High sensitivity detection of specific ultra low-frequency somatic mutations for minimal residual disease (MRD) monitoring

Tina Han[1], Tong Liu[1], Michael Bocek[1], Patrick Cherry[1], Shawn Gorda[1], Nairi Pezeshkian[2], Dan Nasko[2], Po-Yuan Tung[1], Derek Murphy[1] and Esteban Toro[1]
[1]Twist Bioscience, South San Francisco, California, USA [2]Pacific Biosciences, Menlo Park, California, USA

Abstract

Minimal/Molecular residual disease (MRD) refers to the small number of tumor cells which may remain within a patient after therapeutic intervention. The detection of these remnants and monitoring of their abundance is a promising prognostic marker to identify individuals at risk of recurrence or in need of adjuvant therapy. Due to the low abundance of ctDNA present in samples obtained during remission, MRD assays need to be highly sensitive. In addition, each individual will have a different set of somatic variants, requiring personalized solutions for detection. Therefore, personalized NGS assays with high sensitivity and specificity are needed for MRD diagnostics.

To address this need and empower accurate assessments of MRD, Twist Bioscience has developed the MRD Rapid 500 Panels. This product enables customers to design, manufacture and ship fully personalized MRD panels (up to 500 targets) in as little as six days.

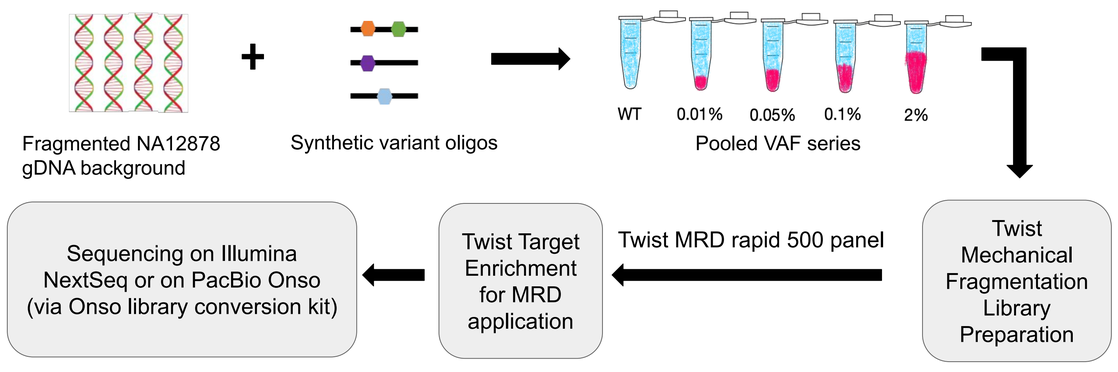
To demonstrate the detection sensitivity of Twist Rapid 500 MRD panels, we designed five custom MRD panels which specifically target somatic variants found in Breast, Lung, CRC, Melanoma and Renal Cell Carcinoma. Each of these MRD panels were designed to include 197 targets, with 3-5 variants per tissue origin and a selection of passenger mutations. Probe sequences of each panel were designed to incorporate the variant allele in the test sample set. To create the sample set, we blended synthetic variant sequences with fragmented cell line gDNA (NA12878) to form a contrived specimen which approximates the profile cell-free and circulating tumor DNA. Five frequency levels were created with average variant allele frequencies (VAFs) of 0% (WT), 0.01%, 0.05%, 0.1% and 2%. Libraries were prepared with UMI adapters and target enrichment was performed using the MRD panels. Same set of samples were sequenced on both Illumina NextSeq and PacBio Onso platforms.

With a sequencing depth of 80,000x, variant calling results revealed that an average of 20 SNV targets can be detected with confidence in the 0.01% VAF samples for each MRD panel, clearly distinguishable from the WT control samples. In addition to demonstrating the accuracy of variant calling by targeting the alternate allele, we showcase the utility of targeting a large number of variants for the detection of an MRD signature at very low levels (e.g. 0.01% VAF).

In summary, the performance of the Twist MRD Rapid 500 Panels showed high detection sensitivity of ultra low-frequency somatic mutations.

