14H30 - 16H00

SESSION 3

DATA TRACKS

SESSION PLAN

image template from previous session

histograms

heat maps
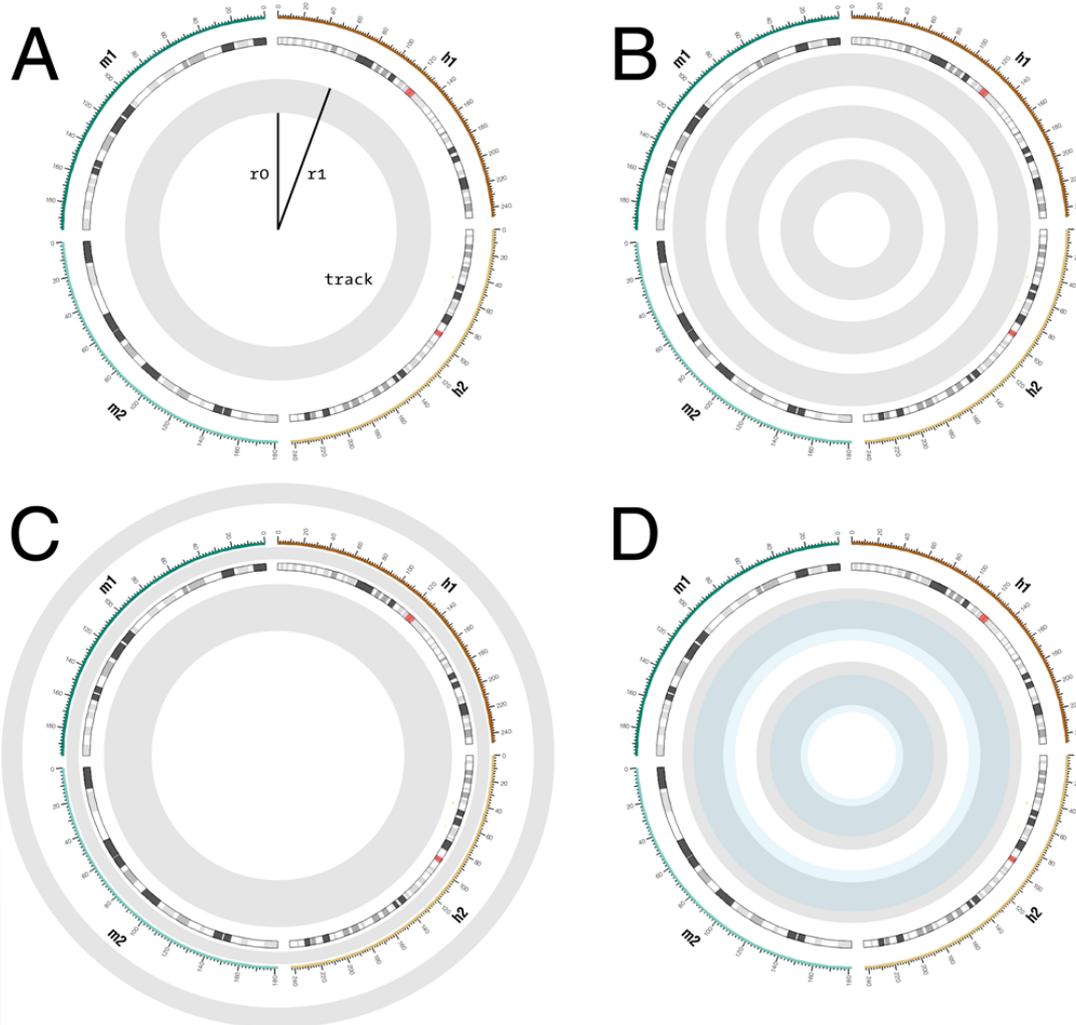
links

density and stacked histograms

tiles

highlights

introduction to rules

DATA TRACKS

A each data track confined to an annulus bounded by radii *r0* and *r1*, **B** any number of tracks can be placed on the figure, **C** tracks can be placed at any radial position, including inside/outside ideogram circle and inside/outside ticks, **D** tracks can be made to overlap and can be drawn in any order.

LESSON 1

IDEOGRAM, TICK, LABEL AND GRID LAYOUT

1 IMAGE TEMPLATE FROM PREVIOUS SESSION

sessions/3/1/circos.conf

```
karyotype =
   ../data/karyotype.human_mouse_labels.txt

chromosomes_units           = 1000000
chromosomes_display_default = no

chromosomes         = hs1;hs2;mm1;mm2
chromosomes_order   = hs1;hs2;mm2;mm1
chromosomes_reverse = mm2;mm1
chromosomes_scale   =
   hs2:1.018;mm1:1.254;mm2:1.360

<highlights>
<highlight>
file = ../data/highlight.txt
r0   = 1r+40p
r1   = 1r+45p
</highlight>
# additional <highlight> blocks will go here
</highlights>

# data tracks will go here
```
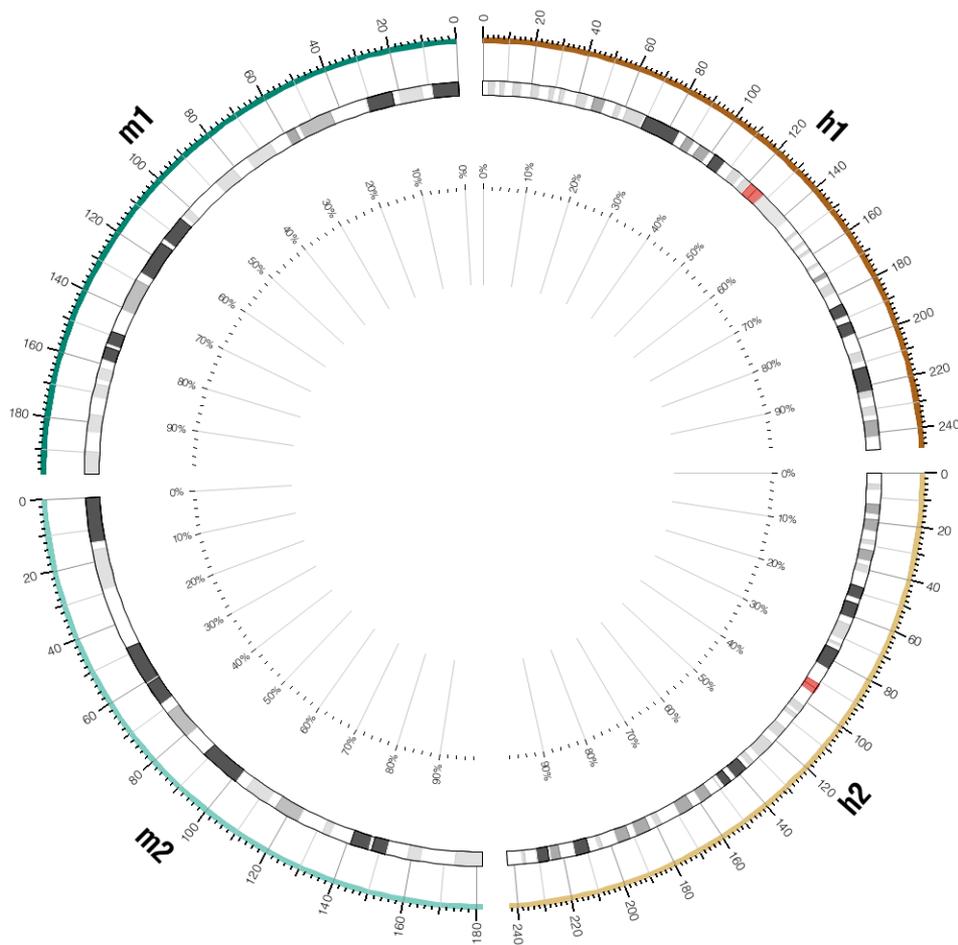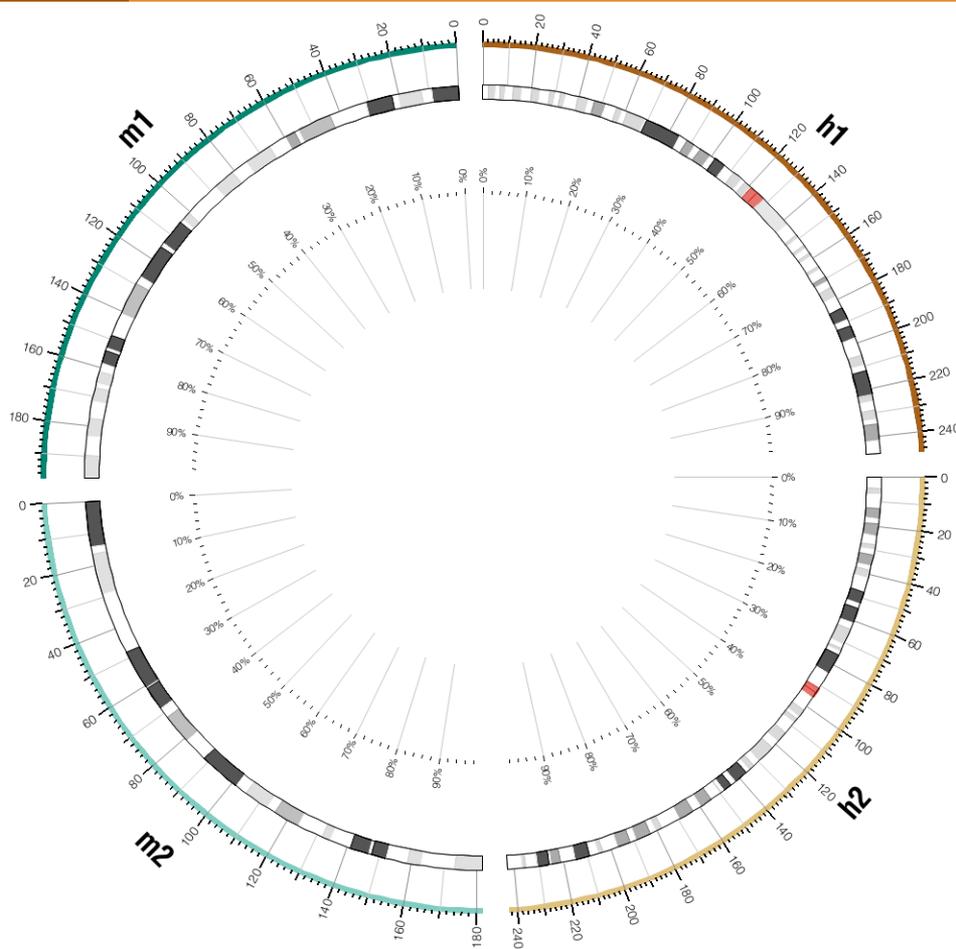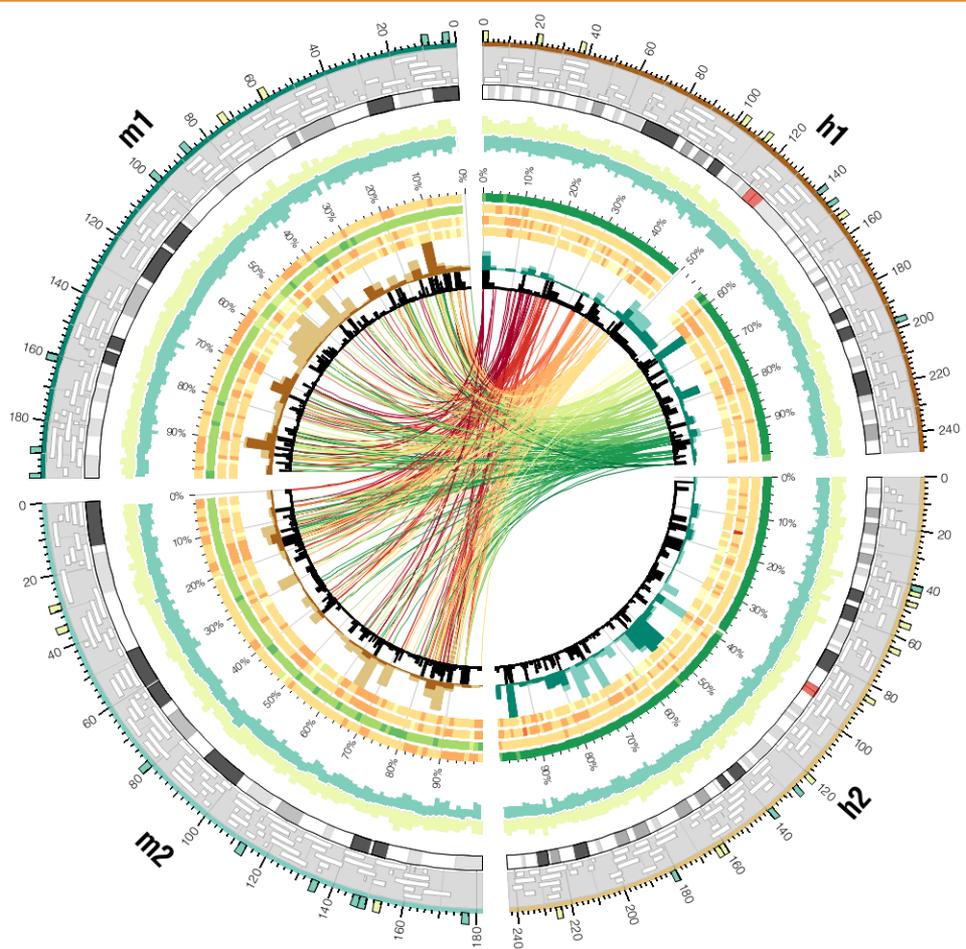

Ideogram, grid and tick layout that will be used for data tracks in this series of lessons. Human chr1 and chr2 (labeled h1 and h2) and mouse chr1 and chr2 (labeled m1 and m2) are shown. Note the orientation of the scale – clockwise for human and counter-clockwise for mouse ideograms.

