Molecular Targets Platform (MTP)

User Guide

08/04/2022

Version 1.0.0

Version	Date	Description	Author
1.0.0	08/04/2022	Initial Release	MTP Team

Table of Contents

Introduction and Overview

The Molecular Targets Platform (MTP) is being developed as an instance of the widely known and respected EMBL Open Targets platform as a part of the National Cancer Institute's (NCI) Childhood Cancer Data Initiative (CCDI). This initiative is tasked with building a set of tools and resources centered around childhood cancer research and patient care. We believe that through enhanced data sharing, we can improve our understanding of cancer biology so that new preventative measures and treatments may be uncovered. Our goal is to ensure that researchers learn from every child with cancer in order to extend the survivorship and quality of life for children with pediatric cancers.

Specifically, MTP is a collaborative effort between the Children's Hospital of Philadelphia (CHOP) and the Frederick National laboratory (FNL) with input from the National Cancer Institute (NCI) and Food Drug Administration (FDA). The MTP, a searchable graphical user interface, is being developed to allow systematic drug-target identification and prioritization based on existing evidence to ultimately impact childhood cancer treatment. The initial goal of MTP is to integrate FDA's RMTL (Relevant Molecular Target List) with pre-clinical and clinical pediatric cancer datasets having somatic alteration data (SNVs, CNVs and gene fusions) and gene expression data. Open Targets native data is also present within MTP. Our integration of GTEx data enables comparisons of Pediatric versus Adult data. Our use of common ontological hierarchies has aided in harmonizing disease, gene (target) and agent (drug) terms. A few examples of pediatric data resources that have been added to MTP are TARGET, KidsFirst, and OpenPedCan. Other data resources, such as PIVOT, will be added in subsequent releases.

This document describes a high-level view of the features and functionalities within MTP. Investigators are encouraged to explore MTP for themselves using this guide as a primer.

Main Landing Page

The main MTP landing page is where users can navigate to the one of several sites, namely: the FDA PMTL Page, an About Page and the Pediatric Cancer Data Navigation Page. Note: RMTL (Relevant Molecular Target List) has been renamed to PMTL (Pediatric Molecular Target List).

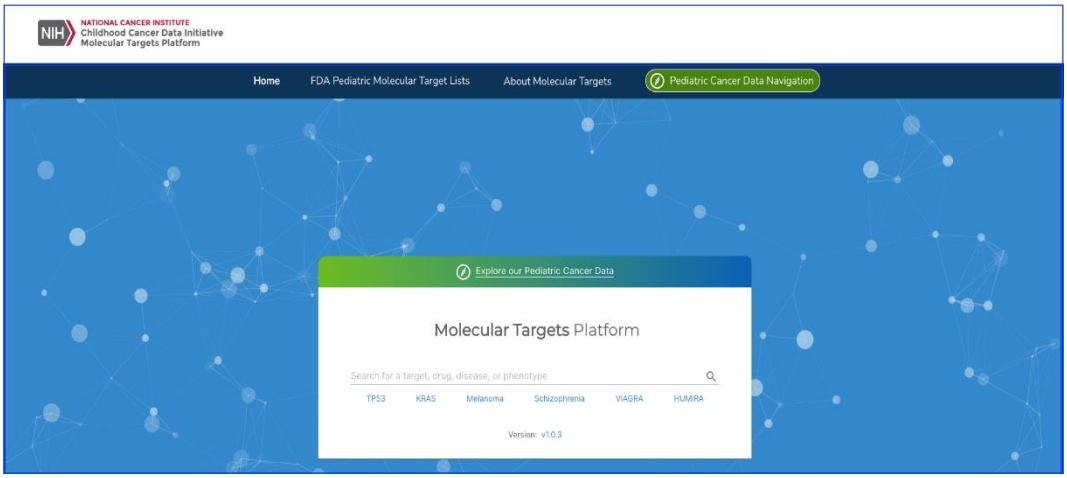

Figure 1: MTP Home Page

Pediatric Cancer Data Navigation Page

Users can find specific pediatric cancer data by clicking the link in the main page menu bar or the title bar to get to the Pediatric Cancer Data Navigation Page.

