

研究论文

基于转录终点序列特征预测大肠杆菌sRNA

刘倩, 应晓敏, 吴佳瑶, 查磊, 李伍举

军事医学科学院基础医学研究所计算生物学中心, 北京 100850

摘要:

细菌sRNA是一类长度在40~500 nt的调控RNA, 在细菌与环境相互作用中发挥重要功能, 因此, 细菌sRNA识别研究具有重要意义。然而, 与蛋白编码基因具有易于识别的特征不同, 目前细菌sRNA识别仍是一件比较困难的事。此方法介绍了一个基于已知细菌sRNA转录终点的碱基频率矩阵来识别sRNA的预测策略, 并在大肠杆菌K-12 MG1655中进行了sRNA的预测。结果表明, 该模型在独立测试集中具有较高的特异性和阳性检出率, 因此, 这一方法将为实验发现细菌sRNA提供较好的生物信息学支持。

关键词: 细菌sRNA 终点信号 预测

Prediction of sRNA in *E. coli* Using The Sequence Features Derived from Transcriptional Termination

LIU Qian, YING Xiaomin, WU Jiayao, ZHA Lei, LI Wujun

Center of Computational Biology, Beijing Institute of Basic Medical Sciences, Beijing 100850, China

Abstract:

Bacterial sRNAs are an emerging class of regulatory RNAs, 40~500 nt in length. These sRNAs play a key role in bacteria-environments interaction, therefore, the identification of sRNAs is very important. Compared to the prediction of protein-coding genes with distinguished features, however, prediction of sRNA-coding genes is still challenging. Here we report on a strategy to predict sRNA-coding genes using the base distribution frequency matrix derived from the known sRNA genes, and *E. coli* K-12 MG1655 is exemplified to demonstrate the performance of the strategy. The results indicate that the presented model has both higher specificity and higher positive prediction value compared to the previous models. The model provides better support for sRNA genes discovery using experimental protocols.

Keywords: Bacterial sRNA Transcriptional termination Prediction

收稿日期 2010-12-06 修回日期 2011-02-25 网络版发布日期

DOI: 10.3724/SP.J.1260.2011.00257

基金项目:

“863”计划项目(2006AA02Z323), 国家自然科学基金项目(31071157)

通讯作者: 李伍举, 电话: (010)66931324, E-mail: liwj@nic.bmi.ac.cn

作者简介:

作者Email: liwj@nic.bmi.ac.cn

参考文献:

- Huang HY, Chang HY, Chou CH, Tseng CP, Ho SY, Yang CD, Ju YW, Huang HD. sRNAMap: Genomic maps for small non-coding RNAs, their regulators and their targets in microbial genomes. *Nucleic Acids Res*, 2009, 37: D150~154
- Wanger EG. Kill the messenger: Bacterial antisense RNA promoters mRNA decay. *Nat Struct Mol*

扩展功能

本文信息

- Supporting info
- PDF(881KB)
- [HTML全文]
- 参考文献[PDF]
- 参考文献

服务与反馈

- 把本文推荐给朋友
- 加入我的书架
- 加入引用管理器
- 引用本文
- Email Alert
- 文章反馈
- 浏览反馈信息

本文关键词相关文章

- 细菌sRNA
- 终点信号
- 预测

本文作者相关文章

- 刘倩
- 应晓敏
- 吴佳瑶
- ??磊
- 李伍举

PubMed

- Article by Liu, Q.
- Article by Ying, X. M.
- Article by Wu, J. Y.
- Article by Cha, L.
- Article by Li, W. J.

3. Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol*, 2009, 16(8): 840~806
4. Tjaden B. TargetRNA: A tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res*, 2008, 36:W109~113
5. Busch A, Richter AS, Backofen R. IntaRNA: Efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics*, 2008, 24(24):2849~2856
6. Fortune DR, Suyemoto M, Altier C. Identification of CsrC and characterization of its role in epithelial cell invasion in *Salmonella enterica* serovar Typhimurium. *Infect Immun*, 2006, 74(1): 331~339
7. Babitzke P, Romeo T. CsrB sRNA family: Sequestration of RNA-binding regulatory proteins. *Curr Opin Microbiol*, 2007,10(2): 156~163
8. Guillier M, Gottesman S. Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol Microbiol*, 2006, 59(1): 231~247
9. Vecerek B, Moll I, Bläsi U. Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J*, 2007, 26(4): 965~975
10. Massé E, Vanderpool CK, Gottesman S. Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol*, 2005, 187(20): 6962~6971
11. Massé E, Salvail H, Desnoyers G, Arguin M. Small RNAs controlling iron metabolism. *Curr Opin Microbiol*, 2007, 10(2): 140~145
12. Lenz DH, Miller MB, Zhu J, Kulkarni RV, Bassler BL. CsrA and three redundant small RNAs regulate quorum sensing in *Vibrio cholerae*. *Mol Microbiol*, 2005, 58(4): 1186~1202
13. Tu KC, Bassler BL. Multiple small RNAs act additively to integrate sensory information and control quorum sensing in *Vibrio harveyi*. *Genes Dev*, 2007, 21(2): 221~233
14. Toledo-Arana A, Repoila F, Cossart P. Small noncoding RNAs controlling pathogenesis. *Curr Opin Microbiol*, 2007,10(2): 182~188
15. Romby P, Vandenesch F, Wagner EG. The role of RNAs in the regulation of virulence-gene expression. *Curr Opin Microbiol*, 2006, 9(2): 229~236
16. Cao Y, Wu J, Liu Q, Zhao Y, Ying X, Cha L, Wang L, Li W. sRNATarBase: A comprehensive database of bacterial sRNA targets verified by experiments. *RNA*, 2010, 16(11):2051~2057
17. Pichon C, Felden B. Small RNA gene identification and mRNA target predictions in bacteria. *Bioinformatics*, 2008, 24(24): 2807~2813
18. Livny J. Efficient annotation of bacterial genomes for small, noncoding RNAs using the integrative computational tool sRNAPredict2. *Methods Mol Biol*, 2007, 395: 475~488
19. Livny J, Teonadi H, Livny M, Waldor MK. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PLoS One*, 2008, 3(9): e3197
20. Rivas E, Eddy SR. Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics*, 2001,2: 8
21. Washietl S, Hofacker IL, Stadler PF. Fast and reliable prediction of noncoding RNAs. *Proc Natl Acad Sci USA*, 2005, 102(7): 2454~2459
22. Carter P, Liu J, Rost B. PEP: Predictions for entire proteomes. *Nucleic Acids Res*, 2003, 31(1): 410~413
23. Wang C, Ding C, Meraz RF, Holbrook SR. PSoL: A positive sample only learning algorithm for finding non-coding RNA genes. *Bioinformatics*, 2006, 22(21):2590~2596
24. Tran TT, Zhou F, Marshburn S, Stead M, Kushner SR, Xu Y. De novo computational prediction of non-coding RNA genes in prokaryotic genomes. *Bioinformatics*, 2009, 25(22):2897~2905
25. Salari R, Aksay C, Karakoc E, Unrau PJ, Hajirasouliha I, Sahinalp SC. smyRNA: A novel Ab initio ncRNA gene finder. *PLoS One*, 2009, 4(5): e5433
26. Sridhar J, Narmada SR, Sabarinathan R, Ou HY, Deng Z, Sekar K, Rafi ZA, Rajakumar K. sRNAscanner: A computational tool for intergenic small RNA detection in bacterial genomes. *PLoS One*, 2010, 5(8): e11970

