



PARAGON GENOMICS

快速简易的全序列新冠病毒测序及 新病源体发现

刘至东 PhD, CTO

Paragon Genomics, CA, USA

Feb. 17, 2020

快速简易的全序列新冠病毒测序及新病源体发现

内容

- 探讨用简单快速的超高超洁多重PCR技术检测新冠病毒，以提高灵敏度的方法。
- 介绍可以用来发现新冠病毒或未知病源体的简易宏基因组方法。

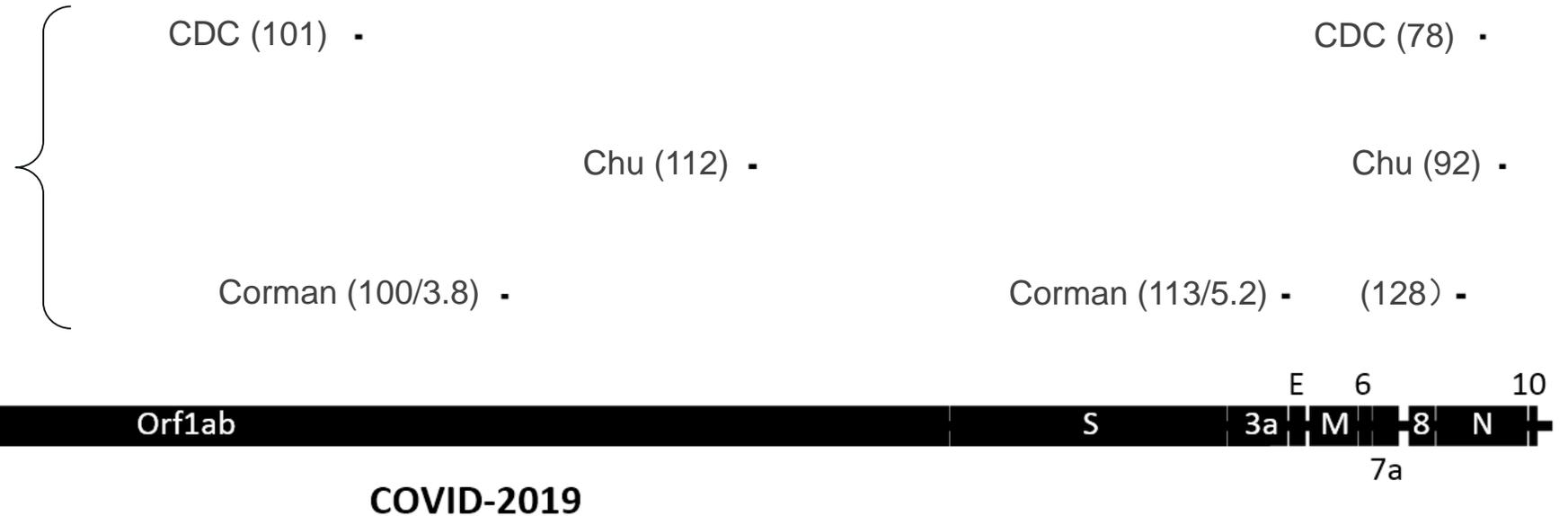
背景介绍

- Paragon Genomics拥有实现简单快速的超高多重PCR，并同时降低或消除多重PCR背景噪音的专利技术。
- 有为美国CDC和Broad Institute订制检测细菌和病毒的超高超洁多重PCR技术的经验。
- 愿以积累的技术，经验和力量，与国内同仁共同攻克检测新冠病毒的难题。

一，探讨用多重PCR技术检测新冠病毒的方法

RT-PCR是目前最主要的检测新冠病毒的方法

已经报导的RT-PCR扩增子位置，大小 (bp) 及最低可检病毒数。



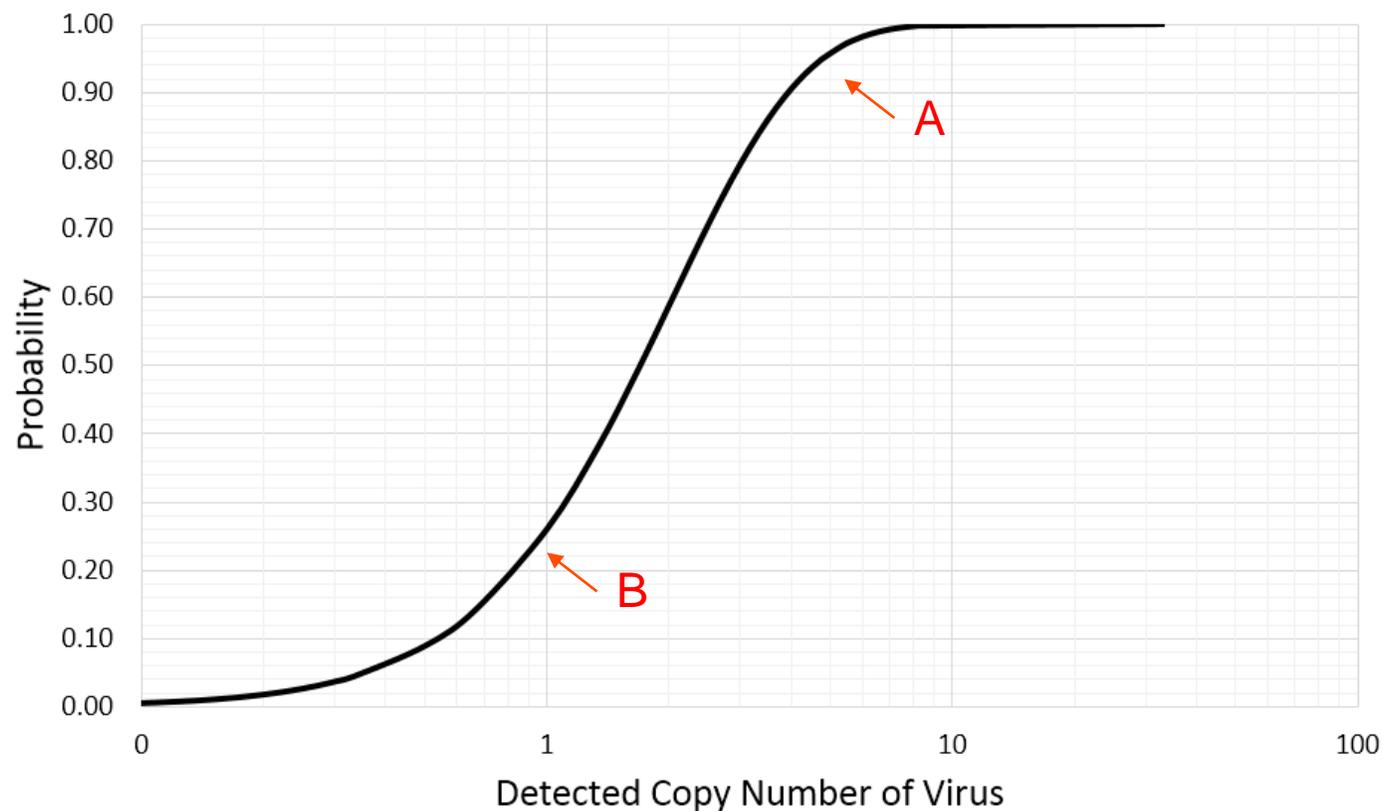
已经报导的RT-PCR扩增子平均长度为103bp，处于ORF1ab及S基因区域的扩增子可在95%几率下检测到4-5个病毒颗粒，处于N基因区域的扩增子(Corman)表现不良。在RT-PCR中，一个扩增子对应于一个新冠病毒。由此可以看出，如果用多个扩增子来检测新冠病毒，可以提高灵敏度。

Chu, DKW, et al, Clinical Chemistry, hvaa029, <https://doi.org/10.1093/clinchem/hvaa029>.

Corman, VM, et al, Euro Surveill. 2020;25(3):pii=2000045. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>.

