5-base genome: Simultaneous detection of genomic and methylation signatures using a concise, integrated multiomic library prep

Suzanne Rohrback[1], Huy Lam[1], Danielle Goldberg[1], James Baye[1], Joshua Beytebiere[1], Eric Brustad[1], Kayla Busby[1], Ev Chekalin[1], Bridget Dorsey[1], Sophie Fertel[1], Stephen Gross[1], William Jordan[1], Valerie Montel[1], Oscar Murillo[1], June Pais[1], Angelica Schalembier[1], Konrad Scheffler[1], David Tse[1], Kat Tzeng[1], Ali Crawford[1], Lisa Watson[1], Allison Yunghans[1], Fiona Kaper[1]

[1]Illumina, Inc., San Diego, CA 92122, USA

BCS02 Convergent Science, poster 3159

Introduction

- DNA molecules contain multiomic information, with nucleotide sequence informing genomic variation and modifying marks informing epigenome status
- Capturing both signals from the same sample would simultaneously inform predisposition and response, accelerating discovery and informing interpretations
- We have engineered a bespoke enzyme which selectively converts 5-methyl cytosine to thymidine, retaining high genome complexity
- Streamlined sample preparation protocol enables single shift library preparation for WGS sequencing and is compatible with low input samples
- DRAGEN algorithm innovations maximize alignment rates and provide small variant calling which approaches the accuracy of whole genome sequencing
- Illumina's 5-base technology introduces a streamlined, end to end workflow that provides high quality genome and methylome annotations from a single sample

Introducing the Illumina 5-base solution: genome and methylome readout from a single library preparation

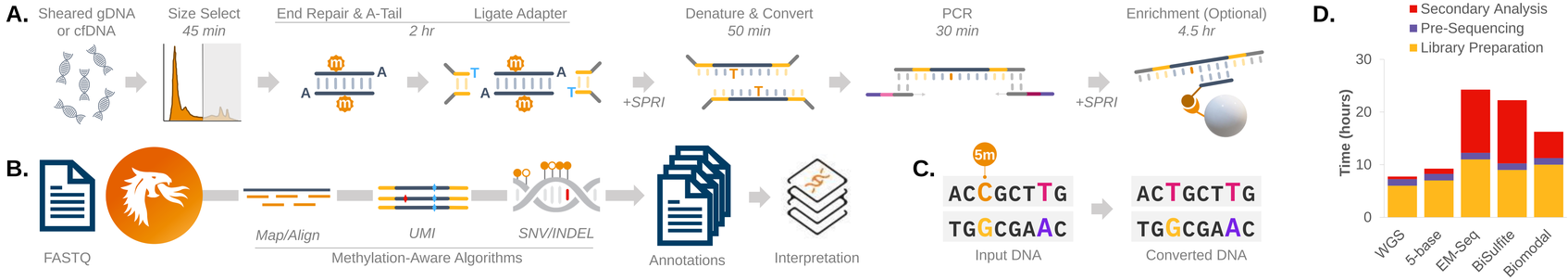

Figure 1: 5-base technology workflow overview. (A) Methyl-conversion is a modular addition to library preparation. After adapter ligation, DNA is denatured (20 min) and incubated with the 5-methyl cytosine conversion enzyme (30 min). (B) DRAGEN analysis pipelines are methyl-conversion aware providing optimal analysis efficiency and accuracy. Output is compatible with diverse interpretation software, such as Illumina Connected Multiomics. (C) 5mC conversion retains 4-base complexity and generates distinct signal from DNA variants. (D) Time required for key workflow stages, compared to orthogonal methodologies.

Concordant with orthogonal technologies

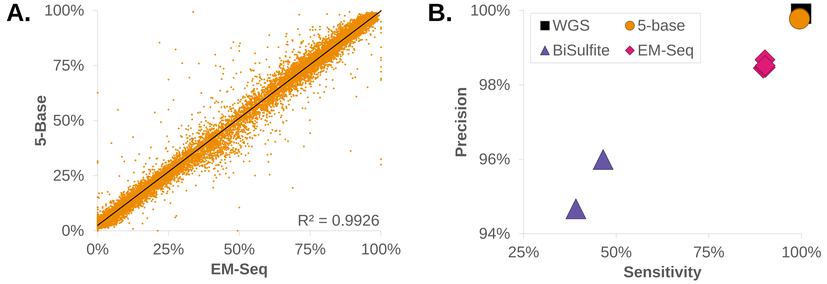

Figure 2: Accurate methylome and genome annotations. (A) Average methylation in CpG islands of NA12878 using EM-Seq or 5-base sample preparations. (B) Germline SNV calling accuracy for NA12878 (EM-Seq and BiSulfite processed by BiSNP, 5-base and WGS by DRAGEN).