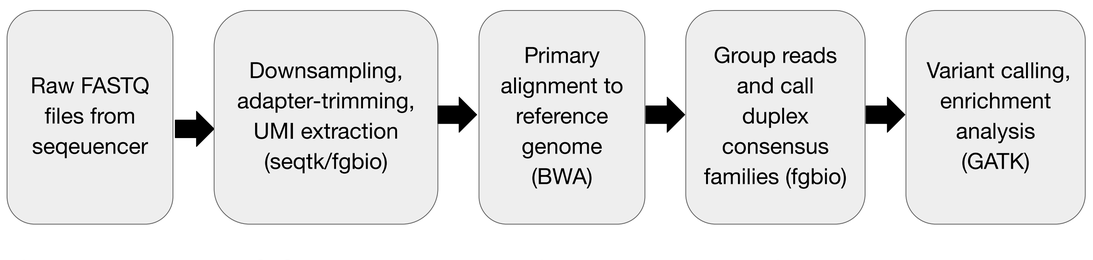
Methods

Experimental methods



To evaluate the detection sensitivity of MRD panels, we generated this pooled VAF series consisting of synthetically designed variant sequences that mimic ctDNA, combined with background gDNA that was fragmented, end-repaired, A-tailed, and purified by Twist EF 2.0 library prep kit and closely mimics the DNA size profile of naive cfDNA. The ctDNA sequences are designed as a tiled pool of ~167 bp sequences that closely mimic natural ctDNA and cover 458 individual mutations including single nucleotide variation, small (2-4bp), medium (5-9bp) and large (10+bp) insertions and deletions. Samples of 5 contrived VAF levels (0% (WT), 0.01%, 0.05%, 0.1%, 2%) were created and QCed by ddPCR, then library prepared with Twist MF kit and UMI. These UMI libraries were captured with 5 Twist Rapid 500 MRD panels (targeted both reference and alternative alleles) using Standard Hybradiation v2 protocol for MRD application. The MRD target enrichment libraries were sequenced using Illumina NextSeq and analyzed using Twist UMI pipeline.

Bioinformatic methods



After base calling and FASTQ generation, reads were first downsampled to a fixed depth based on the target space of the panel. Reads were then pre-processed to mark adapter sequences (Picard) and to isolate UMI sequences (fgbio) into an unaligned BAM file. Raw reads were aligned to the human reference genome (hg38/GRCh38) using BWA, and were merged with the unaligned BAM to provide UMI information. After alignment, UMIs were error-corrected and grouped based on strand and UMI sequence, and consensus reads were called with a duplex strategy (fgbio). Unless otherwise specified, reads were subsequently filtered to keep only duplex consensus families, or those with at least one supporting read derived from each strand. After consensus calling, raw allele counts were obtained using samtools, or variant calls were obtained using Mutect2 (GATK) depending on the specific needs

MRD rapid 500 Panel Performance

Twist designed five 200-probe MRD panels with proprietary algorithms. These five panels specifically target somatic variants found in Breast, Lung, CRC, Melanoma and Renal Cell Carcinoma. To demonstrate panel performance, libraries were prepared with the Twist mechanical library preparation kit with 30ng of Twist WT cfDNA Pan-cancer Reference Standard and the Twist UMI Adapter System for target enrichment and duplex sequencing. A standard hybridization v2 protocol for MRD applications was also developed and optimized to further improve the MRD panel performance.

The Illumina NextSeq sequencing results showed that this upgraded system is able to dramatically improve small panel performance and can reduce the off-target rate in each panel to as low as 10-15% (**Fig. 1A**) with uniform coverage across all targets of interest (**Fig. 1B**). The off target rate didn't greatly increase with UMI consensus analysis pipeline, suggesting the successful probe design kept specific off target rate low.

In addition, no targets dropped out in MRD panels (**Fig. 1C),** indicating the MRD panel manufacture process and standard hybridization v2 protocol for MRD application workflow are highly robust and accurate. With 80000x downsampling, the mean target coverage of each panel after UMI deduplication is ~3000x, generating enough coverage for variant calling analysis (**Fig. 1D**).

