SeqMonk tools for methylation analysis
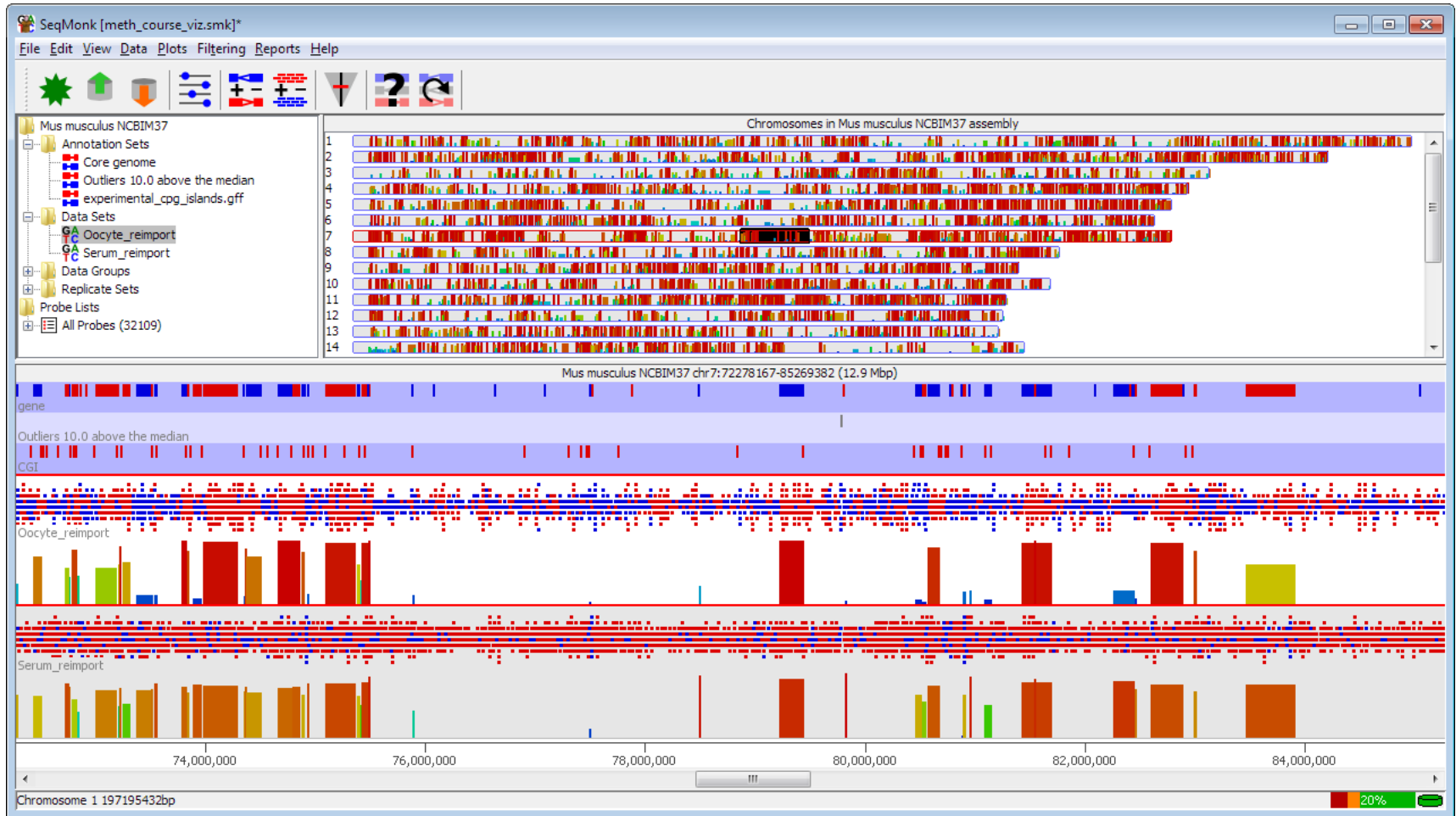
Simon Andrews
simon.andrews@babraham.ac.uk
@simon_Andrews
2016-11

Babraham
Bioinformatics

SeqMonk

SeqMonk Data Model

- Conventional data (ChIP-Seq, RNA-Seq etc)
  - Data is reads (BAM files etc)
  - Strand indicates genomic strand

- BS-Seq and related data
  - Data is methylation calls
  - All 'reads' are 1bp in length
  - Strand indicates meth state (+=meth -=unmeth)
  - Original strand comes from the imported file

