PARTNERS HEALTHCARE | FOUNDED BY BRIGHAM AND WOMEN'S HOSPITAL AND MASSACHUSETTS GENERAL HOSPITAL

INVENTION DISCLOSURE FORM

For Internal Use Only • Invention #:

CONTACT/SUBMISSION INFO

Please address all questions and return both a signed original and electronic copy of the completed form to the **Senior Invention Administrator - Ms. Michele O'Shea** *E-mail:* PHSPatents@partners.org; *Ph#:* (617) 954-9345 *Fax:* (617) 954-9361

(While filing form on-line, please have 'insert' key inactive)

1. TITLE OF INVENTION

A next-generation sequencing assay for targeted sequencing

2. CATEGORY OF INVENTION

☒ Patent ☐ Material ☐ Software ☐ Copyright ☐ Trademark

3. DESCRIPTION OF THE INVENTION

Describe the Invention to the extent known at this time

Key concepts of **Invention**, including nature, stage, purpose and operation of the invention including **technical characteristics**:
(In addition, please attach manuscript, presentation, poster or other documents, including any public disclosure documents)

The invention describes a method for the efficient and simple production of assays that can analyze the sequence of targeted gene sequences using "next generation" (aka "Deep"; aka "massively parallel") sequencing.

The assay involves the use of primer extension with reverse transcriptase with two nested primers hybridizing to the region of interest, using double stranded cDNA as a template. The cDNA will have linkers ligated first (with a univeral primer 1 for donwstream amplification), and the extension/RT primer will be tailed with a universal primer 2. See attached summary for details. The assay will also work with double stranded genomic DNA as a template.

Distinguishing novel features of Invention:

The nested primer extension has not been used in next generation sequencing, and allows not only for the enrichment of certain targets for downstream sequencing, but allows one to map gene rearrangements only knowing one of the partners.

The assay approach is very simple and very inexpensive, and could replace many/most targeted library construction methods.

3. Description of the invention

Key concepts

In accordance with the invention, an assay using a novel half-truncated "Y" shape adapter for next generation sequencing library construction followed by two single-end nested polymerase chain reactions enables rapid and efficient target enrichment using RNA or genomic DNA as template from fresh or formalin-fixed and paraffin-embedded (FFPE) specimens. This enrichment method enables targeted resequencing of cDNA or gDNA for potential detection of genetic alterations (single nucleotide variants, insertions/deletions, and copy number), epigenetic alterations (methylation), gene expression, and genomic rearrangements.

Target enrichment prior to next-generation sequencing is more cost-effective than whole genome, whole exome, and whole transcriptome sequencing and therefore more practical for broad implementation both for research discovery and clinical applications. High coverage depth afforded by target enrichment approaches enables wider dynamic range for allele counting (in gene expression and copy number assessment) and detection of low frequency mutations, a critical feature for evaluating somatic mutations in cancer. Before broad implementation of whole genome or whole exome sequencing is possible, the mainstay of clinical next generation sequencing will involve select disease panels with a discrete number of gene targets. Likewise, research studies for analysis of large sample sizes based on defined gene targets will also require an economical means for genotyping. Our assay would be applicable in both of these settings.

For the detection of inter and intra-chromosomal rearrangements, whole genome or transcriptome sequencing is useful for novel discovery of genomic rearrangements and does not require prior knowledge of the involved gene/chromosomal partners. However, these whole genome and transcriptome approaches are not practical currently in the clinical setting due to high cost, low sequencing depth resulting in poor sensitivity, and highly demanding bioinformatics analysis. Fluorescence in situ hybridization (FISH) has been the gold standard assay for the detection of genomic rearrangements in clinics; however, the assay is low-throughput and its implementation requires special equipment and expert knowledge/experience. RT-PCR is also effective at detecting such rearrangemens but requires knowledge of both the 5' and 3' partners and is not scalable for a large number of targets/samples.

Examples of commonly used preparatory enrichment assays for next generation sequencing include hybridization-based capture assays (TruSeq Capture, Illumina; SureSelect Hybrid Capture, Agilent) and polymerase chain reaction (PCR)-based assays (HaloPlex, Agilent; AmpliSeq, Ion Torrent; TruSeq Amplicon, Illumina; emulsion/digital PCR, Raindance). Hybridization-based approaches capture not only the targeted sequences covered by the capture probes but also near off-target bases that consume additional sequencing capacity. In addition, these methods are relatively time-consuming, labor-intensive, and less specific. A PCR amplification based approach is simpler and faster but by conventional design requires both forward and reverse primers flanking target loci. In particular for detection of genomic rearrangements with unknown fusion partners, PCR is not applicable.

We disclose a target enrichment assay using a novel half-truncated "Y" shape adapter for next generation sequencing library construction, enabling rapid sequencing of cDNA or gDNA from fresh or FFPE specimens. Figure 1 outlines the library construction for target enrichment using a half-truncated Y

adapter for Ion Torrent sequencing as an example. Importantly, this method is adaptable for any other next generation sequencing platform, including but not limited to Illumina, SOLiD, and 454. In short a randomly sheared, double-stranded cDNA or gDNA template is end-repaired, adenylated, and then ligated on one or both ends with the universal Y adapter to create a universal sequence for initiating PCR and subsequent sequencing. An initial round of PCR using a target specific primer tagged with a stuffer tail (the stuff tail aids in multiplexing a high number of targets) and a primer same as the 20 bases at 5' Y adapter overhang. A second round of PCR is carried out with a primer same as the 30 bases at the 5' Y adapter overhang and a second tandem nested target specific primer that anneals 3' downstream of the first target specific primer. The second tandem nested primer is 5' tagged with the second primer sequence required for downstream emulsion PCR or clustering depending on the sequencing technology. High specificity in this system is achieved with the unidirectional tandem, nested primers which will effectively cover a target sequence of 36-40 bases or more (depending on the spacing between the primers). Of note, the number of PCR cycles may be optimized depending on how much starting nucleic acid material is used, the number of pooled samples, and the number of targets.

In summary, the method utilizes a half-truncated Y adapter to universally tag all dsDNA fragments with a common 3' end which is utilized for two rounds of PCR using two unidirectional, nested target specific primers for specificity. The application of tagged nested primers also avoids the effects of primer homodimerization and heterodimerization when targeting many different sites in the genome or transcriptome. In addition to multiplexing targets, our method also allows for multiplex sample pooling after the Y adapter ligation (Step 2) during which individual samples are ligated with unique barcoded adapters. Once barcoded via Y adapter ligation, multiple samples may be pooled into one reaction tube for downstream target enrichment.

In contrast to detection of genomic rearrangements which is more tolerant insertion and deletion sequencing errors such as those found on Ion Torrent and 454 sequencing as a result of homopolymers, detection of single base and multiple base mutations (including insertions and deletions) requires higher accuracy sequencing achievable with the Illumina sequencing platform. For this purpose, the half-truncated Y adapted library may be converted to an Illumina library by amplifying the overhanging, truncated 3' arm of the Y adapter using a primer tagged with the Illumina forward adapter sequence (barcoded or non-barcoded), and tagged gene specific nested primers (GSP2s tagged with Illumina reversed primer). Similarly, Bi-directional sequencing is also achievable by introducing a sequencing primer via the overhanging, truncated 3' Y adapter arm using a tagged primer during the first PCR, Step 3.

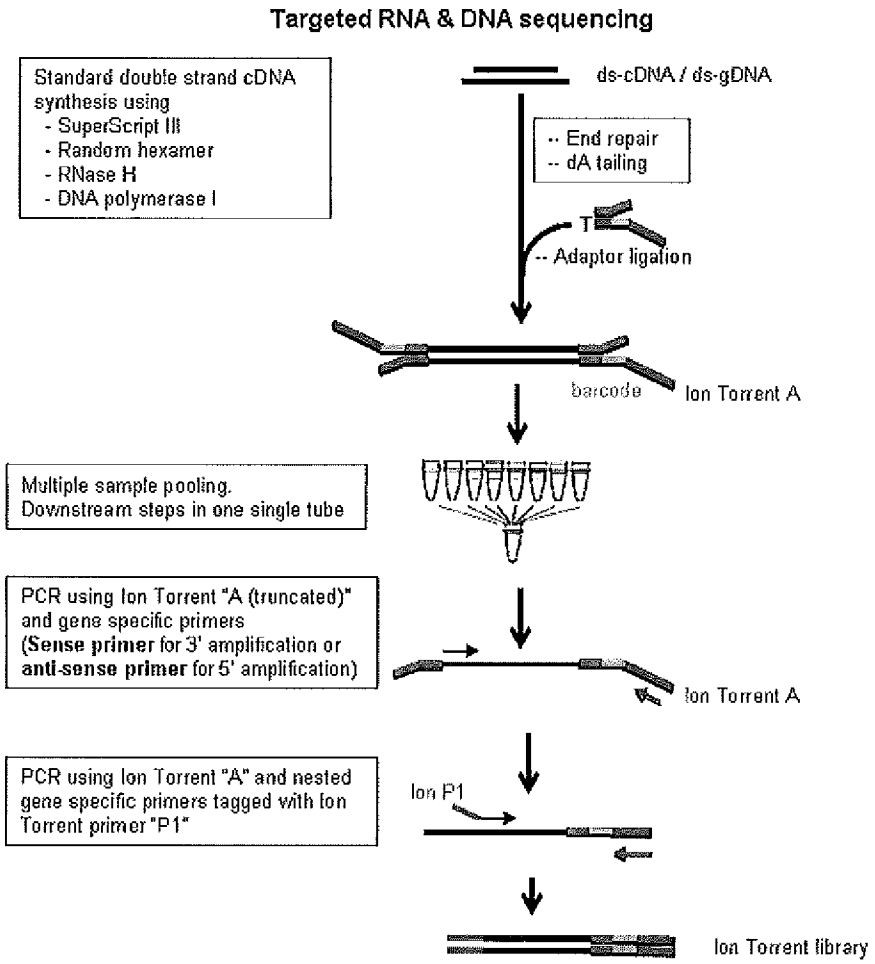


Figure 1 Schematic illustration of library construction for targeted RNA and DNA sequencing. 1) A standard procedure of double-stranded cDNA synthesis is applied using total nucleic acid from FFPE specimen as starting material. Alternatively, the starting material may also be sheared gDNA. 2) After cleanup, the double stranded cDNA or gDNA is subjected to end-repair and dA tailing, directly followed by ligation of a half-truncated Y adapter without cleanup in between. 3) After SPRI cleanup, the ligated sample is subjected to 14 cycles of PCR amplification using multiplex gene specific primers (GSP1s) and Ion Torrent short length Forward primer A 5' 20-mer (A20), with annealing temperature at 65°C. 4). After a second SPRI cleanup, the sample is subjected to an additional 14 cycles of PCR amplification using multiplex nested gene specific primers (3' downstream of GSP1 and in the same direction) tagged with Ion Torrent Reverse primer (P1_GSP2s) and Forward primer A (full length 30-mer), with annealing temperature at 65°C. 5) After a final third SPRI cleanup, the product is Ion Torrent library ready for downstream emulsion PCR and sequencing. Sample pooling may be carried out after Step 2.

We have applied this assay on cDNA derived from FFPE specimens for the detection of *ROS1* gene fusions using seven gene specific primers targeting seven exons comprising the *ROS1* kinase domain. We detected *ROS1* genomic rearrangements with known and previously unknown fusion partner such as "*SLC34A2* Exon 12 – *ROS1* Exon 34" and "*EZR* Exon 9 – *ROS1* Exon 34" fusions. The assay achieved high on-target specificity (~85-95% when mapping using human genome hg19 reference), which enables the sequencing of multiple samples even with the smallest scale sequencing platforms (and thus least

expensive) such as the Ion Torrent PGM 314 chip with high coverage results (7 targeted loci, 5 samples, >1000 X coverage for each target per sample). Figure 2 shows mapping of sequencing reads to the target loci in gene *ROS1* kinase domain.

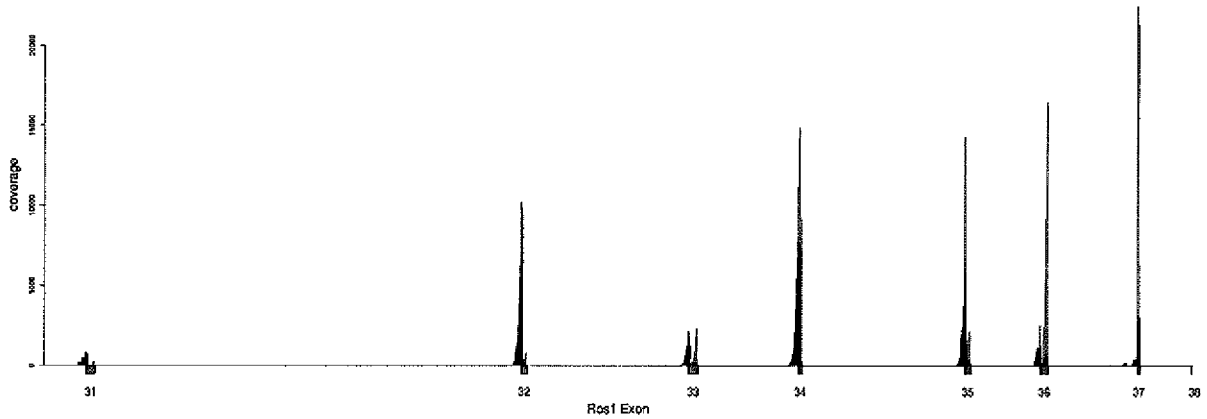


Figure 2. Mapping results showed different primer (green dots) extension products in one sequencing run, including reads that corresponding to genomic DNA (black, spanning intron-exon boundaries), cDNA (red, spanning exons) and fusion cDNA (on exon 34, mapping on the fusion partner SLC34A2 not shown here).

Distinguishing novel features

1. Our method may be utilized for target enrichment of double stranded cDNA or gDNA from fresh or FFPE samples, allowing mapping of both 5' and 3' ends of target cDNA or gDNA fragments, using next-generation sequencing. Current hybridization-based approaches require days of preparation and hybridization, and show lower specificity with capture of near off-target sequences. Current amplification based approaches require known forward and reverse primers by design and are not amenable for high scale multiplexing of targets.

2. Simple Bioinformatics analysis. Depending on the number of selected targets, analysis of data generated from this assay is simple and fast. In addition, it may be set up to be compatible with current existing bioinformatics tools. Thus, small-size clinical labs may be able to perform data analysis without the need for significant investment in bioinformatics, which will be a limitation for broad clinical implementation of next generation sequencing.

3. High specificity (~85-95%). A conventional Y adapter construct having both components of forward and reverse primers (5' and 3' overhangs) for routine library construction will introduce a high level of background off-target sequencing due to carry over of "sequenceable" starting genomic or transcriptome library material. For instance, if the target is of size of 100 bp and present in two copies in one genome, the ratio of target-to-genome is about 1:3x10^7. Hybridization-based approaches normally employ biotinylated oligonucleotide baits to pull out hybridized target fragments by streptavidin coated magnetic beads. Just one non-specific binding event out of the possible 3x10^7 possibilities could dramatically result in a 50% off-target rate. The use of a half-truncated Y adapter effectively avoids this background

carry over issue, since the starting library material cannot be amplified in subsequent preparatory steps for sequencing. Additional specificity is achieved by using two unidirectional primers, with GSP2 3' downstream of GSP1. Effectively, the use of the two tandem primers yields a target priming site of 40 (assume two 20 base pair length primers) and higher specificity than with just one primer. The use of 5' 20-mer primer in the first PCR and the full 30-mer primer (serves as a nested primer for the universal priming site) in the second PCR further increases specificity. Finally, additional specificity is achieved with the use of tagged primers for both PCR steps. Tagged primers, through intramolecular hairpin formation, prevent the propagation of primer homodimers and heterodimers which could overtake the PCR reactions and lead to nonspecific and undesired artifactual products; therefore, tagged primers allow the ability to multiplex many targets while avoiding primer dimers.

