



**HKU  
Med** School of Clinical Medicine  
Department of Surgery  
香港大學外科學系

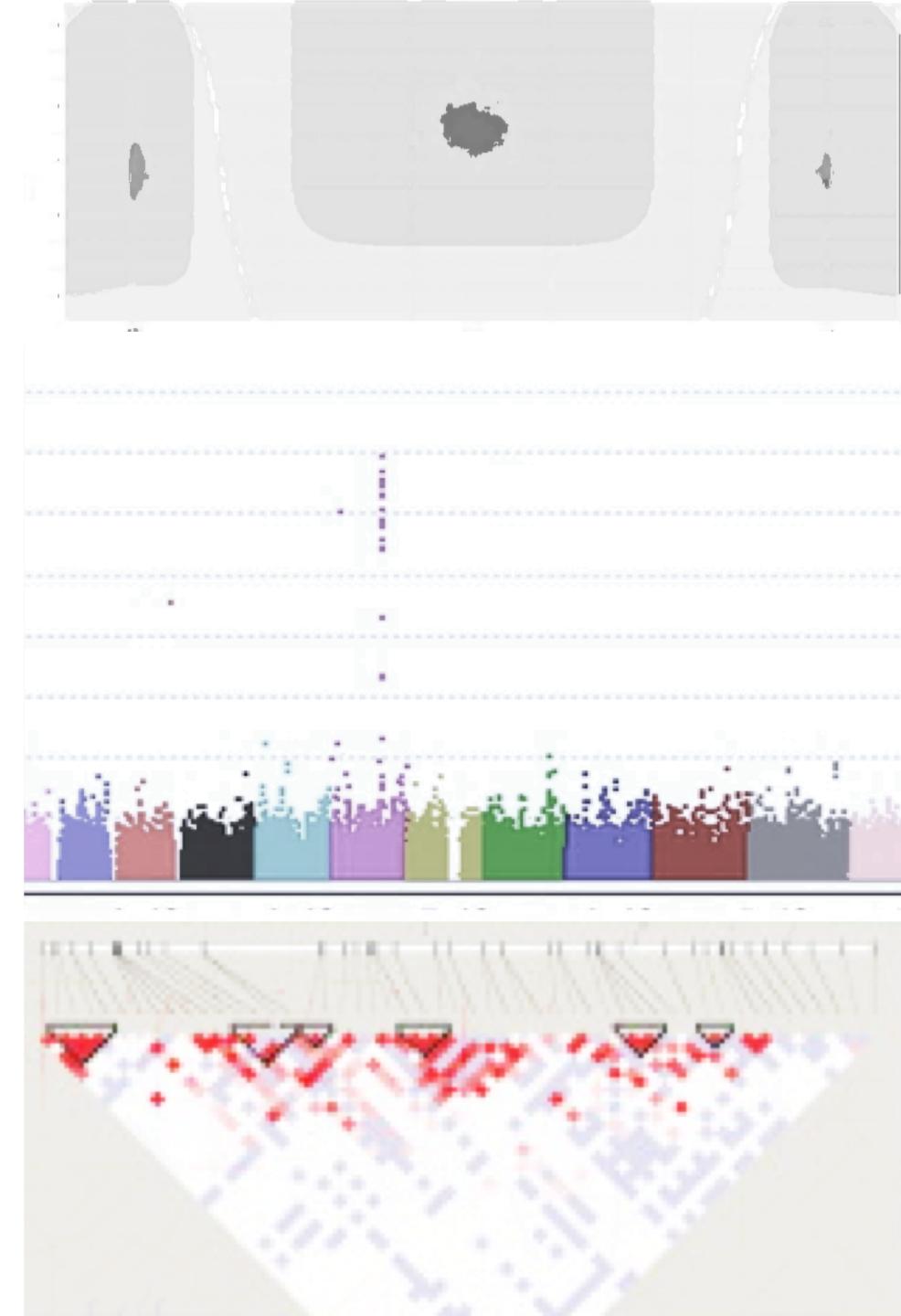


Dr Li Dak-Sum Research Centre  
The University of Hong Kong - Karolinska Institutet  
Collaboration in Regenerative Medicine  
香港大學 - 卡羅琳斯卡學院再生醫學合作計劃  
李達三博士研究中心

# Genotype imputation: Tools, panels and online resources

Clara Tang

June 2022



- **What is genotype imputation?**
- **Why do we need imputation?**
- **How to perform imputation?**



# Genotype imputation

- Genotype imputation refers to the process of predicting the genotypes that are not directly assayed
- **Major purposes:**
  - To recover variants not included or not called in the SNP array

# Typical imputation scenario

1000 Genome Phase

3 v5 (n=2,504):

~49M nonsingleton polymorphic markers

TOPMed vR2 (n= 97,256):

308M polymorphic markers

0	0	1	1	1	0	0	1	1	0	0	0	1	1	1
0	0	0	0	0	1	1	1	0	1	1	1	0	0	1
1	1	1	1	1	0	0	0	1	0	0	0	0	0	0
1	0	1	1	0	0	0	1	1	1	1	1	0	0	1

→

Reference haplotypes

1	?	?	?	2	?	0	?	?	?	?	0	1	?	1
1	?	?	?	1	?	0	?	?	?	?	?	0	?	0
0	?	?	?	1	?	1	?	?	?	?	1	0	?	1
1	?	?	?	2	?	0	?	?	?	?	0	1	?	1
?	?	?	?	2	?	0	?	?	?	?	0	0	?	0
1	?	?	?	1	?	1	?	?	?	?	1	0	?	?
0	?	?	?	2	?	0	?	?	?	?	0	1	?	1
1	?	?	?	1	?	1	?	?	?	?	1	1	?	2

→

Study genotypes



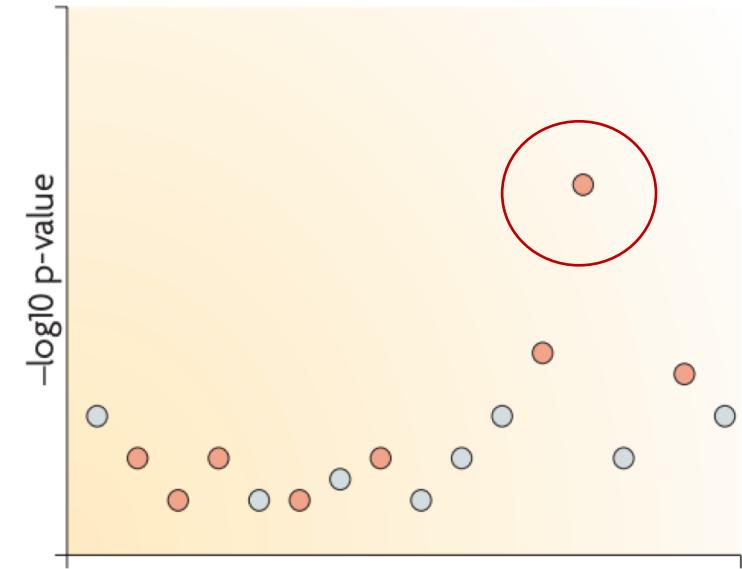
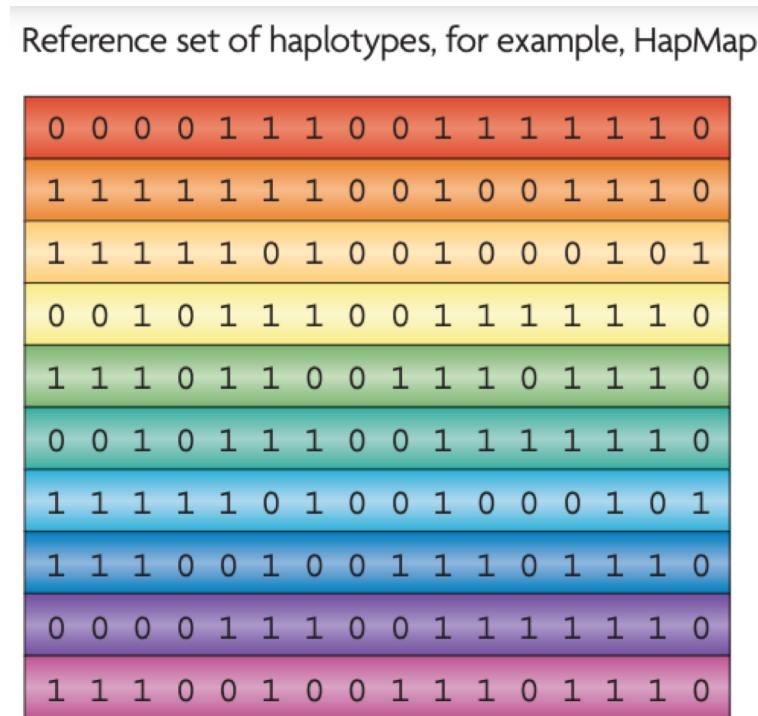
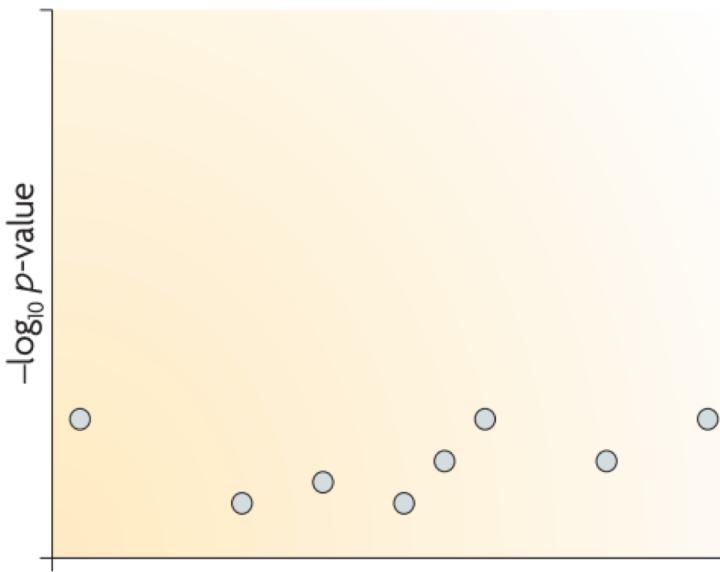
Cases and controls typed on SNP chip

