



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 universal primer 1 for downstream 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 rearrangements 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 stuffer 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.

Targeted RNA & DNA sequencing

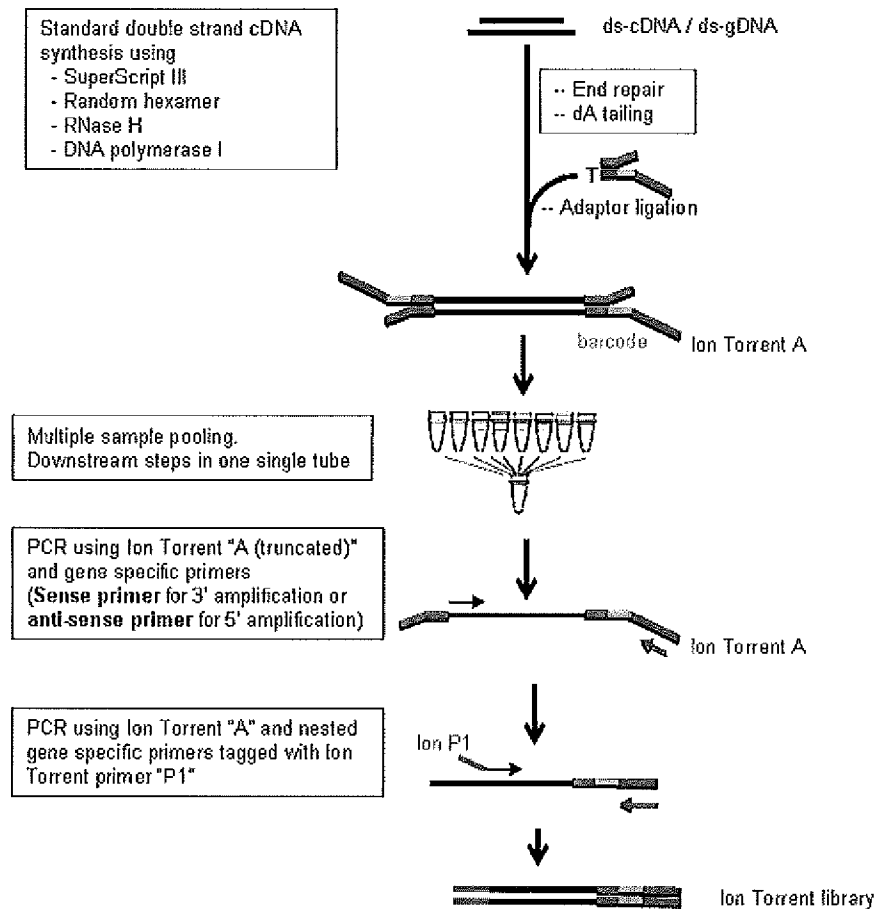


Figure 1 Schematic illustration of library construction for targeted RNA and DNA sequencing. 1) A standard procedure of double-stranded cDNA synthesis is applied using total nucleic acid from FFPE specimen as starting material. Alternatively, the starting material may also be sheared gDNA. 2) After cleanup, the double stranded cDNA or gDNA is subjected to end-repair and dA tailing, directly followed by ligation of a half-truncated Y adaptor 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.

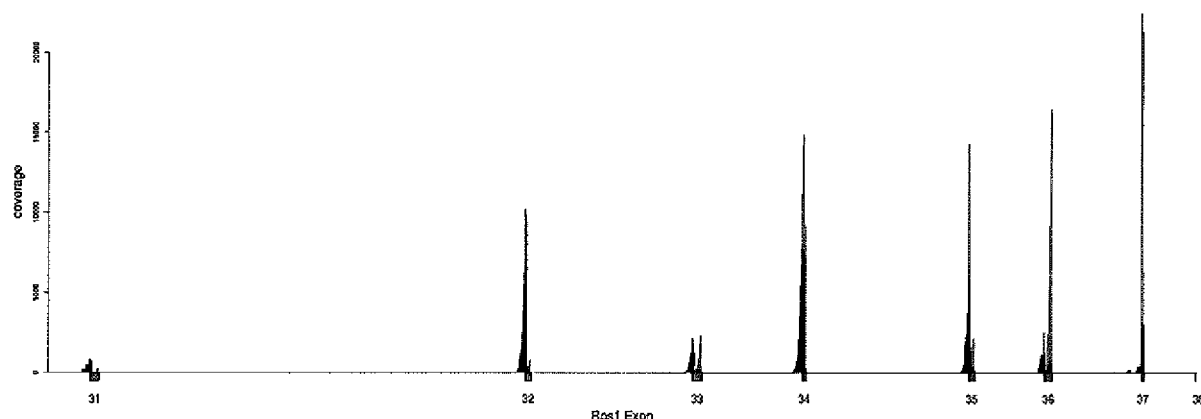


Figure 2. Mapping results showed different primer (green dots) extension products in one sequencing run, including reads that corresponding to genomic DNA (black, spanning intron-exon boundaries), cDNA (red, spanning exons) and fusion cDNA (on exon 34, mapping on the fusion partner SLC34A2 not shown 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 $1:3 \times 10^7$. 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 3×10^7 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

Lung Cancer: Prevalence and Survival

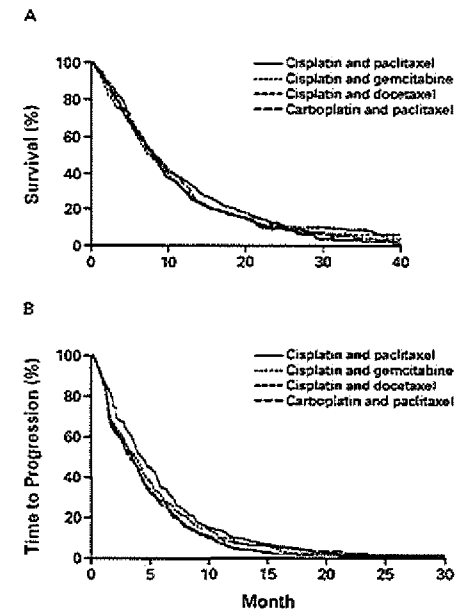
Leading New Cancer Cases and Deaths – 2012 Estimates

Estimated New Cases*	
Male	Female
Prostate 241,740 (29%)	Breast 226,870 (29%)
Lung & bronchus 116,470 (14%)	Lung & bronchus 109,690 (14%)
Colon & rectum 73,420 (9%)	Colon & rectum 70,040 (9%)
Urinary bladder 55,600 (7%)	Uterine corpus 47,130 (6%)
Melanoma of the skin 44,250 (5%)	Thyroid 43,210 (5%)
Kidney & renal pelvis 40,250 (5%)	Melanoma of the skin 32,000 (4%)
Non-Hodgkin lymphoma 38,160 (4%)	Non-Hodgkin lymphoma 31,970 (4%)
Oral cavity & pharynx 28,540 (3%)	Kidney & renal pelvis 24,520 (3%)
Leukemia 26,830 (3%)	Ovary 22,280 (3%)
Pancreas 22,090 (3%)	Pancreas 21,830 (3%)
All sites 848,170 (100%)	All sites 790,740 (100%)

Estimated Deaths	
Male	Female
Lung & bronchus 87,750 (29%)	Lung & bronchus 72,590 (26%)
Prostate 28,170 (9%)	Breast 39,510 (14%)
Colon & rectum 26,470 (9%)	Colon & rectum 25,220 (9%)
Pancreas 18,850 (6%)	Pancreas 18,540 (7%)
Liver & intrahepatic bile duct 13,980 (5%)	Ovary 15,500 (6%)
Leukemia 13,500 (4%)	Leukemia 10,040 (4%)
Esophagus 12,040 (4%)	Non-Hodgkin lymphoma 8,620 (3%)
Urinary bladder 10,510 (3%)	Uterine corpus 8,010 (3%)
Non-Hodgkin lymphoma 10,320 (3%)	Liver & intrahepatic bile duct 6,570 (2%)
Kidney & renal pelvis 8,650 (3%)	Brain & other nervous system 5,980 (2%)
All sites 301,820 (100%)	All sites 275,370 (100%)

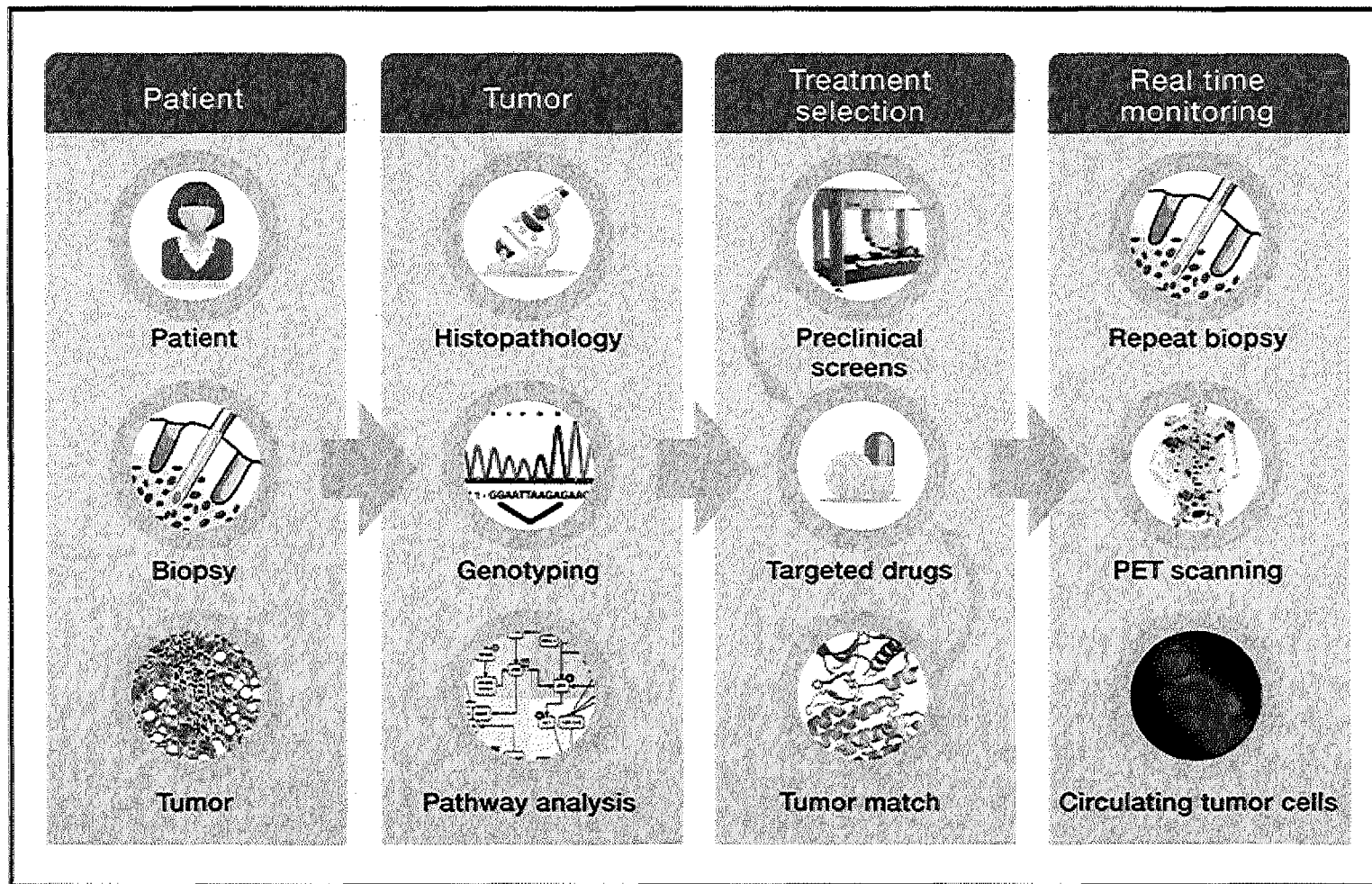
*Excludes basal and squamous cell skin cancers and in situ carcinoma except urinary bladder.

