Novel Cutting Edge Applications of Ion AmpliSeq™ RNA Technology

Kelli Bramlett, Jeoffrey Schageman, Efren Ballesteros-Villagrana, Tom Bittick, Joel Brockman, Angie Cheng, Nathan Harris, Natalie Hernandez, Brian Sanderson, Daniel Williams, Susan Magdaleno, and Bob Setterquist, Life Technologies, 2130 Woodward Street, Austin, Texas, USA, 78744

ABSTRACT

Life Technologies recently released the first Ion AmpliSeq™ offering for RNA which included a library construction kit as well as an integrated design system on Ampliseq.com. This platform allows researchers to simultaneously interrogate gene expression of up to 1200 targets in a single tube, thus combining the sensitivity of qPCR with sequence information gained from RNA sequencing. Since this launch, we have developed a program referred to as white glove to push the limits of what AmpliSeq™ RNA currently offers. Here, we describe several examples of how these white glove initiatives have resulted in technological advances which will expand the repertoire of targeted RNA sequencing applications.

We adapted our Ion RNA AmpliSeq™ primer designer to enable design to multiple RNA transcript configurations addressing several biological phenomena. In particular, we have successfully designed and tested panels to investigate gene fusion detection in cancer, characterize B-cell receptor IgH stereotypes/mutational burden, and interrogate both cancer fusions as well as hotspot mutations in a single panel. In addition, we have developed software to design AmpliSeq™ RNA fusion detection assays given COSMIC annotation. To facilitate easy reuse of AmpliSeq™ RNA fusion primer designs, we created a database to store designs as they are created. Currently, this database contains over 400 fusion detection designs.

INTRODUCTION

RNA-seq experiments have traditionally started with a whole-transcriptome library preparation that produces a sequencing template from all RNA species in a sample. However, in many cases, only a handful of the genes present are necessary to make a statistically significant conclusion confirming a hypothesis, validating expression data derived from an orthologous platform, or informing a clinical treatment decision.

We have demonstrated new technology that allows for targeted RNA-seq from a panel of directed amplicons using Ion AmpliSeq™ technology with Ion Torrent™ semiconductor sequencing. This approach offers many advantages over microarray or qPCR data such as faster turnaround and data analysis, sample multiplexing, lower RNA inputs, and ability to use degraded or FFPE-derived samples. In addition, the technique simultaneously provides digital, quantitative gene expression information and gene sequence at the single nucleotide level. This capability enables not only gene expression measurements, but also identification of fusion genes as well as gaining nucleotide level information along the designed amplicon enabling single nucleotide variant identification. Here, we describe unique applications of this targeted RNA sequencing approach taking advantage of some of these capabilities by sequencing RNA in a custom designed and targeted fashion.

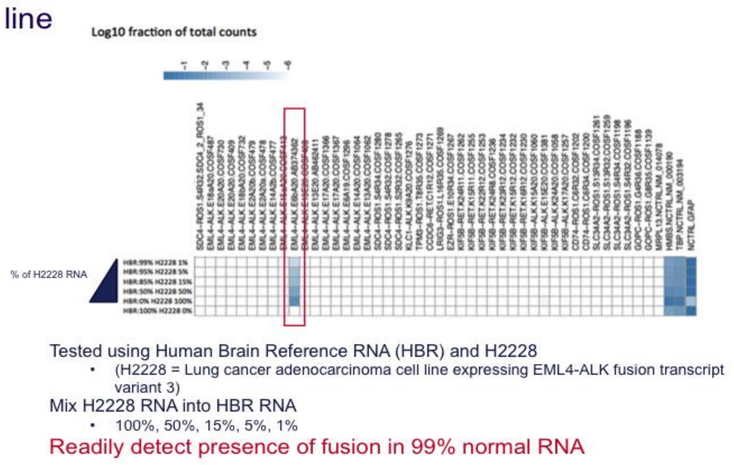
MATERIALS AND METHODS

The White Glove RNA AmpliSeq(™) process requires an application be completed describing the detailed technical needs of this targeted RNA sequencing approach. We then custom design AmpliSeq RNA primers for each customer to meet the specific experimental aims. Additionally, we provide analysis recommendations and the appropriate reference files needed to calculate the number of sequencing reads that align to each custom designed oligo.

