

個別化医療に向けた ヒトゲノムバリエーションデータベース

東京大学大学院医学系研究科

徳永勝士

共同研究機関

東大附属病院、遺伝学研究所、日立製作所

背景

ゲノムワイドなSNPタイピング技術と次世代シーケンサ (NGS) 技術の革新的向上と普及

新規発見の疾患関連/薬剤応答遺伝子・変異が増加



- 1) データの半永続的な管理と研究者間での情報共有
- 2) 変異と表現型との関係の体系化

- ・ 複数遺伝子/変異の疾患への複雑な関与
- ・ 同一疾患における変異部位・種類特異的症状の存在
- ・ 同一遺伝子/変異の複数疾患への関与
- ・ スタディー間の結果の不一致
(集団間差位、検出力不足、表現型定義のゆれ)

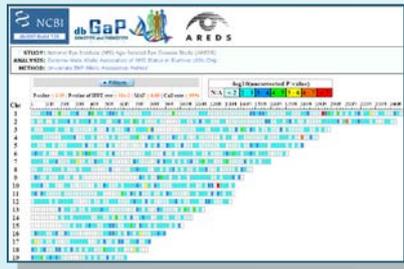
海外でのGWAS/NGSデータ共有化

1) NCBI

- ・ dbSNP SNP情報を蓄積
- ・ dbVAR 構造多型のデータを蓄積
- ・ dbGaP GWAS, 次世代シーケンサー結果を含むgenotype-phenotype に関するデータを蓄積

2) EBI

- ・ EGA GWAS, 次世代シーケンサー結果を含むgenotype-phenotype に関するデータを蓄積



3) Gen2Phen

- ・ GWAS, 次世代シーケンサー結果を含むgenotype-phenotype に関するデータを蓄積

海外での変異データの体系化

- 1) HGMD 疾患関連遺伝子変異（生殖細胞のみ）を網羅的に収集
（有料 登録後3年を経過したデータはアカデミック無料）
- 2) NCBI
 - ・ ClinVar dbSNP, dbVAR, dbGapのデータ及び、臨床検査などの
サブミットデータを精査し登録したgenotypeとphenotypeの関係をDB
- 3) HGVS (human genome variation society)
 - ・ LSDB (locus specific database)のリンク集
- 4) COSMIC: がん体細胞変異のDB
- 5) ICGC, TCGA Project DB がん体細胞変異の体系化と
ゲノム配列データアーカイブ

ヒトゲノムバリエーションDBの目的

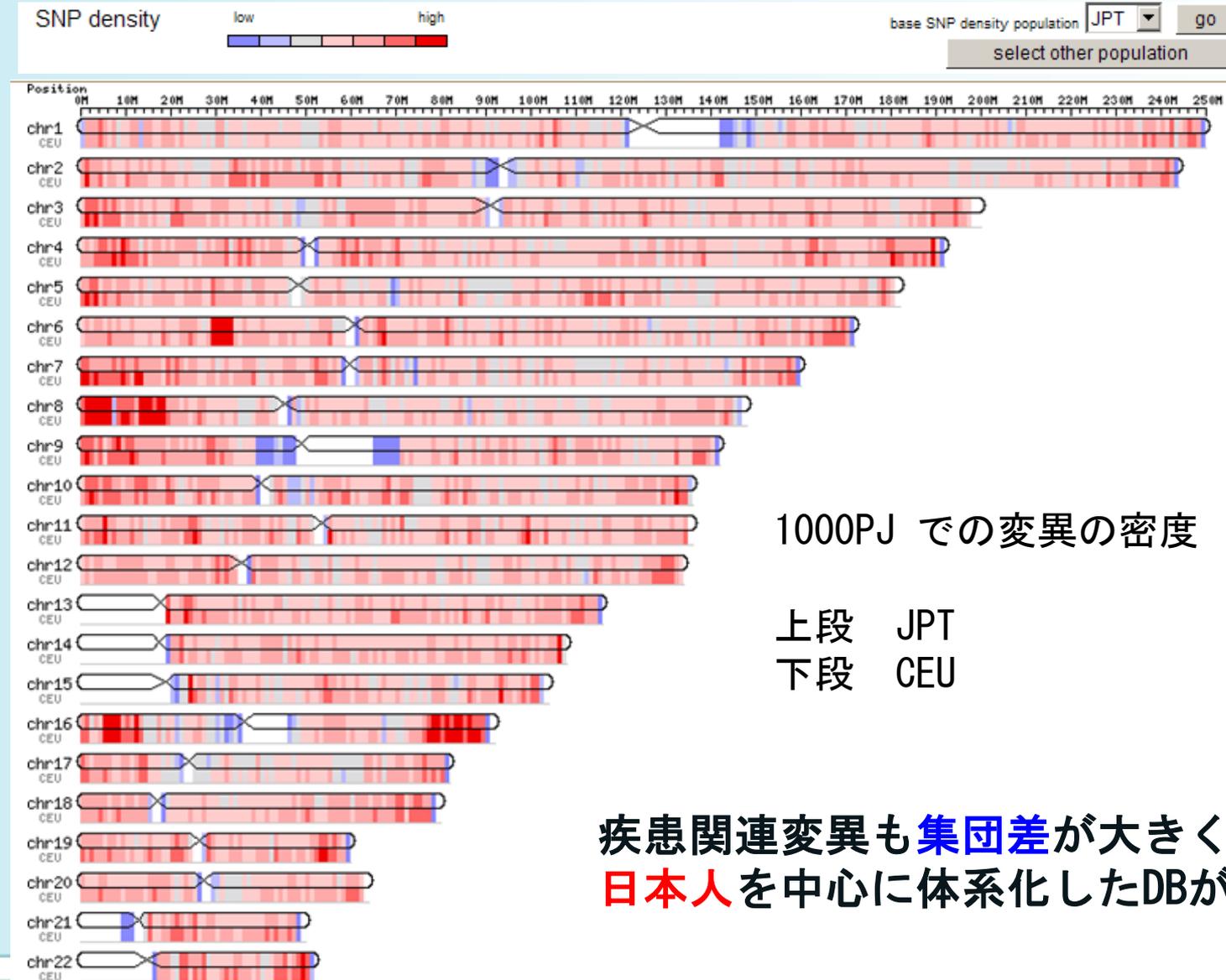
目的：

変異・疾患・薬剤応答・ウィルス耐性・臨床情報を整理・体系化し、俯瞰可能とすると共に、表現型の機序理解促進と個別化医療の礎とする

構想：

1. 次世代シーケンサー(NGS)および、その他の解析法 (GWASを含む) によって発見される多型・変異情報の預入れ
2. 文献情報を含め過去産出された疾患・薬物応答・ウィルス耐性に関する多型・変異データの収集 (統計値と各サンプルデータ)
3. オミックスデータとの連携による疾患・薬物応答の機序理解の促進
4. 既知変異の情報 (配列上の性質) を元に、新規変異の表現型を予想

Human Variation DB - 分布の集団間比較



原発性胆汁性肝硬変 (PBC) に見る 日本人とヨーロッパ系集団の共通性と異質性

日本人で再現された感受性遺伝子

Gene	OR	P value
<i>HLA-DQB1</i>	1.94	8.30×10^{-13}
<i>IL7R</i>	1.47	4.13×10^{-4}
<i>IKZF3</i>	1.44	6.77×10^{-4}
<i>STAT4</i>	1.35	8.11×10^{-4}
<i>RAD51L1</i>	1.07	0.0023
<i>CXCR5</i>	1.42	0.0027
<i>PLCL2</i>	1.07	0.0030
<i>rs6974491</i>	1.33	0.010
<i>IL12RB2/SCHIP1</i>	1.12	0.013
<i>NFKB1</i>	1.35	0.019
<i>CD80</i>	1.48	0.022
<i>DENND1B</i>	1.14	0.046

