

埃可病毒 18 型构象表位的生物信息学研究

王 惠^{1,2}, 贾永涛^{1,2}, 方雨露^{1,2}, 董长征^{1,2*}

(1. 宁波大学 医学院, 浙江 宁波 315211; 2. 浙江省病理生理学技术研究重点实验室, 浙江 宁波 315211)

摘要: 埃可病毒 18 型(Echovirus 18, E18)可引起手足口病、病毒性脑膜炎和急性胃肠炎等疾病, 近 20 年来开始在全球流行. 本文首先使用前期研发的人肠道病毒构象表位的生物信息学预测算法, 对肠道病毒 B 种的 E18 构象表位进行系统性预测, 发现其表位分布模式呈三簇分布, 其中 VP1 BC 环和 C-端、VP2 EF 环是 E18 表位的主要构成区域; 其次基于 E18 的 VP1 和基因组序列的分子进化分析发现, 进化分支的形成伴随表位处氨基酸突变, 而 VP1 C-端和 VP2 EF 环是突变热点区域; 最后通过足迹图发现, 大多数 E18 衣壳表面的分支突变都位于表位上或附近, 受体结合位置处的突变罕见, 提示表位处的突变产生了新的进化分支, 但分支突变尽量避免受体结合位置, 以免影响病毒与宿主细胞的结合能力. E18 构象表位的生物信息学研究, 为进一步通过实验鉴定表位、病毒的监测和预警以及抗病毒药物和疫苗的研发提供了重要支持.

关键词: 埃可病毒 18 型; 肠道病毒; 构象表位; 抗原表位; 分子进化分析; 生物信息学

中图分类号: R373.2; R511

文献标志码: A

文章编号: 1001-5132 (2023) 01-0075-09

人肠道病毒属小 RNA 病毒科(Picornaviridae), 肠道病毒属(*Enterovirus*), 包括肠道病毒 A 种(EV-A)、B 种(EV-B)、C 种(EV-C)和 D 种(EV-D). EV-A 代表性血清型包括肠道病毒 A 组 71 型(*Enterovirus* A71, EV-A71)和柯萨奇病毒 A 组 16 型(*Coxsackievirus* A16, CV-A16)等; EV-B 代表性血清型包括埃可病毒 18 型(*Echovirus* 18, E18)等; EV-C 代表性血清型包括脊髓灰质炎病毒 1 型(*Poliovirus* 1, PV1)、PV2 和 PV3 等; EV-D 代表性血清型为 EV-D68. E18 可引起手足口病、病毒性脑膜炎和急性胃肠炎等疾病^[1-4], 通常感染后症状较为轻微, 但严重时也会危及生命^[5]. 自 1955 年在美国首次分离病毒株后, E18 感染鲜有报道, 直到 21 世纪 E18 开始在全球流行, 并且引起多起较大规模的疫情^[6-9]. 2014—2016 年美国肠道病毒监测报告显示, E18 在最常引起疫情的肠道病毒中排第 4 位^[10]. Chen 等^[11]在 2015—2016 年从我国 6 个省份(山东、河北、山西、黑龙江、江苏和云南)的手足口病、病毒性脑炎和病毒性脑膜炎病例中分离出 34 株

E18, 测定其结构蛋白(Viral Protein, VP)VP1 序列, 并完成了其中 6 株的全基因组测序. 2015—2020 年, E18 在我国广东^[4,12]、河北^[13]、山东^[14]和云南^[15]等省大量检出, 提示我国存在 E18 暴发风险.

与其他肠道病毒相似, E18 有一个长度约为 7.4 kb 的单股正链 RNA 基因组, 其病毒衣壳由 60 个亚单位构成^[16-18]. 每个非对称亚单位由 VP1~VP4 构成, 其中 VP4 位于衣壳内表面, VP1~VP3 位于衣壳外表面, 后者是病毒构象表位的所在区域. 抗原表位的确定对掌握病毒的致病机制^[19-20]、监测病毒的变异和进化^[21]、研发抗病毒药物^[22-23]和疫苗^[24-25]都有重要作用. 然而, 目前尚无针对 E18 表位的研究报道.

本研究首先利用前期发展的生物信息学算法^[17-18]对 E18 构象表位进行系统性预测, 分析表位的分布规律; 然后基于 Nextstrain 平台构建 E18 的分子进化树, 确定进化分支和决定进化分支的标志性氨基酸突变(分支突变); 最后探索分支突变与表位以及病毒受体结合位置的关系, 分析表位在

收稿日期: 2022-08-08.

宁波大学学报(理工版)网址: <http://journal.lg.nbu.edu.cn/>

基金项目: 浙江省基础公益研究计划项目(LGF18C060001); 宁波市公益项目(2021S135); 宁波大学研究生科研创新基金(IF2022190).

第一作者: 王惠(1996-), 女, 山西大同人, 在读硕士研究生, 主要研究方向: 肠道病毒生物信息学. E-mail: wang_2091@163.com

*通信作者: 董长征(1978-), 男, 浙江丽水人, 博士/副教授, 主要研究方向: 生物信息学. E-mail: dongchangzheng@nbu.edu.cn

E18 分子进化过程中的作用. 研究结果可为 E18 的监测和预警提供参考.

1 材料与方法

1.1 E18 结构蛋白的序列和结构特征分析

从 RCSB PDB 数据库^[26](<https://www.rcsb.org>)和 NCBI Nucleotide 数据库^[27](<https://www.ncbi.nlm.nih.gov/nucleotide>)中分别下载 E18 天然成熟颗粒的结构蛋白的三维结构文件(PDB ID: 6HBG^[16]; 病毒株名: Metcalf, 为简便以下用 PDB ID 指代病毒株)和其氨基酸序列(Accession ID: AAL37163). 将 PDB 文件和氨基酸序列导入在线工具 ESPript 3.0^[28](<http://esprict.ibcp.fr>)中注释二级结构信息. 使用 PyMOL^[29]绘制病毒衣壳的表面结构图. 所有软件和在线工具都采用默认参数.

