

DOI:10.13350/j.cjpb.230425

• 综述 •

全基因组测序技术在鼠疫分子溯源中的应用*

杨荣^{1,2}, 王鹏¹, 钟佑宏¹, 石丽媛^{1**}

(1. 云南省地方病防治所, 云南省自然疫源性疾病预防控制重点实验室, 云南大理 671000; 2. 昆明医科大学公共卫生学院)

【摘要】 高通量测序技术的发展使得测序成本大幅降低, 全基因组测序(whole genome sequencing, WGS)技术已成为追溯传染性病原菌和研究细菌遗传进化的主要技术手段。鼠疫菌基因组序列的测定在鼠疫疫情溯源及探索病原菌遗传进化中发挥至关重要的作用。细菌表型分析虽能体现一定的遗传关系, 但不能直接评估所测定表型的遗传基础, 不宜用于鼠疫菌的溯源, WGS技术因其具有获取遗传信息全面、精确、高通量及高分辨率等优势, 现已在鼠疫的分子溯源工作中被广泛应用。本文简要综述了WGS技术及其在鼠疫分子溯源中的应用进展。

【关键词】 鼠疫菌; 全基因组测序; 基因组; 分子溯源; 综述

【中图分类号】 R516.8

【文献标识码】 A

【文章编号】 1673-5234(2023)04-0489-06

[Journal of Pathogen Biology. 2023 Apr;18(4):489-494.]

Application of whole genome sequencing in molecular tracing of *Yersinia pestis*

YANG Rong^{1,2}, WANG Peng¹, ZHONG You-hong¹, SHI Li-yuan¹ (1. Yunnan Provincial Key Laboratory for Zoonosis Control and Prevention, Yunnan Institute for Endemic Disease Control and Prevention, Dali 671000, Yunnan, China; 2. Public Health School, Kunming Medical University)

【Abstract】 The development of high-throughput sequencing has greatly reduced the sequencing cost. Whole genome sequencing (WGS) has become the main technical means to trace the pathogens of infectious diseases and to study the genetic evolution of bacteria. The sequencing of *Yersinia pestis* genome plays an important role in tracing the source of plague and exploring the genetic evolution of pathogenic bacteria. Although phenotypes can reflect certain genetic evolutionary relation, it cannot directly evaluate the genetic basis of phenotypic, therefore, it should not be used for the tracing of *Yersinia pestis*. In contrast, WGS has been widely used in the molecular tracing of *Yersinia pestis* since its advantages of obtaining comprehensive, accurate, high-throughput and high-resolution genetic information. We discussed the advances of WGS and its application in the molecular tracing of *Yersinia pestis*.

【Key words】 *Yersinia pestis*; whole genome sequencing; genome; molecular tracing; review

***鼠疫是由鼠疫菌引起的人兽共患烈性传染病, 曾发生三次世界性大流行, 造成了数百万死亡, 是历史上包括黑死病在内的数次毁灭性流行病的罪魁祸首^[1]。由于鼠疫菌易从广泛分布的鼠疫自然疫源地获得, 鼠疫菌被列为生物恐怖主义病原体^[2]。早在青铜时代鼠疫菌就被证明能引起人类感染, 鼠疫是一种古老的传染病, 如今仍对人类健康造成重大威胁, 例如2017年8~11月马达加斯加爆发了大规模的肺鼠疫^[3], 2019年11月, 我国内蒙古自治区及北京报告的人间鼠疫^[4]。然而, 对鼠疫菌的溯源研究能及时阻断疫情的蔓延以及遏制生物恐怖事件的进一步发展, 可见对鼠疫菌的研究具有重要意义。

WGS技术是鼠疫等各类传染性疾病预防溯源工作的重要基础。全基因组序列能在传染性疾病暴发流行时提供可操作信息并加快病原体的追踪和识别能力^[5]。近年来, 高通量测序技术和分析工具的进步迅速提高了WGS的产量和分析速率, 并降低了成本^[6-7], 使得WGS技术在鼠疫分子溯源工作中得到了广泛的应用, 并有望成为鼠疫等各类传染性疾病预防溯源的主流技术^[8]。此外, 该技术的发展将便于对世界各地鼠疫流行菌株进行快速调查、群体结构分析和聚类分析等研究^[9]。

1 鼠疫菌的进化

对鼠疫菌基因组进化的研究是鼠疫溯源工作的重要基础。

鼠疫菌进化较晚, 无法积累广泛的序列多态性, 遗传变异水平相对较低^[10]。了解其进化背景是阐明鼠疫发病机制及传播流行模式所必需的。尽管早在青铜时代就已证明鼠疫菌能引起人类感染, 但直到1894年微生物学家才将其确定为鼠疫的病原菌^[11]。鼠疫菌和假结核菌的DNA序列高度同源, 从DNA序列相似性来看, 鼠疫菌和假结核菌满足了物种相似性的共同标准, 但由于其毒力、传播途径以及对类历史的影响存在巨大差异, 二者被列为不同的物种^[12]。从rDNA序列获得的数据也支持了这两个物种之间的近缘关系, 二者的16S rDNA序列完全相同, 在23S rDNA高度可变区域内只有一个碱基差异^[13]。为了推断鼠疫菌的起源和进化史, Achtman等^[12]将来自3个经典生物变种的36个分离菌株基因进行了测序, 均未

