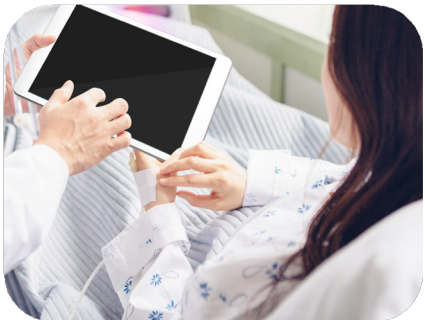


건강한 국민, 안전한 사회

한국인 참조패널 기반 imputation 분석 서비스

미래의료연구부 유전체연구기술개발과



The background features two large, expressive brushstrokes. A vibrant red stroke starts from the top left and curves downwards towards the center. A bright blue stroke starts from the top right and curves downwards towards the center. The two strokes meet at the bottom, creating a frame for the central text.

**한국인 참조패널 기반
Imputation 분석 서비스**

- Preprocessing input data
- Running Imputation Service
- Checking output data

The background features two large, expressive brushstrokes. A vibrant red stroke starts at the top left and curves downwards towards the center. A bright blue stroke starts at the top right and curves downwards towards the center. The two strokes meet at the bottom, creating a frame around the central text. The overall aesthetic is artistic and modern.

Preprocessing Input Data

Introduction : Bcftools

- Utilities for variant calling and manipulating VCFs
- You can see a detailed description of bcftools in
(<https://samtools.github.io/bcftools/bcftools.html>)

Commands	Mean	Basic usage
bcftools view	Subset, filter and convert VCF files	bcftools view FILE less -S
bcftools sort	Sort VCF files	bcftools sort FILE
bcftools index	Index VCF	bcftools index FILE
bcftools concat	Concatenate VCF files from the same set of samples	bcftools concat FILE1 FILE2 bcftools concat --file-list

- Checking the human genome reference version
- GRCh37 : b37, hs37, hg19, grch37, contig ID = 1 ... 22, etc.
- GRCh38 : b38, hs38, hg38, grch38, contig ID = chr1 ... chr22, etc.

```
bgzip -c KBAv2.0_KOGO_QC.vcf > KBAv2.0_KOGO_QC.vcf.gz
```

```
bcftools view KBAv2.0_KOGO_QC.vcf.gz | less -S
```

- The task of Imputation requires sequential genomic positions
- Sorting VCF file by genomic position

```
bcftools sort KBAv2.0_KOGO_QC.vcf.gz \  
--output-type z --output KBAv2.0_KOGO_QC_SORT.vcf.gz
```

- The removal of variants with a low Minor Allele Frequency (MAF) benefits the computational and multiple testing correction burden in Association

```
bcftools view KBAv2.0_KOGO_QC_SORT.vcf.gz \  
--exclude 'MAF < 0.01' \  
--output-type z --output KBAv2.0_KOGO_QC_SORT_MAF0.1.vcf.gz  
  
tabix -p vcf KBAv2.0_KOGO_QC_SORT_MAF0.1.vcf.gz
```


- Imputation starts with the chromosome
- Splitting VCF file by chromosome

```
for chr in `seq 1 22`; do
  echo $chr;
  bcftools view KBAv2.0_KOGO_QC_SORT_MAF0.1.vcf.gz \
  --regions $chr \
  --output-type z --output KBAv2.0_KOGO_QC_SORT_MAF0.1_CHR$chr.vcf.gz;
Done
```

The background features two large, expressive brushstrokes. A vibrant red stroke starts at the top left and curves downwards towards the center. A bright blue stroke starts at the bottom right and curves upwards towards the center. The two strokes meet at the center, framing the text. The overall effect is dynamic and artistic.

Running Imputation Service

Access to Imputation Service website 1



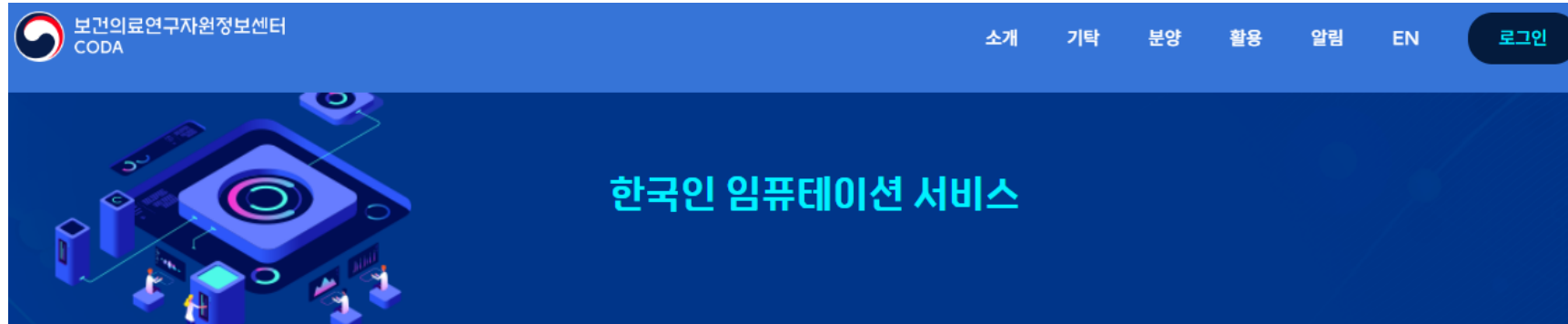
National Institute of Health
Republic of Korea

- <https://coda.nih.go.kr/frt/index.do>
- Click [활용] - [한국인 임퓨테이션 서비스]

The screenshot shows the CODA website homepage. The main header includes the CODA logo and navigation links: 소개, 기탁, 분양, 활용, 알림, EN, and 로그인. A red arrow points to the '활용' (Utilization) menu item, which is highlighted with a red dashed box. Inside this box, the link '한국인 임퓨테이션 서비스' (Korean Imputation Service) is visible. Below the navigation, the main content area features the CODA logo and the text '보건의료연구자원정보센터에서는 연구데이터를 안전하게 관리합니다'. A search bar is present with the text '원하시는 데이터를 키워드로 검색해보세요.' and a search icon. At the bottom, there are social media tags: #코로나, #KoGES, #희귀질환, #자폐증, #대장암. On the right side, there is a statistics section with a DNA double helix icon, showing '151개' (151 items) and '151,133명' (151,133 participants).

