

使用手册

NGenomeSyn

Any Way to Show

Multi genomic Synteny

任意展示多基因组共线性的作图工具



Version 0.17 使用说明文档

2022-04-20

hewm2008@gmail.com / hewm2008@qq.com

微信 打赏



QQ 入群: **125293663**



群名称: Reseqtools (tools)
群 号: 125293663

微信公众号



目录

NGenomeSyn.....	0
1.简介	1
各版本新功能	1
2.应用场景示例	1
2.1 两个基因组.....	2
2.2 三个 (N) 基因组 竖排.....	2
2.3 三个任意分布	4
2.4 任意基因组任意分布	6
3. 下载与安装	8
3.1 下载网址	8
3.2 预先安装	8
3.3 安装	8
4. 用法和参数说明	9
4.1 NGenomeSyn 参数	9
4.1.1 主要配置参数	9
4.1.2 其它配置参数	10
4.2 输入文件	12
4.2.1 数据文件(必须).....	12
4.2.2 配置颜色(可选).....	13
4.3 输出文件	13
5.实例	13
6.优势	14
7.常见问题	14
7.1 NGenomeSyn 和 RectChr 的中的共线性差异.....	14
7.2 NGenomeSyn 的 chr 边框会被覆盖, 如何调作?	14
7.3 联系与打赏.....	15

1.简介

NGenomeSyn 是于基于多个基因组共线性关系的可视工具，对多个基因组自主排列位置，扩大收缩，旋转角度 等，结合颜色，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。

亮点：

- 1 任意多个基因组（目前我限 12 个，可以取消）
- 2 各个基因组可以自己调顺序 颜色和添加自己属性等等功能
- 3 各个基因组可以移动 旋转 扩大和收缩等
- 4 Link 信也可以自定义颜色和添加特定属性(如透明度)等等功能

其中 3 可以调好相关参数 可以出现 三角形 五角形 四边形 等等组合（后面将有空升级调定这些参数）

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。

[程序提供](#)了 3 不同的应用场景实例之后，[知乎专栏](#)也收集了一些用户的配置和说明。

各版本新功能

- A. Beta 版本，大家试用 有 bug 和更的需求有空再丰富
- B. 后面将加一个新的功能参数 可以直接出 三角形 等的排布

2.应用场景示例

两个 三个 等 N 个基因组均可以画。

各个基因组的关系颜色也可以配置，后面再更新

默认分布是 1 各基因组 从上到下 居中排列分布

-

NGenomeSyn 各位画时

- 1 遇到问题可以反馈与我
- 2 同时我希望有好图可以帮发个贴 教与别人
- 3 若图很好很美 可以把物种名 等改一下，[隐藏一些信息](#) 我把其写入 [exampleXX](#) 并 [收藏进使用说明里面](#)。
- 4 NGenomeSyn 若没有问题的话 稳定后再把 beta 版改为稳定版本



2.1 两个基因组

最简易用法。输入两个基因组的坐标 和 link 文件即可 (0.17 后的版本提供了一个小 perl 流程, 可以直接由两个 fa 文件生成这三个文件, [具体说明点击这儿查看](#), 方便更多小白操作)

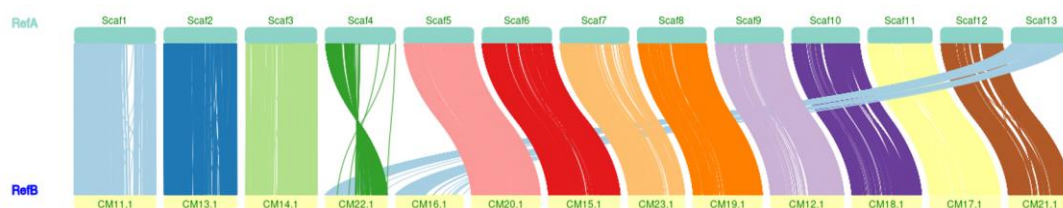
如 example1 的配置文件如下(数据来自 RectChr example6):

```
SetParaFor = global
GenomeInfoFile1=RefA.info
GenomeInfoFile2=RefB.info
LinkFileRef1VsRef2=syn1.data
```

后图如下:



