

CFVisual_V2.1

用户手册

作者：陈辉龙，尚千，葛伟娜

邮箱：chenhuilong131@163.com, Qian Shang16302@163.com

著作权证书号：软著登字第 6089858 号



目录

CFVisual_V2.1.....	1
目录.....	2
1. 软件简介.....	3
2. 编写目的.....	4
3. 系统操作.....	5
3.1 首页.....	5
3.2 操作.....	6
3.2.1 基因结构图绘制.....	7
3.2.2 Motif 图绘制.....	12
3.2.3 结构域图绘制.....	15
3.2.4 启动子图绘制.....	17
3.2.5 基因结构与 Motif 的组合图.....	19
3.2.6 基因结构和结构域的组合图.....	21
3.2.7 Motif 和结构域的组合图.....	23
3.2.8 基因结构、Motif 和结构域的组合图.....	24
3.2.9 其他组合.....	25
3.3.0 树图的交互、分类及美化.....	26
3.3.1 历史记录（File-History）.....	36
3.3.2 反馈功能（Setting-Feedback）.....	36
3.3.3 检查更新功能（Setting-Update）.....	36
4. 后序.....	37

Visual

1. 软件简介

CFVisual_V2.1 软件是一款借助 Python 语言的 matplotlib 库和 PySide2 库进行生物序列结构绘图的数据可视化软件，主要用于生物信息学分析中的基因结构分析，相当于 GSDS 的本地版 (<http://gsds.gao-lab.org/index.php>)，也能绘制进化树图，Motif 图，结构域图，启动子图，BED 文件等多种图表的绘制和组合绘制。



2. 编写目的

该用户手册是用户使用 CFVisual_V2.1 软件的操作说明，用户通过阅读用户手册，能够直接对 CFVisual_V2.1 软件的功能和操作有一定的了解，按照用户手册上的说明，通过实际操作，用户能够迅速掌握 CFVisual_V2.1 软件的使用方法。



3. 系统操作

3.1 首页

CFVisual_V2.1 可以免费从 GitHub 网址（<https://github.com/ChenHuilong1223/CFVisual>）下载。在获取该安装程序后，双击运行安装在用户电脑相应的位置即可。或点击 CFVisual_V2.1 软件依赖文件包中的 CFVisual.exe 即可进入到 CFVisual_V2.1 软件的首页界面。如下图所示：

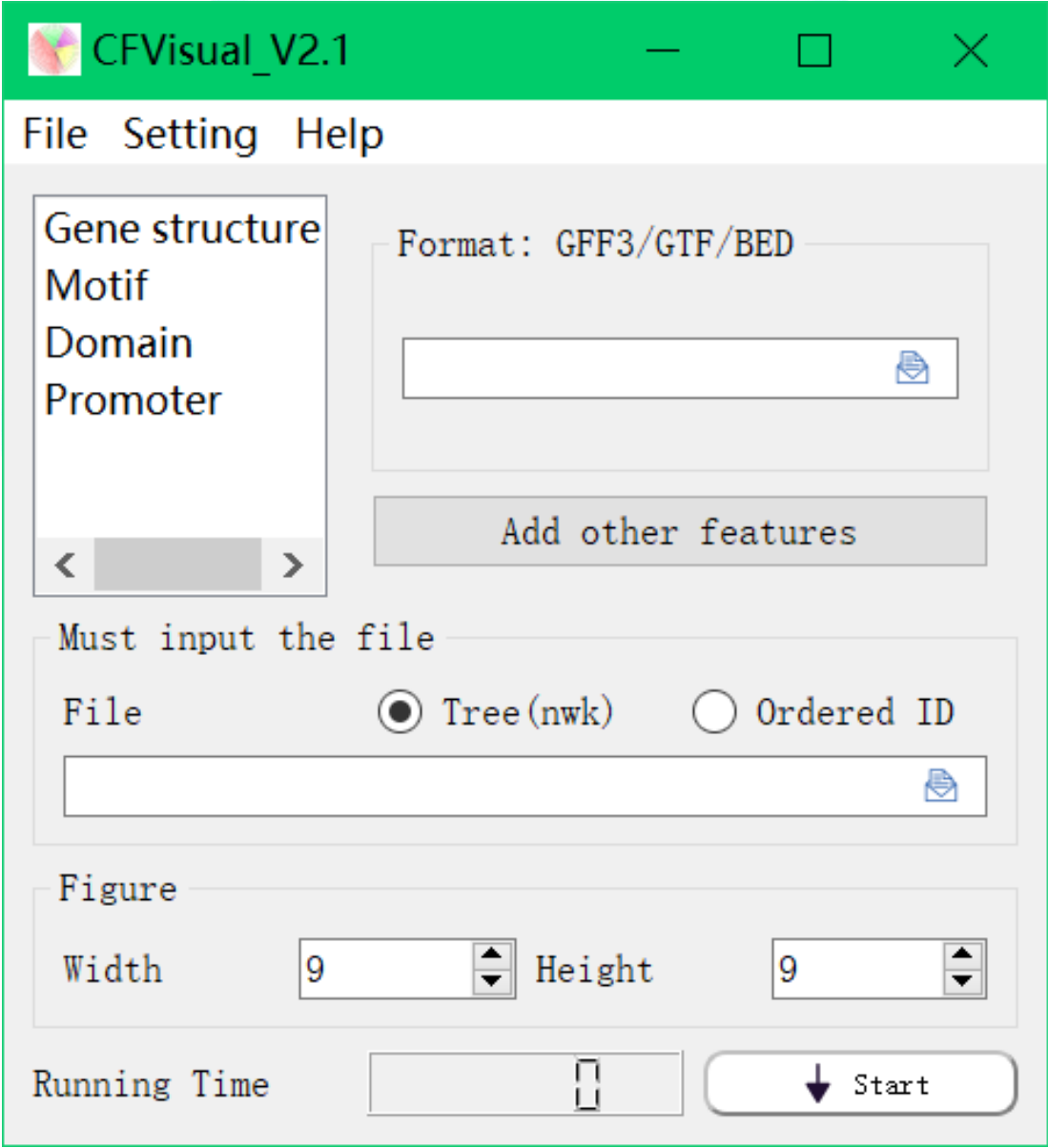


图 1 CFVisual_V2.1 首页面

3.2 操作

使用须知:

- 1) 用户在首页按照软件设计规则可以进行文件的预选工作，我们默认您必须准备 Newick format (后面简称 nwk 格式) 进化树文件 (Si.PME.Tree.Topology.Bootstrap.BranchLength0.nwk、Si.PME.Tree.Topology.Bootstrap1.nwk、Si.PME.Tree.Topology.BranchLength2.nwk 或 Si.PME.Tree.Topology3.nwk) 或者控制基因 ID 展示顺序文件 (Si.PME.Ordered.ID.txt)。
- 2) CFVisual 中 Gene structure 和 Motif 功能中增加的信息 gff 格式文件不是标准的 GFF3 文件，而是 Gene, Start, End, Feature 的这样从左到右的四列——相当于 GFF3 文件中的这四列。(Si.PME.Structure.and.Motif.add.gff.txt)



3.2.1 基因结构图绘制

支持 GFF3 (Si.PME.Structure.GFF3.txt), GTF 和 BED (Si.PME.Structure.BED.txt) 文件格式, 输入树文件和基因结构文件, 点击 Start 按钮, 稍作等待即可出图 (如图 2, 图 3)。

