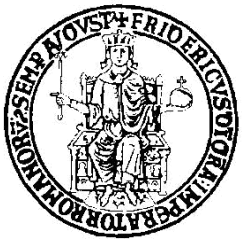
UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II



DIPARTIMENTO DI SANITA' PUBBLICA

DOTTORATO DI RICERCA IN SANITA' PUBBLICA E MEDICINA PREVENTIVA - XXXII CICLO

Next Generation Sequencing for gene fusions detection in Non Small Cell Lung Cancer

Tutor:

Ch.mo Prof. Giancarlo Troncone

Candidato:

Dott. ssa Lorenza Greco

Coordinatore:

Ch.mo Prof. Giancarlo Troncone

ANNO ACCADEMICO 2019- 2020

1 Abstract

Non Small Cell Lung Cancer (NSCLC) represents the most leading cause of death for cancer worldwide.[1] In addition to standard care (chemotherapy), several clinical trials demonstrated the efficacy in advanced stage (IIIB-IV) NSCLC patients of targeted treatment, represented by tyrosine kinase inhibitors (TKIs).[2] For this reason, the College of American Pathologists (CAP), the International Association for the Study of Lung Cancer (IASLC) and the Association for Molecular Pathology (AMP) established, in addition to epidermal growth factor receptor (*EGFR*) molecular assessment, a panel of "must test genes" that includes also gene fusions, such as Anaplastic Lymphoma Kinase (*ALK*) and ROS Proto-Oncogene 1 Receptor Tyrosine Kinase (*ROS1*).[3] On the overall, other clinically relevant gene rearrangements involving REarranged during Transfection (*RET*) and neurotrophic receptor tyrosine kinase (*NTRK*) genes 1, 2, and 3 showed their feasibility as targetable biomarkers in NSCLC patients.[4,5]

For predictive molecular analysis, tissue represents the most suitable sample type in diagnostic routine, but in NSCLC setting tissue specimens are often characterized by scant amount of nucleic acids on which molecular analysis may be performed. For this reason, very sensitive methodologies such as Next Generation Sequencing (NGS) and nanofluidic color-code barcode systems (Nanostring) should be implemented in clinical practice. Unfortunately, despite of NGS approach, a not negligible percentage of NSCLC patients (20-25%) cannot be analyzed for the quality and quantity limitations of "scant" samples. In this setting, "liquid biopsy" may represent a valid diagnostic tool to satisfy clinical needs.[6,7]

Liquid biopsy was clinically approved for the detection of sensitive mutations in EGFR when tissue is not available at basal or for acquired resistance mutation p.T790M after first line of TKIs treatments in NSCLC patients, unfortunately tissue specimen represents the only biological source to test clinically relevant gene fusions in clinical practice, this issues contributes to generate an incomplete molecular profile for the NSCLC patients.

For all these reason, the aim of this project is to validate the possibility to detect clinically relevant gene fusions in advanced NSCLC patients by adopting next generation sequencing (NGS) platform starting from RNA extracted from different specimens; secondary aim is the evaluation of the concordance rate between the two different NGS gene panels on RNA samples and an orthogonal platform with the same reference range.

Results will be compared with those obtained from RNA extracted on corresponding tissue specimens for each patient analyzed by a validated customized multiplex panel on nCounter platform (Nanostring Technologies, Seattle, WA), in Pangaea institute.

2 Introduction

Cancer is a major public health problem worldwide and is the second leading cause of death in the United State.[8] An estimated 606,880 Americans will die from cancer in 2019, corresponding to almost 1,700 deaths per day and one-quarter of all cancer deaths are due to lung cancer.[8] Lung cancer is a heterogeneous disease comprising several subtypes with pathologic and clinical relevance, is divided in small cell carcinoma (SCLC) and in Non-small cell lung cancer (NSCLC), the second one is the predominant subtype, which is further divided based on histological growth pattern.[9] The two major NSCLC histological subtypes are adenocarcinoma (ADC) and squamous cell carcinoma (SqCC).[10] NSCLC comprises approximately 80–85% of all lung cancers[11,12] with adenocarcinoma, approximately 40–50% of cases, and squamous cell carcinoma, approximately 20–30% of cases.[10–12] In particular, lung ADC is considered a cluster of discrete molecular subtypes, with most being defined by a single oncogenic driver alteration that mainly result in a downstream activation of canonical mitogen-activated protein kinases (MAPKs)/extracellular signal–regulated kinases or phosphatidylinositol 3 kinase (PI3K)/protein kinase B cancer pathways. The transition from empirical to mechanism-based biomarker-driven therapeutic decisions has had a profound impact on patients clinical outcomes, in fact in addition to chemotherapy, several clinical

trials demonstrated the efficacy in advanced stage (IIIB-IV) NSCLC patients of targeted treatment, represented by tyrosine kinase inhibitors (TKIs). In order to administrate these drugs, it is mandatory to analyze different biomarkers.[10,13] (Figure 1) For this reason, the College of American Pathologists (CAP), the International Association for the Study of Lung Cancer (IASLC) and the Association for Molecular Pathology (AMP) established, in addition to epidermal growth factor receptor (*EGFR*) molecular assessment, a panel of "must test genes" that includes also gene fusions, such as Anaplastic Lymphoma Kinase (*ALK*) and *ROS* Proto-Oncogene 1 Receptor Tyrosine Kinase (*ROS1*).[14] *ALK* and *ROS1* are examples of oncogenic gene rearrangements, can lead to expression of oncogenic fusion protein when a 5' partners forms an in-frame gene fusion with a 3' protoncogene.[15] (Figures 2,3) The increasing number of clinically relevant biomarkers is showed with gene rearrangements involving REarranged during Transfection (*RET*) and neurotrophic receptor tyrosine kinase (*NTRK*) genes 1, 2, and 3 that could predict patients sensitivity to new generations TKIs. To date, tissue represents the gold standard for the assessment of clinical relevant biomarkers mutational status, gene rearrangements included.[14] But several limitations affect the use of tissue specimen in clinical setting: the discomfort suffered by the patient, clinical risks, tumor heterogeneity, potential surgical complications and economic considerations meaning that multiple or serial biopsies are often impractical.[16] Indeed, a relevant problem associated with lung cancer is delayed diagnosis, a high percentage of NSCLC patients are diagnosed at more advanced stages (IIIB–IV).[17] When this happens, the only material available for morphological diagnosis and molecular assessment is either cytological specimens or small tissue histological biopsies.[18,19] However, because these specimens are often paucicellular, the evaluation of predictive biomarkers is very challenging. The adoption of very sensitive methodologies able to detect clinically relevant mutations should encourage predictive molecular analysis on tissue specimens by reducing samples classified as "inadequate".

In particular, FISH is the gold standard technique for *ALK* and *ROS1* gene fusion detection because was used to detect positive patients for these mutations during Crizotinib clinical trials,[20,21] (Figure

4) but the use of FISH as a screening tool for all patients with NSCLC presents numerous limits: the long turn-around time, high costs, high level of experience and also morphology cannot be appreciated after hybridization.[20,22,23] Indeed, FISH showed limitations for testing *ALK* gene fusions, because the intra-chromosomal inversion event only involves a small part of the short arm of chromosome 2, therefore, technical issues may afflict gene fusions identification by distinguishing red and the green fluorescent signals.[22,24] To overcome these problems and also to allow a contemporary evaluation of the clinically relevant gene rearrangement on same lung cancer tissue samples, nCounter technology (NanoString Technologies) is emerging as a promising method.[25]

This methodology is able to perform the simultaneous detection of *ALK, ROS1* gene fusions events, and other NSCLC gene rearrangements like *RET* proto-oncogene and the *MET* proto-oncogene skipping transcript; in addition, this platform provide a quantitative result about aberrant fusion protein quantification.[25] This technique is a relatively cheap and fast to detect gene fusions by starting from a minute amount of RNA, although the failure rate for the nCounter RNA gene fusion assay is not negligible (20%), numerous studies[26,27] confirm that nCounter analysis is highly concordant with FISH and IHC and thus is very specific.[28–30]

