DATA SHEET

TwinStrand DuplexSeq™ AML-29 MRD v1.0 Assay

High-sensitivity detection of ultra low-frequency mutations in adult acute myeloid leukemia (AML) for residual disease monitoring.

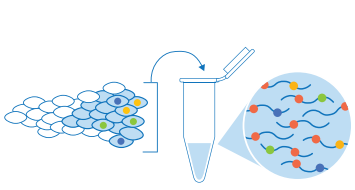
Ultra-Sensitive	DuplexSeq™ technology eliminates the background noise of other NGS methods to enable consistent sub-1/10,000 MRD detection.
Broad Applicability	The 29 gene panel targets loci mutated in 90-95% of adult AMLs.
Standard Workflow	Seamlessly integrates into existing NGS workflows without any major modifications or special equipment.
Bioinformatics Solution	Upload raw sequence files to a secure, cloud-based software pipeline for automated and accurate mutation detection.
Innovative Clinical Research	The presence of MRD after AML treatment is a powerful predictor of long-term outcomes. Streamlined and accurate assessment is an essential tool for developing new therapies.

Introduction

Measurable residual disease (MRD), also called minimal or molecular residual disease, is a prognostic indicator that is important for risk stratification, treatment planning, and relapse monitoring in AML and other malignancies. However, current MRD detection methods are either difficult to standardize, error-prone, or require custom assays to produce reliable results.

Next-generation sequencing (NGS) is a powerful method for detecting residual disease following treatment based on tracking of mutations found in a cancer at time-of-diagnosis (TOD), but its sensitivity is limited due to background noise generated by technical errors.

TwinStrand's AML-29 assay (Table 1) is a research use only MRD solution that detects mutations observed in 90-95% of adult AML cases. Featuring our proprietary DuplexSeq technology, the AML-29 assay yields a sensitivity and specificity markedly higher and more consistent than traditionally-used technologies such as multiparameter flow cytometry (MFC) or, more recently, standard NGS.



Purified DNA

You bring:

- DNA sample
- Water and ethanol

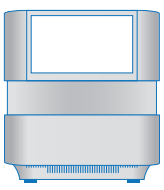
Ideal input material includes fresh/frozen samples such as peripheral whole blood, buffy coat, or bone marrow aspirates.

Library Prep and Target Enrichment

We provide the library prep kit:

- Duplex-optimized library preparation reagents
- DuplexSeq™ adapters
- 48 unique dual indexes
- PCR reagents
- Hybrid capture reagents
- Paramagnetic beads
- TwinStrand DNA technical control

Multi-box configuration facilitates best practices, separating pre- and post-PCR protocol steps.



Sequencing

You bring:

- Illumina® sequencer

Recommended instruments include the Illumina NovaSeq 6000, NextSeq 2000, or HiSeq 4000.

Analysis

We provide the bioinformatics solution:

- Purpose-built for DuplexSeq™ data
- Easy-to-use, secure cloud-based software hosted on DNAnexus®
- Processes raw read files to produce duplex error-corrected sequences, variant calls, and assay quality metrics
- Industry standard formats compatible with existing annotation

Inputs: demultiplexed FASTQ files
Outputs: assay performance report, error corrected BAM, VCF, and CSV files

Figure 1: The AML-29 assay integrates seamlessly into standard NGS lab workflows.