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



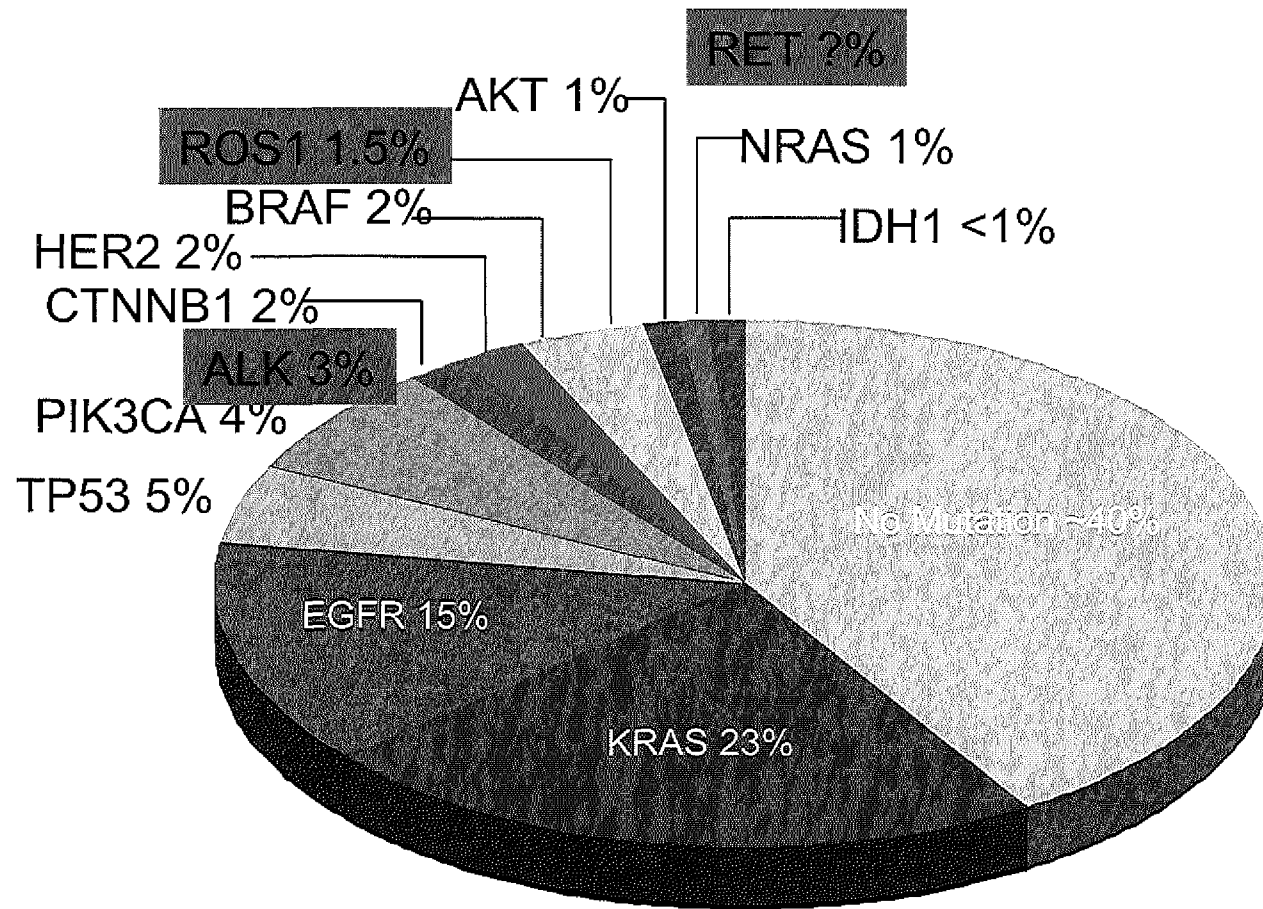
Comparison of Four Chemotherapy Regimens for Advanced Non-Small-Cell Lung Cancer
Schiller et al., NEJM 2002

A New Paradigm in Cancer Treatment



Haber, Gray, Baselga Cell 2011

Mutational profiling in lung cancers

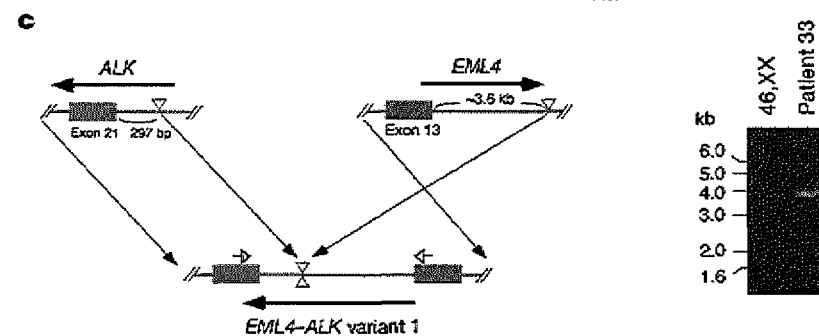
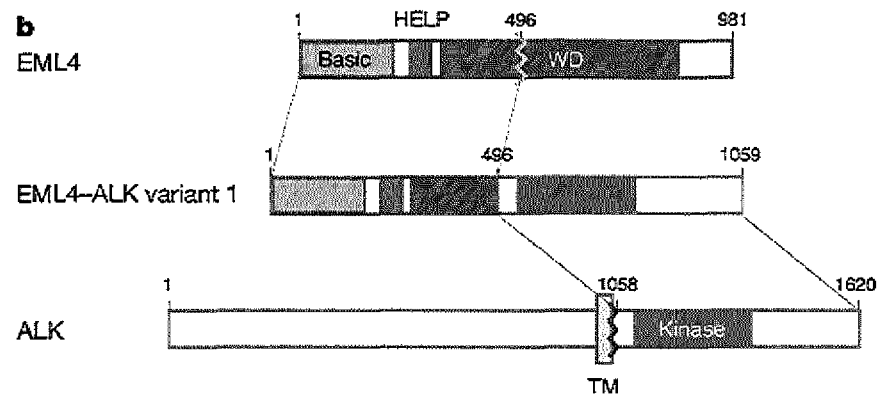


ARTICLES

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

Manabu Soda^{1,2}, Young Lim Choi¹, Munehiro Enomoto^{1,2}, Shuji Takada¹, Yoshihiro Yamashita¹, Shunpei Ishikawa⁵, Shin-ichiro Fujiwara¹, Hideki Watanabe¹, Kentaro Kurashina¹, Hisashi Hatanaka¹, Masashi Bando², Shoji Ohno², Yuichi Ishikawa⁶, Hiroyuki Aburatani^{5,7}, Toshiro Niki³, Yasunori Sohara⁴, Yukihiko Sugiyama² & Hiroyuki Mano^{1,7}

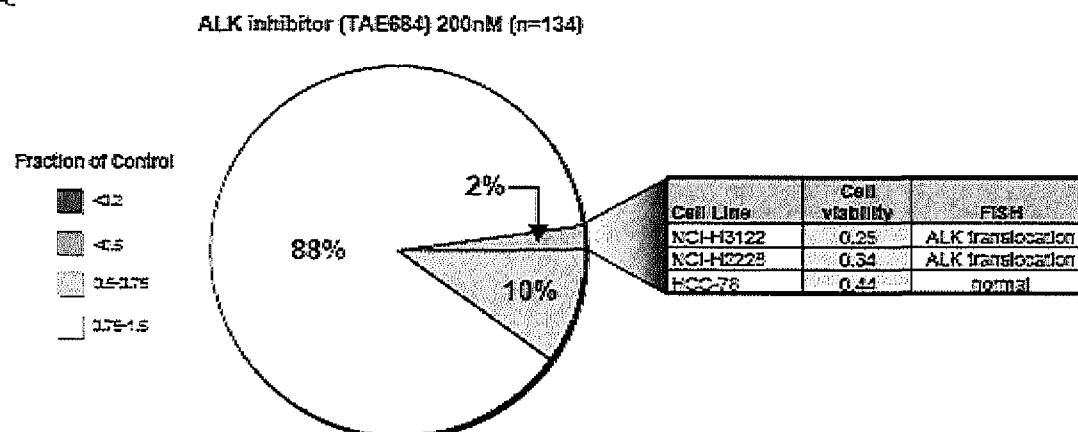
Improvement in the clinical outcome of lung cancer is likely to be achieved by identification of the molecular events that underlie its pathogenesis. Here we show that a small inversion within chromosome 2p results in the formation of a fusion gene comprising portions of the echinoderm microtubule-associated 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.



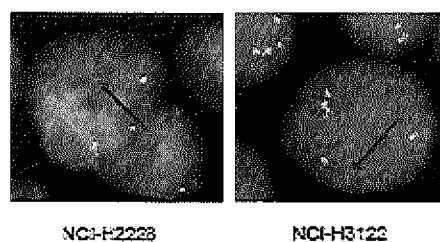
	Vector	EML4	ALK	EML4-ALK	K589M	NPM-ALK	v-Ras
3T3							
Nude mice							
Tumour/injection	0/8	0/8	0/8	8/8	0/8	8/8	2/2

EML4-ALK+ Cell Lines Respond to ALK Inhibitor

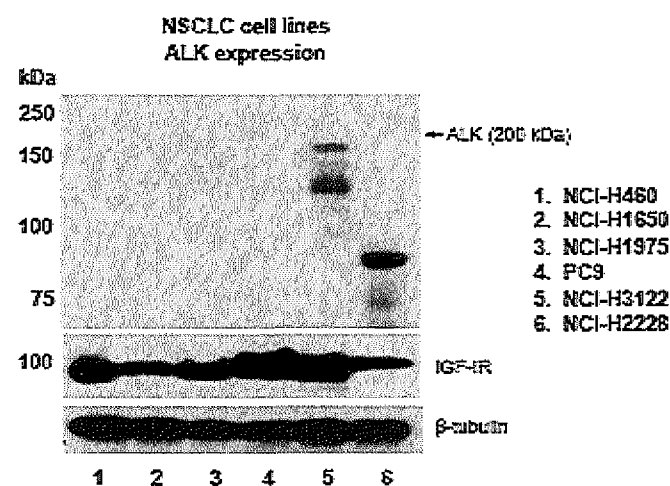
A.



B.



C.