Th, B, Immune system

日本人で見出された遺伝子

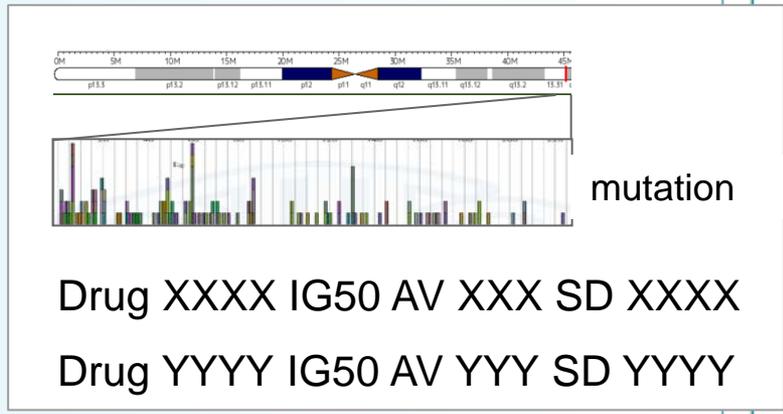
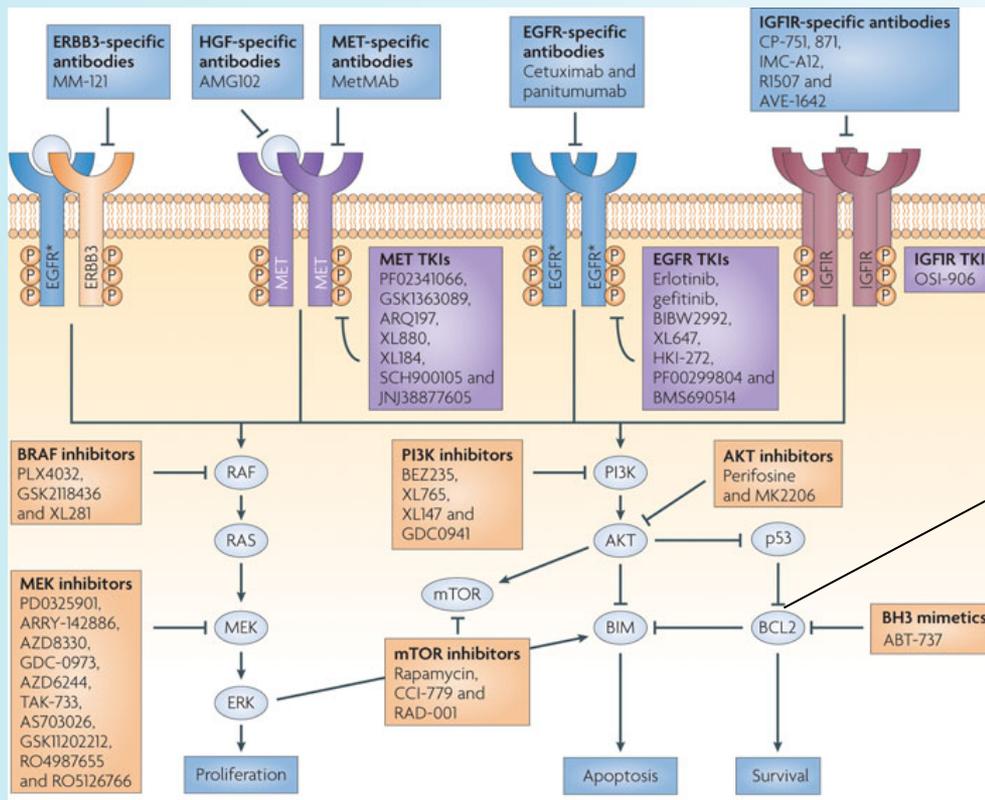
<i>TNFSF15</i>	1.56	2.84×10^{-14}
<i>PAU2AF1</i>	1.39	2.38×10^{-8}

日本人で再現されなかった遺伝子

Gene	OR	P
<i>IRF5/TNPO3</i>	1.08	0.068
<i>rs11117432</i>	1.21	0.10
<i>SPIB</i>	1.08	0.10
<i>RPS6KA4</i>	1.12	0.10
<i>CLEC16A</i>	1.10	0.15
<i>TNFRSF1A</i>	1.12	0.16
<i>IL12A</i>	1.02	0.37
<i>MAP3K7IP1 (TAB1)</i>	1.29	0.43
<i>MMEL1</i>	1.07	0.78
<i>TNFAIP2</i>	1.22	0.98

Nakamura et al. Am J Hum Genet (2012)

パスウェイ、薬剤応答性と変異



Nature Reviews | Cancer

Pao & Chmielecki Nat. Rev. Cancer. 2010

開発データベース

Human Variation DB

HLA DATABASE

SNP CONTROL DATABASE

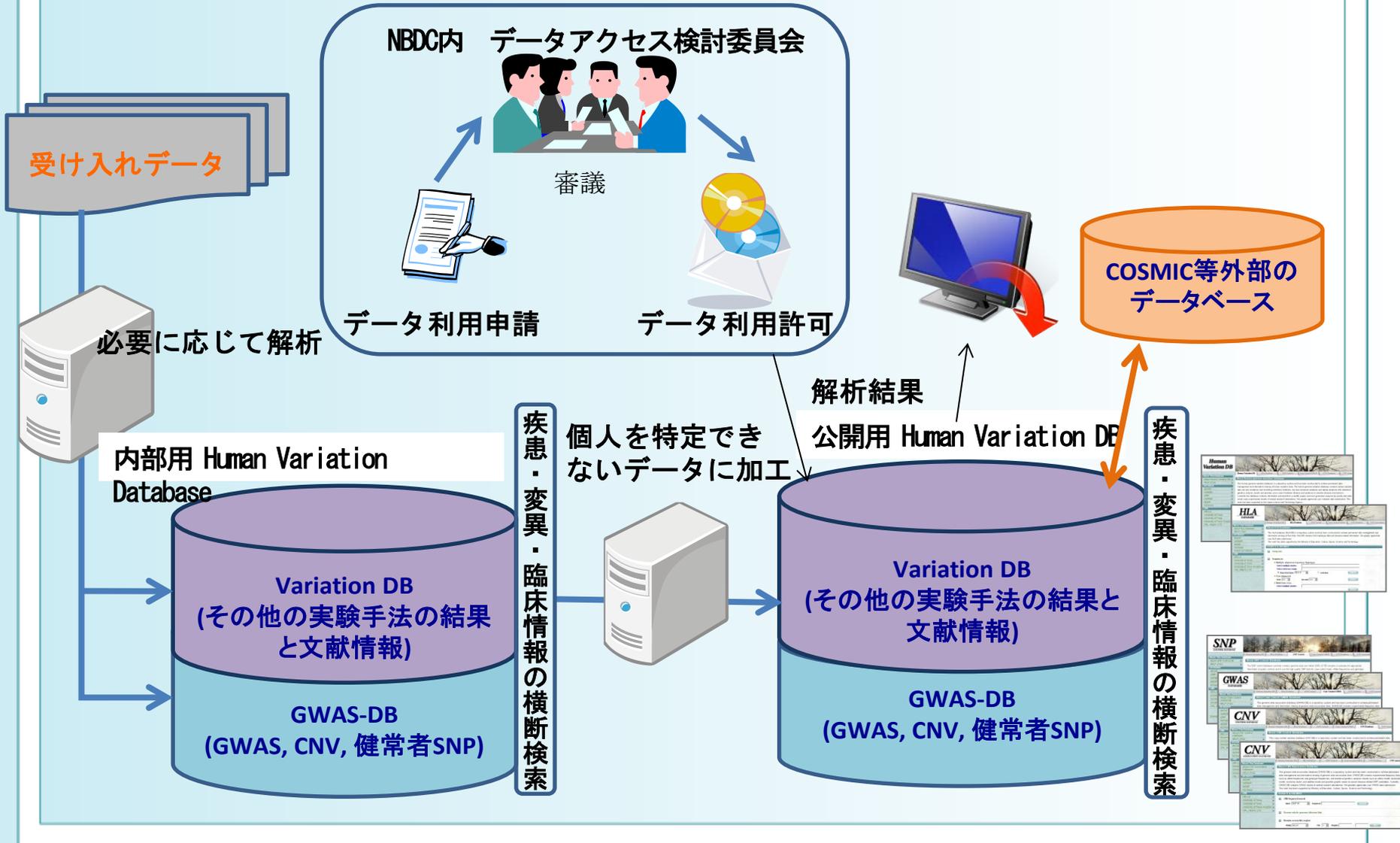
GWAS DATABASE

CNV ASSOCIATION DATABASE

NGSなど
対応DB

GWAS
対応DB

DBシステムの概要



標準SNP DB

▶標準 SNP DB: 健常者のSNPデータ (GWAS チップ用) のデータ

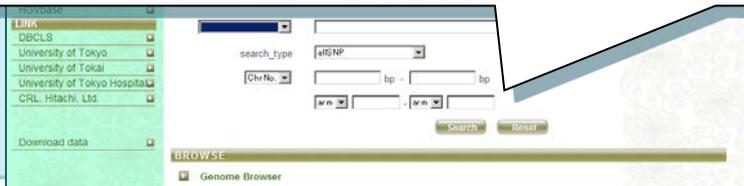
Affy500K 約500検体、Affy6.0 約600検体、Affy
Axiom 約420検体、Illumina OMNI-2.5 約420検体

