

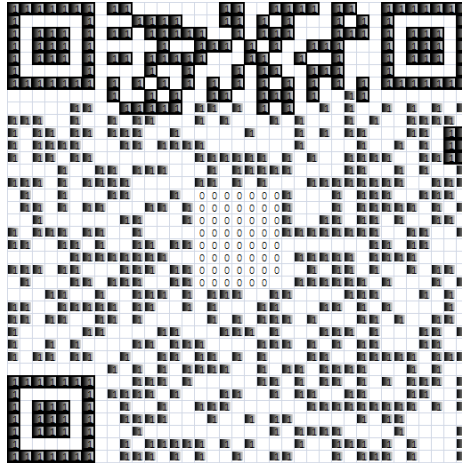
RectChr

使用手册

Visualizing genomic Statistical variables along Rectangular
chromosome

一种可视全基因组上各种统计量的工具

类 *Circos* 的功能, 把圆改为长方型 (Rect)



Version1.12 beta tmp 使用说明文档

2021-01-20

hewm2008@gmail.com / hewm2008@qq.com

目录

RectChr 使用手册	1
1.简介	1
2.应用场景示例	1
2.1 Scaf2chr 画图.....	1
2.2 遗传图谱画图	2
2.3 SNP/indel/gene/missing/GC 等密度图.....	3
2.4 BinMap+maker 画图	4
2.5 遗传统计量如 $P_i/F_{st}/t_{ajima}'D$ /杂合度等等群体相关变量图	5
2.6 多个群体同一统计量比.....	6
2.7 受选择 region 展示	7
2.8 两/三个基因组共线性	8
2.9 一个基因组 pairwise 关系	9
2.10 GWAS 的 Manhattan 图	10
3. 下载与安装	10
3.1 下载网址	10
3.2 预先安装	10
3.3 安装	11
4. 用法和参数说明	11
4.1 RectChr 参数.....	11
4.1.1 主要配置参数	12
4.1.2 其它配置参数	12
4.2 输入文件	15
4.2.1 数据文件(必须).....	15
4.2.2 配置颜色(可选).....	15
4.3 输出文件	15
5.实例	15
6.优势	16
7.常见问题	16
7.1 RectChr 和 Cmplot.....	16
7.2 line 线高于背景线如何办? Or 其它 lines 太低	16
7.3 text 文本颜色 or hightline 颜色自不想渐变? 如何设	17
7.4 对 line 和 hist 而言, 可以画坐标轴吗?	17
7.5 chr 背影条两端 有弧度 切换为 无弧度的方条? 如下图.....	17
7.6 联系与打赏.....	18

1.简介

RectChr 主要用于基于 Chr 染色体水平上的可视工具，对一些统计变量用 点，线，柱状和 heatmap、高亮，文本文字 以及结合颜色【线，散点，直方图，热图,文本,line,scatter/point, histogram ,PairWiseLink, link, heatmap(highlights)和 text 】等 可视化各 chr 上各区域这个统计量，达到快速一眼看出规律，识别结果。 并且各种可以自己组合 自由修改相关参数，使用方法极像 circos 的一样。

简单点说 circos 可以画的，这儿均可以画，只是把圈圈图改为长方型的。其中自己搭配层颜色等，同时也比 circos 多了一些默认配置，用起来十分简单，如 SNP GC 密度 直接输入文件即可。

