

# simulation\_summary

August 16, 2021

## 1 Simulation 1

### 1.1 Code

```
ts = msprime.simulate(sample_size=10000, Ne=5e4,recombination_rate=1.1e-8, mutation_rate=2e-8,
```

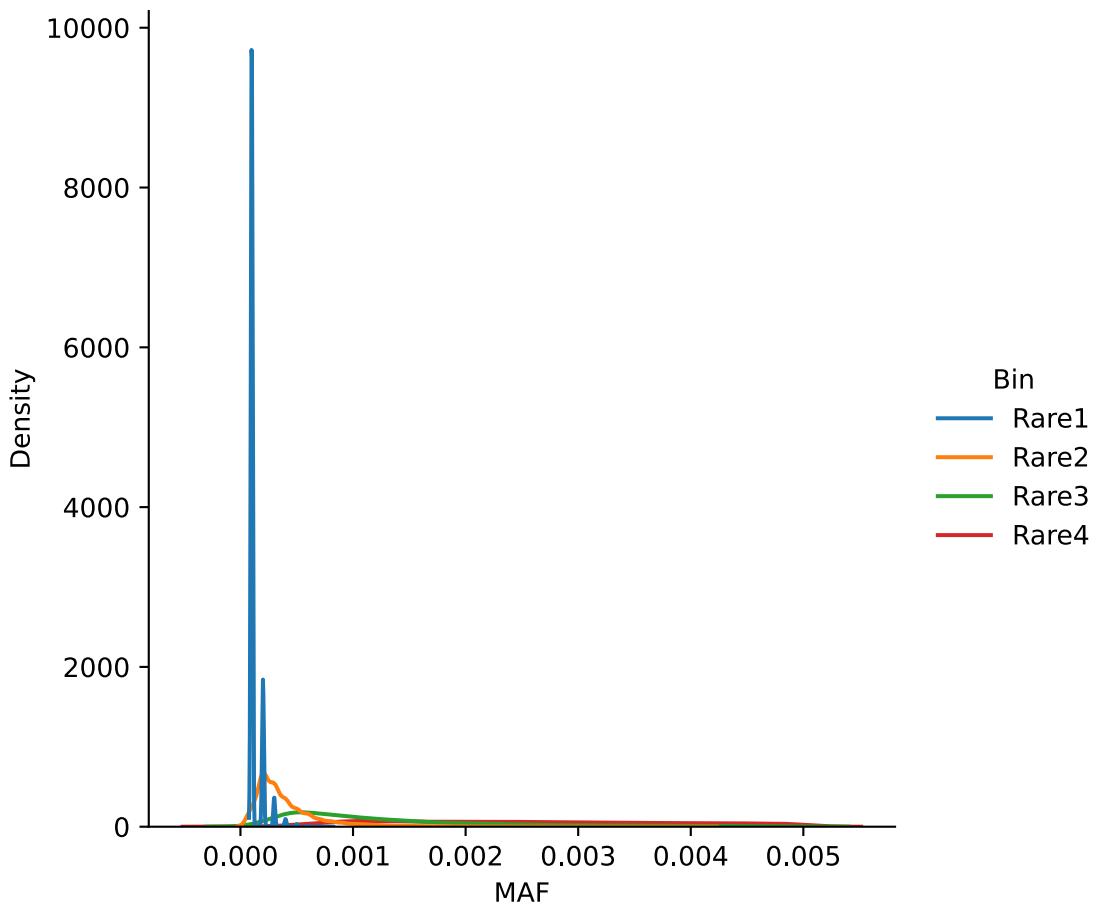
### 1.2 Input Summary

- Sample size = 10,000
- Effective population size,  $N_e = 5 \times 10^4$
- Recombination rate =  $1.1 \times 10^{-8}$
- Mutation rate =  $2 \times 10^{-8}$
- Chromosome length  $5 \times 10^6$

```
<tskit.trees.TreeSequence at 0x7fd467722e10>
```