4. Economical cost. The key components in our method are conventional, unmodified tagged primers, standard PCR reagents, and routine thermocycling. Unlike hybridization based capture methods or microfluidic digital PCR setups, our target enrichment protocol avoids the use of relatively expensive biotinylated oligonucleotides, streptavidin coated magnetic beads, and special equipment. Once manufactured, pooled GSP1 and GSP2 primer mixes may be used for thousands of reactions.

5. Automation. Because our method relies on standard PCR techniques and SPRI cleanup, there is potential for facile automation in high throughput applications. Volumes may be adjusted for 384 well plates for ultra-highthroughput implementation.

Envisioned commercial products and applications:

1. A lung cancer translocation panel consisting of known therapeutic targets including genes *ALK, ROS1* and *RET*.

2. Hematological malignancy panels including those to detect genomic rearrangements in lymphomas and leukemias.

3. Sarcoma genomic rearrangement panel.

4. IGH/TCR gene rearrangement panel for lymphoma testing.

5. Targeted disease gene panel for resequencing (10-100 genes).

6. Targeted resequencing for confirming variants from large scale sequencing projects.

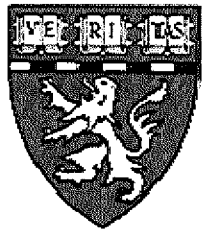
7. Potential for whole exome targeted resequencing.

8. Targeted resequencing for detection of single nucleotide variants, multiple nucleotide variants, insertions, deletions, copy number changes, and methylation status.

9. Microbiota sequencing.

10. Ancient sample sequencing.

11. New variant virus genotyping.



Next-Generation Sequencing Assay Development for Detection of Gene Fusions in Lung Cancer

Zongli Zheng, PhD.
Department of Pathology
Massachusetts General Hospital

MASSACHUSETTS
GENERAL HOSPITAL
PATHOLOGY

IDT Exhibit 1048.0007

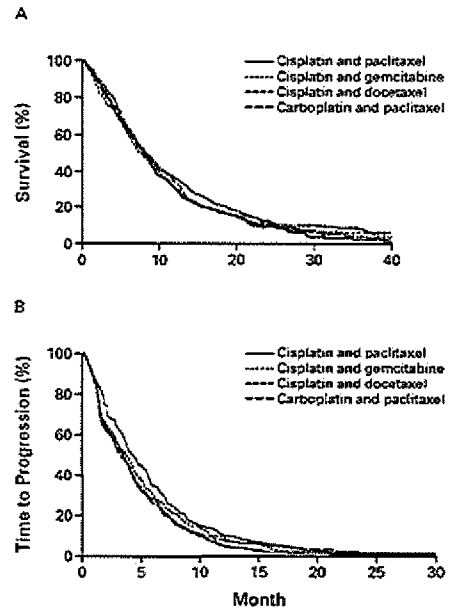
Lung Cancer: Prevalence and Survival

Leading New Cancer Cases and Deaths – 2012 Estimates

Estimated New Cases*		Estimated Deaths	
Male	**Female**	**Male**	**Female**
Prostate 241,740 (29%)	Breast 226,870 (29%)	Lung & bronchus 87,750 (29%)	Lung & bronchus 72,590 (26%)
Lung & bronchus 116,470 (14%)	Lung & bronchus 109,690 (14%)	Prostate 28,170 (9%)	Breast 39,510 (14%)
Colon & rectum 73,420 (9%)	Colon & rectum 70,040 (9%)	Colon & rectum 26,470 (9%)	Colon & rectum 25,220 (9%)
Urinary bladder 55,600 (7%)	Uterine corpus 47,130 (6%)	Pancreas 18,850 (6%)	Pancreas 18,540 (7%)
Melanoma of the skin 44,250 (5%)	Thyroid 43,210 (5%)	Liver & intrahepatic bile duct 13,980 (5%)	Ovary 15,500 (6%)
Kidney & renal pelvis 40,250 (5%)	Melanoma of the skin 32,000 (4%)	Leukemia 13,500 (4%)	Leukemia 10,040 (4%)
Non-Hodgkin lymphoma 38,160 (4%)	Non-Hodgkin lymphoma 31,970 (4%)	Esophagus 12,040 (4%)	Non-Hodgkin lymphoma 8,620 (3%)
Oral cavity & pharynx 28,540 (3%)	Kidney & renal pelvis 24,520 (3%)	Urinary bladder 10,510 (3%)	Uterine corpus 8,010 (3%)
Leukemia 26,830 (3%)	Ovary 22,280 (3%)	Non-Hodgkin lymphoma 10,320 (3%)	Liver & intrahepatic bile duct 6,570 (2%)
Pancreas 22,090 (3%)	Pancreas 21,830 (3%)	Kidney & renal pelvis 8,650 (3%)	Brain & other nervous system 5,980 (2%)
All sites 848,170 (100%)	All sites 790,740 (100%)	All sites 301,820 (100%)	All sites 275,370 (100%)

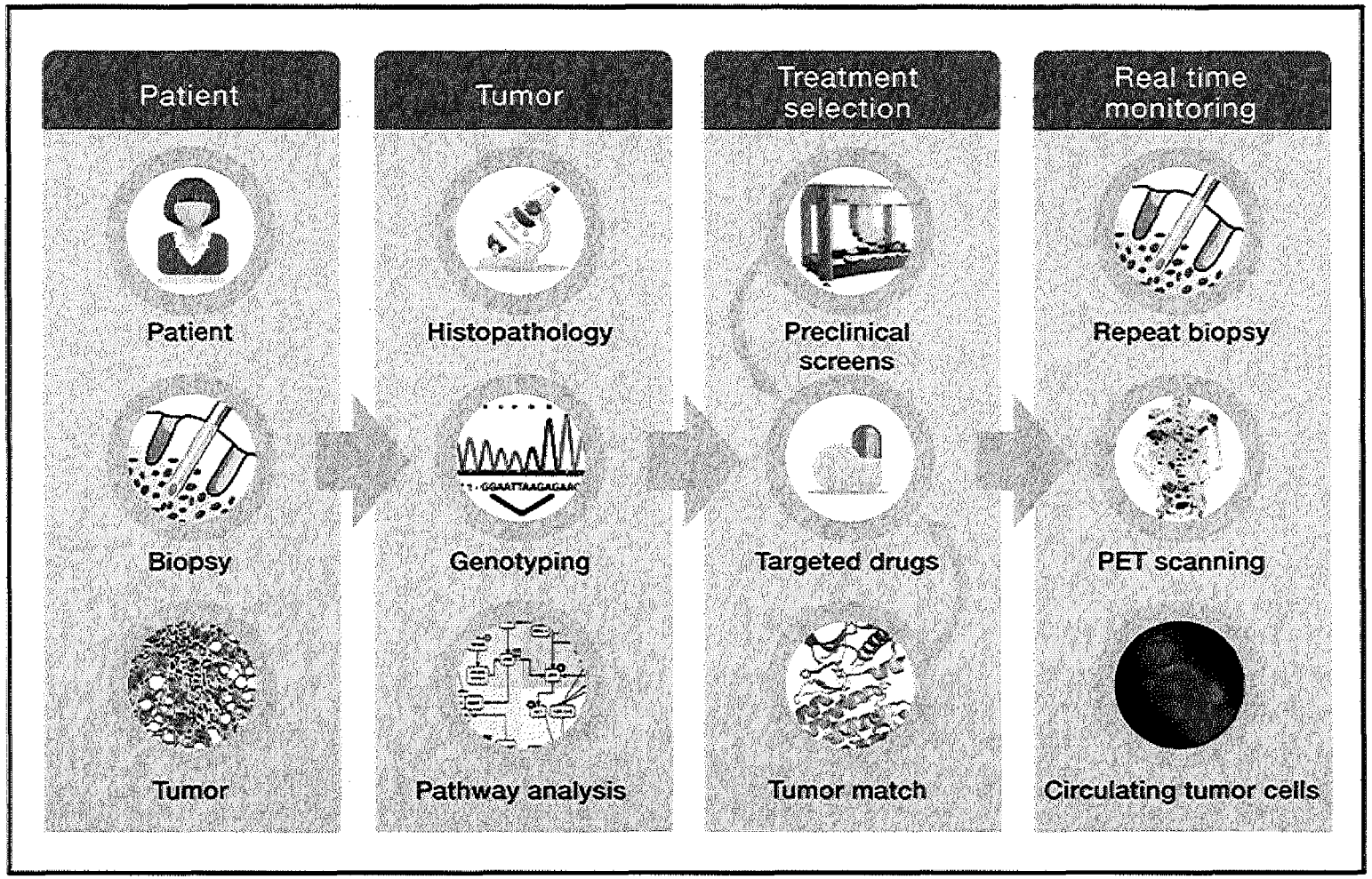
*Excludes basal and squamous cell skin cancers and in situ carcinoma except urinary bladder.

©2012, American Cancer Society, Inc., Surveillance Research



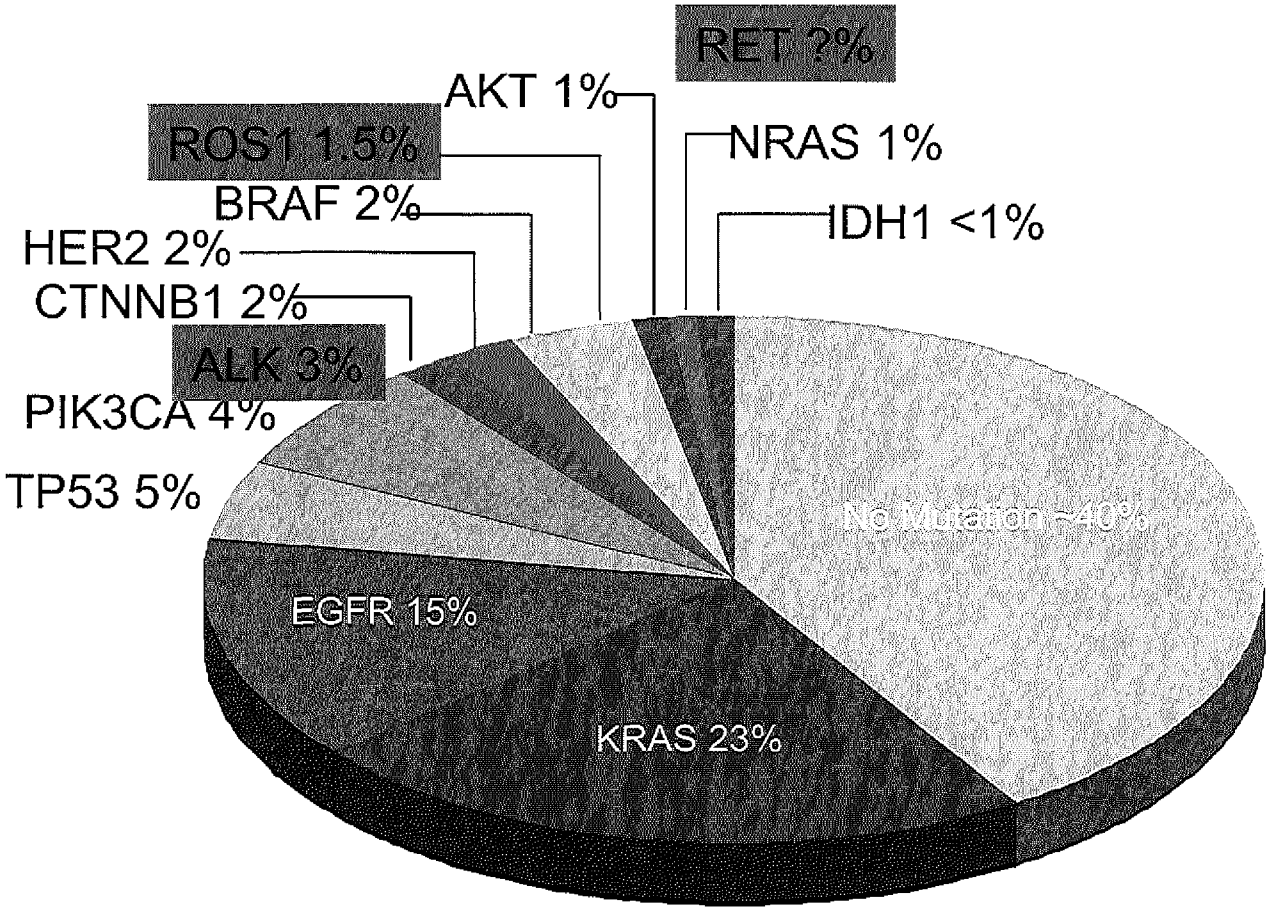
Comparison of Four Chemotherapy Regimens for Advanced Non–Small-Cell Lung Cancer
Schiller et al., NEJM 2002

A New Paradigm in Cancer Treatment



Haber, Gray, Baselga Cell 2011

Mutational profiling in lung cancers



MASSACHUSETTS GENERAL HOSPITAL
PATHOLOGY

Vol 448|2 August 2007|doi:10.1038/nature05945 nature

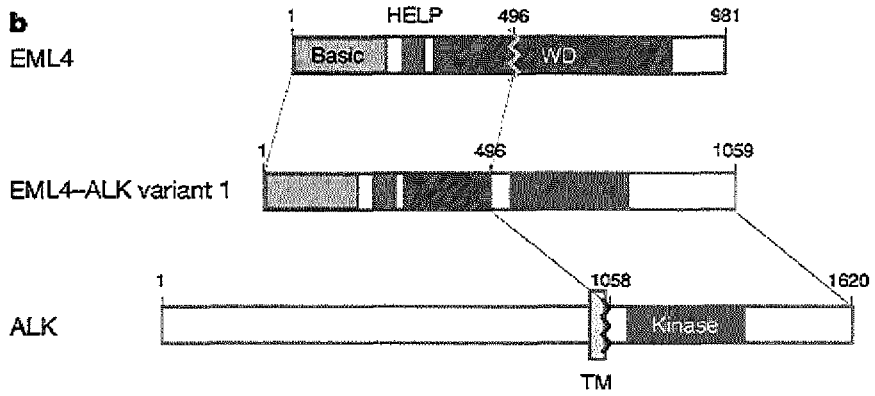
ARTICLES

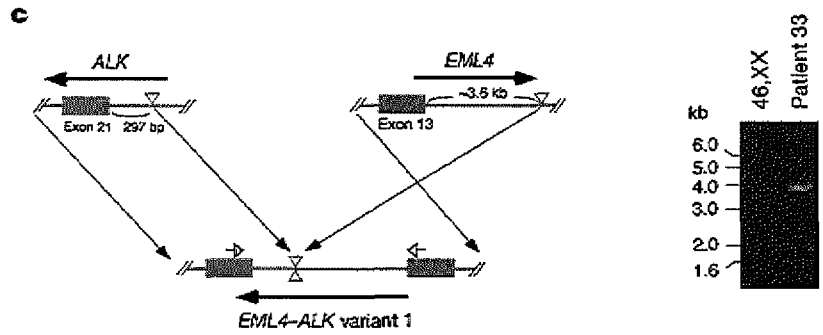
Identification of the transforming *EML4–ALK* fusion gene in non-small-cell lung cancer

