Clinical interpretation of genomic variants

Harriet Feilotter, PhD, FCCMG, FACMG

Professor, Dept of Pathology and Molecular Medicine, Queen's U

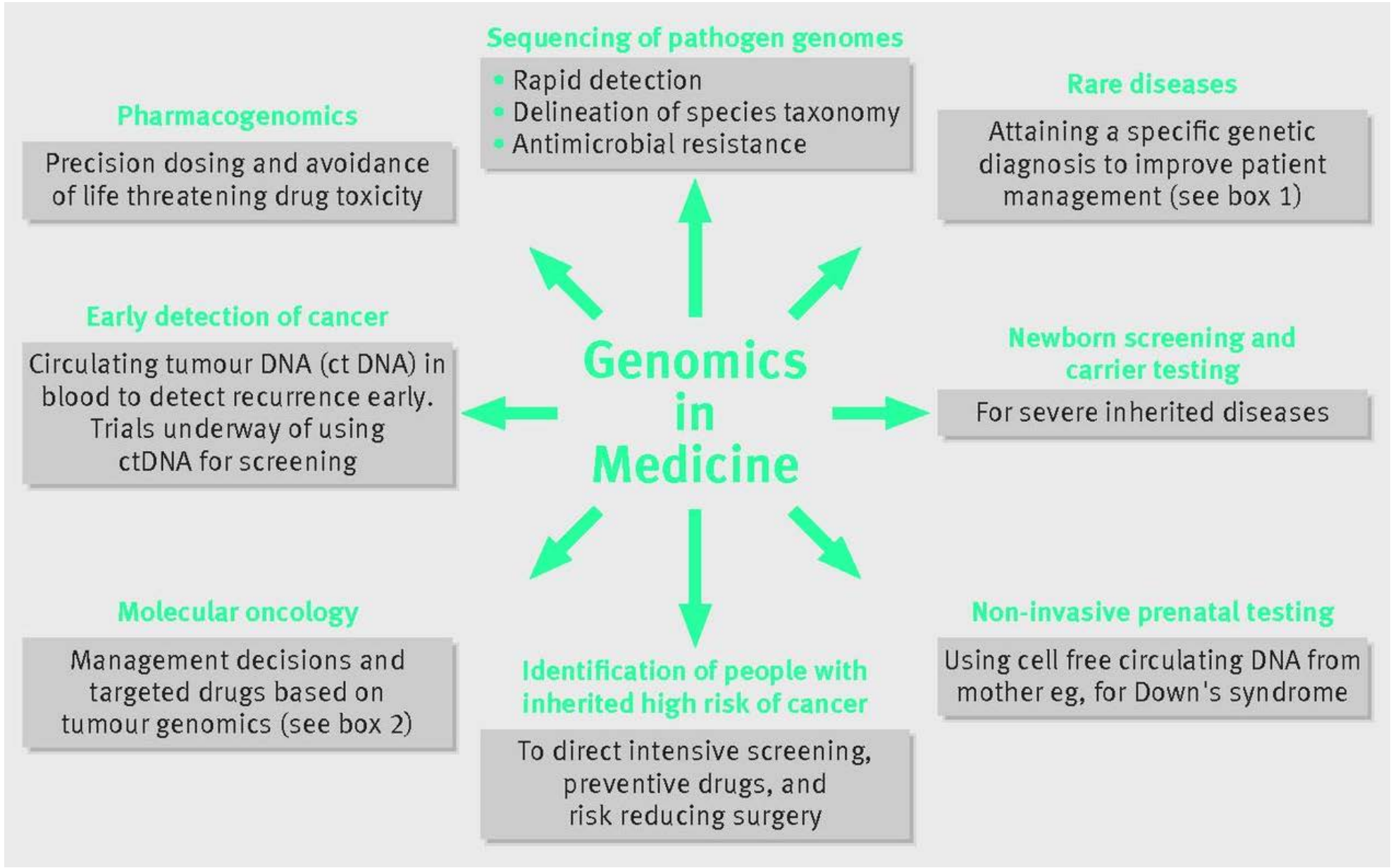
Director, Molecular Genetics, KHSC

Disclosures

- I have the following financial relationships to disclose:
  - Consultant for Janssen, Thermo Fisher, Roche, Astra Zeneca, Precision Rx, Diaceutics
  - Research Support from Thermo Fisher and Astra Zeneca
- I will not discuss off label use and/or investigational use in my presentation.

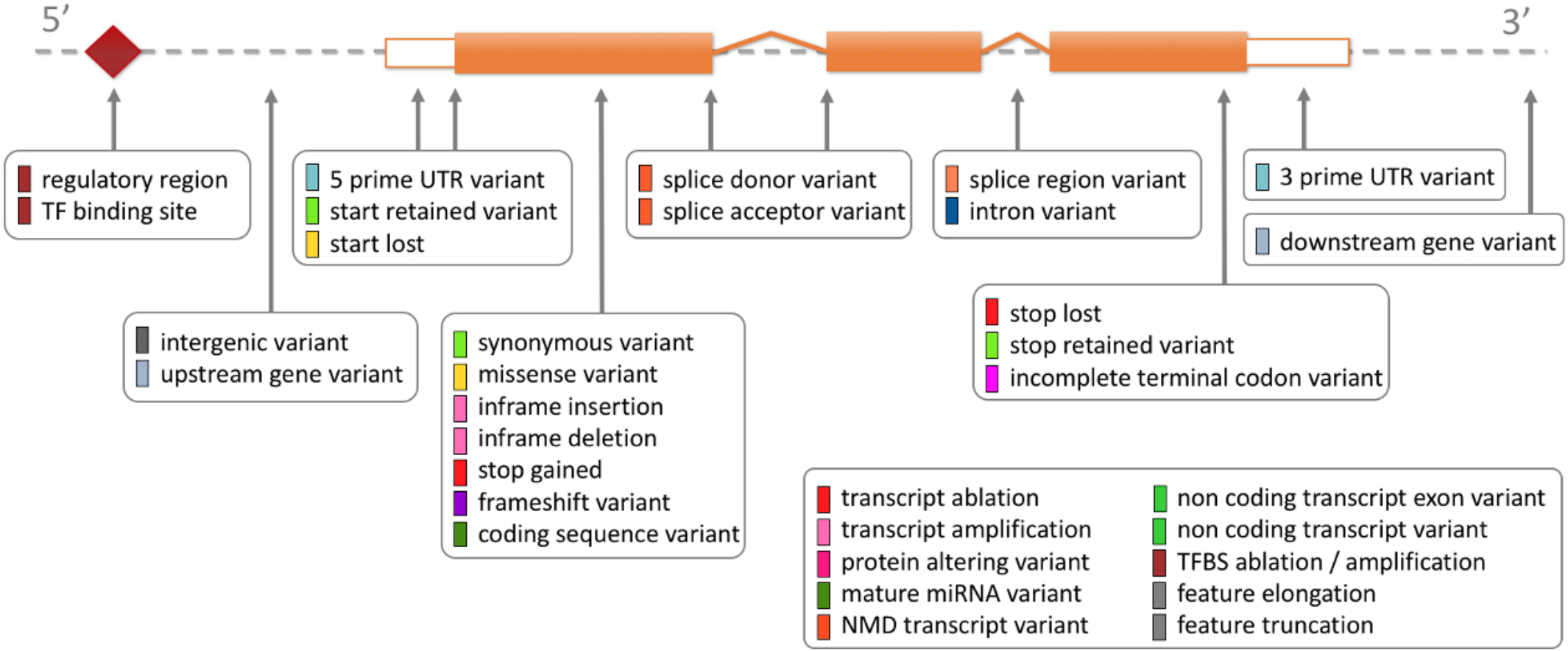
Learning objectives

- Distinguish between the interpretation of germline and somatic variants
- Determine the elements involved in interpretation of germline and somatic variants
- Understand the process of clinical characterization of variants in the context of tissue or disease site



Taken from BMJ 2018 361:k1687

Variant (mutation) types- not all variants are the same



https://useast.ensembl.org/info/genome/variation/prediction/predicted_data.html

Germline vs somatic variants

- Initiates in germ cell- therefore every cell in the body
- Often contribute to monogenic disease
- May contribute to risk of developing disease
- When linked to disease often involved in disease pathogenesis
- Concept of polymorphism
- Concept of benign, pathogenic

