

2008-08-07



User manual version 1.0

Contents

1. System of VarySysDB

1.1 VaryGene2

1.2 LD Search system

1.3 Genome Browser (GBrowse)

2. Acknowledgements and Reference

1. System of VarySysDB



Figure 1-1 Home

Table 1. Subsystems and Pages included in VarySysDB

Subsystem	Page	Function
VaryGene 2	Polymorphism Search	Retrieving and displaying genetic polymorphism.
	Polymorphism Table	Displaying detailed information on polymorphisms.
	Transcript Search	Retrieving and displaying transcript information.
	Transcript Table	Displaying detailed information on transcripts.
	Sequence View	Displaying cDNA sequence with information on polymorphism and functional domains.
	STR/SAR Search	Retrieving and displaying short tandem repeats (STRs) and simple amino acid repeats (SARs).
	CNV Search	Retrieving and displaying Copy Number Variations (CNVs).
	CNV Table	Displaying detailed information on CNVs.
	Keyword Search	Retrieving and displaying by ID, gene name, or Definition.
	System Information	Displaying summary table showing total numbers of transcripts and polymorphisms in VaryGene2.
LD-Search	-	Retrieving and displaying LD-bins within the specified region.
GBrowse	-	Displaying genomic region specified with HITs, HIXs, and polymorphisms.

The results of searches can be downloaded and easily displayed on the computer screen.

1.1. VaryGene2

A menu bar of Varygene2 is designed to select search pages from “Polymorphisms” (i.e., search by feature of SNPs and DIPs), “Transcripts” (i.e., by feature of HITs), “STRs/SARs” (i.e., by feature of STRs and SARs), and “CNVs” (i.e., by feature of CNVs) (Figure 1-1).

1.1.1. Polymorphism Search

The screenshot displays the Polymorphism Search web interface. It includes search criteria for position (Chromosome, Band, Genome Start, Genome End), Polymorphism Features (SNP, DIP, Validated, Heterozygosity), Polymorphism classification (Region in transcript, Type(CDS)), and Search for Analysis Result (Effect on Functional Domains, OMIM Allelic Variant, Effect on Protein 3D Structure). The interface also features search buttons (Search, Download) and a Reset button.

The results table below shows the following data:

dbSNP ID	Position	Allele	Strand	Validated	Heterozygosity	Link
rs10000	63281873..3226	A/G	-	Yes	0.2	dbSNP
rs10000	75768364..57638	C/T	-	Yes	0.43	dbSNP
rs10000	76564143..55614	C/T	+	Yes	0.43	dbSNP
rs100000	31157919..11579	C/T	-	Yes	?	dbSNP
rs10000111	443615..43615	C/T	+	No	?	dbSNP
rs10000147	45718188..5710	A/G	+	No	?	dbSNP
rs1000015	518208394..182	A/G	-	Yes	0.05	dbSNP
rs1000057	415245178..1524	A/G	+	Yes	0.25	dbSNP

Figure 1-2 Polymorphism Search

In Polymorphism Search page, users can search polymorphism by features, classification, and our analysis results such as effects on functional domains and protein 3D structures. Search criteria “between” boxes are searched by “AND” search, which offers polymorphisms that include all criteria. Search criteria “within” box are searched by “OR” search, except the “Polymorphism Classification” box and “Search for Analysis Result” box, in which user can chose search conditions either “AND” or “OR” searches. Clicking “db SNP” of the retrieved results in “Polymorphism” table below leads users to “Polymorphism Table” page. User can download the retrieved data by clicking Download button.

1.1.2. Polymorphism Table

The screenshot displays a web interface for a Polymorphism Table. At the top, there is a navigation bar with "Home > Polymorphism Search > Polymorphism Table" and a "Download" button. Below this, a text prompt reads: "Select one from the search result list to see features of polymorphism in 'Polymorphism Table'".

