Find Peaks with MACS Element

Performs peak calling for ChIP-Seq data.

Parameters in GUI

Parameter	Description	Default value
Output directory	Directory to save MACS output files.	
Name	The name string of the experiment. MACS will use this string NAME to create output files like 'NAME_peaks.xls', 'NAME_negative_peaks.xls', 'NAME_peaks.bed', 'NAME_summits.bed', 'NAME_model.r' and so on. So please avoid any confliction between these filenames and your existing files (--name).	
Wiggle output	If this flag is on, MACS will store the fragment pileup in wiggle format for the whole genome data instead of for every chromosomes (--wig) (--single-profile).	hg19
Wiggle space	By default, the resolution for saving wiggle files is 10 bps,i.e., MACS will save the raw tag count every 10 bps. You can change it along with '--wig' option (--space).	3000
Genome size (Mbp)	Homo sapience - 2700 Mbp Mus musculus - 1870 Mbp Caenorhabditis elegans - 90 Mbp Drosophila melanogaster - 120 Mbp It's the mappable genome size or effective genome size which is defined as the genome size which can be sequenced. Because of the repetitive features on the chromosomes, the actual mappable genome size will be smaller than the original size, about 90% or 70% of the genome size (--gsize).	50
P-value	P-value cutoff. Default is 0.00001, for looser results, try 0.001 instead (--pvalue).	3000
Tag size (optional)	Length of reads. Determined from first 10 reads if not specified (input 0) (--tsize).	5000
Keep duplicates	It controls the MACS behavior towards duplicate tags at the exact same location -- the same coordination and the same strand. The default auto option makes MACS calculate the maximum tags at the exact same location based on binomal distribution using 1e-5 as pvalue cutoff; and the all option keeps every tags. If an integer is given, at most this number of tags will be kept at the same location (--keep-dup).	3000
Use model	Whether or not to use MACS paired peaks model (--nomodel).	
Model fold	Select the regions within MFOLD range of high-confidence enrichment ratio against. Model fold is available when Use model is true, which is the foldchange to chose paired peaks to build paired peaks model. Users need to set a lower (smaller) and upper(larger) number for fold change so that MACS will only use the peaks within these foldchange range to build model (--mfold).	
Shift size	An arbitrary shift value used as a half of the fragment size when model is not built. Shift size is available when Use model is false, which will represent the HALF of the fragment size of your sample. If your sonication and size selection size is 300 bps, after you trim out nearly 100 bps adapters, the fragment size is about 200 bps, so you can specify 100 here (--shiftsize).	
Band width	The band width which is used to scan the genome for model building. You can set this parameter as the sonication fragment size expected from wet experiment. Used only while building the shifting model (--bw).	
Use lambda	Whether to use local lambda model which can use the local bias at peak regions to throw out false positives (--nolambda).	
Small nearby region	The small nearby region in basepairs to calculate dynamic lambda. This is used to capture the bias near the peak summit region. Invalid if there is no control data (--slocal).	
Large nearby region	The large nearby region in basepairs to calculate dynamic lambda. This is used to capture the surround bias (--llocal).	
Auto bimodal	Whether turn on the auto pair model process.If set, when MACS failed to build paired model, it will use the nomodelsettings, the Shift size parameter to shift and extend each tags (--on-auto).	
Scale to large	When set, scale the small sample up to the bigger sample.By default, the bigger dataset will be scaled down towards the smaller dataset,which will lead to smaller p/qvalues and more specific results.Keep in mind that scaling down will bring down background noise more (--to-large).	

Parameters in Workflow File

Type: macs-id

Parameter	Parameter in the GUI	Type
output-dir	**Output directory**	*string*
file-names	**Name**	*string*
wiggle-output	**Wiggle output**	*boolean*
wiggle-space	**Wiggle space**	*numeric*
genome-size	**Genome size (Mbp)**	*numeric*
p-value	**P-value**	*numeric*
tag-size	**Tag size (optional)**	*numeric*
keep-duplicates	**Keep duplicates**	*string*
use-model	**Use model**	*boolean*
model-fold	**Model fold**	*string*
shift-size	**Shift size**	*numeric*
band-width	**Band width**	*numeric*
use-lambda	**Use lambda**	*boolean*
small-nearby	**Small nearby region**	*numeric*
large-nearby	**Large nearby region**	*numeric*
auto_bimodal	**Auto bimodal**	*boolean*
scale_large	**Scale to large**	*boolean*

Input/Output Ports

The element has 1 *input port*:

Name in GUI: MACS data

Name in Workflow File: in-data

Slots:

Slot In GUI	Slot in Workflow File	Type
Treatment features	**_treatment-ann**	*ann-table-list*
Control features	**control-ann**	*ann-table-list*

And 1 *output port*:

Name in GUI: MACS output data

Name in Workflow File: out-data

Slots:

Slot In GUI	Slot in Workflow File	Type
Peak regions	**peak-regions**	*ann-table-list*
Peak summits	**peak-summits**	*ann-table-list*
Treatment fragments pileup	**wiggle-treat**	*string*