程序是给一些有基础的生信朋友用的，若是小白看不懂就算了。

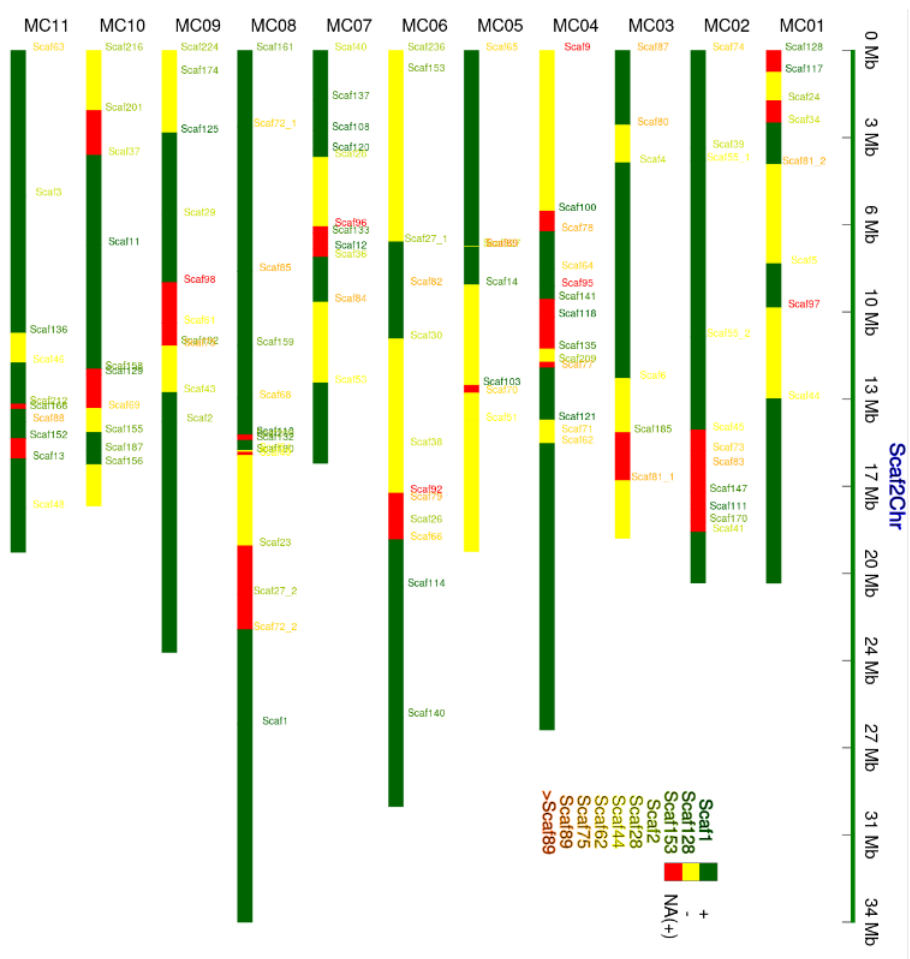
2.应用场景示例

程序可以由用户根据自己的想法应用在多个地方，应用多种不同场景。 如下是我近期应用到的场景，希望不是以为只可在这些场景应用，可以自己设计。

2.1 Scaf2chr 画图

输入格式：#CHR Start End ScaName 正/负

如下主要画两层 text 和 highlight（高亮）两种配合 。 然后旋转一下。

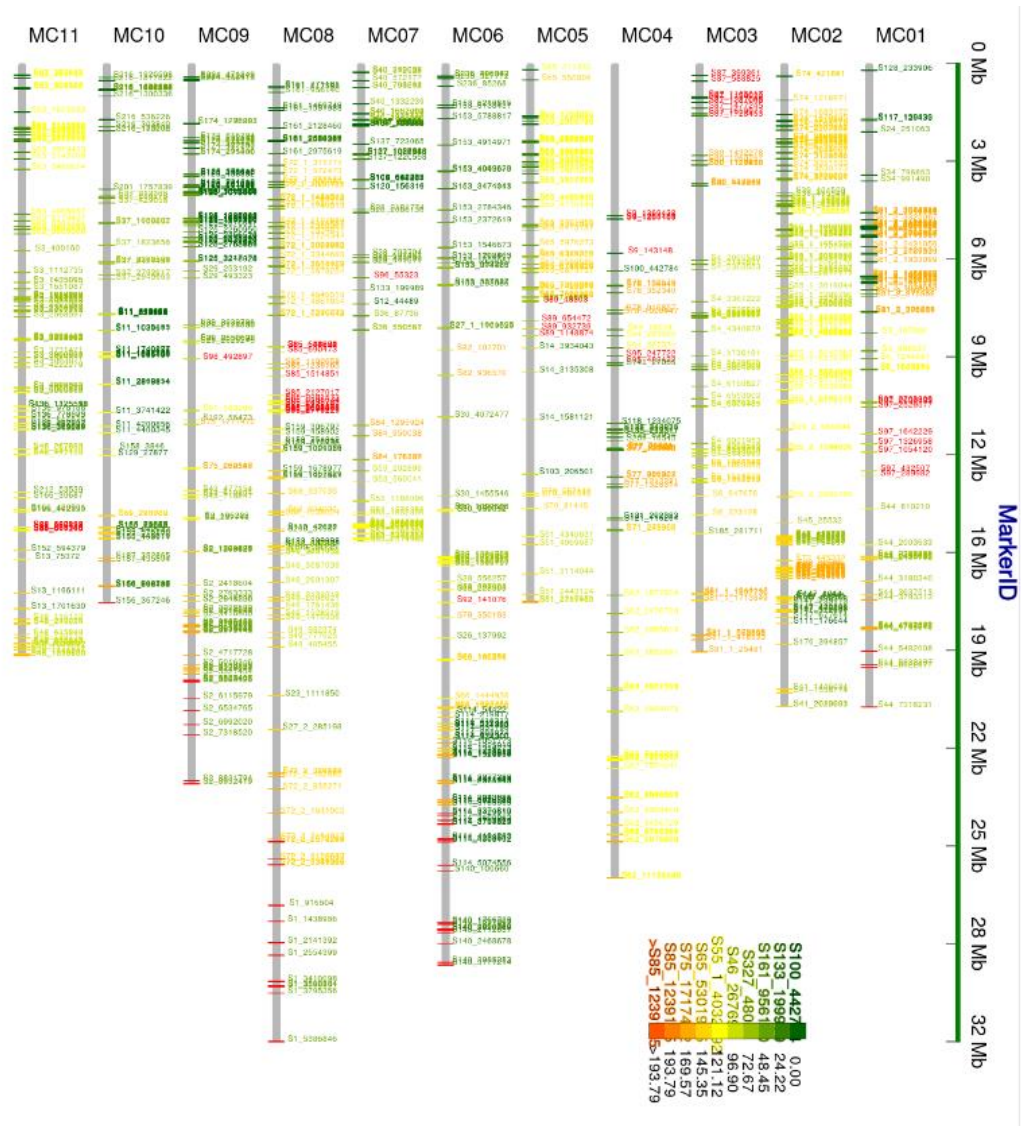


见 example1 里面的配置和用法 极输入文件格式。Text 的颜色可以通过配置设为统一颜色

2.2 遗传图谱画图

输入格式: #CHR Start End ScaName 正/负

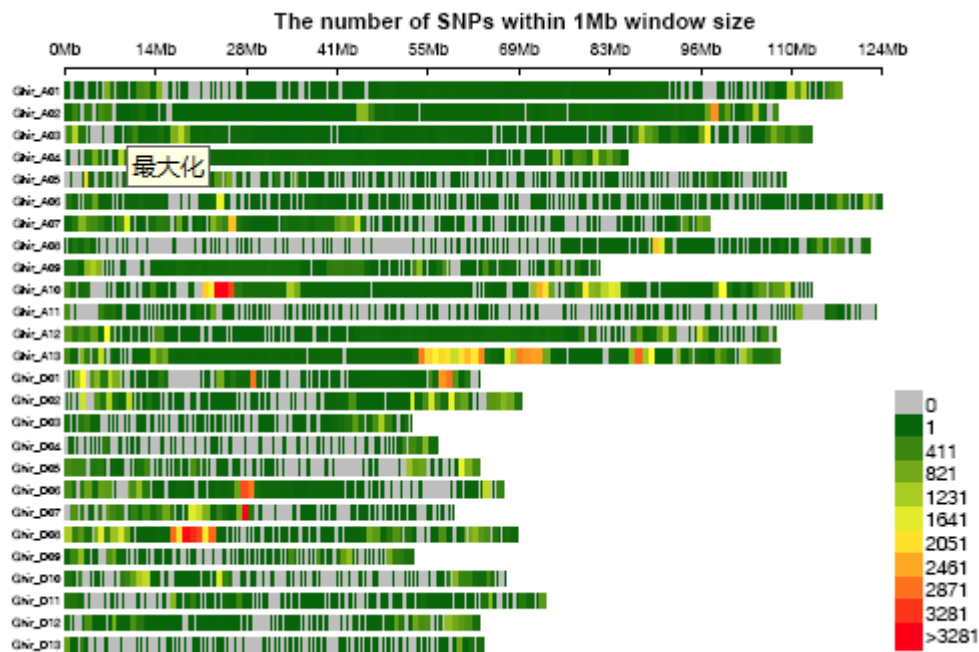
两层 即 画 text 层和 高亮层两种配合。 其中高亮层的背色条宽度缩小了点.