CDC: 新型冠状病毒感染的肺炎实验室检测技术指南。

在已经报导的条件下用RT-PCR检测新冠病毒的理论极限值



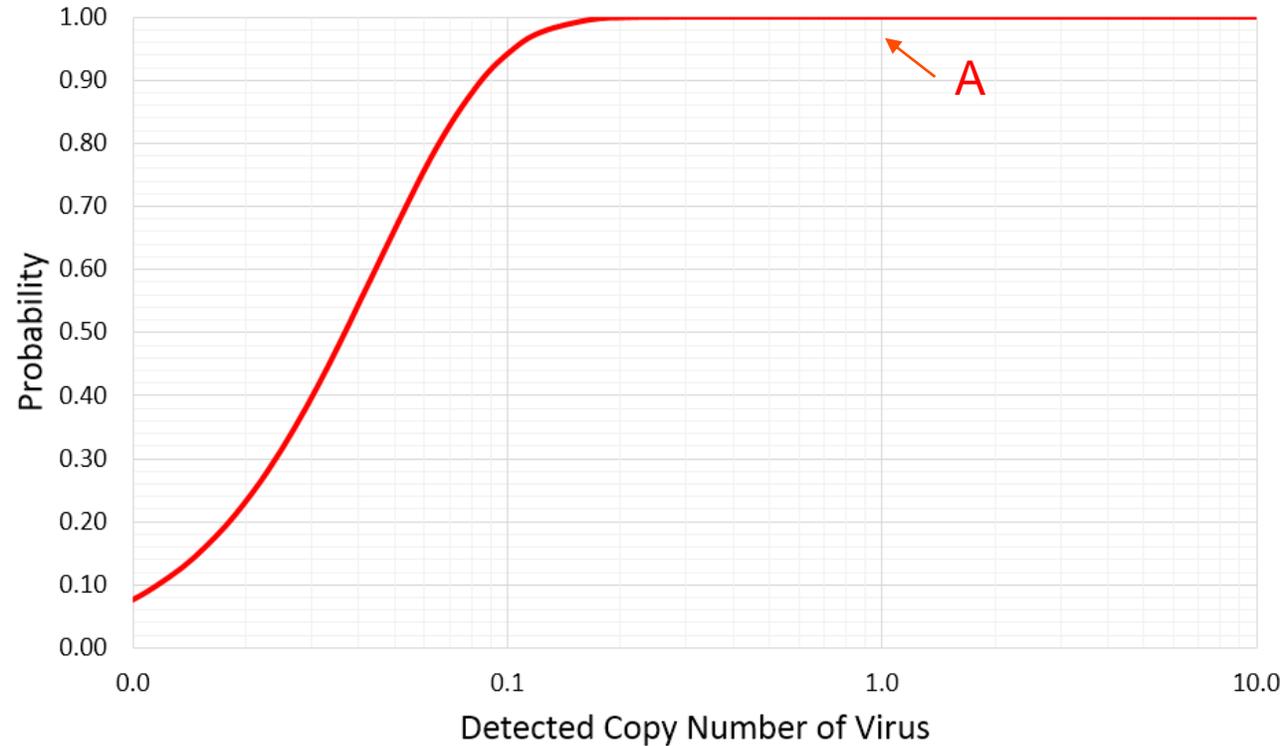
PCR可视为一个取样过程，其结果可用Poisson建模。取平均长度为100bp的扩增子，检测30Kb长的新冠病毒，RT-PCR的扩增效率为99%，理论上在95%几率下可检测到4.8个病毒颗粒(A)。对于1个病毒颗粒的极限，仅有26%几率可检测到(B)。在实际检测操作中，因病毒RNA降解，逆转录反应的效率过低等因素，可能会使入检的病毒更低。对于优化RT-PCR方法，建议参考循环逆转录(Sheng, K., et al, Nat Methods. 2017 Mar;14(3):267-270. doi: 10.1038/nmeth.4145)或许对提高灵敏度有所帮助。

我们提出用多重PCR扩增新冠病毒的cDNA以提高检测灵敏度



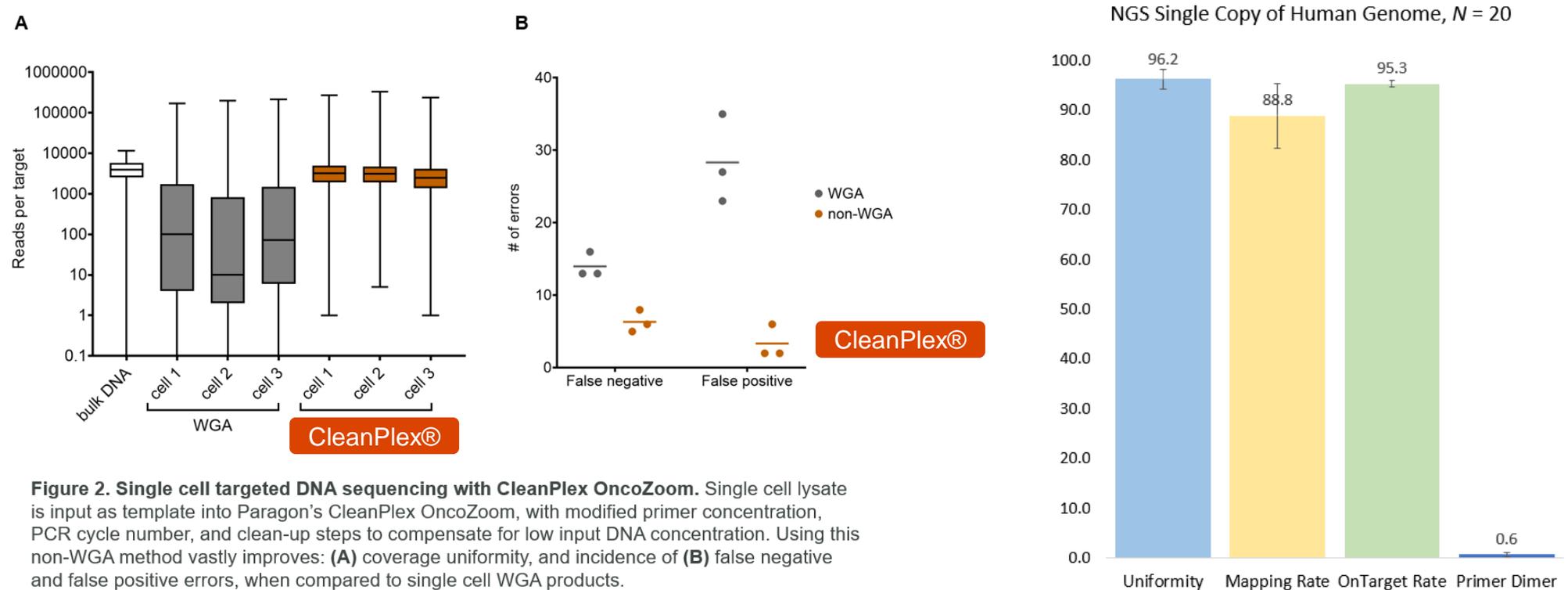
与其用RT-PCR扩增一个目标，我们提出用多重PCR扩增新冠病毒的全部cDNA。我们设计了覆盖全部新冠病毒基因组的扩增子。这些扩增子的平均长度为100bp，分为2个相互重迭的扩增子池，总数为343个。多重PCR扩增之后产生特征性的DNA文库峰形，并可进一步在华大MGI测序仪上核实新冠病毒的核酸序列。多重PCR方法不受病毒核酸序列变异的影响，但除可检测到COVID-2019之外，还可检测到SARS病毒，SARS-like病毒等，需用NGS测序来弥补其特异性的不足。

用多重PCR检测新冠病毒的理论极限值



取平均长度为100bp的343个扩增子，检测30Kb长的新冠病毒，多重PCR的扩增效率为24%，同时有21%的扩增子不能实现扩增，理论上在100%几率下可检测到1个病毒颗粒(A)。（如此高的灵敏度表明没有必要使用343个扩增子。在此条件下，用50个扩增子即可检测到一个新冠病毒。但是，实际应用要求更加苛刻的条件。）