* **【基金项目】** 国家自然科学基金项目(No. 82260659, No. 31660043); 云南省高层次卫生健康技术人才培养项目(No. L-2019001, D-2019002, H-2019003); 云南省科技厅面上项目(No. 2018FB030)。

** **【通讯作者】** 石丽媛, E-mail: 15087261364@163.com

【作者简介】 杨荣(1998-), 女, 云南德宏人, 硕士研究生, 主要研究方向为鼠疫菌基因组学。E-mail: running1678@163.com

发现序列多态性,且这些等位基因与假结核菌几乎相同,从而证实了鼠疫菌是由假结核菌进化而来^[14-15];Cui等^[16]比较了133株鼠疫菌和4株假结核菌的基因组特征,发现假结核菌基因组与鼠疫菌存在相同的单核苷酸多态性(SNP),并且其最大似然树证实了这133个鼠疫菌基因组都表现出有限的多态性,再次证实了鼠疫菌是假结核菌的后代。

随着测序技术的快速发展,更多的耶尔森菌基因组被破译。一项基于全基因组的研究分析了18个不同物种的241株耶尔森菌,基于核心基因的系统发育结果显示,鼠疫菌与其他耶尔森菌的遗传多态性非常有限,然而其与假结核菌的遗传谱系却形成了一个紧密的群体,进一步证明了鼠疫菌是假结核菌的克隆体^[13,17]。

2 鼠疫菌表型分析

甘油发酵(*glpD*+)和硝酸盐还原实验(*napA*+)是鼠疫菌最早也是传统上使用最广泛的表型分析方法。Devignat等^[18]根据鼠疫菌对硝酸盐还原和甘油醇解能力,将鼠疫菌分为3个生物型:古典型(*glpD*+,*napA*+)、中世纪型(*glpD*+,*napA*-)和东方型(*glpD*-,*napA*+)。Zhou等^[20]根据甘油发酵和硝酸盐还原试验以及发酵阿拉伯糖的能力(*araC*+) ,发现了第4种生物型,即田鼠型(*glpD*+,*napA*-,*araC*-),经典的三种生物型鼠疫菌阿拉伯糖发酵均为阳性(表1)。Devignat等^[21]根据表型分析和史料记载,提出鼠疫菌的三个生物型分别与历史上的三次鼠疫大流行相关,古典型菌株与查士丁尼瘟疫有关,中世纪型菌株与第二次大流行有关,东方型菌株与第三次大流行有关,而田鼠型菌株对人类没有致病性,仅引起动物间鼠疫的发生和流行^[22]。

表1 不同生物型鼠疫菌的生化特征

Table 1 Biochemical characteristics of different biological strains of *Yersinia pestis*

鼠疫菌生物型 Biovar	生化特征 Biochemical characteristics		
	甘油发酵 (<i>glpD</i>) Fermenting glycerol	硝酸盐还原实验 (<i>napA</i>) Reducing nitrate	阿拉伯糖发酵 (<i>araC</i>) Fermenting arabinose
古典型	+	+	+
中世纪型	+	-	+
东方型	-	+	+
田鼠型	+	-	-

注:“+”表示阳性,“-”表示阴性。

目前鼠疫菌一系列的表型分析方法被发掘并应用^[23]。然而物种的表型不具稳定性,如由于独立的突变事件导致出现相同表型,菌株可能被错误地分型^[24];不同菌株在某些表型测定时可能看起来是相同的,但基因型不同。全基因组序列能提供整个基因组的概况,能从分子水平阐明全基因组范围内的变异及关联。若将表型分析方法与更可靠的如WGS分析等分子分析方法相比较,上述错误可能会被识别^[25]。

表型分析表明一定的遗传特征,但不能直接评估所测定表型的遗传基础,相比之下,基于WGS的全基因组序列分析能够检测到基因组绝大多数类型的变异情况,便于了解整个基因组的特征^[23,26],从而能更准确地分析细菌的进化来源。

