

Illumina TruPath™ Genome

更简单的全基因组测序工作流程



体验更简单的从样本到数据工作流程，
仅需10分钟手动操作时间

TAGCA
CGTAG
ATAGC

借助短读长序列及邻近数据，
结合NovaSeq™ X系列测序仪，
解析基因组中难以比对的区域



通过提升结构变异检测与超长定相
能力，揭示长距离基因组关联洞察

变革性的基因组可及性

短读长新一代测序（NGS）一直是研究人员分析大部分基因组的主要方法。然而，基因组中仍有部分区域和变异类型的表征极具挑战，例如存在高度重复序列、高序列同源性或复杂结构的区域。近期研究表明，这些区域的变异在人类遗传病中具有潜在作用^{1,2}，因此在探寻遗传病根本病因时，解析这些区域至关重要。

长读长测序等替代方法能够提升分辨率，有助于我们深入了解许多难以比对的区域和变异。但长读长测序方法的实验流程通常较为复杂，且存在公认准确性难题，易导致检测结果出现偏差。

搭载邻位映射读取技术的Illumina TruPath Genome改变了全基因组测序（WGS）的工作流程模式。这款创新检测方案仅需10分钟手动操作步骤即可捕获长DNA片段。在现有NovaSeq™ X系列测序仪上完成测序，并依托前沿的DRAGEN™技术进行二级分析。最终可获得一个准确、全面的

基因组，助力研究人员突破长读长方法本身读长的限制，发现跨更远距离的基因组规律。

借助TruPath Genome，如今我们可充分发挥短读长测序的简便性与准确性，解析基因组中难以比对的区域，提升结构变异（SV）检测能力，并生成定相序列（reads）和变异检出结果。

极简工作流程

TruPath Genome在很大程度上简化了WGS的实施流程，将手动操作时间缩短至10分钟（图1）。经过验证的因美纳转座酶化学技术，省去了传统的文库制备步骤。前沿的生物信息学算法可利用纳米孔邻近信息，结合标准双端短读长数据*，解析被数百万碱基间隔开的变异之间的关联。

* 标准双端短读长数据是指采用手动文库制备和标准边合成边测序（SBS）方法进行全基因组测序所产生的数据。



样本

用标准或高分子量（HMW）市售试剂盒提取纯化的双链DNA（dsDNA）。建议起始量为350 ng DNA，用TruPath转座酶片段化缓冲液稀释后加入文库管条。起始DNA可以从细胞系或全血中分离。不支持从福尔马林固定石蜡包埋（FFPE）和游离DNA（cfDNA）样本中分离的DNA。为确保最佳性能，建议使用能保留大片段的方法提取DNA，要求40%的片段长度 >60 kb。³

如需了解更多样本类型、不同样本质量及最优起始量的使用指南，请参考《TruPath™ Genome 在不同样本类型和样本质量中的性能表现》技术说明。

测序

将含稀释DNA的文库管条与两种TruPath Genome试剂加入NovaSeq™ X试剂卡盒，再将卡盒与NovaSeq™ X C2或C8流动槽一同装载至NovaSeq™ X系列测序仪，两种流动槽可分别一次测2个人类基因组或8个人类基因组。

流动槽表面的纳米孔中预先装载了因美纳转座酶复合物。[†] 加入流动槽的DNA由分布在多个孔中的转座酶复合物捕获，形成一种形成类似星座的纳米孔播种分布模式（图2A）。转座酶切割捕获的DNA，并在片段化加标签[‡]的过程中将片段连接到流动槽表面（图2B）。通过清洗步骤去除未结合的DNA和转座酶。

在纳米孔内部，被片段化的DNA会用于生成测序所需的簇[§]。来自同一条原始长DNA片段的DNA会在邻近的纳米孔中形成簇。测序在NovaSeq™ X系列测序仪上完成，采用成熟的Illumina XLEAP边合成边测序（XLEAP-SBS™）化学技术，运行模式为2×150 bp。

分析

测序完成后，DRAGEN™ Germline分析流程会采用一种全新算法，将测序数据与纳米孔邻近信息相结合，完成DNA与参考基因组的比对。借助邻近信息，我们可以对那些之前无法比对到参考基因组，或仅能以低置信度完成比对的read进行准确定位。通过这种方式，可基于概率将标准短读长彼此关联，

[†] 转座酶是指以二聚体形式存在的DNA转座酶复合物。

[‡] 酶法片段化加标签是指转座酶切割DNA片段并添加接头序列（标签）的过程。

[§] 簇是指在流动槽上形成的DNA扩增点，用于后续测序。

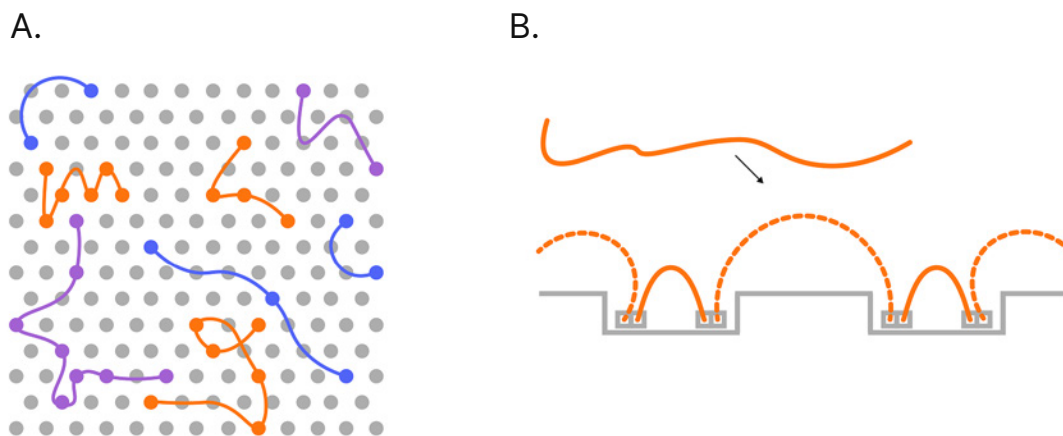


图2：邻近映射读取技术的工作原理

(A) DNA在多个相邻的纳米孔间流动，利用邻近信息生成成长距离基因组信息。(B) DNA被纳米孔（灰色方框）中的转座酶捕获，并在其中完成片段化加标签和簇生成。

从而实现各类变异的全面检测：小型变异（单核苷酸变异（SNV）、插入/缺失（indel））、大型变异（>50 bp）、已知同源区域的变异检测，同时对间隔达数百万个碱基对的小片段变异检测结果进行定相。

