Poster Number: 9063F

Design and development of 2 Mb custom pan-cancer panel to achieve improved sequencing coverage and economy

Peilin Chen[1*], David Wang[1], Katica Ilic[1], Michael Salgado[1], Ekaterina Star,[1] Longhui Ren[2], Shengyao Chen[1] and Bosun Min[1]

[1] Research & Development, Integrated DNA Technologies, Redwood City, CA [2]Research & Development, Integrated DNA Technologies, Coralville, IA

* Corresponding author: pchen@idtdna.com

Introduction

Hybridization capture technology has been widely used for enrichment of specific targets of interests from biological specimens. The continuously evolving list of onco-biomarkers requires researchers to quickly update the target regions for their applications. Hybridization technologies face challenges in designing proper probes for difficult targets and capturing off-target regions.

We leveraged our internal proprietary algorithm to design a pan-cancer panel targeting 536 cancer-relevant genes, selected non-coding regions, and actionable biomarkers. The panel design took challenging genomic regions into consideration to improve sequencing coverage and uniformity, across a total of 2.04 Mb of target space. The new panel design, when tested on commercial reference specimens (*n* = 13), FFPE DNA specimens with DIN 1.8-4.8 (*n* = 30), and tissue specimens (*n* = 24), demonstrated improved panel performance in terms of on target rate (>95 % selected bases) and uniformity (fold-80 base penalty of ≤1.3 for cell line and reference DNA samples) enabling us to improve on the sequencing economy. IDT's proprietary design panels combined with NGS library prep and hybridization kits produce reliable NGS data from highly degraded DNA specimens at lower input.

Table 1: Genes included in IDT's pan-cancer panel. The panel includes 529 cancer genes covering 226 cancer types and 7 pharmacogenomics genes

