マイクロアレイ染色体（CMA）検査の結果解釈に必要なデータベースの使用方法

埼玉県立小児医療センター　遺伝科
大場大樹

Webinarの構成

- イントロダクション
  - マイクロアレイ染色体検査（CMA）検査　超入門
  - BEDファイル形式データの取扱い
  - マイクロアレイ染色体検査の結果解釈の補助ソフトウェアツール（CAS）の使用方法
  - マイクロアレイ（CMA）検査の結果解釈に必要なデータベースの使用方法
- 実践編
  - 解析例１：CASを使用して疾患関連性を調べたLossの例
  - 解析例２：CASと各種データベースを使用して疾患関連性を調べたLossの例
  - 解析例３：CASと各種データベースを使用して疾患関連性を調べたGainの例
  - 解析例４：疾患関連性領域近傍にあるBenignのLossの例
- Advanced編
  - データベース「UCSC」使用方法
  - 解析例５：UCSCを用いて、Uncertain Significance と推定される例
  - インプリンティング疾患の解釈について

CMAの結果とその内容

① どの領域に：染色体領域、ゲノムポジション

② なにが　　：コピー数低下、コピー数上昇

ゲイン/ロス領域表
Gain/Loss Intervals Table

染色体 Chr	最小始点-終点(bp) 最大始点-終点(bp) Min Start-Stop(bp) Max Start-Stop(bp)	最小サイズ(kb) 最大サイズ(kb) Min Size(kb) Max Size(kb)	最小サイトバンド 最大サイトバンド Min Cytoband Max Cytoband	ゲイン/ロス Gain/Loss	アノテーション Annotations
chr1	746608-3375421 647288-3404363	2,628.814 2,757.076	p36.33 - p36.32 p36.33 - p36.32	0.564656	FAM87B, LINC00115, LINC01128, FAM41C, LINC02593, SAMD11, NOC2L, KLHL17, PLEKHN1, PERM1, HES4, ISG15, AGRN, LOC100288175, RNF223, C1orf159, LINC01342, MIR200B, MIR200A, MIR429, TTLL10, TNFRSF18, TNFRSF4, SDF4, B3GALT6, C1QTNF12, UBE2J2, SCNN1D, ACAP3, MIR6726, SNORD167, PUSL1, INTS11, MIR6727, CPTP, TAS1R3, DVL1, MIR6808, MXRA8, AURKAIP1, CCNL2, MRPL20-AS1, MRPL20, ANKRD65, TMEM88B, LINC01770, VWA1, ATAD3C, ATAD3B, ATAD3A, TMEM240, SSU72, FNDC10, LOC105378586, MIB2, MMP23B, MMP23A, CDK11B, SLC35E2B, CDK11A, SLC35E2A, NADK, GNB1, CALML6, TMEM52, CFAP74, GABRD, LOC105378591, PRKCZ, PRKCZ-AS1, FAAP20, SKI, MORN1, LOC100129534, RER1, PEX10, PLCH2, PANK4, HES5, TNFRSF14-AS1, TNFRSF14, LOC100996583, PRXL2B, MMEL1, TTC34, ACTRT2, PRDM16-DT, PRDM16, MIR4251, ARHGEF16

① ②

CMAの結果解釈に必要な理解

- **ハプロ不全（Haploinsufficiency：HI）**

  領域や遺伝子が欠失することによる疾患発症

- **重複感受性（Triplosensitivity：TS）**

  領域や遺伝子が重複することによる疾患発症

ハプロ不全（HI）①

- **Haploinsufficiency Score（HI Score）**

  ClinGen（Clinical Genome Resource）が提唱するハプロ不全の指標

HI/TS Score	Description	Predicted Pathogenicity
3	Sufficient evidence	Pathogenic
2	Emerging evidence	Likely Pathogenic
1	little evidence	Uncertain
0	no evidence	Uncetain or Likely Benign
40	Dosage sensitivity unlikely	Benign
30	Gene associated with autosomal recessive phenotype	

ハプロ不全（HI）②

- **pLI, LOEUF**

  WESやWGSデータから算出されたHIの指標

- **%HI**

  文献やCNVデータなどを基にしたHIの指標

HIを示唆するCut off値

- **pLI ≧ 0.9**
- **LOEUF < 0.35**
- **%HI ≦ 10%**

pLI, LOEUF, %HIはHIを確定するものではない

重複感受性（TS）

➢ Triplosensitivity Score（TS Score）

ClinGen（Clinical Genome Resource）が提唱する重複感受性の指標

HI/TS Score	Description	Predicted Pathogenicity
3	Sufficient evidence	Pathogenic
2	Emerging evidence	Likely Pathogenic
1	little evidence	Uncertain
0	no evidence	Uncetain or Likely Benign
40	Dosage sensitivity unlikely	Benign
30	Gene associated with autosomal recessive phenotype	