1 SESSION IMAGE WITH DATA

We are starting with this image.

We will end up with this image.

LESSON 2

HISTOGRAMS

2 DATA TRACK DEFINITION

```
<plots>

<plot>
# data track definition

...
</plot>

<plot>
# another data track

...
</plot>


</plots>
```

All 2D data tracks are defined within `<plot>` blocks.
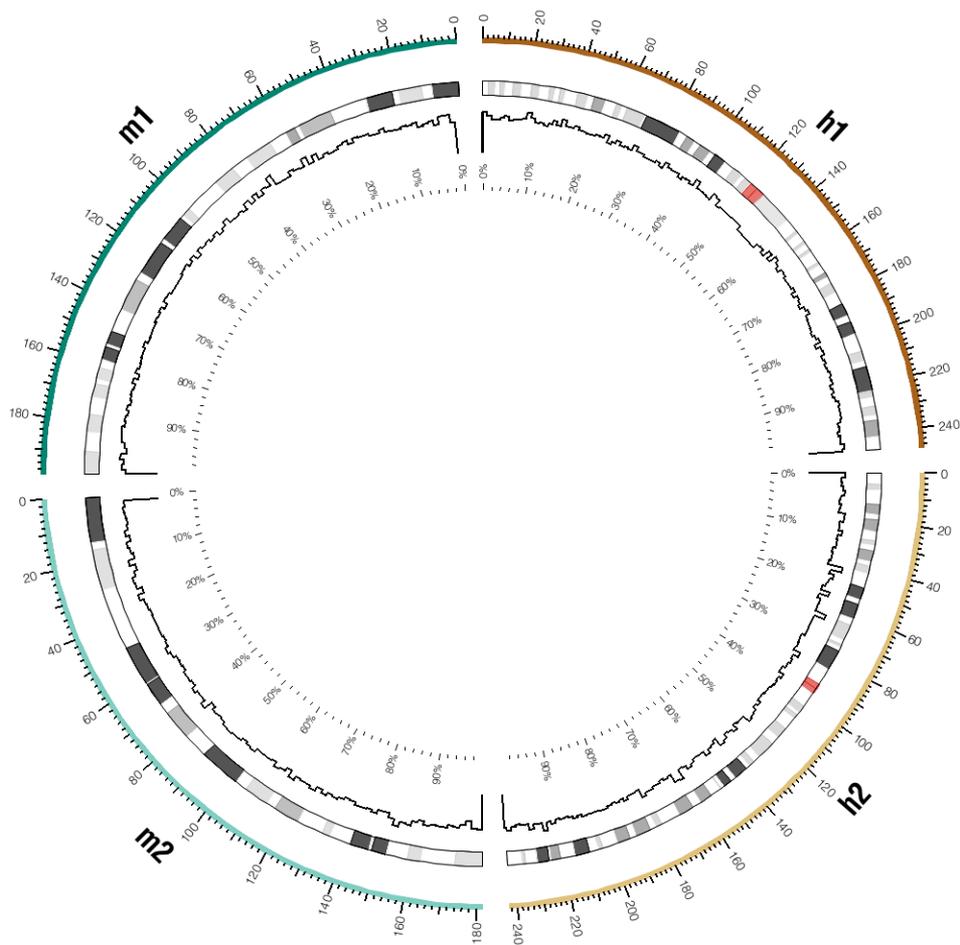
2   DRAWING A HISTOGRAM

sessions/3/2/etc/circos.conf

```
<plot>
type        = histogram
file        = ../data/both.cons.2e6.max.txt
min         = 0
max         = 1
r0          = 0.85r
r1          = 0.975r
color       = black
thickness   = 2p
#fill_under  = yes
#fill_color  = bs8-1
</plot>
```


A histogram data track located within the annulus bounded by *r0 = 0.85r* and *r1 = 0.975r*.

2 HISTOGRAM DATA FILE

sessions/3/data/both.cons.2e6.avg.txt

```
# chr start end value
...
mm1 10000000 11999999 0.439079022807017
mm1 12000000 13999999 0.44630188700565
mm1 14000000 15999999 0.452856295597484
...
```

To learn about the details of this data set, and all data tracks for all the lessons, see the *"What is shown in the figure"* section in the handout for the corresponding lesson.
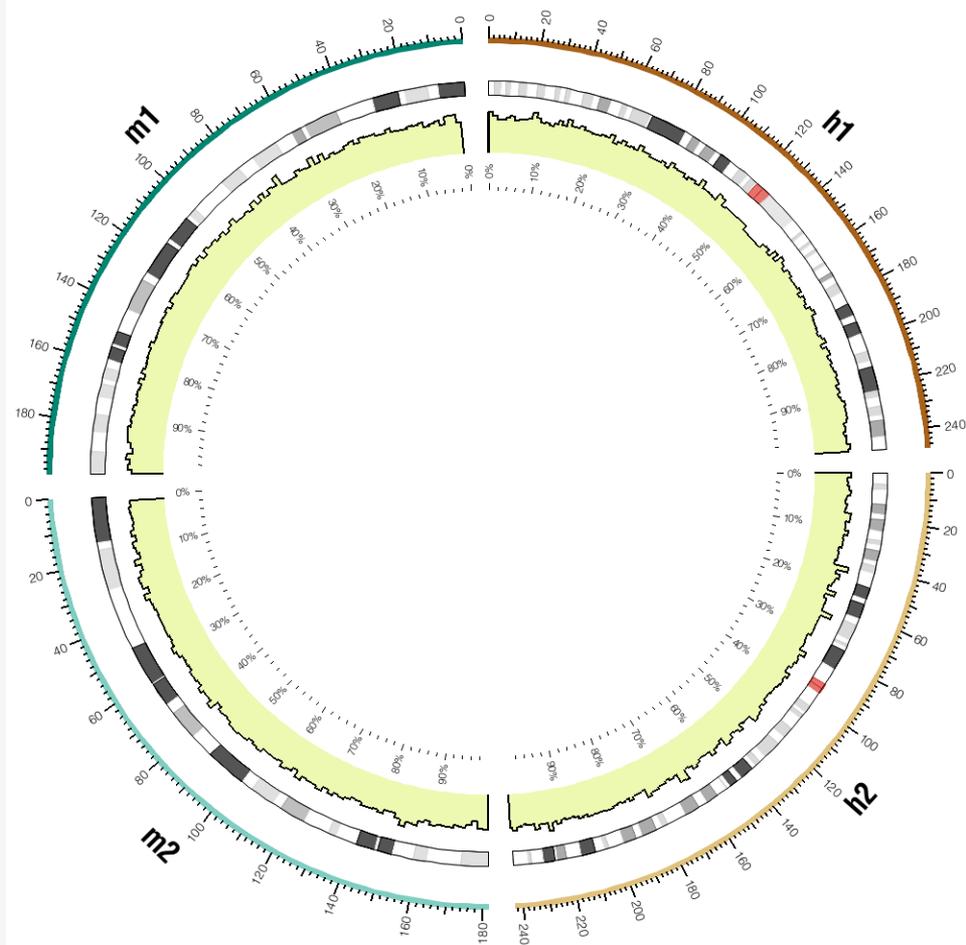
You can recreate the data for each session from the original raw data using

```
sessions/3/data/ucsc/create.tracks
```

2 FILLING HISTOGRAM BINS

```
fill_under  = yes
fill_color  = bs8-1
```

sessions/3/2/etc/circos.conf

The histogram has been adjusted to include a fill color, using `fill_under` and `fill_color` parameters.
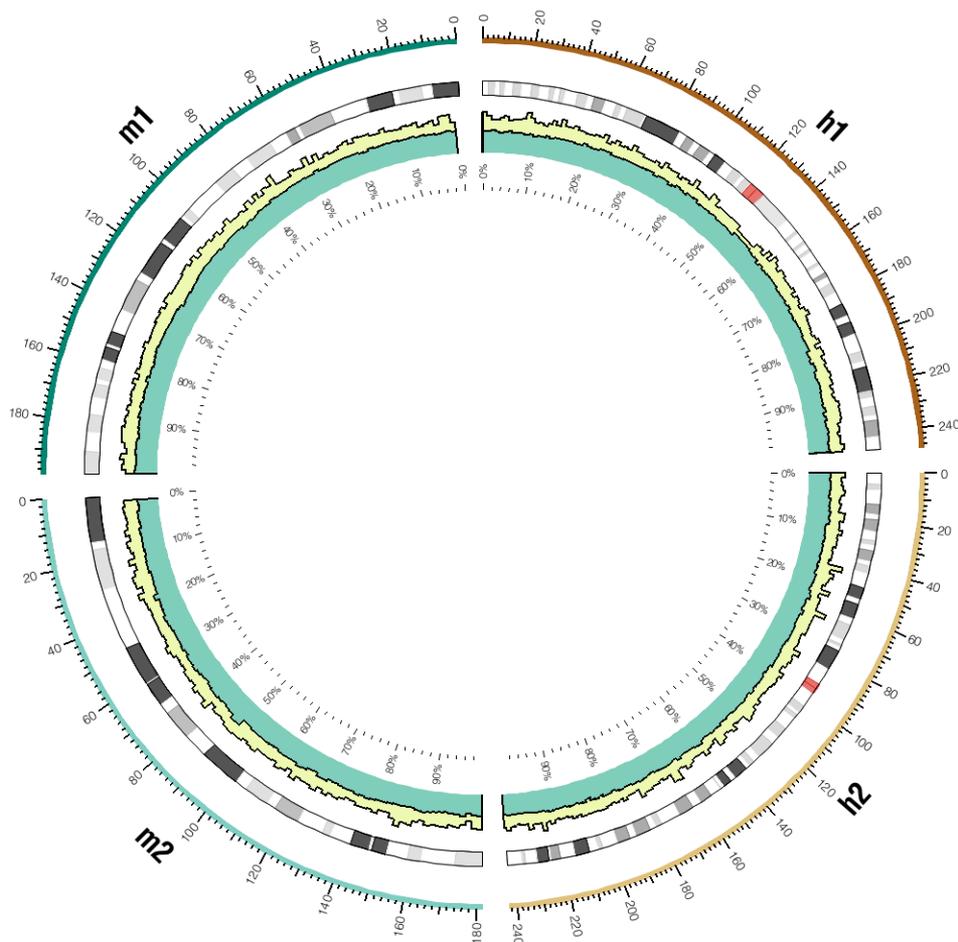
2 OVERLAPPING TWO HISTOGRAMS

sessions/3/2/etc/circos.conf

```
<plot>
show        = yes
type        = histogram
file        = ../data/both.cons.2e6.avg.txt
min         = 0
max         = 1
r0          = 0.85r
r1          = 0.975r
color       = black
#color       = white
thickness   = 2p
fill_under = yes
fill_color = bs8-2
z           = 5
</plot>
```


