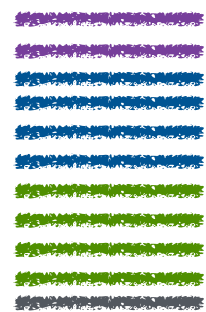


## Step 1



miRNA and isomiR  
counts matrix

import to DESeq2  
object

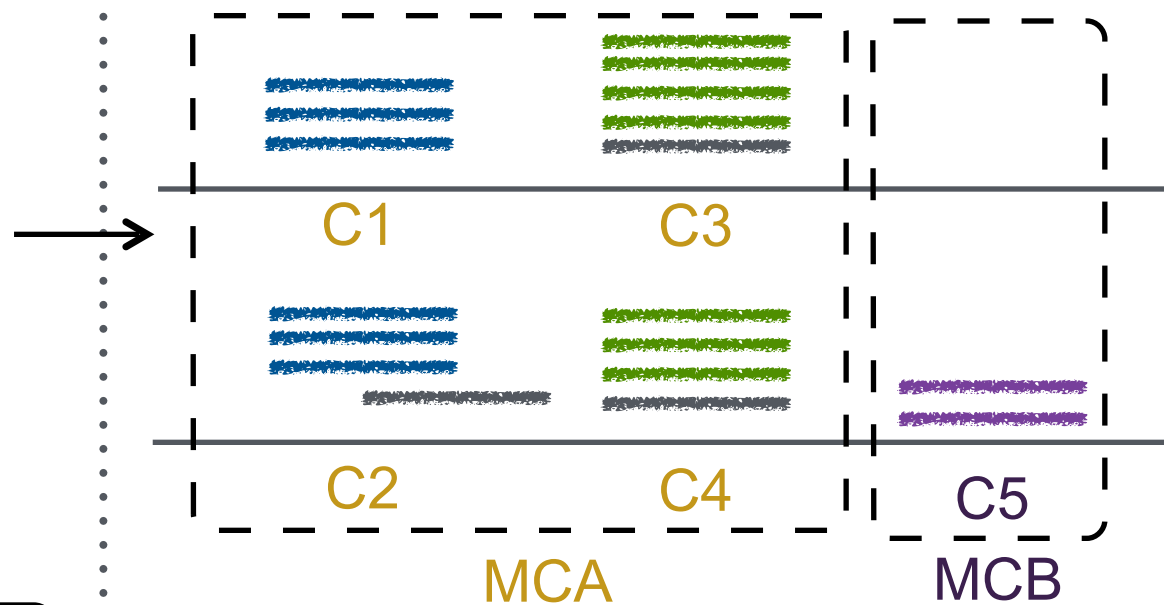
DESeq2  
normalization

PLS-DA

Differential  
expression

## Step 5

Clustering and  
Differential Expression



detecting **clusters** on the genome

## Step 2

## Step 4

Output: generation of sRNA clusters  
with counts matrix and annotation

MC1

C1

C2

MC2

C3

C4

MC3

C5

Initialize new meta-cluster with 0 clusters (ex:MC1)

calculate similarity matrix among all pairs of  
clusters.  
Iterate through pairs of clusters from most similar  
to less similar.

Check if pair is  $\geq 60\%$   
similar  
C1 and C2 in this case

Yes

No

Add both cluster to MC1  
if they weren't.  
ex C'12 and C'34

mark as conflict.  
ex: C'34 with C'12

Remove pair from MCA

Are there more  
clusters in the  
primary meta  
cluster?

Yes

No

sequences to the most  
populated C.  
ex: grey goes with C'34

Yes

any conflict?  
ex: C'34 with C'12

No

create second cluster.  
ex MCB with C'34

Define clusters  
ex. MC1 and MC2

## Step 3

Reduction + Correction