Babraham
Bioinformatics

Raw Data

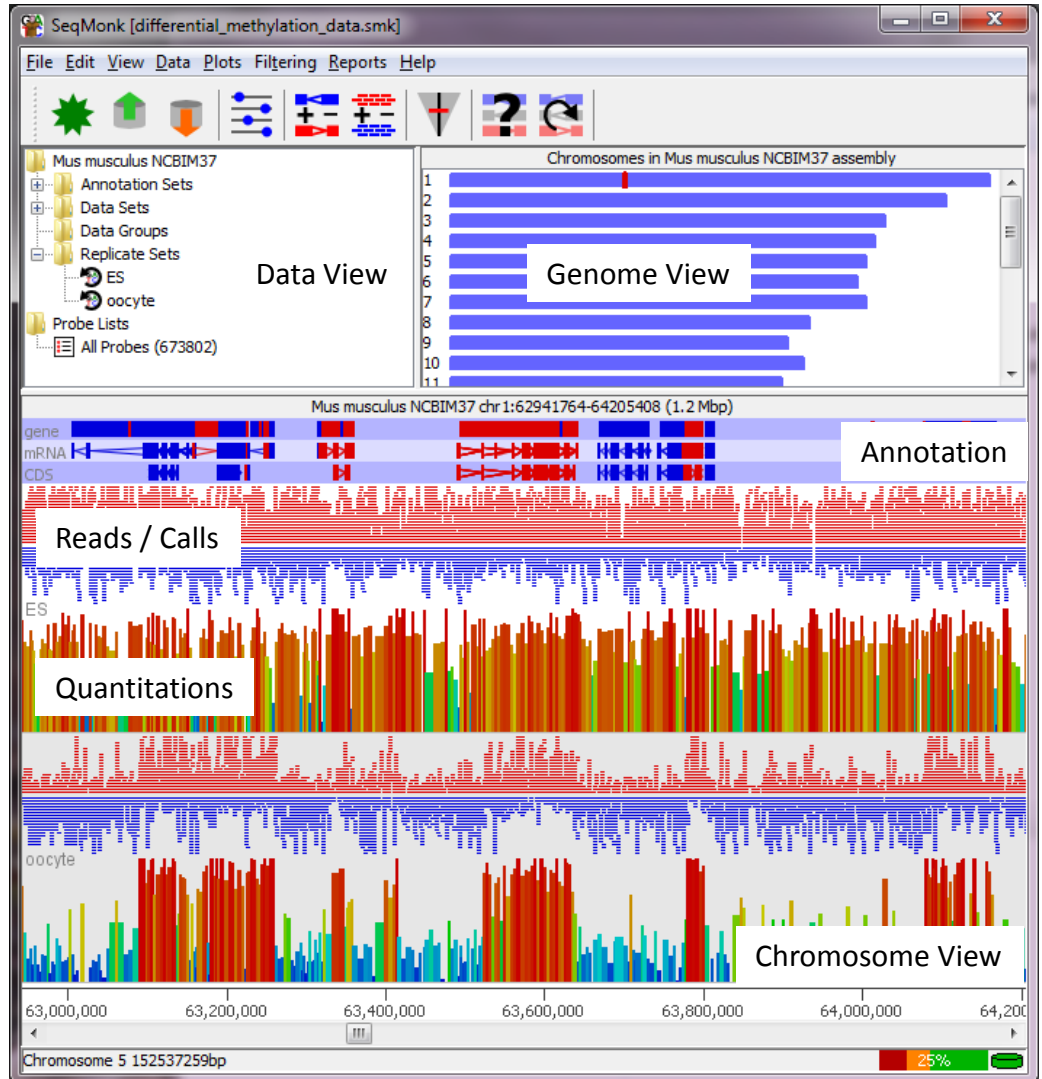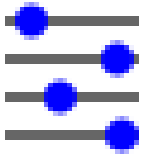Mus musculus NCBIM37 chr7:88379990-88381219 (1.2 kbp)
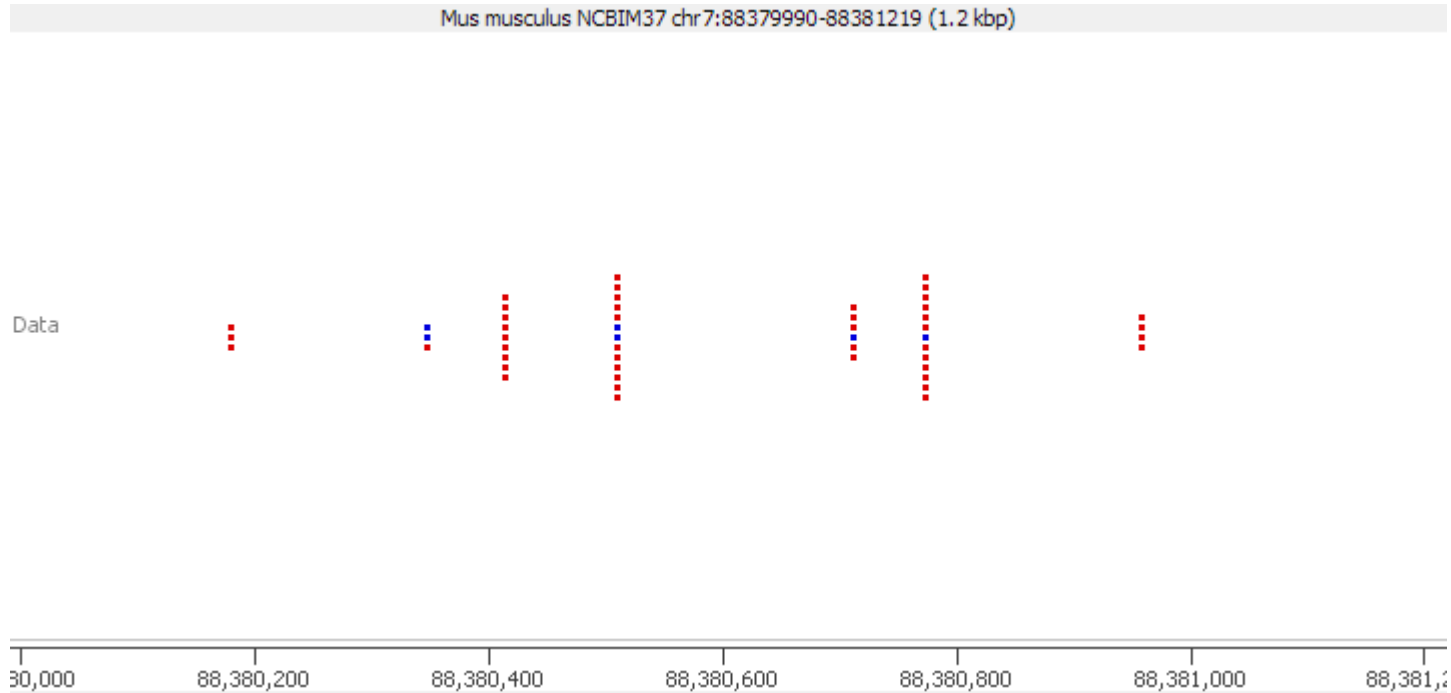
Red = Meth

Blue = Unmeth

Basic Movement Controls

- Move left right
  - Drag bottom scrollbar
  - Mouse scroll wheel
  - Left/Right arrows
- Zoom In
  - Drag a box and release
  - Up arrow
- Zoom Out
  - Right mouse button
  - Down arrow
- Find a feature
  - Edit > Find Feature
  - Control+F
- Change chromosome
  - Edit > Goto Position
  - Drag a box in the genome view