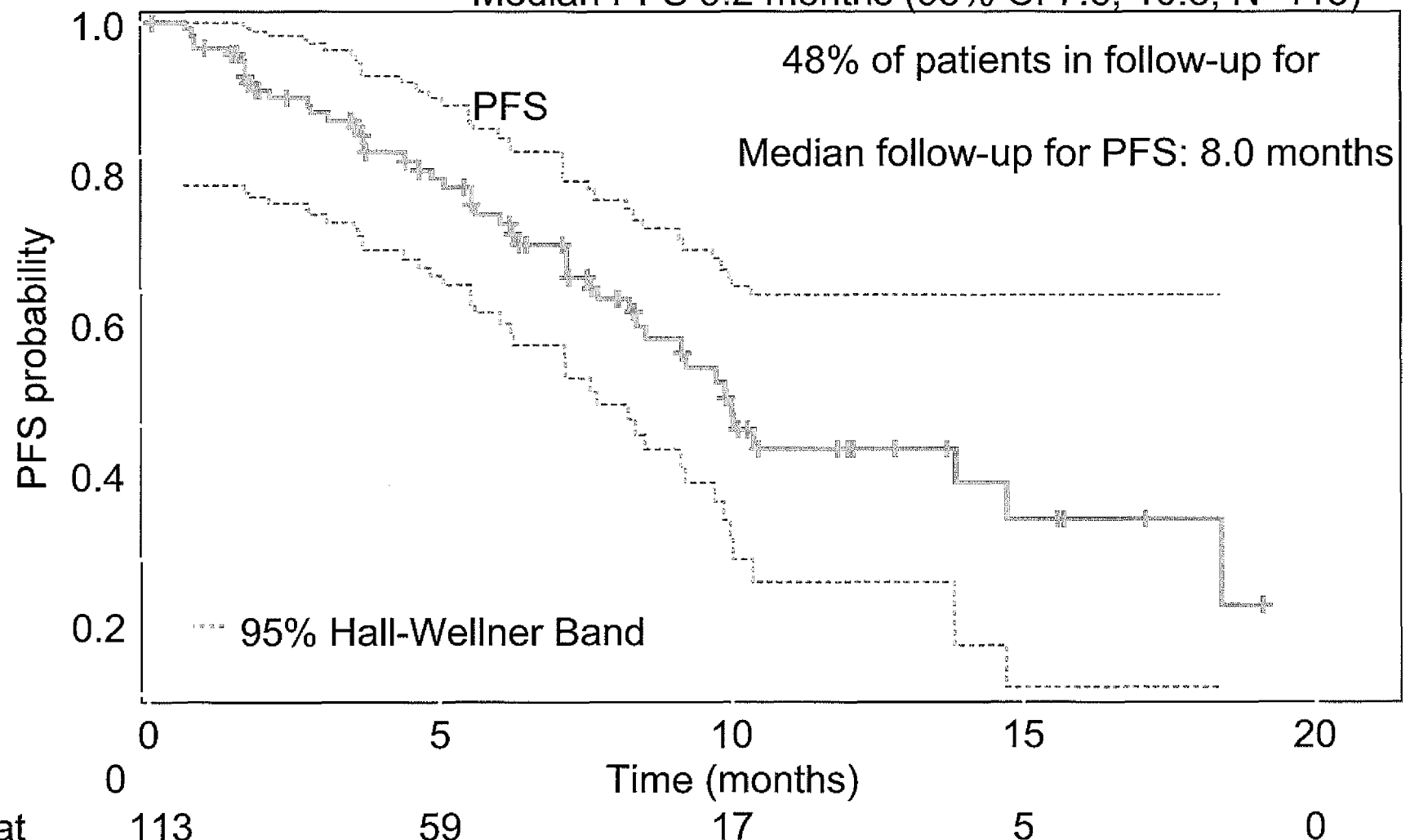
D.

NCI-H2228

NCI-H3122

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

Median PFS 9.2 months (95% CI 7.6, 10.3; N=113)



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

ROS1 rearrangement in NSCLC



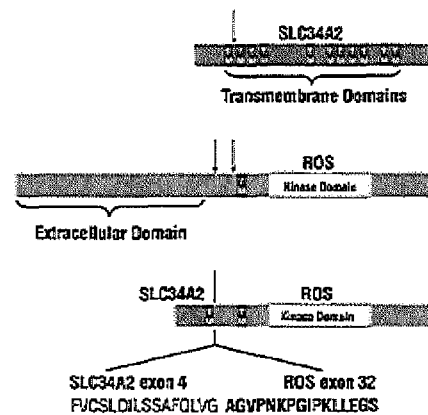
Resource

Global Survey of Phosphotyrosine Signaling Identifies Oncogenic Kinases in Lung Cancer

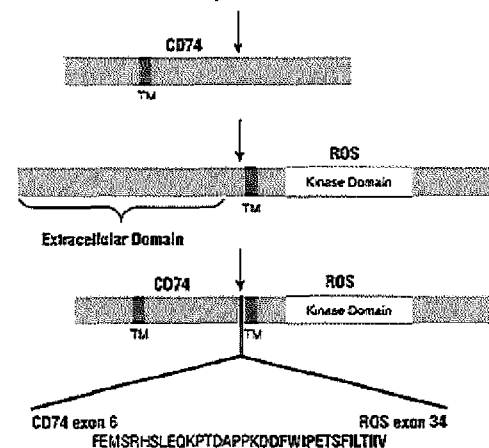
Klarisa Rikova,¹ Allan Guo,¹ Qingfu Zeng,¹ Anthony Possemato,¹ Jian Yu,¹ Herbert Haack,¹ Julie Nardone,¹ Kimberly Lee,¹ Cynthia Reeves,¹ Yu Li,¹ Yierong Hu,² Zhiping Tan,³ Matthew Stokes,¹ Laura Sullivan,¹ Jeffrey Mitchell,¹ Randy Wotzel,¹ Joan MacNeill,¹ Jian Min Ren,¹ Jin Yuan,¹ Corey E. Bakalarski,² Judith Vilian,² Jon M. Kornhauser,¹ Bradley Smith,¹ Daigang Li,⁴ Xinmin Zhou,² Steven P. Gygi,² Ting-Lai Gu,¹ Roberto D. Polakiewicz,¹ John Rush,¹ and Michael J. Comb^{1,2}

¹Cell Signaling Technology, 3 Trask Lane, Danvers, MA 01923, USA
²Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA
³Department of Cardiothoracic Surgery
⁴Department 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

SLC34A2-ROS Fusion in HCC78 cell line



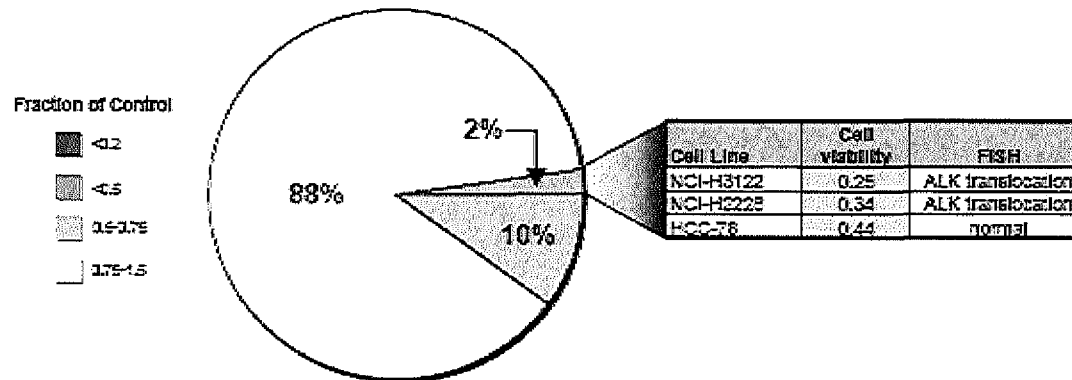
CD74-ROS Fusion in patient CS042



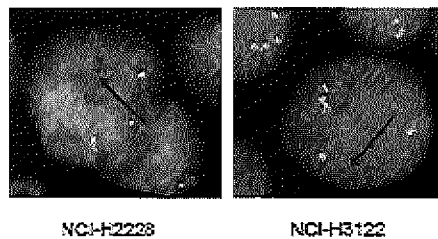
ROS1 rearrangement in NSCLC

A.

ALK inhibitor (TAE684) 200nM (n=134)

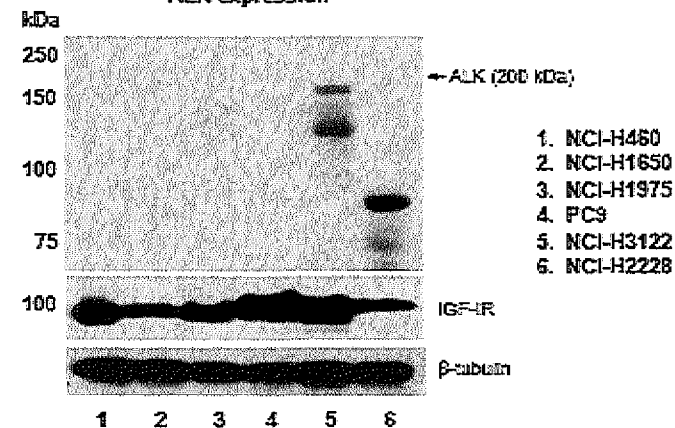


B.



C.

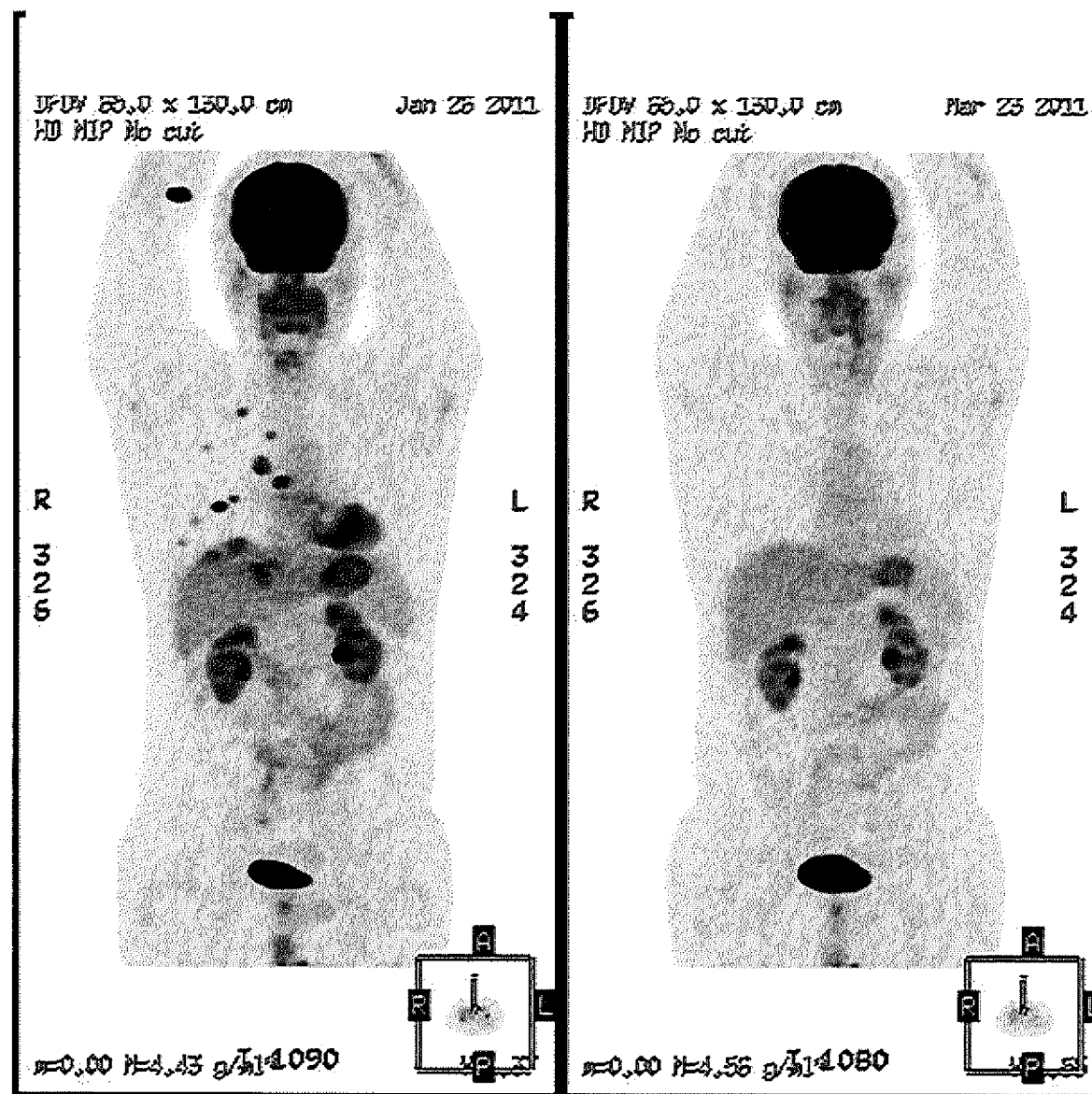
NSCLC cell lines
ALK expression



D.