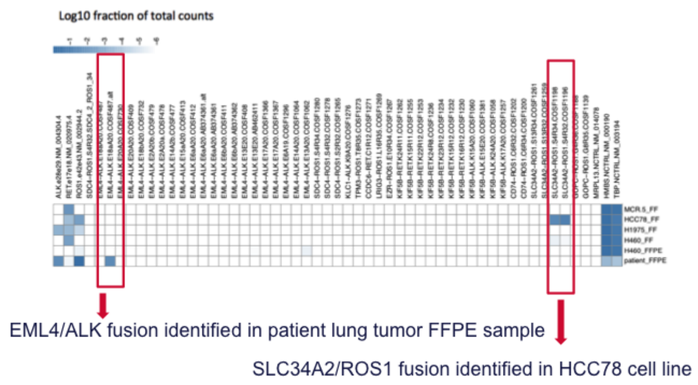
AmpliSeq™ RNA Custom Design Project #1

Challenge:
Create a custom RNA AmpliSeq design for a complete analysis of RNA fusion genes implicated in lung cancer.

Historically, the fluorescent in situ hybridization (FISH) technique has been widely employed to study genomic variation in malignant cells. Recently the field of molecular cytogenetics is investigating next generation sequencing technologies to get even more complete information about the genomic variation in malignancies (1). Recently a gene fusion between echinoderm microtubule-associated protein-like 4 (EML4) and anaplastic lymphoma kinase (ALK) was identified in a NSCLC patient with a smoking history leading to a phase I clinical trial with an inhibitor of ALK (2,3). Subsequently, additional genes have been identified to be involved in lung cancer and targeted for potential for therapeutic intervention, such as RET and ROS1 (4). The FISH technique is labor intensive and experimentally demanding limiting the number of individual fusion genes that can be reasonably assayed for in any given sample. Due to the inherent limitations of FISH and the increased identification of chimeric fusion kinases with potential for oncogenic transformation the desire was to find a technology that would allow for testing for multiple fusion genes with limited tissue. With the new advancements in targeted RNA sequencing, we attempted to design a single panel targeting over 40 fusion genes involving ALK, RET and ROS1. Testing tissue samples for these fusion genes can be accomplished in one amplification reaction from as little as 5 ng of FFPE RNA.



To determine the sensitivity of detection of this custom designed AmpliSeq™ RNA panel, we used a well characterized cell line H2228 containing a known EML4-ALK. We successfully detected this EML4-ALK variant when only 1% of the total RNA used in the library procedure was from the fusion gene containing H2228 cell line in the background of the Human Brain Reference (HBR) RNA.