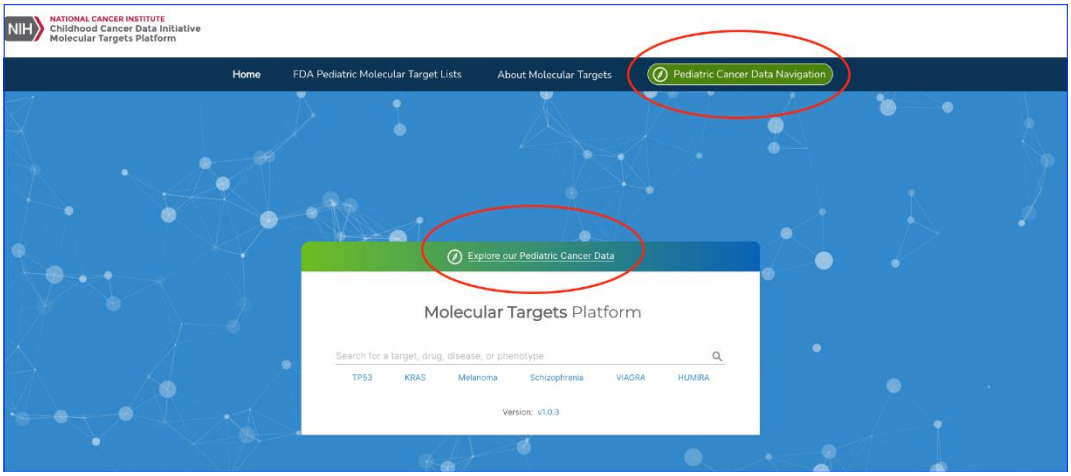

Figure 2: Accessing the Pediatric Cancer Data Navigation Page

On the Pediatric Cancer Data Navigation Search Page, users can query a Gene (target) and/or Disease by selecting their entry from the drop-down list and clicking "Search". Note that the initial release of MTP represents 41 pediatric diseases.

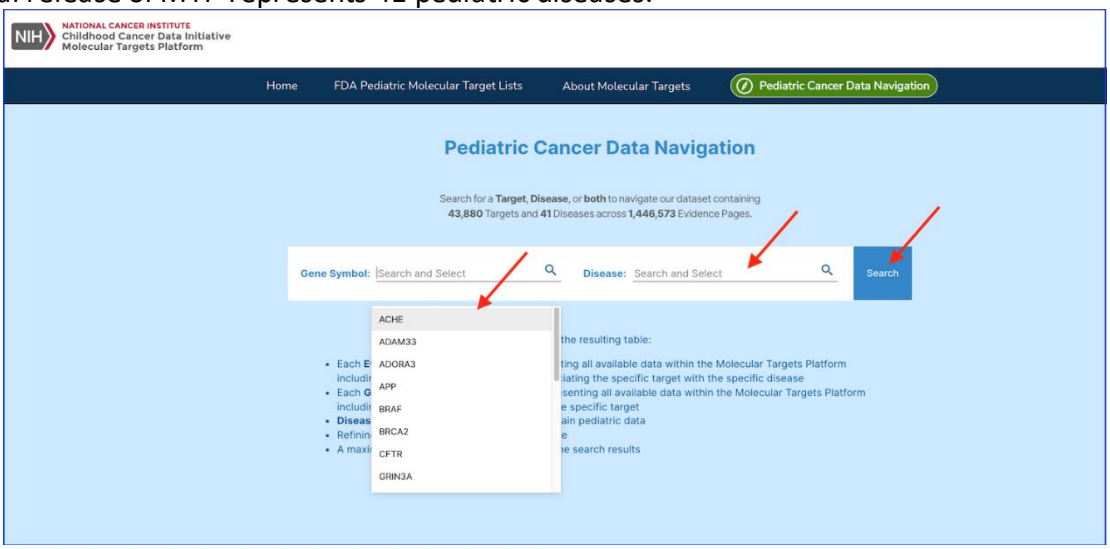

Figure 3: Performing a search on the Pediatric Cancer Data Navigation Page

Users view the returned search results on the Evidence page. The summary display allows the user to quickly assess which diseases have data for a given data type, indicated by the green checkmarks. Note: the presence of a green checkmark only indicates that data is present but does not indicate whether the data is normal or abnormal. This figure shows the results returned when searching for PTEN. Next, the desired Evidence Page for a particular disease can be displayed by clicking on the "Evidence Page". Links are also provided for overview information about the gene and disease as well.