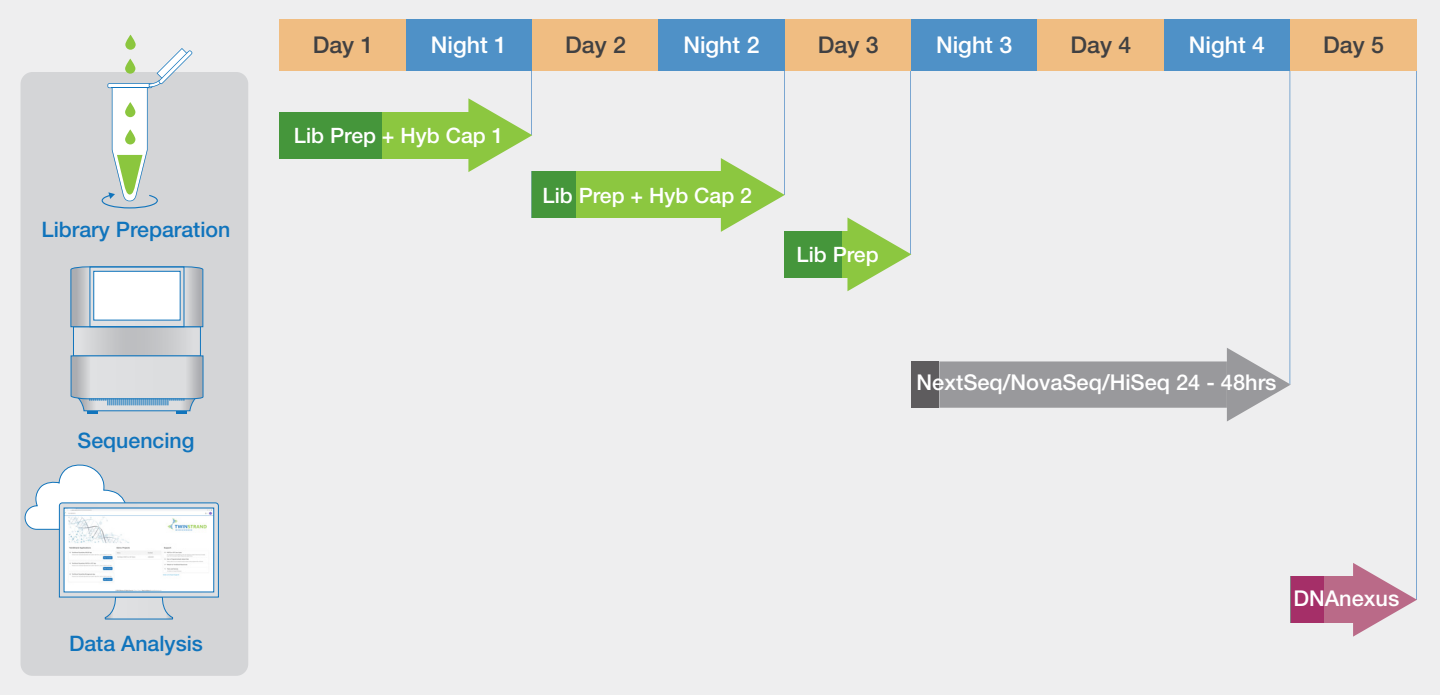


Figure 2: The AML-29 assay workflow provides data within a typical 5 day work week schedule. The hands-on time portion of the workflow fits within normal daytime work hours.

* Arrow length = total time, darker segment of arrow = proportion of total time that is hands-on time. Sample extraction and tertiary analysis (e.g. variant annotation and/or comparison with TOD genotype and interpretation) not included. See Illumina.com for estimated sequencing times by instrument.

Table 1: AML-29 Assay Specifications

Parameter	Specification
Platform	Illumina[a]
Target enrichment method	Hybrid capture
Total probe footprint	58.6 kb
Number of target genes	29
Number of target regions	150
Reference genome	hs38DH (human)
DNA input	50-2,000 ng[b]
Total assay time	3 days[c]
Sequencing run	2 x 150 cycles[d]
Variant types	SNVs, indels, and some structural variants

a. Although our kits are compatible with any Illumina instrument, due to the sequencing depth required we recommend higher capacity instruments such as the Illumina NovaSeq 6000, NextSeq 2000, or HiSeq 4000.

b. The amount of input DNA combined with the number of mutations per AML will determine the maximum achievable MRD sensitivity of the assay.

c. Three days account for library preparation only, including DNA shearing, two hybrid captures, and library quantification and pooling.

d. 2 x 150 cycles is recommended for maximum depth and sensitivity.

Panel Content and Design Strategy

TwinStrand partnered with recognized thought leaders in the hematologic oncology community and considered multiple global guidelines[1,2,3,4] to determine the content for the AML-29 panel. The selected targets were further refined based on peer-reviewed publications and public databases to ensure that the panel targeted key mutations (Table 2) and applied to a significant proportion of the adult AML population. The resultant panel comprises 150 target regions across 29 genes known to be recurrently mutated in adult AML (Table 3). The aggregate 58.6 kb panel (Table 1) balances breadth of applicability across adult AMLs with sequencing cost-effectiveness. The selected regions enable the detection of mutations such as single nucleotide variants (SNVs), insertions and deletions (indels), and some structural variants common in adult AML, clonal hematopoiesis (i.e., CHIP), and myelodysplastic syndromes (MDS).

