

使用 Illumina NGS 系统检测冠状病毒的富集工作流程

可高度灵敏地检测和鉴定冠状病毒等常见呼吸道病毒的靶向富集测序工作流程。

简介

病毒感染是一个重大的全球性健康问题，新发传染性疾病层出不穷。2019 年最早发现于中国武汉并迅速蔓延至多个国家的新型冠状病毒（SARS-CoV-2）疫情就是一个特别令人担忧的事例。冠状病毒（CoV）属于一个庞大的病毒家族，可能引起一系列从普通感冒到中东呼吸综合征（MERS-CoV）以及严重急性呼吸综合征（SARS-CoV）等更加严重的人类呼吸系统疾病。SARS-CoV-2 是一种新的病毒株，之前从未在人类身上发现过。由于世界各地已经出现了数以万计的确诊病例，其死亡人数也超过了 2003 年的 SARS 疫情，世界卫生组织（WHO）宣布 SARS-CoV-2 相关疾病（COVID-19）为国际关注的突发公共卫生事件¹，这突显出快速、准确地检测病毒的必要性。

新一代测序（NGS）提供了一种有效的新方法，在无任何病原体信息的情况下，检测样本和发现病毒。² 靶向富集是一种重测序方法，可通过与靶点特异性生物素化探针杂交来捕获目标基因组区域。通过杂交捕获方法进行靶向富集可实现高灵敏度的检测，而无需鸟枪法宏基因组测序所需的高 read 深度。此外，靶向富集 NGS 工作流程能够获得近乎完整的目标序列数据，特别适于病毒基因组进化或变异监测等应用。³ 与其他靶向重测序方法相比，比如扩增子测序，通过杂交捕获进行富集可实现更大的探针检测范围，能对目标区域进行更全面的分析。此外，即使在高度变异区域，杂交捕获方案使用的寡核苷酸探针仍然有效，能靶向快速进化的病毒，例如 RNA 病毒。

病毒富集工作流程

本应用说明着重介绍了使用 Nextera™ Flex for Enrichment Library Preparation Kit 结合病毒靶向试剂盒、成熟的 Illumina 测序技术和简化的数据分析来检测和分析冠状病毒的一体化工作流程（图 1）。此工作流程旨在从总核酸提取物中富集病毒 DNA 和 RNA 靶点。该实验方案首先会提取核酸样本，然后将提取的 RNA 逆转录为双链 cDNA。文库制备使用 Nextera Flex for Enrichment Kit 和 IDT for Illumina Nextera UD Indexes 来完成。在此实验方案中，DNA 和 cDNA 会进行酶切法片段化、纯化和富集前扩增。扩增后可混合多达 12 个样本，并使用一组靶向病毒序列的寡核苷酸来对混合后的样本进行 1 个富集反应。探针杂交之后的步骤为探针捕获、富集扩增、定量和测序。

方法

样本准备

为了证明病毒靶向富集试剂盒的性能，本研究使用了一株灭活的冠状病毒样本，来自 Microbiologics 的 CoV 病毒株 OC43（QC Sets 和 panel：Helix Elite；货号 8217）。冠状病毒培养物样本* 使用 QIAGEN QIAmp Viral Mini Kit（QIAGEN，货号 52904）在 BSL2 实验室环境中提取（表 1）。使用 2 种不同的工作流程将提取的 150 ng RNA 逆转录为 cDNA：一个源自 Illumina TruSeq RNA 试剂，另一个使用 Thermo Scientific Maxima H Minus Double-Stranded cDNA Synthesis Kit（Thermo Scientific，货号 K2561）。病毒样本中还添加了 95% 的通用人类参考（UHR）背景 RNA（Agilent Technologies，货号 74000），