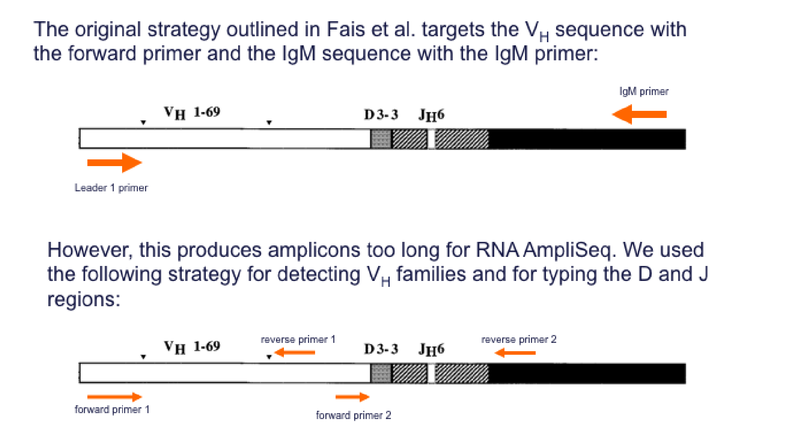


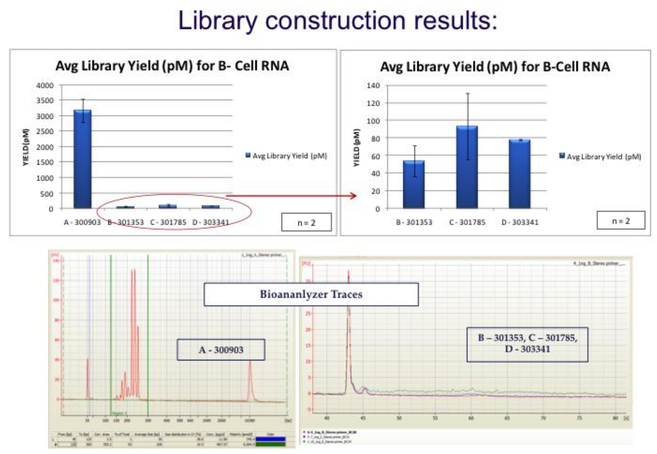
After determining the sensitivity of the AmpliSeq™ RNA fusion panel sensitivity to be ≤ 1% fusion containing RNA, we expanded the testing to include other cell line RNA and human tissue RNA harboring previously characterized fusion genes. As shown in the graph above we were able to successfully detect a second EML4-ALK fusion in RNA isolated from human tissue that was formalin fixed paraffin embedded (FFPE). Additionally, we successfully amplified a fusion gene from different fusion partners, SLC34A2 and ROS1 from HCC78 cell line total RNA.

AmpliSeq™ RNA Custom Design Project #2

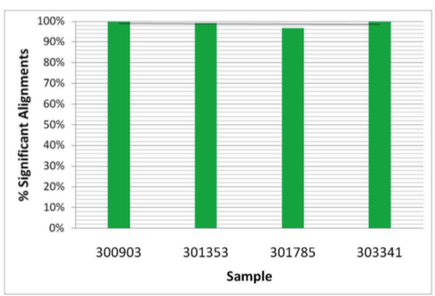
Challenge:
Create a custom RNA AmpliSeq design to determine the dominant B cell Receptor Prototype as well as mutational burden of the variable region (V_H) of the receptor

A seminal publication from 1998 by Fais et al. describes the analysis of rearranged V_H genes expressed by B-type chronic lymphocytic leukemia (B-CLL)cells and the mutations within the V_H family (5). This publication describes primers to amplify long amplicons to determine the correct V_H family assignment, D and J_H sequence identification, and analysis of HCDR3 rearrangements by cDNA sequencing from 1 µg of RNA. We attempted to create custom designed AmpliSeq™ RNA primers to amplify the V_H, D and J_H regions of the B cell receptor in one multiplexed amplification reaction requiring only 5 ng of total RNA.





Using the design strategy highlighted above, we created a multiplexed AmpliSeq™ RNA primer pool that would provide receptor typing information as well as mutational analysis capabilities of the V_H region of the B cell receptor. We received 4 B-CLL RNA samples from our customer and successfully created targeted RNA sequence libraries as shown by the Agilent2100 Bioanalyzer traces above. Although the yield was low for sample B, C, and D, we were able to get viable libraries in sufficient quantity for template prep and sequencing on the Ion Torrent PGM 318™ Chip. We also identified a unique data analysis tool that proved incredibly helpful in evaluating the resulting amplicon information. This analysis tool, IgBLAST (http://www.ncbi.nlm.nih.gov/igblast/) was developed specifically to map sequencing reads to the V_H, D, and J domains of the B cell receptor and also display the variants identified in reads mapping to the VH region (6). Using this tool on 1000 reads of each of the 4 RNA libraries we were able to see significant alignments to the reference available with the online IgBLAST tool. For sample A-300903, we also identified the predominant B Cell Receptor form.



V, D, and J regions summary - 300903 (BC25)

IGHV	count	IGHD	count	IGHJ	count
IGHV3-13*01	16	IGHD3-10*01	453	IGHJ4*01	19
IGHV3-13*05	16			IGHJ4*02	20
IGHV3-48*03	249			IGHJ4*03	20
IGHV3-74*01	414			IGHJ5*02	1
IGHV3-74*03	424			IGHJ6*01	31
IGHV3-h*03	12			IGHJ6*02	403
IGHV6-1*02	11			IGHJ6*03	15
				IGHJ6*04	37

