Basic Molecular Pathology

(for surgical pathologists)

Prepared by Dr. Kurt Schaberg

Last updated: 9/15/2024

A brief word from Dr. Juan Rosai

From, Rosai and Ackerman's Surgical pathology, 10th edition

"In this time and age, the overriding challenge comes from molecular biology. It will not make morphology disappear, as some of its more strident apologists have announced in apocalyptic terms, but it will certainly change dramatically the way we practice the specialty. Actually, it has happened already. It would be a mistake for surgical pathologists to not embrace it, as we have done before with other special techniques. It will not be as natural as it was for electron microscopy or immunohistochemistry, with, after all, a retained morphologic basis. Molecular biology is a seemingly different world, with a different language and pictorial representation. Yet the link is there… Future discovers will not be likely made by morphologists ignorant of molecular biologic findings or biologists scornful of morphologic data but by those willing and capable of integrating them through a team approach."

Use in Practice

Molecular information helps us and the patient in 2 main ways:

1) Neoplasm classification

Increasingly, it is being recognized that many tumors have characteristic underlying molecular alterations. In some instances, these alterations are becoming even definitional. This is particularly true in the CNS. For example, while oligodendroglioma was previously defined solely by morphology, it is now further defined by IDH mutation and 1p/19q-codeletion.

2) Therapy selection (and prognosis)

Precision Oncology is the molecular profiling of tumors to identify targetable alterations and select specific cancer treatments.

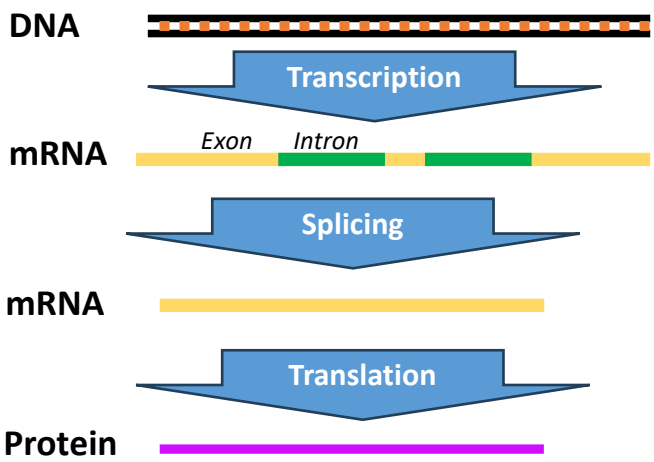
The prototypical example of this is treatment of chronic myelogenous leukemia (CML), which has a characteristic BCR-ABL rearrangement, with the receptor tyrosine kinase small molecule inhibitor imatinib. This has transformed CML from a death sentence to a mostly chronic disease.

Unfortunately, this success of a single type of precision targeting agent achieving near-universal durable benefit is mostly unparalleled in solid tumor oncology (with a few notable exceptions), at least so far. This is likely because most solid tumors (which are often currently only treated with expensive "precision" therapies in the advanced, metastatic setting) often have additional co-driver and passenger mutations (which are not present in normal, early CML). So, while these therapies achieve some "response" in solid tumors, it's often relatively temporary and partial, at least for now.

Tumor-agnostic therapy—treatment based on molecular testing/biomarkers rather than tissue-type/histology ("*pancancer*").
Main examples: 1) *Immunotherapy* like the PD-1 antibody pembrolizumab, which is approved in all, PD-L1 IHC +, mismatch repair-deficient (dMMR), and Tumor Mutation Burden high (TMB-high) tumors.
2) *Gene-targeted agents* (like NTRK, BRAF, RET, HER2 directed therapy)

This seems to work for some tumors/biomarkers, but not others. For example, NTRK inhibitors seem to work in a wide variety of tumors, while BRAF inhibitors seem to work in BRAF-mutated melanoma and hairy cell leukemia, but not BRAF-mutated colorectal cancer (due to EGFR alternate pathway activation).

The Very Basics: DNA→ RNA → Protein



The full details of genetics are beyond the scope of these notes, but to dramatically oversimplify:

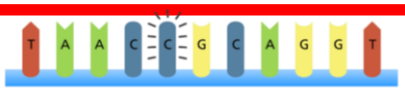
Deoxyribonucleic acid (DNA) encodes through its sequence of 4 different base pairs the genetic information of the body.

This is then transcribed into messenger ribonucleic acid (mRNA), which undergoes splicing to remove the non-coding introns.

This is then translated into protein, a polypeptide chain composed of amino acids.

Each step involves complexes of enzymes that work together in coordination.

Types of Genetic Changes



Mutation/Variant—***any* change** in the **DNA sequence** of a cell (can be neutral, good, or bad).
Often, mutation is used in "somatic" setting and "variant" in germline setting.
Somatic mutations—*acquired* mutations that are not inherited or passed on. Mostly studied in tumors.
Germline variations—*heritable* sequence changes that are in germ cells and passed between generations.
A variant that is expressed in at least 1-2% of the population is called a *polymorphism*.

Types of mutations:
Missense mutation: change in 1 DNA base pair → change in 1 amino acid (may or may not alter function)

Synonymous mutations: result in no change in the amino acid sequence.

Nonsense mutation: change in 1 DNA base pair → changes 1 amino acid to a stop codon → premature stoppage of translation → short protein (often non-functional)

Insertion (ins): addition of at least one base to DNA sequence

Deletion (del): removal of at least one DNA base pair (and sometimes entire sections/genes)

Duplication (dup): section of DNA copied several times

Frameshift: Addition or loss of DNA base pair→ alters reading frame → alters subsequently translated amino acids and resulting protein (can also cause premature stop) → often non-functional protein

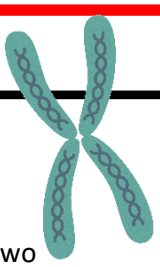
Repeat expansion: expansion in the number of times a DNA sequence is repeated

Conversion (con): Sequence change where nucleotides are replaced by a sequence from elsewhere in the genome

Deletion/Insertion (Indel): sequence where ≥1 nucleotide is replaced by ≥1 nucleotide (essentially a big substitution).

Important to keep in mind: Non-coding mutations can also have implications by altering mRNA splicing, promoter function, and miRNA binding sites.

Types of Genetic Changes *(continued)*



Structural variants

Large changes (often >50Mb) at the subchromosomal level.

Examples: Translocations, Deletions, Insertions, Inversions, Gene amplification.

Translocation: Chromosomal rearrangement between parts of nonhomologous chromosomes. The two chromosomes break and the fragments rejoin at breakpoints.

Transposition: nucleotides move from one position to another in the genome (deletion from one point with an insertion at another point)

Inversion (inv): a fragment of a chromosome breaks off and reattaches in the opposite direction (flips)

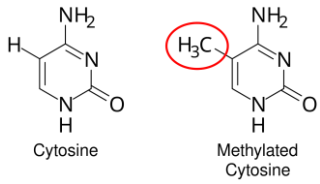
Ploidy: gain or loss of the *whole* complement of chromosomes (e.g., hyperdiploid, hypodiploid)

Aneuploidy: the presence of an abnormal number of chromosomes in a cell that is not a multiple of the haploid state (23 in humans). Present in the majority (~70%) of cancers.

Epigenetics

Heritable traits, or a stable change of cell function, that happen *without changes to the DNA sequence*

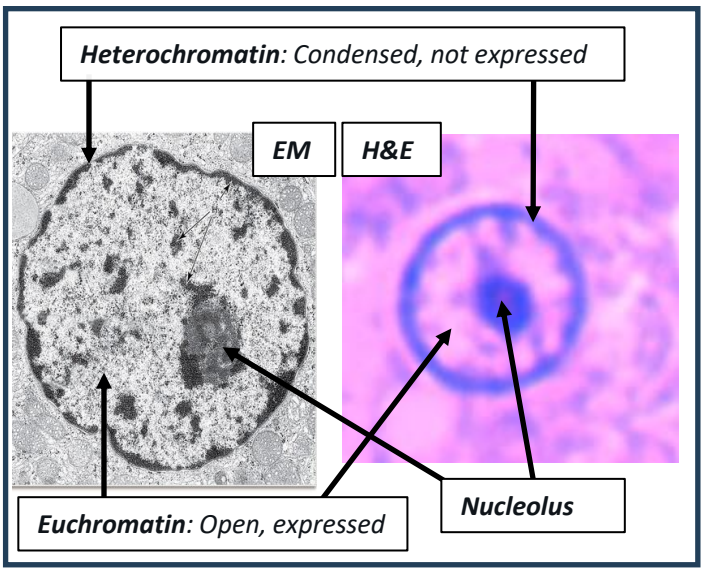
DNA methylation: addition of a methyl group to cytosine.
Occurs in Cytosine-guanine-rich areas (CpG islands).

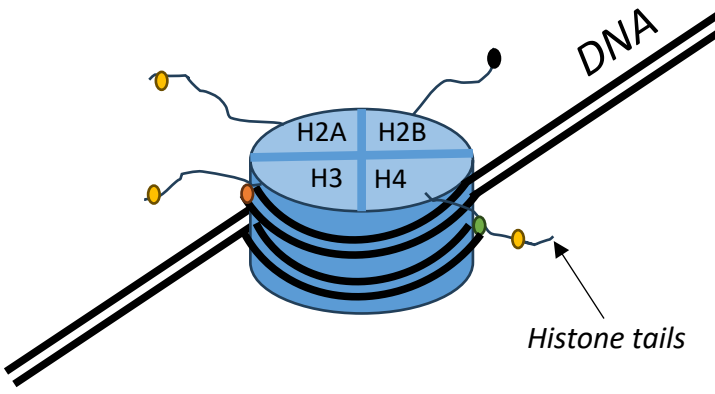


Histone post-translational modifications: modifications of specific amino acids on the tails of histone proteins (H2A, H2B, H3, H4). Depending on the modifications, the associated DNA will form closed, non-expressed *heterochromatin*, or open, actively transcribed, *euchromatin*.