HI関連データ	HIの適用対象 遺伝子	領域
HI Score（ClinGen）	○	○
pLI, LOEUF	○	×
%HI	○	×

TS関連データ	TSの適用対象 遺伝子	領域
TS Score（ClinGen）	○ *LMNB1, PLP1*のみ	○

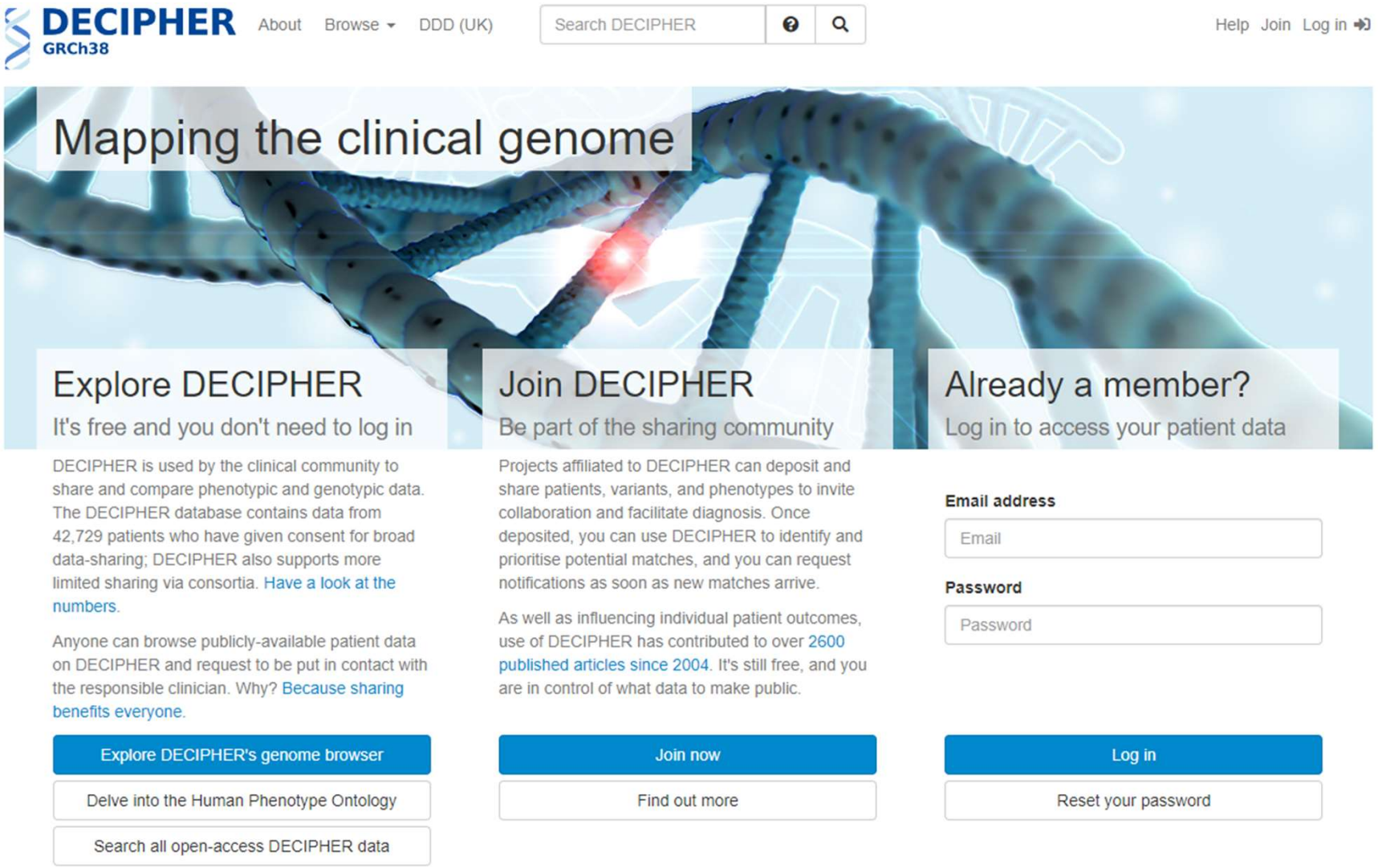
※pLI, LOEUF, %HIと対比可能なTS関連データはない

HI, TSの検討に有用なデータベース

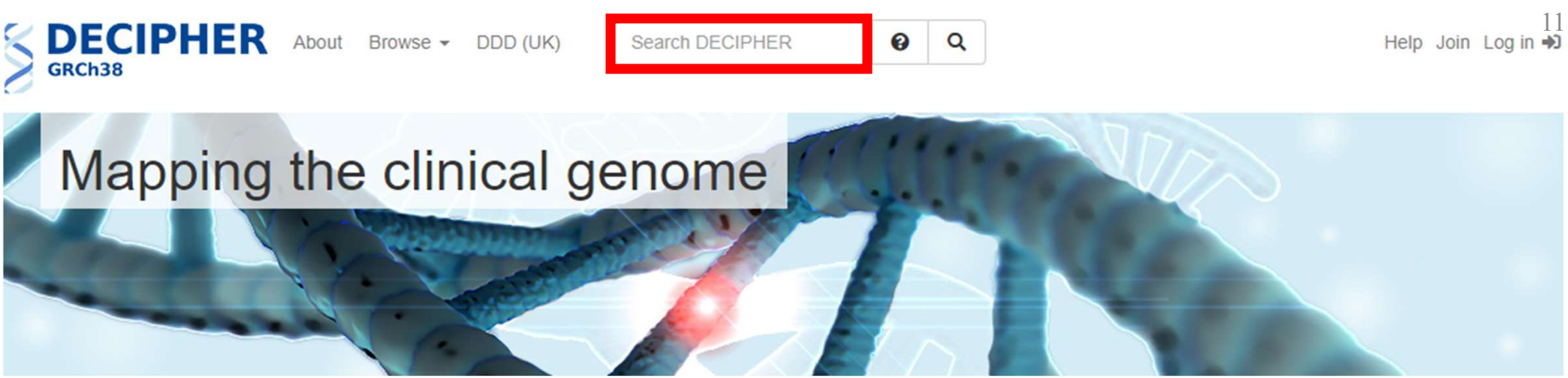
- DECIPHER
- ClinGen
- UCSC Genome Browser