用户可以 自主 加属性 变透明 调颜色 等,
如加上 ChrNameShow=1 显示 chr 名, 具体见 example1



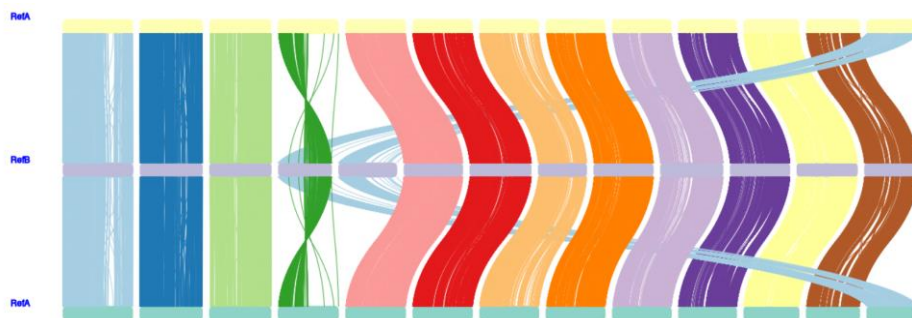
2.2 三个(N)基因组 竖排

我们用的还是上面的数据, 用三个基因示例如下:

```
SetParaFor = global
GenomeInfoFile1=RefA.info
GenomeInfoFile2=RefB.info ## GenomeInfoFile2 就表示有 2 个基因组
GenomeInfoFile3=RefA.info ## GenomeInfoFile3 就表示有 3 个基因组
LinkFileRef1VsRef2=syn1.data
LinkFileRef3VsRef2=syn1.data
```

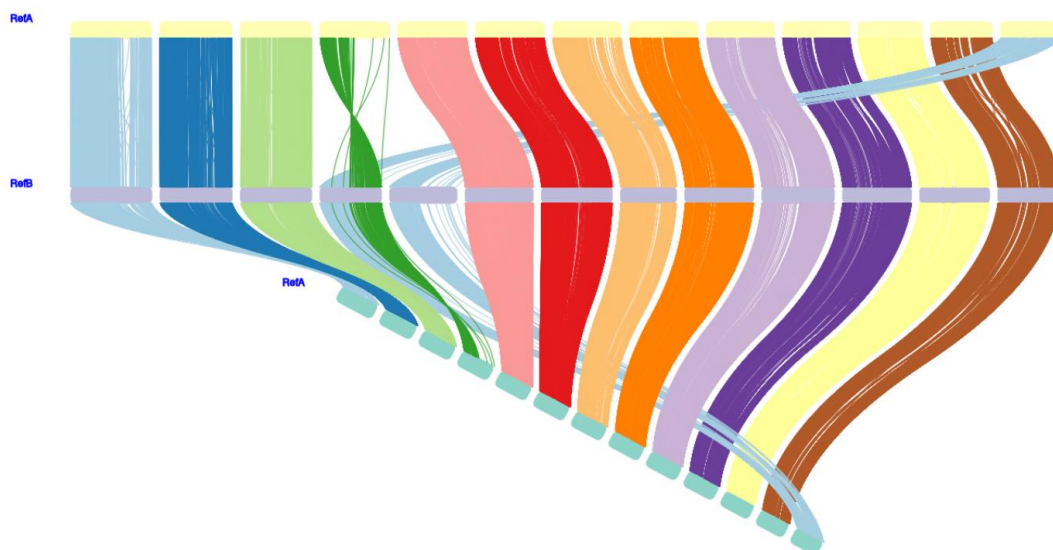
即得到如下图:





那么 我们对第三个基因组进行 缩小 旋转等，在上面的配置后面加上如下信息：

```
CanvasHeightRatio=1.6 # 画布高度多点
SetParaFor = Genome3 # ALL/GenomeX X 第X个基因组
ZoomChr=0.5          ## chr 长度 等大 缩小 or 扩大
RotateChr=28          ## chr 的起点 顺时针 旋转 28 度
ShiftX=60
ShiftY=-80            ##对这个基因组移动
```