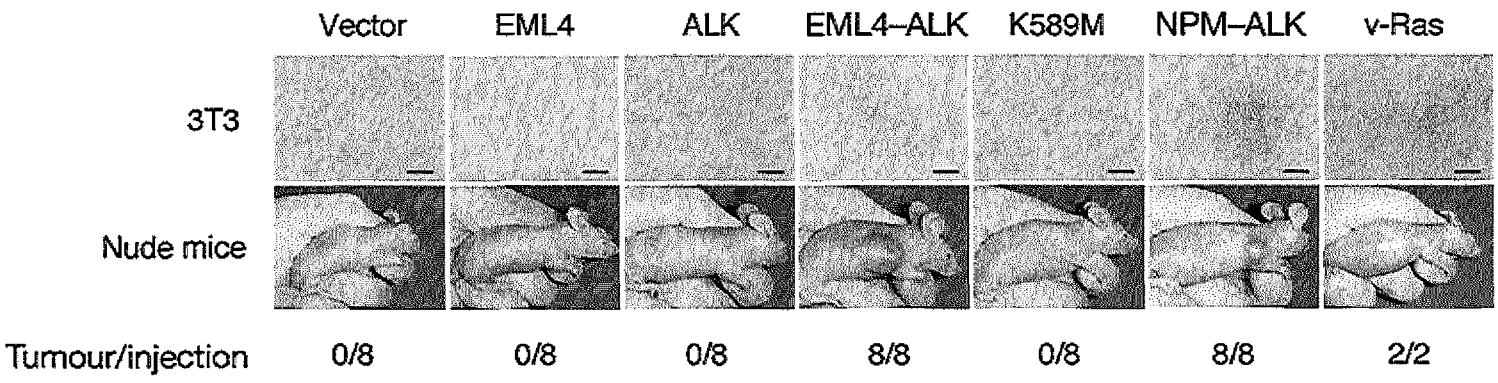
Manabu Soda[1,2], Young Lim Choi[1], Munehiro Enomoto[1,2], Shuji Takada[1], Yoshihiro Yamashita[1], Shunpei Ishikawa[5], Shin-ichiro Fujiwara[1], Hideki Watanabe[1], Kentaro Kurashina[1], Hisashi Hatanaka[1], Masashi Bando[2], Shoji Ohno[2], Yuichi Ishikawa[6], Hiroyuki Aburatani[5,7], Toshiro Niki[3], Yasunori Sohara[4], Yukihiko Sugiyama[2] & Hiroyuki Mano[1,7]

Improvement in the clinical outcome of lung cancer is likely to be achieved by identification of the molecular events that underlie its pathogenesis. Here we show that a small inversion within chromosome 2p results in the formation of a fusion gene comprising portions of the echinoderm microtubule-associated protein-like 4 (*EML4*) gene and the anaplastic lymphoma kinase (*ALK*) gene in non-small-cell lung cancer (NSCLC) cells. Mouse 3T3 fibroblasts forced to express this human fusion tyrosine kinase generated transformed foci in culture and subcutaneous tumours in nude mice. The *EML4–ALK* fusion transcript was detected in 6.7% (5 out of 75) of NSCLC patients examined; these individuals were distinct from those harbouring mutations in the epidermal growth factor receptor gene. Our data demonstrate that a subset of NSCLC patients may express a transforming fusion kinase that is a promising candidate for a therapeutic target as well as for a diagnostic molecular marker in NSCLC.

IDT Exhibit 1048.0011

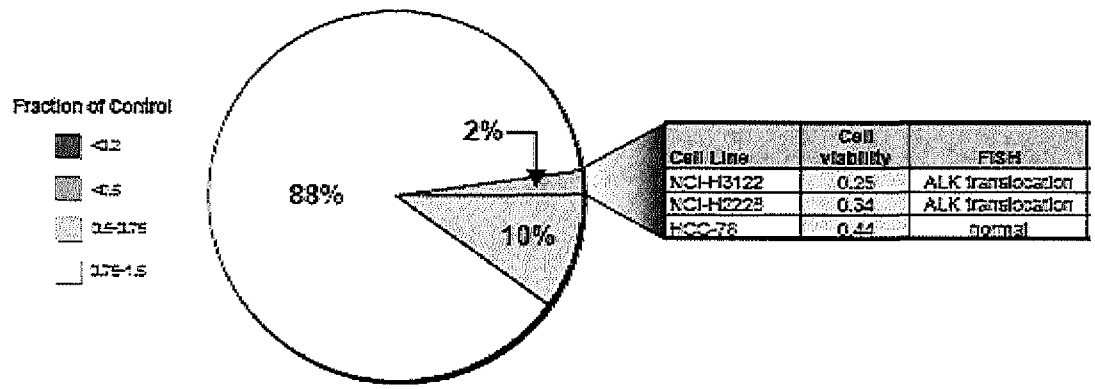

b
EML4
1
HELP
496
981
Basic
WD
EML4–ALK variant 1
1
496
1059
ALK
1
1058
1620
Kinase
TM


c
ALK
EML4
Exon 21
297 bp
~3.5 kb
Exon 13
EML4–ALK variant 1
kb
46,XX
Patient 33
6.0
5.0
4.0
3.0
2.0
1.6


Vector
EML4
ALK
EML4–ALK
K589M
NPM–ALK
v-Ras
3T3
Nude mice
Tumour/injection
0/8
0/8
0/8
8/8
0/8
8/8
2/2

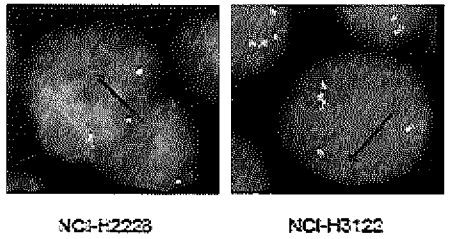
EML4-ALK+ Cell Lines Respond to ALK Inhibitor

A.

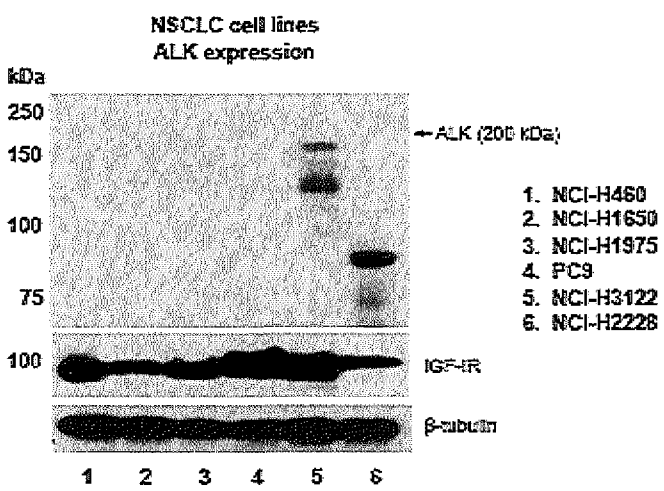


Cell Line	Cell viability	FISH
NCI-H3122	0.25	ALK translocation
NCI-H2228	0.34	ALK translocation
HCC-78	0.44	normal

B.



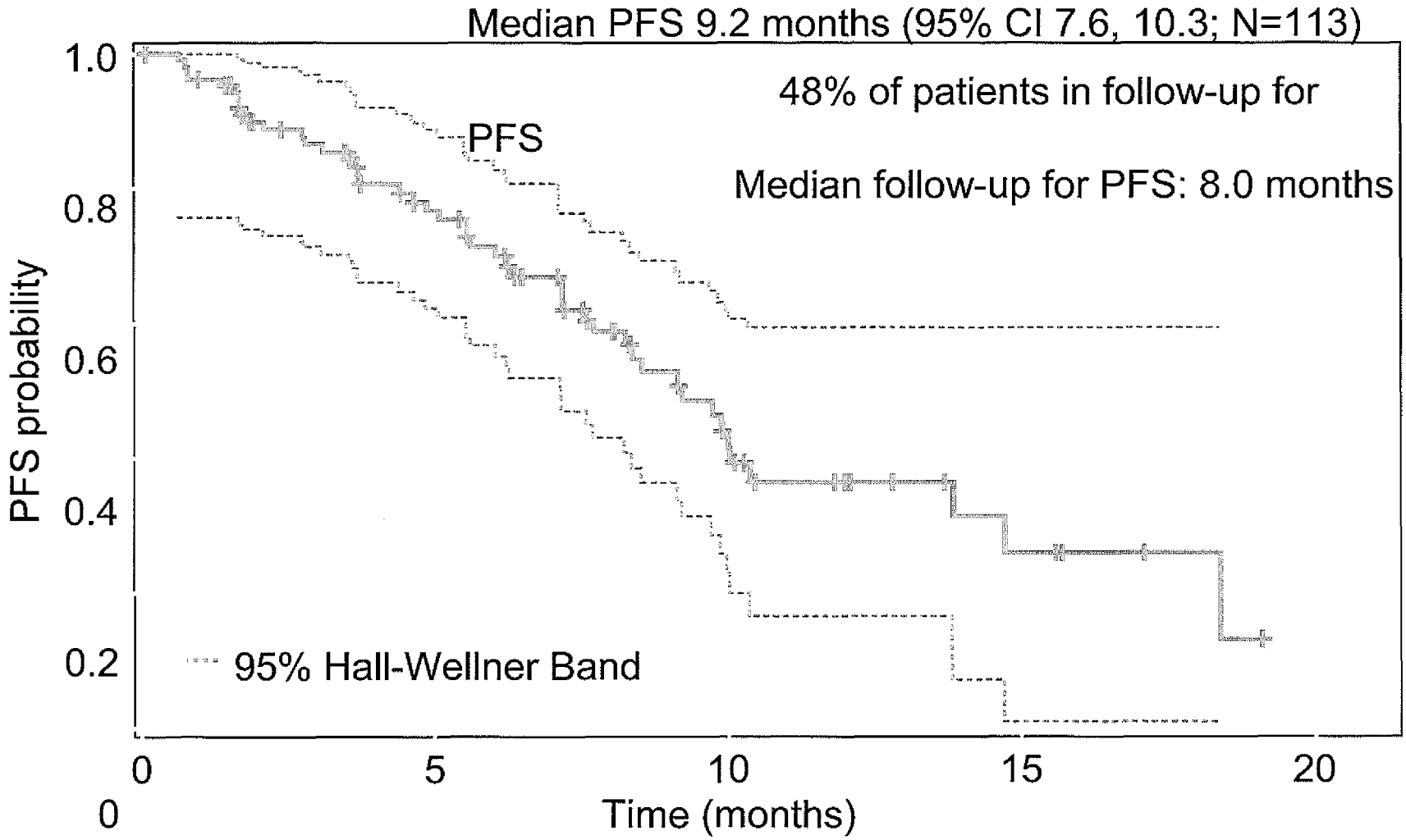
C.



D.

NCI-H2228 NCI-H3122

Progression Free Survival in *ALK*-positive NSCLC Patients Receiving Crizotinib



No. at risk	0	5	10	15	20
	113	59	17	5	0

Camidge R et al. Poster 366 presented at the 35th ESMO, 2010

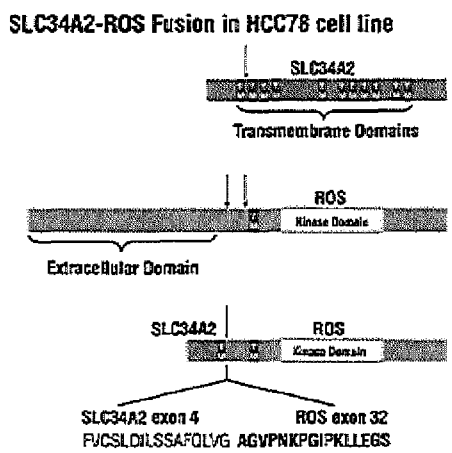
ROS1 rearrangement in NSCLC

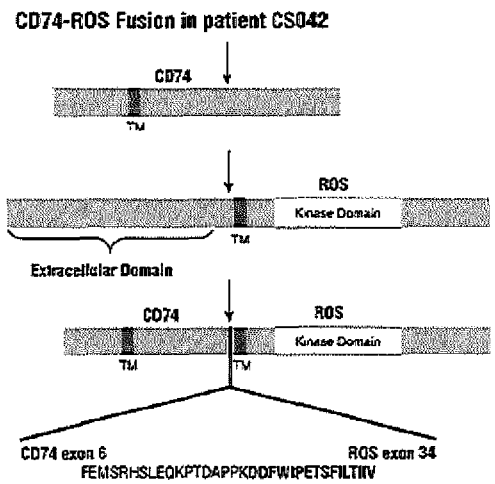
Cell

Resource

Global Survey of Phosphotyrosine Signaling Identifies Oncogenic Kinases in Lung Cancer

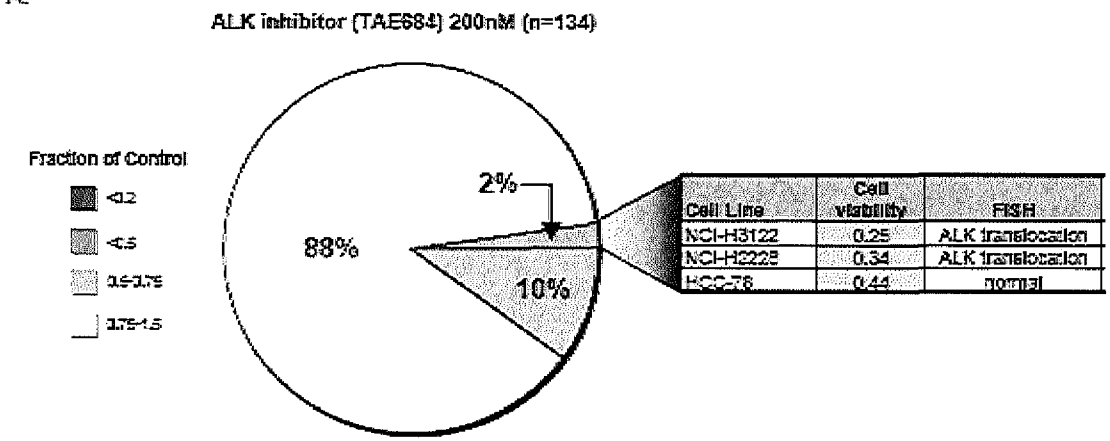
Klarisa Rikova,[1] Allan Guo,[1] Qingfu Zeng,[1] Anthony Possemato,[1] Jian Yu,[1] Herbert Haack,[1] Julie Nardone,[1] Kimberly Lee,[1] Cynthia Reeves,[1] Yu Li,[1] Yerong Hu,[3] Zhiping Tan,[3] Matthew Stokes,[1] Laura Sullivan,[1] Jeffrey Mitchell,[1] Randy Wetzel,[1] Joan MacNeill,[1] Jian Min Ren,[1] Jin Yuan,[1] Corey E. Bakalarski,[2] Judit Villen,[2] Jon M. Kornhauser,[1] Bradley Smith,[1] Daiqiang Li,[4] Xinmin Zhou,[3] Steven P. Gygi,[2] Ting-Lei Gu,[1] Roberto D. Polakiewicz,[1] John Rush,[1] and Michael J. Comb[1,*]
[1]Cell Signaling Technology, 3 Trask Lane, Danvers, MA 01923, USA
[2]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA
[3]Department of Cardiothoracic Surgery
[4]Department of Pathology
The Second Xiangya Hospital, Central South University, Changsha, Hunan, China 410011
*Correspondence: mcomb@cellsignal.com
DOI 10.1016/j.cell.2007.11.025





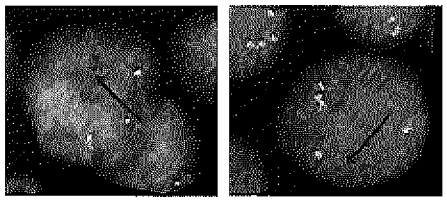
ROS1 rearrangement in NSCLC

A.