* 提取使用的是病毒培养物样本，不是纯病毒 RNA。

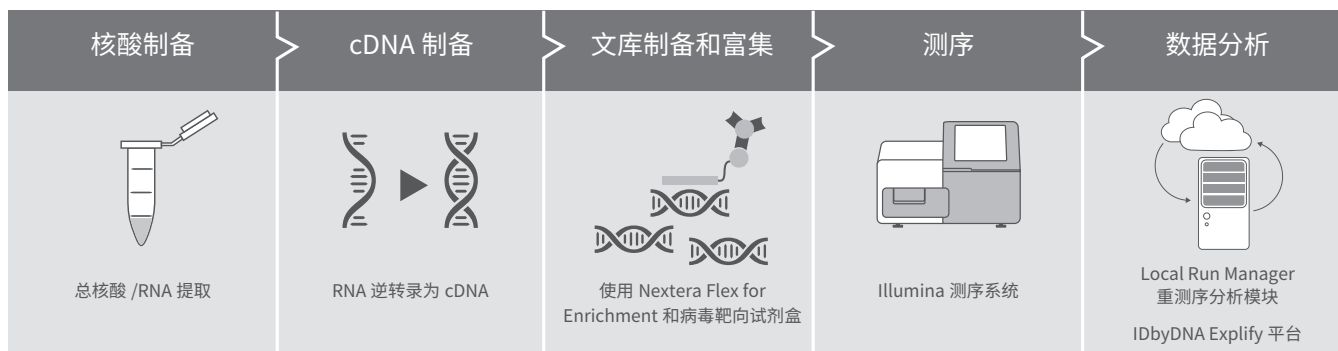


图 1：用于冠状病毒检测的富集工作流程 - 用于冠状病毒检测的一体化 NGS 工作流程整合了样本准备、文库制备、靶向富集、测序和数据分析。

并逆转录为 cDNA（使用上述两种方法），以此模拟来自患者样本的测序结果（表 1）。

此外，我们还使用 QIAGEN AllPrep PowerViral DNA/RNA Kit (QIAGEN, 货号 28000-50) 在 BSL2 实验室环境中提取了含有 4 种 RNA 病毒和 1 种 DNA 病毒的呼吸道病毒混合物（表 2）。我们使用源自 Illumina TruSeq RNA 试剂的工作流程将最大体积的总核酸逆转录为 cDNA（无需 DNase 处理）。病毒混合物样本中还添加了 95% 的通用人类参考 (UHR) 背景 RNA (Agilent Technologies, 货号 74000)，并逆转录为 cDNA，以此模拟来自存在多种病原体的患者样本的测序结果（表 1）。

表 1：用于分析的病毒样本的组成

样本	组成 ^a	参考基因组
CoVOC43	150 ng CoV OC43 RNA	AY391777.1 人冠状病毒 OC43
CoVOC43_95UHR	7.5 ng CoV OC43 RNA 和 142.5 ng UHR RNA	AY391777.1 人冠状病毒 OC43
ViralPool	最大体积的病毒混合物总核酸 (8.5 μL)	见表 2
ViralPool_95UHR	5% 病毒混合物总核酸和 95% UHR RNA (按体积混合)	见表 2

a. 逆转录推荐的最小 RNA/ 总核酸起始量是 10 ng。为获得最佳结果，逆转录应使用新鲜提取的核酸样本。

表 2：病毒混合物组成

病毒	核酸类别	参考基因组
甲型流感病毒 (H1N1)	RNA	Influenza A/Michigan/45/2015
乙型流感病毒	RNA	Influenza B/Colorado/06/2017
人副流感病毒 3 型	RNA	NC_001796.2
RSV B9320	RNA	AY353550.1
腺病毒 7 型	DNA	AY594255.1

文库制备

使用 Nextera DNA Flex Pre-Enrichment Library Prep and Enrichment Reagents Kit (Illumina, 货号 20025524) 和 IDT for Illumina Nextera DNA UD Indexes (Illumina, 货号 20027213)，对来自 CoV 样本 (CoVOC43)、病毒混合物样本 (ViralPool) 和两种加标样本 (CoVOC43_95UHR 和 ViralPool_95UHR) 的 cDNA 进行文库制备，以备上机测序使用。Nextera Flex for Enrichment 酶切片段化推荐的总 DNA 起始量为 10–1000 ng。

