UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Zentrum für Experimentelle Medizin, Institut für Tumorbiologie

Direktor der Einrichtung
Prof. Dr. med. Klaus Pantel

Identification of predictive and prognostic markers for oligometastatic disease in NSCLC patients with brain metastases

Dissertation

zur Erlangung des Doktorgrades Dr. rer. biol. hum. / PhD
an der Medizinischen Fakultät der Universität Hamburg

vorgelegt von:

Mais Alsousli
aus Damaskus, Syrien

Hamburg 2023

Angenommen von der
Medizinischen Fakultät der Universität Hamburg am: 26.06.2024

Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Harriet Wikman

Prüfungsausschuss, zweite/r Gutachter/in: Prof. Dr. Sabine Windhorst

Table of Contents

1. Introduction

1.1. Lung Cancer

Lung cancer is the leading cause of cancer-related deaths worldwide with an estimated 1.8 million deaths, accounting for 18% of total cancer deaths. In men, lung cancer represents the most frequently diagnosed cancer (14.3%) with also the highest mortality rates (21.5%). In women, it is the third most diagnosed cancer (8.4%) and the second in mortality (13.7%) (Figure 1) [1].

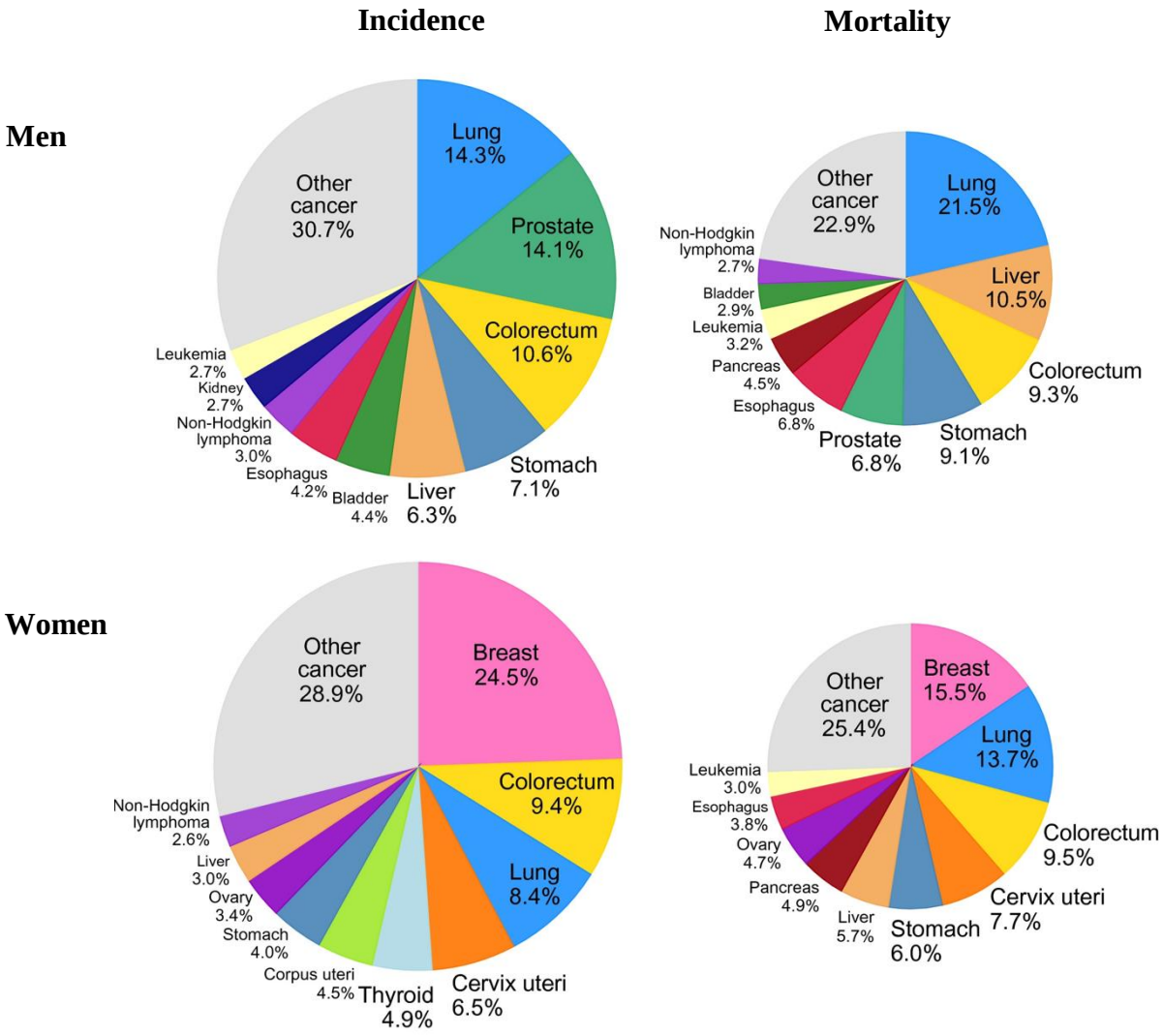


Figure 1: Distribution of cases and deaths for the top 10 most common cancers in 2020 for men and women. (Source: GLOBOCAN 2020 [1] with modification).

Despite improvement in patient outcomes in most cancer types in the last decades, the overall survival rate among lung cancer patients is still dismal (19%). This is mostly because lung cancer is often diagnosed at a later stage of the disease when it is disseminated to other organs

and is harder to treat [4]. Hence, the 5-year survival rate drops from 65% when detected at Stage I to only 5% for Stage IV [5].

While smoking remains the main risk factor for lung cancer, with approximately 80% mortality rate caused by tobacco consumption [6], up to 25% of lung cancer patients are passive smokers [7]. Other risk factors include exposure to radon and other carcinogenic chemicals or ionizing radiation and air pollution [8]. The resulting alterations can be inherited or obtained through DNA copying errors known as "somatic genomic alterations" which are necessary for oncogenesis [9]. The impact of these risk factors differs based on geographic location, gender, age, lifestyle, and race characteristics as well as their combined effects [10].

1.1.1. Non-small cell lung cancer (NSCLC)

Lung cancer is divided into two main types: non-small cell lung cancer (NSCLC) and small-cell lung cancer (SCLC). NSCLC represents approximately 80-85% of all cases and can be further classified based on its histology. The most common subtypes of NSCLC include adenocarcinoma (LUAD), squamous cell carcinoma (LUSC), and large cell carcinoma (LCLC). LUAD originates from alveolar/bronchial cells and represents 40% of NSCLC cases. LUSC arises from the bronchial epithelium of the larger, more central airways and is associated with a higher aggressiveness and a stronger correlation with smoking, constituting 25% of NSCLC cases. LCLC is diagnosed by excluding the other subtypes mentioned above and accounts for approximately 15% of NSCLC cases (Figure 2) [4, 11, 12].

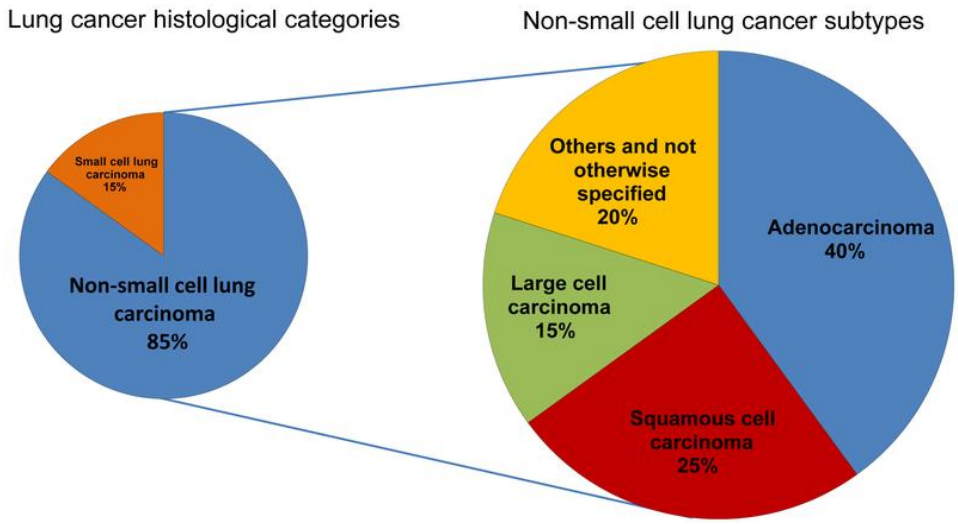


Figure 2: Histological classification of lung cancer. (Source: Schabath and Cote [4])

1.1.2. TNM staging and treatment strategies

The TNM system established by the Union for International Cancer Control (UICC) is used to stage different types of cancer including lung cancer, which helps guide treatment decisions

and predict prognosis [13]. This system is continuously reevaluated and the 8th edition of TNM staging was published in January 2017 [14].

TNM stands for "Tumor-Nodes-Metastasis", where (T) describes the size and extent of the primary tumor ranging from T1-T4, which can be further subdivided based on its dimension and location. The (N) describes the involvement of regional lymph nodes, classifying the degree of spread from N1-N3 based on location and metastatic nodes. Whereas (M) staging is defined by the absence (M0) or presence of distant metastases (M1) which is subdivided into M1a-c based on the extent and number of sites or organs involved. Moreover, the term Tis used for tumor *in situ* and T1mi refers to minimally invasive, while Tx and Nx are used to indicate when the primary tumor or lymph nodes cannot be accessed [13]. After determining the TNM descriptors, their combination can serve to define the overall stage grouping ranging from stage 0 (carcinoma *in situ*) to stage IV (metastatic disease) which is the most severe stage [15] (Table 1).

Table 1: Stages of lung cancer

Stage	Description
Stage 0	The earliest stage where cancer is confined to the innermost layer of the lung or bronchus.
Stage I	Divided into two sub-stages: 1A and 1B, based on the tumor size. Cancer is localized within the lung and has not extended to lymph nodes or distant sites.
Stage II	Divided into two sub-stages, IIA and IIB, each with further sub-types based on tumor's size, location and spread. Cancer remains within the lung but may involve nearby lymph nodes, or multiple tumors may appear in the same lung lobe.
Stage III	Divided into stages IIIA, IIIB, or IIIC, based on tumor size, location, and spread. Cancer has spread to lymph nodes, or multiple tumors are detected in different lobes of the same lung.
Stage IV	The most advanced stage, where cancer has metastasized to the opposite lung, distant organs, or tissues.

1.1.3. Treatment approaches

Due to the heterogeneous nature of NSCLC, treatment options vary depending on the histological type, stage of the tumor, patient's performance status (PS), and presence of certain molecular aberrations. These options usually include surgery, chemotherapy, radiotherapy, immunotherapy, and molecular targeted therapy, either alone or in a combined modality (Figure 3).

Patients diagnosed with early stages of NSCLC (stage I, II, and IIIA) typically undergo surgical resection as first-line treatment, depending on the patient's physical condition and whether the tumor is resectable. Additionally, adjuvant chemotherapy is recommended for stages II-IIIA to eliminate any remaining cancer cells and enhance the survival rate. Other adjuvant therapies might include radiation and targeted therapy (see Chapter 1.1.4.) [16, 17]. For non-surgical

patients with stage III disease, concurrent chemoradiotherapy followed by immunotherapy is a common treatment approach [18].

For advanced and metastatic disease, evaluating genetic alterations or protein expression in specific genes such as *EGFR*, *ALK*, *BRAF*, *KRAS*, *ROS1*, and *PD-L1* has significantly contributed to the effectiveness of additional treatment options, including targeted and immune checkpoints inhibitor (ICI) therapies [19, 20]. In some advanced-stage cases, ICI can also be considered as a first-line treatment [21]. For selected patients with stage IV oligometastatic disease localized in the brain or adrenals, surgery may also be a consideration [22].

The treatment landscape of NSCLC is rapidly evolving, with ongoing research and clinical trials aiming to uncover new insights into the disease and develop more effective therapies.

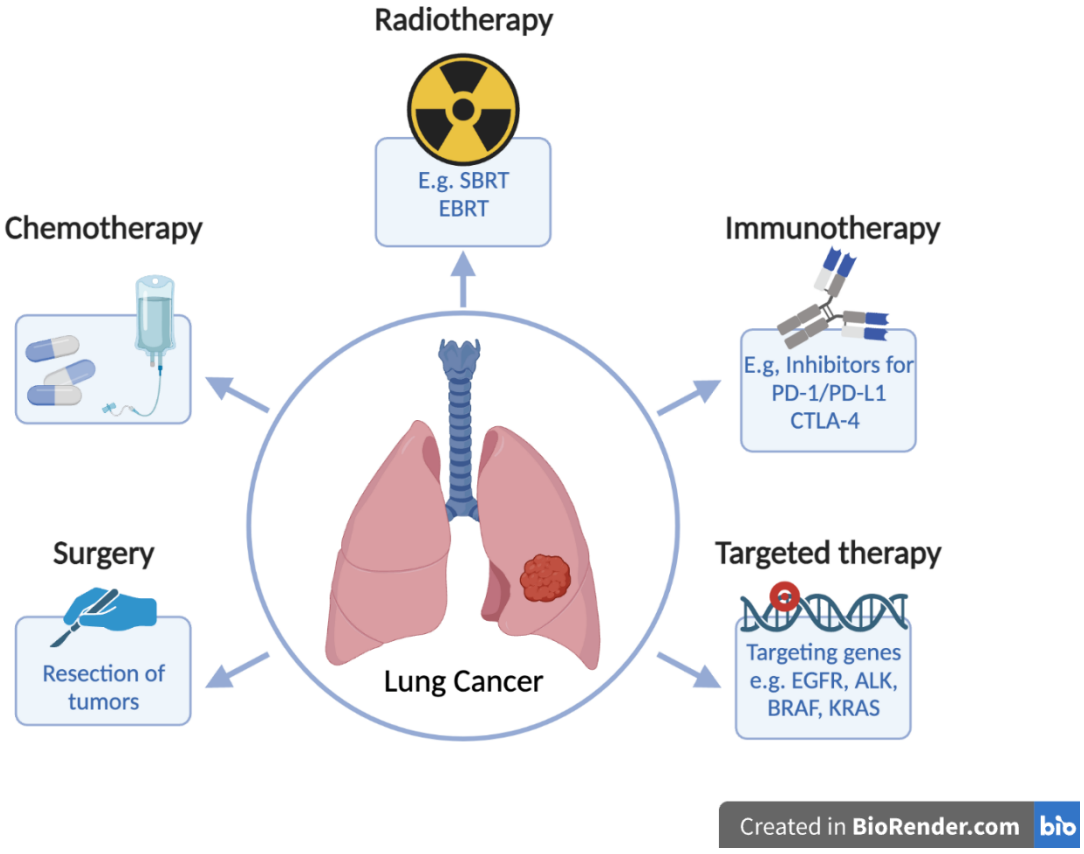


Figure 3: Lung cancer treatment options
SBRT: Stereotactic body radiation therapy, EBRT: External Beam Radiotherapy (Created with BioRender.com) (adapted from [3])

1.1.4. The genetic landscape of lung cancer

Lung cancer is a molecularly heterogeneous disease characterized by extensive genomic instability occurring at different levels, ranging from mutations in single or few nucleotides to gains or losses of entire chromosomes [23]. In the management of lung cancer, treatment decisions have been determined based on histology classification and genetic aberrations [24].

Importantly, lung cancer is known to have one of the highest mutational burdens, which is largely associated with smoking history [25]. This elevated mutational burden observed in tumors presents a challenge in clinical practice and limits the available treatment options [26]. However, recent clinical trials have suggested tumor mutational burden (TMB) could serve as a potential biomarker for immunotherapy. These trials have shown that patients with a high TMB (>10 mutations/Mb) treated with ICIs exhibited higher objective response rates and longer progression-free survival (PFS) compared to those with a TMB of <10 mutations/Mb [27]. Furthermore, somatic mutations in specific genes have shown to drive mutation progression. Consequently, an increase in somatic mutations would likely be accompanied by an increase in driver mutations. Once these genes mutate, they stay in all stages as cancer develops. As a result, the gene expression levels change in the bloodstream, affecting protein and metabolite levels [28].

NSCLC is characterized by high frequencies of certain genetic alterations or "hotspots", which may play important roles in tumor initiation, development, or metastasis [29]. Notably, there are distinct genetic differences among NSCLC subsets, but there are also shared genetic alterations between them. For example, the tumor suppressor gene *TP53* is the most frequently mutated gene in lung cancer. Inactivation mutations of *TP53* are observed in roughly 50% of LUAD and 80% of LUSC [30, 31]. In addition, NFE2L2/KEAP1 pathway which plays a critical role in responding to oxidative and toxic stresses, is frequently altered in both subtypes [32]. Mutations in *NFE2L2* or *KEAP1* occur in approximately 10% of LUAD and 17% of LUSC cases [33].

In LUAD, several targeted therapy options have emerged for patient treatment. The most prominent targetable mutations with oncogenic features in LUAD include *EGFR* and *KRAS* gene mutations, *ALK* and *ROS1* genes rearrangements (occurring in approximately in 1% of patients, and *BRAF* point mutations [25, 34-36]. These oncogenic drivers can control critical functions in cancer cells, such as tumor growth, cell proliferation and survival. Targeting these specific mutations has the potential to suppress tumor growth [37]. For example, several EGFR-specific tyrosine kinase inhibitors (TKIs) have been developed in recent decades targeting different EGFR mutations [25, 34]. Moreover, the most recently approved targeted drug in the EU was for tumors with *KRAS* G12C mutation [38].

Unfortunately, despite the initially promising outcomes, the development of drug resistance is almost inevitable following these treatments, and disease progression is frequently observed within a year [39-41]. Multiple mechanisms of acquired resistance have been identified. For

instance, in case of EGFR, the most common cause of resistance is the emergence of an additional mutation in *EGFR* known as *T790M* in more than 50% of patients [42-47].

On the other hand, the genomic landscape of LUSC exhibits distinct differences [30]. In contrast to LUAD, LUSC rarely harbors *EGFR* and *KRAS* mutations, as well as *ALK* rearrangement. Instead, *ERBB* genes, *FGFR1*, the tyrosine kinase *DDR2*, *SOX2* and JAK/STAT pathway are more frequently altered by mutations or amplifications [48, 49]. LUSC displays a higher genetic complexity, which has presented challenges for targeted therapy, and early-phase studies have shown negative outcomes [50]. Therefore, to date no approved target therapies exist for LUSC, leading most patients to be treated with immunotherapy. However, it's important to note that only a small proportion of patients (~30%) with NSCLC benefit from immunotherapy [51].

In summary, despite the clinical advancements and improved survival rates in NSCLC, less than 25% of patients benefit from targeted therapy, and drug resistance commonly emerges during treatment [52]. Therefore, the discovery of new therapeutic targets is of utmost importance in order to enhance treatment outcomes and overcome drug resistance.

1.2. Metastases in NSCLC

Metastasis is a hallmark of malignant tumors, causing around 70% of cancer-related deaths [53, 54]. Since lung cancer is mostly asymptomatic or lacks specific symptoms at earlier stages, it is mainly discovered after its spreading [55]. Research has indicated that >40% of NSCLC patients have metastasis at the time of diagnosis [56].

Metastatic cancer involves a heterogeneous collection of cells with different genetic and phenotypic characteristics, which drive progression, metastasis, and drug resistance [57]. Primary tumor dissemination is promoted by significant processes such as epithelial-mesenchymal transition (EMT), angiogenesis, immune evasion, tumor microenvironment (TME) remodeling, chromosomal instability, and somatic mutations [53, 58].

NSCLC exhibits preferences when metastasizing to distant organs. Previous studies provided approximate frequencies of metastasis, although these rates display variation among studies and are not absolute, as the reported frequencies may be influenced by factors such as sample size, patient selection, and study design. However, bone metastasis for instance is prevalent in advanced NSCLC patients at initial diagnosis, followed by brain or lung, liver, and adrenal glands [59-61]. Furthermore, the number of affected organs also influences the metastatic pattern. Brain and bone metastases were more commonly observed in metastasis to a single

organ, while in the case of three or more organs, liver metastasis was more frequently involved [62].

1.2.1. Seed and soil theory

The concept of organ-specific metastasis was first proposed by Stephan Paget with the so-called "seed and soil" theory [63]. He suggested that successful metastasis requires two factors: tumor cells with a metastatic ability (seed) and a suitable host organ (soil). Hence, primary tumors selectively modify a permissive microenvironment of a target organ to promote metastasis by establishing a pre-metastatic niche (PMN) [64, 65]. For instance, the bone marrow has been implicated as a favorable metastatic niche, harboring often disseminated tumor cells that can later disseminate further to other organs [66].

An alternative hypothesis was introduced by James Ewing in 1928, suggesting that the site of metastasis is determined by the circulatory pattern between the primary tumor and the target secondary organ [67]. Although the so-called tumor draining blood can explain much of preferential metastasis sites, still different types of tumors have different metastatic patterns that cannot be explained by blood flow alone. Duda et al., presented a newer theory, proposing that metastatic cells can bring their own "soil" (stromal component including activated fibroblasts) from the primary site to the secondary site [68]. Consequently, ongoing research attempts to understand the molecular and genetic features of primary and secondary lesions, as well as the tumor microenvironment and pathways that influence cancer dissemination. Obviously, both tumor-intrinsic factors, as well as host factors, play important roles in this process. Moreover, specific mutations have been found to be associated with distinct patterns of metastatic spread in NSCLC [69]. For example, *EGFR* mutations were linked to a higher incidence of developing brain metastasis [70, 71].

1.2.2. Brain metastasis in NSCLC

Brain metastases (BrM) are the most frequently accruing intracranial tumors in adults which are associated with an extremely unfavorable outcome [72]. They are often accompanied by disabling neurological complications that significantly impact the patient's quality of life [73]. Symptoms, such as headaches, altered mental status, focal motor or sensory deficits, and ataxia, can manifest, with variations depending on factors like size, location, and the degree of edema linked with BrM [74]. The median survival rate of untreated patients with BrM ranges from 1-6 months [75, 76]. This rate can be significantly extended depending on various factors, including the disease stage and therapy effectiveness. Nevertheless, despite the advanced

treatment strategies of BrM, the prognosis remains poor, especially in the advanced stage of NSCLC [77, 78].

A key event in the development of BrM is the migration of tumor cells passing through the Blood-brain barrier (BBB), which ultimately leads to its partial damage [79]. The BBB is a unique feature that acts as a selective barrier separating systemic circulation and the brain. It regulates the flow of ions and nutrients, provides protection from infections, harmful substances, and cancer cells [2], and plays a crucial role in the immune response [80]. As a result of metastasis, the BBB might change its function from inhibiting cancer to becoming a barrier against therapy, thereby limiting the chemotherapeutic drug delivery [81].

The vast majority of BrM cases are observed in lung cancer, with approximately 40% of occurrences reported in NSCLC patients throughout the disease course [76, 82, 83]. The presence of BrM significantly worsens the prognosis of patients [84]. Among NSCLC subtypes, adenocarcinoma (LUAD) exhibits a notably higher propensity for BrM compared to other histological subtypes [85].

The estimated incidence rates of BrM between 1973 and 2001 ranged from 8.3 to 11 per 100,000 persons [82]. However, in 2021, Habbous et al. reported a higher rate of 24.2 per 100,000 persons per year [86]. The observed increase in BrM incidence may be attributed to the longer lifespan of cancer patients resulting from improved treatments of extracranial metastases and the growing number of diagnosed lung cancer cases [84, 87]. On the other hand, the rising frequency of diagnosed BrM may not necessarily be caused by an increase in the number of cases, but rather due to improved imaging techniques that facilitate their detection [88].

1.2.2.1. Oligometastasis and polymetastasis in NSCLC

NSCLC patients with BrM exhibit a significant heterogeneity [89]. Treatment is chosen based on the number of BrM and other extracranial metastases, as the extent of the disease affects survival rates. In general, the term “oligometastasis” refers to patients with a limited number of metastases involving a single or few organs [90, 91]. However, defining an oligometastatic state is still controversial, concerning the maximum number of detected metastatic lesions, their maximum size, and their distribution among other organs [92]. On the other hand, “polymetastasis” refers to patients with widespread metastatic dissemination to multiple sites [93]. Approximately 30 – 45% of NSCLC patients develop BrM as the only affected organ [94-96]. Within oligometastasis, two different cohorts have been identified: patients who present with oligometastatic disease at initial diagnosis of the primary tumor (“synchronous”

oligometastasis), and those who induced oligometastasis after treatment of primary tumor ("metachronous" oligometastasis) [97]. In NSCLC, 5-10% of patients present with synchronous oligo brain metastasis at initial diagnosis, and 20-40% of patients develop BrM oligometastasis during the course of the disease (metachronous) [78, 98].

Patients with oligometastasis generally have a longer lifespan than those with polymetastasis [90]. Previous data have shown that patients with oligometastasis who received radical treatment have better survival rates [89, 99], while patients with polymetastasis are often limited to palliative systemic therapy [100]. Furthermore, the differences between oligo-synchronous and oligo-metachronous have not been clearly defined overall in terms of biology and clinical outcome. While certain studies have reported a longer survival for selective metachronous BrM patients compared to synchronous cases [75, 101], these results have not been consistently observed [102-105].

Two hypotheses have been proposed to explain the clinical significance of oligometastatic disease. The first indicates that oligometastasis and polymetastasis are two distinct phenotypes, with different molecular characteristics and metastatic potentials [106]. The second hypothesis suggests that oligometastasis represent an intermediate state between the primary tumor and the poly-metastatic phenotype on a spectrum of disease progression (Figure 4). This difference between these classifications can be pivotal in determining predictive and prognostic markers of metastatic disease and establishing new therapy targets [90, 91, 106, 107].

Molecular and genomic investigations comparing different oligometastasis and polymetastasis diseases remain even more limited. Some studies have indicated molecular differences, such as miRNA expression, among these metastatic types [99, 106-108]. However, additional data is required to enhance our understanding of these metastatic patterns.

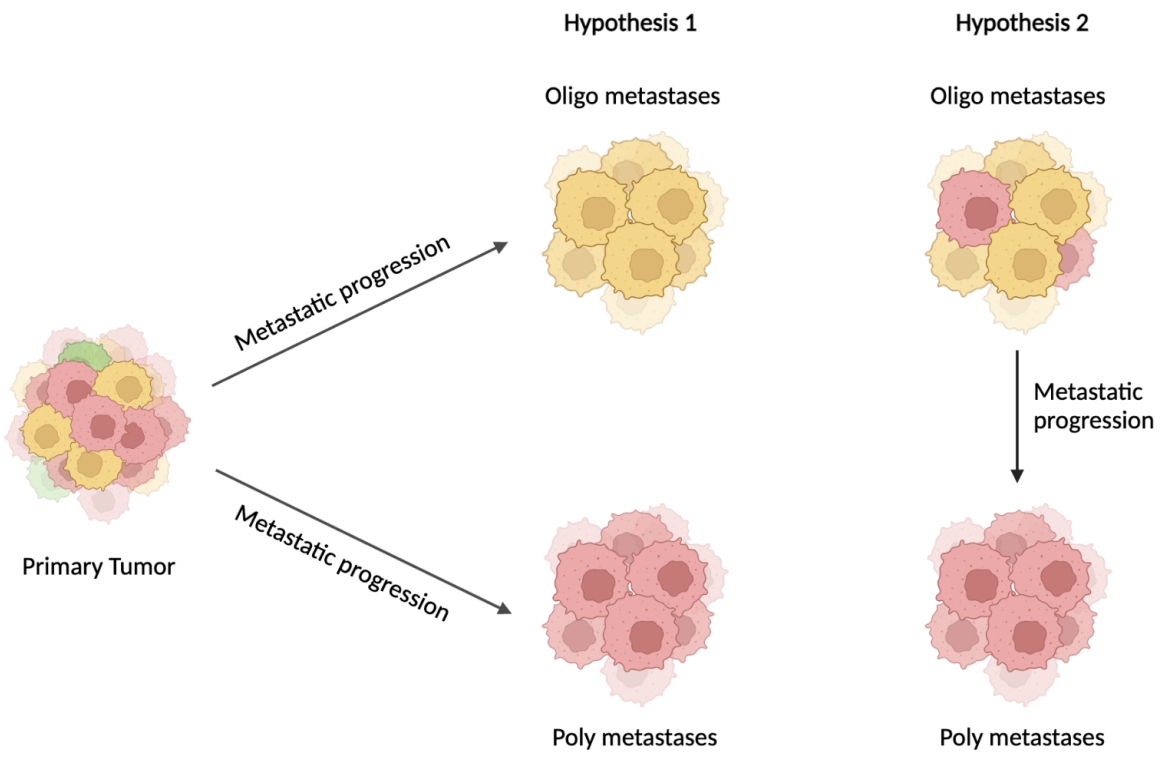


Figure 4: The two hypotheses of Oligometastatic Disease: Hypothesis 1 Oligometastasis and Polymetastasis may originate from clones with different molecular and phenotypic features. Hypothesis 2 oligometastasis may be an early phase of polymetastatic disease (Created with BioRender.com)

1.3. Analysis of host- and tumor-related factors in BrM

The biology behind BrM is highly complex and is influenced by various host and tumor-related factors that affect tumor dormancy, development, and progression. Although previous studies have reported significant heterogeneity in the immune microenvironment of BrM, it is still unclear if specific genetic profiles are associated with distinct immune states, and their correlation with metastatic patterns requires further investigation [108, 109]. Therefore, characterizing both the genomic and immune profiles can provide more comprehensive insights into this disease and its diverse patterns of development in lung cancer.

1.3.1. Genomic profile of lung cancer with BrM

The development of BrM from lung cancer is a complex multi-step process. Genomic characterization of BrM has been instrumental in identifying potential driver mutations and therapeutic targets [110, 111]. In lung cancer, certain driver mutations, such as *EGFR* and *ALK*, have been shown to be homogeneous among primary tumors and BrM. Patients with these mutations tend to have a better survival rate compared to those without these alterations [112]. Additionally, it has been observed that some genes exhibit a higher frequency of mutations in BrM compared to primary tumors in LUAD such as *TP53*, *ATR*, and *APC* [113]. However,

significant heterogenicity and genomic complexity have been reported between BrM and matched primary tumors in terms of somatic mutations and copy number events [114-116]. This heterogeneity may arise due to the independent genetic events occurring in both the lung and the brain [117], indicating novel and unique alterations in the metastatic site that may facilitate the metastatic process.

While numerous mechanisms and genes/proteins were suggested to be involved in the process, it is still unknown if there are specific additional genetic alterations that contribute to the development of BrM. Therefore, there are ongoing efforts to discover new possible biomarkers. Some studies by Stella et al. have indicated that proto-oncogene *MET* activation may play a substantial role in BrM from lung cancer [118, 119]. Another sequencing study on BrM-LUAD identified potential metastatic drivers in three regions with higher amplification frequencies, including *MYC*, *YAP1*, and *MMP13* [110]. Furthermore. PI3K and WNT/TCF signaling pathways have been linked to an increased risk of developing BrM [120, 121].

Many studies have investigated the mechanisms of BrM-NSCLC by comparing the tumor from the BrM site with the primary tumor of the same patient [116, 117, 122]. This approach can be useful in identifying important genes and pathways involved in the metastasizing process. However, it does not help determine patients with a high risk of developing a distinct metastatic pattern. Therefore, to enhance the management of BrM-NSCLC, it is important to find novel prognostic biomarkers that can predict an oligo-brain metastatic disease. This would enable the identification of high-risk patients, who would benefit from a regular brain imaging schedule to detect the metastatic disease earlier. Such a possibility holds tremendous importance in delineating the patients deemed suitable for preventive or radical treatments, especially for individuals who are not eligible for targeted therapy.

1.3.2. The immune system

The immune system is a complex network of cells, tissues, and organs, comprising two distinct compartments: the innate and the adaptive immune systems. They work together to defend the body against infections, pathogens, toxins, and cancer cells. The innate system is the body's first line of internal defense, carrying out nonspecific immediate immune responses. It consists of dendritic cells (DC), macrophages, neutrophils, basophils, eosinophils, and natural killers (NK). The adaptive system is the second line of defense, responsible for antigen-specific long-term responses, it includes B and T lymphocytes as well as NK-T cells, which are a population of cells that share properties from both T-cells and NK cells [123, 124].

The initial adaptive immune responses are typically slower than the innate immune responses. This is primarily because lymphocytes require time to proliferate and differentiate into effector cells capable of fighting the invading antigen during the first encounter. Once the antigen has been successfully eliminated, a small population of lymphocytes differentiates into long-lived memory cells. These memory cells can respond specifically and more rapidly upon subsequent exposure to the same antigen [125, 126].

1.3.2.1. Role of the immune system in NSCLC cancer

The main role of the immune system in cancer is to inhibit tumor progression and maintain cellular homeostasis in a process called cancer immune surveillance, where the tumors produce antigens that in turn evoke an immune response. However, avoiding immune destruction is one of the cancer hallmarks [127]. In fact, the immune system has a dual role in cancer, known as cancer immunoediting, where it can either restrain or promote tumor progression. This process progresses from immune surveillance to immune escape and involves three main phases: elimination, equilibrium, and escape [128, 129].

In the elimination phase, both innate and adaptive immune systems collaborate to identify the transformed cells and protect the host from cancer cells before they develop clinically detectable tumor lesions, yet some tumors manage to survive this phase and progress into the equilibrium phase, where the adaptive immune system is not capable of clearing the tumor cells completely, but still limits the tumor invasion and outgrowth [130, 131]. This incomplete elimination process is caused by the balance between anti-tumor and pro-tumor cytokines, leading to reduced cellular immunogenicity in the tumor, forcing the tumor itself to stay in this intermediate phase between elimination and escape [132]. However, if the immune system can no longer control tumor outgrowth, resistant cancer cells can escape and continue to grow uncontrollably into a visible disease with an immunosuppressive microenvironment [133].

1.3.2.2. Interaction of the host immune system and BrM

The brain is a special organ when it comes to interaction with immune cells. For a long time, the brain was considered “immune-privileged” due to the presence of BBB, which could block the immune cells coming from the blood circulation system. However, with later investigations of the lymphatic system in the brain and the discovery of lymphatic ducts in the meninges, this concept was revised [134]. In general, the more accurate description of the central nervous system (CNS) including the brain is rather “immune distinct” as approximately 80% of its immune cells are microglia [135-137]. These microglia with other immune cells from the

circulatory system keep the CNS under continual immune surveillance in a healthy brain, which primarily occurs within the meningeal compartment [136, 138]. During brain metastasis, tumor cells engage in continuous interaction with resident cells, such as microglia, to exploit their function and evade antitumor responses [139]. In such scenarios, microglia can undergo a transformation towards a pro-invasive and immunosuppressive phenotype to counterattack the anti-tumor immunity and develop resistance to treatments, thereby aiding tumor progression across various metastatic stages [140]. Recent experimental investigations have illuminated the presence of tumor-infiltrating macrophages, T lymphocytes, and other immune cell populations within BrM in addition to the brain residential glial cells [141].

During the metastatic cascade, cancer cells interact closely with the immune system in a complex and dynamic manner, exerting mutual influence, both in the tumor microenvironment (TME) and throughout the systemic circulation. This interaction between tumor and immune cells plays an important role in BrM initiation and immune escape. However, it also presents new therapeutic approaches for cancer patients [142, 143]. Furthermore, cancer development and its response to treatment are also influenced by the peripheral immune system of the host through circulating blood and lymphatic vessels.

The TME is a dynamic network that contains cancer cells and the surrounding environment, including cellular and non-cellular components, such as extracellular matrix, signaling molecules, blood vessels, fibroblasts, and immune cells [144]. Notably, the TME in the brain is unique and different from elsewhere, due to the presence of specialized resident immune cells such as microglia and astrocytes [145, 146]. Microglia, which belong to the monocyte-macrophage system, are highly dynamic innate immune cells. Apart from their roles in antigen presentation and cytokine production, microglia perform various other functions [147]. Whereas astrocytes, another type of glial cells in the CNS, become activated in response to stimuli such as injuries or tumors [148].

Within brain tumors, most of the immune cells are macrophages and microglia [149], followed by astrocytes and in smaller proportions, dendritic cells, neutrophils, and lymphoid cells including T-cells, B-cells and NK-cells [2, 145] (Figure 5). These immune cells can promote tumor progression by exerting a pro-invasive and immunosuppressive effect [150]. Moreover, tumor cells affect their microenvironment by releasing cellular signaling molecules that facilitate angiogenesis and induce immune tolerance [151].

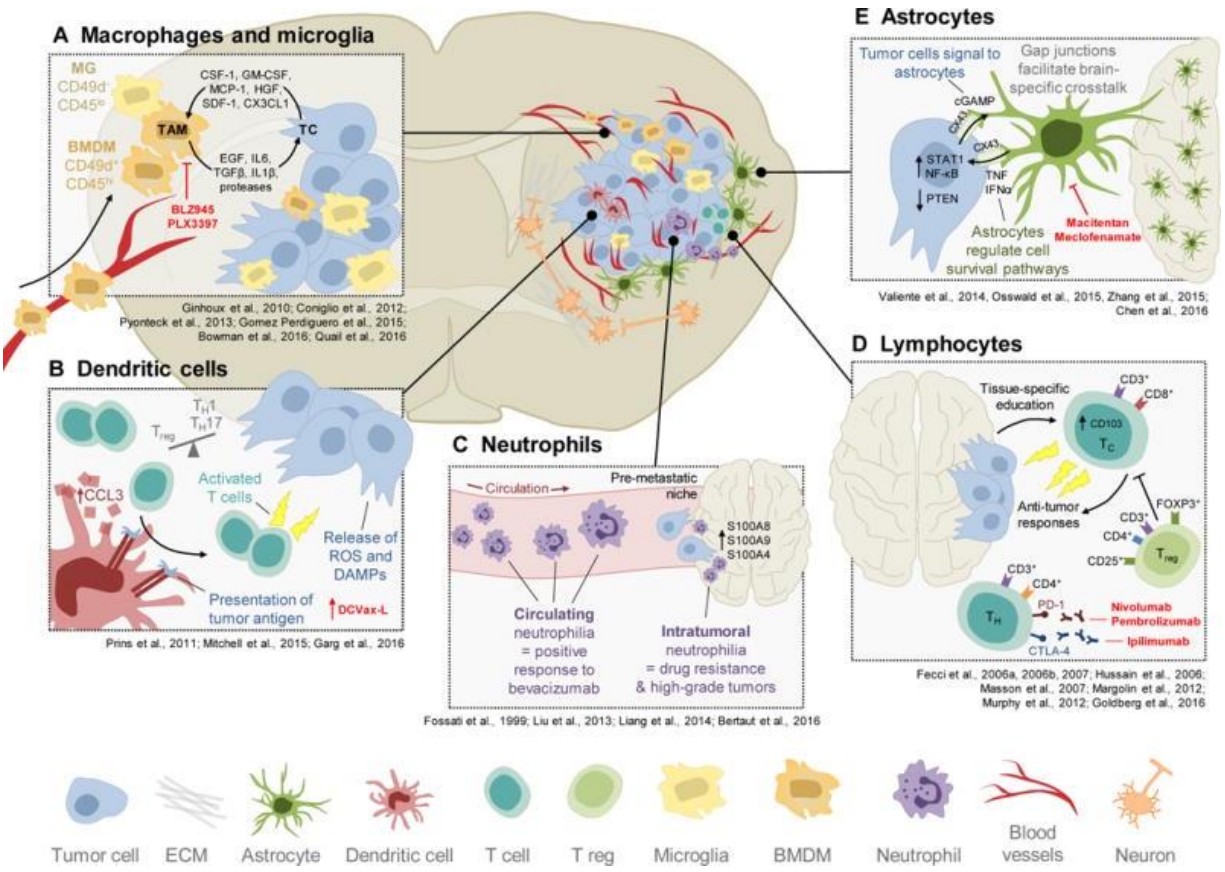


Figure 5: cellular components of tumor microenvironment (TME)
Brain tumors have a unique microenvironment that includes (A) macrophages and microglia, (B) Dendritic cells, (c) neutrophils, (d) lymphocytes and (E) astrocytes (source: Quail and Joyce [2])

The systemic immune environment consists of immune cells such as myeloid cell lineages (monocytes, neutrophils, basophils, and eosinophils), lymphoid cell lineages (T-cells, B-cells, and NK-cells), as well as immune modulators including cytokines and metabolites. It is essential for communication between the primary tumor site and distant organs, as well as the host immune organs [152, 153].

During metastasis formation, tumor cells face the challenge of surviving within the circulation system and interacting spatiotemporally with the various immune compartments [143]. By extravasation and infiltration, tumor cells can pass through BBB and help infiltrated immune cells from the systemic environment to influx to TME, promoting metastasis to the brain [154, 155] (Figure 6).

This constant interaction and coordination between cancer cells, TME, and systemic immune environment, are involved in shaping the host's response. In BrM, the heterogeneity and immunosuppressive nature of TME and systemic tumor immune environment, lower the overall efficacy of therapies [153]. Thus, more promising biomarkers for BrM still need to be identified [117].

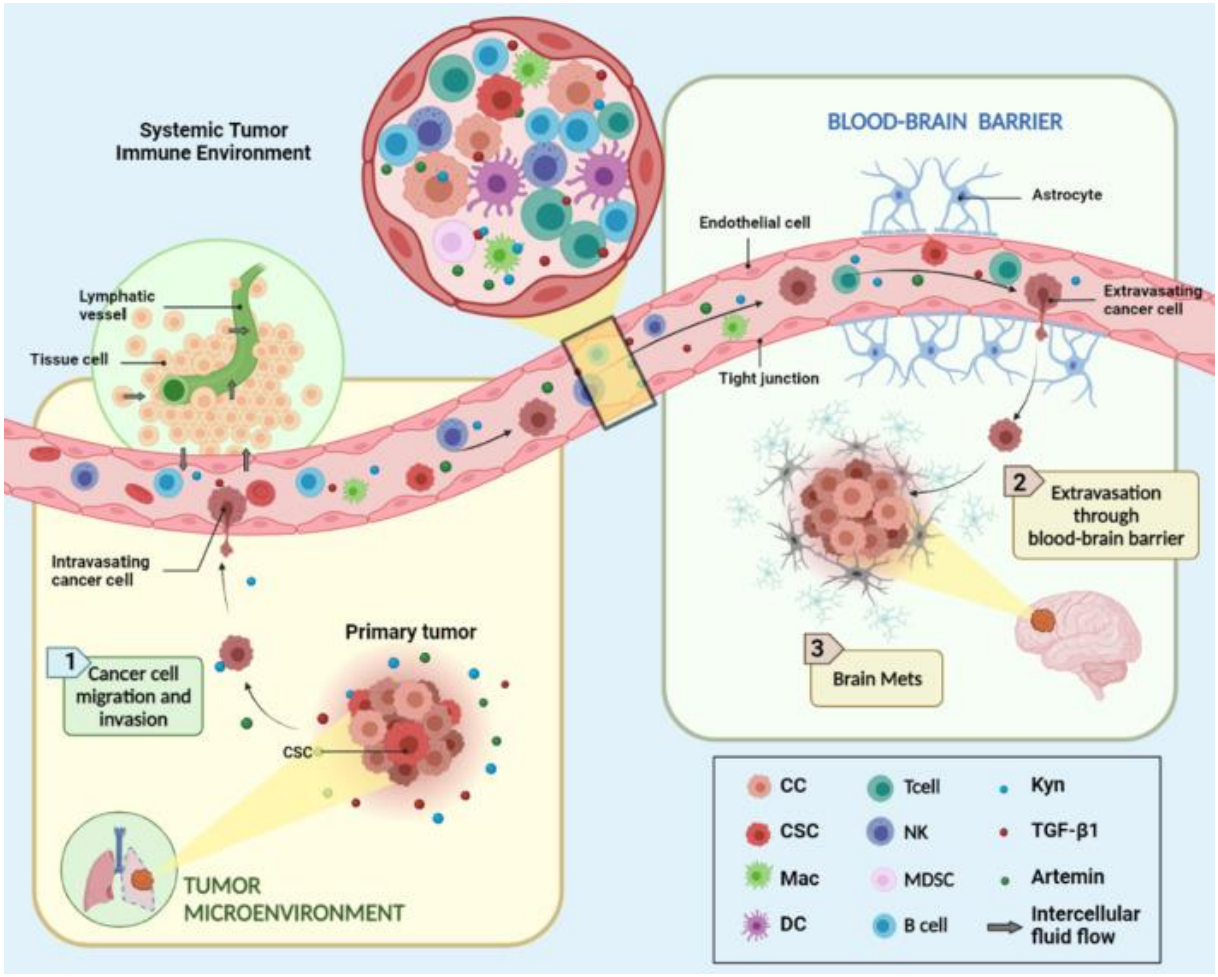


Figure 6: TME and systemic tumor immune environment in brain metastasis.
Tumor cells from the primary site infiltrate the circulation system and immune cells from the systemic tumor immune environment help them to pass through BBB to invade the brain (source: Xu et al., 2022 [153])

1.3.2.3. Tumor-infiltrating lymphocytes (TILs) and NK cells

As mentioned earlier, the TME comprises the innate immune system in BrM, primarily microglia and macrophages, and the adaptive immune system represented by tumor-infiltrating lymphocytes (TILs) which are the main component of TME involved in the immune response and positively correlated with prognosis [145, 156]. TILs mostly consist of T and NK cells, and they can either promote or inhibit tumor cell proliferation. In fact, T-cells are considered to be the center of immunology. They are functionally subdivided into two main subtypes, which are identified by their surface protein expression: $CD8^+$ cytotoxic T-cells which play a crucial role in tumor cell-killing through secreting cytokines and cytotoxic enzymes [157], and $CD4^+$ T cells which can play both promoting and suppressing roles in tumors. Some $CD4^+$ helper T-cells can assist with $CD8^+$ T-cell activation and mediate immune responses to eliminate cancer cells, while other $CD4^+$ T-cells such as regulatory T-cells (Treg) have an immunosuppressive function and hinder $CD8^+$ T-cell function, thus indirectly promoting tumor growth. Depending

on their function and secreted cytokines, $CD4^+$ helper T-cells can be classified into Th1, Th2, Th17, and T follicular helper cells (Tfh) [158].

