

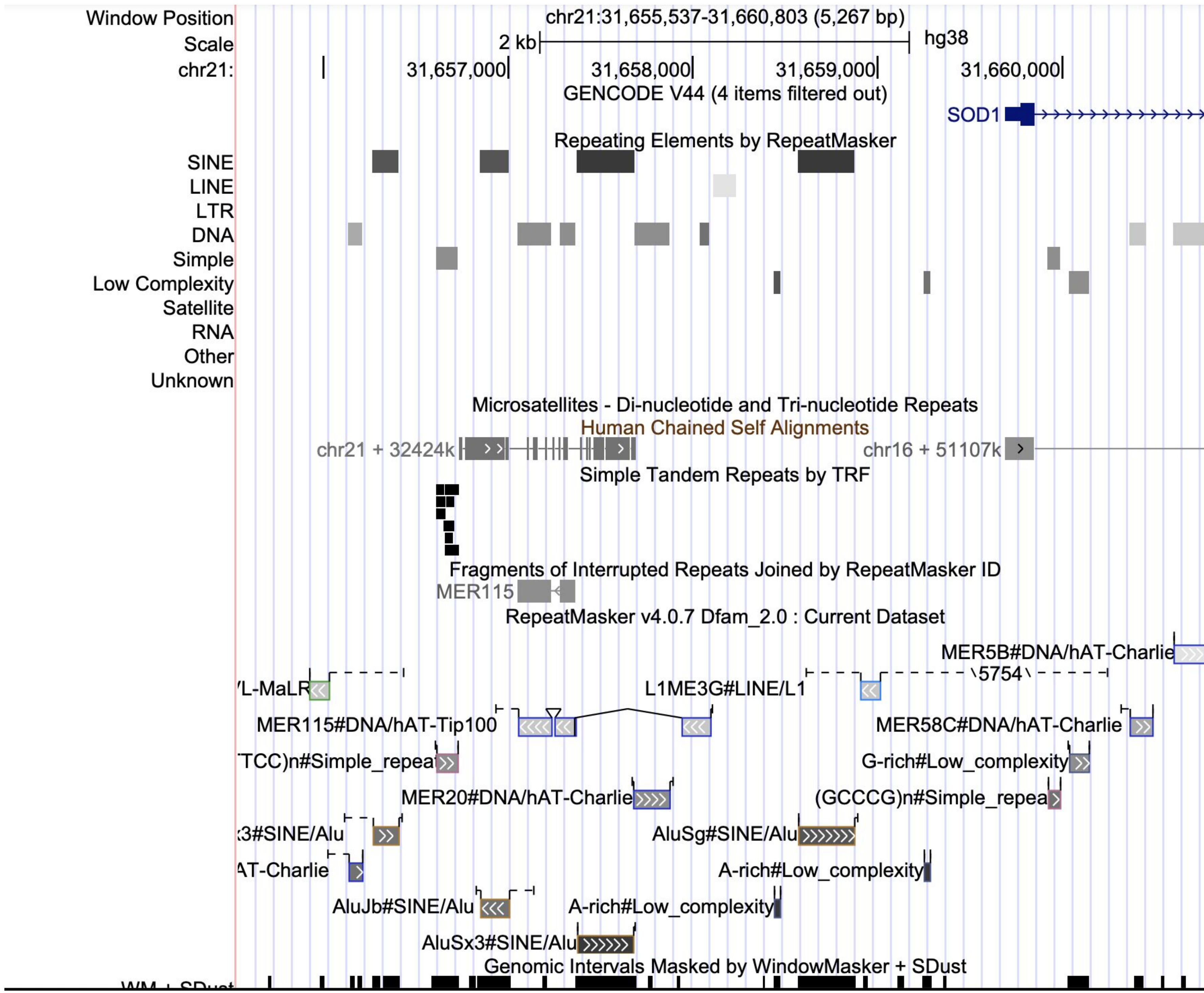
Visualizing transposable elements with the UCSC Genome Browser

