

OncoKB Curation Standard Operating Procedure

**Version 1.0
February 2020
OncoKB.org**

Table of Contents

Chapter 1: Introduction	5
OncoKB Oversight and Governance	5
OncoKB Staff	5
OncoKB Data Sources	6
OncoKB Access	6
Conflicts of Interest	6
Chapter 2: OncoKB Concepts	8
I. Concepts in OncoKB	8
II. The OncoKB Curation Platform	9
A. Genes Homepage	9
B. Tools	10
C. Feedback	11
Chapter 3: Gene Curation	12
I. Gene Summary	12
II. Gene Background	12
III. Gene Classification	12
A. Oncogene	12
B. Tumor Suppressor	12
Chapter 4: Alteration Curation	14
I. Nomenclature and Technical Rules for Alteration Curation	14
A. General Curation Rules	14
B. Missense Mutations	14
C. Truncating Mutations	14
D. Fusions	15
E. Copy Number Aberrations	15
F. In-frame Deletions or Insertions	15
G. Oncogenic Mutations	16
H. Tumor Suppressors and Oncogenes	16
I. Hard-coded alteration names	16
J. Hotspot Mutations	16
II. Evidence-based Alteration Curation	17
A. Oncogenic Effect	17
B. Biological Effect	18
C. Description of Evidence	19
D. Additional Information (Optional)	20
III. Tumor Type Curation	20
Tumor Type	20
Chapter 5: Curation of Tumor Type-Specific Clinical Implications	21
I. Clinical Summary	22

II.	Diagnostic Implications	23
A.	Level of Evidence	23
B.	Description of Evidence	23
C.	Additional Information (Optional)	23
III.	Prognostic Implications	23
A.	Level of Evidence	23
B.	Description of Evidence	24
C.	Additional Information (Optional)	24
IV.	Standard Implications for Sensitivity to Therapy	24
A.	Highest Level of Evidence	24
B.	Level of Evidence for Alternate Oncogenic Alleles	24
C.	Level of Evidence in Other Tumor Types	24
D.	Description of Evidence	24
E.	Additional Information (Optional)	25
V.	Standard Implications for Resistance to Therapy	25
A.	Level R1	25
B.	Description of Evidence	25
C.	Level R1 does not use the Alternate Oncogenic Alleles and Other Tumor Types rules.	25
D.	Additional Information (Optional)	25
VI.	Investigational Implications for Sensitivity to Therapy	25
A.	Highest Level of Evidence	25
B.	Level of Evidence for Alternate Oncogenic Alleles	26
C.	Level of Evidence in Other Tumor Types	26
D.	Description of Evidence	26
E.	Additional Information (Optional)	26
F.	Alterations that are Level 1, 2A or 3A in another tumor type are considered Level 2B or 3B in the indicated tumor type	26
VII.	Investigational Implications for Resistance to Therapy	26
A.	Level R2	26
B.	Description of Evidence	27
C.	Level R2 does not use the Alternate Alleles and Other Tumor Types rules.	27
D.	Additional Information (Optional)	27
VIII.	Variants of Unknown Significance (VUS)	27
IX.	Reanalysis and Reevaluation	28
Chapter 6: Annotation of Variants in Patient Tumor Samples		29
I.	Variant Annotation Process	29
A.	Match gene.	29
B.	Retrieve gene summary.	29
C.	Match curated alterations.	29
D.	Retrieve mutation summary, oncogenic and biological effects for the alteration.	29
E.	Match curated tumor types.	29

F.	Retrieve tumor type summary and clinical implications.	29
II.	Nomenclature and Rules Related to Annotation	31
A.	OncoKB Cancer Gene List	31
B.	Curated Genes	31
C.	Matched Genes	31
D.	Matched Curated Alterations	31
1.	Overall Matching Logic	31
2.	Special Rules for Alterations	31
E.	Alternate Oncogenic Alleles	32
F.	Hotspots	32
G.	Matched Curated Tumor Types	32
H.	OncoKB Therapeutic Implication Levels of Evidence Priorities	32
III.	Annotation Summaries	32
A.	Gene summary	32
B.	Variant Biological Summary	33
C.	Clinical Summary	33
D.	Resistance Mutations	34
Chapter 7: OncoKB Data Access		35
I.	The OncoKB API	35
II.	The OncoKB Website: www.oncokb.org	35
A.	Data summary	35
B.	Search bar	35
C.	Levels of Evidence	36
D.	Actionable Genes	36
E.	Data Access	36
G.	News	37
H.	Usage Terms	38
I.	OncoKB Cancer Gene List	38
J.	About OncoKB	39
K.	Team	39
L.	Gene Pages	40
M.	Alteration Pages	43
III.	OncoKB content accessible through cBioPortal	46
A.	OncoKB icons in cBioPortal	46
B.	OncoKB Cards in cBioPortal	46
APPENDIX		49
REFERENCES		51

Chapter 1: Introduction

OncoKB is a Precision Oncology Knowledgebase that contains information about the biological effects and treatment implications of specific cancer genes and their somatic alterations. OncoKB is developed and maintained by the Knowledge Systems group in the Marie Josée and Henry R. Kravis Center for Molecular Oncology at the Memorial Sloan Kettering Cancer Center (MSK).

In OncoKB, genes are classified as either oncogenes or tumor suppressors based on the curated evidence. Alterations included in OncoKB are genetic changes that arise as a result of DNA-level variants in cancer: non-synonymous mutations, translocations, rearrangements / fusions, copy number amplifications and deletions. This document uses “Alterations”, “Mutations” and “Variants” interchangeably. All alterations in OncoKB are classified according to 1) their oncogenic effect and 2) their biological effect, based on the curated evidence (discussed in Chapter 2). The oncogenic effect of an alteration is an evidence-based assertion that classifies whether the mutation is oncogenic, likely oncogenic, neutral or inconclusive. The biological effect of an alteration is an evidence-based assertion that classifies whether the mutation is gain-of-function, loss-of-function, neutral or inconclusive.

If a cancer alteration in OncoKB is associated with clinical implications, these implications are also curated in OncoKB (discussed in Chapter 2). Alterations with clinical implications are further assigned a Therapeutic (Chakravarty et al. 2017), Diagnostic and/or Prognostic level of evidence. Each Level of Evidence assignment in OncoKB defines the strength of the evidence that supports the alteration as being a diagnostic, prognostic or therapeutic biomarker.

OncoKB Oversight and Governance

Oversight and governance of OncoKB is under the purview of the Lead Scientist and the Clinical Genomics Annotation Committee (CGAC). The Lead Scientist and CGAC are responsible for establishing standards and oversight of all processes in the scope of OncoKB. CGAC provides expertise in cancer variant interpretation, and, in particular, the assignment of the OncoKB Levels of Evidence to specific alterations. CGAC is comprised of “Core” members and “Extended” members. Core CGAC members guide OncoKB development, are at the forefront of clinical management and research and have translational cancer biology expertise in their respective major disease entities. Extended members are selected physicians and scientists who represent the broader MSK clinical leadership across departments and services, including service chiefs, physicians with clinical expertise in their fields, and scientists with specific gene or pathway expertise. Core members, in addition to responding to requests regarding clinical consensus, also maintain an active and responsive dialogue with the Lead Scientist, providing insight or updates regarding genomic biomarker-based clinical data.

OncoKB Staff

The OncoKB staff consists of the following:

1. *The OncoKB Lead Scientist* creates and maintains general oversight and governance procedures for the OncoKB staff including the development, approval, and coordination of all variant assessment activities. The Lead Scientist also liaises between the variant curation processes and their oversight and governance by CGAC. The OncoKB Lead Scientist does not have any relevant conflicts of interest.
2. *The Scientific Content Management Team (SCMT)*, which is comprised of the following: 1) OncoKB Scientists: Two Ph.D-level scientists with translational cancer biology expertise that provide day-to-day guidance and management of the OncoKB Curators regarding appropriate curation, editorial and scientific content review; 2) Lead Software Engineer: Executes database governance and data preservation as well as feature development and maintenance of the OncoKB Curation Platform (curation platform); 3) Lead OncoKB Data Curator: Liaises between the Lead Software Engineer and OncoKB Scientists to ensure seamless data maintenance, updates and access, and is responsible for database operations. No member of the SCMT has any relevant conflicts of interest.
3. *OncoKB Curators* comprise of pre-doctoral graduate students, postdoctoral fellows and clinical fellows. They assess and curate alterations, their biological effects, and associated treatment implications in cancer in compliance with the procedures described by the OncoKB SOP. OncoKB Curators are specifically trained in evaluating evidence from various sources and entering appropriate information into the curation platform.

OncoKB Data Sources

Four primary data sources are used to identify and curate cancer variants and their biological and clinical therapeutic implications (Fig. 1):

1. Public cancer variant databases of alterations identified in tumor sequencing studies, e.g., cBioPortal and COSMIC (Catalogue of Somatic Mutations in Cancer).
2. Statistically significant and recurrent variants identified based on 24,592 sequenced tumors using methods described in Chang et al., 2018.
3. Disease-specific treatment guidelines such as those provided by the National Cancer Compendium Network (NCCN) and proceedings of major scientific and/or clinical conferences such as the American Society of Clinical Oncology (ASCO) and the American Association of Cancer Research (AACR).
4. General scientific literature accessed through PubMed.

When external databases (such as the IARC TP53 Database, the ARUP Laboratories BRCA1 and BRCA2 Mutation Database) are used as references for curation, OncoKB does not import data from these external databases directly into its database. Rather, variants from these sources are used mainly as candidate variants for independent curation in OncoKB. Therefore, variant data from external databases is versioned using the same protocols in place for all OncoKB variants.

OncoKB Access

Data from OncoKB is used in four ways (Fig. 1):

1. OncoKB data is publicly available for personal and research purposes through an interactive website at www.oncokb.org. Usage terms of OncoKB are specified at <https://www.oncokb.org/terms> (Fig. 19).
2. The curated data is also available programmatically through the OncoKB application program interface (API). The different ways to access OncoKB data are documented at www.oncokb.org/DataAccess (Fig 17).
3. The cBioPortal for Cancer Genomics (<https://www.cbioperl.org>) uses the OncoKB API for annotating cancer variants in its database.
4. OncoKB data is used to annotate the patient reports of the results from MSK-IMPACT, a targeted tumor sequencing test available to MSK patients.

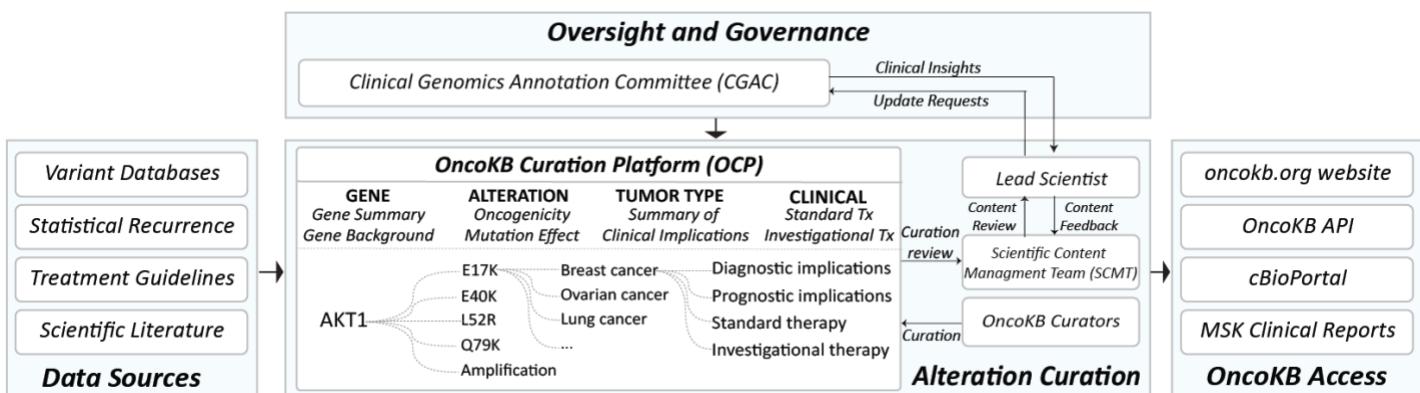


Fig. 1: Summary of OncoKB processes. The schematic shows a summary of the data sources, knowledgebase architecture and processes that compose the OncoKB workflow.

Conflicts of Interest

Evidence-based assertions of the oncogenic and biological effect of an alteration are not considered to be subject to any conflicts of interest. The evidence used to support specific assertions of oncogenic and biological effects are displayed on the website and link to the appropriate reference in PubMed or to the abstract website. Variant assertions are re-analyzed and reevaluated by the OncoKB team in specific review cycles (see Chapter 5, Section X and Table 1) and any new content or inconsistencies are corrected at that time. Additionally, feedback regarding updated content or inconsistencies from users of OncoKB either through the website or via cBioPortal are addressed within 48 hours of receipt (see Chapter 2, Section II.C and Chapter 7, Sections II.M.11 and III.B.6).

A subset of alterations in OncoKB are considered biomarkers that are predictive to response to certain drugs. Some of these drugs are FDA-approved and the biomarker is a consideration in standard care. In these cases, the biomarker is associated with either Level of Evidence 1 or 2A (refer to Chapter 5 and Fig. 7). However, some of these drugs are either 1) FDA-approved, but the biomarker is in an off-label setting or 2) not FDA-approved and instead are being tested in clinical trials, and for these, conflicts of interest may arise. In both of the latter scenarios, the biomarkers and drugs are considered investigational and are associated with a Level of Evidence 2B, 3A or 4 (refer to Chapter 5 and Fig. 7).

To address and resolve potential COI, any new level assignments or changes to an existing level have to be approved unanimously by all CGAC members. To this end, we have implemented the following process: CGAC members are responsible for advising the OncoKB team and entering into consensus regarding the assignation of a level of evidence to a biomarker that is predictive of response to either standard care or investigational drug. Requests for advice and consensus from CGAC occur in the form of periodic emails from the Lead Scientist to all CGAC members and are typically prompted by new FDA-approvals, FDA-breakthrough designations, or newly reported results of major clinical trials from clinical oncology conferences or publications. Consensus emails have the following structure: 1) A statement describing the reason for a proposed new assignation of a level of evidence to an alteration or for changing the current level of evidence for a specific alteration and consequent change to OncoKB data. 2) A summary of the clinical data supporting the proposed assignation of a Level of Evidence to a specific alteration. 3) A sample Clinical Summary that includes the new OncoKB statement that is prompted by the new clinical data. 4) A request for feedback regarding the change to OncoKB data, in the form of a response within two weeks of receipt of the request. 5) A statement that should no feedback be received by the set deadline two weeks from the date of request, approval of the change in OncoKB annotation is inferred and subsequently included into the OncoKB database, the outputs of which will be seen in the Clinical Summaries in the website, the cBioPortal and the MSK-IMPACT reports (see OncoKB Access below).

Financial conflicts of interest for all OncoKB personnel including CGAC are disclosed publicly on the OncoKB website, www.oncokb.org/team (Fig. 22) and reported in publications or in conferences as appropriate. In the event of a conflict of interest arising for a specific CGAC member with regards to a Level of Evidence assignation, he or she is asked to recuse themselves from the consensus request. In the event that consensus cannot be immediately reached, the Lead Scientist is responsible for mediating between conflicting advice to resolve any discrepancy. Should consensus not be reached, the proposed change in the Level of Evidence is rejected.

Chapter 2: OncoKB Concepts

I. Concepts in OncoKB

To curate the clinical implications associated with an alteration in OncoKB in a structured way, each clinical implication must be associated with a specific gene, one or multiple alterations, and one or multiple tumor types. The following is the nested organization of key concepts for each gene in OncoKB (Fig. 2):

Gene

1. Summary
2. Background
3. Alteration
 - i. Mutation Effect
 - ii. Tumor type
- Clinical Implications
 1. Diagnostic Implications
 2. Prognostic Implications
 3. Therapeutic Implications
 4. Standard Sensitivity
 5. Standard Resistance
 6. Investigational Sensitivity
 7. Investigational Resistance
4. Variants of Unknown Significance

The screenshot shows the OncoKB interface for the gene A123B. At the top, there are tabs for 'Review', 'Citations', and 'Download PDF'. Below the tabs, the gene name 'A123B' is shown with a mutation count of '1x TT'. The 'Summary' section is collapsed. The 'Background' section is collapsed. The 'Alteration' section is expanded, showing a mutation entry for 'A123B > 1x TT'. This entry includes a 'Mutation Effect' section with 'No Entry' highlighted in orange, and dropdown menus for 'Oncogenic' and 'Mutation effect' (both with 'No Entry'). Below this is a 'Description of Evidence' and an 'Additional Information (Optional)' field, both of which are collapsed. The 'Tumor type' section is expanded, showing 'All Tumors' selected. This section includes a 'Tumor Type Summary (Optional)' and a list of clinical implications: 'Diagnostic implications: No Entry', 'Prognostic implications: No Entry', 'Standard implications for sensitivity to therapy: No Entry', 'Standard implications for resistance to therapy: No Entry', 'Investigational implications for sensitivity to therapy: No Entry', and 'Investigational implications for resistance to therapy: No Entry'. At the bottom, there are fields for 'Add tumor type(s)', 'Cancer Type' (dropdown), 'Subtype' (dropdown), 'Mutation Name' (text input), and 'Add Mutation' (button). A footer at the bottom says 'Variants of Unknown Significance (Investigated and data not found)' with buttons for 'Variant Name' and '+ Add Variant'.

Fig. 2: OncoKB is hierarchically organized by its key concepts. Any clinical implication, including drugs that show activity in tumors carrying a specific mutation, is always nested under a specific Mutation and Tumor type within a gene.

II. The OncoKB Curation Platform

Variant information is entered into the OncoKB curation platform, a custom web-based application that allows manual curation and review of variant information. All information entered into the curation platform are structured in a hierarchy of gene, alteration, tumor type and clinical implications. The latter include diagnostic, prognostic, and therapeutic implications. The OncoKB Lead Scientist requests periodic disease-specific content updates from individual CGAC members regarding genomic biomarker-based clinical data. The Lead Scientist also oversees and is responsible for all curation processes to ensure consistency and quality of variant curation and assertions by OncoKB Curators and curation review by SCMT. Addition of new or changes to the existing clinical implications in OncoKB may be prompted by new FDA approvals, FDA-breakthrough designations, and newly reported results of major clinical trials from clinical oncology conferences or publications, requiring clinical consensus among all members of CGAC. CGAC consensus feedback, clinical insights and recommendations are communicated to the Lead Scientist, then conveyed to the SCMT, and subsequently incorporated into OncoKB by the SCMT. All new content (including any changes or additions) is reviewed by the SCMT in the Review interface of the curation platform (Fig. 3). New content is reviewed, approved and incorporated into the OncoKB Database. The Lead Scientist receives weekly summaries from the SCMT about new content and prioritizes weekly curation deliverables.

(a) Curator Interface:

Gene: EGFR Last edit was made on Jan 11, 5:59 PM 2019 by Debayan Chakravarty. Last update to database was made on Dec 5, 3:29 PM by Moran Nissan.

Summary: EGFR, a receptor tyrosine kinase, is altered by amplification and/or mutation in lung and brain cancers among others.

Background: EGFR (Epidermal Growth Factor Receptor) is a transmembrane receptor that is activated by EGFR family extracellular ligands (PMID: 24691165). EGFR is a member of the EGFR family of receptors, including the receptors ERBB2, ERBB3, and ERBB4. Binding of EGFR by its ligands, including EGFR ligands and transforming growth factor alpha (TGFα), activates downstream signaling pathways including the canonical MAPK and PI3K/AKT/mTOR signaling cascades (PMID: 22239436). EGFR can homodimerize or heterodimerize with other EGFR family members to initiate signaling (PMID: 25621509). Activation of EGFR-mediated signaling ultimately results in cellular proliferation, migration, and differentiation (PMID: 18045542). While EGFR is usually expressed at low levels in normal adult tissues, hyperactivation of this receptor by somatic mutations and/or amplification of the EGFR gene is now considered to be a key driver of cancer development, particularly in non-small cell lung cancer (NSCLC) (PMID: 18045542). In addition to the constitutively active form of the receptor that is sensitive to EGFR tyrosine kinase inhibitors (PMID: 18324119), Tyrosine kinase inhibitors targeting EGFR, including afatinib, erlotinib, and gefitinib, have been approved for first-line treatment of non-small cell lung cancer patients (PMID: 20906020). Osimertinib is a second-line tyrosine kinase inhibitor that has been FDA approved for relapsed patients with non-small cell lung cancer with the EGFR resistance mutations T790M, L858R, and exon 19 deletions (PMID: 27923840). Additionally, copy number amplification of the EGFR gene result in receptor overexpression in several cancer types, including brain and colorectal cancers, and these cancers may also be sensitive to EGFR inhibition (PMID: 11426640).

