CAD Laboratory

Feb 12th, 2010

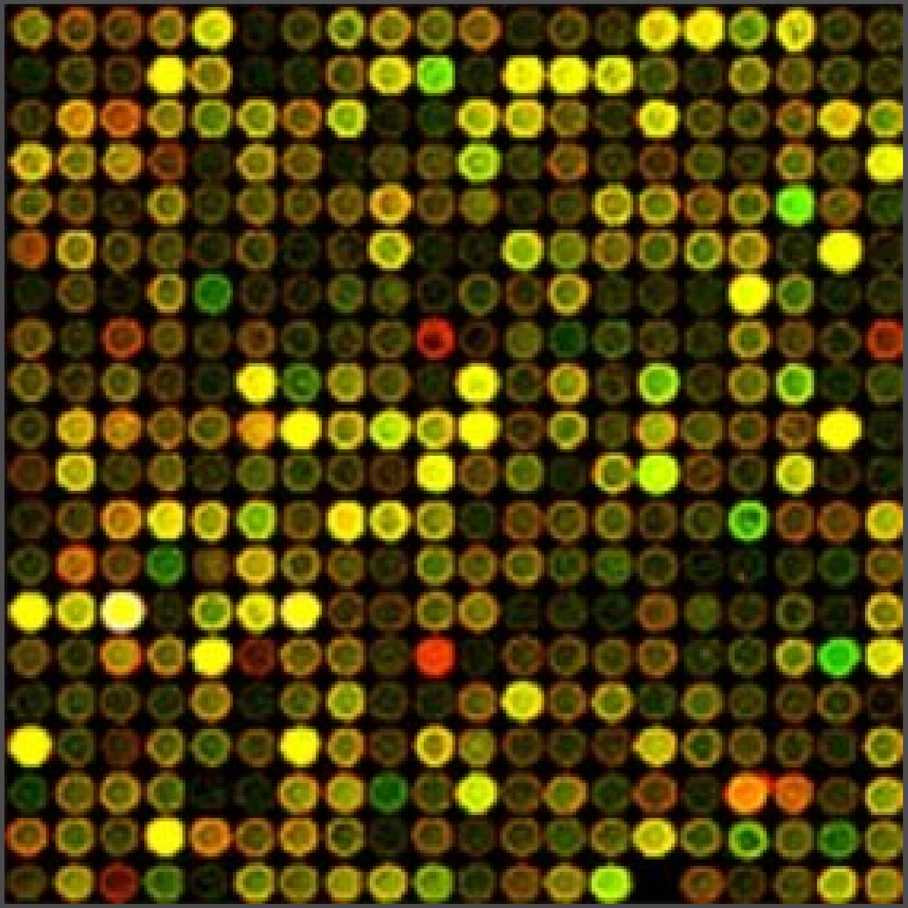
structural variation

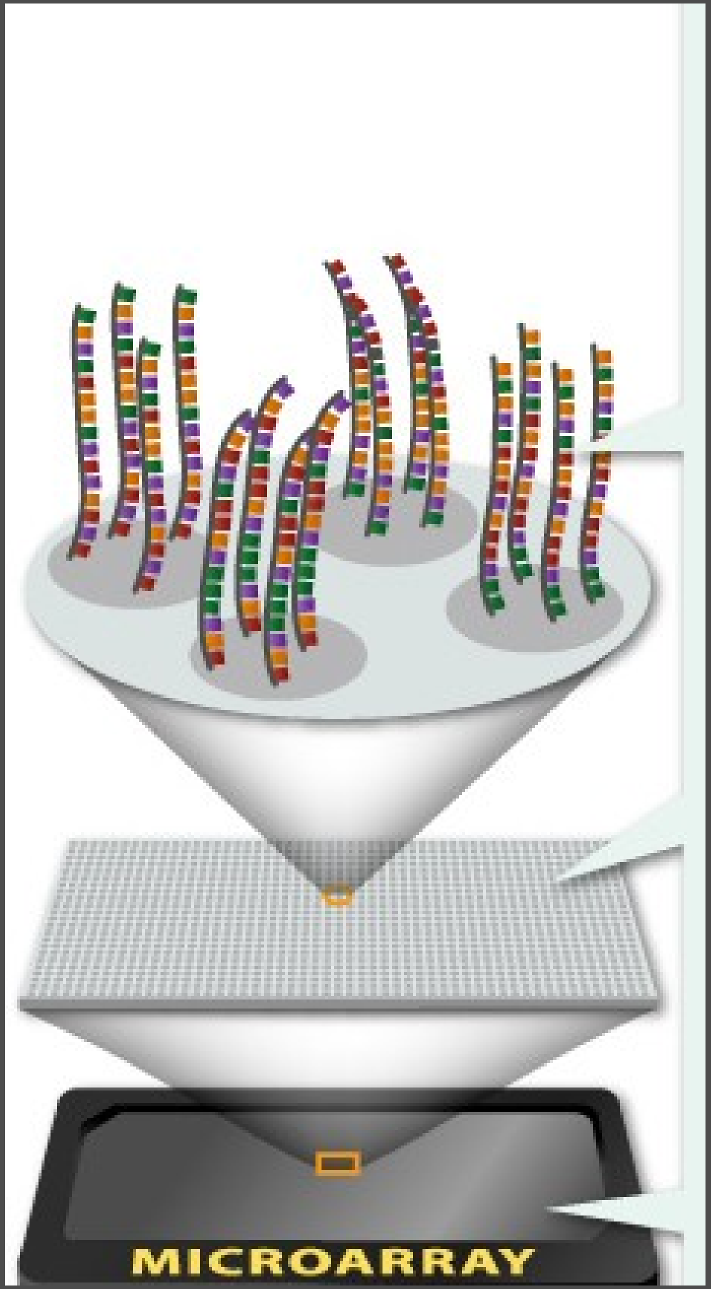
copy number

aCGH

oligonucleotide

chip




MICROARRAY

