GRC: Genome Reference Consortium

Coordinating and updating the assembly and annotation of complex model genomes

https://genomereference.org/ & https://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

GRC: Organization And Function

The Genome Reference Consortium (GRC) is an international collaboration consisting of NCBI, the Wellcome Trust Sanger Institute, the McDonnell Genome Institute at Washington University, the European Bioinformatics Institute (EBI), and the Zebrafish Model Organism Database (ZFIN). The consortium is tasked with ensuring that the reference assemblies for human, mouse, zebrafish and chicken continue to grow as our understanding of these genomes evolve. In order to accomplish this, the GRC uses an assembly model that enables the inclusion of variant sequences in the reference assembly [1]. In this model, "alternate loci" are provided for complex genomic regions. Alternate loci are included in the assembly as accessioned scaffold sequences with an alignment to the corresponding reference chromosome, which provides their chromosome context. The alternate loci therefore represent genomic variation relative to the reference chromosomes. As a result, the reference assembly is not just a linear assortment of haploid chromosomes, but can provide a multi-allelic representation of the genome.

[1] Church DM, et. al. Modernizing reference genome assemblies. PLoS Biol. 2011 Jul;9(7):e1001091.

Curation Efforts And Sequence Access

The GRC site (A) includes organism-specific pages to allow easy access to information on recent releases, assembly statistics, and genome regions under review through links and tabs. For example, in the Human Genome Overview (B), you can access the sequences through FTP using links provided at the top (C). The "Human Genome Issues page (D) sums up issues by chromosome with regions affected indicated by arrows (E). The table below lists issues and their resolution status (F). Filters in the left-hand column (G) narrow down the list by gene, region, and issue types.
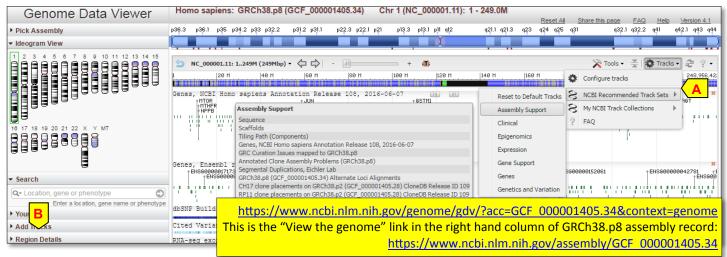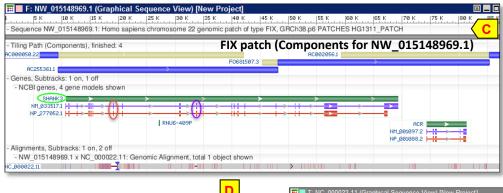
Curation Efforts And Sequence Access (cont.)

The GRC also provides genome browser tracks that describe GRC curation efforts and details of assembly construction and quality. In NCBI genome browsers, these are available via the "Assembly Support" track set in the "Tracks" menu (**A**), or the Region Details widget (**B**). A GRC track hub is also available at UCSC and Ensembl.

https://www.ncbi.nlm.nih.gov/genome/gdv/?acc=GCF_000001405.34&context=genome
This is the "View the genome" link in the right hand column of GRCh38.p8 assembly record:
https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.34

Assembly Updates

Assembly updates are categorized as minor or major releases. Minor releases to GRC assemblies are known as patch releases and provide the most up-to-date assembly data with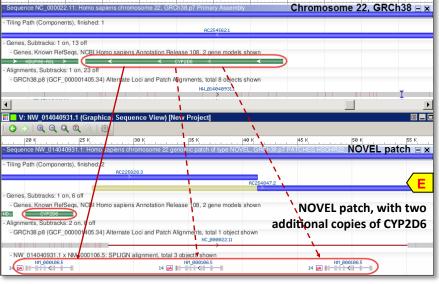out disrupting the chromosome coordinates. Like the alternate loci, patches are accessioned scaffold sequences and are given chromosome context via alignment. FIX patches (such as **C**, resulted from tiling path change in **D**) represent updates containing error corrections and sequence additions and will be incorporated in the chromosomes at the next major release. These typically include gap closures, corrections of local mis-assemblies and sequence errors.

FIX patch (Components for NW_015148969.1)

GRCh38 chr. 22 (NC_000022.11) components

FIX patch (NW_015148969.1) components

Chromosome 22, GRCh38

NOVEL patch, with two additional copies of CYP2D6

NOVEL patches (such as **E**) represent sequence variants. In the next major assembly release, the FIX patch scaffolds will be retired while the NOVEL patch scaffolds will become alternate loci scaffolds. Patch releases for an assembly are cumulative, so the latest patch release contains the sequences for all prior releases. Major releases to GRC assemblies change chromosome coordinates, occur infrequently, and are announced in advance on the GRC website.

Reporting Assembly Problems

To report assembly problems to the GRC, please use the "Report Issues" tab of the homepage. Its direct link is:
https://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/ReportAnIssue.shtml