Master's Thesis

Implementation and Comparative Assessment of Diagnostic Cancer Gene Panels in the Molecular Pathology Laboratory

University of Luxembourg

Faculty of Science, Communication and Technology

Master in Integrated Systems Biology

by

Ben Flies

(010081174D)

Abstract

Contents

1 Introduction

Cancer represents a huge burden for health care systems worldwide and is one of the leading death causes. Scientific discoveries in the last decade have had an enormous impact on our understanding of the underlying causes of cancer. The development of omics techniques, in combination with advanced computational power, has lead to an explosion of biological data. It has become clear that cancer is an incredibly complex malignancy, which is affected by genetic, environmental and behavioural factors. The research community is trying to interprete this vast amount of data with the goal to get a deeper understanding of cancer and to cure it eventually. In recent years, several drugs have been approved, which target proteins needed for cancer development, proliferation or metastasis. Molecular testing is employed to check whether these targeted drugs would be of benefit. In that regard, Next-Generation Sequencing (NGS) is an interesting method to gain deep insights into the genetic information of a tumor and to guide personalized therapy.

1.1 Cancer Genetics

DNA undergoes continuous damage. In cancer cells, the equilibrium between DNA damage and repair systems is dysbalanced [1]. Genetic and epigenetic alterations, in combination with several environmental factors, such as inflammation, enable the hallmarks of cancer [2]. These include replicative immortality, resistance to cell death, sustained proliferative signaling, growth suppressor evasion, inducement of angiogenesis, energy metabolism reprogramming, evasion of immune destruction, invasion, and metastasis.

It is widely accepted that tumors accumulate somatic mutations during their progression towards more malignant stages [3] [4]. Genomic DNA can be damaged by endogenous and environmental agents [5]. Carcinogenic substances produced by industry [6] [7] or present in tobacco smoke [8] are known to increase cancer risk. Cellular metabolic processes also produce DNA-damaging products that induce cancer, such as reactive oxygen species [9] [10]. DNA lesions can be missed DNA repair mechanisms if the damage happens in an inaccessible region of the DNA or if the DNA repair system is defective [11].

The process of DNA replication is not free from errors. It has been estimated that DNA polymerase has error rates ranging from 10^{-4} to 10^{-6}. This is followed by mismatch repair, which is often intrinsic to the polymerase. This allows excision of 90-99% of misincorporated bases, decreasing the overall error rate to 10^{-6} – 10^{-8}. Additionally, several DNA polymerases exist, which differ in their

error rates. DNA polymerase β has a much worse error rate than DNA polymerase δ or ϵ. There is evidence that these enzymes can be used interchangeabily, and that DNA β is increased in some tumors. This leads to more errors, resulting in increased mutagenesis.

Processes affecting chromosomal and microsatellite integrity contribute to genomic instability in cancer cells. Microsatellite instability (MSI) is caused by inactivation or loss of DNA mismatch repair [12]. Microsatellite elongation or shortening is a consequence of defective or inactive DNA mismatch repair (MMR), which corrects base replication errors [13]. DNA polymerase has a higher error rate in repetitive DNA sequences. When MMR genes are inactivated or defective, the replication mistakes in microsatellites cannot be corrected: MSI is the consequence. In some cancers, MSI can occur despite functional MMR through frameshift mutations at microsatellites [13]. MSI is often associated with cancers harboring mutations in TGFβRII, EGFR, PTEN, and BAX, as they contain such microsatellites [14]. Microsatellite and mismatch errors by DNA polymerase usually result in insertions or deletions and in point mutations, which can be missense, nonsense or silent. These mutations occur at the nucleotide level and might affect the protein structure, leading to defective or overactive enzymes.

Chromosomal instability (CIN) is a common observation in solid tumors, especially in colorectal cancer [13]. Defects in proteins needed for segregation of chromosomes result in chromosome missegregation. This leads to telomere dysfunction, faulty cohesion of sister chromatids, loss of heterozygosity (LOH), hypo– or hyperactive spindle assembly checkpoint or defective centrosome duplication and aneuploidy. [13]. About 70% of solid tumors are aneuploid [15]. Another chromosomal instability process has been described recently: chromothripsis happens when chromosomes are fragmented [16] [17] [17]. The cell tries to repair the chromosomes, but this process is far from being perfect, leading to massive chromosomal rearrangements.

Epigenetic changes do not involve changes in the coding sequence, but affect gene expression. Gene transcription can be influenced by histone modifications, dysregulation of DNA–binding of transcription factors, microRNAs or altered CpG island methylation. In normal cells, CpG islands in gene promoters are usually unmethylated. This is associated with active transcription. Other CpG islands across the genome are methylated. In cancer cells, this situation is often inverted. CpG islands in tumor suppressor promoters are found to be hypermethylated in many tumors. Their gene expression is thus drastically decreased.

1.2 Driver and Passenger Mutations

Cancer progression is a process that recognizes basic Darwinian evolution principles [18] [19] [20] [21]. The population of cancer cells harbors heritable genetic variation. These mutations may be of germline origin or may occur through somatic processes. If the occurring mutations are non–deleterious, they can be passed on to the next generation of cells. The second process, which has to take place in Darwinian evolution, is natural selection. Each cell exhibits a unique combination of genetic and environmental perturbations. Cells are in competition for a variety of resources in their microenvironment, which include space, oxygen and nutrients. Eventually, cells with the best fitness, e.g. with the highest proliferative potential and the lowest death rate, are selected through natural selection principles and will outlast less fit cells. Additionally, these cells will continue to accumulate new mutations. This results in sequential waves of clonal expansion [18].

Genomic instability in cancerous cells becomes a critical mechanism if it affects oncogenes or tumor suppressor genes. These genes have the potential to be causative tumor 'driver' mutations. The identification of driver mutations has been a central aim of cancer research. Mutations in at least 350 human genes are found recursively in cancer genomes and are believed to contribute to cancerogenesis [22]. These driver mutations are positively selected during cancer progression and confer a selective advantage to the cells harboring them. Many alterations found in cancer cells are passenger mutations, which occur coincidally or subsequently to driver mutations. These mutations are defined to not contribute to the selective fitness of the cell, even though this conception has been challenged by stochastic tumor progression simulations [23].

Estimating the number of somatic driver and passenger mutations and the rate at which they occur is not well established [24]. Some studies have reported that cancer cells carry 40–80 somatic mutations, and only 5–15 of them are driver mutations [25]. Two tumors, even though histologically indistinguishable, might present different subsets of mutations [24] [26]. This observation has been defined as inter-tumor heterogeneity. Additionally, tumors present heterogeneity at the intra–tumor level [27]: two subclones of the tumor might present different mutations.

Tumor instability and the accumulation of driver mutations present a major problem in cancer therapy. Chemotherapy, for instance, can induce an environment where chemotherapy–sensitive cells are killed, while resistant cells can survive. These resistant clones are often the source of relapses and the chemotherapeutic agent will lose its efficiency. It is thereby of utmost necessity to hit the tumor as fast, as hard and as targeted as possible. Due to the cancer's heterogeneity, a combination of several therapeutic methods and agents often yields the best results.