Despite of the implementation of very sensitive methodologies, in 30% of advanced NSCLC patients tissue specimen is not available.[31] A rapidly emerging and valid approach to overcome these issues and to extend predicting biomarker testing to the greatest possible number of NSCLC patients is non invasive "liquid biopsy." Although relevant molecular information regarding cancer driver genes can be obtained from a number of different liquid biopsy approaches, including circulating tumor cells, tumor RNA sequestered by platelet and exosomes, to date plasma-derived circulating tumor DNA constitutes the only sample source that has been officially approved for clinical use in NSCLC patients.[7,32–38] Generally, upon NSCLC diagnosis and before treatment (basal setting), a molecular analysis of EGFR mutational status on plasma-derived ctDNA is required to guide administration of first- and second-generation tyrosine kinase inhibitors (TKIs)

when tissue specimens are unavailable or inadequate for molecular testing.[7,32–34] Moreover, resistance exon 20 EGFR point mutation (p.T790M) should be analyzed in plasma-derived ctDNA generation TKI treatments.[7,34] In effect, the strong rationale behind this approach is that cancer DNA enters the bloodstream either by apoptosis and necrosis of tumor cells (passive mechanism) or by a spontaneous release of DNA fragments into the circulation from primary tumor tissues or from circulating tumor cells (active mechanism).[32]

Similar to ctDNA, RNA derived from tumor cells (ctRNA) is present in the plasma of cancer patients and can be used for detection of the clinically relevant *ALK*, *ROS1*, *RET* and *METΔ14* splicing variant. However, genetic analyses in cell-free RNA (cfRNA) present specific challenges and have not been widely used. Unlike cfDNA, cfRNA degrades very quickly and needs to be purified rapidly after blood extraction. Despite these limitations, studies have demonstrated the goodness of these resourse.[39]

Accordingly, with previous observations, on the one hand, high sensitivity techniques are required to avoid the risk of false negatives, on the other hand, a careful clinical validation is mandatory to minimize the generation of false-positive results.

Next Generation Sequencing (NGS) techniques may increase analytical sensitivity but also the examined genomic regions amplitude, enabling the simultaneous detection of several mutations in multiple genes of different samples by the parallel sequencing of millions of DNA fragments.[40] Each nucleotide is read several times, ensuring a high degree of sensitivity. To increase this analytical value, considering that circulating tumour nucleic acids represents only a small fraction (<0.5%) in most patients with solid tumours,[41] an 'ultra-deep sequencing' strategy performed by a small NGS panels could be an effective tool in clinical practice, based on cfRNA analysis.[24,42,43] (Figure 5)

For all these reasons, my PhD project purpose is to experiment the detection of clinically relevant gene fusions in advanced NSCLC patients by adopting NGS platform, starting from RNA extracted from different blood elements. Ion Torrent S5 GS instrument (ThermoFisher Scientific, Waltham,

MA) with the customized RNA SiRe® Fusion panel in predictive molecular pathology laboratory at the University of Naples "Federico II" (Naples, Italy). In a second in progress step, results will be compared with those obtained from RNA extracted on paired tissue specimens for each patient analyzed by a validated customized multiplex panel on nCounter platform (Nanostring Technologies, Seattle, WA) in Pangaea institute. Another key point raised in this project is related to the evaluation of the concordance rate between the two different NGS techniques on RNA samples.

3 Material and Methods

3.1 Design of the customized SiRe® RNA fusion panel

The Ion AmpliSeq Designer suite v5.3.1 with hg19 was used as reference genome to develop a customized panel targeting the most clinical relevant rearrangements in four genes (*ALK, ROS1, RET, NTRK*,) and exon skipping in *METΔ14* that play a key role to become NSCLC patients eligible at target therapy. A single primer pool leading to the selection of 91 amplicons (ranging from125 to175 bp) enabled us to cover selected genomic alterations in the target genes. The amplicon design was optimized for the simultaneous analysis of 16 samples with the *510 chip* (Thermofisher, Foster City, CA, USA) on Ion Torrent S5 GS instrument (ThermoFisher Scientific, Waltham, MA).

3.2 Study design

At University of Naples "Federico II" (Naples, Italy) the evaluation of analytical performance of customized SiRe® RNA fusion panel on Ion Torrent S5 GS instrument (ThermoFisher Scientific, Waltham, MA) was executed by analyzing different control specimens constituted by RNA isolated from mutant cell lines that harboured some of the most clinically relevant rearrangements in NSCLC (*EML4-ALK vr. 1, vr. 3a, 3b, SLC34A2-ROS1, CCDC6-RET, TPM3-NTRK1)* exon

skipping *METΔ14* and *MET* amplification. In the first step, we performed an NGS experiment with RNA isolated from 9 mutant cell lines to analyze the feasibility of the panel to detect mutations covered by reference range. In the second phase the technical sensitivity of the gene fusion panel was investigated by evaluating RNA pool, created by mixing RNA extracted from mutant cell lines, at 5 different dilutions points (20 ng/µl, 10 ng/µl, 2 ng/µl, 0,5 ng/µl, 0,1 ng/µl). Finally, a set of quantitative reference standard in cytological form (customized Horizon Diagnostics Multiplex RNA reference standard) was evaluated in order to verify if different pre-analytical approaches may influence the molecular analysis performed by using SiRe RNA fusion panel. Moreover, mutant cell lines were previously analysed in Pangaea Institute (Barcelona, Spain) by using GeneReader Platform (Qiagen, Germany) in combination with QIAact Lung RNA Fusion UMI Panel able to cover most clinically relevant fusion in lung cancer setting, shown in the following table, following manufacturer instructions. Data are not available.

AGTRAP-BRAF	DCTN1-ALK	FGFR1-ZNF703	HOOK3-RET	NCOA4-RET	RNF130-BRAF	TPM3-NTRK1
AKAP9-BRAF	EML4-ALK	FGFR3-TACC3	KIF5B-ALK	SDC4-ROS1	TPM3-ROS1	ATIC-ALK
ERC1-RET	FN1-ALK	KIF5B-RET	NTRK1-TPM3	SEC31A-ALK	TPM4-ALK	CCDC6-RET
ERC1-ROS1	GATM-BRAF	KLC1-ALK	PCM1-RET	SLC34A2-ROS1	TRIM24-RET	CD74-NRG1
ESRP1-RAF1	GNAI1-BRAF	LMNA-NTRK1	PPFIBP1-ALK	SLC45A3-BRAF	TRIM33-RET	CD74-NTRK1
EZR-ROS1	GOLGA5-RET	LRIG3-ROS1	PPFIBP1-ROS1	SQSTM1-ALK	UBE2L3-KRAS	CD74-ROS1
FAM131B-BRAF	GOPC-ROS1	LSM14A-BRAF	PRKAR1A-RET	STRN-ALK	VCL-ALK	CEP89-BRAF
FCHSD1-BRAF	HACL1-RAF1	MET exon14 skipping	PWWP2A-ROS1	TFG-ALK	ZSCAN30-BRAF	CLCN6-BRAF
FGFR1-PLAG1	HERPUD1-BRAF	MKRN1-BRAF	RAF1-DAZL	TFG-NTRK1	CLTC-ALK	FGFR1-TACC1
HIP1-ALK	MYO5A-ROS1	RANBP2-ALK	TPM3-ALK	NPM1-ALK		

Table 1. Reference Range QIAact Lung RNA Fusion UMI Panel

3.3 RNA extraction and cDNA synthesis

Cell lines were purchased from the American Type Culture Collection and cultured in RPMI medium 10% fetal bovine serum under standard conditions. (Table 2)

Cell pellets derived from a minimum of 5 T-75 flasks. Cells were counted using a Neubauer Chamber. and RNA was isolated from cell pellets with a *High Pure RNA Isolation kit* (Roche Diagnostic, Penzberg, Germany) according to the manufacturer's instructions.[42] The RNA was eluted in 30 µl of *RNAsi/DNAsi free water* (Ambion, Thermofisher, USA).

