**MakeHub Creating Individual Assembly Hubs for Display with the UCSC Genome Browser**

**Katharina J. Hoff**

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

Introduction

Pipeline

Connecting Your Hub

References

# **MakeHub**

# Creating Individual Assembly Hubs for Display with the UCSC Genome Browser

Plant and Animal Genomes XXVIII, January 14th 2020

Katharina J. Hoff

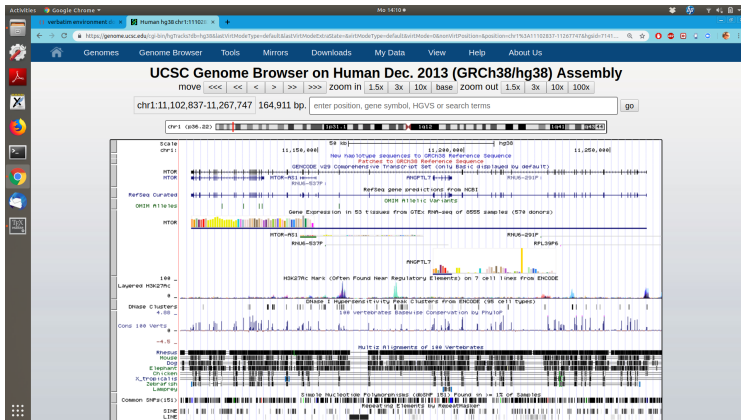Contact: katharina.hoff@uni-greifswald.de

# Contents

**1** **Introduction**

**2** **Pipeline**

**3** **Connecting Your Hub**

# UCSC Genome Browser



- Browser is hosted by UCSC → no installation required
- Assembly Hubs can be locally generated and hosted

MakeHub
Creating Individual
Assembly Hubs for
Display with the
UCSC Genome
Browser

Katharina J. Hoff

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

Introduction

Pipeline

Connecting Your Hub

References

**MakeHub**
**Creating Individual**
**Assembly Hubs for**
**Display with the**
**UCSC Genome**
**Browser**

**Katharina J. Hoff**

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

Introduction

Pipeline

Connecting Your Hub

References

# UCSC Genome Browser & MakeHub
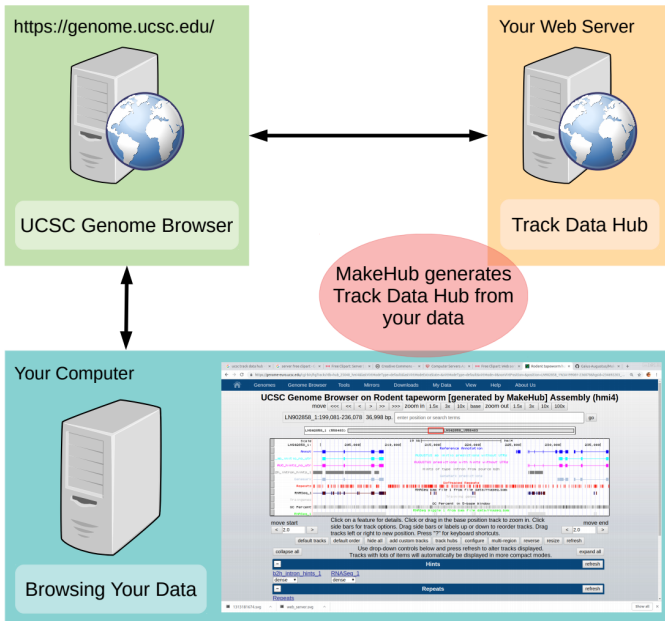
## Prerequisites

**Unix-System**

- MakeHub from `https://github.com/Gaius-Augustus/MakeHub`
  Obtain with
  `git clone https://github.com/Gaius-Augustus/MakeHub.git`
- Python 3.5 including biopython & numpy
- SAMtools
- optional: bam2wig from AUGUSTUS
- UCSC tools:
  - bedToBigBed
  - genePredCheck
  - faToTwoBit
  - gtfToGenePred
  - hgGcPercent
  - ixIxx
  - twoBitInfo
  - wigToBigWig
  - genePredToBed

# MakeHub Workflow



**Software:** https://github.com/Gaius-Augustus/MakeHub
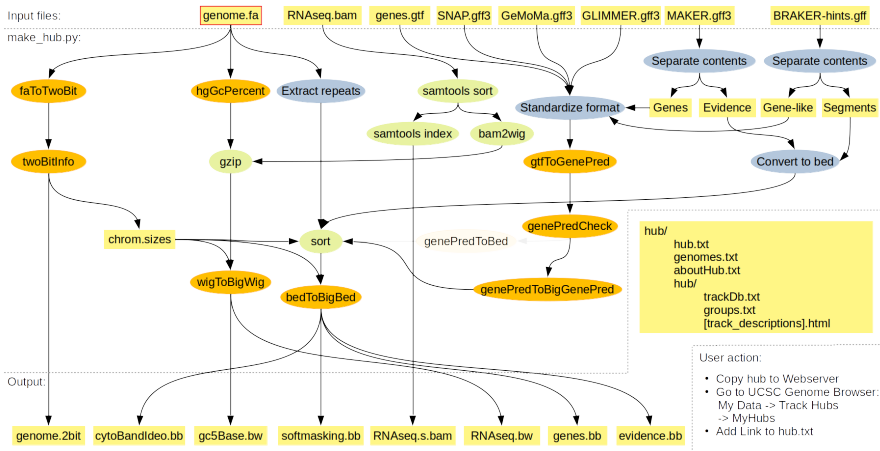
# MakeHub: A Command Line Tool