ABL1	ATRX	BRCA2	CDK6	CTNNB1	EIF1AX	ETV1	FGF23	FYN	HGF	HSP90AA1	JAK3	LYN	MITF	NEGR1	PAK7	PIK3CG	PTCH1	REL	SGK1	SRC	TFRC	XIAP
ABL2	AURKA	BRD4	CDK8	CUL3	EIF4A2	ETV4	FGF3	G6PD	HIST1H1C	ICOSLG	JUN	LZTR1	MLH1	NF1	PALB2	PIK3R1	PTEN	RET	SH2B3	SRSF2	TGFBR1	XPO1
ACVR1	AURKB	BRIP1	CDKN1A	CUX1	EIF4E	ETV5	FGF4	GABRA6	HIST1H2BD	ID3	KAT6A	MAF	MLL	NF2	PARK2	PIK3R2	PTPN11	RFWD2	SH2D1A	STAG1	TGFBR2	XRCC2
ACVR1B	AXIN1	BTG1	CDKN1B	CXCR4	EML4	ETV6	FGF5	GATA1	HIST1H3A	IDH1	KDM5A	MAGI2	MLLT3	NFE2L2	PARP1	PIK3R3	PTPRD	RHEB	SHQ1	STAG2	TMEM127	YAP1
AKT1	AXIN2	BTK	CDKN2A	CYLD	EMSY	EWSR1	FGF6	GATA2	HIST1H3B	IDH2	KDM5C	MALT1	MPL	NFKBIA	PAX3	PIM1	PTPRS	RHOA	SLIT2	STAT3	TMPRSS2	YES1
AKT2	AXL	C11ORF30	CDKN2B	CYP2C9	EP300	EZH2	FGF7	GATA3	HIST1H3C	IFNGR1	KDM6A	MAP2K1	MRE11A	NKX2-1	PAX5	PLCG2	PTPRT	RICTOR	SLX4	STAT4	TNFAIP3	ZBTB2
AKT3	B2M	CALR	CDKN2C	CYP2D6	EPCAM	FAM175A	FGF8	GATA4	HIST1H3D	IGF1	KDR	MAP2K2	MSH2	NKX3-1	PAX7	PLK2	QKI	RIT1	SMAD2	STAT5A	TNFRSF14	ZBTB7A
ALK	BAP1	CARD11	CEBPA	DAXX	EPHA3	FAM46C	FGF9	GATA6	HIST1H3E	IGF1R	KEAP1	MAP2K4	MSH3	NOTCH1	PAX8	PMAIP1	RAB35	RNF43	SMAD3	STAT5B	TOP1	ZFHX3
ALOX12B	BARD1	CASP8	CENPA	DCUN1D1	EPHA5	FANCA	FGFR1	GEN1	HIST1H3F	IGF2	KEL	MAP3K1	MSH6	NOTCH2	PBRM1	PMS1	RAC1	ROS1	SMAD4	STK11	TOP2A	ZNF217
AMER1	BBC3	CBFB	CHD2	DDR2	EPHA7	FANCC	FGFR2	GID4	HIST1H3G	KBKE	KIF5B	MAP3K13	MSI	NOTCH3	PDCD1	PMS2	RAD21	RPS6KA4	SMARCA4	STK40	TP53	ZNF703
ANKRD11	BCL10	CBL	CHD4	DDX41	EPHB1	FANCD2	FGFR3	GLI1	HIST1H3H	KZF1	KIT	MAP3K14	MST1	NOTCH4	PDCD1LG2	PNRC1	RAD50	RPS6KB1	SMARCB1	SUFU	TP63	ZRSR2
ANKRD26	BCL2	CCND1	CHEK1	DHX15	EPHB4	FANCE	FGFR4	GNA11	HIST1H3I	IL10	KLF4	MAP3K4	MST1R	NPM1	PDGFRA	POLD1	RAD51	RPS6KB2	SMARCD1	SUZ12	TPMT	
APC	BCL2L1	CCND2	CHEK2	DICER1	ERBB2	FANCF	FH	GNA13	HIST1H3J	IL7R	KLHL6	MAPK1	MTOR	NRAS	PDGFRB	POLE	RAD51B	RPTOR	SMC1A	SYK	TRAF2	
AR	BCL2L11	CCND3	chrY	DIS3	ERBB3	FANCG	FLCN	GNAQ	HIST2H3A	INHA	KMT2B	MAPK3	MUTYH	NRG1	PDK1	PPARG	RAD51C	RUNX1	SMC3	TAF1	TRAF7	
ARAF	BCL2L2	CCNE1	CIC	DNAJB1	ERBB4	FANCI	FLI1	GNAS	HIST2H3D	INHBA	KMT2C	MAX	MYB	NSD1	PDPK1	PPM1D	RAD51D	RUNX1T1	SMO	TBX3	TSC1	
ARFRP1	BCL6	CD274	CREBBP	DNMT1	ERCC1	FANCL	FLT1	GPR124	HIST3H3	INPP4A	KMT2D	MCL1	MYC	NSD2	PGR	PPP2R1A	RAD52	RYBP	SNCAIP	TCEB1	TSC2	
ARID1A	BCOR	CD276	CRKL	DNMT3A	ERCC2	FAS	FLT3	GPS2	HLA-A	INPP4B	KRAS	MDC1	MYCL1	NTRK1	PHF6	PPP2R2A	RAD54L	SDHA	SOCS1	TCF3	TSHR	
ARID1B	BCORL1	CD74	CRLF2	DNMT3B	ERCC3	FAT1	FLT4	GREM1	HLA-B	INSR	LAMP1	MDM2	MYCN	NTRK2	PHOX2B	PPP6C	RAF1	SDHAF2	SOX10	TCF7L2	U2AF1	
ARID2	BCR	CD79A	CSF1R	DOT1L	ERCC4	FBXW7	FOXA1	GRIN2A	HLA-C	IRF2	LATS1	MDM4	MYD88	NTRK3	PIK3C2B	PRDM1	RANBP2	SDHB	SOX17	TEK	UGT1A1	
ARID5B	BIRC3	CD79B	CSF3R	DPYD	ERCC5	FGF1	FOXL2	GRM3	HNF1A	IRF4	LATS2	MED12	MYOD1	NUDT15	PIK3C2G	PREX2	RARA	SDHC	SOX2	TERC	VEGFA	
ASXL1	BLM	CDC73	CSNK1A1	E2F3	ERG	FGF10	FOXO1	GSK3B	HNRNPK	IRS1	LMO1	MEF2B	NAB2	NUP93	PIK3C3	PRKAR1A	RASA1	SDHD	SOX9	TERT	VHL	
ASXL2	BMPR1A	CDH1	CTCF	EED	ERRFI1	FGF14	FOXP1	H3F3A	HOXB13	IRS2	LoH	MEN1	NBN	NUTM1	PIK3CA	PRKCI	RB1	SETBP1	SPEN	TET1	VTCN1	
ATM	BRAF	CDK12	CTLA4	EGFL7	ESR1	FGF19	FRS2	H3F3B	HRAS	IAK1	LRP1B	MET	NCOA3	PAK1	PIK3CB	PRKDC	RBM10	SETD2	SPOP	TET2	WISP3	
ATR	BRCA1	CDK4	CTNNA1	EGFR	ETS1	FGF2	FUBP1	H3F3C	HSD3B1	IAK2	LTK	MGA	NCOR1	PAK3	PIK3CD	PRSS8	RECQL4	SF3B1	SPTA1	TFE3	WT1	

