

# comparing genomes— orthology

## SESSION 7

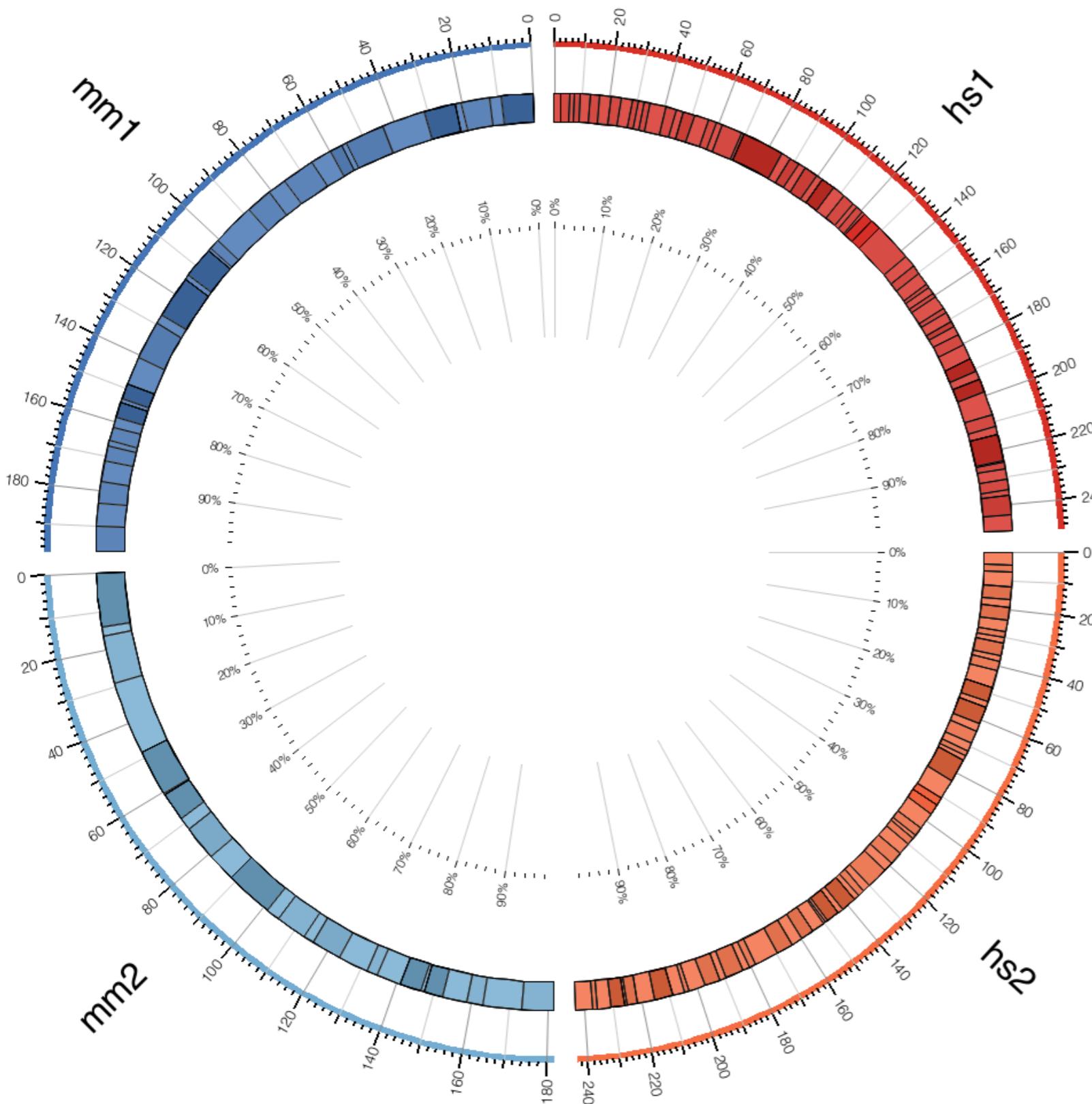
**MARTIN KRZYWINSKI**

Genome Sciences Center  
BC Cancer Agency  
Vancouver, Canada

EMBO PRACTICAL COURSE:  
BIOINFORMATICS GENOME ANALYSES

Centre for Research & Technology - Hellas, Thessalonica, Greece  
June 5–17, 2017

# SESSION FINAL IMAGE



drawing and spacing ideograms  
relative ideogram spacing  
changing ideogram scale  
ideogram selection  
ideogram order  
drawing ideogram regions  
chromosome breaks  
ordering ideogram regions  
cytogenetic bands  
drawing multiple genomes  
ideogram progression and orientation  
relative and absolute ticks

This is the image you will create during this session. It contains chrs 1 & 2 from human and mouse genomes. Each chromosome occupies 1/4 of the figure.

# CIRCOS TOOLS

If you would like, install the Circos tools from

*<http://www.circos.ca/distribution/circos-tools-0.22.tgz>*

to make utilities like bundlelinks and binlinks available.

Some of the slides optionally ask you to run these tools, e.g.

`$CIRCOS/tools/binlinks/bin/binlinks`

where `$CIRCOS` is the directory where you installed the tools.

If you feel inclined to experiment, perform those steps.

If you choose not to, that's fine. All data files to generate the images are already present.

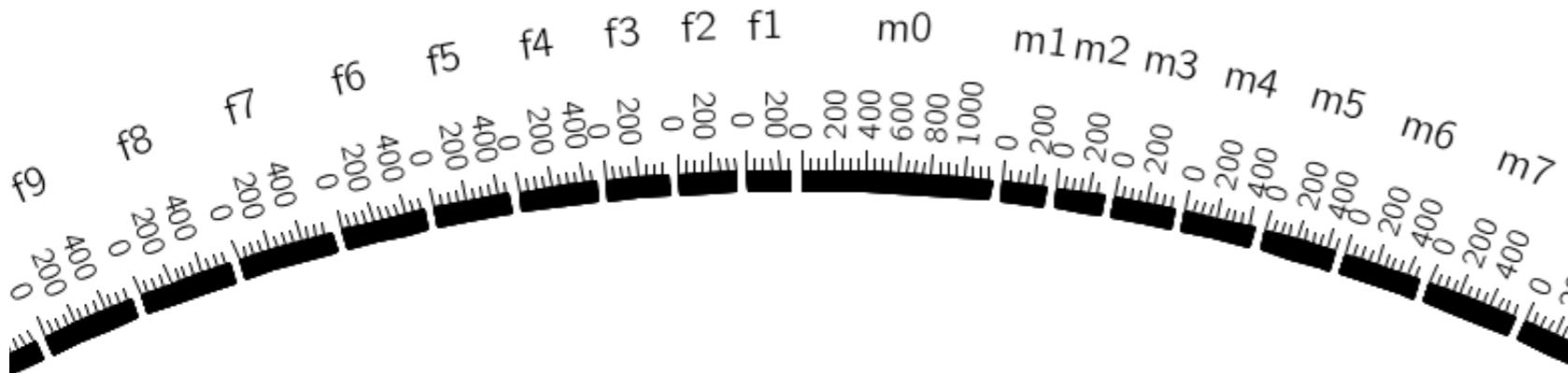
# **multiple genomes**

## **LESSON 1**

# L. MEXICANA AND L. MAJOR KARYOTYPES

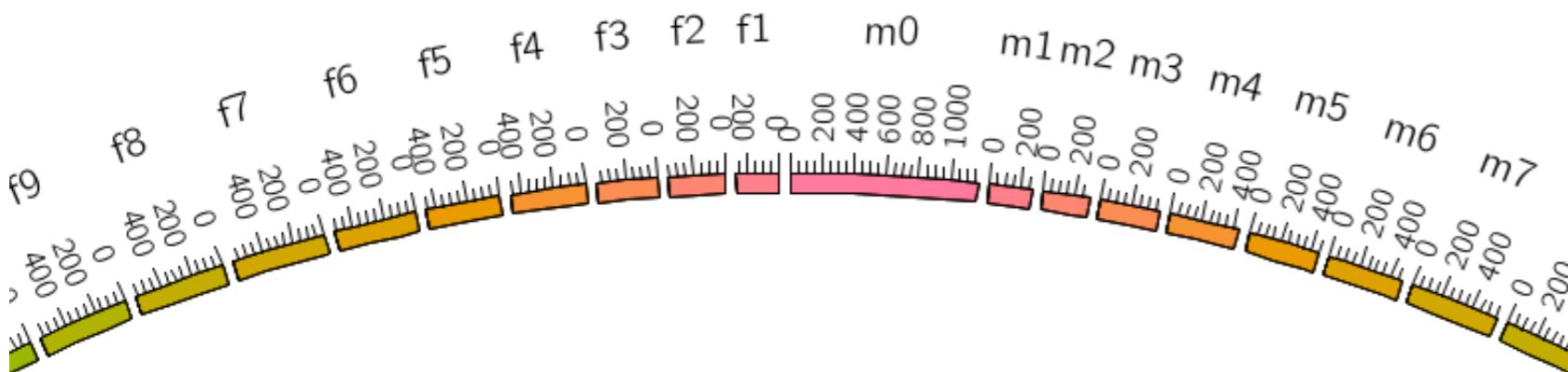
```
# ./data/lm.karyotype.txt
chr - LmxM.00 00 0 1171052 black
chr - LmxM.01 01 0 273291 black
chr - LmxM.02 02 0 298030 black
chr - LmxM.03 03 0 375930 black
chr - LmxM.04 04 0 438817 black
...
# ./data/lf.karyotype.txt
chr - LmjF.36 36 0 2682151 black
chr - LmjF.35 35 0 2090474 black
chr - LmjF.34 34 0 1866748 black
chr - LmjF.33 33 0 1583653 black
chr - LmjF.32 32 0 1604637 black
chr - LmjF.31 31 0 1484328 black
chr - LmjF.30 30 0 1403434 black
chr - LmjF.29 29 0 1212663 black
chr - LmjF.28 28 0 1160104 black
chr - LmjF.27 27 0 1130424 black
...
```

# MULTIPLE KARYOTYPES



```
karyotype      =
  ./data/lm.karyotype.txt,
  ./data/lf.karyotype.txt
chromosomes_units = 1000
#chromosomes_reverse = /jF/
#chromosomes_color   = ./=var(chr)
```

# IDEOGRAM COLORS



```
karyotype      =
  ./data/lm.karyotype.txt,
  ./data/lf.karyotype.txt
chromosomes_units = 1000
chromosomes_reverse = /jF/
chromosomes_color   = ./=var(chr)
```

# L. MEXICANA AND L. MAJOR KARYOTYPES

# 7/2/circos.conf

...

luminance = 70

chroma = 100

<colors>

lmjf.00=lch(conf(luminance),conf(chroma),0)

lmjf.01=lch(conf(luminance),conf(chroma),9)

lmjf.02=lch(conf(luminance),conf(chroma),19)

lmjf.03=lch(conf(luminance),conf(chroma),29)

lmjf.04=lch(conf(luminance),conf(chroma),38)

lmjf.05=lch(conf(luminance),conf(chroma),48)

lmjf.06=lch(conf(luminance),conf(chroma),58)