From Horyzon Discovery (Cambridge,UK) we obtained 4 slides sample set, in two slide set were fixed in ethanol and the other ones in methanol to verify RNA stability in relation to fixative agents adopted. For each fixation point, a negative and a positive slide for the following translocations (*EML4-ALK, CCDC6-RET, SLC3A2-ROS1, TPM3-NTRK1, ETV6-NTRK3)* were respectively tested. (Table 3)

For each fixation modality the cells were directly scraped by using a blade from two unstained slides. RNA extracted from customized Horizon Diagnostics Multiplex RNA reference, was processed after 1 month upon the receipt of the samples, using the *All Prep DNA - RNA Mini Kit* (Qiagen, Crawley, West Sussex, UK) following the manufacturer instructions and RNA was re-suspended in 50 µl of RNAsi/DNAsi free water (Ambion, Thermofisher, USA).

TapeStation 4200 (Agilent Technologies, Milan, Italy) a microfluidic based technology, was used to evaluate RNA concentration (ng/µl) and RNA integrity number (RIN) in order to define the optimal RNA concentration to perform cDNA synthesis.

The Retrotranscription was carried out using by *SuperScript IV VILO Master Mix* (ThermoFisher Scientific, Waltham, MA) according to the manufacturer's instructions.[44]

cell line	mutation
H3122	*EML4-ALK* vr 1
H2228	*EML4-ALK* vr 3a, 3b
Hs746T	*exon splicing MET14*

H596	*exon splicing MET14*
HCC-78	*SLC34A2-ROS1*
LC2-ad	*CCDC6-RET*
EBC-1	*MET Amplification*
SUDHL-1	*NPM1-ALK*
NTRK1cl	*TPM3-NTRK1*

Table 2. Molecular assessment of mutant cell lines

Item number	Description	Numb. slides	Internal reference
HD – D187	Cytology Fusion Negative ETOH fixed RNA	5	FN-ETOH
HD – D188	Cytology Fusion Positive ETOH fixed RNA	5	FP-ETOH
HD – D197	Cytology Fusion Negative MEOH fixed RNA	5	FN-MEOH
HD – D198	Cytology Fusion Positive MEOH fixed RNA	5	FP-MEOH

Table 3. Customized Horizon Diagnostics Multiplex RNA reference standard.

3.4 Libraries preparation and NGS analysis with QIAact Lung RNA Fusion UMI Panel

QIAact Lung RNA Fusion UMI Panel is a two primer mix tube designed to analyse most conventional translocations that play a clinical role in lung cancer patients. Libraries were generated starting from100 ng of extracted RNA on Gene Reader platform (Qiagen, Germany). After the end.repair step, the adaptor ligation was performed according to manufacturer instructions, then fragments were enriched in a PCR- enrichment step and finally sequenced on Gene Reader instrument. Results were carried out om Gene Reader analysis software following manufacturer instructions

3.5 Libraries preparation and NGS analysis with SiRe RNA fusion panel

The SiRe® RNA fusion primers pool was designed to reveal the most clinical relevant translocations in *ALK, ROS1, RET, NTRK* and *MET exon 14* skipping alterations. Libraries were constructed and purified on the Ion Chef instrument (Thermofisher) following the manufacturer procedures. Library generation was as follows: 6 microliters of cDNA (with an optmal concentration of 2 ng/microliter) were dispensed on Ion Code plates and amplified using Ion AmpliSeq DL8 (Thermofisher). We used 24 cycles for cfDNA amplification and 7 cycles for library re-amplification after barcoding, under the thermal conditions defined by the manufacturer. Purified libraries derived from RNA samples were diluted to 60pM and pooled. The pooled libraries were re-loaded into the Ion Chef instrument, and templates were prepared using the S5 510-520-530 chef Kit (Thermofisher). Finally, templates were loaded into the 520 chip and sequenced on S5 NGS platform (Thermofisher).

The results interpretation was carried out by using a proprietary pipeline developed by the Department of Public Health on IonReporter Software (Thermofisher).

4 Results

4.1 SiRe RNA fusion panel technical feasibility

RNA extracted evaluation

RNA concentration and integrity, in terms of RNA Integrity Number (RIN), was performed on microfluidic platform TapeStation 4200 (Agilent) on RNA extracted from mutant cell lines and results are reported in tables below.(Table 4) The average of the RNA concentration was 15,3 ng/µL, the minimum value was 2,12 ng/µL (RNA amount of NTRK1cl) and the maximum was 31,4 ng/µL (RNA amount of H596).

The average of the RIN was 2.9 the minimum value was 1.9 (RIN of LC2-ad) and the maximum was 5.5 (RIN of H596).

Sample	Concentration (ng/µL)	RNA Integrity Number-RIN
H3122	20,9	3.3
H2228	15,2	2.9
Hs746T	28,3	2.3
H596	31,4	5.5
HCC78	3,45	4.2
LC2-ad	8,4	1.9
EBC1	23,2	2.9
SUDHL-1	4,6	1.6
NTRK1cl	2,12	2.0

Table 4. Quantitative and qualitative parameters of RNA extracted from mutant cell lines evaluated by using TapeStation 4200 (Agilent).

Sequencing results

All samples processed for the feasibility test passed quality filters. NGS run parameters in mutant cell lines were evaluated in order to assess the analytical performance of the SiRe RNA fusion panel to detect clinically relevant rearrangments. Taking into account all 9 mutant cases, an average of 99,70% (ranging from 99,13% to 99,98%) reads on target was obtained. The median read length was of 103,55 bp (ranging from 93,00 to 114,00). Concerning the number of mapped reads, an average of 311798,44 (ranging from 60530 to 948223) was obtained. Considering uniformity of coverage, an average uniformity of coverage of 318415,89 (ranging from 60562 to 96378) was evaluated.

The feasibility of the SiRe RNA fusion panel for the detection of the most clinical relevant rearrangements in *ALK, ROS1, RET, NTRK,* and in *METΔ 14* exon skipping, was confirmed by the

results of the first experiment. In particular, two samples that harboured *MET 14* exon skipping, 5 fusion events with the corresponding fusion partner and one ALK fusion with unknown partner were correctly identified. (Table 5)

Cell line	Locus	Type	Filter	Genes (exons)	Reads count	Detection
H3122	chr2:42522656 – chr2:29446394	Fusion	PASS	EML4(13) – ALK(20)	13235	Present
H2228	chr2:42491871 - chr2:29446394	Fusion	PASS	EML4(6) - ALK(20)	2007	Present
HS746T			PASS			None
H596			PASS			None
HCC78	chr4:25665952 - chr6:117650609	Fusion	PASS	SLC34A2(4) - ROS1(32)	12271	Present
LC2ad	chr10:61665880 - chr10:43612032	Fusion	PASS	CCDC6(1) - RET(12)	26725	Present
EBC-1			PASS			None
SUDHL-1	chr2:29551347 - chr2:29430138	Fusion	PASS	Unknown-ALK	0,59	Present
NTRK1cl	chr1:154142878 - chr1:156844363	Fusion	PASS	TPM3(8) - NTRK1(10)	227248	Present-Non-Targeted

Table 5. NGS results of mutant cell lines performed by using Ion S5 platform (Thermofisher). Results showed technical performance of SiRe RNA fusion panel for clinically relevant rearrangements evaluation in mutant cell lines.

