

# ゲノムワイド関連解析（GWAS）の基礎

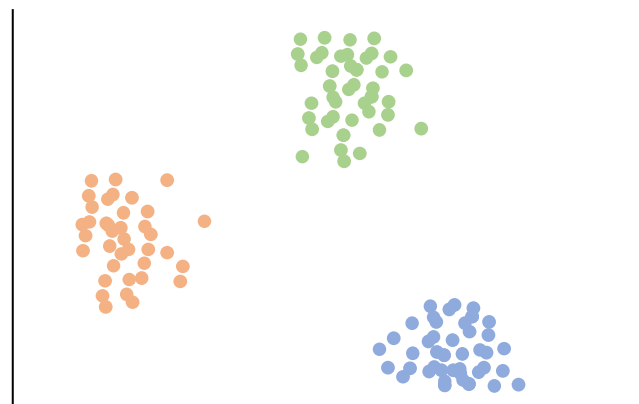
フィルジエン株式会社 バイオインフォマティクス部  
(biosupport@filgen.jp)

## GWAS ワークフロー

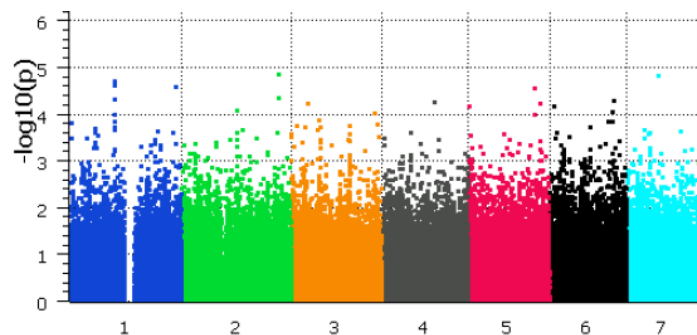
### ✓ SNPデータ

	SNP1	SNP2	SNP3
Sample1	A/G	C/T	G/G
Sample2	A/G	C/T	G/G
Sample3	A/A	C/T	G/G
Sample4	A/G	C/T	G/G
Sample5	A/G	T/T	G/G

### ✓ クオリティコントロール



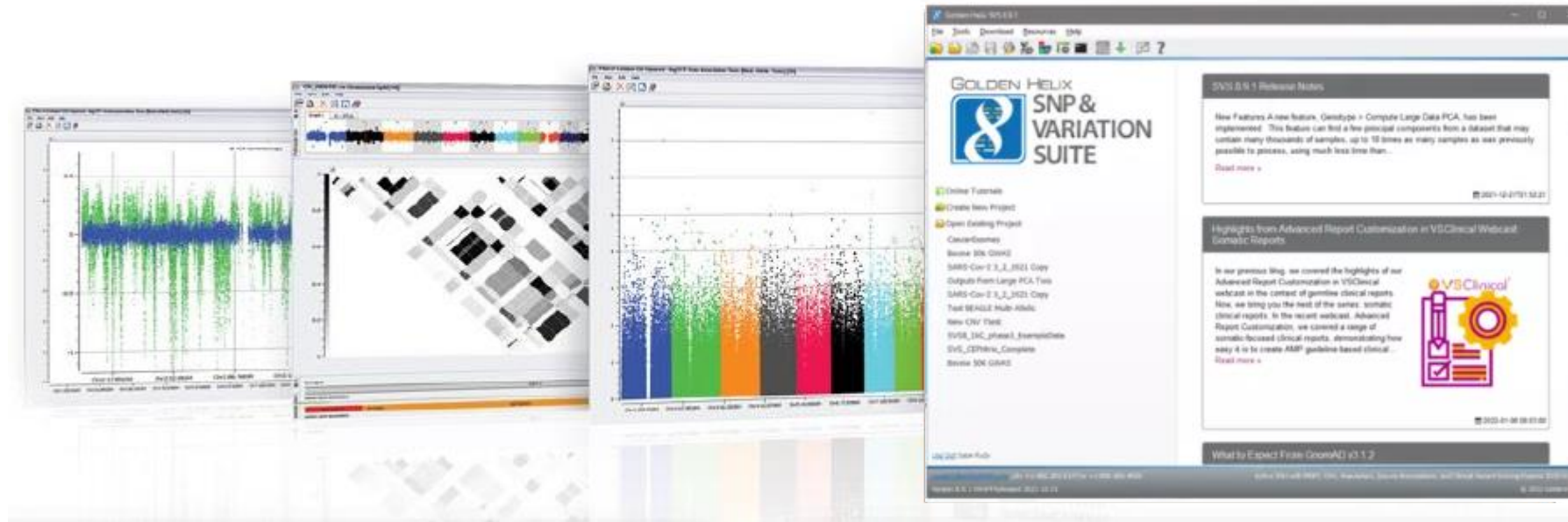
### ✓ 可視化



### ✓ 関連解析

	A/A	A/G	G/G
Case	96	165	217
Control	236	110	89

- おもに多因子疾患における、症例対照研究やコホート研究で利用されるゲノムワイド関連解析（GWAS）では、SNPデータに様々なデータ処理を行いながら解析を行う必要がある
- Golden Helix社SNP & Variation Suite (SVS) では、一般的なGWAS解析に必要な機能がすべて搭載されていて、簡単な操作で実行できる

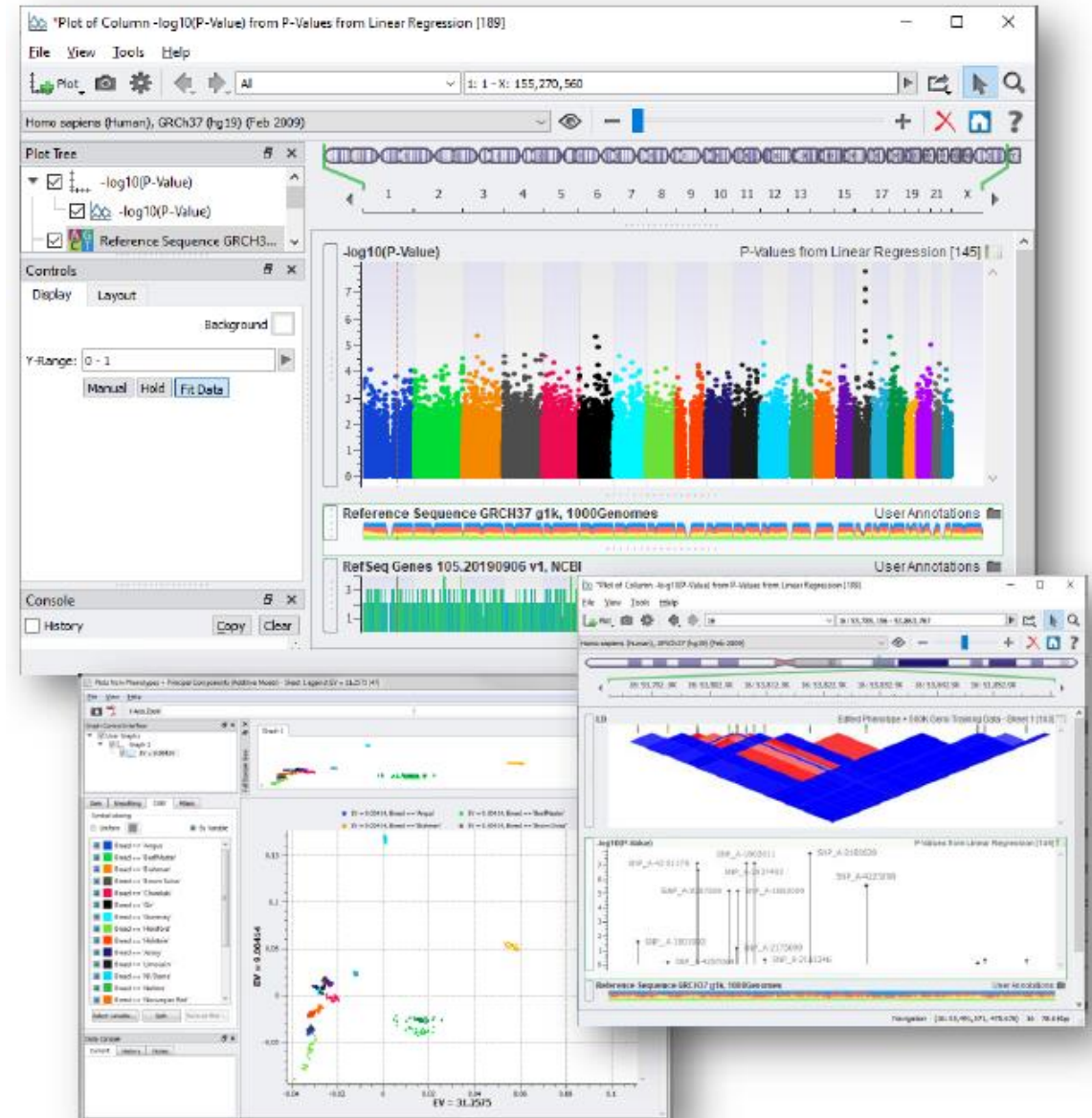
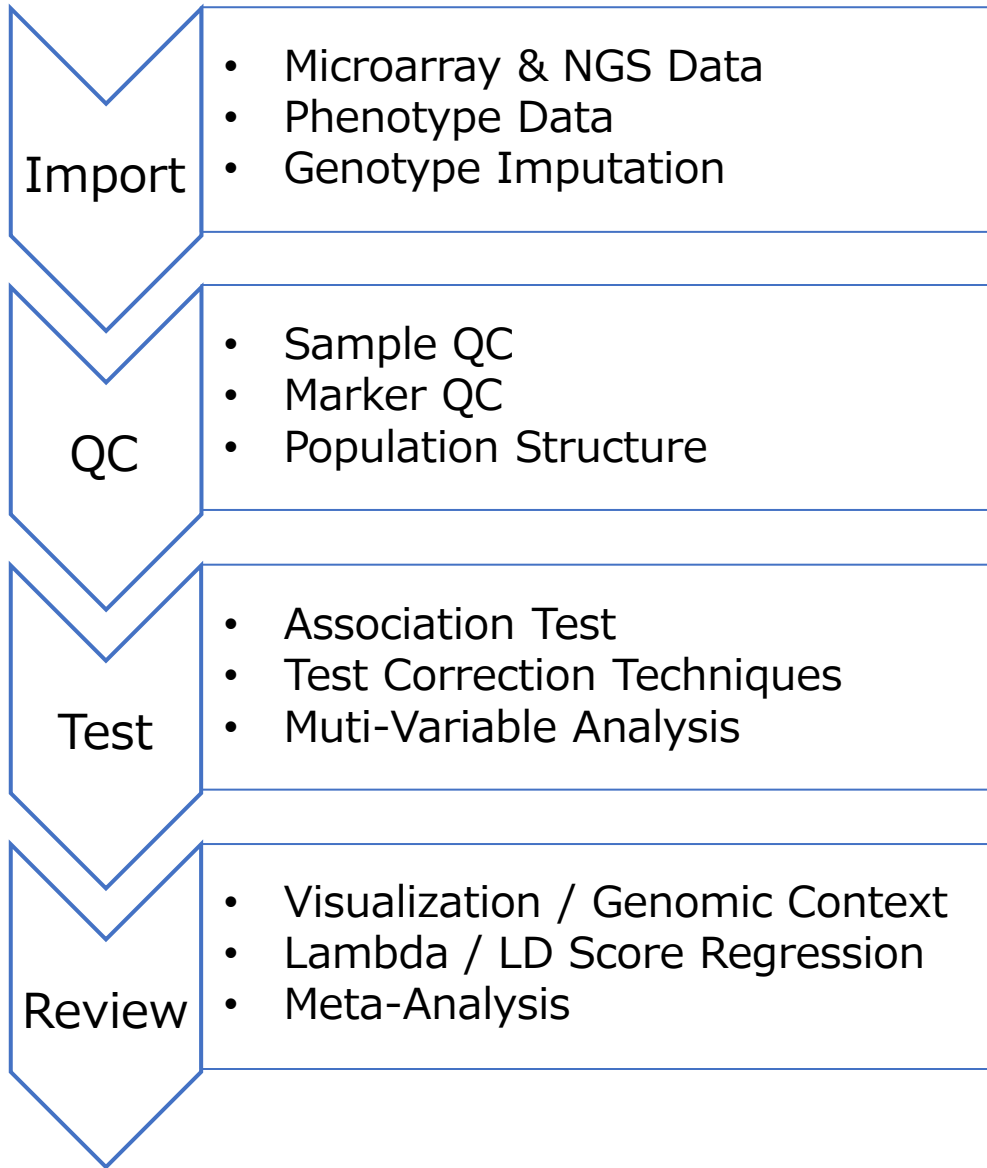


