

MapDraw, 在 Excel 中绘制遗传连锁图的宏

刘仁虎, 孟金陵

(华中农业大学作物遗传改良国家重点实验室, 武汉 430070)

摘要: MAPMAKER 是现今广泛使用的遗传连锁数据分析软件, 然而其广泛使用的 DOS 版本却不具有连锁图绘制功能, 给连锁作图工作带来了相当大的麻烦。为了解决这一问题, 我们以大家广泛使用的数据处理软件 Microsoft Excel 为平台, 编写了一个 Excel 宏——MapDraw 来在轻松的操作中实现遗传连锁图的绘制。

关键词: 遗传连锁图; Excel 宏

中图分类号: Q343.2⁺2

文献标识码: A

文章编号: 0253-9772(2003)03-0317-05

MapDraw: A Microsoft Excel Macro for Drawing Genetic Linkage Maps Based on Given Genetic Linkage Data

LIU Ren-Hu, MENG Jin-Ling

(National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: MAPMAKER is one of the most widely used computer software package for constructing genetic linkage maps. However, the PC version, MAPMAKER 3.0 for PC, could not draw the genetic linkage maps that its Macintosh version, MAPMAKER 3.0 for Macintosh, was able to do. Especially in recent years, Macintosh computer is much less popular than PC. Most of the geneticists use PC to analyze their genetic linkage data. So a new computer software to draw the same genetic linkage maps on PC as the MAPMAKER for Macintosh to do on Macintosh has been crying for. Microsoft Excel, one component of Microsoft Office package, is one of the most popular software in laboratory data processing. Microsoft Visual Basic for Applications (VBA) is one of the most powerful functions of Microsoft Excel. Using this program language, we can take creative control of Excel, including genetic linkage map construction, automatic data processing and more. In this paper, a Microsoft Excel macro called MapDraw is constructed to draw genetic linkage maps on PC computer based on given genetic linkage data. Use this software, you can freely construct beautiful genetic linkage map in Excel and freely edit and copy it to Word or other application. This software is just an Excel format file. You can freely copy it from ftp://211.69.140.177 or ftp://brassica.hzau.edu.cn and the source code can be found in Excel's Visual Basic Editor.

Key words: genetic linkage map; Microsoft Excel macro

随着分子生物学的飞速发展, 分子标记遗传连锁作图方面的工作做得越来越多。要研究一个性状的遗传规律, 要采用图位克隆技术克隆一个基因, 其首要的工作就离不开遗传连锁作图了。现在, 已有各种各样的软件用于标记间遗传连锁距离的计算, 然而, 大多数软件却不能满足人们灵活绘制遗传连锁图的需要。MAPMAKER^[1] 的苹果机版本 MAPMAKER 3.0 for Macintosh 是一款优秀的遗传距离计

算和连锁图绘制合二为一的软件。其绘出的连锁图美观大方, 渐渐被科学家认为是连锁图绘制的一种标准样式。然而, 大家广泛使用的 DOS 版本 MAPMAKER 3.0 for PC 却只能得到连锁关系数据, 不能绘出美观适用的连锁图, 不无给广大 PC 机用户留下了无限的遗憾! 为此, 有些软件开发者也专门开发了一些软件^[2] 来专门绘制遗传连锁图, 以弥补这些软件在绘图方面的不足, 然而该软件绘出的遗传连锁图

收稿日期: 2002-07-10; 修回日期: 2002-09-28

基金项目: 国家转基因植物研究与产业化专项(J00-A-008-01); 高等学校博士学科点专项基金(960403)资助

作者简介: 刘仁虎(1975-), 男(汉), 2000级博士生, 籍贯: 湖北襄樊, 专业方向: 发育生物学。E-Mail: liurenhu@21cn.com

在外观和适用程度上还有待提高。

在 PC 机广泛流行的今天, Microsoft Office 办公套件已成为广大科学工作者最为熟悉的应用软件之一, 而且其内嵌的宏功能支持 Visual Basic 编程语言, 它使得该办公套件的功能更加强大。特别是其中的 Excel 组件就是专门为数据处理而设计的, 已在各学科的数据分析领域获得广泛应用。既然 Excel 这么普及, 功能又这么强大, 如果能用它来完成遗传连锁图的绘制, 使用起来不是非常方便吗? 经过仔细阅读 Excel 宏功能的在线帮助后, 我们对 Excel 进行了二次开发, 编写了这个绘制遗传连锁图的宏 MapDraw。将在其他软件中得到的连锁数据输入 Excel 后, 执行该宏就立即能得到像 MAPMAKER 3.0 for Macintosh 一样绘制的遗传连锁图(图

