VCF-Server 用户手册

（第一版）

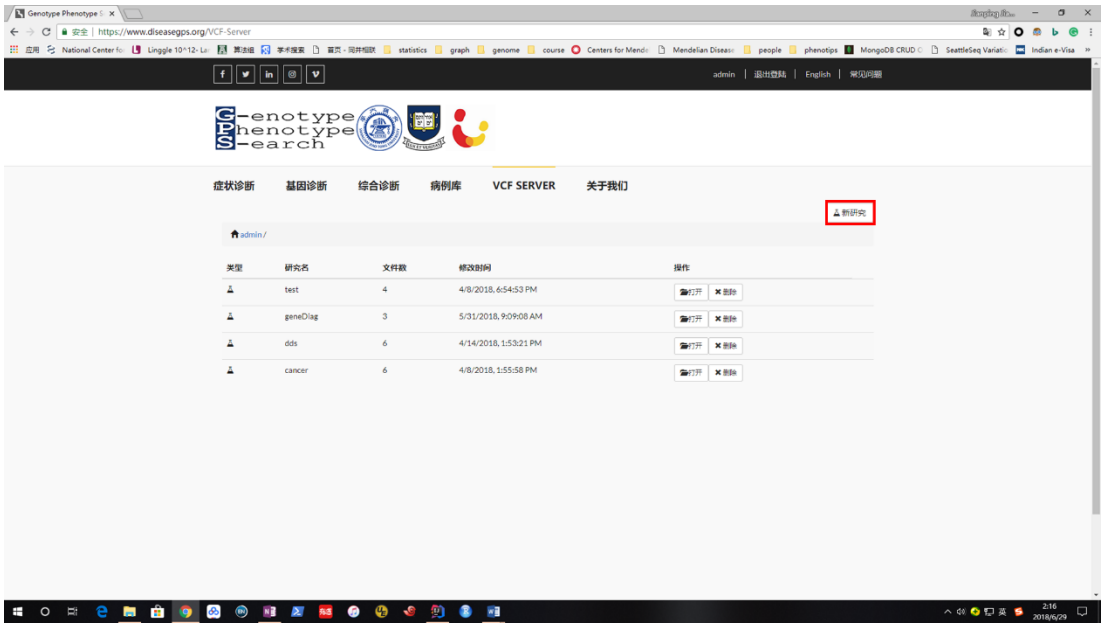
2018.6

目录

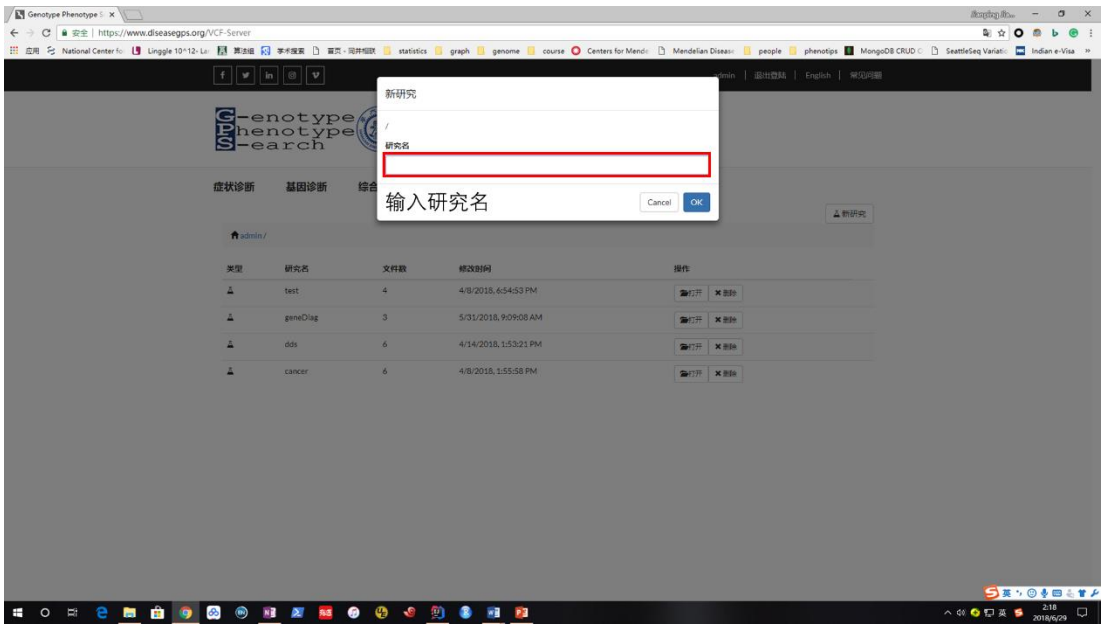
研究项目管理

研究项目管理是为了将不同研究项目中产生的 VCF 文件分别放置到不同的研究项目中，以便于统一有效的对 VCF 文件进行管理，这样可以大大提高 VCF 文件查找和处理的效率。