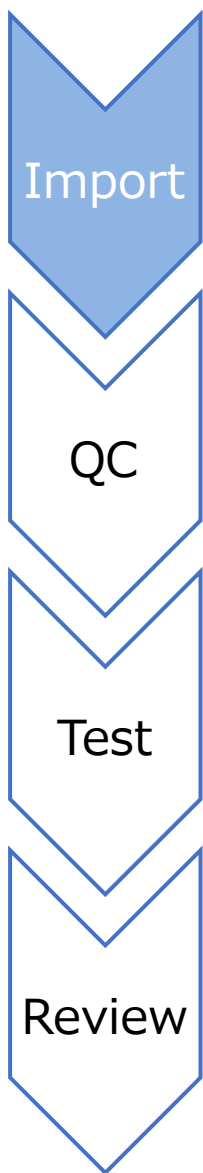
## Core Features

- Powerful Data Management
- Rich Visualizations (GenomeBrowse)
- Robust Static
- Flexible

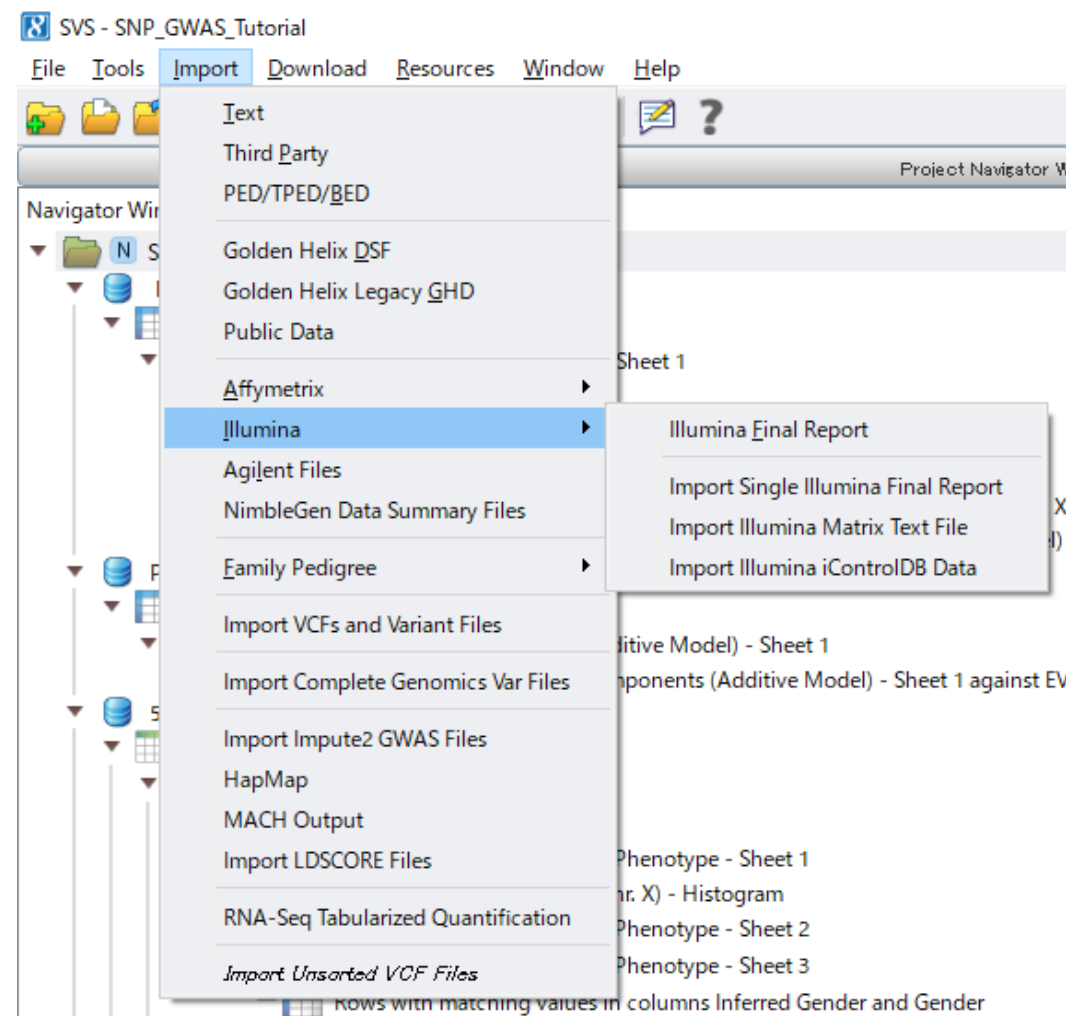
## Applications

- Genotype Analysis
- Agrigenomics Analysis
- DNA/RNA Sequence Analysis
- CNV Analysis





- Microarray Data
  - Affymetrix (e.g. CHP)
  - Illumina (e.g. Final Report, Matrix Text)
  - Agilent
  - NimbleGen
- NGS Data
  - VCF
- Phenotype Data
  - Text
  - Third Party (e.g. Excel)
- Others
  - PED/TPED/BED
  - Family Pedigree



## ✓ SNPデータのテキストファイル

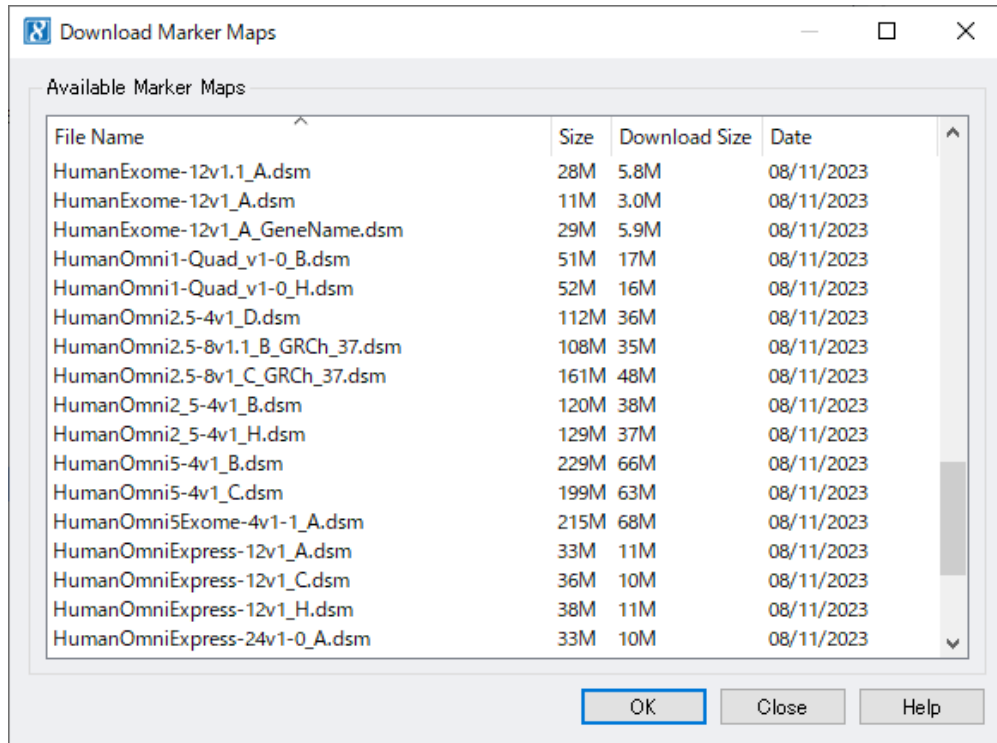
	A	B	C	D	E	F
1	Sample	SNP_A-1909444	SNP_A-4303947	SNP_A-1886933	SNP_A-2236359	SNP_A-2205441
2	GSM233256	T_T	A_A	G_G	T_T	C_C
3	GSM233258	C_T	A_A	G_G	T_T	C_C
4	GSM233260	C_C	?_?	A_A	T_T	C_C
5	GSM233262	C_T	A_G	A_G	T_T	C_C
6	GSM233264	T_T	A_A	G_G	T_T	C_C
7	GSM233266	T_T	A_A	G_G	T_T	C_C
8	GSM233268	C_T	A_G	A_G	T_T	C_C
9	GSM233270	T_T	A_A	G_G	T_T	C_C
10	GSM233272	T_T	A_A	G_G	T_T	C_C
11	GSM233274	T_T				
12	GSM233276	T_T				
13	GSM233278	T_T				
14	GSM233280	T_T				
15	GSM233282	T_T				
16	GSM233284	C_T				
17	GSM233286	T_T				
18	GSM233288	C_T				
19	GSM233290	C_C				
20	GSM233292	T_T				