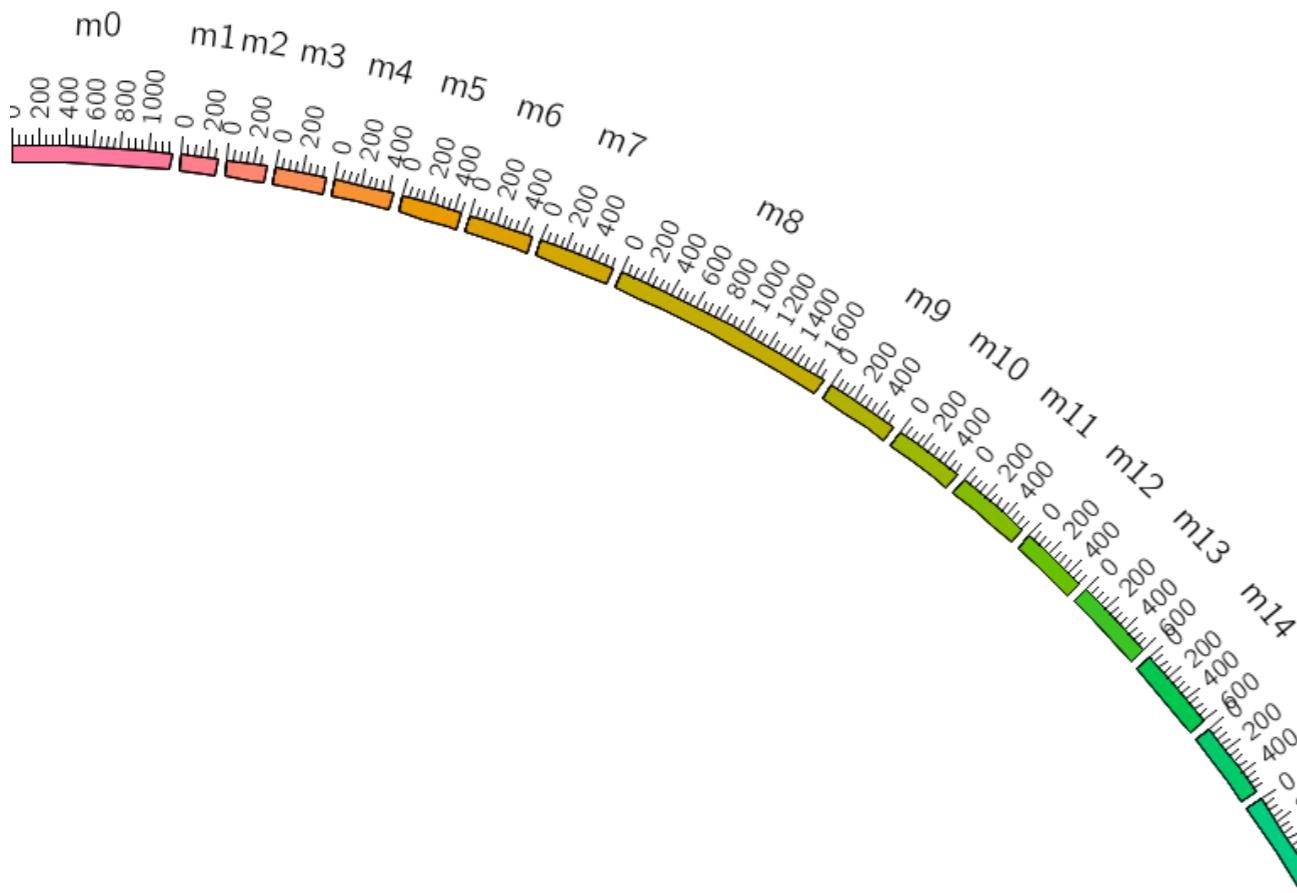
lmjf.07=lch(conf(luminance),conf(chroma),68)

lmjf.08=lch(conf(luminance),conf(chroma),77)

...

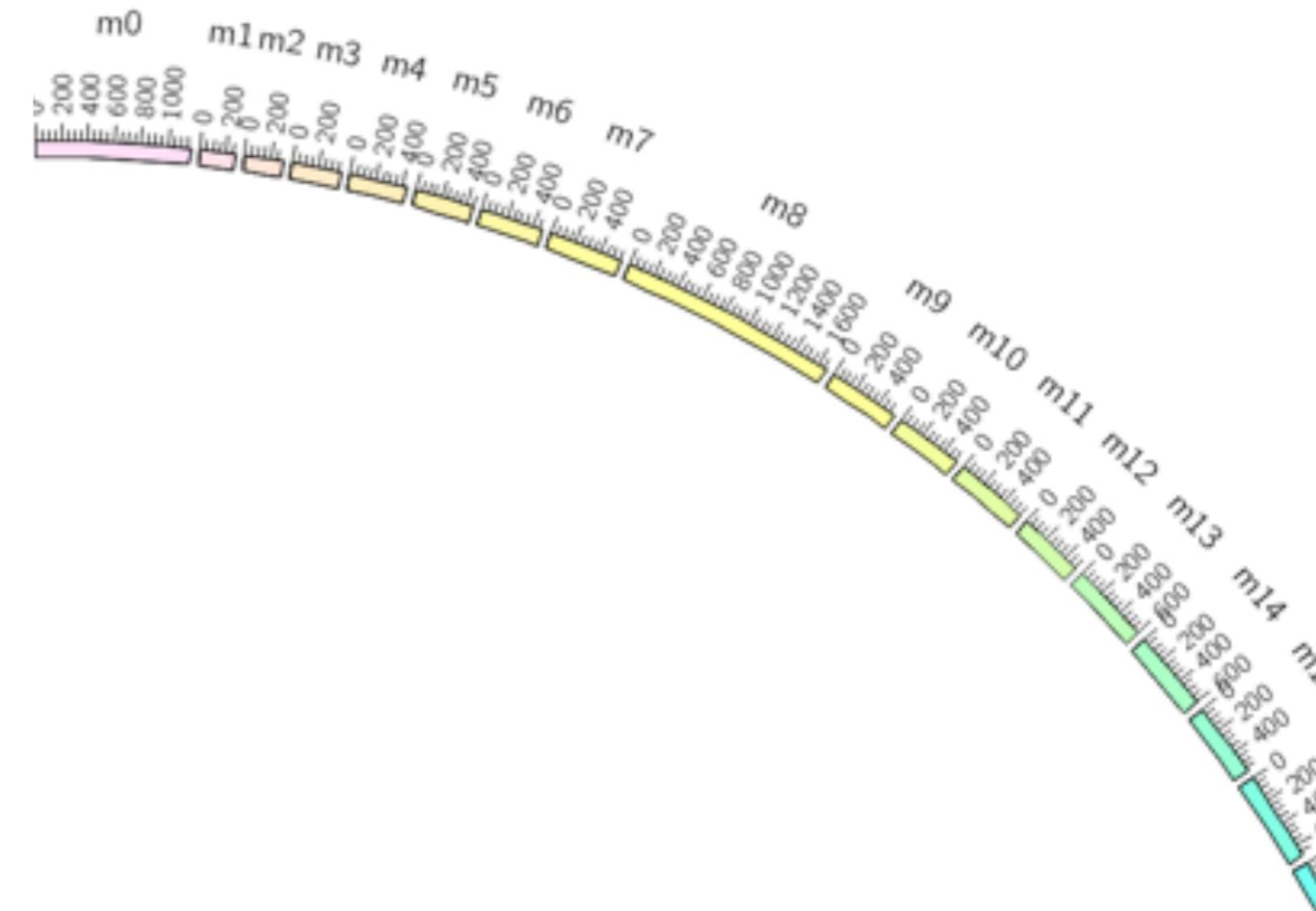
</colors>

# LUMINANCE AND CHROMA



luminance = 70  
chroma = 100

`circos -param luminance=70,chroma=100`



luminance = 100  
chroma = 70

`circos -param luminance=100 -param chroma=70`

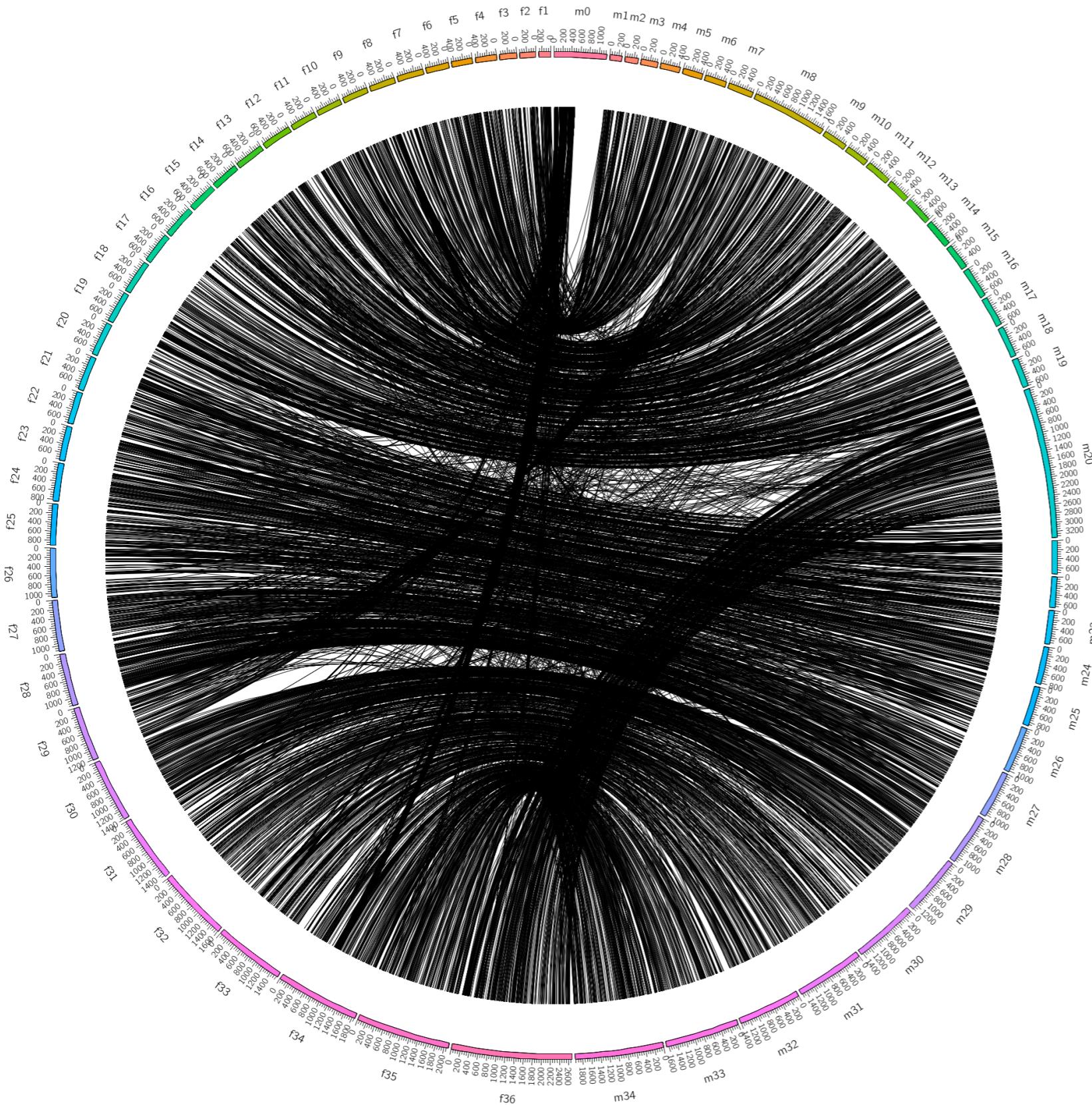
# links

## LESSON 2

# LINKS FILE—GENE ORTHOLOGUES

LmxM.01 9205 11223 LmjF.01 9061 11067 gene1=LmjF.01.0030,gene2=LmxM.01.0030  
LmxM.13 567145 569319 LmjF.01 9061 11067 gene1=LmjF.01.0030,gene2=LmxM.13.1610  
LmxM.01 14846 16843 LmjF.01 15025 17022 gene1=LmjF.01.0050,gene2=LmxM.01.0050  
LmxM.30 1416292 1418355 LmjF.01 15025 17022 gene1=LmjF.01.0050,gene2=LmxM.30.3130  
LmxM.01 27727 28314 LmjF.01 28521 29108 gene1=LmjF.01.0110,gene2=LmxM.01.0110  
LmxM.01 42656 44968 LmjF.01 44148 46466 gene1=LmjF.01.0180,gene2=LmxM.01.0180  
LmxM.01 55167 56171 LmjF.01 57026 58033 gene1=LmjF.01.0220,gene2=LmxM.01.0220  
LmxM.01 60622 63405 LmjF.01 60056 62833 gene1=LmjF.01.0240,gene2=LmxM.01.0240  
LmxM.01 69317 69640 LmjF.01 69584 69907 gene1=LmjF.01.0270,gene2=LmxM.01.0270  
...