Tumor suppressor genes protect a cell from entering the path to cancer. They comprise genes encoding for cell adhesion proteins, DNA repair proteins, proteins acting in apoptosis pathways, or cell cycle proteins [28]. The action of these proteins inhibits metastasis, excessive cell survival or proliferation. Tumor suppressors mostly follow the two-hit hypothesis [29]: to inactivate the tumor-protecting role of tumor suppressors, two genetic events, often LOH in combination with silencing point mutations or silencing of both alleles by somatic events, are necessary to inactivate both alleles of the gene. Compared to dominant oncogenes, tumor suppressor genes are often considered to be recessive. Alternatively, tumor progression can be influenced by functional haploinsufficiency of tumor suppressors [30]. According to this conception, a disease state can emerge if a cell / organism has only one functional copy of a given gene and if it cannot produce enough of a gene product to establish a wild-type condition. APC and TP53 are amongst the best known tumor suppressors.

In the canonical Wnt signaling pathway, a destruction complex, including APC, leads to β–catenin phosphorylation, followed by ubiquitination, marking it for degradation in the proteasome. Activation of Wnt signaling inhibits the destruction complex. Consequently, β–catenin is no longer marked for degradation and can migrate to the nucleus, where it acts on gene expression of target genes [31]. In many tumors, loss or dysfunction of APC leads to β–catenin accumulation in the nucleus, resulting in increased cell migration and decreased cell adhesion and apoptosis [32].

TP53 is the master guardian of the genome [33]. In normal situations, p53, the protein encoded by TP53, is targeted for degradation in the proteasome by ubiquitination [34]. In case of cellular stress, p53 is no longer ubiquitinated. p53 can then stop the cell cycle at the G1/S and G2/M transitions, induce DNA repair, and induce apoptosis if the damage cannot be repaired [35]. One mechanism by which p53 acts on cell-cycle arrest is by activating expression of p21. p21 binds to the G1/S transition complex and inhibits its activity, leading to cell-cycle arrest [35]. Inactivation or mutation of TP53 is a crucial step in many cancers, leading to a loss of control over DNA stability [36].

Oncogenes comprise several GTPases, transcription factors, receptor tyrosine kinases and growth factors [37]. Overexpressed or overactive versions of these proteins lead to increased mitogenic signals, causing increased cell growth or proliferation. Two important oncogenic pathways include the RAS–RAF–MEK–ERK and PTEN–PI3K–AKT pathways, which can both be activated by ligand–binding on Epidermal Growth Factor (EGFR) (cf. section 1.4.).

1.3 Molecular Profiling of Solid Tumors

Lung cancer, melanoma and colorectal cancer are amongst the most common cancers worldwide. Classically, diagnosis has been made by observing histologic, anatomic and pathologic alterations. Molecular profiling of tumors by several methods has lead to a better understanding of cancer development and progression and to the identification of some recursively found driver mutations, which may be potential anti-cancer targets.

Lung cancer is the most common cancer worldwide, both in terms of new cases (1.8 million) and deaths (1.6 million) (cancer.org). Smoking is a widely accepted risk factor. Multiple chemical carcinogens in tobacco smoke induce several genetic mutations [8]. Lung cancer can be divided into two subtypes based on their histology: small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Over the last decade, it has become clear that these subtypes can be classified into additional classes based on the mutational status of recurrent driver mutations.

A combination of oncogenic triggers cause cells of the normal bronchial epithelium to proliferate, giving rise to meta–, hyper– and dysplastic epithelial lesions. Genomic events in early stages of lung cancer giving rise to atypical adenomatous hyperplasia include LOH on chr.3p, $p16^{INK4a}$ or RB inactivation, as well as mutations in KRAS or in β–catenin. TP53 inactivation and LOH on chr.13q are believed to favor progression into the primary adenocarcinoma stage. After that stage, major chromosomic instability is often detected, giving rise to metastatic adenocarcinoma. These chromosomic events include LOH on chr.2q, chr.9p, chr.18q, and chr.22q. Additionally, the oncogene c-myc can be amplified in late stages [38].

Other frequent oncogenic mutations in several oncogenes can be found at each stage of NSCLC and in all histological types, in both smokers and never-smokers. Frequent mutations in NSCLC affect EGFR (10–35%), KRAS (15–25%), PTEN (4–8%), HER2 (2–4%), DDR2 (4%), PIK3CA (1–3%), BRAF (1–3%), AKT1 (1%), MEK (1%) and NRAS (1%). Additionally, rearrangement of ALK (3–7%), RET (1%) and ROS1 (1%) and amplifications of FGFR1 (20%) and MET (2–4%) are found recursively. These mutations are rarely observed together in the same tumor (mycancergenome.org).

Melanoma develops from melanocytes in the basal epidermal layer of the skin. Exposure to UV light, immunosuppression, fair-skin and multiple nevi are risk factors. UV radiation causes cyclobutane pyrimidine dimers (CPDs) [39]. T–T, C–C or C–T dimers (UV fingerprints) are formed, leading to direct DNA damage. People diagnosed with rare genetic disorders like xeroderma pigmentosum are at great risk [40]. Traditionally, melanoma has been classified based on histologic and pathologic properties, such as the thickness of the tumor, ulceration or the anatomic location of the tumor.

Spontaneous mutations in BRAF and NRAS and epigenetic modulations of APC promote progression of the normal epithelium into a dysplastic nevus. Mutations affecting genes PTEN and CDKN2A and altered gene expression of MGMT or RASSF1 then favor invasion of the dermis, which is underlying to the epidermis, during the radial growth phase. Genetic and epigenetic alterations in AKT, MTAP, APAF, CDH1 and CDH2 then result in the vertical growth phase, where the tumor invades surrounding tissues. Finally, loss of TP53 functionality and further epigenetic modulations of MTA2 and MAGE enable metastasis [41] [42].

Driver mutations frequently found in melanoma occur on BRAF (37–50%), NRAS (13–25%), MEK (6–7%), NF1 (11.9%), CTNNB1 (2-4%), GNAQ (1.3%) and GNA11 (1.2%) (mycancergenome.org).

Colorectal cancer (CRC) is one of the best studied cancers. The development of colorectal adenocarcinomas occurs over many years. Caused by the acquisition and accumulation of driver mutations, a normal colorectal epithelium can progress to adenoma, which develops into carcinoma, which can eventually mestastasize.

The molecular progression models in CRC depend on the underlying instability process (chromosomal vs. microsatellite). In CIN CRCs, loss of the tumor suppressor gene APC is often causing the evolution from a normal to a hyperproliferative epithelium. Progression to adenoma stages are associated with DNA hypomethylation, KRAS activation and loss of chr.18q. Mutations in TGFβRII and PIK3CA and loss of TP53 by LOH on chr.17p then lead to the final carcinoma stage. In MSI CRCs, mutations and hypermethylations in MMR genes result in an hyperproliferative epithelium. BRAF mutations, followed by PIK3CA mutations, loss of TP53 and frameshift mutations affecting TGFβRII, BAX or IGF2R are associated with CRC progression towards the carcinoma stage [13] [43] [44].