- 多くの場合は、SNPと表現型それぞれのタブ区切りテキストファイルを作成し、SVSにインポートして使用
- SNPデータは、AB（メジャーアレルとマイナーアレル）形式のものを利用することも可能
- 表現型データには、バイアス補正などに利用する共変量も含めることが可能

## ✓ 表現型データのテキストファイル

	A	B	C	D	E	F	G
1	Sample	Phenotype 1	Phenotype 2	Phenotype 3	SBP	BMI	Gender
2	GSM233256	1	8.96	1	128	29.9	M
3	GSM233258	0	12.18	0	137	32.8	M
4	GSM233260	1	11.45	0	138	31.6	F
5	GSM233262	0	11.01	1	117	27.4	M
6	GSM233264	0	9.25	0	140	29.7	M
7	GSM233266	1	6.85	0	113	27.7	M
8	GSM233268	1	8.77	1	121	27.7	F
9	GSM233270	1	11.65	0	125	29.2	M
10	GSM233272	0	13.67	1	120	29	M
11	GSM233274	0	10.61	0	149	34.1	M
12	GSM233276	0	10.82	1	88	20	M
13	GSM233278	1	9.62	0	131	31.2	M
14	GSM233280	1	9.02	1	124	26.2	M
15	GSM233282	0	7.94	0	142	31.3	M
16	GSM233284	0	11.73	1	111	22.8	M
17	GSM233286	0	12.81	1	114	26.1	F
18	GSM233288	0	9.46	1	107	23.5	M
19	GSM233290	1	8.55	0	146	37.4	M
20	GSM233292	1	13.28	0	122	27.6	M





- SNPデータがマイクロアレイより得られたもの場合は、別途プローブごとのマーカーマップデータが必要
- 既製品の一部のマイクロアレイでは、SVSのダウンロードマネージャーよりマーカーマップデータをダウンロード可能
- SVSダウンロードマネージャーからダウンロードできないものは、マイクロアレイメーカー提供のマーカーマップManifestファイルなどを利用可能

CSV形式のManifestファイルをダウンロード



## Infinium Asian Screening Array v1.0 Product Files

Support Center:  
Infinium Asian Screening Array v1.0 Product Files

Manifest, cluster, and LIMS product descriptor files for the Infinium Asian Screening Array v1.0.

### Files

NAME	TYPE & SIZE	DATE
<a href="#">Infinium Asian Screening Array v1.0 Manifest File (CSV Format - GRCh37)</a>	ZIP(70 MB)	May 7, 2018

## ✓ インポートされたSNP+マーカーマップデータ

Unsort		G 1	G 2	G 3	G 4	G 5	G 6	G 7
Map	Sample	SNP A-1909444	SNP A-4303947	SNP A-1886933	SNP A-2236359	SNP A-2205441	SNP A-2116190	SNP A-4291020
	Chromosome	1	1	1	1	1	1	1
	Position	752566	779322	785989	792480	799463	1003629	1097335
	dbSNP RS ID	rs3094315	rs4040617	rs2980300	rs2905036	rs4245756	rs4075116	rs9442385
	Associated Gene	?	?	?	?	?	?	?
	Cytoband	p36.33	p36.33	p36.33	p36.33	p36.33	p36.33	p36.33
	Reference Alleles A/B	[C/T]	[A/G]	[A/G]	[C/T]	[C/T]	[A/G]	[G/T]
	Top Alleles	[G/A]	[A/G]	[T/C]	[C/T]	[C/T]	[T/C]	[G/T]
	Bottom Alleles	[C/T]	[T/C]	[A/G]	[G/A]	[G/A]	[A/G]	[C/A]
	Strand	-	+	-	+	+	-	+
	Strand Versus dbSNP	same	same	reverse	reverse	same	same	same
1	GSM233256	T_T	A_A	G_G	T_T	C_C	A_A	G_G
2	GSM233258	C_T	A_A	G_G	T_T	C_C	A_A	G_G
3	GSM233260	C_C	?_?	A_A	T_T	C_C	A_A	G_G
4	GSM233262	C_T	A_G	A_G	T_T	C_C	A_G	G_T
5	GSM233264	T_T	A_A	G_G	T_T	C_C	A_A	G_G
6	GSM233266	T_T	A_A	G_G	T_T	C_C	G_G	G_G
7	GSM233268	C_T	A_G	A_G	T_T	C_C	A_G	G_G
8	GSM233270	T_T	A_A	G_G	T_T	C_C	A_A	G_G
9	GSM233272	T_T	A_A	G_G	T_T	C_C	A_A	G_G
10	GSM233274	T_T	A_A	G_G	T_T	C_C	A_A	G_G
11	GSM233276	T_T	A_A	G_G	T_T	C_C	A_A	G_G
12	GSM233278	T_T	A_A	G_G	T_T	C_C	A_A	G_G
13	GSM233280	T_T	A_A	A_G	T_T	C_C	A_G	G_G
14	GSM233282	T_T	A_A	G_G	T_T	C_C	A_A	G_G
15	GSM233284	C_T	A_G	A_G	T_T	C_C	A_G	G_G

- SNPと表現型のデータファイルをインポートすると、それぞれシートが作成される
- SNPデータのシートには、マーカーマップデータを加えることが可能

## ✓ インポートされた表現型データ

Unsort		B 1	R 2	B 3	I 4	R 5	C 6
Map	Sample	Phenotype 1	Phenotype 2	Phenotype 3	SBP	BMI	Gender
1	GSM233256	1	8.96	1	128	29.9	M
2	GSM233258	0	12.18	0	137	32.8	M
3	GSM233260	1	11.45	0	138	31.6	F
4	GSM233262	0	11.01	1	117	27.4	M
5	GSM233264	0	9.25	0	140	29.7	M
6	GSM233266	1	6.85	0	113	27.7	M
7	GSM233268	1	8.77	1	121	27.7	F
8	GSM233270	1	11.65	0	125	29.2	M
9	GSM233272	0	13.67	1	120	29	M
10	GSM233274	0	10.61	0	149	34.1	M
11	GSM233276	0	10.82	1	88	20	M
12	GSM233278	1	9.62	0	131	31.2	M
13	GSM233280	1	9.02	1	124	26.2	M
14	GSM233282	0	7.94	0	142	31.3	M
15	GSM233284	0	11.73	1	111	22.8	M





## ✓ Sample QC

- Genotype Statistics by Sample
  - Call Rate
  - 性別不一致のチェック

- LD Pruning
  - 連鎖不平衡SNPの刈り込み

- Principle Component Analysis
  - 集団の遺伝的背景の推定

- Identity by Descent
  - 隠れた血縁関係の推定

## ✓ Marker QC

- Genotype Filtering by Marker
  - Call Rate
  - ハーディー・ワインベルグ平衡
  - マイナーアレル頻度

共変量に使用

GWASの実行

# Genotype Statistics by Sample

Genotype Statistics by Sample

(No variable is set as dependent.)

Genotype Count Statistics

NOTE: Call rate and heterozygosity are always output.

Number and fraction of genotypes with a minor allele (as determined from sample data)

Variant Statistics (Reference Field in Map: "Reference Alleles A/B")

Number of variant genotypes (non reference)

Number of singletons (variant genotype present only in given sample)

Mean Ti/Tv of variant genotypes

Autosomal Statistics

Hardy-Weinberg Thw P-Value (taken over all autosomal chromosomes and all samples)

Gender Chromosome Statistics

Gender inference:

Select chromosome to use for gender inference: X

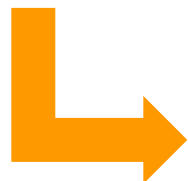
Threshold of heterozygosity for calling M/F 0.02

Additional Outputs (Verbose Output)

Output count and variant statistics for each autosomal chromosome

Help Restore Options Save Options Run Cancel

- Sample QCでは、サンプルごとにSNP Call Rateやヘテロ接合体の割合、ハーディー・ワインベルグ平衡の評価値などをまとめたシートが出力される
- X染色体のヘテロ接合体の割合から、サンプルの性別推定が可能
- シート上よりデータ列を指定し、任意の条件でフィルタリングを行うことが可能 (Call Rate  $\geq 0.95$  など)



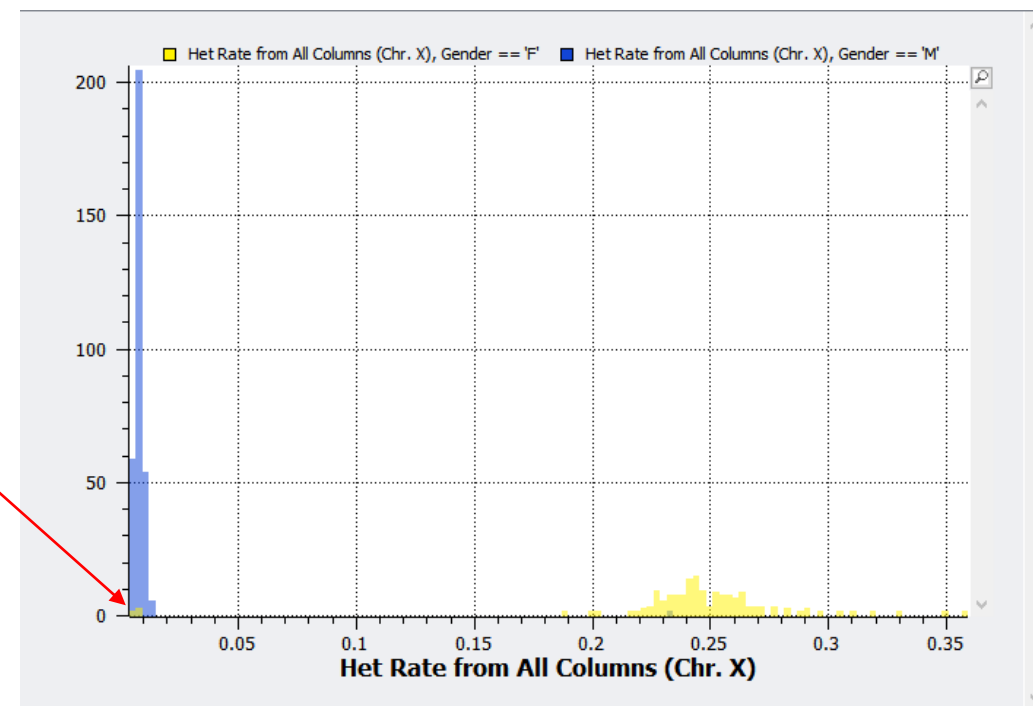
Unsort		R	9	I	10	R	11	I	12
Map	NSP_STY		Call Rate (Autosomes)		# Hets from All Columns (Autosomes)		Het Rate from All Columns (Autosomes)		# Called from Bi-Allelic/Monomorphic (Autosomes)
1	GSM233256 GSM233257		0.973327325974319		132626		0.279056360017338		475266
2	GSM233258 GSM233259		0.964365438571341		121715		0.258478625581346		470890
3	GSM233260 GSM233261		0.939566650965615		113603		0.247619234449552		458781
4	GSM233262 GSM233263		0.980589403837883		127981		0.267288622674453		478812
5	GSM233264 GSM233265		0.973171680763481		125696		0.264517350954355		475190
6	GSM233266 GSM233267		0.97876466853714		127305		0.266372475785747		477921
7	GSM233268 GSM233269		0.931827397653034		109923		0.24158794906396		455002
8	GSM233270 GSM233271		0.958782690614184		119666		0.255607009509488		468164
9	GSM233272 GSM233273		0.924180302688976		109541		0.242740455782373		451268
10	GSM233274 GSM233275		0.918300600053247		106102		0.236625133531223		448397