### 1.3 Density plot of MAF of rare bins

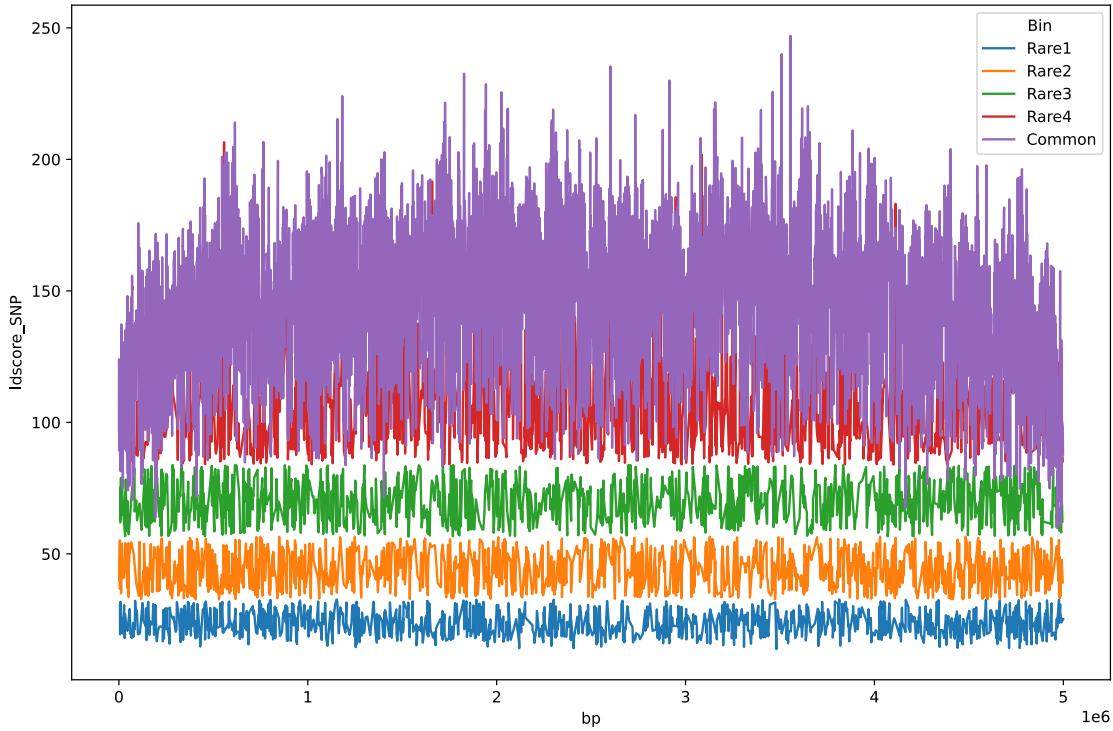
```
<seaborn.axisgrid.FacetGrid at 0x7fd45fe78e10>
```



#### 1.4 LD over the genome for variants in different Bins

The plot uses only a subset size 10,000 of all the variants.