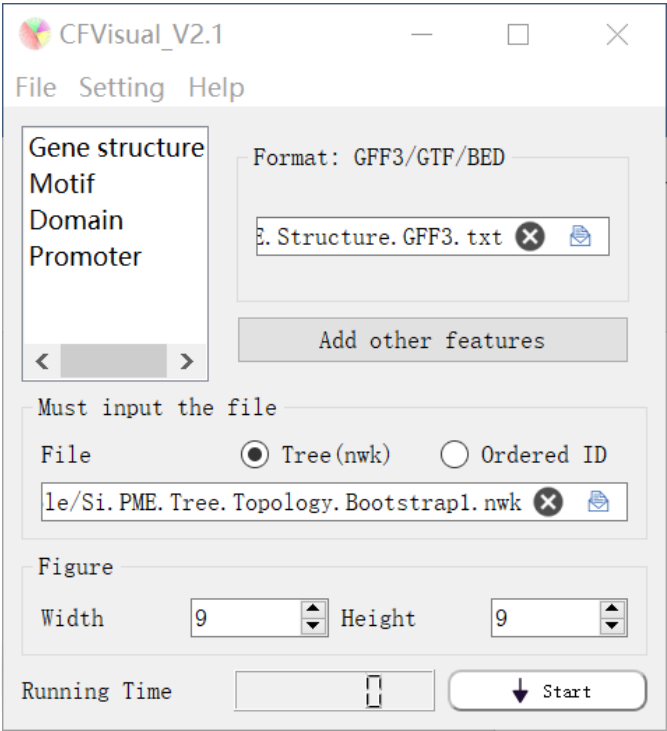


图 2 输入基因结构图准备文件

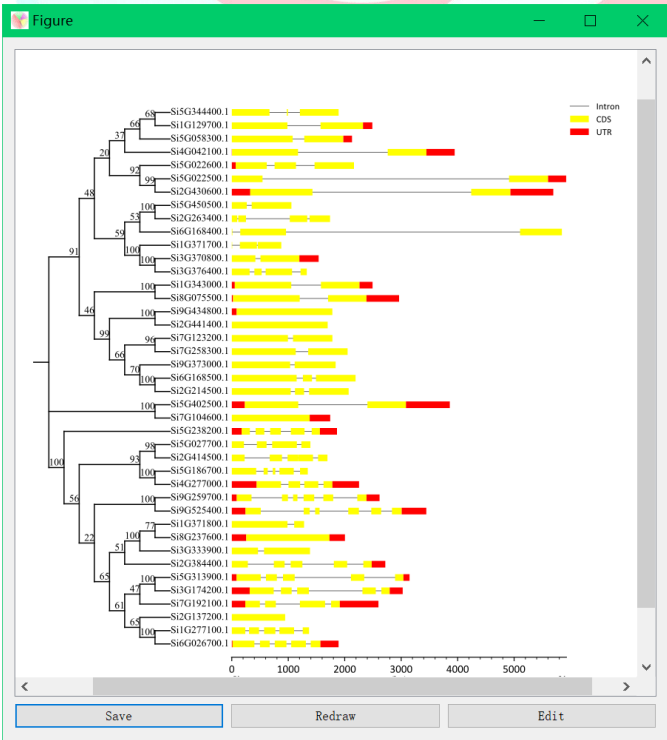


图 3 弹出显示基因结构图的窗口

点击 Edit 按钮，弹出编辑图片的窗口（图 4，图 5）

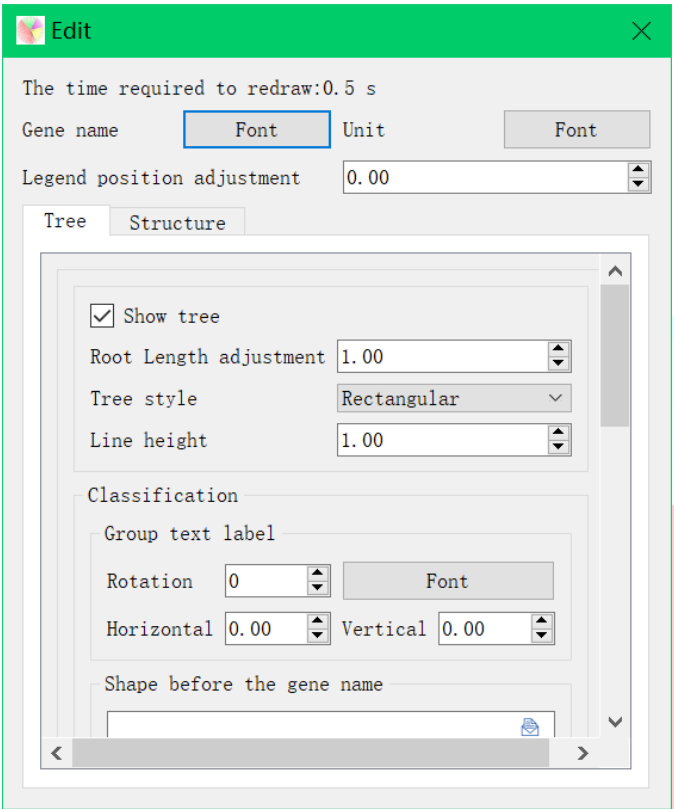


图 4 编辑树图

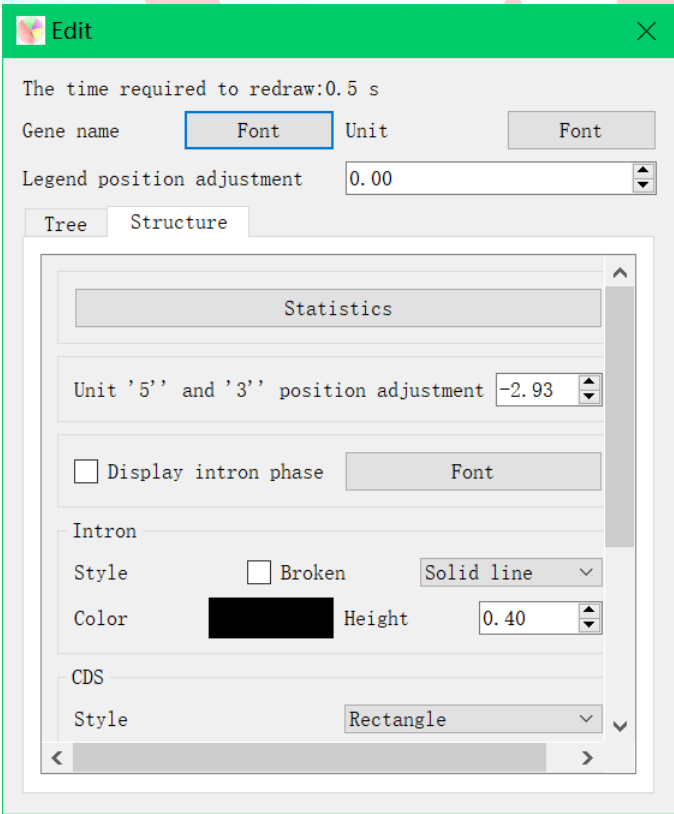
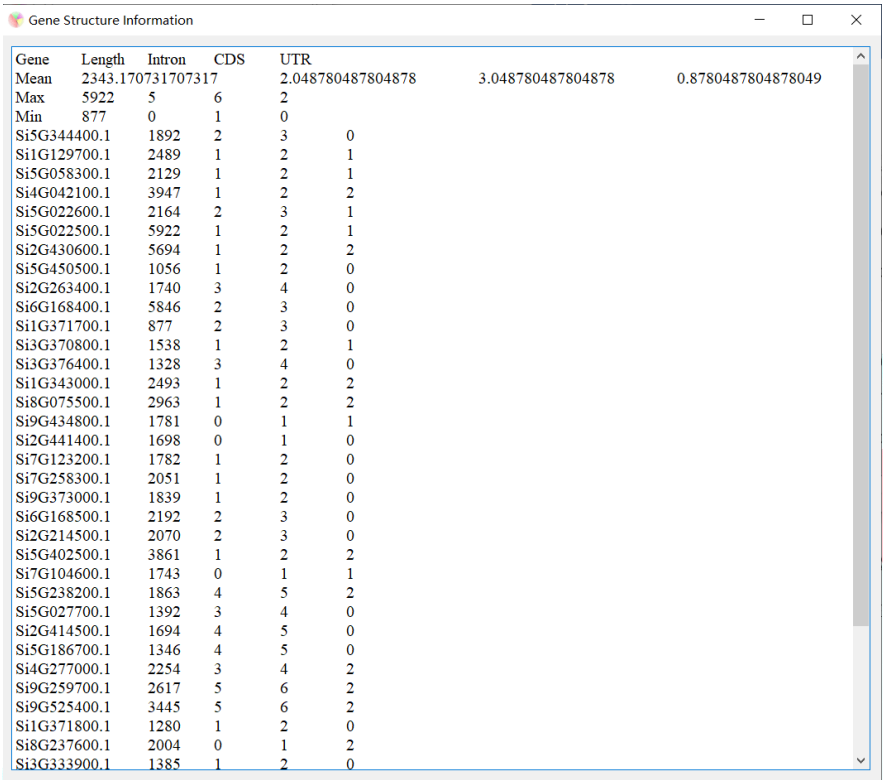


图 5 编辑基因结构图

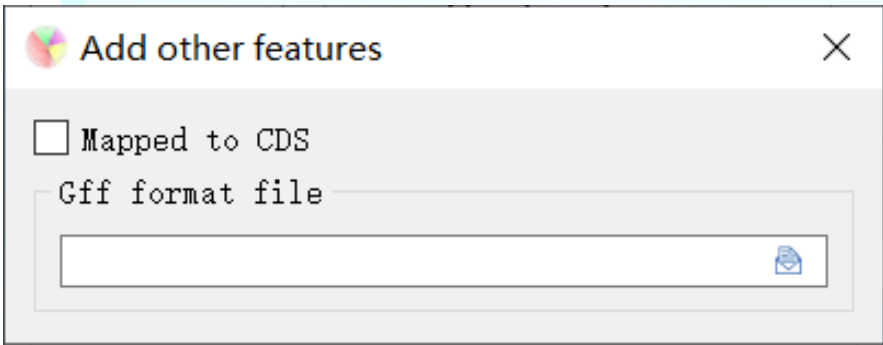
编辑基因结构图特色：点击 **Statistics** 按钮，弹出显示基因结构长度，内含子个数，UTR，CDS，exon 等数量信息（图 6）。（方便写论文时，对基因结构图的数字描述。）



Gene	Length	Intron	CDS	UTR
Mean	2343.170731707317		2.048780487804878	3.048780487804878
Max	5922	5	6	2
Min	877	0	1	0
Si5G344400.1	1892	2	3	0
Si1G129700.1	2489	1	2	1
Si5G058300.1	2129	1	2	1
Si4G042100.1	3947	1	2	2
Si5G022600.1	2164	2	3	1
Si5G022500.1	5922	1	2	1
Si2G430600.1	5694	1	2	2
Si5G450500.1	1056	1	2	0
Si2G263400.1	1740	3	4	0
Si6G168400.1	5846	2	3	0
Si1G371700.1	877	2	3	0
Si3G370800.1	1538	1	2	1
Si3G376400.1	1328	3	4	0
Si1G343000.1	2493	1	2	2
Si8G075500.1	2963	1	2	2
Si9G434800.1	1781	0	1	1
Si2G441400.1	1698	0	1	0
Si7G123200.1	1782	1	2	0
Si7G258300.1	2051	1	2	0
Si9G373000.1	1839	1	2	0
Si6G168500.1	2192	2	3	0
Si2G214500.1	2070	2	3	0
Si5G402500.1	3861	1	2	2
Si7G104600.1	1743	0	1	1
Si5G238200.1	1863	4	5	2
Si5G027700.1	1392	3	4	0
Si2G414500.1	1694	4	5	0
Si5G186700.1	1346	4	5	0
Si4G277000.1	2254	3	4	2
Si9G259700.1	2617	5	6	2
Si9G525400.1	3445	5	6	2
Si1G371800.1	1280	1	2	0
Si8G237600.1	2004	0	1	2
Si3G333900.1	1385	1	2	0

图 6 基因结构数值信息

增加其它信息：用户可以增加其它信息与基因结构图进行结合（图 7），增加信息文件格式为 gff 格式文件。如果用户勾选中 **Mapped to CDS**（图 8），那么位置信息将映射到 CDS 结构位置上（图 9）（方便了解哪段 CDS 编码该信息）。



Add other features

×

☐ Mapped to CDS

Gff format file

图 7

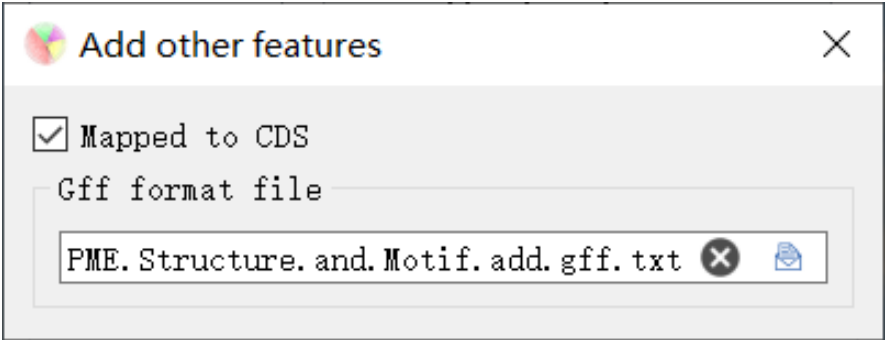


图 8

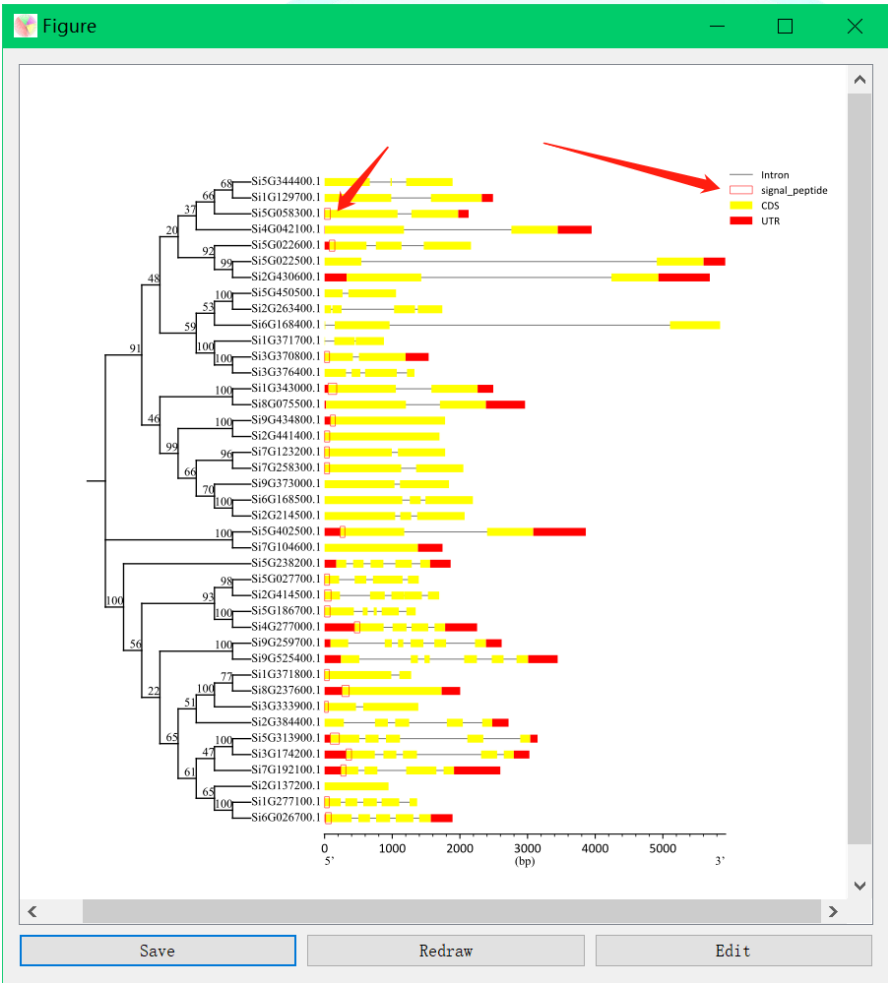


图 9

保存图片：目前支持 png（500dpi），pdf，svg，svgz，raw，rgba 这六种图片个格式（图 10）。注意：如果用户不选择路径，CFVisual 软件自动保存到用户输入的树文件或 ID 顺序文件所在文件夹中（图 11）（省去选文件夹的揪心操作）。

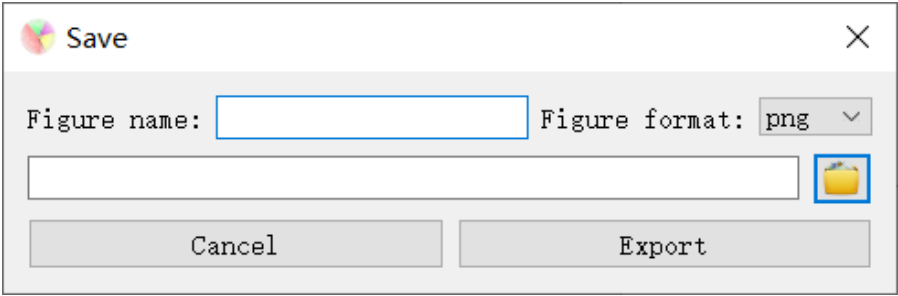


图 10 CFVisual 的图片输出窗口

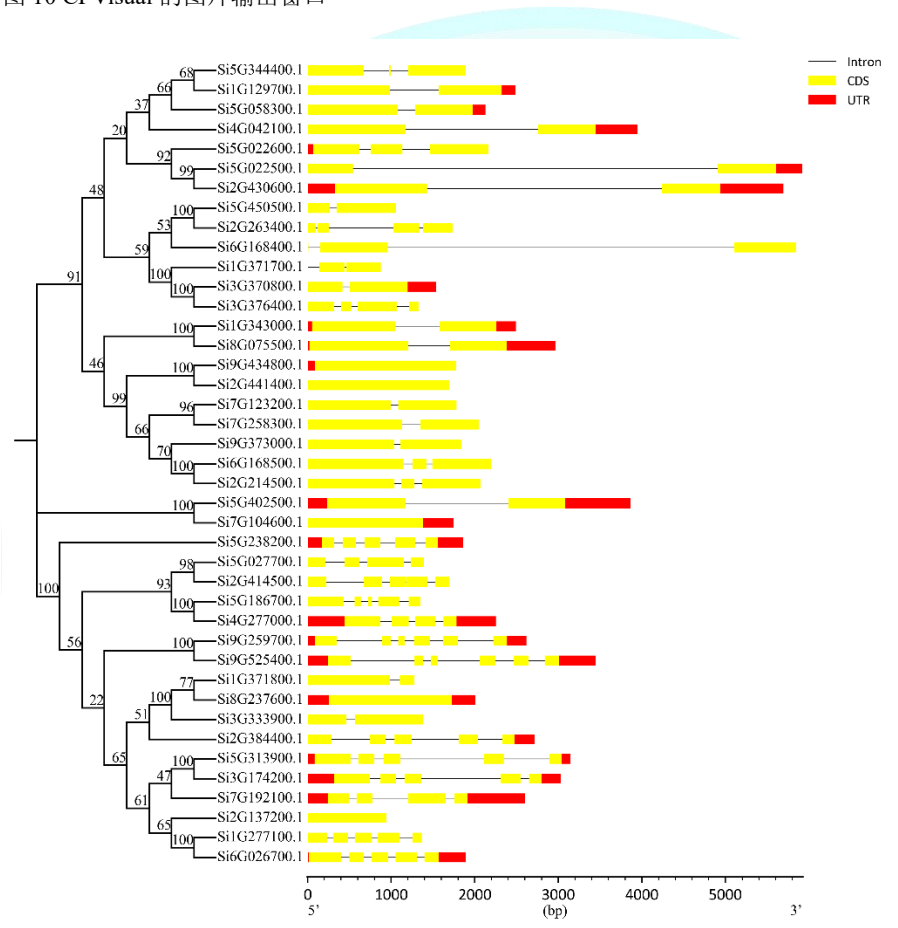


图 11 保存之后的 png 格式图片

3.2.2 Motif 图绘制

准备文件为 MEME 网站 (<https://meme-suite.org/meme/>) 或软件分析之后获取的 meme.xml (Si.PME.Motif.meme.xml) 或者 mast.xml (Si.PME.Motif.mast.xml) 文件。

注意: CFVisual 完全还原了 MEME 官网的 Motif 图结果。当用户输入的文件为 meme.xml 将有两种结果的展示, 分别为: Only Motif Sites (图 12) 和 Motif Sites+Scanned Sites (图 13)。可通过 Motif 编辑选项卡上的控件进行切换 (图 14)。另外, CFVisual 也支持用户选择性地隐藏用户不想要展示的 motif。