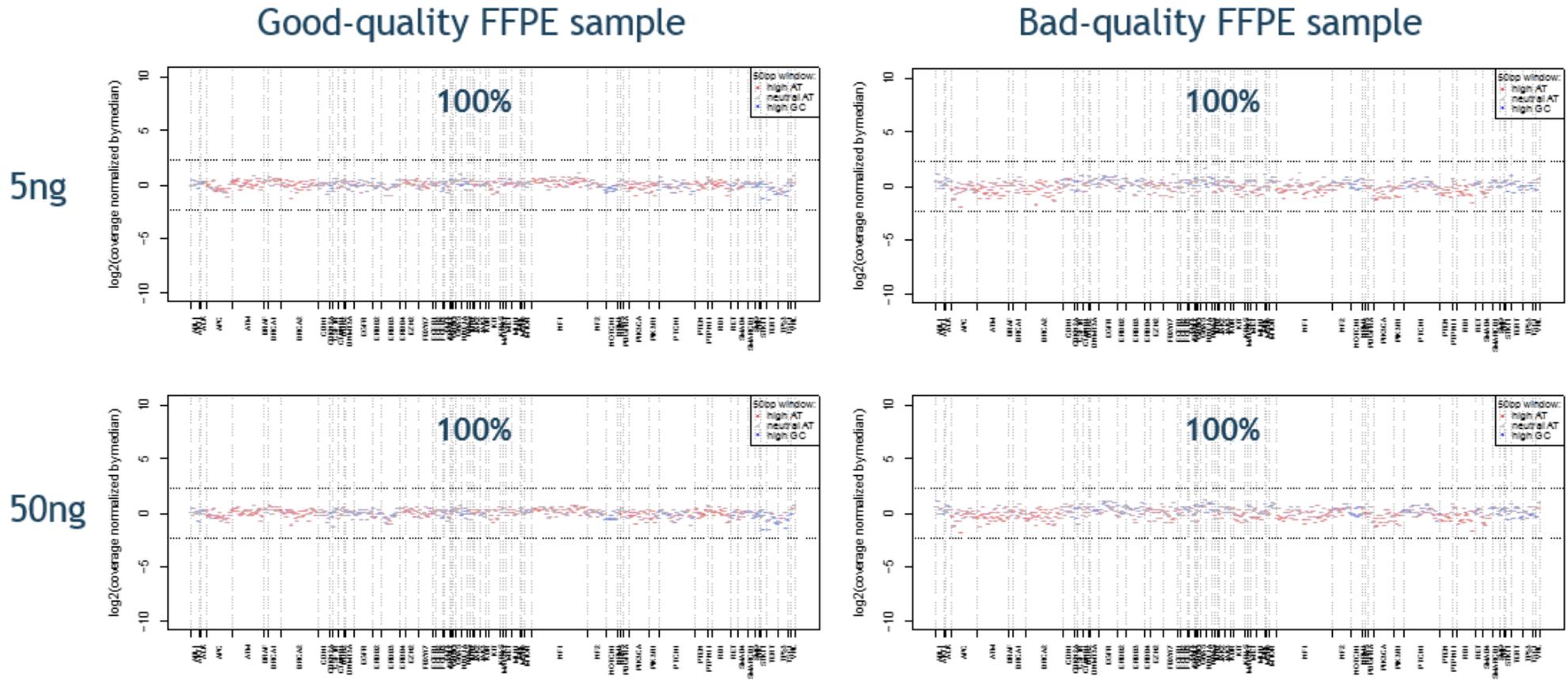
多重PCR能够用来扩增一个COPY的基因组吗？



我们与RareCyte, Mayo Clinic证明, 用Paragon Genomic的CleanPlex®多重PCR技术可以直接扩增一个细胞中的一个基因组上的600个扩增子。经过NGS测序, 发现均一性, 假阳性, 假阴性均优于WGA法, 并成功检测到79%的预知目标。但是, 一个细胞中有4条DNA链, 而新冠病毒基因组只有1条cDNA链。由此, 我们假设扩增新冠病毒基因组的效率只有6% (24% ÷ 4)。

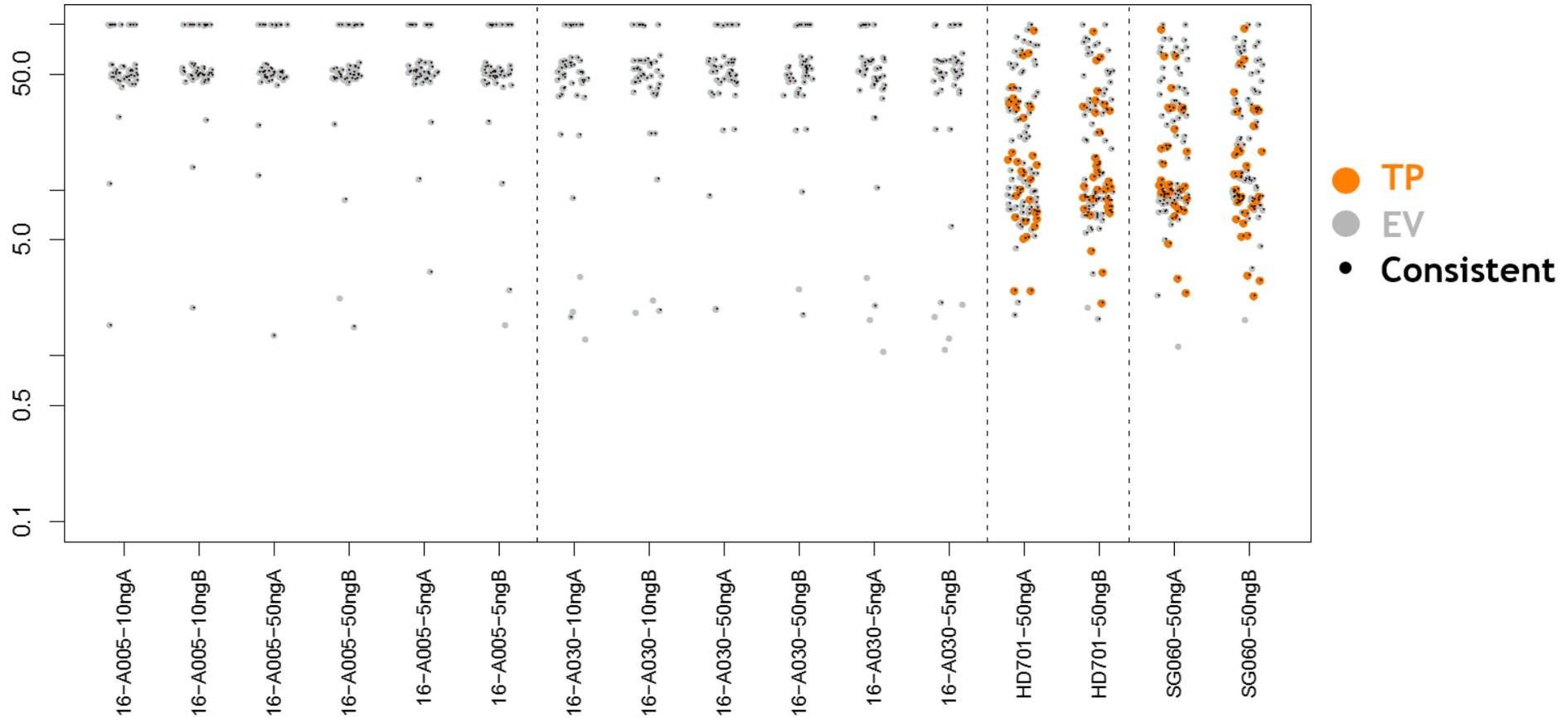
Ericson, NG, et al, Amplicon-based targeted sequencing of single circulating tumor cells. <http://rarecyte.com/publicationsandposters>.

CleanPlex®多重PCR技术可以低质量的临床样品吗？



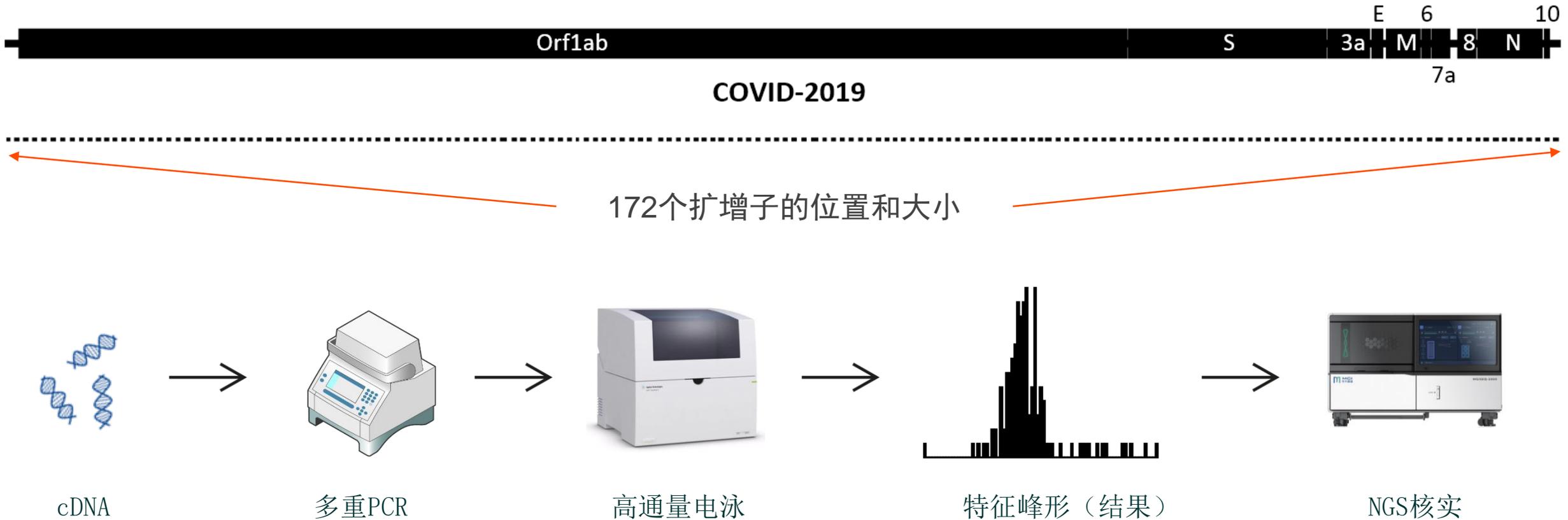
Sophia Genetics证明，用CleanPlex®多重PCR技术可以直接扩增低质量和低含量的FFPE DNA样品中的600个扩增子，经过NGS测序，发现其均一性与高质量和高含量的FFPE DNA样品相近。在上图中，每一个点代表一个扩增子，纵坐标显示每个扩增子的数目与平均数的距离，横坐标显示每个扩增子的标记。