-Originates in a non-germline cell- therefore, not every cell in the body

-Most often in the context of cancer

-When linked to tumours may be diagnostic, predictive, prognostic

-Interpretation must include the tissue type

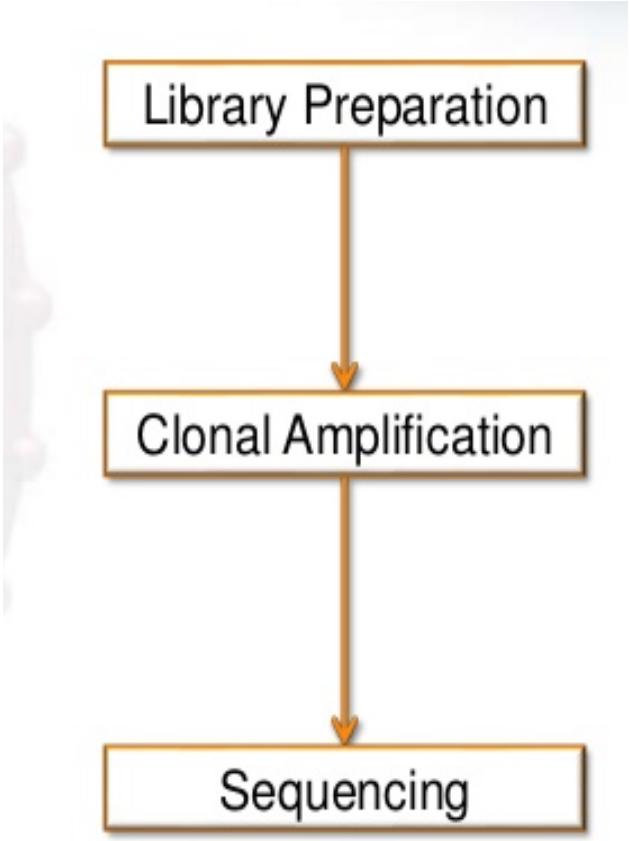
-Concept of drivers and passengers

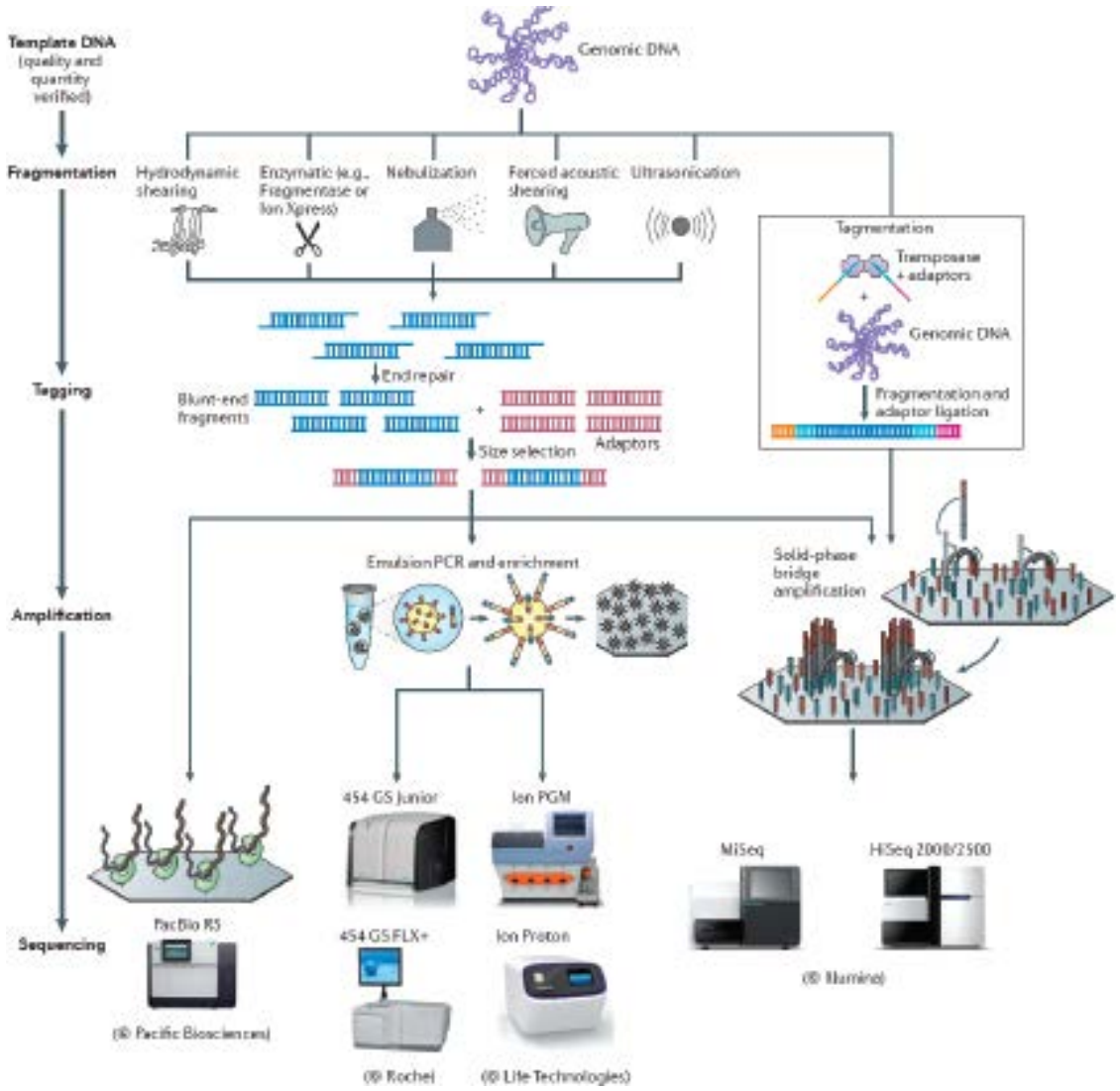
-Concept of actionable or not