Babraham Bioinformatics

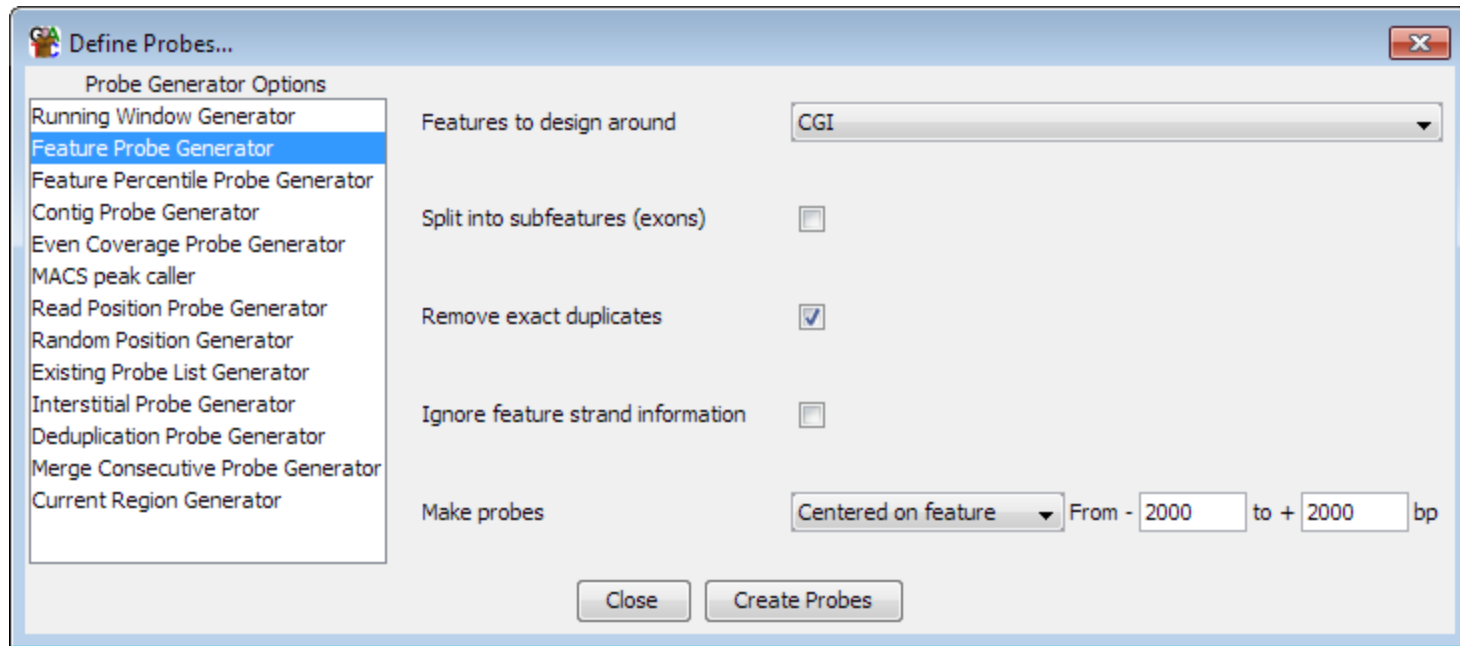# Raw Data Display

View > Data Track Display
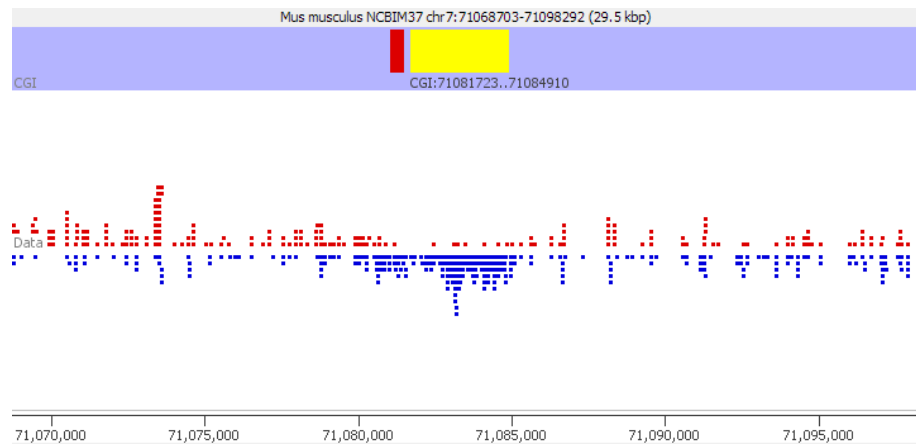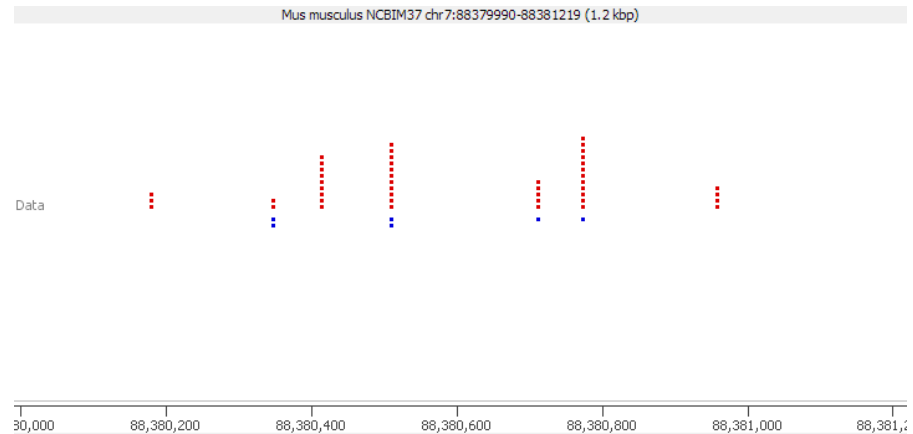
Quantitation Model

- Probe = Location to make a measurement

- ProbeSet = Collection of probes

- Quantitation associates a value with each probe for each data set.