**Print Help:**

```
make_hub.py -h
```
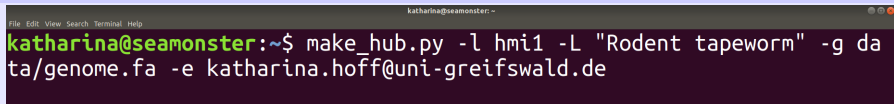
**Print Usage Examples:**

```
make_hub.py -p
```

**Required Input Parameters:**

```
make_hub.py -e EMAIL
            -l SHORT_LABEL
            -L LONG_LABEL
            -g GENOME
```

**Example: Make Simple Hub**



```
katharina@seamonster:~$ make_hub.py -l hmi1 -L "Rodent tapeworm" -g da
ta/genome.fa -e katharina.hoff@uni-greifswald.de
```

```
...
Hub is ready, please copy to a web server, e.g.
"scp -r ./hmi1 user@server:/target/location"
Feed the link to hub.txt to UCSC genome browser:
        My Data -> Track Hubs -> MyHubs
```

Hosted at:

```
http://augustus.uni-greifswald.de/bioinf/makehub/examples/
hmi1/hub.txt
```

# MakeHub: A Command Line Tool

## Example: Make Simple Hub



```
katharina@seamonster:~$ make_hub.py -l hmi1 -L "Rodent tapeworm" -g da
ta/genome.fa -e katharina.hoff@uni-greifswald.de
```

## Resulting Simple Hub

# MakeHub: A Command Line Tool

**Options for RNA-Seq**
```
-b BAM [BAM ...]    → wiggle track
-d                  → bam track
-i HINTS            → (RNA-Seq) hints
```

**Parallelization option:**
```
-c CORES
```

**Gene annotation input options:**
```
-X BRAKER_OUT_DIR    → all BRAKER gene prediction & hints tracks
-M MAKER_GFF         → all MAKER gene predicton & hints tracks
-E GEMOMA_FILTERED_PREDICTIONS    -w AUG_AB_INITIO
-m GENEMARK                       -y AUG_AB_INITIO_UTR
-I GLIMMER_GFF                    -x AUG_HINTS
-S SNAP_GFF                       -z AUG_HINTS_UTR
```

**Options to make hub prettier:**
```
-N LATIN_NAME
-V ASSEMBLY_VERSION
```

# MakeHub: A Command Line Tool

## Example: Make Hub with Many Annotation Tracks



```
katharina@seamonster:~/git/MakeHub$ make_hub.py -l hmi4 -L "Rodent tap
eworm" -g data/genome.fa -e katharina.hoff@uni-greifswald.de -a data/a
nnot.gtf -b data/rnaseq.bam -d -X data -M data/maker.gff -E data/gemom
a.gff -I data/glimmer.gff -S data/snap.gff -N "Hymenolepsis microstoma
" -V GCA_000469805.2
...
Hub is ready, please copy to a web server, e.g.
"scp -r ./hmi4 user@server:/target/location"
Feed the link to hub.txt to UCSC genome browser:
        My Data -> Track Hubs -> MyHubs
```

- 9 gene prediction tracks
- 11 evidence tracks
- Hosted at:
  `http://augustus.uni-greifswald.de/bioinf/makehub/examples/hmi4/hub.txt`

# MakeHub: A Command Line Tool

## Example: Make Hub with Many Annotation Tracks

**MakeHub
Creating Individual
Assembly Hubs for
Display with the
UCSC Genome
Browser**

**Katharina J. Hoff**

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

## Connecting Your Hub

1. copy hub to a web server
2. remember the hyperlink to `hub.txt` file on web server
3. go to `https://genome.ucsc.edu/`
4. click `My Data`
5. click `Track Hubs`
6. click `My Hubs`
7. insert link to `hub.txt`
8. click `Add Hub`
9. ... browse happily ever after

**Connecting Your Hub**

**MakeHub
Creating Individual
Assembly Hubs for
Display with the
UCSC Genome
Browser**

**Katharina J. Hoff**

Introduction

Pipeline

Connecting Your Hub

References

# Connecting Your Hub



Insert link to your custom
Track Data Hub

# Connecting Your Hub

**MakeHub
Creating Individual
Assembly Hubs for
Display with the
UCSC Genome
Browser**

**Katharina J. Hoff**

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

1.13

# References

- Raney BJ et al. (2014) Track Data Hubs.
- Kent WJ et al. (2002) UCSC Genome Browser.
- Hoff KJ (2019) MakeHub: Fully automated generation of UCSC Genome Browser Assembly Hubs
- Li H et al. (2009) The sequence alignment/map format and SAMtools
- Stanke M et al. (2008) Using native and syntenically mapped cDNA alignments to improve de novo gene finding.
- Hoff KJ et al. (2015) BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS.
- Hoff KJ et al. (2019) Whole-Genome Annotation with BRAKER.
- Holt C & Yandell M (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects
- Majoros WH & Salzberg SL (2004) TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders.
- Korf I (2004) Gene finding in novel genomes.
- Keilwagen J et al. (2018) Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi.
- Lomsadze A et al. (2014) Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm.
- Bruna T et al. (2020) GeneMark-EP and -EP+: automatic eukaryotic gene prediction supported by spliced alignments.

MakeHub
Creating Individual
Assembly Hubs for
Display with the
UCSC Genome
Browser

**Katharina J. Hoff**

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456

Introduction

Pipeline

Connecting Your Hub

References

## Acknowledgements

Mario Stanke
Malte Wellnitz
Matthis Ebel
UCSC Genome Browser Group

**MakeHub is Available for Download at**

`https://github.com/Gaius-Augustus/MakeHub`