Complex with lots of interactions. For example:
Acetylation→ lowers positive charge on histones→ decreased binding to negatively charged DNA → Easier to access for transcription factors and RNA polymerase → more transcription.

Methylation→ attracts enzymes that further methylate DNA → decreased gene expression





How mutations cause tumors

Tumorigenesis is often a *multistep* process, where cells acquire successive mutations.

Two main carcinogenic gene changes:
1) Increased protooncogene function *(Think: A car gas pedal getting stuck down → acceleration)*
2) Loss of tumor suppressor function *(Think: Cutting the brake lines on a car → can't stop)*

→ Result in **uncontrolled cell growth** and **failure to respond to normal signals** that lead to cell death

Oncogenes

Protooncogenes are genes involved in *normal* cell growth.
Oncogenes are the *mutated* version of a protooncogene → increase expression or activity of protooncogene through "gain of function" → increased growth (cell division) and inhibition of cell death. Mostly transcription factors.

Only need one copy to cause disease state→ act in a dominant fashion.
If a tumor is profoundly dependent on a specific mutation/protein it is called an "oncogene addiction" and is a great target for therapy (e.g., BCR-ABL with imatinib)

Common examples:
Point mutations, deletions, insertions → increased transcription
Amplification → increased transcriptions
Translocation brings protooncogene to a site with increased expression → increased transcription

Tumor Suppressors

Tumor suppressors encode proteins the *repress the cell cycle* (growth) and *promote apoptosis* (cell death) → "loss of function" due to mutation → no longer inhibits uncontrolled growth.

Since we're diploid (have two copies of every gene), you need "**Two Hits**" to deactivate *both* copies (act in a recessive fashion).

Common tumor suppressor roles:
DNA repair enzymes
Repression of genes essential for continuation of cell cycle.
Coupling of cell cycle to presence of DNA damage with proapoptotic signaling if DNA damage cannot be repaired
Cell adhesion molecules that help prevent metastasis

Driver vs Passenger Mutations

Occur *simultaneously* in the *same tumor* genome.

Driver mutations—mutation(s) that alter the function of cancer genes and directly contribute to development and progression. Tightly clustered within cancer genes. *More* relied upon by cancer cells. Often a small subset of overall mutations and often mutually exclusive. Provide a selective advantage. More often targets of therapy. Examples: EGFR, EML4-ALK

Passenger mutations—mutations that are *less* important for growth and survival (more neutral). Often more numerous, random, sprinkled throughout the genome. While neutral at first, they may provide a selective advantage in the setting of therapy.

Genetics → Morphology

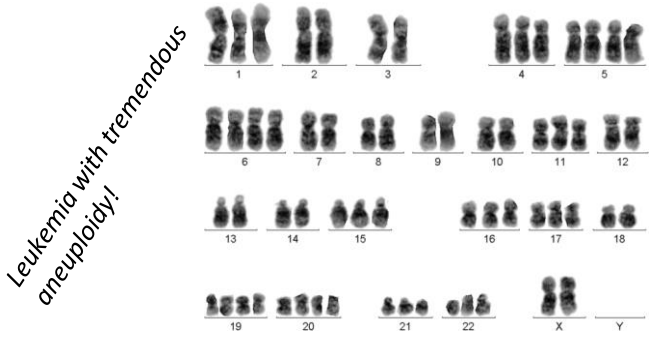
Broadly speaking, one can get a *hint* as to the genetics of a tumor by the basic morphology.
The more genetic changes a tumor has, often the more pleomorphic, heterogeneous, and "atypical" it will look. In contrast, tumors with a recurrent driver mutation/fusions often appear relatively bland and homogeneous/monotonous, with all cells having similar genetics.

"Complex" Genetics

Overall, **More common.**

Examples: Undifferentiated pleomorphic sarcoma, Myxofibrosarcoma, High-grade carcinoma.

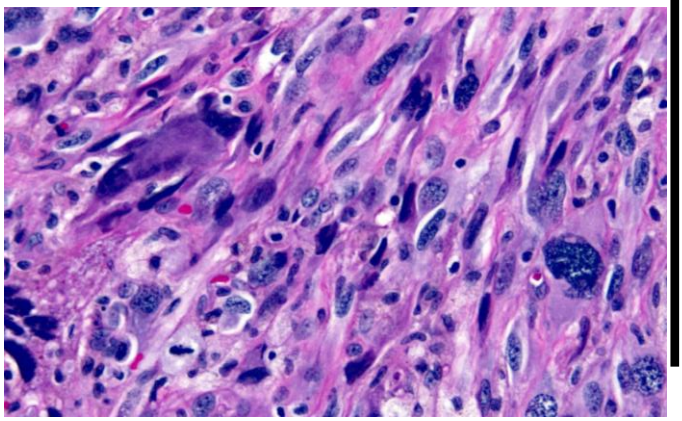
***Lots* genetic changes!**
Often with many passengers and subclones



Very abnormal Karyotype with frequent **Aneuploidy**

Complex Genetics → Pleomorphic, heterogeneous Morphology

Significant **pleomorphism** and **heterogeneity**

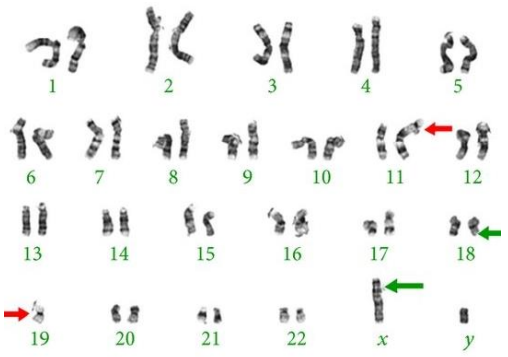


"Simple" Genetics

Overall, **Less common**
Most common in Sarcomas (~20%) and hematolymphoid tumors.

Examples: Synovial sarcoma, Infantile fibrosarcoma, Secretory carcinoma, Adenoid cystic carcinoma.

***Fewer* genetic changes**!
Often one main driver mutation, frequently a **fusion protein**.

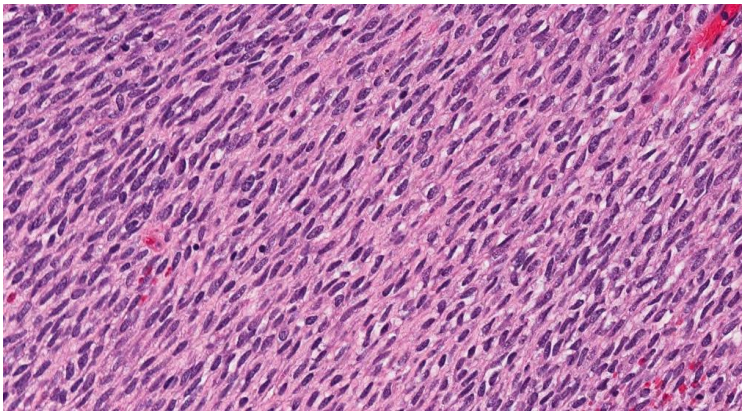


Synovial Sarcoma with X;18 translocation (green arrows) as well as association of chromosome 19 with telomere of chromosome 11 (red arrows).

Minimal Karyotype changes

Simple/monotonous Genetics → Monotonous, Blander Morphology

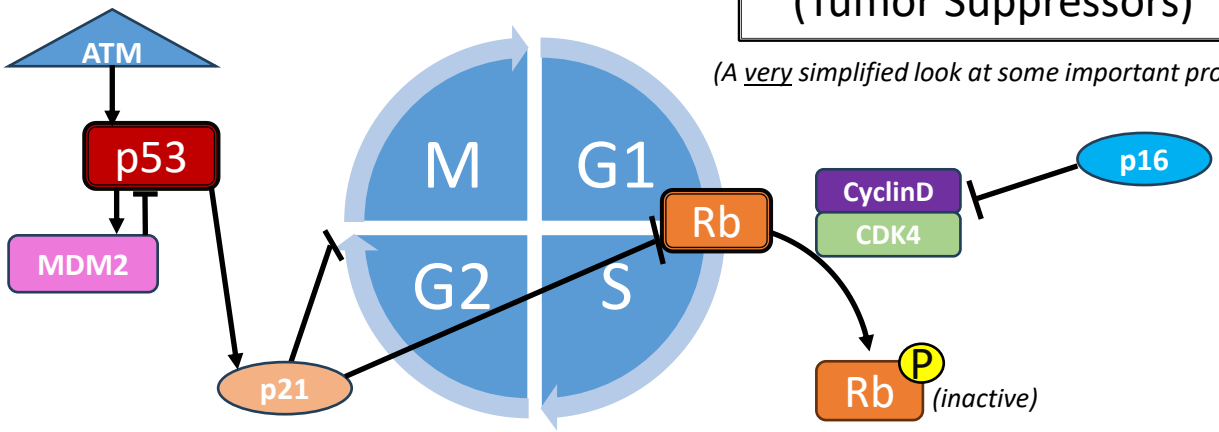
Monotonous morphology with more bland cytology



Common Pathways/alterations

Cell Cycle Checkpoints (Tumor Suppressors)

(A very simplified look at some important proteins)



TP53 — *"The Guardian of the Genome"*

Gene: TP53 Protein: p53

Most commonly mutated gene in human cancer (~70% of cancers).
p53 is activated by stresses, such as DNA damage, and causes G1 arrest, induces expression of DNA repair enzymes, and induces apoptosis if repair fails, through many complex pathways/interactions.

