

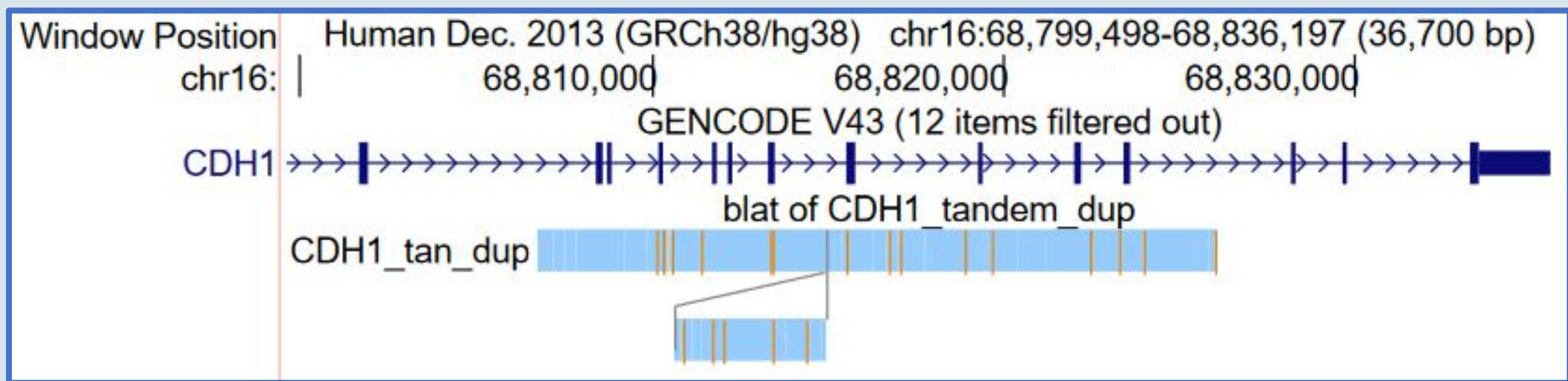
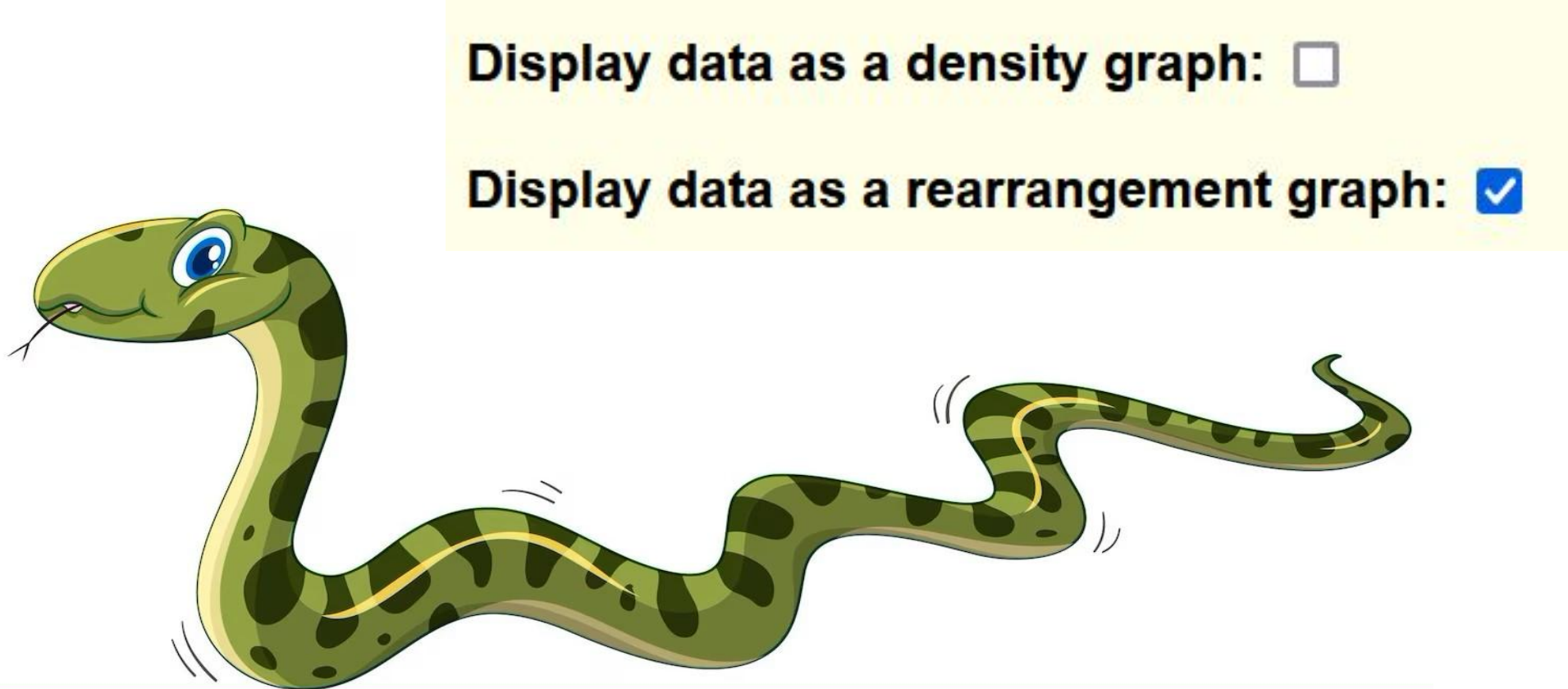
Visualizing genomic rearrangements with the UCSC Genome Browser snakes display

