Analyze 3D-seq data

MATLAB script for modeling & detecting peaks in 3D-seq conversion ratio data.

Paul Wiggins (University of Washington) 04/19/2022

Contents

Set things up.

```
% This is where I am now
disp( pwd );

% What file is running:
disp( [mfilename('fullpath'),'.m'] );
```

```
/Users/wiggins/Google Drive/PAW/peakDetection/scripts/clean/clean/3Dseq
/Users/wiggins/Google Drive/PAW/peakDetection/scripts/clean/clean/3Dseq/analyze_3D_seq_data.m
```

```
clear all;
close all;
colordef white
```

Load in data from a file.

We need to get the genomic positions into vector x and the conversion fraction at position x into D

```
data_ = load( '3Dseq_data.mat' );
D = data_.D;
x = data_.x;
```

Construct data

Get rid of SNP bands run data decimation to make a reduced resolution dataset.

```
data = makeData3D( x, D );
```

```
nbin =

       25058
```

Fit peaks

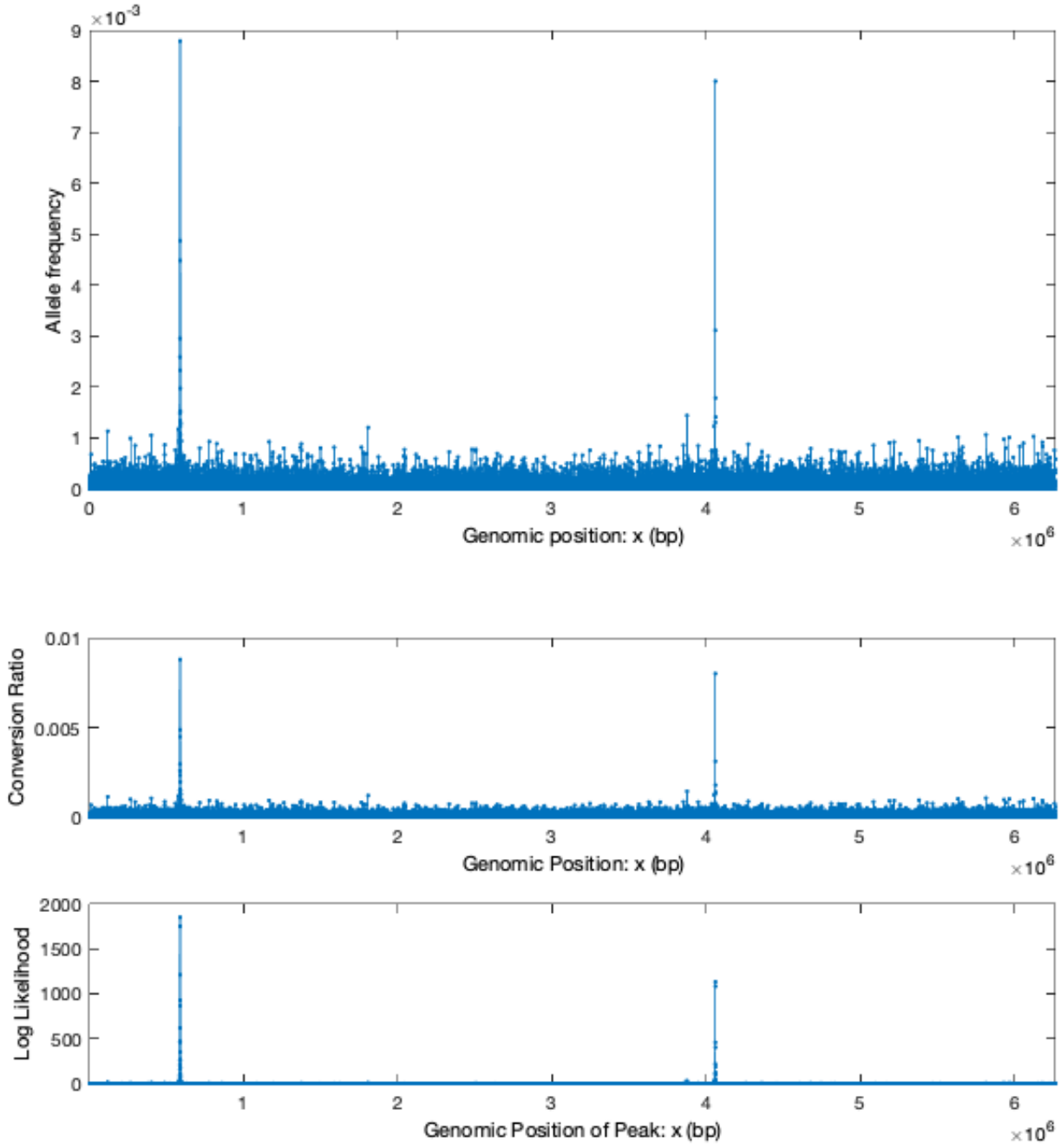
```
data.T0.D0 = 1;
data.T0.L0 = 200;
data =  makeAnalysis3D( data );
```