A second histogram has been added, drawn on top of the first histogram.
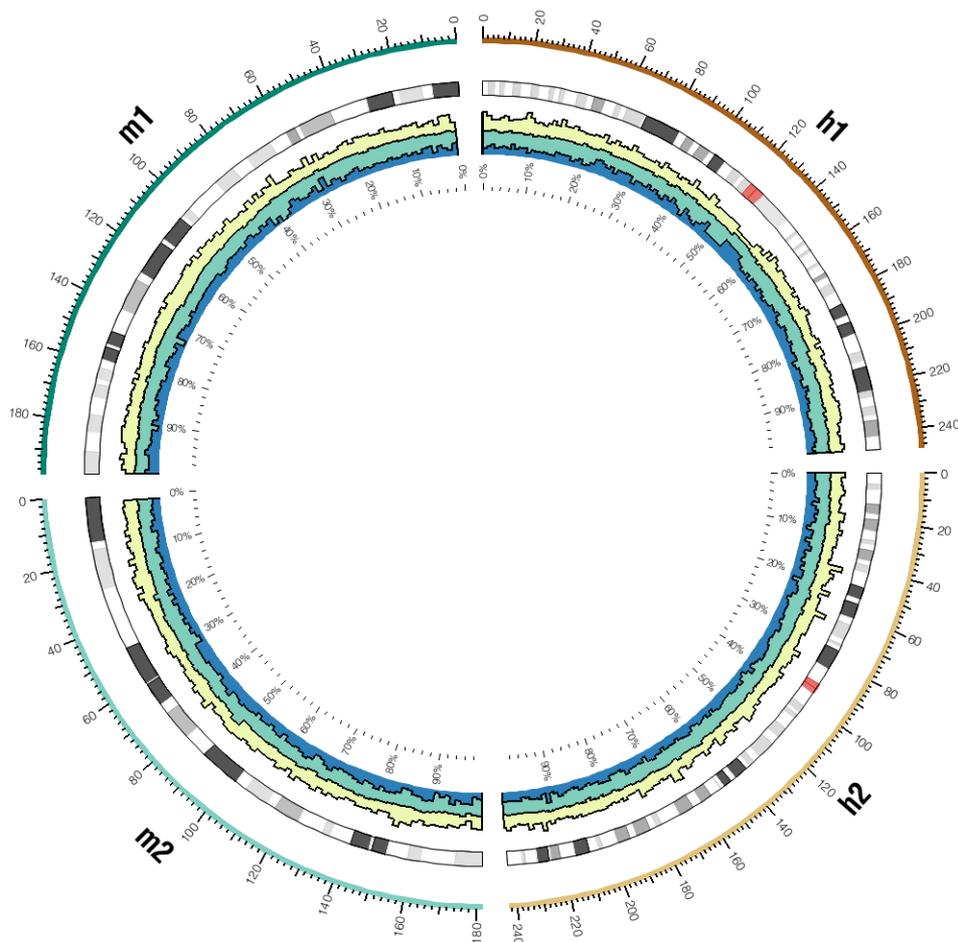
2 OVERLAPPING THREE HISTOGRAMS

sessions/3/2/etc/circos.conf

```
<plot>
show          = yes
type          = histogram
file          = ../data/both.cons.2e6.min.txt
min           = 0
max           = 1
r0            = 0.845r
r1            = 0.975r
color         = black
thickness     = 2p
fill_under    = yes
fill_color    = bs8-3
z             = 10
</plot>
```


A third histogram has been added, drawn on top of the first two histograms.

The region between the relative ticks and the ideogram now shows three histograms, each with a different fill color.

2 FORMATTING OVERLAPPING TRACKS FOR EFFECT

sessions/3/2/etc/circos.conf

```
<plot>
# comment out color and thickness to
# remove the outline on this histogram

#color       = black
#thickness   = 2p
</plot>

<plot>
# make the outline of this histogram white
color        = white
</plot>

<plot>
# comment out color and thickness
# to remove outline on this histogram
#color       = black
#thickness   = 2p
# make the fill color white
fill_under   = yes
fill_color   = white
</plot>
```
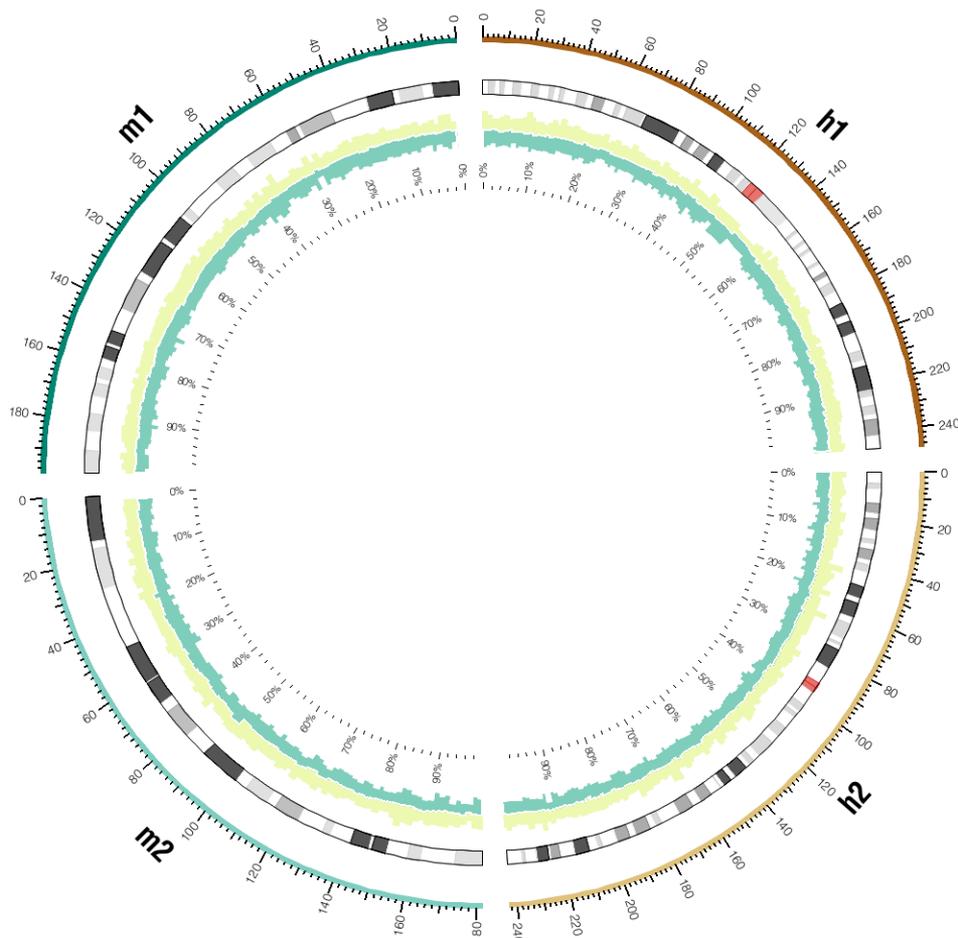

By removing the outline on all histograms and setting the fill of the bottom histogram to white, we can create a compound data tracks which effectively shows low/mid/high range.

LESSON 3
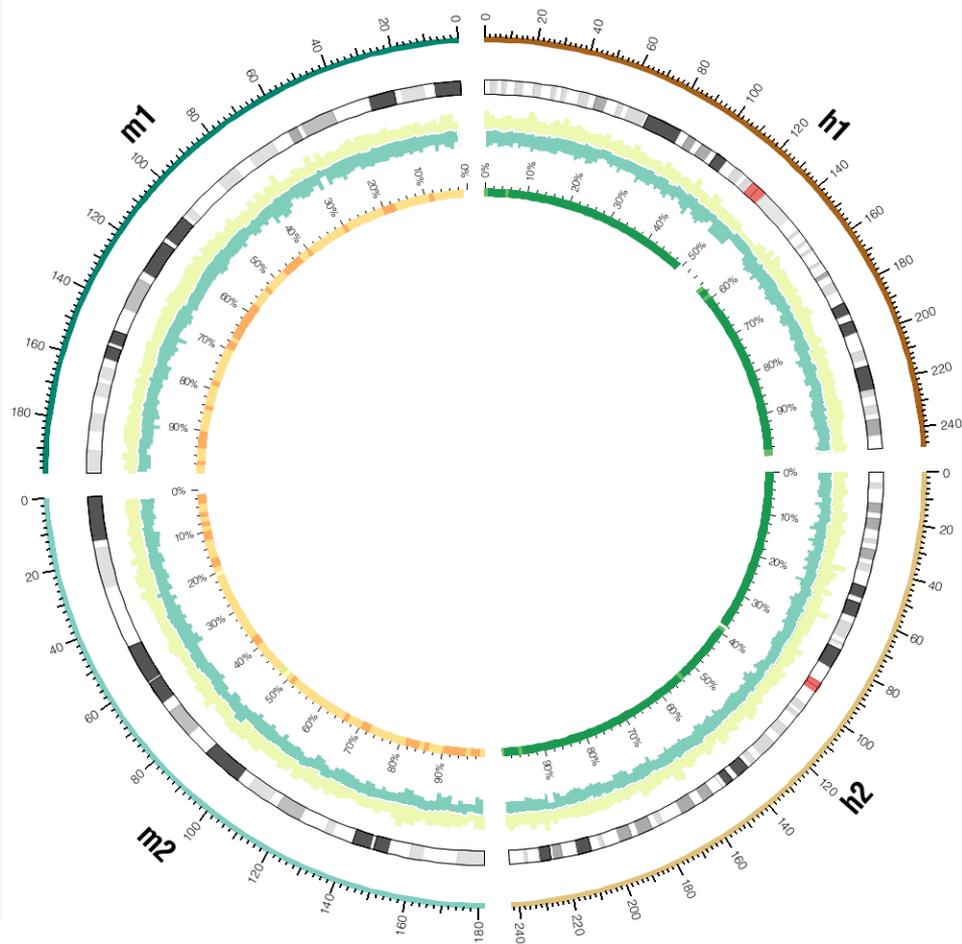
HEAT MAPS

3 HEAT MAPS

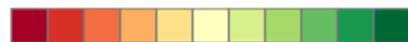sessions/3/3/etc/circos.conf

```
<plot>

type = heatmap

min  = 0.1

max  = 0.9

r0   = 0.73r

r1   = 0.75r

file = ../data/both.cons.2e6.rhe.avg.txt

color =
   bd3-1,bd3-2,bd3-3,bd3-4,bd3-5,bd3-6,bd3-7,
   bd3-8,bd3-9,bd3-10,bd3-11

</plot>
```

11-color Sequential Brewer Palette

red-yellow-green

bd3-1   bd3-11

http://www.colorbrewer.org

A heat map shows the average conservation score in 2Mb bins relative to the rhesus genome. Note that the human genome has a high degree of conservation with rhesus, since both are primates. Mouse, on the other hand, has a significantly lower conservation.