3 WGS技术

WGS技术是对物种个体或群体的基因组进行测序以获得

基因组图谱,分析其基因组间的差异,旨在全基因组水平内探索物种进化规律和筛选功能基因的研究技术^[27]。传统的基因分析技术分辨能力有限,仅能反映病原菌基因组部分遗传信息,在克隆化菌株研究方面存在诸多不足,WGS技术则可了解病原菌整个基因组的核苷酸序列,并进行科学的基因分型,支持对庞大的基因组数据的计算和解释,利用NCBI公共数据库共享的基因组序列信息,比对出基因组之间SNP差异,清晰地反映病原菌的遗传和变异特征,大大提升了对菌株和溯源的分辨能力,进而实现对病原体的追踪^[28]。相比传统的分子分型技术,WGS技术分辨率高,在公共卫生领域已成为公认的传染性疾病的病前监测工具,其将取代微生物实验室中如血清分型、毒力分析、抗菌素耐药性测定及分子分型方法等大多数鉴定和表征方法^[29]。由于细菌基因组学研究的不断深入,测序平台和生物信息学分析方法也在不断进步,极大提升了细菌基因组测序数据的准确性与实用性,为有效监测疾病的传播提供了新思路^[30-31]。

目前细菌基因组学研究领域采用的测序平台有基于短片段测序技术的二代测序平台,如Illumina系列、Ion Torrent系列和BGI系列等;以及基于长片段测序技术的三代测序平台,如Oxford Nanopore Technologies 纳米孔(Nanopore sequencing)和Pacific Biosciences (PacBio)的SMRT (Single Molecule, Real-Time)单分子测序技术等^[32](表2),这些测序平台的发展与改善,为基因组学研究提供了可靠数据。

4 WGS技术在鼠疫分子溯源中的应用

高通量测序技术的进步,使得测序成本大幅降低,目前WGS技术已被广泛使用,并快速超越传统的基因分析技术,成为当前研究物种群体进化、变异分析和功能基因挖掘的主流技术^[33]。因WGS技术具有很高的分辨率^[34],已被广泛应用在鼠疫菌群体进化和溯源研究中^[35]。该技术的快速发展及普及应用使鼠疫菌基因组学方面的研究取得了众多成果^[36]。

4.1 鼠疫菌全基因组序列的测定 自1995年对第一株细菌基因组进行WGS以来,WGS技术已取得了长足的发展,基于高通量测序技术的兴起,越来越多的细菌基因组序列被测定^[38-39]。2001年Parkhill等^[40]将来自北美的强毒菌株CO92进行WGS,CO92是首例完成WGS的鼠疫菌分离菌株,次年Deng等^[41]完成了强毒株KIM的WGS,Song等^[42]于2004年完成了我国布氏田鼠鼠疫疫源地菌株91001的WGS工作。Chain等^[43]于2006年完成Nepal516菌株和Antiqua菌株的WGS工作。随着WGS技术的快速发展,越来越多的鼠疫菌全基因组序列已被测定(表3),截止目前,NCBI公布的已进行全基因组测序的鼠疫菌共657株,这些基因组序列信息的共享为后续研究及溯源工作提供了便利。

4.2 鼠疫菌群体结构特征 根据WGS分析,可将鼠疫菌分为5个主要分支^[44]。基于世界各地不同鼠疫疫源地菌株的WGS结果,Morelli等^[45-47]将各疫源地菌株与CO92、KIM和91001菌株进行全基因组序列的SNP分析,从而将鼠疫菌的种群结构分为3个分支,分别命名为Branch 0、Branch 1和Branch 2。Cui等^[16]对133株鼠疫菌进行全基因组SNP分析,发现了另外2个新的分支,分别是Branch 3和Branch 4。Branch 0包括0. ANT1、0. ANT2、0. ANT3、0. PE2、0. PE3、0. PE4A、0. PE4B、0. PE4C和0. PE7,9个分支;Branch 1包括1. IN1、1. IN2、1.

IN3、1. ANT、1. ORI1、1. ORI2、1. ORI3、Ancient genomes, 8个分支; Branch 2包括 2. ANT1、2. ANT2、2. ANT3、2. MED1、2. MED2、2. MED3, 6个分支; Branch 3包括 3. ANT1、3. ANT2、2个分支; Branch 4包括 4. ANT1, 1个分支(图1)。近年 Shi

等^[48]对10株分离自云南鹤庆和10株分离自丽江的鼠疫菌的基因组进行了测序,并将这些基因组序列与368个已公布的代表全球多样性的基因组序列进行了比较,结果表明鹤庆分离菌株基因组在系统发育树中属于新的谱系分支1. IN5。