Predominant stereotype		
IGHV	IGHD	IGHJ
IGHV2-74*03	IGHD3-10*01	IGHJ6*02

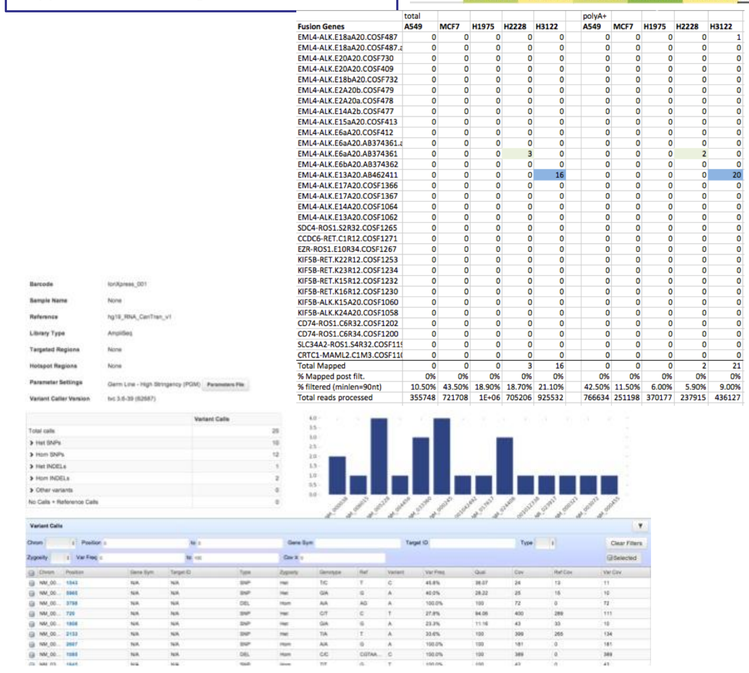
chain type	stop codon	V-J frame	Productive	Strand
VH	NO - 408	In-frame - 294	N0 - 193	neg - 198
	YES - 54	Out-of-frame - 164	YES - 265	pos - 264

AmpliSeq™ RNA Custom Design Project #3

Challenge:
Design a AmpliSeq™ RNA panel that will detect specific fusion genes in addition to evaluating single nucleotide variants in well studies genes shown to be involved in oncogenic transformation.

The challenge in this project was getting as much molecular cytogenetic information as possible from very limited amounts of nucleic acid. The sample type used for targeted RNA sequencing was from a very limited number of precious cells isolated from blood. In this case, the goal was to create one pool that would amplify over 20 fusion genes from the ALK, RET and ROS1 families as well as provide the best coverage of the region of 36 oncogenes and tumor supressor genes for hot spot mutational analysis. The 36 genes targeted for hot spot mutation analysis were chosen based on the content of the DNA AmpliSeq™ Cancer Hotspot Panel. Because so few cells are recovered from blood following this unique sample enrichment method, there is not sufficient material for a DNA and a separate RNA AmpliSeq™ library preparation. So, the goal of this custom design was to use targeted RNA designs to identify fusion genes as well as obtain information on hotspot mutations and gene expression information on the 36 targeted tumor suppressor and oncogenes.