T-lymphocytes can be further characterized based on their differentiation stage/activation state into naïve, effectors, and memory T-cells. Upon antigen stimulation, naïve T-cells which are the resting form get activated and differentiate into specific effector T-cells. They migrate to the antigen site to attack and kill the infected or transformed cells. After antigen elimination, most effector cells undergo apoptosis, and only a small portion turns into memory cells which can respond directly when exposed to the same antigen again [159]. Previous studies have shown a correlation between the infiltration rate of TILs in TME and patient survival in certain cancers, indicating that TIL infiltration may serve as a promising biomarker. In NSCLC, $CD8^+$ T-cells are associated with good prognosis [143], while the effect of $CD4^+$ T-cells on prognosis showed conflicting results in different studies [160]. This could be explained due to the diverse functions of different $CD4^+$ T-cell populations. It can also be influenced by the type of cancer, the components of the TME, and the disease stage [161].

NK cells belong to the innate immune system since they can function without the need for antigen presentation. They play an important role in antitumor immune response, by directly killing tumor cells or by enhancing the antitumorigenic activities of other immune cells. NK cells have the ability to control tumor growth and the early stages of metastatic dissemination [160-162].

In this thesis, we focused on T-cells and NK cells immunophenotyping using flow cytometry to identify and distinguish different cell populations.

1.4. High throughput tools to analyze host and tumor-related factors

High throughput tools have transformed the field of cancer research by enabling comprehensive analyses of host and tumor-related factors on a large scale. These advanced technologies have facilitated the exploration of interactions between the tumor microenvironment and the host immune system, shedding light on the complex mechanisms underlying tumor biology, tumorigenesis, and tumor progression. In this study, we employed a range of high-throughput tools to gain valuable insights into the molecular landscape of our cohort. The tools utilized included next-generation sequencing (NGS) for comprehensive profiling of genetic alterations, RNA sequencing for transcriptome analysis, and flow cytometry for analyzing certain immune cell populations in peripheral blood samples.

Additionally, we utilized bioinformatics pipelines and advanced data analysis methods to process the vast amounts of data generated by these tools. Together, these high throughput approaches helped enhance our understanding of brain metastatic NSCLC.

1.4.1. Immuno-profiling of peripheral blood

The process of immunophenotyping involves identifying and quantifying immune cell populations using specific antibodies as markers to bind to antigens expressed by these cells. Immunophenotyping assay may include the simultaneous use of multiple antibodies to identify different cell types by detecting multiple antigens simultaneously [163].

In addition, analyzing peripheral blood offers a less invasive and more accessible approach to monitor the immune profile of the patient, including immune mediators such as cytokines, and immune cells that may infiltrate the TME.

Flow cytometry is a powerful technique used to identify single cells in a heterogeneous population based on multiple parameters such as size, granularity, and protein markers. Flow cytometry-based immunophenotyping is commonly used in clinical research to study the immune system, including immune cell subsets and protein expression. It is also used in diagnostics to monitor various diseases. This method uses a laser as a light source to illuminate the cells passing through the laser beam in a fluidic stream, producing scattered and fluorescent light signals that are captured by optical detectors. These signals are converted and analyzed by computer software to generate numeric data about different cell populations based on their fluorescent or light-scattering characteristics. Fluorescently labeled antibodies or dyes are used to detect different cell types [163], and it should be noted that multiplexing is easily achieved here compared to tissue analysis.

1.4.2. Next-generation sequencing of DNA and RNA

Next-generation sequencing (NGS) is a high-throughput sequencing method used for the analysis of DNA and RNA. A variety of applications exist, including whole-genome sequencing (WGS), whole-exome sequencing (WES), transcriptome sequencing, sequencing of non-coding RNA such as miRNA and lncRNA, sequencing of targeted genes, epigenomics sequencing, and others [164]. Consequently, NGS can provide specific information on point mutations, small indels, copy number alterations (CNAs), structural variations as well as differential gene expressions. The advantages of NGS make it the optimal approach for assessing a multitude of biomarkers simultaneously in a relatively short time with high sensitivity and specificity, using low amounts of samples today at a rather cost-effective price [165].

There are multiple NGS platforms that use different approaches. In this thesis, the Illumina sequencing platform was used to produce data from DNA and RNA samples, which is the most popular technology in sequencing by synthesis (SBS) method [166]. Prior to sequencing, Illumina NGS starts with library preparation by random fragmentation of DNA or cDNA samples, followed by ligating adapters to both DNA fragment ends. Subsequently, the library is loaded into a flow cell that displays oligonucleotides complementary to adapter sequences, and each fragment is amplified into distinct clonal clusters through bridge amplification PCR. Sequencing reagents with fluorescently labeled nucleotides (dNTPs) are added in the presence of DNA polymerase, and a laser is used to excite the label dyes incorporated into the DNA fragments, detecting the location on the flow cell by a camera. The resulting data can then be analyzed and assembled to determine the sequence of the original DNA or RNA sample (Figure 7) [167].

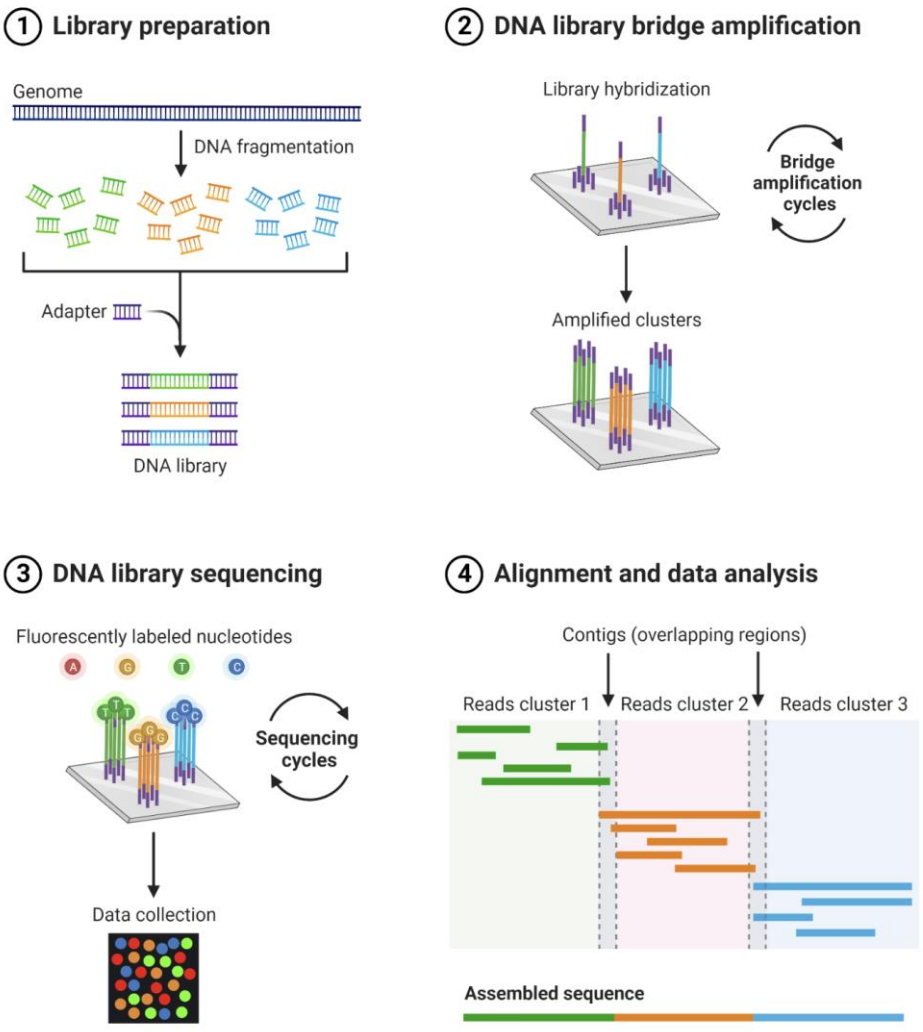


Figure 7: Overview of Illumina NGS workflow
It includes four steps: (1) library preparation, (2) bridge amplification, (3) library sequencing and (4) alignment and data analysis (source: BioRender.com)

1.4.3. NGS Data analysis and their challenges

NGS techniques have revolutionized cancer research, enabling comprehensive molecular profiling and paving the way for enhanced clinical approaches. The decreasing costs and improved accessibility of NGS have led to a recent trend towards comprehensive molecular profiling, including large-scale analyses such as WGS and WES. These efforts aim to gain a better understanding of the genetic basis of cancer biology.

However, connecting cancer phenotypes or metastatic patterns to molecular alterations identified by NGS is a complex process that involves multiple steps, each with its own challenges and potential difficulties. One sample-specific challenge is the quality and quantity of tumor samples. Conducting such studies requires obtaining a large collection of high-quality tumor samples, which can be particularly difficult in biopsy samples and certain cases, such as multiple brain metastases or multi-organ metastases, which can lead to loss of sensitivity of detection such as signal loss in somatic mutations [168].

Additionally, comprehensive NGS techniques like whole-genome and whole-transcriptome sequencing generate large and complex data that need high-performance computing resources, including large data storage as well as longer turnaround times [168-170]. Furthermore, various technical and biological challenges exist, such as sequencing error rate, chimeric reads, insufficient read length or insufficient or biased coverage, and results interpretation [170, 171]. An important example is the determination of driver genes and passenger genes, i.e., cancer-relevant alterations as well as the phenotypic influence of CNAs in potential driver genes [172]. Therefore, there is a great need to develop more efficient and standardized bioinformatical tools for filtering, analyzing, and interpreting the resulting variants with less biased databases and fewer artifacts. In this project, we aimed to employ the standard bioinformatic workflow, employing diverse tools and approaches, to efficiently process and interpret our NGS data.

1.5. Aim of the study

Despite several large-scale investigations on molecular characteristics of BrM in NSCLC and numerous studies comparing the primary lung tumors and their metastatic sites in the brain, there are very few studies on different BrM patterns. Previous research implicates that oligo-brain metastatic disease, which often carries a more favorable prognosis represents a different form of cancer spread than a rapidly spread disease to many different sites (poly-metastatic) [173]. Clinical data have shown that the number of metastatic sites in NSCLC is associated with survival [174]. It showed that patients diagnosed with oligometastatic NSCLC who received aggressive therapy have demonstrated improved progression-free and overall survival rates [175]. Whether there are host-related (immunological) factors or tumor-specific factors regulating the extent of the disease spread has not been investigated today in detail yet. Therefore, the overall aim of this comprehensive study is to have a better understanding of BrM and identify novel markers that can predict an oligo- or poly-metastatic disease. This could enable early detection of high-risk patients and help clinicians to define those who would benefit most from radical treatment.

The specific aims are:

1. To assess the role of mutational and aberration patterns in different cohorts of brain metastasis using large-scale NGS approaches.
2. To investigate gene expression differences at the RNA level between different BrM groups.
3. To identify potential markers that can differentiate tumors in the brain based on their specific patterns of metastasis.
4. To characterize the different immune cell populations in TME including intra-tumoral and peri-tumoral lesions and in peripheral blood between NSCLC patients with different BrM patterns.

2. Materials

2.1. Laboratory devices

Table 2: Laboratory devices used in this study

Device	Company	Office
Analytical scale BP610	Sartorius	Göttingen, DE
Analytical scale BP6100	Sartorius	Göttingen, DE
Axioplan2 imaging with AxioCam MRm and light sourceHXP120V	Carl Zeiss	Jena, DE
Centifuge Multifuge 3 S-R	Heraeus Holding	Hanau, DE
Centrifuge 5417R	Eppendorf	Hamburg, DE
Centrifuge Rotofix 32	Hettich	Villingen- Schwenningen, DE
CO2-cell culture incubator HERAcell150	Thermo Fisher Scientific	Waltham, MA, US
Electrophoresis power source 250 V	VWR International	Radnor, PA, US
Flow cytometry LSR Fortessa	BD Biosciences	Franklin Lakes NJ, US
Gel documentation system GeneGenius	Syngene	Cambridge, UK
Icemaker FM-120 DE	Hoshizaki	Amsterdam, NL
Mastercycler gradient	Eppendorf	Hamburg, DE
Microwave 800	Severin	Sundern, DE
NanoDrop 1000 spectrophotometer	Thermo Fisher Scientific	Waltham, MA, US
NanoSight LM10instrument	NanoSight Technology	Malvern, UK
Tissue Drying Oven TDO 66	Medite Medical	Burgdorf, DE
Pipette boy	Hirschmann Laboratory equipment	Eberstadt, DE
Pipettes (2,5µl, 10µl, 200µl, 1000µl)	Eppendorf	Hamburg, DE
Qubit 4 Fluorometer	Thermo Fisher Scientific	Waltham, MA, US
Sterile hood Herasafe HS12	Heraeus Kendro	Langenselbold, DE
TapeStation RNA system	Agilent Technologies	Santa Clara, CA, US
Thermomix comfort	Eppendorf	Hamburg, DE
Ventana BenchMark ULTRA	Roche Diagnostics	Mannheim, DE
Vi-CELL® XR Cell Viability Analyzer	Beckman Coulter	Brea, CA, US
Vortex-Genie 2	Scientific Industries	New York, NY, US
Water bath GFL-1002/03	GmbH für Labortechnik	Burgwedel, DE

2.2. Chemicals and reagents

Table 3: Chemicals and reagents used in this study

Chemical	Company	Office
Agarose LE	Genaxxon Bioscience	Ulm, DE
AmpliTaq Gold polymerase	Applied Biosystems	Darmstadt, DE
Aqua	B. Braun Melsungen	Melsungen, DE
Brefeldin A Solution (1000X)	Thermo Fisher Scientific	Waltham, MA, US
DMSO (dimethyle sulfoxide)	Serva	Heidelberg, DE
DNA-Marker GeneRuler 1 kb DNA Ladder	Thermo Fisher Scientific	Waltham, MA, US
DNase I from bovine pancreas	Roche Diagnostics	Mannheim, DE
dNTPs (desoxyribonucleoside triphosphate set)	Roche Diagnostics	Mannheim, DE
DPBS (1X), no calcium, no magnesium	Gibco	Billings, MT, US
Eosin Solution	Carl Roth	Karlsruhe, DE
Ethanol absolute	Merck	Darmstadt, DE
FcR Blocking Reagent, human	Miltenyi Biotec	Bergisch Gladbach, Germany
FCS (fetal calf serum)	PAA Laboratories	Pasching, A

Fermacidal D2	IC products SA	Gordola, CH
Ficoll® Paque Plus	Sigma-Aldrich	St. Louis. MO , US
Flow-cytometry staining buffer	Thermo Fisher Scientific	Waltham, MA, US
IC Fixation Buffer	Thermo Fisher Scientific	Waltham, MA, US
Ionomycin calcium salt	Sigma-Aldrich	St. Louis. MO , US
Mayer's hemalum Solution	Sigma-Aldrich	St. Louis. MO , US
MgCl2 solution	Applied Biosystem	Waltham, MA, US
PCR Buffer (TAE)	Sigma-Aldrich	St. Louis. MO , US
Permeabilization Buffer (10x)	Thermo Fisher Scientific	Waltham, MA, US
PMA (Phorbol 12-Myristate 13-Acetate)	Sigma-Aldrich	St. Louis. MO , US
Primers	Eurofins	Ebersberg, DE
Protease	Qiagen	Hilden, DE
RPMI 1640 Medium	PAN Biotech	Aidenbach, DE
SeaKem LE Agarose	Biozym	Hamburg, DE
X-VIVO 15, serum-free, hematopoietic cell medium	Biozym	Hamburg, DE
Xylol	Chemsolute	Renningen, DE
Zombie NIR	Biolegend	San Diago, CA, US

2.3. Consumables

Table 4: Consumables used in this study

Consumable	Company	Office
Cryogenic Tubes (2ml)	Thermo Fisher Scientific	Waltham, MA, US
Falcon tubes (15ml)	Falcon	AZ, US
Falcon tubes (50 ml)	Falcon	AZ, US
Flow cytometry tubes	Sarstedt	Nümbrecht, DE
Microscopic slides	Langenbrinck GmbH	Emmendingen, DE
Pipette tips	Sarstedt	Nümbrecht, DE
Plastic Transfer Pipettes	Thermo Fisher Scientific	Waltham, MA, US
Safe-lock tubes (0.5ml, 1.5ml, 2ml)	Eppendorf	Hamburg, DE
Serological pipettes	Sarstedt	Nümbrecht, DE
Sterile Disposable Serological Pipettes	Thermo Fisher Scientific	Waltham, MA, US

2.4. Commercially available kits

Table 5: Commercially available kits used in this study

Kit	Company	Office
All prep® DNA/RNA/miRNA universal kit	Qiagen	Hilden, DE
QIAamp® DNA blood mini kit (50)	Qiagen	Hilden, DE
RNA TapeStation System (4200)	Agilent Technologies	Santa Clara, CA, US

2.5. Antibodies

Table 6: antibodies used in this study

Antibody	clone	Company	Office
2B4	C1.7	Biolegend	San Diago, CA, US
Antibody diluent green		DAKO, Agilent Technologies	Santa Clara, CA, US
BTLA	MIH26	Biolegend	San Diago, CA, US
CCR4	L291H4	Biolegend	San Diago, CA, US
CCR6	G034E3	Biolegend	San Diago, CA, US

CCR7	G043H7	Biolegend	San Diago, CA, US
CD127	A019D5	Biolegend	San Diago, CA, US
CD161	HP-3G10	Biolegend	San Diago, CA, US
CD25	BC96	Biolegend	San Diago, CA, US
CD27	O323	Biolegend	San Diago, CA, US
CD28	CD28.2	Biolegend	San Diago, CA, US
CD3	UCHT1	Biolegend	San Diago, CA, US
CD3 (M3074)	SP7	Zytomed	Berlin, DE
CD39	A1	Biolegend	San Diago, CA, US
CD4	RPA-T4	Biolegend	San Diago, CA, US
CD4 (M731001-2)	4B12	DAKO, Agilent Technologies	Santa Clara, CA, US
CD45RA	HI100	Biolegend	San Diago, CA, US
CD45RO	UCHL1	Biolegend	San Diago, CA, US
CD56	5.1H11	Biolegend	San Diago, CA, US
CD57	QA17A04	Biolegend	San Diago, CA, US
CD68 (M087629-2)	KP1	DAKO, Agilent Technologies	Santa Clara, CA, US
CD71	CY1G4	Biolegend	San Diago, CA, US
CD73	AD2	Biolegend	San Diago, CA, US
CD8 (C1008C01)	SP16	DCS	Hamburg, DE
CD8a	RPA-T8	Biolegend	San Diago, CA, US
CD95 (Fas)	DX2	Biolegend	San Diago, CA, US
CD98	MEM-108	Biolegend	San Diago, CA, US
CRTH2	BM16	Biolegend	San Diago, CA, US
CXCR3	G025H7	Biolegend	San Diago, CA, US
CXCR5	J252D4	Biolegend	San Diago, CA, US
FITC Mouse Anti-Human TCR	WT31	BD Biosciences	Franklin Lakes NJ, US
FoxP3 (98377)	D2W8E	Cell Signaling Technology	Danvers, MA, US
GLUT1	#202915	R&D Systems	Minneapolis, MN, US
HLA-DR	L243	Biolegend	San Diago, CA, US
IFNg	4S.B3	Biolegend	San Diago, CA, US
IL-10	JES3-9D7	Biolegend	San Diago, CA, US
IL-17A	BL168	Biolegend	San Diago, CA, US
IL-2	MQ1-17H12	Biolegend	San Diago, CA, US
IL-4	MP4-25D2	Biolegend	San Diago, CA, US
Ki67 (275R-15)	SP6	Cell Marques	Rocklin, CA, US
KLRG1	2F1/KLRG1	Biolegend	San Diago, CA, US
Mouse Anti-Human CD68	Y1/82A	BD Biosciences	Franklin Lakes NJ, US
PD-1	EH12.2H7	Biolegend	San Diago, CA, US
S1PR1	SW4GYPP	Thermo Fisher Scientific	Waltham, MA, US
Tgd	11F2	BD Biosciences	Franklin Lakes NJ, US
TIGIT	A15153G	Biolegend	San Diago, CA, US
Tim-3	F38-2E2	Biolegend	San Diago, CA, US
TNFa	MAb11	Biolegend	San Diago, CA, US

2.6. Software and bioinformatic tools

Table 7: Software and bioinformatic tools used in this study

Software	Application	Source
ANNOVAR	Annotating genetic variants	https://github.com/WGLab/doc-ANNOVAR
Axiovision	Image processing	www.zeiss.de
BioRender	Online tool for creating scien- tific figures	https://www.biorender.com/

Burrows-Wheeler Aligner (BWA)	Mapping DNA sequences to reference	https://github.com/lh3/bwa
Control-FREEC	Assessing CNAs	https://github.com/BoevaLab/FREEC
DAVID	Functional annotation tool	https://david.ncifcrf.gov/home.jsp
DESeq2	RNA sequencing analysis	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Enricher	Gene set enrichment analysis	https://maayanlab.cloud/Enrichr/
FACSDiva Software	Analysis of FACS data	BD Bioscience, Franklin Lakes, NJ, US
fastp	Removing adapters	https://github.com/OpenGene/fastp
FastQC	quality control tool	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
g:Profiler	Functional enrichment analysis	https://bio.tools/gprofiler
GraphPad Prism	statistical and graphing software	https://www.graphpad.com/
MutEnricher	somatic mutation enrichment analysis	https://github.com/asoltis/MutEnricher
Manta	InDels caller	https://github.com/Illumina/manta
NCBI	Database for literature (PubMed)	www.ncbi.nlm.nih.gov
R program	Statistical analysis	https://www.r-project.org/
R Studio	Integrated development environment for R	https://rstudio.com/
Ruby on rails	web application framework	https://rubyonrails.org/
samtools	Removing low-quality and duplicates	https://github.com/samtools/samtools
snpEff	Genetic variant annotation	http://pcingola.github.io/SnpEff/
STAR Aligner	aligning RNA sequencing data	https://github.com/alexdobin/STAR
Strelka	SNVs caller	https://github.com/Illumina/strelka

2.7. Databases

Table 8: Databases used in this study

Database	**Source**
cBioPortal	https://www.cbioportal.org/
Human reference genome, hg38 (GRCh38.100)	https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.26/
The Cancer Genome Atlas Program (TCGA)	https://portal.gdc.cancer.gov/
UCSC Genome Browser	https://genome.ucsc.edu/

3. Methods

3.1. Patient cohorts

All patients included in this study were operated for lung cancer brain metastases at the department of Neurosurgery at University Hospital Hamburg-Eppendorf (UKE) between 2014-2020. The main clinical characteristics of 74 NSCLC BrM patients are summarized in Table 10 including tissues and liquid biopsy samples. Of these patients, 61 tissue samples were used for NGS analyses, 56 for tissue immune-profiling and 54 blood samples were used for immunophenotyping including 34 cases with both tissues and blood available. Figure 8 displays an overview of the study design. This study was approved by the local ethical committee of the Hamburg chamber for physicians and was performed in accordance with the Helsinki Declaration of 1975 under numbers Nr. PV5392 & Nr. PV4954.

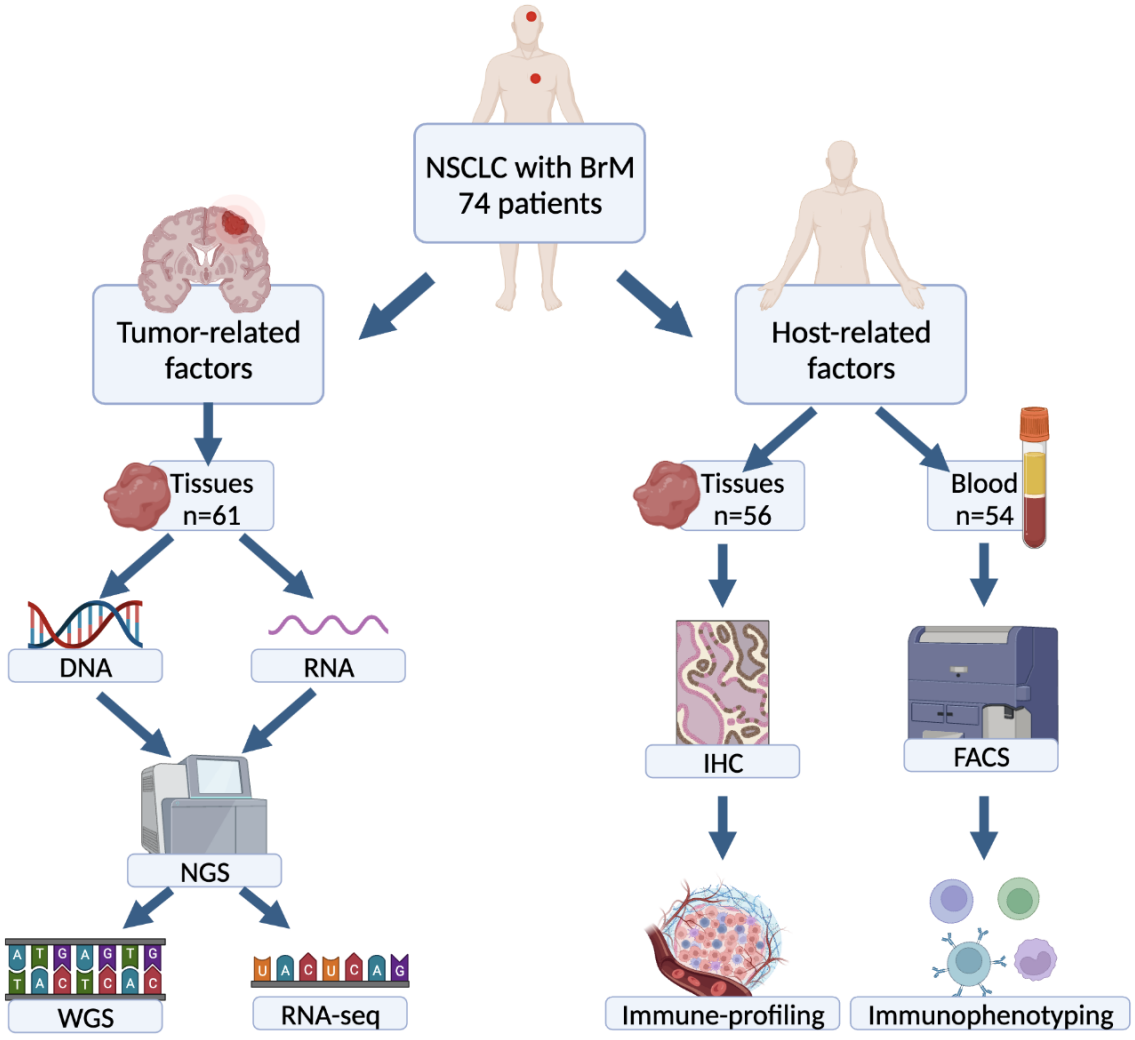


Figure 8: Flowchart of the study design: 74 patients including 61 tissue and 54 blood samples: tissue samples were employed to investigate tumor-related factors through for NGS of DNA and RNA. To investigate host related factors, immune-profiling was performed on tissue samples via Immunohistochemistry (IHC), while immunophenotyping was conducted on blood samples using Flow Cytometry (FACS).

3.2. DNA and RNA extraction

For NGS analysis, 61 fresh-frozen BrM tissue samples were obtained from surgically resected BrM of NSCLC patients. Both DNA and RNA were extracted from these samples using the All prep® DNA/RNA/miRNA universal kit according to the manufacturer's instructions. However, RNA extraction was successful for only 40 of these samples. Among the total 61 patients, 45 matching blood samples were available and utilized to isolate germline DNA using the QIAamp® DNA blood mini kit, which served as normal controls for mutation analysis. The quality of DNA samples was assessed by conducting multiplex GAPDH PCR, and they were deemed of sufficient quality if at least one detectable band within the range of 200-400 bp was observed. The concentration of DNA samples was quantified using the Qubit 4 Fluorometer. Similarly, the quality, quantity, and integrity of the total RNA were evaluated using the NanoDrop1000 spectrophotometer and TapeStation RNA ScreenTape assay. Samples were considered qualified for sequencing when they obtained a RIN (RNA Integrity Number) score greater than 6.0.

3.3. Next-generation sequencing

Qualified DNA (≥ 12.5ng/µl) and RNA (≥ 10ng/µl) samples were shipped to BGI (Hong Kong, China) and (Copenhagen, Denmark), for library preparation, quality control, and sequencing. Subsequently, only samples meeting sufficient quality criteria according to BGI were selected for sequencing. Specifically, 61 DNA samples successfully met the quality standards, while some RNA samples were degraded, resulting in only 28 viable samples. Whole-genome sequencing (WGS) was performed for the qualified DNA samples, and whole transcriptome RNA-seq was conducted for the RNA samples. Both sequencing procedures were carried out using Illumina platforms, achieving a median coverage of 30X.

3.4. DNA whole genome sequencing (WGS)

Whole genome sequencing is a valuable method to detect the complete genome sequence of a human sample in a single analysis, providing a comprehensive insight into its genetic variation compared to a human reference genome.

DNA WGS was performed with the DNBSEQ sequencing system by BGI company. Briefly, library preparation included DNA random fragmentation with an average size of 200-400 bp. These fragments were then end-repaired and 3'adenylated. Subsequently, adaptors were ligated to the ends of these fragments. Following that, the fragments with adapters were amplified by PCR, and the double-stranded PCR products were heat-desaturated and circularized by the

splint oligo sequence. The final library was prepared in single-strand circular DNA (ssCir DNA) format. The qualified libraries were then subjected to sequencing. Every ssCir DNA molecule formed a DNA nanoball (DNB) containing more than 300 of the original DNA fragments tightly packed together. The DNBs were loaded onto a sequencing flow cell and processed for 150 bp paired-end sequencing on the platform.

Genomic data obtained from WGS process were summarized in FASTQ files, which comprised raw genomic information in the form of short sequences referred to as reads. A bioinformatic pipeline was established to analyze somatic mutations including single nucleotide variants (SNVs), small insertions/deletions (InDels), and copy number alterations (CNAs).

The initial step of data pre-processing involved quality-control checks by FastQC v0.11.9 and adapters trimming from the ends of the reads using fastp v0.20.1[176]. Sequence reads were aligned and mapped to the human reference genome (GRCh38) using Burrows-Wheeler Aligner (BWA) v0.7.17. The aligned reads were sorted and converted to BAM files using samtools v1.10. Low-quality reads and duplicates were marked and discarded, ending with high-quality genome sequencing data ready for subsequent analysis.

3.4.1. Prediction of tumor mutational burden (TMB) and mutational signatures

For TMB and mutaional signatures analyses, only tumor samples with matched blood samples were included. TMB determination was accomplished for each individual sample using SnpEff tool. This involved the computation of the mutation count divided by the estimated exome length in megabases (mutation/Mb). Specifically, non-synonymous mutations were considered for this calculation, ensuring a focused assessment of functionally relevant genetic alterations. Mutational signatures profiling performed using the R package “signature.tools.lib” [177], by analyzing single base substitution (SBS) mutational signatures v3.3 obtained from the COSMIC database.

3.4.2. Variant calling

The last part of primary NGS data analysis is known as variant calling, referring to the identification of somatic variants in the cancer genome and its distinction from benign (healthy) germline. This can be done by aligning the mutational tumor profile and comparing it with the existing published databases, or by using a matching non-cancerous sample from the same patient as a control [178]. For the unpaired samples in this thesis, a “panel of normals” was created by using an assembled mix of available germline DNA reads from the other patients, this mix was used as a reference to pair it with tumor samples for variant calling.

3.4.2.1. SNVs/InDels calling

Short somatic variants including single SNVs and small InDels were called for each sample using Strelka v2.9.10 and Manta v1.6.0. Variants not passing the tools internal filters were not considered for further analysis. The remaining detected mutations were annotated with ANNOVAR to observe their functional and clinical significance. The resulting mutation list was further filtered to remove common artifacts, synonymous variants were excluded, and common single-nucleotide polymorphisms (SNPs) were filtered out if they were present at >1% in either 1000genome project or the Genome Aggregation Database (gnomAD). Other exclusion criteria included mapping quality ≤40, a read depth <10 in tumor DNA and <20 in normal/germline DNA, alternative allele count <4 in the tumor, variant allele frequency (VAF) <5% in the tumor and >20 in the normal. The Clinvar and Intervar databases were also used to remove the variants marked as "benign" or "likely benign" to narrow down the final number of somatic mutations and identify cancer-relevant variations.

The oncoprint heatmap was generated using the R package "ComplexHeatmap". The mutation frequencies of genes were extracted from the Pan-Lung Cancer TGCA dataset in cBioPortal. Additionally, putative driver genes identified by the MutEnricher tool were considered. Only variants with high or moderate effected predicted by snpEff in at least one sample were considered.

3.4.2.2. CNA calling

CNA calling and analyses were performed on the DNA-WGS data using Control-FREEC v11.6 [179]. This software utilizes aligned reads as input, and subsequently constructs and normalizes a copy number profile. These profiles are then segmented, and genotype status is assigned to each segment based on copy number. Finally, genomic alterations are annotated as either gains, losses or unchanged.

For this analysis, we used equally sized, non-overlapping windows of 300 kb to compute the read-depth ratio for each window and estimate gains, losses, and neutral regions. A focused analysis was also conducted to look for amplifications/deletions in genes that frequently harbor oncogenic alterations. A gene copy number >5 was considered as high-level amplification, and copy number 0 or 1 as deletion. Subsequently, frequencies were calculated along the whole genome for each sample group. We considered gains or losses within each group when either of them was observed in more than 30% of the samples. The p-values were calculated using Fisher's exact test. CNAs were deemed significant when p-values were less than 0.05 for initial assessment and subsequently when p-values were less than 0.01 for a more stringent criterion.

We identified altered regions by selecting observed alterations with at least five consecutive significant p-values in a row when comparing between groups. To visualize CNAs, copy number frequency plots were generated for all samples combined and for samples from each BrM group.

In summary, Figure 9 presents the comprehensive bioinformatic pipeline employed for NGS analysis.

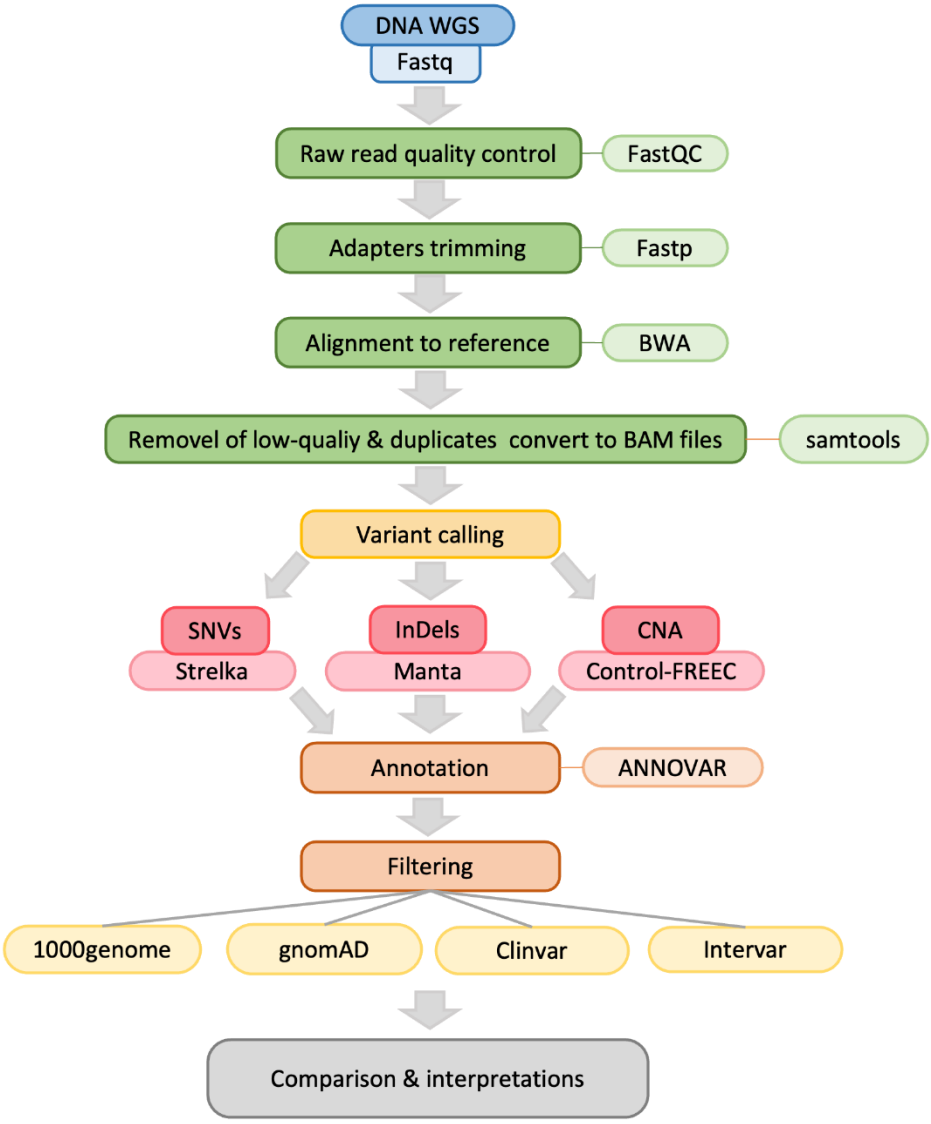


Figure 9: Bioinformatic pipeline for NGS analysis, encompassing quality control, alignment, and pre-processing of individual samples. Subsequently, NGS sequencing variant calling, which includes SNVs, InDels, and CNVs, followed by somatic variant annotation and filtering, leading to the final interpretation of results.

3.4.3. Transcriptome RNA-sequencing

RNA-seq is a powerful tool that offers valuable insights into gene expression patterns. It can help to investigate the impact of DNA-level gene mutations on gene expression.

A total of twenty-eight qualified RNA samples were sequenced using DNBSEQ-G400 platform by BGI. In summary, library preparation started with fragmenting RNA into small pieces. followed by reverse transcription to generate complementary DNA (cDNA) strands. Similar to DNA WGS, adapters were ligated to the ends of these 3' adenylated cDNA fragments, and PCR amplification was performed. Resulting double-stranded PCR products were heat denatured and the final library was formatted as ssCir DNA. The DNBs were loaded onto a sequencing flow cell and single-end 50 (pair-end 100) bases reads were generated using sequencing by synthesis.

Sequenced RNA data processing and filtering followed the same steps as DNA data. Consequently, RNA-seq reads were aligned against the human reference genome (GRCh38) with ENSEMBL annotation using STAR Aligner v2.7.3a [180]. To avoid the number of false-positive differentially expressed genes (DEGs) with low abundance, only transcripts with a median read count >10 in at least one group of samples were retained.

In total, 17478 genes were tested in DE analyses. Principal component analysis (PCA), gene expression-based pairwise distances between samples, and plots were performed on regularized log transformation (rlog) and transformed data were implemented in DESeq2 [181]. Differential mRNA expression between groups was calculated using negative binomial generalized linear models implemented in DESeq2. Genes with a Benjamini-Hochberg-based false discovery rate (FDR) < 0.05 were considered significant.

3.4.3.1. Gene ontology and pathway enrichment analysis

Differentially expressed genes (DEGs) from RNA-seq analysis were subjected to Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis using mainly g:Profiler [182], in addition to DAVID (v6.7) [183] and Enricher [184]. The goal of these analyses was to gain insights into the functional annotation and biological pathways associated with the DEGs.

3.5. Immunohistochemistry (IHC)

Immunohistochemical staining (IHC) was conducted on 2 µm thick sections of formalin-fixed, paraffin-embedded (FFPE) tissue samples obtained from 56 BrM cases from the Institute of Neuropathology, University Hospital Hamburg-Eppendorf (UKE). The following antibodies were used: CD3 (clone SP7; 1:100; M3074, Zytomed), CD8 (clone SP16; 1:500; C1008C01, DCS), CD4 (clone 4B12; 1:50; M731001-2, DAKO), FOXP3 (clone D2W8E; 1:100; 98377, Cell Signaling), CD68 (clone KP1; 1:100; M087629-2, DAKO), and Ki67 (clone SP6; 1:750;

275R-15, Cell Marques). All staining was performed on a Ventana benchmark XT autostainer following the manufacturer's recommendations.

TILs scoring was assessed based on the proportion of positively stained cells as a percentage within two distinct regions: the intra-tumoral region (within the solid tumor tissue area) and the peri-tumoral region (the adjacent area around the tumor tissue). The expression of each marker was evaluated semi-quantitatively by a pathologist as previously described [185]. Briefly, for CD3, CD8, CD4, CD68, and Ki67, a scoring system was used, with scores of 0 (negative: no stained cells), 1 (low: <10%), 2 (moderate: 10-40%), or 3 (high: >40%), based on the estimation of positively stained cells relative to the total number of cells in the investigated intra-tumoral or peri-tumoral region. To facilitate statistical analysis, the parameters were grouped into high (including moderate/high) and low (including negative/low) categories. Due to the relatively low number of $FoxP3^+$ cells in all BrM groups, the samples were scored as either negative (0%) or positive (≥1%) in both the intra-tumoral and peri-tumoral regions.

3.6. Immune cell isolation

Peripheral blood samples were collected from 54 BrM-NSCLC patients using 7.5 ml EDTA-containing tubes prior to surgical removal of the tumor. PBMCs were isolated within 2 hours of blood draw using Ficoll® gradient centrifugation, following the established protocol [186]. Subsequently, the isolated PBMCs were stored in RPMI/10% DMSO (P04-17500) at -80°C until further use.

3.7. Multicolor flow cytometry

Frozen PBMCs were thawed in a water bath at 37°C, washed with 4°C cold 10% FBS in RPMI, resuspended in a cold medium (RPMI, 10% FCS and DNAse (1:1000)) (4°C) and cells were counted using Vi-CELL® XR Cell Viability Analyzer (383556, Beckman Coulter). Five different panels (43 antibodies) were used to analyze the different immune cell states; T-cell exhaustion, T-cell differentiation, T helper cell subsets, T-cell metabolism, and cytokines secretion (as listed in Table 9).

Table 9: Antibodies used for the peripheral blood immuno-phenotyping

Panel 1 **T-cell exhaustion**		**Panel 2** **T-cell subtypes**		**Panel 3** **Th subsets**		**Panel 4** **T-cell metabolism**		**Panel 5** **Cytokines**	
Epitope	Clone	Epitope	Clone	Epitope	Clone	Epitope	Clone	Epitope	Clone
CD25	BC96	CD45RA	HI100	CD45RA	HI100	CD45RA	HI100	CD45RO	UCHL1
KLRG1	2F1/KLRG1	CD127	A019D5	CXCR3	G025H7	CD4	RPA-T4	TNFa	MAb11
CD127	A019D5	CD27	O323	CD161	HP-3G10	FAS	DX2	CD3	UCHT1
CD57	QA17A04	CD4	RPA-T4	CD4	RPA-T4	CD3	UCHT1	IL-2	MQ1-17H12

CD4	RPA-T4	CD25	BC96	CRTH2	BM16	CD98	MEM-108	CD4	RPA-T4
BTLA	MIH26	CD3	UCHT1	CD3	UCHT1	CD71	CY1G4	IFNg	4S.B3
CD3	UCHT1	HLA-DR	L243	Tgd	11F2	GLUT1	202915	CD56 (surf)	5.1H11
2B4	C1.7	CD39	A1	CCR4	L291H4	Zombie NIR	-	IL-4	MP4-25D2
CD56	5.1H11	CD28	CD28.2	CXCR5	J252D4	CD8	RPA-T8	IL-10	JES3-9D7
TIGIT	A15153G	CD73	AD2	CCR6	G034E3	PD-1	EH12.2H7	Zombie NIR	-
Tim-3	F38-2E2	Zombie NIR	-	Zombie NIR	-			CD8 (surf)	RPA-T8
Zombie NIR	-	CD8	RPA-T8	CD8	RPA-T8			IL-17A	BL168
CD8	RPA-T8	CCR7	G043H7	S1PR1	SW4GYPP				
PD-1	EH12.2H7								

For cell surface staining, samples were resuspended in flow-cytometry staining buffer (eBioscience) containing Fc-block (Miltenyi Biotec) and incubated for 45 minutes in the dark at room temperature with the antibody cocktails. For intracellular staining (cytokines), samples were incubated for 4 hours at 37°C and 5% CO_2. Subsequently, a stimulation mix consisting of Phorbol 12-Myristate 13-Acetate (PMA) (1 µg/ml; P1585, Sigma), Ionomycin calcium salt (1 mg/ml; 10634, Sigma), Brefeldin A Solution 1000X (3 mg/ml; 00-4506, Invitrogen) and X-Vivo 15 serum-free (881024, Biozym) was added to each sample. These samples were incubated for an additional 5 hours under the same previous conditions. After incubation, samples were washed, resuspended in flow-cytometry staining buffer, and stained with a surface antibody cocktail for 10 minutes followed by the addition of live/dead staining, and further incubation for 20 minutes in the dark at room temperature. Subsequently, the cells were subjected to additional washing steps and then fixed and permeabilized before staining with the cytokine (intracellular) antibody cocktail for 30 minutes in the dark at room temperature. After the staining procedure, the cells were washed and resuspended in the flow cytometry staining buffer. The analysis was conducted using a BD LSR Fortessa flow cytometer and data were exported as .fcs files. Manual gating was performed using the FACSDiva software (version 9.1 Becton Dickinson). The gating strategies used in the analysis are illustrated in figures 10-14.