扩增完成后，根据样本类型 (CoVOC43、CoVOC43_95UHR、ViralPool 和 ViralPool_95HR) 将样本分为不同的混合文库并进行富集。使用两个探针池中的一个来进行富集反应（表 3）。第一个探针池包含 Pan-Viral Panel (Twist Biosciences, 货号 100516)，另外还添加了额外的寡核苷酸来检测 SARS-CoV-2。寡核苷酸的设计使用 80 bp 的叠瓦式探针覆盖了完整的 SARS-CoV-2 基因组序列（可在 NCBI 上获取，MN908947.3）（表 3）。

第二个探针池是 Respiratory Virus Oligos Panel (Illumina, 货号 20042472)，约有 7800 个探针，这些探针能检测呼吸道病毒、最近的流感病毒株和 SARS-CoV-2，其中还包含作为阳性对照的人类探针（表 3 和表 6）。CoVOC43 和 CoVOC43_95UHR 混合样本使用两个 panel 在两个单独的富集反应中进行富集。ViralPool 和 ViralPool_95UHR 混合样本仅使用 Respiratory Virus Oligos Panel 进行富集。富集后（图 2），将制备好的文库进行定量、混合并加载到 MiSeq™ 系统中进行测序。

表 3：用于病毒富集的探针池

探针池	组成
Pan-Viral/ SARS-CoV-2	Pan-Viral Panel 以及用于检测 SARS-CoV-2 的探针
Respiratory Virus Oligos Panel	试剂盒中的探针可检测呼吸道病毒、当前的流感病毒株和 SARS-CoV-2，还包含作为阳性对照的人类探针

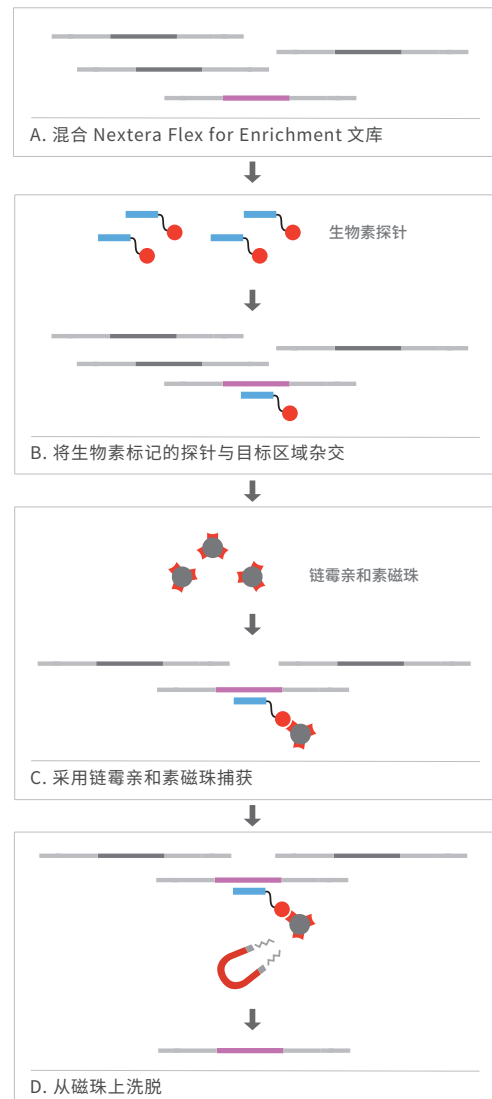


图 2：富集化学技术 - 制备好的文库先用病毒靶向试剂盒进行富集，然后再进行扩增、定量和测序。

测序

制备好的文库可在任何 Illumina 仪器上测序。这些样本对 reads 数要求不高，因此特别适合在台式 iSeq™ 100、MiniSeq™ 和 MiSeq 系统上测序。本研究根据 MiSeq 系统文库变性与稀释指南（文件号 15039740 v10），对病毒样本制备的文库进行了变性，并稀释至最终上样浓度（10 pM），在 MiSeq 系统上使用 MiSeq v3 试剂以 2×151 bp 的读长[†]进行了测序。在探针杂交之前，也对所有文库的等分试样进行测序，以确定富集反应的倍数变化。