用高准确度的短读长解锁完整基因组

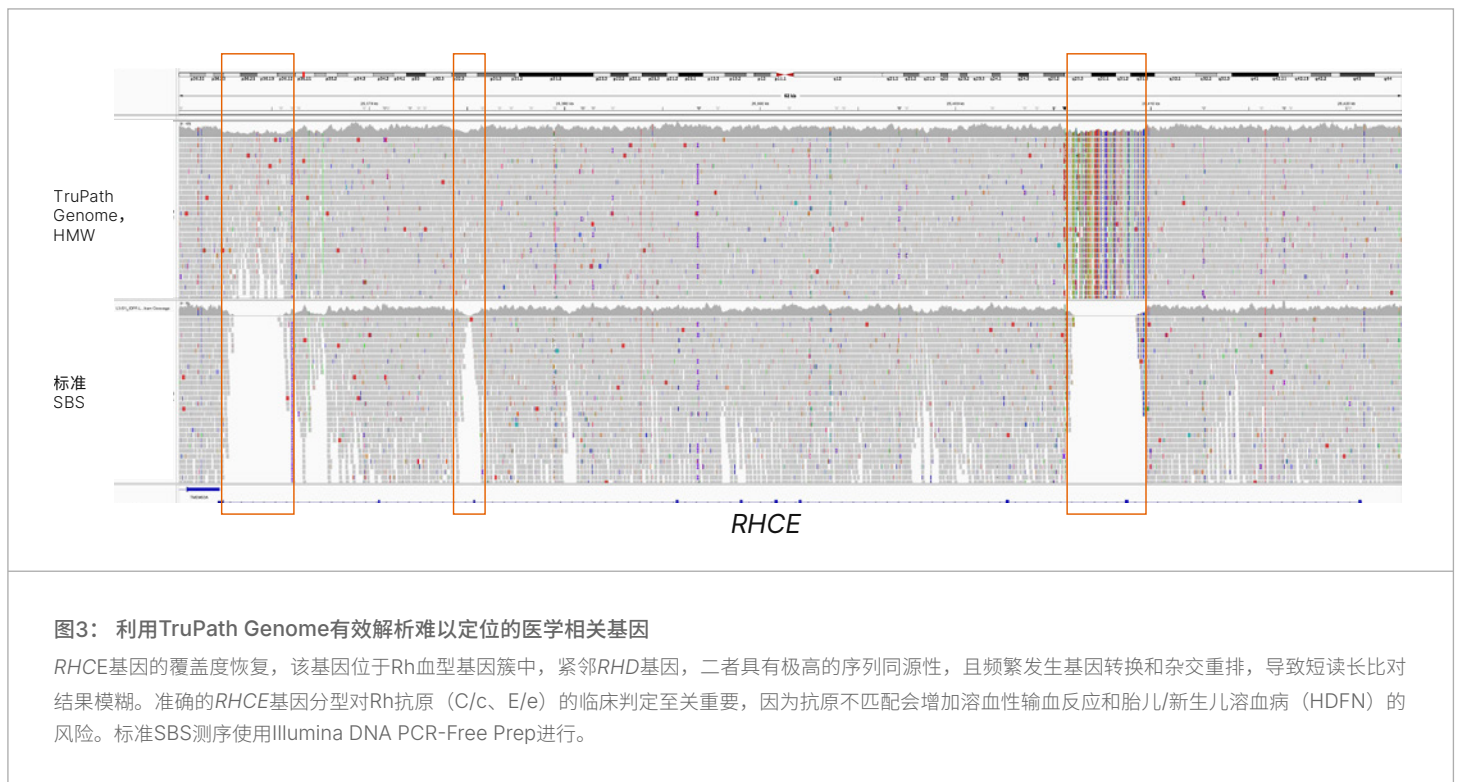
标准边合成边测序（SBS）技术已被证明对小型变异（如SNV、indel和拷贝数变异（CNV））的检测高度准确，但在特定基因组区域和变异类型（如结构变异）的检测上仍存在局限。TruPath Genome为短读长测序赋予了全新能力，可实现更全面的基因组解析。如今，短读长可用于攻克以往难以解析的基因组区域，并针对更多变异类型实现了检测准确度的提升。

提高难比对区域的覆盖度

基因组中复杂、高度多态性或重复的区域，由于可比对性差（低覆盖度或比对质量（MapQ）低），往往难以组装或与参考基因组进行比对。以往，这些“暗区”⁴的覆盖度不足，会导致MapQ评分偏低、基因组序列数据存在缺口，以及变异检测的整体分辨率受限。为填补这些缺口，TruPath Genome将利用来自相邻簇的邻近信息，为序列（reads）分配正确的基因组位置。从而实现对这些模糊比对序列（reads）的高置信度定位（图3和图4）。

更高准确度的变异检测

TruPath Genome使用HG002基因组和端粒到端粒（T2T）联盟的T2T-Q100 v1.1 v0.019基准真值集进行了性能评估，以展示其在小变异检测、结构变异检测和定相测序方面的表现。邻



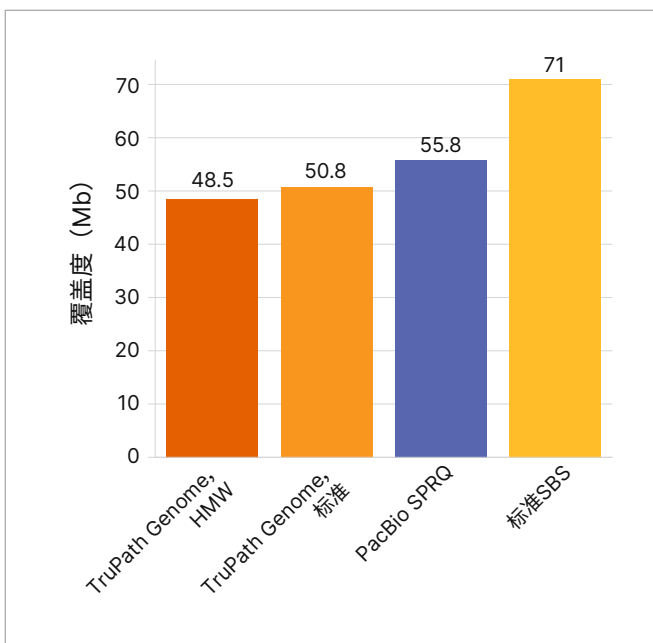


图4: TruPath Genome提升了对基因组“暗区”的覆盖度

TruPath Genome提升了对“暗区”的覆盖度。这里的“暗区”定义为Ebbert等人¹提出的“dark-by-MAPQ”区域，即90%的序列 (reads) 比对质量值 (MAPQ) 低于10的区域。TruPath Genome通过改善难比对区域的read定位，缩小了这类“暗区”在基因组中的占比。标准SBS测序使用Illumina DNA PCR-Free Prep进行。

