Broadening the input application of ResolveDNA genomic amplification to detection of SNV in circulating DNA

Jon S. Zawistowski, Swetha D. Velivela, Isai Salas-Gonzalez, Megan Justice, Tia A. Tate, Victor J. Weigman

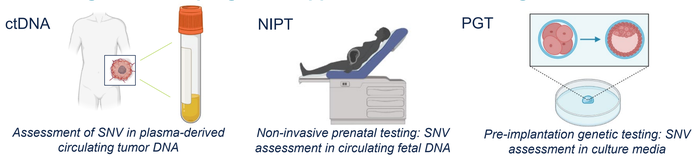
BioSkryb Genomics, Durham, NC, USA

Abstract

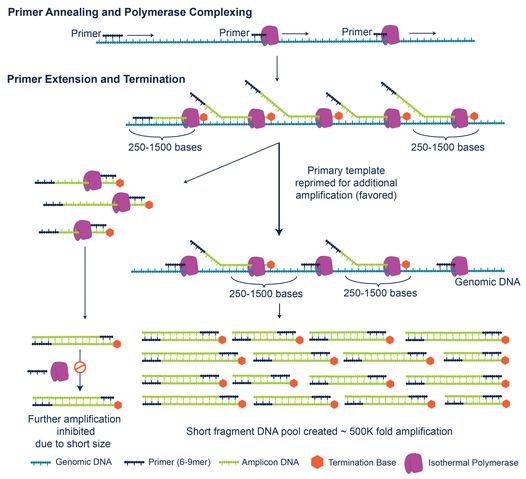
Uniform, allelically-balanced, and near-complete amplification of single-cell genomes has been enabled by ResolveDNA technology powered by primary template-directed amplification (PTA). This is achieved by attenuation of amplicon size, whereby the short amplicons have a low propensity to reamplify, manifesting with quasi-linear vs exponential amplification. Despite the power of this approach, this presents an inherent challenge for the amplification of cell-free and circulating tumor DNAs (cfDNA/ctDNA), which continue to hold promise as non-invasive diagnostic and monitoring tools for a multitude of disorders, due to their small size of only several hundred base pairs available as PTA input template. To provide a solution to the chief cf/ctDNA problem of template limitation, yet maintain the fidelity of PTA, we designed a scheme to concatenate short DNA fragments prior to amplification, increasing the size of input template to that which is amenable and optimal for PTA. This was undertaken using a size titration of synthetic DNA fragments as well as with cfDNA reference standards. Upon low-coverage sequencing, extending the template size of small fragments by ligation improved uniformity of exonic coverage as well as the number of loci detected within the cfDNA reference standard. Importantly, upon high-pass sequencing, while only 5/386 cfDNA reference standard loci were exposed with one detected single nucleotide variant (SNV) in the absence of our ligation strategy, 167 sites with 43 verified SNVs were detected post-ligation. Studies are underway to extend this proof of principle experiment to cfDNA isolated from plasma of a patient harboring a late-stage solid tumor with a defined missense mutational profile from a pan-cancer targeted panel, as well as to contrast the mutation detection sensitivity of existing methodology of cfDNA sequencing in the absence of amplification. In summary, the strategy presented here exposes single nucleotide variation from a realm of template previously not exposed with PTA and holds potential for assessment not only for non-invasive detection of tumor mutations but for a myriad of additional cfDNA applications including prenatal screening and heart disease.

Introduction

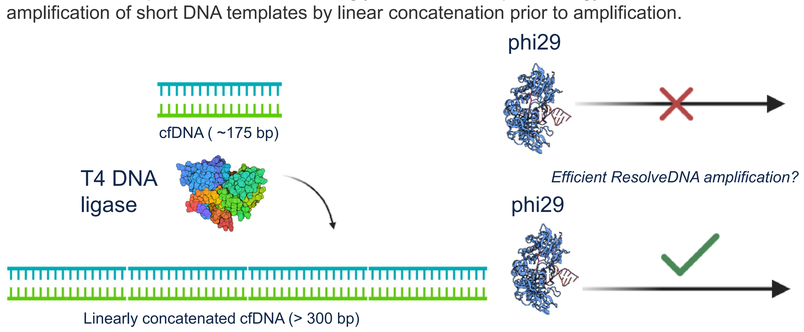
Diagnostic and prognostic applications of circulating / cell-free DNA



ResolveDNA single-cell / low-input DNA amplification powered by Primary Template-directed Amplification



Can we create a more ideal template for ResolveDNA amplification application to short DNA fragments? Near-complete, uniform, and allelically-balanced amplification from single cells is achieved by attenuating the size of amplicons with irreversible terminators. The small amplicons have a low-propensity to re-copy, favoring amplification from the primary template—yet this by definition presents an inherent challenge for ResolveDNA amplification of short DNA templates like cfDNA. In the present study we sought to define the lower limit of template size for ResolveDNA amplification, and accordingly devised a simple strategy to enable more efficient amplification of short DNA templates by linear concatenation prior to amplification.