1)。该软件及其源代码可通过 ftp://211.69.140.177 或 ftp://brassica.hzau.edu.cn 免费下载, 也可与作者联系索要。

1 MapDraw 的原理和特点

MapDraw 是依据现有的遗传连锁数据, 如 MAPMAKER for PC 或其他连锁分析软件得到的数据, 利用 Microsoft Excel 内嵌 Visual Basic 语言开发出来的具有绘制遗传连锁图功能的 Excel 宏。它继承了宏程序固有的优点: 兼容性和可移植性好, 只依赖于应用系统而与操作系统无关。它使用简便灵活, 绘出的连锁图美观整洁, 而且在 Windows 各应用程序中兼容性也很好, 可轻松地对绘出的图形进行各种编辑和修改。

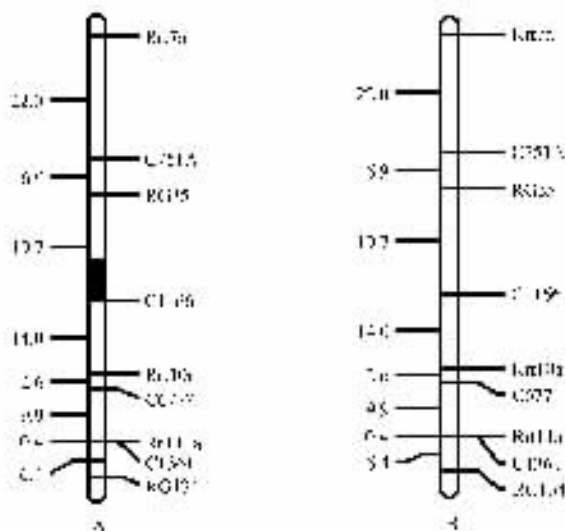


图 1 用 MAPMAKER 3.0 for Macintosh 在苹果机上绘制的水稻第 10 号染色体^[3]连锁图(A)与用 MapDraw 在 Excel 中绘制的同一遗传连锁图(B)的相似性比较

Fig. 1 MapDraw got the same genetic linkage map (B) as MAPMAKER 3.0 for Macintosh did on Macintosh (A)^[3]

2 MapDraw 的使用方法

2.1 数据准备

用常用遗传作图软件(如 MAPMAKER 3.0 for PC)分析得到各标记的遗传连锁数据。

2.2 输入连锁数据

打开具有 MapDraw 宏的 Excel 文件, 添加一张空的电子表格, 在电子表格的第一列输入遗传距离, 第一格输 0 或留为空白, 第二列输入标记的名称, 使第一列的数据为第二列相邻两标记的遗传距离, 若有多个连锁群请相继在后面输入(图 2a)。

2.3 运行宏

按 Alt+F8 快捷键或者从 Microsoft Excel 菜单中顺序

选择: 工具|宏|宏(图 2b)就可打开宏的列表(图 2c), 执行 Mapdraw 宏即可得到美观的遗传连锁图(图 2d)。若还想对得到的连锁图进行适当的编辑和修饰, 例如将其中的某个标记名变为黑体或红色以突出显示, 你可以将该图取消组合, 将某部分修改后再重新组合起来。

2.4 将得到的连锁图拷到 Microsoft Word 或其他系统中排版

运行 MapDraw 得到的遗传连锁图可以像 Microsoft Excel 中其他图形对象一样使用“拷贝”“粘贴”的方法与其他应用程序共享。但要将其“粘贴”到 Word 文件中, 最好使用“选择性粘贴”命令, 按照“增强型图元文件”的格式进行粘贴, 这样可让该图形免受 Word 文档模板格式的影响, 又不影响图片质量。