Driver mutations recursively detected in CRC occur on KRAS (36–40%), SMAD4 (10–35%), PIK3CA (10–30%), BRAF (8–15%), PTEN (5–14%), NRAS (1–6%), and AKT1 (1–6%) (mycancergenome.org).

The mentioned molecular progression profiles are likely to be an oversimplification. These models are based on frequently found alterations in the respective cancers. Due to tumor heterogeneity, these alterations do not always have to be observed within the tumor, and the chronological appearance of these alterations may vary from one tumor to another. Pharmacologically, only a few of these driver mutations are clinically actionable, e.g. can be targeted with drugs. In recent years, major progress has been made in the treatment of solid tumor patients by targeting the EGFR signaling pathway.

1.4 Targeting the EGFR Signaling Pathway

EGFR is a cell surface tyrosine kinase receptor [45]. It is anchored in the cytoplasmic membrane and is composed of an intracytoplasmic tyrosine kinase domain, a short hydrophobic transmembrane domain and an extracellular ligand-binding domain [46]. Ligand binding causes a conformational change of the receptor, which leads to homo– or heterodimerization, followed by an auto– and cross–phosphorylation of key tyrosine residues on its cytoplasmic domain [46]. This forms docking sites for cytoplasmic adaptor proteins that contain phosphotyrosine-binding and Src homology 2 domains. Ligand–binding to EGFR activates several signaling pathways, including the important oncogenic PI3K–AKT and RAS–RAF–MEK–ERK pathways.

Signaling through the PI3K–AKT pathway leads to cell growth, proliferation and survival. The signaling cascade is initiated by G–protein coupled receptors, integrins, cytokine receptors, and receptor tyrosine kinases, such as EGFR [47]. Activation of the receptor results in production of PIP3 by activation of PI3K. PIP3 is anchored in the cell membrane and acts as docking site for proteins containing PH domains, such as PDK1. PIP3–bound PDK1 partially activates Akt by phosphorylation [48]. Full activation of Akt is achieved by phosphorylation of PDK1 and mTORC2 [49]. Activated Akt then acts on a variety of proteins necessary for protein synthesis, glucose metabolism, cell survival / death and proliferation. The phosphatases PP2A and PHLPP can dephosphorylate and thereby inactivate Akt [49]. Additionally, PTEN dephosphorylates PIP3 and indirectly also inactivates Akt [49]. Dysregulation of the PI3K–AKT pathway has been associated with several human diseases including neurological diseases, diabetes and cancer [50]. In cancer, inactivation of PTEN and kinase activity activating mutations on PI3K and Akt are found recursively, leading to enhanced signaling, leading to inhibition of apoptosis and increased proliferation [51].

In the RAS–RAF–MEK–ERK pathway, ligand binding on cell surface receptor tyrosine kinases activates the receptor. GRB2 binds to Tyr1068 of EGFR through its SH2 domain and recruits SOS, a guanine nucleotide exchange factor [52]. Grb2 and SOS then form a complex with the activated EGFR, which activates SOS [52]. Activated SOS promotes recruition of Ras proteins to the activated EGFR. Through its GEF activity, SOS then induces GDP removal from Ras proteins, which can subsequently bind GTP and become active. Activated Ras recuits Raf proteins to the cell membrane and binds to their N-terminus. The activation of Raf, serine/threonine kinase proteins, is complex. In fact, Raf proteins are considered as gatekeepers of the RAS–RAF–MAPK pathway. In its inactive form, Raf is present in a 'closed' conformation, in which an autoinhibitory domain blocks the catalytic kinase domain [53]. Recruitment to the plasma membrane of Raf proteins by Ras results in a conformational change [54], which disrupts the autoinhibitory interaction of Raf. Rafs then form

homo– or heterodimers, which leads to partial activation by allostery. Transphorylative events, with optional phosphorylation by other kinases, such as PAK1 [55], then fully activates Raf. Activated Raf can now bind to MEKS, which are tyrosine/threonine kinases. MEKs phosphorylate ERKSs, which are also serine/threonine kinase enzymes. ERKs then translocate to the cell nucleus, where they influence expression of target genes [55]. RAS–RAF–MEK–ERK signaling promotes cell-cycle progression, cell differentiation, growth and survival [55].

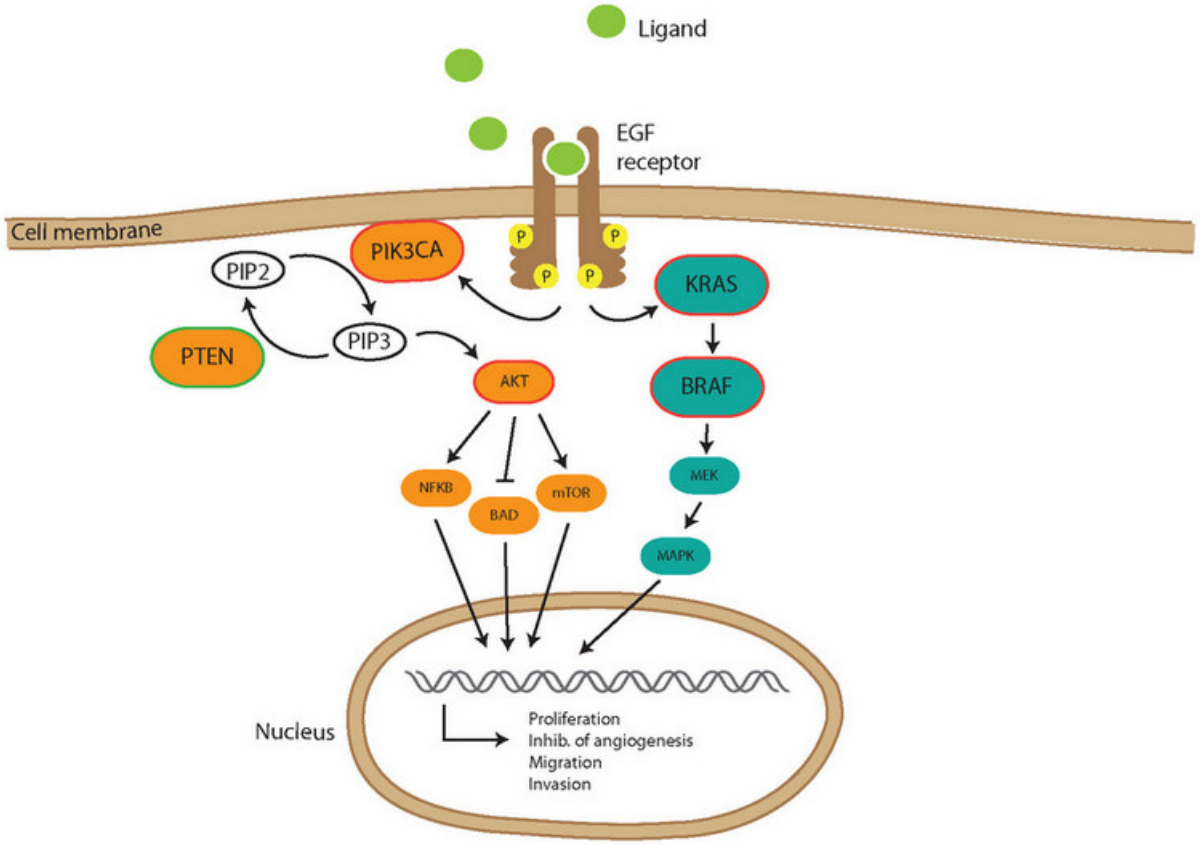

Figure 1: Schematic representation of the EGFR signaling cascade

1.4.1 Biological Role of EGFR in Solid Tumors

The EGFR pathway is long known to be dysregulated in most solid tumors and thereby presents a rational target for cancer therapy [56]. In normal cells, the EGFR signaling pathway drives cell-cycle progression, affects differentiation and migration and acts as a survival signal. EGFR levels have been shown to be higher in tumor samples than in surrounding tissues. Also, more of EGFR's ligands EGF and TGFα are found in these tumors. EGFR overexpression may result from epigenetic alterations or gene amplification.