新建研究项目（目前只对登录用户开放）

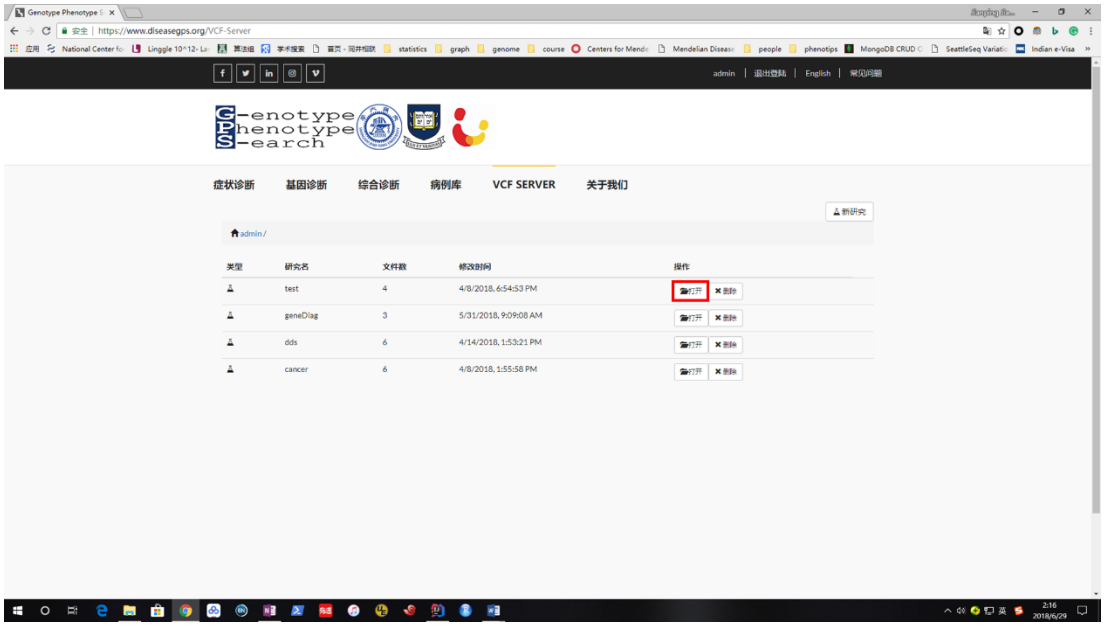


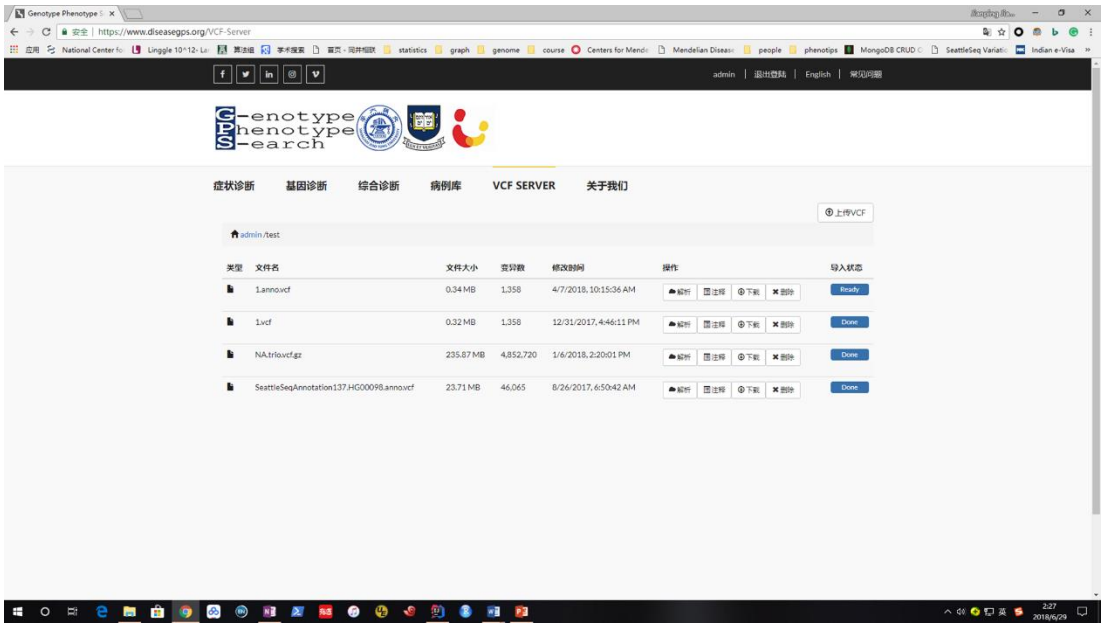
用户登陆后，点击“新研究”图标，可以新建一个研究项目。



输入研究项目名并点击 OK 后，在页面中就可以看到新建的研究项目了。

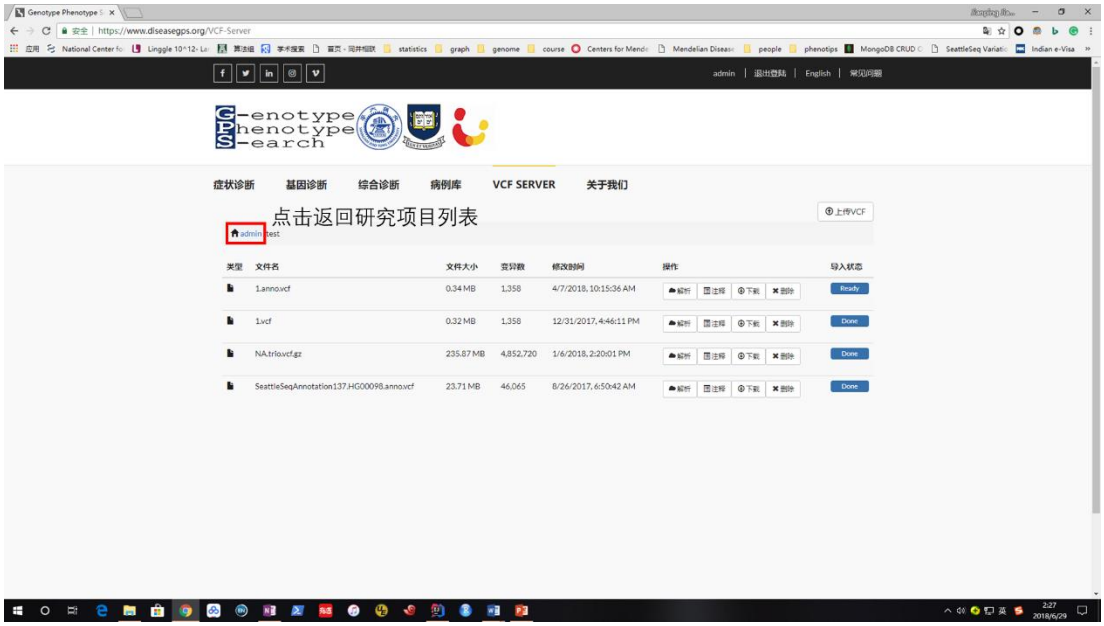
打开研究项目



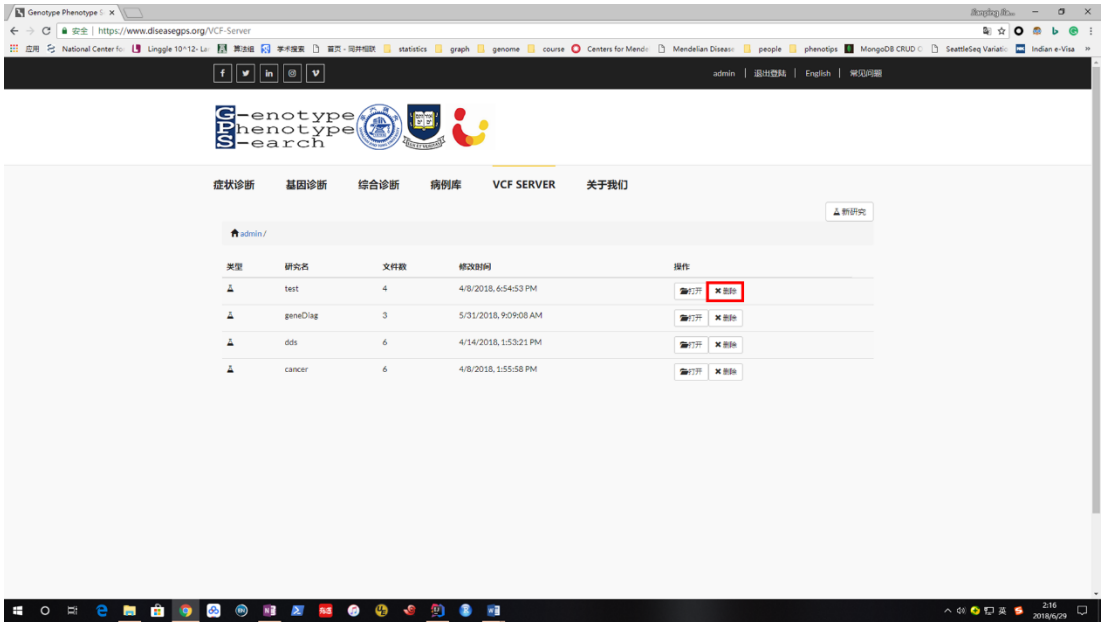


打开研究项目后，显示该项目中所有的 VCF 文件。

返回研究项目列表