Luis R Nassar¹, Brian J Raney¹, Galt P Barber¹, Jonathan Casper¹, Hiram Clawson¹, Mark Diekhans¹, Clay Fischer¹, Jairo Navarro Gonzalez¹, Angie S Hinrichs¹, Christopher M Lee¹, Gerardo Perez¹, Maximilian Haeussler¹, William J Kent¹
¹University of California Santa Cruz, Genomics Institute, Santa Cruz, CA

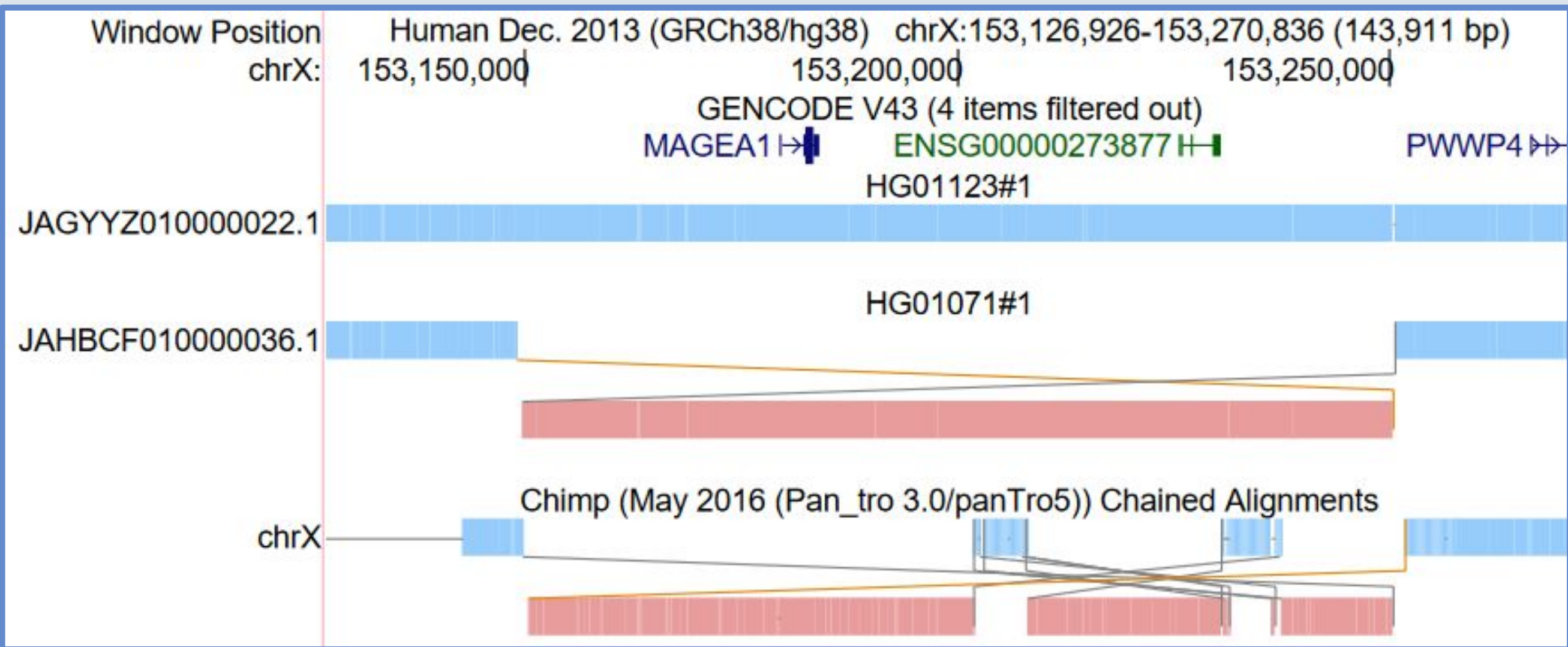
Visualizing genomic rearrangements

The UCSC Genome Browser's new **snakes rearrangements display** provides a more intuitive visualization of pairwise genome alignments that builds on its existing 'chains' display. New features include:

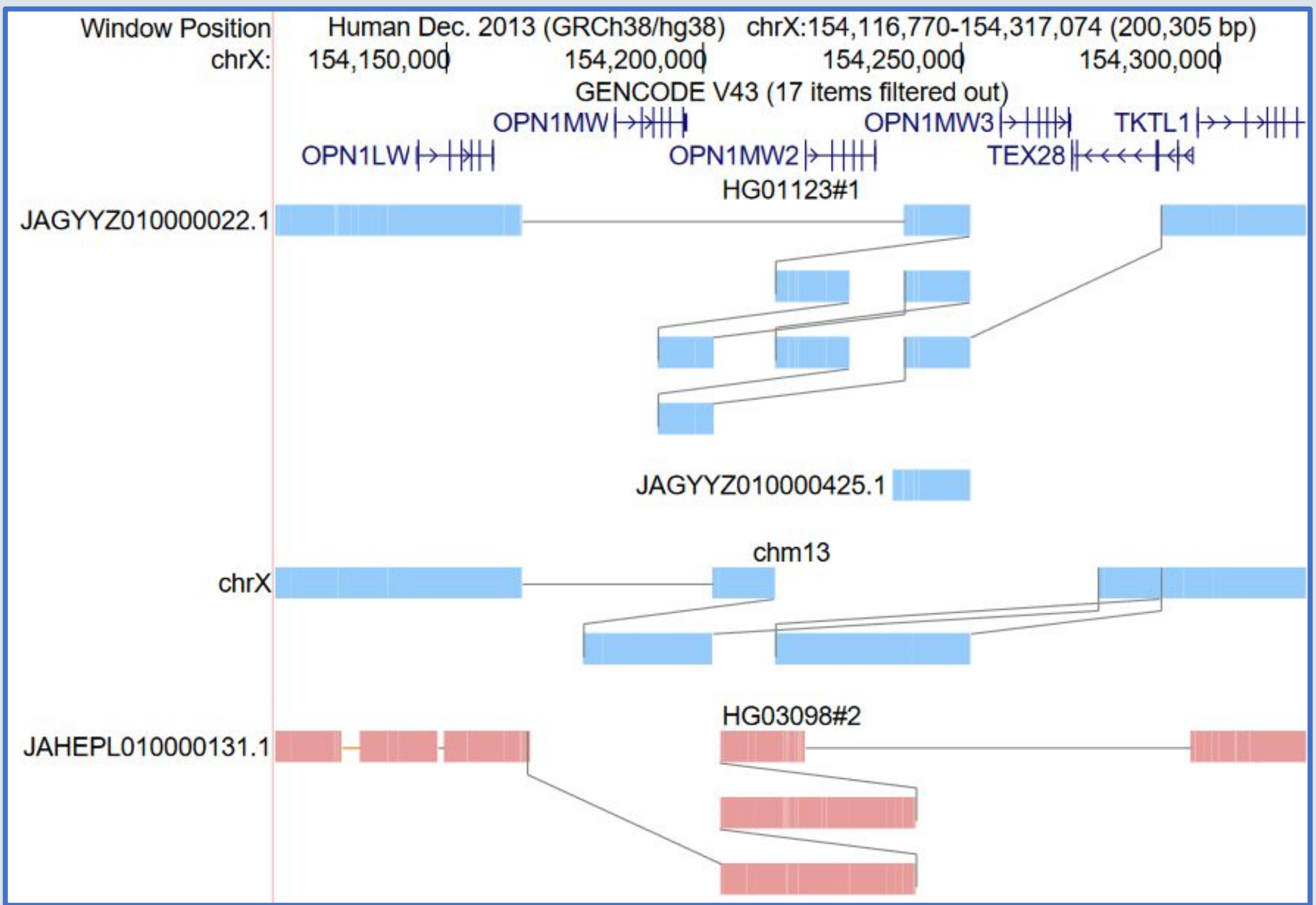
- All aligning segments are now connected by lines in the order that they appear on the target genome
- Tiles connected by **grey lines** represent **deleted or missing** DNA in target genome
- Tiles connected by **orange lines** represent **insertions** in the target genome
- Mismatched bases** are represented as **orange ticks**
- Matches on the target (+) strand are **blue**, and matches on the (-) strand are **rose**
- Snakes rearrangement display is available for chain and PSL data types, which includes Genome Browser **BLAT results**.
- This display can be switched on from the **track description page** of any eligible track type



