

Resource Summary Report

Generated by SPARC SAWG on Jul 29, 2026

Human Genome Segmental Duplication Database

RRID:SCR_007728

Type: Tool

Proper Citation

Human Genome Segmental Duplication Database (RRID:SCR_007728)

Resource Information

URL: <http://projects.tcag.ca/humandup/>

Proper Citation: Human Genome Segmental Duplication Database (RRID:SCR_007728)

Description: THIS RESOURCE IS NO LONGER IN SERVICE, documented on July 17, 2013. It contains information about segmental duplications in the human genome. The criteria used to identify regions of segmental duplication are: Sequence identity of at least 90, Sequence length of at least 5 kb, Not be entirely composed of repetitive elements. Background Previous studies have suggested that recent segmental duplications, which are often involved in chromosome rearrangements underlying genomic disease, account for some 5% of the human genome. We have developed rapid computational heuristics based on BLAST analysis to detect segmental duplications, as well as regions containing potential sequence misassignments in the human genome assemblies. Results Our analysis of the June 2002 public human genome assembly revealed that 107.4 of 3,043.1 megabases (Mb) (3.53) of sequence contained segmental duplications, each with size equal or more than 5 kb and 90 identity. We have also detected that 38.9 Mb (1.28) of sequence within this assembly is likely to be involved in sequence misassignment errors. Furthermore, we have identified a significant subset (199,965 of 2,327,473 or 8.6) of single-nucleotide polymorphisms (SNPs) in the public databases that are not true SNPs but are potential paralogous sequence variants. Conclusion Using two distinct computational approaches, we have identified most of the sequences in the human genome that have undergone recent segmental duplications. Near-identical segmental duplications present a major challenge to the completion of the human genome sequence. Potential sequence misassignments detected in this study would require additional efforts to resolve. The segmental duplication data and summary statistics are available for download. Data for Human Genome (based on the May 2004 Human Genome Assembly (hg17)) Visualize duplication relationships in GBrowse (GBrowse) Duplicon Pair relationships (GFF) Genes within duplication regions (HTML) Genome duplication content (MS Excel) The segmental duplication data can be visualized in a genome browser in the GBrowse section. Selected human genome annotation tracks (except the segmental duplication track) have also been obtained from UCSC and loaded into the genome browser. Detailed information (e.g. overlapping genes, overlapping clones, detailed

alignment) can be obtained by clicking on a duplication cluster in GBrowse. Both keyword search and BLAT search are available. Analyses based on previous human genome assemblies can be found in the Previous Analyses section. Acknowledgments We thank The Centre for Applied Genomics at the Hospital for Sick Children (HSC) as well as collaborators worldwide. Supported by Genome Canada the Howard Hughes Medical Institute International Scholar Program (to S.W.S.) and the HSC Foundation.

Synonyms: Human Genome Segmental Duplication Database

Resource Type: data or information resource, database

Keywords: genes, genome, chromosome, dna, human, segmental duplication

Availability: THIS RESOURCE IS NO LONGER IN SERVICE

Resource Name: Human Genome Segmental Duplication Database

Resource ID: SCR_007728

Alternate IDs: nif-0000-02993

Record Creation Time: 20220129T080243+0000

Record Last Update: 20260729T044152+0000

Ratings and Alerts

No rating or validation information has been found for Human Genome Segmental Duplication Database.

No alerts have been found for Human Genome Segmental Duplication Database.

Data and Source Information

Source: [SciCrunch Registry](#)

Usage and Citation Metrics

We found 2 mentions in open access literature.

Listed below are recent publications. The full list is available at [SPARC SAWG](#) .

Roberti MC, et al. (2011) Complex chromosome rearrangements related 15q14 microdeletion plays a relevant role in phenotype expression and delineates a novel recurrent syndrome. Orphanet journal of rare diseases, 6, 17.

Feuk L, et al. (2005) Discovery of human inversion polymorphisms by comparative analysis of human and chimpanzee DNA sequence assemblies. PLoS genetics, 1(4), e56.