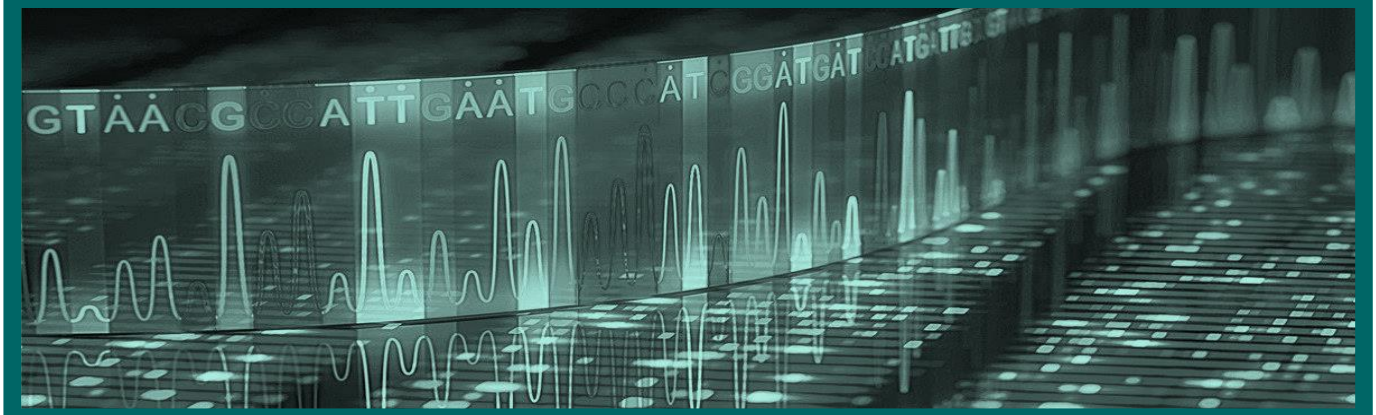
FACULTAD DE MEDICINA
UNIVERSIDAD DE CANTABRIA

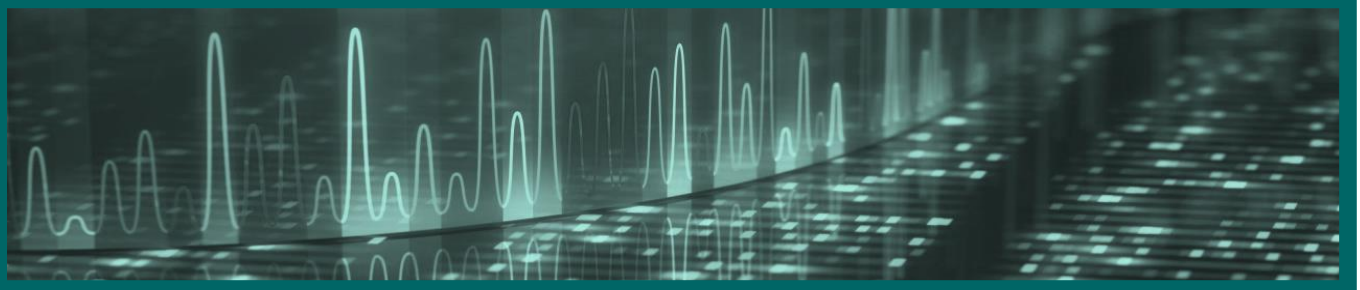
GRADO EN MEDICINA

TRABAJO DE FIN DE GRADO



Advances in the molecular diagnosis of cancer

Avances en el diagnóstico molecular del cáncer



Autora: Natalia Cadrecha Sánchez

Directora: Mª Dolores Delgado Villar

Codirectora: Ana Batlle López

Santander, junio 2018

SUMMARY

Cancer is a heterogeneous disease, classically classified and diagnosed according to its morphological characteristics, which does not reflect its great complexity. The knowledge of cancer has been enormously driven in recent decades due to the advances on molecular biology techniques used for the diagnosis, classification and treatment of cancer. In the first part of this work, we will review the hallmarks of cancer, as well as the genes that are mutated and the cellular pathways altered in tumor cells, recently identified by international consortiums investigating in cancer genomics. Next, advances in the molecular diagnosis of hematological and solid tumors will be addressed, focusing in chronic myeloid leukemia and breast cancer as representative examples. Finally, next generation sequencing techniques and liquid biopsies are highlighted as new tools that are driving the field of precision oncology. These techniques are allowing a more precise characterization of neoplasia that will improve its management and will lead to treatment based on targeted therapies, more effective than conventional treatments.

Keywords: cancer, molecular diagnosis, biomarkers, precision oncology.

RESUMEN

El cáncer es una enfermedad heterogénea que clásicamente se ha clasificado y diagnosticado según sus características morfológicas, lo cual no refleja su gran complejidad. El conocimiento del cáncer ha sido enormemente impulsado en las últimas décadas gracias a los avances en las técnicas de biología molecular, utilizadas para el diagnóstico, clasificación y tratamiento del cáncer. En la primera parte de este trabajo se revisan las principales características del cáncer, los genes mutados y las rutas celulares alteradas en las células tumorales, identificadas recientemente por consorcios internacionales investigando en genómica del cáncer. A continuación se detallan los avances en el diagnóstico molecular de tumores hematológicos y tumores sólidos, los principales biomarcadores analizados y técnicas utilizadas, y se profundiza en la leucemia mieloide crónica y el cáncer de mama como ejemplos representativos. Finalmente, se destacan las técnicas de secuenciación de nueva generación y las biopsias líquidas como nuevas herramientas que están impulsando el campo de la oncología de precisión. Estas técnicas están permitiendo una caracterización más precisa que mejorará el manejo de las neoplasias y llevará al tratamiento mediante terapias dirigidas, más efectivas que los tratamientos convencionales.

Palabras clave: cáncer, diagnóstico molecular, biomarcadores, oncología de precisión.

INDEX

ABBREVIATIONS

aCGH: array-based Comparative Genomic Hybridization

aCML: atypical Chronic Myeloid Leukemia

ALL: Acute Lymphoid Leukemia

AML: Acute Myeloid Leukemia

ARGO: Accelerate Research in Genomic Oncology Project

ccfDNA: circulating cell-free DNA

cDNA: complementary DNA

CEBPA: CCAAT/enhancer-binding protein alpha

CEP17: centromeric probe for chromosome 17

CGH: Comparative Genomic Hybridization

CGP: Cancer Genome Project

CML: Chronic Myeloid Leukemia

CMML: Chronic Myelomonocytic Leukemia

CSF: cerebrospinal fluid

ctDNA: circulating tumor DNA

EFS: Event-Free Survival

ER: Estrogen Receptor

FACS: Fluorescence-Activated cell sorting

FISH: Fluorescence *In Situ* Hybridization

GRCh38.p12.: Genome Reference Consortium Human Build 38 patch release 12

HBC: Hereditary Breast CancerICGC: International Cancer Genome Consortium

IHC: Immunohistochemistry

JMML: Juvenile Myelomonocytic Leukemia

MALT: Mucosa-Associated Lymphoid Tissue

MAP-kinase: Mitogen-Activated Protein kinases

MDS: Myelodysplastic Syndrome

MPN: Myeloproliferative Neoplasia

NGS: Next Generation Sequencing

OS: Overall Survival

PCR: Polymerase Chain Reaction

pCR: pathologic Complete Response

PR: Progesterone Receptor

qPCR: quantitative Polymerase Chain Reaction

RT-PCR: Reverse Transcription Polymerase Chain Reaction

TCGA: The Cancer Genome Atlas

TKI: Tyrosine Kinase Inhibitors

TNM: Tumor, Nodes, Metastasis

TSG: Tumor Suppressor Gene

VEGF: Vascular Endothelial Growth Factor

WBC: White Blood Cell

1. INTRODUCTION

Cancer is the main cause of death in developed countries. Lung, colorectal, breast and prostate tumors are the most common causes of cancer deaths. Cancer is a group of highly heterogeneous diseases in which genetic alterations lead to the transformation of normal cells into tumor cells. During carcinogenesis, deregulation of certain genes due to a number of different mechanisms, cause a normal cell to grow in an uncontrolled manner, giving rise to an abnormal proliferation (Cooper & Hausman, 2014; Weinberg, 2014). Also, very frequently, genetic aberrations induce abnormalities of the differentiation and or in the apoptosis process as part of the malignant transformation.

The molecular study of cancer was delayed in comparison to other pathologies due to the great complexity of this disease. Classically, the approach to these diseases has been based in the classification and treatment on the tissue from which the neoplasia was arisen (Weinberg, 2014). Molecular techniques are changing this approach, leading to understanding cancer as a disease of the genome, with great differences between altered genes among cancers that belong to the same category in the classifications and understanding the mechanisms by which mutations in these genes cause malignancy (Müllauer, 2017). Advances in the molecular diagnosis of cancer will eventually lead to the development of personalized cancer therapy, in which each patient will be treated with drugs that target the specific type of tumor, considering the genetic alterations involved (Chmielecki & Meyerson, 2014; Khotskaya, 2017).

The first part of this work consists in an overview of the molecular bases of cancer, how tumors are traditionally classified and their impact on the overall mortality. The hallmarks of cancer, i.e. the characteristics that cancer cells progressively acquire and cause the malignancy are summarized (Hanahan & Weinberg, 2011). Then, the main features of genes related to cancer are explained: oncogenes, which promote tumorigenesis, and tumor suppressor genes that, under normal conditions, protect cells from overproliferation and promote their death if their genome is altered. The most important international consortiums that emerged from the Human Genome Project to investigate cancer as a genetic disease are presented, including their aims, the progress achieved in the characterization of cancer, as well as the cellular pathways that these genomic studies have been demonstrated to be altered.

The second part of this work deals with the molecular diagnosis of cancer. First, the advances in the molecular diagnosis of hematological tumors are reviewed, focusing in the techniques and the most important biomarkers used on these malignancies. Chronic myeloid leukemia is more deeply analyzed as an example of the approach to these neoplasias in terms of diagnosis, prognosis, treatment and follow-up analysis. Then the molecular diagnosis of solid tumors is addressed, taking breast cancer as a representative example.

Finally, new molecular techniques that are dramatically modifying the field of precision oncology are reviewed, highlighting the Next Generation Sequencing (NGS techniques. NGS has provided a rapid analysis of a big number of tumor genomes, that have been compared with normal genomes, allowing to discover a large number of

genes that could potentially cause cancer. NGS can be used in clinical practice for diagnostic and prognostic purposes and to establish potential therapeutic targets. Liquid biopsies are finally explained as a new technique of great utility in cancer that promises to reduce the need for solid biopsies. This technique not only facilitates establishing the genetic characteristics of the cancer at diagnosis, but also provides the possibility of monitoring of the changes that occur throughout the treatment.

2. MOLECULAR BASES OF CANCER

2.1. TYPES OF CANCER

2.1.1. Classic classifications

There are more than 100 types of cells in the human body and all of them might potentially develop a malignant tumor. Therefore, many types of cancer can be distinguished and it is necessary to systematize its nomenclature (Weinberg, 2014).

Malignant tumors are classified by the type of parenchyma's cells that are composed in four large groups: carcinomas, sarcomas, hematopoietic tumors and neuroectodermal tumors (Weinberg, 2014):

- Carcinomas are spawned by epithelial cells. They are the most common malignant tumor and the ones that more deaths cause (about 80%).
- Sarcomas are the first group or non-epithelial tumors, constituting about the 1% of diagnosed cancers. They arise from mesenchymal cells.
- Hematological tumors are non-epithelial malignancies that emerge from hematopoietic tissues and account for approximately 8% of human cancers.
- Neuroectodermal tumors derive from cells of both peripheral and central nervous system. They comprise around 1,3% of all diagnosed cancers.

Furthermore, inside those major groups the tumors can be caused by different cell types, so cancer can further be classified according to the specific cell of origin *(Table 1):*

-Carcinomas are mainly divided into squamous cells carcinomas and adenocarcinomas, both arising from different epithelial cells types. While squamous cells carcinomas derive from protective cells above the epithelium, adenocarcinomas are caused by missproliferation of cells specialized in secretion. They are also frequently classified by the organ where the missproliferation occurs because they show different histology, treatment and prognosis depending on the organ involved.

-Mesenchymal cells include many different types such as fibroblasts, adipocytes, osteoblasts or myocytes *(Table 1)*. There are also tumors that are formed from precursors of mesenchymal cells, such as hemangiomas.

- Generally speaking, hematological neoplasia could be globally differentiated in four main groups: Precursor lymphoid neoplasms, Mature lymphoid cell neoplasms,

acute myeloid leukemia and related precursor neoplasms, myeloproliferative neoplasms, these last ones arising from mature myeloid cells. Within each group, a variable number of different subtypes can be distinguished based on specific genetic abnormalities (e.g. B Lymphoblastic leukemia/lymphoma with t(9;22)(q34.1;q11.2); *BCR-ABL1*) and/or on the specific cell of origin (e.g. Splenic marginal zone lymphoma) and/or on the area from where they arise (e.g. primary diffuse large B cell lymphoma of the CNS), and even on the age (e.g. pediatric type follicular lymphoma) (Swerdlow et al., 2017).

-Neuroectodermal tumors are mainly divided into gliomas, that derive from glial cells, and meningiomas, arising from arachnoidal cells of meninges (Weinberg, 2014).

Table 1. Common types of cancer classified by cell of origin (Weinberg, 2014).

		Type of tumor	Presumed cell lineage
Epithelial tumors	**Carcinomas (Epithelial cells)**	Squamous cell carcinoma	Squamous cell
		Adenocarcinoma	Secretory cell
Non-epithelial tumors	**Sarcomas (mesenchymal cells)**	Osteosarcoma	Osteoblast
		Liposarcoma	Adipocyte
		Leiomyosarcoma	Smooth muscle cell
		Rhabdomyosarcoma	Striated/skeletal muscle
		Malignant fibrous histiocytoma	Adipocyte/muscle cell
		Fibrosarcoma	Fibroblast
		Angiosarcoma	Endothelial blood vessels' lining cells
		Chondrosarcoma	Chondrocyte
	Hematological tumours	Leukemia "liquid hematological tumor"	White blood cells (bone marrow)
		Lymphoma "solid hematological tumor"	B and T lymphocytes (lymphatic tissues)
		Multiple myeloma	Plasma cells
	Neuroectodermal tumors	Glioblastoma multiforme	Astrocyte (high progression)
		Astrocytoma	Astrocyte (low progression)
		Meningioma	Arachnoidal cells
		Schwannoma	Schwann cell
		Retinoblastoma	Cone cell in retina
		Neuroblastoma	Cells of peripheral nervous system
		Ependymoma	Cells lining ventricles
		Oligodendroglioma	Oligodendrocyte
		Medulloblastoma	Cerebellum's granular cells

2.1.2. Other neoplasia

Some types of tumors have special features and cannot be classified in these four groups:

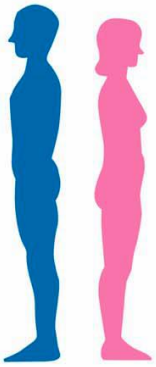
-Melanomas: they derive from melanocytes of skin or retina, that embryonically arise from neural crest, so they primary assemble to neuroectodermal cells but lack of connections with the nervous system.

-Small-cell lung carcinomas: tobacco makes lung's epithelial cells to change its characteristics making them more similar to neurosecretory cells, like those in the adrenal glands. This change of phenotype is known as transdifferentiation (Weinberg, 2014).

2.1.3. Cancer statistics

Cancer is one of the main causes of death worldwide, 1 in 6 deaths is caused by cancer (Ferlay et al., 2014). While breast and prostate cancer are the most incident *(Figure 1),* lung cancer is the most mortal and is the 4th cause of death globally (Vos et al., 2016).

Estimated New Cases

			Males	Females			
Prostate	161,360	19%			Breast	252,710	30%
Lung & bronchus	116,990	14%			Lung & bronchus	105,510	12%
Colon & rectum	71,420	9%			Colon & rectum	64,010	8%
Urinary bladder	60,490	7%			Uterine corpus	61,380	7%
Melanoma of the skin	52,170	6%			Thyroid	42,470	5%
Kidney & renal pelvis	40,610	5%			Melanoma of the skin	34,940	4%
Non-Hodgkin lymphoma	40,080	5%			Non-Hodgkin lymphoma	32,160	4%
Leukemia	36,290	4%			Leukemia	25,840	3%
Oral cavity & pharynx	35,720	4%			Pancreas	25,700	3%
Liver & intrahepatic bile duct	29,200	3%			Kidney & renal pelvis	23,380	3%
All Sites	**836,150**	**100%**			**All Sites**	**852,630**	**100%**

Estimated Deaths

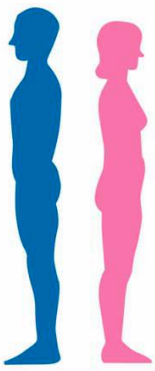
			Males	Females			
Lung & bronchus	84,590	27%			Lung & bronchus	71,280	25%
Colon & rectum	27,150	9%			Breast	40,610	14%
Prostate	26,730	8%			Colon & rectum	23,110	8%
Pancreas	22,300	7%			Pancreas	20,790	7%
Liver & intrahepatic bile duct	19,610	6%			Ovary	14,080	5%
Leukemia	14,300	4%			Uterine corpus	10,920	4%
Esophagus	12,720	4%			Leukemia	10,200	4%
Urinary bladder	12,240	4%			Liver & intrahepatic bile duct	9,310	3%
Non-Hodgkin lymphoma	11,450	4%			Non-Hodgkin lymphoma	8,690	3%
Brain & other nervous system	9,620	3%			Brain & other nervous system	7,080	3%
All Sites	**318,420**	**100%**			**All Sites**	**282,500**	**100%**

Figure 1. Cancer statistics in the USA 2017. The figure shows the incidence of cancer and the estimated deaths caused by it, depending on the organ of origin and the sex (Siegel et al., 2017).

2.2. THE HALLMARKS OF CANCER

Although cancer is a heterogeneous entity, there are at least ten features that are acquired progressively during the tumor development. These "Hallmarks of Cancer" were described by Hanahan and Weinberg in their seminal publication in 2011 and are proposed as causative of cell's malignancy and of the tumoral microenvironment in which malignant cells and normal cells, that form tumor stroma, interact between them in order to sustain tumorigenesis.

These traits are acquired due to two enabling characteristics: i) the genomic instability of cancer cells makes it easier for them to mutate, facilitating the acquisition of new hallmarks and the ii) inflammation in both premalignant and malignant lesions that attracts immune system cells that eventually facilitate neoplastic progression. These two characteristics are considered the primary hallmarks that allow the other features to take place (Hanahan & Weinberg, 2011).

In this section, we will summarize the main properties of the cancer cells proposed by Hanahan and Weinberg *(Figure 2)*.

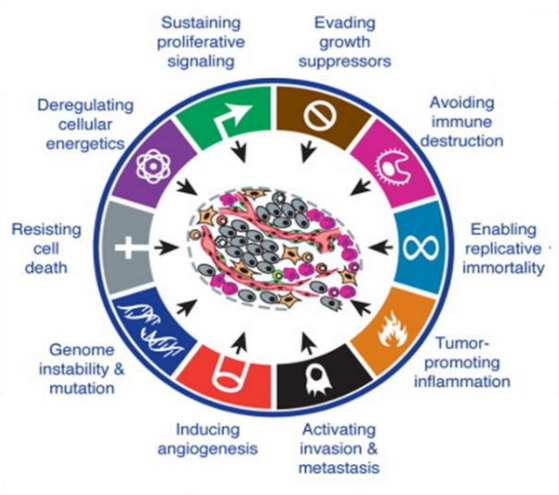


Figure 2.Hallmarks of cancer, proposed for Hanahan and Weinberg, which characterize cancer phenotype (Modified from Hanahan and Weinberg, 2011).