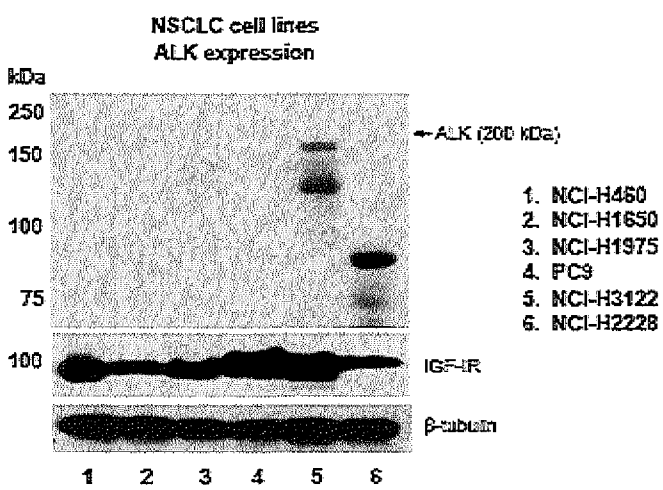


Cell Line	Cell viability	FISH
NCI-H3122	0.25	ALK translocation
NCI-H2228	0.34	ALK translocation
HCC-78	0.44	normal

B.



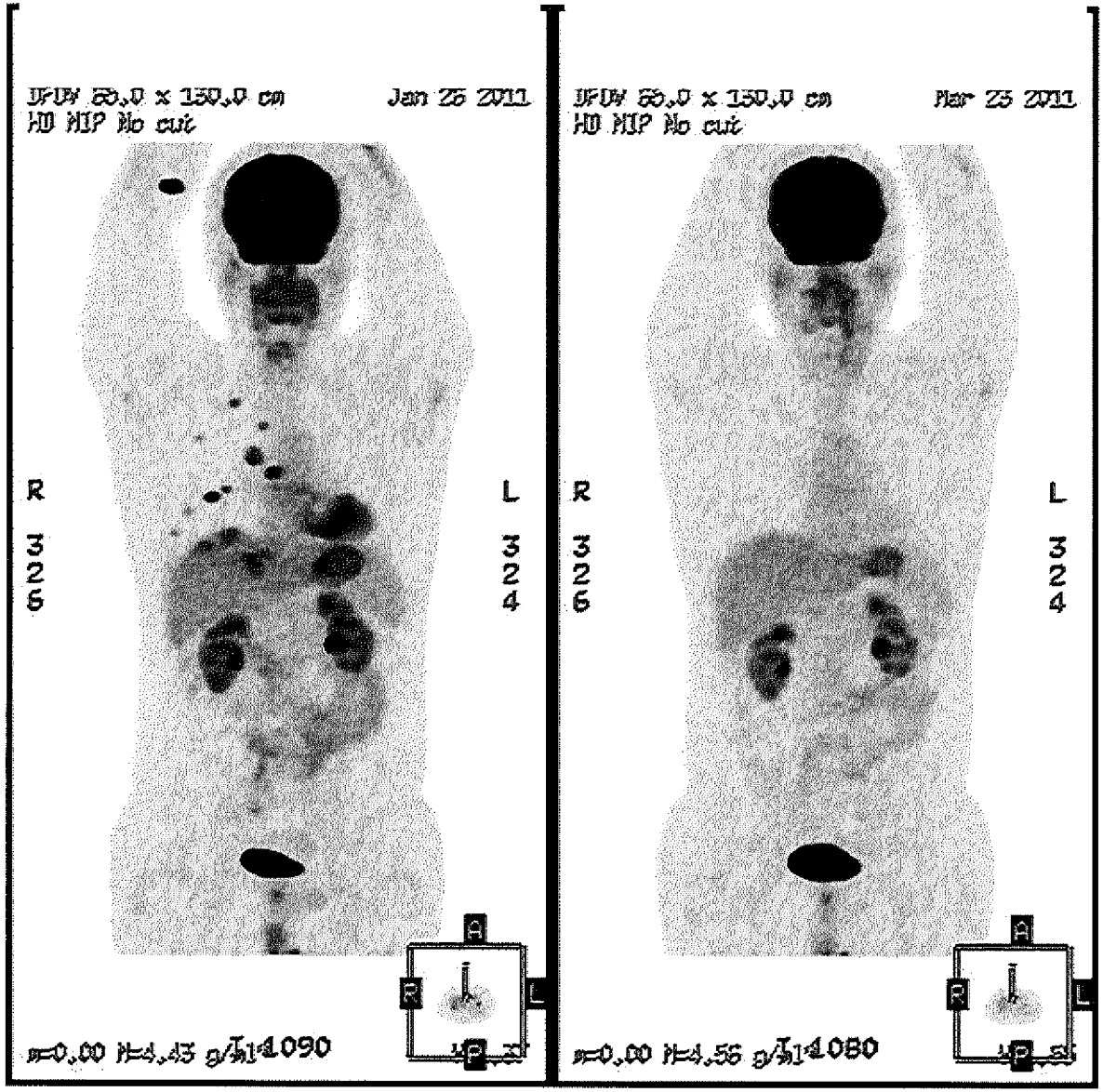
C.



D.

NCI-H2228 NCI-H3122

ROS1-positive patients respond to crizotinib

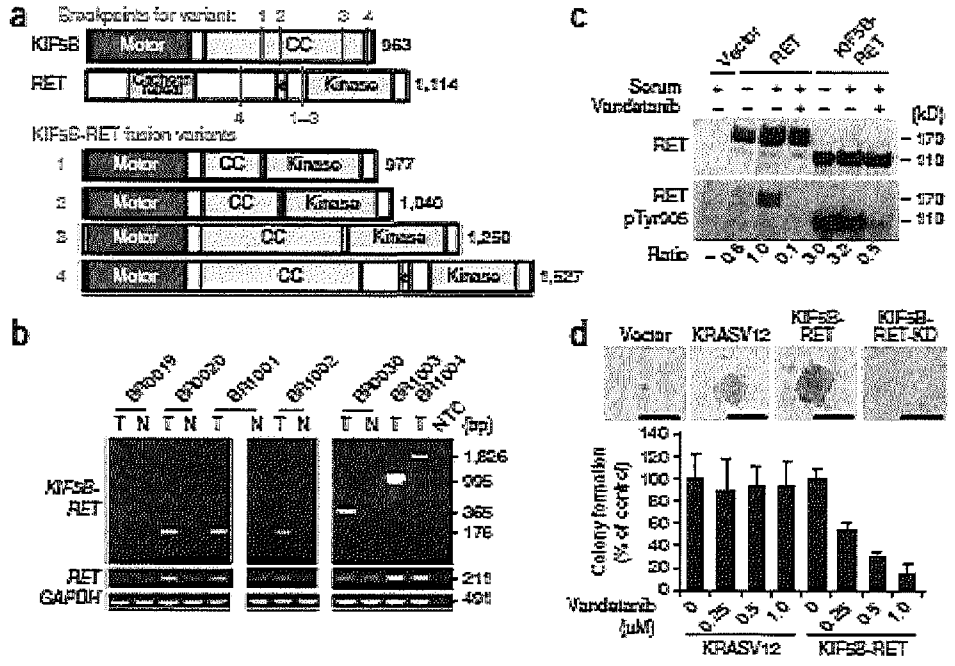


nature medicine

KIF5B-RET fusions in lung adenocarcinoma

Takashi Kohno[1,15], Hitoshi Ichikawa[2,15], Yasushi Totoki[3], Kazuki Yasuda[4], Masaki Hiramoto[4], Takao Nammo[4], Hiromi Sakamoto[2], Koji Tsuta[5], Koh Furuta[5], Yoko Shimada[1], Reika Iwakawa[6], Hideaki Ogiwara[1], Takahiro Oike[6], Masato Enari[7], Aaron J Schetter[8], Hirokazu Okayama[6,8], Aage Haugen[9], Vidar Skuterud[9], Suenori Chiku[10], Itaru Yamanaka[11], Yasuhito Arai[3], Shun-ichi Watanabe[12], Ikuo Sekine[13], Seishi Ogawa[14], Curtis C Harris[8], Hitoshi Tsuda[5], Teruhiko Yoshida[2], Jun Yokota[6] & Tatsuhiro Shibata[3]

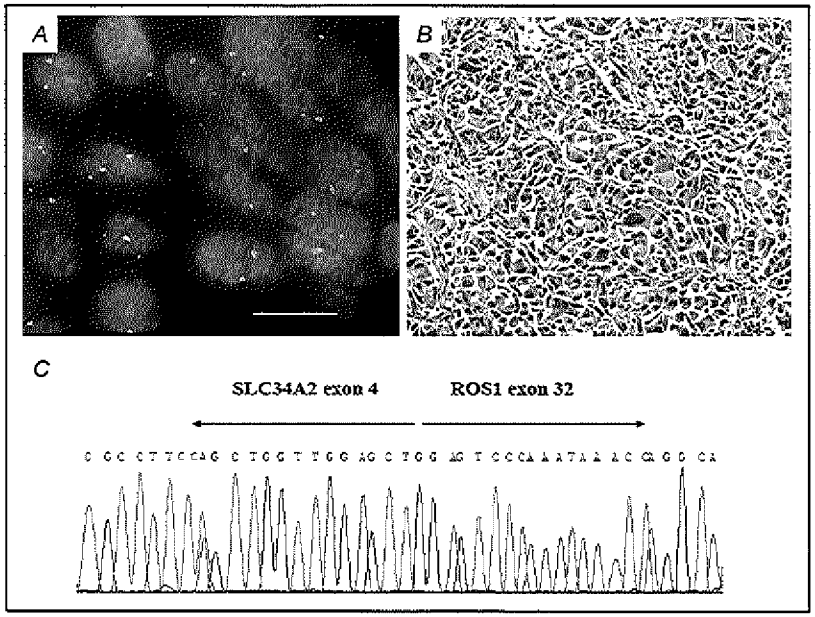
We identified in-frame fusion transcripts of *KIF5B* (the kinesin family 5B gene) and the *RET* oncogene, which are present in 1–2% of lung adenocarcinomas (LADCs) from people from Japan and the United States, using whole-transcriptome sequencing. The *KIF5B-RET* fusion leads to aberrant activation of RET kinase and is considered to be a new driver mutation of LADC because it segregates from mutations or fusions in *EGFR*, *KRAS*, *HER2* and *ALK*, and a RET tyrosine kinase inhibitor, vandetanib, suppresses the fusion-induced anchorage-independent growth activity of NIH3T3 cells.

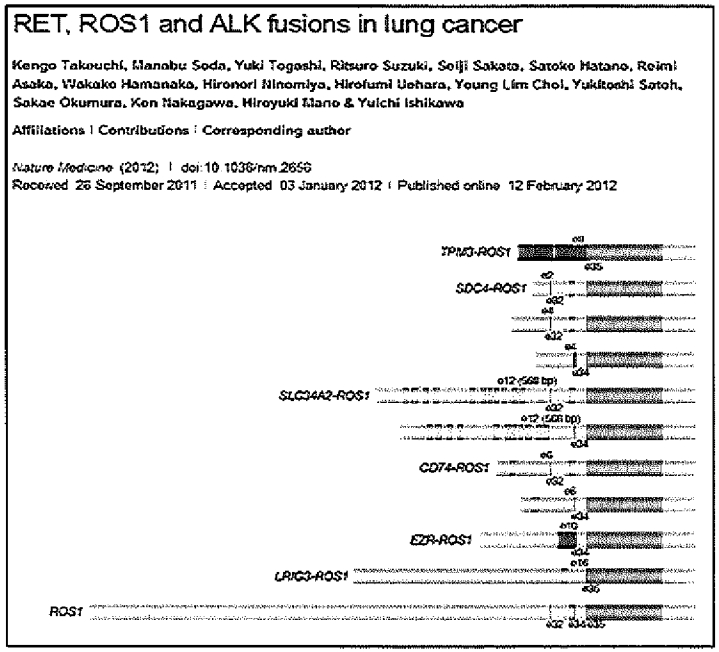


IDT Exhibit 1048.0018

ALK, ROS1 Detection

	Concerns
- FISH	Expertise, difficulty in close proximity fusion
- IHC	Antibody, specificity
- RT-PCR	Requires known fusion partners
	(All are low throughput)




RET, ROS1 and ALK fusions in lung cancer

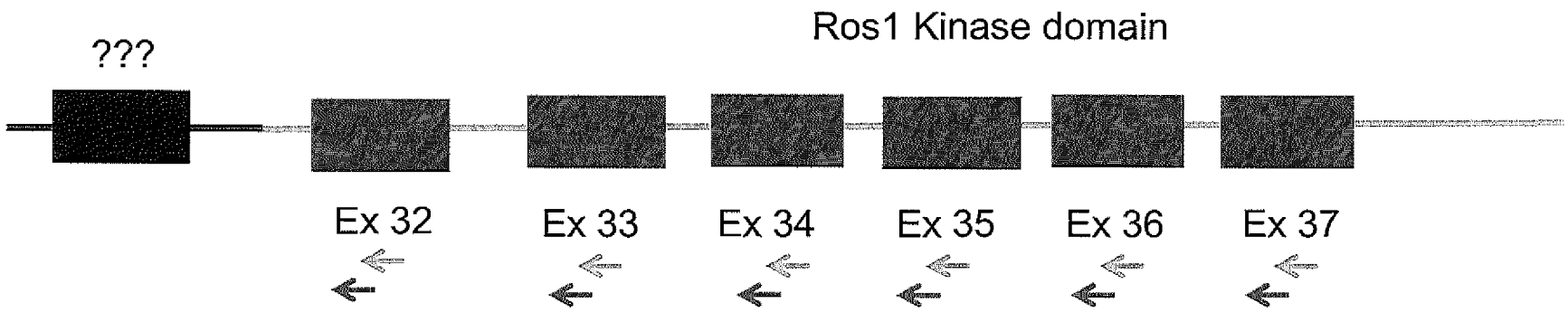
Kengo Takeuchi, Manabu Soda, Yuki Togashi, Ritsuro Suzuki, Seiji Sakata, Satoko Hatano, Reimi Asaka, Wakako Hamanaka, Hironori Ninomiya, Hirofumi Uehara, Young Lim Choi, Yukitoshi Satoh, Sakae Okumura, Ken Nakagawa, Hiroyuki Mano & Yuichi Ishikawa

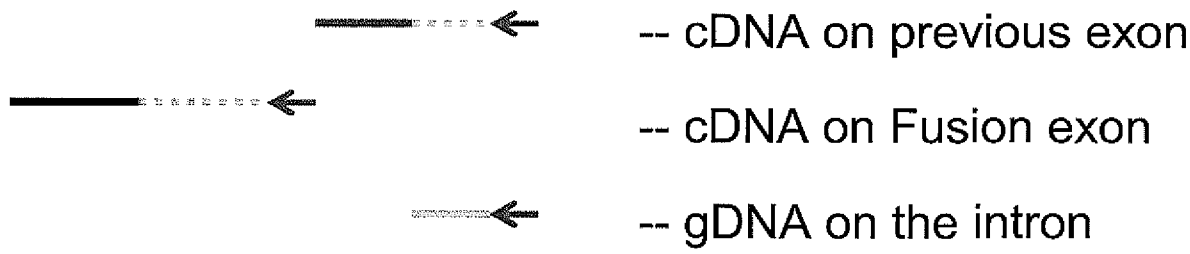
Affiliations | Contributions | Corresponding author

Nature Medicine (2012) | doi:10.1038/nm.2658
Received 26 September 2011 | Accepted 03 January 2012 | Published online 12 February 2012

