|  |
| --- |
| 江西中医药大学 |
| Perl脚本使用说明 |
|  |

|  |
| --- |
| 李元  2018/1/27 |

# 一、配置Perl运行环境

## 1. 获取操作系统类型（32bit或64bit）



图1 查看当前系统的类型

## 2. Windows下Perl的下载与安装

2.1 Perl的下载

由于max os以及linux自带了Perl，因此我们演示如何在windows下载及安装Perl环境。如果系统类型是64位的，就选择64bit安装包，32位的即选择32bit安装包。

官方下载网址为：<http://www.perl.org/get.html>，当进入下载网址后，windows系统中存在两种Perl环境，这里我们选择ActiveState Perl (图2)，并点击进入下载页面。根据当前操作系统的类型选择合适的ActiveState Perl，点击图中的“download”（图3）即可自动开始下载。

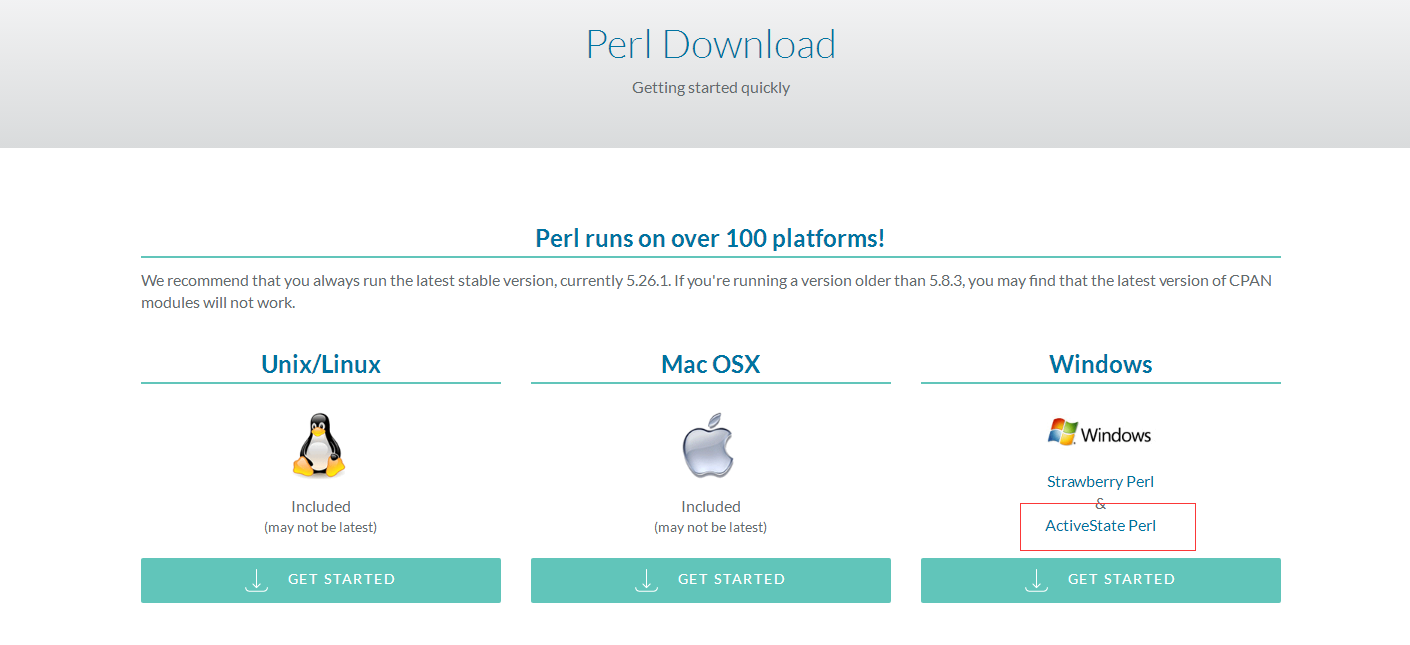