- Define Probes > Quantitate Probes > Visualise
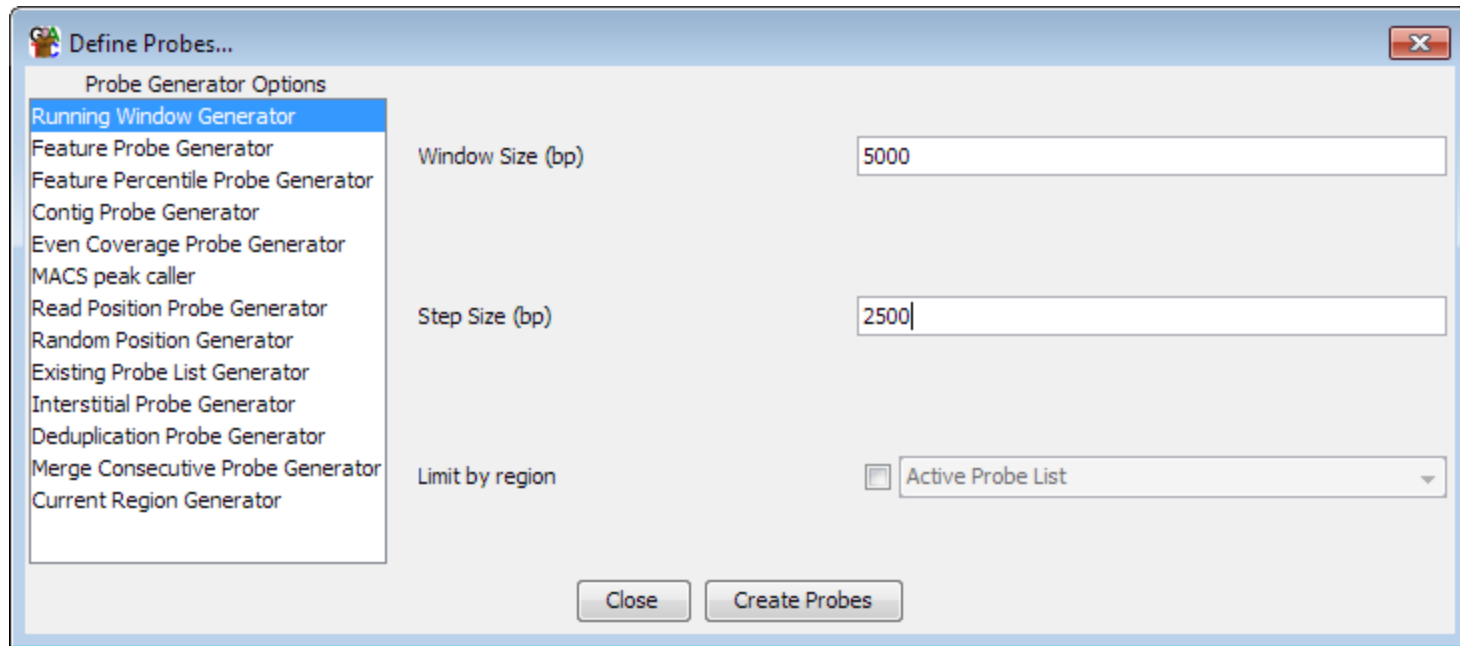
Probe Generation