DECIPHER

DatabasE of genomiC varIation and Phenotype in Humans using Ensembl Resources



https://www.deciphergenomics.org/ （2022.5.2アクセス）



- **赤枠に検出されたCNV領域を入力する**

（例）BEDファイルのデータ

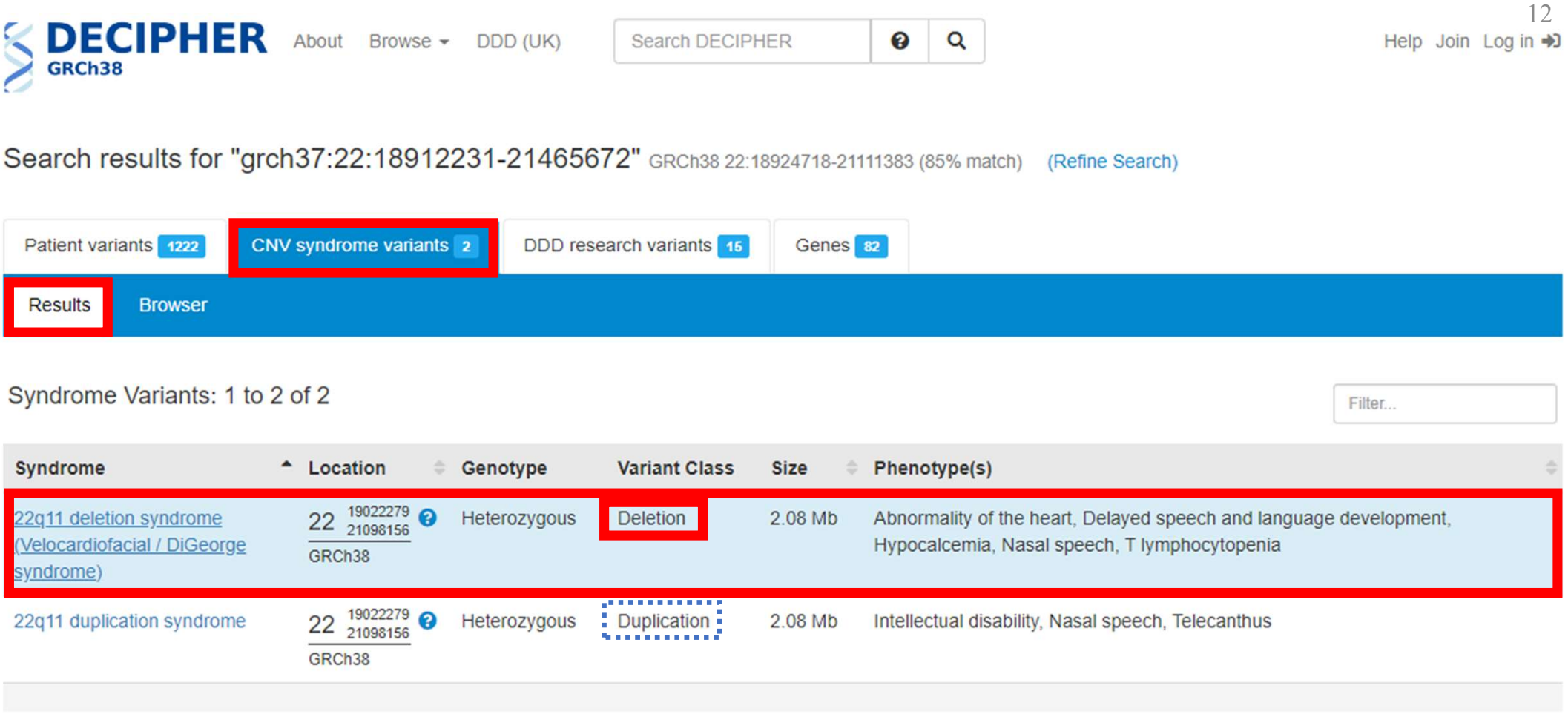
22　18912231　21465672　Loss

入力方法

検出したCNVの染色体番号

“grch37:22:18912231-21465672”と記入

使用した標準配列　　CNVのGenome position



CNV syndrome variantsのResultsに22q11.2 deletion syndromeの表示

22q11.2欠失症候群