240K to 5 million variants

# Genotype imputation

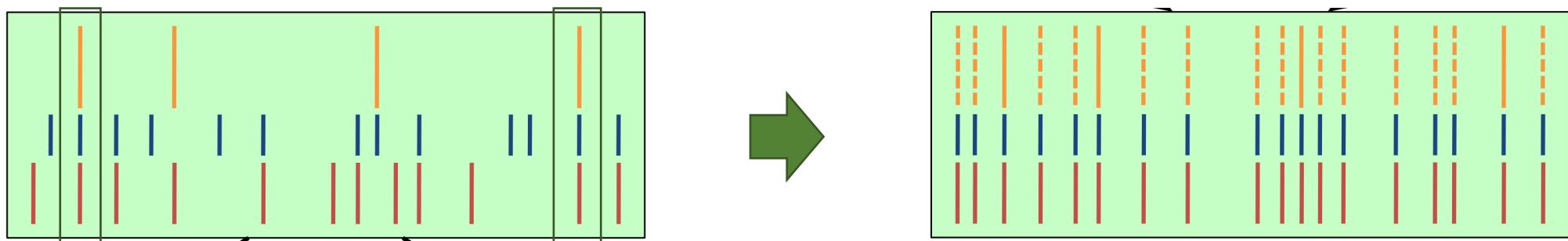
- Genotype imputation refers to the process of predicting the genotypes that are not directly assayed
- **Major purposes:**
  - To recover variants not included or not called in the SNP array
  - To finemap the causal variant

# Imputation to finemap the causal variant



# Genotype imputation

- Genotype imputation refers to the process of predicting the genotypes that are not directly assayed
- **Major purposes:**
  - To recover variants not included or not called in the SNP array
  - To finemap the causal variant
  - To increase the number of variants overlapped between studies assayed using different SNP arrays for meta-analysis



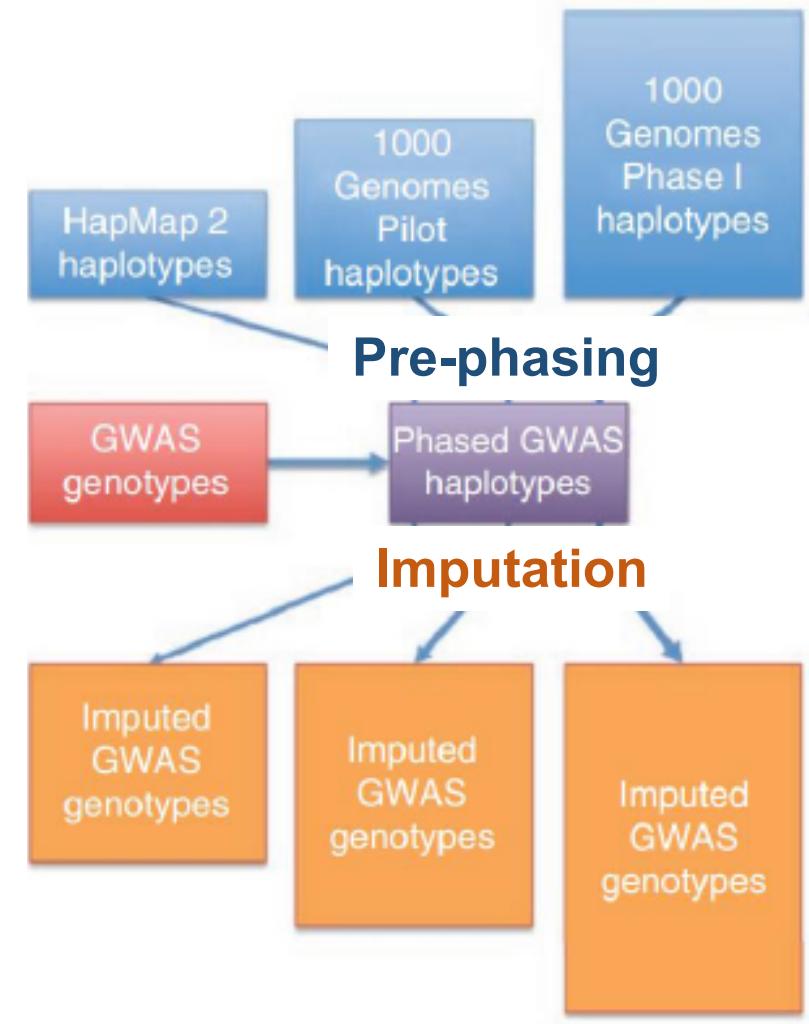
Taking **intersection** will reduce the number of overlapping variants for meta-analysis

# Genotype imputation

- Genotype imputation refers to the process of predicting the genotypes that are not directly assayed
- **Major purposes:**
  - To recover variants not included or not called in the SNP array
  - To finemap the causal variant
  - To increase the number of variants overlapped between studies assayed using different SNP arrays for meta-analysis
- **Other purposes:**
  - Imputation of non-SNP variations
  - Correction/Detection of genotyping errors
  - Missing data imputation

# Two basic steps in imputation

- **Pre-phasing** for haplotype estimation
- **Imputation** to impute the genotype probabilities for missing variants into the estimated haplotype



# Prephasing : Identify shared chromosomal regions

Heterozygous genotypes at 3 sites

**AC      TG      AT**

.	.	A	.	.	G	.	A	.
.	.	C	.	.	T	.	T	.

Three other possible haplotypes

.	.	A	.	.	T	.	T	.
.	.	C	.	.	G	.	A	.

.	.	A	.	.	T	.	A	.
.	.	C	.	.	G	.	T	.

.	.	A	.	.	G	.	T	.
.	.	C	.	.	T	.	A	.

Reference  
haplotypes

A	G	A	T	C	T	C	T	T
A	G	C	T	C	T	C	A	T
A	G	A	T	C	<b>G</b>	C	T	T
A	G	A	T	C	T	A	T	T

haplotype 1 frequency

haplotype 2 frequency

haplotype 3 frequency

haplotype 4 frequency

# Prephasing : Identify shared chromosomal regions

Heterozygous genotypes at 3 sites

AC      TG      AT

.	.	A	.	.	G	.	A	.
.	.	C	.	.	T	.	T	.

.	.	A	.	.	T	.	T	.
.	.	C	.	.	G	.	A	.

.	.	A	.	.	T	.	A	.
.	.	C	.	.	G	.	T	.

Best guess haplotypes

.	.	A	.	.	G	.	T	.
.	.	C	.	.	T	.	A	.

Reference  
haplotypes

A	G	A	T	C	T	C	T	T
A	G	C	T	C	T	C	A	T
A	G	A	T	C	G	C	T	T
A	G	A	T	C	T	A	T	T

haplotype 1 frequency

haplotype 2 frequency

haplotype 3 frequency

haplotype 4 frequency

# Prephasing : Identify shared chromosomal regions

Heterozygous genotypes at 3 sites

AC      TG      AT

Best guess haplotypes

.	A	.	G	.	T	.
.	C	.	T	.	A	.

Reference  
haplotypes

A	G	A	T	C	T	C	T	T
A	G	C	T	C	T	C	A	T
A	G	A	T	C	G	C	T	T
A	G	A	T	C	T	A	T	T

haplotype 1 frequency

haplotype 2 frequency

haplotype 3 frequency

haplotype 4 frequency

# Imputation : Fill in missing genotypes

Heterozygous genotypes at 3 sites

AC      TG      AT

Best guess haplotypes

A	G	A	T	C	G	C	T	T
A	G	C	T	C	T	C	A	T

Reference  
haplotypes

A	G	A	T	C	T	C	T	T
A	G	C	T	C	T	C	A	T
A	G	A	T	C	G	C	T	T
A	G	A	T	C	T	A	T	T

haplotype 1 frequency

haplotype 2 frequency

haplotype 3 frequency

haplotype 4 frequency

# Workflow for imputation

## Using imputation server

- Choose the imputation server and reference panels
- Prepare the genotype data in required format
- Upload the genotype data to imputation server
- Download the imputed data
- Perform association analysis

Alternatively, you can carry out the imputation locally, using tools such as IMPUTE2, BEAGLE, MACH, minimac, and SHAPEIT2

# Step 1: Choose the reference panel

- Publicly available for download
  - HapMap2 and 3
  - 1000G Phase I and 3
- Others only available for imputation servers

## Sanger Imputation Service

This is a free genotype **imputation** and **phasing** service provided by the [Wellcome Sanger Institute](#). You can upload GWAS data in VCF or 23andMe format and receive imputed and phased genomes back. Click [here](#) to learn more and [follow us on Twitter](#).

## Michigan Imputation Server

Free Next-Generation Genotype Imputation Service

Sign up now    Login

87.5M	9025	17
Imputed Genomes	Registered Users	Running Jobs

BioData **CATALYST**  
TOPMed Imputation Server

Home About Help Contact

## TOPMed Imputation Server

Free Next-Generation Genotype Imputation Service

Sign up now    Login

31.1M	2642	6
Imputed Genomes	Registered Users	Running Jobs

# Reference panels in imputation servers

The screenshot shows the TOPMed Imputation Server interface. At the top, there's a blue header bar with the server's name and a search bar labeled "Search docs". Below the header is a dark grey sidebar containing links to "Home", "Getting Started", "Data Preparation", and a collapsed section titled "Reference Panels". Under "Reference Panels", there's a list of reference panels: "TOPMed (Version R2 on GRC38)", "HRC (Version r1.1 2016)", "HRC (Version r1 2015)", "1000 Genomes Phase 3 (Version 5)", "1000 Genomes Phase 1 (Version 3)", "CAAPA - African American Panel", and "HapMap 2".

