# 用R程序包plantlist查询和处理植物学名

张金龙

[jinlongzhang01@gmail.com](mailto:jinlongzhang01@gmail.com)

嘉道理农场暨植物园植物保育部

# 1 简介

plantlist是用来查询和处理植物学名的R程序包。类似的R程序包还有taxize、Taxonstand等, 相比之下plantlist的主要优点在于：使用本地自带的数据从而不依赖网络，优化了算法从而速度更快，给出的信息更加简洁明确，同时能够批量查询中文学名。

### plantlist主要功能为：

1. 批量查询植物科属，内置的数据主要来源于theplantlist网站， 其中被子植物采用APGIII分类系统，同时提供维管植物每个科的编号，方便标本馆管理植物标本。
2. 直接生成科/属/种的列表, 以便通过Phylomatic等软件生成进化树；
3. 用中文名批量查询植物学名以及科属等。
4. 查询学名的接受状态以及完整学名

### plantlist包内的函数：

CTPL 通过中文名查询植物的完整学名和科属

CTPL2 通过中文名查询每个种的科、属、分布、海拔、IUCN等级数据

Status 查询每个学名在The Plant List 1.1 接受状态

taxa.table 基于TPL查询结果制作科属种列表， 以便用于Phylomatic软件建树

TPL 提供学名， 查询目、科、属、以及科的编号

### plantlist包内的数据集：

acc\_dat数据集 The Plant List网站上所有接受名

cnplants数据集 中国植物名录及每个种的科、属、分布、海拔、IUCN等级数据

syn\_dat 数据集 The Plant List网站上异名数据库

# 2 软件安装

plantlist是R程序包， 所以必须要先安装R才能使用。由于plantlist内部函数CTPL2函数需要使用openxlsx读取xlsx文件, 所以本程序包在安装时需要安装openxlsx所依赖的Rtools以及Rcpp，并配置好启动路径。本部分介绍如何安装。

## 2.1 安装R软件

R软件下载的地址为: <http://cran.r-project.org/bin/windows/base/>，请按照默认路径安装R软件。因为涉及到UFT8字符转换，R版本不能低于3.0.3。