CleanPlex®多重PCR技术用于临床样品的高重复性



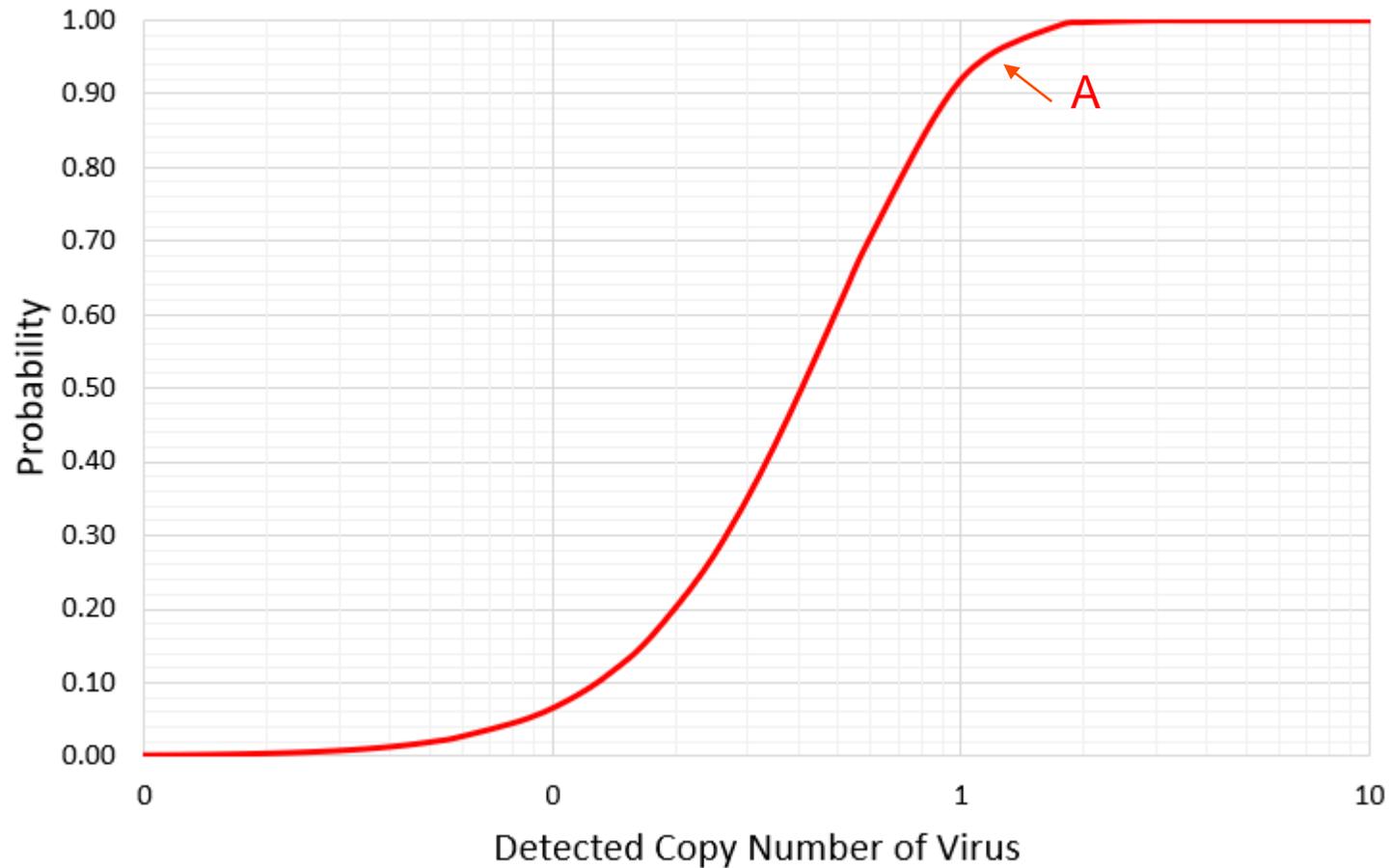
Sophia Genetics同时证明，用CleanPlex®多重PCR技术扩增5-50ng FFPE临床样品中的600个扩增子，可以达到极高的可重复性。在上图中，每一个点代表一个检测到的核酸变异，纵坐标显示每个核酸变异的频率，横坐标显示每个样品的标记。

用一个引物池多重PCR扩增新冠病毒的一半基因组



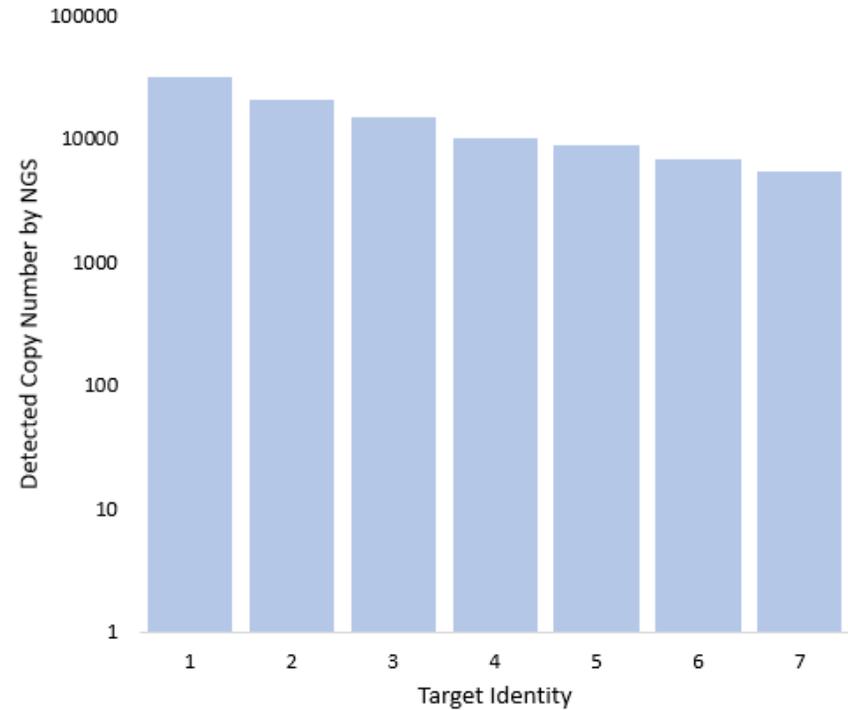
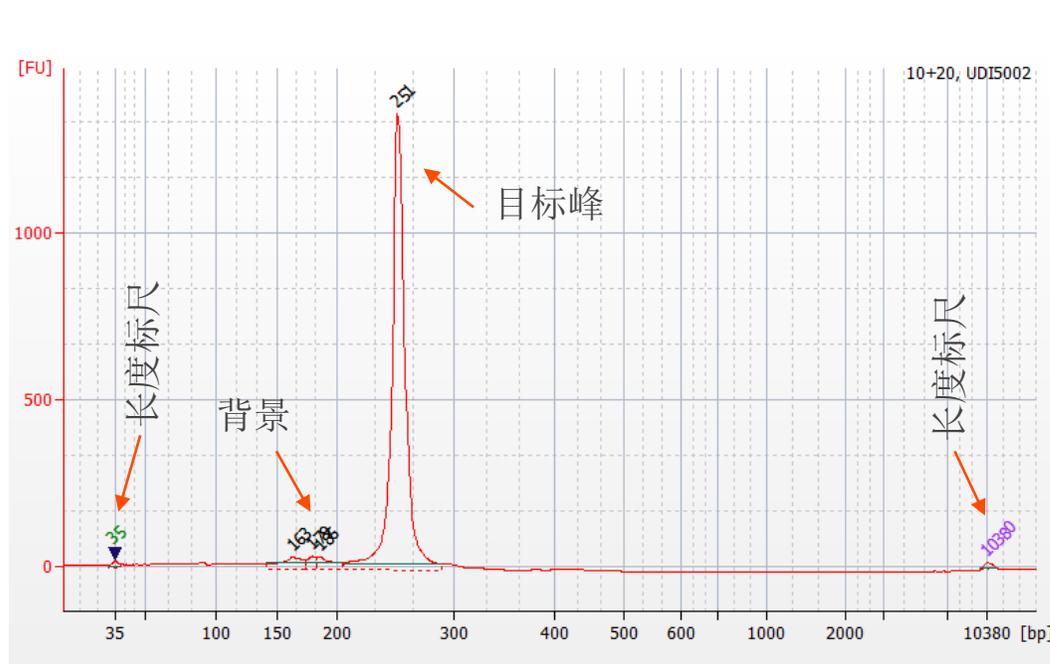
用一个引物池多重PCR扩增新冠病毒的一半基因组，可以使整个操作流程更加简易块速。从多重PCR到电泳结果约需4小时。阳性样品出现峰值为230bp的特征峰形，正常样品无扩增产物DNA或有少量背景扩增产物。阳性及可疑样品可进一步集中测序核实。