- **TOPMed (Version R2 on GRC38)**
  - 194,512 haplotypes
  - 308,107,085 sites
- **Haplotype Reference Consortium (HRC; Version r1.1 2016)**
  - 64,940 haplotypes of predominantly European ancestry
  - 39,635,008 sites
- **CAAPA - African American Panel**
  - 1766 haplotypes included in the Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA)
  - 31,163,897 sites

# Reference panels in imputation servers

TOPMed Imputation Server

Search docs

Home

Getting Started

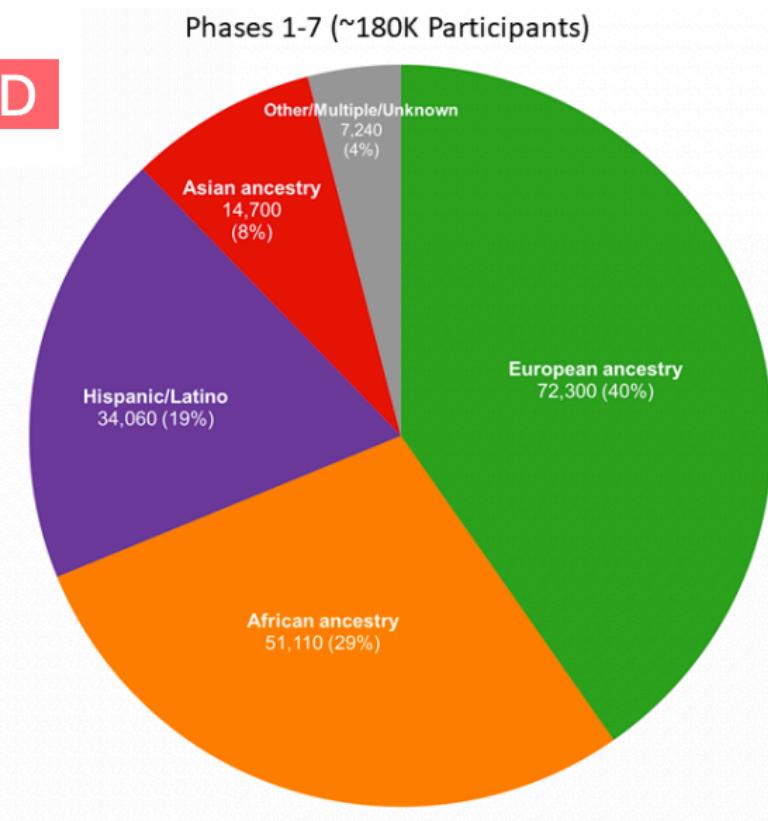
Data Preparation

Reference Panels

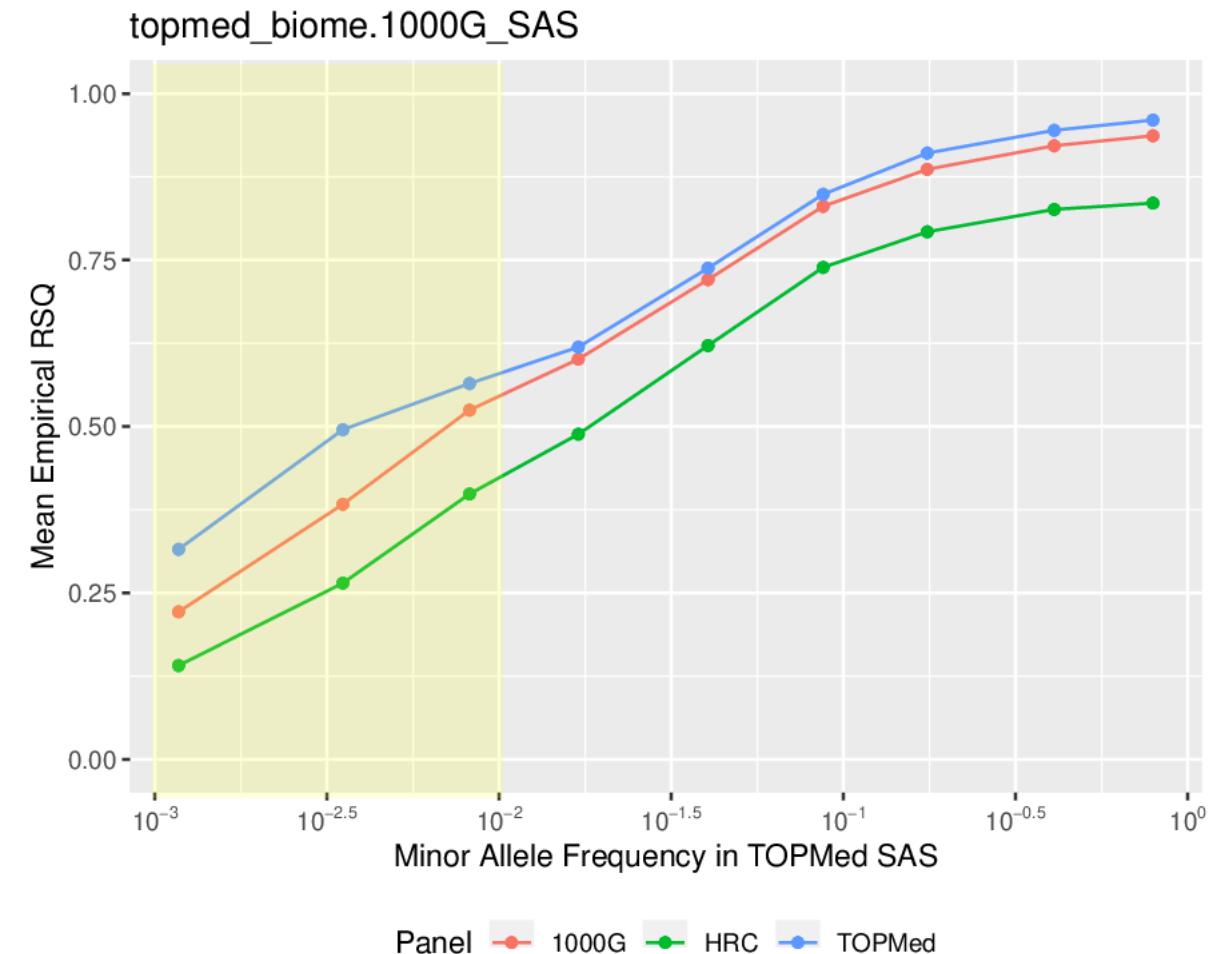
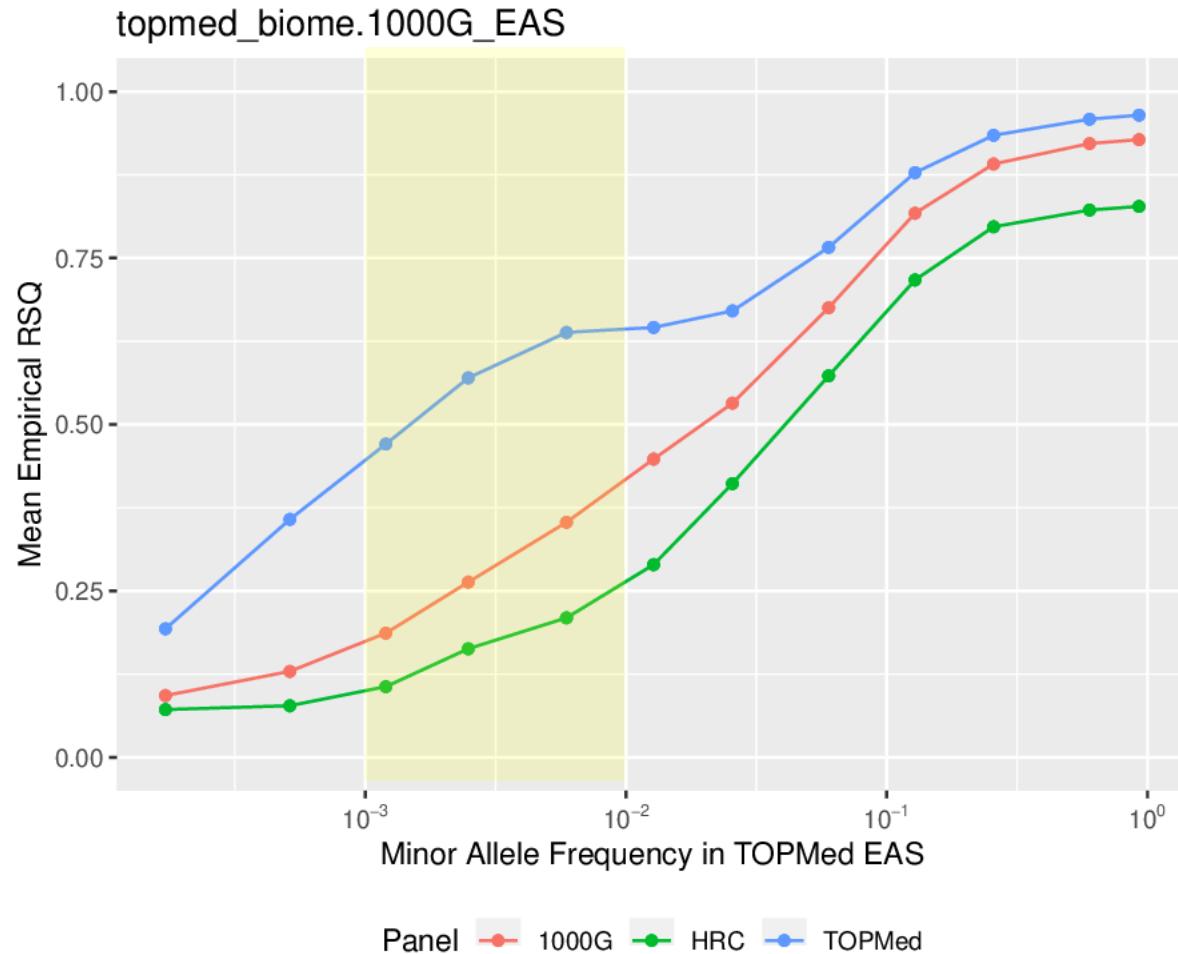
- TOPMed (Version R2 on GRC38)
- HRC (Version r1.1 2016)
- HRC (Version r1 2015)
- 1000 Genomes Phase 3 (Version 5)
- 1000 Genomes Phase 1 (Version 3)
- CAAPA - African American Panel
- HapMap 2

- **TOPMed (Version R2 on GRC38)**
  - including 2394 Asians
  - 194,512 haplotypes
  - 308,107,085 sites

T O P L D



# TOPMed imputation



# Reference panels in imputation servers

 Michigan Imputation Server

Search docs

Reference Panels

- Reference Panels
  - HRC (Version r1.1 2016)
  - 1000 Genomes Phase 3 (Version 5)
  - Genome Asia Pilot - GAsP
  - Genome Asia v2 - GAsP
  - Four-digit Multi-ethnic HLA v1 (2021)
  - Four-digit Multi-ethnic HLA v2 (2022)
  - CAAPA - African American Panel
  - HRC (Version r1 2015)
  - 1000 Genomes Phase 1 (Version 3)
  - HapMap 2