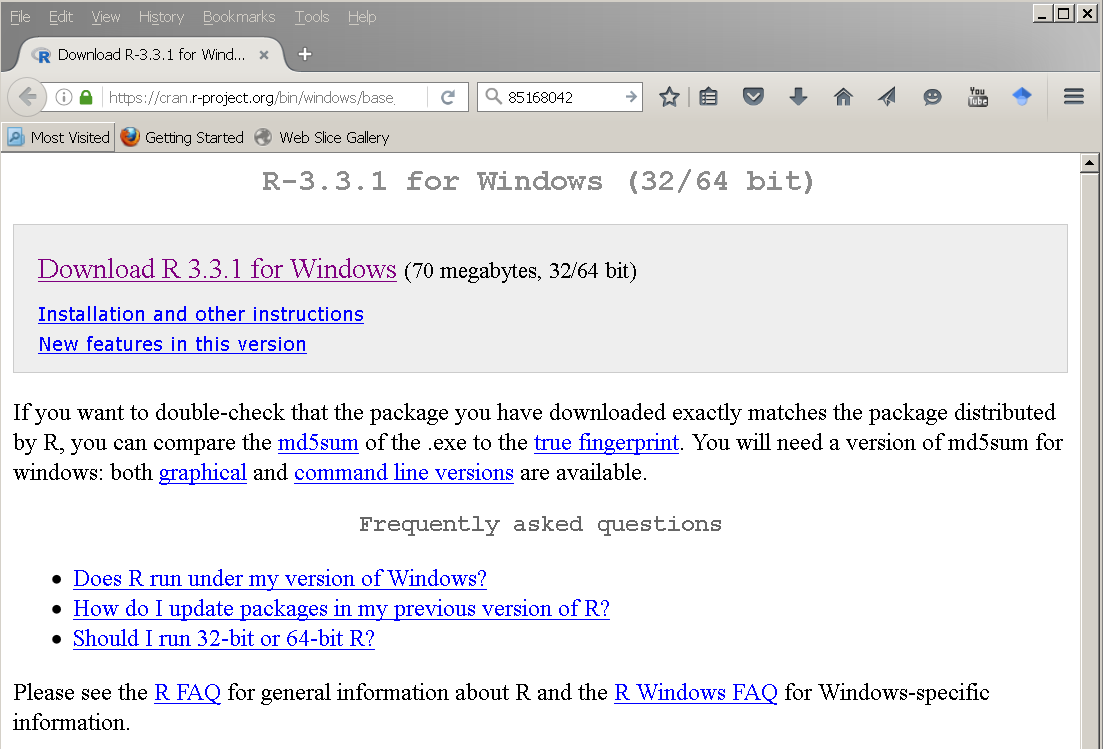


图1. R软件windows版本下载页

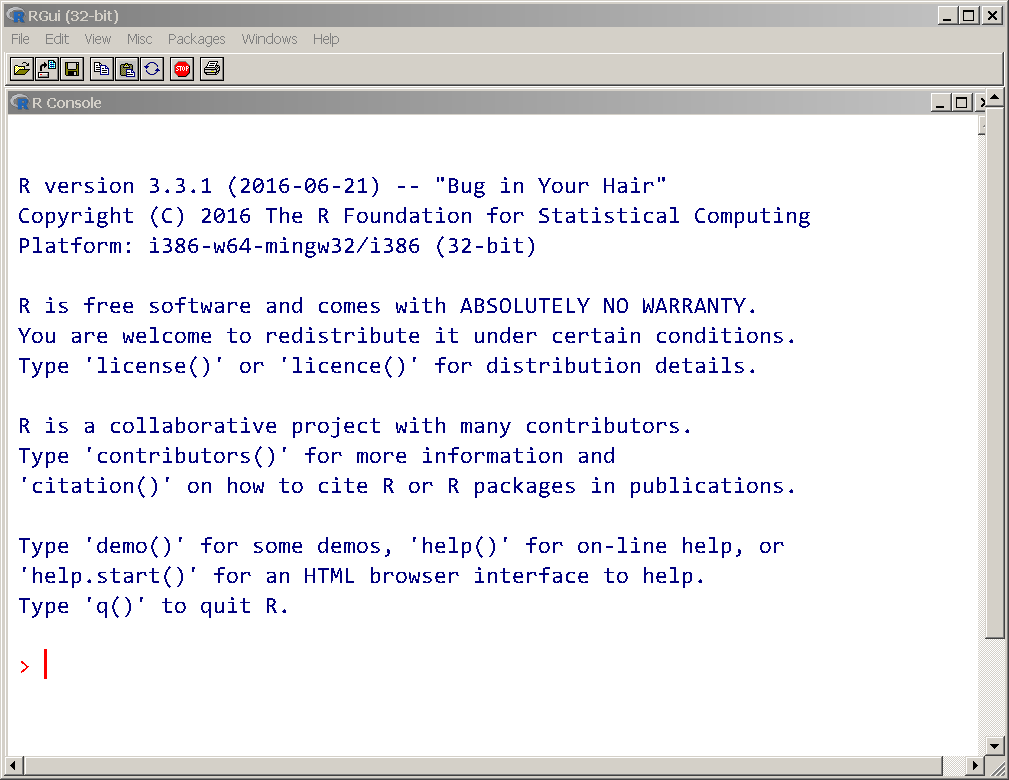


图2. R软件3.3.1的登录界面

## 2.2 安装Rtools

Rtools是编写R程序包的工具软件，含有读写xlsx文件所需的unzip和zip函数，安装时必须允许Rtools修改启动路径，以便从控制台中调用zip等命令读写Excel模板。Rtools的下载地址为<https://cran.r-project.org/bin/windows/Rtools/>。安装Rtools时须允许其修改系统路径system PATH，以便从控制台中可以直接使用Rtools的zip等命令。