Targeting Measurement Features
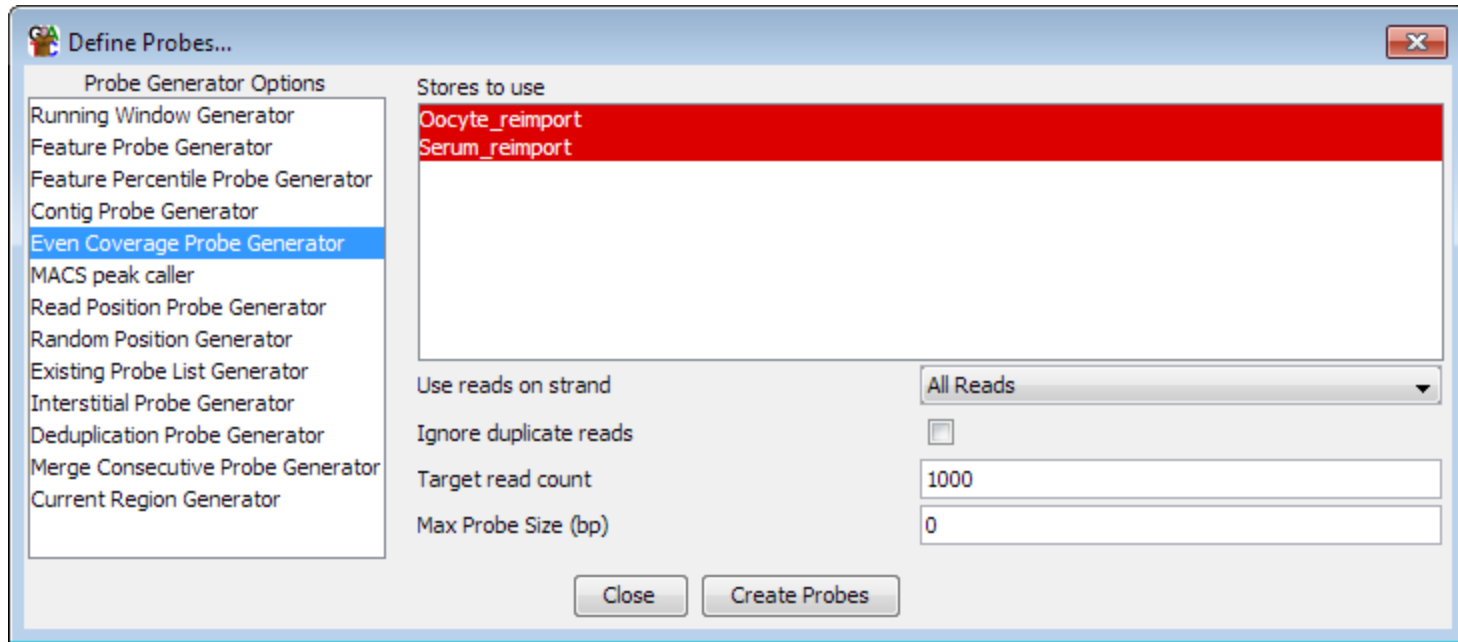

Data > Define Probes > Feature Probe Generator

Targeting Measurement
Fixed Windows

Data > Define Probes > Running Window Generator
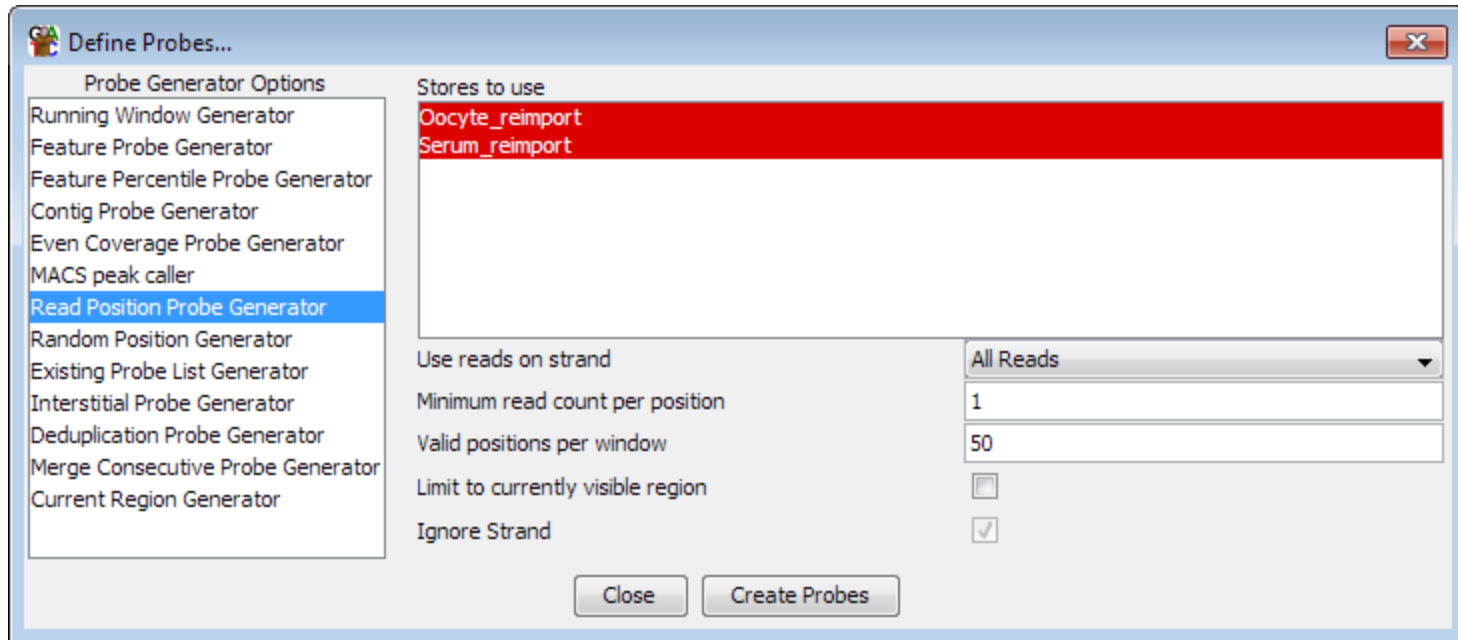
Babraham Bioinformatics

Targeting Measurement
Fixed number of calls

Data > Define Probes > Even Coverage Probe Generator

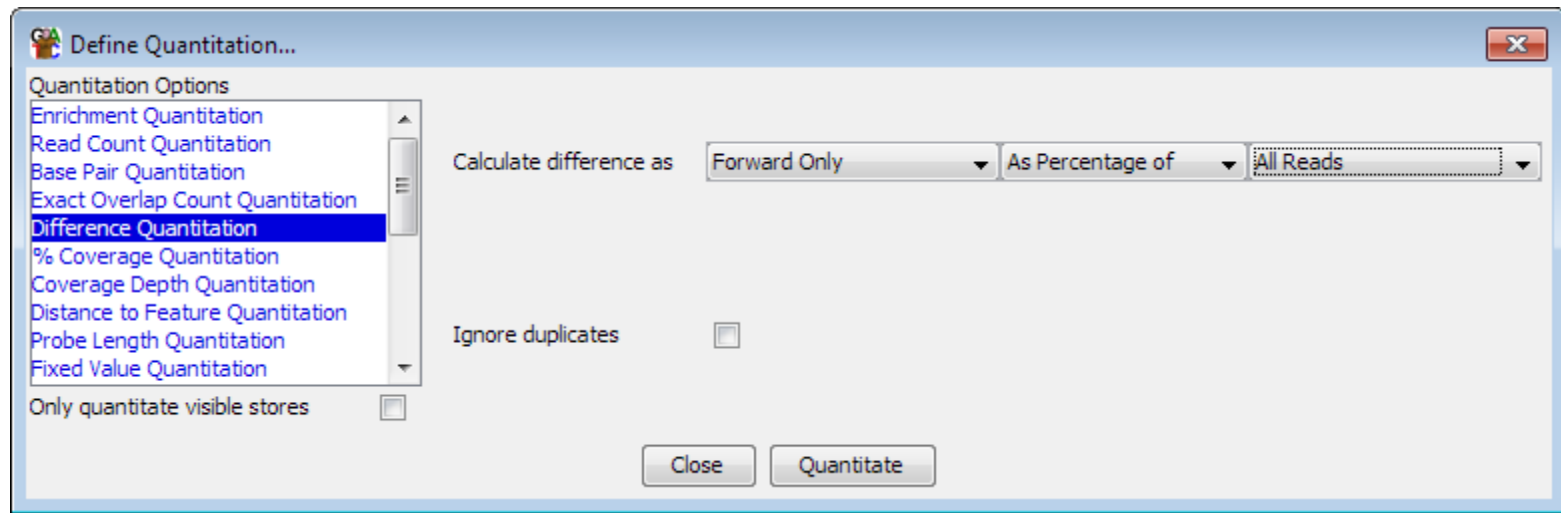# Targeting Measurement
Fixed number of call positions