Contents:

- ・30-250万SNPの遺伝子型頻度、アレル頻度、ハーディーワインバーク平衡検定値、Call rate等
- ・genotypic test, allelic test, additive risk model, recessive model, dominant model のP-value, OR, 95% CI, AICなどの遺伝統計値
- ・SNPのアノテーション

SNP

SNPの検索 (アクセッション番号、染色体上の位置、機能、疾患との関連性などで検索可能)



SNP search

SNP ID: [NRS6663840](#)

dbSNP ID(rs): [rs6663840](#)

dbSNP ID(ss): [ss16429890](#) [ss19129725](#) [ss19855219](#) [ss20488566](#) [ss23157850](#) [ss9823292](#)

JSNP ID:

HGVbase ID: [SNP006996858](#)

Chromosome: 1

Variation Class: SNP

SNP type: iSNP

Allele: A/G
NM_014704.2[A/G]:forward

Amino acid change:

Affymetrix: SNP_A-1960639 A/G

Illumina: A/G

Array kind	Ethnic group	Individual Num.	Call Rate	Genotype detail			HWP	Allele	
				A/A	A/G	G/G		A	G
Illumina317K	Japanese	200	1.000	0.19	0.51	0.3	0.774	0.450	0.550
Affy500K	Japanese	471	0.965	0.188	0.505	0.305	0.641	0.440	0.560
HAPMAP	Japanese	44	1.000	0.272	0.431	0.295	0.376	0.490	0.510

Haplotype frequencies

Affy500K NRS12563491- NRS39424283- NRS7543006- NRS2154068- NRS36702916- NRS36702935- NRS36703035- NRS6663840- NRS39424310- NRS17403773- NRS2298225- NRS2298224- NRS17404435- NRS6683156;

AAAAATAGCAAACT	0.403
GGAGGCGAAAAATC	0.393
AGAGATGGAAGGTC	0.104
GGGGCGAAGAATC	0.033
AGGGCGAAGAATC	0.015
AGAGGCGAAAAATC	0.011

Gene Name: KIAA0562

EntrezGene ID: [9731](#)

Gene Symbol: [KIAA0562](#)

Refseq ID (NM-ID): [NM_014704](#)
(NP-ID): [NP_055519](#)

Gene ontology (process):

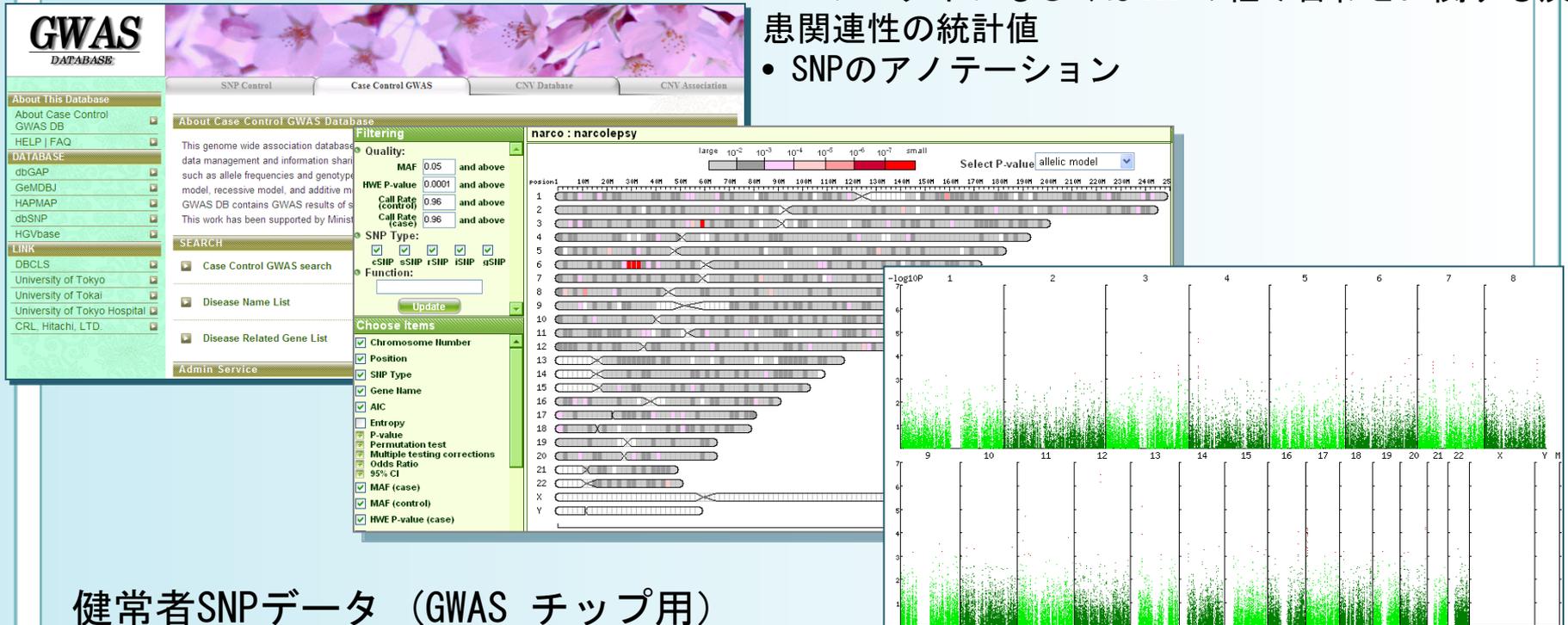
検索結果例

SNP-based GWAS DB

GWAS DB: GWASデータ
19疾患/28スタディー (内部用DB登録)
17形質 (内部用DB登録)
11疾患/13スタディー (公開データ)

Contents:

- 30-100万SNPの遺伝子型頻度、アレル頻度、ハーディー・ワインベルク平衡検定値、Call rate等
- P値 (2df, 1df), Additive risk model, recessive model, dominant modelのP-value, OR, 95% CI, AICなどの遺伝統計値
- ハプロタイプもしくはSNPの組み合わせに関する疾患関連性の統計値
- SNPのアノテーション



健常者SNPデータ (GWAS チップ用)

Affy500K 約500検体、Affy6.0 約600検体, Illumina OMNI-2.5 約420検体など

SNP-based GWAS DB

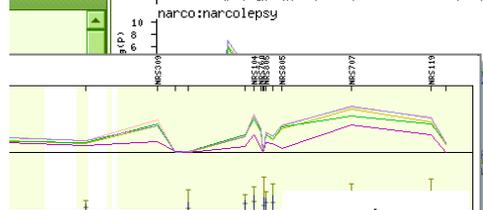
領域表示とepistasis表示

narco : narcolepsy

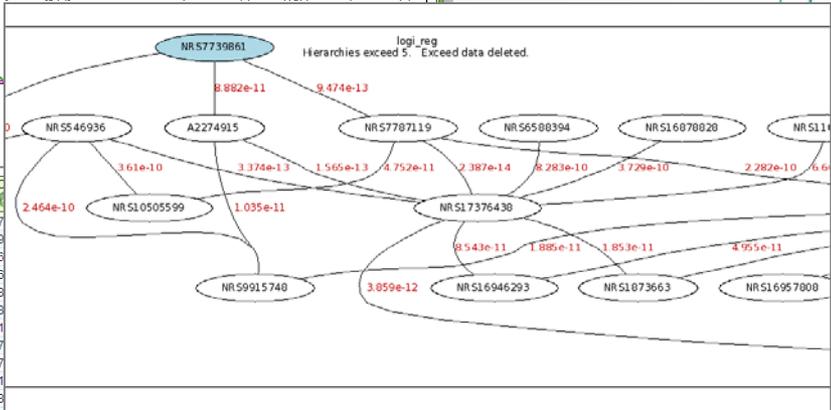
Chromosome 6 Position 30000001 - 32000000