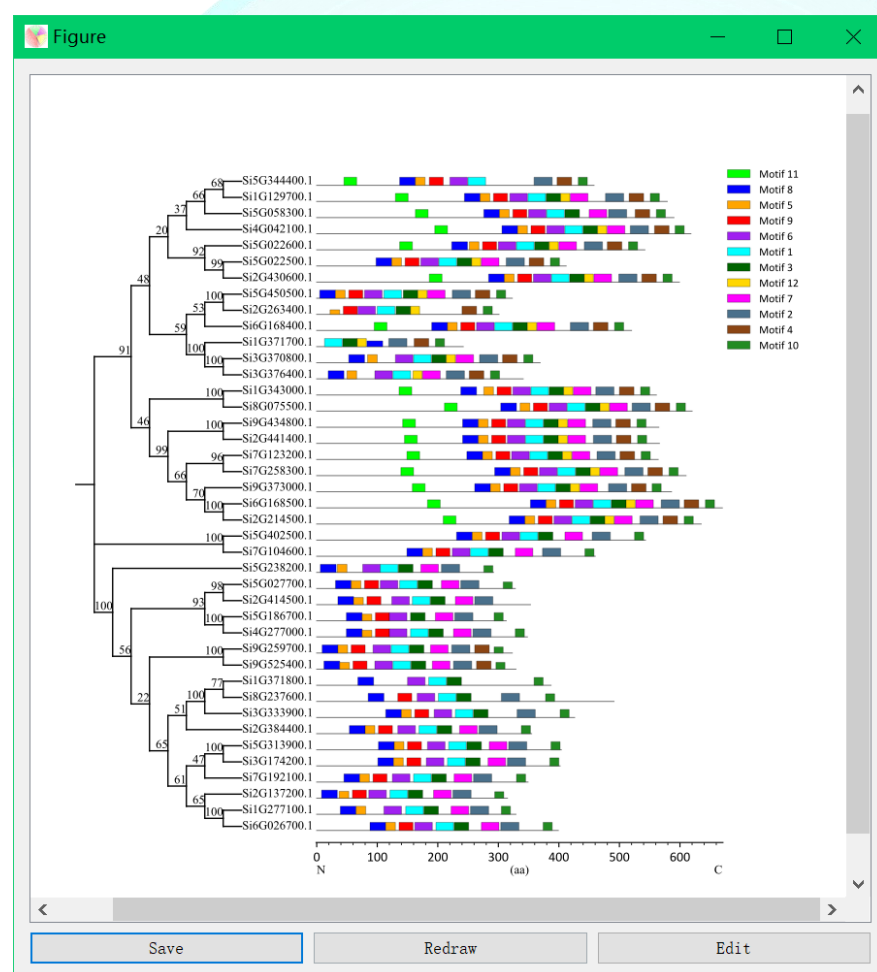


图 12 跟 MEME 官网一模一样的初始化结果 (Only Motif Sites)。

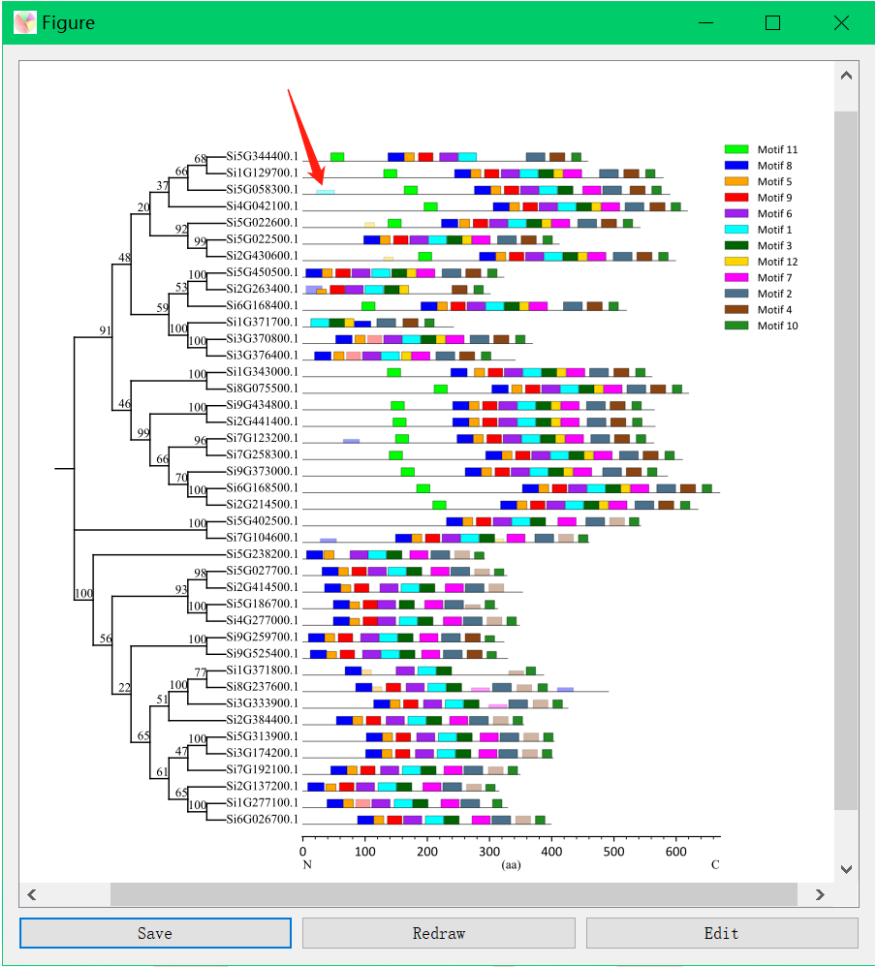


图 13 Motif Sites+Scanned Sites 的结果

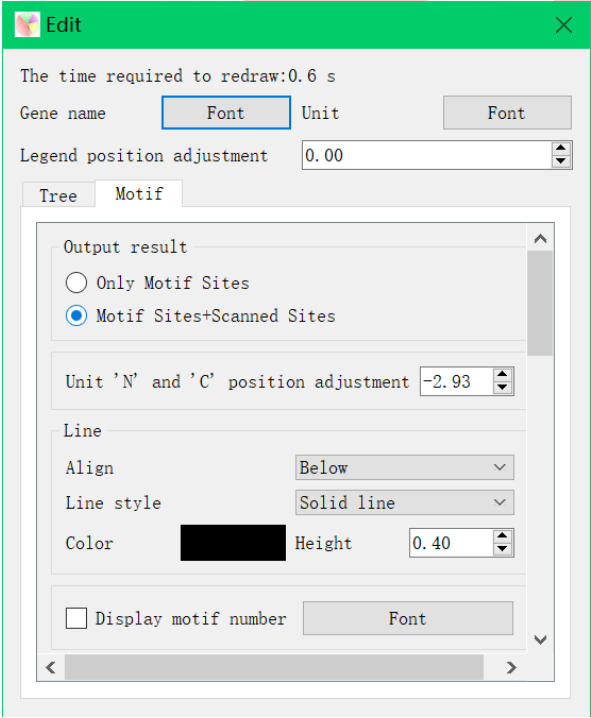


图 14

增加其它信息：用户可以增加其它信息与 Motif 图进行结合（图 15，图 16），增加信息文件类型为 gff 格式文件。

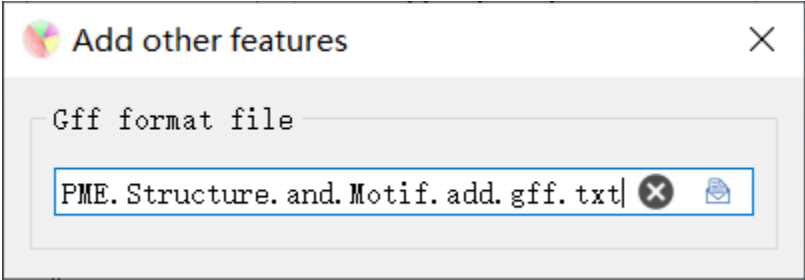


图 15

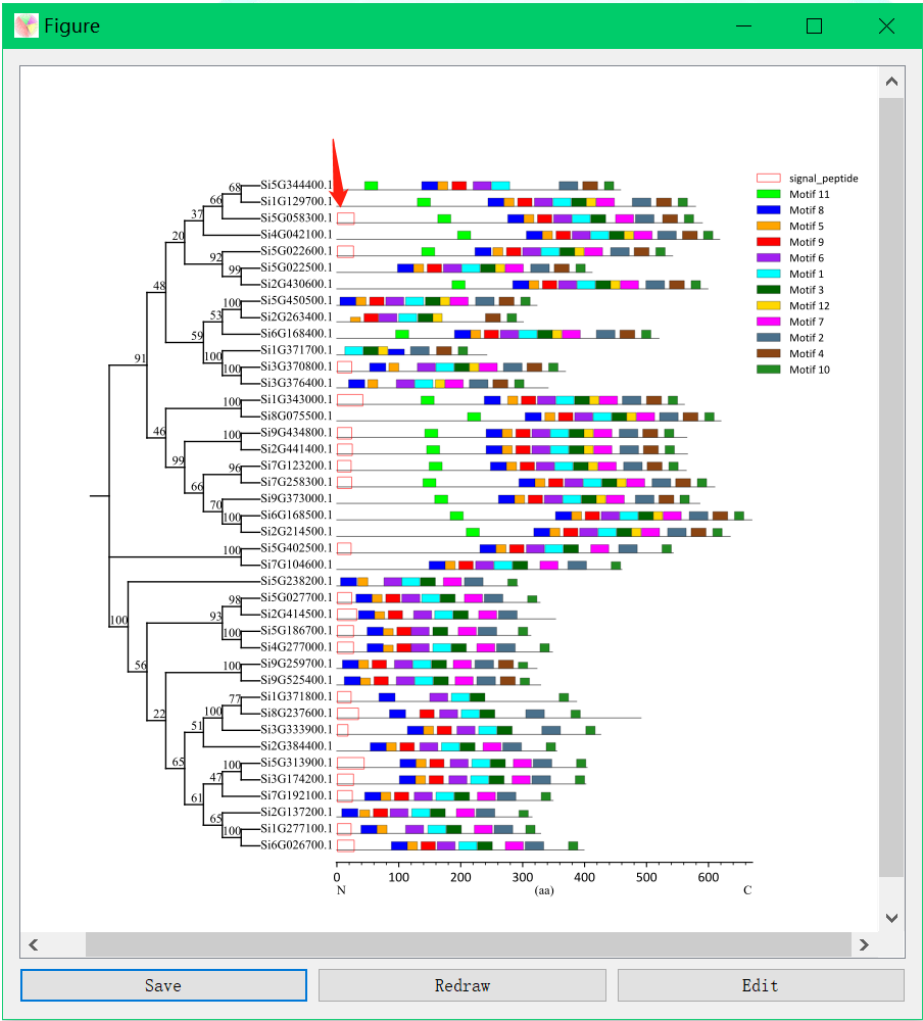


图 16

3.2.3 结构域图绘制

准备文件为 Pfam 网站 (<http://pfam.xfam.org/search#tabview=tab1>) 分析之后从邮箱获取的 Pfam 结果文件(Si.PME.Domain.Pfam.txt)或者 CDD 网站(<https://www.ncbi.nlm.nih.gov/cdd>) 分析之后下载的 CDD 结果文件 (Si.PME.Domain.CDD.txt) (图 17)。注意：在使用 CDD 结果文件的时候,建议用户先把 Short name 这一列的同一类结构域名字手动改成统一的名字。