A next-generation sequencing based assay

Primer design





Possible types of extension products

gDNA & RNA from FFPE

Standard double strand cDNA synthesis using
- SuperScript III
- Random hexamer
- RNase H
- DNA polymerase I

gDNA / cDNA

A-B library

End polishing, 5′ phosphorylation

Ligation of adapters

A B

Nick repair

Two effective molecules

Y library

End polishing, 5′ phosphorylation, adenylation and ligation of Y adapter

Without cleanup in between

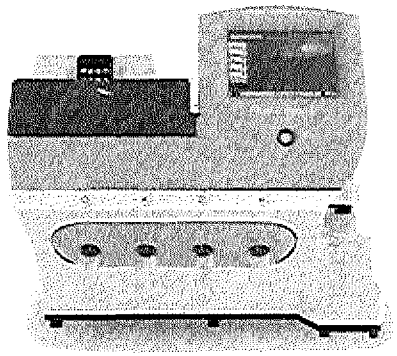
Eight effective molecules

One of the potential forms, if no cleanups in between

Target enrichment for ALK, ROS1, RET

2 rounds of PCR

Ion Torrent library

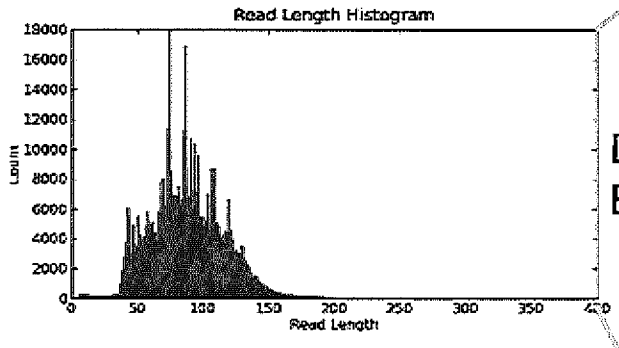


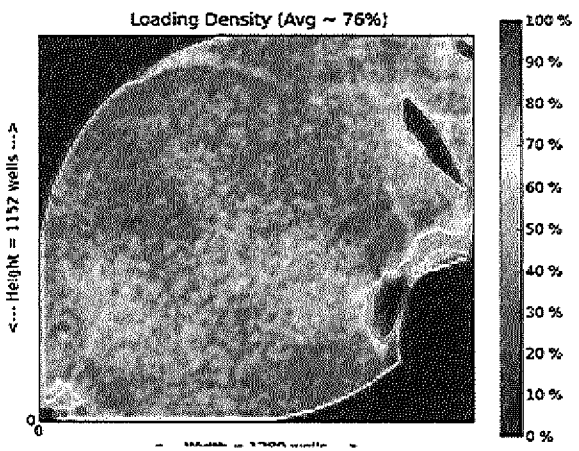
Ion PGM™ Sequencer

IDT Exhibit 1048.0021

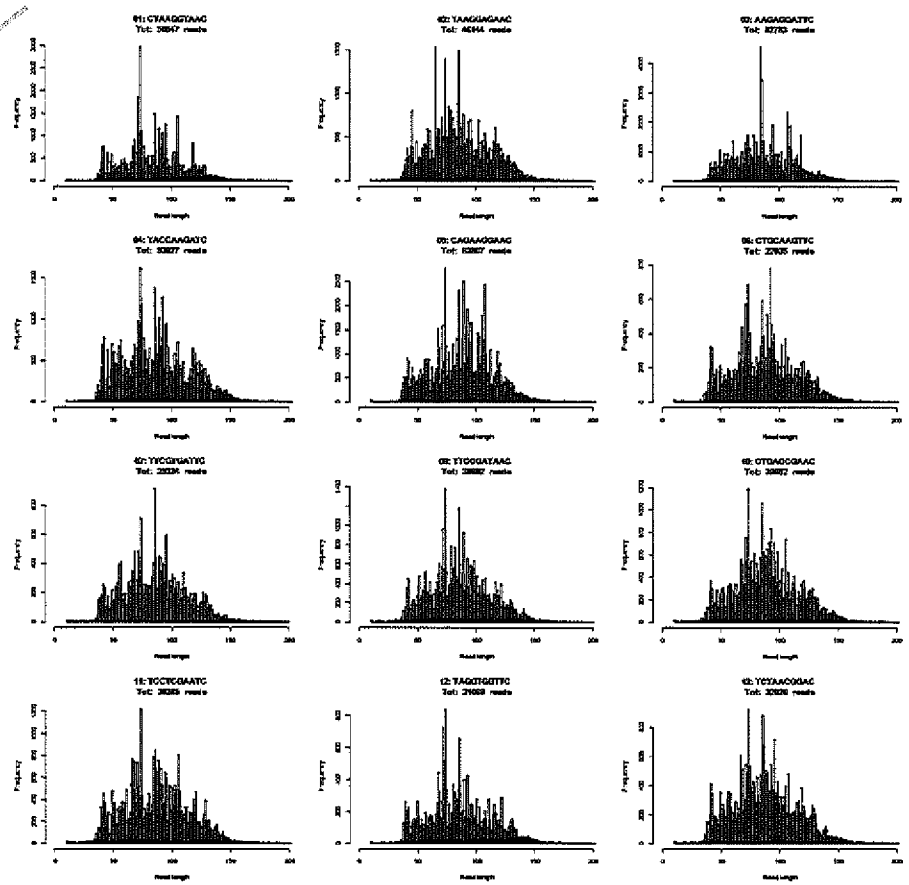
Sequencing results

Total Number of Bases [Mbp]	49.19
▸ Number of Q17 Bases [Mbp]	43.86
▸ Number of Q20 Bases [Mbp]	41.58
Total Number of Reads	563,255
Mean Length [bp]	87
Longest Read [bp]	203

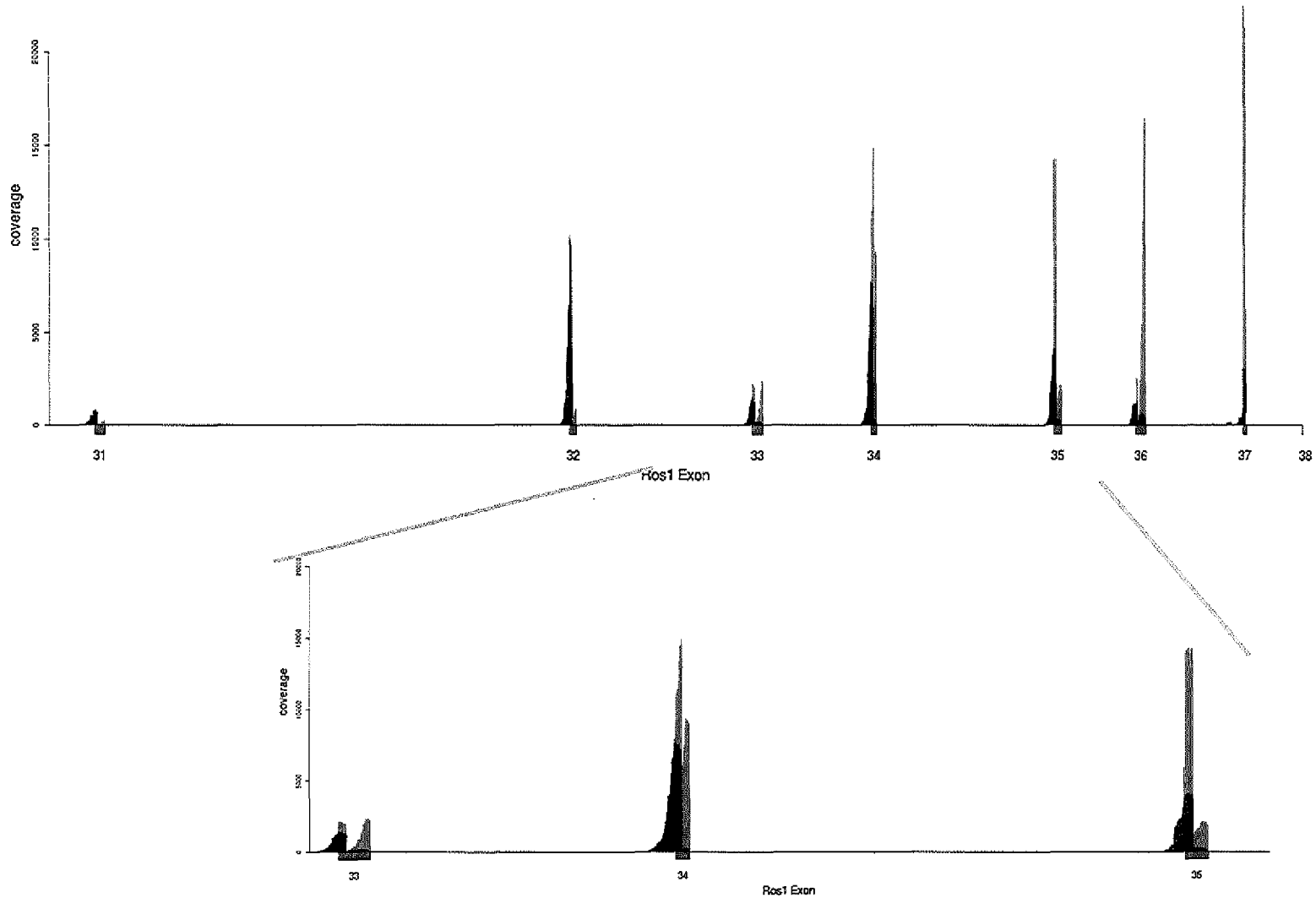


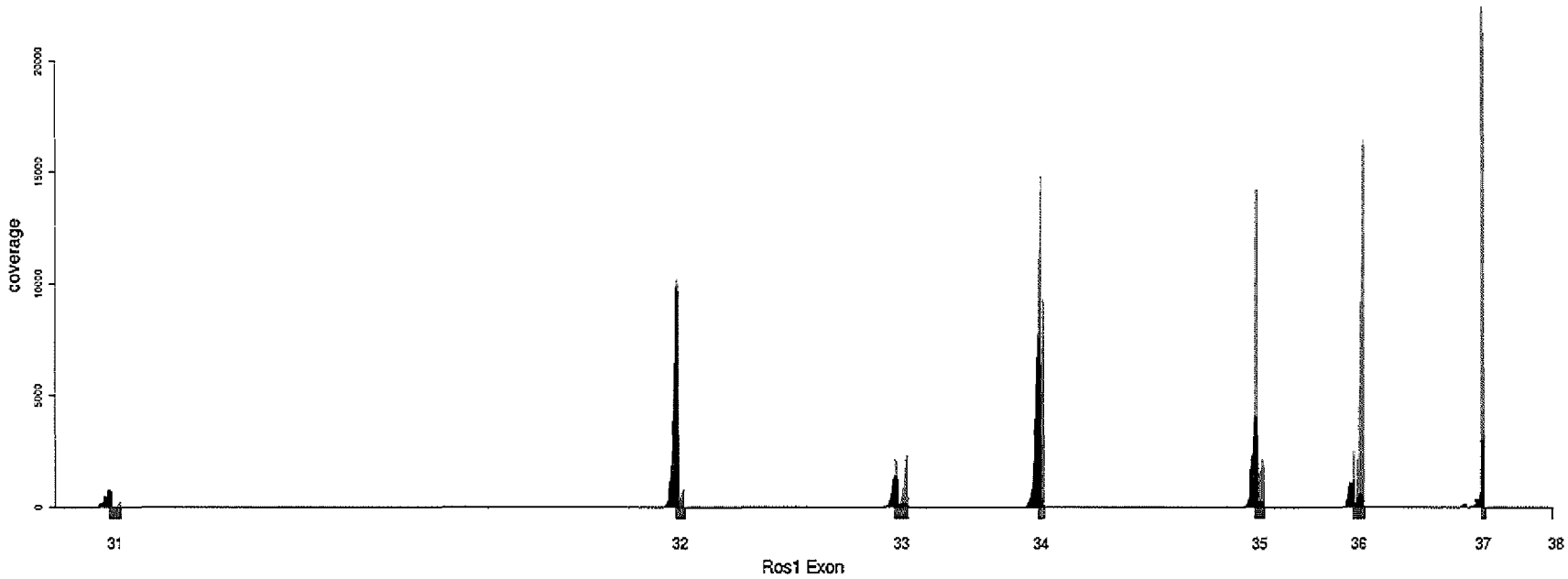


De-Barcode



One example - Coverage on ROS1





Alternative splicing

SLC34A2 Exon 13 ROS1

AGGATGTCCCTGTCAAGGCTCCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAAATAAACCAGGCATTCCCAAATTACTAGA
TCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCAAATCAAACAGGCATTCCCAAATACTAGA
GAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAAATAAACCAGGCATTCCCAAAATTACTAGA
GTCCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAAATAAACCAGGCATTCCCAAATTACTAGA
} Exon 32

GCAGGATGTCCCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGATGATTTTTGGATACCAGAAACAAGTTTCATACTTACTA
GGCGCAGGAGGGGCAGGATGTCCCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGATGATTTTTGGATACCAGAAACAAGTTTCATACTTACTA
} Exon 34

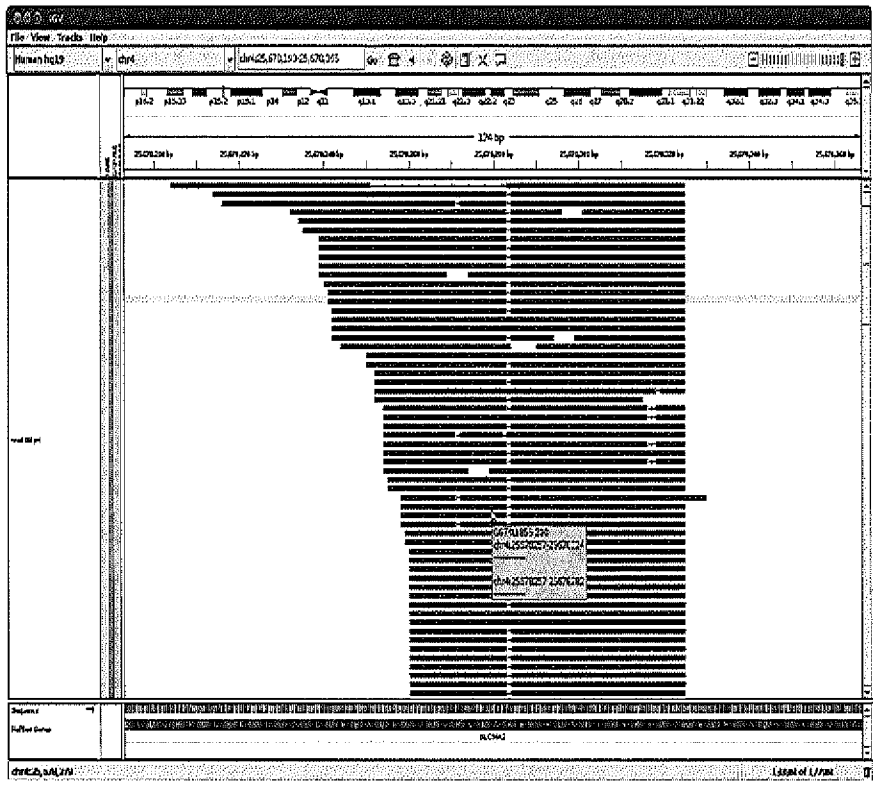
Specificity: 91.8% (= 127,446 / 138,787)

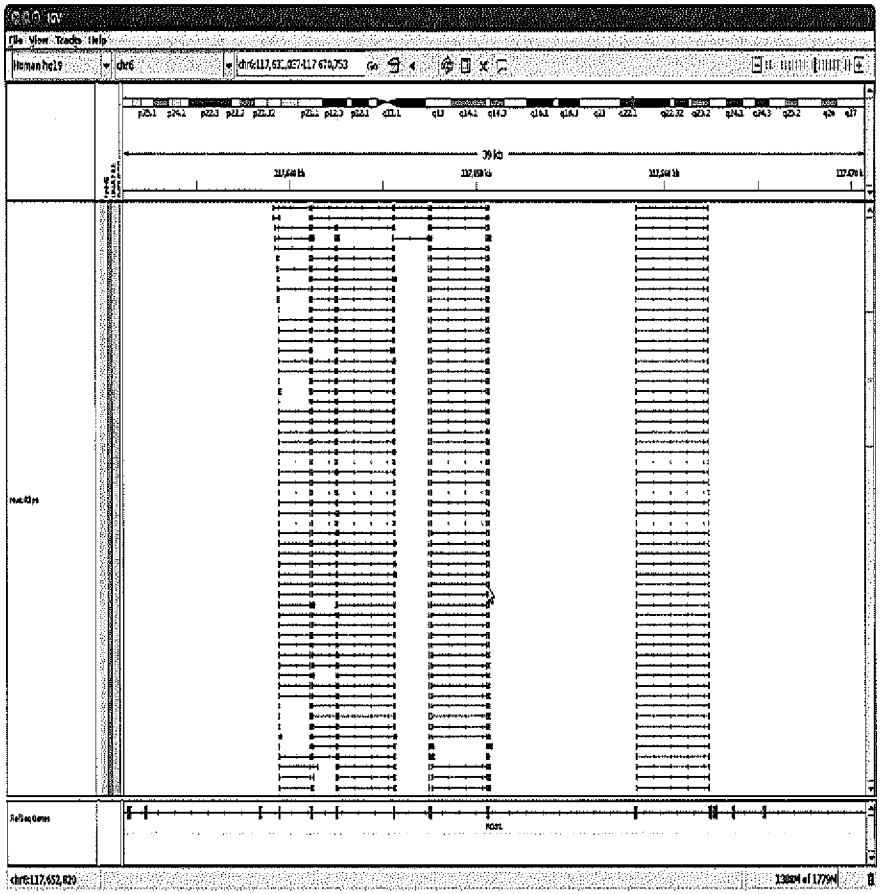
(On-target: have not yet included reads mapped solely to the fusion gene).