Table 2: Mutation Coverage Within the AML-29 Assay by Database[a]

Database	Mutation Coverage
COSMIC	95.9%
ClinVar	94.4%
TCGA	97.1%

a. For the 29 target genes, these values represent the percentage of AML-associated mutations in the database that are included in the 150 target regions.

Table 3: Genomic Regions (hs38DH) Included in the AML-29 Panel

Gene	Target Region
ASXL1	a.a. 363-1542
CBL	a.a. 366-477
CEBPA	complete CDS
DNMT3A	a.a. 286-913
EZH2	a.a. 87-208, 244-302, 503-752
FAM5C (BRINP3)	a.a. 80-142, 396-767
FLT3	a.a. 569-647, 807-847
GATA2	a.a. 77-481
HNRNPK	a.a. 21-85, 173-215, 319-336, 371-453
IDH1	a.a. 106-138
IDH2	a.a. 126-178
KIT	a.a. 412-448, 788-828
KRAS	a.a. 1-96
MLL-X (KMT2A-X)	*MLL* intron 9
MYH11-CBFB	*MYH11* intron 30, exon 31
NPM1	a.a. 258-295
NRAS	a.a. 1-96
PHF6	complete CDS
PTEN	a.a. 258-337, 385-440
PTPN11	a.a. 47-110, 484-533
RAD21	complete CDS
RUNX1	complete CDS
SMC1A	a.a. 38-99, 447-515, 578-637, 687-732, 772-902, 1096-1145
SMC3	a.a. 184-268, 365-435, 656-705, 882-1035, 1100-1158
STAG2	a.a. 42-128, 155-297, 436-472, 513-546, 578-675, 787-844, 892-1155
TET2	a.a. 1-2003
TP53	complete CDS
U2AF1	a.a. 16-44
WT1	a.a. 372-523

a.a. = amino acids
CDS = Coding Sequence

Integrated Library Preparation and Data Analysis Workflow

The AML-29 assay seamlessly integrates with commonly used Illumina NGS instruments and workflows, allowing labs to begin collecting informative data soon after obtaining their first assay kit (Figure 1). The workflow does not require any major modifications or special equipment.

Library preparation begins with input of 50-2,000 ng of high quality genomic DNA (Table 1), ideally extracted from fresh/frozen samples such as peripheral whole blood, buffy coat or bone marrow aspirates, that has been sheared to ~300 bp. The kit includes all reagents necessary for library preparation* and hybrid capture, including duplex-barcoded adapters and 48 unique dual indexes. TwinStrand designed the workflow for use with 96 (or 48) well plates and is compatible with liquid-handling platforms. Automated library preparation protocols enable increased throughput and reproductivity while reducing hands-on time. A typical library preparation requires 3 days and could be sequenced in 1-2 days on average. Thus, results can be available within a standard 5 day work week, with hands-on time only required during normal 8 hr daytime work hours (Figure 2).

* Except water and ethanol

Sequencing data generated from AML-29 is analyzed using TwinStand's cloud-based data analysis software, hosted on the secure DNAnexus[5] platform. Following the upload of demultiplexed FASTQ files, the software performs alignment, inter-species decontamination, double strand consensus making with our proprietary DuplexSeq technology, and variant calling to generate error-corrected VCF, MUT, and BAM files, as well as an assay performance report in html and CSV format. The output reports SNVs, indels (inclusive of *NPM1*), internal tandem duplications (i.e. in *FLT3*), and mappable structural variants. Output files are compatible with third-party tertiary analysis software for functional annotation and clinical reporting. The MUT file is provided in an easy-to-read tab-delimited file format that can be viewed using standard desktop spreadsheet software. It includes information such as the base position, variant type, reference and alternative alleles present at each site, as well as the depth of the alternate allele and total allelic depth at a given position, which can be used to calculate the variant allele frequency (VAF). The VCF file is provided in an industry-standard format compatible with third-party tertiary analysis and clinical reporting tools.

