Unifying Paradigms of Oncology

SESSION OBJECTIVES

1. Recall the definition of neoplasia.
2. Differentiate between “driver” & “passenger” mutations.
3. Explain the difference between tumor suppressors and (proto)oncogenes; provide examples and recognize their associated malignancies.
4. Discuss the “Hallmarks of cancer” and their role in the development of neoplasia
5. Describe the roles that inflammation, infection, and genomic instability play in enabling cells to become malignant using specific examples.
6. Describe the relationship between Lynch syndrome and microsatellite instability.

OPTIONAL PRE-CLASS MATERIALS FOR THIS SESSION:

- Skim the **section titles**, **bolded terms,** and **image captions** from Robbins & Kumar 11th edition: Chapter 6 to fill in any knowledge gaps you need.

A brief forward: Many thanks to Dr. Natasha Hunter, MD for her substantial contributions to this chapter.

OVERVIEW:

For as long as humans have existed, cancer has co-existed with us. Genetic mutations accumulate as we age, so as humans have improved their hygiene and nutrition, and gained better control over infectious diseases, cancer has emerged as a more prominent health issue. Today, cancer is the second leading cause of death in the United States, and a leading cause of death worldwide.

In this course, we will provide a basic framework in which to understand how cancer develops, and how our understanding of cancer fuels the development of the tools used to combat it. As you move through the Foundations curriculum, you will encounter cancers that develop in almost every organ and system you study. We hope the concepts you grasp in this introductory course will give you a roadmap for delving deeper into specific oncologic conditions.

KEY CONCEPTS IN CANCER GENOMICS:

A brief review of the fundamental processes of cellular biochemistry provides a framework for understanding the molecular alterations that can occur in human cancers.

DNA Replication and Repair: DNA replication is a highly regulated process that copies DNA from one cellular generation to the next. Multiple safeguards, including proofreading and mismatch repair functions, ensure that an error-free copy of the genomic DNA is generated with each cell division. Cancer-related genomic changes often stem from failures of these safeguards, including mutations in DNA polymerases *(e.g. POLD1* and *POLE)*,

genes involved in mismatch repair (*MLH1*, *MSH2*, *MSH6*, *PMS1*, and *PMS2*), and genes involved in double-strand break repair (e.g. *BRCA1*, *BRCA2*, *CHEK2*, and *ATM*).

Gene Structure and Mutations: Genes have a number of essential components, include exons, intervening introns, 5' and 3' untranslated regions (UTRs), promoters, and enhancers. Mutations in exons (missense or nonsense) lead to abnormal or missing protein products, while mutations outside of exons can have profound effects on gene expression by altering gene transcription or effects on the transcribed mRNA itself.

mRNA Processing and Transcription: DNA transcription generates an mRNA version of the gene. This primary mRNA molecule is further processed to generate a mature mRNA molecule, a step that may be disrupted in cancer, leading to aberrant mRNA splicing, to mutations in mRNA that are not present in the DNA template, or to changes in mRNA stability.