1.2 E18 构象表位的生物信息学预测

实验室前期在 Borley 算法^[30]基础上发展了人肠道病毒构象表位的生物信息学预测算法, 并成功地应用于肠道病毒 A 种(EV-A71、CV-A16^[17])和 D 种(EV-D68^[31])的表位预测. 预测算法主要包括以下 3 步(详见文献[17-18]):

(1)在 PDB 文件中删除 VP4 并生成复合链(图

1), 将复合链整体看作一个蛋白质, 代替单个结构蛋白进行表位预测.

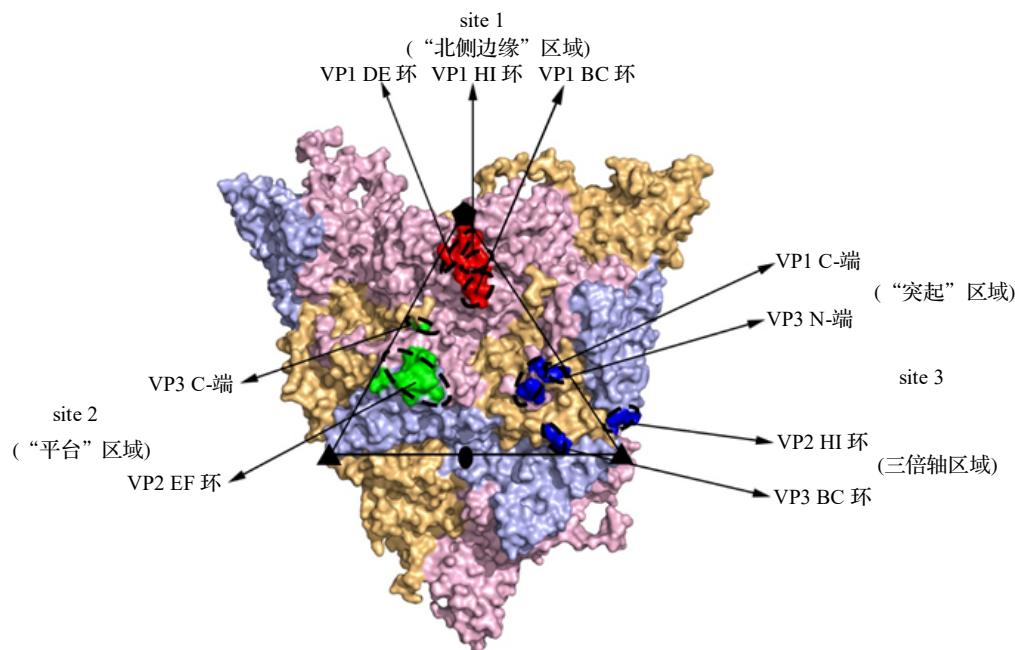
(2)利用 3 个表位预测工具 Epitopia^[32](<http://epitopia.tau.ac.il>)、Ellipro^[33](<http://tools.iedb.org/ellipro>)以及 DiscoTope^[34](<http://www.cbs.dtu.dk/services/DiscoTope>)分别预测复合链的表位. 三者的阈值均使用默认参数(0.174、0.3、-10.7). 采用投票法, 将同时被 3 种工具预测为表位的氨基酸残基作为一致性表位.

(3)提取中心链 Chain 1(一个由 VP1~VP3 组成的亚单位)上的一致性表位, 并筛选处于病毒衣壳相对暴露面的残基(其 C_{α} 到衣壳中心的距离超过所有 C_{α} 到衣壳中心的平均距离), 获得最终的预测结果.

1.3 基于 Nextstrain 的 E18 分子进化分析

Nextstrain^[35]是免费的在线(<https://nextstrain.org>)或本地开源平台, 针对病毒基因/基因组数据, 通过强大的分子进化分析和数据可视化功能, 追踪病毒的分子进化和时空传播路线. 本研究通过本地化部署 Nextstrain 平台, 利用 Nextstrain 平台的 augur 病原体生物信息分析包, 分别对 E18 的 VP1 和全基因组序列进行分子进化分析, 步骤如下:

(1)从 VIPR 病毒数据库^[36](<https://www.viprbrc>.



黑色三角形框为中心链, 外围有 3 个三角形形状的周围链, 每条链均为由 VP1~VP3 组成的非对称亚单位. 五角形、三角形和椭圆形分别代表五倍轴、三倍轴和二倍轴. 构象表位聚集成三簇, 分别用 site 1、site 2 和 site 3 表示.

图 1 E18 复合链和衣壳表面结构

org/brc)和 NCBI Nucleotide 数据库^[27](<https://www.ncbi.nlm.nih.gov/nucleotide>)中下载所有长度>700 bp 的 E18 VP1 序列和长度>7 000 bp 的基因组序列, 剔除信息不全序列(无病毒毒株分离国家或时间信息等). 采用 MAFFT 7^[37]在线平台(<https://mafft.cbrc.jp/alignment/server>)进行多重序列比对, 人工剔除异常序列(例如多重序列比对后通过人工核查发现存在蛋白质翻译错误, 序列中存在超过 3 个模糊碱基). 最终获得 285 条 E18 的 VP1 序列和 63 条基因组序列.

(2)利用 augur 病原体生物信息分析包进行生物信息学和分子进化分析. align 命令调用内嵌的 MAFFT 软件^[37]进行多重序列比对, tree 命令调用 IQ-tree 工具^[38]构建分子进化树(最大似然法), refine 命令调用 TreeTime 工具^[39]构建基于时间尺度的分子进化树并对进化树进行优化, clade 命令标注进化分支(clade)和确定分支突变. 由于表位仅分布在 VP1~VP3 上, 因此只标注了 VP1~VP3 上的分支突变. 最后利用 Nextstrain 平台的 auspice 工具进行基于 JavaScript 技术的网页互动可视化展示.