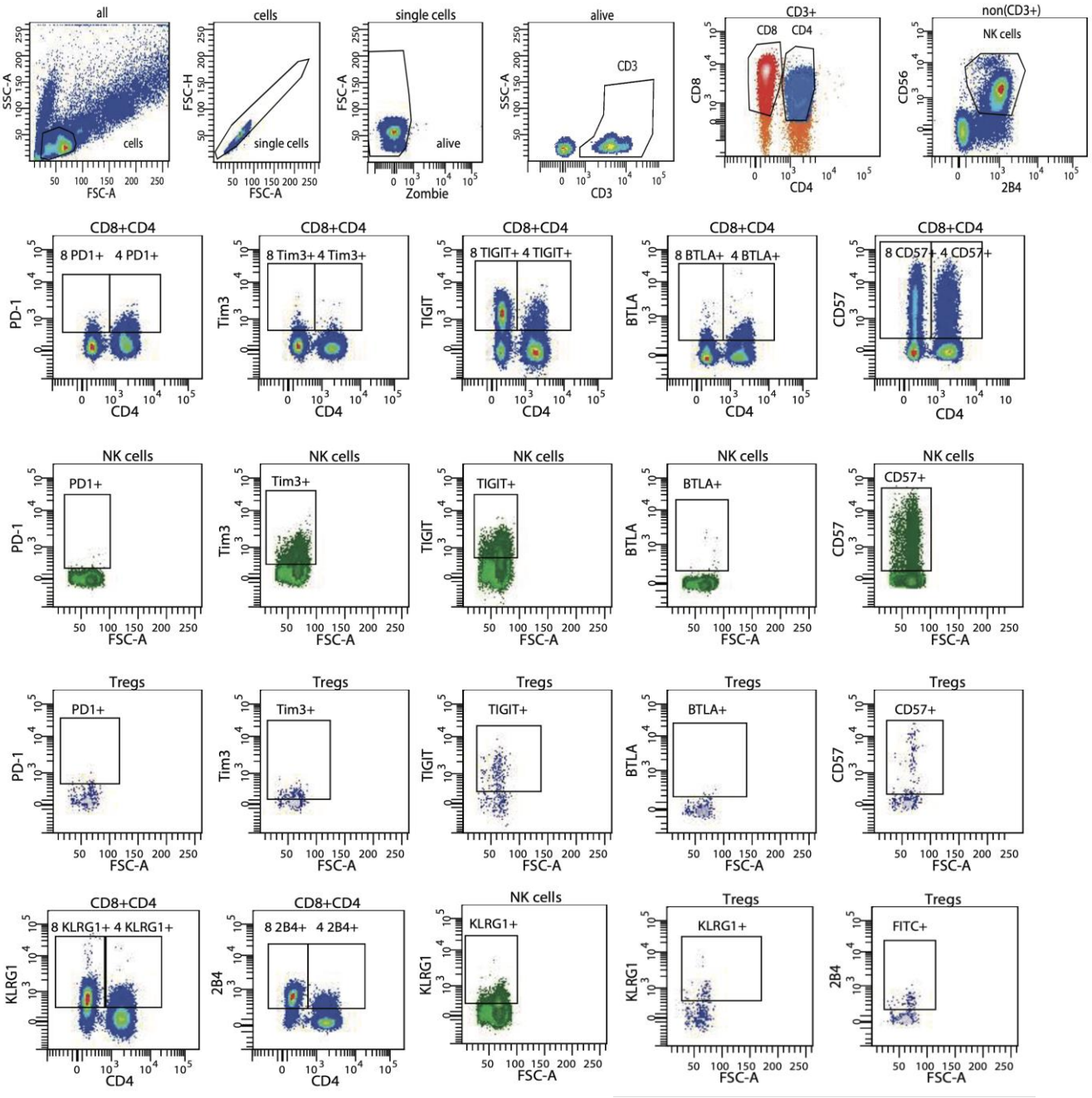


Figure 10: Flow cytometry gating strategy for T-cell exhaustion markers (panel 1)

Lymphocytes were gated based on Forward Scatter (FSC-A) against Side Scatter (SSC-A), and single cells were selected using FSC-A versus Forward Scatter Height (FSC-H). Live cells were identified based on their negative stain for Zombie NIR. T-cells were selected by gating on $CD3^+$ cells, which were further categorized based on markers expressions into $CD4^+$ and $CD8^+$ T-cells. NK cells were identified as $CD56^+$ and $2B4^+$ cells, and Treg cells were selected as CD4+ CD25+ CD127-. To assess T-cell exhaustion, various cell surface markers were used on $CD4^+$, $CD8^+$, Treg and NK cells; PD1, Tim3, TIGIT, BTLA, CD57, KLRG1 and 2B4.

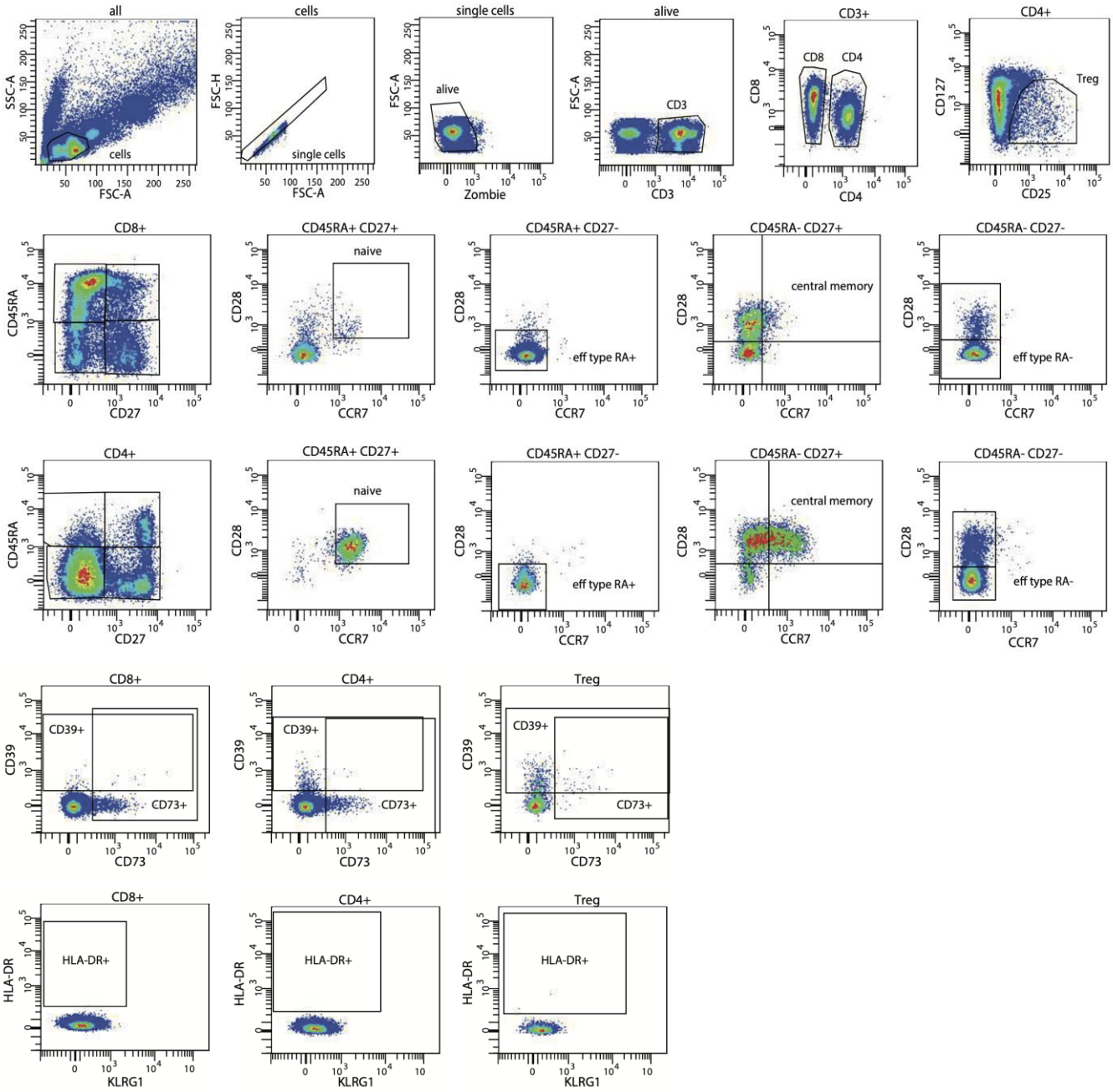


Figure 11: Flow cytometry gating strategy for T-cell differentiation markers (panel 2)

Lymphocytes were gated based on FSC-A against SSC-A, and single cells were selected using FSC-A versus FSC-H. Live cells were identified based on their negative stain for Zombie NIR. T-cells were selected by gating on $CD3^+$ cells, which were further categorized based on markers expressions into $CD4^+$ and $CD8^+$ T-cells. Treg cells were selected as CD4+ CD25+ CD127-. T-cells populations were defined based on CD45RA, CD27, CD28 and CCR7 marker expression as follows: naïve (CD45RA+ CD27+ CD28+ CCR7+), T effectors type RA+ (CD45RA+, CD27−, CD28−, CCR7−), Early memory (CD45RA-, CD27+, CD28+, CCR7-), Early-like memory (CD45RA-, CD27-, CD28+, CCR7+), Intermediate (CD45RA-, CD27+, CD28-, CCR7-), Central memory (CD45RA−, CD27+, CD28+, CCR7+), and TeffRA- (effector memory cells: CD45RA−, CD27−, CD28−, CCR7−). $CD73^+$, $CD39^+$ and HLA-DR^+ cells were identified using their markers expressions on $CD4^+$, $CD8^+$ and Treg.

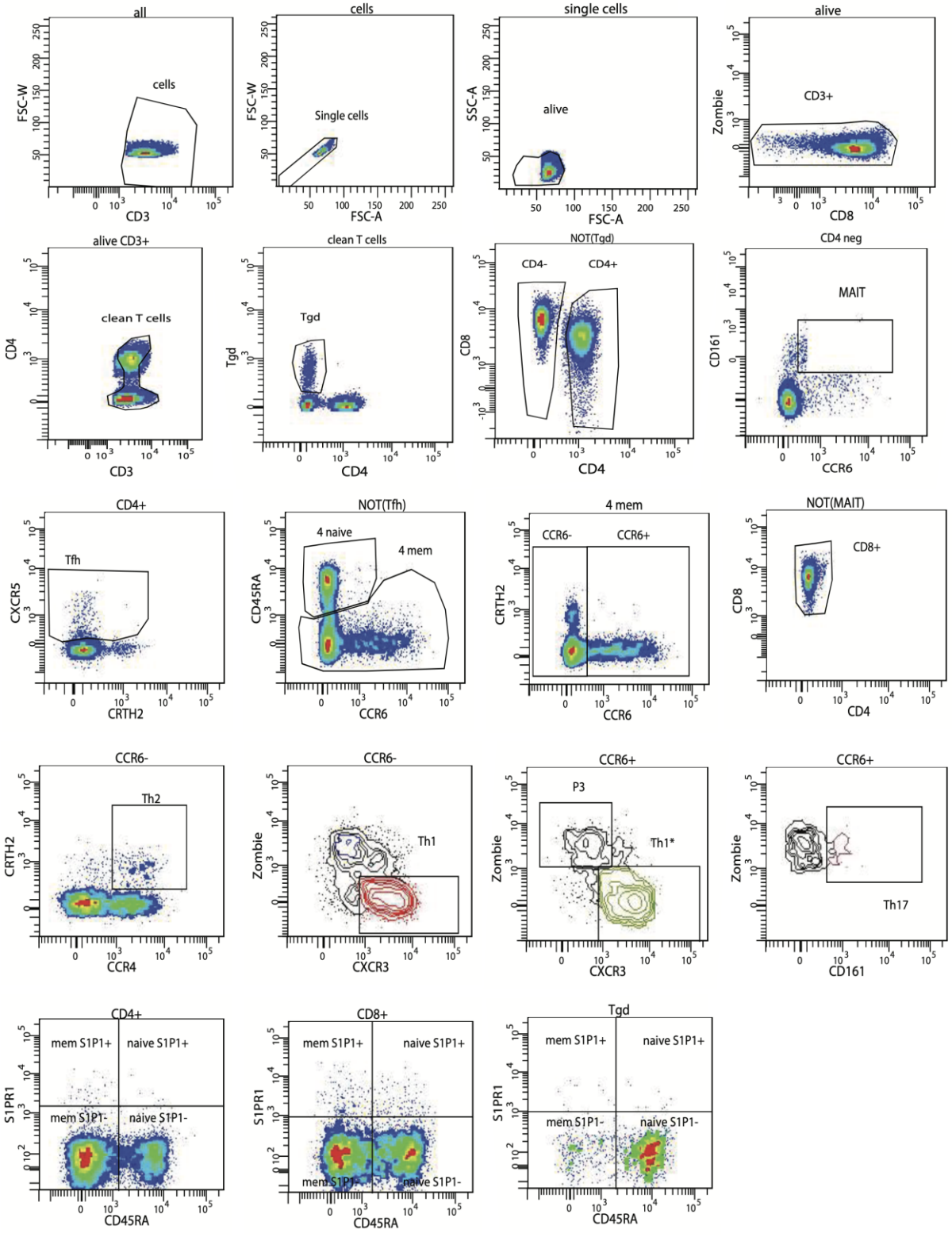


Figure 12: Flow cytometry gating strategy for T helper cell subsets markers (panel 3)

Specific T-cell subsets were identified based on marker expressions. Gamma delta T-cells (Tgd) were determined using the corresponding marker. T Follicular Helper cells (Tfh) were identified by their CXCR5+ expression, while Mucosal-associated invariant T (MAIT) cells were characterized as CD4- CCR6+ CD161+. The CD45RA marker was employed to differentiate between naïve and memory T-cells, and CCR6 served as a marker for T helper cells (Th). Th2 cells were defined as CCR6- CCR4+ CRTH2+, Th1 cells as CCR6- CXCR3+, Th1* cells as CCR6+ CXCR3+, and Th17 cells as CCR6+ CD161+. Additionally, the S1PR1 marker was used to monitor T-cell dynamics within CD4+, CD8+, and Tgd cell populations.

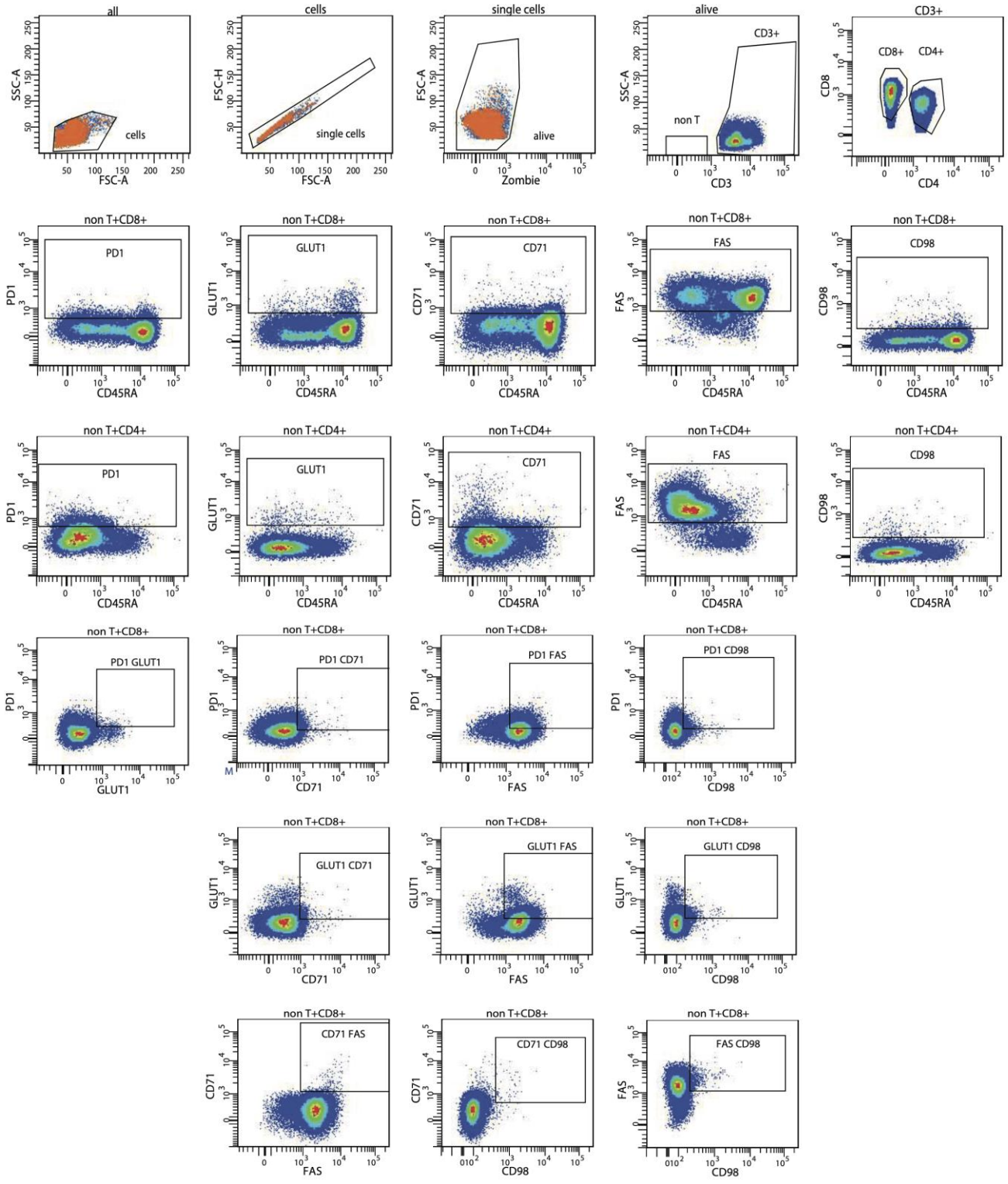


Figure 13: Flow cytometry gating strategy for T-cell metabolism markers (panel 4)
Lymphocytes were gated based on FSC-A against SSC-A, and single cells were selected using FSC-A versus FSC-H. Live cells were identified based on their negative stain for Zombie NIR. T-cells were selected by gating on $CD3^+$ cells, which were further categorized based on markers expressions into $CD4^+$ and $CD8^+$ T-cells. $PD\text{-}1^+$, $GLUT1^+$, $CD71^+$, FAS^+ and $CD89^+$ expressing cells are identified by their corresponding markers on $CD4^+$ and $CD8^+$ cells.

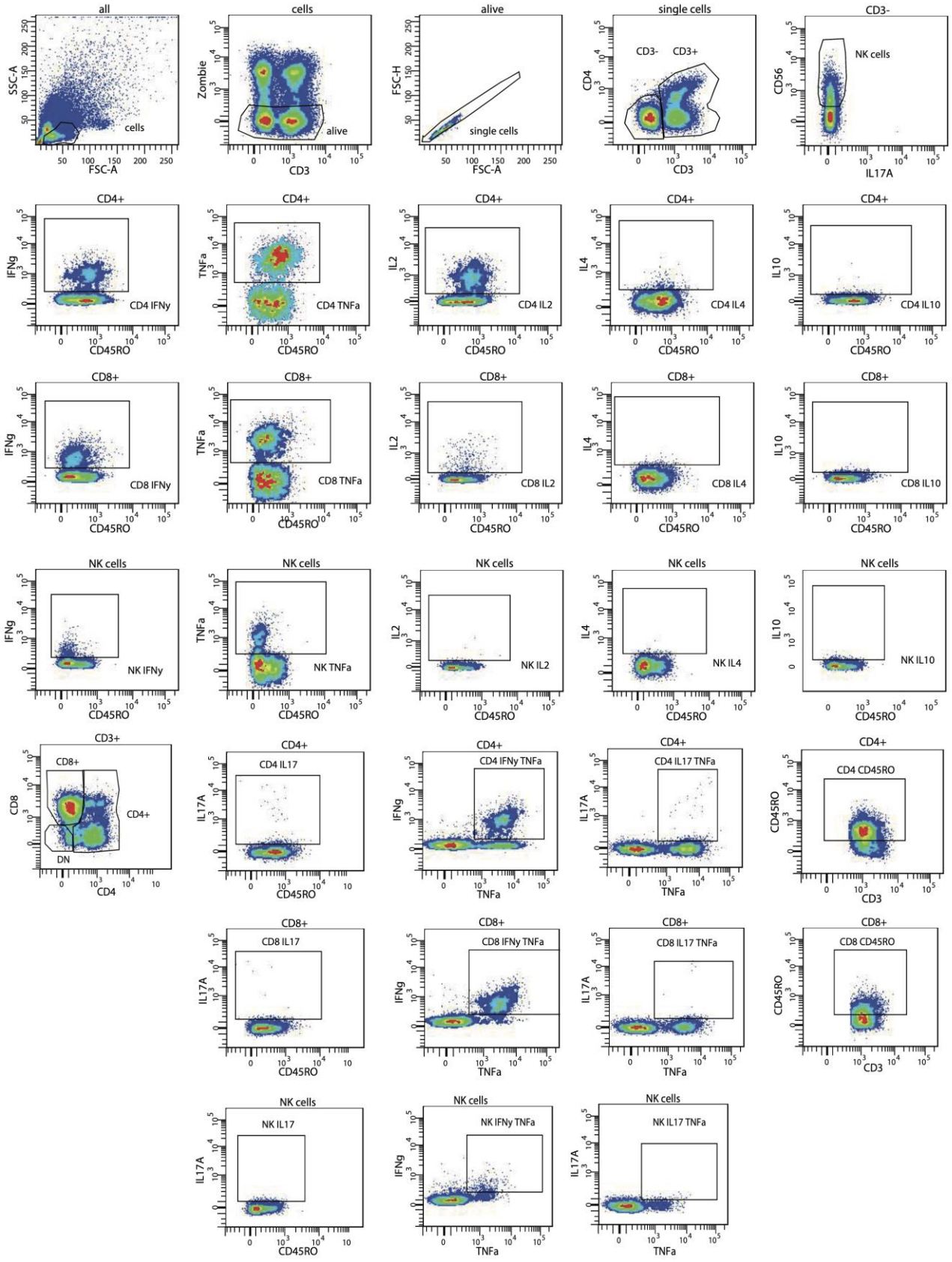


Figure 14: Flow cytometry gating strategy for cytokines markers (panel 5)

Lymphocytes were gated based on FSC-A against SSC-A, and single cells were selected using FSC-A versus FSC-H. Live cells were identified based on their negative stain for Zombie NIR. T-cells were selected by gating on $CD3^+$ cells, which were further categorized based on markers expressions into $CD4^+$ and $CD8^+$ T-cells. NK cells were identified as $CD56^+$ cells. The cytokines markers, including IFNy, TNFa, IL2, IL4, IL10 and IL17 were determined based on their corresponding markers expressions on $CD4^+$, $CD8^+$ and NK cells.

Samples with a count of less than 1000 living T cells were excluded from the analyses, leading to variations in the number of samples included in each panel. Specifically, panel 1 comprised 45 patients, panel 2 included 44 patients, panel 3 encompassed 47 patients, panel 4 consisted of 31 patients and finally, panel 5 included 28 patients.

3.8. Statistical analyses

Data analysis was performed with In-Silico Online v2.3.1 [187], GraphPad Prism Software (v9.5.0, GraphPad Inc., Boston, MA, USA), and R version 4.1.3 [188]. The R-packages "ggplot2" and "dakl/oncoprint" were employed to generate graphs. Fisher's exact test was utilized to analyze associations between independent nominal data, while McNemar's test was used for dependent data. Kruskal-Wallis and Wilcoxon's tests were employed to assess the significance of differences between medians, and ANOVA was used to calculate the significance of mean differences. A significance level of 0.05 was considered statistically significant and was adjusted for multiple testing when necessary. Survival analysis was conducted using the Kaplan-Meier formula from the "survival" R package. The visualization of the survival curves was achieved through the utilization of the ggsurvplot function from the "survminer" R package. The statistical significance of the survival differences was assessed using the log-rank test. For the survival analysis, patients were divided into high and low-expression groups based on median expressions, except for FOXP3, where groups were categorized based on positive or negative staining. In the peripheral-blood immunophenotyping, 48 patients were included in the survival analysis out of 56, as we excluded patients who had no follow-up or were deceased during the perioperative period of 3 months. Overall survival (OS) was estimated from the date of brain surgery until death or censored at the last follow-up. To calculate the correlation between mutational signatures, the spearman correlation test was used. Hazard ratios (HR) were estimated with 95% confidence intervals using Cox proportional regression model.

For panel 2 of flow cytometry, UMAP unsupervised clustering analysis was performed. Compensated and normalized flow data were used to generate UMAP dimensionality reduction (using the R-package “umap”) and PhenoGraph clustering of all groups together (using R-package “Rphenograph”). The heatmap of marker expression was calculated as the median fluorescence intensity of each marker for each phenoGraph cluster. Subsequently, comparisons were calculated with the Kruskal-Wallis test.

4. Results

4.1. Clinical characteristics of patients

The main clinico-pathological features of 74 NSCLC BrM patients are presented in table 10. Patients harboring known driver mutation in *EGFR, ALK* or *ROS* were excluded from the cohort. The samples were categorized into two main groups: patients with oligometastatic NSCLC who had metastasis limited to the brain (lymph node dissemination allowed) and patients with polymetastatic NSCLC who had metastases in the brain and other extracranial organs. The oligometastatic BrM group was further divided into two sub-groups: patients with synchronous oligometastasis (only BrM at time of primary diagnosis) and patients with metachronous oligometastatic disease (only BrM after primary diagnosis). Regarding the main clinical factors considered for the three groups, neoadjuvant treatment was the only variable that showed a significant difference between the groups ($p < 0.0001$, Fisher's exact test). Among the oligo-metachronous patients, the majority had received prior treatment (20/21), while only two patients in the oligo-synchronous group (2/27) and ten patients in the poly-metastatic group (10/26) were pre-treated before the brain metastasis operation.

Table 10: Patient characteristics

Characteristics		All BrM	Oligo-synchronous	Oligo-metachronous	Poly-met	Statistical test	P-value
Samples, n (%)		74	27 (36.5)	21 (28.4)	26 (35.1)	Kruskal Wallis	0.368
Gender, n (%)	F	41 (55.4)	13 (48.1)	11 (52.4)	17 (65.4)	Fisher's exact	0.449
	M	33 (44.6)	14 (51.9)	10 (47.6)	9 (34.6)		
Age at PD (y)	Range	33 – 83	45 – 80	34 – 76	33 – 82	ANOVA	0.333
	Mean	60.7	62.3	61.5	58.3		
Age at BM-OP (y)	Range	34 – 83	45 – 83	34 – 78	35 – 83	ANOVA	0.238
	Mean	61.6	62.8	63.6	58.8		
Histology, n (%)	LUAD	70	23	19	26	Fisher's exact	0.141
	LUSC	4	2	2	0		
Time btw PD & BM OP (months)	Range	0 – 94	0 - 94	2 – 70	0 – 30	Kruskal Wallis	0.965
	Mean	11	4.5	24.6	6.8		
Neoadjuvant treatment	none	42	25	1	16	Fisher's exact	<0.0001
	IT	8	1	1	6		
	RCT	24	1	19	4		
Survival (m)	alive	23	8	5	10	Fisher's exact	0.568
	dead	51	19	16	16		
FUP after BM OP (m)	Range	0 - 83	0 - 65	0 - 59	0 - 83	ANOVA	0.678
	Mean	16.8	18.5	15.7	15.8		

BrM: brain metastases, PD: Primary diagnosis, BM-OP: brain metastases operation, F: female. M: male, LUAD: adenocarcinoma, LUSC: Squamous cell carcinoma, none: treatment-naive, IT: immunotherapy, RCT: Radio-/Chemotherapy

4.2. Molecular characterization of BrM tumor tissues in NSCLC tumors

4.2.1. The mutational landscape of BrM-NSCLC groups

Surgically resected BrM tumor tissues were collected from 61 NSCLC patients. Among these patients, 45 matching blood samples were available as germline controls, while the remaining samples with no germline controls were compared to an established "panel of normal" as a reference. Raw sequencing data were processed and analyzed using a well-established bioinformatic pipeline, to investigate the molecular characteristics and compare the mutational landscape in different BrM groups (see Chapter 3.4 and Figure 9).

Somatic non-synonymous mutations, including high- and moderate-impact SNVs and InDels were called and analyzed using Strelka and Manta. After the filtering process, 3821 genes were identified to be mutated in ≥3 samples, with 296 of these genes annotated as cancer-relevant in cBioPortal data. Figure 15 displays the mutation frequencies of the most commonly found genes across all samples.

We compared the somatic mutation landscape in our BrM-NSCLC (UKE) cohort to the published data by The Cancer Genome Atlas (TCGA) on primary NSCLC (Figure 15, left panel). The mutation frequencies in our cohort were comparable yet higher than TCGA primary tumor data, suggesting that the same mutations occur in metastases but with increased frequencies. The majority of identified mutations were high or moderate impact SNVs, while only a few were InDels. The most frequently mutated gene was *TP53*, observed in 72% of samples. Of the *TP53* mutations, 80% were SNVs, while only 20% were InDels. The 15 most frequently mutated genes, along with their respective mutation frequencies, were as follows: *TP53* (72%), *LRP1B* (59%), *CSMD3* (56%), *RYR2* (56%), *USH2A* (49%), *ZNF536* (38%), *XIRP2* (34%), *COL11A1* (34%), *OTOGL* (34%), *HMCN1* (33%), *RP1L1* (33%), *KRAS* (31%), *SMARCA4* (31%), *ERICH3* (28%), *CNTNAP2* (26%).

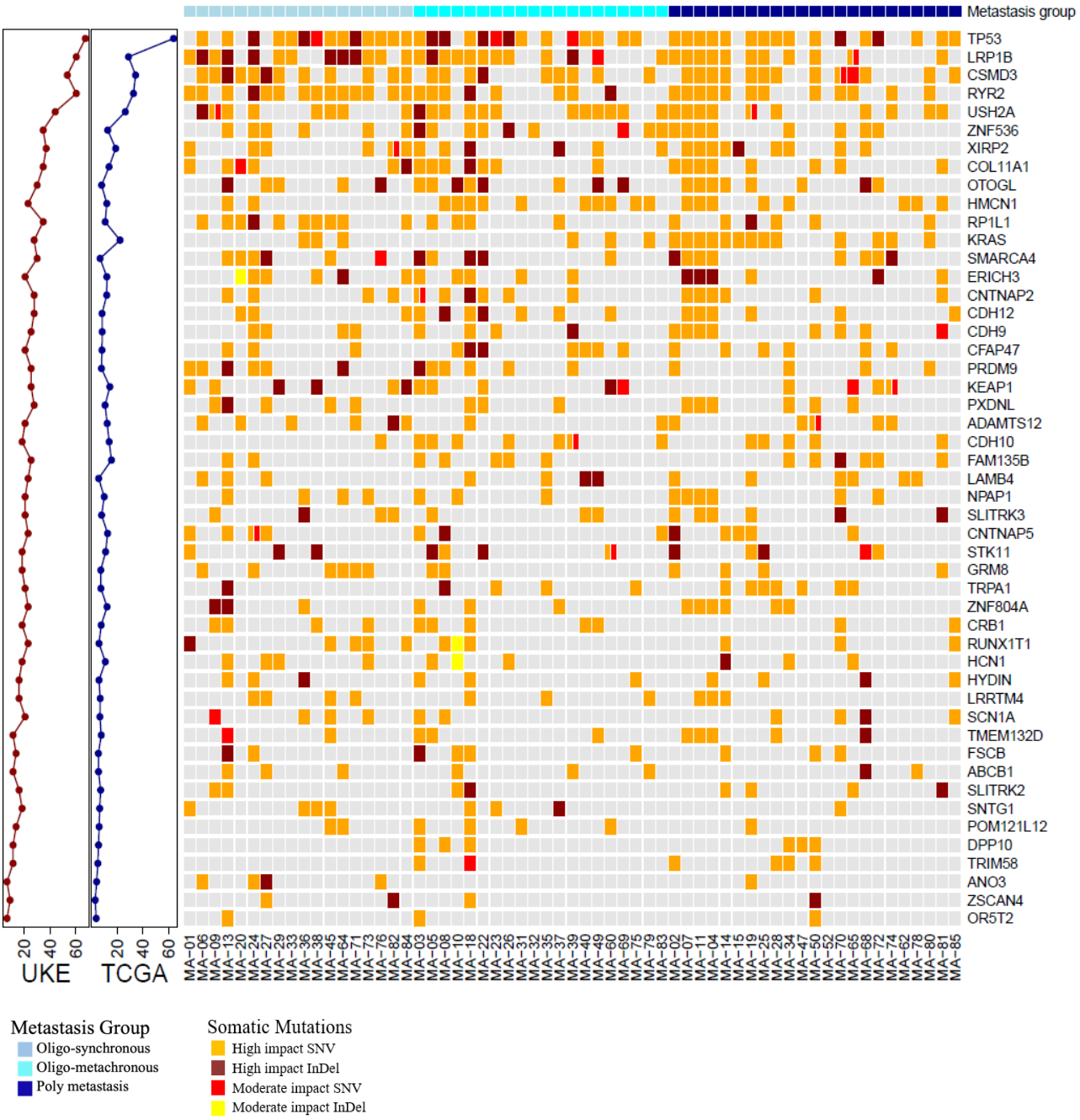


Figure 15: The Mutational landscape of somatic alterations in BrM groups.
The top 50 most frequently mutated genes were selected. Each column represents a single sample and each row a gene, with samples grouped by BrM subtype (synchronous, metachronous and poly metastasis). On the left: somatic mutational landscapes of our BrM-NSCLC (UKE) and TCGA data on primary NSCLC.

In the analysis of mutation frequencies among the three groups of BrM, a total of 241 genes were identified as showing significant differences in mutation rates. However, after applying a correction for multiple testing using the Benjamini-Hochberg method, these significances were no longer observed. To provide an overview of the findings, table 11 presents the top 25 genes that exhibited the most pronounced differential mutation patterns. Notably, Hierarchical clustering analyses of the top 50 mutated genes did not either show a separation for the three BrM groups (Figure 15).

Table 11: Top 25 differently mutated genes among BrM groups (n=61)

Gene	Detected-in-N-samples	Mut % Sync	Mut % Meta	Mut % Poly	p-value	FDR
OBSCN	16	11	60	9	0.0002	0.940
ATP10A	11	6	45	4	0.0006	0.940
GPATCH1	6	0	30	0	0.001	0.940
SLC6A15	5	28	0	0	0.001	0. 940
EPS8L1	8	0	35	4	0.001	0.940
GOLGA6L2	21	17	65	22	0.002	0.940
PEAR1	6	0	0	26	0.002	0.940
ALMS1	13	17	45	4	0.003	0.940
DOCK3	5	0	25	0	0.004	0.940
INPP5B	5	0	25	0	0.004	0.940
MAGEA10	5	0	25	0	0.004	0.940
OAS3	5	0	25	0	0.004	0.940
MYO7B	7	6	30	0	0.004	0.940
OTOL1	7	6	30	0	0.004	0.940
TSHZ3	7	6	30	0	0.004	0.940
MDN1	11	22	35	0	0.004	0.940
PCLO	16	56	10	17	0.004	0.940
CDCP2	4	22	0	0	0.005	0.940
OR1I1	4	22	0	0	0.005	0.940
SLC4A8	4	22	0	0	0.005	0.940
PXDN	6	28	5	0	0.005	0.940
KRAS	19	17	15	57	0.006	0.940
ZNF229	8	0	5	30	0.007	0.940
ZNF462	9	17	30	0	0.008	0.940
BCAS3	6	28	0	4	0.009	0.940

Mut: mutation, Sync: oligo-synchronous BrM, Meta: oligo-metachronous BrM, Poly: poly-metastasis.

To investigate the TMB, nucleotide changes, and mutation signatures, only matched tumor-germline (blood) DNA samples were included (n= 45). Our analysis revealed varying degrees of overall mutational load across the different BrM patients, with median total TMB 4.680 mutation/Mb (range 0.493-31.347 mutation/Mb) (Figure 16, top panel), However, these differences did not show significant differentiation between the BrM groups (p= 0.870, Kruskal-Wallis test).

Furthermore, we examined the specific nucleotide changes present in each tumor sample and performed mutational signatures analysis of single base substitution (SBS) (Figure 16, central and lower panels). The nucleotide changes detected within the tumors predominantly manifested as C > A mutations, which are typically associated with smoking-related diseases such as NSCLC. Subsequently, C > T mutations emerged as the next prevalent mutational type. Therefore, not surprisingly, among the analyzed samples, the most frequently detected mutational signatures were SBS4 (related to smoking: 93%), SBS8 (common in cancer but its etiology is debated: 87%), SBS18 (reactive oxygen species: 49%), SBS2 and SBS13 (APOBEC

signatures: 33% and 42% respectively) and SBS3 (associated with defective homologous recombination-based DNA repair: 29%).

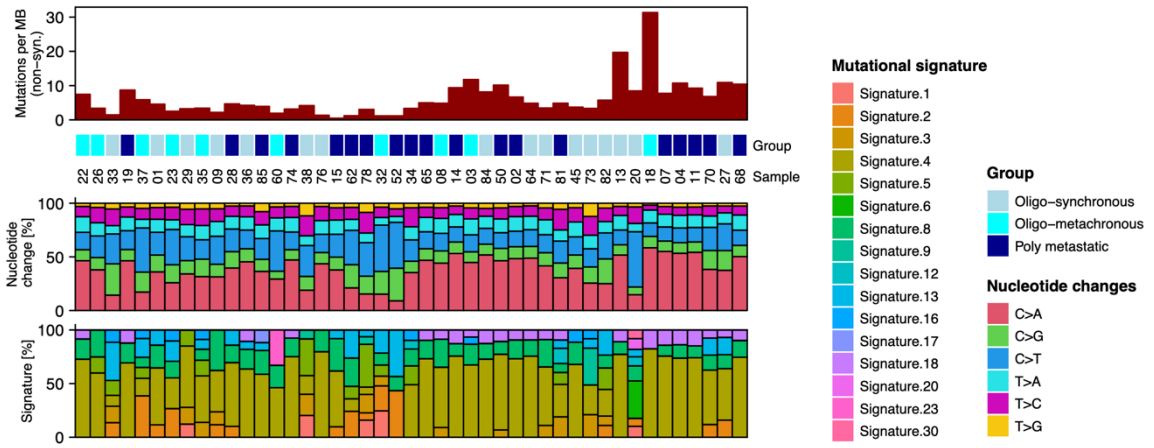


Figure 16: Tumor mutational burden and mutational signatures profiles of paired Tumor-Germline samples (n=45): Top panel illustrates the mutational load (TMB), the central panel shows nucleotide changes in individual samples, and the lower panel presents the mutational signatures.

Table 12: Mutational signatures in BrM groups

Signature	Sync (%)	Meta (%)	Poly (%)	p_value	p-adj
Signature.1	19	10	5	0.509	0.903
Signature.2	38	30	32	0.923	1
Signature.3	44	30	16	0.196	0.902
Signature.4	94	90	95	1	1
Signature.5	25	40	21	0.620	0.903
Signature.6	6	0	0	0.577	0.903
Signature.8	75	90	95	0.281	0.902
Signature.9	13	0	5	0.592	0.903
Signature.12	6	0	11	0.785	1
Signature.13	44	40	42	1	1
Signature.16	19	40	5	0.061	0.902
Signature.17	0	0	5	1	1
Signature.18	38	40	63	0.260	0.902
Signature.20	6	0	0	0.577	0.903
Signature.23	0	10	0	0.222	0.902
Signature.30	6	0	0	0.577	0.903

It is expected to find the smoking signature highly enriched in NSCLC patients, as tobacco smoking is the main risk factor for this cancer, which originates from epithelia directly exposed to tobacco smoke. A strong negative correlation was observed between the smoking signature and the APOBEC signatures (SBS2: r= -0.678, p= 2.973e-07, Spearman correlation test) and (SBS13: r= -0.719, p= 2.541e-08, Spearman correlation test). This suggests an inverse relationship between these mutational processes. Additionally, a strong positive correlation was observed between the smoking signature and the reactive oxygen species signature (r= 0.724,

p= 1.869e-08, Spearman correlation test), confirming the interplay between these mutational processes which are both related to smoking. On the other hand, only a moderate negative correlation was observed between smoking signature and homologous recombination deficiency (HRD) (r= -0.487, p= 0.0006, Spearman correlation test). Moreover, a moderate positive correlation was detected between SBS4 (smoking) and TMB (r= 0.531, p= 0.0001, Spearman correlation test).

One sample exhibited enrichment in the mutational signature SBS6 which is associated with DNA mismatch repair (MMR) deficiency. The patient harbored a somatic mutation in the *MLH1* gene, which is a known driver of MMR. Moreover, upon comparing these mutational signatures among the three BrM groups, no significant differences were observed (Table 12).

Overall, our findings indicate that the landscapes of small mutations and mutational burden do not distinguish between the various patterns of BrM in our cohort.

4.2.2. Copy number analysis

Copy number alterations (CNAs) occur due to changes to DNA structure that lead to the gain/amplification or loss/deletion of copies of DNA sections from a normal genome.

Our brain metastatic samples displayed typical chromosomal aberrations previously found in the primary NSCLC [189, 190]. Among all combined BrM samples, the most frequently observed arm-level alterations included gains along chromosomes 1q, 5p, 7p, 8q,12p, 14q, 17q, and 20q and losses along 3p, 5q, 8p, 13q, 15p, 15q, 18q and 21q (Figure 17a).

In our BrM-NSCLC samples, we identified high-level amplification (copy number > 5) at seventeen distinct genomic regions in more than five samples, and deletions (copy number ≤ 1) at five regions (Table 13). Some of these alterations were consistent with previous findings in NSCLC studies, including amplifications in *TERT, CCND3, MYC, CCND1, NKX2-1,* and *FOXA1,* as well as deletions in *CDKN2A/B* and *PTEN*. Additionally, specific high-level amplifications were detected in only one sample, including *MET* (copy number =24) and *CCNE1* (copy number =8), or in two samples for genes such as *MDM2* (copy number =9 and 39), or in three samples for *CDK4* (copy number =11, 21 and 26).

Table 13: High-level amplifications and deletions in all BrM-NSCLC samples

Chr.	Cytoband	Cytoband size	Type	No. of samples	% of samples	No. of genes	Frequent reported genes

1	p36.13	600,001 bp	amp	9	14.8%	14	
1	q21.3	300,001 bp	amp	10	16.4%	16	
1	q22	300,001 bp	amp	10	16.4%	26	
1	q23.3	300,001 bp	amp	11	18.0%	11	
4	p11	900,001 bp	amp	16	26.2%	4	
5	p15.33	1,500,001 bp	amp	12	19.7%	40	*TERT*
5	p15.31-p15.2	2,700,001 bp	amp	12	19.7%	26	
5	p13.2	300,001 bp	amp	17	27.9%	6	
6	p21.1	1,200,001 bp	amp	5	8.2%	26	*CCND3*
8	p23.3	300,001 bp	del	11	18.0%	4	
8	p22	300,001 bp	del	9	14.8%	15	
8	q21.13	900,001 bp	amp	8	13.1%	18	
8	q24.21	1,500,001 bp	amp	10	16.4%	17	*MYC*
8	q24.3	300,001 bp	amp	10	16.4%	16	
9	p11.2	300,001 bp	amp	10	16.4%	3	
9	p21.3	900,001 bp	del	15	24.6%	10	*CDKN2A/B*
10	q23.31	300,001 bp	del	5	8.2%	3	*PTEN*
11	q13.3	300,001 bp	amp	6	9.8%	6	*CCND1*
14	q13.2-q21.1	1,800,001 bp.	amp	9	14.8%	24	*NKX2-1, FOXA1*
15	q11.2	600,001 bp	del	11	18.0%	12	
18	p11.32	300,001 bp	amp	15	24.6%	8	
20	p11.1	300,001 bp	amp	13	21.3%	1	

amp: amplification, del: deletion

Furthermore, upon separate analysis of the different BrM groups, we detected significant alterations in ten regions across the genome with p-values < 0.05 (Table 14). However, to refine our findings, we focused on CNAs with p-values < 0.01 (determined by Fisher's exact test). This stricter criterion provided an enhanced overview of the most pronounced alterations. Consequently, we identified significant aberrations in four chromosomal regions that exhibited notable differences between the BrM groups (Table 15). Specifically, a substantial segment on chromosome 4q31.3-q35.2 (including 359 annotated genes) displayed significant distinctions in the oligo-synchronous BrM group, demonstrating losses in half of the cases (50.3%) (Figure 17b). Conversely, the oligo-metachronous BrM group demonstrated a mix of samples with both gains and losses (29% each) (Figure 17c). In contrast, the poly-metastatic group displayed minimal aberrations along this specific chromosomal region (9.7% for gains and 10.9% for losses) (Figure 17d). For contextual comparison, we turned to the Pan-Lung Cancer dataset from TCGA (Nat Genet 2016) and metastatic NSCLC (MSK, Nature Medicine 2022), where we identified three annotated genes—*FAT1*, *IRF2*, and *SFRP2*—that were inherently cancer-relevant. However, alterations in these genes did not exhibit significant differences among the three BrM groups.