4.2 SiRe RNA fusion panel technical sensitivity

RNA extracted evaluation

RNA concentration and integrity, in terms of RNA Integrity Number (RIN), was performed on microfluidic platform TapeStation 4200 (Agilent) on RNA extracted from pooled mutant cell lines and results are reported in tables below. (Table 6)

Results showed that each dilution point was confirmed. As shown in the table below, a medium RIN value was 2.8 (from 3.6 to 4.0); the last three dilution points didn’t pass the RIN evaluation

Sample	Concentration (ng/µL)	RNA Integrity Number-RIN
Dil.1	19,3	3.6
Dil.2	9,4	4.0
Dil.3	1,8	n.a.
Dil.4	0,4	n.a
Dil.5	0,8	n.a.

Table 6. Quantitative and qualitative parameters of RNA extracted from a pool of mutant cell lines at different dilution point evaluated by using TapeStation 4200 (Agilent).

Sequencing results

All samples, evaluated to establish the analytical sensitivity of the panel, passed quality filters. NGS run parameters in pooled mutant cell lines were evaluated in order to assess the limit of detection of the SiRe RNA fusion panel to detect clinically relevant rearrangements Taking into account all dilution point, an average of 73,97% (ranging from 64,51% to 85,60%) reads on target was obtained. The median read length was of 93,00 bp (ranging from 48 to 117). Concerning the number of mapped reads, an average of 21427,80 (ranging from 13642,00 to 27230,00) was obtained. Considering uniformity of coverage, an average uniformity of coverage of 20432,40 (ranging from 28229 to 13784) was evaluated.

Results showed that The SiRe RNA fusion panel was able to correctly detect all fusions and the splicing event harboured by RNA pool until the dilution point of 5 ng/µl. In details, for the first and second diluition point six alterations of pooled mutant cell lines were carried out : *EML4-ALK vr 1, EML4-ALK vr 3a, 3b,* exon splicing *MET 14, SLC34A2-ROS1, CCDC6-RET, MET* Amplification, *NPM1-ALK, TPM3-NTRK1*.This results showed that 5 ng/µl is the minimum imput of starting material required to correctly identify translocations covered by the panel. All results are reported in Table 7.

RNA pool dilution point	Filter	Read counts	Fusion Detected	Splicing Detected
20 ng/µl	PASS	247573	6/6	1/1
10 ng/µl	PASS	274581	6/6	1/1
5 ng/µl	PASS	246518	6/6	1/1
0,5 ng/µl	PASS	366710	n.a.	n.a
0,1 ng/µl	PASS	173875	n.a.	n.a

Table 7. NGS results of mutant pooled mutant cell lines performed by using Ion S5 platform (Thermofisher). Results showed that 5 ng/µl represents the lowest starting material concentration able to correctly identify rearrangments harboured by pooled mutant cell lines.

4.3 SiRe RNA fusion panel analytical performance cytological reference standard

RNA extracted evaluation

RNA concentration and integrity, in terms of RNA Integrity Number (RIN), was performed on microfluidic platform TapeStation 4200 (Agilent) on RNA extracted from cytological reference standard slide and results are reported in tables below. (Table 8)

The average of the RNA concentrations was 21,7 ng/µL, the minimum value was 7,7 ng/µL (Cytology Fusion Positive MEOH fixed RNA) and the maximum was 37,6 ng/µL (Cytology Fusion Negative ETOH fixed RNA). RIN average was 1.8 (from a minimum of 1 in Cytology Fusion Negative ETOH and MEOH fixed to a maximum value of 2.9 in Cytology Fusion Positive MEOH fixed and the maximum was 2.9 (Cytology Fusion Positive MEOH fixed RNA).

Sample	Concentration (ng/µL)	RNA Integrity Number-RIN
HD – D187	37,6	1.0
HD – D188	8,4	2.5
HD – D197	33,0	1.0
HD – D198	7,7	2.9

Table 8. Quantitative and qualitative parameters of RNA extracted cytological reference

standard slide by using TapeStation 4200 (Agilent).

Sequencing results

All samples, evaluated to establish the analytical performance of the panel, passed quality filters. NGS run parameters in cytological reference standard slides were evaluated in order to assess the diagnostic feasibility of the SiRe RNA fusion panel to detect clinically relevant rearrangements. An average of 99,65% (ranging from 99,49% to 99,81%) reads on target was obtained. The median read length was of 97,25 bp (ranging from 95 to 99). Concerning the number of mapped reads, an average of 976,14 (ranging from 917630 to 1027576) was obtained. Considering uniformity of coverage, an average uniformity of coverage of 20432,40 (ranging from 28229 to 13784) was evaluated.

The SiRe RNA fusion panel correctly recognized all translocations in the cytological standard slide reference range (*EML4-ALK, CCDC6-RET, SLC3A2-ROS1, TPM3-NTRK1, ETV6-NTRK3)* in both positive artificial cytological smear slides, without any significant difference in relation to fixation modality. All results are reported in the following tables. (Table 9-12)

The FP – MEOH sample showed an average of 1171,00 reads count from minimum of 61 for EML4 – ALK fusion detection to 33224 reads count for SLC34A - ROS1 fusion detection.

The FP – ETOH sample showed an average of 307,69 reads count, from a minimum of 52 for the EML4 – ALK fusion detection to a maximum of 46965 reads count for the SLC34A - ROS1 fusion detection.

The FN – MEOH sample showed an average of 96326,20 reads count; from a minimum of to 11732 to 155179 reads count for house-keeping detection.

The FN – ETOH sample showed an average of 74790,4 reads count ; from a minimum of 4321 to 130763 reads count for house-keeping detection.

Locus	Type	Filter	Genes (exons)	Read counts	Detection
chr6:170871321	EXPR_CON-	PASS	TBP	52448	Present
chr12:53585787	EXPR_CON-	PASS	ITGB7	1874	Present
chr8:128751265	EXPR_CON-	PASS	MYC	163240	Present
chr11:118960975	EXPR_CON-	PASS	HMBS	87238	Present
chr1:156104319	EXPR_CON-	PASS	LMNA	174150	Present
chr4:25665952 – chr6:117650609	FUSION	PASS	SLC34A2(4) - ROS1(32)	33224	Present
chr10:61665880 – chr10:43612032	FUSION	PASS	CCDC6(1) - RET(12)	31148	Present
chr2:42522656 – chr2:29446394	FUSION	PASS	EML4(10) - ALK(20)	61	Present
chr4:25665952 – chr6:117645578	FUSION	PASS	SLC34A2(4) - ROS1(34)	402	Present
chr1:154142878 – chr1:156844363	FUSION	PASS	TPM3(8) - NTRK1(10)	1161	Present-Non-Targeted

Table 9. NGS results of FP – MEOH sample performed by using Ion S5 platform (Thermofisher).

Locus	Type	Filter	Genes (exons)	Read counts	Detection
chr6:170871321	EXPR_CON-	PASS	TBP	63629	Present
chr12:53585787	EXPR_CON-	PASS	ITGB7	4434	Present
chr8:128751265	EXPR_CON-	PASS	MYC	108560	Present
chr11:118960975	EXPR_CON-	PASS	HMBS	68122	Present
chr1:156104319	EXPR_CON-	PASS	LMNA	131058	Present
chr4:25665952 – chr6:117650609	FUSION	PASS	SLC34A2(4) - ROS1(32)	46965	Present
chr10:61665880 – chr10:43612032	FUSION	PASS	CCDC6(1) - RET(12)	33085	Present
chr2:42522656 – chr2:29446394	FUSION	PASS	EML4(10) - ALK(20)	52	Present
chr4:25665952 – chr6:117645578	FUSION	PASS	SLC34A2(4) - ROS1(34)	910	Present
chr1:154142878 – chr1:156844363	FUSION	PASS	TPM3(8) - NTRK1(10)	1591	Present-Non-Targeted

Table 10. NGS results of FP – ETOH sample performed by using Ion S5 platform (Thermofisher).