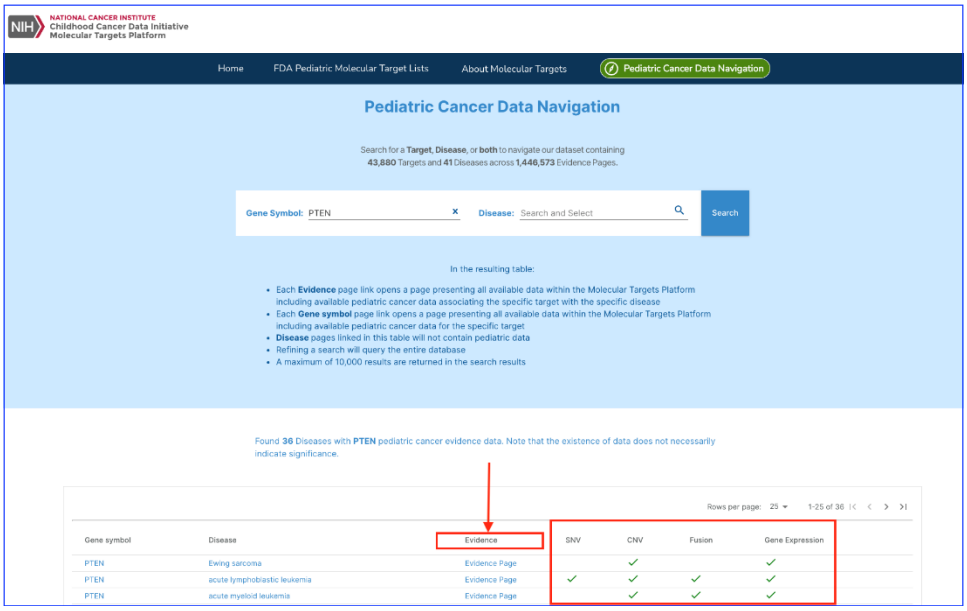

Figure 4: Results returned from a search on the Pediatric Cancer Data Navigation Page

Here, the evidence for PTEN in Acute Lymphoblastic Leukemia was chosen. Now the user can see the specific data types that are available within "widgets". Blue widgets are native to Open Targets while green widgets represent pediatric data that has been added and does not exist within native Open Targets. Gray widgets have no data.

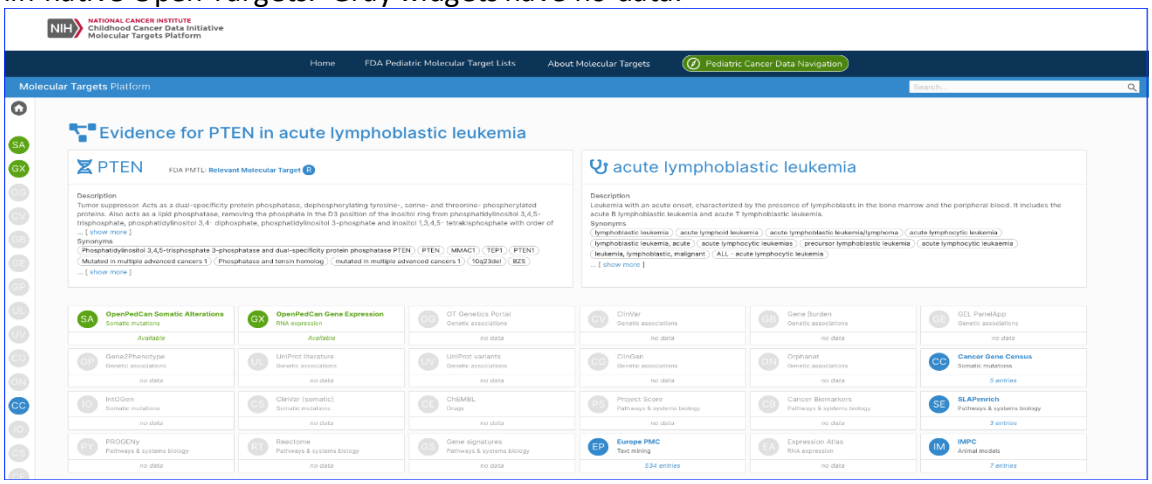

Figure 5: General Widget view of the returned Evidence for a given gene and disease

The user can click on the Somatic Alterations widget.