Access to Imputation Service website 2

- Click [서비스로 이동]



🏠 > 활용 > 한국인 임퓨테이션 서비스

한국인 임퓨테이션 서비스 (Korean Imputation Service, 이하 KIS)란?

한국인 4,799명 참조패널을 제공하는 웹 기반 임퓨테이션 서비스

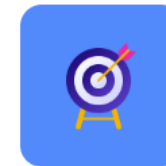
주요 기능



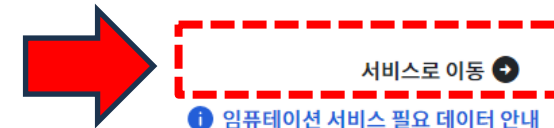
웹 기반 임퓨테이션 서비스를
지원하여
비전문가도 쉽게 분석 가능



KISTI 슈퍼컴퓨터 기반
임퓨테이션
분석 지원으로 빠르게 분석
가능



한국인 4,799명 전염기서열
기반
참조패널 지원으로 정확도
향상

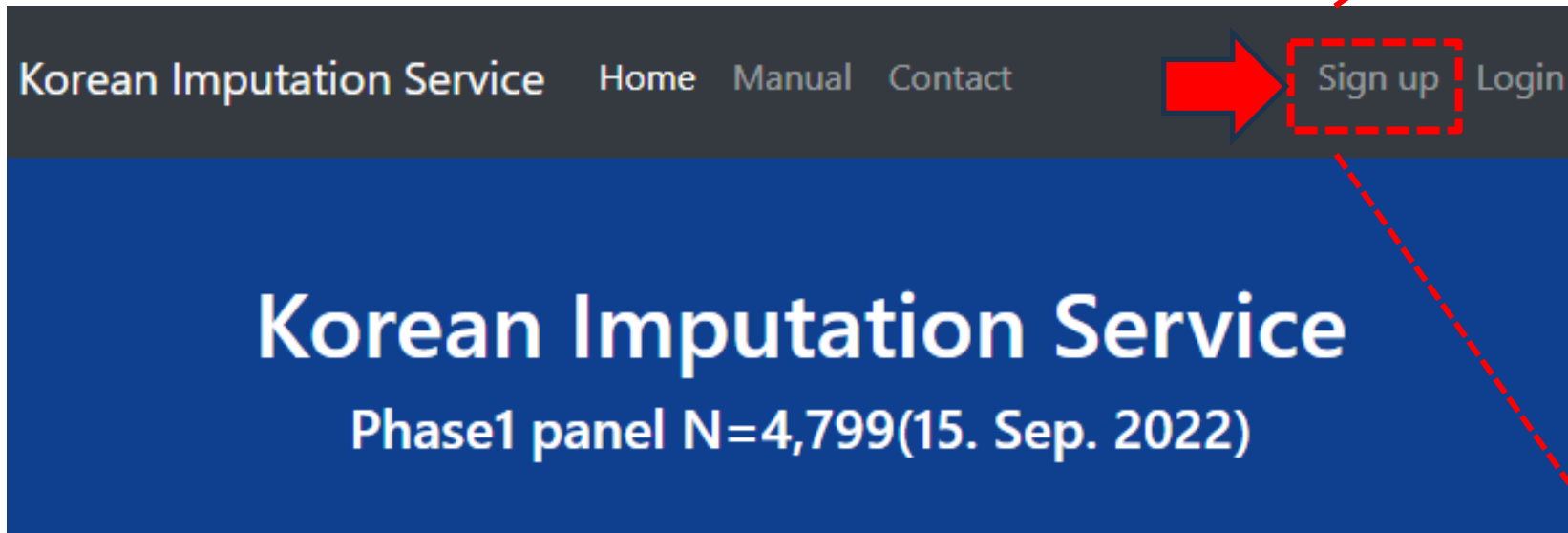


📄 임퓨테이션 서비스 필요 데이터 안내

유전체연구기술개발과 | 최낙현 (043-719-8878)

Sign up for Imputation Service

- Click [Sign up]
- Complete the blanks & Click [Register]



Sign up

Username:

Full Name:

E-Mail:

Password:

Confirm password:

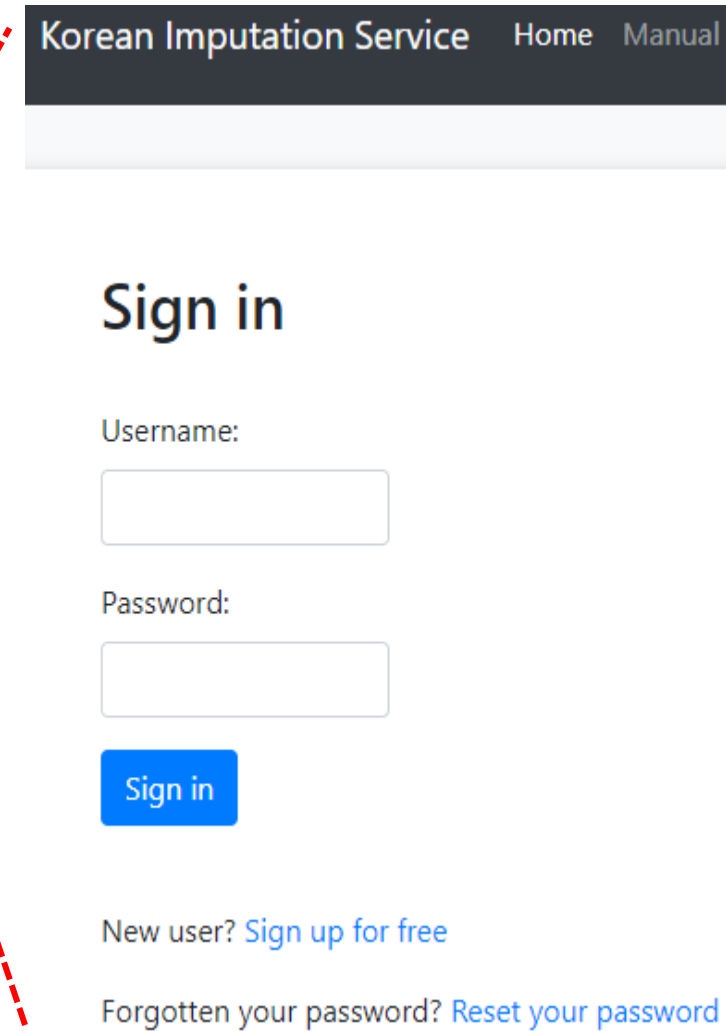
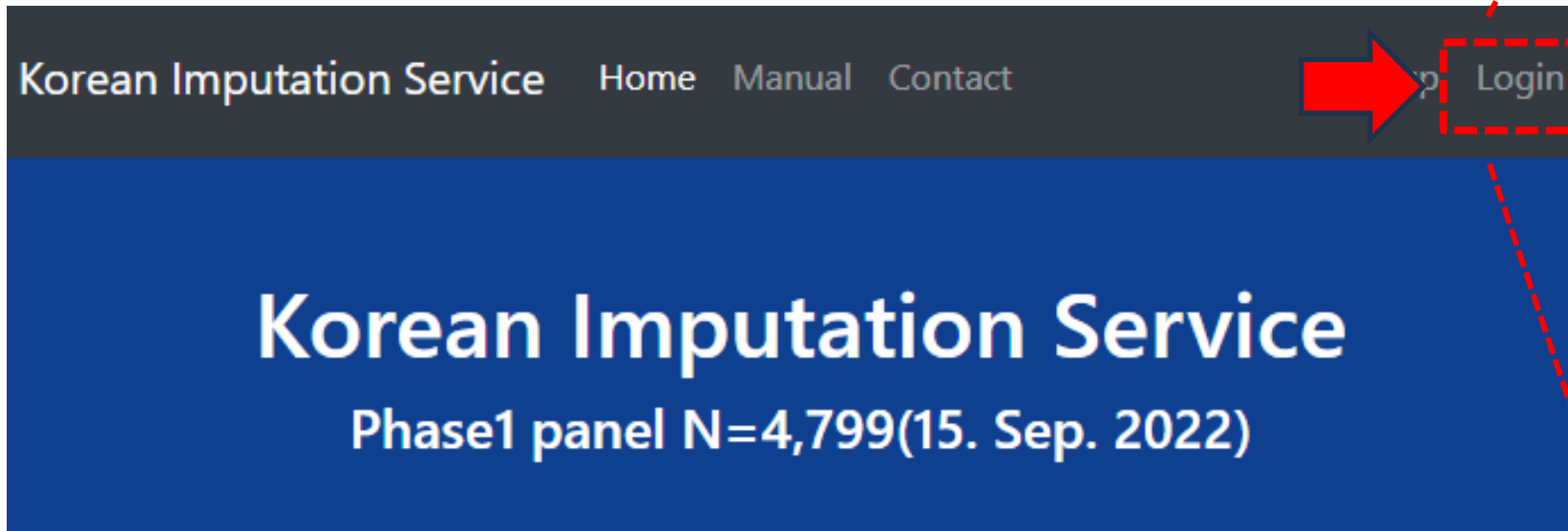
Register

Login for Imputation Service



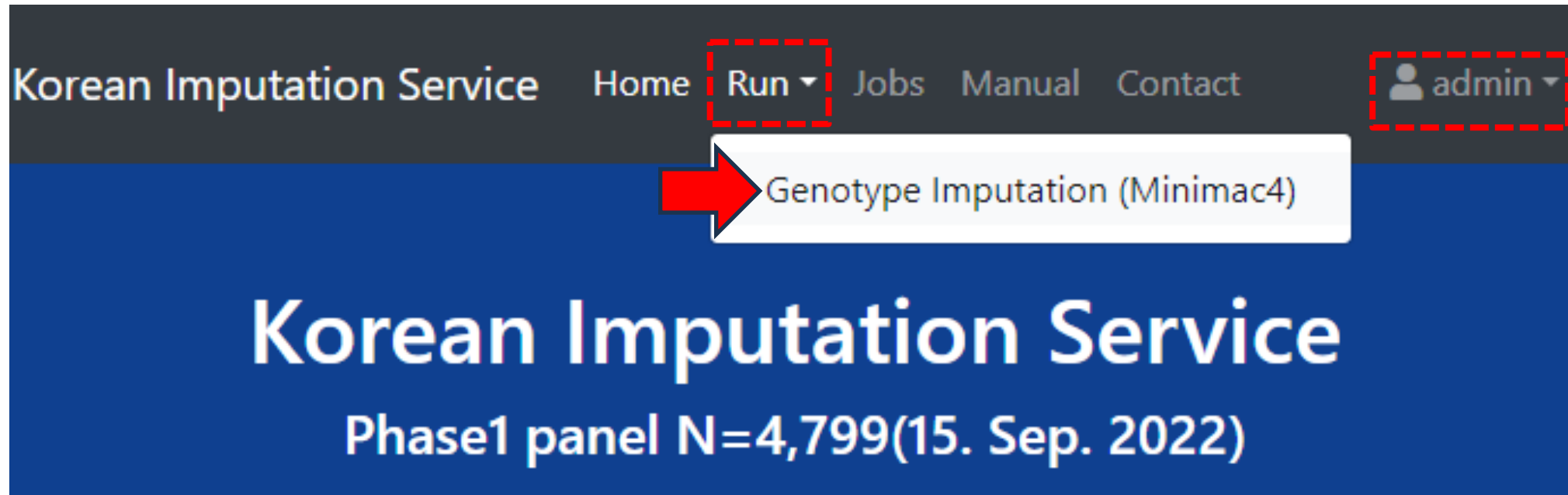
National Institute of Health
Republic of Korea

- Click [Login]
- Complete the blanks & Click [Sign in]



Start Imputation Service : [Run]

- Click [Run] & [Genotype Imputation (Minimac4)]



Korean Imputation Service Home Run Jobs Manual Contact admin

Genotype Imputation (Minimac4)

Korean Imputation Service

Phase1 panel N=4,799(15. Sep. 2022)

Start Imputation Service : Overview



National Institute of Health
Republic of Korea

- Imputation service are based on the Michigan Imputation Server
- The service provide RUN, Details, Results, Logs information

Korean Imputation Service Home Run Jobs Manual Contact admin

Genotype Imputation (Minimac4) 1.6.8

This is the new Michigan Imputation Server Pipeline using Minimac4. Documentation can be found here.

If your input data is **GRCh37/hg19** please ensure chromosomes are encoded without prefix (e.g. **20**).
If your input data is **GRCh38hg38** please ensure chromosomes are encoded with prefix 'chr' (e.g. **chr20**). <https://imputationserver.readthedocs.io>

Run

Name optional job name

Reference Panel (Details) -- select an option --

Input Files (VCF) File Upload

Select Files

Multiple files can be selected by using the **ctrl** / **cmd** or **shift** keys.