用一个引物池多重PCR检测新冠病毒的理论极限值



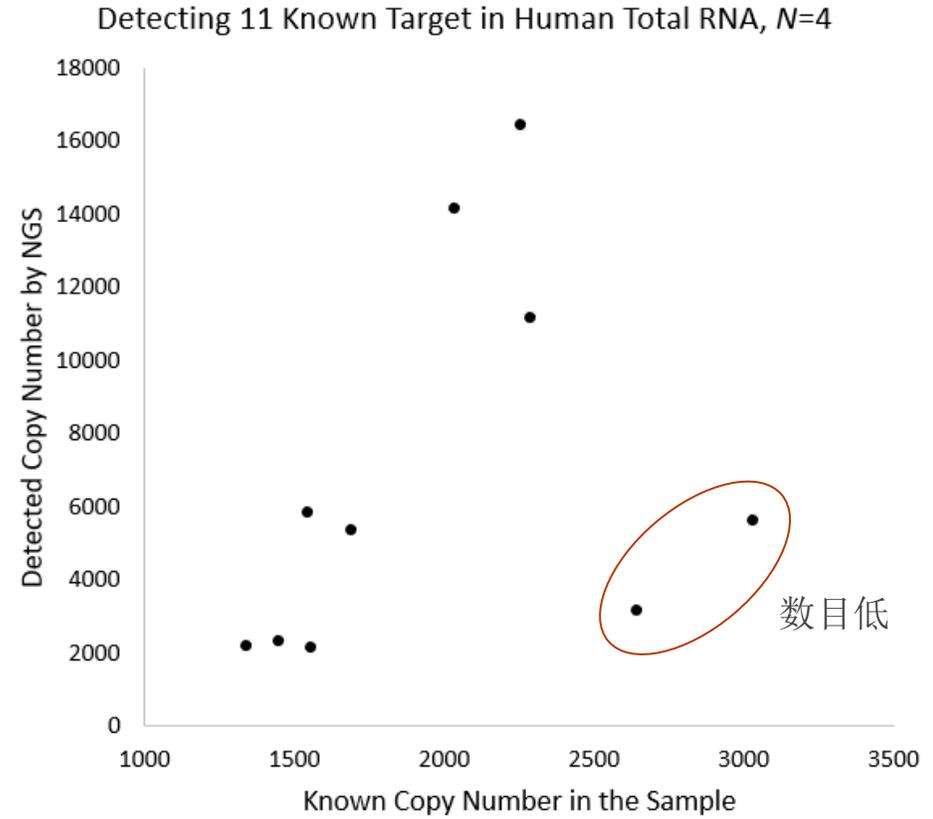
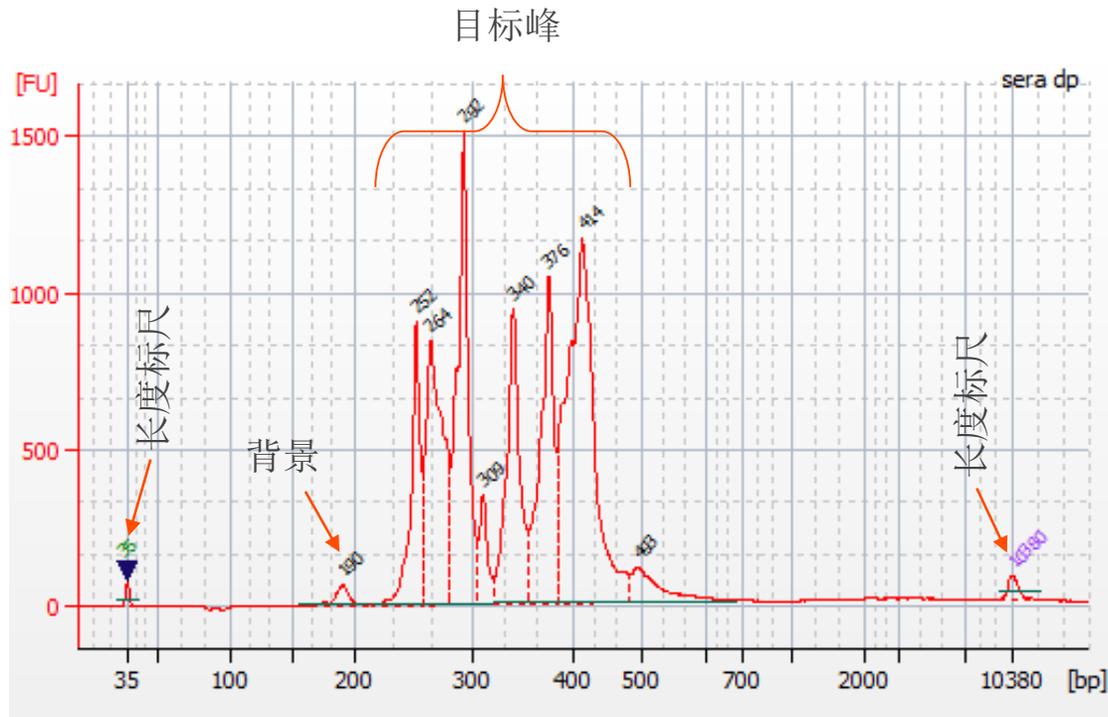
取平均长度为100bp的172个扩增子，检测30Kb长的新冠病毒，多重PCR的扩增效率为6%，同时有20%的扩增子不能实现扩增，理论上在95%几率下可检测到1.15个病毒颗粒(A)。

验证用多重PCR扩增人类全RNA片段中的7个特定目标



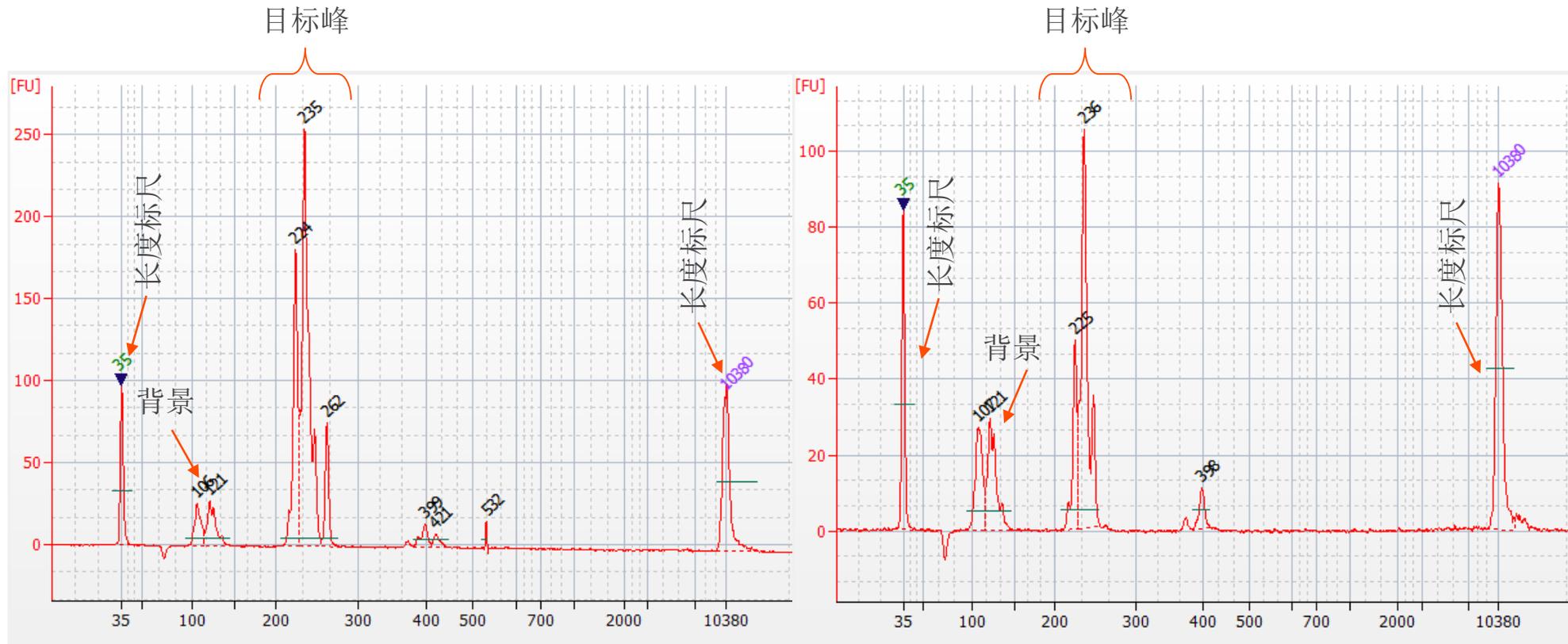
由于需要从提取的全RNA开始进行新冠病毒的检测，而全RNA含有大量的人类RNA和rRNA，新冠病毒的RNA只占极少数，这些人类RNA和rRNA可能造成很高的非特异性背景噪音。我们验证用多重PCR扩增人类全RNA中的特定目标的可行性。我们用Random Hexamer逆转录Agilent Human Reference RNA合成cDNA，选用一个含有97个扩增子的引物池，加入7个特定的人类RNA中的目标。由左图看到多重PCR并未可能造成显著的非特异性背景噪音。右图显示NGS测序核实的7个（且只有7个）特定目标。

