Joint Operational Programme
Romania-Republic of Moldova 2014-2020

Research Report regarding theoretical research directions that will base future projects in lung cancer care

This project is funded by the European Union

Romania-Republic of Moldova
ENI-CROSS BORDER COOPERATION

Institutul Regional de Oncologie Iasi IRO
A project implemented by Regional Institute of Oncology Iasi

ONCO
PMSI Institute of Oncology

www.ro-md.net

Research Report regarding theoretical research directions that will base future projects in lung cancer care

Content

Introduction

Cancer is the second leading cause of mortality after cardiovascular diseases with over 1.5 million deaths annually. Each year, about 11,000 Romanians are diagnosed with lung cancer. In Moldova lung cancer is the third most common type of cancer diagnosed and represents more than 10 percent of the total malignant tumors detected in our country, with over 800 people diagnosed last year. Most of affected persons are men and among the male population lung tumors are on the first place having an incidence of 18.8%. While there are multiple types of the disease, non-small cell lung cancer (NSCLC) comprises 85% of all lung cancers. The majority of NSCLC are locally advanced or metastatic disease and most new diagnoses survive less than a year, and among all NSCLC cases. The 5-year survival rate is 16%. The main involved target gene is EGFR, whose mutations were demonstrated to have a very important role in pathogenesis of lung cancer.

Besides EGFR, many other genes, including ALK, RET and ROS1, called driver oncogenes, are critical to the survival, growth and proliferation of cancer cells. Driver oncogene profiles are unique for each cancer type, and even histologically similar cancers can be stratified into molecularly distinct subsets of a given type of cancer. Many of these driver oncogenes (including ALK, RET and ROS1), encode tyrosine kinases, and researchers have identified tyrosine kinase inhibitors (TKIs) that selectively inhibit the activity of some oncogene translocations.

Access to innovative therapies that double the chance of survival for patients with this disease, has to be facilitated by introducing molecular testing in the routine.

Research goal

The main objective of the project was to create and develop a transnational clinical research and diagnostic network in the field of lung cancer. This was established by developing regional education capacities by training a body of proactive health professionals highly skilled and competitive in the management of lung cancer in the cross-border region. This specialist core will have the capacity to implement research and development projects in lung cancer domain.

Implementation of quality management system integrated with European standards for the research and clinical practice in lung cancer, involved a series of activities accomplish in proper timing. These steps consist in optimization of the technique, cross-validation of diagnosis for identification of therapeutic markers (EGFR mutations), and introduction of a supplementary panel of genes (ALK, ROS, RET etc.) in order to provide a reliable work tool for lung cancer diagnosis, as well as for documentation of the evolution of the disease.

This led to establishing the realistic and trusted technical tool for investigation in routinely diagnosis and in researching in lung cancer domain. Moreover, this conducted towards obtaining a trained multidisciplinary team capable to work with this diagnosis tool and to continue with forward research directions, for medium and long term.

The project contributed to fulfilling of basic principles of National Health strategy: equitable access to essential services, cost-effectiveness, optimization of health services, with emphasis on services and preventive interventions.

The research team has integrated specialists with experience from the two institutions that can cover the entire range of investigations necessary for the project and have sufficient experience in interpreting and integrating the data obtained. Also, the multidisciplinary team from both institutions drew conclusions from the data obtained during its development and through comparisons with the literature and the state of current knowledge. Besides the multidisciplinary work team of the project, the collateral beneficiaries were the employees of the two laboratories (doctors, biologists, chemists, laboratory assistants) and the young researchers, as well as the residents of the clinic. They participated in the workshops to promote the project and to exchange the experience in which very concentrated, but well-structured information was taught.

A continuous training was also performed regarding the development of the laboratory techniques and the interpretation of the results.

Research methodology

Samples from patients with diagnostic of lung cancer - 300 from each region/city — Iasi, Chisinau were subjected to Real-Time QPCR analysis by using different equipment and reagents. A number of samples (10) were exchanged between the teams and the results were compared regarding the mutational status of each sample. A number of samples (60) from each institution were subjected to NGS analysis by using different methodology. The workflow of Illumina sequencing and Ion Torrent sequencing was compared and the results were discussed extensively. Both teams exchanged knowledge about optimization, implementation and validation of NGS technique for driver genes mutation testing on Illumina MiSeq platform and data interpretation on Illumina software and the data was compared with the IonTorrent platform and software interpretation at the training sessions.

Methodology

Patients

This study included both patients from the Regional Oncology Institute from Iasi (IRO-Is) and the Oncology Institute from Chisinau (OI-Ch), diagnosed with non-microcellular lung cancer (NSCLC) or squamous lung cancer (young patients and non-smokers).

The samples enrolled in the study were retrospective samples as well as prospective ones.

Inclusion criteria:

- Adult patient with non-microcellular lung carcinoma (NSCLC) or squamous cell lung cancer (young and non-smoking patients)
- Signing the informed consent to take sections of paraffin tissue or peripheral blood for processing

Exclusion criteria:
- Paraffin blocks with low percentage of tumor cells (depleted blocks)
- Small cell lung cancer (SCLC)

Mutational analysis of EGFR

Epidermal growth factor receptor (EGFR) activating mutations are found in exons 18 to 21 of the EGFR gene, which is part of the gene coding for the tyrosine kinase domain of the EGFR protein.

Exon 19 deletions and an L858R point mutation in exon 21 account for around 90% of all EGFR mutations in advanced non-small cell lung cancer (NSCLC). EGFR tyrosine kinase inhibitors (TKIs) have high affinity for mutant EGFR and block aberrant EGFR signaling pathways to inhibit proliferation, tumor growth, cell survival and angiogenesis.

Treatment of EGFR mutation-positive advanced NSCLC tumors with EGFR-TKIs results in tumor shrinkage in the majority of patients. Several randomized trials have demonstrated the importance of epidermal growth factor receptor (EGFR) mutation status in determining the most appropriate first-line treatment for patients with advanced non-small cell lung cancer (NSCLC). Clinical trials have shown that patients with EGFR mutation-positive tumors have significantly improved progression-free survival (PFS) and objective response rate (ORR) when treated with first-line EGFR tyrosine kinase inhibitors (TKIs) compared with standard platinum-based doublet chemotherapy regimens.

DNA extraction

The **cobas® DNA Sample Preparation Kit** is used for manual sample preparation to isolate genomic DNA from formalin-fixed paraffin-embedded tumor tissue (FFPET) samples. FFPET samples are processed and genomic DNA isolated using the cobas® DNA Sample Preparation Kit, a generic manual sample preparation based on nucleic acid binding to glass fibers. A deparaffinized 5-micron (µm) section of an FFPET sample is lysed by incubation at an elevated temperature with a protease and chaotropic lysis/binding buffer that releases nucleic acids and protects the released genomic DNA from DNases. Subsequently, isopropanol is added to the lysis mixture that is then centrifuged through a column with a glass fiber filter insert. During centrifugation, the genomic DNA is bound to the surface of the glass fiber filter. Unbound substances, such as salts, proteins and other impurities, are removed by centrifugation. The adsorbed nucleic acids are washed and then eluted with an aqueous solution. The amount of genomic DNA can be spectrophotometrically determined and adjusted to a fixed concentration.

The **NucleoSpin FFPE DNA kit** is designed to purify total DNA from paraffin-fixed samples (FFPE). After deparaffining, the tissue is subjected to the action of proteinase K, so the DNA is released into solution. The release of DNA from different bonds is done by incubation at a high temperature in the presence of the decross-link buffer and precipitation is done with absolute alcohol. The entire lysate is loaded onto a NucleoSpin FFPE DNA Column and the DNA is attached to the silicon membrane. After two washes, the salts, metabolites, macromolecular cellular components are removed and the DNA is eluted under conditions of

low ionic strength. The DNA obtained, with dimensions of 100-300 bp, is suitable for PCR and other enzymatic reactions.

The **RecoverAll™ Total Nucleic Acid Isolation Kit** is designed to extract total nucleic acids (RNA, miRNA, and DNA) from formaldehyde- or paraformaldehyde-fixed, paraffin-embedded (FFPE) tissues. Up to four 20 µm sections, or up to 35 mg of unsectioned core samples, can be processed per reaction. The RecoverAll advantage The RecoverAll Total Nucleic Acid Isolation procedure requires about 45 minutes of hands-on time and can easily be completed in less than one day when isolating RNA. FFPE samples are deparaffinized using a series of xylene and ethanol washes. Next, they are subjected to a rigorous protease digestion with an incubation time tailored for recovery of either RNA or DNA. Nucleic acids are purified using a rapid glass-fiber filter methodology that includes an on-filter nuclease treatment and are eluted into either water or the low salt buffer provided. The recovered nucleic acids are suitable for downstream applications such as microarray analysis, quantitative real-time RT-PCR, and mutation screening. However, sample fixation and storage typically cause nucleic acid fragmentation and modification. The RecoverAll procedure does not affect nucleic acid fragmentation, and some chemical modifications may remain after the procedure, therefore some downstream procedures may require modification for best results.