$\log_2$ ratio

$\log_2$ ratio

zero

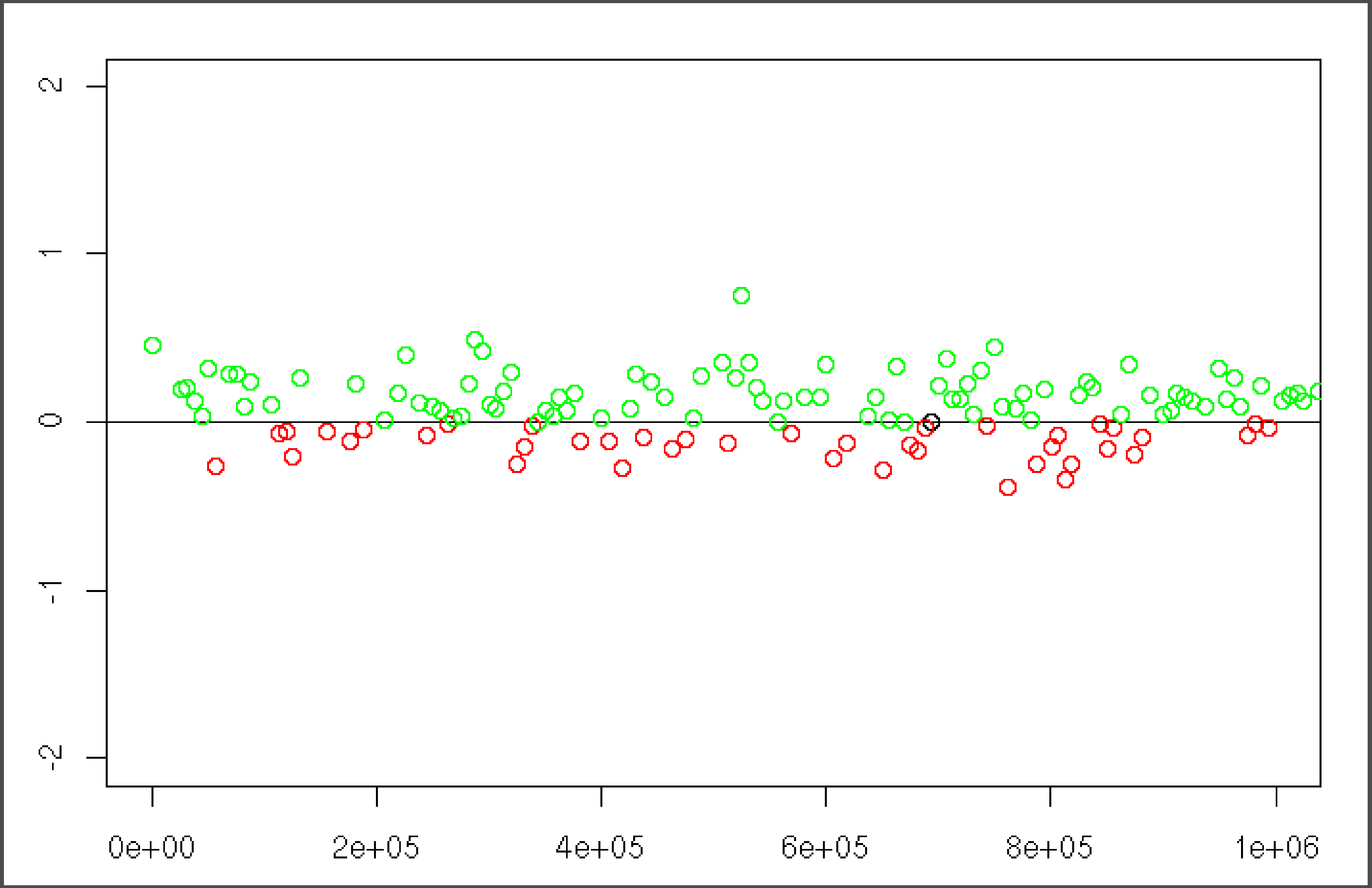
$\log_2$ ratio

positive

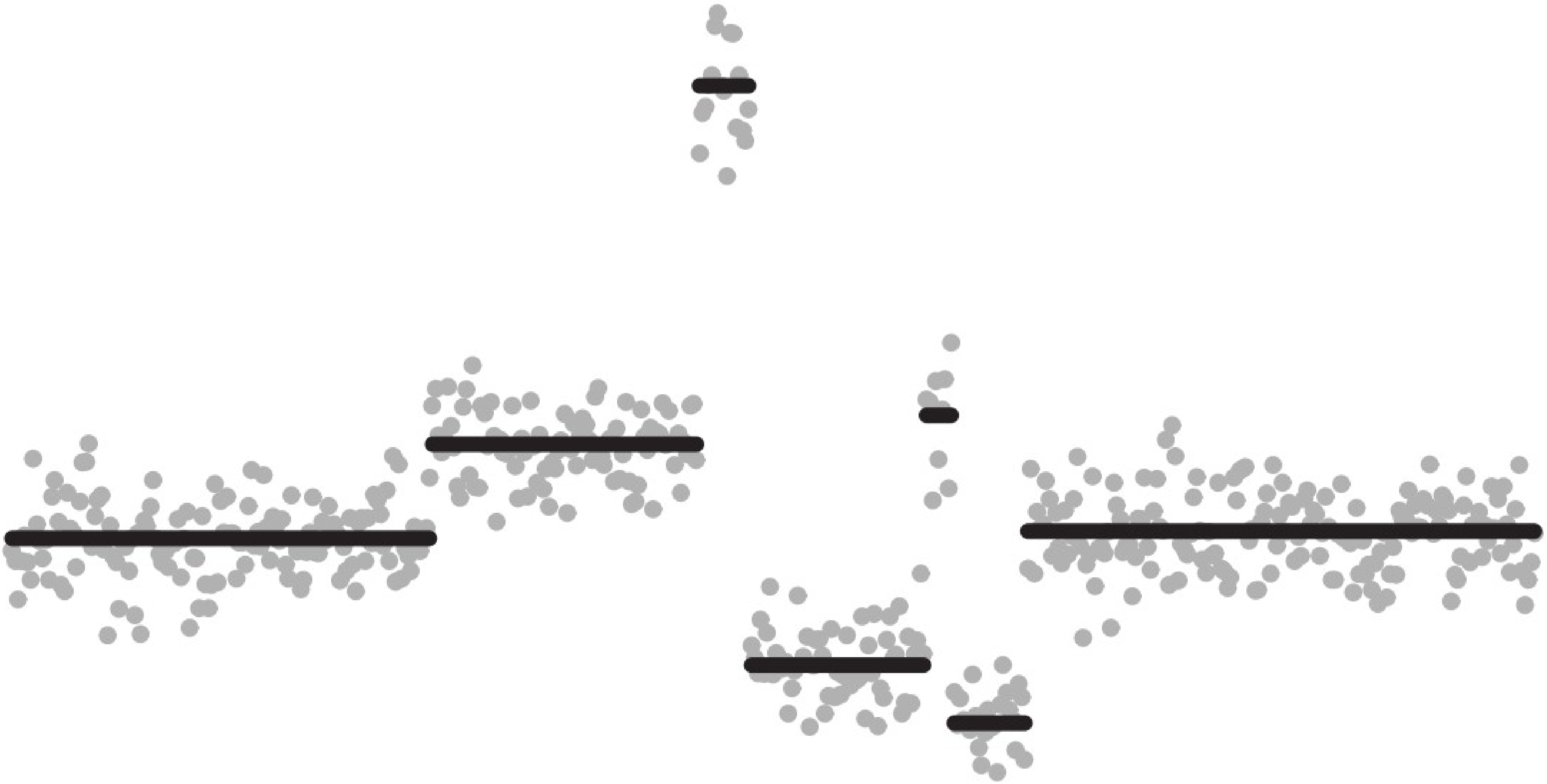
$\log_2$ ratio

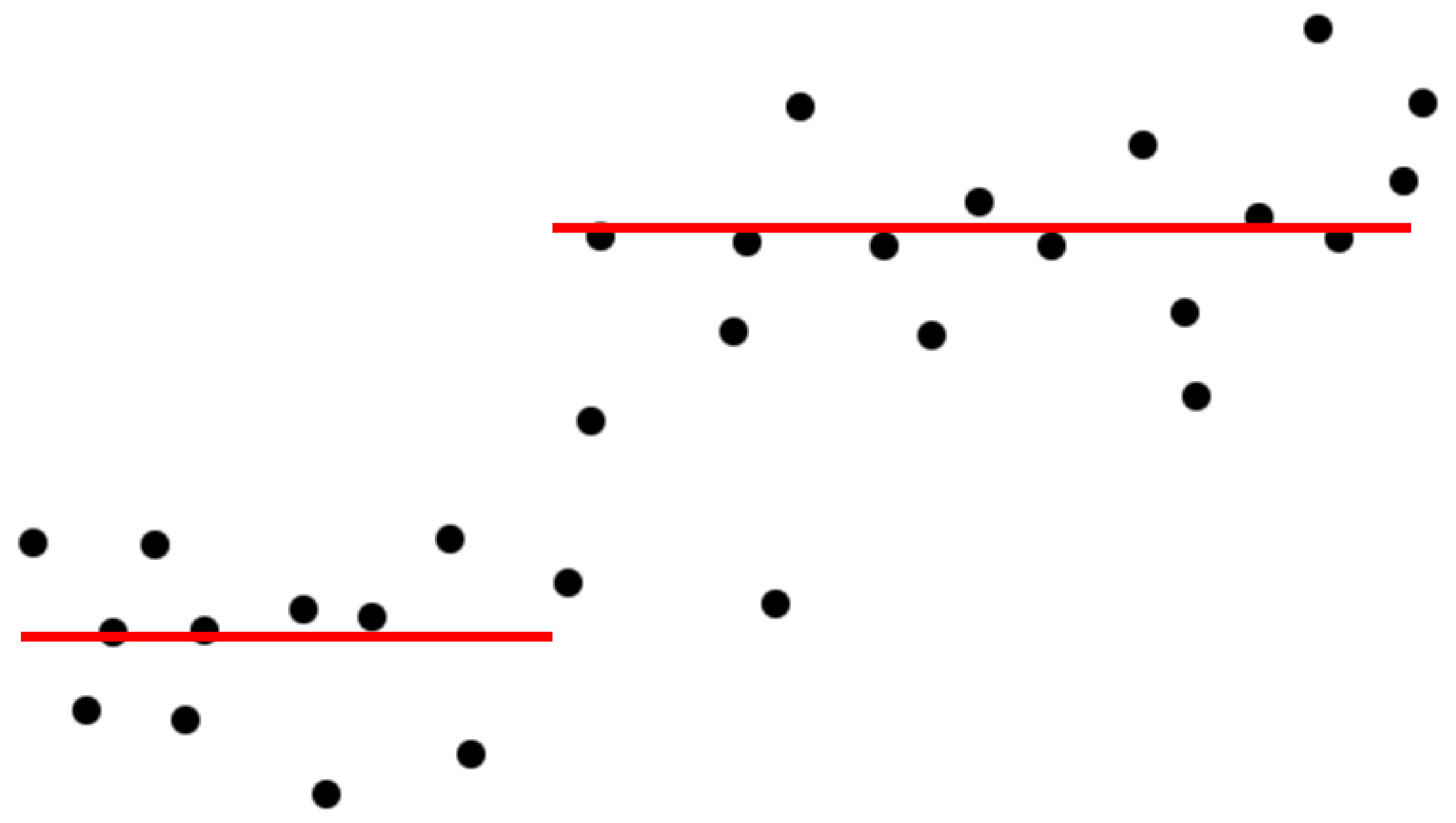