# LINKS



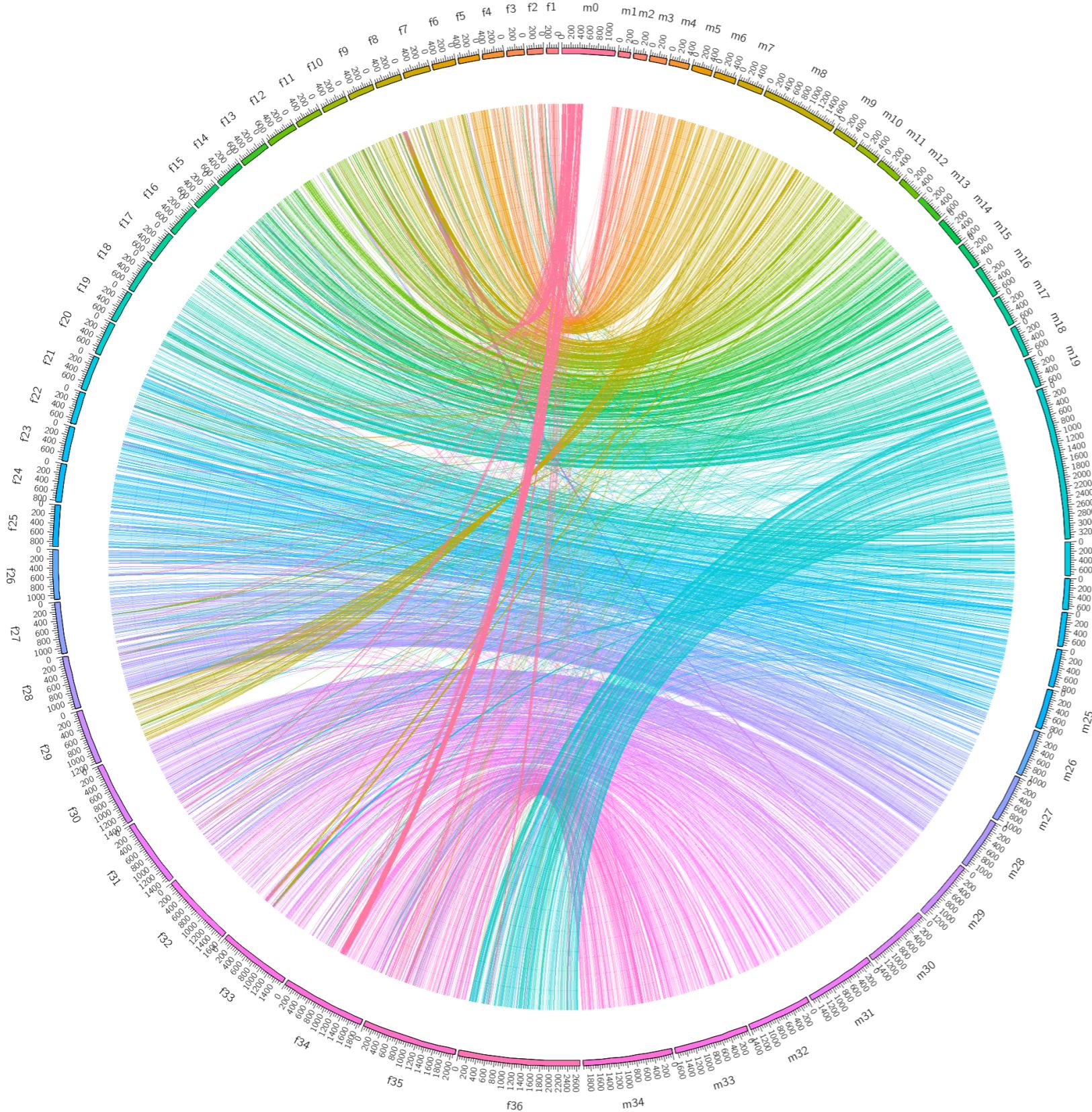
# 7/2/etc/circos.conf

```
<links>
<link>
file  = conf(datadir)/ortho.links.txt
radius = 0.90r

<rules>
use = no
<rule>
condition =
substr(var(chr1),-2) eq substr(var(chr2),-2)
show   = no
</rule>
<rule>
use = no
condition = 1
color   =
eval(sprintf("%s_a4",lc var(chr1)))
</rule>
</rules>

</link>
</links>
```

# LINKS WITH TRANSPARENCY



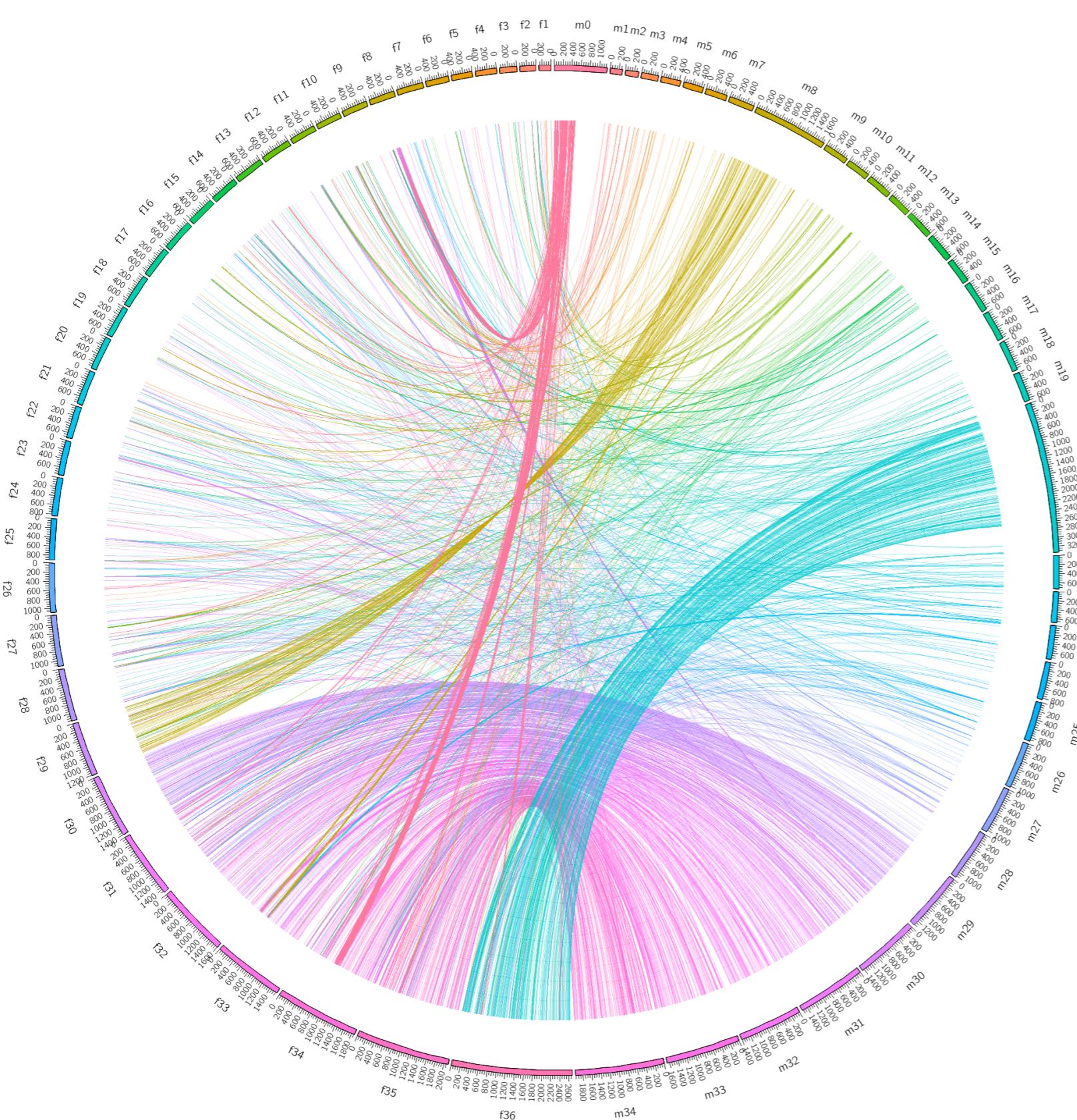
# 7/2/etc/circos.conf

```
<links>
<link>
file    = conf(datadir)/ortho.links.txt
radius = 0.90r

<rules>
use = no
<rule>
condition =
substr(var(chr1),-2) eq substr(var(chr2),-2)
show    = no
</rule>
<rule>
use = yes
condition = 1
color    =
eval(sprintf("%s_a4",lc var(chr1)))
</rule>
</rules>

</link>
</links>
```

# LINKS BETWEEN CHRS WITH DIFFERENT INDEX



# 7/2/etc/circos.conf

```
<links>
<link>
file    = conf(datadir)/ortho.links.txt
radius  = 0.90r

<rules>
use = yes
<rule>
condition =
substr(var(chr1),-2) eq substr(var(chr2),-2)
show    = no
</rule>
<rule>
use = yes
condition = 1
color    =
eval(sprintf("%s_a4",lc var(chr1)))
</rule>
</rules>

</link>
</links>
```

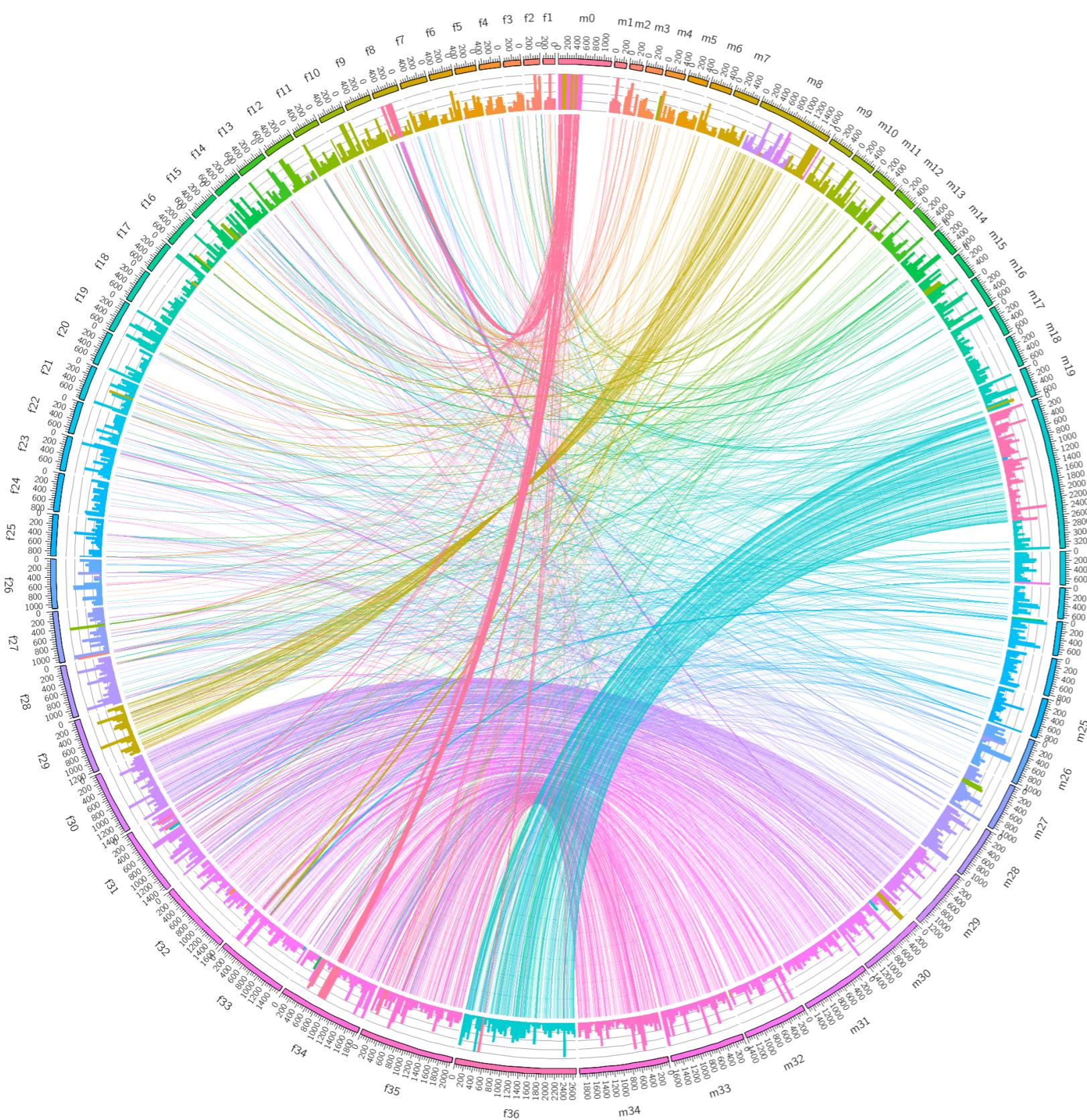
# **counting links, consensus chromosomes**