图2 不同操作系统的Perl环境

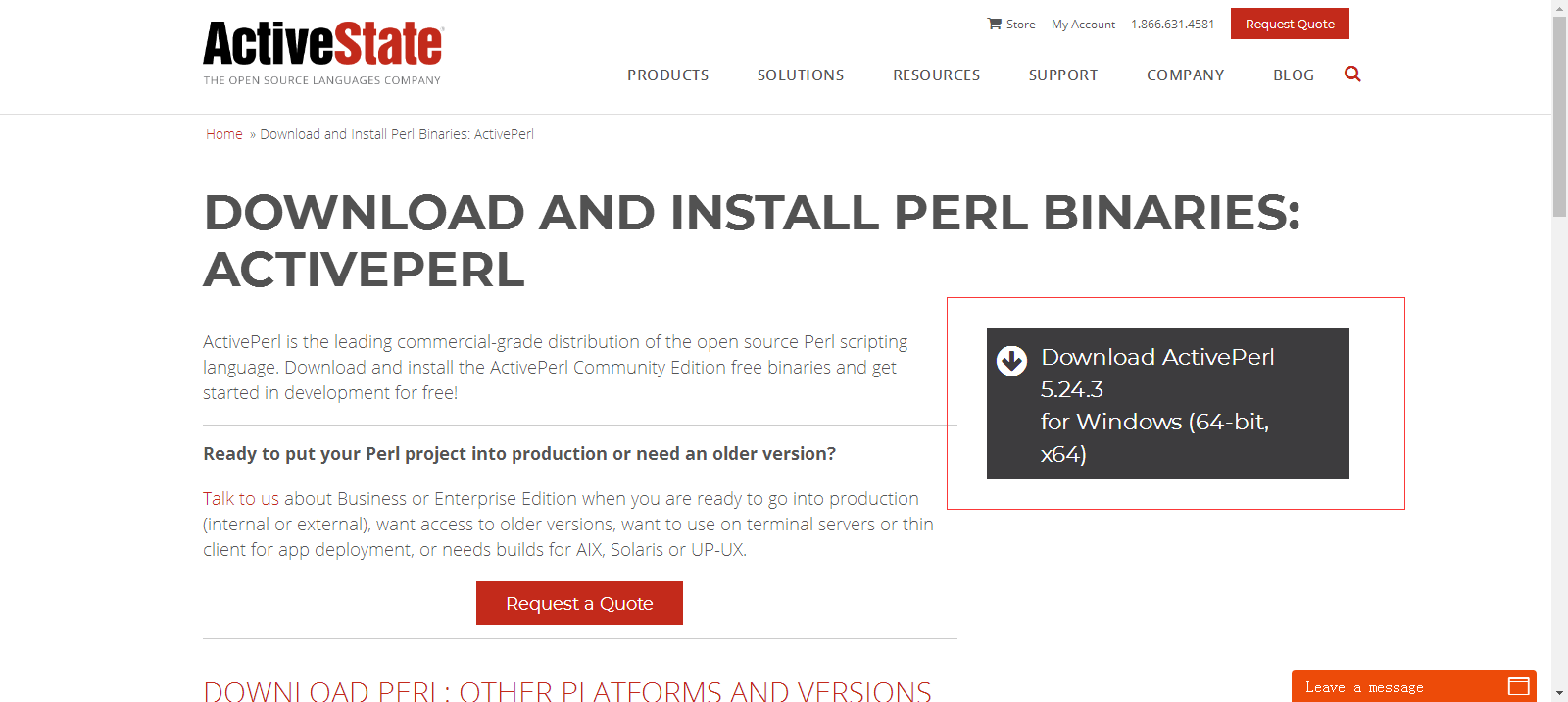


图3 下载Windows系统的Perl

2.2 Perl的安装

当我们下载完毕后，双击执行，当弹出安装欢迎界面（图4），我们的具体步骤如下：

1. 我们选择“next”进行下一步操作。

2. 显示的是软件使用协议，我们选择“accept..”选项后继续选择“next”（图5）。

3. 显示的是不同安装类型，这里我们选择“typical”（图6）。

4. 显示的是相关信息配置界面，这里我们勾选前两个框，分别让自动配置path路径以及关联.perl文件。第3个框为安装使用的编译器，有需求的读者也可使用（图7）。

5. 安装完成后，在键盘同时按下“win”+ “R” 组合键弹出“运行窗口”（图7），在窗口中输入“cmd”，再按回车即可进入dos操作界面。在dos操作界面输入“perl -version”，当弹出一系列信息后，证明ActiveState Perl安装成功（图8）。