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd45fe76f90>
```



<pandas.io.formats.style.Styler at 0x7fd46261c610>

## 2 Simulation 2

### 2.1 Code

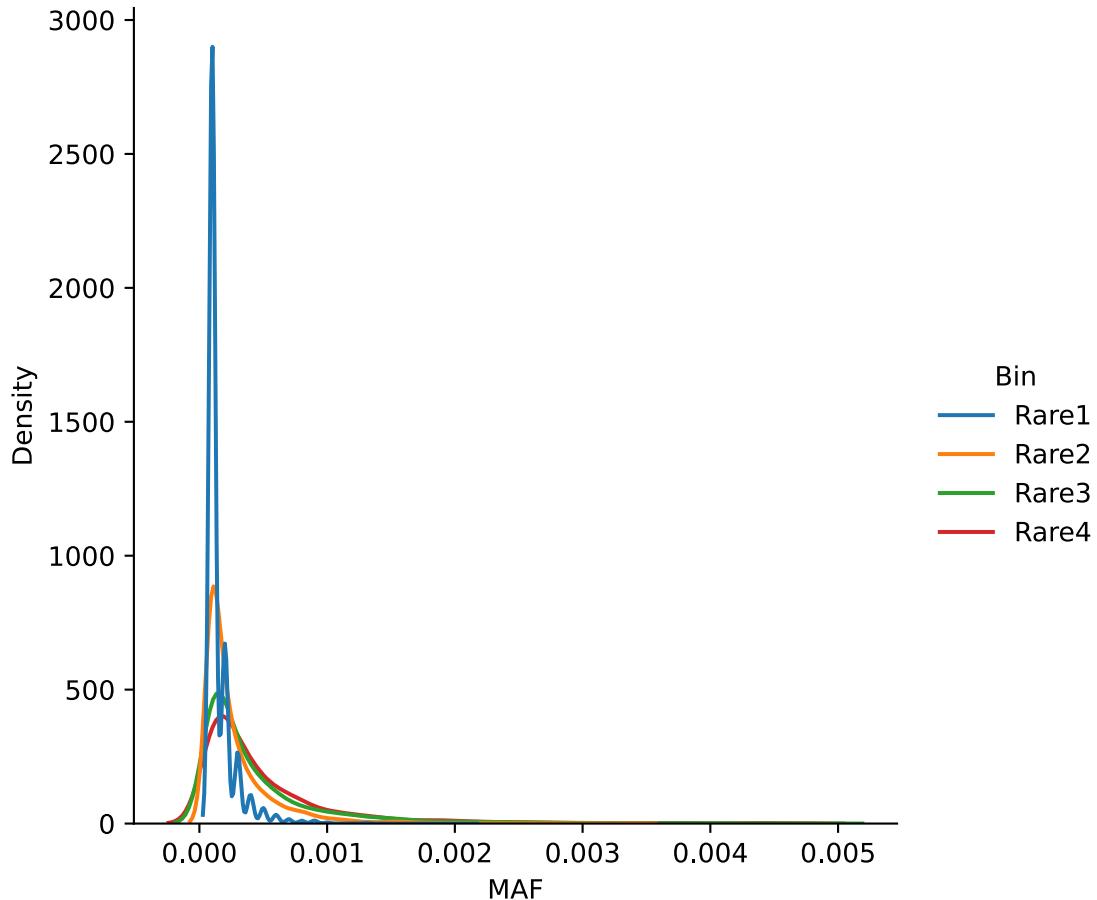
```
ts = msprime.simulate(
    population_configurations = [
        msprime.PopulationConfiguration(sample_size=10000, growth_rate=0, initial_size=100000)
    ],
    demographic_events = [
        msprime.PopulationParametersChange(time=200, growth_rate=0.05)
    ],
    length = 1e6, random_seed = 101,recombination_rate=1.1e-8, mutation_rate=2e-8
)
```

### 2.2 Input Summary

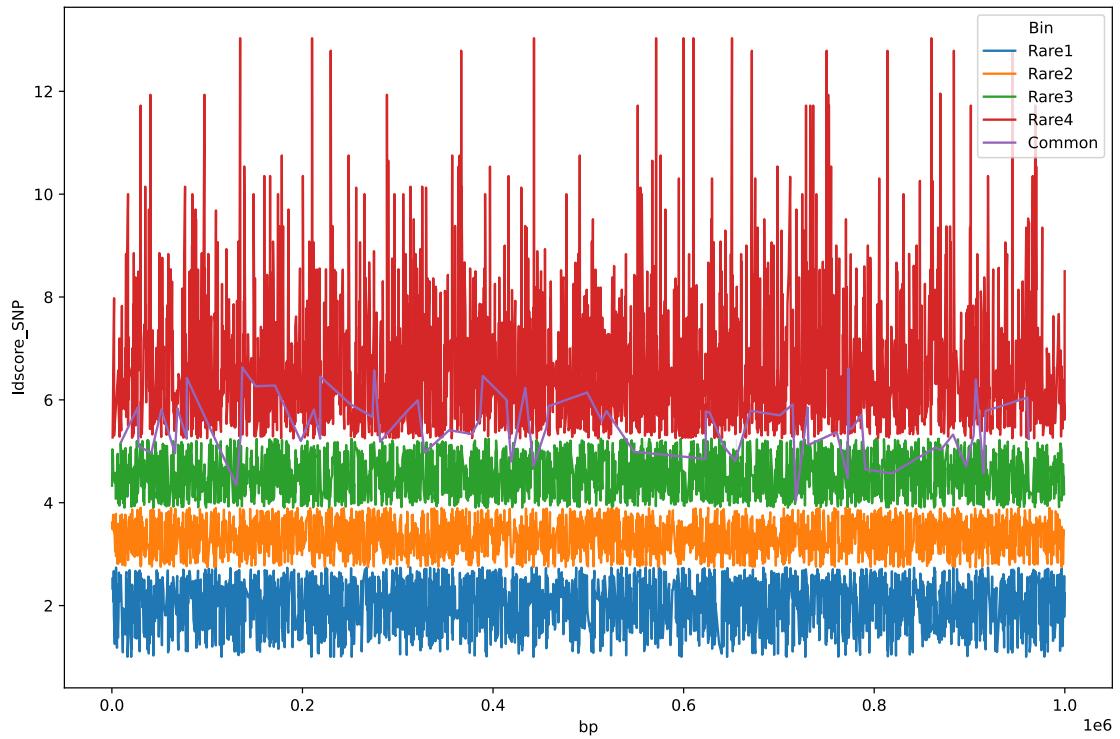
- Sample size = 10,000
- Initial population with constant growth rate and last 200 generation with exponential growth with rate 0.05
- Recombination rate =  $1.1 \times 10^{-8}$
- Mutation rate =  $2 \times 10^{-8}$
- Chromosome length  $5 \times 10^6$