IDT Exhibit 1048.0025

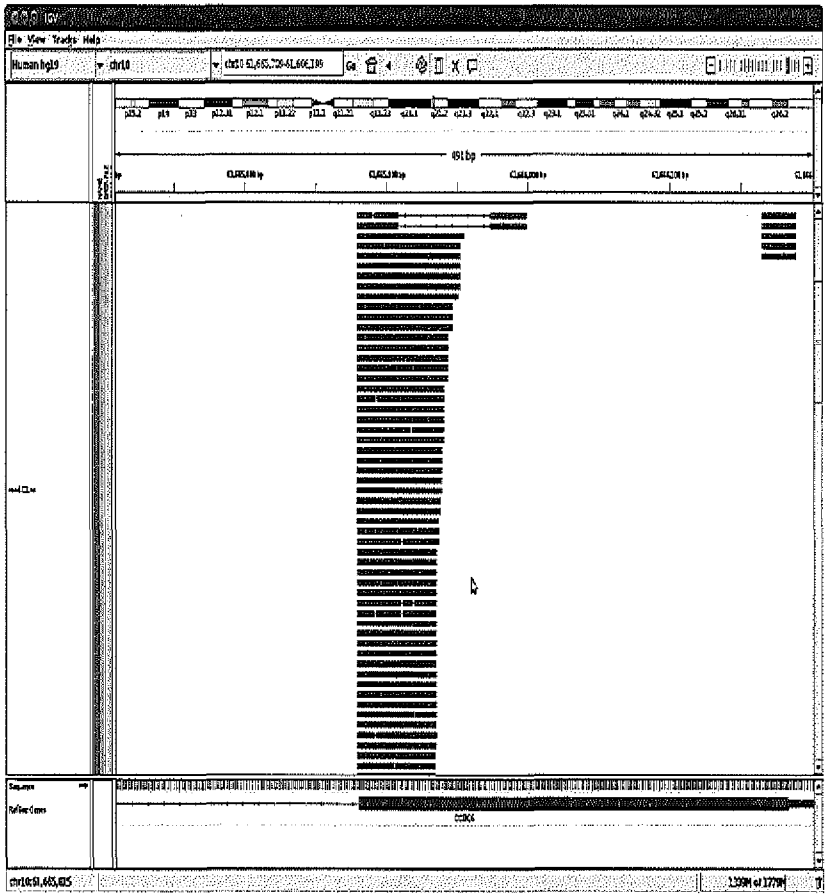
Mapping visualization in Integrated Genome Viewer

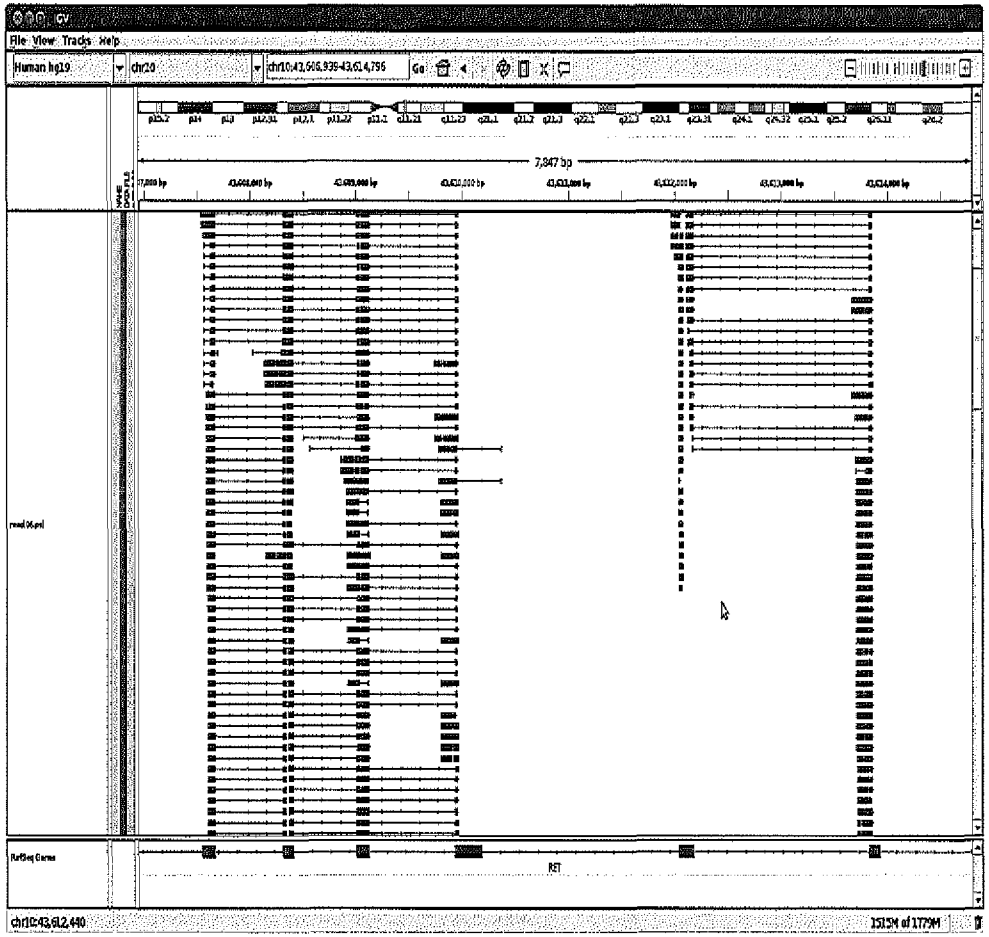
SLC34A2—ROS1





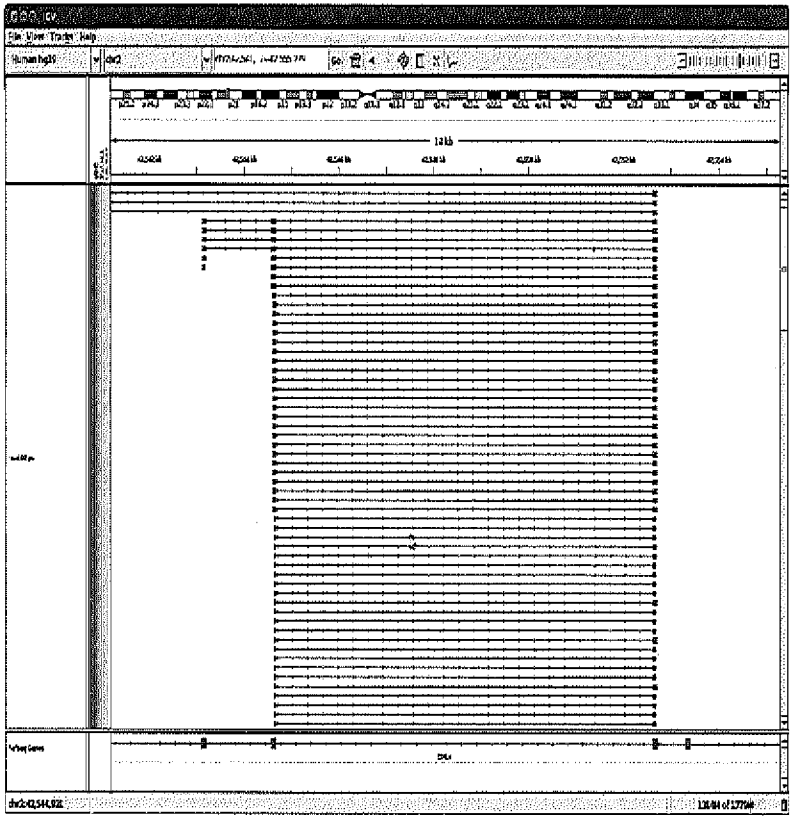
CCDC6---RET

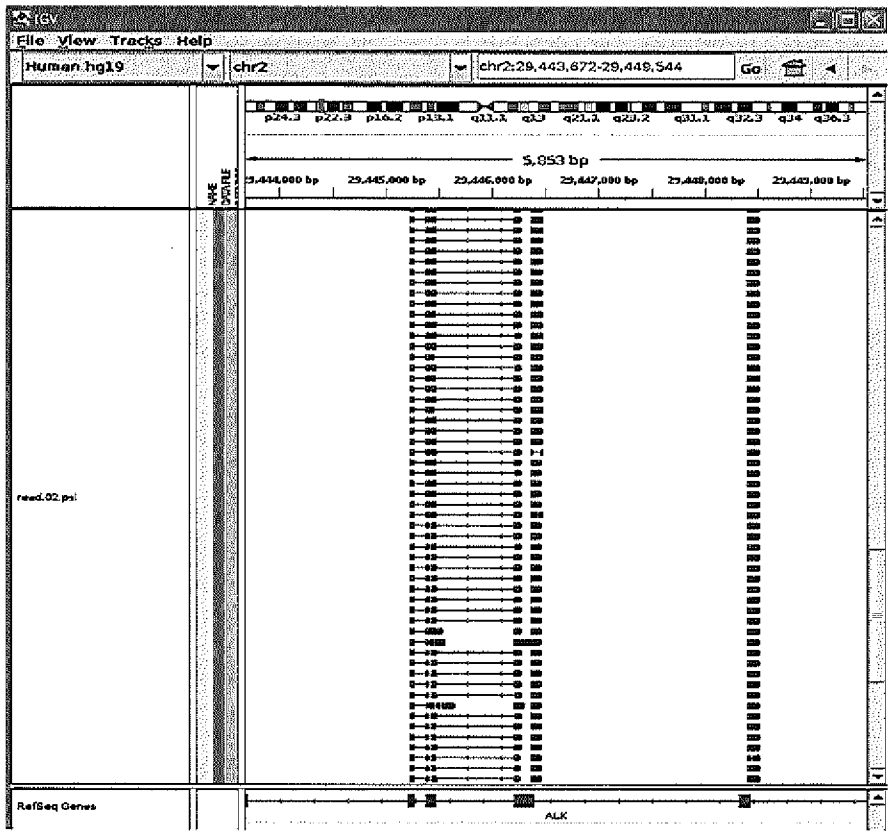




IDT Exhibit 1048.0026

EML4---ALK





IDT Exhibit 1048.0027

Summary

This targeted next-gen sequencing assay

-Can detect *ALK, ROS1, RET* gene fusions using FFPE
-Even target distribution
-High coverage (high analytical sensitivity)
-Low cost, high throughput
-Simple bioinformatics

Clinical Sensitivity and Specificity
-- ongoing study in a lung cancer cohort

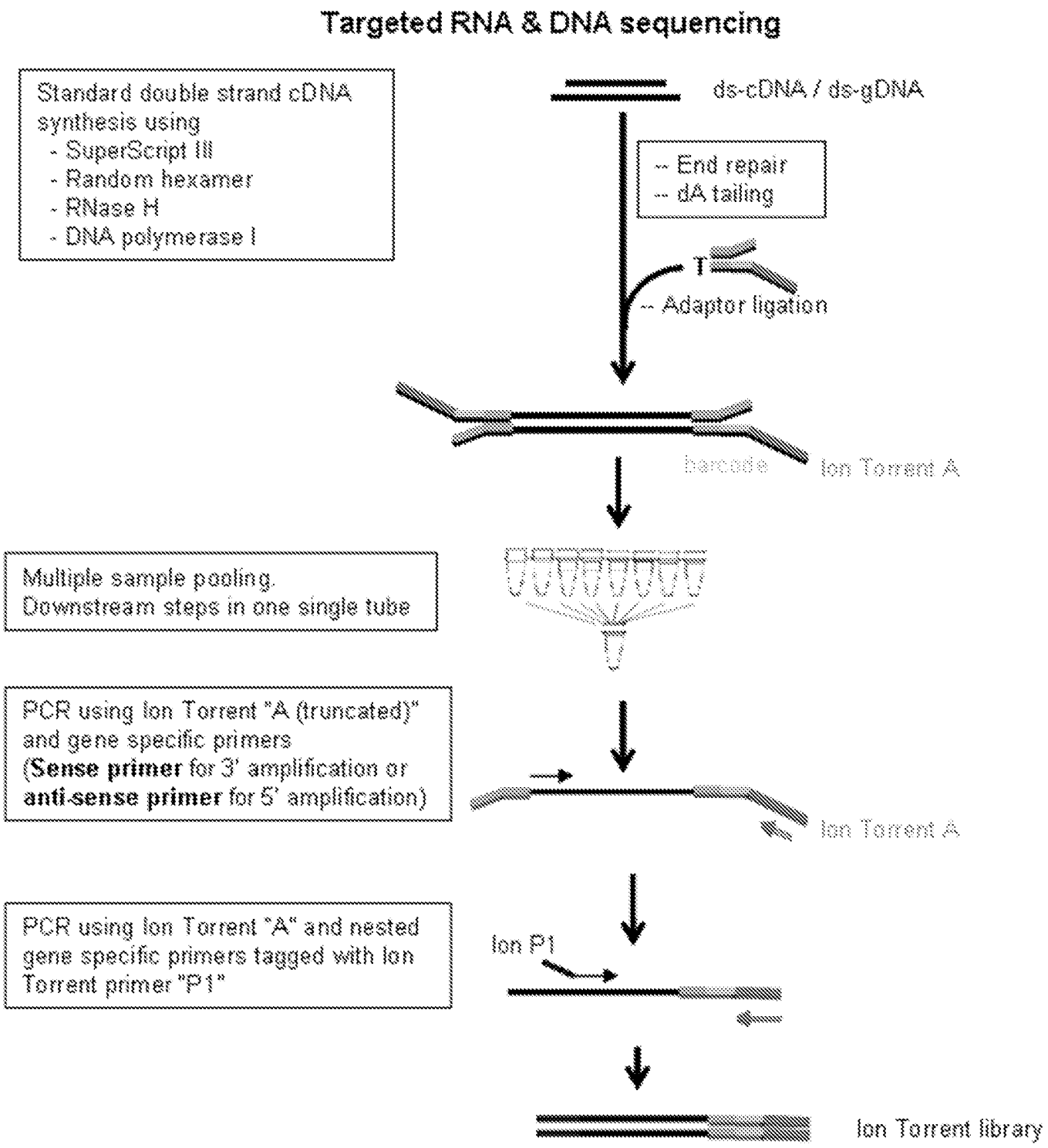

Targeted RNA & DNA sequencing
Standard double strand cDNA synthesis using
- SuperScript III
- Random hexamer
- RNase H
- DNA polymerase I
ds-cDNA / ds-gDNA
-- End repair
-- dA tailing
T
-- Adaptor ligation
barcode
Ion Torrent A
Multiple sample pooling.
Downstream steps in one single tube
PCR using Ion Torrent "A (truncated)" and gene specific primers
(Sense primer for 3' amplification or anti-sense primer for 5' amplification)
Ion Torrent A
PCR using Ion Torrent "A" and nested gene specific primers tagged with Ion Torrent primer "P1"
Ion P1
Ion Torrent library