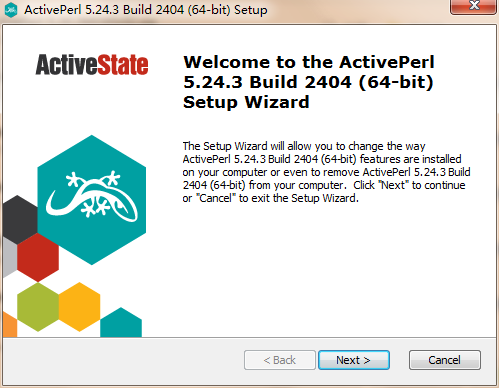


图4 安装ActiveState Perl-1

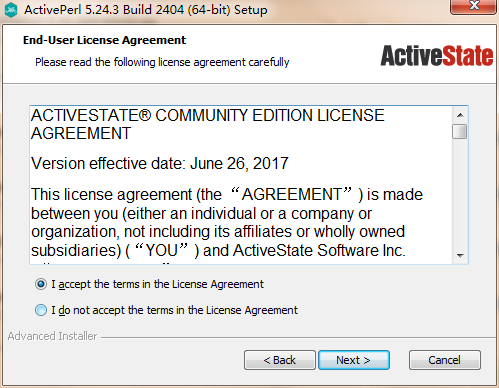


图5安装ActiveState Perl-2

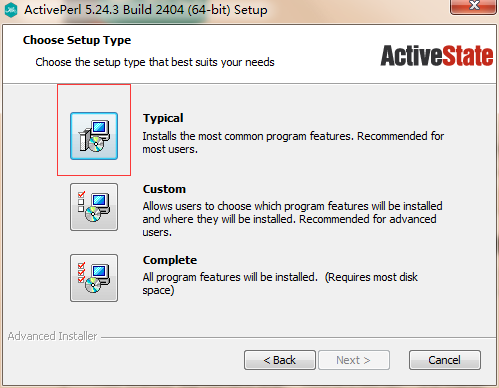


图6 安装ActiveState Perl-3

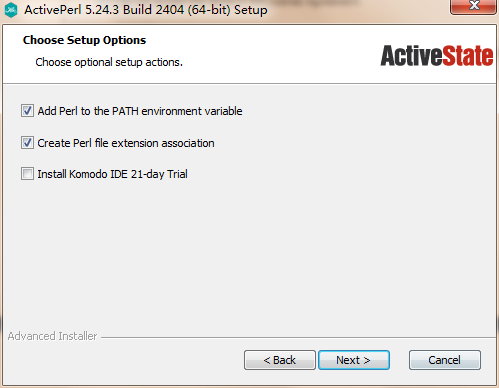


图7 安装ActiveState Perl-4

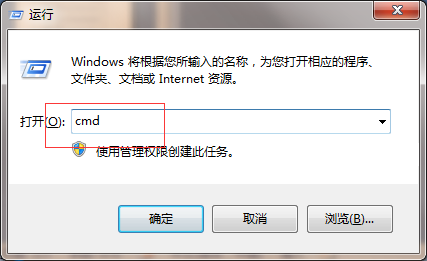


图8 打开运行窗口

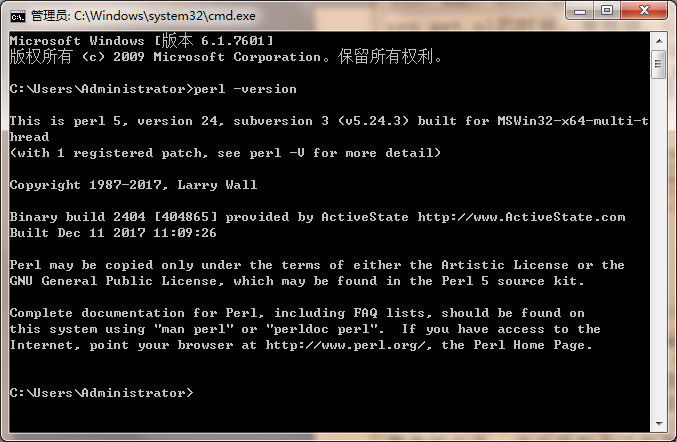


图9 测试ActiveState Perl

# 二、运行程序准备工作

## 1. 准备测试文件

a. 准备tax\_get.pl测试文件，名字为“test.taxon”，文件内容如下，并置于当前路径。

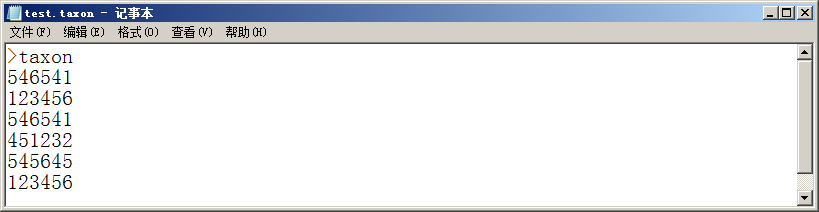


图10-1 taxon测试文件

b. 准备seq\_get.pl测试文件，名字为“16S\_sequences(for\_test).txt”，文件内容如下，并置于当前路径。