结果 (图 18) 和用户交互功能 (图 19) 同上。

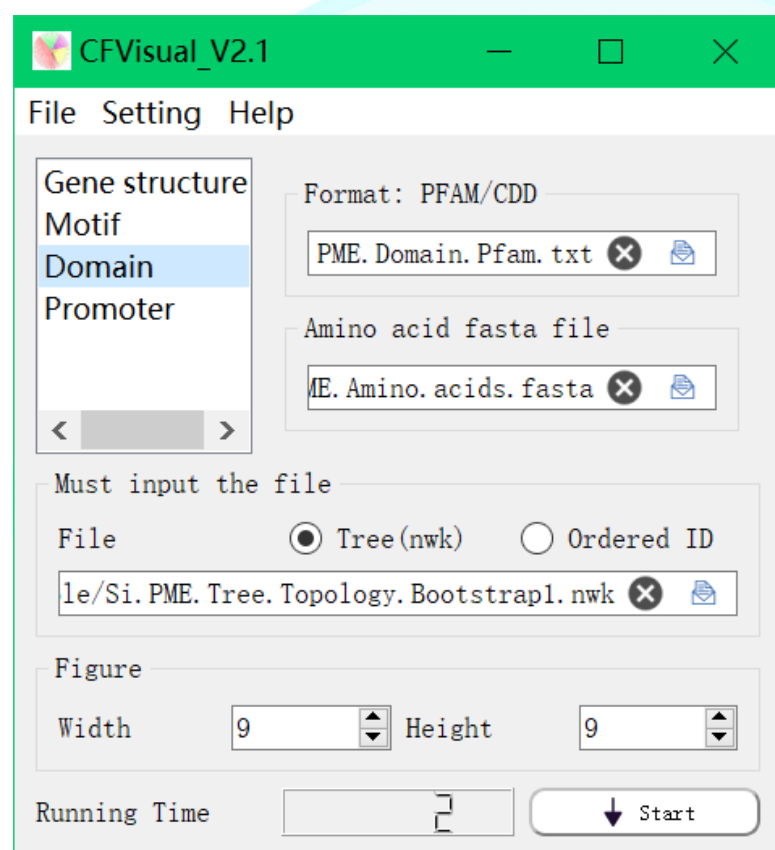


图 17

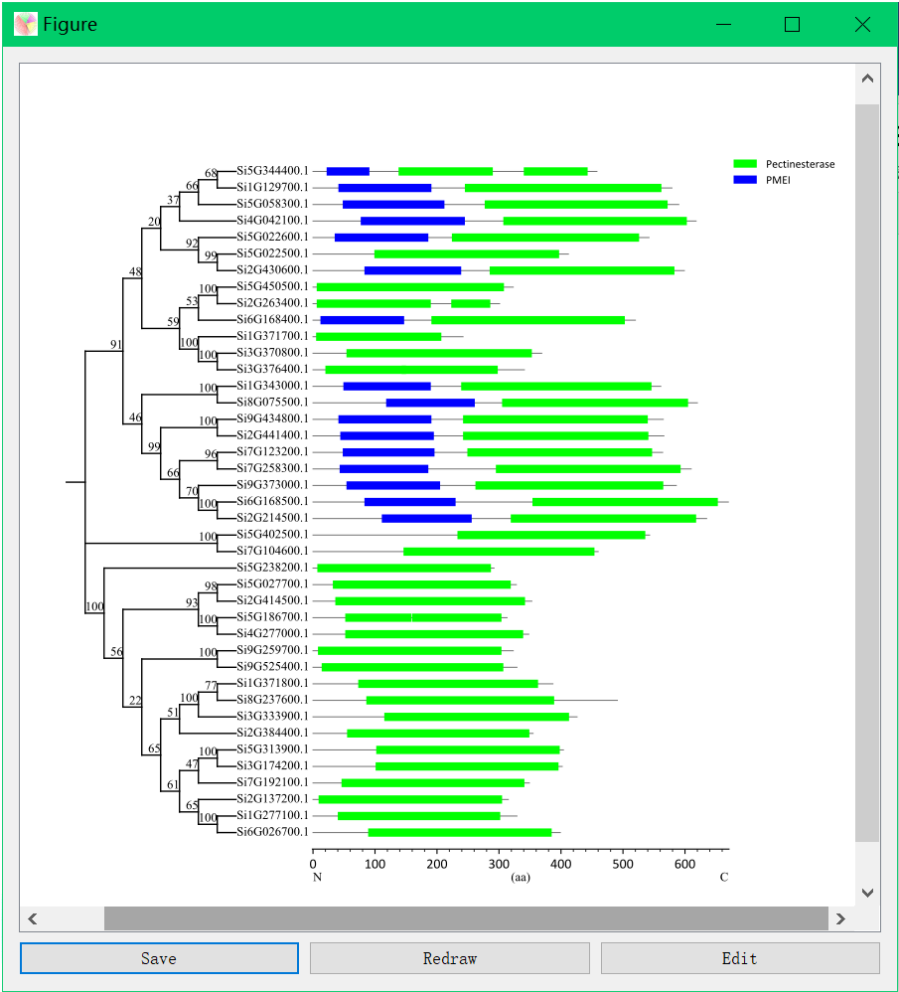


图 18

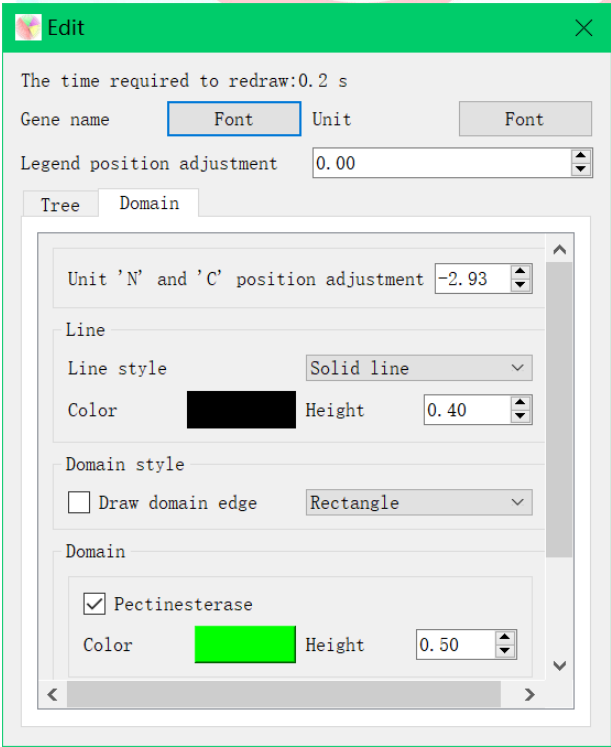


图 19

3.2.4 启动子图绘制

准备文件为 PlantCare 网站 (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) 分析之后获取的结果文件中的四列 (Si.PME.PlantCare.result.txt) (图 20) 或者其他启动子预测网站下载的结果, 只要结果格式能够处理成 Si.PME.PlantCare.result.txt 文件类型, 都能绘制 (作者没有用过其他启动子预测网站, 欢迎补充。)。注意: 使用 PlantCare 网站分析结果中, 要根据自己的课题需求稍作处理, 且只保留+链上的顺式作用预测结果。灵活地使用 CFVisual 的启动子绘图功能中的用户选择性地隐藏用户不想要展示的顺式作用元件控件, 可以达到用户筛选顺式作用元件的效果——换句话说, 用户获得 PlantCare 的结果之后, 只需要删掉多余的列和-链所在行的结果, 即可直接用 CFVisual 绘制, 不再需要花费繁琐的功夫去筛选顺式作用元件类别了。

结果 (图 21) 和用户交互功能 (图 22) 同上。

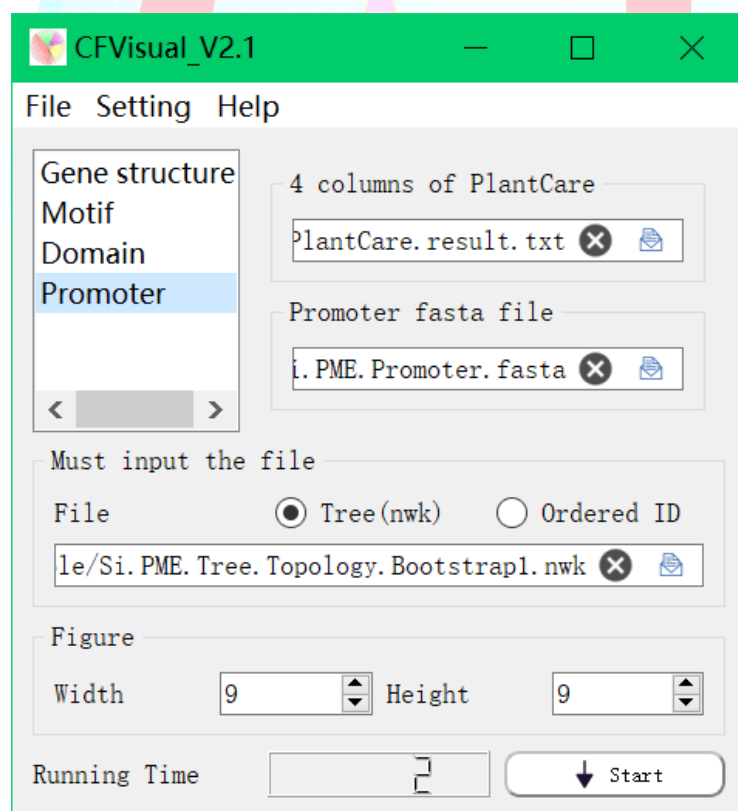


图 20

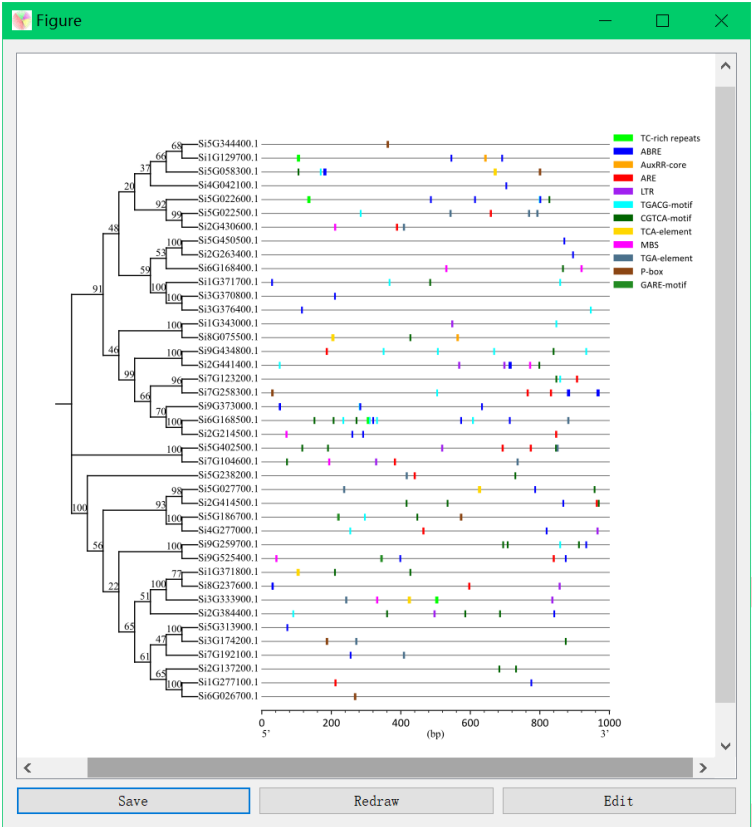


图 21

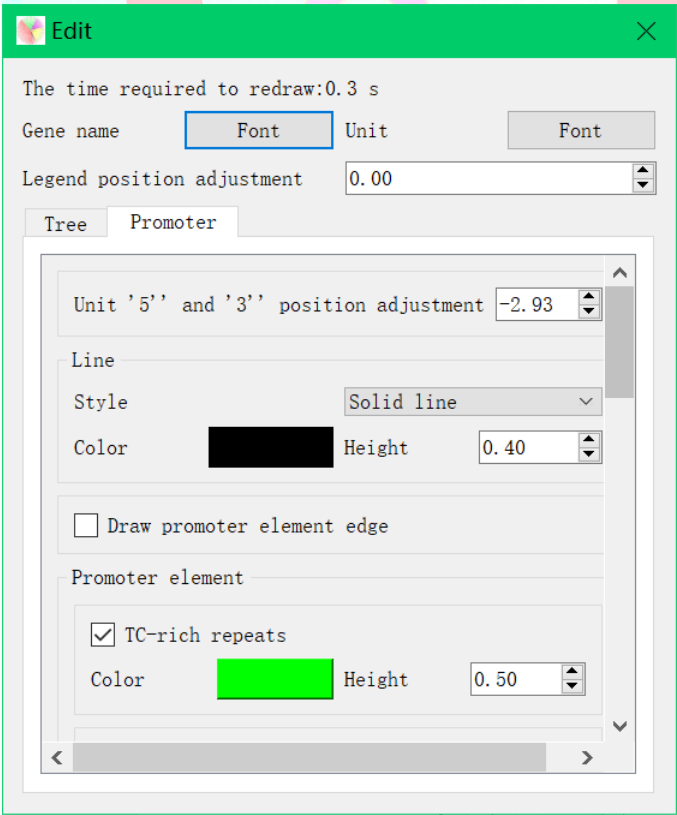


图 22

3.2.5 基因结构与 Motif 的组合图

作者读研三年，做的最多的生信分析就是基因家族分析了，深知做基因家族分析少不了这两个分析，而大多数人最喜欢的还是将树图、基因结构和 Motif 图进行组合成一张图展示，我们一般都是用 Ai 或 Ps 手动进行拼接，作者深知修图的繁琐和痛苦，因此就自己编程实现了这组合图的展示，一是方便自己，也方便师兄妹们。

准备文件：同基因结构图和 Motif 图绘制准备文件（图 23，图 24）。编辑方法同基因结构图和 Motif 图绘制（图 25，图 26）。作者考虑到这种情况，有的人喜欢把基因结构图放在中间，因此，特意增加该功能（Motif 编辑选项卡中的 Exchange coordinates）（图 25）。

