+------------------------------------------------------------------------------+

| 转录组项目信息分析结果说明文档

+------------------------------------------------------------------------------+

目录结构及文件说明：

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

|-- BMK\_4\_Unigene\_Anno #Unigene功能注释分析结果目录

| |-- All\_Database\_annotation.xls #功能注释整合文件（excel格式）

| |-- Function\_Annotation.stat.xls #Unigene在各个数据库注释结果统计文件

| |-- BMK\_1\_annotation #各个数据库详细注释结果目录

| |-- BMK\_2\_statistic #各个数据库统计图目录

| |-- BMK\_3KEGG\_map #注释到的KEGG通路图目录

| | |-- ko00010.png

| | |-- ... ...

| | `-- ko04626.png

===================文件夹====================

BMK\_1\_annotation :各个数据库详细注释结果目录,里面包含里各个数据库比对结果的详细信息。

BMK\_2\_statistic:基因在各个数据库统计结果和作图结果目录  
BMK\_3\_Kegg\_map:基因所在的所有通路图。

##############################################################

#文件：Function\_Annotation.stat.xls

#描述；Unigene在各个数据库注释结果统计文件

#第一列Anno\_Database：注释的数据库

#第二列Annotated\_Number：此数据库注释的Unigene个数

#第三列300<=length<1000：此数据库注释的长度在300到1000bp的Unigene个数

#第四列length>=1000：此数据库注释的长度大于等于1000bp的Unigene个数

注：第二列数字不等于第三列和第四列之和，因为没有考虑小于300bp的Unigene个数

###############################################################

# 文件：All\_Database\_annotation.xls

# 描述：功能注释、富集和统计整合文件（6个sheets）

# 字段解释：

###############################sheet

# INTegrated\_Function.anno: 所有基因的注释信息

# #GeneID：基因名称

# COG\_class：COG库注释的功能分类

# COG\_class\_annotation：COG库功能分类描述

# GO\_annotation：GO注释

# KEGG\_annotation：KEGG库注释

# KOG\_class：KOG库注释的功能分类

# KOG\_class\_annotation：KOG库功能分类描述

# Swissprot\_annotation：Swissprot库注释

# NR\_annotation：NR库注释

# Pfam\_annotation：Pfam库注释

###############################sheet

# Function\_anno.stat：注释结果统计

# #Anno\_Database：注释用的数据库

# Annotated\_Number：注释到的基因数目

# 300<=length<1000：注释到的基因长度在300到1000的数目

# length>=1000：注释到的基因长度大于等于1000的基因数目

###############################sheet

# GO.list：每个基因注释到GO数据库上的编号

###############################sheet

# GO\_tree.stat：对于各个GO功能分类，注释到的基因统计

# #GO\_Function：GO功能分类

# Unigene\_number：注释到的基因数目

# Unigene\_ID：注释到的基因名称

###############################sheet

# Kegg.pathway：Kegg上注释到的基因信息

# #pathway：pathway 名称

# pathway\_id：pathway 编号（一个ko编号表示一个通路）

# Gene\_number：注释到某pathway上的基因数目

# Gene\_id：注释到某pathway上的基因名称

# KOs：KO分类信息（一个代谢通路上有哪些KO分类，它是蛋白质（酶）的一个分类体系，序列高度相似，并且在同一条通路上有相似功能的蛋白质被归为一组，然后打上K标签）

###############################sheet

# Kegg.ko：每个基因的kegg注释信息

# #Gene\_id：基因编号

# KO|e\_value|Database\_Genes|Anno：四列值，分别表示基因的注释到的K编号，blastx比对上基因Ｅ值，比对上的基因ID以及它的功能注释

################################################################