EGFR signaling as a survival signal: Cells communicate with their environment. Without extracellular survival signals, a cell undergoes apoptosis. Signaling pathways induced by cell–matrix interactions, cell-cell interactions, and soluble survival factors act on a variety of genes and proteins. Loss of matrix attachment leads to growth arrest and death in normal epithelial cells, a process called anoikis [57]. The cell–matrix interaction acts as a safeguard against inappropriate proliferation and migration [58].

Activation of the EGFR signaling pathway protects normal epithelial cells against anoikis in the suspended state. EGFR blockade sensitizes normal epithelial cells to apoptosis, but the effect is much more pronounced in the suspended state than in the attached state [57] [59]. The redundancy of cell survival signals makes normal epithelial cells relatively resistant to EGFR-blockade in their normal microenvironment.

Tumor cells are often in transit or at sites with inadequate matrix composition [60]. They are often provided with inadequate or missing cell–cell and cell–matrix interactions. They are thereby more dependent on survival signals propagated by soluble mediators, such as EGF or TGFα [58]. Consequently, tumor cells are relatively sensitive against blockade of EGFR. This is counterbalanced by an upregulation of cell surface receptors that activate anti–apoptotic pathways. MAPK activation in the EGFR–activated RAS–RAF–MEK–ERK pathway is required for high expression of Bcl–XL, an anti-apoptotic protein of the Bcl–2 family [61]. Bcl–2 proteins regulate liberation of cytochrome c from the mitochondria, which is essential in the apoptotic caspase pathway. Additionally, EGFR signaling leads to post-transcriptional phosphorylation on the pro-apoptotic Bad protein, which is thereby functionally inactivated [59].

1.4.2 EGFR-targeted drugs

The observations that EGFR is recursively upregulated in many cancers and that EGFR is such an important mediator of cell-cycle progression, cell growth, and cell survival, has lead to the development of agents that block this pathway. Pharmacologically, these agents can be classed by their mode of action: EGFR-targeted monoclonal antibodies and EGFR-specific tyrosine kinase inhibitors (TKIs) [58]. Several proteins acting downstream of EGFR are often found to be mutated. This lead to the development of other targeted drugs, such as BRAF– or MEK–inhibitors. As many tumors are 'EGFR–addicted' and normal cells usually are provided with a redundancy of survival signals, these agents have less side–effects than classical cytotoxic chemotherapeutic agents [58].

Anti–EGFR monoclonal antibodies bind to the extracellular domain of EGFR in its inactive state

and thereby compete for EGF or TGFα binding. They inhibit ligand-induced EGFR tyrosine kinase activation. The most popular anti–EGFR monoclonal antibody is cetuximab, which binds to EGFR with a higher affinity than the natural ligands EGF or TGFα [62]. Binding of the antibody induces internalization of EGFR, followed by its degradation. Cetuximab also binds to EGFRvIII, a constantly active version of EGFR [63]. Cetuximab blocks cell–cycle progression at the G0/G1 boundary, inhibits cell proliferation and induces cancer cell death [64].

EGFR-specific tyrosine kinase inhibitors comprise three classes, which include first generation reversible EGFR-inhibitors (gefitinib, erlotinib), second generation irreversible inhibitors (afatinib, dacomitinib, neratinib) and third generation mutant-selective inhibitors (brigatinib, osimertinib, rociletinib) (mycancergenome.org). Third generation agents have a better sensitivity against mutated than wild-type EGFR and have been designed to further decrease treatment-associated side effects [65]. TKIs are low molecular weight molecules that are mainly derived from quinazoline [66]. These compounds bind to EGFR and block ligand-induced receptor phosphorylation by occluding the ATP–binding site.

1.4.3 Predictive markers

The mutational status of EGFR and downstream proteins is predictive of the potential success of EGFR-targeted therapy. Demonstration of the predictive value of these markers is not trivial and has to be proved in clinical trials. In many cases the predictive value of a marker, even though theoretically reasonable, has not yet been established. Currently, EGFR, KRAS, NRAS and BRAF are used as predictive markers for EGFR–targeted therapy in solid tumors.

EGFR is a strong predictive biomarker for the success of the administration of EGFR-specific TKIs. EGFR activating mutations are observed in 10–35% of NSCLC [67]. 90% of EGFR mutations are exon 19 deletions (48%) and exon 21 L585R (c.2573T$>$G) (43%) point mutations [68]. In melanoma and CRC, EGFR mutations are seldom. These mutations confer increased sensitivity to EGFR–specific tyrosine kinase inhibitors [68]. Patients with EGFR–mutated tumors have a longer progression–free survival if treated with TKIs than those treated with traditional chemotherapy [68]. Also, patients with EGFR–mutated tumors display a better prognosis if treated with EGFR TKIs compared to patients with wild–type EGFR cancers [68].

KRAS mutations are found in 36–40% of CRCs [69], 15–25% of NSCLC [70], and in 2% of melanomas [71]. Critical mutations in the KRAS gene include point mutations in codons 12, 13 and 61 [70] [69]. Amongst these, the G12C variant is the most common. These mutations lock KRAS in its GTP-bound state, resulting in a constantly active protein, leading to a constantly active signal transduction [72]. Blocking EGFR in that case is useless, as KRAS acts downstream of EGFR. KRAS activating mutations have been shown to confer reduced sensitivity to EGFR-targeted monoclonal antibodies in CRC and EGFR–TKIs in NSCLC [70] [69].

NRAS is an isoform of KRAS. Activating mutations in NRAS codons 12 and 61 are found in 1–6% of CRCs [73], 13–25% of melanomas [74] and 1% of NSCLCs [75]. NRAS mutations have been associated with reduced sensitivity to EGFR monoclonal antibodies in CRC [73]. The predictive value of the influence of NRAS mutations in NSCLC and melanoma has not yet been demonstrated in large clinical trials at this time.

BRAF mutations are very common in melanoma (37–50%) [76]. 8–15% of CRCs [77] and 1–4% of NSCLCs [78] are BRAF–mutant positive. Amongst BRAF-mutated melanomas, the V600E variant is found in 80–90% cases [76]. V600E affects the activation segment of the BRAF kinase domain and results in an increased kinase activity [79]. BRAF mutations usually confer a resistance to EGFR-targeted therapy in KRAS wild–type tumors. BRAF V600E mutations have been associated with increased sensitivity to BRAF inhibitors in melanoma and NSCLS and MEK inhibitors in melanoma [76].