(3)将获得的分支突变与表位、病毒受体结合位置利用 RIVEM 工具^[40-41]绘制在 E18 衣壳表面结构图上, 即足迹图(roadmap), 分析分支突变与表位以及受体结合之间的关系. 新生儿 Fc 受体 (Neonatal Fc Receptor, FcRn)是埃可病毒共用的脱衣壳受体, 埃可病毒借助 FcRn 感染宿主细胞并释放病毒基因组. 虽然 E18 的结合位置尚未确定, 但 E6 和 E30 与 FcRn 的结合位置已经确定^[19-20], 可以通过序列比对的方式确定 E18 的模拟结合位置.

2 结果

2.1 E18 的结构蛋白特征和构象表位预测结果

E18 病毒颗粒 6HBG 的结构蛋白 VP1、VP2 和 VP3 的长度分别为 287、260 和 239 个氨基酸残基, 由 8 条反向平行的 β 链(β B- β I)构成 β 桶(β -barrel), 链之间的部分为环区(loop), 两端为 N-端(N-terminus)和 C-端(C-terminus), 如图 1 和图 2 所标注. E18 的衣壳表面结构如图 1 所示, 衣壳表面有峡谷(canyon)、峡谷两侧的“边缘”(rim)、“平台”(puff)和“突起”(knob)等结构特征以及五倍轴、三倍轴和二倍轴等三维结构标记.

E18 构象表位预测结果见表 1 及图 1 和图 2. E18 共有 27 个氨基酸残基预测为表位, 分布在 VP1 (BC 环、DE 环、HI 环和 C-端)、VP2 (EF 环和 HI 环)和 VP3 (N-端 knob 区域、BC 环和 C-端). 与 EV-A^[17]和 EV-D^[31]相似, E18 的构象表位也聚集成三簇(表 1 和图 1): site 1、site 2 和 site 3, 分别位于峡谷的“北侧边缘”区域、峡谷南侧的“平台”区域、峡谷南侧的“突起”区域和三倍轴区域. 其中 VP1 BC 环和 C-端、VP2 EF 环是 E18 表位的主要构成区域.