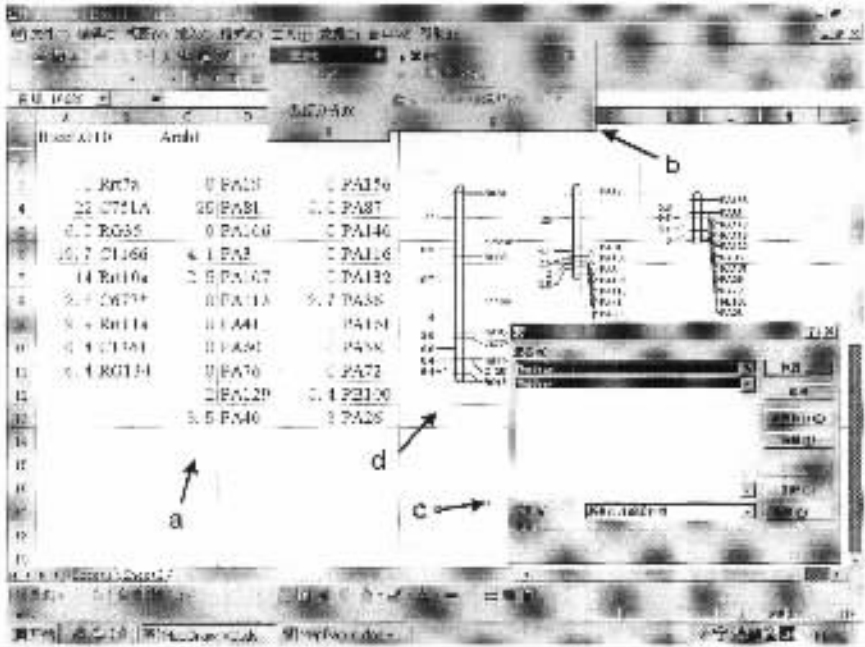


图 2 利用宏 MapDraw 在 Excel 中绘遗传连锁图的示意图

a: 数据输入格式; b, c: 运行宏; d: 连锁图输出。

Fig. 2 Procedure of linkage map construction using MapDraw

a: genetic linkage data input; b c: find your Macro and run it; d: linkage map output.

3 MapDraw 的源代码

程序源代码比较简单, 不方便上网的读者将这些代码输入到 Excel 的“Visual Basic 编辑器”中就可立即运行(以'号开头的行为程序注释, 可不输入)。

```

'*****程序开始*****
Option Explicit

Sub MapDraw()
'本程序用于遗传连锁图的绘制!
'Renhu Liu 记录的宏 1999-11-21
'***** 定义变量 variation definition *****
    Dim Row As Integer, Col As Integer, TempName1 As String,
    TempName2 As String
    Dim LineD As Double, SumDist As Double, FontS As Integer
    Dim i As Integer, j As Integer, CurMar As String, CurRange As
    Range, a As Integer
    Dim Dist As Double, TextP1 As Double, TextP2 As Double
    Dim cM As Double, StartPx As Integer, StartPy As Integer,
    Wide As Double, Head As Double, LineW As Double, dec As
    Integer
'***** 参数初始化 Parameters below, you can
change the numbers *****
    cM = 3; StartPx = 200; StartPy = 20; Wide = 7; Head = 7;
    LineW = 1.5; dec = 1
'Sometime, you have to change the value of "cM" larger, 5 for exam-
ple, if your average linkage distance less than 8 cM.
'Please change the value of Dec to 2 when your linkage data have two
decimals.
'各常量分别表示: 每 cM 长度(P), 1P=0.3517mm; 连锁图位置的横
坐标; 纵坐标; 连锁图顶端的长;
'连锁图线条的粗细; 遗传距离小数点位数

