MDR-CaseCount-And-FreqPercentage

February 25, 2020

```python
[1]: # Imports
     import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import matplotlib.ticker as ticker
```

```python
[2]: def MDR_plot_geno_freq(drug, drug_res):
         # Create Data Frame
         df = pd.read_csv(drug, sep='\t')
         # plot
         plt.plot(df['current_time'], df.iloc[:,0:150].filter(regex=drug_res,␣
      ↪axis=1).sum(axis=1))
```

```python
[3]: def MDR_case_count(drug, drug_res):
         # Create Data Frame
         df = pd.read_csv(drug, sep='\t')
         # Total genotype count
         total = df.iloc[:,22:150].sum(axis=1).tail(1).values
         # Total population count
         total_pop = df['population'].tail(1).values
         # Filter out target genotypes
         mdr_count = df.iloc[:,0:150].filter(regex=drug_res, axis=1).sum(axis=1).
      ↪tail(1).values
         percentage = mdr_count[0] / total_pop[0] * 100
         return [mdr_count[0], total_pop[0], percentage]
```

```python
[9]: def MDR_percentage_time(drug, drug_res, threshold):
         # Create Data Frame
         df = pd.read_csv(drug, sep='\t')
         # Total genotype count by time
         df['total_geno_freq'] = df.iloc[:,22:150].sum(axis=1)
         # Filtered out genotype count by time
         df['mdr_count'] = df.iloc[:,0:150].filter(regex=drug_res, axis=1).
      ↪sum(axis=1)
         # Calculate the percentage by time
         df['current_percentage'] = df['mdr_count'] / df['total_geno_freq']
         # Get first row # that's bigger than threshold
         frn = df[df['current_percentage'].gt(threshold)].index[0]
```

```
    # Get which year that is and function returns
    return (df.loc[frn].iloc[0] / 365)
```

1 Main Function

```
[5]: # Figure Size
    plt.rcParams['figure.figsize'] = [20, 10]
```

```
# Define Drug and Resistance Pattern
drug = 'al.txt'
drug_res = 'KNF..Y..'

#drug = 'asaq.txt'
#drug_res = 'TYY..Y..'

#drug = 'dhappq.txt'
#drug_res = '.....Y2.'
```
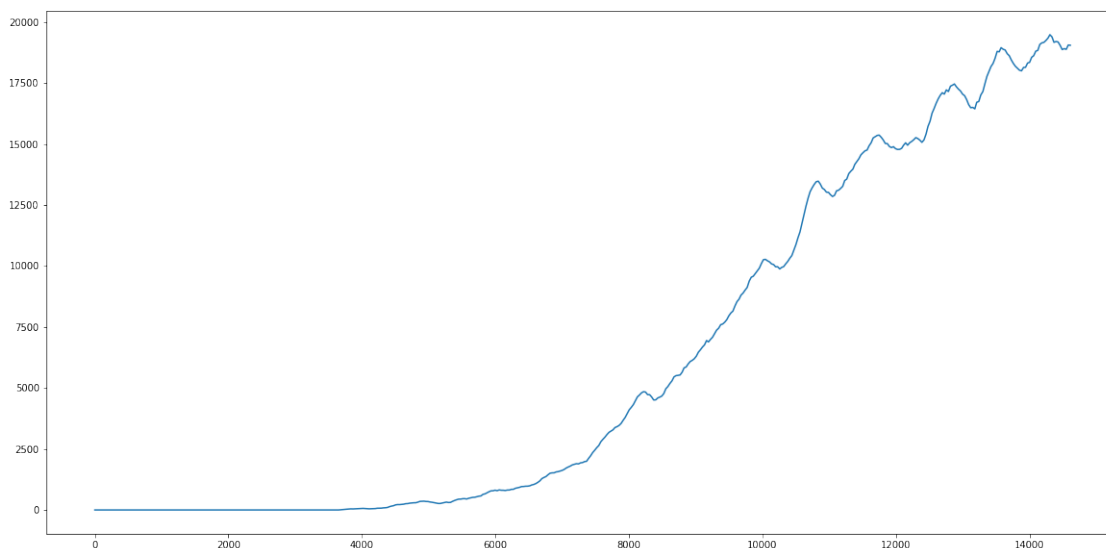
1.1 AL

```
[6]: MDR_case_count('al.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|KNF)..Y..$')
```

```
[6]: [19044, 148249, 12.845955116054746]
```

```
[7]: MDR_plot_geno_freq('al.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|KNF)..Y..$')
```



```
[10]: MDR_percentage_time('al.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|KNF)..Y..$', 0.01)
```

```
[10]: 10.841095890410958
```

```
[11]: MDR_percentage_time('al.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|KNF)..Y..$', 0.02)
```