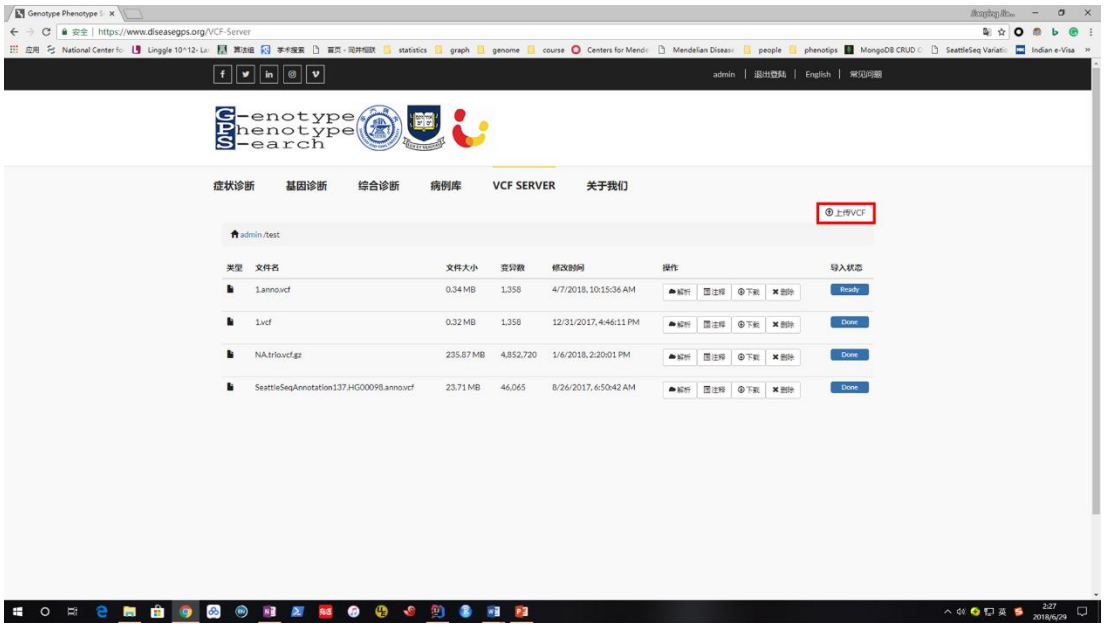
删除研究项目



只有当研究项目中没有文件时，才可以删除该研究项目。

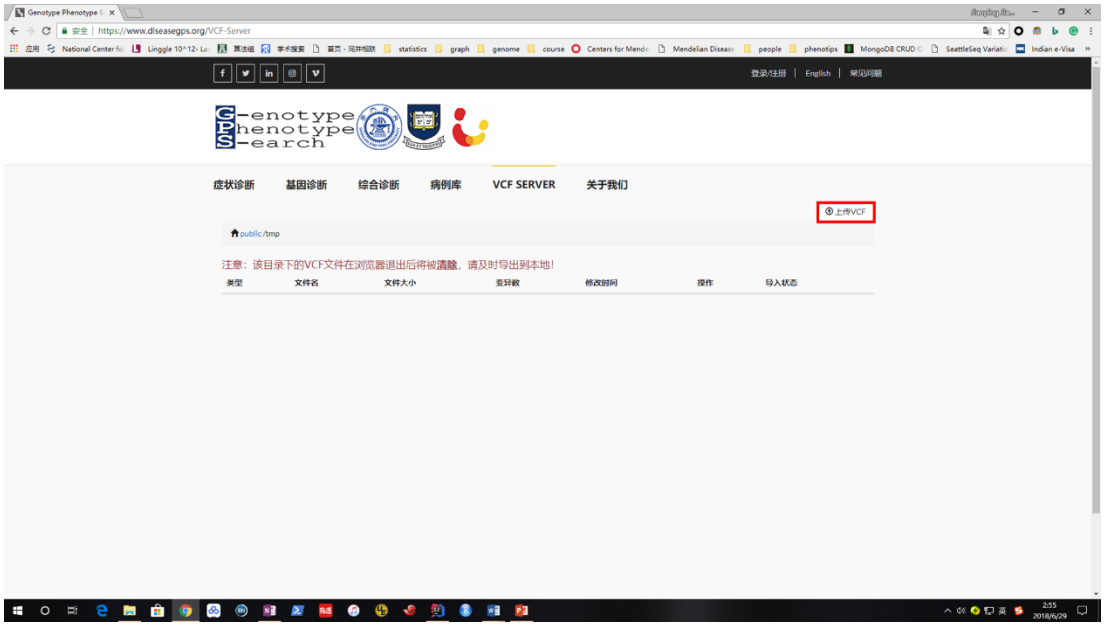
VCF 文件上传

- 登陆用户



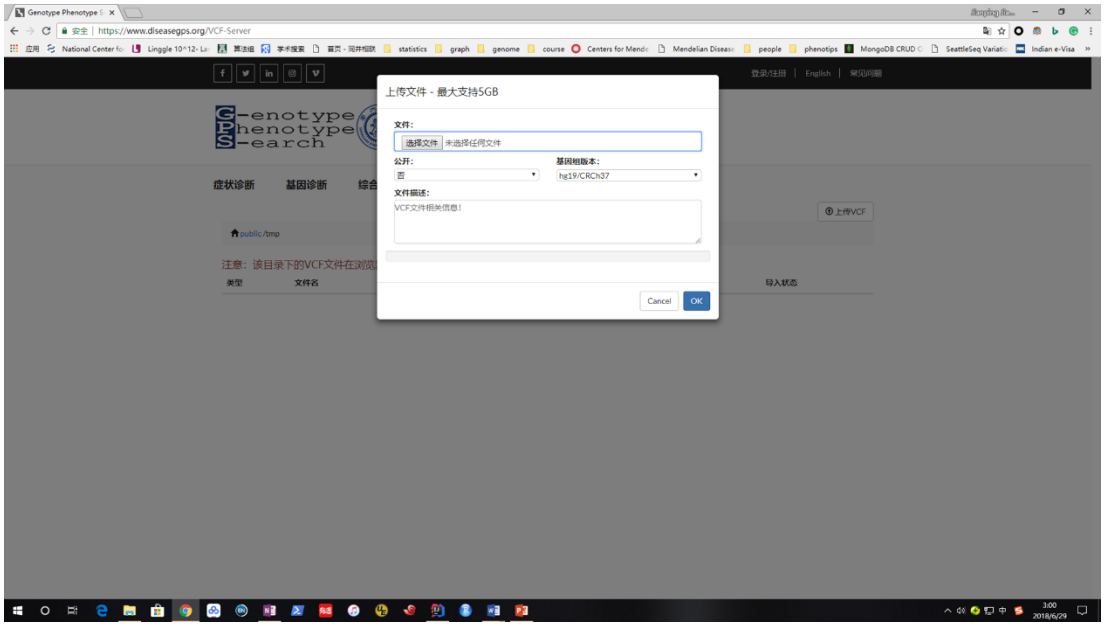
对于登陆用户，可以将 VCF 文件上传至新建的研究项目中。

- 未登陆用户



对于未登陆用户，只能将 VCF 文件上传至 tmp 研究项目中，并且浏览器关闭后上传的 VCF 文件及处理后的 VCF 文件会被清除。

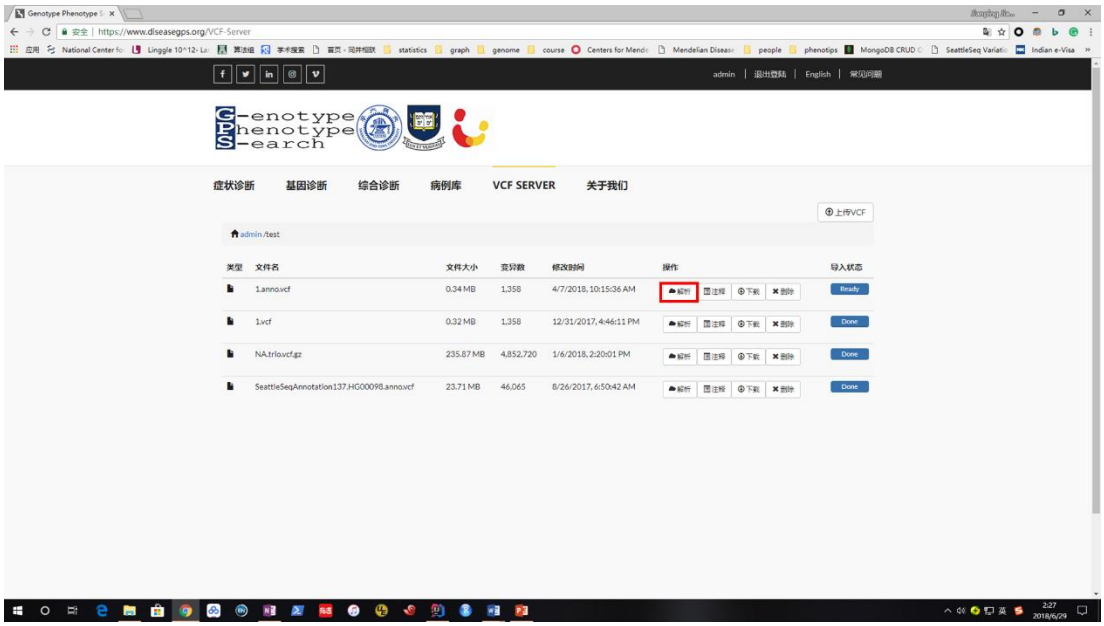
