

Disclosure Slide

Financial Disclosure for:

Christian Fuchsberger

Sebastian Schönherr

Lukas Forer

Cassie Spracklen

Albert Smith

Sarah Hanks

We have nothing to disclose

Section 3

Tracking runs and downloading data



Lukas Forer

Medical University of Innsbruck

lukas.forer@i-med.ac.at

@lukfor



MICHIGAN
IMPUTATION SERVER

Learning objectives

Participants will

1. Learn how to monitor jobs and their progress
2. Learn how to download imputation results
3. Learn how to decrypt results using standard software

Life Cycle of an Imputation Job



- Job passed Quality Control
- Job scheduled in imputation queue

job-20201009-043422-854

Fri Oct 09 2020 10:34:22 14 sec admin Genotype Imputation (Minimac4) 1.4.1

Job is in queue on position 2.

- Waits until resources are available

Life Cycle of an Imputation Job



- Phasing and Imputation starts



- Waiting
- Running
- Complete

Life Cycle of an Imputation Job



- Data is encrypted
- Email with one time password is sent to user

Dear Lukas,
the password for the imputation results is: pp09Z0KeQvQMc

The results can be downloaded from <https://imputationserver.sph.umich.edu/start.html#!jobs/job-20190919-112230-581/results>

Life Cycle of an Imputation Job



- After 7 days the job is retired
- All results are deleted
- We will send you an email 2 days before

Dear Lukas Forer,
Your job retires in 2 days! All imputation results will be deleted at that time.

Please ensure that you have downloaded all results from
<https://imputationserver.sph.umich.edu/start.html#!jobs/job-20191011-124306-370>

How to download
the imputed genotypes?

Option 1: Web-Interface

A screenshot of a web-based application interface. At the top, there are three tabs: "Details", "Results", and "Logs". The "Results" tab is highlighted and has a red circle with the number "1" above it. Below the tabs, the title "Imputation Results" is displayed. A list of files is shown, each with a download icon and a file name followed by its size. A red circle with the number "2" is placed over the list of files.

| File | Size |
|------------|----------|
| chr_1.zip | (893 MB) |
| chr_10.zip | (617 MB) |
| chr_11.zip | (587 MB) |
| chr_12.zip | (576 MB) |
| chr_13.zip | (467 MB) |
| chr_14.zip | (390 MB) |

A screenshot of a download progress dialog box. It shows a file named "chr_1.zip" with a download link below it. A progress bar indicates the download speed is 5.3 MB/s and 69.2 MB of 893 MB have been downloaded, with 3 mins left. Two buttons, "Pause" and "Cancel", are at the bottom. A red circle with the number "3" is placed above the download link.

Option 1: Web-Interface

- + Easiest way to download data
- Requires manual interaction

Option 2: Batch Download

- + Most flexible way to download data
- + No browser needed
- Requires basic knowledge of the command line

Option 2: Batch Download

Imputation Results 

-  chr_1.zip (469 MB)
-  chr_10.zip (287 MB)
-  chr_11.zip (281 MB)
-  chr_12.zip (269 MB)
-  chr_13.zip (195 MB)
-  chr_14.zip (192 MB)
-  chr_15.zip (181 MB)
-  chr_16.zip (203 MB)
-  chr_17.zip (187 MB)
-  chr_18.zip (160 MB)
-  chr_19.zip (170 MB)
-  chr_2.zip (471 MB)
-  chr_20.zip (129 MB)

Download data 

wget (22) [URLs \(22\)](#)

```
 wget https://imputationserver.sph.umich.edu/share/results/1fc3d1b4
 wget https://imputationserver.sph.umich.edu/share/results/3d9f5f0c
 wget https://imputationserver.sph.umich.edu/share/results/528e411a
 wget https://imputationserver.sph.umich.edu/share/results/ed598ab4
 wget https://imputationserver.sph.umich.edu/share/results/7c818b4d
 wget https://imputationserver.sph.umich.edu/share/results/1c1e651a
```

Use the following command to download all results at once:

```
curl -sL https://imputationserver.sph.umich.edu/get/1584555, 
```



**ASHG**
American Society of Human Genetics

fantasia:~> █

```
fantasia:~> curl -sL https://imputationserver.sph.umich.edu/get/1584555/675233d69  
db57b793589b916f2a81cb8 | bash
```

```
fantasia:~> curl -sL https://imputationserver.sph.umich.edu/get/1584555/675233d69  
db57b793589b916f2a81cb8 | bash
```

Downloading file chr_1.zip (1/22)...

| % Total | % Received | % Xferd | Average Speed | Time | Time | Time | Current | |
|---------|------------|---------|---------------|--------|-------|-------|---------|----------------------------------|
| | | | Dload | Upload | Total | Spent | Left | Speed |
| 100 | 185 | 100 | 185 | 0 | 0 | 21087 | 0 | --:--:-- --:--:-- --:--:-- 30833 |
| 100 | 469M | 100 | 469M | 0 | 0 | 116M | 0 | 0:00:04 0:00:04 --:--:-- 167M |

Downloading file chr_10.zip (2/22)...

| % Total | % Received | % Xferd | Average Speed | Time | Time | Time | Current | |
|---------|------------|---------|---------------|--------|-------|-------|---------|----------------------------------|
| | | | Dload | Upload | Total | Spent | Left | Speed |
| 100 | 185 | 100 | 185 | 0 | 0 | 23886 | 0 | --:--:-- --:--:-- --:--:-- 37000 |
| 100 | 287M | 100 | 287M | 0 | 0 | 87.4M | 0 | 0:00:03 0:00:03 --:--:-- 138M |

Downloading file chr_11.zip (3/22)...

