmd entry(s)
Extract sequence(s) from BLAST database
Flye de novo assembler for single molecule sequencing reads
Merge collections
velvetg Velvet sequence assembler for very short reads
velveth Prepare a dataset for the Velvet velvetg Assembler
FASTA-to-Tabular converter
VelvetOptimiser Automatically optimize Velvet assemblies
Bowtie2 - map reads against reference genome

unique_line remove duplicate lines (Galaxy Version 1.0.0)

Original file 6: Concatenate datasets on data 2 and data 1

Job Resource Parameters

Specify job resource parameters

Memory (GB) Maximum Job Memory

Time (hours) Maximum job time

History

search datasets

L05-A

15 shown
7.92 MB

15: unique_line on data 14
14: Sort on data 13
13: Cut on data 12
12: Join two Datasets on data 11 and data 10
11: Cut on data 2
10: Cut on data 9
9: Paste on data 8 and data 7
8: Cut on data 1
7: Cut on data 1
6: Concatenate datasets on data 2 and data 1
17,606 regions
format: bed, database: hg38

1.Chrom 2.Start 3.End 4
chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_105387
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_105411
chr21 10542394 10542448 ENST00000622113.4_cds_5_0_chr21_105423
chr21 10552656 10552716 ENST00000622113.4_cds_6_0_chr21_105526
chr21 10559493 10559544 ENST00000622113.4_cds_7_0_chr21_105594

5: Remove beginning on data 1
4: Select last on data 1
3: Select first on data 1
2: Hs_Chromosome_B
1: Hs_Chromosome_A

Introduction to Text Manipulation(s)

Unique occurrences of each record

HPC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

- search tools
- [Upload Data](#)

HPRC

- [Get Data](#)
- [Send Data](#)
- [Collection Operations](#)
- [Lift-Over](#)
- [Text Manipulation](#)
- [Convert Formats](#)
- [Filter and Sort](#)
- [Join, Subtract and Group](#)
- [Fetch Alignments/Sequences](#)
- [Operate on Genomic Intervals](#)
- [Statistics](#)
- [Graph/Display Data](#)
- [Phenotype Association](#)
- [BED](#)
- [Annotation](#)
- [Multiple Alignments](#)
- [NCBI BLAST+](#)
- [Mapping](#)
- [SAM/BAM](#)
- [Assembly](#)
- [FASTQ Quality Control](#)
- [FASTA/FASTQ](#)
- [RNA-seq](#)
- [CD-HIT](#)
- [Datamash](#)
- [EMBOSS](#)
- [MUMmer4](#)
- [Nanopore](#)

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+			
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+			
chr21	10538723	10538734	ENST00000618007.5_cds_3_0_chr21_10538724_f	0	+			
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+			
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+			
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+			
chr21	10541111	10541165	ENST00000618007.5_cds_4_0_chr21_10541112_f	0	+			
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+			
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+			
chr21	10542394	10542448	ENST00000618007.5_cds_5_0_chr21_10542395_f	0	+			
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+			
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+			
chr21	10543328	10543382	ENST00000618007.5_cds_6_0_chr21_10543329_f	0	+			
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+			
chr21	10552656	10552716	ENST00000618007.5_cds_7_0_chr21_10552657_f	0	+			
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+			
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+			
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+			
chr21	10559493	10559544	ENST00000618007.5_cds_8_0_chr21_10559494_f	0	+			
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+			
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+			
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+			
chr21	10561029	10561191	ENST00000618007.5_cds_9_0_chr21_10561030_f	0	+			
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+			
chr21	10561159	10561191	ENST00000612746.1_cds_4_0_chr21_10561160_f	0	+			
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+			
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+			
chr21	10567669	10567789	ENST00000612746.1_cds_5_0_chr21_10567670_f	0	+			
chr21	10567669	10567789	ENST00000618007.5_cds_10_0_chr21_10567670_f	0	+			
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+			
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+			
chr21	10569436	10569536	ENST00000612746.1_cds_6_0_chr21_10569437_f	0	+			
chr21	10569436	10569536	ENST00000618007.5_cds_11_0_chr21_10569437_f	0	+			
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+			
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+			
chr21	10569682	10569746	ENST00000612746.1_cds_7_0_chr21_10569683_f	0	+			
chr21	10569682	10569746	ENST00000618007.5_cds_12_0_chr21_10569683_f	0	+			
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+			
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+			
chr21	10570484	10570549	ENST00000612746.1_cds_8_0_chr21_10570485_f	0	+			
chr21	10570484	10570549	ENST00000618007.5_cds_13_0_chr21_10570485_f	0	+			
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+			
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+			
chr21	10577459	10577520	ENST00000612746.1_cds_9_0_chr21_10577460_f	0	+			