Maximilian Haeussler¹, Luis R Nassar¹, Galt P Barber¹, Jonathan Casper¹, Clay Fischer¹, Jairo Navarro Gonzalez¹, Angie S Hinrichs¹, Christopher M Lee¹, Gerardo Perez¹, Brian J Raney¹, Hiram Clawson¹, Arian Smith², Robert Hubley²
¹University of California Santa Cruz, Genomics Institute, Santa Cruz, CA, USA ² Institute of Systems Biology, 401 Terry Avenue North, Seattle, WA, USA

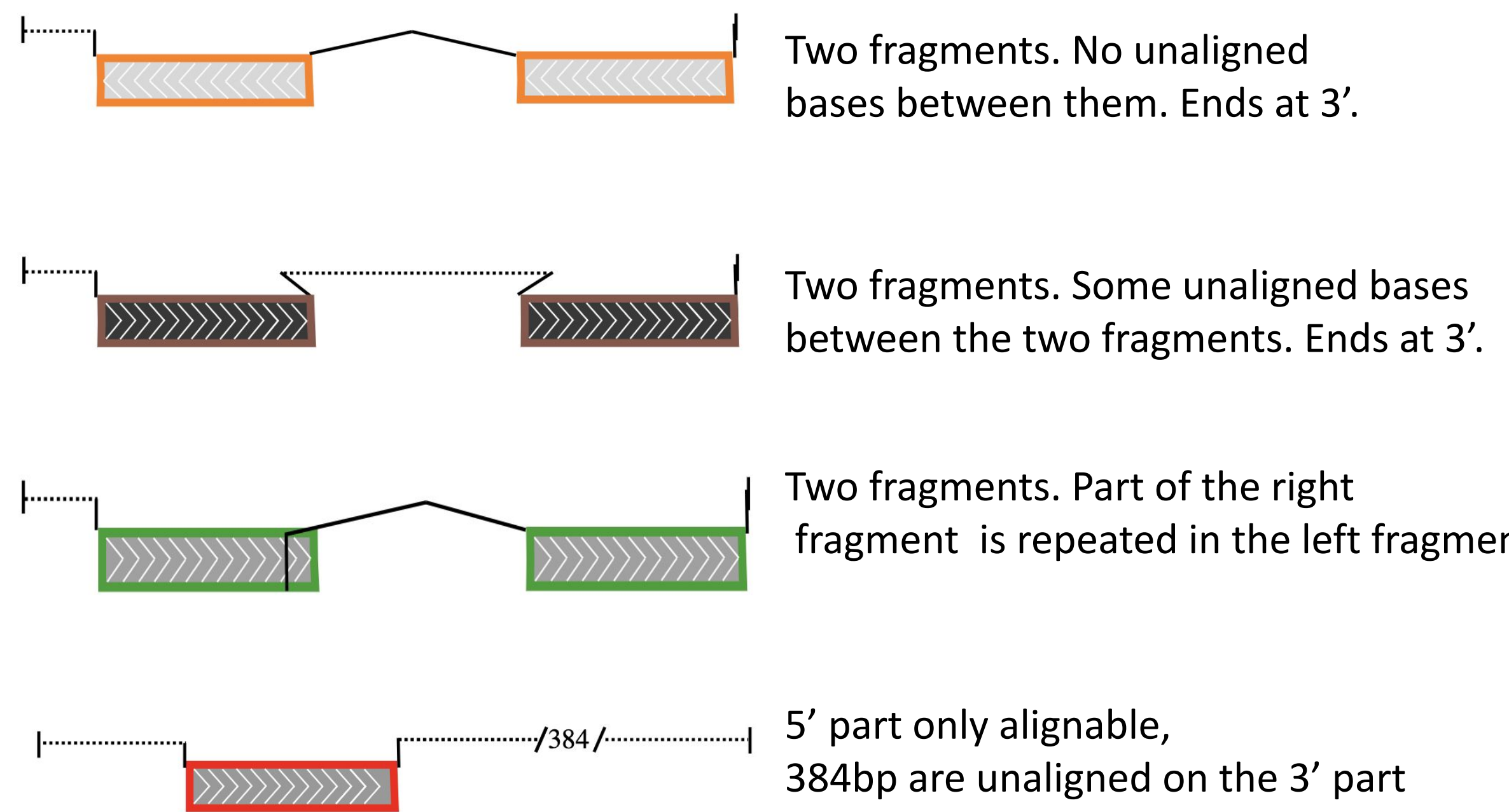
Visualizing transposable elements

- The UCSC Genome Browser's **new Repeat Viz display** provides a more intuitive visualization of repeats. New features include:
- All aligning consensus fragments are connected by lines that indicate if the two query parts have unalignable sequence between them
 - Unaligned sequence lengths at the 5' or 3' side are indicated
 - Color indicates the repeat class

- As before, the Browser has tracks for:
- Self-chain: alignment of human genome to itself, to find non-repeat copies of sequence
 - WindowMaster, TRF finder for simple repeats
 - Microsatellites track



“Repeatmasker Viz.” track on hg38 in “Robert Hubley Display”



Color	Repeat Class
Orange	SINE - Short Interspersed Nuclear Element
Green	LINE - Long Interspersed Nuclear Element
Blue	LTR - Long Terminal Repeat
Red	DNA - DNA Transposon
Purple	Simple - Single Nucleotide Stretches and Tandem Repeats
Brown	Low_complexity - Low Complexity DNA
Pink	Satellite - Satellite Repeats
Grey	RNA - RNA Repeats (including RNA, tRNA, rRNA, snRNA, scRNA, srpRNA)
Yellow	Other - Other Repeats (including class RC - Rolling Circle)
Light Blue	Unknown - Unknown Classification

Genome Browser with most of the tracks in the “Repeats” group set to “pack”. Traditional “Repeatmasker” visualisation at the top, new display at the bottom. Session can be opened via <https://genome.ucsc.edu/s/Max/hubley>

All figures copied from the track documentation page of the track “RepeatMasker Viz.” on hg38