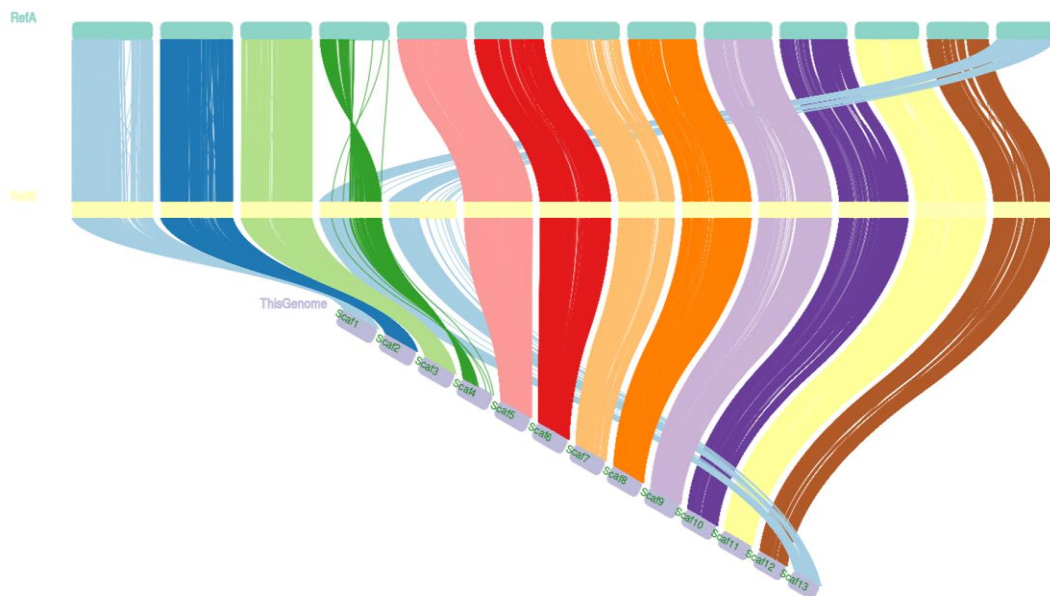


另外可以加上一些其它信息 显示 chr 名 颜色 和稍移动

```
ChrNameShow=1
ChrNameColor=green
ChrNameShiftX=10
ChrNameShiftY=30
GenomeName="ThisGenome" ##自定义基因组名字
```

后见下面：





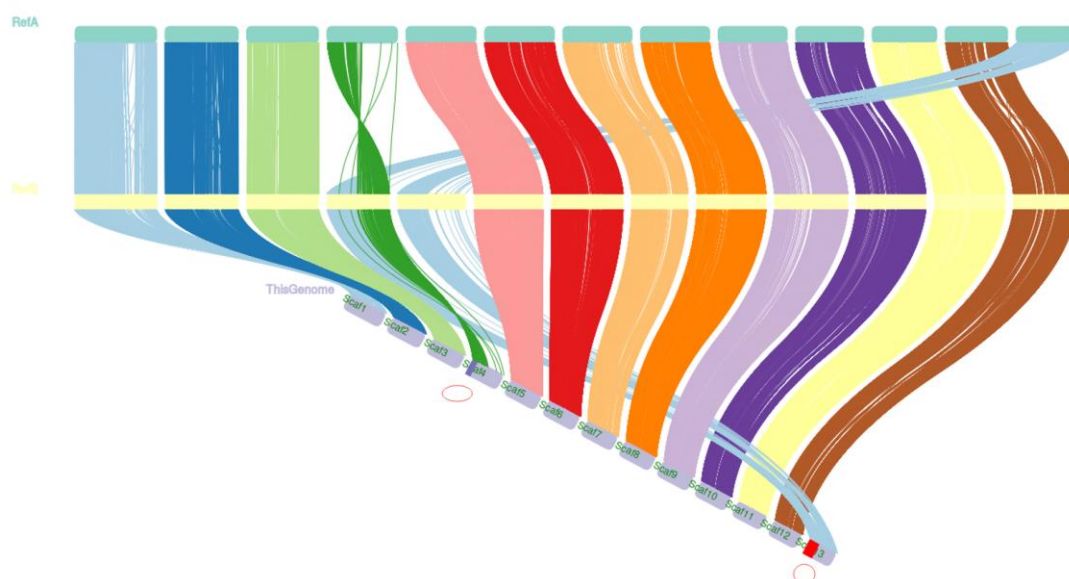
其中想把特别的区域表示出如如，文件 Spe.region 格式 如下

```
Scaf4    1000000 8000000
Scaf13   1000000 15000000 fill=red stroke=red
```

在对应的 Genome3 加上如下一行

SpeRegionFile=Spe.region ## 文件,把特别区域表示出来[格式 chr start End]

则如下:



2.3 三个任意分布

共四种作 link 方式

DownDown UpUp DownUp UpDown 及设置颜色和透明度等 示例:



