Technical note

biomodal

Using UMIs with duet multiomics solution evoC reduces sequencing requirements and increases coverage

Introduction

DNA contains molecular information encoded in both genetic and epigenetic bases, essential for understanding biology. However, interrogating both of these informational layers simultaneously has traditionally been challenging. With the recent introduction of duet multiomics solution evoC, which uniquely reads genetic and epigenetic cytosine modifications, this challenge has been overcome. duet evoC enables the simultaneous identification of A, C, G, T, 5mC, and 5hmC within a single DNA molecule with high precision. However, obtaining this comprehensive information across the entire genome requires extensive sequencing, especially for in-depth analysis with unique dual indexing.

Using Unique Molecular Identifiers (UMIs) for sequencing enhances the accuracy of genome interrogation, whether examining the whole genome or specific regions with a target enrichment panel. This method increases the number of unique reads, reducing both sequencing requirements and costs.

Here, we employ UMIs with duet evoC to deliver a full 6-base genome, integrating genetic and epigenetic information in a single read while reducing sequencing needs. We demonstrate that using duet evoC with UMIs significantly reduces the read duplication rate, thereby increasing the number of genome equivalents recovered per experiment.

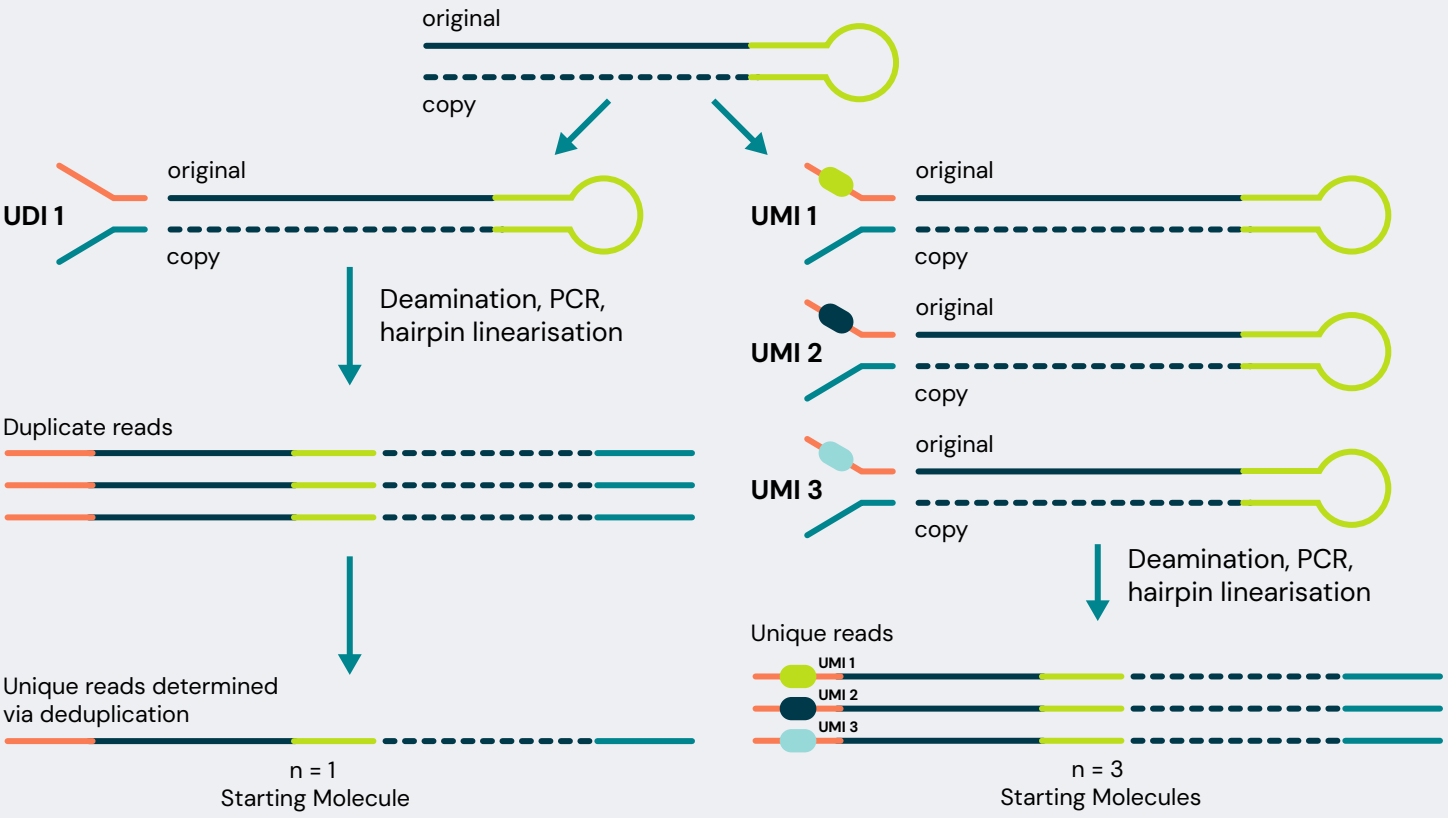

Figure 1: Workflow schematics for duet evoC with UDIs (left) and duet evoC with UDI-UMIs (right).

Whole genome sequencing with UMI Adapters

Whole genome sequencing (WGS) investigates the entirety of a sample genome, but the large amount of data required to sequence an entire genome can restrict the achievable depth of coverage, making it difficult to detect low-frequency genetic variants and abnormalities. Reducing sequencing read requirements can enable deeper analysis of genomic loci. To address this, we employed UMI adapters in the duet evoC workflow to prepare WGS libraries. This unique index-corrected WGS method reduced duplicate reads, thereby increasing overall genomic coverage. Our results demonstrate that using UMI libraries in WGS reduces the sequencing requirements per sample. At a target coverage of 30x, the inclusion of UMIs enabled an approximately 10% reduction in the required number of sequencing read pairs compared to an approach without UMIs.

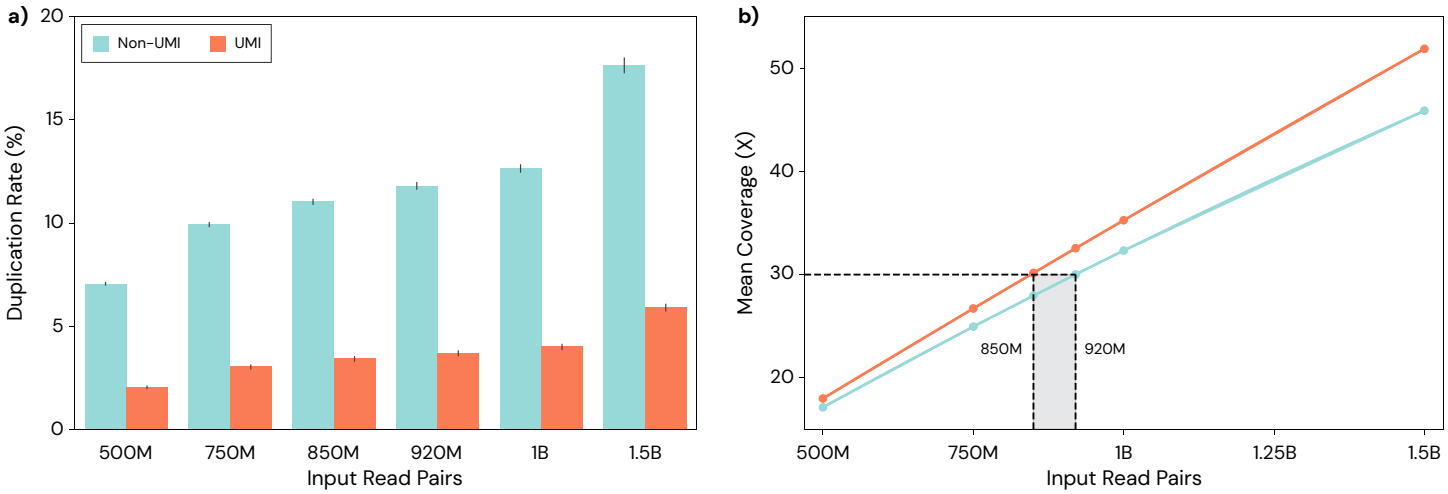

Figure 2: duet evoC UMI Libraries Reduce Sequencing Requirements. (a) Duplication Rate (%) or **(b)** Mean coverage (x), achieved with or without UMIs for a range of input sequencing read pairs.