History

search datasets

L05-A

16 shown
8.52 MB

16: unique_line on data 6

8,803 regions
format: bed, database: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom 2. Start 3. End 4

chr21 10538723 10538734 ENST00000427445.6_cds_3_0_chr21_105387
chr21 10538723 10538734 ENST00000612957.4_cds_3_0_chr21_105387
chr21 10538723 10538734 ENST00000618007.5_cds_3_0_chr21_105387
chr21 10538723 10538734 ENST00000622113.4_cds_3_0_chr21_105387
chr21 10541111 10541165 ENST00000427445.6_cds_4_0_chr21_10541112
chr21 10541111 10541165 ENST00000612957.4_cds_4_0_chr21_10541112
chr21 10541111 10541165 ENST00000618007.5_cds_4_0_chr21_10541112
chr21 10541111 10541165 ENST00000622113.4_cds_4_0_chr21_10541112

15: unique_line on data 14

14: Sort on data 13

13: Cut on data 12

12: Join two Datasets on data 11 and data 10

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

17,606 regions
format: bed, database: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom 2. Start 3. End 4

Introduction to Text Manipulation(s)

Add column to an existing dataset

What it does: You can enter any value and it will be added as a new column to your dataset

Example: If your original data looks like this:

```
chr1 10 100 geneA  
chr2 200 300 geneB  
chr2 400 500 geneC
```

Typing **+** in the text box will generate:

```
chr1 10 100 geneA +  
chr2 200 300 geneB +  
chr2 400 500 geneC +
```

You can also add line numbers by selecting **Iterate: YES**. In this case if you enter **1** in the text box you will get:

```
chr1 10 100 geneA 1  
chr2 200 300 geneB 2  
chr2 400 500 geneC 3
```

Introduction to Text Manipulation(s)

Add column to an existing dataset

HPC Kaiser Galaxy Using 0%

Tools

add

Upload Data

Show Sections

AddCommentsToBam add comments to BAM dataset

RevertOriginalBaseQualitiesAndAddMateCigarTag revert the original base qualities and add the mate cigar tag

AddOrReplaceReadGroups add or replaces read group information

Add column to an existing dataset

History Divider

Fetch closest non-overlapping feature for every interval

ClustalW multiple sequence alignment program for DNA or proteins

FixMateInformation ensure that all mate-pair information is in sync between each read and it's mate pair

CleanSam perform SAM/BAM grooming

bedtools WindowBed find overlapping intervals within a window around an interval

SortSam sort SAM/BAM dataset

Merge Columns together

Filter BAM datasets on a variety of attributes

CollectHsMetrics compute metrics about datasets generated through hybrid-selection (e.g. exome)

ReplaceSamHeader replace header in a SAM/BAM dataset

BedToIntervalList convert coordinate data into picard interval list format

NormalizeFasta normalize fasta datasets

CollectWgsMetrics compute metrics for evaluating of whole genome sequencing experiments

Add column to an existing dataset (Galaxy Version 1.0.0)

Add this value: 1

to Dataset: 16: unique_line on data 6

Dataset missing? See TIP below

Iterate?: NO

Job Resource Parameters

Specify job resource parameters

Memory (GB): 7

Time (hours): 24

Execute

TIP: If your data is not TAB delimited, use *Text Manipulation->Convert*

What it does: You can enter any value and it will be added as a new column to your dataset