GENOME INSTABILITY AND MUTATION

To develop the other hallmarks and eventually become malignant, tumor cells must accumulate alterations in their genome that provide them an advantage over normal cells, that make them proliferate without control. Mutated genotypes are acquired by chance, and triggers successive clonal expansions that involve tumor progression. Non-mutational epigenetic changes such as DNA methylation and histone modifications that alter the gene expression are also involved in tumor development.

Genome is generally protected by the DNA repair system, triggered upon defects on DNA, eliminating most of the random mutations. Tumor progression is able to overcome the DNA repair system by different mechanisms including an increased mutation rate on the transforming cells more sensitive to the mutational agents and also due to the mutation on the genomic maintenance systems or "caretakers". The

caretaker genes act like tumor suppressor genes, and their alteration can involve fails in the detection on DNA damage and the lack of DNA repair.

TUMOUR-PROMOTING INFLAMATION

Every malignant tumor includes, in different grade, immune cells infiltrated in its environment. Inflammation promotes other hallmarks by providing different molecules to the tumor microenvironment, such as growth factors, survival factors, proangiogenic factors or enzymes that modify extracellular matrix. This inflammation is especially important in the first steps of tumor progression and promotes de development of earlier stage neoplasia into cancer.

SUSTAINING PROLIFERATIVE SIGNALLING

The most characteristic feature of tumors is its ability to proliferate without control, due to the alteration of the pathways that control the division cycle and/or external signals that under normal conditions maintain its order and function. Cancer cells influence the signals to perform they own regulation by different ways. Cells proliferation is regulated by growth factors that bind cell-surface receptors. Those receptors typically contain intracellular tyrosine kinase domains that emit signals that regulate progression through the cell cycle as well as cell growth, sometimes also cell survival and energy metabolism. Paracrine signals are also transmitted between the cells. Cancer cells can produce growth factors ligands that upregulate receptors on their surface, making them more numerous and more affine to the ligands, so they overstimulate themselves (autocrine signals). They also stimulate their normal stromal cells that respond producing growth factors that also stimulate the cancer cells.

Finally, some mutations detected in tumors result in growth factors receptors constitutively activated instead of triggered by ligand-receptor binding. That involves the activation of downstream pathways that promote proliferation. Proliferative signalling pathways such as MAP-kinase or PI3K and signal transducers (e.g. RAS or RAF) are frequently altered in cancer.

EVADING GROWTH SUPRESSORS

As well as cancer cells are able to promote proliferative signals, they can also avoid anti-proliferative pathways, which are usually determined by suppressor genes. The most prominent oncosuppressors are TP53 and RB proteins (see next section). The contact inhibition mechanism is also avoided in some tumors by not well-known mechanisms related with the inhibition of adhesion molecules in the cell membrane by oncogenes products and disorganization of epithelia caused by the lack of suppressor genes.

ACTIVATING INVASION AND METASTASIS

The ability of invade other tissues and metastasize is a main feature of malignant tumors. It is known that high grade carcinomas' cells alter their architecture and the way they relate with other cells and the extracellular matrix. The best-known alteration is the loss of adhesion molecules, such as E-cadherin, that leads to an increased capacity

of invasion. The complete process that results in the activation of invasion and metastasis is a complex cascade which is not known completely. The most studied field is the invasion and metastasis in carcinomas, mediated by the called epithelial-mesenchymal transition. It is controlled by transcriptional factors (e.g. Snail, Slug, Zeb1/2) that act in embryogenesis to provoke migration. These factors are overexpressed in some malignant tumors and stimulate invasion by causing the loss of adhesion in the tissues and changing its architecture from epithelial to mesenchymal.

After this local invasion, some tumors can start the metastatic colonization. Cells travel to distant tissues where they adapt to the tissue performing colonization. Some genes related with the appearance of metastasis have been identify, although the full regulation of metastasis is unknown.

ENABLING REPLICATIVE IMMORTALITY

Normal cells can only replicate a limited number of cycles, and when they are completed, the cell suffers senescence or dies. However, cancer cells are able to replicate in an unlimited way. This immortality seems to be related with the action of telomerase, that in normal adult cells has a low activity that results in the loss of the telomeres which protect the chromosomes ends. Telomeres diminish at each cycle, until they cannot protect the chromosome which ends fusion and form unstable chromosomes that imply cell unviability. The telomerase is a polymerase that adds telomere units at the end of chromosomes and it is abnormally expressed in cancer cells, avoiding senescence or apoptosis. It is thought to be a lately-acquired hallmark because it appears in carcinomas, not in premalignant lesions.

INDUCING ANGIOGENESIS

All tissues need a vascular network to receive nutrients and oxygen and throw away CO_2 and metabolism wastes. The more active is the tissue, the more productive is the metabolism and the more wastes are generated, so, in cancerous tissues a wide vascular network is early needed, even at the premalignant phase.

In normal conditions, angiogenesis is important in embryogenesis, when all the vasculature is formed. After that, it remains as a quiescent function, that is just reactivated transiently to heal wounds or in reproductive cycling. However, in tumors angiogenesis keeps activated and generates new vessels to sustain the increasing tumor. Two well-known regulators of angiogenesis are VEGF-A (activator) and Thrombospondin-1 (inhibitor).

RESISTING CELL DEATH

Apoptosis has been widely studied and established as a protection mechanism against neoplasia. It is induced by cell stress during tumorigenesis, especially upon high oncogene signaling and DNA damage. The main mechanism to avoid apoptosis by cancer cell is loss of TP53, that detects DNA damage and activates the cell death. In absence of this factor the aberrant DNA is not detected, so the apoptotic cascade is avoided. Other mechanisms include overexpression of antiapoptotic factors (eg Bcl2 or BclXL) or survival signals, or decreasing proapoptotic factors (eg Bax, Bim, PUMA).

DEREGULATING CELLULAR ENERGETICS

In aerobic conditions, cells obtain energy by degrading glucose to pyruvate in the cytosol and then pyruvate to CO_2 and H_2O in the mitochondria. The process of cellular respiration requires oxygen, so in anaerobic conditions the mitochondrial activity is very low. In this situation, the anaerobic glycolysis to lactate in the cytosol is favored. Cancer cells under aerobic conditions reprogram their metabolism and obtain energy just via glycolysis, a less effective via to generate ATP in comparison to cellular respiration. Cancer cells overexpress glucose transport, increasing the glucose in the cytoplasm. This increased glycolysis is even more active in anaerobic conditions, frequently present in tumors. Additional mechanism to reprogram metabolism are being identify in cancer cells.

AVOIDING IMMUNE DESTRUCTION

Immune system has a role in control neoplasia and micrometastases. About 80% of emerging neoplasia would be controlled by immune system. However, tumors slip past its vigilance to grow up. This hallmark is very relevant nowadays because the emerging knowledge about the mechanisms alterations that provoke tumoral evasion are an important target in the development of new treatments (Spranger & Gajewski, 2018).

2.3. ONCOGENES AND TUMOR SUPPRESSOR GENES

Cancer is caused by alterations in regulatory genes that control all the cell's activities. A human genome would suffer about 10^{10} mutations in an average life, but to cause neoplasia many regulatory pathways must fail. Mutations in different genes that affect cell proliferation accumulate during the tumor progression (Cooper & Hausman, 2014). Specific genes called oncogenes induce the cells to acquire the traits explained above giving rise to neoplasia. These oncogenes have its origin in proto-oncogenes that when mutated or overexpressed, make those pathways more active. On the other hand, the tumor suppressor genes avoid missproliferation, especially in situations in which the cell shows aberrant characteristics. Mutations on suppressor genes make them to be underexpressed or inactive, promoting neoplastic changes.

ONCOGENES

Proto-oncogenes are present in the normal human genome, and they encode proteins that assure normal cells proliferation and differentiation. When these genes mutate into oncogenes, their expression is altered, and cell proliferation is deregulated. The proto-oncogene transformation into an oncogene usually causes deregulation of the protein that they encode by different mechanisms *(Figure 3),* causing either overexpression or changes in the protein structure. Overexpression is usually caused by amplification or translocations, while structural changes are caused by translocations or point mutations.

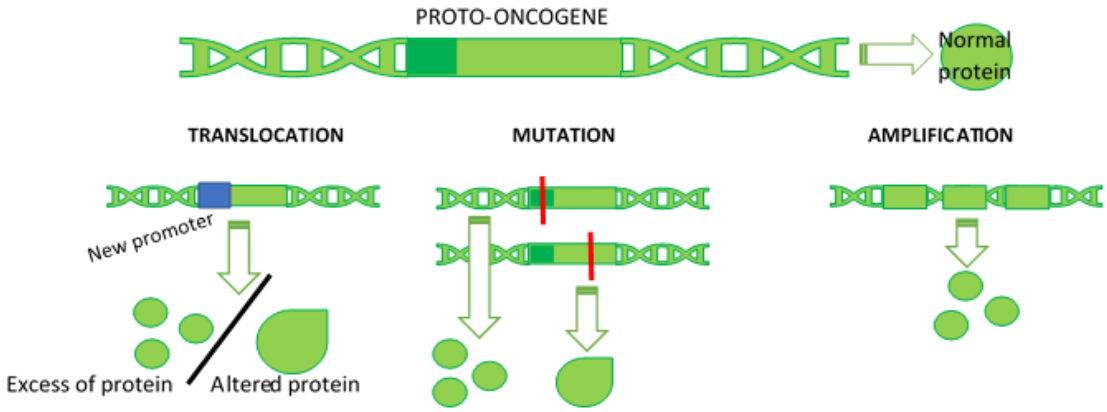


Figure 3. Types of mutations that turn proto-oncogene into oncogene. Translocations, amplifications and other mutations cause either overexpression of the protein or changes in its structure (McCance & Huether, 2018).

Most oncogenes are dominant, so an alteration in one of the copies of the gene is usually enough to provoke the activation, although more than one oncogene was believed to be necessary to provoke cancer (Cooper & Hausman, 2014). Some examples of well-known oncogenes are shown in the *table 2*.

Table 2. Examples of proto-oncogenes and its mutation into oncogenes. Proto-oncogenes are mutated into oncogenes by translocations (typical fusion gene expressed below), amplification or other mutations. One proto-oncogene can suffer more than one type of mutation that turns it into an oncogene (Cooper & Hausman, 2014; Weinberg, 2014).

Oncogene	Protein function	Types of cancer	Mutation
BCL-2 BCL2/Ig heavy chain t(14;18)	Apoptosis regulator	CLL, lymphomas, carcinomas, melanomas	Translocation, other
BCR-ABL1 t(9;22)	Serine/threonine-protein kinase	CML, ALL	Translocation
CCND1,2,3 *IGH/CCND1* t(11;14)	Cyclin (cell cycle control)	Carcinoma (ovary, testicle, breast), lymphoma	Translocation, amplification
EGFR	Receptor tyrosine kinase	Carcinomas (lung, breast), glioblastoma.	Other
erbB-2 (HER2)	Receptor tyrosine kinase	Breast cancer	Amplification, other
EWSR1 *EWS/FLI-1* t(11;22)	RNA binding protein	Ewing sarcoma	Translocation
hTERT	Telomerase reverse transcriptase	Carcinoma (pancreas, lung, ovary)	Other
KIT	Receptor tyrosine kinase protein	Sarcoma, GIST	Other
MYC	Transcription factor	Burkitt lymphoma. Carcinomas	Translocation, amplification
RAF	Serine/threonine-protein kinases (A-raf, B-raf, C-raf)	Carcinomas (breast)	Other
PML-RARA *t(15;17)*	Retinoic acid receptor alpha (transcription factor)	APL	Translocation
RAS (H-Ras, N-Ras, K-Ras)	GTPase small G protein	Carcinoma (90% pancreas)	Other
SRC	Tyrosine-kinase	Carcinomas, sarcomas, neuroblastomas	Other
TAL1,2 *SIL-TAL1*	Transcription factors	T-cell leukemia	Other

TUMOR SUPPRESSOR GENES

While oncogenes promote cell proliferation, tumor suppressor genes (TGS) counterbalance them, avoiding excessive growth. Therefore, their inactivation may induce neoplasia development.

The proteins encoding by TSGs have a role in suppression of cell proliferation, inhibiting growth, controlling differentiation signals or inhibiting the cell cycle upon DNA damage or metabolic alterations. As result of TSGs alteration, the proteins they encode will not be expressed or will do it in an inactive form.

Contrarily to oncogenes, TSGs are recessive, so both copies are usually required to be altered to provoke the disease. In addition, altered gene can recombine with the normal one during mitosis and cause that both copies are affected. That occurs more frequently than another mutation that affect the normal gene.

While oncogenes frequently appear in sporadic cancers, TSGs are also strongly related with many types of familiar cancers when the altered copies are inherited. In this case the individuals will be strongly predisposed to different types of cancer *(Table 3).*

TSGs can be inactivated by deletions or point mutations that make them nonfunctional *(Table 4)*. Other mechanism is the hypermethylation of their promoters that turns them inactive. Methyl groups can bond covalently to the cytosine base in the sequence CpG and repress the transcription of the gene by that inactivation of its promoter. Such methylation is transmitted to descendant cells, preserving it. Apparently, this epigenetic mechanism could be as important as somatic mutations to develop cancer (Weinberg, 2014). Finally, microRNAs that block translation, also have a role in TSG inactivation (Svoronos et al., 2016).

With some exceptions, analysis of TSGs in molecular diagnosis has lower value than oncogenes. However, they are important in the study of individuals whose family shows a history of cancers related with the same mutations, in order to prevent cancer by stricter screening or prophylactic treatments.

Table 3. Examples of TSGs (Weinberg, 2014) (Cancer gene census(https://cancer.sanger.ac.uk/cosmic)).

TSGs	Protein function	Familial cancer	Spontaneous cancer (e.g.)
APC	β-catenin degradation	Familial adenomatous polyposis coli	Colorectal, pancreatic, and stomach carcinomas; prostate carcinoma.
BRCA1	Proteins involved in DNA reparation	Hereditary breast–ovarian cancer syndrome	Breast cancer, epithelial ovarian cancer
BRCA2			Ovary, lung, breast
CDH1	cell–cell adhesion receptor	Hereditary diffuse gastric cancer	Bladder, breast, colon, gastric
CDKN2A	Cyclin-dependent-kinase inhibitor	Familial melanoma	Many types
MLH1	DNA mismatch repair enzymes	Lynch syndrome/HNPCC (colorectal, endometrial carcinoma, digestive adenoma, ovarian serous cystadenocarcinoma)	Carcinoma (stomach, esophageal, HNSCC, NSCLC, colorectal)
MSH2,6			ALL, carcinomas (renal, esophageal, lung…), soft tissue sarcoma
PTEN	PIP3 phosphatase	Cowden's disease, breast and gastrointestinal carcinomas	Glioblastoma; prostate, breast, and thyroid carcinomas
pVHL	Ubiquitylation of hypoxia-inducible factor	Von-Hipple-Lindau syndrome	Renal cell carcinoma
RB	Transcriptional repression; control of E2Fs	Retinoblastoma, osteosarcoma	Retinoblastoma; sarcomas; bladder, breast, esophageal, and lung carcinomas
TP53	Transcription factor	Li–Fraumeni syndrome	50% of human tumors
WT1	Transcription factor	Wilms tumor	Wilms tumor

Table 4. Types of mutation in TSGs (Cancer gene census (https://cancer.sanger.ac.uk/cosmic)).

TSGs	Deletion	Point mutation	Methylation	MicroRNas
APC				
BRCA1				
BRCA2				
CDH1				
CDKN2A				
MLH1				
MSH2,6				
PTEN				
pVHL				
RB				
TP53				
WT1				

2.4. CANCER GENOMICS

DNA sequencing has revolutionized our understanding of the cancer genome. Cancer genomics is defined as "The study of the totality of DNA sequence and gene expression differences between tumor cells and normal host cells. It aims to understand the genetic basis of tumor cell proliferation and the evolution of the cancer genome under mutation and selection by the body environment, the immune system and therapeutic interventions." (https://www.nature.com/subjects/cancer-genomics).

To develop this field of investigation, the presence of a human genome of reference was required. Therefore, when in 2003 the complete genome was presented as result of Human Genome Project, the comparison of cancer genomes with the reference evidenced gene variants that were possibly related to the development of malignancy. The epidemiological and functional study of those genes resulted in the establishment of many genes as related to cancer, both oncogenes and TGSs. Later on, the existence of faster and cheaper methods of sequencing led to the creation of international consortiums exclusively directed to study cancer genomics (Khotskaya et al., 2017).

2.4.1. Cancer genome projects

Nowadays, the most important International Consortiums in cancer genomics are: Cancer Genome Project (CGP), International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA).

The Cancer Genome Project (https://cancer.sanger.ac.uk/cosmic) studies the genetic changes acquired through life by sequencing human genomes (over 32000) and identify mutations that led to cancer and its impact. CGP generated a Cancer Gene Census that, for the moment, has registered 719 cancer-related genes. Of those, 577 have documented activity related to cancer, while the rest have an observed strong role in cancer but with less extensive evidence. CPG locate these mutations in the genome and study what was the mechanism that produce them.

International Cancer Genome Consortium (https://dcc.icgc.org/) coordinates over 88 different projects focused on different cancers and 16 countries. Nowadays the ICGC have analyzed the genome of 24077 cancer patients. Their goal is "to obtain a comprehensive description of genomic, transcriptomic and epigenomic changes in 50 different tumor types and/or subtypes which are of clinical and societal importance across the globe and make the data available to the entire research community as rapidly as possible, and with minimal restrictions, to accelerate research into the causes and control of cancer". This year, they have started the Accelerate Research in Genomic Oncology Project (ARGO) that aims to analyze the genome of 100000 patients until 2018 and include analysis of lifestyle, co-morbidity, diagnostics, toxicity response to therapy and survival.

The Cancer Genome Atlas (https://cancergenome.nih.gov/) described the 33 most prevalent tumor types based on the comparison between normal and cancer tissue of 11000 patients. The TCGA improved the knowledge about molecular basis of cancer,

changed cancer classifications by identifying different subtypes and stablished therapeutic targets. Their program ended this year with the publication of the Pan-Cancer Atlas (https://www.cell.com/pb-assets/consortium/pancanceratlas/pancani3/index.html?code=cell-site) which reclassifies tumor types and describes cell-of-origin patterns (Hoadley et al., 2018), oncogenic processes (Ding et al., 2018), and signaling pathways (Sanchez-Vega et al., 2018).

2.4.2. Cellular pathways altered in human cancer

Genomic studies have identified multitude of mutations in human tumors. Variations in genes do not always imply changes in cell behavior. While some mutations do not have impact on cellular pathways, meaningful mutations may alter them. The development of the international cancer consortiums and the improvement of DNA sequencing techniques led to the knowledge of meaningful mutations and the alterations that they cause in those cellular pathways *(Figure 4)*. The detection of these alterations allows the design of drugs that target these aberrant pathways, as the genes cause them are considered "actionable genes".

Large-scale studies have discovered multiple mutations in signal transduction molecules *(Figure 4, panel A)*. Abnormal activation of these molecules can cause uncontrolled cell proliferation. They are relevant therapeutic targets due to its crucial function in cell survival. Mutations in MAPK signaling cause its constitutive activation. Mutations in EGFR, ERBB2, BRAF, KIT, PDGFRα, FGFR2/3, DDR, JAK2 or MET alter this signaling pathway causing overproliferation. Mutations of RAS are the most frequent alteration of this mechanism (Sanchez-Vega et al., 2018). In addition, mutations in the parallel PI3K/AKT signaling pathway, like those in catalytic subunit of PI3K or of the downstream molecule AKT, have been also observed in multiple cancer types, due to the continuous activation of AKT that promotes cell survival.

Cancer cell's metabolism is altered due to the mutations in metabolism genes *(Figure 4, panel B)*. E.g. mutations in *IDH1* and *IDH2* disrupt Krebs cycle and deregulate downstream pathways. Mutations in the TSG *STK11* cause a decrease on AMPK signaling that stimulates glycolysis.

Genes that control chromatin modifications *(Figure 4, panel C)* are also altered and cause changes in patterns of DNA methylation or histone modifications. Those epigenetic changes are caused mainly by inactivation of methyltransferases.