Locus	Type	Filter	Genes (exons)	Read counts	Detection
chr6:170871321	EXPR_CON-	PASS	TBP	51348	Present
chr12:53585787	EXPR_CON-	PASS	ITGB7	11732	Present
chr8:128751265	EXPR_CON-	PASS	MYC	155179	Present
chr11:118960975	EXPR_CON-	PASS	HMBS	86799	Present
chr1:156104319	EXPR_CON-	PASS	LMNA	176573	Present

Table 11. NGS results of FN – MEOH sample performed by using Ion S5 platform (Thermofisher).

Locus	Type	Filter	Genes (exons)	Read counts	Detection
chr6:170871321	EXPR_CON-	PASS	TBP	64528	Present
chr12:53585787	EXPR_CON-	PASS	ITGB7	4321	Present
chr8:128751265	EXPR_CON-	PASS	MYC	107905	Present
chr11:118960975	EXPR_CON-	PASS	HMBS	66435	Present
chr1:156104319	EXPR_CON-	PASS	LMNA	130763	Present

Table 12. NGS results of FN – ETOH sample performed by using Ion S5 platform (Thermofisher).

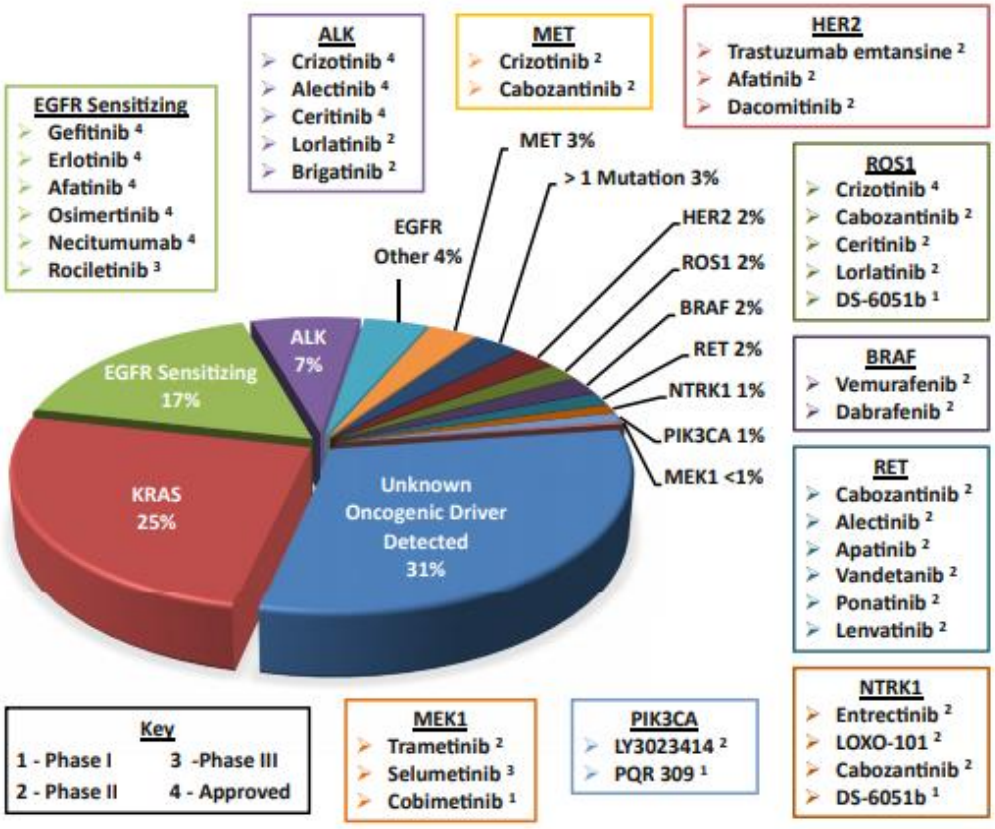
5 Discussion

In the era of personalized medicine is mandatory to test approved biomarkers to address NSCLC patients to the most tailed therapeutic approach. In the last decade the increasing number of biomarkers to test in order to better define the molecular profile of NSCLC patient, in addition to conventional *EGFR* molecular assessment, showed all critical issues regarding molecular analysis performed on NSCLC specimens. Generally, the most representative NSCLC specimen is a "scant "sample (small biopsy or cytological specimen) in terms of quality and quantity of nucleic acids available from this sample. In this setting where large part of samples are characterized by a small amount of nucleic acids implementation of high sensitive approach (such as ddPCR) is encouraged to avoid false negative results. In addition, single text technology seems to be inadequate to cover the molecular alterations of a panel of must test genes defined by the international societies. To overcome this limitation, multi test technology

will be largely adopted for molecular analysis in several hotspots in several genes in a single experimental procedure. Next Generation Sequencing is the most attractive and fashionable technology that may be approached to resolve this technical issue. Despite of NGS progressive introduction in clinical practice, another unsolved question is represented by breakpoint evaluation of less frequent recurrent clinical relevant translocations that may predict a positive response to TKIs in NSCLC patients. The first issue regarding commercial panel is represented by a very large amount of starting material to correctly identify clinical relevant translocations, a condition that in 85% of cases may not be verified in NSCLC setting. This evidence was full revealed in NGS experiments performed by using commercial panel on Gene Reader platform. The high amount of starting material required to successfully carry out molecular analysis did not allow us to correlate results produced from a commercial and a custom NGS panel to assess technical feasibility of NGS analysis routine scenario. This limitation revealed very low attraction of Gene Reader platform adoption for detection of rearrangements in lung cancer patients. . In addition, a full closed analytical pipeline does not allow to elucidate chromosomal alterations without a very high computational background. For all these reasons the built and validation of a custom RNA panel able to detect the clinical relevant translocations in *ALK, ROS, RET* genes is mandatory to correctly evaluate molecular assessment of NSCLC patients in clinical practice. SiRe RNA fusion panel showed high feasibility in detection of NSCLC recurrent chromosomal alterations by starting from a very scant amount of nucleic acids, in addition the creation of a specific pipeline for data analysis contribute to reduce misinterpretation in sequencing results and to generate a report in short time by integrating variant caller inspection in workflow analysis. The main limitation of the study is the lack of an orthogonal platform adopted to confirm NGS results, this step will be carried out in a second part of the project. The next step for this project is represented by its application for molecular analysis in diagnostic routine tissue samples. In addition, a very relevant perspective is to investigate SiRe RNA fusion panel performance to identify chromosomal alterations by starting from liquid

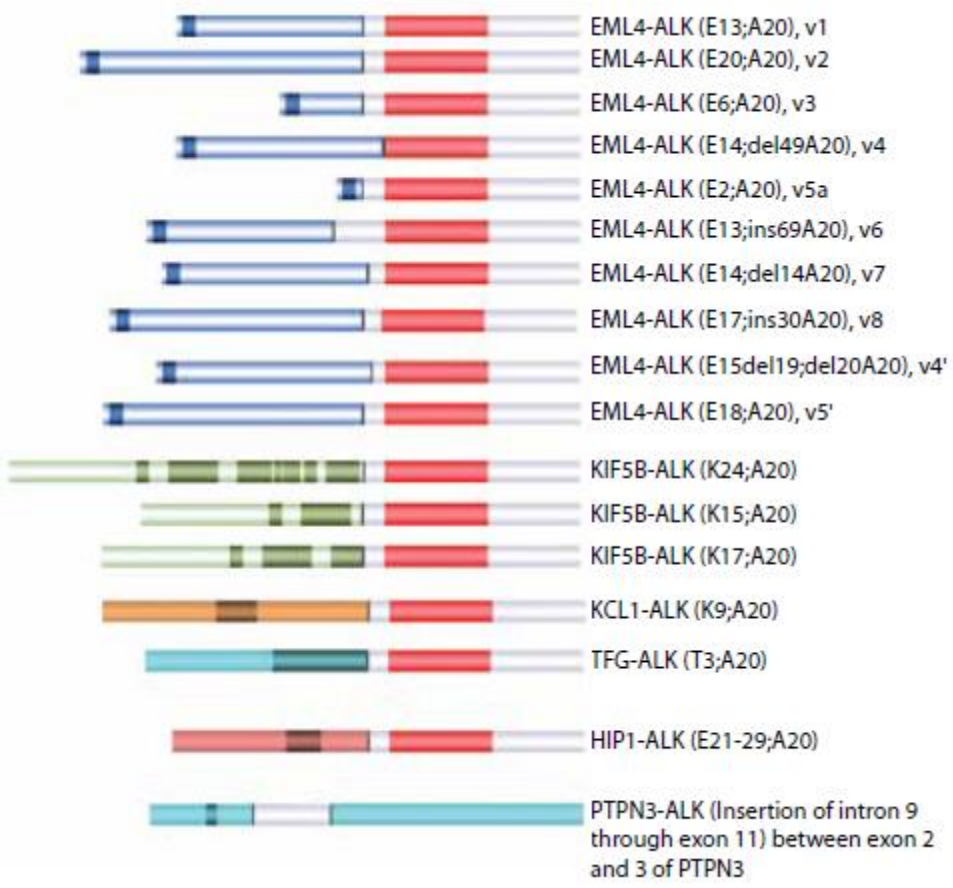
biopsy samples. This project should lead, in my hypothesis, to increase liquid biopsy attitude to be adopted in diagnostic routine for the definition of molecular NSCLC patient assessment.