Generic Massively Parallel Sequencing

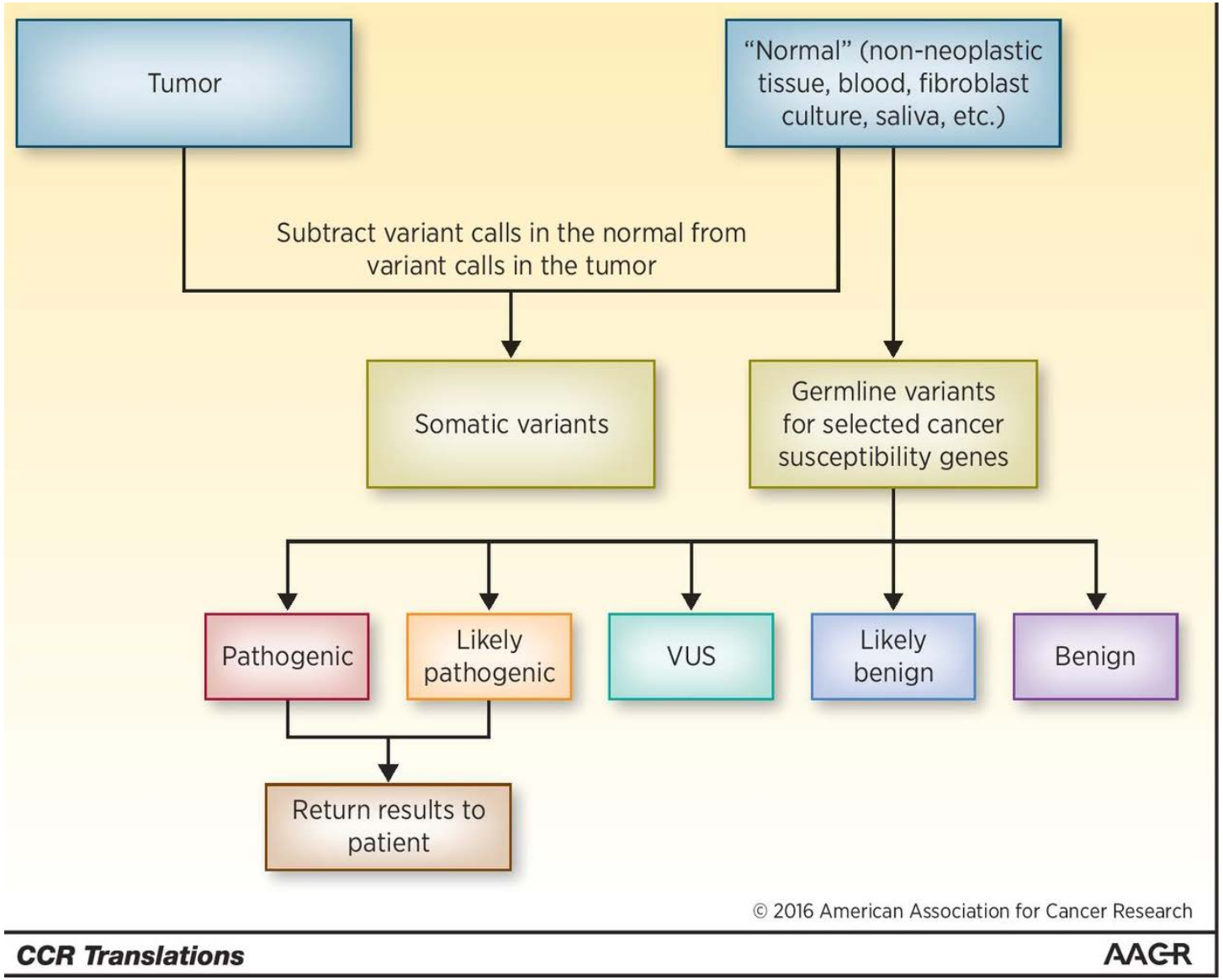
TUMOUR
DNA/RNA

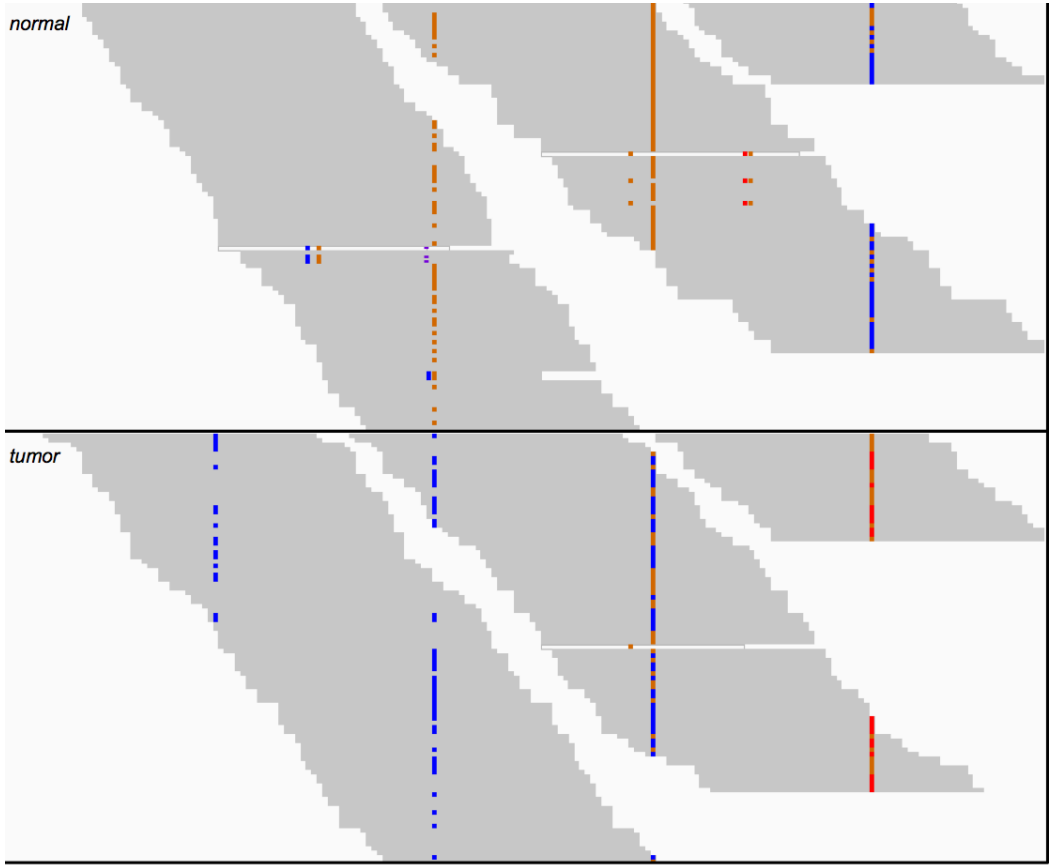
GERMLINE
DNA/RNA





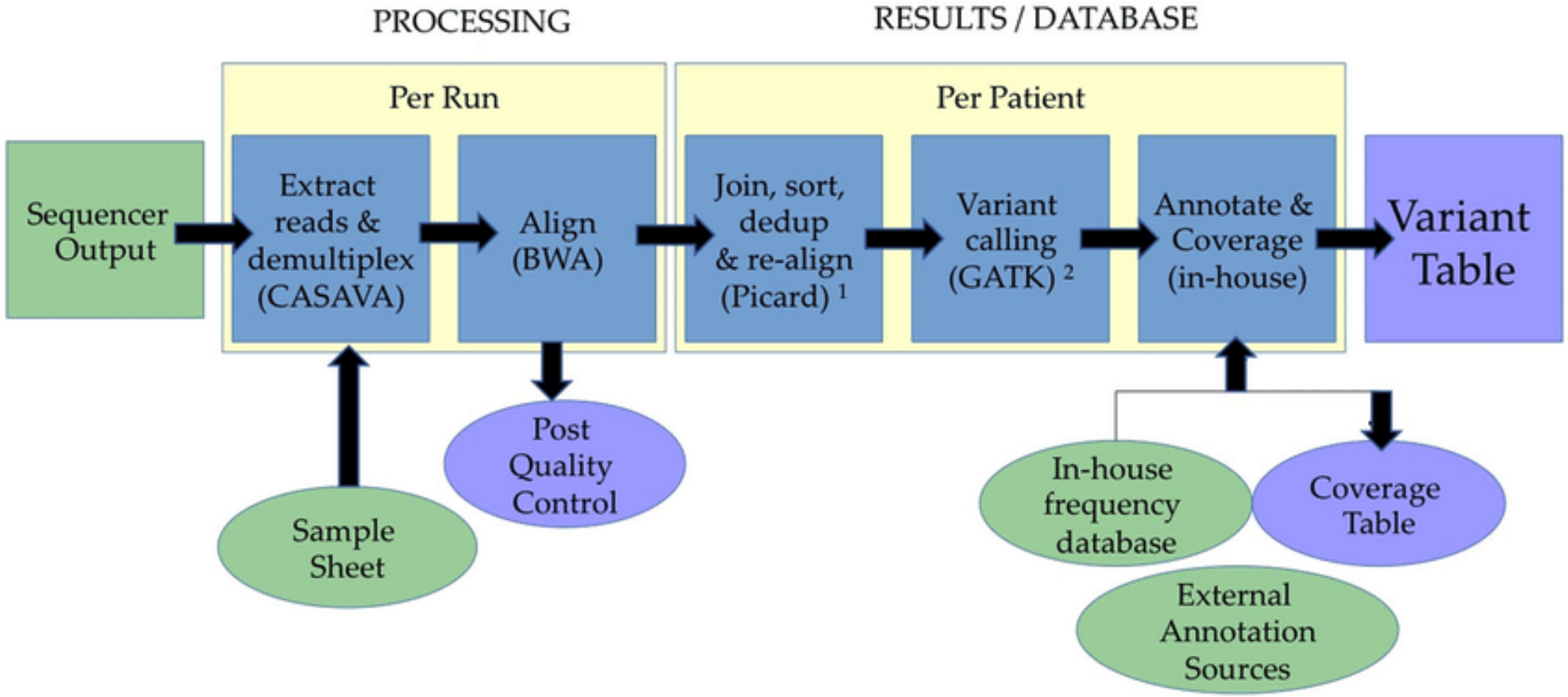
Germline vs somatic variants





https://gatkforums.broadinstitute.org/gatk/discussion/11127/somatic-calling-is-not-simply-a-difference-between-two-callsets