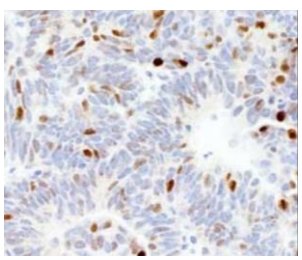
Germline mutation: Li-Fraumeni syndrome
Can be inactivated by viruses (e.g., E6 protein in HPV), causing carcinogenesis

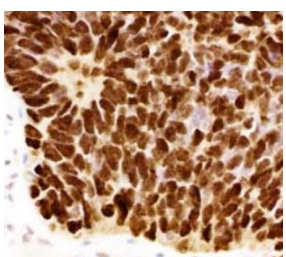
MDM2 inhibits p53, so it's amplification (e.g. in liposarcomas) leads to loss of p53 function.

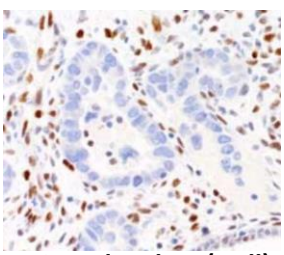
p53 immunohistochemistry is a quick and easy surrogate marker for TP53 mutation in most circumstances (Very good specificity, reasonable sensitivity). The most common pattern is diffuse strong nuclear positivity (overexpression) involving at least 80% of the tumor cells (but usually almost 100%). Wild-type staining is characterized by an admixture of negative cells, weakly and strongly positive cells. Notably, some splice site mutations or truncating mutations can result in detectable (but nonfunctional) p53 protein, yielding a normal wild-type staining pattern. Interpretation of complete loss requires good internal controls (e.g., lymphocytes and fibroblasts) to accurately interpret.

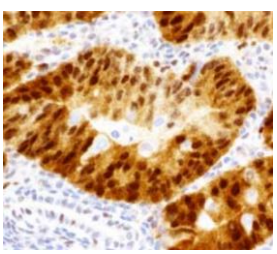
Staining pattern	p53 interpretation	TP53 status
Wild-type	Normal/Wild-type	No mutation
Overexpression	Abnormal/aberrant/mutant	Nonsynonymous missense mutation
Complete loss (Null)	Abnormal/aberrant/mutant	Loss of function mutation
Cytoplasmic	Abnormal/aberrant/mutant	Loss of function mutation, disrupting nuclear localization domain
Wild-type	Normal/Wild-type	Truncating mutation

Modified from: Köbel M, et al. Int J Gynecol Pathol. 2019 Jan;38 Suppl 1(Iss 1 Suppl 1):S123-S131. PMID: 29517499


Wild-type

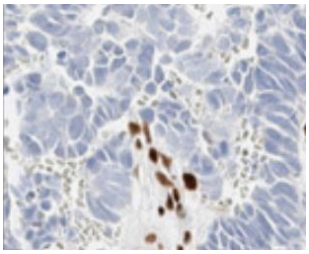

Overexpression


Complete loss (Null)


Cytoplasmic

RB1

"Governor of the cell cycle" ***Gene: RB1*** ***Protein: Rb***



Loss of nuclear Rb expression by IHC in small cell carcinoma

Key negative regulator of the cell cycle.
Directly or indirectly inactivated in most human cancers.
Regulates G1/S checkpoint (cells must pass to replicate DNA)

First tumor suppressor to be discovered. Protype for family cancer genes after found to be mutated in familial *R*etino*b*lastoma (in which one germline allele is inactive at birth; "Second hit"→ cancer).

RB mutations common in small cell carcinoma, breast, and bladder.
Functionally inactivated by E7 protein in HPV to cause cervical SCC.

Detection: **IHC** (looking for loss of staining), NGS

Cyclin D

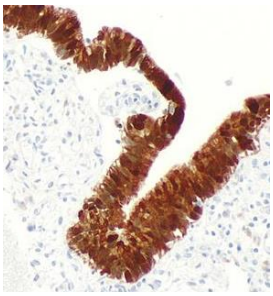
Mutations and amplification → RB phosphorylation → RB inactivation → No checkpoint!
Genes are overexpressed in many breast, esophagus, liver cancers as well as lymphomas and plasma cell neoplasms.
By IHC, typical staining pattern: Rb loss, P16 block positive, CyclinD1 low/loss

CDK4

Cyclin-Dependent Kinase (CDK) which helps regulate cell cycle.
Mutations and amplification → RB phosphorylation → RB inactivation → No checkpoint!
Amplifications common in melanomas, sarcomas, and glioblastomas

CDKN2A

Gene = CDKN2A ***Protein = p16***



"Block Positive" p16 IHC as a surrogate marker of HPV infection in endocervical AIS

Tumor suppressor. CDK inhibitor (CDKI) frequently disabled in cancer by mutation or gene silencing.

In many cancers, including HPV-mediated squamous cell carcinoma, High-grade serous carcinoma, small cell carcinoma, Liposarcoma, etc. abnormally accumulates→ **Diffuse, strong "block positive" staining by IHC**
Notably, this is only a *surrogate* often, for other abnormalities (like HPV infection), which can be investigated separately (e.g., by HPV ISH)

CDKN2A deletion is also seen in mesothelioma and can be detected by FISH and can be helpful in differentiating benign from malignant mesothelial proliferations

ALK

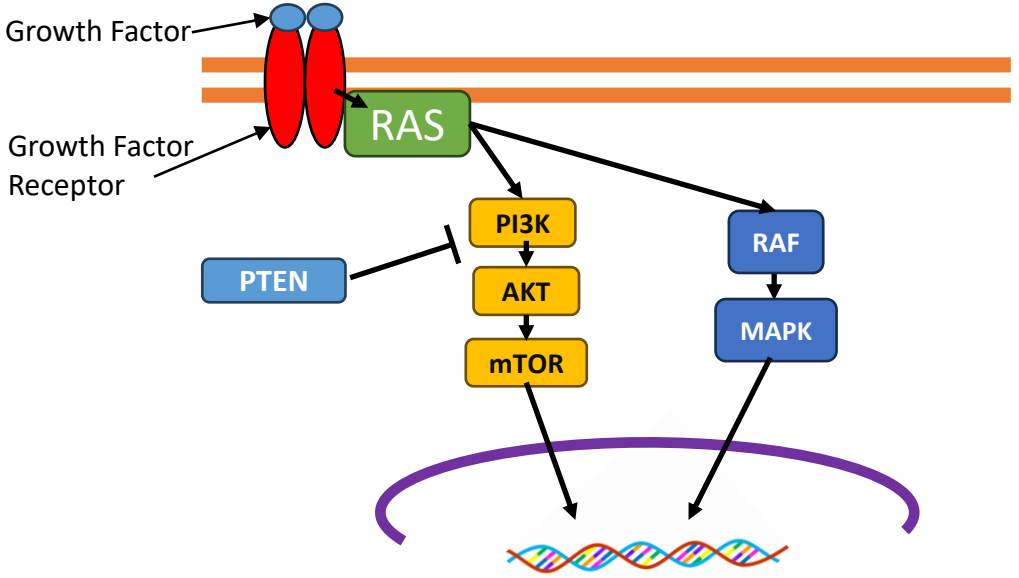
*Prior name: **A**naplastic **L**ymphoma receptor tyrosine **K**inase*

Encodes ALK **receptor tyrosine kinase**.
Binding of ligand induces homodimerization → activation of *multiple* pathways (RAS, JAK, PI3K) → all lead to cell proliferation and survival.

3 types of alterations:
ALK Translocations: Fusion proteins lead to ligand-independent constant activation.
Usually good prognosis. Common examples: NPM1-ALK in Anaplastic Large Cell Lymphoma (ALCL).
EML4-ALK in lung cancer (mutually exclusive of EGFR or KRAS) → Highly sensitive to ALK inhibitors crizotinib and ceritinib.
ALK Mutations: Many reported. Usually confer resistance to tyrosine kinase inhibitors
ALK Amplifications: Present in many tumor types

Alteration detection: IHC often used for diagnostic purposes (e.g., ALK+ ALCL and IMT)
For targeted therapy, FISH or NGS is often done (more specific)

RAS Pathway Oncogenes



Activation on Transcription → proteins like MYC → **Cell cycle progression & growth** → ***Cancer***

RAS

Three RAS genes: **KRAS, HRAS, NRAS**

Most commonly mutated oncogene(s) in human tumors.

Activated RAS stimulates downstream regulators of proliferation → uncontrolled growth.

KRAS mutations: Most commonly activated by point mutations (either at GTP-binding or enzymatic region, either way doesn't breakdown GTP and stays activated instead of normal cycling on/off).
About 30% of all tumors, higher in GI and Lung. For example, >90% of pancreatic ductal carcinomas, ~50% of colorectal. In lung, associated with smoking and confers a worse prognosis.
Confers a resistance to anti-EGFR therapy (cetuximab).
More commonly seen in metastatic cancers→ worse prognosis.

KRAS amplification: can be seen in a variety of cancers, including colorectal.

Detection: NGS

BRAF

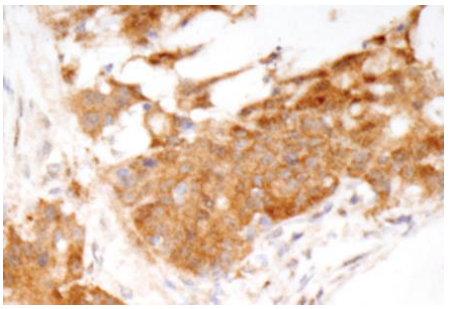
Serine/threonine kinase in RAF/ERK/MAP kinase pathway

Usually activated by somatic missense mutations → **constantly active**
Most common = BRAF V600E
Fusions can be seen.