病毒滴度、RNA 质量和每个样本的 reads 数会影响获得的病毒特异性 reads 数和病毒基因组的覆盖度。作为通用指南，此工作流程推荐的 reads 数是每个样本 500k reads，该数字只是推荐起点，是可调整的。

数据分析

对于自动化机载分析，可使用重测序分析模块在 Local Run Manager (LRM) 中设置测序运行。该模块允许输入所有运行信息和参考基因组选择，用于后续序列比对。用户可以将参考基因组直接上传到仪器，易于自定义设置。测序完成后分析会自动开始，以使用户能尽快解读结果。重测序模块提供了比对、覆盖度和小变异数据，以及 FASTQ、BAM 和 VCF 文件，可以在其他数据分析流程中使用（如有需要）（图 3）。

IDbyDNA Explify 平台为深度的数据分析提供了易于使用的解决方案，该方案提供强大的数据质量控制（QC）、标准化的结果解读、精心管理的数据库和定制的报告生成。数据分析基于 k-mers 和比对步骤，包括病毒的蛋白水平检测，这能提高识别新型病毒和高度差异化病毒的能力。此平台可在 BaseSpace™ Sequence Hub 中使用。

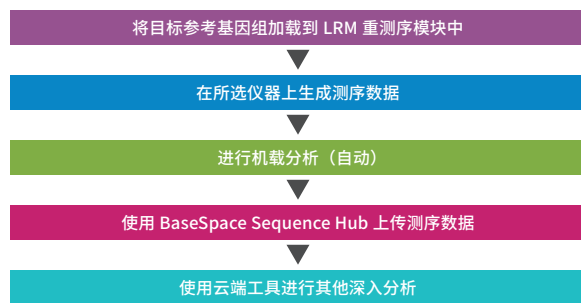


图 3：富集分析工作流程 - 测序运行在 LRM 中设置，运行完成后数据分析会自动开始。如有需要，可使用第三方应用进行其他分析。

结果

完成文库制备和测序后，使用 LRM 重测序模块（v2.5.56.11）将每个样本与冠状病毒 OC43 的参考基因组进行比对，具体步骤参见《Local Run Manager 重测序分析模块工作流程指南》（文件号 1000000002705 v01）。

[†] Nextera Flex for Enrichment 实验方案建议至少使用 2×101 bp 的读长。

使用 LRM 重测序模块评估逆转录和 CoV 检测

本研究使用了 2 种不同的逆转录试剂盒来生成用于文库制备的 cDNA：一个来自 Thermo Scientific 的商业化试剂盒和一个来自 Illumina TruSeq RNA 试剂的工作流程（详细信息参见“方法”部分）。对于 CoVOC43 样本，Illumina 逆转录方法获得的序列与参考基因组的匹配率为 >99%，Thermo Scientific 方法获得的序列与参考基因组的匹配率为 >97%，两种病毒试剂盒的水平类似（图 4）。Respiratory Virus Oligos Panel 和 Pan-Viral/SARS-CoV-2 Panel 对加标样本的匹配率分别为 >64% 和 >19%，两种逆转录方法的结果类似（图 4）。