```

```

a = 2
FontS = 8
'上两个变量分别表示相邻两标记在图上的相对位置和标记字体的
大小
'***** 参数初始化完毕 *****
ActiveCell.CurrentRegion.Select
Set CurRange = Selection
Row = CurRange.Rows.Count
Col = CurRange.Columns.Count

If Row < 2 Or Col < 2 Then
    MsgBox "Data invalid or not selected!"
    GoTo End3
End If

TempName1 = "ATTENTION; This macro is used to linkage
map drawing! Some of the following parameters may be important
for you." & Chr(10) & "cM: Chromosome length (Points) per centi-
Mogen; Dec: Decimal digits; Font: Font size; Wide: Wide of your
Chromosome; Head: Head length of your chromosome; LineW: Points
of your chromosome line" & Chr(10) & Chr(10) & "Please contact
with liurenhu@21cn.com for technique support" & Chr(10) & Chr
(10)

Again1:
    CurMar = InputBox(TempName1 & "Please Input Paramet-
ers; cM, Dec, Font, Wide, head, LineW", "Parameters Input Box",
    cM & "," & dec & "," & FontS & "," & Wide & "," & Head &
    "," & LineW)
    If CurMar = "" Then GoTo End3
    cM = Val(CurMar); CurMar = Mid(CurMar, InStr(CurMar,
    ",")) + 1)
    dec = Val(CurMar); CurMar = Mid(CurMar, InStr(CurMar,
    ",")) + 1)
    FontS = Val(CurMar); CurMar = Mid(CurMar, InStr(Cur-
    Mar, ",") + 1)

```

```

Wide = Val(CurMar): CurMar = Mid(CurMar, InStr(CurMar, ",") + 1)
Head = Val(CurMar): CurMar = Mid(CurMar, InStr(CurMar, ",") + 1)
LineW = Val(CurMar): LineD = (FontS + a) / cM
If cM * FontS * Wide * (Head + 1) * (dec + 1) * LineW <= 0 Then
    MsgBox "Parameter error! Please input them again!"
GoTo Again1
End If
j = 1
TempName1 = ActiveSheet.Name
Worksheets.Add
TempName2 = ActiveSheet.Name
Do While j < Col
    Dist = WorksheetFunction.Sum(CurRange.Columns(j))
    If CurRange.Cells(1,j).Value <> 0 Then
        MsgBox ("Your data have no 0 in the first row! Please add it on and try again!")
        CurRange.Cells(1,j).Select
        GoTo End1
    End If
    ActiveSheet.Shapes.AddShape(msoShapeRoundedRectangle, StartPx, StartPy, Wide, Head * 2 + Dist * cM).Select
    Selection.ShapeRange.Adjustments.Item(1) = 0.5
    Selection.ShapeRange.Line.Weight = LineW
    ActiveSheet.Shapes.AddLine(StartPx - Wide/4 - LineW/2, StartPy + Head, StartPx - 0.4 * Wide, StartPy + Head).Select
    TextP2 = -99: TextP1 = -99
    Dist = 0
    CurMar = CurRange.Cells(1, j + 1).Text
    ActiveSheet.Shapes.AddLine(StartPx - 0.4 * Wide, StartPy + Head, StartPx + 3.4 * Wide, StartPy + Head).Select
    With ActiveSheet.Shapes.AddTextbox(msoTextOrientationHorizontal, StartPx + 3.4 * Wide, StartPy + Head - FontS / 2 - a / 2, 72, FontS + a).TextFrame
        .Characters.Text = CurMar
        .Characters.Font.Name = "Arial"
        .Characters.Font.Size = FontS
        .HorizontalAlignment = xlLeft
        .AutoSize = True
        .Parent.Line.Visible = msoFalse
        .Parent.Fill.Transparency = 1
    End With
    TextP2 = Dist
    SumDist = 0
    For i = 2 To Row
        CurMar = CurRange.Cells(i, j + 1).Text
        If Not IsNumeric(CurRange.Cells(i, j)) Then CurRange.Cells(i, j) = 0
        Dist = WorksheetFunction.Fixed(CurRange.Cells(i, j), dec)
        SumDist = SumDist + Dist
        If SumDist - TextP2 >= LineD And CurMar <> Empty Then
            ActiveSheet.Shapes.AddLine(StartPx - 0.4 * Wide, StartPy + SumDist * cM + Head, StartPx + 3.4 * Wide, StartPy + SumDist * cM + Head).Select
            With ActiveSheet.Shapes.AddTextbox(msoTextOrientationHorizontal, StartPx + 3.4 * Wide, StartPy + SumDist * cM + Head - FontS / 2 - a / 2, 72, FontS + a).TextFrame
                .Characters.Text = CurMar
                .Characters.Font.Name = "Arial"