位映射读取技术提高了比对分辨率，从而提升了小型变异检测的性能，尤其在基因组中那些历来难以比对的区域（如同源或重复区域）中表现更为明显。与标准的短读长测序和变异检测相比，TruPath Genome通过大幅减少假阴性 (FN) 和假阳性 (FP) 变异检出，展现出更高的准确性 (图5)。

与标准短读长方法相比，TruPath Genome在SV (>50 bp) 检测性能上有明显提升。SV的查全率从采用标准SBS技术时的86%提升至采用邻位映射读取技术时的94% (图6)。

超长定相

定相测序可提供单倍型信息，可用于区分来自母源和父源染色体的等位基因。¹¹ 这类信息能帮助研究人员解析复合杂合变异，这对解读常染色体隐性遗传疾病至关重要。

邻位映射读取技术特别适用于人类基因组定相。在流动槽上捕获的DNA可生成跨越数百千碱基到数兆碱基的定相区块 (图7和图8)，从而助力深入了解单倍型与复合杂合变异。

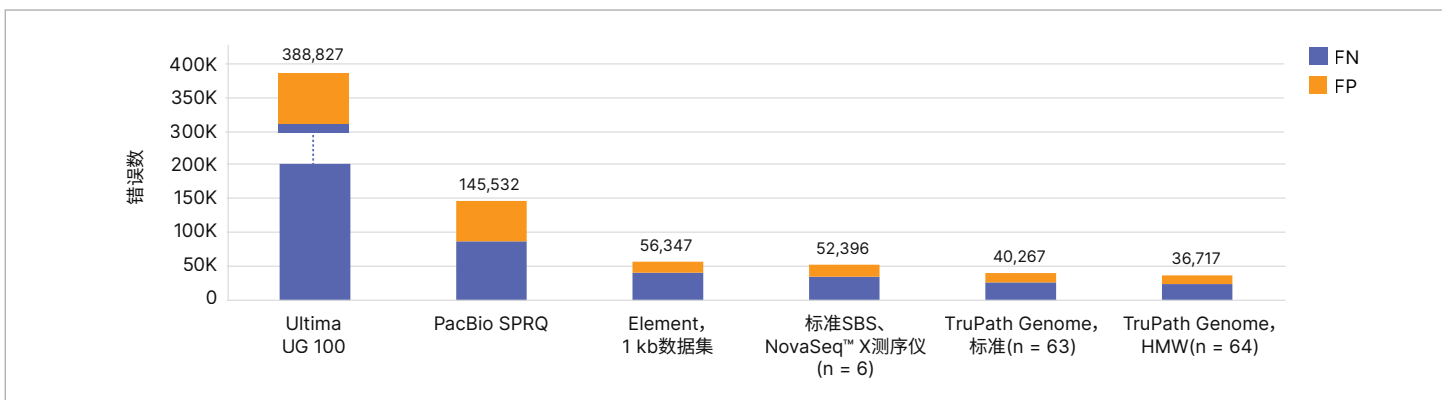


图5: 利用TruPath Genome提高小型变异检出的准确度

本研究以瓶中基因组NIST T2T-Q100 HG002 SV v1.1真值集为参照，对不同NGS测序仪及检测方法的小型变异检出性能进行了对比。Ultima UG 100、搭配SPRQ化学技术的PacBio平台、以及搭配1 kb数据集的Element AVITI相关数据均来自于已发表文献⁵⁻⁷。标准SBS数据的生成：采用Illumina DNA PCR-Free Prep构建文库，在搭载v1.4版本软件的NovaSeq™ X测序仪上结合10B流动槽完成测序，通过DRAGEN Germline v4.5.2进行数据分析（6次技术重复）。TruPath Genome标准数据生成：采用标准方法提取DNA，在搭载v1.4版本软件的NovaSeq™ X测序仪上结合C8流动槽完成测序，通过DRAGEN Germline v4.5.2进行数据分析（63次技术重复）。TruPath Genome HMW数据生成：采用HMW提取方法获得DNA，在搭载v1.4版本软件的NovaSeq™ X测序仪上结合C8流动槽完成测序，通过DRAGEN Germline v4.5.2进行数据分析（64次技术重复）。FN，假阴性变异检出；FP，假阳性变异检出。

