UNIVERSITÀ DI SIENA 1240

Department of Medical Biotechnologies

Doctorate in Genetics, Oncology and Clinical Medicine

Cycle XXXVIII°

Coordinator: Professor Meloni Ilaria

Liquid biopsy for the evaluation of molecular alterations in human cancer

Disciplinary Scientific Sector: MEDS-02/A General Pathology

Candidate
Michele Santillo
Department of Experimental and Clinical Medicine, University of Florence

Firma digitale del/della candidato/a

Supervisor
Dott. Lastraioli Elena
Department of Experimental and Clinical Medicine, University of Florence

26/11/2025 17:28:07
MICHELE SANTILLO

Academic year of obtaining the degree of PhD
2025/26

"I...a universe of atoms, an atom in the universe."
— Richard P. Feynman

Alla mia famiglia
e a chi non ha mai smesso di credere in me

Table of Contents

ABSTRACT

Liquid biopsy can detect molecular biomarkers in body fluids. The profile of circulating tumour DNA (ctDNA) allows to identify the risk of recurrence, therapeutic targets and drug resistance. This project was aimed to analyse the mutational status of *RAS*, *BRAF* and *EGFR* genes from blood samples of patients with metastatic colorectal and lung cancer, to characterise cancers through metabolites and extracellular vesicles (EVs) to exploit novel potential biomarkers. For these purposes, blood samples were analysed using the BEAMing technology and qPCR with Idylla ™ Mutation Test. Concordance analyses between tissue sample plasma generally showed high concordance and substantial agreement. Moreover, *RAS* and *EGFR* mutational profiles were analysed during therapy and several associations with clinical features emerged. We then investigated the profile of circulating miRNAs in patients with early and established form of Idiopathic Pulmonary Fibrosis (IPF) along with lung cancer patients, through small RNA sequencing and uncovered important differences between the groups and with healthy controls: in particular, 30 differentially expressed miRNAs, 21 up-regulated and 9 down-regulated in IPF compared to healthy subjects; in patients with E-IPF compared to IPF 127 up-regulated and 109 down-regulated differentially expressing miRNAs were identified while 20 ci-miRNA were down-regulated in E-IPF compared to controls and moreover 43 downregulated in LC patients compared to controls. Moreover, a functional analysis using Gene Ontology revealed altered terms related to signalling mechanisms mediated by GTPases, cell adhesion molecules, MAPKs and Ras signalling. Analysis through KEGG ontology showed significant enrichment for the term NSCLC, specifically relevant genes (ALK, *EGFR* kinases, etc.). EVs in the plasma of mCRC and mLC patients are significantly more abundant than in controls. Overall, these findings suggest that mutational testing of plasma is a valuable tool for the clinical management of lung cancer and colorectal cancer. Validation in larger cohorts is essential to evaluate concordance rates, establish sensitivity thresholds, and identify correlations with clinical outcomes that will guide best treatment option for patients. In conclusion, our data supports the integration of liquid biopsy in the routine of clinical practice for patients with metastatic colorectal and lung cancers. Early detection of resistance through liquid biopsy may change treatment paradigms, enabling more precise decisions during treatment decisions and improving patient outcomes.

I – INTRODUCTION

I-1 Cancer Hallmarks

The "Hallmarks of Cancer" are biological characteristics that tumour cells acquire to become malignant and grow uncontrollably. These hallmarks provide a conceptual framework for understanding the complexity of cancer biology. Originally identified as six (Hanahan & Weinberg, 2000) the hallmarks have been expanded to include new features.

The classic criteria are (**Figure 1**):
-Autonomy in growth signals, cancer cells may be able to create their own growth signals without consideration for the body's usual limitations;
-Resistance to negative growth modulation, cancer cells ignore signals that normally limit cell division and survival;
-Evasion of apoptosis, cancer cells avoid cell death even when they should be killed off because of mutations or damage;
- Unlimited replicative potential, cancerous cells can divide endlessly beyond normal cell boundaries;
-Continuous angiogenesis: the tumour-cells can produce new blood-vessels for its supporting as they grow more after vessels;
-Tissue invasion and metastasis, when tumour cells penetrate other tissues escaping from primary site of growth and can spread to organs and tissues throughout the body, leading to new secondary growths.

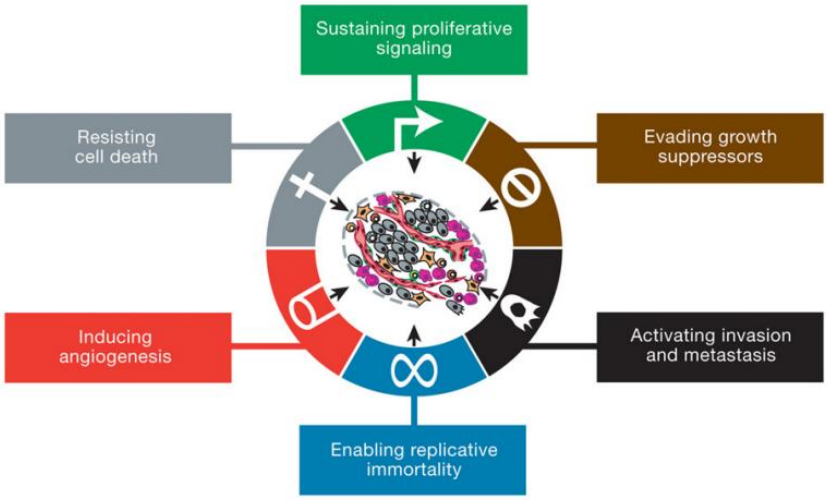


Figure 1: Acquired Capabilities of Cancer cells (Hanahan & Weinberg, 2000).

Subsequently additional hallmarks were proposed in the following years (Hanahan & Weinberg, 2000, **Figure 2**): -Genomic and Epigenetic Instability, cancerous cells can have distortedness in their DNA and gene activity that participate for its development and spread;
-Immune Evasion: cancer behaviours can escape from immuno-surveillance so that the growth and metastasis of the tumour develop without being inhibited;
-Inflammation, especially chronic inflammation caused in the promotion and progression of cancer;
-Phenotypic Plasticity: cancer cells can shift from one phenotype (phenotype: a state of appearance and behaviour) to another, becoming more aggressive or adapting to the new environment;

-Reprogramming Metabolism Energy, cancer cells may reprogram their metabolism to further fuel the rapid and abnormal growth;
-Tumour Microenvironment, interactions between cancer cells and the surrounding tissue (microenvironment) can contribute to their growth, survival and metastasis.

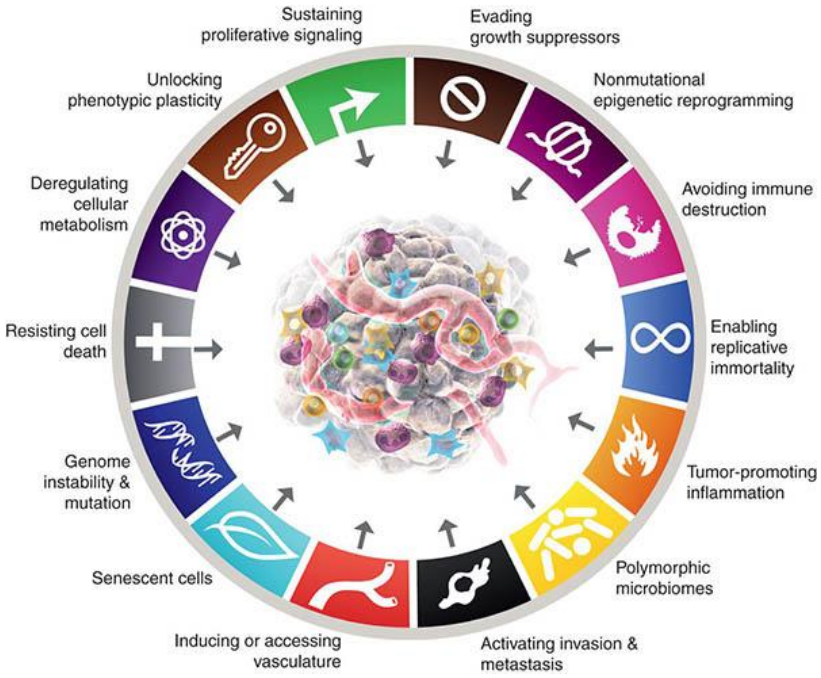


Figure 2: Hallmarks of Cancer—new additions (Hanahan et. al, 2022).

These novel and additional hallmarks are based on a complex interaction of activating signals regardless (exceptions e.g. aneuploidy) these are extremely fundamental concepts that contribute to the understanding of how normal cells can be transformed into cancer cells, as well as, in part, to clarify the type of biological obstacles that the tumours should overcome when managing them. To design new therapies and attack cancer new mechanisms are needed to be understood.

I-2: The Shifting Landscape of Cancer Diagnostics from Biopsies to the Bloodstream

Nowadays, the two kinds of biopsies are being used to identify cancer at an earlier stage and with higher precision than previous methods. Improvements in technology to read DNA now enable scientists to find individual signatures caused by different carcinogens. These molecular fingerprints, as on ribosomes or DNA, can be diagnosed and small devices would be able to read these fingerprints with extremely high fidelity which allow for novel non-invasive cancer tests that work in few hours (Milenkovic et al., 2025).

The transition from tissue-based biopsies to the bloodstream marks one of the most significant transformations in cancer diagnostic strategies, enabling blood tests which are minimally invasive and test for circulating biomarkers. Biopsies of tissue are like snapping photographs, where we can get detailed images of the tumour once, maybe twice in a lifetime, while liquid biopsies are more analogous to video recording; they can provide both an ongoing summary as well as unique real-time insights that cannot be obtained from one

sample or scan alone. The potential for non-invasive, repeatable, and comprehensive insights makes them a revolutionary tool for personalized medicine (Abdelwahab et al., 2024). On one side the Traditional Biopsy requires a surgical procedure to remove a tissue sample for analysis, providing detailed but limited information about the tumour. On the other side Liquid Biopsyanalyses body fluids, primarily blood, for cancer-related biomarkers that shed from tumours into the circulation. Both the tests are supplementary to each other but it's a liquid biopsy which has gone leaps and bounds apart because of several advantages like (i) early detection, detecting cancer before symptoms develop or tumours can be detected in imaging scans; (ii) real-time monitoring, can monitor disease progression and therapeutic efficacy at regular intervals to observe dynamic changes in tumour; (iii) precision treatment, identify precise genetic mutations in guiding which targeted therapies to use and enable treatment plans to adjust as tumours evolve; (iv) monitoring Resistance, follows the evolving changes in the genetic mutations that cause drug resistance, which can be invaluable in refining treatment strategies; (v) reduced risk, provides a safer and less uncomfortable alternative to invasive procedures which improves the patient's safety and comfort (L. Ma et al., 2024).

I-3: The Burden of Colorectal and Lung Cancers

I-3.1 Colorectal Cancer: Epidemiology, Characteristics, and Risk Factors

Colorectal cancer (CRC) is a significant worldwide health problem because it is the third most common malignant tumour and account for approximately 10% of all cancer diagnoses, it is even the second major cause of death related to cancer, responsible for nearly 9–10% of global cancer mortality. According to GLOBOCAN 2022 (**Figure 3**), around 1.9 million new CRC cases and 935,000 deaths were estimated in 2020, with slight differences between men (10.6% incidence, 9.3% mortality) and women (9.4% incidence, 9.5% mortality) (Ferlay et al., 2021; Sung et al., 2021).

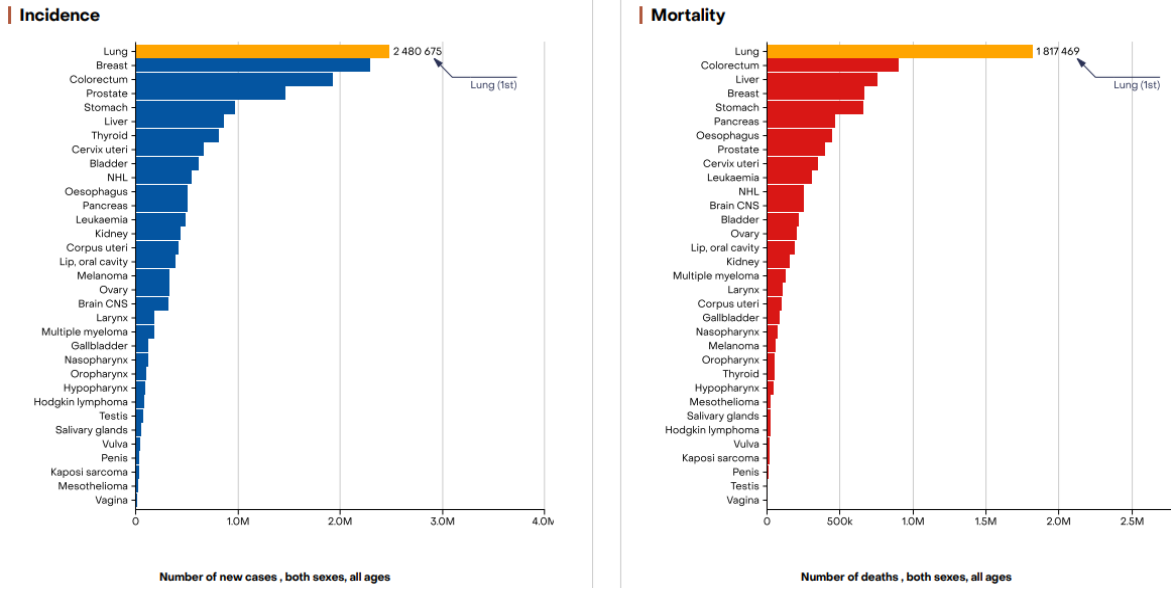


Figure 3: **Cancer site ranking** - Colorectum cancer number of new cases and number of deaths in both sexes and all ages from all over the world (International Agency for Research on Cancer- World Health Organization; Global Cancer Observatory, GCO- https://gco.iarc.who.int).

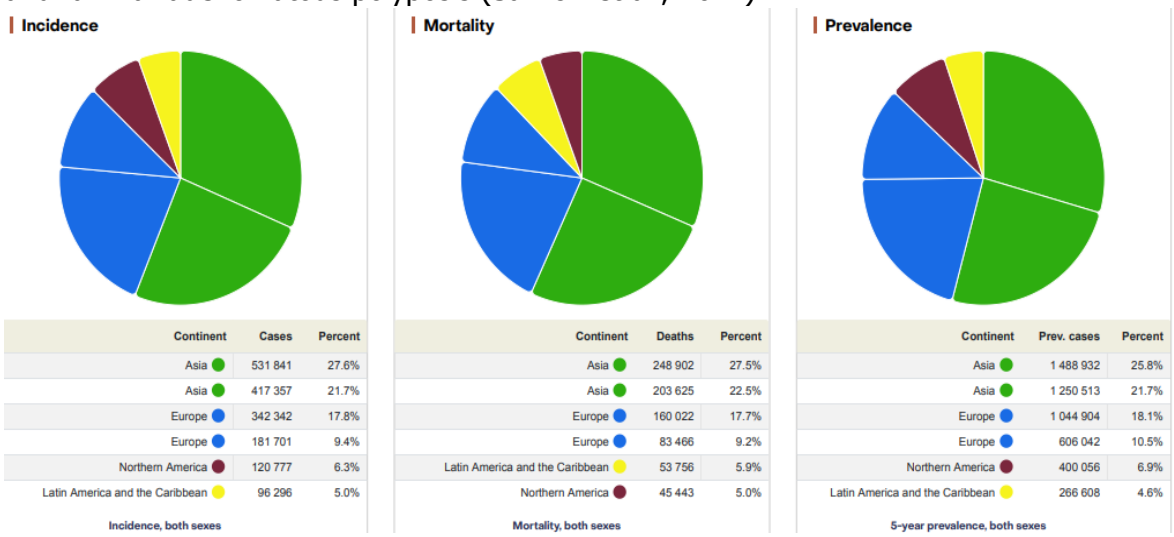
CRC may originate from the colon, the longest portion of the large intestine, or in the rectum the last intestinal part located near the anus. Anatomical location influences the clinical onset and treatment: e.g. cancers that arise in the proximal (right) colon differ biologically and symptomatically from those that arise in the distal (left) colon or rectum (Dekker et al., 2019).

CRC incidence and mortality are higher in high- and upper-middle income countries, which together account for approximately 88% of cases and 85% of deaths. Despite this, high-income countries report lower mortality rates, probably due to better access to screening and treatment services. Conversely, countries that are in development have experienced a rise in incidence and mortality in recent years. CRC is therefore also considered a marker of socio-economic development, with risk factors that drives lifestyle westernization (e.g. obesity, physical inactivity, alcohol, and tobacco consumption) (Sung et al., 2021).

The disease affects primarily individuals aged between 60–75 years, with the highest burden seen in Asia (44%) and Europe (32%) (**Figure 4**).Some countries have reported a 1–4% increase in incidence per year amongst those under the age of 50, suggesting evolving risk exposures in younger populations. With the contribution of raising risk factors and enhanced screening programs are estimates that the global annual CRC incidence could increase from 2 to 5 million by 2035 (Ferlay et al., 2021; Sung et al., 2021).

Risk Factors

Also familiar history represents a relevant risk factor for CRC, as CRC are related to positive familial history and 6–10% is connected with hereditary syndromes such as Lynch syndrome and familiar adenomatous polyposis (Sawicki et al., 2021).



Continent	Cases	Percent
Asia	531 841	27.6%
Asia	417 357	21.7%
Europe	342 342	17.8%
Europe	181 701	9.4%
Northern America	120 777	6.3%
Latin America and the Caribbean	96 296	5.0%

Incidence, both sexes

Continent	Deaths	Percent
Asia	248 902	27.5%
Asia	203 625	22.5%
Europe	160 022	17.7%
Europe	83 466	9.2%
Latin America and the Caribbean	53 756	5.9%
Northern America	45 443	5.0%

Mortality, both sexes

Continent	Prev. cases	Percent
Asia	1 488 932	25.8%
Asia	1 250 513	21.7%
Europe	1 044 904	18.1%
Europe	606 042	10.5%
Northern America	400 056	6.9%
Latin America and the Caribbean	266 608	4.6%

5-year prevalence, both sexes

Figure 4: CRC worldwide - Incidence**,** Mortality and 5-year prevalence in both sexes (https://gco.iarc.who.int).

However, most CRC cases are sporadic, that arise without a clear inherited component. That is due mainly because CRC is a multifactorial disease, influenced so by many factors like genetics, lifestyle, and environmental components. The main risk factors include (Sawicki et al., 2021):

-Age, most diagnoses occur in individuals over 50;
-Family history, first-degree relatives with CRC significantly have increased risk;
-Diet, red meat high consumption and low fibre intake contribute to higher risk;
-Alcohol and tobacco, both are established carcinogens for CRC;
-Obesity and sedentary lifestyle, associated with chronic inflammation and insulin resistance;
-Diabetes mellitus, or hyperinsulinemia and metabolic changes increase carcinogenesis risk;
-Ethnicity, black populations show a 50% higher risk compared to other groups;
-Inflammatory bowel disease, Ulcerative colitis and Crohn's disease elevate the risk through chronic intestinal inflammation;
- Intestinal microbiota, changes in the composition and function of the microbiota can lead to the onset and even progression of cancer (Sawicki et al., 2021).

CRC can be categorized as sporadic, hereditary, or familial based on the type of genetic alteration. For example, Sporadic CRC often arises through a heterogeneous pathogenesis, where point mutations affect various target genes, compromising individual cells and their progeny. Hereditary CRC, on the other hand, results from germline mutations in DNA repair or tumour suppressor genes, as observed in familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC, or Lynch syndrome). Familial cases instead are thought to emerge from inherited susceptibility combined with environmental factors (Mármol et al., 2017).

CRC is one of the cancers with the highest mutational burden and is therefore often described as hypermutated (>12 mutations per 10^6 bases) or non-hypermutated (<8.24 mutations per 10^6 bases). Genetically, CRC is highly heterogeneous, with genomic instability recognized as a major driver. Three principal pathways may cause this instability: chromosomal instability (CIN), microsatellite instability (MSI), and epigenetic instability associated with the CpG island methylation phenotype (CIMP). CIN, responsible for ~85% of sporadic CRC cases, involves alterations in chromosome number or structure that lead to deregulated expression of genes which control cell proliferation. MSI arises from defects in mismatch repair (MMR) genes, which generate hypermutable phenotypes that affect oncogenes and tumour suppressors. CIMP instead is driven by CpG island hypermethylation in promoter regions, causing the silencing of some tumour suppressor genes (Hossain et al., 2022) (Sawicki et al., 2021).

The cells of origin of most CRC cases, are thought to be represented by Colorectal stem cells, located at the crypt bases. Their capacity for self-renewal facilitates mutation accumulation, eventually leading to neoplastic transformation then CRC typically develops through four stages: initiation (irreversible genetic damage), promotion (abnormal cell growth), progression (acquisition of malignant traits), and metastasis (invasion of blood and distant organs, frequently the liver and lungs). In fact, precursor lesions as Adenomas, emerge through DNA repair defects and abnormal proliferation, but often requires 10–15 years to progress into invasive carcinoma (Sawicki et al., 2021).

CRC development can be classified in two main histogenetic pathways (**Figure 5**): (i) the adenoma–carcinoma sequence (70–90% of cases), where progressive genomic instability drives adenoma transformation into invasive carcinoma; (ii) the serrated neoplasia pathway (10–20% of cases), characterised instead by serrated polyps (e.g., sessile serrated adenomas), different genetic mutations and CpG methylation (Testa et al., 2018).

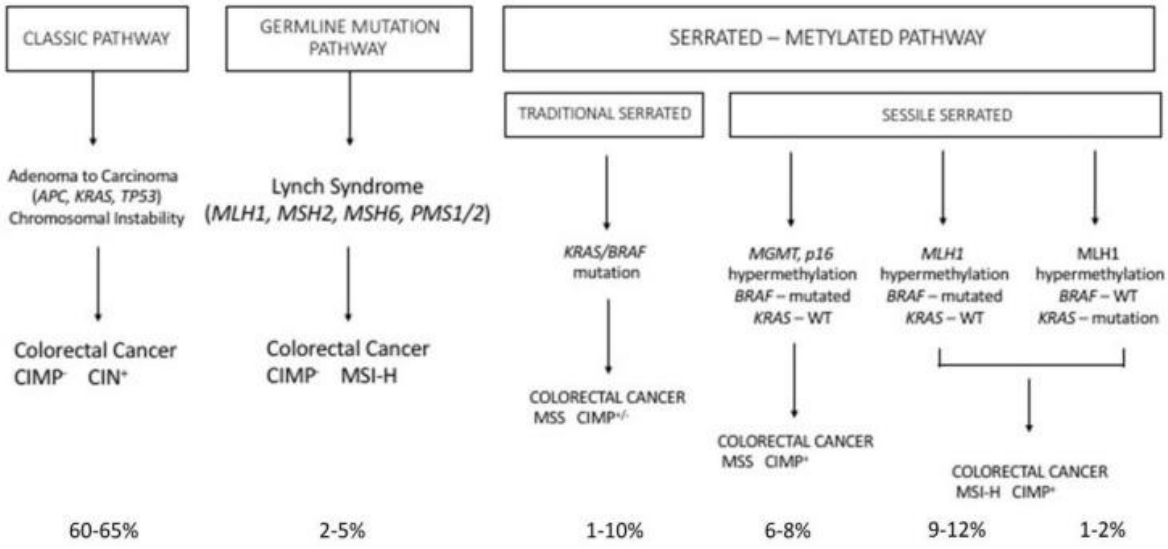


Figure 5: Outline of the main pathways of colorectal carcinogenesis (Testa et al., 2018).

Beyond histological pathways, CRC is also classified at the molecular level. According to the CRC Subtyping Consortium (2015), CRC can be grouped into four consensus molecular subtypes (CMS): CMS1 (Immunogenic) with high immune infiltration and often linked to MSI; CMS2 (Canonical) associated with WNT and RAS pathway mutations; CMS3 (Metabolic), characterized by metabolic reprogramming and frequent BRAF mutations; CMS4 (Mesenchymal) that involves TGF-β activation, with stromal infiltration and frequent metastasis (Guinney et al., 2015).

These molecular subtypes have prognostic value and help guiding therapeutic strategies, but histological grading may help further to refine prognosis which reflect better the degree of tumour differentiation: Grade 1, well-differentiated and slow growth; Grade 2, moderately differentiated; Grade 3, poorly differentiated and rapid growth; Grade 4 (Anaplastic), undifferentiated and highly aggressive. Taken together, this molecular and clinical information provide a more comprehensive understanding of CRC onset and progression, to enable at the same time some more precise diagnostic, therapeutic, or preventive approaches (Brierley J. et al. 2006).

I-3.2 Lung Cancer: Epidemiology and Risk Factors

Currently, lung cancer is the most frequent diagnosed malignancy worldwide and remains the leading cause of cancer-related death (Smolarz et al., 2025). In 2022, it accounted approximately 2.5 million new cases (12.4% of all cancers) and almost 1.8 million deaths (18.7%), confirming its position as the deadliest cancer globally (**Figure 6**) (Ren et al., 2024).

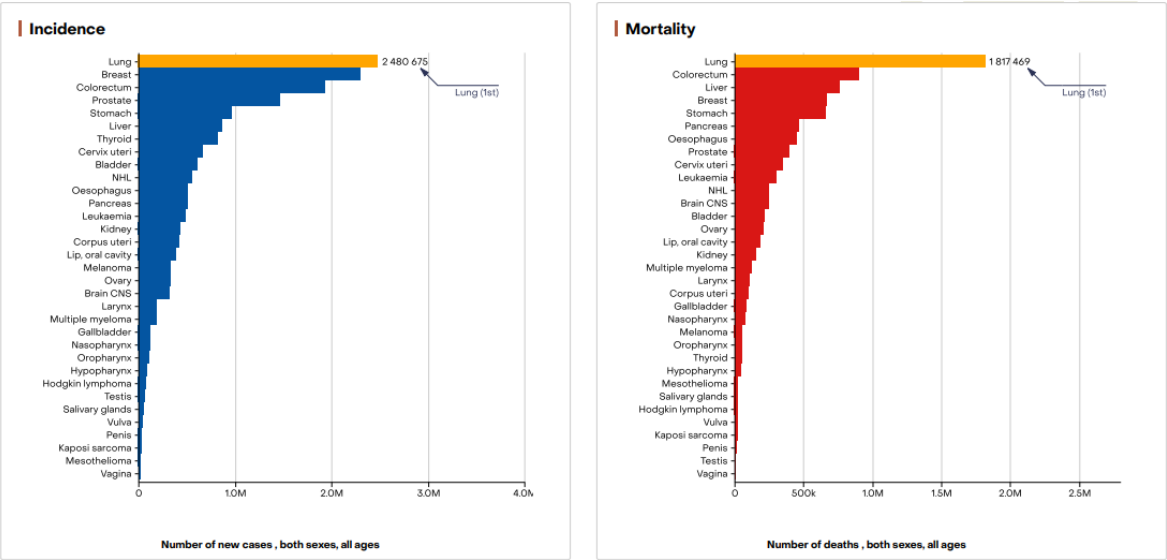


Figure 6: Cancer site ranking – Lung Cancer number of new cases and number of deaths in both sexes and all ages from all over the world (International Agency for Research on Cancer- World Health Organization; Global Cancer Observatory, GCO- https://gco.iarc.who.int).

Histological Subtypes

Clinically, lung cancers are classified into two major groups but together represent over 95% of all cases. Small cell lung cancer (SCLC), also known as microcytoma, constitutes 10–15% of cases. It usually originates in the bronchi of large calibre, is strongly associated with smoking and has a particularly poor prognosis because of its rapid progression and early dissemination to distant organs. On the other side, non-small cell lung cancer (NSCLC) accounts for about 85% of cases and comprehend three main histological subtypes (**Figure 7**):

-Squamous cell carcinoma: representing 25–30% of NSCLCs, it typically arises in medium/large bronchi and is considered the subtype with most favourable prognosis;

-Adenocarcinoma: the most frequent subtype among non-smokers, often associated with pulmonary scars, and more commonly located in peripheral and smaller bronchi;

-Large cell carcinoma: less frequent (~10%) but usually aggressive, with a tendence to grow and spread rapidly across different lung regions (Dekker et al., 2019).

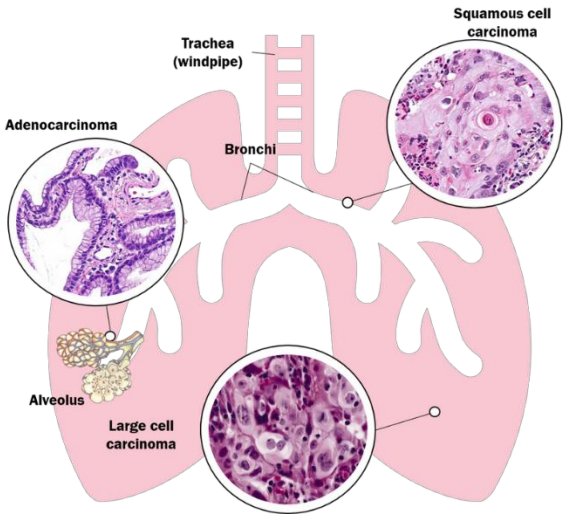


Figure 7: Types of non-small cell lung cancer subtypes. Images adapted from Wikimedia Commons (https://www.ourcancerstories.com/lung-cancer/general/large-cell-carcinoma).

All these histological variants arise from the epithelial lining of pulmonary structures, as shown in histological and anatomical studies of the lung parenchyma.

Epidemiology

According to IARC's Global Cancer Observatory, lung cancer was the most diagnosed cancer worldwide in 2022. Lung cancer incidence shows evident geographical variability with rates that exceed 500 cases per 100,000 inhabitants in Oceania, compared to less than 100 per 100,000 in West Africa. Among women, incidence ranges from over 400 per 100,000 in Oceania to approximately 100 per 100,000 in East-Central Asia (Smolarz et al., 2025).

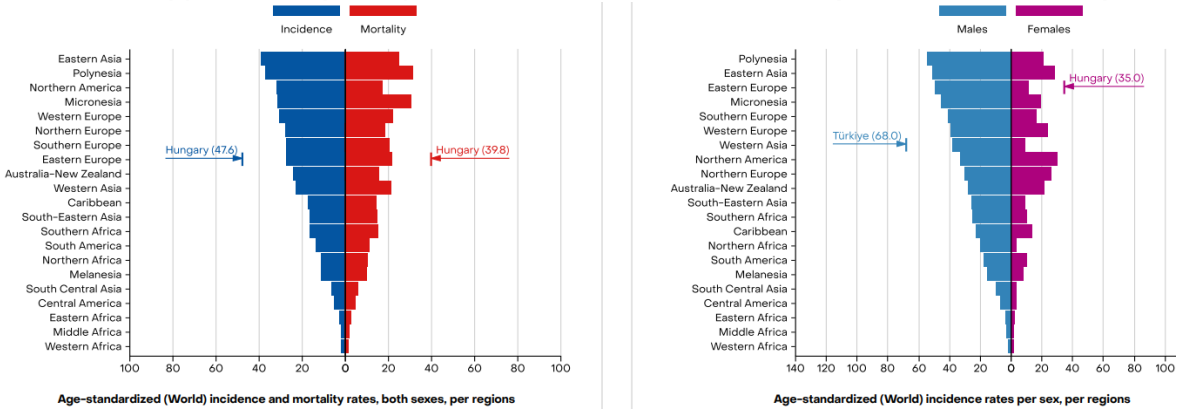


Figure 8: **(L) Age-standardized (World) Incidence and mortality rates**, both sexes, per regions; **(R)** Age-standardized (World) incidence rates, per sex, per regions (International Agency for Research on Cancer- World Health Organization; Global Cancer Observatory, GCO- https://gco.iarc.who.int).

Higher incidence rates are observed in highly developed regions, such as North America and Europe, whereas lower rates are reported in low-income regions. Yet, many countries still have a high mortality due to cancer because of delays in diagnosis and lack of access to treatments (Global Cancer Observatory, 2022).

Risk Factors

Tobacco exposure is the leading cause of lung cancer and responsible for most cases. Nevertheless, as many as 30% of lung cancers are diagnosed in never smokers (NSCLC), particularly in Asia, where EGFR mutations occur more frequently than in other populations. These results indicate alternative carcinogenic pathways, with air pollution being one of the greater contributors.