References

1. Gonzalez-Pena, V. *et al.* Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc Natl Acad Sci U S A* 118, e2024176118 (2021).

Methods

To generate linear concatemers of short synthetic DNA fragments or cfDNA NGS standards, we optimized a workflow that begins with T4 Polynucleotide Kinase (NEB, M0201S, 10,000 U/ml) treatment for phosphorylation of 5' ends. This reaction occurs in the presence of NEB T4 Ligase Buffer, to facilitate downstream ligation in the absence of buffer exchange. A 65°C step prior to ligation ensures heat inactivation of T4 PNK. Ligation then proceeds overnight at 16 °C. Ligation products (prior to ResolveDNA amplification) were ascertained on an Agilent TapeStation D5000 tape using 2 ng/µl dilutions of the ligation reaction.

Reaction setup:

cfDNA template:	1.0µl
10X T4 DNA ligase buffer (NEB)	1.0µl
T4 Polynucleotide kinase (NEB M0201S, 10,000 U/ml)	0.5µl
nuclease free water	7.0µl
	9.5µl

The reaction was incubated at 37°C for 30 min followed by 65°C for 20 minutes, followed by 4°C cooling. 0.5 µl of T4 DNA Ligase (NEB M0202T, 2E6 U/ml) was added to the above reaction and mixed by pipetting, followed by 15h of 16°C incubation and 10 minutes 65°C incubation, with a 4°C hold. With this protocol, we have tested the following amounts of ligation reaction as input into ResolveDNA amplification: 10pg, 50 pg, 100 pg, 1 ng, 10 ng, 25 ng. SPRI cleanup of the ligation reaction is not required for input into ResolveDNA amplification.

Library preparation was performed using BioSkryb ResolveDNA Library Preparation kit and sequencing was performed on a NextSeq1000 instrument, using 2X150 cycling.

Results

Horizon OncoSpan cfDNA reference panel

cfDNA NGS control; cell line-derived
386 high-confidence variants; 46 variants with <20% AF
236 COSMIC variants; 27 indels

ABL1, AKT1, AKT2, ALK, APC, AR, ARID1A, ATR, ATRX, AXL, BARD1, BCL6, BLM, BMPR1A, BRAF, BRCA1, BRCA2, BTK, BUB1B, CARD11, CCND1, CCND3, CCNE1, CD79B, CDH2, CDK12, CDK4, CEP57, CFH, CREBBP, CSF1R, CTNNB1, DDR2, DIS3L2, DNMT3A, EGFR, EML4, EP300, EPCAM, ERBB2, ERBB3, ERCC1, ERCC2, ERCC4, ERCC5, ERG, ETS1, ETV4, EWSR1, EXT1, FANCA, FANCD2, FANCE, FANCG, FANCI, FANCM, FBXW7, FGF10, FGF2, FGF3, FGF6, FGFR1, FGFR3, FLCN, FLI1, FLT1, FLT3, FZR1, GATA2, GATA3, GEN1, GNA11, GNAS, HNF1A, HRAS, IDh2, IDh4, JAK1, JAK2, JAK3, KDR, KIT, KRAS, LDLR, MAGI1, MAP2K1, MAP2K2, MAX, MDM4, MED12, MET, MLh2, MLLT3, MMAB, MRE11, MSh4, MSH6, MTOR, NBN, NF1, NFE2L2, NOTCh1, NOTCh4, NOTCh6, NRAS, NRG1, NTRK1, NTRK3, PDGFRA, PDGFRB, PIK3CA, PIK3CD, PIK3CG, PIK3R1, PMS2, PPARG, PPP2R2A, PRKAR1A, PROC, PTCh2, PTPN11, RAD51B, RAD54L, RAF1, RB1, RBM45, RECQL4, RET, RHBDF2, ROS1, RPS6KB1, SDHB, SF3B1, SF3B2, SLTM, SLX4, SMARCB1, SMO, SMOX, STK11, TERT, TET2, TFRC, TP53, TP53BP1, TSC1, TSC2, WRN, XPA, XPC, ZNF395

In the present study we employed either 1) synthetic *FLT3* DNA fragments of varying size or 2) a cfDNA reference panel (above) as template for ResolveDNA amplification reactions.