Pathway to variant calling



Annotation = identification of variant characteristics

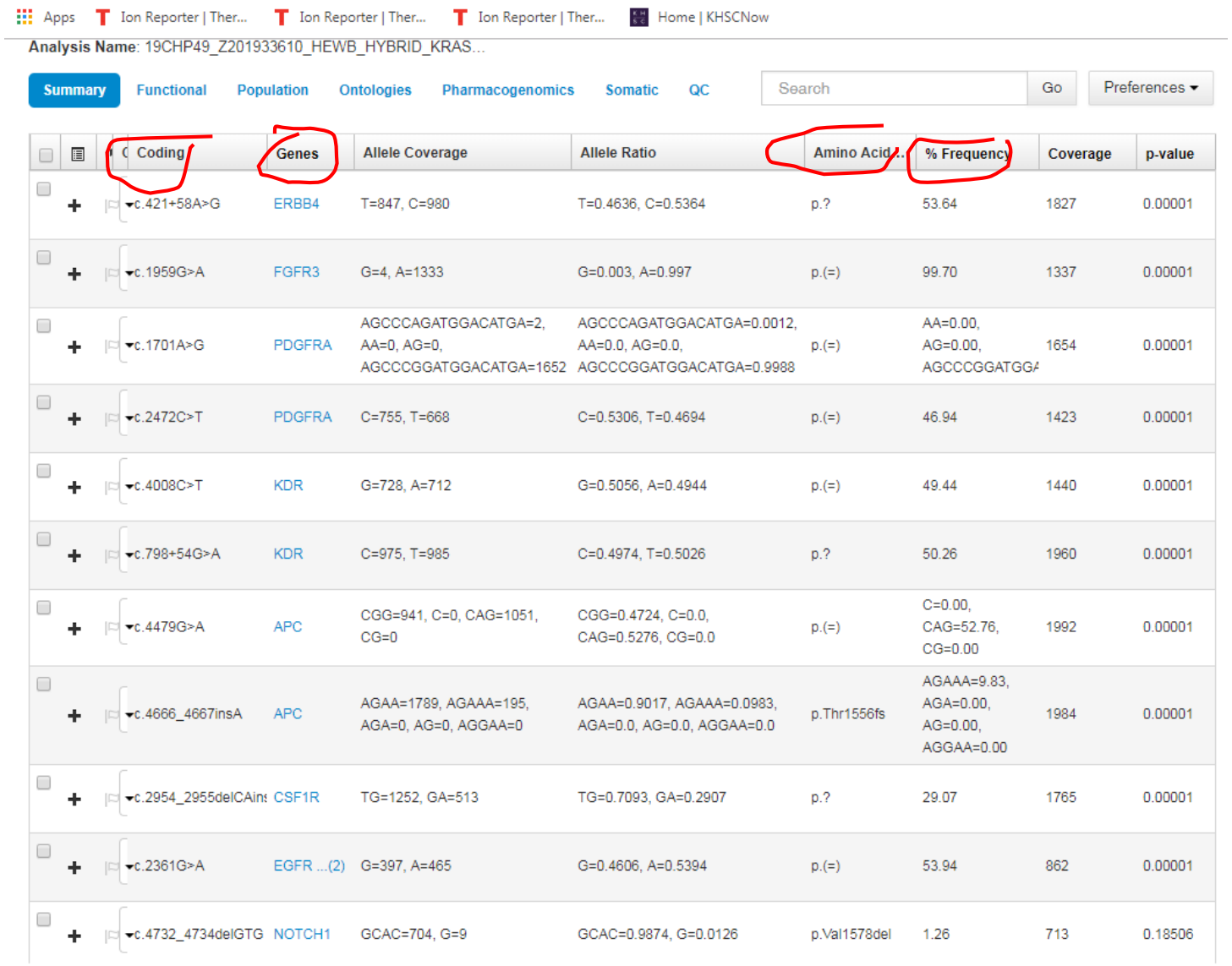
Interpretation = synthesis of characteristics and relation to clinical context

Taken from Kamps et al 2017 Int J Mol Sci 18:308

Example from Hereditary cancer testing case

Coverage	Ref	Alt	Ref#(F;R)	Alt#(F;R)	Alt%	Ins%	Del%	Overall Score	Mutation Call: HGVS Coding	Function	Zygosity	Gene	Strand	Exon	SNP db_xref	Amino Acid Change
504	C	G	1;1	262;240	99.60	0.00	0.00	21.5	NM_000251.2:c.[211+9C>G];[(211+9C>G)]	Noncoding	Homozygous	MSH2	+		rs2303426	
1152	A	T	325;248	317;262	50.26	0.00	0.00	24.3	NM_000251.2:c.[1511-9A>T];[=]	Noncoding	Heterozygous	MSH2	+		rs12998837	
822	G	A	0;1	302;518	99.76	0.00	0.00	23.1	NM_000251.2:c.[1661+12G>A];[(1661+12G>A)]	Noncoding	Homozygous	MSH2	+		rs3732183	
857	T	C	240;187	248;182	50.18	0.00	0.00	23.3	NM_000251.2:c.[2006-6T>C];[=]	Noncoding	Heterozygous	MSH2	+		rs2303428	
689	G	A	199;184	148;158	44.41	0.00	0.00	22.6	NM_000179.2:c.[116G>A];[=]	Missense	Heterozygous	MSH6	+	1	rs1042821	p.G39EG
448	G	A	139;95	129;85	47.77	0.00	0.00	21.1	NM_000465.2:c.[1568+14C>T];[=]	Noncoding	Heterozygous	BARD1	-		rs5031011	
832	CA	TG	236;197	228;170	47.84			23.3	NM_000465.2:c.1518_1519delTGinsCA	Missense	Heterozygous	BARD1	-	6		p.HV506_507HM
1213	C	G	1;1	632;579	99.84	0.00	0.00	24.5	NM_000465.2:c.[1134G>C];[(1134G>C)]	Missense	Homozygous	BARD1	-	4	rs2229571	p.R378S
689	G	A	196;202	136;155	42.24	0.00	0.00	22.6	NM_000465.2:c.[70C>T];[=]	Missense	Heterozygous	BARD1	-	1	rs1048108	p.P24SP
1144	A	G	309;303	271;260	46.42	0.00	0.00	24.3	NM_000249.3:c.[655A>G];[=]	Missense	Heterozygous	MLH1	+	8	rs1799977	p.I219IV
896	C	T	278;208	228;182	45.76	0.00	0.00	23.5	NM_000535.5:c.[2006+6G>A];[=]	Noncoding	Heterozygous	PMS2	-		rs111905775	
1376	T	C	1;2	657;716	99.78	0.00	0.00	24.9	NM_000535.5:c.[1621A>G];[(1621A>G)]	Missense	Homozygous	PMS2	-	11	rs2228006	p.K541E
1185	G	T	323;310	268;280	46.24	0.00	0.00	24.5	NM_000535.5:c.[1454C>A];[=]	Missense	Heterozygous	PMS2	-	11	rs1805323	p.T485TK
789	G	C	0;0	432;357	100.00	0.00	0.00	23.0	NM_000535.5:c.[780C>G];[(780C>G)]	Synonymous	Homozygous	PMS2	-	7	rs1805319	p.S260S
577	TA	TAA,AA	9;2	312;177,31;46	84.75,13.34			21.9	NM_000051.3:c.3403-14_3403-13insA,c.3403-15[T>A];[(T>A)]	Noncoding	Heterozygous	ATM	+			
764	A	G	0;0	394;369	99.87	0.00	0.00	22.9	NM_000051.3:c.[5948A>G];[(5948A>G)]	Missense	Homozygous	ATM	+	40	rs659243	p.N1983S
1230	A	C	324;303	320;283	49.02	0.00	0.00	24.7	NM_000059.3:c.[1114A>C];[=]	Missense	Heterozygous	BRCA2	+	10	rs144848	p.N372NH
1354	A	G	1;2	694;657	99.78	0.00	0.00	24.9	NM_000059.3:c.[4563A>G];[(4563A>G)]	Synonymous	Homozygous	BRCA2	+	11	rs206075	p.L1521L
1081	G	C	0;0	548;533	100.00	0.00	0.00	24.1	NM_000059.3:c.[6513G>C];[(6513G>C)]	Synonymous	Homozygous	BRCA2	+	11	rs206076	p.V2171V
935	T	C	0;0	421;514	100.00	0.00	0.00	23.6	NM_000059.3:c.[7397T>C];[(7397T>C)]	Missense	Homozygous	BRCA2	+	14	rs169547	p.V2466A
644	T	C	207;129	185;120	47.36	0.00	0.47	22.4	NM_000059.3:c.[7806-14T>C];[=]	Noncoding	Heterozygous	BRCA2	+		rs9534262	
437	C	T	125;119	95;95	43.48	0.00	0.00	21.0	NM_004360.3:c.[48+6C>T];[=]	Noncoding	Heterozygous	CDH1	+		rs3743674	
961	T	C	226;233	252;250	52.24	0.00	0.00	23.7	NM_004360.3:c.[2076T>C];[=]	Synonymous	Heterozygous	CDH1	+	13	rs1801552	p.A692AA
1078	G	C	264;253	269;291	51.95	0.00	0.09	24.1	NM_000546.5:c.[215C>G];[=]	Missense	Heterozygous	TP53	-	4	rs1042522	p.P72RP
1072	A	G	253;265	268;285	51.59	0.00	0.00	24.1	NM_032043.2:c.[3411T>C];[=]	Synonymous	Heterozygous	BRIP1	-	20	rs4986763	p.Y1137YY
1164	A	G	276;302	276;310	50.34	0.00	0.00	24.4	NM_032043.2:c.[2755T>C];[=]	Missense	Heterozygous	BRIP1	-	19	rs4986764	p.S919SP
1244	T	C	316;333	295;300	47.83	0.00	0.00	24.6	NM_032043.2:c.[2637A>G];[=]	Synonymous	Heterozygous	BRIP1	-	19	rs4986765	p.E879EE
1769	C	T	513;442	467;347	46.01	0.00	0.00	25.8	NM_032043.2:c.[69G>A];[=]	Synonymous	Heterozygous	BRIP1	-	2	rs45458996	p.P23PP
1369	G	C	341;399	293;336	45.95	0.00	0.00	25.0	NM_000455.4:c.[920+7G>C];[=]	Noncoding	Heterozygous	STK11	+		rs2075607	