Assay Performance

AML-29 validation experiments have shown repeatable and reproducible performance metrics with high and reliable on-target rate (~95%) and uniformity (100% of target intervals have a mean duplex depth ≥20% of panel-wide mean duplex depth, Table 4). Uniformity was consistent across a six-fold range of DNA inputs (Figure 3). To demonstrate the superior sensitivity of DuplexSeq technology at lower limits of detection, an analytical validation experiment was performed after diluting cell line genomic DNAs containing specific AML-associated mutations into peripheral blood-derived DNA from a healthy young donor at controlled ratios. The results show that AML-29 can reliably detect SNVs and indels (including *FLT3* ITDs) at variant allele frequencies (VAF) from 1% to less than 0.01% (1 in 10,000) (Table 5).

Table 4: Assay Performance Metrics

Parameter	Metric
On-target bases[a]	95.1%
Target bases with ≥1,000x duplex depth	100%[b]
Uniformity (targets)[c]	• 100% of target intervals have mean duplex depth ≥20% of panel-wide mean duplex depth • 74% of target intervals have mean duplex depth ≥80% of panel-wide mean duplex depth

a. On-target percentage is derived from raw reads.

b. The minimum duplex depth recorded from target regions was at least 2,500x for 250 ng and at least 10,000x for 1,500 ng.

c. Uniformity is typically defined as the percentage of bases in all targeted regions that is covered by at least 20% of the average base coverage depth reads - we include here a more stringent threshold of 80% is included as well.