Significant differences were also noted in smaller regions of chromosomes. Loss of 5q13.2-q13.3 was primarily observed in oligo-metachronous BrM (70%). a significantly higher proportion compared to oligo-synchronous (38.9%) and poly-metastasis (28.4%) cases. Moreover, the altered pattern of 8 p23.1-p22 was characterized by both loss and gain in the oligo-synchronous BrM group (44.4% and 33.3%, respectively), while both oligo-metachronous and poly-metastasis groups exhibited only loss in this region (55% and 43.5%, respectively). Additionally, a combination of loss and gain was observed in oligo-synchronous BrM for 12p13.2-p12.1 (33.3% and 44.4% respectively), whereas oligo-metachronous BrM showed minor alterations (18.9% loss and 14.5% gain). In contrast, poly-metastasis demonstrated gain in 43.5% of the samples, with loss observed in only 4.3% of samples. Figure 17e shows these four significantly differentiated regions highlighted in dark blue.

In summary, these findings revealed significant variations in the CNA landscapes among the different metastatic patterns, providing novel insights, particularly within the oligo-synchronous BrM group.

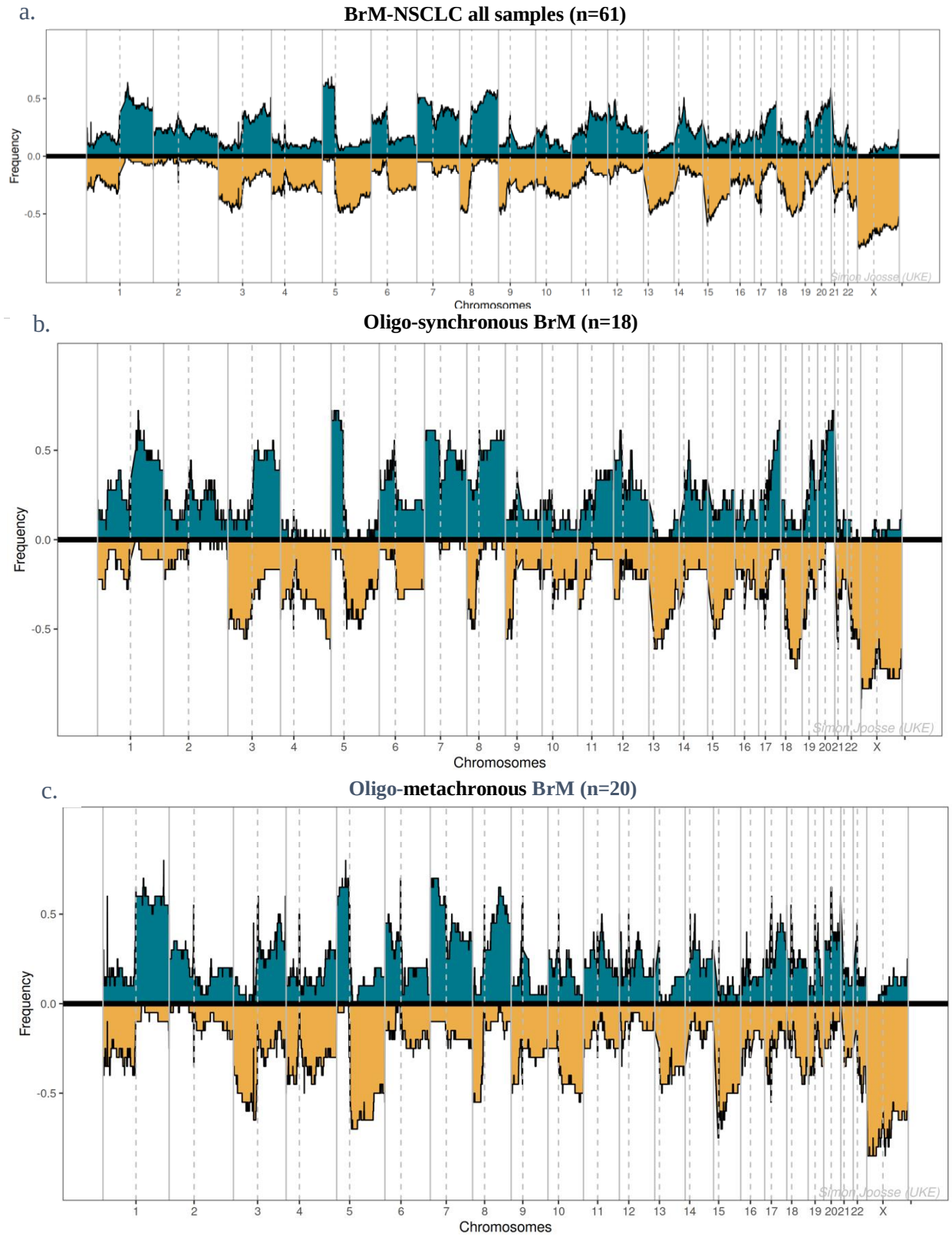
Table 14: Significant CNA regions between BrM groups ($p<0.05$)

Description	Cytoband	Cytoband size	Frequency Sync (n=18)	Frequency Meta (n=20)	Frequency Poly (n=23)	p-value	No. of genes	No. of cancer relevant genes
Loss3p	p25.3-p13	7,787,673 bp	39.6%	43.2%	18.3%	0.039	179	11
Loss4q	q31.21-q35.2	43,774,064 bp	45.6%	30.3%	11.0%	0.010	359	4
Loss5q	q12.1-q33.3	42,900,008 bp	36.9%	62.7%	27.4%	0.032	508	10
Gain7p	p22.3-p14.3	29,369,943 bp	59.6%	66.8%	25.1%	0.019	346	14
Loss8p	p23.3-p21.3	19,737,394 bp	42.5%	54.5%	43.1%	0.017	184	3
Loss10q	q22.1	2,719,729 bp	25.6%	40.0%	34.8%	0.036	36	3
Gain12p	p13.31-p12.1	17,087,074 bp	45.1%	16.4%	44.5%	0.018	296	7
Loss18q	q12.2	1,826,081 bp	66.7%	30.0%	40.6%	0.025	4	0
Loss22q	q13.2-q13.31	6,080,623 bp	59.0%	35.2%	20.5%	0.042	1.53	4

Sync: oligo-synchronous BrM, Meta: oligo-metachronous BrM, Poly: poly-metastases

Table 15: Significant CNA regions between BrM groups (p<0.01)

Description	Cytoband	Cytoband size	Frequency Sync (n=18)	Frequency Meta (n=20)	Frequency Poly (n=23)	p-value	No. of genes	No. of cancer relevant genes
Loss4q	q31.3-q35.2	37,489,761 bp.	50.3%	29.0%	10.9%	0.004	245	3
Loss5q	q13.2-13.3	6,426,767 bp	38.9%	70.0%	28.4%	0.008	87	1
Loss/Gain 8q	p23.1-p22	3,428,575 bp	44.4%/ 33.3%	55%/ 5%	43.5%/ 0%	0.009	24	1
Loss/Gain12p	p13.2-p12.1	11,578,285 bp	33.3%/ 44.4%	18.9%/ 14.5%	4.3%/ 43.5%	0.011	85	3



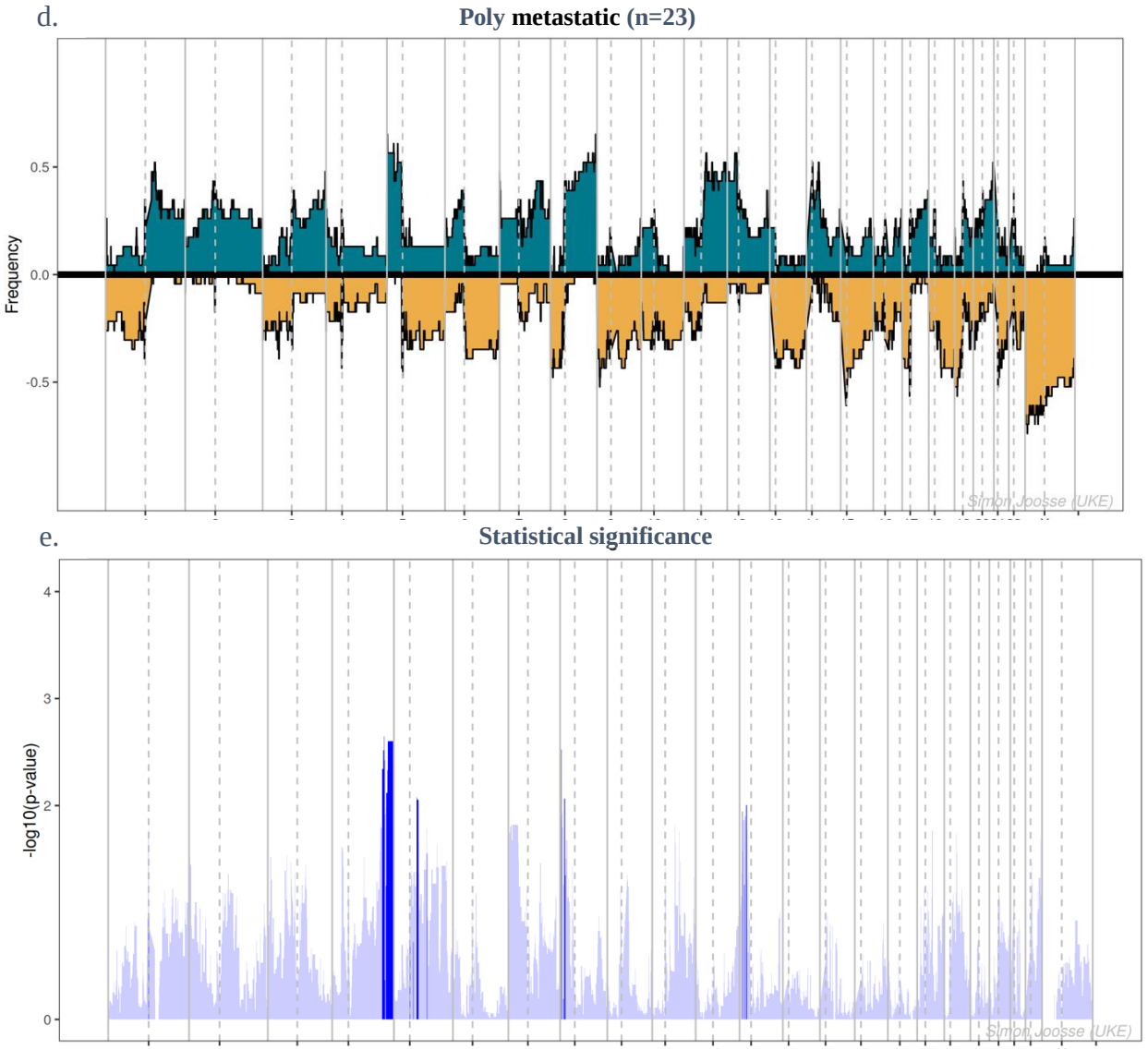


Figure 17: Copy number profiles BrM-NSCLC groups.
Frequency plots of CNAs distribution in (a.) all BrM-NSCLC samples (n=61), (b.) oligo-synchronous BrM group (n=18), (c.) oligo-metachronous BrM group (n=20) and (d.) poly metastatic group (n=23). (Green and yellow indicate gain and loss, respectively). (e.) Statistical significance between BrM groups ($p<0.01$).

4.2.3. The RNA expression profile of BrM

Total RNA was extracted from 40 samples. Good quality RNA samples suitable for NGS were obtained from 28 of these samples (see methods chapter 3.4). The Venn diagram (Figure 18) shows the intersection among significant differentially expressed genes (DEGs) (FDR< 0.05, Wald test). It revealed that there were no common DEGs across the three BrM groups. However, a comparison between the oligo-synchronous and oligo-metachronous groups identified a total of 41 genes that exhibited significant differential expression. Among these genes, 21 were found to be significantly lower and 20 were higher expressed in the oligo-metachronous group compared to the oligo-synchronous group. Furthermore, a comparison between the oligo-metachronous and poly-metastatic samples revealed 105 genes that were

significantly differently expressed, with 64 genes being lower and 41 genes being higher in expression in the poly-metastatic group compared to the oligo-metachronous group. Notably, when comparing the oligo-synchronous group with the poly-metastatic group, a total of 200 genes exhibited significant differential expression, with 111 genes being lower and 89 genes being higher in expression in the poly-metastatic group compared to the oligo-synchronous group. Gene lists of top 40 DEGs from each comparison are listed in supplement tables 1-5 (Appendix). These findings emphasize the presence of substantial molecular distinctions between the oligo BrM groups and the poly metastatic group.

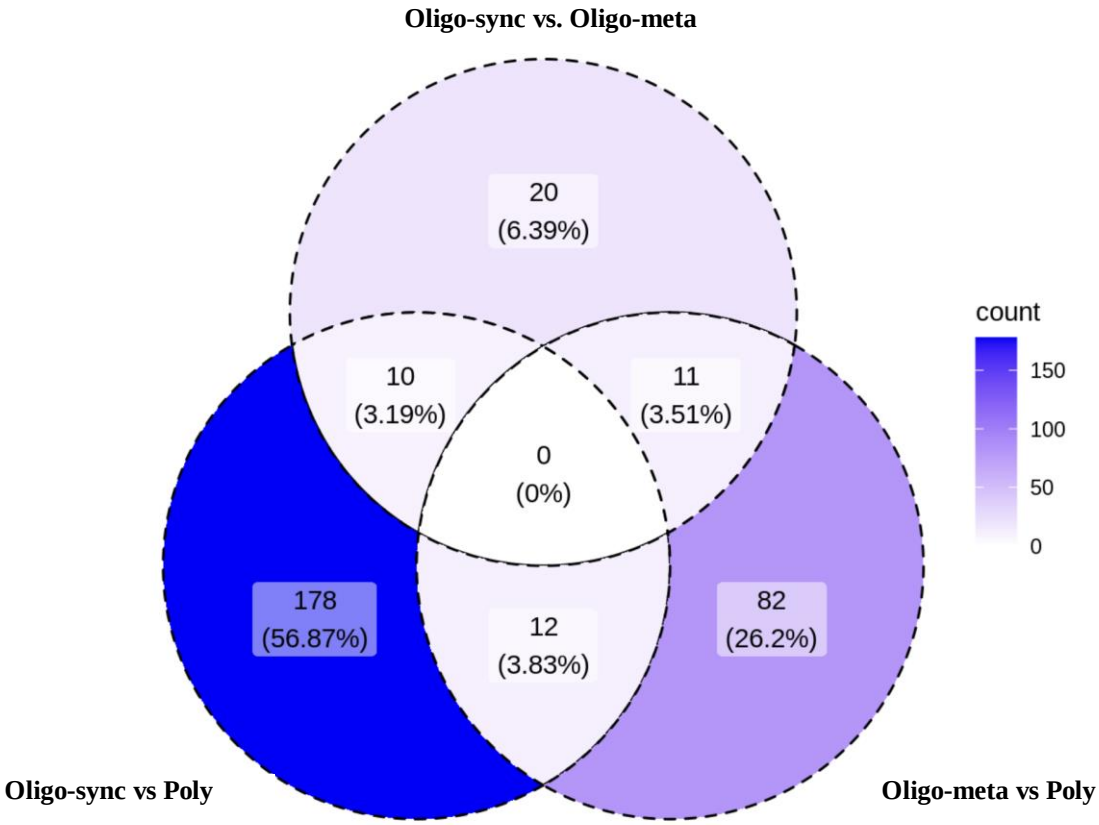


Figure 18: Venn diagram illustrating the intersections of the DEGs from the three BrM groups.
The numbers in overlapping arcs indicate the number of genes shared between BrM groups (oligo-sync: oligo-synchronous BrM, oligo-meta: oligo-metachronous BrM and poly: poly metastasis).

4.2.4. Pathway enrichment analysis

To understand the potential function and biological significance of the resulting genes, various online tools were employed for GO functional annotation, including biological process (BP), molecular function (MF), and cellular component (CC), as well as KEGG pathway enrichment analysis. We mainly used the g:profiler but for validation of main pathways the data was also analyzed by Enricher and DAVID. Through the utilization of these tools, our objective was to

identify pathways and central genes functionally associated with our results. Pathway enrichment analysis was performed pairwise between BrM groups. The chapters below shows the results for g:profiler, mentioning other tools when similar results were obtained.

4.2.4.1. Oligo-synchronous vs. oligo-metachronous BrM groups

GO enrichment analysis (Table 16) revealed significant enrichment of genes related to the regulation of cellular processes and inflammatory response, which are known to be involved in cancer. Additionally, central nervous system myelination was enriched in the BP category, potentially influenced by tumor growth in the brain. Genes were also found to be associated with cytoplasm, vesicles, and axons in the cellular CC category, while in the MF category, enrichment was observed for protein binding (also in DAVID) and extracellular matrix structural constituent (also in DAVID and Enricher).

Table 16: Pathway enrichment analysis of DESs between oligo-synchronous and oligo-metachronous BrM (g:profiler)

Source	Term id	Term name	Adjusted p-value
GO:BP	GO:0050794	regulation of cellular process	2.83E-22
GO:BP	GO:0022010	central nervous system myelination	1.49E-02
GO:BP	GO:0006954	inflammatory response	3.99E-02
GO:BP	GO:1990961	xenobiotic detoxification by transmembrane export across the plasma membrane	4.40E-02
GO:CC	GO:0005737	cytoplasm	5.98343E-24
GO:CC	GO:0031982	vesicle	9.58277E-05
GO:CC	GO:0030424	axon	7.11E-03
GO:CC	GO:0005583	fibrillar collagen trimer	9.89E-03
GO:CC	GO:0005667	transcription regulator complex	3.10E-02
GO:MF	GO:0005515	protein binding	1.68527E-24
GO:MF	GO:0005201	extracellular matrix structural constituent	4.28E-05
GO:MF	GO:0003824	catalytic activity	2.38E-02

4.2.4.2. Oligo-metachronous BrM vs. poly-metastatic groups

Go enrichment analysis (Table 17) using g:profiler and DAVID identified multiple biological processes related to metastasis, including cell migration, cell adhesion, cell motility, regulation of cell migration, and extracellular matrix organization. The immune system process pathway highlighted the interplay between gene expression and the immune system. Interestingly, enrichment in the regulation of response to reactive oxygen species was observed, which corresponds to an observed mutational signature (SBS18) at the DNA level. Moreover, the Enricher tool showed enrichment in the canonical Wnt signaling pathway (GO:0060070) (q-value= 0.028) and Regulation of angiogenesis (GO:0045765) (q-value= 0.043), both of which play roles in cancer progression. In the CC category, genes were primarily involved in the

cytoplasm, receptor complex, and Wnt signalosome (according to g:profiler and DAVID), with Enricher showing involvement in Collagen-Containing Extracellular Matrix (GO:0062023) (q-value = 2.10E-10).

Table 17: Pathway enrichment analysis of DESs between oligo-metachronous BrM and poly-metastasis (g:profiler)

Source	Term id	Term name	Adjusted p-value
GO:BP	GO:0048518*b	positive regulation of biological process	1.90E-26
GO:BP	GO:0032501*bc	multicellular organismal process	1.69E-19
GO:BP	GO:0016477*bc	cell migration	3.27E-11
GO:BP	GO:0030198	extracellular matrix organization	1.12E-10
GO:BP	GO:0007155	cell adhesion	2.97E-10
GO:BP	GO:0019538*b	protein metabolic process	9.74E-10
GO:BP	GO:0048870*bc	cell motility	1.18E-09
GO:BP	GO:0048585*b	negative regulation of response to stimulus	2.86213E-09
GO:BP	GO:0048519*b	negative regulation of biological process	1.90572E-08
GO:BP	GO:0030334*b	regulation of cell migration	4.55002E-08
GO:BP	GO:0023057*b	negative regulation of signaling	6.4459E-08
GO:BP	GO:0001667*bc	ameboidal-type cell migration	1.7872E-07
GO:BP	GO:0010648*b	negative regulation of cell communication	2.31126E-07
GO:BP	GO:2000145*b	regulation of cell motility	2.95E-07
GO:BP	GO:0006066	alcohol metabolic process	1.64141E-06
GO:BP	GO:1901031	regulation of response to reactive oxygen species	0.009820509
GO:BP	GO:0002376	immune system process	0.022825226
GO:BP	GO:0031667	response to nutrient levels	0.031787373
GO:BP	GO:0022010	central nervous system myelination	3.32E-02
GO:BP	GO:0045055	regulated exocytosis	3.69E-02
GO:CC	GO:0005737*b	cytoplasm	3.79022E-33
GO:CC	GO:0043235	receptor complex	9.54651E-05
GO:CC	GO:0030424	axon	0.003425881
GO:CC	GO:1903439	calcitonin family receptor complex	0.035099912
GO:CC	GO:1990909	Wnt signalosome	0.037745849
GO:MF	GO:0005515*bc	protein binding	1.25997E-21
GO:MF	GO:0005201	extracellular matrix structural constituent	9.37634E-08
GO:MF	GO:0050840	extracellular matrix binding	3.04086E-05
GO:MF	GO:0038024	cargo receptor activity	0.000114278
GO:MF	GO:0098772	molecular function regulator activity	0.008366254
GO:MF	GO:0005539	glycosaminoglycan binding	0.008403064
GO:MF	GO:0001730	2'-5'-oligoadenylate synthetase activity	0.014714746
GO:MF	GO:0003824*b	catalytic activity	2.12E-02
GO:MF	GO:0005041	low-density lipoprotein particle receptor activity	2.13E-02
GO:MF	GO:0033218	amide binding	2.67E-02
GO:MF	GO:0030169	low-density lipoprotein particle binding	3.55E-02
GO:MF	GO:0043167*c	ion binding	4.70E-02

*genes of interest: (b) *SMURF2*, (c) *S100A2*

4.2.4.3. Oligo-synchronous BrM vs. poly-metastatic groups

In this comparison, similar enriched pathways were observed for biological processes (BP) and cellular components (CC), as previously mentioned (Table 18). Additionally, KEGG analysis using g:profiler, DAVID, and Enricher showed significant enrichment in pathways related to cancer, Wnt signaling, metabolic pathways, and Proteoglycans in cancer.

Table 18: Pathway enrichment analysis of DESs between oligo-synchronous BrM and poly-metastasis (g:profiler)

Source	Term id	Term name	Adjusted p-value
GO:BP	GO:0050794*ab	regulation of cellular process	1.40E-33
GO:BP	GO:0048518*ab	positive regulation of biological process	3.04E-31
GO:BP	GO:0048522*ab	positive regulation of cellular process	8.31E-25
GO:BP	GO:0007165*b	signal transduction	1.86E-15
GO:BP	GO:0010646*b	regulation of cell communication	5.93E-15
GO:BP	GO:0007166*b	cell surface receptor signaling pathway	1.47E-14
GO:BP	GO:0023051*b	regulation of signaling	5.27E-14
GO:BP	GO:0019538*b	protein metabolic process	2.091E-13
GO:BP	GO:0030334*b	regulation of cell migration	2.09756E-10
GO:BP	GO:0016477*bc	cell migration	7.76901E-10
GO:BP	GO:0048870*bc	cell motility	1.04087E-09
GO:BP	GO:0009987*abc	cellular process	5.88823E-07
GO:BP	GO:0048585*b	negative regulation of response to stimulus	1.56729E-06
GO:BP	GO:0086001	cardiac muscle cell action potential	6.52E-04
GO:BP	GO:0006629*a	lipid metabolic process	0.001284996
GO:BP	GO:1902305	regulation of sodium ion transmembrane transport	0.002453445
GO:BP	GO:0044419	biological process involved in interspecies interaction between organisms	0.004518603
GO:BP	GO:0002221	pattern recognition receptor signaling pathway	0.015885443
GO:BP	GO:1990778	protein localization to cell periphery	2.00E-02
GO:BP	GO:0086003	cardiac muscle cell contraction	3.27E-02
GO:BP	GO:0006941	striated muscle contraction	0.041747685
GO:BP	GO:0036010	protein localization to endosome	0.048686499
GO:CC	GO:0005737*ab	cytoplasm	3.33521E-44
GO:CC	GO:0005667	transcription regulator complex	0.002099935
GO:CC	GO:1990907	beta-catenin-TCF complex	0.008496222
GO:CC	GO:0031528	microvillus membrane	0.026663999
GO:MF	GO:0005515*abc	protein binding	2.39353E-39
GO:MF	GO:0016491*a	oxidoreductase activity	4.89659E-06
GO:MF	GO:0003824*ab	catalytic activity	9.09447E-05
GO:MF	GO:0042802*bc	identical protein binding	0.000857528
GO:MF	GO:0003714	transcription corepressor activity	0.001631819
KEGG	KEGG:05200	Pathways in cancer	8.09627E-05
KEGG	KEGG:04310	Wnt signaling pathway	1.23E-02
KEGG	KEGG:01100*a	Metabolic pathways	2.01E-02
KEGG	KEGG:05205	Proteoglycans in cancer	2.32E-02

* genes of interest: (a) *AKR1C1* (b) *SMURF2* (c) *S100A2*

4.2.4.4. Oligo-synchronous or oligo-metachronous BrM vs. poly-metastatic groups

According to g:profiler results (Table 19), the BP section showed significant enrichment of the developmental process when comparing either of the oligo-BrM groups against the poly-metastatic group. Multiple regulatory pathways, cellular response to stimulus, cell migration, and cell motility were also among the significantly enriched terms. Notably, the interleukin-27-mediated signaling pathway, which plays a role in the immune system, was enriched. Additionally, DAVID software revealed enrichment in angiogenesis and cell-cell adhesion (FDR= 0.014 for both pathways) and the apoptotic process (FDR=0.042).

KEGG analysis indicated involvement in Wnt signaling, metabolic pathways, and pathways in cancer, as revealed by both DAVID and g:profiler. Moreover, Enricher showed enrichment in the Hedgehog signaling pathway (q-value= 0.023) and Hippo signaling pathway (q-value= 0.030).

Table 19: Pathway enrichment analysis of DESs between oligo-synchronous or oligo-metachronous BrM and poly-metastasis (g:profiler)

Source	Term id	Term name	Adjusted p-value
GO:BP	GO:0032502*ab	developmental process	1.08879E-34
GO:BP	GO:0048518*ab	positive regulation of biological process	1.54388E-33
GO:BP	GO:0050794*ab	regulation of cellular process	1.77285E-32
GO:BP	GO:0030154*a	cell differentiation	2.19207E-26
GO:BP	GO:0050896*ab	response to stimulus	1.99435E-25
GO:BP	GO:0065007*ab	biological regulation	2.76847E-24
GO:BP	GO:0051716*ab	cellular response to stimulus	9.39E-20
GO:BP	GO:0007154*b	cell communication	1.0889E-18
GO:BP	GO:0009893*a	positive regulation of metabolic process	1.75065E-16
GO:BP	GO:0016477*bc	cell migration	3.66012E-13
GO:BP	GO:0048870*bc	cell motility	1.82174E-12
GO:BP	GO:2000145*b	regulation of cell motility	2.15E-10
GO:BP	GO:0030198	extracellular matrix organization	9.47E-07
GO:BP	GO:0090257	regulation of muscle system process	0.000527014
GO:BP	GO:0009615	response to virus	0.003743878
GO:BP	GO:0010876*a	lipid localization	0.003780289
GO:BP	GO:0002691	regulation of cellular extravasation	0.00458446
GO:BP	GO:0070106	interleukin-27-mediated signaling pathway	0.004646573
GO:BP	GO:0044419	biological process involved in interspecies interaction between organisms	0.007846168
GO:BP	GO:0072577	endothelial cell apoptotic process	0.013894001
GO:BP	GO:0071359	cellular response to dsRNA	0.022210392
GO:BP	GO:0006996	organelle organization	0.025289328
GO:BP	GO:0097435	supramolecular fiber organization	0.031456183
GO:CC	GO:0005737*ab	cytoplasm	2.19663E-39
GO:CC	GO:0016020*b	membrane	4.74162E-20
GO:CC	GO:0005581	collagen trimer	1.71E-03
GO:CC	GO:0055037	recycling endosome	2.46E-02
GO:MF	GO:0005515*abc	protein binding	1.61E-30

GO:MF	GO:0042802*bc	identical protein binding	7.58E-07
GO:MF	GO:0016491*a	oxidoreductase activity	9.26E-07
GO:MF	GO:0050839	cell adhesion molecule binding	6.95E-03
GO:MF	GO:0004028	3-chloroallyl aldehyde dehydrogenase activity	4.10E-02
GO:MF	GO:0001730	2'-5'-oligoadenylate synthetase activity	4.10E-02
GO:MF	GO:0016829	lyase activity	4.55E-02
KEGG	KEGG:04310	Wnt signaling pathway	6.00E-04
KEGG	KEGG:05205	Proteoglycans in cancer	7.14E-04
KEGG	KEGG:01100*a	Metabolic pathways	9.26E-03
KEGG	KEGG:05200	Pathways in cancer	4.23E-02

* genes of interest: (a) *AKR1C1* (b) *SMURF2* (c) *S100A2*

4.2.4.5. Oligo-synchronous vs oligo-metachronous BrM or poly-metastatic groups

Comparing oligo-synchronous BrM against one of the other groups also showed enrichments in pathways previously mentioned (Table 20).

Notably, in CC and MF categories, cytoplasm and protein binding were enriched in all sets of comparisons. Moreover, cell motility and cell migration were enriched in all comparisons except for oligo-synchronous vs. poly metastasis comparison. Additionally, significant enrichment in extracellular matrix-related pathways and different regulatory pathways were observed in almost all comparisons. These pathways can indicate different contributions to the spread of cancer cells in different metastatic patterns. Moreover, metabolic pathways were also enriched especially in the comparison between oligo-synchronous and poly metastatic groups.

Table 20: Pathway enrichment analysis of DESs between oligo-synchronous and oligo-metachronous BrM or poly-metastasis (g:profiler)

Source	Term id	Term name	Adjusted p-value
GO:BP	GO:0050794*ab	regulation of cellular process	1.48E-22
GO:BP	GO:0065007*ab	biological regulation	3.85E-19
GO:BP	GO:0007154*b	cell communication	4.44454E-15
GO:BP	GO:0050896*ab	response to stimulus	1.50263E-14
GO:BP	GO:0051716*b	cellular response to stimulus	3.28853E-14
GO:BP	GO:0048523*b	negative regulation of cellular process	1.73076E-13
GO:BP	GO:0023051*b	regulation of signaling	7.56453E-12
GO:BP	GO:0009893*a	positive regulation of metabolic process	3.94033E-11
GO:BP	GO:0031323*ab	regulation of cellular metabolic process	7.84E-11
GO:BP	GO:0019538*b	protein metabolic process	9.90112E-11
GO:BP	GO:0007165*b	signal transduction	1.39014E-10
GO:BP	GO:0019222*ab	regulation of metabolic process	5.74286E-08
GO:BP	GO:0031325*a	positive regulation of cellular metabolic process	4.16818E-07
GO:BP	GO:0016477*bc	cell migration	1.27E-05
GO:BP	GO:0048870*bc	cell motility	4.77E-05
GO:BP	GO:0044248	cellular catabolic process	0.000641202
GO:BP	GO:0002028	regulation of sodium ion transport	0.004440799
GO:BP	GO:0086001	cardiac muscle cell action potential	0.008624692
GO:CC	GO:0005737*ab	cytoplasm	3.80608E-33

GO:CC	GO:1990907	beta-catenin-TCF complex	0.004399878
GO:CC	GO:0045334	clathrin-coated endocytic vesicle	0.01473982
GO:CC	GO:0097728	9+0 motile cilium	0.048596878
GO:MF	GO:0005515*abc	protein binding	1.97E-30
GO:MF	GO:0003824*ab	catalytic activity	2.17E-04
GO:MF	GO:0070566	adenylyltransferase activity	8.29E-04
GO:MF	GO:0043167*ac	ion binding	9.91E-04
GO:MF	GO:0003714	transcription corepressor activity	1.03E-02
KEGG	KEGG:04360	Axon guidance	0.00389482
KEGG	KEGG:04390	Hippo signaling pathway	0.009103414
KEGG	KEGG:05200	Pathways in cancer	0.042854973

* genes of interest: (a) *AKR1C1* (b) *SMURF2* (c) *S100A2*

4.2.5. Candidate gene selection

In this study, we employed a large screening approach to identify potential candidate genes associated with the investigated BrM phenotypes in NSCLC. Due to the limitations of a small sample size, it is crucial to acknowledge the possibility of chance findings in such exploratory analyses. Therefore, validation of these candidate genes becomes imperative to ascertain the robustness and reliability of our findings.

First, in order to narrow down the potential target genes in the significant CNA regions, we combined the CNA results with expression data. To accomplish this, we utilized the UCSC Genome Browser database to download the gene annotations corresponding to the identified CNA regions. Subsequently, we analyzed the RNA expression levels of the genes located exclusively within these specific CNA regions. As a result, 15 significant differently expressed genes at three chromosomal locations were only found in the oligo-synchronous BrM vs. poly metastasis comparison (FDR< 0.05), while the other comparisons showed no significance. To refine our selection further, we applied a log2 fold change cutoff |log2FC|≥2, focusing on the most significantly differentially expressed genes (Table 21 highlighted in green). The final selection of two candidate genes was based on pertinent literature.

Table 21: DEG list resulted from integrating CNA and RNA-seq data

Oligo-synchronous vs. Poly-metastasis			
Genes with (p<0.05)	Log2FC	FDR (RNA)	Position
TLR3	2.32	9.29E-04	chr4 (q35.1)
CPE	2.49	7.75E-03	chr4 (q32.3)
TRIM61	2.61	3.20E-02	chr4 (q32.3)
DDX60	1.49	3.92E-02	chr4 (q32.3)
KLHL2	0.91	3.20E-02	chr4 (q32.3)
SNX25	1.24	4.61E-02	chr4 (q35.1)
C4orf47	1.30	3.73E-02	chr4 (q35.1)
GCNT4	2.26	3.20E-02	chr5 (q13.3)
CERT1	0.83	3.20E-02	chr5 (q13.3)

PDE3A	-2.36	3.71E-02	chr12 (p12.2)
RERG	1.94	3.92E-02	chr12 (p12.3)
RP11-174G6.5	1.38	3.71E-02	chr12 (p12.3)
GPRC5D	1.51	4.61E-02	chr12 (p13.1)
EMP1	1.65	3.20E-02	chr12 (p13.1)
BCL2L14	4.13	9.60E-05	chr12 (p13.2)

The first candidate gene identified was *TLR3* (Toll Like Receptor 3) located on Chr. 4q35.1. *TLR3* belongs to the Toll-like receptor (TLR) family, which plays a crucial role in antimicrobial proinflammatory immune responses by inducing the secretion of cytokines and chemokines [191]. *TLR3* showed downregulation at the RNA level, simultaneously to a deletion at DNA level in the oligo-synchronous compared to poly-metastatic group (FDR< 0.001) (Figure 19). *TLR3* was significantly differential expressed also in the whole transcriptome analysis (p= 0.003, Wilcoxon test).

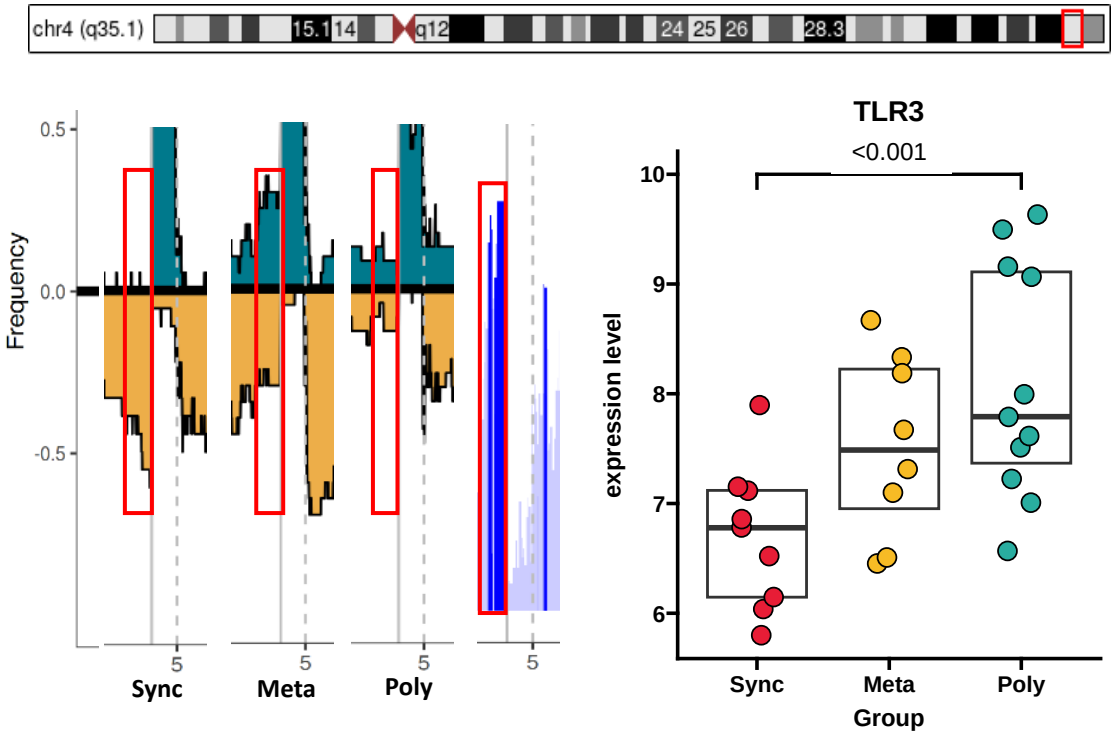


Figure 19: On the left side: a magnified view of the copy number alteration (CNA) frequency plot is shown, focusing on the region Chr4:(q35.1) where the *TLR3* gene is located, Indicating a loss of this region in the oligo-synchronous BrM in comparison to poly metastasis. On the right side, a box plot illustrates the RNA expression levels of *TLR3* in the three groups of brain metastasis (Sync: oligo-synchronous, Meta: oligo-metachronous, Poly: poly metastasis) (FDR<0.001).

On Chr.12p12.2, *PDE3A* (Phosphodiesterase 3A) was chosen as the second candidate gene for validation based on relevant literature connecting its enzymatic functions to cancer progression and metastasis [192]. In our study, *PDE3A* exhibited upregulation in oligo-synchronous BrM where gain was observed in 44.4% and loss in 33.3% of cases in the corresponding region of chr.12. This increased expression was significantly higher than that observed in poly-metastatic

group (FDR= 0.037) (Figure 20). This marked variation in *PDE3A* expression further underscores its potential relevance within the context of the complex molecular landscape of BrM.

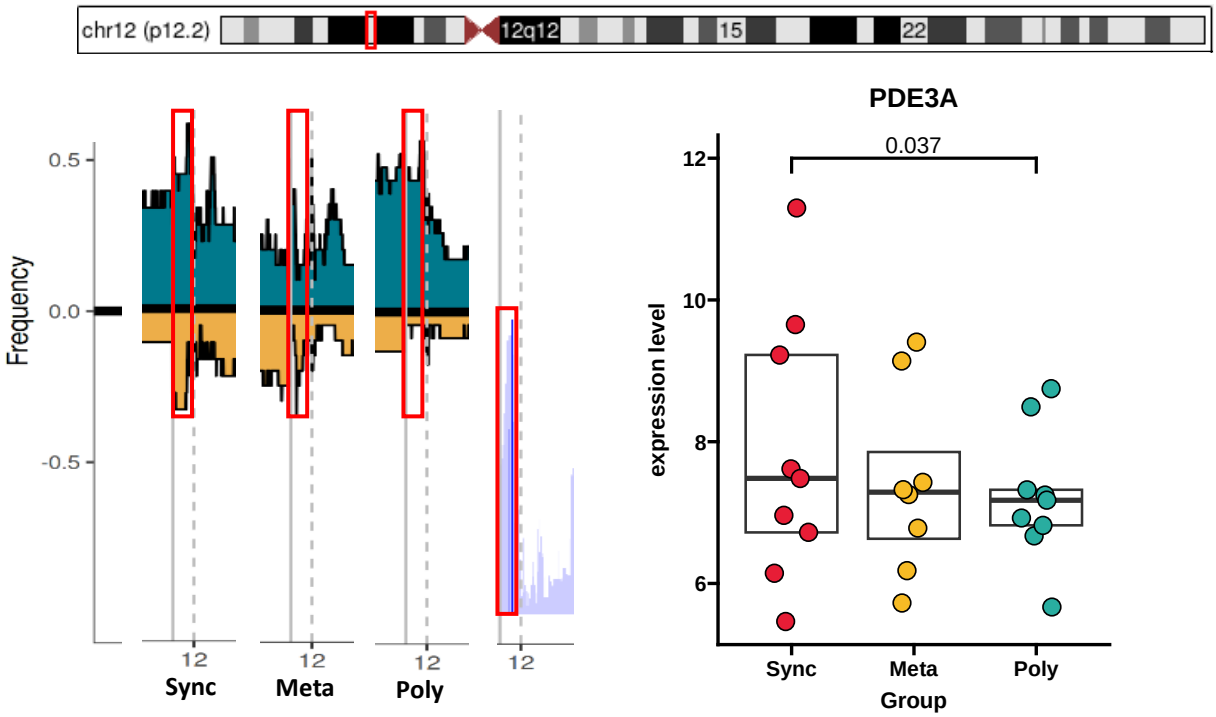


Figure 20: On the left side: a magnified view of the copy number alteration (CNA) frequency plot is shown, focusing on the region Chr12:(p12.2) where the *PDE3A* gene is located. On the right side, a box plot illustrates the RNA expression levels of *PDE3A* in the three groups of brain metastasis (Sync: oligo-synchronous, Meta: oligo-metachronous, Poly: poly metastasis) (FDR=0.037).

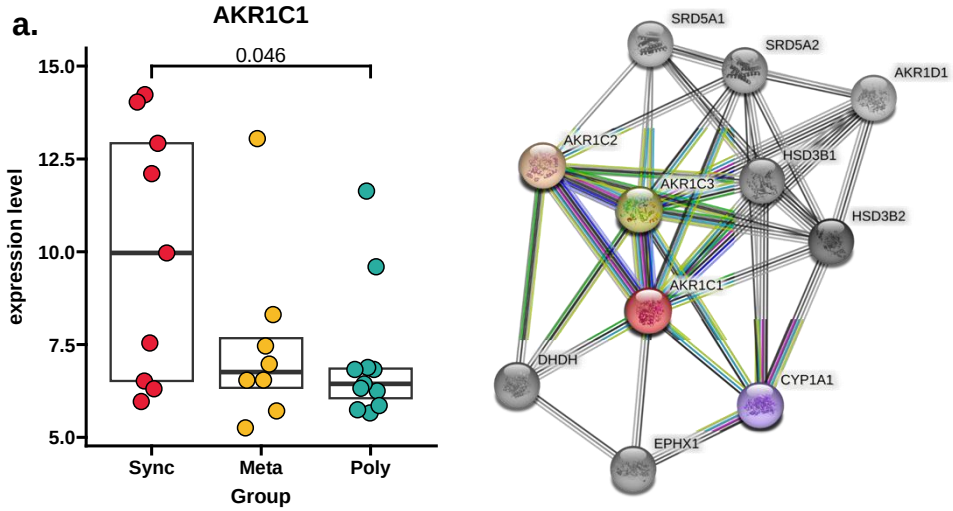
Among the numerous differently expressed genes evaluated with pathway enrichment analysis through whole transcriptome RNA-seq data examination, three target genes warranted particular attention. These genes are *AKR1C1* (Aldo-keto reductase family 1 member C1), *SMURF2* (SMAD Specific E3 Ubiquitin Protein Ligase 2) and *S100A2* (S100 calcium-binding protein A2) stood out in at least two of the mentioned analysis tools (marked with Asterisks ”*“ in tables 16-20). These genes exhibited differential expression patterns among BrM groups and their involvement in specific pathways suggests their potential relevance to the metastatic process.