NCI-H2228 NCI-H3122

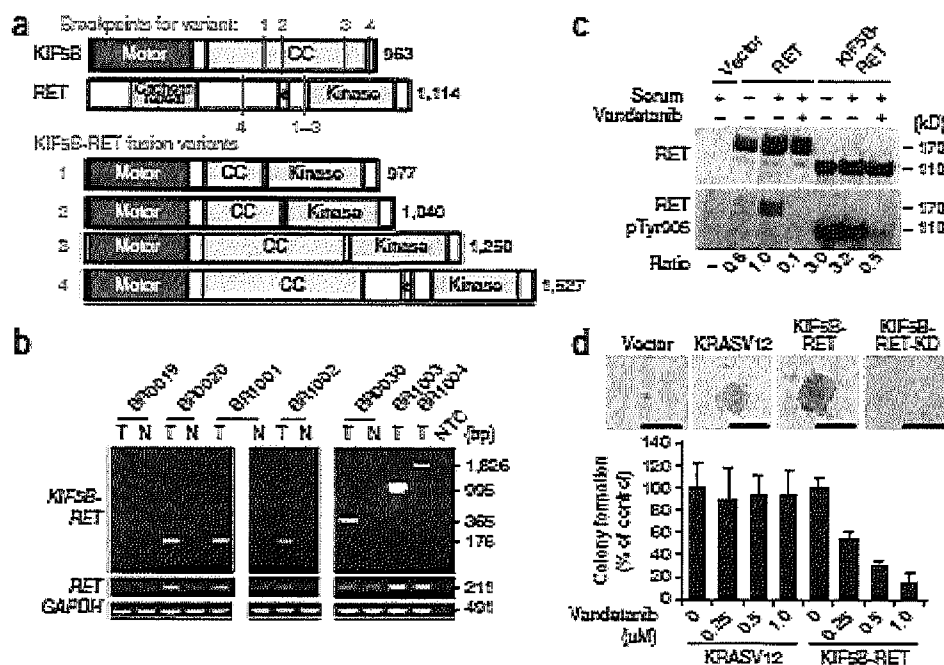
ROS1-positive patients respond to crizotinib



KIF5B-RET fusions in lung adenocarcinoma

Takashi Kohno^{1,15}, Hitoshi Ichikawa^{2,15}, Yasushi Totoki³, Kazuki Yasuda⁴, Masaki Hiramoto⁴, Takao Nanno⁴, Hiromi Sakamoto², Koji Tsuta², Koh Furuta², Yoko Shimada¹, Reika Iwakawa⁶, Hideaki Ogiwara¹, Takahiro Oike⁶, Masato Enari², Aaron J Schetter⁸, Hirokazu Okayama^{6,9}, Aage Haugen⁹, Vidar Skang⁹, Suenori Chiku¹⁰, Itaru Yamanaka¹¹, Yasuhito Arai³, Shun-ichi Watanabe¹², Isao Sekine¹³, Seishi Ogasawa¹⁴, Curtis C Harris⁸, Hitoshi Tsuda⁵, Teruhiko Yoshida², Jun Yokota⁶ & Tatsuhiko Shibata³

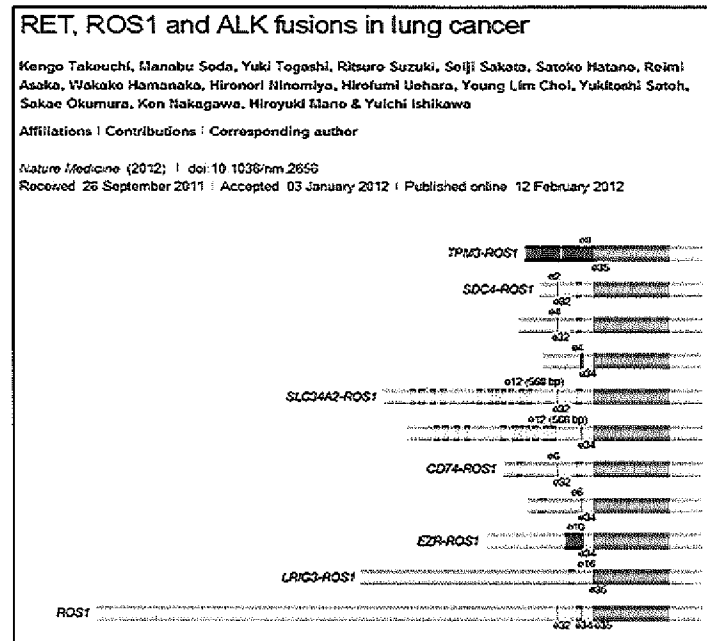
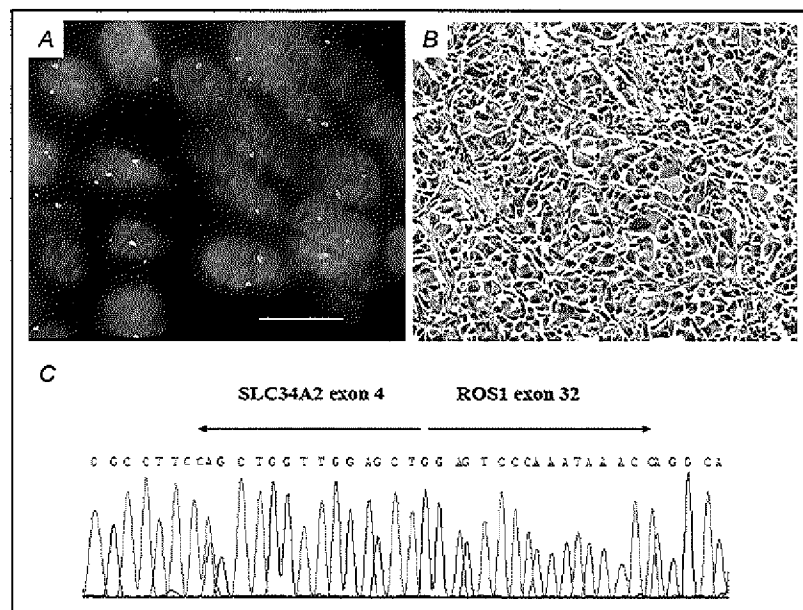
We identified in-frame fusion transcripts of *KIF5B* (the kinesin family 5B gene) and the *RET* oncogene, which are present in 1–2% of lung adenocarcinomas (LADCs) from people from Japan and the United States, using whole-transcriptome sequencing. The *KIF5B-RET* fusion leads to aberrant activation of 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.



ALK, ROS1 Detection

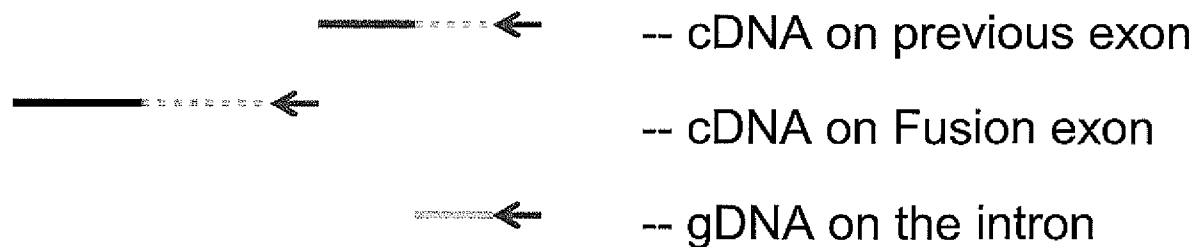
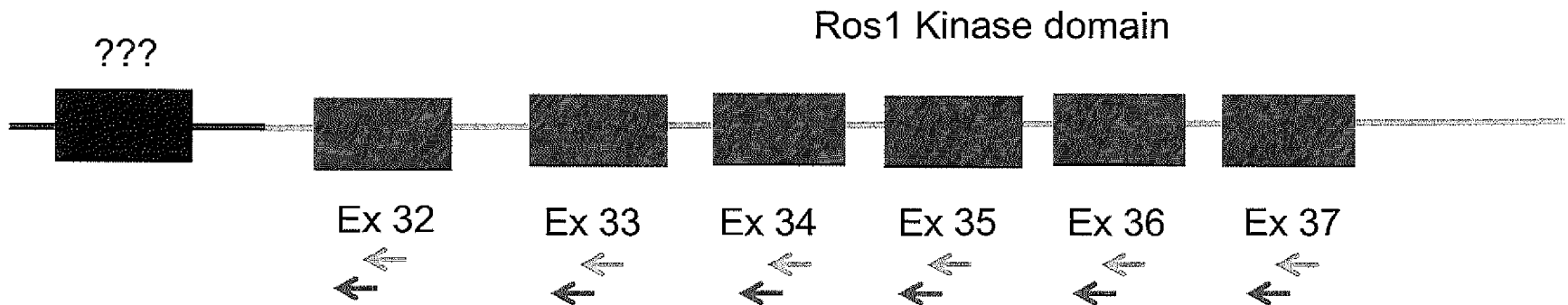
Concerns