Mutations inactivate genes that control genome integrity and act as cell cycle checkpoints *(Figure 4, panel D)*. Cancer cells avoid apoptosis and replicate uncontrolledly despite the DNA damage. The best-known TSG gene that is usually mutated in different cancers is *TP53*, that inhibits cell division if there are breaks in double-strand DNA. Other genes involved in this pathways that are altered in some types of cancer are *ATM*, that under normal conditions repairs DNA, and *CDKN2A*/*B*, that regulate cell cycle.

One of the latest alterations discovered are mutations in transcription, translation and processing genes *(Figure 4, panels D and E)*, which were not considered

in previous target sequencing studies but were evidenced with massive sequencing techniques.

Spliceosome mutations *(Figure 4, panel F)* are also present in cancer, deregulating mRNA processing, and promoting tumorigenesis (Chmielecki & Meyerson, 2014).

57% of the samples analyzed in TCGA program Pan-Cancer atlas contained potentially actionable mutations on these pathways. The most studied as therapeutic target are RKT/RAS, PI3K, cell cycle and p53 pathways (Sanchez-Vega et al., 2018).

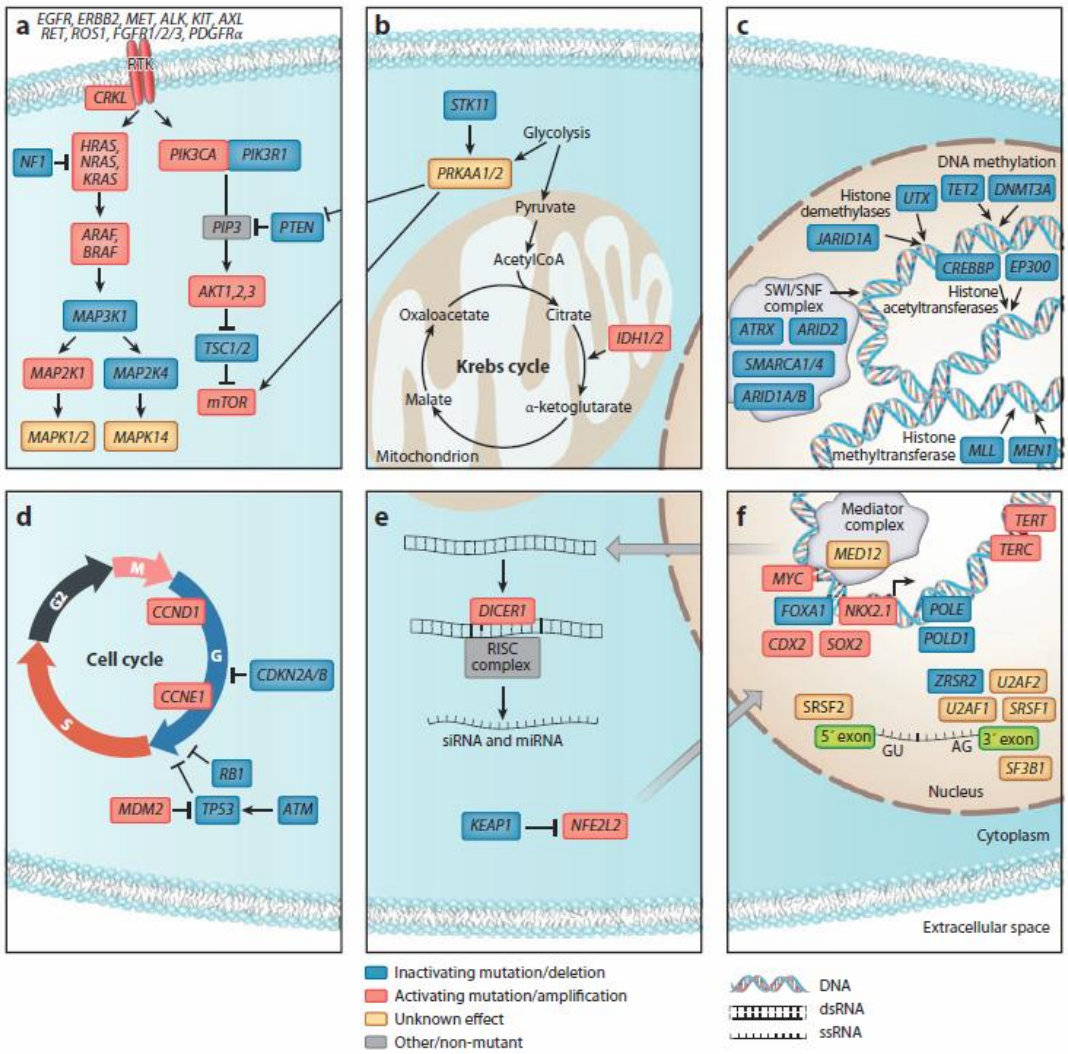


Figure 4. Cellular pathways altered in cancer, identified by genomic studies, include signal transduction (a), metabolism (b), chromatin remodeling (c), cell cycle and DNA repair (d), transcription, translation and processing (d and e) and splicing (f) (Chmielecki & Meyerson, 2014).

3. MOLECULAR DIAGNOSIS OF CANCER

Although most of statistics of prevalence and mortality of cancer are based on the primary organ where the tumor arose or as much on the type of tumor depending on the cells that missproliferate and characterize the tumor, now we know that apparently phenotypically similar tumors can be very different depending on the genes they express. In widely studied neoplasia was seen that tumor's molecular profile modifies the treatment and prognosis of cancer. E.g., chronic myeloid leukemia cells express the fusion gene *BCR-ABL1*, and the good prognostic associated nowadays with the disease is due to the existence of a specific treatment for the disease. Moreover, mutations on the tyrosine kinase inhibitor domain, blocking the inhibitor to bind will be associated to a worse prognosis. On the other hand, if the *ATM* or *TP53* genes are not expressed, the leukemia will be resistant to treatment and the prognosis would be worse (Hallek, 2015).

Therefore, apparently similar malignant tumors, that appear in the same organ and have the same histological classification, can be very different according to the mutations they present. They have different prognosis, treatment, response to it, so it is important to develop the complete cancer genes map and to have diagnostic tools to characterize each tumor, in order to design the best strategy to confront it. Due to the development of molecular biology techniques, the study of cancer has evolved from a model focused mostly on morphology to a new approach including molecular aspects.

Knowledge of genes that provoke neoplasia is useful not just for detection of predisposed individuals to develop a certain tumor but also for diagnose, stablish its prognosis, predict its response to treatment and monitor the evolution of the disease upon treatment. Some examples of molecular biomarkers used in clinical practice to guide diagnosis and therapeutic decisions are shown in the *Table 5.* The complete list of FDA approved targeted therapies for tumors that have an associated biomarker can be found in the *Annexe 1* (Kurnit et al., 2018). The importance of detecting these genes has increase the demand of molecular diagnosis techniques.

This section will address first the molecular diagnosis of hematological tumors, explaining the main techniques used and the most significant biomarkers analyzed nowadays. We will focus on the chronic myeloid leukemia (CML) as representative example of how molecular approaches are used for the diagnosis and treatment monitoring of hematological neoplasias. Then, the biomarkers and methods used for solid tumors diagnosis will be summarized. As a paradigm, breast cancer molecular diagnosis and treatment will be analyzed deeper. Finally, new tools such as next generation sequencing techniques and the use of liquid biopsies in the field of molecular diagnosis will be reviewed.

Table 5. Examples of molecular biomarker used in clinical practice to guide diagnosis and therapeutic decisions (Gonzalez de Castro et al., 2013).

Diagnostic		
Acute leukemias	*PML-RARA*	WHO 2008 classification of leukemias
	BCR-ABL1	
	CBFB-MYH11	
	ETV6-RUNX1	
	RUNX1-RUNX1T1	
	MLL-rearranged	
	TCF3-PBX1	
	RBM15-MKL1	
MPD	*JAK2*	Mutations confirm diagnosis of clonal MPD
Sarcomas	*SS18-SSX1/SSX2*	Synovial sarcoma
	PAX3/ PAX7-FOXO1A	Alveolar rhabdomyosarcoma
	EWSR1-FLI1	Ewing's sarcoma
	EWSR1-ERG	
	EWSR1-NR4A3	Extraskeletal myxoid chondrosarcoma
	TAF15-NR4A3	
	EWSR1-ATF1	Clear cell sarcoma (and angiomatoid fibrous histiocytoma)
	EWSR1-CREB1	
	ASPSCR1-TFE3	Alveolar soft-part sarcoma (and renal cell carcinoma)
	FUS-DDIT3	Myxoid liposarcoma
	FUS-CREB3L2	Low-grade fibromyxoid sarcoma
	JAZF1-SUZ12	Endometrial stromal sarcoma
	ETV6-NTRK3	Congenital fibrosarcoma (and secretory breast carcinoma)
Predictive		
NSCLC	*EGFR*	Mutations predict response to TKI
	ALK	Rearrangements predict response to ALK-inhibitors
GIST	*KIT* and *PDGFRA*	Mutations predict response to c-KIT/PDGFRA inhibitors
mCRC	*KRAS*	Mutations predict lack of response to anti-EGFR antibodies
Melanoma	*BRAF*	Mutations predict response to specific BRAF inhibitors
Breast cancer	*HER2*	Amplifications predict response to anti-HER2 antibodies
Prognostic		
CLL	*TP53*	Mutations are indicative of poor outcome
	IGHV	Lack of mutations is indicative of poor outcome
AML	*FLT3-ITD*	Mutations are indicative of poor outcome
mCRC	*BRAF*	Mutations are indicative of poor outcome
Breast cancer	*OncotypeDx*	Risk stratification (21-gene expression signature)
	Mammaprint	Risk stratification (70-gene expression signature)
	IHC4	Risk stratification (4-protein IHC expression)
Disease monitoring		
CML	*BCR-ABL1*	Minimal residual disease detection
APML	*PML-RARA*	Minimal residual disease detection
ALL	*IGHV-TCR* rearrangements	Minimal residual disease detection

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukaemia; APML, acute promyelocytic leukemia; CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; EGFR, epidermal growth factor receptor; GIST, gastro-intestinal stromal tumors; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; mCRC, metastatic colorectal cancer; MPD, myeloproliferative diseases; NSCLC, non-small-cell lung cancer; TKI, tyrosine kinase inhibitor; WHO, World Health Organization.

3.1. HEMATOLOGICAL TUMORS

The molecular diagnosis of cancer has always been more developed in hematological tumors than in solid tumors, in part due to the easier extraction of samples and the possibility to easily expand them in vitro, the clinical implication at early stages of the disease, and the presence of recurrent translocations. Some hematological tumors subtypes, such as many of the precursors neoplasms, are defined by specific genetic alterations. In these subtypes not only is possible to effectively detect and monitor them with molecular techniques but also, many of them have effective target treatment.

The molecular characterization of these diseases, required for the establishment of the prognosis and ultimately defining the treatment, is based on the combination of different genetic techniques that, at the moment, are complementary such as karyotyping, FISH, PCR, and more recently NGS. Some genetic alterations usually imply good prognosis and a longer survival, while others involve bad prognosis, due to either higher aggressiveness or resistance to treatment.

3.1.1. Biomarkers analyzed in hematological tumors

The most important genes altered in hematological neoplasias that are used as biomarkers for the molecular diagnosis of these diseases are shown in the *Table 6*.

Table 6. Markers in molecular diagnosis of the most well defined hematological neoplasias (Batlle-López et al., 2016).

TYPE OF CANCER	GENE	ALTERATION	TECHNIQUE
CML	BCR-ABL1	t(9;22)	FISH, qPCR, Karyotype
AML			
AML with t(8;21)(q22;q22.1)	RUNX1-RUNX1T1	t(8;21)(q22;q22.1)	RT-PCR-FISH-Karyotype
AML with t(16;16)(p13.1;q22) or inv(16)(p13.1q22)	CBFB-MYH11	Inv(16) or t(16;16)	
Acute promyelocytic leukemia	PML-RARA	t(15;17)(q24;q21)	
AML with t(9;11)(p21.3;q23.3)	MLLT3-KMT2A	t(9;11) (p21.3;q23.3)	
AML with t(6;9)(p23;q34.1)	DEK-NUP214	t(6;9) (p23;q34.1)	
AML with inv(3)(q21.3;q34.1)	RPN1-MECOM	Inv(3) or t(3;3)	
AML with t(1;22)(p13.3;q13.1)	RBM15-MKL1	t(1;22) (p13.3;q13.1)	
AML with BCR-ABL1	BCR-ABL1	t(9;22)(q34;q22)	
AML with gene mutations	NPM1, CEBPA, RUNX1, FLT3	NPM1, CEBPA, RUNX1 mutations	
Mixed phenotype AML	BCR-ABL1	t(9;22)	
	KMT2A (MLL1)	t(11q23;v)	

Table 7 (cont.). Markers in molecular diagnosis of the most well defined hematological neoplasias (Batlle-López et al., 2016).

MDS	SF3B1, TET2,ASXL1,SRSF2, DNMT3A, RUNX1, TP53,U2AF1	Mutations	Sequencing (mostly NGS panels)
MPN Ph-	JAK2	Mutations	PCR
	CALR		
	MPL		
B ALL	BCR-ABL1	t(9;22)(q34.1;q11.2)	PCR-FISH
	KMT2A	t(v;11)(v;q23.3)	FISH-PCR
	EVT6-RUNX1	t(12;21)(p13.2;q22.1)	FISH-PCR
	IGH-IL3	t(5;14)(q31.1;q32.1)	FISH-PCR
	TCF3-PBX1	t(1;19)(q23;p13.3)	FISH-PCR
		iAMP21	FISH
	BCR-ABL1 like mutations: CRFL2,PDGFRB,etc	Mutations, translocations, deletions,	FISH-PCR-Arrays
T ALL	TCRA/D	Translocations with different genes t(14;v)	FISH
	TCRB	t(7;v)	FISH
	SIL-TAL	Deletion in 1p causes the fusion gene SIL-TAL	FISH-PCR
	BCR-ABL1	t(9;22)(q34.1;q11.2)	FISH-PCR
	PICALM-MLLT10	t(10;11)(p13;q14)	FISH-PCR
	ETV6-JAK2	t(9;12)(p24;p13)	FISH-PCR
	CDKN2A/B	Del(9p)	FISH
	Variable	Complex Karyotype	Karyotype
Follicular lymphoma	BCL2	t(14;18)(q32;q21)	FISH
Mantle cell lymphoma	CCND1	t(11;14)(q13;q32)	FISH, conventional cytogenetics
High Grade lymphoma with BCL2,MYC and or BCL6 rearrangements	MYC with BCL2 and or BCL6		FISH
Burkitt-Like Lymphoma with 11q aberration		Ain and losses on 11q	Array
MALT	BIRC3-MALT1	t(11;18)(q21;q21)	FISH
	IGH-MALT	t(14;18)(q32;q21)	FISH
	BCL10-IGH	t(1;14)(p22;q32)	FISH
Hairy cell leukemia	BRAF-V600E	Mutation	PCR, Sequencing
Wäldenstrom	MYD88	Mutation	PCR, Sequencing

3.1.2. Techniques for molecular diagnosis in oncohematology

MORPHOLOGY:

The first approach to the diagnosis of a hematological neoplasia in the laboratory is the analysis under the microscope of peripheral blood and bone marrow samples by cytology and cytochemistry (Bain et al., 2011).

IMMUNOPHENOTYPE:

Flow cytometry is used to analyze the markers that define a specific cell type, and for this, it requires cells to be on a suspension. Nodes and mases can be firstly mechanically disrupted to obtain a suspension of cells that will be then incubated with a combination of specific monoclonal antibodies depending on the tumor type suspected and those will induce a signal on the cytometer. Immunophenotype analysis is not only useful to both stablish the cell lineage and maturation stage of the tumoral cells but it will also provide information of the expression of abnormal markers (Davids, 2018). Together with the morphology will provide a first orientation of which molecular markers should be investigated. These techniques will however not be able to provide information of the specific genetic alterations involved.

CYTOGENETICS:

-Conventional cytogenetics studies the chromosomal alterations evaluable under a microscope on a karyotype. It consists in the analysis of the chromosomes in metaphase, previously stained, usually with Trypsin and Giemsa, which generates a characteristic band pattern.

Usually 20 metaphases are evaluated in order to identify translocations, inversions, deletions and many other abnormalities. This technique provides a global information of the genome although the resolution is low (~5Mb). One of the most important limitations of this technique is that requires fresh viable cells in order to obtain metaphases (Schrijver, 2017).

-Fluorescent In Situ Hybridization (FISH): in this technique, chromosomes in both metaphase or interphase, are hybridized with specific fluorescent probes. FISH can be used to confirm specific diagnostics based on the presence of genetic characteristic aberrations and will also provide prognosis information. It can sometimes be used to monitoring the disease although it has a low sensitivity (it only evaluates 200-1000 nuclei). The resolution is also low (~50Kb) so it will not be appropriate to detect point mutations. Compared to karyotype, it will gives less general information because the probes are specific to particular alterations (Zhang, 2017).

MOLECULAR BIOLOGY: different molecular techniques are used for the analysis of cancer genes altered by translocation, mutation, overexpression, etc. *(See Table 6)*. The most important are:

-PCR (Polymerase chain reaction): with this technique a specific sequence in the DNA can be amplified to study different alterations that typically appears in hematological malignancies such as gene amplification, point mutations or chromosomal translocations (Raby, 2016). There are different types of PCRs.

-RT-PCR (Reverse-transcriptase polymerase chain reaction): messenger RNA is converted into complementary DNA (cDNA) by a reverse transcriptase. The double-stranded cDNA becomes the template for a subsequent PCR reaction allowing the analysis of gene overexpression or chromosomal translocations (Stock & Estrov, 2017).

-qPCR (quantitative real-time polymerase chain reaction) is the most useful molecular technique in hematological for disease monitoring since is very sensitive. This technique has a high sensitivity for the quantification of a specific sequence by the addition of fluorescent probes that hybridize between the primes and, for this reason, is the best method for the monitoring of residual disease in patients after treatment (Raby, 2016).

-dPCR (digital polymerase chain reaction) is an adaptation of conventional PCR that does not need multiple PCR cycles to make a quantification. The sample is separated into packets that contain a small number of nucleic acid sequences from the patient. PCR is applied on each packet. The presence or absence of the target nucleic acid is expressed in binary language. The starting amount of target nucleic acid is estimated by the device without the reliance on multiple amplification cycles. The dPCR achieves an absolute measure of target sequence, while traditional PCR compares the amount of target DNA on the amplified sample with standard thresholds (Raby, 2016).

-DNA sequencing: Sanger technique sequences large DNA fragments by its polymerization with template DNA. In the first Sanger method, nucleotides were marked with radioactive molecules. Later, they were substituted by fluorescent probes. It is used when the test targets a specific gene. It cannot analyze large portions of the genome due to its cost and long timeframe, but its precision makes it valuable as a confirmation technique of gene variations detected by other techniques (Hulick, 2018). Next generation sequencing techniques are being implemented for the analysis of mutations in an increasing number of hematological malignancies. These techniques will be further explained in the section 3.3 “New tools for molecular diagnosis”.

3.1.3. Molecular diagnosis of chronic myeloid leukemia

Chronic myeloid leukemia (CML) is a myeloproliferative neoplasm originated in a pluripotent hematopoietic stem cell that can give rise to all myeloid lineage cells (erythrocytes, granulocytes, monocytes or megakaryocytes). Unlike the other myeloid neoplasm, stem cells that origin CML can also differentiate into lymphoid cells (Arber et al., 2016).

CML is characterized by the presence of a *BCR-ABL1* fusion gene, present in 90-95% of the cases *(Figure 5)*. This gene is formed by a reciprocal translocation between chromosomes 9 and 22, originating the t(9;22)(q34;q11) (Arber et al., 2016).