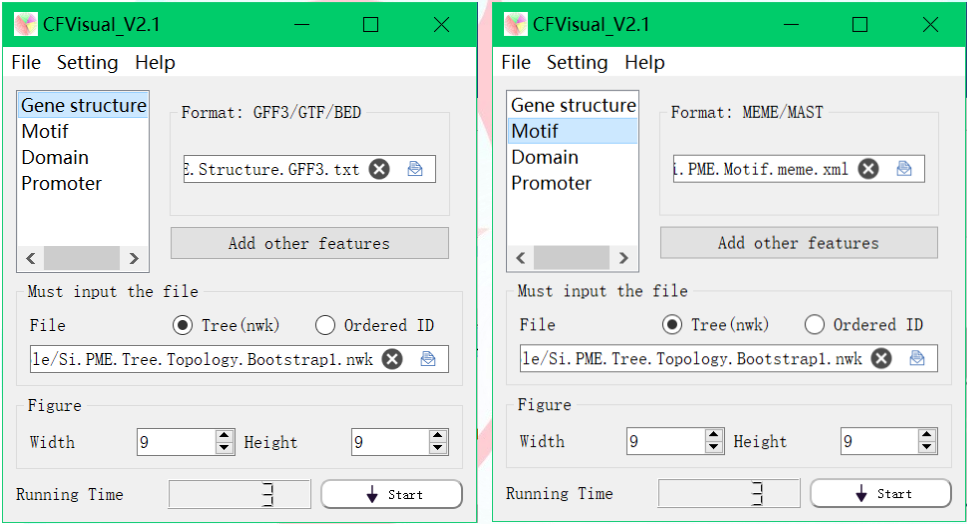


图 23 和图 24

Visual

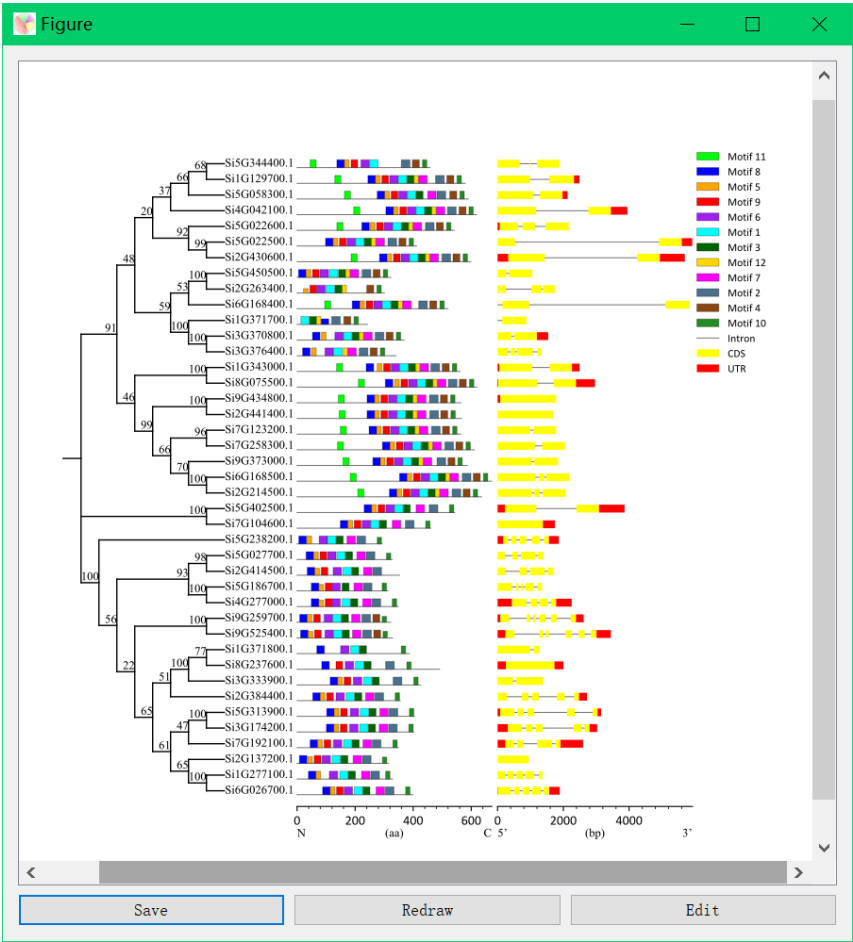


图 25

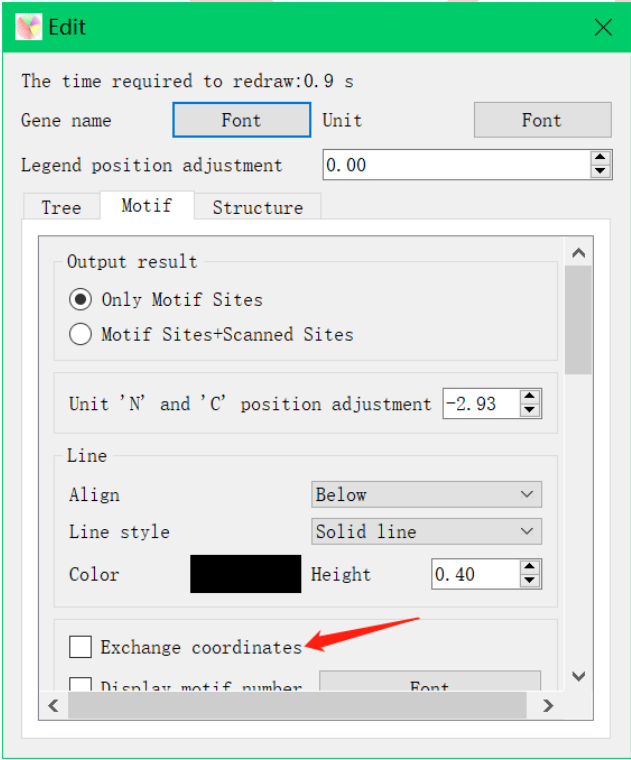


图 26

3.2.6 基因结构和结构域的组合图

这种组合图的作用在于方便科研人员了解是哪些 cds 编码结构域。准备文件同基因结构图和结构域图绘制（图 27），但是要注意一下，CFVisual 绘制组合图的时候，准备结构域文件时，不再需要增加氨基酸 fasta 文件（图 27）。基因结构结合结构域的功能 GSDS 其实已经有了，但是得自己准备氨基酸位置文件，不够方便，作者这里直接读取并操作 Pfam 或 CDD 结果文件获取位置信息，最需要注意的一点是，CFVisual 的结构域表现形式与 GSDS 不太一样，但本质一样，区别点在于，CFVisual 的结构域用的方框表示，把区间的内含子也框中了（图 28），而 GSDS 是实心区块并不给内含子上色。作者这样做的理由是试图打开一个新的基因结构分析思路，由于导师比较关注内含子，因此，作者认为，这样展示，在不失去原本用途的前提下，能比较方便地查看编码这段结构域的 CDS 中内含子的插入情况或者说丢失情况。

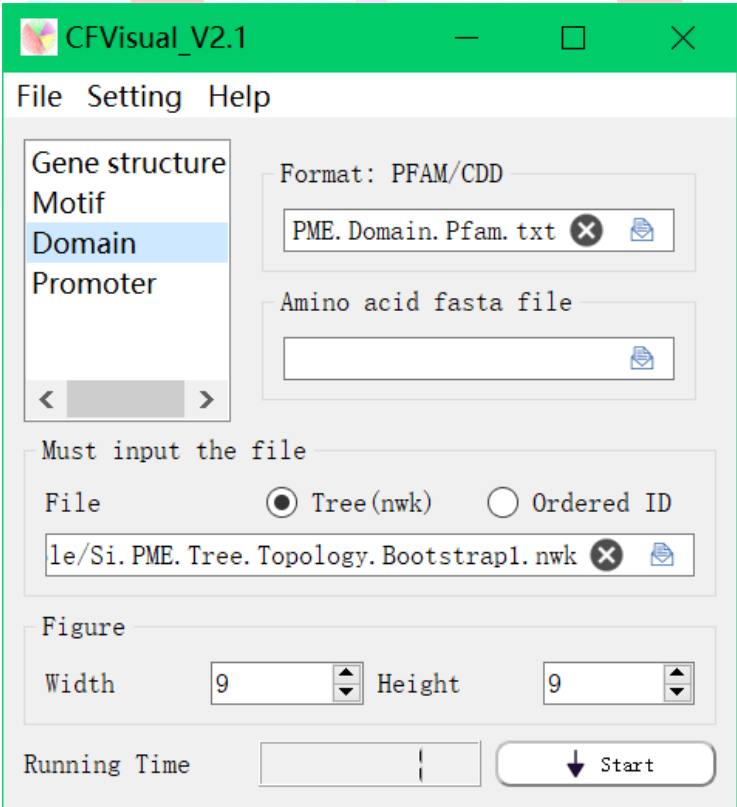


图 27

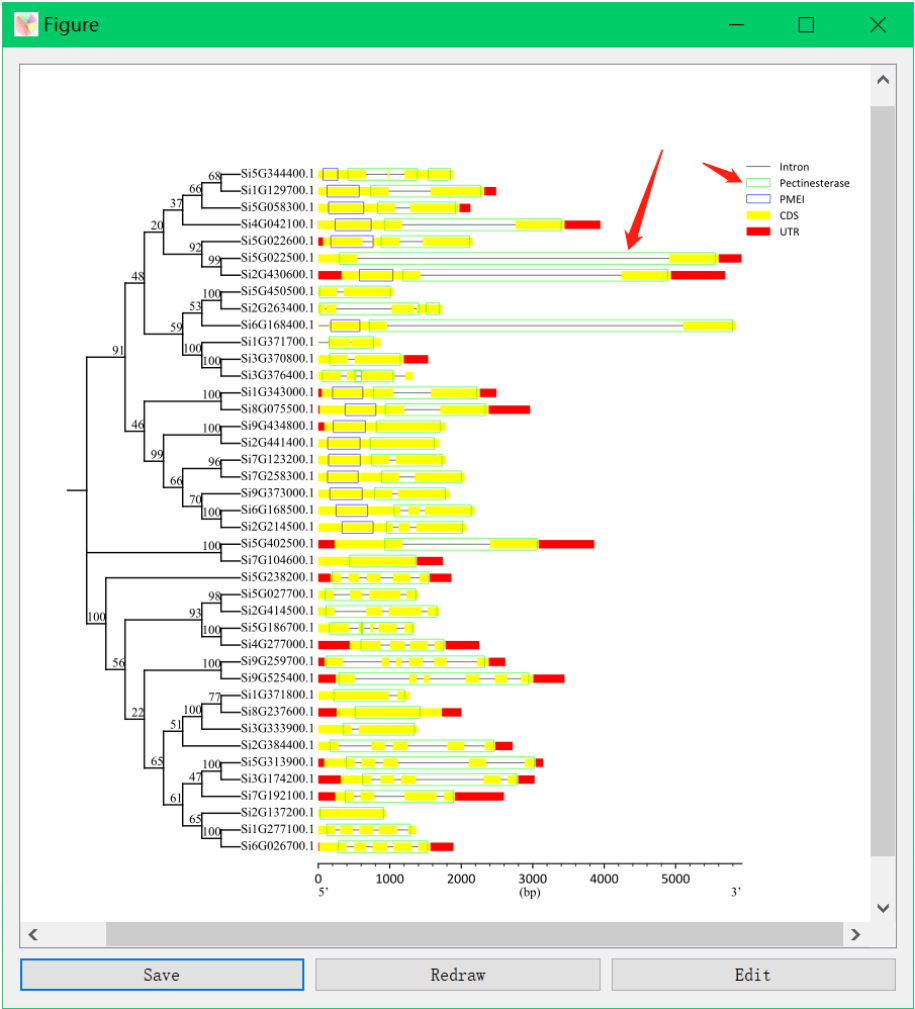


图 28

Visual

3.2.7 Motif 和结构域的组合图

这种组合源于一个师妹做抗病基因 NBS-LRR 家族，审稿人要求她把 motif 与结构域组合分析，看看结构域由哪些 motif 包含或者说结构域与 motif 的位置分布关系。由于目前没有找到拥有此功能的软件，她只能肉眼根据结构域位置标记，这样做，既麻烦，位置锚定还不一定准确。好在审稿人只要求她标记几个位置作为示意图即可，不然抗病基因那么大的家族，得标记到何年何月。基于此，作者实现了这种组合需求（图 29）。

准备文件同 Motif 图和结构域图绘制（图 27），同样要注意一下，CFVisual 绘制组合图的时候，准备结构域文件时，不再需要增加氨基酸 fasta 文件（图 27）。

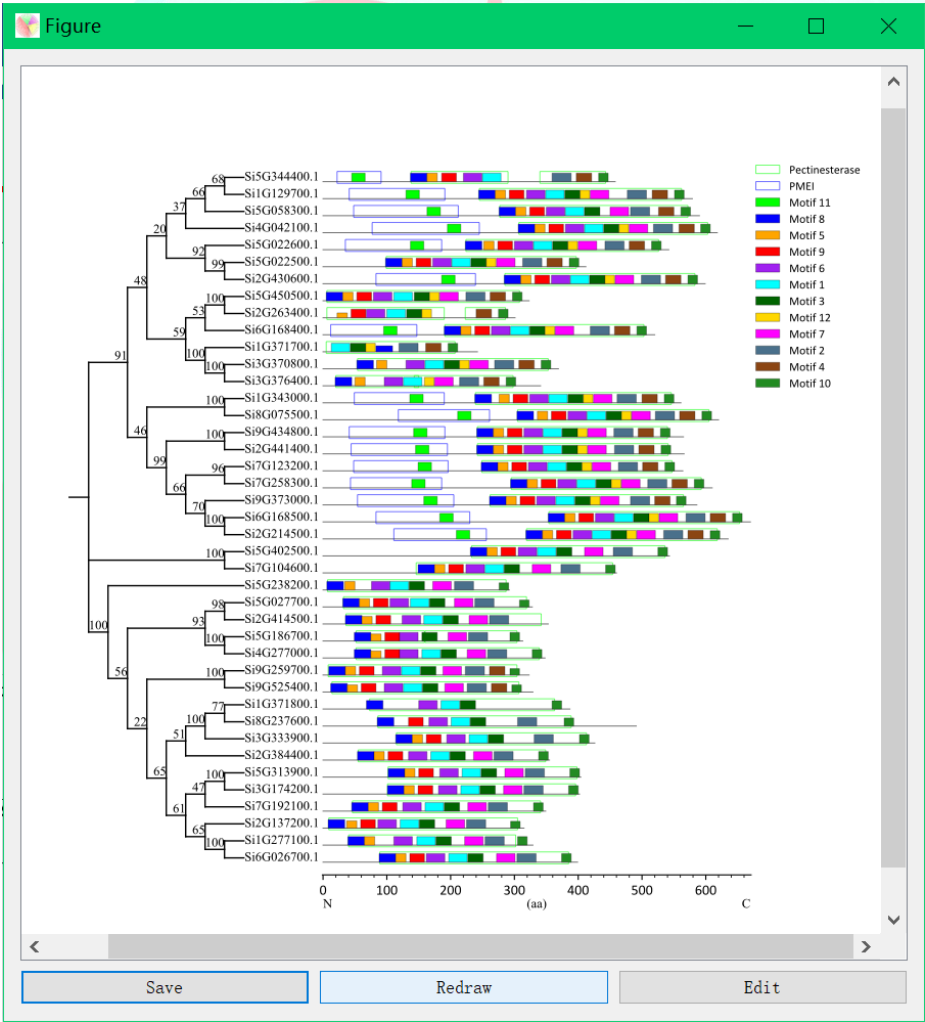


图 29

3.2.8 基因结构、Motif 和结构域的组合图

准备文件同基因结构图绘制，Motif 图绘制和结构域图绘制，同样的，绘制组合图时，结构域图不需要再准备氨基酸 fasta 文件。绘制结果如图（图 30），组合图的解读同上。