※下段は22q11.2 duplication syndrome

22q11.2重複症候群

欠失・重複を
間違えない！！

DECIPHER GRCh38 About Browse DDD (UK) Search DECIPHER Help Join Log in

Search results for "grch37:22:18912231-21465672" GRCh38 22:18924718-21111383 (85% match) (Refine Search)

Patient variants 1226 | CNV syndrome variants 2 | DDD research variants 15 | Genes 82

Genes: 21 to 30 of 82

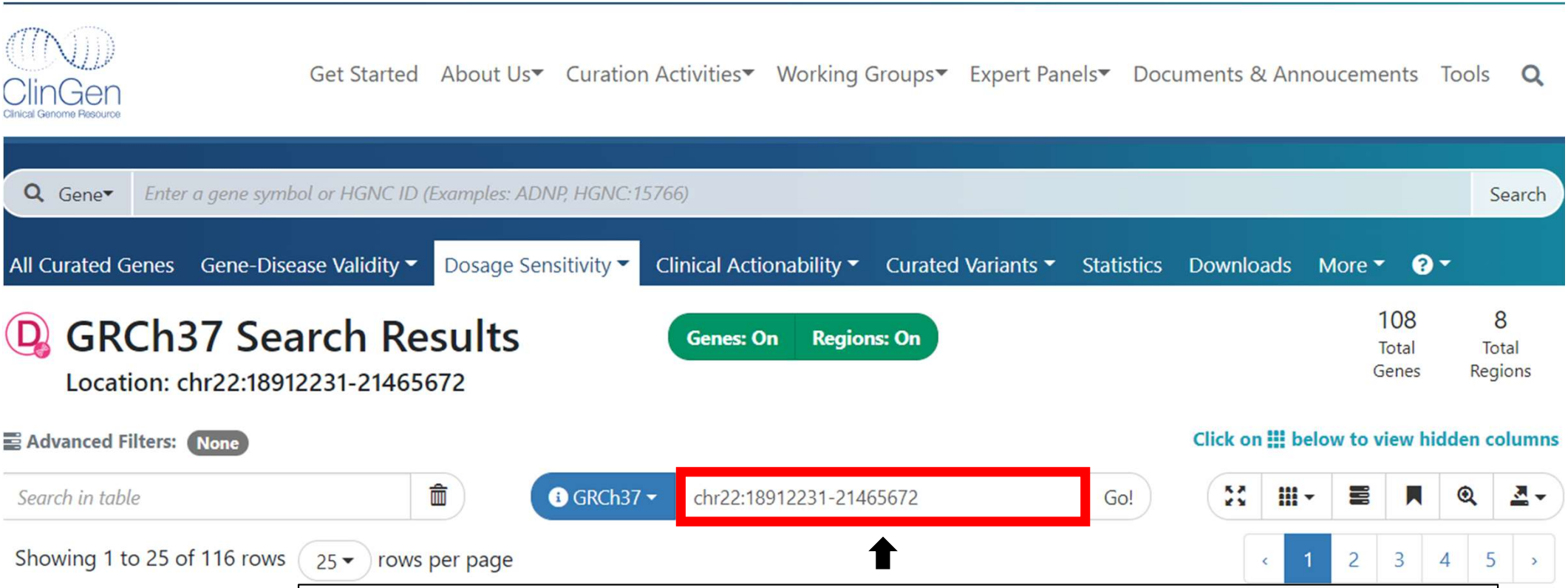
Show: All genes Filter...