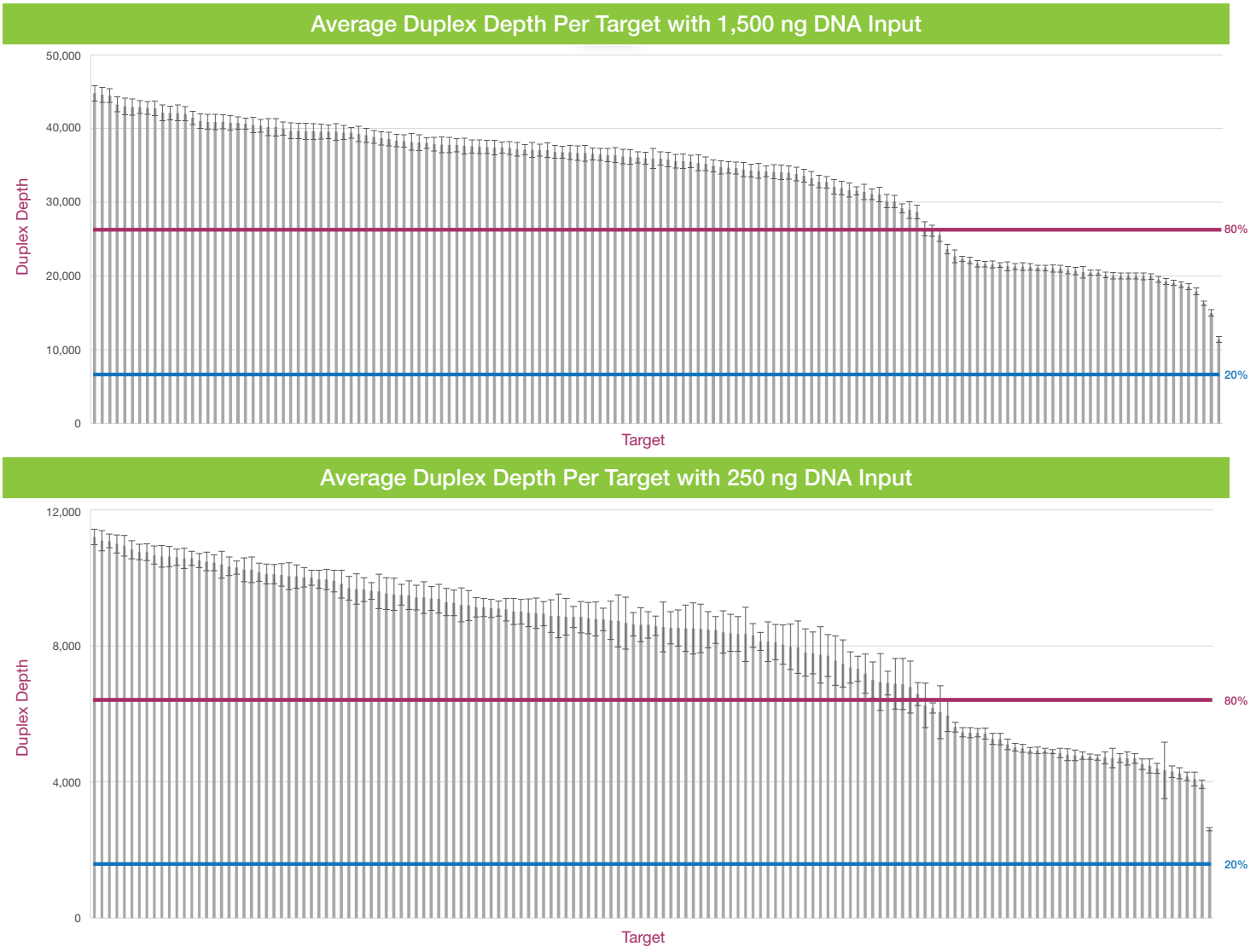


Figure 3: The mean duplex depth per target interval is plotted for AML-29 libraries using an input of 1,500 ng (top) and 250 ng (bottom) of DNA input (pre-fragmented mass) as prepared per the TwinStrand DuplexSeq Kit™ Manual. The mean duplex depth (± standard deviation) is represented as plots from 4 replicates at each of the 150 target intervals. The horizontal lines in each plot indicate 80% (top) and 20% (bottom) of the overall mean duplex depth.

Table 5: Experimental Performance Metrics

Metric		Result
Analytical sensitivity[a]	SNV	100% for predicted VAF down to $3.9x10^{-6}$ (0.00039%) VAF
	FLT3 ITD (21 bp)	100% for predicted VAF down to $3.0x10^{-5}$ (0.003%) VAF
	NPM1 insTCTG	100% for predicted VAF down to $3.0x10^{-5}$ (0.003%) VAF
	Additional indels	100% for predicted VAF down to $5.0x10^{-5}$ (0.005%) VAF
Analytical specificity[b]	SNVs and indels (including ITDs)	100% for mutant alleles down to $1.3x10^{-5}$ (0.0013%) VAF

a. Sensitivities are shown for the predicted VAF of the 21 investigated variants from the following mean duplex depths: SNV: 138,964x, *FLT3* ITD: 133,466x, *NPM1* insTCTG: 64,422x, additional indels 137,070x. Duplex depth corresponds to the number of unique DNA molecules interrogated at the specified nucleotide position.

b. Mutant alleles at the 21 investigational sites were detected in only 4 molecules out of 2,993,429x combined duplex molecular depth in the 1,500 ng input negative control libraries (overall $1.3x10^{-6}$). The highest single VAF was $1.3x10^{-5}$ (approximately 1 in 80,000). These frequencies are consistent with age-associated biological background mutation levels[6].

The analytical validation experiment further demonstrated high accuracy and precision throughout 4 serial dilutions (for *FLT3* and *NPM1*) or a single mix (for SNVs and other indels) each sequenced in 4 replicates. These results demonstrated confident detection for 21 known variants (SNVs, indels, and ITDs in 13 genes commonly mutated in adult AML including *TP53, DNMT3A, NMP1*, and *FLT3*) at expected VAFs from $1.0x10^{-2}$ to $3.9x10^{-6}$ (Figure 4). The complete set of mutations is listed in Table 6.