The ability to isolate nucleic acid that is suitable for molecular analysis from archived tissue samples provides a powerful tool in retrospective studies of diseased tissue at both the genomic and gene expression level. Standard preservation techniques for storage of biological tissue samples use formaldehyde or, less frequently, paraformaldehyde. The high reactivity of these chemicals makes them ideal for maintaining tissue structure and preventing putrefaction, however, tissues preserved in this fashion have historically been thought to be unusable for molecular analysis. Nucleic acid is both trapped and modified by extensive protein-protein and protein-nucleic acid crosslinks. RNA (and DNA) in histological samples typically is fragmented and chemically modified to a degree that renders it incompatible with many molecular analysis techniques. The degree of RNA fragmentation that has already occurred in FFPE samples prior to nucleic acid isolation cannot be changed. However, the protease digestion conditions of the RecoverAll Total Nucleic Acid Isolation Kit are designed to release, in a relatively short period of time (30 min), a maximal amount of RNA fragments of all sizes, including miRNA. The recovered RNA can be used in real-time RT-PCR.

DNA tends not to fragment as easily as RNA. However, the nucleo-histone matrix is quite dense and also appears to be much more reactive to formaldehyde. A much longer protease digestion time (overnight instead of 30 min) is required to release substantial amounts of DNA from the tissue extract. After purification on the glass-fiber filter, DNA recovered with the RecoverAll Total Nucleic Acid Isolation Kit can typically be used for PCR and other downstream applications.

Approaches to liquid biopsy analysis

Circulating tumor cells (CTCs)

Recent advances in technology now permit robust and reproducible detection of circulating tumor cells (CTCs) from a simple blood test. Standardization in methodology has been instrumental in facilitating multicenter trials with the purpose of evaluating the clinical utility of CTCs.

CTC are epithelial tumor cells that have been expunged into the bloodstream and can settle at a secondary site to form metastases. Their presence has been reported to be an independent prognostic marker of relative short overall survival (OS) in several types of cancer, including NSCLC. It is possible that the presence of CTC is a reflection of the tumor burden or invasiveness causing them to be associated with worse survival. These characteristics allow them to be used as a liquid biopsy in a less invasive approach to obtain information on prognosis and treatment management.

Cell free DNA (cfDNA)

cfDNA is extracellular DNA detectable in blood. The presence of cfDNA can be found in patients with malignant pathologies, but also in healthy individuals and in patients with non-malignant diseases, namely, erythematic lupus, rheumatoid arthritis, lung embolism, myocardium infarction, traumas, or invasive therapeutics practices. The size of the DNA released from dead cancer cells varies between small fragments of 70 to 200 base pairs and large fragments of about 21 kb and is longer than that of non-neoplastic DNA. In cancer patients the release of cfDNA may be the result of apoptotic and necrotic processes, which are characteristic of tumors with a high cellular turnover. The concentration of cfDNA in serum of cancer patients is about four times higher than that of healthy controls.

Several driver mutations have been found in NSCLC, such as *EGFR* (epidermal growth factor receptor), *KRAS* (V-Ki-ras2 Kirsten rat sarcoma), *BRAF* (B-Raf proto-oncogene, serine/threonine kinase) or *HER2* (erb-b2 receptor tyrosine kinase 2).

The cobas® cfDNA Sample Preparation Kit is used for manual sample preparation to isolate circulating cell-free DNA (cfDNA) from plasma samples. Plasma samples are processed and cfDNA isolated using the cobas® cfDNA Sample Preparation Kit, a generic manual sample preparation based on nucleic acid binding to glass fibers. Two milliliters (mL) of plasma are processed with a protease and a chaotropic binding buffer that protects the DNA from DNases. Subsequently, isopropanol is added to the binding mixture that is then centrifuged through a column with a glass fiber filter insert. During centrifugation, the cfDNA is bound to the surface of the glass fiber filter. Unbound substances, such as salts, proteins and other impurities, are removed by centrifugation. The adsorbed nucleic acids are washed and then eluted with an aqueous solution.

The **NucleoSpin® cfDNA XS** kit is designed for the efficient isolation of circulating DNA from human blood plasma or other types of cell-free fluids. Fragmented DNA as small as 50–1000 bp can be purified with high efficiency. The special funnel design of the NucleoSpin® cfDNA XS Columns allows very small elution volumes (5–30 µL) resulting in highly concentrated DNA. The protocol follows state-of-the-art bind-wash-elute procedures: After mixing of the sample with the binding buffer, the mixture is applied to the NucleoSpin® cfDNA XS Column. Upon loading of the mixture DNA binds to a silica membrane. Two subsequent washing steps efficiently remove contaminations and highly pure DNA is finally eluted with 5–30 µL elution buffer (5 mM Tris-HCl, pH 8.5).

A good portion of the cell-free DNA in plasma results from apoptotic cells. Therefore, a considerable percentage of this circulating nucleosomal DNA is known to be highly

fragmented. However, the degree of fragmentation and the ratio of fragmented DNA to high molecular weight DNA depends on several parameters like origin of the DNA (e.g., fetal, tumor, microbial DNA), health of the blood donor, blood sampling procedure, and handling of the sample.

The NucleoSpin® cfDNA XS kit is recommended for the isolation of fragmented cell-free DNA from human EDTA plasma, serum, and bronchial lavage. It has been successfully used for other cell-free fluids including urine and follicular fluid. The NucleoSpin® cfDNA XS kit is designed for high recovery, especially of fragmented DNA in a range of 50–1000 bp. Up to 240 µL sample can be used with a single column loading step. DNA yield strongly depends on the individual sample. For plasma it is typically in the range of 0.1 ng up to several 100 ng DNA per mL sample. Up to 720 µL sample can be used with three column loadings. If more than 240 µL sample is processed, additional Lysis Buffer BB is required (see ordering information). Elution can be performed with as little as 5–30 µL elution buffer. DNA is ready to use for downstream applications like real time PCR or others.

In Vitro Amplification (real time-PCR)

The **TRUPCR® EGFR Kit** is an in vitro diagnostic test intended for the qualitative detection of 32 somatic mutations in exons 18-21 of epidermal growth factor receptor (EGFR) gene from tumor tissue DNA (fresh, frozen or formalin fixed paraffin-embedded tissue) or liquid biopsy. Results are intended to aid the clinician in identifying patients with lung cancer who may benefit from treatment with EGFR tyrosine kinase inhibitors.

The TRUPCR® EGFR Kit is based on allele specific amplification and is achieved by ARMS PCR. Taq DNA polymerase is extremely effective at distinguishing between a match and a mismatch at the 3’-end of a PCR primer. Specific mutated sequences can be selectively amplified, even in samples where the majority of the sequences do not carry the mutation.

The kit is designed to selectively amplify mutant specific sequences in samples that contain a mixture of wild-type and mutated DNA. In real-time PCR, the fluorescent signal is generated from the presence of an oligonucleotide probe specific for target DNA sequence. The probe contains a fluorescent dye molecule on its 5’ end and a quencher molecule on its 3’ end. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence primarily by Fluorescence resonance energy transfer (FRET). The probe hybridizes with one of the chains of the amplified fragment. During synthesis of a complementary chain, Taq DNA polymerase which possesses 5' - 3' exonuclease activity cleaves the probe. As a result, the fluorescent dye and quencher dye are separated, and the total fluorescence of reaction volume increases in direct proportion to the number of amplicon copies synthesized during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. The fluorescent signal is measured in each cycle of reaction, and the threshold cycle value is determined from the obtained curve. The threshold cycle is proportional to the initial number of DNA copies in a sample and its value allows qualitative comparisons of analyzed and control samples. The detection is achieved in multiplex reaction using fluorescent probes labelled with FAM, VIC /HEX and ROX/TEXAS RED. The TRUPCR® EGFR Kit comprises of five separate PCR amplification reaction mixes: where mutation-specific reactions (exons 18, 19, 20, and 21 of the EGFR gene) and a reference (wild

type control in exon 2 without any known polymorphism/mutation) is amplified simultaneously in each tube.

The **Cobas® EGFR Mutation Test v2** is a real-time polymerase chain reaction (PCR) test that identifies 42 mutations in exons 18, 19, 20 and 21 of the epidermal growth factor receptor (EGFR) gene, including the T790M resistance mutation. The cobas® EGFR Mutation Test v2 has been clinically validated in multiple clinical trials as a companion diagnostic (CDx) for both 1st and 2nd line EGFR TKI therapy in patients with advanced non-small cell lung cancer (NSCLC). The exceptional robustness and repeatability of the cobas® EGFR Mutation Test v2 has been demonstrated in real-world intra and inter-laboratory studies,3 providing high confidence for labs in using the test. It is designed to enable fast time-to-results testing of both tissue and plasma specimens simultaneously with a single kit (tissue results < 8 hrs, plasma results < 4 hrs). For optimized workflow results, Roche has developed the cobas® DNA Sample Preparation Kit for DNA extraction from formalin-fixed paraffin-embedded tissue (FFPET) and the cobas® cfDNA Sample Preparation Kit for extraction of DNA from plasma.