具体见 example3 里面的

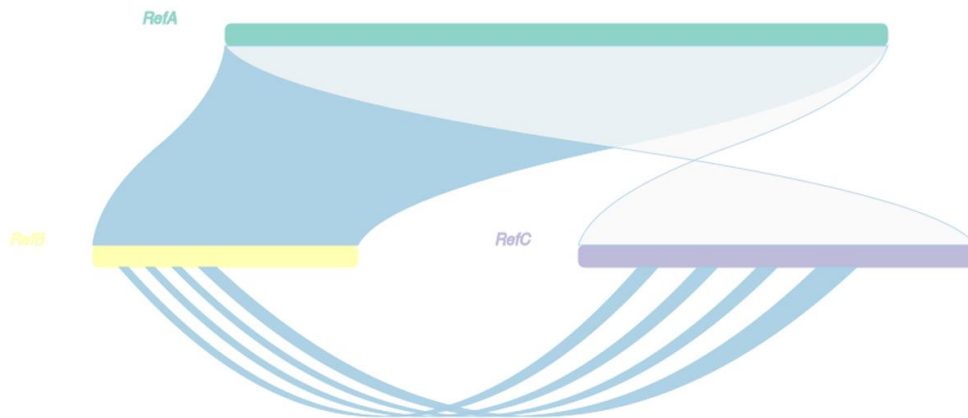
SetParaFor =Link3

HightRation=1.5

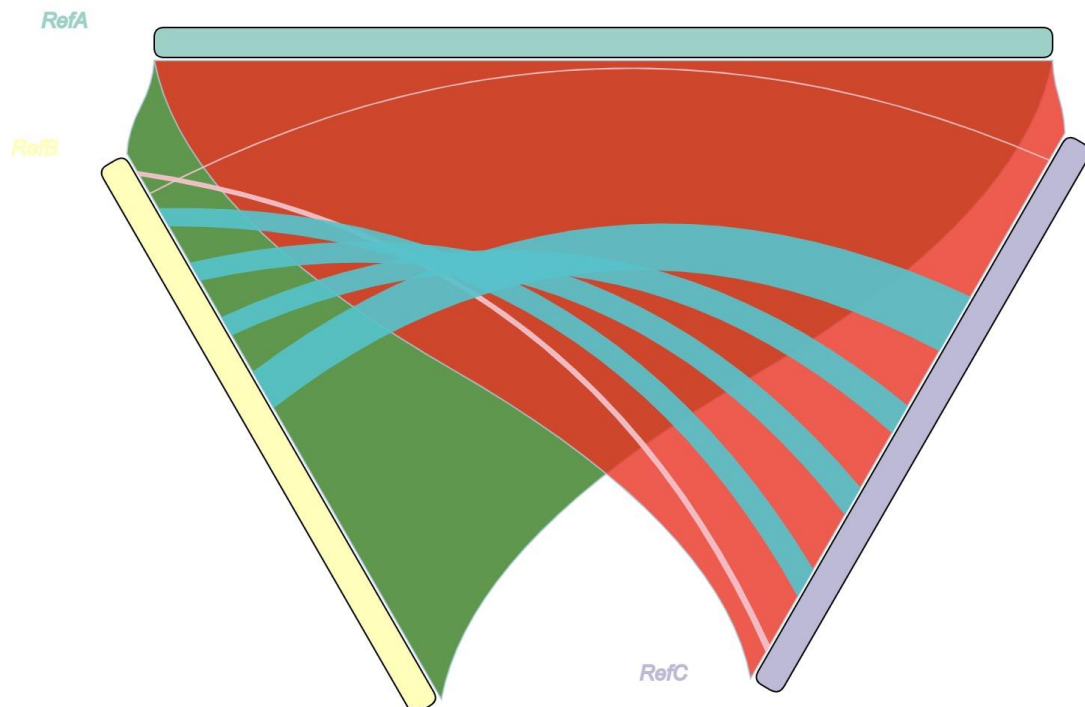
StyleUpDown=DownDown ## DownDown UpUp DownUp UpDown