6 Images



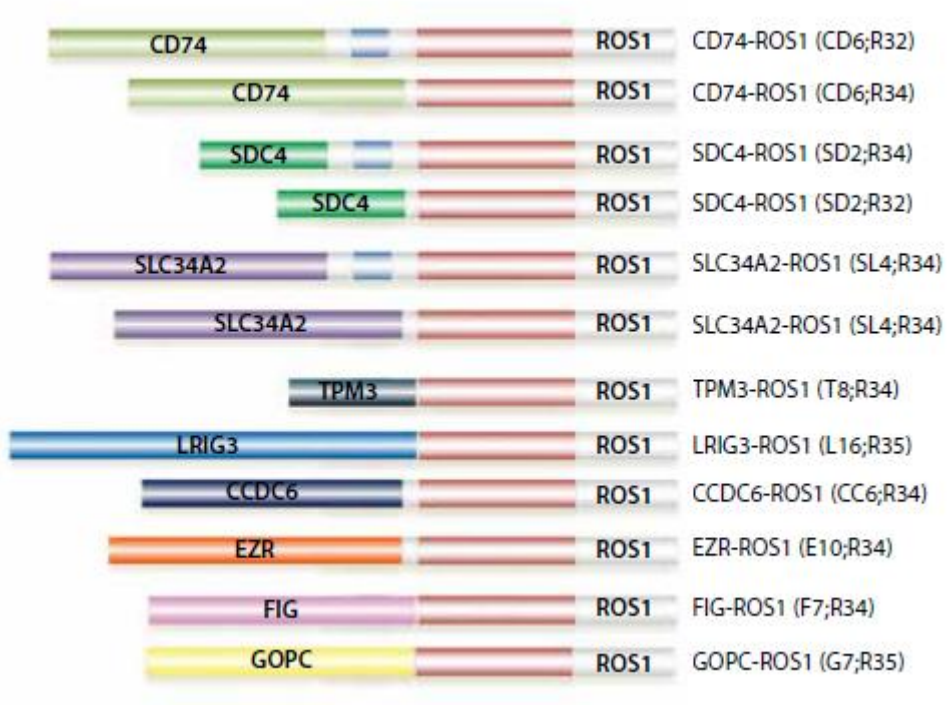
Tsao AS, Scagliotti GV, Bunn PA Jr et al. Scientific Advances in Lung Cancer 2015. J Thorac Oncol. 2016;11:613-38.

Figure 1. Frequency of molecular aberrations in various driver oncogenes in lung adenocarcinomas and current available drugs against these oncogenic proteins. EGFR, epidermal growth factor receptor; ALK, anaplastic lymphoma receptor tyrosine kinase; MET, mesenchymal-toepithelial transition factor; HER2, erb-b2 receptor tyrosine kinase 2; ROS1, ROS proto-oncogene 1, receptor tyrosine kinase; BRAF, B-Raf proto-oncogene, serine/threonine kinase; RET, ret proto-oncogene; NTRK1, neurotrophic tyrosine kinase receptor type 1; PIK3A, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha; MEK1, mitogen-activate protein kinase kinase 1; KRAS, Kirsten rat sarcoma viral oncogene homolog.



Tsao MS, Hirsch FR, Yatabe Y. *Iaslc Atlas of ALK and ROS1 Testing in Lung Cancer*.2016;

Figure 2. Schematic diagram of *ALK* rearrangement. The genes and domains are highlighted in different colors. Darker regions represent coil-coil domains in the fusion partner genes (*EML4, KIF5B, KLC1, TFG*), and the kinase domain in *ALK* (red).



Tsao MS, Hirsch FR, Yatabe Y. *Iaslc Atlas of ALK and ROS1 Testing in Lung Cancer*.2016;

Figure 3. Schematic diagram of *ROS1* rearrangement. *ROS1* kinase and transmembrane domains are highlighted in brown and blue, respectively. Partner genes are shown as different colors.