Firstly, on the early 70s Philadelphia chromosome, which is the 22 abnormal chromosome, was discovered, and it was proposed as a cause of pathogenesis. Later the t(9;22) translocation originating the Ph chromosome was detected. Finally, the fusion oncogene *BCR-ABL1* was established as causative of malignancy and was proposed as a possible therapeutic target (Druker & Lydon 2000) (Bennour et al., 2016). Based on this knowledge, Imatinib, a tyrosine kinase inhibitor, was designed as a selective drug for BCR-ABL1 oncoprotein (Deininger & Druker, 2003).

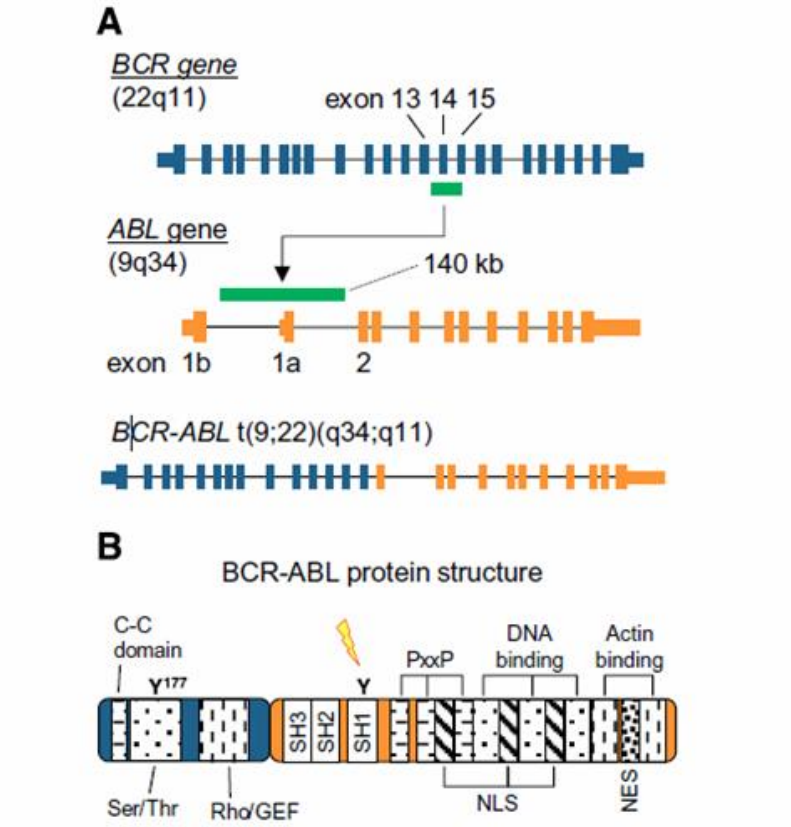


Figure 5. BCR-ABL gene and protein structure of BCR-ABL. a) BCR-ABL is generated by the fusion of intron 13 or 14 of BCR with a region between exons 1b and 2 of ABL1. b) The protein expressed contains the dimerization of coleidcoil, Ser/Thr kinase and Rho/GEF domains from BCR and the SH, proline rich, nuclear localization signal, DNA-binding, nuclear export signal and Actin binding domains from ABL (Chereda & Melo, 2015).

Diagnosis and monitoring of CML involves the detection of the *BCR-ABL1* fusion gene. The BCR-ABL1 protein has a constitutively activated tyrosine kinase activity, triggering downstream pathways that involve overproliferation, producing the chronic phase of CML *(Figure 6)*. This chronic phase usually lasts 3-4 years without treatment, that can end in a blast crisis if it is not treated, becoming highly mortal. In some patients, there is an accelerated phase that rapidly leads to blast crisis (Chereda & Melo, 2015).

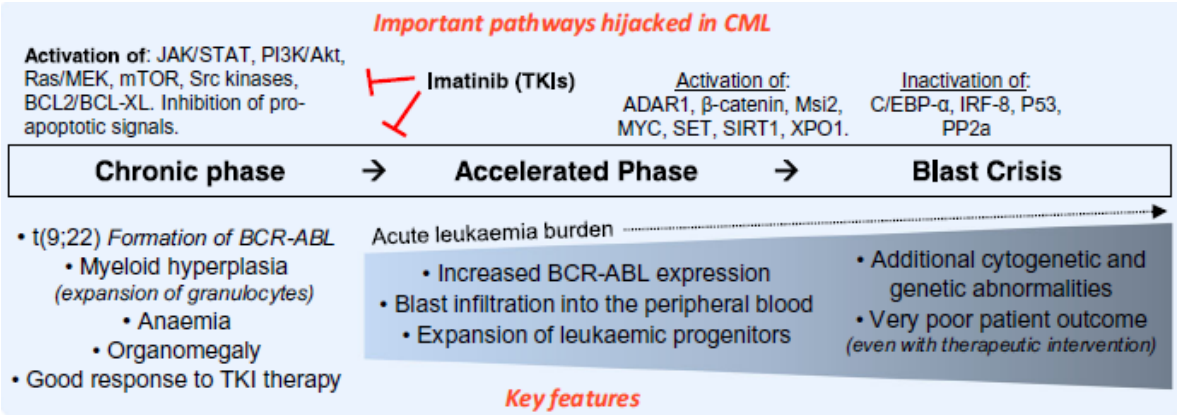


Figure 6. Important events in CML (Chereda & Melo, 2015)

a) Diagnosis

The diagnosis of CML is frequently established in the chronic phase of the disease. Frequent symptoms are fatigue and abdominal discomfort, and in the physical examination, the splenomegaly is quite prevalent. Asymptomatic patients in which CML is diagnosed as a casual finding are also common (Chereda & Melo, 2015).

When a diagnostic suspicion exists, the morphological analysis of the cells is useful as an initial approach (Sangle, 2011):

-Cytology: an increase in WBC with neutrophilia, basophilia, eosinophilia and mild monocytosis in peripheral blood is observed. In half of the patients it exists thrombocytosis. Bone marrow shows 100% cellularity with an increment in granulocytes precursors and in the mature forms. Erythrocytes and its precursors are usually unaltered *(Figure 7)*.

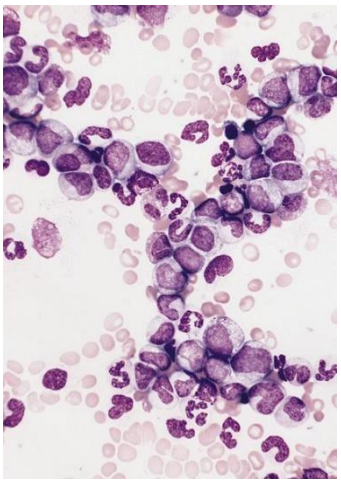


Figure 7. Peripheral blood sample of a 44 years old woman with a WBC count of 425k. Chronic phase of CML http://www.pathologyoutlines.com/topic/myeloproliferativecml.html

Even if these techniques are used as an initial orientation to CML diagnosis, its confirmation is based in the detection of the *BCR-ABL1* gene, which is considered constant in all CML cases (Chereda & Melo, 2015).

Karyotype analysis: it is performed in order to detect the Philadelphia chromosome associated with the fusion gene *(Figure 8)*.

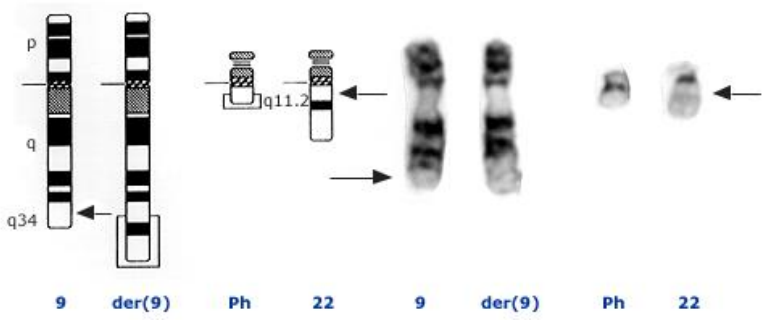


Figure 8. Philadelphia chromosome. t(9;22)(q34;q11.2). G-band ideograms (left) and partial karyotype (right). The translocations result in a longer chromosome 9 (der9) and shorter chromosome 22 (Ph chromosome). The breakpoints are indicated with arrows (Van Etten, 2015).

FISH uses DNA probes labelled with fluorophores. *BCR* is marked with a color probe and *ABL1* with other different, when they are located in the same place it suggests the existence of the fusion gene *(Figure 9)*. With this technique, *BCR-ABL1* fusion gene can be detected in either metaphase or interphase cells (Van Etten, 2015).

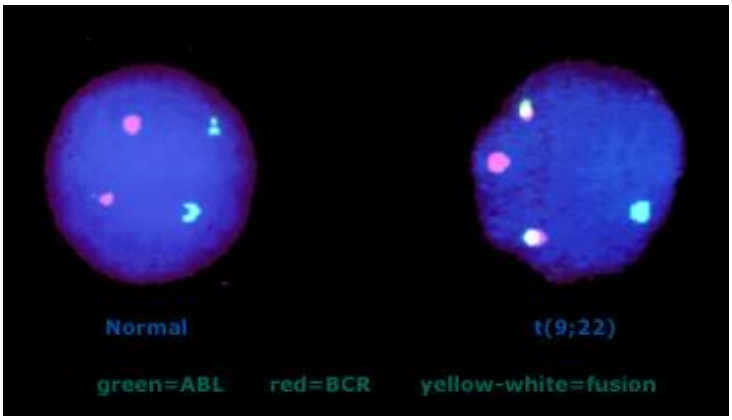


Figure 9. FISH: the image in the left shows the separated ABL1 and BCR signals (ABL: green; BCR: red) in a normal sample. The FISH on the right shows how the fluorescent signal turns yellow when both genes are fused (Van Etten, 2015).

RT-PCR (Reverse-transcriptase PCR): This technique uses primers to amplify a cDNA fragment generated by reverse transcription from *BCR-ABL1* mRNA. It is possible to quantify the *BCR-ABL1* copy number by RT-qPCR with high sensitivity. Using different primers, alternative breakpoints that caused fusion gene can be detected.

RT-PCR detects one cell positive for chromosome Philadelphia in 10^5-10^6 normal cells. Due to this sensitivity, its low-cost and rapidity, it is chosen as diagnostic test of Philadelphia-positive leukemia (Van Etten, 2015), specially to monitor response to treatment measuring minimal residual disease.

b) Treatment

ABL protein is a tyrosine kinase that, in normal conditions, is activated only upon specific stimulus in order to maintain cell's function. When the *BCR-ABL1* fusion gene is generated, the protein that it encodes (p210) also has tyrosine kinase activity but it is constitutively activated by the replacement of the region that inactivates tyrosine kinase by a BCR protein fragment (*Figure5)*. This leads to the activation of intracellular signaling pathways that induce overproliferation and cause the disease (Van Etten, 2017).

As CML is caused by a deregulated protein tyrosine kinase, kinase inhibitors (TKIs) were studied as possible therapeutic weapons to control the disease. Imatinib was designed as a specific inhibitor of the tyrosine kinase enzyme *(Figure 10)* and due to its effectivity, it was stablished as first line treatment of CML. Other drugs such as Nilotinib, Bosutinib, Ponatinib, Dasatinib and Omacetaxine (new generation TKIs) inhibits other pathways as well as BCR-ABL1 and are used as treatment in cases of resistance to Imatinib (Bennour et al., 2016; Negrin & Schiffer, 2017; Van Etten, 2017).

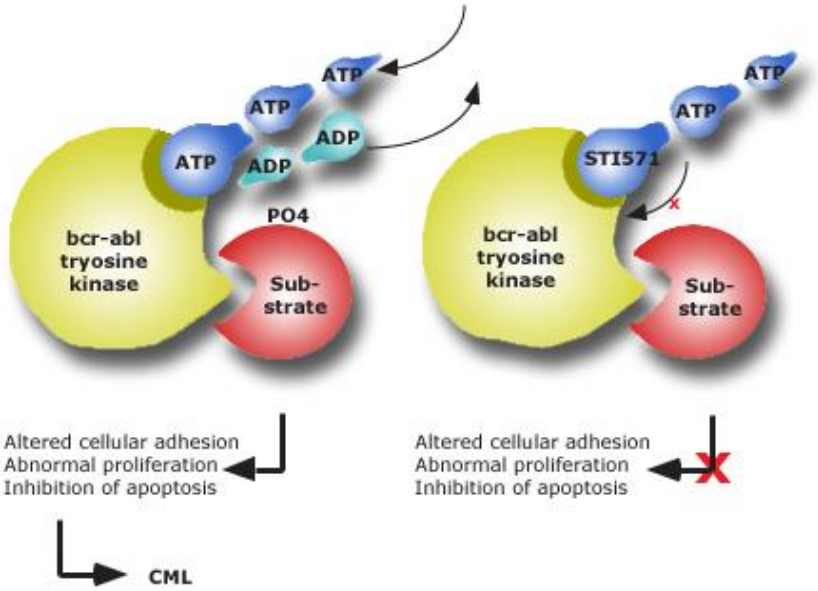


Figure 10. Mechanism of action of Imatinib (STI571). The drug blocks the ATP binding to the tyrosine kinase, inhibiting its activity and avoiding the activation of the pathway (Mauro & Druker, 2001).

c) Monitoring

It is important to monitor the expression of the genetic alteration during the treatment of the hematological neoplasia, to avoid relapses. They are caused by the so called "minimal residual disease" i.e. the presence of residual leukemic cells during or after the treatment, that are undetectable by the conventional morphological techniques (Stock & Estrov, 2017). The monitoring of the response in CML is based in blood counts (hematologic response), presence of Ph (cytogenetic response) and expression of BCR-ABL1 by qPCR (molecular response) *(Table 7)*.

Table 7. Definitions of hematologic, cytogenetic, and molecular response (MR) in chronic myeloid leukemia (Negrin & Schiffer, 2018)

RESPONSE BY TYPE	DEFINITIONS
Hematologic	
Complete	WBC<10x10^9/L
	Basophils<5%
	No myelocytes, promyelocytes, myeloblasts in the differential
	Platelet count<450 x10^9/L
	Spleen nonpalpable
Cytogenetic	
Major	Complete: No Ph+ metaphases or <1 percent BCR-ABL1-positive nulcie of at least 200 nuclei on fish
	Partial: 1 to 35 percent Ph+ metaphases
Minor	36 to 65 percent Ph+ metaphases
Minimal	66 to 95 percent Ph+ metaphases
None	>95% percent Ph+ metaphases
Molecular	
MR[5]	Detectable disease with ratio of BCR-ABL1 to ABL1 (or other housekeeping genes) ≤0.0001 or Undetectable disease in cDNA with ≥100,000 ABL1 transcripts
MR[4,5]	Detectable disease with ratio of BCR-ABL1 to ABL1 (or other housekeeping genes) ≤0.0032 percent (≥4.4 log reduction) on the international scale (IS) or Undetectable disease in cDNA with ≥32,000 ABL1 transcripts
MR[4]	Detectable disease with ratio of BCR-ABL1 to ABL1 ≤0.01 percent (≥4 log reduction) on the IS or Undetectable disease in cDNA with ≥32,000 ABL1 transcripts
MR[3]	Detectable disease with ratio of BCR-ABL1 to ABL1 (or other housekeeping genes) ≤0.1 percent (≥3 log reduction) on the IS Undetectable disease in cDNA with ≥10,000 ABL1 transcripts

Nowadays, most patients achieve cytogenetic response when they are treated with TKIs. RT-qPCR is sensitive enough to detect the minimal residual disease and to establish the probability of relapses on each patient, so it is the main technique in CML

monitoring, and determinates the most convenient therapeutic measures, as it is the basis of guidelines (Deininger, 2015).

Even if the results of BCR-ABL1 expression by PCR are compared with international standards nor with patient's previous expression ratio, an initial qPCR is performed when the patient is diagnosed of CML. Then, the patient is monitored firstly every 3 months and the results stablish the treatment *(Figure 11)*.

The failure in the treatment can be primary, if the patient never responds to the TKIs, or secondary, when it has responded but has not achieved the desired levels that assure non-relapsing. When the failure is secondary, a new cytogenetic evaluation is performed, and *BCR-ABL1* kinase mutations are investigated, since they cause 40-50% of TKIs therapy resistance. A detection of a specific mutation in *BCR-ABL1* implies a change in treatment, that could be a dose increase or a change in the TKI administered. Eventually, once the patient is on Complete molecular remission, treatment can be discontinued. As half of patients have relapses upon treatment discontinuation, narrow monitoring by qPCR is mandatory. If a relapse is detected TKIs will be then reintroduced (Negrin & Schiffer, 2018).

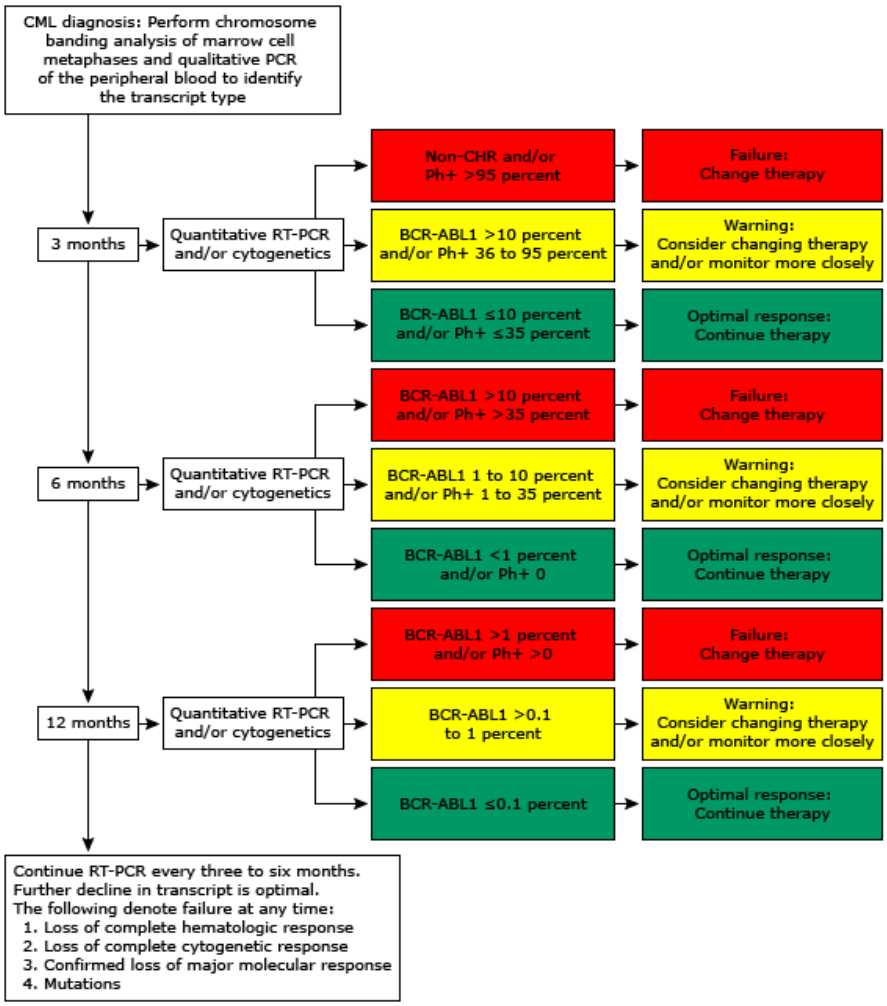


Figure 11. Algorithm for monitoring disease response to therapy in CML (Negrin & Schiffer, 2018).

3.2. SOLID TUMORS

Molecular techniques have therefore changed the approach to classify and understand hematological neoplasms, from a morphological classification to a more complex system which integrates different phenotypic, clinic and genetic alterations, that provide a deeper knowledge of the malignancy. In contrast, the molecular study of carcinomas had been less applied. However, molecular approaches have been recently boosted by the new techniques that have contributed to stablish solid tumor's molecular features (Pareja et al., 2018).