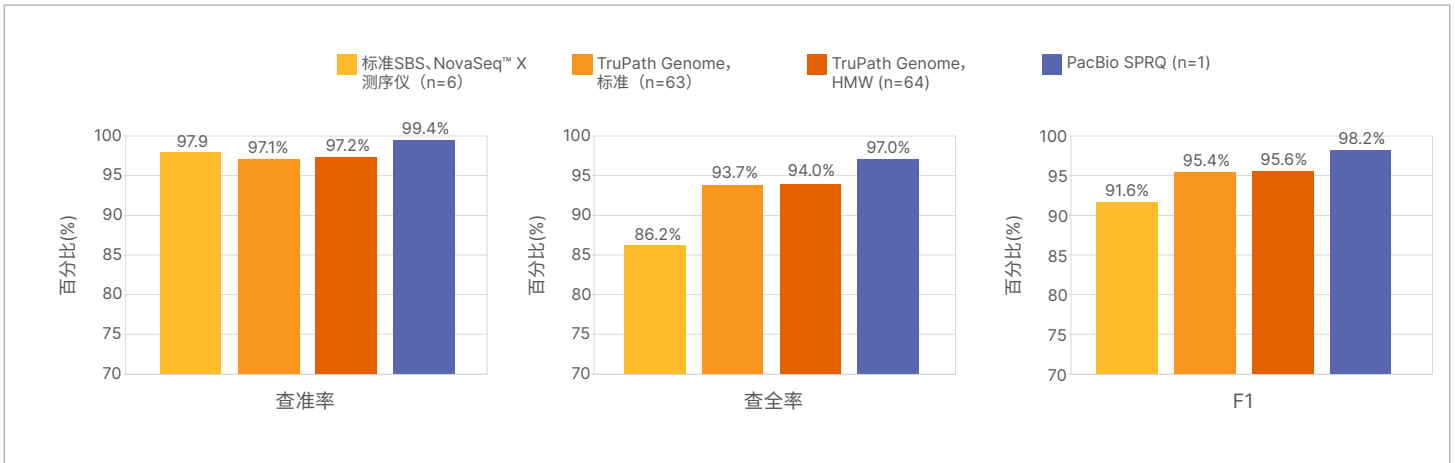


图6：利用TruPath Genome提升结构变异检出性能

TruPath Genome的结构变异 (SV) 性能，与采用DRAGEN Germline v4.5.2分析的标准SBS技术进行对比。本次分析使用了瓶中基因组NIST T2T-Q100 HG002 SV v1.1真值集及SV高置信BED文件。基准测试严格遵循瓶中基因组的结构变异基准测试指南，通过Truvari v4.2.2工具的“bench”与“refine”命令完成。标准SBS数据生成：采用Illumina DNA PCR-Free Prep构建文库，在搭载v1.4版本软件的NovaSeq™ X测序仪上结合10B流动槽完成测序（6次技术重复）。TruPath Genome标准数据生成：采用标准方法提取DNA，在搭载v1.4版本软件的NovaSeq™ X测序仪上结合C8流动槽完成测序（63次技术重复）。TruPath Genome HMW数据生成：采用HMW提取方法获得DNA，在搭载v1.4版本软件的NovaSeq™ X测序仪上结合C8流动槽完成测序（64次技术重复）。搭配SPRQ化学技术的PacBio相关数据均来自于已发表文献。⁶ 查准率用于衡量真阳性结果的占比。查全率为真阳性结果占样本中变异总数的百分比。F1评分（%）是根据真阳性和真阴性结果占总结果的比例来计算的。数值越高，表明基于参考数据的检测准确性越优。

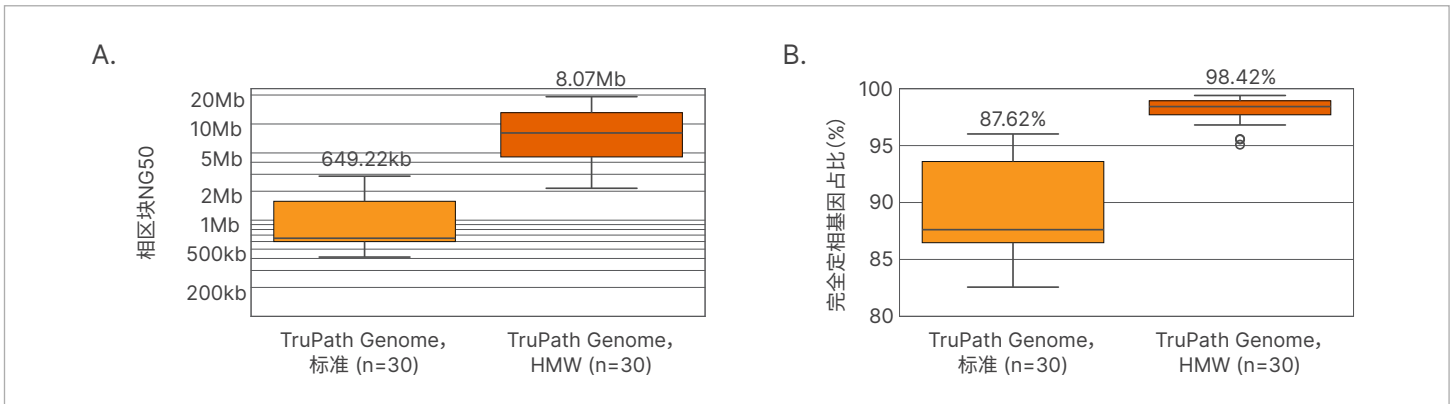


图7：利用TruPath Genome对30组配对的HMW与标准提取细胞系进行超长定相

采用DRAGEN Germline v4.5.2中的DRAGEN定相算法，在1~22号染色体上测定相区块长度。TruPath Genome数据来自30个不同的Coriell细胞系样本，每个样本均同时采用高分子量（HMW）和标准两种方法提取DNA。相区块NG50指当目标区域（如基因组）的50%完成定相时，对应的相区块长度。若某项技术无法对50%的目标区域完成定相，则其NG50为0碱基对。完全定相基因占比指在指定基因列表（Gencode v44 genes.gtf）中，完全包含在单个相区块内的基因区域所占的百分比。