Most common oncogenic driver in melanoma (~60%).
BRAF V600E mutation seen in **Papillary thyroid carcinoma** (~70%), **Hairy cell leukemia** (nearly 100%), Langerhans cell histiocytosis (~60%), Colorectal cancer (~10%, usually sporadic MSI-high), and NSCLC (~5%).

Detection: Usually, NGS
BRAF V600E mutation-specific IHC → shows cytoplasmic staining if positive (high concordance)

Therapy: BRAF inhibitors (e.g., Vemurafenib).
Often develop resistance.



Positive BRAF V600E IHC: diffuse cytoplasmic staining of tumor cells with or without membranous accentuation, and staining of nearly uniform intensity.

Note*: some normal tissue show some degree of staining, including normal cilia.*

EGFR

Epidermal Growth Factor Receptor (EGFR)

Oncoprotein that functions as a tyrosine kinase in EGFR pathway → cell proliferation
Altered in 30% of epithelial tumors→ correlate with tumor progression & resistance to chemotherapy.

EGFR Mutation: Found in ~15% of NSCLC with adenocarcinoma morphology (~40% in Asia). More common in female nonsmokers. Better prognosis.
Most activating mutations are in exons 18-21→ Sensitive to EGFR inhibitors (Erlotinib and Gefitinib)
T790M (exon 20) mutations confer resistance to inhibitors (can develop pre or post inhibitor therapy)

EGFR Amplification: Can see in NSCLC (often SCC), glioblastoma, and others

Detection: NGS
Therapy: EGFR Small molecule inhibitors (Erlotinib and Gefitinib), used in advanced/metastatic NSCLC
Anti-EGFR monoclonal antibody (Cetuximab) used in metastatic colorectal carcinoma (if RAS or BRAF mutated, then resistant)

HER2

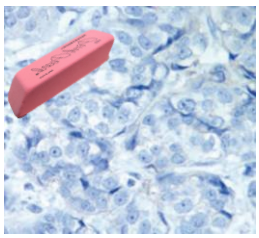
Gene: ERBB2 Protein: HER2/neu

EGFR family: Tyrosine kinase inhibitor involved in growth and signaling

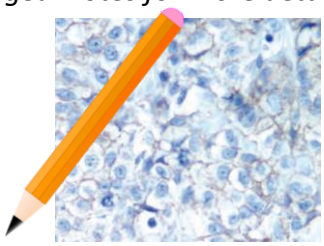
HER2 Amplification and overexpression: most common in Breast and Gastroesophageal carcinoma.
Smaller proportion of GYN, Salivary, Lung.
HER2 Mutations: less common than amplification. Seen in Lung, Breast, other

Can treat with **Trastuzumab** (Herceptin)—monoclonal antibody directed against HER2, or
Trastuzumab-deruxtecan (Enhertu)—Trastuzumab conjugated with the chemotherapy drug deruxecan.

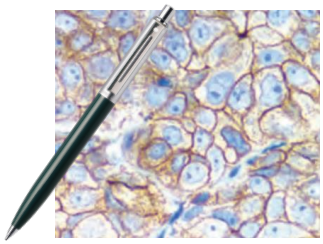
Detection: IHC, FISH
HER2 amplification/overexpression can be detected by IHC, with equivocal cases being analyzed by FISH
(See Breast and Esophageal notes for more details on HER2 IHC interpretation, as criteria vary by site)



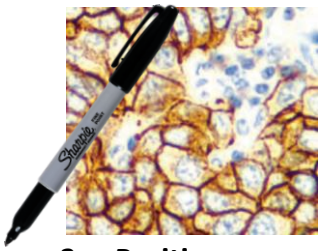
0, Negative
No staining is observed



1+, Negative
A faint/barely perceptible membrane staining



2+, Equivocal
A weak to moderate **membrane staining**



3+, Positive
Strong *complete* membrane staining

RET

Encodes transmembrane receptor tyrosine kinase → activates many pathways (e.g., RAS/MAPK and PI3K/AKT) → cell growth and proliferation.

RET fusions: Result in ligand-independent activation. ~1-2% NSCLC and 10-20% of Papillary thyroid carcinoma. Detect with NGS or FISH

RET Mutations: Germline mutations in MEN 2A and 2B. Somatic mutations in sporadic medullary thyroid carcinoma. Detect with NGS

Several multikinase inhibitors (e.g., sunitinib) have activity with both alterations

MET

Gene: MET Protein: HGFR

Receptor tyrosine kinase→ Various pathways (e.g., PI3K-Akt and RAS-MAP) → cell proliferation

MET overexpression (through amplification or epigenetic changes) is the most common alteration, particularly in lung cancer → often correlates with poor prognosis in many tumor types.

MET gene amplification can be secondary to treatment with anti-EGFR tyrosine kinase inhibitors

MET mutations: Exon 14 skipping mutations can be seen in NSCLC and have specific targeting drugs (Capmatinib and Tepotinib)

Detection: FISH or NGS

Can treat with **MET inhibitor**, but often develop resistance to these.

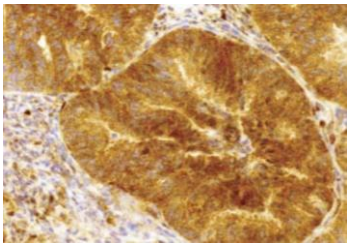
PTEN

Tumor suppressor, virtually ubiquitously expressed, regulates cell proliferation and apoptosis through inhibition of AKT via the PI3K/AKT pathway → inactivation leads to cell proliferation.

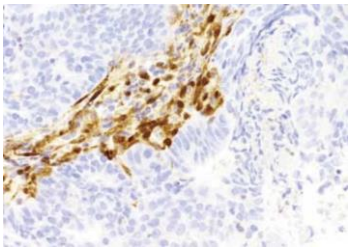
Germline mutations: PTEN hamartoma tumor (Cowden) syndrome

Mutations frequently seen in Endometrioid carcinoma (and useful diagnostically to in this setting to separate from mimics, like serous carcinoma), Prostate cancer, Glioblastoma

Detection: **IHC** for diagnostic purposes, but not high enough sensitivity/specificity (~80%) to be absolute surrogate marker, so may need NGS in some circumstances.



Retains/intact cytoplasmic staining (Normal)



Absent/complete loss (Abnormal)

MYC

Functions mainly by activating transcription of *other genes*, including those that drive the cell cycle.

Dysregulation promotes tumorigenesis by simultaneously promoting the progression of the cell cycle and enhancing alterations in metabolism that support cell growth.

Translocation of MYC with immunoglobulin gene loci in Burkitt lymphoma (and other lymphomas, like DLBCL, and leukemias)
Amplified in Breast, colon, lung, neuroblastoma.

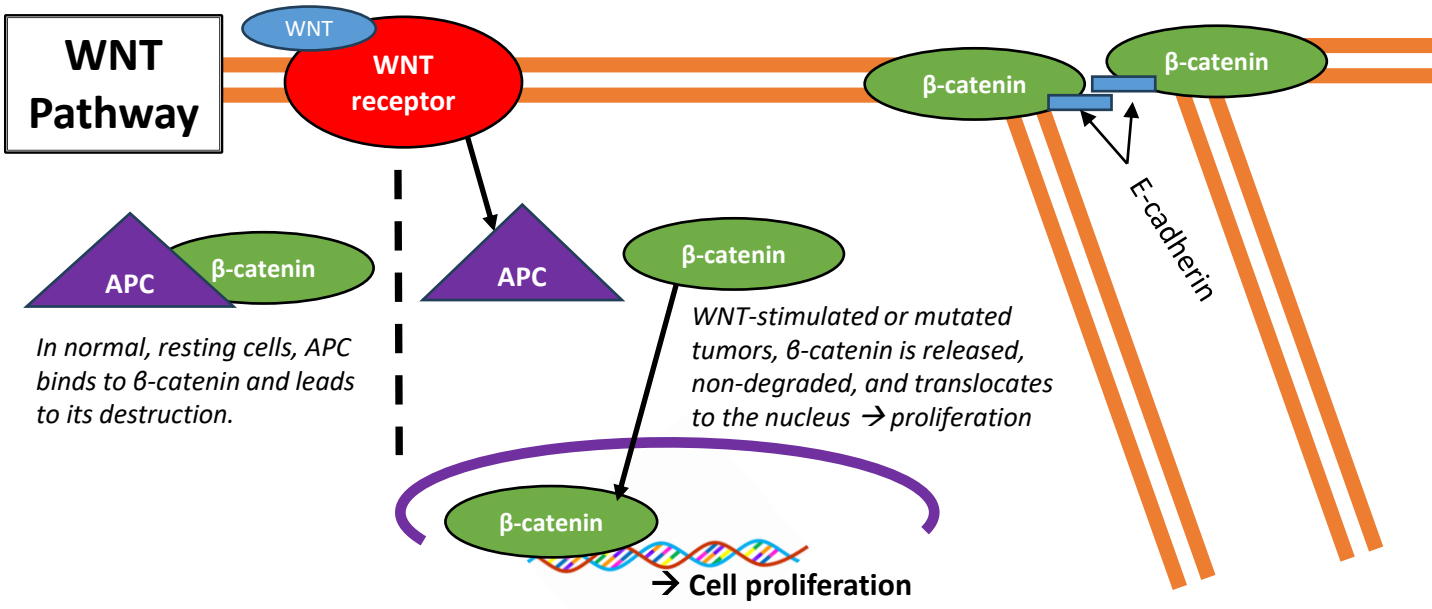
Detection: MYC translocations often confirmed by FISH break-apart assay
Although IHC exists, and protein expression is often upregulated with translocation, this is not considered an adequate surrogate marker.