Downloading file chr_7.zip (20/22)...

| % Total | % Received | % Xferd | Average Speed | Time | Time | Time | Current |
|---------|------------|---------|---------------|-------|-------|-------|---------|
| | | | Dload Upload | Total | Spent | Left | Speed |
| 100 | 185 | 100 | 185 | 0 | 0 | 24189 | 0 |
| 100 | 337M | 100 | 337M | 0 | 0 | 101M | 0 |

Downloading file chr_8.zip (21/22)...

| % Total | % Received | % Xferd | Average Speed | Time | Time | Time | Current |
|---------|------------|---------|---------------|-------|-------|-------|---------|
| | | | Dload Upload | Total | Spent | Left | Speed |
| 100 | 185 | 100 | 185 | 0 | 0 | 24529 | 0 |
| 100 | 306M | 100 | 306M | 0 | 0 | 94.9M | 0 |

Downloading file chr_9.zip (22/22)...

| % Total | % Received | % Xferd | Average Speed | Time | Time | Time | Current |
|---------|------------|---------|---------------|-------|-------|-------|---------|
| | | | Dload Upload | Total | Spent | Left | Speed |
| 100 | 185 | 100 | 185 | 0 | 0 | 22486 | 0 |
| 100 | 245M | 100 | 245M | 0 | 0 | 82.0M | 0 |

All 22 file(s) downloaded.

fantasia:~> █

Data Decryption

- All imputed genotypes are in **encrypted zip files** (e.g. chr_1.zip)
- We send you an email with a password

Dear Lukas,
the password for the imputation results is: pp09Z0KeQvQMc

The results can be downloaded from <https://imputationserver.sph.umich.edu/start.html#!jobs/job-20190919-112230-581/results>

- You need this password to **decrypt** your genotypes
- Decryption with standard zip programs (e.g. WinZip, 7zip or gunzip)
- AES Encryption: Needs additional software to decrypt (e.g. 7z)

What is in each zip file?

```
chr_20.zip
├── chr20.dose.vcf.gz
└── chr20.info.gz
```

What is in each zip file?

chr_20.zip

```
|__ chr20.dose.vcf.gz  
|__ chr20.info.gz
```

| #CHROM | POS | ID | REF | ALT | QUAL | FILTER | INFO | |
|--------|--------|---------------|-----|-----|------|--------|------------------------|--------------------|
| 20 | 61795 | 20:61795:G:T | G | T | . | PASS | AF=0.26318;MAF=0.26318 | R2=0.54658;IMPUTED |
| 20 | 63231 | 20:63231:T:G | T | G | . | PASS | AF=0.03843;MAF=0.03843 | R2=0.67736;IMPUTED |
| 20 | 63244 | 20:63244:A:C | A | C | . | PASS | AF=0.16132;MAF=0.16132 | R2=0.49907;IMPUTED |
| 20 | 68749 | 20:68749:T:C | T | C | . | PASS | AF=0.59894;MAF=0.40106 | R2=0.98392;TYPED |
| 20 | 161502 | 20:161502:C:T | C | T | . | PASS | AF=0.05882;MAF=0.05882 | TYPED_ONLY |

What is in each zip file?

chr_20.zip

```
|__ chr20.dose.vcf.gz  
└__ chr20.info.gz
```

| #CHROM | POS | ID | REF | ALT | ... |
|--------|--------|---------------|-----|-----|-----|
| 20 | 61795 | 20:61795:G:T | G | T | ... |
| 20 | 63231 | 20:63231:T:G | T | G | ... |
| 20 | 63244 | 20:63244:A:C | A | C | ... |
| 20 | 68749 | 20:68749:T:C | T | C | ... |
| 20 | 161502 | 20:161502:C:T | C | T | ... |

| FORMAT | Sample1 |
|----------|-----------------------------|
| GT:DS:GP | 1 0:1.126:0.100,0.673,0.226 |
| GT:DS:GP | 0 0:0.002:0.998,0.002,0.000 |
| GT:DS:GP | 0 0:0.285:0.723,0.270,0.008 |
| GT:DS:GP | 1 1:1.999:0.000,0.001,0.999 |
| GT:DS:GP | 0 0:0:1,0,0 |

What is in each zip file?

chr_20.zip

```
|__ chr20.dose.vcf.gz  
└__ chr20.info.gz
```

| SNP» | REF(0)» | ALT(1)» | ALT_Frq» | MAF» | AvgCall» | Rsq» | Genotyped» | ... | ... |
|---------------|---------|---------|----------|----------|----------|----------|------------|-----|-----|
| 20:61795:G:T» | G» | T» | 0.26318» | 0.26318» | 0.88455» | 0.54658» | Imputed» | ... | ... |
| 20:63231:T:G» | T» | G» | 0.03843» | 0.03843» | 0.98342» | 0.67736» | Imputed» | ... | ... |
| 20:63244:A:C» | A» | C» | 0.16132» | 0.16132» | 0.91761» | 0.49907» | Imputed» | ... | ... |

Summary

- Monitor imputation process online
- Different options to download data
- One zip file for each chromosome
- Data decryption using one-time password

More info and FAQ can be found here:
<https://imputationserver.readthedocs.io>