

Introducing the *NEW* Database of Genomic Variants

Thursday December 8th, 2011

http://dgvbeta.tcag.ca/dgv/app/home

The Centre for Applied Genomics (TCAG) Genome Canada Science & Innovation Centre



- I. DNA Sequencing & Synthesis including NGS (Illumina HiSeq, 5 SOLiD, Roche 454), capillary, IDT portal, modified oligos, HPLC purification
- 2. Gene Expression & Microarray -Affymetrix, Agilent, Illumina; genotyping, SNP/CNV, expression, tiling

clone libraries

3. Cytogenomics & Genome Resources karyotyping/SKY, FISH mapping, probe labeling, cDNA and

- 4. Databasing & Biobanking cell line immortalization and banking, population controls
- 5. Genetic & Statistical Analysis project design, statistical genetics, microarray analysis, pathway analysis
- **6. Informatics** project consultation, sophisticated project design supporting NGS sequencing analysis

More than 1,600 laboratories worldwide have used TCAG services for research in human and animal health, plant/forestry genomics and environment.



Overview

- Introduction to the *New* DGV
- Genome Browser
- Query Tool
- Examples of how to use the data
- Summary/Questions



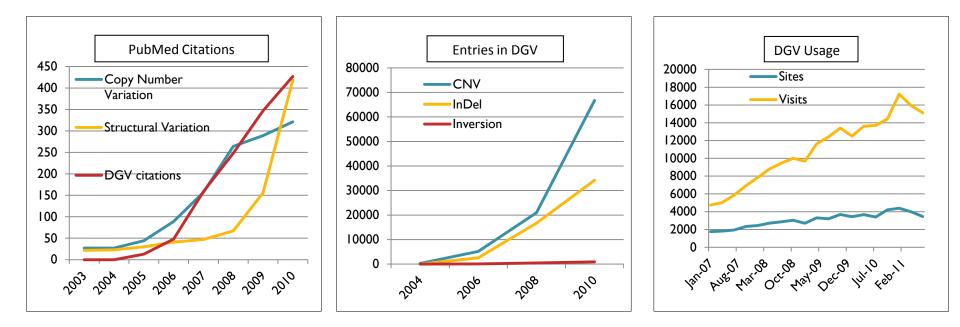
Overview

- Introduction to the *New* DGV
- Genome Browser
- Query Tool
- Examples of how to use the data
- Summary/Questions

The *New* DGV

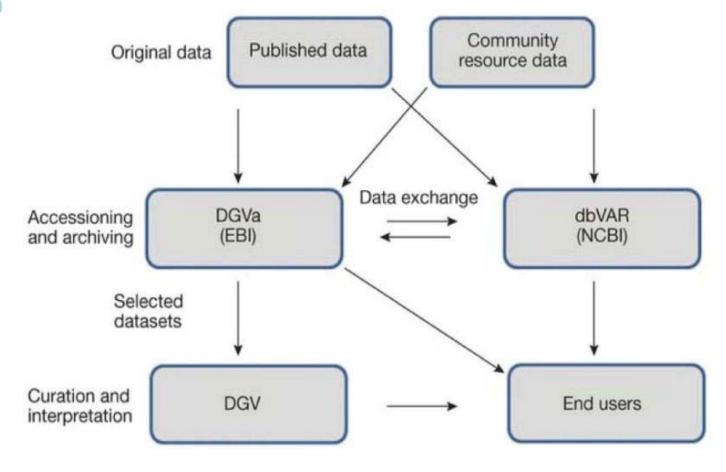
http://dgvbeta.tcag.ca/dgv/app/home

- The Database of Genomic Variants (DGV) has been working in partnership with the new database archives (DGVa and dbVar).
- In 2008, a collaboration betweenTCAG and EMBL-EBI, was established to collect, organize and curate genome-wide information on copy number variation. The goal was to expand the data curation capabilities to accommodate higher resolution array and sequencing data.



Public data archives for genomic structural variation

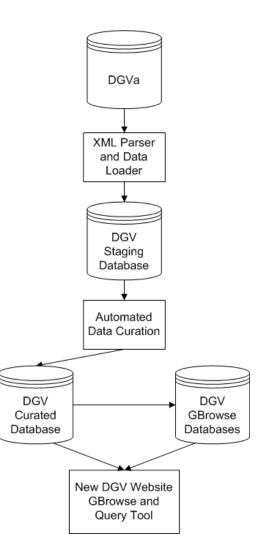
O



Church DM et. al. Nat Genet. 2010 Oct;42(10):813-4



Data Flow Diagram



Data Model/Definitions

Logical model

Project details ER diagram Entities Attributes Relationships Domains Views Procedures Sequences

List of entities Entity details Analysis Cohort **Entity: Analysis** Copy Number Coverage Cytoband Overlap Entity details: Dataset This table captures information about each Analysis that is performed in one Study. A Study may perform one or more Analyses within that Dataset Analysis Description Study. An Analysis is the comparison of the Study's Dataset to a Reference. Dataset Xref Feature Overlap Primary key constraint name PK Analysis Filter Gene Overlap Merged Analysis Attributes: Method Platform Sample Method Study Not Key Attribute name Data type Description null Platform Study Reference SERIAL This is the unique database identifer for one Analysis performed within a Study. PK Analysis ID Yes Reference Assembly This is the unique identifier from the 'Study' table and represents the Study in which an instance of an Analysis was Sample Study ID INTEGER Yes FK employed. Sample Pooled Sample Study This field captures a brief description of the Analysis employed in a Study. This information is not always provided in a Sample Xref CHARACTER No Study. Note: specific criteria used during the Analysis should be captured in the 'Analysis_Criteria' field and a general Analysis Description VARYING Study description captured in this field. Study File Prefix Supporting Merged Variant If this information is provided, the Analysis Type will be captured in this field; this information is not always provided in CHARACTER No Analysis Type VARYING a Study. An example of an Analysis Type is "split-read mapping". Translocation Mapping Variant CHARACTER This field captures the name of the Analysis Tool employed in a given Analysis, Examples of Analysis Tools are Variant Analysis Analysis Tool No VARYING "Birdsuite", "CNAG", "Genemapper". Variant Mapping Variant Type CHARACTER This is a version number for the Analysis Tool that was used in a given Study. This information is not always provided Analysis Tool Version No Variant Xref VARYING for an Analysis Tool. This is a description of the Analysis Tool that was used in a given Study. This information is not always provided for an CHARACTER No Analysis Tool Description VARYING Analysis Tool. CHARACTER This field captures whether an Analysis Tool is an Algorithm or a Software Suite. This information is not always provided Analysis Tool Type No VARYING for a given Analysis Tool in a Study; DGVa and dbVar currently do not capture this information.

Aims of DGV

- To provide information about genomic variants identified in control samples
- To serve as a resource to both clinical and research labs
- To show variation in genomic context in a simple genome browser
- To transparently provide access to underlying source data (downloadable)
- To be publicly available at no cost

Database Content

- The majority of data sets in DGV are from five different types of studies:
- Array based comparative genomic hybridization (CGH) and comparative intensity analysis (SNP/CNV arrays)
- Identification of deletions based on statistical analysis of SNP data
- Clone end sequence mapping
- Sequence trace mapping
- Next generation sequencing approaches (NGS)



What is new in the *New* DGV

- Eight new studies added
- New genome browser
- New query tool
- New database schema/backend
- New servers to improve speed/performance
- Why have we launched the new DGV as a BETA site?

New Studies Added

Author/Study	Year	PubmedID	Sample Size	Variant Count	Study Accession		
Durbin_et_al	2010	PubMedID=20981092	4719	212777	estd59		
Perry et al 2008b	2008b	PubMedID=18775914	62	791	nstd8		
	2010			4467			
Kidd et al 2010b	2010b	PubMedID=21111241	9	1167	nstd47		
Teague et al	2010	PubMedID=20534489	4	4205	nstd49		
Wilson et al	2010	PubMedID=21212237	1	2642	nstd50		
Kidd et al	2010	PubMedID=20440878	9	9825	nstd35		
Columbus et a	2010			107			
Schuster et a	2010	PubMedID=20164927	1	187	nstd39		
Alkan et al	2009	PubMedID=19718026	3	226	nstd31		



What is new in the *New* DGV

- Eight new studies added
- New genome browser
- New query tool
- New database schema/backend
- New servers to improve speed/performance
- Why have we launched the new DGV as a BETA site?

DGV Beta Home Page

http://dgvbeta.tcag.ca/dgv/app/home

 ${m D}$ atabase of ${m G}$ enomic ${m V}$ ariants

A curated catalogue of human genomic structural variation

About the ProjectDownloadsLinksStatisticsFAQGenome BrowserQuery ToolSubmissionsContact UsTutorial

Keyword, Landmark or Region Search:

BETA

Search NCBI36/hg18 -

RFTA

Examples: RP11-34P13; CFTR, 7q11.21; chr7:71890181-72690180

Find DGV Variants

by Study	by Sample
by Method	by Variant
by Platform	by Chromosome

Summary Statistics

Stat	Merged-level	Sample-level
CNVs:	136804	277913
Inversions:	521	1096
Number of	f Studies: 37	

News: October 19, 2011 The new DGV website has been launched.