Data > Define Probes > Read Position Probe Generator

Quantitation

Methylation Measurement
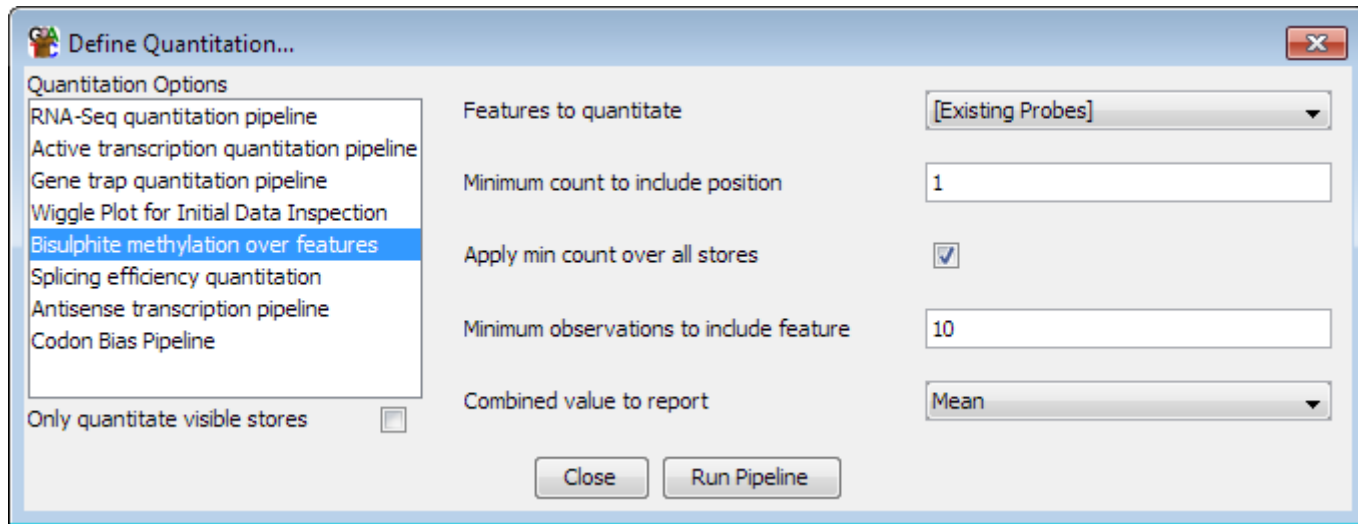# Simple percentage of all calls

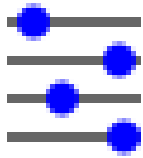Data > Quantitate Existing Probes > Difference Quantitation

Methylation Measurement
More complex corrected measure

Data > Quantitation Pipelines > Bisulphite methylation over features

Babraham
Bioinformatics

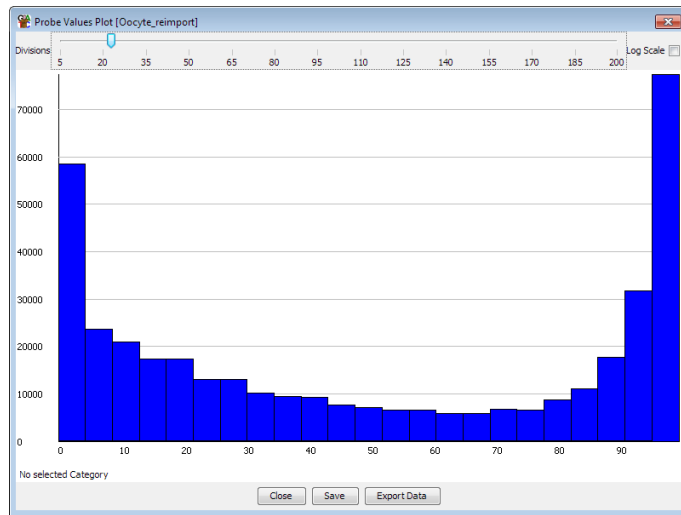# Visualisation of quantitated methylation
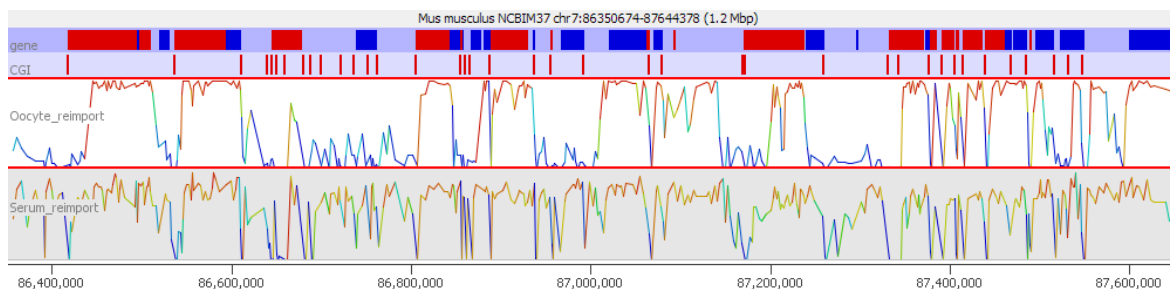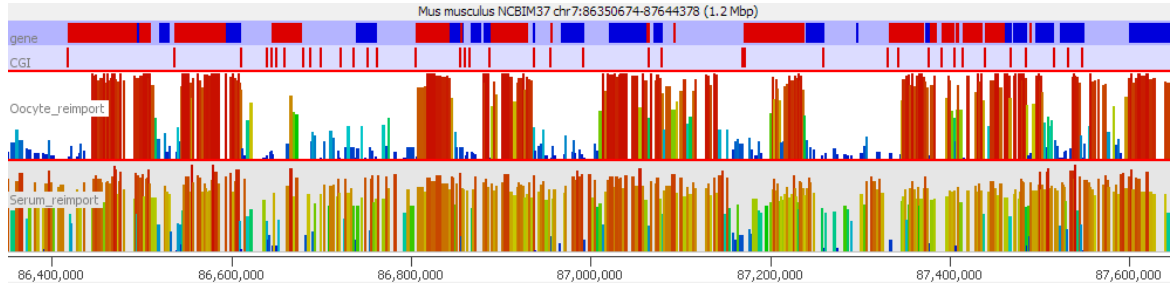
View > Data Track Display