Reverse=1

后图如下：



同样的数据，用 MoveToX/MoveToY ZoomChr RotateChr=-60 /60 构成三角型后如下：



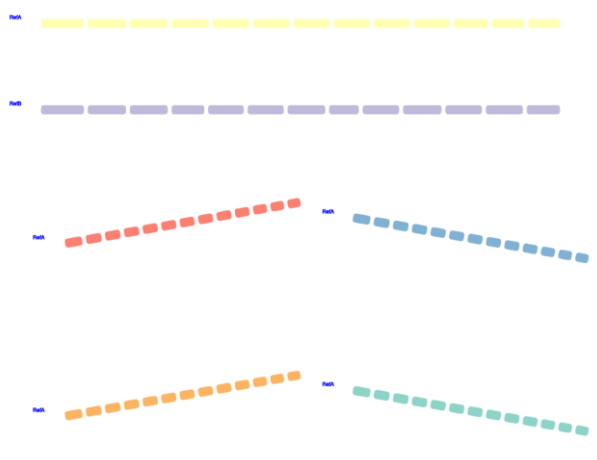
2.4 任意基因组任意分布

亮点:

- 1 任意多个基因组 (目前我限 12 个, 可以取消)
- 2 各个基因组可以自己调顺序 颜色和添加自己属性等等功能
- 3 各个基因组可以移动 旋转 伸长和收缩等
- 4 Link 信也可以自定义颜色和添加特定属性(如透明度)等等功能

其中 3 可以调好相关参数 可以出现 三角形 五角形 四边形 等等组合 (后面将有空升组调定这些参数)

现我这试调如下图:



配置如下:

SetParaFor = global

GenomeInfoFile1=RefA.info

GenomeInfoFile2=RefB.info	##	GenomeInfoFile X	就表示有 X 个基因组
GenomeInfoFile3=RefA.info	##	GenomeInfoFile X	就表示有 X 个基因组
GenomeInfoFile4=RefA.info	##	GenomeInfoFile X	就表示有 X 个基因组
GenomeInfoFile5=RefA.info	##	GenomeInfoFile X	就表示有 X 个基因组
GenomeInfoFile6=RefA.info	##	GenomeInfoFile X	就表示有 X 个基因组

#LinkFileRef1VsRef2=syn1.data

#LinkFileRef3VsRef2=syn1.data

SetParaFor = Genome3 # ALL/GenomeX X 第 X 个基因组

ZoomChr=0.4 ## chr 长度 等大 缩小 or 扩大

RotateChr=-10 ## chr 的起点 顺时针 旋转 28 度

ShiftY=110 ## 也可以用 MoveToX / MoveToY

ShiftX=-270

SetParaFor = Genome4 # ALL/GenomeX X 第 X 个基因组



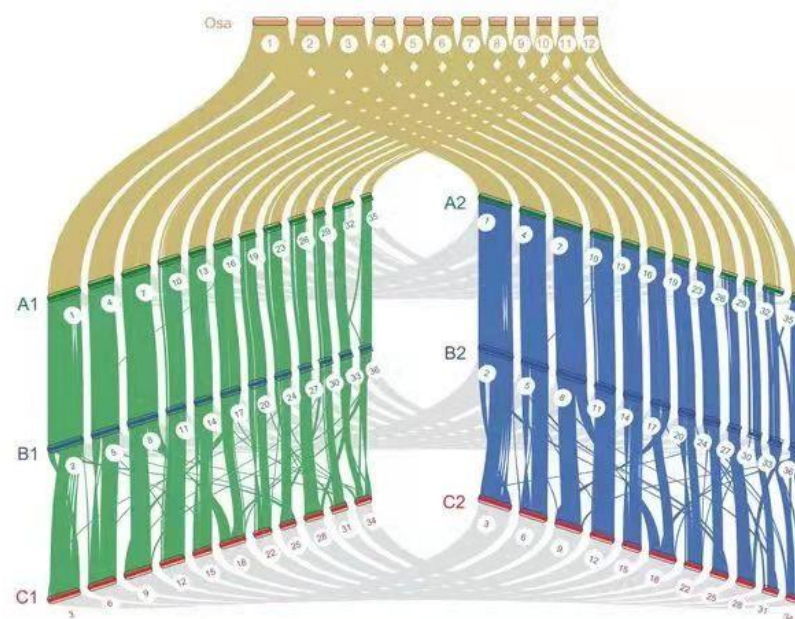

```
ZoomChr=0.4      ## chr 长度 等大 缩小 or 扩大
RotateChr=10     ## chr 的起点 顺时针 旋转 28 度
ShiftX=400
ShiftY=-150
```

```
SetParaFor = Genome5 # ALL/GenomeX X 第 X 个基因组
ZoomChr=0.4      ## chr 长度 等大 缩小 or 扩大
RotateChr=-10    ## chr 的起点 顺时针 旋转 28 度
ShiftY=110
ShiftX=-270
```

```
SetParaFor = Genome6 # ALL/GenomeX X 第 X 个基因组
```

```
ZoomChr=0.4      ## chr 长度 等大 缩小 or 扩大
RotateChr=10     ## chr 的起点 顺时针 旋转 28 度
ShiftX=400
ShiftY=-150
```