## GenomeAsia 100K Project

### Genome Asia Pilot - GAsP

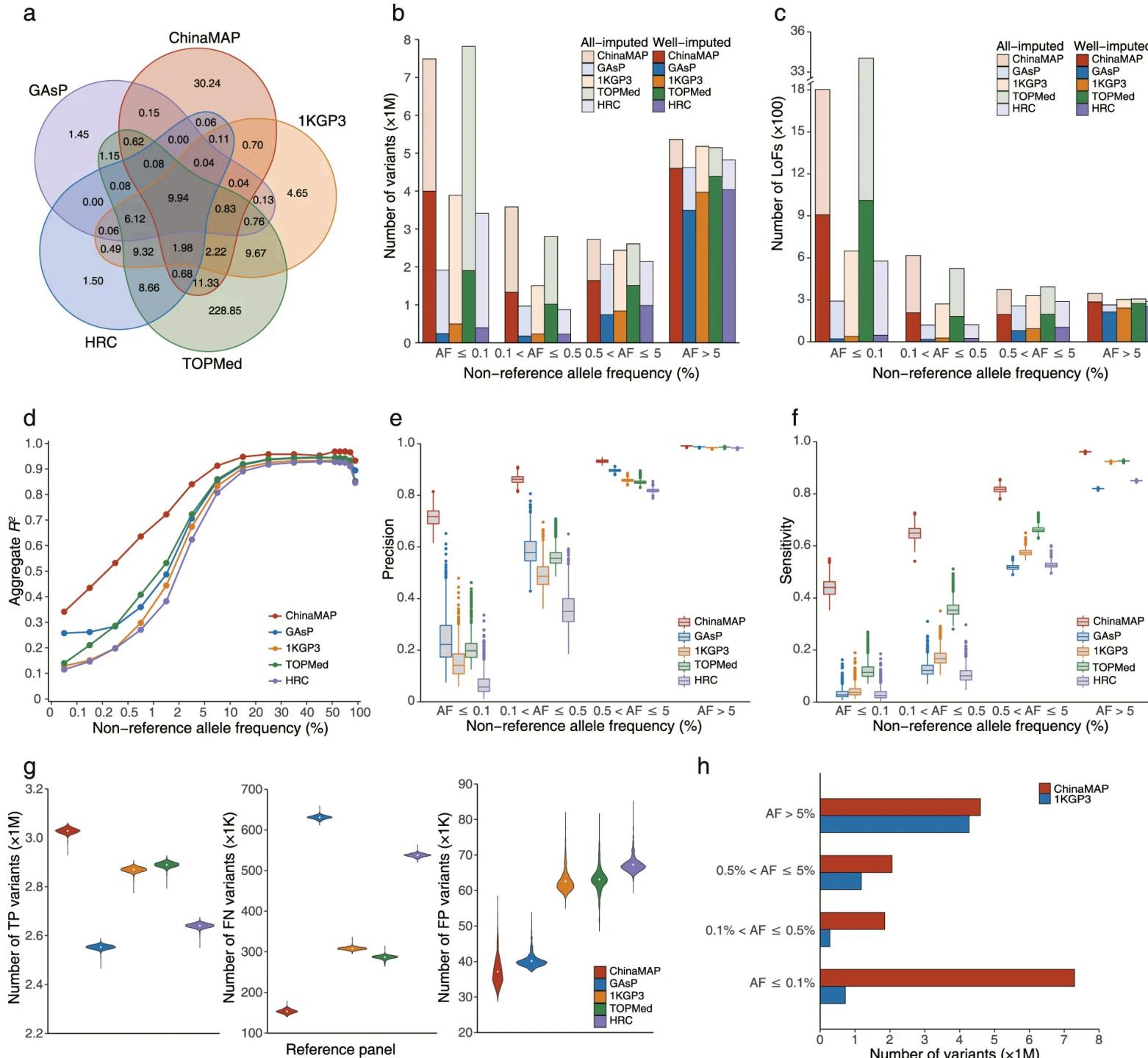
Number of Samples	1,654
Sites (chr1-22)	21,494,814
Chromosomes	1-22
Publication	<a href="https://www.nature.com/articles/s41586-019-1793-z">https://www.nature.com/articles/s41586-019-1793-z</a>

### Genome Asia v2 - GAsP

Number of Samples	6,461
Sites (chr1-22)	-
Chromosomes	1-22
Publication	<a href="https://www.nature.com/articles/s41586-019-1793-z">https://www.nature.com/articles/s41586-019-1793-z</a>

<http://mbiobank.com>

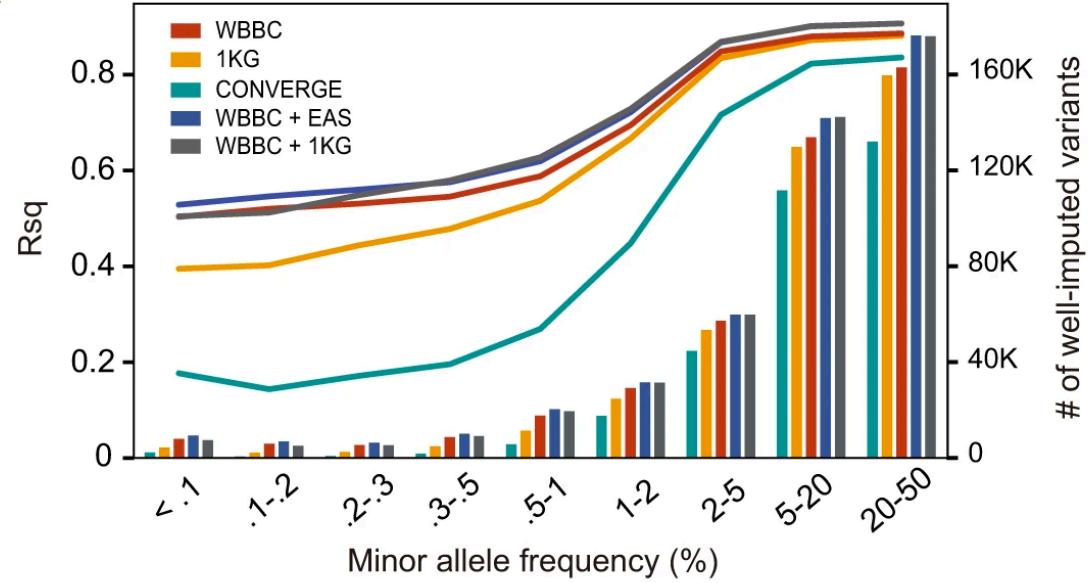
- 10,155 unrelated Chinese
- 59.01 M SNVs