Methods

Thirteen commercial reference specimens, 30 FFPE DNA specimens (DIN 1.8-4.8) and 24 tissue specimens were included in the study. For IDT workflow, 25 ng DNA input was taken into IDT xGen™ cfDNA & FFPE DNA library prep kit workflow followed by target enrichment by IDT xGen Hybridization Wash kit V2 using the panels designed through our proprietary design tool. The data analyses were conducted through internal NGS data analysis pipelines, where we focused on panel performance defined by Picard metrics. On-target rate was measured by % selected bases, and uniformity was measured using fold-80 base penalty and % target base coverage within 2X of the mean. For the Illumina® TruSight™ Oncology 500 DNA kit (TSO500), protocols recommended by Illumina® were employed using 40 ng DNA as input, and the sequencing data was compared to IDT workflow.

Results

Robust NGS metrics across diverse specimens.

IDT's pan-cancer panel was tested on five distinct sample types by three users, demonstrating consistent performance across all specimens (Figure 1). NGS metrics showed high target enrichment; % selected base for Reference, FFPE and Tissue specimens were 95.8% ± 0.5%, 95.4% ± 0.9%, and 95.3% ± 0.3%, respectively. Fold-80 base penalties were 1.18 ± 0.03, 1.57 ± 0.21, and 1.27 ± 0.13, respectively. Additionally, over 97% of bases were within 2X of the mean, indicating high uniformity and robustness across users and sample types.

Enhanced sequencing efficiency over TSO500

To compare our in-house panel with TSO500, matched FFPE samples were processed using each workflow—TSO500 libraries were prepared according to the manufacturer's recommended protocol, while the same specimens were processed using IDT's reagents and pipelines (Figure 2). The TSO500 libraries showed an average of 85.4% selected bases and a fold-80 base penalty of 1.42. In contrast, our workflow's improved upon on-target rate (% selected bases >95) and uniformity (Fold-80 base penalty 1.57 ± 0.21). These improved metrics enabled sequencing of 24 DNA specimens on a single NextSeq550 High Output Flowcell while achieving coverage equivalent to the TSO500 assay that would achieve 2-3-fold fewer samples per flowcell.

Figure 1: NGS metrics demonstrate consistent performance of IDT's pan-cancer panel across five distinct sample types, including cfDNA, reference, tissue, and FFPE specimens tested by three users. All sample types exhibited comparable results, indicating robustness of the panel regardless of specimen origin.