Array Build GRCh37/hg19

Please note that the final SNP coordinates always match the reference build.

rsq Filter off

Start Imputation Service : [Name]



- Set study name

Korean Imputation Service Home Run ▾ Jobs Manual Contact

Genotype Imputation (Minimac4) 1.6.8

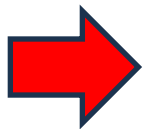
This is the new Michigan Imputation Server Pipeline using Minimac4. Documentation can be found here.

If your input data is **GRCh37/hg19** please ensure chromosomes are encoded without prefix (e.g. **20**).

If your input data is **GRCh38hg38** please ensure chromosomes are encoded with prefix 'chr' (e.g. **chr20**). <https://imputationserver.readthedocs.io>

▶ Run

Name

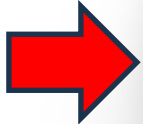


Start Imputation Service : [Reference Panel]

- 2024 KOGO Reference Panel was created by scaling down KIS Phase1 Panel(GRCh37/hg19)
- Select [Reference Panel] - [2024 KOGO(GRCh37)]

Name

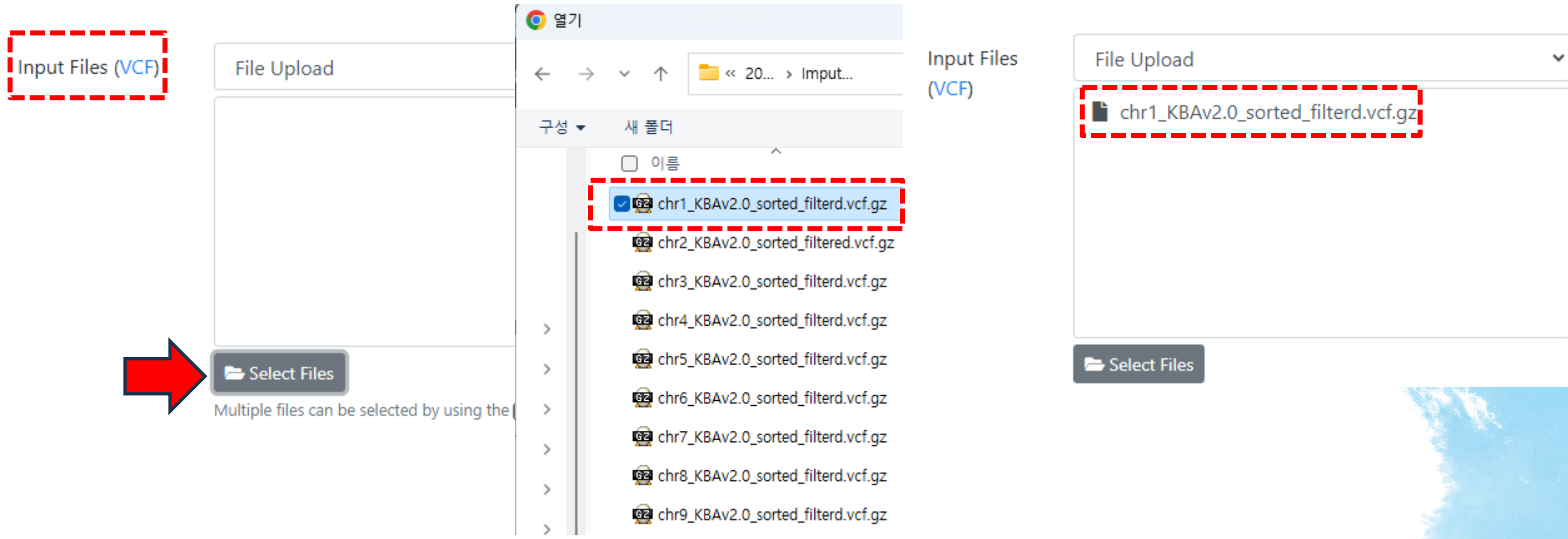
Reference Panel ([Details](#))

Input Files (VCF) 

- 2023 KOGO(GRCh37)
- KIS Phase1 Panel(GRCh37/hg19)

Start Imputation Service : [File Upload]

- Select input data



Input Files (VCF)

File Upload

Select Files

Multiple files can be selected by using the

열기

< > << 20... >> Imput...

구성 새 폴더

이름

chr1_KBAv2.0_sorted_filterd.vcf.gz

chr2_KBAv2.0_sorted_filterd.vcf.gz

chr3_KBAv2.0_sorted_filterd.vcf.gz

chr4_KBAv2.0_sorted_filterd.vcf.gz

chr5_KBAv2.0_sorted_filterd.vcf.gz

chr6_KBAv2.0_sorted_filterd.vcf.gz

chr7_KBAv2.0_sorted_filterd.vcf.gz

chr8_KBAv2.0_sorted_filterd.vcf.gz

chr9_KBAv2.0_sorted_filterd.vcf.gz

Input Files (VCF)