背影宽度 层次的配置 颜色若不要渐变 均可以配置 自己组合

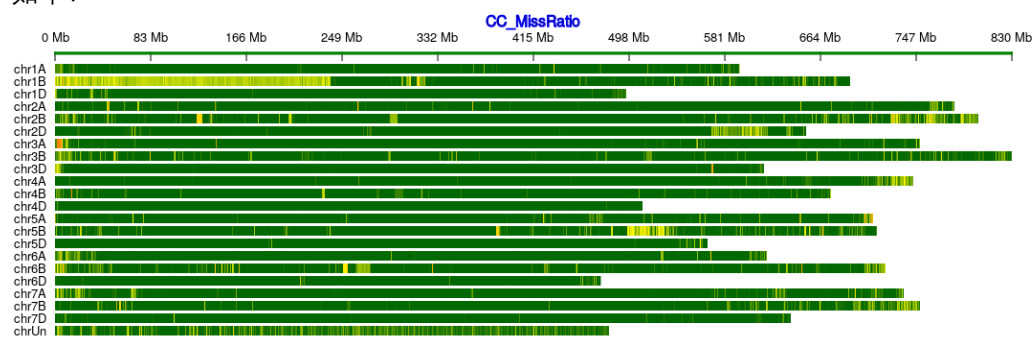
2.3 SNP/indel/gene/missing/GC 等密度图

输入格式: [#ChrStat End Flag] : Flag 为这个区域的统计量的值, 可以为整数可以为小数可以为字符串. 其中下面图的都是用 binbox 可以用颜色表示大小数. 用户可以结果数据用 point 高低 柱状图 和 lines 及结合颜色来画. **-Ptype** <s> : plot type (line/point/hist/heatmap/text)[heatmap]