AKR1C1 showed frequent enrichment in various pathways in our analysis (see chapter 4.2.4), including metabolic pathways like lipid metabolic pathway and NADP+ 1-oxidoreductase activity. Additionally, *AKR1C1* was associated with regulatory pathways such as regulation of the metabolic process, positive regulation of cellular and biological processes and regulations, in addition to response to stimulus, cell differentiation, ion binding, and developmental process. Among the BrM groups, *AKR1C1* expression was significantly upregulated in oligo-

synchronous BrM compared to poly metastatic group (p= 0.046, Wilcoxon test) (Figure 21a). Moreover, *AKR1C1* was one of the most significant DEGs in oligo-synchronous or oligo-metachronous vs poly metastasis comparison (adj.p< 0.001, Wald test). According to the existing literature, *AKR1C1* has shown involvement in cancer development and metastasis in different cancer types [193-195], and contribution to cancer chemotherapeutic drug resistance [196-198]. Hence, we selected *AKR1C1* as a potential candidate for validation.

SMURF2 was another interesting gene due to its enrichment in cell migration, ameboidal-type cell migration, cell motility, as well as its involvement in regulatory pathways such as the regulation of cell migration, cell motility, cellular, metabolic, and biological processes, cell communication, and signaling. Additionally, *SMURF2* was implicated in metabolic pathways specifically protein metabolic process. In our RNA expression data, *SMURF2* showed significant upregulation in the poly metastatic group compared to oligo-synchronous BrM (p= 0.0002, Wilcoxon test) and oligo-metachronous (p= 0.015, Wilcoxon test) (Figure 21b).

S100A2, similar to *SMURF2*, was enriched in cell migration and cell motility pathways, and it was also one of the most significant DEGs between oligo-metachronous and poly metastasis groups (adj.p= 0.001, Wald test). In our findings, *S100A2* exhibited a significant upregulation in the poly metastatic group compared to the oligo-metachronous BrM group (p=0.012, Wilcoxon test) (Figure 21c). *S100A2* captured our attention not only due to its observation in various malignancies, but also because of its close association with the tumor immune microenvironment.



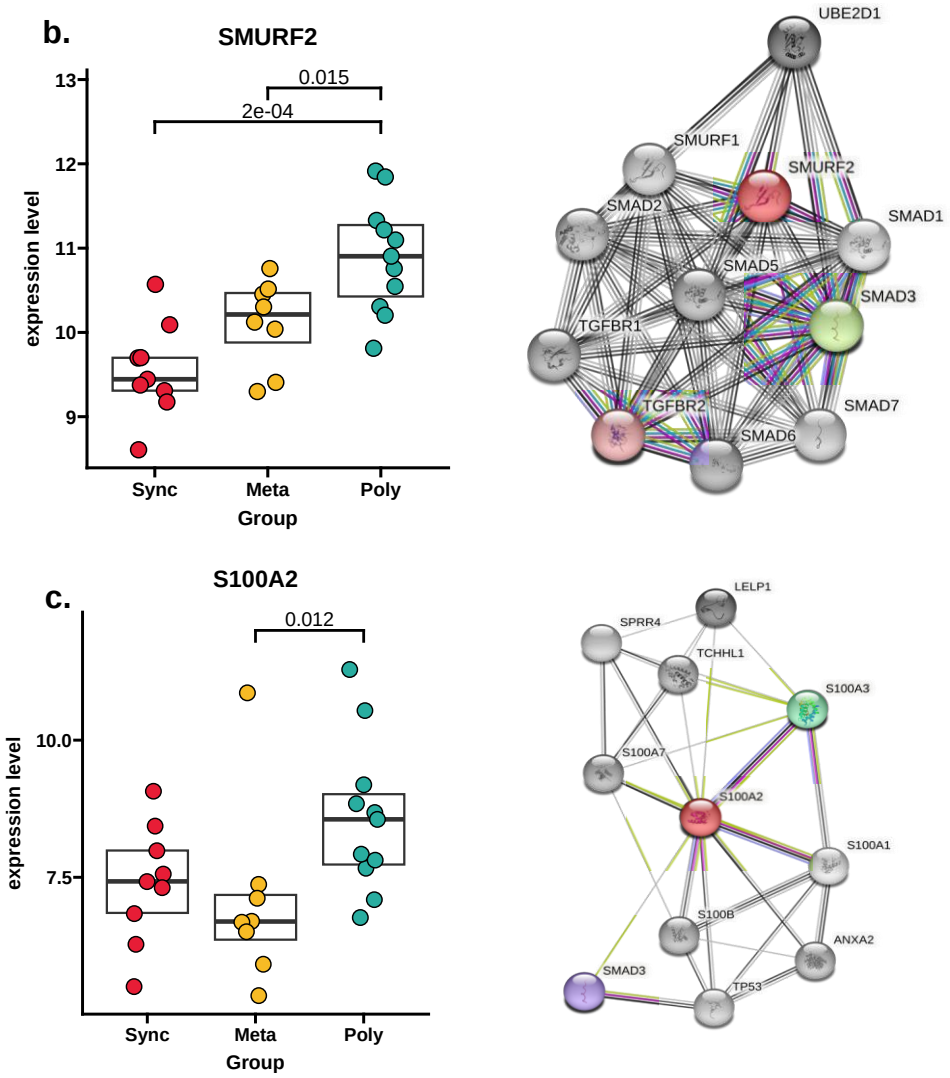


Figure 21: Boxplots of (a) *AKR1C1*, (b) *SMURF2*, and (c) *S100A2* expression levels in RNA-seq results with statistical significance (p-values). On the right, protein-protein interaction (PPI) networks of the three candidate genes generated by STRING tool, highlighting the genes present in our gene lists.

Interestingly, a high expression level of *AKR1C1* exhibited a significant positive correlation with overall survival rate (p= 0.01, log-rank test) (Figure 22). Moreover, the cox proportional hazard regression analysis revealed that low expression levels of *AKR1C1* substantially increased the risk of death compared to high expression (HR: 11.557, 95% CI: 1.802-74.083, p=0.009; Cox proportional hazard ratio). However, no significant correlation was observed between *TLR3*, *PDE3A*, *SMURF2*, or *S100A2* expression and overall survival (p>0.05, log-rank test).

In summary, genetic, genomic, and transcriptomic analyses have revealed that different metastatic patterns exhibited differences on large-scale alteration landscapes but not on single mutation levels. And on these CNA regions we could identify genes that have different expression on RNA level.

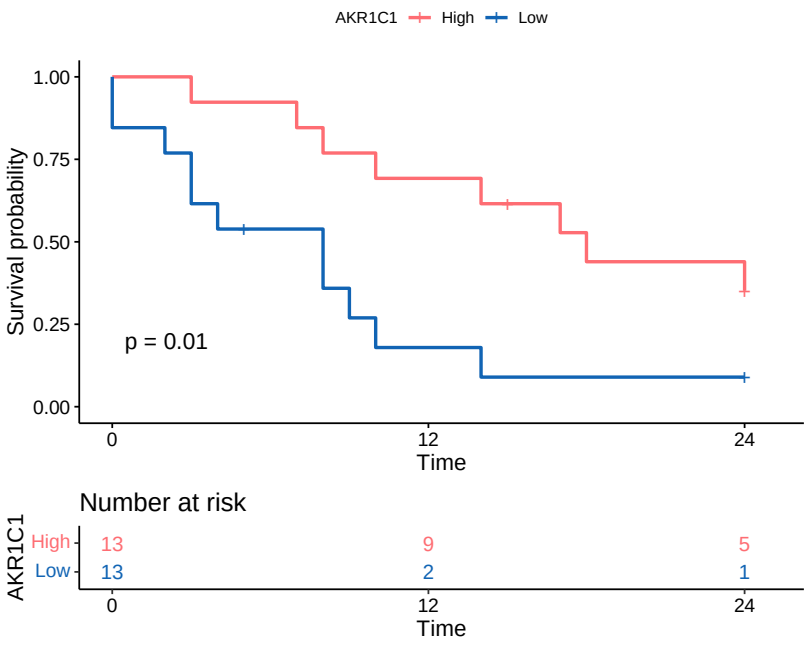


Figure 22: Kaplan-Meier curve of *AKR1C1* expression (High: expression>median, low: expression<median)

4.3. Immuno-profiling of NSCLC brain metastasis

To investigate the immune profile associated with BrM in NSCLC patients, we evaluated the percentages of infiltrating immune cells expressing CD3, CD8, CD4, FOXP3, and CD68 in peri-tumoral and intra-tumoral regions of BrM tissues obtained from 56 NSCLC patients (60 samples were used for FOXP3 analysis). Additionally, we assessed tumor cell proliferation using Ki67 staining (Figure 23a and c).

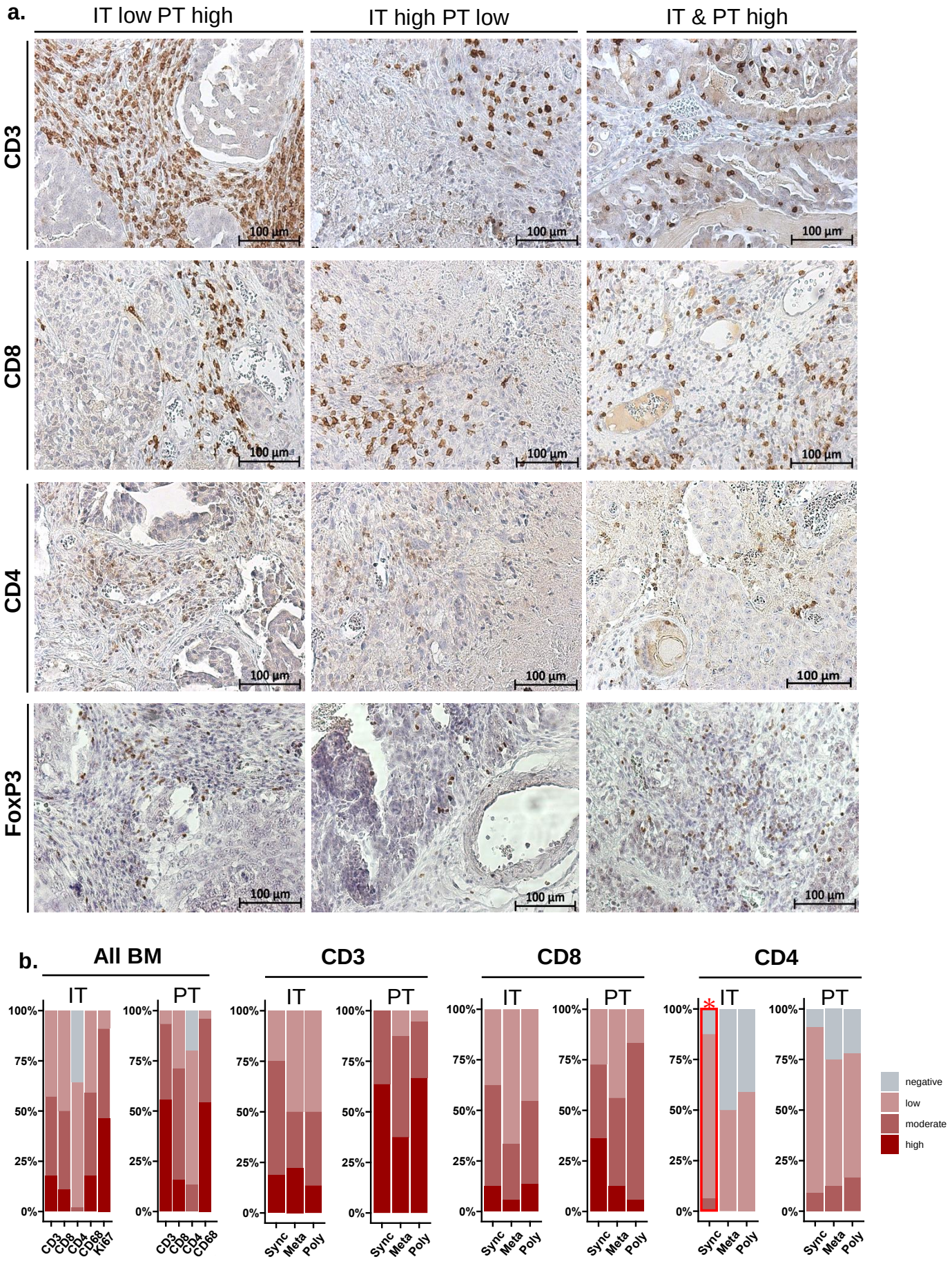
Our study revealed a significantly higher frequency of $CD3^+$ T-cell infiltration and $CD68^+$ microglia/macrophages in peri-tumoral lesions (93.3% and 95.7%, respectively) compared to intra-tumoral regions (57.1% and 58.9%) ($p< 0.0001$ for both markers, Fisher's exact test). Moreover, within $CD3^+$ T-cell population, the percentage of $CD8^+$ cells was significantly higher (>10% positively stained) in both intra-tumoral and peri-tumoral regions (50% and 71.1% respectively) compared to $CD4^+$ cells (1.8% and 13.3%, respectively) with $p< 0.0001$ in both markers (McNemar's test). Notably, the average proliferation rate of the tumors was 44% (range 5-90%), and no correlation was found between the proliferation rate and the immune cell profiles (Figure 23b). These findings suggest that various immune cell types can infiltrate the brain, with a higher prevalence observed in the peri-tumoral compared to intra-tumoral area.

Further analysis was conducted to compare immune cell infiltrations in oligo- and poly-metastatic cases. No significant differences were observed in the number of $CD3^+$ or $CD8^+$ cells between the groups in either region (Figure 23b). However, in oligo-synchronous BrM group,

there was a significantly higher infiltration of $CD4^+$ cells in the intra-tumoral region compared to oligo-metachronous BrM and poly metastasis groups (p= 0.044, G-test). Specifically, 87.5% of oligo-synchronous cases showed $CD4^+$ T-cell infiltration within the tumor tissues, while only 60% in oligo-metachronous and 50% in poly metastasis cases had detectable $CD4^+$ T-cells in this region (Figure 23b). On the other hand, there were no significant differences observed in the proportions of microglia/macrophages ($CD68^+$ cells), or $Ki67^+$ tumor cells among the three BrM groups (Figure 23d).

Given the high presence of $CD4^+$ T-cells in the oligo-synchronous BrM, the frequency of regulatory T-cells (Tregs) defined by FOXP3 expression was further investigated (Figure 23a). Statistical comparisons between the different BrM groups indicated no significant differences in the percentage distribution of positive or negative FOXP3 expression within Tregs, both in the intra-tumoral or peri-tumoral region (p= 0.524 and 0.831, respectively, Fisher's exact test; Figure 23d). However, frequency of $FOXP3^+$ infiltrating cells in intra-tumoral area exhibited a positive correlation with overall survival (p= 0.045, log-rank test). In contrast, this correlation was not observed in peri-tumoral area (p= 0.600, log-rank test) (Figure 23e). In multivariable analysis, the presence of intra-tumoral FOXP3-expressing cells was independently associated with survival (HR: 0.415, 95% CI: 0.210-0.821, p= 0.011; Cox proportional hazard ratio). No significant differences in survival were observed related to the number of CD3, CD8, CD4, or CD68 cells in either the intra-tumoral or peri-tumoral regions (p>0.05, log-rank test, Figure 24).

Taken together, these data indicate that oligo-synchronous BrM exhibits a distinct immune landscape in the TME compared to other metastatic patterns, highlighting the potential significance of intra-tumoral CD4+ T-cell infiltration in the context of BrM in NSCLC.


a.
IT low PT high
IT high PT low
IT & PT high
CD3
CD8
CD4
FoxP3
100 µm
b.
All BM
CD3
CD8
CD4
IT
PT
100%
75%
50%
25%
0%
CD3
CD8
CD4
CD68
Ki67
Sync
Meta
Poly
negative
low
moderate
high

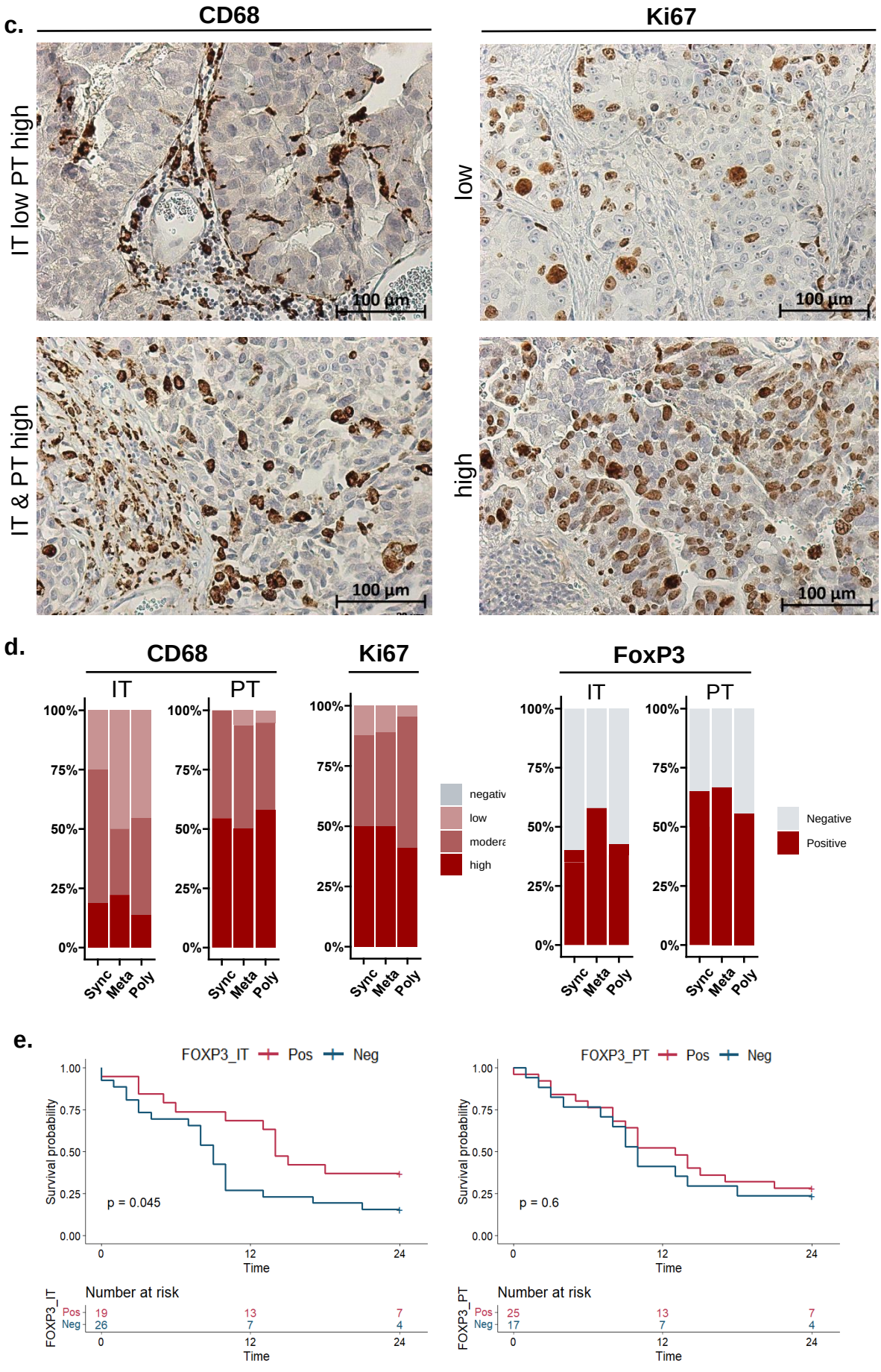


Figure 23: Immunohistochemical detection and estimation of positive staining for CD3, CD8, CD4, and

FOXP3 on tumor-infiltration lymphocytes (TILs) in brain metastasis (BrM) tissues. The staining is shown for both intra-tumoral (IT) and peri-tumoral (PT) regions.
Representative IHC staining of CD3, CD8, CD4, and FoxP3 respectively in IT and PT regions (high and low infiltration) Magnification: x20. (b.) Comparison of staining scores for CD3, CD8, CD4, CD68 and Ki67 across all groups combined. Additionally, a comparison of the staining scores for CD3, CD8, and CD4 between oligo-synchronous BrM (Sync), oligo-metachronous BrM (Meta), and poly BrM (Poly) individually in both IT and PT regions. (c.) Representative IHC staining of CD68 (in IT and PT regions) and Ki67 on BrM tissues (high and low infiltration) Magnification: x20. (d.) Comparison of the staining scores for CD68 in both IT and PT regions, and Ki67 in tumor tissues between BrM groups. Comparison of the staining scores for FOXP3 (negative and positive) between BrM groups in IT and PT regions. (e.) Kaplan-Meier survival curves (24m) of FOXP3 in IT and PT regions. Sync: oligo-synchronous BrM, Meta: oligo-metachronous BrM, and poly: poly metastasis), (negative: no stained cells, low: <10%, moderate:10-40%, and high: >40%.

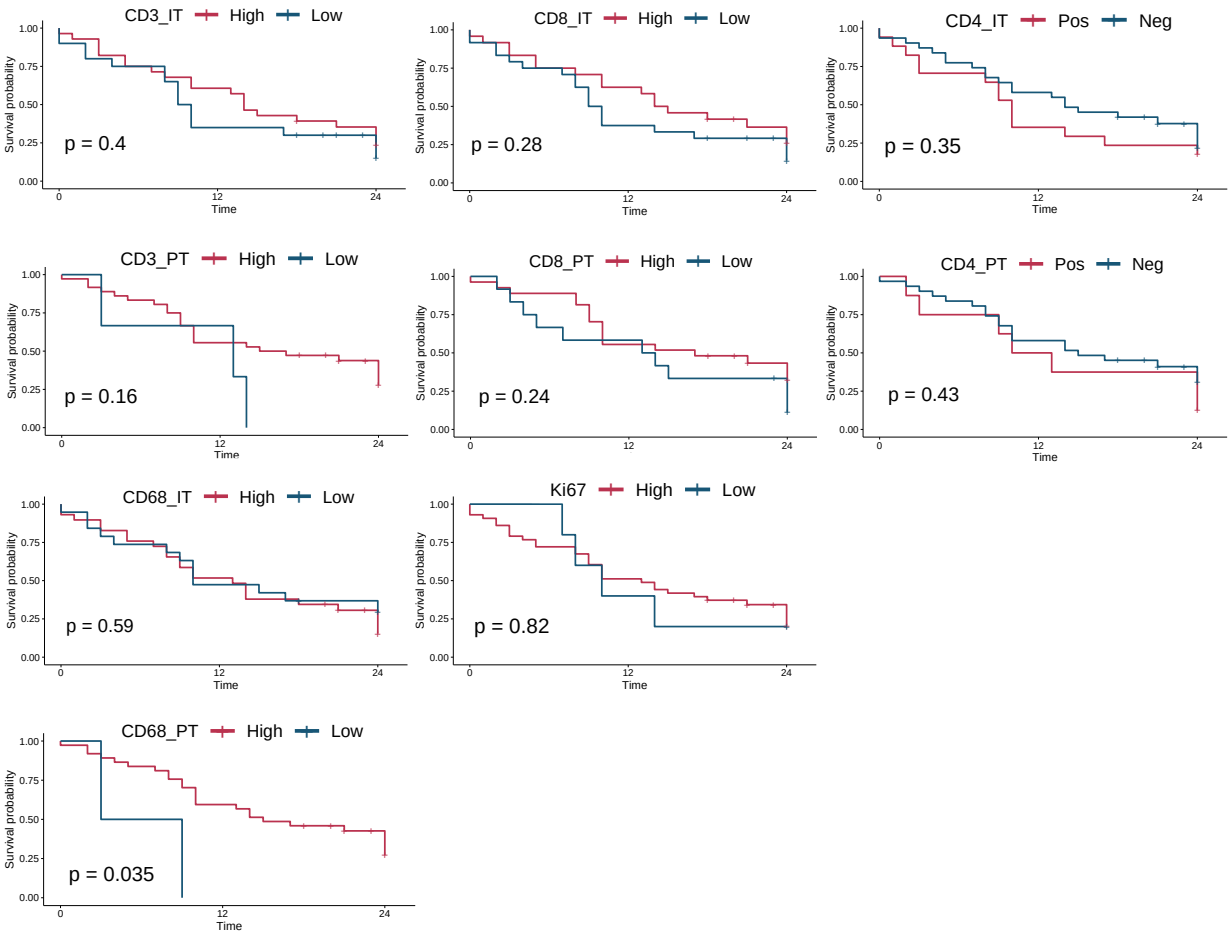


Figure 24: Kaplan-Meier survival curves (24m) of CD3, CD8, CD4 and CD68 in intra-tumoral (IT) and peri-tumoral (PT) regions, and Ki67.

4.4. Peripheral blood immune profiles of BrM-NSCLC groups

Flow cytometry analyses were performed on peripheral blood samples obtained from patients with BrM-NSCLC to assess the T- and NK cell immunophenotypes within the mononuclear cell fractions. We utilized five different multicolor antibody staining cocktails to investigate the profiles of NK and T-cells, focusing on specific aspects such as T-cell exhaustion, T-cell differentiation subtypes, T helper cell subsets, T-cell metabolism, and cytokine secretion.

When comparing the different metastatic phenotypes, significant findings were observed specifically in the T-cell subtypes differentiation based on their activation status. Following established flow cytometry guidelines from Cossarizza et al., frequencies of distinct stages of T-cell development were assessed. These stages include: Naïve: (CD45RA+, CD27+, CD28+, CCR7+), TeffRA+ (T effectors type RA+: CD45RA+, CD27−, CD28−, CCR7−), Early memory (Early: CD45RA-, CD27+, CD28+, CCR7-), Early-like memory (Early-like: CD45RA-, CD27-, CD28+, CCR7+), Intermediate (CD45RA-, CD27+, CD28-, CCR7-), Central memory: (CD45RA−, CD27+, CD28+, CCR7+), and TeffRA- (effector memory cells: CD45RA−, CD27−, CD28−, CCR7−) [199].

Supervised analysis revealed noticeable differences in $CD4^+$ T-cell subsets, while no differences were observed in $CD8^+$ T-cells ($p > 0.05$ in all subsets, Wilcoxon test) (Figure 25a). Specifically, oligo-synchronous patients exhibited a higher median percentage of CD4 naïve cells compared to the oligo-metachronous group ($p = 0.011$, Wilcoxon test). Moreover, the oligo-synchronous BrM group demonstrated a lower percentage of $CD4^+$ effector type RA+ cells ($Teff_{RA+}$) compared to both oligo-metachronous ($p = 0.052$, Wilcoxon test) and poly brain metastatic patients ($p = 0.019$, Wilcoxon test). Similarly, the median percentage of $CD4^+$ effector type RA- cells ($Teff_{RA-}$), also known as effector memory cells, was significantly lower in oligo-synchronous patients compared to the other two groups ($p = 0.021$, and $p = 0.002$, respectively, Wilcoxon test). In contrast, no noteworthy distinctions were observed among the three brain metastatic groups concerning CD4 early, early-like, intermediate, or central memory cells (Figure 25b).

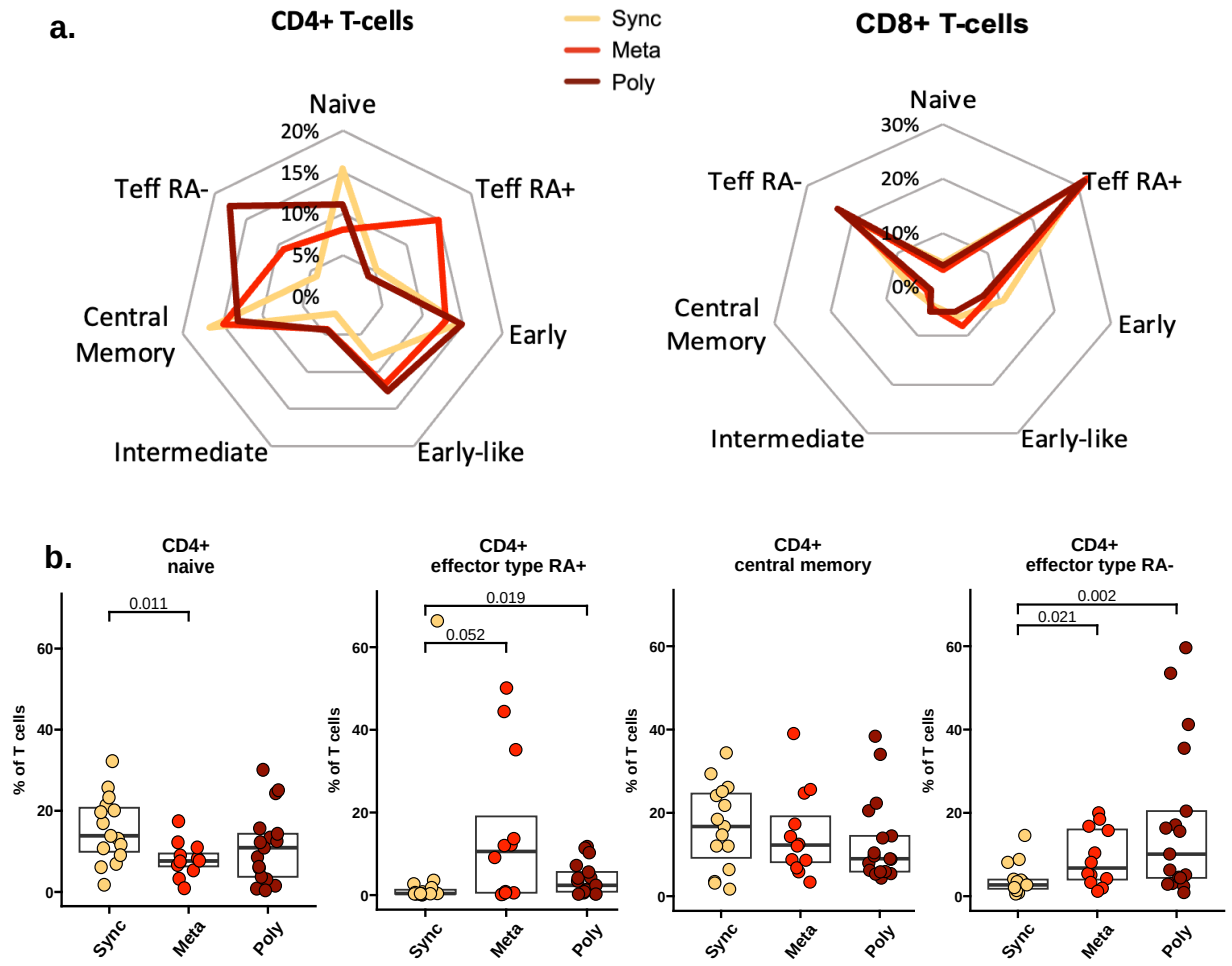


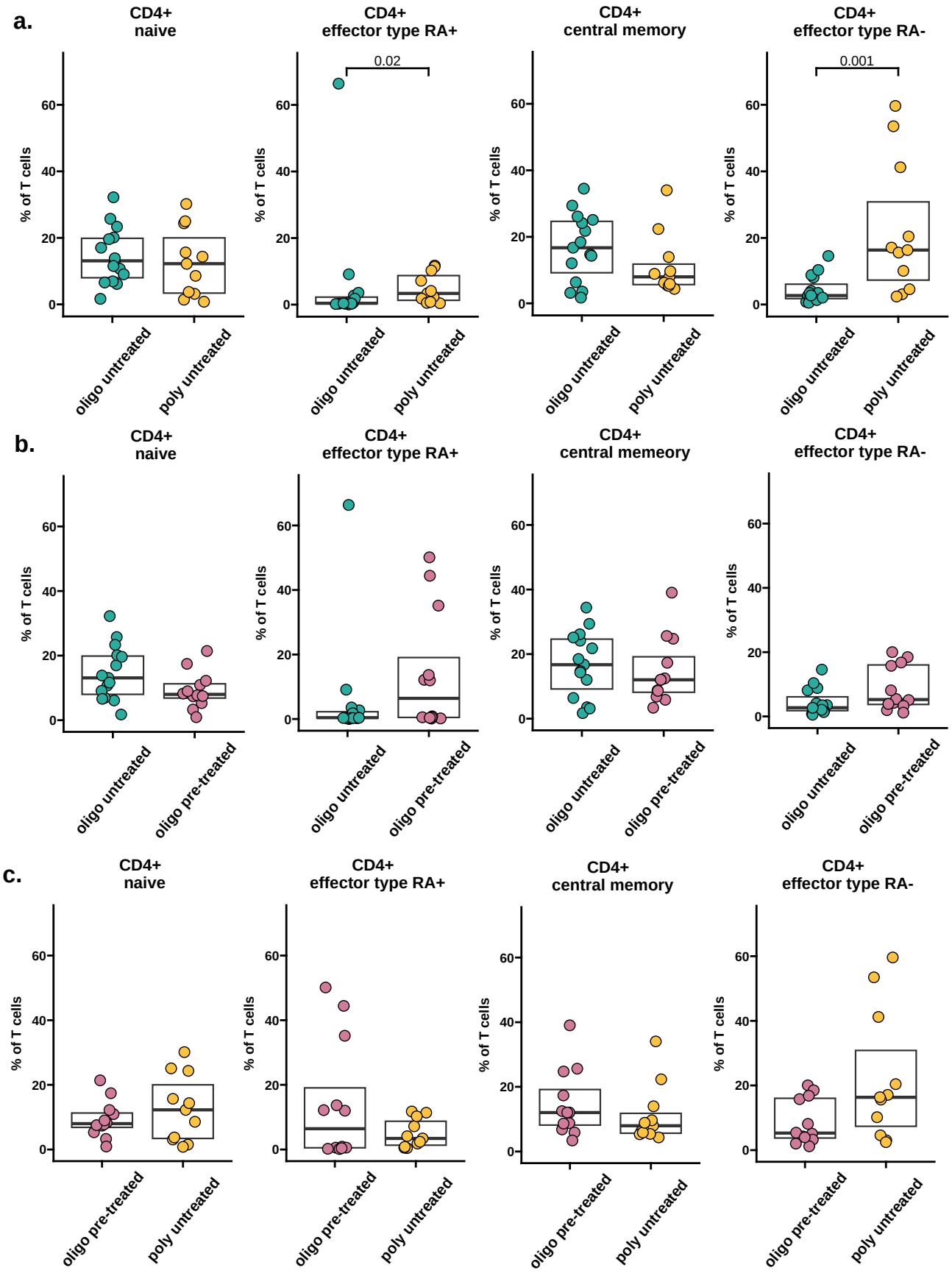
Figure 25: Flow cytometric immunophenotyping of T-cells differentiation

(a.) Spider plots of $CD4^+$ and $CD8^+$ differentiation phenotypes of peripheral blood in BM groups. Populations are defined by expressions of CD45RA, CD27, CD28, and CCR7 markers. Naïve: (CD45RA+, CD27+, CD28+, CCR7+), $Teff_{RA+:}$ T effectors type RA+ (CD45RA+, CD27−, CD28−, CCR7−), Early (CD45RA-, CD27+, CD28+, CCR7-), Early-like (CD45RA-, CD27-, CD28+, CCR7+), Intermediate (CD45RA-, CD27+, CD28-, CCR7-), Central Memory: (CD45RA−, CD27+, CD28+, CCR7+), and $Teff_{RA-}$: effector memory cells (CD45RA−, CD27−, CD28−, CCR7−). (b.) Box plots depicting expression of $CD4^+$ naïve, $Teff_{RA+}$, central memory and $Teff_{RA-}$ expressions between BM groups. Sync: oligo-synchronous BrM, Meta: oligo-metachronous BrM and Poly: poly metastasis.

On the other hand, no significant differences were observed in the T-cell exhaustion, T helper cell subsets, T-cell metabolism, or cytokine secretion profiles between BrM-NSCLC groups.

To investigate the potential impact of pre-treatment on the T-cell differentiation profiles of the BrM groups, a comparison was conducted between untreated samples from the oligo-synchronous and poly-metastatic groups. The results of this comparison confirmed the persistence of the significant differences observed earlier, indicating that treatment did not influence the T-cell differentiation patterns between these BrM groups. However, when untreated oligo-synchronous samples were compared to pre-treated oligo-metachronous samples, the previously observed significant differences were no longer evident. (Figure 26).

This finding reinforces the notion that treatment is not the underlying factor contributing to the observed differences among the BrM groups.



Wilcoxon test (p-value)	naive	Teff type RA+	Central memory	Teff type RA-
oligo_untreated vs poly_untreated	ns	0.02	ns	0.001
oligo_untreated vs oligo_pretreated	0.113	0.152	ns	0.059
poly_untreated vs oligo_pretreated	ns	ns	ns	ns

Figure 26: Treatment impact on T-cell functional differentiation profiles
(a.) comparing untreated samples from oligo- and poly-BrM groups did not alter the significant differences observed between the groups. (b.) comparing untreated samples only with pre-treated samples only in oligo-BrM groups causes loss of all significant differences. (c.) comparing pre-treated oligo-BrM samples with untreated poly metastasis did not show any significant differences.

In this study, we also evaluated the expression of the surface ectonucleotidases CD73 and CD39 in BrM-NSCLC patients in both $CD4^+$ and $CD8^+$ T-cells using the same flow cytometry panel that assessed T-cell differentiation. Interestingly, a significant mean difference in CD73 expression was observed in $CD4^+$ cells between the oligo-synchronous BrM and poly metastasis groups ($p= 0.021$, Wilcoxon test), while no difference was found in CD73 expression by $CD8^+$ T-cells among any of the BrM patient groups. Notably, this difference in CD73 expression was even more pronounced in the subset of Tregs between the oligo-synchronous and poly metastasis groups ($p= 0.023$, Wilcoxon test) and between the oligo-synchronous and oligo-metachronous BrM groups ($p= 0.016$, Wilcoxon test) (Figure 27a). However, no significant difference in the percentage of any T-cell subgroups expressing CD39 was observed between the BrM patient groups. Furthermore, utilizing unsupervised UMAP analysis, we identified a distinct role of CD73 in the oligo-synchronous group, as indicated by a notable difference in the percentage of CD73-expressing cells at the tip of cluster 2 (Figure 27b). This cluster predominantly comprised CD4+, CD45RA+, CD127+, and CD73+ cells in the oligo-synchronous group.

Taken together, these results strongly suggest that the metastatic phenotype of oligo-synchronous metastasis is associated with altered T-cell differentiation, particularly involving CD73 expression.

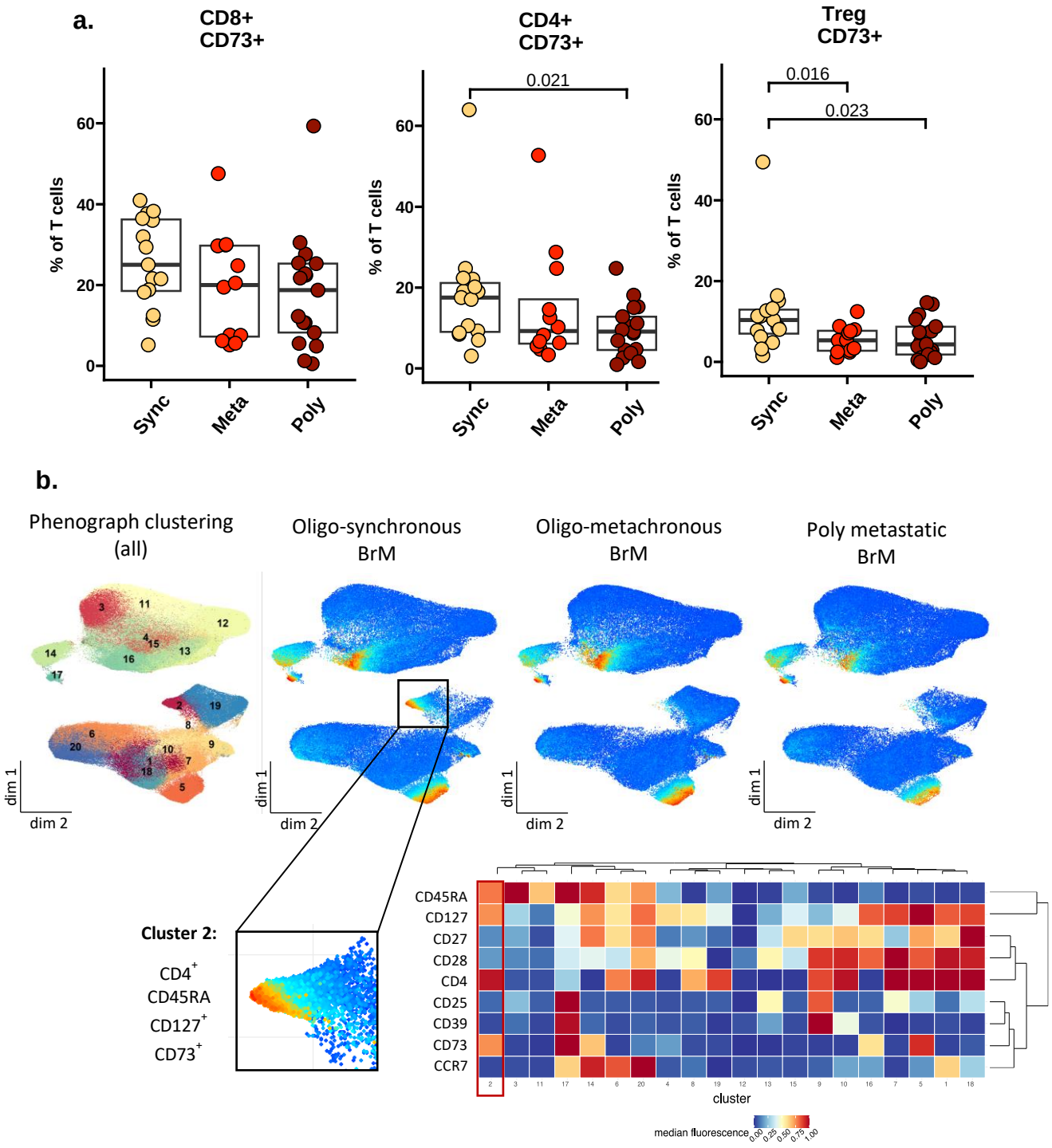


Figure 27: (a.) Box plots of CD73 expressions between BrM groups in $CD8^+$, $CD4^+$, and Treg cells. (b.) Phonograph clustering of different markers in the three groups combined, and individual UMAPs display CD73 expression circled in red (cluster2) in Sync, Meta, and Poly. A Heatmap of markers expression in cluster 2: CD4+, CD45RA+, CD127+ and CD73+. Values describe population size as percentages. (Sync: oligo-synchronous BrM, Meta: oligo-metachronous BrM, and Poly: poly metastasis).

5. Discussion

The intricate process of metastasis development includes multiple complex steps, as illuminated in this thesis. Only a few cells can successfully survive all the steps of metastatic cascade and evade the host's anti-tumor immune response [200]. Moreover, tumor-related alterations in genes play a crucial role in enhancing cancer cell tropism for specific organs and are often linked to the cancer cell's ability to overcome specific obstacles such as the Blood-Brain Barrier (BBB) in the case of the brain. Additionally, these alterations enable the creation of a permissive niche in an unfavorable environment, thereby aiding the progression of metastasis [139, 201]. Given the distinct characteristics of the brain—an organ with limited immune cell access due to the protective mechanism of the BBB—it becomes a conducive environment for the establishment of a tumor niche [202].

5.1. A genetic insight into brain metastasis patterns in NSCLC

Multi-omics approaches including genomics and transcriptomics offer a comprehensive way to comprehend the progression of metastasis. Thus, in our investigation, we integrated DNA and RNA sequencing data of the brain tumors to have a deeper look into the biology of BrM and to pinpoint potential indicators that could distinguish between cases of limited metastasis to the brain (oligo-BrM) and more extensive spreading (poly-metastasis).

Numerous cancers exhibit signs of genomic instability, leading to an increased frequency of somatic mutations and copy number alterations [203]. While somatic genetic alterations are widely acknowledged for their role in initiating primary tumor formation, there remains a gap in our understanding regarding to what extent further genetic mutations play a role in the development of BrM in NSCLC. It has been suggested that the evolution of some tumors might be confined to mutations in specific driver genes while comprehending other types of cancer may require an exploration of large-scale alterations and copy number events [204]. To gain insights into tumor evolution, it is crucial to consider both microevolutionary events, such as driver gene mutations, and macroevolutionary phenomena, including chromosomal aberrations [203].

5.1.1. The Mutational Landscape of BrM in NSCLC

Our findings exhibited good concordance and resemblance in terms of the genomic landscape, effectively aligning with previously published data. The presence of known drivers in primary NSCLC, including *TP53*, *KRAS*, *KEAP1*, and *STK11* indicates that our BrM-NSCLC was representative of NSCLC. The most frequent mutation was observed in the *TP53* gene (detected

in 72% of samples), which is consistent with published data on BrM [98, 117, 205]. Moreover, similar to previously reported BrM-NSCLC data, mutations in genes such as *KRAS* (31%), *STK11* (20%), *KEAP1* (25%), *LRP1B* (59%), and *SMARCA4* (31%) were among the most frequently mutated genes. Some of these mutations were similarly identified in the matched primary NSCLC samples from these earlier investigations, albeit reported to be with higher frequencies in BrM [98, 110, 117, 206, 207]. This observation aligns with our mutation pattern, which is also in accordance with the TCGA data on primary NSCLC. The identification of these alterations implies their probable retention within the metastatic sites, regardless of the specific stages of cancer progression that these mutations facilitated. Consequently, somatic alterations driving cancer progression are expected to exhibit elevated mutational frequencies in brain metastases. As anticipated, none of the patients in our study had mutations in *EGFR, ALK*, or *ROS* genes, as these patients were excluded from the study due to the highly specific oncogenic nature of their cancers.