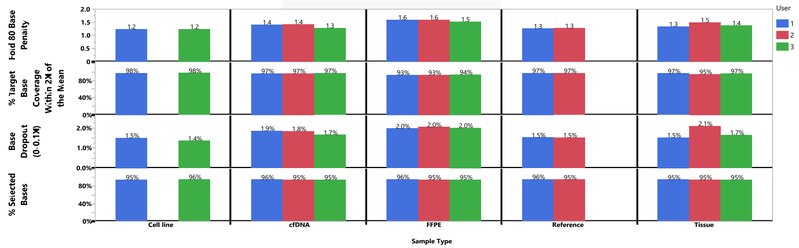
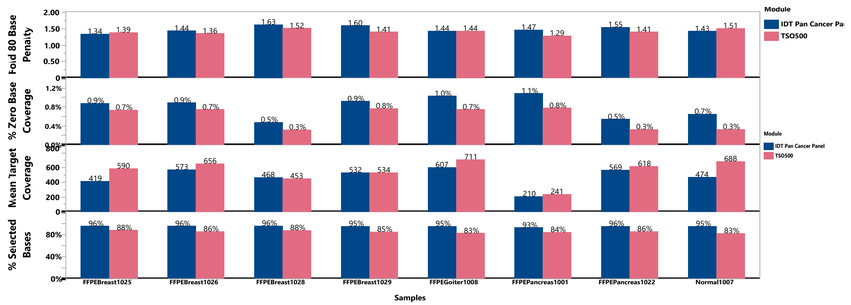


Figure 2: NGS metrics comparison of 8 DNA samples between IDT pan-cancer panel and TSO500 demonstrate performance equivalence and sequencing economy when 24 libraries prepared with IDT pan-cancer panel were loaded on one NextSeq550 High Output flowcell.



Mutation detection concordance and accuracy

SNV, indel and CNV are used to demonstrate mutation detection concordance and accuracy.

• SNV detection

Eight samples (4 frozen tissue, 2 highly degraded FFPE (DIN 1.8 and 4.1) and 2 cell line DNA samples) were tested with the panel at 25 ng DNA input. The same samples were tested with TSO500 kit at 40 ng DNA input. The results are shown in Table 2 and Figure 2.

Table 2. SNV detection concordance with TSO500

	Total SNVs	Concordance with TSO500	95% CI
LOD	6939	1%	0.34%-1.66%
PPA		99.7%	99.5%-99.97%
PPC		98.4%	98.1%-98.6%

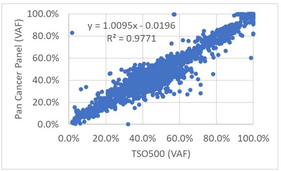


Figure 3. Correlation of SNV detection between IDT's pan-cancer panel with TSO500

• Indel detection

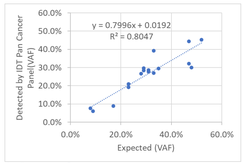


Figure 4. Indel detection
A reference HD827 (Horizon Discovery) was used to evaluate indel detection for the pan-cancer panel. Seventeen of 18 indels in the sample were detected (PPA 94.4% and PPV 100%). The correlation between the expected and detected VAF is R=0.897.

• CNV detection

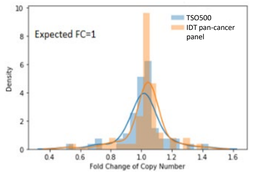


Figure 5. Fold change distribution of a cell line sample (KM12L4). Comparison between IDT pan-cancer panel and TSO500

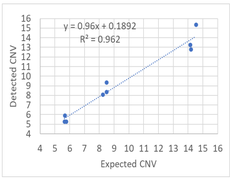


Figure 6. Correlation of CNV calling. SeraSeq Breast CNV Mixes (3, 6 and 12 copies) were used to evaluate IDT pan-cancer panel CNV calling accuracy. The correlation with the expected copy number (R) is 0.992.

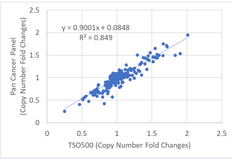


Figure 7. CNV calling of IDT pan-cancer panel vs. TSO500. The same 8 DNA samples were tested using both IDT pan-cancer panel and TSO500. The correlation of the CNV calling (R) is 0.92.

Conclusion

IDT's pan-cancer panel demonstrated consistent and high-quality performance across diverse sample types and users. Compared to the TSO500 assay, our workflow showed improved on-target rates and uniformity, enabling higher throughput without compromising coverage and cancer mutation detection accuracy.