Publication IDs: PMID:24691165 PMID:22239436 PMID:25621509 PMID:18045542 PMID:10886430 PMID:17318210 PMID:16977817 PMID:24688098 PMID:2603630 PMID:23983089 PMID:23983040 PMID:11426640

cBioPortal link: <https://cbioportal.mskcc.org/lncg/EGFR>

COSMIC link: <http://cancer.sanger.ac.uk/cosmic/gene/overview?ln=EGFR>

(b) Lead Scientist/SCMT Interface:

Gene: EGFR Last edit was made on Jan 11, 5:59 PM 2019 by Debayan Chakravarty. Last update to database was made on Dec 5, 3:29 PM by Moran Nissan. **Debayan Chakravarty is reviewing this gene.**

You are currently in "Review" mode. Click the "Review Complete" button to exit.

Accept All Changes from Kinisha Gala Accept All Changes from Debayan Chakravarty Accept All Changes from Sarah Phillips

Mutation: T790M

Tumor type: Non-Small Cell Lung Cancer

Standard implications for sensitivity to therapy:

Therapy: Osimertinib

Description of Evidence: Updated by Kinisha Gala at Jan 7, 12:22 AM 2019 ✓ X

New Content:

Osimertinib is third generation EGFR tyrosine kinase inhibitor (TKI) that inhibits T790M-mutant EGFR and is FDA-approved for the treatment of patients with metastatic EGFR T790M mutation-positive non-small cell lung cancer (NSCLC) who have progressed on prior EGFR TKI therapy. FDA-approval is based on the results of the Phase I AURA study of osimertinib in 127 patients with T790M mutation-positive NSCLC (PMID: 25923549) and the Phase II AURA2 study of osimertinib in 210 patients with T790M mutation-positive NSCLC (PMID: 27751847). In the Phase I dose-escalation and dose-expansion studies, the response rate was 61% (95% CI 52-70) among patients with T790M mutations, with a median progression-free survival (PFS) of 9.6 months (95% CI 8.3-11) versus 2.8 months (95% CI 2.1-4.3) in patients without T790M mutations (PMID: 25923549). In the Phase II single-arm cohort of patients with T790M-positive NSCLC who progressed on prior EGFR TKI therapy, 106 of 198 patients (53%) achieved a partial response or better, with a median PFS of 9.8 months (95% CI 8.5-12.3) (PMID: 27751847). Since its FDA-approval, Osimertinib has also been administered in patients with EGFR exon 19 deletion or L858R mutations-positive NSCLC. Osimertinib showed significantly longer PFS compared with a control in

Fig. 3: Curation Platform Review Interface. (a) The curation platform interface for curators. (b) The curation platform interface for the Lead Scientist and SCMT with administrative privileges including “Review interface” for reviewing and approving new content curated by OncoKB Curators.

The OncoKB Curation Interface Homepage is divided into the following pages:

A. Genes Homepage

The Genes page (Fig. 4) is displayed upon entering the OncoKB curation interface and is the main homepage of the curation interface. This page lists all genes (linking to its own Gene Curation Page) in the OncoKB curation system, along with the following information for each gene:

1. *Last modified*: Timestamp indicating when the Gene Curation Page was last modified
2. *Last modified by*: Name of the last user to edit the page
3. *Needs to be reviewed*: Indicates if there is new content in the Gene Curation Page that needs to be reviewed by the SCMT.
4. *Search Box*: Allows the user to search for their gene of interest.

Gene	Last modified	Last modified by	Needs to be reviewed	# of articles to curate
EGFR	Jan 11, 5:59 PM 2019	Debyani Chakravarty	No	3
RET	Dec 19, 12:18 PM 2018	Sarah Phillips	No	2
BRAF	Dec 18, 10:28 AM 2018	Sarah Phillips	Yes	1
BRD4	Jun 7, 1:58 PM 2018	Moriah Nissan	No	1
CARD11	Jan 17, 1:32 PM 2019	Hannah Johnsen	Yes	1
EPHA3	Oct 17, 11:56 AM 2018	Moriah Nissan	No	1
EPHA5	Feb 22, 2:19 PM 2018	Sarah Phillips	No	1
ERBB2	Jan 16, 2:06 PM 2019	Sarah Phillips	Yes	1
KEAP1	Jul 2, 1:18 PM 2018	Moriah Nissan	No	1
NF1	Jan 16, 11:53 AM 2019	Sarah Phillips	No	1

Fig. 4: Gene homepage in the OncoKB Curation Platform. The Genes homepage lists all genes in the curation system.

B. Tools

The purpose of the Tools page is to provide data validation checks to the SCMT (Fig. 5a). This page is divided into several sections:

1. *Review History:* Allows the SCMT to visualize reviewed changes made to a specific Gene Page. Once a gene is specified, the following outputs are displayed:
 - a. Gene Name: The name of the queried gene
 - b. Reviewed By: The SCMT member who reviewed the data in question
 - c. Records: The specific section within the Gene Page that was reviewed by the SCMT member (i.e., Background, Mutation Effect), and the action taken (Added, Deleted or Updated)
 Each discrete piece of reviewed data within a Gene Page is displayed as its own entry.
2. *Query Reviewed Data:* Allows the SCMT to visualize the following outputs in a table format. These outputs are chosen from a drop-down list and can be downloaded as an XLS file by clicking the 'Download' button.
 - a. Oncogene/Tumor Suppressor: Lists all genes in OncoKB and their classification as an oncogene or tumor suppressor. The table also indicates whether the following alterations are curated for each gene: Truncating Mutations, Deletion, and Amplification.
 - b. Mutation Effect: Lists all alterations in OncoKB and provides the following information (extracted from the database) for each alteration: Associated Gene, Oncogenic Effect, Mutation Effect, Description of Mutation Effect, Citations.
 - c. Tumor Type Summary: Lists all Tumor Type Summaries in OncoKB and indicates the gene-alteration-tumor type combination for which they are associated.
 - d. Therapeutics: Lists all alterations associated with a Level of Evidence in OncoKB and provides the following information (extracted from the database) for each alteration: Associated Level of Evidence, Therapeutic, Therapeutic Description of Evidence.
3. *Additional Validation Checks:* SCMT can also query the following two validation questions:
 - a. Are truncating mutations curated for tumor suppressor genes?

This query returns a list of genes in OncoKB that have Truncating Mutations curated as an alteration but are not marked as Tumor Suppressors.
 - b. Do all tumor suppressor genes have truncating mutations curated?

This query returns a list of genes in OncoKB that are marked as Tumor Suppressors but do not have Truncating Mutations curated as an alteration. For some tumor suppressor genes, such as POLE, truncating mutations are purposely not curated as they have no oncogenic effect. However, for the majority of tumor suppressors, truncating mutations are assumed to result in the loss-of-function of the protein and therefore considered oncogenic. Exceptions apply here as well, like in the case of BRCA2, where truncating mutations close to the C-terminus, such as K3326*, are known not to have an inactivating effect.

C. Feedback

The purpose of the Feedback page is to collate all user feedback received about specific OncoKB annotations from a feedback form within the cBioPortal. The feedback form in cBioPortal is also described in Chapter 4, Section III, B.4.f. In brief, the feedback form records the following user inputs (if applicable): gene, alteration, feedback, reference(s), user email address, and cBioPortal link. The Feedback page in the curation platform includes a “Complete” column, in which SCMT members can add the status of the response to the feedback, and a “Comments” column, in which SCMT members can add notes or comments regarding the feedback (Fig. 5b).

(a)

Review History

Genes: Select Some Options Include UUID

Date:

Type: update name change add delete

Query Reviewed Data

Query Type:

Are all truncating mutations curated under tumor suppressor genes?

Do all tumor suppressor genes have truncating mutation curated?

(b)

OncoKB Annotation Feedback (Responses)

File Edit View Insert Format Data Tools Form Add-ons Help

	Timestamp	Gene	Alteration	Feedback
1	12/18/2018 16:00:16	NRAS	Q61K	Not oncogenic? Alternate allele?
2	10/21/2018 6:59:20	AKT3	V722I	Polymorphism? ES7_Pas / 1000G_Freq
3	4/28/2018 7:14:23	ERBB2	L755S	The mutation also appears to be resistant to lapatinib, compared to neratinib.
4	4/28/2018 7:14:23	BRAF	R462T	This variant is in COSMIC 5 times ... and I think described here: http://www.ccl.com/cancer-cell/tumors/S1103-6108(14)00299-2
5	4/28/2018 7:14:23	CREBBP	R144E	This is a highly recurrent mutation.
6	4/28/2018 7:14:23	CREBBP	R144E	I found this: http://www.nature.com/nature/journal/v471/n7337/full/nature09727_F1.html
7	4/28/2018 7:14:23	CREBBP	R144E	CREBBP R144E (equivalent to EP300 R141Q) contacts phosphate of the CoA moiety of the inhibitor (salt bridges are shown)
8	10/27/2018 19:27:10	MYB	MYB-NPB Fusion	Can we add MYB as a new gene, only to annotate the fusions. A good description is here: https://www.acrf.org/research/the-myc-oncogene-project
9	10/27/2018 19:27:10	MYB	MYB1	MYB1 also needs to be annotated.
10	4/28/2018 9:28:51	AR	H87Y	This mutation is missing annotation. It is in My Cancer Genome.
11	1/25/2019 22:14:51	AR	L70H	Also missing and in My Cancer Genome. Who curated AR?
12	1/15/2019 14:04:24	CDH1	A0000000000000013	CDH1 missense mutations should be annotated as oncogenic
13	1/15/2019 14:04:24	PHOX2A	E45K	This hotspot mutation should be annotated as oncogenic
14	1/15/2019 14:05:57	KIF3A	G35V	This is a common variant that needs to be added. Also known as Q34.
15	1/15/2019 14:43:40	NFI	M203I/R747	NFI truncating mutations should be annotated as oncogenic.
16	1/16/2019 14:28:19	NFI	P184L/T194M	This should not be annotated as nononcotic.

+ Form Responses 1 Export

Fig. 5: OncoKB Curation Platform Tools and Feedback Pages. (a) Includes ability to look up curation review history, query specific data and check the annotation of tumor suppressor genes. (b) All feedback received through cBioPortal is fed to a Google sheet that is accessible through the Curation Platform.

Chapter 3: Gene Curation

OncoKB uses the following standardizations for each gene:

- The HUGO gene symbols are used for gene names. We update to the latest HUGO symbols periodically.
- For each gene, a canonical transcript is selected for annotation. Both Ensembl and RefSeq transcript IDs are provided per gene.

The OncoKB Gene Curation Page contains the biological and clinical implications of each gene and its alterations. The Gene Curation Page contains the following sections (ordered by the hierarchy specified in the concept hierarchy section II):

I. Gene Summary

Provides a brief overview of the gene and its role in cancer. This section is free text and contains a 1-2 sentence summary. For the majority of genes, the summary is one sentence that describes the gene function and the cancer types in which it is most frequently altered, e.g., “EGFR, a receptor tyrosine kinase, is altered by amplification and/or mutation in lung and brain cancers among others.”

II. Gene Background

Provides a detailed overview of the biological function of the gene/protein in the normal cell, its role in cancer development and progression, and its clinical significance. The background section is free text and contains 6-10 sentences, although some genes with little published information may have shorter background sections. The background should contain sufficient detail to thoroughly explain the above-mentioned information but should not include minute details and extraneous information. The references used in this section should primarily come from high impact journals (i.e., New England Journal of Medicine, Journal of Clinical Oncology, Journal of Clinical Investigation, Cell, Cancer Discovery, Science, Nature, etc.).

III. Gene Classification

Genes in OncoKB can be classified as oncogenes (e.g., BRAF), tumor suppressors (e.g., PTEN) or both (e.g., NOTCH1), using the following criteria:

A. Oncogene

In OncoKB, an oncogene may be defined by any of the following four criteria:

1. (1) A cancer-inducing gene. (2) A gene that can transform cells (Weinberg, p.G:20, 2014).
2. A gene that, when activated by mutation, increases the selective growth advantage of the cell in which it resides (Vogelstein et al., 2013).
3. A gene that, in tumor samples, has i) higher functional impact and higher amplification frequency in comparison to those observed in neutral genes, and ii) lower loss-of-function mutations, splicing mutations and frequency of deletions and increased frequency of amplification compared to tumor suppressors (Davoli et al., 2013).
4. A gene that is a mutated (changed) form of a gene involved in normal cell growth. Oncogenes may cause the growth of cancer cells. Mutations in genes that become oncogenes can be inherited or caused by being exposed to substances in the environment that cause cancer (NCI Dictionary of Cancer Terms, cited 2019, Feb 01).

B. Tumor Suppressor

In OncoKB, a tumor suppressor may be defined by any of the following four criteria:

1. (1) A gene whose partial or complete inactivation, occurring in either the germline or the genome of a somatic cell, leads to an increased likelihood of cancer development. (2) Such a gene that is responsible for constraining cell proliferation. (3) Also known as a gatekeeper, a gene that operates to hinder cell multiplication or to further cell differentiation or cell death and in this way prevents the appearance of populations of neoplastic cells (Weinberg, p.G:28, 2014).
2. A gene that, when inactivated by mutation, increases the selective growth advantage of the cell in which it resides. Mutated through protein-truncating alterations throughout their length (Vogelstein et al., 2013).
3. A gene that, in tumor samples, has i) higher frequencies of loss-of-function and splicing mutations, higher functional impact, and higher frequency of deletions compared to those found in neutral genes, and ii)

- higher frequencies of loss-of-function and splicing mutations, higher deletion frequency and lower amplification frequency compared to those found in oncogenes (Davoli et al., 2013).
4. A type of gene that makes a protein called a tumor suppressor protein that helps control cell growth. Mutations (changes in DNA) in tumor suppressor genes may lead to cancer. Also called antioncogene (NCI Dictionary of Cancer Terms, cited 2019, Feb 01).

There are two check boxes under the gene summary with which a curator may assign whether the gene is an oncogene and/or a tumor suppressor. If there is no clear evidence that a gene falls into either category based on the criteria defined by OncoKB (see above), both boxes may be left unchecked.

Chapter 4: Alteration Curation

I. Nomenclature and Technical Rules for Alteration Curation

Specific nomenclature when curating alterations in OncoKB must be used to allow for seamless annotation of variants with its oncogenic and biological effects and clinical implications when using the OncoKB API.

A. General Curation Rules

1. Multiple mutations may be grouped together (comma separated) for curation of shared clinical implications and/or tumor type summaries. The oncogenic and mutation effect of each of the mutations should be curated separately.
2. Mutation ranges, which capture all amino acid substitutions in a specified amino acid range, can be used (e.g., TP53_102_292mis [TP53 DNA binding domain mutations], KIT_C788_N828mut [KIT Exon 17 non-truncating mutations]). Mutation ranges must have an associated oncogenic effect, mutation effect, and description of evidence based on the available evidence. Clinical implications and/or tumor type summaries can also be curated under mutational ranges.
3. **Alteration Codes** – the following are codes that can be used for naming alterations in the OncoKB curation platform:
 - a. mis = missense mutation - e.g., 102_292mis [DNA binding domain missense mutations]
 - b. dup = duplication of a specified range - e.g., S501_A502dup
 - c. del = in-frame deletion of a specified range - e.g., P551_E554del
 - d. ins = in-frame insertion - e.g., W557_V559delinsC; e.g. T574insTQLPYD
 - e. delins = in-frame alteration - whether it's in-frame insertion or deletion, will be interpreted by the number of amino acid changes. e.g., V600_K601delinsE = inframe deletion - e.g., R435_K436delinsKKR = in-frame insertion
 - f. nontrunc = any non-truncating mutation - e.g., R449_E514 nontrunc
 - g. fs = frameshift - e.g., N457Mfs*22
 - h. _splice = splice mutations - e.g., X963_D1010splice or X963_splice
 - i. trunc = truncating mutation - e.g., D286_L292trunc
 - j. 1? = start lost - e.g., M1?
 - k. * = stop gained - e.g., R2019*

4. **Brackets and Parentheses in the Mutation Header**

- a. Square Brackets [] - used in the mutation header to rename a curated alteration. For example, to curate a specific insertion, amino acid positions are written in the mutation header to indicate the protein change (e.g., 729_761ins). However, for the purpose of displaying this alteration on the OncoKB website, the SCMT may want to refer to this alteration as "Exon 19 insertion". By using square brackets in the mutation header as follows: "729_761ins [Exon 19 insertion]", the OncoKB website will display the alteration as "Exon 19 insertion" instead of 729_761ins.
- b. Parentheses () - used in the mutation header to leave comments. Any text in () in the mutation header is for administrative purposes only and can only be viewed within the OncoKB curation interface. It will not affect the output of how a mutation is displayed on any output platform (cBioPortal, MSK-IMPACT Reports or OncoKB Website).

B. Missense Mutations

1. The naming convention for missense mutations is <ref_allele><position><tumor_allele> (e.g., V600E)
2. Every missense mutation needs to be separately curated with respect to its oncogenic and mutation effect.
3. Positional variants, which capture all amino acid substitutions at a given position, can be used for curation of shared clinical implications and/or tumor type summaries (e.g., KRAS G12, BRAF V600). Positional variants do not include curation of oncogenic effect or mutation effect, as this information should be captured under each allele-specific missense mutation for which there is functional data.

C. Truncating Mutations

"Truncating Mutations" can be curated as a specific alteration within a Gene Page. "Truncating Mutations" must have an associated oncogenic effect, mutation effect, and description of evidence.

1. Since "Truncating Mutations" captures all truncating alterations within the gene (some of which have not been functionally characterized), its oncogenic and mutation effect should be marked as "Likely Oncogenic" and "Likely Loss of Function" respectively.

2. Clinical implications and/or tumor type summaries can also be curated under “Truncating Mutations.”
3. The oncogenic effect, mutation effect and clinical implications associated with “Truncating Mutations” can be limited by defining a range for the truncation (e.g., “CCND1 256_286trunc [C Terminal Truncating Mutations]”). Truncating mutations outside this range will not be associated with the designated oncogenic effect, mutation effect and clinical implication of those in the defined range.
4. “Truncating Mutations” include the following based on the [Sequence Ontology](#):
 - a. Stop_lost: A sequence variant where at least one base of the terminator codon (stop) is changed, resulting in an elongated transcript
 - b. Start_lost: A codon variant that changes at least one base of the canonical start codon
 - c. Stop_gained: A sequence variant where at least one base of a codon is changed, resulting in a premature stop codon and leading to a shortened transcript
 - d. TFBS_ablation: A feature ablation where the deleted region includes a transcription factor binding site
 - e. Feature_truncation: A sequence variant that causes the reduction of a genomic feature, with regard to the reference sequence
 - f. Frameshift_variant: A sequence variant which causes a disruption of the translational reading frame, i.e., the number of nucleotides inserted or deleted is not a multiple of three
 - g. Transcript_ablation: A feature ablation whereby the deleted region includes a transcript feature
 - h. Splice_donor_variant: A splice variant that changes the 2 base region at the 5' end of an intron
 - i. Splice_region_variant: A sequence variant in which a change has occurred within the region of the splice site, either within 1-3 bases of the exon or 3-8 bases of the intron
 - j. Stop_retained_variant: A sequence variant where at least one base in the terminator codon is changed, but the terminator remains
 - k. Splice_acceptor_variant: A splice variant that changes the 2 base region at the 3' end of an intron
 - l. Incomplete_terminal_codon_variant: A sequence variant where at least one base of the final codon of an incompletely annotated transcript is changed.

D. Fusions

“Fusions” can be curated as a specific gene alteration within a Gene Page, and include any fusion that involves the specified gene.

1. “Fusions” must have an associated oncogenic effect, mutation effect, and description of evidence.
2. Since “Fusions” captures all fusions within the gene (some of which have not been functionally characterized), its oncogenic and mutation effect should be marked as “Likely Oncogenic” and “Likely Gain of Function” respectively.
3. Clinical implications and/or tumor type summaries can also be curated under “Fusions.”
4. Specific fusions, in which both fusion partners are specified, can be curated as separate alterations if there is functional evidence in the literature describing their oncogenic and/or mutation effect (e.g., “EML4-ALK fusion”). The oncogenic effect, mutation effect, and clinical implications of the specific fusion alteration will be prioritized over those of the “Fusions” alteration.
5. Although a specific fusion names two gene partners, the alteration is only curated in one Gene Page - the gene that is the main driver (or hypothesized to be the main driver) of the fusion oncoprotein (e.g., BCR-ABL1 is curated in the ABL1 Gene Page).