Example from somatic cancer



Coding	Genes	Allele Coverage	Allele Ratio	Amino Acid...	% Frequency	Coverage	p-value
c.421+58A>G	ERBB4	T=847, C=980	T=0.4636, C=0.5364	p.?	53.64	1827	0.00001
c.1959G>A	FGFR3	G=4, A=1333	G=0.003, A=0.997	p.(=)	99.70	1337	0.00001
c.1701A>G	PDGFRA	AGCCCAGATGGACATGA=2, AA=0, AG=0, AGCCCGGATGGACATGA=1652	AGCCCAGATGGACATGA=0.0012, AA=0.0, AG=0.0, AGCCCGGATGGACATGA=0.9988	p.(=)	AA=0.00, AG=0.00, AGCCCGGATGGA	1654	0.00001
c.2472C>T	PDGFRA	C=755, T=668	C=0.5306, T=0.4694	p.(=)	46.94	1423	0.00001
c.4008C>T	KDR	G=728, A=712	G=0.5056, A=0.4944	p.(=)	49.44	1440	0.00001
c.798+54G>A	KDR	C=975, T=985	C=0.4974, T=0.5026	p.?	50.26	1960	0.00001
c.4479G>A	APC	CGG=941, C=0, CAG=1051, CG=0	CGG=0.4724, C=0.0, CAG=0.5276, CG=0.0	p.(=)	C=0.00, CAG=52.76, CG=0.00	1992	0.00001
c.4666_4667insA	APC	AGAA=1789, AGAAA=195, AGA=0, AG=0, AGGAA=0	AGAA=0.9017, AGAAA=0.0983, AGA=0.0, AG=0.0, AGGAA=0.0	p.Thr1556fs	AGAAA=9.83, AGA=0.00, AG=0.00, AGGAA=0.00	1984	0.00001
c.2954_2955delCAins	CSF1R	TG=1252, GA=513	TG=0.7093, GA=0.2907	p.?	29.07	1765	0.00001
c.2361G>A	EGFR ...(2)	G=397, A=465	G=0.4606, A=0.5394	p.(=)	53.94	862	0.00001
c.4732_4734delGTG	NOTCH1	GCAC=704, G=9	GCAC=0.9874, G=0.0126	p.Val1578del	1.26	713	0.18506

Variant Interpretation Classifications
Germline Variants

Pathogenic	Likely Pathogenic	Uncertain Significance (VUS)	Likely Benign	Benign

Benign		Pathogenic			
Strong	Supporting	Supporting	Moderate	Strong	Very Strong

Richards et al, *Genet Med*, 2015

ACMG STANDARDS AND GUIDELINES

Allele frequency in large studies of control chromosomes- is the allele frequency too high? Has the allele been seen in homozygous state? Is it particular to some ethnic group? Is the prevalence in disease increased over controls?

Computational predictions are just that- predictions. Generally not assigned a high impact and can only be used if all tools point in the same direction, which becomes more unlikely the more tools one tries. However some can be very strong- ie null variants

Functional data requires the interpreter is expert in all fields of physiology, biochemistry, metabolomics, genomics, statistics, analytics and assessing the work-up of the variant

De novo- did the variant appear in the affected individual for the first time or has the parent been unaffected in this is confirmed? Depending on inheritance mode, this will have different impact

Segregation data can be highly impactful if sufficient meiosis are examined- the more affected family members segregating the variant the stronger the evidence

Has the variant been seen in trans or in cis with another variant?

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very strong
Population data	MAF is too high for disorder BA1/BS1 **OR** observation in controls inconsistent with disease penetrance BS2			Absent in population databases PM2	Prevalence in affecteds statistically increased over controls PS4	
Computational and predictive data		Multiple lines of computational evidence suggest no impact on gene /gene product BP4 Missense in gene where only truncating cause disease BP1 Silent variant with non predicted splice impact BP7 In-frame indels in repeat w/out known function BP3	Multiple lines of computational evidence support a deleterious effect on the gene /gene product PP3	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before PM5 Protein length changing variant PM4	Same amino acid change as an established pathogenic variant PS1	Predicted null variant in a gene where LOF is a known mechanism of disease PVS1
Functional data	Well-established functional studies show no deleterious effect BS3		Missense in gene with low rate of benign missense variants and path. missenses common PP2	Mutational hot spot or well-studied functional domain without benign variation PM1	Well-established functional studies show a deleterious effect PS3	
Segregation data	Nonsegregation with disease BS4		Cosegregation with disease in multiple affected family members PP1	Increased segregation data →		
De novo data				De novo (without paternity & maternity confirmed) PM6	De novo (paternity and maternity confirmed) PS2	
Allelic data		Observed in *trans* with a dominant variant BP2 Observed in *cis* with a pathogenic variant BP2		For recessive disorders, detected in trans with a pathogenic variant PM3		
Other database		Reputable source w/out shared data = benign BP6	Reputable source = pathogenic PP5			
Other data		Found in case with an alternate cause	Patient's phenotype or FH highly specific for			

NM_005228.3(EGFR):c.2361G>A(p.=)

- Synonymous variant- does not change amino acid
- No indication that it could affect splicing either
- Population data says it is found across all tested populations at a frequency of 40%, reaching 53% in some populations
- This is sufficient to know that this is not a pathogenic germline variant
- Falls clearly into the neutral or benign polymorphism category
- No need to do additional work