表2 各测序平台对比
Table 2 Comparison of sequencing platforms

测序平台 Sequencing platform	原理 Sequencing principle	读取长度 Average read length	优点 Advantages	缺点 Disadvantages
Illumina	利用荧光标记的核苷酸来确定DNA片段的遗传序列 ^[32] 。采用可逆性测序末端边合成边测序;在碱基延伸过程中,每个循环反应只延伸一个互补的带荧光碱基,通过不同的荧光信号确认碱基种类。	≤300 bp	高精度度、高通量、高自动化	读长短、费用高、耗时长
Ion Torrent	半导体测序技术,通过半导体芯片直接将化学信号转化为数字信号。利用测序过程DNA合成时所释放H ⁺ 导致的测序环境中pH的变化进行碱基的识别 ^[33] 。	≤400 bp	精确度高、耗时短、费用低	读长短、通量低、操作较繁琐、对操作人员的依赖性较强
BGI	采用DNA纳米球测序技术,利用滚环式复制让待测的DNA扩增形成DNA纳米球,通过组合探针锚定连接法进行边合成边测序 ^[33] 。	≤300 bp	精确度高	读长短、费用高、耗时长
Pacific Bioscience	当荧光标记的脱氧核苷酸与待测DNA链形成化学键时,荧光被切除而消失,通过检测记录荧光强度的变化来获得相应的碱基序列 ^[33] 。	≤25 kb	读长长、耗时短	精确度低、费用高
OxfordNanopore	纳米孔测序,纳米孔直径微小,仅允许单个核苷酸通过,而ATCG单个碱基的带电性质不同,利用其带电性质的差异来检测碱基类别 ^[34] 。	≤2 Mb	读长长、费用低、耗时短、实时分析数据	精确度低、随机错误而非聚集在读取两端

表3 鼠疫菌各谱系分支菌株全基因组测序信息
Table 3 Whole gene sequencing information of each branch strain of Yersinia pestis

菌株名称 Strain name	菌株来源 Source	NCBI 登录号 Accession number	基因组长度(bp) Genome size	(G+C)% GC content	生物型 Biovar	谱系分支 Node
42013	中国新疆	ADPG00000000	4495101	44.10	古典型	0. ANT1
B42003004	中国新疆	NZ_AAYU00000000	4841690	47.71	古典型	0. ANT2
CMCC38001	中国新疆	ADQU00000000	4483258	52.48	古典型	0. ANT3
Pestoides F	前苏联	NC_009381	4517345	48.40	田鼠型	0. PE2
Angola	安哥拉	NC_010159	4504254	47.50	古典型	0. PE3
12	中国青海	ADOV00000000	4368767	47.89	古典型	0. PE4A
PestoidesA	FSU	NZ_ACNT00000000	4750606	47.62	田鼠型	0. PE4B
91001	中国内蒙古	NC_005810	4595065	47.65	田鼠型	0. PE4C
CMCC05009	中国青海	ADQE00000000	4501730	43.30	古典型	0. PE7
Antiqua	非洲刚果	NC_008150	4702289	47.70	古典型	1. ANT
CMCC11001	中国青海	ADQK00000000	4423348	46.84	古典型	1. IN1
30017	中国台湾	ADPC00000000	4591472	43.76	古典型	1. IN2
F1954001	中国云南	ADSC00000000	4521000	46.65	古典型	1. IN3
CO92	美国科罗拉多州	NC_003143	4556337	47.64	东方型	1. ORI1
YN2179	缅甸	ADTE00000000	4307368	46.04	东方型	1. ORI2
MG05	马达加斯加	NZ_AAYS00000000	4977609	48.96	东方型	1. ORI3
EV76	马达加斯加	NZ_ADSA00000000	4336975	50.93	东方型	1. ORI3
Nepal516	尼泊尔	NZ_ACNQ00000000	4629080	47.58	古典型	2. ANT1
G1996006	中国台湾	ADSE00000000	4542684	49.05	古典型	2. ANT2
CMCC92010	中国内蒙古	ADRL00000000	4384955	44.12	古典型	2. ANT3
KIM	伊朗库尔德	NC_004088	4600755	47.64	中世纪型	2. MED1
7338	中国新疆	ADPQ00000000	4344510	42.82	中世纪型	2. MED2
J1963002	中国甘肃	ADSP00000000	4434765	46.30	中世纪型	2. MED3
CMCC71001	中国甘肃	ADRC00000000	4254189	54.82	古典型	3. ANT1
MGJZ6	中国内蒙古	ADSX00000000	4368072	44.35	古典型	3. ANT2
MGJZ12	中国内蒙古	ADSV00000000	4438965	36.52	古典型	4. ANT1