E. Copy Number Aberrations

“Amplification” and “Deletion” can be curated as specific gene alterations within a Gene Page if appropriate functional data exists:

1. “Amplification” and “Deletion” must have an associated oncogenic effect, mutation effect, and description of evidence.
2. Prognostic implications, clinical implications and/or tumor type summaries can also be curated under “Amplification” and “Deletion.”

F. In-frame Deletions or Insertions

In-frame deletions or insertions can be curated as a specific gene alteration within a Gene Page (see section IV.E.1).

1. “del” = in-frame deletion (e.g., P551_E554del, P191del)
2. “ins” = in-frame insertion (e.g., T574insTQLPYD)
3. “delins” = a specified in-frame alteration. Whether the alteration is an in-frame deletion or in-frame insertion is determined by the specified number of amino acid changes. For example:

- a. *V600_K601delinsE* is an in-frame deletion because the number of amino acids deleted (2) is greater than the number of amino acids inserted (1).
- b. *R435_K436delinsKKR* is an in-frame insertion because the number of amino acids inserted (3) is greater than the number of amino acids deleted (1).
- 4. Each curated alteration must have an associated oncogenic effect, mutation effect, and description of evidence.
- 5. Clinical implications and/or tumor type summaries can also be curated under an in-frame deletion or insertion.

G. Oncogenic Mutations

“Oncogenic Mutations” can be curated as a specific gene alteration within a Gene Page.

- 1. “Oncogenic Mutations” is used when there is tumor-specific information that applies to ALL functional (oncogenic/likely oncogenic) alterations within a Gene Page. The tumor-specific information will automatically get linked to all mutations in the Gene Page that have the “Yes” or “Likely” boxes checked next to the Oncogenic label.
- 2. “Oncogenic Mutations” does not include curation of oncogenic effect, mutation effect, and description of evidence, as this information should be captured under each individual variant in the Gene Page for which “Oncogenic Mutations” applies.
- 3. If a gene has “Amplification” curated as “Oncogenic” or “Likely Oncogenic”, this alteration will NOT be associated with the tumor-type specific information under “Oncogenic Mutations.”

H. Tumor Suppressors and Oncogenes

For genes marked as Tumor Suppressors:

- 1. The alteration “Truncating Mutations” should be curated.
- 2. The alteration “Deletion” may be curated, but this is dependent on the data available in the literature.
- 3. For Oncogenes: Truncating Mutations in oncogenes are often nonfunctional/not oncogenic. However, there are some examples in which they are functional including the genes CCND1 and CALRX. In these cases, truncating mutations in the protein are often activating via loss of C-terminal negative regulatory domains and in these cases, truncating mutations are restricted to a specific range.

I. Hard-coded alteration names

Alterations that do not follow the above nomenclature are not supported unless they are hard coded. Examples of such alterations include:

- 1. FLT3: internal tandem duplication
- 2. EGFR: vIII
- 3. EGFR: Kinase domain duplication
- 4. EGFR: C-terminal domain

J. Hotspot Mutations

Mutational hotspots are defined as mutant residues arising more frequently than expected in the absence of selection based on the analysis by Chang et al., 2016 and Chang et al., 2018. In this analysis 24,592 cancers including 10,336 prospectively sequenced patients with advanced disease were analyzed, and the authors identified 1,165 statistically significant missense or in-frame insertion or deletion hotspot mutations, of which 80% arose in 1 in 1,000 or fewer patients.

- 1. If there is functional data in the literature describing the oncogenic and/or mutation effect of an allele-specific hotspot, the hotspot should be curated as an individual variant within the appropriate Gene Page.
- 2. Curated hotspots must have an associated oncogenic effect, mutation effect, and description of evidence based on the available evidence.
- 3. Alternate oncogenic alleles (which are not themselves specifically curated, see **Chapter 6.I**) of a curated hotspot (parent hotspot allele) will be associated with an oncogenic and mutation effect according to the rules outlined in A.2 above.
- 4. If no allele-specific variants are curated for a hotspot (including if variants are only located in the VUS section of the Gene Page), the hotspot’s oncogenic effect will be automatically designated as “predicted oncogenic” in any output platform (cBioPortal, MSK-IMPACT Reports or OncoKB website).

II. Evidence-based Alteration Curation

Alterations included in OncoKB are genetic changes that arise as a result of DNA-level variants in cancer: non-synonymous mutations, translocations, rearrangements / fusions, copy number amplifications and deletions. This document uses “alterations”, “mutations” and “variants” interchangeably. OncoKB describes alterations by their effect on the protein and not at the DNA level. All alterations in OncoKB are classified according to 1) their oncogenic effect and 2) their biological effect, based on the curated evidence (Fig. 6).

The oncogenic and biological effects of a mutation are curated based on the properties of transformed cells described in the second edition of “The Biology of Cancer” by Robert Weinberg and the hallmarks of cancer described by Douglas Hanahan and Robert Weinberg in their manuscript “Hallmarks of cancer: the next generation.” published in Cell in 2011 (Hanahan and Weinberg, 2011).

The screenshot shows the OncoKB curation interface for a BRAF V600E mutation. At the top, it says "Mutation: V600E". Below that, under "Mutation Effect", there is a section titled "Oncogenic" with checkboxes for "Yes" (checked), "Likely", "Likely Neutral", and "Inconclusive". A red box highlights the "Yes" checkbox, and a red arrow points to the text "Oncogenic Effect". Below this, a section titled "Mutation effect" contains checkboxes for "Gain-of-function" (checked), "Likely Gain-of-function", "Loss-of-function", "Likely Loss-of-function", "Switch-of-function", "Likely Switch-of-function", "Neutral", "Likely Neutral", and "Inconclusive". A red box highlights the "Gain-of-function" checkbox, and a red arrow points to the text "Biological Effect". Further down, there is a "Description of Evidence" section with a detailed paragraph about the BRAF V600E mutation. At the bottom, it lists publication IDs and an "Additional Information (Optional)" section.

Fig. 6: Curation of the Oncogenic and Biological effects of an alteration in OncoKB. An alteration is described by two assertions: 1) The Oncogenic Effect of the mutation and 2) The Biological Effect of the mutation. Every variant in OncoKB must be curated with both of these assertions or placed in the Variants of Unknown Significance section of the curation platform. Otherwise entry of the variant is not allowed into the OncoKB database.

A. Oncogenic Effect

In OncoKB, “oncogenic” is defined as “referring to the ability to induce or cause cancer” as described in the second edition of The Biology of Cancer by Robert Weinberg (2014). OncoKB distinguishes between five possible evidence-based assertions to describe the oncogenic effect conferred by the alteration when it is present in cells. An alteration may be asserted as oncogenic, likely oncogenic, likely neutral or inconclusive based on the following criteria:

1. **Oncogenic**

Evidence shows the alteration promotes cell proliferation or other hallmark of cancer as defined by Douglas Hanahan and Robert Weinberg (Hanahan and Weinberg, 2011).

- Multiple experimental studies (in one or multiple publications) provide evidence that the alteration is oncogenic.
- The alteration is a known hotspot (Chang et al., 2018) AND there is at least one experimental study suggesting the alteration is oncogenic.
- The alteration has been identified in a patient who responded to a targeted inhibitor, AND at least one experimental study provides strong evidence that the alteration is oncogenic.
- The alteration is classified as either gain/loss-of-function, likely gain/loss-of-function, switch-of-function or likely switch-of-function AND there is at least one experimental study suggesting the alteration is oncogenic.

2. **Likely Oncogenic**

Evidence suggests the alteration likely promotes cell proliferation or other hallmark of cancer as defined by Douglas Hanahan and Robert Weinberg (Hanahan and Weinberg, 2011).

- a. At least one experimental study provides reasonable evidence suggesting the alteration is oncogenic.
- b. The alteration is a known hotspot (Chang et al., 2016; Chang et al., 2018), AND there are no known functional studies describing the oncogenic potential of the alteration.
- c. The alteration is classified as either gain/loss-of-function or switch-of-function or likely gain/loss-of-function, or likely switch-of-function AND there are no known functional studies describing the oncogenic potential of the alteration.

3. *Likely Neutral*

Evidence suggests the alteration does not alter protein activity or does not confer growth or survival advantage when expressed in cells.

- a. The mutation effect of the alteration is neutral or likely neutral.
- b. At least one experimental study provides reasonable evidence suggesting the alteration is likely neutral.

4. *Inconclusive*

There is conflicting and/or weak data describing the oncogenic effect of the mutant alteration

- a. Conflicting data exists as to the oncogenic effect of the alteration.
- b. Data is limited to “weak” experimental data describing the oncogenic effect of the alteration (small, under-powered experimental studies in one or multiple publications).
- c. Data is limited to studies demonstrating either patient and/or in vitro sensitivity/resistance to a targeted drug.
- d. Data is limited to in silico studies that predict the oncogenic effect of the alteration.

B. Biological Effect

In OncoKB, the Biological Effect is defined as the biological effect of a mutation/alteration on the protein function that gives rise to changes in the biological properties of cells expressing the mutant/ altered protein compared to cells expressing the wildtype protein.

- Transformed cells are characterized by the following properties (Weinberg,p.82, Table 3.2, 2014):
 - Altered morphology (rounded shape, refractile in phase-contrast microscope)
 - Loss of contact inhibition (ability to grow over one another)
 - Anchorage independence (ability to grow without attachment to solid substrate)
 - Ability to proliferate indefinitely
 - Reduced requirement of mitogenic growth factors
 - High saturation density (ability to accumulate large numbers of cells in culture dish)
 - Inability to halt proliferation in response to deprivation of growth factors
 - Increase transport of glucose
 - Tumorigenicity (ability to form tumors *in vivo* following injection into appropriate host animals)
- The hallmarks of cancer comprise the biological capabilities acquired during the multistep development of human tumors. Mutations when expressed in cells may exhibit any one of these hallmarks of cancer in cells expressing the altered protein. Published experimental measurement of any of one these hallmarks of cancer may be taken as evidence that the mutation is oncogenic:
 - Sustaining proliferative signaling
 - Evading growth suppressors
 - Resisting cell death
 - Enabling replicative immortality
 - Inducing angiogenesis
 - Activating invasion and metastasis
 - Genome instability and mutation
 - Tumor-promoting inflammation
 - Deregulated cellular energetics
 - Evading immune destruction

OncoKB distinguishes between five possible evidence-based assertions to describe the biological effect conferred by the alteration when it is present in cells. An alteration may be asserted as known or likely gain-, loss-, or switch-of-function, neutral, likely neutral, or inconclusive based on the following criteria.

1. *Known Gain/Loss/Switch-of-function*

- a. **Gain-of-function:** Strong evidence-based data demonstrating that the alteration increases the function of the protein

- b. **Loss-of-function:** Strong evidence-based data demonstrating that the alteration decreases the function of the protein
 - c. **Neutral:** Strong evidence-based data demonstrating that the function of the protein is unchanged by the alteration
 - d. **Switch-of-function:** Strong evidence-based data demonstrating that the alteration causes the protein to acquire a new function
 - e. **Rules for classifying an alteration with a known function**
 - i. Multiple experimental studies (in one or multiple publications) provide evidence that the alteration confers gain-, loss-, or switch-of-function.
 - ii. The alteration is a known hotspot (Chang et al., 2016. Chang et al., 2018) AND at least one experimental study provides strong evidence that the alteration confers gain-, loss-, or switch-of-function.
 - iii. The alteration has been identified in a patient who responded to a targeted inhibitor AND at least one experimental study provides strong evidence that the alteration confers gain-, loss-, or switch-of-function.
- 2. **Likely Gain/Loss/Switch-of-function**
 - a. **Likely Gain-of-function:** Probable, possible, and/or evidence-based data suggesting that the alteration likely increases the function of the protein
 - b. **Likely Loss-of-function:** Probable, possible, and/or evidence-based data suggesting that the alteration likely decreases the function of the protein
 - c. **Likely Switch-of-function:** Probable, possible, and/or evidence-based data suggesting that the alteration likely causes the protein to acquire a new function
 - d. **Rules for classifying an alteration with a probable function**
 - i. At least one experimental study provides reasonable evidence suggesting the alteration confers gain-, loss-, or switch-of-function.
 - ii. The alteration is a known hotspot (Chang et al., 2016. Chang et al., 2018), and there are no known functional studies describing the mutation effect of the alteration.
 - iii. While conflicting evidence may exist, there is a reasonable assumption based on the data suggesting the alteration confers gain-, loss-, or switch-of-function.
- 3. **Neutral**

There is strong evidence the function of the protein is unchanged by the alteration and the following are the **rules for classifying an alteration as known Neutral function**

 - i. Strong evidence-based data demonstrating that there is no difference in measurable cell attributes expressing either the wildtype or mutant form of the gene.
 - ii. Rescue experiment provides strong evidence that the alteration is neutral.
- 4. **Likely Neutral**

Evidence suggests the alteration likely does not alter the function of the protein and the following are the **rules for classifying an alteration as likely Neutral function**

 - i. Probable, possible, and/or evidence-based data suggesting that there is no difference in measurable cell attributes expressing either the wildtype or mutant form of the gene.
- 5. **Inconclusive**

There is conflicting and/or weak data describing the mutation effect of the alteration and the following are the **rules for classifying an alteration as Inconclusive**

 - i. Conflicting data exists as to the mutational effect of the alteration.
 - ii. Data is limited to “weak” experimental data describing the mutational effect of the alteration (small, under-powered experimental studies in one or multiple publications).
 - iii. Data is limited to studies demonstrating patient and/or in vitro sensitivity/resistance to a drug.
 - iv. Data is limited to in silico studies that predict the mutation effect of the alteration.

C. Description of Evidence

Describes the functional evidence supporting the oncogenic and mutation effects of the alteration. This section is free text and typically contains 2-6 sentences. This section must contain supporting reference(s) from the literature (either a PMID referencing a peer-reviewed journal or an abstract from a scientific conference such as ASCO or AACR). Only the PubMed Identifiers (PMIDs) and abstract citations are made publicly available to OncoKB users. Curated assertions are for evaluation purposes by the SCMT and OncoKB Lead Scientist.

D. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT but may not necessarily be included in the final output. The information in this section will only be accessible from the OCP and therefore will not be displayed on downstream systems such as the OncoKB website and cBioPortal.

III. Tumor Type Curation

Tumor Type

Below each alteration in the curation interface, the user must choose one or multiple Tumor Type(s) for the purpose of curating alteration- and tumor type-specific clinical implications, if any. OncoKB uses OncoTree (<http://oncotree.mskcc.org>) to manage the vocabulary of tumor types. Currently OncoTree version 2018_06_15 is being used. The user may choose a main cancer type and/or subtype from the dropdown list. In addition to the Oncotree nodes, the dropdown list also contains the following categories:

- A. *All Solid Tumors*: Includes all solid tumors within the Oncotree
- B. *All Liquid Tumors*: Includes all liquid tumors (from the myeloid and lymphoid branches) within Oncotree
- C. *All Tumors*: Includes all solid and liquid tumors within the Oncotree
- D. *Other Tumors*: This tumor classification is a special case and is only utilized for the purpose of incorporating Tumor Type Summaries.

Chapter 5: Curation of Tumor Type-Specific Clinical Implications

A subset of alterations in OncoKB are considered biomarkers that are predictive to response to certain drugs. Some of these drugs are FDA-approved and the biomarker is a consideration in standard care. Alternatively, some of these drugs are either 1) FDA-approved, but the biomarker is in an off-label setting or 2) not FDA-approved and instead are being tested in clinical trials. In both of the latter scenarios, the biomarkers and drugs are considered investigational.

The original Levels of Evidence system was developed by OncoKB to rank the therapeutic implications associated with an alteration found in a patient tumor by the relative weight of the evidence (Fig. 7). For example, an alteration that is recognized by the FDA to be predictive of response to an FDA-approved drug would have a higher Level of Evidence (Level 1) compared to an alteration that has been shown in preclinical studies to be sensitizing to an investigational drug that's being tested in a clinical trial (Level 4). Accordingly, the highest levels of evidence, Levels 1 and 2A refer to the standard implications for sensitivity to an FDA-approved drug. Additionally, Level R1 refers to the standard implications for resistance to an FDA-approved drug. Levels 2B, 3A and 4 refer to the investigational implications for sensitivity to either an FDA-approved drug (in the off-label setting, Level 2B) or an investigational drug (Levels 3A and 4). Level R2 includes investigational implications for resistance to either an FDA-approved or investigational drug. Since the FDA does not endorse off-label use of drugs, the scope of FDA-recognition sought for the clinical implications of OncoKB is restricted for Level 1 (FDA-recognized variants that are biomarkers predictive of response to FDA-approved drugs) and Level 2A (NCCN-listed variants that are biomarkers predictive of response to FDA-approved drugs) variants only. Each of these different sets of clinical implications are described in greater detail in Sections IV to VII below.

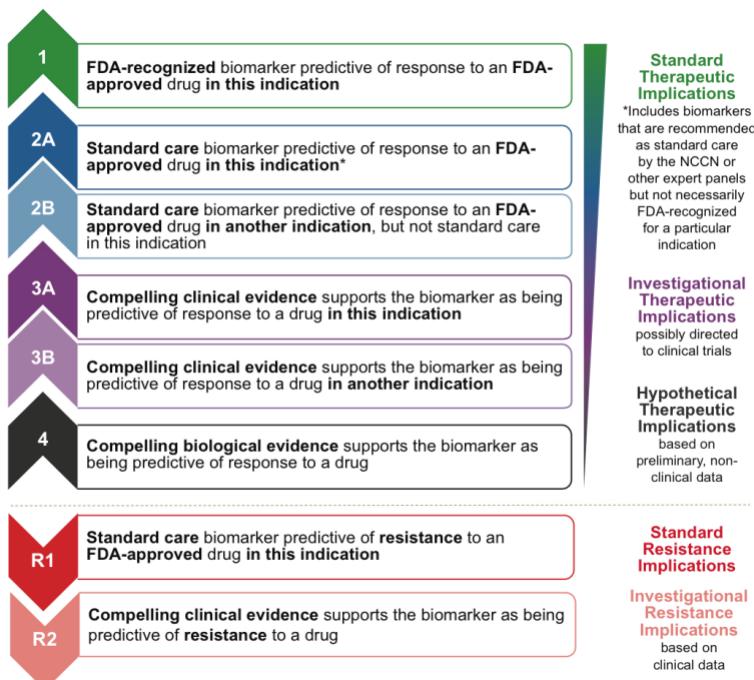


Fig. 7: OncoKB (Therapeutic) Levels of Evidence (Chakravarty et al. 2017).

Similarly, to rank the diagnostic and prognostic implications of an alteration found in a specific tumor type, the OncoKB Diagnostic and Prognostic Levels of Evidence schema were developed (see below, Fig. 9 and 10).

The potential clinical implications of the alteration meets the following criteria:

CGAC members are responsible for advising the OncoKB team and entering into consensus regarding the assignation of a level of evidence to a biomarker. Requests for advice and consensus from CGAC occur in the form of periodic emails from the Lead Scientist to all CGAC members and are typically prompted by new FDA-approvals, FDA-breakthrough designations, or newly reported results of major clinical trials from clinical oncology conferences or publications.

Consensus emails have the following structure:

1. A statement describing the reason for a proposed new assignation of a level of evidence to an alteration or for changing the current level of evidence for a specific alteration and consequent change to OncoKB data.

2. A summary of the clinical data supporting the proposed assignation of a Level of Evidence to a specific alteration.
3. A sample Clinical Summary that includes the new OncoKB statement that is prompted by the new clinical data.
4. A request for feedback regarding the change to OncoKB data, in the form of a response within two weeks of receipt of the request.
5. A statement that should no feedback be received by the set deadline two weeks from the date of request, approval of the change in OncoKB annotation is inferred and subsequently included into the OncoKB database, the outputs of which will be seen in the Clinical Summaries in the website, the cBioPortal and the MSK-IMPACT reports (see OncoKB Access below).

Financial conflicts of interest for all OncoKB personnel including CGAC are disclosed publicly on the OncoKB website, www.oncokb.org/team (Fig. 22) and reported in publications or in conferences as appropriate (described in detail in Chapter 1). In the event of a conflict of interest arising for a specific CGAC member with regards to a Level of Evidence assignation, he or she is asked to recuse themselves from the consensus request. In the event that consensus cannot be immediately reached, the Lead Scientist is responsible for mediating between conflicting advice to resolve any discrepancy. Should consensus not be reached, the proposed change in the Level of Evidence is rejected.