[Study ID] [Study Name]

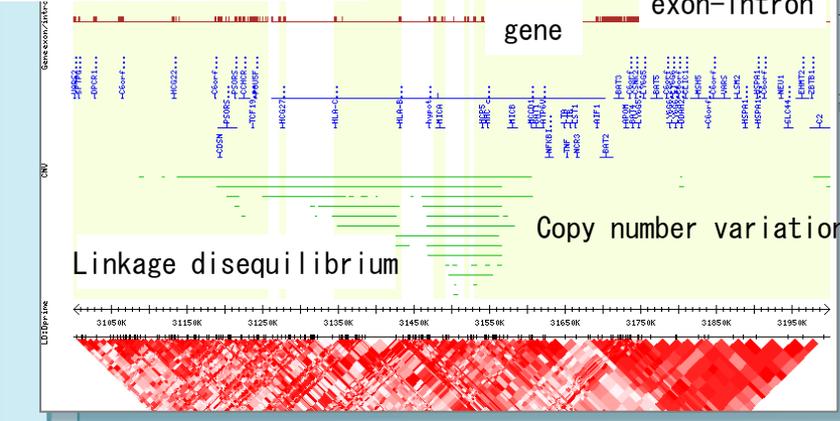
SNP ID	Chr	Position	SNP Type	Gene Name	AIC	Allelic P-value	Genotypic P-value	Dom P-value	Rec I
NRS9261282	6	30143436	ISNP	PPP1R11	3.4280	0.2219	0.4749	0.2656	0.65
NRS9261301	6	30149538	ISNP	RNF39	-23.9060	9.495e-08	5.181e-07	8.82e-07	0.00
NRS2523990	6	30185208	ISNP	TRIM31	-7.2210	0.0009156	0.002192	0.03042	0.00
NRS9261471	6	30213328	ISNP	TRIM40	1.6410	0.1852	0.171	0.4167	0.07
NRS2857435	6	30214003	ISNP	TRIM40	2.3400	0.1734	0.2508	0.3165	0.17
NRS2857439	6	30214275	ISNP	TRIM40	-1.3270	0.07401	0.0365	0.2354	0.01
NRS9261485	6	30216730	ISNP	TRIM40	0.9960	0.1745	0.1255	0.4145	0.05
NRS9261488	6	30217391	ISNP	TRIM40	0.8200	0.146	0.1114	0.3656	0.05



Int Kind	Gene A/SNP-A	C
logi_reg	NRS7787119	7
logi_reg	A2274915	9
logi_reg	NRS546936	6
logi_reg	NRS7739861	6
logi_reg	NRS17824132	8
logi_reg	NRS17824132	8
logi_reg	NRS17376438	1
logi_reg	NRS7787119	7
logi_reg	NRS7787119	7
logi_reg	NRS1945652	1
logi_reg	NRS17824132	8
logi_reg	A2274915	9
logi_reg	NRS7787119	7
logi_reg	NRS1945652	11
logi_reg	80905066	11
logi_reg	80905066	11
logi_reg	73317985	10
logi_reg	NRS4682115	3
logi_reg	NRS17376438	10
logi_reg	NRS17824132	8
logi_reg	NRS7787119	7
logi_reg	NRS4700811	5



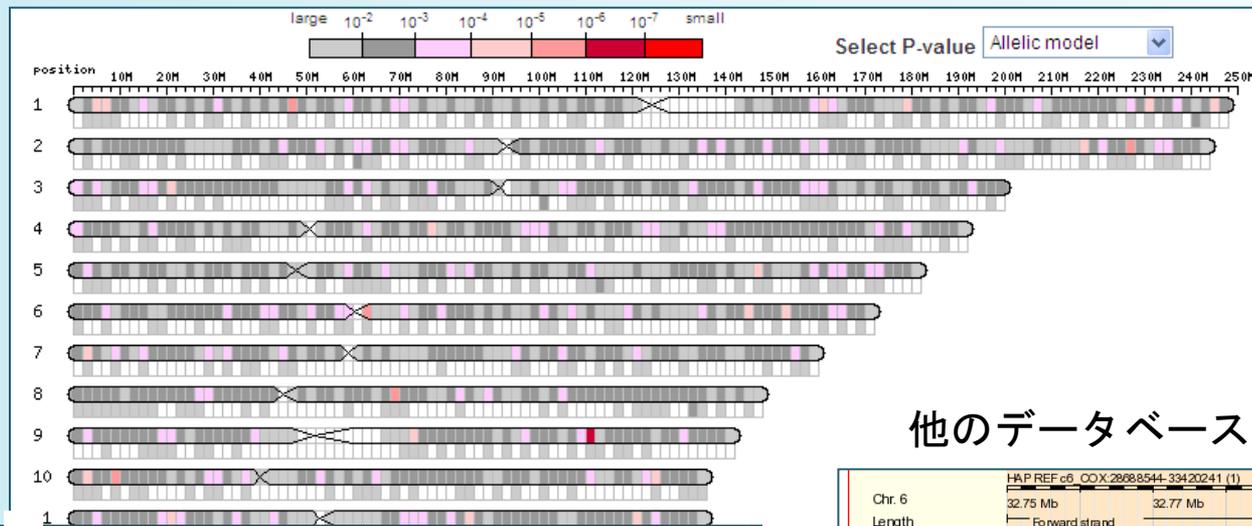
Allelic test
 Genotypic test
 Additive model
 Recessive model
 上記モデルの permutation
 多重検定の補正など



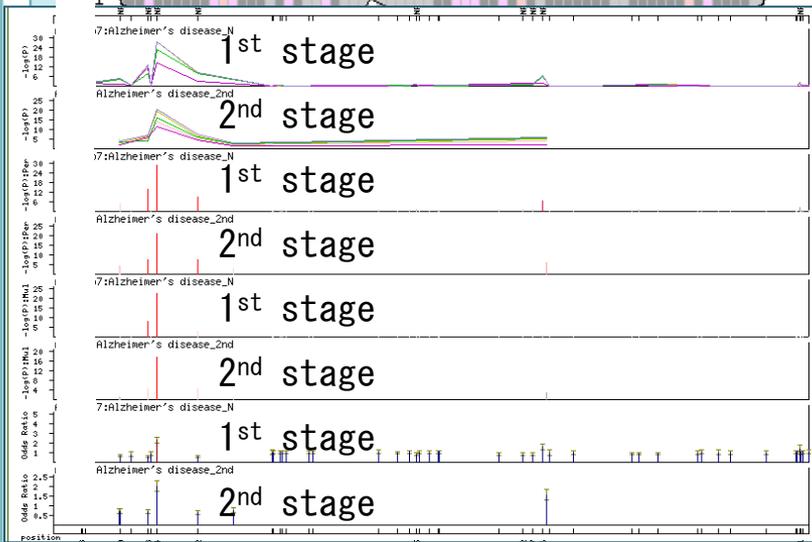
Logistic regressionをはじめとしたEpistasisの登録、表示

SNP-based GWAS DB - 比較機能

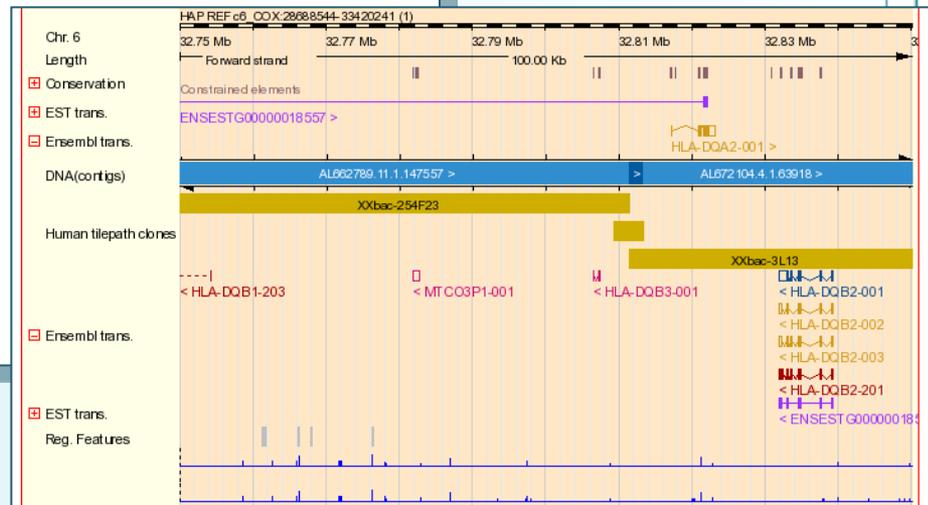
複数の実験結果の比較



上が1st stage
下が2nd stage



他のデータベースからの呼び出し (DAS)