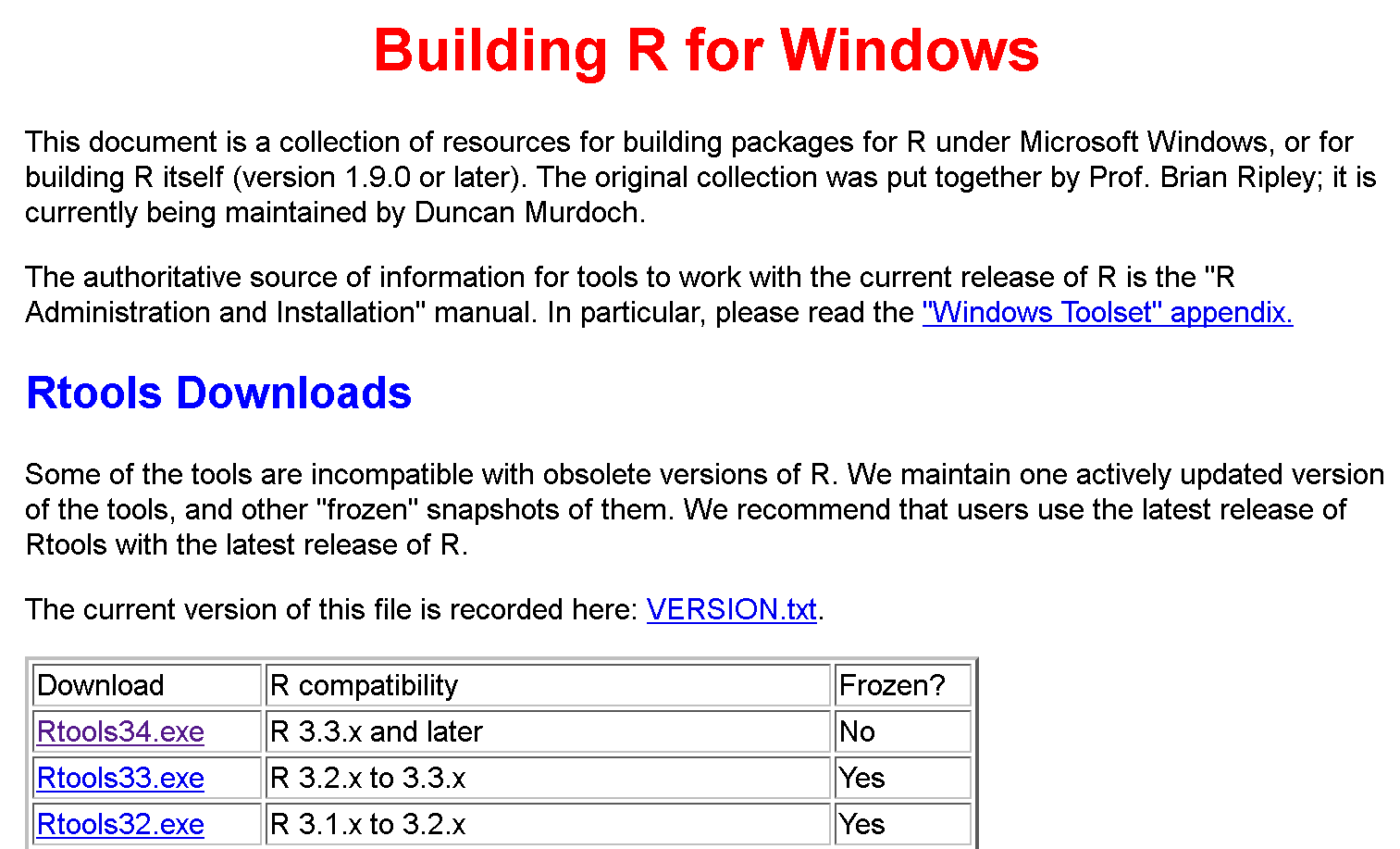


图3. Rtools下载页面，以未冻结的版本为准

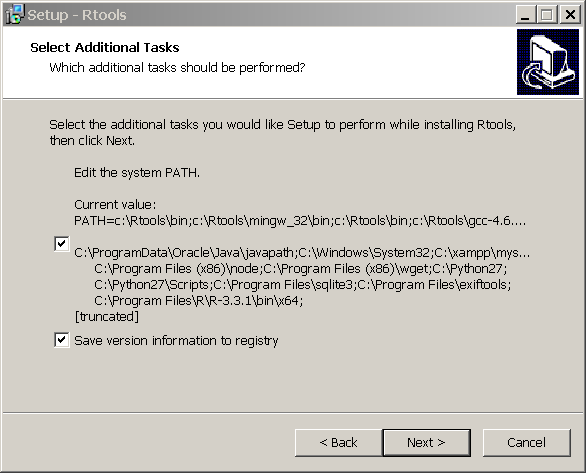


图4. 安装Rtools时应该允许其修改system PATH

## 2.4 检查Rtools是否已正确配置

Rtools安装完成后可以通过以下方式检查是否安装成功:

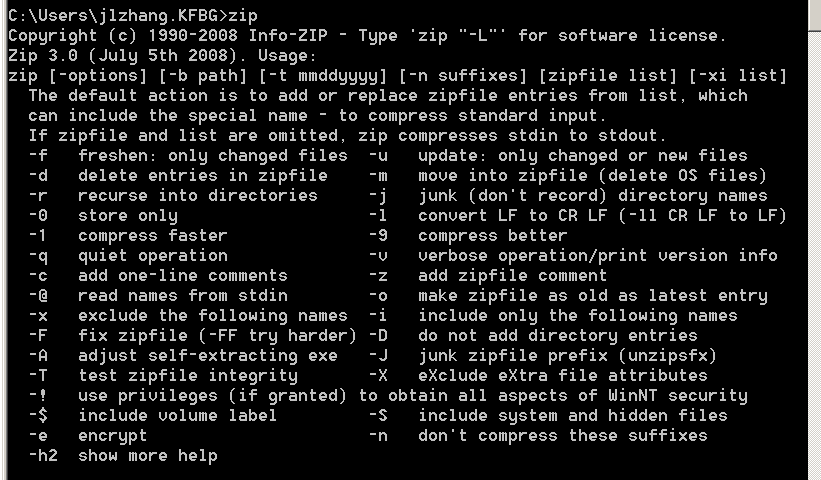


图5. 在控制台中输入zip, 如图所示，若未提示错误，则Rtools配置成功

## 2.5 安装openxlsx程序包

openxlsx是一个R程序包，用来读写标本数据的xlsx模板， 在RGUI的命令行中输入 install.packages("openxlsx")， 选择距离较近的CRAN镜像，openxlsx程序包会自动下载并安装好。

C:\Users\jlzhang.KFBG\Desktop\save\QQ图片20160819122741.tif

图6. 通过install.packages命令安装openxlsx程序包

## 2.6 从Rforge 安装 plantlist

网址是 <https://r-forge.r-project.org/R/?group_id=2052>

命令是 install.packages("plantlist", repos="http://R-Forge.R-project.org")

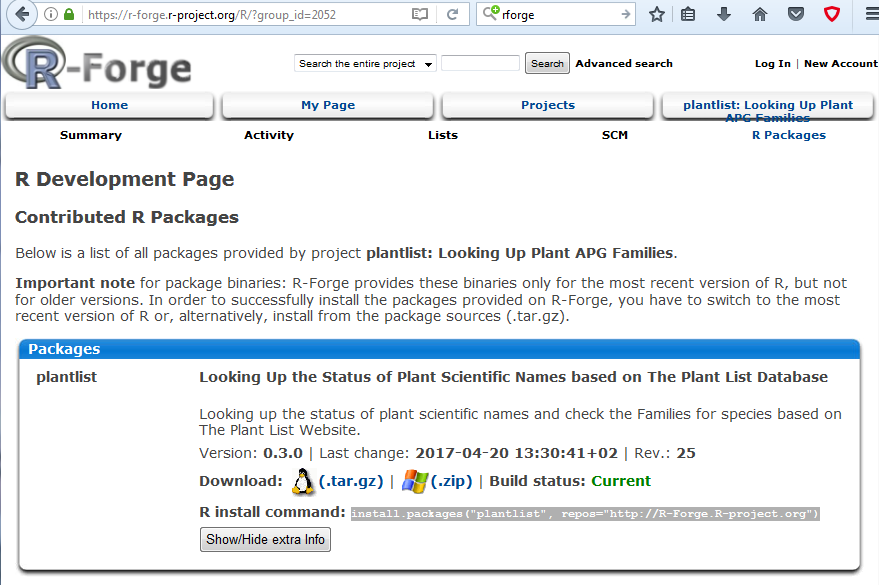


图7. Rforge网站上plantlist的页面

plantlist源代码也保存在github， 网址是 <https://www.github.com/helixcn/plantlist>

至此plantlist及其依赖的程序包已经完成安装。

# 3 plantlist各函数的使用方法

在使用plantlist之前，必须用library(plantlist) 加载plantlist。

每个函数的使用方法， 请输入"?函数名"， 例如： ?TPL 查询TPL函数的用法。

## 3.1 CTPL 通过中文名查询植物的完整学名和科属

（该函数现已不再维护，用户应该考虑尽量使用CTPL2）

## 3.2 CTPL2 通过中文名查询每个种的科、属、分布、海拔、IUCN等级数据

例如 Excel文件保存在 D:/data/ 文件夹下，文件名为 guangdong\_plants.xlsx

用法

library(plantlist)

setwd("D:/data/")