The clinical implications of an alteration may be curated in one or more of seven sections (summarized in Fig. 8):

1. Tumor Type Summary
2. Diagnostic Implications
3. Prognostic Implications
4. Standard Implications for Sensitivity to Therapy
5. Standard Implications for Resistance to Therapy
6. Investigational Implications for Sensitivity to Therapy
7. Investigational Implications for Resistance to Therapy

The screenshot shows a user interface for curating clinical implications. At the top, it says "Tumor type: Melanoma" with "1x TTS, 2x Level 1". Below this is a "Tumor Type Summary (Optional)" section with text about BRAF inhibitors. To the right, a vertical red box highlights the "Clinical Implications" section, which contains expandable sections for "Diagnostic implications", "Prognostic implications", "Standard implications for sensitivity to therapy", "Standard implications for resistance to therapy", "Investigational implications for sensitivity to therapy", and "Investigational implications for resistance to therapy".

Fig 8. The Clinical Implications of the mutation. If a mutation has a clinical implication, it is described within the context of the tumor type in which the clinical implication is relevant. If a mutation has a diagnostic clinical implication, it must be associated with a Diagnostic Level of Evidence. Similarly, if the tumor type-specific clinical implication is prognostic or therapeutic, it must be associated with a Prognostic or Therapeutic Level of Evidence respectively.

I. Clinical Summary

The clinical implications of an alteration is summarized in 1-2 sentences. These sentences describe the therapeutic, diagnostic and/or prognostic implications for alterations with a level of evidence. This section is free text codes may be used for curating tumor type summary in order to include patient's variant and tumor type in the sentence, since they may be different from the curated data, e.g., V600E in patient will be matched to V600.

- A. [[variant]]: "gene" "mutation" mutant "tumor type" - e.g., BRAF V600E mutant melanoma
- B. [[tumor type]]: "tumor type" - e.g., melanoma
- C. [[gene]] - Adds the "gene" name - e.g., BRAF
- D. [[mutation]] - Adds the "mutation" name - e.g., V600E
- E. [[mutation]] [[mutant]] - Adds: "mutation" name and "mutant" - e.g., V600E mutant

II. Diagnostic Implications

The purpose of this section is to curate alterations which have tumor type specific diagnostic implications.

A. Level of Evidence

This section includes a drop-down list that allows a curator to choose the appropriate diagnostic Level of Evidence associated with the alteration in a specific tumor type. The drop-down list includes the following choices (Fig. 9):

Dx1	FDA and/or professional guideline-recognized biomarker required for diagnosis in this indication
Dx2	FDA and/or professional guideline-recognized biomarker that supports diagnosis in this indication
Dx3	Biomarker that may assist disease diagnosis in this indication based on clinical evidence

Fig. 9: OncoKB Diagnostic Levels of Evidence Schema.

1. *Dx1* defined as “FDA and/or professional guideline-recognized biomarker required for diagnosis in this indication.”
2. *Dx2* defined as “FDA and/or professional guideline-recognized biomarker that supports diagnosis in this indication.”
3. *Dx3* defined as “Biomarker that may assist disease diagnosis in this indication based on clinical evidence.”

B. Description of Evidence

This section is free text and contains 4-6 sentences and describes an overview and results from clinical studies describing the prevalence of the gene-alteration in the specified disease including the cohort size, the genetic criteria for patient selection, and the total number and percent of patients with the specified gene-alteration.

C. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT, but may not necessarily be included in the final output. The information in this section will only be accessible from the OncoKB curation interface and therefore will not be displayed on other platforms (ie. OncoKB website, cBioPortal, MSK-IMPACT Reports). This section is free text.

III. Prognostic Implications

The purpose of this section is to curate alterations which have tumor type specific prognostic implications.

A. Level of Evidence

This section includes a drop-down list that allows the user to choose the appropriate prognostic Level of Evidence associated with the alteration in a specific tumor type. The drop-down list includes the following choices (Fig. 10):

Px1	FDA and/or professional guideline-recognized biomarker prognostic in this indication based on well-powered study(s)
Px2	FDA and/or professional guideline-recognized biomarker prognostic in this indication based on a single or multiple small studies
Px3	Biomarker is prognostic in this indication based on clinical evidence in well-powered studies

Fig.10: OncoKB Prognostic Levels of Evidence Schema.

1. *Px1* defined as “FDA and/or professional guideline-recognized biomarker prognostic in this indication based on well-powered studies.”
2. *Px2* defined as “FDA and/or professional guideline-recognized biomarker prognostic in this indication based on a single or multiple small studies.”

3. *Px3* defined as “*Biomarker is prognostic in this indication based on clinical evidence in well-powered studies.*”

B. Description of Evidence

An overview and results from clinical studies describing the prognostic implications of the gene-alteration in the specified disease including the cohort size, the genetic criteria for patient selection, the percent of patients with and without the specified gene-alteration, and the endpoints used to predict clinical benefit or harm (e.g., overall survival (OS), progression-free survival (PFS), disease-free survival (DFS) and associated p-values). This section is free text and contains 4-6 sentences.

C. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT but may not necessarily be included in the final output. The information in this section will only be accessible from the OncoKB curation interface and therefore will not be displayed on other platforms (ie. OncoKB website, cBioPortal, MSK-IMPACT Reports). This section is free text.

IV. Standard Implications for Sensitivity to Therapy

The standard therapeutic implications for sensitivity of alterations that are FDA- or NCCN- recognized as biomarkers predictive of response to FDA-approved therapies in specific tumor types are curated in this section (Fig. 7, above). Here, a curator may enter the name of the standard sensitivity therapy as free text in the “Therapy.” box. Once a therapy is entered, the following sections become available for curation:

A. Highest Level of Evidence

This section includes a drop-down list that allows the user to choose the appropriate standard Level of Evidence. The drop-down list includes the following choices (refer to Fig. 7):

1. *Level 1* defined as “*FDA-recognized biomarker predictive of response to an FDA-approved drug in this indication.*”
2. *Level 2A* defined as “*Standard care (NCCN or other expert panels) biomarker predictive of response to an FDA-approved drug in this indication.*”

B. Level of Evidence for Alternate Oncogenic Alleles

If a curated allele (e.g., BRAF V600E) is considered Level 1 or 2A and predictive of sensitivity to a standard care therapy, then all alternate oncogenic alleles at the same amino acid (e.g., BRAF V600G) will be associated with the same level of evidence and will be considered predictive of sensitivity to the same standard care therapies.

C. Level of Evidence in Other Tumor Types

This section includes a drop-down list that allows the user to decide if the standard therapy evidence should be propagated to other tumor types. The drop-down list includes the following choices (refer to Fig. 7):

1. *Level 2B* defined as “*Standard care biomarker predictive of response to an FDA-approved drug in another indication, but not standard care in this indication.*” By definition, any Level 1 or Level 2A gene-alterations are propagated to Level 2B in other tumor types.
2. *Level 4*: defined as “*Compelling biological evidence supports the biomarker as being predictive of response to a drug.*” A Level 1 or 2A alteration found in another tumor type can be associated with Level 4 (and not 2B), if there is evidence in the specified tumor-type that is conflicting or inconclusive.
3. *No Level*: The curated standard therapeutic evidence will not be propagated to other tumor types. Therefore, if the gene-alteration combination is found in a tumor-type other than the one specified, it will not receive a Level of Evidence.

D. Description of Evidence

This section is 4 to 6 sentences, comprised of free text and describes the following:

1. The therapy and its targets.
2. Overview and results from clinical studies testing the drug in patient populations including the cohort size, the genetic criteria for patient selection, and the results of the study (e.g., response rates and statistical analysis).

3. Description and results from studies testing the therapy in in vitro and/or in vivo models, if relevant. For Level 1 and 2A therapies, the curated studies reflect those referenced by the FDA and/or NCCN Compendium.

E. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT but may not necessarily be included in the final output. The information in this section will only be accessible from the OncoKB curation interface and therefore will not be displayed on other platforms (i.e., OncoKB website, cBioPortal, MSK-IMPACT Reports). This section is free text.

V. Standard Implications for Resistance to Therapy

The standard therapeutic implications for resistance of alterations that are NCCN- recognized as biomarkers predictive of resistance to FDA-approved therapies in specific tumor types are curated in this section (refer to Fig. 7). Here, a curator can enter the name of the standard resistance therapy as free text in the “Therapy:” box. Once a therapy is entered, the following sections become available for curation:

A. Level R1

The highest and only standard level of resistance, Level R1. It is defined as “*Standard care biomarker predictive of resistance to an FDA-approved drug in this indication.*”

B. Description of Evidence

This section is free text and contains 4-6 sentences that describes the following:

1. The drug and its genetic targets.
2. Overview and results from clinical studies and/or case studies documenting resistance to the therapy.
3. The following information should be documented from clinical studies: cohort size, the genetic criteria for patient selection, and the clinical results of the study (e.g., response rates and statistical analysis).
4. Description and results from studies documenting resistance to the therapy in in vitro and/or in vivo models, if relevant. For Level R1 therapies, the curated studies reflect those referenced by the NCCN Compendium.

C. Level R1 does not use the Alternate Oncogenic Alleles and Other Tumor Types rules.

Note, alterations considered Level R1 are allele- and tumor type-specific. Therefore, if not specifically curated, positional variants of a Level R1 alteration are considered VUS. Similarly, Level R1 alterations when found in a tumor type outside of the Level R1 indication, are not associated with a Level of Evidence unless otherwise specified.

D. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT but may not necessarily be included in the final output. The information in this section will only be accessible from the OncoKB curation interface and therefore will not be displayed on other platforms (ie. OncoKB website, cBioPortal, MSK-IMPACT Reports). This section is free text.

VI. Investigational Implications for Sensitivity to Therapy

The investigational therapeutic implications for sensitivity of alterations for which there is published clinical (Level 3A) or preclinical (Level 4) data that supports that the alteration may serve as a biomarker that is predictive of response to an investigational therapy in specific tumor types are curated in this section (Fig. 7, above). Here, a curator may enter the name of the investigational sensitivity therapy as free text in the “Therapy:” box. Once a therapy is entered, the following sections become available for curation:

A. Highest Level of Evidence

This section includes a drop-down list that allows the user to choose the appropriate standard Level of Evidence. The drop-down list includes the following choices:

1. *Level 3A* defined as “*Compelling clinical evidence supports the biomarker as being predictive of response to a drug in this indication, but neither biomarker nor drug are standard care.*”
2. *Level 4* defined as “*Compelling biological evidence supports the biomarker as being predictive of response to a drug, but neither biomarker nor drug are standard care.*”

B. Level of Evidence for Alternate Oncogenic Alleles

If a curated allele (e.g., PIK3CA E542K) is considered Level 3A or 4 and predictive of sensitivity to an investigational therapy, then all alternate oncogenic alleles (e.g., PIK3CA E542A) will be associated with the same level of evidence and will be considered predictive of sensitivity to the same investigational therapies.

C. Level of Evidence in Other Tumor Types

This section includes a drop-down list that allows the user to decide if the investigational therapy evidence should be propagated to other tumor types. The drop-down list includes the following choices:

1. Level 3B: If the alteration is found in another tumor type (i.e., *not* in the tumor type for which compelling clinical evidence documents the alteration being predictive of response to a drug), the level of evidence which defines the alteration as being predictive of response to a specific therapy is by default Level 3B.
2. Level 4: If the alteration is found in another tumor type (i.e., *not* in the tumor type for which compelling clinical evidence documents the alteration being predictive of response to a drug), the level of evidence which defines the alteration as being predictive of response to a specific therapy may be defined as Level 4 if there is conflicting or inconclusive evidence in the specified tumor-type.¹
3. No Level: If the alteration is found in another tumor type (i.e., *not* in the tumor type for which compelling clinical evidence documents the alteration being predictive of response to a drug), the level of evidence which defines the alteration as being predictive of response to a specific therapy is not propagated to other tumor types. Therefore, if the gene-alteration combination is found in a tumor-type other than the one specified, it will not receive a Level of Evidence.

D. Description of Evidence

This section is free text and contains 4-6 sentences that describes the following:

1. The drug and its targets.
2. Overview and results from clinical studies and/or case studies testing the drug in patient populations including the cohort size, the genetic criteria for patient selection, and the results of the study (e.g., response rates and statistical analysis) (Level 3A only).
3. Description and results from studies testing the therapy in in vitro and/or in vivo models.

E. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT, but may not necessarily be included in the final output. The information in this section will only be accessible from the OncoKB curation interface and therefore will not be displayed on other platforms (ie. OncoKB website, cBioPortal, MSK-IMPACT Reports). This section is free text.

F. Alterations that are Level 1, 2A or 3A in another tumor type are considered Level 2B or 3B in the indicated tumor type

Level 2B and 3B evidences are not curated directly into OncoKB, but are automatically propagated. These levels of evidence are defined as the following:

1. Level 2B “Standard care (NCCN or other expert panels) biomarker predictive of response to an FDA-approved drug in another indication, but not standard of care for this indication.”
2. Level 3B “Compelling clinical evidence supports the biomarker as being predictive of response to a drug in another indication, but neither biomarker nor drug are standard care.”

VII. Investigational Implications for Resistance to Therapy

The investigational therapeutic implications for resistance of alterations are those for which there is compelling clinical data that supports that the alteration may serve as a biomarker predictive of resistance to FDA-approved or investigational therapies in specific tumor types are curated in this section (refer to Fig. 7). Here, a curator can enter the name of the investigational resistance therapy as free text in the “Therapy:” box. Once a therapy is entered, the following sections become available for curation:

A. Level R2

The highest and only investigational level of resistance, Level R2. It is defined as “*Compelling clinical evidence supports the biomarker as being predictive of resistance to a drug.*”

B. Description of Evidence

This section is free text and contains 4-6 sentences that describes the following:

1. The drug and its targets.
2. Overview and results from clinical studies and/or case studies (if applicable) documenting resistance to the therapy.
3. The following information should be documented from clinical studies: cohort size, the genetic criteria for patient selection, and the clinical results of the study (e.g., response rates and statistical analysis).
4. Description and results from studies documenting resistance to the therapy in in vitro and/or in vivo models.

C. Level R2 does not use the Alternate Alleles and Other Tumor Types rules.

Note, alterations considered Level R2 are allele- and tumor type-specific. Therefore, if not specifically curated, positional variants of a Level R2 alteration are considered VUS. Similarly, Level R2 alterations when found in a tumor type outside of the Level R2 indication, are not associated with a Level of Evidence unless otherwise specified.

D. Additional Information (Optional)

Provides the curator with space to add additional information and/or references to information that may be of relevance to the SCMT but may not necessarily be included in the final output. The information in this section will only be accessible from the OncoKB curation interface and therefore will not be displayed on other platforms (ie. OncoKB website, cBioPortal, MSK-IMPACT Reports). This section is free text.

VIII. Variants of Unknown Significance (VUS)

VUS are added to a unique section within the OncoKB Gene Curation Page called “Variants of Unknown Significance (Investigated and data not found)” (Fig. 11). Once a VUS is entered, it is linked to a timestamp displaying the date the VUS was last edited. If a VUS on the Gene Curation Page is investigated at a future date and still no data is found, the “Refresh” button can be clicked to update the timestamp associated with the VUS in question.

VUS are alterations for which limited, or no information is publicly available and falls into one of three possible classes:

1. No data exists.
2. The variant has been identified within a tumor, but not functionally tested (in this case, the comment bubble for each variant lists the appropriate publications for SCMT reference).
3. The variant has been described ONLY in the context of conferring sensitivity or resistance to a drug (in this case, the comment bubble for each variant lists the appropriate publications for SCMT reference).

A VUS on the Gene Curation Page entered:

1. Grey = Curated < 3 months prior to the current date.
2. Yellow = Curated 3 > 6 months prior to the current date.
3. Red = Curated > 6 months prior to the current date.

Variants of Unknown Significance (Investigated and data not found)
L707W
K642N
A245
D41N
G63K
C231H
C571S
L730F
G857R
A120P
V786L
A1118S
N103I
X297_splice
T302Qfs*39
A439P
S308L
G221V
D1014V
G63R
R108Q
P252C
I598S
P741L
R222H
H988P
A291T
C1089T
R1052G
ZCC108-EGFR
MD277Y
Q911A
E967T
D1058Y
E967K
D1014V
G698A
R146W
P252P
C620W
C779Y
H304Y
L858K
G8A
E1005V
TNS3-EGFR
L747Qfs*16
A955T
SV766L
D849L
D1009Y
M945I
E282D
A137
EGFR-TNS3
G598E
E160K
A289N
C103Y
F252H
T363S
K754I
G776F
G1185S
SV766L
X210_splice
M800I
S752Efs*9
D256Y
N1054Y

Fig. 11: Variants of Unknown Significance section in OncoKB Curation Platform.

IX. Reanalysis and Reevaluation

All reanalysis and reevaluation of OncoKB assertions and data is executed by the SCMT under the guidance of the Lead Scientist and occurs every 4 to 6 weeks. Each OncoKB data release is logged in the OncoKB github data repository and accessible to all users through the OncoKB website.

Re-analysis and re-evaluation of potential data discrepancies are identified using the following six database queries:

1. Variants with conflicting/inconclusive assertions of oncogenic/biological effect
2. Variants without oncogenic or mutation effect assertions
3. Variants with oncogenic/mutation effect assertions but without curated evidence (i.e., absence of PMIDs)
4. Variants with oncogenic and mutation effect assertions but without curated Evidence (i.e., absence of PMIDs)
5. Variants with multiple oncogenic/mutation effect assertions
6. Comparison of all variants associated with a Level of Evidence between previous and about-to-be released website versions

Any discrepancies identified through these queries are corrected in real-time prior to OncoKB data release.

In addition, to ensure that all variant assertions are accurate and the evidence supporting an assertion is up-to-date, comprehensive re-evaluation and re-analysis of genes and their associated variants occurs in review cycles specified in Table 1. The SCMT may execute the review themselves or assign specific gene(s) as needed for re-evaluation to curators.

Table 1. OncoKB data as of 2/1/2019.

	Genes (%)	Variants (%)	Review Cycle
# with a Level of Evidence	81 (14)	161 (4)	Every 4 to 6 weeks
# with Variant Assertions	311 (54)	4220 (96)	~50 genes every 4 months (all genes evaluated in ~2 years)
# without Variant Assertions	187 (32)	N/A	All gene summary and backgrounds reviewed every 2 years
Total	579 (100)	4381 (100)	-

In the OncoKB curation platform, all variant assertions in the OncoKB website are associated with a Description of Evidence that have been curated by OncoKB curators and/or SCMT with links to the supporting evidence sources (e.g., PMIDs or Abstracts). Per specific review cycle, these descriptions of evidence for the set of genes being re-evaluated can be downloaded for review (Fig. 12). Should the SCMT find that the Description of Evidence or sources supporting a variant assertion is inaccurate, the SCMT, in consultation with the Lead Scientist, makes the appropriate changes.

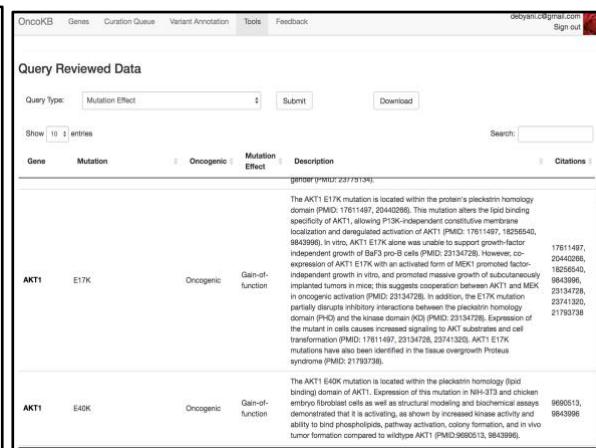
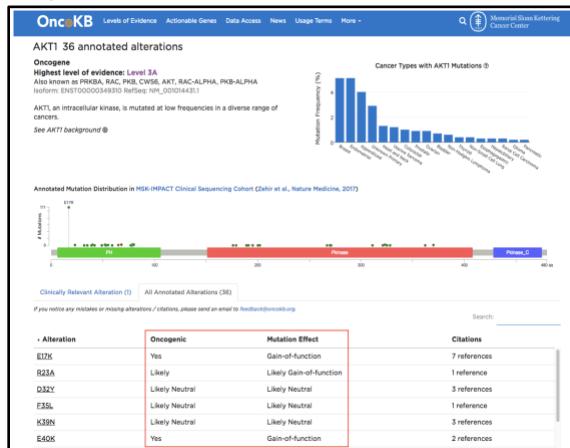


Fig. 12: OncoKB Data Reanalysis and Reevaluation. Variant assertions in the OncoKB website (left panel, boxed in red) have curated descriptions of the evidence supporting the assertion in the curation platform. These are used by the Lead Scientist and SCMT to reevaluate assertions.