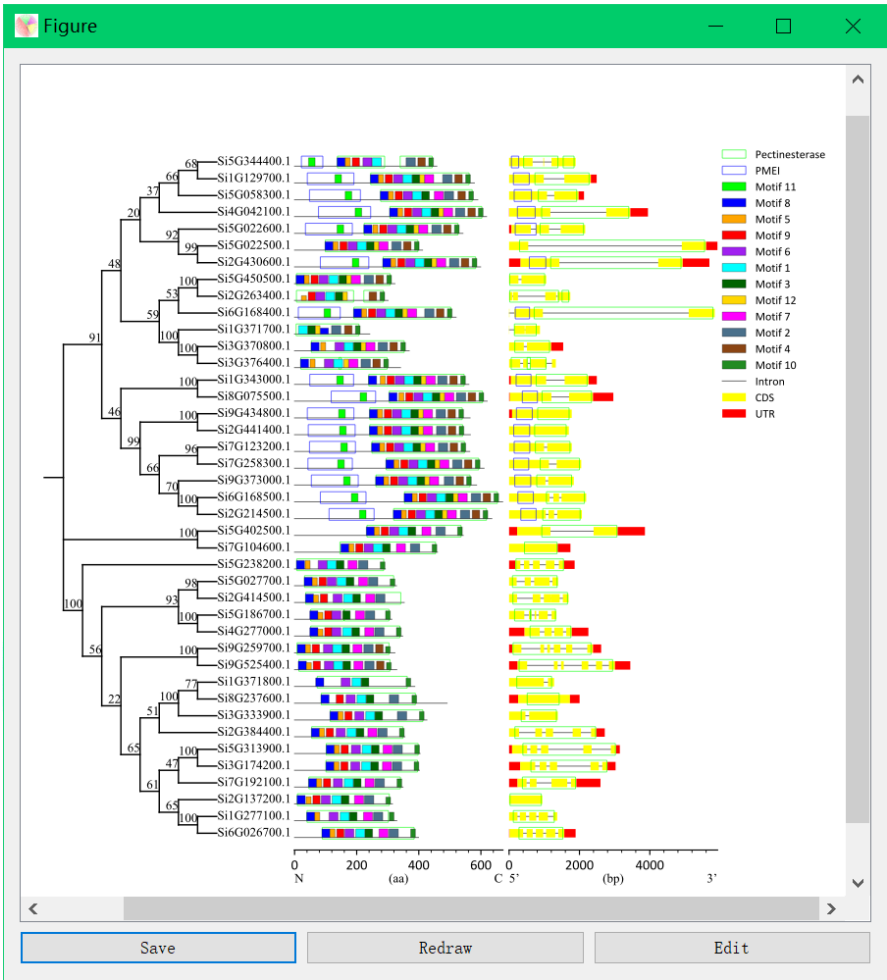


图 30

3.2.9 其他组合

作者想过其他组合图，比如启动子图与基因结构图的组合等，但是考虑到这种组合不常见，因此，暂不支持这种组合。



3.3.0 树图的交互、分类及美化

CFVisual 仅支持 nwk 格式的树文件，如果用户的树文件格式不是 nwk 文件，请使用其他软件把树文件转换成 nwk 格式的树文件，这里推荐 Figtree 软件 (<http://tree.bio.ed.ac.uk/software/figtree/>)，如图（图 31 和图 32）：

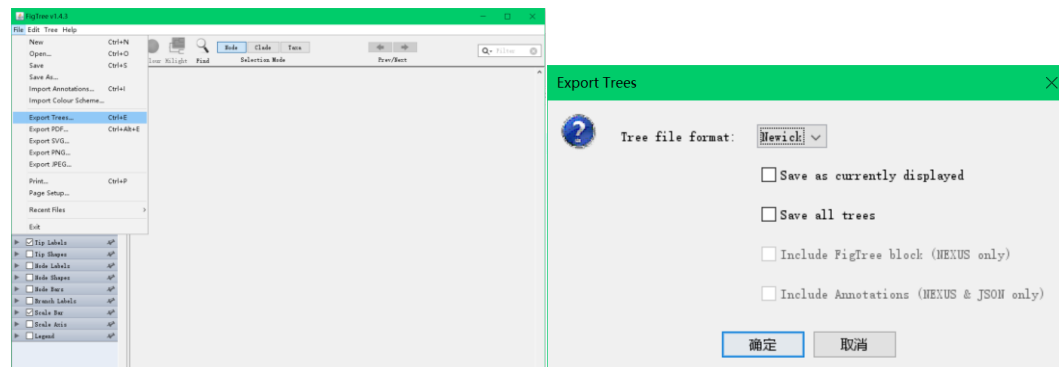


图 31 和图 32 用 Figtree 把树文件转换成 nwk 格式的步骤

nwk 格式的树文件从包含信息来讲，作者认为可以分为四类，分别为包含 bootstrap 值和枝长信息的 nwk 文件（Si.PME.Tree.Topology.Bootstrap.BranchLength0.nwk）（图 33）、只包含 bootstrap 值的 nwk 文件（Si.PME.Tree.Topology.Bootstrap1.nwk）（图 34）、只包含枝长信息的 nwk 文件（Si.PME.Tree.Topology.BranchLength2.nwk）（图 35）和最简单的拓扑结构的 nwk 文件（Si.PME.Tree.Topology3.nwk）（图 36）。

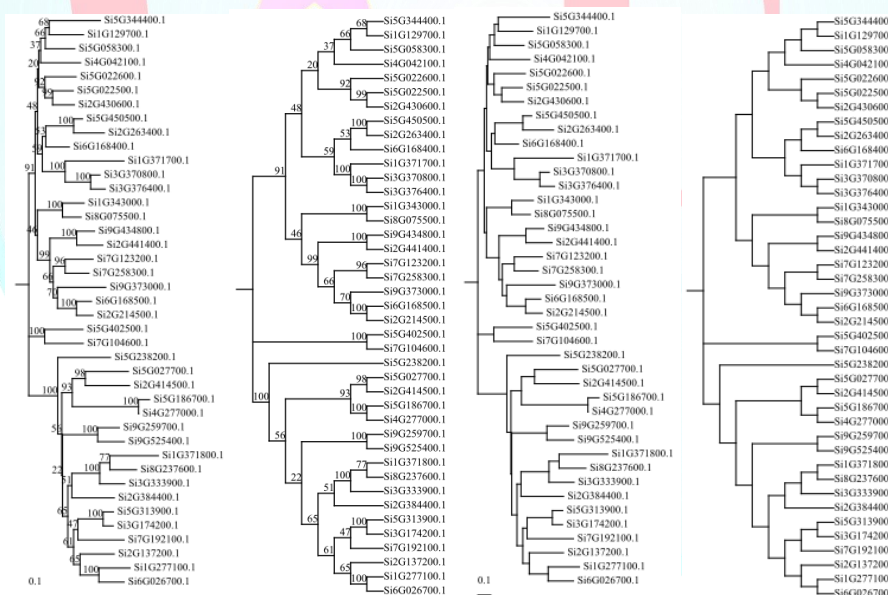


图 33、图 34、图 35 和图 36 nwk 格式的四类文件

与树图交互的边界窗口如下图所示（图 37 和图 38），用户可以自行体验：

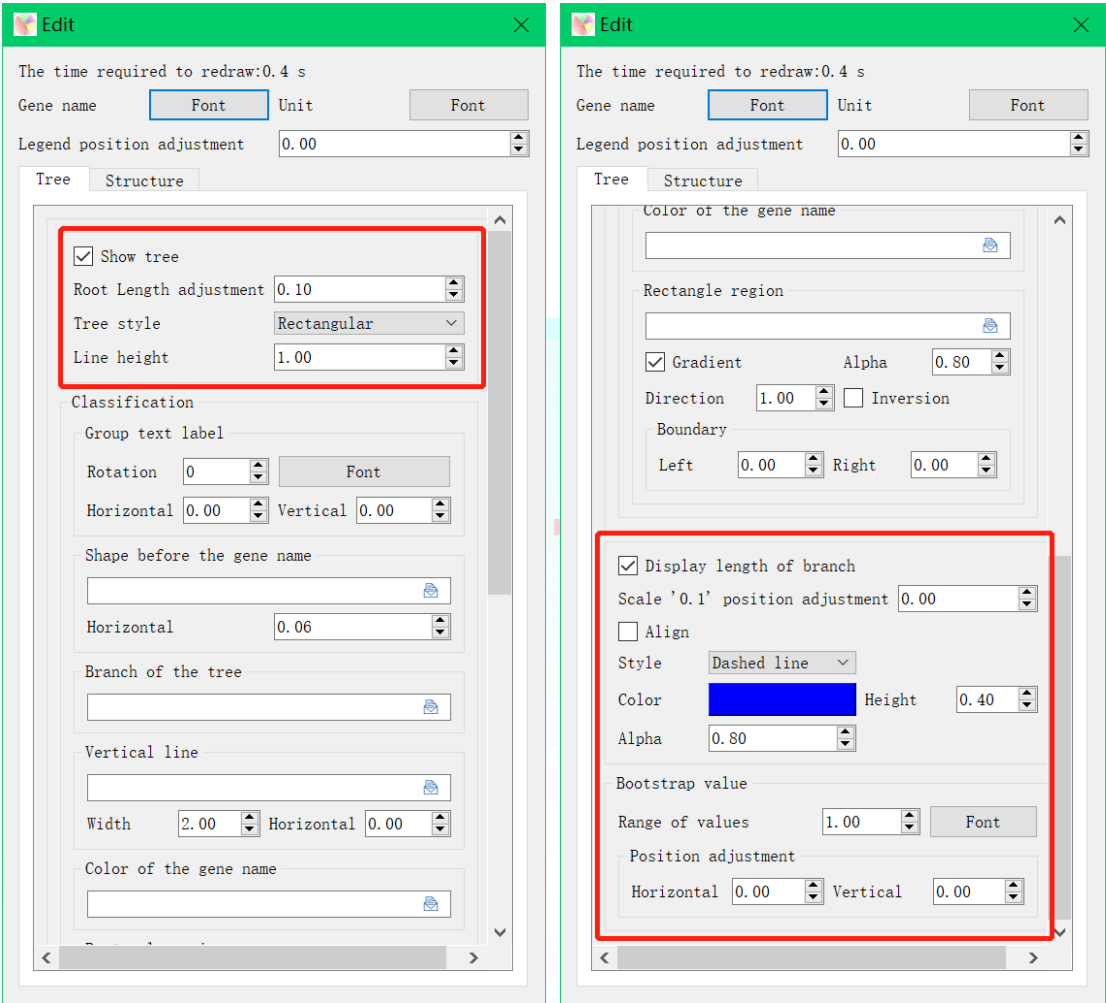


图 37 和图 38 树图交互界面

Visual

由于作者深知要真正实现基因家族分析中这个多图合一，树图的分类及美化功能是必须得有的，这里值得仔细说明的是关于**树图的分类及美化**。

首先，CFVisual 支持两种可以识别的分类及美化文件，分别为**基因名前的形状控制文件**（**Si.PME.Tree.Shape.before.gene.name.txt**）和**树图分类控制文件**（**Si.PME.Tree.Classification.txt**）。

基因名前的形状控制文件格式如下图（图 39）。用户根据自己的需求，利用 EXCEL 制作一个 CFVisual 可以识别的交互格式文件：**从左到右的列分别表示为：基因名、形状符号、形状尺寸、形状填充颜色和形状轮廓颜色**。其中，形状符号理解如图 40，颜色参数支持常用的 8 个表示颜色的单词或简写（图 41）和十六进制的颜色编码字符串（图 42）——这里作者推荐菜鸟工具中的一个板块（<https://c.runoob.com/front-end/55>），因为不仅支持颜色查找，也支持 RGB 格式和十六进制颜色编码字符串的相互转换，值得收藏。

Si1G371700.1	o	5	blue	#00FF00
Si2G414500.1	v	10	cyan	#00FF01
Si9G259700.1	^	15	green	#00FF02
Si1G277100.1	<	20	black	#00FF03
Si2G263400.1	>	25	magenta	#00FF04
Si3G333900.1	8	30	red	red
Si2G384400.1	s	35	white	white
Si5G450500.1	p	40	yellow	yellow
Si9G525400.1	*	50	b	b
Si6G026700.1	h	55	c	#00FF01
Si8G237600.1	D	60	g	#00FF02
Si5G186700.1	X	55	k	#00FF03
Si3G376400.1	o	50	m	#00FF04
Si5G022500.1	v	45	r	#00FF05
Si2G137200.1	^	40	w	#00FF06
Si5G313900.1	<	35	y	y
Si5G027700.1	>	30	#FF0000	#FF0000
Si7G192100.1	8	25	#00FF00	red
Si3G370800.1	s	20	#0000FF	red
Si1G371800.1	p	15	#FFA500	#FFA500
Si5G238200.1	*	10	#A020F0	#FFA500
Si7G104600.1	h	5	#00FFFF	red
Si3G174200.1	D	10	#006400	red
Si4G277000.1	X	15	#FFD700	#FFD700
Si5G402500.1	o	20	#FFD700	#FFD700
Si5G344400.1	v	25	#FFD700	#FFD700
Si5G058300.1	^	30	#FFD700	#FFD700
Si5G022600.1	<	35	#FFD700	#FFD700
Si9G434800.1	>	40	#FFD700	b
Si9G373000.1	8	45	#FFD700	b

图 39

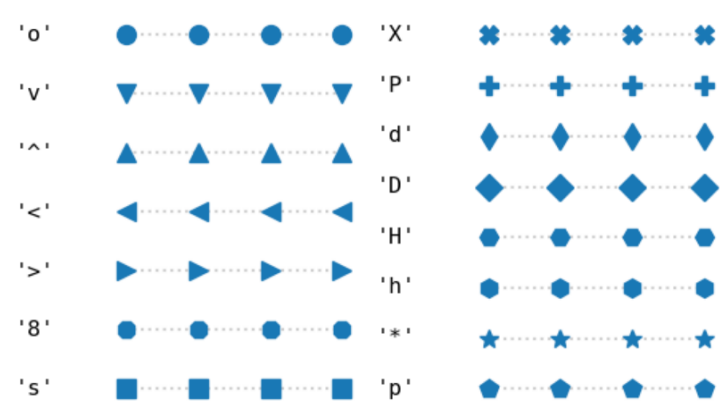


图 40



图 41

RGB转16进制工具用于将RGB颜色值与十六进制字符串相互转换，工具使用简单，你只需要在以下三个输入框：红(R)、绿(G)、蓝(B)中输入RGB的颜色值及会自动回转换十六进制的字符串，反之你只想要在右侧十六进制输入框输入16进制字符串即可转换为 RGB 数值。