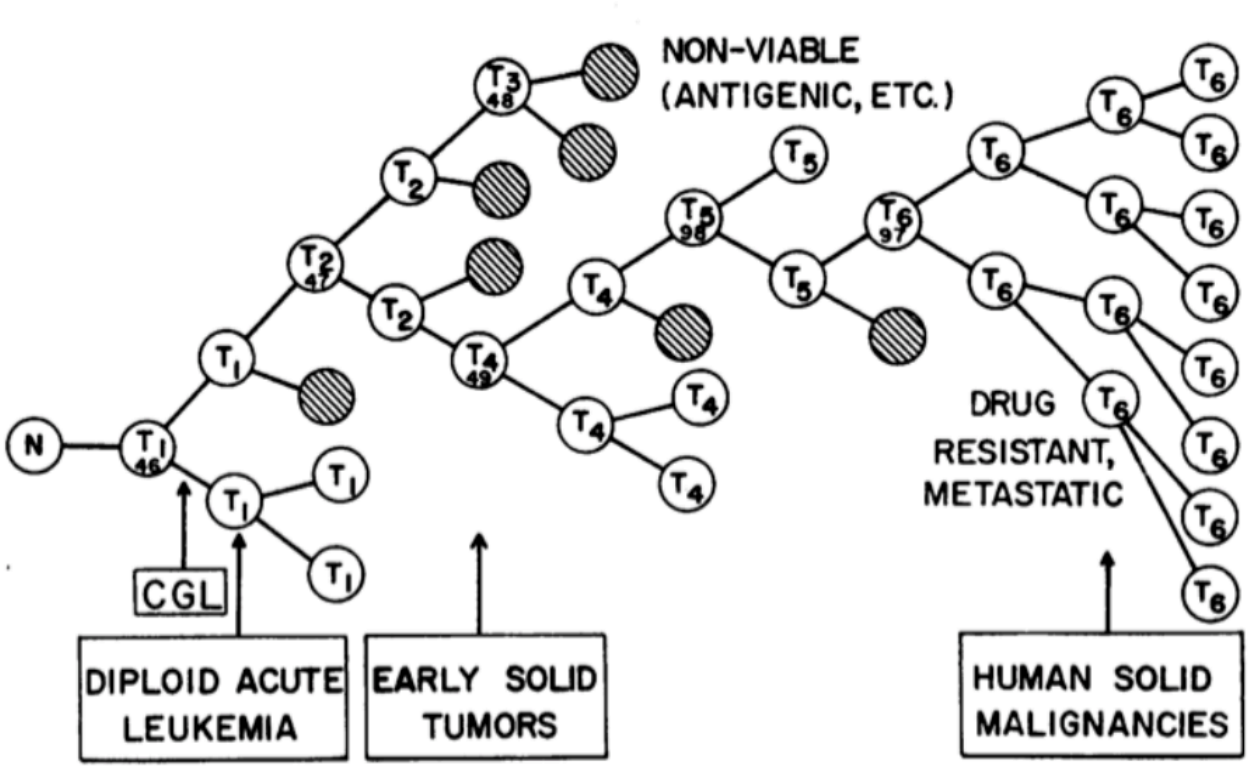
Protein Translation and Modification: Translation converts the mature mRNA into a protein product, which then is folded into a mature protein. Post-translational protein modifications, including phosphorylation, acetylation, and ubiquitination, provide further modulation of protein function, by affecting activity, stability, cellular localization, regulated turnover, and other processes.

Beyond Coding DNA and Epigenetics: While we tend to focus on the mutations within genes as the main drivers of cancer, the vast majority of genomic DNA does not code for genes. Mutations outside of genes can drive cancer as well. Small non-coding RNAs, such as microRNAs (miRNA), are able to regulate gene networks and are often altered in human cancer. Whole genome changes, including DNA methylation and histone modification, can likewise alter whole networks of gene expression. The term epigenetics is often used to refer to such changes in DNA that alter gene function without causing changes in the nucleotide sequence of DNA itself.