Chapter 6: Annotation of Variants in Patient Tumor Samples

With the curated content as the foundation, OncoKB has implemented tools for annotating variants detected in sequenced patient tumors (including a web application programming interface and an annotator tool, both described in the chapter 7). OncoKB annotates variants with assertions of its oncogenic and biological effects, and with its tumor type-specific clinical implications using automation based on specific rules described below. These rules are in place to simplify the curation process when possible and provide annotation to variants for which there may not be specific functional data, but whose oncogenic and mutation effect can be inferred from other functionally validated variants or through its statistical recurrence in cancer.

I. Variant Annotation Process

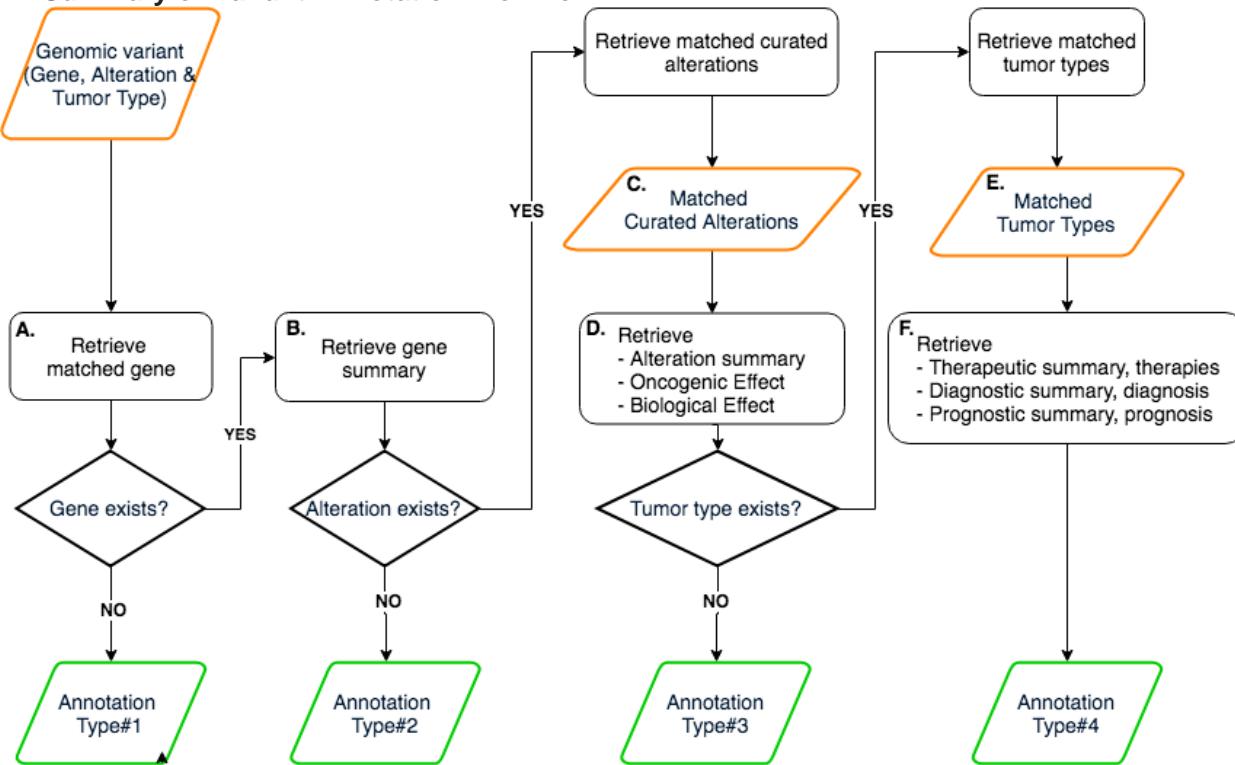
In cBioPortal, OncoKB data is used to annotate alterations found in individual patient tumor samples. These annotations comprise of three brief statements:

1. **Gene summary:** One to two sentences detailing the functional role of the gene in a cell and in which tumor types it is frequently altered. e.g., *BRAF, an intracellular kinase, is frequently mutated in melanoma, thyroid and lung cancers among others.*
2. **Oncogenic summary:** An evidence-based assertion that defines the oncogenic effect of the alteration. Possible assertions include Oncogenic, Likely Oncogenic, Neutral, Likely Neutral, or Inconclusive. (See Chapter 4, Section II.A) e.g., *The BRAF V600E mutation is known to be oncogenic.*
3. **Clinical Summary:** The clinical summary is a one or two sentences summarizing the therapeutic implications of the queried alteration with a therapeutic level of evidence in a specific tumor type. e.g *The RAF-inhibitors encorafenib, dabrafenib and vemurafenib alone or in combination with the MEK-inhibitors binimetinib, trametinib and cobimetinib, respectively, are FDA-approved for the treatment of patients with BRAF V600E/K mutant melanoma.*

When an alteration in a patient tumor sample is queried, the clinical implications associated with all matched curated alterations and all matched curated tumor types are matched to the queried alteration and tumor type. However, only one oncogenic effect, mutation effect, and description of evidence can be associated with the queried alteration and tumor type. Therefore, to assign the specific oncogenic effect, biological effect, and description of evidence to a queried alteration, the process described in Fig. 13 is used:

- A. Match gene.
Curated genes can be queried by HUGO symbols or Entrez Gene IDs.
- B. Retrieve gene summary.
The curated gene summary will be retrieved to annotate the queried variant.
- C. Match curated alterations.
The process to match curated alterations is described in the Nomenclature and Rules section.
- D. Retrieve mutation summary, oncogenic and biological effects for the alteration.
This is based on matched curated alterations (see Nomenclature and Rules section).
- E. Match curated tumor types.
Refer to Section II. Nomenclature and Rules.
- F. Retrieve tumor type summary and clinical implications.
A tumor type summary will be generated for the queried variant. All clinical implications related to matched curated alterations and tumor types will be pulled. These implications will then be sorted by OncoKB level priorities (defined above). The resistance level implication has a higher priority than sensitivity levels if they are associated with the same therapy.

A. Summary of Variant Annotation Workflow



B. Sample Annotation Workflows

Genomic Variant #1 = PHF6 C242Y Acute Myeloid Leukemia

Annotation Type#1 = There is currently no information about this gene in OncoKB.

Genomic Variant #2 = BCL2 A131D Glioblastoma Multiforme

Annotation Type#2 = BCL2, an anti-apoptotic protein, is frequently altered in non-Hodgkin lymphomas. As of 01/03/2019, there was no available functional data about the BCL2 A131D mutation. However, it has been identified as a statistically significant hotspot and is predicted to be oncogenic (<http://cancerhotspots.mskcc.org>). There are no FDA-approved or NCCN-compendium listed treatments specifically for patients with BCL2 A131D mutant glioblastoma multiforme.

Genomic Variant #3 = EWSR1 ZNRF3-EWSR1 fusion Lung Adenocarcinoma

Annotation Type#3 = EWSR1, a multi-functional protein that binds DNA and RNA, is altered by chromosomal rearrangement in various cancer types, most frequently in Ewing Sarcoma. The ZNRF3-EWSR1 mutation has not specifically been reviewed by the OncoKB team, and its oncogenic function is considered unknown. There are no FDA-approved or NCCN-compendium listed treatments specifically for patients with ZNRF3-EWSR1 mutant lung adenocarcinoma.

Genomic Variant #4 = BRAF V600E Melanoma

Annotation Type#4 = BRAF, an intracellular kinase, is frequently mutated in melanoma, thyroid and lung cancers among others. The BRAF V600E mutation is known to be oncogenic. The RAF-inhibitors encorafenib, dabrafenib and vemurafenib alone or in combination with the MEK-inhibitors binimetinib, trametinib and cobimetinib, respectively, are FDA-approved for the treatment of patients with BRAF V600E/K mutant melanoma.

Fig. 13: Variant and Sample Annotation Workflow. A, Summary: To annotate variants found in patient tumor samples with its oncogenic and biological effects, and with tumor type-specific clinical implications OncoKB uses semi-automation summarized by this workflow. **A: Match gene;** Curated genes can be queried by HUGO symbols or Entrez Gene IDs. **B: Retrieve gene summary;** The curated gene summary will be retrieved to annotate the queried variant. **C: Match curated alterations;** The process to match curated alterations is described in the Nomenclature and Rules section. **D: Retrieve mutation summary, oncogenic and biological effects for the alteration;** This is based on matched curated alterations (see Nomenclature and Rules section). **E: Match curated tumor types;** Refer to Section II. Nomenclature and Rules. **F: Retrieve tumor type summary and clinical implications.** A tumor type summary will be generated for the queried variant. All clinical implications related to matched curated alterations and tumor types will be pulled. These implications will then be sorted by OncoKB level priorities (defined above). The resistance level implication has a higher priority than sensitivity levels if they are associated with the same therapy. Orange rhombus = Input; Green rhombus = Output; Rectangle = Process; Diamond = Decision. **B, Examples:** Shown are sample annotation for the four different annotation types shown in part A of the figure.

II. Nomenclature and Rules Related to Annotation

A. OncoKB Cancer Gene List

OncoKB maintains a list of genes we consider as cancer genes based on their inclusion in various different sequencing panels, the Sanger Cancer Gene Census, or Vogelstein et al. (2013).

B. Curated Genes

Not every gene in the OncoKB Cancer Gene list has been curated by the team. We release new genes incrementally and refer to these genes as Curated Genes.

C. Matched Genes

OncoKB accepts gene HUGO symbols, Entrez gene IDs and gene aliases in the query to identify Curated Genes.

D. Matched Curated Alterations

When an alteration is queried in the OncoKB database, it may be associated with several alterations curated in the Gene Page and their associated annotations which includes their oncogenic and biological effect, clinical implications and tumor type summary. The various curated alterations in OncoKB that match the queried alteration are referred to as Matched Curated Alterations.

1. Overall Matching Logic

Each queried alteration may be associated with one oncogenic effect and one biological effect. Therefore, the biological effect can be automatically associated with the queried alteration. The order of retrieving the information is the following:

- a. Exact Match (single mutation header, e.g., V600E)
- b. Exact Match (mutation in a string, e.g., V600E, V600K)
- c. Positional Variant Match (e.g., V600)
- d. Alternate Allele (e.g., V600K)
- e. Range Mutations (e.g., V600_K601delinsEQ)
- f. Fusions
- g. Deletion
- h. Truncating Mutations
- i. Oncogenic Mutations
- j. Gain of Function Mutations
- k. Loss of Function Mutations
- l. Special Rules for Alterations

2. Special Rules for Alterations

- a. **Missense Mutations:** If a specific missense mutation (e.g., BRAF V600E) is queried, it will be mapped to all curated mutations that reference the specific mutation position. This may include:
 - i. the exact mutation match (V600E)
 - ii. the exact mutation match in a list of mutations (V600E, V600K)
 - iii. the positional variant match (V600)
 - iv. alternate oncogenic alleles of a curated oncogenic mutation (V600K)
 - v. a missense mutation range that includes the queried mutation (V600_K601mut)
- b. **In-frame Mutations:** OncoKB can curate in-frame mutations within an amino acid range. In-frame mutations will be mapped when the queried alteration position intersects within a curated range.
- c. **Oncogenic Mutations:** Any queried alteration that is annotated as “Oncogenic” or “Likely Oncogenic” in the OncoKB database, will be mapped to “Oncogenic Mutations”.
- d. **Fusions:** If a specific fusion is queried, it will be mapped to: 1) the specific fusion and 2) “Fusions” if curated.
- e. **Truncating Mutations:** If a truncating alteration is queried, it will be mapped to: 1) the specific truncating alteration and 2) “Truncating Mutations” if curated.
- f. **Duplications:** For small tandem duplications (dups), the queried alteration must be an exact match to get mapped.

- g. ***Deletion:*** If a deletion event is queried, it will be mapped to: 1) “Deletion” and 2) “Truncating Mutations” if curated. If a deletion event is queried, and “Truncating Mutations” but not “Deletion” is curated.

E. Alternate Oncogenic Alleles

1. In OncoKB, when an alteration is curated as oncogenic or likely oncogenic (based on the rules described in Chapter 2, Section IV), alterations with variant alleles at the same position are considered likely oncogenic.
2. “*Curated oncogenic alleles*” are defined as alterations that are curated as oncogenic or likely oncogenic (based on the rules described in Chapter 2, Section IV). Examples of curated oncogenic alleles are BRAF V600E/K/R or BRAF L595Q/R.
3. “*Alternate oncogenic alleles*” are defined with respect to a curated allele and refer to missense variants at the same allelic protein position that is curated as oncogenic or likely oncogenic. The “alternate oncogenic allele” rule applies to VUS and variants that have not been investigated or curated by the SCMT. Examples of alternate oncogenic alleles are BRAF V600L or BRAF L595M.
4. Importantly, the term “alternate oncogenic alleles” does NOT apply for alterations that are considered likely neutral or inconclusive.
5. If there is only one curated allele in OncoKB, the oncogenic and mutation effect of the alternate allele will be as follows:
 - a. *Oncogenic Effect:* For a curated Oncogenic or Likely Oncogenic allele, its alternate oncogenic alleles are Likely Oncogenic.
 - b. *Biological Effect:*
 - i. If a curated oncogenic allele is Gain of Function or Likely Gain of Function, then its alternate oncogenic alleles are Likely Gain of Function
 - ii. If curated oncogenic allele is Loss of Function or Likely Loss of Function, then its alternate oncogenic alleles are Likely Loss of Function
 - iii. If a curated oncogenic allele is Switch of Function or Likely Switch of Function, then its alternate oncogenic alleles are Likely Switch of Function
6. If there are two or more curated oncogenic alleles on the same position with different oncogenic effects, the curated oncogenic allele with the highest-ranked oncogenic effect (Oncogenic > Likely Oncogenic) will be used to assign the biological effect to the alternate oncogenic allele.

F. Hotspots

Mutational hotspots are defined as mutant residues arising more frequently than expected in the absence of selection based on the analysis by Chang et al., 2016 and Chang et al., 2018.

G. Matched Curated Tumor Types

Clinical implications are matched based on the patient’s tumor type. Queried tumor type will be associated with curated tumor types for the summary and clinical implication. As long as the curated tumor type is the same as or the parent node (base on OncoTree definition) of the query tumor type, it will be matched as a matched curated tumor type. We also include few general tumor types (All Tumors, All Solid Tumors, All Liquid Tumors) and they will be mapped accordingly.

H. OncoKB Therapeutic Implication Levels of Evidence Priorities

Multiple therapeutic implications may be matched to a variant in a patient. When ranking them, we use the following order to keep the highest level of the implications.

Level R1 > Level 1 > Level 2A > Level 2B > Level R2 > Level 3A > Level 3B > Level 4 > Level R3

We also recommend Level 3A treatment when the highest level is 2B and Level 3A implications also exist.

III. Annotation Summaries

A. Gene summary

Gene summary will be retrieved as curated in the system, e.g., “BRAF, an intracellular kinase, is frequently mutated in melanoma, thyroid and lung cancers among others.”

B. Variant Biological Summary

The biological summary is one sentence that describes the oncogenic effect of the queried alteration. This sentence is programmatically generated based on the oncogenicity of the genetic alteration (refer to Table 2). The mutation summary is included in the variant-annotation endpoints of the OncoKB API.

Table 2. Example mutation summaries.

Mutation	If the alteration selected in cBioPortal is...	The sentence in the OncoKB card will be...
BRAF V600E	Oncogenic	The BRAF V600E mutation is known to be oncogenic.
BRAF T241P	Likely Oncogenic	The BRAF T241P mutation is likely oncogenic.
BRAF V600L	Alternate Allele	The BRAF V600L mutation has not been functionally or clinically validated. However, BRAF V600D/E/K/R are likely oncogenic, and therefore V600L is considered likely oncogenic.
BRAF R509Q	Likely Neutral	The BRAF R509Q mutation is likely neutral.
BRAF Q201H	Inconclusive	There is conflicting and/or weak data describing the oncogenic function of the BRAF V600X mutation
BRAF A762V	Variant of Unknown Significance (VUS) assessed by SCMT	As of 10/17/2018, there was no available functional data about the BRAF A762V alteration.
BRAF P318S	VUS not assessed by SCMT	The BRAF P318S mutation has not specifically been reviewed by the OncoKB team, and its oncogenic function is considered unknown.
ARID1A G2087V	Hotspot (VUS not assessed by SCMT)	The ARID1A G2087V mutation has been identified as a statistically significant hotspot and is predicted to be oncogenic.
[Gene] [Mutation]	Hotspot (VUS assessed by SCMT)	As of [date], there was no available functional data about the [gene] [mutation] mutation. However, it has been identified as a statistically significant hotspot and is predicted to be oncogenic (http://cancerhotspots.mskcc.org).
DAXX Duplication	Structural variant within a gene that has “Truncating Mutations” curated as likely oncogenic	This DAXX duplication may be a truncating alteration may be a truncating alteration and is likely oncogenic.
BRAF Q201*	Truncating mutation in an oncogene	BRAF is considered an oncogene and truncating mutations in oncogenes are typically nonfunctional.

C. Clinical Summary

The clinical summary is a one or two sentences summarizing the therapeutic implications of the queried alteration with a therapeutic level of evidence in a specific tumor type. For example, “The RAF-inhibitor dabrafenib in combination with the MEK1/2-inhibitor trametinib is FDA-approved for the treatment of patients with BRAF V600E mutant anaplastic thyroid cancer.”

When a specific alteration in a patient tumor sample is queried for annotation, multiple curated alterations may be matched, and each matched curated alteration may have its own clinical summary. However, only one clinical summary will be associated with each specific alteration in a patient of a specific tumor type.

Therefore, in order to assign the clinical summary, the matched curated alterations are prioritized based on the order below (using BRAF V600E in a Colorectal Cancer (CRC) patient as an example):

1. Clinical summary under the exact match alteration (V600E) for the tumor type in question (CRC) (so in the example of V600E in CRC, we will stop here because we have curated the alteration and tumor type specific clinical summary)
2. Clinical summary under curated alternate alleles (e.g., V600K) for the tumor type in question (CRC)
3. Clinical summary under the relevant positional variant (V600) for the tumor type in question (CRC)
4. Clinical summary under the exact match alteration (V600E) for “Other Tumor Type”
5. Clinical summary under curated alternate alleles (V600K) for “Other Tumor Type”
6. Clinical summary under the relevant positional variant (V600) for “Other Tumor Type”
7. Clinical summary under the highest priority relevant alteration (see above for prioritization of matched curated alterations) for the tumor type in question (CRC)
8. Search under the highest priority relevant alteration (see above for prioritization of MCAs) for the other tumor type and use that summary (if present)
9. Continue steps 7-8 until all matched curated alterations have been evaluated for clinical summaries
10. If the queried alteration is associated with an “Oncogenic” or “Likely Oncogenic” mutation effect, search under “Oncogenic Mutations” for the tumor type in question (CRC)
11. If the queried alteration is associated with an “Oncogenic” or “Likely Oncogenic” mutation effect, search under “Oncogenic Mutations” for “Other Tumor Types”

D. Resistance Mutations

For alterations with an associated Level R1 or R2, the specified therapy (i.e., the therapy to which the alteration is considered a biomarker of resistance) will ONLY be associated with resistance (and NOT sensitivity).

Chapter 7: OncoKB Data Access

There are three ways that the public may access OncoKB data:

1. Through the OncoKB API
2. Through the publicly available website www.oncokb.org
3. Through cBioPortal

I. The OncoKB API

The OncoKB data can be accessed through REST API (<https://oncokb.org/api/v1/swagger-ui.html>). The API is defined and organized using swagger annotation. You could also annotate your MAF file by using oncokb-annotator (<https://github.com/oncokb/oncokb-annotator>) which is fully supported by using OncoKB REST APIs.

II. The OncoKB Website: www.oncokb.org

The OncoKB.org website (www.oncokb.org) was first released to the public at the American Society of Clinical Oncology (ASCO) Annual Meeting in 2016. This website provides to the clinical and scientific community worldwide the current and detailed annotation of the oncogenic effects and therapeutic implications of alterations observed in cancer.

As of 01/2019, the website has information about 4474 variants annotated in 595 genes across 38 tumor types, with therapeutic information for 79 drugs.