见 example 3

如下：



2.4 BinMap+maker 画图

输入文件 A 格式： [#Chr Stat End Flag] 其中 Flag 为： P1/P2/H

输入文件 A 格式： [#chr site Flag] 其中 Flag 为： P1/P2

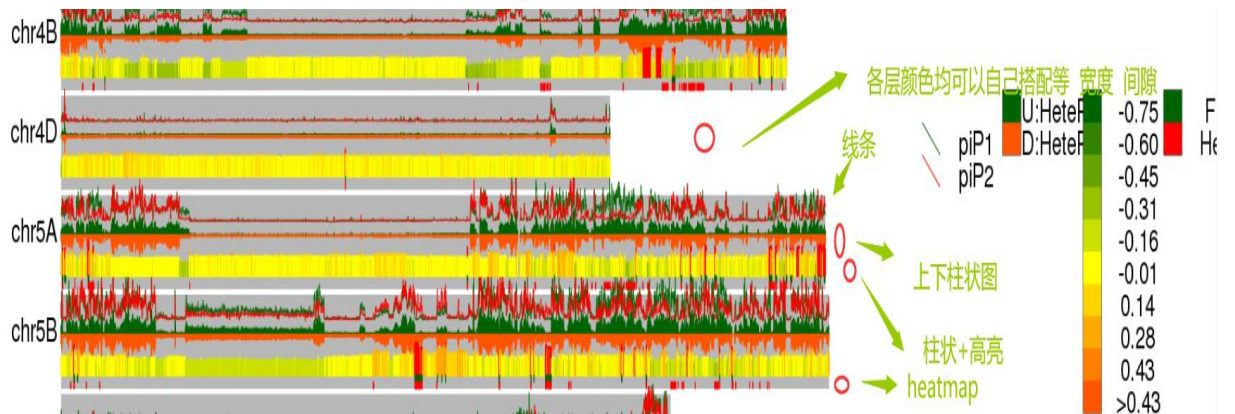
如下图

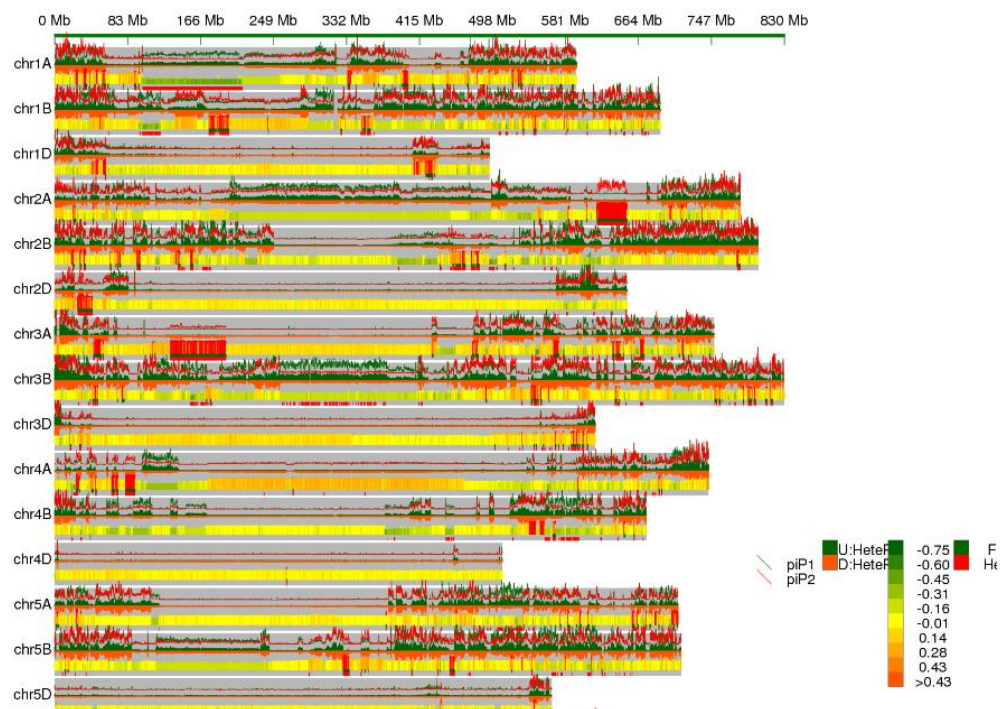


见 example4 里面的配置
主要是两层 heatmap

2.5 遗传统计量如 P_i /Fst/tajima'D/杂合度等等群体相关变量图

自由组合，如下 具体见 example5 用 line 表示线 柱+高亮表示 杂合
Fst 用 hist sele region 用热度图



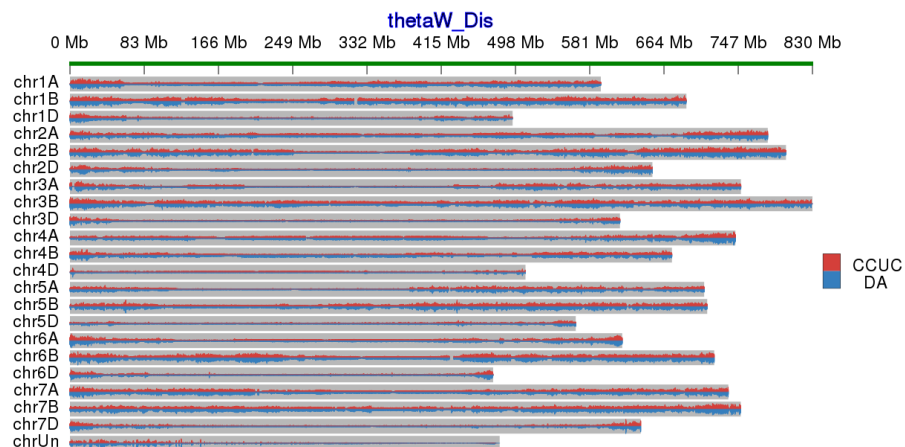


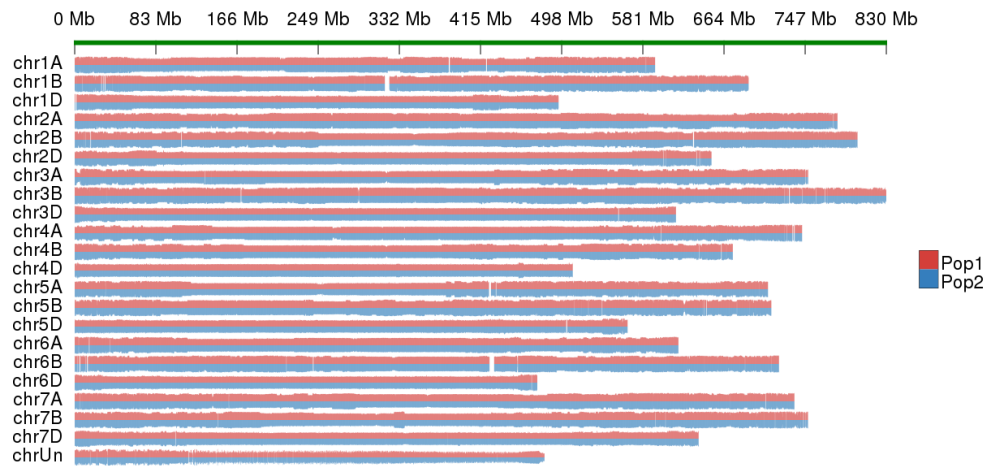
2.6 多个群体同一统计量比

上面都是同一个群体的，若多个群体也是可以的

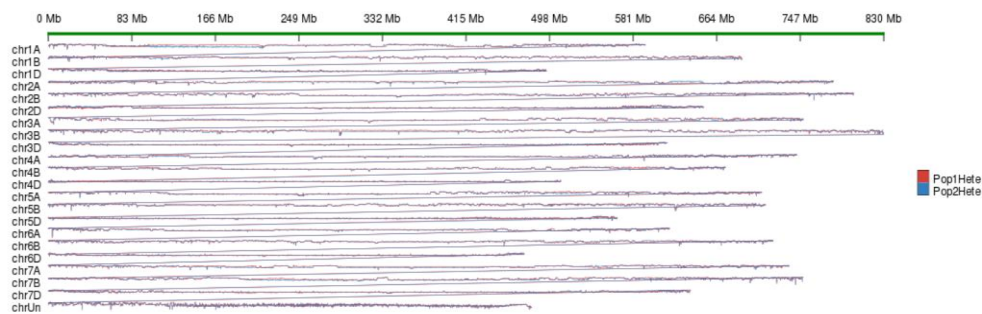
格式为：【chr Start End Flag1 Flag2 Flag3 ...】 FlagN 表示第 N 个群体这个统计量的值 没有的话用 NA 。 参数记得用 **-Row5Type** **MutiValue**