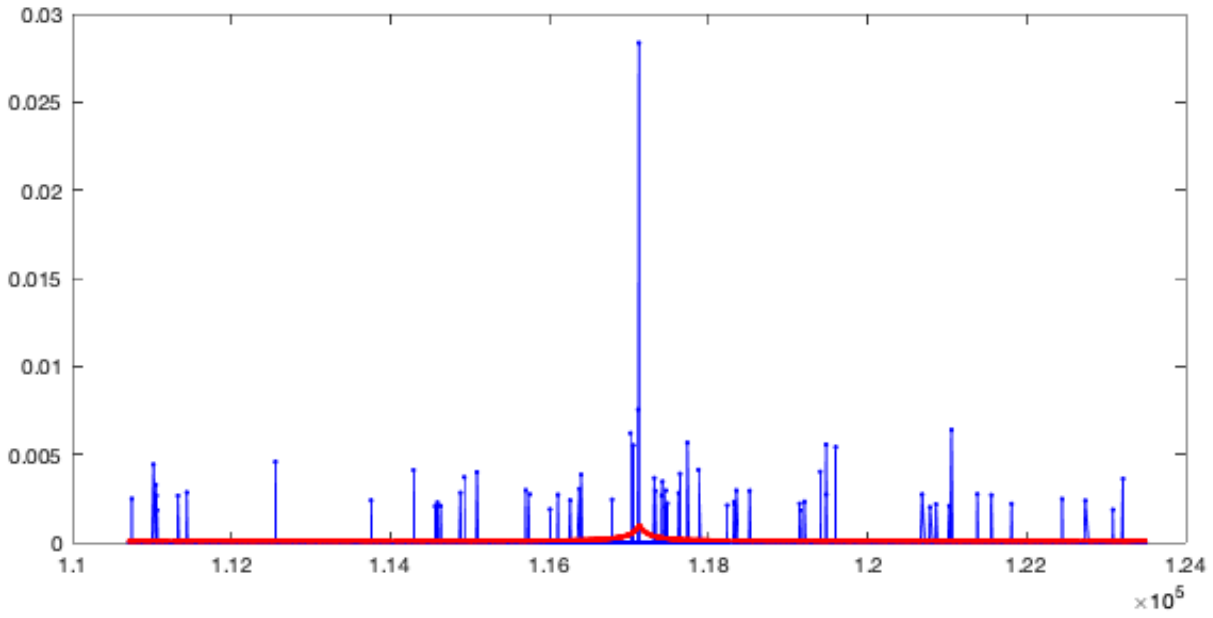
```
Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.


Local minimum possible.

lsqnonlin stopped because the final change in the sum of squares relative to
its initial value is less than the value of the function tolerance.

P value is
     0

Test statistic is
   1.9046e+03


Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum possible.

lsqnonlin stopped because the final change in the sum of squares relative to
its initial value is less than the value of the function tolerance.

P value is
     0

Test statistic is
   1.1953e+03


Local minimum found.
```

```
Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
     1

Test statistic is
  -22.8539


Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
    0.9074

Test statistic is
   20.4385


Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
    0.9232

Test statistic is
   20.0218
```

```
Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
    0.9993

Test statistic is
   14.3524


Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
     1

Test statistic is
   -8.1413


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
     1

Test statistic is
   -3.9435
```

```
Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
    1.0000

Test statistic is
    6.6331


Initial point is a local minimum.

Optimization completed because the size of the gradient at the initial point
is less than the value of the optimality tolerance.


Local minimum found.

Optimization completed because the size of the gradient is less than
the value of the optimality tolerance.

P value is
    1.0000

Test statistic is
   11.2242
```