# Genotype Statistics by Sample

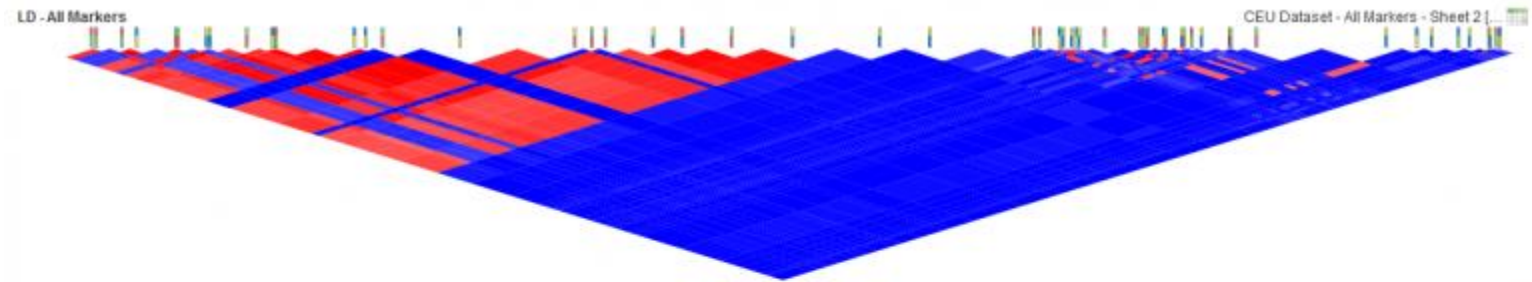
	R	18	C	19	B	20	R	21	B	22	I	23	R	24	C	25
	NSP_STY	Het Rate from All Columns (Chr. X)		Inferred Gender	Phenotype 1		Phenotype 2		Phenotype 3		SBP		BMI		Gender	
43	GSM233362 GSM233363	0.00694100148735746		M	0		8.29		1		134		29.6		M	
44	GSM233366 GSM233367	0.265898751330946		F	1		10.85		1		?		?		F	
45	GSM233368 GSM233369	0.00942655145326002		M	1		13.13		1		125		32.9		M	
46	GSM233370 GSM233371	0.0083015919523391		M	0		7.17		1		134		29.8		M	
47	GSM233372 GSM233373	0.00888932304386051		M	0		11.49		0		138		30.9		F	
48	GSM233374 GSM233375	0.00918686473807662		M	1		7.87		1		105		23.5		M	
49	GSM233378 GSM233379	0.00668436056227268		M	0		11.36		1		133		30.7		M	
50	GSM233382 GSM233383	0.00490966221523959		M	0		10.03		1		130		30.3		M	
51	GSM233384 GSM233385	0.008756038647343		M	0		6.69		0		140		31.3		M	
52	GSM233386 GSM233387	0.223991297468354		F	1		8.07		0		129		31.4		F	

- 性別推定を実行すると、Sample QCシートに推定されたサンプルごとの性別が書き込まれ、表現型データシートに性別データがあれば、不一致の確認を行うことが可能
- X染色体のヘテロ接合体の割合をヒストグラムにプロットし、表現型データシートの性別データで色付けを行うことで、性別不一致サンプルを視覚的に確認することが可能

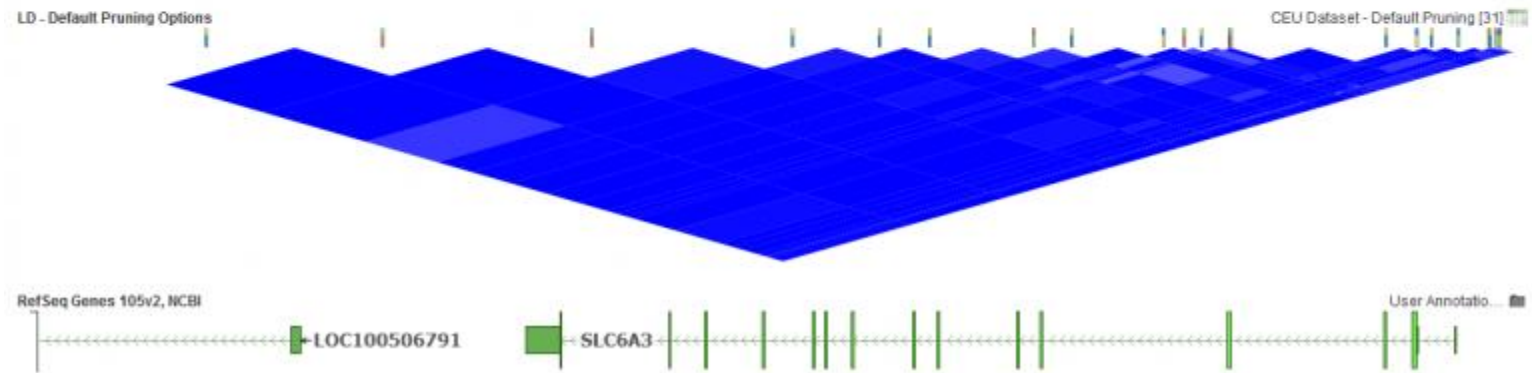
表現型データの性別と推定された性別が一致しないサンプル



LD Pruning前のSNP



LD Pruning後のSNP

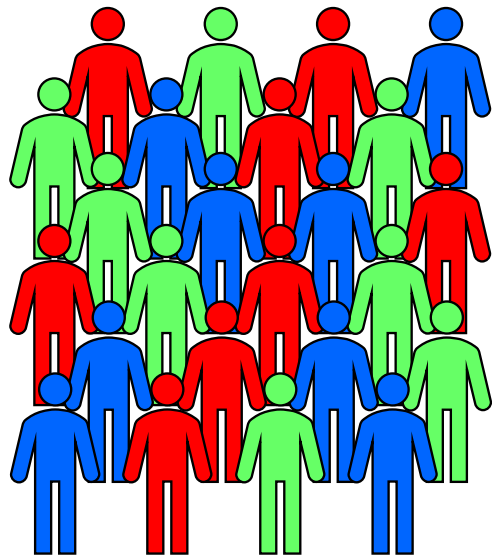


フィルターされたSNP

- 連鎖不平衡（Linkage Disequilibrium）の関係にあるSNPの中から、1つだけを残し残りのSNPをすべてフィルターすることによって、代表的なSNPのみに絞り込む手法
- PCAやIBD計算に用いるSNPデータの前処理として実行

All: 464 x 498,784  
Active: 464 x 212,282

G	214819	G	214820	G	214821	G	214822
	SNP A-2289687		SNP A-1874178		SNP A-2168040		SNP A-2180172
	C_T		A_A		A_A		G_G
	C_T		G_G		C_C		A_G
	T_T		G_G		C_C		A_G
	C_T		G_G		C_C		A_A
	C_C		A_G		A_C		A_G
	?_?		G_G		C_C		A_G



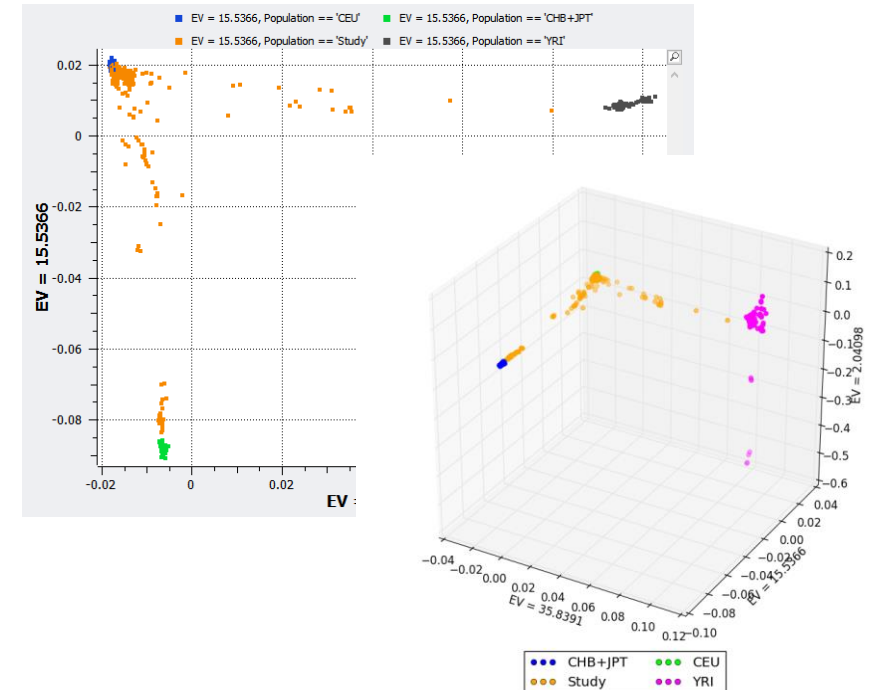
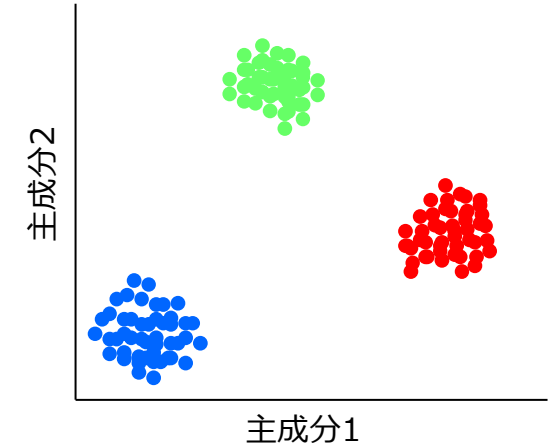
サンプル名	主成分1	主成分2	主成分3
サンプル1	-0.0091	0.015	-0.023
サンプル2	-0.013	0.016	0.086
サンプル3	-0.017	0.014	-0.031
サンプル4	-0.0073	-0.079	-0.054
サンプル5	0.0078	0.0059	0.069