如下 是仅用两个群体的用 hist 柱状图表示画。：



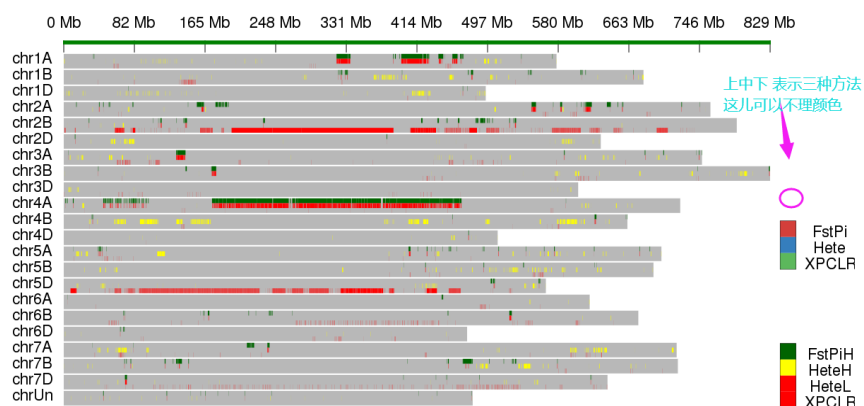


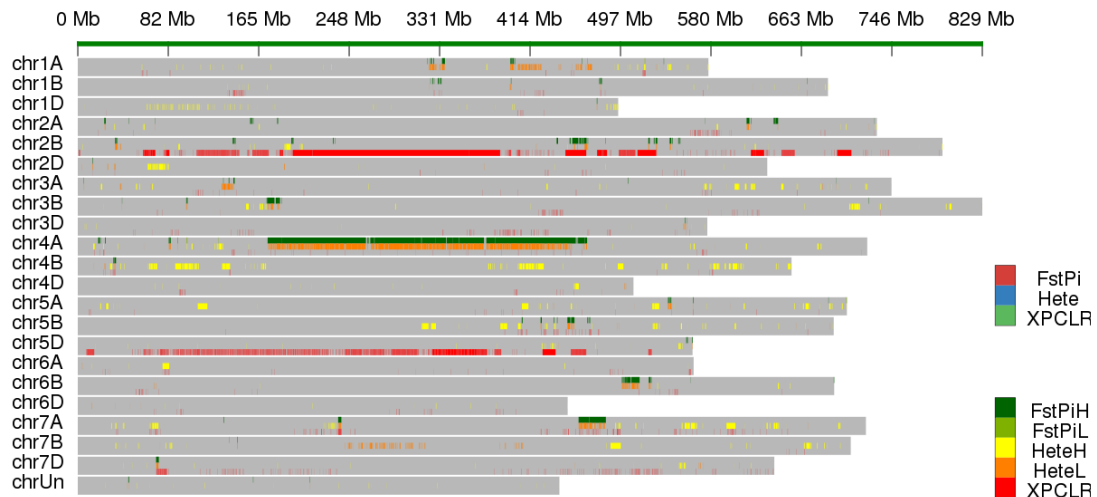
Lines 画



2.7 受选择 region 展示

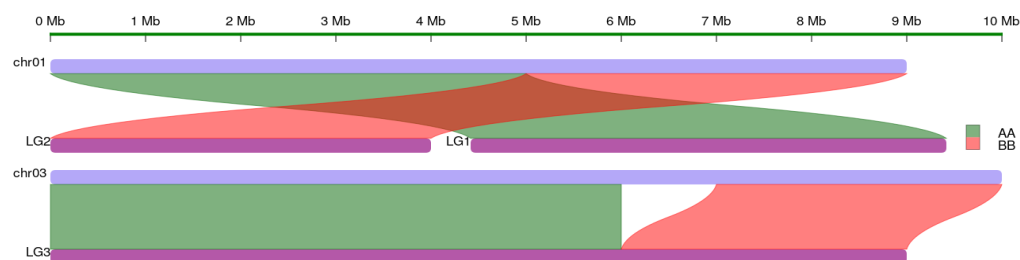
几种受选择分别找到受选择的方法，查看这些 region 的 overlap 情况。如下
 格式为：【chr Start End Flag1 Flag2 Flag3 ...】 FlagN 为方法 N 找到的区域 若不是
 用 NA 表示，其中 Flag 可以是字体字符串 or 数值。如下：
 如下用的是 boxbin，chr 分成了三份表示三种方法。用户可以用 lines point 画看等



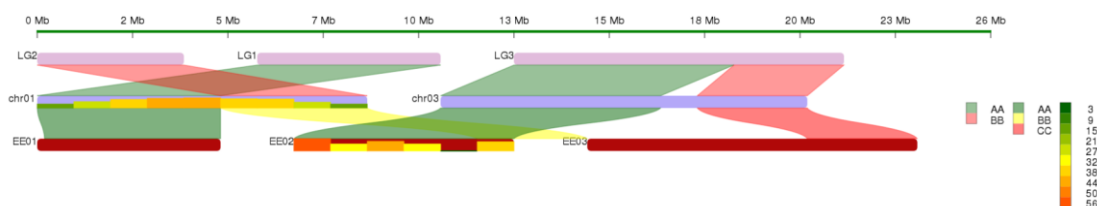


2.8 两/三个基因组共线性

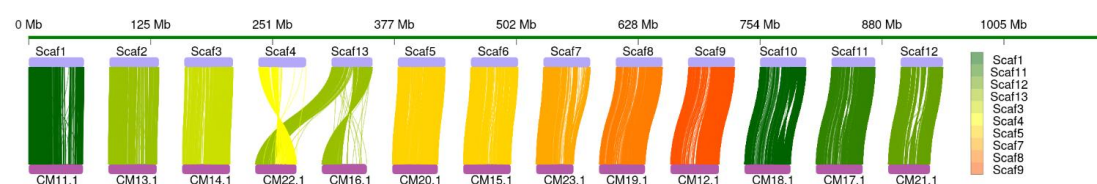
格式为：【chr Start End Flag1 chrB StartB EndB】 ,设置 link 时, 不能和其它结合。可以三个基因组一起画, 自动判断画在一起还是各 chr 分开 也可以选择是 chr 横向并行排列还是竖直排列, 同时用了直线和均匀曲线(贝塞尔) (可以选和搭配)。 现效果如下:



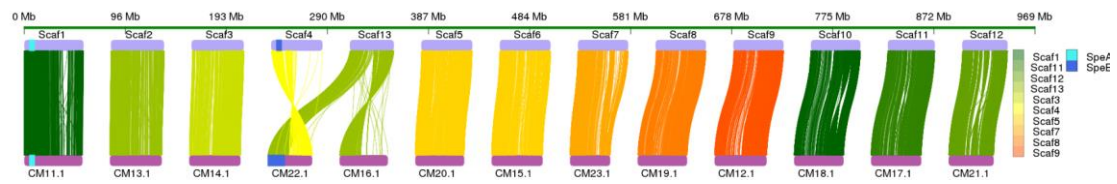
#lineType = curve ## line/curve 直线还是曲线,默认为 curve (见 example 6 :out2 和 out5)



真实数据画如下图 (配置见 example 6 :out3):



其中共线性可以对各 chr 添加一层 其它信息的如 gene 密度 GC 等（用 point hist line 等）
如下用了 heatmap 表示某几个 SpeA/B 的区域。见（配置见 example 6 :out3）：