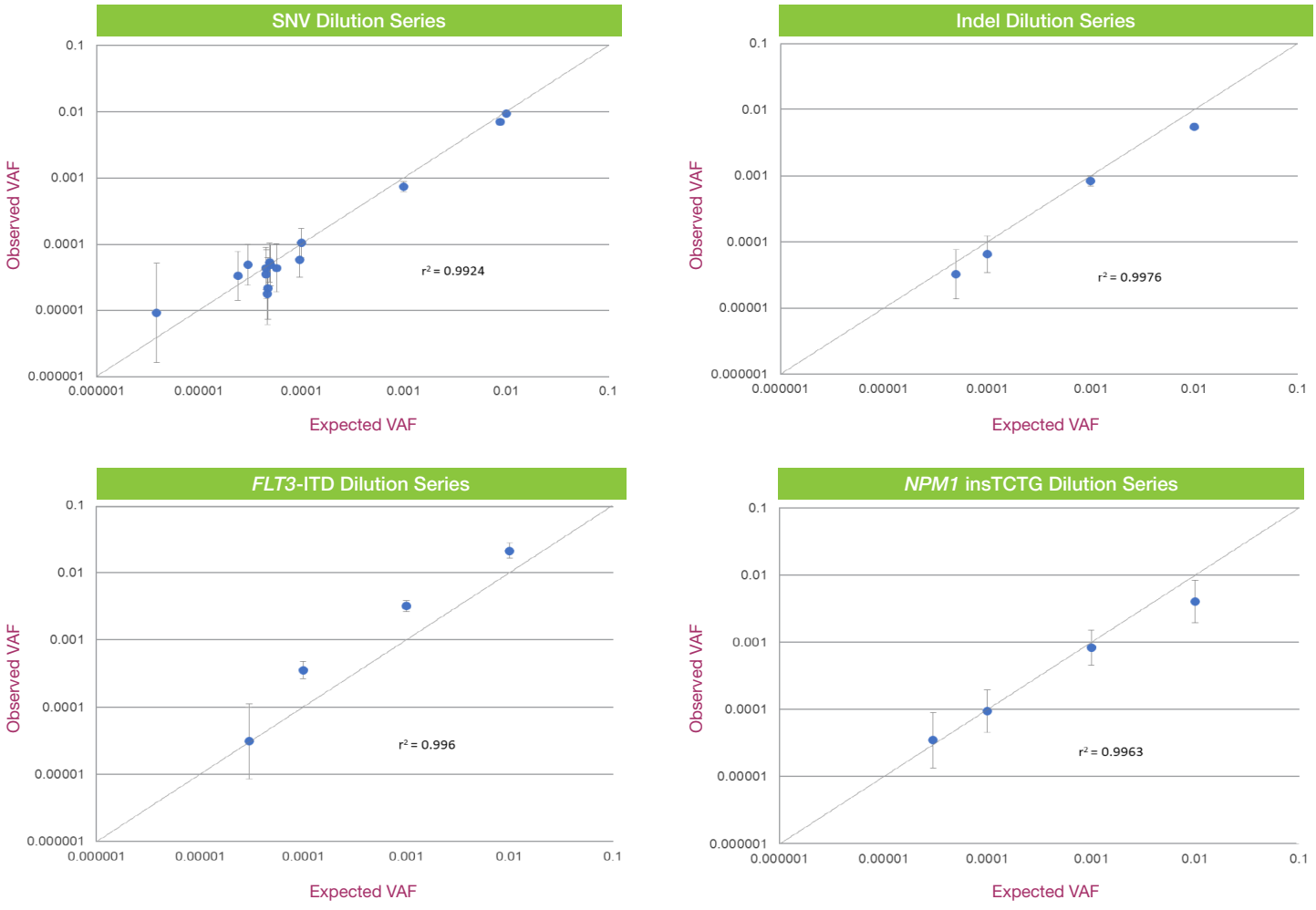


Figure 4: Concordance between a dilution series from a reference standard with known spiked-in mutations (expected) and the VAF measured by duplex sequencing with the AML-29 panel (observed). Data points represent the total VAF from 4 combined technical replicate libraries and are shown with 95% Wilson binomial confidence intervals. Zygosity was independently verified in all cell lines prior to dilution when available (notably this excludes *FLT3*-ITD, which may explain the slight deviation from expectation).