3 ADDING MORE HEAT MAPS

sessions/3/3/etc/circos.conf

```
<plot>
type = heatmap
min  = 0.1
max  = 0.9
r0   = 0.70r
r1   = 0.72r
file = ../data/both.cons.2e6.rn.avg.txt
color = bd3-1,bd3-2,bd3-3,bd3-4,bd3-5,bd3-6,bd3-7,
        bd3-8,bd3-9,bd3-10,bd3-11
</plot>
<plot>
type = heatmap
min  = 0.1
max  = 0.9
r0   = 0.67r
r1   = 0.69r
file = ../data/both.cons.2e6.danrer.avg.txt
color = bd3-1,bd3-2,bd3-3,bd3-4,bd3-5,bd3-6,bd3-7,
        bd3-8,bd3-9,bd3-10,bd3-11
</plot>
<plot>
type = heatmap
min  = 0.1
max  = 0.9
r0   = 0.64r
r1   = 0.66r
file = ../data/both.cons.2e6.fr.avg.txt
color = bd3-1,bd3-2,bd3-3,bd3-4,bd3-5,bd3-6,bd3-7,
        bd3-8,bd3-9,bd3-10,bd3-11
</plot>
```
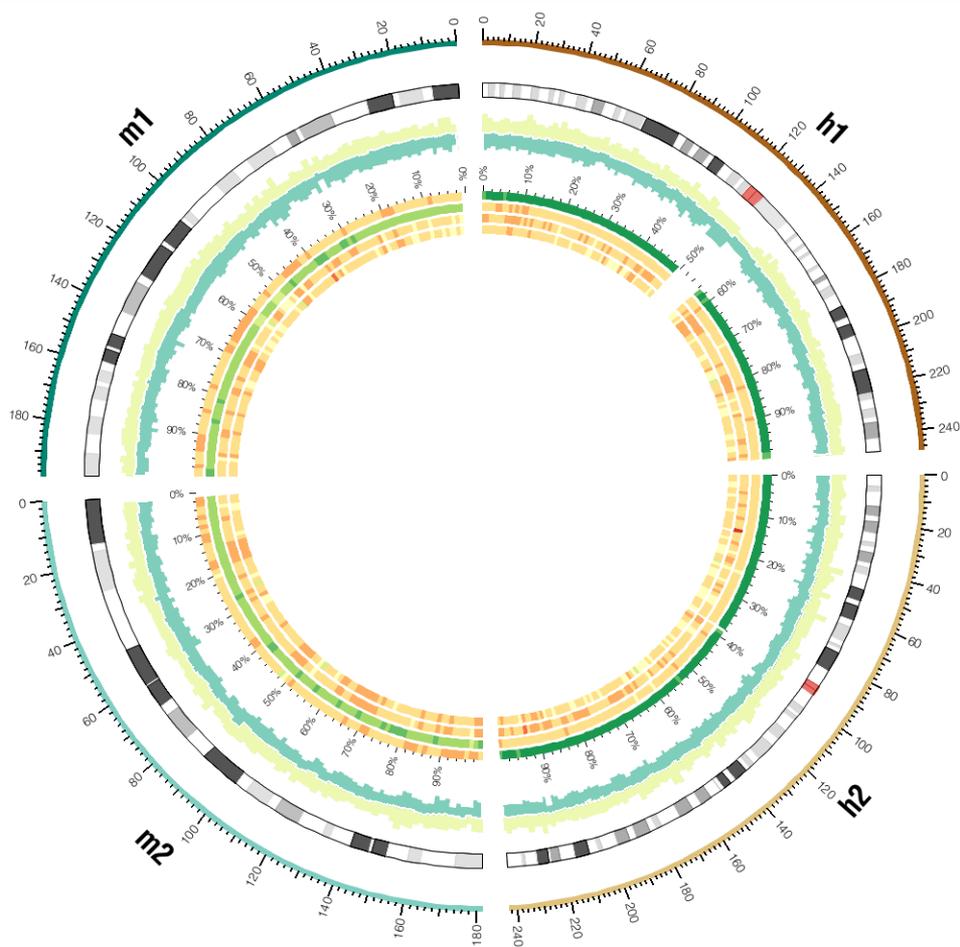

The heat maps show the average conservation score in 2Mb bins relative to the rhesus (outer heat map), rat, zebra fish and fugu (inner heat map) genomes. The human genome has a high conservation with rhesus, since both are primates. Mouse, on the other hand, has high conservation with rat (second heat map).
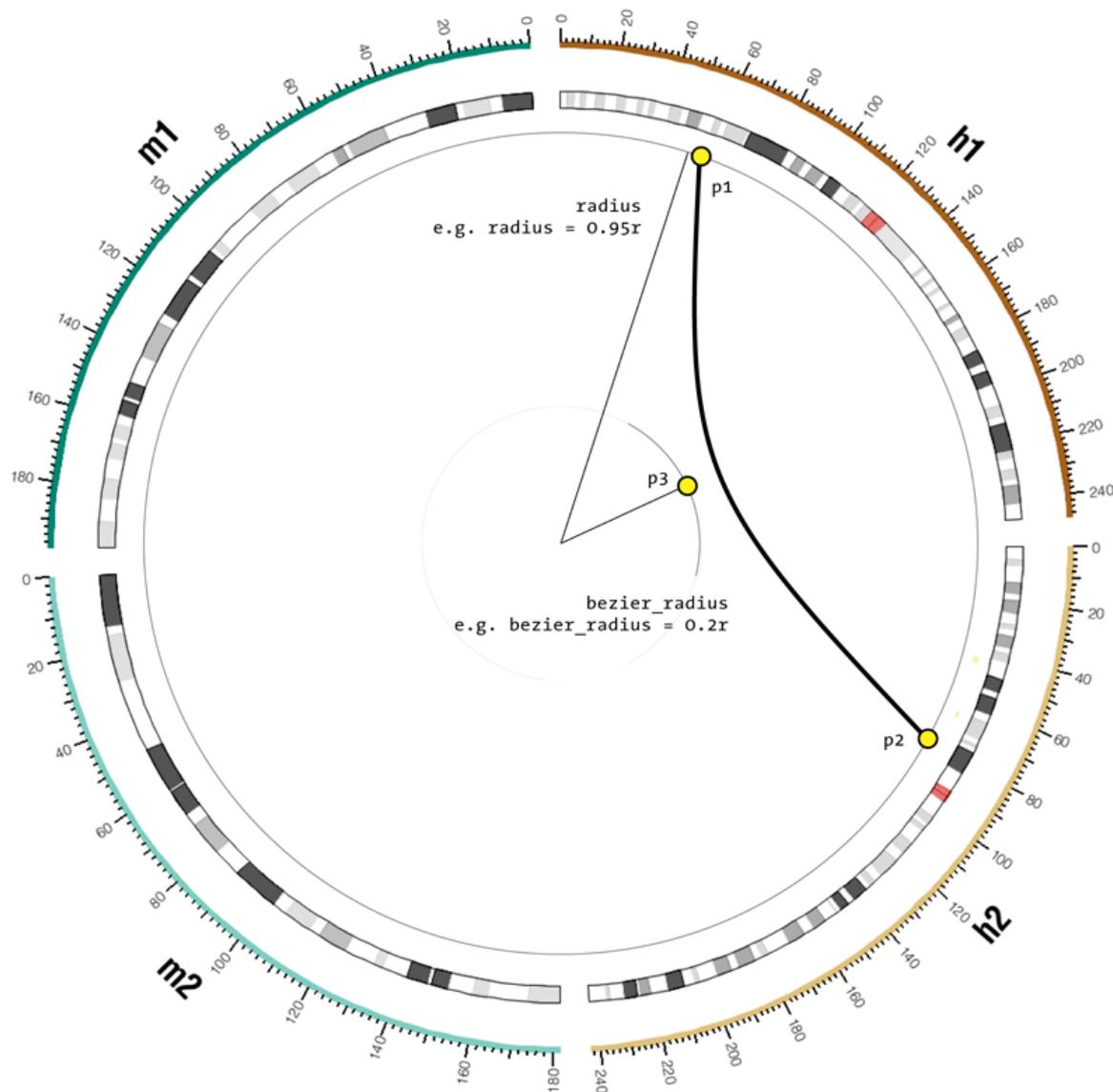
3   LOGARITHMIC COLOR MAPPING
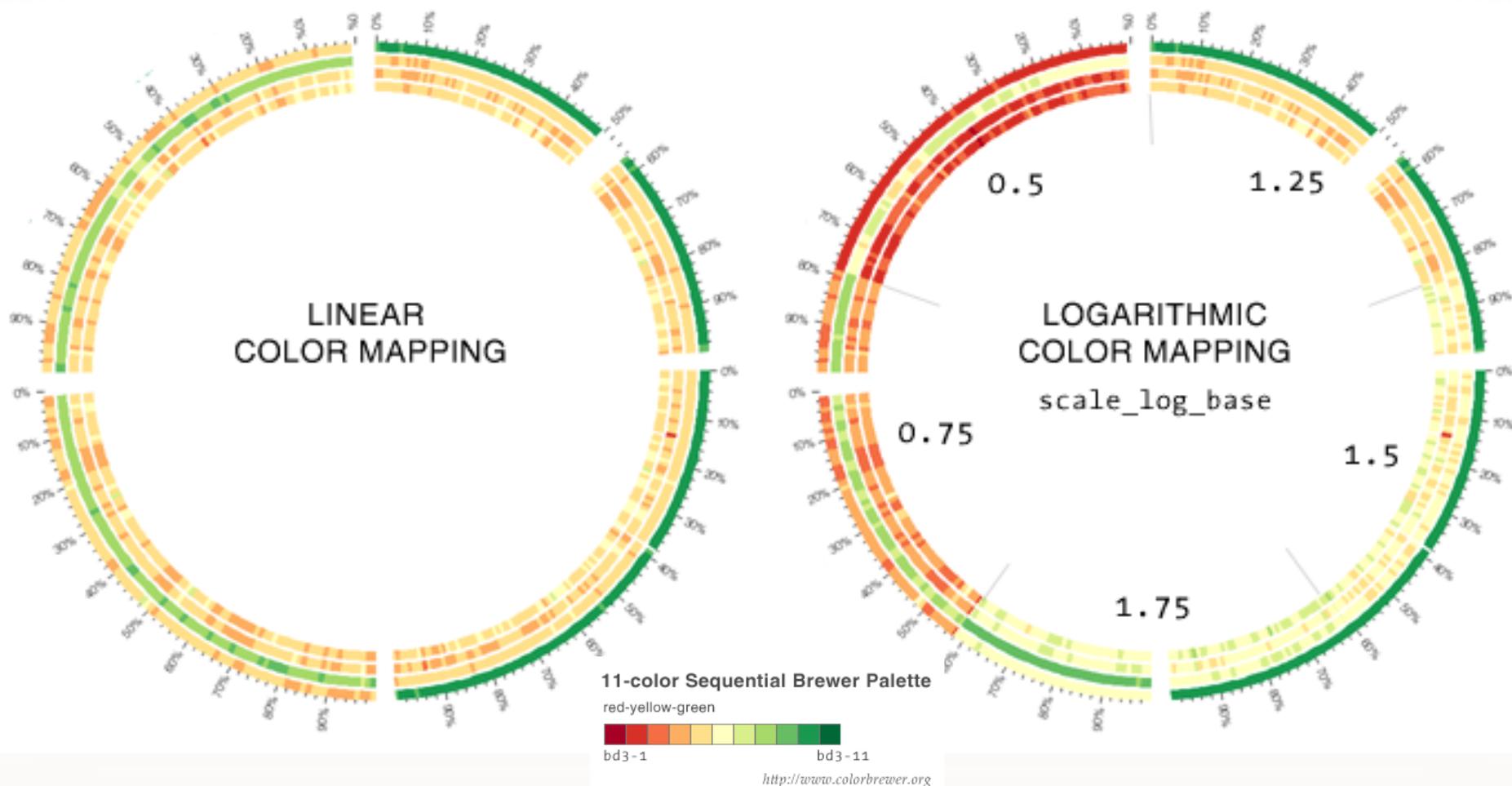
Color mapping in a heat map can be altered from its default linear mapping (**left**) to logarithmic (**right**) using the `scale_log_base` parameter.

LESSON 4

LINKS

4 QUADRATIC BEZIER CURVES

Links in Circos are drawn as quadratic Bezier curves which are specified by two ends (*p1*, *p2*) and a control point (*p3*).