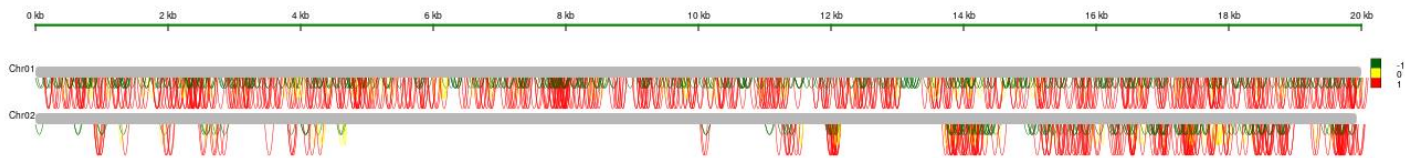


2.9 一个基因组 pairwise 关系

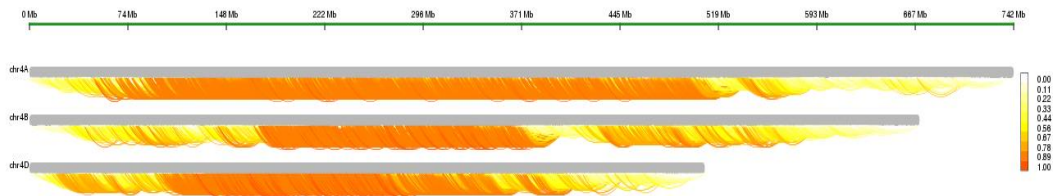
格式为：格式为：【chr Start End Flag】其中 当 PType = PairWiseLink
如下图见 example7 OUT1 ,可以看 sca 上比对的 paired 的距离。

其中

##SameHightForPairWiseLink=1 # 在 PairWiseLink 时, 这一参数设为 1 时 可以一样高
##radianUpDown=0 ## 在 PairWiseLink 时, 这一参数设为 1 时 弧度向下 默认为向上



再如下图 R^2 图的每取 1M 取一个 snp 看附近 30M-40M 和他的关系 R^2。如下图所示 见 example7 OUT2



如上主要是两层，第一层只画背影条，第二层则背影条为白色（即不画）用 PairWiseLink。 第二层高低可以通过配置自定义。

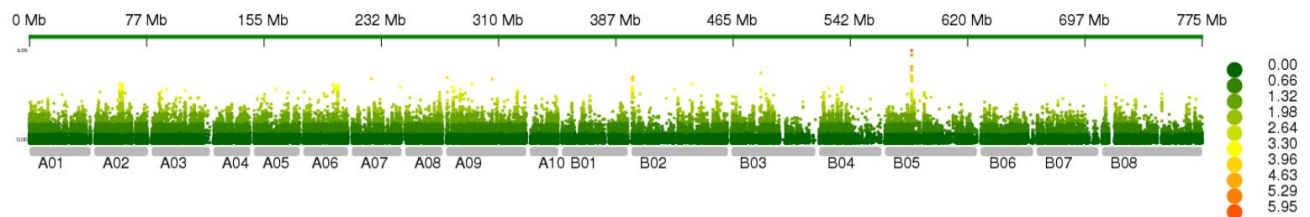
2.10 GWAS 的 Manhattan 图

如全局参数 **ChrArrayDirection = horizontal** 即把 chr 按横排列
(默认为 horizontal 竖排列)

两层上层为点图，层高度调高点；下层为 chr 啥都不画仅背景条。具体配置文件和例子见 example8

```
SetParaFor=Level1  ## 下面开始处理第 1 层 参数处理
PType = point      ## 散点
ShowColumn = File0:4  ## 把 file0 的第四列用散点图形式画出来)
ChrWidth=100        ## 这一层宽度调高点
LogP=1 # 0-log() 对数值进行取 log
crGB="#FFFFFF" # 背景为白色，实没有画
TopVHigh=1.1 # 只要大于 1 时就会取最高点
ShowYaxis=1 # 显示左边数值
ValueSpacingRatio=0.04 # 和一层空一点

SetParaFor=Level2 # 这一层啥都不画 仅用背景条
ShowColumn = File0:4
ChrWidth=10
PType = EEE #啥都不画 空
```



3. 下载与安装

3.1 下载网址

暂时无对外公开，须要的话再联系 NA 。主要还是我自己在用

对外 基因库 试用 **1 个月**的版本，(20210120 后不能用)

若有需求再更新后再正式公布。

这主要是防止大家还在用 beta 版本，还在用可能存在 bug 的程序，强迫大家定期更新。

/hwfssz4/BC_PUB/Software/07.User-defined/03.Animal_Plant/RectChr-*

里面有 example1-7

3.2 预先安装

RectChr 适用于 Linux/Unix/macOS 系统。在安装之前，请先安装以下使用条件：

- 1) Perl: Perl 内的[模块 SVG.pm](https://metacpan.org/release/SVG) (<https://metacpan.org/release/SVG>)。可以预先安装。尽管我们安装包已经内置了 SVG.pm 包了
- 2) [convert](#): 系统命令，可以将 svg 转 png。莫有也无关系，有则更佳

3.3 安装

使用者可采用以下直接 chmod 755 运行：

1)

```
chmod 755 bin/*  
./bin/RectChr -h # 直接运行
```

4. 用法和参数说明

4.1 RectChr 参数

程序 RectChr 很简单，一个输入和一个输出。具体如下。

```
[heweiming@cngb-ologin-25 bin]$ ./bin/RectChr  
Version:1.4          hewm@genomics.cn  
  
Options  
  
-InPut      <s> : InPut Configuration File  
-OutPut     <s> : OutPut svg file result  
  
-help      : Show more help with more parameter
```

- InPut 输入文件配置好的文件
-OutPut 输出文件， out.svg

-help 具体见配置参数

4.1.1 主要配置参数

SetParaFor 是设置参数 可以是 global 全局变量 ALL 是所有层都起作用 LevelX 是对 X 层的参数
配置文件主要参数解释有如下几个

序号	示例	解析说明
参数 1	FileN = /path/in.data	(N=0,1,2...) 输入的是第 N 个文件的路径
参数 2	PType = heatmap	[关键词中的一种: heatmap/ highlights/link/ pairwiselink/ histogram/hist/ scatter/ point/ text/line/lines, 即共有 7 种画法] 若不是关键, 则不画, 仅背景条 heatmap/highlights 对区域用 binbox 型式结合颜色的不同显示统计量的差异 link : 对关联在一起的区域用颜色(贝塞尔曲线/直线)块结合颜色连接在一起, pairwiselink: 起终点 用 贝塞尔曲线 结合高低和颜色 展示统计量(又叫彩虹线) histogram/hist : 用柱状图高低结合颜色 展示 统计量 scatter/ point: 用散点的高低结合颜色 展示 统计量 text : 用文本形式 结合颜色 展示 统计量 lines:用线条及结合颜色 展示 统计量
参数 3	ShowColumn = File0:8 File1:5,6	某一文件的第 X 列作为统计量