CTPL2(infile = "guangdong\_plants.xlsx", outfile = "results.xlsx")

注意， 输入的中文名不能有空格， 否则将不能正确匹配。

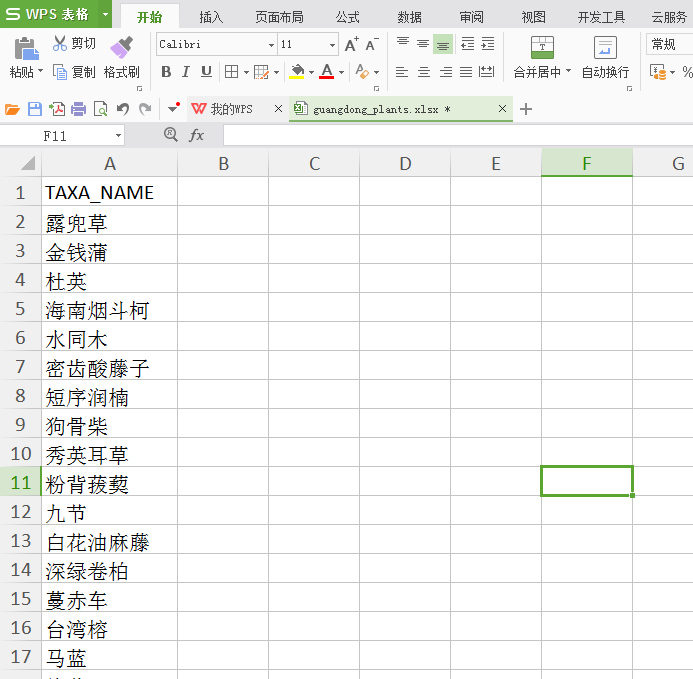
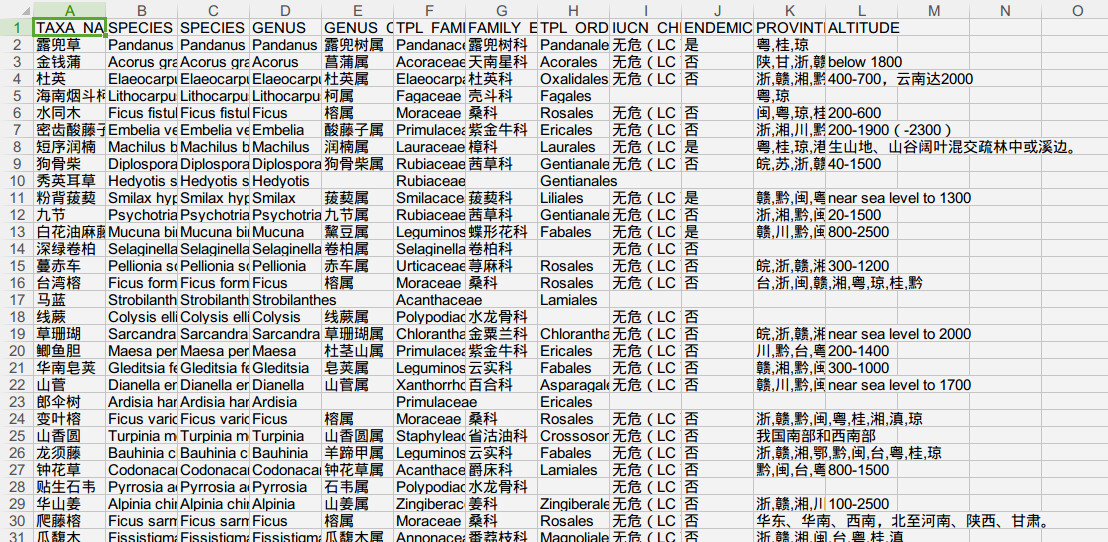