The main section is titled "Polymorphism" and contains a table with the following data:

dbSNP ID	Position	Allele	Strand	Validation	Heterozygosity	Link
rs2000	6,32261873..32261	A/G	-	Yes	0.2	dbSNP

Below the Polymorphism table is a "Classification" section with a table containing the following data:

H-Inv ID	Position in Transcript	Region	Type	Codon	Effect on Domain	Effect on Protein 3D	OMM Allele	Variant	Link
rs2000	1898..1898	3'UTR	+						Sequence View GBrowse
rs2000	1818..1818	3'UTR	+						Sequence View GBrowse
rs2000	1853..1853	3'UTR	+						Sequence View GBrowse

At the bottom of the interface, there are links for "Genome Browser" and "LD Search System", and a footer with "AIST JBIC JBIRC H-Inv" and "Copyright (c) 2007, 2008 JBIC, AIST, and NIG. All Rights Reserved."

Figure1-3 Polymorphism Table

Click somewhere on a row in "Polymorphism Table" and user can browse classification and analysis results of the selected polymorphism, such as Position in H-Inv Transcript and effects on domain. "Classification" table contains links to the "Transcript Table", "Sequence View", "GBrowse", "Transcript View" and "Locus View." User can download the retrieved data by clicking Download button.

1.1.3. Transcript Search

Home > Transcript Search

Search by position
 Chromosome Band Genome Start Genome End

Selection of Representative Transcript
 Representative

Search by Similarity Category
 Category 1 Category 2 Category 3 Category 4 Category 5 Category 6 Category 7

H-Inv ID	Position	Accession No.	H-Inv cluster ID	Representative Trans Gene Name	Category	Definition	Link
HInv000025	15:20822968..3093	HT000957098	H93012142	hinv	serine peptidase w/ 1	Hunt2-type protein	Sequence View GBrowse
HInv000026	21:30348186..3841	HT000957099	H93043034	hinv	Down syndrome cr	Down syndrome cr	Sequence View GBrowse
HInv000014	9:92295693..92268	HT000957090	H93025694	hinv	osteomodulin	Osteomodulin prec.	Sequence View GBrowse
HInv000015	1:70739163..79029	HT000957090	H93000722	hinv	interferon-induced	Similar to Interferon	Sequence View GBrowse
HInv000020	7:80016836..88193	HT000957094	H93008098	hinv	sexo domain, vsmu	Sexuprotein 3C, pre	Sequence View GBrowse
HInv000021	17:31415753..3143	HT000957095	H93029017	hinv	shendrin [C-C to	Small inducible cyt	Sequence View GBrowse
HInv000063	1:10444542..10445	HT000957096	H93000113	hinv	collistatin	Collistatin precursor	Sequence View GBrowse
HInv000068	11:3793147..37224	HT000957098	H93000377	hinv		Similar to Nuclear p	Sequence View GBrowse
HInv000076	8:1902423..163718	HT000957098	H93038299	hinv	discs, large (Dros	Similar to discs leg	Sequence View GBrowse
HInv000076	18:3488837..38253	HT000957098	H93014387	hinv	discs, large (Dros	Discs large-antoci	Sequence View GBrowse
HInv000077	18:3488837..38706	HT000957098	H93014387	hinv	discs, large (Dros	Discs large-antoci	Sequence View GBrowse

Page 1 of 3000

Displaying items 1 - 100 of 167982

[Genome Browser](#) [LD Search System](#)

[JST](#) [JIRC](#) [JIRIC](#) [NIG](#)

Copyright (c) 2007, 2008 JST, JIRC, and NIG. All Rights Reserved.

Figure1-4 Transcript Search

In Transcript Search page, users can search transcript by features. Search condition concerning between and within box is the same as Polymorphism search page. Clicking “H-Inv ID” of the retrieved results in “Transcript” table below leads users to “Transcript Table” page. “Transcript” table showing retrieved results contains links to the “Sequence View”, “GBrowse”, “Transcript View”, and “Locus View.” User can download the retrieved data by clicking Download button.