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



A next-generation sequencing based assay

Primer design



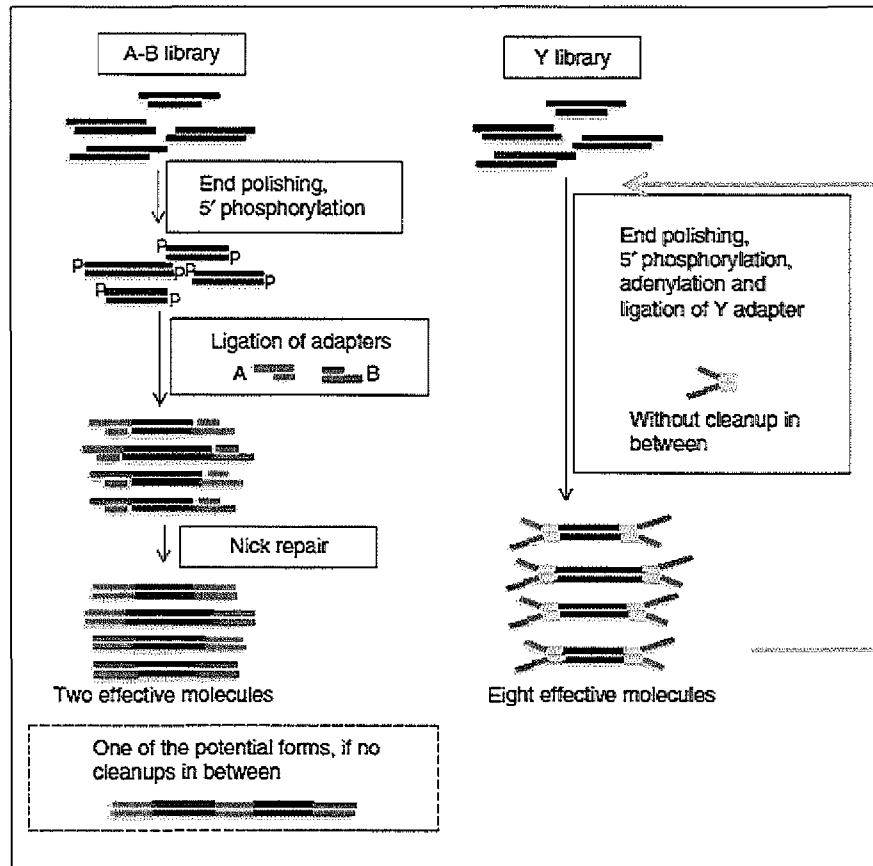
Possible types of extension products

gDNA & RNA from FFPE

Standard double strand cDNA synthesis using

- SuperScript III
- Random hexamer
- RNase H
- DNA polymerase I

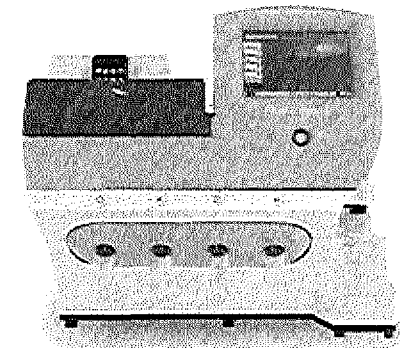
gDNA / cDNA



Target enrichment for
ALK, ROS1, RET

2 rounds of PCR

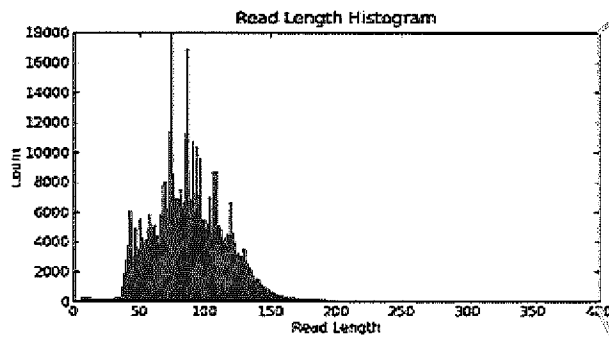
Ion
Torrent
library



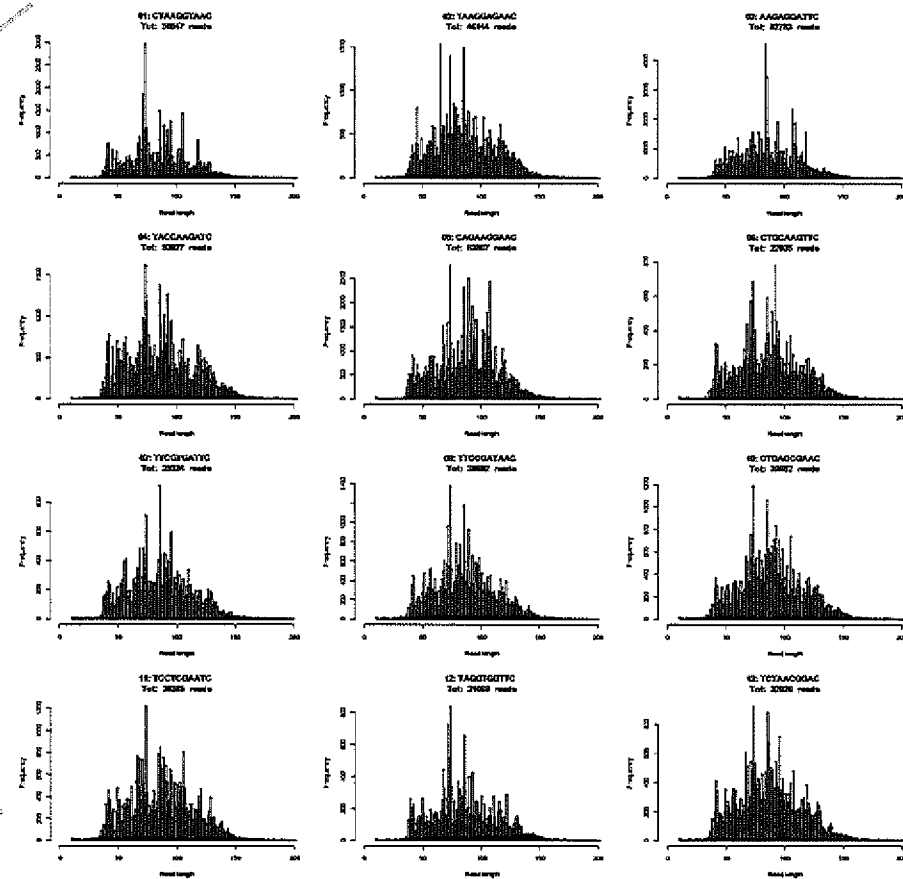
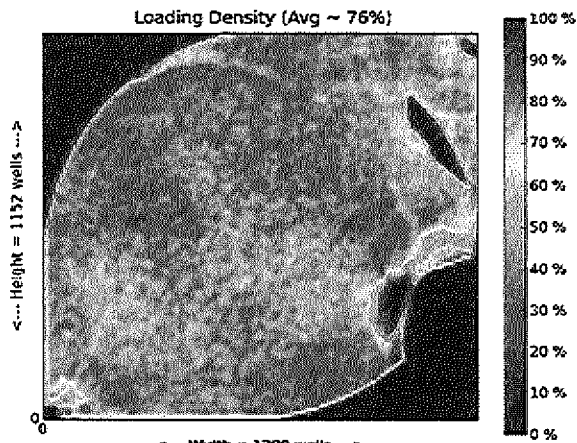
Ion PGM™ Sequencer

Sequencing results

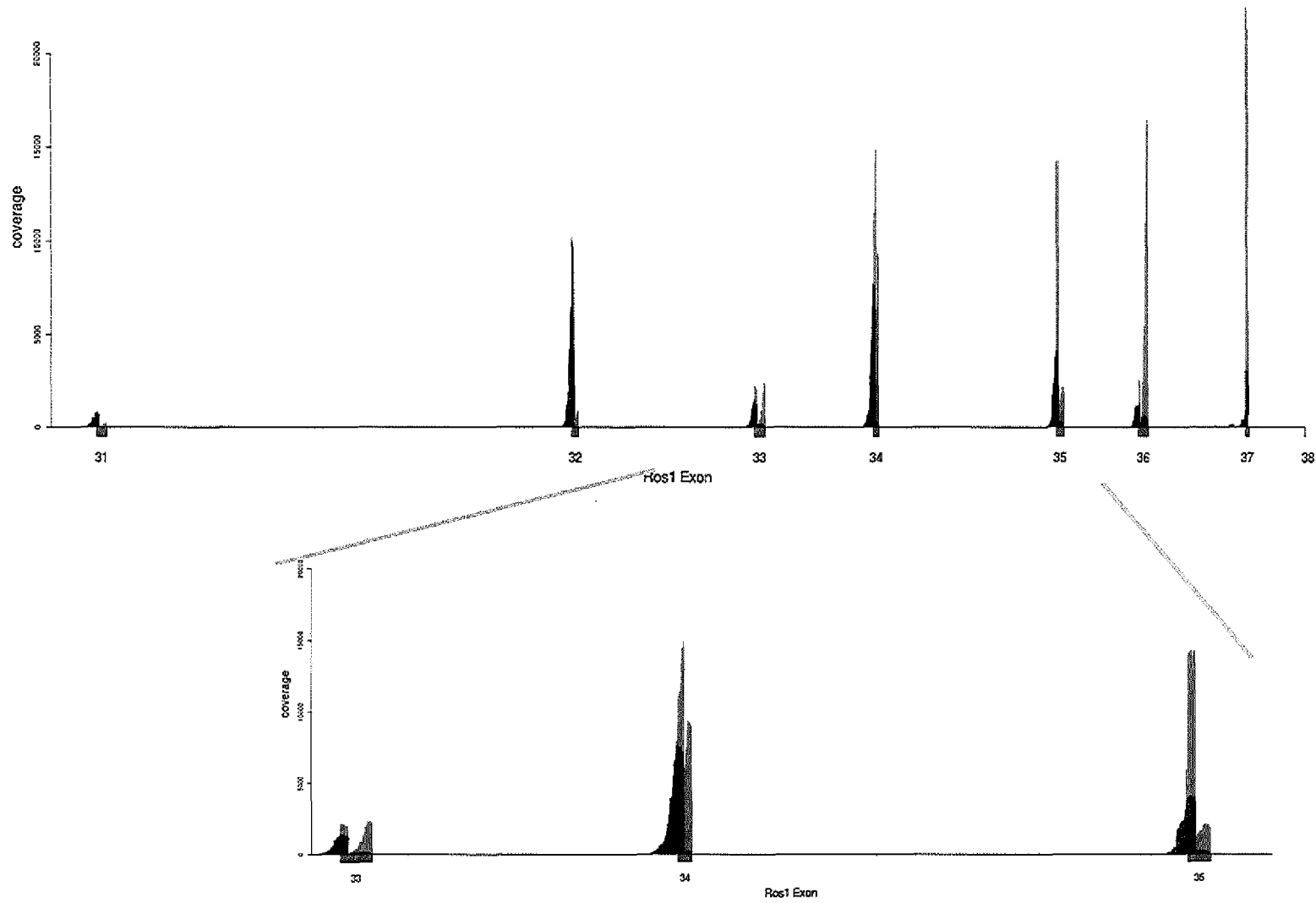
Total Number of Bases [Mbp]	49.19
• Number of Q17 Bases [Mbp]	43.86
• Number of Q20 Bases [Mbp]	41.58
Total Number of Reads	563,255
Mean Length [bp]	87
Longest Read [bp]	203

