

研究报告

Research Report

密码子使用模式和成分分布在霍乱弧菌 *N16961* 全基因组和毒力基因之间的比较性研究

Sushanta Deb , Surajit Basak 

分子生物学与生物信息学系, 特里普拉大学, Suryamaninagar, 印度

 通讯作者: basakurajit@gmail.com  作者

计算分子生物学, 2015 年, 第 4 卷, 第 12 篇 doi: 10.5376/cmb.cn.2015.04.0012

本文首次发表在 *Computational Molecular Biology* 上。现依据版权所有人授权的许可协议, 采用 Creative Commons Attribution License 协议对其进行授权, 再次发表与传播。只要对原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。

建议最佳引用格式:

Deb S. and Basak S., 2015, Comparative study of Codon usage pattern and compositional distribution between whole genome and virulence gene set of *Vibrio cholera* N16961., *Computational Molecular Biology*, Vol.5, No.6 1-4 (doi: 10.5376/cmb.2015.05.0006)

摘要 霍乱弧菌是一种能够引起霍乱、严重腹泻疾病的致病性微生物, 经常发生在南亚。霍乱弧菌其致病和非致病性主要区别在于其菌株所含有的毒力基因不同。已有大量不同生物型的霍乱弧菌被发现。这些菌株包含不同的致病因素, 例如接收和传递毒素的基因, 定植因子, 抗生素耐药性, 为氯 7 提供抗性的荚膜多糖抗, 以及例如 O139 唇多糖和 O 抗原胶囊等新的表面抗原。霍乱弧菌的这些致病基因会在不同的阶段发生不同模式的转移, 例如侧向或横向转移。收集这些致病基因和其他辅助遗传因子, 有助于我们了解细菌病原体的致病机理以及其是如何变成一个新的菌株。为了了解这些致病基因的遗传特性和致病基因簇 (VGS) 全部密码子的使用模式, 我们测量了 VGS 的 GC 含量, 结果表明全基因组和 VGS 在 GC 含量上没有任何区别。同时也发现, 全基因组的编码序列 (CDS) 和致病基因簇 (VGS) 在 GC 含量的分布上是相似的。A3s、T3s、G3s、C3s 和 GC3s, 以及 ENC 值, 核苷酸含量 (A%, T%, C%, G%, 和 GC%) 的研究结果显示, 突变偏好性在 VGS 的密码子偏好性塑造中扮演着重要角色。

关键词 密码子使用模式, 全基因组, 致病基因, 霍乱弧菌

Comparative study of Codon usage pattern and compositional distribution between whole genome and virulence gene set of *Vibrio cholera* N16961

Sushanta Deb , Surajit Basak 

Department of Molecular Biology and Bioinformatics, Tripura University, Suryamaninagar, India

 Corresponding author: basakurajit@gmail.com  Author

Abstract *Vibrio cholerae* is the pathogenic organism causes cholera, a severe diarrheal disease. Occurs frequently in southern Asia. *Vibrio cholerae* has both pathogenic and nonpathogenic strains that vary in their virulence gene content. Great variety of strains and biotypes of *Vibrio cholerae* are found. These varieties are involve in shuffling of different pathogenic factors among them such as receiving and transferring genes for toxins, colonization factors, antibiotic resistance, capsular polysaccharides which giving resistance to chlorine 7 and new surface antigens, such as the O139 lip polysaccharide and O antigen capsule. Different mode of transfer of these virulence gene i.e. lateral and horizontal transfer by phase, collection of pathogenic genes and other accessory genetic element, pave the way to understand how bacterial pathogen develop its Pathogenicity and become a new strain. To provide a insights into the genetic features and the relationship between the overall codon usage pattern of virulence gene set (VGS). We measure the GC content of VGS which shows that there is no any difference between GC content of whole genome and VGS. It also has been found that GC content shows the similar distribution among the CDS of both whole genome and Virulence gene set. A correlation analysis between the A3s, T3s, G3s, C3s, and GC3s, the ENC values, and the nucleotide contents (A%, T%, G%, C%, and GC %) indicated that mutational bias plays role in shaping the VGS codon usage bias.