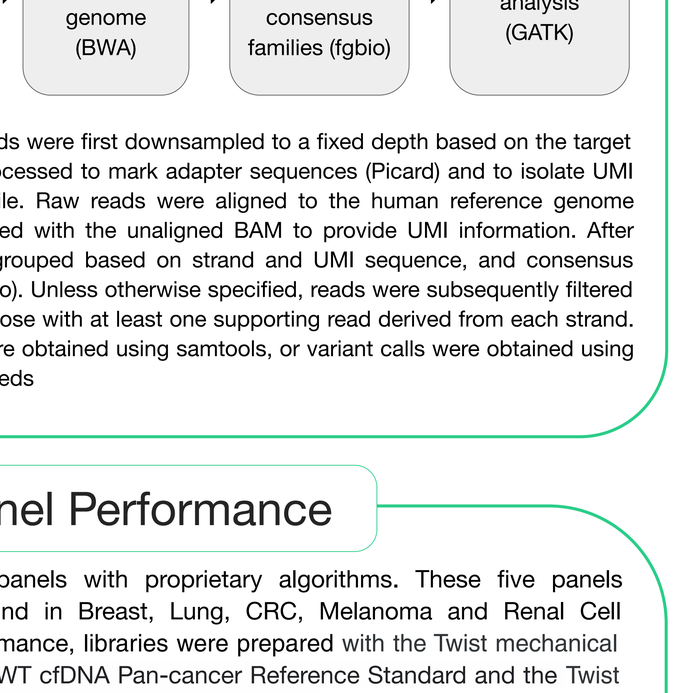


Figure 1. MRD rapid 500 panel performance showing in picard metrics. (A) Off-target rate. (B) Fold-80 uniformity. (C) Mean target coverage. (D) Zero coverage target rate.

MRD rapid 500 panel detection sensitivity

Across all panel types, we detected variant allele frequencies approximately at 60% of the expected dilution frequency (**Fig. 2A**), likely from a combination of capture and alignment bias. The mean error rate for the WT samples was about 10ppm (0.001%), and the margin of error between the WT and the lowest VAF sample (0.01%) was about 7-fold, indicating that WT and 0.01% VAF samples could be clearly distinguished. Five MRD panels targeting different somatic variants also showed very similar performance, demonstrating the consistency of the design strategy across different targets. Incorporating the variant allele rather than the reference allele into the probes improved detection sensitivity by 5-10% across different panel types at the 0.05% and 0.1% VAF levels. At 0.01% VAF level, different MRD panels showed inconsistent positive sites calling results between targeting reference allele and alternative allele, possibly due to sampling error at this low VAF level.

In terms of target sites recall rate (**Fig. 2B)**, almost 100% of targets had detected at 2% VAF level. While VAF impacted sensitivity, variant calling results revealed an average of at least 20/200 SNV targets detected at 0.01% VAF with 80000x sequencing depth across all panels. At 0.05%-0.1% VAF level we saw small but consistent increases in the recall rate (detection defined as at least one supporting duplex consensus read). By targeting the alternate allele in MRD panels, we obtained a ~5-10% improvement in recall for 0.05% samples, but little improvement at 0.01% likely due to sampling effects as discussed above.