Example: If your original data looks like this:

```
chr1 10 100 geneA
chr2 200 300 geneB
chr2 400 500 geneC
```

Typing + in the text box will generate:

```
chr1 10 100 geneA +
chr2 200 300 geneB +
chr2 400 500 geneC +
```

You can also add line numbers by selecting **Iterate: YES**. In this case if you enter 1 in the text box you will get:

```
chr1 10 100 geneA 1
chr2 200 300 geneB 2
chr2 400 500 geneC 3
```

History

search datasets

L05-A

16 shown
8.52 MB

16: unique_line on data 6

15: unique_line on data 14

14: Sort on data 13

13: Cut on data 12

12: Join two Datasets on data 11 and data 10

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1

2: Hs_Chromosome_21_Table_B

1: Hs_Chromosome_21_Table_A

Introduction to Text Manipulation(s)

Add column to an existing dataset

HPC Kaiser Galaxy

Workflow Visualize Shared Data Admin Help User

Using 0%

Tools

- x
- [Upload Data](#)
- [Show Sections](#)

AddCommentsToBam add comments to BAM dataset

RevertOriginalBaseQualitiesAndAddMateCigarTag revert the original base qualities and add the mate cigar tag

AddOrReplaceReadGroups add or replaces read group information

Add column to an existing dataset

History Divider

Fetch closest non-overlapping feature for every interval

ClustalW multiple sequence alignment program for DNA or proteins

FixMateInformation ensure that all mate-pair information is in sync between each read and it's mate pair

CleanSam perform SAM/BAM grooming

bedtools WindowBed find overlapping intervals within a window around an interval

SortSam sort SAM/BAM dataset

Merge Columns together

Filter BAM datasets on a variety of attributes

CollectHsMetrics compute metrics about datasets generated through hybrid-selection (e.g. exome)

ReplaceSamHeader replace header in a SAM/BAM dataset

BedToIntervalList convert coordinate data into picard interval list format

NormalizeFasta normalize fasta datasets

CollectWgsMetrics compute metrics for evaluating of whole genome sequencing experiments