Keywords Codon usage pattern, whole genome, virulence gene, *Vibrio cholera*

收稿日期: 2015 年 10 月 27 日

接受日期: 2015 年 10 月 27 日

发表日期: 2015 年 10 月 27 日

Copyright © 2015 BioPublisher

早在 1910 年, 厄尔托型霍乱杆菌就引起了零星的感染和霍乱流行, 这种生物型菌株在 1961 年又出现, 并引起了第七次流行性疾病。霍乱弧菌是一种革兰氏阴性细菌, 厄尔托霍乱杆菌 N16961 完整的基因组序列含有 4 033 460 个碱基对(bp)。厄尔托霍乱杆菌 N16961 的全基因组被分为两个环状的染色体, 大小分别为 2 961 146 bp 和 1 072 314 bp, 共编码 3885 个 ORF。主要部分的可识别的基因起着主要的作用, 例如细胞功能(如 DNA 复制, 转录, 蛋白质合成和细胞壁生物合成)和致病性(例如毒素, 表面抗原和粘连), 这些基因主要是位于染色体。霍乱弧菌基因组序列有助于我们了解一种游离环境微生物如何进化成为一个重大的人类细菌性病原体。

病原细菌在人类宿主内有很多种致病机制, 细菌病原体有广泛的分子去结合宿主靶细胞以产生不同类型的宿主响应。病原菌与宿主之间作用的分子机制在每个病原菌中是独特的, 或者在几个不同的物种之间是保守的。几种细菌病原体的完整基因组序列的可用性有助于揭示细菌用于感染宿主的分子策略背后的奥秘。“水平基因转移”是以快速和戏剧性方式改变细菌基因组的基因组特征的主要因素之一。最近的研究表明, 水平基因转移在新型细菌病原体的分子进化中发挥重要作用。在细菌基因组中, 可能存在含有大部分毒力决定因子(粘连, 侵入, 毒素, 抗生素抗性蛋白等)的区域, 因此被称为致病岛。已有报道, 部分生物因子在基因组特性和致病基因水平上, 影响细菌的致病性。

1 方法

表 1 毒性基因的编码序列的密码子组成 (A3s, T3s, G3s, C3s 和 GC3s), ENC 值, 核苷酸组成 (A%, T%, G%, C% 和 GC%)

	A%	T%	C%	G%	GC%
A3s	0.8041	0.0707	-0.4902	-0.5241	-0.6393
T3s	0.4029	0.6070	-0.5921	-0.6762	-0.3935
C3s	-0.2190	-0.6002	0.6797	0.1969	0.8144
G3s	-0.6766	-0.1759	0.3012*	0.6587	-0.5269
GC3s	-0.4680	-0.2950 *	0.6952	0.4453	0.5445

注: * 0.01 < P < 0.05

霍乱弧菌 N16961 全基因组的编码基因从 NCBI 数据库 (<http://www.ncbi.nlm.nih.gov/Ftp/>) 获取, 致病基因簇从致病岛数据库 (http://www.paidb.re.kr/about_paidb.php) 下载得到。从全基因组 CDS 中剔除毒力基因组下的基因, 以避免基因组中的 CDS 发生重复。毒性基因组的密码子组成 (A3s, T3s, G3s, C3s 和 GC3s) 使用 John Peden 编写的软件 Codon W 获得 (fppp: <http://molbiol.ox.ac.uk/cu/codon.tar.Z/>)。利用 MEGA 4.0 生物软件分析 VGS 的每个 CDS 的核苷酸含量 (A%, T%, G% 和 C%)。使用统计软件 (www.statsoft.com/Products/STATISTICA/Data-Miner) 进一步统计所获得的数据以获得统计测量的值。我们还测量了此细菌基因组中密码子使用偏好性的程度 (NCdiff)。为了测量 NCdiff, 我们下载了来自 NCBI 的所有核糖体蛋白的编码基因, 我们产生两组编码序列以评价 NCdiff, 即核糖体蛋白编码基因和其余基因。使用 Codon W 软件, 我们已经分析了核糖体基因和其余基因的 ENC 值。