```
<tskit.trees.TreeSequence at 0x7fd45fe74c50>
```

```
<seaborn.axisgrid.FacetGrid at 0x7fd46267d190>
```

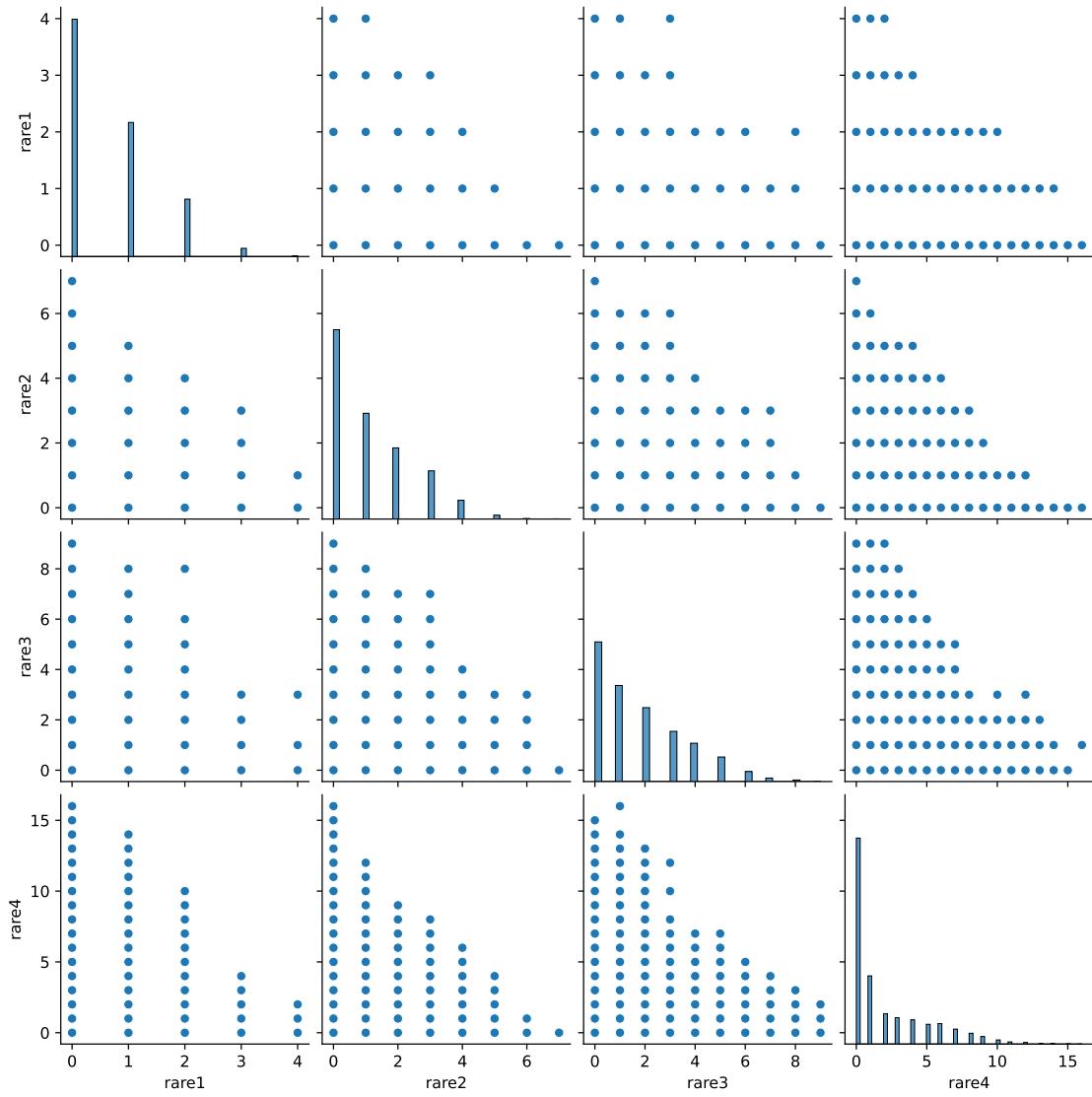


```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd4626487d0>
```



<pandas.io.formats.style.Styler at 0x7fd462625210>

<seaborn.axisgrid.PairGrid at 0x7fd4622b4090>



### 3 Simulation 3

#### 3.1 Code

