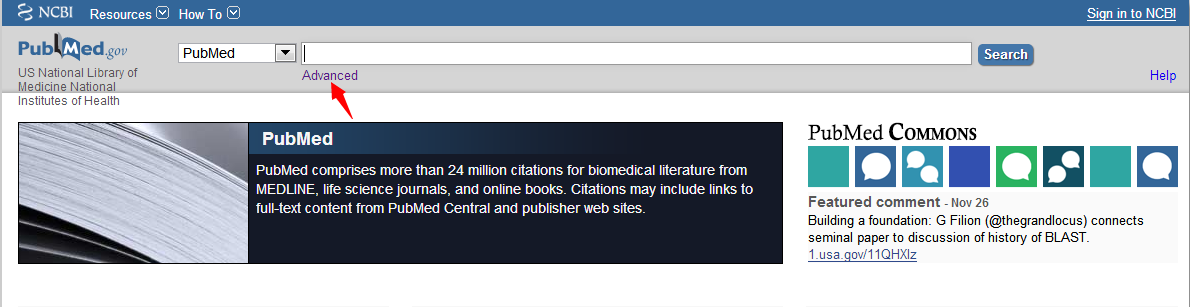
棕榈酰化修饰数据收集需求

棕榈酰化修饰（Palmitoylation）是蛋白质脂质化修饰的一种。棕榈酸盐是丰度最高的脂肪酸，可被共价修饰到蛋白质上，根据连接方式不同，棕榈酰化修饰可分为N型( 通过酰胺键连接到Gly /Cys残基上) 和S型，其中S-Palmitoylation是最常见的形式。S-Palmitoylation是指饱和的16个碳的棕榈酸盐通过硫酯键共价修饰到蛋白质Cys(简写C)的巯基上， 它是一种动态可逆的翻译后修饰形式，可在时间和空间上调节蛋白质的功能。

之前，我们实验室开发了一款预测棕榈酰化修饰位点的软件CSS-Palm，该软件可以预测发生棕榈酰化修饰的蛋白质位点。该软件的形成过程首先是通过收集被实验验证的发生棕榈酰化修饰的蛋白质位点，然后以这些修饰位点为训练集，用我们实验室独立开发的GPS算法进行数据训练，获得棕榈酰化修饰位点的预测模型，然后对用户提交的数据进行预测。

由于生物文章在不断地增加，发生棕榈酰化修饰的位点也在不断增加，因此我们必须隔一段时间进行数据更新及收集。我们数据收集的方式是这样的：

1. 在Pubmed数据库中搜索摘要中包含关键字Palmitoylation or Palmitoylated的文献



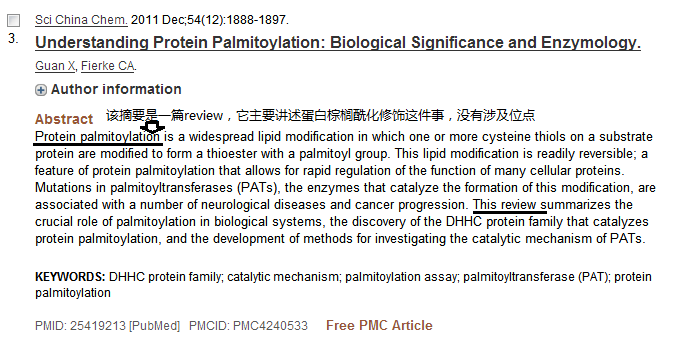




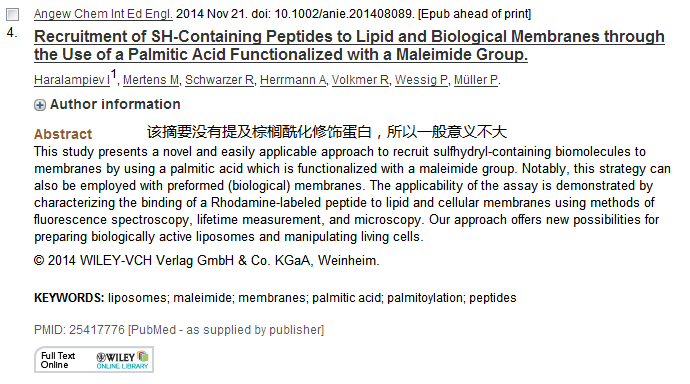
结果获得了2129篇文献。

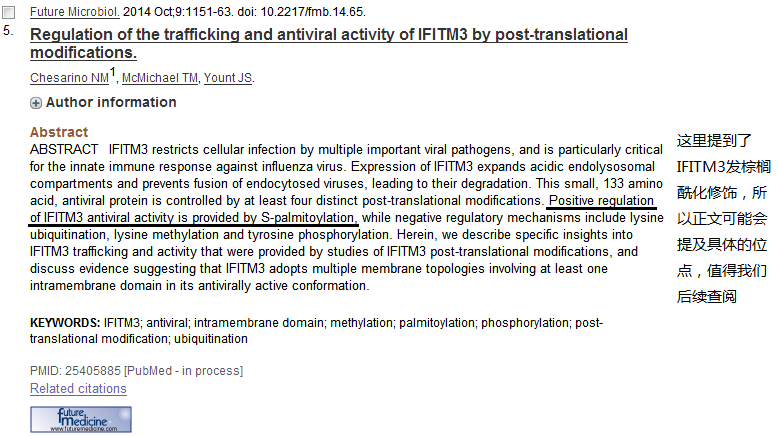
1. 查阅文献：通过查阅文献获取哪个蛋白发生棕榈酰化修饰，它发生修饰的位点在哪里及它发生在哪个物种上。下面主要列举了不同文章abstract的情况：

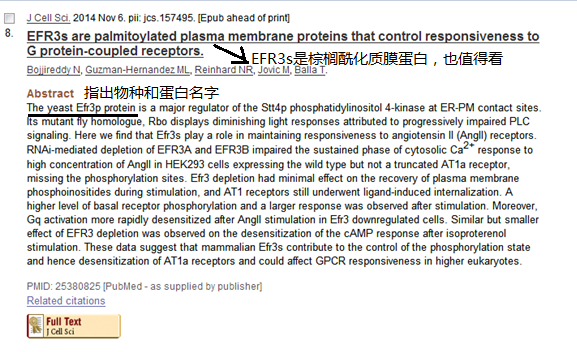




这种综述性的文章一般只做参考







以上是列出的文献摘要的几种情况。

总而言之，我们收集的文献最好不是综述性文章（但可以将所有相关的综述性文章放在一个文件夹，便于以后了解背景的了解），文献abstract中必须出现哪个蛋白发生了棕榈酰化修饰，至于位点信息可以有，也可以没有，因为位点信息可以查阅整篇文献得知。

利用关键字在Pubmed数据库搜索相关文献，它只能相当于一个初筛，而且查阅很多文献后发现至少有1/3的文献并没有我们要的位点信息，还有相当一部分文献强调的是同一个蛋白，如果能够通过数据挖掘的方式对文献进行归类，那对我们生物学者们来说将节约到很多时间与精力。