2 结果和讨论

为了研究突变压力是否对 VGS 中的密码子使用偏好有任何可能的影响, 在不同密码子位置 (A3, T3, G3, C3 和 GC3) 的组成, 核苷酸组成 (A%, T%, G%, C% 和 GC%) 和 ENC 值 (表 1) 方面做了相关的分析。结果表明大多数密码子组成与核苷酸组成相关, 此外, ENC 值总是显示与核苷酸组成没有相关性。这些结果证实 VGS 的密码子使用偏好性受核苷酸组成影响, 并因此受突变偏好性的影响。

我们还建立 CDS 集合, 即全基因组 CDS 和 VGS 的 CDS 关于 GC12 VS GC3 图, 在两种情况下, 相关值或多或少相似, 表明类似地影响两个基因集的突变压力。GC12 是 GC1、GC2 和 GC3 的平均值, 针对该平均值作图, 并找出相关性以预测在 CDS 集合中形成密码子使用偏好的突变力之间是否存在任何差异。GC3 对 GC12 在毒力基因集合中的显示相对较弱, 但具有显著的相关性 ($r=0.2749$, $p<0.1$)。上述发现表明, 形成全基因组和 VGS 的组成模式的元件对于所有密码子位置是相同的, 并以相似的方式作用于三个密码子位置。

2.1 全基因组和致病基因簇中的 GC 含量分布

从图 1 和图 2, 可以假定 GC 含量或多或少地均匀分布在全基因组和 VGS 的 CDS 中。因此, 可以预测, 全基因组和 VGS 可以具有相似种类的核苷酸组成以及可以具有相同的密码子使用模式。

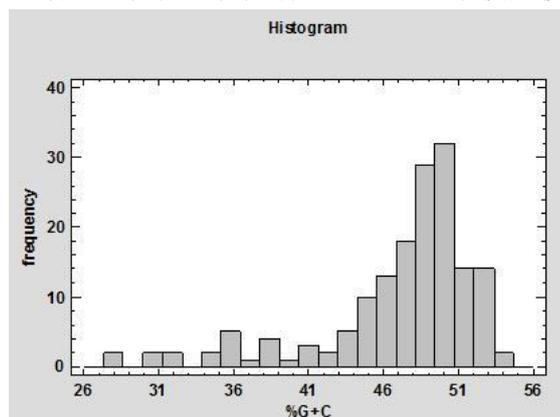


图 1 VGS 的基因中 GC 含量的分布

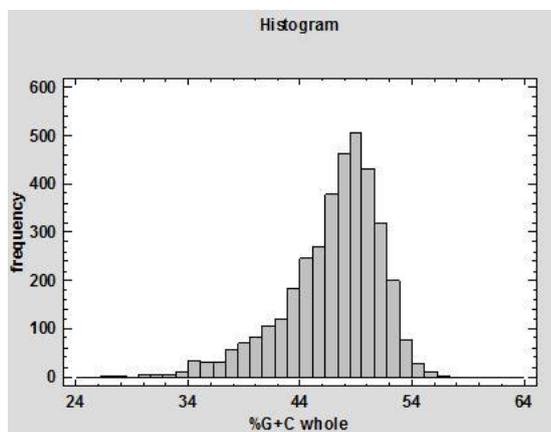


图 2 全基因组的基因中 GC 含量的分布

2.2 密码子偏好性程度

如果高度表达的基因显示出与基因组中其他基因不同的同义密码子分布, 则我们就需要烤炉是否与微生物体内的密码子使用的偏好性有关。提出了几种方法来评估生物体水平上密码子使用偏好的程度, 以确定生物体是否具有偏好性。NCdiff 是一种最广泛使用的判断生物体是否存在偏好性的方法。在这里, 我们使用该措施来评估霍乱弧菌 N16961 的密码子偏好性的程度。核糖体蛋白编码基因的平均 NC 值与基因组中其余基因的平均 NC 值的差异称为 Ncdiff, Ncdiff 的高值和低值分别表示生物体密码子偏好性程度的大小。