```

```

                .Characters.Font.Size = FontS
                .HorizontalAlignment = xlLeft
                .AutoSize = True
                .Parent.Line.Visible = msoFalse
                .Parent.Fill.Transparency = 1
            End With
            TextP2 = SumDist
        Else
            If SumDist - TextP2 < LineD And CurMar <> Empty Then
                TextP2 = TextP2 + LineD
                ActiveSheet.Shapes.AddLine(StartPx - 0.4 * Wide, StartPy + SumDist * cM + Head, StartPx + 2 * Wide, StartPy + SumDist * cM + Head).Select
                ActiveSheet.Shapes.AddLine(StartPx + 2 * Wide, StartPy + SumDist * cM + Head, StartPx + 3 * Wide, StartPy + TextP2 * cM + Head).Select
                ActiveSheet.Shapes.AddLine(StartPx + 3 * Wide, StartPy + TextP2 * cM + Head, StartPx + 3.4 * Wide, StartPy + TextP2 * cM + Head).Select
                With ActiveSheet.Shapes.AddTextbox(msoTextOrientationHorizontal, StartPx + 3.4 * Wide, StartPy + TextP2 * cM + Head - FontS / 2 - a / 2, 72, FontS + a).TextFrame
                    .Characters.Text = CurMar
                    .Characters.Font.Name = "Arial"
                    .Characters.Font.Size = FontS
                    .HorizontalAlignment = xlLeft
                    .AutoSize = True
                    .Parent.Line.Visible = msoFalse
                    .Parent.Fill.Transparency = 1
                End With
            End If
        End If
    End If
    If Dist <> 0 And Dist <> Empty Then
        If SumDist - Dist / 2 - TextP1 >= LineD Then
            TextP1 = SumDist - Dist / 2
            ActiveSheet.Shapes.AddLine(StartPx - Wide / 4 - LineW / 2, StartPy + (SumDist - Dist / 2) * cM + Head, StartPx - 2.4 * Wide, StartPy + (SumDist - Dist / 2) * cM + Head).Select
            With ActiveSheet.Shapes.AddTextbox(msoTextOrientationHorizontal, StartPx - 2.4 * Wide - (Len("A" & Dist)) * FontS, StartPy + (SumDist - Dist / 2) * cM + Head - FontS / 2 - a / 2, (Len("A" & Dist)) * FontS, FontS + a).TextFrame
                .Characters.Text = Dist
                .Characters.Font.Name = "Arial"
                .Characters.Font.Size = FontS
                .HorizontalAlignment = xlRight
                .AutoSize = False
                .Parent.Line.Visible = msoFalse
                .Parent.Fill.Transparency = 1
            End With
        Else
            TextP1 = TextP1 + LineD
            ActiveSheet.Shapes.AddLine(StartPx - Wide / 4 - LineW / 2, StartPy + (SumDist - Dist / 2) * cM + Head, StartPx - Wide, StartPy + (SumDist - Dist / 2) * cM + Head).Select
            ActiveSheet.Shapes.AddLine(StartPx - Wide, StartPy + (SumDist - Dist / 2) * cM + Head, StartPx - 2 * Wide, StartPy + TextP1 * cM + Head).Select
            ActiveSheet.Shapes.AddLine(StartPx - 2 * Wide, StartPy + TextP1 * cM + Head, StartPx - 2.4 * Wide,

```

```

StartPy + TextP1 * cM + Head). Select
    With ActiveSheet. Shapes. AddTextbox(msoTextOrientationHorizontal, StartPx - 2.4 * Wide - (Len("A" & Dist)) * FontS, StartPy + TextP1 * cM + Head - FontS / 2 - a / 2, (Len("A" & Dist)) * FontS, FontS + a), TextFrame
        . Characters. Text = Dist
        . Characters. Font. Name = "Arial"
        . Characters. Font. Size = FontS
        . HorizontalAlignment = xlRight
        . AutoSize = False
        . Parent. Line. Visible = msoFalse
        . Parent. Fill. Transparency = 1
    End With
End If
End If
Next i
End2:
ActiveSheet. Shapes. SelectAll
Selection. ShapeRange. Group. Cut
Worksheets(TempName1). Paste
j = j + 2
StartPx = StartPx + 150
Loop
End1:
Application. DisplayAlerts = False
Worksheets(TempName2). Delete
Application. DisplayAlerts = True
End3:
End Sub
/****** 程序结束 *****/