```

recomb_map = msprime.RecombinationMap.read_hapmap('./Simu3/genetic_map_GRCh37_chr22.txt')
ts = msprime.simulate(
    population_configurations = [
        msprime.PopulationConfiguration(sample_size=10000, growth_rate=0, initial_size=100000)
    ],
    demographic_events = [
        msprime.PopulationParametersChange(time=200, growth_rate=0.05)
    ],
    random_seed = 101, recombination_map=recomb_map, mutation_rate=2e-8
)

```

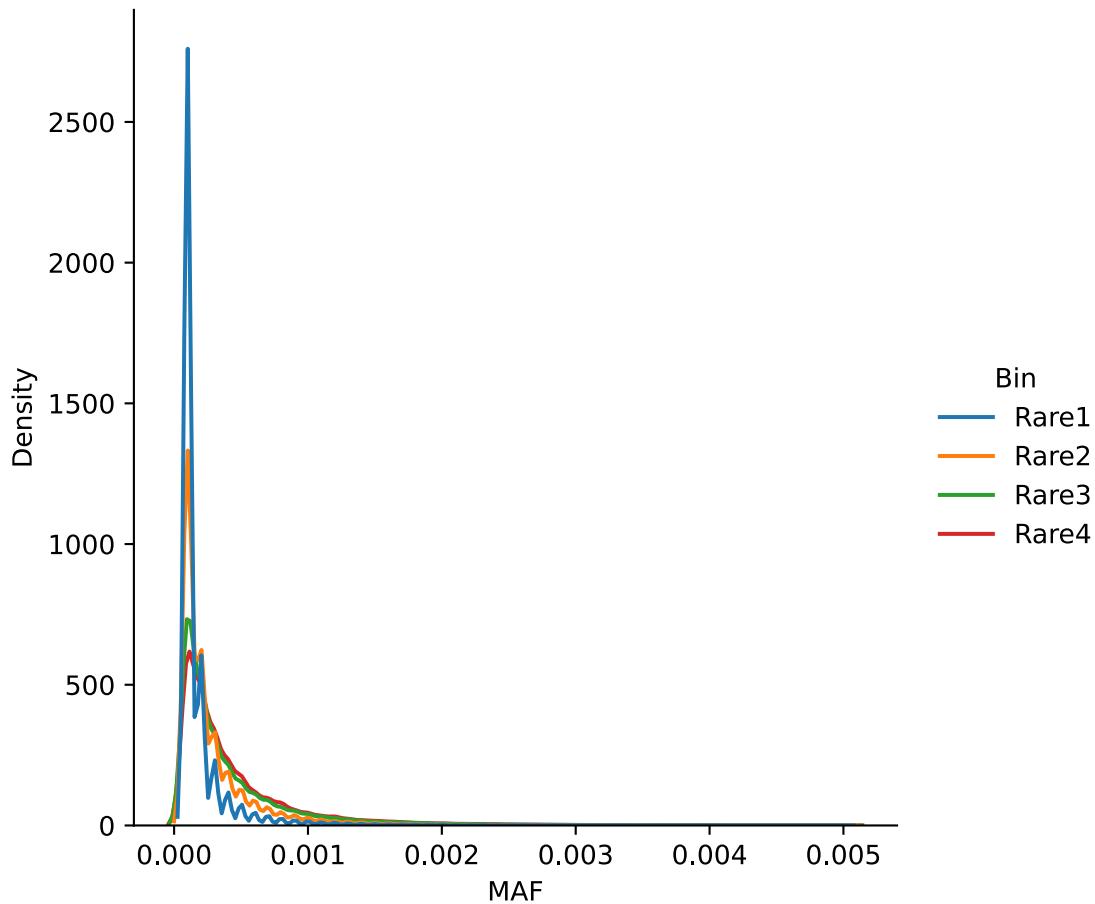
)

### 3.2 Input Summary

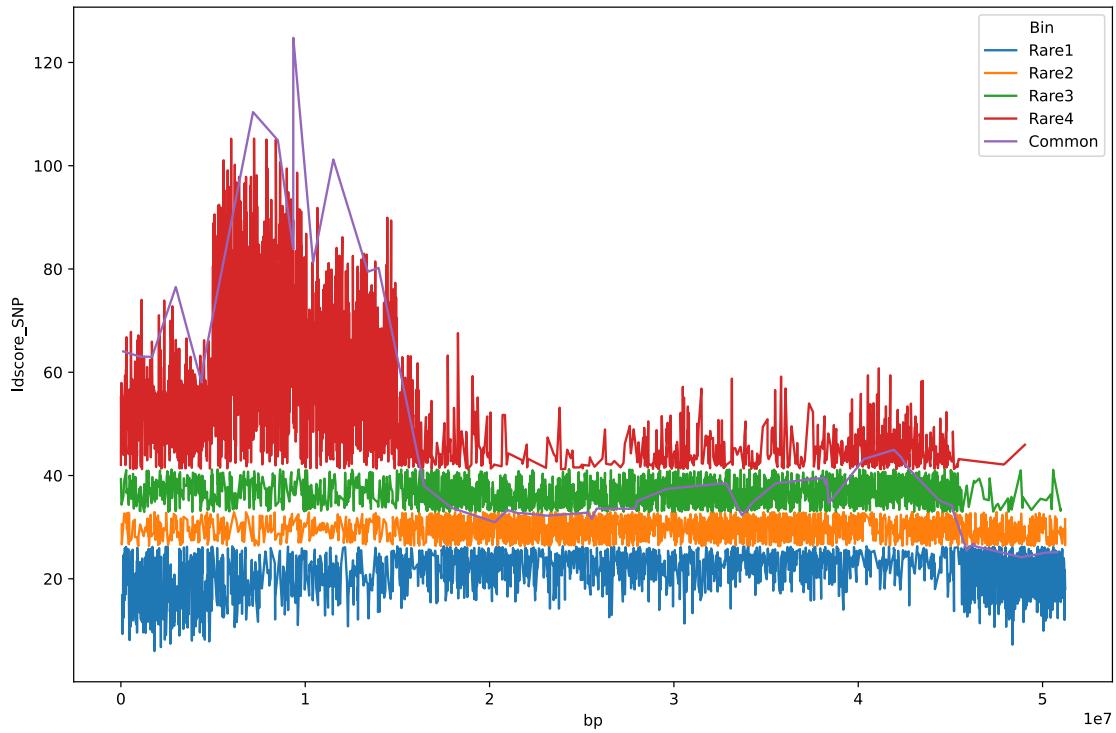
- Instead of fixed recombination rate, in this simulation, hapmap chromosome 22 recombination map is used
- Sample size = 10,000
- Initial population has a fixed growth rate but last 200 generation exponential growth with rate = 0.05
- Mutation rate =  $2 \times 10^{-8}$