Example 1. BLAT result of a **tandem duplication** on CDH1 which duplicates 3 exons.



Example 2. Polymorphic **inversion** in human populations that occurred after the split from chimp.



Example 3. Tandem opsin gene cluster showing different humans with varying rearrangements containing different number and order of genes, including a fusion in HG01123 (top) and inversions in HG03098 (bot). This highly variable rearrangement region on chrX can result in color blindness.

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 **3,100 genomes** ready to display.

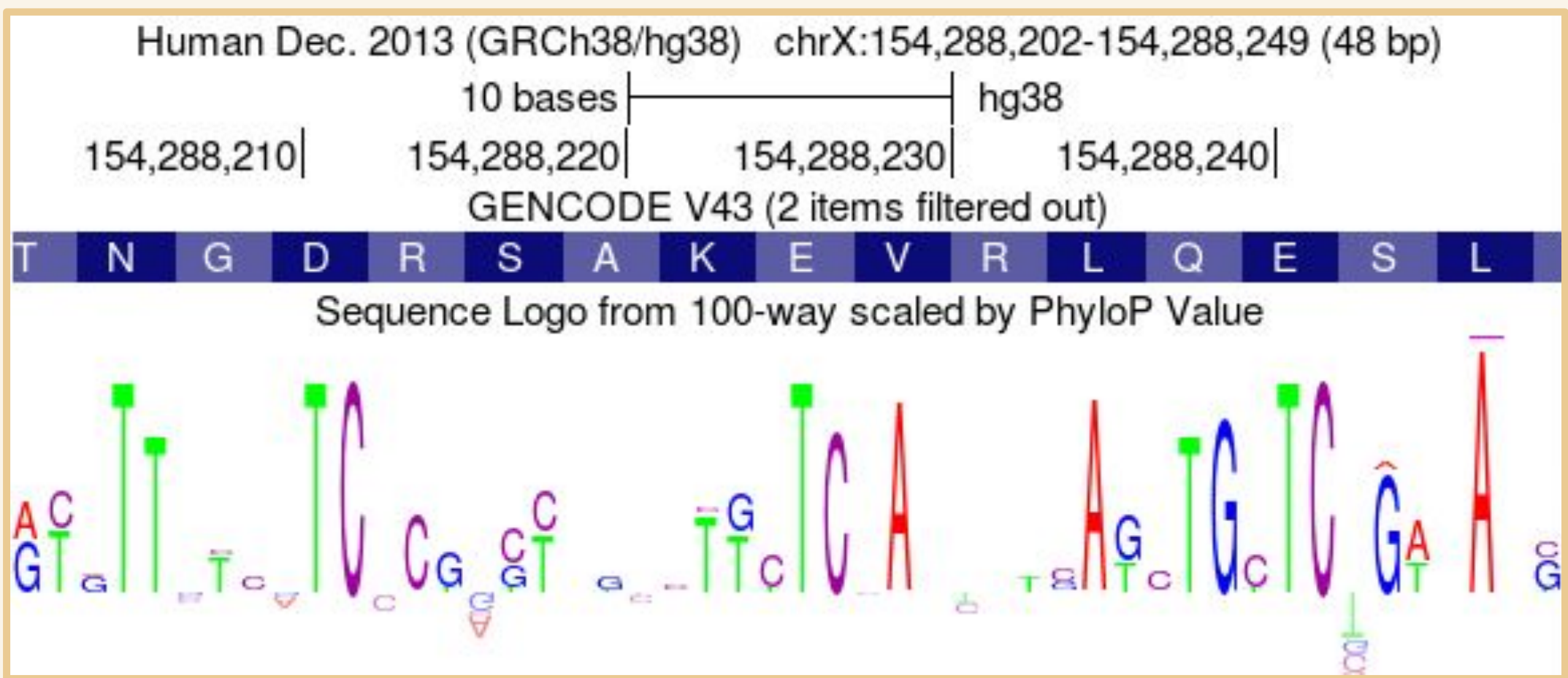
Have a genome you would like to visualize? Use our **assembly request page** (<https://genome.ucsc.edu/assemblyRequest.html>)!