1.1.4. Transcript Table

Home > Transcript Search > Transcript Table

Download

Transcript

Hinv ID	Position	Accession No.	Hinv cluster ID	Representative Tran	Gene Name	Category	Definition	Link
H000001	632261278..32265	HF000261876	H00005705	tabe	pre-B-cell leukemia 1	1	Pre-B-cell leukemia	Sequence View GBrowse

Classification

dbSNP ID	Position	Allele	Strand	Validation	Heterozygosity	Position in Transcrip	Region	Type	Code
rs1500	632261673..32265	A/G	-	Yes	0.2	1508..1508	3'UTR	-	
rs1500	632261681..32265	C/T	-	Yes	0.5	1502..1502	3'UTR	-	
rs1500	632261695..32265	C/G	-	No	0	1458..1458	3'UTR	-	
rs1500001	632262080..32263	C/G	-	Yes	0.48	1302..1302	CDS	non-synonymous	

Page 1 of 1

Displaying topics 1 - 19 of 19

Domain

Position in Transcrip	InterPro	Name
801..802	IP0001250	Homeobox
758..805	IP0001250	Homeobox
753..850	IP0001250	Homeobox
758..805	IP0001250	Homeobox
803..802	IP0001250	Homeobox

Page 1 of 1

Displaying topics 1 - 10 of 10

STR

Region	Position in Transcrip	Repeat Unit	Repeat Number	Polymorphism
3'UTR	1501..1508	ca	1	false
3'UTR	2038..2033	ag	1	false
3'UTR	2068..2068	ggf	1	true

Page 1 of 1

Displaying topics 1 - 3 of 3

SAR

Region	Position in Transcrip	Repeat Unit	Repeat Number	Polymorphism
CDS	422..458	A	1	false

Page 1 of 1

Displaying topics 1 - 1 of 1

Genes Browser LD Search System

Figure1-5 Transcript Table

Transcript Table page displays information of transcript, polymorphisms, functional domains, STRs, and SARs. “Transcript” table above contains links to the “Sequence View”, “GBrowse”, “Transcript View”, and “Locus View” of H- InvDB. User can download the retrieved data by clicking Download button.

1.1.5. Sequence View



Figure1-6 Sequence View

Sequence View page displays transcript sequence with polymorphism classification and location of CDS and functional domain. Click on colored area and users can see detailed information in a separate window. User can download the displayed data by clicking Download button.

1.1.6. STR/SAR Search

The screenshot displays the STR/SAR Search interface. At the top, there are search criteria fields for Chromosome, Band, Genome Start, and Genome End. Below these are options for 'Selection of Representative Transcript' and 'Repeat Search' (Microsatellite or Single Amino acid Repeat). A 'Repeat Polymorphism' checkbox and a 'Unit' dropdown are also present. A 'Repeat Count' field with a range selector is visible. A search bar contains the text 'Search' and a 'Download (Invit 10000)' button. Below the search bar is a 'Repeat' table with the following columns: H-Inv ID, Region, Position in Transcript, Repeat Unit, Repeat Number, Polymorphism, H-Inv cluster ID, Transcript Position, Representative Tra, and Deln.

H-Inv ID	Region	Position in Transcript	Repeat Unit	Repeat Number	Polymorphism	H-Inv cluster ID	Transcript Position	Representative Tra	Deln
HInv000020	CDG	1671-1680	#	5	false	HV0006006	7:20019326-20193	true	S
HInv000430	CDG	3065-3077	gn	5	false	HV0004339	4:2604305-277140	false	re
HInv000450	CDG	3065-3077	gn	5	false	HV0004339	4:2604305-277140	true	SI
HInv000481	3'UTR	3016-3030	gn	5	false	HV0004339	4:2604305-277140	false	re
HInv000484	3'UTR	3070-3081	ag	6	false	HV0004040	4:2574278-290570	false	SI
HInv000509	3'UTR	2520-2529	at	5	false	HV0007585	1:287685574-2879	true	TI
HInv000509	3'UTR	2530-2539	at	8	false	HV0007585	1:287685574-2879	true	TI
HInv000620	CDG	1988-1987	cg	6	false	HV0003058	7:181521807-1816	true	PI
HInv000714	5'UTR	62-181	gtcc	5	false	HV0006750	7:72627078-72829	false	CI
HInv000714	3'UTR	982-902	ccc	7	false	HV0006750	7:72627078-72829	false	CI
HInv000714	3'UTR	904-818	ccc	6	false	HV0006750	7:72627078-72829	false	CI