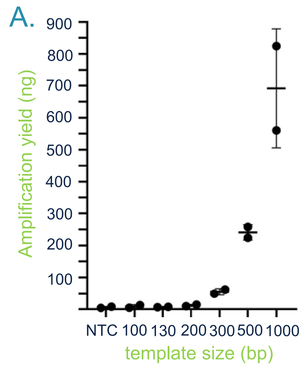


Figure 1: (A) PTA template size titration and corresponding amplification yield. 1 ng of synthetic DNA fragments (*FLT3* gene) of varying size were subjected to ResolveDNA amplification reactions. Efficient amplification begins to occur with template size ~ >400 bp. **(B) Visualization of short DNA fragment ligation products**. 20 ng of either OncoSpan cfDNA NGS control or synthetic *FLT3* DNA fragments of the indicated size were resolved by D5000 TapeStation with or without ligation treatment as indicated. For all templates analyzed, the ligation conditions yielded a broad range of higher molecular weight species relative to the input template size. **(C) Ligation product profile as a function of PNK treatment**. 25 ng of 130 bp synthetic template was utilized in ligation reactions, in the presence or absence of polynucleotide kinase treatment prior to ligation. PNK treatment yielded a broader range of higher molecular weight ligation product relative to reactions with ligase alone.

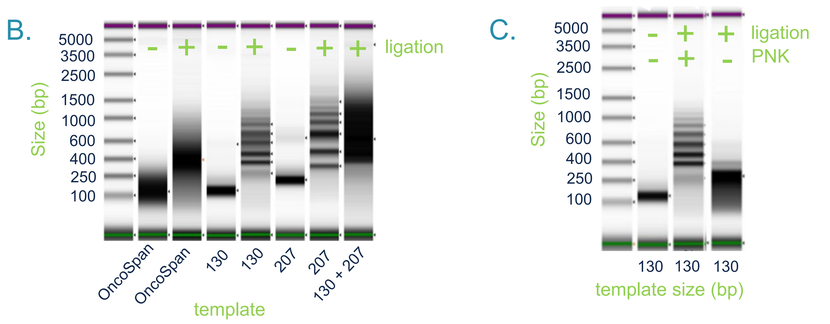


Figure 2: IGV views of cfDNA reference panel in the presence or absence of ligation treatment. Low-pass NextSeq 1000 sequencing (5M total 2X50 reads per sample) was performed on Horizon Discovery OncoSpan cfDNA NGS reference standard that was either ligated to concatenate the small fragments or unligated prior to ResolveDNA amplification. Duplicate experiments are shown to illustrate heterogeneity between reactions. A general increase in read density (including at intronic regions) for OncoSpan-represented genes is observed for the ligated vs. unligated samples. Some loci in the OncoSpan panel were poorly covered (*TP53),* while we additionally detected multiple loci that were not defined as loci represented in the reference standard (*TMEM179*).

