

使用 MolAICal 进行配体相似性比较或搜索的操作教程

作者：MolAICal (update 2020-08-04)

更多教程（含英文教程）请见如下：

MolAICal 官方主页：<https://molaical.github.io>

MolAICal 文章介绍：<https://doi.org/10.1093/bib/bbaa161>

MolAICal 中文博客：<https://molaical.github.io/tutorial.html>

MolAICal blogspot：<https://qblab.blogspot.com>

1. 简介

某些情况下配体的相似性比较和搜索可以帮助科学家快速找到有潜力的化合物。本教程介绍了分子指纹相似性和 3D 结构相似性的比较和搜索，本教程可以用于分子相似性比较，和基于指定配体结构的简单虚拟筛选。

2. 工具

2.1. 所需软件下载地址

1) MolAICal: <https://molaical.github.io>

2.2. 操作示例文件

所有用到的操作教程文件均可在下面的网站下载：

<https://github.com/MolAICal/tutorials/tree/master/011-similaritySearch>

3. 操作流程

3.1. 指纹相似性搜索

转至“011-similaritySearch/fingerprint”文件目录下：

```
#> cd 011-similaritySearch/fingerprint
```

1. 在 Linux 控制台中使用“ls > listfiles.dat”命令，或者在 windows 系统的 DOS 控制台使用命令“dir /b > listfiles.dat”命令。打开生成的“listfiles.dat”文件，删除无用描述，确保“listfiles.dat”中只包含配体的名称。

2. 运行以下命令：

```
#> molaical.exe -tool finger -i listfiles.dat -o result.dat
```

命令运行结束后会生成“result.dat”文件，该文件包含了与“listfiles.dat”中名称顺序相对应的结构相似预测数据。

注意：“result.dat”文件中除了以“#”开头的注释行，第一行是被用于比较的配体。

3.2. 3D 结构相似性搜索及比较

转至“011-similaritySearch/3Dsimilarity”文件目录下：

```
#> cd 011-similaritySearch/3Dsimilarity
```

1. 在 Linux 控制台中使用“ls > listfiles.dat”命令，或者在 windows 系统的 DOS 控制台使用命令“dir /b > listfiles.dat”命令。打开生成的“listfiles.dat”文件，删除无用描述，确保“listfiles.dat”文件中只包含配体的名称。

2. 运行以下命令：

```
#> molaical.exe -tool 3Dcompare -i listfiles.dat -s result.dat -f mol2list
```

命令运行结束后会生成“result.dat”文件，该文件包含了与“listfiles.dat”中除去第一行的名称顺序相对应的结构相似预测数据。

注意：“result.dat”文件中第一行不是被比较的配体。“result.dat”中的第一行与“listfiles.dat”中的第二行相对应，“result.dat”中的第二行与“listfiles.dat”中的第三行相对应，以此类推。

此外，MolAICal 可以比较两个配体的 3D 结构相似性，命令如下：

```
#> molaical.exe -tool 3Dcompare -i ligand.mol2 -s 1_10023_out.mol2
```

3.3. 将结果排序

排序选用 3D 结构相似性搜索的结果，你也可以参考这个例子对指纹相似性搜索的结果进行排序。

如果你是用的是 Linux 系统，首先“listfiles.dat”文件需要和“result.dat”相对应。删除“listfiles.dat”中的第一行，执行以下命令来完成此操作：

```
#> molaical.exe -tool col -f listfiles.dat -l result.dat -s " " -o merge.dat
```

```
#> sort -n -t ' ' -k 2r merge.dat > rank.dat
```

此步所生成的“rank.dat”文件包含排序的结果。

如果你用的是 Window 系统，首先“listfiles.dat”文件需要和“result.dat”相对应。删除“listfiles.dat”中的第一行，通过以下步骤来完成此操作：

打开 Excel，将“result.dat”中的内容复制到第一列，将“listfiles.dat”中的内容复制到第二列。选择第一列中所有的数据，然后点击 Excel 上的“Sort Largest to Smallest”(如图 1 所示)。

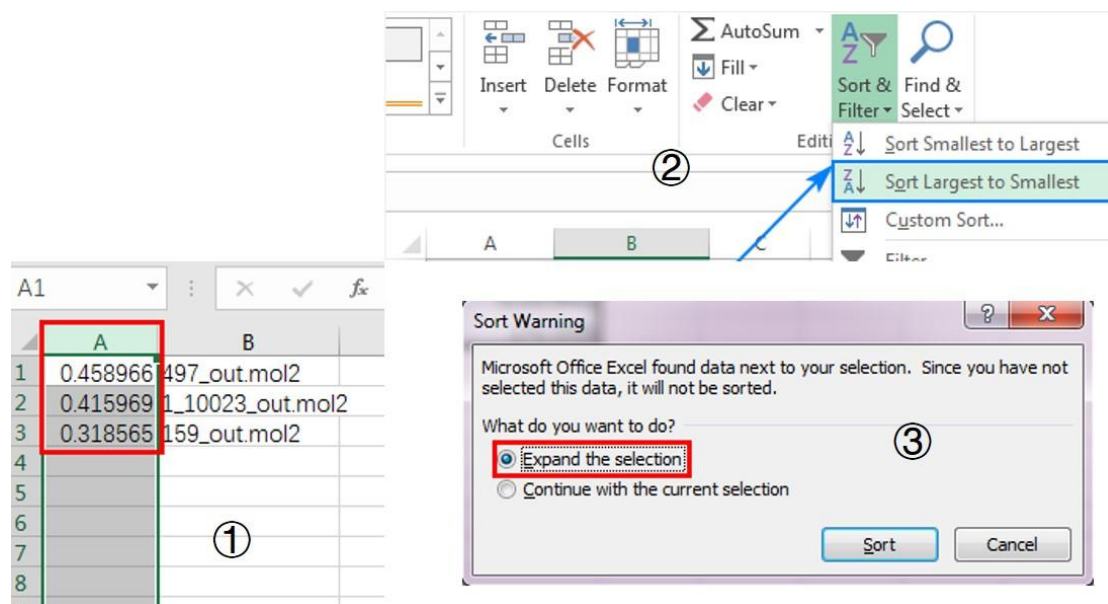


图 1. Excel 中排序操作示意图