- 上传 VCF 文件信息

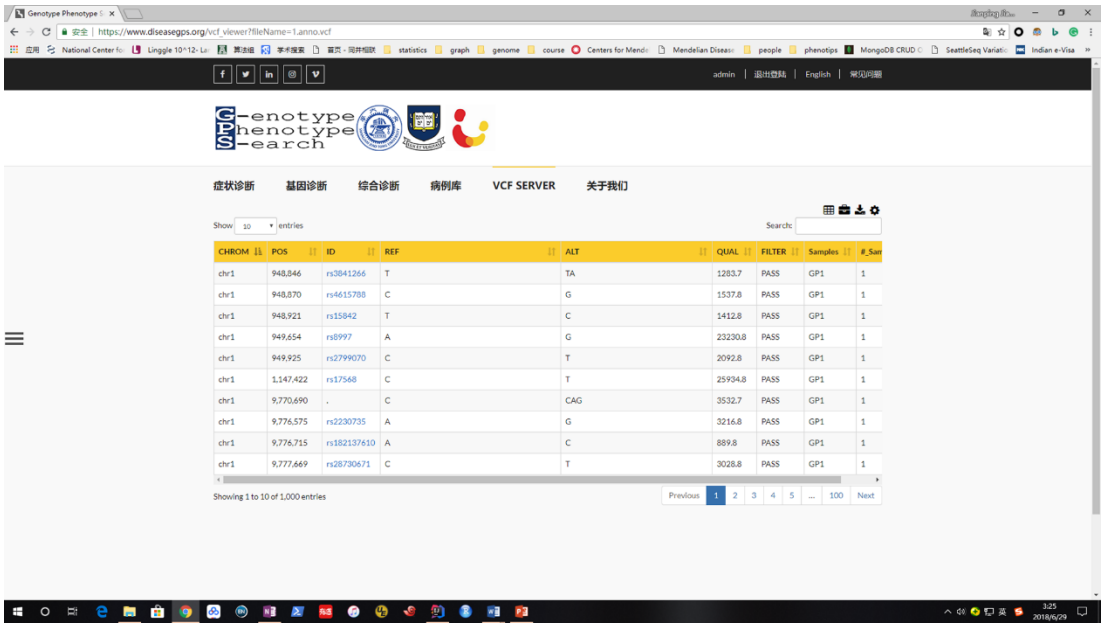


上传 VCF 文件时需要注明 VCF 使用的基因组版本号。

VCF 文件处理

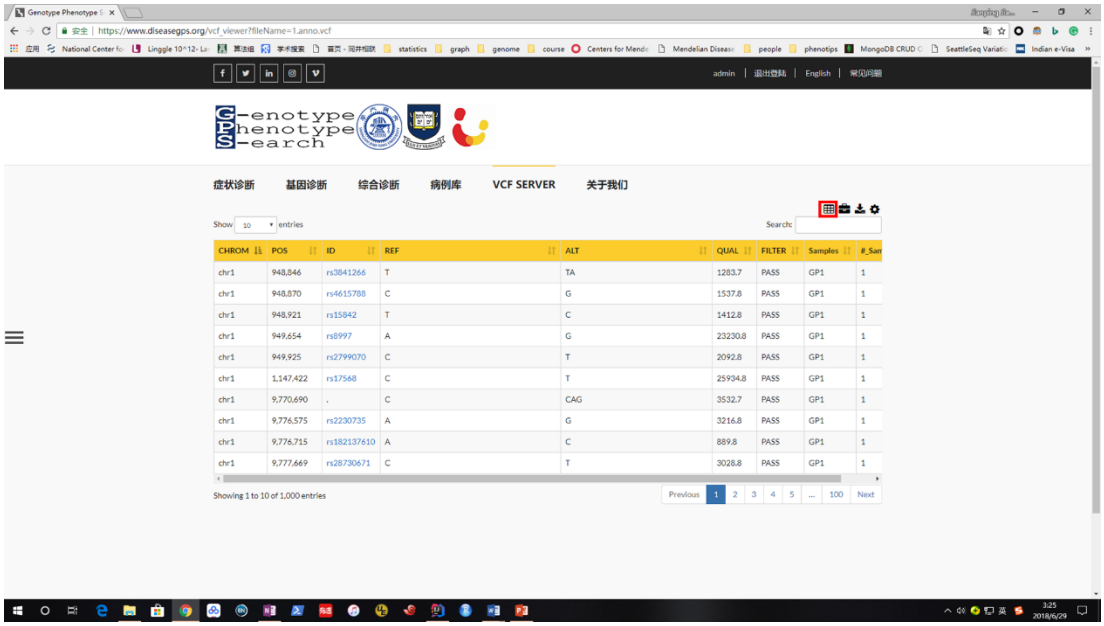
VCF 文件解析

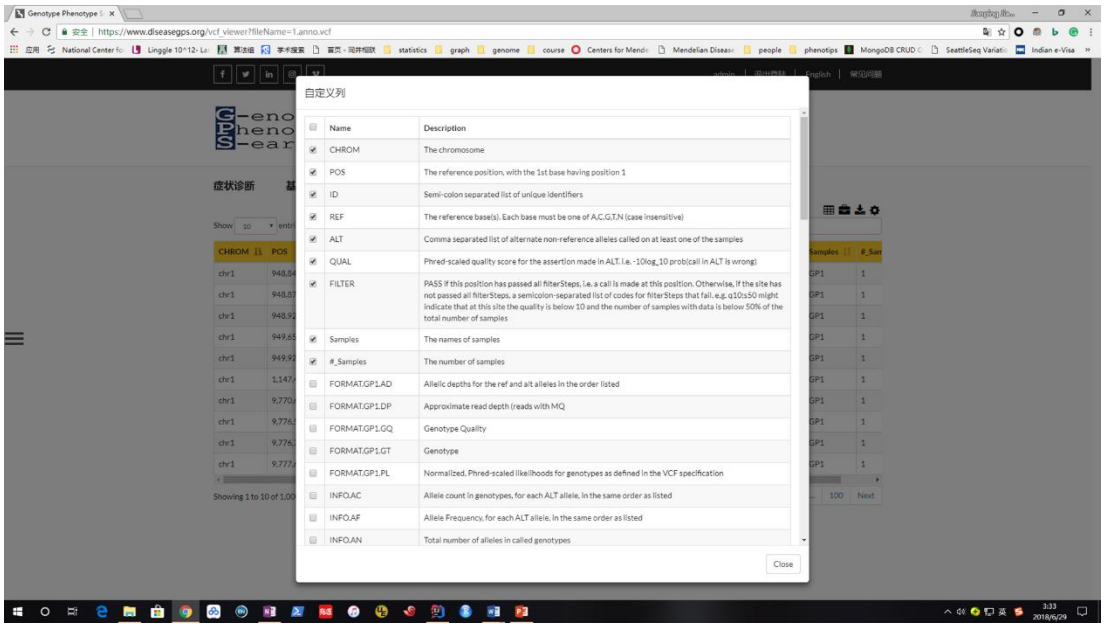




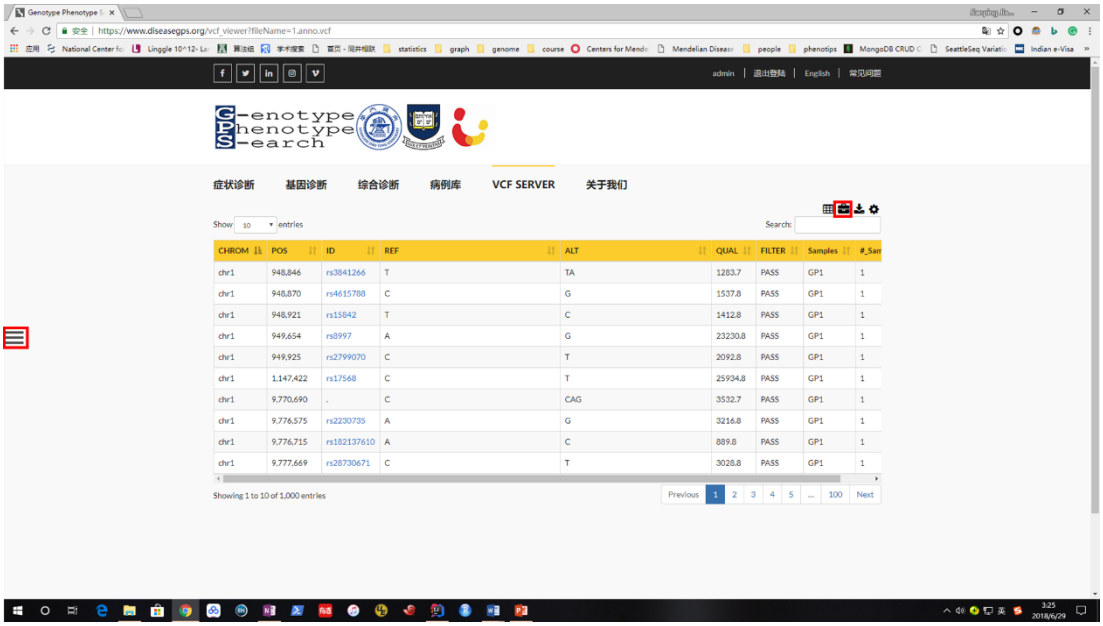
将 VCF 文件进行在线的解析，并显示 VCF 中的基本信息，如染色体，位置，dbSNP 的 ID 等等。