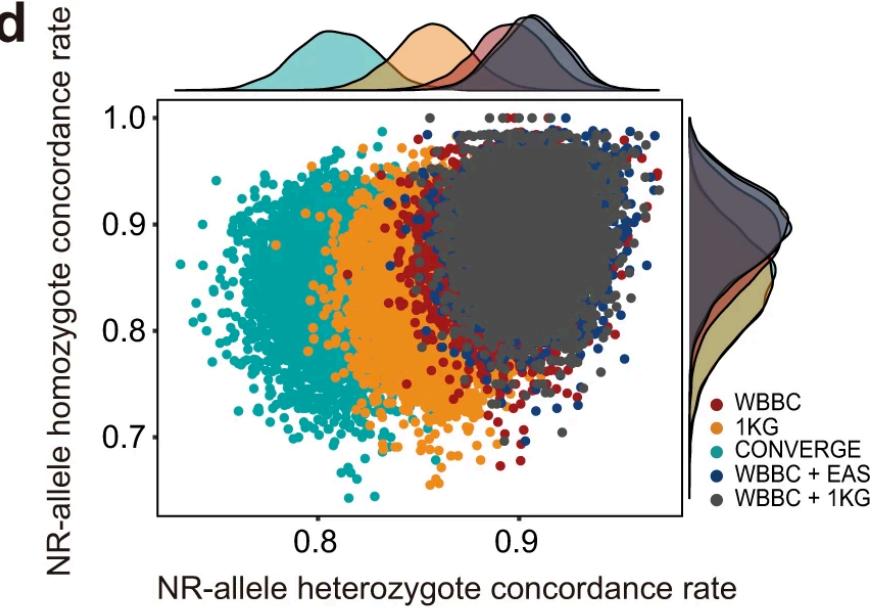
# Westlake Imputation Server

- 4,535 whole-genome sequencing Chinese samples

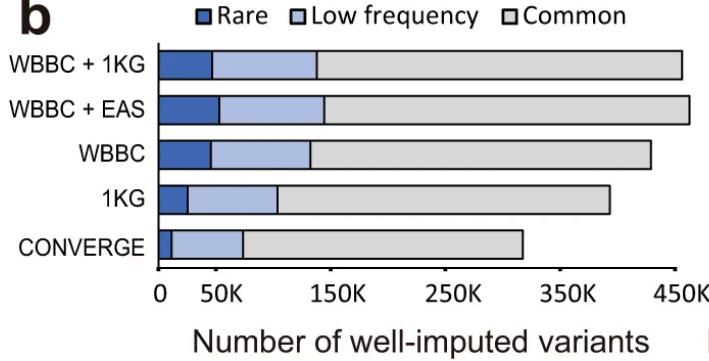
**a**



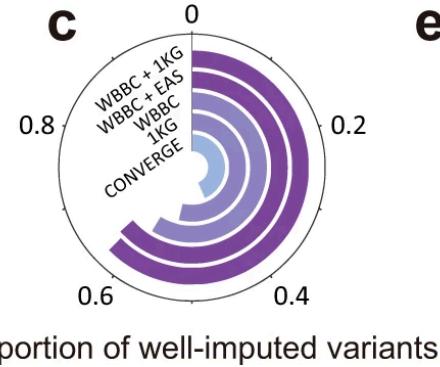
**d**



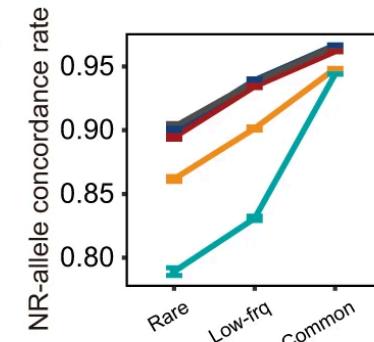
**b**



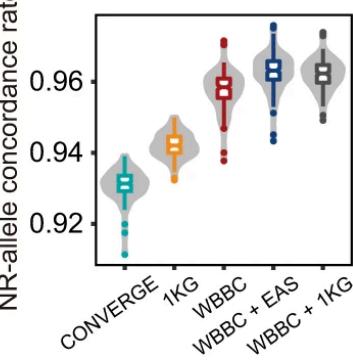
**c**



**e**



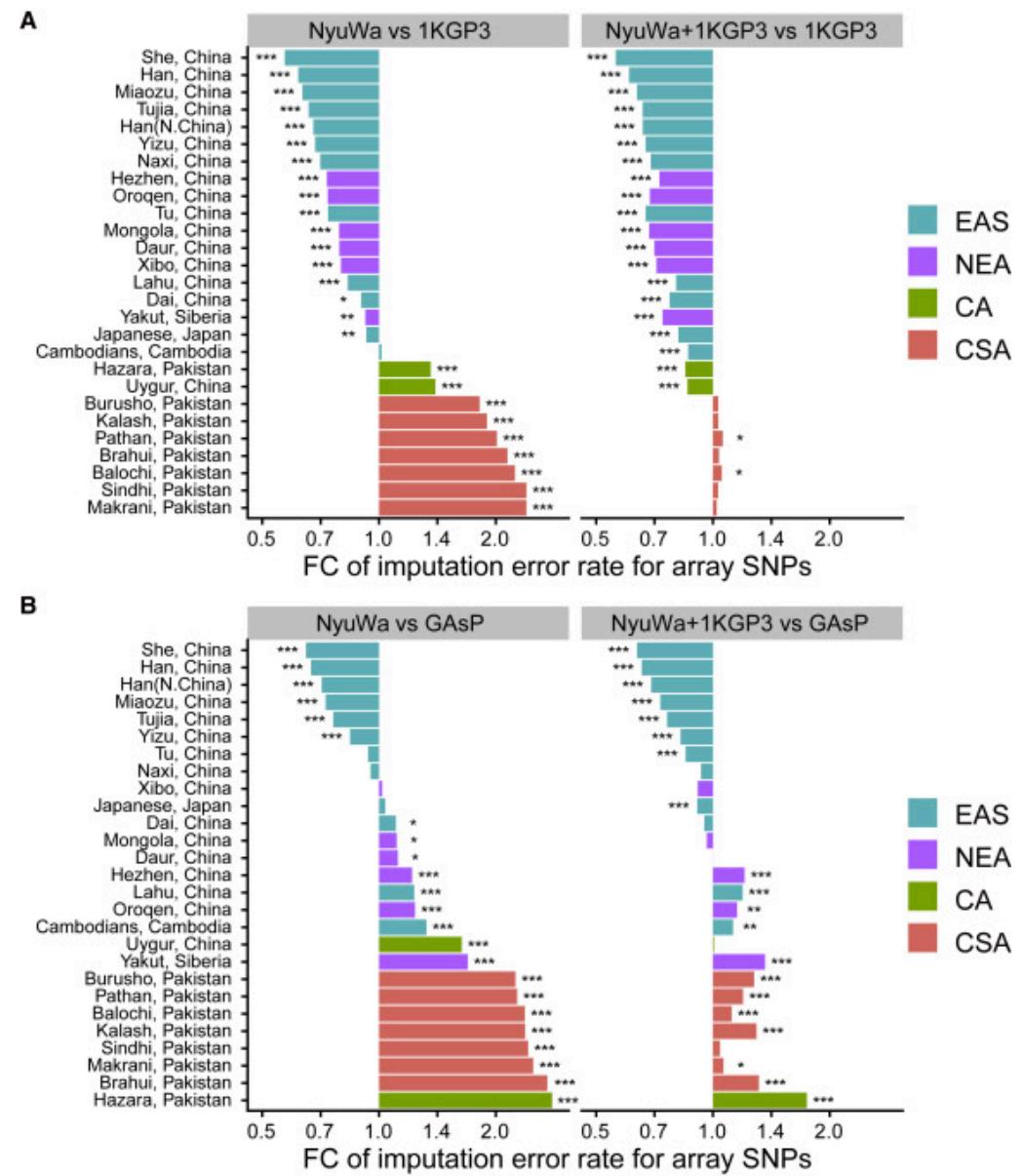
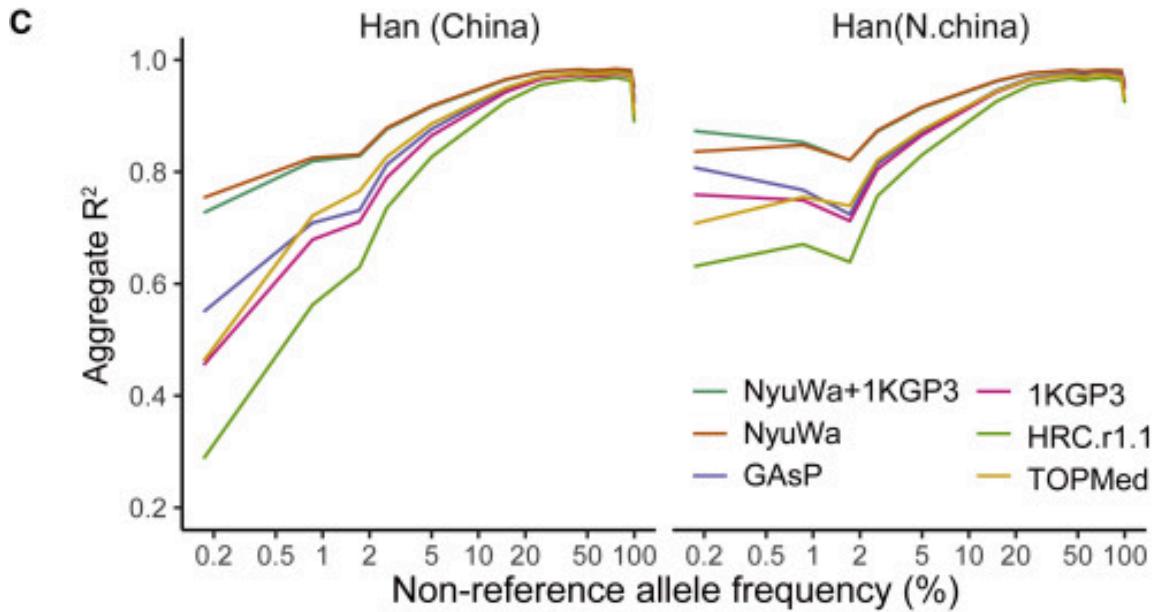
**f**



# NyuWa Server

Imputation service with Chinese population genome resource phase 1

- 2902 Chinese whole genome sequenced samples



# Meta-imputation

- Impute target samples against two or more different reference panels
- Combine the imputation results using weights guided by the empirical performance of each of the panels in stretches of each individual genome
  - weights are individual and region specific
  - weights are estimated through a hidden Markov model (HMM)
- Meta-imputed result at each marker is then a weighted average of the estimated allele counts from imputation against each panel

AJHG

Volume 109, Issue 6, 2 June 2022, Pages 1007-1015



Article

Meta-imputation: An efficient method to combine genotype data after imputation with multiple reference panels

Ketian Yu <sup>1</sup> Sayantan Das <sup>1, 2</sup>, Jonathon LeFaive <sup>1</sup>, Alan Kwong <sup>1</sup>, Jacob Pleiness <sup>1</sup>, Lukas Forer <sup>3</sup>,  
Sebastian Schönherr <sup>3</sup>, Christian Fuchsberger <sup>1, 3, 4</sup>, Albert Vernon Smith <sup>1</sup>, Gonçalo Rocha Abecasis <sup>1</sup>,

# Step 2: Data preparation

- **Variant-based QC**
  - Exclude variants with high missingness (>5%), violating HWE, or having Mendelian errors
  - Exclude variants with low MAF (<1%)
  - Drop strand ambiguous SNVs, i.e. A/T or C/G SNPs
  - Update genome build (GRCh37 vs GRCh38) using UCSC LiftOver
  - Check and update strand, alleles, position, Ref/Alt assignments and frequency differences with reference to the reference panel  
(<https://www.well.ox.ac.uk/~wrayner/tools/>)
  - Output the data in required format (e.g. separate vcf.gz file for each chromosome for Michigan imputation server)

## Step 3: Upload the data to imputation server

- If input genotype data is unphased, pre-phasing will be performed
- Michigan imputation server
  - Eagle v2.4, Beagle (for HRC and 1000G)
  - PBWT (Positional Burrows-Wheeler Transform) for genotype imputation
- TOPMed imputation server
  - Eagle v2.4
  - Minimac4
- Sanger imputation server
  - Eagle2, SHAPEIT2
  - PBWT

# Step 3: Upload the data to imputation server

The screenshot shows the BioData CATALYST TOPMed Imputation Server interface. The 'Run' tab is selected. On the left, there are input fields for 'Name' (optional job name), 'Reference Panel' (set to 'TOPMed r2'), and 'Input Files (VCF)' (File Upload). On the right, configuration options include 'Array Build' (set to 'GRCh38/hg38'), 'rsq Filter' (set to 'off'), 'Phasing' (set to 'Eagle v2.4 (phased output)'), 'Population' (set to '-- select an option --'), 'Mode' (set to 'Quality Control & Imputation'), and checkboxes for 'AES 256 encryption' and 'Generate Meta-imputation file'.

National Heart, Lung,  
and Blood Institute | BioData **CATALYST** TOPMed Imputation Server

Home Run ▾ Jobs About Help ☰ C

Run

Name optional job name

Reference Panel TOPMed r2 (Details)

Input Files (VCF) File Upload

Select Files

Array Build GRCh38/hg38

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

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 standard unzip programs. Use 7z instead.

Generate Meta-imputation file

# Step 4: Download the imputed data

- Two types of output
  - INFO file records the confidence of imputation
    - Variants imputed with Rsq / INFO score >0.3 to >0.5 are usually considered for downstream analysis

SNP	REF(0)	ALT(1)	ALT_Frq	MAF	AvgCall	Rsq	Genotyped	LooRsq	EmpR	EmpRsq	Dose0	Dose1	
chr22:16388891:G:A		G	A		0.01449	0.01449	1.00000	0.99995	Genotyped	0.993	1.000	0.99999	0.99907 0.00009
chr22:16388903:C:T		C	T		0.00000	0.00000	1.00000	0.00045	Imputed	-	-	-	-
chr22:16388907:C:A		C	A		0.00000	0.00000	1.00000	0.00003	Imputed	-	-	-	-
chr22:16388913:C:T		C	T		0.00000	0.00000	1.00000	0.00020	Imputed	-	-	-	-
chr22:16388914:C:T		C	T		0.00001	0.00001	0.99999	0.00140	Imputed	-	-	-	-
chr22:16388915:G:A		G	A		0.00009	0.00009	0.99991	0.01665	Imputed	-	-	-	-
chr22:16388931:G:A		G	A		0.00001	0.00001	0.99999	0.00605	Imputed	-	-	-	-
chr22:16388942:C:G		C	G		0.00001	0.00001	0.99999	0.00009	Imputed	-	-	-	-
chr22:16388944:G:C		G	C		0.00002	0.00002	0.99998	0.00598	Imputed	-	-	-	-

## Imputation quality evaluation

[https://genome.sph.umich.edu/wiki/Minimac:\\_Tutorial](https://genome.sph.umich.edu/wiki/Minimac:_Tutorial)

Minimac hides each of the genotyped SNPs in turn and then calculates 3 statistics:

- looRSQ - this is the estimated rsq for that SNP (as if SNP weren't typed).
- empR - this is the empirical correlation between true and imputed genotypes for the SNP. If this is negative, the SNP alleles are probably flipped.
- empRSQ - this is the actual R<sup>2</sup> value, comparing imputed and true genotypes.