RGB 与 HEX 颜色对照表							
实色效果	英文名称	R.G.B	16色	实色效果	英文名称	R.G.B	16色
	Snow	255 250 250	#FFFAFA		PaleTurquoise1	187 255 255	#BBFFFF
	GhostWhite	248 248 255	#F8F8FF		PaleTurquoise2	174 238 238	#AEEEEE
	WhiteSmoke	245 245 245	#F5F5F5		PaleTurquoise3	150 205 205	#96CDCD
	Gainsboro	220 220 220	#DCDCDC		PaleTurquoise4	102 139 139	#668888
	FloralWhite	255 250 240	#FFFAF0		CadetBlue1	152 245 255	#98F5FF
	OldLace	253 245 230	#FDF5E6		CadetBlue2	142 229 238	#8EE5EE
	Linen	250 240 230	#FAF0E6		CadetBlue3	122 197 205	#7AC5CD
	AntiqueWhite	250 235 215	#FAEBD7		CadetBlue4	83 134 139	#53868B
	PapayaWhip	255 239 213	#FFEFD5		Turquoise1	0 245 255	#00F5FF
	BlanchedAlmond	255 235 205	#FFEBCD		Turquoise2	0 229 238	#00E5EE

图 42

准备好的基因名字前形状设置文件，即可在在树图编辑窗口输入文件（图 43），再点击 Figure 窗口的 Redraw 按钮进行重新绘制（图 44）。

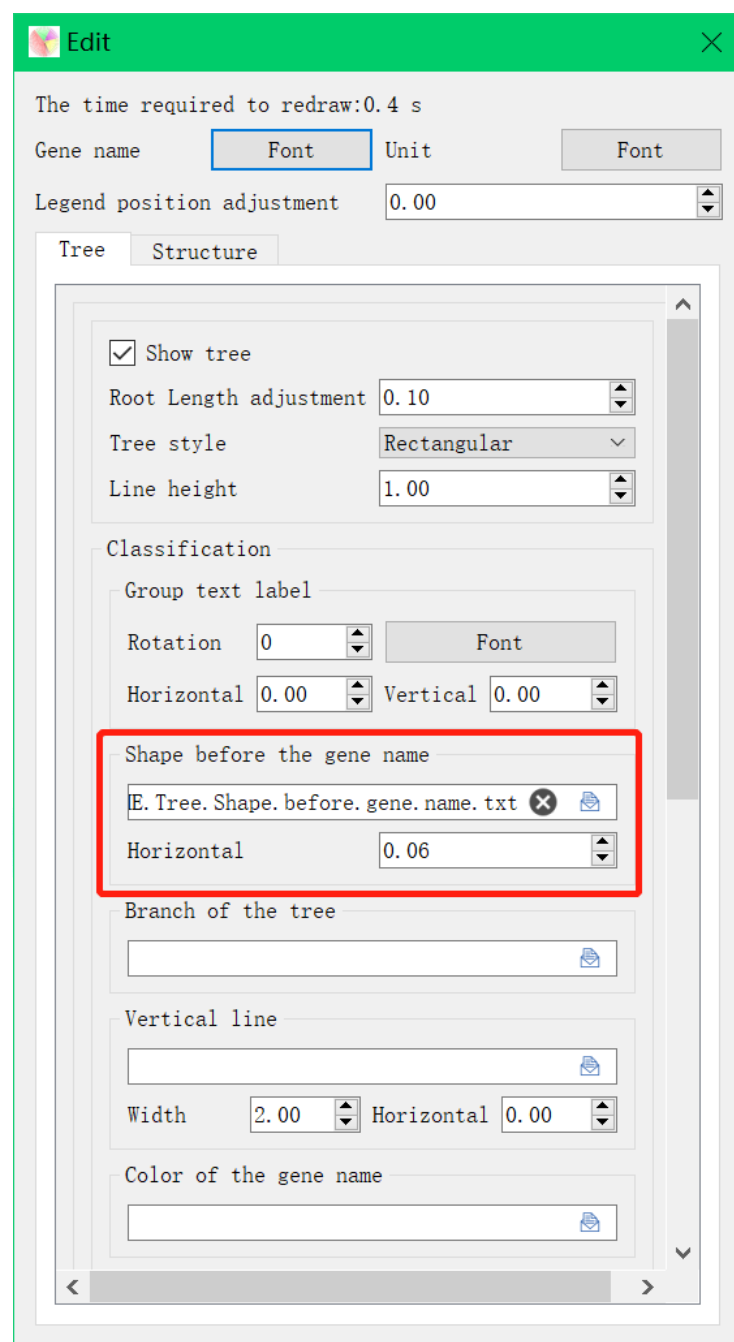


图 43

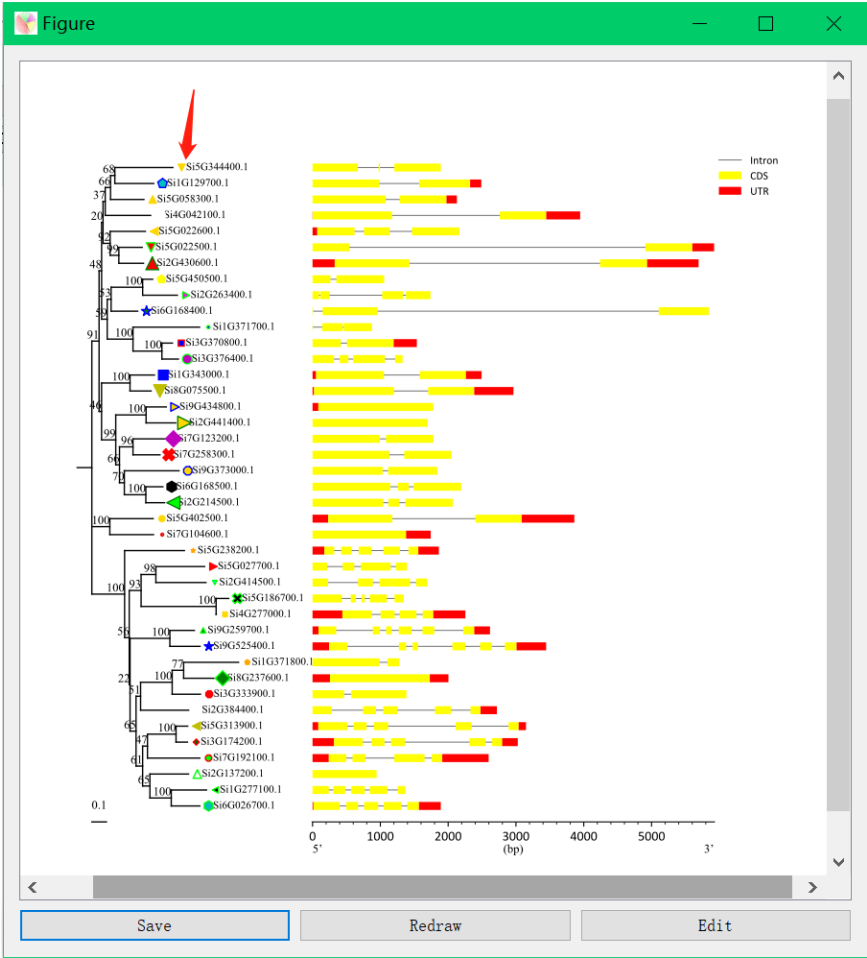


图 44

Visual

树图分类控制文件的格式如下（图 45）：从左到右的列分别表示为：分类区域上端或下端的基因名、分类区域下端或上端对应的基因名、颜色和分类的标签。注意：每一行对应一个分类区域；第四列可以不提供。

Si5G344400.1	Si4G042100.1	r	Group 1
Si5G022600.1	Si2G430600.1	blue	Group 2
Si5G450500.1	Si6G168400.1	#FFA500	
Si1G371700.1	Si3G376400.1	green	Group 4
Si1G343000.1	Si7G258300.1	red	Group 5
Si9G373000.1	Si2G214500.1	#00FF05	Group 6
Si5G402500.1	Si7G104600.1	b	Group 7
Si5G238200.1	Si7G192100.1	yellow	Group 8
Si2G137200.1	Si6G026700.1	#FFD700	Group 9

图 45

支持四种分类形式（如图 46），分别为树枝颜色、分类竖线颜色、基因名字颜色和分类区域颜色。用户可以把准备好的树图分类控制文件拖进用户喜欢的任意一种或多种分类形式所对应的文本框中（图 47）。设置好分类参数之后，点击 Redraw 按钮，结果如图 48。值得一提的是，充分利用分类区域颜色界面中的 Direction 和 Inversion 的设置，可以调节区域渐变的方向，例如，默认的 Direction 设置为 1，Inversion 不选中，区域渐变方向为从右到左；Inversion 选中，区域渐变方向为相反的从左到右。Direction 设置为 0，Inversion 不选中，区域渐变方向为从下到上；Inversion 选中，区域渐变方向为相反的从上到下。另外，把 Gradient 控件取消选中，区域为完全填充，不显示渐变的效果。

Branch of the tree

Vertical line

Width2.00Horizontal0.00

Color of the gene name

Rectangle region

☒ GradientAlpha0.80

Direction1.00☐ Inversion

Boundary

Left0.00Right0.00

图 46

Branch of the tree

le/Si. PME. Tree. Classification. txt

Vertical line

le/Si. PME. Tree. Classification. txt

Width2.00Horizontal0.00

Color of the gene name

le/Si. PME. Tree. Classification. txt

Rectangle region

le/Si. PME. Tree. Classification. txt

☒ GradientAlpha0.40

Direction1.00Inversion

Boundary

Left0.00Right0.00

图 47

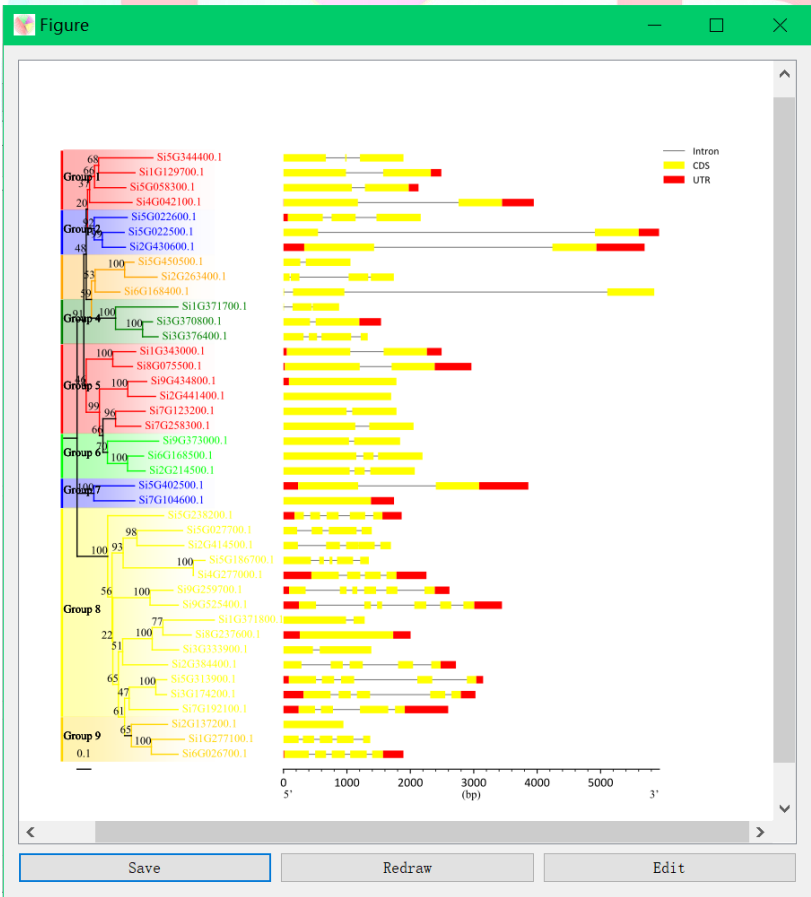


图 48

总而言之，只要用户充分利用 CFVisual 的树图分类功能，真的可以不用再用 Ps 或 Ai 等修图工具去修饰了。然后，作者展示两张样图以供参考（图 48 和图 49）。