Hosted by The Centre for Applied Genomics Grant support for DGV Please read the usage <u>disclaimer</u>



Download Data

The contents of the Database of Genomic Variants can be downloaded as tab delimited text files. The first line of each file is the column description. Each variant as reported in the original study is given a VariationID and this represents a stable identifier. Since variants may overlap between studies we also cluster overlapping variants into loci. Note that the loci are redefined with each database update and therefore do not represent stable identifiers. Studies where CNV coordinates are based on the NCBI 36 (hg18) assembly are not mapped back to older builds. The Build 35 download file will therefore contain less data than the Build 36 download file.

Release Versions

Release Date	Build 35 (hg17)	Build 36 (hg18)	GRCh 37 (hg19)
2011-10-17	link	link	link
2011-10-14	link	link	link
2011-10-11	link	link	link

Filtered Variants

Release Date	GFF3 File
2011-10-17	link
2011-10-14	link
2011-10-11	link

Studies included in DGV2 and links to the original data

The table below outlines the studies that have been used as sources for the data and the number of regions included from each study. Here you can also download the original dataset for studies with more than ten CNV entries.

Study Summary

how 50 ✔ entries howing 1 to 38 of 38 entries (filt	tered from 76 total entrie	es)				Co	py Print	CSV	Excel	PDF FA
lit. reference	▲ year	regions	pubmed id	assembly	÷	link to archive data		\$	comme	nts
Ahn et al 2009	2009	4091/4298	19470904	NCBI36/hg18	estd19					
Alkan et al 2009	2009	821/226	<u>19718026</u>	NCBI35/hg17	nstd31			descr this s	nber of variar iptions are m tudy but will t ext DGV upda	issing from e included in
Conrad et al 2006	2006	935/935	16327808	NCBI34/hg16	nstd17					
Conrad et al 2009	2009	72004/20206	19812545	NCBI36/hg18	estd20					
Cooper et al 2008	2008	368/368	18776910	NCBI35/hg17	nstd14					
De Smith et al 2007	2007	10093/1469	17666407	NCBI35/hg17	estd24					
Durbin_et_al_2010	2010	0/173	20981092		estd59					
Durbin_et_al_2010	2010	170048/212604	20981092	NCBI36/hg18	estd59					
Giglio et al 2002	2002	1/2	12058347	NCBI34/hg16	estd50					
Gusev et al 2009	2009	209/196	18971310	NCBI36/hg18	estd49					
Hinds et al 2006	2006	100/100	16327809	NCBI35/hg17	nstd34					
lafrate et al 2004	2004	244/255	<u>15286789</u>	GRCh37/hg19	<u>nstd41</u>			patie found kept patie	study identifie nts and control in control ind and variants in nts were only also found in ples.	ols. Variants lividuals were ound in kept if they
Itsara et al 2009	2009	11166/13843	<u>19166990</u>	NCBI35/hg17	<u>nstd27</u>			samp HGDF Varia samp were	variants detec les from the l are included nts found in t les were only also found in or HGDP samp	NINDS and in DGV. he PARC kept if they the NINDS
Jakobsson et al 2008	2008	4836/1428	18288195	NCBI36/hg18	nstd30					
Kidd et al 2008	2008	18013/7458	18451855	NCBI35/hg17	nstd2					
Kidd et al 2010	2010	20/9825	20440878	NCBI36/hg18	nstd35					
Kidd et al 2010b	2010	1496/1167	21111241	NCBI36/hg18	nstd47					
Korbel et al 2007	2007	1139/1297	17901297	NCBI36/hg18	nstd16					
Levy et al 2007	2007	44/45	17803354	NCBI36/hg18	estd22					
Locke et al 2006	2006	338/388	16826518	NCBI34/hg16	nstd29					



${\mathcal D}$ atabase of ${\boldsymbol G}$ enomic ${\boldsymbol {\mathcal V}}$ ariants

A curated catalogue of human genomic structural variation

About the Project Downloads Links Statistics FAQ Genome Browser Query Tool Submissions Contact Us Tutorial Keyword, Landmark or Region Search: Search NCBI36/hg18 -

Examples: RP11-34P13; CFTR, 7q11.21; chr7:71890181-72690180

Find DGV Variants

by Study	by Sample
by Method	d by Variant
by Platfor	m by Chromosome

Summary Statistics

1erged-level	Sample-level
36804	277913
521	1096
	36804

Number of Studies: 37

News: October 19, 2011 The new DGV website has been launched.

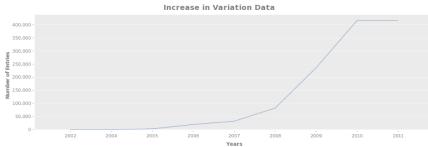
Hosted by The Centre for Applied Genomics Grant support for DGV Please read the usage <u>disclaimer</u>



ΒΕΤΑ

Content Growth

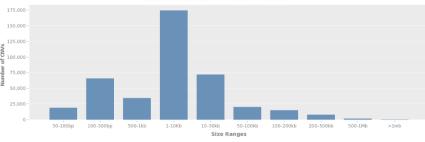
This graph shows the increase in published structural variation data that have been added to the database since its start in 2004; the numbers reflect the year of publication.



Size distribution (CNV)

The graph displays the size distribution of CNVs in the database.

Size Distribution of CNVs in DGV2



Coverage (CNV):

The table shows the data underlying the graph showing the non-redundant CNV coverage per chromosome. Note that there are several biases going into the underlying data. still a lot of smaller CNVs (<30kb) that remain to be identified.

Chr	Nucleotides Covered	Chromosome Length	% Coverage
chr1	121227647	247249719	49.03
chr2	125984828	242951149	51.86
chr3	108996479	199501827	54.63
chr4	100193894	191273063	52.38
chr5	91036041	180857866	50.34
chr6	82312288	170899992	48.16
chr7	101083954	158821424	63.65
chr8	75697497	146274826	51.75
chr9	73156027	140273252	52.15
chr10	76578489	135374737	56.57
chr11	70474623	134452384	52.42
chr12	69013530	132349534	52.14
chr13	52173934	114142980	45.71
chr14	47038666	106368585	44.22
chr15	50921933	100338915	50.75
chr16	51184213	88827254	57.62
chr17		78774742	63.43
chr18	39134218	76117153	51.41
chr19	37979515	63811651	59.52
chr20	32418199	62435964	51.92
chr21	19108870	46944323	40.71
chr22	24255876	49691432	48.81
chrX	102696418	154913754	66.29
chrY	20698862	57772954	35.83
Total	1623335318	3080419480	52.7

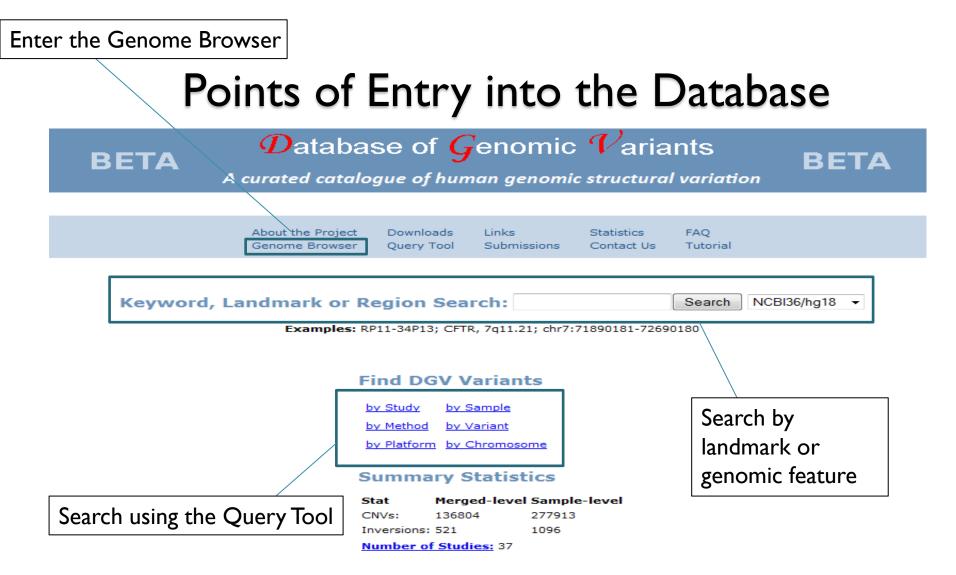
Overlap:

The following table shows the overlaps of CNVs with relevant genomic features. It is important to remember that many CNVs currently have overestimated boundaries, which category includes genes that are represented in the OMIM Morbid Map.