即可以如上排布，再结果自己定义颜色和属性 加上 links 数据必可以出下图：



3. 下载与安装

3.1 下载网址

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/08.Centos7/NGenomeSyn-*

里面有 example1-3

持续更新在这: <https://github.com/hewm2008/NGenomeSyn> 下载时记得加个星星哦

3.2 预先安装

NGenomeSyn 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)

```
git clone https://github.com/hewm2008/NGenomeSyn
cd NGenomeSyn; chmod 755 bin/*
./bin/NGenomeSyn -h # 直接运行
```



4. 用法和参数说明

4.1 NGenomeSyn 参数

程序 NGenomeSyn 很简单，一个输入和一个输出. 具体如下。

```
[heweiming@cngb-ologin-25 bin]$ ./bin/NGenomeSyn
Version:0.1      hewm2008@gmail.com

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

其中程序会对输入的配置文件进行简单的检测，若异常的配置如文件格式和行数 不符合时会报异常，麻烦检查.

另外提供了一个小 [perl GetTwoGenomeSyn.pl](#) 可以快速生成相关输入文件和图的， [该程序 GetTwoGenomeSyn.pl](#) 的参数说明和用法见[这儿](#)， 和 [这儿](#)

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 GenomeX 是对第 X 个的基因组进行配置
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	GenomeInfoFileX= /path/Ref.info	(X=1,2,3...) 输入的是第 N 个基因组 文件格式见下面
参数 2	LinkFileRefNVsRefM =./in.link	N,M <=X 第 N 个基因和 M 的 link 文件，可重复输入，格式见下面
参数 3	#ZoomChr=1.0	该 X 的基因组等比例放长 or 缩短

参数 4	#RotateChr=0	该 X 的基因组 在画布上 旋转角度
参数 5	#ShiftX=0 #ShiftY=0 MoveToX/Y	该 X 的基因组在画布上 移动 指定位置
参数 6	#ChrWidth=20	## 这个基因组 chr 的在画布的宽度
参数 7	#LinkWidth=180	## 这个基因组和下一个 link 的高度
参数 8	#ChrSpacing=10	## 这个基因组 chr 之间的空隙

4.1.2 其它配置参数

SetParaFor = GenomeX 为设置第 (X) 个基因组的一个 Flag 参数 global 是全局变量

ALL 是所有基因组都能管用, 各层没有配置的话, 则默认为全局(ALL)的参数。

下面列出一个示例 大家可以看后面##的解析

```

1. SetParaFor = global
2.
3. ##### 全局参数 #####
4.
5. GenomeInfoFile1=./RefA.info
6. ##### Format (chr Start End ...其它属性) chr 顺序和这文件一致 若是 End Start 则这条 chr 反向互补
7. ## 其它属性 如 fill=red stroke-width=0 stroke=black stroke-opacity=1 fill-opacity=1 等等可以不同行不同属性
8.
9. GenomeInfoFile2=./RefB.info ## GenomeInfoFile X 就表示有 X 个基因组
10.
11. LinkFileRef1VsRef2=./Link.file
12. ##### Format (chrA StartA EndA chrB StartB End ...其它属性)
13. # 可以多次 Ref1VsRef2 LinkFileRef2VsRef1 等
14.
15. #Main = "main_Figure" ## the Fig Name :MainRatioFontSize MainColor ShiftMainX ShiftMainY
16. # font-size
17.
18. ##### Figure #####
19.
20.
21. ##### 画布 和 图片 参数配置 #####
22. #body=1200 ## 默认是 1200, 主画布大小设置 另外: up/down/left/right) = (55,25,100,120); #CanvasHeightRitao=1.0 CanvasWidthRitao=1.0
23. ##RotatePng = 0 ## 对 Figure 进行旋转的角度

```