The **TaqMan® Mutation Detection Assay Competitive Allele-Specific TaqMan® PCR** detect and measure somatic mutations in genes that are associated with cancer. The assays are compatible with different samples types, such as cell lines, FFPE tissue samples, and fresh frozen tissue samples. The TaqMan® Mutation Detection Assays are powered by castPCR™ technology, which refers to Competitive Allele-Specific TaqMan® PCR. The castPCR™ technology is highly specific and sensitive, and it can detect and quantitate rare amounts of mutated DNA in a sample that contains large amounts of normal, wild type gDNA. Mutant allele assays target key somatic mutations in oncogene and tumor suppressor genes. All mutation targets are from the comprehensive Sanger COSMIC database. Target selection was based on frequency of occurrence and input from leading cancer researchers. Gene reference assays detect the genes that the target mutations reside in. They are designed to amplify a mutation-free and polymorphism-free region of the target gene. Each assay contains:

- a locus-specific pair of forward and reverse primers
- a locus specific TaqMan® FAM™ dye-labeled MGB probe

In a mutation detection experiment, a sample of unknown mutation status is run in individual real-time PCRs with one or more assays that target mutant alleles within a gene and the corresponding gene reference assay. After amplification, the CT values are determined by the Applied Biosystems® real-time PCR instrument software. Data files containing the sample CT values can be exported from the real-time PCR instrument software and imported into a data analysis tool. Life Technologies Mutation Detector™ Software is recommended for post-PCR data analysis of mutation detection experiments. In mutation analysis calculations, the difference between the CT value of each mutant allele assay and the CT value of the gene reference assay is calculated. This ΔCT value represents the quantity of the specific mutation allele detected within the sample; it is used to determine sample mutation status by comparison to a predetermined detection ΔCT cutoff value.

Mutation detection and RNA profiling

RNA extraction

The **NucleoSpin® totalRNA FFPE XS** kit provides a convenient, reliable, and fast method to isolate RNA from formalin-fixed, paraffin-embedded (FFPE) tissue. Odorless and non-toxic Paraffin Dissolver replaces the flammable and odorous xylene or d-limonene commonly used for deparaffinization. The tissue sample is then heat incubated with Proteinase K to digest the fixed tissue, release nucleic acids, and gently remove crosslinks. Optimal binding conditions for even small RNA (e.g., miRNA) are adjusted and the lysate is applied to the NucleoSpin® RNA Column / NucleoSpin® RNA FFPE XS Column. RNA is bound to the silica membrane. Residual DNA remaining on the membrane is removed by convenient on-column rDNase digestion. Washing steps remove salts, metabolites, and macromolecular cellular components. Pure RNA is finally eluted in a small volume of RNase-free water, yielding highly concentrated RNA. Nucleic acid preparation using NucleoSpin® totalRNA FFPE (XS) can be performed at room temperature. The eluate, however, should be treated with care. RNA is very sensitive to trace contaminations of RNases, often found on general lab ware, fingerprints, and dust. To ensure RNA stability keep RNA frozen at -20 °C for short-term or -70 °C for long-term storage.

NucleoSpin® totalRNA FFPE (XS) is recommended for the isolation of total RNA from formalin-fixed, paraffin-embedded (FFPE) tissue samples, typically as thin sections (approx. 3–20 µm thickness). Formalin-fixed samples which are not embedded in paraffin can also be used as sample material by omitting the deparaffinization steps. The sample size can be up to ~10 sections (1–20 µm) of FFPE. The amount of embedded tissue can be up to 50 mg for NucleoSpin® totalRNA FFPE or up to 5 mg for NucleoSpin® totalRNA FFPE XS (1x 10 µm section with 1 cm2 tissue is approximately 1 mg). RNA yield strongly depends on sample type, quality, and amount. Furthermore, the procedures of tissue sampling, post sampling delay before fixation, fixation time, embedding, and storage conditions have a high impact on RNA quality and yield. RNA concentration: RNA can be eluted highly concentrated and ready-to-use in a small volume of 30–50 µL (NucleoSpin® totalRNA FFPE) or even 5–30 µL (NucleoSpin® totalRNA FFPE XS). RNA size distribution: RNA isolated from formalin-fixed, paraffin-embedded tissue shows size distribution from 15 to 5,000 bases. Often short sized RNA from ca. 100– 300 bases predominate, especially when the sample material is old. However, samples which were subjected to good tissue fixation, embedding, and storage conditions can yield RNA even larger than 5,000 bases. RNA integrity: RNA Integrity Numbers (RIN) according to Agilent 2100 Bioanalyzer assays depend on sample type and quality. In general, the quality of RNA extracted from FFPE samples is poor. Typical RIN of RNA isolated with NucleoSpin® totalRNA FFPE (XS) kits are in range of 2–6.

RecoverAll™ Total Nucleic Acid Isolation Kit

The RecoverAll™ Total Nucleic Acid Isolation Kit is designed to extract total nucleic acids (RNA, miRNA, and DNA) from formaldehyde- or paraformaldehyde-fixed, paraffin-embedded (FFPE) tissues. Up to four 20 µm sections, or up to 35 mg of unsectioned core samples, can be processed per reaction. The RecoverAll advantage The RecoverAll Total Nucleic Acid Isolation procedure requires about 45 minutes of hands-on time and can easily

be completed in less than one day when isolating RNA (see Figure 1 on page 3). FFPE samples are deparaffinized using a series of xylene and ethanol washes. Next, they are subjected to a rigorous protease digestion with an incubation time tailored for recovery of either RNA or DNA. Nucleic acids are purified using a rapid glass-fiber filter methodology that includes an on-filter nuclease treatment and are eluted into either water or the low salt buffer provided. The recovered nucleic acids are suitable for downstream applications such as microarray analysis, quantitative real-time RT-PCR, and mutation screening. However, sample fixation and storage typically cause nucleic acid fragmentation and modification. The RecoverAll procedure does not affect nucleic acid fragmentation, and some chemical modifications may remain after the procedure, therefore some downstream procedures may require modification for best results.

The ability to isolate nucleic acid that is suitable for molecular analysis from archived tissue samples provides a powerful tool in retrospective studies of diseased tissue at both the genomic and gene expression level. Standard preservation techniques for storage of biological tissue samples use formaldehyde or, less frequently, paraformaldehyde. The high reactivity of these chemicals makes them ideal for maintaining tissue structure and preventing putrefaction, however, tissues preserved in this fashion have historically been thought to be unusable for molecular analysis. Nucleic acid is both trapped and modified by extensive protein-protein and protein-nucleic acid crosslinks. RNA (and DNA) in histological samples typically is fragmented and chemically modified to a degree that renders it incompatible with many molecular analysis techniques. The degree of RNA fragmentation that has already occurred in FFPE samples prior to nucleic acid isolation cannot be changed. However, the protease digestion conditions of the RecoverAll Total Nucleic Acid Isolation Kit are designed to release, in a relatively short period of time (30 min), a maximal amount of RNA fragments of all sizes, including miRNA. The recovered RNA can be used in real-time RT-PCR.

DNA tends not to fragment as easily as RNA. However, the nucleo-histone matrix is quite dense and also appears to be much more reactive to formaldehyde. A much longer protease digestion time (overnight instead of 30 min) is required to release substantial amounts of DNA from the tissue extract. After purification on the glass-fiber filter, DNA recovered with the RecoverAll Total Nucleic Acid Isolation Kit can typically be used for PCR and other downstream applications.

NGS sequencing

The **Archer FusionPlex Protocol** is intended for use with Archer reagent kits and corresponding target-enrichment panels to produce high-complexity libraries for use with Illumina next generation sequencing (NGS) platforms. FusionPlex sequencing data produced by this method should be processed using Archer Analysis software - a complete bioinformatics suite that leverages Anchored Multiplex PCR (AMP™) chemistry to detect unique sequence fragments, thus enabling error correction, read deduplication, and ultimately high-confidence alignment and mutation calling. Archer Analysis takes demultiplexed FASTQ files straight from the sequencer as input, and produces both high level and detailed mutation reporting, as well as raw text and BAM outputs for full transparency of the pipeline.

Test Principle AMP is a rapid and scalable method to generate target-enriched libraries for NGS. AMP technology can be used for applications in targeted RNA sequencing, genomic DNA sequencing and genotyping applications to generate a sequencing library in a matter of hours. Designed for low nucleic acid input, this process delivers robust performance across a variety of sample types. AMP utilizes unidirectional gene-specific primers (GSPs) that enrich

for both known and unknown mutations. Adapters that contain both molecular barcodes and sample indices permit quantitative multiplex data analysis, read deduplication and accurate mutation calling.

The Archer Analysis software utilizes these molecular barcodes for duplicate read binning, error correction and read deduplication to support quantitative multiplex data analysis and confident mutation detection. Analysis reports both sequencing metrics and number of unique observations supporting called variants.

FusionPlex Comprehensive Thyroid and Lung Kit from ArcherDX, is a targeted NGS panel to detect gene fusions, SNV/indels, splicing and gene expression. This panel is expertly designed to cover relevant exons in 34 genes when used in combination with the VariantPlex CTL Panel. Libraries are created by using the FusionPlex assay in conjunction with the Archer MBC Adapters. This panel complements the Archer VariantPlex® CTL Panel for comprehensive mutation profiling of fusions, CNVs and variants.