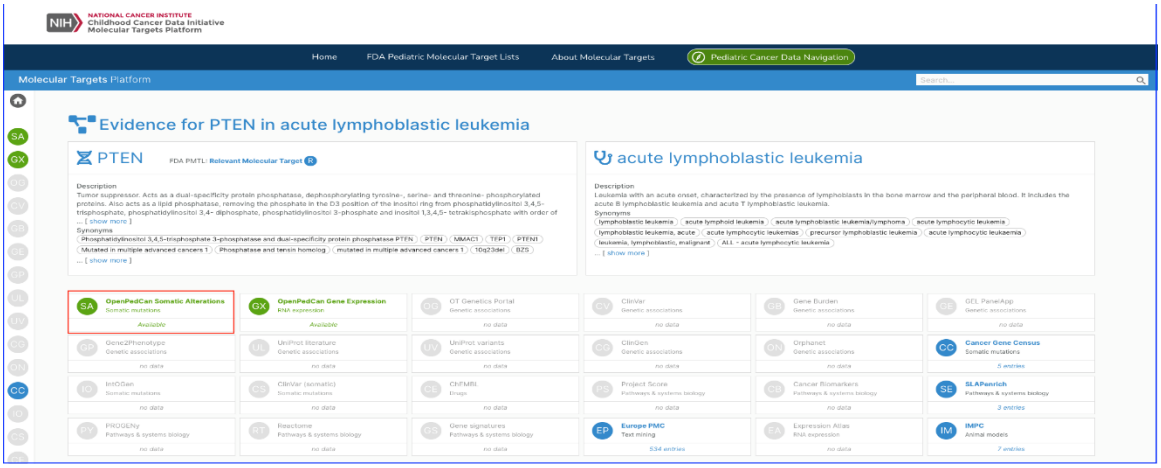

Figure 6: Somatic Alteration Widget View

Next, the user will see multiple tabs: SNV by Gene, SNV by Variant, CNV by Gene, Fusion by Gene and Fusion. Clicking on any tab shows the corresponding data for PTEN in Acute Lymphoblastic Leukemia along with many useful columns such as frequency in the dataset, % in the cohort and, in some cases, links to PedcBioPortal information. Active tabs are colored green. If a tab is not highlighted, there is no data for that data type. Some columns are filterable. If a gene is on the PMTL list, it is indicated by a blue circle with a white "R" inside. The data is downloadable in json and tsv formats and by API query.

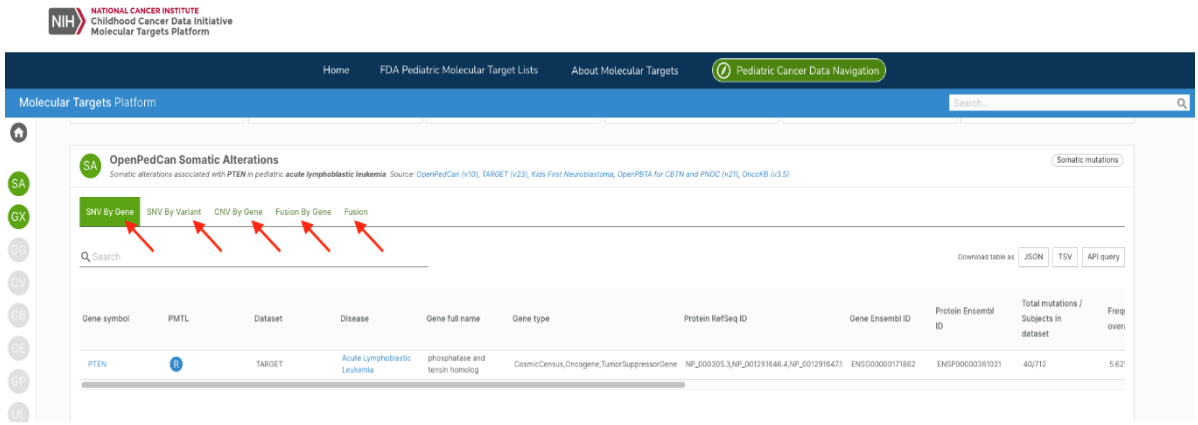

Figure 7: Table indicating Data Types within the Somatic Alteration Widget

Additionally, the user can click on the Gene Expression Widget.