Narcolepsy data view from Ensembl

標準CNV DB および CNV-based GWAS DB

➤ Control CNV DB

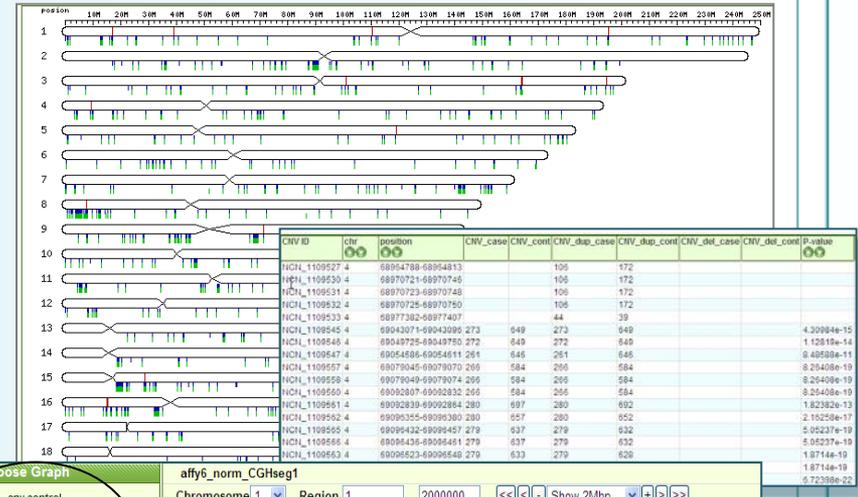
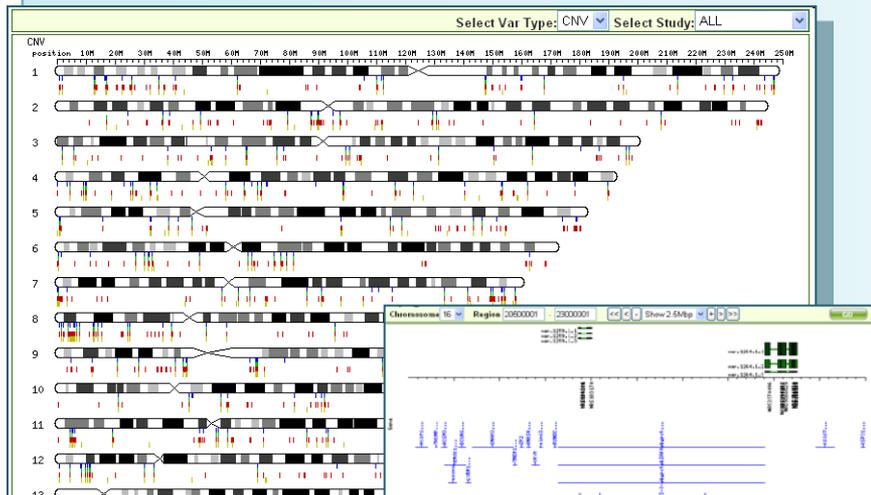
健康者 CNV DB 約160検体
登録、公開

➤ CNV association DB

Case-control association 5疾患 (内部用)
1疾患 (公開データ)

複数の計算手法の結果を比較表示

Case-controlのCNVを比較表示



Var. Alias	Sub. identifier	Region	Variations	Type	Frequency	Genotype patterns	CNV #/2 copies	Gene Name
var.1259	1	Chr16 21441805-21493293	14158					
var.1259	1.1	Chr16 21441805-21493293	2198	CNV		NRS150235_4 NRS163170_4 NRS163204_4 NRS150238_4 NRS154661_4		
var.1259	1.1.1	Chr16 21441805-21493293	1198	CNV		NRS150235_0A4D NRS163170_4C1T NRS163204_3D1T NRS150238_0A4D NRS154661_0A4D		
var.1259	1.1.2	Chr16 21441805-21493293	1198	CNV		NRS150235_0A4D NRS163170_2C2T NRS163204_2C2T NRS150238_2A2D NRS154661_1A2D		
var.1259	1.2	Chr16 21441805-21493293	3198	CNV		NRS150235_3 NRS163170_3 NRS163204_3 NRS150238_3 NRS154661_3		
var.1259	1.2.1	Chr16 21441805-21493293	1198	CNV		NRS150235_0A3D NRS163170_1C2T NRS163204_1C2T NRS150238_0A3D NRS154661_3A0D		
var.1259	1.2.2	Chr16 21441805-21493293	1198	CNV		NRS150235_0A3D NRS163170_2C1T NRS163204_2D1T		

Choose Graph

- cnv control
- Details
- Simple
- Without clustering
- With clustering
- exon/intron
- Gene
- CNV