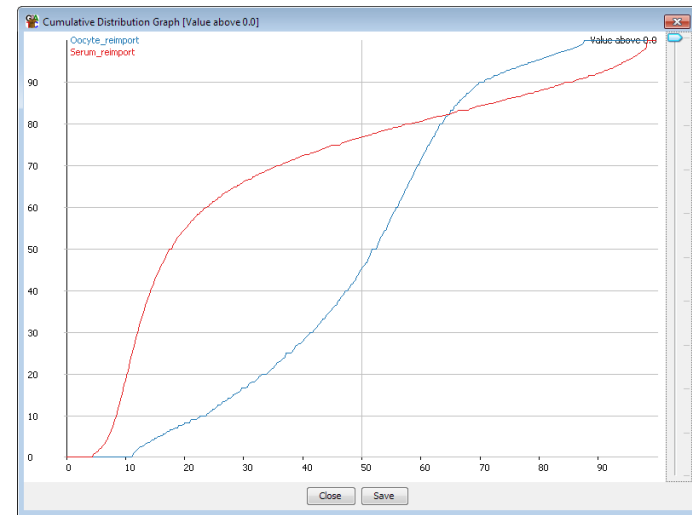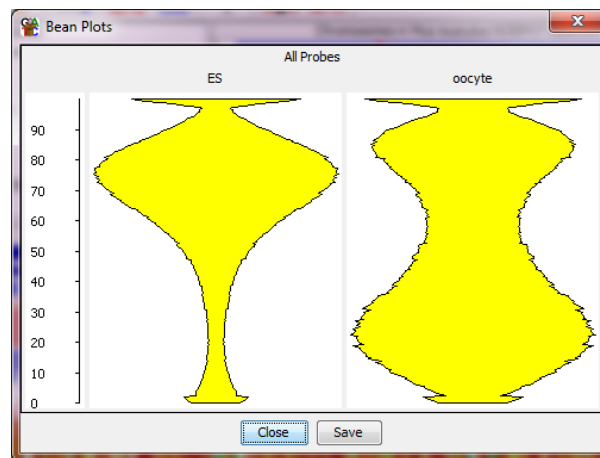View > Set Data Zoom Level
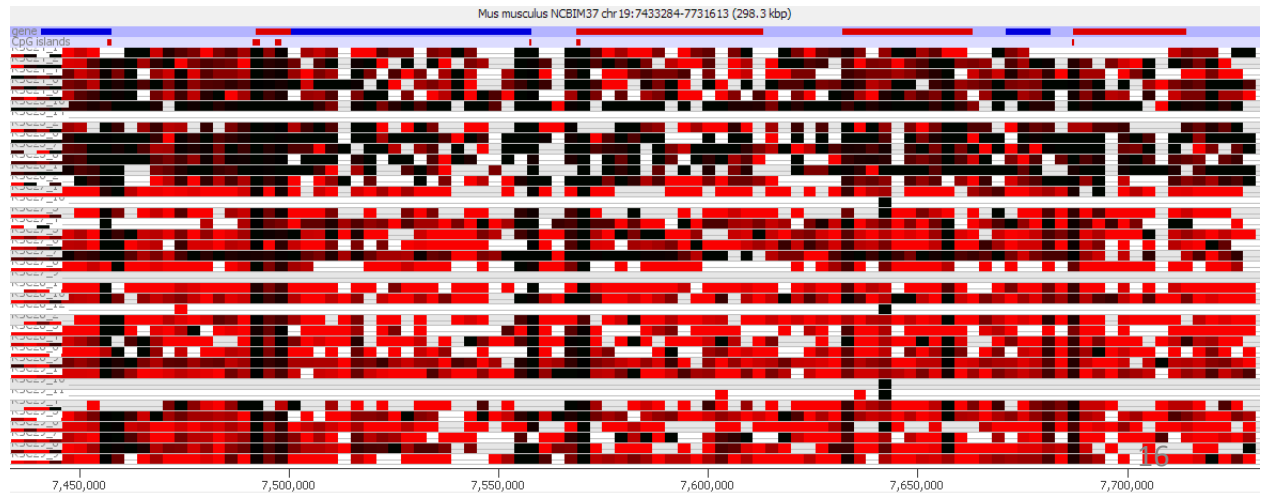
Distributions
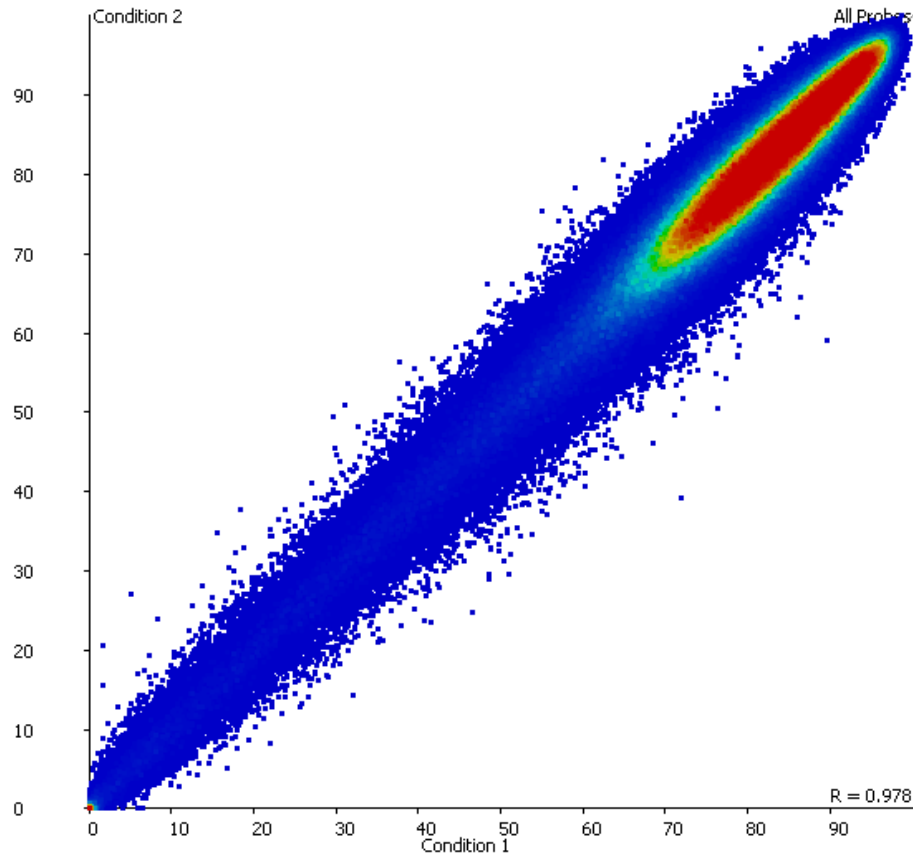