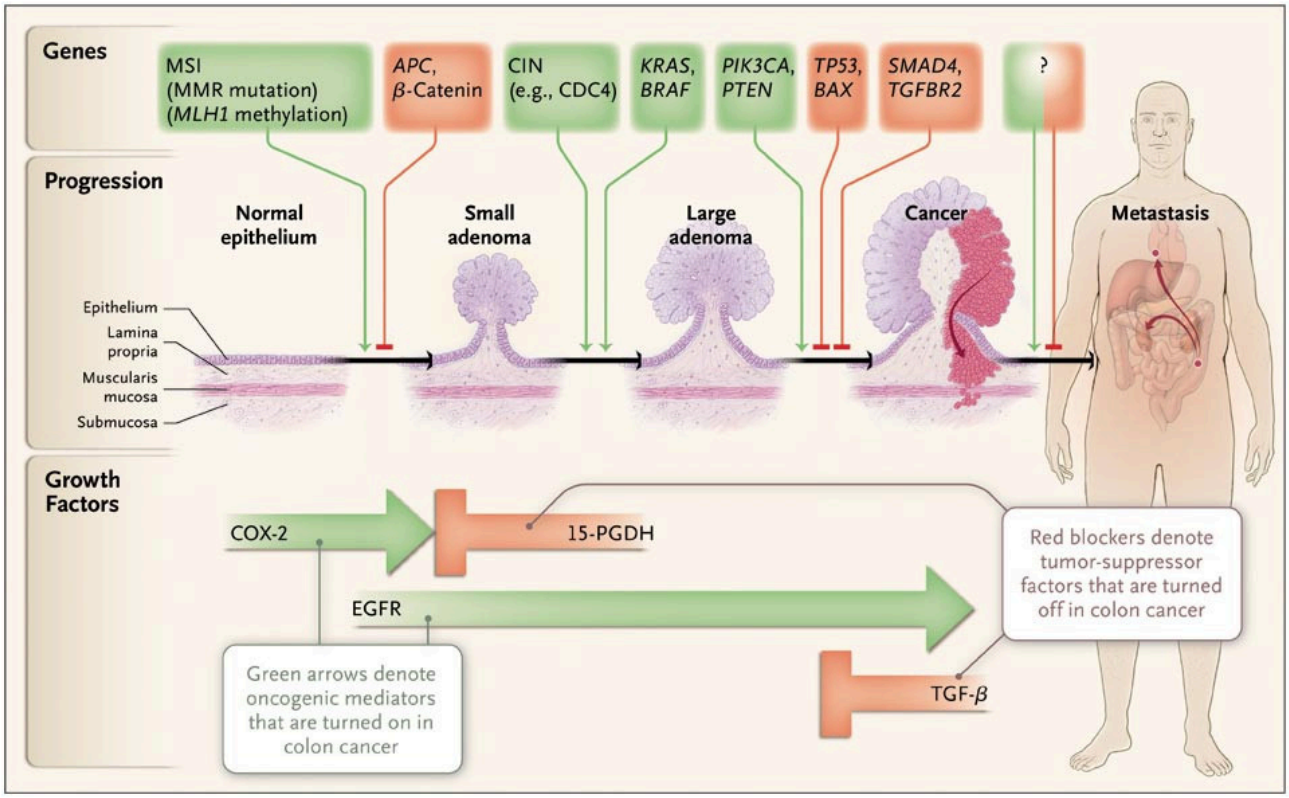
Cell-Cell Interactions and Immunotherapy: Finally, it is important to consider how whole cells interact with each other. Cancer cells are intimately associated with their microenvironment and crosstalk between cell types can dictate cancer growth and spread. An increasingly important aspect of cancer biology is the interaction of the immune system with cancer, in particular how cancers are able to avoid detection and destruction by the immune system. As we learn more about this process, powerful new "immunotherapy" cancer treatments have been developed and have already had a significant impact on patient treatment.

TUMOR EVOLUTION AND GENETIC DRIVERS OF CANCER PROGRESSION:

Although single molecular events, such as the t(9;22) "Philadelphia Chromosome" translocation of chronic myelogenous leukemia, are rarely sufficient to cause cancer, most cancers arise from multiple genetic changes. These alterations provide cells with progressive selective advantages, driving them along the neoplasia continuum: from normal to premalignant, invasive, metastatic, and ultimately drug resistant. Modern molecular techniques suggest that a surprisingly small number of genetic changes are required to drive cancer development (generally less than 10); however, ***it is important to remember that there is no single path of genetic mutations that lead to cancer, even within a single tissue type.*** The sequence of events that can transform normal cells into cancer cells is often depicted as an evolutionary tree (Nowell, Science, 1976, PMID 959840), as depicted below.



The molecular events driving colorectal cancer evolution, gathered from biopsy and resections of lesions at various stages of development (premalignant adenomas, malignant but localized carcinomas, and metastatic carcinomas), serve as a model of this concept. A **gatekeeper mutation** refers to mutations that are universally required to initiate cancer in a specific tissue. In colorectal cancer, mutations in the *APC* genes or mismatch repair genes are the most common gatekeeper mutations. Subsequent mutations in other cancer genes, such as *TP53*, *KRAS*, and *FBXW7*, accumulate as tumors progress. Diagrams of this process are commonly called "**Vogelgrams**", in reference to Bert Vogelstein and his lab, who published the first of these in 1996. An updated review of this topic can be found in Markowitz, NEJM, 2009 (PMID 20018966). Nowell's proposed process of tumor evolution has now been validated in numerous human cancers, through whole genome sequencing of primary and relapsed leukemia (see, for example, PMID 22237025) and of primary and metastatic prostate cancer (PMID 26855148).