GenArk: Genomic Archive

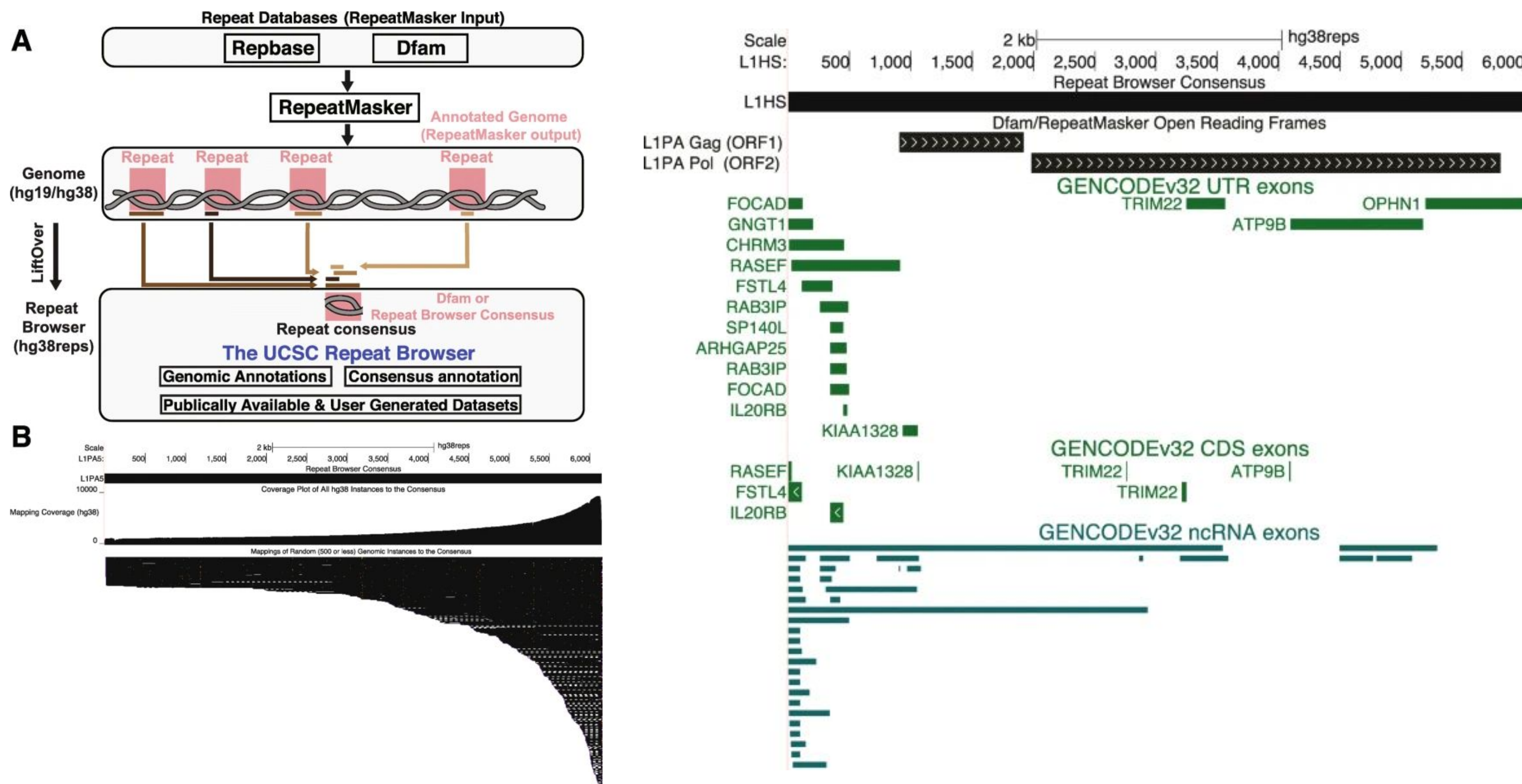
GenArk (<https://hgdownload.soe.ucsc.edu/hubs/>) is a collection of genome browser assemblies from NCBI resources “RefSeq” and “GenBank”. It currently includes over **3,885 genomes** ready to display. **All genomes come with RepeatModeler and RepeatMasker results.** Your assembly is not there? **Use the Assembly request page** <https://genome.ucsc.edu/assemblyRequest.html>

hub gateway	description
primates	NCBI primate genomes (152 assemblies)
mammals	NCBI mammal genomes (398 assemblies)
birds	NCBI bird genomes (264 assemblies)
fishes	NCBI fish genomes (260 assemblies)
vertebrate	NCBI other vertebrate genomes (111 assemblies)
invertebrate	NCBI invertebrate genomes (530 assemblies)
fungi	NCBI fungi genomes (537 assemblies)
plants	NCBI plant genomes (226 assemblies)
viral	NCBI virus genomes (256 assemblies)
bacteria	NCBI bacteria genomes (64 assemblies)
collections below are subsets of the assemblies above	
VGP	Vertebrate Genomes Project collection (435 assemblies)
CCGP	The California Conservation Genomics Project (126 assemblies)
HPRC	Human Pangenome Reference Consortium (96 assemblies)
globalReference	Global Human Reference genomes, January 2020 (10 assemblies)
mouseStrains	16 mouse strain assembly and track hub, May 2017
legacy	NCBI genomes legacy/superseded by newer versions (301 assemblies)

Please note: text file [listing](#) of 3102 genome assembly hubs

See Clawson et al. Genome Biology 2023.
<https://dx.doi.org/10.1186/s13059-023-03057-x>

A Repeat Consensus Browser at <https://repeatbrowser.ucsc.edu/> Various hg19 annotations are “lifted” to these consensus sequences



RepeatBrowser of L1HS consensus sequence, with annotations from hg19 “lifted” to this consensus through BLAT alignments. Hundreds of Chip-seq peaks are available as tracks. For details see Fernandes et al, Mobile DNA 2020 <https://mobilednajournal.biomedcentral.com/articles/10.1186/s13100-020-00208-w>

More information

Questions? Email us at genome@soe.ucsc.edu

Funding provided by NHGRI [U41HG002371]