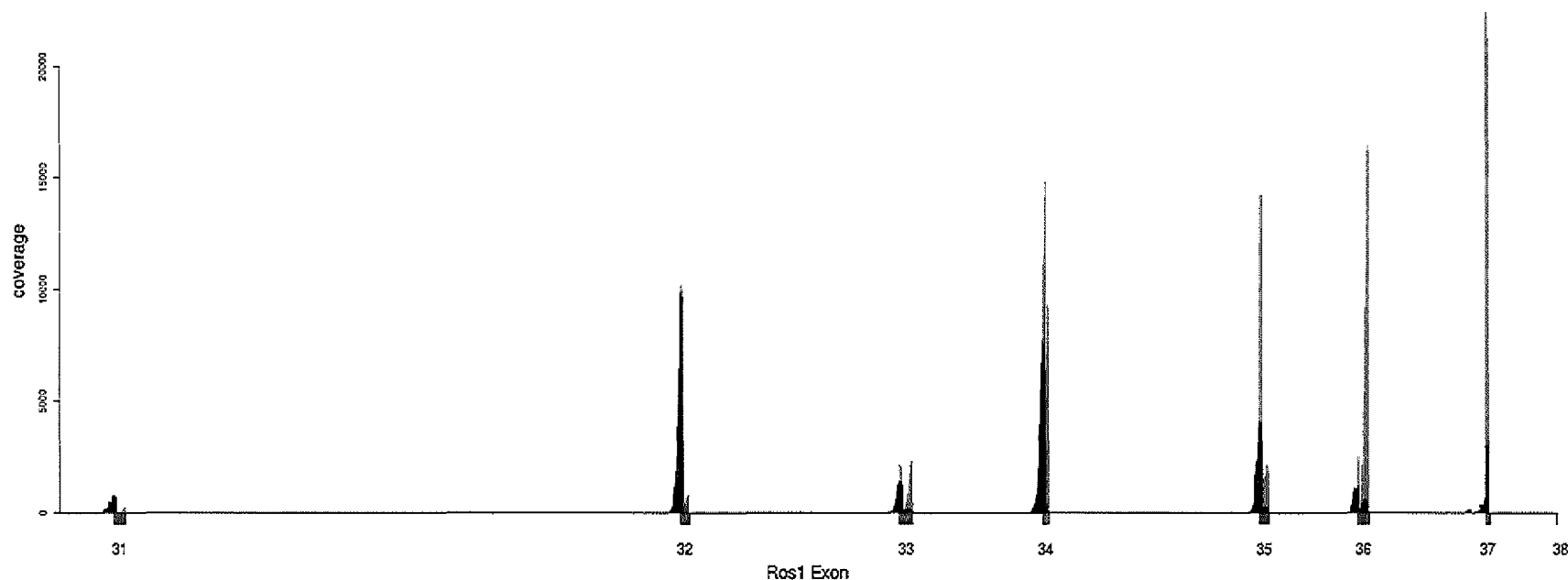


De-Barcode



One example - Coverage on ROS1





SLC34A2 Exon 13 ROS1 Alternative splicing

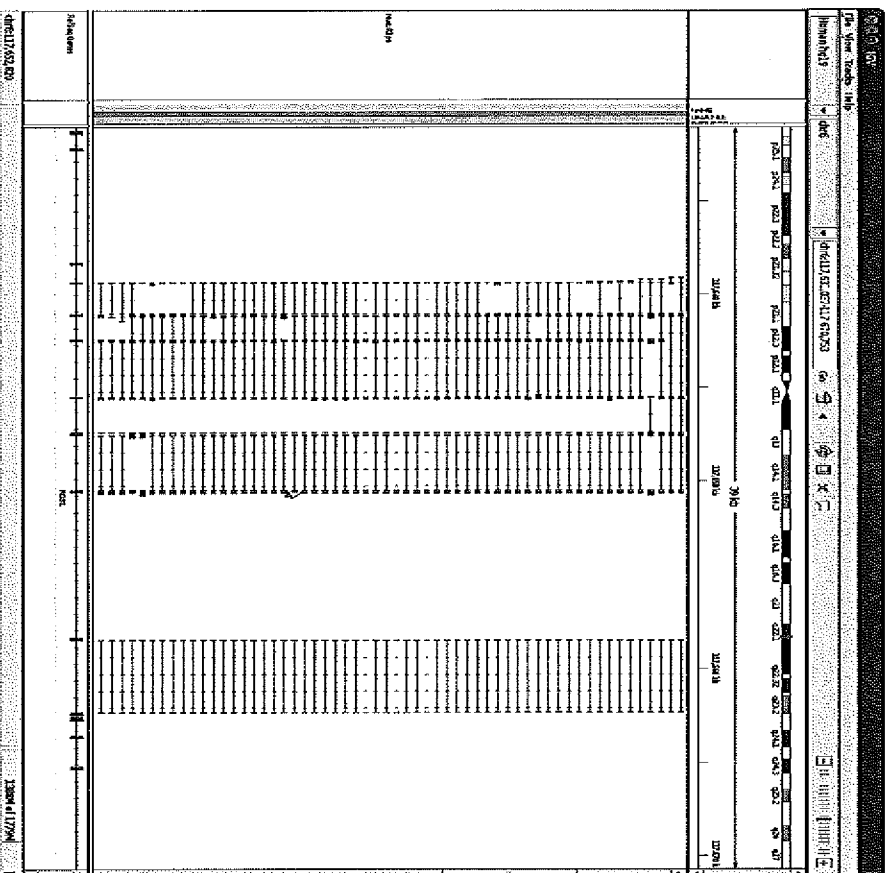
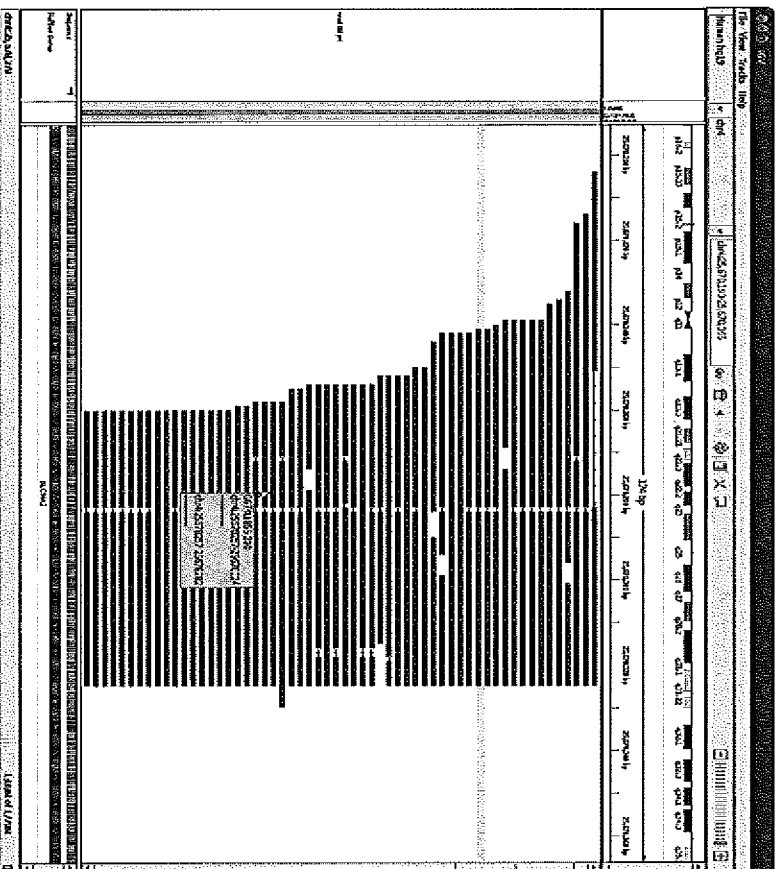
AGGATGTCCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAATAAACACAGGCATTCCCAATTACTAGA TCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAATCAACACAGGCATTCCCAATTACTAGA GAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAATAAACACAGGCATTCCCAATTACTAGA GTCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGCTGGAGTCCCAATAAACACAGGCATTCCCAATTACTAGA GCAGGATGTCCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGATGATTTTTGGATACCAGAAACAGTTTCATACTTACTA GCGCAGGAGGGGACAGGATGTCCTGTCAAGGCTCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGATGATTTTTGGATACCAGAAACAGTTTCATACTTACTA	} }	Exon 32 Exon 34
---	--	--

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

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

Mapping visualization in Integrated Genome Viewer

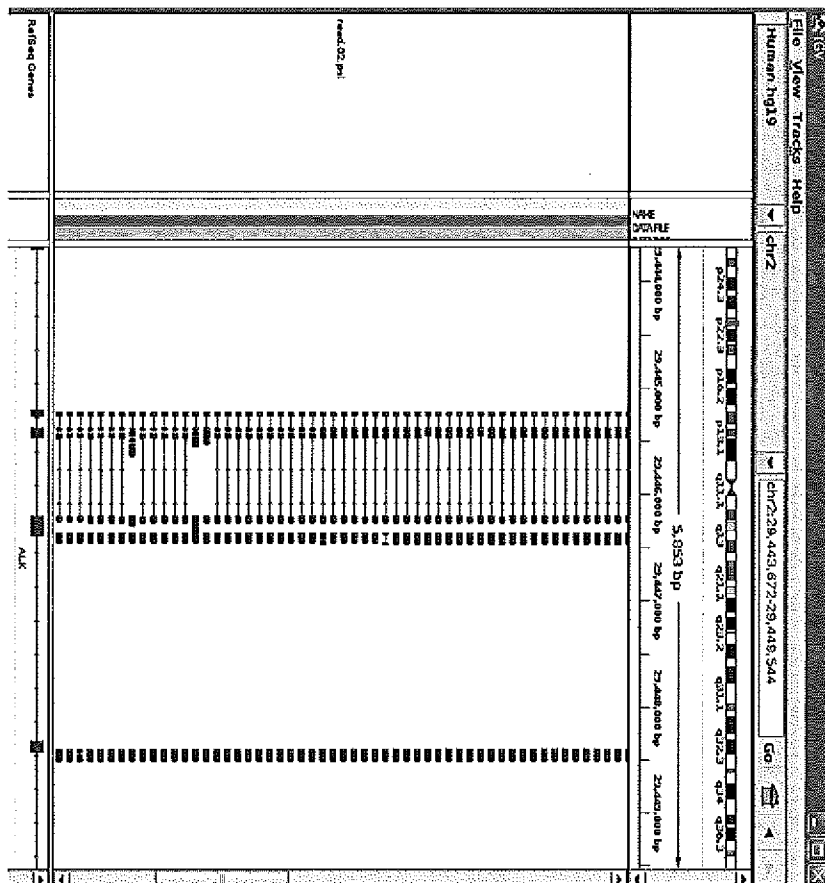
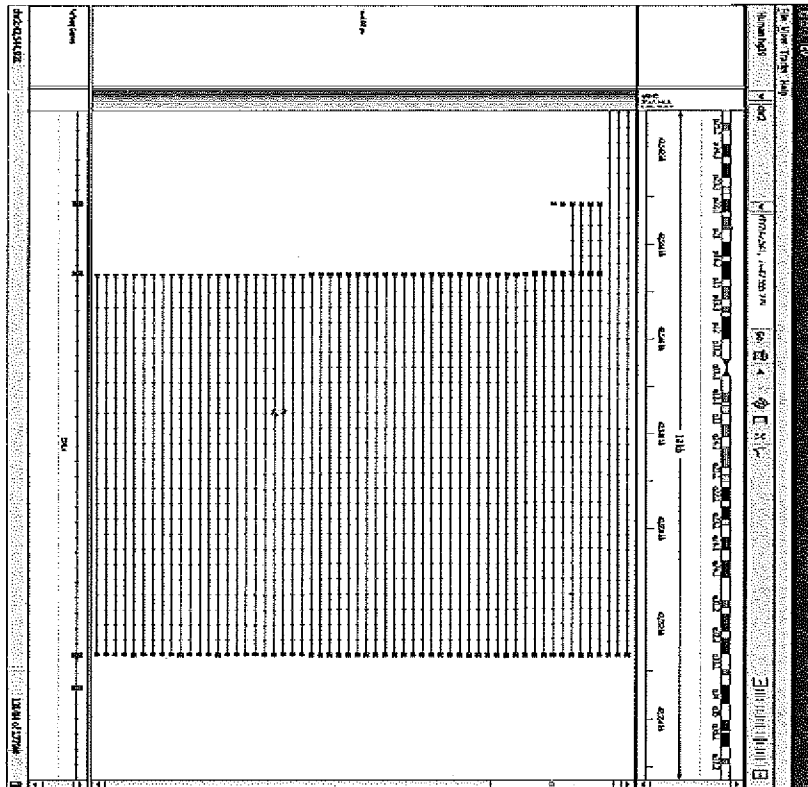
SLC34A2—ROS1



IDT Exhibit 1048.0026

[illegible]