Update

表示オプションの変更

affy6_norm_CGHseg1

Chromosome 16 Region 1 - 200000

affy6_norm_CGHseg1

var.-800000.1

var.-800000.2

var.-800000.3

var.-800000.4

var.-800000.5

var.-800000.6

var.-800000.7

var.-800000.8

var.-800000.9

var.-800002.1

var.-800002.2

var.-800002.3

var.-800000.1

var.-800000.1

var.-800000.1

クラスタリングして類似のCNVをマージ

Human Variation DB トップページ

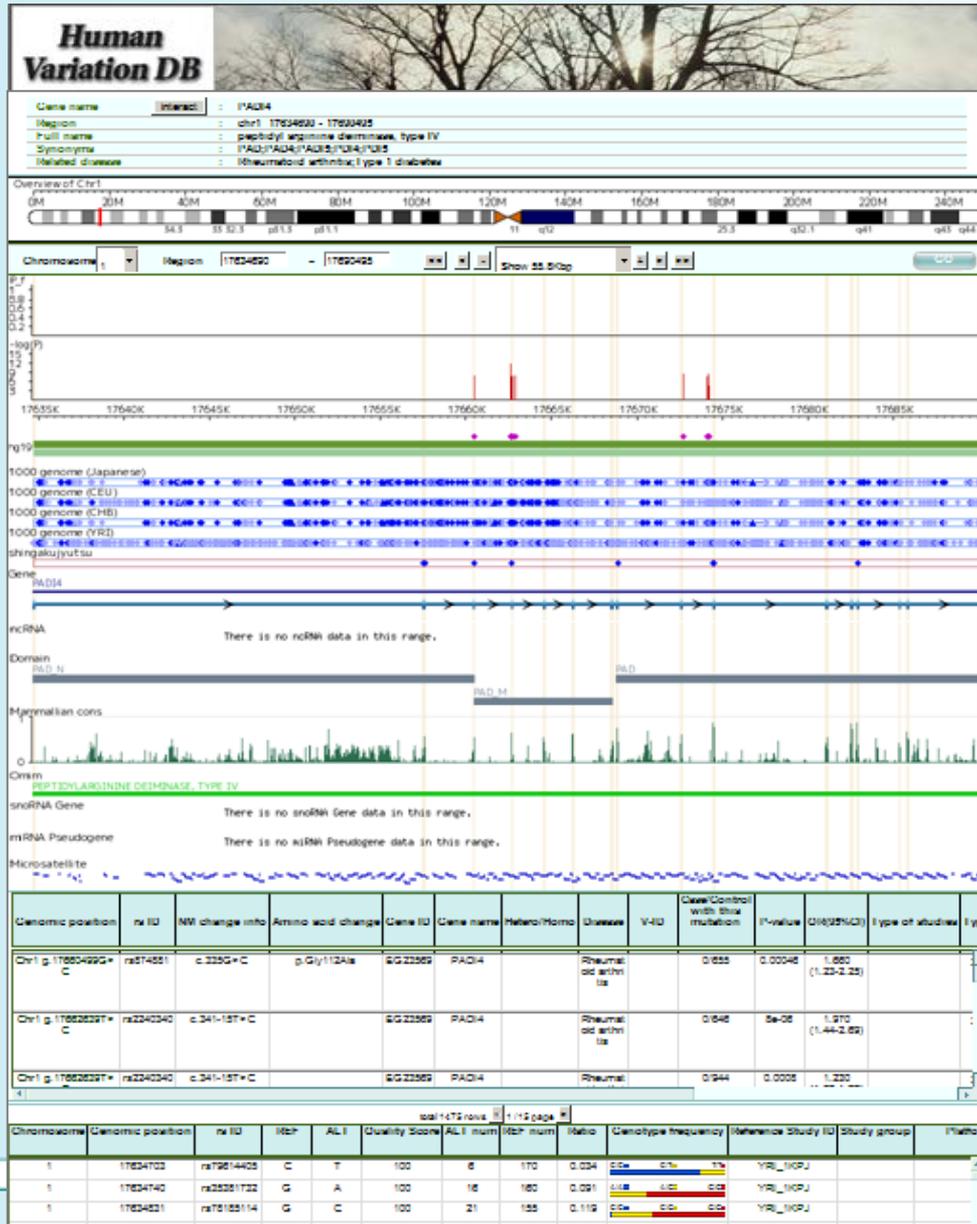
<NGSデータへの対応>

The screenshot shows the Human Variation DB homepage. On the left is a navigation menu with sections: 'About this Database' (including About Human Variation DB, HELP | FAQ, DATABASE, dbGAP, GeMDBJ, JSNP, HAPMAP, dbSNP, HGVDbase), 'LINK' (including NBOC, DBCLS, University of Tokyo, National Institute of Genetics, University of Tokyo Hospital, CRL, Hitachi, LTD.), and 'BROWSE' (including Browse by disease name, Browse by gene name, Browse by chromosome, Overview of SNP density, Summary of all diseases, Genome Browser (GBrowse)). The main content area has a header with 'Human Variation DB' and sub-headers for 'HLA Database', 'SNP Control', 'Case Control GWAS', 'CNV Database', and 'CNV Association'. Below the header is an 'ABOUT HUMAN genome variation database' section with a descriptive paragraph. A 'SEARCH' section follows, with five search methods: 'Search by gene name' (keyword field), 'Search by disease name' (keyword field), 'Search by genomic position' (ChrNo. dropdown, Region field), 'Search by SNP number' (kind dropdown set to 'SNP-ID', keyword field), and 'Interaction view by gene name' (keyword field). Each search method has a 'Search' button. At the bottom of the main content area is a 'BROWSE' section with the same five options as the search section, each with a dropdown arrow.

遺伝子検索
疾患検索
領域検索等が可能

遺伝子名,
疾患名でのブラウズ
集団ごとの変異分布の鳥瞰図
疾患ごとの変異分布の鳥瞰図
DASとしての図

Human Variation DB - 領域表示 (1)



遺伝子名

(何を表示するか
選択可能)

P-value

Reference Genome

Domain情報

Conservation score

miRNA等情報

変異詳細情報

Reference情報 (アレル頻度)

Human Variation DB - 領域表示 (2)

The screenshot displays the Human Variation DB interface. On the left, a sidebar shows the 'Human Variation DB' logo and a search bar. The main area shows a genomic region with a detailed view of reference genomes and a table of variant data.

On the right side of the interface, there are two text annotations:

- 多数のreference genome登録
随時追加可能
- 個体ごとのgenome登録

The table below shows the variant data for the region:

Chromosome	Genomic position	rs ID	REF	ALT	Quality Score	ALT num	REF num	Ratio	Genotype frequency	Reference Study ID	Study group	Platform
1	17862839	rs2240340	T	C	100	97	81	0.545		Japanese_1KPJ	1KPJ	Illumina
1	17862839	rs2240340	T	C	100	92	78	0.541		CEU_1KPJ		
1	17862839	rs2240340	T	C	100	127	67	0.655		CHB_1KPJ		
1	17862839	rs2240340	T	C	100	67	109	0.381		YRI_1KPJ		
1	17862862	rs1748033	T	C	100	105	73	0.590		Japanese_1KPJ	1KPJ	Illumina
1	17862862	rs1748033	T	C	100	110	60	0.647		CEU_1KPJ		
1	17862862	rs1748033	T	C	100	135	59	0.696		CHB_1KPJ		

Reference genomeのアリル頻度情報

Human Variation DB - 領域表示 (3)

変異のゲノム上の位置、 SNPの種類、アミノ酸置換情報、症例-対照関連P値、オッズ比、実験手法、臨床情報等

NGS解析データも、文献データも同時に表示



Genomic position	rs ID	NM ID	NM change info	NP ID	NP change info	Amino acid change	Gene ID	Gene name	var_type	SNP type	Hetero/Homo	Ethnic	N_fa
C		387.2		519.2									
Chr1 g.17862639T>C	rs2240340	NM_012387.2	c.341-15T>C	NP_036519.2			EG23569	PADI4	mutation	iSNP			
Chr1 g.17862682T>C	rs1748033	NM_012387.2	c.349T>C	NP_036519.2	p.Leu117=	TTG>CTG	EG23569	PADI4	mutation	sSNP		Japanese	
Chr1 g.17862682T>C	rs1748033	NM_012387.2	c.349T>C	NP_036519.2	p.Leu117=	TTG>CTG	EG23569	PADI4	mutation	sSNP		Japanese	