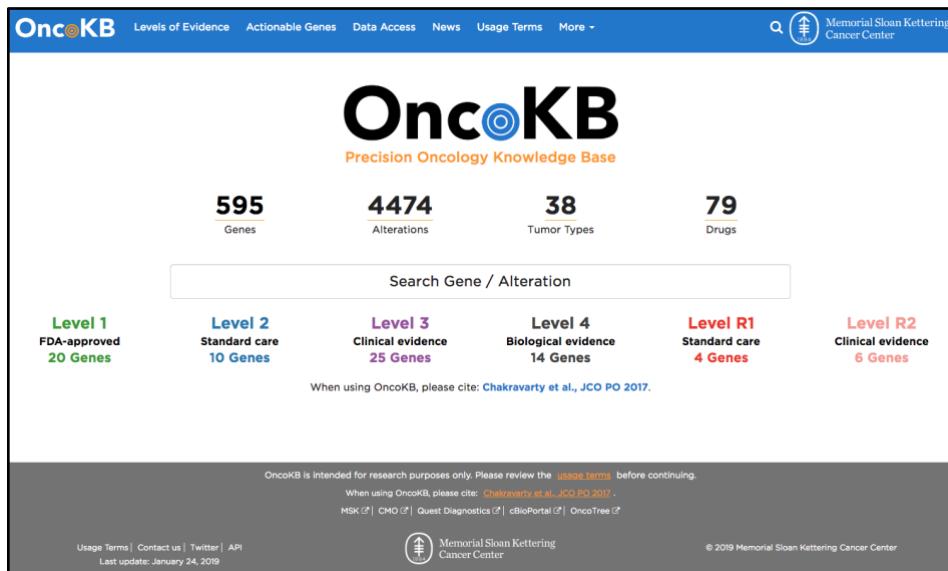


Fig. 14: OncoKB.org Homepage.

The homepage of [oncokb.org](http://www.oncokb.org) (Fig. 14) displays the following sections and functionalities:

A. Data summary

The website shows the current number of genes (clickable), alterations, actionable tumor types and drugs curated in OncoKB. The “genes” number links to the OncoKB Cancer Gene List page. Below the search bar, the number of genes with alterations associated with a level of evidence are summarized. The number of genes below each Level of Evidence links to the Actionable Genes page.

B. Search bar

Queries can be entered in the search box to lookup genes, aliases, EntrezID or gene-variant combinations in OncoKB. Upon entering a query, a drop-down menu will automatically appear listing possible gene and variant matches. Additionally, each suggested variant in the drop-down menu will be associated with an oncogenicity and (if relevant) the highest associated level of evidence. Clicking on a variant in the drop-down menu links to the variant page. Currently only one gene and/or one variant can be queried at a time.

C. Levels of Evidence

The Levels of Evidence page (Fig. 15) shows the hierarchy and definitions of the OncoKB Levels of evidence, as described in Chapter 3. This schematic can be downloaded in PDF or PPT format.

D. Actionable Genes

The Actionable Genes page (Fig. 16) lists all the gene-alteration-tumor type combinations that are associated with a level of evidence (Sensitivity Levels 1-4 and Resistance Levels R1-R2). The table is divided into four columns: Gene, Alterations, Disease, and Drugs. Clicking on the entry under “Gene” will bring the user to the respective gene page. Clicking on the entry under “Alteration” will bring the user to the variant page. Each level of evidence has its own search bar in the upper right-corner of the table. However, searching for information across levels of evidence requires using the search function of the browser.

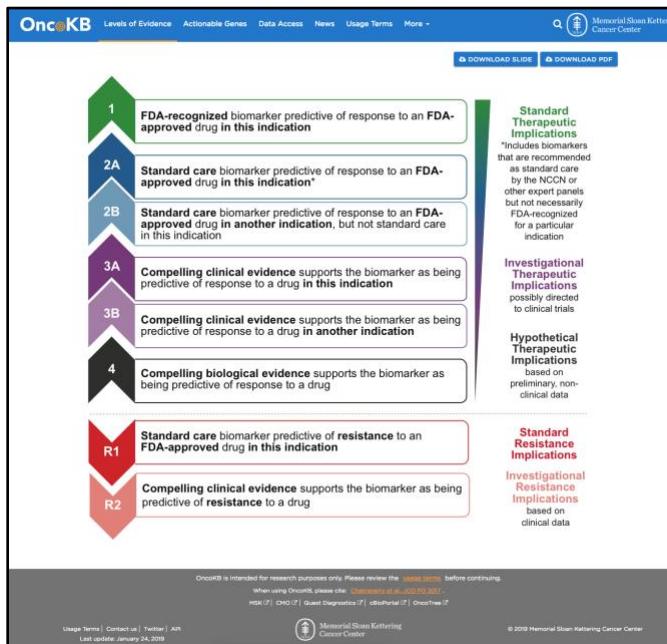


Fig. 15: Levels of Evidence page in oncokb.org.

The Actionable Genes page displays a table of gene-alteration-disease-drug associations. The table includes the following columns:

Gene	Alterations	Disease	Drugs
ABL1	BCR-ABL1 Fusion	B-Lymphoblastic Leukemia/Lymphoma with t(9;22)(q34.1;q12);BCR-ABL1	Imatinib, Dasatinib
ABL1	BCR-ABL1 Fusion	Chronic Myeloid Leukemia, BCR-ABL1+	Imatinib, Nilotinib, Dasatinib
ALK	Fusions	Non-Small Cell Lung Cancer	Crizotinib, Ceritinib, Alectinib
ALK	Oncogenic Mutations	Non-Small Cell Lung Cancer	Lorlatinib
BRAF	V600	Erdheim-Chester Disease	Vemurafenib
BRAF	V600E	Anaplastic Thyroid Cancer	Dabrafenib+Trametinib
BRAF	V600E	Melanoma	Dabrafenib, Vemurafenib
BRAF	V600E, V600K	Non-Small Cell Lung Cancer	Dabrafenib+Trametinib
BRAF	V600E, V600K	Melanoma	Dabrafenib+Trametinib, Cobimetinib+Vemurafenib, Binimetinib+Encorafenib
BRCA1	Oncogenic Mutations	Ovarian Cancer	Rucaparib, Niraparib
BRCA2	Oncogenic Mutations	Ovarian Cancer	Rucaparib, Niraparib

Showing 1 to 34 of 34 entries

Fig. 16: Actionable Genes page in oncokb.org.

E. Data Access

The OncoKB Data Access page (Fig. 17) provides all access to OncoKB data in three ways:

1. A programmatic web application interface (API), which is accessible at <https://oncokb.org/api/v1/swagger-ui.html>
2. The OncoKB Annotator, a tool that utilizes the OncoKB API and has embedded technical rules as described in Chapter 3
3. Via downloadable text files:

Three static downloadable text files are accessible in the Data Access page:

1. **All Curated Genes.** The text file contains the following information about every gene curated in OncoKB:
 - Hugo Symbol
 - Entrez Gene ID
 - Is Tumor Suppressor (YES/NO)
 - Is Oncogene (YES/NO)
 - Isoform
 - Refseq
 - Highest level of Evidence (sensitivity)
 - Highest level of Evidence (resistance)
2. **All Curated Alterations.** The text file contains the following information about every alteration curated in OncoKB:

- a. Gene
 - b. Alteration
 - c. Oncogenicity
 - d. Mutation Effect
 - e. PMIDs for Mutation Effect
 - f. Abstracts for Mutation Effect
3. Actionable Alterations. The text file contains the following information about every alteration associated with a level of evidence:
- a. Gene
 - b. Alteration
 - c. Cancer Type
 - d. Level of Evidence
 - e. Drugs(s)
 - f. PMIDs for drug
 - g. Abstracts

Fig. 17: Data Access page in oncokb.org.

G. News

The News page (Fig. 18) contains: 1) details of any new data and/or updates added at each OncoKB version release, 2) the date of each release, and 3) a link to sign up to receive low-volume OncoKB email updates. Website updates are released approximately monthly.

Specifically highlighted in the news are:

1. Changes to actionable alterations, levels of evidence or therapeutics
2. Addition of new genes
3. Changes to any functions on the website

Not highlighted are:

1. Changes to mutation effect or oncogenic effect of alterations
2. Changes to citations
3. Addition or subtraction of alterations
4. Changes to a gene's designation as tumor suppressor or oncogene

January 24, 2019

Updated Actionable Genes

Level	Update
Level 1	November 2, 2018: the FDA approved lorlatinib for patients with anaplastic lymphoma kinase (ALK)-positive metastatic non-small cell lung cancer (NSCLC) whose disease has progressed on crizotinib and at least one other ALK inhibitor or whose disease has progressed on alectinib or ceritinib for metastatic disease.
Addition of 16 new heme-associated genes: EC2L, RELN, TALI, MLLT10, TLX3, TLX1, TRA, TRB, TRD, TRG, EPOR, ABL2, MECOM, DEK, RBMS1, BCL11A	

December 14, 2018

Inclusion of NTRK1 and NTRK3 Level R2 alterations to the Actionable Genes page

Updated Actionable Genes

Level	Update
Level 1	November 26, 2018: the FDA approved larotrectinib for adult and pediatric patients with solid tumors that have an NTRK1, -2, or -3 gene fusion without a known acquired resistance mutation.
Level 2	BRCA1/2 - Oncogenic Mutations - Breast Cancer - Talazoparib (new association) RET - Fusions - Non-Small Cell Lung Cancer - BLU-667 (new association)
Level 3	BRAF - V600E - Colorectal Cancer - Encorafenib + Binimetinib + Cetuximab (new association) ERBB2 - Oncogenic Mutations - Non-Small Cell Lung Cancer - Ado-trastuzumab Emtansine (new association) RET - Oncogenic Mutations - Medullary Thyroid Cancer - BLU-667 (new association)
Level 4	KDM6A - Oncogenic Mutations - Bladder Cancer - EZH2 inhibitors (new association)

Addition of 25 new heme-associated genes:
KSR2, LCK, LTB, MGAM, MOB3B, MPEG1, NCOR2, PIGA, PLCG1, POTI, ROBO1, RUNX1IT1, SAMHD1, SETD1A, SGK1, SMC1A, SMC3, SMG1, SP140, STAT6, TBX1X, UBR5, VAV1, VAV2, XBPI

Fig. 18: News page in oncokb.org.

H. Usage Terms

This page contains OncoKB licensing and data usage terms and guidelines (Fig. 19). The usage guidelines must be read and understood before using the data in OncoKB. Any additional inquiries about OncoKB usage terms may be directed to contact@oncokb.org.

Usage Terms

OncoKB is a precision oncology knowledge base maintained by Memorial Sloan Kettering Cancer Center (MSK). MSK reserves the right to update http://oncokb.org at any time without notice. MSK makes no warranties or representations, express or implied, with respect to any of the content, including as to the present accuracy, completeness, timeliness, adequacy, or usefulness of any of the content, and by using this website you agree that MSK will not be liable for any loss or damages arising from your use of or reliance on information contained in this site or other sites that may be linked to from our site. This information is not intended as a substitute for medical professional help or advice. A physician should always be consulted for any health problem. If you have concerns about the content, please contact us. You may view the content and use it on http://oncokb.org solely for your own personal or research purposes. You may not use any part of the http://oncokb.org content ("Content") for any commercial purpose, including the distribution, licensing or sale of the Content to any other person or entity, whether alone or in combination with other materials, or the incorporation of the Content into any commercial product. You may copy, reproduce, or create derivative works of the Content only if:

- you are a researcher or a non-profit entity; **and**
- your use of the Content complies with all of the following requirements:
 - you are using the Content for personal or research purposes **only**; **and**
 - you are using the Content only to replicate OncoKB locally, whether in whole or in part; **or**
 - you are aggregating the Content with other data of similar nature for the purposes of advancing cancer research. You must credit the source and reference these usage terms.

OncoKB is intended for research purposes only. Please review the [Usage Terms](#) before continuing.
When using OncoKB, please cite Chakravarty et al. JCO PO 2017.
MSK () | CMO () | Guest Diagnostic () | cBioPortal () | OncoTree ()

Usage Terms | Contact us | Twitter | API
Last update: January 24, 2019

Fig. 19: Usage Terms in oncokb.org.

I. OncoKB Cancer Gene List

The OncoKB Cancer Gene List page (Fig. 20) contains the genes considered by OncoKB to be cancer genes and indicates with a check mark their inclusion in a specified resource, including:

1. MSK-IMPACT
2. MSK-IMPACT Heme
3. Foundation One
4. Foundation One Heme
5. Sanger Cancer Gene Census
6. Vogelstein et al., 2013.

Each gene is further classified as an Oncogene or Tumor Suppressor based on the criteria defined by OncoKB, which can be found in Chapter 3. The data on this page can be downloaded as a tab delimited file by clicking on the button in the upper right-hand corner of the page.

Fig. 20: OncoKB Cancer Gene List in oncokb.org.

J. About OncoKB

The About page (Fig. 21) provides information about the history of OncoKB, and provides a schema delineating its oversight and governance, inputs, workflow and outputs.

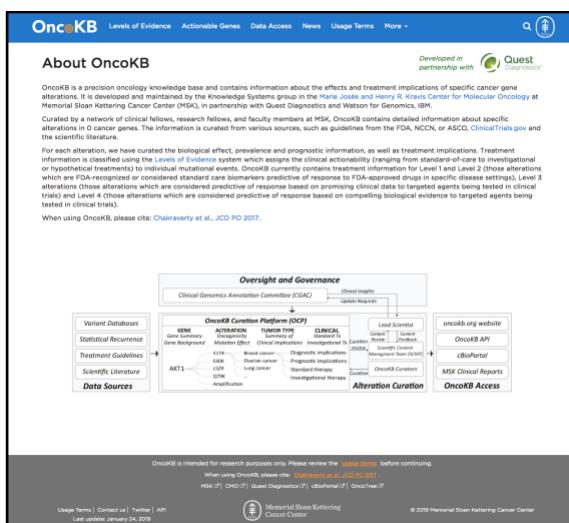


Fig. 21: The About Page in oncokb.org.

K. Team

The Team page (Fig. 22a) lists the names of the individuals involved in the creation, development and maintenance of OncoKB, including:

1. Design & Development Team (including members of the Lead Scientist, SCMT members and Leadership)
2. Current OncoKB Curators
3. Past Contributors to OncoKB
4. Clinical Genomics Annotation Committee

Note, financial conflicts of interest for all OncoKB personnel are disclosed publicly on the OncoKB website via linking to an online spreadsheet that lists all relevant relationships (Fig. 22b).


OncokB Levels of Evidence Actionable Genes Data Access News Usage Terms More
 
 Memorial Sloan Kettering Cancer Center

OncokB Team

OncokB is developed and maintained by the Knowledge Systems group in the Marie Javis and Henry R. Krieger Center for Molecular Oncology at Memorial Sloan Kettering Cancer Center. Disclosure of conflicts of interest of all OncokB contributors is available [here](#).

Design & Development

Debanya Chakravarty, PhD
 Jianping Ge, PhD
 Samir Desai, PhD
 Hongyan Zhang, MSc
 Ritika Kundra, MSc
 Nikiolaus Nissen, PhD
 Jing Su, PhD
 Ezzatullah Paraisa, MPA
 Julia Rudolph, MPA
 David Solit, MD
 Paul Sabatini, MD
 Nikolaius Schultz, PhD

Cure Curator

Krishna Gala, PhD
 Hannah Rosenman
 Phillip S. Thomas, PhD
 Lindsey M. Lafferty, PhD
 David Knorr, MD, PhD
 Linde Miles

Past Contributors

Tara Soumelli, MD
 Aphrothi Hennrich, PhD
 Anton Hennsen, MD
 Trupti Shrestha Bhattachari, PhD
 Neel Shah, PhD
 Rupa Lanahan, PhD
 Neel Shah, PhD
 Eneda Toksa, PhD
 Jiaqiao Wang, MSc

Clinical Genomics Annotation Committee

Carol Aghajanian, MD
 Maria Arribalzaga, MD
 Michael Berman, PhD
 Margaret Callahan, MD, PhD
 Timothy A. Chan, MD, PhD
 Sarat Chandranasamy, MD, PhD
 Peleg Chou, PhD
 Daniel Cohen, MD
 Lisa DeAngelis, MD
 Luis Alberto Diaz, Jr., MD
 Ahmet Dogan, MD, PhD
 Alexander Drilon, MD
 James A. Egner, MD
 Minral M. Gounder, MD
 James J. Hardin, MD
 Matthew D. Hellmann, MD
 Alan L. Holden, MD, PhD
 David Hwang, MD
 Gopu Iyer, MD
 Edgar A. Jimenez, MD
 Yelena Y. Janjigian, MD
 Philip Kantoff, MD
 David Klimstra, MD
 Andrew Kung, MD, PhD

Clinical Genomics Annotation Committee (continued)

Mark Ladanyi, MD
 C. Todd Andringra, MD, PhD
 Hugo K. Kahl, MD
 Kenneth Offit, MD
 Paul K. Palk, MD
 David G. Pfister, MD
 Dana E. Reiss, MD
 Gregory J. Rhee, PhD
 Mark E. Robinson, MD
 Neal Rosen, MD, PhD
 Leonard Saltz, MD
 Maurizio Santini, PhD
 Howard Scher, MD
 Sobhan Shah, PhD
 Alexander N. Shoushtari, MD
 Neeraj N. Shukla, MD
 Michael S. Sabel, MD
 William D. Stewart, MD
 Barry S. Taylor, PhD
 Tiffany A. Traina, MD
 Martin H. Voss, MD
 Jeld D. Welchek, MD, PhD
 Roni D. Yeager, MD
 Akira Younes, MD

OncokB is intended for research purposes only. Please review the [Usage Terms](#) before continuing.

When using OncokB, please cite: [OncokB \(2019\). doi:10.1101/2019.07.16.201035](#)
 HasC (2) | CMG (2) | Guest Interpreters (2) | Observers (2) | OncokB (2)

Usage Terms | Contact us | Twitter | API
 Last update: January 24, 2019


 Memorial Sloan Kettering Cancer Center

© 2019 Memorial Sloan Kettering Cancer Center

Onc0808 Contributor Disclosure of Potential Conflicts of Interest Q2017-2018 - Collected 2/24/2019										
Name	Employment	Leadership & Scientific Advisory Board	Stock or Ownership	Honoraria	Consulting or Advisory Role	Speakers' Bureau	Research Funding	Patients Royalties, other Intellectual Property (P)	Expert Testimony	Travel, Entertainment, Expenses
Carol Aphiczon, MD	Durata Pharma, Davis, Unincorporated, Calif.			AstraZeneca, Genentech, Merck, Novartis			AstraZeneca, Atrieps, Genentech, Otsuka			
Maria Aviles, MD	Invescience Inc., Biscayne			Invescience						Invescience Inc., Biscayne
Michael Berger, PhD				Roth						
Margaret Callahan, MD, PhD	Bristol Myers Squibb (Family Member)			AstraZeneca/MedImmune, Imclone, Merck, Novartis and Teva			Bristol Myers Squibb			
No disclosures										
Delphy Chabot, PhD										
Timothy A. Chan, MD, PhD	Bristol Myers Squibb Advisory Board Administrative Activities Audit Committee	Gillette Orchard Park, N.Y. Bristol Myers Squibb Co.	Bristol Myers Squibb Executive Committee Administrative Activities Audit Committee		Bristol Myers Squibb Administrative Activities Audit Committee			Use of TNM for prediction of intraoperative regional lymph node - presented to ASCO		Bristol Myers Squibb Administrative
Sonal Chaudhary, MD, PhD				Noxalis, Semera, Lilly, Genentech Therapeutics, Chugai			Genentech			Noxalis, Lilly, Chugai
Ping Chi, MD, PhD				Despax			Noxalis	U.S. Department of Defense, American Society of Clinical		U.S. Department of Defense, American Society of Clinical

Fig. 22: OncoKB Team List in oncokb.org. (a) All OncoKB personnel including past contributors are listed here. (b) The word “here” in the introduction statement “Disclosure of conflicts of interest of all OncoKB contributors is available here.” links to a spreadsheet that lists the relevant financial conflicts of all OncoKB personnel.

L. Gene Pages

Gene-specific data in OncoKB can be found on individual gene pages (Fig. 23). Note: Not all genes in the OncoKB Cancer Genes List have gene pages in OncoKB. Gene pages include the following information:

Oncokb

Levels of Evidence Actionable Genes Data Access News Usage Terms More +

Memorial Sloan Kettering Cancer Center

Oncokb

Precision Oncology Knowledge Base

595 Genes

4474 Alterations

38 Tumor Types

79 Drugs

Level 1 FDA-approved 20 Genes

Level R2 Clinical evidence 6 Genes

BRAF

BRAF (Entrez Gene: 673) Highest level of evidence: Level 1

BRAF / BRAF-CCDC fusion

The BRAF-CCDC fusion is likely oncogenic.

BRAF / BRAF-COKSRAF2 fusion

The BRAF-COKSRAF2 fusion is likely oncogenic.