3.2.1. Biomarkers analyzed in solid tumors

Despite the genetic complexity of carcinomas, the molecular techniques have contributed, as in hematological neoplasia, to the development of targeted therapies based on the mutations present in the tumor. Some representative examples are shown in *Table 8*. A complete list of FDA approved targeted therapies for tumors that have an associated biomarker can be found in the *Annexe 1* (Kurnit et al., 2018). In addition, a list of actionable genes and their alterations in common solid tumors are shown in the *Annexe 2* (Kurnit et al., 2018).

Table 8. Some genes routinely analyzed for mutations in solid tumors in diagnostic molecular pathology (Müllauer, 2017).

Tumor	Altered genes	Therapy (selection)
Lung adenocarcinoma	EGFR	Gefitinib, erlotinib, afatiniv, osimertinib
	ALK	Crizotinib, ceritinib
	ROS1	Crizotinib
	MET exon 14	Crizotinib
Gastrointestinal stroma tumor (GIST)	KIT	Imatinib, sunitinib
	PDGFRA	Imatinib, sunitinib
Colorectal carcinoma	KRAS, NRAS, BRAF, Microsatellite instability	Cetuximab, panitumumab, immune checkpoint inhibitor
Malignant melanoma	BRAF	Vemurafenib, dabrafenib, trametinib, cobimetinib
	KIT	Imatinib, sunitinib, dasatinib
Breast carcinoma	HER2	Trastuzumab, pertuzumab
Ovarian Carcinoma; triple negative breast cancer	BRCA1/2	Olaparib
Medullary thyroid cancer	RET	Vandetanib

3.2.2. Techniques for molecular diagnosis in solid tumors

CYTOLOGY, CYTOCHEMISTRY, HISTOLOGY AND IMMUNOHISTOCHEMISTRY

- Histological diagnosis by analysis under microscope of tissue's slides has been the most used technique in samples obtained by biopsy, and to study the affection of lymph nodes to determine tumor's extension yet, even if now these techniques are frequently supported by molecular studies in a growing number of malignancies (Müllauer, 2017).

-Cytological techniques are widely used in the diagnostic of many carcinomas, such as cervix, lung, bladder or thyroid, in which samples are obtained as isolated cells from smear, spontaneous secretions or fine needle aspiration (Gray & Kocjan, 2010).

CYTOGENETICS:

-Conventional cytogenetics (karyotype analysis) is not so frequently used in solid tumors, being the main cause for this the difficulty to obtain appropriated metaphase chromosomes. Furthermore, in solid tumors, single characteristic mutations and recurrent chromosomal translocations are usually rare, so karyotype has relatively low applications for most solid tumors (Grade et al., 2015).

-FISH is applied for the detection of some gene abnormalities such as *ALK* on lung cancer and *CISH* in breast cancer tissue samples for *HER-2/neu* status.

MOLECULAR BIOLOGY:

-Comparative Genomic Hybridization (CGH) allows to detect gain and loses on the genome. Two DNA samples are extracted and isolated, one from the patient's tumor and other for a non-affected tissue. Both are marked with different fluorochromes, mixed together and hybridized to normal lymphocyte metaphasic chromosomes. Depending on the fluorochrome showing the higher intensity, gain or loss of genetic material is detected. As it still needed metaphasic chromosomes and that limited the resolution to the size of a band, metaphasic chromosomes were substituted for short DNA genomes that represent the entire genome and are used in different slides. This technique is the array-based CGH (aCGH) which provides higher resolution and generates a map of fluorescence that shows gains and losses of DNA of all the genome *(Figure 12)*. This allows the localization of the alterations, even shows which allele is preferably affected (Grade et al., 2015).

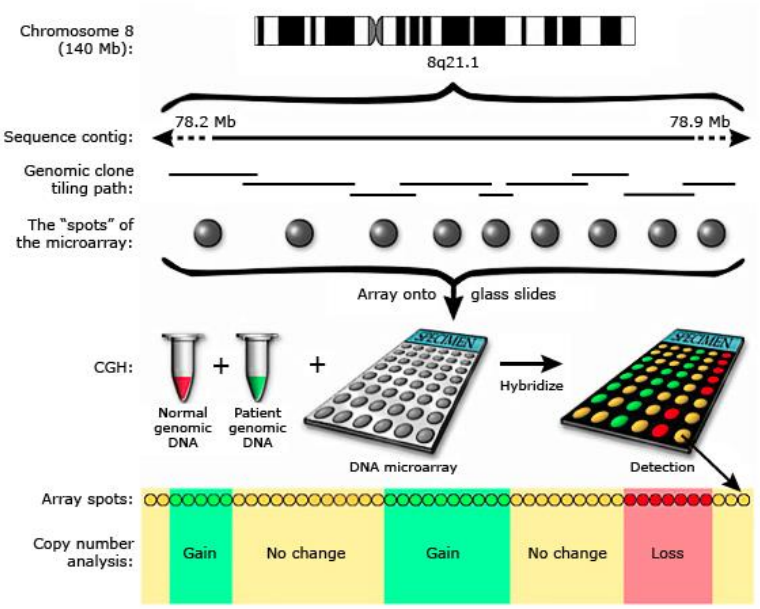
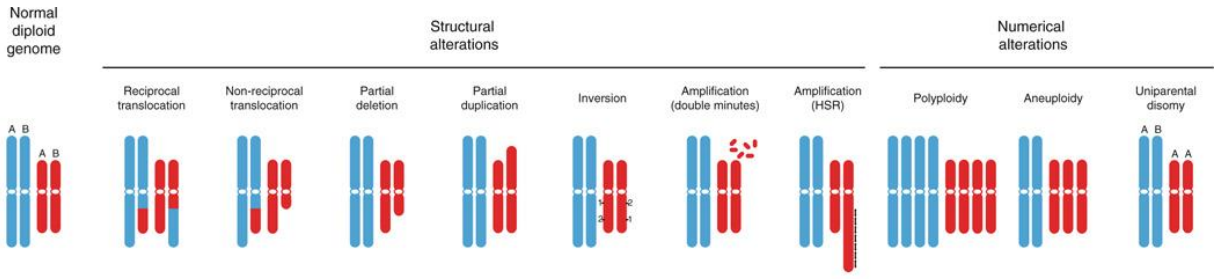


Figure 12. Array-based Comparative Genomic Hybridization aCGH. Representation of aCGH technique to analyze copy number imbalances. A genomic clone tiling path covers the region of interest (e.g., 8q21.1). After extraction and purification, the genomic DNA targets are arrayed. Normal and tumor DNA are hybridized into the microarray and detected using a scanner. Each array can be analyzed by the fluorescence ratio to identify changes in copy number (Bacino, 2017)

-PCR: as in hematological tumors, different PCR-based techniques (see above section), are used to detect mutations that frequently appear in the solid neoplasia (Igbokwe & Lopez-Terrada, 2011).

-Sanger sequencing and Next generation sequencing: NGS can analyze a full human genome in a short period of time. It determines the DNA sequence of the genome or the exome, detecting amplifications, deletions, fusions, methylation and gene expression. With this information cancer genetics can be accurately described, creating a gene panel for each malignancy (Müllauer, 2017). These techniques are further explained in the next section

A summary of the techniques used for the detection of common chromosomal alterations is shown in *Figure 13.*



Technique	Detection									
Banding	+	+	+	+	+	+	+	+	+	-
SKY/M-FISH	+	+	+	+	-	+	+	+	+	-
CGH/aCGH	-	+	+	+	-	+	+	-	+	-
SNP arrays	-	+	+	+	-	+	+	+	+	+
NGS	+	+	+	+	+	+	+	+	+	+

Figure 13. Common chromosomal alterations in solid tumors and techniques that detect them. Chromosome banding, FISH and CGH are low resolution techniques, while aCGH, SNP arrays and NGS are much higher resolution methods (Grade et al., 2015).

3.2.3. Molecular diagnosis of breast cancer

Breast cancer is a complex disease that presents a wide heterogeneity. Despite this, it has been the most deeply studied carcinoma, and molecular techniques have achieved a wide and accurate characterization of this neoplasia. This information has led to methods applicable in clinical practice that have changed the approach to the disease (Pareja et al., 2018).

a) Markers

-Hormone receptors: Breast cancer depends on estrogen and progesterone to grow, and the receptors to these hormones are overexpressed in some patients with breast cancer in the malignant breast tissue. Estrogen receptor (ER) and progesterone receptor belong to nuclear hormone receptor superfamily and act as ligand-dependent transcription factors (Hammond, 2016).

ER and PR expression are associated with better prognosis as the survival is longer and the early recurrence is lower. However, after five years, recurrences are more frequent than in ER-negative tumors. They are usually well differentiated and unfrequently associated with mutations, loss or amplification of genes associated with worse prognosis, like *HER2* or *TP53*.

These tumors tend to metastasize in bone, soft tissue and genital tract. On the other hand, ER-negative tumors metastasize in lower-survival organs, such as brain and liver (Foukakis & Bergh, 2018).

-HER2 oncogene encodes a transmembrane receptor with intracellular tyrosine kinase activity. It belongs to the EGFR family receptor *(Figure 14)* that activates pathways controlling cell proliferation, differentiation and angiogenesis (Yamauchi & Hayes, 2017).

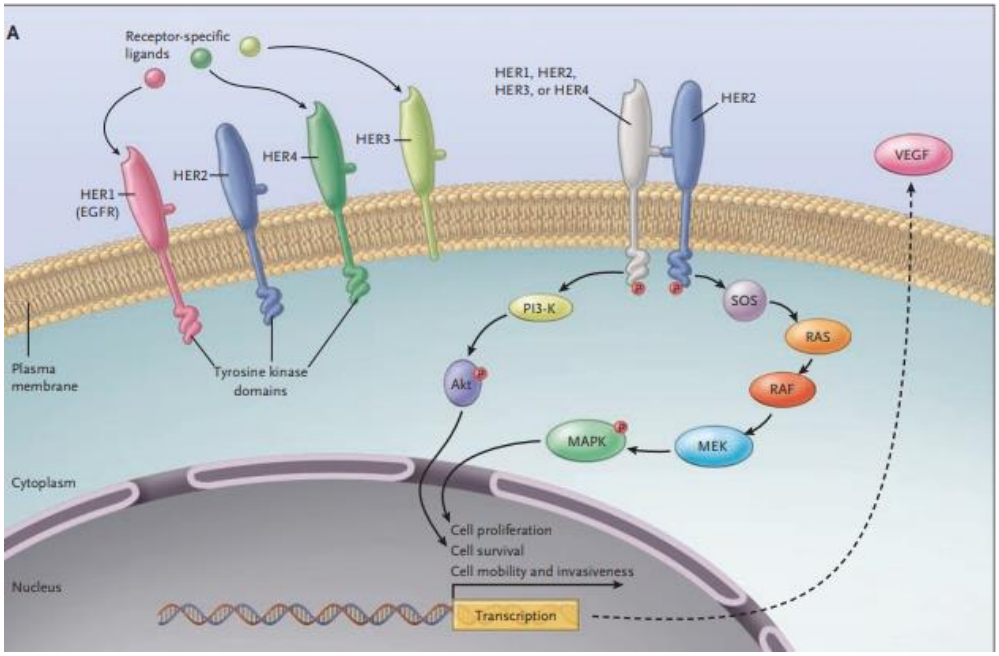


Figure 14. Signal Transduction by the HER family of growth factor receptors. Upon binding of specific ligands, phosphorylation of its intracellular tyrosine kinase domains causes dimerization, inducing cell proliferation and survival signaling. HER2 is the preferred dimerization partner. The activated tyrosine residues activate PI3-K that phosphorylates a phosphatidylinositol that phosphorylates the enzyme Ak, driving cell survival. At the same time, a guanine nucleotide exchange factor activates RAS that activates RAF and then the MAPK and MEK, driving cellular proliferation. One of the downstream effects is the production of VEGF, that supports angiogenesis (Hudis, 2007).

The analysis of overexpression or amplification of human epidermal growth factor receptor 2 (HER2) is performed in all diagnosed breast cancer samples. It involves worse prognosis itself, but chemotherapy and HER2 targeted therapies have supposed an increase in survival (Foukakis & Bergh, 2018).

b) Molecular subtypes

The study of the molecular alterations led to the classification of breast cancer into four subtypes *(Figure 15)*, that imply different prognosis and therapeutic targets (Foukakis & Bergh, 2018):

-Luminal subtypes are the most common. They are associated with estrogen receptor (ER) expression. The gene expression of these tumors is similar to the luminal epithelium of the breast. They express not just ER but also progesterone receptor and other genes

that promote ER activation. This group is differentiated into two subtypes with different molecular characteristics and prognosis:

-Luminal A: comprises 40% of breast tumors. They have high expression of ER and related genes, low expression of HER2 genes and of proliferation genes. This subtype has the best prognosis.

-Luminal B: 20% of breast tumors. The expression of ER related genes is lower, HER2 genes are expressed variably and proliferation genes are expressed higher than in luminal A. The prognosis is worse, presenting an elevated recurrence rate.

-*HER2*-enriched subtype: 10-15% of breast tumors fit into this subtype. They present high expression of *HER2* and proliferation genes and low expression of luminal and basal genes, so they are usually negative for ER and PR. This subtype differs from the clinical category HER2-positive, in which half of tumors are HER2-enriched but the other half is composed by other subtypes, especially by HER2-positive luminal subtypes.

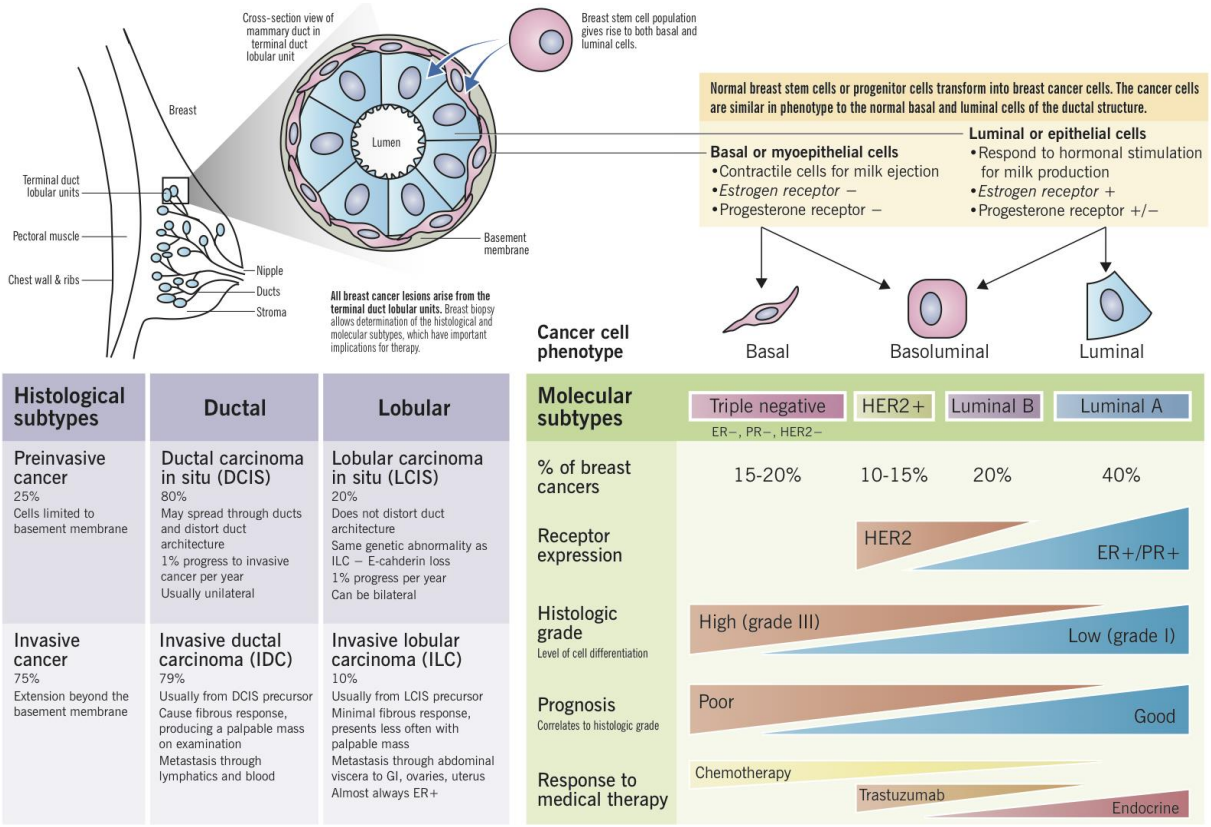


Figure 15. Breast cancer pathogenesis and histologic vs. molecular subtypes (Wong & Rebelo, 2012).

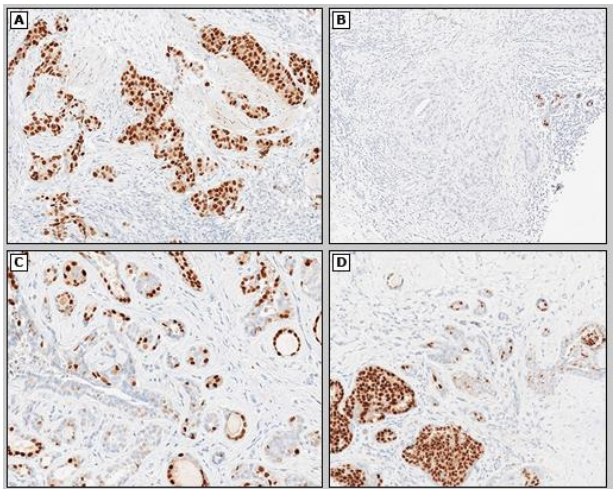
Furthermore, about one third of HER2-enriched tumors are HER2-negative in the clinical classification.

-ER-negative subtypes: include many subtypes, such as basal-like, claudin-low and interferon-rich. Most of them are categorized as triple-negative breast cancer due to negativity also to PR and HER2.

c) Molecular techniques

-Hormone receptor tests: The prognosis of the breast cancer is related to the expression of hormone receptors and there exist therapies that target ER, so these tests are performed in all breast cancers. PR detection tests are also performed because when it is positive, ER-negative results can be false, and patient would benefit from endocrine therapy.

Quantitation of ER and PR is performed by immunohistochemical (IHC) methods that detect the receptors by the binding of antibodies *(Figure 16)*.



a) Strong ER positivity. >90% stained cells

b) Tumor cells with no ER activity.

c) Breast cancer with variable ER activity. Some areas are negative while others show strong positivity. As>1% cells are stained is considered positive.

d)Intrinsic controls: shows ductal carcinoma in situ that stain strongly in the lower left and normal ducts at the upper right.

Figure 16. IHC: Estrogen receptor stain in breast cancer (Hammond, 2016).

The IHC tests are considered positive to ER or PR if more than 1% of cells are stained.

Even if immunohistochemical assays are still the recommended tests, there are newer techniques that detect ER mRNA by RT-PCR or microarrays *(Table 9)*. These techniques are often performed as part of commercial multigene assays (Hammond, 2016).

Table 9. List of commercially available prognostic gene signature assays that are clinically useful in the context of ER+/HER2- (Pareja et al., 2018).

	Mammaprint	Oncotype Dx	Prosigna	EndoPredict	Breast Cancer Index
Method	Microarrays	qRT-PCR	NanoString	qRT-PCR	qRT-PCR
Feasibiity on FFPE samples	Yes	Yes	Yes	Yes	Yes
Type of assessment	Central laboratory	Central laboratory	Local laboratory	Local laboratory	Central laboratory
Level I evidence	Yes, IA	Yes, IA	Yes, IB	Yes, IB	Yes, IB
Information regarding the molecular subtype	No	No	Yes	No	No

-*HER2* expression tests: since the presence of *HER2* implies changes in the prognosis and the treatment these tests are performed routinely in the newly diagnosed breast cancer.

Even if the activity of *HER2* can be detected by many methods, two assays are mainly validated and performed in clinical practice: detection of overexpression of

HER2 protein by immunohistochemistry *(Figure 17)* and detection of HER2 gene amplification by in situ hybridization *(Figure 18).*

Positivity in the test performed predict that the patient would be benefited by anti-HER2 therapy. Tests are taken as positive if the result of IHC score is 3+, FISH Her2/CEP17 ratio is >2 or ratio<2 with average *HER2* copy number is ≥2 signals/cell.