Feature	# of Features Overlapped by CNVs	# of CNVs Overlapped by Features
Transcripts (39426)	32797 (83.19%)	381752 (45.93%)
Exons (382693)	229150 (59.88%)	185170 (22.28%)
microRNA (1120)	713 (63.66%)	6734 (0.81%)
OMIM (2785)	2399 (86.14%)	62358 (7.5%)
Segmental Duplications (51809)	42108 (81.28%)	167112 (20.11%)

Summary of Filtered Variants

Filtered Summary									
Show 50 🗸 entries Showing 1 to 33 of 33 entries				Cop	y Print	CSV	Excel	PDF	FAQ
study 🔺		number of filtered variants	\$	reasons	or filtering				0
Ahn et al 2009	8596		Inferred variant Variant is too sr Variant had all ii			ered out			
Conrad et al 2006	1870		Inferred variant Reported merge	d variant has had all its supporting	variants filt	ered out			
Conrad et al 2009	94		Overlaps 50% o Overlaps 50% o Reported merge	of a decipher region. (hg18) of a decipher region. (hg19) of a gap region. (hg19) di variant has had all its supporting ant had its parent variant filtered.	variants filt	ered out			
Cooper et al 2008	736		Inferred variant Reported merge	d variant has had all its supporting	variants filt	ered out			
De Smith et al 2007	388		Overlaps 50% o Supporting varia Variant is too sr Variant had all i Reported merge Supporting varia	of a decipher region. (hg18) of a decipher region. (hg19) ant had its parent variant mapping mall (< 50 bp) Its reported mappings filtered out d variant has had all its supporting ant had its parent variant filtered. irge (> 300000 bp)		ered out			
Durbin_et_al_2010	2270079		Overlaps 50% o Overlaps 50% o Inferred variant Supporting varia Variant is too sr Overlaps 50% o Variant had all i Reported merge Supporting varia	ant had its parent variant mapping		ered out			



News: October 19, 2011 The new DGV website has been launched.

Hosted by The Centre for Applied Genomics Grant support for DGV Please read the usage <u>disclaimer</u>





Overview

- Introduction to the *New* DGV
- Genome Browser
- Query Tool
- Examples of how to use the data
- Summary/Questions

Genome Browser

Many features of the genome browser are the same, with some functional differences and some new tools/options.

	BETA D atabase of G	Genomic Variants BETA
	A curated catalogue of hur	man genomic structural variation
File 🕆 Help 👻		
Genomic Variants in Human Genome (Build 36: Mar. 2006, hg18) Browser Select Tracks Custom Tracks Preferences	: 800 kbp from chr7:71,890,18172,690,180	
Search		
Landmark or Region: chr7:71,890,18172,690,11 Search Examples: chr7:7189018172690180, CFTR, AC108171.3, dgv_3535539. Data Source Genomic Variants in Human Genome (Build 36: Mar. 2006, hg18)		Scroll/Zoom: K 🤇 🚥 Show 800 kbp 💌 📫 🎽 🎾 🗆 Flip
Filter variants sample id Filter Reset		

		BETA	${\mathcal D}$ atabase of ${\mathcal G}$ enomic ${\mathcal V}$ ariants	BETA
			A curated catalogue of human genomic structural variation	
File 🔻 Help 🔻				
Bookmark this Share these tracks	n Genome (Build 36: Mar. 2	2006, hg18): 800 kbp from	chr7:71,890,18172,690,180	
Export as Get chrom sizes	low-res PNG image editable SVG image			
Reset to defaults Examples: chr7:71890181726	GFF annotation table	535539.		
Data Source Genomic Variants in Human Ge	enome (Build 36: Mar. 2006, hg18))	Scroll/Zoom: < 🗹 💻 Show 800 kbp	🗸 🕂 🔰 🔀 🕞 Flip
Filter variants sample id • = •	+ .			

Genome Browser Track Options

Genomic Variants in Human Genome (Build 36: Mar. 2006, hg18): 50 kbp from chr7:116,976,604..117,026,603

Browser Select Tracks Custom Tracks Preferences		
<< Back to Browser Show Favorites Only 🚖	Clear All Favorites 🛱	
Tracks		
Breakpoints All on All off		
☆ Breakpoint annotations from Lam et al (Blue:Gain;Red:Loss;Indigo:Inversion;Brown:Complex) [?]	☆ □ Annotations from Conrad et al [?]	
CGH Arrays All on All off		
☆ ☐ Clones on WGTP Array [?]	😭 🖬 Cytochip V3.0 [?]	
☆	🟠 🗆 NimbleGen 385K Whole Genome Array [?]	
Chromosome All on All off		
☆ ✓ Cytogenetic Bands [?]	😭 🗆 Assembly [?]	🟠 🗆 Gap [?]
Clones All on All off		
☆ ✓ Clone [?]	😭 🖬 BAC End Pairs [?]	☆ 🛛 Fosmid End Pairs [?]
Disease All on All off		
☆ ☐ ISCA Curated clinically relevant regions [?]	☆ ✓ Disease Genes (OMIM) [?]	☆
G ISCA Clinical cytogenetic testing (Blue:Gain;Red:Loss;Brown:Complex;Black:NA) [?]	🟠 🖬 DECIPHER Genomic Disorders [?]	
Gene All on All off		
☆ ✓ RefSeq Genes [?]	☆	🟠 🗆 microRNA [?]
General All on All off		
☆ 🗆 dbRIP [?]	😭 🗆 Recombination hotspots [?]	
☆ 🗆 SNPs [?]	🟠 🗆 RepeatMasker [?]	
SNP Arrays All on All off		
☆	🟠 🗆 ILMN HumanHap 550 [?]	☆ 🗆 ILMN HumanHap 1M [?]
☆	😭 🗆 ILMN HumanHap 650Y [?]	
☆	☆ 🗆 ILMN Human 660W [?]	
Segmental Duplications All on All off		
☆	☆ ✓ UCSC segmental duplications [?]	
Study Variants All on All off		
☆ ✓ DGV Structural Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inversion;Black:Unkn	town) [?] Supporting Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inverison;Black:Unknown) [?]	☆ ✓ DGV Version 1 Structural Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inversion;Black:Unknown) [?]

Genome Browser Track Options

Genomic Variants in Human Genome (Build 36: Mar. 2006, hg18): 50 kbp from chr7:116,976,604..117,026,603

Browser Select Tracks Custom Tracks Preferer	nces	
<< Back to Browser Show Favorites Only 🚖	Clear All Favorites ☆	
Tracks Breakpoints All on		
Breakpoint annotations from Lam et al (Blue:Gain;Red:Loss;Indigo:Inversion;Brown:Complex) [?	Annotations from Conrad et al [?]	
CGH Arrays All on All off		
☆ □ Clones on WGTP Array [?]	☆ 🗆 Cytochip V3.0 [?]	
☆	☆ □ NimbleGen 385K Whole Genome Array [?]	
Chromosome All on All off		
☆ ✓ Cytogenetic Bands [?]	🟠 🗆 Assembly [?]	☆ 🗆 Gap [?]
Clones All on All off		
☆ ✓ Clone [2]	🟠 🗆 BAC End Pairs [?]	🛱 🗆 Fosmid End Pairs [?]
Disease All on All off		
SCA Curated clinically relevant regions [?]	☆ ✓ Disease Genes (OMIM) [?]	🟠 🗆 DECIPHER: Chromosomal Imbalance and Phenotype in Humans (Blue:Gain;Red:Loss;Brown:Complex;Black:NA) [?]
Glue:Gain;Red:Loss;Brown:Complex;Black:NA) [?]	☆ 🖨 DECIPHER Genomic Disorders [?]	
Gene All on All off		
☆ ✓ RetSeq Genes [?]	☆ 🗆 mRNA [?]	🟠 🗆 microRNA [?]
General All on All off		
☆	🛱 🗆 Recombination hotspots [?]	
☆	🟠 🗆 RepeatMasker [?]	
SNP Arrays All on All off		
☆	☆ 🗆 ILMN HumanHap 550 [?]	🟠 🖂 ILMN HumanHap 1M [?]
AFFY SNP Array 6.0 [?]	🟠 🗆 ILMN HumanHap 650Y [?]	
🟠 🗆 ILMN HumanHap 300 [?]	☆ 🗆 ILMN Human 660W [?]	
Segmental Duplications All on All off		
☆	☆ ✓ UCSC segmental duplications [?]	
Study Variants All on All off		
☆ ✓ DGV Structural Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inversion;Bl	lack:Unknown) [?] 🛛 Cupporting Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inverison;Bla	ack:Unknown) [?] ↓ GV Version 1 Structural Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inversion;Black:Unknown) [?]

nstd45

Study type: Curated

Submitter: ISCA Consortium

International Standard Cytogenetic Array Consortium Curated Dataset

PubMed: ISCA - Curated 2011

Name: ISCA curated clinically relevant regions for cytogenetic testing Description:

Annotated clinically relevant regions used in interpretation of cytogenetic testing

Size: 57 (sample data not available)