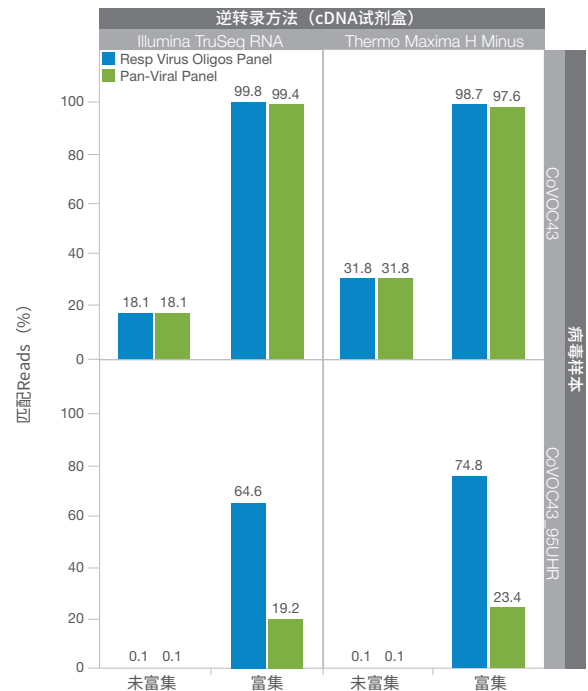


图 4：不同 cDNA 试剂盒的性能类似 - 图中展示了使用和未使用 Respiratory Virus Oligos Panel 和 Pan-Viral Panel 进行富集的情况下，CoVOC43 和 CoVOC43_95UHR 病毒样本通过两种不同的 cDNA 试剂盒（参见“方法”部分）来进行逆转录时的匹配 Reads 百分比。

我们比较了富集前后样本的匹配 Reads 百分比，计算了富集的倍数变化，将其作为衡量靶向捕获和杂交反应是否成功的指标。结果表明，无论使用哪种逆转录方法，CoV 样本的富集倍数都有显著的增加。尤其是 CoVOC43_95UHR 样本，它在 Respiratory Virus Oligos Panel 中展示了 >640 倍的富集，而在 Pan-Viral Panel 中展示了 >190 倍的富集（表 4）。

平均覆盖度提供了参考基因组中每个碱基覆盖深度的概况，能识别出可能未被测序的区域。无论使用哪种逆转录方法，使用病毒靶向探针进行富集都极大地提高了所有样本中病毒序列的平均覆盖度（表 5）。

表 4：使用 LRM 重测序模块获得的富集指标

Respiratory Virus Oligos Panel				
Illumina TruSeq RNA		Thermo Maxima H Minus		
富集倍数	CoVOC43	CoVOC43_95UHR	CoVOC43	CoVOC43_95UHR
	5.5x	646x	3.1x	748x
Pan-Viral/SARS-CoV-2 Panel				
Illumina TruSeq RNA		Thermo Maxima H Minus		
富集倍数	CoVOC43	CoVOC43_95UHR	CoVOC43	CoVOC43_95UHR
	5.5x	192x	3.1x	234x

表 5：使用 Respiratory Virus Oligos Panel 时 CoVOC43 的平均覆盖度

cDNA 试剂盒	CoVOC43		CoVOC43_95UHR	
	未富集	富集	未富集	富集
Illumina TruSeq RNA	427.1	3283.5	2.6	2437.3
Thermo Maxima H Minus	987.0	3861.4	4.2	3035.0

平均覆盖度参数已归一化为每个样本 1M reads。

使用 IDbyDNA Explify 进行分析

Explify 平台鉴定出了加入的冠状病毒 (图 5) 和病毒混合物中的全部 5 种病毒 † (图 6)。Explify 平台提供了病毒共有基因组序列、覆盖度曲线, 具有检测与其他病毒、细菌、真菌或寄生虫合并感染的能力 (图 5 和图 6)。

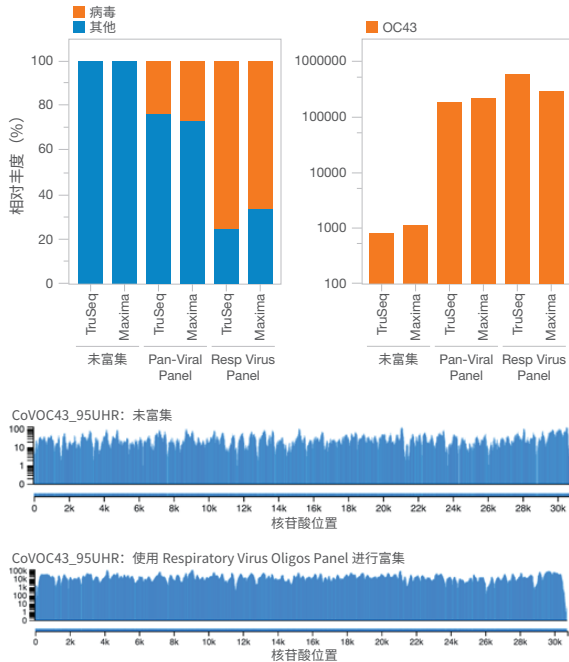


图 5：OC43 的鉴定 - 无论使用何种 cDNA 试剂盒或病毒探针池, IDbyDNA 都能以出色的病毒基因组覆盖度鉴定添加到 UHR RNA 中的 OC43。数据被降采样至每个样本 1M reads。