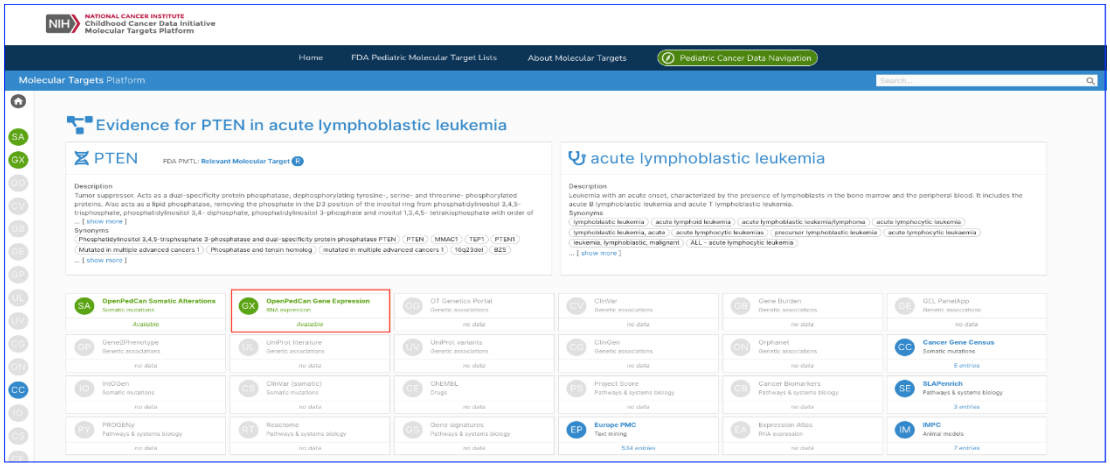

Figure 8: Gene Expression Widget

Next, the user will see choices between Linear and Log10 views of the graph of PTEN expression in Acute Lymphoblastic Leukemia versus the expression of GTEx normal tissue. The data is downloadable in json and tsv formats.

Figure 9: Gene Expression Widget view with graphical display choice tabs

Open Targets Data Search Page

At any time, the user can navigate back to the main landing page by clicking on the persistent title at the top left of the screen. Diseases, targets and drugs can be searched on this main page. Open Targets data will be displayed in the resulting screens. The native Open Targets data is mostly adult data. Note: To see the added Pediatric datasets, this page provides two links to the Pediatric Cancer Data Navigation page for a view of the Pediatric data, as mentioned above.