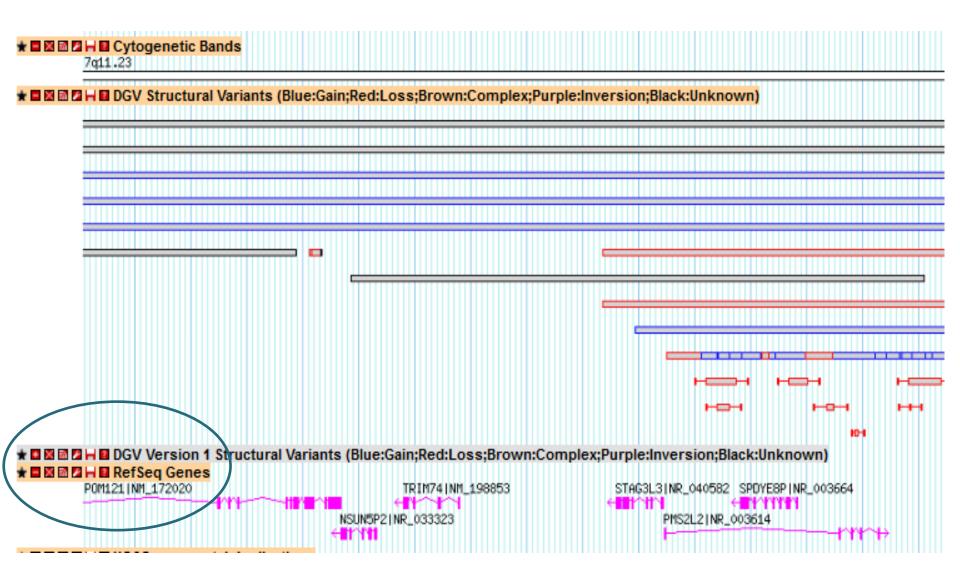
Subject Phenotype(s): Developmental Delay and additional significant developmental and morphological phenotypes referred for genetic testing

nstd37

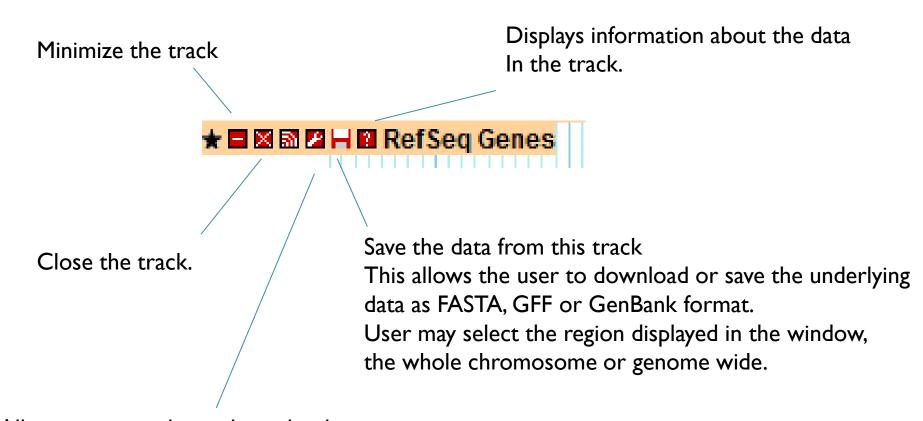
Study type: Case-Set Submitter: ISCA Consortium Description: International Standard Cytogenetic Array Consortium PubMed: ISCA 2011 dbGaP:To gain access to subject level data, see <u>dbGaP</u>. (dbGap release date: June 2011) Name: Clinical_cytogenetic_testing_referral (All) All patients referred for cytogenetic testing due to clinical phenotypes Size: 7605 (sample data available via <u>dbGaP</u>)

Subject Phenotype(s): Dev. Delay, or other clinically referred phenotypes

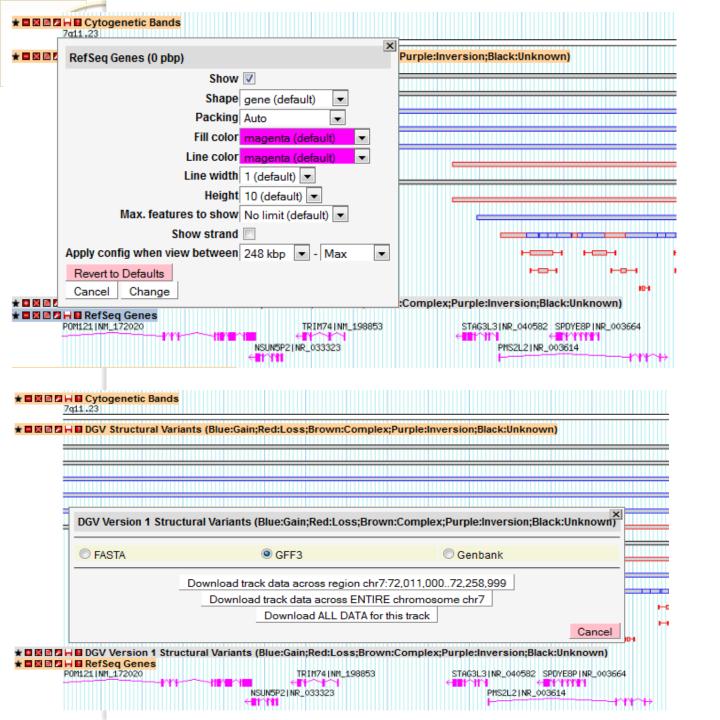
Genome Browser Track Options



Genome Browser Track Options



Allows users to change how the data are displayed.



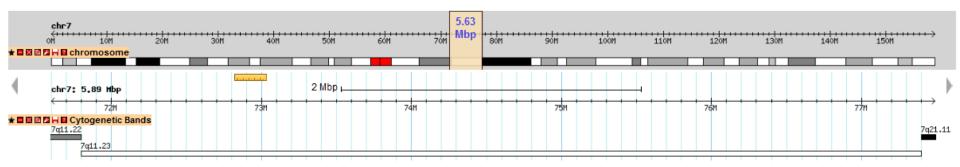
Change Display



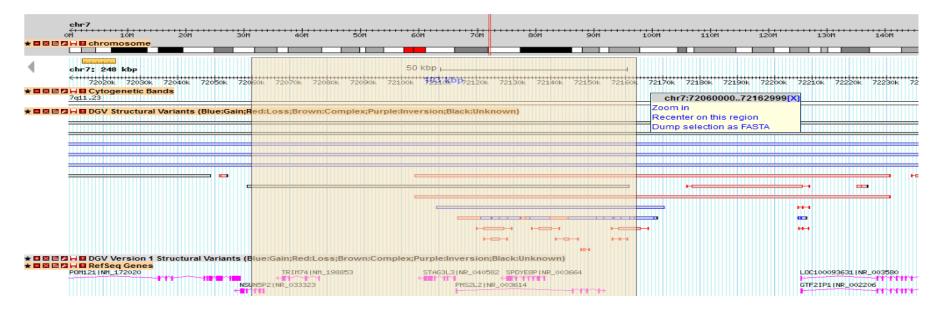
		★ 🖬 🕅 🖬 🖬 🖬 Cyt 7q11.23	ogenetic Bands										
			V Structural Vari	ants (Blue:Gain;Red:Lo	ss;Brown:	Complex;	Purple:Inve	ersion;Black:Unknown)					
		DGV Version 1 Structural Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inversion;Black:Unknown)											
		© F/	ASTA	GFI	F3			© Genbank					
				Download track data	across reg	ion chr7:7	72,011,000	72.258.999					
				Download track	data across	ENTIRE	chromoso		H				
				Down	load ALL D	ATA for th	his track		H-1				
								Cancel					
		★ ■ ₩ ■ Z + ■ DG\ ★ ■ ₩ ■ 2 + ■ Ref		ctural Variants (Blue:G	ain;Red:Los	ss;Brown	:Complex;	Purple:Inversion;Black:Unknown)					
			M_172020		TRIM74 NM_	198853		STAG3L31NR_040582 SPDYE8P1NR_003664					
				NSUN5P211	IR_033323			PMS2L2 INR_003614					
		/							IIII				
/													
##_55													
##gii- ##date	version 3	7 10:24:05 201:	1										
		e gbgff gff3 du											
		on chr7:73536000											
chr7	DGVA	merged_variant	73547601	73547601				Name=dgv_634129;ID=58156;metho					
chr7	DGVA	merged_variant		73550100				Name=dgv_634129;ID=58156;varia					
chr7	DGVA	merged_variant		73548800				Name=dgv_634129; ID=58156; varia					
chr7 chr7	DGVA DGVA	<pre>merged_variant merged_variant</pre>		73546251 73550850	•			Name=dgv_634351;ID=58368;metho Name=dgv_634351;ID=58368;varia					
chr7	DGVA DGVA	merged_variant		73548551				Name=dgv_634351;ID=58368;varia	nt_sub_type=Loss;method_na				
chr7	DGVA	sample_variant		73546851				Name=dgv_2997458; ID=333879; met	hod name=Seguencing;variar				
chr7	DGVA	sample_variant	73550050	73550050				Name=dgv_2997458;ID=333879;var	iant_sub_type=Loss;method_				
chr7	DGVA	sample_variant	73548050	73548851				Name=dgv_2997458;ID=333879;var	iant_sub_type=Loss;method_				