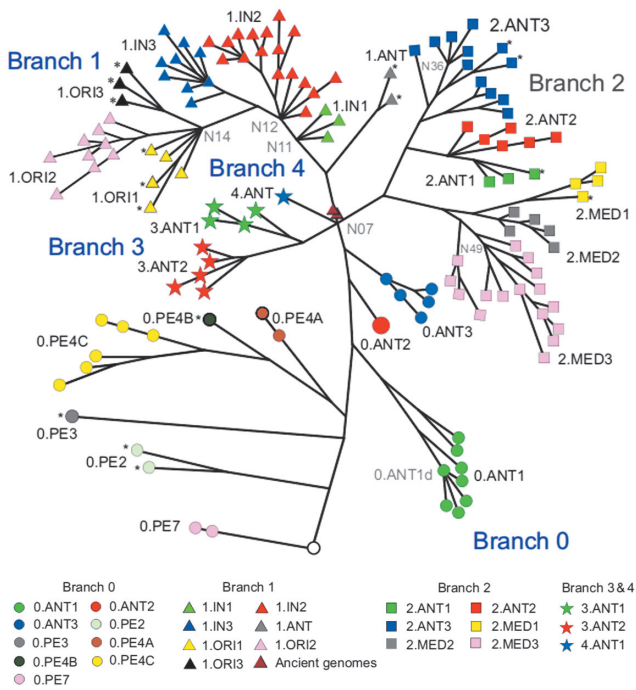


图 1 鼠疫菌谱系分支^[16]
Fig.1 Population structure of *Y. pestis*

4.3 基于 WGS 技术的 SNP 分析 SNP 在系统发育分析中具有分辨率高、易鉴定等优势^[8]。基于 WGS 的 SNP 分析,通过同类致病菌基因的横向比对,加速了致病菌来源的测定速度,可达到及时阻断传染源的传播途径,加速疫情控制的目的^[49]。目前,包括鼠疫菌在内的大多数细菌群体基因组研究都是基于 SNP 的变异^[50]。Morelli 等^[51]根据全球 286 株鼠疫菌分离株中 17 个完整基因组序列和 933 个 SNP 的序列变异,重建了鼠疫的历史传播模式,其结果表明鼠疫菌在中国或其附近进化,并通过多种途径传播,包括通过丝绸之路传播到西亚和通过中国海上航行传播到非洲等,同时该研究也进一步验证了鼠疫菌进化种群结构框架的准确性和可行性。Spyrou 等^[52]重建了 34 个古代鼠疫菌基因组,利用最大似然法构建 SNP 系统发育,来推断新重建和以往发表的鼠疫菌分离株之间的遗传关系,结果显示所有新重建的基因组都出现在 Branch 1 上,且证实其与之前来自欧洲第二次鼠疫大流行的分离株密切相关。Cui 等^[16]将来自中国、蒙古、前苏联、缅甸和马达加斯加的 118 株鼠疫菌进行全基因组测序,并结合已公布测序的 15 株鼠疫菌全基因组序列,重建基于 SNP 序列的系统发育关系,结果表明最深的分支谱系 0. PE7 只在中国的青藏高原被分离,鼠疫可能起源于中国青藏高原或青藏高原附近^[51]。Eroshenko 等^[53]对 50 多年来在吉尔吉斯斯坦高山区分离到的 56 株鼠疫菌进行了鉴定,结果发现 42 株菌株分别属于系统发育分支的 0. ANT2、0. ANT3 和 0. ANT5 谱系,其中 0. ANT5 是首次检出的鼠疫菌系统发育分支谱系,从而也证实了 0. ANT2 和 0. ANT3 分支不仅是中国菌株独有的谱系;同时该研究还将 7 株分离菌株进行 WGS,全基因组 SNP 分析结果表明,0. ANT5 菌株是与引起查士丁尼鼠疫关系最密切的菌株之一,鼠疫菌的多样性和 0. ANT 分支的广泛传播证实了鼠疫源地的古老性,表明所有强毒鼠疫菌 0. ANT 分支的菌株都起源于天山山脉一代。

Eroshenko 等^[54]对分离自里海北部和咸海北部疫区的 38 株鼠疫菌的进行全基因组 SNP 分析,确定了该地区一个中世纪生物变种 2. MED 种群的时空结构,系统发育分支 2. MED4 先于其后分化的 2. MED1 分支,并证实 2. MED1 菌株是 20 世纪初导致里海北部地区高死亡率鼠疫暴发的病原,20 世纪后期,2. MED1 分支广泛分布在里海地区、高加索地区和中亚的广大地区。Shi 等^[48]将分离自云南鹤庆和丽江的鼠疫菌基因组序列与 368 个已公布在 NCBI 数据库中的基因组序列进行了比较,构建基于 SNP 的最大似然树,结果显示鹤庆菌株在系统发育树较深分支上含有丽江分离菌株,表明鹤庆分离菌株可能起源于丽江地区,且鹤庆和丽江均位于横断山脉南段,两地相距仅 35 公里,因此,鹤庆野鼠鼠疫源地的出现可能是丽江鼠疫源地向南扩大的结果。Li 等^[4]将 2019 年内蒙古 4 例(分别为 A、B、C、D)人间鼠疫分离株与代表全球鼠疫菌物种多样性的 11 个鼠疫菌基因组序列及 GenBank 数据库中分离自内蒙古和中国其他省份的 54 株鼠疫菌全基因组序列,采用全基因组 SNP 分析建立系统发育,进行病原体来源追踪,结果显示,当地内蒙古分离菌株与患者 C 中分离出的菌株聚为 2MED3m 和 2MED3 种群中的两个新谱系(2MED3q、2. MED3r),与 A、B 和 D 患者聚为 2. MED3m。此外,来自患者住所附近的啮齿动物分离菌株与患者聚在同一个谱系中,结果表明,此次内蒙古人间鼠疫来源于当地鼠间鼠疫源地。