## **LESSON 3**

# IMPORTING COMMON TRACK SETTINGS

```
> cat ortho.links.txt |  
$CIRCOS/circos/tools/binlinks/bin/binlinks -bin 50000 -link_end 2 -num -color_by_chr  
  
LmjF.01 0 49999 6.0000 fill_color=lmxm.01  
LmjF.01 50000 99999 6.0000 fill_color=lmxm.01  
LmjF.01 100000 149999 25.0000 fill_color=lmxm.01  
LmjF.01 150000 199999 7.0000 fill_color=lmxm.01  
LmjF.01 200000 249999 5.0000 fill_color=lmxm.01  
LmjF.01 250000 299999 4.0000 fill_color=lmxm.01  
LmjF.02 0 49999 19.0000 fill_color=lmxm.02  
LmjF.02 50000 99999 11.0000 fill_color=lmxm.02  
LmjF.02 100000 149999 20.0000 fill_color=lmxm.02  
LmjF.02 150000 199999 7.0000 fill_color=lmxm.02  
LmjF.02 250000 299999 1.0000 fill_color=lmxm.02  
LmjF.02 300000 349999 6.0000 fill_color=lmxm.02  
...
```

# LINK COUNTS AS HISTOGRAM



# 7/3/etc/circos.conf

```
<plot>
type = histogram
file = conf(datadir)/hist.links.txt
r1 = 0.98r
r0 = 0.91r
color = undef
min = 0
max = 20

<axes>
<axis>
spacing = 5
</axis>
</axes>

</plot>
```

# LINK COUNTS AS HISTOGRAM



# 7/3/etc/circos.conf

```
<plot>
type = histogram
file = conf(datadir)/hist.links.txt
r1 = 0.98r
r0 = 0.91r
color = undef
min = 0
max = 50
```

```
<axes>
<axis>
spacing = 5
</axis>
</axes>
```

```
</plot>
```

# **highlights from binned links**

## **LESSON 4**

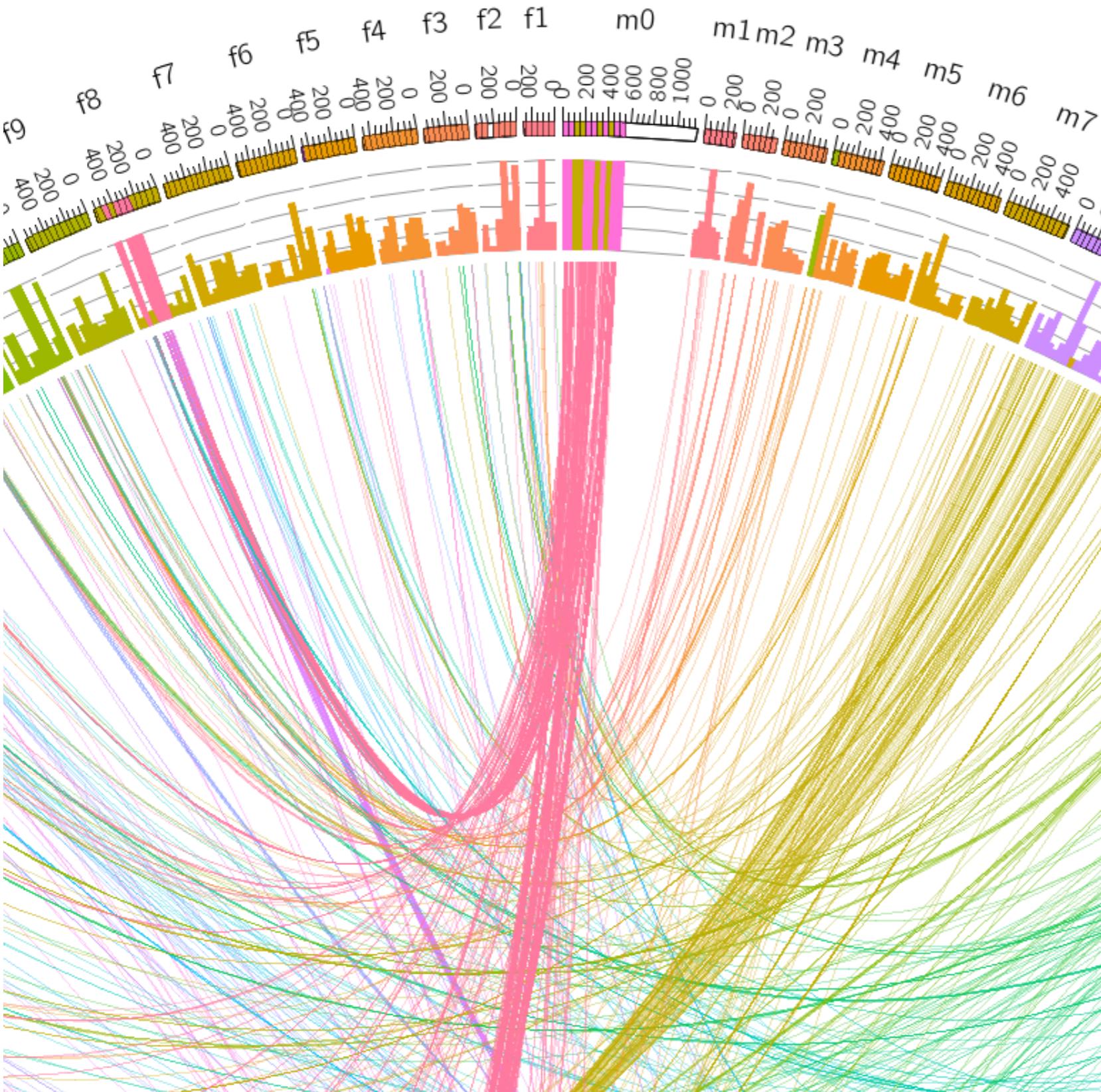
# BINNED LINK COUNTS

...

LmjF.**16** 550000 599999 12.0000 fill\_color=lmxm.**08**  
LmjF.**16** 600000 649999 15.0000 fill\_color=lmxm.**14**  
LmjF.16 650000 699999 5.0000 fill\_color=lmxm.16  
LmjF.17 0 49999 57.0000 fill\_color=lmxm.17  
LmjF.17 50000 99999 36.0000 fill\_color=lmxm.17  
LmjF.17 100000 149999 5.0000 fill\_color=lmxm.17  
LmjF.17 150000 199999 2.0000 fill\_color=lmxm.17  
LmjF.17 200000 249999 1.0000 fill\_color=lmxm.17  
LmjF.17 250000 299999 3.0000 fill\_color=lmxm.17  
LmjF.**17** 300000 349999 4.0000 fill\_color=lmxm.**07**  
LmjF.17 350000 399999 7.0000 fill\_color=lmxm.17  
LmjF.17 400000 449999 3.0000 fill\_color=lmxm.17  
LmjF.17 450000 499999 2.0000 fill\_color=lmxm.17  
LmjF.17 500000 549999 6.0000 fill\_color=lmxm.17  
LmjF.17 550000 599999 6.0000 fill\_color=lmxm.17  
LmjF.17 600000 649999 2.0000 fill\_color=lmxm.17  
LmjF.17 650000 699999 4.0000 fill\_color=lmxm.17  
LmjF.18 0 49999 4.0000 fill\_color=lmxm.18  
LmjF.18 50000 99999 1.0000 fill\_color=lmxm.18

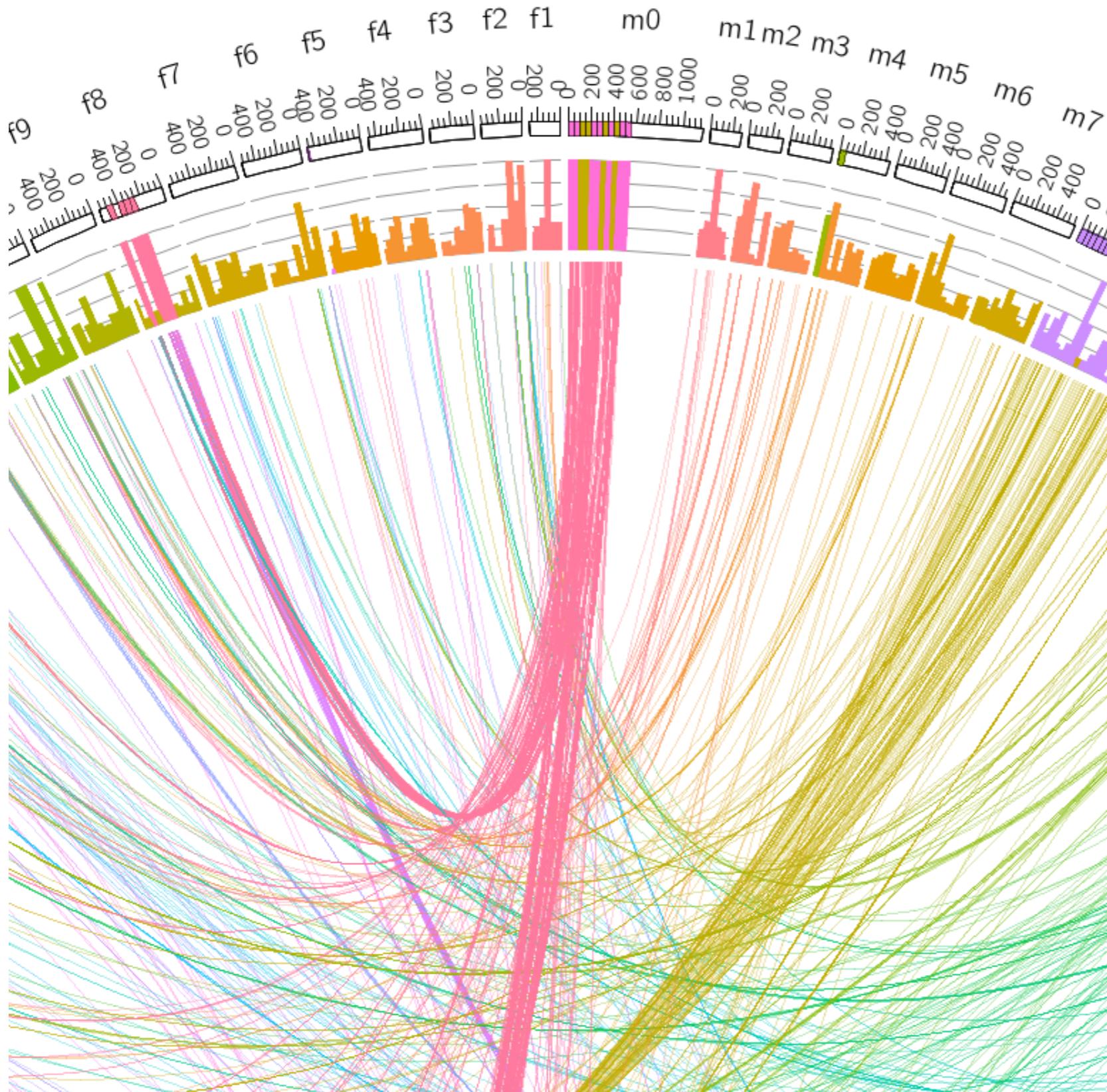
...