4.1.2 其它配置参数

SetParaFor = LevelX 为设置各层(X)的一个 Flag 参数 .global 是全局变量
ALL 是所有层都能管用, 各层没有配置的话, 则默认为全局的参数。第 X 层, 相当 circos 的第 X 个圈
下面列出一个示例 大家可以看后面##的解析

```

1. SetParaFor = global
2.
3. File0 = ./in.data          ## 这个是必须输入参数, 并且尽量放在最前, 格式为
   [Chr Start End Value1 Value2 ...ValueN]
4.          ## 其中用 NA 表示不画, chr End End NA 不画但 End 可以用来记
   记为 chr 的长度
5. #ValueX =          ## 多少层, 类同 circos 多少个圈, 这不设默认是 N, 即根据
   File0 的格式来的, 可以自己设
6. #ChrSpacingRatio =0.2 ## 不同染色体 chr 之间的间隔比例
   (ChrWidth*ChrSpacingRatio)
7. #Main = "main_Figure" ## the Figtur Name #font-
   size strokewidth=1; fill="green"
8. ColorsConf = col.file ## 通过在主配置文件 input 自定义颜色和 Value 的对应关
   系;( P1 = "#FE0808" )
9. ##### Figure #####
   #####

```

```

10.
11.
12.
13.
14.
15. #####          画布 和 图片 参数配
    置 #####
16. #Chromosomes_order =    ## chr 的顺序和只列某些 chr 出来画，若没有配置，程序会按
    chr 名自动排序 chr1,chr2,chr3
17. #body=1200    ##    默认是 1200，主画布大小设置    另外：
    up/down/left/right) = (55,25,100,120);
18. #RotatePng    = 0    ## 对 Figure 进行旋转的角度
19. #RotateChrName = 0    ## 旋转 chr 名字 text
20. #ChrSpacingRatio=0.2    ## 不同染色体 chr 之间的间隔比例
    (Sum(ChrWidthX*X)*ChrSpacingRatio)
21.
22.
23.
24.
25.
26. #####    默认各层的配置参数 若各层没有配置的会，则会用这儿的参数 #####
27.
28. SetParaFor = LevelALL    ## 下面是处理初始化参数 SetParaFor 参数处理,若
    为 LevelALL，即先为所有层设置的默认值
29. #File1    =    ## 可以输入别的文件
30. PType = heatmap    ## 线，散点，直方图，热图,文
    本, line, scatter, histogram , heatmap(highlights)和 text 新的功能 links
31. #ShowColumn =    ## 若 SetParaFor 为 LevelALL 时，N 层的 ShowColumn 默认为
    File0 的第 ValueN 所的 Column(N+3)
32.    ## 参数格式可以设为 ShowColumn=File0:4 File1:4,5
33.    ## File1:4,5 表示 file1 的第四和第五列用 heatmap 表示
34. #crBegin="#006400"    ## 此层(ValueX)最低值 Value 的配色
35. #crMid="#FFFF00"    ## 此层(ValueX)中间值 Value 的配色
36. #crEnd="#FF0000"    ## 此层(ValueX)最大值 Value 的配色
37. #crGB="#B8B8B8"    ## 此层(ValueX)背景色 的配色
38. #TopVHigh=0.95    ## 此层 Top of ValueX 用最高点颜色[0.95],其它再等分
39. #TopVLow=0    ## 此层 Top of ValueX 用最低点颜色[0],其它再等分
40. ##YMax=    ## 设置此层(ValueX)的最大值,默认自动
41. ##YMin=    ## 设置此层(ValueX)的最小值,默认自动
42. #Gradien=10    ## 此层(ValueX)多少等分颜色
43. #ChrWidth=20    ## 此层(ValueX)在画布的宽度
44. #BGWidthRatio =1    ## 此层(ValueX)的背景(backgroup)的宽度默认和 ChrWidth
    一样(0-1])
45. #LogP=0    ## 此层(ValueX)不作 0-log10(Value) 处理

```



```

46. #ValueSpacingRatio=0    ## 同一染色体中此层(ValueX)之间的间隔比例
    (ChrWidth*ValueSpacingRatio)
47. #SizeGradienRatio=    ##设置渐变条的大小
48. #ShowYaxis=0          ## 是否显示所有层的 Y axis 的起终点值,默认值此:0 不显示
49. ##### 更多配置的参数 可以自己设,没有的话会自动设置 #####
50. ##Rotate/fill/Cutline/strokewidth/stroke/stroke-width/font-
    size/fontfamily/ =    ### 等等
51. ##ShiftGradienX=0 ## 渐变条左右移动    ##ShiftGradienY=0 ## 渐变条上下移动
52. ##ShiftChrNameX/ShiftChrNameY 移动 chr 名
53.
54.
55. ##### 各层的参
    数 #####
56. ### 具体某层的具体配置 把 DealLevePara 设为具体正数(<=ValueX),然后可以具体
    修改此层要改变的部分
57.
58.
59. SetParaFor=Level2    ## 下面开始处理第 2 层 参数处理
60. #File1 =            ## 可以输入别的文件 file1
61. PType = hist    ## 散点
62. #ShowColumn = File1:5 ## 把 file1 的第五列用散点图形式画出来)
63.
64.
65. SetParaFor=Level3    ## 下面开始处理第 3 层 参数处理
66. PType = lines    ## 线条
67. ShowColumn = File0:5,6 ## 把 file1 的第五,第六列用线条图形式画出来)

```

其中也提供有了更多参数,只是大家不常用到,就隐藏起来,如 X 轴坐标系的上移下移
global : ShiftAxisY=0 等。若须要更多参数,可以反馈~ 程序设计时就考虑到随机加
减参数。现列出几个重要的其它参数