Occupational and environmental cancerous (carcinogens such as asbestos, radon and heavy metals) also contribute substantially especially in those who were exposed through their occupation. Other risk factors are positive family history, previous pulmonary diseases and thoracic radiotherapy that predispose to lung cancer (Shankar et al., 2019).

Diagnostic Approaches

Many efforts have led to large-scale screening programs just to improve early detection, such as the National Lung Screening Trial (NLST) in the United States during 2002–2004. Using low-dose computed tomography (LDCT) in high-risk populations (long-term heavy smokers), the trial demonstrated a ~20% reduction in lung cancer mortality compared to the conventional one. Beyond LDCT, other imaging techniques include spiral CT, MRI, and ultrasound-guided fine-needle aspiration (EUS-FNA), while interventional pulmonology has advanced tissue acquisition methods, making diagnosis safer and more precise (Chudgar et al., 2015).

Histopathological examination remains the gold standard, but all current diagnostic tools face limitations—ranging from invasiveness and cost to radiation exposure and false positives. Biomarkers, such as CEA, have been extensively investigated but their lack of specificity precludes them from being a standard test. Therefore, it is recommended to combine all the analysis of biomarkers with imaging and histopathology as well for clinical implications (Ren et al., 2024).

Treatment Modalities

Surgery, chemotherapy and radiotherapy (the standard treatments) remain the base of therapy, but all too frequently prove to be inadequate for cancers at late stages. Targeted therapies and immunotherapy have been achieved relevant gains in this field, notably in non-small cell lung cancer (NSCLC). Targeted therapies are meant to directly target molecular changes while immunotherapies, especially immune checkpoint inhibitors (ICIs), have revolutionized the treatment regimens of both NSCLC patients and those with small cell lung cancer (SCLC) (Deb et al., 2022).

Emerging Therapies and Research Directions

Antibody-drug conjugates (ADCs) provide a novel option to targeted drug delivery and lower systemic toxicity. Meanwhile, the knowledge of the cancer microenvironment is a rapidly growing areas in biomedicine because immune cells and stromal elements have a profound impact on treatment response. Thus, the assessment of T-lymphocyte subsets (CD3+ CD4+, CD3+, CD8+) is now being considered as a predictor factor for the clinical effect of checkpoint inhibitors and contributes to patient staging information (He et al., 2025).

Taking together these innovations in diagnostics, immunotherapy, targeted therapy and ADCs, they are changing the face of lung cancer but unfortunately, success relies also on advancing early diagnosis and drugs that match the molecular and immunological profile of each patient (He et al., 2025).

I-6.2 Key Molecular Drivers and Therapeutic Targets in NSCLC and SCLC

Lung cancer is a result of a complex interaction between genetic and epigenetic alterations such as point mutations, amplifications, insertions, deletions, and chromosomal translocations. These alterations trigger molecular cascades that stimulates abnormal growth and silence tumour suppressor genes. Over time, the progressive accumulation of genetic abnormalities disrupts many biological processes like cell cycle regulation, DNA repair, apoptosis, and signalling pathways, leading in conclusion at tumour initiation and progression (Smolarz et al., 2025).

Key Alterations in NSCLC

Non-Small Cell Lung Cancer (NSCLC) is characterised by frequent mutations in oncogenes and less to tumour suppressor genes. *ALK*, *MET*, *BRAF*, *EGFR*, *KRAS*, *RET*, *ROS1*, *HER2*, and *NTRK* are frequently altered in NSCLC, particularly adenocarcinomas and these changes stimulate proliferative signalling, angiogenesis, and evasion of apoptosis. While mutations or deletions in *TP53*, *RB1* and *CDKN2A* can contribute to genomic instability and may disable cell cycle checkpoints.

The *EGFR* pathway is particularly important in NSCLC pathogenesis because dysregulation of *EGFR* promotes cell proliferation and survival but fortunately *EGFR* targeted therapies

have transformed treatment approaches. Similarly, abnormal angiogenesis driven by VEGF overexpression supports tumour vascularization (Grodzka et al., 2023).

Key Alterations in SCLC

Small Cell Lung Cancer (SCLC) displays a different molecular profile due mainly by loss-of-function mutations in tumour suppressor genes other than dysregulated signalling pathways. Tumour suppressors like *TP53* and *RB1* deletions are considered fundamental events in SCLC development; *MYC* family amplification and aberrant Notch signalling also contribute to tumour growth and neuroendocrine differentiation while epigenetic regulators act on dysregulation of *SOX2*, *EZH2*, *KDM5A* and *HDAC* further drives tumour progression and plasticity.

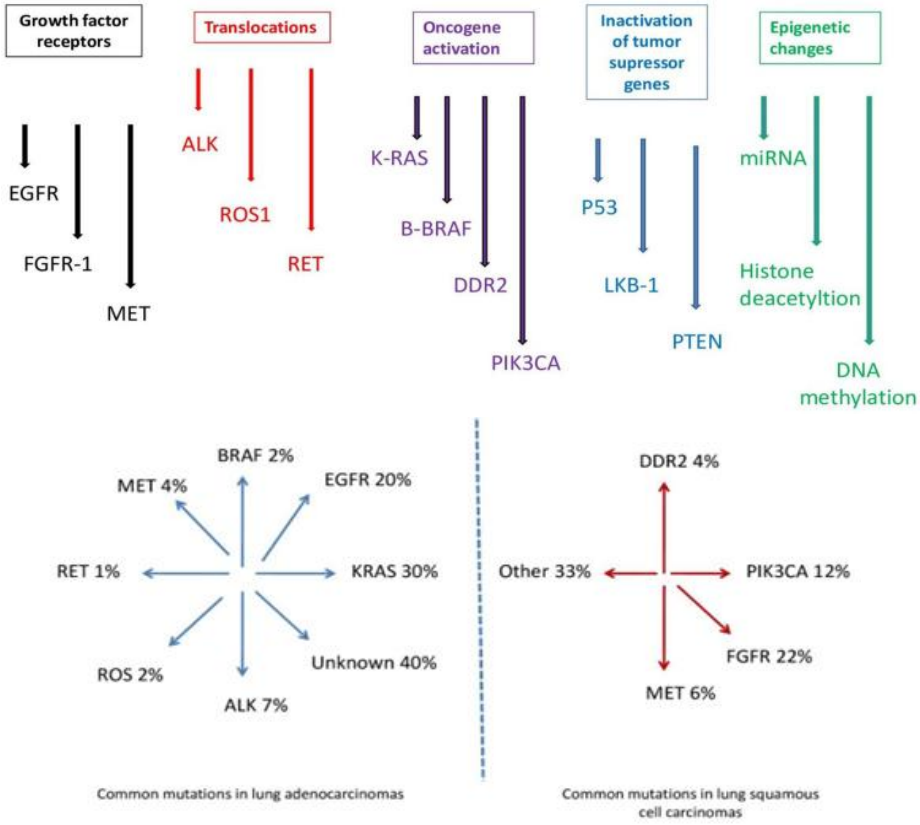


Figure 9: Representation of the main molecular alterations in lung cancer in general) and respectively common mutations in Lung Adenocarcinomas (LUAD) or common mutations in Lung Squamous Cell Carcinoma (SCC) (Smolarz et al., 2025).

Across both NSCLC and SCLC (**Figure 9**), other several processes are central to malignant transformation and can represent the hallmarks of molecular pathogenesis like impaired DNA repair and genomic instability; enhanced angiogenesis through VEGF signalling; evasion of apoptosis and increased telomerase activity and tissue invasion and metastasis (Smolarz et al., 2025).

These findings lead to therapeutic implicationsfor both of LC subtypes. In NSCLC, precision medicine strategies have led to the development of tyrosine kinase inhibitors (TKIs) targeting *EGFR*, ALK, *ROS1* and other oncogenic drivers. ADCs that deliver cytotoxic agents specifically to tumour cells and ICIs were also integrated to treatment. While in SCLC, historically related on chemotherapy, is now being targeted with ICIs, ADCs directed against novel surface markers (e.g., B7H3, Trop-2), or PARP inhibitors that exploit defective DNA repair and kinase inhibitors (Patel & Das, 2023).

Altogether, the molecular landscape of lung cancer highlight the heterogeneity between NSCLC and SCLC, while underlining actionable vulnerabilities that are reshaping in a continuous way the therapeutic strategies.

I-6.2.1 Monitoring EGFR Mutations, Detection of ALK/ROS1 Rearrangements and Other Fusions for Personalized Medicine

Some of the most clinical important biomarkers are EGFR mutational status and ALK/ROS1 rearrangements, not only because they influence the therapy choices but also for the requirement of repeated treatment assessment using more advanced diagnostic methods.

Mutations in EGFR (e.g., exon 19 deletion and L858R substitution) are correlated with sensitivity to EGFR tyrosine kinase inhibitors (TKIs) thus real-time surveillance of these changes is important to take treatment decisions and to determine mechanisms of acquired resistance. Patients presenting sensible EGFR mutations respond favourably to TKIs, making molecular profiling critical at the diagnosis and guiding therapy while Secondary mutations, as EGFR T790M, frequently drive resistance to first- or second-generation TKIs. Detection of T790M allows then for in time initiation of third-generation inhibitors such as Osimertinib. This allow to track alternative pathways with other resistance mechanisms, such as MET amplification, which can be detected by analysis of circulating tumour DNA (ctDNA). This may facilitate combination strategies (e.g., adding MET inhibitors) to overcome resistance mechanisms (Deng et al., 2019).

Detecting ALK and ROS1 Rearrangements and Other Fusions
Rearrangements involving ALK and ROS1 genes are key oncogenic drivers in NSCLC, particularly in younger and non-smoking patients.Their recognition is therefore essential to choose therapy which may include ALK- and ROS1-inhibitors with significant clinical efficacy.

I-4 Molecular Pathology

The molecular pathology emerges as a product of remarkable developments in molecular biology and biotechnology, revolutionising disease diagnosis from simple examination into tissue-based approaches to detection of specific molecular or genetic material. Central drivers are the development of technologies such as PCR and, more recently, NGS that allow for accurate gene-based analysis. In addition, beyond diagnosis, molecular pathology allows personalized therapy based on recognition of specific genetic mutations and genes to be targeted. It has a role in risk estimation and prevention by detecting predisposed individuals to inherited diseases, providing time for preventive measures. Lastly, molecular biomarkers for therapeutic monitoring can contribute to assessing the treatment response and detecting the disease at an early stage.

I-4.1 Oncogenes and Tumour Suppressors

An Oncogene is the mutated form of a 'normal' gene, (proto-oncogene) which is inappropriately activated with stimuli causing unchecked cellular proliferation and neoplastic transformation with resultant cancer establishment. This activation may arise from genetic mutations, environmental exposures, or viral infections.

Oncogenes, together with tumour suppressor genes, constitute the two major classes of cancer-related genes. Not only has their discovery advanced the knowledge of cancer causation, but it has also set the stage for modern precision medicine where targeted therapies are developed to block the function of individual oncogenes as a way to treat cancer (National Human Genome Research Institute, NIH. gov).

Tumour suppressor genes are critical drivers of cellular homeostasis that serve to regulate functions such as cell proliferation, division and programmed cell death (apoptosis). They function as a barrier to halt the uncontrolled proliferation and malignancy of the cell (**Figure 10**). However, if the activity of these genes is inactivated or disrupted through mutation, loss of this control can lead to out-of-control cell division and tumour formation.

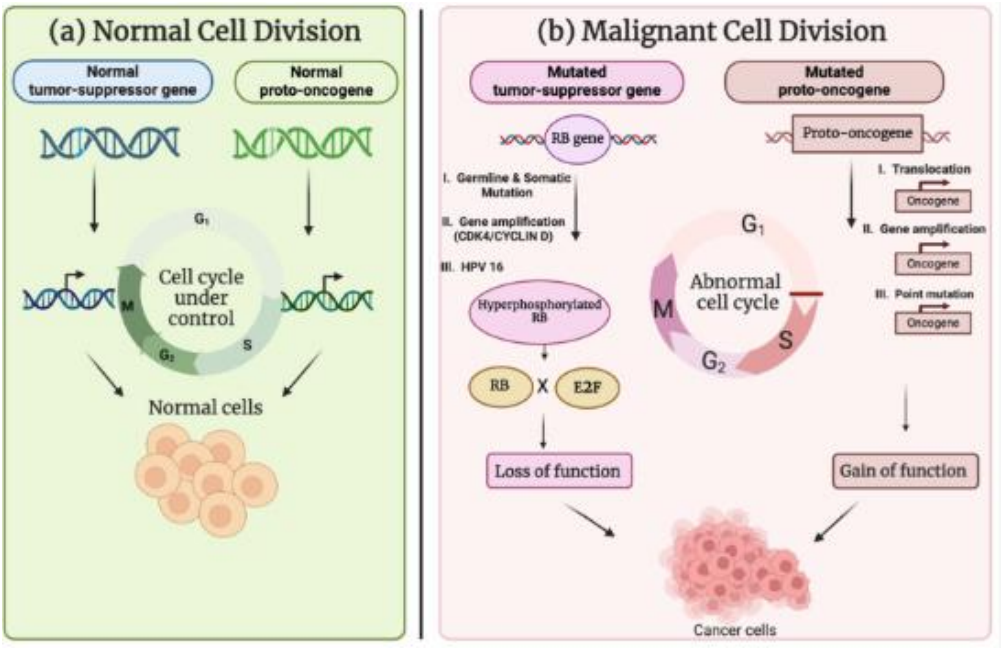


Figure 10: Comparison between normal cell division and malignant cell division with the support of typical or dysregulated cell divisions due to the function of proto-oncogenes and tumour suppressor genes, (Singh, S.R. et al.; Cancers 2025).

Taken together, tumour suppressors along with oncogenes compose the core molecular blueprint moulding cancer biology. Their disruption not only explains much of tumorigenesis but also provides critical targets for cancer risk assessment, prevention, and therapy (American Cancer Society, ACS; https://www.ncbi.nlm.nih.gov/sites/books/NBK532243/).

In Precision Oncology, "actionable alterations" refer to specific genomic changes within a tumour that can be directly targeted by tailored therapies. These alterations hold the potential to improve patient outcomes with a workflow that starts with genetic profiling (most commonly NGS), to analyse tumour DNA and identify driver mutations. The next step is to identify actionable targets since that not every alteration detected can be addressed for therapy. Once the alteration that contributes to tumour growth is confirmed, clinicians can proceed to match the right therapies, selecting from existing targeted drugs or experimental agents available through clinical trials (Chin-Yee & Plutynski, 2024).

Several mutations today can illustrate how genomic findings translate into therapeutic opportunities. For instance, BRAF V600E and V600K mutations in melanoma can be effectively treated with BRAF/MEK inhibitors. NTRK gene fusions represent another actionable category, where multiple tumour types are targeted by NTRK inhibitors. While historically considered undruggable, KRAS mutations are now partially actionable, with specific inhibitors emerging for defined variants (Clinical Use of Precision Oncology Decision Support, 2017).

I-5 Liquid Biopsy

A non-invasive revolution of cancer care is focused on technologies such as liquid biopsy, next-generation imaging and AI to detect disease earlier, more accurate diagnose and monitor treatment response in real time. These techniques are less invasive and a more comfortable approach than conventional tissue biopsies, leading to better patient outcomes and personalized treatment strategies with the detection of biomarkers in body fluids.

I-5.2 Liquid Biopsy Components and Limitations

In the past few years liquid biopsy has become a minimally invasive tool with tremendous clinical potential and is of considerable interest because of its capacity to offer real-time molecular information on tumours. Unlike tissue biopsies, which obtain invasive samples of tumour tissue, liquid biopsies examine components that come from the tumour and are in body fluids such as blood (or other specimen types, including urine, saliva, ascitic fluid, cerebrospinal fluid or pleural effusions). The approach has the added advantage of sampling at multiple time points, which permits for dynamic tracking of disease progression and treatment response (Ren et al., 2024c).

Many types of biomarkers can be used trough liquid biopsy, to assay which collectively contribute to our knowledge of the genomic, transcriptomic and proteomic features of tumours (**Figure 11**). These include: -Circulating Tumour DNA (ctDNA), that are small DNA fragments of tumour cells released into circulation, providing direct genetic information on driver mutations and tumour burden; -Circulating Tumour Cells (CTCs), that are intact malignant cells shed from the primary tumour, which may reflect metastatic potential; -Extracellular Vesicles (EVs), which are nanosized particles containing DNA, RNA, and proteins that offer hints about tumour aggressiveness; -Other Biomarkers, like circulating RNA species (ctRNA, microRNAs, and non-coding RNAs), tumour-associated antigens (TAA), and even tumour-educated platelets (TEPs) (Pandey & Yadav, 2025).

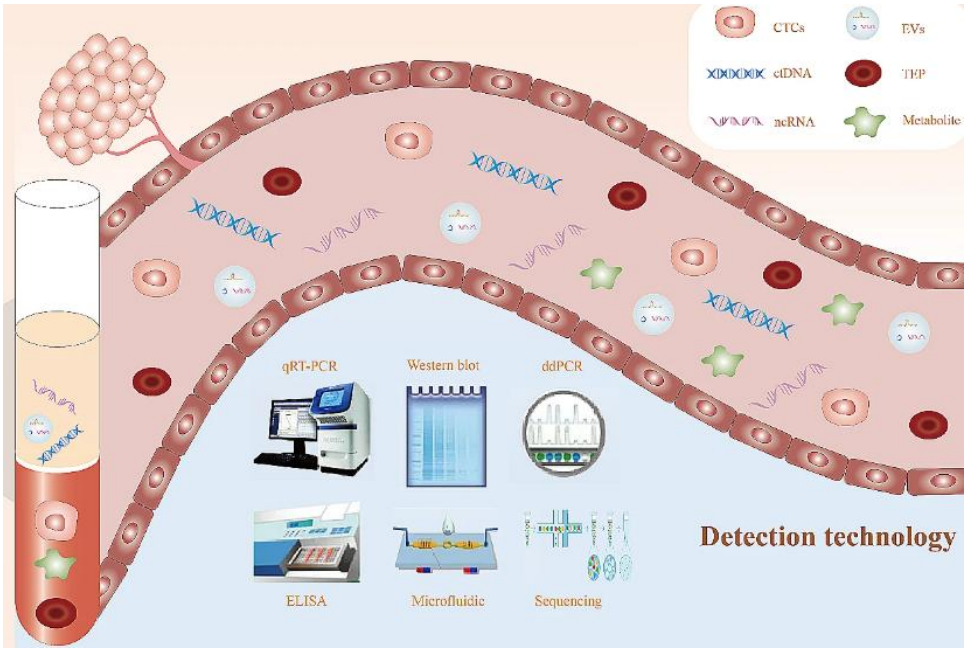


Figure 11: Liquid biopsy Components and Techniques for diagnosis, evaluations and monitoring (Ren et al., 2024b).

To detect and analyse these constituents, numerous and very sensitive molecular techniques can be used, such as real-time PCR (RT-PCR), droplet digital PCR (ddPCR), BEAMing (Beads, Emulsion Amplification and Magnetics), ELISA (Enzyme-linked immunosorbent assay), microfluidics and next generation sequencing (NGS). These methods not only allow identification of tumour-specific mutations, but also to evaluate in quantitative manner levels of ctDNA; particularly the latter have been found to strongly correlate with tumour stage and burden.

The versatility of liquid biopsy enables multiple clinical applications like early diagnosis and screening for detecting cancers at earlier stages. At the same time enables treatment monitoring, evaluating therapy response often identifying changes before radiological evidence is available and Minimal Residual Disease (MRD), detecting small amounts of residual tumour material after surgery or chemotherapy. Additionally allows tumour profiling, providing a more comprehensive overview of tumour evolution (Pandey & Yadav, 2025).

The clinical data support the robustness of liquid biopsy. For example, one research found that 100% stage II–IV patients of non-small cell lung cancer (NSCLC) and 50% of stage I NSCLC patients were positive for ctDNA in peripheral plasma, with a substantial relationship between ctDNA quantification and tumour burden by volume (Newman et al., 2014).These results provide its diagnostic utility as well as monitoring the therapeutic response and predicting relapse.

Limits and Considerations

Liquid biopsy also presents challenges, for example the levels of ctDNA in early-stage cancers are very low and it is difficult to detect them, making more frequent false negatives. In many cases indicative findings have as to be confirmed by imaging or histological testing for a reliable diagnostic characterization. There are further complications with clonal haematopoiesis, where mutations occur in the blood cells instead of the tumours and could

affect findings. Moreover, technology and interpretative skill delivery imbalances make routine use in clinical practice further challenging (Pandey & Yadav, 2025).

Overall, liquid biopsy is probably better thought of as an adjunct rather than a substitute for tissue biopsy. Its increasing utility in all aspects of the cancer continuum, including screening, diagnosis, therapeutic decision and longitudinal monitoring make it a key component for precision oncology in the near future.

I-5.3 cfDNA, cfRNA, EVs and Metabolites

I-5.3.1 Cell-Free DNA (cfDNA)

Cell-free DNA (cfDNA) consists in short DNA fragments released into body fluids (most commonly blood, but also urine, saliva, cerebrospinal fluid, and pleural effusion) through processes such as apoptosis, necrosis, NETosis, or active secretion via EVs. But cfDNA originates from both normal and diseased tissues, making it a heterogeneous pool of genetic material. Its fate is influenced by fragment size and molecular associations, with clearance mechanisms involving enzymatic degradation by DNases and removal via the reticuloendothelial system, though the exact pathways remain still not completely understood (Ren et al., 2024c) (Stejskal et al., 2023).

Origins of cfDNA
-Apoptosis, the major source in healthy individuals, producing fragments of ~167 base pairs with characteristic nucleosomal patterns;
-Necrosis, where releases larger and less uniform fragments due to uncontrolled cell death, common in damaged or hypoxic tissues;
-NETosis, where Neutrophils expel DNA in the form of neutrophil extracellular traps (NETs) and contribute to circulating cfDNA in inflammatory settings (Grabuschnig et al., 2020);
-Active Secretion of living cells can release DNA actively through EVs (e.g., exosomes, microvesicles);
-Tissue Turnover, by which rapidly renewing tissues such as the gastrointestinal tract, lung, and kidney contribute significantly to the cfDNA pool.

Within this pool, a tumour-specific subset known as circulating tumour DNA (ctDNA) can be identified because ctDNA is shed by malignant cells during some same processes as apoptosis, necrosis, or secretion, but carries the genetic and epigenetic alterations characteristic of the tumour.

I-5.3.1.1 Tumour-Derived cfDNA (ctDNA)

The size of ctDNA typically falls into 160–200 base pairs, though variability exists depending on tumour biology. ctDNA can represent as little as 0.01% to as much as 90% of total cfDNA in peripheral blood, depending on tumour stage (**Figure 12**) and burden (Ren et al., 2024) but in particular ctDNA contains tumour-derived genetic material, including somatic mutations, copy number variations, DNA methylation patterns, and microsatellite instability (Stejskal et al., 2023).

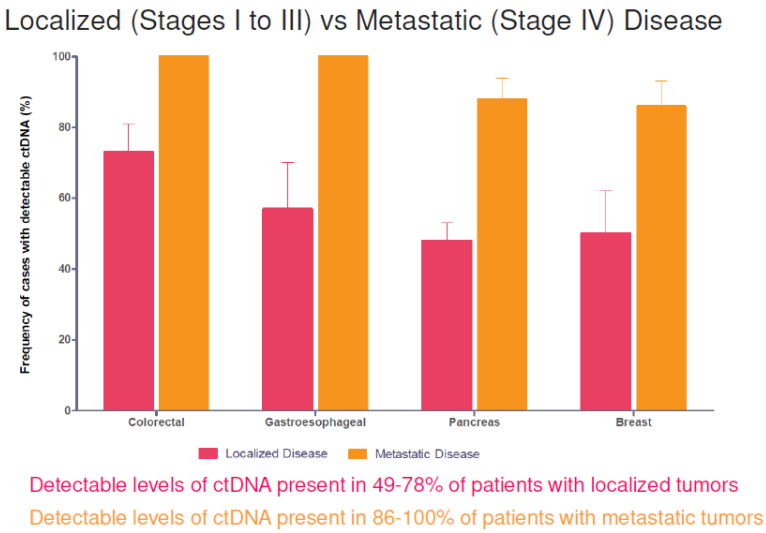


Figure 12: ctDNA frequency in different cancer types (Bettegowda et al, 2014).

Often ctDNA has low abundance in samples, given that highly sensitive technologies and analytical techniques are required. These may include Allele-specific PCR methods such as ARMS (Amplification Refractory Mutation System), qPCR for mutation detection and quantification, dPCR/ddPCR for precise quantification of low-frequency variants and NGS for a broader genomic profiling and discovery of rare variants.

Therefore cfDNA and ctDNA have a dynamic nature, they are part of a continuous cycle of release, degradation, and clearance (mostly by kidneys). Their short half-life ranges from minutes to hours and makes them excellent indicators biological and pathological processes in real-time, including tumour evolution under therapeutic pressure.

I-5.3.2 Cell-Free RNA (cfRNA)

cfRNA refers to fragmented RNA molecules that are released from various cell types including tumour cells, into body fluids as mentioned before. Despite their instability and susceptibility to degradation, cfRNAs carry many biological information, making them a promising non-invasive biomarker for serious processes in cancer management as detection, prognosis, and treatment monitoring (Larson et al., 2021).

cfRNAs are RNA fragments that freely circulates in biofluids and they can arise as other cell-free nucleic acids from cell death processes (apoptosis, necrosis), active secretion through EVs, or other cellular mechanisms (Larson et al., 2021) as mentioned above. Most cfRNAs are fragmented by RNase activity, though some remain protected within EVs or bound to proteins, which increases their stability.

cfRNA has its own unique role in cancer, it contains tumour-specific information, offering other molecular signatures to tumour type and state other than ctDNA. cfRNA profiling enables then another level of non-invasive detection of cancers and helps distinguish malignant from premalignant or healthy conditions. As well for prognostic and monitoring applications**,** cfRNAs may predict patient outcomes and are useful for monitoring treatment response, detecting relapse, or tracking disease progression (Jin et al., 2023).

The research and analytical approaches are similar to the previous one for cfDNA and for other circulating nucleic acids starting by the sample collection**,** cfRNAs are typically extracted from plasma or other biofluids obtained from patients then Next-generation

sequencing (NGS), particularly RNA sequencing, is applied to characterise the full repertoire of cfRNA molecules and quantify transcript abundance. Finally, the data analysis with computational pipelines, often incorporating machine learning, are used to identify cancer-specific cfRNA signatures and clinically relevant biomarkers.

On the other hand, the are some limitations for these biomarkers like degradation**,** the short half-life of RNA and the activity of RNases in biofluids complicate cfRNA detection and analysis but more challenging is the standardization with reliable use of cfRNA in clinical practice, that requires standardized protocols for sample handling, processing, and data interpretation to ensure reproducibility across laboratories (Bao et al., 2025).

I-5.3.2.1 Circulating microRNAs

Biogenesis and Release

MicroRNAs (miRNAs) are small non-coding RNAs generated in a canonical pathway that begins with transcription of primary miRNAs (pri-miRNAs) by RNA polymerase II. In the nucleus, the processor complex that comprehend Drosha and DGCR8, processes pri-miRNAs into precursor miRNAs (pre-miRNAs). These are exported in the cytoplasm by Exportin-5, where the enzyme Dicer cleaves them into mature miRNA duplexes. One strand of the duplex is then loaded onto Argonaute (Ago) proteins to form the RNA-induced silencing complex (RISC), which mediates post-transcriptional gene silencing (O'Brien et al., 2018) (**Figure 13**). Once processed, miRNAs can then be released in the extracellular space in several forms e.g. enclosed in extracellular vesicles such as exosomes or microvesicles, linked to proteins or associated with high-density lipoproteins (HDL). In circulation, they can be internalized in recipient cells through endocytosis, micropinocytosis, or direct cell-to-cell interactions, functioning as well as intercellular signalling molecules (O'Brien et al., 2018).

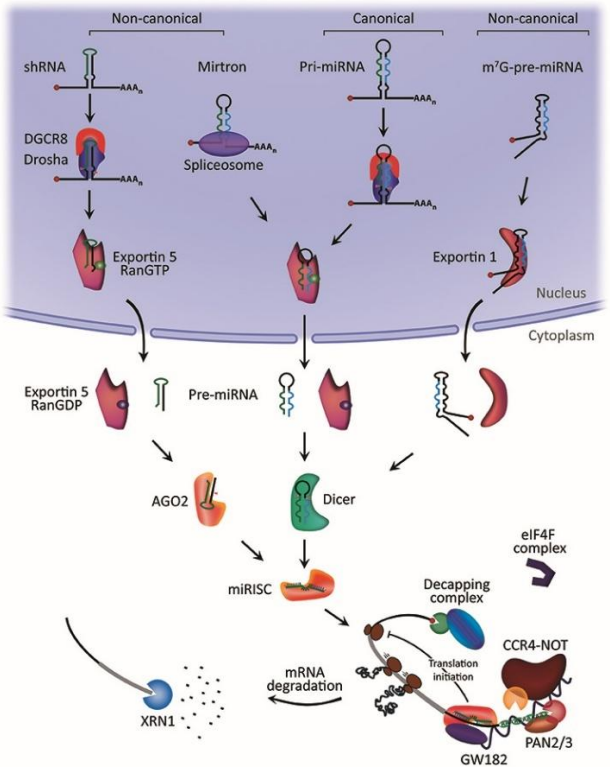


Figure 13: MicroRNA biogenesis and mechanism of action (modified from O'Brien, Hayder H et. al 2018).

Role in Cancer
Circulating miRNAs (ci-miRNAs) have emerged as key players in cancer biology, with a dual role that depends on their target pathways: - OncomiRs, such as miR-21 and miR-155, act by downregulating tumour-suppressive genes, thereby promoting proliferation, angiogenesis, and metastasis; - Tumour-Suppressor miRNAs that instead inhibit oncogenic pathways, are able to restraining malignant transformation and progression; - Tumour Microenvironment, some circulating miRNAs also shape the tumour microenvironment, where their influence may help immune evasion, angiogenesis, and interactions between cancer and stromal cells (R. Ma et al., 2012).

miRNAs are be proposed as diagnostic and prognostic biomarkers because distinct circulating miRNA and their expression patterns have been linked to specific cancer types and stages, offering again another non-invasive tools for early diagnosis and prognosis. In the meantime, they also have a therapeutic potential thanks to strategies under investigation that include delivering tumour-suppressive miRNAs to malignant cells or inhibiting oncomiRs with antagomiRs (Ardila-Molano et al., 2015).

The instability of RNA in biofluids though necessitates stabilization and a reliable detection method. Nowadays there is a lack of standardized protocols for sample handling and profiling, interfering with reproducibility across studies. Moreover, we have to consider the complexity of miRNA networks where individual miRNAs can target multiple genes and pathways, this requires integration with other omics data and advanced computational approaches. Future research combining circulating miRNAs with multi-omics platforms and machine learning may significantly improve their diagnostic, prognostic, and therapeutic utility in precision oncology (O'Brien et al., 2018).

I-5.3.3 Extracellular Vesicles (EVs)

EVs are nanoparticles made up by a lipid bilayer and are secreted by virtually all cell types, playing a central role in intercellular communication and regulation of diverse biological processes. They transport a variety of molecular cargo, including proteins, lipids, and nucleic acids such as mRNA, miRNA, lncRNA, and are actively involved in tumorigenesis, cancer progression or modulation of the immune microenvironment. Importantly, EVs can provide real-time information about tumour activity that can be more accurate than ctDNA alone, making their genetic content, in particular miRNAs and lncRNAs, highly promising biomarkers for detection (De Sousa et al., 2023).

EVs are highly heterogeneous particles different ways, in origin, size, and composition, and are generally classified into three main subtypes (**Figure 14**):
-Exosomes (30–150 nm), formed by the maturation of late endosomes into multivesicular bodies, which fuse together with plasma membrane to release intraluminal vesicles. Due to their stability, exosomes are the most widely studied EVs, even for their small size, and ability to penetrate tissues, as well as their enrichment in specific proteins and nucleic acids;
-Microvesicles (100–1000 nm) instead are generated by the direct budding of the outer plasma membrane, carrying proteins, lipids, and RNA cargo reflective of their parent cell;
-Apoptotic bodies (50–5000 nm), at last are larger vesicles formed during programmed cell death, containing for this reason some fragments of organelles and cellular debris (De Sousa et al., 2023).