4.4 基于 WGS 技术的比较基因组学分析 Bos 等^[55]将 1722 年法国马赛最后一次欧洲大瘟疫爆发期间的 5 株鼠疫菌样本进行 WGS,并将全基因组序列与现代鼠疫菌以及其他历史菌株样本的基因序列进行比较,结果表明,引起马赛鼠疫爆发的流行菌株很可能是 14 世纪导致黑死病流行菌株的后代。秦婧靓等^[56]将两株已进行 WGS 的苏拉特鼠疫菌基因组与 366 株鼠疫全球代表菌株的基因组进行比较基因组学分析,聚类结果显示 2 株苏拉特菌株均为鼠疫菌 EV76 减毒活疫苗株的后代,其基因型为 EVT03/EVT04,该结果表明,引发 1994 年苏拉特鼠疫疫情的菌株属于疫苗株 EV76 种系。

4.5 基于 WGS 技术的其他分析 Pisarenko 等^[37]将 16 株分离自高加索和外高加索自然疫源地的鼠疫菌进行 WGS,并结合 NCBI 数据库中的 329 株鼠疫菌基因组序列,进行进化系统地理学分析,结果表明,所研究的高加索和外高加索自然疫源地的鼠疫菌属于遗传谱系 0. PE2、2. MED0 和 2. MED1。张恩民等^[57]将 2 株分离自云南并完成全基因组测序的鼠疫菌与 NCBI 公布已完成全基因组测序的 8 株鼠疫菌基因组的编码序列进行比较,建立了鼠疫菌“默认编码序列”数据库,该数据库在分析鼠疫菌的遗传进化特征和溯源方面提供了可靠信息,例如将分离的鼠疫菌株全基因组序列与该数据库进行比对,可快速确定鼠疫感染的来源,并及时作出相应决策^[58]。Yan 等^[59]基于 WGS 信息,提出了鼠疫菌两步溯源策略,通过这一策略分析发现,青海省星海县虽自 1956 年以来已多次分离出鼠疫菌株,但 2009 年暴发的分离株并非这些历史分离株的直系后代,而 1971 年从邻近的湟源县分离到的 1. IN2f 谱系的菌株与星海县暴发株之间仅存在 2 个 SNP 位点的差异,据此推测 2009 年星海鼠疫暴发极有可能是周边湟源县疫源地鼠疫菌的扩散感染所致。

5 小结

目前鼠疫在我国乃至全球仍是重大的公共卫生问题之一,弄清鼠疫疫情的起源及流行菌株的基因组学特征,是预防与控制鼠疫进一步扩散蔓延的科学手段。近年来,WGS技术的应用越来越广泛,其在国内外鼠疫溯源工作中的广泛应用,使得追溯鼠疫疫情的来源以及了解致病菌的进化历程成为了可能。然而,WGS技术也存在一定的不足之处。全基因组序列的测定以及基于全基因组序列的系统发育分析需要耗费大量的时间和计算资源;其次通过全基因组序列测定来分析菌株间的亲缘关系,还需结合生物信息学分析工具才能完成。因此,未来发展更简便、快捷的细菌基因组分析策略,是鼠疫暴发期间快速溯源的迫切需要,同时也是为后续进行基因组学研究提供简便实用的技术方法的需要。