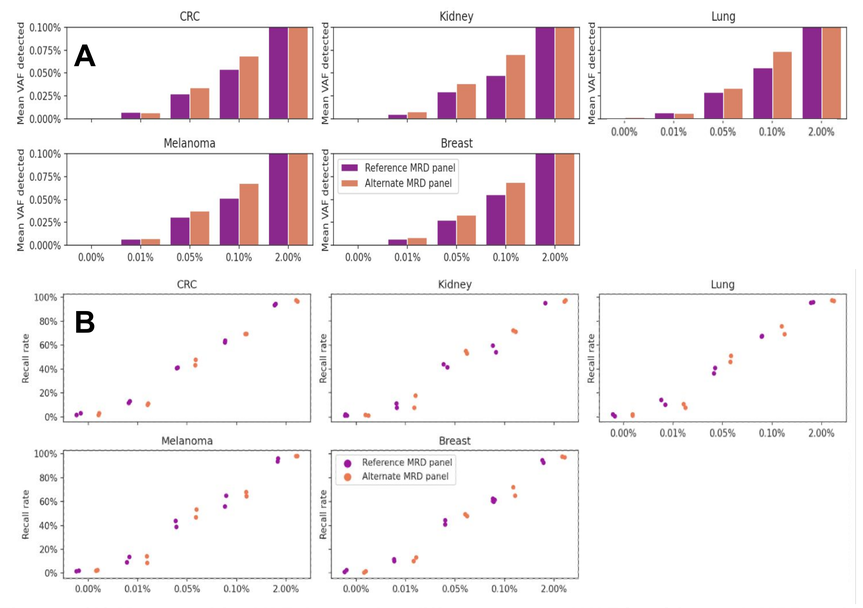


Figure 2 (A) Mean positive VAF detected at different targeted VAF level for each MRD rapid 500 panel. (B) Variant recall rate at different VAF level for each MRD rapid 500 panel

Improving detection sensitivity of INDELs

Insertion–deletion mutations (indels) are supremely important in clinical NGS, as they are implicated as drivers in many cancers. The indel detection rate is affected by the mapping parameters of a short read aligner, normalization scheme for representing indel alignments, as well biases resulting from the use of targeted capture sequencing. As a result, the concordance rate for indel detection tools from short read targeted sequencing can be low.

To further demonstrate the indel detection sensitivity of MRD panels, we performed variant calling by both kmer based searches and raw pileups from duplex-consensus read alignments, and collapsed all the detected variants from each experiment into different conditions.The results showed the combination of targeting alternate alleles for enrichment and K-mer based search methods dramatically improved the detection sensitivity of larger indels (2+bp). ~10% of total indels could be called at a 0.01% VAF level, clearly distinguishable from the WT control samples.

Differences in indel recall rate tend to be most obvious for 0.05-0.1% VAF samples. At 0.1% VAF level, the recall rate can be improved from ~25% to ~75% for large events (>10bp). In addition, medium and large indels are penalized, so even with the alternative panel the improvement can be hard to predict in this range. At 2% VAF level 100% of indels were detected by targeting alternative alleles along with K-mer based search method indicating the advantage of this approach (**Fig. 3A).** Focusing on the mean VAF detection rate at each VAF level, application of K-mer based search and targeting alternative alleles achieves consistent VAFs, which is similar to the target VAFs. Particularly in 2% VAF condition, targeting the alternate allele instead of the reference allele showed much more obvious improvement. Rates of random errors are quite high for SBS and single-base indels, but low for the larger indels. (**Fig. 3B).**

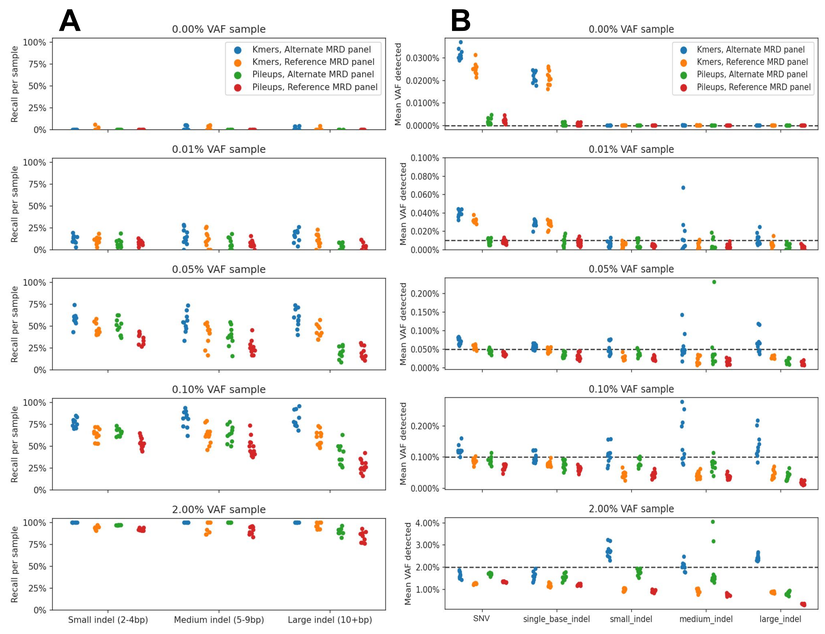


Figure 3. Variant calling analysis split into different variant types: single nucleotide variation (SNV), single-base indel, small indel (2-4bp), medium indel (5-9bp) and large indel (10+bp) categories. Each point represents an experiment, panels are collapsed together into conditions. (A) Different types of INDEL recall rate in each VAF levels. (B) Mean VAF detection rate at each VAF level.