Studies are performed from RNA extracted from several types of starting sample. Using Archer's proprietary Anchored Multiplex PCR (AMP™)-based enrichment, fusions of all genes in this kit can be identified in a single sequencing assay, even without prior knowledge of fusion partners or breakpoints.

The Archer® FusionPlex® Lung Kit is focused to detect EGFR and MET exon 14 skipping events along with prominent ALK, BRAF, FGFR, NRG1, NTRK, RET, and ROS1 fusions and select point mutations in 14 key gene targets associated with lung cancer.

Targeted RNA-seq to identify oncogenic driver mutations from low-quality RNA

- Detect known and novel gene fusions

- Confirm key point mutations (SNVs/indels)

- Capture RNA abundance & expression imbalance

NGS sequencing targets the following genes: AKT1, ALK, AXL, BRAF, CALCA, CCND1, CTNNB1, DDR2, EGFR, ERBB2, FGFR1, FGFR2, FGFR3, GNAS, HRAS, IDH1, IDH2, KRAS, KRT20, KRT7, MAP2K1, MET, NRAS, NRG1, NTRK1, NTRK2, NTRK3, PIK3CA, PPARG, PTH, RAF1, RET, ROS1, SLC5A5, THADA, TTF1.

Anchored Multiplex PCR (AMP™) chemistry relies on MBC adapters for target amplification. These partially-functional adapters are ligated to cDNA fragments and contain a universal primer binding site that permits amplification of both known and unknown genomic regions of interest. This approach generates libraries with random start sites and varying lengths, increasing library complexity and retaining sample heterogeneity. AMP chemistry can capture both 5' and 3' fusions, including novel fusions that would be missed by opposing primer-based methodologies.

Relative RNA abundance can be determined for select genes, because molecular barcodes are ligated to input material prior to amplification. Knowing RNA abundance helps with tissue of origin identification, expression signature-based differentiation of diffuse large B-cell lymphoma (DLBCL) subtypes and relative expression level detection in critical genes. Relative RNA abundance can be used to measure CD274 (PD-L1) expression levels across 4

FFPE expression standards (Horizon® CD274 (PD-L1) Reference Standard cat# HD787) of varying PD-L1 status and across different DLBCL subtypes.

The analysis platform produces highly confident gene mutation detection from RNA, DNA and ctDNA sources. The simple user interface paired with in-line visualization enables powerful and clear reporting.

Ion AmpliSeq RNA Fusion Lung Cancer Research Panel

The **Ion AmpliSeq Colon and Lung Cancer Research Panel v2** has primer pairs in a single pool for hotspots and targeted regions for 22 known genes associated with colon and lung tumor tissue. In addition, the panel now also includes primers for 3 additional amplicons covering target regions of the NRAS and ALK genes:

- NRAS exon 4 variants (p.117, p.146)
- ALK variants (G1269A, p.S1206Y)

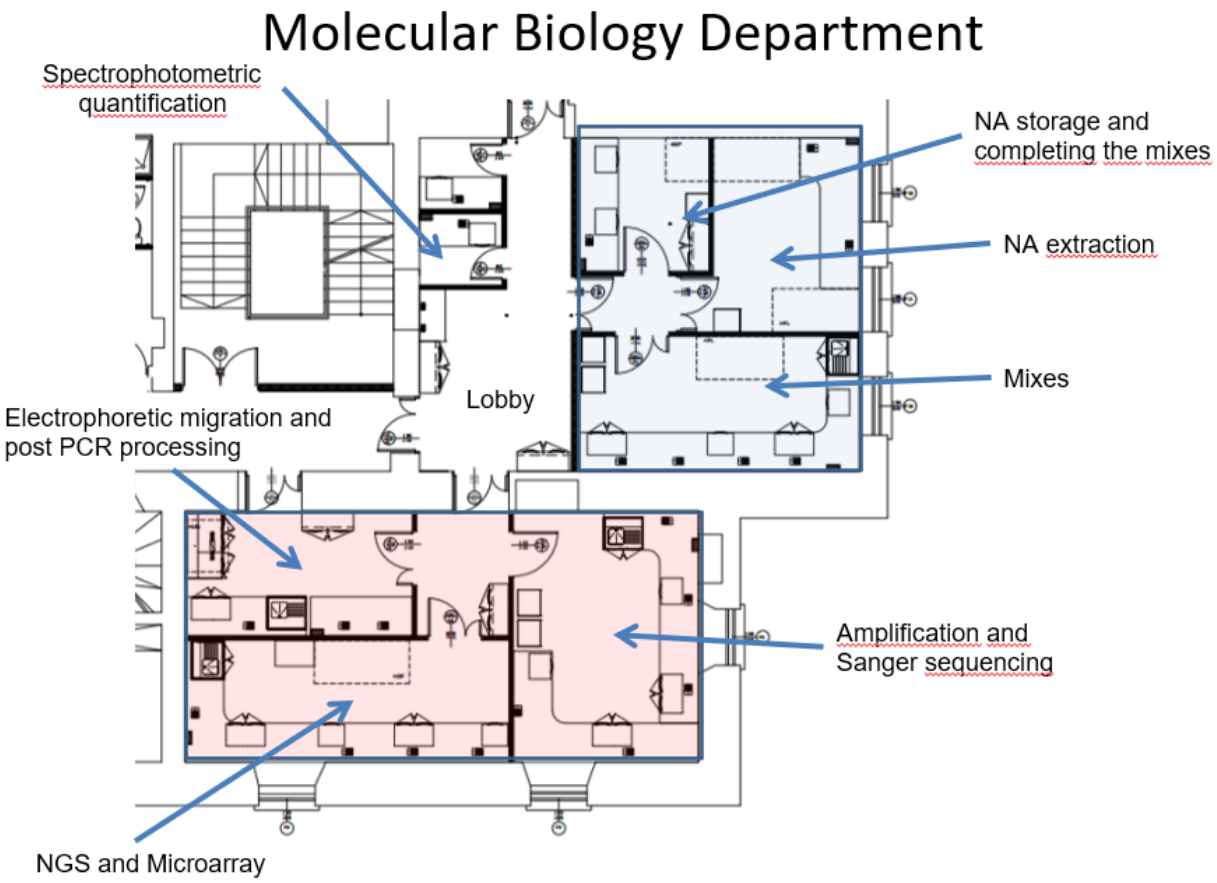
These known rare variants are important in colon and lung cancer research. The first version of the panel, the Ion AmpliSeq Colon and Lung Cancer Research Panel, was tested and verified by the OncoNetwork Consortium, with 155 unique FFPE samples. The Ion AmpliSeq Colon and Lung Cancer Research Panel v2 is ready-to-use and optimized for data analysis with Torrent Suite and Ion Reporter Software.

Table 1. Ion AmpliSeq Colon and Lung Cancer Research Panel v2.

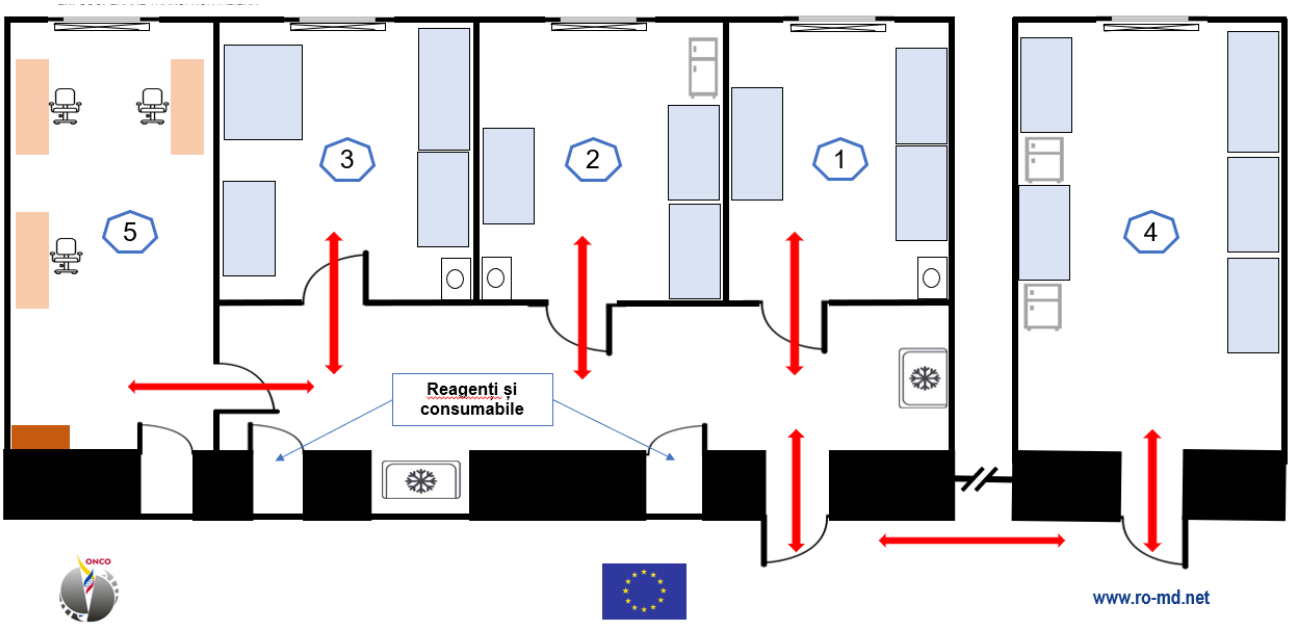
Sample type	FFPE samples
Application	Somatic mutation detection
Genes	*KRAS*, *EGFR*, *BRAF*, *PIK3CA*, *AKT1*, *ERBB2*, *PTEN*, *NRAS*, *STK11*, *MAP2K1*, *ALK*, *DDR2*, *CTNNB1*, *MET*, *TP53*, *SMAD4*, *FBX7*, *FGFR3*, *NOTCH1*, *ERBB4*, *FGFR1*, and *FGFR2*
Primer pairs, amplicon length	92 pairs of primers in a single pool 92 amplicons with an average length of 162 bp
Input DNA required	10 ng
Observed performance	Percent of amplicons with the target base coverage at 500x: >95% Average panel uniformity: 95% Average percent reads on target: 98%
Multiplexing	2 samples per Ion 314™ Chip with at least 500x sequencing coverage 8 samples per Ion 316™ Chip with at least 500x sequencing coverage 16 samples per Ion 318™ Chip with at least 500x sequencing coverage