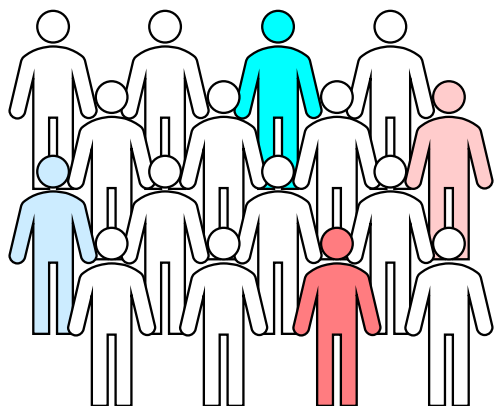
GWASのデータ補正に利用

- サンプル間の関係の主成分 (Principal Component) を計算することによって、サンプルの遺伝的背景を調べると同時に、必要に応じてGWAS計算時のデータ補正に利用
- SVSでは、主成分のデータシート出力に加え、2次元/3次元のグラフ表示が可能

グラフ表示

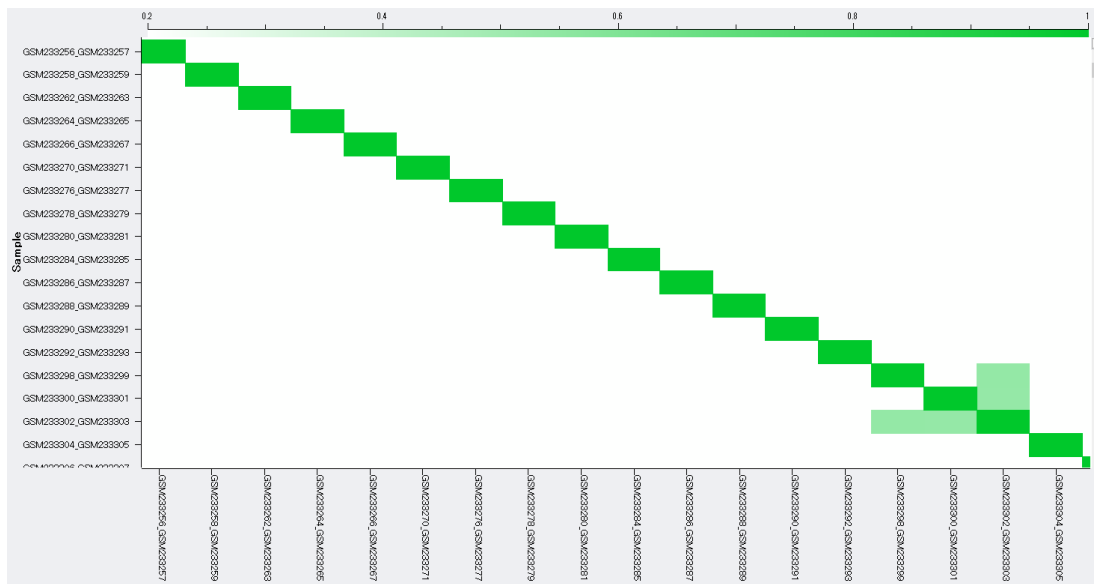


# Identity by Descent (IBD)



	サンプル1	サンプル2	サンプル3	サンプル4
サンプル1	1	0.01	0	0.06
サンプル2	0.01	1	0.04	0
サンプル3	0	0.04	1	0.03
サンプル4	0.06	0	0.03	1

グラフ表示



- SNPデータをもとに、すべてのサンプル同士の関係を数値化データで、サンプル間の血縁関係の推定に用いる
- GWASでは、血縁関係のない独立したサンプルを使うことが前提のため、IBDの数値の大きい場合は、サンプルペアのどちらか一方を解析から除外する
- 計算したデータは、ワンクリックでヒートマップ表示に切り替え、関係の強度をグラフィカルに表示することが可能



Genotype Filtering by Marker

(No variable is set as dependent.)

Classify alleles by allele frequency  Classify alleles by reference/alternate  
(Reference field in map: "Reference Alleles A/B")

Filter Genotype Columns

General Statistics Filtering

Drop if call rate < 0.9

Drop if number of alleles > 2

Drop if Minor Allele Frequency (MAF) < 0.01

Drop if carrier count < 10

Hardy Weinberg Equilibrium (HWE) Filtering

Perform HWE filtering based on: All

Drop if Hardy Weinberg Equilibrium (HWE) P-Value < 0.001

Drop if Fisher's exact test for HWE P-Value < 0.0001

Drop if signed HWE R (positive if more homozygous) > 0.2

Actions

Inactivate genotype columns that meet above criteria for filtering

Output spreadsheet with marker statistics and 'Drop?' columns

Additional Output

Output -log10(Value)

Help Restore Options Save Options Run Cancel

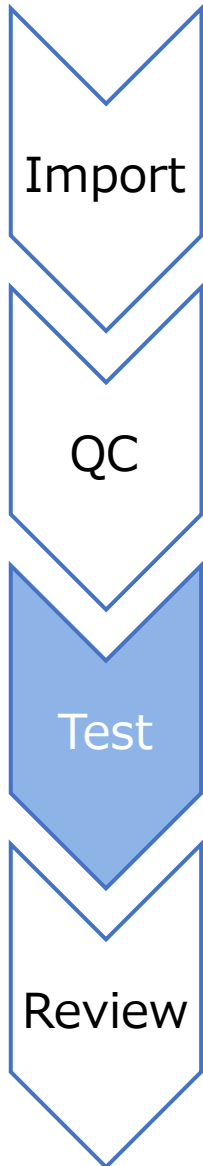
フィルターされたSNPデータシート

Unsort		G 59	G 60	G 61	G 62	G 63	G 64	G 65
Map	sub	SNP A-4248380	SNP A-4249114	SNP A-4273679	SNP A-2135694	SNP A-2233800	SNP A-4220764	SNP A-2232054
1	GSM233256 GSM233257	C_T	T_T	C_T	G_G	C_C	G_G	C_C
2	GSM233258 GSM233259	T_T	C_C	?_?	?_?	C_C	A_A	C_C
3	GSM233262 GSM233263	C_C	T_T	T_T	G_G	C_C	?_?	C_G
4	GSM233264 GSM233265	T_T	T_T	C_C	?_?	C_C	A_A	C_C
5	GSM233266 GSM233267	T_T	T_T	C_C	?_?	C_C	A_A	C_C
6	GSM233270 GSM233271	C_C	T_T	T_T	G_G	C_C	G_G	C_C
7	GSM233276 GSM233277	C_T	T_T	C_T	G_G	C_C	?_?	C_C
8	GSM233278 GSM233279	C_T	C_T	T_T	G_G	C_C	A_G	C_C
9	GSM233280 GSM233281	C_T	T_T	C_T	A_G	C_C	A_A	C_C
10	GSM233284 GSM233285	T_T	T_T	C_C	G_G	C_C	A_G	C_C

SNPごとのQCシート

Unsort		B 1	R 2	B 3	R 4	B 5	R 6
Map	Marker	Drop?	Call Rate	Call Rate < 0.9	MAF	MAF < 0.01	Fisher's HWE P (Controls)
1	SNP A-1909444	0	1	0	0.172413793103448	0	0.637770213047421
2	SNP A-4303947	0	0.978448275862069	0	0.111233480176211	0	0.746665776740478
3	SNP A-1886933	0	0.978448275862069	0	0.177312775330396	0	0.187512269449798
4	SNP A-2116190	0	0.997844827586207	0	0.251619870410367	0	0.494572258078673
5	SNP A-4291020	0	0.969827586206897	0	0.0888888888888889	0	1
6	SNP A-1902458	0	0.995689655172414	0	0.114718614718615	0	0.769526106706426
7	SNP A-2109914	0	1	0	0.0980603448275862	0	0.70145721300437
8	SNP A-2291997	0	0.974137931034483	0	0.0597345132743363	0	0.596759476599105
9	SNP A-4277872	0	0.991379310344828	0	0.1	0	0.0161970528095594
10	SNP A-4221087	0	0.980603448275862	0	0.010989010989011	0	0.9999999999999807

- Marker QCでは、SNPごとのCall Rateやマイナーアレル頻度、ハーディー・ワインベルグ平衡の評価値などの閾値を設定し、SNPデータシートからSNPのフィルタリングを直接実行する
- Sample QCと同様に、SNPごとのQCデータをまとめたシートも別途出力される



## ■ 表現型データ

- 疾患の有無や量的形質など

## ■ 共変量

- 性別や集団名、主成分データなど

## ■ SNPデータ

- QCをパスしたデータのみを使用

- Create Imputation Reference Panel
- Genotype Imputation with BEAGLE
- PBAT Family-Based QA
- PBAT Genotype Analysis
- Genotype Association Tests
- **Genotypic Regression Analysis**
- Haplotype Association Tests
- Haplotype Block Detection
- Haplotype Trend Regression
- Runs of Homozygosity for GWAS