```

24.
25. SetParaFor = Genome1 # ALL/GenomeX X第X个基因组
26.
27. #ZoomChr=1.0          ## chr 长度 等大 缩小 or 扩大
28. #RotateChr=30         ## chr 的起点 顺时针 旋转 xx 度
29. #ShiftX=0
30. #ShiftY=0             ##对这个基因组移动
31. #MoveToX MoveToY      ## 直接移动到指定的点
32.
33.
34. #ChrWidth=20           ## 这个基因组 chr 的在画布的宽度
35. #LinkWidth=180         ## 这个基因组和下一个 link 的高度
36. #ChrSpacing=10         ## 这个基因组 chr 之间的空隙
37. #NormalizedScale=0     ## 用自己的标尺 这个相当该基因组与默认的基因组变成等长
38. #SpeRegionFile         ## 文件,标记特别区域[格式 chr start End (xx=yy 加属性等)]
39.
40. ## GenomeNameRatio GenomeName
41. ##其它当很少用到的参数 EndCurveRadian=3/ 等等
42. ## GenomeName GenomeNameSizeRatio GenomeNameColor GenomeNameShiftX
    GenomeNameShiftY
43. ## ChrNameShow 显示 chr 名及其属性 ChrNameShiftX ChrNameShiftY ChrNameSizeRatio
    ChrNameColor ChrNameRotate
44.
45. SetParaFor = Genome2 # ALL/GenomeX X第X个基因组
46.
47. SetParaFor=Link1 # 对第X个 Link X File 进行设置
48. #StyleUpDown=          ## UpDown DownUp UpUP DownDown 四种形式
49. #Reverse=1             ## 反向 link
50. #HightRation=1.0       ## links 的高比例 扩大 or 缩小
51. ##### fill/ stroke/stroke-opacity/fill-opacity/stroke-width 等可设##
52.

```

其中也提供有了更多参数，只是大家不常用到，就隐藏起来

如对 Main 标题的一些控制参数隐藏起来：在 global 下：

移动位置：ShiftMainX=0 /ShiftMainY=0; 字体颜色为：MainCol="blue"；字体大小：
MainRatioFontSize=1.0

ChrColorBrewer 和 GenomeColorBrewer



4.2 输入文件

两种输入文件

- A 一种是基因组信息, GenomeInfoFile
主要存 chr 的长度 和条数. 至少要两个基因组即两个.
- B LinkFile 两个基因组的 link 关系

4.2.1 数据文件(必须)

GenomeInfoFile 文件格式

```
chrB1    1    1000000
chrB3    1    1000000    fill="#F8F8F8"    stroke="#F8F8F8"
chrB2    1000000 1    fill=red    fill-opacity=0.5    ...
```

如上

- A 格式固定为前三列 Chr Start End , 其中 会 check Start End 其中一个是否为 1,后面若有 =号的为 作图属性, 不同行可以用不同属性
- B chr 的顺序为文件从上到下的序列, 即上面的序列为 B1 B3 B2 其中 B2 的 1 在 end 上, 即会对 B2 进行反向互补的作图,
- C 其它属性 fill 和 stroke 是颜色, stroke-opacity 和 fill-opacity 是透明度 B3 用了 "#F8F8F8" 填补颜色, B2 用 red. 颜色可以同时 fill=grey stroke=grey
- D 其他属性可以是 stroke-width=0 stroke=black ; stroke-opacity=1 fill-opacity=1 等等

LinkFileRef1VsRef2 文件格式

- A LinkFile 两个基因组的 link 关系 Ref XX Vs Ref YY . 表示第 XX 个基因组和 YY 个基因组的 link 文件
- B Ref XX Vs RefYY XX 和 YY 可以多次输入, 越后出来的就是层次越高
- C 前 6 列为 chrA StartA EndA chrB StartB End 后面若有=号为其属性
- D 示例文件如下:

```
chrB1    1    1000000    LG01    100000 1
chrB3    1    1000000    LG02    1 1000000    fill="#F8F8F8"
chrB2    1001    LG03    500    8000    fill=red    fill-opacity=0.5
...
```

重要提示 :



可以在 SetParaFor=Link1 # 对 fill 和 stroke , stroke-opacity 和 fill-opacity 和 stroke-width 对该 link 层所有 属性 进行设置, 其它属性只能在 文件里面设置。

4.2.2 配置颜色(可选)

颜色和数值 对应关系 #####
在这也顺列强调说明一下: 在没有自定义颜色时, Value 对应的颜色 是根据 1 颜色的颜色画板 [ColorBrewer](#) 先自动分配颜色 (点击可以查看更多), 2 Value 的种数 和 2 等分颜色数 及 3 Value 的 Count 数目 以及 4 最大最小值, 自动按一定的规则分配。 [更多详情点击这儿查看](#)

	chr	Start	End	Flag(Value)
这儿	ChrColorBrewer			和 GenomeColorBrewer 先对 Ref 和 link 默认配色
ChrColorBrewer				设置 chr 的默认颜色条 Paired
GenomeColorBrewer				设置不同基因组的默认颜色条 Set3
SpeRegionColorBrewer				设置特别区域的默认颜色条 Dark2

等我其它有空等再开放一些参数 和细节优化 ,暂时内定一些颜色。

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/NGenomeSyn -InPut in.cofi -OutPut out.svg
```