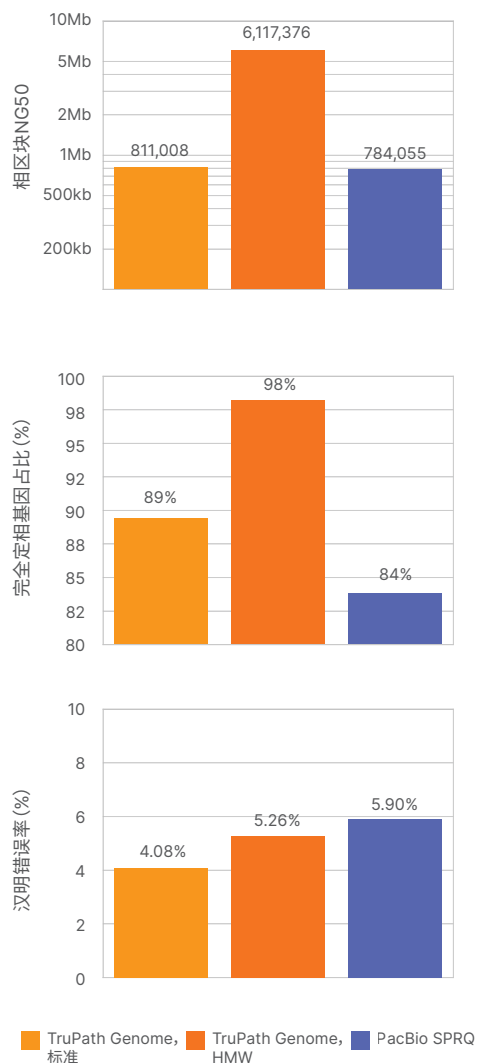


图8：TruPath Genome的定相测序性能（基于HG002真值资料库验证）

采用DRAGEN Germline v4.5.2中的DRAGEN定相算法，在1~22号染色体上测定相区块长度。HiFi数据（PacBio SPRQ）的定相VCF文件来源于：https://downloads.pacbcloud.com/public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2/。TruPath Genome数据来自HG002细胞系样本的测序重复实验，该实验同时采用了HMW提取和标准提取两种样本处理方式。相区块NG50指当目标区域（如基因组）的50%完成定相时，对应的相区块长度。完全定相基因占比指在指定基因列表（Gencode v44 genes.gtf）中，完全包含在单个相区块内的基因区域所占的百分比。定相汉明错误率基于T2T Q100真值集⁹进行验证，该数据集下载自：https://ftp-trace.ncbi.nlm.nih.gov/ReferenceSamples/giab/data/AshkenazimTrio/analysis/NIST_HG002_DraftBenchmark_defrabbV0.020-20250117/GRCh38_HG2-T2TQ100-V1.1_smvar.vcf.gz。

采用标准DNA提取方法时，TruPath Genome可对约91%的基因实现完全定相，采用高分子量（HMW）DNA提取方法时，这一比例可提升至约97%。此外，无论采用标准还是HMW DNA提取方法，TruPath Genome均可对约98%⁹的杂合子SNV进行定相。

对于HMW DNA，结合DRAGEN定相分析的TruPath Genome基因组数据，将汉明错误率**降至约4%，明显优于HapCut2等公开工具（约6%）。¹⁰

高同源性区域的变异检测

由于序列同源性高，使用标准短读长在旁系同源区域进行准确的变异检测十分困难，这会导致read比对模糊和变异检测错误。解决方案通常依赖成本高昂的反复检测，如多重连接依赖性探针扩增技术（MLPA）或长片段聚合酶链式反应（LR-PCR）。TruPath Genome结合了流动槽邻近信息与全新的多区域联合检测（MRJD）算法，可在旁系同源基因中实现可靠的从头单倍型解析变异检出（图9和图10）。

TruPath Genome靶向基因组中已知存在串联或非串联拷贝分布的11个基因（表1），采用MRJD技术实现高精度序列比对，同时完成单倍型解析重建与小片段变异检出（图10）。

表1：用于MRJD分析的TruPath Genome靶向基因

<i>CFHR</i>	<i>NCF1</i>	<i>SMN1/2</i>
<i>CYP2D6</i>	<i>PMS2</i>	<i>STRC</i>
<i>CYP11B1</i>	<i>RCCX (CYP21A2 和TNXB)</i>	<i>USP8</i>

⁹ 杂合变异定相率指已完成定相的杂合子SNV占所有杂合SNV的比例，计算公式为：定相SNV数量/杂合子SNV总数。

**汉明错误率指变异被错误分配到母源或父源染色体链上的情况。

提升短串联重复序列（STR）的解析能力

STR是基因组中的重复区域，其长度可能扩增至正常范围之外，引发与多种遗传病相关的突变，包括脆性X综合症、肌萎缩侧索硬化症和亨廷顿病。¹¹ 准确估算这类区域的长度并恢复其read一直颇具挑战，因为最具临床意义的扩增等位基因长度往往超出了标准短读长测序数据的读长范围。通过结合邻近信息与高质量短读长数据，TruPath Genome可明显提升这类区域的read恢复率，为准确的STR长度测定提供支持，并借助定相信息对两个拷贝进行基因分型，从而提升解析能力（图11）。

复杂SV的超长距离基因组信息

TruPath Genome能够解析大规模染色体重排，其全新能力超越了传统变异检出的性能基准。TruPath Genome可捕获基因组任意两个区域间之间邻近簇^{††}的read相关信息，从而构建高分辨率的基因组结构可视化图谱——即“协同定位图”（图12）。