本刊中的类似文章

1. 孙之荣,赵卫东,赵南明,杨存荣.用模式识别方法预测膜蛋白RC、BR和RH的二级结构[J]. *生物物理学报*, 1993,9(1): 143-148
2. 陈洁,丁达夫.蛋白质侧链的预测[J]. *生物物理学报*, 1993,9(4): 651-656
3. 汤海旭, 丁达夫.用于蛋白质同源模建及三维结构预测的结构比较方法[J]. *生物物理学报*, 1995,11(1): 60-66
4. 王彦力, 来鲁华, 韩玉真, 徐筱杰, 唐有祺.新型平均势函数在蛋白质反向折叠中的应用[J]. *生物物理学报*, 1995,11(1): 67-74
5. 李宏, 罗辽复.基因表达水平与同义密码子使用关系的初步研究[J]. *生物物理学报*, 1995,11(2): 237-244
6. 江凡, 罗宇, 来鲁华, 徐筱杰.应用死端排除法考虑分子对接中的侧链柔性[J]. *生物物理学报*, 1995,11(2): 249-254
7. 孙之荣, 饶晓谦.用人工神经网络方法预测蛋白质超二级结构[J]. *生物物理学报*, 1995,11(4): 570-574

8. 李伍举, 吴加金. 基于螺旋区随机堆积的RNA二级结构预测[J]. 生物物理学报, 1996,12(2): 213-218
9. 张彦, 刘次全. P53蛋白质N末端的二级结构预测及其三维构象[J]. 生物物理学报, 1998,14(4): 743-749
10. 李萍, 过涛, 李衍达, 孙之荣. 基于小波分析的膜蛋白跨膜区段序列分析和预测[J]. 生物物理学报, 2000,16(3): 577-585
11. 蔡禄, 罗辽复. 三核苷酸重复及相关序列宏观弯曲的预测[J]. 生物物理学报, 2000,16(3): 595-601
12. 陈卫国, 李鹏程, 卢广, 胡波. 用近红外光拓扑图技术实时跟踪脑血流变化[J]. 生物物理学报, 2000,16(3): 613-617
13. 胡磊, 乔立安, 公衍道, 赵南明. 利用支持向量机预测II类MHC分子结合多肽[J]. 生物物理学报, 2001,17(4): 669-675
14. 王娴, 李鹭, 王明会, 冯焕清. 基于支持向量机方法的蛋白可溶性预测[J]. 生物物理学报, 2005,21(1): 60-64
15. 陈双平, 郑浩然, 刘海燕, 王煦法. 蛋白质序列中的关联规则发现及其应用[J]. 生物物理学报, 2006,22(3): 171-176
16. 郭建秀, 马彬广, 张红雨. 蛋白质折叠速率预测研究进展[J]. 生物物理学报, 2006,22(2): 89-95
17. 姜小莹, 朱俊东, 李晓波, 张同亮. 使用伪氨基酸组成和模糊支持向量机预测蛋白质结构类[J]. 生物物理学报, 2008,24(1): 43-48
18. 晋宏营, 罗辽复, 张利绒. 使用估计的反应自由能预测组成性和可变剪接位点[J]. 生物物理学报, 2009,25(1): 57-64
19. 王立贵, 应晓敏, 曹源, 查磊, 李伍举. sRNASVM——基于SVM方法构建大肠杆菌sRNA预测模型[J]. 生物物理学报, 2009,25(4): 287-293
20. 刘建国, 刘建荣, 刘明, 闫蓬勃. 基于进化信息改进蛋白质定点突变稳定性预测准确率[J]. 生物物理学报, 2009,25(5): 343-348
21. 蔡禄, 赵秀娟. 核小体定位研究进展[J]. 生物物理学报, 2009,25(6): 385-395
22. 黄俊峰, 段鹏, 吴文言. 基于模板的蛋白质结构预测[J]. 生物物理学报, 2011,27(1): 28-37
23. 刘辉, 张纪阳, 孙汉昌, 徐长明, 张伟, 马海滨, 朱云平, 谢红卫. 基于马尔科夫链的胰蛋白酶肽段蛋白酶切位点概率预测[J]. 生物物理学报, 2011,27(6): 528-536

文章评论

反 馈 人	<input style="width: 95%;" type="text"/>	邮箱地址	<input style="width: 95%;" type="text"/>
反 馈 标 题	<input style="width: 95%;" type="text"/>	验证码	<input style="width: 50px;" type="text"/> 6390