Several international projects compile information on the effects of genomic variants, based on findings continuously reported by clinical and preclinical studies. Most of these knowledgebases provide this information through publicly accessible websites.

In this hands-on, we will first explore two widely used knowledgebases: **ClinVar** and **OncoKB**.

ClinVar

Some databases are mainly focused on the relevance of **germline variants** and their potential pathogenicity—namely, their role in causing (or predisposing) individuals to disease. One such database is **ClinVar** (https://www.ncbi.nlm.nih.gov/clinvar/), a resource developed by the U.S. **National Center for Biotechnology Information (NCBI)**.

Although ClinVar has recently begun including annotations related to the oncogenicity of tumor somatic variants in specific cancer types, it is still primarily used to review evidence regarding the potential of germline variants for leading to a pathogenic phenotype. Of note, this includes variants associated with cancer predisposition—i.e. germline mutations that increase the risk of developing cancer.

ClinVar is a community-based resource in which registered users can submit their assertions regarding the pathogenicity of a given variant. The classification provided in the assertion must follow the ACMG/AMP framework, which includes germline phenotype-specific criteria, such as how the variant segregates across related (un)affected individuals. These user assertions are not independently reviewed, but ClinVar provides an aggregated quality score (see here). This score (named as '**review status**' in the database, and represented by a four star-tier system) takes into account, among other factors, the extent to which multiple submitters have provided consistent classifications for the same variant and/or whether the submitter is a recognized expert panel.

In ClinVar, you can search at the **gene level,** meaning that you can explore all variants with available assertions in the database for that gene. To do so, simply type the gene symbol in the search bar. You can then use the filter options on the left side of the results page to narrow down the displayed variants.

On the other hand, you can also direcly search for a **specific variant.** To do so, you can type the variant in the search bar using various variant nomenclatures (see here). Unfortunately, note that one common headache in the field is that different resources may require different formats to query the variant data.

Exercise:

Search for the following variant in ClinVar (stated as cDNA syntax):

TP53:c.818G>A (see here)

→ Discuss the following– what is the 'level of trust' that you would have to call this variant as pathogenic according to ClinVar available assertions? Which is the specific supporting evidence to classify the variant as such by the corresponding assertion(s) (you check these details in the 'more information' column of a given submission)?

Do the same for:

BRCA2:c.9285C>A (see here)

BRCA1:c.71G>C (see here)

OncoKB

Some databases are mainly focused on the relevance of cancer mutations, i.e. their potential to drive tumorigenesis and/or their associated clinical actionability for diagnosis, prognosis and therapy selection. One such database is **OncoKB** (https://www.oncokb.org/), developed by the **Memorial Sloan Kettering Cancer Center**.

OncoKB is curated by an in-house team. The content of the database is freely available for research use but requires a paid license for clinical applications. The database uses its own classification framework both to determine whether a variant is tumorigenic (categorized as *oncogenic* or *likely oncogenic*) and to rank its associated actionability, if applicable (see here and here).

In OncoKB, you can search at the **gene level**, allowing you to explore all variants with available annotations for that gene. Simply enter the gene symbol in the search bar. The resulting assertions are organized into several tables:

- **Annotated Alterations**: Variants with evidence of tumorigenicity (aka oncogenicity)
- **Therapeutic**: Variants with predictive value for treatment response
- **Diagnostic**: Variants with diagnostic significance
- **Prognostic**: Variants with implications for prognosis

Recently, OncoKB also introduced a separate section gathering annotations approved by the FDA.

In addition, you can also search directly for a **specific variant** in OncoKB. To do so, you can type the variant in the search bar by using different nomenclatures; again, it's important to note that naming conventions may vary across databases. In the case of this database, you can search for variants by typing the gene symbol and the protein change separated by a space (e.g. “BRAF V600E” or “EFGR T790M”).

Exercise:

Check the annotations available for the BRAF gene, and discuss the following:

→ In the ‘Therapeutic” table, what do you think that the term ‘fusions’ means here? i.e. does it mean that any fusion involving BRAF that is observed in gliomas should match with the Tovorafenib biomarker?

→ Similarly, what ‘oncogenic mutations’ means here?

Bonus: why do you think that the biomarkers for BRAF mutations are separated between BRAF V600E/K mutations and other BRAF mutations?

Check for the “FGFR2-PDHX Fusion” – how confident you would be to match it with a FGFR inhibitor when it is found in a cholangiocarcinoma as compared when it is found in a esophagus carcinoma?