1.5 Tumor DNA Sequencing

The completion of the Human Genome Project in 2001 resulted in a massive boost in molecular medicine. New high-throughput techniques, in combination with advanced computational performance and storage capacities, lead to an explosion of biological data. Amongst the many mutation detection techniques, Next–Generation Sequencing (NGS) constitutes the most powerful method and allows deep insights into the underlying causes of diseases. Even though advances in sequencing technology and computational power and tools have decreased the time and cost of a sequencing experiment, NGS is still mainly used in research, with only a few laboratories using this technique in diagnostics. NGS has profoundly impacted the field of oncology. A wide variety of NGS applications have been applied to study the genetics and epigenetics of cancer. ChIP–Seq and FAIRE–Seq allow determination of DNA–protein interactions and identification of DNA regulatory

elements, respectively. The cancer transcriptome can be studied with RNA–Seq experiments. NGS has massively accelerated discovery of genetic and epigenetic alterations in tumors.

The development of targeted NGS approaches have made it possible to implement NGS in clinical diagnostics laboratories. NGS allows deeper insights into the genetic sequences of oncogenes and tumor suppressor genes than array–based and single–gene approaches, which are currently used in most laboratories. Multiple genes can be studied in a single experiment, while classical methods are much more restricted in that regard. Targeted NGS experiments differ from whole–genome or whole–exome sequencing, as only a selection of regions of interest (ROIs) is captured and sequenced. This approach increases the efficiency of the experiment and the number of samples per run. Additionally, coverage, e.g. the number of sequencing reads that align to a specific base of the reference genome, is drastically increased in targeted NGS. This offers the sensitivity to detect low–frequency mutations in the tumor sample. Targeted NGS methods can be applied to study insertions, deletions and point mutations in genes of interest. NGS can thereby guide personalized cancer therapy by identifying the mutational status of genetic markers, which are predictive of the potential success of targeted cancer treatment.

1.5.1 Practical implications in the laboratory

Several studies have demonstrated the power of targeted NGS in the identification of the mutational status of EGFR, KRAS and BRAF. Comparison of NGS and real time PCR–based approaches showed high concordance of results. NGS was able to detect mutations that were not detectable by qPCR. Still, the implementation of new techniques into the workflow of molecular diagnostics laboratories requires a careful assessment of the sensitivity and sensibility of the method [80]. The quality of tumor sequencing experiments is affected by several factors. These include the content of tumor cells in the sample, the quality of the tissue material, the choice of the sequencing library preparation kit and the bioinformatic pipeline.

The tumor biopsy usually consists of an admixture of normal and cancer cells. The sensitivity of tumor variant detection is linked to the tumor cell content of the specimen. In addition, cancers are highly heterogeneous: a small subpopulation might present mutations, which provide resistance to the treatment. Detecting these low-frequency mutations and clearly delineating them from possible sample processing or sequencing induced artifacts presents an important challenge [81].

Tumor biopsies usually yield a limited amount of tissue, therefore it is important to optimize sample usage by multiplexing analysis. In Luxembourg, all relevant tumor biopsies are sent to the

Laboratoire National de Santé (LNS) to the Service of Pathologic Anatomy. Here, the biopsy is fixed in formalin and embedded in paraffin (FFPE). FFPE preserves the tissue morphology and thereby allows histological analysis. In addition, it allows specimen storage for decades. Sample quality, however, is influenced by this fixation method and the fixation time [80]. DNA extraction from FFPE samples is difficult and yields low amounts of DNA [82]; formaldehyde leads to cross-linking of nucleic acids and proteins [83]; FFPE introduces fixation artifacts into DNA sequences, for instance C¿T transitions [84]. These circumstances complicate sample processing as well as NGS data interpretation. Though, FFPE samples have been shown to be still suitable for downstream PCR–based analyses [85].

Several NGS bench-top devices have become available in the last decade [86]. These instrumentations differ in their underlying chemistry that influences the instrument's performance, accuracy, output and time per run. Common sequencing principles include pyrosequencing (454), sequencing by ligation (SOLiD), ion semiconductor sequencing (Ion Torrent) and sequencing by synthesis (Illumina).

Sequencing library preparation also affects the final NGS result. Several technologies for target enrichment exist and are available for different sequencing instruments [87]. Two steps are essential for all these enrichment methods: target sequences have to be enriched and samples have to be multiplexed, which requires the incorporation of a unique index combination for each sample. Target enrichment methods can be separated into three basic groups: targeted circularization, hybrid capture of target fragments and PCR-based enrichment methods. Multiplex PCR–based approaches produce short DNA fragments of target regions in a first PCR. In a second PCR, adapters and indices are added. Hybridization–based methods require a so-called shotgun library construction before target regions can be captured. During this process, genomic DNA is randomly sheared into small fragments and an adapter– and index–linked library is produced. Biotinylated baits are added that bind to target sequences. Target regions can then be captured using streptadivin coated magnetic beads. Targeted circularization methods rely on a digestion of DNA by restriction enzymes. The produced DNA fragments are circularized and captured. Only circularized target are then amplified by PCR.

The establishment and validation of a bioinformatic NGS data analysis pipeline still presents a challenge in diagnostics. After generation of FASTQ files of the sequencer, data generally undergo quality control, followed by trimming of low quality bases, alignment to the reference genome, variant calling and variant annotation. For each of these steps, several bioinformatic algorithms and tools exist [88]. The computational pipeline of the molecular pathology laboratory has to incorporate the

tools that allow the most sensitive and sensible analysis of data. For instance, quality trimming influences the mapping to the reference genome. The mapping, in turn, strongly affects the variant calling. In fact, variant calling is a critical step in NGS data analysis. Several variant calling tools exist, but vary in their false-positive and false-negative detection rates [89] [90]. These tools have to be carefully assessed, as false-positives or false-negatives should absolutely be avoided when it comes to the prescription of a targeted chemotherapeutic agent.

To facilitate interpretation of NGS data, variants have to be annotated and their clinical actionability has to be identified. Several databases have emerged in this field and numerous tools allow to automatize variant annotation. Here again, the choice of the database and the variant annotator is important.

Finally, the sample–to–results time is a very pragmatic, but important factor. The time from the biopsy to the potential start of an administration of a targeted chemotherapeutic drug should be reduced to a reasonable minimum. For instance, in case of late–stage cancer patients, it would be unacceptable if analysis would take several weeks. To reduce the sample–to–results time, the sample processing workflow should be as short as possible, while still yielding high quality sequencing libraries. The bioinformatic pipeline should not only incorporate the best tools, but should also be automatized to further reduce the time of analysis.

1.6 Aims of the Thesis

References

[1] M. Berwick and P. Vineis, “Markers of DNA Repair and Susceptibility to Cancer in Humans: an Epidemiologic Review,” *JNCI J Natl Cancer Inst*, vol. 92, no. 2, pp. 874–897, 2000.

[2] D. Hanahan and R. A. Weinberg, “Hallmarks of Cancer: The Next Generation,” *Cell*, vol. 114, no. 5, pp. 646—-674, 2011.