Results

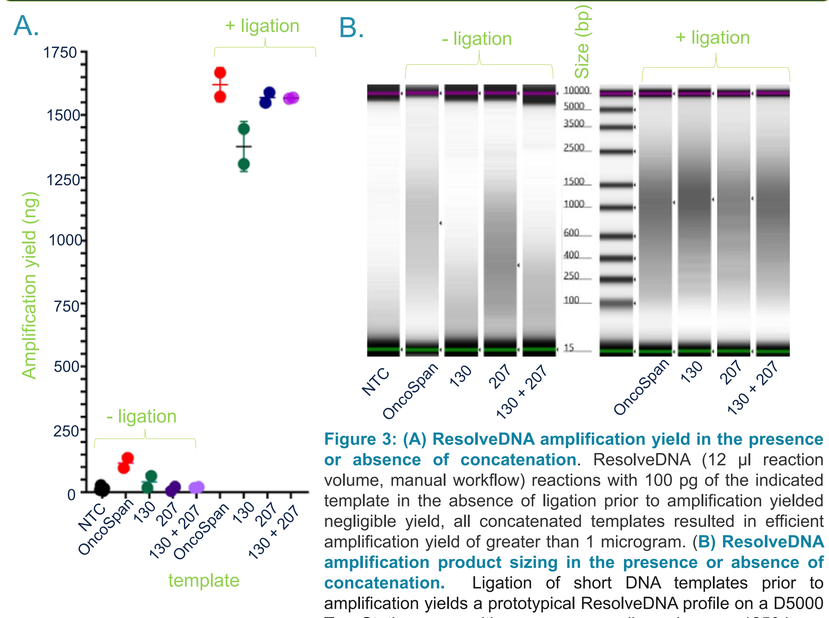
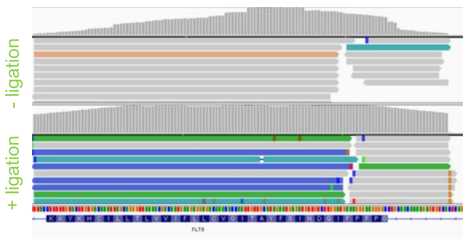


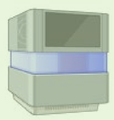
Figure 3: (A) ResolveDNA amplification yield in the presence or absence of concatenation. ResolveDNA (12 µl reaction volume, manual workflow) reactions with 100 pg of the indicated template in the absence of ligation prior to amplification yielded negligible yield, all concatenated templates resulted in efficient amplification yield of greater than 1 microgram. **(B) ResolveDNA amplification product sizing in the presence or absence of concatenation.** Ligation of short DNA templates prior to amplification yields a prototypical ResolveDNA profile on a D5000 TapeStation assay with an average amplicon size near 1250 bp.

Figure 4: Read distribution profile in the presence or absence of ligation. For a synthetic 130 bp template of the *FLT3* gene mimicking a cfDNA fragment, the 5' and 3' ends of the fragment are more highly covered by reads after amplification relative to non-ligated template. An increase in chimeric rate is noted for ligated samples; but this does not preclude accurate SNV calling (below).



Single nucleotide variant calling ability is enhanced with template concentration

A.



	- Ligation	+ Ligation
OncoSpan loci detected	5/386	167/386
OncoSpan SNVs detected	1	43

Figure 5: (A) Number of variants detected with or without ligation prior to ResolveDNA in cfDNA reference panel. 100 pg of ligated or unligated cfDNA reference standard was subjected to ResolveDNA amplification and subsequent NextSeq 1000 2X150 sequencing, downsampling to 100M total paired-end reads. Ligation prior to amplification resulted in more 1X coverage of OncoSpan loci and in the number of OncoSpan validated single nucleotide variants detected. **(B) IGV views of OncoSpan variants.** An OncoSpan validated *MSH3* missense mutation is detected in both ligated and unligated ResolveDNA templates, while a validated *NOTCH2* variant was only detected in the ligated sample. Allelic frequency of the panel variants is indicated in the table.

B.

Gene	Variant Type	COSMIC ID	dbSNP ID	dbSNP Amino Acid (v151)	NGS allele frequency
MSH3	SNP	COSM4416274; COSM4416275	rs26279	A1045T	100.0%
NOTCH2	SNP	NA	NA	NA	42.50%

Benefits of ResolveDNA amplification of cfDNA over standard practice? Current commercial solutions exist to purify circulating DNA, and to subsequently directly ligate the purified nucleic acid with sequencing adapters. These library preparation approaches, in the absence of PCR amplification, require 10-100 ng/µl cfDNA concentration while library preps utilizing PCR amplification require 1 – 10 ng of cfDNA input. We have shown the ability to input as low as 10 pg of template DNA, and desire to leverage this inherent sensitivity of ResolveDNA for cfDNA applications. Studies currently underway include assaying plasma-derived patient cfDNA with a known primary tumor mutation, and comparing the read depth required for ResolveDNA to detect the mutation relative to direct-ligation approaches with no amplification.

Conclusions

- Defining the lower-bound of ResolveDNA template size by titration necessitated the implementation of new strategy for optimal amplification of small DNA fragments / cell-free DNA
- Concatenation of short DNA fragments as cfDNA mimics resulted in efficient ResolveDNA amplification and the ability to better expose single nucleotide variation in an cfDNA reference panel with oncogenic variants