The control point determines the direction of the curve at its ends and is placed midway (in angle) between the ends.
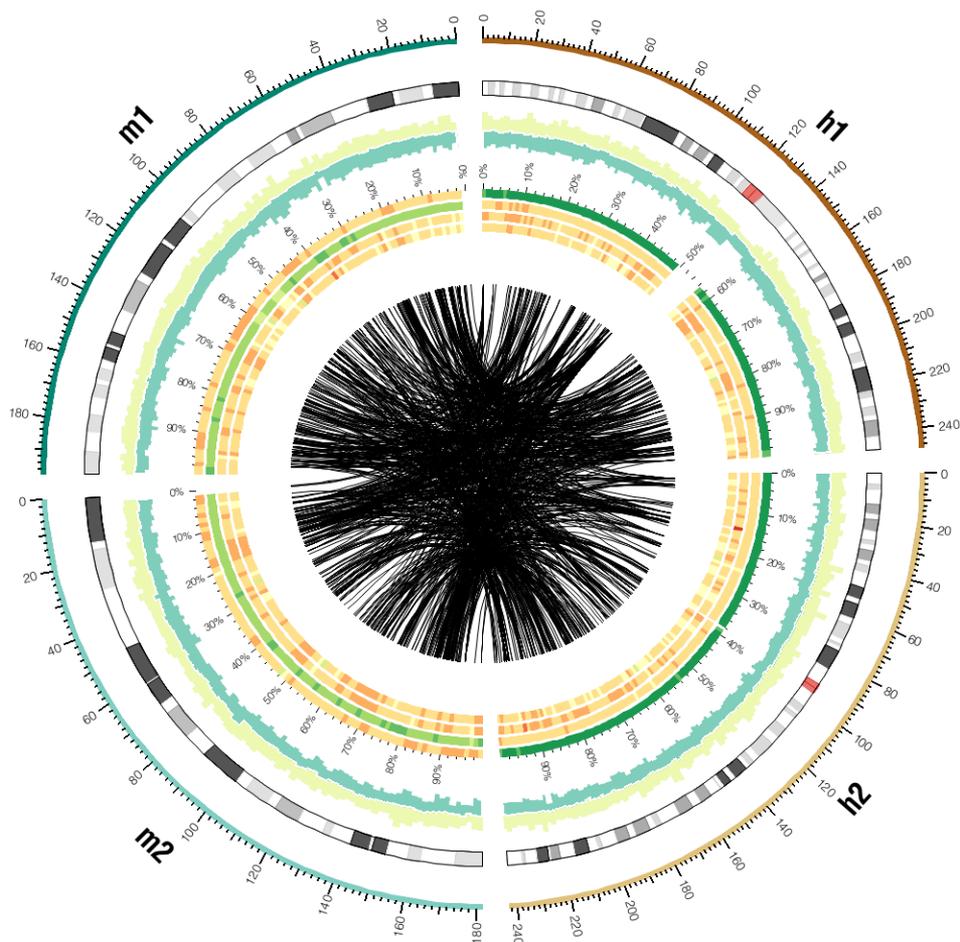
4 DRAWING LINKS

sessions/3/4/etc/circos.conf

```
<links>

<link chain>

file           = ../data/links.txt
# radial position of control point
bezier_radius = 0r
# radial position of link ends
radius         = 0.5r
thickness      = 1p
color          = black
#color          = black_a5

</link>

</links>
```


Links connecting regions with high sequence similarity between human and mouse chromosomes.
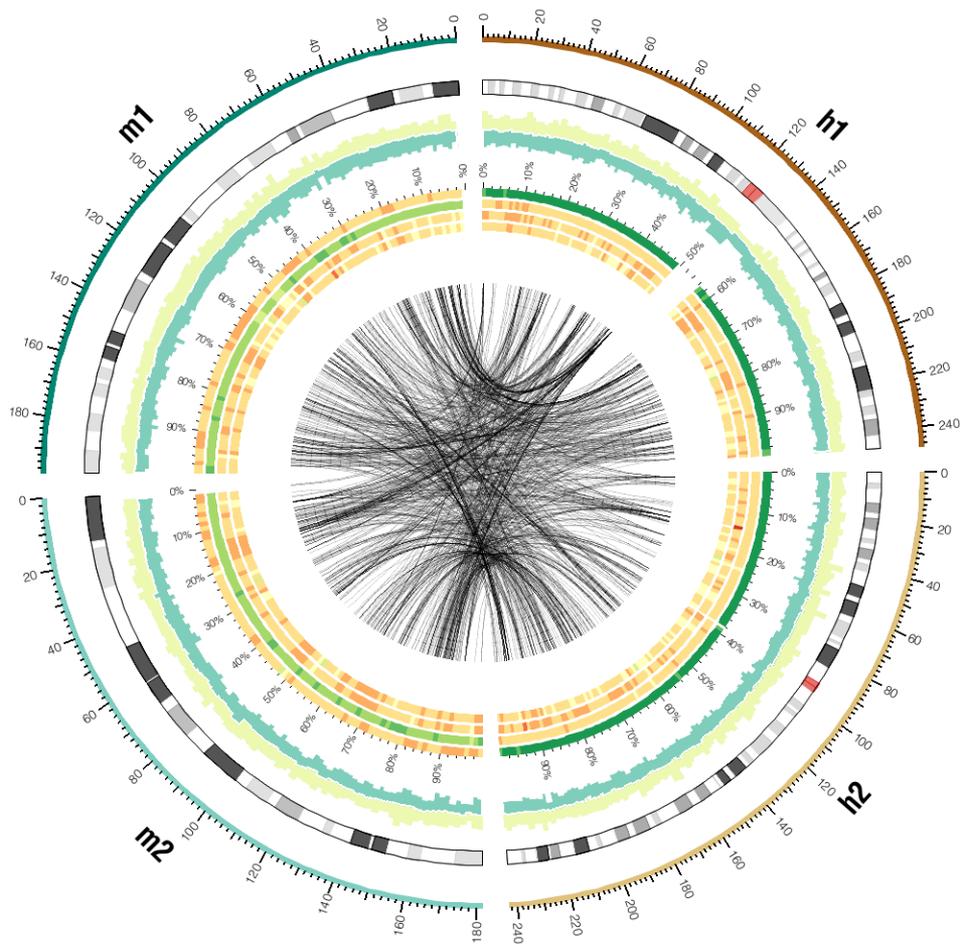
4   LINK DATA FILE

sessions/3/data/links.txt

```
# ID chr start end
link7572 hs1 120788758 120834144
link7572 mm1 63903740 63957877
link54010 hs2 190247704 190320005
link54010 mm1 124448312 124506291
...
```

4 ADDING TRANSPARENCY TO LINKS
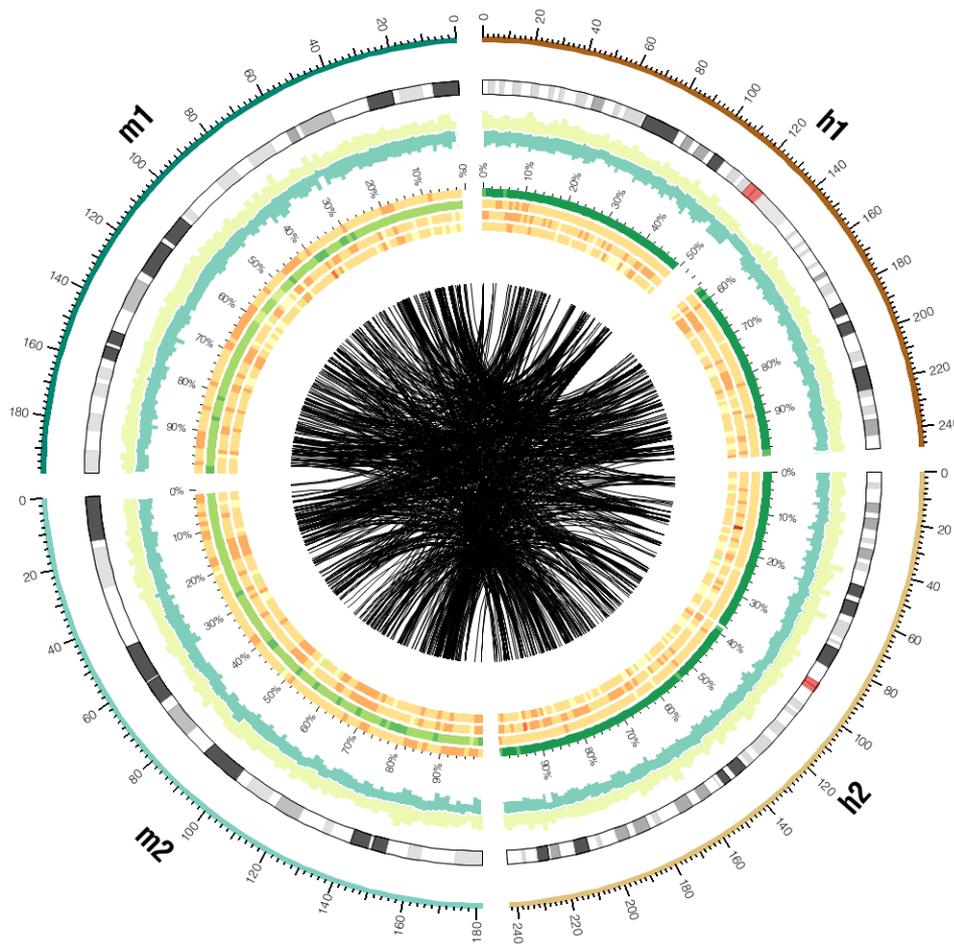
sessions/3/4/etc/circos.conf

```
#color          = black
color           = black_a5
```
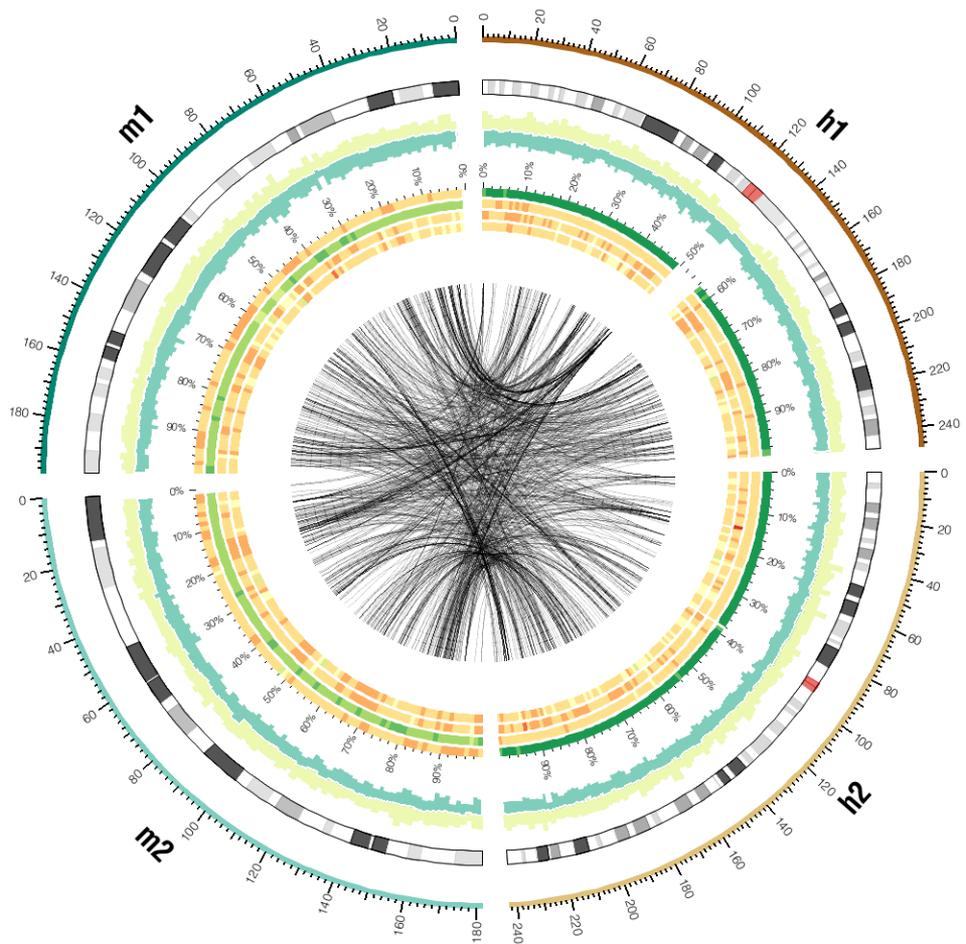

By making the link curves transparent, patterns in the data can be more easily discerned when the links are dense.

4   ADDING TRANSPARENCY TO LINKS

Links drawn without transparency.

By making the link curves transparent, patterns in the data can be more easily discerned when the links are dense.

4   TRANSPARENCY

```
# explicitly defined in color.conf
red     = 255,0,0 # fully opaque


# implicitly defined, once auto_alpha_colors=yes
#
# if, auto_alpha_steps=5
red_a1 # 16% transparent
red_a2 # 33% transparent
red_a3 # 50% transparent
red_a4 # 62% transparent
red_a5 # 87% transparent
```

The number of transparency steps depends on `auto_alpha_steps` parameter in

`sessions/etc/image.conf`

A reasonable default is 5.