EML4---ALK



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

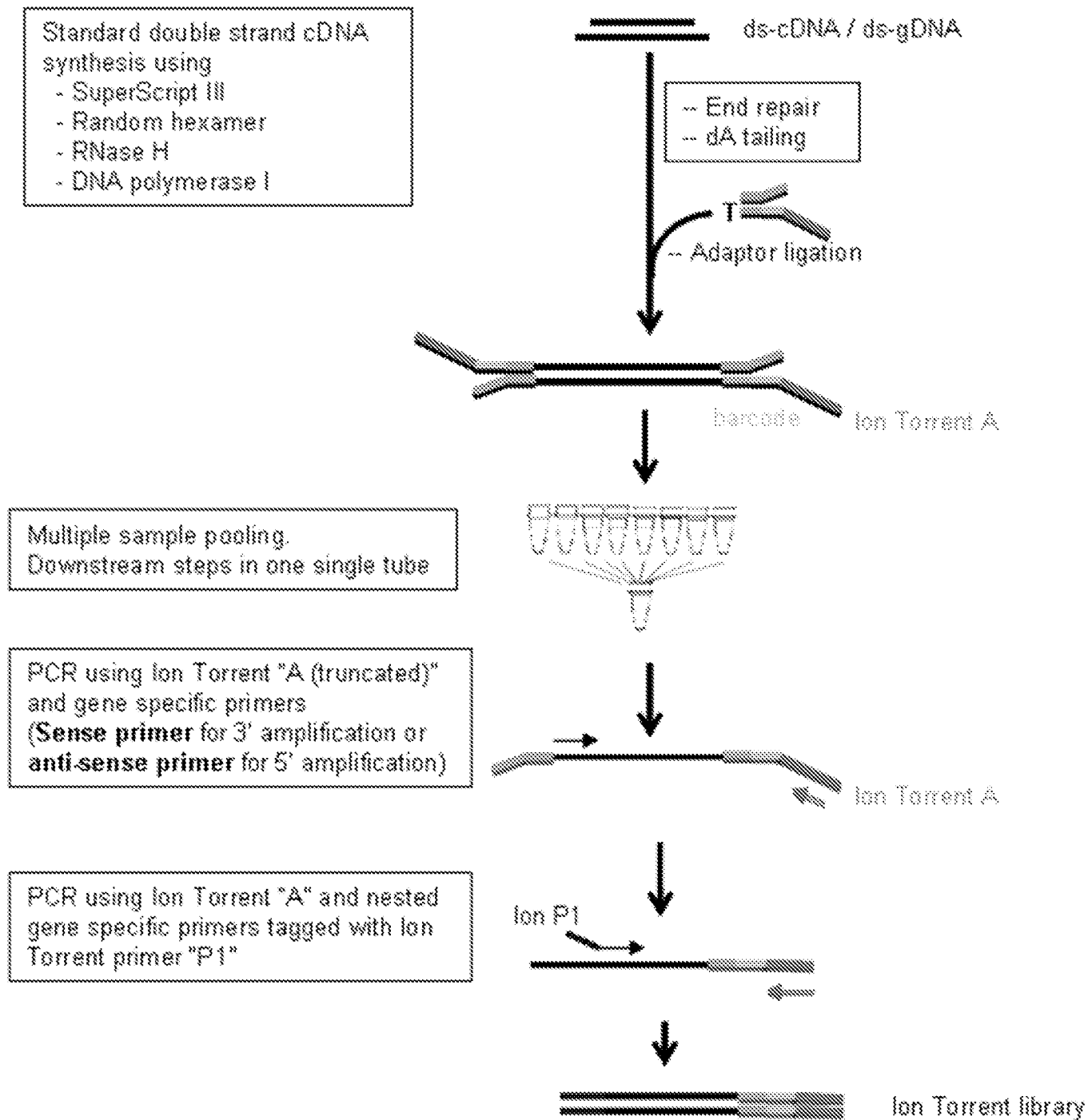


Figure 1

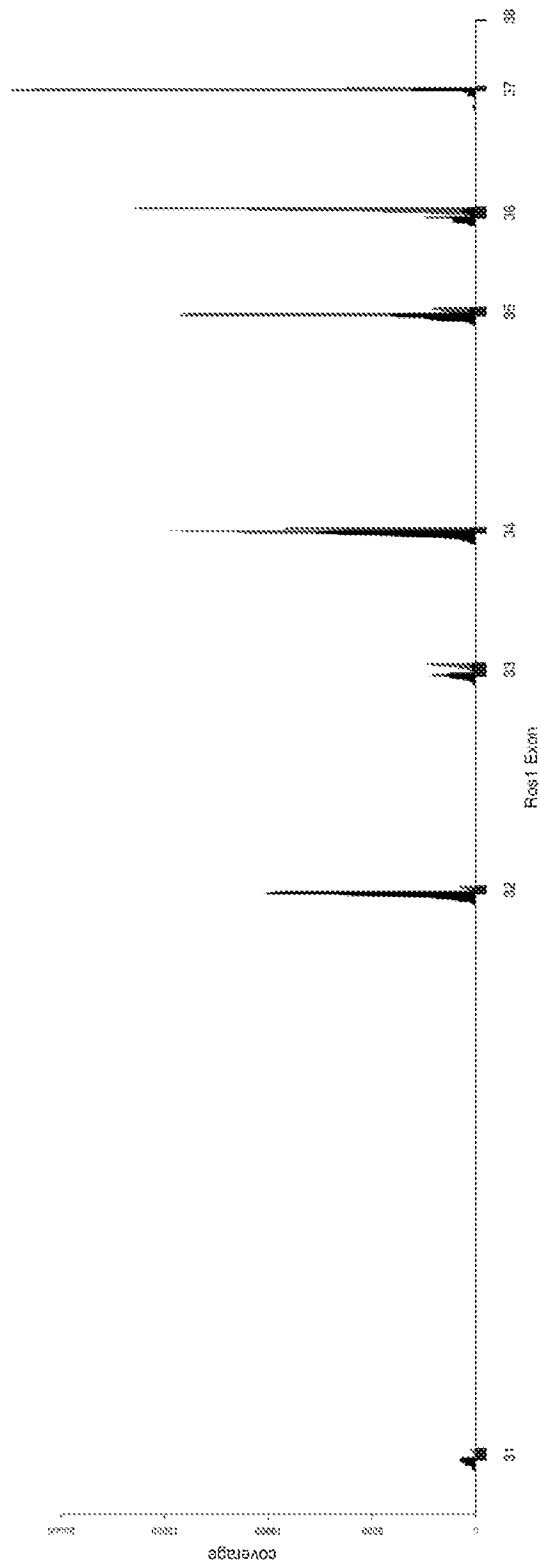
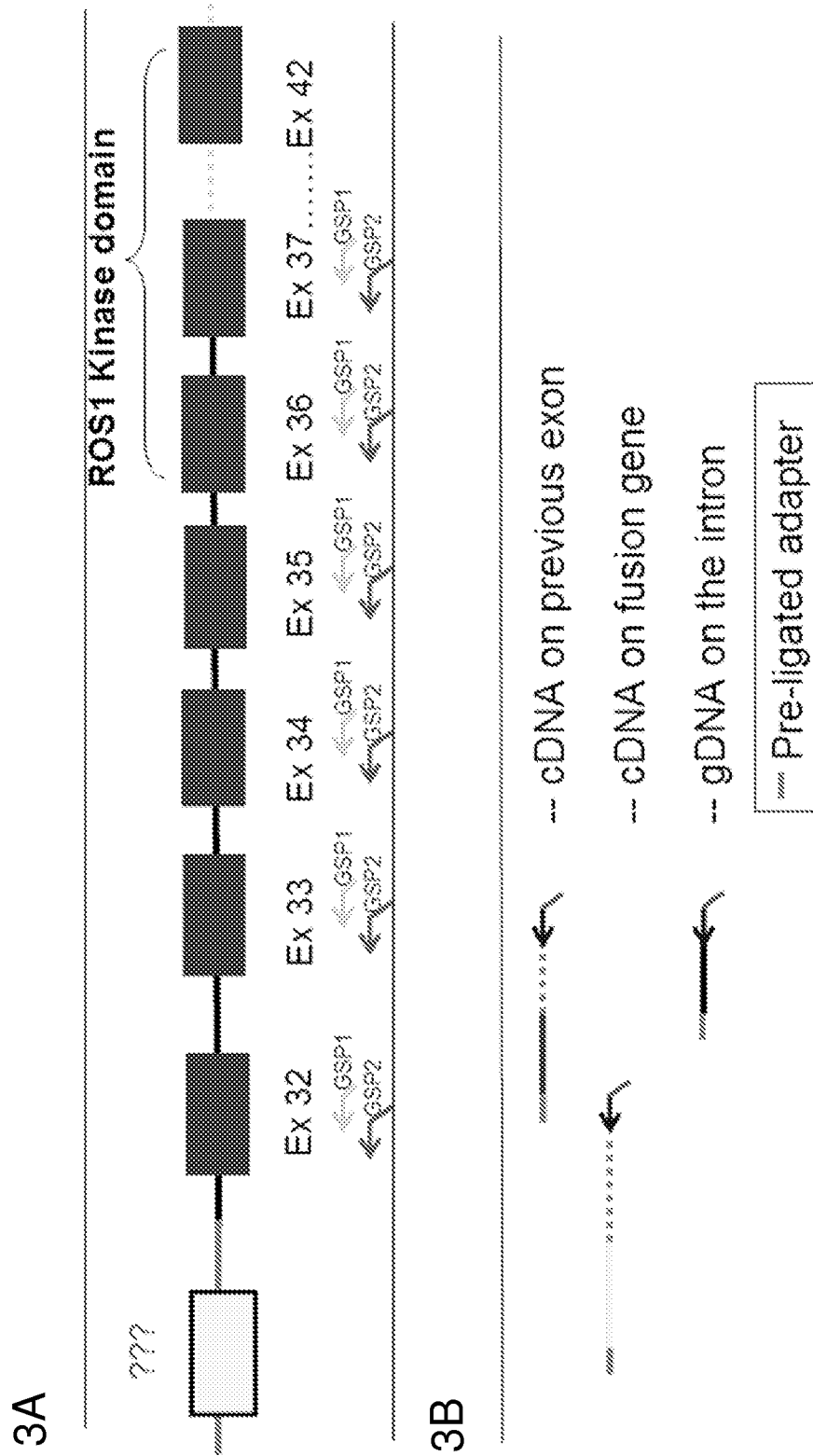
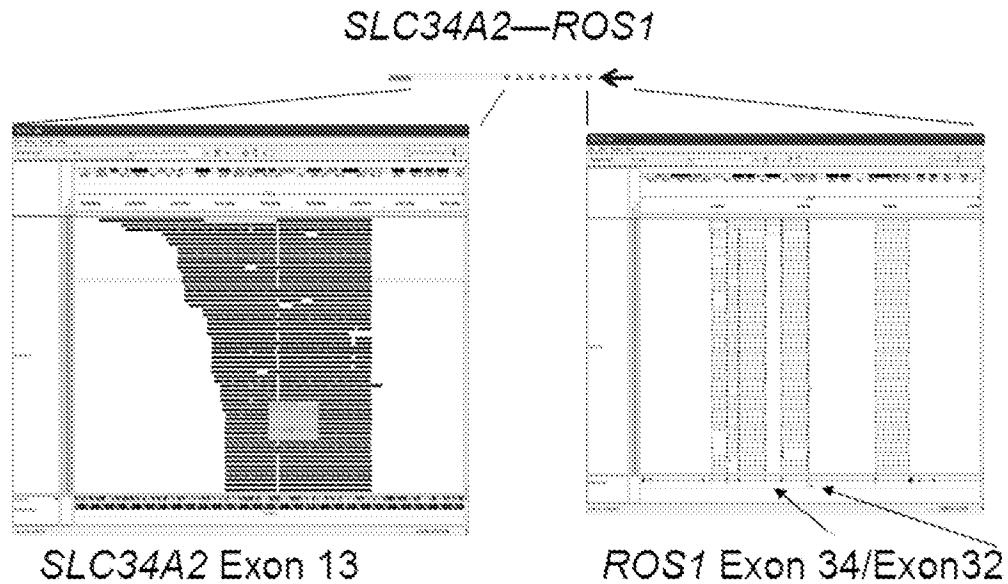


Figure 2



Figures 3A-3B

4A



4B

SLC34A2 Exon 13 ROS1 Alternative splicing

Exon 32

Exon 34

4C

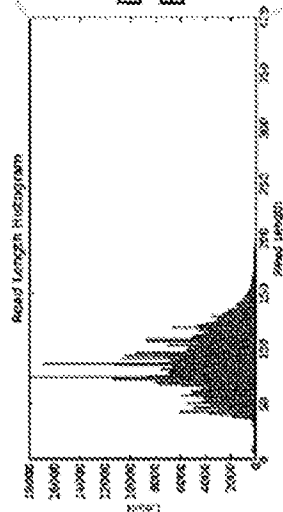
```
> Sample_03
Freq Fusion
168 C1C2A2A+Em13_12--+R031-Em32_11
103 C1C2A2A+Em13_12--+R031-Em34_11

> Sample_15
Freq Fusion
41 E2B-Em10_12--+R031-Em24_11
8 E2B-Em10_12--+R031-Em33_11
```

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

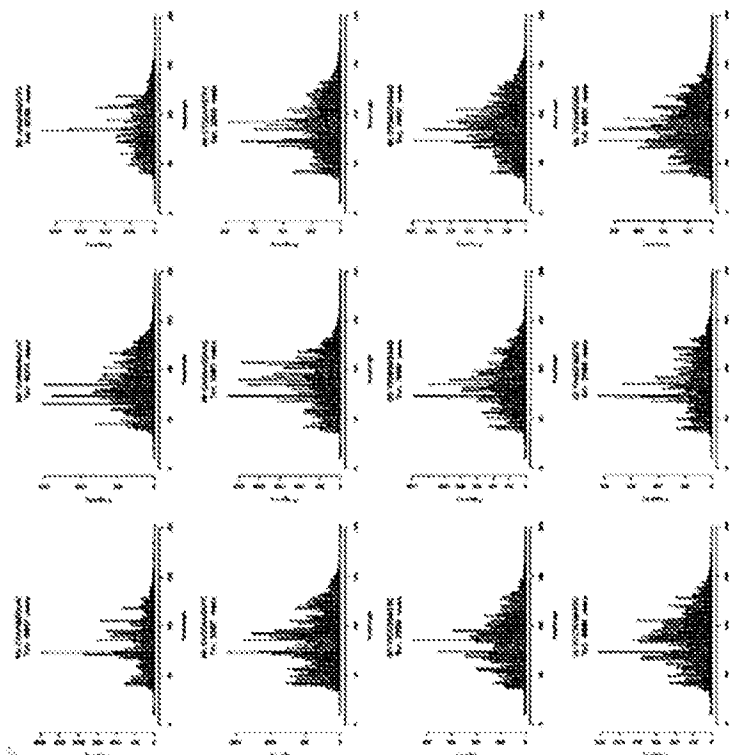
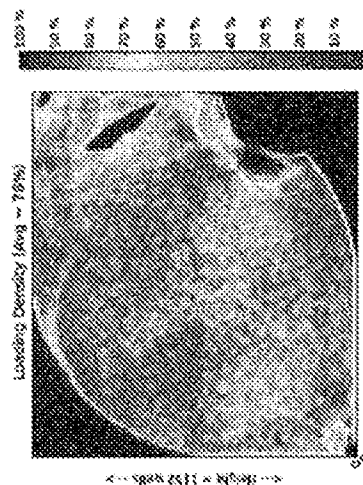