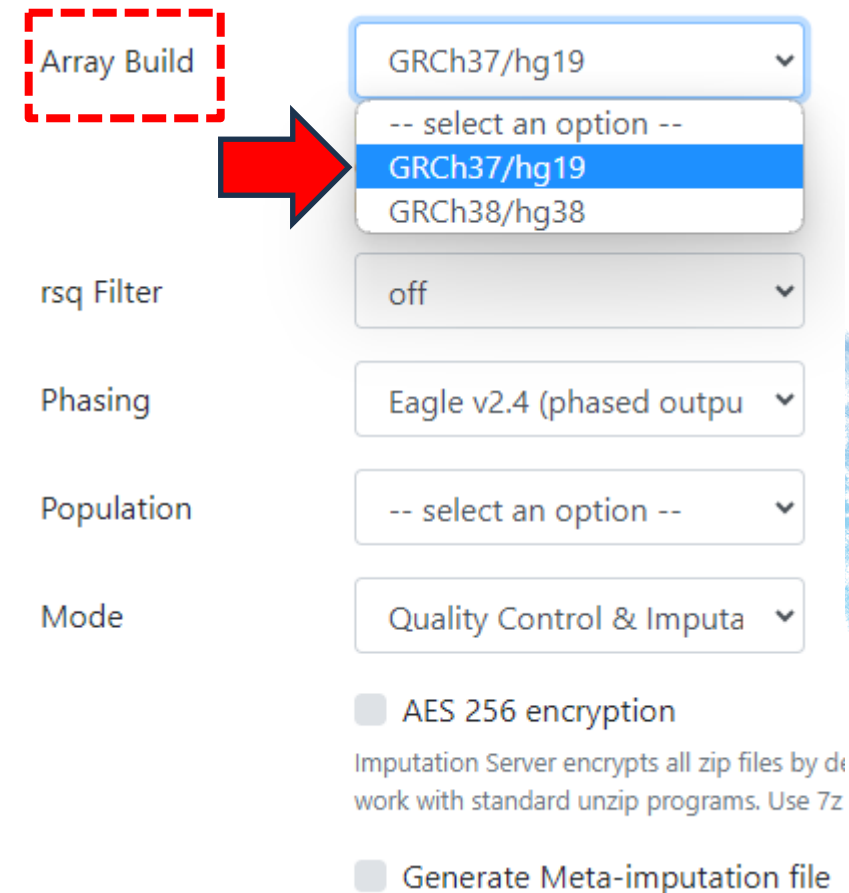
File Upload

chr1_KBAv2.0_sorted_filterd.vcf.gz

Select Files

Start Imputation Service : [Array Build option]

- Set [Array Build] for the build of your data according to the human genome reference version
- Input data (GRCh37/hg19) : chromosomes are encoded without prefix (e.g. 1 ... 22)
- Input data (GRCh38/hg38) : chromosomes are encoded with prefix (e.g. chr1 ... chr22)
- Server automatically updates the genome positions (liftOver) of your data



Array Build

GRCh37/hg19

-- select an option --

GRCh37/hg19

GRCh38/hg38

rsq Filter off

Phasing Eagle v2.4 (phased output)

Population -- select an option --

Mode Quality Control & Imputation

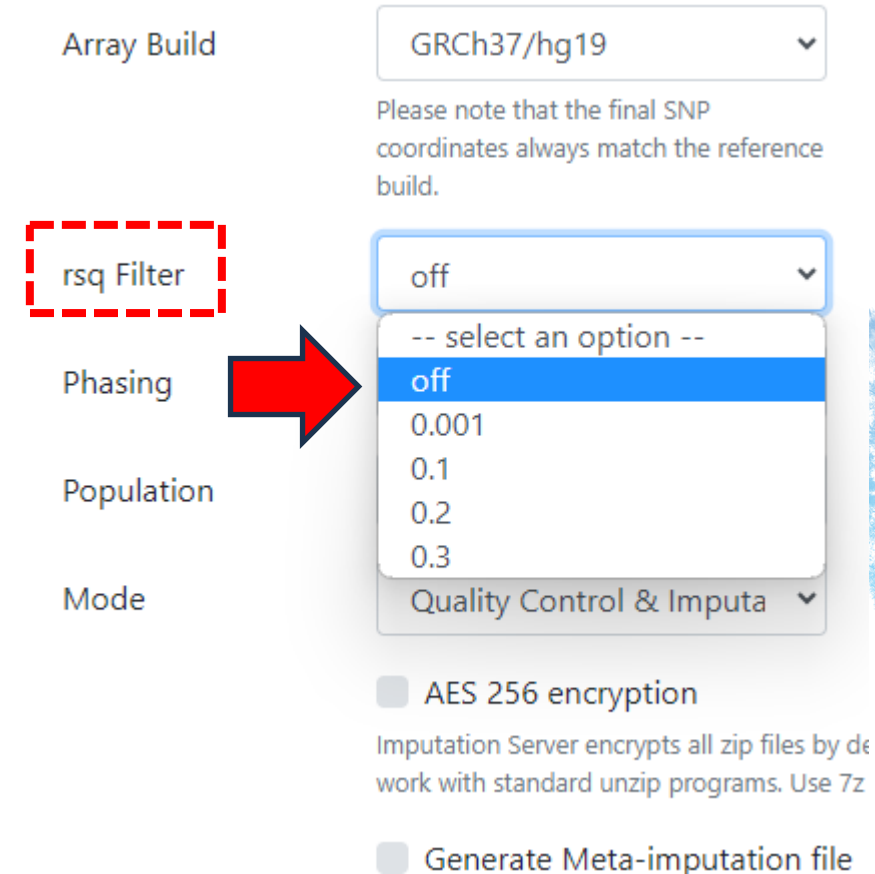
AES 256 encryption

Imputation Server encrypts all zip files by default to work with standard unzip programs. Use 7z

Generate Meta-imputation file

Start Imputation Service : [rsq Filter option]

- Set [rsq Filter] for output
- Purpose to minimize the file size
- Removing all imputed SNPs with a rsq-value(=imputation quality) smaller than the specified value
- Well-imputed variants (estimated rsq ≥ 0.8)



Array Build

Please note that the final SNP coordinates always match the reference build.

rsq Filter

Phasing

Population

Mode

AES 256 encryption
Imputation Server encrypts all zip files by default to work with standard unzip programs. Use 7z

Generate Meta-imputation file

A red dashed box highlights the 'rsq Filter' dropdown, and a red arrow points to the 'off' option in the expanded menu.

Start Imputation Service : [Phasing option]


- Set [Phasing] for your data according to phase status
- Unphased data(0/0) : selecting [Eagle v2.4]
- Phased data(0|0) : selecting [No phasing]

Array Build

Please note that the final SNP coordinates always match the reference build.

rsq Filter

Phasing
-- select an option --
Eagle v2.4 (phased output)
No phasing

Population 

Mode

AES 256 encryption
Imputation Server encrypts all zip files by default to work with standard unzip programs. Use 7z

Generate Meta-imputation file

Start Imputation Service : [Population option]

- Set [Population] for your data
- Selecting the population of your uploaded samples
- The option [all] is only supported

Array Build

Please note that the final SNP coordinates always match the reference build.

rsq Filter

Phasing

Population

Mode

 AES 256 encryption

Imputation Server encrypts all zip files by default to work with standard unzip programs. Use 7z

Generate Meta-imputation file

Start Imputation Service : [Mode option]



National Institute of Health
Republic of Korea

- Set [Mode] for your data
- Mode : [Quality Control & Imputation],
[Quality Control & Phasing Only],
[Quality Control Only]
- Selecting [Quality Control & Imputation]

Array Build

Please note that the final SNP coordinates always match the reference build.

rsq Filter

Phasing

Population

Mode

Generate Meta-imputation file

Start Imputation Service : [Other options & Submit Job]

- [AES 256 encryption] : Output is encrypted by default. If you want to use AES 256 encryption, tick this checkbox
- [Generate Meta-imputation file] : Generating empirical.dose.vcf.gz for meta-imputation
- [Submit Job] : Start your imputation job

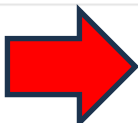
AES 256 encryption

Imputation Server encrypts all zip files by default. Please note that AES encryption does not work with standard unzip programs. Use 7z instead.