```
<tskit.trees.TreeSequence at 0x7fd4622fefd0>
```

```
<seaborn.axisgrid.FacetGrid at 0x7fd4676e0710>
```

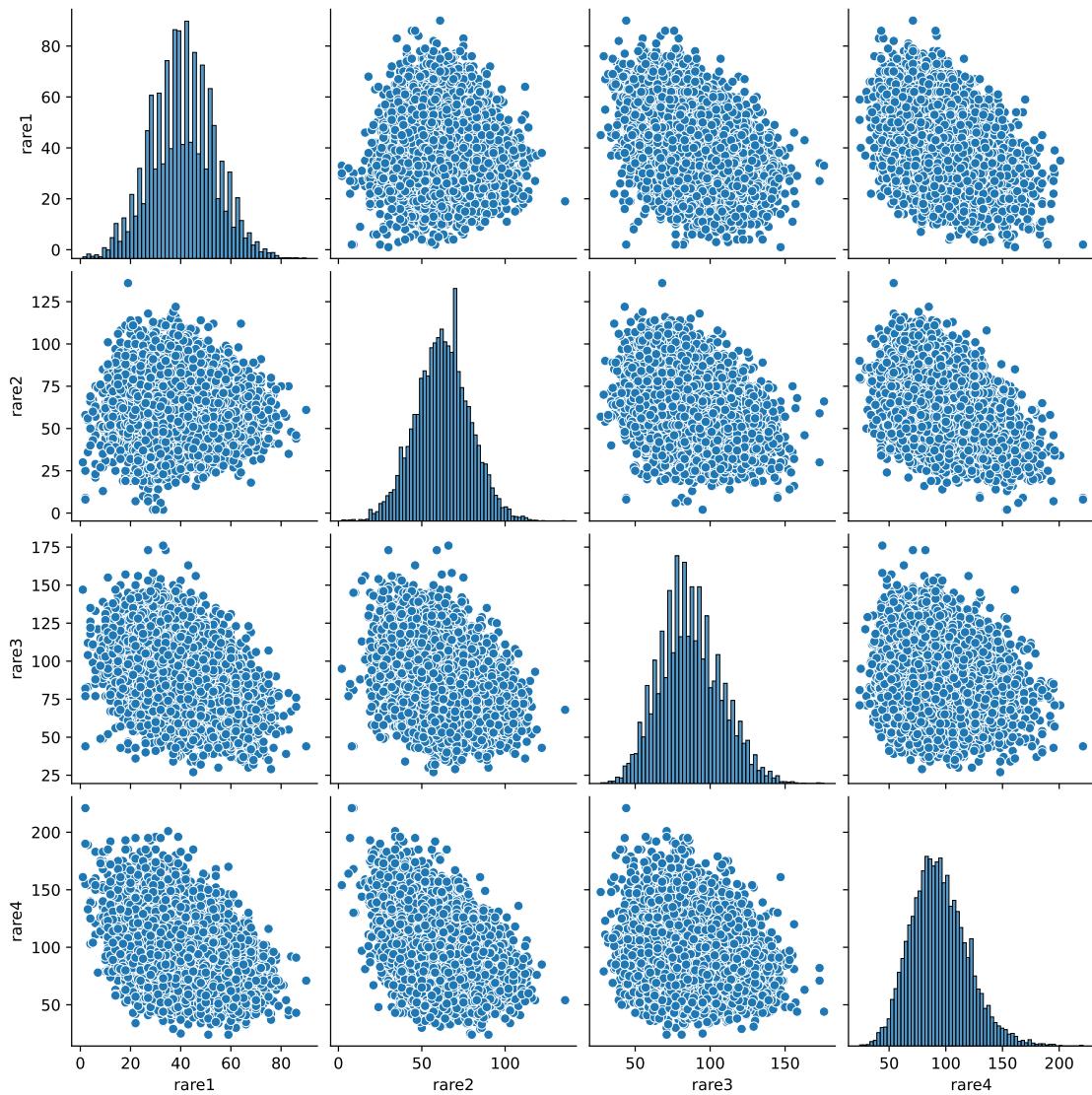