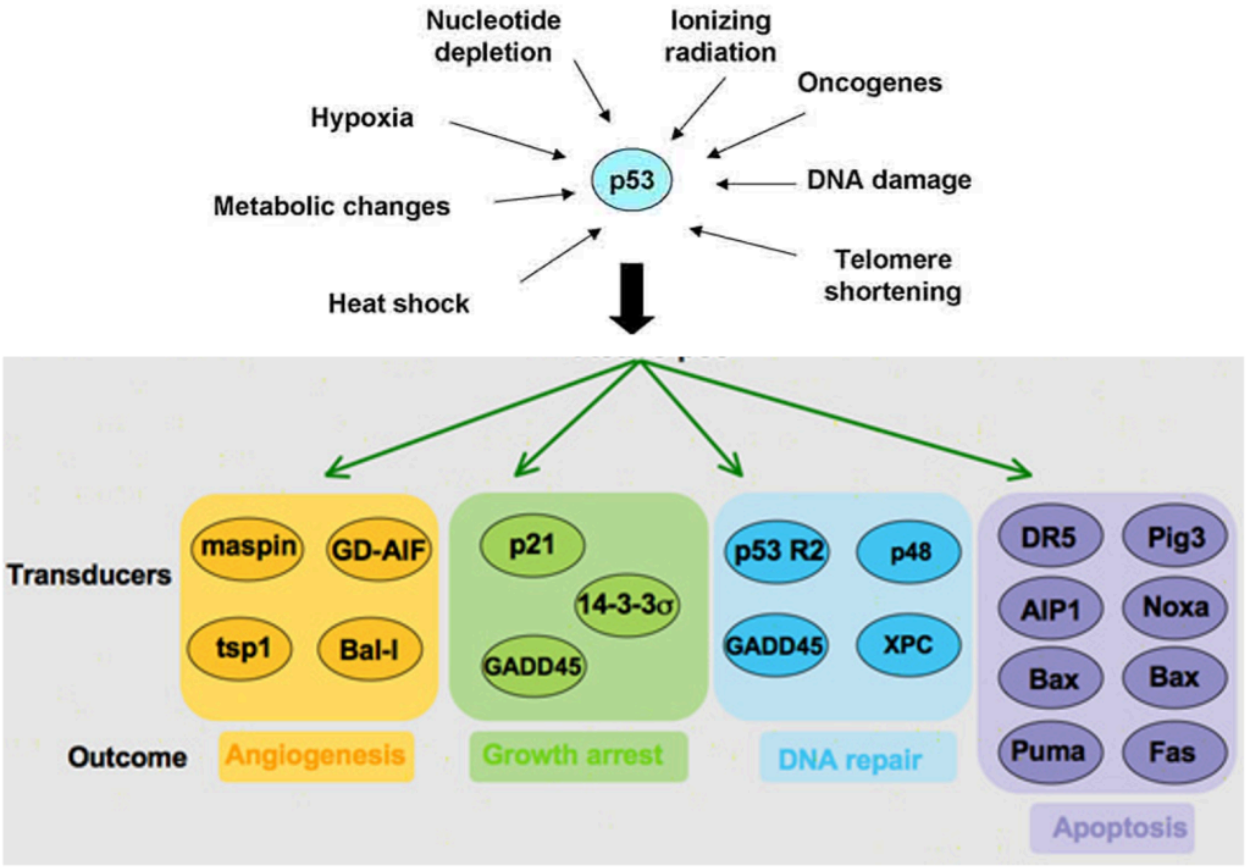


TUMOR SUPPRESSORS AND ONCOGENES:

Although external agents such as viruses were once thought to be the main cause of cancer, it is now clear that the majority of cancers are caused by mutation or deregulation of endogenous cellular genes that normally control fundamental cellular processes. Genes implicated in cancer development are generally categorized as either tumor suppressor genes or oncogenes. A major goal of cancer research is to determine the function of oncogenes and tumor suppressors and to use this knowledge to develop new targeted therapies for cancer patients.