Name / Description	Location	pLI	LOEUF	sHet	%HI	GenCC	OMIM / Morbid	G2P	ClinGen	Links
FAM230G family with sequence similarity 230 member G	22 20338805 20354372	0.40	1.69	-	-	-	-	-	-	View
FAM246C family with sequence similarity 246 member C (gene/pseudogene)	22 19029524 19031242	-	-	-	-	-	-	-	-	View
GNB1L G protein subunit beta 1 like	22 19783223 19854939	0.00	1.05	0.005	48.56	-	OMIM	-	-	View
GP1BB glycoprotein Ib platelet subunit beta	22 19723539 19724771	0.51	1.33	-	78.00	Definitive: 1 Supportive: 2	OMIM Morbid	-	Definitive: AR Haploinsufficiency: 30	View
GSC2 goosecoid homeobox 2	22 19146993 19150292	0.00	1.91	-	65.74	-	OMIM	-	-	View
HIRA histone cell cycle regulator	22 19330698 19447450	1.00	0.14	0.380	19.47	-	OMIM	Limited: Monoallelic	Haploinsufficiency: 0 Triplosensitivity: 0	View

Geneのタブでは各遺伝子のHI/TS関連データを表示

ClinGen：Dosage Sensitivity

Clinical Genome Resource



chr22:18912231-21465672と入力して”Go!”をクリック
（染色体番号のところに”chr”を必ず入れる。）

https://search.clinicalgenome.org/kb/gene-dosage?page （2022.5.7アクセス）

ClinGen
Clinical Genome Resource

Get Started | About Us | Curation Activities | Working Groups | Expert Panels | Documents & Annoucements | Tools

Gene | Enter a gene symbol or HGNC ID (Examples: ADNP, HGNC:15766) | Search

All Curated Genes | Gene-Disease Validity | Dosage Sensitivity | Clinical Actionability | Curated Variants | Statistics | Downloads | More

GRCh37 Search Results

Location: chr22:18912231-21465672

Genes: On | Regions: On

108 Total Genes | 8 Total Regions

領域に含まれる遺伝子 ＋ HI/TS関連データのある領域がすべて表示される

Click on ⋮⋮⋮ below to view hidden columns

Showing 1 to [illegible] of 116 rows | 25 rows per page

‹ 1 2 3 4 5 ›

Gene/Region	GRCh37	HI Score	TS Score	OMIM	Morbid	%HI	pLI	LOEUF	Report
G O PRODH	22 18900287 18924066	Not Yet Evaluated	Not Yet Evaluated	✓	✓	45.38	0	1.1	Reopened
R C 22q11.2 recurrent (DGS/VCFS) region (proximal, A-B) (includes TBX1)	22 18912231 20287208	3 (Sufficient Evidence)	3 (Sufficient Evidence)			-	-	-	Complete
R C 22q11.2 recurrent (DGS/VCFS) region (proximal, A-D) (includes TBX1)	22 18912231 21465672	3 (Sufficient Evidence)	3 (Sufficient Evidence)			-	-	-	Complete
G C DGCR5	22 18958011 19018755	Not Yet Evaluated	Not Yet Evaluated	✓		-	-	-	Awaiting Review
P C LOC100287576	22 18985531 18987264	-1 (Pseudogene)	-1 (Pseudogene)			-	-	-	Not Reviewable

GRCh37 Search Results

Location: chr22:18912231-21465672

Genes: Off | Regions: On

108 Total Genes　8 Total Regions

Genesをクリックしてoffにするとregionのみの表記となる

Showing 1 to 8 of 8 rows

Gene/Region	GRCh37		HI Score	TS Score	OMIM	Morbid	%HI	pLI	LOEUF	Report
22q11.2 recurrent (DGS/VCFS) region (proximal, A-B) (includes TBX1)	22	18912231 20287208	3 (Sufficient Evidence)	3 (Sufficient Evidence)			-	-	-	Complete
22q11.2 recurrent (DGS/VCFS) region (proximal, A-D) (includes TBX1)	22	18912231 21465672	3 (Sufficient Evidence)	3 (Sufficient Evidence)			-	-	-	Complete
22q11.21 population region (gnomAD-SV_v2.1_DEL_22_181446)	22	20366424 20371500	40 (Dosage Sensitivity Unlikely)	0 (No Evidence)			-			
22q11.21 population region (DGV_Gold_Standard_June_2021_gssvL76981)	22	20428070 20499197	40 (Dosage Sensitivity Unlikely)	0 (No Evidence)			-	-	-	Complete
22q11.21 population region (gnomAD-SV_v2.1_DEL_22_181452)	22	20451599 20457600	40 (Dosage Sensitivity Unlikely)	0 (No Evidence)			-	-	-	Complete
22q11.21 population region (gnomAD-SV_v2.1_DEL_22_181459)	22	20693999 20699000	40 (Dosage Sensitivity Unlikely)	0 (No Evidence)			-	-	-	Complete
22q11.2 recurrent region (central, B/C-D) (includes CRKL)	22	20731986 21465672	2 (Emerging Evidence)	1 (Little Evidence)			-	-	-	Complete
22q11.2 recurrent region (central, C-D) (includes CRKL)	22	21092338 21465672	Not Yet Evaluated	Not Yet Evaluated			-	-	-	Under Secondary Rev