When one of the techniques provides an equivocal result of HER2 (IHC 2+ or FISH ratio<2 and average HER2 copy number ≥4.0 and <6.0 signals/cell) it is indicated to perform the other test or an alternative ISH method with the same sample or a new test with a different sample (Yamauchi & Hayes, 2017).

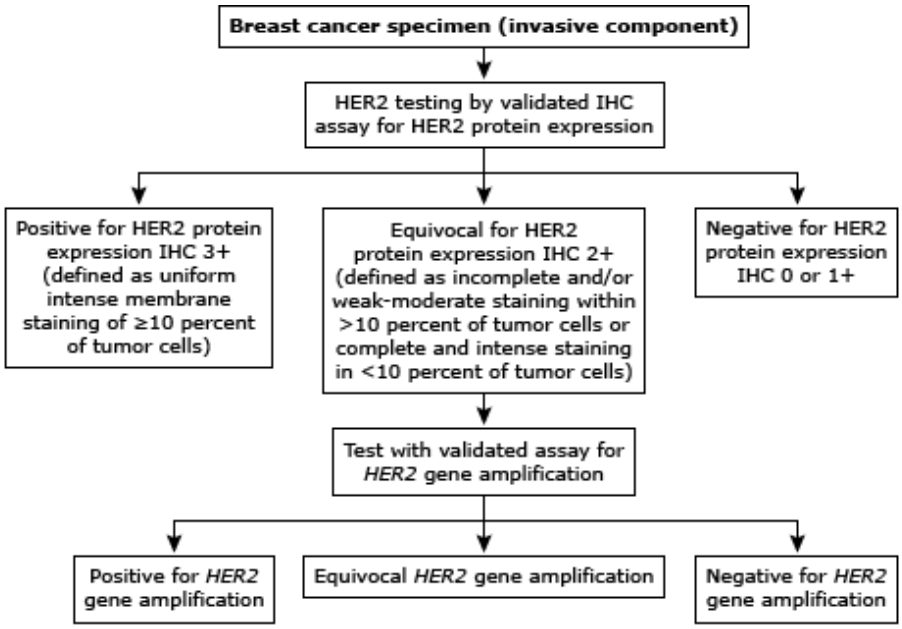


Figure 17. Algorithm for immunohistochemistry (IHC) in assessment of HER2 status (Yamauchi & Hayes, 2017).

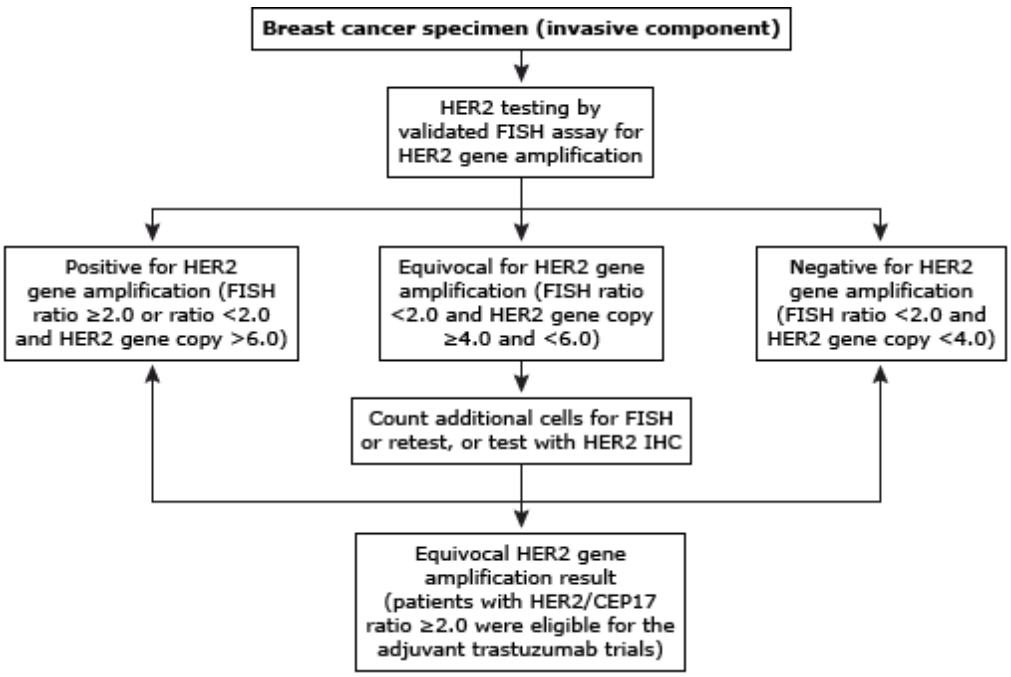


Figure 18. Algorithm for FISH in assessment of HER2 status (Yamauchi & Hayes, 2017).

d) Prognosis

The prognosis of breast cancer was established based on TNM staging, that, based on the anatomic extent of the neoplasia. However, since 2018 a new prognostic staging system that incorporates the prognostic biomarkers has been stablished (Hayes, 2018). This new TNM staging can be found in the *Annexe 3*.

The markers incorporated to the prognostic staging system are:

-ER and PR expression detected by IHC.

-HER2 assayed by IHC or ISH.

-Histologic grade determined by morphologic features that assign a score that goes between 1 (favorable) and 3 (not favorable)

-Recurrence score: based on Oncotype Dx (RT-PCR)

Absence of hormonal receptors and HER2 expression (triple-negative) implies worse prognosis in tumors that share the TNM staging and histologic grade. On the other hand, triple positive breast cancer has better prognosis in advanced breast cancer than hormone receptor positive or HER2 positive but similar in lowest and highest staging cancer.

e) Treatment

The treatment of breast cancer is sustained by a multidisciplinary approach by surgery, radiotherapy and chemotherapy.

Surgical treatment is considered the definitive therapy, but it needs neoadjuvant systemic therapy that reduces the risk of recurrence and reduce the extension of the surgery.

Chemotherapy was the classical neoadjuvant therapy, but endocrine and HER2-directed therapy can be applied in selected patients, improving survival.

The standard neoadjuvant therapy consists in chemotherapy and antiHER2-targeted drugs if the tumor is HER2 positive. If the tumor is hormone receptor-positive, chemotherapy is still the standard, but in selected patients endocrine therapy can be chosen (Sikov, 2017).

HER2-positive breast cancer therapy

Amplification of HER2 provokes an overactivation of signaling pathways, that involves higher aggressiveness but also an increased sensitivity to chemotherapy. It is the subtype that most frequently gets a pathologic complete response (pCR) to neoadjuvant chemotherapy, even if targeted therapies are not administered. If anti-HER2 targeted therapy is used, the response is even higher.

Neoadjuvant chemotherapy is considered in locally advanced breast cancer of any subtype and in earlier-stage HER2-positive neoplasia if it allows the breast-conserving surgery. Standard neoadjuvant therapy consists on chemotherapy (typically Docetaxel and carboplatin) and HER2-directed drugs. The anti-HER2 drug that is administered is Trastuzumab, which can be boosted with Pertuzumab.

Trastuzumab is a monoclonal antibody that binds to HER2 extracellular domain and inhibits proliferation of cells that overexpress HER2 protein by antibody-dependent cellular cytotoxicity *(Figure 19)*. It improves pCR rate, event-free survival (EFS) and overall survival (OS)

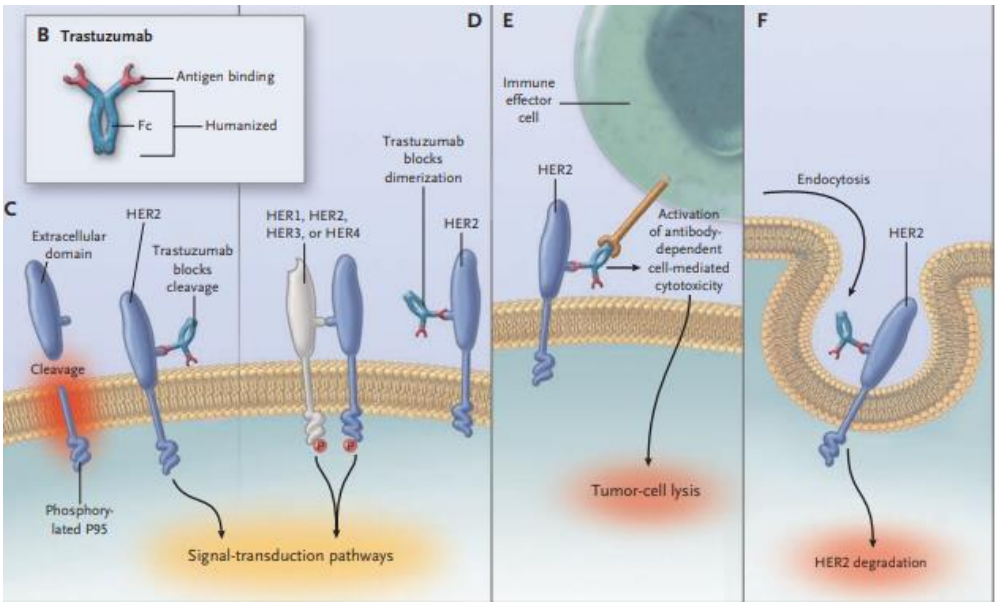


Figure 19. Mechanisms of action of Trastuzumab. b) When the extracellular domain of HER2 cleavages, it leaves a membrane-bound phosphorylated p95 that activates signal pathways. c) Trastuzumab binds to a juxtamembrane domain of HER2 and reduces shedding of the extracellular domain, reducing p95. d) Trastuzumab reduces dimerization of HER2, decreasing the signaling. e) Trastuzumab triggers cytotoxicity, leading to cell death. f) Trastuzumab down-regulates HER2 by endocytosis (Hudis, 2007).

Pertuzumab is also a monoclonal antibody, it binds to a different epitope on HER2, blocking the formation of HER2-HER3 dimers that seem to cause resistance to trastuzumab. It improves the locoregional response, but due to this toxicity and secondary effects it is not used in low-risk disease and patients with wide comorbidities.

These treatments are administered in combination with chemotherapy, but the non-chemotherapy treatment with combinations of HER2-targeted drugs are object of investigations in exceptional cases in which patient cannot receive chemotherapy (Sikov, 2018).

Hormone receptors-positive breast cancer therapy

Neoadjuvant therapy in HR-positive cancer is typically chemotherapy or endocrine therapy that facilitates posterior surgery and improves it results. Endocrine therapy is elected in HER2-negative HR-positive tumors with strong ER positivity, and the response to the therapy is positively correlated with the levels of ER expression.

Endocrine therapy gets worse results than chemotherapy in premenopausal women, but gets similar response rates in postmenopausal, so it is chosen in these patients. Endocrine drugs used are aromatase inhibitors (AIs). Different AIs show similar results but Letrozole has shown a higher suppression of estrogen levels (Bardia & Dixon, 2017).

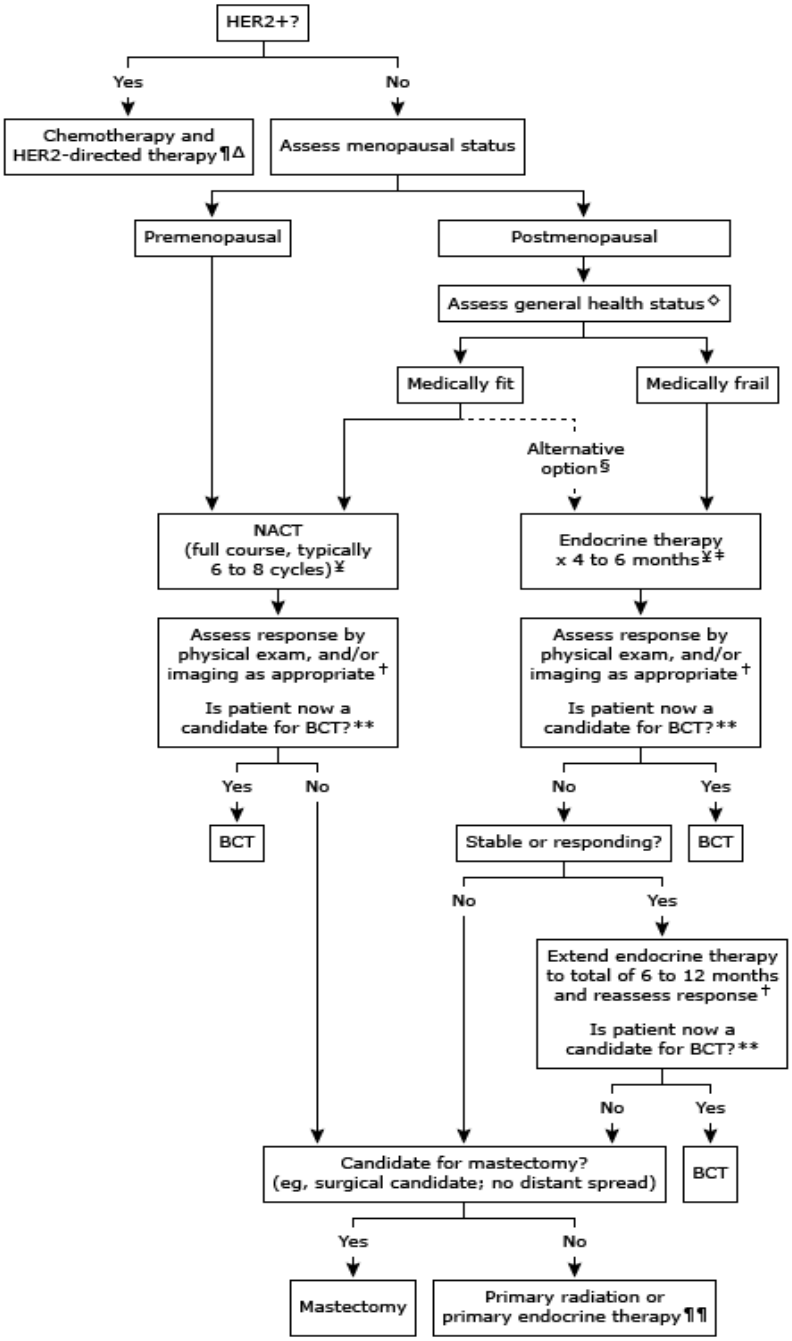


Figure 20. Approach for patients with locally advanced, hormone-positive breast cancer receiving neoadjuvant therapy (Bardia & Dixon, 2017).

3.3. NEW TOOLS FOR MOLECULAR DIAGNOSIS

In order to extend our knowledge on the molecular characterization of cancer, which could lead to the development of efficient directed therapies, new technologies and tools are being developed. As mentioned before, genomic studies of cancer have been greatly improved by next generation sequencing (NGS) methodologies (Müllauer, 2017). In addition, the analysis of circulating tumor DNA (ctDNA) from "liquid biopsies" has promising applications in clinical oncology (Donaldson & Ho Park, 2018). In this section these new tools for the molecular diagnosis of cancer are reviewed.

3.3.1. Next Generation Sequencing

Cancer research achieved the identification of cancer-related genes, but the greatest impulse to the genetic characterization of tumors was the development of next generation sequencing (NGS) technology.

a) Traditional sequencing methods vs NGS

NGS can rapidly sequence tumor genome at relative low cost. While the first sequencing with Sanger technology costed 3 billion US dollars and took 13 years, NGS can analyze a complete human genome in a week for 1000 dollars. Smallest NGS instruments sequence 3-15 gigabases per day while current Sanger sequencer analyze 1-2 megabases (Müllauer, 2017).

b) Technique

NGS technique sequences multiple small fragments of DNA in parallel. Patient's DNA is prepared by purification, amplification and fragmentation. Then the fragments are physically isolated and fixed to solid surfaces. These fragments are sequenced, and the results are compared with a reference sequence, currently the GRCh38.p12. (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.38/).

Different NGS assays can be used to sequence either the whole genome, the exome or a preselected gene panel:

-Sequencing the entire genome is more expensive but it provides information about non-coding DNA human which role is becoming increasingly important.

-The exome represents about 2% of the 3,3 gigabases that compose the genome. Most of mutations that are known to cause diseases are contained in it.

-Gene panels are typically composed of 10 to 200 genes, they can be used to diagnose diseases based on known mutations that are too many to be sequenced individually. For most clinical application it is the method of choice because it is cheaper, it reduces the casual finding of alterations that are unrelated to the disease and it improves sensitivity. The results obtained are easier to analyze.

Accuracy of NGS for the entire genome and exome is around 95%, which is higher for gene panels. This loss of precision compared to traditional methods is mostly caused by the amplification process. Since Sanger sequencing gets >99,99% accuracy it is still considered as the gold standard for diagnosis, so it was used to confirm variants reported as pathogenic. This use of Sanger as secondary validation of the results has been challenged by the improvement of NGS technology, that allows the establishment of a definitive diagnosis without performing Sanger techniques.

Furthermore, Third Generation Sequencing is being developed. It uses single DNA molecules, instead of amplified DNA, to overcome this loss of accuracy due to the amplification techniques. Different software are used for the interpretation being the databases consulted for the definition of the pathogenic implication critical for an accurate classification (Hulick, 2018).

c) Interpretation

The biggest challenge is the clinical interpretation of the huge amount of data provided by NGS. Also, it reports variants on multiple genes with unknown significance. Gene variations can be reported as:

-Pathogenic: strong evidence of causing disease.

-Likely pathogenic: apparently implicated in pathogenesis but there is not conclusive evidence.

-Likely benign: most evidences suggest that the effect is not pathogenic.

-Benign: variants do not alter function.

-Unknown clinical significance: data that suggests gene function alteration but without evidence of being benign or pathogenic.

While most laboratories agree on the variants established as pathogenic or benign, there is a debate about the variations that belong to the other categories. Consistence in the interpretation of the signification of the variants between laboratories will be important in order to guide clinicians in the genetic counseling and the consequent clinical decisions (Hulick, 2018). Different softwares are used for the interpretation being the databases consulted for the definition of the pathogenic implication critical for an accurate classification. The high ratio of errors detected in some databases have led to the creation of some specific databeses e.g a specific TP53 database has been created and is constantly updated and validated.

d) Applications in oncology

As there are techniques that analyze a few genes more accurately, rapidly and at a lower cost, NGS is used when many genes have to be screened.

It is indicated for diagnosing complex diseases in which many genes may be responsible and NGS would be cheaper and more efficient than sequencing each gene individually. It can be also applied when most common genes have been tested and have

not shown any pathologic variant, so the study has to be expanded to less-likely-pathogenic genes.

In many cancers, especially in those in which there exist different genetic variants, targeted panels have proven its usefulness (Hulick, 2018; Serratì et al., 2016). They have a wide use in inherited cancer syndromes:

-Hereditary breast cancer (HBC): conventional DNA sequencing of *BRCA1* and *BRCA2*, that cause about 30% of HBC, requires long time and it is expensive, because they are long genes. Gene panels include *BRCA1* and *BRCA2*, and other genes associated with HBC such as *ATM, BRIP1, CDH1, CHEK2, FANC1, MSH2, MUTYH, NF1, PALB2, RAD51C* and *TP53.*

-Inherited gastrointestinal cancer: panels that include *APC, BMPR1A, EPCAM, MLH1, MSH2, MSH6, MUTYH, PMS2, PTEN, SMAD4, STK11, TP53, BLM, CHEK2, GALNT12, GREM1, POLD1,* and *POLE.*

-Diagnosis of familiar acute leukemia and myelodysplastic syndromes. Longwood panel (http://www.dlongwood.com/a/catalogo/ficha/699/Panel-Mieloide-de-Sophia-Genetics-CEIVD-NGS.html) is one of the most used panels in these pathologies and it includes *ABL1, ASXL1, BRAF, CALR, CBL, CEBPA, CSF3R, DNMT3A, ETV6, EZH2, FLT3, HRAS, IDH1, IDH2, JAK2, KIT, KRAS, MPL, NPM1, NRAS, PTPN11, RUNX1, SETBP1, SF3B1, SRSF2, TET2, TP53, U2AF1, WT1* and *ZRSR2.*

-Categorization of prognostic groups in AML, with gene panels that include *ASXL1, ASXL2, CEBPA, DNMT3A, FLT3, IDH, KIT, NPM, TP53* and *WT1.*

NGS is useful to identify variations that can be the target of directed therapies and to analyze pathogenic changes that can exist in tumors and that differentiate them from adjacent normal tissue, which could guide therapeutic decisions based on genetic abnormalities (Hulick, 2018; Serratì et al., 2016).