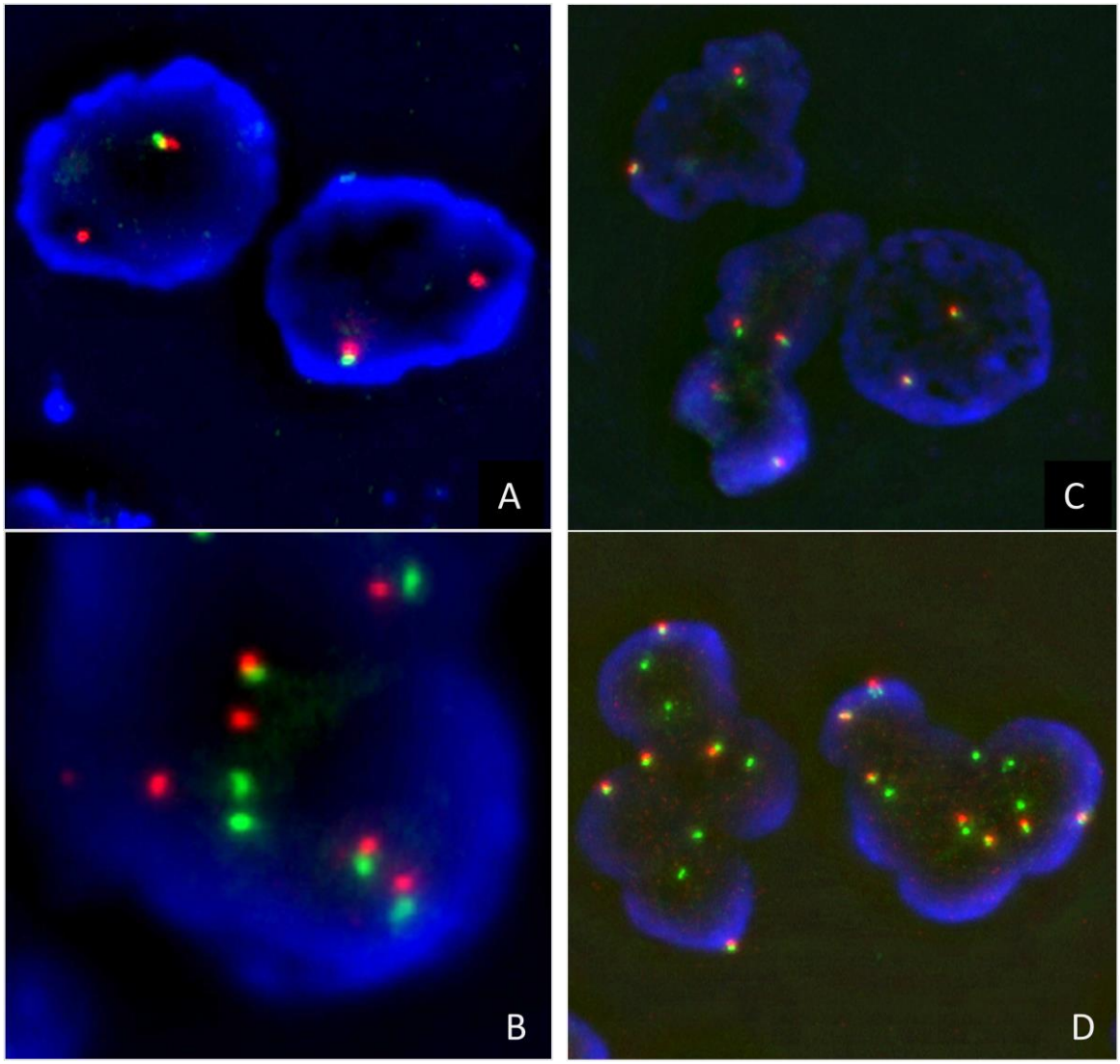
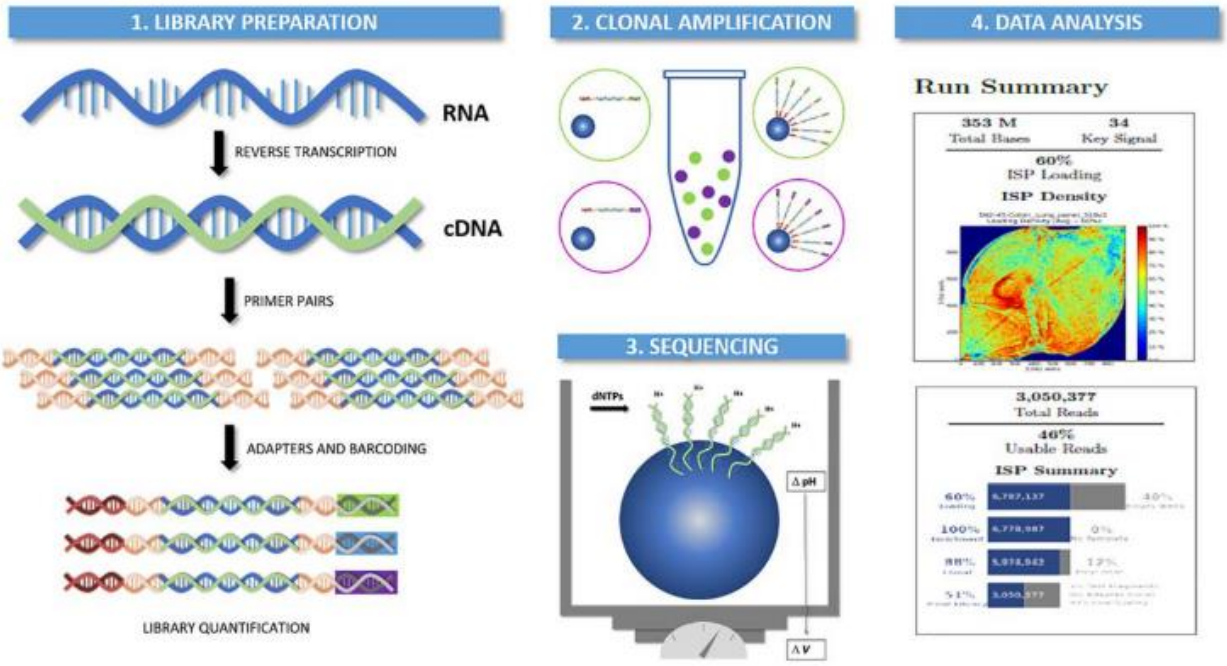


Figure 4. Examples of FISH with break-apart probe for ALK and ROS1 fusions detection. ALK-rearranged lung cancer (A-B) and ROS1-rearranged lung cancer (C-D) microscopic fields, A-C and B-D are a cytological and histological specimen respectively.



Pisapia P, Lozano MD, Vigliar E, et al. ALK and ROS1 testing on lung cancer cytologic samples: Perspectives. *Cancer Cytopathol*. 2017;125(11):817-830

Figure 5. Workflow of the next-generation sequencing assay performed with the Ion Torrent platform is illustrated. The procedure consists of 4 sequential phases: 1. Library preparation: The RNA to be sequenced is reverse transcribed into combinational DNA (cDNA), amplified, and used to construct a library, ie, a collection of cDNA fragments, each with barcodes for platform process- ing and patient identification. The library concentration is then quantified. 2. Clonal amplification: A single cDNA fragment of the library mixture is isolated by limiting dilution and clonally amplified by an emulsion polymerase chain reaction on beads. 3. Sequencing: Individual nucleotides are left to flow over the open wells of an Ion Torrent Chip. Upon the incorporation of each nucleotide, the chip detects the pH, and voltage changes consequent to the release of a hydrogen ion (H1) within any individual well. 4. Data analysis: In the run summary shown, 60% of wells in a 316 Ion Torrent Chip were loaded with template beads to gen- erate a read. After automated processing of the bioinformatics pipeline, more than 3 million high-quality sequencing reads are produced.

7 Bibliography

1. Siegel R, Naishadham D, Ahmedin J. Cancer Statistics, 2012 Rebecca. *CA Cancer J Clin*. 2012;62(1):10-29. doi:10.3322/caac.20138.Available
2. Bunn PA, Thatcher N. Systemic treatment for advanced (stage IIIb/IV) non-small cell lung cancer: more treatment options; more things to consider. Introduction. *Oncologist*. 2008;13(Supplement 1):1-4. doi:10.1634/theoncologist.13-s1-1
3. Lenz C, Steffen JA, Engelsberg A. Molecular testing guidelines for selection of lung cancer patients for ALK tyrosine kinase inhibitors. *J fur Pharmakologie und Ther*. 2013;22(2):45-48. doi:10.1016/j.jmoldx.2013.03.001
4. Drilon A, Rekhtman N, Arcila M, Al. E. A Phase 2 Single Arm Trial of Cabozantinib in Patients with Advanced RET-Rearranged Lung Cancers. *lancet oncol*. 2016;17(12):1653-1660. doi:10.1016/j.physbeh.2017.03.040
5. Drilon A, Laetsch TW, Kummar S, et al. Efficacy of Larotrectinib in TRK Fusion–Positive Cancers in Adults and Children. 2018;378(8):731-739. doi:10.1056/NEJMoa1714448.Efficacy
6. Pisapia P, Malapelle U, Troncone G. Liquid Biopsy and Lung Cancer. *Acta Cytol*. 2018. doi:10.1159/000492710
7. Malapelle U, Pisapia P, Rocco D, et al. Next generation sequencing techniques in liquid biopsy: Focus on non-small cell lung cancer patients. *Transl Lung Cancer Res*. 2016;5(5):505-510. doi:10.21037/tlcr.2016.10.08
8. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019 (US statistics). *CA Cancer J Clin*. 2019;69(1):7-34. doi:10.3322/caac.21551
9. Fujimoto J, Wistuba I. Current concepts on the molecular pathlogy of NSCLC. *Semin Diagn Pathol*. 2014;31(4):306-313. doi:10.1053/j.semdp.2014.06.008.Current
10. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization

Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances since the 2004 Classification. *J Thorac Oncol*. 2015;10(9):1243-1260. doi:10.1097/JTO.0000000000000630