## ■ Genotype Regression Analysis

- SNPデータから回帰分析を実行するためのツール
- カイ二乗検定などの従来の手法と違い、量的な表現型データ（量的形質）や共変量によるバイアス補正が可能

# 解析用シートの作成



各データシートを一つに統合

Unsort		R	1	R	2	C	3	R	4	R	5	R	6	G	7	G	8	G	9	G	10	G	11
Map	sub		Phenotype 2		BMI		Gender		EV = 35.8391		EV = 15.5366		EV = 2.04098		SNP A-1909444		SNP A-4303947		SNP A-1886933		SNP A-2116190		SNP A-4291020
1	GSM233256 GSM233257		8.96		29.9		M		-0.00913848082717825		0.0152483756818808		0.00079288192320567		T_T		A_A		G_G		A_A		G_G
2	GSM233258 GSM233259		12.18		32.8		M		-0.0132832140660563		0.0160274568117261		-0.00267927413469164		C_T		A_A		G_G		A_A		G_G
3	GSM233262 GSM233263		11.01		27.4		M		-0.0171541561898054		0.0144654467729679		-0.00151115502578382		C_T		A_G		A_G		A_G		G_T
4	GSM233264 GSM233265		9.25		29.7		M		-0.0175477947358796		0.0179457088844222		-0.000371874804655187		T_T		A_A		G_G		A_A		G_G
5	GSM233266 GSM233267		6.85		27.7		M		-0.0176869143656866		0.0192942175196264		-9.8021270587903e-06		T_T		A_A		G_G		G_G		G_G
6	GSM233270 GSM233271		11.65		29.2		M		-0.0174722583523377		0.0185435996780324		-0.000304887286060429		T_T		A_A		G_G		A_A		G_T
7	GSM233276 GSM233277		10.82		20		M		-0.0171744721818751		0.0193516580266033		0.000390985401467564		T_T		A_A		G_G		A_A		G_T
8	GSM233278 GSM233279		9.62		31.2		M		-0.0167299709588903		0.0188172288894126		0.000398072606222266		T_T		A_A		G_G		A_A		G_G
9	GSM233280 GSM233281		9.02		26.2		M		-0.00731550512550411		-0.0790740316310559		0.00125077230443915		T_T		A_A		A_G		A_A		G_T
10	GSM233284 GSM233285		11.73		22.8		M		-0.0176815540204872		0.0187010220735615		0.000158925142877242		C_T		A_G		A_G		A_G		G_G
11	GSM233286 GSM233287		12.81		26.1		F		-0.017094187251538		0.0187133319025732		-0.000186116671362646		T_T		A_A		G_G		A_A		G_G
12	GSM233288 GSM233289		9.46		23.5		M		-0.016657779126341		0.0195007943419922		0.0015462503981008		C_T		A_G		A_G		A_G		G_G

その他の表現型データ  
(共変量に使用)

SNPデータ

GWASに使用する表現型データ  
(従属変数)

主成分データ  
(共変量に使用)

Genotypic Regression Analysis

Quantitative dependent variable: Phenotype 2 (464 active samples) Linear Regression

Regression Parameters Output Parameters Genotypic Parameters

Selection Parameters

Regress on each of the 384255 genotypic columns

Correct for covariate(s)

Add additional full model covariate(s)

Regress on a moving window with parameters:

Fixed window over 384255 genotypic columns

# of columns in window: 1

Dynamic window over 384255 mapped genotypic columns

Size: 10000 base pairs  with max markers: 20

Correct for covariate(s)

Add additional full model covariate(s)

Perform single regression with selected covariates

Correct for covariate(s)

Regress on covariate-column interactions (on 384255 genotypic cols)

Correct for additional covariate(s)

Regression Options

Stepwise Regression P-value cutoff: 0.01

Backward elimination  Forward selection

Output residual spreadsheet

Full Model Covariates

(The current genotypic column)

Add Covariate

Add Interaction

Remove Selected

Clear List

Reduced Model Covariates

Gender

Gender:F

EV = 35.8391

EV = 15.5366

EV = 2.04098

Add Covariate

Add Interaction

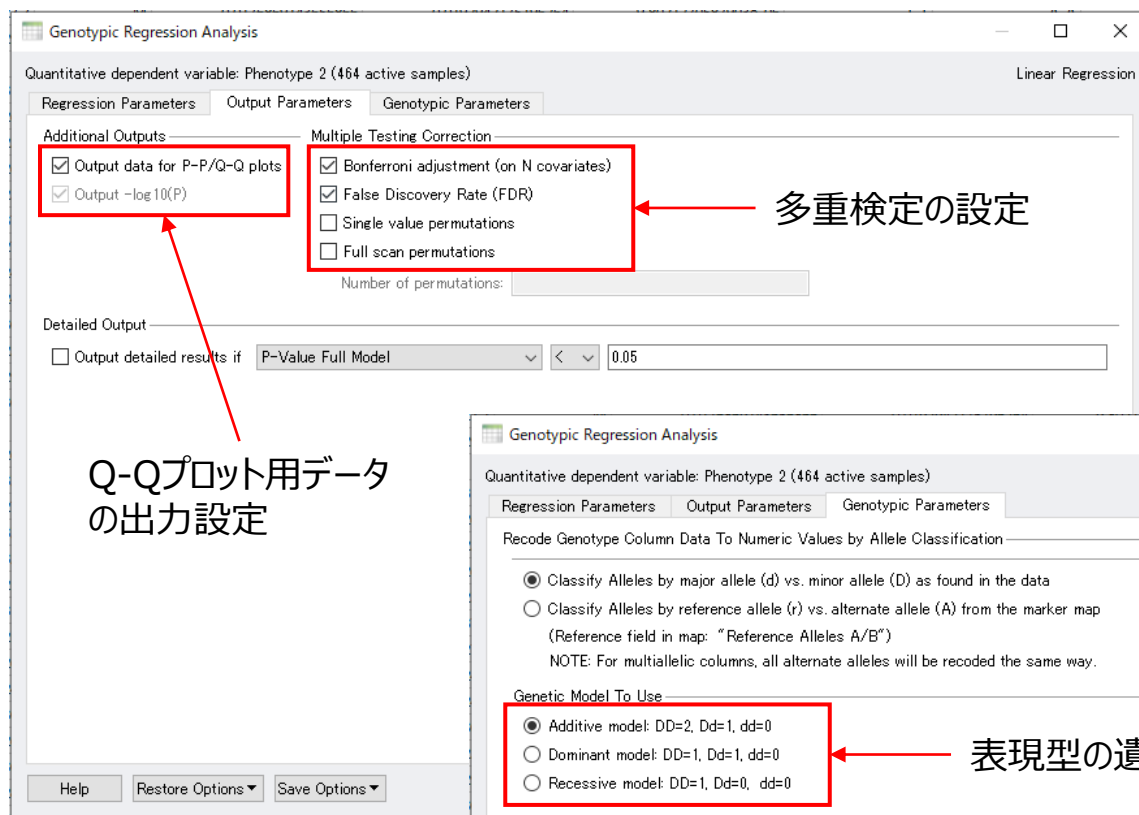
Remove Selected

Clear List

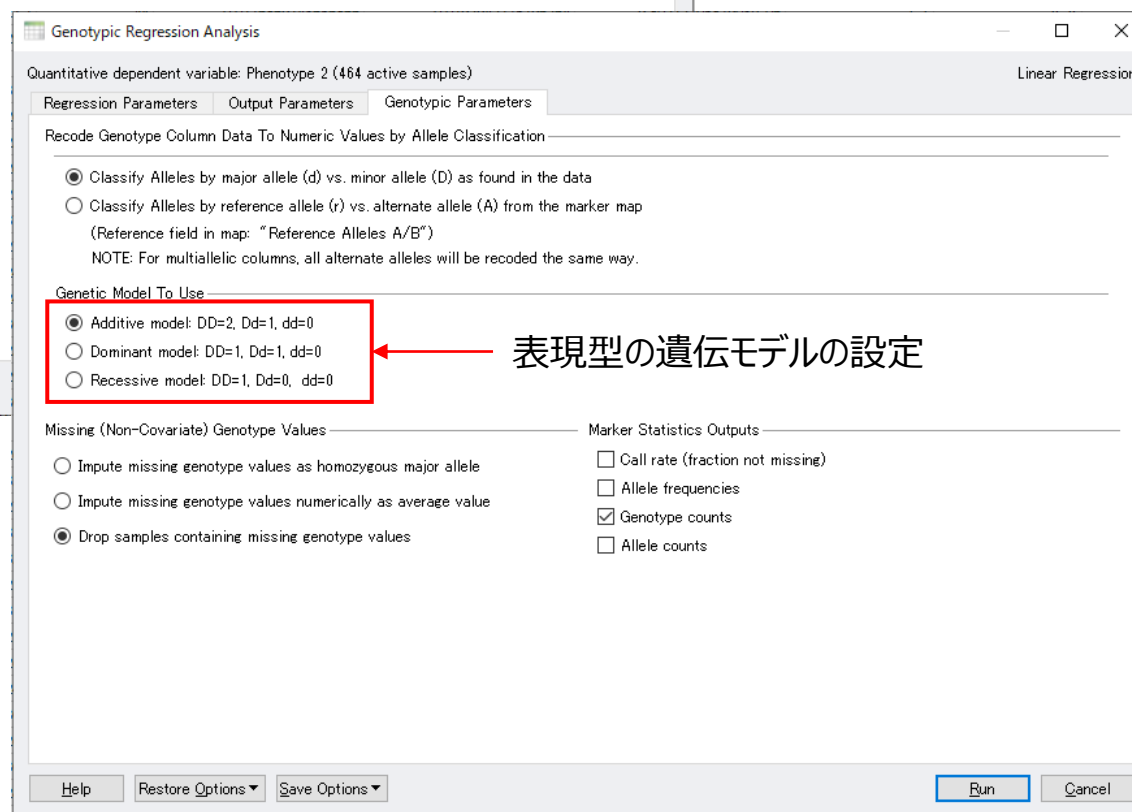
Help Restore Options Save Options Run Cancel

補正に用いる共変量  
(シート内より選択)

- Genotype Regression Analysisツールを起動し、シート内の補正に用いる共変量の列名を選択
- シートで選択した表現型がバイナリデータ（ケース／コントロールなど）の場合はロジスティック回帰、定量データの場合は線形回帰の手法が自動で実行される