Research steps

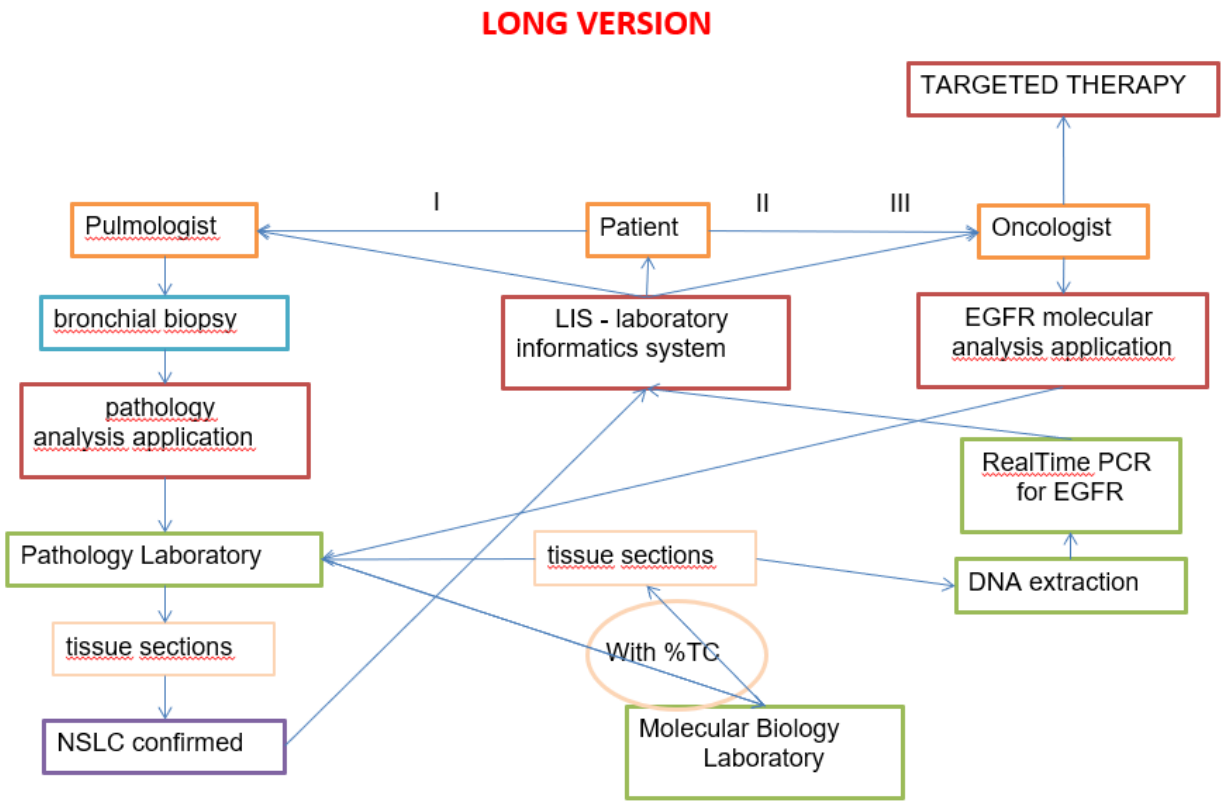
1. **Molecular biology department organization at IRO Iasi**