You do not have to define transparent colors. They are defined for you automatically.

LESSON 5

DENSITY AND STACKED HISTOGRAMS

5   STACKED HISTOGRAMS

sessions/3/data/links.density.stacked.txt

```
hs2 5000000 9999999 1.0000,0.0000,0.0000,1.0000
hs2 10000000 14999999 1.0000,0.0000,0.0000,2.0000
...
```

A stacked histogram data file is analogous to a histogram data file, but the value is now a comma-separated list.

5   STACKED HISTOGRAMS

sessions/3/5/etc/circos.conf

```
<plot>
type        = histogram
min         = 0
max         = 20
r0          = 0.5r
r1          = 0.55r
file        = ../data/link.density.txt
fill_under  = yes
fill_color  = black
</plot>


<plot>
type        = histogram
min         = 0
max         = 40
r0          = 0.55r
r1          = 0.65r
file        =
  ../data/link.density.stacked.txt
fill_under  = yes
fill_color  = bd2-4,bd2-1,bd2-2,bd2-3
</plot>
```
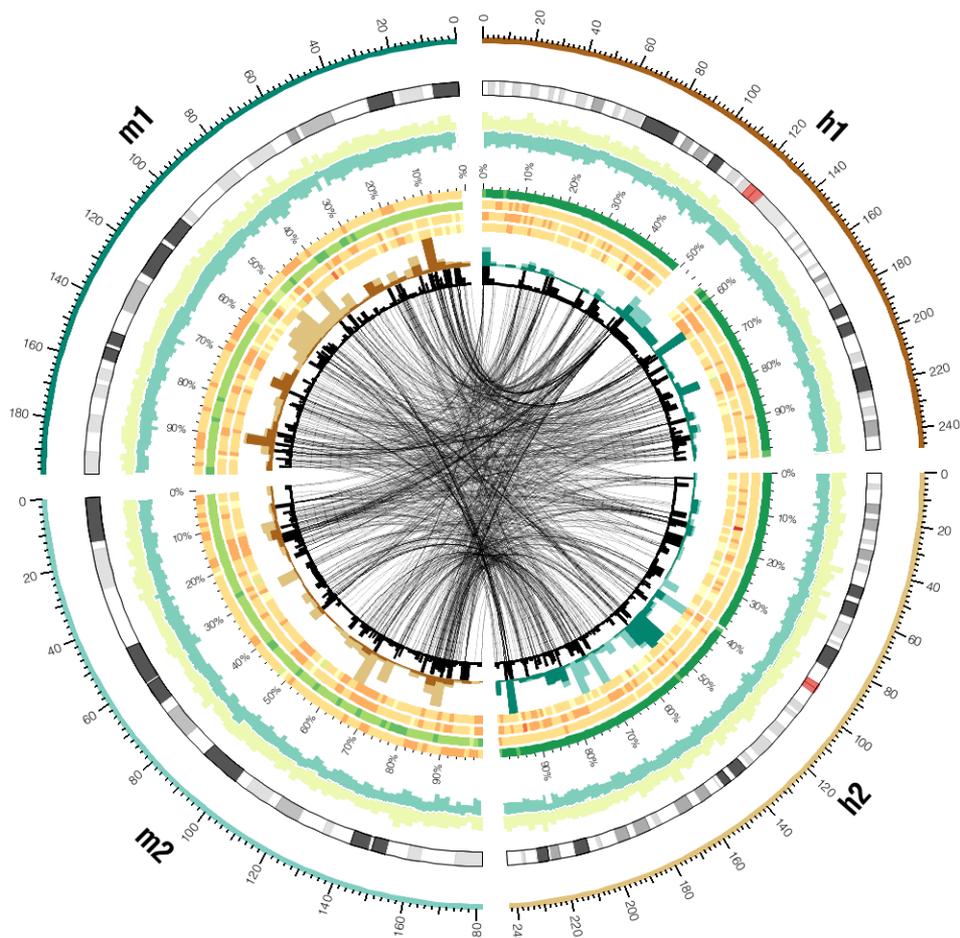

The stacked histograms were created with the `binlinks` tool.
The inner histogram (black) shows the number of link ends
within a 1Mb bin. The outer histogram (multi-colored, 5Mb bin)
shows the total size of links outgoing from a bin for each target
chromosome.

5 INTERPRETING DENSITY HISTOGRAMS
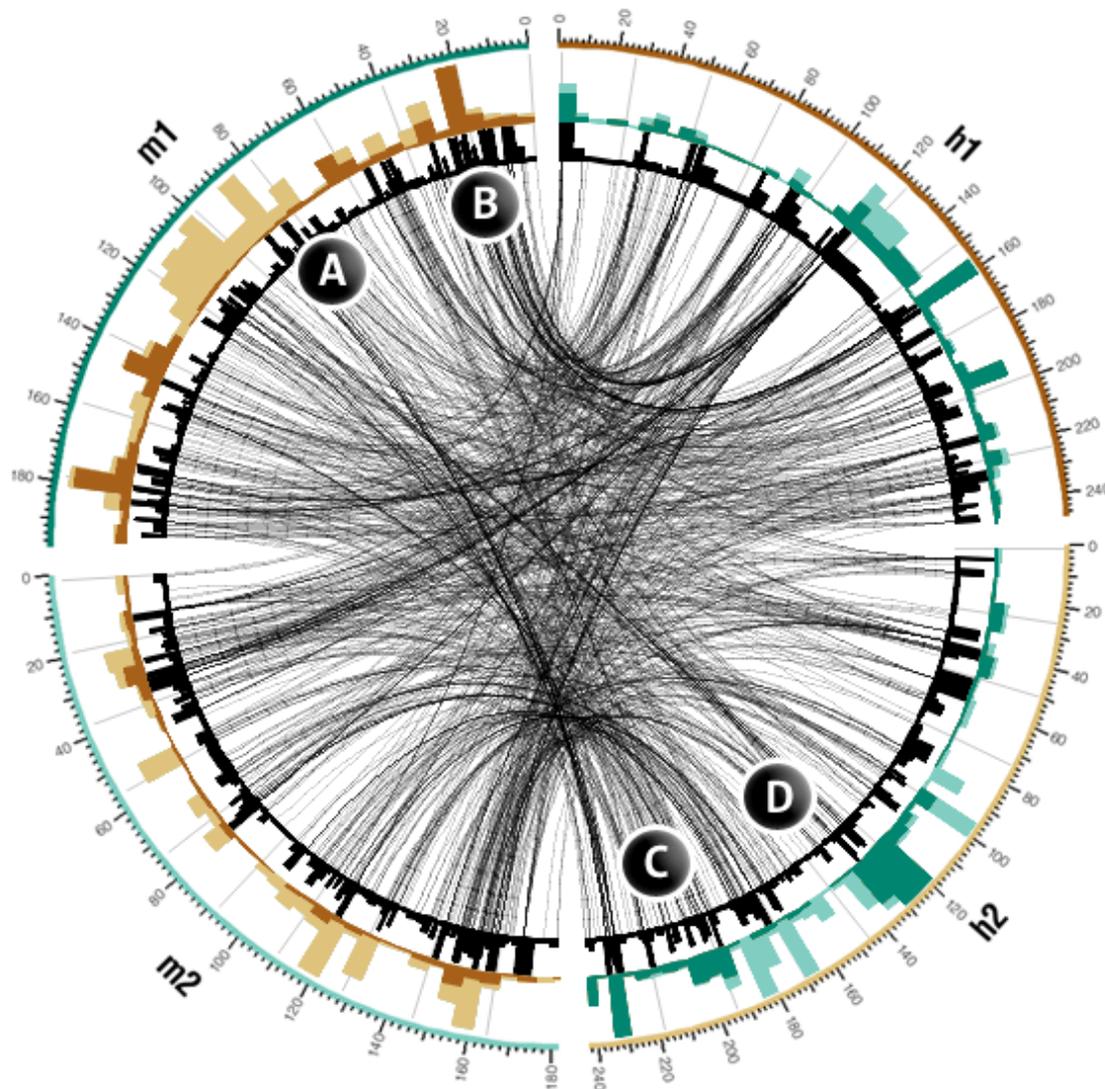
The stacked histograms were created with the `binlinks` tool. See handout for details.

The inner histogram (black) shows the number of link ends within a 1Mb bin.

The outer histogram (multi-colored, 5Mb bin) shows the total size of links outgoing from a bin for each target chromosome.

LESSON 6

TILES

6 TILES

sessions/3/6/etc/circos.txt

```
<plot>
type    = tile
file    = ../data/tiles.txt
r0      = 1r+3p
r1      = 1r+40p


layers                = 5
layers_overflow       = hide
layers_overflow_color = red


margin    = 1u
thickness = 5
padding   = 2


orientation = out


color            = white
stroke_thickness = 1
stroke_color     = dgrey


background = yes
background_color = grey_a2


</plot>
```
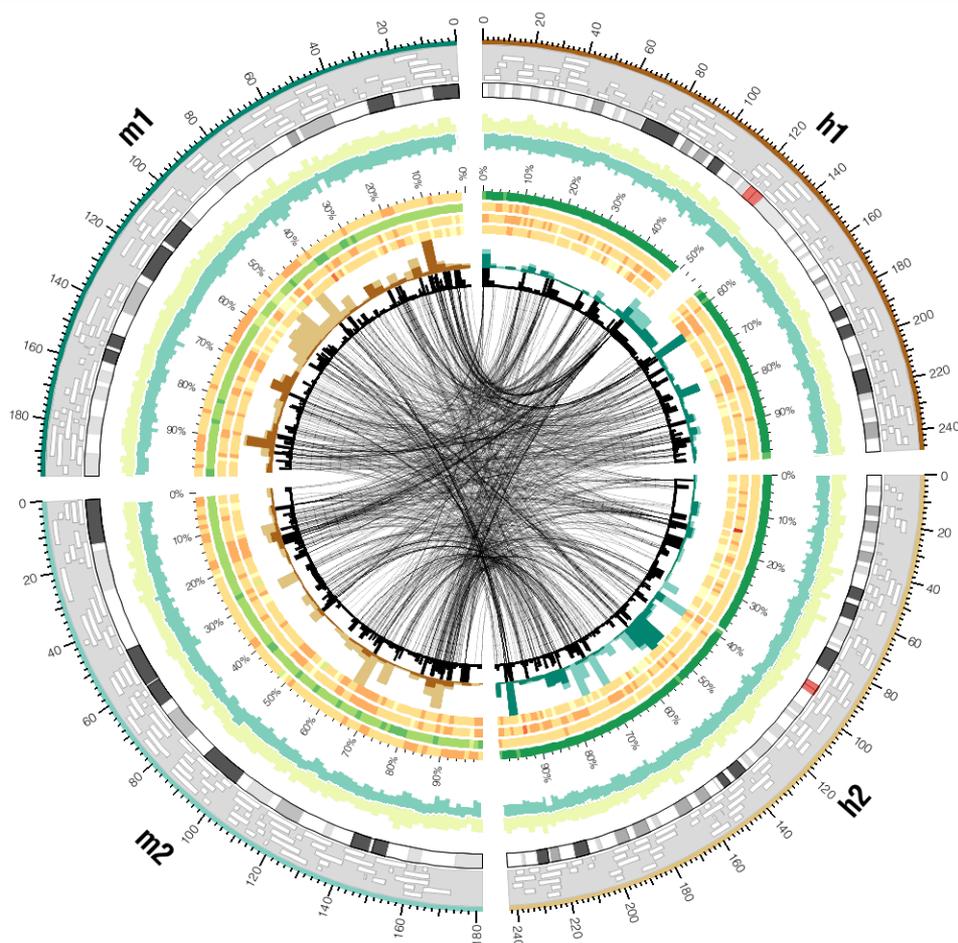

A tile track is shown outside the ideogram circle.