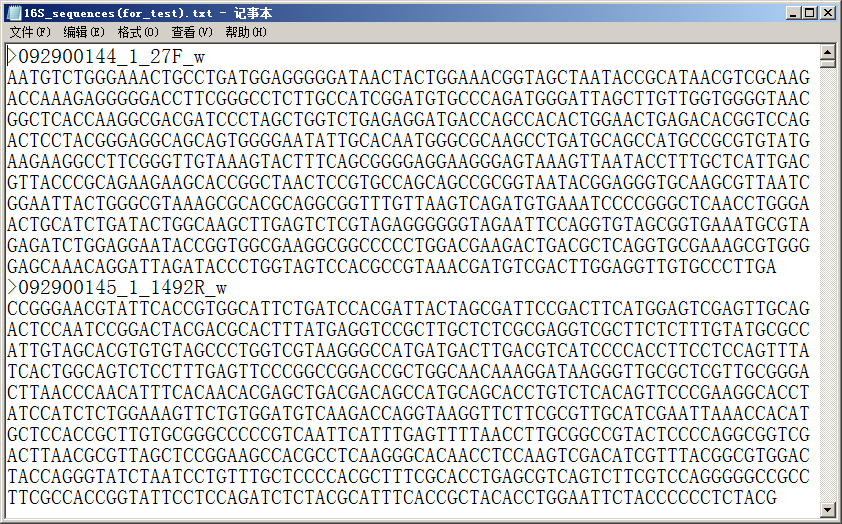


图10-2 blast测试文件

## 2. 在dos命令行进入当前路径（图11）

a. 在dos命令窗口输入命令“cd D:\\TestDir”，接着回车确认，其中D:\\TestDir为笔者测试的目录，读者可自行调整。

b. 因为笔者测试文件在D盘，而系统盘在C盘，因此需要切换盘符才能进入目的路径。继续在dos命令窗口输入“cd D:”，输完继续回车确认。

c. 输入完成后会转换成当前路径“D:\TestDir”。



图11 进入当前路径

## 3. 运行tax\_get.pl

a. 在当前路径下输入“perl tax\_get.pl”，回车执行。执行完毕后，会显示程序使用菜单（图12）。

b. 在光标出继续输入文件名“input.taxon”，回车执行（图13）。

c. 执行时，会先对taxon进行除重操作，接着动态打印当前查找的taxon，如果未发现taxon对应的物种分类，则会提示信息并跳过。执行完毕后，会在当前路径产生一个文件“ncbi.reg.out”，用于记录taxon对应的物种分类信息。该信息与NCBI的taxonomy网站“界门纲目科属种”分类一致（图14）。