negative

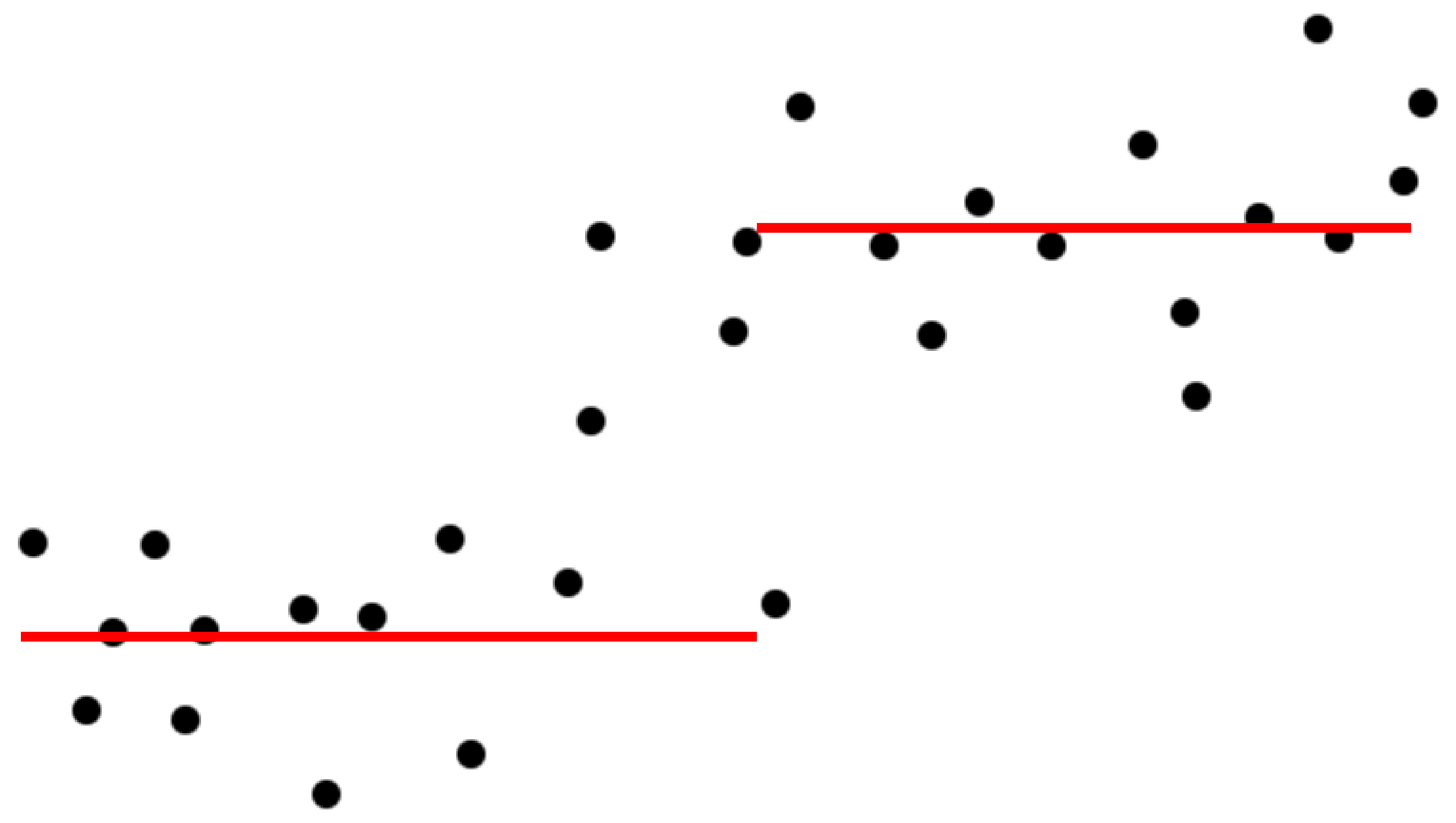
chr1:1-1,000,000
H1 cell line

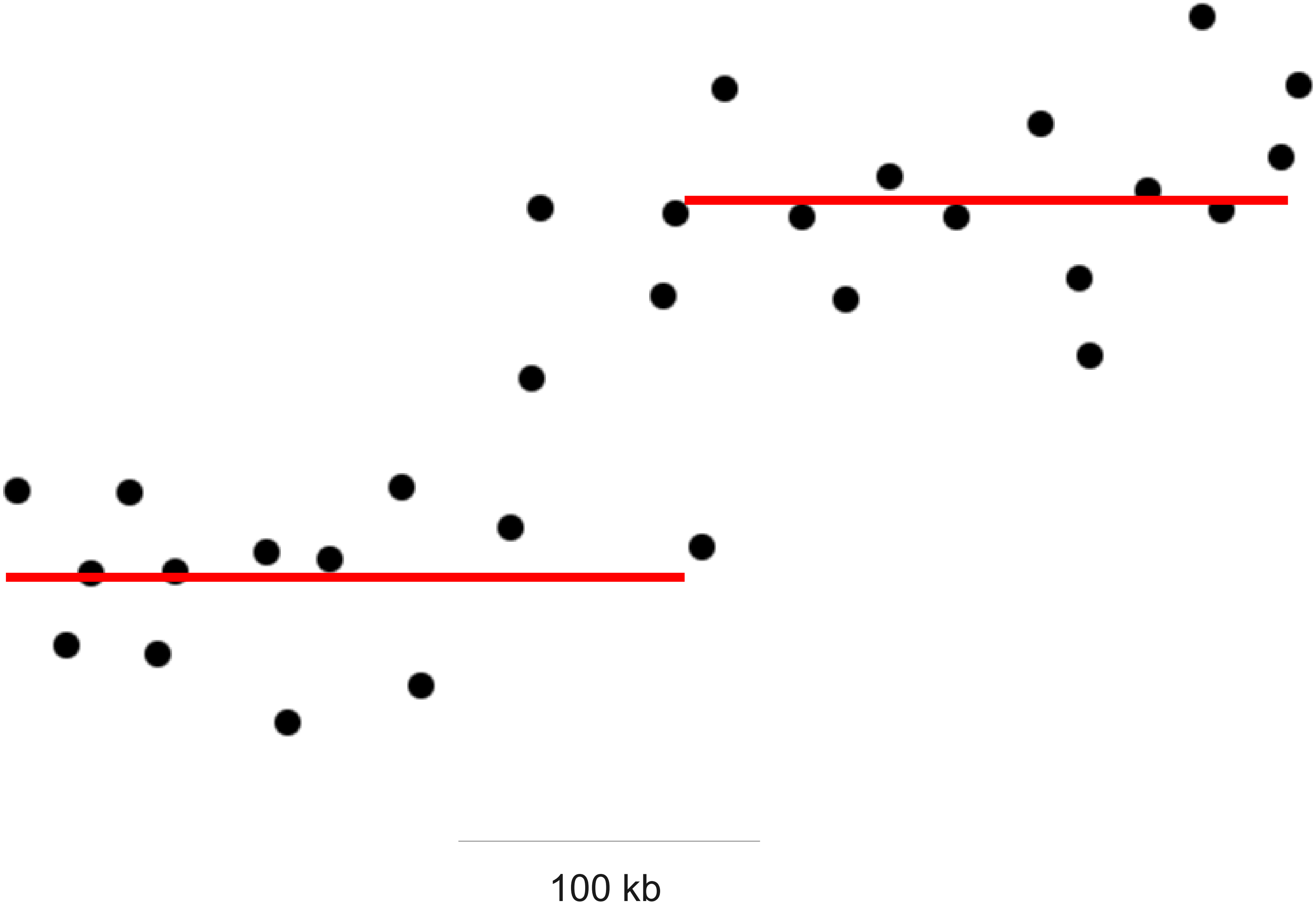


segmentation



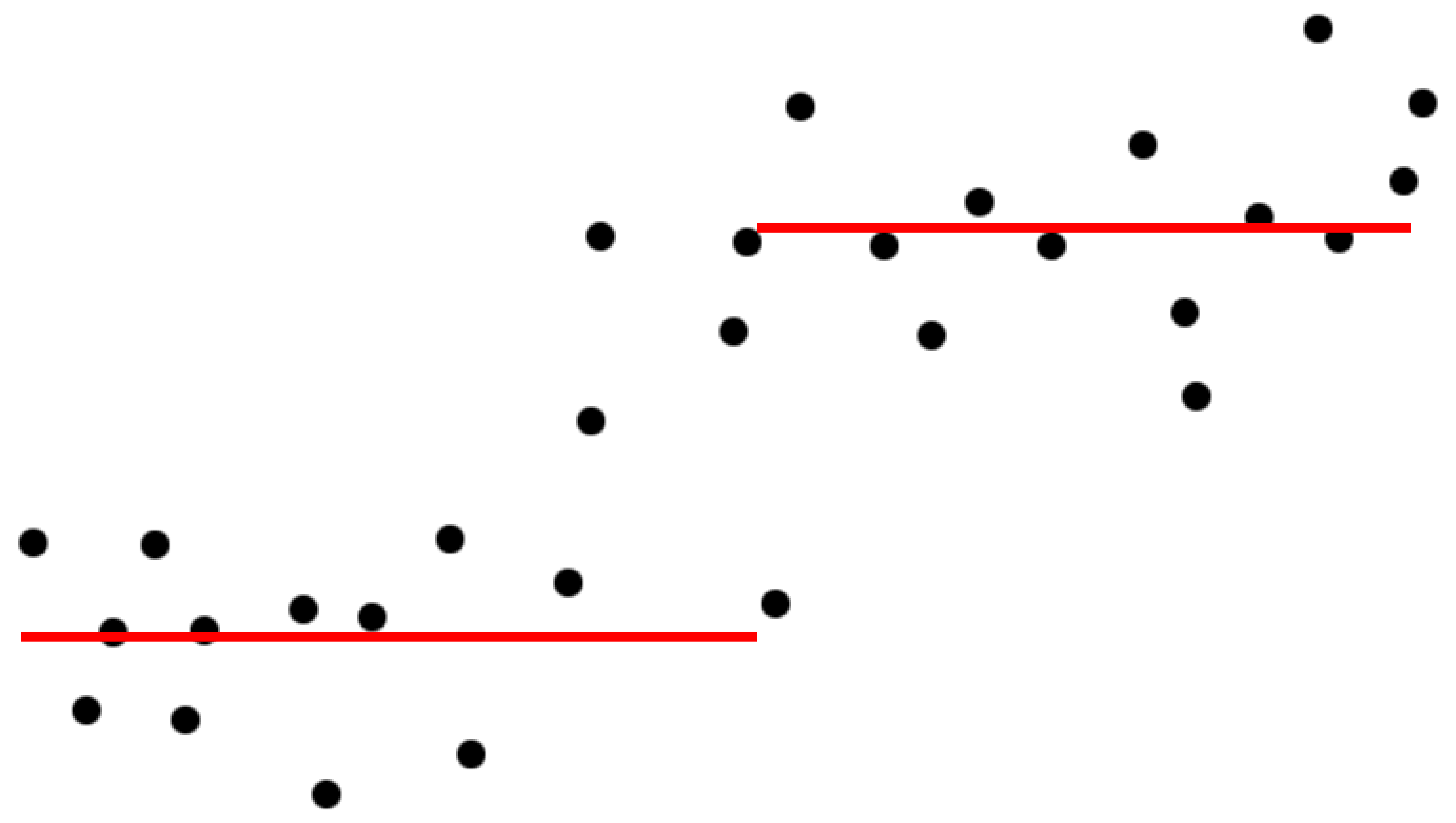