†† 邻近簇指在流动槽上物理位置彼此邻近的簇。

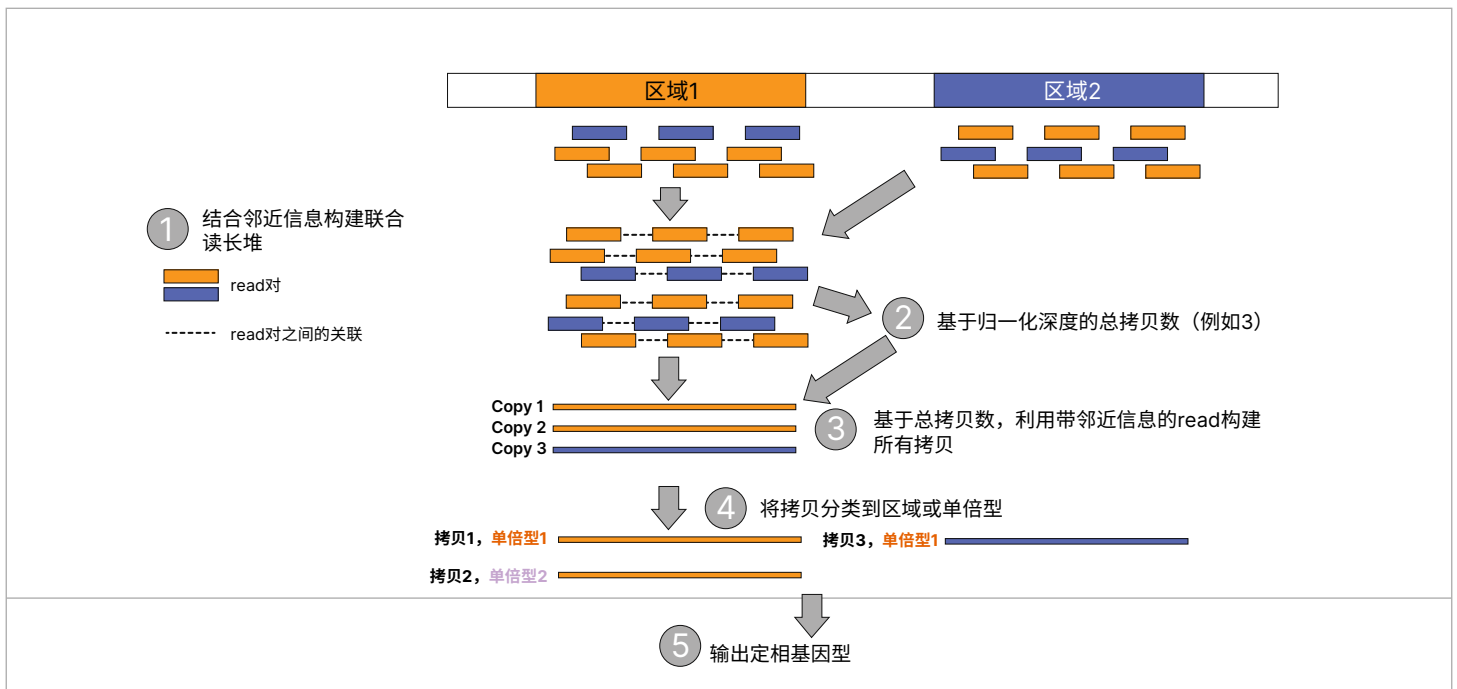
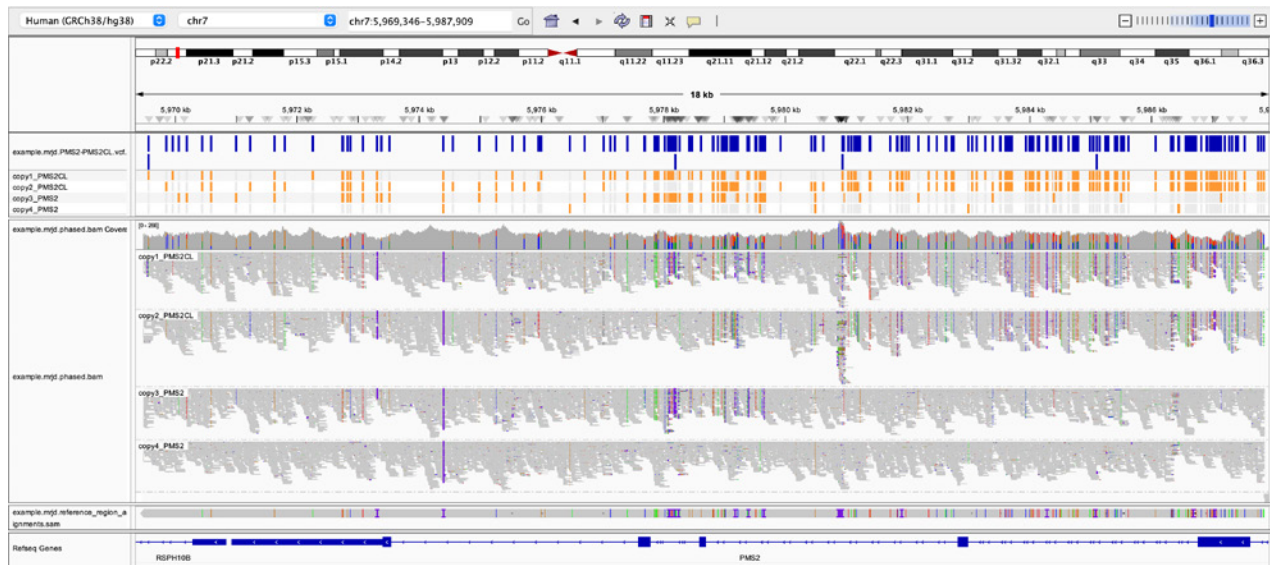


图9：用于检测片段重复区域变异的MRJD方法示意图

1. 针对一组片段重复区域，MRJD会提取所有比对到该区域内任意拷贝的read，进行联合分析。利用邻近信息建立这些read间的关联，识别出源自同一原始分子的read集合；2. 然后对步骤1中提取的所有read的计数进行归一化，经GC校正后，用于估计起始样本中该片段重复序列的总拷贝数；3. 基于步骤2得到的拷贝数，采用联合基因分型方法，结合所有带邻近信息的read，识别并定相起始样本中该片段重复序列每个拷贝的所有变异。4. 当重复区域在参考基因组中位置较远时，将每个完全定相的拷贝分配到特定的参考基因组位置；（仅）当重复区域在参考基因组中呈串联或邻近分布时，则将其定相到其他拷贝并构建为单倍型。5. 将所有拷贝的完全定相变异、每个拷贝的单倍型分配结果、拷贝到参考基因组位置的对应关系，以及分配到每个拷贝的序列（reads），输出至MRJD专用的VCF、JSON和BAM文件。