这儿共提示了 3 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: exampl1-3
具体见程序目录里面的 Example/example* 的数据和配置
在这搜索了一些在网上的其它人的配置和示意,点击可以找开网页, 查看

[NGenomeSyn 总简介](#)

NGenomeSyn 之 两个基因组

NGenomeSyn 各位画时

- 1 同时我希望有好图可以帮发个贴 教与别人
- 2 若图很好很美 可以把物种名 等改一下，隐藏一些信息 我把其写入 exampleXX 并 收藏进使用说明里面。

更多实例 随时更新，见 website 网页，具体见这：

<https://github.com/Hewm2008/NGenomeSyn> 里面的

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 应用场景广泛，应用场景多种多样，自主性较大，用户可以结合自己的数据画图。
- 4 免安装，使用方便

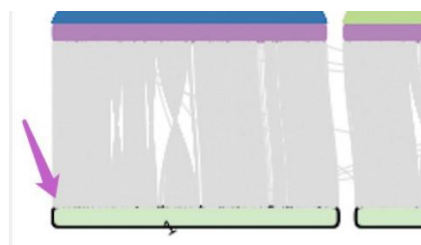
7.常见问题

7.1 NGenomeSyn 和 RectChr 的中的共线性差异

NGenomeSyn 主要可以任意个数，同时可以任意移动放在画布的任意位置，可以旋转等

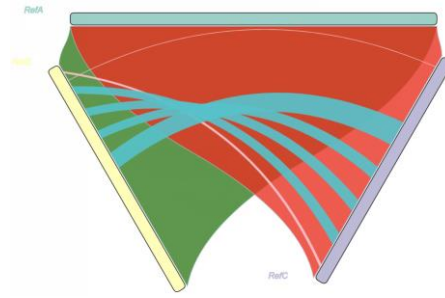
7.2 NGenomeSyn 的 chr 边框会被覆盖，如何调作？

如下图：

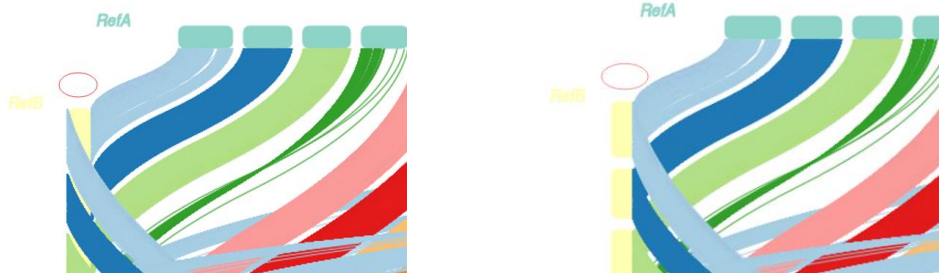


这个**边框会被覆盖**的原因，是 chr 的 `stroke-width=xx`，xx 外边框偏大的原因，那么此时须要调作的方法是把 link 层的上下缩小这个距离， 在对应的 link 层设置中 加上**隐藏参数** `OutsideWidthDeta=xx` （xx 建议和 `stroke-width` 同一数），如对 example3 各 link 层添加 `OutsideWidthDeta=2` 后如下：





另外提供了更多两个隐藏参数 ShortLinkLineRefA 和 ShortLinkLineRefB 即对 link 上下的两个基因组连接处缩短 or 变长。如对 link 添加参数 *ShortLinkLineRefB=20* 的前后变化如下：

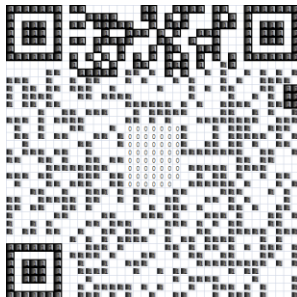


7.3 联系与打赏

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

微信公众号