Tumor Suppressor Genes (TSGs) are involved in preventing the development of cancer and are often represented as the "brake pedal" guarding cells from transformation. They do so by maintaining the integrity of the genome (e.g., *TP53*), regulating the cellular response to external stimuli (e.g., *PTEN*), inhibiting proliferation or cell division (e.g, *CDKN2A* or p16), and promoting apoptosis (e.g., TP53). TSGs are often targeted in hereditary cancer syndromes, where individuals are born with single allele mutations that predispose to cancer. Many well-known TSGs were first identified through the study of cancer-prone families. Now, TSGs are identified through large-scale sequencing efforts such as the Cancer Genome Atlas (further described below). These studies have revealed TSGs in unexpected pathways, such as in RNA splicing machinery (e.g., *SF3B1*) or factors that control chromatin dynamics (e.g., *ARID1A*). In sporadic (non-hereditary) cancers, TSGs are frequently inactivated through missense or nonsense mutation or through gene deletion.

- *TP53* (aka p53) is the most commonly mutated TSG across all human tumors. Often called "***the guardian of the genome***", the p53 protein is the main cellular defense against a diverse array of potentially damaging insults. These can be extrinsic insults, such as DNA damaging radiation or chemicals, or intrinsic processes, such as abnormal growth factor signaling. Such insults trigger cellular pathways that converge to activate p53, which, in turn, causes a transcriptional program that either blocks the cell cycle (to allow repair of DNA damage) or, in more extreme cases, promotes cell death.



Oncogenes are altered versions of cellular genes that become overactive and subsequently promote or accelerate cancer. Oncogenes can be thought of as a "gas pedal" that drives cancer. These genes are often involved in growth factor signaling (e.g., *ERBB2*, Her2/Neu), cell cycle control (e.g., *CCND1*, cyclin D), gene transcription (e.g., MYC), anti-apoptosis (e.g., *BCL2*), and migration/metastasis (e.g., *SNAIL*). These genes are commonly over-expressed in cancers through amplifications, translocations, or activating mutations.

HALLMARKS OF CANCER:

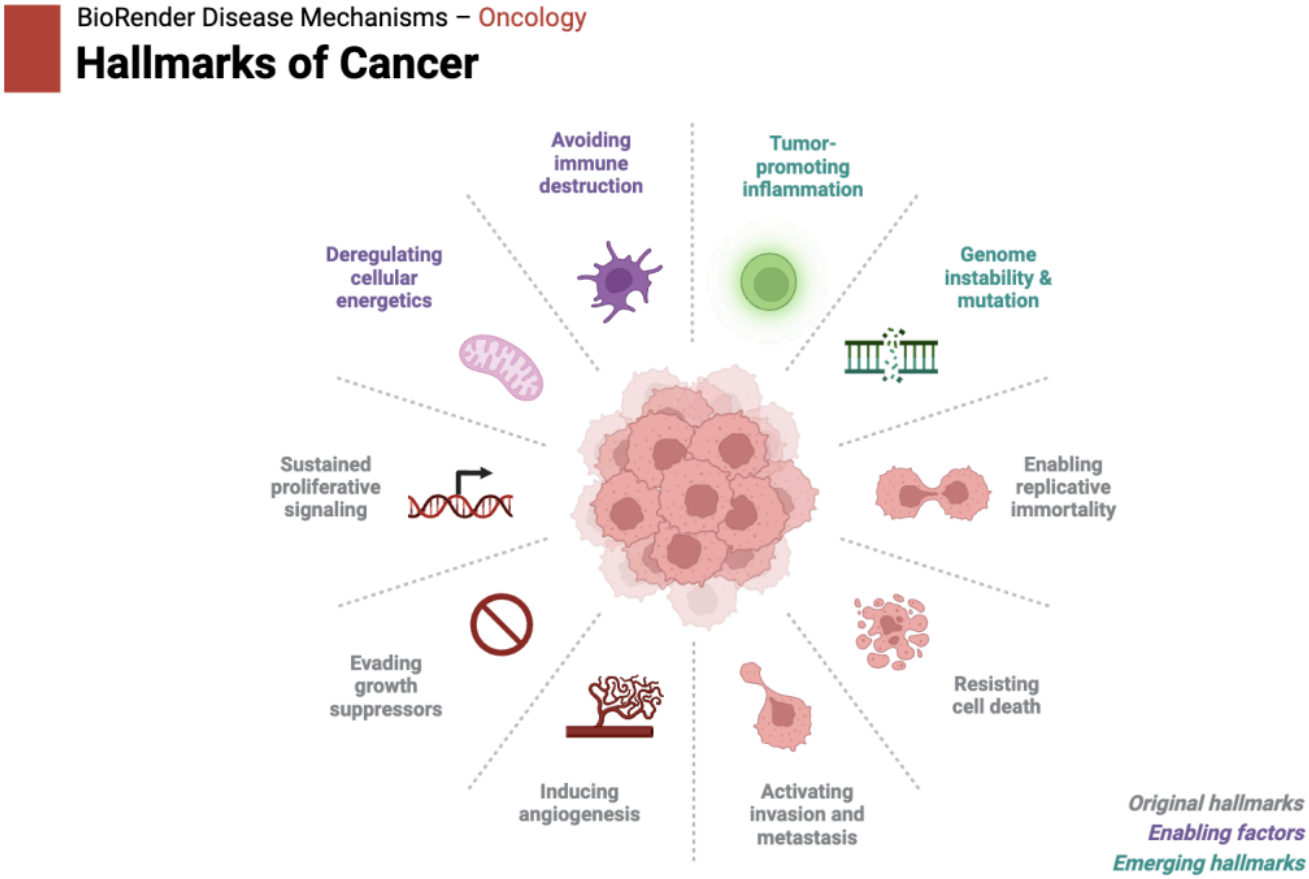
Over the last 20 years, our increased knowledge of cancer genetics and genomics has yielded incredibly large datasets that are daunting in their complexity. Too often, this data acquisition has helped little to answer a fundamental question: what are the essential molecular and biological features of cancer? Hanahan and Weinberg addressed this paradox head-on in two seminal publications in the journal Cell (*The hallmarks of cancer*, 2001, PMID 10647931; and *Hallmarks of cancer: the next generation*, 2011, PMID 21376230). In the first paper, the authors suggest that the vast number of mutated cancer genes can be better understood if put into a framework of molecular processes. In their words:

> "We suggest that the vast catalog of cancer cell genotypes is a manifestation of six essential alterations in cell physiology that collectively dictate malignant growth: ***self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis***."

As normal cells transform to cancer cells, they must acquire most, if not all, of these characteristics. The order in

which they are acquired is not important and may in fact be stochastic or arbitrary. These authors go through these six cancer hallmarks one by one, giving specific examples of molecules and pathways that contribute to each hallmark, then suggesting how each could be exploited therapeutically.

Roughly 10 years later, they published an update. This second paper shows the extensive experimental progress made in validating the 6 original hallmarks. It also defines 2 "emerging" hallmarks as well as two enabling characteristics that help cells to acquire the previously defined hallmarks. Finally, it gives an update on the progress made in developing targeting therapies to each of these pathways. Together, these papers provide an incredibly useful framework to consider the biology of cancer development and metastasis.



- **Sustained Proliferative Signaling**: Cancer cells maintain chronic proliferation by sustaining growth signals, such as through mutations in oncogenes like *KRAS* or *ERBB2*. These mutations cause cells to keep dividing uncontrollably.
- **Evading Growth Suppressors**: Tumor suppressor genes, like *TP53* and *RB1*, normally prevent excessive cell growth. In cancers, mutations in these genes allow cells to bypass growth-inhibitory signals, promoting uncontrolled division.
- **Resisting Cell Death (Apoptosis)**: Cancer cells avoid programmed cell death (apoptosis), enabling their survival. The over-expression of anti-apoptotic proteins like *BCL-2* or mutations in *TP53* are common examples of this hallmark.
- **Enabling Replicative Immortality**: Normal cells can divide a limited number of times due to the shortening of telomeres. Cancer cells bypass this limit by activating telomerase (*TERT*), which maintains telomere length, allowing indefinite replication.
- **Inducing Angiogenesis**: Tumors stimulate the growth of new blood vessels to ensure a supply of oxygen and nutrients. A key regulator of this process is *VHL* (von Hippel-Lindau tumor suppressor gene), which controls hypoxia-induced factors like *HIF-1α*. Mutations in *VHL* lead to uncontrolled blood vessel formation,

often seen in renal cell carcinomas.