ROS1

Encodes cell surface receptor tyrosine kinase → activation signals through multiple pathways (MAPK and PI3K) → cell growth and proliferation

Involved in fusion proteins with a variety of partners → signaling in the absence of a ligand
Fusions in ~1-2% of Non-small cell lung cancer → Classically, young, female, never smokers with adenocarcinoma → treat with crizotinib.

Alteration detection: NGS or FISH usually.

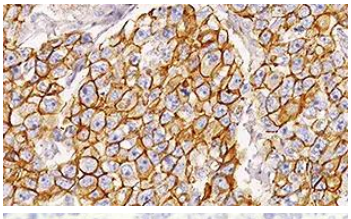


WNT Pathway

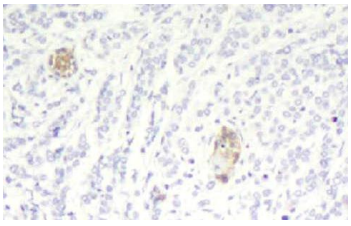
E-cadherin maintains cell "contact inhibition" in epithelium, which can be lost in malignancy. (Normal cells stop proliferating when they lose contact with each other, but cancer cells keep growing).

Lost in lobular breast cancer and diffuse gastric cancer.

Detection: IHC for E-cadherin shows membranous staining normally, which is lost in some tumors.



Intact membranous E-cadherin (Normal)



Lost membranous E-cadherin

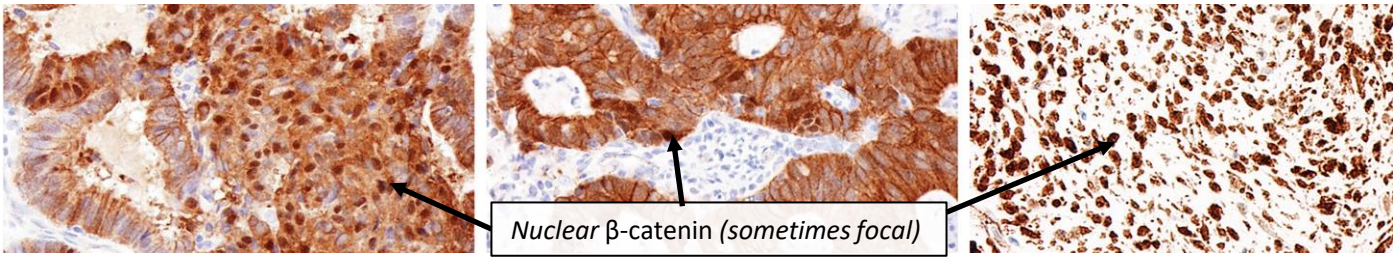
APC tumor suppressor discovered in Adenomatous Polyposis Coli Syndrome.
Encodes a cytoplasmic protein whose main function is to promote degradation of β-catenin.

Patients with APC germline mutations develop multiple adenomatous polyps in the colon and have a high rate of colon cancer.

β-catenin (*CTNNB1* gene) has 2 main roles:

1) Binds E-cadherin and is involved in cell adhesion and contact inhibition
2) When translocated to the nucleus, it binds transcription factors and coactivates genes that promote the cell cycle and proliferation

Detection: IHC for β-catenin shows *membranous* staining in many cells. With WNT pathway mutations (such as CTNNB1), *nuclear* β-catenin staining can be seen (warning: can be very focal, so look thoroughly!), which can be useful diagnostically in tumors such as desmoid fibromatosis, nasopharyngeal angiofibromas, and certain hepatic adenomas. IHC shows very good specificity, but only ~80% sensitivity, so NGS may be necessary in negative cases if it's important to confirm a mutation



SWI/SNF Complex

Chromatin-remodeling complexes→ regulate nucleosome positioning → alter transcription and also likely play a role in DNA repair.
Complicated macromolecular assemblies consisting of *many* subunits.
Nearly 25% of all cancers harbor aberrations in one or more of these genes→ linked to a **worse prognosis** across several cancer types (but not universal), particularly true for SMARCB1 (INI1)

Common/noteworthy subunits associated with cancer:
SMARCB1 (protein: **INI1**)—mutated in epithelioid sarcoma, ATRT, rhabdoid tumor *(all aggressive)*
SMARCA4 (protein: **BRG1**)—mutated in SMARCA4-undifferentiated thoracic tumor, Small-cell carcinoma of the ovary, hypercalcemic type, Undifferentiated endometrial carcinoma *(all aggressive)*
ARID1A—mutated frequently in clear cell and endometrioid carcinomas of GYN tract
PBRM1—mutated often in clear cell renal cell carcinoma
SS18—SS18–SSX fusion found in synovial sarcoma

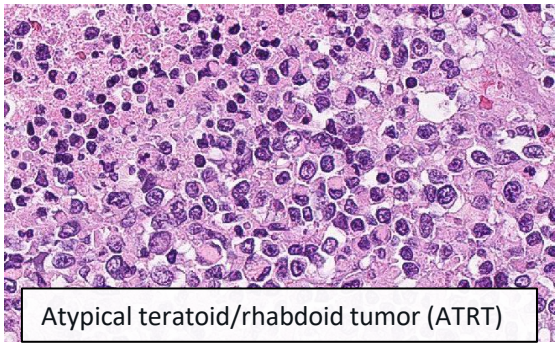
Detection: Varies by subunit
IHC exists for INI1, BRG1, ARID1A, where ***loss*** is indicative of mutation.
SS18-SSX fusion was historically tested for by FISH. Now, there is an IHC stain for the SS18-SSX chimeric protein that is a useful tool for synovial sarcoma differential diagnosis, but unusual immunophenotype should trigger molecular genetic testing.

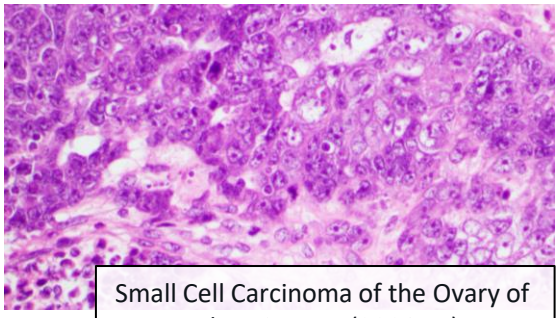
Genetics → Morphology → Behavior

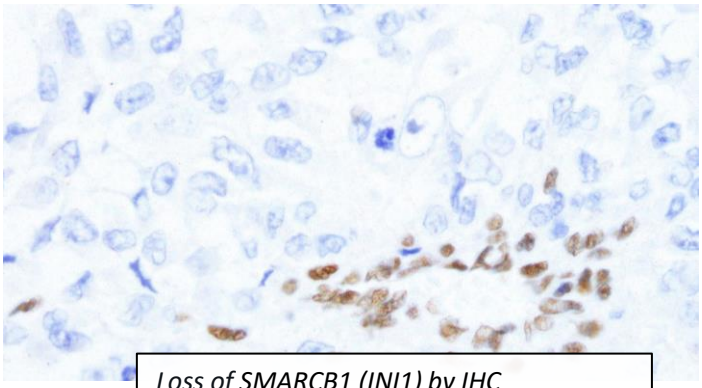
Particularly, for SMARCB1 (INI1) and SMARCA4 (BRG1) deficient tumors, these tumors are often very ***aggressive*** and often have a similar morphologic appearance, with poorly-differentiated, somewhat discohesive, rhabdoid to basaloid cells with vesicular chromatin, prominent nucleoli, and abundant necrosis.

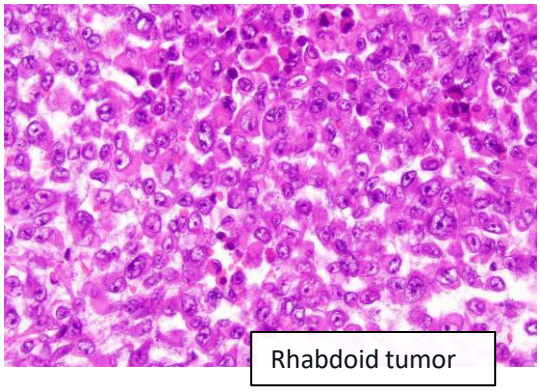
These tumors show variable keratin staining, and sometimes stain with CD34.

If you ever have a "bad" looking tumor with morphology like this, consider getting INI1 and/or BRG1 IHC to evaluate for one of these tumors as they can be seen in most organs.


Atypical teratoid/rhabdoid tumor (ATRT)


Small Cell Carcinoma of the Ovary of Hypercalcemic Type (SCCOHT)

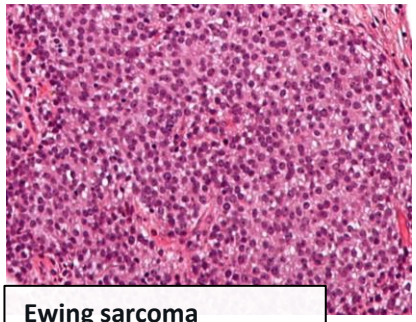

Loss of SMARCB1 (INI1) by IHC


Rhabdoid tumor

EWSR1

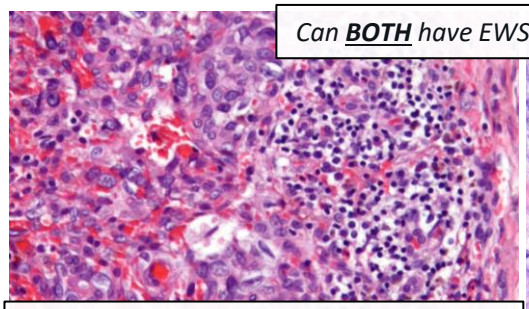
Discovered in Ewing sarcoma→ Ewing **s**arcoma breakpoint **r**egion **1** (EWSR1)
Involved in RNA binding, transcription, and RNA metabolism.
Role in cancer cell progression is still unclear, although it may play a critical role in DNA damage response and cell division. Nevertheless, fusion oncogenes can act as aberrant transcription factors, deregulating genes involved in tumorigenesis.