Although the univariate statistical test showed significance in mutations frequency among BrM groups, this significance could not be maintained following False Discovery Rate (FDR) correction. This is most likely related to the relatively limited size of our cohort due to the scarcity of available BrM samples, which emphasizes the need for careful interpretation. As the complexities of genetic alterations in BrM continue to unfold, larger cohort sizes will provide more robust insights into these complexities.

In the context of our investigation into tumor mutational burden (TMB), it serves as an indicator of the neoantigens of a tumor and can provide insights into tumor heterogeneity and potential therapeutic strategies [208]. Specifically, TMB has been associated with immunotherapy response in various solid tumors including NSCLC [209-211]. As TMB accumulates, so does the abundance of tumor neoantigens, facilitating immune recognition and targeting of the tumor, particularly in the presence of immune checkpoint inhibitors (ICIs) [212].

To explore TMB, in addition to nucleotide changes and mutational signatures, we selectively included samples with matching germline DNA obtained from patient blood samples. Our findings revealed varying levels of overall mutational load across the different BrM patients, ranging from 0.493-31.347 mutation/Mb, with no significant differences observed among the BrM groups (medians: oligo-synchronous= 4.213, oligo-metachronous= 4.133, and poly-metastasis= 4.987). This inter-patient mutational variability within lung cancer BrM cases is consistent with recent literature findings [206, 213, 214]. Similarly, in agreement with our

outcomes, Song et al. also reported the absence of significant TMB differences between synchronous and metachronous BrM groups [117].

Furthermore, studies comparing TMB profiles between BrM and primary lung tumors have yielded heterogeneous outcomes. Some investigations documented heightened TMB levels (>10 mutations/Mb) in metastatic NSCLC in comparison to primary tumors, with this difference being most pronounced in BrM [98, 215]. In contrast, Alvarez-Prado et al. observed marginally elevated TMB in BrM relative to primary tumors [206]. Conversely, other studies have observed greater TMB in primary tumors compared to BrM [117, 213]. Van den Heuvel and collogues have noted variations in the median total TMB across distinct disease stages, with lower TMB in stage IV tumors compared to earlier disease stages [216].

Moreover, there is evidence supporting a correlation between specific genetic alterations and TMB. For example, *PIK3CA* mutations exhibit a positive association with TMB, whereas *TP53* and *EGFR* mutations display a negative correlation [217]. Additionally, BrM-LUAD cases with a TMB ≥ 10 mutations/Mb are less likely to harbor *STK11* mutations [215]. However, our own data did not replicate these correlations ($p > 0.05$ for all mentioned genes).

In addition, our analysis demonstrated a moderate positive correlation between TMB and smoking signature, which is consistent with observations made by Ernst and colleagues, who reported a strong association between TMB and smoking signature [218]. Elevated TMB levels have consistently been linked to the smoking history of patients compared to non-smokers in the context of NSCLC [219-221]. It is worth noting that the moderate correlation observed in our results may be influenced by the inclusion of some patients with missing smoking status, categorized as non-smokers in our cohort, which introduces a degree of uncertainty, as some of these individuals can be smokers.

Taken together, TMB levels seem to be influenced by various factors, including disease stage, smoking history, and specific somatic mutations. Although TMB has been proposed as a site-specific marker in NSCLC and its metastases, including BrM, it does not seem to differentiate between oligo-BrM and poly metastasis.

As expected in most lung cancer cases, the most frequently detected single base substitution (SBS) mutational signature in our BrM cohort was SBS4 (present in 93% of samples), which is characterized by C>A mutations. This signature is strongly associated with tobacco smoking [25, 218]. Moreover, SBS2 predominantly characterized by C>T mutations and SBS13 by C>A and C>G [222], were also prevalent in our cohort (33% and 42% of cases respectively). These two signatures are linked to the activity of AID/APOBEC enzymes [222]. These enzymes

induce cytosine-to-uracil conversion in single-stranded DNA, leading to base substitutions and strand breaks [223]. Their activation is stimulated by cytokines produced during the body's inflammatory response to infections, contributing to infection combat through diverse mechanisms. Activation of APOBEC has been noted in various cancers and is often linked to tissue inflammation [25, 224, 225].

Interestingly, SBS2 and SBS13 exhibited a robust negative correlation with the smoking-associated signature SBS4. This opposing relationship has also been identified in LUAD patients, where APOBEC signatures were enriched in nonsmokers, implying that different mutation patterns exist depending on tobacco exposure [218, 226, 227]. While our cohort's smoking history information is incomplete, this correlation couldn't be verified. Previous investigations into BrM-NSCLC have also reported the presence of tobacco signature [205, 228]. These findings collectively emphasize that lung-brain metastases maintain the mutational profile inherited from their primary tumor source.

In contrast, the reactive oxygen species signature (SBS18) which is related to DNA damage, was observed in almost half of the BrM samples, and showed a strong correlation with smoking signature. Reactive oxygen species are known carcinogens strongly associated with tobacco consumption [229]. This observation aligns with findings by Ernst et al., who identified SBS18 in smoking-associated NSCLC [218]. Additionally, the SBS3 signature was identified in 29% of our BrM cases. This signature is linked to homologous recombination deficiency (HRD) and exhibited a moderate negative correlation with the smoking signature within the studied BrM groups. HRD arises from specific DNA aberrations and has previously been detected in certain lung cancer cases, which can potentially benefit from PARP inhibitor therapy [230]. Moreover, SBS3 has been strongly associated with somatic and germline BRCA1/2 variants and BRCA1 promoter methylation in breast, ovarian, and pancreatic cancers [213]. Elevated levels of genomic aberration-based HRD have been observed in breast cancer BrM [231-233], suggesting that heightened HRD levels might enhance tumor cells' adaptation to CNS microenvironment [232]. In the context of BrM-NSCLC, SBS3 has been identified as one of the prevalent mutational signatures, with higher levels in BrM compared to the primary tumor [213, 228].

Notably, SBS8 signature was found in 87% of our BrM cohort. This signature is common in various cancers and tends to escalate during cancer progression. However, its origin and underlying causes are not well understood [234]. Some studies have proposed an association between SBS8 and genomic instability, including the HRD signature SBS3 [235]. Nevertheless,

SBS8 can also be detected in non-HRD-related genomic instability, underscoring the ongoing debate surrounding its mechanistic foundation; it could potentially result from uncorrected late replication errors during cancer progression [234].

Similar to TMB, the mutational signature profile did not show significant differences among BrM groups. Despite the absence of distinctive mutational patterns, the overall similarity in mutational signatures across the BrM groups underscores the concept that most mutational signatures established in the primary tumor are largely conserved in the metastatic lesions or acquired prior to metastasis [205]. This retention of mutational signatures suggests that the driving genetic alterations initiating primary tumor growth continue to exert their influence throughout the metastatic process.

5.1.2. Copy number alteration (CNA) profile of BrM-NSCLC

The phenotypes associated with metastases are complex and may not solely result from single gene mutations but might rather require significant genomic aberrations [203]. In this study, we aimed to explore the CNA profile of BrM-NSCLC and investigate the potential distinctions among the studied metastatic patterns.

Our BrM samples displayed highly altered copy numbers and showed diversity in their CNA profiles. The chromosome arm-level copy number events in our BrM samples are closely resemble previously published datasets [190, 214]. Many of the genes frequently reported to be highly amplified in lung cancer were observed in our BrM-NSCLC cohort [31, 190, 236], including *TERT* (19.7%), *MYC* (16.4%), *NKX2*-1and *FOXA1* (both 24%), as well as *CCND1* (9.8%) and *CCND3* (8.2%). Most of these genes were also detected in prior studies investigating Brain metastatic NSCLC [110]. Notably, *MYC* amplification is associated with multifocal regional failure [214], and it has been suggested that *MYC* overexpression plays a role in promoting brain metastasis by mitigating oxidative stress caused by activated microglia [237].

Additionally, our CNA results showed deletions within regions containing *CDKN2A*, *CDKN2B* and *PTEN*. These genes are known for their roles in negatively regulating cell proliferation and their deletions have been consistently documented in prior studies of BrM [238]. Specifically, *CDKN2A* copy number deletion which has been reported as one of the most frequent events [110, 120, 214, 228], was also prevalent in our cohort, affecting 24.6% of BrM samples.

Overall, the somatic CNAs identified in our study encompass genes with a credible potential to act as metastatic drivers. *MYC* and *CDKN2A/B* exhibited recurrent genomic amplifications and

deletions, respectively, consistent with findings from a previous sequencing study of brain metastases originating from various primary cancer types, including LUAD [111].

One of the most interesting results of our CNA analysis was the significant alterations found among the different BrM groups. Specifically, four chromosomal regions showed significantly different alterations among these groups. The largest altered regions were located on chromosome 4q31.3-q-35.2 where the oligo-synchronous BrM group exhibited a distinct loss compared to the other groups. Within this region of chromosome 4, we identified three genes (*FAT1*, *IRF2*, and *SFRP2*) that have been previously associated with cancer-related processes, according to published data from cBioPortal and implicated in tumor progression and metastasis [239-241]. However, these genes did not exhibit different mutational frequencies among the various BrM groups. Nevertheless, when we integrated RNA and CNA data, we identified other genes with significantly different expressions on chromosomes 4, 5, and 12, as shown in Table 21. Among them, we selected *TLR3* and *PDE3A* as potential candidate genes. Additionally, smaller chromosomal regions also displayed differences ($p<0.05$) in CNAs among the BrM groups, as detailed in Table 14.

One paper published by Lee et al., conducted a comparative analysis somewhat similar to ours, but focusing on primary LUAD and its relationship with developing BrM. Their study found that primary LUAD cases with early BrM development may harbor more CNAs predictive of metastatic potential or aggressive transformation compared to those with metachronous BrM [242]. Despite the relatively small cohort included in their study, their findings align with our own, underscoring the importance of further investigating the CNA landscapes of both primary and BrM tumors and comparing them between the two oligometastatic subtypes, synchronous and metachronous.

In summary, our results suggest that metastatic patterns in BrM-NSCLC may be distinguished by CNA profiles rather than single mutations. This finding aligns with the theory proposed by Gerlinger et al. in their review, suggesting that large genomic aberrations, such as CNAs, could play a significant role in driving metastasis in certain tumors more than point mutations [203].

5.2. Candidate genes identified as potential biomarkers

5.2.1. By integrating CNA and RNA-seq data

The integration of copy number profiles with RNA-seq data has unveiled two promising gene candidates potentially linked to the pattern of BrM. These candidates are *TLR3* located on Chromosome q435.1 and *PDE3A* located on Chromosome 12p12.2.

Our CNA analysis highlighted a notable aberration in the oligo-synchronous BrM group, characterized by a loss in a relatively large segment of chromosome 4. Building upon this observation, we hypothesize that this genetic aberration may also be reflected in the RNA expression level of *TLR3*, which exhibited lower expression in the oligo-synchronous group when compared to the other two groups. However, this expression difference was particularly significant in comparison to poly metastasis ($p= 0.003$, Wilcoxon test) (as shown in Figure 19). This observation provides support for exploring the role of *TLR3* in the context of brain metastasis.

TLRs are recognized for their presence on immune cells, where they activate the innate immune system, promoting an anti-tumor immune response [243, 244]. *TLR3* is also expressed on epithelial cells, displaying its most prominent expression in the brain, particularly among astrocytes, glial cells, and neurons [245]. While *TLR3* has been found to be expressed in cancer cells of various histotypes and involved in mediating apoptosis through the intrinsic pathway [246], recent evidence suggests its association with tumor progression, metastasis, and therapy resistance [247]. Alkurdi et al. have proposed TLR3 as a potential therapeutic target for treating lung cancer using a combination of paclitaxel chemotherapy and TLR3-ligand, as *TLR3* is frequently overexpressed in NSCLC in contrast to normal bronchial epithelium [248].

However, the role of *TLR3* can be paradoxical based on cancer type, stage, and immune microenvironment context. In the pre-metastatic phase, *TLR3* signaling promotes apoptosis in lung and breast cancer, as well as head and neck squamous cell carcinoma. Prior research has emphasized that *TLR3* activation can induce cancer cell apoptosis in both human and murine models, or even suppress cancer cell migration depending on the tumor stage. However, post-initiation of the metastatic process, *TLR3* activation takes on an opposing role, enhancing tumor migration [249-251]. In mouse models of breast and lung cancer, the activation of *TLR3* by tumor RNA in metastasis triggers the expression of *SLIT2*, which subsequently promotes the migration of cancer cells toward endothelial cells and facilitates intravasation [252]. In human breast cancer, elevated TLR3 expression in tumor cells, as observed through IHC has been

significantly linked to a higher rate of distant metastasis [253]. Conversely, in colorectal cancer, the absence of TLR3 expression on IHC has been correlated with lymph node metastasis [254].

In the brain, *TLR3* is also expressed on the surface of microglia, which play a pivotal role in clearing tumors and metastatic cells [200, 251, 255]. This indicates the complex nature of *TLR3* and its intricate role in cancer. In our results, we observed different expression patterns of *TLR3* in oligo-synchronous BrM. However, its precise pro- or anti-tumor role in this metastatic pattern remains unclear. In conjunction with the previously mentioned insights, these findings underscore the significance of investigating into the multifaceted involvement of *TLR3* within the complex landscape of brain metastasis, making it a strong candidate for further investigation in our study cohort. Additionally, including both metastatic and non-metastatic samples in future analyses could help provide a clearer understanding of the direction in which *TLR3* is implicated.

The other candidate gene was *PDE3A*, which is a member of the phosphodiesterase superfamily. Phosphodiesterases play a crucial role in regulating intracellular concentrations of cyclic nucleotides, including cyclic AMP (cAMP) and cyclic GMP (cGMP). Through their enzymatic activity, they catalyze the hydrolysis of these second messengers, thereby modulating various intracellular signal transduction pathways and cellular activities [256]. Furthermore, *PDE3A* activates inflammatory pathways linked to cancer cell stemness by suppressing cAMP/PKA pathway [257]. *PDE3A* has been implicated in cancer cell invasion and cell motility, suggesting its involvement in tumor progression [258]. Notably, upregulation of *PDE3A* has been correlated with metastasis in breast cancer [257] and SCLC [259, 260].

In NSCLC, De Waal et al. revealed a novel role for *PDE3A* in cancer maintenance, which its function can be modified by a subset of *PDE3* inhibitors, resulting in toxicity to lung cancer cell lines expressing elevated levels of *PDE3A* [261]. In NSCLC-LUAD, it emerges as a prominent player. Following radiotherapy, *PDE3A* is one of the most frequently mutated cuproptosis-related genes in LUAD. Cuproptosis, a newly described non-apoptotic cell death mechanism, holds potential as a therapeutic modality for LUAD patients who develop resistance to conventional treatments [262].

While no existing publications have directly linked *PDE3A* alterations to brain metastasis, our analysis showed a significant upregulation in *PDE3A* gene expression within the oligo-synchronous BrM when compared to the poly metastatic group (FDR=0.037). Despite the gene's complex loci, where both loss and gain were observed, *PDE3A* expression exhibited a notable elevation in oligo-synchronous BrM group. This finding suggests that *PDE3A* could be

an interesting candidate gene for further investigation and validation in this context, even in the absence of established connections with brain metastasis in the literature.

Of particular importance is the fact that PDE3A's involvement in a multitude of cellular pathways, as elucidated earlier, suggests its plausible implication in biological pathways that hold relevance to brain metastasis. Specifically, our results indicate that *PDE3A* upregulation contribute to the metastatic process by enhancing cancer cell invasiveness in oligo-synchronous BrM.

5.2.2. By considering DEGs and pathway enrichment analysis

Based on the analysis of the top differentially expressed genes across the studied groups, as well as pathway enrichment analyses revealing their involvement in critical pathways and supported by relevant findings in the published literature, three genes have emerged as promising candidates for investigating their potential roles in brain metastasis in NSCLC: *AKR1C1*, *SMURF2*, and *S100A2*.

AKR1C1 gene encodes an enzyme belonging to the aldo-keto reductase 1C protein family (*AKR1C1-AKR1C4*), which catalyzes NADP+-dependent reduction and plays essential roles in the metabolism of steroid hormones, prostaglandins, and polycyclic aromatic hydrocarbons [263]. Chien et al. suggested the potential oncogenic function of *AKR1C1*, reporting that upregulation of this gene leads to neoplastic transformation and tumor formation in nude mice [264]. Earlier studies have also documented the upregulation of AKR1C1 in various cancers, including lung cancer [265, 266].

Due to *AKR1C1* role in the metabolic activation of polycyclic aromatic hydrocarbons (PAH), known lung carcinogens, it has been implicated in contributing to smoking-related lung cancer. This is especially relevant as its overexpression has been observed in the bronchial epithelial cells of tobacco smokers [265, 267, 268]. Furthermore, *AKR1C1* has been reported to promote cell proliferation and metastasis in NSCLC, where its overexpression in non-metastatic cancer cells has been found to notably promote metastasis both *in-vitro* and *in-vivo*. Zhu et al. demonstrated that this pro-metastatic activity is linked to acetylated *AKR1C1*'s activation of the STAT3 pathway [269, 270]. While Fu et al, revealed that targeting *AKR1C1* via ALA inhibits this pathway, effectively suppressing NSCLC proliferation and metastasis [271].

Chang et al., on the other hand, presented evidence of *AKR1C1*'s role in promoting cell proliferation through crosstalk between hypoxia-inducible factor 1-alpha (HIF-1α) and metabolic reprogramming [196]. These results collectively highlight the complex role of

AKR1C1 in carcinogenesis and suggest its potential as an important target for cancer therapy [271].

In our investigation of BrM groups, the oligo-synchronous group exhibited a significant upregulation of *AKR1C1* expression compared to the poly-metastasis group. Furthermore, our pathway analysis revealed associations of *AKR1C1* with metabolic and enzymatic activities, as well as the regulation of biological and cellular processes. These findings underscore the unique molecular landscape of oligo-synchronous BrM, indicating a potential predominance of pro-tumor progression activities in this specific metastatic pattern.

Notably, we observed an interesting contrast in the case of *AKR1C1* expression levels. High expression levels of *AKR1C1* across all BrM cases were positively correlated with the survival rate, which differs from previously published data [270-272]. This discrepancy could be attributed to the specific characteristics of our small RNA cohort, emphasizing the necessity for including a larger sample size in future studies.

The candidate gene *SMURF2* encodes an E3 ubiquitin ligase pivotal in targeted ubiquitin tagging and the regulation of TGF-β signaling. It exhibits diverse effects on various cellular processes, including the DNA damage response, preservation of genomic stability, modulation of chromatin modifications, and regulation of the cell cycle control [273, 274]. In cancer, *SMURF2* has been correlated with the development and progression of tumors. However, divergent findings in various studies depict its role as either promoting or suppressing tumors in different cancers, as Fu et al. explained in their review [275].

For instance, certain investigations have demonstrated that elevated *SMURF2* levels can foster invasion, migration, and metastasis in human breast cancer tissues and MDA-MB-231 cell lines [276, 277]. Similarly, *SMURF2* overexpression has been linked to esophageal squamous cell carcinoma and pancreatic carcinoma, with specific implications for tumor invasion and lymph node metastasis [278, 279]. In colorectal carcinoma, *SMURF2*'s potential oncogenic role has been suggested, as it was found to enhance the invasion and migration of tumor cells [280]. However, a contrasting perspective was provided by Fukunaga et al., who demonstrated that knockdown of Smurf2 in MDA-MB-231 cells promoted cell migration in vitro and led to bone metastasis in vivo, indicating a possible tumor suppressor function [281].

Notably, *SMURF2*'s functional diversity extends to hepatocellular carcinoma, where it was reported to suppress metastasis through the ubiquitin degradation of Smad2 [282]. In medulloblastoma, the intricate role of *SMURF1* and *SMURF2* was demonstrated, as their knockdown or overexpression promotes or inhibits cell proliferation, and colony formation

respectively, by modulating RNF220 protein levels, thus affecting Shh signaling [283]. Additionally, a noteworthy context arises in lung cancer, where the role of *SMURF2* as a tumor suppressor was highlighted. Depletion of SMURF2 was shown to promote cell proliferation and tumorigenesis *in-vitro* and *in-vivo* in nude mice [284].

Accumulating evidence indicates that *SMURF2* regulates a wide spectrum of physiological processes, including cell proliferation, invasion, and migration through its regulatory functions within various signaling pathways [277, 285, 286]. This observation aligns with our pathway enrichment analysis results.

In our investigation, we observed a significant upregulation of *SMURF2* expression in the poly-metastasis group compared to both the oligo-synchronous and oligo-metachronous BrM groups. While the direct link between *SMURF2* and brain metastasis remains underexplored in existing literature, the multifaceted involvement of *SMURF2* in regulating cell migration and invasion suggests its potential significance in the poly metastatic process. Thus, despite the limited specific studies, our results fostered us to consider *SMURF2* as a potential player in BrM, encouraging further investigation and validation of its role within this context.

S100A2 is a member of the S100 family of proteins with a wide expression distribution across various organs, and engaging in interactions with a broad spectrum of molecules [287-289]. Its presence has been identified in organs such as the lungs, kidneys, prostate, and various glands, in addition to being found in mammary epithelial cells [290]. Notably, this protein family has been linked to pivotal roles in both physiological and pathological processes, contributing to human diseases [291].

In the context of cancer, *S100A2* plays a significant role in tumor pathogenesis. Research has demonstrated that S100A2 is induced by the p53 activator etoposide and positively regulated by p53 [292-294]. Furthermore, *BRCA1* has been implicated in the interaction with ΔNp63, leading to the positive regulation of *S100A2* expression and consequent tumor growth [295]. Interestingly, there are controversial roles of *S100A2* in cancerogenesis, wherein it functions as a tumor suppressor in certain cancer types [290], such as breast cancer [296], esophagus squamous cell carcinoma [297] and oral cancer [298]. In contrast, it acts as tumor promoter in others [290], with instances of *S100A2* overexpression observed in epithelial skin tumors [299] and ovarian cancer [300].

The dual role of *S100A2* was specifically observed in lung cancer [301]. Feng et al. proposed *S100A2* as a tumor suppressor in the early stages of human lung carcinogenesis [302]. Conversely, Heighway et al. [303] and Wang et al. [304] reported frequent overexpression of

S100A2, which and linked with worst prognosis in NSCLC patients with stage I. Moreover, elevated levels of *S100A2* mRNA expression have been correlated with unfavorable clinical outcomes in patients with NSCLC who have undergone surgical resection [305].

The significance of *S100A2* expression has been observed in various malignancies, reflecting alterations in its expression during the transformation and metastasis of different cell types and tumors [306]. Naz et al. reported a protumorigenic action of *S100A2* and its involvement in EMT and TGF-β-mediated cancer cell invasion in A549 lung cancer cells [307]. Moreover, earlier studies have underscored the crucial role of S100A2 as a positive driver of tumor development and distant metastasis in NSCLC [305, 308]. In addition, overexpression of *S100A2* in LUAD tissues has been associated with lymph node metastasis [309]. We found that *S100A2* was upregulated in the poly-metastasis group when compared to the oligo-BrM groups. However, this expression difference reached significance only within the oligo-metachronous BrM subgroup.

Interestingly, *S100A* protein family has been also implicated to participate in innate and adaptive immune responses, cell migration, tissue development, repair mechanisms, and tumor cell invasion [287]. These proteins were suggested to function through IL-17 signaling pathway. In low-grade glioblastoma, *S100A2* was shown to be positively correlated with the infiltration of $CD4^+$ T-cells, B-cells, macrophages, neutrophils, and dendritic cells [310]. Similarly, this positive correlation between *S100A2* and M0 macrophages and activated dendritic cells was also observed in pancreatic cancer. However, it showed a remarkable negative correlation with $CD8^+$ T-cells and NK cells [311]. These associations highlight the potentially pivotal role of *S100A2* in orchestrating interactions within the tumor microenvironment and the immune response.

Given these findings, *S100A2* emerges as an attractive candidate gene for validation. This could potentially shed light on its involvement in the complex mechanism underlying metastasis and its interplay with the immune milieu within the brain microenvironment.

Both *SMURF2* and *S100A2* showed similar expression patterns, with elevated levels in the poly-metastatic group. Furthermore, pathway enrichment analysis indicated their involvement in cell motility and migration. Despite the conflicting findings regarding their roles in cancer, we hypothesize that *SMURF2* and *S100A2* may contribute to the promotion of multi-metastatic dissemination to different organs.

To our current knowledge and search, there is no published data available for the direct comparison between oligo-BrM and poly-metastasis groups. Therefore, the candidate genes we

have identified hold the promise of ser ving as novel biomarkers capable of distinguishing between these distinct metastatic patterns. However, the specific mechanisms and roles of these genes in this process remain unclear and require further investigation. The validation of these potential biomarkers is in progress, with ongoing testing of suitable antibodies to confirm their utility and reliability.

5.3. Immune landscapes of BrM-NSCLC

The immune defense mechanisms in the brain differ from those in most other tissues and play a pivotal role in responding to tumors [312]. On a systemic level, tumor cells and the peripheral immune system mutually influence each other, giving rise to what is commonly referred to as the systemic immune environment [153]. Consequently, conducting a comprehensive analysis of local and peripheral immunological characteristics holds the potential to unveil crucial insights into immune-related parameters associated with BrM.

In this part of our project, we aimed to explore the local and peripheral immunological characteristics of NSCLC patients with BrM. Within the brain microenvironment, we observed that $CD68^+$ cells were the predominant immune cell population both in the peri-tumoral and intra-tumoral regions of the brain. Interestingly, our analysis revealed significantly higher counts of $CD68^+$ cells and lymphocytes (TILs) in the peri-tumoral regions compared to the intra-tumoral areas. Nevertheless, it is noteworthy that, similar to findings from other research studies, the abundance of TILs was notably lower than what has been reported for primary lung tumors. This observation suggests the presence of an immunosuppressive microenvironment in the brain [117, 122, 313, 314].

To my knowledge, no prior studies have investigated the immune profiles among different brain metastatic patterns. In our tissue immunophenotyping, a specific upregulation in $CD4^+$ T-cells was observed in oligo-synchronous BrM compared to the other metastatic groups. This difference was particularly pronounced in the infiltration of intra-tumoral $CD4^+$ T-cells. Earlier *in-vivo* data have demonstrated that $CD4^+$ T-cells, including T helpers and Tregs, can influence metastatic spreading independently of the presence or absence of $CD8^+$ T-cells, illustrating both pro- and anti-metastatic roles of $CD4^+$ T-cells depending on their differentiation [315].

Therefore, to gain further insights into the role of $CD4^+$ T-cells, we evaluated their expression of FOXP3, a marker indicative of Tregs [316]. Our analysis revealed that T-cells expressing FOXP3 did not exhibit significant differences among the various BrM groups. This suggests that perhaps it's not Tregs but other subsets of $CD4^+$ T-cells, such as T-helpers, that hold particular importance for the oligo-synchronous BrM group. Therefore, it is necessary to

conduct more in-depth investigations that can help precisely identify and characterize these T-cell populations. Techniques like spatial single cell sequencing offer the potential to capture the exact phenotypes of these cells.

However, an intriguing finding emerged—there was a positive correlation between the presence of infiltrating cells expressing FOXP3 in intra-tumoral regions and overall survival for the whole study cohort ($p= 0.045$). Although limited data exist on the prognostic significance of FOXP3 in BrM [141, 317], our findings are in alignment with another study focused on BrM-NSCLC. This study also reported a positive correlation between overall survival and the density of FOXP3-expressing cells within tumor lesions, but not in the peri-tumoral regions [122]. This result stands in contrast to the general trend observed in tumors. Therefore, the expression of FOXP3 warrants further investigation to elucidate this phenomenon.

Moreover, metastatic disease has been demonstrated to modulate also the peripheral immune response, which can be monitored in the blood of cancer patients [318]. Therefore, we used flow cytometry to compare the peripheral blood T- and NK-cell profiles, including their subsets in different BrM-NSCLC groups. Consequently, when comparing patients from the three different BrM groups, no significant differences emerged regarding the population sizes of $CD4^+$ and $CD8^+$ T-cells. However, correspondingly to the results from the BrM tissues, a notable shift was identified in the distribution of $CD4^+$ differentiation phenotypes. Exploring the expression patterns of T-cell functional differentiation subtypes, exhibited remarkable differences in $CD4^+$ subtypes patterns among the BrM groups, whereas $CD8^+$ subtypes showed a relative homogeneity. This observed distinctive pattern may potentially imply that metastasis triggers different immune responses specifically on the level of peripheral $CD4^+$ T-cells. Our analysis showed more $CD4^+$ naïve cells in the oligo-synchronous BrM group, while $Teff_{RA+}$ and $Teff_{RA-}$ cells were significantly less observed. Naïve T-cells function as immune surveillance circulating in the blood, promptly reacting to foreign antigens. Subsequently, they get activated and differentiate into effector cells $Teff_{RA+}$ that eliminate or help other immune cells to combat the invader. After the peak of effector expansion and the clearance of antigen, most activated T-cells die, and only a small population transitions into a memory T-cell pool. Part of these memory cells are effector memory cells ($Teff_{RA-}$), which provide a rapid immune response upon restimulation [319-321]. Overall, our results indicate that patients with oligo-synchronous BrM have a less activated immune system suggesting that a brain metastasis alone does not trigger a robust peripheral immune response.

Furthermore, our investigation revealed a noticeable increase in CD73 expression, specifically in $CD4^+$ and Treg cells in oligo-synchronous BrM compared to the other BrM groups. In human peripheral T-cell compartments, CD73 is primarily expressed on naïve $CD8^+$ T-cells and to a lesser extent on memory $CD8^+$ T-cells [322, 323]. Conversely, circulating $CD4^+$ T-cells expressing CD73 are predominantly memory cells [324]. Moreover, a specific subpopulation of effector $CD4^+$ cells enriched in polyfunctional Th1.17 cells has been reported [325]. However, CD73 expression on Treg is relatively low [325, 326].

CD73 is a membrane-bound enzyme expressed on both tumor and immune cells, converting AMP into adenosine [322, 327]. Adenosine interacts with A2AR expressed on immune cells, inhibiting T-cell activation, proliferation, and cytokine secretion [328]. In Murine models, CD73 is mostly expressed by Tregs and involved in suppressing the anti-tumor immune response [329]. However, in humans, the role of adenosine-mediated suppression by Tregs is still unclear [330, 331]. Tolosa and colleagues suggested that human Treg-derived CD73 is not essential for adenosine-mediated suppression of conventional $CD4^+$ T-cells [332]. On the other side, Mandapathil et al. reported that inhibition of CD73 reduced human-mediated suppression *in-vitro* [326]. These differences in CD73 expression between murine and human models, as well as *in-vitro* and *in-vivo* studies make its function unclear. Therefore, it is uncertain whether CD73 expression variations observed in Tregs influence metastatic patterns or the immune response in our NSCLC patients.

In our unsupervised analysis, a difference in cluster 2 of oligo-synchronous BrM emerged, encompassing CD4+ CD45RA+ CD127+ CD73+ cells (Figure 27b). This population may belong to early-memory $CD4^+$ cells that could have migrated from the TME into circulation. a previous study has reported a higher proportion of CD73 on circulating $CD4^+$ in patients with chronic inflammation [333], indicating that CD73 on $CD4^+$ might be part of an immune inflammatory response. Beyond its enzymatic function, CD73 has been identified as a lymphocyte differentiation antigen, suggesting its involvement in lymphocyte maturation, development, and T-cell activation. Additionally, CD73 serves as an adhesion molecule facilitating the binding of lymphocytes to the endothelium [334-336].

While studies on CD73 in the peripheral immune system are very scarce, its impact on tumor progression and anti-tumor responses in NSCLC has been observed in the TME [337, 338]. Given the complex roles and mechanisms of CD73, caution is needed when interpreting our results. While the exact function of CD73 in BrM-NSCLC requires further investigation, our data suggest an intriguing difference among BrM cohorts. We observed an elevated presence

of $CD73^+$ T-cells in oligo-synchronous BrM, particularly within CD4 and Treg cell populations. These findings align with the observed differentiation phenotypes of $CD4^+$ T-cells observed in the oligo-synchronous BrM, indicating a more pronounced immune suppressive environment in this BrM group.

In summary, it has been suggested that the differentiation patterns of peripheral CD4 cells can serve as an independent predictor of tumor progression in NSCLC [339]. Our data provides additional evidence that this approach also yields valuable insights into the pattern of metastatic spread among BM patients.

5.4. Does the tumor's genomic landscape shape the host's immune profile or vice versa?

The mutational characteristics of cancer cells can influence the immune phenotype of the tumor microenvironment (TME) by initiating immunosuppressive signals that contribute to the formation of a tumor-supportive TME. Conversely, TME compartments can also in turn impact the genomic landscape of the tumor. This impact is not limited to simply detecting and eliminating immunogenic clones but extends to promoting the outgrowth of clones capable of evading immune responses [340]. This intricate relationship has been observed in various cancer types, including NSCLC [341], pediatric and adult brain tumors [342-344].

Interestingly, oligo-synchronous BrM displayed the most pronounced alterations in both molecular and immunological aspects. The chromosomal aberrations observed in this group might exert an influence on the presence of $CD4^+$ T-cells within the TME or their differentiation profile in peripheral blood. For instance, studies have noted that TP53-mutant lung BrMs exhibit increased infiltration and activation of $CD8^+$ T-cells, along with a more immunosuppressive myeloid compartment [206]. Among our candidate genes, *TLR3* and *S100A2* have been implicated in innate and adaptive immune responses, as mentioned earlier. This suggests a possible connection between molecular alterations and immune profiles, however, confirming this hypothesis requires further investigation.

5.5. Conclusion and outlook

The high mortality rate associated with lung cancer is primarily attributed to early metastasis to various organs, including the brain, which occurs in more than 40% of cases [345]. Brain metastases pose a significant clinical challenge in the management of lung cancer patients, often resulting in a median survival rate of less than six months without treatment [76]. Therefore, there is an urgent need to identify reliable markers that can aid in predicting brain metastasis

and the extent of tumor spread. Finding such markers would enable clinicians to identify patients with a high risk of developing brain metastasis and determine those who could benefit from radical treatment.

In this study, our primary objective was to gain a deeper understanding of the biology underlying different patterns of brain metastasis in NSCLC. To achieve this, we conducted a novel and comprehensive analysis, including molecular and immunological aspects of our BrM cohort. With that, we wanted to find potential biomarkers that could differentiate between oligo-BrM and poly-metastatic disease. Notably, while numerous studies have recently explored BrM-NSCLC, none have delved into this particular aspect of the disease.

Prior research has suggested that oligometastatic disease might manifest as a distinctive form with specific molecular characteristics [106, 346]. To delve deeper into this hypothesis, our study included two distinct groups within our cohort: oligo-BrM patients, where brain metastasis was the sole affected organ, and patients with multiple distant metastases, including brain metastasis, termed the poly-metastatic group. Furthermore, the oligo-BrM group was in turn divided into synchronous and metachronous oligo-BrM. It's important to note that our cohort is not a purely random selection of samples, as all patients underwent surgery at the UKE hospital. Consequently, future investigations may benefit from including samples from diverse sources to enhance the robustness of our findings.

When examining the mutational landscape across various BrM subtypes, we found that small somatic mutations (SNVs and InDels), tumor mutational burden (TMB), and mutational signatures did not exhibit significant differences among the three BrM patterns. However, the profiles of CNA unveiled substantial distinctions, particularly within four chromosomal regions on chromosomes 4, 5, 8, and 12. Specifically, the oligo-synchronous BrM group exhibited a significant loss in more than half of the samples within a >37Mb segment on chromosome 4q. Additionally, this subgroup displayed a notable gain in 8p and a loss in 12p, although these regions were comparatively smaller in size.

From these observations, it can be inferred that the average frequency of the small somatic mutations is generally higher across all BrM patterns when compared to published data on primary NSCLC, which aligns with previous research findings. However, it's important to note that these mutations do not appear to influence the metastatic patterns or distinguish between the BrM subgroups. It should be emphasized that one of the limitations of our study was the relatively small cohort size, underscoring the need for further confirmation with a larger sample pool. Nonetheless, it is evident that chromosomal aberrations play a role in characterizing

synchronous brain metastasis in NSCLC. Copy number alterations have previously been linked to an increased risk of distant metastases at the time of diagnosis in a prior investigation [347].

Furthermore, it's worth noting that RNA data faced limitations, with even fewer samples included due to degradation issues in some samples during the sequencing process. Additionally, the majority of oligo-metachronous BrM samples had received prior treatment, potentially impacting the results and making direct comparisons challenging.

Based on NGS data obtained from DNA and RNA, and the subsequent bioinformatic analyses, we have identified five genes that hold promise as potential markers capable of distinguishing between the distinct BrM subgroups. These candidate genes have previously been implicated in tumor progression and metastasis, offering intriguing avenues for further exploration. Validation of these candidate genes is necessary to confirm their impact at the protein level, involving IHC staining of the target protein on an extensive set of samples, including both metastatic and non-metastatic cases to validate sensitivity and specificity. In addition, *in-vitro* functional experiments on cell lines focusing on these target genes are essential to elucidate their functions and assess the feasibility of testing them in different *in-vivo* models.

In clinical diagnosis and treatment, identifying predictive or prognostic markers can empower clinicians to categorize patients at risk for either oligo- or poly-metastatic disease, thereby aiding in the identification of individuals who can benefit from specific treatments.

Our study provides valuable insights into the immune landscape of brain metastasis in NSCLC. Notably, the oligo-synchronous BrM group exhibited specific alterations in the $CD4^+$ T-cell population, both within the tumor microenvironment and in peripheral blood. To the best of our knowledge, no prior research has explored the relationship between $CD4^+$ T-cells and BrM. Therefore, further investigations are warranted to unravel the intricate connection between NSCLC and its distinct BrM phenotypes.

While our project focused on T-cells and NKs, it's essential to acknowledge that other immune cell populations likely influence BrM formation. Therefore, future studies should delve into the roles of immune cells such as B-cells and macrophages. Additionally, broadening the scope to include more samples is imperative to define the biological significance of the observed differences, particularly in the context of driving oligo-metastatic disease.

In conclusion, this thesis sheds light upon the molecular and immunological characteristics of different patterns in NSCLC, uncovering novel alterations at both levels. Furthermore, our research has identified five potential biomarkers, with two of them implicated in the immune system. However, validation of these findings is still essential. Our data represents a significant

step forward in comprehending BrM disease, laying the foundation for further investigations and advancing our knowledge in this field.

6. Abstract

Brain metastases (BrMs) pose a severe complication in cases of lung cancer, significantly impacting both patient prognosis and their quality of life. Approximately 5-10% of individuals with non-small cell lung cancer (NSCLC) experience the presence of BrMs at the time of their initial diagnosis, and up to 40% will develop BrMs as the disease progresses, resulting in significantly reduced overall survival rates. However, patients with oligometastasis (involving only the brain) tend to exhibit a more favorable prognosis and may represent a distinct disease manifestation compared to patients with multiple metastases affecting distant organs (poly-metastases).

The influence of host-related (immunological) factors and tumor-specific variables in governing the extent of metastatic spread to the brain has remained relatively unexplored. In this thesis, we aimed to gain a better comprehension of BrM in NSCLC. Our objectives were twofold: first, to investigate and compare the molecular characteristics and immunological phenotypes of patients with three different BrM patterns—oligo-synchronous (BrM at the initial diagnosis), oligo-metachronous (BrM occurring later after the initial diagnosis), and poly-metastasis (involving multiple organs, including the brain); second, to identify new markers capable of predicting oligo- or poly-metastatic disease, which could assist clinicians in identifying patients who would benefit from more aggressive treatment strategies.

Tumor-specific factors were investigated through DNA and RNA next-generation sequencing (NGS) of brain metastases. While the landscape of copy number alterations (CNAs) showed significant variation among the studied groups, point-mutation profiles remained consistent across the patient groups.

Through mRNA-seq analysis, we revealed distinct gene expression patterns among the BrM patterns. The comprehensive data analyses identified five potential candidate genes (*TLR3*, *PDE3A*, *AKR1C1*, *SMURF2*, and *S100A2*) that may contribute to the metastatic process or growth patterns. These genes hold promise as biomarkers to distinguish between oligo- and poly-metastatic disease.

To investigate variations in immune phenotypes among our BrM cohorts, tissue samples were analyzed for tumor-infiltrated lymphocytes (TILs) and microglia/macrophages within the tumor microenvironment (TME). In parallel, blood samples were examined to probe the peripheral immune landscape of the studied groups, focusing on T-cells and NK cells. Interestingly, the oligo-synchronous BrM group exhibited significant distinctions in $CD4^+$ T-cell populations in both TME and peripheral blood when compared to the other groups. Additionally, we detected

a higher frequency of CD73-expressing $CD4^+$ T-cell populations in the oligo-synchronous BrM group. Taken together, these findings indicate a unique immune profile at the $CD4^+$ T-cell level in the oligo-synchronous BrM group.

Taken together, the current study represents the first comprehensive investigation into the distinct patterns of brain metastasis in NSCLC disease. Our findings unveil intriguing differences in the molecular and immunological profiles of oligo-synchronous BrM, providing essential insights into the complexity of BrM in NSCLC.

Zusammenfassung

Gehirnmetastasen (BrMs) stellen eine schwerwiegende Komplikation bei Lungenkrebs dar und beeinflussen signifikant die Prognose der Patienten sowie deren Lebensqualität. Etwa 5-10% der Personen mit nicht-kleinzelligem Lungenkrebs (NSCLC) weisen bei ihrer erstmaligen Diagnose BrMs auf, und bis zu 40% entwickeln BrMs im Verlauf der Krankheit, was zu erheblich reduzierten Überlebensraten führt. Allerdings neigen Patienten mit Oligometastasen (nur im Gehirn) dazu, eine bessere Prognose zu haben und könnten eine eigenständige Krankheitsmanifestation im Vergleich zu Patienten mit multiplen Metastasen in entfernten Organen (Poly-Metastasen) darstellen.

Der Einfluss von wirtsspezifischen (immunologischen) Faktoren und tumorbezogenen Variablen auf das Ausmaß der metastatischen Ausbreitung im Gehirn ist weitgehend unerforscht geblieben. In dieser Dissertation hatten wir zwei Hauptziele: Erstens wollten wir eine bessere Verständnis für BrM im NSCLC entwickeln. Wir untersuchten und verglichen die molekularen Merkmale und immunologischen Phänotypen von Patienten mit drei verschiedenen BrM-Mustern: oligosynchron (BrM bei der erstmaligen Diagnose), oligometachron (BrM traten später nach der erstmaligen Diagnose auf) und Polymetastasen (mit multiplen Organen, einschließlich des Gehirns). Zweitens wollten wir neue Marker identifizieren, die in der Lage sind, oligo- oder polymetastatische Erkrankungen vorherzusagen und Ärzten bei der Identifizierung von Patienten unterstützen könnten, die von aggressiveren Behandlungsstrategien profitieren würden.

Tumorspezifische Faktoren wurden durch die DNA- und RNA-Sequenzierung der nächsten Generation (NGS) von Gehirnmetastasen untersucht. Während die Landschaft der Kopienzahlveränderungen (CNAs) erhebliche Unterschiede zwischen den untersuchten Gruppen aufwies, blieben die Punktmutationsprofile zwischen den Patientengruppen konsistent.

Durch mRNA-Sequenzierung zeigten sich unterschiedliche Genexpressionsmuster zwischen den BrM-Mustern. Die umfassende Datenanalyse identifizierte fünf potenzielle Kandidatengene (*TLR3*, *PDE3A*, *AKR1C1*, *SMURF2* und *S100A2*), die zum metastatischen Prozess oder den Wachstumsmustern beitragen könnten. Diese Gene bergen das Potenzial als Biomarker zur Unterscheidung zwischen oligo- und polymetastatischen Erkrankungen.

Um Unterschiede in den Immunphänotypen unserer BrM-Kohorten zu untersuchen, wurden Gewebeproben auf tumorinfiltrierende Lymphozyten (TILs) und Mikroglia/Makrophagen innerhalb des Tumormikroenvironments (TME) analysiert. Parallel dazu wurden Blutproben

untersucht, um die periphere Immunlandschaft der untersuchten Gruppen zu untersuchen, wobei der Fokus auf T-Zellen und NK-Zellen lag. Interessanterweise zeigte die Gruppe der oligosynchronen BrM signifikante Unterschiede in $CD4^+$ T-Zell-Populationen im TME und im peripheren Blut im Vergleich zu den anderen Gruppen. Darüber hinaus stellten wir eine höhere Frequenz von CD73-expressierenden $CD4^+$ T-Zell-Populationen in der Gruppe der oligosynchronen BrM fest. Zusammenfassend deuten diese Befunde auf ein einzigartiges Immunprofil auf $CD4^+$ T-Zellebene in der Gruppe der oligosynchronen BrM hin.