验证用多重PCR扩增人类全RNA片段中的11个已知目标



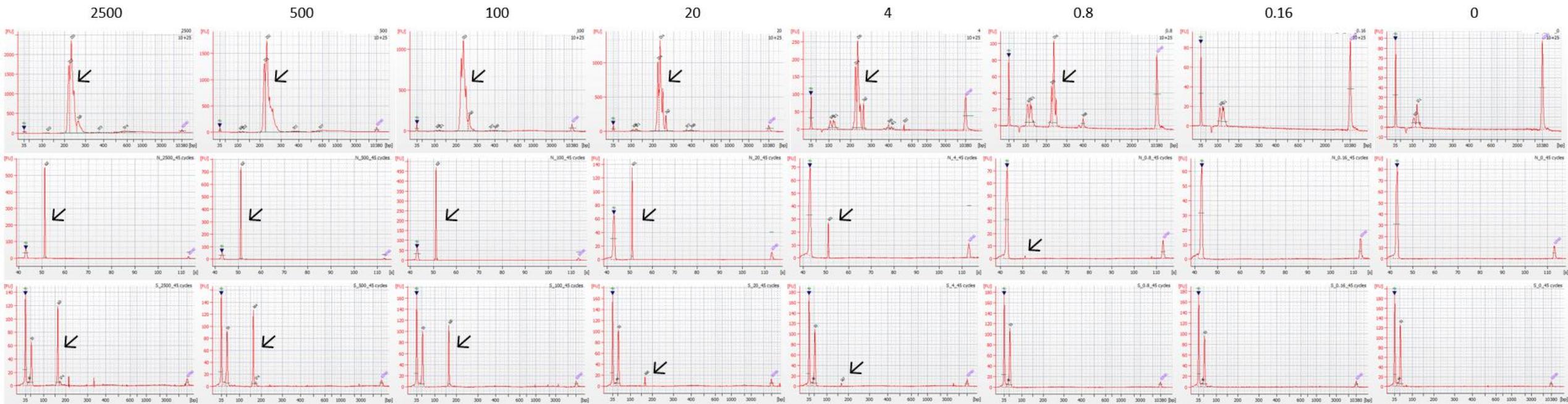
我们再次验证用多重PCR扩增人类全RNA中的特定目标的定量的可行性。我们在引物池中加入11个扩增子，用SeraSeq® Fusion RNA Mix V4为目标，扩增其中的11个有已知数量（浓度）的特定目标。由左图看到多重PCR未可能造成显著的非特异性背景噪音。右图显示NGS测序核实的11个特定目标的数目与已知数量（浓度）的关系。除2个扩增子数目较低之外，其余9个与已知数量的相关度为92%。由此我们估计，在检测1-2个新冠病毒时，至少有部分扩增子可以实现扩增。

验证用多重PCR-电泳法检测4个和1个新冠病毒



用多重PCR扩增新冠病毒的S和N基因。S和N基因被置于质粒上，质粒数经digital PCR定量。被扩增的扩增子预计为30个，峰值预计为234bp。PCR循环数为35。虽有172对引物参与扩增，并未引起严重的背景噪音。S和N基因拷贝数为1时，特征性的阳性峰型（峰值235bp）仍可扩增并检测。

比较多重PCR-电泳法和RT-PCR检测新冠病毒



Copies	mPCR Noise	N & S by mPCR	N gene by PCR	S gene by PCR
2500	53	121621	106980	15753
500	831	94163	154980	21445
100	1303	49261	92113	13740
20	3619	41856	27652	1786
4	2393	5992	5034	519
0.8	3282	2324	349	0
0.16	2413	12	0	0
0	1937	0.5	0	0

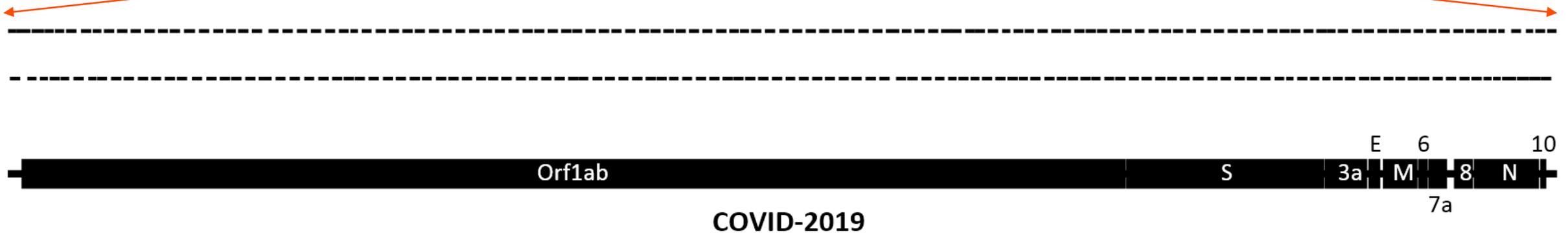
第一排数字为S和N基因的拷贝数，第二排图为多重PCR（循环数为35）电泳结果，第三排图为N基因PCR电泳，第四排图为S基因PCR电泳（循环数均为45），左为浓度列表。箭头示阳性DNA电泳峰型。由此可见：

- 1, 多重PCR-电泳法可以检测1个新冠病毒。
- 2, 如已有报导和数学模型所示，RT-PCR可以检测4个新冠病毒。
- 3, N基因RT-PCR扩增子在一个新冠病毒仍有极小但可见的峰。N基因最小峰在4个新冠病毒。