Table 6: Concordance of Expected VAF in a DNA Standard vs Observed Duplex Sequencing VAF

Gene and Variant	Expected	Observed	Gene and Variant	Expected	Observed
TP53 C277F	$1.0x10^{-2}$	$9.4x10^{-3}$	*EZH2* S734S	$5.7x10^{-5}$	$4.4x10^{-5}$
FLT3 ITD	$1.0x10^{-2}$	$2.1x10^{-2}$	*TP53* S90fs	$5.0x10^{-5}$	$3.2x10^{-5}$
NPM1 insTCTG	$1.0x10^{-2}$	$4.1x10^{-3}$	*RAD21* D116E	$5.0x10^{-5}$	$4.8x10^{-5}$
TP53 Q136fs	$1.0x10^{-2}$	$5.5x10^{-3}$	*HNRNPK* Q182H	$4.9x10^{-5}$	$5.2x10^{-5}$
NRAS Q61R	$8.8x10^{-3}$	$7.1x10^{-3}$	*DNMT3A* M761I	$4.7x10^{-5}$	$2.1x10^{-5}$
TET2 R881G	$1.0x10^{-3}$	$7.4x10^{-4}$	*ASXL1* W796C	$4.7x10^{-5}$	$1.8x10^{-5}$
FLT3 ITD	$1.0x10^{-3}$	$3.2x10^{-3}$	*KRAS* G13D	$4.6x10^{-5}$	$3.5x10^{-5}$
NPM1 insTCTG	$1.0x10^{-3}$	$8.4x10^{-4}$	*BRINP3* Q493K	$4.5x10^{-5}$	$4.3x10^{-5}$
TP53 V218fs	$1.0x10^{-3}$	$8.4x10^{-4}$	*FLT3* ITD	$3.0x10^{-5}$	$3.1x10^{-5}$
TET2 L1329M	$1.0x10^{-4}$	$1.0x10^{-4}$	*NPM1* insTCTG	$3.0x10^{-5}$	$3.5x10^{-5}$
FLT3 ITD	$1.0x10^{-4}$	$3.5x10^{-4}$	*KRAS* G12V	$3.0x10^{-5}$	$4.9x10^{-5}$
NPM1 insTCTG	$1.0x10^{-4}$	$9.4x10^{-5}$	*TP53* R273H	$2.4x10^{-5}$	$3.3x10^{-5}$
PTEN R130fs	$1.0x10^{-4}$	$6.6x10^{-5}$	*RUNX1* Q264R	$3.9x10^{-6}$	$9.1x10^{-6}$
BRINP3 C726*	$9.6x10^{-5}$	$5.8x10^{-5}$			

Determining Assay Sensitivity

Determining the sensitivity for detecting AML-MRD using the TwinStrand DuplexSeq AML-29 panel is based upon two key variables: DNA input and the number of informative AML mutations. An informative AML mutation is a nucleotide variant within the target region that identifies a DNA molecule derived from AML cells, as determined by sequencing of a disease sample obtained at the TOD or relapse. As DNA input or number of informative mutations are increased, sensitivity for detecting the MRD clone also increases (Table 7). Data are shown for both the conservative situation of only one informative mutation and for 3 mutations (the average number of mutations expected for adult AML patients using this panel). The achievable, sufficiently-powered sensitivity is an interplay between DNA input (i.e. material available, which dictates potential duplex depth of sequencing), and the number of mutations that mark a given individual's AML.

Table 7: Estimated Sensitivity and Number of Samples Achievable per Flow Cell Based on DNA Input

DNA Input (ng)[a]	Number of Clusters (M)[b]	Mean Duplex Depth[c]	Number of Samples by Instrument and Flowcell[d]			AML Clone Frequency that can be Detected with 95% Power[e]			
			NovaSeq 6000 S4	NextSeq 2000 P3	NextSeq 550 HO	One Informative Mutation		Three Informative Mutations	
50	15	700	660[d]	66	26	1 in 233	0.43%	1 in 700	0.14%
250	75	5,500	132	13	5	1 in 1,833	0.055%	1 in 5,500	0.018%
500	150	11,000	64	6	2	1 in 3,667	0.027%	1 in 11,000	0.009%
1,000	300	22,000	32	3	1	1 in 7,333	0.014%	1 in 22,000	0.005%
1,500	450	28,000	20	2	0	1 in 9,333	0.0107%	1 in 28,000	0.0036%
2,000	600	33,000	16	1	0	1 in 11,000	0.0091%	1 in 33,000	0.0030%

a. These DNA requirements assume a high molecular weight genomic DNA quantified pre-fragmentation. DNA requirements will differ if using cell-free DNA (higher efficiency) or heavily damaged DNA (lower efficiency) such as that from FFPE.