A. PMS2基因



B. SMN1基因

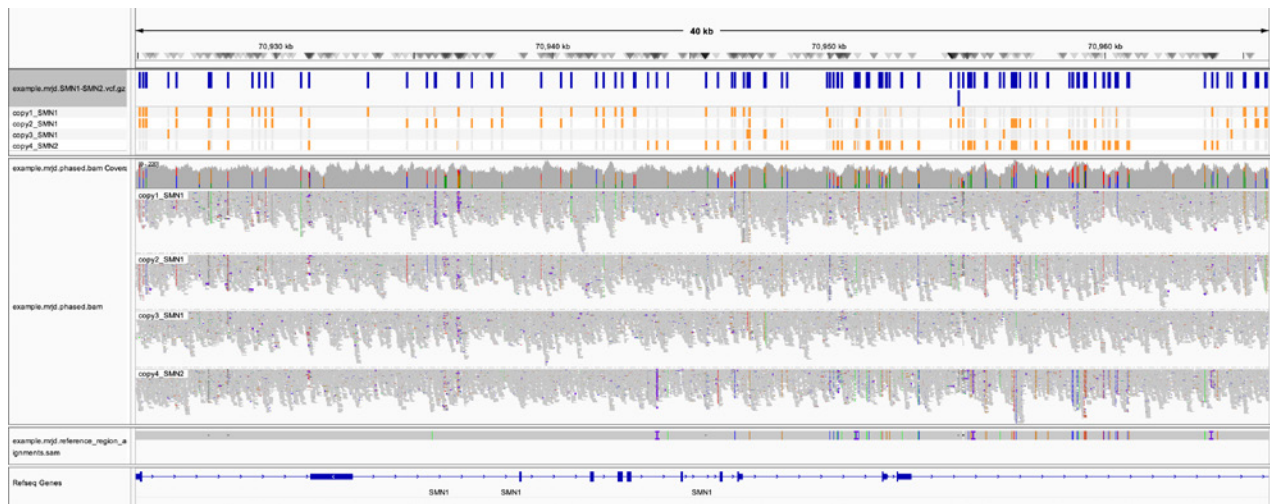


图10：利用TruPath Genome与MRJD在旁系同源基因中得到的变异检出与read分配示例

(A) 在PMS2参考基因组位置上，展示了PMS2/PMS2CL基因-假基因对所有拷贝的定相变异检出及对应的支持read。根据与PMS2/PMS2CL侧翼序列的邻近关联，四个拷贝被分为两类：两个为PMS2CL（假基因）拷贝，两个为PMS2（活性基因）拷贝。拷贝3（PMS2）在第13和14外显子区域，存在类似基因转换事件的变异，其序列与PMS2CL参考序列匹配。从顶部到底部的第三个面板中，展示了PMS2CL参考序列与PMS2参考序列的比对结果。(B) 在SMN1参考基因组位置上，展示了SMN1/SMN2基因-假基因对所有拷贝的定相变异检出及对应的支持read。根据每个定相拷贝中SMN1 c.840位点的碱基，四个拷贝被明确分为两类：三个为SMN1（活性基因）拷贝，两个为SMN2（失活基因）拷贝（SMN1拷贝该位点为C碱基，SMN2拷贝该位点为T碱基）。

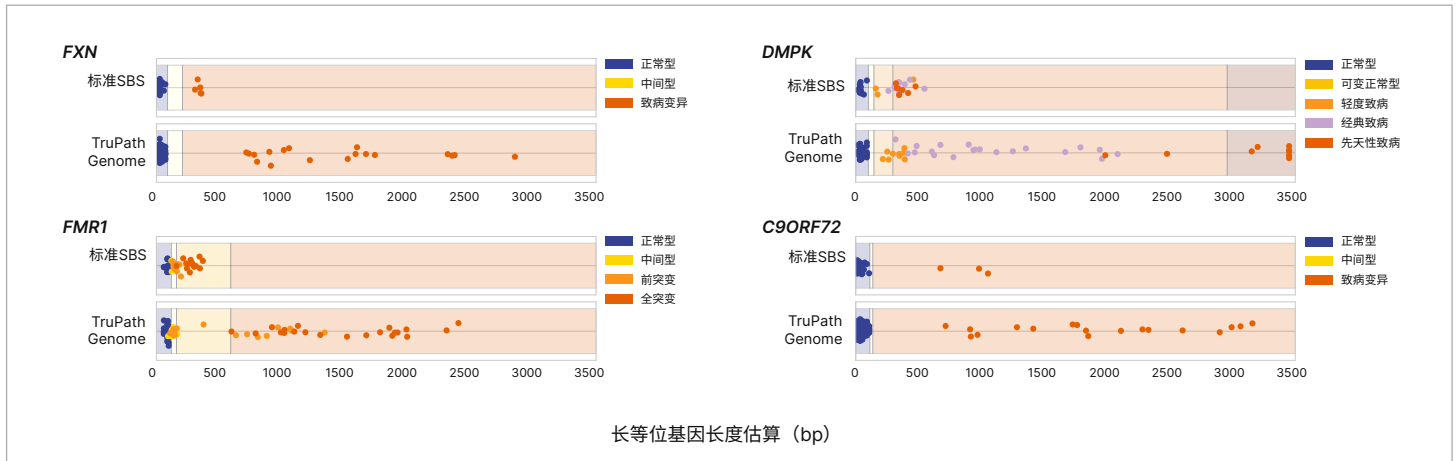


图11: 利用TruPath Genome实现更准确的STR扩增检测

每个点的颜色代表STR的真实分类。彩色条带定义了每个位点中不同扩增类型的不同分类区间。X轴表示使用TruPath Genome或标准SBS (Illumina DNA PCR-Free Prep) 作为输入时, 由DRAGEN STR分析得到的STR长度估算值。TruPath Genome的分类结果与真实分类的一致性明显更高, 且覆盖的STR长度范围更广。标准SBS的总重复数: 78份生物样本 (78个独特个体)。TruPath Genome的总重复数: 135份生物样本 (42个独特个体)。