# BINNED LINK HIGHLIGHTS



```
# 7/4/etc/circos.conf  
  
chromosomes_color = ./=white  
  
<plot>  
type = highlight  
file = conf(datadir)/hist.links.txt  
r1 = dims(ideogram,radius_outer)  
r0 = dims(ideogram,radius_inner)  
stroke_color = black_a3  
</plot>  
  
...
```

# BINNED LINK HIGHLIGHTS



```
# 7/4/etc/circos.conf

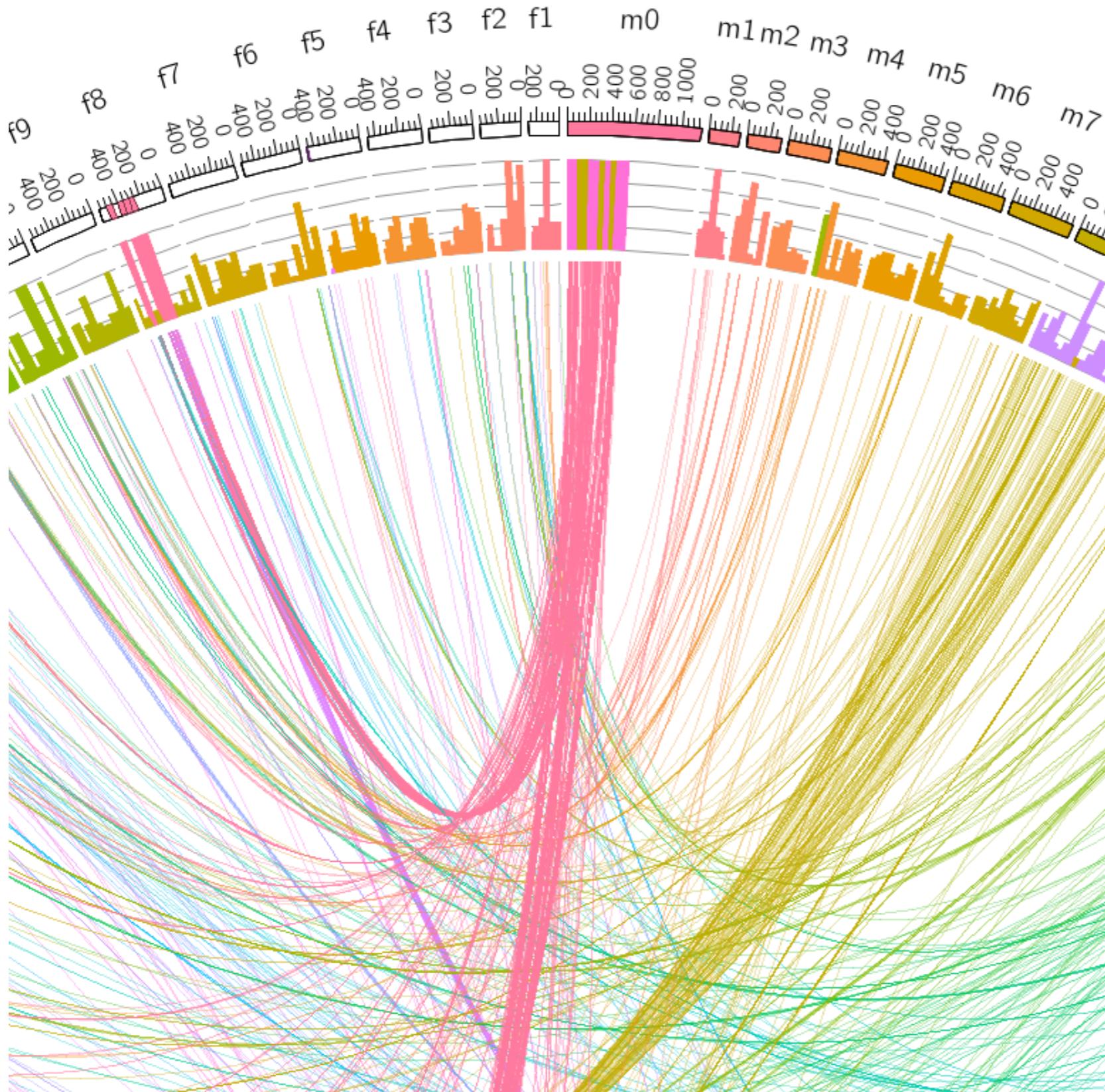
chromosomes_color = ./=white

<plot>
type = highlight
file = conf(datadir)/hist.links.txt
r1 = dims(ideogram,radius_outer)
r0 = dims(ideogram,radius_inner)
stroke_color = black_a3

<rules>
use = yes
<rule>
#condition = var(chr) =~ /xm/i
condition =
    substr(var(fill_color),-2) eq
    substr(var(chr),-2)
show = no
</rule>
</rules>
</plot>

</plot>
...
```

# BINNED LINK HIGHLIGHTS



```
# 7/4/etc/circos.conf

chromosomes_color = ./=var(chr);jF=/white

<plot>
type = highlight
file = conf(datadir)/hist.links.txt
r1 = dims(ideogram,radius_outer)
r0 = dims(ideogram,radius_inner)
stroke_color = black_a3

<rules>
use = yes
<rule>
condition = var(chr) =~ /xm/i ||
substr(var(fill_color),-2) eq
substr(var(chr),-2)
show = no
</rule>
</rules>
</plot>

</plot>
...
```

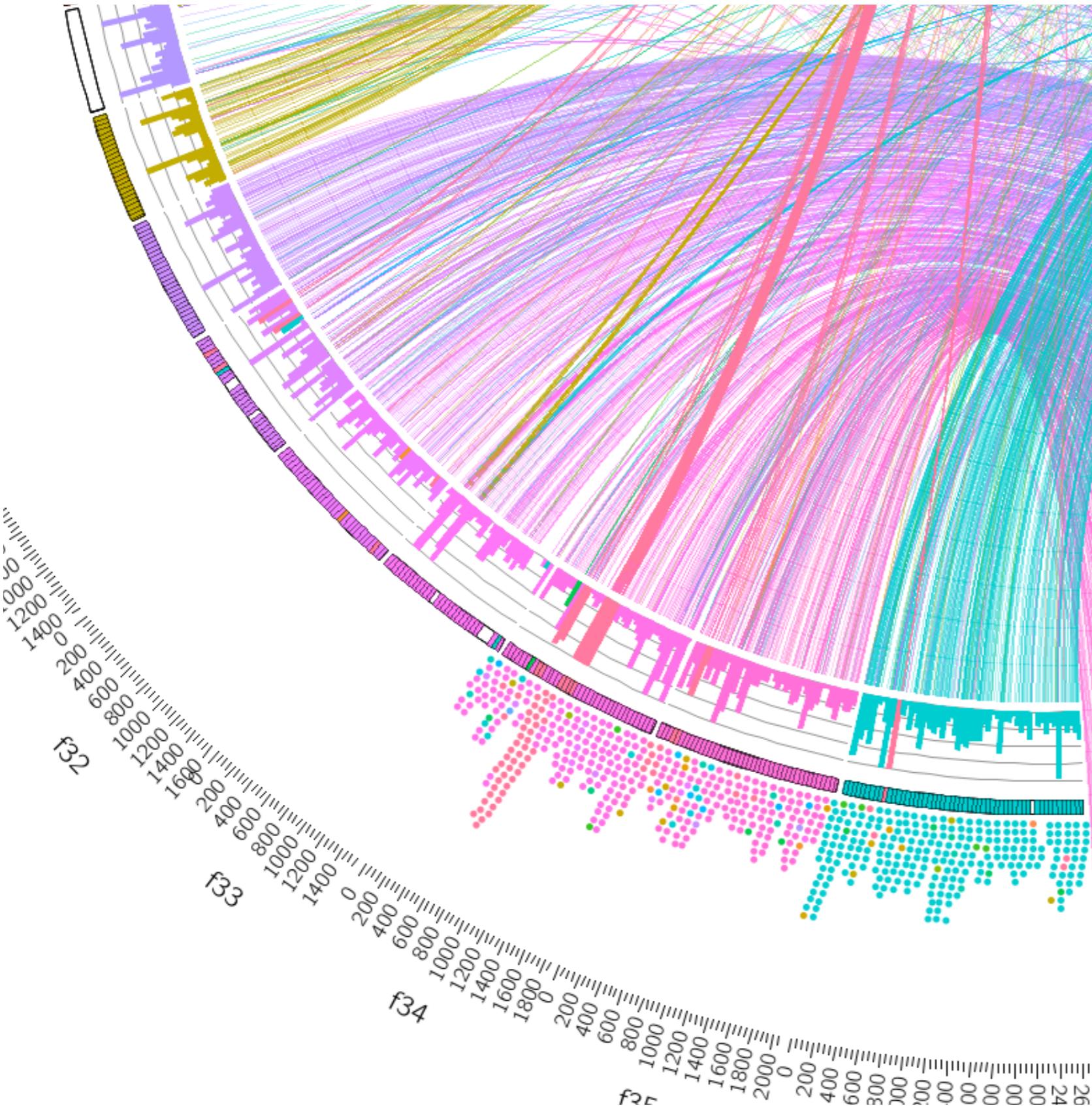
# **bubble track—(text track with symbol font**

## **LESSON 5**

# ORTHOLOGUE CHROMOSOME RELATIONSHIPS

```
> cat ortho.links.txt | awk '{print $4,$5,$6,$1,$7}' | sort -u -k 2,2 | sort > ortho.labels.txt
LmjF.01 100528 101130 LmxM.01 gene1=LmjF.01.0410,gene2=LmxM.01.0410
LmjF.01 102444 103046 LmxM.01 gene1=LmjF.01.0420,gene2=LmxM.01.0410
LmjF.01 112139 112795 LmxM.01 gene1=LmjF.01.0450,gene2=LmxM.01.0450
LmjF.01 121759 123849 LmxM.01 gene1=LmjF.01.0470,gene2=LmxM.01.0470
LmjF.01 124711 125790 LmxM.01 gene1=LmjF.01.0480,gene2=LmxM.01.0480
LmjF.01 127704 129800 LmxM.01 gene1=LmjF.01.0490,gene2=LmxM.01.0470
LmjF.01 130947 133070 LmxM.01 gene1=LmjF.01.0500,gene2=LmxM.01.0470
...
...
```

# GENE BUBBLES COLORED BY LM CHR



```
# 7/5/etc/circos.conf

chromosomes_radius = /jF/:0.8r

<plots>

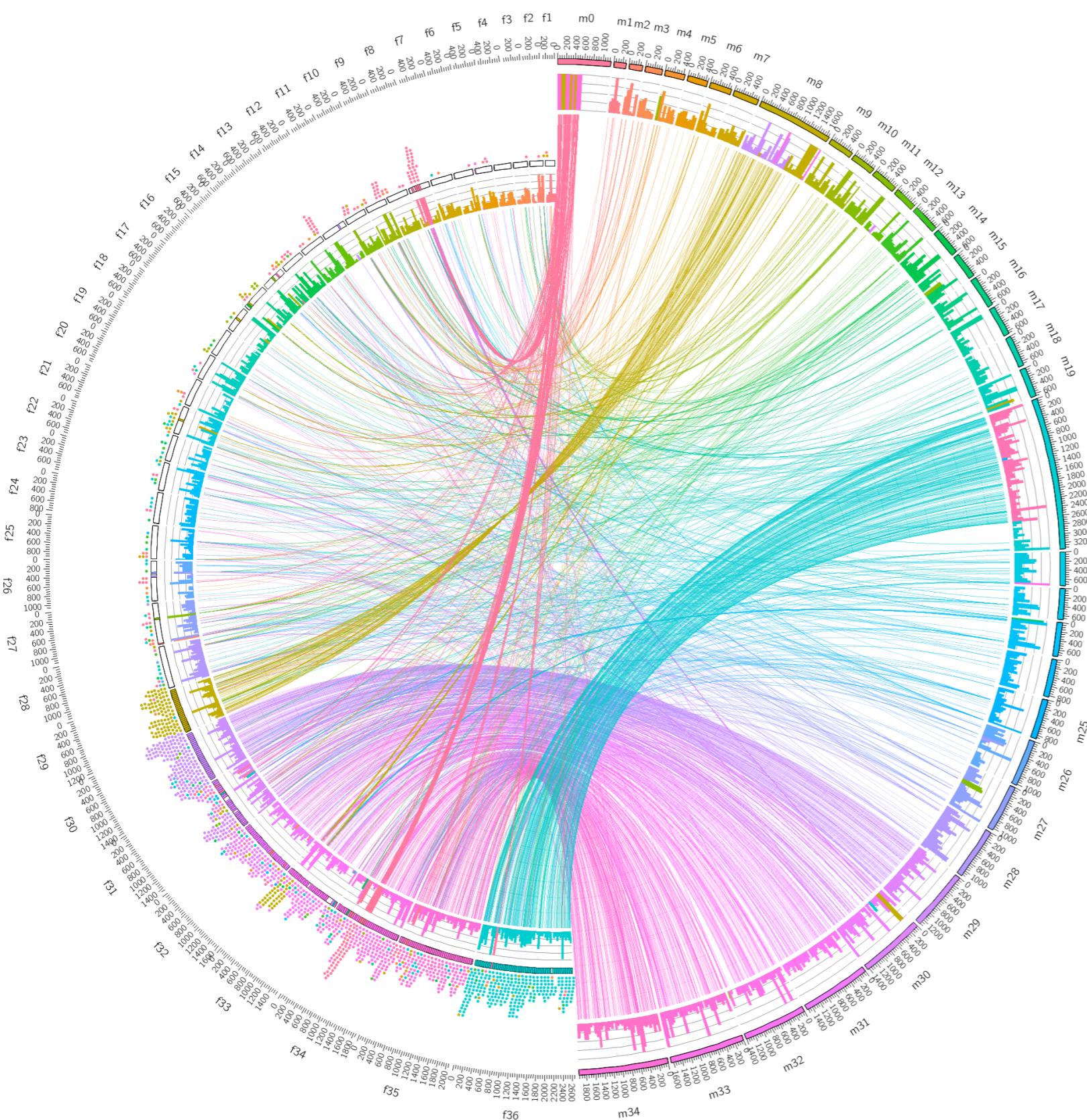
<plot>
type = text
file = conf(datadir)/ortho.labels.txt
r0 = 1r+5p
r1 = dims(image,radius)-100p
label_font = glyph

label_snuggle = yes

<rules>
<rule>
condition = var(chr) !~ /\.3[456]/
#condition = abs(substr(var(chr),-2) - substr(var(value),-2)) == 0
show = no
</rule>
<rule>
condition = 1
color = eval(lc var(value))
</rule>
</rules>

</plot>
```