Figure 1

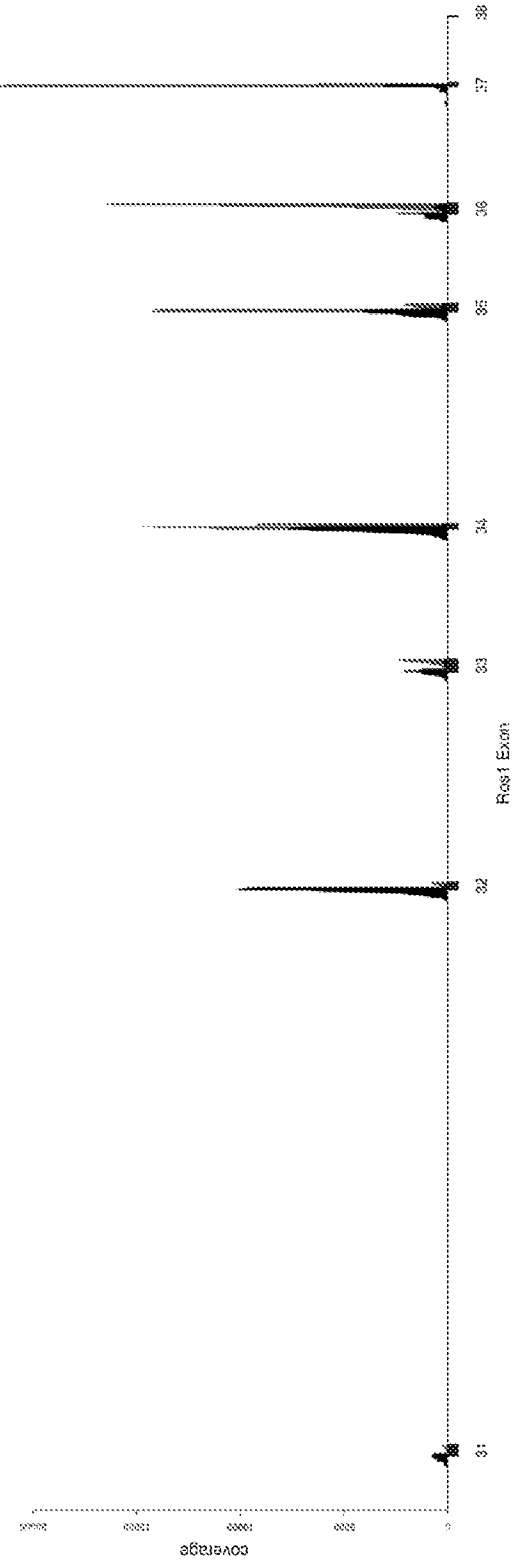
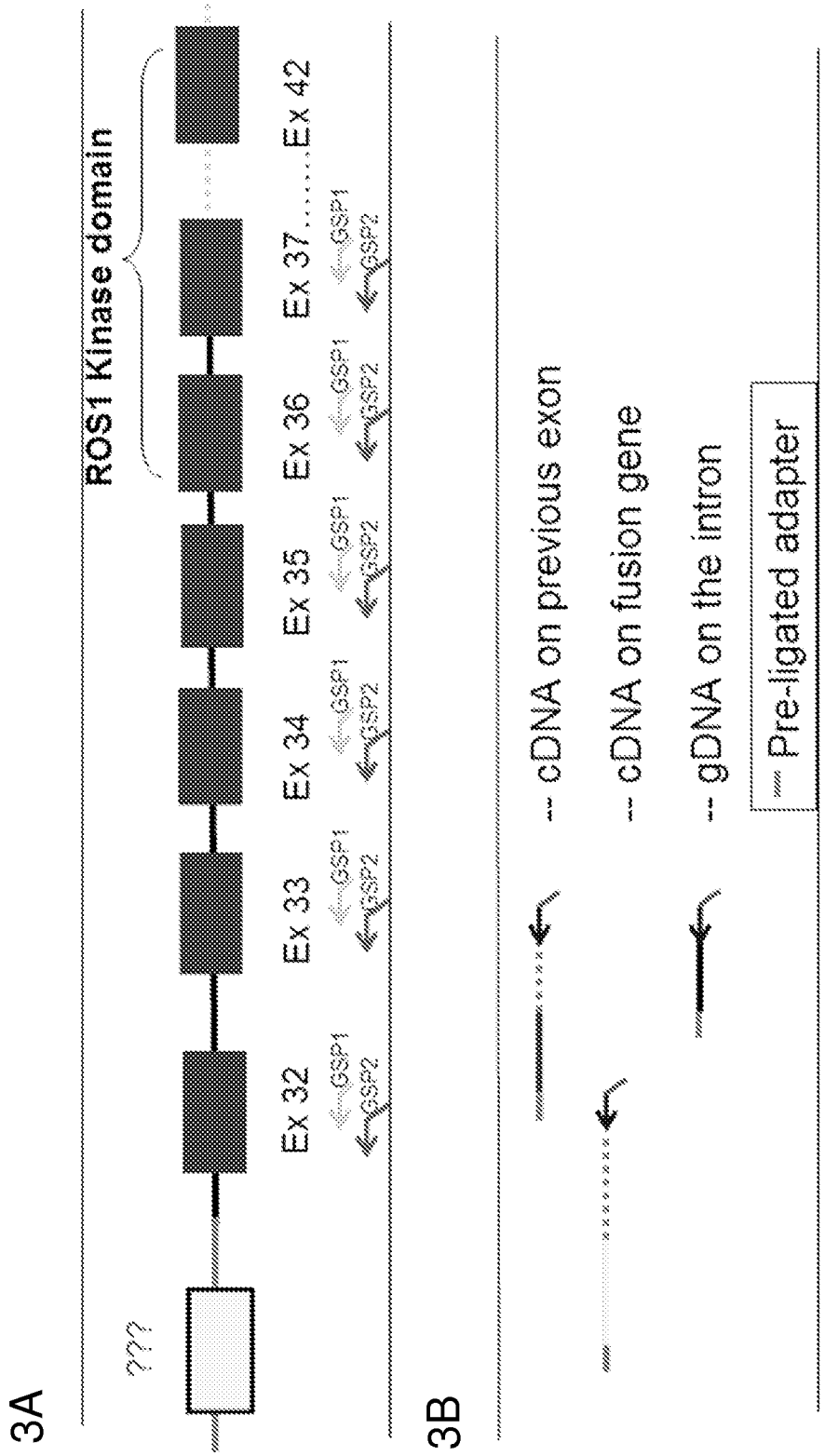
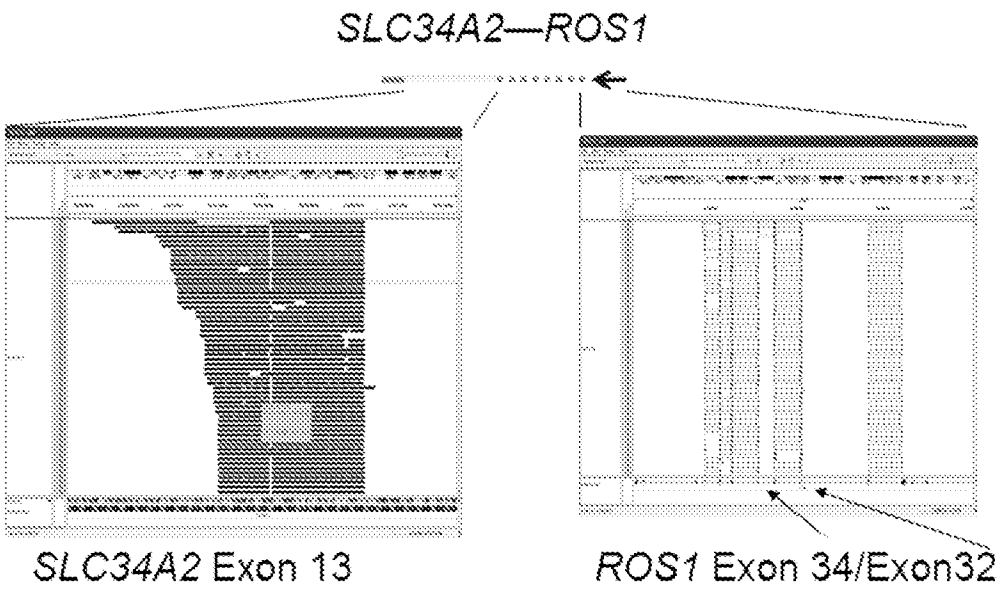


Figure 2

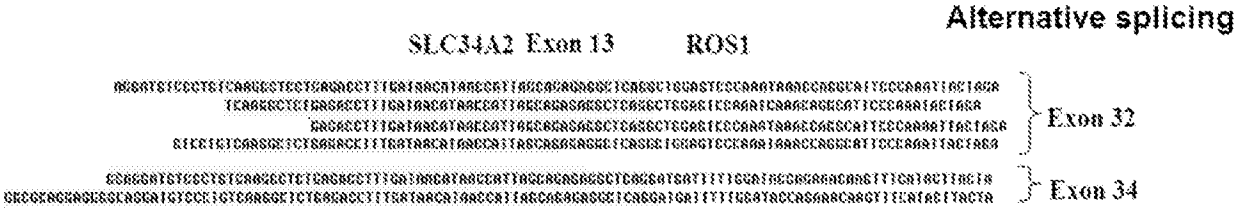


Figures 3A-3B

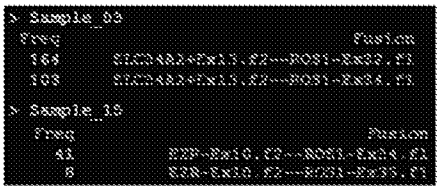
4A



4B



4C



Figures 4A – 4C

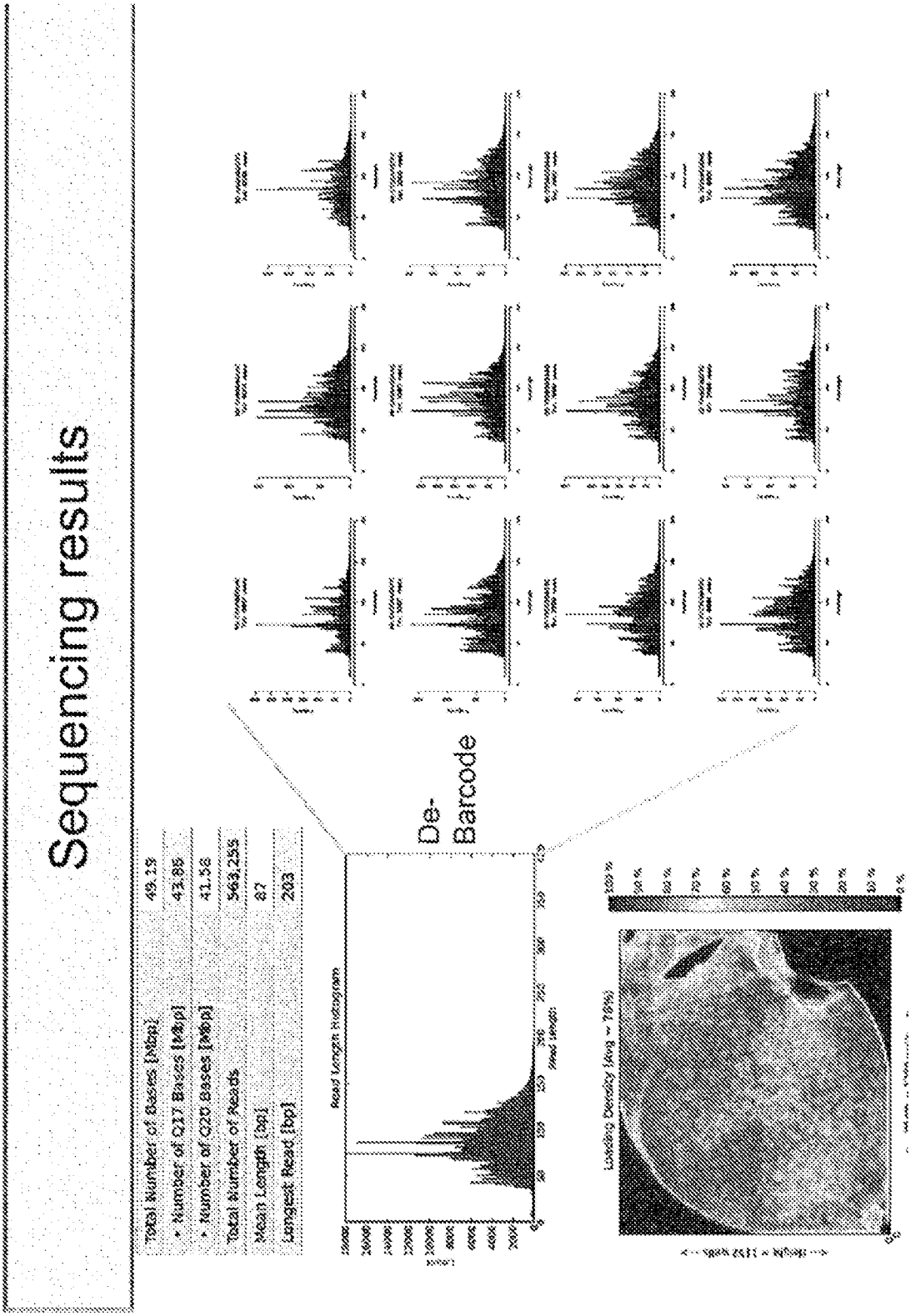

Sequencing results
Total Number of Bases [Mbp] 49.19
Number of Q17 Bases [Mbp] 43.85
Number of Q20 Bases [Mbp] 41.58
Total Number of Reads 563,255
Mean Length [bp] 87
Longest Read [bp] 203
De-Barcode
Read Length Histogram
Loading Density (Avg = 76%)