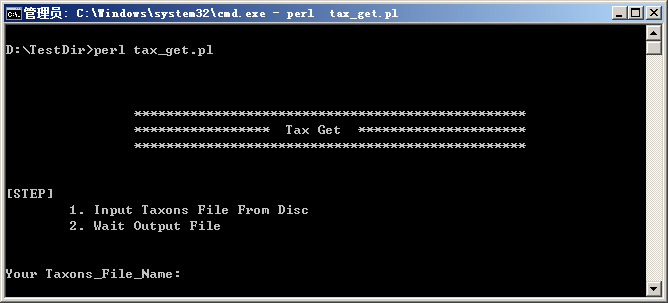


图12 tax\_get.pl 使用说明

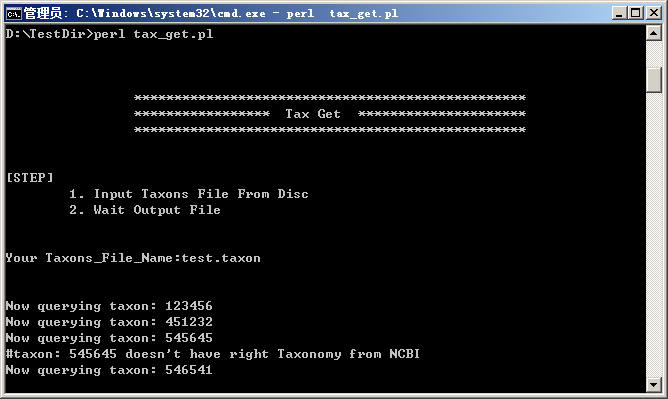


图13 输入taxon文件

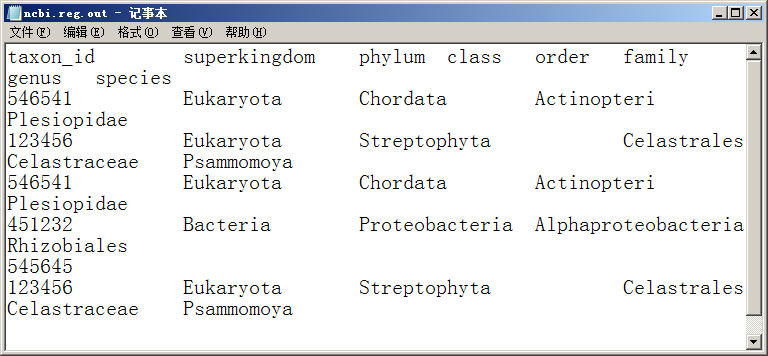


图14 tax\_get.pl测试结果

## 4. 运行seq\_get.pl

a. 在当前路径下输入“perl seq\_get.pl”，回车执行。执行完毕后，会显示程序使用菜单（图15）。

b. 在光标出继续输入待测序列文件名（FASTA格式）“input.fasta”，回车执行（图16）。

c. 接着会弹出下一个菜单（图17）提示选择WEB-BLAST方式，因为我们测试使用的是16s序列，这里我们输入“0”。接着还需要设置数据库，我们选择“3”。

d. 接着会提示输入输出文件名（图18），我们输入的是“output.file”，回车后继续提示输入rank的值，我们输入2，意味着保存查找序列的前2个最匹配的结果（图19）。

e. 最后我们回车确定，此时会自动提交到ncbi服务器进行查找、自动化下载。控制台会实时跟进当前信息，结果会保存到”output.file”中,我们将结果文件直接复制到excel中，并进行了相应处理（图20-1, 图20-2）。