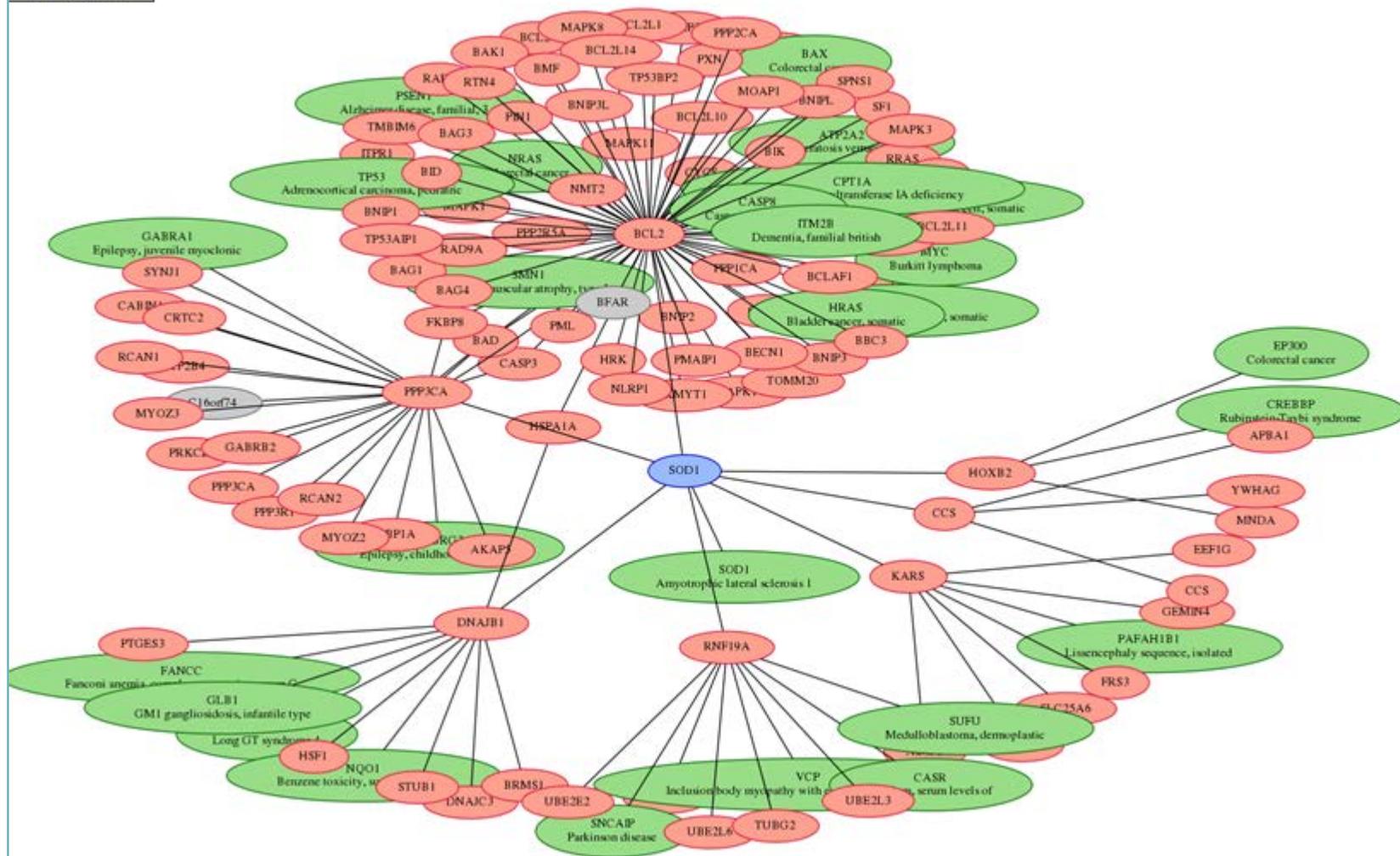
Human Variation DB - OMICSデータとの連携

蛋白質総合作用のネットワーク上に疾患情報を投影

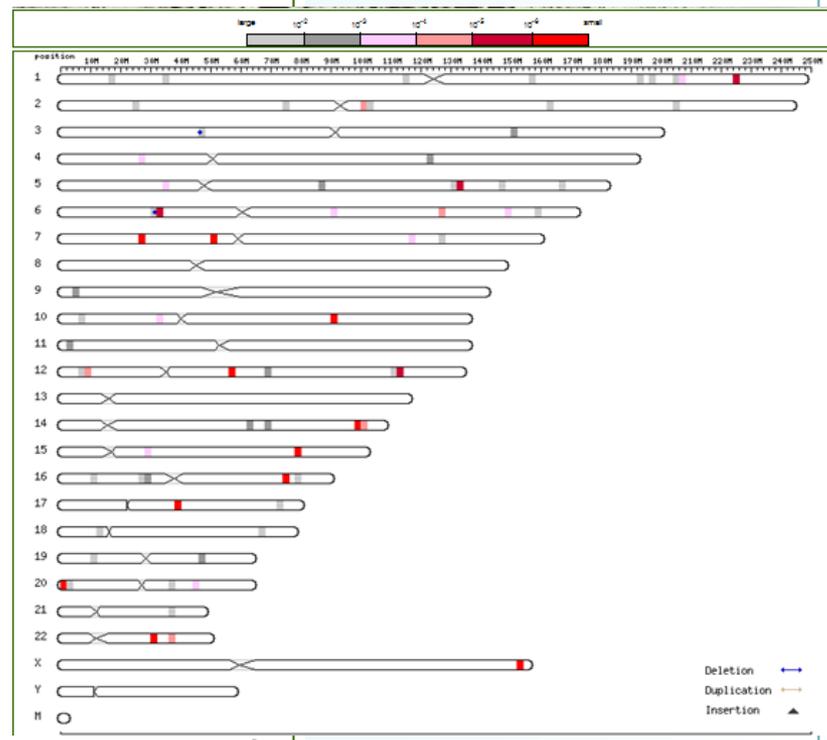
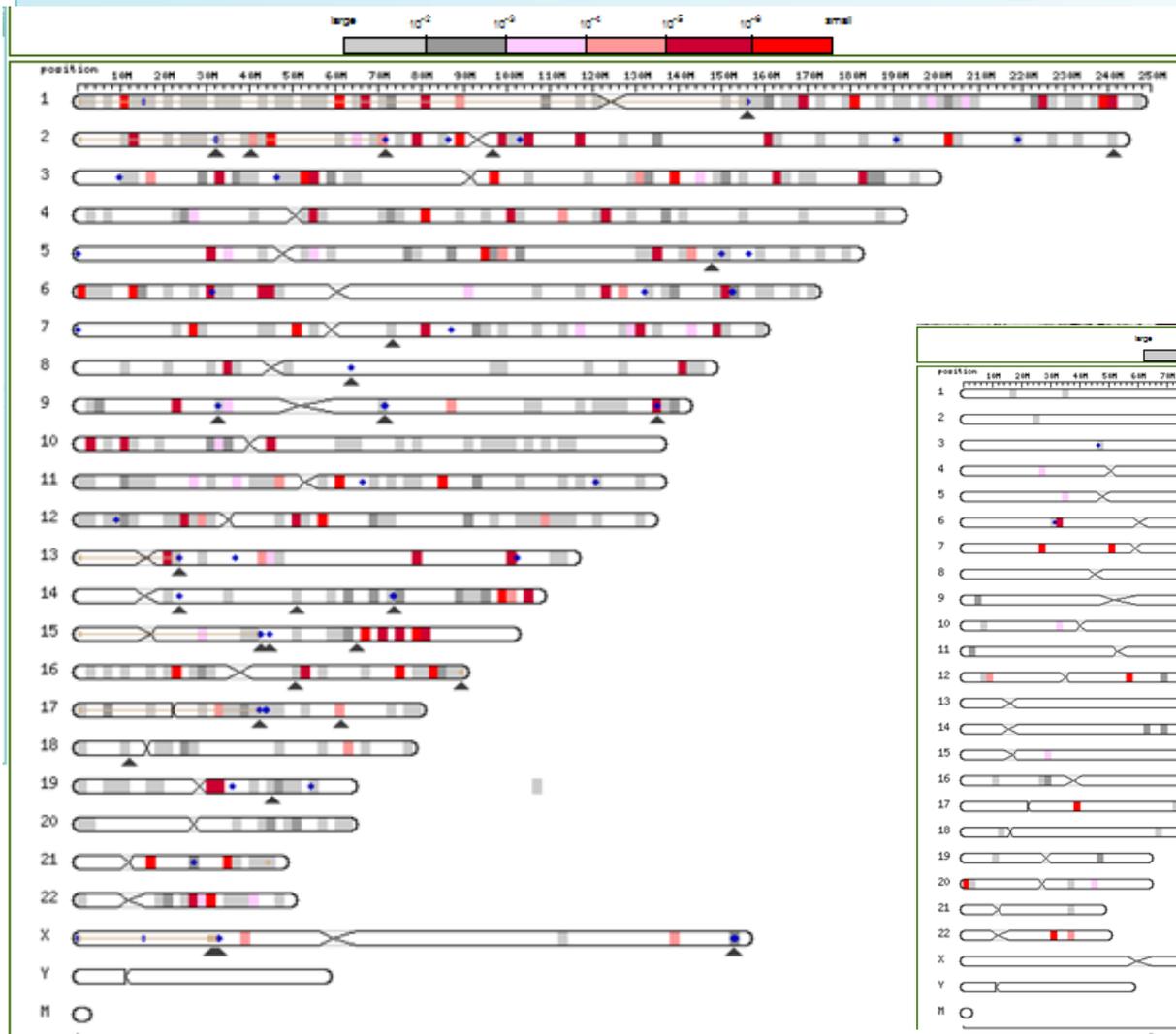
Query String : SOD1 Step : 2

zoom in zoom out

draw 1st step only



Human Variation DB - 疾患ごとの変異分布



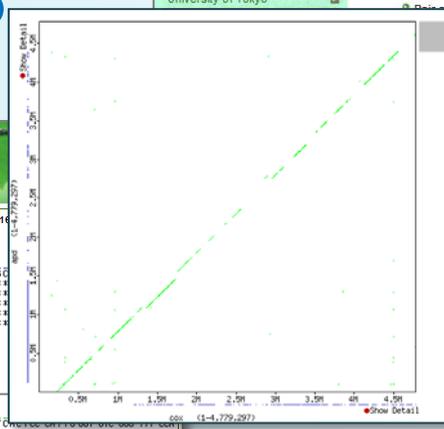
登録している全ての遺伝子での変異分布

ある疾患での変異分布

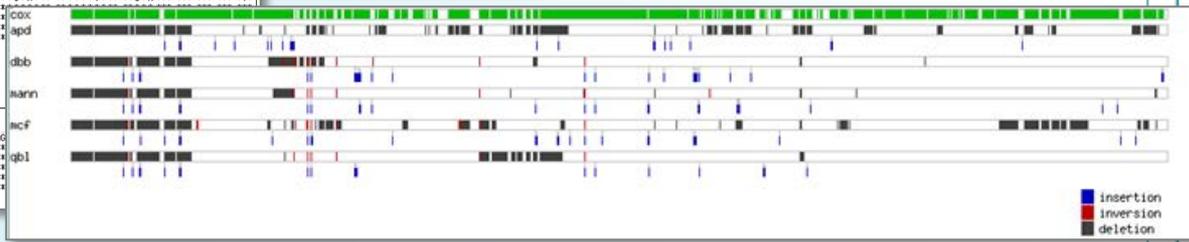
HLA DB

- HLAのハプロタイプごとの変異の登録
- HLAの多型と疾患感受性、免疫応答性、薬剤過敏症との関連を俯瞰可能

500レコード程度文献から抽出蓄積
NGSデータから変異データ抽出の
パイプラインを実装



異なるHLA型間での相同性
(遺伝子ごと)