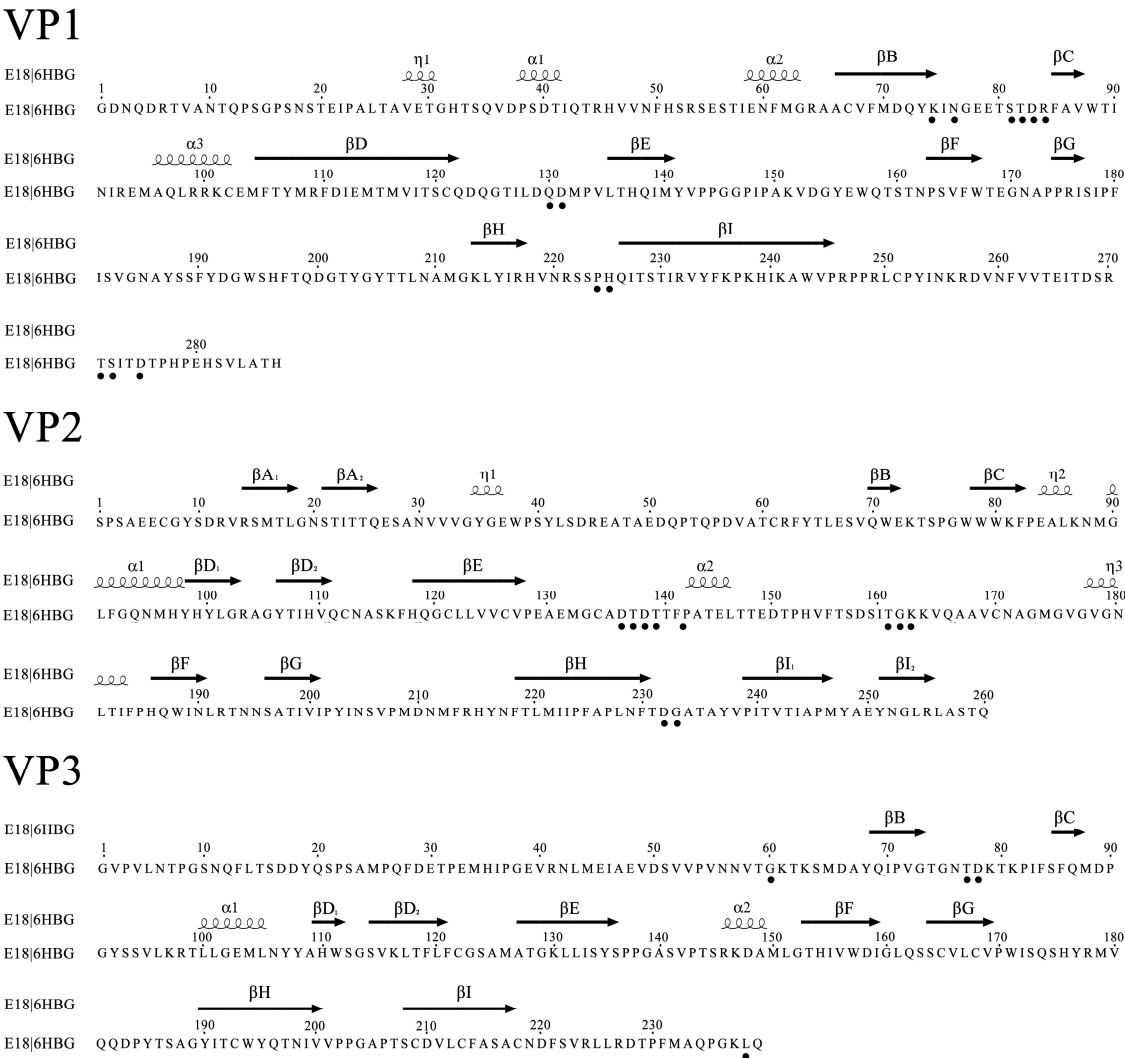
2.2 E18 的分子进化分析

利用 Nextstrain 平台的 augur 病原体生物信息分析包, 分别构建了基于 E18 基因组和 VP1 序列的时间尺度的分子进化树, 分别简称为基因组进化树(图 3(a))和 VP1 进化树(图 3(b)), 两者具有一致的拓扑结构, 都分为 A、B、C 三个进化分支, C 分支又分为 C1 和 C2 两个子分支, 原株 Metcalf 不属于任何一个分支. 以序列数较多的 VP1 进化树为例, 大约 1946 年 Metcalf 从 A、B、C 三个分支的共同祖先 A~C 中分歧出来, 并于 1955 年在美国首次分离. 1970 年分子进化树首先分歧出 A 分支; 1979 年又分歧出 B 和 C 分支; 最后在 1989 年 C 分支分歧成 C1 和 C2 两个子分支. C 分支尤其是 C2 分支是目前流行的病毒分支. A 分支由 2 株来自中国的病毒毒株构成. B 分支有 1 株来自埃塞俄比亚, 其余都来自印度. C1 分支主要来自法国、德国、俄罗斯、瑞典和澳大利亚. C2 分支全球分布广泛, 其中分离病毒毒株数最多的国家为中国(168 株)、法国(14 株)、澳大利亚(14 株)、日本(13 株)和美国(7 株).

每个进化分支都有若干个决定进化分支的分支突变. 以包含 VP1~VP3 的基因组进化树为例(图 3(a)), A 分支的分支突变为 VP1 C-端的 T271A 和 D275E, B 分支的分支突变为 VP1 C-端的 K271S、A285V 和 T286S. C1 的分支突变包括 VP1 C-端的 D275E 以及 R6K、VP2 EF 环的 S159P、VP3 的 V2I 和 H182N. C2 的分支突变包括 VP1 C-端的 G257S、A262V、A285V 以及 I42L 和 I92V, VP2 的 T74S, VP3 knob 的 V58I 以及 N11T. 从图 3(a)可以发现, 分子进化树上多数(61.2%)的分支突变都位于表位处, 即分支进化伴随着表位处的氨基酸突变, 尤其

是 VP1 C-端是突变热点区域. 此外, 还有 VP2 的 EF 环. VP1 进化树(图 3(b))也表现为高度一致, 绝大多数分支突变在两种进化树上分布一致, 但由

于序列数目差异较大(285 比 63), 个别分支突变产生了差异, 例如 VP1 进化树 B~C 分支上的突变 D129E 成为基因组进化树 C 分支上的突变.



→β 代表片层, § 代表螺旋, 箭头之间的区域代表环区, • 为构象表位标注.

图 2 E18 结构蛋白的一级结构和二级结构注释

表 1 E18 构象表位的预测结果

表位名称	结构特征	二级结构	氨基酸残基
site 1	“北侧边缘”区域	VP1 BC 环	74K、76N、81S、82T、83D、84R
		VP1 DE 环	130Q、131D
		VP1 HI 环	224P、225H
site 2	“平台”区域	VP2 EF 环	136D、137T、138D、139T、142P、161T、162G、163K
		VP3 C-端	238L
site 3	“突起”区域	VP1 C-端	271T、272S、275D
		VP3 N-端	60G
	三倍轴区域	VP2 HI 环	232D、233G
		VP3 BC 环	77T、78D

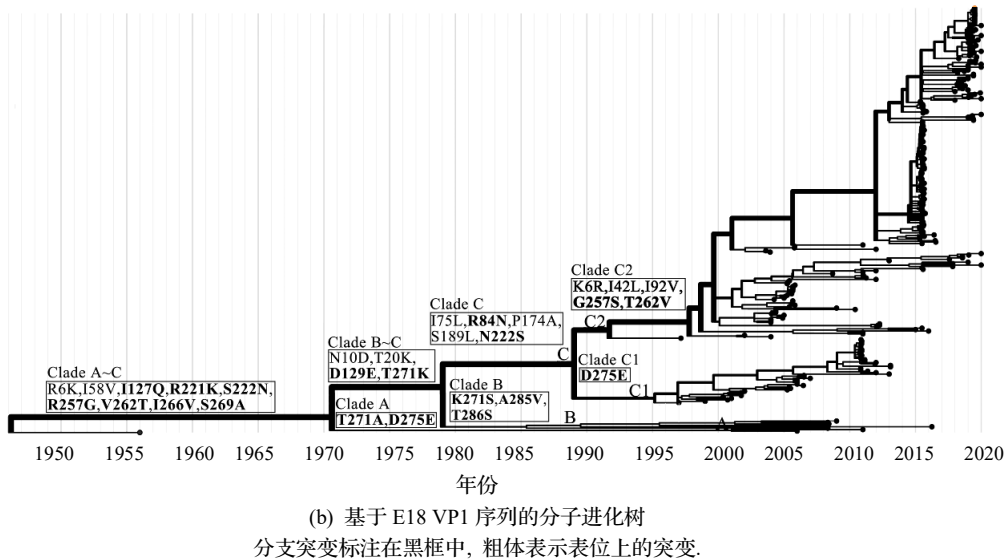
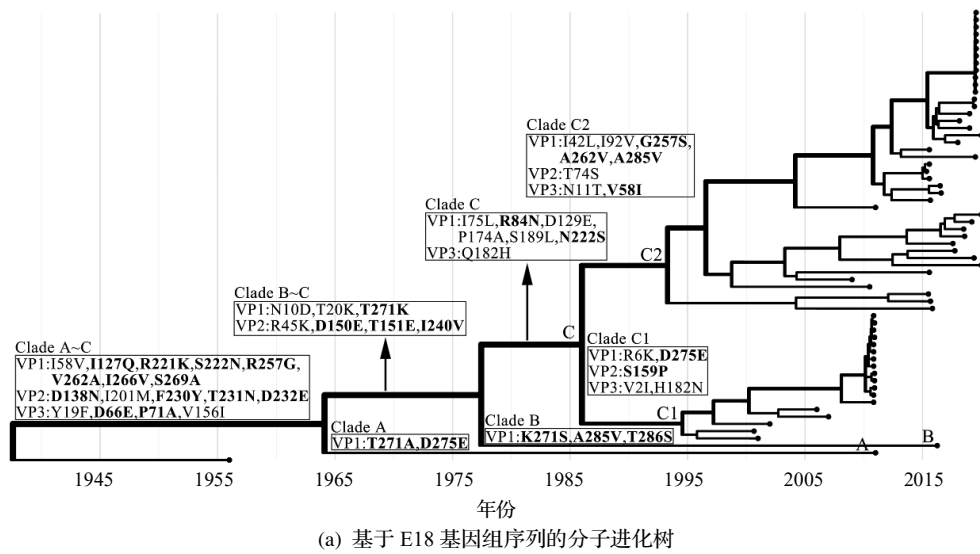