	totalRNA	reads per gene, normalized				
BC/LC	A549	MCF7	H1975	H2228	H3122	
	BC6	BC2	BC3	BC4	BC5	HotSpot Genes
	0.0123084	0.01495079	0.01151407	0.02016814	0.01236835	APC
	0.03483165	0.01763771	0.04039996	0.06421415	0.02625685	ATM
	0	0	0.00118068	0	1.1137E-06	CDKN2A
	0.00047928	0.00016702	0.00076397	0.00072335	0.00025057	JAK3
MET	0.07286977	0.0080251	0.07251336	0.13415846	0.08817879	MET
	0.06894732	0.1012713	0.09387467	0.07172364	0.05003024	PTEN
	0.0190345	0.02855405	0.00431096	0.00299799	0.0301157	RB1
	0.00244345	0.00375067	0.00418892	0.00393922	0.0020703	STK11
	0.03692225	0.05179101	0.0661893	0.00332625	0.0242478	TP53
	0.01074265	5.9998E-05	0	1.162E-05	4.4547E-06	NTRK3
	0.14164802	0.08554379	0.08537976	0.14022706	0.02847639	NF1
	0.19125798	0.14969596	0.13812126	0.05088734	0.25572952	CTNNB1
FLT1	0.00114527	0.0018502	0.00011807	0.00038492	0.00072277	FLT1
	0.00038225	0.0004962	0.00196548	0.00160212	0.00151124	KDR
	1.4702E-05	8.1078E-06	1.4882E-05	5.8101E-06	0	NTRK1
	0.00300359	0.03758128	0.04975409	0.01589485	0.03098324	SMARCA4
	0.00021318	0.00058538	3.671E-05	0.00010313	9.6889E-05	FLT3
	0.00202151	0.00128914	0.00153587	0.00613251	0.02355065	ALK
	0.01021779	0.0134249	0.01023418	0.00929318	0.01074908	BRAF
	0.00046752	0.00024496	7.1436E-05	0.00013073	0.00012028	CTNNA2
	0.00614538	0.00573546	0.00290605	0.00537285	0.00376084	EPHA5
EZH2	0.03112384	0.06265385	0.02304005	0.02335205	0.02129214	EZH2
	0.00032638	0.00059998	0.00387341	0.00030503	0.00026282	CSF1R
EGFR	0.06717574	0.00224424	0.07808139	0.11314059	0.03279629	EGFR
	0.00161867	0.00297556	0.00089394	0.00092235	0.0016382	EPHA3
	0.00506185	0.01345733	0.00141383	0.00345117	0.00203689	ERBB4
	0.02405077	0.01595615	0.01932936	0.01892043	0.01662589	SMAD4
	0.10938773	0.13529002	0.08177324	0.08305615	0.18880723	ARID1A
	0.0054838	0.0004589	0.00011906	0.00025709	0.00015146	PDGFRA
	0.0021597	0.00496197	0.00438041	0.0095343	0.00402367	PIK3CA
	0.00011467	0.00015243	5.7546E-05	5.5195E-05	5.5683E-05	CTNNA3
	0.01295234	0.02007816	0.02266402	0.00971005	0.01398762	NOTCH1
RET	0.00024552	0.0251877	0.00013692	0.00107341	0.0001058	RET
	0.09394195	0.16432081	0.16246898	0.17706569	0.09601119	NOTCH2
	0.03210592	0.02151648	0.01202206	0.02488881	0.02741618	KRAS
	0.00409006	0.0074835	0.0046721	0.00297039	0.00556386	FBXW7



CONCLUSIONS –

The AmpliSeq™ RNA custom design “white glove” process has been utilized for unique applications of targeted RNA sequencing technology. Being able to get quantitative gene expression information in addition to detect fusion transcripts and single nucleotide variants within the targeted amplicon is ideal for some of the unique applications described here.

- White Glove custom design capabilities are available to meet individual experimental goals with AmpliSeq™ RNA technology
- This process provides custom primer design as well as specific reference and analysis files with guidance on recommended analysis methods
- Through white glove design species other than human can also be amplified in a targeted fashion
- As little as 5 ng of total RNA is required for AmpliSeq™ RNA sequencing of up to 1200 targets
- This technology is ideal for fusion gene detection alone or in combination with gene expression measurements and mutation analysis

REFERENCES

1. Das K, Tan P. Molecular cytogenetics: recent developments and applications in cancer. *Clin Genet 2013: 84: 315-325.*
2. Stephens PJ, Tarpey PS, Davies H et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature 2007: 448: 561-566.*
3. Kimura H, Nakajima T, Takeuchi K et al. ALK fusion gene positive lung cancer and 3 cases treated with an inhibitor for ALK kinase activity. *Lung Cancer 2012: 75: 66-72.*
4. Gainor J, Shaw A. Novel Targets in Non-Small Cell Lung Cancer: ROS1 and RET Fusions. *The Oncologist 2013: 18: 865 – 875.*
5. Fais F, Ghiotto F, Hashimoto S et al. Chronic Lymphocytic Leukemia B Cells Express Restricted Sets of Mutated and Unmutated Antigen Receptors. *J Clinical Invest 1998: 102: 1515-1525.*
6. Ye J, Ma N, Madden TL, and Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res 2013: 41: 34-40.*

ACKNOWLEDGMENTS

We would like to thank Tom Bittick, the product manager for AmpliSeq™ RNA products and the White Glove Design team for AmpliSeq™ DNA for valuable insight and discussion. We also appreciate the

TRADEMARKS/LICENSING

For Research Use Only. Not intended for any animal or human therapeutic or diagnostic use.