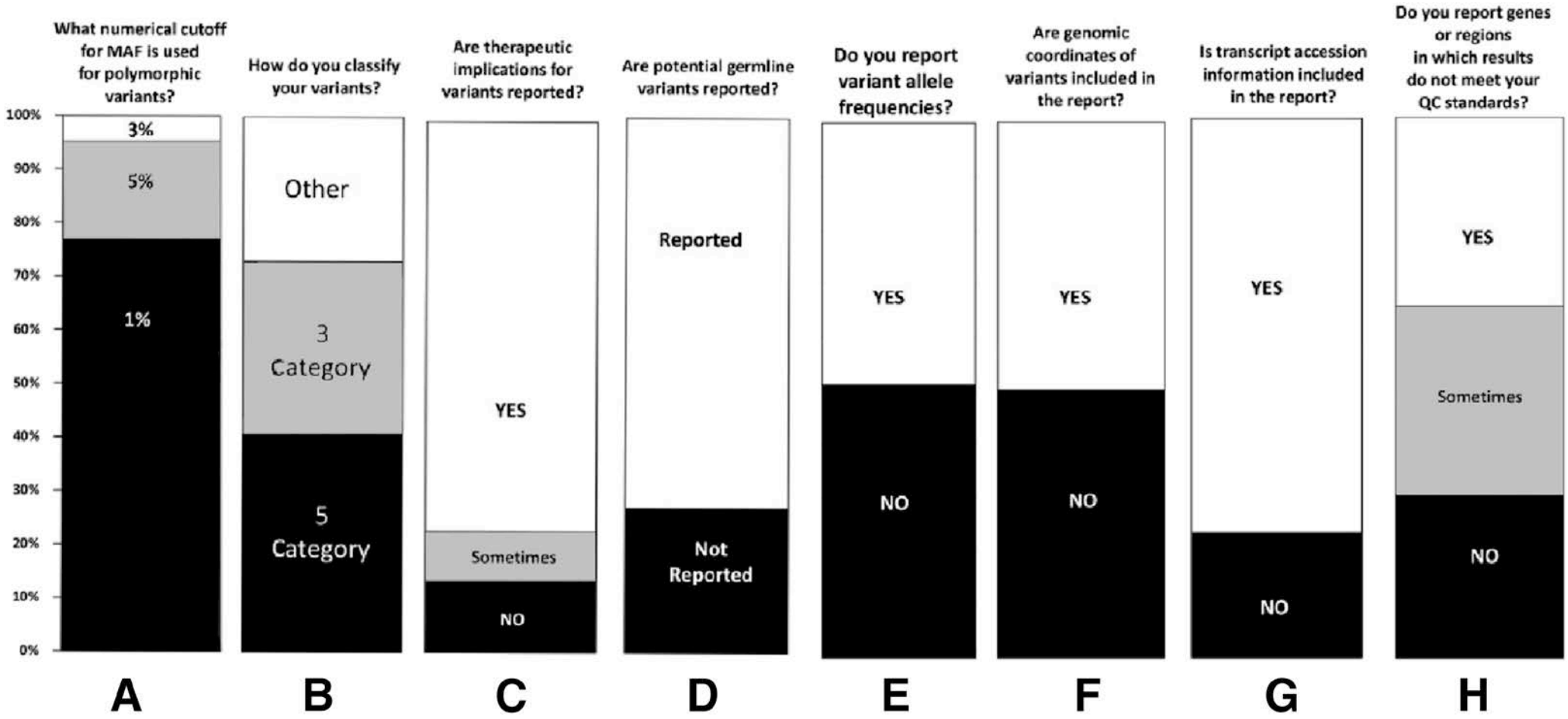
NM_000059.2(BRCA2):9976A>T(p.Lys3326*)

- Truncating variant- in a disease where this is a known mechanism - but the variant is in the last exon, towards the end- the last amino acid should be 3419
- Population data shows that the variant has been identified across all populations at a frequency of ~0.6%- much higher than the frequency of hereditary breast/ovarian cancer syndrome
- In some populations, frequency >1%, defining this as a polymorphism
- More recently, literature shows this variant impacts HR ability, although does not abolish it
- May be associated with specific cancers (i.e. not necessarily breast or ovarian)
- Large study of ~400,000 cases and controls suggests it is associated with cancers with environmental genotoxic risk factors

Somatic genomics is different

- The goal of interpretation of somatic genomic variants is to guide/change patient management
- Unlike germline variants, won’t necessarily see variant allele fractions of 50% or 100%, although that’s possible
- Have issues of sensitivity that are different from germline analysis
- Issues of tumour heterogeneity as well