Using `background_*` parameters, the track is given a grey background.
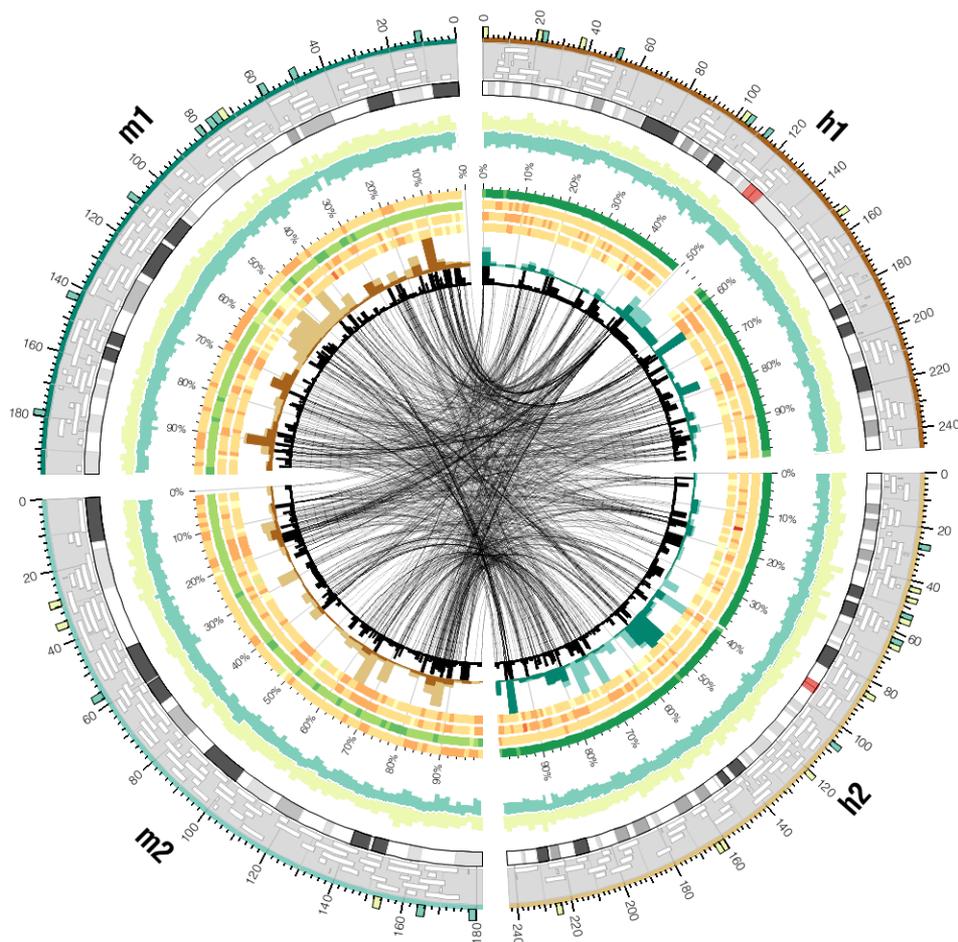
LESSON 7

HIGHLIGHTS

7 HIGHLIGHTS

sessions/3/7/etc/circos.conf

```
<highlight>
file = ../data/scatter.max.top20.txt
r0    = 1r+45p
r1    = 1r+57p
fill_color        = bs8-1
stroke_thickness = 1p
stroke_color      = black
</highlight>

<highlight>
file = ../data/scatter.min.top20.txt
r0    = 1r+45p
r1    = 1r+57p
fill_color        = bs8-2
stroke_thickness = 1p
stroke_color      = black
</highlight>
```


Two highlight tracks have been added to indicate the position of regions with very low and very high conservation.

The grid has been also toggled on. Notice how the transparent background of the tile track allows the grid to show through.

LESSON 8

INTRODUCTION TO RULES

8 RULES

sessions/3/8/etc/circos.conf

```
<links>

<link chain>

...

<rules>
<rule>
# each link is tested with a rule
# if the condition passes...
condition    = _CHR1_ eq "hs1"
# the color is changed to bd2-1
color        = bd2-1
</rule>
</rules>

</link>

</links>
```
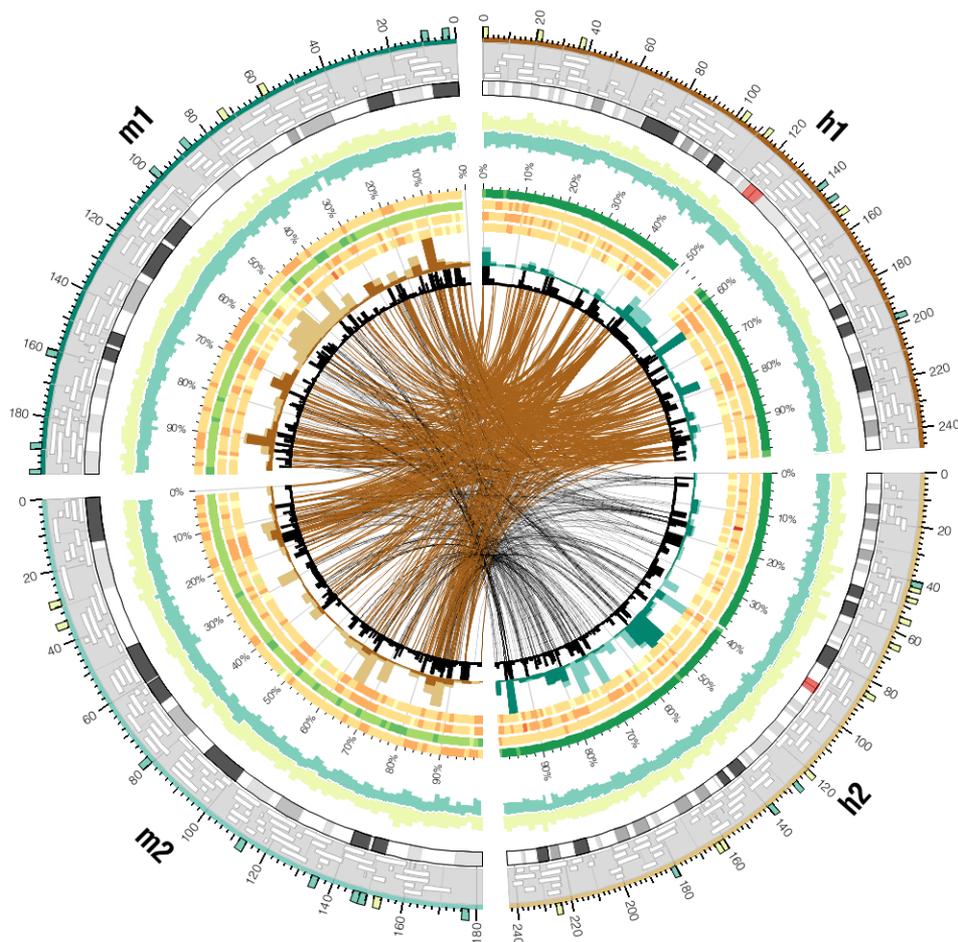

Using a rule, all links that start on hs1 are colored brown.

8 MULTIPLE RULES

sessions/3/8/etc/circos.conf

```
<rule>
importance = 100
condition  = _CHR1_ eq "hs1"
color      = bd2-1
</rule>

# any link that passed the
# above rule is not tested further

<rule>
importance = 90
# this rule always matches,
# since its condition is always true
# any link that failed the previous
# rule is matched by this one
condition  = 1
# the link will be hidden
show       = no
</rule>
```
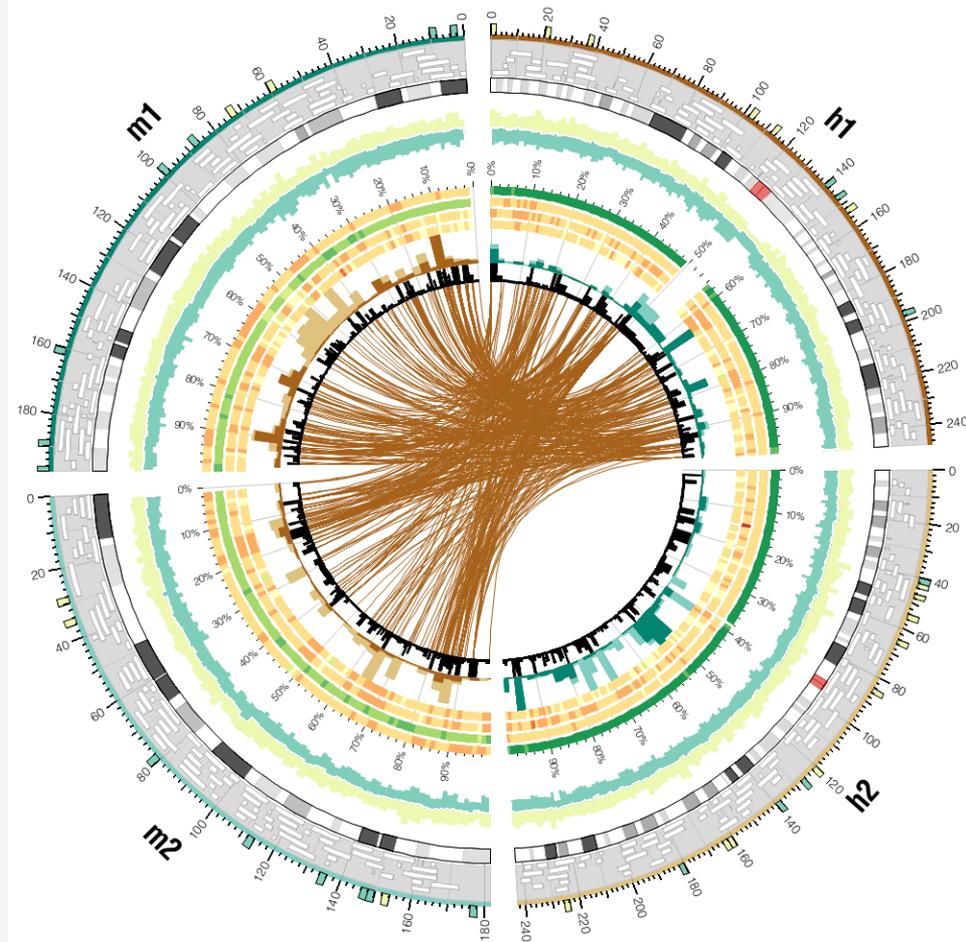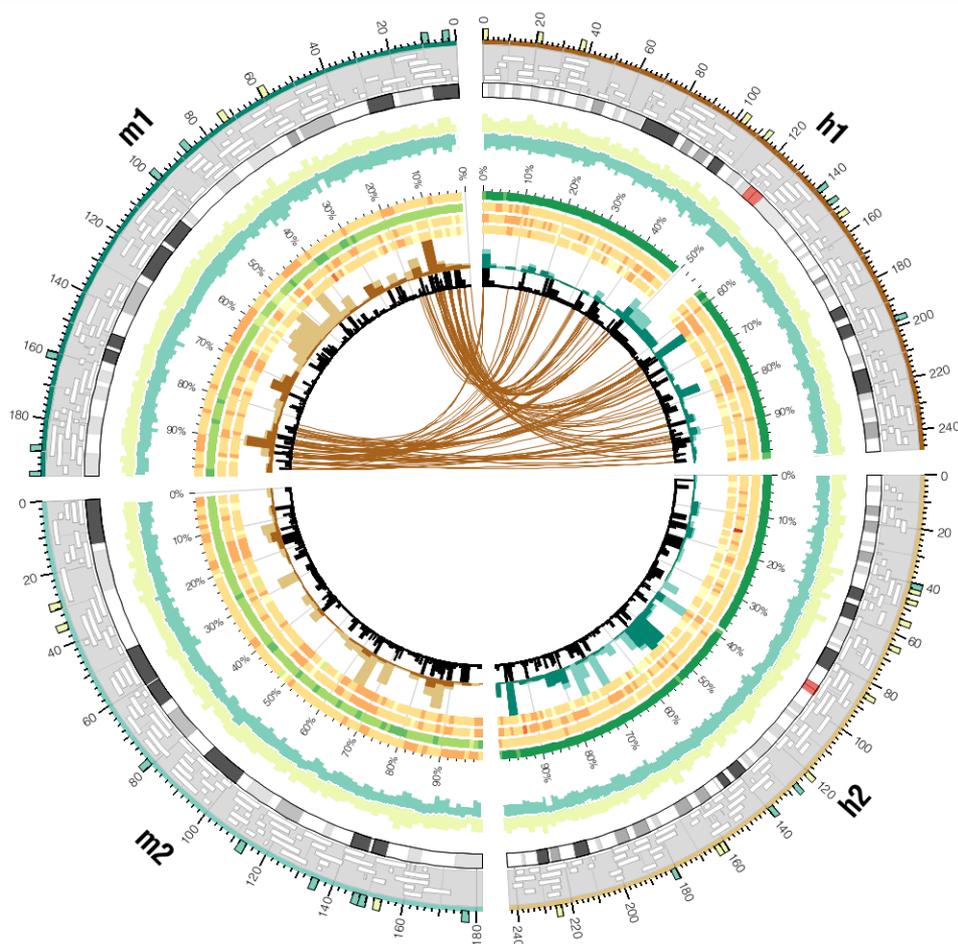
By adding a second rule that hides links, links that fail to match the first rule are not shown.