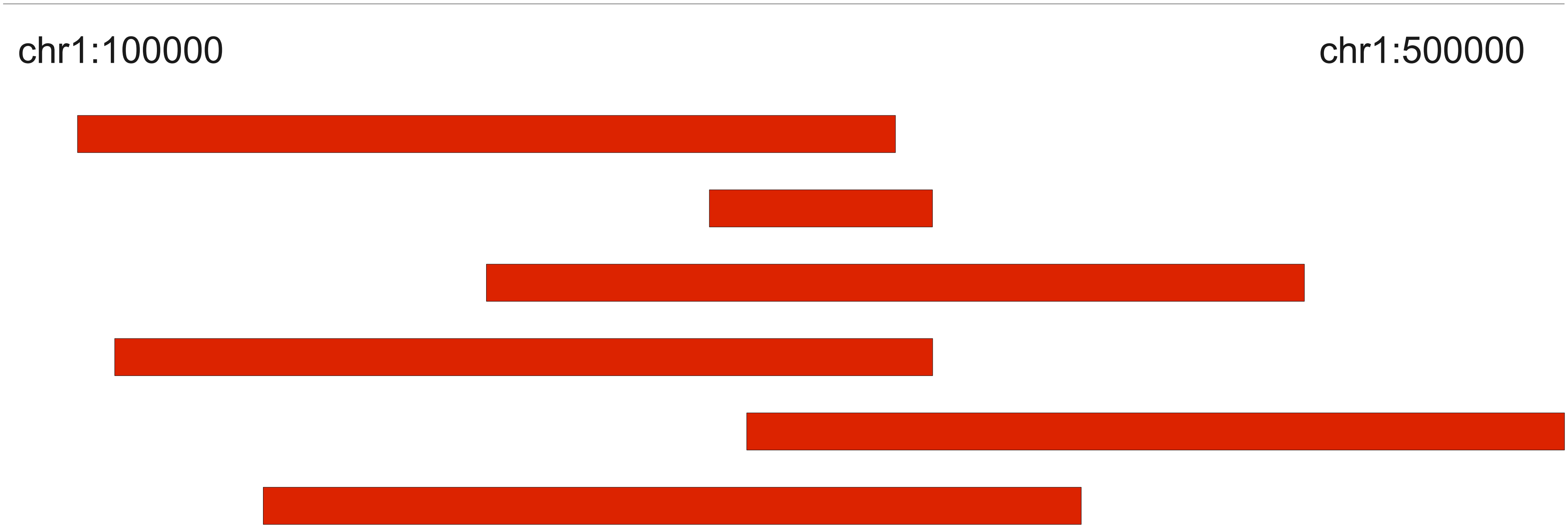

100 kb

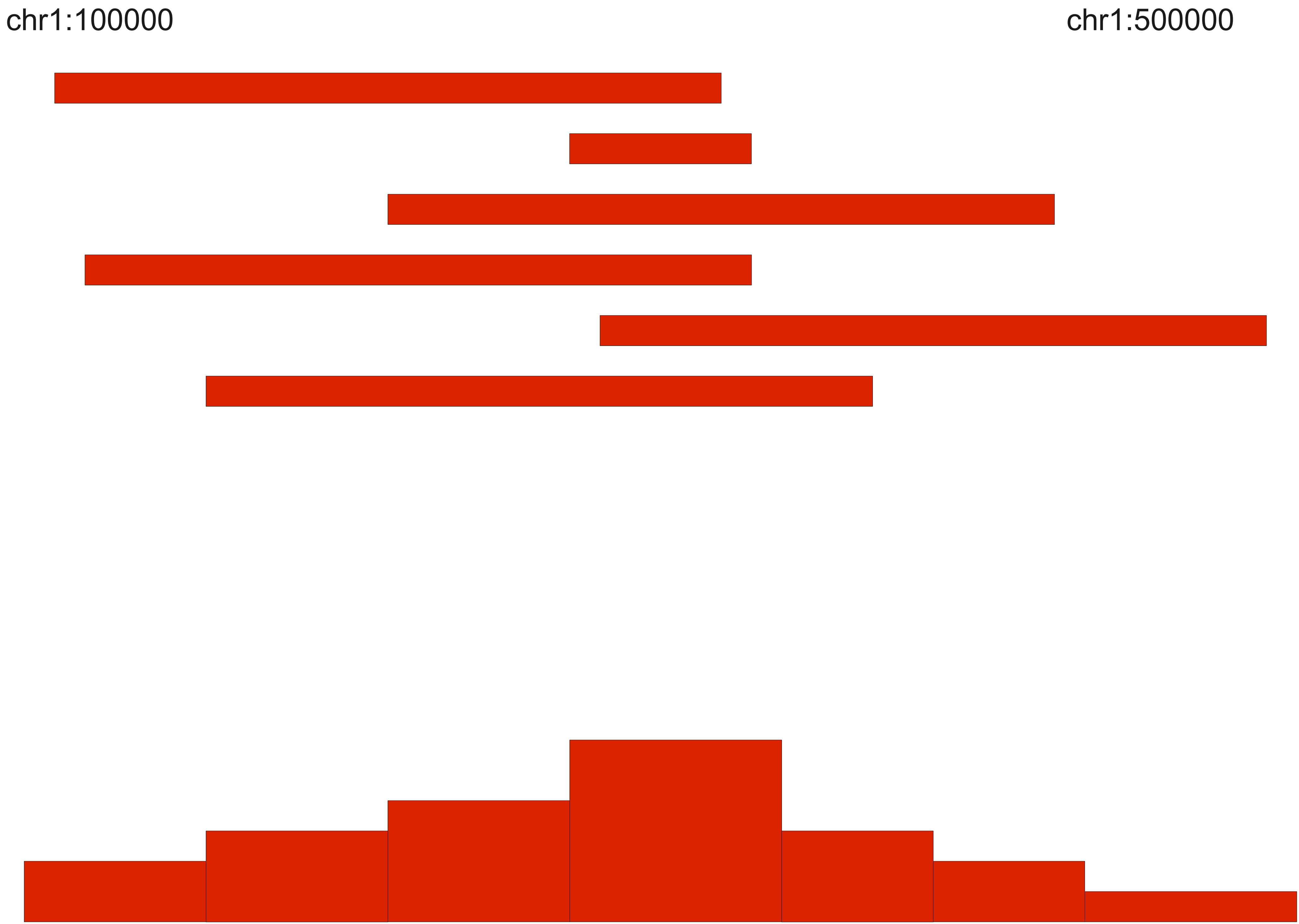
circular binary segmentation

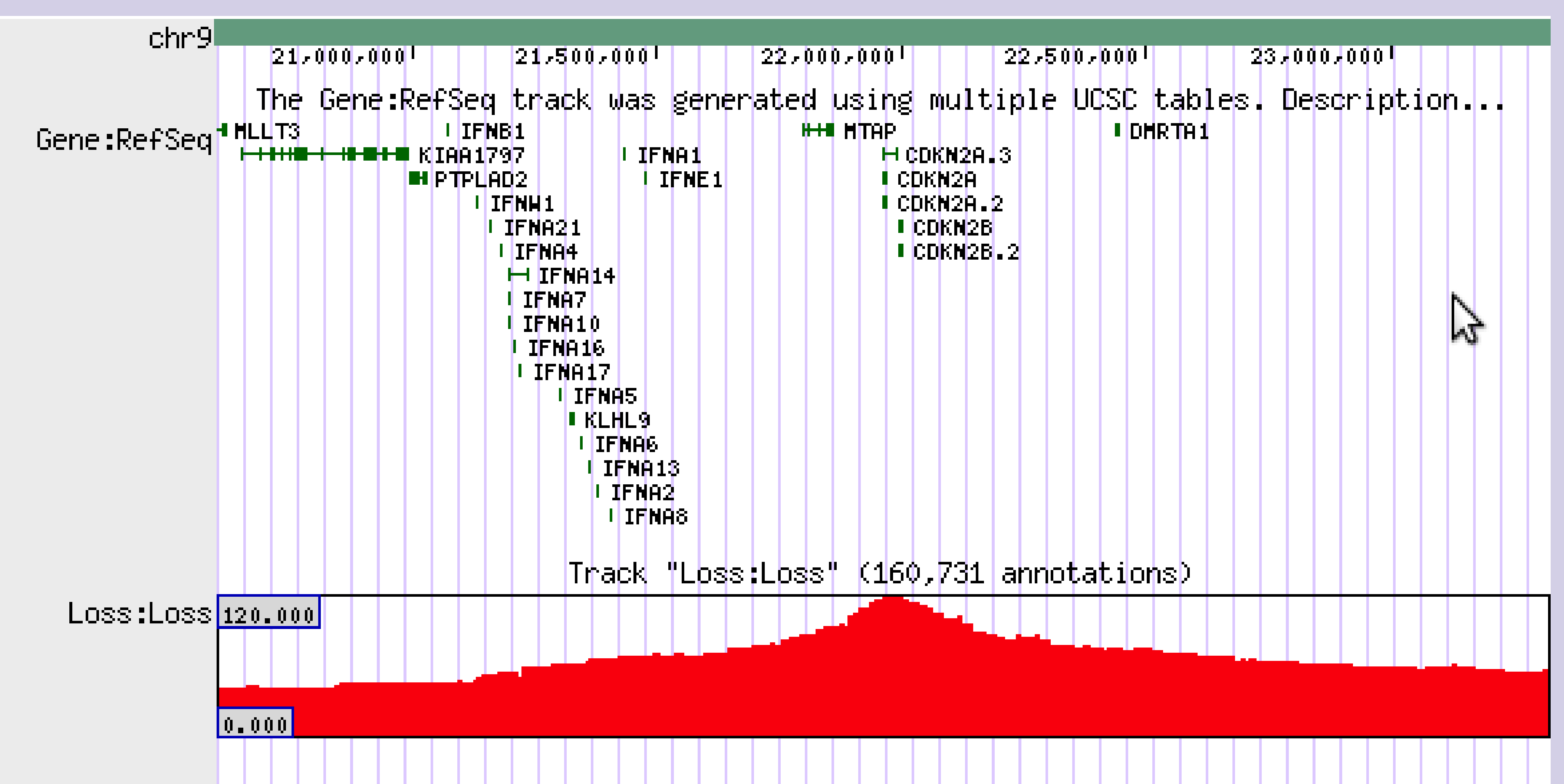
change-points