Generate Meta-imputation file

I will not attempt to re-identify or contact research participants.

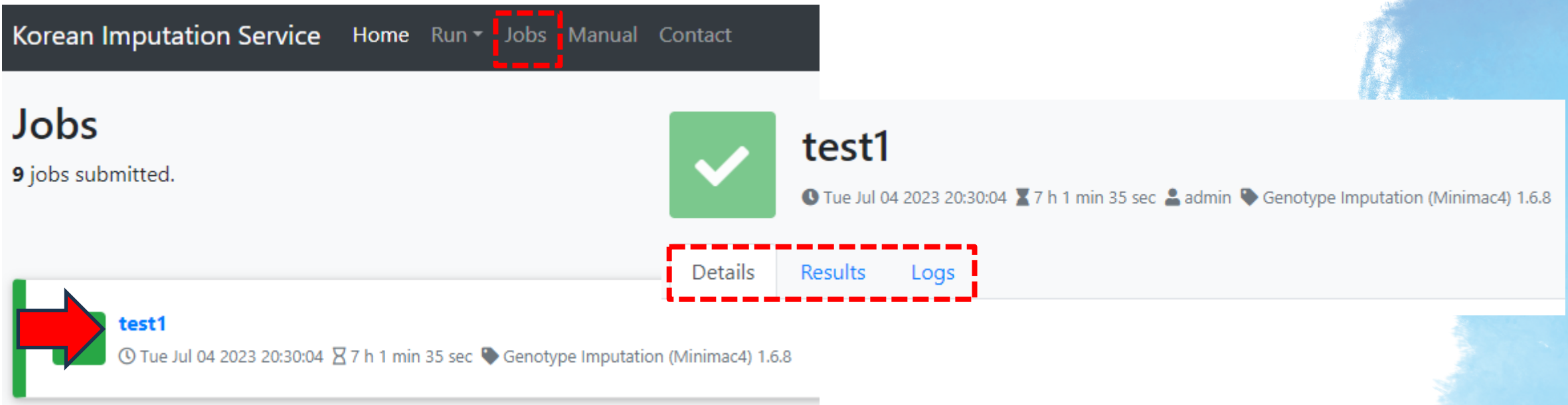
I will report any inadvertent data release, security breach or other data management incident of which I become aware.



The background features two large, expressive brushstrokes. A vibrant red stroke starts at the top left and curves downwards towards the center. A bright blue stroke starts at the bottom right and curves upwards towards the center. The two strokes meet at the center, framing the text. The overall style is artistic and hand-drawn.

Checking Output Data


- Click [Job] - Your study name
- [Details] : Showing current status of your job
- [Results] : Showing output of your job
- [Logs] : : Showing current workflow of your job




Korean Imputation Service Home Run **Jobs** Manual Contact

Jobs

9 jobs submitted.

**test1**
🕒 Tue Jul 04 2023 20:30:04 ⌚ 7 h 1 min 35 sec 👤 admin 📁 Genotype Imputation (Minimac4) 1.6.8

[Details](#) [Results](#) [Logs](#)

**test1**
🕒 Tue Jul 04 2023 20:30:04 ⌚ 7 h 1 min 35 sec 📁 Genotype Imputation (Minimac4) 1.6.8

- [Input Validation] : Check uploaded files
- Calculate some basic statistics such as amount of samples, chromosomes and SNPs

Input Validation

22 valid VCF file(s) found.

Samples: 504

Chromosomes: 10 11 12 13 14 15 16 17 18 19 1 20 21 22 2 3 4 5 6 7 8 9

SNPs: 1094021

Chunks: 154

Datatype: phased

Build: hg19

Reference Panel: apps@2023-kogo(grch37)@1.0.0 (hg19)

Population: all

Phasing: eagle

Mode: imputation

Details : [Quality Control]



- [Quality Control] : Check each variant and exclude it (Invalid alleles, duplicates, indels, monomorphic sites, allele mismatch between reference panel and uploaded data, SNP call rate < 90%)
- All filtered variants are listed in files (chunks-excluded.txt, snps-excluded.txt)

Quality Control

Calculating QC Statistics

Statistics:

Alternative allele frequency > 0.5 sites: 178,441

Reference Overlap: 97.55 %

Match: 1,006,905

Allele switch: 0

Strand flip: 0

Strand flip and allele switch: 0

A/T, C/G genotypes: 0

Filtered sites:

Filter flag set: 0

Invalid alleles: 20,554

Multiallelic sites: 0

Duplicated sites: 0

NonSNP sites: 0

Monomorphic sites: 41,164

Allele mismatch: 63

SNPs call rate < 90%: 0

Excluded sites in total: 61,781

Remaining sites in total: 1,006,905

See [snps-excluded.txt](#) for details

Typed only sites: 25,335

See [typed-only.txt](#) for details

Warning: 1 Chunk(s) excluded: reference overlap < 50.0% (see [chunks-excluded.txt](#) for details).