[3] D. Hao, L. Wang, and L. Di, “Distinct mutation accumulation rates among tissues determine the variation in cancer risk,” *Scientific Reports*, vol. 6, 2016.

[4] I. Tomlinson, P. Sasieni, and W. Bodmer, “How Many Mutations in a Cancer?,” *Am J Pathol.*, vol. 160, pp. 755–758, 2002.

[5] K. R. Loab and L. A. Loeb, “Significance of multiple mutations in cancer,” *Carcinogenesis*, vol. 21, pp. 376–385, 2000.

[6] T. Kauppinena, J. Toikkanena, D. Pedersenb, R. Youngb, W. Ahrensc, P. Boffettad, J. Hansene, H. Kromhoutf, J. M. Blascog, D. Mirabellih, V. Orden-Riverag, B. Pannetti, N. Platoj, A. Savelaa, R. Vincentk, and l. M Kogevinasd, “Occupational exposure to carcinogens in the European Union,” *Occup Environ Med*, vol. 57, pp. 10–18, 2000.

[7] M. Kogevinas, M. Sala, P. Boffetta, N. Kazerouni, H. Kromhout, and S. Hoar-Zahm, “Cancer risk in the rubber industry: a review of the recent epidemiological evidence,” *Occup Environ Med*, vol. 55, pp. 1–12, 1998.

[8] J. Cornfield, W. Haenszel, E. C. Hammond, A. M. Lilienfeld, M. B. Shimkin, and E. L. Wynder, “Smoking and lung cancer: recent evidence and a discussion of some questions,” *Int. J. Epidemiol.*, vol. 38, pp. 1175–1191, 2009.

[9] P. T. Schumacker, “Reactive oxygen species in cancer cells: Live by the sword, die by the sword,” *Cancer Cell*, vol. 10, pp. 241–252, 2006.

[10] G. Waris and H. Ahsan, “Reactive oxygen species: role in the development of cancer and various chronic conditions,” *Journal of Carcinogenesis*, vol. 5, 2006.

[11] F. Dietlein, L. Thelen, and H. C. Reinhardt, “Cancer-specific defects in dna repair pathways as targets for personalized therapeutic approaches,” *Trends in Genetics*, vol. 30, no. 8, 2014.

[12] C. Boland and A. Goel, “Microsatellite instability in colorectal cancer,” *Gastroenterology*, vol. 138, pp. 2073–2087, 2010.

[13] R. Kanthan, J. Senger, and S. C. Kanthan, “Molecular Events in Primary and Metastatic Colorectal Carcinoma: A Review,” *Pathology Research International*, 2012.

[14] V. Deschoolmeester, M. Baay, P. Specenier, F. Lardon, and J. B. Vermorken, “A review of the most promising biomarkers in colorectal cancer: one step closer to targeted therapy,” *Oncologist*, vol. 15, pp. 699–731, 2010.

[15] M. Giam and G. Rancati, “Aneuploidy and chromosomal instability in cancer: a jackpot to chaos,” *Cell Div*, vol. 10, 2015.

[16] J. Forment, A. Kaidi, and S. Jackson, “Chromothripsis and cancer: causes and consequences of chromosome shattering,” *Nature Reviews Cancer*, vol. 12, pp. 663–670, 2012.

[17] A. Rode, K. Maass, K. Willmund, P. Lichter, and A. Ernst, “Chromothripsis in cancer cells: An update,” *Int J Cancer*, vol. 138, pp. 2322–2333, 2016.

[18] M. Greaves and C. C. Maley, “Clonal Evolution in Cancer,” *Nature*, vol. 481, pp. 306–313, 2012.

[19] M. Gerlinger and C. Swanton, “How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine,” *British Journal of Cancer*, vol. 103, pp. 1139—1143, 2010.

[20] J. Breivik, “Don’t stop for repairs in a war zone: Darwinian evolution unites genes and environment in cancer development,” *PNAS*, vol. 98, no. 10, pp. 5379–5381, 2001.

[21] M. P. Little, “Cancer models, genomic instability and somatic cellular Darwinian evolution,” *PNAS*, vol. 5, no. 19, 2010.

[22] M. R. Stratton, P. J. Campbell, and P. A. Futreal, “The cancer genome,” *Nature*, vol. 458, pp. 719–724, 2009.

[23] C. D. McFarland, K. S. Korolev, G. V. Kryukov, S. R. Sunyaev, and L. A. Mirny, “Impact of deleterious passenger mutations on cancer progression,” *Proc Natl Acad Sci USA*, vol. 110, pp. 2910—-2915, 2013.

[24] S. K. Merid, D. Goranskaya, and A. Alexeyenko, “Distinguishing between driver and passenger mutations in individual cancer genomes by network enrichment analysis,” *BMC Bioinformatics*, vol. 15, 2014.

[25] I. Bozic, T. Antal, H. Ohtsuki, H. Carter, D. Kim, S. Chen, R. Karchin, K. W. Kinzler, B. Vogelstein, and M. A. Nowak, “Accumulation of driver and passenger mutations during tumor progression,” *Proc Natl Acad Sci USA*, vol. 107, pp. 18545—18550, 2010.

[26] M. Cusnir and L. Cavalcante, “Inter-tumor heterogeneity,” *Human Vaccines and Immunotherapeutics*, vol. 8, 2012.

[27] F. Michor and K. Polyak, “The Origins and Implications of Intratumor Heterogeneity,” *Cancer Prev Res*, vol. 3, pp. 1361–1364, 2010.

[28] W. Sun and J. Yang, “Functional Mechanisms for Human Tumor Suppressors,” *J Cancer*, vol. 1, pp. 136–140, 2010.

[29] H. Chial, “Tumor Suppressor (TS) Genes and the Two-Hit Hypothesis,” *Nature Education*, vol. 1, 2010.

[30] M. Santarosa and A. Ashworth, “Haploinsufficiency for tumour suppressor genes: when you don’t need to go all the way,” *Biochim Biophys Acta.*, vol. 1654, pp. 105–122, 2004.

[31] B. T. MacDonald, K. Tamai, and X. He, “Wnt/β-catenin signaling: components, mechanisms, and diseases,” *Dev Cell*, vol. 17, no. 1, pp. 9–26, 2009.

[32] P. Polakis, “Wnt signaling and cancer,” *Genes Dev*, vol. 14, pp. 1837–1851, 2000.

[33] M. Farnebo, V. J. Bykov, and K. G. Wiman, “The p53 tumor suppressor: a master regulator of diverse cellular processes and therapeutic target in cancer,” *Biochem Biophys Res Commun.*, vol. 396, no. 1, pp. 85–89, 2010.

[34] Y. Yang, C. H. Li, and A. M. Weissman, “Regulating the p53 system through ubiquitination,” *Oncogene*, vol. 23, pp. 2096–2106, 2004.

[35] J. T. Zilfou and S. W. Lowe, “Tumor Suppressive Functions of p53,” *Cold Spring Harb Perspect Biol.*, vol. 1, no. 5, 2009.