本程序是基于NCBI-BLAST进行的自动化操作，提交的序列及相关信息也均以NCBI为准，用户也可以通过更改源代码，直接输入RID进行自动化下载。

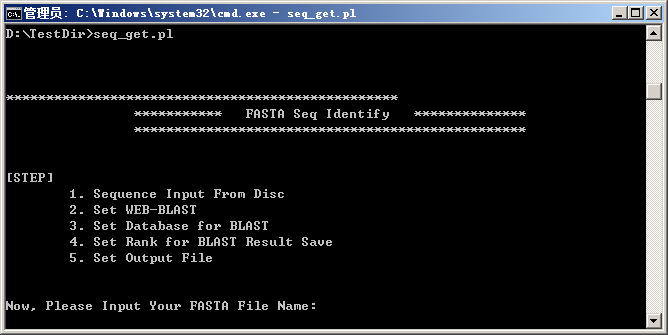


图15 tax\_get.pl运行

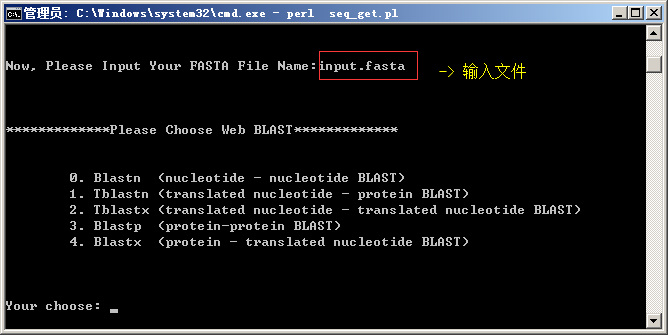


图16 input文件

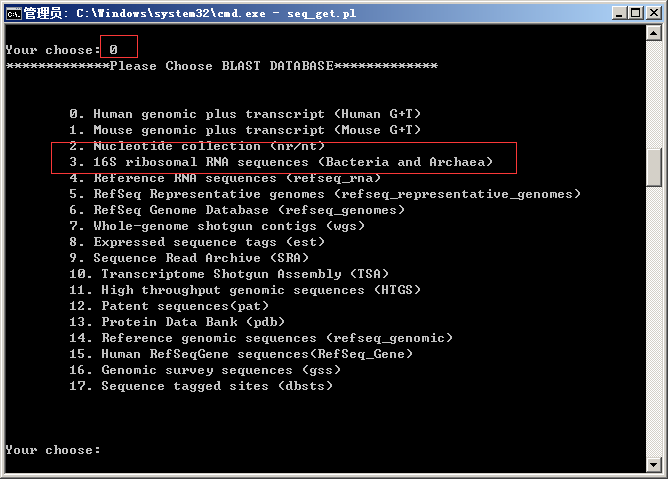


图17 设置blast方式

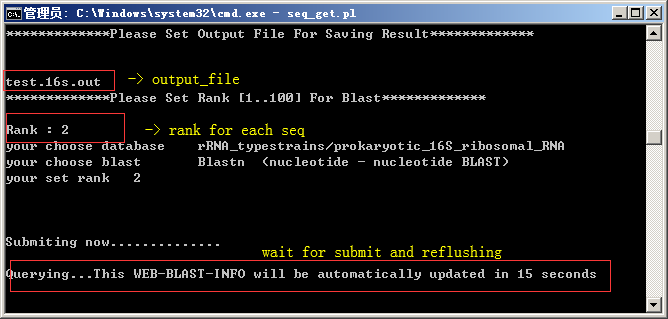


图18 设置output以及rank

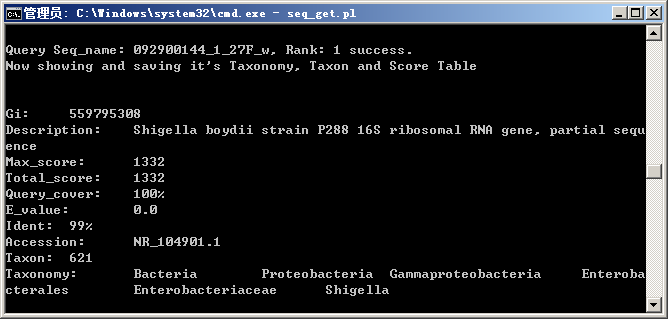
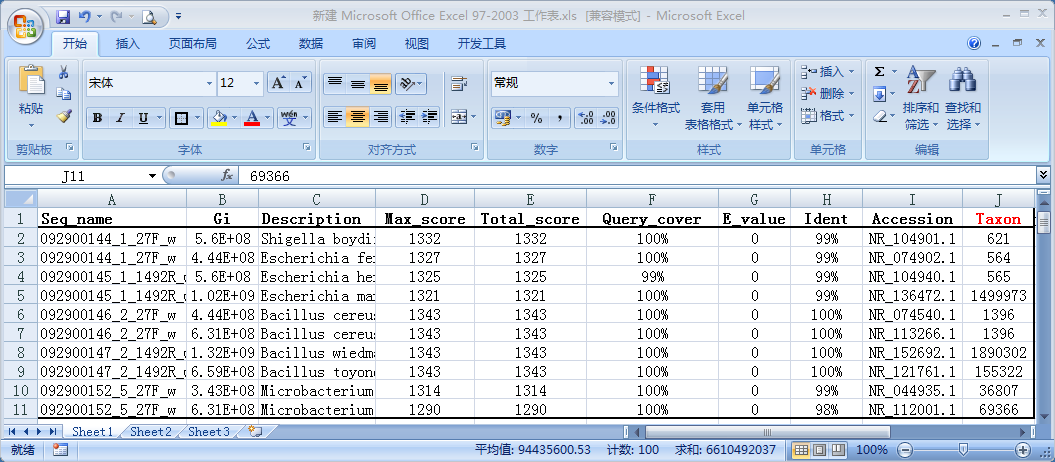


图19 动态打印结果

图20-1 保存结果1

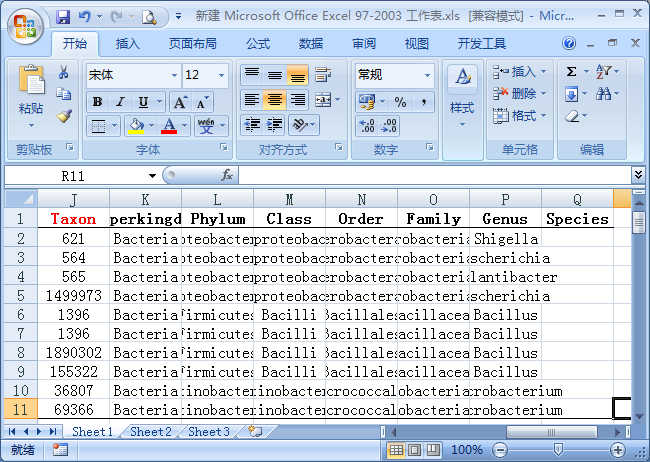


图20-2 保存结果2