```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd409b10250>
```



<pandas.io.formats.style.Styler at 0x7fd409248a50>

<seaborn.axisgrid.PairGrid at 0x7fd407dc0150>



## 4 Simulation 4

### 4.1 Code

```

species = stdpopsim.get_species('HomSap')
model = species.get_demographic_model('OutOfAfrica_3G09')
contig = species.get_contig("chr22")
samples = model.get_samples(0,10000,0)
engine = stdpopsim.get_engine('msprime')
ts = engine.simulate(model, contig, samples)

```

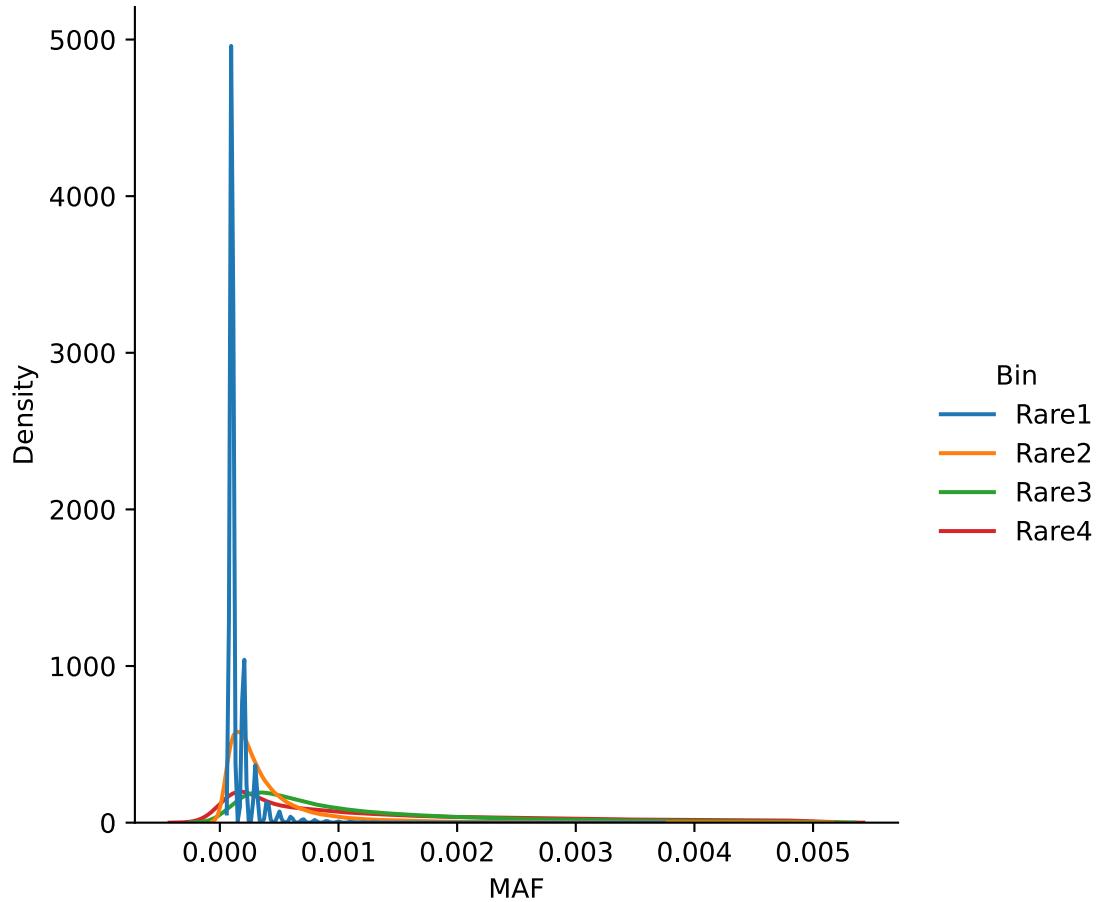
## 4.2 Description

Simulation is done using `stdpopsim` package. This simulation package is a community-maintained standard library of population genetic models. It uses `msprime` for many simulations. Advantage of using `stdpopsim` is that it has some standard built-in models that have been published already. In this simulation, I have used the demographic model from Gutenkunst et al. 2009 : Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data.

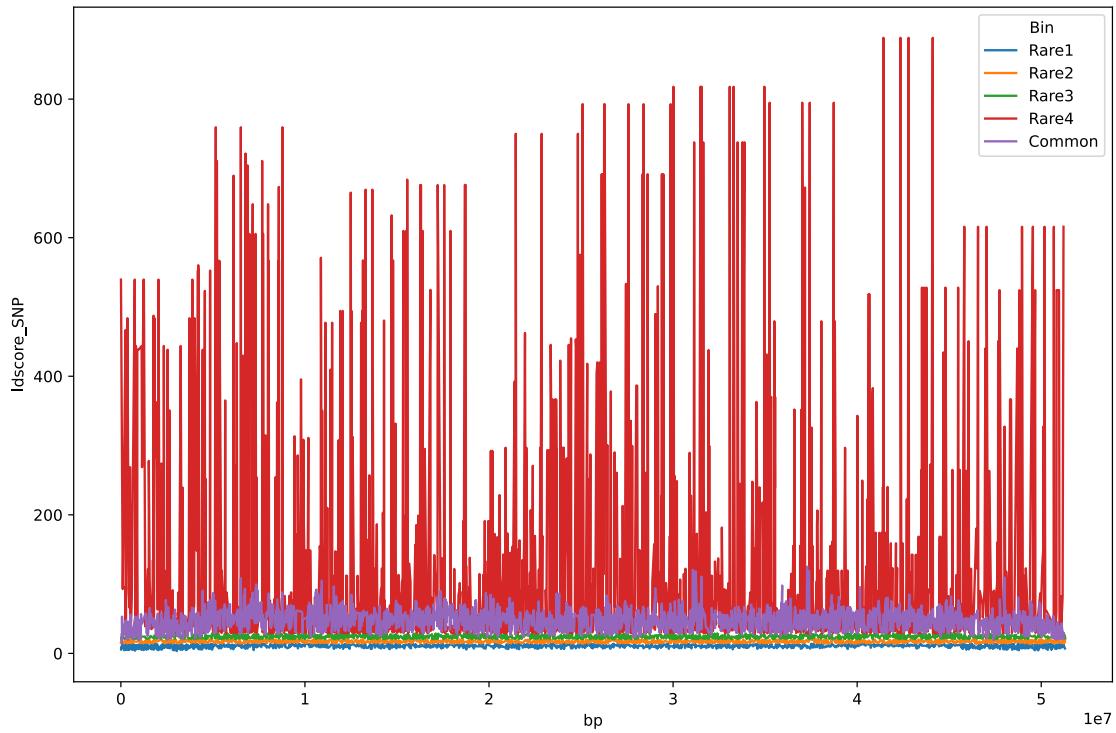
The three population Out-of-Africa model from Gutenkunst et al. 2009.  
It describes the ancestral human population in Africa, the out of Africa event, and the subsequent European-Asian population split.  
Model parameters are the maximum likelihood values of the various parameters given in Table 1 of Gutenkunst et al.

```
[MassMigration(time=848.0, source=2, dest=1, proportion=1.0),  
 MigrationRateChange(time=848.0, rate=0, source=-1, dest=-1),  
 MigrationRateChange(time=848.0, rate=0.00025, source=0, dest=1),  
 MigrationRateChange(time=848.0, rate=0.00025, source=1, dest=0),  
 PopulationParametersChange(time=848.0, initial_size=2100, growth_rate=0,  
 population=1),  
 MassMigration(time=5600.0, source=1, dest=0, proportion=1.0),  
 MigrationRateChange(time=5600.0, rate=0, source=-1, dest=-1),  
 PopulationParametersChange(time=8800.0, initial_size=7300, growth_rate=None,  
 population=0)]
```

```
<seaborn.axisgrid.FacetGrid at 0x7fd405a3ee90>
```



```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd40af11050>
```



<pandas.io.formats.style.Styler at 0x7fd409cb8810>

<seaborn.axisgrid.PairGrid at 0x7fd40975a950>