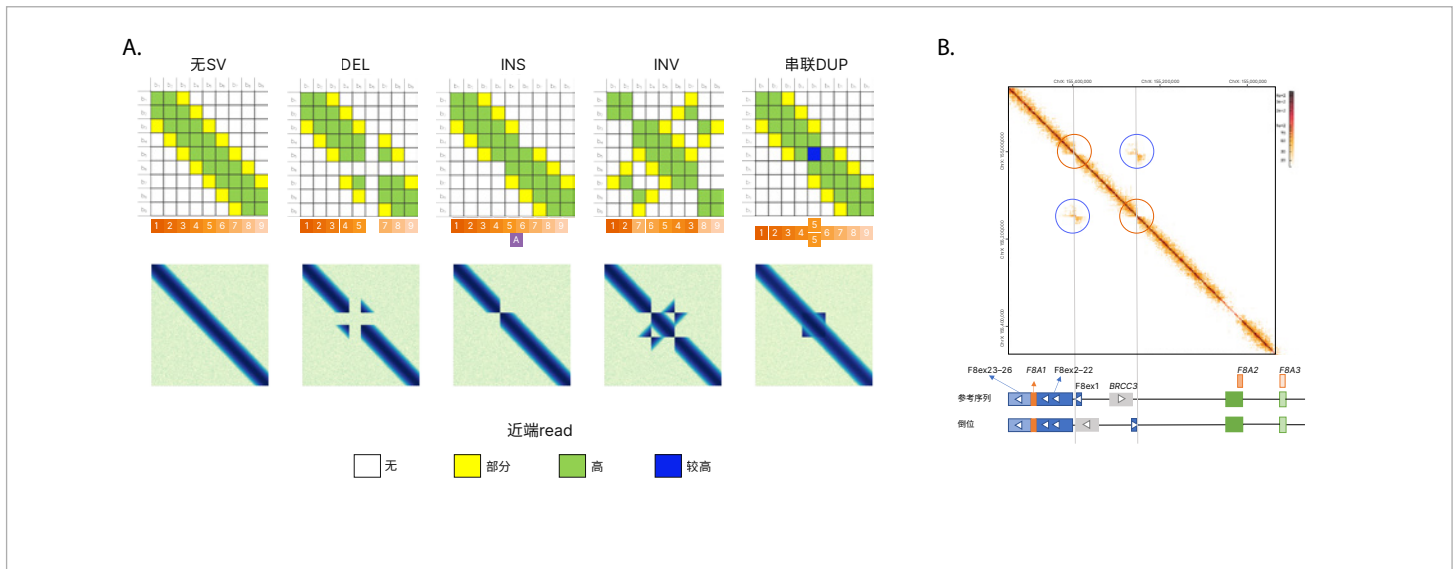


图12: 使用TruPath Genome协同定位图可视化结构变异事件

(A) 通过统计比对到所有基因组区间对的邻近read (流动槽空间) 数量, 生成二维协同定位图。X轴和Y轴均代表参考基因组区间。两个区间之间的大量连接, 表示这些区间在起始样本的基因组空间中彼此邻近。在没有结构变异的区域中, 参考基因组中邻近的区间在起始样本中也邻近, 因此在协同定位图的对角线附近会出现大量连接。当存在结构变异时, 参考基因组中邻近的基因组区间在起始样本中可能不再邻近, 而参考基因组中距离较远的基因组区间在起始样本中可能变得邻近。这种连接性的变化会在协同定位图上体现为不同形状的非对角线信号, 以及对角线连接性的缺失。(B) F8基因位点内含子1基因倒位的示例, 该变异与重型血友病A高度相关。请注意事件边界处预期呈现的沙漏状信号 (位于非对角区域, 蓝色圆圈所示), 以及对角区域信号缺失 (橙色圆圈所示)。图谱下方为该位点在参考基因组中的示意图, 以及基因倒位区域的示意图。

全基因组测序的极简工作流程

Illumina TruPath Genome凭借核心的邻位映射读取技术，彻底革新了人类基因组测序。这种创新的工作流程摒弃了传统的文库制备步骤，实现前所未见的操作简便性，同时利用邻近序列信息，在高准确度的标准短读长测序数据基础上，完整保留DNA链的长距离基因组信息。最终打造出一款兼具短读长测序数据质量优势、小片段变异高准确度检测能力的检测技术，不仅提升了人类基因组的比对覆盖度，更为遗传性疾病的研究提供了强大工具。

了解更多 →

[Illumina TruPath Genome](#)

[邻位映射读取技术](#)

[NovaSeq™ X系列测序仪](#)

订购信息	
产品	货号
Illumina TruPath Genome (NovaSeq™ X C8流动槽)	20157405
Illumina TruPath Genome (NovaSeq™ X C2流动槽)	20157406

参考文献

1. Ebbert MTW, Jensen TD, Jansen-West K, et al. **Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight.** *Genome Biol.* 2019;20(1):97. Published 2019 May 20. doi:10.1186/s13059-019-1707-2
2. Ryan NM, Corvin A. **Investigating the dark-side of the genome: a barrier to human disease variant discovery?.** *Biol Res.* 2023;56(1):42. Published 2023 Jul 20. doi:10.1186/s40659-023-00455-0
3. Illumina. TruPath Genome performance with samples of varying type and quality. [illumina.com/content/dam/illumina-marketing/documents/products/technotes/trupath-genome-sample-info-tech-note-m-gl-03932](https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/trupath-genome-sample-info-tech-note-m-gl-03932). Published February 2026. Accessed February XX, 2026.
4. Wadsworth ME, Page ML, Aguzzoli Heberle B, Miller JB, Steely C, Ebbert MTW. **Sequencing the gaps: dark genomic regions persist in CHM13 despite long-read advances.** Preprint. *bioRxiv.* 2025;2025.05.23.655776. Published 2025 May 28. doi:10.1101/2025.05.23.655776
5. Ultima Genomics. Genome-in-a-Bottle (GIAB) samples HG001-HG005. <https://cdn.sanity.io/files/l7780ks7/production-2024/0a1b6a62a6da3e3fcafb81cad4c8ff2ffe85dd41.pdf>. Published February 2025. Accessed January 29, 2026.
6. PacBio. Index of /public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2. https://downloads.pacbcloud.com/public/revio/2024Q4/WGS/GIAB_trio/HG002_rep1/analysis/v3.0.2/. Accessed January 29, 2026.
7. Asri M, Chang PC, Mier JC, et al. **Pangenome-aware DeepVariant.** Preprint. *bioRxiv.* 2025;2025.06.05.657102. Published 2025 Jun 6. doi:10.1101/2025.06.05.657102
8. Browning SR, Browning BL. **Haplotype phasing: existing methods and new developments.** *Nat Rev Genet.* 2011; 12(10):703-714. doi: 10.1038/nrg3054
9. Hansen NF, Dwarshuis N, Ji HJ, et al. **A complete diploid human genome benchmark for personalized genomics.** Preprint. *bioRxiv.* 2025;2025.09.21.677443. Published 2025 Sep 21. doi:10.1101/2025.09.21.677443
10. Data on file. Illumina, Inc., 2026.
11. Steely CJ, Watkins WS, Baird L, Jorde LB. **The mutational dynamics of short tandem repeats in large, multigenerational families.** *Genome Biol.* 2022;23(1):253. Published 2022 Dec 12. doi:10.1186/s13059-022-02818-4

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) 8441-6900 • 传真 (010) 8455-4855
 技术支持热线 400-066-5835 • chinasupport@illumina.com
 市场销售热线 400-066-5875 • china_info@illumina.com • www.illumina.com.cn

© 2026 Illumina, Inc. 保留所有权利。所有商标均为因美纳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.cn/company/legal.html。
 M-GL-03931 v1.0



因美纳企业动态



因美纳讲堂



因美纳中国官方商城