Plots > Probe value histogram

Plots > Cumulative Distribution Plot
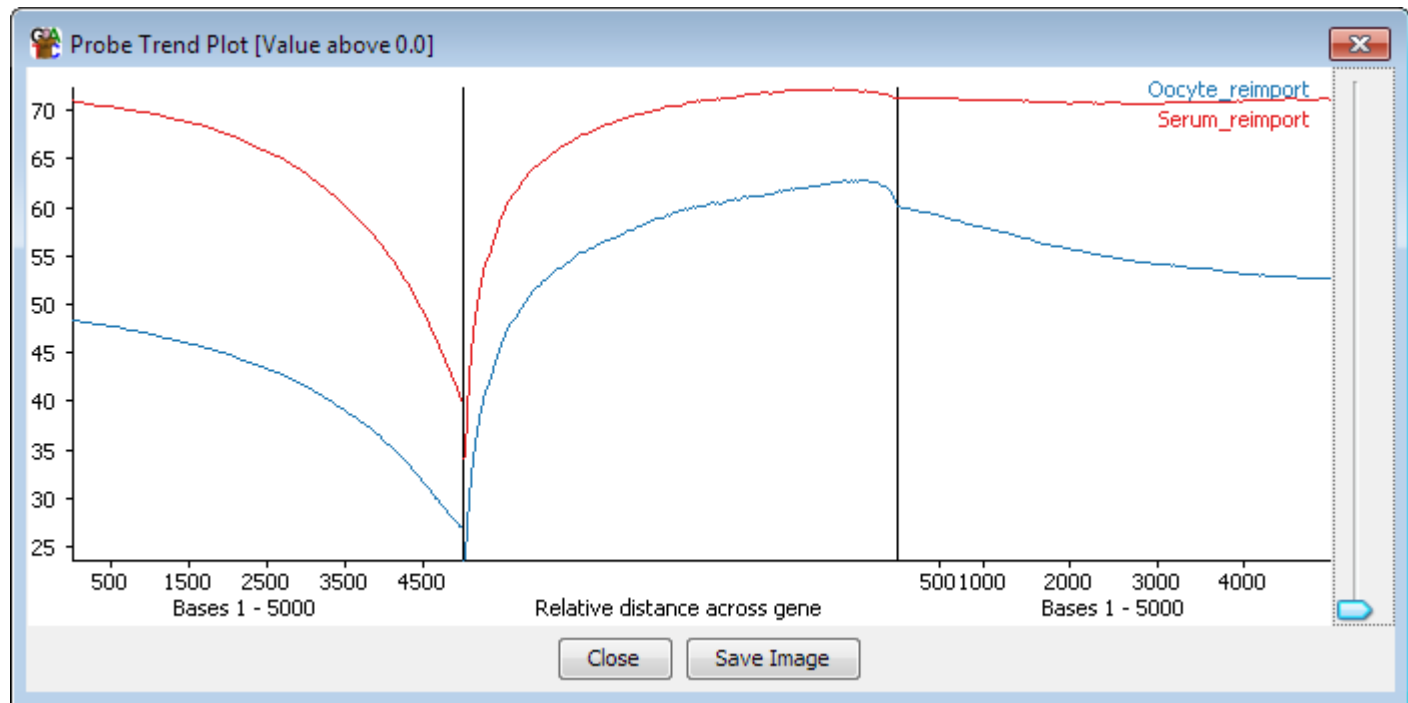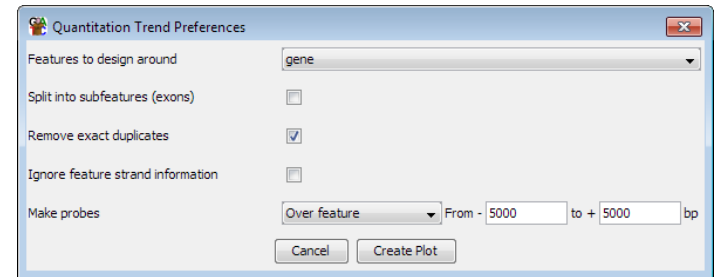

Plots > Beanplot

Comparisons

Plots > Scatter plot

Plot is interactive – mouse over a point for information
Double click on a point to move there in the chromosome view

Trend Plots

Plots > Quantitation Trend Plot

Clustering

Correlation based (per probe normalised)   Euclidean

Plots > Hierarchical Clusters