- 必要に応じて、出力データの選択や、解析実行時のオプション設定を調整

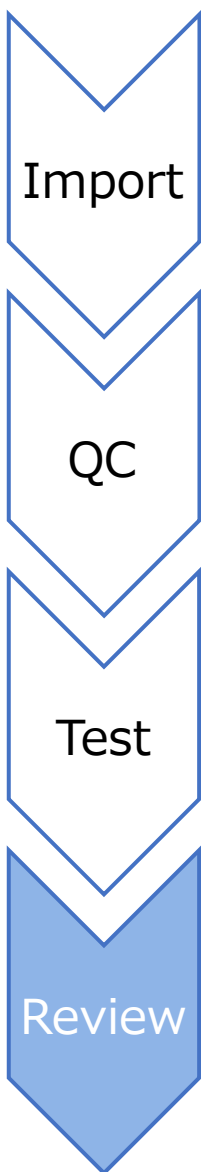


マーカーマップ表示ボタン

Unsort	R	1	R	2	R	3	R	4	R	5	I	6	R	7	R	8
Map	Predictor	FvR Model P-Value	-log10 FvR Model P	expected FvR Model P	expected -log10 FvR Model P	Predictor Beta	Sample Size	Bonferroni P	FDR							
1	SNP A-1909444	0.846928278871993	0.0721533658309954	0.847974652249157	0.0716171295273073	-0.0386061416380235	464	1	0.998764499639262							
2	SNP A-4303947	0.696968128321045	0.156787081314872	0.697808746795748	0.156263591119894	-0.0950116568866571	454	1	0.998793482988185							
3	SNP A-1886933	0.321287856958079	0.493105688254137	0.321584624793431	0.492704723378167	-0.204406897693895	454	1	0.999073127800428							
4	SNP A-2116190	0.357700657404535	0.446480261676863	0.358292019622386	0.445762864936203	-0.154918001699734	463	1	0.998345870819747							
5	SNP A-4291020	0.50297133440925	0.29845676576442	0.503830789449714	0.297715295989301	0.177532959639651	450	1	0.99829158111274							
6	SNP A-1902458	0.527754778527308	0.277567825706434	0.528543545301948	0.276919226511155	-0.146032706881106	462	1	0.998505201594373							
7	SNP A-2109914	0.690381124220186	0.160911091219672	0.691211565236627	0.160389003925533	-0.103084042161614	464	1	0.99879669161839							
8	SNP A-2291997	0.688603463568493	0.162030797192046	0.689449714382376	0.161497403980733	-0.128838432677009	452	1	0.998770685641262							
9	SNP A-4277872	0.767088480402878	0.11515453913515	0.767697232306671	0.114810025103855	-0.0818344150063285	460	1	0.999205348067771							
10	SNP A-4221087	0.563698637088136	0.248953015216742	0.564493630531808	0.248340954076471	0.414263245752205	455	1	0.998589367914349							
11	SNP A-1866065	0.936095131040668	0.0286800136693436	0.935911048652587	0.0287654257438163	-0.0139473824233057	430	1	1							
12	SNP A-2288244	0.580502678209219	0.23619577225822	0.581326202651885	0.23558010124685	0.0881060979396464	464	1	0.998581134289337							
13	SNP A-1884606	0.159505324081015	0.797224816177858	0.159699938842696	0.796695250175136	-0.295968602189654	460	1	0.998773234441719							
14	SNP A-1783407	0.276797305655153	0.557838141616994	0.277160999856866	0.557267880620717	0.178604063617939	455	1	0.998683098604904							
15	SNP A-2082515	0.955279470875209	0.0198695552191927	0.955267986102978	0.0198747765217964	0.0132014512781551	464	1	1							

- 計算が完了すると、SNPごとのP値などをまとめたシートが出力される
- SNPデータシートにマーカーマップデータを加えてあれば、「Map」ボタンをクリックすることで、各SNPの詳細データを表示可能





## ■ データの可視化

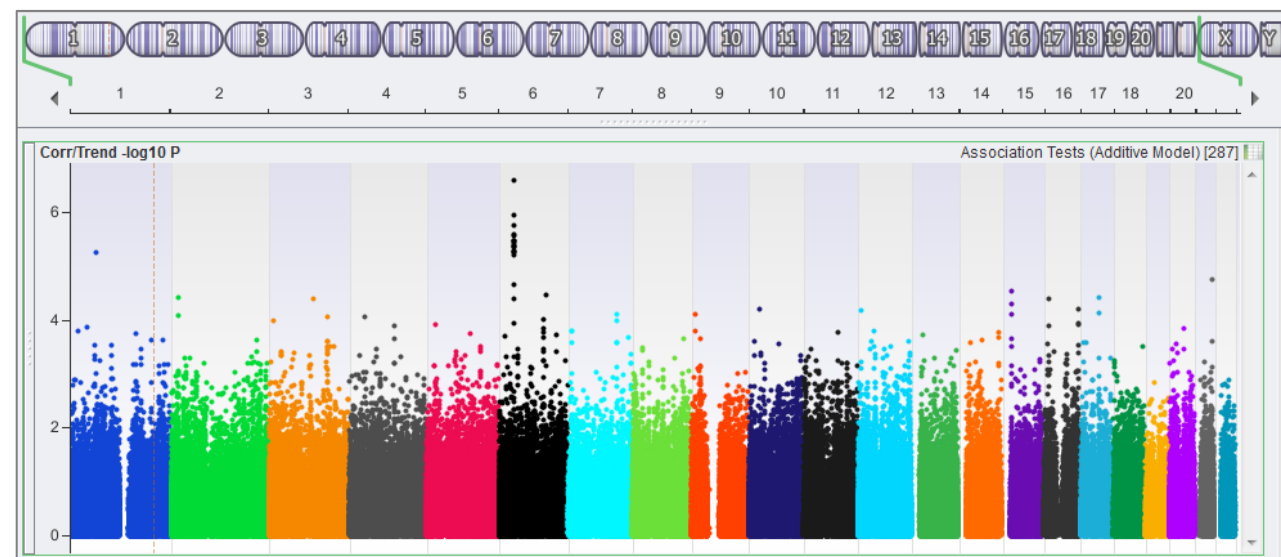
- Manhattan Plot
- LD Plot
- Q-Q Plot ...

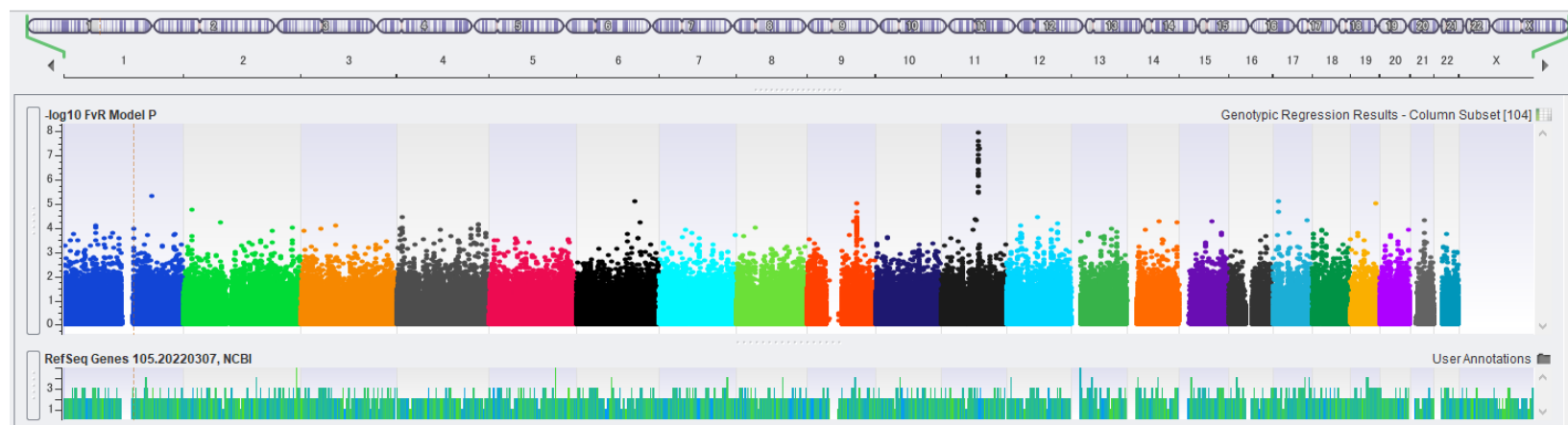
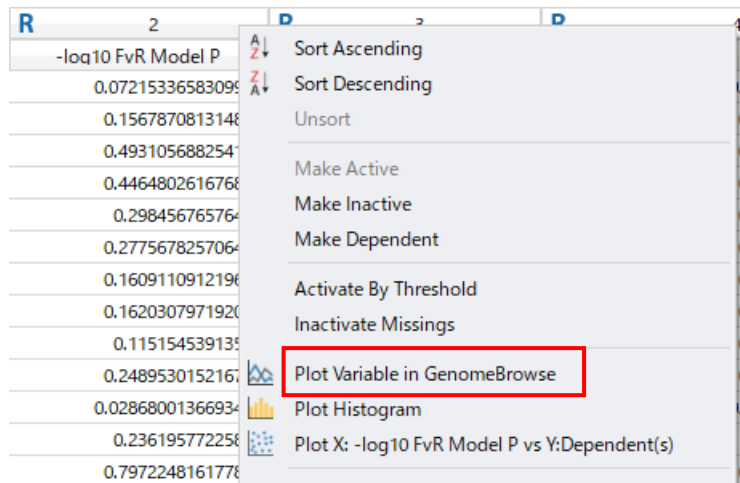
## ■ 集団の階層化の評価、遺伝率の計算など