Genome Browser Navigation Options

I. Click and drag on the chromosome to select a particular region



2. Click and drag on the genomic position bar to zoom in or centre on a region of interest.

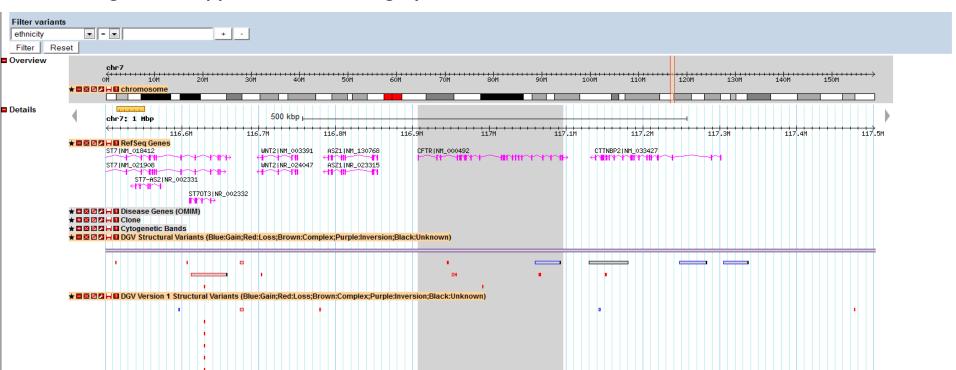


Genome Browser Visualization Options

To highlight a region of interest in the browser, select the Preferences tab, and add the target region into this box.

Genomic Variants in	n Human Genome (Build 36: Mar. 2006, h	18): 1000 kbp from chr7:116,501,604117,501,603
Browser Select Track	ks Custom Tracks Preferences	
Show grid	Image Width ◎ 600 ◎ 760 ◎ 980 ම 1240	Highlight feature(s) (feature1 feature2) Clear highlighting
Cache tracks		Highlight regions (region1:startend region2:startend) chr7:116907253117095954 Clear highlighting
Show tooltips		Region Size (bp)
		Update Appearance

The region will appear as a shaded grey box as seen below.



Genome Browser Visualization Options

	BETA \mathcal{D} atabase of \mathcal{G} enomic \mathcal{V} ariants BETA	
	A curated catalogue of human genomic structural variation	
File 🕆 Help 🔹		
Genomic Variants in Humar Browser Select Tracks Custo	n Genome (Build 36: Mar. 2006, hg18): 1000 kbp from chr7:116,501,604117,501,603	
	Image Width Highlight feature(s) (feature1 feature2)	
✓ Show grid✓ Cache tracks	© 600 © 760 © 980 ● 1240 CFTR Clear highlighting Highlight regions (region1:start.end region2:start.end) chr7:116751604117251603 Clear highlighting	
Show tooltips	Region Size (bp)	Update Appearance
For questions about the data at this	site, please contact us	
Genomic Variants in Huma	an Genome (Build 36: Mar. 2006, hg18): 1000 kbp from chr7:116,501,604117,501,603	
	tom Tracks Preferences	
Search Landmark or Region: chr7:116,501,604117,501 Searcher Examples: chr7:71890181726 Data Source		
Filter Reset	chr7 <	
Details	0f1 10f 20f 30f 40f 50f 60f 70f 80f 90f 100f 110f 120f 130f 140f 150f * B3 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
H	★ 🛙 🖾 🗗 📙 DGV Structural Variants (Blue:Gain;Red:Loss;Brown:Complex;Purple:Inversion;Black:Unknown)	

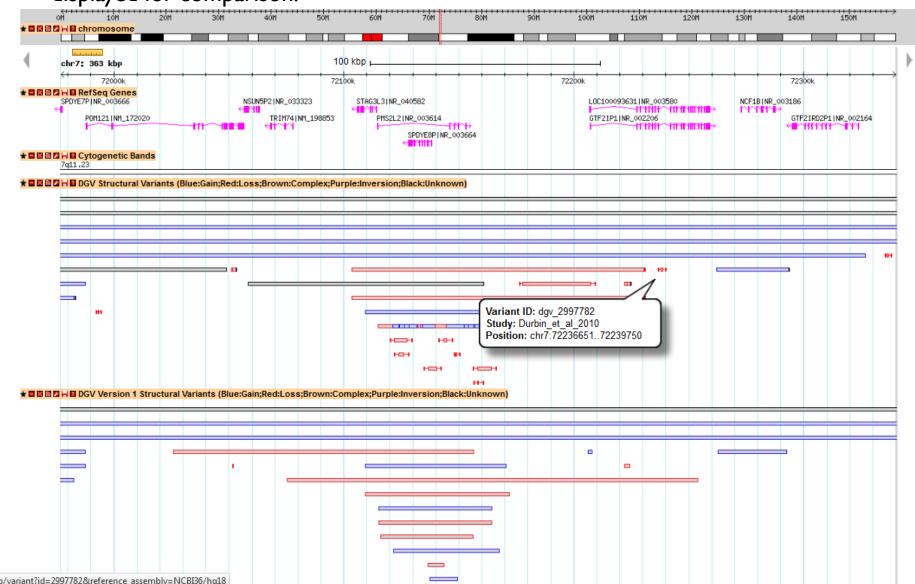


DGV Genomic Variant Data

- We have a total of 275,536 sample level variants and 137,493 merged variants in DGV from 37 different studies from a total of 5,709 samples.
- We have both the original and new data displayed as separate tracks for completeness and to allow users to compare/contrast the entries in both. We have maintained active links to the old DGV as well..
- The underlying supporting variants are also displayed in a separate track so users can visualize the contributing variants which comprise the merged calls.
- Variant types now include Inversions, CNV and Other/Complex.Variant subtypes are used to describe gain/loss/insertion/deletion.
- DGV now contains variants that are 50bp and larger and no longer differentiate InDels and CNVs.

DGV Genomic Variant Data

The original (DGV Version I) and new DGV structural variation data are included and displayed for comparison.



Variant Details

Variant: dgv_2997782

	Cytogenetic Bands 7q11.23
	DGV Structural Variants (Blue;Gain;Red;Loss;Brown;Complex;Purple;Inversion;Black;Unknown)
	DGV Version 1 Structural Variants (Blue;Gain;Red;Loss;Brown;Complex;Purple;Inversion;Black;Unknown)
	UCSC segmental duplications 210:0.9976271Lehgth:187552
	Z10:0.995369 Length:105582
	DECIPHER Genomic Disorders chr7:74254837-2284159 Williaws-Beuren Syndrowe (WBS)
	chr7:74254837-228415917q11.23 duplication syndrowe
dgv_2997782	
chr7:7223865172237750 (UCSC Ensembl)	
chr7:7223655172239750 (UCSC Ensembl)	
7011.23	
CNV	
5	
-	
Sequencing	
Analysis structural variant merging of all supporting struct	ural variant calls generated by type of computational approach: read depth analysis.
Durbin_et_al_2010	
20981092	
<u>esv82353</u>	

72238k

72239k

chr7:72236651..72239750 72237k

RefSeq Genes L0C100093631|NR_003580 GTF21P1|NR_002206 -

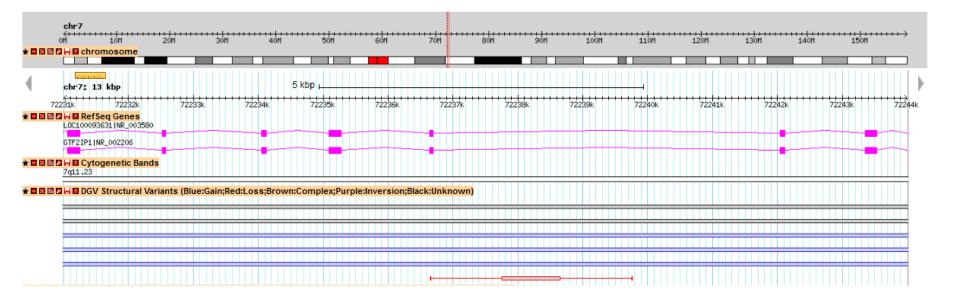
Variant ID Landmark

Cytoband Variant Type Merged Status Merged Variants Supporting Variants Samples Known Genes Method Analysis Comments Reference Pubmed ID

Genomic Position Genomic Position (Inner) Genomic Position (Outer)

External IDs Frequency

There are a few different types if images and different colours used to display structural variants in the genome browser.