Zusammenfassend stellt diese Studie die erste umfassende Untersuchung der unterschiedlichen Muster von Hirnmetastasen in der NSCLC-Krankheit dar. Unsere Erkenntnisse enthüllen faszinierende Unterschiede in den molekularen und immunologischen Profilen von oligosynchronen BrM, die wesentliche Einblicke in die Komplexität von BrM im NSCLC bieten.

7. List of abbreviations

AKR1C1	Aldo-keto reductase family 1 member C1
ALK	Anaplastic lymphoma kinase
BBB	Blood-brain barrier
BRAF	B-raf proto-oncogene
BrM	Brain metastasis
CNA	Copy number alteration
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
DC	Dendritic cells
DEG	Differently expressed gene
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
FFPE	Formalin fixed paraffin embedded
FGFR1	Fibroblast growth factor receptor 1
GO	Gene Ontology
HRD	Homologous Recombination Deficiency
ICIs	Immune checkpoints inhibitors
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRAS	Kirsten rat sarcoma viral oncogene homolog
LCLC	Large cell carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MMR	DNA mismatch repair
NK	Natural killers
NSCLC	Non-small cell lung cancer
PBMC	Peripheral blood mononuclear cells
PD-1	Programmed death protein 1
PD-L1	Programmed death-ligand 1
PDE3A	Phosphodiesterase 3A
PMN	Pre-metastatic niche
S100A2	S100 calcium-binding protein A2
SBS	Single Base Substitution
SCLC	Small cell lung cancer
SMURF2	SMAD Specific E3 Ubiquitin Protein Ligase 2

TCGA	The Cancer Genomic Atlas
TILs	Tumor-infiltrating lymphocytes
TKI	Tyrosine kinase inhibitors
TLR3	Toll Like Receptor 3
TMB	Tumor mutational burden
TME	Tumor microenvironment
TP53	Tumor protein p53
Tregs	Regulatory T cells
UICC	Union for International Cancer Control
WES	Whole-exome sequencing
WGS	Whole-genome sequencing

8. Appendix

Supplement table 1: Top 40 DEGs in oligo-synchronous vs. oligo-metachronous BrM

Gene.symbol	EnsemblID	log2FoldChange	lfcSE	pvalue	padj
SCGB3A1	ENSG00000161055	7.181.995.463	1.181.327.334	1.20E-09	2.110E-05
LINC00942	ENSG00000249628	-5.249.826.575	1.035.214.175	3.95E-07	3.454E-03
LXN	ENSG00000079257	279.123.441	0.568699305	9.20E-07	4.651E-03
MAST4-AS1	ENSG00000229666	3.564.502.757	0.730528944	1.06E-06	4.651E-03
BMP5	ENSG00000112175	4.890.127.136	1.026.424.114	1.90E-06	4.686E-03
SLC26A9	ENSG00000174502	-4.747.995.357	0.99714778	1.92E-06	4.686E-03
PCSK9	ENSG00000169174	-4.853.378.772	1.022.489.976	2.07E-06	4.686E-03
TDRD12	ENSG00000173809	478.763.916	1.010.209.045	2.15E-06	4.686E-03
CYS1	ENSG00000205795	4.339.375.354	0.924987601	2.72E-06	5.273E-03
MOBP	ENSG00000168314	-4.190.720.827	0.900739222	3.28E-06	5.731E-03
IGFL2	ENSG00000204866	4.530.689.482	0.986257754	4.35E-06	6.915E-03
AZGP1	ENSG00000160862	5.143.454.353	1.135.583.142	5.92E-06	8.369E-03
PCSK2	ENSG00000125851	-775.961.914	1.717.263.105	6.23E-06	8.369E-03
PPEF1	ENSG00000086717	-3.567.108.646	0.799582599	8.15E-06	1.017E-02
TPSB2	ENSG00000197253	3.850.160.064	0.87821647	1.16E-05	1.296E-02
GLB1L3	ENSG00000166105	664.306.336	151.717.173	1.19E-05	1.296E-02
INHA	ENSG00000123999	-5.567.018.707	1.274.820.743	1.26E-05	1.296E-02
TGFBR3	ENSG00000069702	3.018.060.592	0.710725422	2.17E-05	2.039E-02
GNLY	ENSG00000115523	-2.789.606.582	0.657637466	2.22E-05	2.039E-02
SPTSSB	ENSG00000196542	43.136.098	1.020.695.739	2.38E-05	2.077E-02
H3C10	ENSG00000278828	2.923.278.696	0.693873193	2.52E-05	2.097E-02
TEK	ENSG00000120156	2.477.108.348	0.591428689	2.81E-05	2.232E-02
OTUD7A	ENSG00000169918	-3.247.185.679	0.782736789	3.35E-05	2.519E-02
TCIM	ENSG00000176907	-3.294.626.287	0.795620932	3.46E-05	2.519E-02
DMBT1	ENSG00000187908	4.374.306.774	1.068.317.046	4.23E-05	2.957E-02
SFTPA2	ENSG00000185303	4.876.441.215	1.193.691.263	4.40E-05	2.961E-02
PIPSL	ENSG00000180764	-1.440.521.715	0.353850295	4.68E-05	2.974E-02
DSCAML1	ENSG00000177103	-4.325.716.445	1.063.633.858	4.76E-05	2.974E-02
NPW	ENSG00000183971	3.787.150.698	0.945949033	6.24E-05	3.761E-02
RASD1	ENSG00000108551	-3.485.190.723	0.87255783	6.49E-05	3.781E-02
GON4L	ENSG00000116580	-1.101.548.387	0.277412269	7.16E-05	4.039E-02
EYA2	ENSG00000064655	3.193.096.095	0.808652761	7.86E-05	4.117E-02
MFAP5	ENSG00000197614	4.044.063.792	1.025.339.928	8.01E-05	4.117E-02
CDK5R2	ENSG00000171450	-2.802.615.646	0.714039872	8.67E-05	4.293E-02
CBR1	ENSG00000159228	-2.455.250.736	0.626284996	8.84E-05	4.293E-02
NEFM	ENSG00000104722	-4.413.294.017	113.319.309	9.84E-05	4.511E-02
BOC	ENSG00000144857	3.195.195.793	0.821548537	0.00010056	4.511E-02
FZD10	ENSG00000111432	-3.410.916.247	0.877532776	0.000101516	4.511E-02
POLRMTP1	ENSG00000266066	-2.116.981.343	0.545214591	0.000103243	4.511E-02
AKR1B10	ENSG00000198074	-5.515.408.881	1.426.433.661	0.000110376	4.705E-02

Supplement table 2: Top 40 DEGs in oligo-synchronous BrM vs. poly-metastases

Gene.symbol	EnsemblID	log2FoldChange	lfcSE	pvalue	padj
FZD10	ENSG00000111432	-5.270.424.696	0.817072079	1.12E-10	0.00000195
CYP11A1	ENSG00000140459	-4.914.428.604	0.904967036	5.62E-08	0.000491085
AKR1B10	ENSG00000198074	-700.872.236	1.320.618.497	0.000000111	0.000617892
SMURF2	ENSG00000108854	1.917.728.895	0.364351261	0.000000141	0.000617892
PPEF1	ENSG00000086717	-3.785.065.492	0.739781098	0.000000311	0.001088222
SPINK2	ENSG00000128040	5.197.208.404	1.022.699.658	0.000000374	0.001088799
BCL2L14	ENSG00000121380	4.128.037.629	0.818649932	0.00000046	0.001147312
GATA2	ENSG00000179348	-3.441.385.902	0.696526779	0.000000778	0.001699964
ATP10A	ENSG00000206190	3.384.188.028	0.698500786	0.00000127	0.002459655
HEPHL1	ENSG00000181333	-4.523.212.906	0.943744478	0.00000164	0.002874077
RP11-964E11.2	ENSG00000258661	3.877.171.504	0.816107105	0.00000203	0.003219151
MT1E	ENSG00000169715	3.055.791.643	0.653781299	0.00000295	0.00394275
SYT12	ENSG00000173227	3.594.414.599	0.769552915	0.000003	0.00394275
RP11-705C15.3	ENSG00000257027	2.220.854.762	0.479047139	0.00000355	0.00394275
ASCL2	ENSG00000183734	-3.439.935.147	0.742433766	0.0000036	0.00394275
LINC01116	ENSG00000163364	3.959.506.202	0.858554701	0.00000399	0.004070673
H3P6	ENSG00000235655	-1.130.278.454	0.245626397	0.00000419	0.004070673
KCNE4	ENSG00000152049	-2.507.782.574	0.548556241	0.00000484	0.004229695
CBR1	ENSG00000159228	-2.638.409.035	0.579311342	0.00000525	0.004354073
ACKR3	ENSG00000144476	-282.655.847	0.623485733	0.0000058	0.004354073
ITGA10	ENSG00000143127	335.077.899	0.739957591	0.00000594	0.004354073
DIO2	ENSG00000211448	-294.267.881	0.650008775	0.00000598	0.004354073
RNF150	ENSG00000170153	-3.402.659.693	0.754318896	0.00000646	0.004512967
SPRR2D	ENSG00000163216	6.124.630.305	1.370.138.869	0.00000782	0.004963545
RSU1	ENSG00000148484	-0.871381966	0.194946371	0.00000783	0.004963545
GLTPD2	ENSG00000182327	-3.571.765.485	0.799682212	0.00000795	0.004963545
HLA-G	ENSG00000204632	379.666.615	0.854465222	0.00000886	0.005180539
TLR3	ENSG00000164342	2.315.724.465	0.521266454	0.00000889	0.005180539
BPIFB1	ENSG00000125999	-5.998.796.164	1.353.264.342	0.0000093	0.00524354
SEMA3A	ENSG00000075213	3.219.128.218	0.729498958	0.0000102	0.005404693
RARG	ENSG00000172819	2.083.424.434	0.474630854	0.0000114	0.005477062
DNMT3B	ENSG00000088305	-2.645.978.474	0.6030847	0.0000115	0.005477062
ALPL	ENSG00000162551	-3.577.638.825	0.815750362	0.0000116	0.005477062
DNAJC22	ENSG00000178401	26.307.456	0.602935275	0.0000128	0.005692752
SPATA18	ENSG00000163071	-3.153.625.865	0.722925356	0.0000129	0.005692752
HS3ST3B1	ENSG00000125430	-3.462.217.082	0.79491391	0.0000133	0.005692752
SOX7	ENSG00000171056	-2.558.657.541	0.587626935	0.0000134	0.005692752
ADAMTSL1	ENSG00000178031	-3.518.932.599	0.81452621	0.0000156	0.006486567
COL26A1	ENSG00000160963	-30.262.734	0.704158267	0.0000173	0.006854784
KSR1	ENSG00000141068	-2.225.752.935	0.518902369	0.0000179	0.00696006

Supplement table 3: Top 40 DEGs in oligo-metachronous BrM vs. poly-metastases

Gene.symbol	EnsemblID	log2FoldChange	lfcSE	pvalue	padj
SCGB3A1	ENSG00000161055	-7.416.251.206	1.129.673.023	5.2E-11	0.00000091
LXN	ENSG00000079257	-2.932.180.295	0.543760336	6.95E-08	0.000607516
S100A2	ENSG00000196754	3.441.742.575	0.676925993	0.000000369	0.001611678
FNDC1	ENSG00000164694	-3.428.407.431	0.681464177	0.000000488	0.001706109
DSCAML1	ENSG00000177103	5.032.498.069	1.020.048.479	0.000000807	0.002351752
TPSB2	ENSG00000197253	-4.026.914.358	0.839365142	0.00000161	0.004009973
TGFBR3	ENSG00000069702	-3.190.233.703	0.679603949	0.00000268	0.005845978
EYA2	ENSG00000064655	-3.597.172.393	0.773720328	0.00000333	0.00647156
L1CAM	ENSG00000198910	383.288.717	0.845134414	0.00000575	0.009384673
MFAP5	ENSG00000197614	-4.443.907.318	0.981056964	0.00000591	0.009384673
KLF7	ENSG00000118263	1.406.182.924	0.313796899	0.00000742	0.01031395
C8orf34-AS1	ENSG00000248801	-6.359.286.513	1.421.338.245	0.00000767	0.01031395
FCGRT	ENSG00000104870	-1.384.766.386	0.311982558	0.00000905	0.011303909
KCNJ5	ENSG00000120457	-283.461.907	0.642647134	0.0000103	0.011545691
FMO1	ENSG00000010932	-3.214.506.786	0.729709771	0.0000106	0.011545691
LYVE1	ENSG00000133800	-3.383.899.161	0.77193752	0.0000117	0.011999215
TEK	ENSG00000120156	-2.460.759.533	0.565146999	0.0000134	0.012075448
GLT8D2	ENSG00000120820	-2.251.483.655	0.518187197	0.0000139	0.012075448
NYAP2	ENSG00000144460	3.713.395.246	0.857290731	0.0000148	0.012075448
MYBPH	ENSG00000133055	4.896.114.881	113.304.779	0.0000155	0.012075448
HLA-G	ENSG00000204632	380.671.654	0.883264635	0.0000163	0.012075448
MYRF	ENSG00000124920	2.776.965.386	0.645120546	0.0000167	0.012075448
CLEC14A	ENSG00000176435	-1.626.824.566	0.378376682	0.0000171	0.012075448
CYP2T1P	ENSG00000233622	-2.872.412.006	0.668389051	0.0000173	0.012075448
STRN3	ENSG00000196792	1.128.813.113	0.265276806	0.0000209	0.012923009
TMEM156	ENSG00000121895	333.109.813	0.78325188	0.0000211	0.012923009
TDRD12	ENSG00000173809	-4.067.988.964	0.957863517	0.0000217	0.012923009
PRDM6	ENSG00000061455	-2.366.315.174	0.558635287	0.0000228	0.012923009
SPINK2	ENSG00000128040	4.462.312.845	1.053.835.659	0.0000229	0.012923009
BZW1	ENSG00000082153	1.050.901.471	0.249077996	0.0000245	0.013119979
LINC02701	ENSG00000250508	-3.754.539.467	0.890361691	0.0000248	0.013119979
RAMP2	ENSG00000131477	-1.456.848.096	0.347950463	0.0000283	0.014534455
CPA3	ENSG00000163751	-427.939.442	1.028.327.285	0.0000316	0.015787626
AJM1	ENSG00000232434	1.709.959.149	0.412698719	0.0000342	0.016616471
RNF157	ENSG00000141576	2.610.041.084	0.634021497	0.0000384	0.017813461
PDCL3P4	ENSG00000244119	3.327.625.153	0.809735413	0.0000396	0.017813461
EXTL1	ENSG00000158008	281.196.151	0.684352637	0.0000397	0.017813461
CORIN	ENSG00000145244	-2.964.117.487	0.723489773	0.0000419	0.01816271
COL11A1	ENSG00000060718	-3.166.795.637	0.773734245	0.0000426	0.01816271
LMX1B	ENSG00000136944	-391.197.774	0.957688601	0.0000441	0.018357121

Supplement table 4: Top 40 DEGs in oligo-synchronous or -metachronous BrM vs. poly-metastases

Gene.symbol	EnsemblID	log2FoldChange	lfcSE	pvalue	padj
SPRR2D	ENSG00000163216	5.520896019	0.933668927	3.35688E-09	5.35905E-05
RSPO4	ENSG00000101282	-5.069386252	0.876322795	7.25818E-09	5.35905E-05
CDH18	ENSG00000145526	5.21782473	0.908256211	9.19851E-09	5.35905E-05
AKR1C1	ENSG00000187134	-6.236576792	1.09585562	1.26255E-08	5.51673E-05
SPINK2	ENSG00000128040	4.805292743	0.87193327	3.56642E-08	0.000124668
SCGB3A1	ENSG00000161055	-6.351084063	1.174840051	6.44788E-08	0.000187827
FZD10	ENSG00000111432	-3.925272534	0.7424648	1.24466E-07	0.000310773
HLA-G	ENSG00000204632	3.801475921	0.724185188	1.52666E-07	0.000333537
C8orf34-AS1	ENSG00000248801	-5.807316623	1.111718437	1.75355E-07	0.000340539
COL26A1	ENSG00000160963	-2.967535512	0.601322973	8.01462E-07	0.001400796
ZPLD1	ENSG00000170044	3.570269669	0.751904179	2.05129E-06	0.002987711
ACKR2	ENSG00000144648	-3.58132342	0.764507641	2.80667E-06	0.003085176
H3P6	ENSG00000235655	-0.996160581	0.212756015	2.83849E-06	0.003085176
STRN3	ENSG00000196792	1.022420869	0.218789991	2.96731E-06	0.003085176
RSU1	ENSG00000148484	-0.779677012	0.167683234	3.32416E-06	0.003085176
SMURF2	ENSG00000108854	1.51682431	0.32694021	3.4933E-06	0.003085176
LMX1B	ENSG00000136944	-3.685144338	0.794598725	3.52235E-06	0.003085176
BZW1	ENSG00000082153	0.954478182	0.205827455	3.53035E-06	0.003085176
ALDH3A1	ENSG00000108602	-4.093984427	0.886142305	3.83725E-06	0.003193692
FNDC1	ENSG00000164694	-2.768367712	0.603912949	4.56053E-06	0.003623134
NINJ1	ENSG00000131669	-1.142805721	0.253309125	6.43689E-06	0.004689736
ACKR3	ENSG00000144476	-2.455859019	0.544365871	6.43973E-06	0.004689736
S100A2	ENSG00000196754	2.604547852	0.579733796	7.03385E-06	0.004917502
L1CAM	ENSG00000198910	3.141080305	0.706037731	8.63174E-06	0.005802523
OAS3	ENSG00000111331	1.513388599	0.342500186	9.93191E-06	0.006143497
ANO3	ENSG00000134343	-3.261955126	0.738305744	9.95433E-06	0.006143497
CORIN	ENSG00000145244	-2.672331863	0.605555115	1.01935E-05	0.006143497
SLF1	ENSG00000133302	1.051003855	0.238807393	1.07727E-05	0.006276181
ASCL2	ENSG00000183734	-2.892611289	0.662162	1.2514E-05	0.006625491
MESP1	ENSG00000166823	-2.146179332	0.491632444	1.26882E-05	0.006625491
SOX7	ENSG00000171056	-1.780826881	0.408260367	1.28886E-05	0.006625491
RAMP2	ENSG00000131477	-1.263722697	0.291452492	1.45134E-05	0.00681481
CLU	ENSG00000120885	-2.29018516	0.529308758	1.51328E-05	0.00681481
MFAP5	ENSG00000197614	-4.185666435	0.96818449	1.53774E-05	0.00681481
DNAJC22	ENSG00000178401	2.245969221	0.519551724	1.53993E-05	0.00681481
MT1A	ENSG00000205362	3.478766486	0.805060006	1.55234E-05	0.00681481
AJM1	ENSG00000232434	1.479295886	0.342422198	1.55963E-05	0.00681481
DLX5	ENSG00000105880	-2.821483457	0.656246822	1.71242E-05	0.007126132
SPATA18	ENSG00000163071	-2.72668764	0.635076538	1.75895E-05	0.007149523
PCDHB3	ENSG00000113205	2.740772606	0.642164473	1.97213E-05	0.007833845

Supplement table 5: Top 40 DEGs in oligo-synchronous vs. oligo-metachronous BrM or poly-metastases

Gene.symbol	EnsemblID	log2FoldChange	lfcSE	pvalue	padj
PPEF1	ENSG00000086717	-3.689.543.155	0.655338374	0.000000018	0.000315004
AKR1B10	ENSG00000198074	-6.189.483.015	1.194.752.211	0.000000221	0.00193327
PCSK2	ENSG00000125851	-6.074.722.175	12.208.483	0.00000065	0.002267758
CBR1	ENSG00000159228	-2.558.443.143	0.51538073	0.00000069	0.002267758
SP8	ENSG00000164651	448.355.459	0.904305403	0.000000712	0.002267758
SPRR2D	ENSG00000163216	5.670.569.921	1.162.709.595	0.00000108	0.002551215
LINC00942	ENSG00000249628	-4.218.645.903	0.871351657	0.00000129	0.002551215
LINC01116	ENSG00000163364	3.729.034.629	0.770918032	0.00000132	0.002551215
BCL2L14	ENSG00000121380	3.667.327.249	0.763571002	0.00000156	0.002551215
DMBT1	ENSG00000187908	4.556.932.233	0.952356618	0.00000171	0.002551215
SFTPA1	ENSG00000122852	514.449.909	1.076.220.284	0.00000175	0.002551215
UPK3B	ENSG00000243566	3.736.164.864	0.786337996	0.00000202	0.002716697
GLB1L3	ENSG00000166105	5.110.670.986	1.081.580.707	0.0000023	0.002870185
MAST4-AS1	ENSG00000229666	3.031.540.699	0.652568009	0.00000339	0.003951928
SCGB3A1	ENSG00000161055	5.945.651.122	1.291.365.479	0.00000414	0.004320699
RASD1	ENSG00000108551	-330.350.415	0.719758273	0.00000444	0.004320699
SLC1A7	ENSG00000162383	4.695.709.973	1.025.737.282	0.0000047	0.004320699
SYT12	ENSG00000173227	317.679.537	0.706446324	0.0000069	0.006021497
GJB6	ENSG00000121742	4.090.548.794	0.911713885	0.00000723	0.006021497
CYP11A1	ENSG00000140459	-3.746.482.081	0.840304974	0.00000825	0.006381424
TMEM100	ENSG00000166292	3.752.895.857	0.843503329	0.00000862	0.006381424
F5	ENSG00000198734	3.066.974.671	0.689882289	0.00000876	0.006381424
POLRMTP1	ENSG00000266066	-195.751.192	0.444643895	0.0000107	0.007483871
GLTPD2	ENSG00000182327	-3.121.683.101	0.71966558	0.0000144	0.009106018
DDO	ENSG00000203797	3.118.566.155	0.719422355	0.0000146	0.009106018
ATP10A	ENSG00000206190	2.871.968.918	0.664139355	0.0000153	0.009214408
DIO2	ENSG00000211448	-2.542.405.301	0.58918986	0.000016	0.009214408
KCNE4	ENSG00000152049	-2.148.688.729	0.498563476	0.0000163	0.009214408
ZNF687	ENSG00000143373	-1.234.010.533	0.287310059	0.0000175	0.009501558
ACADL	ENSG00000115361	43.541.708	1.015.170.961	0.0000179	0.009501558
SFTPC	ENSG00000168484	5.833.424.583	1.375.346.493	0.0000222	0.011417672
MT1E	ENSG00000169715	2.598.491.625	0.615348741	0.0000241	0.012049394
PIPSL	ENSG00000180764	-1.182.320.532	0.280500794	0.000025	0.012124876
SECISBP2L	ENSG00000138593	-1.153.830.257	0.275551948	0.0000282	0.012676677
SEMA6C	ENSG00000143434	-1.499.958.541	0.358248894	0.0000283	0.012676677
H3C10	ENSG00000278828	2.458.941.391	0.587302375	0.0000283	0.012676677
CYP4B1	ENSG00000142973	-3.984.807.346	0.955160372	0.0000302	0.013200548
SMURF2	ENSG00000108854	153.367.507	0.368694628	0.0000319	0.013581908
LGMN	ENSG00000100600	1.325.583.936	0.319208988	0.0000329	0.013672536
IGFL2	ENSG00000204866	3.641.482.784	0.87840664	0.0000339	0.013778844

9. References

1. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. **71**(3): p. 209-249.
2. Quail, D.F. and J.A. Joyce, *The Microenvironmental Landscape of Brain Tumors.* Cancer Cell, 2017. **31**(3): p. 326-341.
3. Ashrafi, A., et al., *Current landscape of therapeutic resistance in lung cancer and promising strategies to overcome resistance.* Cancers, 2022. **14**(19): p. 4562.
4. Schabath, M.B. and M.L. Cote, *Cancer Progress and Priorities: Lung Cancer.* Cancer Epidemiol Biomarkers Prev, 2019. **28**(10): p. 1563-1579.
5. Miller, K.D., et al., *Cancer treatment and survivorship statistics, 2022.* CA Cancer J Clin, 2022. **72**(5): p. 409-436.
6. Society, A.C. *Lung Cancer Risk Factors*. Available from: https://www.cancer.org/cancer/lung-cancer/causes-risks-prevention/risk-factors.html Accessed February 22, 2023.
7. Scagliotti, G.V., M. Longo, and S. Novello, *Nonsmall cell lung cancer in never smokers.* Curr Opin Oncol, 2009. **21**(2): p. 99-104.
8. de Alencar, V.T.L., et al., *Lung cancer in never smokers: Tumor immunology and challenges for immunotherapy.* Front Immunol, 2022. **13**: p. 984349.
9. Chalela, R., et al., *Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy.* Journal of thoracic disease, 2017. **9**(7): p. 2142.
10. Barta, J.A., C.A. Powell, and J.P. Wisnivesky, *Global Epidemiology of Lung Cancer.* Ann Glob Health, 2019. **85**(1).
11. Wood, S.L., et al., *The role of the tumor-microenvironment in lung cancer-metastasis and its relationship to potential therapeutic targets.* Cancer Treat Rev, 2014. **40**(4): p. 558-66.
12. Davidson, M.R., A.F. Gazdar, and B.E. Clarke, *The pivotal role of pathology in the management of lung cancer.* J Thorac Dis, 2013. **5 Suppl 5**(Suppl 5): p. S463-78.
13. Lim, W., et al., *The 8th lung cancer TNM classification and clinical staging system: review of the changes and clinical implications.* Quantitative imaging in medicine and surgery, 2018. **8**(7): p. 709.
14. Brierley, J.D., M.K. Gospodarowicz, and C. Wittekind, *TNM classification of malignant tumours*. 2017: John Wiley & Sons.
15. Society, A.C. *Treatment Choices for Non-Small Cell Lung Cancer, by Stage*. Available from: https://www.cancer.org/cancer/lung-cancer/treating-non-small-cell/by-stage.html Accessed February 22, 2023.
16. Group, N.-s.C.L.C.C., *Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials.* Bmj, 1995. **311**(7010): p. 899-909.
17. Zappa, C. and S.A. Mousa, *Non-small cell lung cancer: current treatment and future advances.* Translational lung cancer research, 2016. **5**(3): p. 288.
18. Alexander, M., S.Y. Kim, and H. Cheng, *Update 2020: management of non-small cell lung cancer.* Lung, 2020. **198**: p. 897-907.
19. Jianguo, Z. and J. Xiong, *Advances on driver oncogenes of non-small cell lung cancer.* Zhongguo Fei Ai Za Zhi, 2015. **18**(1).
20. Walsh, G.L. and R.J. Winn, *Baseline institutional compliance with NCCN guidelines: non-small-cell lung cancer.* Oncology (Williston Park, NY), 1997. **11**(11A): p. 161-170.
21. Facchinetti, F., et al., *First-line immunotherapy in non-small cell lung cancer patients with poor performance status: a systematic review and meta-analysis.* Translational Lung Cancer Research, 2021. **10**(6): p. 2917.

22. Novoa, N.M., G. Varela, and M.F. Jiménez, *Surgical management of oligometastatic non-small cell lung cancer.* Journal of thoracic disease, 2016. **8**(Suppl 11): p. S895.
23. Varella-Garcia, M., *Chromosomal and genomic changes in lung cancer.* Cell adhesion & migration, 2010. **4**(1): p. 100-106.
24. Marino, F.Z., et al., *Molecular heterogeneity in lung cancer: from mechanisms of origin to clinical implications.* International journal of medical sciences, 2019. **16**(7): p. 981.
25. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer.* Nature, 2013. **500**(7463): p. 415-421.
26. Kruglyak, K.M., E. Lin, and F.S. Ong, *Next-generation sequencing and applications to the diagnosis and treatment of lung cancer.* Lung Cancer and Personalized Medicine: Novel Therapies and Clinical Management, 2016: p. 123-136.
27. Samstein, R.M., et al., *Tumor mutational load predicts survival after immunotherapy across multiple cancer types.* Nature genetics, 2019. **51**(2): p. 202-206.
28. Zhao, T., et al., *Prediction and collection of protein–metabolite interactions.* Briefings in Bioinformatics, 2021. **22**(5): p. bbab014.
29. Nana-Sinkam, S.P. and C.A. Powell, *Molecular biology of lung cancer: Diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines.* Chest, 2013. **143**(5): p. e30S-e39S.
30. Network, C.G.A.R., *Comprehensive genomic characterization of squamous cell lung cancers.* Nature, 2012. **489**(7417): p. 519.
31. Network, C.G.A.R., *Comprehensive molecular profiling of lung adenocarcinoma.* Nature, 2014. **511**(7511): p. 543.
32. Drilon, A., et al., *Squamous-cell carcinomas of the lung: emerging biology, controversies, and the promise of targeted therapy.* The lancet oncology, 2012. **13**(10): p. e418-e426.
33. Project, C.L.C.G. and Network Genomic Medicine *A genomics-based classification of human lung tumors.* Science translational medicine, 2013. **5**(209): p. 209ra153-209ra153.
34. Dong, J., et al., *Advances in targeted therapy and immunotherapy for non-small cell lung cancer based on accurate molecular typing.* Frontiers in pharmacology, 2019. **10**: p. 230.
35. Shaw, A.T., et al., *Crizotinib in ROS1-rearranged non–small-cell lung cancer.* New England Journal of Medicine, 2014. **371**(21): p. 1963-1971.
36. Kwak, E.L., et al., *Anaplastic lymphoma kinase inhibition in non–small-cell lung cancer.* New England Journal of Medicine, 2010. **363**(18): p. 1693-1703.
37. Gower, A., Y. Wang, and G. Giaccone, *Oncogenic drivers, targeted therapies, and acquired resistance in non-small-cell lung cancer.* Journal of molecular medicine, 2014. **92**: p. 697-707.
38. Skoulidis, F., et al., *Sotorasib for lung cancers with KRAS p. G12C mutation.* New England Journal of Medicine, 2021. **384**(25): p. 2371-2381.
39. Mitsudomi, T., et al., *Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial.* The lancet oncology, 2010. **11**(2): p. 121-128.
40. Maemondo, M., et al., *Gefitinib or chemotherapy for non–small-cell lung cancer with mutated EGFR.* New England Journal of Medicine, 2010. **362**(25): p. 2380-2388.
41. Tan, D.S., et al., *The international association for the study of lung cancer consensus statement on optimizing management of EGFR Mutation–Positive Non–Small cell lung cancer: Status in 2016.* Journal of Thoracic Oncology, 2016. **11**(7): p. 946-963.

42. Sequist, L.V., et al., *Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors.* Science translational medicine, 2011. **3**(75): p. 75ra26-75ra26.
43. Yu, H.A., et al., *Analysis of Tumor Specimens at the Time of Acquired Resistance to EGFR-TKI Therapy in 155 Patients with EGFR-Mutant Lung CancersMechanisms of Acquired Resistance to EGFR-TKI Therapy.* Clinical cancer research, 2013. **19**(8): p. 2240-2247.
44. Kuiper, J., et al., *Incidence of T790M mutation in (sequential) rebiopsies in EGFR-mutated NSCLC-patients.* Lung cancer, 2014. **85**(1): p. 19-24.
45. Friboulet, L., et al., *The ALK Inhibitor Ceritinib Overcomes Crizotinib Resistance in Non–Small Cell Lung CancerCeritinib and Crizotinib Resistance.* Cancer discovery, 2014. **4**(6): p. 662-673.
46. Gadgeel, S.M., et al., *Safety and activity of alectinib against systemic disease and brain metastases in patients with crizotinib-resistant ALK-rearranged non-small-cell lung cancer (AF-002JG): results from the dose-finding portion of a phase 1/2 study.* The Lancet Oncology, 2014. **15**(10): p. 1119-1128.
47. Shaw, A.T., et al., *Lorlatinib in non-small-cell lung cancer with ALK or ROS1 rearrangement: an international, multicentre, open-label, single-arm first-in-man phase 1 trial.* The Lancet Oncology, 2017. **18**(12): p. 1590-1599.
48. Hammerman, P.S., et al., *Mutations in the DDR2 Kinase Gene Identify a Novel Therapeutic Target in Squamous Cell Lung CancerDDR2 Mutations in SCC of the Lung.* Cancer discovery, 2011. **1**(1): p. 78-89.
49. Dutt, A., et al., *Inhibitor-sensitive FGFR1 amplification in human non-small cell lung cancer.* PloS one, 2011. **6**(6): p. e20351.
50. Niu, Z., et al., *Signaling pathways and targeted therapies in lung squamous cell carcinoma: mechanisms and clinical trials.* Signal Transduction and Targeted Therapy, 2022. **7**(1): p. 353.
51. Rodak, O., et al., *Current landscape of non-small cell lung cancer: epidemiology, histological classification, targeted therapies, and immunotherapy.* Cancers, 2021. **13**(18): p. 4705.
52. Boumahdi, S. and F.J. de Sauvage, *The great escape: tumour cell plasticity in resistance to targeted therapy.* Nature reviews Drug discovery, 2020. **19**(1): p. 39-56.
53. Fares, J., et al., *Molecular principles of metastasis: a hallmark of cancer revisited.* Signal transduction and targeted therapy, 2020. **5**(1): p. 28.
54. Dillekås, H., M.S. Rogers, and O. Straume, *Are 90% of deaths from cancer caused by metastases?* Cancer medicine, 2019. **8**(12): p. 5574-5576.
55. Society, A.C. *Signs and Symptoms of Lung Cancer*. Available from: https://www.cancer.org/cancer/lung-cancer/detection-diagnosis-staging/signs-symptoms.html Accessed February 23, 2023.
56. Tamura, T., et al., *Specific organ metastases and survival in metastatic non-small-cell lung cancer.* Molecular and clinical oncology, 2015. **3**(1): p. 217-221.
57. Lawson, D.A., et al., *Tumour heterogeneity and metastasis at single-cell resolution.* Nature cell biology, 2018. **20**(12): p. 1349-1360.
58. Wu, Y., et al., *Driver and novel genes correlated with metastasis of non-small cell lung cancer: A comprehensive analysis.* Pathology-Research and Practice, 2021. **224**: p. 153551.
59. Hess, K.R., et al., *Metastatic patterns in adenocarcinoma.* Cancer, 2006. **106**(7): p. 1624-1633.
60. Wang, L., et al., *The Prognostic value of baseline distant metastasis in icotinib-treated patients with EGFR-mutated stage IV non-small cell lung cancer.* Cancer Management and Research, 2021: p. 2613-2622.

61. Wang, X., et al., *Patterns of extrathoracic metastases in different histological types of lung cancer.* Frontiers in Oncology, 2020. **10**: p. 715.
62. Sato, H., et al., *Analysis of bevacizumab treatments and metastatic sites of lung cancer.* Cancer Treatment and Research Communications, 2021. **26**: p. 100290.
63. Paget, S., *The distribution of secondary growths in cancer of the breast.* Cancer Metastasis Rev, 1989. **8**: p. 98-101.
64. Kaplan, R.N., et al., *VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche.* Nature, 2005. **438**(7069): p. 820-827.
65. Hiratsuka, S., et al., *MMP9 induction by vascular endothelial growth factor receptor-1 is involved in lung-specific metastasis.* Cancer cell, 2002. **2**(4): p. 289-300.
66. Sowder, M.E. and R.W. Johnson, *Bone as a preferential site for metastasis.* JBMR plus, 2019. **3**(3): p. e10126.
67. Ewing, J., *Neoplastic diseases*. 1922: Wb saunders.
68. Duda, D.G., et al., *Malignant cells facilitate lung metastasis by bringing their own soil.* Proceedings of the National Academy of Sciences, 2010. **107**(50): p. 21677-21682.
69. Doebele, R.C., et al., *Oncogene status predicts patterns of metastatic spread in treatment-naive nonsmall cell lung cancer.* Cancer, 2012. **118**(18): p. 4502-4511.
70. Han, G., et al., *A retrospective analysis in patients with EGFR-mutant lung adenocarcinoma: is EGFR mutation associated with a higher incidence of brain metastasis?* Oncotarget, 2016. **7**(35): p. 56998.
71. Bhatt, V.R., et al., *Epidermal growth factor receptor mutational status and brain metastases in non–small-cell lung cancer.* Journal of global oncology, 2017. **3**(3): p. 208-217.
72. Fox, B.D., et al., *Epidemiology of metastatic brain tumors.* Neurosurgery Clinics, 2011. **22**(1): p. 1-6.
73. Lee, J.S., et al., *The impact of systemic treatment on brain metastasis in patients with non-small-cell lung cancer: a retrospective nationwide population-based cohort study.* Scientific Reports, 2019. **9**(1): p. 18689.
74. Myall, N.J., et al., *Management of brain metastases in lung cancer: evolving roles for radiation and systemic treatment in the era of targeted and immune therapies.* Neuro-Oncology Advances, 2021. **3**(Supplement_5): p. v52-v62.
75. Abrahams, J.M., et al., *Risk factors affecting survival after brain metastases from non—small cell lung carcinoma: a follow-up study of 70 patients.* Journal of neurosurgery, 2001. **95**(4): p. 595-600.
76. Villano, J.L., et al., *Incidence of brain metastasis at initial presentation of lung cancer.* Neuro Oncol, 2015. **17**(1): p. 122-8.
77. Schiller, J.H., et al., *Comparison of four chemotherapy regimens for advanced non–small-cell lung cancer.* New England Journal of Medicine, 2002. **346**(2): p. 92-98.
78. Ali, A., et al., *Survival of patients with non-small-cell lung cancer after a diagnosis of brain metastases.* Current Oncology, 2013. **20**(4): p. 300-306.
79. Tominaga, N., et al., *Brain metastatic cancer cells release microRNA-181c-containing extracellular vesicles capable of destructing blood–brain barrier.* Nature communications, 2015. **6**(1): p. 6716.
80. Fares, J., et al., *Emerging principles of brain immunology and immune checkpoint blockade in brain metastases.* Brain, 2021. **144**(4): p. 1046-1066.
81. Eichler, A.F., et al., *The biology of brain metastases—translation to new therapies.* Nature reviews Clinical oncology, 2011. **8**(6): p. 344-356.
82. Barnholtz-Sloan, J.S., et al., *Incidence proportions of brain metastases in patients diagnosed (1973 to 2001) in the Metropolitan Detroit Cancer Surveillance System.* Journal of clinical oncology, 2004. **22**(14): p. 2865-2872.

83. Quint, L.E., et al., *Distribution of distant metastases from newly diagnosed non-small cell lung cancer.* Ann Thorac Surg, 1996. **62**(1): p. 246-50.
84. He, J., et al., *Risk factors for brain metastases from non-small-cell lung cancer: a protocol for observational study.* Medicine, 2021. **100**(9).
85. Waqar, S.N., et al., *Non–small-cell lung cancer with brain metastasis at presentation.* Clinical lung cancer, 2018. **19**(4): p. e373-e379.
86. Habbous, S., et al., *Incidence and real-world burden of brain metastases from solid tumors and hematologic malignancies in Ontario: a population-based study.* Neuro-Oncology Advances, 2021. **3**(1): p. vdaa178.
87. Jablonska, P.A., et al., *Challenges and novel opportunities of radiation therapy for brain metastases in non-small cell lung cancer.* Cancers, 2021. **13**(9): p. 2141.
88. Palmieri, D., et al., *The biology of metastasis to a sanctuary site.* Clinical Cancer Research, 2007. **13**(6): p. 1656-1662.
89. Ernani, V. and T.E. Stinchcombe, *Management of brain metastases in non–small-cell lung cancer.* Journal of oncology practice, 2019. **15**(11): p. 563-570.
90. Hellman, S. and R.R. Weichselbaum, *Oligometastases.* Journal of Clinical Oncology, 1995. **13**(1): p. 8-10.
91. Weichselbaum, R.R. and S. Hellman, *Oligometastases revisited.* Nature reviews Clinical oncology, 2011. **8**(6): p. 378-382.
92. Chaung, K.V., A.E. Sloan, and S. Choi, *Limited brain metastases: a narrative review.* Ann Palliat Med, 2021. **10**(5): p. 6016-6027.
93. Al-Shafa, F., et al., *A review of ongoing trials of stereotactic ablative radiotherapy for oligometastatic cancers: Where will the evidence lead?* Frontiers in Oncology, 2019. **9**: p. 543.
94. Shi, A.A., et al., *Does initial staging or tumor histology better identify asymptomatic brain metastases in patients with non–small cell lung cancer?* Journal of Thoracic Oncology, 2006. **1**(3): p. 205-210.
95. Chen, A.M., et al., *Risk of cerebral metastases and neurological death after pathological complete response to neoadjuvant therapy for locally advanced nonsmall-cell lung cancer: clinical implications for the subsequent management of the brain.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2007. **109**(8): p. 1668-1675.
96. Hung, J.-J., et al., *Prognostic factors of postrecurrence survival in completely resected stage I non-small cell lung cancer with distant metastasis.* Thorax, 2010. **65**(3): p. 241-245.
97. Palma, D.A., et al., *The oligometastatic state—separating truth from wishful thinking.* Nature reviews Clinical oncology, 2014. **11**(9): p. 549-557.
98. Huang, R.S., et al., *Clinicopathologic and genomic landscape of non-small cell lung cancer brain metastases.* The Oncologist, 2022. **27**(10): p. 839-848.
99. De Ruysscher, D., et al., *Radical treatment of non–small-cell lung cancer patients with synchronous oligometastases: long-term results of a prospective phase II trial (Nct01282450).* Journal of Thoracic Oncology, 2012. **7**(10): p. 1547-1555.
100. Bauman, G.S., et al., *Ablative radiation therapy to restrain everything safely treatable (ARREST): Study protocol for a phase I trial treating polymetastatic cancer with stereotactic radiotherapy.* BMC cancer, 2021. **21**(1): p. 1-10.
101. Ashworth, A.B., et al., *An individual patient data metaanalysis of outcomes and prognostic factors after treatment of oligometastatic non–small-cell lung cancer.* Clinical lung cancer, 2014. **15**(5): p. 346-355.
102. Fleckenstein, J., et al., *Long-term outcomes in radically treated synchronous vs. metachronous oligometastatic non-small-cell lung cancer.* BMC cancer, 2016. **16**(1): p. 1-8.

103. Sharma, A., et al., *Survival and prognostic factors of pulmonary oligometastases treated with stereotactic body radiotherapy.* Acta Oncologica, 2019. **58**(1): p. 74-80.
104. Getman, V., et al., *Prognosis of patients with non-small cell lung cancer with isolated brain metastases undergoing combined surgical treatment.* European journal of cardio-thoracic surgery, 2004. **25**(6): p. 1107-1113.
105. Burt, M., et al., *Resection of brain metastases from non-small-cell lung carcinoma: results of therapy.* The Journal of thoracic and cardiovascular surgery, 1992. **103**(3): p. 399-411.
106. Gobbini, E., et al., *Epidemiology of oligometastatic non-small cell lung cancer: results from a systematic review and pooled analysis.* Translational Lung Cancer Research, 2021. **10**(7): p. 3339.
107. Uppal, A., et al., *Towards a molecular basis of oligometastatic disease: potential role of micro-RNAs.* Clinical & experimental metastasis, 2014. **31**: p. 735-748.
108. Xiao, G., et al., *Heterogeneity of tumor immune microenvironment of EGFR/ALK-positive tumors versus EGFR/ALK-negative tumors in resected brain metastases from lung adenocarcinoma.* Journal for Immunotherapy of Cancer, 2023. **11**(3).
109. Álvaro-Espinosa, L., et al., *Brain microenvironment heterogeneity: Potential value for brain tumors.* Frontiers in Oncology, 2021. **11**: p. 714428.
110. Shih, D.J., et al., *Genomic characterization of human brain metastases identifies drivers of metastatic lung adenocarcinoma.* Nature genetics, 2020. **52**(4): p. 371-377.
111. Brastianos, P.K., et al., *Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic TargetsGenomics of Brain Metastases.* Cancer discovery, 2015. **5**(11): p. 1164-1177.
112. Sperduto, P.W., et al., *The effect of gene alterations and tyrosine kinase inhibition on survival and cause of death in patients with adenocarcinoma of the lung and brain metastases.* International Journal of Radiation Oncology* Biology* Physics, 2016. **96**(2): p. 406-413.
113. Dono, A., et al., *Differences in genomic alterations between brain metastases and primary tumors.* Neurosurgery, 2021. **88**(3): p. 592-602.
114. Wang, H., et al., *Driver genes as predictive indicators of brain metastasis in patients with advanced NSCLC: EGFR, ALK, and RET gene mutations.* Cancer medicine, 2020. **9**(2): p. 487-495.
115. Tomasini, P., et al., *EGFR and KRAS mutations predict the incidence and outcome of brain metastases in non-small cell lung cancer.* International journal of molecular sciences, 2016. **17**(12): p. 2132.
116. Jiang, T., et al., *Characterization of evolution trajectory and immune profiling of brain metastasis in lung adenocarcinoma.* NPJ Precision Oncology, 2021. **5**(1): p. 6.
117. Song, Z., et al., *Genomic profiles and tumor immune microenvironment of primary lung carcinoma and brain oligo-metastasis.* Cell Death & Disease, 2021. **12**(1): p. 106.
118. Stella, G.M., et al., *MET mutations are associated with aggressive and radioresistant brain metastatic non-small-cell lung cancer.* Neuro-oncology, 2016. **18**(4): p. 598-599.
119. Stella, G.M., et al., *Brain metastases from lung cancer: Is MET an actionable target?* Cancers, 2019. **11**(3): p. 271.
120. Wang, H., et al., *Genes associated with increased brain metastasis risk in non–small cell lung cancer: Comprehensive genomic profiling of 61 resected brain metastases versus primary non–small cell lung cancer (Guangdong Association Study of Thoracic Oncology 1036).* Cancer, 2019. **125**(20): p. 3535-3544.
121. Nguyen, D.X., et al., *WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis.* Cell, 2009. **138**(1): p. 51-62.