Figure 5

CCDC6--RET

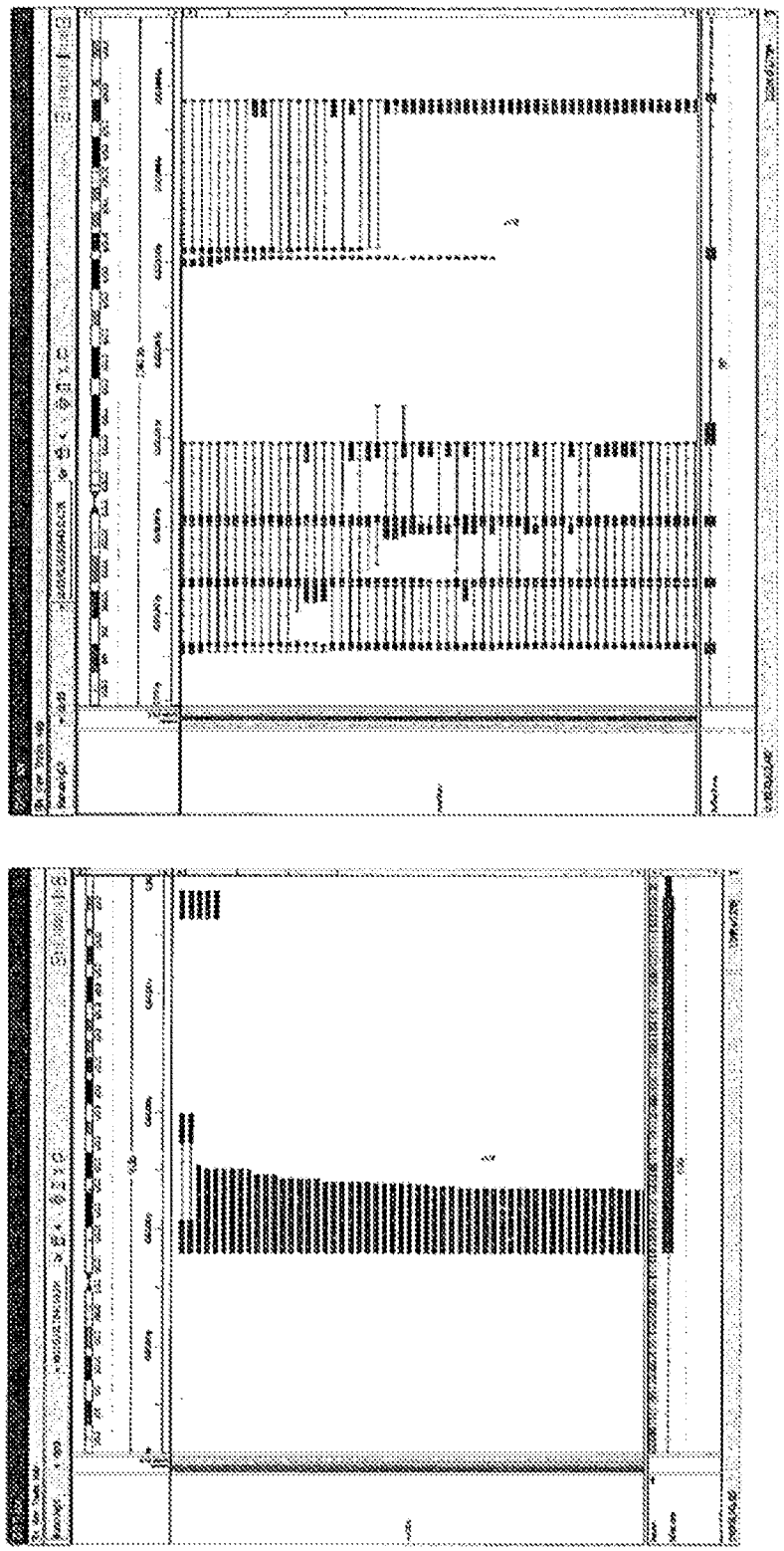


Figure 6

EML4--ALK

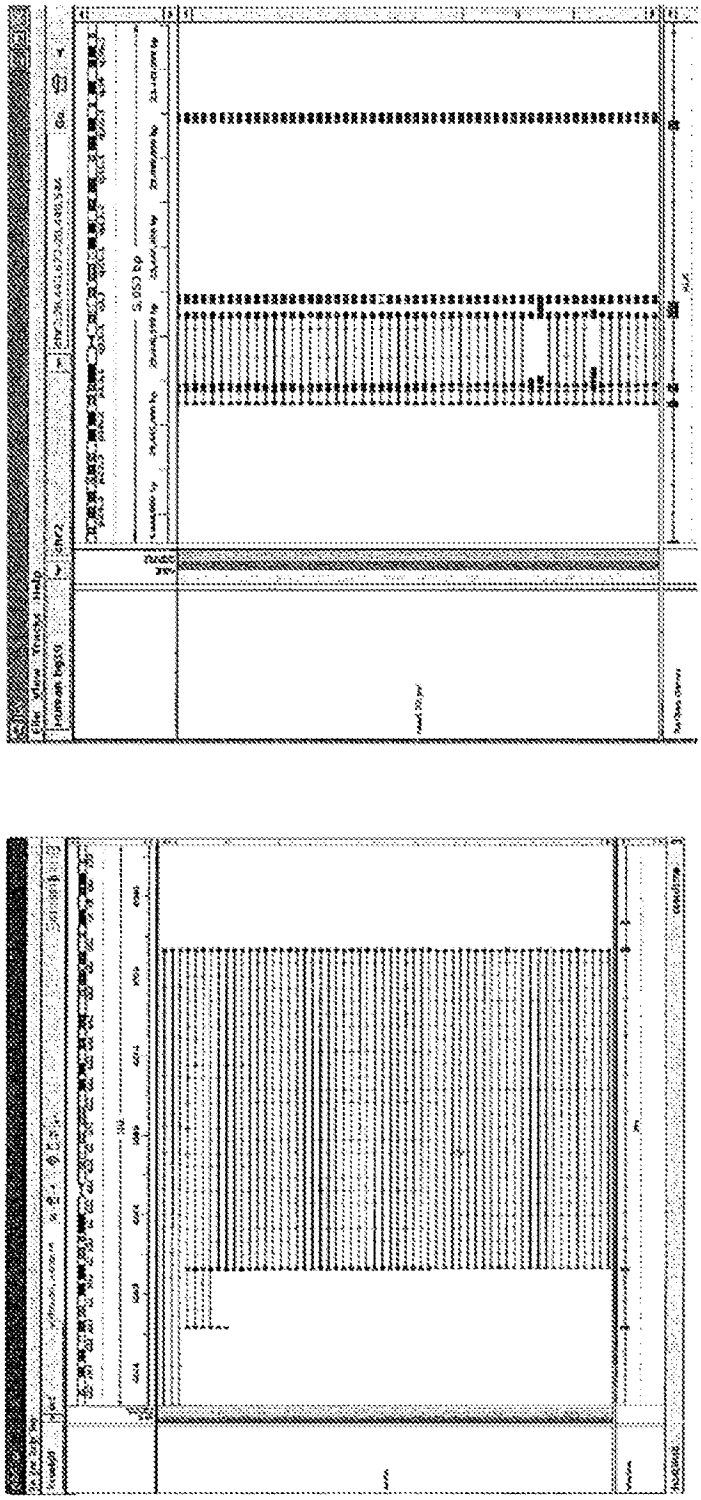


Figure 7

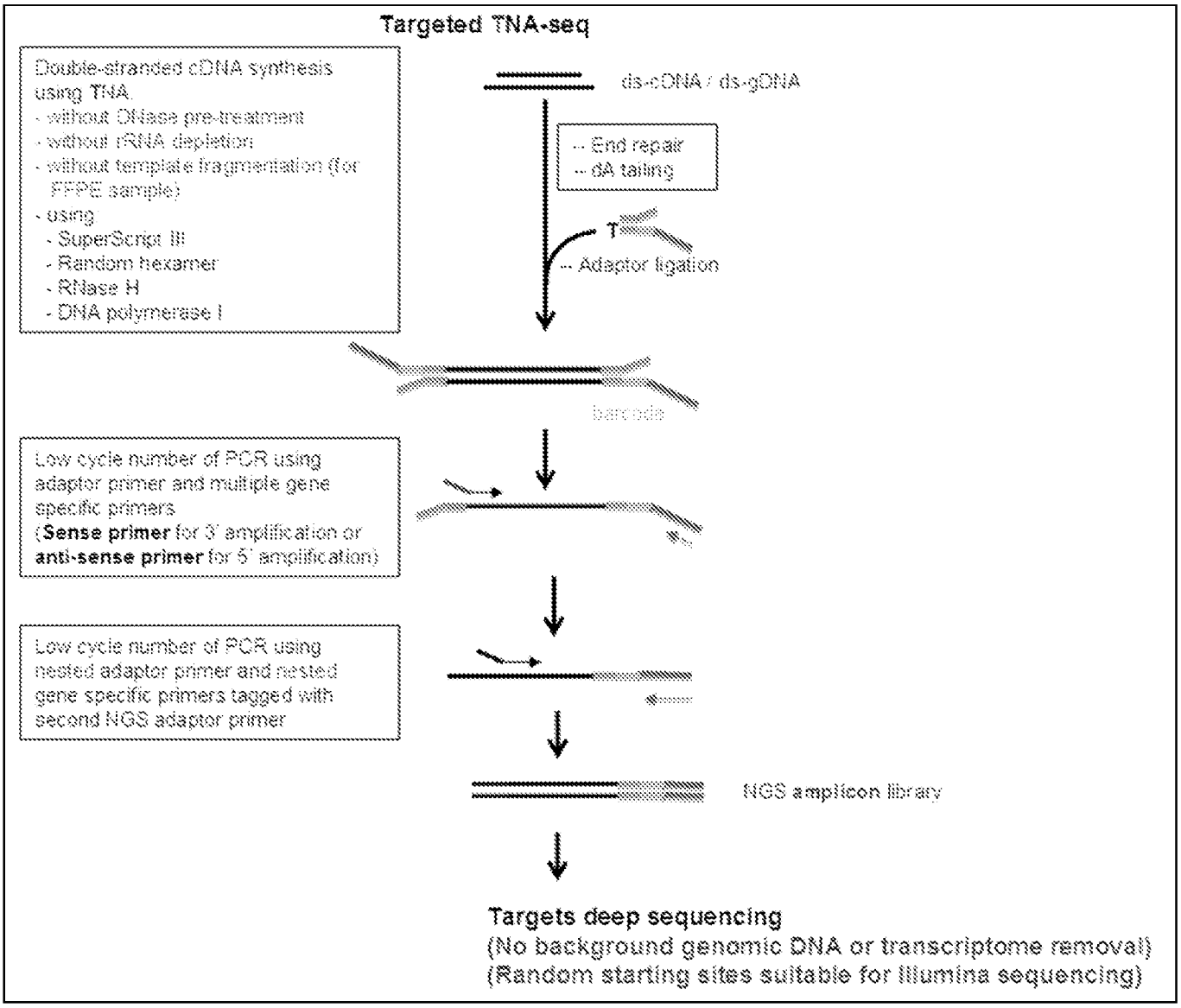

Targeted TNA-seq
Double-stranded cDNA synthesis using TNA
- without DNase pre-treatment
- without rRNA depletion
- without template fragmentation (for FFPE sample)
- using
- SuperScript III
- Random hexamer
- RNase H
- DNA polymerase I
ds-cDNA / ds-gDNA
-- End repair
-- dA tailing
T
-- Adaptor ligation
barcode
Low cycle number of PCR using adaptor primer and multiple gene specific primers
(Sense primer for 3' amplification or anti-sense primer for 5' amplification)
Low cycle number of PCR using nested adaptor primer and nested gene specific primers tagged with second NGS adaptor primer
NGS amplicon library
Targets deep sequencing
(No background genomic DNA or transcriptome removal)
(Random starting sites suitable for Illumina sequencing)

Figure 8

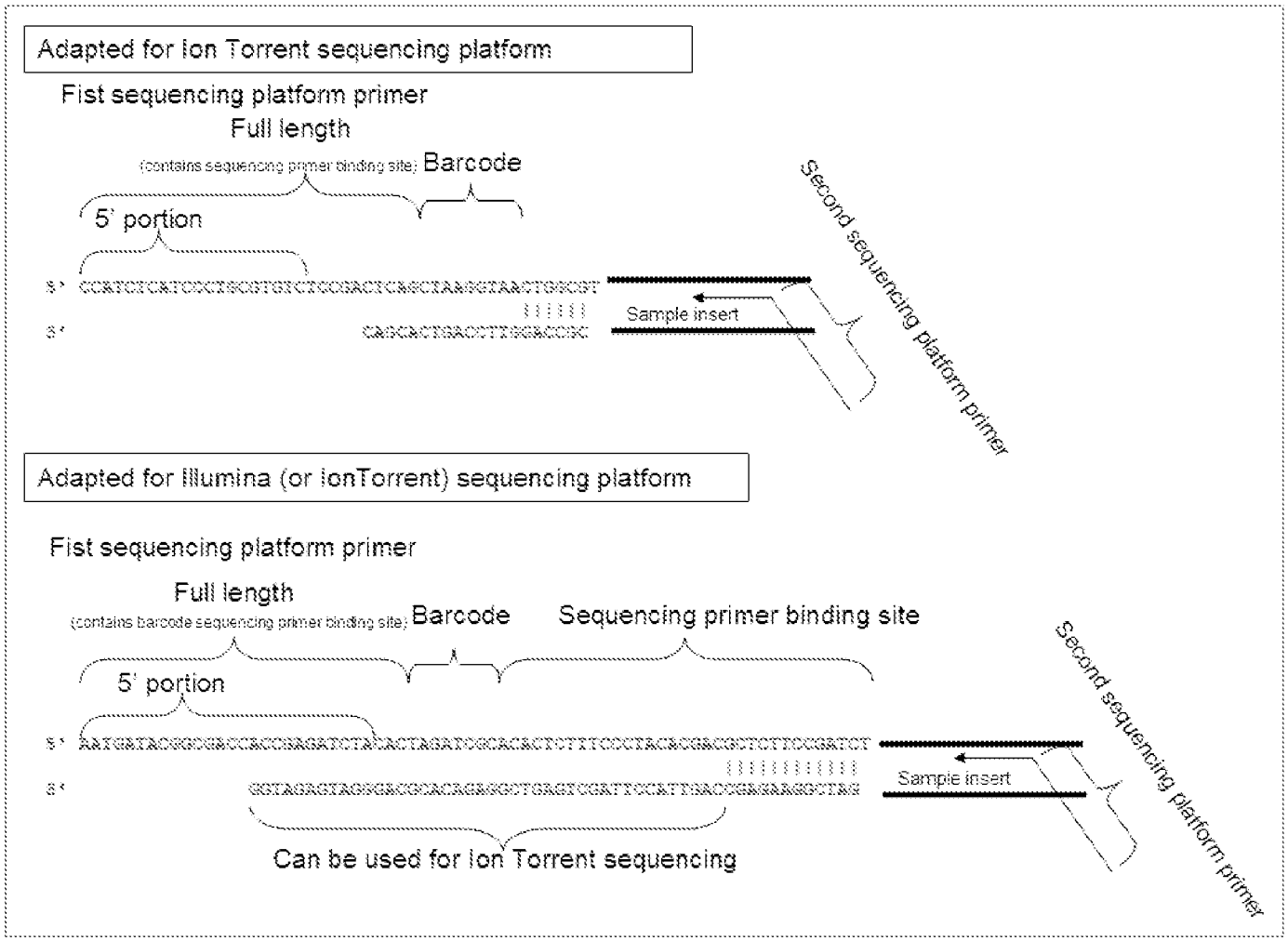

Adapted for Ion Torrent sequencing platform
Fist sequencing platform primer
Full length
(contains sequencing primer binding site)
Barcode
5' portion
Sample insert
Second sequencing platform primer
Adapted for Illumina (or IonTorrent) sequencing platform
Fist sequencing platform primer
Full length
(contains barcode sequencing primer binding site)
Barcode
Sequencing primer binding site
5' portion
Sample insert
Second sequencing platform primer
Can be used for Ion Torrent sequencing

Figure 9