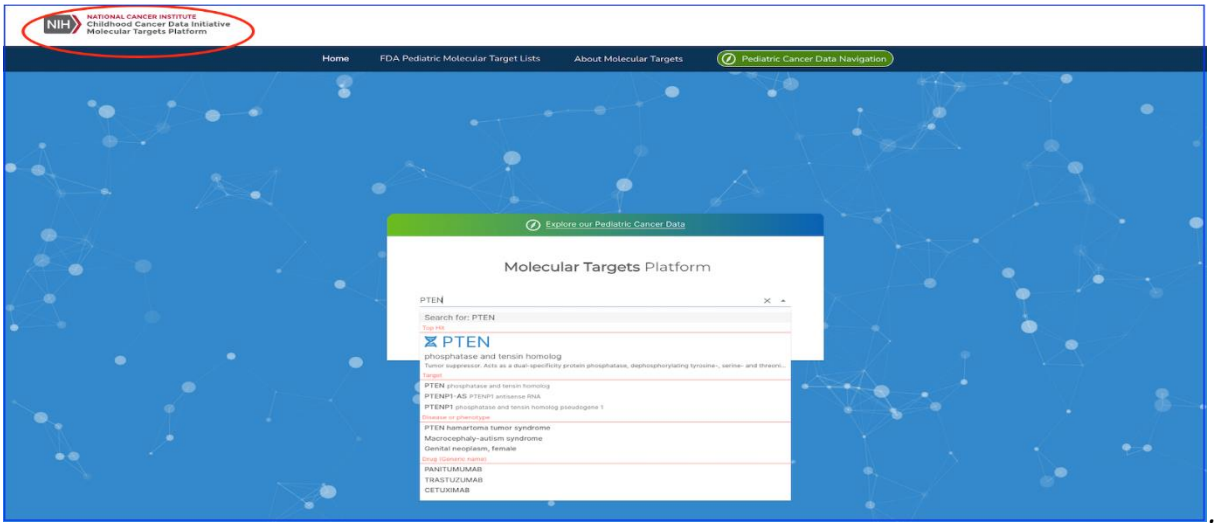

Figure 10: Open Targets Search page

When searching by the gene PTEN, the user can see two tabs presented. The "Associated Disease" tab shows all diseases ranked by descending evidence (highest ranking at the top) in a heatmap. The heatmap also shows what kind of data led to the score as well as an overall association score. The view allows searching by a specific gene in the search bar above and to the left of the heatmap. Alternately, one can begin by searching for a specific disease and an Associated genes heatmap will be displayed.

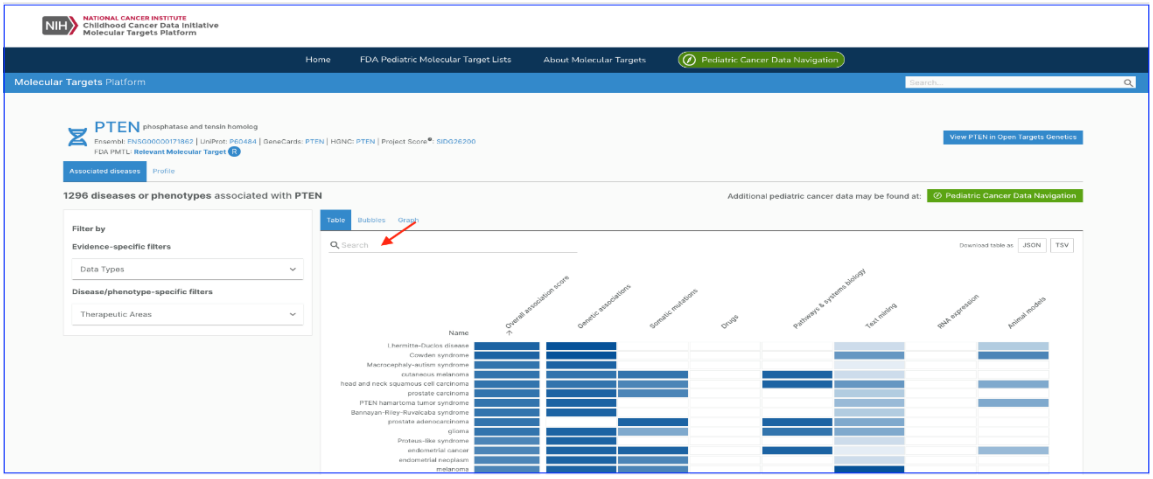

Figure 11: Open Targets Search page

About Molecular Targets Page

The About Page can be navigated to from the main landing page and has two pull-down choices. The choice of "About Molecular Target" will display a page with broad sections describing the derivation of the platform, Pediatric Cancer Diseases, Pediatric Cancer Data Sources, Data Processing methods and Pediatric Cancer Data Visualizations as well as other information.