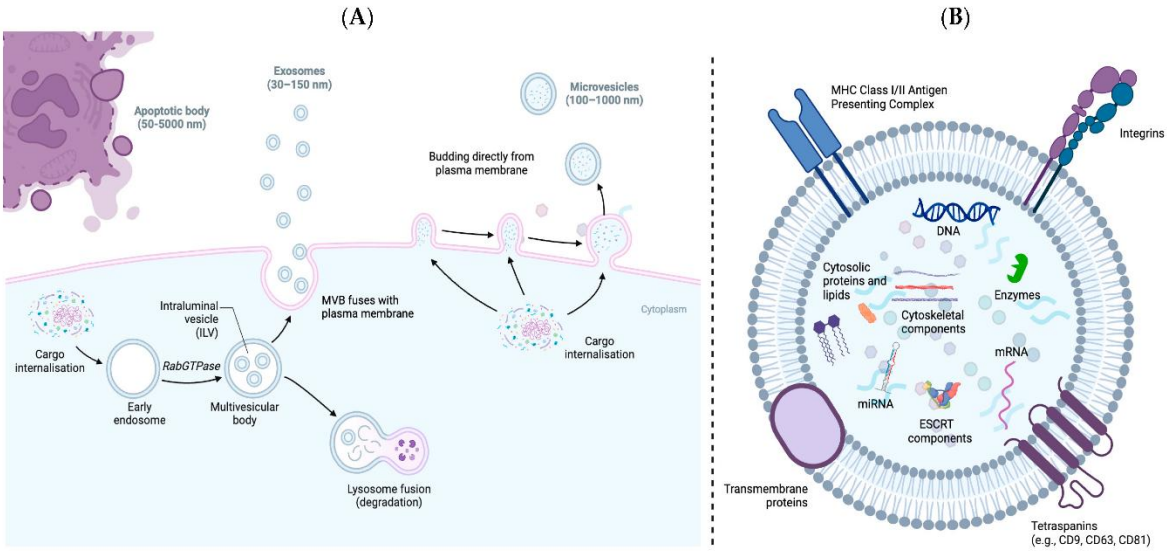


Figure 14: (A) Illustration of the biogenesis pathway for EV and the secretion of exosomes, through multivesicular bodies within endosomes, budding of the plasma membrane that generate microvesicles and apoptotic bodies. **(B)** Representation of a typical EV that contain various macromolecules including coding and non-coding RNAs, DNA, signalling proteins, lipids, and transcriptional regulators (Lawrence & Shah, 2024).

Implications in Cancer

The heterogeneity EVs has profound implications in cancer biology and clinical applications. They can have a tumour promotion effect**,** carrying pro-tumorigenic signals that enhance cancer cell proliferation, survival, and invasive potential. They can also hold metastatic potential by transferring oncogenic molecules to recipient cells and reshaping the tumour microenvironment. EVs modulate immune activity helping immune evasion thus tumour cells can escape immune surveillance and establish a more permissive microenvironment (Jabalee et al., 2018). On the other hand, they are therapeutic targets, since the specific molecular cargo and surface markers of distinct EV subpopulations, even at the single-vesicle level, provide novel opportunities for the development of targeted therapies and cancer vaccines. They may serve the cause as diagnostic tools when the analysis of tumour-derived EVs (tdEVs) circulating in biofluids enables the identification of unique molecular signatures, supporting the development of highly sensitive and non-invasive liquid biopsy tool for early cancer detection or longitudinal monitoring of disease (Lawrence & Shah, 2024).

Isolation and Characterization of EVs: Challenges and Innovations

Isolating and characterizing EVs remains technically a challenge due to their heterogeneity**,** because they present a substantial variability in size, density, and cargo that further complicates the purification of homogeneous populations. In addition, blood contain a variety of lipoproteins, protein aggregates, and platelets that overlap in density and size with various EVs, making it more difficult to achieve specificity with conventional ultracentrifugation.

The lack of standardization reflects the differences in sample collection, processing, and analytical approaches limit reproducibility across studies because of the analytical variation**,** inconsistencies in instrument calibration and workflow introduce further variability (Escudero-Cernuda et al., 2025).

I-5.3.4 Role of Circulating Metabolites in Cancer

Circulating metabolites are small molecules present in blood and other biofluids that reflects cellular and systemic metabolic processes and for that they are recognized as valuable biomarkers for cancer risk, progression, and treatment. We know that cancer cells undergo profound metabolic reprogramming to sustain the rapid growth rate and survival, thus changes in metabolite profiles can provide insights into tumour biology and patient-specific responses (Wang et al., 2023).

Significance

Circulating metabolites can serve as non-invasive biomarkers in liquid biopsies, aiding in cancer detection, disease monitoring, and evaluation of response to treatment as for the previous markers and provide etiological insights**,** characterising metabolite alterations associated with cancer risk. Specific metabolites might affect the tumour microenvironment directly and thus play a role in cancer initiation, progression and immune escape but these individual metabolic profiles could be indicators or A indicators for treatment response so that tailored therapeutic concepts for personalized medicine can be chosen.

Examples of Key Metabolites

-Short-chain fatty acids (SCFAs), like Acetate and other SCFAs can be used by tumor cells as substrate for lipid biosynthesis to support their growth;
-Lipids, as circulating lipids, including HDL cholesterol and triglycerides differ in their associations with risk of cancer. High HDL might be protective while high triglycerides are often associated with higher risk;
-Amino Acids: Aberrant metabolism of amino acids promotes the synthesis of protein and production of energy for fast multiplying tumor cells resulting in progression;.
-Ketone Bodies: Changes in ketone body metabolism reflect the metabolic plasticity of cancer cells and have roles in energy supply and signalling;
-Inflammation: New markers such as GlycA have related the two broad to cancer incidence and death, illustrating a connection between systemic inflammation and tumorigenesis (Jokela et al., 2024).

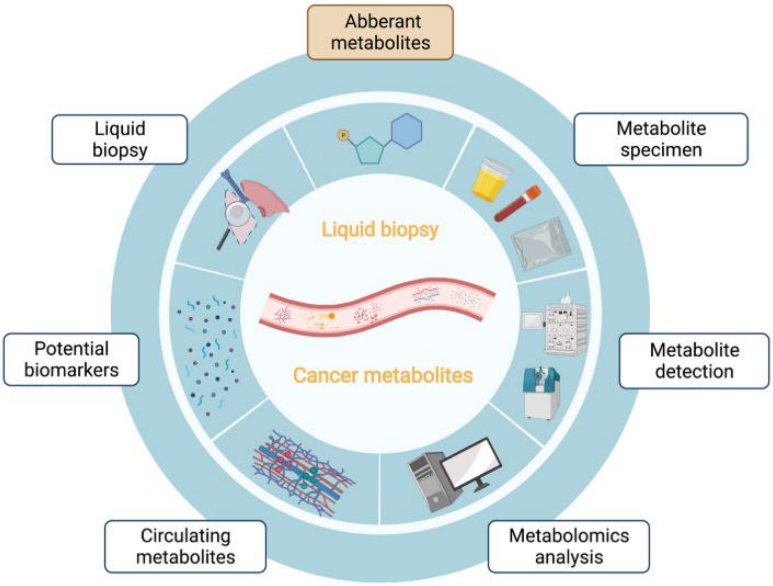


Figure 15: Overview of Metabolite contents, which may help discovery of total abnormal metabolites in cancer and the development of high-throughput metabolomics. Circulating metabolites are then potential biomarkers for early detection and diagnosis of cancer in the context of liquid biopsy (Wang et al., 2023).

The field is gaining technological advancements, especially in Metabolomics (**Figure 15**), using high-throughput techniques to analyse hundreds or thousands of metabolites in biological samples. Combined with advanced computational tools, including Machine Learning, metabolomics enables the identification of cancer-specific metabolic signatures and the development of diagnostic and prognostic tools with high precision (Wang et al., 2023).

I-7.2 Guiding Treatment Decisions and Surveillance in CRC

I-7.2.1 KRAS/NRAS/BRAF Mutations: Predicting Response to Anti-EGFR Therapies

The proximal and distal large intestines have different embryonic origins, as mentioned previously and these differences are associated with distinct gene expression patterns. In particular, proximal colon tumours generally show higher TMN stages and are associated with microsatellite instability and the CIMP phenotype, with hypermethylated promoters and more frequent mutations in *KRAS* and *BRAF*; conversely, distal colon tumours are more frequently associated with chromosomal instability (Li J et al., 2021).

KRAS and BRAF

Approximately 40% of patients with colorectal cancer have a missense mutations in *KRAS*, " Kirsten rat sarcoma", that is also one of the most frequent mutated oncogenes in this type of cancer. The *KRAS* gene belongs to the *RAS* family genes, which play a crucial role in regulating cellular processes like proliferation, differentiation and survival. The encoded protein is a GTPase that consist of two principal domains, the G domain for GTP/GDP exchange, which activates and deactivates it respectively, and the C-terminal domain that is responsible for translocation and anchorage (G. Zhu et al., 2021). *KRAS* is involved in the RAF-MEK-ERK pathway, a pathway implicated in sequential phosphorylation and subsequent activation of a series of proteins, which results at the end in the activation of nuclear substrates that control other proteins involved in cell cycle regulation. Among these are ARAF, BRAF, and RAF1 (all belonging to the Raf family), which are involved in turn in the activation of MEK1 and MEK2, which, in turn again phosphorylate ERK1 and ERK2; the latter are responsible for the activation of nuclear and cytosolic substrates/enzymes such as cyclin D1, that control the cell cycle (Pino & Chung, 2010).

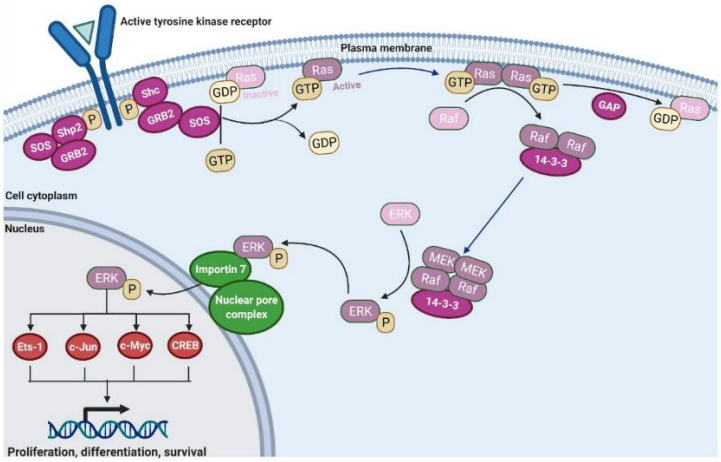


Figure 16: Pathway cascade of Ras-MAPK/ERK (adapted from Dillon, Cancers 2021).

Mutations in *KRAS* are responsible for the earliest stages of carcinogenesis and occur following mutations in *APC*. In most cases, mutations in *KRAS* occur at codons 12 and 13 and less frequent in exons 3 and 4 (Testa U et al., 2018). Generally, these mutations alter the protein's GTPase activity, preventing it from returning to its inactive form; this results in a constitutive activation of the entire pathway and, therefore, of cell proliferation and division signals.

The KRAS protein is involved in several upstream signalling pathways, which involve distinct growth factors. Of particular importance among these is the EGFR (epidermal growth factor receptor), which determines the activation of the RAS signalling pathway and PI3K, phosphatidylinositol-3-kinase. Important kinases are found downstream of this latter pathway. Important targets, such as mTOR and AKT/PKB, implicated in promoting cell proliferation and survival, are found downstream of this pathway. Thus, in CRC mutations are often found in genes involved in these pathways, except for *EGFR*, which is mutated in only 1% of colorectal cancer cases. Another signalling cascade of which KRAS is a part is that of the transcription factor NF-κB, responsible for regulating cell survival, as well as inflammatory and immune responses (Pino MS et al., 2010) (J. Li et al., 2021). Mutations in *KRAS* are highly correlated with the CIN phenotype, and the CMS3 subtype in which deregulations in metabolic pathways are also usually identified, including carbohydrates, lipophospholipids, and fatty acids; these deregulations indeed appear to contribute to tumour growth (Nguyen et al., 2020).

BRAF (B-type RAF), is a serine/threonine kinase belonging to the Raf family, involved in the MAPK pathway (mitogen-activated protein kinase) therefore mutations that constitutively activate this protein are responsible for continuous cell proliferation. The most common mutation affecting this gene is the *BRAF* V600E mutation observed in 20% of patients with cancer (Ros et al., 2021).

Predicting Response

Predicting response to anti-EGFR therapy primarily involves testing for *KRAS* mutations, as their presence is a negative predictor of response to these treatments. Other factors like *NRAS*, *BRAF*, and *PIK3CA* mutations, as well as *PTEN* expression, are also investigated as potential markers to stratify patients, even if their clinical utility is still under development. Novel approaches involve novel imaging probes to assess EGFR internalization and degradation, which show promise in predicting cetuximab response (Shankaran et al., 2010).

It must be mentioned the synonymous polymorphism (rs1050171) in the *EGFR* gene has been identified in some studies as a factor predicting better clinical outcomes for anti-EGFR treatment, independent of *RAS* status. Also, overexpression of the *MET* gene has been identified as a new marker that can predict response to anti-EGFR therapy (Kishiki et al., 2014).

Future directions will include clinical trials, and correlative analyses to confirm the predictive role of these potential markers and to develop better strategies for patients who may not benefit from current anti-EGFR therapies. Elucidating the integrated approaches combining genetic, protein expression, and novel imaging data may lead to more accurate and personalised predictions of treatment response (Shankaran et al., 2010).

I-7.2.2 Microsatellite Instability, Tumour Mutational Burden and Minimal Residual Disease Detection with cfDNA

Microsatellites are short tandem repeats of 1-4 bp, widely distributed across both coding and non-coding regions of the genome. These sequences are particularly prone to replication errors, often caused by DNA polymerase slippage during synthesis, that may physiologically occur but due to their repetitive nature, it occurs way more frequently. In physiological conditions, such errors are corrected by the mismatch repair (MMR) system, a protein complex that includes MLH1, PMS2, MSH2, MSH6, and MSH3. Upon detecting an error, MSH2 forms a heterodimer with MSH3 or MSH6, which then recruits the MLH1/PMS2 heterodimer to excise and repair the erroneous DNA segment (Tieng et al., 2021).

Mutations or epigenetic silencing of these MMR genes result in defective DNA repair mechanism, leading to the accumulation of replication errors and the development of a hypermutated phenotype, known as microsatellite instability (MSI). Germline mutations have a role in MMR genes, accounting for ~20% of MSI cases, while a large ~80% frequency is linked to epigenetic inactivation, particularly hypermethylation of the MLH1 promoter (Testa et al., 2018).

MSI thus may serves as a key molecular hallmark of hereditary non-polyposis colorectal cancer (HNPCC, or Lynch syndrome), seen that is being present in ~95% of such cases, but is also observed in ~15% of sporadic CRCs. MSI tumours can then be stratified into three groups for this feature: MSI-High (MSI-H) >30% of unstable loci; MSI-Low (MSI-L) 10–29% of unstable loci; Microsatellite Stable (MSS) with no instability detected. MSI-H tumours often exhibit accelerated progression (1–3 years vs. 10–15 years for CIN-positive CRC) due to their hypermutable context. They are frequently associated with mutations in oncogenes such as BRAF or KRAS, with KRAS mutations being more common in MSS tumours, while BRAF mutations are enriched in sporadic MSI cancers driven by MLH1 hypermethylation (Vilar & Gruber, 2010).

Alongside MSI, another important genomic characteristic is Tumour Mutational Burden (TMB), which quantifies the total number of somatic mutations within a tumour. Both MSI and TMB influence the behaviour of tumour and responsiveness to immunotherapy, making them clinically relevant as biomarkers (Tieng et al., 2021).

In this perspective, ctDNA provide a non-invasive tool of detecting molecular residual disease (MRD) after treatment. While MSI and TMB are rarely measured in cfDNA for MRD assessment, ctDNA analysis mainly focuses on identifying tumour-specific mutations shed into the bloodstream. This approach though offers multiple clinical benefits with the capacity of detecting ctDNA that indicates the presence of MRD and a higher risk of recurrence. An MRD-positive result can suggest therapeutic adjustments and guide treatment refining because patients with persistently undetectable MRD are more likely to achieve a long-term remission or cure (Gristina et al., 2023).

I-8 Idiopathic Pulmonary Fibrosis (IPF)

Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive lung disease of unknown etiology, characterized by scarring (fibrosis) of lung tissue that makes the lungs stiff and impairs breathing. It affects more often people around 70–75 years old and typically presents with progressive shortness of breath, a persistent dry cough, associated with fatigue and sometimes an unexplained weight loss, presented with joints and muscles that ache, and rigidity of the fingertips.

I-8.1 Clinical Features

The excessive fibrosis leads to loss of function of lung tissue with reduced compliance and progressive respiratory failure. Thus diagnosis requires a multidisciplinary approach, typically including clinical evaluation like age, symptoms, and bibasilar inspiratory crackles; imaging with High-resolution computed tomography (HRCT) and the scans have to show the pattern of usual interstitial pneumonia (UIP); exclusion, ruling out alternative causes of lung scarring (Barratt et al., 2018).
Since there is no proper cure, the actual treatment focuses on slowing progression, relieving symptoms and improving quality of life of patients with antifibrotic drugs as Pirfenidone and Nintedanib that reduce the rate of fibrosis. Therefore, the patient management remain still an unmet need due to the lack of supportive care**,** although smoking cessation or pulmonary rehabilitation with supplemental oxygen may help. As last call, lung transplantation is considered only for eligible patients while palliative care with early integration of drugs may help manage symptoms (Raghu et al., 2011).

In this scenario, understanding the molecular pathways that drive fibrosis become crucial to develop new targeted therapies.

I-8.2 Pathogenesis and Molecular Mechanisms

The aberrant activation and/or injury of alveolar epithelial cells originates the disease, which fail to repair themselves properly and instead trigger persistent fibroblast proliferation and extracellular matrix (ECM) deposition. Several interconnected molecular pathways are implicated (**Figure 17**) (Mustafin, 2022):
-Epithelial dysfunction, caused by chronic or repeated epithelial injury may initiates abnormal repair responses;
-Epithelial–mesenchymal transition (EMT), injured epithelial cells reprogram into mesenchymal-like cells, producing an excess of ECM (Kinoshita & Goto, 2019);
-Pro-fibrotic cytokines, as TGF-β that plays a central role by driving fibroblast differentiation into myofibroblasts and stimulating collagen synthesis while PDGF promotes fibroblast proliferation and contributes to ECM accumulation;
-Signalling pathways, Wnt/β-catenin normally active during development, becomes dysregulated in IPF, promoting EMT and fibrogenesis while AKT/STAT signalling in abnormal way can enhance pro-fibrotic factor activity;
-Extracellular matrix remodelling, made up by fibroblasts and myofibroblasts that produce large amounts of collagen, resulting in dense scar tissue that disrupts lung architecture and gas exchange.

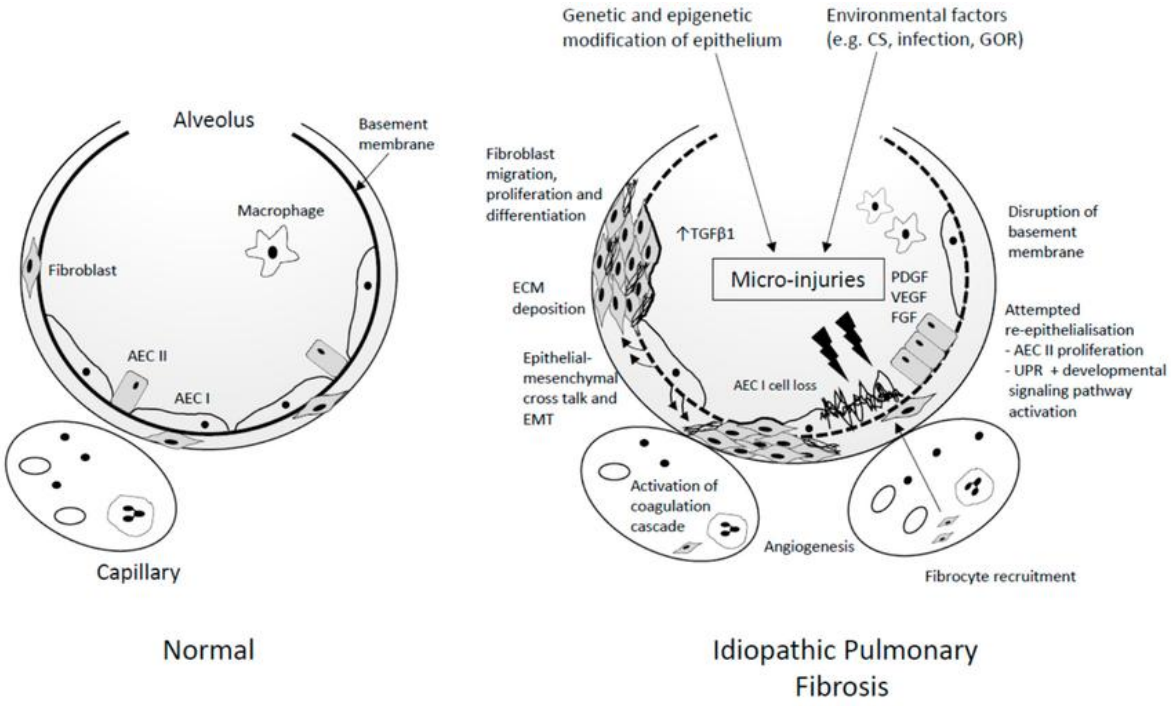


Figure 17: Schematic view of IPF pathogenesis. Genetic and epigenetic phenomena contribute to the development of a dysfunctional epithelium (Barratt et al., 2018).

Genetic and epigenetic factors also contribute, like variants in genes as desmoplakin that weaken epithelial integrity, while epigenetic modifications link genetic predisposition to environmental triggers (Mustafin, 2022).

I-8.3 Early Idiopathic Pulmonary Fibrosis

Early Idiopathic Pulmonary Fibrosis (E-IPF) refers instead to the initial stage of the disease, which may be asymptomatic or present with mild symptoms in patient such as short breath and dry cough. At this stage, lung function tests and high-resolution CT (HRCT) scans may still appear normal or show only minimal abnormalities. Although early diagnosis is really challenging and at the same time it is crucial because treatment at the right timing can slow disease progression and improve quality of life.

Patients in this condition may have mild or absent symptoms or report exertional dyspnoea, dry cough, and fatigue. In early stages, pulmonary function tests often appear normal or show only minor impairment, such as a slightly reduced forced vital capacity (FVC) or diffusing capacity for carbon monoxide (DLCO) resulting as slightly reduced lung function. HRCT may reveal no clear changes or only subtle fibrosis as minimal radiological abnormalities, which does not yet meet the full criteria for nonspecific interstitial pneumonia (NSIP) (Mori & Kondoh, 2021).

In this context the importance of early diagnosis finds its benefits in therapeutics and improved quality of life. Current antifibrotic therapies are more effective in slowing lung damage when started in the early stages of disease and early intervention can help relieve symptoms and preserve the ability to carry out daily activities (Alsomali et al., 2023).

The concept of early-IPF is complex, as its presentation can differ among patients while the subject need for further research. Therefore, we need a great deal more work to find reliable biomarkers which will accurately show at an early phase IPF patient's progress toward symptom-free state (particularly those cases where lung functions have not yet become seriously impaired).

I-8.4 IPF, miRNAs and Lung Cancer

Since microRNAs are involved in both IPF and lung carcinogenesis (e.g. pathways related to cell proliferation, differentiation) they are then implicated in the onset and progression of both diseases. In IPF, various miRNAs are an important cause of fibrosis as they make EMT happen more frequent and more intense as mentioned above; some can make fibroblasts more active than normal while others promote development of abnormal ECM proteins. miRNAs' participation in major biological processes together plus the overlapped risk factors for IPF and lung cancer (like smoking and advancing years) encourages speculation that they could also be responsible for IPF's higher incidence and correlation with lung cancer (Angulo et al., 2012; Ortiz-Quintero et al., 2020).

Role of miRNAs in Idiopathic Pulmonary Fibrosis

miRNAs are involved in regulating crucial cellular processes and key pathways such as inflammation, cellular senescence and proliferation, which are all relevant to IPF pathogenesis. Some miRNAs, like miR-21, have been shown to promote EMT and the deposition of extracellular matrix, contributing to the fibrosis seen in IPF having thus pro-fibrotic effects. Anyway, the specific patterns of dysregulated miRNAs in IPF make them potential targets for novel diagnostic and therapeutic strategies to treat this severe lung disease as promising therapeutics (Cadena-Suárez et al., 2022).

Role of miRNAs in Lung Carcinogenesis

miRNAs are involved in tumorigenesis in general, in a wide range of cellular processes that are vital in the development of lung cancer, including proliferation, differentiation, and apoptosis. Depending on the specific miRNA and its target genes, they can act as either oncogenes or tumour suppressors. Some miRNAs also influence metastasis and angiogenesis (Angulo et al., 2012).

The Connection Between IPF, miRNAs, and Lung Cancer

Patients with IPF have a higher risk of developing lung cancer, which can be attributed to overlapping risk factors such as older age, smoking history and genetic predisposition. This must due also to shared pathways in development of fibrosis and pathogenesis as shown **Figure 18**, including the extracellular matrix organization, negative regulation of cell migration and relevant inflammatory response for fibrosis (Angulo et al., 2012). Elucidating miRNA networks in IPF and lung cancer is crucial and can provide us with new diagnostic tools to detect high-risk cancer patients and novel therapeutic approaches targeting these overlapping molecular pathways.

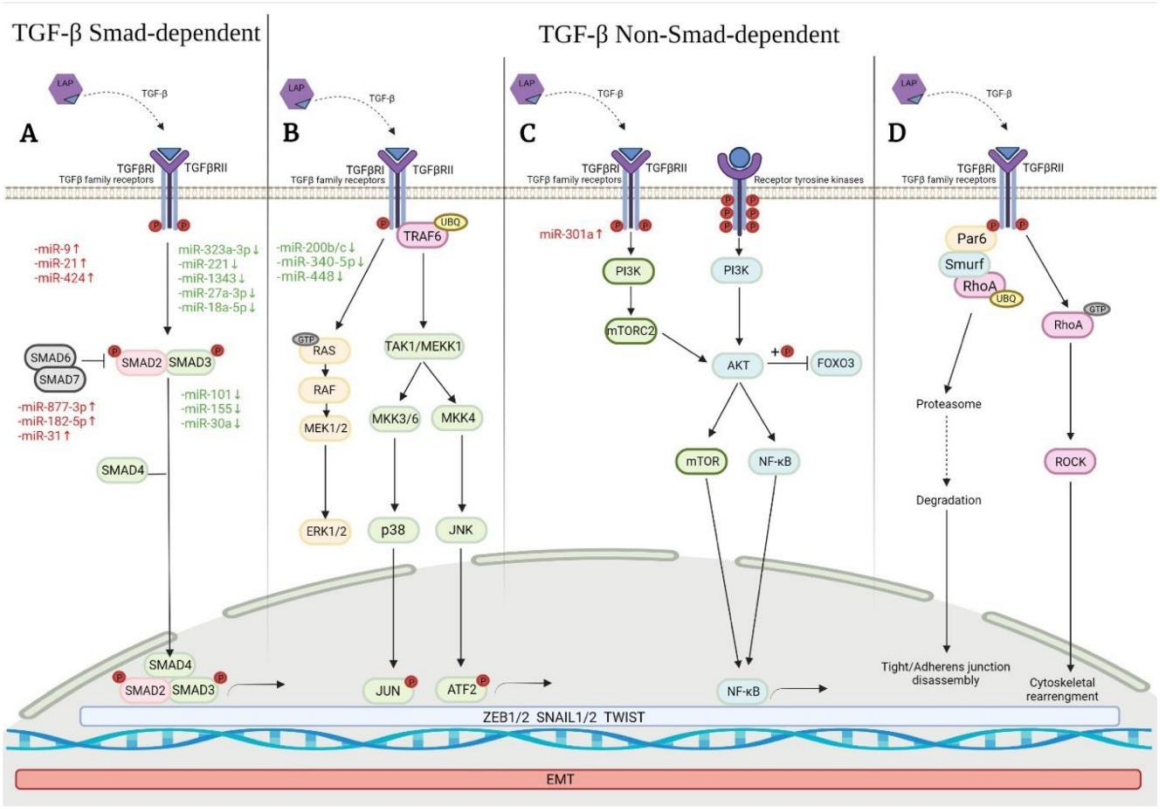


Figure 18: Signalling pathways related to Epithelial Mesenchymal Transition (EMT) in IPF, and examples of some associated miRNAs. Up-regulated miRNAs are indicated in red, while green colour indicates downregulated miRNAs (Cadena-Suárez et al., 2022).

II – AIMS OF THE PROJECT

The aim of this thesis is to investigate liquid biopsy as a minimally invasive tool for the molecular characterisation and monitoring of cancer patients. Specifically, the study focuses on analysing mutations in *KRAS*, *NRAS*, *BRAF* and *EGFR* in ctDNA derived from plasma samples of patients with mCRC and LC. Plasma samples were analysed using BEAMing technology, qPCR and in the case of EVs characterization with Nanoparticle Tracking Analysis.

The mutational status identified in plasma was compared with results from tissue biopsy to evaluate concordance. In mCRC and mLC patients, the mutational profile was monitored longitudinally, at baseline and after 4, 8, 12 and 24 weeks of therapy, in order to check treatment response and detect the emergence of new mutations that may drive resistance.

The study also was meant to consider the mutant allele fraction in plasma and evaluate possible associations or correlations with clinical features.

In LC patients, additional analyses included a gene fusion panel (*ALK*, *ROS1*, *RET*, *MET* and *EGFR* mutations in both plasma and bronchial lavage samples.

Moreover, the project investigated the role of microRNAs in early and established IPF, to compare their patterns with lung cancer.

Finally, metabolites and EVs are evaluated as potential novel biomarkers because taken together, these integrative approaches aim to expand the clinical utility of liquid biopsy for cancer diagnosis, therapy monitoring, and biomarker discovery.

III - MATERIALS AND METHODS

III-1 Studies Design, Population and Settings

The data presented in this thesis originate from two biological, observational, prospective, multi-centre, open-label translational studies that involved collecting blood specimens and clinical information from patients with metastatic colorectal cancer (mCRC), lung cancer (LC) undergoing treatment for metastatic disease, and individuals with idiopathic pulmonary fibrosis (IPF). Recruitment took place at the Medical Oncology Units of Careggi University Hospital (Florence, Italy), S. Jacopo Hospital (Pistoia, Italy), S.S. Cosma e Damiano Hospital (Pescia, Italy), and the Pneumology and Respiratory Intensive Care Unit of Careggi University Hospital (Florence, Italy) between 2017 and 2025. Eligibility required a histological diagnosis of TNM stage IV colorectal adenocarcinoma or lung carcinoma, no prior systemic treatment (treatment-naïve), and measurable disease per Response Evaluation Criteria in Solid Tumours (RECIST v1.1) (Eisenhauer et al., 2009).

For the IPF study, patients were considered eligible if presented with radiological pattern (high-resolution CT) compatible with usual interstitial pneumonia (UIP) defined or probable for IPF.

The studies were approved by the local ethical committee (BIO.16.028 released on October 5th 2016 for Careggi hospital and 15858_bio, released on March 5 th 2020 for Pistoia and Pescia hospitals for mCRC study; 25189_bio released on November 29 th 2023 for Pistoia and Pescia hospitals and released on February 2nd 2025 for Careggi hospital).

Each patient provided written informed consent at the enrolment according to the Helsinki Declaration (World Medical Association, WMA).

III-1.1 Patients' Assessment and Follow-Up

Demographic, clinical, and therapeutic information were obtained from clinical records at enrolment. The cohort included both male and female patients, all older than 40 years. Tumour tissue status for each patient had been previously assessed on FFPE biopsies from either primary lesions or metastases using NGS and Mass Spectrometry, performed by experienced laboratory staff at the participating hospitals as part of routine diagnostics.