Rationale for developing a new scheme for somatic variant interpretation



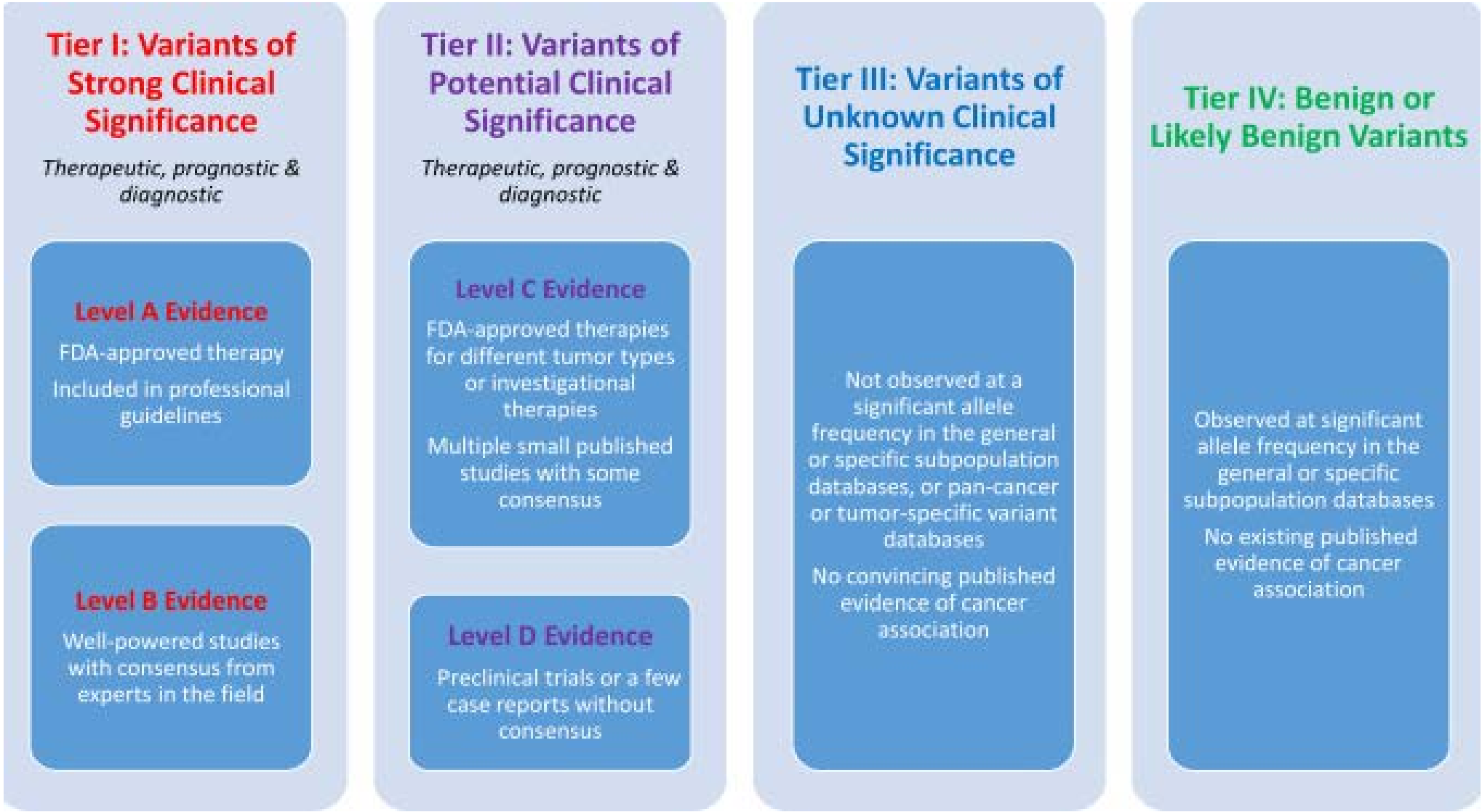
AMP/ASCO/CAP levels of evidence

Table 3 Categories of Clinical and/or Experimental Evidence

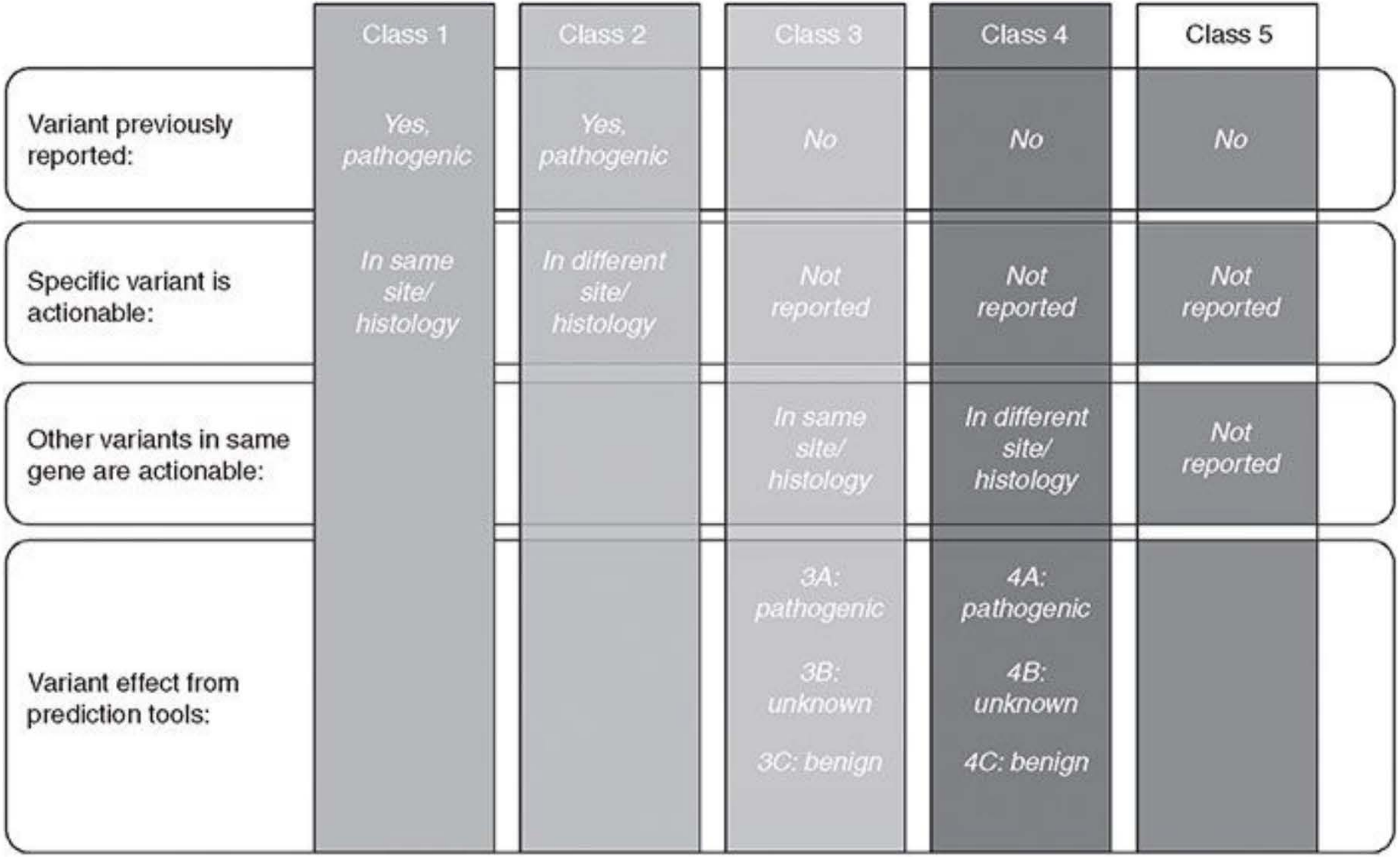
Category	Therapeutic	Diagnosis	Prognosis
Level A	1. Biomarkers that predict response or resistance to FDA-approved therapies for a specific type of tumor 2. Biomarkers included in professional guidelines that predict response or resistance to therapies for a specific type of tumor	Biomarkers included in professional guidelines as diagnostic for a specific type of tumor	Biomarkers included in professional guidelines as prognostic for a specific type of tumor
Level B	Biomarkers that predict response or resistance to therapies for a specific type of tumor based on well-powered studies with consensus from experts in the field	Biomarkers of diagnostic significance for a specific type of tumor based on well-powered studies with consensus from experts in the field	Biomarkers of prognostic significance for a specific type of tumor based on well-powered studies with consensus from experts in the field
Level C	1. Biomarkers that predict response or resistance to therapies approved by the FDA or professional societies for a different type of tumor 2. Biomarkers that serve as inclusion criteria for clinical trials	Biomarkers of diagnostic significance based on the results of multiple small studies	Biomarkers of prognostic significance based on the results of multiple small studies
Level D	Biomarkers that show plausible therapeutic significance based on preclinical studies	Biomarkers that may assist disease diagnosis themselves or along with other biomarkers based on small studies or a few case reports	Biomarkers that may assist disease prognosis themselves or along with other biomarkers based on small studies or a few case reports

Taken from Li et al 2017 JMD 19:4

AMP/ASCO/CAP Interpretation scheme



Taken from Li et al 2017 JMD 19:4



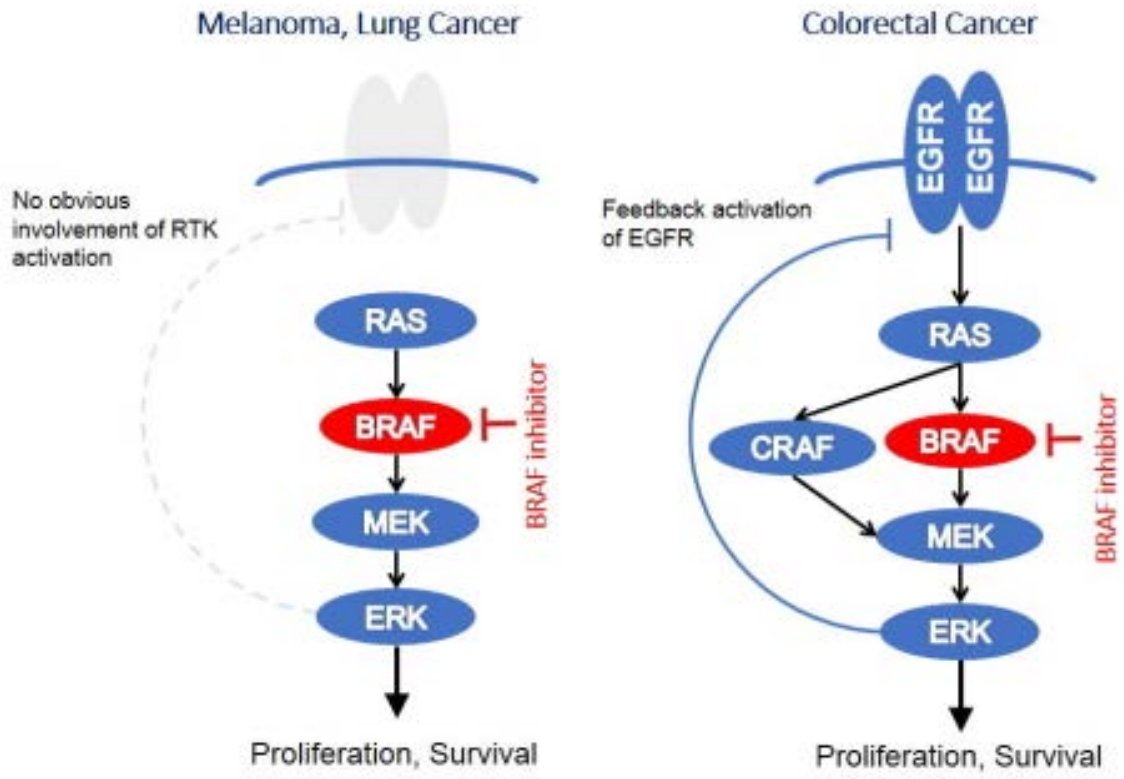
Taken from Sukhai et al 2016 Genet Med 18:128

NM_004333.4(BRAF):c.1799T>A(p.Val600Glu)

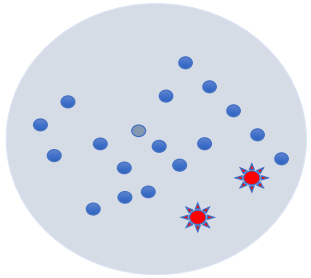
Obvious clinically relevant call in the context of melanoma

But what about other tumours?