Remaining chunk(s): 153

Details : [Pre-phasing & Imputation & Data Compression]

- [Pre-phasing] : Execute Eagle2 algorithm for each 2Mb chunk with 5Mb window size
- [Imputation] : Execute Minimac4 algorithm for each 2Mb chunk with 500 kb window size
- [Data Compression and Encryption] : Keep your password for output

Pre-phasing and Imputation

Chr 11 Chr 22 Chr 12 Chr 13 Chr 14 Chr 15
Chr 16 Chr 17 Chr 18 Chr 19 Chr 1 Chr 2
Chr 3 Chr 4 Chr 5 Chr 6 Chr 7 Chr 8
Chr 9 Chr 20 Chr 21 Chr 10

■ Waiting
■ Running
■ Complete

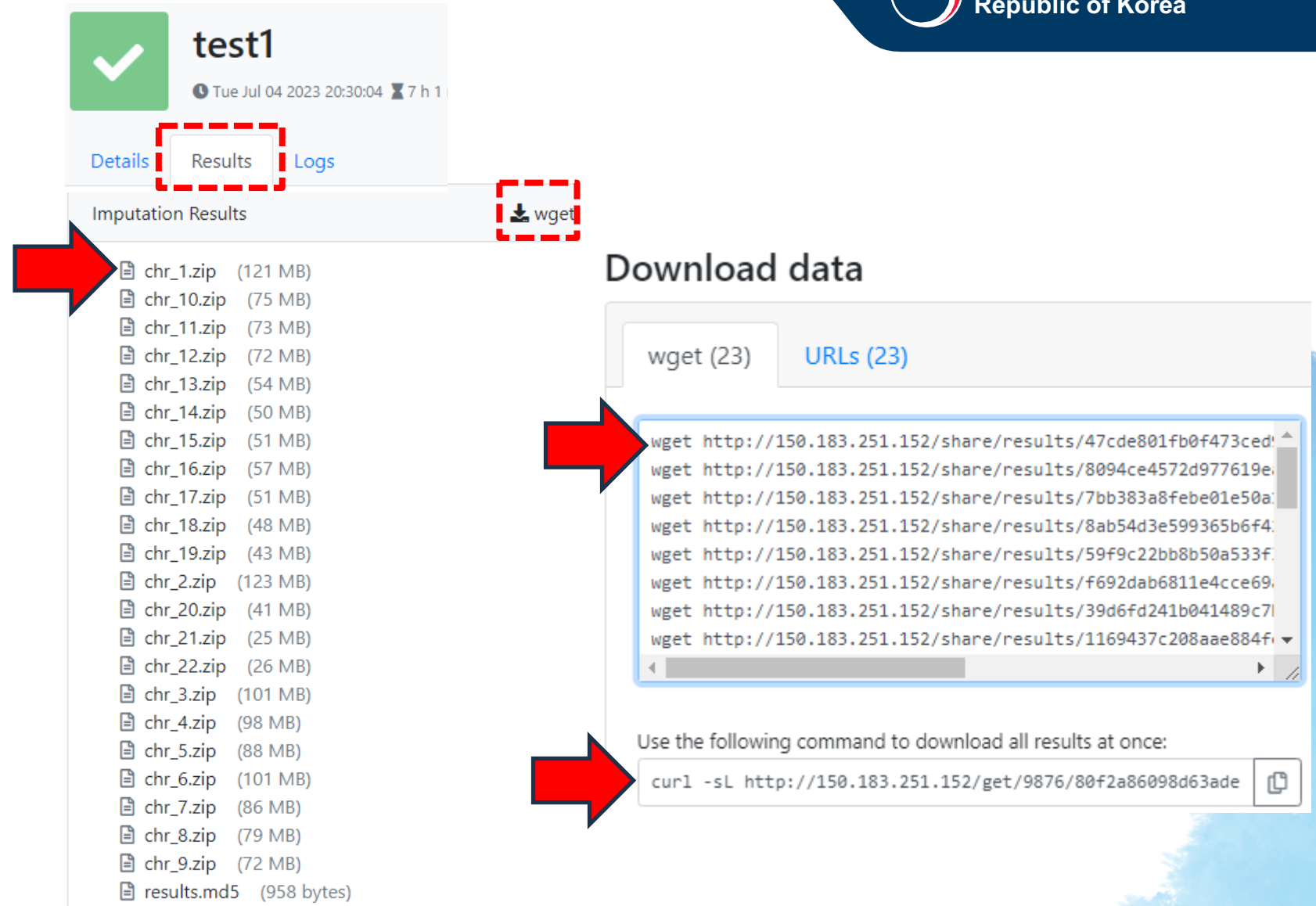
Data Compression and Encryption

Exported data.

Email notification is disabled. All results are encrypted with password **L6(vr3t)SwTMKd**

Results : [Imputation Results]

- [Imputation Results] :
Download your output
- Click each output
- Copy commands &
paste in command
window



test1
Tue Jul 04 2023 20:30:04 7 h 1

Details Results Logs

Imputation Results wget

- chr_1.zip (121 MB)
- chr_10.zip (75 MB)
- chr_11.zip (73 MB)
- chr_12.zip (72 MB)
- chr_13.zip (54 MB)
- chr_14.zip (50 MB)
- chr_15.zip (51 MB)
- chr_16.zip (57 MB)
- chr_17.zip (51 MB)
- chr_18.zip (48 MB)
- chr_19.zip (43 MB)
- chr_2.zip (123 MB)
- chr_20.zip (41 MB)
- chr_21.zip (25 MB)
- chr_22.zip (26 MB)
- chr_3.zip (101 MB)
- chr_4.zip (98 MB)
- chr_5.zip (88 MB)
- chr_6.zip (101 MB)
- chr_7.zip (86 MB)
- chr_8.zip (79 MB)
- chr_9.zip (72 MB)
- results.md5 (958 bytes)

Download data

wget (23) URLs (23)

```
wget http://150.183.251.152/share/results/47cde801fb0f473ced'  
wget http://150.183.251.152/share/results/8094ce4572d977619e.  
wget http://150.183.251.152/share/results/7bb383a8febe01e50a:  
wget http://150.183.251.152/share/results/8ab54d3e599365b6f4.  
wget http://150.183.251.152/share/results/59f9c22bb8b50a533f.  
wget http://150.183.251.152/share/results/f692dab6811e4cce69.  
wget http://150.183.251.152/share/results/39d6fd241b041489c7/  
wget http://150.183.251.152/share/results/1169437c208aae884f
```

Use the following command to download all results at once:

```
curl -sL http://150.183.251.152/get/9876/80f2a86098d63ade
```

- Decompress output with password using unzip
- chr#.dose.vcf.gz : Output for association and other analysis
- chr#.empiricalDose.vcf.gz : Output for meta-imputation with MetaMinimac2
- chr#.info.gz : Summary statistics from chr#.dose.vcf.gz

Data Compression and Encryption

Exported data.