HLA型間の塩基配列の違い

異なるHLA型間の相同性 (HLA領域全体)

関連HLA-DPB1アリルに見られる共通性と異質性

日本					韓国				
	case (n=488)	control (n=464)				case (n=251)	control (n=140)		
	freq.	freq.	P-value	OR		freq.	freq.	P-value	OR
02:01	0.186	0.245	0.00213	0.71	02:01	0.191	0.239	0.1191	0.75
04:01	0.020	0.058	2.38x10⁻⁵	0.34	04:01	0.040	0.039	1	1.01
04:02	0.039	0.099	1.59x10⁻⁷	0.37	04:02	0.028	0.129	1.27x10⁻⁷	0.19
05:01	0.467	0.386	0.00036	1.40	05:01	0.538	0.368	5.13x10⁻⁶	2.00
09:01	0.157	0.087	3.70x10⁻⁶	1.94	09:01	0.040	0.018	0.136	2.28
香港					タイ				
	case (n=280)	control (n=156)				case (n=369)	control (n=122)		
	freq.	freq.	P-value	OR		freq.	freq.	P-value	OR
02:01	0.089	0.167	0.00092	0.49	02:01	0.084	0.107	0.3014	0.77
04:01	0.039	0.048	0.5997	0.81	04:01	0.065	0.119	0.00891	0.52
-					04:02	0.018	0.025	0.591	0.71
05:01	0.552	0.481	0.04769	1.33	05:01	0.360	0.320	0.2785	1.20
-					09:01	0.011	0.016	0.5056	0.66
13:01	0.100	0.074	0.2193	1.40	13:01	0.220	0.115	0.00025	2.17
(有意なアリルのみ表示)					31:01	0.008	0.041	0.00143	0.19

NGSと文献登録データ

疾患原因変異にも特徴

データ登録数の現状

遺伝子のどの部位に生じても疾患発症

遺伝子の特定ドメインに生じた場合に発症

フレームシフト変異がホモ接合になった場合に発症

遺伝子の特定の位置で特定の置換をした場合に発症

集団により発症リスクが異なる遺伝子など

体系化することにより
疾患の**発症機序**の理解
新規変異の**表現型への影響**の予測
への貢献を期待

➤NGS公開データ

健常者：1000 genome data exome 98検体

➤NGS内部登録データ

健常者：exome 21検体、68検体、健常者：
HLA 1検体、HLA 33検体（セルライン）

➤NGS内部登録準備データ

疾患遺伝子：4遺伝子変異（新規）
2遺伝子変異（既知）

➤文献公開データ

Common disease, 神経変性変異データを
中心に変異と付随情報の登録15,000レコード

➤文献内部データ

変異と付随情報の登録 28,000レコード
HLAハプロタイプと付随情報の登録 500レコード

今後論文発表
次第随時公開するデータ

	名称	件数	概要
1	健常者データ	2 studies	ゲノムワイドSNPデータ: 数千人規模
2	B型肝炎	1 study	GWAS: 数千人規模
3	C型肝炎	1 study	GWAS: 数千人規模
4	原発性胆汁性肝硬変	1 study	GWAS: 数千人規模
5	自己免疫性肝炎	1 study	GWAS: 数百人規模
6	真性過眠症	1 study	GWAS: 数百人規模
7	ナルコレプシー	1 study	GWAS: 千人規模
8	変形性関節症	1 study	GWAS: 数百人規模
9	パニック障害	1 study	GWAS: 千人規模
10	パーソナリティ、気分障害	1 study	GWAS: 数百人規模
11	1型糖尿病	1 study	GWAS: 数百人規模
12	2型糖尿病	1 study	GWAS: 数百人規模
13	不育症	1 study	GWAS: 数百人規模
14	インスリン自己免疫症候群	1 study	GWAS: 数百人規模
15	スチーブンスジョンソン症候群	1 study	GWAS: 数百人規模
16	自閉症	1 study	GWAS: 数百人規模
17	多系統萎縮症	2 studies	GWAS: 数百人規模
18	脳動脈瘤	1 study	GWAS: 数百人規模
19	緑内障	1 study	GWAS: 数百人規模
20	放射線被ばく	1 study	GWAS: 数百人規模
21	肺がん	1 study	GWAS: 数百人規模
22	地域集団	2 study	GWAS: 数百人規模
23	HLA配列データ1	7 検体	HLA領域ホモ接合セルライン
24	HLA配列データ2	10検体	完全胞状奇胎
25	HLA配列データ3	1検体	HLA領域ホモ接合セルライン化前DNA
26	脾類上皮嚢腫のエクソーム解析	2検体	エクソーム解析に伴う変異データ(pool)
27	四肢突発性筋痛症のエクソーム解析	2検体	エクソーム解析に伴う変異データ(pool)
28	永久歯萌出不全のエクソーム解析	2検体	エクソーム解析に伴う変異データ(pool)
29	家族性腎臓病のエクソーム解析	4検体	エクソーム解析に伴う変異データ
30	IgA腎症のエクソーム解析	52検体	エクソーム解析に伴う変異データ (ゲノム支援)
31	慢性関節リウマチのエクソーム解析	59検体	エクソーム解析に伴う変異データ (ゲノム支援)
32	大腸癌のエクソーム解析	50検体	エクソーム解析に伴う変異データ (ゲノム支援)
33	永久歯欠如のエクソーム解析	48検体	エクソーム解析に伴う変異データ (ゲノム支援)
34	MODYのエクソーム解析	59検体	エクソーム解析に伴う変異データ (ゲノム支援)
35	脾類上皮嚢腫のエクソーム解析	8検体	エクソーム解析に伴う変異データ
36	家族性腎臓病のエクソーム解析	4検体	エクソーム解析に伴う変異データ
37	コーデン病のエクソーム解析	5検体	エクソーム解析に伴う変異データ
38	間質性腎炎のエクソーム解析	8検体	エクソーム解析に伴う変異データ
39	ウォルフラム症候群のエクソーム解析	4検体	エクソーム解析に伴う変異データ
40	卵巣癌のエクソーム解析	48検体	エクソーム解析に伴う変異データ
41	日本人健常者	400検体	エクソーム解析に伴う変異データ
42	多系統萎縮症のエクソーム	600検体	エクソーム解析に伴う変異データ (ゲノム支援)
43	日本人ゲノム全ゲノム配列	10検体	全ゲノム配列

NBDC ヒトデータ共有ガイドライン

ヒトデータを公開・共有するための運用ルールとしてのガイドラインを策定
<http://humandbs.biosciencedbc.jp/guidelines/data-sharing-guidelines>

公的資金を用いて産生されたヒトに関するデータ一般に適用
データの適正に管理しつつ、収集したデータをなるべく広く共有
公開の有無、アクセス制限のレベルによって以下の4つに分類

オープンデータ：アクセスに制限を設けることなく、利用することが可能な公開データ。頻度データ、統計解析結果、すでに発表された論文の参照データなど。

制限公開データ：利用者、利用目的等を明らかにしたうえで、関連研究に実績のある研究者が研究のために利用可能な公開データ。個人毎の解析結果など。

公開待機データ：論文発表や知的財産権取得等、データ提供者による成果公開の後、オープンデータあるいは制限公開データとして公開。

匿名化前・公開留保データ：各プロジェクトや実施機関が保有する匿名化前のデータ。

今後の展望

Human Variation DB の充実：次世代シーケンサーなどによる新規多型・変異情報と、文献からの多型・変異情報を情報量、機能ともに一層充実

がん体細胞変異・オミックスデータ等（菅野グループDB）との連携

変異の表現型への影響を解釈・予想できるような知識型DBへ発展

データの受け入れ、再配布に関し、NBDCによる新たなガイドラインに基づいた運用を継続

ゲノム支援、新学術、その他のプロジェクト、外部機関からの**NGS**データ登録、**SNP-GWAS**データ登録のための活動を継続

海外の関連するデータベースとの連携