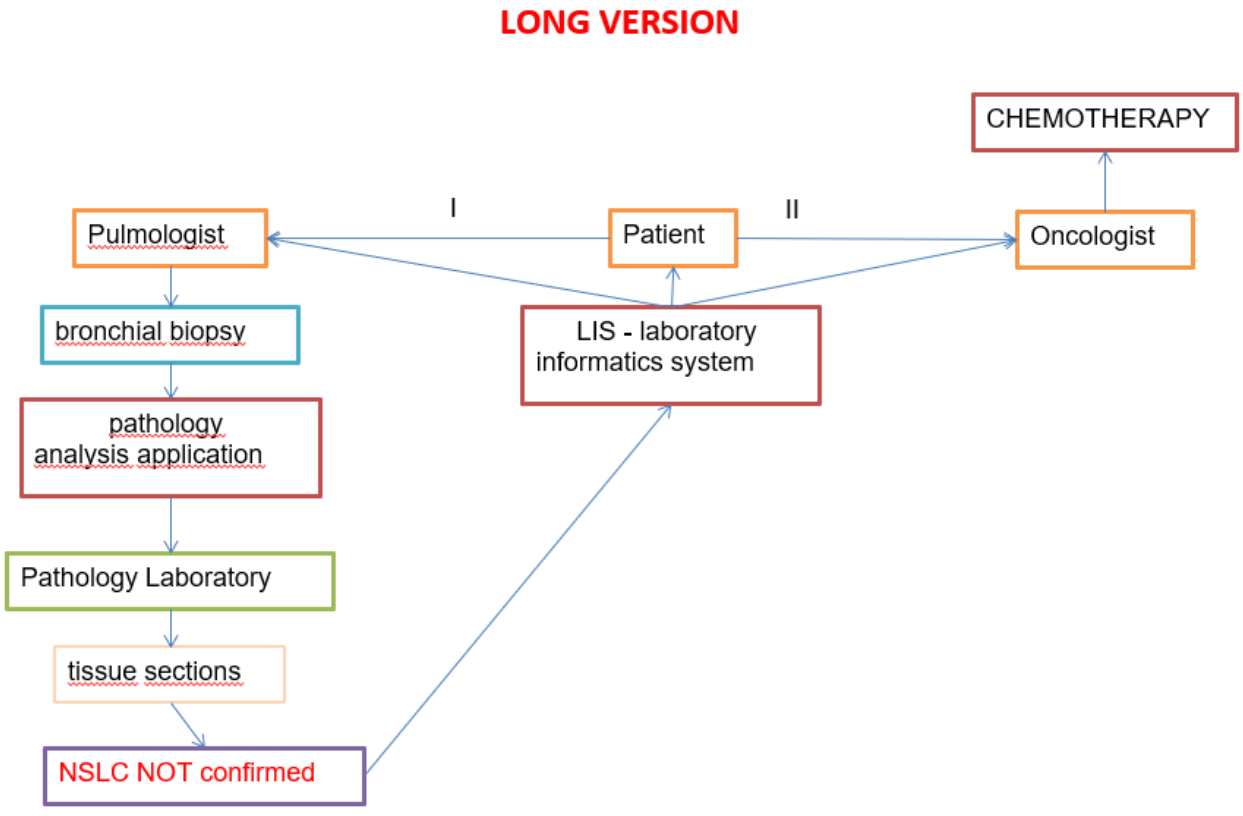


2. **Molecular biology department organization at IOM Chisinau**

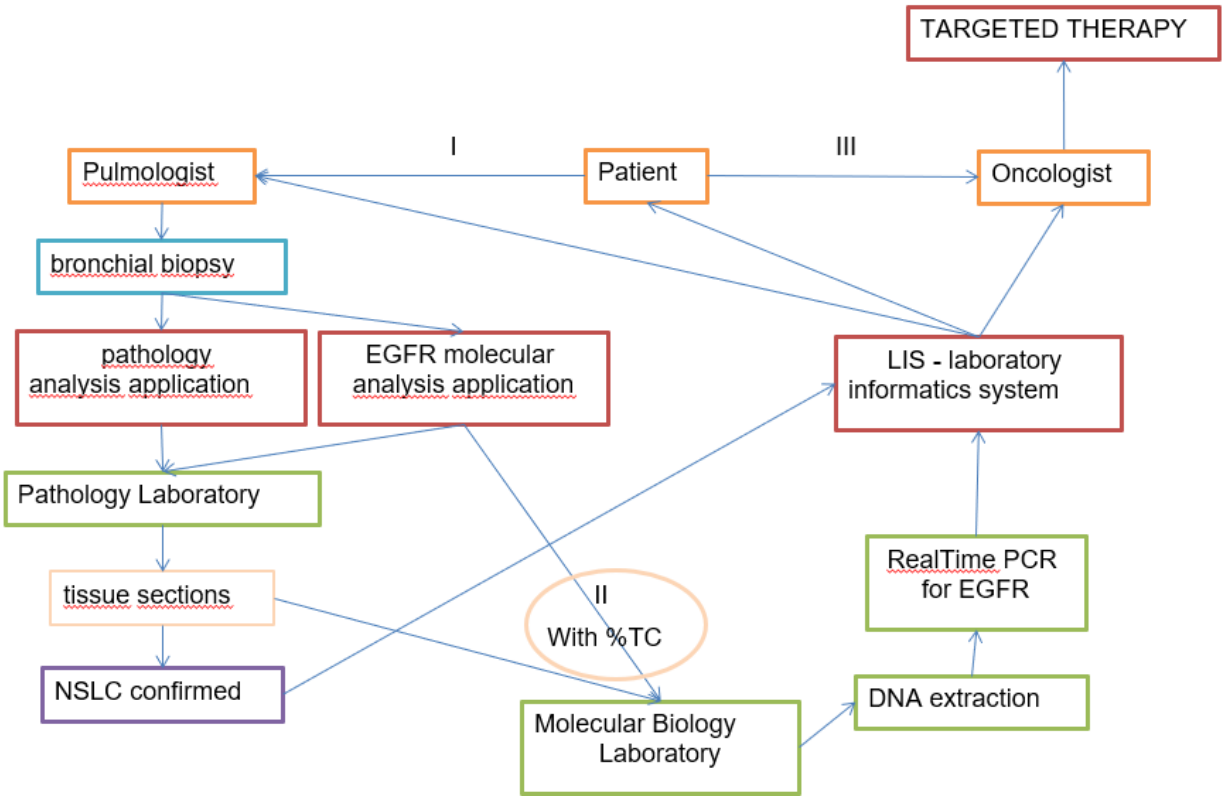


3. Biological sample circuit from sample to result validation

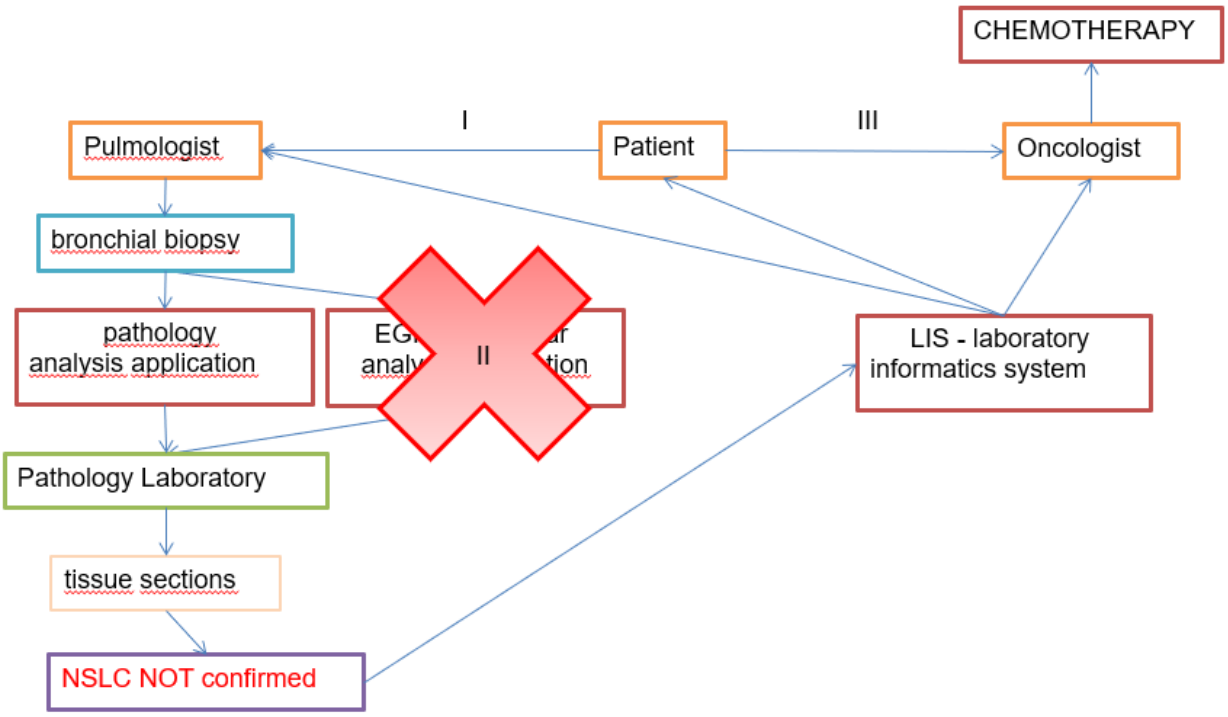


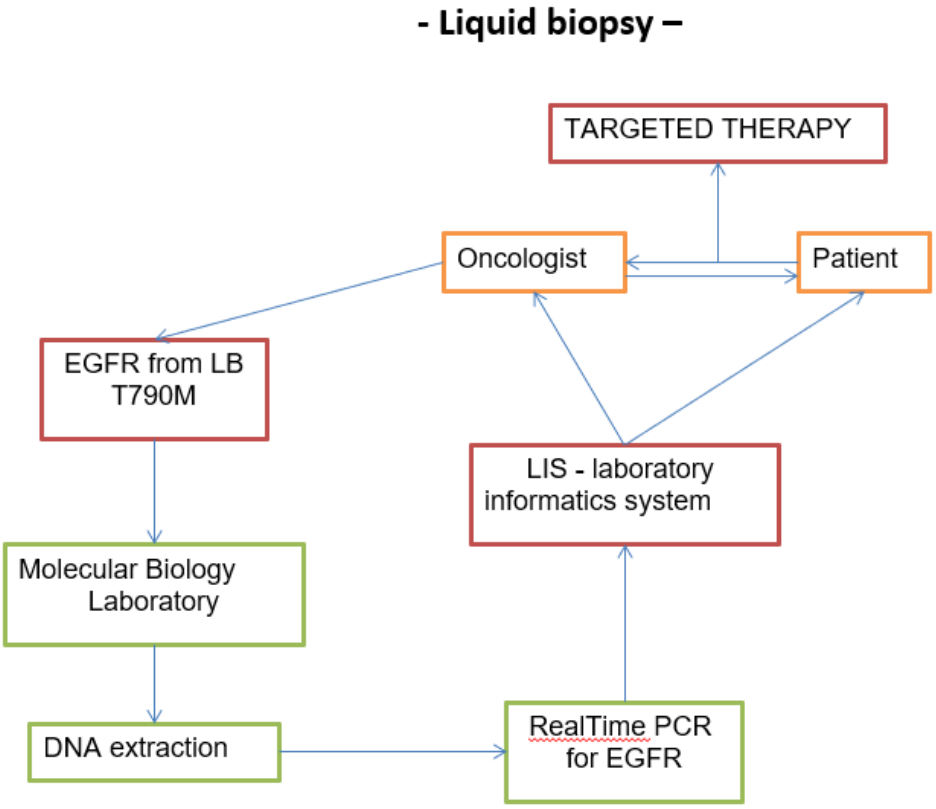


OPTIMIZED VARIANT



OPTIMIZED VARIANT





Results

Performance of DNA and RNA Isolation

DNA was extracted from FFPE tumor samples (*3 paraffin sections of 10 μm*) using the RecoverAll Total Nucleic Acid Isolation Kit (Invitrogen, ThermoFisher Scientific) and from plasma (2 – 4 mL) using MagMAX Cell-Free DNA Isolation Kit (Applied Biosystems, ThermoFisher Scientific) according to the manufacturer's instructions.

RNA was extracted from FFPE tumor samples (*4 paraffin sections of 15 μm*) using the RecoverAll Total Nucleic Acid Isolation Kit (Invitrogen, ThermoFisher Scientific) and from fresh cancer tissue (15 – 25 mg) using the TRIzol (Invitrogen, ThermoFisher Scientific) according to the manufacturer's instructions.

The concentration of isolated DNA and RNA had the following parameters:

- DNA from FFPE: Average – 7.57 ng/μL, Minimum – 0.27 ng/μL, Maximum – 22.4 ng/μL.
- RNA from FFPE: Average – 21.3 ng/μL, Minimum – 4.52 ng/μL, Maximum – 59.80 ng/μL
- RNA from Fresh cancer tissue: Average – 16.31 ng/μL, Minimum – 4.24 ng/μL, Maximum – 59.80 ng/μL

PCR-based testing for therapy-related EGFR mutations

Was tested the following EGFR mutation: ex19del (c.2240_2251del12, c.2237_2255>T, c.2239_2258>CA, c.2240_2257del18, c.2240_2254del15, c.2236_2253del18,

c.2239_2251>C, c.2239_2248TTAAGAGAAG>C, c.2238_2248>GC, c.2236_2250del15, c.2237_2251del15, c.2238_2252>GCA, c.2237_2254del18, c.2239_2247delTTAAGAGAA, c.2235_2252>AAT, c.2238_2255del18, c.2239_2256del18, c.2235_2249del15, c.2238_2252del15), L858R and T790M.

