

# Ploting with hail

In [1]:

```
import numpy as np
import pandas as pd
import hail as hl
from hail.plot import show
hl.init(quiet=True)
hl.plot.output_notebook()
```

<https://docs.hail.is/> successfully loaded.

In [3]:

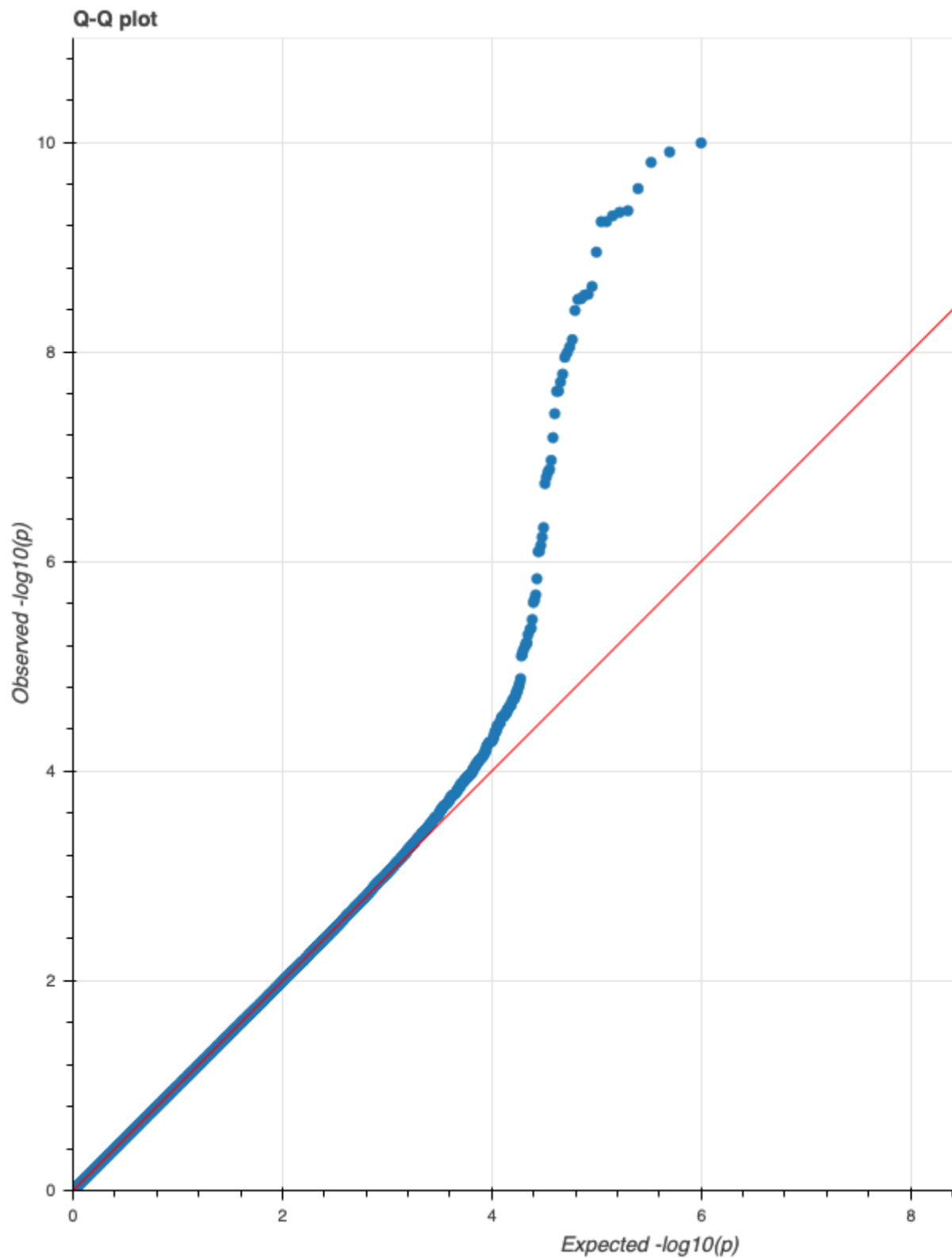
```
results = hl.import_table("/Users/bambrozi/Downloads/1miSNP5kSamples.T2D.glm.logistic_regression_results.txt",
    # key='ID',
    # impute=True
    types={
        '#CHROM': hl.tstr,
        'POS': hl.tint32,
        'OBS_CT': hl.tint64,
        'OR': hl.tfloat64,
        'LOG(OR)_SE': hl.tfloat64,
        'Z_STAT': hl.tfloat64,
        'P': hl.tfloat64,
    })
```

In [7]:

```
# There's too many simulated significant SNP's. e.g. 3127e-249.
# No sense! Let's remove them to have a "prettier Plot".
# Cut-off: 1e-10
results = results.filter(results.P > 1e-10)
```

In [8]:

```
# QQ-Plot  
p = hl.plot.qq(results['P'])  
show(p)
```



In [9]:

```
# Manhattan plot
p = hl.plot.manhattan(
    pvals=results['P'],
    # locus=hl.locus(results['mock_chr'], results['POS']),
    locus=hl.locus(results['#CHROM'], results['POS']),
    title="SUNGDL-496 - Dataset 20k Samples & 8million SNPs",
    hover_fields=dict(Chromosome=results['#CHROM'],
                      ID=results["ID"],
                      REF=results["REF"],
                      ALT=results["ALT"],
                      A1=results["A1"],
                      p_value=results["P"]),
    # collect_all=True,
    # n_divisions=500,
    significance_line=5e-8
)
show(p)
```