# BUBBLES ON LF CHRS WITH DIFFERENT INDEX



# 7/5/etc/circos.conf

chromosomes\_radius = /jF/:0.8r

<plots>

<plot>

type = text

file = conf(datadir)/ortho.labels.txt

r0 = 1r+5p

r1 = dims(image,radius)-100p

label\_font = glyph

label\_snuggle = yes

<rules>

<rule>

#condition = var(chr) !~ /\.[3][456]/  
condition = abs(substr(var(chr),-2) - substr(var(value),-2)) == 0

show = no

</rule>

<rule>

condition = 1

color = eval(lc var(value))

</rule>

</rules>

</plot>

# bundling links

## LESSON 5

# BUNDLING LINKS

```
> cd ~/circos/data  
> cat ortho.links.txt | $CIRCOS/tools/bundlelinks/bin/bundlelinks  
    -min_bundle_membership 3
```

```
num_links 5665  
num_initial_bundles 160  
num_links_in_initial_bundles 5135
```

LmxM.00 24301 504351 LmjF.08 286603 479327

nlinks=227,bsize1=18876,bsize2=20488,bidentity1=0.039321,bidentity2=0.106307,  
depth1=2,depth2=0,gene1=LmjF.08.0820,gene2=LmxM.33.1740

LmxM.00 280060 419889 LmjF.10 7849 561941

nlinks=12,bsize1=2850,bsize2=23652,bidentity1=0.020382,bidentity2=0.042686,de  
pth1=10,depth2=1,gene1=LmjF.10.0460,gene2=LmxM.10.0465a

LmxM.00 173261 508096 LmjF.11 253778 551310

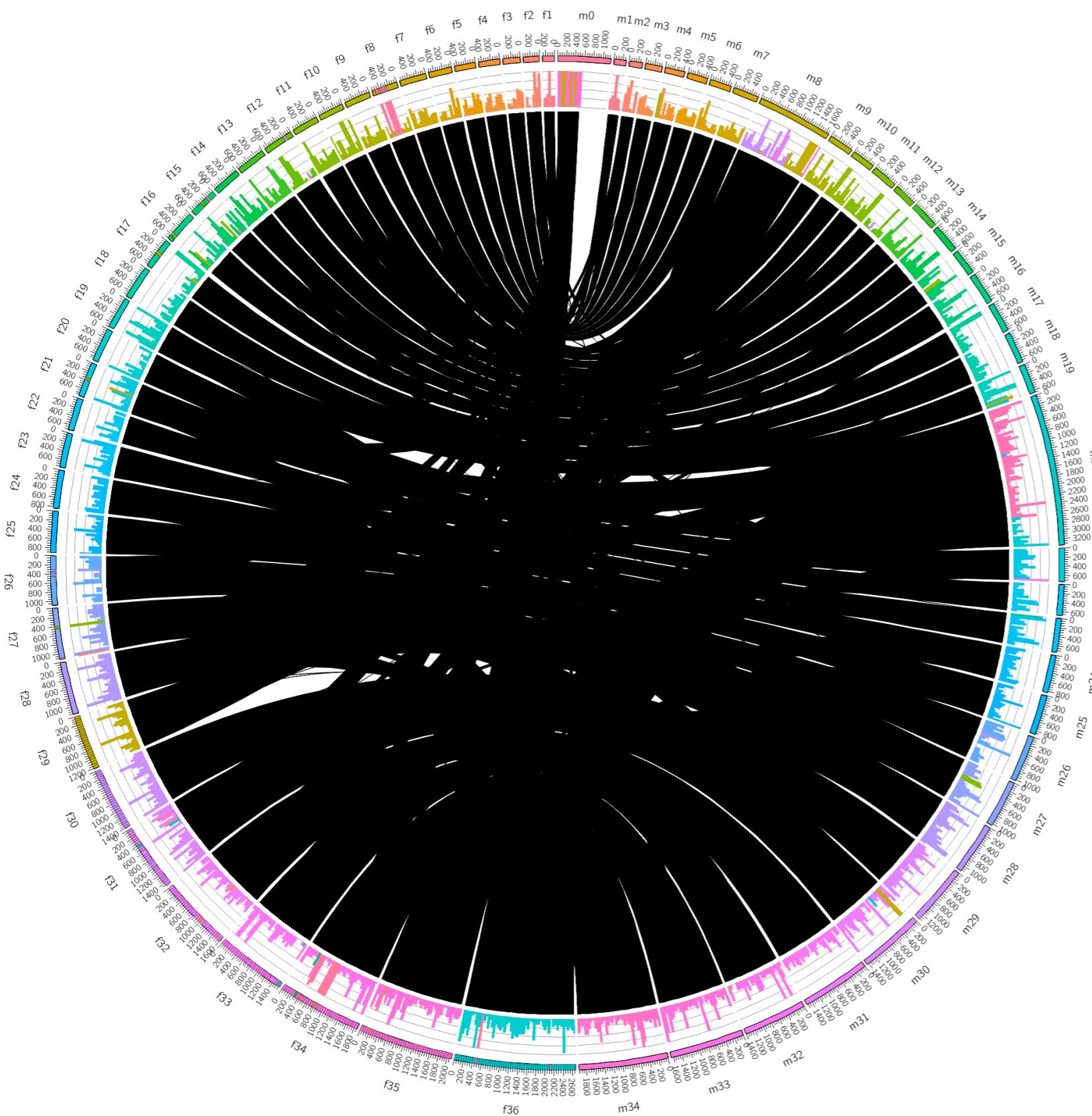
nlinks=6,bsize1=2967,bsize2=29076,bidentity1=0.008861,bidentity2=0.097724,dep  
th1=5,depth2=2,gene1=LmjF.11.1220,gene2=LmxM.11.1220a

...

num\_passed\_bundles 160 (100.00%)

num\_links\_in\_passed\_bundles 5135 (100.00%)

# BUNDLED LINKS – CAN'T SEE ANYTHING

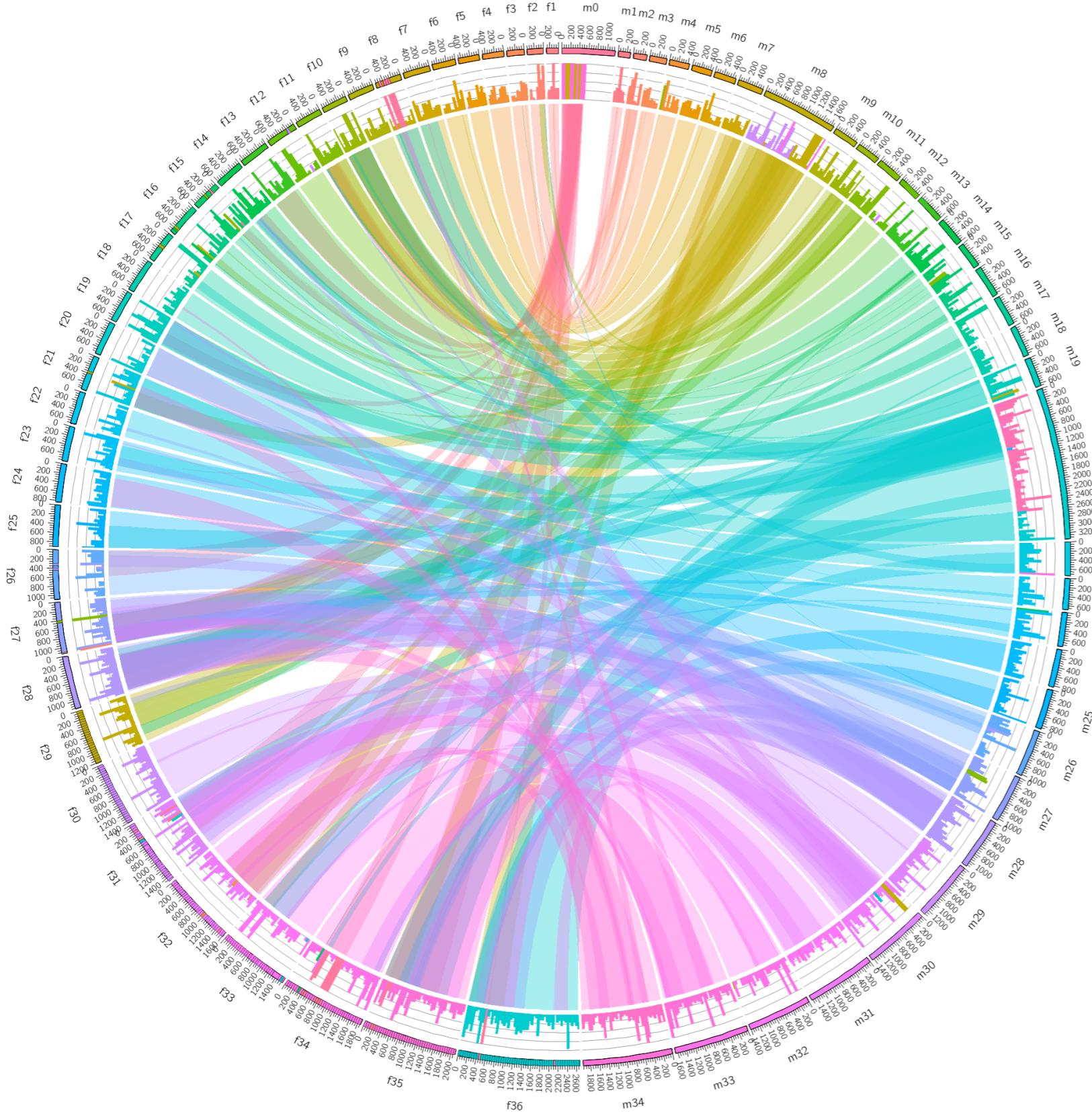


# 7/6/etc/circos.conf

```
<link>
file  = conf(datadir)/ortho.bundles.txt
radius = 0.90r
ribbon = yes

<rules>
use = no
<rule>
condition = 1
color    =
eval(sprintf("%s_a4",lc var(chr1)))
</rule>
<rule>
condition =
substr(var(chr1),-2) eq substr(var(chr2),-2)
show      = no
</rule>
</rules>
```