8 TESTING LINK CHROMOSOME AND POSITION

sessions/3/8/etc/circos.conf

```
condition = _CHR1_ eq "hs1"
            &&
            _CHR2_ eq "mm1"
            &&
            ( _START2_ < 40e6 ||
              _START2_ > 160e6    )
```


Rules can be used to select links based on position. Here, rules are used to show only links between hs1 and mm1, which start before 40Mb or after 160Mb on mm1.

8 ADJUSTING THICKNESS, COLOR AND LINK DEPTH

sessions/3/8/etc/circos.conf

```
<rule>
importance = 100
condition  = _CHR1_ eq "hs1"
             &&
             _CHR2_ eq "mm1"
             &&
             ( _START2_ < 40e6 ||
               _START2_ > 160e6   )

thickness  =
  eval(
        min(5,max(1,int(_SIZE1_/5e3)))
        . "p"
        )

color      =
  eval( "bd3-"
        .
        max(min(11,int(_SIZE1_/5e3)),1)
      )

z          = eval( int(_SIZE1_/5e3) )

</rule>
```
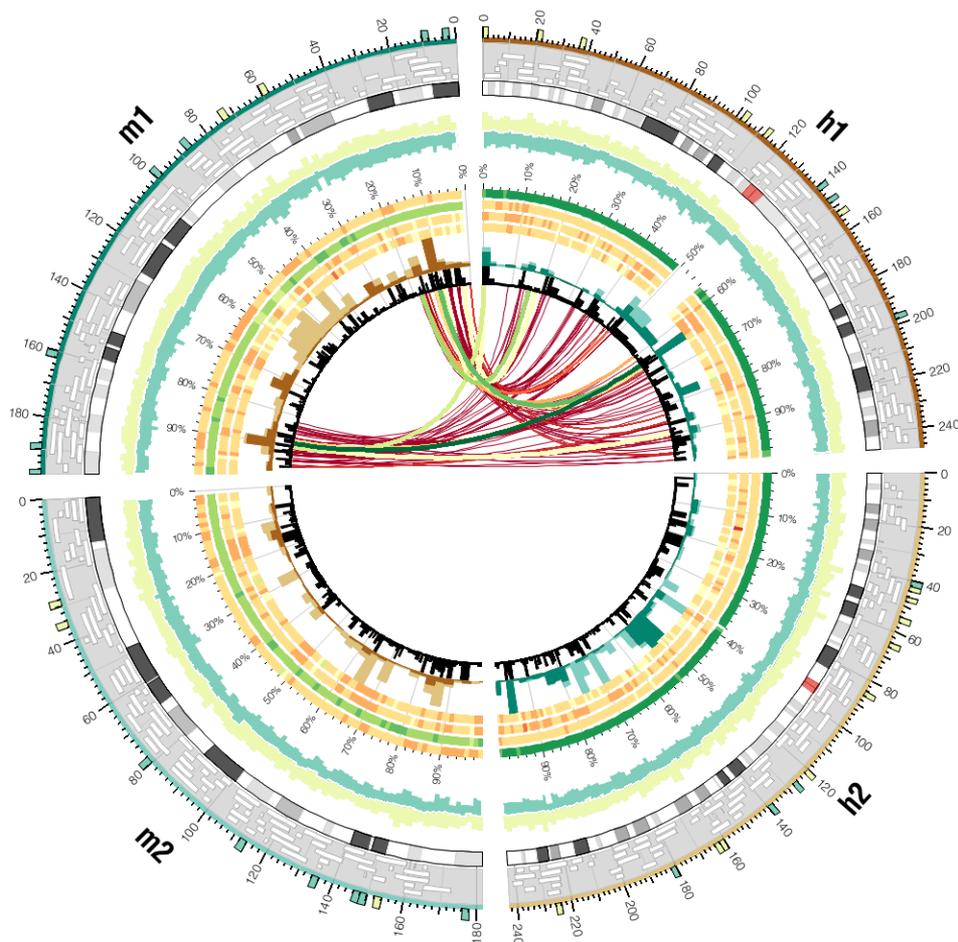

Using Perl code in the rule parameters, we can change the format of a link based on its properties.
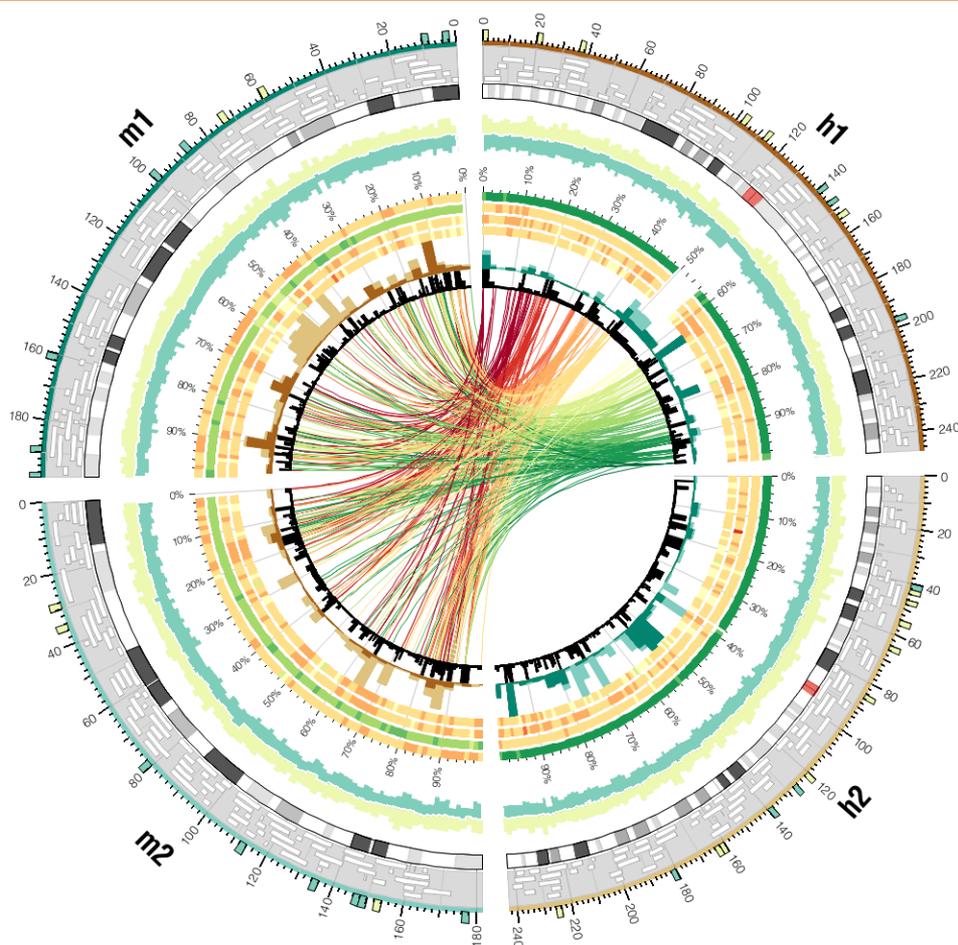
Here, the color, thickness and z parameters are changed based on the size of the link.

8

```
                                  sessions/3/8/etc/circos.conf
condition = _CHR1_ eq "hs1"

z           = eval( int(_SIZE1_/5e3) )

color       = eval(
                   "bd3-"
                   .
                   max( 1,int(_START1_/220e5))
                 )
```
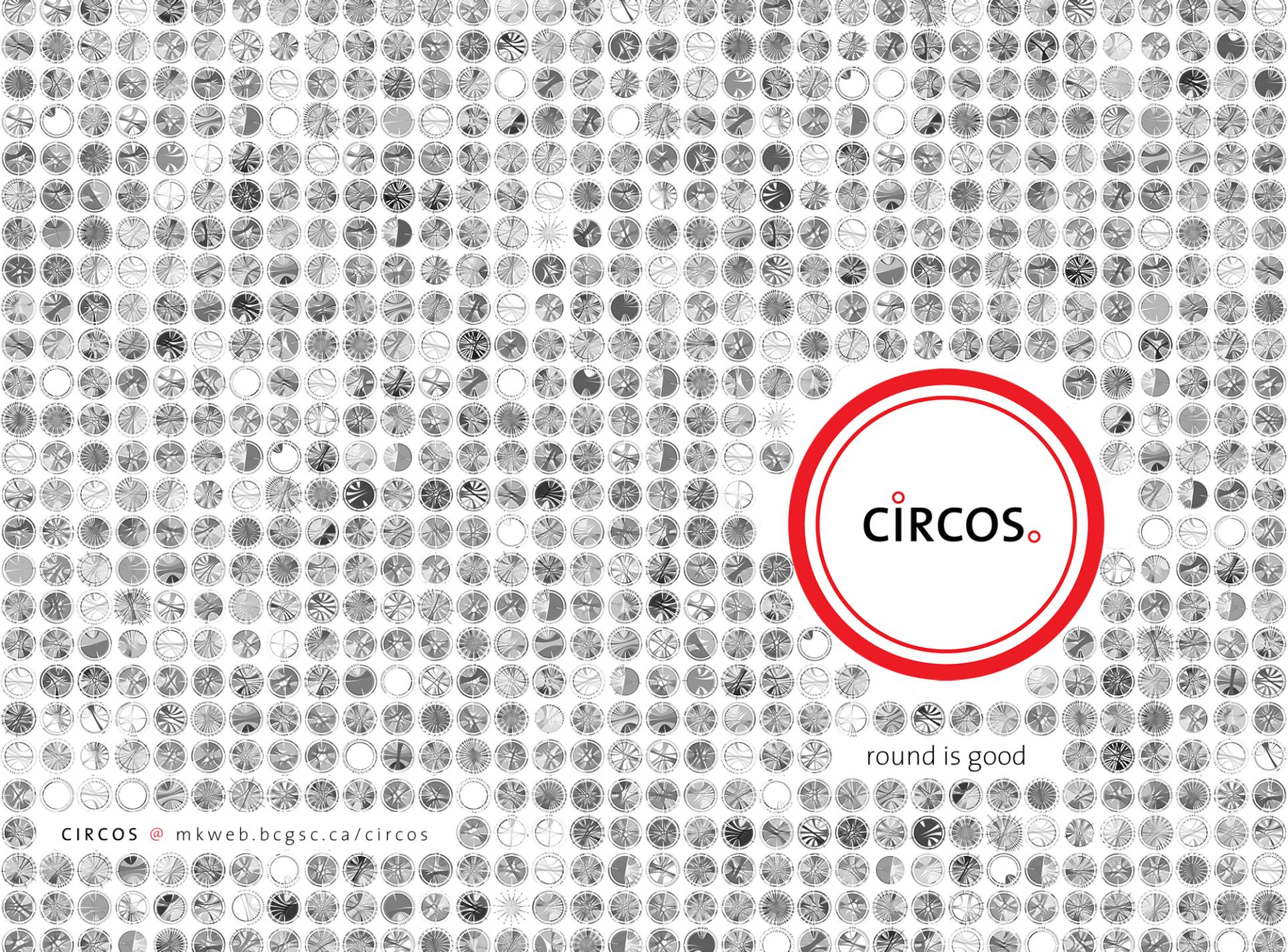

Using Perl code in the rule parameters, we can change the format of a link based on its properties.

Here, the color of the link is changed based on the position of its start.

CIRCOS.

round is good

CIRCOS @ mkweb.bcgsc.ca/circos