$$Ncdiff = (NC(all) - NC(rib)) / NC(all)$$

这种细菌的 Ncdiff 值非常低, 这意味着该细菌在其全基因组中显示出非常小的密码子偏好性。

2.3 致病基因簇和全基因组的 GC 含量

测定致病基因簇和全基因组基因组 (缺乏致病基因) 的总 GC 含量, 并且已经发现在这两个基因组之间存在均匀分布 GC, 这意味着致病基因簇不受突变压力以及其他因素的影响, 几乎没有水平基因转移到霍乱弧菌病原基因组的机会。全基因组和 VGS(致病簇)的总 GC 含量分别为 47% 和 48%, 这两个基因组之间的 GC 含量的变化是可忽略的。

3 结论

从该研究可以观察到, 在霍乱弧菌 N16961 的 VSG 和全基因组中, 密码子的使用模式没有任何差异, 且致病基因簇具有与全基因组的核苷酸成分分布的相同模式。选择力即突变压力以相似的方式在霍乱弧菌 N16961 的基因组中产生影响。这一发现也支持几乎不存在任何基因组密码子使用差异, 因为我们发现密码子使用偏差的程度在整个基因组中非常低, 这表明在这种致病菌中, 不同的功能分类基因之间可能存在极小的密码子使用差异的机会。

参考文献

- Blake P.A., and Olsvik Ø., eds., 1994, In *Vibrio cholerae and Cholera: Molecular to Global Perspectives*, ASM Press, pp.293-295
- Lawrence J.G., and Roth J.R., 1996, *Selfish Operons: Horizontal Transfer May Drive the Evolution of Gene*

- Clusters, *Genetics*, 143(4): 1843–1860
- Ochman H., and Moran N., 2001, Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis, *Science*, 292: 1096–9
- Liu Q.P., 2006, Analysis of codon usage pattern in the radioresistant bacterium *Deinococcus radiodurans*, *BioSystems*, 85(2): 99–106
- Read A.F., 1994, The evolution of virulence, *Trends Microbiol*, 2: 73–81
- Waldor M.K., Colwell R., and Mekalanos J.J., 1994, The *Vibrio cholerae* O139 serogroup antigen includes an O-antigen capsule and lipopolysaccharide virulence determinants. *Proc. Natl Acad. Sci. USA* 91, 91(24): 11388–11392
- Ziebuhr W., Ohlsen K., Karch H., Korhonen T., and Hacker J., 1999, Evolution of bacterial pathogenesis, *Cell Mol Life Sci*, 56:719–28
- Botzman M. and Margalit H., 2011, Variation in global codon usage bias among prokaryotic organisms is associated with their lifestyles, *Genome Biology*, 12: R109
- Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K., Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S., Yasunaga T., Honda T., Shinagawa H., Hattori M., and Iida T., 2003, Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholera*, *Lancet*. 1, 361(9359):743–9
- Jazel D. and Satchell K.J.F., 2013, Analysis of *Vibrio cholerae* Genome Sequences Reveals Unique *rtxA* Variants in Environmental Strains and an *rtxA*-Null Mutation in Recent Altered El Tor Isolates, *mBio*, 00624–12
- Blokesch M. and Schoolnik G.K., 2007, Serogroup conversion of *Vibrio cholerae* in aquatic reservoirs, *PLoS Pathog.*, 3: e81.
- Safa A., Nair G.B., and Kong R.Y., 2010, Evolution of new variants of *Vibrio cholerae* O1. *Trends Microbiol.*, 18: 46–54
- Siddique A.K., Nair G.B., Alam M., Sack D.A., Huq A., Nizam A., Longini IM Jr., Qadri F., Faruque S.M., Colwell R.R., Ahmed S., Iqbal A., Bhuiyan N.A. and Sack R.B., 2010, El Tor cholera with severe disease: a new threat to Asia and beyond, *Epidemiol. Infect.*, 138: 347–352