```

4 MapDraw 参数的使用

一般而言,不同的连锁群会有不同的密度,采用程序默认的参数有的人得到的连锁图太密,各标记挤成一团,有的人又觉得太稀疏,星星点点几个标记,很不协调。所以本程序设计了 6 个参数来控制连锁图的输出,用于打造自己个性化的遗传连锁图谱。程序运行时会产生一个参数设置窗口(图 3),表 1 列出了各个参数的意义,根据需要修改相应的参数即可实现对连锁图输出的控制。

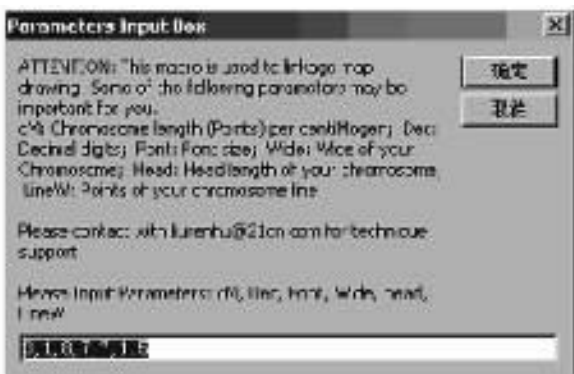


图 3 MapDraw 的参数设置窗口

Fig. 3 Parameters input box of MapDraw

5 讨 论

Microsoft Excel 是大家都很熟悉的数据处理应用软件,利用 Excel 的宏功能设计的应用程序,源代码开放,可移植性好,又是免安装的绿色软件,拷过来就能直接运行,再加上绘出的图形是 Windows 平台下的矢量图形,兼容性和可编辑性都很好,所以受到了很多科学工作者的欢迎。然而,该软件还不具有 QTL 曲线图的绘制功能,许多人都是将 QTL 峰值和区间手工标上连锁图,很不方便,若能将软件 Map-Plotter^[2]的 QTL 曲线绘制功能用 Excel 的 VBA 语言改写一下,使绘出的 QTL 曲线为 Windows 矢量图形,则可能使该软件的功能更完善。

表 1 MapDraw 的参数

Table 1 Using parameters of MapDraw to construct linkage map with your own style

参数名称	参数意义	参数用法
cM	每单位遗传距离(cM)所画的染色体长度*	连锁图作得很密时,可将默认值改大一些,标记比较稀疏时,可将该值改小些,以得到更协调的连锁图
Dec	遗传距离要精确的小数位数	遗传距离要精确到 2 位小数就将该参数的值修改为 2
FontS	连锁图上字体的大小*	想加大字体就加大该参数值,反之,降低该值
Wide	染色体的宽度*	想把染色体绘粗些就加大该值,反之降低该值
Head	两端标记距染色体端部的距离*	想把染色体顶端留长些就加大该值,反之降低该值
LineW	绘染色体所用线的粗细*	绘染色体想用粗线就加大到 2 或 3,想用细线就降低到 1

* 单位为“磅”,1 磅约为 0.35 毫米。

利用 Excel 的宏来进行各种数据分析的各种优点是显而易见的,今后各种遗传学中的数据分析和图形绘制若都能在 Excel 中用其宏功能实现,将大大方便用户使用。

参 考 文 献 (References):

- [1] Lander E S, Green P, Abrahamson J, Barlow A, Daly M J, Lincoln S E, Newburg L. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations [J]. Genomics, 1987, 1, 174~181.
- [2] 沈利爽, 郑先武, 朱立煌. Mapplotter——一个输出遗传图谱、图示基因型和 QTL 曲线图形的软件[J]. 遗传, 2000, 22: 172~174.
- [3] Wang S, Liu N, Peng K, Zhang Q. The distribution and copy number of copia-like retrotransposons in rice (*Oryza sativa* L.) and their implications in the organization and evolution of the rice genome [J]. Proc Natl Acad Sci USA, 1999, 96: 6824~6828.