Significant "***fusion gene promiscuity***" as EWSR1 can fuse with ***many*** other genes.
One interesting aspect of this "promiscuity" is that identical fusion genes are present in tumor types that have entirely ***different*** histologic features and biologic potential—it's possible that tumorigenesis depends on the differentiation state of a cell and its cellular compartment.

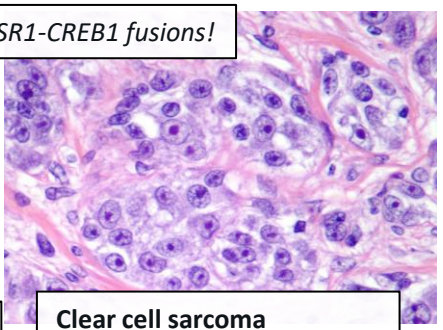


Ewing sarcoma
The prototypical "small round blue cell tumor"

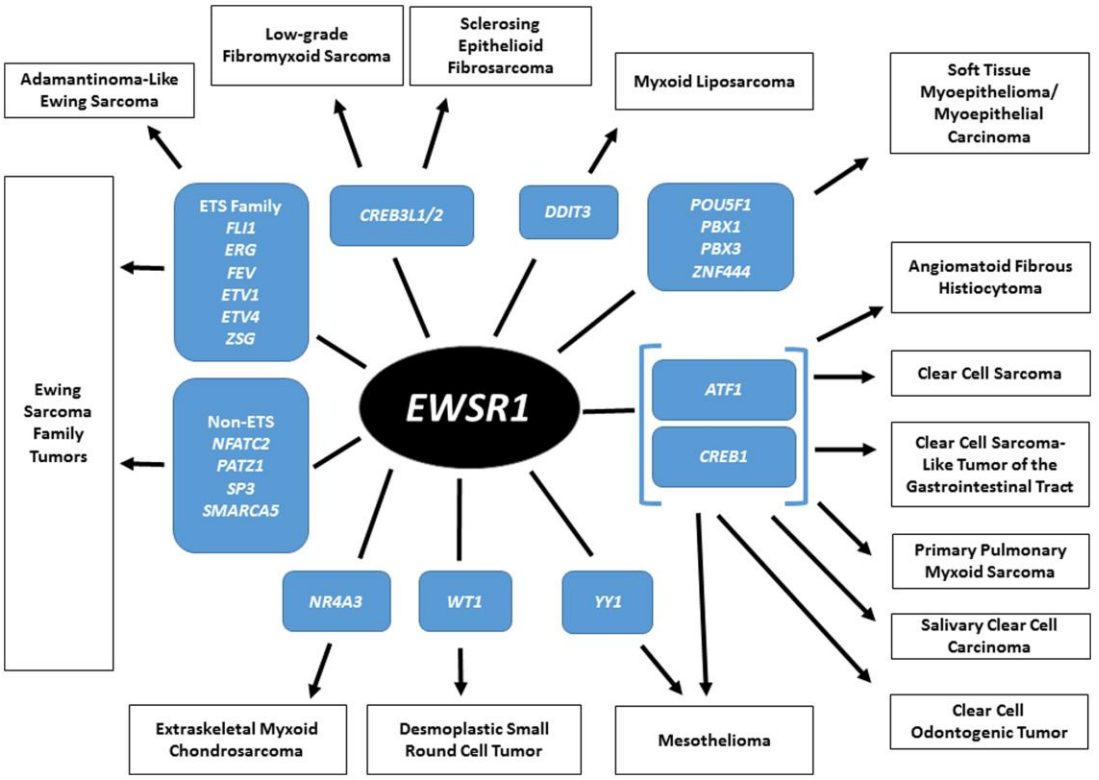
*Can **BOTH** have EWSR1-CREB1 fusions!*



Angiomatoid fibrous histiocytoma
(Intermediate behavior, <5% metastasize)



Clear cell sarcoma
(Malignant, poor prognosis)



From: Jo VY. EWSR1 fusions: Ewing sarcoma and beyond. Cancer Cytopathol. 2020 Apr;128(4):229-231. PMID: 31995669.

Detection: FISH, RNA-NGS, or RT-PCR panel
If I'm confident of the Dx on H&E, I usually get FISH as it's faster.
If I'm unsure, I do NGS or RT-PCR to identify the fusion partner, which can be diagnostically helpful.

KIT *aka C-kit, CD117*

Receptor Tyrosine Kinase: Homodimerization induces a signalizing cascade that has a roll in cell survival, proliferation, and differentiation.

Activating mutations are found in many tumors including **GISTS** (~80%), systemic mastocytosis (~75%), AML (~25%), and a subset of mucosal melanomas (~20%).
Treat with Receptor Tyrosine Kinase small molecule inhibitors, like **imatinib**, with often good response.

The specific exon mutated often impacts prognosis and RTK-inhibitor (RTKI) sensitivity, so tumors are often sequenced to determine therapy and prognosis.

Detection: NGS
No correlation between immunohistochemical expression of CD117 and KIT mutations

PDGFRA

***Platelet-derived growth factor receptor alpha*: Receptor Tyrosine Kinase**, in the same family as KIT

PDGFRA mutations: Found in GISTS (~10%).
Some, like D842V, indicate resistance to imatinib, while others do not.

PDGFRA fusions: Seen in various hematolymphoid tumors. Most common fusion is FIP1L1-PDGFRA, which most commonly presents as chronic eosinophilic leukemia. Responds to RTKI's, like imatinib.

Detection: Mutations by NGS. Fusions by FISH or RNA-NGS or RT-PCR.

FGFR

Fibroblast Growth Factor Receptors: family of Receptor tyrosine kinase that bine fibroblast growth factors. Downstream signaling via the MAP kinase, JAK/STAT, and PI3K/AKT pathways.
Alterations are most common in FGFR1, followed by FGFR2, FGFR3, and FGFR4, respectively.

FGFR Amplification: Seen in ~20% of lung SCC. Also some urothelial, breast, etc..

FGFR Mutation: Many different point mutations described in varied tumors.

FGFR Fusion: Seen in some cholangiocarcinoma, gliomas, and hematolymphoid malignancies.

Can be treated with multi-target TKIs (like above) or FGFR selective TKIs.

NTRK

Neurotrophic Receptor Tyrosine Kinases: Activate the downstream MAPK pathway.
Three: NTRK1, NTRK2, NTRK3

NTRK fusions: Seen in a variety of tumors of soft tissue, thyroid, salivary, GYN, gliomas, etc…
Most common→ETV6-NTRK3—congenital mesoblastic nephroma, infantile fibrosarcoma, secretory carcinoma
NTRK-rearranged solid tumors, regardless of tumor type, usually respond to treatment with a targeted inhibitor of the NTRK tyrosine kinases, larotrectinib and entrectinib.

NTRK mutations: Seen in a variety of tumors. Can contribute to drug resistance.

Detection: Pan-TRK IHC shows some utility as a diagnostic and surrogate marker for NTRK *screening*, however, physiologic or non-specific expression may lead to false-positive results (reasonable sensitivity, suboptimal specificity), so RNA testing is often required to confirm.

IDH1/2

Isocitrate dehydrogenase 1/2: Enzymes involved in Krebs cycle and metabolism.

IDH mutations: Present in the vast majority of astrocytomas and oligodendrogliomas, usually R132H. Also, Chondrosarcomas (~50%), Cholangiocarcinomas (~20%), Hematolymphoid tumors (~10% AML, ~20% AITL).

Inhibitors of mutant IDH1/2 enzymes are beginning use (e.g., Vorasidenib).

Detection: Mutations-specific antibody exists for IDH1 R132H, which is the most common mutation in gliomas, and works well. Genetic analysis of IDH1/2 genes by sequencing is recommended for diffuse gliomas when immunohistochemistry for IDH1 R132H is negative.

Microsatellite Instability (MSI) / Mismatch Repair (MMR) Deficiency

DNA mismatch repair (MMR) is a tumor suppressor system for recognizing and repairing errors that can arise during DNA replication and recombination, as well as repairing some forms of DNA damage.

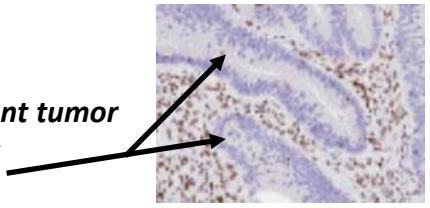
If these genes are mutant/lost/deficient→ defective DNA repair → tons of mutations accumulate ("hypermutated") → gain or loss of an entire short tandem repeat microsatellite fragments → **microsatellite instability** ("*MSI-high*")

Deficiency in MMR (dMMR) is seen frequently in colorectal cancer (CRC; ~15%), endometrioid carcinoma (~25%), other GI adenocarcinomas, and sebaceous carcinoma.

Can be *germline* (Lynch syndrome, less common) or *sporadic* (more common).
In colon, sporadic cases are often due to MLH1 promoter hypermethylation→ epigenetic silencing of MLH1 → dMMR

Detection: Universal screening of all new CRC and Endometrioid carcinomas. Can do via **IHC** and/or **PCR**. Most institutions do **IHC** first given its ease and availability.
Both methods have ~90% sensitivity for identifying lynch syndrome.