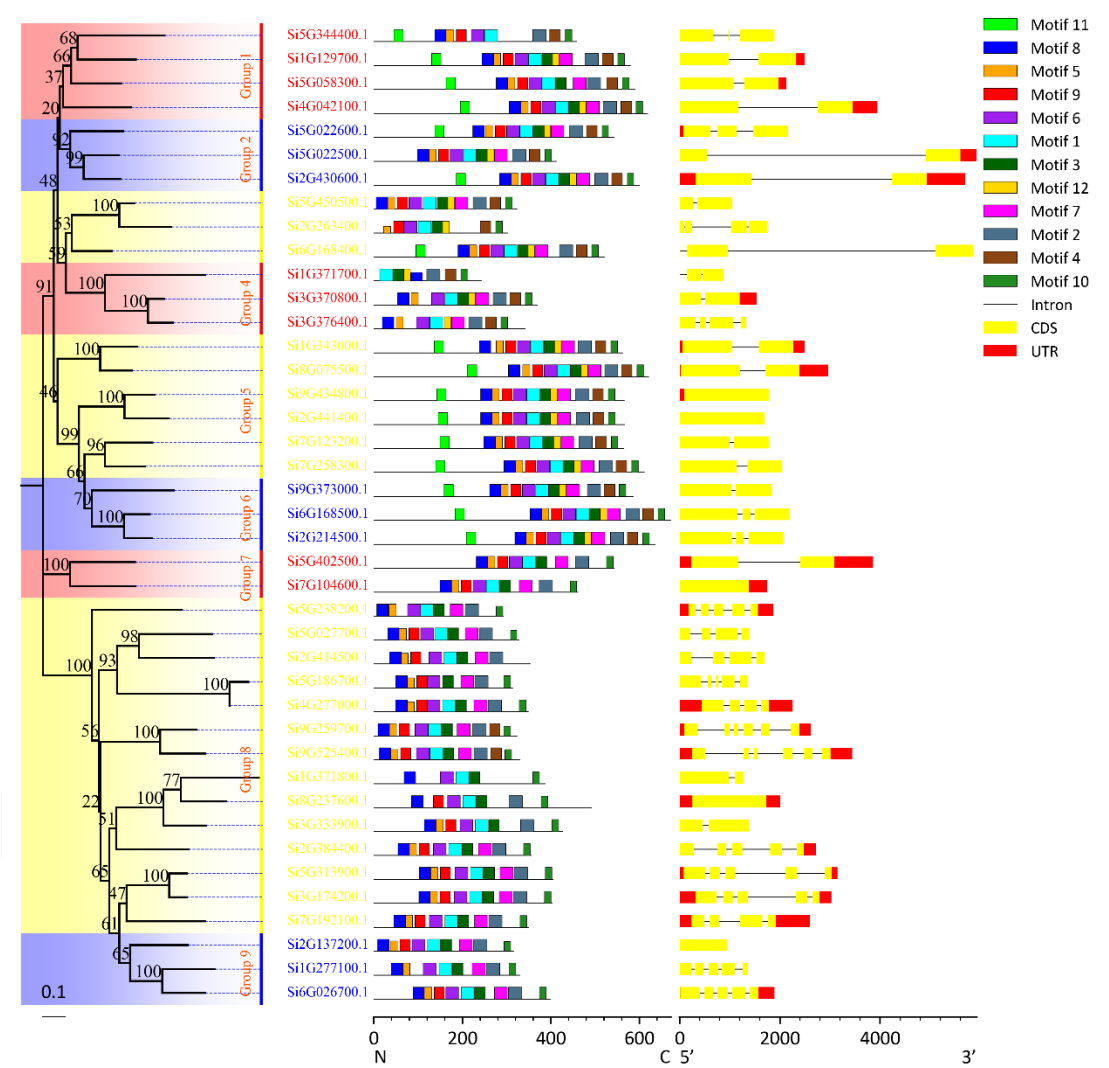


图 49

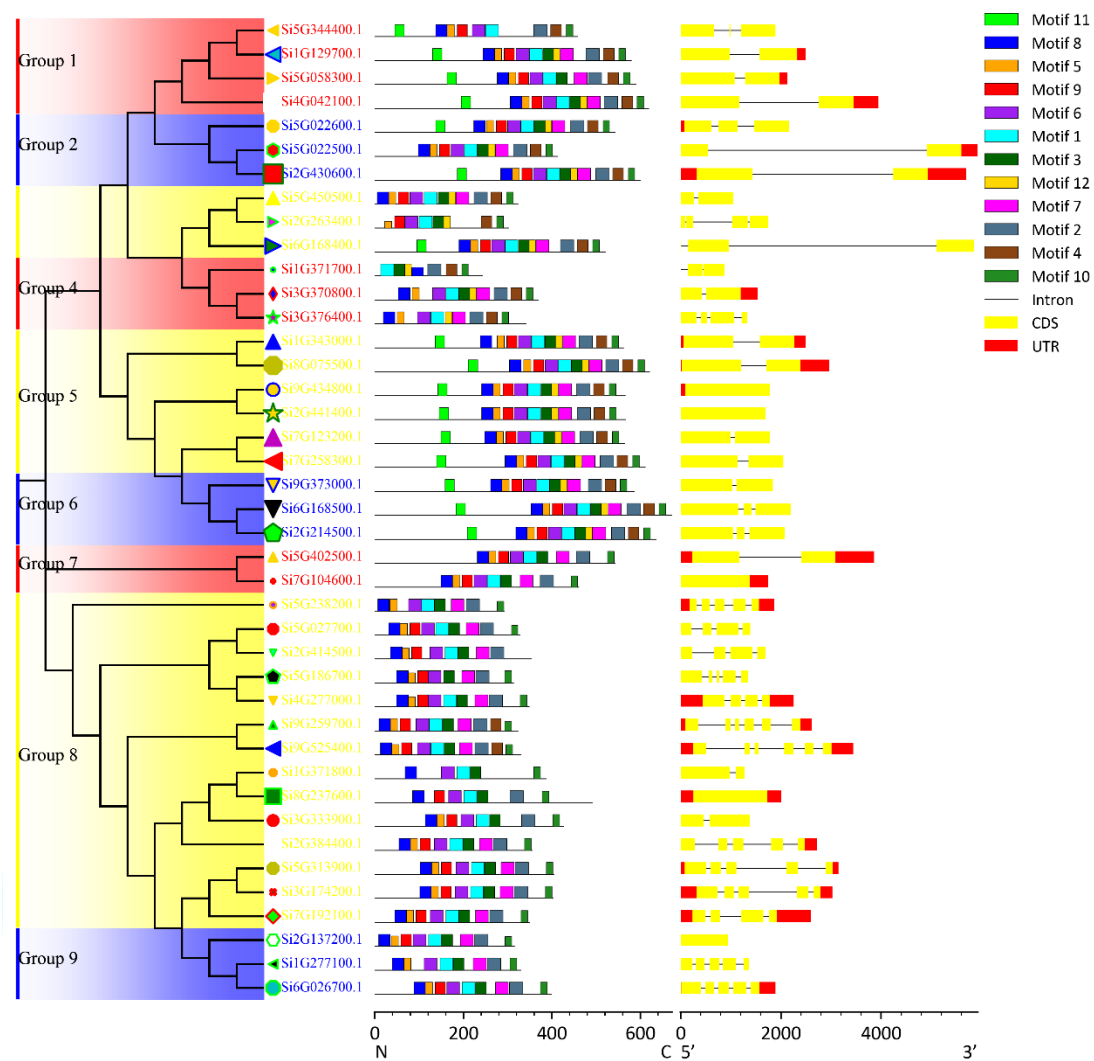


图 50

Visual

3.3.1 历史记录（File-History）

作者在新的版本中添加了历史记录功能，来增加用户的使用体验，历史记录功能可以保存用户上一次绘制成功时，所用到的绘图文件，用户选择后点击“Click and fill it”按钮后，软件将会自动填充上一次用户所用到的文件路径（包括一些额外的特征的文件），用户只需要确认文件类型无误后，直接点击 Start 按钮即可运行，以节省用户的时间。

软件最多保存您 20 条的使用记录。点击 Clear 按钮可以清空您的日志文件。

另外软件的工作原理是将您的运行记录储存在文件中，但是由于作者的开发水平有限，在您卸载软件时，建立的日志储存文件不会被删除，需要您手动删除。

3.3.2 反馈功能（Setting-Feedback）

作者认为本软件仍然是不够完善的有很多值得改进的地方，所以作者开发了 Feedback 模块，当用户在使用过程遇到任何无法理解的情况或者有关于完善软件的想法都可以都通过反馈模块来与作者取得及时的联系。作者在这里欢迎广大用户反映您所遇到的问题、提出您宝贵的意见。

在反馈的过程中作者提供了上传文件选项，来帮助作者继续完善该软件，默认情况是不会上传您所使用的文件的，但是当您认为，上传文件可以帮助作者改进软件，并且您愿意与我们分享交流时，可以上传您所使用的文件，**作者保证您的文件仅作为软件调试的材料并严格保护您的数据隐私，且不会有任何其他目的的使用。**

这里说明一下反馈模块的使用方法，当您不发送文件时只需要填写完全您的姓名，联系方式，遇到的问题或者建议即可。为了更方便的沟通，您不能遗落上面提到的任何一项数据，如果一项数据为空，软件将会给您提示，而不执行发送命令。为了服务器的安全作者限制了反馈的次数，**您每次启动软件最多可以发送三次反馈。**

当您选择发送文件时，除上面的要求外，您只需将您的文件拖动到开始页面的文件路径读取框内即可，软件在发送文件的过程将会根据您的文件所在的位置和您选择的地址读取框的类型而上传您的文件，比如当您的树图文件绘制出错，只需将您的树图文件拖动到树图文件的地址读取框内即可。

同样的为了服务器的安全，**软件不能上传单个文件大小超过 10Mb 的文件。**如果您有这样需求请尝试在软件官网（GitHub）上联系作者。

3.3.3 检查更新功能（Setting-Update）

为了满足更多用户的使用需求，作者开发了在线更新功能，该功能既方便了软件不断的更新与进步，也方便了用户，提升了用户的使用体验。当软件有小更新时，用户不再需要再次忍受 GitHub 缓慢网速的困扰，用户只需要做的就是在网络连接的情况下检查是否有新版本软件，确定更新后，稍等片刻即可。

4. 后序

开发 CFVisual 心路历程，作者以前从来没想到过写软件，因为基因结构图和 Motif 图分别可以通过 GSDS 和 MEME 绘制并通过 Ai 或 Ps 拼接修图，虽然过程繁琐，但是效果满意。是研二之初，要做启动子分析，发现市面上并没有一款软件可以绘制启动子图。因此，无奈只能寻找已有的工具，作者小恋一下，通过 EXCEL 成功实现了启动子图的绘制，作者很高兴，但是绘制过程太麻烦，太繁琐，耗费自己太多时间，又不容易传授给师弟师妹。因此下定决心，自己去写一个绘制启动子图的程序。当写出来了第一个启动子图，慢慢想着自己不能去实现其他做基因家族的图，经过努力和做家族的经验，绘图原理都已破解，需要的就是时间，于是乎，不到一年，作者成功将基因家族常见的图都自己用 Python 编程实现，比如 Circos，染色体定位图，聚类热图，共线性图等等，所以作者现在做基因家族已经不再需要使用其他的绘图的软件了。这里，作者想说，经历三年，作者已然成为了一个基因家族专家（专家们心里想着，可笑的菜逼），对，low 到爆的基因家族分析专家，熟练几乎基因家族可见分析中一切的分析 and 软件的安装使用。

基因家族分析在很多大牛眼里就是一个灌水的 low 到爆的分析，但是很多时候，当人们需要的时候，艾玛，真香，几乎 0 成本，适合培养小白，方便毕业，还能发 SCI，做的好的都能发 BMC Genomic 等这样的二区顶刊。所以，这样 low 到爆的分析依旧每年还是有很多专家去发表，毕竟是香饽饽。作者从来没想到会去读一个研究生，但是大四那年，因为一些原因，还是选择了考母校的硕士，也就是朋友戏称的，辉龙，你在这续命三年呢！作者读硕士三年，很多时候都是做的 low 到爆的基因家族分析，毕竟作者的读研门槛就在那，做不了太多牛气的研究。由于作者也是一名魔术爱好者，喜欢破解一些有趣事物的原理，本科的时候是学校小有名气的 magician，以偶像刘谦的魔术为目标，尝试破解他所有的魔术原理，并加以勤学苦练，对魔术的热爱使作者遇到了一生的魔术导师，并加入了他的门下。后来，作者成功拿到了唐山魔术协会颁发的青年魔术师授牌，成为一个正式的半职业魔术师。）因此，既然作者选择暂时放下热爱的魔术，拿起笔做回初中时期的学霸，那我就得认真对待每一件事物，本着干一行，爱一行的原则，作为小导师门下的大师兄，作者凭借自己的专研劲，自个成功自学几乎所有的基因家族分析套路和流程，但毕竟是被视作 low 的东西，终究上不了门面，但作者阳光的性格和曾经的经历使得作者是永远不会堕落和放弃的，并一如既往的保持着自信，作者就是一个不可多得的人才，只是机遇未到。因此，作者依旧认真的对待每一个基因家族的分析，既然会流程了，就继续走下去专研，潜心研究，终于，好事多磨，

对基因家族的每种分析的可视化原理似乎都破解了。

当作者写出来第一个启动子图的代码之后，有了当初破解魔术的喜悦感和自我满足，实现了很多自己想实现的基因家族可视化的图之后，回顾使用的那些生信软件，作者又尝试着去学习写第一个软件。因此，所有的一切，源于作者对魔术的热爱，对有趣事物的好奇，源于对自学能力的经验自信。作者还是顺利很短时间就自学会了如何写 GUI 界面。成功写出了第一个 GUI 程序——启动子图的可视化界面。课题组的师弟师妹用着虽然不顺手，但是似乎很有用，接着……

作者实现了结构方面的可视化界面（基因结构、motif、结构域和启动子图等），本质上绘图原理一致，因此，需要的只是时间。有实现想法，无奈，读研期间的空闲时间并不多，因此作者只能利用法定假日啊等别人去玩耍的时间去实践自己想展现的 idea。期间，多亏小导师葛维娜老师，引进了一个也特别能专研的大一小师弟尚千。终于，有了第一个不成熟的 GUI 软件 CFVisual_V1.1,并申请了软件著作权，嗯，一件有趣的事情，自己曾经的想法，有了初步的成果变现。满足，因为零计算机基础的作者凭借着自学的本领，拿到了软件著作权。

随着时间的推移，作者已经毕业，终于结束了那段努力也得不到自己追求的奖赏的压抑科研生活，毕竟这三年是自己的人生赌注和选择，早已料到，但勇往之前，一直坚持，所以无悔。终于，作者已经长大，成为一个真正的成熟的成年人，也选择了自己追求的生活理念，做自己想做的，要活得开心，要对世界有点用处。作者成为了一所高校的教师。

利用毕业到现在的自我做主的空余时间，作者又实现了一直没有时间去研究的树图的绘制和交互，终于，努力是不会白费的，作者实现了，并完成了最初想实现的真正的多图合一和交互的功能，真正的做到展示 low 到爆的家族分析中的树图+基因结构+motif+结构域等结构可视化的多图合一过程中不再需要繁琐的拼图和修图了。

至此，作者今日推出的 CFVisual_V2.1 也算是对这个领域做出了实质性的贡献吧。希望这个 low 到爆的软件可以成为人们做基因家族分析中结构展示的首选软件。最后，毕竟这所有的一切都是自学的，没有专业人士指导，作者是业余的，做这个软件呢，写着玩，仅仅是为了向曾经的自己证明，作者还是那个想做一件事，依旧可以努力往极致去发展，依旧可以满足曾经给自己定下的自我合理要求。作者不是计算机专业出身，知道自己菜，比不了那些真正的生信大佬，但就是这么不惧，希望聪明绝顶的高智商大佬们多骂骂我，让我成长的更迅速一些，谢谢！

陈辉龙

2021 年 8 月 8 日