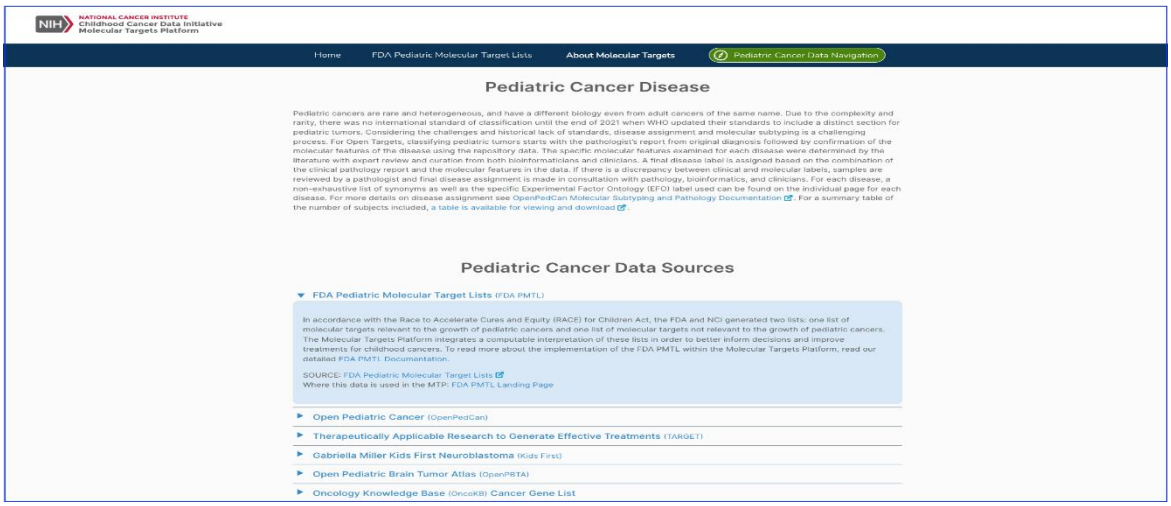

Figure 12: About Page content

The "Change Log" will provide the user with each version of a particular MTP release.

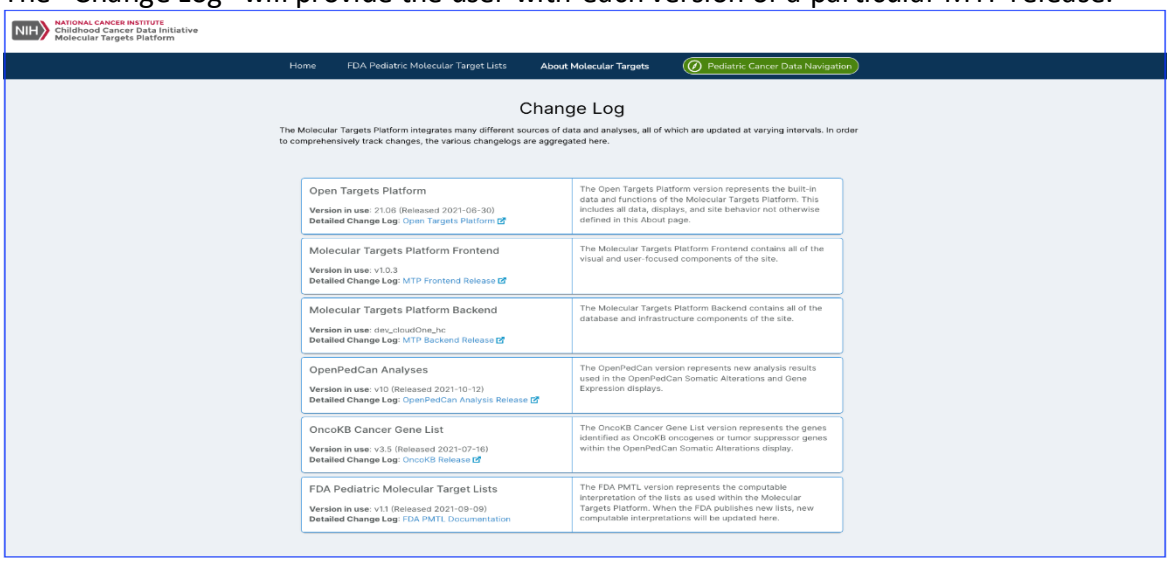

Figure 13: About Page --> Change Log content

FDA Pediatric Molecular Targets Page

The Pediatric Molecular Targets (PMTL) page provides links to the FDA documentation and displays all the Relevant Molecular Targets with corresponding FDA-derived information in sortable columns. The PMTL list is exportable in json, tsv and csv formats.