recurrence


chr1:100000
chr1:500000

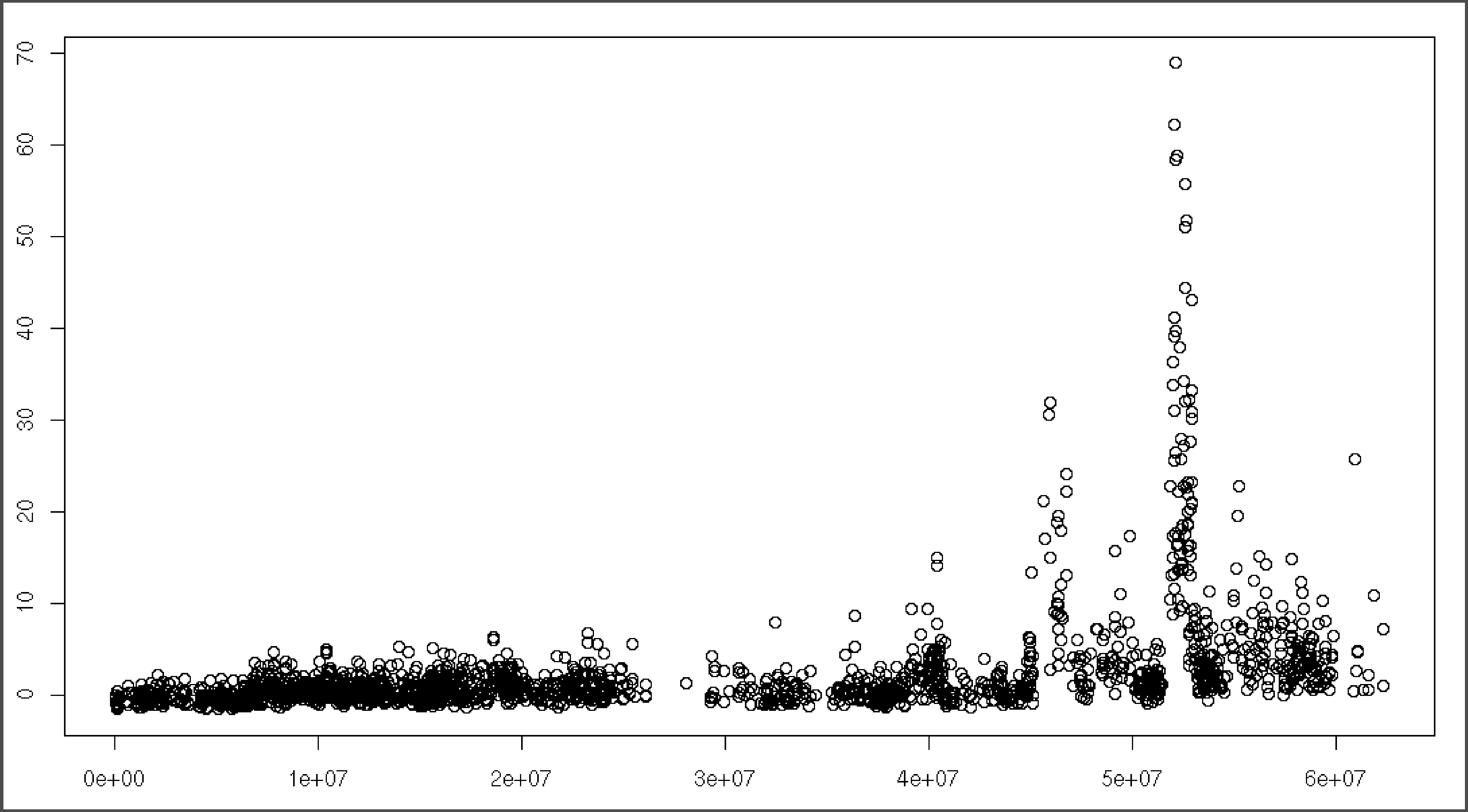

chr1:100000
chr1:500000



The Gene:RefSeq track was generated using multiple UCSC tables. Description...