实验结果持续更新中

用多重PCR检测新冠病毒，亦可在Illumina测序仪上核实

241个扩增子的位置和大小



COVID-2019



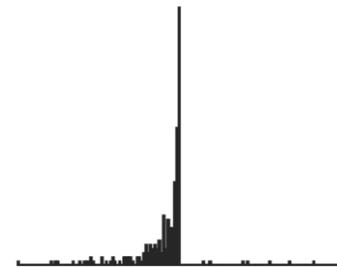
cDNA



多重PCR



高灵敏度电泳



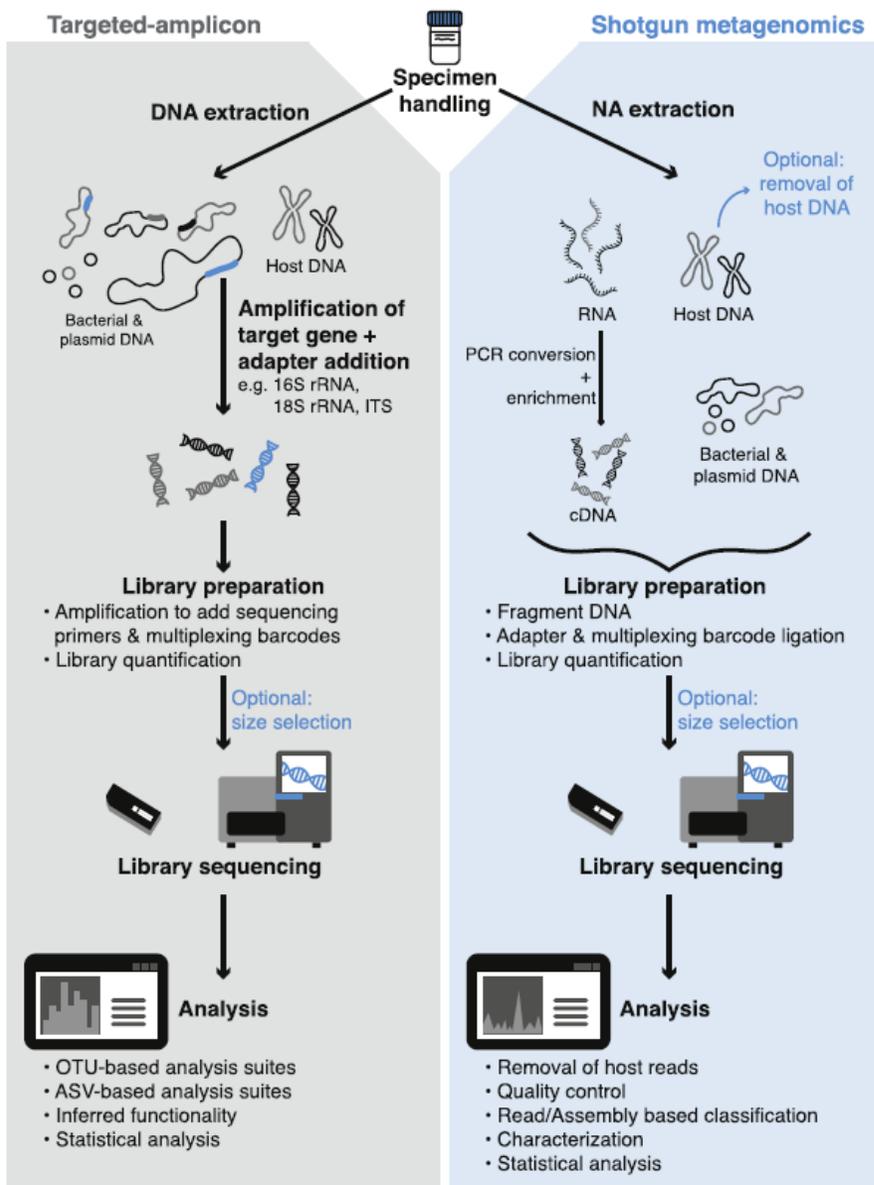
特征峰形（结果）



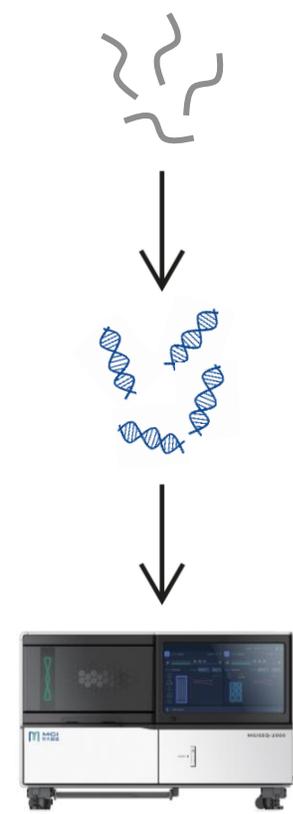
NGS核实

二，发现新型冠状病毒或未知病原体

普通的宏基因组方法



Paragon Genomics多重PCR宏基因组方法



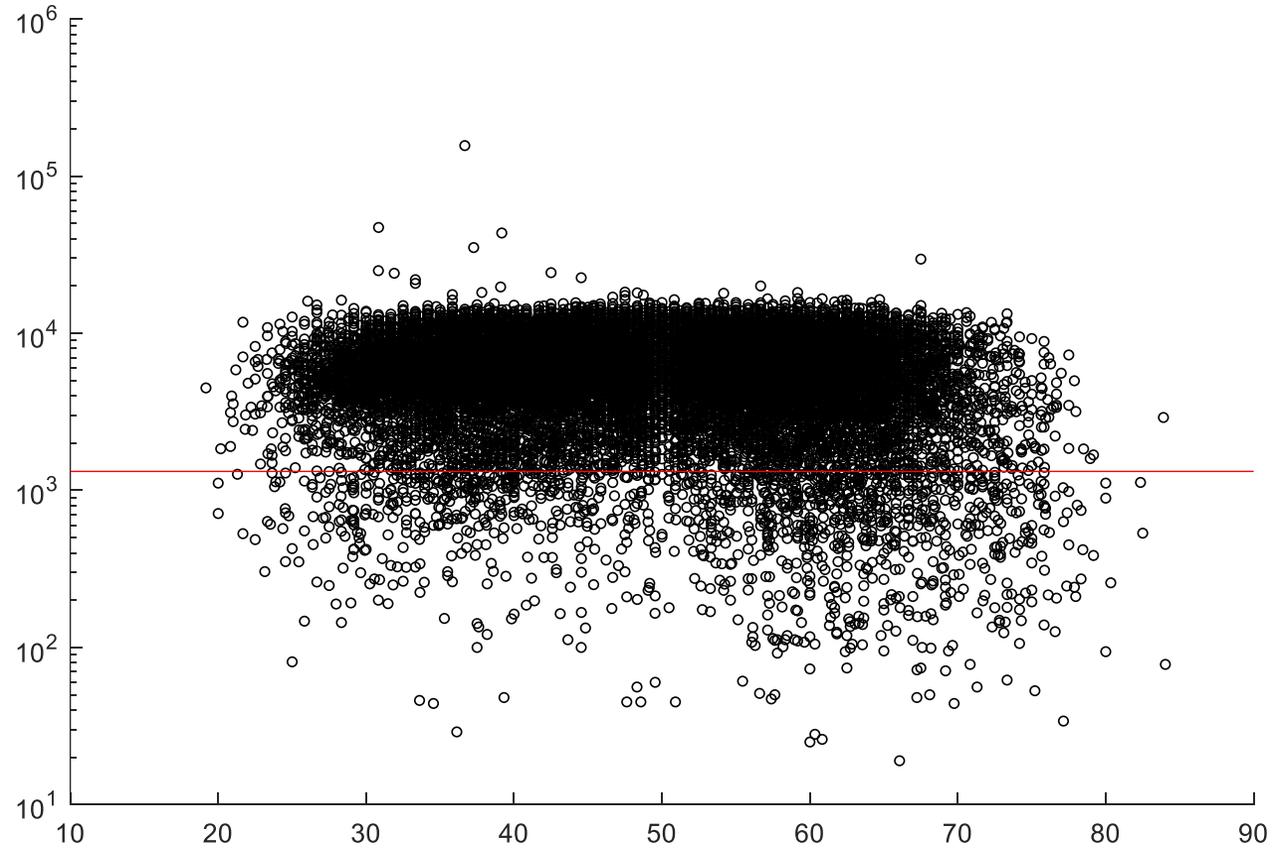
以Paragon方法逆转录RNA生成cDNA，以减少人类rRNA含量（样平中的人类rRNA含量可达90%以上）。

以带通用接头的Random Hexamer实现超高多重PCR

J.D. Forbes et al. / Computational and Structural Biotechnology Journal 16 (2018) 108–120