Intact nuclear expression of MLH1, MSH2, MSH6, PMS2→ ***MMR-proficient tumor***
Loss of nuclear staining with positive internal control → ***MMR deficiency***
(See GI tumor syndrome notes for more info on interpretation)



To identify Lynch syndrome cases: deficient tumors must undergo MLH1 promoter methylation testing (if present, methylation indicates a sporadic tumor). In colon cancer, BRAF mutations correlate very closely with MLH1 promoter methylation, so BRAF testing has become routine before MLH1 methylation testing as it is technically less demanding/costly.

Generally, MSI-high/dMMR tumors have lots of mutations→ cytologically higher grade → lots of neoantigens presented to immune system → lymphocytic infiltrates → better prognosis.

Note: DNA polymerase (POL) ε and δ mutations can result with a similar clinical phenotype

Therapy implications: Automatically approved for anti-PDL1 therapy (e.g., pembrolizumab)

Tumor Mutation Burden (TMB): calculated by counting the number of likely somatic coding mutations within a proportion of the tumor genome.

Although threshold for a "high" TMB score varies between panels, TMB-H is generally defined as ≥10 mutations/megabase (mut/Mb). dMMR (and POLE-mutated) tumors have high TMB.

Regardless of tissue type: TMB-high tumors all qualify for immunotherapy with pembrolizumab.

Histone Mutations/Modifications

Protruding from the nucleosome (DNA + Histones) are the terminal tails of the histone proteins, which can undergo post-translational modifications and/or mutations→ impact gene expression by 1) altering charge of histones (and their attraction to negatively charged DNA) 2) altering binding of other proteins.

Common post-translational changes: Methylation, Acetylation, Phosphorylation, Ubiquitination.

Specific histone modifications found in tumors:

H3 K27me3: Definitional of the pediatric infiltrative midline glioma "Diffuse midline glioma, H3 K27-altered" (aka Diffuse Intrinsic Pontine Glioma, DIPG) Shows *loss* of the methyl group on H3-K27 due to a mutation that replaces the lysine (K) with a methionine (M). There is an accompanying IHC stain, H3 p.K28me3 (K27me3), which shows *loss*.

MPNST's are characterized by many mutations including: NF1, CDKN2A and/or CDKN2B deletions, and inactivation of SUZ12 or EED (core components of PRC2). PRC2 mediates the deposition of H3 p.K28me3 (K27me3), an important repressive mark that plays a critical role in cellular differentiation. Inactivation of PRC2 leads to a complete global loss of H3 p.K28me3 (K27me3) in tumor cells, which can be demonstrated by IHC. About 80% of conventional high-grade MPNSTs show loss of H3 K27me3.

H3 G34: Definitional of "Diffuse hemispheric glioma, H3 G34-mutant," an infiltrative glioma involving the cerebral hemispheres of kids. Missense mutation of the H3-3A gene results in one of several substitutions on the histone H3 protein (usually G34R). Detect by NGS.

H3 G34W: Almost all (>95%) of Giant cell tumors of bone harbor pathogenic H3-3A (H3F3A) gene mutations, approximately 90% of which are H3.3 p.Gly34Trp (G34W), which can be detected by IHC (other mutations can be tested for by NGS).

H3 K36M: The vast majority of chondroblastomas harbor a p.Lys36Met (K36M) substitution in one of the genes that encode H3.3. This inhibits the H3K36 methyltransferases, which results in reduced global methylation. IHC using an antibody against H3.3B (H3F3B) p.Lys36Met (K36M) shows diffuse nuclear expression in > 96% of the cases.

DNA Methylation

Epigenetic modifications of DNA by methylation of cytosine nucleotides.

Cancer cells show global hypomethylation *and* selective promoter-localized hypermethylation.

Genome-wide hypomethylation → chromosomal instability → Increased copy number of oncogenes and decreased copy of tumor suppressors → Carcinogenesis

Methylation of promoter DNA leads to decreased gene transcription and expression
(attract proteins that cause chromatin condensation and block transcription)
Can block transcription of tumor suppressors→ cancer
Methylation of specific gene promoters has prognostic significance (e.g., promoter methylation of MGMT (DNA repair gene) in high-grade gliomas predicts response to alkylating agents).

Methylation profiling can be useful in classifying tumors and will likely be of increasing importance given its ability to identify new cancer classes and to consolidate histologically disparate cancers, particularly of the CNS and soft tissue. The NIH has a large "Methylscape Analysis" project and will perform free analysis of any CNS or soft tissue tumors sent to them.

Common Techniques

Cytogenetics

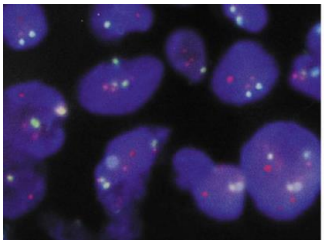
Allows to look at entire Karyotype
Low resolution (chromosome level changes only; 5-10 Megabase)
Examples of findings: Aneuploidy, large chromosome gain/loss, translocations
Requires living cells. Relatively slow.

Grow cells in culture → adding colchicine causes "metaphase arrest" → cells are dropped on a glass slide and stained to reveal band pattern.

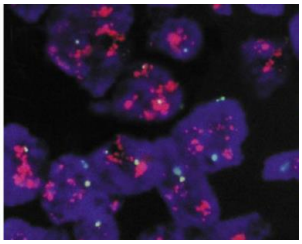
Balanced reciprocal translocation: 46,XY,t(9;22) (q11; p11).

FISH

"Florescence In Situ Hybridization"



Non-amplified

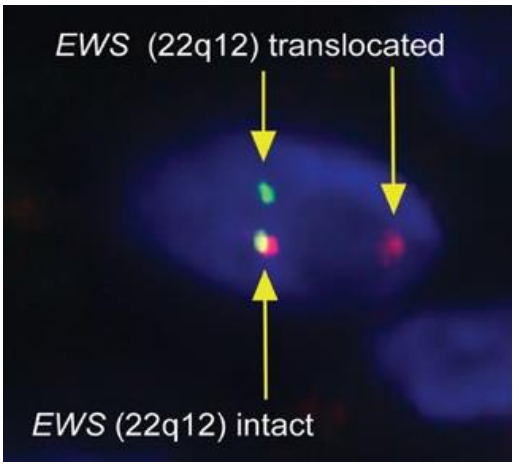


Amplified

Higher resolution (~40kb), but limited scope of analysis—only evaluates area(s) identified by sequence specific probes
Can be performed on FFPE, smears, etc..
Relatively fast (a 1-2 days)
Can evaluate multiple cells

Fluorescently labeled probe specific to DNA target sequence applied→ incubate to allow to hybridize to cell DNA → apply DAPI counterstain (to see DNA) → view under fluorescent microscope

Examples of findings: Fusion, Amplification



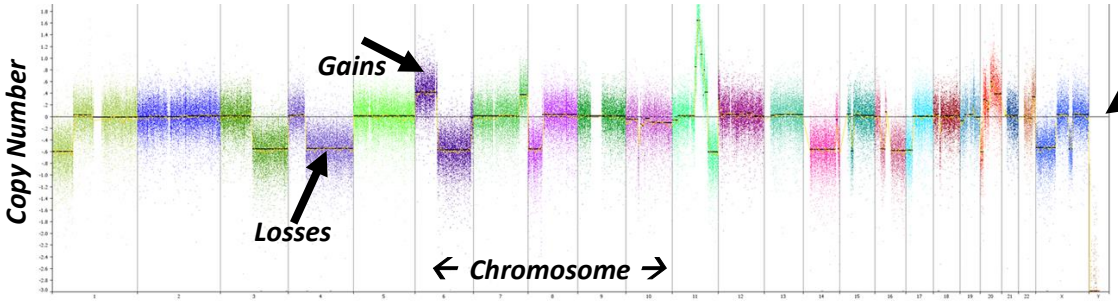
<u>Three probe strategies:</u>

1) Chromosome/Gene number
Detect presence or absence of specific chromosomes/genes

2) Fusion probes
If 2 genes involved in a translocation are known, gene #1 is bracketed by red probes and gene #2 is bracketed by green probes → translocation results in swaps of red and green with a fused yellow signal (highly specific)

3) Break-apart probes
If only one gene in fusion is know, it can be bracketed by a red probe on one side and green on the other (forming a yellow signal)→ Translocations result in separation of the probes (with no yellow) → doesn't identify gene partner

Chromosomal Microarray



A case of melanoma with numerous gains and losses of segments of chromosomes

Shows **Copy Number Variations (CNVs)** = genetic copy number losses or gains
Relative ploidy level of DNA in sample is compared to reference sample/database.
Improved resolution for microdeletions/duplications (~30kb)
No prior knowledge of chromosomal aberrations necessary (unlike FISH)
Can use FFPE tissue.

Types of Microarrays:

1) **Array-based comparative genomic hybridization (aCGH):**
Just detects copy number changes.
Microarray platform contains probes covering the whole genome. Test and control DNA are labeled with different fluorescent dyes→ both hybridize on a DNA chip → relative signal intensity for each probe is analyzed

2) **Single Nucleotide Polymorphism (SNP)-Chromosomal Microarray (SNP-CMA):**
Detects CNV, LOH, Uniparental disomy
Allows to detect copy-neutral changes not detected by aCGH
Contains BOTH representative probes from throughout the genome (nonpolymorphic copy number probes) and SNP probes→ only test DNA is added to array → relative signal intensity is compared to "in silico" (digital reference) to determine gains/losses

Polymerase Chain Reaction (PCR)

Targeted amplification of nucleic acids of interest.
Many methods rely heavily on PCR—Foundation of majority of clinical testing!
Highly sensitive & Specific
Allows production of large quantities of product from rare targets, allowing for analysis.