- Genomic Control
- LD Score Regression

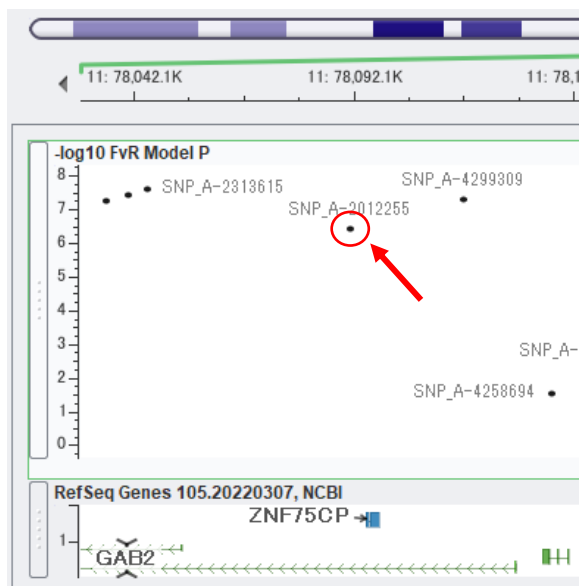
## ■ メタ解析

- Meta-Analysis





- $-\log_{10}$ 変換されたP値の列を右クリックし、メニューからゲノムブラウザーを起動させることで、マンハッタンプロットを表示



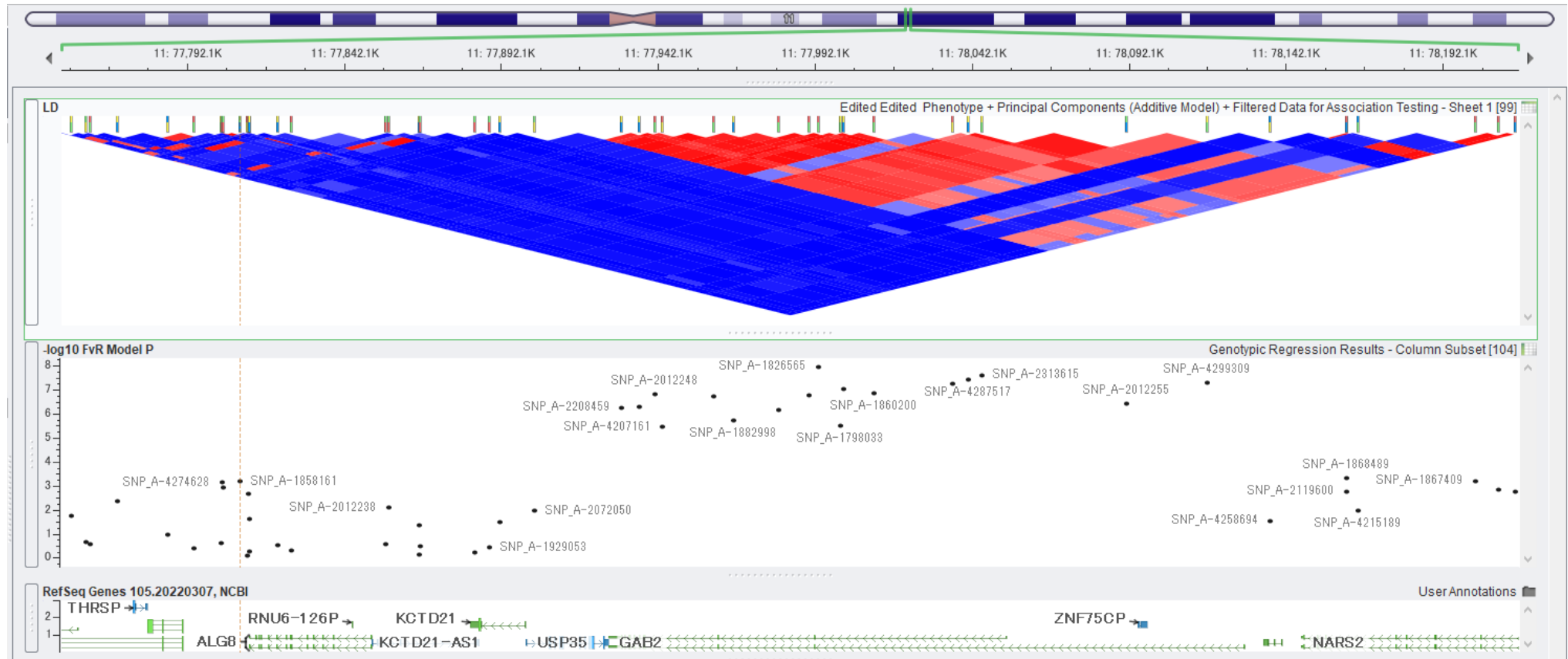
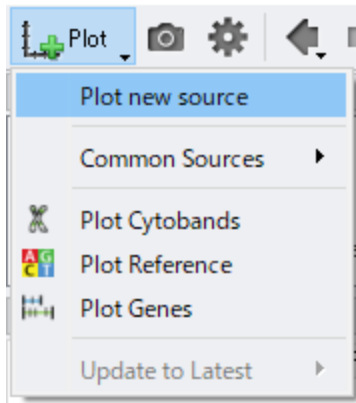
Chr11: 78,091,150 - 78,091,150

$-\log_{10}$  FvR Model P 6.39313312151234

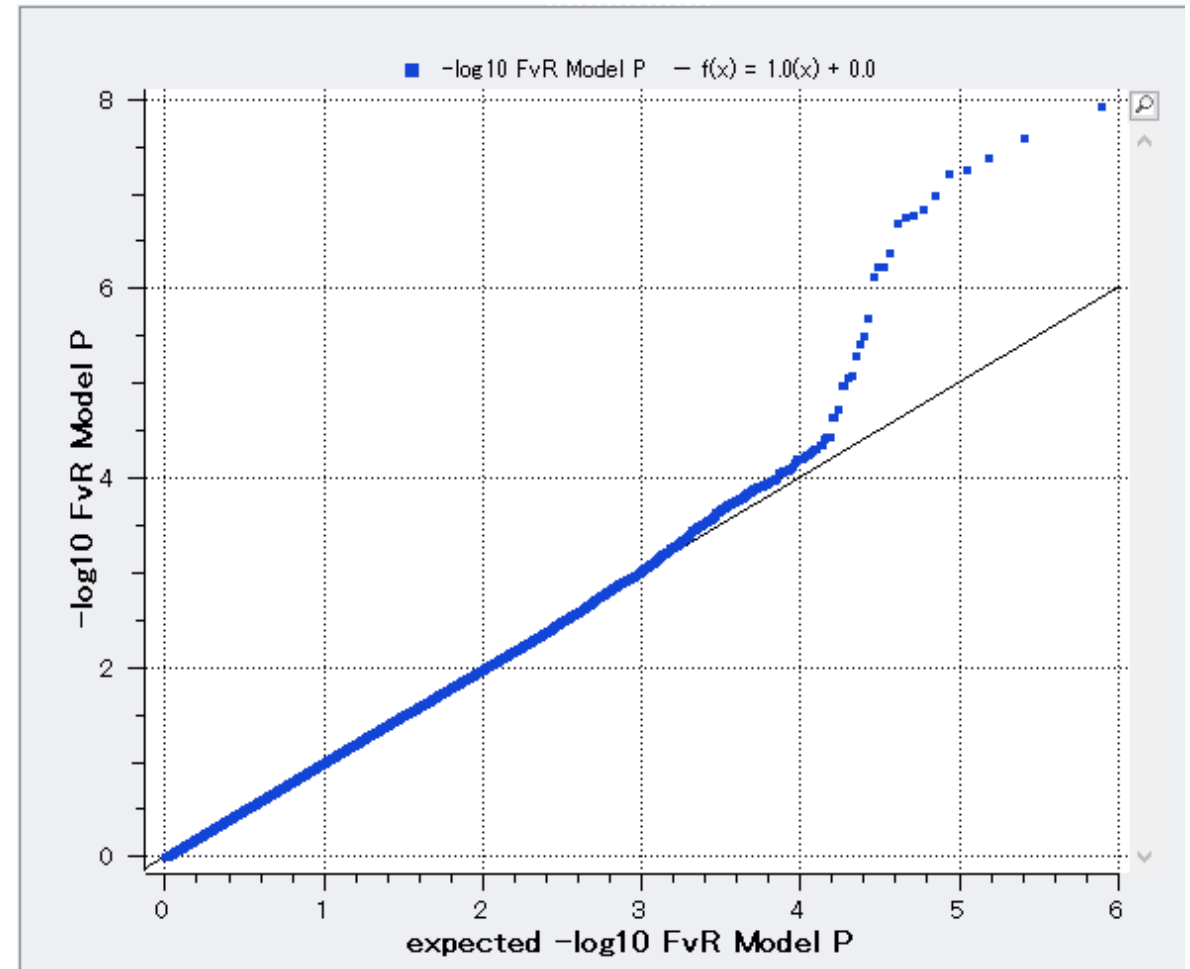
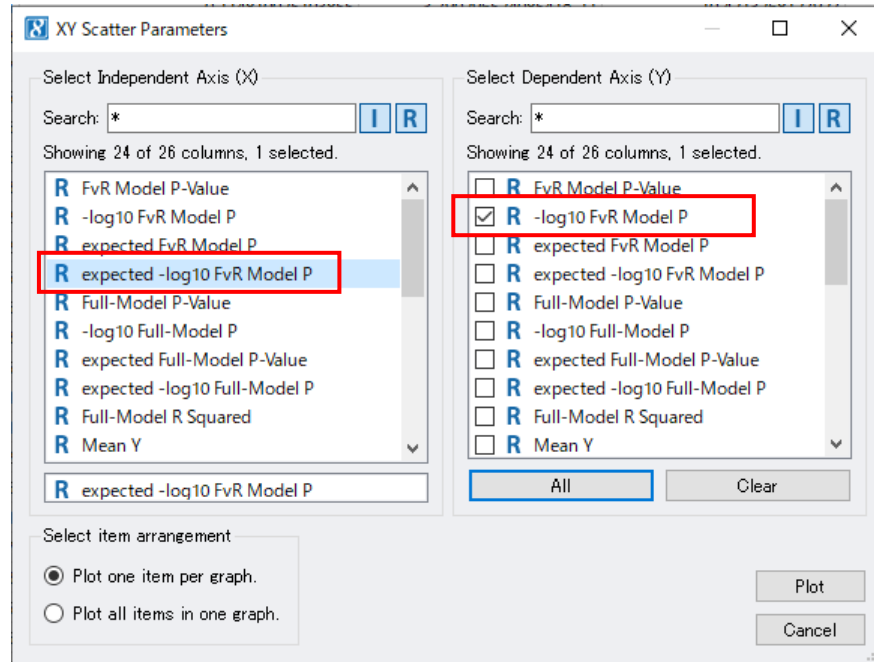
Extended Fields

Name	SNP_A-2012255
dbSNP RS ID	rs2373115
Associated Gene	?
Cytoband	q14.1
Reference Alleles A/B	[G/T]
Top Alleles	[C/A]
Bottom Alleles	[G/T]
Strand	-
Strand Versus dbSNP	same
Probe Count	12
FvR Model P-Value	4.04451898639024e-07
$-\log_{10}$ FvR Model P	6.39313312151234
expected FvR Model P	2.73256040910333e-05
expected $-\log_{10}$ FvR Model P	4.56343022826198

- ゲノム上の任意の領域を拡大表示させ、SNPのプロットをマウスクリックすると、SNPの詳細データ（マーカーマップやGWAS結果など）が表示される



- ゲノムブラウザには任意のアノテーションリソース（リファレンスゲノム配列、RefSeq, dbSNPなど）を選択して表示が可能
- サンプル集団のSNPデータシートを選択した場合は、LDプロットが表示される



- 解析結果シートより、X軸に $-\log_{10}$ 変換されたP値の予測値、Y軸に $-\log_{10}$ 変換されたP値の実測値を選択して、Q-Qプロットを作図
- プロットが全体的に、傾き45°の直線から上側にずれている場合は、サンプル集団の階層化などのバイアスの存在が疑われる

お問い合わせ先：フィルジエン株式会社

TEL: 052-624-4388 (9:00～18 : 00)

FAX: 052-624-4389

E-mail: [biosupport@filgen.jp](mailto:biosupport@filgen.jp)