At the bottom of the table, there is a pagination control showing 'Page 1 of 306' and a 'Displaying topics 1 - 100 of 38435' indicator. Below the table are links for 'Genome Browser' and 'LD Search System', and a footer with 'AIST JBIC JBIRC N-Inv' and 'Copyright (c) 2007,2008 JBIC, AIST, and NIG. All Rights Reserved.'

Figure1-7 STR/SAR Search

In STR/SAR Search page, user can search STRs and SARs with transcript information. Clicking “H-Inv ID” of the retrieved results in “Repeat” table below leads users to “Transcript Table” page. “Repeat” table contains links to the “Transcript View”, and “Locus View” of H-InvDB. User can download the retrieved data by clicking Download button.

1.1.1.7. CNV Search

The screenshot displays the CNV Search web interface. At the top, there is a navigation bar with "Home" and "CNV Search" links. Below this is a "Search by position" section with input fields for "Chromosome", "Band", "Genome Start", and "Genome End". A "Detection" section contains checkboxes for "Division A" through "Division F". The "CNV Class" section has checkboxes for "Copy number variation" and "Inversion". A dark bar at the bottom of the search section contains "Search", "Download (limit 10000)", "OK", and "Reset" buttons.

Below the search section is a table titled "Structural variation". The table has five columns: "Variation ID", "Position", "CNV Class", "Method", and "Link". The table contains 13 rows of data, all with "Copy number variant" as the CNV Class and "Amev CGH" as the Method. Each row has a "Link" column with a "SubBrowse" button. The table is paginated, showing "Page 1 of 128" and "Displaying topics 1 - 100 of 1194".

At the bottom of the page, there are links for "Genome Browser" and "LIP Search System", and a footer with "AIST JBRC JBIRC H-Inv" and "Copyright (c) 2007,2008 JBRC, AIST, and NIG. All Rights Reserved."

Figure1-8 CNV Search

In CNV Search page, user can search CNVs. Clicking “Variation ID” of the retrieved results in “Structural variation” table below leads users to “CNV Table.” User can download the retrieved data by clicking Download button.

1.1.8. CNV Table

The screenshot shows a web browser window with the following elements:

- Navigation: Home > CNV Search > CNV Table
- Buttons: Download
- Section: Structural variations
- Table:

Variant ID	Position	CNV Class	Method	Link
Variant_000	1,108,885,120,000	Copy number neutral	Array CGH	Cross Deleted

Page 1 of 1 | Displaying topics 1 - 1 of 1

Genome Browser | LD Search System

AIST JBIC JBIRC H-Inv

Copyright (c) 2007, 2008 JBIC, AIST, and NIG. All Rights Reserved.

Figure1-9 CNV Table

CNV Table page displays information of CNVs. User can download the retrieved data by clicking Download button.