- **Activating Invasion and Metastasis**: Cancer cells gain the ability to invade neighboring tissues and metastasize to distant sites. Mutations in genes like E-cadherin (*CDH1*) can disrupt cell adhesion, promoting invasion and metastatic spread.
- **Deregulating Cellular Energetics**: Cancer cells often switch to a metabolic process known as aerobic glycolysis (the Warburg effect), allowing them to rapidly generate energy even in the presence of oxygen. This altered metabolism supports their rapid growth.
- **Avoiding Immune Destruction**: Cancer cells develop mechanisms to evade the immune system. For example, they can up-regulate immune checkpoint proteins like PD-L1, which inhibit T-cell activation and allow tumors to evade immune surveillance.
- **Genome Instability and Mutation**: Defects in DNA repair mechanisms, such as in MMR genes (e.g., *MLH1*, *MSH2*, *MSH6*, *PMS2*) and BRCA1/2, can result in the accumulation of genetic changes that drive cancer progression.
- **Tumor-Promoting Inflammation**: Chronic inflammation can promote tumor development by providing growth factors, pro-angiogenic factors, and extracellular matrix-modifying enzymes.

PREPARING FOR DIAGNOSTICS AND MULTI-DISCIPLINARY CARE

This section provides an overview of various molecular tests commonly used in pathology and oncology. While you won't be tested on this material, it may help enhance your understanding and prepare you for the *Path/Onc: Diagnostics, Assay Selection, and Multidisciplinary Care* lecture. Familiarizing yourself with these concepts will give you a strong foundation for discussing the selection of diagnostic tests and the roles of the clinical team in optimizing patient outcomes.

ONCOGENOMICS

Oncogenomics is a term used to refer to the in-depth molecular characterization of cancer tissues (cell lines grown in culture or tumor tissues harvested from patients). The methods used for oncogenomics are expanding every year, but generally include high-throughput analysis of DNA, RNA, and protein. Some of the common technologies used in these studies are outlined below.

1) Genomic DNA Analysis (Genomics): These techniques are used to identify single base pair mutations, genetic translocations (where two pieces of separate chromosomes have swapped places), and copy number variations (including amplifications or deletions of sections of chromosomes). They can also be used to look at epigenetic regulators of gene expression, such as methylation.

- ***Targeted sequencing/Panel Testing*** is the most common approach used to analyze specific genes or regions of interest in the genome. It allows for efficient, cost-effective identification of mutations in clinically relevant genes, but can miss important findings that are "off target."
- ***Exome sequencing*** is most useful for identification of mutations within coding regions or in nearby regulatory regions (i.e. splice sites). Because the exome is <5% of the genome, it misses a huge amount of genetic information.
- ***Whole genome sequencing (WGS)*** is less commonly used. As the name implies, this technique provides a comprehensive sequence of the entire genome, not just the exons. This remains very expensive and is much more difficult to analyze statistically. Most large-scale genome sequencing studies perform WGS on a subset of samples (~10%).
- ***Copy number variation (CNV) Analysis:*** CNVs can be studied in a variety of ways. Older low-resolution techniques include standard karyotyping (chromosomal spreads) or conventional cytogenetics. These

have largely been replaced by microarray technologies such as array comparative genomic hybridization (which requires a normal comparator) or single nucleotide polymorphism (SNP) arrays (which analyze DNA by assessing allele frequency patterns). These high resolution approaches allow more precise identification of specific genes or regions that may be affected by copy number gains and losses.

- ***Genome Methylation*** is a regulated process that results in gene silencing. Genome wide methylation patterns are often altered in cancers. Genome methylation arrays can identify areas of the genome that are hypo- or hyper-methylated. Whole genome methylation can be studied with a modified DNA sequencing procedure (Whole Genome Bisulfite Sequencing).

2) RNA Analysis (Transcriptomics): Changes in genomic DNA can cause mutations in protein coding regions, which ultimately leads to the production of abnormal mutant proteins. However, it can also cause alterations in overall levels of gene expression. These changes in gene expression can drive cancer phenotypes, are often characteristic of specific cancer subtypes, and can even dictate cancer prognosis.

- ***Gene Expression Arrays*** are similar to arrays outlined above, but measure levels of specific mRNAs in the cell. In this case, mRNA (rather than genomic DNA) is hybridized to the array (after reverse transcription to create cDNA) allowing measurement of gene expression across the entire transcriptome. These types of assays are sometimes used clinically to assess risk of recurrence for cancers treated with curative intent (e.g. Oncotype DX or Mammaprint). Results of such assays can influence decisions for chemotherapy by identifying patients at low or high risk of recurrence.
- ***RNA-sequencing (aka RNA-Seq)*** is a more precise method to measure gene expression. Based on sequencing of mRNA transcripts (again, after cDNA conversion), RNA-seq quantifies gene expression levels, detects allele-specific expression, and relative abundance of alternatively spliced transcripts. It also enables detection of gene fusions resulting from chromosomal translocations (including the precise identity of the precise transcript).