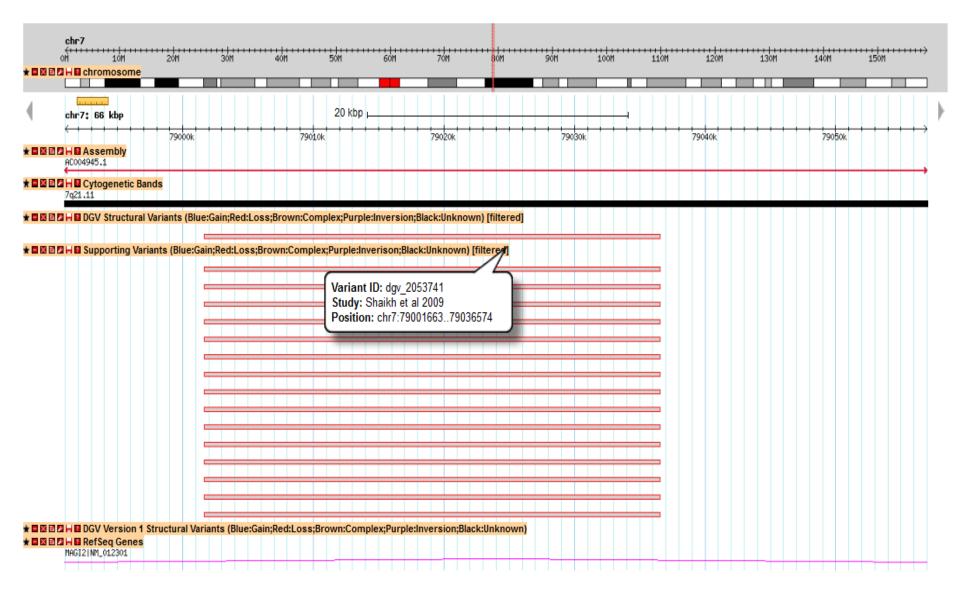


		Variant ID	dgv_380175
/ariant ID	dgv_2993496	Landmark	
andmark		Genomic Position	
Genomic Position		Genomic Position	
Genomic Position (Inner)	chr7:116381201116380950 (UCSC Ensembl)	(Inner)	
Genomic Position (Outer)	chr7:116379801116382650 (UCSC Ensembl)	Genomic Position (Outer)	<u>chr7:112147851121027925</u> (UCSC Ensembl)
Cytoband	7q31.2	Cytoband	7q31.1
Variant Type	CNV	Variant Type	OTHER
Merged Status	S	Merged Status	S

Applying Filters to DGV Variants

Genomic Variants in H	uman Ger	nome (Build	36: Mar. 2	006, hg1	8): 388 kbp	from chr	1:72,115,000	072,502,9	99										
Browser Select Tracks																			
Search Landmark or Region: chr1:72,115,00072,502.9 Examples: chr7:71890181	Search	CETR AC10817	713 day 35	35530															
Data Source Genomic Variants in Huma				•					5	Scroll/Zoon	n: < <	Show 3	88 kbp 💌	-+ 2 22	🗖 Flip				
Filter variants ethnicity	•		+ -																
Filter Reset						1													
overview	* 🖬 🖾 📾	chr1 off 10M	2011 30 some	m 40M	50M 60		80M 90M	100M		1 130M				LBOM 190M			230M		
Details	4	chr1: 388 kl					100 kbp												
		NEGR1 NM_1738	08		72200k				72300k					72400k				72500k	
	* 🗖 🖾 🖻	1p31.1	etic Bands																
	* 🖬 🔛 📾	DGV Stru	ictural Varia	nts (Blue:G	ain;Red:Loss;	Brown:Comp	lex;Purple:Inv	ersion;Black	:Unknown)										
		нн								н						нан	-	-	
		•			• =				1		н					н	•	t	
																н	н		
									· .					-					
									·										
Genomic Variants in Hui Browser Select Tracks C Search				6, hg18): (888 kbp from	n chr1:72,1	15,00072,5	02,999											
Landmark or Region: chr1:72,115,00072,502,94 Examples: chr7:7189018172 Data Source		TR, AC108171.3), dgv_35355	39.															
Genomic Variants in Human	Genome (Bu	ild 36: Mar. 200	6, hg18)	•					Scroll/Zo	om: <u> </u>	Show	388 kbp 💌] <mark>- 2 2</mark> 2	E Flip					
Filter variants	Durki	+																	
study Filter Reset	Durbin	+	-																
Overview		chr1																	
	of * = = = = *	10M 20	e		iom 60m	70M 80M	90M 100			1401		M 170M				2301 2			
Details	4	chr1: 388 kbp					100 kbp 📖										•		
	•	RefSeq Gene	es		72200k			72	300k	_	_		72400k			_	72500k		
	* 🗆 🖾 🖬 🗖	- Cytogenetic		4				_											
	1	ng31.1		Blue Gain D	ed:Loss:Brown	Complex-Pu	rnle:Inversion:	Blacktlinkpor	vn) [filtered]										
				onder Ganij K	cu.coss,browl	I I	rpicanversion;	H	M) (Intered) H						нан				
		++		81		•						н			· · · ·				
																•			
								'											

DGV Supporting Variant Data





Overview

- Introduction to the *New* DGV
- Genome Browser
- Query Tool
- Examples of how to use the data
- Summary/Questions

The Query Tool

	BETA	$\begin{array}{c} \mathbf{D} \text{atabase of } \mathcal{G} \text{enomic } \mathcal{V} \text{ariants} \\ A \text{ curated catalogue of human genomic structural variation} \end{array} \qquad \textbf{BETA} \end{array}$												
		About the Project Genome Browser	Downloads Query Tool	Links Submissions	Statistics Contact Us	FAQ Tutorial								
embly	▼ NCBI36/hg18 ▼ + -													
udy Variants Samples	Methods Platforms Analyses											Filtered	Variar	
how 50 👻 entries howing 1 to 35 of 35 entries (filter	ed from 37 total entries)							Сору	Print	CSV	Excel	PDF	FA	
accession	▲ study	\$ pub	med id	\$		sample size	;	¢		variant co	unt		_	
estd1	Redon et al 2006	17122850		2	/1			3103						
estd19	Ahn et al 2009	19470904		1				4091						
estd20	Conrad et al 2009	19812545		4	51			72004						
std21	Wheeler et al 2008	18421352		3				23						
std22	Levy et al 2007	17803354		2				44						
std24	De Smith et al 2007	17666407		5	L			10093						
std3	Wang et al 2008	18987735		1				2663						
std48	Stefansson et al 2005	15654335		1				1						
std49	Gusev et al 2009	18971310		2	70			209						
std50	Giglio et al 2002	12058347		0				1						
std55	Pinto et al 2007	17911159		7	15			974						
	Durbin_et_al_2010	20981092		1	36			170048						
	Balbin_et_al_2010			1				296						
std59	Tuzun et al 2005	15895083												
std59 std1		<u>15895083</u> <u>18776910</u>		9				368						
std59 std1 std14	Tuzun et al 2005			9				368 1139						
std59 std1 std14 std16	Tuzun et al 2005 Cooper et al 2008	18776910												
std59 std1 std14 std16 std17	Tuzun et al 2005 Cooper et al 2008 Korbel et al 2007	<u>18776910</u> <u>17901297</u>		2)			1139						
std59 std1 std14 std16 std17 std2	Tuzun et al 2005 Cooper et al 2008 Korbel et al 2007 Conrad et al 2006	<u>18776910</u> <u>17901297</u> <u>16327808</u> <u>18451855</u>		2 6 9)			1139 935						
estd59 nstd1 nstd14 nstd16 nstd17 nstd2 nstd20 nstd21	Tuzun et al 2005 Cooper et al 2008 Korbel et al 2007 Conrad et al 2006 Kidd et al 2008	<u>18776910</u> <u>17901297</u> <u>16327808</u>		2 6 9 2)			1139 935 18013						

Introduction to the Query Tool

assembly **v** = **v** NCBI36/hg18 **v** + -

Filter query Reset

Show 50 • entries Copy Pint CSV Excel PDF accession + study 0 pubmed id 0 sample size 0 variant count extd1 Redon et al 2006 17122500 271 3103 4091 500	
Showing 1 to 35 of 35 entries (likered from 37 total entries)	
accession * study 0 pubmed id 0 sample size 0 variant count edd1 Redon et al 2006 17/22250 271 3103 4091	
etd19 Ah et al 2009 19470504 1 4051 eterduery •••••• NCB136hg18 ••• •• ••••• Filter terduery Reset Copy Print CSV Excel PDF thow 50 • entries Copy Print CSV Excel PDF thow 50 • entries Copy Print CSV Excel PDF thow 50 • entries Copy of als5,556 entries (filtered from 1,963,157 total entries) Filter PDF thow 50 • entries Copy Start 0 start <td></td>	
edd 9 Ah et a 2009 19470924 1 4051	Redon et al 2006 17122850 271 3103
er query Reset	
howing 1 to 50 of 415,566 entries (filtered from 1,983,157 total entries) chromosome start choo study variant id M=Merged, S=Sample Call ethnicity gender assembly variant type variant t	Ples Methods Platforms Analyses
chromosome ★ start 0 study 0 variant id M=Merged, S=Sample Call 0 ethnicity 0 gender 0 assembly 0 variant type 0 study 0 study 0 study 0 variant type 0 study st	
1 27337400 27338705 Conrad et al 2009 dav 2691686 S YRI Female NCB136/hg18 CNV Loss 1 61855378 61856336 Durbin_et_al_2010 dav 3014505 M NCB136/hg18 CNV Loss hbly • • = • NCB136/hg18 • + • query Reset by Variants Samples Methods Platforms Analyses w 50 • entries w 10 • of 2,830 entries (filtered from 26,060 total entries) study • external sample id ¢ family id ¢ source ¢ sample description ¢ ethnicity ¢ gender ¢ cohort ne n et al 2009 SX DNA KOREAN Male	
hbly • = • NCBI36/hg18 • + • query Reset dy Variants Samples Methods Platforms Analyses w 50 • entries wing 1 to 50 of 2,830 entries (filtered from 26,060 total entries) study * external sample id ¢ family id ¢ source ¢ sample description ¢ ethnicity ¢ gender ¢ cohort na an et al 2009 SJK DNA KOREAN Male	
query Reset	61855378 61856336 Durbin_et_al_2010 <u>dqv 3014505</u> M NCBI36/hg18 CNV Loss
study external sample id family id source sample description ethnicity gender ¢ cohort na nn et al 2009 SJK DNA KOREAN Male V V V V V	
wing 1 to 50 of 2,830 entries (filtered from 26,060 total entries) study external sample id family id source sample description et al 2009 SJK DNA KOREAN Male KOREAN Male M	is Methods Platforms Analyses Filter
In et al 2009 SJK DNA KOREAN Male	
nnrad et al 2006 NA19094 DNA YORUBA Female 30 Yoruba trio	ies (filtered from 26,060 total entries)
onrad et al 2006 NA19208 DNA YORUBA Male 30 Yoruba trio	ies (filtered from 26,060 total entries) external sample id family id for source DNA korean