[11]: 12.093150684931507

```
[12]: MDR_percentage_time('al.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|KNF)..Y..$', 0.03)
```
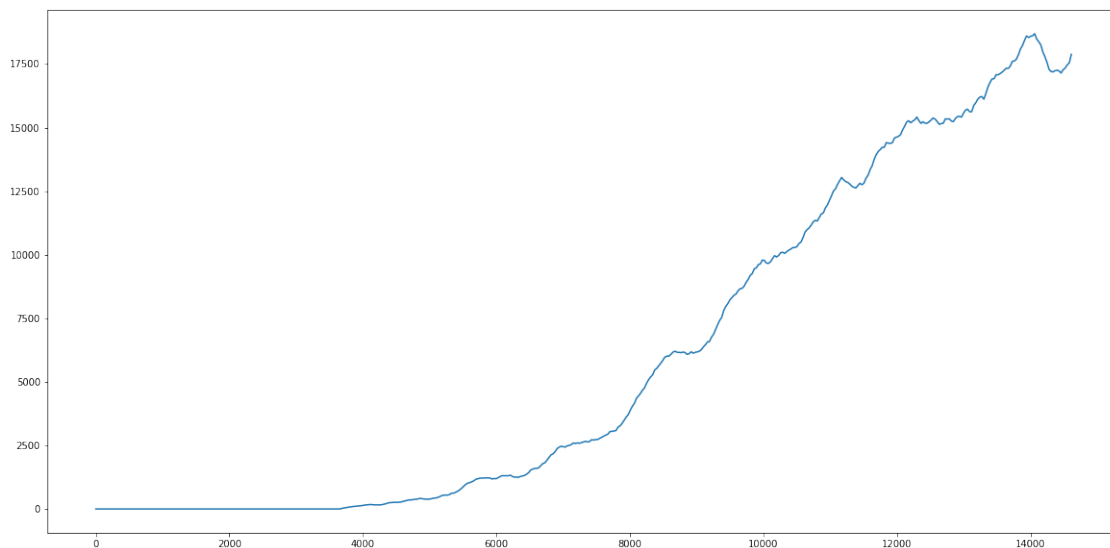
[12]: 12.756164383561643

1.2 AS-AQ

```
[13]: MDR_case_count('asaq.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|TYY)..Y..$')
```

[13]: [17880, 149123, 11.990102130456066]

```
[14]: MDR_plot_geno_freq('asaq.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|TYY)..Y..$')
```



```
[15]: MDR_percentage_time('asaq.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|TYY)..Y..$', 0.01)
```

[15]: 10.504109589041096

```
[16]: MDR_percentage_time('asaq.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|TYY)..Y..$', 0.02)
```

[16]: 10.841095890410958

```
[17]: MDR_percentage_time('asaq.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|TYY)..Y..$', 0.03)
```
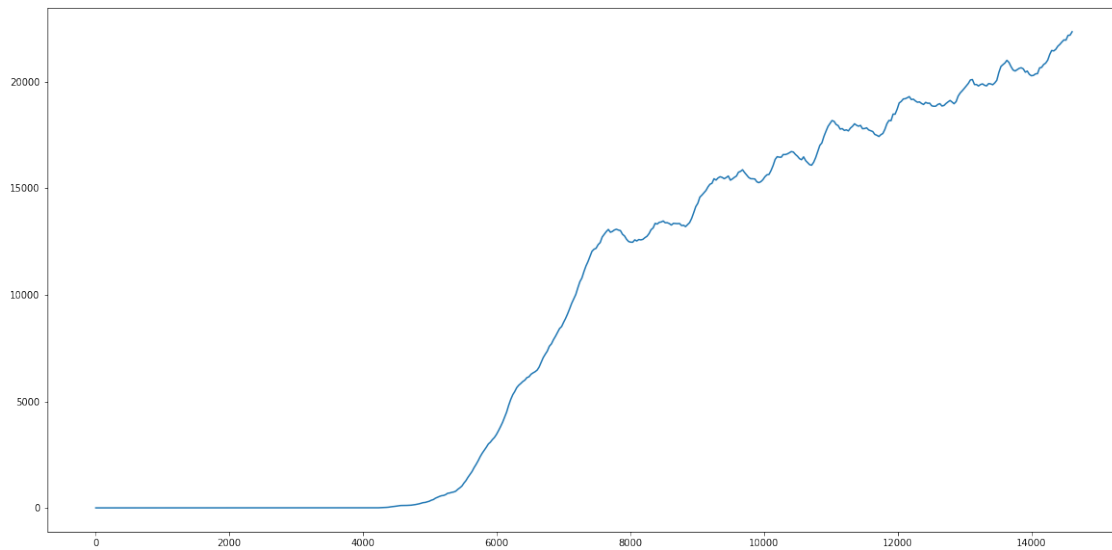
[17]: 11.093150684931507

1.3 DHA-PPQ

```
[18]: MDR_case_count('dhappq.txt', '^.....Y2.$')
```

[18]: [22344, 149582, 14.93762618496878]

```
[19]: MDR_plot_geno_freq('dhappq.txt', '^.....Y2.$')
```



```
[20]: MDR_percentage_time('dhappq.txt', '^.....Y2.$', 0.01)
```

```
[20]: 12.254794520547945
```

```
[21]: MDR_percentage_time('dhappq.txt', '^.....Y2.$', 0.02)
```

```
[21]: 13.254794520547945
```

```
[22]: MDR_percentage_time('dhappq.txt', '^.....Y2.$', 0.03)
```

```
[22]: 13.504109589041096
```

```
[ ]:
```