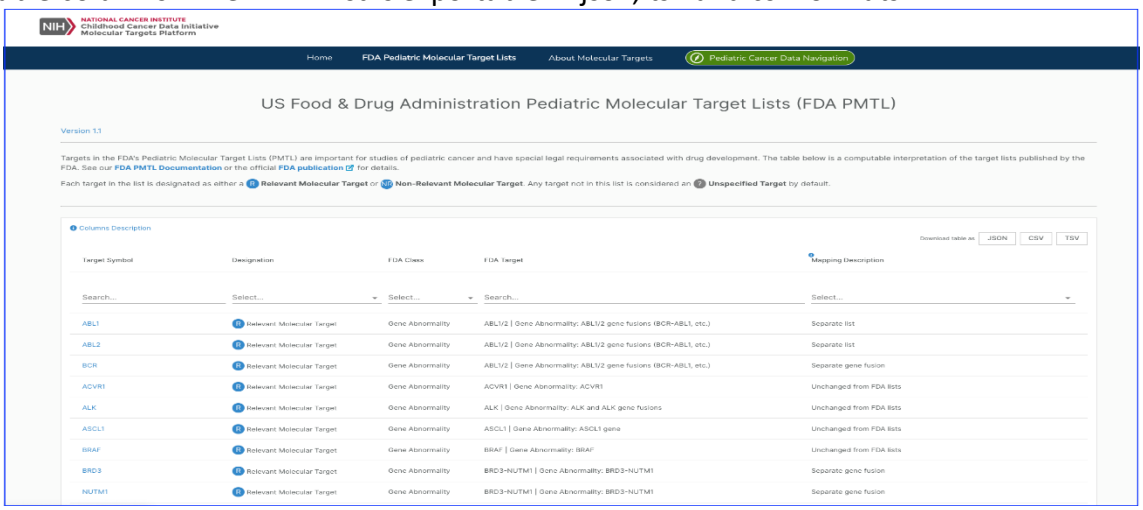

Figure 14: FDA Pediatric Molecular Targets Page

Step-by-Step guidance on how an investigator might use MTP

Question: What is the landscape of FGFR3 mutations in all pediatric cancers?

- Do mutations exist and are there specific histologies in which they are enriched?
- Can I find out if there are any gene amplifications in any pediatric cancers?
- There is an ongoing debate as to whether this receptor is highly over-expressed at the RNA level. Can I query RNA-seq datasets to look at RNA expression?
- Can I find out whether there are any splice-site variants present in order to see whether a protein-domain specific adult cancer drug might be adapted to pediatric cancer use?

Answers from MTP searching:

- **We can view Somatic Mutations that exist and their frequencies and in what specific histologies they are found:** From the main menu, click on the link for "Pediatric Cancer Data Navigation". Searching by Gene symbol "FGFR3" returns a list of 36 diseases with pediatric cancer evidence data, 11 of which have SNV data. If, for example, you click the "Evidence Page" link for osteosarcoma, then click on the OpenPedCan Somatic Alterations green widget button, then choose the tab "SNV By Variant", you will see that there are 2 mutations, both from patients in the TARGET dataset. The tab "SNV by Gene" shows the number of FGFR3 mutations/osteosarcoma total in the dataset, and if available, a link-out is provided for PedcBioPortal data.
- **We can view Gene Amplifications (CNVs):** From the main menu, click on the link for "Pediatric Cancer Data Navigation". Searching by Gene symbol "FGFR3" returns a list of 36 diseases with pediatric cancer evidence data, 13 of which have Copy Number data. Click on the kidney Wilms tumor Evidence Page, then click on the OpenPedCan Somatic Alterations green widget button, then choose the tab "CNV By Gene" which shows the types of CN alterations in this disease along with the frequency in the overall dataset.
- **We can view Histology-specific RNA expression across all pediatric diseases and also compared to public RNA-seq datasets such as GTEx**: From the main menu, click on the link for "Pediatric Cancer Data Navigation". Search by Gene symbol "FGFR3". In this Evidence view, click on the FGFR3 link, then the Profile tab, then the GX (Gene Expression) widget which takes the user to linear and log10 graphs of the RNA expression levels of FGFR3 (Y- axis) across all pediatric cancers (X-axis). Next, going back to the main menu, click on the link for "Pediatric Cancer Data Navigation". Searching again by Gene symbol "FGFR3" takes you to the results page. Selecting any of the diseases from the "Evidence Page" links will take you to Gene Expression. Click on the Gene Expression widget to see linear and log10 views of RNA Expression of one particular Pediatric tumor histology (highlighted in blue and located on the far left) versus GTEx normal adult tissue.
- **We can view Somatic Mutations including Splice-site mutations:** From the main menu, click on the link for "Pediatric Cancer Data Navigation". Searching by Gene

symbol “FGFR3” takes you to a results page. In this view, click on the FGFR3 link, then the Profile tab, then the SM (Somatic Mutations) “SNV by Variant” tab which shows a listing of all specific genomic mutation by Dataset and Disease and frequency observed. Searching by “Splice_Site” (a choice for the field “Variant classifications”), will return 6 FGFR3 splice-site mutations across various diseases in one particular dataset.