3) Protein Analysis (Proteomics): Most genomic/transcriptomic changes must ultimately alter protein abundance or function to drive cancer growth. Gene mutations within exons can lead to the production of mutant proteins. However, alterations in the level of wild-type cellular proteins can come about through genomic amplifications, deletions, and translocations. Thus, the precise measurement of proteins is essential to understanding cancer. Unfortunately, while some moderate throughput approaches exist, cancer proteomics remains suboptimal, especially for the study of fresh tumor samples. Additionally, the wide range of post-translational modifications (phosphorylation, acetylation, etc.) that alter protein function remain difficult to study.

- ***Mass Spectrometry (MS)*** can identify specific proteins within complex protein mixtures and is suitable for analysis of many post-translational modifications. Total protein is isolated from tumor tissue, digested into peptides, then “shot” through the mass spectrometer to identify and quantify individual peptides. Issues with sample preparation, reproducibility, and limits of detection have hampered this technology, although there is much enthusiasm for MS-based methods to study serum and other fluids to allow the early detection of cancer.
- ***Reverse Phase Protein Arrays (RPPA)*** are another variation on the microarray where antibodies recognizing specific proteins are arrayed on slides. Solubilized protein from tumors are then hybridized to these arrays to look at the relative expression levels of a moderate number of different proteins (hundreds rather than thousands). This technique is stymied by the limited availability of suitable antibodies to individual protein targets and by non-specific background binding; it is currently best suited to examining the activation of specific pathways (such as kinase networks) across cancer samples.

ONCOGENOMICS: The Cancer Genome Atlas (TCGA)

TCGA is an ongoing, large-scale effort to interrogate the genomics, transcriptomics, and proteomics of over 30 different human cancers, both common (breast, lung, prostate, and colorectal cancers) and uncommon (uveal melanoma, cholangiocarcinoma). It is a direct descendent of the Human Genome Project and in many ways mirrors and expands on that previous effort. The central goals of TCGA are to map the shared and unique molecular features of individual tumor types to better understand cancer biology; to define prognostic subgroups within pathologically similar cancers; to identify new and cooperating components of genetic pathways; and to identify specific vulnerabilities that may be used to inform cancer treatments. Each individual TCGA study collects data from genome sequencing, CNV, epigenetics, gene expression, and, more recently, proteomics to gain a comprehensive snapshot of that particular tumor. TCGA datasets include analysis of large numbers of tumors (hundreds) to provide adequate statistical power to identify true cancer drivers, and TCGA publications generally include hundreds of pages of data. A list of primary publications from the TCGA network can be found at http://cancergenome.nih.gov/publications. After initial publication, all TCGA data is publicly available, so that researchers worldwide can perform analyses of their own. Furthermore, pooled TCGA datasets across tumor types allow "meta-analyses" (PanCanAtlas) that have been helpful in identifying rare but universal cancer drivers and in pinpointing specific genes targeted by recurrent CNVs.

Some examples of new knowledge gained from TCGA includes:

- The mutation rate varies considerably across tumor types.
- A finite number of cancer driver genes exist and we have likely identified the majority.
- Mutations in DNA polymerase genes (POLD1/POLE) can lead to very high mutation rates.
- Unexpected genes and processes are recurrently altered in cancer, such as RNA splicing, RNA editing, and genome remodeling factors.
- Tumors arising from different tissues can look very similar at the genetic level (serous ovarian, serous endometrial, and basal breast cancers).
- Tumors caused by viruses have distinct genetic features compared to their non-virally mediated counterparts (e.g., HPV+ and HPV- squamous cell cancers).

THE IMPACT OF MOLECULAR DIAGNOSTICS:

The knowledge gained and the approaches taken in TCGA will profoundly impact how we treat cancer patients in the near future. Already, molecular techniques are being used to guide patient therapies in real time and adaptive clinical trial designs allow rapid amendments to add molecularly defined patient subgroups.

This Chapter's PDF

LINK

- Note: The interactive features of this chapter are not reproducible in this PDF format.