# COLORED BY CHROMOSOME

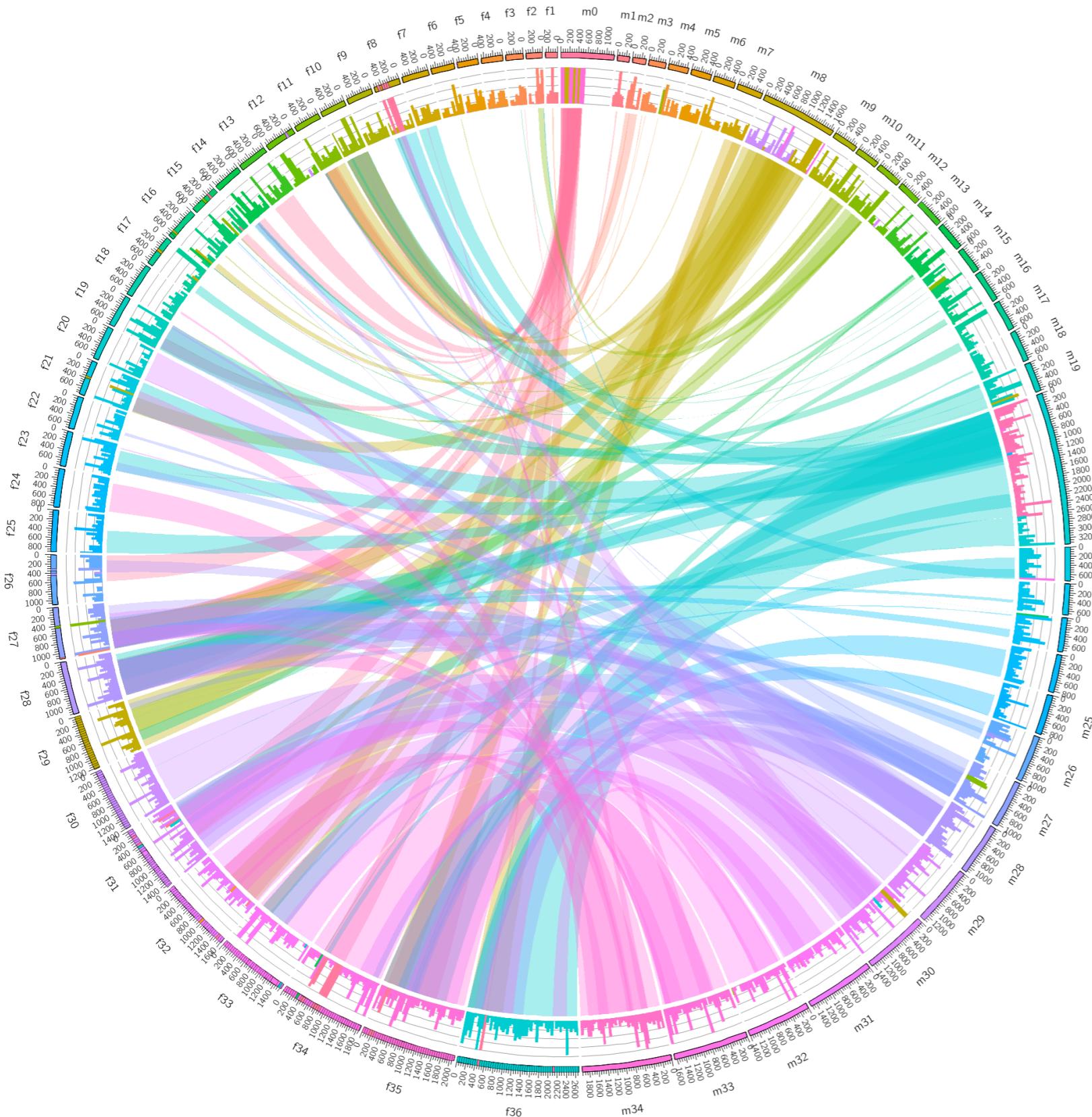


# 7/6/etc/circos.conf

```
<link>
file  = conf(datadir)/ortho.bundles.txt
radius = 0.90r
ribbon = yes

<rules>
use = yes
<rule>
condition = 1
color   =
    eval(sprintf("%s_a4",lc var(chr1)))
</rule>
<rule>
condition =
    substr(var(chr1),-2) eq substr(var(chr2),-2)
show    = no
</rule>
</rules>
```

# COLORED BY CHROMOSOME



# 7/6/etc/circos.conf

<link>

file = conf(datadir)/ortho.bundles.txt

radius = 0.90r

ribbon = yes

<rules>

use = yes

<rule>

condition = 1

color =

eval(sprintf("%s\_a4",lc var(chr1)))

flow = continue

</rule>

<rule>

condition =

substr(var(chr1),-2) eq substr(var(chr2),-2)