GenArk: UCSC Genome Archive

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

Other new visualizations

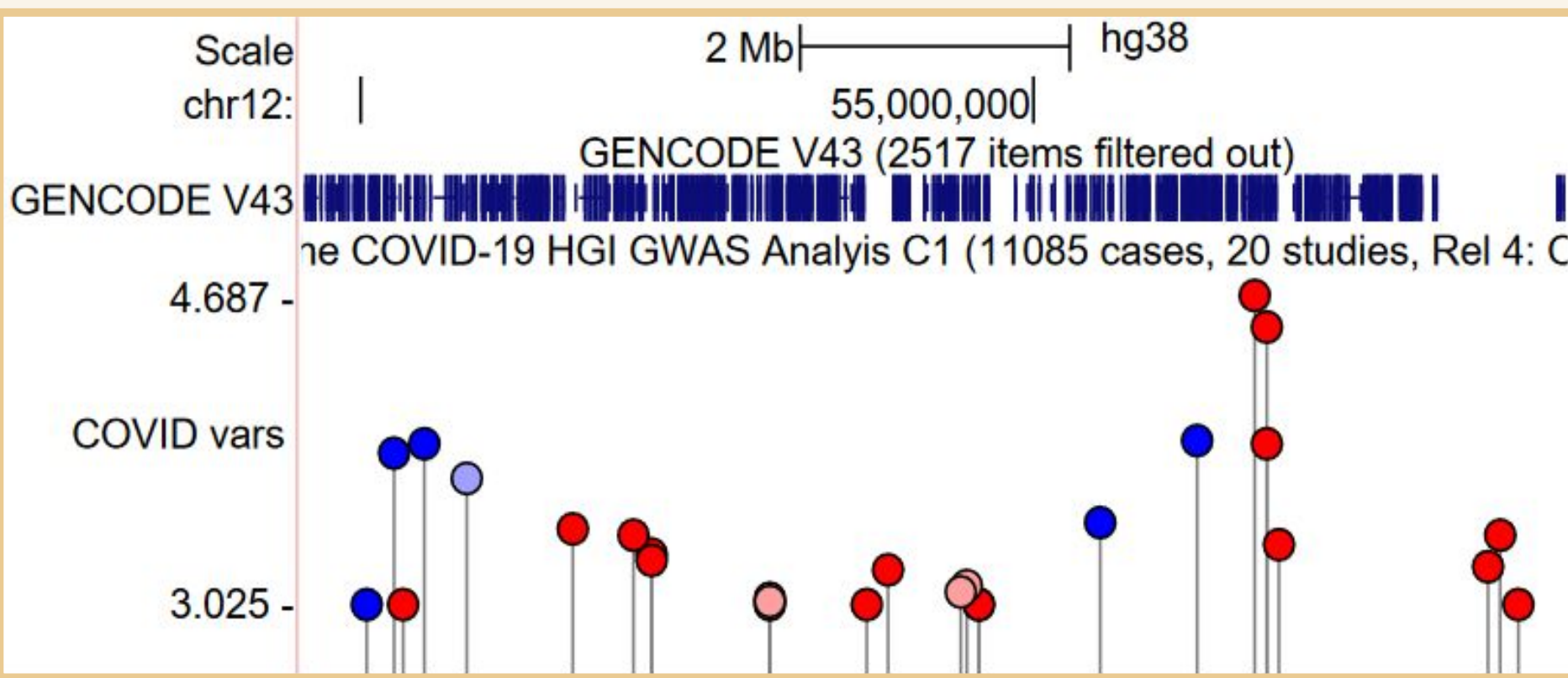


Dyseq logo display. Dyseq display scales nucleotide characters by user-specified, base-resolution scores and was developed by the Kundaje Lab. Frequently used to convey nucleotide conservation.

Learn more: <https://genome.ucsc.edu/goldenpath/help/bigWig.html#Ex4>

bigLolly display. The bigLolly (lollipop) display allows for customization of height, color, and radius of the item. Frequently used in variant display.

Learn more:
<https://genome.ucsc.edu/goldenPath/help/bigLolly.html>



More information

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

CSHL: Biology of Genomes 2023
Funding provided by NHGRI
[U41HG002371]



UNIVERSITY OF CALIFORNIA
SANTA CRUZ Genomics
Institute