The study included 155 eligible patients with Non-Small Cell Lung Cancer Treatment (NSCLC). Overall, median age at diagnosis was 69.0 years. Among these patients tested, 39.4% (61/155) patients were EGFR mutation-positive. The two most common mutations, L858R (28/61 cases, 45.9%) and DEL19 (31/61 cases, 50.8%). Two from all 61 mutation-positive were T790M positive.

Next-generation sequencing (NGS) Result

Summary samples statistics

*- NOCALL samples excluded

Total sequenced samples	85
Total RNA sequenced samples	55
Unique RNA sequenced samples	40
Samples with NEGATIVE call for fusions	24
Samples with NOCALL for fusions	11
Samples with POSITIVE call for fusions	5
Average number of fusions per sample*	0.17
Tissue samples tested for both: mutations and fusions	9
Total DNA sequenced samples	30
Unique DNA sequenced samples	30
Total mutated genes	18
Average number of mutations per sample	4.9

Summary of sequencing hotspots and targeted regions for 22 known genes associated with lung tumor tissue

Total mutated genes: 18

EGFR, TP53, NOTCH1, CTNNB1, MET, BRAF, KRAS, STK11, DDR2, SMAD4, NRAS , ERBB4, PTEN, AKT1, PIK3CA, FGFR3, FGFR2, FGFR1.

Total number of genes that contains key variants: 13

EGFR, TP53, NOTCH1, CTNNB1, MET, BRAF, KRAS, STK11, ERBB4, PTEN, AKT1, PIK3CA, FGFR3.

Types of variants detected: 91

Key variants: 77

Non-key variants: 14

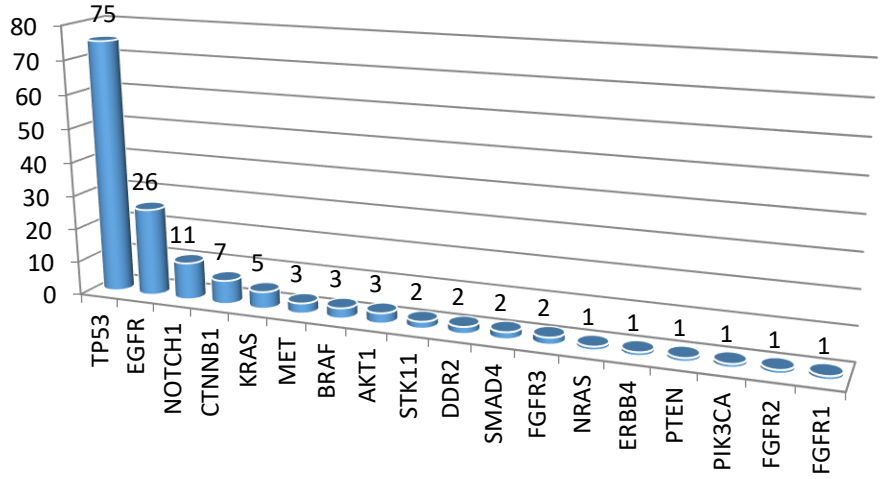
Total variants detected: 147

Average number of variants per sample: 4.9

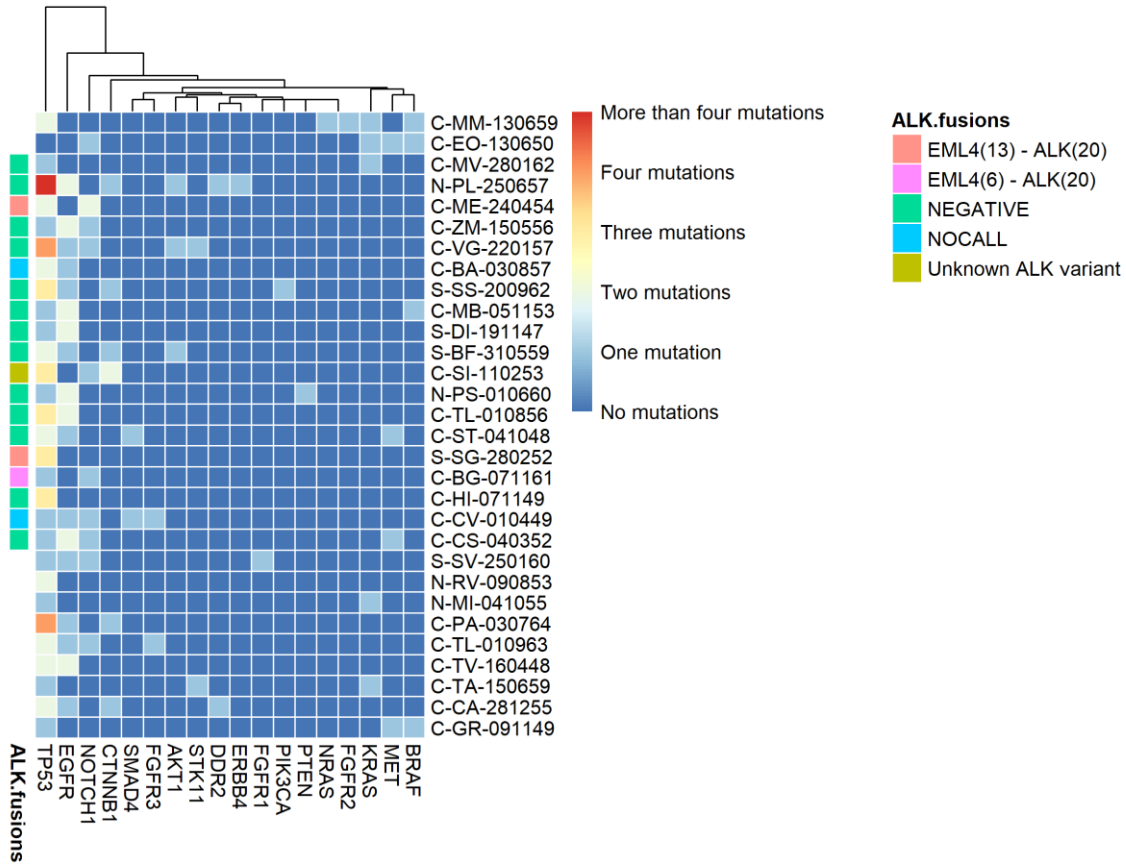
The most genetic variants detected per sample: 28

The fewest genetic variants detected per sample: 2

The most commonly mutated gene in the studied group is TP53 which contains around 51% of all detected variants followed by EGFR with 17.7% and NOTCH1 with 7.5%.



The number of mutations per gene for each analyzed sample (Annotation is performed with results for ALK fusions)



The results show that TP53 is afected in 29 out of 30 samples which represents 96.66%. The sample with the most mutated TP53 contains a number of 22 mutations, while there is only one sample in which were not detected any TP53 variants.

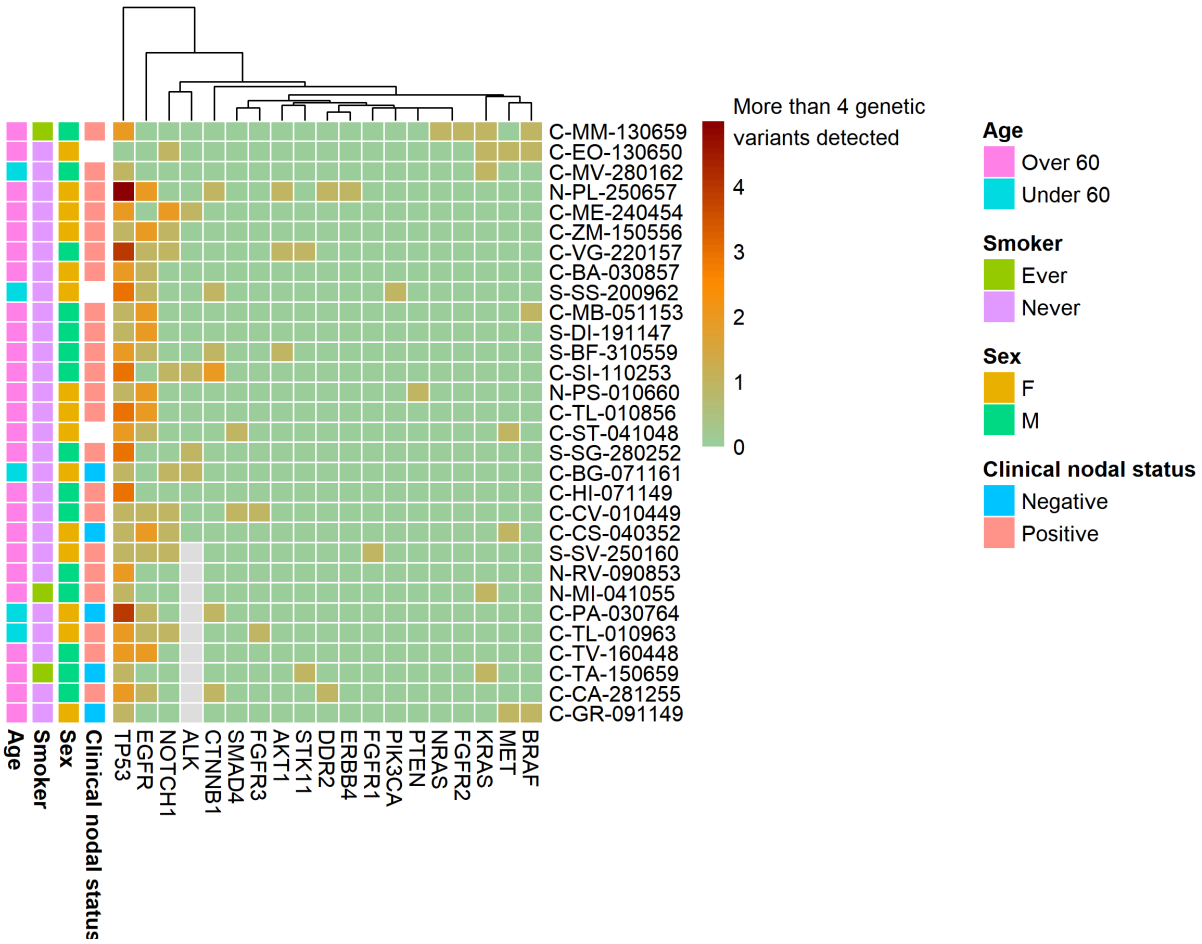
#	DNA samples	Mutations detected	RNA samples	Fusions Overall Call
1	C-ME-240454_dna_v1	*Key variants:* TP53 R213Q, TP53 P278L, NOTCH1 V1578del, NOTCH1 L1585P	C-ME-240454_ft_RNA_v1	POSITIVE EML4(13) - ALK(20)
2	C-SI-110253_dna_v1	*Key variants:* TP53 P77S, TP53 E271K, CTNNB1 A20V, CTNNB1 G34E, NOTCH1 V1578del *Non-key variants:* TP53 p.(P72R)	C-SI-110253_RNA_v1	POSITIVE Unknown ALK variant
3	C-SG-280252_dna_v1	*Key variants:* TP53 E285K, TP53 H193Y *Non-key variants:* TP53 p.(P72R)	C-SG-280252_RNA_v1	POSITIVE EML4(13)-ALK(20) COSF408.1

