

# Manual setting of k-mer coverage peaks

Afiahayati, Sato K, Namiki T, Hachiya T, Tanaka H, Sakakibara Y.

## Abstract

Contrary to single-genome assemblers (e.g., Velvet and SOAPdenovo), MetaVelvet assumes metagenomics settings. Accordingly, *k*-mer coverage histogram might be multi-modal rather than uni-modal. Please note that the coverage peak parameters can largely affect MetaVelvet assembling results. Although simple and automatic peak detection algorithm is implemented in **meta-velvetg**, we strongly recommend manual inspection of *k*-mer coverage peaks and manual setting of the coverage peak parameters. For manually setting the coverage peak parameters, please execute the following procedures.

**Citation:** Afiahayati, Sato K, Namiki T, Hachiya T, Tanaka H, Sakakibara Y. Manual setting of k-mer coverage peaks. **protocols.io**

dx.doi.org/10.17504/protocols.io.d7k9kv

**Published:** 14 Jan 2016

## Guidelines

For installation and getting started protocol click [here](#).

## Frequently Asked Questions

- **Q: meta-velveth** is not generated in the new version ( $\geq 1.1.01$ ). Is it problem?  
**A:** This is not problem. The usage of MetaVelvet is changed when the version 1.1.01 is released, and the new version does not include **meta-velveth**. Instead, please use **velveth**, **velvetg**, and **meta-velvetg**.
- **Q:** When only one coverage peak is detected (or manually input), is there any difference between MetaVelvet and Velvet algorithms?  
**A:** There is no substantial difference. In such cases, **meta-velvetg** moves to "single-peak mode" and graph splitting functions in **meta-velvetg** is not called. Instead of graph splitting functions, standard velvet functions are called in such cases.
- **Q:** What's the difference in working procedures between **velvetg**, **meta-velvetg** ( $\leq 0.3.1$ ), and **meta-velvetg** ( $\geq 1.1.01$ )?  
**A:** The following is the working procedure of **velvetg**, **meta-velvetg** ( $\leq 0.3.1$ ), and **meta-velvetg** ( $\geq 1.1.01$ ):

### **velvetg & meta-velvetg ( $\leq 0.3.1$ ) :**

- Load Sequences & Roadmaps file
- > Generate PreGraph file
- > Generate Graph or Graph2 file

-> Generate contigs.fa and LastGraph

**meta-velvetg (>= 1.1.01):**

Load Sequences & Roadmaps & Graph2 file

-> Generate meta-velvetg.contigs.fa and meta-velvetg.LastGraph

- **Q:** Is version compatibility between Velvet and MetaVelvet fully tested?  
**A:** Version compatibility between Velvet-1.0.06 and MetaVelvet-1.1.01 is fully tested.

## Troubleshooting

**Trouble:** When drawing *k*-mer coverage histogram (as in the ["Advanced topics 1"](#) section), the following warning messages is appeared:

```
> weighted.hist(data$shot1_cov,data$lgth,breaks=seq(0,200,1))
Warning messages:
1: In min(x, na.rm = na.rm) :
no non-missing arguments to min; returning Inf
2: In max(x, na.rm = na.rm) :
no non-missing arguments to max; returning -Inf
3: In weighted.hist(data$shot1_cov, data$lgth, breaks = seq(0, 200, :
Areas will not relate to frequencies
```

**Solution:** This warning (error) is caused by "Inf" values in the Graph2 node stats. Accordingly, by removing "Inf" values from the Graph2 stats, the error is resolved:

```
$ head -n 1 meta-velvetg.Graph2-stats.txt \
> meta-velvetg.Graph2-stats.rmInf.txt
$ perl -ne '{print $_ unless /Inf/;}' \
meta-velvetg.Graph2-stats.txt \
>> meta-velvetg.Graph2-stats.rmInf.txt
$ R
> library(plotrix)
> data = read.table("meta-velvetg.Graph2-stats.rmInf.txt", header=TRUE)
> weighted.hist(data$shot1_cov,data$lgth,breaks=seq(0,200,1))
```

## Protocol

### Step 1.

Execute velveth, velvetg, and meta-velvetg as in the "Getting started" protocol. Please check that "out-dir/meta-velvetg.Graph2-stats.txt" is created.

### Step 2.

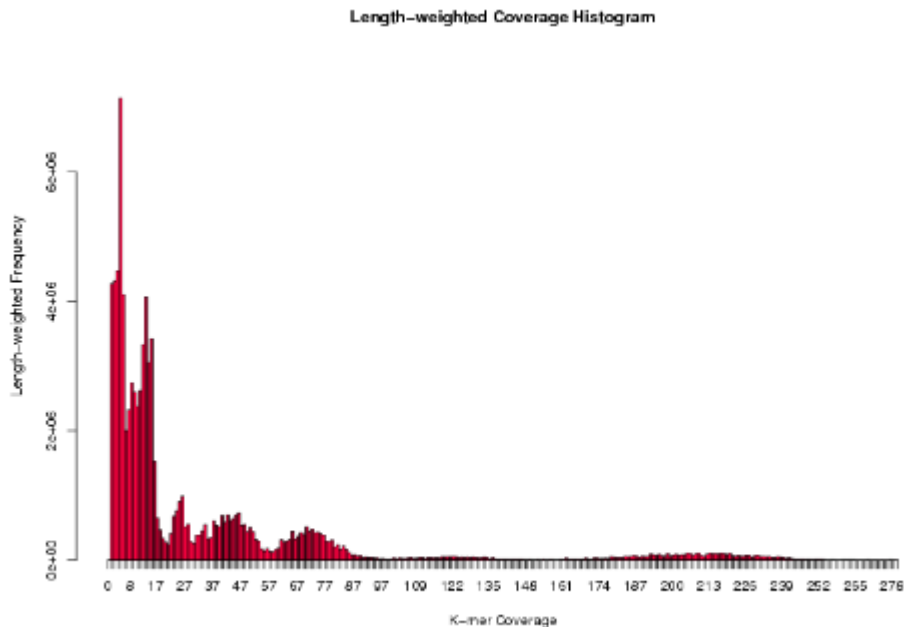
Draw a *k*-mer coverage histogram and manually determine the *k*-mer coverage peaks.

cmd **COMMAND**

```
~$ R
(R) > install.packages("plotrix")
(R) > library(plotrix)
(R) > data = read.table("out-dir/meta-velvetg.Graph2-stats.txt", header=TRUE)
(R) > weighted.hist(data$short1_cov, data$lgth, breaks=seq(0, 200, by=1))
```

### Step 3.

Please determine the coverage peak values from the histogram.



### 📌 NOTES

**Bonnie Hurwitz** 16 Nov 2015

For example, from the above histogram, seven coverage peaks are observed (around 210x, 120x, 70x, 45x, 23x, 12x, and 6x).

If errors are occurred in this step, please see the ["Trouble shootings"](#) section.

### Step 4.

Run **meta-velvetg** with manual setting of coverage peaks.

cmd **COMMAND**

```
~$ meta-velvetg out-dir -ins_length 260 \  
-exp_covs 214_122_70_43_25_13.5
```

Please note that meta-velvetg assumes that the peak values are sorted in a descending order.