In clinical practice, mCRC patients with wild-type (wt) RAS received anti-EGFR therapy with or without chemotherapy at the treating physician's discretion. Patients with RAS-mutant mCRC were treated with anti-VEGF agents, with or without cytotoxic chemotherapy, according to physician choice. LC patients were managed with chemotherapy and immunotherapy as appropriate. First-line therapy continued until radiological progression or unacceptable toxicity. CT scans were performed at baseline (prior to initiating first-line therapy) and then every three months to monitor response per clinical standards. All-cause mortality was recorded prospectively.

III- 2 Sample Collection and Processing

Peripheral Blood and Plasma

8-12 mL blood samples of each patient enrolled in the studies, were collected in K2 EDTA BD Vacutainer® collection tubes (BD, Franklin Lakes, NJ, USA) or Cell Free DNA BCT collection tubes® (Streck, La Vista, NE, USA) (**Figure 19**) by the nurses of Medical Units of the abovementioned hospitals, immediately before starting therapy. For ctDNA analysis were collected prior to starting first-line treatment (T0), 4, 8 weeks after starting treatment, and every 12 weeks thereafter until disease progression. An additional plasma sample was collected at the time of radiological progression according to RECIST v1.1 (Eisenhauer et al., 2009).

Figure 19: K2 EDTA BD Vacutainer® collection tubes (BD, Franklin Lakes, NJ, USA) on the left, with capacity of 4 ml and plasma preparation must begin ≤ 4 hours from time of blood draw. On the right, Cell Free DNA BCT collection tubes® (Streck, La Vista, NE, USA) with capacity of 10 ml; plasma preparation must begin ≤ 3 days from time of blood draw.

Plasma was then prepared within 4 or 72 h, depending on the tube type. Plasma preparation was carried out according to the following protocol through two consecutive centrifugation steps (**Figure 20**):

- Centrifugation (Rotofix 32A, Hettich Zentrifugen) at 2000 rpm, 10 minutes, room temperature to separate plasma from cellular components; carefully collect the supernatant (containing plasma), avoiding the buffy coat and cellular fraction, and transfer it into a clean tube;
- Second Centrifugation at 3000 rpm, 10 minutes, room temperature to remove residual platelets and cellular debris; collect the supernatant avoiding disturbing the pellet at the bottom of the tube;
- Aliquots and Storage : purified plasma was splitted into cryogenic tubes in appropriate aliquots (1-2 mL each) and store samples at –80 °C for long-term preservation.

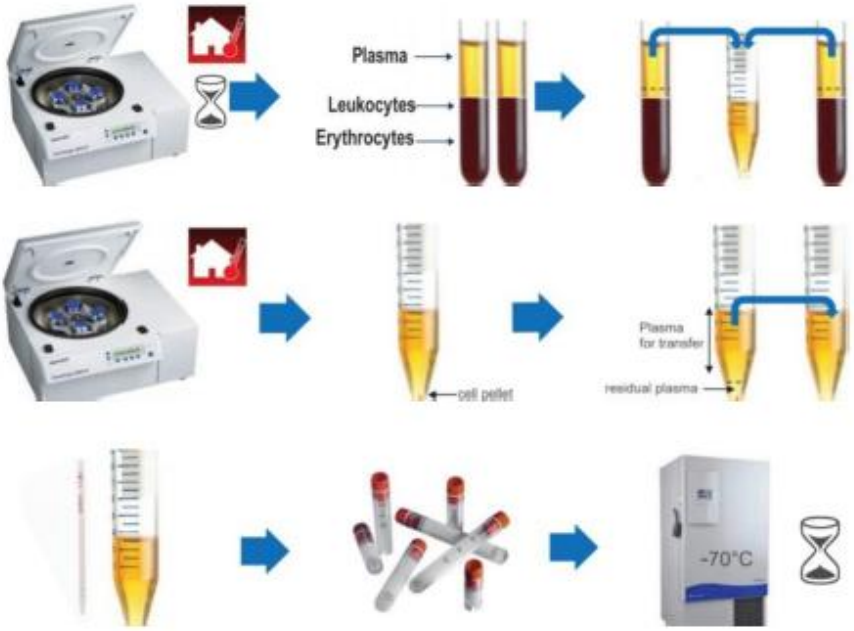


Figure 20: Plasma preparation workflow.

Urine

Samples were collected, in the morning after a minimum of 8 hours fasting. Sample processing was carried out through a centrifugation at 3000 rpm, 5 minutes, followed by filtration (0,20 µm cut-off filter). For long-term storage, the aliquots of processed samples were stored into cryo-vials at -80°C.

Bronco-alveolar Lavage (BAL)

The collection procedure was carried out at the Pneumology and Respiratory Intensive Care Unit of Careggi University Hospital in a sterile tube, with variable volume of BAL and CytoLyt or without fixative and appropriately labelled. Then the sample was transferred immediately to the laboratory where the following protocol was carried out:
-Centrifugation at 2000 rpm, 10 minutes at 4°C to remove mucus and large debris; the supernatant was transferred to another tube and stored at -80°C;
-If the pellet appeared red or dark coloured Red Cell Lysis Buffer treatment was carried out to eliminate RBCs while preserving tumour and epithelial cells:
resuspend the pellet in 1-5 ml of cold RBCLB (Red Blood Cell Lysis Buffer, Roche) and incubate 5-10 minutes on ice;
-Centrifuge at 2000 rpm, 5 minutes at 4°C and discard the supernatant; repeat the previous step if the pellet still appears red;
-Wash the pellet, resuspending in 500µl-1mL of PBS (Dulbecco-PBS®), filter the sample using a 40-70 µm filter;
-Centrifugation at 3000 rpm, 10 minutes at 4°C:discard the supernatant and resuspend the pellet in 20-50 µL of PBS;
-Store at -80°C or proceed with downstream analyses.

III-3 Mutational Tests

III-3.1 BEAMing Digital PCR Technology

The BEAMing method (Beads, Emulsion, Amplification, and Magnetics) is a highly sensitive digital PCR–based technology designed to detect rare mutations in ctDNA. The OncoBEAM™ RAS CRC Assay (Sysmex Inostics, Hamburg, Germany) can identify 34 mutations across the *RAS* genes: 16 in *KRAS* and 18 in *NRAS*, targeting codons 12, 13, and 61 (exons 2, 3, and 4) (**Figure 21**). With a sensitivity as high as 0.01% (detecting one mutant DNA molecule among 10,000 wild-type), this assay provides a powerful tool for non-invasive mutation profiling in plasma samples.

The OncoBEAM™ RAS CRC Assay workflow follows eight steps in according to manual protocol:

1. Blood collection and Plasma preparation
2. ctDNA extraction and purification
3. Multiplex PCR pre-amplification
4. Emulsion PCR (water-in-oil bead-based amplification)
5. Denaturation and allele-specific hybridization
6. Flow cytometric analysis
7. Data interpretation and reporting

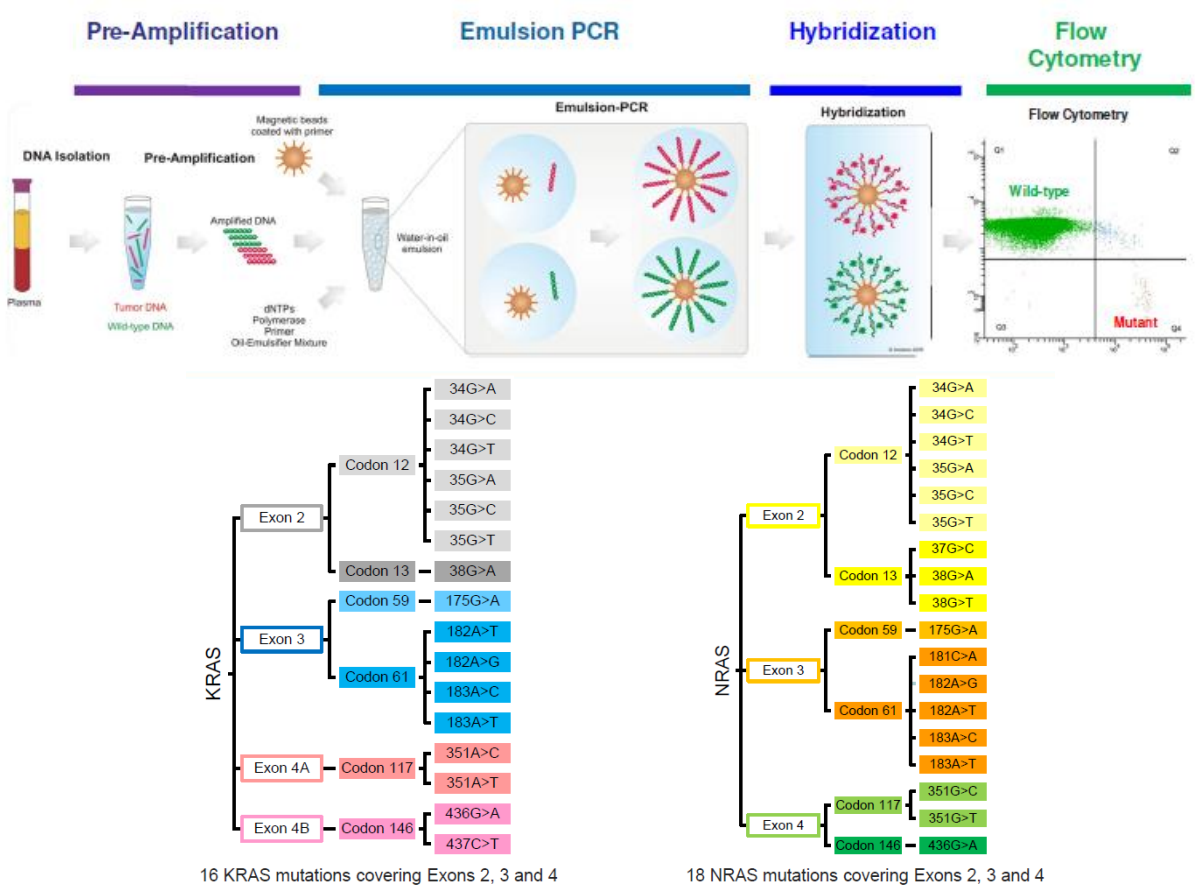


Figure 21: BEAMing workflow overview and OncoBEAM RAS Mutation Panel for CRC (Sysmex Inostics).

ctDNA Extraction and Purification

Plasma samples stored at –80 °C were thawed at room temperature and ctDNA was then isolated using the QIAamp® Circulating Nucleic Acid Kit in combination with the QIAvac 24 Plus system (Qiagen, Hilden, Germany).

Multiplex PCR Amplification

Amplification of *KRAS* (exons 2, 3, 4A, 4B) and *NRAS* (exons 2, 3, 4) was performed using Multiplex PCR (mPCR) in a 96-well plate and run in a thermal cycler (ThermoFisher VeritiDX).

Emulsion PCR (ePCR)

For the emulsion PCR step, an oil–water emulsion was prepared generating millions of microdroplets per reaction. Each droplet contained a single DNA fragment bound to an individual magnetic bead, thereby compartmentalizing millions of parallel PCR reactions within a single tube. Amplification was performed in a thermal cycler (ThermoFisher VeritiDX), producing beads coated with thousands of identical copies of the original DNA template.

Denaturation and Hybridization

Following amplification, the droplets were broken, and the DNA-beads were recovered using a magnetic plate. Denaturation was carried out to remove bound DNA in non-covalent mode, leaving beads coated with single-stranded DNA. The purified beads were then incubated with fluorescent labelled oligonucleotide probes specific to the codons of interest.

Flow Cytometry Analysis

Labelled beads were then analysed using the CyFlow Cube 6i system (Sysmex Inostics).

Data Analysis and Reporting

Fluorescence signals were processed with FCS Express software (version 5.0, De Novo Software).

Results were considered valid only when both positive and negative controls reached the quality thresholds. Mutations were classified if the proportion of mutant DNA–positive beads exceeded the predefined detection threshold.

III-3.2 Idylla™ ctDNA Mutation Tests

The Biocartis Idylla™ system performs automated *in vitro* testing of ctDNA from plasma using a q-PCR. This platform is designed to detect a wide range of clinically relevant mutations in key oncogenes, including *KRAS*, *NRAS*, *BRAF*, and *EGFR*, making it a valuable tool in the management of patients with advanced cancers such mCRC and NSCLC.

Depending on the cartridge used the assay has different targets, the Idylla™ system in our case was used to detect:

- *KRAS*: 21 mutations across codons 12, 13, 59, 61, 117, and 146;
- *NRAS*: 18 mutations in codons 12, 13, 59, 61, 117, and 146;
- *BRAF*: 5 mutations in codon 600
- *EGFR*: mutations in exons 18, 20, and 21, exon 19 deletions, and exon 20 insertions.

The process is semi-automated and performed within a single-use, genes-specific cartridge (**Figure 22**) that contains all required reagents for ctDNA extraction and purification, amplification by allele-specific q-PCR and mutation detection thanks to fluorescent probes.

Figure 22: Idylla ™ System of Biocartis (Belgium). Image shows the platform and a cartridge for plasma with plasma loading (https://www.biocartis.com/).

Data analysis and reporting

To begin the analysis, the operator selects the sample type (plasma or tissue) on the touchscreen interface, scans the barcode on the cartridge (which identifies the gene panel to be analysed), and loads the sample together with Proteinase K into the cartridge. For RAS/BRAF testing, 1 mL of plasma and 30 µL of Proteinase K were required and for EGFR testing, 2 mL of plasma plus 60 µL of Proteinase K were used.
Inside the cartridge, ctDNA is automatically lysed, purified, and bound to a solid support. Amplification is performed in dedicated chambers using allele-specific primers and finally the emitted fluorescence is captured and analysed by the Idylla™ software.

Turnaround Time and Output

The entire analysis takes approximately 110-130 minutes depending on the cartridge. At the end of the run, results are displayed directly on the system screen and can also be exported as a PDF report. The report specifies whether mutations are present or absent, the exact codon and mutation type identified along with test validity.

III-3.2 Idylla™ Gene Fusion Panel

Another type of cartridge was used to analyse cell pellets of broncho-alveolar lavage from NSCLC patients. Stored pellets resuspended in PBS were thawed and added to the cartridge without additives with limited sample input. The Gene Fusion Panel cartridge has different specimen requirements (≥20 mm² tissue area as one FFPE section of 5 µm; or <20 mm² as 3 FFPE sections of 5 µm; ≥10% neoplastic cells) and can detect ALK, ROS1 & RET fusions and MET exon 14 skipping. The assay turnaround time is approximately 180 minutes and combines two detection technologies:
- Highly sensitive detection of the relevant gene fusions directly from RNA transcripts by qPCR;
- Expression imbalance detects gene fusions, respective of the fusion partner, based on the 3' kinase overexpression caused by that partner gene.

III-4 Nuclear Magnetic Resonance-based Metabolomics

Plasma and urine samples of mCRC patients were analysed and studied in collaboration with Giotto Biotech S.r.l., located at the Magnetic Resonance Center (CERM) of the University of Florence. NMR was performed with a Bruker IVDr 600 MHz spectrometer (Bruker BioSpin) operating at 600.13 MHz proton Larmor frequency, equipped with a 5 mm PATXI H/C/N with 2H-decoupling probe including a z axis gradient coil, an automatic tuning-matching (ATM) and an automatic refrigerated sample changer (SampleJet). The workflow combined the nature of biological samples with spectral acquisition, then Data processing and results provided with a table of metabolites (mmol/L, accuracy% and LOD).

III-5 Isolation of Extracellular Vesicles and Nanoparticle Analysis

III-5.1 Purification of Extracellular Vesicles with Ultracentrifuge

This procedure involved several preliminary steps. Plasma samples were thawed and filtered through a 0,22 µm pore membrane (Filtropur S, Sarstedt) using a 3 mL sterile syringe. An initial centrifugation (Centrifuge 5415 R, Eppendorf) was performed at 10'000 rpm for 45 minutes at 4°C. In the meantime, the ultracentrifuge was prepared, inserting the rotor (SW 32 Ti Rotor, Beckman Coulter™) for medium-large volumes and settled the temperature and vacuum for cooling. Post-centrifugation supernatant was recovered and kept on ice, for sample dilution with cold Dulbecco®-PBS to reduce density and viscosity. Then the samples were placed in the appropriate tubes (Ultra-Clear™ Centrifuge Tubes, Beckman Coulter) and accurately weighted. Ultracentrifugation (Optima™ L-90K Ultracentrifuge, Beckman Coulter) was carried out at 30,000 rpm for 1 hour at 4°C and the supernatant was discarded while the pellet was resuspended in DPBS (50-200 µL) by repeated pipetting and scraping. The resuspended samples were stored at -80°C or immediately used for downstream analyses.

III-5.2 Nanoparticle Tracking Analysis (NTA)

Malvern Panalytical's NanoSight Pro® (**Figure 23**), the instrument used for Nanoparticle Tracking Analysis (NTA), provided high-quality and reliable data. The software, after video recordings, perform measurements, eliminates subjectivity, and provides size and concentration data for light scattering. Indeed, the main principle of the technique is the capture of light scattered by particles undergoing Brownian motion; NTA allows individual particle tracking to obtain high-resolution size and concentration data. The NTA measurement technique is absolute, so it requires no calibration. Each particle is analysed individually and simultaneously, giving us an in-depth understanding even of complex samples. The Nanosight measures nanoscale particles in the range of 10 nm to 1000 nm, including biological and low-scattering particles.

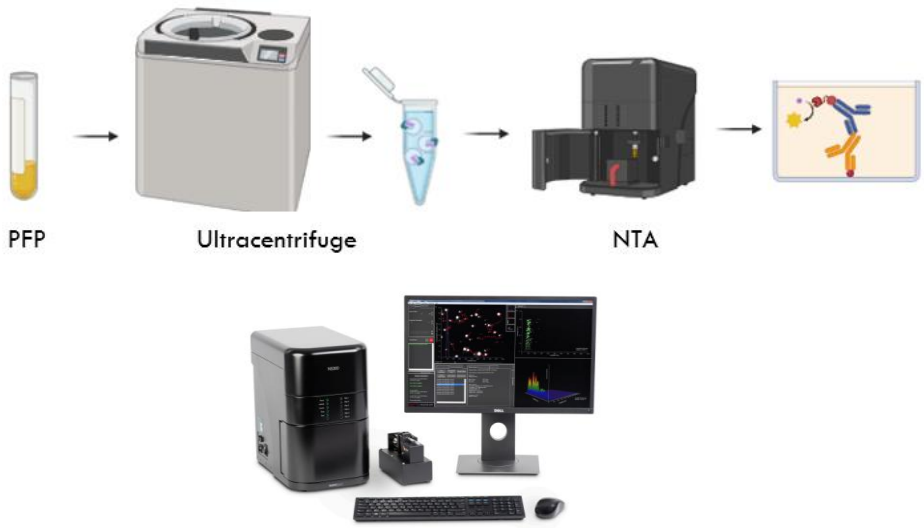


Figure 23: Experimental workflow to isolate and quantify vesicles from plasma aliquots (created with BioRender) and image representing Malvern Panalytical's NanoSight Pro® with syringe pump and software used (https://www.malvernpanalytical.com/).

For this purpose, samples of diluted EVs were thawed or prepared (1:50-1:200) for the analysis. First, the instrument must be prepared by inserting the Quartz chamber and check the connection of fluidics. At this point, the light source (laser) illuminates the particles, making them visible. The goal is to obtain 30-150 particles/frame (<30 is too few particles, greater than or equal to 200 is excessive concentration, causing suboptimal reading); after loading the sample and insert the syringe into the fluidic system, the quartz chamber must be tilt of 45° and apply controlled pressure to avoid bubbles. The analysis was run with the following workflow and parameters: focus, initial infusion rate 100-500 (to check flow), analysis infusion rate 30; 5 captures of 60 seconds, Detection Threshold 5, Temperature 25°C; then the analysis occurred and the validity was checked calculating:

$$Traces\ Ratio = Total\ number\ of\ traces\ /Number\ of\ correct\ traces$$

The analysis was considered Valid if <5 or Invalid if ≥5. Post-analysis cleaning was mandatory after each sample, flushing with 2 mL of pure distilled water, followed by 3 mL of air in the fluidic system while for the final cleaning of the instrument and the quartz chamber 10% ethanol was used.

III-6 RNA Extraction and Transcriptomics

III-6.1 Total RNA extraction and Purification

Total RNA, including small RNA species, was extracted from plasma using Invitrogen™ TRIzol™ LS Reagent, formulated for liquid samples. TRIzol™ LS contains phenol and guanidine isothiocyanate to lyse samples and inhibit RNases while enabling isolation of RNAs of varying sizes, following the single-step method of Chomczynski and Sacchi (Chomczynski and Sacchi, 1987). After homogenization with TRIzol™ LS, chloroform is added and the mixture separates into an aqueous phase (RNA), an interphase, and an organic phase (DNA and proteins). RNA is then precipitated with isopropanol, washed with ethanol, and resuspended for downstream use.

In our case, plasma samples were thawed and a ratio of 1:4 volume (0.25 mL of sample per 0.75 mL of TRIzol™ LS Reagent) was used; the isolation or purification steps were performed in according to Invitrogen™ protocol.

III-6.2 Transcriptomics and RNA-Sequencing

RNA sequencing (RNA-Seq) also known as Transcriptome sequencing or massively parallel RNA-Seq, is a powerful tool that offers various of information in gene expression and how changes drive biological processes after expression. However, RNA-Seq experiments generate a large amount of data that must be managed and analysed with careful. RNA-Seq data analysis is indeed a complex procedure requiring several types of data processing, quality control (QC), and statistical analysis during each phase.

The workflow (**Figure 24**) begins with sample QC to confirm suitability for small RNA-Seq, followed by library preparation and QC. Libraries were sequenced using a single-end 50 bp strategy, and sequence data underwent QC before bioinformatic analyses.

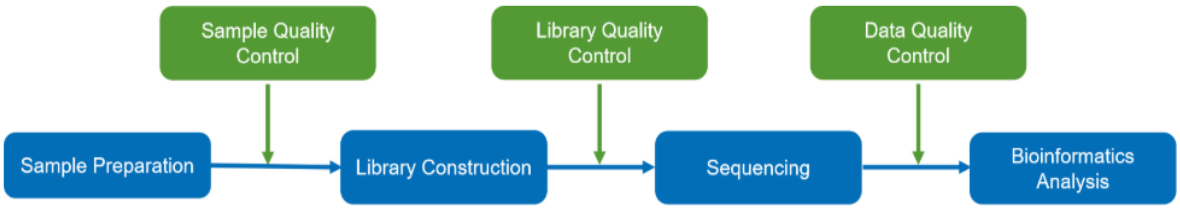


Figure 24: Scheme on **procedural workflow for small RNA sequencing** projects (https://www.novogene.com/).

Small RNA Method (Illumina): Experimental Procedure

Sample Quality Control

Depending on the Sample Type, the amount (µg), volume (µL) concentration**,** RIN (Agilent 2100™) and purity (NanoDrop™) are checked, to verify degradation or contamination.

Library Construction, Quality Control and Sequencing

Total RNA served as input. Briefly, 3' and 5' adapters were ligated to small RNAs, followed by first-strand cDNA synthesis using a reverse transcription primer. PCR enrichment generated double-stranded cDNA libraries. Size selection yielded libraries with inserts of 18–40 bp. Libraries were quantified and pooled for sequencing on an Illumina NovaSeq 6000 with single-end 50 bp reads, aiming for ≥10 million reads per sample depending on library concentration and project requirements.

Bioinformatics Analysis Pipeline and Data Quality Control

Raw Data

Illumina output fluorescence images are converted to short reads by base calling and stored in FASTQ format (Cock et al., 2009). Raw reads were processed with fastp to remove adapter contamination, reads with poly-N, and low-quality reads, producing clean reads. Q20, Q30, and GC content were computed. All downstream analyses used these high-quality clean reads (**Figure 25**).

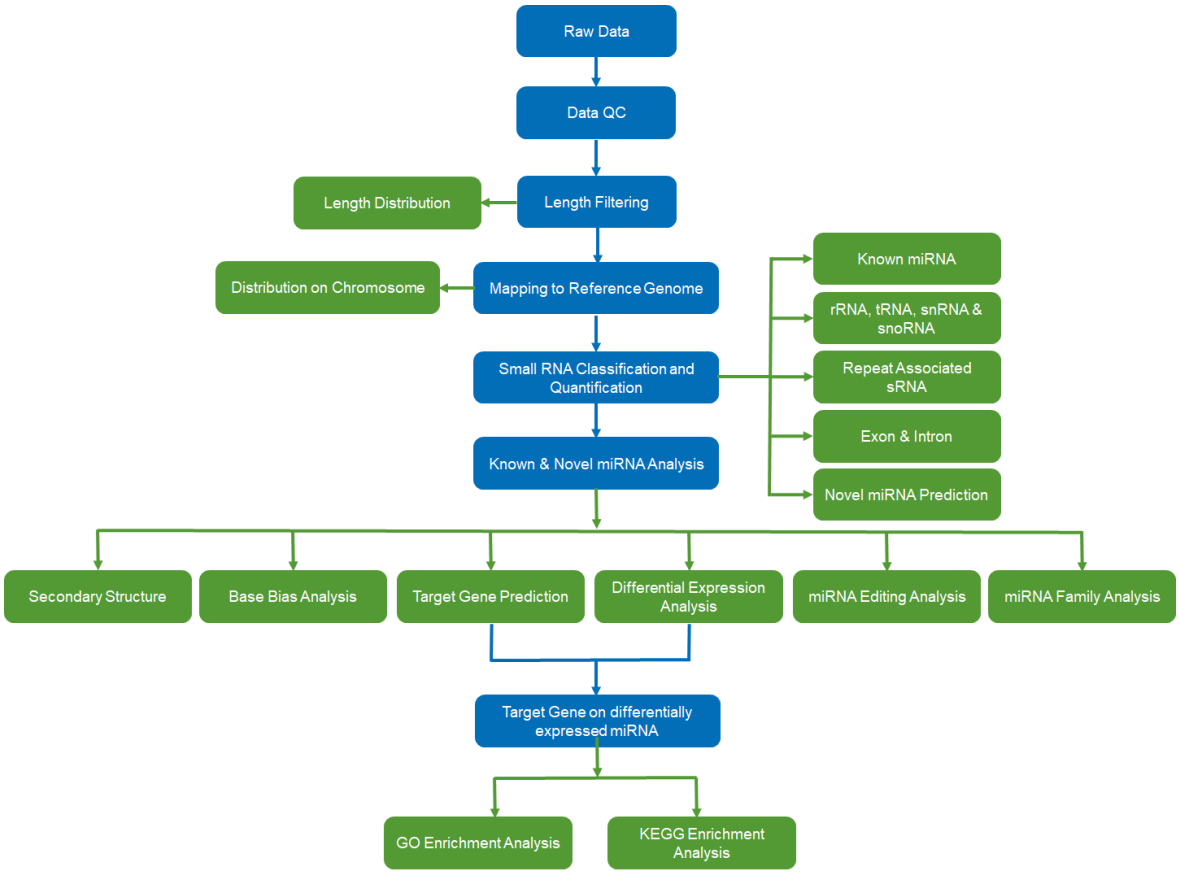


Figure 25: Bioinformatics Analysis Pipeline; Considering that data is obtained from animal with reference genome, the small RNA bioinformatic analysis process so occurred as is shown.

Evaluation of Data (Data Quality Control)

Quality control criteria included removing paired reads if either read contained adapter contamination, if >10% of bases were ambiguous in either read, or if >50% of bases had Phred <5.

Mapping to Reference Genome

To locate sRNA was used Bowtie (Langmead et al., 2009) and after length screening to the reference sequence with analysis of the distribution of sRNA on the reference sequence occurred.

Analysis of Known miRNA

Reads matched to reference sequences were compared with sequences in the specified range in miRBase. The sRNA information of each sample is obtained by mirdeep2 (v2.0.1.3) and SRna-Tools-CLI. Including the known miRNA secondary 3 structure, miRNA sequence, length, frequency of occurrence, etc.

Remove Reads from These Sources

Annotate sRNA using the ncRNA sequences of the species, or select rRNA, tRNA, snRNA, and snoRNA from RFAM to annotate sRNA, identifying and removing potential rRNA, tRNA, snRNA, and snoRNA. Using species-specific repeat sequence annotation information, or reference sequence information for de novo prediction of repeat sequences, align sRNA with repeat sequences to identify and remove potential repeat sequences, and statistically analyse the types and quantities of sRNA matching various repeat types.

Novel miRNA Prediction

The hairpin structure of miRNA precursors was exploited to predict novel miRNAs. Using miREvo and mirdeep2, sRNA sequences of a certain length are extracted and aligned with the reference genome. These sequences are then analysed for secondary structure, Dicer cleavage site information, energy characteristics, and other features to predict novel miRNAs. Statistically analyse the aligned sRNA sequences for their sequence, length, occurrence frequency, the distribution of the first nucleotide at different miRNA lengths, and the nucleotide distribution at each position for all miRNAs.

Small RNA Annotation

Summarize the alignment and annotation of all small RNAs with various types of RNAs. Since a single sRNA can match multiple different annotation categories, to ensure that each unique small RNA is assigned to only one annotation, it was followed the priority order: known miRNA > rRNA > tRNA > snRNA > snoRNA > repeats > genes > NAT-siRNA > novel miRNA.

miRNA Base Edit

MicroRNA may undergo nucleotide editing at certain positions, leading to changes in the seed sequence and subsequently altering target genes (Wei et al., 2009). By aligning the sRNA sequences from each sample with the detected known and novel mature miRNAs as well as their precursors, potentially mutated miRNAs can be identified.

miRNA Family Analysis

A family analysis was conducted of the detected known and novel miRNAs to explore the presence of their miRNA families in other species. Known miRNAs can be identified using miFam.dat to determine their family origin, while novel miRNAs can be classified using RFAM to determine their RFAM family.

miRNA Expression and Differential Expression

The expression levels were statistically analysed of known and novel miRNAs in each sample and normalized the expression levels using TPM (Zhou et al., 2010). TPM= (read count * 1,000,000) / libsize (libsize: total miRNA read count). For samples without biological replicates, the edgeR (v4.0.16) TMM algorithm was used to normalize read count data for analysis.

Target Gene Prediction for Known and Novel miRNA

Being animal samples, was used miRanda and RNAhybrid to predict miRNA target genes, taking the intersection as the final targeting result and performed a target gene prediction for the analysed known and novel miRNAs to obtain the relationships between miRNAs and their target genes.

Enrichment Analysis

It was used clusterProfiler (v4.8.1) to achieve functional enrichment analysis of differentially expressed genes in GO (Gene Ontology). Differentially expressed genes are significantly enriched, and padj < 0.05 as the threshold of significant enrichment. KEGG (Kyoto Encyclopedia of Genes and Genomes) is the main public database of pathway significant enrichment analysis, hypergeometric test is applied to identify the pathway of significant enrichment in candidate target genes, and padj < 0.05 as the threshold of significant enrichment. The differentially expressed genes in the KEGG pathway are analysed by clusterProfiler.

III-7 Statistical Analysis

In this thesis, categorical variables are presented as counts and percentages while continuous variables as medians and interquartile ranges (IQR). The Shapiro–Wilk test assessed normality. Diagnostic performance of plasma BEAMing and qPCR was evaluated against tissue KRAS, NRAS, and EGFR testing as the reference; sensitivity, specificity, and 95% confidence intervals (CI) were estimated. Agreement between plasma and tissue analyses was measured with Cohen's kappa statistic and its 95% CI. The Youden index was used to empirically estimate optimal cut-points for KRAS and NRAS mutant allele fraction (MAF) as potential diagnostic thresholds.

To compare EV levels between groups, an F test assessed variance differences and Welch's-corrected unpaired t-tests compared means. Correlations were evaluated using Spearman's correlation coefficient (r).

Analyses were performed using GraphPad Prism (v10). Two-sided p-values < 0.05 were considered statistically significant.

IV – RESULTS

IV - 1 Metastatic Colorectal Cancer Findings

In the study cohort were enrolled sixty-two mCRC patients; 56.5% were males and 43.5% females, with a median age at enrolment of 67 years (IQR 61–74). Demographic and clinical patient characteristics are summarized in **Table 1**.

Table 1: Clinical and demographic features of the enrolled mCRC patients in the study.