Finally, NGS has a great potential as a research tool. It can identify new genes that could explain syndromes and determinate the genetic changes during the development of cancer. This would lead to the understanding of the pathogenesis and the development of diagnostic tests and therapies based on the genetics (Hulick, 2018).

3.3.2. Liquid biopsy

Molecular pathology of cancer is not just being widely developed by the methods of analysis but could also adopt change in the way that samples are taken. The analysis of the DNA that release the cells when they die could eventually be a less invasive substitute of conventional biopsy. This circulating cell-free DNA (ccfDNA) is obtained with the sampling by "liquid biopsy". Part of this ccfDNA will be circulating tumor DNA (ctDNA) and can be measured and analyzed to determine tumor's features *(Figure 20)* (Bardelli & Pantel, 2017).

CcfDNA can be found in blood, lymph, urine, saliva and CSF. These fluids would contain other circulating genetic material. It is usually collected from peripheral blood

sample, obtaining serum that contains the ccfDNA. NGS techniques have increased the sensitivity and specificity in the analysis of ccfDNA, achieving reproducibility that makes liquid biopsy applicable in clinical practice. Analysis of ctDNA detected a difference in the methylation patterns and length between ctDNA and ccfDNA. CtDNA supposes a variable amount of cfDNA, from 0,01% to more than 90% (median of 0,18%) and the detection of the alterations is even more difficult because of the presence of subclonal mutations, undetectable by most of techniques.

The development of high sensitivity tests that could detect early cancer by a low aggressive method like liquid biopsy, that could be widely used in the population or risk individuals, would suppose a milestone in tumor screening. Other application of liquid biopsies would be the monitor of the minimal residual disease in solid tumors, as presence of high levels of ctDNA is associated with worse prognosis, high incidence of metastasis and lower survival. It would prevent metastatic disease and recurrences by the early adjuvant therapy. The periodic analysis of ctDNA during the therapy would inform on the response to the treatment. Moreover, resistance to the drug that could have been developed during treatment could also be detected.

One already approved application of ctDNA-based tests is the guide of anti-EGFR therapy in non-small-cell lung cancer with mutated EGFR. Traditional tissue biopsies are still the standard in the initial tumor diagnosis and its substitution for the analysis of ctDNA is currently limited. Beside this, ctDNA gives wider information, not just the obtained by the analysis of the point biopsied that can be less representative of the whole tumor, and allows a real time following of tumor's molecular features, so it is expected to be the gold standard in the future, although now it cannot substitute other techniques and its clinical use is limited (Aravanis et al., 2017; Bardelli & Pantel, 2017; Donaldson & Ho Park, 2018; Müllauer, 2017).

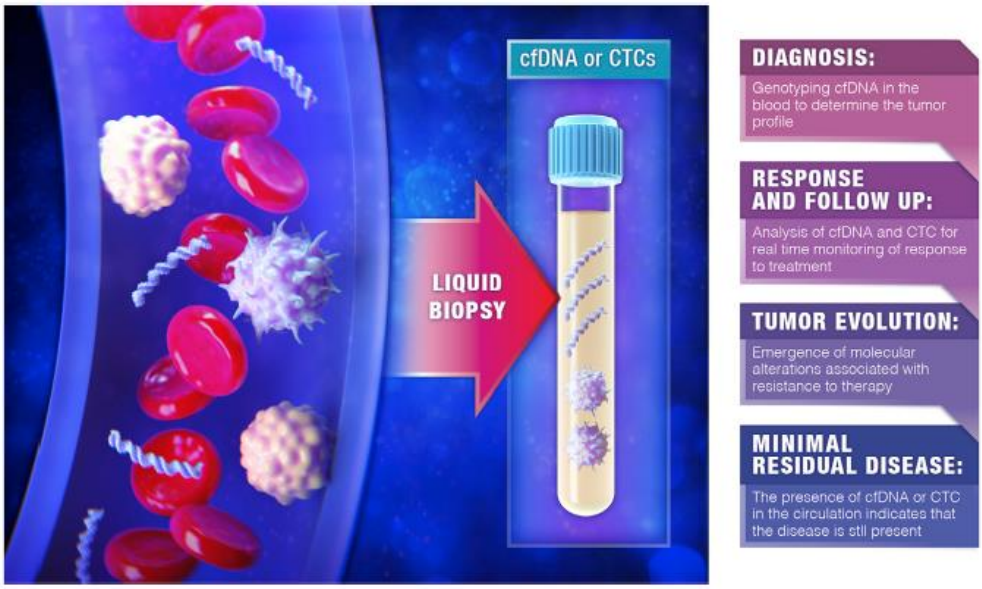


Figure 20. Biological basis of liquid biopsies based on circulating tumor DNA and circulating tumor cells (CTCs), that are present in fluids. Liquid biopsies could be useful in tumor diagnosis and determination of tumor profile, response to therapy and follow up, monitoring the alterations that associate with resistance and tracking minimal residual disease (Bardelli & Pantel, 2017).

4. CONCLUSIONS AND FUTURE PERSPECTIVES

Advances in the molecular techniques, especially NGS, have achieved the establishment of a wide number of genes involved in cancer. The relative lower cost, higher velocity and availability of these techniques make them very useful to the clinical oncology. The farthest goal of tumor profiling is the identification of its vulnerabilities and the development of precision therapies and multidisciplinary treatment approach against it (Senft et al., 2017). Particularly, the pathogenic gene report should inform of those genetic alterations that being relevant for the maintenance of the tumor might be druggable (Khotskaya, 2017).

Many targeted therapies have already shown higher efficacy than the equivalent non-targeted treatment with apparently less secondary effects (Khotskaya, 2017). It is however urgent to identify biomarkers that predict response to targeted therapies and immunotherapy, because very frequently treatments show impressive efficacy but just in a limited number of patients or for a limited period of time. For the treatment of solid tumors there are many targeted therapies approved, but only half of this therapies have a predictive biomarker that allows to select the patients that may benefit from the specific therapy (Kurnit et al., 2018).

In the *Figure 20,* a scheme of the future precision oncology service is shown. Precision oncology starts with the identification of the tumor sample that would be analyzed, considering the combination of a liquid biopsy to test ccfDNA with traditional methods that are more representative. Then, the sample will be sequenced by NGS techniques to detect molecular alterations. The results will be interpreted, reporting the molecular profile of the tumor. Eventually, the precision oncology service will not only be able to determine what alterations are actionable and what targeted therapies are available for this gene, but also which abnormalities should be addressed to eliminate the tumor. In an advanced tumor more than 10 druggable mutations can be found and there is still uncertain which one should be chosen as target of therapy, or if it should be composed of a combination of the treatments against each mutation. This determines the low efficacy of personalized medicine nowadays. Oncologist will be counseled on which approach is most appropriate. This will lead to genome driven oncology, in which every patient is managed according to its cancer genome characteristics (Kurnit et al., 2018).

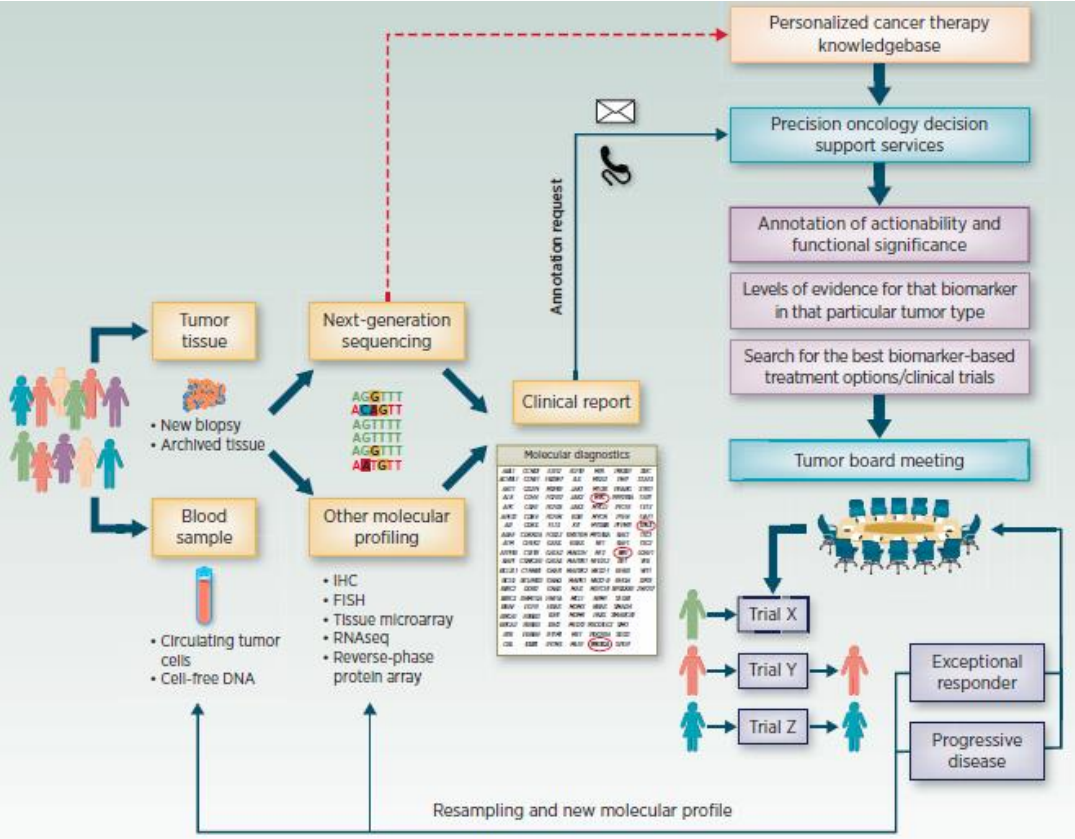


Figure 20. Flow diagram of precision oncology services. It is a complex process that starts with the solid and liquid biopsy of the tumor. Then, NGS techniques and other additional molecular techniques are performed in order to obtain a molecular profile of the patient's tumor. The results are communicated to the precision oncology decision services that interpret the profile, highlighting which genes are actionable and useful as biomarkers and select the best therapy to the patients. The results are explained to the clinical services in a tumor board meeting. Then, the patient is treated with available treatments or unrolled in a clinical trial (Kurnit et al., 2018).

5. BIBLIOGRAPHY

Aravanis A, Lee M, Klausner R. Next-Generation Sequencing of Circulating Tumor DNA for Early Cancer Detection. Cell. 2017;168:571-4.

Arber D, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-405.

Bacino C. Genomic disorders: An overview 2017 [cited 2018]. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/genomic-disorders-an-overview?search=cgh&source=search_result&selectedTitle=1~88&usage_type=default&display_rank=1.

Bain B, Clark D, BS W. Bone Marrow pathology. 4 ed: Wiley-Blackwell; 2011.

Bardelli A, Pantel K. Liquid Biopsies, What We Do Not Know (Yet). Cancer cell. 2017;31(2):172-9.

Bardia A, Dixon M. Neoadjuvant therapy for newly diagnosed hormone-positive breast cancer 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/neoadjuvant-therapy-for-newly-diagnosed-hormone-positive-breast-cancer?topicRef=14225&source=related_link.

Batlle-López A, Costa D, Martínez-Laperche C, Anguita E, Espinet B, Buño I. Genética hematológica: Manual práctico. 1 ed2016.

Bennour A, Saad A, Sennana H. Chronic myeloid leukemia: Relevance of cytogenetic and molecular assays. Critical reviews in oncology/hematology. 2016;97:263-74.

Cooper G, Hausman R. The cell: a molecular approach. 6th ed: ASM Press; 2014.

Chereda B, Melo J. Natural course and biology of CML. Annals of hematology. 2015;94 Suppl 2:107-21.

Chmielecki J, Meyerson M. DNA Sequencing of Cancer: What Have We Learned? Annual Review of Medicine. 2014;65:63-79.

Davids M. Approach to the adult with lymphocytosis or lymphocytopenia 2018. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/approach-to-the-adult-with-lymphocytosis-or-lymphocytopenia?search=flow%20cytometry&source=search_result&selectedTitle=2~150&usage_type=default&display_rank=2.

Deininger M. Molecular monitoring in CML and the prospects for treatment-free remissions. Hematology American Society of Hematology Education Program. 2015;2015:257-63.

Deininger M, Druker B. Specific targeted therapy of chronic myelogenous leukemia with imatinib. Pharmacological reviews. 2003;55(3):401-23.

Ding L, et al. Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. Cell. 2018;173(2):401-23.

Donaldson J, Ho Park B. Circulating Tumor DNA: Measurement and Clinical Utility. Annual Review of Medicine. 2018;69:223-34.

Druker B, Lydon N. Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukemia. . Journal of Clinical Investigation. 2000;105(1):3-7.

Ferlay J, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer 2014.

Foukakis T, Bergh J. Prognostic and predictive factors in early, nonmetastatic breast cancer 2018. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/prognostic-and-predictive-factors-in-early-nonmetastatic-breast-cancer?sectionName=GENOMIC%20PROFILES&topicRef=744&anchor=H697669959&source=see_link#H697669959.

Gonzalez de Castro D, Clarke P, Al-Lazikani B, Workman P. Personalized cancer medicine: molecular diagnostics, predictive biomarkers, and drug resistance. Clinical pharmacology and therapeutics. 2013;93(3):252-9.

Grade M, Difilippantonio M, Camps J. Patterns of Chromosomal Aberrations in Solid Tumors. Recent results in cancer research Fortschritte der Krebsforschung Progres dans les recherches sur le cancer. 2015;200:115-42.

Gray W, Kocjan G. Cytopathology: the history, the present and the future direction Diagnostic Cytopathology. 3rd ed: Elsevier; 2010. p. 3-13.

Hallek M. Chronic lymphocytic leukemia: 2015 Update on diagnosis, risk stratification, and treatment. American journal of hematology. 2015;90(5):446-60.

Hammond HE, M. Hormone receptors in breast cancer: Clinical utility and guideline recommendations to improve test accuracy 2016. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/hormone-receptors-in-breast-cancer-clinical-utility-and-guideline-recommendations-to-improve-test-accuracy?topicRef=782&source=see_link.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

Hayes D. Tumor, Node, Metastasis (TNM) staging classification for breast cancer 2018. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/tumor-node-metastasis-tnm-staging-classification-for-breast-cancer?topicRef=737&source=see_link.

Hoadley K, Yau C, Hinoue T, Wolf D, Lazar AE, Shen R. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell. 2018;173(2):291-304.

Hudis CA. Trastuzumab--mechanism of action and use in clinical practice. The New England journal of medicine. 2007;357(1):39-51.

Hulick P. Next-generation DNA sequencing (NGS): Principles and clinical applications 2018. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/next-generation-dna-sequencing-ngs-principles-and-clinical-applications?search=ngs&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1.

Igbokwe A, Lopez-Terrada D. Molecular Testing of Solid Tumors. Arch Pathol Lab Med. 2011;135:67-82.

Khotskaya Y, Gordon B, Mills S, Kenna R. Next-Generation Sequencing and Result Interpretation in Clinical Oncology: Challenges of Personalized Cancer Therapy. Annual Review of Medicine. 2017;68:113-25.

Kurnit K, et al. Precision Oncology Decision Support: Current Approaches and Strategies for the Future. Clinical cancer research : an official journal of the American Association for Cancer Research. 2018;24(12):1-13.

Mauro M, Druker B. STI571: Targeting BCR-ABL as therapy for CML. Oncologist. 2001;6:233.

McCance K, Huether S. Pathophysiology: The Biologic Basis for Disease in Adults and Children. 8 ed: Elsevier; 2018. p. 350.

Müllauer L. Milestones in pathology—from histology to molecular biology. memo - Magazine of European Medical Oncology. 2017;10(1):42-5.

Negrin R, Schiffer C. Clinical use of tyrosine kinase inhibitors for chronic myeloid leukemia 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/clinical-use-of-tyrosine-kinase-inhibitors-for-chronic-myeloid-leukemia?topicRef=4496&source=see_link.

Negrin R, Schiffer C. Overview of the treatment of chronic myeloid leukemia 2018. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/overview-of-the-treatment-of-chronic-myeloid-leukemia?topicRef=4537&source=related_link.

Pareja F, Caterina M, Reis-Filho J. Molecular diagnosis in breast cancer. Diagnostic histopatology. 2018;24(2):71-82.

Raby B. Tools for genetics and genomics: Polymerase chain reaction 2016. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/tools-for-genetics-and-genomics-polymerase-chain-reaction?search=pcr&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1.

Sanchez-Vega F, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell. 2018;173(2):321-37.e10.

Sangle N. Chronic Myeloid Neoplasms . Myeloproliferative neoplasms (MPN). Chronic myelogenous leukemia (CML) 2011. Available from: http://www.pathologyoutlines.com/topic/myeloproliferativecml.html.

Schrijver IZ, James L; Cherry, Athena M. Tools for genetics and genomics: Cytogenetics and molecular genetics 2017 [updated Sept 19, 2017; cited 2018]. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/tools-for-genetics-and-genomics-cytogenetics-and-molecular-genetics?search=karyotype&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1.

Senft D, Leiserson M, Ruppin E, Ronai Z. Precision Oncology: The Road Ahead. Trends Mol Med. 2017;23(10):874-98.

Serratì S, et al. Next-generation sequencing: advances and applications in cancer diagnosis. Onco Targets Ther. 2016;9:7355-65.

Siegel R, Miller K, Jemal A. Cancer Statistics, 2017. CA: a cancer journal for clinicians. 2017;67(1):7-30.

Sikov W. Neoadjuvant therapy for patients with HER2-positive breast cancer 2018. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/neoadjuvant-therapy-for-patients-with-her2-positive-breast-cancer?topicRef=14225&source=related_link.

Sikov WM. General principles of neoadjuvant therapy for breast cancer 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/general-principles-of-neoadjuvant-therapy-for-breast-cancer?topicRef=800&source=related_link.

Spranger S, Gajewski TF. Impact of oncogenic pathways on evasion of antitumour immune responses. Nature reviews Cancer. 2018;18(3):139-47.

Stock W, Estrov Z. Detection of minimal residual disease in acute lymphoblastic leukemia 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/detection-of-minimal-residual-disease-in-acute-lymphoblastic-leukemia?search=reverse%20pcr&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1.

Svoronos A, Engelman D, Slack F. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. Cancer research. 2016;76(13):3666-70.

Swerdlow S, et al. WHO classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th ed: WHO; 2017.

Van Etten R. Molecular genetics of chronic myeloid leukemia 2015. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/molecular-genetics-of-chronic-myeloid-leukemia?search=philadelphia%20kariotype&source=search_result&selectedTitle=5~101&usage_type=default&display_rank=5.

Van Etten R. Cellular and molecular biology of chronic myeloid leukemia 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/cellular-and-molecular-biology-of-chronic-myeloid-leukemia?sectionName=BCR-ABL1&topicRef=4537&anchor=H12&source=see_link#H12.

Vos T, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet (London, England). 2016;388(10053):1545-602.

Weinberg RA. The biology of cancer. 2nd ed: Garland science; 2014.

Wong E, Rebelo J. Breast cancer pathogenesis and hystologic vs. molecular subtypes McMaster Pathophysiology Review2012. Available from: http://www.pathophys.org/breast-cancer/breastcancer-copy/.

Yamauchi H, Hayes D. HER2 and predicting response to therapy in breast cancer 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/her2-and-predicting-response-to-therapy-in-breast-cancer?search=her2&source=search_result&selectedTitle=1~143&usage_type=default&display_rank=1.