×10^-3
Allele frequency
Genomic position: x (bp)
×10^6
Conversion Ratio
Genomic Position: x (bp)
×10^6
Log Likelihood
Genomic Position of Peak: x (bp)
×10^6



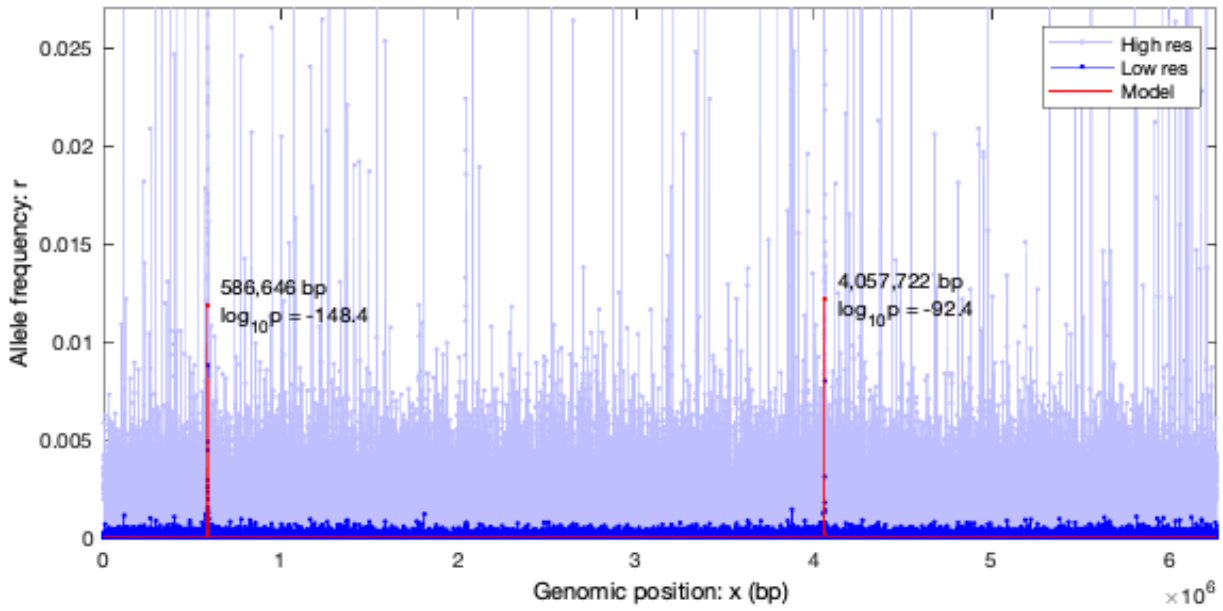
Visualize model

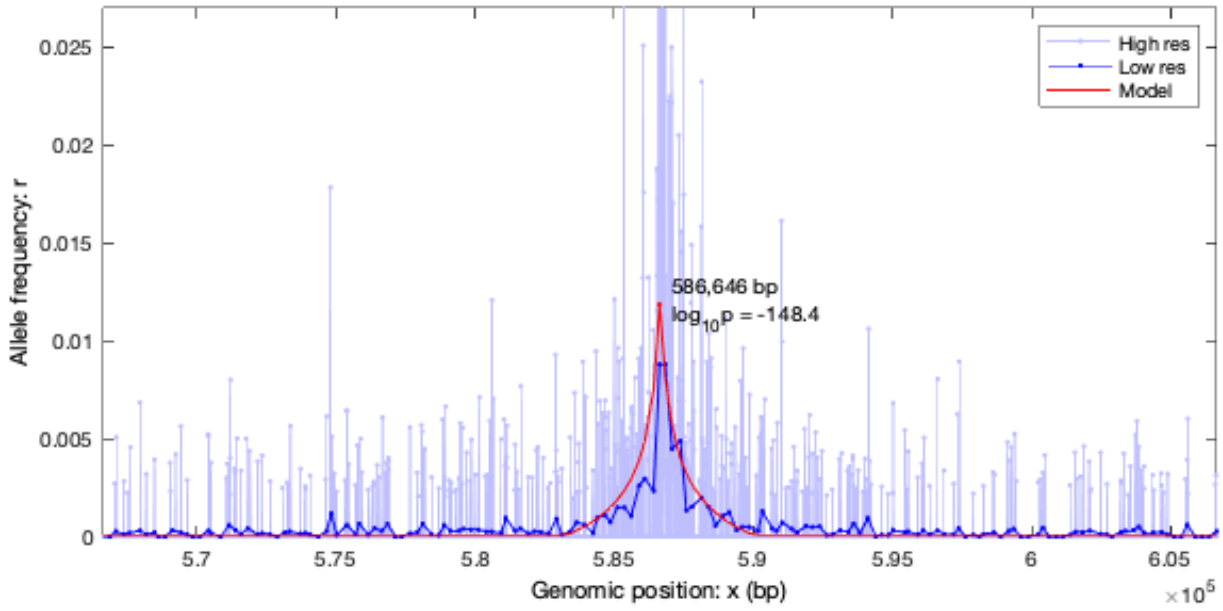
Show plots of the entire genome as well as individual peaks, and generating excel files of data and models