图 3 E18 的分子进化树

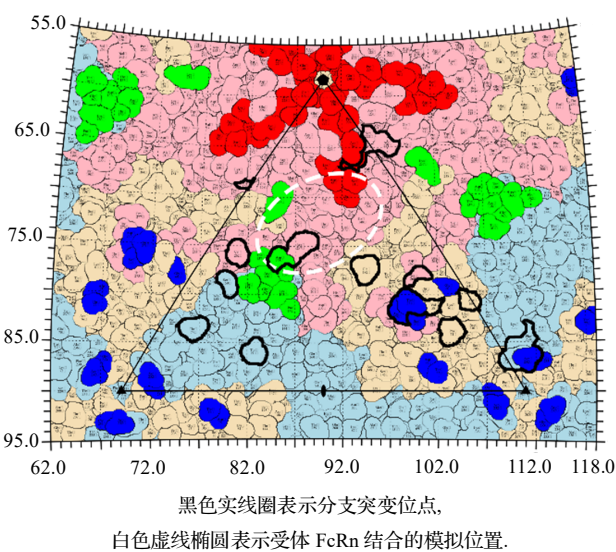


图 4 E18 足迹图

利用足迹图(图 4)将上述分支突变与表位、受体结合位置都标注在 E18 的衣壳表面结构图上, 分析三者之间的关系可以发现, 大多数 E18 衣壳表面的分支突变(黑色实线图)都位于表位上或附近, 而且表位的 3 个 site 均被突变覆盖; 受体 FcRn 模拟结合位置(白色虚线椭圆)仅有 2 个突变。这提示表位处的突变产生了新的进化分支, 但尽量避开受体结合位置, 以免影响病毒与宿主细胞的结合能力。这两个突变位点对受体结合能力的影响需要进一步实验研究。

3 讨论

病毒构象表位测定的金标准是冷冻电镜技术^[19-20,23], 但冷冻电镜技术门槛高, 现阶段难以常

规应用. 免疫逃避实验是传统测定构象表位的常用方法^[42-43], 但只能通过突变体确定部分表位, 同时也费时费力. 生物信息学算法具有高通量和相对准确的特点, 能够为实验性研究提供候选靶标, 辅助实验性研究. 传统的生物信息学表位预测算法具有普适性, 但如果直接用来预测肠道病毒的结构蛋白的表位, 由于未考虑“结构蛋白嵌在衣壳中”这个生物学结构特征, 预测结果必然有大量的假阳性^[17,30,44]. 本研究使用实验室前期发展的表位预测算法, 其最大优势在于将生物信息学算法和肠道病毒的衣壳结构特征密切结合, 大大提高了算法的可靠性^[17]. 在不同肠道病毒种型(EV-A 和 EV-D)表位预测的实际应用中, 通过对比实验结果可以确认算法具有较高的准确性^[17,31]. 当然, 生物信息学算法不可避免地存在自身的缺陷和限制. 例如, 算法仅能对已测定三维结构的肠道病毒进行表位预测; 算法的精确性受到三维结构精确性的严重影响; 算法的预测原理基于表位的病毒学假设(衣壳表面突出区域更易与抗体或受体结合), 而这种假设需要在实践中通过实验性研究不断验证和改进, 算法也随之更新.

作为 EV-B 代表性血清型之一的 E18, 它的构象表位与 EV-A(EV-A71 和 CV-A16^[17])和 EV-D(EV-D68^[31])一样呈现三簇分布规律. VP1 BC 环和 C-端、VP2 EF 环是肠道病毒共有的表位, VP1 GH 环是 EV-A 特有的表位, VP1 DE 环是 EV-D 特有的表位, E18 则未发现特异性表位. 即使表位分布高度相似, 但表位上的氨基酸突变决定了不同血清型肠道病毒的抗原性差异^[17,45]. 与 EV-D68^[31]相似, VP1 C-端和 VP2 EF 环是 E18 的突变热点区域, 每一个 E18 进化分支的形成都伴随着氨基酸突变. 目前尚未能见中和 E18 的单克隆抗体(单抗)的相关报道, 但同属 EV-B 的 E30 已报道有两个单抗 4B10 和 6C5^[46]. 单抗 4B10 主要结合 E30 峡谷南侧的 VP1 C-端和 VP2 EF 环, 单抗 6C5 结合位点包括峡谷北侧的 VP1 BC 环、DE 环、EF 环和 HI 环, 而这些位点大多数都是 E18 的表位区域, 提示这些表位区域易与单抗结合.