Introduction to the Query Tool (2)

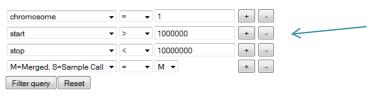
assembly • = • NCBI36/hg18 • + -

Filter query Reset

Study Variants Samples Methods Platforms Analyses								Filtered	Variants
Show 50 - entries				Сору	Print	CSV	Excel	PDF	FAQ
Showing 1 to 43 of 43 entries (filtered from 496 total entries)				0007			Excer	1.51	17102
method name	study	method type	\$		method de	scription			\$
BAC_BCGH	Locke et al 2006	Discovery	Array hybridization: use of the segmen consists of 2,007 E complex regions o duplications. All 26 experiments, to th reference individua	tal duplica ACs, spot f the geno 9 individua e segmen	tion array. ed in triplic me and fla als were hy tal duplicat	The segm ate, that with nked by in bridized, with ion array with	ental duplic were targete itrachromos ith dye-swa vith use of a	cation arra ed to 130 somal seg op replicate	ay mental
BAC_aCGH	Redon et al 2006	Discovery	Comparative hybrid total of 26,678 larg Path to cover the h screened for T1 ph finger printing and	ge insert d iuman gen age and P	ones were Iome in tilir Seudomon	selected fr ng path res	om the put solution. Cl	blished Go ones were	lden
ssembly									
Study Variants Samples Methods Platforms Analyses								Filtered	l Variants
Ahn et al 2009 Illumina Genome Analyz	orm name Image: platform type er er and Roche/454 sequencer.	e ¢ platform version	\$ geo acces	Copy	Print ≎	CSV	Excel	PDF	FAQ \$
Conrad et al 2006 Not Provided									
Conrad et al 2009 Sanger H. Sapiens Nimbl					E-M	TAB-40			
Cooper et al 2008 Illumina Human1Mv1 DN (Human1Mv1_C)	A Analysis BeadChip		GPL6983						
udy									
Study Variants Samples Methods Platforms Analyses								Filtered V	ariants
Show 50 - entries Showing 1 to 3 of 3 entries (filtered from 1,557 total entries)				Сору	Print	CSV	Excel	PDF	FAQ
study 🗘 analysis type 🗘 analysis tool	 analysis description 	reference type	reference	¢		reference	e descriptio	n	\$
Alkan et al 2009 Detection mrFAST	We constructed duplication maps for each o the three genomes and estimated the absolute copy number of each duplication interval larger than 20 kb in length. We considered a given segment to be duplicate within an individual if the median estimated	f Ref_sequence	NCBI35/hg17						

Query Tool Functionality

	iants Samples	Methods	Platforms	Analyses			Use ~	to perforn	n a wildcaı	rd searc	h				Filtered	Variants
		Reset ed from 38 f	total entries)								Сору	Print	CSV	Excel	PDF	FAQ
	accession		put	blication	¢		pubmed id	\$	sample size	e	÷		varia	nt count		÷
nstd2		к	idd et al 2008			18451855		9			74	458				
nstd35		к	idd et al 2010		2	20440878		9			98	825				
nstd47		К	idd et al 2010b		1	21111241		9			11	167				
Showing 1 to	3 of 3 entries (filtere	ed from 38 t	total entries)										First	Previous	1 Nex	: Last



Use the "+" button to add additional terms Use the "-" button to remove search terms

udy Variants	Samples	Methods	Platfor	rms Analyses											Filtered \	/ariants
iow 50 👻 entrie	es										Сору	Print	CSV	Excel	PDF	FAQ
owing 1 to 50 of 1	1,846 entries (filtered fro	om 1,983,1	57 total entries)												
chromosome		\$	stop 🗘	study 🗘	variant id	\$	M=Merged, S=Sample Call \$	ethnicity	\$ gender	assembly	\$	variant type	\$	variant	t subtype	
	14711	38 14	75687	Durbin_et_al_2010	<u>dqv 148358</u>	м				GRCh37/hg19	CNV		Lo	oss		
	18561	57 18	56942	Ahn et al 2009	<u>dqv 519875</u>	м		KOREAN	Male	NCBI36/hg18	CNV		Lo	oss		
	32156	07 32	17688	Wilson et al 2010	<u>dqv 2509135</u>	М		European caucasian	Female	GRCh37/hg19	CNV					
	39244	54 39	91153	Itsara et al 2009	dgv 702033	м		Druze	Female	NCBI35/hg17	CNV					
	13749	32 13		Schuster et al 2010	<u>dqv 3022768</u>	м		San	Male	NCBI36/hg18	CNV					
	13509	76 13	54825	Durbin_et_al_2010	<u>dqv 148078</u>	м				NCBI36/hg18	CNV		Lo	oss		
	10413	88 10	45987	Durbin_et_al_2010	<u>dqv 147697</u>	м				GRCh37/hg19	CNV		Lo	oss		
	13741	13 13	76662	Durbin_et_al_2010	<u>dqv 148260</u>	м				GRCh37/hg19	CNV		Lo	oss		
	36201	74 36	94948	Itsara et al 2009	<u>dqv 704561</u>	м		NAN Melanesia	n Female	NCBI35/hg17	CNV					
	10411	76 10	43725	Durbin_et_al_2010	<u>dqv 147834</u>	м				NCBI36/hg18	CNV		Lo	oss		
	20143	28 20	17283	Durbin_et_al_2010	<u>dqv 3014878</u>	м				NCBI36/hg18	CNV		Lo	DSS		
	32948	91 32	97640	Durbin_et_al_2010	<u>dqv 3021430</u>	м				GRCh37/hg19	CNV		Lo	oss		
	16152	56 16	15349	Ahn et al 2009	<u>dqv 518008</u>	м		KOREAN	Male	GRCh37/hg19	CNV		G	ain		
	12400	38 12	44287	Durbin_et_al_2010	<u>dqv 149215</u>	м				GRCh37/hg19	CNV		Lo	oss		

Query Tool

Users can save, copy or print the output using these options.

Study Variants S	Samples M	lethods P	Platforms A	nalyses					Filtered Varian
chromosome assembly Filter variant Filter	• ~	 Y NCBI36/hg eset 		+ -					
Show 10 👻 entries							Сор	y Print CSV	Excel PDF FAQ
Showing 1 to 10 of 1,22	9 entries (filte	ered from 1,2	201,742 total e	entries)					
chromosome 🔺	start 🗘	stop 🗘	study 🗘	variant id	\$	M=Merged, S=Sample Call \diamond ethnicity \diamond gender \diamond	assembly	variant type	variant subtype
Y	18507723	18508445	Perry et al 2008	<u>dqv 346155</u>	S	NC	CBI36/hg18	CNV	
Υ	22495042	22496257	Perry et al 2008	<u>dqv 346207</u>	s	NC	CBI36/hg18	CNV	
Y	10625601	10641271	Perry et al 2008	<u>dqv 346268</u>	s	NC	CBI36/hg18	CNV	
Υ	19153972	19168216	Perry et al 2008	<u>dqv 346288</u>	s	NC	CBI36/hg18	CNV	
Y	1895307	1901245	Perry et al 2008	<u>dqv 346327</u>	М	NC	CBI36/hg18	CNV	
Υ	57758500	57772954	Perry et al 2008	<u>dqv 346385</u>	s	NC	CBI36/hg18	CNV	
Y	24104834	24106049	Perry et al 2008	<u>dqv 346466</u>	s	NC	CBI36/hg18	CNV	
Y	10475956	10477314	Perry et al 2008	<u>dqv 346542</u>	s	NC	CBI36/hg18	CNV	
Y	22720984	22743122	Perry et al 2008	<u>dqv 346647</u>	м	NC	CBI36/hg18	CNV	
Υ	24116299	26530061	Perry et al 2008	<u>dqv 346892</u>	м	NC	CBI36/hg18	CNV	
Showing 1 to 10 of 1,22	9 entries (filte	ered from 1,2	201,742 total e	entries)				First Previous	1 2 3 4 5 Next Last



Overview

- Introduction to the *New* DGV
- Genome Browser
- Query Tool
- Examples of how to use the data
- Summary/Questions

Examples of Common Searches

•There are a few common types of searches that individuals may use to query or search the database.