These statistics can be found in the \*.info file

Be aware that, unfortunately, imputation quality statistics are not directly comparable between different imputation programs (MaCH/minimac vs. Impute vs. Beagle etc.).

## Step 4: Download the imputed data

- Two types of output
  - INFO file records the confidence of imputation
  - Genotype output
- Three main genotype output formats
  - Best guess genotypes
  - Dosage data
  - Genotype probability data

```

##fileformat=VCFv4.1
###FILTER=<ID=PASS,Description="All filters passed">
###pipeline=michigan-imputationserver-1.3.3
###imputation=minimac4-1.0.2
###phasing=eagle-2.4
###r2Filter=0.0
###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">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT sample1
chr1 80346 chr1:80346:C:G C G . PASS AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED
chr1 722408 chr1:722408:G:C G C . PASS AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED
chr1 727717 chr1:727717:G:C G C . PASS AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED
chr1 732994 chr1:732994:G:A G A . PASS AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED
chr1 769828 chr1:769828:G:C G C . PASS AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED
chr1 770988 chr1:770988:A:G A G . PASS AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED
chr1 771265 chr1:771265:A:C A C . PASS AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED
chr1 778597 chr1:778597:C:T C T . PASS AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED
chr1 779059 chr1:779059:G:A G A . PASS AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED
chr1 783006 chr1:783006:A:G A G . PASS AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED
chr1 784860 chr1:784860:T:C T C . PASS AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED
chr1 785417 chr1:785417:G:A G A . PASS AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED
chr1 792275 chr1:792275:C:T C T . PASS AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED
chr1 792779 chr1:792779:C:T C T . PASS AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED
chr1 792862 chr1:792862:C:G C G . PASS AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED
chr1 805887 chr1:805887:T:C T C . PASS AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED
chr1 806199 chr1:806199:T:C T C . PASS AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED
chr1 819894 chr1:819894:C:T C T . PASS AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED
chr1 821723 chr1:821723:C:T C T . PASS AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED
chr1 821740 chr1:821740:A:G A G . PASS AF=0.12484;MAF=0.12484;R2=0.96666;IMPUTED
chr1 822260 chr1:822260:G:A G A . PASS AF=0.64565;MAF=0.35435;R2=0.56299;IMPUTED
chr1 822354 chr1:822354:C:T C T . PASS AF=0.78694;MAF=0.21306;R2=0.40978;IMPUTED
chr1 824457 chr1:824457:T:A T A . PASS AF=0.79042;MAF=0.20958;R2=0.40835;IMPUTED
chr1 825042 chr1:825042:C:T C T . PASS AF=0.12887;MAF=0.12887;R2=0.99988;ER2=0.99954;TYPED
chr1 826940 chr1:826940:C:T C T . PASS AF=0.12869;MAF=0.12869;R2=0.99917;ER2=0.9952;TYPED
chr1 827476 chr1:827476:G:A G A . PASS AF=0.1291;MAF=0.1291;R2=0.99939;ER2=0.99176;TYPED
chr1 829648 chr1:829648:C:T C T . PASS AF=0.12425;MAF=0.12425;R2=0.96595;IMPUTED
chr1 830826 chr1:830826:G:T G T . PASS AF=0.12511;MAF=0.12511;R2=0.9665;IMPUTED
chr1 832873 chr1:832873:A:C A C . PASS AF=0.53547;MAF=0.46453;R2=0.48568;IMPUTED

```

**GT:** Best guess genotype

**DS:** Estimated alternative allele dosage

**GP:** Posterior genotype probabilities

GT:DS:HDS:GP	0 0:0.174:0.143,0.031:0.83,0.165,0.004
GT:DS:HDS:GP	0 1:0.991:0.105,0.885:0.102,0.804,0.093
GT:DS:HDS:GP	0 1:1.05:0.157,0.893:0.091,0.769,0.14
GT:DS:HDS:GP	0 0:0.02:0.02,0.001:0.98,0.02,0
GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0
GT:DS:HDS:GP	0 1:0.946:0.103,0.843:0.141,0.772,0.087
GT:DS:HDS:GP	1 0:0.734:0.733,0.001:0.267,0.733,0
GT:DS:HDS:GP	1 0:0.734:0.733,0:0.267,0.733,0
GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0
GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142
GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142
GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142
GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0
GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142
GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142
GT:DS:HDS:GP	1 0:0.784:0.783,0:0.216,0.783,0
GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142
GT:DS:HDS:GP	0 0:0:0,0:1,0,0
GT:DS:HDS:GP	0 0:0:0,0:1,0,0
GT:DS:HDS:GP	0 0:0:0,0:1,0,0
GT:DS:HDS:GP	1 1:1.784:0.891,0.893:0.012,0.193,0.796
GT:DS:HDS:GP	1 1:1.782:0.891,0.892:0.012,0.194,0.794
GT:DS:HDS:GP	1 1:1.784:0.891,0.893:0.012,0.192,0.796
GT:DS:HDS:GP	0 0:0:0,0:1,0,0
GT:DS:HDS:GP	1 1:1.746:0.868,0.879:0.016,0.221,0.762

```

##fileformat=VCFv4.1
###FILTER=<ID=PASS,Description="All filters passed">
###pipeline=michigan-imputationserver-1.3.3
###imputation=minimac4-1.0.2
###phasing=eagle-2.4
###r2Filter=0.0
###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">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT sample1
chr1 80346 chr1:80346:C:G C G . PASS AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED GT:DS:HDS:GP 0|0:0.174:0.143,0.031:0.83,0.165,0.004
chr1 722408 chr1:722408:G:C G C . PASS AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED GT:DS:HDS:GP 0|1:0.991:0.105,0.885:0.102,0.804,0.093
chr1 727717 chr1:727717:G:C G C . PASS AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED GT:DS:HDS:GP 0|1:1.05:0.157,0.893:0.091,0.769,0.14
chr1 732994 chr1:732994:G:A G A . PASS AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED GT:DS:HDS:GP 0|0:0.02:0.02,0.001:0.98,0.02,0
chr1 769828 chr1:769828:G:C G C . PASS AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 770988 chr1:770988:A:G A G . PASS AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED GT:DS:HDS:GP 0|1:0.946:0.103,0.843:0.141,0.772,0.087
chr1 771265 chr1:771265:A:C A C . PASS AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0.001:0.267,0.733,0
chr1 778597 chr1:778597:C:T C T . PASS AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0:0.267,0.733,0
chr1 779059 chr1:779059:G:A G A . PASS AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 783006 chr1:783006:A:G A G . PASS AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 784860 chr1:784860:T:C T C . PASS AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 785417 chr1:785417:G:A G A . PASS AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792275 chr1:792275:C:T C T . PASS AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 792779 chr1:792779:C:T C T . PASS AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792862 chr1:792862:C:G C G . PASS AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 805887 chr1:805887:T:C T C . PASS AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED GT:DS:HDS:GP 1|0:0.784:0.783,0:0.216,0.783,0
chr1 806199 chr1:806199:T:C T C . PASS AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 819894 chr1:819894:C:T C T . PASS AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821723 chr1:821723:C:T C T . PASS AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821740 chr1:821740:A:G A G . PASS AF=0.12484;MAF=0.12484;R2=0.96666;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 822260 chr1:822260:G:A G A . PASS AF=0.64565;MAF=0.35435;R2=0.56299;IMPUTED GT:DS:HDS:GP 1|1:1.784:0.891,0.893:0.012,0.193,0.796
chr1 822354 chr1:822354:C:T C T . PASS AF=0.78694;MAF=0.21306;R2=0.40978;IMPUTED GT:DS:HDS:GP 1|1:1.782:0.891,0.892:0.012,0.194,0.794
chr1 824457 chr1:824457:T:A T A . PASS AF=0.79042;MAF=0.20958;R2=0.40835;IMPUTED GT:DS:HDS:GP 1|1:1.784:0.891,0.893:0.012,0.192,0.796
chr1 825042 chr1:825042:C:T C T . PASS AF=0.12887;MAF=0.12887;R2=0.99988;ER2=0.99954;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 826940 chr1:826940:C:T C T . PASS AF=0.12869;MAF=0.12869;R2=0.99917;ER2=0.9952;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 827476 chr1:827476:G:A G A . PASS AF=0.1291;MAF=0.1291;R2=0.99939;ER2=0.99176;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 829648 chr1:829648:C:T C T . PASS AF=0.12425;MAF=0.12425;R2=0.96595;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 830826 chr1:830826:G:T G T . PASS AF=0.12511;MAF=0.12511;R2=0.9665;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 832873 chr1:832873:A:C A C . PASS AF=0.53547;MAF=0.46453;R2=0.48568;IMPUTED GT:DS:HDS:GP 1|1:1.746:0.868,0.879:0.016,0.221,0.762