b. A cluster equals two paired-end reads (i.e. 15M clusters = 30M PE reads). Data estimates assume 150 bp paired-end reads.

c. This assumes libraries prepared per the library prep user manual. Represented mean duplex depth may vary.

d. S4 (v1.5), P3, and HO (v2.5) are the flow cell types shown. TwinStrand provides 48 unique dual indexes in assay kits although 96 are available upon request. Thus, for the S4 flow cell with Xp workflow, up to 96 samples can be run on each of 4 lanes for a total of 384 samples total (660 samples is not currently possible).

e. One informative mutation is a conservative estimate while 3 informative mutations is the expected average in this panel for adult AMLs. Table values represent the clone frequencies that can be detected with 95% power assuming Poisson binomial sampling. Certain mutations commonly found in clonal hematopoiesis (CH) may be less reliable indicators of true AML-MRD, as their residual presence after treatment may reflect precursor CH, rather than AML.[7]

Summary

TwinStrand's AML-29 assay is an NGS, hybrid capture-based assay that utilizes our proprietary DuplexSeq™ technology to enable accurate detection of variants present at extremely low levels (sub-1/10,000). The highly-optimized probe set and core chemistry provides comprehensive ultra-deep coverage of regions in 29 genes that are frequently mutated in adult AML, making it ideally suited for clinical research studies assessing MRD following therapy. The assay includes a sophisticated, easy-to-use bioinformatics pipeline that automatically generates data in industry-standard formats. With the AML-29 assay, labs can be confident that they are identifying only true variants down to a predetermined (and adjustable) limit of detection. The exceptionally clean data allows for higher-sensitivity MRD detection than other NGS-based technologies and significant time savings that stem from obviating the need for supplemental bioinformatics and manual review to flag potential artifacts.

Ultra-Sensitive DuplexSeq technology eliminates the background noise of other NGS methods to enable consistent sub-1/10,000 MRD detection.

Broad Applicability The 29 gene panel targets loci recurrently mutated in 90-95% of adult AMLs.

Standard Workflow Seamlessly integrates with existing NGS workflows without any major modifications or special equipment.

Bioinformatics Solution Upload raw sequence files to a secure, cloud-based software pipeline for automated and accurate mutation detection.

Innovative Clinical Research The presence of MRD after AML treatment is a powerful predictor of long-term outcomes and streamlined, accurate assessment is an essential tool for developing new therapies.

Ordering Information

Product Number	Product Description
06-1002-02	TwinStrand Duplex Sequencing™ AML-29 Kit (Human), v1.0 - 24x
06-1002-03	TwinStrand Duplex Sequencing™ AML-29 Kit (Human), v1.0 - 48x

For more information on the TwinStrand DuplexSeq™ AML-29 MRD assay or how to order, contact us at **www.twinstrandbio.com/contact.**

References

1. NCCN Guidelines. https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1411. Accessed November 17, 2021.
2. ASH Guidelines. https://ashpublications.org/bloodadvances/article/4/15/3528/461693/American-Society-of-Hematology-2020-guidelines-for. Accessed November 17, 2021.
3. ESMO Guidelines. https://www.annalsofoncology.org/article/S0923-7534(20)36079-8/fulltext. Accessed November 17, 2021.
4. ELN Guidelines. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5291965/. Accessed November 17, 2021.
5. DNAnexus https://www.dnanexus.com/privacy.
6. Risques RA, Kennedy SR. (2018) Aging and the rise of somatic cancer-associated mutations in normal tissues. PLoS Genet. 4;14(1):e1007108 PMID: 29300727.
7. Jongen-Lavrencic M, et al. (2018) Molecular Minimal Residual Disease in Acute Myeloid Leukemia. N Engl J Med. 29;378(13):1189-1 199. PMID: 29601269.

TWINSTRAND™
BIOSCIENCES

3131 Elliott Ave, Suite 750 | Seattle, WA 98121 | +1(877) 202-TWIN | support@twinstrandbio.com

MKT0022 v1.0