自定义显示列

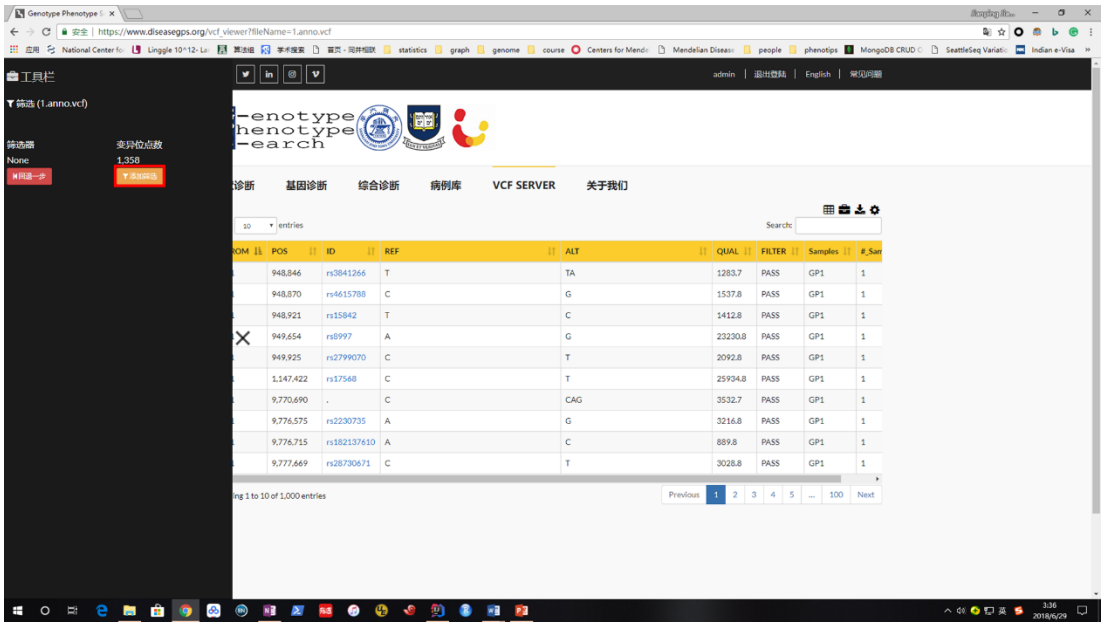




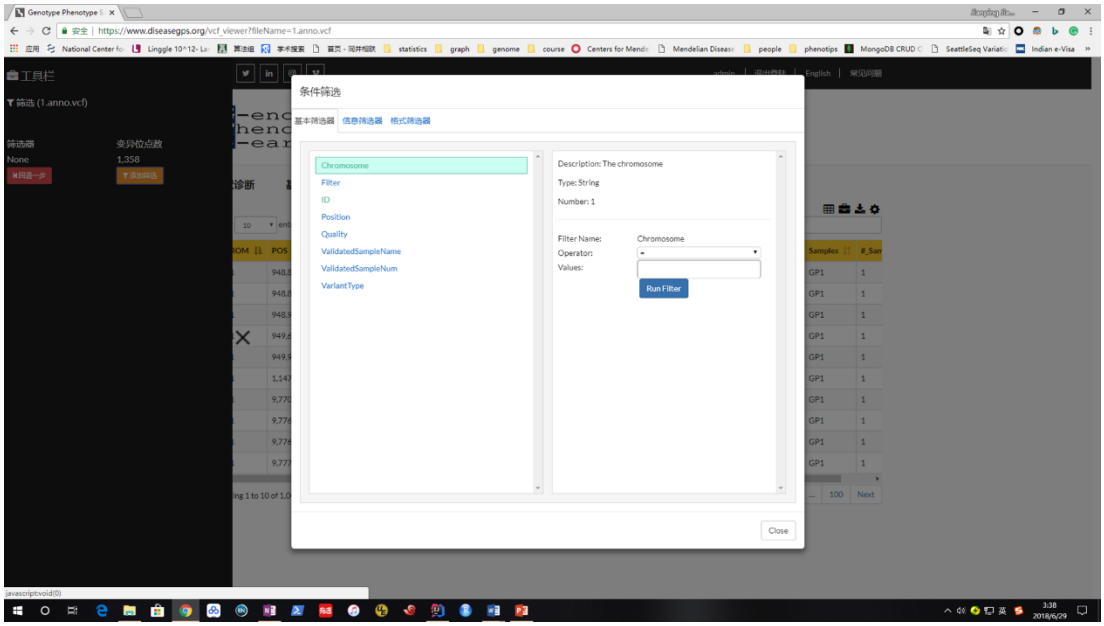
自定义筛选



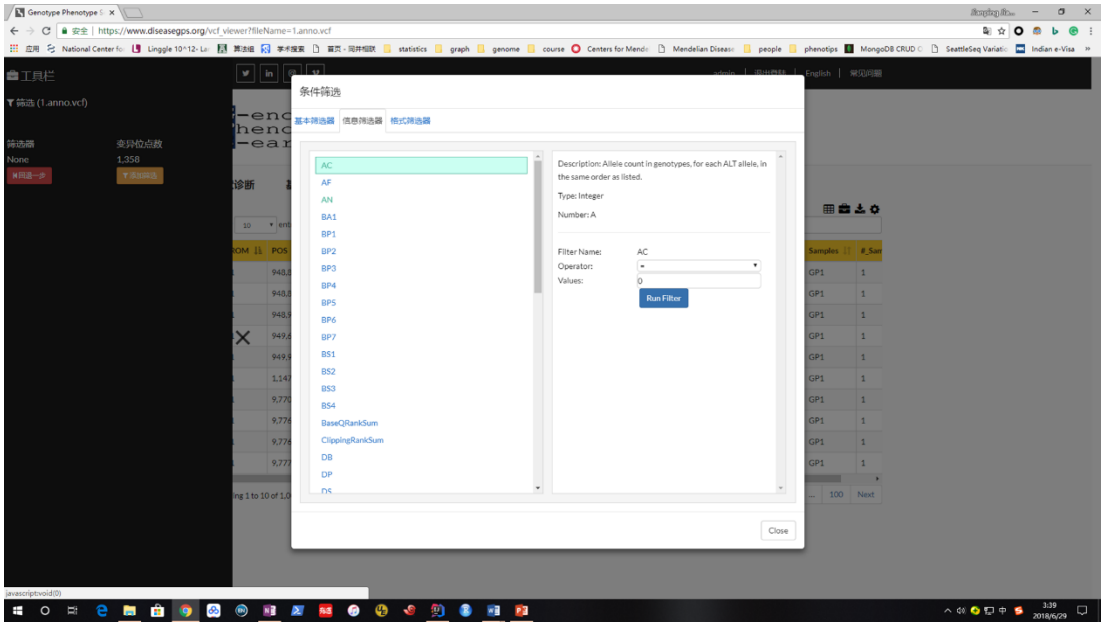
添加筛选条件



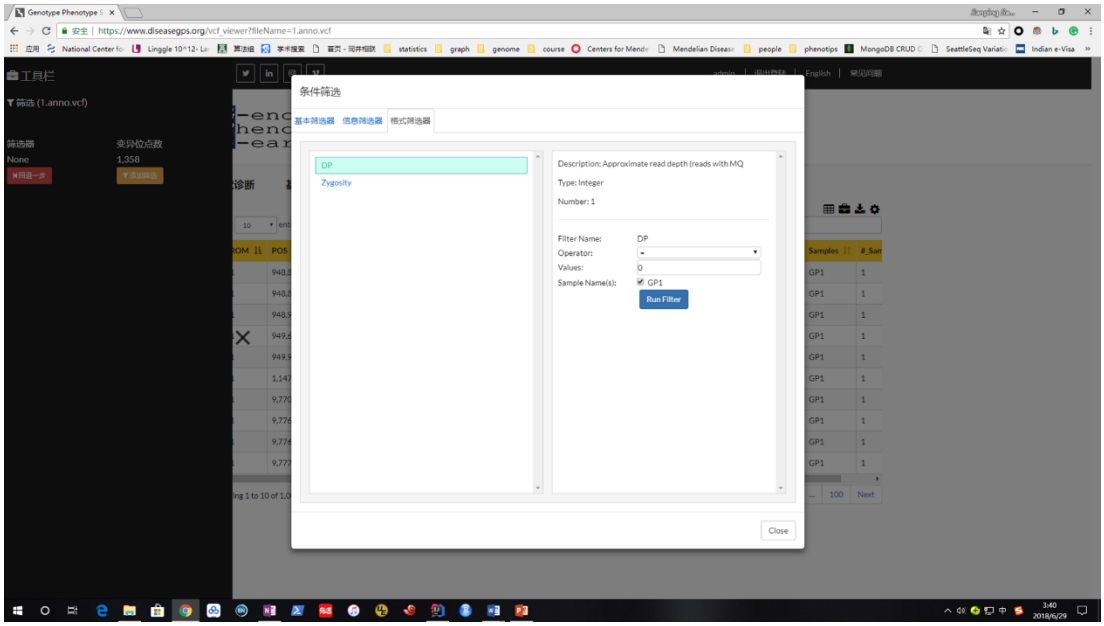
- **基本筛选器**


条件筛选
基本筛选器
信息筛选器
格式筛选器
Chromosome
Filter
ID
Position
Quality
ValidatedSampleName
ValidatedSampleNum
VariantType
Description: The chromosome
Type: String
Number: 1
Filter Name: Chromosome
Operator:
Values:
Run Filter
Close

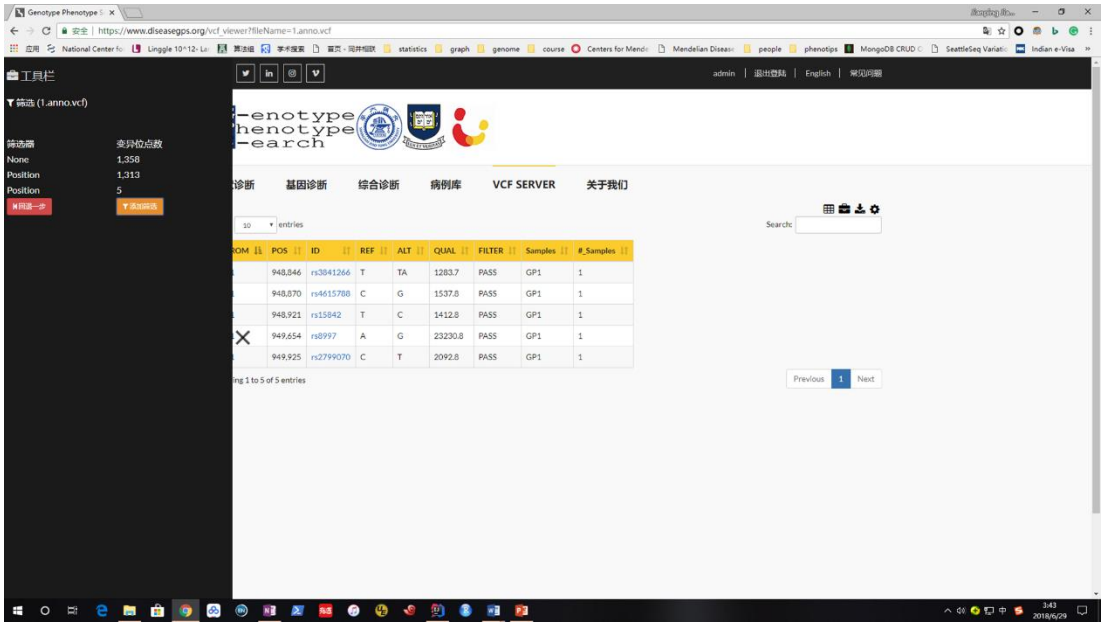
- **信息筛选器**


条件筛选
基本筛选器
信息筛选器
格式筛选器
AC
AF
AN
BA1
BP1
BP2
BP3
BP4
BP5
BP6
BP7
BS1
BS2
BS3
BS4
BaseQRankSum
ClippingRankSum
DB
DP
Description: Allele count in genotypes, for each ALT allele, in the same order as listed.
Type: Integer
Number: A
Filter Name: AC
Operator:
Values:
Run Filter
Close