Assay versatility enables diverse cancer research needs

High complexity across DNA inputs

Capture somatic DNA variants

Retain fragmentome profiles

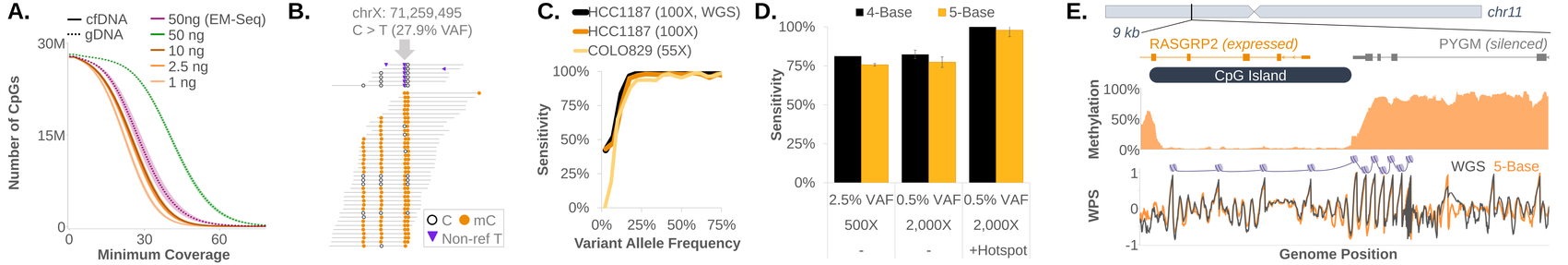

Figure 3: Data quality supports sensitive discovery in cancer applications. (A) Genomic CpG coverage for different DNA types and inputs. (B) Simultaneous reporting of genome and methylome captures somatic allele-specific methylation. (C) Somatic SNV sensitivity across allele frequencies and genomic coverage with whole genome sequencing data. (D) Sensitivity to low variant allele frequency (VAF) SNVs compared to DNA-only (4-base) sequencing with enrichment and UMI collapsing. (E) Overlaid fragmentome and methylome patterning from a healthy cfDNA donor identifies active and silenced genes.

Dual-omic characterization of AML molecular subtypes

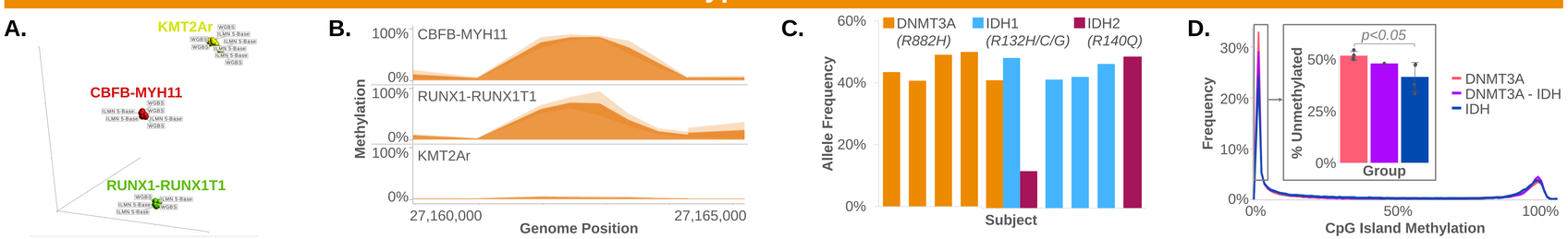

Figure 4: 5-base assay captures DNA and methylation abnormalities in AML subject samples. (A) uMAP clustering of differentially methylated regions distinguish driver mutations. (B) HOXA9 gene shows the expected demethylation in the KMT2Ar subtype. (C) Detection of missense variants in methylation regulatory enzymes. (D) Global CpG island methylation corresponds to DNMT3A and IDH mutation state. Inset shows the percent of islands with <5% methylation (Kruskal-Wallis p<0.01, Wilcoxan test for pairwise comparison). *Data generated in collaboration with David Spencer lab, WashU.*

Conclusion

- Methylome and genome annotation accuracy comparable to single-ome preps
- High complexity data facilitates sensitive somatic investigations
- Retain fragment information from low input cfDNA samples
- Multiomic data stratifies samples with unique biological changes

ONE ASSAY. DUAL INSIGHTS. UNLOCKING DISCOVERY