1.1.9. Keyword Search

rs1000 ID Go

Home > Keyword Search

Keyword Search

ID #	Content type	Content
rs1000	dbSNP ID	rs1000
rs10000	dbSNP ID	rs10000
rs100000	dbSNP ID	rs100000
rs1000000	dbSNP ID	rs1000000
rs1000010	dbSNP ID	rs1000010
rs1000047	dbSNP ID	rs1000047
rs100005	dbSNP ID	rs100005
rs1000057	dbSNP ID	rs1000057
rs1000090	dbSNP ID	rs1000090
rs10000713	dbSNP ID	rs10000713
rs10000027	dbSNP ID	rs10000027
rs10000964	dbSNP ID	rs10000964
rs10001	dbSNP ID	rs10001
rs1000114	dbSNP ID	rs1000114

Page 2 of 2

Displaying results 1 - 100 of 158

Genome Browser LD Search System

AIST JBIC JBIRC NInv

Copyright (c) 2007, 2008 JBIC, AIST, and NIG. All Rights Reserved.

Figure1-10 Keyword Search

In the search box above in each page, user can search IDs, genes, and definitions of H-Inv transcript by keyword. “Keyword Search” table contains links to suitable page according to retrieved keyword types. For example, when “dbSNP ID” is selected for keyword search, link to “Polymorphism” table will be appeared.

1.1.10. System Information

Home > System Information

System Information

Name	Total	Version
Human Transcript	147992	5.0
dbSNP ID	321420	135

Data Information

Name	Total
-	571700
AA-STOP	3006
NRD	7500
Nonzygosity	384307
STOP_AA	880
Synonymous	320875

[Genome Browser](#) [LD Search System](#)

[AJST](#) [JBIC](#) [JBIRC](#) [Home](#)

Copyright (c) 2007, 2008 JBIC, AJST, and NIG. All Rights Reserved.

Figure1-11 System Information

System Information page displays summary table showing total numbers of transcripts and polymorphisms in VaryGene2.

1.2. LD Search System

VarySysDB
genetic polymorphism

LD Search System

Search by position

Chromosome Band Genome Start Genome End

Population

R² Threshold (default 0.8, (range 0.5-1.0))

SNP Marker

rsSNP ID

Search Download (first 10000)

LD

rsSNP ID	Position	Marker	R ²	Link
7:141322-178986		rs6853338-rs7702358	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:638652-881293		rs7706636-rs10281779	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:632585-364417		rs6873278-rs17719	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:637086-336262		rs4719496-rs6461376	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:971049-1069904		rs2297-rs10236055	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:1003325-1251706		rs6591-rs4075837	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:1009826-1899607		rs19298972-rs1007786	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype
7:1046870-1214553		rs2801444-rs850218	MORE THAN 0.8	rsSNP Search Subpopulation Search LD Search Haplotype

Figure2-1 LD Search System

LD Search System is a subsystem to retrieve LD-bin distributed within a specified region of the chromosome. Data of the LD-bins presented here are definitive haplotypes that originate from a single sperm, indicating that they are free from errors, which is typically caused by the inference from diploid genotypes. The data of LD-bin are provided by Prof. Kenshi Hayashi and Dr. Koichiro Higasa at Kyushu University (D-HaploDB Phase II on Build 36). For detail of data, please see web page of D-HaploDB (<http://finch.gen.kyushu-u.ac.jp/>) (1).

1.3. GBrowse



Figure3-1 GBrowse

GBrowse in VarySysDB can be used to navigate positional relationships among HITS, HIXs, and polymorphisms. Since GBrowse is an open source architecture with various functions, users can conveniently download information from the retrieved region and upload their own data to make comparisons with the information in VarySysDB.

URL or GBrowse:

<http://www.gmod.org/wiki/index.php/Gbrowse>

2. Acknowledgements and Reference

Data of LD-bin (Definitive haplotype):

Koichiro Higasa and Kenshi Hayashi

(Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University)

Advice and Discussion:

Shinsei Minoshima

(Photon Medical Research Center, Hamamatsu University School of Medicine)

Satoshi Fukuchi

(Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Shizuoka,)

1. **Higasa, K., Miyatake, K., Kukita, Y., Tahira, T. and Hayashi, K. (2007) D-HaploDB: a database of definitive haplotypes determined by genotyping complete hydatidiform mole samples. *Nucl. Acids Res.*, 35, D685-689.**