4	C-BG-071161_dna_v1	*Key variants:* NOTCH1 V1578del *Non-key variants:* TP53 p.(P72R)	C-BG-071161_ft_RNA_v1	POSITIVE EML4(6) - ALK(20)

Based on 3'/5' Imbalance score, was detected a fusion other than those targeted by the panel but still including ALK driver gene.

Also, it can be observed that mutations in NOTCH1 gene are common in DNA samples of patients tested positive for fusions. The NOTCH1 V1578del mutation is present in 3 out of 4 ALK-positive patients. There are no clear observations regarding the co-occurring or mutually exclusive interactions between studied genes.

An interesting observation is that all smoking patients have mutations in the KRAS gene (KRAS G12C and KRAS G12A variants), which corresponds to the data obtained in the previous studies. According to a study conducted by Riely et al (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754127/), KRAS mutations have been strongly associated with cigarette smoking. These mutations are predictive of poor prognosis in resected disease as well as resistance to treatment with erlotinib or gefitinib.



Also, in the heatmap can be distinguished a cluster that shows already mentioned association between NOTCH1 mutations and ALK fusions. No clear molecular differences were observed between sexes, age groups or clinical nodal status, however, given the relatively small number of samples included in the study, these observations should be verified in larger cohorts.

Conclusions

The above curricula contributed to the familiarization of young lab staff (Junior Scientific Researchers, Resident Doctors) with sequencing molecular biology techniques and metagenomic research. Special molecular techniques for diagnosis and management of lung cancer will be used routinely and patients will have access and benefit from target therapies that prolonged the surviving and augment the life. This project also increased researchers' motivation by opening new professional opportunities. Projects stimulate cross border cooperation between research institutes and business/local authorities - the partnerships between the biomedical, clinical and primary and community care research cores that will set up a network of new multidisciplinary research.

Future research directions

Each year, about 11,000 Romanians are diagnosed with lung cancer. Most of affected persons are men and among the male population lung tumors are on the first place having an incidence of 18.8%. While there are multiple types of the disease, non-small cell lung cancer (NSCLC) comprises 85% of all lung cancers. The majority of NSCLC are locally advanced or metastatic disease and most new diagnoses survive less than a year, and among all NSCLC cases. The 5-year survival rate is 16%. The Federation of Cancer Associations from Romania draws attention to the dramatic situation regarding the access of patients diagnosed with lung cancer to innovative therapies, 75% of all cases diagnosed with this disease being identified as inoperable. Lung cancer is ranked first in Romania, both as a number of cases and as a number of deaths among men and is a common cause of mortality among women, standing on the 4th place after breast cancer, cancer cervical cancer and colorectal cancer. In Europe, more than 250,000 people have died annually as a result of diagnosing lung cancer. According to specialty studies, the survival rate for a patients diagnosed with this disease is extremely low at 5 years - for those in advanced stage, the chances of decay are about 18%, while for patients suffering from metastasized pulmonary cancers, at a very late stage advanced, the survival rate is estimated at only 2%. Otherwise, only 1 out of 3 cases of lung cancer are diagnosed at a stage that allows useful therapies. Only 20% of these patients live longer than one year. In the absence of information campaigns for the at-risk population, 3 out of 4 new cases registered for this disease in our country are found in an inoperative stage.

Although gene fusions occur rarely in NSCLC patients, they represent a relevant target in treatment decision algorithms. To date, immunohistochemistry and fluorescence in situ hybridization are the two principal methods used in clinical trials. However, using these methods in routine clinical practice is often impractical and time consuming because they can only analyze single genes and the quantity of tissue material is often insufficient. Thus, novel technologies, able to test multiple genes in a single run with minimal sample input, are being under investigation. The introduction of a targeted therapy is conditional on the acquirement of a sufficiently large tissue sample of adequate quality and requires molecular analysis to be performed. The optimal method for the latter is next-generation sequencing (NGS), which saves both time and the precious tissue material (assessing all genes at once, not one by one). Importantly, introducing NGS into clinical practice creates an opportunity to detect rare alterations such as the MET exon14 skipping mutation (3–4% of NSCLC patients), RET gene fusions (1–2% of NSCLC patients), NTRK gene fusions (approximately 1% of NSCLC patients) and exon 20 insertions of EGFR gene (1–2%) or HER2 gene mutations (2–4% of NSCLC patients). Although these particular genetic alterations in NSCLC are generally rare, about half of the all-NSCLC patients have molecular alterations and may ultimately receive targeted therapies.

A particularly crucial role in the quality and outcome of the treatment of oncologic patients is played by the early diagnosis (Stage I-II) of malignant tumors but also the molecular subtyping of cancers. Thus, the efficiency of the treatment and the prognosis of the disease rely heavily on an in-depth analysis of the genetic mutational aspects and of the expression patterns of the genes involved in tumor formation and progression.

Selective RET inhibitors are currently studied, but the limitation of various test approaches, coupled with lack of knowledge of acquired resistance mechanisms, and specific patient groups that bear special consideration, consists a challenge. Thus, there is a need of various diagnostic techniques to provide evidence to guide management of RET-fusion-positive Non-Small Cell Lung Cancer (NSCLC) patients, including specific patient groups, such as EGFR-mutant NSCLC patients who acquired RET fusions after resisting EGFR TKIs, and offer a compendium of mechanisms of acquired resistance to RET targeted therapies

Taking into consideration the statistics of cancer cases in Romania, there is an urgent need to develop and implement studies focused on cancer genetics at the level of mutations as well as genetic networks and gene expression patterns. With the experience and promising results acquired in this project, we can further investigate NSCLC biomarkers by using NGS technology in order to develop targeted therapies. Patients with molecularly targetable disease have a chance of avoiding chemotherapy in their first line treatment, and foremost, this treatment is well tolerated, giving them a chance to have a normal family, social and professional life. Searching for molecular aberrations and implementing targeted therapy if appropriate should be a key diagnostic approach for NSCLC patients.

Iași Regional Institute of Oncology, as a beneficiary, launches the project "Network of excellence for diagnosis and research in lung cancer", acronym LUNGNEX-RD - Project funded under the ROMD Program 2014-2020, through the European Neighborhood Instrument (ENI) and co-financed by the project partners, based on the contract.

"This publication/material has been produced with the assistance of the European Union. The contents of this publication are the sole responsibility of Iași Regional Institute of Oncology and can in no way be taken to reflect the views of the European Union or of the Joint Operational Programme Romania-Moldova 2014-2020 management structures."

Institutul Regional de Oncologie Iasi IRO

www.ro-md.net

A project implemented by
Regional Institute of Oncology Iași
Adress: Str. G-ral Henri Mathias Berthelot nr. 2-4
Phone: +40374278810
Fax: +40374278802

This project is funded by the European Union

"The European Union is made up of 28 Member States who have decided to gradually link together their knowhow, resources and destinies. Together, during a period of enlarge- ment of 50 years, they have built a zone of stability, democra- cy and sustainable development whilst maintaining cultural diversity, tolerance and individual freedoms. The European Union is committed to sharing its achievements and its values with countries and peoples beyond its borders".