詳細はここをクリック

Showing 1 to 8 of 8 rows

Haploinsufficiency (HI) Score Details

HI Score: 3

HI Evidence Strength: **Sufficient Evidence for Haploinsufficiency** (Disclaimer)

HI Disease: DiGeorge syndrome Monarch

HI Evidence:

PUBMED: 20301696

McDonald-McGinn, Emanuel and Zackai. GeneReviews: 22q11.2 Deletion Syndrome.

PUBMED: 27189754

McDonald-McGinn et al. (2015) reviewed clinical findings in 22q11.2 deletion syndrome. See also PMID 21200182.

HI Evidence Comments:

Deletion of the 22q11.2 proximal (A-D) region is associated with DiGeorge/Velocardiofacial (DGS/VCFS) syndrome. Clinical findings are variable, but typically include congenital heart disease (particularly conotruncal malformations), palatal abnormalities (particularly velopharyngeal incompetence, cleft palate and bifid uvula), characteristic facial features, DD/ID, behavior problems, immune deficiency, and hypocalcemia. Most 22q11.2 (DGS/VCFS) deletions (>90 percent) are de novo. This deletion is enriched in the clinical population.

Additional relevant literature is summarized below:

Case-control studies:

PMID 25217958:
Coe et al. (2014): In a large-scale case-control comparison study of the relative prevalence of copy number variants in children with ID/DD, MCA, and other developmental phenotypes compared to controls, deletion of the recurrent 22q11.2 (DGS/VCFS) A-B region, the presumed critical region between LCR22A and -D, was observed in 158/29,085 cases versus 0/19,584 controls (p=3.97E-36; LR: Inf, CI: 43.9 to Inf).

GRCh37 Search Results

Location: chr22:18912231-21465672

Genes: On | Regions: Off

108 Total Genes　8 Total Regions

Regionをクリックしてoffにするとgenesのみの表記となる

Advanced Filters: None　　w hidden columns

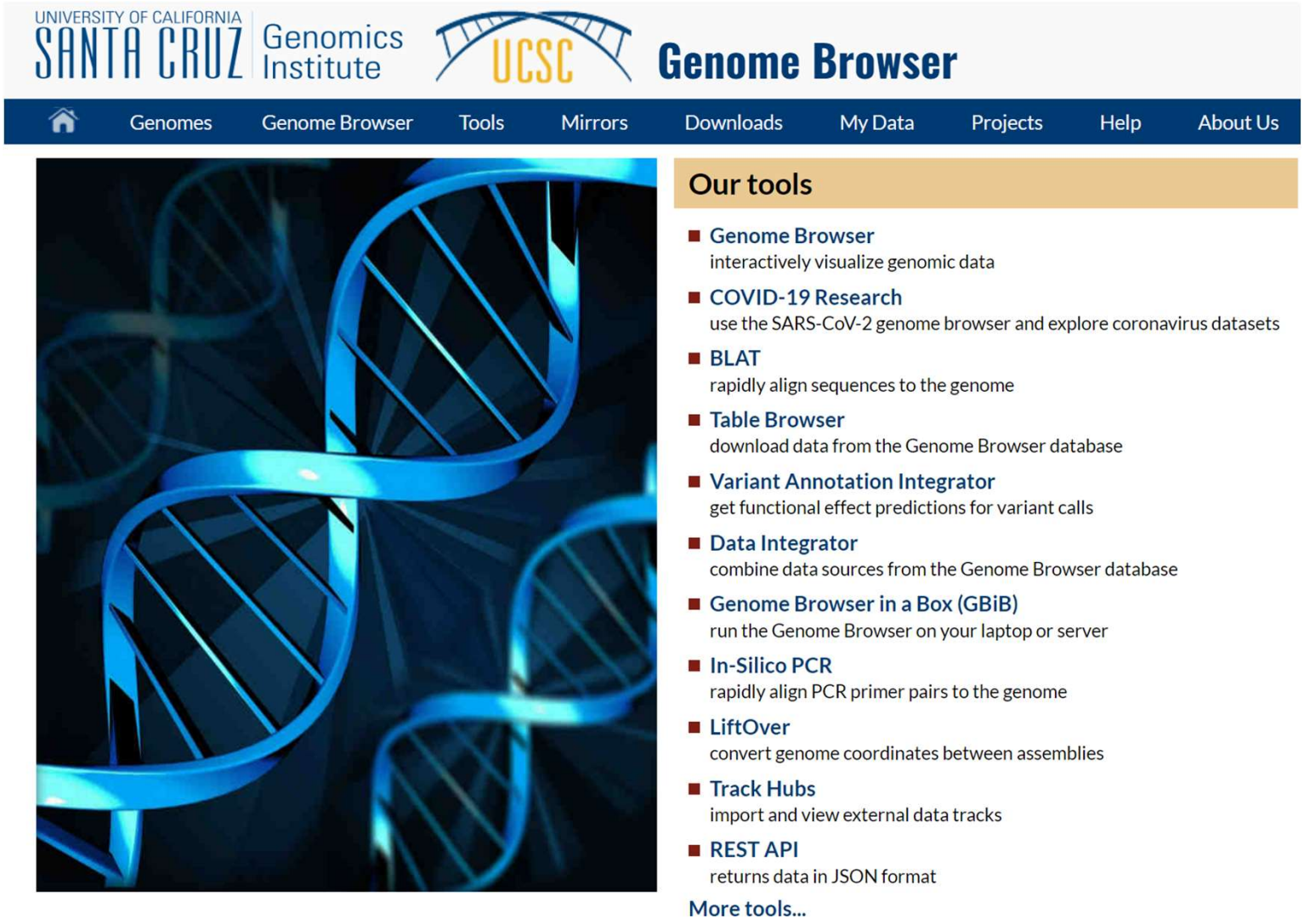
Search in table　　GRCh37　Enter cytoband or genomic coordinates　Go!

