### Conclusions

In this article, we presented an improved algorithm based on the epsm algorithm for string matching problem on biological sequences. Based on characteristics of biological sequences, the new algorithm took the optimization size of the packed string to generate the hash table. Besides, the new algorithm used a 64bit number as a fingerprint of string, the fingerprint was generated by 8 bytes string. We used the fingerprint filtering many comparisons before taking byte-by-byte comparison. And the experiment results proved using these two optimized strategies, our new algorithm was obviously more efficient than the epsm.

Our further studies would apply the new algorithm to the multiply exactlly string matching problem.

本文针对生物信息学中的单串匹配问题，在epsm基础上提出了一个改进的算法。新算法针对生物信息的特单，采用优化的块字符大小来生成hash表。除此之外，新算法采用一个64bit的大整数来表示字符串的指纹，该指纹通过8个字符生成。算法在进行逐字节的比较前通过该指纹过滤掉很多逐字匹配的调用。实验结果表明，以上两个优化策略，使得新算法明显比epsm高效。  
将来的研究包括本算法在多字符串匹配问题中的应用等。