- **格式筛选器**

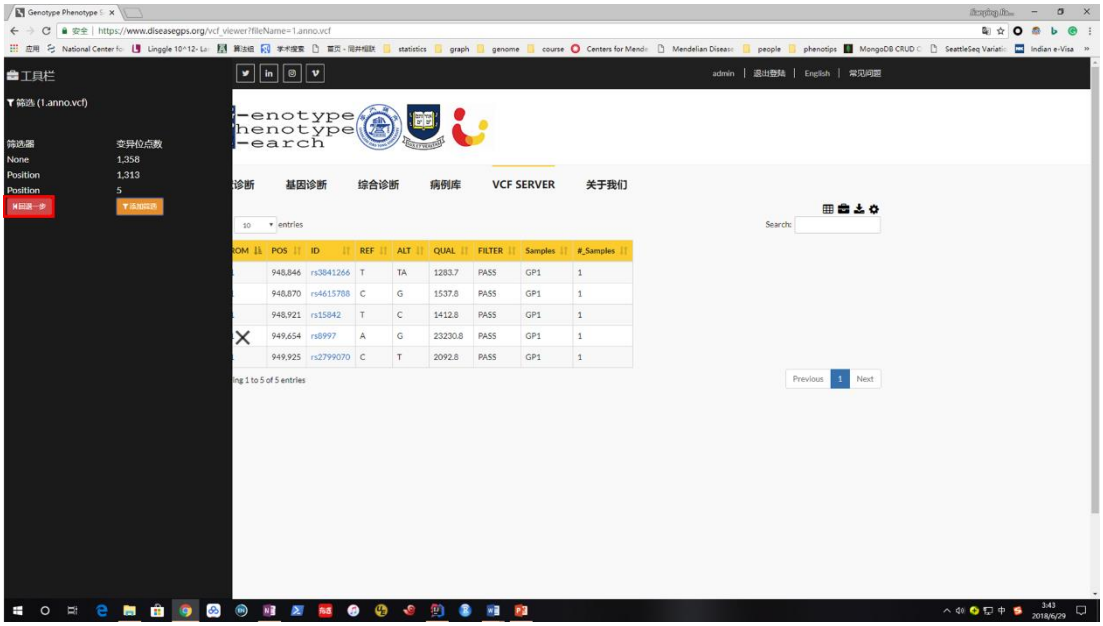


筛选示例

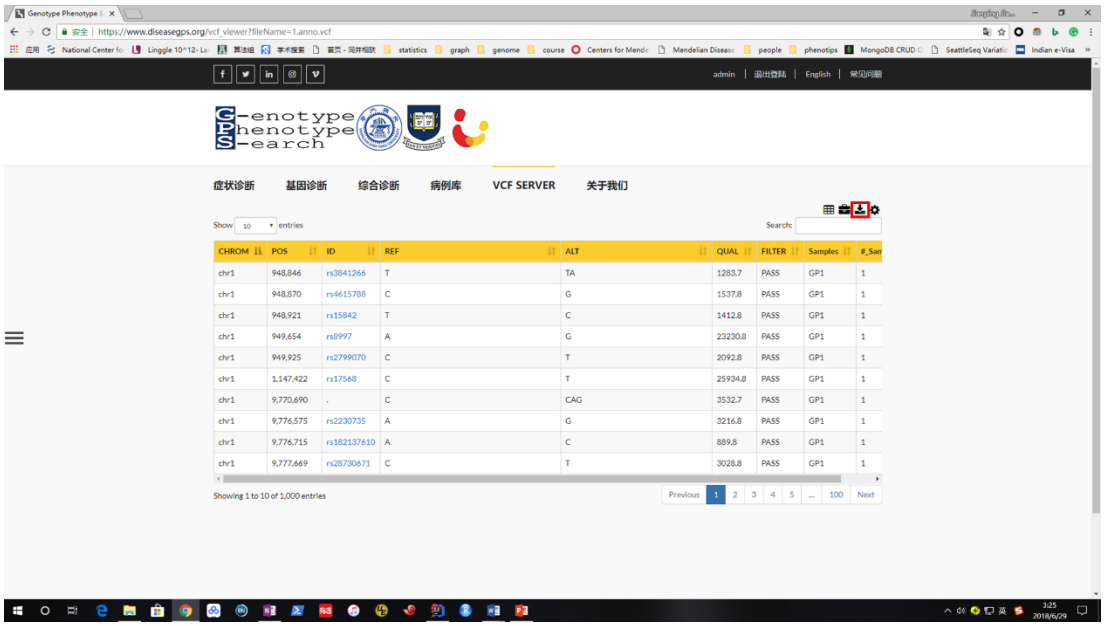


筛选基因组位置 900K-1M 区域内的位点。

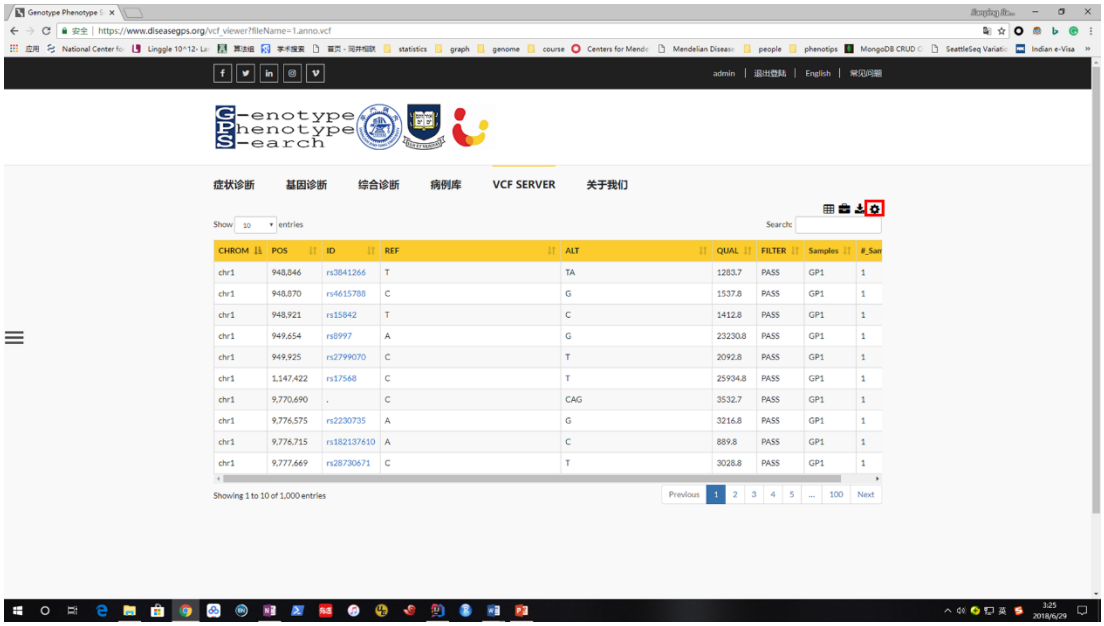
撤销筛选条件

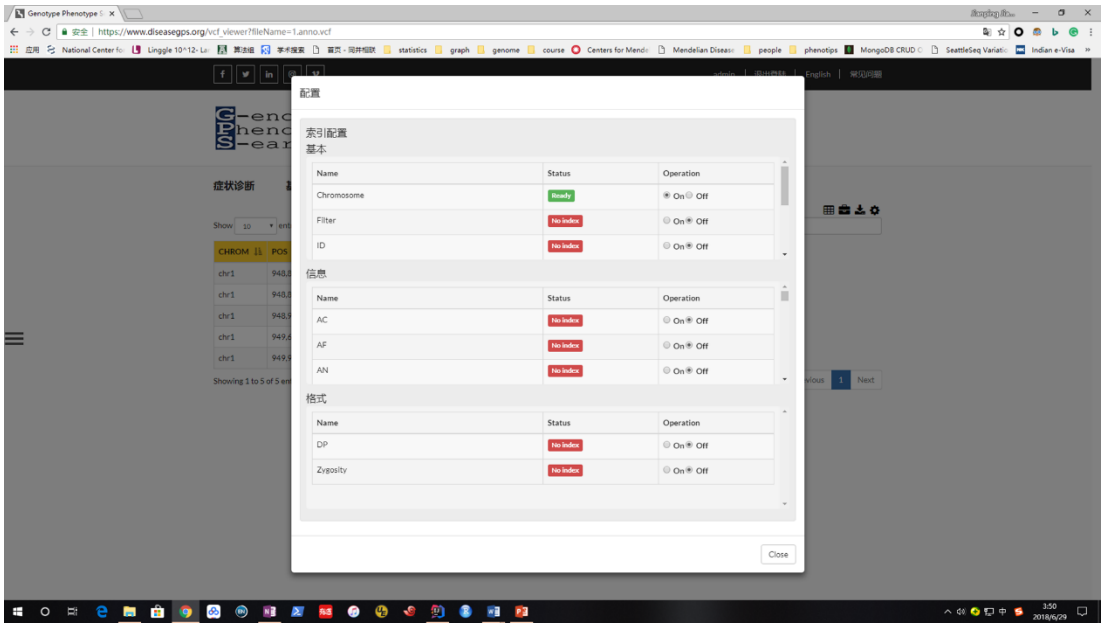


将筛选结果导出为 CSV

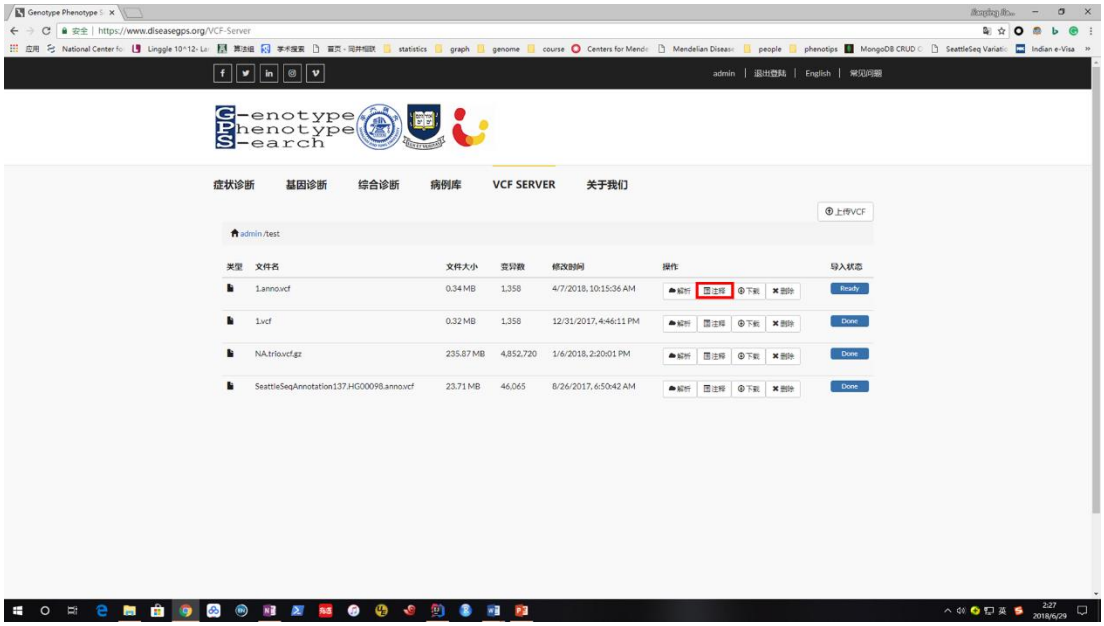


建立列索引(加快筛选速度)