BRAF / BRAF-LUCT2 fusion

The BRAF-LUCT2 fusion is likely oncogenic.

BRAF / BRAF-NETIL2 fusion

Oncokb is intended for research purposes only. Please review the [Usage Terms](#) before continuing.
When using Oncokb, please cite: [Jenner et al. \(2020\) 2021](#)

HSK () CHO () Quest Diagnostics () ctDNA () Oncotree ()

Usage Terms | Contact us | Twitter | API
Last update January 24, 2019

Memorial Sloan Kettering Cancer Center

© 2019 Memorial Sloan Kettering Cancer Center

OncKB Levels of Evidence Actionable Genes Data Access News Usage Terms More +

 Memorial Sloan Kettering Cancer Center

BRAF 119 annotated alterations

Oncogenes

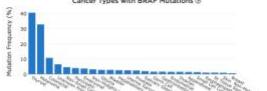
Highest level of evidence: Level 1

Also known as NS7, B-rat, BRAF1, RAF1B1, B-RAF1
Isoform: ENST00000288602 RefSeq: NM_004533.4

BRAF, an intracellular kinase, is frequently mutated in melanoma, thyroid and lung cancers among others.

See BRAF background @

Annotated Mutation Distribution in MSK-IMPACT Clinical Sequencing Cohort (Zehir et al., Nature Medicine, 2017)



Estimated data from chart:

Cancer Type	Estimated Mutation Frequency (%)
Melanoma	~38
Lung Adenocarcinoma	~32
Skin Cutaneous Melanoma	~28
Colon Adenocarcinoma	~18
Esophagus Adenocarcinoma	~15
Stomach Adenocarcinoma	~12
Bladder Transitional Cell Carcinoma	~10
Prostate Adenocarcinoma	~8
Breast Invasive Carcinoma	~7
Head and Neck Squamous Cell Carcinoma	~6
Leukemia Acute Myeloid	~5
Leukemia Chronic Myeloid	~4
Stomach Gastroesophageal Junction Adenocarcinoma	~3
Esophagus Squamous Cell Carcinoma	~2
Other	<2

CLINICAL RELEVANT ALTERATIONS (16) All Annotated Alterations (119)

If you notice any mistakes or missing alterations / citations, please send an email to heatdb@cbioinfo.org.

Search

Alteration	Cancer Type	Drug(s)	Level	Citations
E1001	Ovarian Cancer	Cobimetinib, Tremelimumab	3A	3 references
D287H	All Tumors	PLX4394	4	3 references
V599L	All Tumors	PLX4394	4	3 references
G164	All Tumors	PLX4394	4	3 references
G166	All Tumors	PLX4394	4	3 references
S472L	All Tumors	PLX4394	4	3 references
G169	All Tumors	PLX4394	4	3 references

Fig. 23: BRAF gene Page. (a) Searching for a specific gene will highlight all possible links and take you to the appropriate gene or alteration page. (b) BRAF Gene page shown here as an example.

1 Gene summary:

The gene summary at the top of the gene page contains the following elements (Fig. 24).

- The Summary at the top of the gene page contains the following elements (Fig. 24):

 - a. Gene name and its total number of annotated alterations in OncoKB
 - b. Evidence-based classification of the gene as either oncogene and/or tumor suppressor
 - c. The highest gene-associated Level of Evidence (if any)
 - d. Gene-name aliases
 - e. OncoKB utilized gene isoform and RefSeq ID

Additionally, the gene summary has 1-2 sentences detailing the functional role of the gene in a cell and in which tumor types it is frequently altered.

BRAF 119 annotated alterations

Oncogene

Highest level of evidence: Level 1

Also known as NS7, B-raf, BRAF1, RAFB1, B-RAF1
Isoform: ENST00000288602 RefSeq: NM_004333.4

BRAF, an intracellular kinase, is frequently mutated in melanoma, thyroid and lung cancers among others.

See *BRAF* background

Fig. 24: Gene Summary. BRAF shown as an example.

2. **Gene background:** Clicking the “See [Gene] background” below the Gene Summary expands the “Gene Background” text (see Fig. 25, example gene is BRAF), which describes the role of the gene-encoded protein in normal cells, its function in tumorigenesis, and its prevalence and mutation pattern in relevant tumor types. PMIDs in the gene background link out to the referenced paper abstract in PubMed in a new browser page.

Hide BRAF background

BRAF is a serine/threonine kinase that plays a key role in the regulation of the mitogen-activated protein kinase (MAPK) cascade ([PMID: 15520807](#)), which under physiologic conditions regulates the expression of genes involved in cellular functions, including proliferation ([PMID: 24202393](#)). Genetic alterations in BRAF are found in a large percentage of melanomas, thyroid cancers and histiocytic neoplasms as well as a small fraction of lung and colorectal cancers. The most common BRAF point mutation is V600E, which deregulates the protein's kinase activity leading to constitutive BRAF activation, as BRAF V600E can signal as a monomer independently of RAS or upstream activation ([PMID: 20179705](#)). Other BRAF mutations have been found that affect the protein's propensity to dimerize ([PMID: 16858395, 26343582, 12068308](#)). The product of these alterations is a BRAF kinase that can activate MAPK signaling in an unregulated manner and, in some instances, is directly responsible for cancer growth ([PMID: 15520807](#)). Inhibitors of mutant BRAF, including vemurafenib and dabrafenib, are FDA-approved for the treatment of late-stage or unresectable melanoma.

Fig. 25: Gene Background. BRAF shown as an example.

3. **Gene-specific “Cancer Types” histogram:** Fig. 26 shows the mutation frequency of the gene in different tumor types. The Y-axis shows the percent of samples that carry a mutation in the specific genes (including missense mutations, truncating mutations, frameshift mutations, insertions, amplifications and deletions) and the X-axis specifies tumor type. Data for this histogram is sourced from the ~10,000 tumor samples of the MSK-IMPACT Clinical Sequencing Cohort (Zehir et al., 2017) and does not account for copy number changes, chromosomal translocations or cancer types with fewer than 50 samples. Clicking on a bar in the histogram changes the data in the lollipop plot to reflect the selected tumor type.

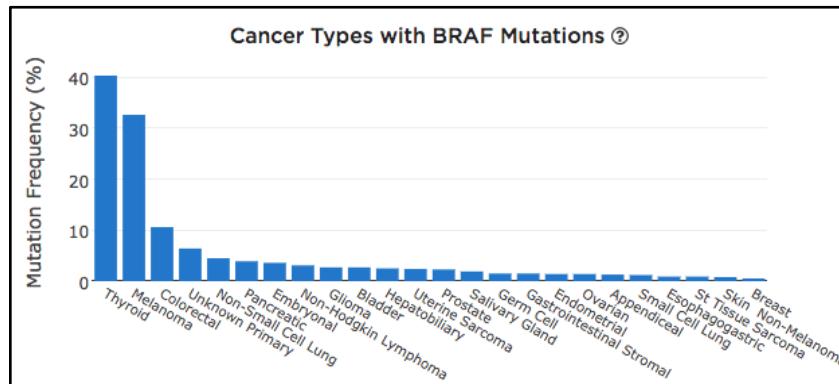


Fig 26: Gene-specific Cancer-Types histogram. BRAF shown as an example.

4. **Gene-specific lollipop plot:** The gene-specific lollipop plot is a schematic that displays the gene-encoded protein (Fig. 27). The X-axis of the plot is the amino acid position in the gene-encoded protein and the Y-axis of the plot is mutation count. On this schematic, the location of each mutation on the protein is indicated by a “lollipop”, and the height of the lollipop signifies the mutational frequency of the mutant allele. Data for this histogram is sourced from the 10,000 tumor samples of the MSK-IMPACT Clinical

Sequencing Cohort (Zehir et al., 2017). Clicking a specific mutation (or clicking a single lollipop) restricts the Alterations table (see below in this section, Item L.5.b) to display oncogenic and actionability information (if any) associated with the selected mutation. Clicking on a tumor type in the “Cancer Types with [Gene] Mutations” histogram will restrict the displayed mutations in the lollipop plot to only those found in the selected tumor type. To undo the tumor type filter, the user can click “Current view shows filtered results. Click here to reset all filters”. The user can customize the plot and download it as a PDF or SVG file using the buttons that appear when the user hovers over the upper right of the lollipop plot.

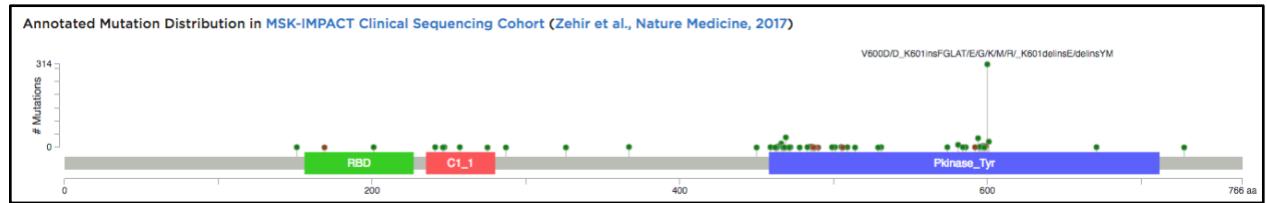


Fig. 27: The gene-specific lollipop plot based on the published MSK clinical sequencing cohort in Zehir et al 2017. BRAF shown as an example.

5. Clinically Relevant and All Annotated Alterations tables: Below the lollipop plot are two tabs, the Clinically Relevant Alterations and the All Annotated Alterations tables (Fig. 28). Both tables are searchable using the search bar indicated on the right-hand side of the table. By default, the Clinically Relevant Alterations table is shown. Each column in both tables is sortable. Clicking on the alteration brings you to the individual alteration page. Hovering over the citation column reveals a dialogue box that lists the title, citation and PMID of each source used to support the association. Clicking on either the title or the PMID will link out to the referenced paper abstract in PubMed in a new browser page.

- a. Clinically Relevant Alterations (# of alterations): Gene-specific alterations associated with a level of evidence indicating potential clinical actionability are shown in this tab (Fig. 28a) which lists:
 - i. Clinically Relevant Alterations: Gene alteration considered clinically relevant
 - ii. Tumor type in which the alteration is considered clinically relevant
 - iii. Drug(s) associated with the clinical relevance of the alteration
 - iv. Level of evidence for the alteration-tumor type-drug association
 - v. Relevant citations
- b. All Annotated Alterations (# of alterations): All OncoKB curated Gene-specific alterations are shown in this tab (Fig. 28b):
 - i. Gene alteration
 - ii. Oncogenic status: Yes, Likely, Neutral, Likely Neutral or Inconclusive
 - iii. Mutation Effect: Gain-, Loss-, Switch-of-function, Neutral, Likely Gain-, Likely Loss-, Likely switch-of-function, Likely Neutral, Inconclusive.
 - iv. Citations: Citation number is listed with mouse-over dialogue box that lists the title, citation and PMID of all references. Clicking on either the title or the PMID links out to the referenced paper abstract in PubMed in a new browser page.
 - v. Clicking on the alteration links to the individual alteration page (discussed in this chapter Section II.M below).

(a)

The table displays 16 entries of clinically relevant alterations, each with a link to the individual alteration page. The columns are: Alteration, Cancer Type, Drug(s), Level, and Citations. The table includes rows for V600E in Non-Small Cell Lung Cancer, V600E in Anaplastic Thyroid Cancer, V600E in Melanoma, V600K in Melanoma, V600 in Erdheim-C Chester Disease, and L597 in Melanoma.

Clinically Relevant Alterations (16)		All Annotated Alterations (119)		
If you notice any mistakes or missing alterations / citations, please send an email to feedback@oncokb.org .				
Search: _____				
Alteration	Cancer Type	Drug(s)	Level	Citations
V600E	Non-Small Cell Lung Cancer	Dabrafenib + Trametinib	1	2 references
V600E	Anaplastic Thyroid Cancer	Dabrafenib + Trametinib	1	1 reference
V600E	Melanoma	Vemurafenib Dabrafenib Dabrafenib + Trametinib Vemurafenib + Cobimetinib Trametinib Encorafenib + Binimatinib	1	16 references
V600K	Melanoma	Dabrafenib + Trametinib Vemurafenib + Cobimetinib Trametinib Encorafenib + Binimatinib	1	11 references
V600	Erdheim-Chester Disease	Vemurafenib	1	2 references
L597	Melanoma	Trametinib	3A	2 references

Clinically Relevant Alterations (16) All Annotated Alterations (119)

If you notice any mistakes or missing alterations / citations, please send an email to feedback@oncokb.org.

Search:

Alteration	Oncogenic	Mutation Effect	Citations
Amplification	Yes	Gain-of-function	5 references
AGK-BRAF Fusion	Yes	Gain-of-function	3 references
Fusions	Likely	Gain-of-function	11 references
V600_K601delinsE	Yes	Gain-of-function	2 references
V600D_K601insFGLAT	Yes	Gain-of-function	1 reference
T599_V600insV	Yes	Gain-of-function	1 reference
T599_V600insEAT	Likely	Likely Gain-of-function	1 reference
T599_V600insETT	Likely	Likely Gain-of-function	1 reference
R506_K507insVLR	Likely	Gain-of-function	1 reference
P490_Q494del	Likely	Likely Gain-of-function	1 reference
T488_P492del	Likely	Likely Gain-of-function	1 reference
Y487_P492delinsA	Likely	Likely Gain-of-function	1 reference
N486_P490del	Likely	Likely Gain-of-function	2 references

(b)

Fig. 28: Clinically Relevant and All Annotated Alterations Tables. BRAF shown as an example. (a) The Clinically Relevant Alterations (with 16 total alterations) is selected by default (in black). The All Alterations table (with 119 total alterations) is clickable in blue. (b) When selected, the All Annotated Alterations tab (with 119 total alterations) is shown.

M. Alteration Pages

Similar to gene-specific data, alteration-specific data in OncoKB can be found on individual alteration pages. Typing the alteration into the homepage or OncoKB header search bars can access these pages (Fig. 29, BRAF V600E example shown). Alterations across all pages in oncokb.org also link to their respective Alteration pages. Note, not all alterations have alteration pages.

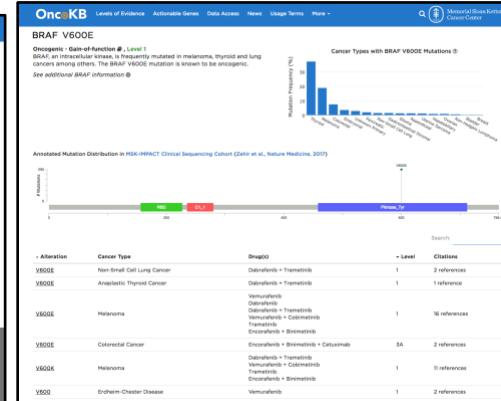


Fig. 29: BRAF V600E Alteration Page. (a) Searching for a specific alteration will highlight all possible alterations to select from and take the user to the appropriate alteration page. (b) BRAF V600E Alteration page shown as an example.

Each Alteration page has the following information:

- Gene and alteration name.**
- Evidence-based classification of the oncogenic effect** of the alteration (see Chapter 4, Section II.A). Possible classifications include Oncogenic, Likely Oncogenic, Neutral, Likely Neutral, Inconclusive
- Evidence-based classification of the biological effect** of the alteration (see Chapter 4, Section II.B). Possible classifications include Gain-, Loss-, Switch-of-function, Neutral, Likely Gain-, Likely Loss-, Likely Switch-of-function, Likely Neutral, Inconclusive.
- Evidence-based classification of the clinical effect** of the alteration **and its highest alteration-associated therapeutic Level of Evidence (if any)**. Possible levels of evidence include the following (refer to Chapter 5):
 - Therapeutic: Levels 1, 2A, 2B, 3A, 3B, 4, R1 and R2
 - Diagnostic: Levels Dx1, Dx2, Dx3
 - Prognostic: Levels Px1, Px2, Px3
- Gene summary:** See Item L.1 in this section and Fig. 24 and 30

- 6. Alteration summary:** Similar to item described in Item M.2 above (in this section) summary of the evidence-based classification of the oncogenic effect of the alteration (see Chapter 4, II.A) is given in sentence form (highlighted in blue in Fig. 30).

BRAF V600E

Oncogenic · Gain-of-function  , **Level 1**

BRAF, an intracellular kinase, is frequently mutated in melanoma, thyroid and lung cancers among others. **The BRAF V600E mutation is known to be oncogenic.**

Fig. 30: Alteration Summary. In addition to the gene summary the alteration summary is also shown in the Alteration page (highlighted in blue). BRAF V600E shown as an example.

- 7. Additional gene information:** Information is described in Items L.1, Fig. 24 and Item L.2 Fig. 25 in this section. Briefly, whether the gene is an oncogene or tumor suppressor, the highest level of evidence associated (if any), the gene aliases and the gene background, with PMIDs that link directly to the reference.
- 8. Alteration-specific “cancer types” histogram:** The Cancer Types histogram (Fig. 31) shows the frequency of the specific alteration in different tumor types. The Y-axis shows the percent of samples that carry the specific alteration and the X-axis specifies the tumor type. Data for this histogram is sourced from the ~10,000 tumor samples of the MSK-IMPACT Clinical Sequencing Cohort (Zehir et al., 2017). Alteration pages for copy number changes or chromosomal translocations do not have this histogram.

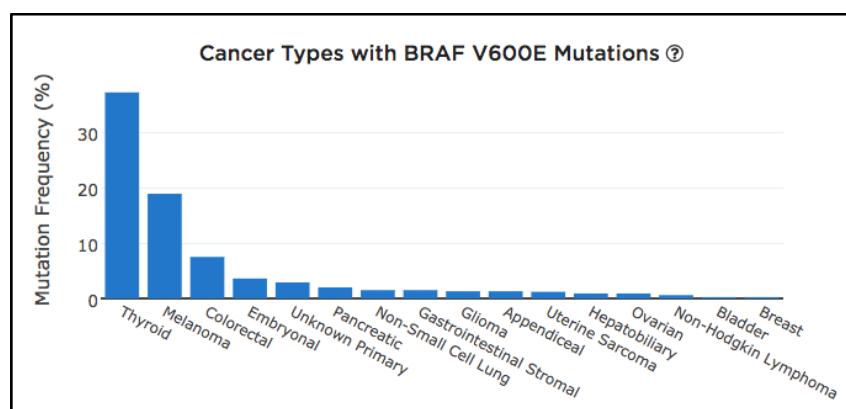


Fig. 31: Alteration-specific Cancer Types histogram. BRAF V600E is shown as an example.

- 9. Alteration-specific lollipop plot schematic:** The alteration-specific lollipop plot shows the position of the alteration in the gene-encoded protein and the tumor-type-specific mutational count of the specific mutant allele (as indicated by the height of the lollipop) (Fig. 32). Similar to the lollipop plot in Item L.4 and Fig. 27 of this section, the X-axis of this schematic is the amino acid position in the gene-encoded protein and the Y-axis of the plot is mutation count. Data for this histogram is sourced from the ~10,000 tumor samples of the MSK-IMPACT Clinical Sequencing Cohort (Zehir et al., 2017).

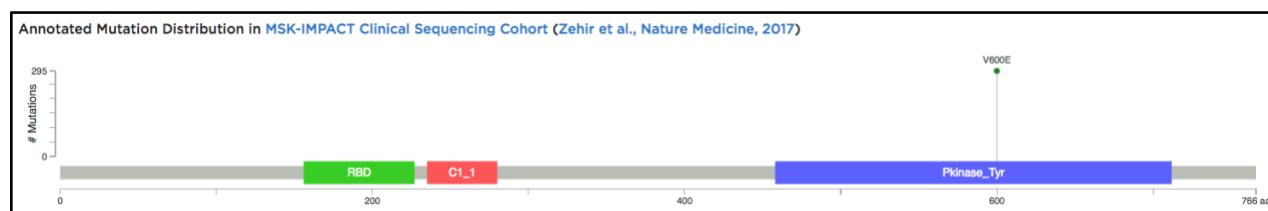


Fig. 32: The alteration-specific lollipop plot based on the published MSK clinical sequencing cohort in Zehir et al 2017. BRAF is shown as an example.

- 10. Alteration-specific table:** Alterations with no associated level of evidence will not have an alteration-specific table. For clinically relevant alterations associated with a level of evidence indicating potential clinical actionability, an alteration-specific table (Fig. 33) will list the following:
- Gene alteration considered clinically relevant
 - Cancer type in which the alteration is considered clinically relevant

- c. Drug(s) associated with the alterations clinical relevance
- d. Level of evidence for the alteration-tumor-type-drug association
- e. Relevant citations

Each column in the table is sortable. Clicking on the alteration brings you to the individual alteration page. Hovering over the citation column reveals a dialogue box that lists the title, citation and PMID of each source used in support of the association. Clicking on either the title or the PMID will link out to the referenced paper abstract in PubMed in a new browser page.

