1. 数据来源

我们从Uniprot数据库中下载了人类、蝙蝠的ACE2氨基酸序列，人类的ACE2氨基酸序列经过可变剪接可分为2个亚型，共计2个序列：Q9BYF1-1（805AA）和Q9BYF1-2（555AA）；蝙蝠的ACE2氨基酸序列根据种属不同可分为6种，共计10个序列。12个氨基酸序列的详细信息储存于[分析过程记录](https://github.com/longqh3/Record-for-bioinformatic-data-preprocessing/blob/master/%E4%BA%BA%E7%B1%BB%26%E8%9D%99%E8%9D%A0ACE2%E8%9B%8B%E7%99%BD%E5%BA%8F%E5%88%97%E5%AF%B9%E6%AF%94.md)中，经汇总后的fasta文件名为”human\_bat\_ACE2\_sequence\_compare\_request.fasta”。

1. 多序列比对分析

应用[CLUSTALW](https://www.genome.jp/tools-bin/clustalw)软件进行多氨基酸序列比对，得到名为”human\_bat\_compare\_clustalw.aln”比对结果，上传结果于[ESPript 3](http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) [1]上进行可视化分析后，得到结果文件” human\_bat\_ACE2\_sequence\_compare.pdf”。现有文献表明，蝙蝠ACE2蛋白序列的差异位点主要位于N-端区域 （N-terminal region），该区域在既往的分子结构研究中被认为是SARS冠状病毒的结合区域，提示并非所有蝙蝠的ACE2蛋白都可作为SARS冠状病毒的功能受体[2]。而在最新的研究中，进化树分析表明当前流行的2019新型冠状病毒（2019-nCoV）起源于蝙蝠，且被证实和SARS冠状病毒同样使用了ACE2蛋白作为其进入细胞的受体[3, 4]，故而有必要基于2019新型冠状病毒的受体结合模体（RBM）信息，研究其与不同种的蝙蝠ACE2蛋白的相互作用，进而帮助进一步确证其传染源头。

1. 进化树构建

应用MEGA-X[5]软件进行进化树构建，采用邻接法（Neighbor-Joining）来建树、bootstrap法对建树的可靠性进行评估，得到进化树对应可视化文件“human\_bat\_ACE2\_sequence\_compare.svg”。节点数据均大于70，提示该分类计算结果可靠性有所保证。

1. 序列一致性分析

应用BLAST对人类和蝙蝠的ACE2氨基酸序列进行比较，发现蝙蝠的ACE2与人类的ACE2高度相关，且氨基酸一致性在79%-81%之间。现有文献表明，不同属蝙蝠间ACE2的氨基酸序列一致性在78%-84%之间，而在菊头蝠（Rhinolophus）属内的四种蝙蝠（中华菊头蝠、马铁菊头蝠、小菊头蝠、大耳菊头蝠）间，ACE2的氨基酸序列一致性则增加至89%-98%。

**参考文献**

[1] Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server[J]. Nucleic Acids Research. 2014, 42(W1): W320-W324.

[2] Hou Y, Peng C, Yu M, et al. Angiotensin-converting enzyme 2 (ACE2) proteins of different bat species confer variable susceptibility to SARS-CoV entry[J]. Archives of Virology. 2010, 155(10): 1563-1569.

[3] Zhou P, Yang X, Wang X, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin[J]. Nature. 2020.

[4] Wan Y, Shang J, Graham R, et al. Receptor recognition by novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS[J]. Journal of Virology. 2020.

[5] Kumar S, Stecher G, Li M, et al. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms[J]. Molecular Biology and Evolution. 2018, 35(6): 1547-1549.