Patients features	Tot (n = 62)
Demographics	
Female	43.5%
Male	56.5%
Age at enrolment, median (IQR)	67 (61–74)
Histology	
Adenocarcinoma	88.7%
Mucinous adenocarcinoma	11.3%
Grading	
G2	41.9
G3	17.7%
G4	1.6%
Missing	38.7%
Site of primary lesion	
Colon	74.2%
Rectum	16.1%
Transverse colon	3.2%
Missing	6.5%
Stage	
IV (new diagnosis)	33.9%
IV (relapse)	22.6%
Missing	43.6%
Number of metastases	
1	45.2%
2	38.7%
3+	16.1%
Site of metastasis	
Liver	59.7%

Lung	29.0%
Loco-regional	19.4%
Lymph nodes	19.4%
Peritoneum	19.4%
Pleura	8.1%
Adrenal gland	4.8%
Bone	3.2%
Kidney	3.2%
Pancreas	1.6%
Endometrium	1.6%
Bladder	1.6%
Brain	1.6%
Surgery on primary tumour	
Yes	77.4%
No	22,6%
Chemotherapies	
Yes	75.8%
No	1.6%
Missing	22.6%
Chemotherapics	
Only synthetic agent	27.4%
Only targeted biologics	8.1%
Combination of biologics and synthetic	40.3%

IV-1.1 KRAS, NRAS, BRAF Mutational Status in Tissue Samples

KRAS, NRAS, and BRAF mutational status data were obtained on tissue using NGS and MALDI-TOF technologies at the hospitals of Pistoia, Pescia, and Careggi (Florence).

From the molecular analysis performed on tissue samples it emerged that: 63.8% of the patients (41 out of 65) have a KRAS mutation; 9.8% of patients (5 out of 51) have a BRAF mutation; only one patient (col008) has a NRAS mutation (**Figure 26**).

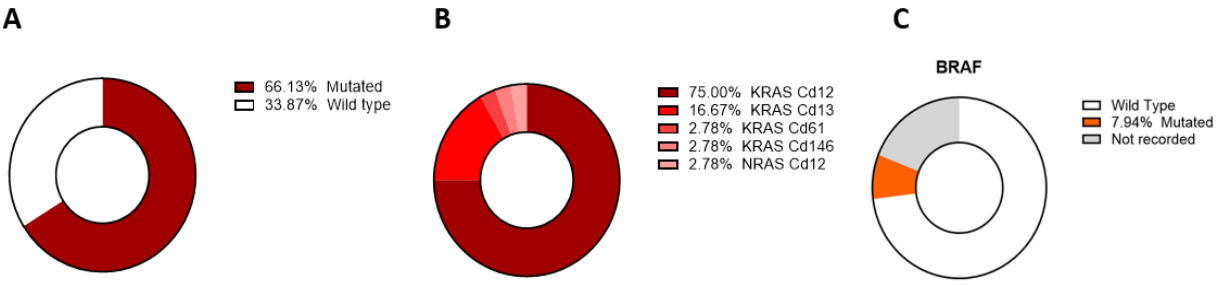


Figure 26: Tumour tissue profile of K-, NRAS and BRAF (A) RAS genotype frequency in the studied cohort; (B) K- and NRAS mutations frequency according to the most relevant codons; (C) Frequency of BRAF genotype in the cohort.

IV-1.2 Mutational Status Evaluation of KRAS and NRAS through BEAMing

Baseline RAS Mutational Analysis with OncoBEAM™

At baseline, the mutational status of KRAS and NRAS genes was evaluated in plasma samples using the OncoBEAM™ RAS CRC assay. Of the 56 plasma samples tested for KRAS, 43 (76.8%) were found to be mutated. Codon 12 was the most frequently affected site, being altered in 38 out of 43 mutated samples. KRAS mutations were also identified at codon 13 (6.1%) and 146 (1.5%).

NRAS mutations were found in 6 of 61 samples (9.8%), more precisely three at codon 12 and three at codon 61. In a few patients (e.g., oncobio021, col002, col018, ob10) concurrent KRAS and NRAS mutations were observed even though alterations in these genes are generally considered to be mutually exclusive (**Figure 27**).

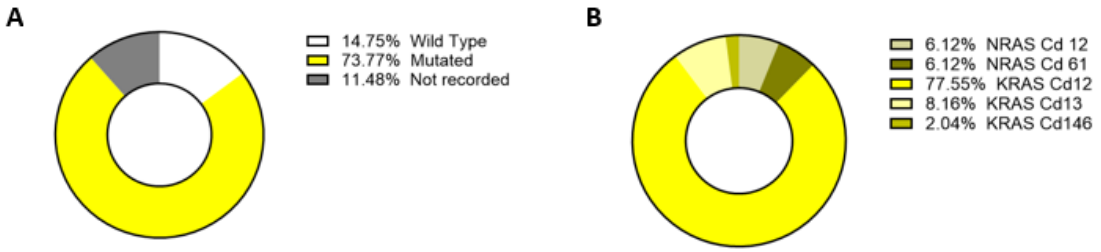


Figure 27: K-, NRAS and profile in plasma samples. (A) Frequency genotype of RAS in the cohort under study (Total= 62); (B) of K- and NRAS frequency mutations (Total = 49).

Representative dot plots obtained through flow cytometry are shown in **Figure 28**. Panel A displays the positive control (PC), Panel B the negative control (NTC), Panel C a KRAS codon 12 mutant and Panel D a KRAS wild-type sample. In these plots, mutant beads are in the bottom right gate, and their proportion varies depending on the mutant allele fraction (MAF).

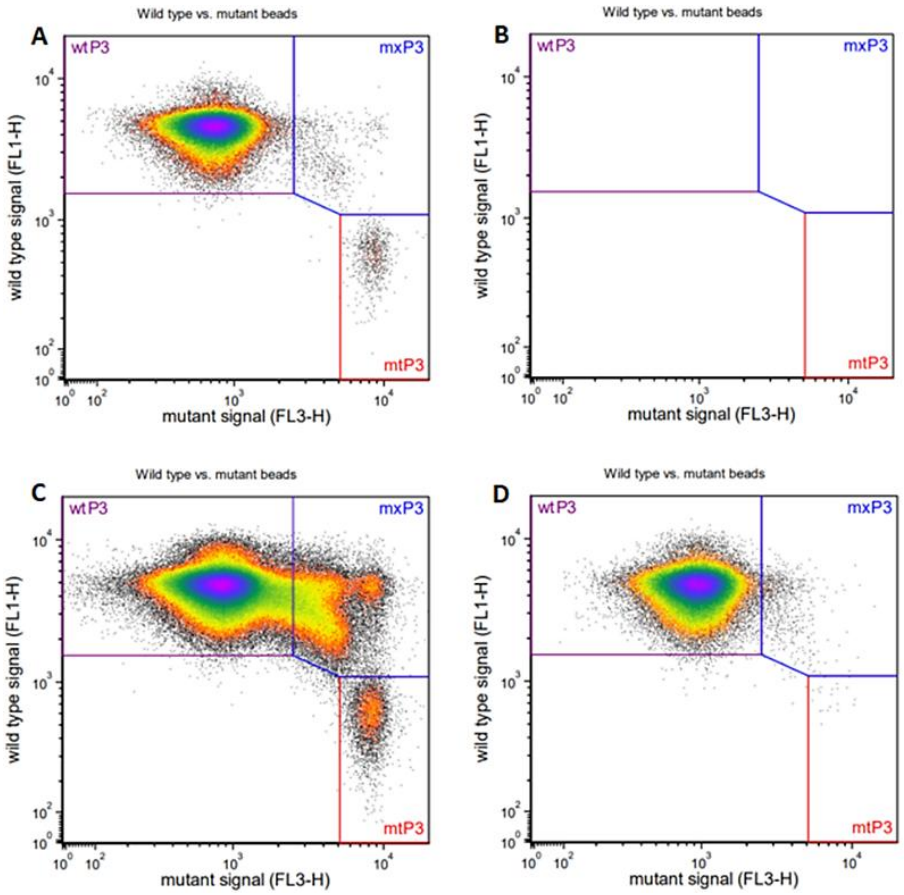


Figure 28: Examples of data obtained from BEAMing analysis: A) Positive control (PC); B) Negative control (NTC); C) KRAS-cd.12 mutant; D) KRAS-cd.12 wt.

Overall, 56.9% of the patients (37/65) carried a *KRAS* codon 12 mutation, confirming its predominance as most frequent in this cohort. A smaller proportion of patients showed *KRAS* codon 13 (6.1%) or codon 146 (1.5%) mutations. For *NRAS*, mutations were distributed between codons 12 and 61.

IV-1.3 Concordance Between Tissue and plasma *RAS* Mutational Analysis carried out through BEAMing

As a preliminary step, the diagnostic performance and concordance of *KRAS* and *NRAS* mutational analysis in plasma were compared with standard tissue-based testing (**Table 2**).

For *KRAS*, BEAMing showed a sensitivity of 89.5% (95% CI: 75.2–97.1%) and a specificity of 50.0% (95% CI: 26.0–74.0%). The overall agreement between tissue and plasma analyses was moderate (76.8%, Cohen's κ = 0.43 [0.17–0.68]), with a similar level of concordance for codon-specific identification (75.0%, Cohen's κ = 0.54 [0.33–0.75]).

For *NRAS*, BEAMing demonstrated high sensitivity (100%; 95% CI: 2.5–100%) and good specificity (91.7%; 95% CI: 81.6–97.2%). The agreement between plasma and tissue testing in distinguishing wild-type from mutant *NRAS* was fair (91.8%, Cohen's κ = 0.27 [–0.15–0.68]), and the concordance at the codon level was similar (93.4%, Cohen's κ = 0.32 [0.15–0.80]).

Table 2: Concordance between results obtained on tissue by NGS or MALDI-TOF and plasma obtained with OncoBEAM™ for *KRAS* e *NRAS* genes.

Gene	% of concordance	Cohen's K	Concordance
KRAS	76,8%	0,43	Moderate concordance
NRAS	91,8%	0,27	Good concordance

Importantly, the rate of concordance between tissue- and plasma-based *RAS* testing did not significantly vary with respect to sex, histology, tumour grade, primary tumour site, stage, metastatic location (liver, peritoneum, lung, locoregional lymph nodes), number of metastatic sites, surgery, type of chemotherapy, or patient outcome.

IV-1.4 *KRAS* and *NRAS* Mutational Status and Clinical Outcomes

We next assessed whether *KRAS* or *NRAS* mutational status was associated with clinical outcomes, including treatment response and survival. Treatment response data were available for 43 out of 62 patients. Among them, 3 patients (7.0%) achieved a complete response, 4 (9.3%) a partial response, and 8 (18.6%) maintained stable disease. In contrast, 20 patients (46.5%) experienced disease progression, while in 8 patients (18.6%) the evaluation could not be performed because it was too early. Survival information was available for 36 out of 62 patients. After a median follow-up of 254 days (IQR 95–447) from enrolment, 25 patients (69.4%) were alive, whereas 11 patients (30.6%) had died. No differences statistically significant were observed in treatment response or survival between patients with wild-type (wt) or mutated *KRAS*/*NRAS*.

As discussed in the introduction, *RAS* mutational status strongly influences therapeutic decisions, particularly the use of anti-EGFR agents, since the presence of *RAS* mutations is typically associated with absent response. For this reason, we specifically evaluated clinical outcomes in patients with wt *RAS* by tissue testing but mutated *RAS* by plasma BEAMing analysis. This subgroup included nine patients for *KRAS* and five for *NRAS* (with two patients showing discordance for both genes).

In this discordant cohort, disease progression occurred in 4 out of 9 patients (44.4%) with *KRAS* discordance and 3 out of 4 patients (75%) with *NRAS* discordance (data were unavailable for the fifth *NRAS* patient). Of the three discordant patients who received anti-EGFR therapy, only one progressed. Interestingly, this patient had wt *NRAS* in tissue testing but mutated *NRAS* at codon 61 in plasma, with a high MAF value (0.516).

IV-1.5 KRAS and NRAS Mutant Allele Fraction (MAF) and clinical features

The MAF median level for KRAS was 0.16 (with IQR 0.01–4.79 and range 0–28.15). Interestingly, patients with G2 tumour grading showed significantly higher MAF values [0.49 (0.02–7.37)] compared with those with G3 [0.01 (0.01–0.14)] or G4 (0.00) tumours (p = 0.025). Higher KRAS MAF values were observed in patients with liver metastases [0.33 (0.02–6.76) vs. 0.05 (0.01–0.44) without; p = 0.049] and in patients who had not undergone surgery for the primary tumour [5.46 (0.07–9.86) vs. 0.06 (0.01–0.92) with surgery; p = 0.010].

For NRAS, the median MAF level across the cohort was 0.007 (IQR 0.003–0.010; range 0.001–0.516). Notably, patients with mucinous adenocarcinoma had significantly higher values [0.027 (0.009–0.310)] compared to those with adenocarcinoma [0.006 (0.002–0.008); p = 0.004]. Higher NRAS MAF was also found in patients with lung metastasis [0.008 (0.006–0.017) vs. 0.005 (0.002–0.009) in those without; p = 0.025].

Based on these findings, we hypothesized that plasma KRAS and NRAS MAF values could serve as early biomarkers to detect liver and lung metastases in CRC patients, respectively. To explore this, we conducted a post hoc analysis. To identify liver metastasis, the performance of KRAS was moderate (AUC = 0.66, 95% CI: 0.51–0.80). The optimal empirical cut-off was 0.196, but sensitivity (0.58) and specificity (0.68) were poor. Similarly, for lung metastasis detection, the AUC for NRAS was 0.68 (95% CI: 0.54–0.82), with an optimal cut-off of 0.006, yielding moderate sensitivity (0.78) but poor specificity (0.51).

IV-1.6 Monitoring of KRAS and NRAS Mutational Status during therapy

In a subset of patients (n = 31), the MAF values of KRAS and NRAS were monitored every four weeks, at baseline and after 4, 8, 12 and 48 weeks from the beginning of the therapy, until eventual disease progression. For most patients, the mutational profile remained stable during treatment, meaning that the presence or absence of mutations in KRAS and NRAS was maintained over time but in other cases, some variations were observed. The distribution of MAF values across timepoints is shown in the scatter plot reported in **Figure 29**. While the majority of samples clustered in the 0.0–0.5 range, considerable variability between patients was evident.

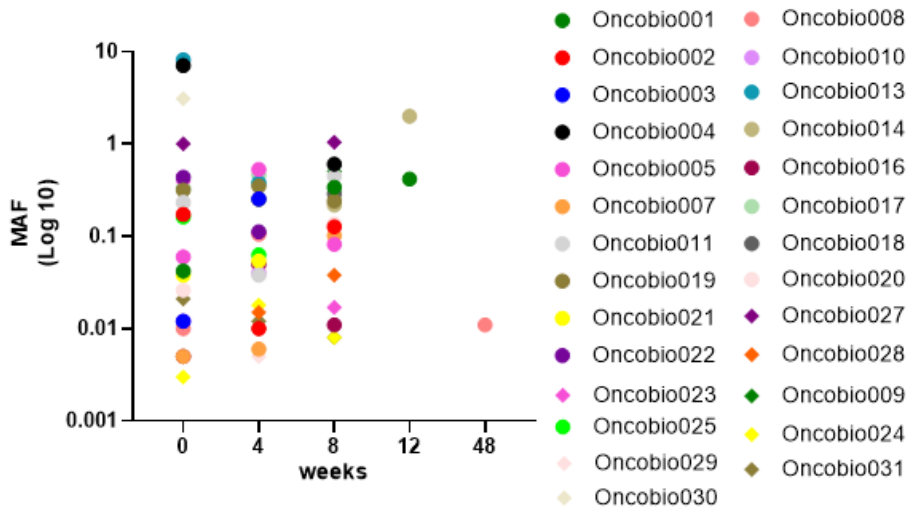


Figure 29: Scatter plot representing the distribution of Mutant Allele Fraction values during the course of therapy (at the baseline and after 4, 8, 12 or 48 weeks).

Representative curves (**Figure 30**) illustrate four different therapeutic responses associated with distinct MAF trends:
-Complete Response (red curve): Low baseline MAF, followed by a sharp increase at four weeks (likely due to the elimination of wild-type clones), then a decline to zero at eight weeks;
-Partial Response (green curve): A similar pattern, but MAF levels never reached zero;
-Stable Disease (brown curve): Comparable MAF values at baseline and after eight weeks;
-Progressive Disease (blue curve): Low baseline MAF that initially dropped to zero with treatment, but then rose sharply after four weeks, anticipating clinical progression.

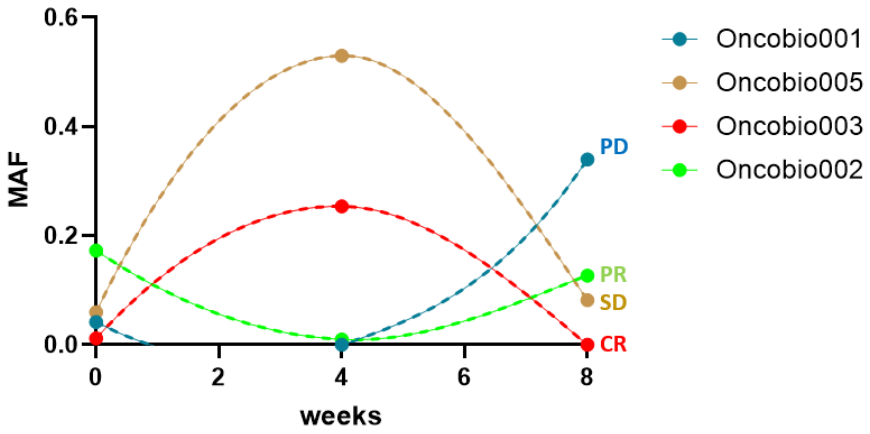


Figure 30: Akima spiline plots that shows the distribution of the Mutant Allele Fraction values over time (at the baseline until 8 weeks of treatment) for four exemplificative patients with different type of response. CR= Complete Response; PR= Partial Response; PD= Progression Disease; SD: Stable Disease.

In fact, when molecular progression (evaluated through MAF increase) was compared with radiological progression, molecular changes consistently appeared earlier, suggesting that plasma MAF could serve as an early biomarker of treatment resistance and disease progression (**Figure 31**).

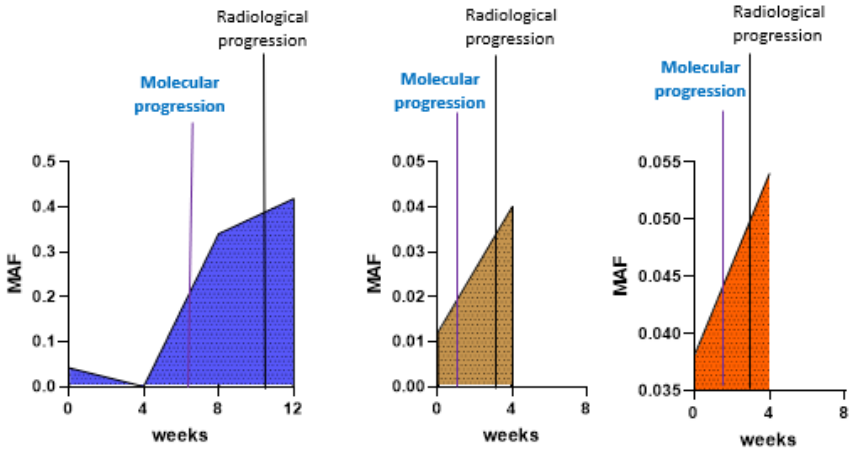


Figure 31: Mutant Allele Fraction values over time (at the baseline and during 4, 8, or 12 weeks of treatment) **for three** representative patients with Progressed Disease

Of particular interest was the detection of double mutations in four patients during therapy: Oncobio001 (12 weeks), Oncobio017 (4 weeks), Oncobio021 (baseline), and Oncobio030 (4 weeks). Oncobio001 and Oncobio017 exhibited two *KRAS* mutations (at codons 12+61 and 12+117, respectively), while Oncobio021 and Oncobio030 carried a *KRAS* codon 12 mutation alongside a *NRAS* codon 12 mutation.

IV-1.7 Mutational Analysis of *RAS* and *BRAF* Genes at Baseline with Idylla™

Using the Idylla™ system, mutational analysis of *RAS* and *BRAF* genes was performed on plasma samples by means of two different cartridges (ctKRAS and ctNRAS-BRAF-EGFR S492R) (**Figure 31**). Data obtained in the whole cohort are summarized in **Figure 33**.

Idylla™ ctKRAS Mutation Assay

KRAS GENOTYPE	MUTATION DETECTED IN KRAS CODON 12
Mutation	G12D
Protein	p.Gly12Asp
Nucleotide Change	c.35G>A
Cq of KRAS control	23.1
G12D Cq	26.7

Idylla™ ctNRAS-BRAF-EGFR S492R Mutation Assay

NRAS GENOTYPE	NO MUTATION DETECTED IN NRAS CODON 12,13,59,61,117,146
CQ OF NRAS CONTROL	34.6
BRAF GENOTYPE	NO MUTATION DETECTED IN BRAF CODON 600
EGFR GENOTYPE	NO MUTATION DETECTED IN EGFR CODON 492

Figure 32: Example of reports of the analysis performed with Idylla™ to search for mutations with ctKRAS or ctNRAS cartridges, showing the presence of absence of mutations.

Consistently with the literature, the most frequently detected alterations involved *KRAS*. Only one *NRAS* mutation was detected across all analysed samples, while a single *BRAF* mutation was observed (patient Oncobio010). For some samples, *NRAS* analysis could not be completed due to technical limitations.

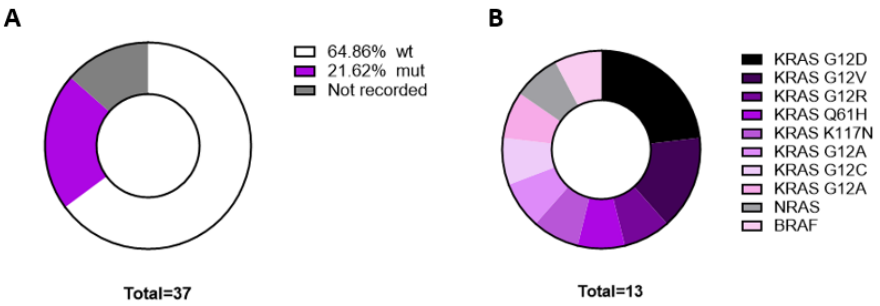


Figure 33: Mutational status of the *KRAS*, *NRAS* and *BRAF* genes from plasma analysis with Idylla™. (A) Genotype frequency and (B) Mutation type frequency with information on the specific type of mutation identified in ctKRAS of mCRC plasma samples.

IV-1.8 Concordance Analysis Between Tissue and plasma mutational status evaluated trough Idylla™

To assess the agreement between the mutational status in tissue (NGS-based) and plasma (Idylla™), Cohen's kappa (κ) was calculated. The concordance results are reported in **Table 3**.

Table 3: Degree of agreement between the results obtained on tissue with NGS or MALDI-TOF methods and those on plasma obtained with Idylla™ for the *KRAS, NRAS* and *BRAF* genes.

Gene	% of concordance	Cohen's K	Concordance
KRAS	54,5%	0,20	Low concordance
NRAS	94,4%	0,64	Substantial concordance
BRAF	84,2%	0,34	Low concordance

From the results of the concordance between the mutational status of the *RAS* and *BRAF* genes in tissue and that in plasma obtained with the Idylla™ system, we can observe a low concordance between the two analyses for the mutational status of *BRAF*, regarding the *RAS* genes, we can detect a substantial concordance for the mutational status of *NRAS*, while for *KRAS* the concordance is low, probably due to the lower capacity of the Idylla ™ system in detecting mutations present at low allele frequency.

IV-1.9 Concordance of *RAS* Mutational Status Between OncoBEAM™ and Idylla™

The mutational status of *RAS* genes obtained with OncoBEAM™ and Idylla™ on plasma samples was compared to evaluate concordance between the two methodologies. As shown in **Table 4**, concordance was substantial for *NRAS*, but low for *KRAS*. This indicates that some mutations detected with OncoBEAM™ were not identified by Idylla™, possibly due to its lower sensitivity (detection limit ~1% vs. 0.01% for BEAMing).

Table 4: Grade of concordance between results obtained with the analysis on plasma sample with OncoBEAM™ and Idylla™ for KRAS and NRAS genes.

Gene	% of concordance	Cohen's K	Concordance
KRAS	51,8%	0,22	Low concordance
NRAS	93,7%	0,63	Substantial concordance

Analysis of Discordant Samples

Several samples yielded discordant results depending on the method used. These discrepancies are mainly attributable to the reduced sensitivity of Idylla™, which may fail to detect mutations when present at very low allele frequencies.

A representative case was the sample Oncobio021 that, using BEAMing, was classified as mutated at codon 12 with a very low MAF (0.038%), whereas Idylla™ classified it as wild type. However, examination of the amplification curves revealed a signal for G12D that was lower than the threshold, as well as an amplification for codon 59, also below threshold.

In other discordant cases, no valid amplification was observed with Idylla™, again reflecting its lower sensitivity compared to OncoBEAM™. Moreover, this phenomenon also highlights tumour heterogeneity, since within a single tumour mass, subclones carrying different genetic alterations coexist. Highly sensitive methods like BEAMing are therefore capable of detecting multiple mutations in plasma at very low frequencies, even within the same gene.

IV-1.10 Comparison of the Mutation Type Between Tissue and Plasma evaluated through Idylla™

The Idylla™ platform not only detects the presence of mutations but also provides information on the specific type of mutation identified.

Table 5 reports representative cases for which both tissue and plasma data were available. In most cases, there was good concordance between tissue and plasma analyses. However, one sample (Oncobio023) was classified as wild type in tissue, while a mutation at codon 117 was detected in plasma. Cohen's k confirmed a moderate-to-substantial agreement overall [88.8% ($\kappa = 0.61$)].

Table 5: Comparison regarding the type of mutation found between NGS (tissue) and Idylla (plasma).

ID Patient	Tissue	Plasma
Oncobio001	KRAS Q61H	KRAS cd 61
Oncobio005	Codone 12	KRAS G12A
Oncobio005 4w		KRAS G12R
Oncobio014	KRAS G12V	KRAS G12V
Oncobio014 12w		KRAS G12V
Oncobio023	WT	KRAS K117N
Oncobio027	KRAS G12D	KRAS G12D
Oncobio027 4w		KRAS G12D
Oncobio027 8w		KRAS G12D
Col002	KRAS G12C	KRAS G12C
Col003	KRAS G12A	KRAS G12A
Col016	KRAS G12D	KRAS G12D

Additionally, the same table includes plasma samples collected at baseline and at 4, 8, and 12 weeks after the beginning of the therapy. In most cases, both the presence and type of mutation were maintained over time. The only exception was Oncobio005, which showed a G12A mutation at baseline and a G12R mutation at 4–8 weeks. This apparent discrepancy can be explained by the high intra-tumoral heterogeneity, where different subclones carrying distinct alterations may emerge or dominate at different time points.

IV-1.11 Metabolomics findings

To study the metabolomic profile of mCRC patients, ^{1}H-NMR spectra were acquired using a high-field 600-MHz NMR machine. The values of the relative plasma or urine concentrations of metabolites were estimated through the integration of the signals in the NMR spectra. From the acquired NMR profiles of urine and blood samples, even if preliminary given the limited number of patients analysed, it emerges that there are some interesting differences (attributable to individuals but also to conditions like post-intervention). As shown in the bar graph in **Figure 34** most metabolites are expressed at similar levels while other display slight differences between patients.

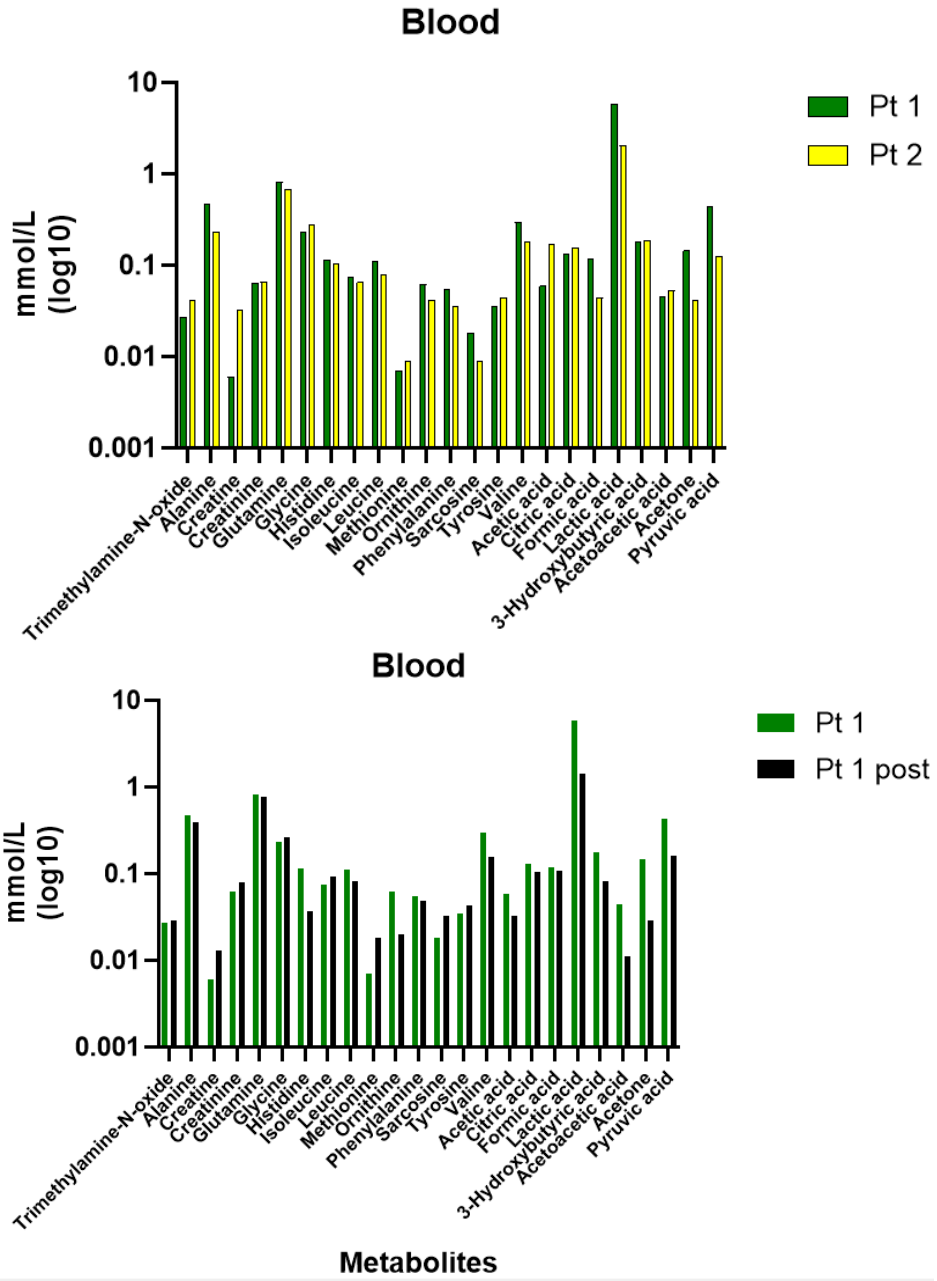


Figure 34: Metabolites distribution of quantities in plasma samples, n=3; Data with % Accuracy of detection <70% and in according with Limit of Detection (LOD) were not included in the graph. Pt1/Pt2 = Patient number 1 / 2; Pt 1 post = patient post chirurgical intervention.

Regarding blood metabolites we can observe in the plot that in the same patient, after surgery, the levels are broadly similar, suggesting stability in metabolic profile after treatment/intervention, although some small decreases are seen post-treatment (e.g., Formic acid, Lactic acid). Notably, these are also high-abundant metabolites, showing the highest levels (~10 mmol/L). Also, Creatinine, Glycine, and Glutamine are consistently elevated compared to others.

Moving to urine metabolites (**Figure 35**), concentrations vary widely (0.001–100 mmol/L). Creatinine is markedly higher in patient 1 (>10 mmol/L) than in patient 2 (~1 mmol/L). Patient 1 also shows higher excretion of several metabolites, including lactic acid, trigonelline, caffeine, and inosine. Distinctive compounds such as creatinine and 1-methylguanidine predominate in patient 1, while trigonelline and lactic acid are notably elevated in patient 1's urine. Differences in urinary metabolite levels, particularly creatinine, may reflect variations in metabolic rate, muscle mass, or renal handling. Post-treatment blood values for patient 1 show only mild changes, suggesting no major systemic metabolic disruption.