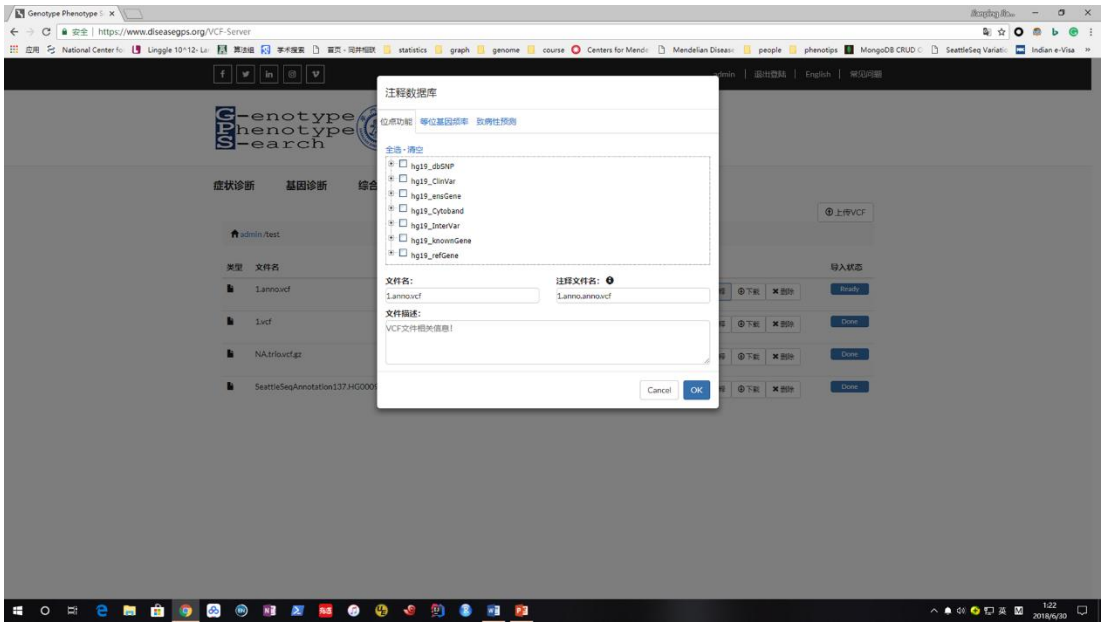


VCF 文件注释

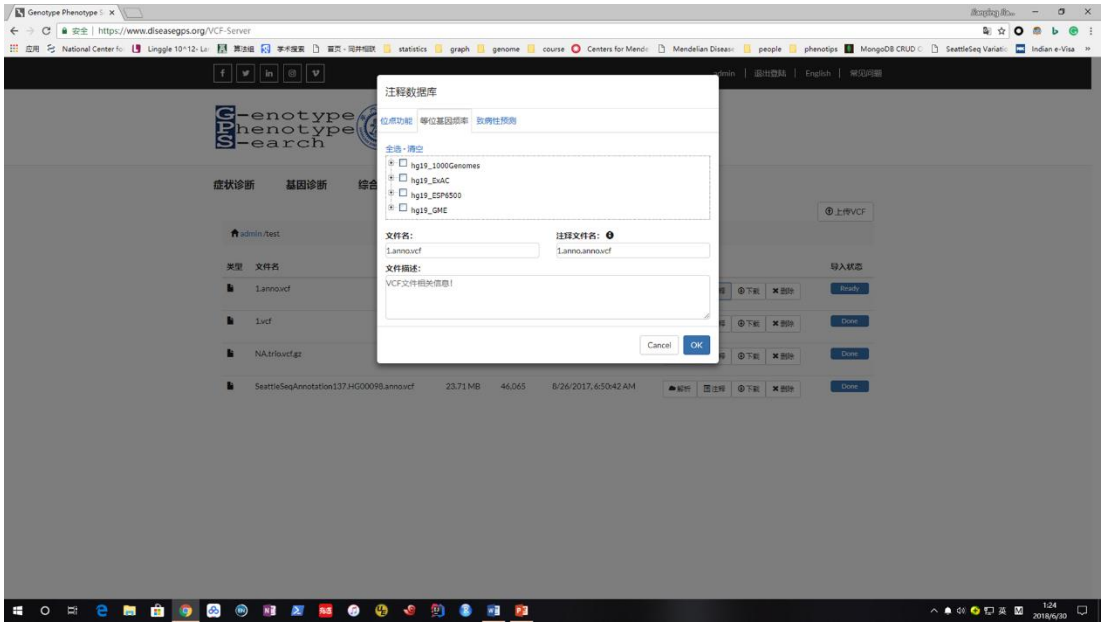


注释数据库选择

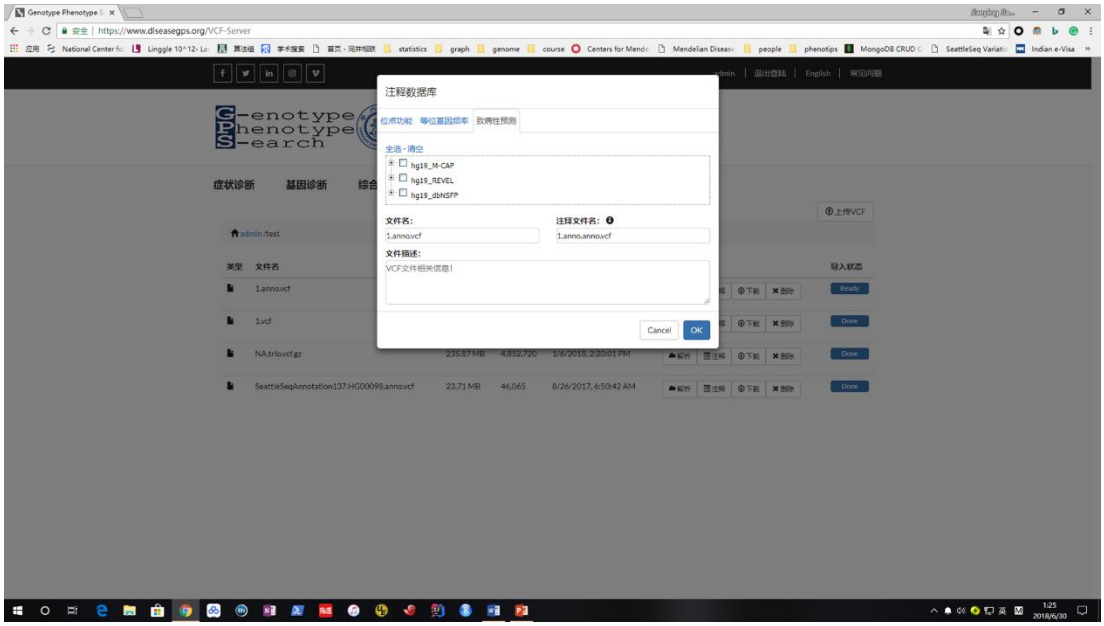
- 位点功能数据库



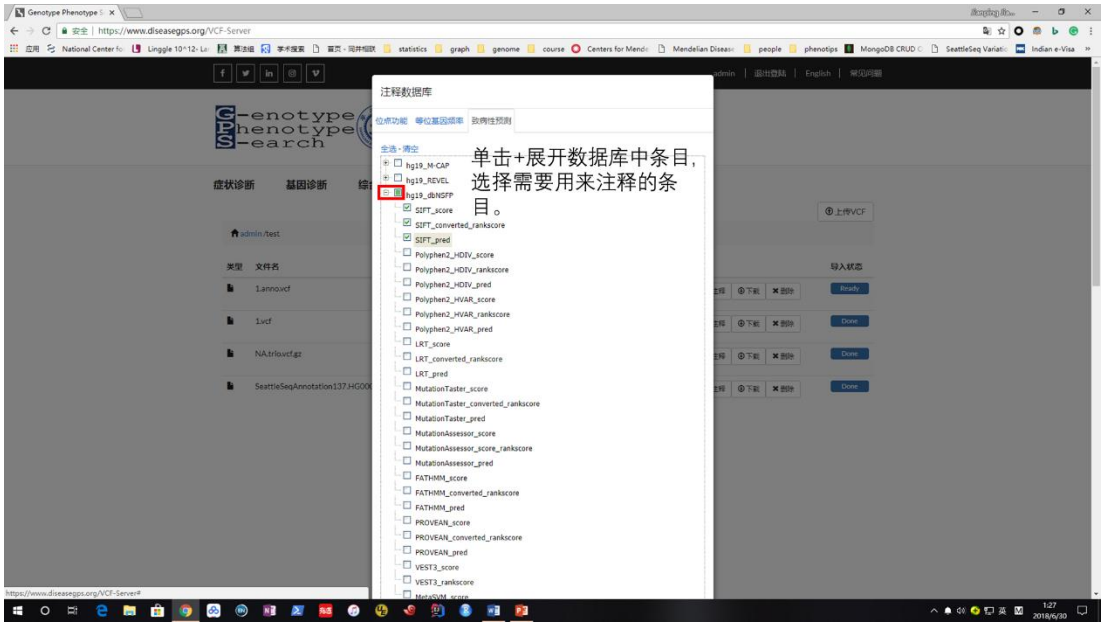
- **等位基因频率数据库**



- **位点致病性预测数据库**



注释数据库中的某些条目

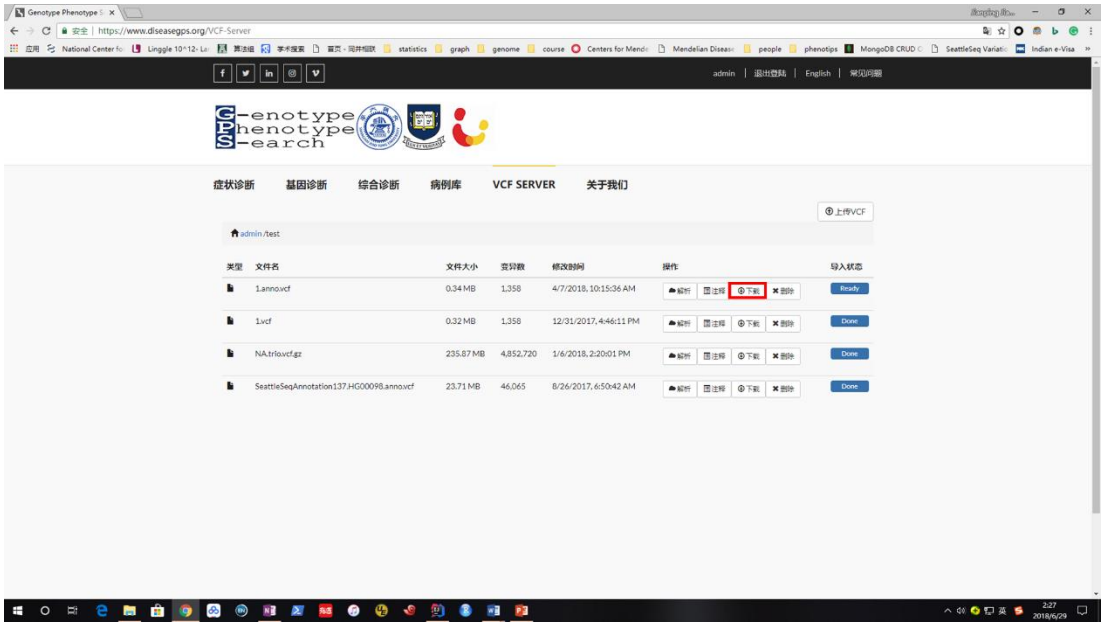


注释数据库信息

Table S1. Summary of databases used in VCF Server annotation

Category	Database	Version/RelaseDate	Description
Variants Function	**dbSNP**	150	The Single Nucleotide Polymorphism database (dbSNP) is a public-domain archive for a broad collection of simple genetic polymorphisms.
	knownGene	201706	The University of California Santa Cruz (UCSC) Known Genes dataset is constructed by a fully automated process, based on protein data from Swiss-Prot/TrEMBL (UniProt) and the associated mRNA data from Genbank.
	ensGene	201706	The Ensembl gene set is based on evidence, and includes manual annotation for our most used species.
	refGene	201706	The RefGene database was created from the UCSC database. RefGene specifies known human protein-coding and non-protein-coding genes taken from the NCBI RNA reference sequences collection (RefSeq).