序号	示例	解析说明
global	ChrArrayDirection= vertical	horizontal/vertical chr 是按纵排列还是横排列
	TopVHigh=0.95	此层 Top of ValueX 用最高点颜色[0.95],其它再等分
	TopVLow=0	设置此层(ValueX)的最小值,默认自动

4.2 输入文件

至少要输入一个数据文件。

4.2.1 数据文件(必须)

输入格式为文本格式, 同一个文件里 各行的列数须相当。确保相当确保前三列为固定格式。
在 PType 在非 link 的情况下, 文件格式如下:

【Chr 名 坐标 1 坐标 2 Flag1 Flag2 ... 】

为 PType=link 时格式为:

【chrA StartA EndA Flag chrB StartB EndB】

其中第一行若为 head 的话, 须以“#”开头

4.2.2 配置颜色(可选)

程序提供了可选的输入文件, **非必须的**, 即可以自己定义配置颜色的文件, 格式很简单, 颜色为 16 进制的 HTML 颜色

```
A="#F8F8F3"  
B="#A3FF88"  
C="#FFFFFF"  
...  
...
```

4.3 输出文件

输出文件	说明
out.svg	输出的 SVG 格式图
out.png	输出的 png 格式图

示例图见上面应用场景给的图

5.实例

上面示例图都有实例, 我这
具体数据格式和配置等见这:

```
./bin/RectChr -InPut in.cofi -OutPut out.svg
```

这儿共提示了 7 个实例，配置文件和输入文件格式等，运行就只须要几分钟。

具体见 软件里面的 **Example** 目录：

exempl1-7

6.优势

- 1 快速少内存 批量画图 瞬间出结果。
- 2 可以自定义各种参数，如 chr 层，可以自己颜色，定义高度，画图设计全开放与用户，可以应用在多种场景
- 3 免安装，使用方便

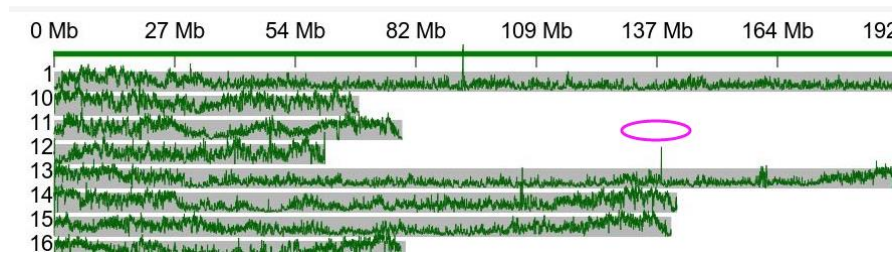
7.常见问题

7.1 RectChr 和 Cmplot

RectChr 和 circos 有点类似，各有优势。你们用之 corcos 是圈 我这是长方型
Cmplot 不能点和线等自由组合。

7.2 line 线高于背景线如何办？ Or 其它 lines 太低

那个线图，高的值超出染色体背景了，这个有参数直接调吗？



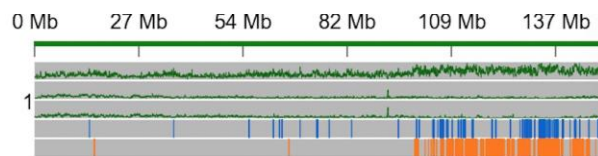
答： 有， 参数-TopVHigh=0.95 修改为 1 即可

此层 Top of ValueX 用最高点颜色[0.95],其它再等分 默认为 0.95 ,。

再问： 但改为 1 后其它地方点得太低了？ 如何办？

答： 主要是数据问题，如 95%的值都在 0-100 之间，另外 5%的最高值为 1000.
此时的建议是对数据作一定的转化，如高于 100 的强制为 100 or 取 log 之类的
再设 TopVHigh=1.0

7.3 text 文本颜色 or hightline 颜色自不想渐变？如何设



A. 方法

参数 crMid="#FFFF00" crEnd="#FF0000" crGB="#B8B8B8" 三者改为同一颜色就行

B. 方法

Gradien 设为 1 修改 crGB="#B8B8B8" 为你要的参数就行

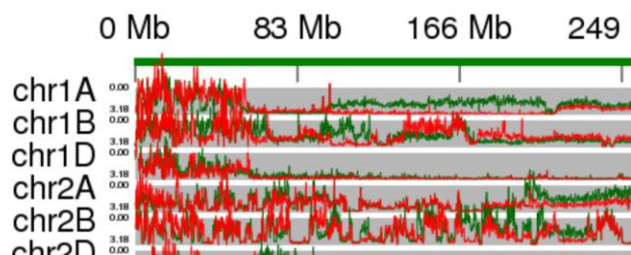
7.4 对 line 和 hist 而言，可以画坐标轴吗？

如上图 line 的左边能加上数值吗？

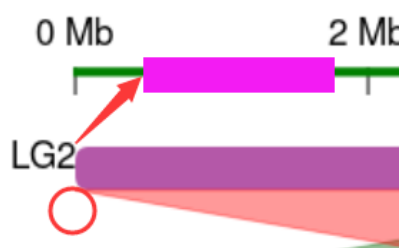
答：已经添加参数

ShowYaxis=0 ## 是否显示所有层的 Y axis 的起终点值,默认值此:0 不显示

把对应的层 ShowYaxis=1 即可



7.5 chr 背影条两端 有弧度 切换为 无弧度的方条？如下图



回可以： 全局变量

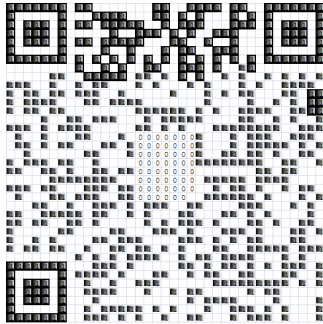
BGChrEndCurve = 0 #即可 默认为 1 即为 曲线

7.6 联系与打赏

随意 是缘是福，一切随风

- [✉ hewm2008@gmail.com](mailto:hewm2008@gmail.com) / hewm2008@qq.com
- join the **QQ Group : 125293663**

微信 打赏



QQ 入群: **125293663**



群名称: Reseqtools (itools)
群 号: 125293663