122. Ikarashi, D., et al., *Comparison of Tumor Microenvironments Between Primary Tumors and Brain Metastases in Patients With NSCLC.* JTO Clin Res Rep, 2021. **2**(10): p. 100230.
123. Hald, S.M., et al., *CD4/CD8 co-expression shows independent prognostic impact in resected non-small cell lung cancer patients treated with adjuvant radiotherapy.* Lung cancer, 2013. **80**(2): p. 209-215.
124. Tomar, N. and R.K. De, *A brief outline of the immune system.* Immunoinformatics, 2014: p. 3-12.
125. Marshall, J.S., et al., *An introduction to immunology and immunopathology.* Allergy, Asthma & Clinical Immunology, 2018. **14**(2): p. 1-10.
126. Bremnes, R.M., et al., *The role of tumor-infiltrating immune cells and chronic inflammation at the tumor site on cancer development, progression, and prognosis: emphasis on non-small cell lung cancer.* Journal of Thoracic Oncology, 2011. **6**(4): p. 824-833.
127. Hanahan, D., *Hallmarks of cancer: new dimensions.* Cancer discovery, 2022. **12**(1): p. 31-46.
128. Dunn, G.P., et al., *Cancer immunoediting: from immunosurveillance to tumor escape.* Nat Immunol, 2002. **3**(11): p. 991-8.
129. Kim, R., M. Emi, and K. Tanabe, *Cancer immunoediting from immune surveillance to immune escape Immunology 121: 1–14.* 2007.
130. Ostroumov, D., et al., *CD4 and CD8 T lymphocyte interplay in controlling tumor growth.* Cell Mol Life Sci, 2018. **75**(4): p. 689-713.
131. Kunimasa, K. and T. Goto, *Immunosurveillance and Immunoediting of Lung Cancer: Current Perspectives and Challenges.* Int J Mol Sci, 2020. **21**(2).
132. Koebel, C.M., et al., *Adaptive immunity maintains occult cancer in an equilibrium state.* Nature, 2007. **450**(7171): p. 903-7.
133. Mittal, D., et al., *New insights into cancer immunoediting and its three component phases—elimination, equilibrium and escape.* Current opinion in immunology, 2014. **27**: p. 16-25.
134. Korn, T. and A. Kallies, *T cell responses in the central nervous system.* Nature Reviews Immunology, 2017. **17**(3): p. 179-194.
135. Ransohoff, R.M. and B. Engelhardt, *The anatomical and cellular basis of immune surveillance in the central nervous system.* Nature Reviews Immunology, 2012. **12**(9): p. 623-635.
136. El Rassy, E., et al., *Non-small cell lung cancer brain metastases and the immune system: from brain metastases development to treatment.* Cancer treatment reviews, 2018. **68**: p. 69-79.
137. Morimoto, K. and K. Nakajima, *Role of the immune system in the development of the central nervous system.* Frontiers in Neuroscience, 2019. **13**: p. 916.
138. Ousman, S.S. and P. Kubes, *Immune surveillance in the central nervous system.* Nature neuroscience, 2012. **15**(8): p. 1096-1101.
139. Sevenich, L., et al., *Analysis of tumour-and stroma-supplied proteolytic networks reveals a brain-metastasis-promoting role for cathepsin S.* Nature cell biology, 2014. **16**(9): p. 876-888.
140. Schulz, M., et al., *Cellular and molecular changes of brain metastases-associated myeloid cells during disease progression and therapeutic response.* IScience, 2020. **23**(6).
141. Berghoff, A.S., et al., *Density of tumor-infiltrating lymphocytes correlates with extent of brain edema and overall survival time in patients with brain metastases.* Oncoimmunology, 2016. **5**(1): p. e1057388.

142. Blomberg, O.S., L. Spagnuolo, and K.E. de Visser, *Immune regulation of metastasis: mechanistic insights and therapeutic opportunities.* Dis Model Mech, 2018. **11**(10).
143. Chen, W.-W., et al., *Immune related biomarkers for cancer metastasis to the brain.* Experimental Hematology & Oncology, 2022. **11**(1): p. 1-19.
144. Spano, D. and M. Zollo, *Tumor microenvironment: a main actor in the metastasis process.* Clinical & experimental metastasis, 2012. **29**(4): p. 381-395.
145. Luo, L., et al., *The immune microenvironment in brain metastases of non-small cell lung cancer.* Frontiers in Oncology, 2021. **11**: p. 698844.
146. Aili, Y., et al., *Tumor microenvironment and exosomes in brain metastasis: Molecular mechanisms and clinical application.* Frontiers in Oncology, 2022. **12**: p. 983878.
147. Michell-Robinson, M.A., et al., *Roles of microglia in brain development, tissue maintenance and repair.* Brain, 2015. **138**(5): p. 1138-1159.
148. Sofroniew, M.V., *Molecular dissection of reactive astrogliosis and glial scar formation.* Trends in neurosciences, 2009. **32**(12): p. 638-647.
149. Graeber, M.B., B.W. Scheithauer, and G.W. Kreutzberg, *Microglia in brain tumors.* Glia, 2002. **40**(2): p. 252-259.
150. You, H., S. Baluszek, and B. Kaminska, *Immune microenvironment of brain metastases—are microglia and other brain macrophages little helpers?* Frontiers in Immunology, 2019. **10**: p. 1941.
151. Wang, Y., et al., *Tumor Immune Microenvironment and Immunotherapy in Brain Metastasis From Non-Small Cell Lung Cancer.* Front Immunol, 2022. **13**: p. 829451.
152. Hiam-Galvez, K.J., B.M. Allen, and M.H. Spitzer, *Systemic immunity in cancer.* Nat Rev Cancer, 2021. **21**(6): p. 345-359.
153. Xu, L., et al., *Reshaping the systemic tumor immune environment (STIE) and tumor immune microenvironment (TIME) to enhance immunotherapy efficacy in solid tumors.* J Hematol Oncol, 2022. **15**(1): p. 87.
154. Weiss, N., et al., *The blood-brain barrier in brain homeostasis and neurological diseases.* Biochim Biophys Acta, 2009. **1788**(4): p. 842-57.
155. Leibold, A.T., G.N. Monaco, and M. Dey, *The role of the immune system in brain metastasis.* Current neurobiology, 2019. **10**(2): p. 33.
156. Fridman, W.H., et al., *The immune contexture in human tumours: impact on clinical outcome.* Nature Reviews Cancer, 2012. **12**(4): p. 298-306.
157. Maimela, N.R., S. Liu, and Y. Zhang, *Fates of CD8+ T cells in tumor microenvironment.* Computational and structural biotechnology journal, 2019. **17**: p. 1-13.
158. Zheng, X., Y. Hu, and C. Yao, *The paradoxical role of tumor-infiltrating immune cells in lung cancer.* Intractable & rare diseases research, 2017. **6**(4): p. 234-241.
159. Pennock, N.D., et al., *T cell responses: naive to memory and everything in between.* Adv Physiol Educ, 2013. **37**(4): p. 273-83.
160. Jian, H., et al., *Th1high in tumor microenvironment is an indicator of poor prognosis for patients with NSCLC.* Oncotarget, 2017. **8**(8): p. 13116.
161. Bruni, D., H.K. Angell, and J. Galon, *The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy.* Nature Reviews Cancer, 2020. **20**(11): p. 662-680.
162. Shimasaki, N., A. Jain, and D. Campana, *NK cells for cancer immunotherapy.* Nature reviews Drug discovery, 2020. **19**(3): p. 200-218.
163. Adan, A., et al., *Flow cytometry: basic principles and applications.* Critical reviews in biotechnology, 2017. **37**(2): p. 163-176.
164. Qin, D., *Next-generation sequencing and its clinical application.* Cancer biology & medicine, 2019. **16**(1): p. 4.

165. Cainap, C., et al., *Next generation sequencing technology in lung cancer diagnosis.* Biology, 2021. **10**(9): p. 864.
166. Pervez, M.T., et al., *A Comprehensive Review of Performance of Next-Generation Sequencing Platforms.* BioMed Research International, 2022. **2022**.
167. Heather, J.M. and B. Chain, *The sequence of sequencers: The history of sequencing DNA.* Genomics, 2016. **107**(1): p. 1-8.
168. Berger, M.F. and E.R. Mardis, *The emerging clinical relevance of genomics in cancer medicine.* Nature reviews Clinical oncology, 2018. **15**(6): p. 353-365.
169. Menaa, F., *Next-generation sequencing or the dilemma of large-scale data analysis: opportunities, insights, and challenges to translational, preventive and personalized medicine.* J Investig Genomics, 2014. **1**(1): p. 00005.
170. Yin, Y., C. Butler, and Q. Zhang, *Challenges in the application of NGS in the clinical laboratory.* Human Immunology, 2021. **82**(11): p. 812-819.
171. Nagarajan, N. and M. Pop, *Sequence assembly demystified.* Nature Reviews Genetics, 2013. **14**(3): p. 157-167.
172. Raphael, B.J., et al., *Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine.* Genome medicine, 2014. **6**(1): p. 1-17.
173. Lussier, Y.A., et al., *Oligo- and polymetastatic progression in lung metastasis(es) patients is associated with specific microRNAs.* PLoS One, 2012. **7**(12): p. e50141.
174. Pirker, R., et al., *Prognostic factors in patients with advanced non-small cell lung cancer: data from the phase III FLEX study.* Lung Cancer, 2012. **77**(2): p. 376-382.
175. Gomez, D.R., et al., *Local consolidative therapy versus maintenance therapy or observation for patients with oligometastatic non-small-cell lung cancer without progression after first-line systemic therapy: a multicentre, randomised, controlled, phase 2 study.* The lancet oncology, 2016. **17**(12): p. 1672-1682.
176. Chen, S., et al., *fastp: an ultra-fast all-in-one FASTQ preprocessor.* Bioinformatics, 2018. **34**(17): p. i884-i890.
177. Degasperi, A., et al., *Substitution mutational signatures in whole-genome–sequenced cancers in the UK population.* Science, 2022. **376**(6591): p. abl9283.
178. Koboldt, D.C., *Best practices for variant calling in clinical sequencing.* Genome Medicine, 2020. **12**(1): p. 1-13.
179. Boeva, V., et al., *Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data.* Bioinformatics, 2012. **28**(3): p. 423-425.
180. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
181. Anders, S. and W. Huber, *Differential expression analysis for sequence count data.* Nature Precedings, 2010: p. 1-1.
182. Raudvere, U., et al., *g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update).* Nucleic acids research, 2019. **47**(W1): p. W191-W198.
183. Dennis, G., et al., *DAVID: database for annotation, visualization, and integrated discovery.* Genome biology, 2003. **4**(9): p. 1-11.
184. Kuleshov, M.V., et al., *Enrichr: a comprehensive gene set enrichment analysis web server 2016 update.* Nucleic acids research, 2016. **44**(W1): p. W90-W97.
185. Camy, F., et al., *Brain metastasis PD-L1 and CD8 expression is dependent on primary tumor type and its PD-L1 and CD8 status.* J Immunother Cancer, 2020. **8**(2).
186. Joosse, S.A., et al., *Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells.* Clin Cancer Res, 2012. **18**(4): p. 993-1003.

187. Joosse, S.A. *In-Silico Online v2.3.1*. accessed on 2022 Jul; Available from: http://in-silico.online.
188. Ooms, J., *curl: A Modern and Flexible Web Client for R.* 2021.
189. Bjaanæs, M.M., et al., *Whole genome copy number analyses reveal a highly aberrant genome in TP53 mutant lung adenocarcinoma tumors.* BMC cancer, 2021. **21**(1): p. 1-12.
190. Han, X., et al., *Comprehensive profiling of gene copy number alterations predicts patient prognosis in resected stages I–III lung adenocarcinoma.* Frontiers in Oncology, 2019. **9**: p. 556.
191. Salaun, B., et al., *TLR3 as a biomarker for the therapeutic efficacy of double-stranded RNA in breast cancer.* Cancer research, 2011. **71**(5): p. 1607-1614.
192. Hao, N., et al., *Phosphodiesterase 3A Represents a Therapeutic Target that Drives Stem Cell-like Property and Metastasis in Breast Cancer.* Mol Cancer Ther, 2020. **19**(3): p. 868-881.
193. Mitsiades, N., et al., *Distinct patterns of dysregulated expression of enzymes involved in androgen synthesis and metabolism in metastatic prostate cancer tumors.* Cancer research, 2012. **72**(23): p. 6142-6152.
194. Ji, Q., et al., *Selective loss of AKR1C1 and AKR1C2 in breast cancer and their potential effect on progesterone signaling.* Cancer research, 2004. **64**(20): p. 7610-7617.
195. Gagliardi, M., et al., *Aldo-keto reductases protect metastatic melanoma from ER stress-independent ferroptosis.* Cell death & disease, 2019. **10**(12): p. 902.
196. Chang, W.-M., et al., *AKR1C1 controls cisplatin-resistance in head and neck squamous cell carcinoma through cross-talk with the STAT1/3 signaling pathway.* Journal of Experimental & Clinical Cancer Research, 2019. **38**(1): p. 1-14.
197. Zhou, C., et al., *Loss of AKR1C1 is a good prognostic factor in advanced NPC cases and increases chemosensitivity to cisplatin in NPC cells.* Journal of Cellular and Molecular Medicine, 2020. **24**(11): p. 6438-6447.
198. Matsumoto, R., et al., *Aldo-keto reductase 1C1 induced by interleukin-1β mediates the invasive potential and drug resistance of metastatic bladder cancer cells.* Scientific reports, 2016. **6**(1): p. 34625.
199. Cossarizza, A., et al., *Guidelines for the use of flow cytometry and cell sorting in immunological studies.* Eur J Immunol, 2017. **47**(10): p. 1584-1797.
200. Caffarel, M.M. and M.S. Braza, *Microglia and metastases to the central nervous system: victim, ravager, or something else?* Journal of Experimental & Clinical Cancer Research, 2022. **41**(1): p. 1-13.
201. Peinado, H., et al., *Pre-metastatic niches: organ-specific homes for metastases.* Nature Reviews Cancer, 2017. **17**(5): p. 302-317.
202. Achrol, A.S., et al., *Brain metastases.* Nature Reviews Disease Primers, 2019. **5**(1): p. 5.
203. Gerlinger, M., et al., *Cancer: evolution within a lifetime.* Annual review of genetics, 2014. **48**: p. 215-236.
204. Ciriello, G., et al., *Emerging landscape of oncogenic signatures across human cancers.* Nature genetics, 2013. **45**(10): p. 1127-1133.
205. Wardell, C.P., et al., *Genomic and Transcriptomic Profiling of Brain Metastases.* Cancers, 2021. **13**(22): p. 5598.
206. Álvarez-Prado, Á.F., et al., *Immunogenomic analysis of human brain metastases reveals diverse immune landscapes across genetically distinct tumors.* Cell Reports Medicine, 2023. **4**(1).
207. Vassella, E., et al., *Mutational profiles of primary pulmonary adenocarcinoma and paired brain metastases disclose the importance of KRAS mutations.* European journal of cancer, 2021. **159**: p. 227-236.

208. Zou, X.-l., et al., *Prognostic value of neoantigen load in immune checkpoint inhibitor therapy for cancer.* Frontiers in Immunology, 2021. **12**: p. 689076.
209. Marabelle, A., et al., *Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study.* The Lancet Oncology, 2020. **21**(10): p. 1353-1365.
210. Greillier, L., P. Tomasini, and F. Barlesi, *The clinical utility of tumor mutational burden in non-small cell lung cancer.* Translational lung cancer research, 2018. **7**(6): p. 639.
211. Aggarwal, C., et al., *Assessment of Tumor Mutational Burden and Outcomes in Patients With Diverse Advanced Cancers Treated With Immunotherapy.* JAMA Network Open, 2023. **6**(5): p. e2311181-e2311181.
212. Giustini, N. and L. Bazhenova, *Recognizing prognostic and predictive biomarkers in the treatment of non-small cell lung cancer (NSCLC) with immune checkpoint inhibitors (ICIs).* Lung Cancer: Targets and Therapy, 2021: p. 21-34.
213. Xie, T., et al., *Evolutionary Characteristics and Immunologic Divergence of Lung and Brain Metastasis Lesions in NSCLC.* Molecular Cancer Research, 2023. **21**(4): p. 374-385.
214. Skakodub, A., et al., *Genomic analysis and clinical correlations of non-small cell lung cancer brain metastasis.* Nature Communications, 2023. **14**(1): p. 4980.
215. Stein, M.K., et al., *Tumor mutational burden is site specific in non–small-cell lung cancer and is highest in lung adenocarcinoma brain metastases.* JCO Precision Oncology, 2019. **3**: p. 1-13.
216. van den Heuvel, G.R., et al., *Mutational signature analysis in non-small cell lung cancer patients with a high tumor mutational burden.* Respiratory Research, 2021. **22**: p. 1-7.
217. Ma, K., et al., *Relationship between tumor mutational burden, gene mutation status, and clinical characteristics in 340 cases of lung adenocarcinoma.* Cancer medicine, 2022. **11**(22): p. 4389-4397.
218. Ernst, S.M., et al., *Tobacco smoking-related mutational signatures in classifying smoking-associated and nonsmoking-associated NSCLC.* Journal of Thoracic Oncology, 2023. **18**(4): p. 487-498.
219. Alexandrov, L.B., et al., *Mutational signatures associated with tobacco smoking in human cancer.* Science, 2016. **354**(6312): p. 618-622.
220. Zehir, A., et al., *Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.* Nat Med, 2017. **23**(6): p. 703-713.
221. Chiu, M., et al., *A detailed smoking history and determination of MYC status predict response to checkpoint inhibitors in advanced non-small cell lung cancer.* Translational Lung Cancer Research, 2020. **9**(1): p. 55.
222. Petljak, M., et al., *Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis.* Cell, 2019. **176**(6): p. 1282-1294. e20.
223. Siriwardena, S., K. Chen, and A.S. Bhagwat, *The functions and malfunctions of AID/APOBEC family deaminases: the known knowns and the known unknowns.* Chemical reviews, 2016. **116**(20): p. 12688.
224. Nik-Zainal, S., et al., *Mutational processes molding the genomes of 21 breast cancers.* Cell, 2012. **149**(5): p. 979-993.
225. Roberts, S.A., et al., *An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers.* Nature genetics, 2013. **45**(9): p. 970-976.
226. Ma, J., et al., *APOBEC mutational signature predicts prognosis and immunotherapy response in nonsmoking patients with lung adenocarcinoma.* Translational Lung Cancer Research, 2023. **12**(3): p. 580.
227. Alexandrov, L.B., et al., *The repertoire of mutational signatures in human cancer.* Nature, 2020. **578**(7793): p. 94-101.

228. Nicoś, M., et al., *Genomic Profiling Identifies Putative Pathogenic Alterations in NSCLC Brain Metastases.* JTO Clinical and Research Reports, 2022. **3**(12): p. 100435.
229. Caliri, A.W., S. Tommasi, and A. Besaratinia, *Relationships among smoking, oxidative stress, inflammation, macromolecular damage, and cancer.* Mutation Research/Reviews in Mutation Research, 2021. **787**: p. 108365.
230. Diossy, M., et al., *A subset of lung cancer cases shows robust signs of homologous recombination deficiency associated genomic mutational signatures.* NPJ Precision Oncology, 2021. **5**(1): p. 55.
231. Diossy, M., et al., *Breast cancer brain metastases show increased levels of genomic aberration-based homologous recombination deficiency scores relative to their corresponding primary tumors.* Annals of Oncology, 2018. **29**(9): p. 1948-1954.
232. Garber, H.R., et al., *Incidence and impact of brain metastasis in patients with hereditary BRCA1 or BRCA2 mutated invasive breast cancer.* NPJ Breast Cancer, 2022. **8**(1): p. 46.
233. Tyran, M., et al., *A comparison of DNA mutation and copy number profiles of primary breast cancers and paired brain metastases for identifying clinically relevant genetic alterations in brain metastases.* Cancers, 2019. **11**(5): p. 665.
234. Singh, V.K., et al., *Mutational signature SBS8 predominantly arises due to late replication errors in cancer.* Communications biology, 2020. **3**(1): p. 421.
235. Nik-Zainal, S. and S. Morganella, *Mutational signatures in breast cancer: the problem at the DNA level.* Clinical Cancer Research, 2017. **23**(11): p. 2617-2629.
236. Deutsch, L., et al., *Opposite roles of FOXA1 and NKX2-1 in lung cancer progression.* Genes, Chromosomes and Cancer, 2012. **51**(6): p. 618-629.
237. Klotz, R., et al., *Circulating tumor cells exhibit metastatic tropism and reveal brain metastasis drivers.* Cancer discovery, 2020. **10**(1): p. 86-103.
238. Li, F., L. Sun, and S. Zhang, *Acquirement of DNA copy number variations in non-small cell lung cancer metastasis to the brain.* Oncology reports, 2015. **34**(4): p. 1701-1707.
239. Pastushenko, I., et al., *Fat1 deletion promotes hybrid EMT state, tumour stemness and metastasis.* Nature, 2021. **589**(7842): p. 448-455.
240. Li, P., S. Zhao, and Y. Hu, *SFRP2 modulates non-small cell lung cancer A549 cell apoptosis and metastasis by regulating mitochondrial fission via Wnt pathways.* Molecular Medicine Reports, 2019. **20**(2): p. 1925-1932.
241. Kriegsman, B.A., et al., *Frequent loss of IRF2 in cancers leads to immune evasion through decreased MHC class I antigen presentation and increased PD-L1 expression.* The Journal of Immunology, 2019. **203**(7): p. 1999-2010.
242. Lee, H.W., et al., *Genomic copy number alterations associated with the early brain metastasis of non-small cell lung cancer.* International journal of oncology, 2012. **41**(6): p. 2013-2020.
243. Le Noci, V., et al., *Poly (I: C) and CpG-ODN combined aerosolization to treat lung metastases and counter the immunosuppressive microenvironment.* Oncoimmunology, 2015. **4**(10): p. e1040214.
244. Le Noci, V., et al., *Reprogramming the lung microenvironment by inhaled immunotherapy fosters immune destruction of tumor.* Oncoimmunology, 2016. **5**(11): p. e1234571.
245. Lafon, M., et al., *The innate immune facet of brain: human neurons express TLR-3 and sense viral dsRNA.* Journal of Molecular Neuroscience, 2006. **29**: p. 185-194.
246. Estornes, Y., et al., *Dual role of TLR3 in inflammation and cancer cell apoptosis.* World's largest Science, Technology & Medicine Open Access book publisher, 2013: p. 247-70.
247. Muresan, X.M., et al., *Toll-like receptor 3 in solid cancer and therapy resistance.* Cancers, 2020. **12**(11): p. 3227.

248. Alkurdi, L., et al., *Release of c-FLIP brake selectively sensitizes human cancer cells to TLR3-mediated apoptosis.* Cell Death Dis, 2018. **9**(9): p. 874.
249. Rydberg, C., et al., *Toll-like receptor agonists induce inflammation and cell death in a model of head and neck squamous cell carcinomas.* Immunology, 2009. **128**(1pt2): p. e600-e611.
250. Cheng, Y.-s. and F. Xu, *Anticancer function of polyinosinic-polycytidylic acid.* Cancer biology & therapy, 2010. **10**(12): p. 1219-1223.
251. Khajeh Alizadeh Attar, M., et al., *Basic understanding and therapeutic approaches to target toll-like receptors in cancerous microenvironment and metastasis.* Medicinal Research Reviews, 2018. **38**(5): p. 1469-1484.
252. Tavora, B., et al., *Tumoural activation of TLR3–SLIT2 axis in endothelium drives metastasis.* Nature, 2020. **586**(7828): p. 299-304.
253. González-Reyes, S., et al., *Study of TLR3, TLR4 and TLR9 in breast carcinomas and their association with metastasis.* BMC cancer, 2010. **10**(1): p. 1-9.
254. Yoshida, T., et al., *Toll-like receptor 3 as a recurrence risk factor and a potential molecular therapeutic target in colorectal cancer.* Clinical and Experimental Gastroenterology, 2020: p. 427-438.
255. Benbenishty, A., et al., *Prophylactic TLR9 stimulation reduces brain metastasis through microglia activation.* PLoS biology, 2019. **17**(3): p. e2006859.
256. Ahmad, F., E. Degerman, and V. Manganiello, *Cyclic nucleotide phosphodiesterase 3 signaling complexes.* Hormone and metabolic research, 2012: p. 776-785.
257. Hao, N., et al., *Phosphodiesterase 3A represents a therapeutic target that drives stem cell–like property and metastasis in breast cancer.* Molecular Cancer Therapeutics, 2020. **19**(3): p. 868-881.
258. Savai, R., et al., *Targeting cancer with phosphodiesterase inhibitors.* Expert opinion on investigational drugs, 2010. **19**(1): p. 117-131.
259. Shafer, S.H., S.H. Phelps, and C.L. Williams, *Reduced DNA synthesis and cell viability in small cell lung carcinoma by treatment with cyclic AMP phosphodiesterase inhibitors.* Biochemical pharmacology, 1998. **56**(9): p. 1229-1236.
260. Peng, T., et al., *Inhibitors of phosphodiesterase as cancer therapeutics.* European journal of medicinal chemistry, 2018. **150**: p. 742-756.
261. De Waal, L., et al., *Identification of cancer-cytotoxic modulators of PDE3A by predictive chemogenomics.* Nature chemical biology, 2016. **12**(2): p. 102-108.
262. Li, G., et al., *Deep learning reveals cuproptosis features assist in predict prognosis and guide immunotherapy in lung adenocarcinoma.* Frontiers in Endocrinology, 2022. **13**: p. 970269.
263. Rižner, T.L. and T.M. Penning, *Role of aldo–keto reductase family 1 (AKR1) enzymes in human steroid metabolism.* Steroids, 2014. **79**: p. 49-63.
264. Chien, C.-W., I.-C. Ho, and T.-C. Lee, *Induction of neoplastic transformation by ectopic expression of human aldo-keto reductase 1C isoforms in NIH3T3 cells.* Carcinogenesis, 2009. **30**(10): p. 1813-1820.
265. Woenckhaus, M., et al., *Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers.* The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland, 2006. **210**(2): p. 192-204.
266. Fukumoto, S.-i., et al., *Overexpression of the aldo-keto reductase family protein AKR1B10 is highly correlated with smokers' non–small cell lung carcinomas.* Clinical cancer research, 2005. **11**(5): p. 1776-1785.
267. Penning, T.M. and C. Lerman, *Genomics of smoking exposure and cessation: lessons for cancer prevention and treatment.* Cancer Prevention Research, 2008. **1**(2): p. 80-83.
268. Penning, T.M., *Human aldo-keto reductases and the metabolic activation of polycyclic aromatic hydrocarbons.* Chemical research in toxicology, 2014. **27**(11): p. 1901-1917.

269. Zhu, H., et al., *The SIRT2-mediated deacetylation of AKR1C1 is required for suppressing its pro-metastasis function in non-small cell lung cancer.* Theranostics, 2020. **10**(5): p. 2188.
270. Zhu, H., et al., *AKR1C1 Activates STAT3 to Promote the Metastasis of Non-Small Cell Lung Cancer.* Theranostics, 2018. **8**(3): p. 676-692.
271. Fu, Z., et al., *Natural product alantolactone targeting AKR1C1 suppresses cell proliferation and metastasis in non-small-cell lung cancer.* Frontiers in Pharmacology, 2022. **13**: p. 847906.
272. Huang, F., et al., *Ferroptosis-related gene AKR1C1 predicts the prognosis of non-small cell lung cancer.* Cancer Cell Int, 2021. **21**(1): p. 567.
273. Kavsak, P., et al., *Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGFβ receptor for degradation.* Molecular cell, 2000. **6**(6): p. 1365-1375.
274. David, D., S.A. Nair, and M.R. Pillai, *Smurf E3 ubiquitin ligases at the cross roads of oncogenesis and tumor suppression.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2013. **1835**(1): p. 119-128.
275. Fu, L., et al. *The functions and regulation of Smurfs in cancers*. in *Seminars in cancer biology*. 2020. Elsevier.
276. Tsao, S.-M. and H.-Y. Hsu, *Fucose-containing fraction of Ling-Zhi enhances lipid rafts-dependent ubiquitination of TGFβ receptor degradation and attenuates breast cancer tumorigenesis.* Scientific Reports, 2016. **6**(1): p. 36563.
277. Jin, C., et al., *Smad ubiquitination regulatory factor 2 promotes metastasis of breast cancer cells by enhancing migration and invasiveness.* Cancer research, 2009. **69**(3): p. 735-740.
278. Wu, B., et al., *Downregulation of Smurf2 ubiquitin ligase in pancreatic cancer cells reversed TGF-β-induced tumor formation.* Tumor Biology, 2016. **37**: p. 16077-16091.
279. Fukuchi, M., et al., *High-level expression of the Smad ubiquitin ligase Smurf2 correlates with poor prognosis in patients with esophageal squamous cell carcinoma.* Cancer research, 2002. **62**(24): p. 7162-7165.
280. Klupp, F., et al., *E3 ubiquitin ligase Smurf2: a prognostic factor in microsatellite stable colorectal cancer.* Cancer management and research, 2019: p. 1795-1803.
281. Fukunaga, E., et al., *Smurf2 induces ubiquitin-dependent degradation of Smurf1 to prevent migration of breast cancer cells.* Journal of Biological Chemistry, 2008. **283**(51): p. 35660-35667.
282. Song, D., et al., *Smurf2 suppresses the metastasis of hepatocellular carcinoma via ubiquitin degradation of Smad2.* Open Medicine, 2022. **17**(1): p. 384-396.
283. Li, Y., et al., *Smurf1 and Smurf2 mediated polyubiquitination and degradation of RNF220 suppresses Shh-group medulloblastoma.* Cell Death & Disease, 2023. **14**(8): p. 494.
284. Han, M., et al., *SMURF2 facilitates ubiquitin-mediated degradation of ID2 to attenuate lung cancer cell proliferation.* International Journal of Biological Sciences, 2023. **19**(11): p. 3324.
285. Wu, Q., et al., *Smurf2 induces degradation of GSK-3β and upregulates β-catenin in chondrocytes: a potential mechanism for Smurf2-induced degeneration of articular cartilage.* Experimental cell research, 2009. **315**(14): p. 2386-2398.
286. Koganti, P., G. Levy-Cohen, and M. Blank, *Smurfs in protein homeostasis, signaling, and cancer.* Frontiers in Oncology, 2018. **8**: p. 295.
287. Donato, R., et al., *Functions of S100 proteins.* Current molecular medicine, 2013. **13**(1): p. 24-57.
288. Zimmer, D.B., et al., *The S100 protein family: history, function, and expression.* Brain research bulletin, 1995. **37**(4): p. 417-429.

289. Lombardi, T. and J. Samson, *S-100 protein in normal and pathologic oral tissues. A review.* Schweizer Monatsschrift fur Zahnmedizin= Revue Mensuelle Suisse D'odonto-stomatologie= Rivista Mensile Svizzera di Odontologia e Stomatologia, 1993. **103**(4): p. 413-418.
290. Wolf, S., C. Haase-Kohn, and J. Pietzsch, *S100A2 in cancerogenesis: a friend or a foe?* Amino acids, 2011. **41**: p. 849-861.
291. Sugino, H. and Y. Sawada, *Influence of S100A2 in Human Diseases.* Diagnostics, 2022. **12**(7): p. 1756.
292. Kirschner, R.D., et al., *Transcriptional activation of the tumor suppressor and differentiation gene S100A2 by a novel p63-binding site.* Nucleic acids research, 2008. **36**(9): p. 2969-2980.
293. van Dieck, J., et al., *Modulation of the Oligomerization State of p53 by Differential Binding of Proteins of the S100 Family to p53 Monomers and Tetramers♦.* Journal of Biological Chemistry, 2009. **284**(20): p. 13804-13811.
294. Tan, M., et al., *Transcriptional activation of the human S100A2 promoter by wild-type p53.* FEBS letters, 1999. **445**(2-3): p. 265-268.
295. Buckley, N., et al., *S100A2 is a BRCA1/p63 coregulated tumour suppressor gene with roles in the regulation of mutant p53 stability.* Cell death & disease, 2014. **5**(2): p. e1070-e1070.
296. Wicki, R., et al., *Repression of the candidate tumor suppressor gene S100A2 in breast cancer is mediated by site-specific hypermethylation.* Cell calcium, 1997. **22**(4): p. 243-254.
297. Kyriazanos, I.D., et al., *Expression and prognostic significance of S100A2 protein in squamous cell carcinoma of the esophagus.* Oncology reports, 2002. **9**(3): p. 503-510.
298. Tsai, W.-C., et al., *Cyclooxygenase-2 is involved in S100A2-mediated tumor suppression in squamous cell carcinoma.* Molecular cancer research, 2006. **4**(8): p. 539-547.
299. Shrestha, P., et al., *Localization of Ca 2+-binding S100 proteins in epithelial tumours of the skin.* Virchows Archiv, 1998. **432**: p. 53-59.
300. Hough, C.D., et al., *Coordinately up-regulated genes in ovarian cancer.* Cancer research, 2001. **61**(10): p. 3869-3876.
301. Hountis, P., et al., *S100A2 protein and non-small cell lung cancer. The dual role concept.* Tumour Biol, 2014. **35**(8): p. 7327-33.
302. Feng, G., et al., *Diminished expression of S100A2, a putative tumor suppressor, at early stage of human lung carcinogenesis.* Cancer research, 2001. **61**(21): p. 7999-8004.
303. Heighway, J., et al., *Expression profiling of primary non-small cell lung cancer for target identification.* Oncogene, 2002. **21**(50): p. 7749-7763.
304. Wang, H., et al., *Overexpression of S100A2 protein as a prognostic marker for patients with stage I non small cell lung cancer.* International journal of cancer, 2005. **116**(2): p. 285-290.
305. Bulk, E., et al., *S100A2 induces metastasis in non-small cell lung cancer.* Clin Cancer Res, 2009. **15**(1): p. 22-9.
306. Liu, D., et al., *Expression of calcium-binding protein S100A2 in breast lesions.* British journal of cancer, 2000. **83**(11): p. 1473-1479.
307. Naz, S., et al., *Protumorigenic actions of S100A2 involve regulation of PI3/Akt signaling and functional interaction with Smad3.* Carcinogenesis, 2014. **35**(1): p. 14-23.
308. Diederichs, S., et al., *S100 family members and trypsinogens are predictors of distant metastasis and survival in early-stage non-small cell lung cancer.* Cancer research, 2004. **64**(16): p. 5564-5569.

309. Wang, T., et al., *Expression and clinicopathological significance of S100 calcium binding protein A2 in lung cancer patients of Chinese Han ethnicity.* Clin Chim Acta, 2017. **464**: p. 118-122.
310. Zhang, Y., et al., *S100A gene family: immune-related prognostic biomarkers and therapeutic targets for low-grade glioma.* Aging (Albany NY), 2021. **13**(11): p. 15459.
311. Chen, Y., et al., *S100A2 is a prognostic biomarker involved in immune infiltration and predict immunotherapy response in pancreatic cancer.* Frontiers in immunology, 2021. **12**: p. 758004.
312. Sampson, J.H., et al., *Brain immunology and immunotherapy in brain tumours.* Nat Rev Cancer, 2020. **20**(1): p. 12-25.
313. Mansfield, A., et al., *Temporal and spatial discordance of programmed cell death-ligand 1 expression and lymphocyte tumor infiltration between paired primary lesions and brain metastases in lung cancer.* Annals of Oncology, 2016. **27**(10): p. 1953-1958.
314. Kudo, Y., et al., *Suppressed immune microenvironment and repertoire in brain metastases from patients with resected non-small-cell lung cancer.* Ann Oncol, 2019. **30**(9): p. 1521-1530.
315. DeNardo, D.G., et al., *CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages.* Cancer Cell, 2009. **16**(2): p. 91-102.
316. Fontenot, J.D. and A.Y. Rudensky, *A well adapted regulatory contrivance: regulatory T cell development and the forkhead family transcription factor Foxp3.* Nat Immunol, 2005. **6**(4): p. 331-7.
317. Harter, P.N., et al., *Distribution and prognostic relevance of tumor-infiltrating lymphocytes (TILs) and PD-1/PD-L1 immune checkpoints in human brain metastases.* Oncotarget, 2015. **6**(38): p. 40836-49.
318. Li, Y.D., et al., *Tumor-induced peripheral immunosuppression promotes brain metastasis in patients with non-small cell lung cancer.* Cancer Immunol Immunother, 2019. **68**(9): p. 1501-1513.
319. Yamaguchi, K., et al., *Activation of central/effector memory T cells and T-helper 1 polarization in malignant melanoma patients treated with anti-programmed death-1 antibody.* Cancer Sci, 2018. **109**(10): p. 3032-3042.
320. Saxena, A., P.K. Dagur, and A. Biancotto, *Multiparametric Flow Cytometry Analysis of Naive, Memory, and Effector T Cells.* Methods Mol Biol, 2019. **2032**: p. 129-140.
321. Masopust, D. and J.M. Schenkel, *The integration of T cell migration, differentiation and function.* Nature Reviews Immunology, 2013. **13**(5): p. 309-320.
322. Allard, D., et al., *Targeting the CD73-adenosine axis in immuno-oncology.* Immunol Lett, 2019. **205**: p. 31-39.
323. Dianzani, U., et al., *Co-stimulatory signal delivered by CD73 molecule to human CD45RAhiCD45ROlo (naive) CD8+ T lymphocytes.* Journal of immunology (Baltimore, Md.: 1950), 1993. **151**(8): p. 3961-3970.
324. Yang, L., J.J. Kobie, and T.R. Mosmann, *CD73 and Ly-6A/E distinguish in vivo primed but uncommitted mouse CD4 T cells from type 1 or type 2 effector cells.* The Journal of Immunology, 2005. **175**(10): p. 6458-6464.
325. Rissiek, A., et al., *The expression of CD39 on regulatory T cells is genetically driven and further upregulated at sites of inflammation.* Journal of autoimmunity, 2015. **58**: p. 12-20.
326. Mandapathil, M., et al., *Generation and accumulation of immunosuppressive adenosine by human CD4+ CD25highFOXP3+ regulatory T cells.* Journal of Biological Chemistry, 2010. **285**(10): p. 7176-7186.
327. Zimmermann, H., *5'-Nucleotidase: molecular structure and functional aspects.* Biochem J, 1992. **285 (Pt 2)**: p. 345-65.

328. Gao, Z.W., K. Dong, and H.Z. Zhang, *The roles of CD73 in cancer.* Biomed Res Int, 2014. **2014**: p. 460654.
329. Borsellino, G., et al., *Expression of ectonucleotidase CD39 by Foxp3+ Treg cells: hydrolysis of extracellular ATP and immune suppression.* Blood, 2007. **110**(4): p. 1225-32.
330. Jarvis, L.B., et al., *Therapeutically expanded human regulatory T-cells are super-suppressive due to HIF1A induced expression of CD73.* Communications biology, 2021. **4**(1): p. 1186.
331. Stagg, J. and M. Smyth, *Extracellular adenosine triphosphate and adenosine in cancer.* Oncogene, 2010. **29**(39): p. 5346-5358.
332. Schneider, E., et al., *CD73-mediated adenosine production by CD8 T cell-derived extracellular vesicles constitutes an intrinsic mechanism of immune suppression.* Nature Communications, 2021. **12**(1): p. 5911.
333. Doherty, G.A., et al., *CD 73 is a phenotypic marker of effector memory T h17 cells in inflammatory bowel disease.* European journal of immunology, 2012. **42**(11): p. 3062-3072.
334. Colgan, S.P., et al., *Physiological roles for ecto-5'-nucleotidase (CD73).* Purinergic Signal, 2006. **2**(2): p. 351-60.
335. Resta, R., Y. Yamashita, and L.F. Thompson, *Ecto-enzyme and signaling functions of lymphocyte CD73.* Immunol Rev, 1998. **161**: p. 95-109.
336. Yamashita, Y., et al., *CD73 expression and fyn-dependent signaling on murine lymphocytes.* Eur J Immunol, 1998. **28**(10): p. 2981-90.
337. Tu, E., et al., *Anti-PD-L1 and anti-CD73 combination therapy promotes T cell response to EGFR-mutated NSCLC.* JCI Insight, 2022. **7**(3).
338. Inoue, Y., et al., *Prognostic impact of CD73 and A2A adenosine receptor expression in non-small-cell lung cancer.* Oncotarget, 2017. **8**(5): p. 8738-8751.
339. Yang, P., et al., *Peripheral CD4+ naive/memory ratio is an independent predictor of survival in non-small cell lung cancer.* Oncotarget, 2017. **8**(48): p. 83650-83659.
340. Mansouri, S., et al., *Cancer genome and tumor microenvironment: Reciprocal crosstalk shapes lung cancer plasticity.* Elife, 2022. **11**: p. e79895.
341. Skoulidis, F. and J.V. Heymach, *Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy.* Nature Reviews Cancer, 2019. **19**(9): p. 495-509.
342. Wang, Q., et al., *Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment.* Cancer cell, 2017. **32**(1): p. 42-56. e6.
343. Klemm, F., et al., *Interrogation of the microenvironmental landscape in brain tumors reveals disease-specific alterations of immune cells.* Cell, 2020. **181**(7): p. 1643-1660. e17.
344. Petralia, F., et al., *Integrated proteogenomic characterization across major histological types of pediatric brain cancer.* Cell, 2020. **183**(7): p. 1962-1985. e31.
345. Kamer, I., et al., *Predicting brain metastasis in early stage non-small cell lung cancer patients by gene expression profiling.* Translational Lung Cancer Research, 2020. **9**(3): p. 682.
346. Rashdan, S., et al., *Narrative review: molecular and genetic profiling of oligometastatic non-small cell lung cancer.* Translational Lung Cancer Research, 2021. **10**(7): p. 3351.
347. Grist, E., et al., *Accumulation of copy number alterations and clinical progression across advanced prostate cancer.* Genome Medicine, 2022. **14**(1): p. 1-15.

10. Acknowledgments

As I reach this pivotal point in my academic journey, I would like to express my profound gratitude to several individuals who have played an instrumental role in the completion of my Ph.D. thesis.

First and foremost, I am deeply indebted to **Prof. Dr. med. Klaus Pantel** for giving me the invaluable opportunity to pursue my Ph.D. in the field of cancer research. Our initial conversation in Freiburg, where you took the time to listen to my aspirations, marked the beginning of this remarkable journey.

I extend my sincere appreciation to my promoter, **Prof. Dr. Harriet Wikman.** Dear Haiju, your decision to include me in your research project has been a turning point in my academic career. Thank you for your constant guidance and support, for troubleshooting my work, and for continually pushing me to do better.

To my Ph.D. supervisor, **Priv.-Doz. Dr. Simon A. Joosse**, I offer my genuine gratitude. Dear Simon, thank you for welcoming me to your team, and for your support and guidance. Your efforts in helping me develop my bioinformatic skills have been invaluable. Your creation of a trusting and open atmosphere allowed me to honestly discuss my difficulties and concerns.

I would also like to acknowledge **Priv.-Doz. Dr. med. Malte Mohme** for his significant contributions to this study. A special thanks go to **Dr. rer. nat. Cecile Maire**, who consistently provided assistance and guidance whenever needed. Thank you, Cecile, for your kindness and warmth.

This research would not have been possible without the contributions of numerous individuals within ITB and UKE. I extend my thanks to **Christian Müller, Priv.-Doz. Dr. med. Jakob Matschke, Prof. Dr. Eva Tolosa, Laura Glau, Merle Reetz, Svenja Schneegans, Gresa Emurlai, Andras Piffko, and Benedikt Asey** for their dedicated efforts. I reserve a special note of gratitude for **Jolanthe Kropidlowski**, whose lab support and unfailing kindness have been truly appreciated.

To all colleagues in the ITB, I express my sincere thanks for all the memories we've created together and for your constant kindness and assistance. Nikhil, Rafael, Hyeongjung, Maria, Max, Falk, Victoria, Yassine, and Annabelle, I'm grateful for the enjoyable lunches and stimulating scientific discussions. Sandra, your vibrant presence has brought life and joy to our lab.

I would also like to offer my thanks to the lung cancer patients who participated in this study. Your willingness to contribute has made this research possible and holds the promise of improving lives.

Finally, I want to express my deepest gratitude to my family. To my parents, my sister Maya, and my two brothers Murhaf and Manhal, your constant love and support have been my anchor. To my husband Mohamad, you have been my greatest supporter, believing in me during moments of self-doubt, providing encouragement during times of frustration, and standing by my side since the day we met. To my soon-to-arrive baby, your presence, despite the morning sickness and back pains, has given me the extra motivation needed to complete this thesis on time.

With love and heartfelt thanks to you all.

11. Curriculum Vitae

CV not included for data protection reasons

12. Affidavit: Own work declaration

Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Mais Alsousli