Figure 5

CCDC6--RET

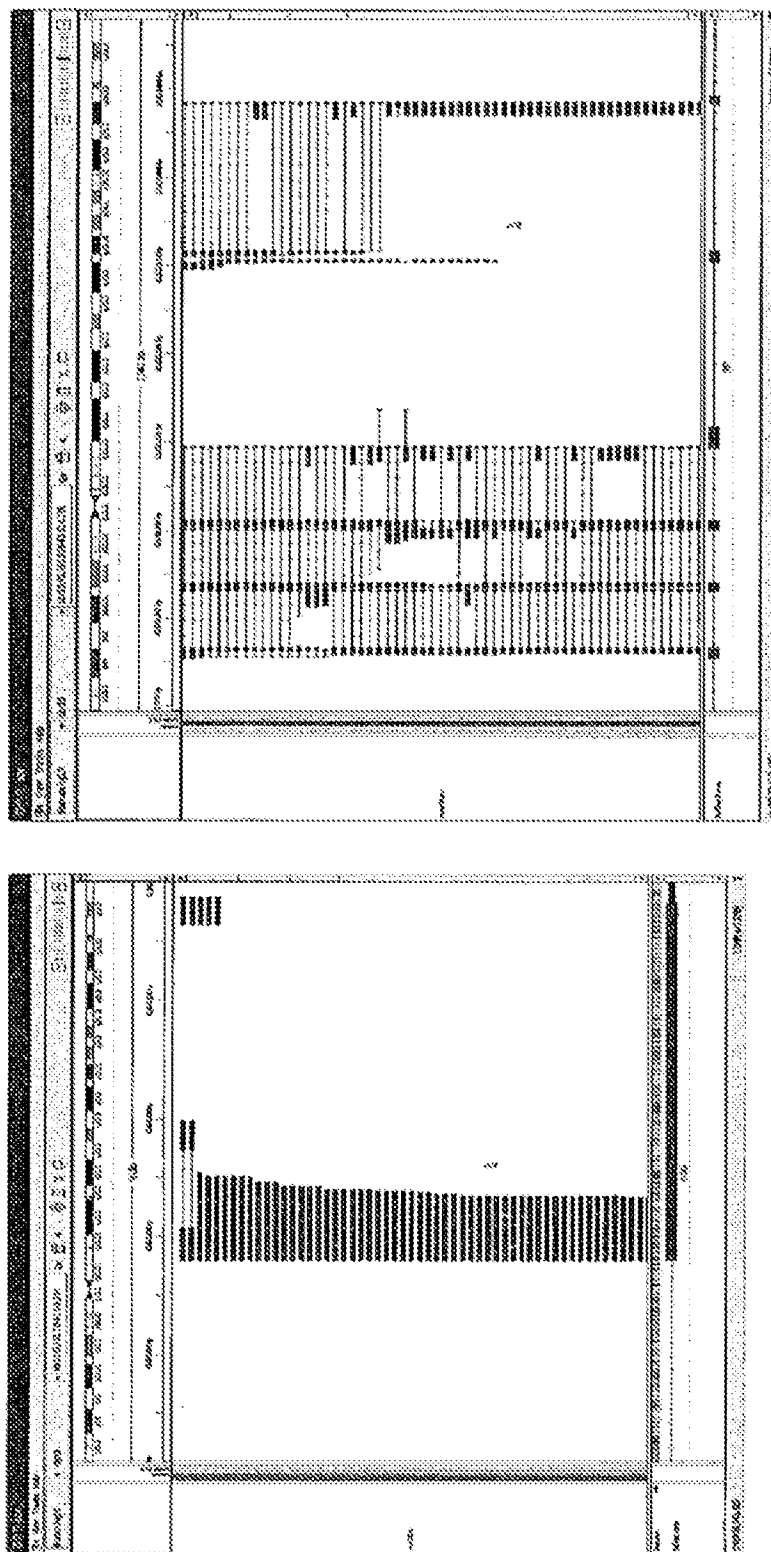


Figure 6

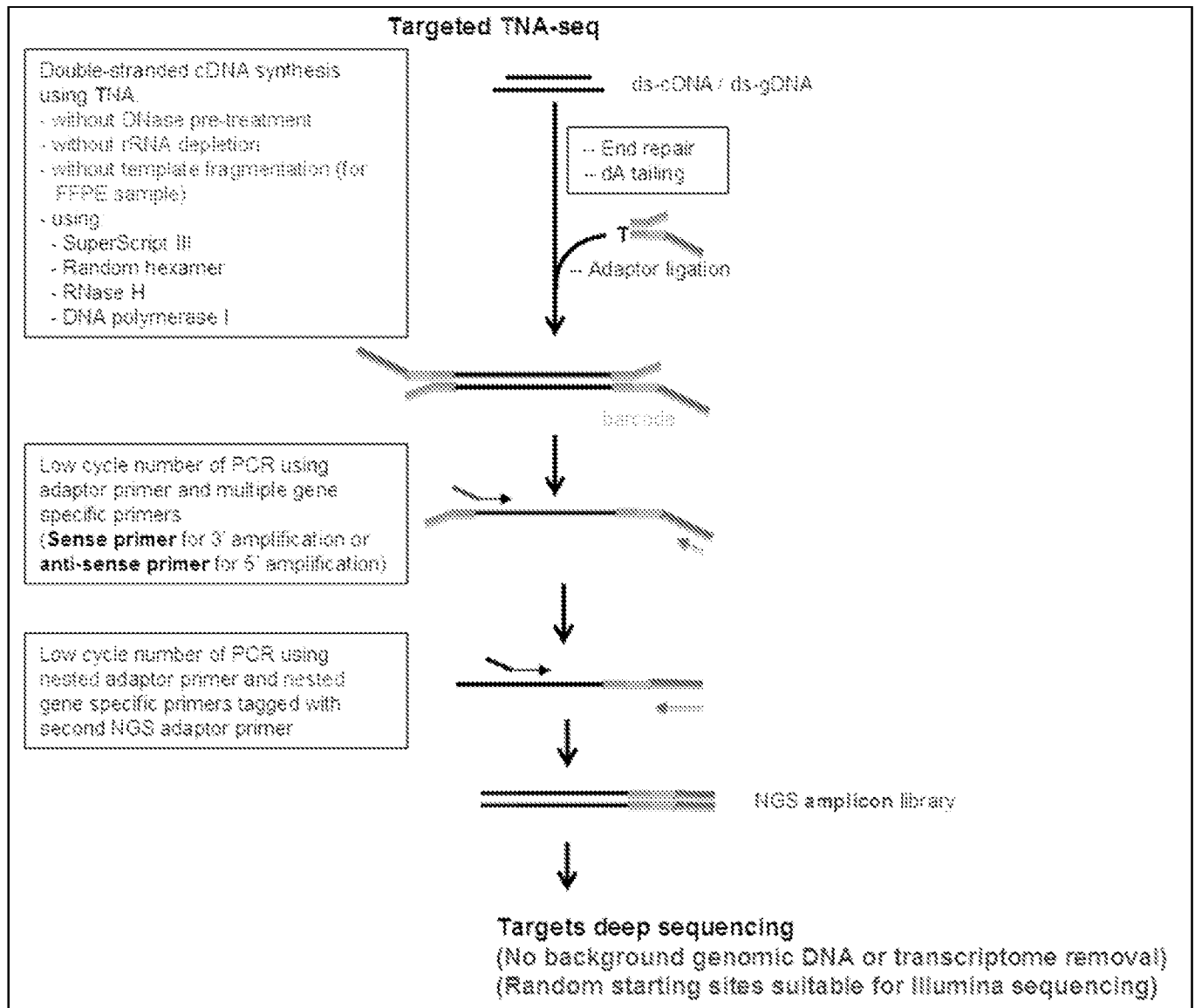


Figure 8

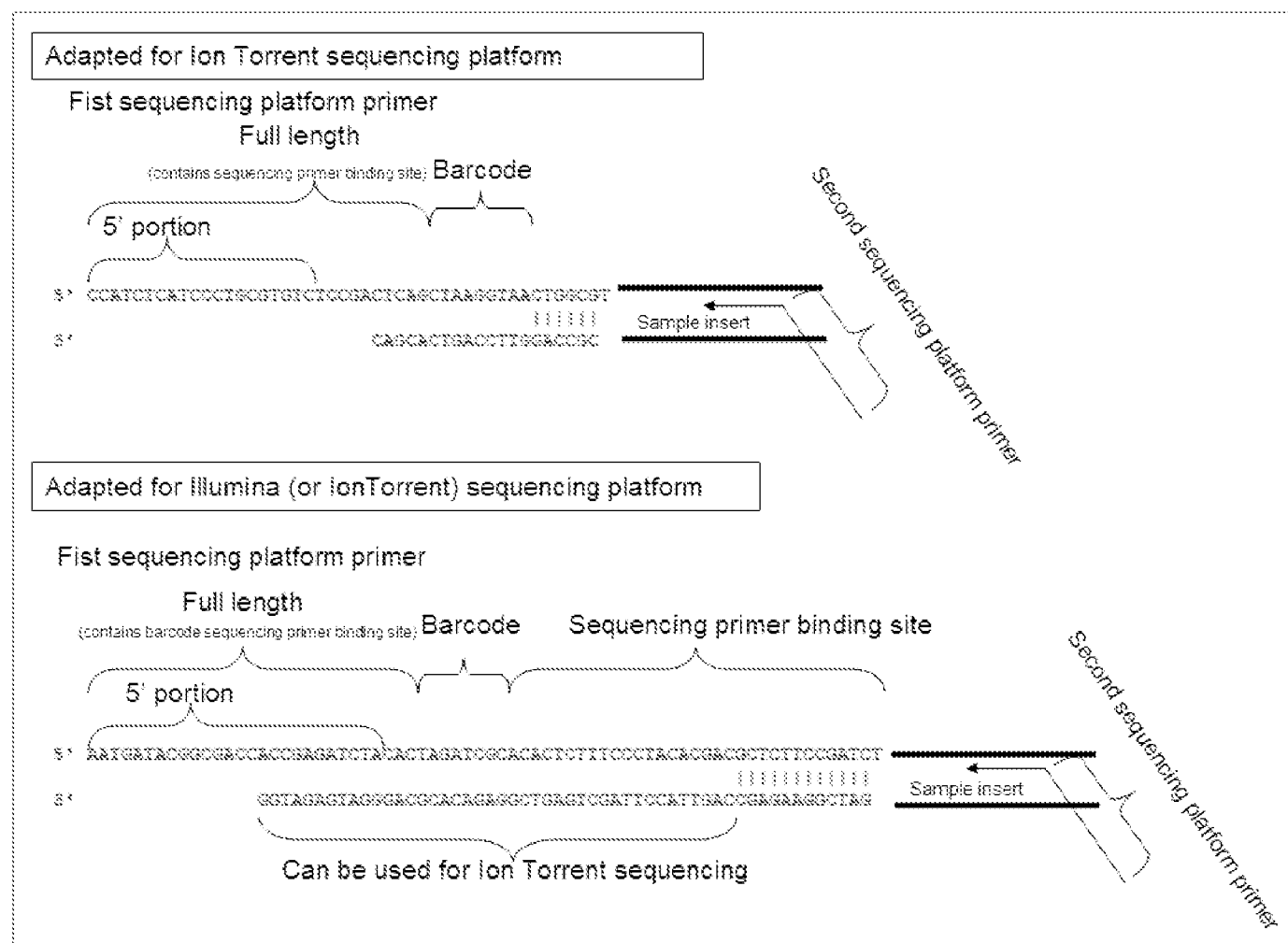


Figure 9