show = no

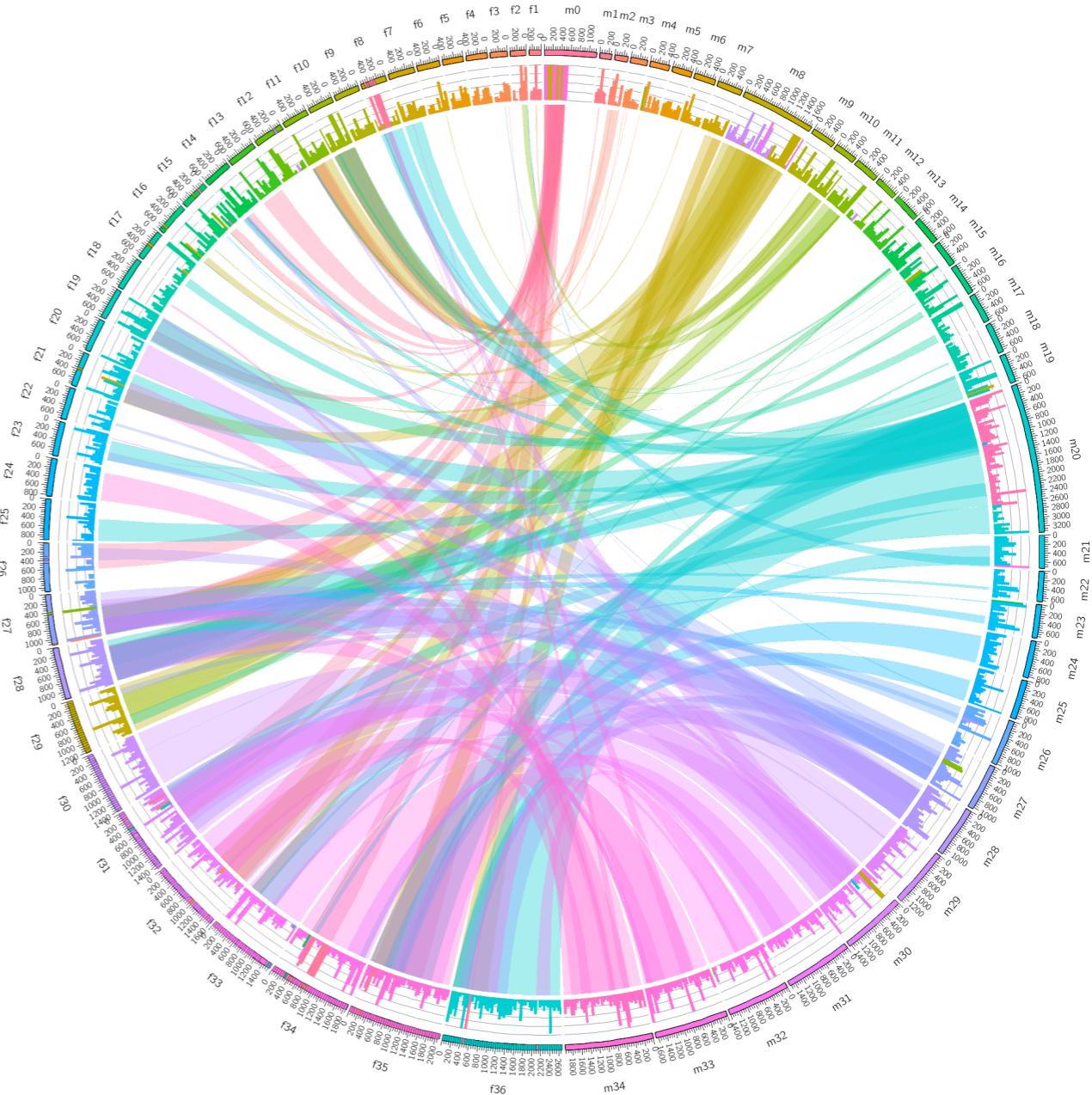
</rule>

</rules>

# COLORED BY CHROMOSOME



all links, as lines

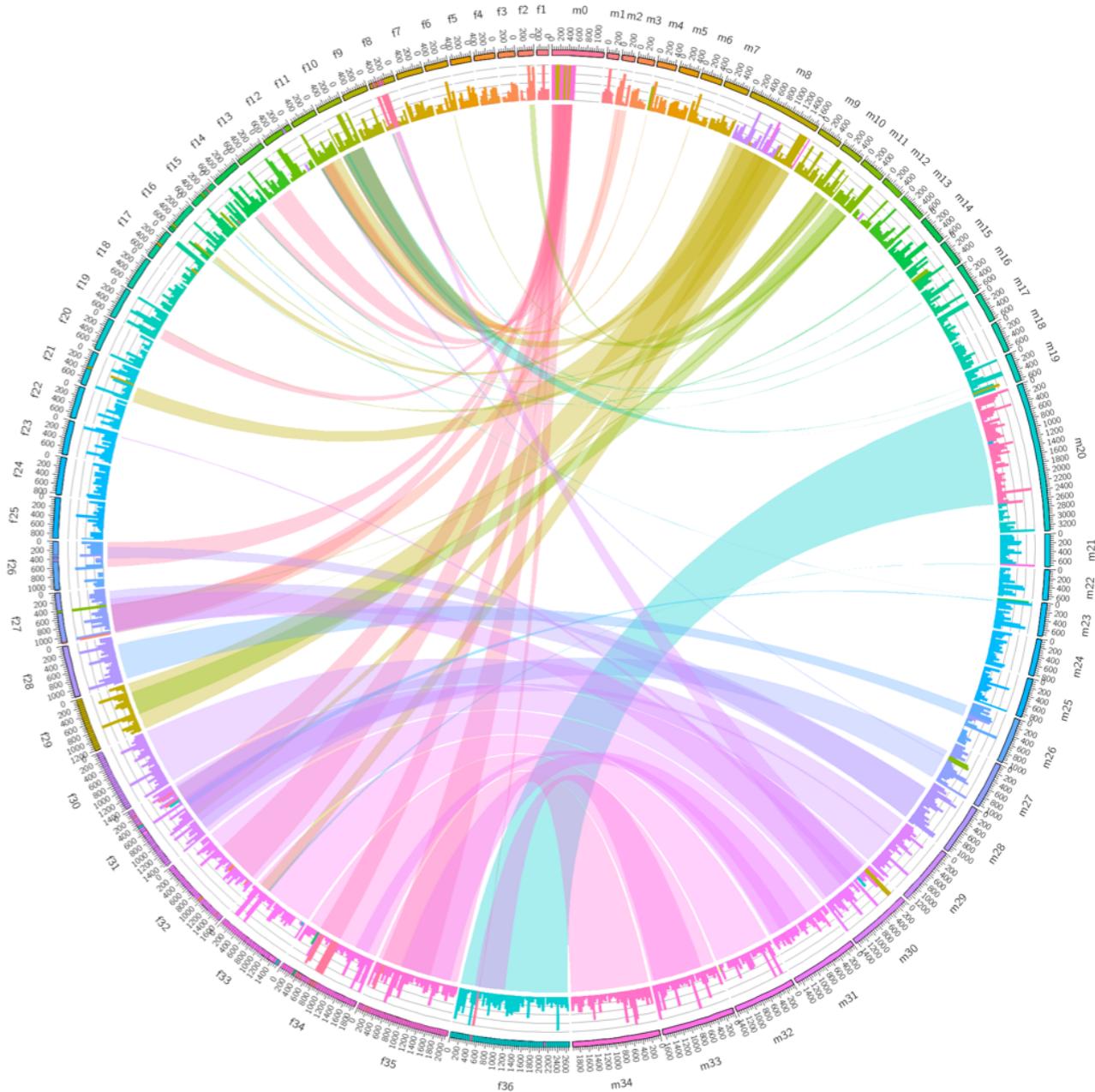


bundled links, min 3 links in a bundle  
max 1 Mb gap between links in a bundle

# COLORED BY CHROMOSOME

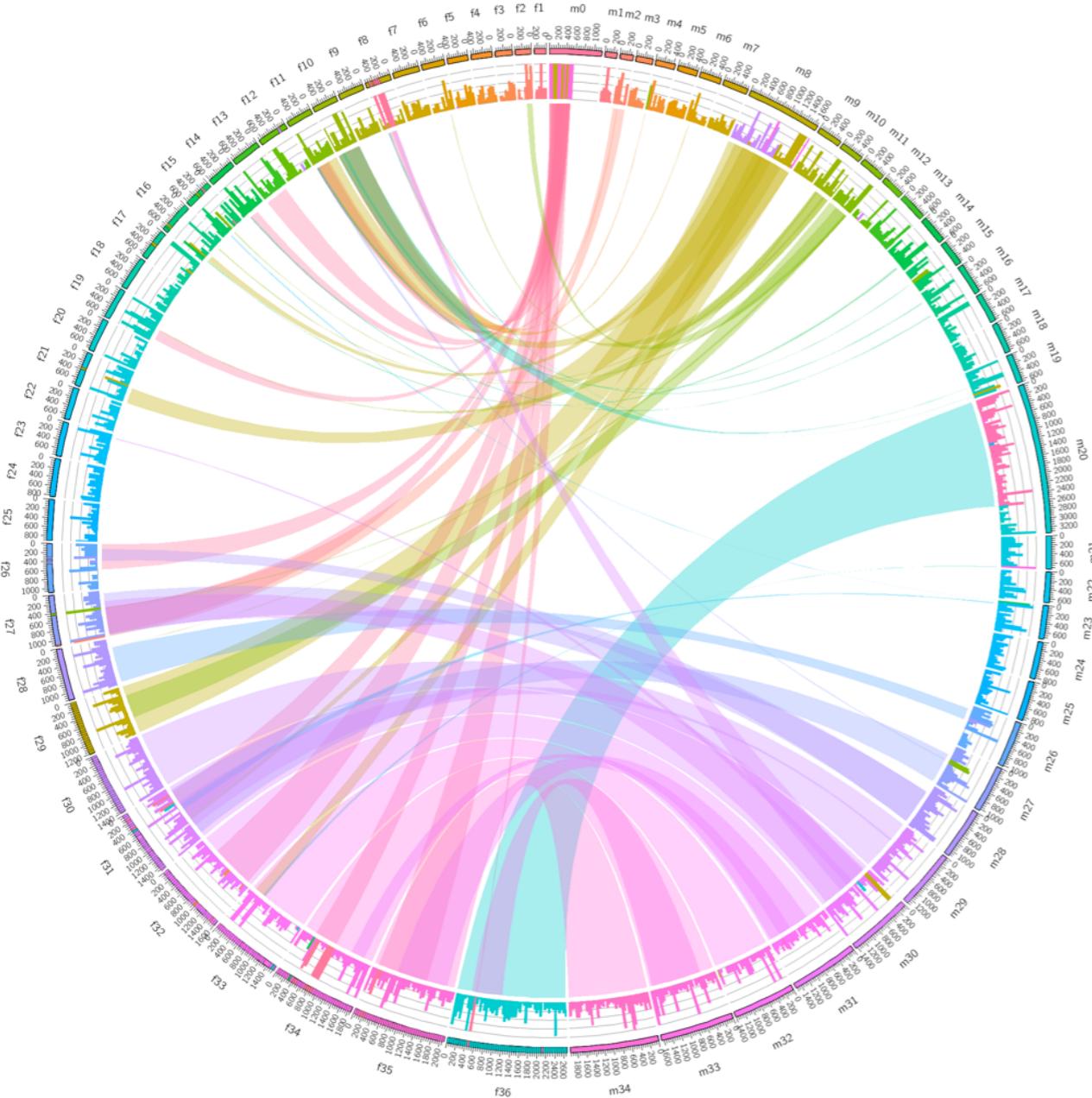


all links, as lines

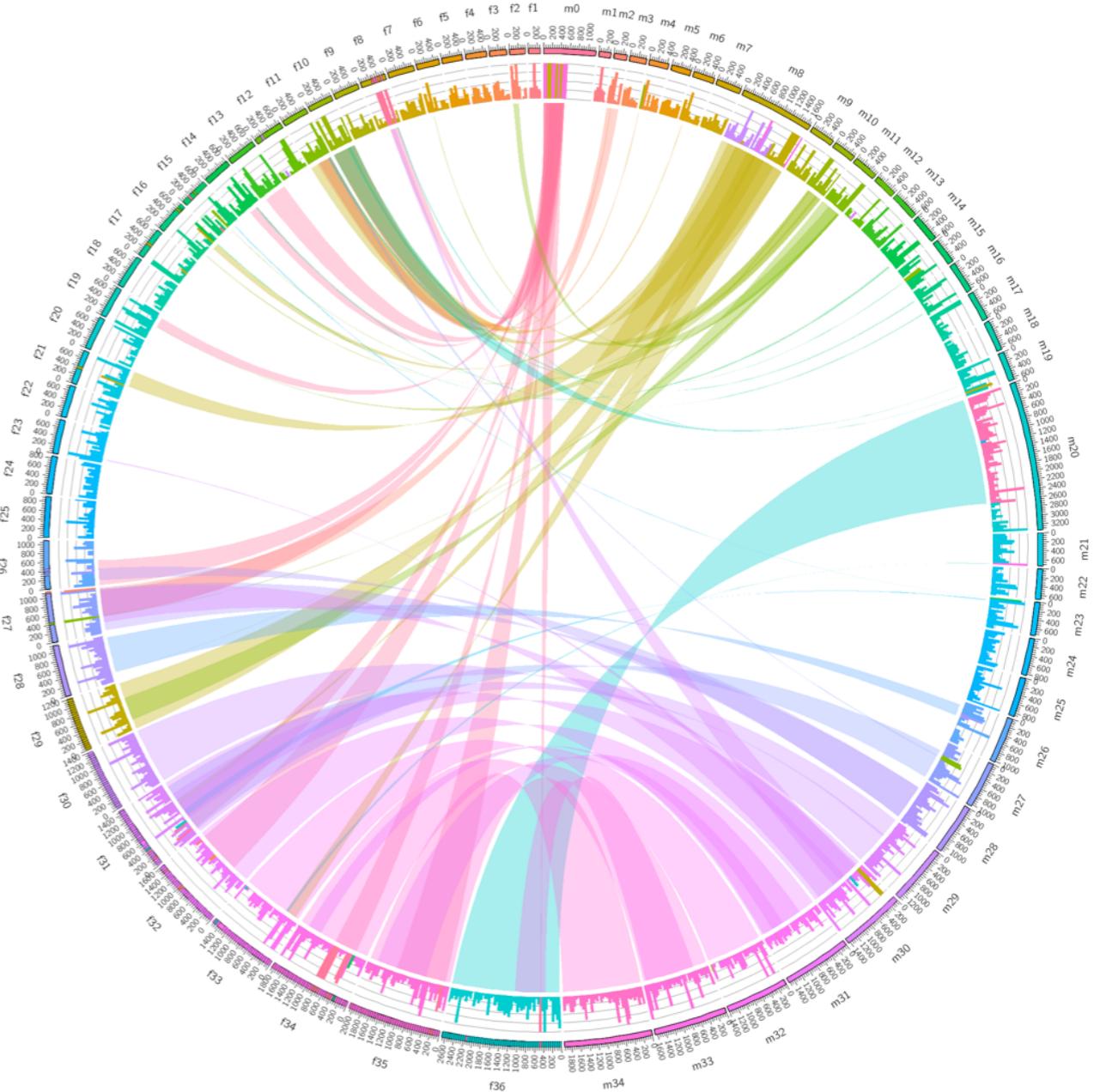


bundled links, min 5 genes per bundle  
max 1 Mb gap between links in a bundle

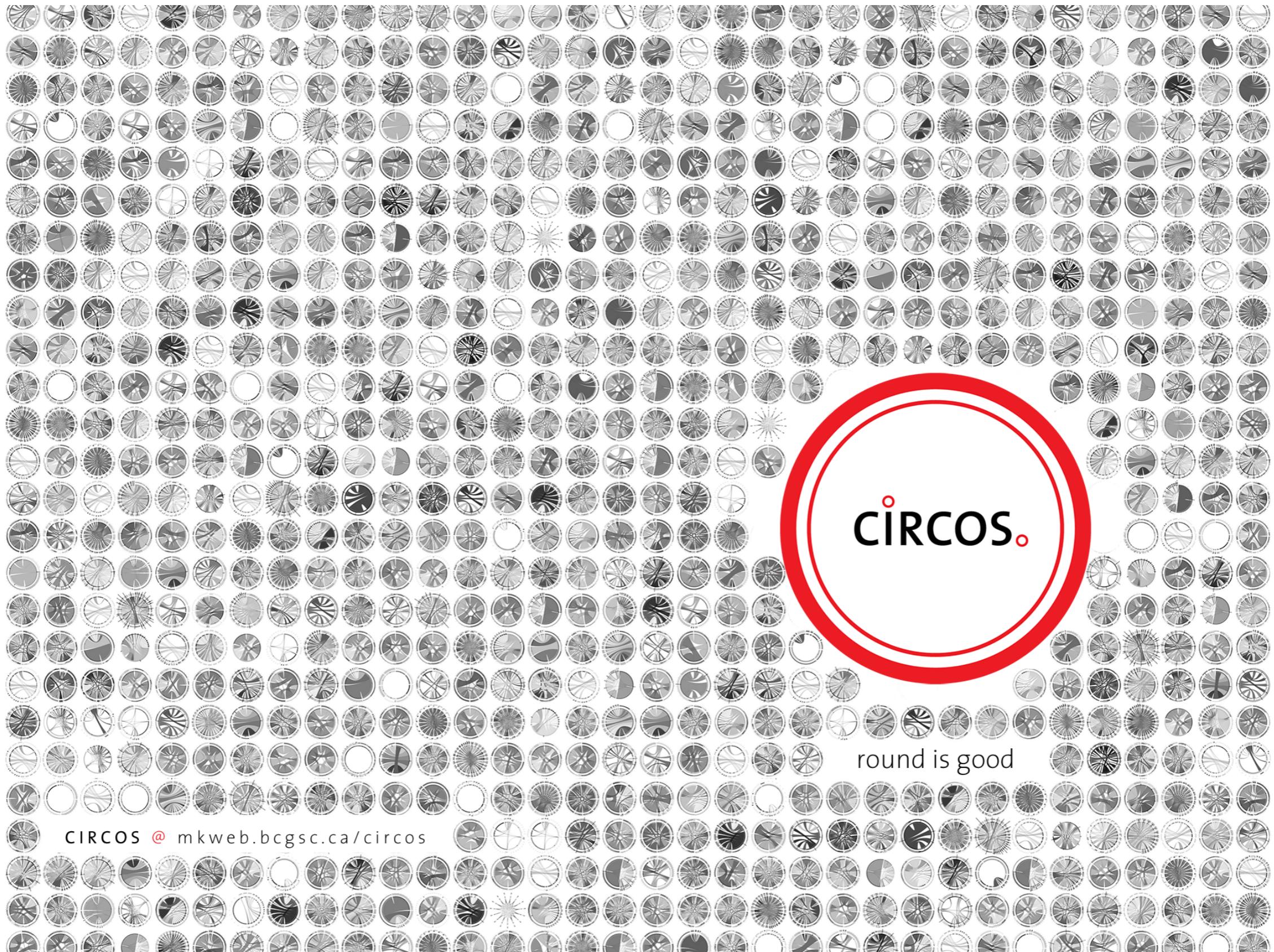
# ORIENTATION OF CHROMOSOMES AVOIDS TWISTS



L. major chromosomes (left) reversed



L. major chromosomes (left) not reversed



CIRCOS @ [mkweb.bcgsc.ca/circos](http://mkweb.bcgsc.ca/circos)