- Extract only variants detected in a specific population (CEU)
- Identifying which studies have detected variants in a specific individual (NAI5510)
- Find variants located in a specific region/locus of interest.
- Locate which studies and which variants are derived from HapMap samples.
- Search for variants that were detected using a sequencing based approach.

Example One Search by sample name: NAI5510

external sample id	▼ =	▼ NA1551	0	+ -											
Filter query Reset															
Study Variants	Samples	Methods	Platforms	Analyses										Filtered V	/ariants
Show 50		ed from 26.06	30 total entries	5)					Co	ору	Print	CSV	Excel	PDF	FAQ
study 🔺	-	external sa		-, ¢	family id	source	\$ sample description	\$ ethnicity	\$	gend	er	\$	cohort	name	\$
Conrad et al 2009	NA15510					DNA				-					
Cooper et al 2008	NA15510					DNA									
Kidd et al 2008	NA15510					DNA									
Kidd et al 2010	NA15510					Cell-culture						Fosmi	d9		
Kidd et al 2010b	NA15510					Cell-culture						Fosmic	d9		
Korbel et al 2007	NA15510					DNA			Ferr	nale					
Mills et al 2006	NA15510					DNA						TSC			
Pinto et al 2007	NA15510					DNA			Male	e		HapMa	эр		
Showing 1 to 8 of 8 e	ntries (filter	ed from 26,06	50 total entries	s)								First	Previous	1 Next	Last

Select the common reference sample NA15510 using external sample ID and filter. -Results returned include a list of all the studies that have used this sample -Select the variants tab to get a list of all variants detected in NA15510

Example One (continued)

external sample id Filter query Reset	▼ = ▼ NA15510) + -												
Study Variants	Samples Methods	Platforms Analyses											Filte	ered Variants
Show 50 ✔ entrie: Showing 1 to 50 of 4,	s ,558 entries (filtered from 1	1,983,157 total entries)							C	Сору	Print	CSV	Excel PD	DF FAQ
chromosome	≎ start	≎ stop ≎	study	variant id	\$	M=Merged, S=Sample Call ♦	ethnicity	\$ gender	\$ asse	mbly		variant typ	e 🗘 varian	t subtype 🗘
1	10499	92013	Conrad et al 2009	<u>dqv 2673064</u>	s				GRCh37/hg	19	c	NV	Gain	
1	444338	449236	Conrad et al 2009	<u>dqv 2672840</u>	s				GRCh37/hg	19	c	NV	Gain	
1	449286	450511	Conrad et al 2009	<u>dqv 2673356</u>	s				GRCh37/hg	19	c	NV	Loss	
1	869415	870347	Conrad et al 2009	<u>dqv 2672675</u>	s				GRCh37/hg	19	c	NV	Gain	
1	1074379	1076019	Conrad et al 2009	<u>dqv 2673401</u>	s				GRCh37/hg	19	c	NV	Gain	
1	1223594	1225599	Conrad et al 2009	<u>dqv 2672664</u>	s				GRCh37/hg	19	c	NV	Loss	
1	1285400	1286900	Conrad et al 2009	<u>dqv 2672237</u>	S				GRCh37/hg	19	c	'NV	Gain	
1	1910377	1911917	Conrad et al 2009	<u>dqv 2672632</u>	s				GRCh37/hg	19	c	NV	Loss	
1	2409767	2450161	Kidd et al 2008	<u>dgv 372002</u>	s				GRCh37/hg	19	c	THER	Inversi	on

We now have a list of variants for NA15510 detected in multiple studies.

Example Two: Search for variants in a defined region

chromosome	▼ =	▼ 7	+
start	• >	▼ 116907253	+ -
stop	▼ <	▼ 117095954	+ -
assembly	▼ =	▼ GRCh37/hg19 ▼	+ -
Filter query Reset			

Study Variants S	amples Meth	ods Platfo	rms Analyses											Filtered Va	ariants
Show 50 👻 entries										0	Duint	001	Freed	205	54.0
Showing 1 to 8 of 8 entri	es (filtered from	n 1,983,157 to	tal entries)							Сору	Print	CSV	Excel	PDF	FAQ
chromosome	start 🗘	stop 🗘	study 🌣	variant id	\$	M=Merged, S=Sample Call 🗘	ethnicity \$	gender	assembly	\$	variant type	¢	variar	t subtype	\$
7	117018506	117063855	Kidd et al 2008	dgv 380508	S				GRCh37/hg19	CNV	1				
7	117035303	117035613	Mills et al 2006	dgv 888512	s				GRCh37/hg19	CNV	1				
7	117018506	117063855	Kidd et al 2008	<u>dqv 367649</u>	S		UTAH/MORMON	Female	GRCh37/hg19	CNV	1		Loss		
7	117081446	117086131	Shaikh et al 2009	<u>dqv 1997989</u>	S				GRCh37/hg19	CNV	1		Loss		
7	117081446	117086131	Shaikh et al 2009	<u>dqv 2003874</u>	S				GRCh37/hg19	CNV	1		Loss		
7	116915618	116915869	Durbin_et_al_2010	<u>dqv 2994296</u>	s				GRCh37/hg19	CNV	1		Loss		
7	117035303	117035613	Mills et al 2006	<u>dqv 1110850</u>	s				GRCh37/hg19	CNV	1		Loss		
7	117081446	117086131	Shaikh et al 2009	<u>dqv 2051043</u>	М				GRCh37/hg19	CNV	1				
Showing 1 to 8 of 8 entri	es (filtered from	n 1,983,157 to	tal entries)									First	Previous	1 Next	Last

Compare a list of variants in DGV to a region defined in your study or in a clinical array report. -Can obtain a list of all variants across studies that fall within a defined region

-Can use this for classifying potential rare/de-novo variants, or as evidence to suggest your variant of interest may or may not be pathogenic.

(Note: it is very important to only use DGV data as a guide. The content is still limited and should not replace proper experimental controls.)



Overview

- Introduction to the *New* DGV
- Genome Browser
- Query Tool
- Examples of how to use the data
- Summary/Questions

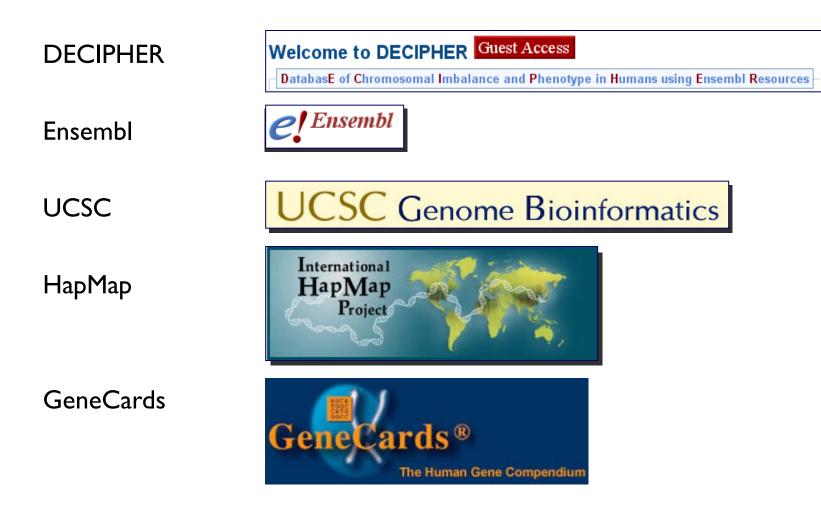


Summary/Questions

- Launched a fully integrated, accessioned and archived database with several new studies including the 1,000 Genomes.
- Introduced a new genome browser and implemented an interactive query tool for improved access to the data.
- Improved the performance and function of the database, while maintaining the user-friendly interface and community database design.

DGV Inter-Operates With Other Genome Databases

• Other sites displaying data from DGV:





Help and Support

http://dgvbeta.tcag.ca/dgv/app/contacts

If you have any questions while using the DGV beta site, or if you notice any errors or bugs in the database, please contact the DGV Team at your earliest convenience.

Email: <u>dgv-contact@sickkids.ca</u>

If you would like to receive updates and notifications about DGV, please sign up for our newsletter

Thank you for taking the time to participate in the webinar today, your support and feedback is greatly appreciated.

Sincerely, The DGV Team

DGV Website: http://dgvbeta.tcag.ca/dgv/app/home