Two samples of NA12878 and two samples of cerebellum genomes were used for library preparation. Libraries were created with IDT UDI-UMI adapters using duet evoC library preparation reagents from biomodal. All four samples were sequenced on an Illumina NovaSeq 6000 S4 flow cell. These libraries were analysed using Picard MarkDuplicates in UMI-aware and UMI-unaware modes. Figure 2a illustrates the impact on duplication rate with various input read amounts, and figure 2b shows the effect on mean coverage at various input reads when UMIs are used.

	Read pairs required for 30X coverage (millions)
With UMI	850
Without UMI	920

Table 1: Number of sequencing reads required to achieve 30X using duet evoC UMI libraries with UMI aware vs. UMI unaware.

Use of UMI adapters helps reduce the amount of sequencing needed to maintain high coverage across the whole genome. Table 1 outlines the sequencing input necessary to achieve 30X coverage using the duet multiomics solution evoC, either with or without UMIs. Utilizing the duet multiomics solution evoC UMI libraries significantly decreases the total number of reads required to reach equivalent coverage.

Target enrichment with duet evoC UMI libraries

Enrichment using target panels aims to screen regions of interest within the sample genome, reducing sequencing requirements by focusing only on areas of importance, and enhancing both accuracy and cost-effectiveness. We demonstrate the performance of duet evoC UMI libraries with the Twist Alliance Pan-cancer Methylation Panel. duet evoC ensures uniformity across various input types and input amounts, while reducing sequencing requirements to achieve high coverage of target regions. Furthermore, using UMIs instead of UDIs, represented by UMI aware and unaware settings respectively, emphasizes the coverage benefits of detecting a higher number of unique molecules at equivalent depth.