Chrom	Start	End	Name	Score	Strand	ThickStart	ThickEnd	ItemRGB
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_10538724_f	0	+	1		
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_10538724_f	0	+	1		
chr21	10538723	10538734	ENST00000618007.5_cds_3_0_chr21_10538724_f	0	+	1		
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_10538724_f	0	+	1		
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_10541112_f	0	+	1		
chr21	10541111	10541165	ENST00000612957.4_cds_4_0_chr21_10541112_f	0	+	1		
chr21	10541111	10541165	ENST00000618007.5_cds_4_0_chr21_10541112_f	0	+	1		
chr21	10541111	10541165	ENST00000622113.4_cds_4_0_chr21_10541112_f	0	+	1		
chr21	10542394	10542448	ENST00000427445.6_cds_5_0_chr21_10542395_f	0	+	1		
chr21	10542394	10542448	ENST00000618007.5_cds_5_0_chr21_10542395_f	0	+	1		
chr21	10542394	10542448	ENST00000622113.4_cds_5_0_chr21_10542395_f	0	+	1		
chr21	10543328	10543382	ENST00000612957.4_cds_5_0_chr21_10543329_f	0	+	1		
chr21	10543328	10543382	ENST00000618007.5_cds_6_0_chr21_10543329_f	0	+	1		
chr21	10552656	10552716	ENST00000612957.4_cds_6_0_chr21_10552657_f	0	+	1		
chr21	10552656	10552716	ENST00000618007.5_cds_7_0_chr21_10552657_f	0	+	1		
chr21	10552656	10552716	ENST00000622113.4_cds_6_0_chr21_10552657_f	0	+	1		
chr21	10559493	10559544	ENST00000427445.6_cds_6_0_chr21_10559494_f	0	+	1		
chr21	10559493	10559544	ENST00000612957.4_cds_7_0_chr21_10559494_f	0	+	1		
chr21	10559493	10559544	ENST00000618007.5_cds_8_0_chr21_10559494_f	0	+	1		
chr21	10559493	10559544	ENST00000622113.4_cds_7_0_chr21_10559494_f	0	+	1		
chr21	10561029	10561191	ENST00000427445.6_cds_7_0_chr21_10561030_f	0	+	1		
chr21	10561029	10561191	ENST00000612957.4_cds_8_0_chr21_10561030_f	0	+	1		
chr21	10561029	10561191	ENST00000618007.5_cds_9_0_chr21_10561030_f	0	+	1		
chr21	10561029	10561191	ENST00000622113.4_cds_8_0_chr21_10561030_f	0	+	1		
chr21	10561159	10561191	ENST00000612746.1_cds_4_0_chr21_10561160_f	0	+	1		
chr21	10567669	10567699	ENST00000612957.4_cds_9_0_chr21_10567670_f	0	+	1		
chr21	10567669	10567789	ENST00000427445.6_cds_8_0_chr21_10567670_f	0	+	1		
chr21	10567669	10567789	ENST00000612746.1_cds_5_0_chr21_10567670_f	0	+	1		
chr21	10567669	10567789	ENST00000618007.5_cds_10_0_chr21_10567670_f	0	+	1		
chr21	10567669	10567789	ENST00000622113.4_cds_9_0_chr21_10567670_f	0	+	1		
chr21	10569436	10569536	ENST00000427445.6_cds_9_0_chr21_10569437_f	0	+	1		
chr21	10569436	10569536	ENST00000612746.1_cds_6_0_chr21_10569437_f	0	+	1		
chr21	10569436	10569536	ENST00000618007.5_cds_11_0_chr21_10569437_f	0	+	1		
chr21	10569436	10569536	ENST00000622113.4_cds_10_0_chr21_10569437_f	0	+	1		
chr21	10569682	10569746	ENST00000427445.6_cds_10_0_chr21_10569683_f	0	+	1		
chr21	10569682	10569746	ENST00000612746.1_cds_7_0_chr21_10569683_f	0	+	1		
chr21	10569682	10569746	ENST00000618007.5_cds_12_0_chr21_10569683_f	0	+	1		
chr21	10569682	10569746	ENST00000622113.4_cds_11_0_chr21_10569683_f	0	+	1		
chr21	10570484	10570549	ENST00000427445.6_cds_11_0_chr21_10570485_f	0	+	1		
chr21	10570484	10570549	ENST00000612746.1_cds_8_0_chr21_10570485_f	0	+	1		
chr21	10570484	10570549	ENST00000618007.5_cds_13_0_chr21_10570485_f	0	+	1		
chr21	10570484	10570549	ENST00000622113.4_cds_12_0_chr21_10570485_f	0	+	1		
chr21	10577459	10577520	ENST00000427445.6_cds_12_0_chr21_10577460_f	0	+	1		
chr21	10577459	10577520	ENST00000612746.1_cds_9_0_chr21_10577460_f	0	+	1		

History

search datasets

L05-A
17 shown
9.14 MB

17: Add column on data 16
8,803 regions
format: bed, database: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main test

1. Chrom	2. Start	3. End	4.
chr21	10538723	10538734	ENST00000427445.6_cds_3_0_chr21_105387
chr21	10538723	10538734	ENST00000612957.4_cds_3_0_chr21_105387
chr21	10538723	10538734	ENST00000618007.5_cds_3_0_chr21_105387
chr21	10538723	10538734	ENST00000622113.4_cds_3_0_chr21_105387
chr21	10541111	10541165	ENST00000427445.6_cds_4_0_chr21_105411

16: unique_line on data 6

15: unique_line on data 14

14: Sort on data 13

13: Cut on data 12

12: Join two Datasets on data 11 and data 10

11: Cut on data 2

10: Cut on data 9

9: Paste on data 8 and data 7

8: Cut on data 1

7: Cut on data 1

6: Concatenate datasets on data 2 and data 1

5: Remove beginning on data 1

4: Select last on data 1

3: Select first on data 1