```

**GT: Best guess genotype**

0|0

0|1

1|1

0 0:0.174:0.143,0.031:0.83,0.165,0.004
0 1:0.991:0.105,0.885:0.102,0.804,0.093
0 1:1.05:0.157,0.893:0.091,0.769,0.14
0 0:0.02:0.02,0.001:0.98,0.02,0
1 0:0.733:0.733,0:0.267,0.733,0
0 1:0.946:0.103,0.843:0.141,0.772,0.087
1 0:0.734:0.733,0.001:0.267,0.733,0
1 0:0.734:0.733,0:0.267,0.733,0
1 0:0.733:0.733,0:0.267,0.733,0
0 1:1.055:0.159,0.896:0.087,0.77,0.142
0 1:1.055:0.159,0.896:0.087,0.77,0.142
0 1:1.055:0.159,0.896:0.087,0.77,0.142
1 0:0.733:0.733,0:0.267,0.733,0
0 1:1.055:0.159,0.896:0.087,0.77,0.142
0 1:1.055:0.159,0.896:0.087,0.77,0.142
1 0:0.784:0.783,0:0.216,0.783,0
0 1:1.055:0.159,0.896:0.087,0.77,0.142
0 0:0:0,0:1,0,0
0 0:0:0,0:1,0,0
0 0:0:0,0:1,0,0
1 1:1.784:0.891,0.893:0.012,0.193,0.796
1 1:1.782:0.891,0.892:0.012,0.194,0.794
1 1:1.784:0.891,0.893:0.012,0.192,0.796
0 0:0:0,0:1,0,0
0 0:0:0,0:1,0,0
0 0:0:0,0:1,0,0
1 1:1.784:0.891,0.893:0.016,0.221,0.762

```

##fileformat=VCFv4.1
###FILTER=<ID=PASS,Description="All filters passed">
###pipeline=michigan-imputationserver-1.3.3
###imputation=minimac4-1.0.2
###phasing=eagle-2.4
###r2Filter=0.0
###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">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT sample1
chr1 80346 chr1:80346:C:G C G . PASS AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED GT:DS:HDS:GP 0|0:0.174:0.143,0.031:0.83,0.165,0.004
chr1 722408 chr1:722408:G:C G C . PASS AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED GT:DS:HDS:GP 0|1:0.991:0.105,0.885:0.102,0.804,0.093
chr1 727717 chr1:727717:G:C G C . PASS AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED GT:DS:HDS:GP 0|1:1.05:0.157,0.893:0.091,0.769,0.14
chr1 732994 chr1:732994:G:A G A . PASS AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED GT:DS:HDS:GP 0|0:0.02:0.02,0.001:0.98,0.02,0
chr1 769828 chr1:769828:G:C G C . PASS AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 770988 chr1:770988:A:G A G . PASS AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED GT:DS:HDS:GP 0|1:0.946:0.103,0.843:0.141,0.772,0.087
chr1 771265 chr1:771265:A:C A C . PASS AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0.001:0.267,0.733,0
chr1 778597 chr1:778597:C:T C T . PASS AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0:0.267,0.733,0
chr1 779059 chr1:779059:G:A G A . PASS AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 783006 chr1:783006:A:G A G . PASS AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 784860 chr1:784860:T:C T C . PASS AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 785417 chr1:785417:G:A G A . PASS AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792275 chr1:792275:C:T C T . PASS AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 792779 chr1:792779:C:T C T . PASS AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792862 chr1:792862:C:G C G . PASS AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 805887 chr1:805887:T:C T C . PASS AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED GT:DS:HDS:GP 1|0:0.784:0.783,0:0.216,0.783,0
chr1 806199 chr1:806199:T:C T C . PASS AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 819894 chr1:819894:C:T C T . PASS AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821723 chr1:821723:C:T C T . PASS AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821740 chr1:821740:A:G A G . PASS AF=0.12484;MAF=0.12484;R2=0.96666;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 822260 chr1:822260:G:A G A . PASS AF=0.64565;MAF=0.35435;R2=0.56299;IMPUTED GT:DS:HDS:GP 1|1:1.784:0.891,0.893:0.012,0.193,0.796
chr1 822354 chr1:822354:C:T C T . PASS AF=0.78694;MAF=0.21306;R2=0.40978;IMPUTED GT:DS:HDS:GP 1|1:1.782:0.891,0.892:0.012,0.194,0.794
chr1 824457 chr1:824457:T:A T A . PASS AF=0.79042;MAF=0.20958;R2=0.40835;IMPUTED GT:DS:HDS:GP 1|1:1.784:0.891,0.893:0.012,0.192,0.796
chr1 825042 chr1:825042:C:T C T . PASS AF=0.12887;MAF=0.12887;R2=0.99988;ER2=0.99954;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 826940 chr1:826940:C:T C T . PASS AF=0.12869;MAF=0.12869;R2=0.99917;ER2=0.9952;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 827476 chr1:827476:G:A G A . PASS AF=0.1291;MAF=0.1291;R2=0.99939;ER2=0.99176;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 829648 chr1:829648:C:T C T . PASS AF=0.12425;MAF=0.12425;R2=0.96595;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 830826 chr1:830826:G:T G T . PASS AF=0.12511;MAF=0.12511;R2=0.9665;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 832873 chr1:832873:A:C A C . PASS AF=0.53547;MAF=0.46453;R2=0.48568;IMPUTED GT:DS:HDS:GP 1|1:1.746:0.868,0.879:0.016,0.221,0.762

```

**DS: Estimated alternative allele dosage**  
 $= 0*P(0|0) + 1*P(0|1) + 2*P(1|1)$   
 $\Rightarrow$  from 0 to 2

```

##fileformat=VCFv4.1
###FILTER=<ID=PASS,Description="All filters passed">
###pipeline=michigan-imputationserver-1.3.3
###imputation=minimac4-1.0.2
###phasing=eagle-2.4
###r2Filter=0.0
###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">

```

CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	sample1	GT:DS:HDS:GP	0 0:0.174:0.143,0.031:0.83,0.165,0.004
chr1	80346	chr1:80346:C:G	C	G	.	PASS	AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED	GT:DS:HDS:GP	0 0:0.174:0.143,0.031:0.83,0.165,0.004		
chr1	722408	chr1:722408:G:C	G	C	.	PASS	AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED	GT:DS:HDS:GP	0 1:0.991:0.105,0.885:0.102,0.804,0.093		
chr1	727717	chr1:727717:G:C	G	C	.	PASS	AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED	GT:DS:HDS:GP	0 1:1.05:0.157,0.893:0.091,0.769,0.14		
chr1	732994	chr1:732994:G:A	G	A	.	PASS	AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED	GT:DS:HDS:GP	0 0:0.02:0.02,0.001:0.98,0.02,0		
chr1	769828	chr1:769828:G:C	G	C	.	PASS	AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED	GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0		
chr1	770988	chr1:770988:A:G	A	G	.	PASS	AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED	GT:DS:HDS:GP	0 1:0.946:0.103,0.843:0.141,0.772,0.087		
chr1	771265	chr1:771265:A:C	A	C	.	PASS	AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED	GT:DS:HDS:GP	1 0:0.734:0.733,0.001:0.267,0.733,0		
chr1	778597	chr1:778597:C:T	C	T	.	PASS	AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED	GT:DS:HDS:GP	1 0:0.734:0.733,0:0.267,0.733,0		
chr1	779059	chr1:779059:G:A	G	A	.	PASS	AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED	GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0		
chr1	783006	chr1:783006:A:G	A	G	.	PASS	AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	784860	chr1:784860:T:C	T	C	.	PASS	AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	785417	chr1:785417:G:A	G	A	.	PASS	AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	792275	chr1:792275:C:T	C	T	.	PASS	AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED	GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0		
chr1	792779	chr1:792779:C:T	C	T	.	PASS	AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	792862	chr1:792862:C:G	C	G	.	PASS	AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	805887	chr1:805887:T:C	T	C	.	PASS	AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED	GT:DS:HDS:GP	1 0:0.784:0.783,0:0.216,0.783,0		
chr1	806199	chr1:806199:T:C	T	C	.	PASS	AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	819894	chr1:819894:C:T	C	T	.	PASS	AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	821723	chr1:821723:C:T	C	T	.	PASS	AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	821740	chr1:821740:A:G	A	G	.	PASS	AF=0.12484;MAF=0.12484;R2=0.96666;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	822260	chr1:822260:G:A	G	A	.	PASS	AF=0.64565;MAF=0.35435;R2=0.56299;IMPUTED	GT:DS:HDS:GP	1 1:1.784:0.891,0.893:0.012,0.193,0.796		
chr1	822354	chr1:822354:C:T	C	T	.	PASS	AF=0.78694;MAF=0.21306;R2=0.40978;IMPUTED	GT:DS:HDS:GP	1 1:1.782:0.891,0.892:0.012,0.194,0.794		
chr1	824457	chr1:824457:T:A	T	A	.	PASS	AF=0.79042;MAF=0.20958;R2=0.40835;IMPUTED	GT:DS:HDS:GP	1 1:1.784:0.891,0.893:0.012,0.192,0.796		
chr1	825042	chr1:825042:C:T	C	T	.	PASS	AF=0.12887;MAF=0.12887;R2=0.99988;ER2=0.99954;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	826940	chr1:826940:C:T	C	T	.	PASS	AF=0.12869;MAF=0.12869;R2=0.99917;ER2=0.9952;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	827476	chr1:827476:G:A	G	A	.	PASS	AF=0.1291;MAF=0.1291;R2=0.99939;ER2=0.99176;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	829648	chr1:829648:C:T	C	T	.	PASS	AF=0.12425;MAF=0.12425;R2=0.96595;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	830826	chr1:830826:G:T	G	T	.	PASS	AF=0.12511;MAF=0.12511;R2=0.9665;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	832873	chr1:832873:A:C	A	C	.	PASS	AF=0.53547;MAF=0.46453;R2=0.48568;IMPUTED	GT:DS:HDS:GP	1 1:1.746:0.868,0.879:0.016,0.221,0.762		

**GP: Posterior genotype probabilities**  
 $P(0|0)$ ,  $P(0|1)$ ,  $P(1|1)$



100

100

100

```

##fileformat=VCFv4.1
###FILTER=<ID=PASS,Description="All filters passed">
###pipeline=michigan-imputationserver-1.3.3
###imputation=minimac4-1.0.2
###phasing=eagle-2.4
###r2Filter=0
###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">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT sample1
chr1 80346 chr1:80346:C:G C G . PASS AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED GT:DS:HDS:GP 0|0:0.174:0.143,0.031:0.83,0.165,0.004
chr1 722408 chr1:722408:G:C G C . PASS AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED GT:DS:HDS:GP 0|1:0.991:0.105,0.885:0.102,0.804,0.093
chr1 727717 chr1:727717:G:C G C . PASS AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED GT:DS:HDS:GP 0|1:1.05:0.157,0.893:0.091,0.769,0.14
chr1 732994 chr1:732994:G:A G A . PASS AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED GT:DS:HDS:GP 0|0:0.02:0.02,0.001:0.98,0.02,0
chr1 769828 chr1:769828:G:C G C . PASS AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 770988 chr1:770988:A:G A G . PASS AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED GT:DS:HDS:GP 0|1:0.946:0.103,0.843:0.141,0.772,0.087
chr1 771265 chr1:771265:A:C A C . PASS AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0.001:0.267,0.733,0
chr1 778597 chr1:778597:C:T C T . PASS AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0:0.267,0.733,0
chr1 779059 chr1:779059:G:A G A . PASS AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 783006 chr1:783006:A:G A G . PASS AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 784860 chr1:784860:T:C T C . PASS AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 785417 chr1:785417:G:A G A . PASS AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792275 chr1:792275:C:T C T . PASS AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 792779 chr1:792779:C:T C T . PASS AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792862 chr1:792862:C:G C G . PASS AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 805887 chr1:805887:T:C T C . PASS AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED GT:DS:HDS:GP 1|0:0.784:0.783,0:0.216,0.783,0
chr1 806199 chr1:806199:T:C T C . PASS AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 819894 chr1:819894:C:T C T . PASS AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821723 chr1:821723:C:T C T . PASS AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0