Showing 1 to 25 of 108 rows　25 rows per page　‹ 1 2 3 4 5 ›

Gene/Region	GRCh37		HI Score	TS Score	OMIM	Morbid	%HI	pLI	LOEUF	Report
G O PRODH	22	18900287 18924066	Not Yet Evaluated	Not Yet Evaluated	✓	✓	45.38	0	1.1	Reopened
G C DGCR5	22	18958011 19018755	Not Yet Evaluated	Not Yet Evaluated	✓		-	-	-	Awaiting Review
P C LOC100287576	22	18985531 18987264	-1 (Pseudogene)	-1 (Pseudogene)			-	-	-	Not Reviewable
P C LOC100506454	22	19001864 19018742	-1 (Pseudogene)	-1 (Pseudogene)			-	-	-	Not Reviewable
G C DGCR5	22	19005347 19007761	Not Yet Evaluated	Not Yet Evaluated	✓		-	-	-	Awaiting Review
G C DGCR5	22	19010137 19011063	Not Yet Evaluated	Not Yet Evaluated	✓		-	-	-	Awaiting Review

UCSC Genome Browser

University of California
SANTA CRUZ



https://http://www.genome.ucsc.edu/ （2022.5.7アクセス）

UCSC Genome Browser使用の流れ

1. Custom TracksにCNVデータを入力

CNV領域が表示される

- ✓ Zoom Outして領域全体を見渡しやすくする
- ✓ HighlightでCNV領域を見やすくする

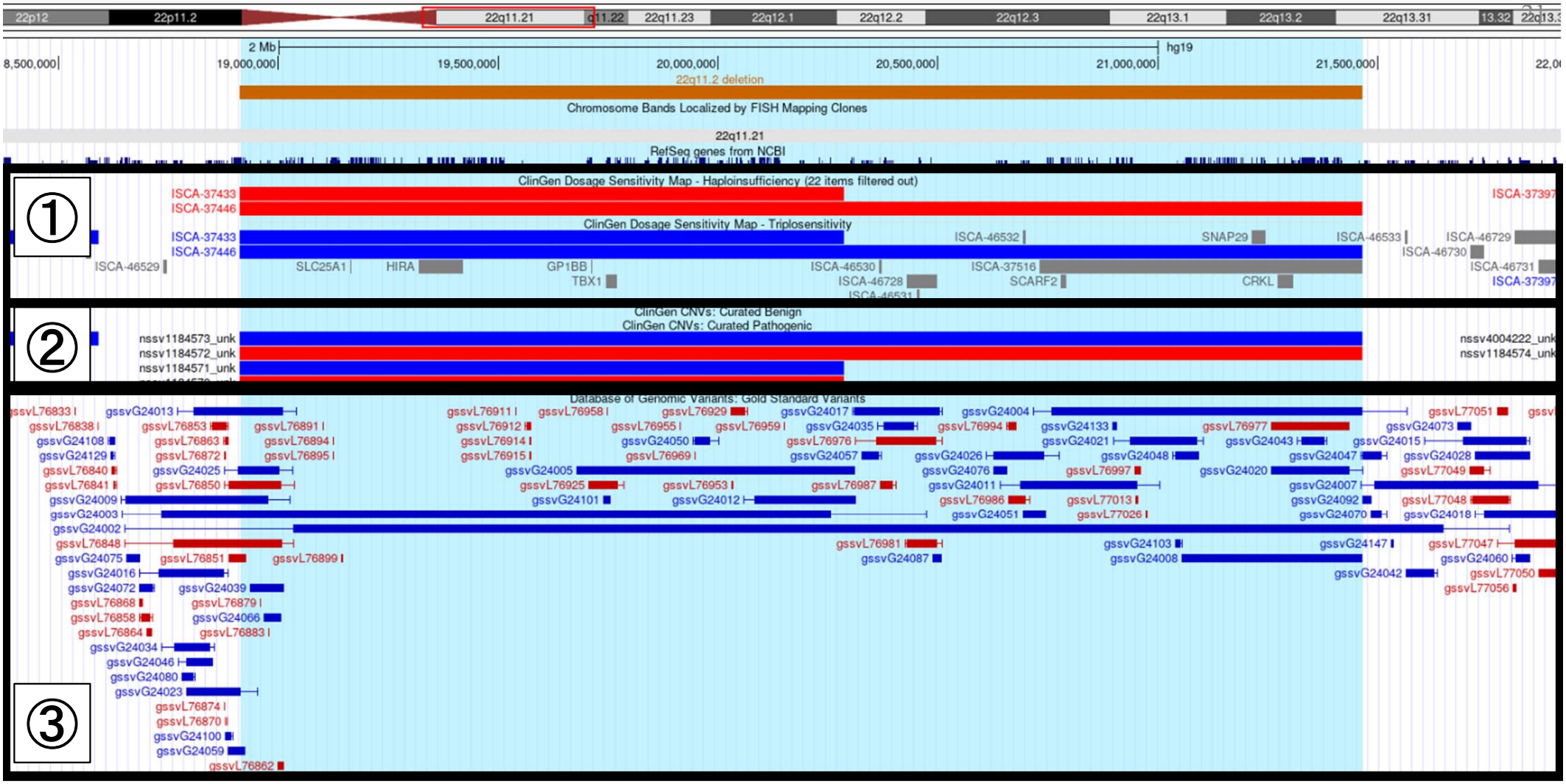
2. 必要な情報表示の選択をする

CNVの病原性評価に有用な表示情報（一例）

- ✓ Base position, Chromosome Band
- ✓ ClinGen（HI/TS Score）
- ✓ ClinGen CNVs（Curated Pathogenic/Benign）