图8. guangdong\_plants.xlsx文件的内容



## 图9. CTPL2查询结果， 该文件保存在D:/data/ 文件夹下。

## 3.3 status 查询每个学名在The Plant List 1.1 接受状态

函数的参数 status(species =NA, exact =TRUE, spell\_error\_max =NULL)

species为输入的字符串向量。

exact 表示是否进行精确匹配，如果不是精确匹配，则所有能用grep正则表达式匹配的结果都会显示。一般建议用精确匹配。

spell\_error\_max 为所允许的最大的错误拼写的字母数量。

status函数对输入的species 物种名大小写不敏感,物种名的前后以及中间允许有多个空格， species可以通过包括或者不包括命名人（变型f.之前的命名人除外）。status函数可以用来查询变种 var. \ 亚种 subsp. 以及变型f. 否接受等信息.

使用方法:

library(plantlist)

## 建立一个物种名单

sp <-c("Elaeocarpus decipiens",

"Syzygium buxifolium",

"Daphniphyllum oldhamii",

"Loropetalum chinense",

"Rhododendron latoucheae",

"Rhododendron ovatum",

"Vaccinium carlesii",

"Schima superba")

### 检查名单内每个种的接受情况以及接受名

status(sp)

### 检查一个种的接受情况

status("Myrica rubra") ## 杨梅

status("Adinandra millettii") ## 杨桐

status("cyclobalanopsis neglecta ") ## 竹叶青冈

status("Lirianthe henryi") ### 大叶木兰

#### 同时检查几个学名的状态

status(c("Myrica rubra", "Adinandra millettii",

"Machilus thunbergii", "Ranunculus japonicus",

"Cyclobalanopsis neglecta"))

### Check the statusof a scientific name (with or without authorship)

### 查询学名是否接受

status("Hypoxis filifolia")

### Subspecies (withor without authorship)

### 查询亚种是否接受

status("Hypoxis kilimanjarica subsp. kilimanjarica")

### Variaty (with orwithout authorship)

### 查询变种是否接受

status("Hypoxis erecta var. aestivalis")

### Form (with orwithout authorship)

### 查询变型是否接受

status("Hypoxis hirsuta f. villosissima")

### 重要提示： 由于表示变型的 f. 有时也用于命名人中，

### 因此，用status函数在查询变型时， 请勿为种添加命名人，

### 但是变型的命名人可以提供或者不提供，具体为:

"Hypoxis hirsuta (L.)Coville f. vollosissima Fernald".（不能处理）

"Hypoxis hirsuta f. vollosissima Fernald"（能处理）

"Hypoxis hirsuta f. vollosissima"（能处理）

taxa.table 基于TPL查询结果制作科属种列表， 以便用于Phylomatic软件建树

举例：

sp <- c( "Ranunculus japonicus", "Anemone udensis",

"Ranunculus repens", "Ranunculus chinensis",

"Solanum nigrum", "Punica sp." )

res <- TPL(sp)

taxa.table(res)

## 3.4 TPL根据拉丁学名，查询目、科、属、以及科的编号

TPL函数 输入的数据必须是字符串格式的向量。 可以查询科、属、种的相应信息， 但是并不查询学名是否有效。 查询学名是否有效， 请用status函数

TPL("Carex") #查询薹草属

TPL("Apple") # 查询苹果英文名

splist <- c( "Ranunculus japonicus",

"Solanum nigrum",

"Punica sp.",

"Machilus", "Today", "####" ) ### 查询多个种

res <- TPL(splist)

# 4. 引用plantlist程序包：

Jinlong Zhang (2017). plantlist: Looking Up the Status of Plant Scientific Names based on The Plant List Database. R package version 0.3.0. <https://github.com/helixcn/plantlist/>

# 致谢

感谢高芳銮、李嵘、胡晓丽、冯嘉恩、黄世芳、俞筱押、胡海花、李家湘、刘水银、鲍志贵、张美霞、葛斌杰、孔德良、刘振稳、龙文兴、金建军等各位老师同学试用本软件并提出宝贵意见。如有遗漏，请尽快与本人联系，以便修改致谢名单。

# 参考文献

(待添加)

plantlist版本0.3.0, 本指南版本0.0.2