本研究分别基于 E18 的 VP1 和基因组序列构建了时间尺度的分子进化树. VP1 进化树包含的序列数较多(285 条), 能够更加准确地确定进化关系

和推断分歧时间, 但 VP1 进化树仅能获得 VP1 上的分支突变. 基因组进化树虽然序列数较少(63 条), 但能获得 VP1~VP3 上的分支突变. 基因组进化树和 VP1 进化树具有高度一致的拓扑结构和 VP1 上的分支突变, 一方面证明本研究确定的分支突变高度可靠, 另一方面基因组进化树补充了 VP1 进化树缺乏的 VP2 和 VP3 上的分支突变, 尤其是 VP2 EF 环和 VP3 knob 这两个重要表位区域. 参照 Chen 等^[11,13]研究结果, 将 E18 分为 A、B、C 三个主要分支, 原株 Metcalf 未划入任何分支. 相较于 A 分支和 B 分支, C 分支是 E18 的主要流行分支, 96.8% 分离的病毒毒株都属于 C 分支. C 分支分为 C1 分支和 C2 分支, 其中 C1 分支在 2013 年后未再流行, 而 C2 分支病毒毒株广泛分离于近年来的病毒性脑膜炎和手足口病疫情中, 提示 C2 分支可能具有较强的传染力和毒力, 但尚未有 E18 传染力和毒力的进一步研究报道. 分子进化树(图 3)显示, E18 进化分支的形成伴随着表位处的氨基酸突变, 尤其是 VP1 C-端和 VP2 EF 环是突变热点区域. 之前的研究发现^[4-5,11]多个 VP1 上的多个重要突变位点, 包括 R6K、N10D、R84N、M104L、Y215F、I216V、V262T/A 和 D275E 等. VP1 BC 环是 E18 重要的表位区域(图 1 和图 2). 有研究发现^[17,31], VP1 BC 环是 EV-A 和 EV-D 共有的重要表位. BC 环上的 R84N 是 C 分支的分支突变, A 和 B 分支均为 R, 几乎所有 C 分支均为 N(2 株为 S), 进一步验证了表位 VP1 BC 环对肠道病毒的重要性. K6R 和 T262V 是 C2 的分支突变, D275E 则是 C1 的分支突变(图 3). 从 Nextstrain 平台构建的 VP1 进化树上可以清晰看到, D10N、Y215F、I216V、S257G 和 H287R 都是 C2 分支中国株的标志性氨基酸突变, M104L 则是其中一大簇病毒毒株上的标志性突变. E18 的分子进化分析, 一方面从另一个角度间接地验证了表位预测的准确性, 另一方面也提示表位处的重要氨基酸突变及相应病毒毒株的时空传播路线是肠道病毒监测和预警分析的重点.

对于流感病毒和冠状病毒, 由于接种疫苗和广泛感染产生群体免疫, 使得病毒处于正选择进化压力下, 因此不仅表位突变频繁, 而且受体结合区域也存在广泛突变, 从而产生免疫逃避^[47-48]. 而对于肠道病毒, 由于新生儿不断补充易感群体^[49],

使得多数肠道病毒并没有面临显著的进化压力, 因此相对突变速率要慢一些(E18 的 VP1 约为 5.6×10^{-3} 替换·(位点·a)⁻¹, 与其他肠道病毒相似), 在表位区域存在较弱的正选择压力, 其他区域则偏向中性进化^[50]. E18 受体结合区域也更加保守, 其足迹图提供了佐证. 但随着 E18 的广泛流行, 人群血清中普遍存在中和抗体, 病毒面临越来越强的正选择压力, 表位也会突变得更加频繁, 需要密切监测其抗原性是否会发生较大改变, 出现类似 EV-A71 那样的大规模疫情^[51].

生物信息学在流感病毒的优势病毒毒株预测方面取得了巨大成功^[52-54], 为流感疫苗的研发和准备提供了重要的技术支撑. 对肠道病毒表位生物信息学的研究, 可为进一步通过实验鉴定构象表位、病毒的监测和预警以及抗病毒药物和疫苗的研发提供重要支持.

参考文献:

- [1] Graf J, Hartmann C J, Lehmann H C, et al. Meningitis gone viral: Description of the echovirus wave 2013 in Germany[J]. BMC Infectious Diseases, 2019, 19(1):1010.
- [2] Tsai H P, Huang S W, Wu F L, et al. An echovirus 18-associated outbreak of aseptic meningitis in Taiwan: Epidemiology and diagnostic and genetic aspects[J]. Journal of Medical Microbiology, 2011, 60:1360-1365.
- [3] Zhang H H, Zhao Y L, Liu H B, et al. Molecular characterization of two novel echovirus 18 recombinants associated with hand-foot-mouth disease[J]. Scientific Reports, 2017, 7:8448.
- [4] 周健明, 谢显清, 方冬, 等. 引起一起暴发性急性胃肠炎疫情的埃可病毒 18 型 VP1 基因特征分析[J]. 病毒学报, 2021, 37(1):133-139.
- [5] Krumbholz A, Egerer R, Braun H, et al. Analysis of an echovirus 18 outbreak in Thuringia, Germany: Insights into the molecular epidemiology and evolution of several enterovirus species B members[J]. Medical Microbiology and Immunology, 2016, 205(5):471-483.
- [6] Kelly H, Brussen K A, Lawrence A, et al. Polioviruses and other enteroviruses isolated from faecal samples of patients with acute flaccid paralysis in Australia, 1996–2004[J]. Journal of Paediatrics and Child Health, 2006, 42(6):370-376.
- [7] Othman I, Mirand A, Slama I, et al. Enterovirus migration patterns between France and Tunisia[J]. PLoS One, 2015, 10(12):e0145674.
- [8] Kusuhara K, Saito M, Sasaki Y, et al. An echovirus type 18 outbreak in a neonatal intensive care unit[J]. European Journal of Pediatrics, 2008, 167(5):587-589.
- [9] Bubba L, Broberg E K, Jasir A, et al. Circulation of non-polio enteroviruses in 24 EU and EEA countries between 2015 and 2017: A retrospective surveillance study[J]. The Lancet Infectious Diseases, 2020, 20(3): 350-361.
- [10] Abedi G R, Watson J T, Nix W A, et al. Enterovirus and Parechovirus surveillance – United States, 2014–2016[J]. Morbidity and Mortality Weekly Report, 2018, 67(18): 515-518.
- [11] Chen X P, Ji T J, Guo J Y, et al. Molecular epidemiology of echovirus 18 circulating in mainland China from 2015 to 2016[J]. Virologica Sinica, 2019, 34(1):50-58.
- [12] 尚永朋, 赵宇曦, 邓名贵, 等. 宏基因组二代测序技术辅助诊断新生儿埃可病毒 18 型聚集性感染[J]. 中国感染控制杂志, 2022, 21(4):317-322.
- [13] Chen X P, Li J J, Guo J Y, et al. An outbreak of echovirus 18 encephalitis/meningitis in children in Hebei Province, China, 2015[J]. Emerging Microbes & Infections, 2017, 6(6):e54.
- [14] Wang J, Meng M, Xu H, et al. Analysis of echovirus genotypes in the cerebrospinal fluid of children associated with aseptic meningitis in Liaocheng, China, from 2018 to 2019[J]. BMC Infectious Diseases, 2021, 21(1):1-9.
- [15] 张名, 许丹菡, 冯昌增, 等. 2019 年云南省埃可病毒 18 型分离株的全基因组分析[J]. 中国病原生物学杂志, 2021, 16(8):888-892.
- [16] Buchta D, Füzik T, Hrebík D, et al. Enterovirus particles expel capsid pentamers to enable genome release[J]. Nature Communications, 2019, 10:1138.
- [17] Wang L P, Zhu M, Fang Y L, et al. Bioinformatics-based prediction of conformational epitopes for enterovirus A71 and coxsackievirus A16[J]. Scientific Reports, 2021, 11:5701.
- [18] 祝苗. 人肠道病毒 71 型和柯萨奇病毒 A16 型抗原表位的生物信息学预测[D]. 宁波: 宁波大学, 2018.
- [19] Wang K, Zhu L, Sun Y, et al. Structures of echovirus 30 in complex with its receptors inform a rational prediction for enterovirus receptor usage[J]. Nature Communications, 2020, 11:4421.
- [20] Zhao X, Zhang G G, Liu S, et al. Human neonatal Fc receptor is the cellular uncoating receptor for enterovirus B[J]. Cell, 2019, 177(6):1553-1565.
- [21] Huang S W, Tai C H, Fonville J M, et al. Mapping

- enterovirus A71 antigenic determinants from viral evolution[J]. *Journal of Virology*, 2015, 89(22):11500-11506.
- [22] Meng B, Lan K K, Xie J, et al. Inhibitory antibodies identify unique sites of therapeutic vulnerability in rhinovirus and other enteroviruses[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2020, 117(24):13499-13508.
- [23] Xu L F, Zheng Q B, Zhu R, et al. Cryo-EM structures reveal the molecular basis of receptor-initiated coxsackievirus uncoating[J]. *Cell Host & Microbe*, 2021, 29(3):448-462.
- [24] Xu L F, He D L, Li Z Q, et al. Protection against lethal enterovirus 71 challenge in mice by a recombinant vaccine candidate containing a broadly cross-neutralizing epitope within the VP2 EF loop[J]. *Theranostics*, 2014, 4(5):498-513.
- [25] Anasir M I, Poh C L. Structural vaccinology for viral vaccine design[J]. *Frontiers in Microbiology*, 2019, 10: 738.
- [26] Burley S K, Berman H M, Bhikadiya C, et al. RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy[J]. *Nucleic Acids Research*, 2019, 47(D1):D464-D474.
- [27] Clark K, Karsch-Mizrachi I, Lipman D J, et al. GenBank[J]. *Nucleic Acids Research*, 2016, 44(D1): D67-D72.
- [28] Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server[J]. *Nucleic Acids Research*, 2014, 42:W320-W324.
- [29] Rigsby R E, Parker A B. Using the PyMOL application to reinforce visual understanding of protein structure[J]. *Biochemistry and Molecular Biology Education*, 2016, 44(5):433-437.
- [30] Borley D W, Mahapatra M, Paton D J, et al. Evaluation and use of in-silico structure-based epitope prediction with foot-and-mouth disease virus[J]. *PLoS One*, 2013, 8(5):e61122.
- [31] Fang Y L, Chen Q, Wang H, et al. The role of conformational epitopes in the evolutionary divergence of enterovirus D68 clades: A bioinformatics-based study[J]. *Infection, Genetics and Evolution*, 2021, 93:104992.
- [32] Rubinstein N D, Mayrose I, Martz E, et al. Eptopia: A web-server for predicting B-cell epitopes[J]. *BMC Bioinformatics*, 2009, 10:287.
- [33] Ponomarenko J, Bui H H, Li W, et al. ElliPro: A new structure-based tool for the prediction of antibody epitopes[J]. *BMC Bioinformatics*, 2008, 9:514.
- [34] Kringelum J V, Lundegaard C, Lund O, et al. Reliable B cell epitope predictions: Impacts of method development and improved benchmarking[J]. *PLoS Computational Biology*, 2012, 8(12):e1002829.
- [35] Hadfield J, Megill C, Bell S M, et al. Nextstrain: Real-time tracking of pathogen evolution[J]. *Bioinformatics*, 2018, 34(23):4121-4123.
- [36] Pickett B E, Sadat E L, Zhang Y, et al. ViPR: An open bioinformatics database and analysis resource for virology research[J]. *Nucleic Acids Research*, 2012, 40: D593-D598.
- [37] Katoh K, Standley D M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability[J]. *Molecular Biology and Evolution*, 2013, 30(4):772-780.
- [38] Nguyen L T, Schmidt H A, von Haeseler A, et al. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies[J]. *Molecular Biology and Evolution*, 2015, 32(1):268-274.
- [39] Sagulenko P, Puller V, Neher R A. TreeTime: Maximum-likelihood phylodynamic analysis[J]. *Virus Evolution*, 2018, 4(1):vex042.
- [40] Xiao C, Rossmann M G. Interpretation of electron density with stereographic roadmap projections[J]. *Journal of Structural Biology*, 2007, 158(2):182-187.
- [41] 高柳莺, 戎浩, 祝苗, 等. 基于 RIVEM 的人肠道病毒衣壳表面结构绘制流程[J]. *病毒学报*, 2018, 34(5): 519-527.
- [42] Kiener T K, Jia Q, Meng T, et al. A novel universal neutralizing monoclonal antibody against enterovirus 71 that targets the highly conserved knob region of VP3 protein[J]. *PLoS Neglected Tropical Diseases*, 2014, 8(5): e2895.
- [43] Buttinelli G, Donati V, Ruggeri F M, et al. Antigenic sites of coxsackie A9 virus inducing neutralizing monoclonal antibodies protective in mice[J]. *Virology*, 2003, 312(1): 74-83.
- [44] Yao B, Zheng D D, Liang S D, et al. Conformational B-cell epitope prediction on antigen protein structures: A review of current algorithms and comparison with common binding site prediction methods[J]. *PLoS One*, 2013, 8(4):e62249.
- [45] Rong H, Wang L P, Gao L Y, et al. Bioinformatics-based prediction of conformational epitopes for human parechovirus[J]. *PLoS One*, 2021, 16(4):e0247423.