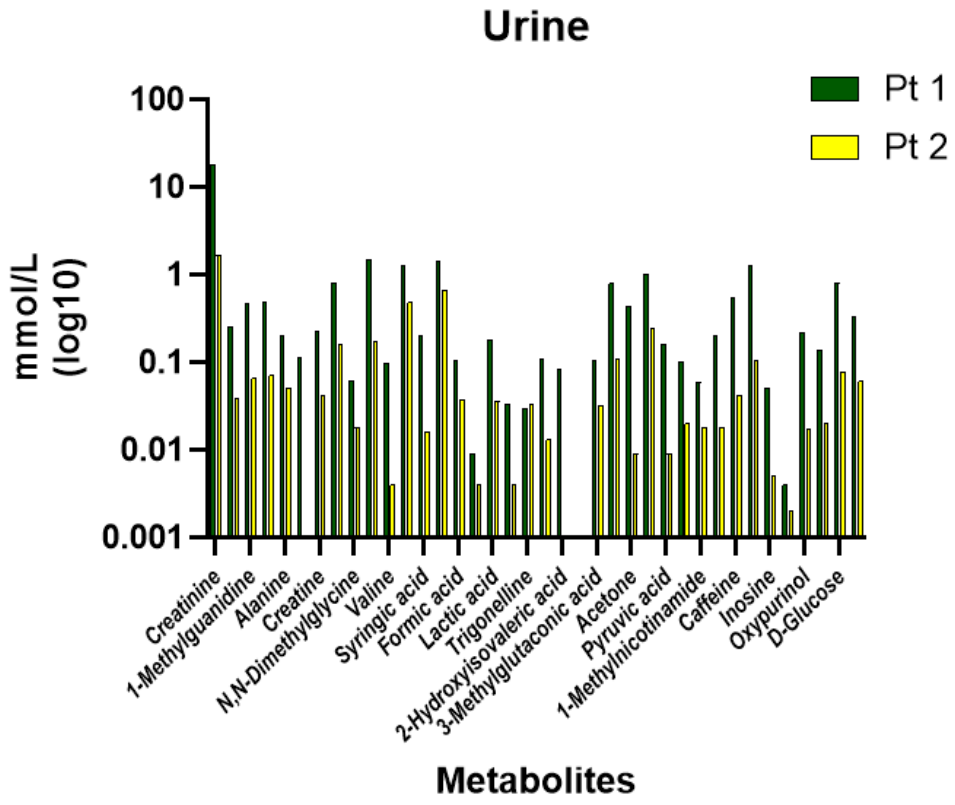


Figure 35: Metabolites distribution of quantities in urine samples, n=2; Data with % Accuracy of detection <70% and in according with Limit of Detection (LOD) were not included in the graph. Pt1/Pt2 = Patient number 1 / 2.

IV-1.12 Evidence of EVs in Patients Plasma

We compared the abundance and the size from 15 cases, 5 mCRC patients and 10 healthy controls.

Regarding the EVs Modes (**Figure 35**), we observed that negative control shows a very small particle size (around 123 nm), which is likely to be due to background noise deriving from the buffer, confirming minimal or absent contamination. The modal diameters of EVs in healthy subjects are broadly distributed, ranging from about 100 nm to more than 180 nm. The group median size is around 158 nm. The EV modal diameters in patients are clustered, ranging from about 160 nm to 175 nm. The group median size is around 163 nm, slightly larger than the healthy group. So, while the range for healthy subjects is broader, the overall median EV size appears very similar between controls and mCRC patients.

Regarding EVs concentrations (**Figure 36**), in negative control the concentration is extremely low confirming the DPBS is not a significant source of particles. The EV concentration in healthy subjects is highly clustered, and the median is approximately 5×10^9 particles/ml. The EV concentration in mCRC patients is slightly higher and more scattered, ranging from about 10^9 to $>10^{11}$ particles/ml and the median is higher than the healthy group. Thus, mCRC patients have a higher concentration of circulating EVs compared to healthy subjects, with some individuals showing concentrations an order of magnitude higher.

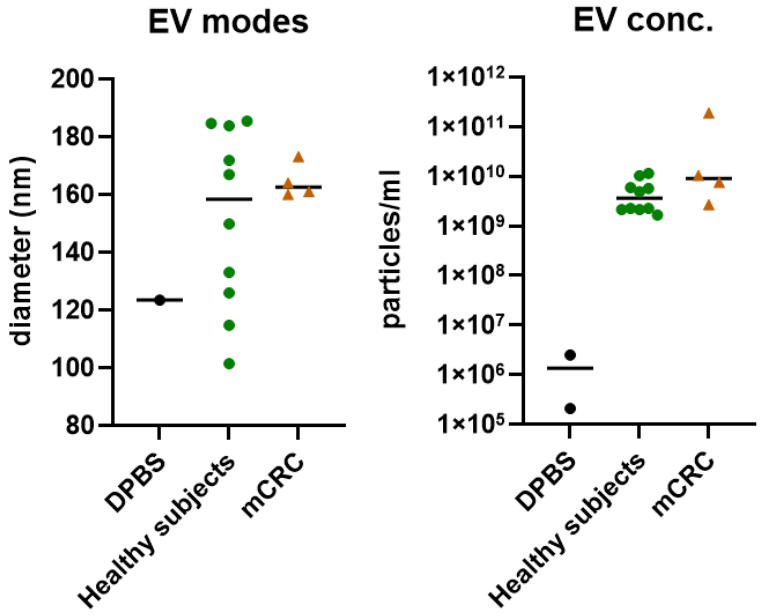


Figure 36: (L) The graph displays the concentration of Extracellular Vesicles in each sample. **(R)**, the graph displays the modal diameter (peak size) of the Extracellular Vesicles for each group. Each circle (green) or triangle (orange) represents the modal diameter measured in an individual subject's sample. The horizontal line represents the group median.

The graph in **Figure 37** provides the full-size distribution profile of EVs, plotting the concentration across the entire range of detected particle sizes, comparing the two biological groups. The green curve shows average distribution for healthy subjects, and the brown curve shows the average distribution for mCRC patients. The shadings of curves represent the variability across the samples within each group.

Both groups show a primary peak in concentration (modal size) between 100 nm and 200 nm. The majority of the particles fall into the exosome/microvesicle size range with a sharp decrease after the main peak. The mCRC curve remains above the control curve across the primary peak (~100–300 nm). The mCRC peak is broader and shows a slight shift toward larger sizes compared with controls, although the primary mode is similar. The shaded area is wider for the mCRC group (brown) in the peak region, indicating greater inter-patient variability in EV concentration and/or size distribution compared with the healthy group.

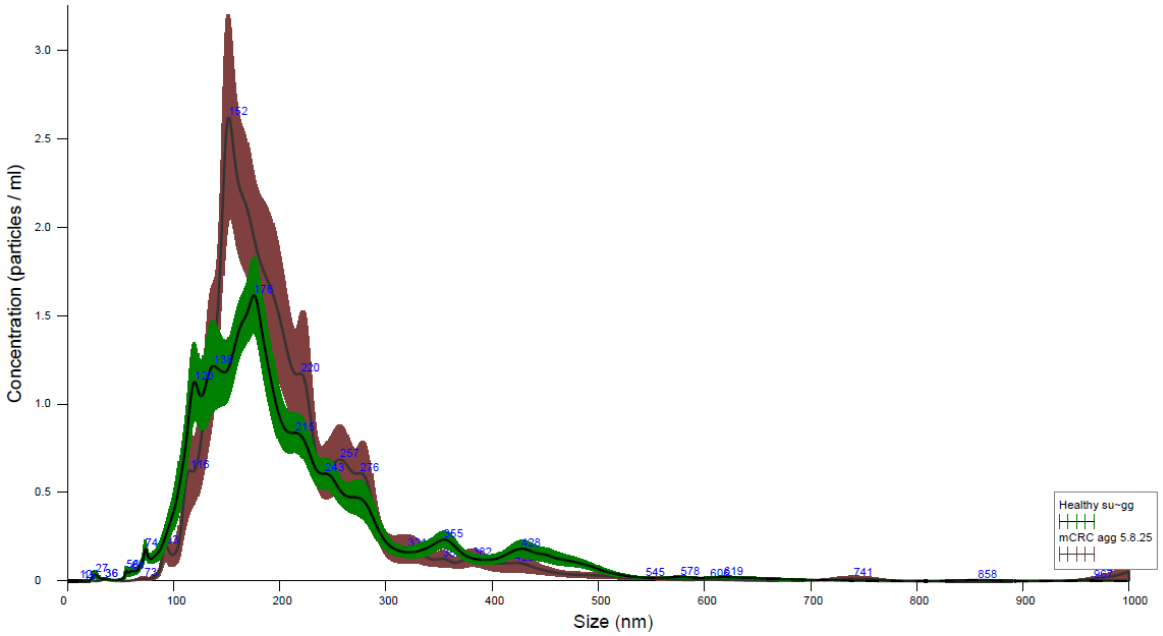


Figure 37: Averaged FTLA Concentration / Size for Experiments. Shading of colours represents error, bars indicate + / -1 standard error of the mean; mCRC patients n=5 vs healthy donors n =10.

Overall, no statistical differences were found (Unpaired t test for different SD with Welchs' correction, p value > 0,05). These preliminary findings suggest that EVs distribution may serve as a potential quantitative liquid biopsy biomarker for the detection or monitoring of metastatic colorectal cancer, while the overall EV concentration profile may be more informative in distinguishing the groups.

IV-3 Lung Cancer Findings

Fifteen patients suffering from metastatic LC were enrolled; 46,7% of them were women, with a median age at inclusion of 70 (61–74) years. The patients' clinical and demographic features are summarized in **Table 6**.

Table 6: Demographic and clinical features of the patients enrolled in the study.

Patient Features	Tot (n = 15)
Demographics	
Female sex	46,7%
Male sex	53.3%
Age at inclusion, median (IQR)	70 (61–74)
SARS-CoV2	
Infection	40%
Mild disease	20%
Serious disease	6% (1)
Long Covid	6% (1)

Histology	
Adenocarcinoma	66,7%
Squamous cell carcinoma	26,7%
Staging	
IV	80%
IV A/B	20%
Number of metastases	
1	13,3%
2	26,6%
3+	40%
Missing	20,1%
Site of metastasis	
Liver	20%
Lung	73%
Lymph nodes	73%
Mediastinum	13%
Pleura	20%
Adrenal gland	6% (1)
Bone	40%
Brain	6% (1)
Surgery on primary tumour	
Yes	33,3%
No	66,7%
Chemotherapy	
Only targeted biologics	20% (3)
Combination of synthetic agents and biologics target	80% (12)

IV-1.1 Mutational Status in Tissue Samples

Mutational status data were obtained on tissue using NGS at the hospitals of Pistoia, Pescia and Careggi (Florence).

From **Figure 38**, we can observe that: 46,7% of patients (7 out of 15) have a KRAS mutation on tissue samples; any of the patients have a BRAF mutation on tissue; NRAS was not evaluated in tissue samples; only two patients have an EGFR mutation and in 26% (4 out of 15) other variants with potential clinical relevance were detected.

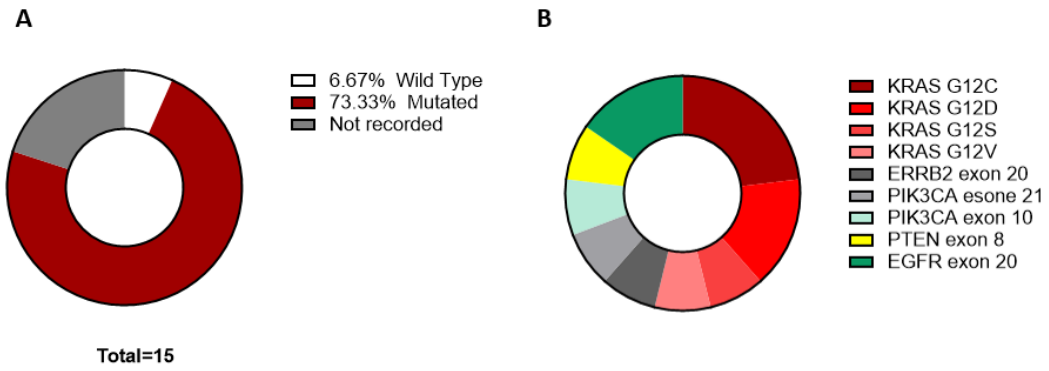


Figure 38: Mutational status in lung tumour tissues. (A) Frequency genotype in the cohort under study; (B) Frequency of specific type of mutations.

IV-1.2 Concordance Analysis Between Lung Tumour Tissue and Idylla™

To assess the agreement between mutational status in tissue (NGS-based) and plasma (Idylla™), Cohen's kappa (κ) was calculated. The results are reported in **Table 7**.

Table 7: Degree of agreement between the results obtained on tissue with NGS methods and those on plasma obtained with Idylla™ for the KRAS, BRAF and EGFR genes.

Gene	% of concordance	Cohen's K	Concordance
KRAS	69,2%	0,40	Moderate concordance
BRAF	100%	1	Perfect concordance
EGFR	91,7%	0,625	Substantial concordance

From results of the comparison between tissue and plasma mutational status for KRAS, BRAF, and EGFR using the Idylla™ system, we observed a generally high overall agreement. Concordance for KRAS was moderate, likely due to the Idylla™ system's reduced sensitivity for detecting mutations at low allele frequencies.

IV-1.3 Comparison of the Mutation Type Between Tissue and Plasma Using Idylla™

The Idylla™ platform as mentioned previously not only detects the presence of mutations but also provides information on the specific type of mutation identified, as highlighted in **Figure 39.**

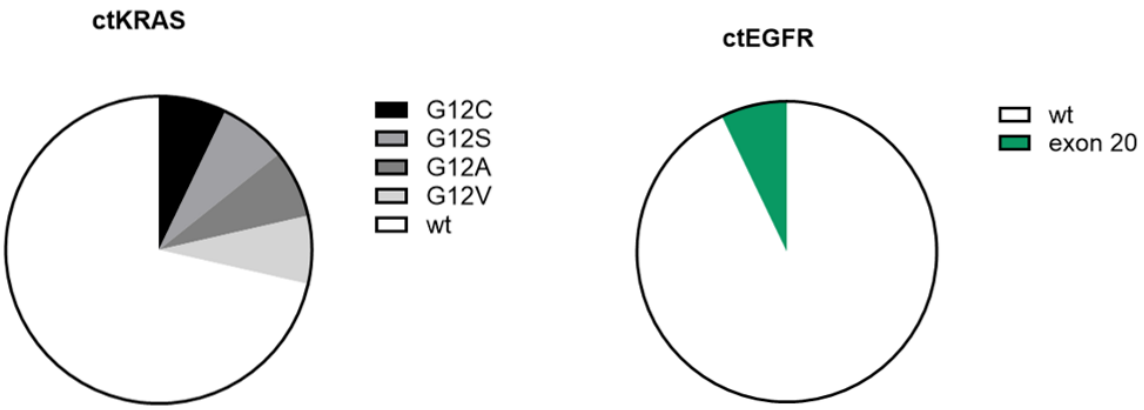


Figure 39: Pie charts of the specific type of mutation identified in ctKRAS and ctEGFR of LC plasma samples. Total number =15

In this subgroup, there was generally a moderate concordance between tissue and plasma: 75% concordance for *KRAS* mutation types with moderate agreement (Cohen's kappa = 0,55), and 50% concordance (Cohen's kappa = 0) for *EGFR* mutation type.

Regarding *EGFR*, an insertion in exon 20 was detected only in the plasma of patient CaPoS010, in agreement with the findings of the molecular analysis conducted on the tumour tissue. Conversely, for patient CaPoS006, the analysis on plasma did not detect the presence of alterations, while an insertion in exon 20 was detected in the tissue. The correlation analysis shows that there is a positive correlation between the results obtained on plasma and tissue, although the p-value is not significant probably due to the small sample size (**Figure 40**).

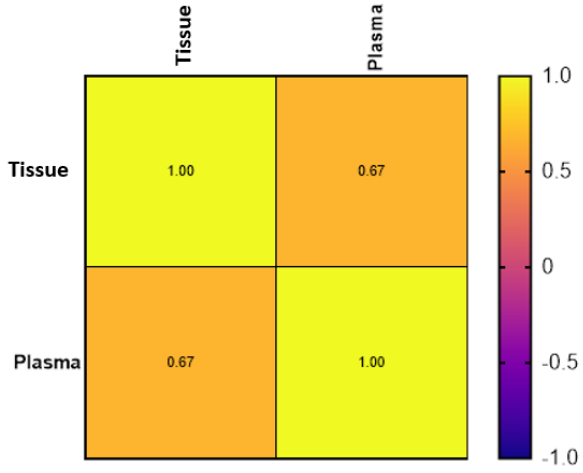


Figure 40: Correlation matrix between the analysis conducted on tissue and plasma samples for the assessment of *EGFR* mutations. Spearman r = 0.6667, p = 0.200.

IV-1.4 Monitoring of *KRAS* and *EGFR* Mutational Status Over Time in plasma

Through sequential sample collection, *KRAS* and *EGFR* mutational status were monitored every four weeks, from baseline and after 4, 8,12 and 24 weeks from the beginning of the therapy, until disease progression or death. The distribution of the expression levels of mutated genes across the timepoints is shown in the scatter plot (**Figure 41**).

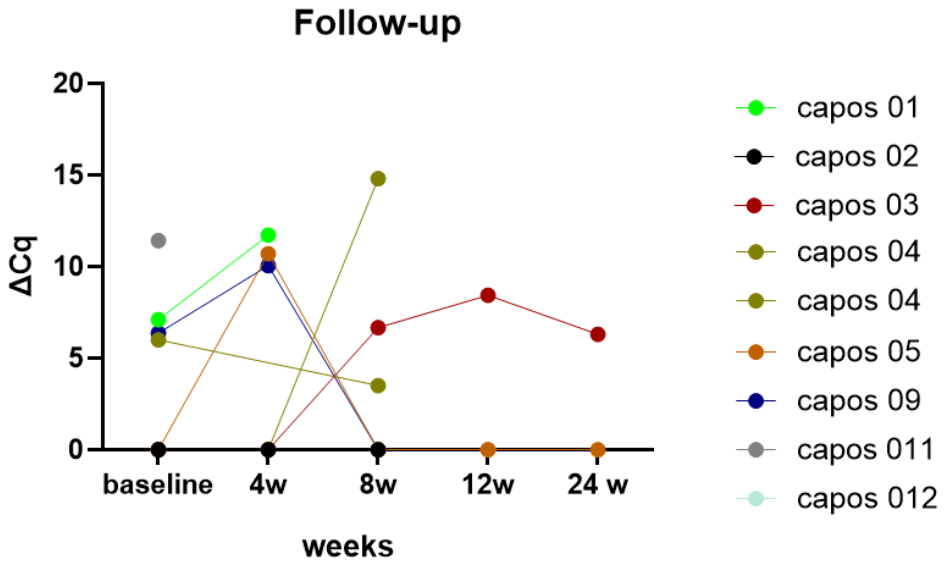


Figure 41: Scatter plot showing the distribution of the expression level of mutated *KRAS* (ΔCq) during the therapy (at the baseline and after 4, 8,12 and 24 weeks).

As can be observed in the plot, variations in mutated *KRAS* levels were observed across several patients undergoing molecular monitoring during treatment. Notably, patients CaPoS001 and CaPoS004 exhibited increased expression of the mutated gene at 4 weeks, followed by patient death (CaPoS001) and disease progression followed to second-line therapy (CaPoS004).

Another notably case is that of patient CaPoS009, who demonstrated a fluctuating mutational status pattern: elevated levels were detected at baseline and at 8 weeks, and undetectable levels at 4 and 12 weeks. This oscillation may suggest a differential therapeutic response, with a stronger effect at 4 weeks leading to the elimination of sensitive clones. Conversely, the rise at 8 weeks may indicate a diminished response, followed by a renewed suppression of the mutated fraction. This pattern may be explained by a delay in the administration of the second treatment cycle due to significant comorbidities, potentially resulting in an accumulation of mutated ctDNA molecules detected at the 8 weeks. Subsequently, a good response to treatment occurred and the mutated fraction decreased sharply until becoming undetectable.

Another relevant finding is the detection of a double mutation in the CaPoS004 sample at 8 weeks, both located on codon 12 of *KRAS* (G12R in addition to the G12C mutation already present at baseline). This observation suggests the emergence of multiple mutated clones and aligns with the clinical progression of the disease.

Patient CaPoS003 was wild type at baseline and at the first monitoring point (4 weeks) but developed a mutation (G12D) at 8 weeks, likely due to a reduced therapeutic response, resulting in the predominance of the mutated clone. Subsequent monitoring revealed the persistence of the mutation at 12 weeks, with increased levels (ΔCq 8.44) and a slight decrease at 24 weeks (**Figure 40**). Similarly, patient CaPoS005, initially wild type, developed a mutation at 4 weeks, followed by a decreased to undetectable levels (wild type) in subsequent monitoring. Finally, patient CaPoS012 remained wild type throughout the study, with no evidence of mutation development.

Analysis performed with the Idylla™ ctEGFRMutation Assay showed that all patients except one did not present mutations or rearrangements of the gene (8.3% mutated, 91.7% wild type) as showed below (**Figure 42**).

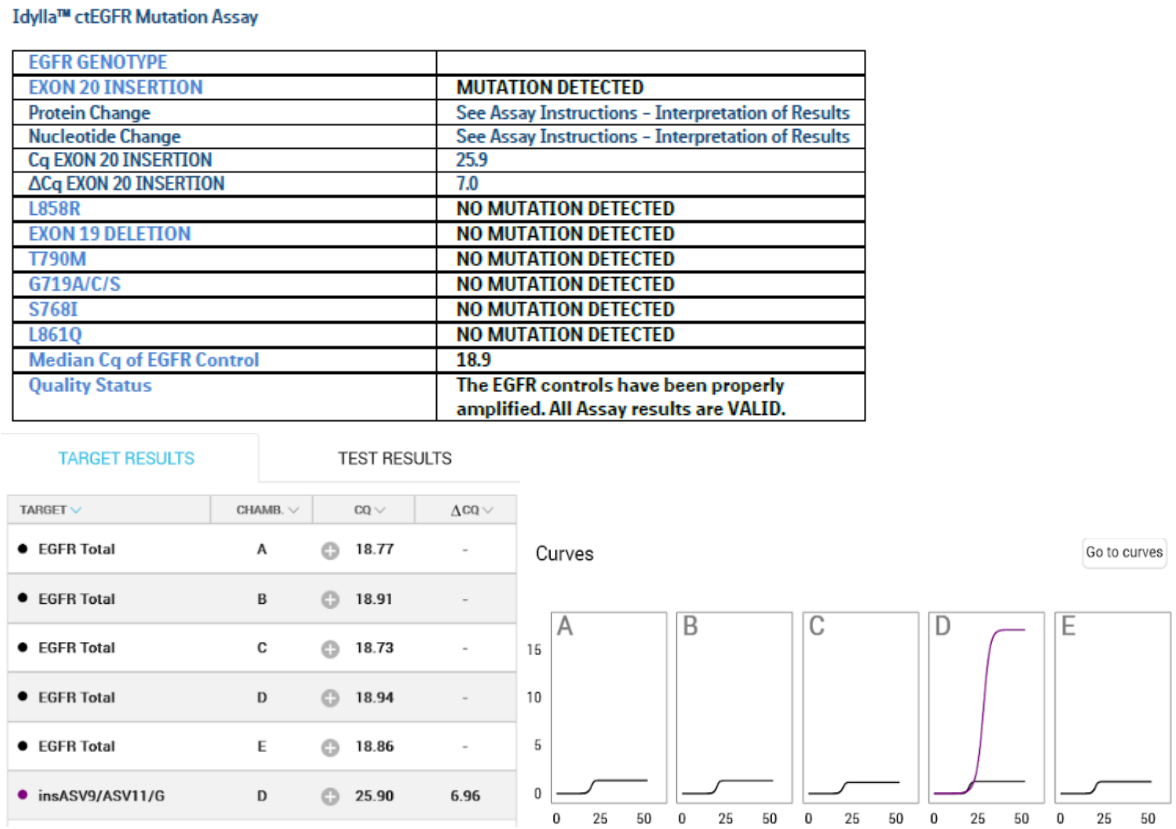


Figure 42: Example of a positive ct EGFRMutation Assay test report performed with Biocartis Idylla™.

High mutational levels were recorded at baseline and at the 4 weeks timepoint, followed by a sharp decline at 8 weeks, reaching undetectable levels. Subsequently, a marked increase was noted at 12 weeks, surpassing all previously recorded levels (**Figure 43**).

This trend aligns with the patient's clinical history. The initial treatment cycle consisted of Carboplatin (CBDCA) combined with Pemetrexed, which did not elicit a satisfactory therapeutic response, as evidenced by the persistence of the mutation in the plasma sample at 4 weeks. Consequently, Amivantamab, a bispecific monoclonal antibody targeting EGFR and MET, was introduced in the second cycle. The response to this regimen was good, resulting in the elimination of the mutated clone by week 8. Unfortunately, due to severe toxicity, Amivantamab had to be discontinued. This interruption was reflected in the plasma sample at 12 weeks, which showed a substantial resurgence in mutation-associated expression levels.

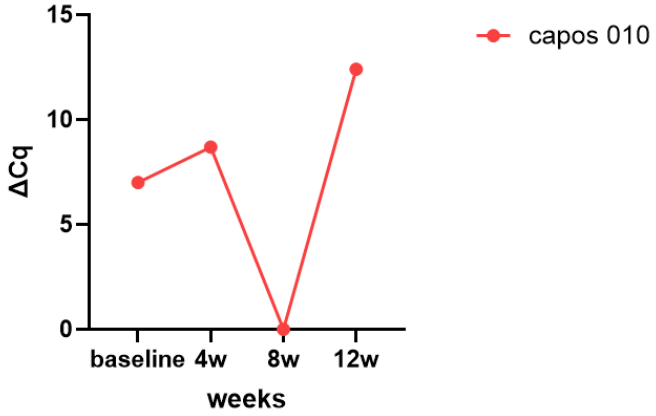


Figure 43: Scatter plot showing the expression levels of mutated EGFR (reported as ΔCq) during the therapy (at the baseline and after 4, 8 and 12 weeks) of a representative patient.

IV-1.7 EVs Distribution in Plasma Samples

With the same approach described for mCRC, using NanoSight Pro® by Malvern Panalytical it was possible to evaluate the profile of EVs in plasma samples, as shown in the graph below (**Figure 44**): it can be observed that there are differences in the distributions between healthy subjects and mLC patients with regards to the size and concentration of extracellular vesicles.

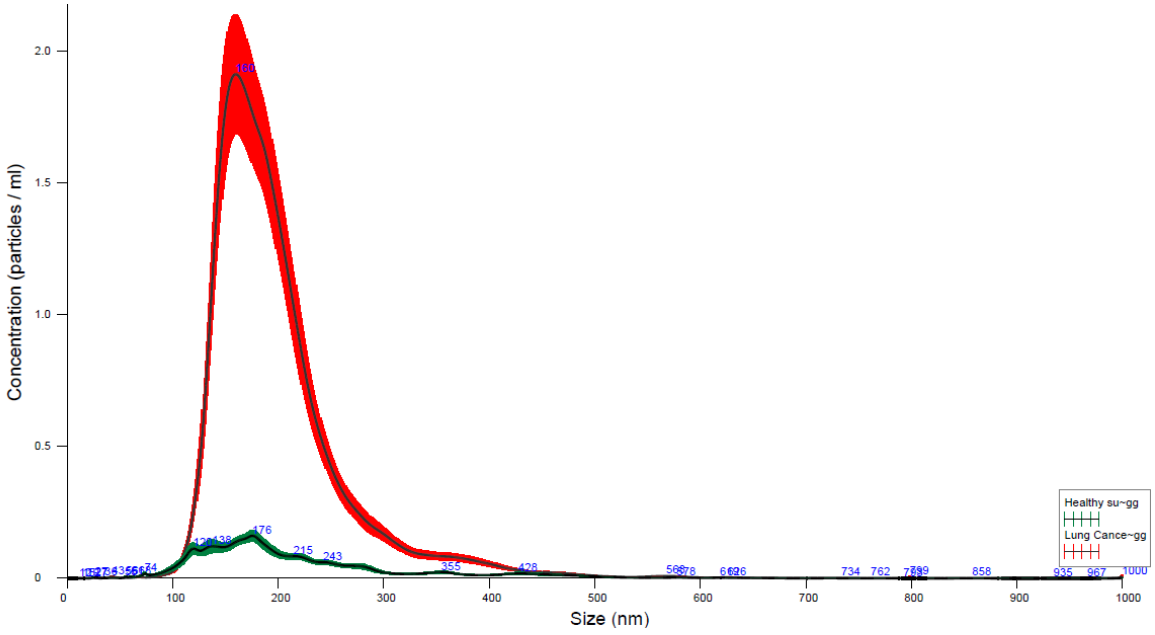


Figure 44: Graph of the size distribution and concentration of EVs. Comparison of the Averaged FTLA (Finite Track Length Adjustment) of mLC patients, n=16 in red versus healthy subjects, n=10 in green. Shading of colours represents error, bars indicate + / -1 standard error of the mean.

The graph shows the size distribution (x-axis) and concentration of extracellular vesicles (y-axis) using NTA in samples from healthy subjects (green curve) and lung cancer patients (red curve). For both groups, a greater distribution of EVs is observed in the size range between 150 and170 nm, with bigger differences in signal concentration. In particular, a higher concentration of EVs is observed in lung cancer patients; the concentration peak is extremely higher (around 2.0×10^8 particles/ml at 166 nm), confirming the overproduction

of EVs. The values for healthy subjects show a lower peak, around 176 nm, and reduced intensity even though the distribution appears broader and irregular. This latter factor may indicate a more heterogeneous vesicle population, a normal and physiological condition in the circulation of healthy subjects. These cohort of patients on the other hand seems to be associated with a strong enrichment of small EVs, like exosomes (~160 nm).

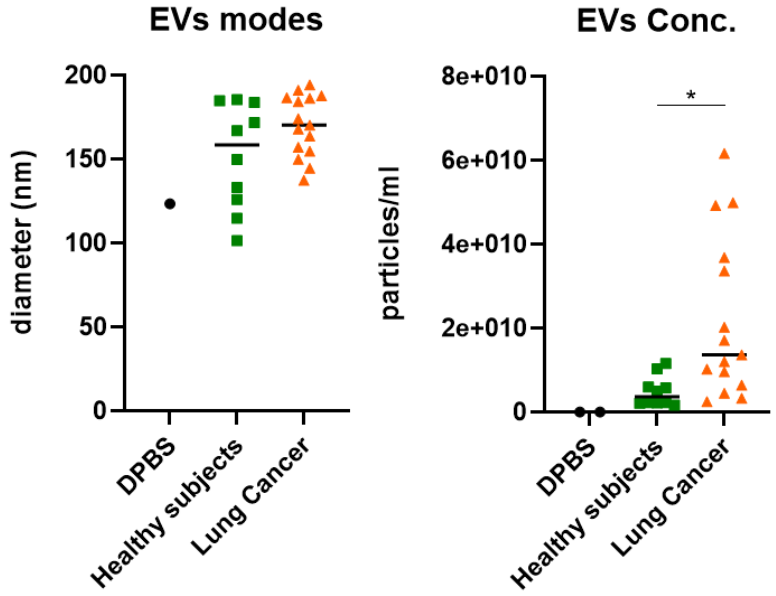


Figure 45: (L) Graph showing the analysis of the median size (mode) of EVs. Healthy control subjects are shown in green, while patients are shown in orange. **(R)** Graph displaying the concentration of extracellular vesicles in lung cancer patients (orange, n=16), versus healthy subjects (green, n=10), and negative control (black); * p value <0.05.

EVs size

The graph shows that the diameter of EVs isolated from lung cancer patients are only optically larger than that of healthy subjects but not statistically different (Student's t-test for unpaired samples, p value 0.069), indicating a possible alteration in vesicle size in the presence of tumour pathology (**Figure 45**-L). Healthy subjects EVs are mainly around ~150 nm in diameter, but with a wider spread (100–180 nm). While in lung cancer patients EVs are also centred around ~160–170 nm, but values are closer and higher (indicating a less variability). Therefore, the mode size of EVs is not drastically different between groups, but lung cancer EVs appear slightly larger and more uniform.