[36] M. Olivier, M. Hollstein, and P. Hainaut, “TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use,” *Cold Spring Harb Perspect Biol.*, vol. 2, no. 1, 2010.

[37] C. M. Croce, “Oncogenes and Cancer,” *N Engl J Med*, vol. 358, pp. 502–511, 2008.

[38] J. Yokota and T. Kohno, “Molecular footprints of human lung cancer progression,” *Cancer Sci.*, vol. 95, no. 3, pp. 197–204, 2004.

[39] S. Mouret, C. Baudouin, M. Charveron, A. Favier, J. Cadet, and T. Douki, “Cyclobutane pyrimidine dimers are predominant DNA lesions in whole human skin exposed to UVA radiation,” *PNAS*, vol. 103, no. 37, pp. 13765–13770, 2011.

[40] K. Paszkowska-Szczur, R. J. Scott, P. Serrano-Fernandez, A. Mirecka, P. Gapska, B. Górski, C. Cybulski, R. Maleszka, M. Sulikowski, L. Nagay, J. Lubinski, and T. D. T, "Xeroderma pigmentosum genes and melanoma risk," *Int J Cancer*, vol. 133, no. 5, pp. 1094–1100, 2013.

[41] R. A. Anvekar, J. J. Asciolla, D. J. Missert, and J. E. Chipuk, "Born to be alive: a role for the BCL-2 family in melanoma tumor cell survival, apoptosis, and treatment," *Frontiers in Oncology*, vol. 1, no. 34, 2011.

[42] B. Bandarchi, L. Ma, R. Navab, A. Seth, and G. Rasty, "From Melanocyte to Metastatic Malignant Melanoma," *Dermatol Res Pract.*, 2010.

[43] S. D. Markowitz and M. M. Bertagnolli, "Molecular Basis of Colorectal Cancer," *The new england journal of medicine*, vol. 361, no. 55, pp. 2449–2460, 2009.

[44] R. J. Davies, R. Miller, and N. Coleman, "Colorectal cancer screening: prospects for molecular stool analysis," *Nature Reviews Cancer*, vol. 5, pp. 199–209, 2005.

[45] R. S. Herbst, "Review of epidermal growth factor receptor biology," *Int J Radiat Oncol Biol Phys.*, vol. 59, pp. 21–26, 2004.

[46] K. M. Ferguson, "Structure-based view of epidermal growth factor receptor regulation," *Annu Rev Biophys.*, vol. 37, pp. 353–373, 2008.

[47] J. LoPiccolo, G. M. Blumenthal, W. B. Bernstein, and P. A. Dennis, "Targeting the PI3K/Akt/mTOR pathway: effective combinations and clinical considerations," *Drug Resist Updat*, vol. 11, pp. 32–50, 2008.

[48] B. A. Hemmings and D. F. Restuccia, "PI3K-PKB/Akt Pathway," *Cold Spring Harb Perspect Biol*, 2012.

[49] J. R. Hart and P. K. Vogt, "Phosphorylation of AKT: a Mutational Analysis," *Oncotarget*, vol. 2, no. 6, pp. 467—-476, 2011.

[50] I. Hers, E. E. Vincent, and J. M. Tavaré, "Akt signalling in health and disease," *Cell Signal*, vol. 23, no. 10, pp. 1515–1527, 2011.

[51] J. F. Vara, A. Casado, J. de Castro, P. Cejas, C. Belda-Iniesta, and M. González-Barón, "PI3K/Akt signalling pathway and cancer," *Cancer Treat Rev.*, vol. 30, pp. 193–204, 2004.

[52] E. H. Lowenstein, R. J. Daly, A. G. Batzer, W. Li, B. Margolis, R. Lammers, A. Ullrich, E. Y. Skolnik, D. Bar-Sagi, and J. Schlessinger, "The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to Ras signaling," *Cell*, vol. 70, no. 3, pp. 431–442, 1992.

[53] N. H. Tran, X. Wu, and J. A. Frost, “B-Raf and Raf-1 are regulated by distinct autoregulatory mechanisms,” *J Biol Chem.*, vol. 280, no. 6, pp. 16244–16253, 2005.

[54] K. Hibino, T. Shibata, T. Yanagida, and Y. Sako, “A RasGTP-Induced Conformational Change in C-RAF Is Essential for Accurate Molecular Recognition,” *Biophys J.*, vol. 97, no. 5, p. 1277–1287, 2009.

[55] M. Zang, C. Hayne, and Z. Luo, “Interaction between active Pak1 and Raf-1 is necessary for phosphorylation and activation of Raf-1,” *J Biol Chem.*, vol. 277, no. 6, pp. 4395–4405, 2002.

[56] C. Artega, “Targeting HER1/EGFR: A molecular approach to cancer therapy,” *Seminars in Oncolog*, vol. 30, no. 3, pp. 3–14, 2016.

[57] U. Rodeck, M. Jost, J. DuHadaway, C. Kari, P. J. Jensen, B. Risse, and D. L. Ewert, “Regulation of Bcl–xL expression in human keratinocytes by cell–substratum adhesion and the epidermal growth factor receptor,” *Proc. Natl. Acad. Sci. USA*, vol. 94, pp. 5067–5072, 1994.

[58] C. Kari, T. O. Chan, M. R. de Quadros, and U. Rodeck, “Targeting the Epidermal Growth Factor Receptor in Cancer: Apoptosis Takes Center Stage,” *Cancer Research*, vol. 63, pp. 1–5, 2003.

[59] A. P. Gilmore, A. J. Valentijn, P. Wanger, A. M. Ranger, N. Bundred, M. J. O’Hare, A. Wakeling, S. J. Korsmeyer, and C. H. Streuli, “Activation of BAD by therapeutic inhibition of epidermal growth factor receptor and transactivation by insulin–like growth factor receptor,” *J. Biol. Chem.*, vol. 277, pp. 27643–27650, 2002.

[60] W. G. Stetler-Stevenson, S. Aznavoorian, , and L. A. Liotta, “Tumor Cell Interactions with the Extracellular Matrix During Invasion and Metastasis,” *Annual Review of Cell Biology*, vol. 9, pp. 541–573, 1993.

[61] M. Jost, T. M. Hugget, C. Kari, L. H. Boise, and U. Rodeck, “Epidermal growth factor receptor–dependent control of keratinocyte survival and Bcl–xL expression through a MEK–dependent pathway,” *J. Biol. Chem.*, vol. 276, pp. 6320–6326, 2001.

[62] J. Baselga, “The EGFR as a target for anticancer therapy—focus on cetuximab,” *European Journal of Cancer*, vol. 37, pp. 16–22, 2001.

[63] D. Patel, A. Lahiji, S. Patel, M. Franklin, X. Jimenez, D. J. Hicklin, and X. Kang, “Monoclonal Antibody Cetuximab Binds to and Down-regulates Constitutively Activated Epidermal Growth Factor Receptor vIII on the Cell Surface,” *Anticancer Research*, vol. 27, pp. 3355–3366, 2007.

[64] D. Raben, B. Helfrich, D. C. Chan, F. C. L. Zhao, W. Franklin, A. E. Barón, C. Zeng, T. K. Johnson, and P. A. Bunn, "The effects of cetuximab alone and in combination with radiation and/or chemotherapy in lung cancer," *Clin Cancer Res*, vol. 11, pp. 795–805, 2005.