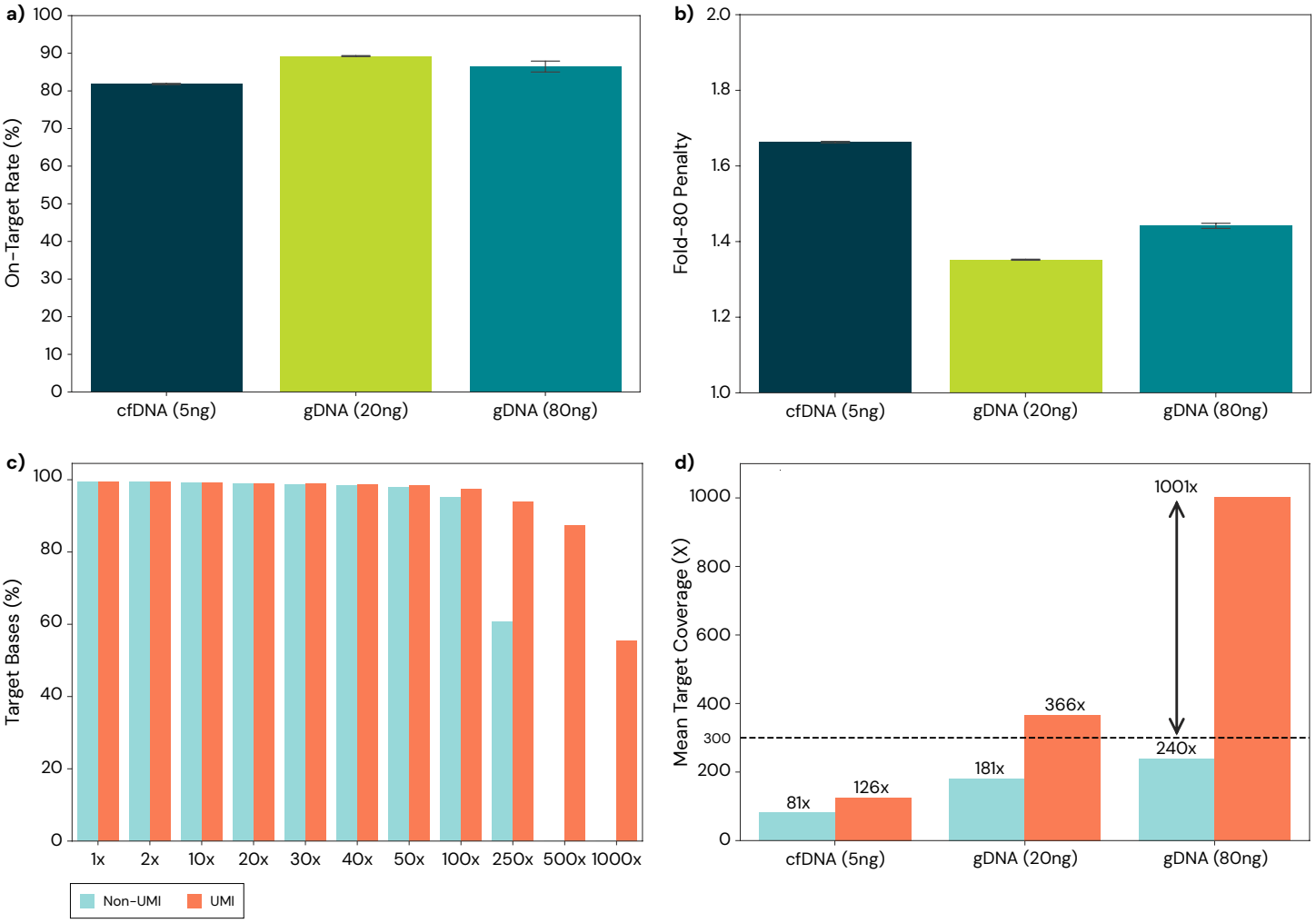

Figure 3: duet evoC using UMIs with hybrid capture method. Hybridization capture was performed on 8-plex pool of WGS UMI libraries constructed using duet evoC reagents. Targeting was performed using the Twist Alliance Pan-cancer Methylation Panel from varying DNA input types. **(a)** On-target rate (%), **(b)** Fold-80 base penalty, **(c)** Target bases (%) with or without UMIs, **(d)** Mean target coverage with or without UMIs. Data was obtained using cell free DNA (cfDNA) isolated from patients with colorectal cancer at 5ng input, and genomic DNA (gDNA) from NA12878 & Cerebellum samples at 20ng and 80ng input respectively.

Conclusion

Here we demonstrate the compatibility of duet evoC sequencing libraries with UMI adapters. Using UMIs to construct whole genome libraries effectively reduced duplication rates while increasing coverage and lowering sequencing requirements. The superior performance of duet evoC UMI library construction was also evident in target enrichment approaches, where the enhanced depth from increased unique reads allowed for a deeper understanding of target areas in both genomic and cell-free DNA. Refining genetic and epigenetic investigations to narrower regions of interest improves the detection of low-prevalence disease-associated biomarkers at significant sequencing depths, and utilising UMIs can enable even deeper interrogation of key areas of interest, all while maintaining the high accuracy and minimal sample input benefits provided by biomodal's duet technology.

Methods

The compatibility of duet evoC 6-base libraries with UMI adapters demonstrated through WGS was performed with UMI adapter libraries constructed using the duet evoC reagents (with protocol adjustments for UMI forkheads). The input consisted of 80 ng of NA12878 genome and cerebellum genome (two samples each, sonicated at 250 bp). All four samples were sequenced on an Illumina NovaSeq 6000 S4 flow cell.

Target enrichment of UMI libraries generated from the duet evoC 6-base workflow was performed using the duet evoC kit with an adjusted target enrichment protocol (biomodal). Library preparation was performed using 5ng of cfDNA isolated from patients with colorectal cancer (two samples), 20ng & 80ng gDNA of Cerebellum (two samples for each input amount) and 80ng of NA12878 (two samples). Genomic DNA was sonicated at 250bp. duet evoC UMI-constructed libraries were then enriched using the Twist Alliance Pan-cancer Methylation Panel (1.5Mb panel size), following a modified Twist methylation sequencing workflow. The target-enriched libraries were sequenced on an Illumina NovaSeq 6000 SP flow cell. These datasets were analysed with Picard MarkDuplicates in UMI-aware and UMI-unaware modes.

biomodal
Chesterford Research Park
Cambridge, UK
CB10 1XL

+44 (0) 1223 800 700
www.biomodal.com
info@biomodal.com