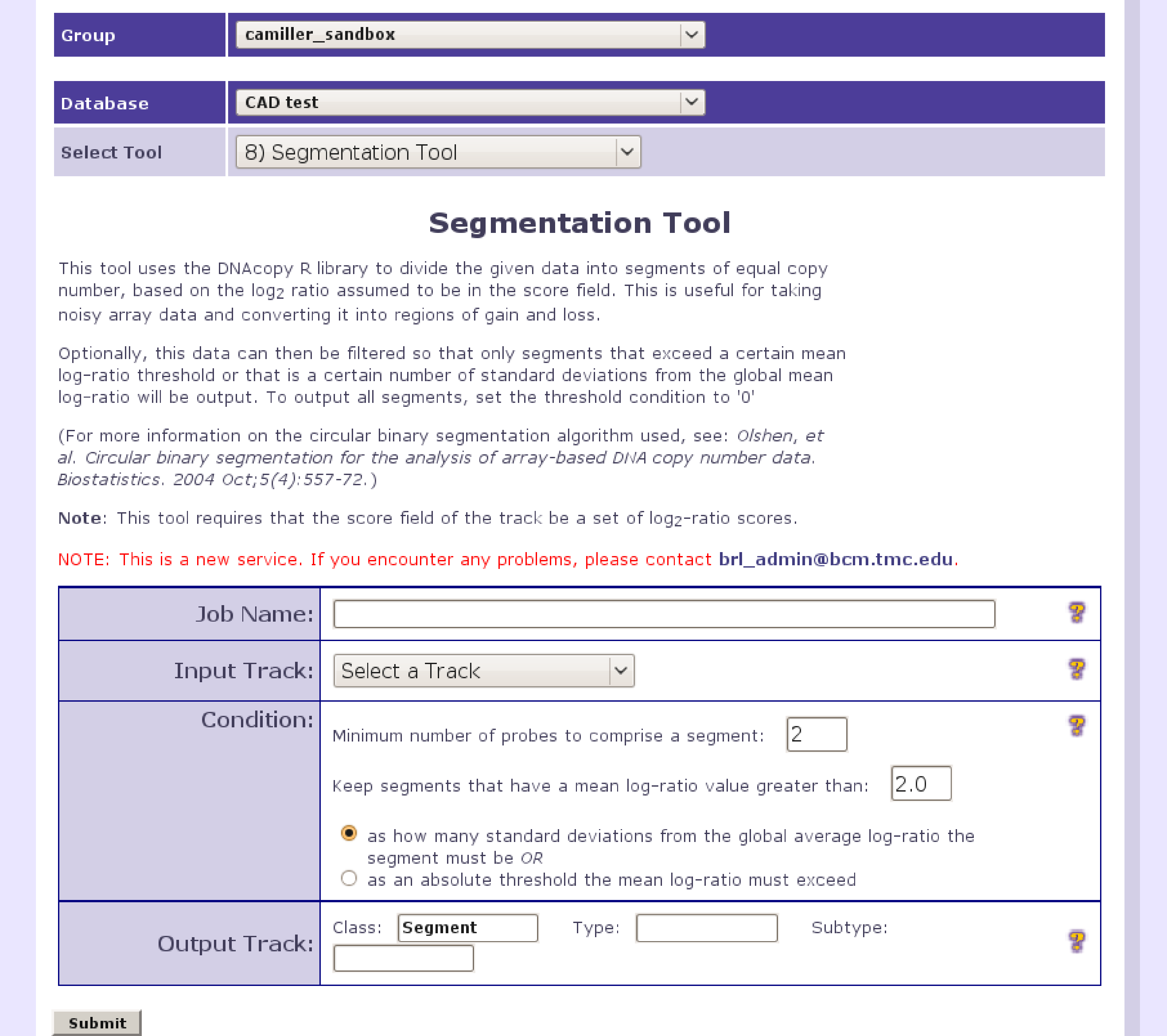
Track "Loss:Loss" (160,731 annotations)

chr20:all
MCF-7 breast cancer cell line



tools

segmentation tool



Group	camiller_sandbox
Database	CAD test
Select Tool	8) Segmentation Tool

Segmentation Tool

This tool uses the DNAcopy R library to divide the given data into segments of equal copy number, based on the $\log_2$ ratio assumed to be in the score field. This is useful for taking noisy array data and converting it into regions of gain and loss.

Optionally, this data can then be filtered so that only segments that exceed a certain mean log-ratio threshold or that is a certain number of standard deviations from the global mean log-ratio will be output. To output all segments, set the threshold condition to '0'

(For more information on the circular binary segmentation algorithm used, see: *Olshen, et al. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics. 2004 Oct;5(4):557-72.*)

Note: This tool requires that the score field of the track be a set of $\log_2$-ratio scores.

NOTE: This is a new service. If you encounter any problems, please contact **brl_admin@bcm.tmc.edu**.

Job Name:

Input Track: Select a Track

Condition:

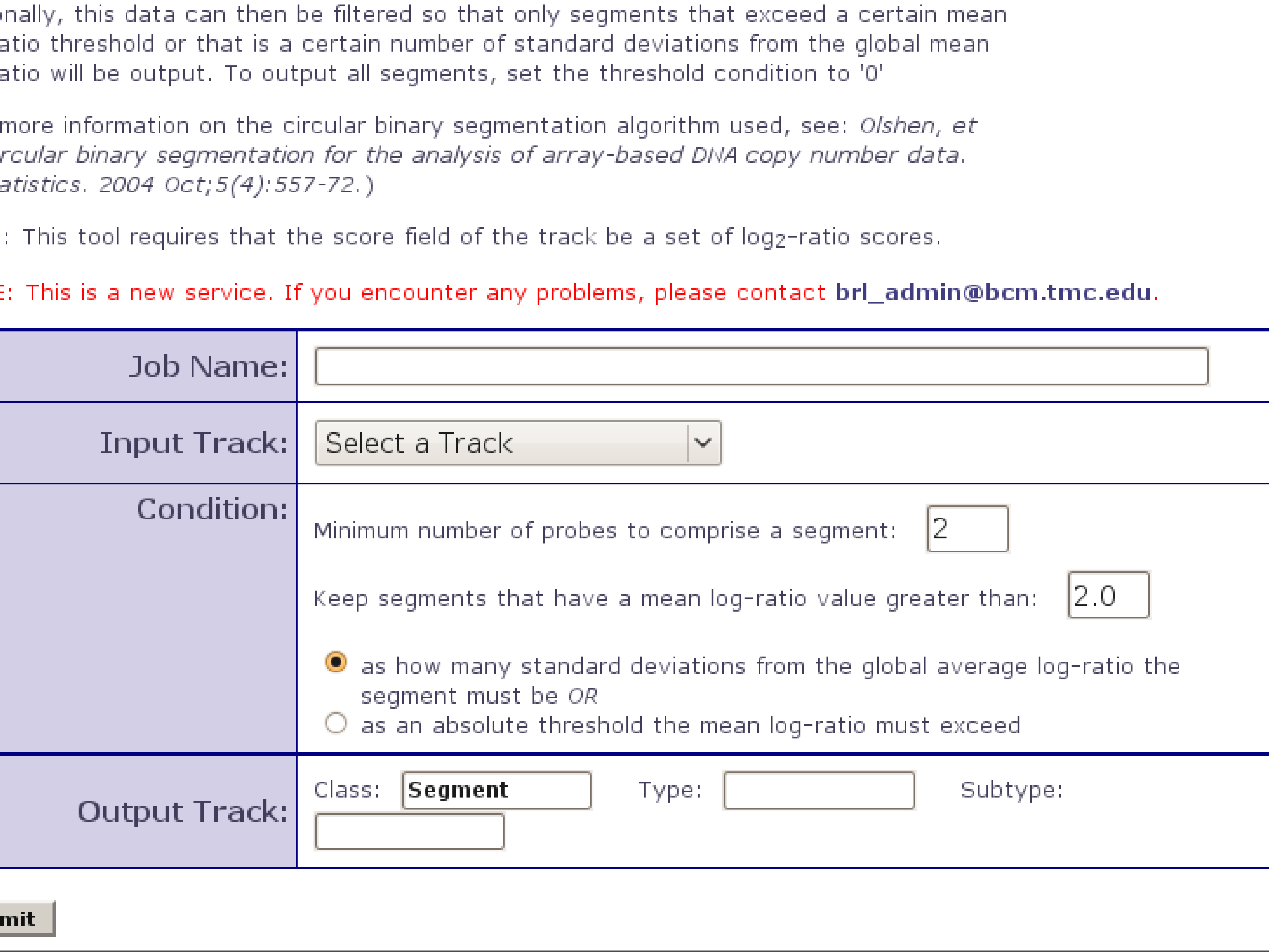
Minimum number of probes to comprise a segment: 2

Keep segments that have a mean log-ratio value greater than: 2.0

● as how many standard deviations from the global average log-ratio the segment must be *OR*

○ as an absolute threshold the mean log-ratio must exceed

Output Track: Class: **Segment** Type: Subtype:



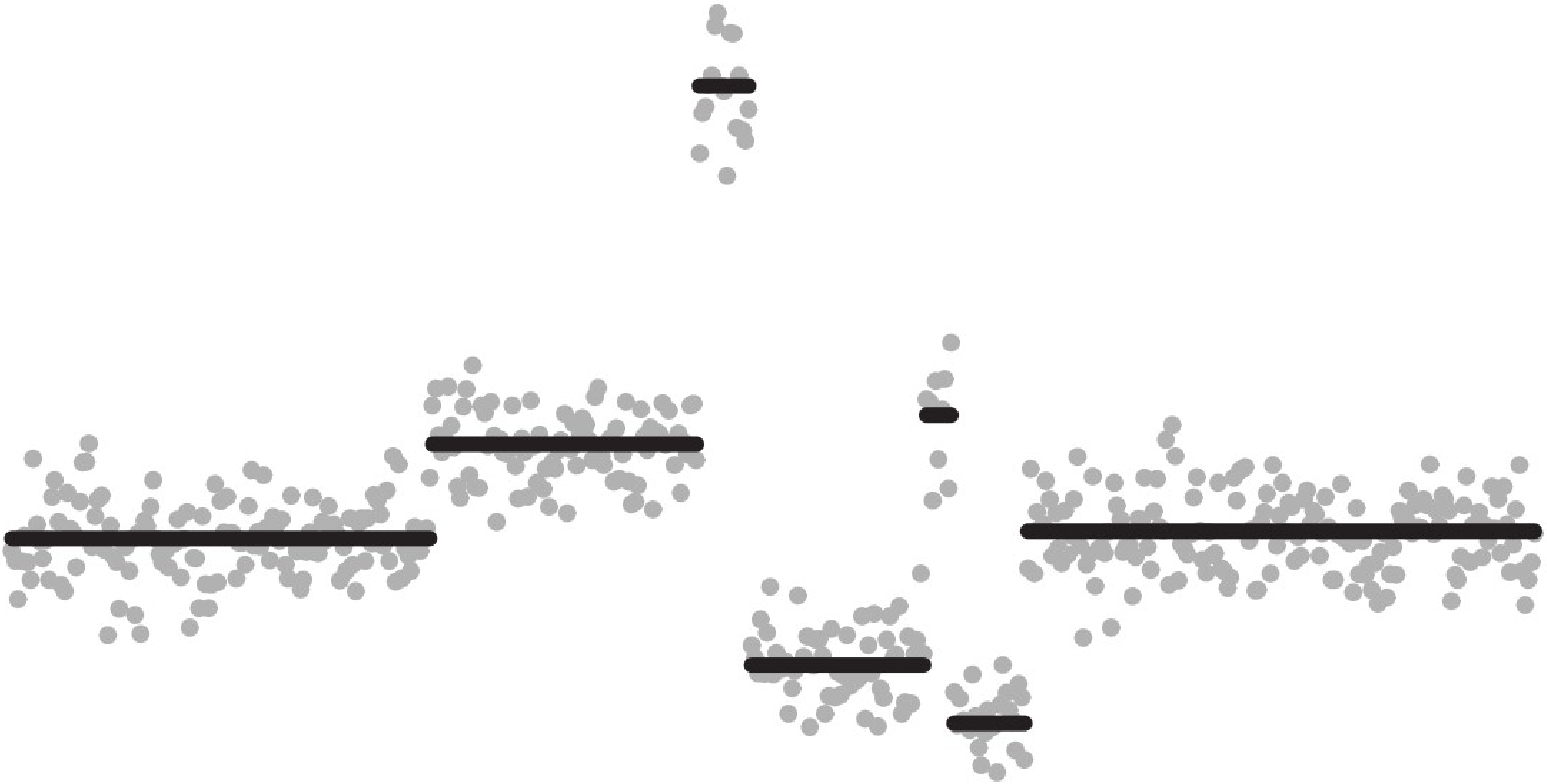
nally, this data can then be filtered so that only segments that exceed a certain mean atio threshold or that is a certain number of standard deviations from the global mean atio will be output. To output all segments, set the threshold condition to '0'

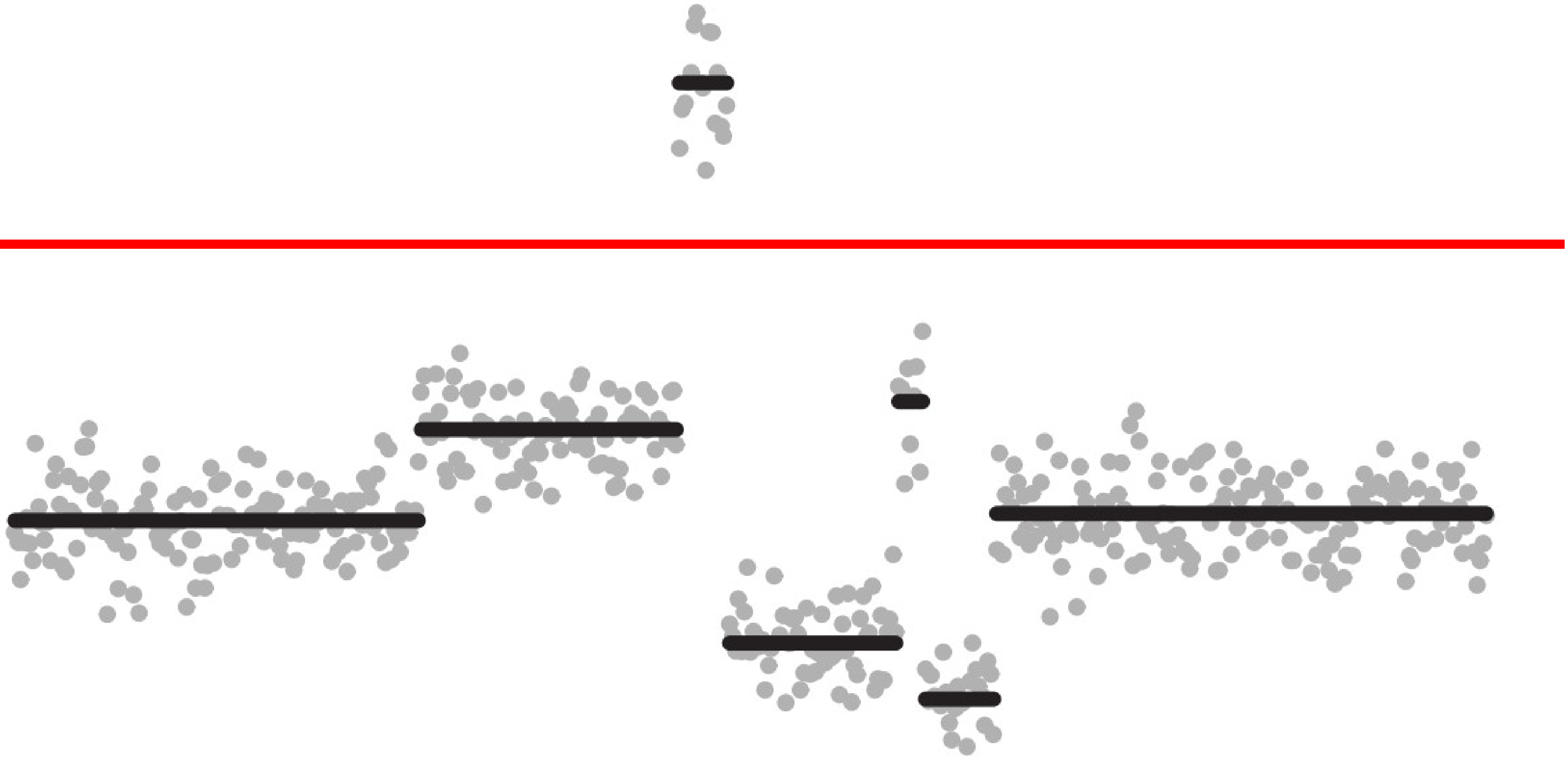
more information on the circular binary segmentation algorithm used, see: *Olshen, et rcular binary segmentation for the analysis of array-based DNA copy number data. atistics. 2004 Oct;5(4):557-72.)*

: This tool requires that the score field of the track be a set of log$_2$-ratio scores.