```

**chr1:821723:C:T**

- Genotype: 0|0 = C|C
- Dosage = 0
- Genotype probabilities = 1,0,0

CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	sample1	GT:DS:HDS:GP	0 0:0.174:0.143,0.031:0.83,0.165,0.004
chr1	80346	chr1:80346:C:G	C	G	.	PASS	AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED	GT:DS:HDS:GP	0 0:0.174:0.143,0.031:0.83,0.165,0.004		
chr1	722408	chr1:722408:G:C	G	C	.	PASS	AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED	GT:DS:HDS:GP	0 1:0.991:0.105,0.885:0.102,0.804,0.093		
chr1	727717	chr1:727717:G:C	G	C	.	PASS	AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED	GT:DS:HDS:GP	0 1:1.05:0.157,0.893:0.091,0.769,0.14		
chr1	732994	chr1:732994:G:A	G	A	.	PASS	AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED	GT:DS:HDS:GP	0 0:0.02:0.02,0.001:0.98,0.02,0		
chr1	769828	chr1:769828:G:C	G	C	.	PASS	AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED	GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0		
chr1	770988	chr1:770988:A:G	A	G	.	PASS	AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED	GT:DS:HDS:GP	0 1:0.946:0.103,0.843:0.141,0.772,0.087		
chr1	771265	chr1:771265:A:C	A	C	.	PASS	AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED	GT:DS:HDS:GP	1 0:0.734:0.733,0.001:0.267,0.733,0		
chr1	778597	chr1:778597:C:T	C	T	.	PASS	AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED	GT:DS:HDS:GP	1 0:0.734:0.733,0:0.267,0.733,0		
chr1	779059	chr1:779059:G:A	G	A	.	PASS	AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED	GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0		
chr1	783006	chr1:783006:A:G	A	G	.	PASS	AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	784860	chr1:784860:T:C	T	C	.	PASS	AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	785417	chr1:785417:G:A	G	A	.	PASS	AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	792275	chr1:792275:C:T	C	T	.	PASS	AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED	GT:DS:HDS:GP	1 0:0.733:0.733,0:0.267,0.733,0		
chr1	792779	chr1:792779:C:T	C	T	.	PASS	AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	792862	chr1:792862:C:G	C	G	.	PASS	AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	805887	chr1:805887:T:C	T	C	.	PASS	AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED	GT:DS:HDS:GP	1 0:0.784:0.783,0:0.216,0.783,0		
chr1	806199	chr1:806199:T:C	T	C	.	PASS	AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED	GT:DS:HDS:GP	0 1:1.055:0.159,0.896:0.087,0.77,0.142		
chr1	819894	chr1:819894:C:T	C	T	.	PASS	AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	821723	chr1:821723:C:T	C	T	.	PASS	AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	821740	chr1:821740:A:G	A	G	.	PASS	AF=0.12484;MAF=0.12484;R2=0.96666;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	822260	chr1:822260:G:A	G	A	.	PASS	AF=0.64565;MAF=0.35435;R2=0.56299;IMPUTED	GT:DS:HDS:GP	1 1:1.784:0.891,0.893:0.012,0.193,0.796		
chr1	822354	chr1:822354:C:T	C	T	.	PASS	AF=0.78694;MAF=0.21306;R2=0.40978;IMPUTED	GT:DS:HDS:GP	1 1:1.782:0.891,0.892:0.012,0.194,0.794		
chr1	824457	chr1:824457:T:A	T	A	.	PASS	AF=0.79042;MAF=0.20958;R2=0.40835;IMPUTED	GT:DS:HDS:GP	1 1:1.784:0.891,0.893:0.012,0.192,0.796		
chr1	825042	chr1:825042:C:T	C	T	.	PASS	AF=0.12887;MAF=0.12887;R2=0.99988;ER2=0.99954;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	826940	chr1:826940:C:T	C	T	.	PASS	AF=0.12869;MAF=0.12869;R2=0.99917;ER2=0.9952;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	827476	chr1:827476:G:A	G	A	.	PASS	AF=0.1291;MAF=0.1291;R2=0.99939;ER2=0.99176;TYPED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	829648	chr1:829648:C:T	C	T	.	PASS	AF=0.12425;MAF=0.12425;R2=0.96595;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	830826	chr1:830826:G:T	G	T	.	PASS	AF=0.12511;MAF=0.12511;R2=0.9665;IMPUTED	GT:DS:HDS:GP	0 0:0:0,0:1,0,0		
chr1	832873	chr1:832873:A:C	A	C	.	PASS	AF=0.53547;MAF=0.46453;R2=0.48568;IMPUTED	GT:DS:HDS:GP	1 1:1.746:0.868,0.879:0.016,0.221,0.762		

```

##fileformat=VCFv4.1
###FILTER=<ID=PASS,Description="All filters passed">
###pipeline=michigan-imputationserver-1.3.3
###imputation=minimac4-1.0.2
###phasing=eagle-2.4
###r2Filter=0.0
###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">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT sample1
chr1 80346 chr1:80346:C:G C G . PASS AF=0.20725;MAF=0.20725;R2=0.30525;IMPUTED GT:DS:HDS:GP 0|0:0.174:0.143,0.031:0.83,0.165,0.004
chr1 722408 chr1:722408:G:C G C . PASS AF=0.69034;MAF=0.30966;R2=0.42713;IMPUTED GT:DS:HDS:GP 0|1:0.991:0.105,0.885:0.102,0.804,0.093
chr1 727717 chr1:727717:G:C G C . PASS AF=0.70029;MAF=0.29971;R2=0.43755;IMPUTED GT:DS:HDS:GP 0|1:1.05:0.157,0.893:0.091,0.769,0.14
chr1 732994 chr1:732994:G:A G A . PASS AF=0.16886;MAF=0.16886;R2=0.59382;IMPUTED GT:DS:HDS:GP 0|0:0.02:0.02,0.001:0.98,0.02,0
chr1 769828 chr1:769828:G:C G C . PASS AF=0.1473;MAF=0.1473;R2=0.8098;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 770988 chr1:770988:A:G A G . PASS AF=0.32021;MAF=0.32021;R2=0.43854;IMPUTED GT:DS:HDS:GP 0|1:0.946:0.103,0.843:0.141,0.772,0.087
chr1 771265 chr1:771265:A:C A C . PASS AF=0.1508;MAF=0.1508;R2=0.78526;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0.001:0.267,0.733,0
chr1 778597 chr1:778597:C:T C T . PASS AF=0.16702;MAF=0.16702;R2=0.81547;IMPUTED GT:DS:HDS:GP 1|0:0.734:0.733,0:0.267,0.733,0
chr1 779059 chr1:779059:G:A G A . PASS AF=0.14286;MAF=0.14286;R2=0.82221;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 783006 chr1:783006:A:G A G . PASS AF=0.73865;MAF=0.26135;R2=0.42769;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 784860 chr1:784860:T:C T C . PASS AF=0.73629;MAF=0.26371;R2=0.43809;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 785417 chr1:785417:G:A G A . PASS AF=0.7356;MAF=0.2644;R2=0.43953;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792275 chr1:792275:C:T C T . PASS AF=0.14537;MAF=0.14537;R2=0.81409;IMPUTED GT:DS:HDS:GP 1|0:0.733:0.733,0:0.267,0.733,0
chr1 792779 chr1:792779:C:T C T . PASS AF=0.73524;MAF=0.26476;R2=0.43906;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 792862 chr1:792862:C:G C G . PASS AF=0.7399;MAF=0.2601;R2=0.42392;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 805887 chr1:805887:T:C T C . PASS AF=0.14527;MAF=0.14527;R2=0.8338;IMPUTED GT:DS:HDS:GP 1|0:0.784:0.783,0:0.216,0.783,0
chr1 806199 chr1:806199:T:C T C . PASS AF=0.73578;MAF=0.26422;R2=0.44644;IMPUTED GT:DS:HDS:GP 0|1:1.055:0.159,0.896:0.087,0.77,0.142
chr1 819894 chr1:819894:C:T C T . PASS AF=0.12886;MAF=0.12886;R2=0.99993;ER2=0.99269;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821723 chr1:821723:C:T C T . PASS AF=0.12485;MAF=0.12485;R2=0.96663;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 821740 chr1:821740:A:G A G . PASS AF=0.12484;MAF=0.12484;R2=0.96666;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0


---


chr1 822260 chr1:822260:G:A G A . PASS AF=0.64565;MAF=0.35435;R2=0.56299;IMPUTED GT:DS:HDS:GP 1|1:1.784:0.891,0.893:0.012,0.193,0.796
chr1 822354 chr1:822354:C:T C T . PASS AF=0.78694;MAF=0.21306;R2=0.40978;IMPUTED GT:DS:HDS:GP 1|1:1.782:0.891,0.892:0.012,0.194,0.794
chr1 824457 chr1:824457:T:A T A . PASS AF=0.79042;MAF=0.20958;R2=0.40835;IMPUTED GT:DS:HDS:GP 1|1:1.784:0.891,0.893:0.012,0.192,0.796
chr1 825042 chr1:825042:C:T C T . PASS AF=0.12887;MAF=0.12887;R2=0.99988;ER2=0.99954;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 826940 chr1:826940:C:T C T . PASS AF=0.12869;MAF=0.12869;R2=0.99917;ER2=0.9952;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 827476 chr1:827476:G:A G A . PASS AF=0.1291;MAF=0.1291;R2=0.99939;ER2=0.99176;TYPED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 829648 chr1:829648:C:T C T . PASS AF=0.12425;MAF=0.12425;R2=0.96595;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 830826 chr1:830826:G:T G T . PASS AF=0.12511;MAF=0.12511;R2=0.9665;IMPUTED GT:DS:HDS:GP 0|0:0:0,0:1,0,0
chr1 832873 chr1:832873:A:C A C . PASS AF=0.53547;MAF=0.46453;R2=0.48568;IMPUTED GT:DS:HDS:GP 1|1:1.746:0.868,0.879:0.016,0.221,0.762

```

## chr1:822260:G:A

- Genotype: 1|1 = A|A
- Dosage = 1.784
- Genotype probabilities = 0.012, 0.193, **0.796**

# Step 5: Perform association analysis

- Post imputation QC
  - Check distribution of Rsq/INFO score
  - Plot imputed MAF against reference MAF
  - Decide the Rsq/INFO score cutoff
- Association analysis
  - PLINK : dosage
  - SNPTEST : genotype probabilities from IMPUTE
  - SAIGE
  - Bolt-LMM
  - BGENIE
- Compare association result before and after imputation

# QUESTIONS?