11. Livstone EM. Colorectal cancer (Colon Cancer) - Merck Manual, Professional Version. *Merck Manual, Prof Version*. 2019;67(1):7-30. doi:10.3322/caac.21387
12. Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG. Introduction to the 2015 World Health Organization Classification of Tumors of the Lung, Pleura, Thymus, and Heart. *J Thorac Oncol*. 2015;10(9):1240-1242. doi:10.1097/JTO.0000000000000663
13. Tsao AS, Scagliotti G V., Bunn PA, et al. Scientific advances in lung cancer 2015. *J Thorac Oncol*. 2016;11(5):613-638. doi:10.1016/j.jtho.2016.03.012
14. Lindeman NI, Cagle PT, Aisner DL, et al. Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors guideline from the college of American pathologists, the international association for the study of lung cancer, and the a. *Arch Pathol Lab Med*. 2018;142(3):321-346. doi:10.5858/arpa.2017-0388-CP
15. Tsao MS, Hirsch FR, Yatabe Y. *IASLC ATLAS OF ALK AND ROS1 TESTING IN LUNG CANCER*.; 2016. doi:10.1055/s-0030-1255864
16. McPherson E. Genetic diagnosis and testing in clinical practice. *Clin Med Res*. 2006;4(2):123-129. doi:10.3121/cmr.4.2.123
17. Pisapia P, Pepe F, Smeraglio R, et al. Cell free DNA analysis by SiRe® next generation sequencing panel in non small cell lung cancer patients: Focus on basal setting. *J Thorac Dis*. 2017;9(Suppl 13):S1383-S1390. doi:10.21037/jtd.2017.06.97
18. Ofiara LM, Navasakulpong A, Ezer N, Gonzalez AV. The importance of a satisfactory biopsy for the diagnosis of lung cancer in the era of personalized treatment. *Curr Oncol*. 2012;19(0):16-23. doi:10.3747/co.19.1062
19. Wang S, Yu B, Ng CC, et al. The suitability of small biopsy and cytology specimens for

EGFR and other mutation testing in non-small cell lung cancer. *Transl Lung Cancer Res*. 2015;4(2):119-125. doi:10.3978/j.issn.2218-6751.2015.01.05

20. Sholl LM, Weremowicz S, Gray SW, et al. Combined use of ALK immunohistochemistry and FISH for optimal detection of ALK-rearranged lung adenocarcinomas. *J Thorac Oncol*. 2013;8(3):322-328. doi:10.1097/JTO.0b013e31827db604
21. Chiari R, Buttitta F, Iacono D, et al. Dramatic response to crizotinib in ROS1 fluorescent in situ hybridization- and immunohistochemistry-positive lung adenocarcinoma: A case series. *Clin Lung Cancer*. 2014;15(6):470-474. doi:10.1016/j.cllc.2014.06.004
22. Savic S, Bubendorf L. Common fluorescence in situ hybridization applications in cytology. *Arch Pathol Lab Med*. 2016;140(12):1323-1330. doi:10.5858/arpa.2016-0202-RA
23. Roy Chowdhuri S, Hanson J, Cheng J, et al. Semiautomated laser capture microdissection of lung adenocarcinoma cytology samples. *Acta Cytol*. 2012;56(6):622-631. doi:10.1159/000342984
24. Pisapia P, Lozano MD, Vigliar E, et al. ALK and ROS1 testing on lung cancer cytologic samples: Perspectives. *Cancer Cytopathol*. 2017;125(11):817-830. doi:10.1002/cncy.21899
25. Reguart N, Teixidó C, Giménez-Capitán A, et al. Identification of ALK, ROS1, and RET fusions by a multiplexed mRNA-based assay in formalin-fixed, paraffin-embedded samples from advanced non-small-cell lung cancer patients. *Clin Chem*. 2017;63(3):751-760. doi:10.1373/clinchem.2016.265314
26. Lira ME, Choi Y La, Lim SM, et al. A single-tube multiplexed assay for detecting ALK, ROS1, and RET fusions in lung cancer. *J Mol Diagnostics*. 2014;16(2):229-243. doi:10.1016/j.jmoldx.2013.11.007
27. Zhao W, Choi Y La, Song JY, et al. ALK, ROS1 and RET rearrangements in lung squamous cell carcinoma are very rare. *Lung Cancer*. 2016;94:22-27. doi:10.1016/j.lungcan.2016.01.011
28. Sgariglia R, Pisapia P, Nacchio M, et al. Multiplex digital colour-coded barcode technology

on RNA extracted from routine cytological samples of patients with non-small cell lung cancer: Pilot study. *J Clin Pathol*. 2017;70(9):803-806. doi:10.1136/jclinpath-2017-204373

29. Lira ME, Kim TM, Huang D, et al. Multiplexed gene expression and fusion transcript analysis to detect ALK fusions in lung cancer. *J Mol Diagnostics*. 2013;15(1):51-61. doi:10.1016/j.jmoldx.2012.08.006

30. Sunami K, Furuta K, Tsuta K, et al. Multiplex diagnosis of oncogenic fusion and met exon skipping by molecular counting using formalin-fixed paraffin embedded lung adenocarcinoma tissues. *J Thorac Oncol*. 2016;11(2):203-212. doi:10.1016/j.jtho.2015.10.005

31. Malapelle U, Bellevicine C, De Luca C, et al. EGFR mutations detected on cytology samples by a centralized laboratory reliably predict response to gefitinib in non-small cell lung carcinoma patients. *Cancer Cytopathol*. 2013;121(10):552-560. doi:10.1002/cncy.21322

32. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: Monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol*. 2013;10(8):472-484. doi:10.1038/nrclinonc.2013.110

33. Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*. 2011;11(6):426-437. doi:10.1038/nrc3066

34. Malapelle U, Mayo De-Las-Casas C, Rocco D, et al. Development of a gene panel for next-generation sequencing of clinically relevant mutations in cell-free DNA from cancer patients. *Br J Cancer*. 2017;116(6):802-810. doi:10.1038/bjc.2017.8

35. Pantel K, Speicher MR. The biology of circulating tumor cells. *Oncogene*. 2016;35(10):1216-1224. doi:10.1038/onc.2015.192

36. Reclusa P, Taverna S, Pucci M, et al. Exosomes as diagnostic and predictive biomarkers in lung cancer. *J Thorac Dis*. 2017;9(3):S1373-S1382. doi:10.21037/jtd.2017.10.67

37. Rolfo C, Giallombardo M, Reclusa P, Sirera R, Peeters M. Exosomes in lung cancer liquid biopsies: Two sides of the same coin? *Lung Cancer*. 2017;104:134-135.

doi:10.1016/j.lungcan.2016.11.014

38. Molina-Vila MA, Mayo-de-las-Casas C, Bertran-Alamillo J, Jordana-Ariza N, González-Cao M, Rosell R. CfDNA analysis from blood in melanoma. *Ann Transl Med*. 2015;3(20):1-6. doi:10.3978/j.issn.2305-5839.2015.11.23

39. Molina-Vila MA, Mayo-de-las-Casas C, Giménez-Capitán A, et al. Liquid biopsy in non-small cell lung cancer. *Front Med*. 2016;3(DEC):1-8. doi:10.3389/fmed.2016.00069

40. Malapelle U, Vigliar E, Sgariglia R, et al. Ion Torrent next-generation sequencing for routine identification of clinically relevant mutations in colorectal cancer patients. *J Clin Pathol*. 2015;68(1):64-68. doi:10.1136/jclinpath-2014-202691

41. Mead R, Duku M, Bhandari P, Cree IA. Circulating tumour markers can define patients with normal colons, benign polyps, and cancers. *Br J Cancer*. 2011;105(2):239-245. doi:10.1038/bjc.2011.230

42. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7521):543–550. doi:10.1038/nature13879

43. Paweletz CP, Sacher AG, Raymond CK, et al. Bias-corrected targeted next-generation sequencing for rapid, multiplexed detection of actionable alterations in cell-free DNA from advanced lung cancer patients. 2016;22(4):915–922. doi:10.1158/1078-0432.CCR-15-1627-T.Bias-corrected

44. Thermo Fisher Scientific. SuperScript IV VILO Master Mix. 2016;(11766050):6. https://www.thermofisher.com/order/catalog/product/11756050.