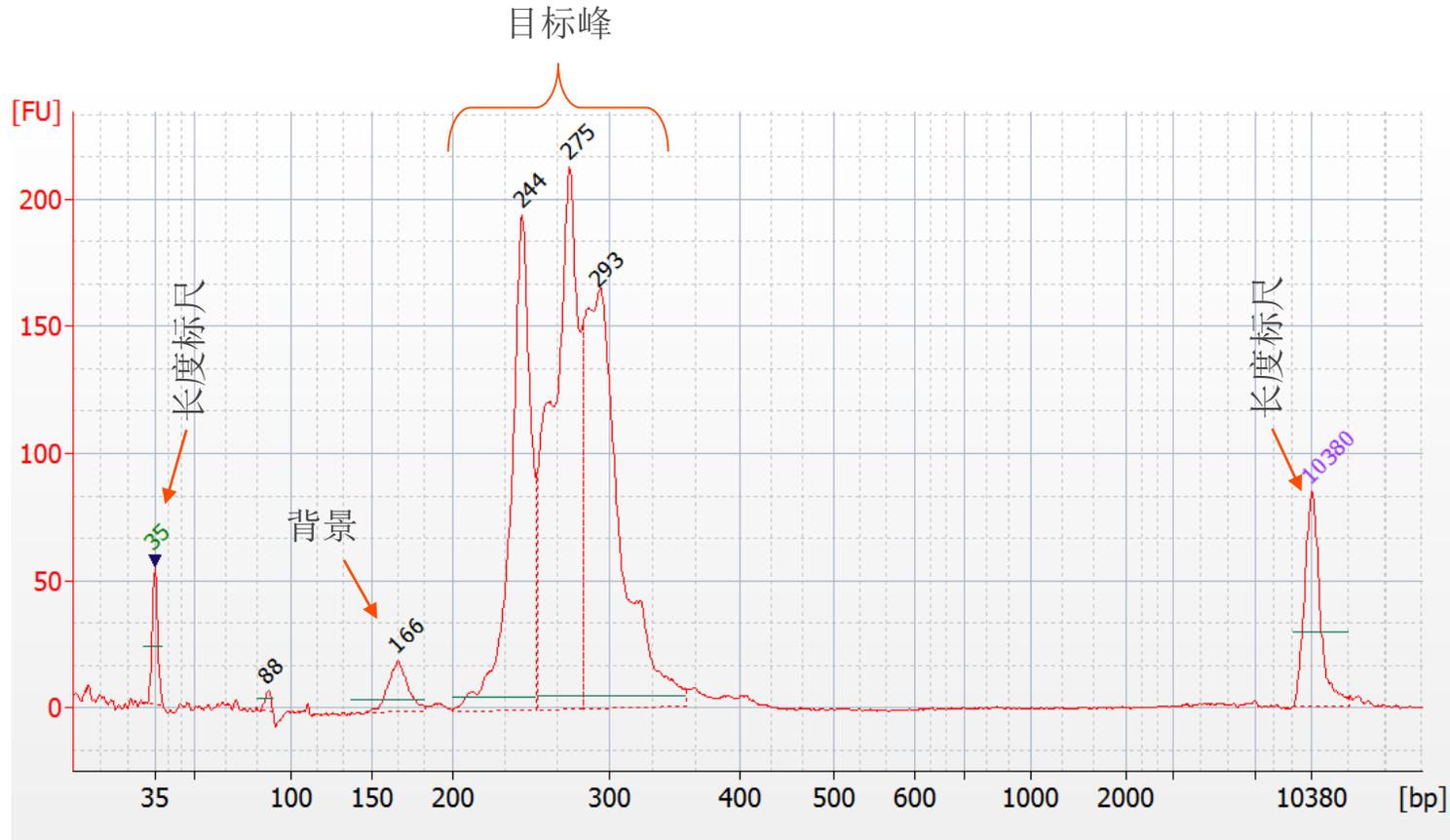


用单管超高多重PCR扩增宏基因组的可行性



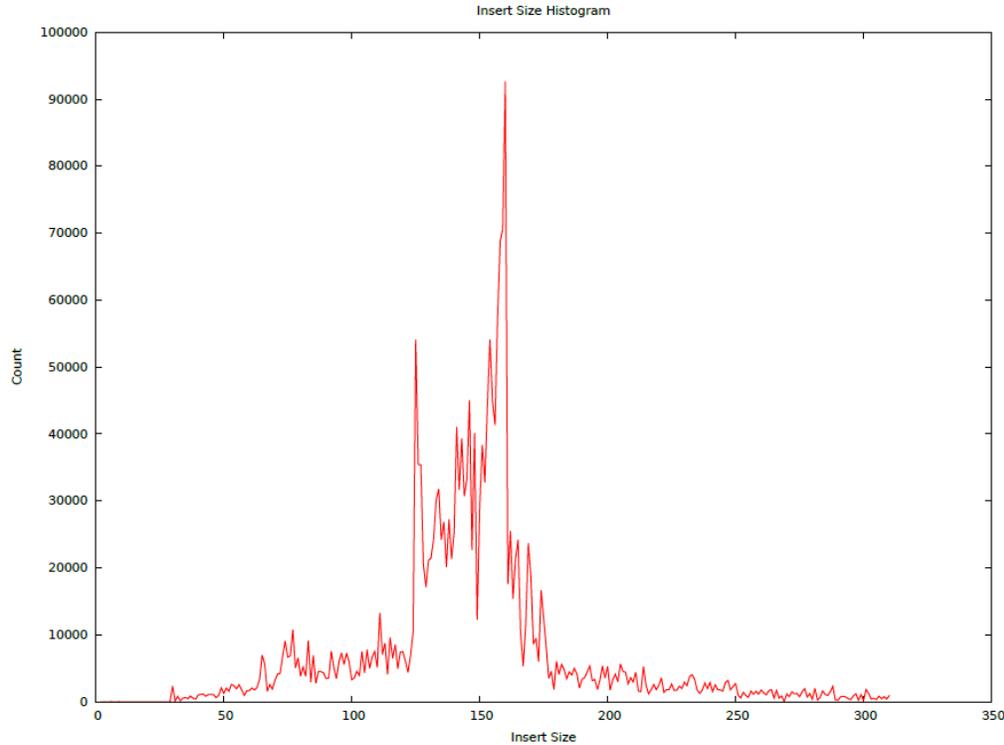
我们用超高多重PCR方法扩增27296个扩增子，以验证用扩增宏基因组的可行性。经NGS测序分析，该方法的均一性(0.2Xmean)为92%。Mapping rate为96%，上靶率为96%，无显著GC Bias，非特异性背景信号为1%。在上图中，每一个点代表一个扩增子，纵坐标显示扩增子的数目，横坐标显示GC含量，图中红线代表0.2Xmean的位置。

用单管超高多重PCR扩增宏基因组的可行性

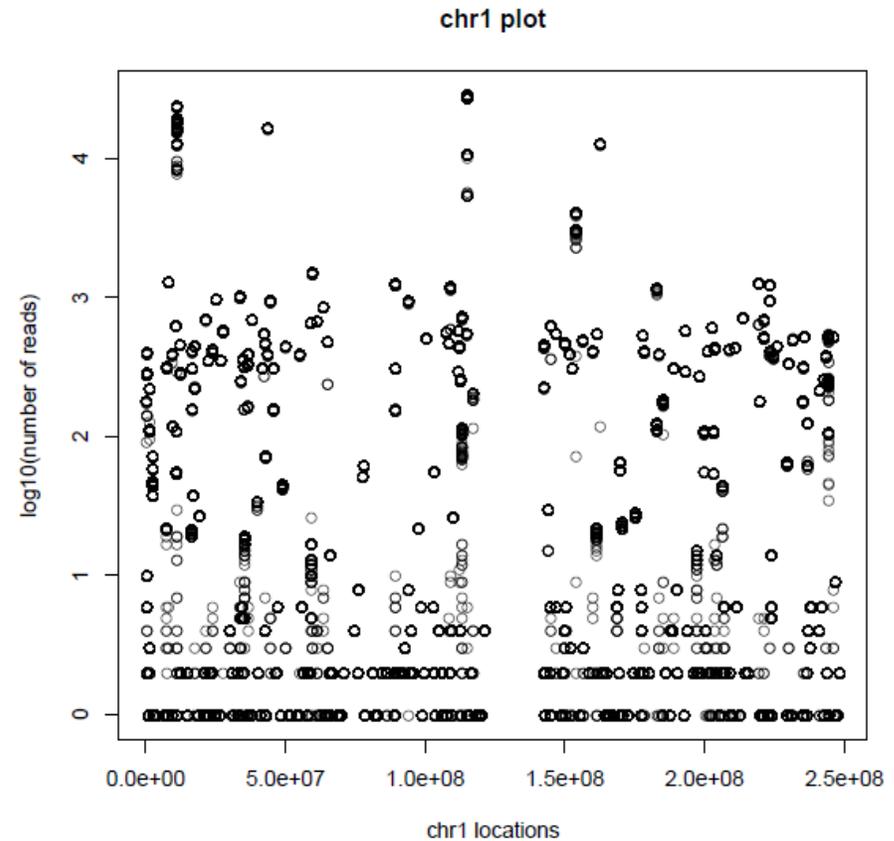


我们用带通用接头的Random Hexamer扩增人类DNA NA12878，以验证用扩增宏基因组的可行性。人类DNA被扩增成约150bp长的片段（带通用接头的长度为240-290bp），无显著的非特异性背景噪音。我们用数学模型和NGS测序优化了该方法。

NGS测序优化结果



扩增子长度分布



NGS测序的结果表明，被扩增的DNA片段呈正态分布，与数学模型相符合（左图）。而且，在极低的测序深度下（平均约3个 reads），扩增子在每条染色体上均匀分布（右图，展示一号染色体）。

实验结果持续更新中

我们将持续优化所报导的方法，并将最新版本上传至本公司网站
www.paragongenomics.com。

中国区联系人：谭行之 tanxingzhi@bcreative.cn
师慧敏 huimin.shi@outlook.com

Paragon Genomics, Inc.
1-510-363-9918
contact@paragongenomics.com
www.paragongenomics.com
3521 Investment Blvd Ste 1, Hayward, CA 94545, USA



PARAGON GENOMICS

陟岵陟屺，心脉相连