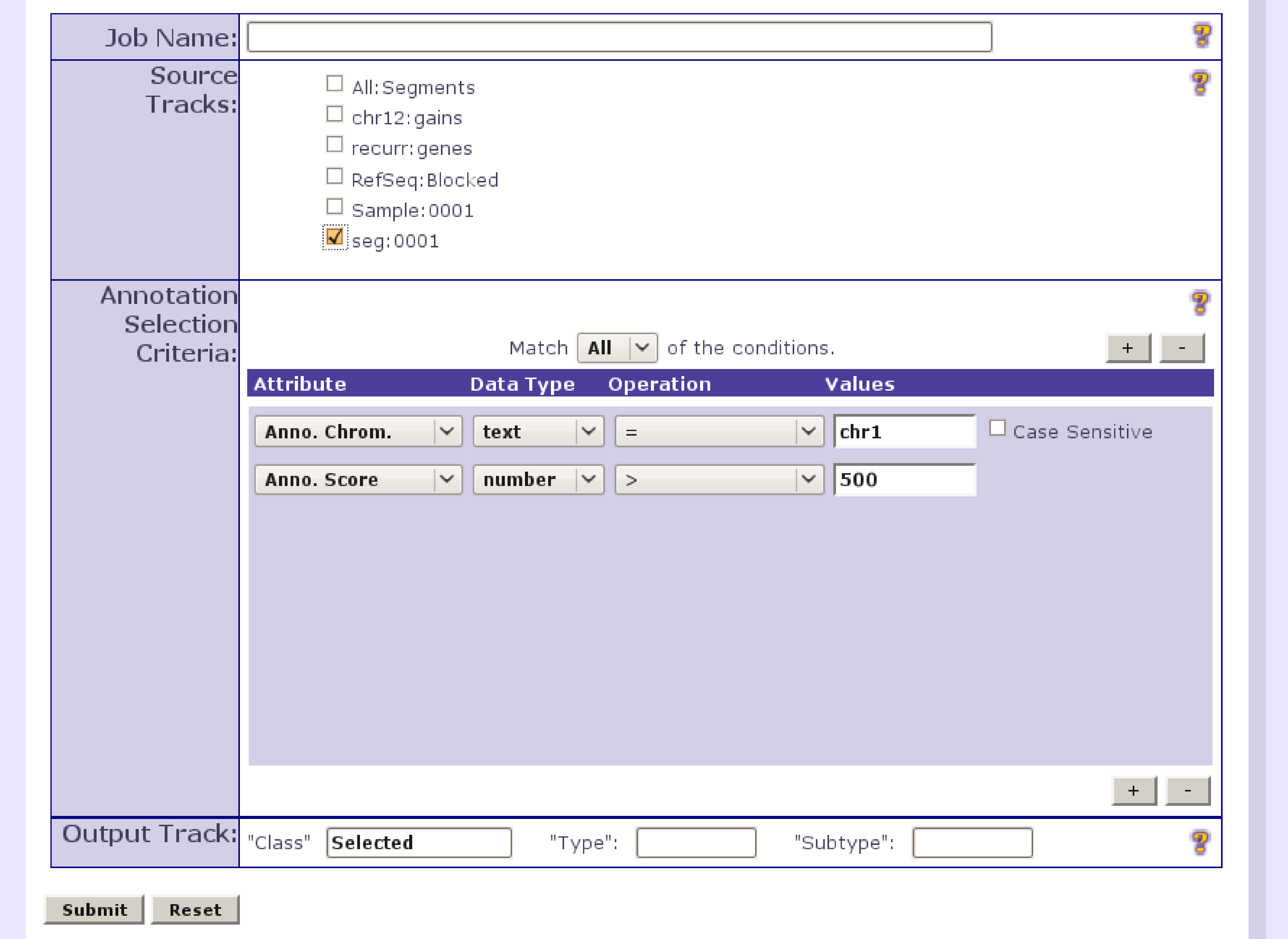
E: This is a new service. If you encounter any problems, please contact **brl_admin@bcm.tmc.edu**.

Job Name:	
Input Track:	Select a Track
Condition:	Minimum number of probes to comprise a segment: 2 Keep segments that have a mean log-ratio value greater than: 2.0 ● as how many standard deviations from the global average log-ratio the segment must be *OR* ○ as an absolute threshold the mean log-ratio must exceed
Output Track:	Class: **Segment** Type: Subtype:





annotation selector



Job Name:

Source Tracks:

☐ All:Segments
☐ chr12:gains
☐ recurr:genes
☐ RefSeq:Blocked
☐ Sample:0001
☑ seg:0001

Annotation Selection Criteria:

Match All of the conditions.

Attribute	Data Type	Operation	Values	
Anno. Chrom.	text	=	chr1	☐ Case Sensitive
Anno. Score	number	>	500	

Output Track: "Class" Selected "Type": "Subtype":

attribute lifter

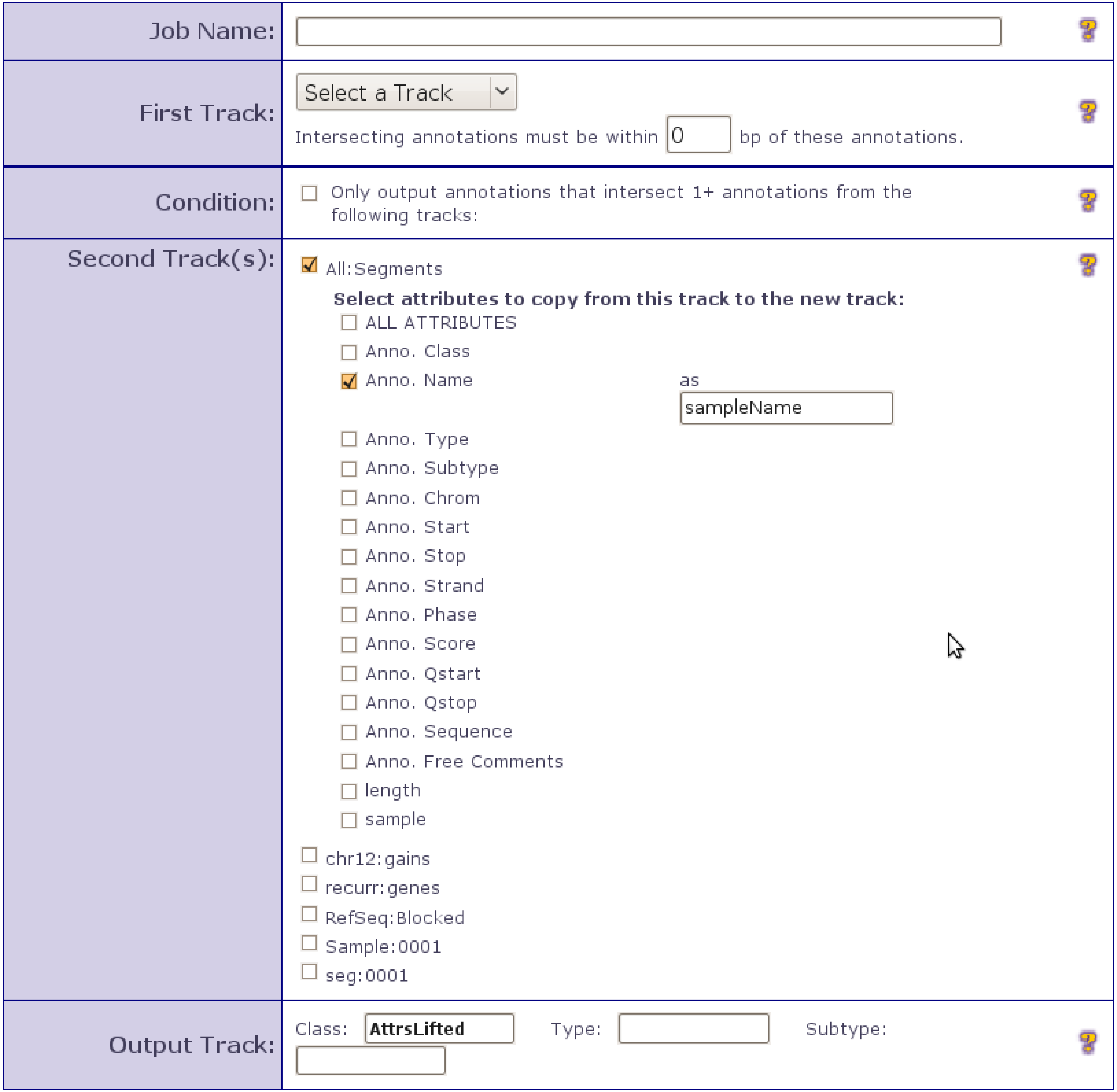
attribute-value pair

gender=male;

gender=male; sampleId=97A;

Track 1: Regions

Track 2: Genes



Job Name:

First Track: Select a Track

Intersecting annotations must be within 0 bp of these annotations.

Condition: ☐ Only output annotations that intersect 1+ annotations from the following tracks:

Second Track(s):

☑ All:Segments

Select attributes to copy from this track to the new track:

- ☐ ALL ATTRIBUTES
- ☐ Anno. Class
- ☑ Anno. Name as sampleName
- ☐ Anno. Type
- ☐ Anno. Subtype
- ☐ Anno. Chrom
- ☐ Anno. Start
- ☐ Anno. Stop
- ☐ Anno. Strand
- ☐ Anno. Phase
- ☐ Anno. Score
- ☐ Anno. Qstart
- ☐ Anno. Qstop
- ☐ Anno. Sequence
- ☐ Anno. Free Comments
- ☐ length
- ☐ sample

☐ chr12:gains

☐ recurr:genes

☐ RefSeq:Blocked

☐ Sample:0001

☐ seg:0001

Output Track: Class: AttrsLifted Type: Subtype:

intGenes=ERBB2, MET, KRAS;

tabular view

assignment

glioblastoma

1 year

TGCA

186 tumors

matched normals

structural variants

data munging

segment

recurrence

genes

expression

correlation

Feb 24th

chrisamiller@gmail.com