【参考文献】

- [1] Morozova I, Kasianov A, Bruskin S, et al. New ancient Eastern European *Yersinia pestis* genomes illuminate the dispersal of plague in Europe[J]. *Philos Trans R Soc Lond B Biol Sci*, 2020, 375(1812):20190569.
- [2] Kaiser J. Select agents. Researchers relieved by final biosecurity rules[J]. *Science*, 2005, 308(5718):31.
- [3] Tsuzuki S, Lee H, Miura F, et al. Dynamics of the pneumonic plague epidemic in Madagascar, August to October 2017 [J]. *Euro Surveill*, 2017, 22(46):17-00710 .
- [4] Li J, Wang Y, Liu F, et al. Genetic source tracking of human plague cases in Inner Mongolia-Beijing, 2019[J]. *PLoS Negl Trop Dis*, 2021, 15(8):e9558.
- [5] Lin Y, Yang L, Qiu S, et al. Rapid identification and source tracing of a *Salmonella Typhimurium* outbreak in china by metagenomic and whole-genome sequencing [J]. *Foodborne Pathog Dis*, 2022, 19(4):259-265.
- [6] Harris SR, Cartwright EJ, Torok ME, et al. Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study[J]. *Lancet Infect Dis*, 2013, 13(2):130-136.
- [7] Quainoo S, Coolen J, van hijum S, et al. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis[J]. *Clin Microbiol Rev*, 2017, 31(1): e00082-17.
- [8] 贾慧琼,阮陟. 全基因组测序在病原菌分型与溯源中的应用研究进展[J]. *微生物学报*, 2022, 62(03):949-967.
- [9] Li W, Cui Q, Bai L, et al. Application of whole-genome sequencing in the national molecular tracing network for foodborne disease surveillance in China[D]. 2021.
- [10] Susat J, Lubke H, Immel A, et al. A 5 000-year-old hunter-gatherer already plagued by *Yersinia pestis*[J]. *Cell Rep*, 2021, 35(13):109278.
- [11] Rasmussen S, Allentoft ME, Nielsen K, et al. Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago[J]. *Cell*, 2015, 163(3):571-582.
- [12] Achtman M, Zurth K, Morelli G, et al. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*[J]. *Proc Natl Acad Sci U S A*, 1999, 96(24):14043-14048.
- [13] Cui Y, Song Y. Genome and evolution of *Yersinia pestis*[M]. Dordrecht; Springer Netherlands, 2016:171-192.
- [14] Andrades VA, Mittik A, Key FM, et al. The stone age plague and its persistence in Eurasia[J]. *Curr Biol*, 2017, 27(23):3683-3691.
- [15] 宋凯, 祁芝珍, 杨瑞馥, 等. 鼠疫历史疫情的考古微生物学研究进展[J]. *中国人兽共患病学报*, 2018, 34(12):1142-1146.
- [16] Cui Y, Yu C, Yan Y, et al. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*[J]. *Proc Natl Acad Sci U S A*, 2013, 110(2):577-582.
- [17] Reuter S, Connor TR, Barquist L, et al. Parallel independent evolution of pathogenicity within the genus *Yersinia* [J]. *Proc Natl Acad Sci U S A*, 2014, 111(18):6768-6773.
- [18] Devignat R. Varieties of *Pasteurella pestis*; new hypothesis[J]. *Bull World Health Organ*, 1951, 4(2):247-263.
- [19] Anisimov AP, Lindler LE, Pier GB. Intraspecific diversity of *Yersinia pestis*[J]. *Clin Microbiol Rev*, 2004, 17(2):434-464.
- [20] Zhou D, Tong Z, Song Y, et al. Genetics of metabolic variations between *Yersinia pestis* biovars and the proposal of a new biovar, *microtus*[J]. *J Bacteriol*, 2004, 186(15):5147-5152.
- [21] Devignat R. Varieties of *Pasteurella pestis*; new hypothesis[J]. *Bull World Health Organ*, 1951, 4(2):247-263.
- [22] Zhou D, Tong Z, Song Y, et al. Genetics of metabolic variations between *Yersinia pestis* biovars and the proposal of a new biovar, *microtus*[J]. *J Bacteriol*, 2004, 186(15):5147-5152.
- [23] Vogler AJ, Keim P, Wagner DM. A review of methods for subtyping *Yersinia pestis*: From phenotypes to whole genome sequencing[J]. *Infect Genet Evol*, 2016, 37:21-36.
- [24] Anisimov AP, Lindler LE, Pier GB. Intraspecific diversity of *Yersinia pestis*[J]. *Clin Microbiol Rev*, 2004, 17(2):434-464.
- [25] Platonov ME, Evseeva VV, Dentovskaya SV, et al. Molecular typing of *Yersinia pestis*[J]. *Mol Gen Mikrobiol Virusol*, 2013(2):3-12.
- [26] Brown AC. Whole-genome sequencing of *Mycobacterium tuberculosis* directly from sputum samples [J]. *Methods Mol Biol*, 2021(2314):459-480.
- [27] 周子超, 何岚, 丁月平. 全基因组测序(WGS)在肠杆菌科细菌中的应用研究[J]. *智慧健康*, 2019, 5(2):70-72.
- [28] 崔玉娟. 基因测序技术在食品安全检测中的应用[J]. *食品安全导刊*, 2021(23):173-174.
- [29] Jagadeesan B, Gerner-smidt P, Allard MW, et al. The use of next generation sequencing for improving food safety: Translation into practice[J]. *Food Microbiol*, 2019, 79:96-115.
- [30] Ruan Z, Yu Y, Feng Y. The global dissemination of bacterial infections necessitates the study of reverse genomic epidemiology[J]. *Brief Bioinform*, 2020, 21(2):741-750.
- [31] Tang P, Croxen MA, Hasan MR, et al. Infection control in the new age of genomic epidemiology[J]. *Am J Infect Control*, 2017, 45(2):170-179.
- [32] Lau KA, Gon alves da Silva A, Theis T, et al. Proficiency testing for bacterial whole genome sequencing in assuring the quality of microbiology diagnostics in clinical and public health laboratories [J]. *Pathology*, 2021 Dec;53(7):902-911.
- [33] 陈慧娟, 刘琪琦. 高通量测序技术在细菌耐药中的应用[J]. *中国生物化学与分子生物学报*, 2022, 38(7):865-874.
- [34] Greninger AL, Naccache SN, Federman S, et al. Rapid

- metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis[J]. *Genome Med*, 2015,7:99.
- [35] Consortium Z. A comparative genomics multitool for scientific discovery and conservation[J]. *Nature*, 2020, 587(7833): 240-245.
- [36] Rascovan N, Sjogren KG, Kristiansen K, et al. Emergence and spread of basal lineages of *Yersinia pestis* during the neolithic decline[J]. *Cell*, 2019, 176(1-2): 295-305.
- [37] Pisarenko SV, Evchenko AY, Kovalev DA, et al. *Yersinia pestis* strains isolated in natural plague foci of Caucasus and Transcaucasia in the context of the global evolution of species [J]. *Genomics*, 2021, 113(4): 1952-1961.
- [38] Didelot X, Bowden R, Wilson DJ, et al. Transforming clinical microbiology with bacterial genome sequencing [J]. *Nat Rev Genet*, 2012, 13(9): 601-612.
- [39] Tagini F, Greub G. Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review [J]. *Eur J Clin Microbiol Infect Dis*, 2017, 36(11): 2007-2020.
- [40] Parkhill J, Wren BW, Thomson NR, et al. Genome sequence of *Yersinia pestis*, the causative agent of plague[J]. *Nature*, 2001, 413(6855): 523-527.
- [41] Deng W, Burland V, Plunkett GR, et al. Genome sequence of *Yersinia pestis* KIM[J]. *J Bacteriol*, 2002, 184(16): 4601-4611.
- [42] Song Y, Tong Z, Wang J, et al. Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to humans[J]. *DNA Res*, 2004, 11(3): 179-197.
- [43] Chain PS, Hu P, Malfatti SA, et al. Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: evidence of gene reduction in an emerging pathogen[J]. *J Bacteriol*, 2006, 188(12): 4453-4463.
- [44] Wang J, Yang X, Zheng H, et al. Whole-genome assemblies for three *Yersinia pestis* strains isolated in erenhot, China [J]. *Microbiol Resour Announc*, 2020, 9(45): e01084-20. .
- [45] Morelli G, Song Y, Mazzoni CJ, et al. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity [J]. *Nat Genet*, 2010, 42(12): 1140-1143.
- [46] Achtman M, Morelli G, Zhu P, et al. Microevolution and history of the plague bacillus, *Yersinia pestis* [J]. *Proc Natl Acad Sci U S A*, 2004, 101(51): 17837-17842.
- [47] Achtman M, Zurth K, Morelli G, et al. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis* [J]. *Proc Natl Acad Sci U S A*, 1999, 96(24): 14043-14048.
- [48] Shi L, Qin J, Zheng H, et al. New genotype of *Yersinia pestis* found in live rodents in Yunnan Province, China [J]. *Front Microbiol*, 2021(12): 628335.
- [49] Spyrou MA, Tukhbatova R, Wang CC, et al. Analysis of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague [J]. *Nat Commun*, 2018, 9(1): 2234.
- [50] Wu Y, Hao T, Qian X, et al. Small insertions and deletions drive genomic plasticity during adaptive evolution of *Yersinia pestis* [J]. *Microbiol Spectr*, 2022, 10(3): e224221.
- [51] Morelli G, Song Y, Mazzoni CJ, et al. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity [J]. *Nat Genet*, 2010, 42(12): 1140-1143.
- [52] Spyrou MA, Keller M, Tukhbatova RI, et al. Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes [J]. *Nat Commun*, 2019, 10(1): 4470.
- [53] Eroshenko GA, Nosov NY, Krasnov YM, et al. *Yersinia pestis* strains of ancient phylogenetic branch 0. ANT are widely spread in the high-mountain plague foci of Kyrgyzstan [J]. *PLoS One*, 2017, 12(10): e187230.
- [54] Eroshenko GA, Popov NV, Al'khova ZV, et al. Evolution and circulation of *Yersinia pestis* in the Northern Caspian and Northern Aral Sea regions in the 20th-21st centuries [J]. *PLoS One*, 2021, 16(2): e244615.
- [55] Bos KI, Herbig A, Sahl J, et al. Eighteenth century *Yersinia pestis* genomes reveal the long-term persistence of an historical plague focus [J]. *Elife*, 2016(5): e12994.
- [56] 秦婧靓, 杨超, 郑宏源, 等. 1994年印度苏拉特鼠疫暴发的回顾性基因溯源研究 [J]. *军事医学*, 2021, 45(10): 757-761.
- [57] 张恩民, 王琪, 申小娜, 等. 鼠疫菌基因组“默认编码序列”数据库的建立与初步研究 [J]. *疾病监测*, 2010, 25(6): 436-438.
- [58] 王梅, 唐新元, 梁莹, 等. 鼠疫菌基因组差异编码序列的筛选研究 [J]. *中国病原生物学杂志*, 2017, 12(2): 115-118.
- [59] Yan Y, Wang H, Li D, et al. Two-step source tracing strategy of *Yersinia pestis* and its historical epidemiology in a specific region [J]. *PLoS One*, 2014, 9(1): e85374.

【收稿日期】 2022-10-14 【修回日期】 2023-01-06