					Search:
Alteration	Cancer Type	Drug(s)	Level	Citations	
V600E	Non-Small Cell Lung Cancer	Dabrafenib + Trametinib	1	2 references	
		Vemurafenib Dabrafenib Dabrafenib + Trametinib Vemurafenib + Cobimetinib Trametinib Encorafenib + Binimatinib	1	16 references	
V600E	Melanoma	Dabrafenib + Trametinib	1	1 reference	
V600E	Anaplastic Thyroid Cancer	Encorafenib + Binimatinib + Cetuximab	3A	2 references	
V600K	Melanoma	Dabrafenib + Trametinib Vemurafenib + Cobimetinib Trametinib Encorafenib + Binimatinib	1	11 references	
V600	Erdheim-Chester Disease	Vemurafenib	1	2 references	
V600	Colorectal Cancer	Panitumumab + Dabrafenib + Trametinib	3A	8 references	

Fig. 33: Alteration-specific tables. Alteration-specific tables are only available for those alterations associated with a level of evidence. The alteration specified as well as leveled therapeutic evidence related to the specified alteration are displayed. BRAF V600E is shown as an example.

11. Feedback through OncoKB.org: Assertion feedback by OncoKB users is an important feature of the knowledgebase. There are two web-based mechanisms through which users may provide feedback on OncoKB content: 1) the OncoKB website, and 2) via the cBioPortal for Cancer Genomics. Any feedback, comments or questions may also be sent via email to contact@oncokb.org, which is provided in multiple places within the OncoKB website (Fig. 34). Emails sent to contact@oncokb.org are received by the Lead Scientist and all SCMT members and answered within 48 hours.

Oncokb Levels of Evidence Actionable Genes Data Access News Usage Terms More

While we aim to keep the information up to date and correct, there will inevitably be gaps or mistakes. Please help us to identify any issues by sending an email to contact@oncokb.org, or use the feedback button that appears next to alterations in cBioPortal.

Stay tuned for future data updates (improved annotations, new alterations), as well as new features. You can follow us on Twitter (@OncoKB) or subscribe to our low-volume email list for updates.

When using OncoKB, please cite: Chakravarty et al., JCO PO 2017.

January 24, 2019

Updated Actionable Genes

Level	Update
Level 1	November 2, 2018: the FDA approved Iorlatinib for patients with anaplastic lymphoma kinase (ALK)-positive metastatic non-small cell lung cancer (NSCLC) whose disease has progressed on crizotinib and at least one other ALK inhibitor or whose disease has progressed on alectinib or ceritinib for metastatic disease.

Addition of 16 new gene-associated genes:
ECT2L RELN TALI MLLT10 TLX3 TLX1 TRA TRB TRD TRG EPOR ABL2 MECOM DEK RBM15 BCL9

Oncokb Levels of Evidence Actionable Genes Data Access News Usage Terms More

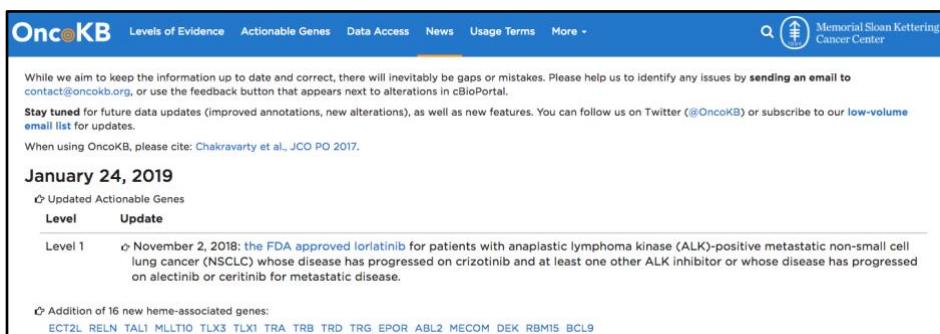
While we aim to keep the information up to date and correct, there will inevitably be gaps or mistakes. Please help us to identify any issues by sending an email to contact@oncokb.org, or use the feedback button that appears next to alterations in cBioPortal.

Stay tuned for future data updates (improved annotations, new alterations), as well as new features. You can follow us on Twitter (@OncoKB) or subscribe to our low-volume email list for updates.

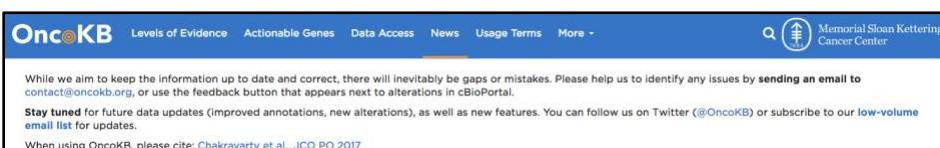
When using OncoKB, please cite: Chakravarty et al., JCO PO 2017.

OncoKB is intended for research purposes only. Please review the [usage terms](#) before continuing.
 When using OncoKB, please cite: Chakravarty et al., JCO PO 2017.
[MSK](#) | [CMO](#) | [Quest Diagnostics](#) | [cBioPortal](#) | [OncoTree](#)

(a)



(b)



(c)

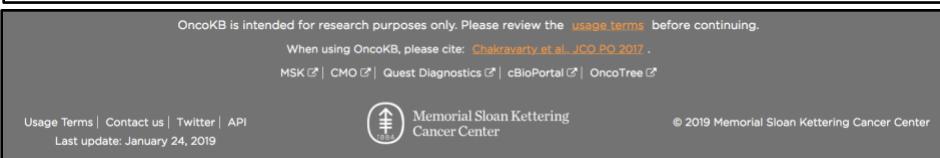


Fig. 34: Feedback through Oncokb.org. Users of Oncokb.org may provide feedback on the website by clicking the email link for contact@oncokb.org (a) In the News section, (b) In the Usage Terms section, or by clicking “Contact Us” in (c) the OncoKB webpage footer.

III. OncoKB content accessible through cBioPortal

The OncoKB knowledgebase is integrated into cBioPortal (cbioportal.org) through annotation of mutation effect, oncogenic effect and level of evidence of alterations visualized on the platform. The portal uses the Mutation Annotation Format or MAF as an input file for genomic data.

- The MAF is then annotated via Genome Nexus to assign an amino acid change based on the genomic coordinates, chromosome number and ref and tumor allele change. This is documented in the MAF via HGVS_P_Short.
- The MAF is then annotated with OncoKB annotations using the tool/oncokb-annotator which is also available on GitHub: <https://github.com/oncokb/oncokb-annotator>.

A. OncoKB icons in cBioPortal

OncoKB icons are coded and used in cBioPortal OncoKB annotation to communicate the oncogenic and biological effect and actionability of a given variant. The following are the rules of the icons used in portal:

1. cBioPortal uses the OncoKB symbols to signify information known about the variant
2. In addition to specifying the oncogenic effect, the portal icon will display the highest levels of evidence for the given variant and the tumor type.
3. “Predicted variants” are mutations that are hotspots from Chang et al., 2016 and Chang et al., 2018 but that are not specifically annotated in the system.

B. OncoKB Cards in cBioPortal

OncoKB information is displayed in cBioPortal in OncoKB cards that appear when the user hovers over the OncoKB icon that is next to an alteration in the mutation table in the “mutations” tab of a gene query or in the Patient View of a sample in the Mutations tab.

The card is divided into the following sections:

1. **Header:** The header lists the gene, alteration, and tumor type of the respective sample
2. **Clinical Implications:** The clinical implications tab (Fig. 35) describes the oncogenicity of the alteration. This section is clickable and changes the information in the “description” space directly below.
3. **Description:** By default, the information displayed in the description section is the “clinical implications” information. The “clinical implications” information includes the:
 - a. Gene summary
 - b. Mutation summary
 - c. Tumor type summary
 - d. Clinical actionability table: The information in this table includes:
 - i. Level of evidence icon: if the user hovers over the icon, the definition of the level is displayed. While the OncoKB icon on the “mutations” tab displays the highest level of evidence for the alteration, the OncoKB Card lists all levels of evidences associated with the alteration.
 - ii. Alteration associated with the level of evidence
 - iii. Drugs associated with the level of evidence
 - iv. Tumor type associated with the level of evidence
 - v. Citation icon: Upon mouse-over, this icon shows sources associated with each leveled evidence.

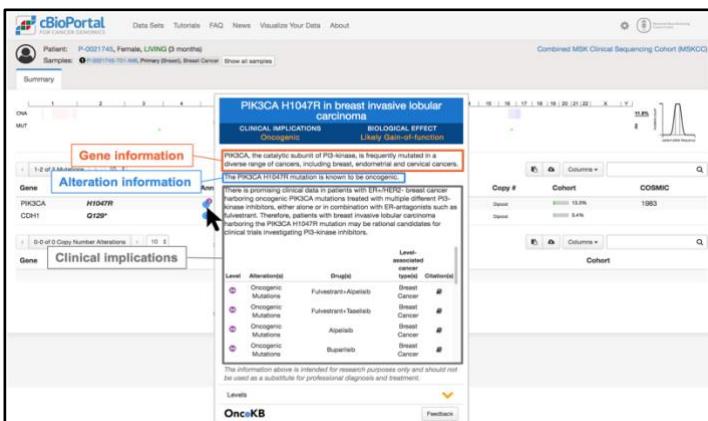


Fig. 35: The OncoKB Card in cBioPortal, Clinical Implications: Hovering over the OncoKB card in the patient view or mutations tab in cBioPortal will display the OncoKB card. Gene-specific information is outlined orange, alteration-specific information is outlined in blue and clinical implications (if relevant for the specified tumor-type) is displayed in grey.

- 4. Biological Effect:** The biological effect tab (Fig. 36) describes the biological effect of the alteration (whether the alteration is gain-of-function, loss-of-function, neutral, etc). Clicking on the biological effect tab in the gene card will switch views to display the biological effect of the alteration. In this section of the OncoKB card, an evidence-based classification of the biological effect of the alteration is provided and the list of references supporting this classification.

Fig. 36: The OncoKB Card in cBioPortal, Biological Effect: Clicking on the biological effect tab in the OncoKB gene card shows a list of references that support the assertion of the biological effect shown in dark blue (example shown here; the PIK3CA H1047 mutation [found in breast invasive lobular carcinoma] is Gain-of-function) and link out to the respective PubMed Abstract page.

- 5. Levels:** Levels in the OncoKB card (Fig. 37) refers to the Levels of Evidence that support the mutation being predictive of response to the targeted therapies. Clicking on the down arrow next to “Levels” reveals a drop down description of all the OncoKB levels of evidence (both sensitivity and resistance).

Level	Alteration(s)	Drug(s)	Level-associated cancer type(s)	Citation(s)
1	Oncogenic Mutations	Alpelisib	Breast Cancer	
1	Oncogenic Mutations	Alpelisib+Fulvestrant	Breast Cancer	
1	Oncogenic Mutations	Buparlisib	Breast Cancer	
1	Oncogenic Mutations	Copanlisib	Breast Cancer	
1	Oncogenic	Fulvestrant+Ranibizumab	Breast	

The information above is intended for research purposes only and should not be used as a substitute for professional diagnosis and treatment.

Levels

1 FDA-recognized biomarker predictive of response to an FDA-approved drug in this indication
 2A Standard care biomarker predictive of response to an FDA-approved drug in this indication
 2B Standard care biomarker predictive of response to an FDA-approved drug in another indication, but not standard care for this indication
 3A Compelling clinical evidence supports the biomarker as being predictive of response to a drug in this indication
 3B Compelling clinical evidence supports the biomarker as being predictive of response to a drug in another indication
 4 Compelling biological evidence supports the biomarker as being predictive of response to a drug
 R1 Standard care biomarker predictive of resistance to an FDA-approved drug in this indication
 R2 Compelling clinical evidence supports the biomarker as being predictive of resistance to a drug

OncokB

Fig. 37: Levels in the OncoKB card: Clicking on the yellow arrowhead in the OncoKB card displays a glossary of the definition of the Levels of evidence.

- 6. OncoKB website and feedback:** Clicking on the OncoKB logo will bring the user to the OncoKB.org website. Clicking on “Feedback” (Fig. 38a) results in a pop-up comment card (Fig. 38b) that allows the user to provide feedback about the gene-alteration combination directly to the OncoKB team via Google forms. In the “OncoKB Annotation Feedback” pop-up form, information about the gene and alteration, the email address used to log into the portal, and the web-address of the specific portal instance will be pre-populated in the feedback form. Users may then enter specific feedback and associated references in the Feedback and References fields before

submitting the feedback. Submission of feedback by a cBioPortal user will auto-populate in a Google spreadsheet with all information entered above. Changes to this Google Sheet will trigger an automatic email sent to the Lead Scientist and SCMT alerting them of user feedback via cBioPortal and will be answered within 48 hours. Upon completion of any necessary deliverables as suggested by the feedback (either curation or software related), the appropriate OncoKB staff member fills in the “Complete” column and adds their initials as well as any comments related to the feedback item (Fig. 38c). The Feedback Page collates all cBioPortal user feedback related to OncoKB assertions and is a log of OncoKB development based on cBioPortal user-feedback.

(a)

(b)

(c)

Fig. 38: OncoKB Feedback through cBioPortal. On cBioPortal, if hovering over the OncoKB icon, a pop up with OncoKB information appears (a), clicking on the OncoKB icon in the pop-up will take users to the OncoKB homepage, clicking on the “Feedback” button in cBioPortal results in a pop-up comment card (b) that allows the user to provide feedback about the OncoKB annotation on the specific variant. User feedback is auto-populated into a google spreadsheet (c) which the OncoKB SCMT accesses and uses to answer user questions with a 48-hour turn-around period.

APPENDIX

Table A1. OncoKB icons in cBioPortal. For each oncogenic effect, the most common biological effects assigned to OncoKB variants are shown.

OncoKB Icon	Oncogenic Effect	Biological Effect
		Gain-of-Function (GOF) / Likely GOF
	Oncogenic	Loss-of-Function (LOF) / Likely LOF
		Switch-of-Function (SOF) / Likely SOF
		Likely GOF
	Likely Oncogenic	Likely LOF
		Likely SOF
		Neutral
	Likely Neutral	Likely Neutral
	Inconclusive	Inconclusive
	SCMT reviewed Variant of Unknown Significance (VUS)	SCMT reviewed VUS
	Unknown (SCMT non-reviewed VUS)	Unknown (SCMT non-reviewed VUS)

Table All. OncoKB Levels of Evidence and their icons in cBioPortal. Variants with clinical implications are given a specific OncoKB icon in cBioPortal as described here.

Level of Evidence (per Chakravarty et al. 2017)	OncoKB Icon in cBioPortal
 Level 1	
 Level 2A	
 Level 2B	
 Level 3A	
 Level 3B	
 Level 4	
 Level R1	
 Level R2	

REFERENCES

1. Chakravarty, D, Gao, JJ, Phillips, SM, Kundra, R, Zhang, H, Wang, J, Rudolph, JE, Yaeger, R, Soumerai, T, Nissan, MH, Matthew T. Chang, Sarat Chandarlapaty, Tiffany A. Traina, Paul K. Paik, Alan L. Ho, Feras M. Hantash, Andrew Grupe, Shrujal S. Baxi, Margaret K. Callahan, Alexandra Snyder, Ping Chi, Daniel C. Danila, Minal Gounder, James J. Harding, Matthew D. Hellmann, Gopa Iyer, Yelena Y. Janjigian, Thomas Kaley, Douglas A. Levine, Maeve Lowery, Antonio Omuro, Michael A. Postow, Dana Rathkopf, Alexander N. Shoushtari, Neerav Shukla, Martin H. Voss, Ederlinda Paraiso, Ahmet Zehir, Michael F. Berger, Barry S. Taylor, Leonard B. Saltz, Gregory J. Riely, Marc Ladanyi, David M. Hyman, José Baselga, Paul Sabbatini, David B. Solit, and Nikolaus Schultz. **OncoKB: A Precision Oncology Knowledge Base** *JCO Precision Oncology* (2017) 1:1, 1-16
2. Chang MT, Bhattacharai TS, Schram AM, Bielski CM, Donoghue MTA, Jonsson P, Chakravarty D, Phillips S, Kandoth C, Penson A, Gorelick A, Shamu T, Patel S, Harris C, Gao J, Sumer SO, Kundra R, Razavi P, Li BT, Reales DN, Socci ND, Jayakumaran G, Zehir A, Benayed R, Arcila ME, Chandarlapaty S, Ladanyi M, Schultz N, Baselga J, Berger MF, Rosen N, Solit DB, Hyman DM, Taylor BS. **Accelerating Discovery of Functional Mutant Alleles in Cancer.** *Cancer Discovery.* (2018) Feb;8(2):174-183. doi: 10.1158/2159-8290.CD-17-0321. Epub 2017 Dec 15. PMID: 29247016
3. Wu K, Hinson SR, Ohashi A, Farrugia D, Wendt P, Tavtigian SV, Deffenbaugh A, Goldgar D, Couch FJ. **Functional evaluation and cancer risk assessment of BRCA2 unclassified variants.** *Cancer Research* (2005) Jan 15;65(2):417-26. PMID: 15695382
4. Chang MT, Asthana S, Gao SP, Lee BH, Chapman JS, Kandoth C, Gao J, Socci ND, Solit DB, Olshen AB, Schultz N, Taylor BS. **Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity.** *Nature Biotechnology* (2016) Feb;34(2):155-63. doi: 10.1038/nbt.3391 Epub 2015 Nov 30. PMID: 26619011
5. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, Srinivasan P, Gao J, Chakravarty D, Devlin SM, Hellmann MD, Barron DA, Schram AM, Hameed M, Dogan S, Ross DS, Hechtman JF, DeLair DF, Yao J, Mandelker DL, Cheng DT, Chandramohan R, Mohanty AS, Ptashkin RN, Jayakumaran G, Prasad M, Syed MH, Rema AB, Liu ZY, Nafa K, Borsu L, Sadowska J, Casanova J, Bacares R, Kiecka IJ, Razumova A, Son JB, Stewart L, Baldi T, Mullaney KA, Al-Ahmadi H, Vakiani E, Abeshouse AA, Penson AV, Jonsson P, Camacho N, Chang MT, Won HH, Gross BE, Kundra R, Heins ZJ, Chen HW, Phillips S, Zhang H, Wang J, Ochoa A, Wills J, Eubank M, Thomas SB, Gardos SM, Reales DN, Galle J, Durany R, Cambria R, Abida W, Cerce A, Feldman DR, Gounder MM, Hakimi AA, Harding JJ, Iyer G, Janjigian YY, Jordan EJ, Kelly CM, Lowery MA, Morris LGT, Omuro AM, Raj N, Razavi P, Shoushtari AN, Shukla N, Soumerai TE, Varghese AM, Yaeger R, Coleman J, Bochner B, Riely GJ, Saltz LB, Scher HI, Sabbatini PJ, Robson ME, Klimstra DS, Taylor BS, Baselga J, Schultz N, Hyman DM, Arcila ME, Solit DB, Ladanyi M, Berger MF. **Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.** *Nature Medicine.* (2017) Jun;23(6):703-713. doi: 10.1038/nm.4333. Epub 2017 May 8. PMID: 28481359
6. Weinberg RA. **The Biology of Cancer.** Garland Science; New York: 2007. p.G:20, G28 and 82.
7. Vogelstein B1, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. **Cancer genome landscape.** *Science.* (2013) Mar 29;339(6127):1546-58. doi: 10.1126/science.1235122. PMID: 23539594
8. Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, Elledge SJ. **Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome.** *Cell.* (2013) Nov 7;155(4):948-62. doi: 10.1016/j.cell.2013.10.011. Epub 2013 Oct 31. PMID: 24183448
9. National Cancer Institute [Internet]. Bethesda, MD: National Cancer Institute (US); [cited 2014 Dec 04]. NCI Dictionaries; [cited 2014 Dec 04]. Available from:<http://www.cancer.gov/dictionary>
10. Hanahan D, Weinberg RA. **Hallmarks of cancer: the next generation.** *Cell.* (2011) Mar 4;144(5):646-74. doi: 10.1016/j.cell.2011.02.013. Review. PMID: 21376230
11. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. **The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.** *Cancer Discovery* (2012) May;2(5):401-4. doi: 10.1158/2159-8290.CD-12-0095. PMID: 22588877
12. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. **Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.** *Science Signaling.* (2013) Apr 2;6(269):pl1. doi: 10.1126/scisignal.2004088. PMID: 23550210