ROC curve analysis

ROC (Receiver Operating Characteristic) analysis was developed as a standard methodology to quantify a signal receiver's ability to correctly distinguish objects of interest from the background noise in the system. ROC curves are frequently used to show the connection/trade-off between clinical sensitivity and specificity for every possible cut-off in a graphical way for a test or a combination of tests. In addition the area under the ROC curve gives an idea about the benefit of using the test(s) in question, which has a meaningful interpretation for disease classification from healthy subjects.

We applied ROC analysis (**Fig. 4**) with MRD target enrichment variant calling dataset to 1) simulate the diagnostic power of Twist MRD panels as compared to approaches that profile smaller numbers of sites, and 2) to find the optimal thresholds of detectable VAF levels and targeted variant sites.

From our ROC analysis, larger target numbers (>50 sites) are beneficial for detection from lower VAF (<0.01%) samples. When the MRD target numbers are lower than 50 sites at 0.01% VAF samples, the simulated ROC curves are close to the diagonal line suggesting the difficulty of distinguishing between true and false positive variants. However, when MRD target numbers are higher than 50 sites at 0.01% VAF level, the MRD test is able to discriminate between true and false positive variants close to 100% sensitivity and 100% specificity. Moreover, even with target numbers as low as 10 sites, the test still shows excellent accuracy and precision at VAFs higher than 0.05%.

In summary, to achieve better detection sensitivity and specificity of MRD test, we suggest to incorporate more target sites (>50 sites) to the MRD panels for samples with 0.01% or lower VAF levels.

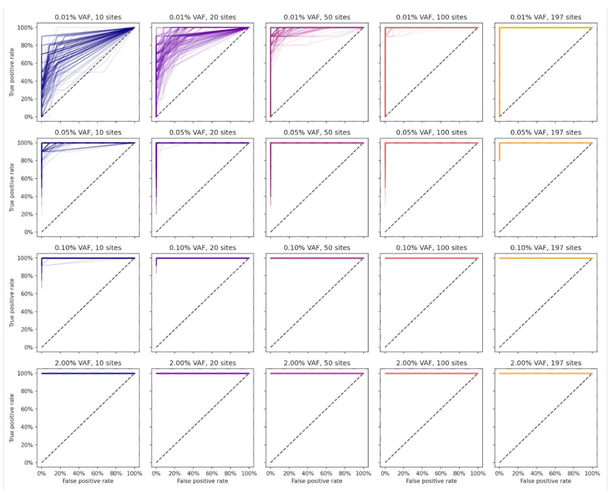
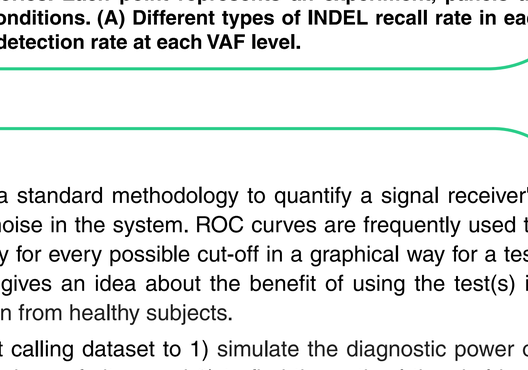


Figure 4. ROC curve analysis to show the connection between sensitivity and specificity with different target sites at each VAF level

PacBio SBB sequencing on the Onso platform

In a head-to-head comparison study at matched coverages using a Twist MRD Rapid 500 panel designed to target somatic variants in breast and lung cancer, SBB achieved significantly higher sensitivity for rare variant detection across all variant allele frequencies tested (WT, 0.01%, 0.05%, 0.1%, and 2.0%) and at all coverage depths as compared to SBS. Further, SBB required lower coverage than SBS to achieve similar sensitivity performance.



References

Li H, Durbin R, Fast and accurate short read alignment with Burrows–Wheeler transform, *Bioinformatics*, Volume 25, Issue 14, 15 July 2009, Pages 1754–1760,

Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 2017 Mar;27(3):491-499.