==========

THE MOLECULAR TUMOR BOARD PORTAL

Variant annotation requires the combination of multiple bioinformatic tools and databases. This process can be time-consuming and prone to inaccuracies/errors when performed manually. **Clinical Decision Support Systems (CDSS)** can automate much of this workflow, including the application of predefined criteria to associate variant annotations with clinical actions of interest, such as ranking drug biomarkers by their clinical evidence or matching with eligibility criteria for clinical trials.

In brief, the input for a CDSS is a list of variants detected by a given NGS diagnostic assay, along with relevant case information (e.g., cancer type). The output is typically a report and/or data file that includes the resulting variants annotation and classification, along with all the supporting references and contextual information.

While several commercial CDSS platforms are available, larger academic centers can also develop their own in-house solutions to better suit their specific clinical and research needs. One such academic solution is the **Molecular Tumor Board Portal (MTBP)**, developed at **Karolinska Institutet** and used in various clinical initiatives.

A publicly accessible version of MTBP is available at http://mtbp.org. This public resource is a **simplified version** of the full production system used in clinical settings. Although it does not include certain features—such as assessment of variant clonality or clinical trial matching—it provides general interpretations of the analyzed variants for **research purposes only**.

With the public MTBP, you can analyse mutations, copy number alterations and/or gene fusions by typing the corresponding alterations in the free text box according to various variant syntax standards. Alternatively, you can also upload a VCF file with mutations.

The MTBP will provide an HTML interactive report as a result, with all the annotation to support:

(1) A classification of the **functional relevance** of the variant, based on three sources of evidence:

(a) variant effects **curated** from knowledge bases used by the MTBP;

(b) variant effects that are **assumed** according to *bona fide* predefined criteria; or

(c) bioinformatic **predictions**, which represent the lowest level of evidence.

(2) An assessment of the **actionability** of the variant, using a modified ESMO–ESCAT scale to tier the clinical relevance of using the variant as a biomarker for the cancer type in the sample under analysis.

All the elements that appear in blue opens popup with additional details. Note that the MTBP reports provides two main views:

- **Alteration centric view:** in this view, gene alterations are grouped in three different tables according to their functional relevance classification (functionally relevant, neutral or unclassifed). Therefore, a given alteration can appear only once in this view. All the biomarkers associated with that alteration (if any) are aggregated according to their actionability tier in the last column.

- **Biomarker centric view:** in this view, gene alterations are grouped in different tables according to the actionability tier of their associated biomarker(s) (if any). Therefore, a given alteration can appear in more than one table if associated with biomarkers of multiple actionability tiers.

Exercise- case 1:

Use the MTBP to discuss drivers and associated clinical actions for the following variants detected by a NGS assay performed in a tumor tissue sample (with blood sample - paired control) of a 44yo Serous Ovarian Cancer female before first-line treatment:

BRCA2:p.Lys2008Ter

BRCA2:p.Cys2689Phe

CDKN2A:DEL

KIT:p.Gln922Leu

APC:p.Arg1171Cys

PTEN:p.His397LeufsTer19

Exercise- case 2:

Discuss drivers and associated clinical actions for the following somatic variants detected by a NGS assay performed in a tumor tissue sample (without a blood-sample paired control) of a 58yo EGFRmut Non-Small Cell Lung Carcinoma male after relapse to 1st generation EGFR inhibitor treatment:

EGFR:p.Leu858Arg

EGFR:p.Thr790Met

TP53:p.Asp186Glu

PTEN:DEL

PALB2:p.Lys30Asn

Exercise- case 3:

A 68yo Colon Adenocarcinoma male, originally diagnosed as RAS wild-type, currently relapsed to EGFR-inhibitor therapy with multiple metastases; discuss drivers and associated clinical actions for the following variants detected by a circulating tumor DNA assay:

APC:p.Arg141Ser

KRAS:p.Gln22Glu

KRAS:p.Gly12Glu

MYC:AMP

PIK3CA:p.Asn157Asp

TSC1:p.Gly1089Ala

Exercise- case 4:

The patient of “case 1” received PARP-inhibitor therapy after diagnosis, with good initial response, but currently shows progressive disease. Therefore, a tumor circulating DNA assay is performed, which detect the same mutations than in the index case plus the following inframe deletion:

BRCA2:p.Ser2001_Asp2023del

Do you think that this new BRCA2 variant has any relevance?