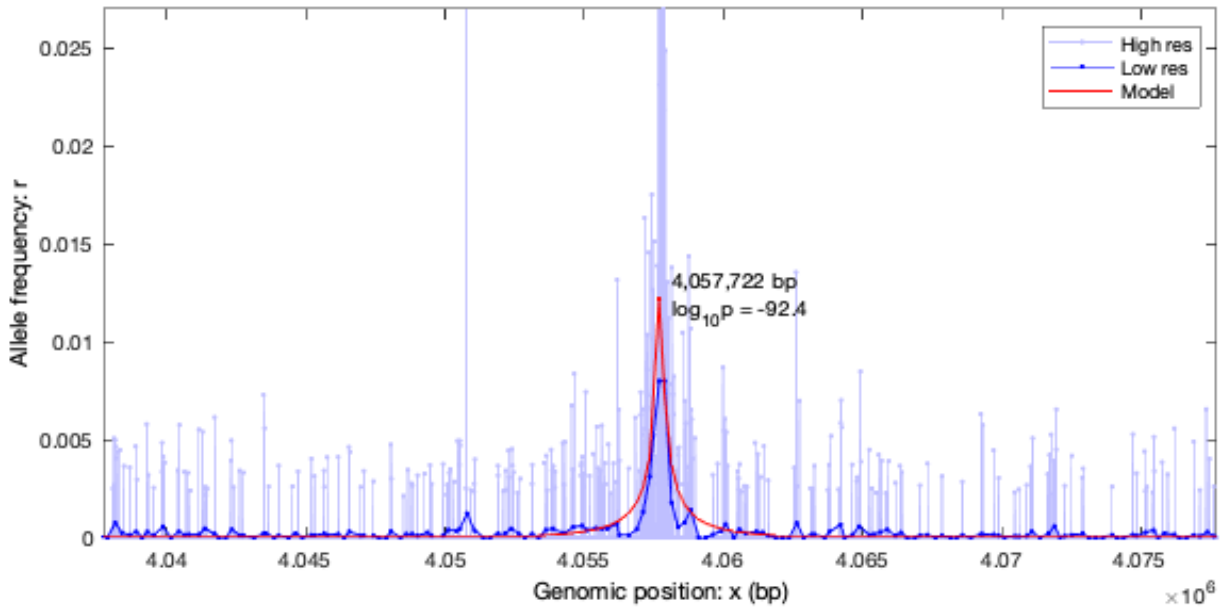
```
data = showAnalysis3D( data );
```

```
filename =

    'Figs/Full_low.xls'
```

```
filename =

    'Figs/Model.xls'
```


High res
Low res
Model
586,646 bp
log10p = -148.4
4,057,722 bp
log10p = -92.4
Allele frequency: r
Genomic position: x (bp)
×10^6


High res
Low res
Model
586,646 bp
log10p = -148.4
Allele frequency: r
Genomic position: x (bp)
×10^5



Published with MATLAB® R2019b