- [46] Wang K, Zheng B Y, Zhang L, et al. Serotype specific epitopes identified by neutralizing antibodies underpin immunogenic differences in enterovirus B[J]. *Nature Communications*, 2020, 11:4419.
- [47] Kistler K E, Bedford T. Evidence for adaptive evolution in the receptor-binding domain of seasonal coronaviruses OC43 and 229e[J]. *eLife*, 2021, 10:e64509.
- [48] Koel B F, Burke D F, Bestebroer T M, et al. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution[J]. *Science*, 2013, 342(6161):976-979.
- [49] Pons-Salort M, Grassly N C. Serotype-specific immunity explains the incidence of diseases caused by human enteroviruses[J]. *Science*, 2018, 361(6404):800-803.
- [50] Tee K K, Lam T T Y, Chan Y F, et al. Evolutionary genetics of human enterovirus 71: Origin, population dynamics, natural selection, and seasonal periodicity of the VP1 gene[J]. *Journal of Virology*, 2010, 84(7):3339-3350.
- [51] Zhang Y, Zhu Z, Yang W Z, et al. An emerging recombinant human enterovirus 71 responsible for the 2008 outbreak of hand foot and mouth disease in Fuyang city of China[J]. *Virology Journal*, 2010, 7(1):94.
- [52] Łuksza M, Lässig M. A predictive fitness model for influenza[J]. *Nature*, 2014, 507(7490):57-61.
- [53] Huddleston J, Barnes J R, Rowe T, et al. Integrating genotypes and phenotypes improves long-term forecasts of seasonal influenza A/H3N2 evolution[J]. *eLife*, 2020, 9:e60067.
- [54] Du X J, Dong L B, Lan Y, et al. Mapping of H3N2 influenza antigenic evolution in China reveals a strategy for vaccine strain recommendation[J]. *Nature Communications*, 2012, 3:709.

Bioinformatics analysis of conformational epitopes for echovirus 18

WANG Hui^{1,2}, JIA Yongtao^{1,2}, FANG Yulu^{1,2}, DONG Changzheng^{1,2*}

(1.School of Medicine, Ningbo University, Ningbo 315211, China;

2.Zhejiang Provincial Key Laboratory of Pathological and Physiological Technology, Ningbo 315211, China)

Abstract: Echovirus 18 (E18) causes aseptic meningitis (AM), hand-foot-mouth disease (HFMD), and acute gastroenteritis, which has been prevalent worldwide for the last two decades. In the present study, a bioinformatics-based algorithm of conformational epitopes for human enteroviruses developed previously was used to systematically predict the conformational epitopes of E18 of enterovirus B. The results show that the epitopes were distributed into three clusters, where the VP1 BC loop, VP1 C-terminus and VP2 EF loop were the main regions of E18 epitopes. Further, molecular evolution analysis based on the VP1 and genome sequences of E18 revealed that the emergence of evolutionary clades was accompanied by amino acid mutations on epitopes, and the VP1 C-terminus and VP2 EF loop were mutation hotspot regions. Finally, the roadmap revealed that most of the clade-specific mutations on the E18 capsid surface were located on or near the epitopes, while mutations on the receptor binding sites were rare. This suggested that the mutations on the epitopes give rise to a new evolutionary clade, but the clade-specific mutations try to avoid the receptor binding site so as not to affect the ability of the virus to bind to the host cells. The bioinformatics study of E18 conformational epitopes provides important support for experimental identification of epitopes, virus surveillance and early warning, and the development of antiviral drugs and vaccines.

Key words: echovirus 18; enterovirus; conformational epitope; antigenic epitope; molecular evolutionary analysis; bioinformatics

(责任编辑 史小丽)