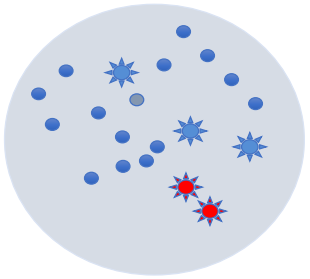
Also need to consider the proportion of the case that was tumour cells versus the variant allele fraction seen in the analysis



Beyond variant interpretation



10% tumour cells
All are heterozygous for variant of interest
Variant is present at 5%



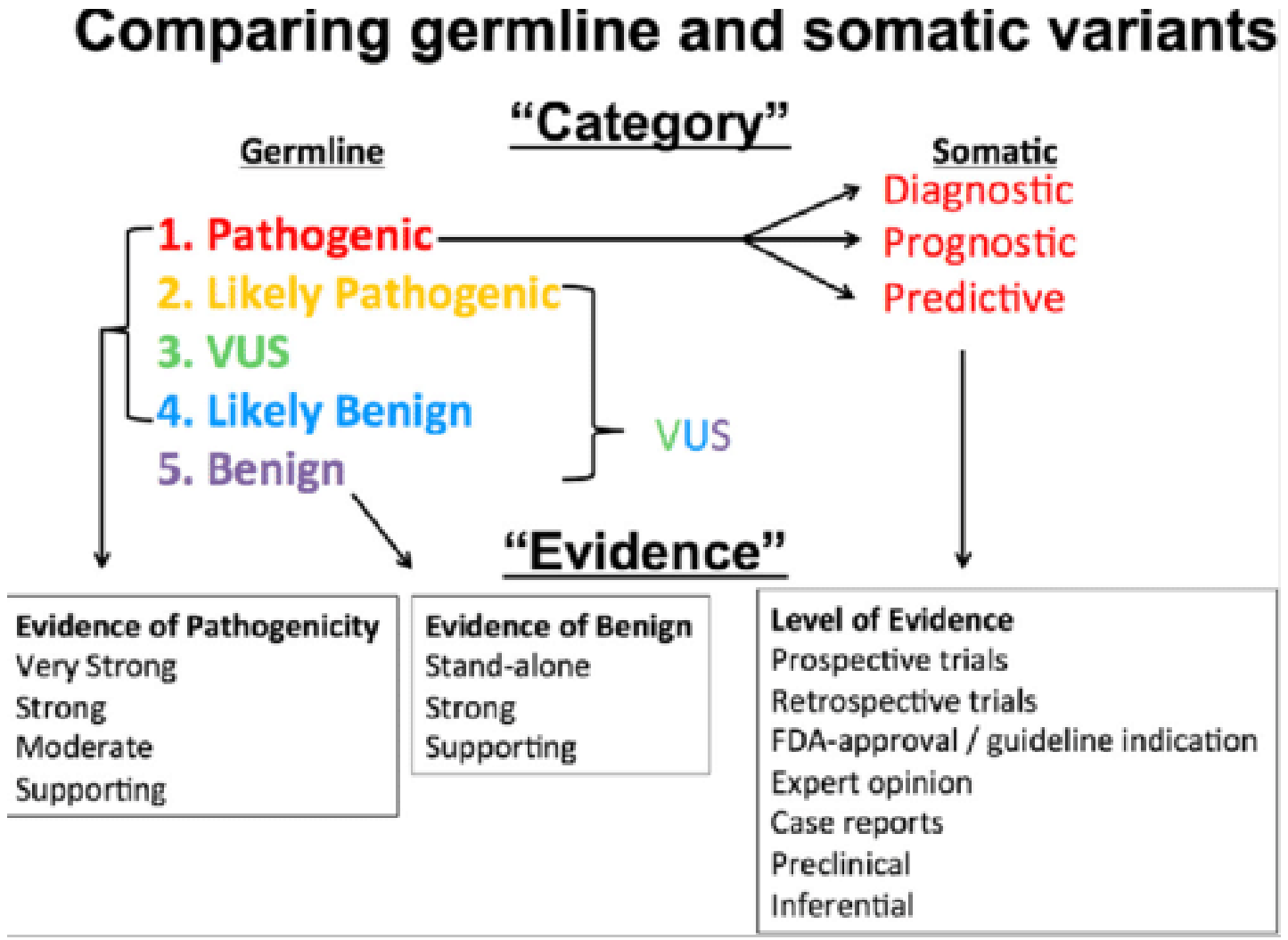
25% tumour cells
40% of them are heterozygous for variant of interest
Variant is present at 5%

Established the assay with a limit of detection for an SNV of 5% at a depth of 400

Identify an EGFR L858R variant present at 5% VAF

Both patients eligible for targeted therapy, but will both respond equally well? If information not conveyed, missing opportunity to look for other strategies? If tumour % estimate way off, is this predictive information less useful?

When two worlds collide

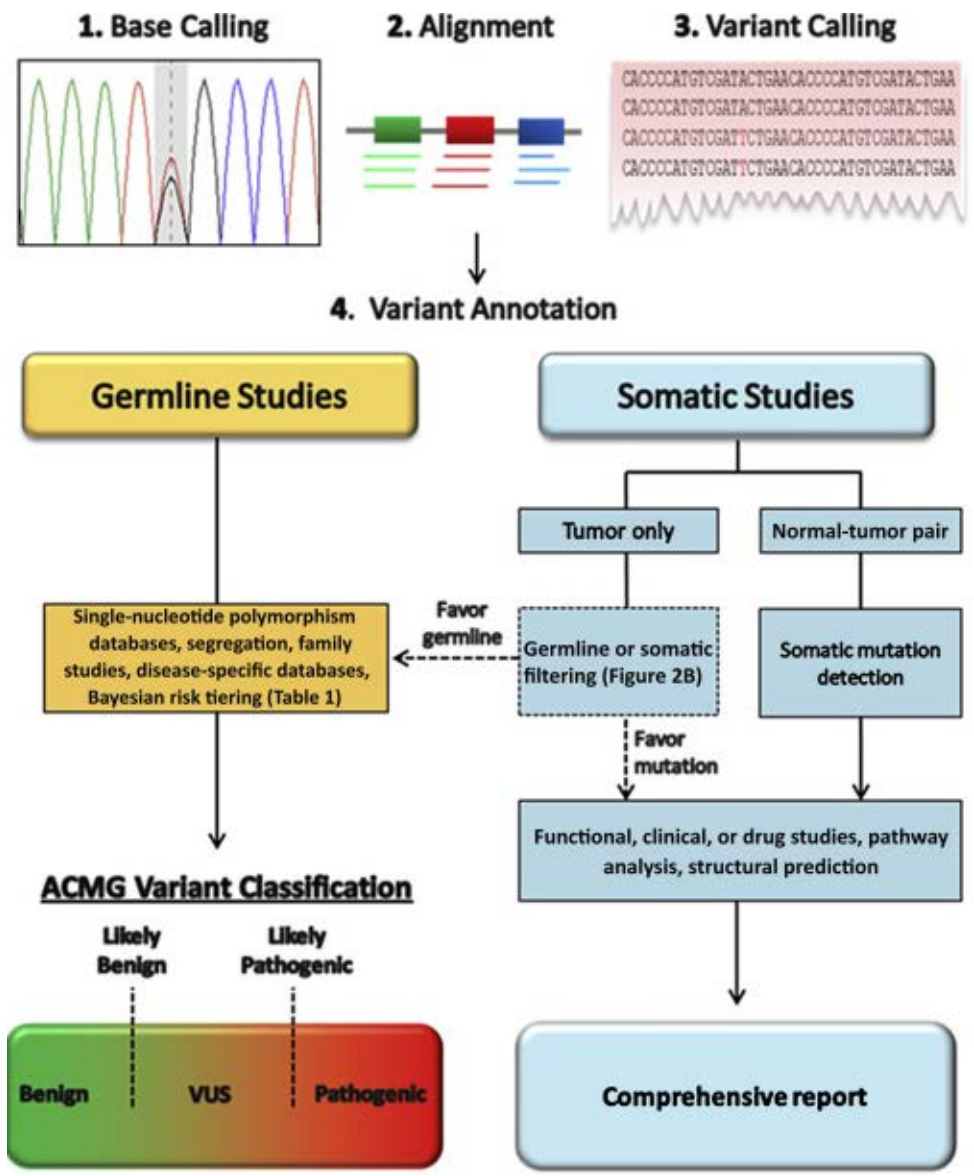


Summary

Interpretation of germline variants generally uses a 5 category schema

All variants are assessed using a combination of population, computational, literature and physical data

Variants are generally associated with diagnosis or prediction of risk- to that extent variants are interpreted in the context of disease



Taken from Lee et al 2015 JMD 17:339

Interpretation of somatic variants is more complex in the sense that they may be more difficult to identify

They require interpretation in the context of clinical management

They require interpretation in the context of disease