EVs Concentration

As expected, no relevant quantities of particles are detected in DPBS (negative control) compared to plasma samples, confirming the absence of contamination (**Figure 45**-R). In healthy subjects EV concentrations are relatively low (mainly below $\sim 2\times10^{10}$ particles/ml) while in lung cancer patients, a great variability in the concentration of extracellular vesicles is observed. In fact, it can be noted that the concentrations are higher, ranging up to $\sim 6\times10^{10}$ particles/ml. This increased concentration in cancer patients is statistically significant (Unpaired t test with Welch's correction, p value <0.05) and suggests that there is a greater release of EVs associated with the presence of tumour.

Thus, these preliminary results suggest that EV concentration may serve as a potential biomarker to distinguish cancer patients from healthy individuals.

IV-1.8 Gene fusion panel rearrangements in Broncho-Alveolar Lavage (BAL)

Since a major cause of NSCLC are chromosomal translocations that generate fusion genes, their fast and accurate diagnosis is crucial for successful treatment. Consequently, every patient diagnosed with NSCLC are recommended to be tested forALK, ROS1, RET, MET EX14 skipping and NTRK re-arrangements following international treatment guidelines as ESMO and NCCN. Here is why we analysed the broncho-alveolar lavage (BAL) of LC patients with the Idylla Gene Fusion™ Assay method. Specifically, this evaluation can be obtained analysing the cell pellet of broncho-alveolar lavage or from RNA extracted by the BAL cells. An example is shown in **Figure 46.**

Idylla™ GeneFusion Assay

OVERALL GENEFUSION RESULT	EQUIVOCAL
ALK FUSION	
ALK Specific Fusion	Not Detected
ALK Expression Imbalance	Indeterminate
ROS1 FUSION	
ROS1 Specific Fusion	Not Detected
ROS1 Expression Imbalance	Indeterminate
RET FUSION	
RET Specific Fusion	Not Detected
RET Expression Imbalance	Not Detected
Cq of the 5' RET	38.1
ΔCq RET Expression Imbalance (3' - 5')	-0.3
MET EXON 14 SKIPPING	
MET Exon 14 Skipping	Not Detected
Cq MET Wild Type	35.2

NTRK1	
NTRK1 Expression Imbalance	Equivocal
ΔCq NTRK1 Expression Imbalance (3' - RNA controls)	2.2
NTRK2	
NTRK2 Expression Imbalance	Not Detected
ΔCq NTRK2 Expression Imbalance (3' - RNA controls)	Not Available
NTRK3	
NTRK3 Expression Imbalance	Not Detected
ΔCq NTRK3 Expression Imbalance (3' - RNA controls)	Not Available
Cq of the RNA controls	35.8
Cq of the DNA controls	35.8
Quality Status	The RNA controls have been properly amplified. All assay results are VALID.

Figure 46: Report of Gene Fusion test result of an exemplificative BAL sample.

The test indicates that only aΔCq NTRK1 Expression Imbalance (3'-RNA controls) was detected and no other specific gene fusions or imbalance expressions.

IV-2 Idiopathic Pulmonary Fibrosis (IPF) Transcriptomic Insights

In this pilot study we analysed the profile of circulating miRNAs through small RNA sequencing (RNA-Seq) as mentioned above, in the plasma of 14 patients with IPF, 11 early IPF (E-IPF) patients, 6 LC patients and11 healthy subjects. All patients aged between 57-88 years and IPF patients had high-resolution CT features suggestive for definite or probable usual interstitial pneumonia (UIP). Total RNA underwent RNA-Seq and differential expression analysis using DESeq2. Functional enrichment of predicted miRNA targets employed GO and KEGG databases, and hierarchical clustering evaluated expression patterns.

IV-2.2 Visual display and normalization of transcript expression

After RNA-Seq data has been generated and the raw sequence reads have been subjected to quality control processing mapping to a reference genome was performed.

Length distribution and Mapping to Genome

The length of sRNA typically ranges from 18 to 40 nucleotides (nt). Analysing length distribution helps determining the composition of small RNA samples. For example, miRNAs are normally 21 or 22 nt, siRNA 24 nt, and piRNAs range between 28 and 30 nt. Length distribution varies between plants or animals but peaks at 22 nt are normal in animals (**Figure 47**).

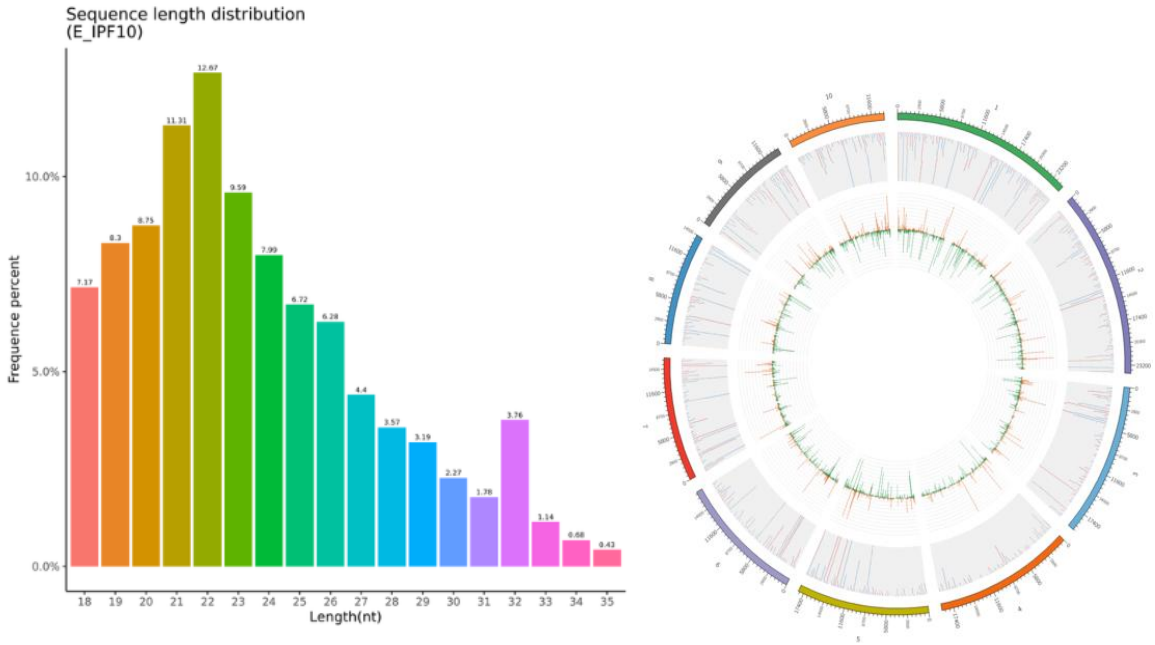


Figure 47: **(L)** Length distribution of total sRNA of a representative E-IPF sample. The x-axis reports the length of sRNA reads; the y-axis shows the percentage of one length read accounted for total sRNA. **(R)** Reads distribution of E-IPF 10 per chromosome.

sRNA mapping across chromosomes was visualised (circos plots) to analyse sense/antisense distribution and read density (Figure 46-R). Known miRNA mapping to miRBase yielded sequence, secondary structure, length, and abundance information. First-base biases and nucleotide distributions were examined. Reads mapping to ncRNAs, repeats, or other classes were annotated and filtered.

Analysis of known miRNA, ncRNA, prediction of novel miRNA and Repeat Sequences Alignment

The reads mapped to the reference genome were compared with specific sequences in the miRBase database to obtain detailed information, including the secondary structure of the mapped miRNAs, their sequence in each sample, length and the frequence. Novel miRNAs were predicted based on precursor hairpin structures (Figure 48-L) and repeat read classification was performed (Figure 48-R).

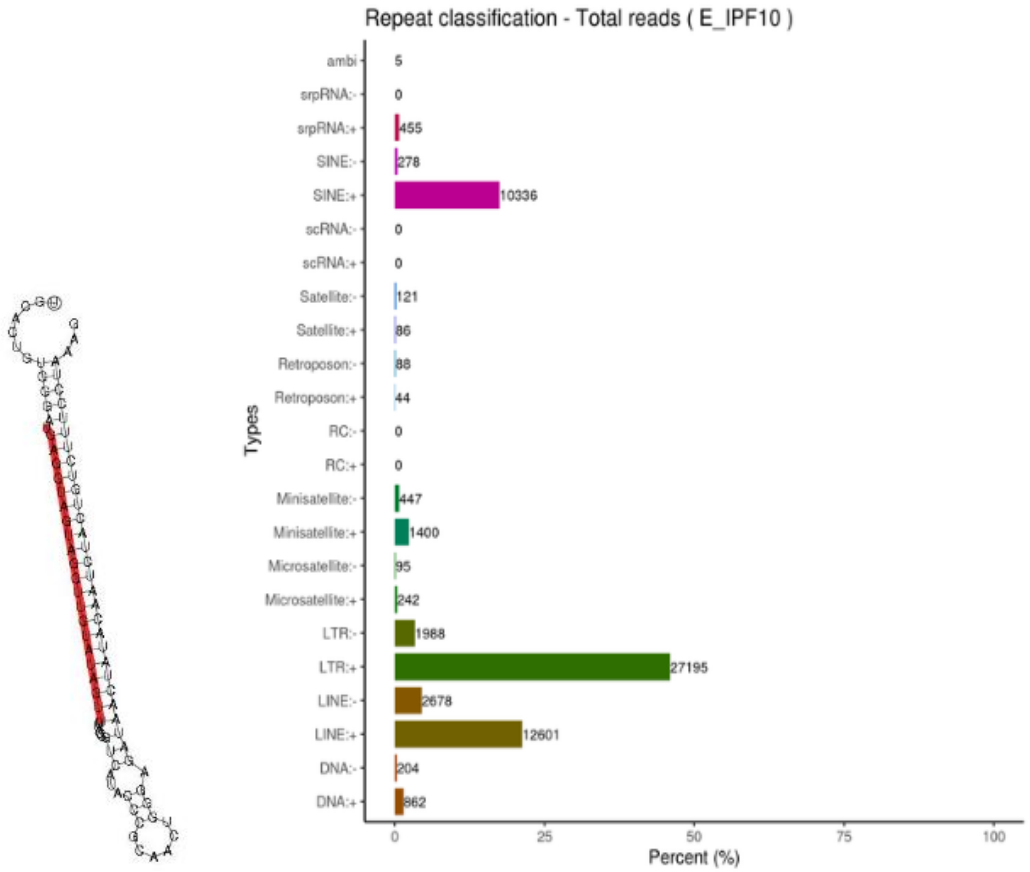


Figure 48: (L) Secondary structure of the known miRNAs on partial schematic matches. The entire sequence showed is a miRNA precursor while only the red sequence is mature miRNA. **(R)** Result of total repeat reads classification of a representative sample.

miRNA Expression and Differential Expression

In each sample the expression of known and unique miRNAs was normalized by TPM and statistically analysed (Zhou et al., 2010). The normalized expression = (read count*1,000,000)/libsize where "Libsize" is the read count of sample miRNAs. TPM density distribution can represent gene expression mode of samples. The results of the analysis in the whole cohort are shown in **Figure 49.**

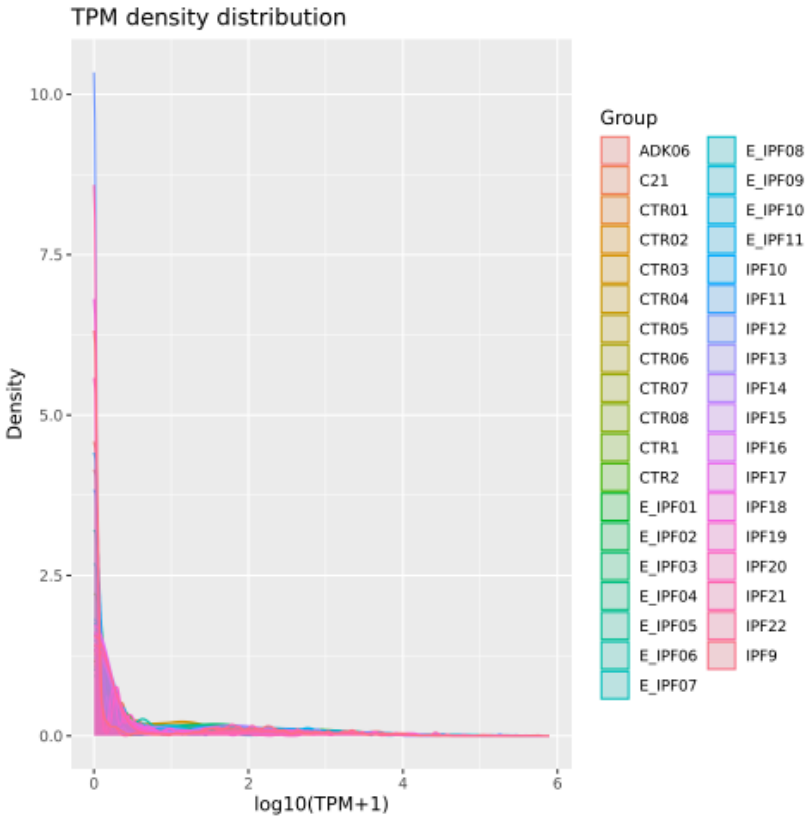


Figure 49: TPM distribution. The y axis report density and x axis is the value of miRNA $\log_{10}$(TPM+1).

RNA-Seq Correlation

Correlation analysis is often performed using clustering analysis methods to compare the transcript profiles of all samples in the data set. The purpose is to determine intra- and inter-group variability for samples and to identify outliers that were not excluded during earlier QC steps.

As shown in **Figure 50**, the heatmap is symmetric and despite the high overall correlation, distinct clusters or groups of samples with even stronger internal correlation are visible: The correlation between the major groups is visible as lighter blue areas, but it is lower than the correlation within the groups, confirming the groups represent different biological conditions with a certain degree of correlation. In the scatter plots it is possible to see how the Rho and Tau association coefficients are lower and this is reasonable since the compared samples (CTR vs IPF; E-IPF vs IPF) are derived from different experimental groups. To conclude, the correlation matrix demonstrates the quality and consistency of our data. The strong correlation within specific sample types suggests a high degree of reproducibility within those experimental conditions.

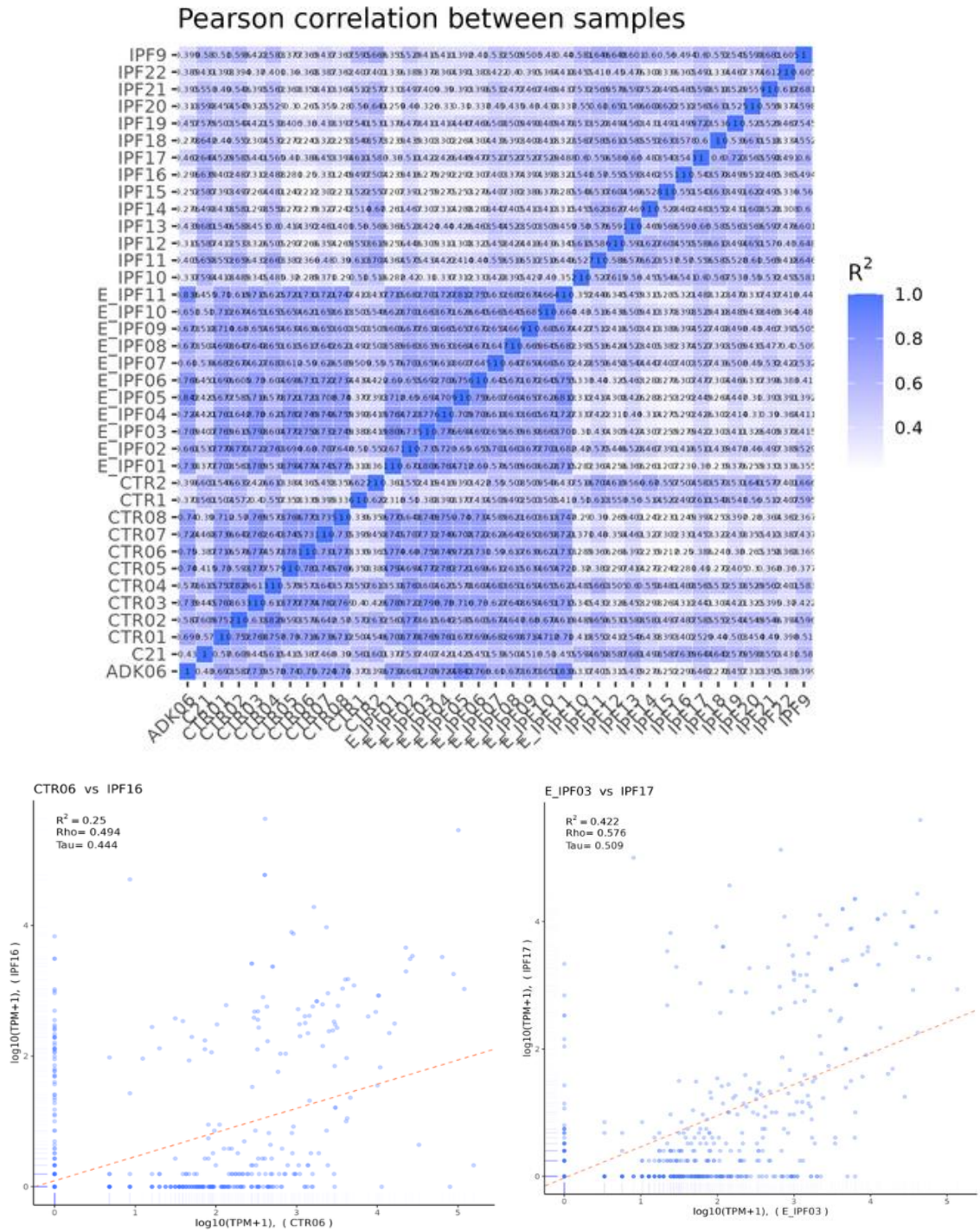


Figure 50: Correlation matrix between the samples under study and scatter plots between sample group types. The x axis and y axis represented the $\log_{10}$(TPM+1) (R^2: pearson RSQ); Rho: spearman coefficient of association ; Tau: kendall-tau coefficient of association).

Filtering the Different Expression miRNA

Volcano plots showing the comparison of the overall distribution of different miRNA expression in the various group of subjects enrolled is reported in **Figure 51**.

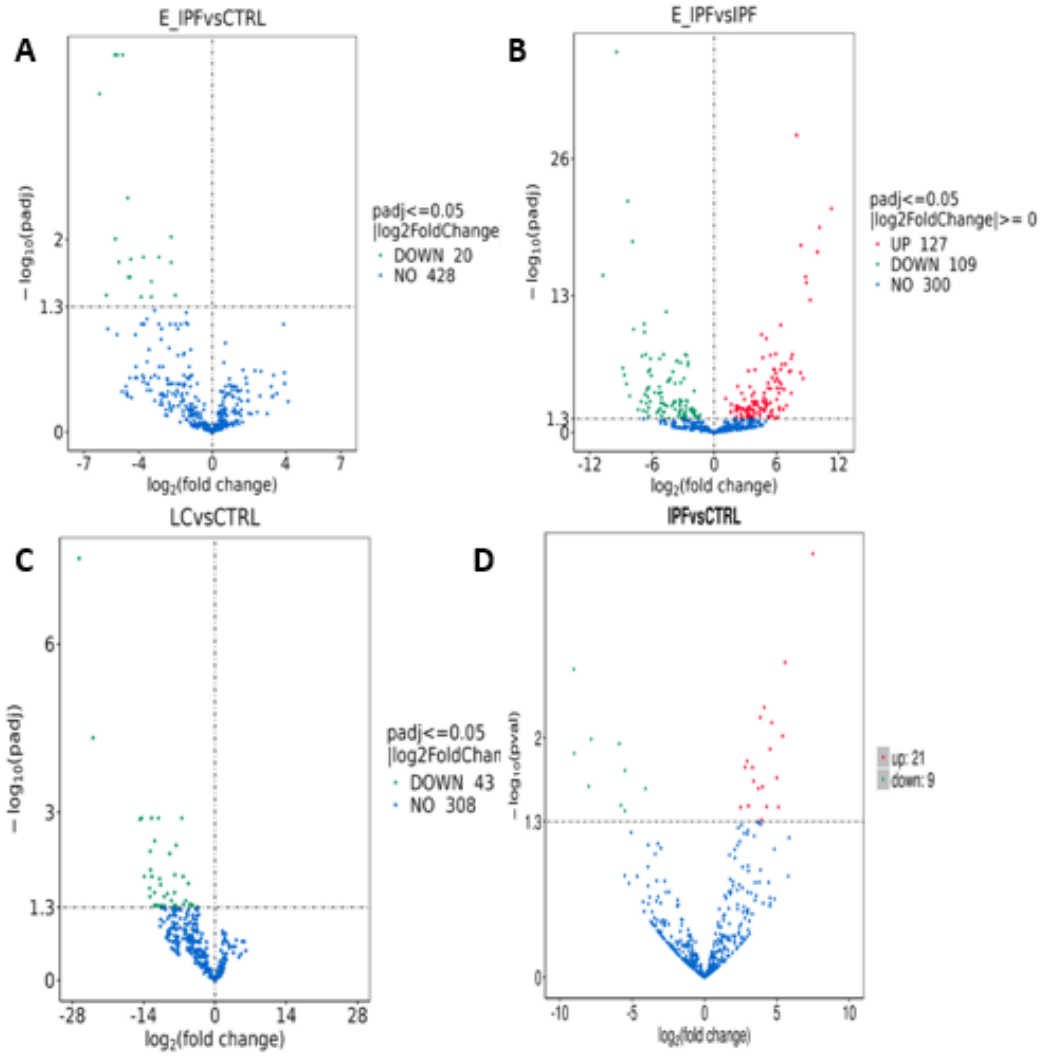


Figure 51: Volcano Plots of the comparisons of miRNA expression in the different groups. The y-axis shows the statistical significance of differences while the x-axis shows the fold change in miRNA expression between samples. Statistically significant differences are represented by red and green dots. A) E_IPF vs controls; B) IPF vs controls; (C) LC vs controls (D) IPF vs controls.

Thirty differentially expressed miRNAs were identified in IPF vs CTR, including 21 up-regulated and 9 down-regulated. This profile was found to be unique for this type of pathology, showing no overlap with circulating miRNAs identified in other fibrotic disease such us liver cirrhosis and autoinflammatory diseases such as Systemic Lupus Erythematosus and vasculitis like Behcet's syndrome (Emmi et al., 2022). Also 20 ci-miRNAs were differentially expressed in early-IPF compared to healthy controls, all significantly downregulated. Even these did not overlap with miRNA signatures from other fibrotic or systemic autoimmune conditions. Regarding LC samples it was possible to identify 43 miRNAs statistically downregulated compared to the control group.

Cluster Analysis of the Differences between miRNA Expression

Cluster analysis was used to evaluate miRNA expression patterns in the four groups. In this hierarchical clustering (**Figure 52**), there is high variability of expression among the various miRNAs in controls and patients, as indicated by the distinct colors. In particular, the left side of the plot shows data from controls and patients with early-stage disease, and greater homogeneity among these subjects is evident, in contrast with patients affected by IPF in a more advanced stage. It can be observed that for some miRNAs the expression profiles

are quite different between IPF and controls and are instead more similar within the members of each category. Surprisingly, the tumor samples display a profile more similar to that of the E_IPF group than to the more advanced forms.

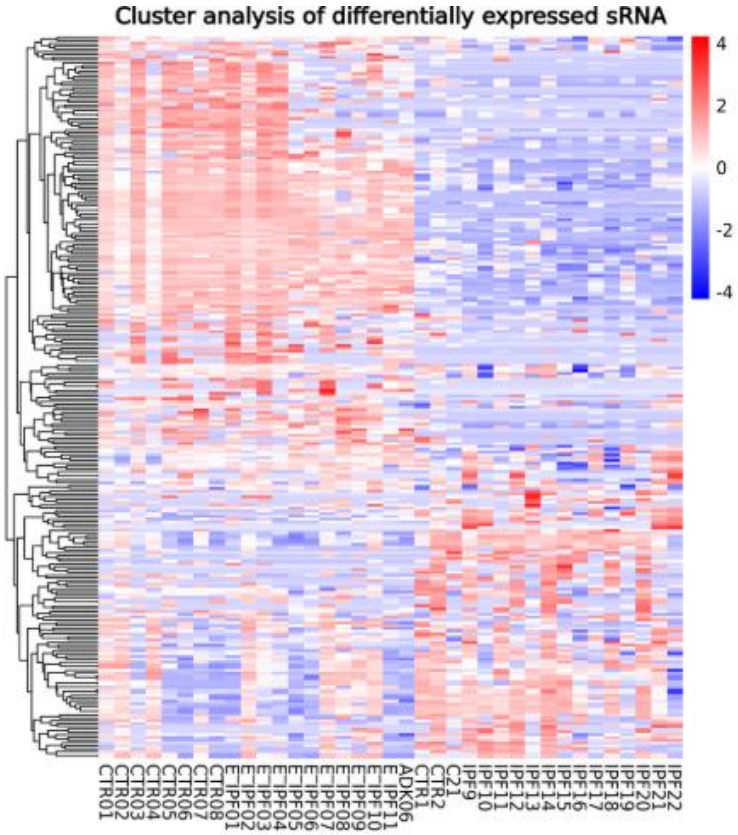


Figure 52: Cluster Analysis of differentially expressed circulating miRNAs in Controls, E_IPF, IPF and ADK. Red: highly expressed miRNAs; Blue: lowly expressed miRNAs. Shades of each color indicate higher or lower expression of individual miRNAs.

Gene Set Enrichment Analysis

The next step in RNASeq, after DGE analysis, was the gene set enrichment analysis, to assign biological meaning to gene sets of interest by placing RNASeq transcript expression data in an appropriate biological context. To investigate whether a gene set is enriched for genes that share a biological feature, such as a protein function, a signalling pathway, or a biological response, was used Enrichment analysis.

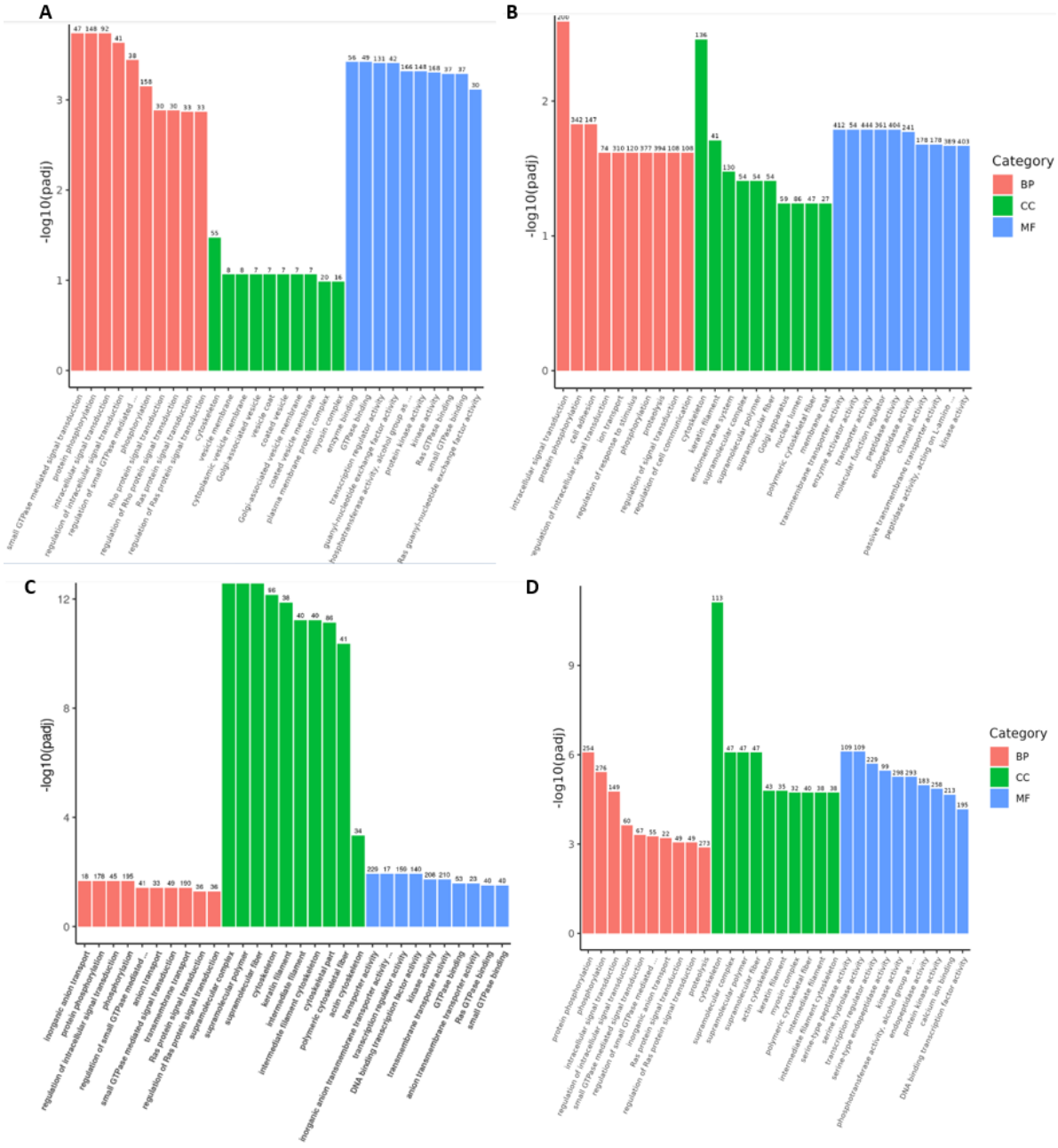


Figure 53: Functional enrichment analysis performed using Gene Ontology (GO). The x axis shows GO terms grouped by category (Red: Biological Processes, BP; Green: Cellular Components, CC; Blue: Molecular Function, MF). The y axis shows significance as -log10(padj).in A) E-IPF VS CTR B) E-IPF VS IPF C) IPF VS CTR (D) LC vs CTR.

Functional analysis using Gene Ontology (**Figure 53**) revealed several terms related to fibrosis, particularly signalling mechanisms mediated by GTPases and cell adhesion molecules. Also, this functional enrichment highlighted pathways related to cytoskeletal remodelling, GTPases-mediated signalling, and cell-matrix adhesion.

KEGG (Kyoto Encyclopaedia of Genes and Genomes) is a collection of manually curated databases related to genomes, biological pathways, diseases, drugs, and chemical substances. With Pathway enrichment analysis we identified significantly enriched metabolic pathways or signal transduction pathways associated with differentially expressed miRNA target genes related to the whole genome background.