Standard PCR: Exponentially amplifies target DNA sequence→ can by analyzed by a variety of methods. Primers created to region of interest (provide specificity).

Reverse Transcriptase PCR: Exponentially amplifies target RNA sequence (i.e., transcriptome instead of genome). Uses reverse transcriptase to convert mRNA → cDNA, which can then be amplified by standard PCR. Useful for evaluating gene expression and fusion genes.

Real-Time PCR: Allows for detection and quantification of target DNA.
Based on detection of fluorescence during PCR (signal increases in proportion to amount of product formed).

Sequencing

Identifying the **nucleotide order** in nucleic acids.

"First Generation" Sequencing: Usually refers to Sanger sequencing.
Based on DNA chain termination→ creates fragments of different lengths, which can be separated by gel electrophoresis. Methodology is largely outdated.

"Next Generation" ("Second Generation") Sequencing (NGS): Massive parallel read analysis of multiple genes simultaneously, physically separated in a small space, starting with a single DNA or cDNA fragment at each location and recorded in real-time.
Increased sensitivity, specificity, and economics.
Principle driver of "genomic revolution" and accessibility of clinical genetic analysis.
General steps: Library preparation → Clonal amplification → Sequence analysis
Several different technologies exist with Illumina and Ion Torrent (Thermo-Fisher) comprising the largest market share.
Illumina—generates sequence by reversible dye termination ("sequence by synthesis")
Ion torrent—generates sequence by measuring change in pH when H+ is released from incorporated bases (ion-semiconductor based)

Regardless, this methodology utilizes extensive computational processing and bioinformatics, involving base calling, read alignment, variant identification, and variant annotation.

Sequencing depth— refers to the number of times a particular nucleotide is read during the sequencing process. Higher sequencing depth increases the confidence in calling a variant at a specific location.
Coverage—the proportion of the genome (or targeted region) that has been sequenced at least once.

"Third Generation" Sequencing: Techniques for single-molecule analysis without clonal amplification. Mostly research use at present.

Applications of NGS *beyond* genomics:
RNA-Seq (transcriptomics): identification and enumeration of many or most of the RNAs expressed in a sample and **Methylomics**.

DNA Methylation Analysis

"Methylomics"

Genome-wide interrogation of DNA methylation signatures, in conjunction with machine learning methods, has allowed for the creation of clinical-grade classifiers, most prominently in CNS and soft tissue tumors. Tumor DNA methylation profiling has led to the identification of new entities and the consolidation of morphologically disparate cancers into biologically coherent entities, and it will likely progressively become mainstream in the future.

Methods of detecting DNA methylation:
Bisulfite Conversion Method: Most common method. Sodium Bisulfite causes chemical conversion of unmethylated cytosines to uracil, which is then replaced by thymine during PCR (C→U→T). Ratio of cytosine to thymine at each potential conversion site is measured.

Non-bisulfate-dependent Methods:
Restriction endonuclease digestion: DNA incubated with methylation-specific enzymes→ digested products analyzed.
Affinity enrichment: Methylated DNA is purified via immunoprecipitation and the products are hybridized via array or NGS.

When **should I order molecular testing?**

The main reason to order molecular testing (or any testing for that reason!) is if it will change management of the patient. The main ways it could impact management are through altering ***Diagnosis*** or by offering ***Prognostic/Predictive*** insight.

Diagnosis**:**

As different tumors are being increasingly defined by specific molecular alterations, it is increasingly necessary to order confirmatory molecular testing. However, just because a confirmatory molecular test *exists*, doesn't mean that it is *necessary* for diagnosis!

Generally, I think testing is ***unnecessary*** when the diagnosis is reasonably certain based on other studies, including morphology. For example, if there is a retroperitoneal adipocytic tumor that looks classic for well-differentiated liposarcoma, ordering FISH for MDM2 amplification is likely an unnecessary cost and effort. But, if it's a fatty tumor with equivocal morphology, MDM2 testing is likely required. That said, some untrusting but savvy clinicians may request it in all cases.

Prognostic/Predictive:

Some genetic changes are **prognostic** and stratify patient outcomes (but do not determine therapy) and/or **predictive** of if someone's cancer will likely respond to a given therapy (e.g., BRAF V600 mutations in melanoma predict a beneficial response to a specific inhibitor, vemurafenib). These are both excellent reasons to get molecular testing, but should ideally be initiated by the clinical team who know if it is necessary for management.

Which **lesion should I test?**

If a stage IV cancer patient is considered for molecularly informed therapy, a frequent issue is whether to analyze the *primary* tumor or a *metastasis*?

No single easy answer and sample **availability and logistics often prevail**. For example, if a patient has a lung mass and liver masses, biopsying a liver mass is likely easier technically (it can often be done under ultrasound as opposed to CT) and it can potentially also help document metastasis.

In the absence of therapy, primary tumors, and metastases show mostly the same driver gene mutations. Therapy often accelerate clonal divergence among the primary and subsequent metastases.

So basically, do what you can with what you have. Practicalities often prevail.

When heterogeneity is suspected, "liquid biopsies" for circulating tumor DNA or circulating tumor cells in the peripheral blood may reflect the clinical situation better, presumably integrating tumor status from a variety of sites and reflecting a composite mutational landscape.

How much **tissue is necessary for Molecular?**

Minimum percent of tumor content needed **depends on sensitivity of method for testing** (so, when in doubt, check for the desired test), but many state a sensitivity of 5% variant allele frequency, which corresponds to **10% tumor content**, assuming heterozygosity.

Notably, this is of all DNA/Nuclei present in the specimen (including inflammatory cells, etc..). If there are a lot of other cells present, Macro or microdissection can allow for tumor enrichment.

FISH often requires fewer cells, but often still requires at least 20 identifiable tumor cells (and ideally more!).

What test should I order?

No easy single answer! It really **depends on what you're looking for** as there is **no universal test**. That said, generally, I use the following tests in the following circumstances:

IHC: As I've tried to highlight throughout, some molecular changes can be interrogated via IHC, particularly with "next generation" stains, like H3G34W for giant cell tumor of bone. IHC is relatively fast, cheap, and easy to interpret, so this is always my first "go to" test, if available/applicable.

FISH: Best for simple gene amplification/translocation questions where you *need limited information*, and would appreciate it *fast*. For example, I often get this diagnostically for MDM2 amplification and EWSR1 translocations in straightforward cases, where I'm pretty sure of the diagnosis going in and just need a quick confirmation. For translocations, if the partner matters, or I'm not certain of the diagnosis, I often get a RNA fusion panel.

RNA NGS or RT-PCR fusion panel: Look at a set selection of *common RNA fusion transcripts*. Best for cases that seem like they have a high likelihood of a fusion protein from a translocation, but where there are several possibilities (so a single FISH study is not ideal). Common scenarios where I get this: Sarcomas and Salivary gland tumors that seem like they have a high likelihood of having an underlying fusion, but where I am perhaps am not absolutely sure of the partner(s), particularly on a small biopsy, where I may have fewer morphologic clues.

***NGS Panel*:** Best for looking for *specific DNA mutations,* currently, based on the testing panel selected (although the technology can be used to query more things). Most commonly employed in the prognostic/predicative setting (particularly looking for druggable targets), but can be useful diagnostically, particularly on small biopsies.

For example, diagnostically, I've had small GI spindle cell neoplasm biopsies that I thought were inflammatory fibroid polyps, but where scant tissue made this definitively diagnosing this challenging. In this case, demonstrating a PDGFRA mutation helped support the diagnosis.

The optimal panel(s) of genes to target by NGS is controversial and evolving.
Many panels focus on "**Hotspots**"—regions within the genome that are commonly mutated, giving rise to specific, clinically significant, phenotypes. Middle ground between single gene and whole exome sequencing. Thought to be most cost effective (for now).
May have different panels for different types of tumors (e.g., solid vs hematologic), based on mutation frequency.

Often include:
Commonly mutated tumor suppressors/oncogenes (e.g., TP53)
Genes mutated in specific neoplasms (assist in diagnosis and subclassification)
Genes for targeted therapy (e.g., EGFR, BRAF)

Future: As sequencing and computation become faster and less expensive, it is possible that comprehensive genomic and transcriptomic sequencing will become standard.

For example, CARIS currently offers a single product with Whole Exome and Whole Transcriptome sequencing and tumor-relevant protein biomarkers.

Additional Resources

I haven't found a perfect molecular book yet (let me know if you find one!).
In my limited experience, *Diagnostic Pathology: Molecular Oncology*, but Vasef and Auerbach is the easiest resource to use, but the field of molecular pathology is growing so fast, I often have to turn to the primary literature.

Additional Internet Resources:

Association for Molecular Pathology
https://www.amp.org/

Databases for interpreting specific genetic alterations:

My Cancer Genome
A precision cancer database organized by disease, biomarker, drug, pathway, and clinical trial.
https://www.mycancergenome.org/

The Clinical Knowledgebase (CKB)
Digital resource for interpreting complex cancer genomic profiles in the context of protein impact, therapies, and clinical trials
https://ckb.jax.org/

OncoKB
MSK's Precision Oncology Knowledge Base—another Human Genetic Variant Database
https://www.oncokb.org/

ClinGen
National Institutes of Health (NIH)-funded resource dedicated to building a central resource that defines the clinical relevance of genes and variants for use in precision medicine and research.
https://clinicalgenome.org/