† 病毒混合物未采用正交技术来验证, 因此不能保证每种病毒的占比相同。

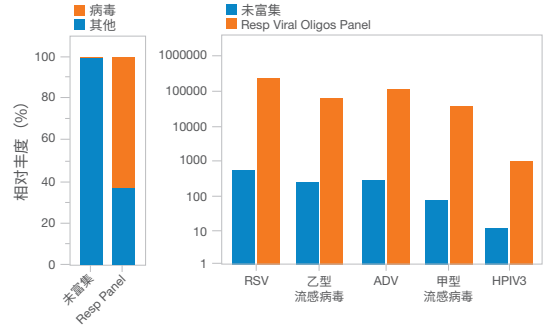


图 6：病毒混合物的鉴定 - 使用 Respiratory Viral Oligos Panel 富集后, IDbyDNA 鉴定出了病毒混合物中所有的 RNA 和 DNA 病毒株。数据被降采样至每个样本 1M reads。

Explify 平台也证明了 Pan-Viral/SARS-CoV-2 Panel 和 Respiratory Virus Oligos Panel 的 CoVOC43_95UHR 富集数据与鸟枪法宏基因组测序数据具有同等的覆盖度⁴ (图 7), 证明了使用病毒靶向探针池的富集工作流程与鸟枪法相比没有引入覆盖度偏差。

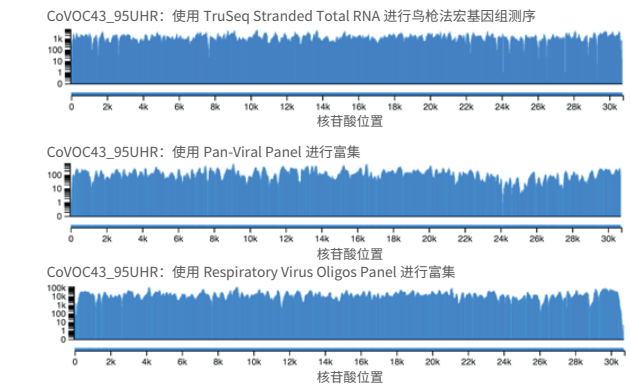
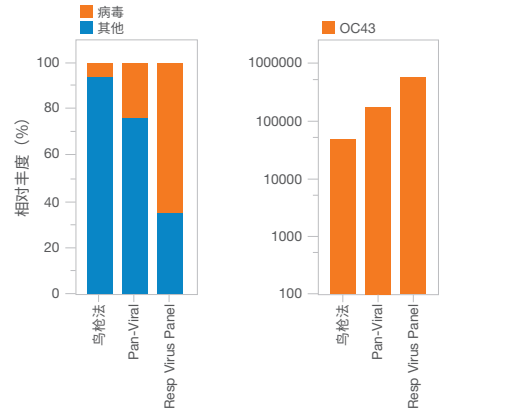