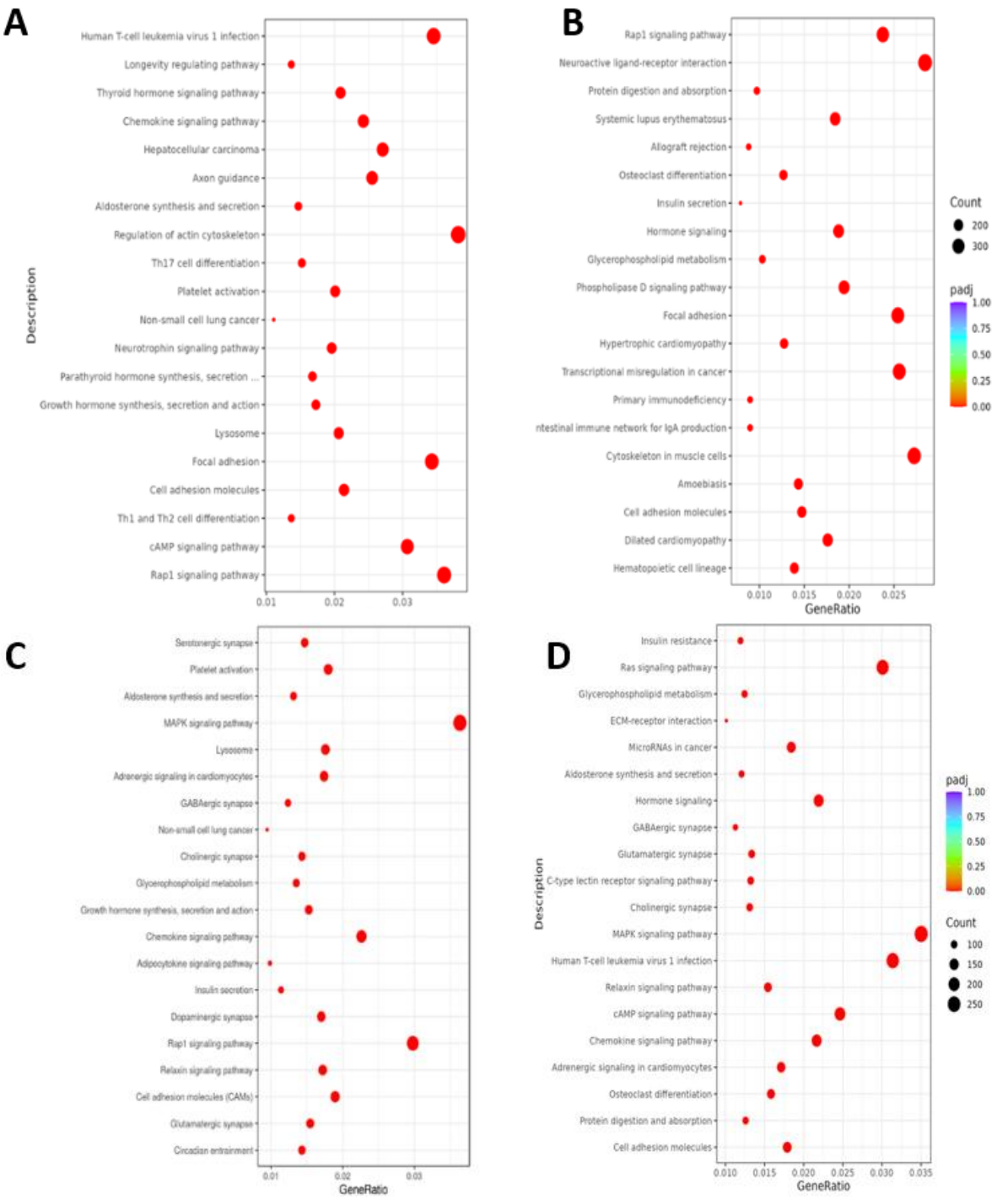


Figure 54: KEGG enrichment scatter plot of DEGs in A) E-IPF VS CTR B) E-IPF VS IPF C) IPF VS CTR D) LC vs CTR. The y-axis shows the pathways, and the x-axis shows the GeneRatio. Number of target genes are represented with Dot Size while padj-value is indicated with different colours.

Analysis conducted through the KEGG ontology (**Figure 54**) showed significant enrichment for the term Non-Small Cell Lung Cancer in IPF vs CTR. Specifically, among the target genes and miRNAs deregulated in patients with IPF we identified genes relevant to the pathophysiology of this type of tumour like ALK, EGFR, numerous kinases, genes related to apoptosis, etc.)

Comparing early and established IPF, 236 differentially expressed ci-miRNAs were enriched in processes tied to fibrosis (e.g., keratin filament, supramolecular complexes), and hierarchical clustering separated groups with early IPF closer to controls, suggesting progressive transcriptional shifts. This discriminatory capacity of ci-miRNAs as biomarkers highlight their potential to stratify patients belonging to different disease stages. Regarding the comparison LC vs CTR the MAPK signalling pathway is the most prominent finding because it is one of the most significant. This suggests the MAPK signalling pathway is the strongest and extensively affected biological process in LC condition. Other highly significant and important pathways include "cAMP signaling pathway," and "Ras signaling pathway."

V – DISCUSSION, CONCLUSIONS AND PERSPECTIVES

V-1 Discussion

Evaluating mutational status (KRAS, NRAS, BRAF, EGFR) in tumour biopsies has become a standard in clinical procedure and in most institutions, with the major implications of selecting the most appropriate therapeutic strategy. Indeed, mutations in the RAS and BRAF family genes are associated with poor response to anti-EGFR therapies, making mutational profiling essential to optimize treatment outcomes. In the cohorts reported in this thesis, molecular analysis of tissue samples confirmed a low frequency of KRAS, NRAS and EGFR mutations, in agreement with the prevalence reported in the literature for mCRC and LC (Afrăsânie et al., 2023; Parra-Medina et al., 2023).

Since mutations in driver oncogenes can affect drug response and resistance, evaluation of both the allelic configuration and the fraction of mutant alleles is clinically relevant. KRAS mutant tumours, for instance, are characterised by enhanced proliferation and sensitivity to MEK inhibitors compared to wild-type tumours. However, while the presence of driver mutations became clearly linked to the therapeutic response, the prognostic role of MAF or expression levels of mutated genes remains still to be fully elucidated, given that no consistent impact on survival or prediction of response has been demonstrated.

Notably, mutational profiles should not be assessed only at baseline but also during follow-up to anticipate treatment outcomes. Tissue biopsies, however, are often unfeasible in metastatic patients due to their poor clinical condition. For this reason, liquid biopsy has emerged as a minimally invasive and reliable alternative for both diagnostic and monitoring purposes in different tumours also comprising mCRC and LC. ctDNA represents a prognostic indicator since higher levels are generally associated with poor outcome, while decreases are observed after surgery but may rise before radiological progression (Chen et al., 2021). Compared to circulating tumour cells, ctDNA analysis offers greater sensitivity and requires only a small blood volume.

Discordant cases, particularly when tumour tissue is wild-type and plasma harbours mutations, may have relevant clinical consequences, as patients could inappropriately receive anti-EGFR therapy despite carrying RAS mutations. Interestingly, the analysis of plasma samples, allowed us to detect subclonal double mutations at very low frequency. Interestingly, in some patients with double mutations, disease progression was subsequently observed, reinforcing their potential clinical significance.

Thus, routine monitoring of the mutational status of a gene panel in plasma is gaining increasing relevance in clinical practice. Our findings confirm that ctDNA-based assays can detect emerging RAS mutations weeks before radiological evidence of progression, supporting the role of liquid biopsy as a valuable tool for early identification of resistance and for optimizing treatment strategies in mCRC and LC patients.

For personalized treatments based on a better understanding of disease biology, different tools must coexist to analyse the increasing amount of patient data that becomes available in order to overcome with the complexity of pathologies and molecular understandings (García-Casas et al., 2017).

In this study, extracellular vesicles (EVs) were quantified in plasma samples from mCRC and LC patients. We observed significant differences in concentration and size distribution of particles compared to healthy controls, suggesting that the EV release may reflect tumour burden and represent a useful biomarker. Further analyses of EV content like proteins, nucleic acids, and metabolites, could provide additional insights in their biological and clinical relevance.

V –2 Conclusions

Overall, our findings confirm that:

- Determine the mutational profile of a tumour is fundamental in the mCRC patients' management. The molecular characterization of genes like RAS, BRAF and EGFR in tissue and plasma at diagnosis, and only in plasma during follow up, provides essential information to guide therapeutic decisions.

-These results highlight the importance of integrating liquid biopsy into clinical practice. Unlike tissue biopsy, which is invasive and not suitable for repeated sampling necessary for patient monitoring, liquid biopsy provides a minimally invasive alternative that can capture tumour heterogeneity and detect subclonal populations. This is very relevant since patients initially classified as wild type may develop new mutations during therapy and ctDNA levels reflect tumour burden. Interestingly, molecular progression was detected before radiological progression, that would allow oncologists to anticipate relapse and adapt treatment strategies earlier.

-In this context, the molecular monitoring in real-time of tumour characteristics represents a promising strategy for precision oncology. It represents a guide for the choice of molecular targeted therapies, such as the use of anti-EGFR monoclonal antibodies in patients with RAS wild type tumours, an example of how liquid biopsy can help clinicians to maximize treatment efficacy, avoid unnecessary toxic effects and reduce costs associated with ineffective therapies. In this context Idylla™, with its ease of use and automation, represent a valuable tool for routine monitoring.

- Preliminary metabolomic analysis provided us evidence that there are signature metabolic fingerprints of the mCRC patients which can be detected by ^{1}H-NMR. Both blood plasma and urine profiles presented a patient-specific variability, although urine showed broader concentration ranges and greater inter-individual variation. Interestingly, metabolites like creatinine and lactic acid were identified as independent variables which can prompt the implementation of metabolomic signatures for patient stratification and disease progression monitoring. These preliminary results demonstrate the utility of metabolomics as an additive tool for molecular fingerprinting but need to be validated in larger cohorts.

- EV analyses in plasma samples from mCRC and mLC patients provided evidence of differences associated to the disease in vesicle profiles compared to healthy controls. Across both cohorts, EVs were consistently enriched in the 100–200 nm range, which correspond mainly to the exosome/microvesicle population. While median sizes between patients and healthy individuals appeared similar in general, patient-derived EVs tended to be slight

larger and more uniform, reflecting a possible alteration in vesicle biogenesis that may associated with the influence of the tumour.

-One of the most relevant emerging features from both datasets was the difference in EV concentrations. In mCRC patients, EV concentrations were more variable and reached higher values than in healthy subjects. In mLC patients, this effect was even more pronounced with significantly higher EV concentrations compared to controls, suggesting an overproduction of vesicles due to tumour presence. These findings support the hypothesis of tumour cells that actively release more vesicles into circulation, potentially reflecting tumour burden or altered metabolism and microenvironmental interactions.

-Although the concentration trends were not statistically significant, together with the observed inter-patient variability, the EV quantification highlighted their potential as a minimally invasive biomarker for cancer detection and monitoring.

-When comparing mutational profiles between tumour tissue and plasma of mLC patients using the Idylla™ system, overall agreement was detected. Notably, the concordance analysis was extended not only to the presence of mutations but also to their specific types.

-Longitudinal monitoring in plasma of *KRAS* and *EGFR* mutational status in lung cancer patients highlighted and confirmed the dynamic nature of ctDNA as a marker of therapeutic response. These findings underlined the utility of liquid biopsy not only for an initial molecular characterization but also as a tool for disease monitoring or for the early detection of resistance mechanisms.

-Finally, circulating microRNAs were highlighted as potential early and accessible biomarkers in Idiopathic Pulmonary Fibrosis. In this study we defined distinct and specific ci-miRNA profiles capable of discriminating healthy individuals from patients with different stages of the disease. Such profiles not only deepen our understanding of the pathophysiological mechanisms of IPF but also lay the groundwork for improving diagnostic accuracy. Last but not least, it confirmed that through the detection of ci-miRNAs it is possible to highlight the most altered pathways in patients affected by LC.

Collectively, these data suggest that plasma-based mutational testing is a valuable tool in the clinical management of LC and CRC and highlights the importance of validation in larger cohorts to identify correlations with disease outcomes in order to choose optimal treatment strategies for patients.

In conclusion, our data support the hypothesis that implementation of liquid biopsy into clinical routine may change the paradigm in treatment of mCRC and mLC patients providing early detection of resistance and achieve more precise therapeutic decision and better patient outcome.

V – 3 Future Directions

The studies on colorectal cancer is and lung cancer patients is still ongoing since they are part of a wider project, and a higher number of patients will be enrolled and analysed in all field mentioned in this thesis (Clinical Data, ctDNA, EVs, ci-miRNAs and NMR-Metabolomics).

Regarding the miRNAs the obtained results will be validated in a larger and independent patients' cohort, including other fibrotic conditions as well as tissue biopsies from patients who developed IPF. The transcriptomic analysis of circulating miRNAs will also be extended to a bigger cohort of patients affected by LC in order to further characterize miRNA signatures for diagnosis and prognosis.

The Lab will also focus on Protein Biomarkers in EVs, identifying novel therapeutic targets and to investigate whether there are differences at the vesicle level of potential biomarkers.

VIII – REFERENCES

Chial, H. (2008) Tumour suppressor (TS) genes and the two-hit hypothesis. Nature Education 1(1):177

Abdelwahab, S. I., Farasani, A., Alfaifi, H. A., & Hassan, W. (2024). Global landscape of liquid biopsy research: Regional contributions and disparities. *The Journal of Liquid Biopsy*, *6*, 100165. https://doi.org/10.1016/j.jlb.2024.100165

Afrăsânie, V. A., Marinca, M. V., Gafton, B., Alexa-Stratulat, T., Rusu, A., Froicu, E. M., Sur, D., Lungulescu, C. V., Popovici, L., Lefter, A. V., Afrăsânie, I., Ivanov, A. V., Miron, L., & Rusu, C. (2023). Clinical, Pathological and Molecular Insights on KRAS, NRAS, BRAF, PIK3CA and TP53 Mutations in Metastatic Colorectal Cancer Patients from Northeastern Romania. *International Journal of Molecular Sciences*, *24*(16). https://doi.org/10.3390/ijms241612679

Alsomali, H., Palmer, E., Aujayeb, A., & Funston, W. (2023). Early Diagnosis and Treatment of Idiopathic Pulmonary Fibrosis: A Narrative Review. In *Pulmonary Therapy* (Vol. 9, Issue 2, pp. 177–193). Adis. https://doi.org/10.1007/s41030-023-00216-0

Angulo, M., Lecuona, E., & Sznajder, J. I. (2012). Role of MicroRNAs in Lung Disease. In *Arch Bronconeumol* (Vol. 48, Issue 9).

Ardila-Molano, J., Vizcaíno, M., & Serrano, M. L. (2015). Circulating microRNAs as potential cancer biomarkers. In *Revista Colombiana de Cancerologia* (Vol. 19, Issue 4, pp. 229–238). Elsevier Doyma. https://doi.org/10.1016/j.rccan.2015.08.002

Bao, P., Wang, T., Liu, X., Xing, S., Ruan, H., Ma, H., Tao, Y., Zhan, Q., Belmonte-Reche, E., Qin, L., Han, Z., Mao, M., Li, M., & Lu, Z. J. (2025). Peak analysis of cell-free RNA finds recurrently protected narrow regions with clinical potential. *Genome Biology*, *26*(1). https://doi.org/10.1186/s13059-025-03590-x

Barratt, S. L., Creamer, A., Hayton, C., & Chaudhuri, N. (2018). Idiopathic pulmonary fibrosis (IPF): An overview. In *Journal of Clinical Medicine* (Vol. 7, Issue 8). MDPI. https://doi.org/10.3390/jcm7080201

Cadena-Suárez, A. R., Hernández-Hernández, H. A., Alvarado-Vásquez, N., Rangel-Escareño, C., Sommer, B., & Negrete-García, M. C. (2022). Role of MicroRNAs in Signaling Pathways Associated with the Pathogenesis of Idiopathic Pulmonary Fibrosis: A Focus on Epithelial-Mesenchymal Transition. In *International Journal of Molecular Sciences* (Vol. 23, Issue 12). MDPI. https://doi.org/10.3390/ijms23126613

Chen, G., Peng, J., Xiao, Q., Wu, H. X., Wu, X., Wang, F., Li, L., Ding, P., Zhao, Q., Li, Y., Wang, D., Shao, Y., Bao, H., Pan, Z., Ding, K. F., Cai, S., Wang, F., & Xu, R. H. (2021). Postoperative circulating tumor DNA as markers of recurrence risk in stages II to III colorectal cancer. *Journal of Hematology and Oncology*, *14*(1). https://doi.org/10.1186/s13045-021-01089-z

Chin-Yee, B., & Plutynski, A. (2024). Concepts of Actionability in Precision Oncology. *Philosophy of Science*, *91*(5), 1349–1360. https://doi.org/10.1017/psa.2023.115

Chudgar, N. P., Bucciarelli, P. R., Jeffries, E. M., Rizk, N. P., Park, B. J., Adusumilli, P. S., & Jones, D. R. (2015). Results of the National Lung Cancer Screening Trial: Where Are We Now? In *Thoracic Surgery Clinics* (Vol. 25, Issue 2, pp. 145–153). W.B. Saunders. https://doi.org/10.1016/j.thorsurg.2014.11.002

Clinical Use of Precision Oncology Decision Support. (2017). http://pct.

Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2009). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, *38*(6), 1767–1771. https://doi.org/10.1093/nar/gkp1137

De Sousa, K. P., Rossi, I., Abdullahi, M., Ramirez, M. I., Stratton, D., & Inal, J. M. (2023). Isolation and characterization of extracellular vesicles and future directions in diagnosis and therapy. In *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* (Vol. 15, Issue 1). John Wiley and Sons Inc. https://doi.org/10.1002/wnan.1835

Deb, D., Moore, A. C., & Roy, U. B. (2022). The 2021 Global Lung Cancer Therapy Landscape. *Journal of Thoracic Oncology*, *17*(7), 931–936. https://doi.org/10.1016/j.jtho.2022.03.018

Dekker, E., Tanis, P. J., Vleugels, J. L. A., Kasi, P. M., & Wallace, M. B. (n.d.). Colorectal cancer. In *www.thelancet.com*. www.thelancet.com

Deng, T., Tang, J., Zhou, L., & Duan, H. (2019). Effective targeted therapy based on dynamic monitoring of gene mutations in non-small cell lung cancer. *Translational Lung Cancer Research*, *8*(4), 532–538. https://doi.org/10.21037/tlcr.2019.08.06

Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., & Verweij, J. (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, *45*(2), 228–247. https://doi.org/10.1016/j.ejca.2008.10.026

Emmi, G., Bagni, G., Lastraioli, E., Di Patti, F., Bettiol, A., Fiorillo, C., Becatti, M., Silvestri, E., Urban, M. L., Emmi, L., Prisco, D., & Arcangeli, A. (2022). A unique circulating miRNA profile highlights thrombo-inflammation in Behcet's syndrome. *Annals of the Rheumatic Diseases*, *81*(3). https://doi.org/10.1136/annrheumdis-2021-220859

Escudero-Cernuda, S., Eiro, N., Fraile, M., Vizoso, F. J., Fernández-Colomer, B., & Fernández-Sánchez, M. L. (2025). Limitations and challenges in the characterization of extracellular vesicles from stem cells and serum. *Microchimica Acta*, *192*(5). https://doi.org/10.1007/s00604-025-07147-4

Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International Journal of Cancer*, *149*(4), 778–789. https://doi.org/10.1002/ijc.33588

Grabuschnig, S., Bronkhorst, A. J., Holdenrieder, S., Rodriguez, I. R., Schliep, K. P., Schwendenwein, D., Ungerer, V., & Sensen, C. W. (2020). Putative origins of cell-free DNA in humans: A review of active and passive nucleic acid release mechanisms. In *International Journal of Molecular Sciences* (Vol. 21, Issue 21, pp. 1–24). MDPI AG. https://doi.org/10.3390/ijms21218062

Gristina, V., La Mantia, M., Peri, M., Iacono, F., Barraco, N., Perez, A., Viscardi, G., Cutaia, S., Russo, T. D. B., Anwar, Z., Incorvaia, L., Fulfaro, F., Vieni, S., Pantuso, G., Graceffa, G., Russo, A., Galvano, A., & Bazan, V. (2023). Navigating the liquid biopsy Minimal Residual Disease (MRD) in non-small cell lung cancer: Making the invisible visible. In *Critical Reviews in Oncology/Hematology* (Vol. 182). Elsevier Ireland Ltd. https://doi.org/10.1016/j.critrevonc.2022.103899

Grodzka, A., Knopik-Skrocka, A., Kowalska, K., Kurzawa, P., Krzyżaniak, M., Stencel, K., & Bryl, M. (2023). MOLECULAR ALTERATIONS OF DRIVER GENES IN NON-SMALL CELL LUNG CANCER – FROM DIAGNOSTICS TO TARGETED THERAPY. In *EXCLI Journal* (Vol. 22, pp. 415–432). Leibniz Research Centre for Working Environment and Human Factors. https://doi.org/10.17179/excli2023-6122

Guinney, J., Dienstmann, R., Wang, X., De Reyniès, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., Bot, B. M., Morris, J. S., Simon, I. M., Gerster, S., Fessler, E., De Sousa E Melo, F., Missiaglia, E., Ramay, H., Barras, D., … Tejpar, S. (n.d.). *The Consensus Molecular Subtypes of Colorectal Cancer*. https://doi.org/10.7303/syn2623706

Hanahan, D., & Weinberg, R. A. (2000). The Hallmarks of Cancer Review evolve progressively from normalcy via a series of pre. In *Cell* (Vol. 100).

He, Y., Huang, W., Hong, H., Li, Y., Shen, Y., & Qu, Y. (2025). Clinical trial landscape of lung cancer treatment with ADCs: current perspectives and future directions. *International Journal of Surgery*, *111*(7), 4906–4909. https://doi.org/10.1097/js9.0000000000002501

Hossain, M. S., Karuniawati, H., Jairoun, A. A., Urbi, Z., Ooi, D. J., John, A., Lim, Y. C., Kaderi Kibria, K. M., Mohiuddin, A. K. M., Ming, L. C., Goh, K. W., & Hadi, M. A. (2022). Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies. In *Cancers* (Vol. 14, Issue 7). MDPI. https://doi.org/10.3390/cancers14071732

Jabalee, J., Towle, R., & Garnis, C. (2018). The role of extracellular vesicles in cancer: Cargo, function, and therapeutic implications. In *Cells* (Vol. 7, Issue 8). MDPI. https://doi.org/10.3390/cells7080093

Jin, N., Kan, C. M., Pei, X. M., Cheung, W. L., Ng, S. S. M., Wong, H. T., Cheng, H. Y. L., Leung, W. W., Wong, Y. N., Tsang, H. F., Chan, A. K. C., Wong, Y. K. E., Cho, W. C. S., Chan, J. K. C., Tai, W. C. S., Chan, T. F., Wong, S. C. C., Yim, A. K. Y., & Yu, A. C. S. (2023). Cell-free circulating tumor RNAs in plasma as the potential prognostic biomarkers in colorectal cancer. *Frontiers in Oncology*, *13*. https://doi.org/10.3389/fonc.2023.1134445

Jokela, T. A., Karppinen, J. E., Kärkkäinen, M., Mecklin, J.-P., Walker, S., Seppälä, T. T., & Laakkonen, E. K. (2024). Circulating metabolome landscape in Lynch syndrome. *Cancer & Metabolism*, *12*(1). https://doi.org/10.1186/s40170-024-00331-9

Kinoshita, T., & Goto, T. (2019). Molecular mechanisms of pulmonary fibrogenesis and its progression to lung cancer: A review. In *International Journal of Molecular Sciences* (Vol. 20, Issue 6). MDPI AG. https://doi.org/10.3390/ijms20061461

Kishiki, T., Ohnishi, H., Masaki, T., Ohtsuka, K., Ohkura, Y., Furuse, J., Watanabe, T., & Sugiyama, M. (2014). Overexpression of MET is a new predictive marker for anti-EGFR therapy in metastatic colorectal cancer with wild-type KRAS. *Cancer Chemotherapy and Pharmacology*, *73*(4), 749–757. https://doi.org/10.1007/s00280-014-2401-4

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3). https://doi.org/10.1186/gb-2009-10-3-r25

Larson, M. H., Pan, W., Kim, H. J., Mauntz, R. E., Stuart, S. M., Pimentel, M., Zhou, Y., Knudsgaard, P., Demas, V., Aravanis, A. M., & Jamshidi, A. (2021). A comprehensive characterization of the cell-

free transcriptome reveals tissue- and subtype-specific biomarkers for cancer detection. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22444-1

Lawrence, S. R., & Shah, K. M. (2024). Prospects and Current Challenges of Extracellular Vesicle-Based Biomarkers in Cancer. In *Biology* (Vol. 13, Issue 9). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/biology13090694

Li, J., Ma, X., Chakravarti, D., Shalapour, S., & Depinho, R. A. (2021). *Genetic and biological hallmarks of colorectal cancer*. https://doi.org/10.1101/gad.348226

Ma, L., Guo, H., Zhao, Y., Liu, Z., Wang, C., Bu, J., Sun, T., & Wei, J. (2024). Liquid biopsy in cancer current: status, challenges and future prospects. In *Signal Transduction and Targeted Therapy* (Vol. 9, Issue 1). Springer Nature. https://doi.org/10.1038/s41392-024-02021-w

Ma, R., Jiang, T., & Kang, X. (2012). *Circulating microRNAs in cancer: origin, function and application*. http://www.jeccr.com/content/31/1/38

Mármol, I., Sánchez-de-Diego, C., Dieste, A. P., Cerrada, E., & Yoldi, M. J. R. (2017). Colorectal carcinoma: A general overview and future perspectives in colorectal cancer. In *International Journal of Molecular Sciences* (Vol. 18, Issue 1). MDPI AG. https://doi.org/10.3390/ijms18010197

Mori, Y., & Kondoh, Y. (2021). What parameters can be used to identify early idiopathic pulmonary fibrosis? In *Respiratory Investigation* (Vol. 59, Issue 1, pp. 53–65). Elsevier B.V. https://doi.org/10.1016/j.resinv.2020.10.008

Mustafin, R. N. (2022). Molecular genetics of idiopathic pulmonary fibrosis. In *Vavilovskii Zhurnal Genetiki i Selektsii* (Vol. 26, Issue 3, pp. 308–318). Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences. https://doi.org/10.18699/VJGB-22-37

Newman, A. M., Bratman, S. V., To, J., Wynne, J. F., Eclov, N. C. W., Modlin, L. A., Liu, C. L., Neal, J. W., Wakelee, H. A., Merritt, R. E., Shrager, J. B., Loo, B. W., Alizadeh, A. A., & Diehn, M. (2014). An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nature Medicine*, *20*(5), 548–554. https://doi.org/10.1038/nm.3519

Nguyen, L. H., Goel, A., & Chung, D. C. (2020). Pathways of Colorectal Carcinogenesis. *Gastroenterology*, *158*(2), 291–302. https://doi.org/10.1053/j.gastro.2019.08.059

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. In *Frontiers in Endocrinology* (Vol. 9, Issue AUG). Frontiers Media S.A. https://doi.org/10.3389/fendo.2018.00402

Ortiz-Quintero, B., Buendía-Roldán, I., Ramírez-Salazar, E. G., Balderas-Martínez, Y. I., Ramírez-Rodríguez, S. L., Martínez-Espinosa, K., & Selman, M. (2020). Circulating microrna signature associated to interstitial lung abnormalities in respiratory asymptomatic subjects. *Cells*, *9*(6), 1–14. https://doi.org/10.3390/cells9061556

Pandey, S., & Yadav, P. (2025). Liquid biopsy in cancer management: Integrating diagnostics and clinical applications. In *Practical Laboratory Medicine* (Vol. 43). Elsevier B.V. https://doi.org/10.1016/j.plabm.2024.e00446

Parra-Medina, R., Pablo Castañeda-González, J., Montoya, L., Paula Gómez-Gómez, M., Clavijo Cabezas, D., & Plazas Vargas, M. (2023). Prevalence of oncogenic driver mutations in

Hispanics/Latin patients with lung cancer. A systematic review and meta-analysis. *Lung Cancer*, *185*. https://doi.org/10.1016/j.lungcan.2023.107378

Patel, S. R., & Das, M. (2023). Small Cell Lung Cancer: Emerging Targets and Strategies for Precision Therapy. In *Cancers* (Vol. 15, Issue 16). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/cancers15164016

Pino, M. S., & Chung, D. C. (2010). The Chromosomal Instability Pathway in Colon Cancer. *Gastroenterology*, *138*(6), 2059–2072. https://doi.org/10.1053/j.gastro.2009.12.065

Raghu, G., Collard, H. R., Egan, J. J., Martinez, F. J., Behr, J., Brown, K. K., Colby, T. V., Cordier, J. F., Flaherty, K. R., Lasky, J. A., Lynch, D. A., Ryu, J. H., Swigris, J. J., Wells, A. U., Ancochea, J., Bouros, D., Carvalho, C., Costabel, U., Ebina, M., … Schünemann, H. J. (2011). An Official ATS/ERS/JRS/ALAT Statement: Idiopathic pulmonary fibrosis: Evidence-based guidelines for diagnosis and management. *American Journal of Respiratory and Critical Care Medicine*, *183*(6), 788–824. https://doi.org/10.1164/rccm.2009-040GL

Ren, F., Fei, Q., Qiu, K., Zhang, Y., Zhang, H., & Sun, L. (2024a). Liquid biopsy techniques and lung cancer: diagnosis, monitoring and evaluation. In *Journal of Experimental and Clinical Cancer Research* (Vol. 43, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13046-024-03026-7

Ros, J., Baraibar, I., Sardo, E., Mulet, N., Salvà, F., Argilés, G., Martini, G., Ciardiello, D., Cuadra, J. L., Tabernero, J., & Élez, E. (2021). BRAF, MEK and EGFR inhibition as treatment strategies in BRAF V600E metastatic colorectal cancer. In *Therapeutic Advances in Medical Oncology* (Vol. 13). SAGE Publications Inc. https://doi.org/10.1177/1758835921992974

Sawicki, T., Ruszkowska, M., Danielewicz, A., Niedźwiedzka, E., Arłukowicz, T., & Przybyłowicz, K. E. (2021). A review of colorectal cancer in terms of epidemiology, risk factors, development, symptoms and diagnosis. In *Cancers* (Vol. 13, Issue 9). MDPI AG. https://doi.org/10.3390/cancers13092025

Shankar, A., Dubey, A., Saini, D., Singh, M., Prasad, C. P., Roy, S., Bharati, S. J., Rinki, M., Singh, N., Seth, T., Khanna, M., Sethi, N., Kumar, S., Sirohi, B., Mohan, A., Guleria, R., & Rath, G. K. (2019). Environmental and occupational determinants of lung cancer. In *Translational Lung Cancer Research* (Vol. 8, pp. S31–S49). AME Publishing Company. https://doi.org/10.21037/tlcr.2019.03.05

Shankaran, V., Obel, J., & Benson, A. B. (2010). Predicting Response to EGFR Inhibitors in Metastatic Colorectal Cancer: Current Practice and Future Directions. *The Oncologist*, *15*(2), 157–167. https://doi.org/10.1634/theoncologist.2009-0221

Smolarz, B., Łukasiewicz, H., Samulak, D., Piekarska, E., Kołaciński, R., & Romanowicz, H. (2025). Lung Cancer—Epidemiology, Pathogenesis, Treatment and Molecular Aspect (Review of Literature). In *International Journal of Molecular Sciences* (Vol. 26, Issue 5). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijms26052049

Stejskal, P., Goodarzi, H., Srovnal, J., Hajdúch, M., van ’t Veer, L. J., & Magbanua, M. J. M. (2023). Circulating tumor nucleic acids: biology, release mechanisms, and clinical relevance. In *Molecular Cancer* (Vol. 22, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12943-022-01710-w

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36

Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Testa, U., Pelosi, E., & Castelli, G. (2018). Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. In *Medical sciences (Basel, Switzerland)* (Vol. 6, Issue 2). NLM (Medline). https://doi.org/10.3390/medsci6020031

Tieng, F. Y. F., Abu, N., Lee, L. H., & Ab Mutalib, N. S. (2021). Microsatellite instability in colorectal cancer liquid biopsy—current updates on its potential in non-invasive detection, prognosis and as a predictive marker. In *Diagnostics* (Vol. 11, Issue 3). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/diagnostics11030544

Vilar, E., & Gruber, S. B. (2010). Microsatellite instability in colorectal cancerthe stable evidence. In *Nature Reviews Clinical Oncology* (Vol. 7, Issue 3, pp. 153–162). https://doi.org/10.1038/nrclinonc.2009.237

Wang, W., Rong, Z., Wang, G., Hou, Y., Yang, F., & Qiu, M. (2023). Cancer metabolites: promising biomarkers for cancer liquid biopsy. In *Biomarker Research* (Vol. 11, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s40364-023-00507-3

Wei, B., Cai, T., Zhang, R., Li, A., Huo, N., Li, S., Gu, Y. Q., Vogel, J., Jia, J., Qi, Y., & Mao, L. (2009). Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (Triticum aestivum L.) and Brachypodium distachyon (L.) Beauv. *Functional and Integrative Genomics*, *9*(4), 499–511. https://doi.org/10.1007/s10142-009-0128-9

Zhu, G., Pei, L., Xia, H., Tang, Q., & Bi, F. (2021). Role of oncogenic KRAS in the prognosis, diagnosis and treatment of colorectal cancer. In *Molecular Cancer* (Vol. 20, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12943-021-01441-4

VIII – ACKNOWLEDGEMENTS