Zhang YLB, Michelle M. General aspects of cytogenetic analysis in hematologic malignancies 2017. Available from: https://www-uptodate-com.scsalud.a17.csinet.es/contents/general-aspects-of-cytogenetic-analysis-in-hematologic-malignancies?search=fish&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1.

ANNEXES

ANNEXE 1. Food and Drug Administration (FDA) approved targeted therapies for tumors that have an associated biomarker (Kurnit et al., 2018).

Preferred Name	Direct drug Target	Company	FDA Approved Indication - Disease(s)	FDA Approved Indication - Biomarker(s)
Abemaciclib	CDK4, CDK6	Eli Lilly and Company	Breast cancer	ER Positive, PR Positive, HER2 Negative
Afatinib	EGFR	Boehringer Ingelheim	Non-small cell lung carcinoma	EGFR Deletion Exon 19, EGFR L858R
Alectinib	EML4-ALK, ALK, INSR, RET	Genentech	Non-small cell lung carcinoma	ALK Fusion
Anastrozole	Aromatase	AstraZeneca Pharmaceuticals	Breast cancer	ER Positive, PR Positive
Bosutinib	ABL1, BCR-ABL1, SRC	Pfizer	Chronic myelogenous leukemia	BCR-ABL1
Brigatinib	ALK, CSF1R, INSR, ABL1, LCK, IGF1R, CAMK2G, FLT4, RET, FGFR1, FGFR2, FGFR3, AURKA, JAK2, FGFR4, FYN, HCK, LYN, SRC, EGFR, EML4-ALK, FER, FES, FLT3, FPS, ROS1, TYK2, YES1, PTK2B, HER4, CAMK2D, CHEK1, CHEK2	Ariad Pharmaceuticals	Non-small cell lung carcinoma	ALK Fusion
Ceritinib	NPM-ALK, ROS1 Fusion, ALK, INSR, IGF1R, TSSK3, FLT3, FGFR2, RET, FGFR3, LCK, JAK2, AURKA, LYN, EGFR, FGFR4	Novartis Pharmaceuticals	Non-small cell lung carcinoma	ALK Fusion
Cetuximab	EGFR	Eli Lilly and Company	Colorectal cancer	KRAS Wildtype, EGFR Positive
Cobimetinib	MAP2K1	F. Hoffmann-La Roche	Melanoma	BRAF V600E, BRAF V600K
Crizotinib	ALK, MET, ROS1, NTRK1	Pfizer	Non-small cell lung carcinoma	ALK Fusion
		Pfizer	Non-small cell lung carcinoma	ROS1 Positive
Dabrafenib	BRAF, RAF1	GlaxoSmithKline	Melanoma	BRAF V600E
		GlaxoSmithKline	Melanoma	BRAF V600E, BRAF V600K
		GlaxoSmithKline	Non-small cell lung carcinoma	BRAF V600E
Dasatinib	ABL1, KIT, BRAF, BCR-ABL1, ABL2, PDGFRA, PDGFRB, SRC, DDR1, DDR2, EPHA3 Amplification, EPHA2, FYN, LCK, LYN, YES1	Bristol-Myers Squibb	Chronic myelogenous leukemia	BCR-ABL1
			Acute lymphoblastic leukemia	BCR-ABL1
Enasidenib	IDH2	Agios Pharmaceuticals	Acute myeloid leukemia	IDH2 Mutation

Preferred Name	Direct drug Target	Company	FDA Approved Indication - Disease(s)	FDA Approved Indication - Biomarker(s)
Erlotinib	EGFR	Genentech	Non-small cell lung carcinoma	EGFR Deletion Exon 19, EGFR L858R
Everolimus	MTOR	Novartis Pharmaceuticals	Breast cancer	ER Positive, PR Positive, HER2 Negative
Exemestane	Aromatase	Pfizer	Breast cancer	ER Positive
Fulvestrant	ER	AstraZeneca Pharmaceuticals	Breast cancer	ER Positive, PR Positive, HER2 Negative
			Breast cancer	ER Positive, PR Positive
Gefitinib	EGFR	AstraZeneca Pharmaceuticals	Non-small cell lung carcinoma	EGFR Deletion Exon 19, EGFR L858R
Gemtuzumab Ozogamicin	CD33	Pfizer	Acute myeloid leukemia	CD33 Positive
Ibrutinib	TEC, ABL1, FYN, RIPK2, SRC, LYN, PDGFRA, HER2, BTK, EGFR, BLK, BMX, CSK, FGR, PTK6, HCK, YES1, ITK, JAK3, FRK, LCK, RET, FLT3	Janssen Biotech|Pharmacyclics	Small lymphocytic lymphoma, Chronic lymphocytic leukemia	c.CHR17p Deletion
Imatinib	PDGFRA, KIT, BCR-ABL1, ABL1, PDGFRB	Novartis Pharmaceuticals	Gastrointestinal Stromal Tumors	KIT Positive
			Chronic myeloid leukemia, Acute lymphoblastic leukemia,	BCR-ABL1
			Myelodysplastic/myeloproliferative diseases	PDGFRA Fusion
			Chronic eosinophilic leukemia	FIP1L1-PDGFRA
Lapatinib	EGFR, HER2, HER4	Novartis Pharmaceuticals	Breast cancer	HER2 Positive
			Breast cancer	ER Positive, PR Positive, HER2 Positive
Letrozole	Aromatase	Novartis Pharmaceuticals	Breast cancer	ER Positive, PR Positive
Midostaurin	KDR, FLT3, PDGFRA, PDGFRB, SYK, AKT1, FLT1, AKT2, AKT3, KIT, SRC, PRKCA, PRKCB, PRKCG, CDK1, FGR, ETV6-NTRK3	Novartis Pharmaceuticals	Acute myeloid leukemia	FLT3 Mutation
Neratinib	HER2, EGFR, KDR	Puma Biotechnology, Inc.	Breast cancer	HER2 Overexpression, HER2 Amplification
Nilotinib	BCR-ABL1	Novartis Pharmaceuticals	Chronic myelogenous leukemia	BCR-ABL1
Olaparib	PARP1, PARP2	AstraZeneca Pharmaceuticals	Ovarian cancer	BRCA1 (any deleterious), BRCA2 (any deleterious)
Osimertinib	EGFR, EGFR T790M, EGFR Exon 19 deletion	AstraZeneca Pharmaceuticals	Non-small cell lung carcinoma	EGFR T790M
Palbociclib	CDK4, CDK6	Pfizer	Breast cancer	ER Positive, PR Positive, HER2 Negative

Preferred Name	Direct drug Target	Company	FDA Approved Indication - Disease(s)	FDA Approved Indication - Biomarker(s)
Panitumumab	EGFR	Amgen	Colorectal cancer	KRAS Wildtype, NRAS Wildtype
Pertuzumab	HER2	Genentech	Breast cancer, Inflammatory breast cancer	HER2 Positive
Ponatinib	PDGFRA, KDR, SRC, ABL1, FGFR1, BCR-ABL1, KIT, RET	Ariad Pharmaceuticals	Acute lymphoblastic leukemia / lymphoblastic lymphoma, Chronic myeloid leukemia	BCR-ABL1 T315I
			Chronic myeloid leukemia, Acute lymphoblastic leukemia	BCR-ABL1
Ribociclib	CDK4, CDK6	Novartis Pharmaceuticals	Breast cancer	ER Positive, PR Positive, HER2 Negative
Rituximab	CD20	Genentech	Non-Hodgkin's lymphoma, Chronic lymphocytic leukemia	CD20 Positive
Rucaparib	PARP1	Clovis Oncology	Ovarian cancer	BRCA1 (any deleterious), BRCA2 (any deleterious)
Tamoxifen	ER	AstraZeneca Pharmaceuticals	Breast cancer	ER Positive (may help predict whether therapy will be beneficial)
Trametinib	MAP2K1, MAP2K2	GlaxoSmithKline	Melanoma	BRAF V600E, BRAF V600K
			Non-small cell lung carcinoma	BRAF V600E
Trastuzumab	HER2	Genentech	Breast cancer, Gastric cancer, Gastroesophageal junction	HER2 Positive
Trastuzumab Emtansine	HER2, p.Tubulin	Genentech	Breast cancer	HER2 Positive
Vemurafenib	BRAF V600E	F. Hoffmann- La Roche	Melanoma	BRAF V600E
Venetoclax	BCL2	AbbVie	Chronic lymphocytic leukemia	c.CHR17p Deletion

ANNEXE 2. Actionable genes, the alteration types, and the alteration frequencies for several common cancer types (Kurnit et al., 2018).

Tumor type	Actionable genes	Alteration type	Frequency	Comments
Non-small cell lung cancer	BRAF	Mutations	5-10%	
	DDR2	Mutations	1-6%	
	EGFR	Mutations	4-18%	
	EML4-ALK	Fusion	4%	
	ERBB2	Mutations	2-3%	
	FGFR1	Amplification	2-17%	
	FGFR3	Fusion	2%	
	KRAS	Mutations	1-32%	1% in adenocarcinoma, 32% in squamous cell carcinoma
	MAP2K1	Mutations	1%	
	MET	Amplification	1-4%	
	MET	Mutations	3-8%	3% MET exon 14 mutation in lung adenocarcinoma
	NF1	Mutations	11%	
	NTRK1	Fusion	2-4%	
	PIK3CA	Mutations	4-16%	
	PTEN	Mutations/ Deletion	1-8%	
	RET	Fusion	2-4%	
	RICTOR	Amplification	2-5%	
	ROS1	Fusion	4-11%	
	STK11	Mutations	2-17%	
Bladder	AKT1	Mutations	3%	
	CDKN2A	Deletion	47%	
	CDKN2A	Mutations	5%	
	EGFR	Amplification	11%	
	ERBB2	Amplification	7%	
	ERBB3	Mutations	11%	
	FGFR3	Mutations	45%	60-80% in non-muscle-invasive; 15–20% in muscle-invasive bladder cancer
	FGFR3	Amplification	3%	
	FGFR3-TACC3	Fusion	5%	

Tumor type	Actionable genes	Alteration type	Frequency	Comments
	KRAS	Mutations	4%	
	MDM2	Amplification	9%	
	PIK3CA	Mutations	20%	
	PTEN	Mutations	3%	
	PTEN	Deletion	13%	
	TSC1	Mutations	9%	
Biliary	BRAF	Mutations	7%	
	EGFR	Mutations/ Amplification	5%	
	ERBB2	Mutations/ Amplification	4-18%	18% in gallbladder carcinoma
	FGF19	Amplification	3%	
	FGFR1	Mutations/ Amplification	4%	
	FGFR2	Fusion	5%	5% in intrahepatic cholangiocarcinoma
	IDH1/2	Mutations	0-6%	4-6% in intrahepatic cholangiocarcinoma
	KRAS	Mutations	18%	
	MDM2	Amplification	5%	
	PIK3CA	Mutations	7%	
	PTEN	Deletion	1-7%	7% in gallbladder carcinoma
Gastric	EGFR	Mutations	3-5%	
	EGFR	Amplification	6%	
	ERBB2	Mutations	5-7%	
	ERBB2	Amplification	13%	
	ERBB3	Mutations	5-11%	
	ERBB3	Amplification	4%	
	FGFR1	Mutations	4%	
	FGFR2	Amplification	5%	
	KRAS	Mutations	6%	
	MET	Mutations	2%	
	MET	Amplification	4%	
	PIK3CA	Mutations/ Amplification	24%	42% and 72% in MSI-H and EBV+ gastric cancer, respectively
	PTEN	Mutations	4-8%	

Tumor type	Actionable genes	Alteration type	Frequency	Comments
	PTEN	Deletion	4%	
Melanoma	BRAF	Mutations	45%	
	CDKN2A	Deletion	13%	
	IDH1	Mutations	6%	
	KDR	Amplification	3%	
	KIT	Amplification	4%	
	MAP2K1	Mutations	5%	
	NF1	Mutations	14%	
	NRAS	Mutations	10-25%	
	PDGFRA	Amplification	3%	
Breast	11q	Amplification	15%	
	AKT1	Mutations	2-4%	
	CDKN2A	Deletion	3-4%	
	ERBB2	Mutations/ Amplification	13%	
	ESR1	Mutations	10%	ER+ breast cancer, metastatic samples and not primary (marker of resistance to antiestrogen therapy)
	FGFR1	Amplification	10-15%	
	FGFR2	Amplification	4%	
	MAP2K4	Mutations	2-7%	
	MAP3K1	Mutations	4-13%	
	NF1	Mutations	2-4%	
	NTRK3	Fusion	92%	Secretory breast cancer
	PIK3CA	Mutations	9-45%	
	PIK3CA	Amplification	4-5%	
	PIK3R1	Mutations	2%	
	PTEN	Mutations/ Deletion	3-8%	
	RB1	Mutations/ Deletion	5-6%	Marker of resistance to CDK 4/6 inhibitors
Colorectal	AKT1	Mutations	1-6%	
	BRAF	Mutations	3-47%	47% in MSI-H colorectal cancer
	ERBB2	Mutations/ Amplification	6-13%	

Tumor type	Actionable genes	Alteration type	Frequency	Comments
	ERBB3	Mutations	4-20%	
	KRAS	Mutations	35%	
	NRAS	Mutations	10%	
	PIK3CA	Mutations	15-37%	37% MSI-H colorectal cancer
	PIK3R1	Mutations	2-17%	
	PTEN	Deletion	4-20%	20% MSI-H colorectal cancer
Ovarian	AKT1	Amplification	3%	
	AKT2	Amplification	2%	
	BRAF	Mutations	2-6% (low grade serous ovary)	Extremely rare in high grade ovarian cancer; 2-6% low grade serous ovarian cancer (excluding borderline tumors)
	BRCA1 (germline or somatic)	Mutations	9%	
	BRCA2 (germline or somatic)	Mutations	5%	
	CCND1	Amplification	20%	
	CDKN2A	Deletion	32%	
	FGFR1	Amplification	5%	
	KRAS	Mutations/ Amplification	19-33% (low grade serous ovary)	Extremely rare in high grade; 19-33% low grade serous ovarian cancer (excluding borderline tumors)
	NF1	Mutations/ Deletion	12%	
	NOTCH3	Mutations/ Amplification	11%	
	PIK3CA	Mutations/ Amplification	18%	
	PTEN	Mutations/ Deletion	7%	
Glioblastoma	BRAF	Mutations	2%	
	CDK4	Amplification	14%	
	CDK6	Amplification	2%	
	CDKN2A/B	Deletion	61%	
	EGFR	Mutations	17-21%	
	EGFR	Amplification	41-44%	
	FGFR1-TACC1	Fusion	NA	

Tumor type	Actionable genes	Alteration type	Frequency	Comments
	FGFR3-TACC3	Fusion	3-7%	
	IDH1	Mutations	5-12%	
	MDM2	Amplification	7%	
	MDM4	Amplification	8%	
	MET	Amplification	2%	
	NF1	Mutations	10%	
	NTRK1	Fusion	1%	
	PDGFRA	Amplification	10%	
	PIK3CA	Mutations/ Amplification	25%	
	PTEN	Mutations/ Deletion	41%	

ANNEXE 3. Breast carcinoma TNM prognostic stage group American Joint Committee on Cancer (AJCC) and Union for International Cancer Control (UICC) (AJCC Cancer Staging Manual 2017)

When TNM is...	And grade is...	And HER2 status is...	And ER status is...	And PR status is...	Then the clinical prognostic stage group is...
Tis N0 M0	Any	Any	Any	Any	0
T1* N0 M0 T0 N1mi M0 T1* N1mi M0	G1	Positive	Positive	Positive	IA
				Negative	IA
			Negative	Positive	IA
				Negative	IA
		Negative	Positive	Positive	IA
				Negative	IA
			Negative	Positive	IA
				Negative	IB
T1* N0 M0 T0 N1mi M0 T1* N1mi M0	G2	Positive	Positive	Positive	IA
				Negative	IA
			Negative	Positive	IA
				Negative	IA
		Negative	Positive	Positive	IA
				Negative	IA
			Negative	Positive	IA
				Negative	IB
T1* N0 M0 T0 N1mi M0 T1* N1mi M0	G3	Positive	Positive	Positive	IA
				Negative	IA
			Negative	Positive	IA
				Negative	IA
		Negative	Positive	Positive	IA
				Negative	IB
			Negative	Positive	IB
				Negative	IB
T0 N1¶ M0 T1* N1¶ M0 T2 N0 M0	G1	Positive	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIA
		Negative	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIA
T0 N1¶ M0 T1* N1¶ M0 T2 N0 M0	G2	Positive	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIA
		Negative	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIB
T0 N1¶ M0 T1* N1¶ M0 T2 N0 M0	G3	Positive	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIA
		Negative	Positive	Positive	IIA
				Negative	IIB
			Negative	Positive	IIB
				Negative	IIB

When TNM is...	And grade is...	And HER2 status is...	And ER status is...	And PR status is...	Then the clinical prognostic stage group is...
T2 N1[Δ] M0 T3 N0 M0	G1	Positive	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIB
		Negative	Positive	Positive	IIA
				Negative	IIB
			Negative	Positive	IIB
				Negative	IIB
T2 N1[Δ] M0 T3 N0 M0	G2	Positive	Positive	Positive	IB
				Negative	IIA
			Negative	Positive	IIA
				Negative	IIB
		Negative	Positive	Positive	IIA
				Negative	IIB
			Negative	Positive	IIB
				Negative	IIIB
T2 N1[Δ] M0 T3 N0 M0	G3	Positive	Positive	Positive	IB
				Negative	IIB
			Negative	Positive	IIB
				Negative	IIB
		Negative	Positive	Positive	IIB
				Negative	IIIA
			Negative	Positive	IIIA
				Negative	IIIB
T0 N2 M0 T1* N2 M0 T2 N2 M0 T3 N1[Δ] M0 T3 N2 M0	G1	Positive	Positive	Positive	IIA
				Negative	IIIA
			Negative	Positive	IIIA
				Negative	IIIA
		Negative	Positive	Positive	IIA
				Negative	IIIA
			Negative	Positive	IIIA
				Negative	IIIB
T0 N2 M0 T1* N2 M0 T2 N2 M0 T3 N1[Δ] M0 T3 N2 M0	G2	Positive	Positive	Positive	IIA
				Negative	IIIA
			Negative	Positive	IIIA
				Negative	IIIA
		Negative	Positive	Positive	IIA
				Negative	IIIA
			Negative	Positive	IIIA
				Negative	IIIB
T0 N2 M0 T1* N2 M0 T2 N2 M0 T3 N1[Δ] M0 T3 N2 M0	G3	Positive	Positive	Positive	IIB
				Negative	IIIA
			Negative	Positive	IIIA
				Negative	IIIA
		Negative	Positive	Positive	IIIA
				Negative	IIIB
			Negative	Positive	IIIB
				Negative	IIIC

When TNM is...	And grade is...	And HER2 status is...	And ER status is...	And PR status is...	Then the clinical prognostic stage group is...
Any T N3 M0				Negative	IIIB
		Negative	Positive	Positive	IIIB
				Negative	IIIB
			Negative	Positive	IIIB
				Negative	IIIC
T4 N0 M0 T4 N1[Δ] M0 T4 N2 M0 Any T N3 M0	G2	Positive	Positive	Positive	IIIA
				Negative	IIIB
			Negative	Positive	IIIB
				Negative	IIIB
		Negative	Positive	Positive	IIIB
				Negative	IIIB
			Negative	Positive	IIIB
				Negative	IIIC
T4 N0 M0 T4 N1[Δ] M0 T4 N2 M0 Any T N3 M0	G3	Positive	Positive	Positive	IIIB
				Negative	IIIB
			Negative	Positive	IIIB
				Negative	IIIB
		Negative	Positive	Positive	IIIB
				Negative	IIIC
			Negative	Positive	IIIC
				Negative	IIIC
Any T Any N M1	Any	Any	Any	Any	IV