图 7：与鸟枪法的性能类似 - 使用病毒试剂盒的富集显示了与鸟枪法宏基因组测序同等的性能和覆盖度。

表 6 : Respiratory Virus Oligos Panel 的靶向病毒

人冠状病毒 229E	人副流感病毒 1 型
人冠状病毒 NL63	人副流感病毒 2 型
人冠状病毒 OC43	人副流感病毒 3 型
人冠状病毒 HKU1	人副流感病毒 4a 型
SARS-CoV-2	人类偏肺病毒 (CAN97-83)
人腺病毒 B1 型	呼吸道合胞病毒 (A 型)
人腺病毒 C2 型	人类呼吸道合胞病毒 9320 (B 型)
人腺病毒 E4 型	甲型流感病毒 (A/Puerto Rico/8/1934(H1N1))
人博卡病毒 1 型 (灵长类博卡病毒 1 型分离株 st2)	甲型流感病毒 (A/Korea/426/1968(H2N2))
人博卡病毒 2c 型 PK 分离株 PK-5510	甲型流感病毒 (A/New York/392/2004(H3N2))
人博卡病毒 3 型	甲型流感病毒 (A/goose/Guangdong/1/1996(H5N1))
人博卡病毒 4 型 NI 病毒株 HBoV4- NI-385	甲型流感病毒 (A/Zhejiang/DITD-ZJU01/2013(H7N9))
KI 多瘤病毒 Stockholm 60	甲型流感病毒 (A/Hong Kong/1073/99(H9N2))
WU 多瘤病毒	甲型流感病毒 (A/Texas/50/2012(H3N2))
人双埃柯病毒 1 型 PicoBank/HPeV1/a	甲型流感病毒 (A/Michigan/45/2015(H1N1))
人双埃柯病毒 6 型	乙型流感病毒 (B/Lee/1940)
人鼻病毒 A89	乙型流感病毒 (B/Wisconsin/01/2010)
人鼻病毒 C 型 (病毒株 024)	乙型流感病毒 (B/Brisbane/60/2008)
人鼻病毒 B14	乙型流感病毒 (B/Colorado/06/2017)
人类肠病毒 C104 病毒株 : AK11	乙型流感病毒 (B/Washington/02/2019)
人类肠病毒 C109 分离株 NICA08-4327	人类对照基因

总结

对新发病毒进行鉴定与生物学特性研究是改善公众健康的核心工作。在这种场景中, NGS 是一种合适且强大的方法, 可以进行广泛的检测来识别已知病毒和新发病毒。利用 Nextera Flex for Enrichment 和靶向病毒病原体的试剂盒能让研究人员获得可确认 CoV 是否存在的基因组数据, 并继续进行基因分型和变异分析等进一步的解析。不可知论的设计能让研究人员在所有感兴趣的样本类型中广泛识别致病性病毒, 并且使用唯一双标签序列来降低多重样本标签信号串扰的风险。这种简单易行的工作流程采用了成熟的 Illumina 测序技术, 使研究人员在疫情爆发时能快速发现并鉴定病原体, 如此次 SARS-CoV-2。

了解更多

如需深入了解病毒测序方法, 请访问 www.illumina.com/areas-of-interest/microbiology/infectious-disease-surveillance.html

如需深入了解 LRM 分析, 请访问 www.illumina.com/products/by-type/informatics-products/local-run-manager.html

参考文献

1. World Health Organization. WHO Director-General's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV). [www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihf-emergency-committee-on-novel-coronavirus-\(2019-ncov\)](http://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihf-emergency-committee-on-novel-coronavirus-(2019-ncov)). Accessed January 30, 2020.
2. Bulcha B. Review on viral metagenomics and its future perspective in zoonotic and arboviral disease surveillance. *J Biol Agric Healthcare*. 2017;7(21):35-41.
3. Gaudin M and Desnues C. Hybrid Capture-Based Next Generation Sequencing and Its Application to Human Infectious Diseases. *Front Microbiol*. 2018;9:2924.
4. Illumina (2020) Comprehensive workflow for detecting coronavirus using Illumina benchtop systems. Accessed March 19, 2020.

Illumina 中国

上海办公室 · 电话 (021) 6032-1066 · 传真 (021) 6090-6279

北京办公室 · 电话 (010) 8455-4866 · 传真 (010) 8455-4855

技术支持热线 400-066-5835 · techsupport@illumina.com · www.illumina.com.cn

© 2020 Illumina, Inc. 保留所有权利。所有商标均为 Illumina 公司或其各自所有者的财产。关于具体的商标信息, 请访问 www.illumina.com/company/legal.html. 1270-2020-002-A QB9563

仅供研究使用。不得用于诊断。



Illumina Academy @illumina

illumina®

1270-2020-002-A | 5