	Cytoband	201707	The chromosome band track represents the approximate location of bands seen on Giemsa-stained chromosomes.
	ClinVar	201709	ClinVar is a freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting evidence.
	InterVar	201702	InterVar: clinical interpretation of missense variants based on ACMG standards and guidelines.
Allele Frequency	**ExAC**	0.3	ExAC 65000 exome allele frequency data for ALL, AFR (African), AMR (Admixed American), EAS (East Asian), FIN (Finnish), NFE (Non-finnish European), OTH (other), SAS (South Asian).
	1000Genomes	201708	1000 Genome allele frequency data for ALL, AFR (African), AMR (Admixed American), EAS (East Asian), EUR (European), SAS (South Asian).
	GME	201610	Great Middle East allele frequency including NWA (northwest Africa), NEA (northeast Africa), AP (Arabian peninsula), Israel, SD (Syrian desert), TP (Turkish peninsula) and CA (Central Asia).
Function Prediction	**M-CAP**	201610	M-CAP scores for non-synonymous variants.
	REVEL	201610	Rare Exome Variant Ensemble Learner scores for all possible human missense variants.
	dbNSFP	0.33a	dbNSFP is an annotation database for non-synonymous SNPs assembled by Xiaoming Liu from the University of Texas School of Public Health.

VCF 文件下载



VCF 文件删除