[65] B. C. Liao, C. C. Lin, and J. C. Yang, "Second and third-generation epidermal growth factor receptor tyrosine kinase inhibitors in advanced nonsmall cell lung cancer," *Curr Opin Oncol.*, vol. 27, pp. 94–101, 2015.

[66] N. Normanno, M. R. Maiello, and A. D. Luca, "Epidermal growth factor receptor tyrosine kinase inhibitors (EGFR-TKIs): simple drugs with a complex mechanism of action?," *J Cell Physiol.*, vol. 194, pp. 13–19, 2003.

[67] T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman, and D. A. Haber, "Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib," *N Engl J Med.*, vol. 350, no. 21, pp. 2129–2139, 2004.

[68] C. Lovly, L. Horn, and W. Pao, "EGFR in Non-Small Cell Lung Cancer (NSCLC)," *My Cancer Genome https://www.mycancergenome.org/content/disease/lung-cancer/egfr/*, 2015, as of 20.4.2016.

[69] E. Chan, "KRAS in Colorectal Cancer.," *My Cancer Genome https://www.mycancergenome.org/content/disease/colorectal-cancer/kras/*, 2015, as of 20.4.2016.

[70] C. Lovly, L. Horn, and W. Pao, "KRAS in Non-Small Cell Lung Cancer (NSCLC)," *My Cancer Genome https://www.mycancergenome.org/content/disease/lung-cancer/kras/*, 2015, as of 20.4.2016.

[71] H. Tsao, L. Chin, L. A. Garraway, and D. E. Fisher, "Melanoma: from mutations to medicine," *Genes and Dev.*, vol. 26, no. 4, pp. 1131–1155, 2012.

[72] S. Schubbert, G. Bollag, N. Lyubynska, H. Nguyen, C. P. Kratz, M. Zenker, C. M. Niemeyer, A. Molven, and K. Shannon, "Biochemical and Functional Characterization of Germ Line KRAS Mutations," *Mol. Cell. Biol.*, vol. 27, no. 22, pp. 7765–7770, 2007.

[73] E. Chan, "NRAS in Colorectal Cancer," *My Cancer Genome https://www.mycancergenome.org/content/disease/colorectal-cancer/nras/*, 2015, as of 20.4.2016.

[74] C. Lovly, W. Pao, and J. Sosman, “NRAS in Melanoma,” *My Cancer Genome https://www.mycancergenome.org/content/disease/melanoma/nras/*, 2015, as of 20.4.2016.

[75] C. Lovly, L. Horn, and W. Pao, “NRAS in Non-Small Cell Lung Cancer (NSCLC),” *My Cancer Genome https://www.mycancergenome.org/content/disease/lung-cancer/nras/*, 2015, as of 20.4.2016.

[76] C. Lovly, W. Pao, and J. Sosman, “BRAF in Melanoma,” *My Cancer Genome https://www.mycancergenome.org/content/disease/melanoma/braf/*, 2015, as of 20.4.2016.

[77] D. Stover, “BRAF in Colorectal Cancer,” *My Cancer Genome https://www.mycancergenome.org/content/disease/colorectal-cancer/braf/*, 2015, as of 20.4.2016.

[78] C. Lovly, L. Horn, and W. Pao, “BRAF in Non-Small Cell Lung Cancer (NSCLC),” *My Cancer Genome https://www.mycancergenome.org/content/disease/lung-cancer/braf/*, 2015, as of 20.4.2016.

[79] E. R. Cantwell-Dorris, J. J. O’Leary, and O. M. Sheils, “BRAF V600E: Implications for Carcinogenesis and Molecular Therapy,” *Mol Cancer Ther*, vol. 10, no. 385, 2011.

[80] R. Simon and S. Roychowdhury, “Implementing personalized cancer genomics in clinical trials,” *Nature Reviews Drug Discovery*, vol. 12, pp. 358–369, 2013.

[81] S. Q. Wong, J. Li, A. Y. Tan, R. Vedururu, J. B. Pang, H. Do, J. Ellul, K. Doig, A. Bell, G. A. MacArthur, S. B. Fox, D. M. Thomas, A. Fellowes, J. P. Parisot, and A. Dobrovic, “Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel sequencing,” *BMC Medical Genomics*, vol. 7, no. 23, 2014.

[82] M. T. Gilbert, T. Haselkorn, M. Bunce, J. J. Sanchez, and S. B. Lucas, “The Isolation of Nucleic Acids from Fixed, Paraffin-Embedded Tissues– Which Methods Are Useful When?,” *PLoS ONE*, vol. 2, no. 6, 2007.

[83] R. Thavarajah, V. K. Mudimbaimannar, J. Elizabeth, U. K. Roa, and K. Ranganathan, “Chemical and physical basics of routine formaldehyde fixation,” *J Oral Maxillofac Pathol.*, vol. 16, no. 3, 2012.

[84] H. Do, S. Q. Wong, J. Li, and A. Dobrovic, “Reducing Sequence Artifacts in Amplicon-Based Massively Parallel Sequencing of Formalin-Fixed Paraffin-Embedded DNA by Enzymatic Depletion of Uracil-Containing Templates,” *Clinical Chemistry*, vol. 59, no. 9, 2013.

[85] N. Ludyga, B. Grènwald, O. Azimzadeh, S. Englert, H. Höfler, S. Tapio, and M. Aubele, “Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses,” *Virchows Arch*, vol. 460, pp. 131–140, 2012.

[86] C. Meldrum, M. A. Doyle, and R. W. Tothill, “Next-Generation Sequencing for Cancer Diagnostics: a Practical Perspective,” *Clin Biochem Rev*, vol. 32, no. 385, 2011.

[87] J. Haas, I. Barb, H. A. Katus, and B. Meder, “Targeted Next-Generation Sequencing: The Clinician’s Stethoscope for Genetic Disorders,” *Personalized Medicine*, vol. 11, no. 6, pp. 581–592, 2014.

[88] S. Pabinger, A. Dander, M. Fischer, R. Snajder, M. Sperk, M. Efremova, B. Krabichler, M. R. Speicher, J. Zschocke, and Z. Trajanoski, “A survey of tools for variant analysis of next-generation genome sequencing data,” *Briefings in Bioinformatics*, vol. 15, no. 2, pp. 256–278, 2013.

[89] J. O’Rawe, T. Jiang, G. Sun, Y. Wu, Q. Wang, J. Hu, P. Bodily, L. Tian, H. Hakonarson, W. E. Johnson, Z. Wei, K. Wang, and G. J. Lyon, “Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing,” *Genome Medicine*, vol. 15, no. 28, 2013.

[90] D. H. Spencer, M. Tyagi, F. Vallania, A. J. Bredemeyer, J. D. Pfeifer, R. D. Mitra, and E. J. Duncavage, “Performance of Common Analysis Methods for Detecting Low-Frequency Single Nucleotide Variants in Targeted Next-Generation Sequence Data,” *The Journal of Molecular Diagnostics*, vol. 16, no. 1, 2014.