Email notification is disabled. All results are encrypted with password **L6(vr3t)SwTMKd**

```
unzip -P c6hsGfDZOtljr2 chr_21.zip
```


Decompressed Output 1

- chr#.dose.vcf.gz : Imputation output for association and other analysis

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	HG00403	HG004
##fileformat=VCFv4.1										
##filedate=2023.7.4										
##contig=<ID=1>										
##INFO=<ID=AF,Number=1,Type=Float,Description="Estimated Alternate Allele Frequency">										
##INFO=<ID=MAF,Number=1,Type=Float,Description="Estimated Minor Allele Frequency">										
##INFO=<ID=R2,Number=1,Type=Float,Description="Estimated Imputation Accuracy (R-square)">										
##INFO=<ID=ER2,Number=1,Type=Float,Description="Empirical (Leave-One-Out) R-square (available only for genotyped variants)">										
##INFO=<ID=IMPUTED,Number=0,Type=Flag,Description="Marker was imputed but NOT genotyped">										
##INFO=<ID=TYPED,Number=0,Type=Flag,Description="Marker was genotyped AND imputed">										
##INFO=<ID=TYPED_ONLY,Number=0,Type=Flag,Description="Marker was genotyped but NOT imputed">										
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">										
##FORMAT=<ID=DS,Number=1,Type=Float,Description="Estimated Alternate Allele Dosage : [P(0/1)+2*P(1/1)]">										
##FORMAT=<ID=HDS,Number=2,Type=Float,Description="Estimated Haploid Alternate Allele Dosage ">										
##FORMAT=<ID=GP,Number=3,Type=Float,Description="Estimated Posterior Probabilities for Genotypes 0/0, 0/1 and 1/1 ">										
##pipeline=michigan-imputationserver-1.6.8										
##imputation=minimac4-1.0.2										
##phasing=n/a										
##panel=apps@2023-kogo(grch37)@1.0.0										
##r2Filter=0.0										
#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	HG00403	HG004
1	10682	1:10682:A:A	A	G	.	PASS	AF=0.00323;MAF=0.00323;R2=0.17328;IMPUTED	GT:DS:HDS:GP	0 0:0.004:0.003,0.001:0.996,0.004,0.000	0 0:0:0
1	10687	1:10687:G:G	G	C	.	PASS	AF=0.00023;MAF=0.00023;R2=0.00428;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0	0 0:0:0
1	54548	1:54548:C:C	C	T	.	PASS	AF=0.00012;MAF=0.00012;R2=0.01052;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0	0 0:0:0
1	69610	1:69610:C:C	C	T	.	PASS	AF=0.01091;MAF=0.01091;TYPED_ONLY	GT:DS:HDS:GP	0 0:0:0,0:1,0,0	0 0:0:0
1	71886	1:71886:T:T	T	C	.	PASS	AF=0.00036;MAF=0.00036;R2=0.02017;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0	0 0:0:0

Decompressed Output 2

- chr#.empiricalDose.vcf.gz : Output for meta-imputation with MetaMinimac2

```
##fileformat=VCFv4.1
##filedate=2023.7.4
##contig=<ID=1>
##INFO=<ID=TYPED,Number=0,Type=Flag,Description="Marker was genotyped AND imputed">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotyped alleles from Array">
##FORMAT=<ID=LDS,Number=1,Type=String,Description="Leave-one-out Imputed Dosage : Estimated Haploid Alternate Allele Dosage assuming site was NOT genotyped ">
##pipeline=michigan-imputationserver-1.6.8
##imputation=minimac4-1.0.2
##phasing=n/a
##panel=apps@2023-kogo(grch37)@1.0.0
##r2Filter=0.0
```

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	HG00403	HG00404	HG00406	HG00407	HG00409	HG00410	HG00411
1	713250	1:713250:C:G	C	C	.	PASS	TYPED	GT:LDS	0 0:0.001 0.001	0 0:0.000 0.000	0 0:0.001 0.001	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.001 0.001	0 0:0.000 0.000
1	729957	1:729957:T:T	C	C	.	PASS	TYPED	GT:LDS	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000
1	738539	1:738539:T:T	C	C	.	PASS	TYPED	GT:LDS	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000	0 0:0.000 0.000
1	739132	1:739132:A:A	C	C	.	PASS	TYPED	GT:LDS	0 0:0.002 0.008	0 0:0.001 0.002	0 0:0.005 0.005	0 0:0.000 0.000	0 0:0.002 0.002	0 0:0.000 0.000	0 0:0.000 0.000

- Integrate the imputed results using MetaMinimac2 (<https://github.com/yukt/MetaMinimac2>)

In terminal

```
MetaMinimac2 -i PanelA.imputed:PanelB.imputed -o A_B.meta.run
```

Decompressed Output 3

- chr#.info.gz : Summary statistics from chr#.dose.vcf.gz

SNP	REF(0)	ALT(1)	ALT_Frq	MAF	AvgCall	Rsqr	Genotyped	LooRsqr	EmpR	EmpRsqr	Dose0	Dose1
1:52955:A:G	A	G	0.00023	0.00023	0.99977	0.00865	Imputed	-	-	-	-	-
1:54548:C:T	C	T	0.00012	0.00012	0.99988	0.01052	Imputed	-	-	-	-	-
1:69610:C:T	C	T	0.01091	0.01091	-	-	Typed_Only	-	-	-	-	-
1:71886:T:C	T	C	0.00036	0.00036	0.99964	0.02017	Imputed	-	-	-	-	-

In terminal

```
pigz -d ./chr#.info.gz
```

Rscript output.R Start

```
[1] "The number of imputed variants ..."  
# A tibble: 3 × 2  
# Groups:   Genotyped [3]  
  Genotyped     n  
  <fct>       <int>  
1 Typed_Only  25325  
2 Genotyped  1006901  
3 Imputed     581332  
[1] "The number of WellImputed variants per AF ..."  
# A tibble: 8 × 3  
# Groups:   WellImputed, Range [8]  
  WellImputed Range     n  
  <chr>       <fct>       <int>  
1 Imputed     AF<=0.1      283230  
2 Imputed     0.1<AF<=0.5  82793  
3 Imputed     0.5<AF<=5    32601  
4 Imputed     5<AF         11913  
5 WellImputed AF<=0.1      4858  
6 WellImputed 0.1<AF<=0.5 23797  
7 WellImputed 0.5<AF<=5    33976  
8 WellImputed 5<AF         108164
```



- Merge all output into one file for Association

```
bcftools concat --file-list merge_list.txt \  
--output-type z --output KBAv2.0AB_KOGO_IMP.vcf.gz --threads 10  
bcftools index KBAv2.0AB_KOGO_IMP.vcf.gz
```

건강한 국민, 안전한 사회