- ✓ DGV Gold Standard（一般集団CNV→”Benign示唆”）

3. 表示内容から病原性を判断する

※詳細はAdvanced編 “データベース「UCSC」使用方法”を参照



① Dosage Sensitivityに関連した情報：ClinGen

② Curated Pathogenic/Benignの情報：CliGen CNVs

③ 一般集団CNVデータベースの情報：DGV Gold Standard

	利点	欠点（留意点）
DECIPHER	➢ HI, TSの根拠がわかりやすい • 数値化されたデータ （HI/TS Score, pLI, LOEUF, %HI） ⬇ 病原性判断がスムーズ ⬆	➢ GRCh38を基準としている • CMA結果はGRCh37準拠 • データ入力時に”grch37”のが必須 ➢ 入力したCNVを一部ないしすべて含むすべてのCNV関連情報が表示される（表示内容の中から取捨選択が必要） ➢ 視覚的にCNV範囲が分かりにくい
ClinGen Dosage Sensitivity	➢ HI, TSの根拠がわかりやすい • 数値化されたデータ （HI/TS Score, pLI, LOEUF, %HI） ➢ GRCh37が選択可能	➢ 入力したCNVを一部ないしすべて含むすべてのCNV関連情報が表示される（表示内容の中から取捨選択が必要） ➢ 視覚的にCNV範囲が分かりにくい
UCSC Genome Browser	➢ 視覚的にCNVを表示できる（CNV） DECIPHERやClinGenで生じる表示特性による誤診が防ぎやすい ➢ HI, TS関連情報のほか、一般集団データベースの情報も表示可能（たくさんの情報を統合表示できる）	➢ 表示したいデータの選択（操作）が煩雑 ➢ 視覚表示に慣れるまでに時間がかかる（通常の表示ではScore表示がない）

追記

ClinGen, DECIPHER, UCSCでは確立されたHI/TS領域の病原性評価が可能

ただし、

「病原性が確立されていないCNVの検出」もありうる

「類似症例の文献検索」「報告症例との対比」

CMA結果解釈の留意点

- コピー数低下（Loss）なのか増加（Gain）なのかを間違えない
- 対応する症候群（疾患）がハプロ不全（欠失）、重複感受性どちらで生じるのかも必ず確認する

Loss, Gainの間違いは誤診に直結！

- データベースはそれぞれ特性あり。最初は「使いやすい」「使い慣れた」からの使用でOK
- 病原性が確立されたCNV以外の評価には文献検索が必要