User Guide

1 Short Description of the Web Application

The MHN History Tree Web App is a web application designed to visualize the most likely evolutionary path of tumors inferred from patient data by the MHN algorithm.
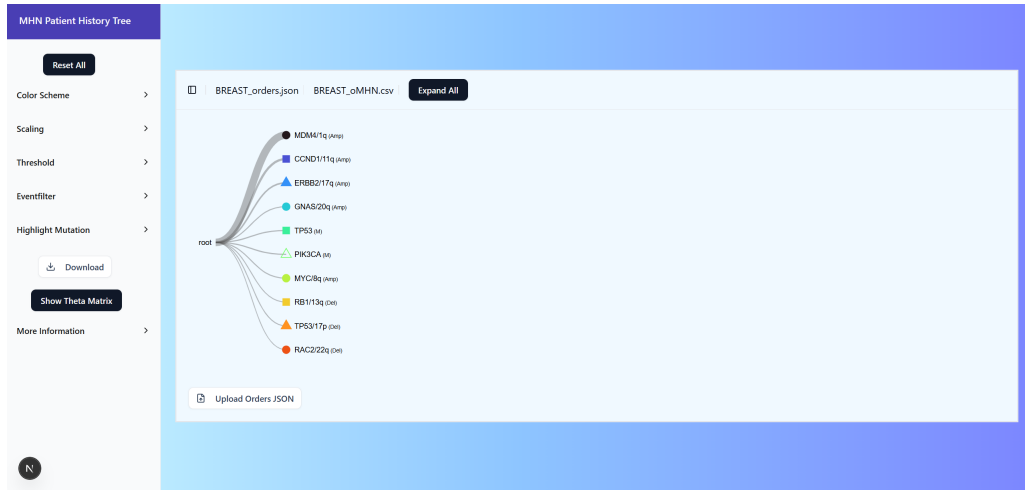

Figure 1: User interface of the MHN History Tree Web App

This tool is primarily intended for scientists and clinicians working with tumor data, regardless of their technical background. With this tool, users can examine how genetic events accumulate over time during the progression of tumors.

The application supports:

- importing data files

- visualization of the Theta Matrix

- interactive visualization of History Trees

- styling options (e.g., colors, edge scaling)

- filtering for parameters (event filtering, threshold)

- downloading the customized History Tree as *SVG* (Scalable Vector Graphics)

2   Requirements

To run the MHN History Tree Web Application locally, the following software components must be installed on the system:

- Node.js (version 18 or later), available at (`https://nodejs.org/en/download`)

- npm (included in Node.js download) - package manager

- Modern web browser (for example, Google Chrome or Mozilla Firefox)

Node.js provides the *runtime environment* necessary for executing the programming language JavaScript outside of the browser, which was used to build this application. *npm*, which comes bundled with Node.js, is a *package manager* responsible for installing, updating, and managing the project's dependencies. These are prebuilt components, such as libraries and modules, that the application requires to work correctly.

In addition to the established runtime environment, these input files can be uploaded for visualization:

- JSON file (orders): This file contains the most likely reconstructed tumor history paths for individual patients. It is required to display the History Tree.

- CSV file (rates): This file contains the multiplicative effects between genetic events, as well as base rates and observation effects. It is used to visualize the Theta Matrix.

3   Installation and Start of the Project

Downloading the Project

Once all system requirements have been met, the web application can be downloaded with the following steps:

1. Visit `https://github.com/Antonia-gthb/historytree`

2. Click on the green button and select "Download ZIP"

3. Unpack the ZIP file (right-click ⇒ "Extract All")

Starting the Project

1. A terminal must be opened in the **unpacked project folder** by right-clicking on the folder and selecting "Open in Terminal".

   As an alternative, a terminal can be opened manually and navigated to the project folder using the `cd` command:

   > Windows: Click on the button or the symbol in the taskbar and search for "Terminal".

   For example: `cd C:\Users\admin\Downloads\historytree-main\historytree`

   > macOS: Open the Launchpad and enter "Terminal" in the search field.

   For example: `cd ~/Downloads/historytree-main/historytree`

2. The app is located in the subordinate directory "historytree", which can be reached by using the `cd` command as shown above

3. The required dependencies are installed by entering and executing:

   ```
   npm install
   ```

   This step is only necessary during the initial setup.

4. The application is compiled with:

   ```
   npm run build
   ```

   This step is only necessary during the initial setup.

5. The application is started in production mode using the command:

   ```
   npm start
   ```

6. Once the server is running, the application can be accessed in a web browser of choice by copying this URL into the address bar:

   ```
   http://localhost:3000
   ```

7. The application can be exited by pressing:
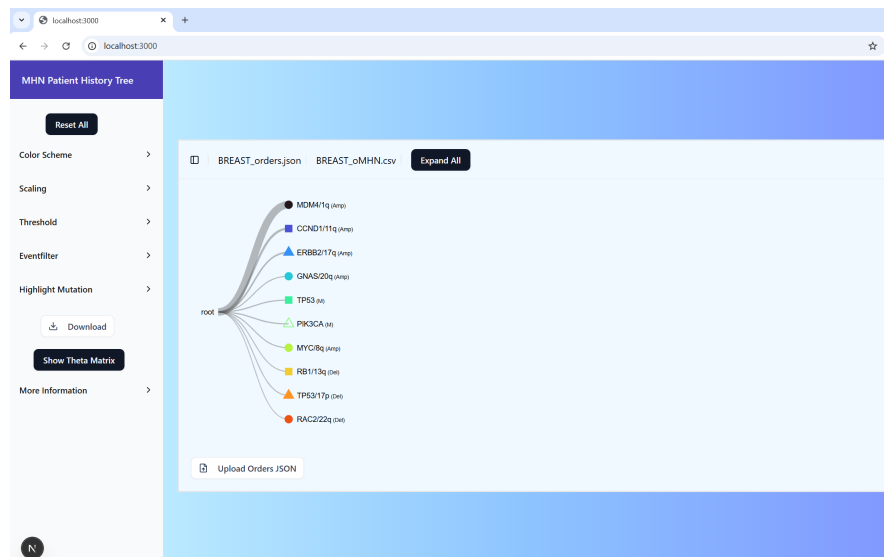
   `strg + c` or `control + c` ⇒ yes

Figure 2: Initial user interface after starting the app locally with `http://localhost:3000` .

4 Example Workflow with Breast Cancer Data

To demonstrate how to use the web application, the breast cancer dataset, which originally stems from the work of Schill et al. [2] on oMHN, is used to present an example workflow in the app.

4.1 Breast Cancer Background

Over the past few decades, biomedical research has made remarkable progress concerning the screening, diagnosis, and treatment of breast cancer patients. Due to improved early detection methods, such as mammography, as well as advanced treatment options, the number of deaths from breast cancer has been decreasing for years [1].

Nevertheless, breast cancer remains a significant health concern. Approximately one in eight women will develop breast cancer during her lifetime [1]. According to the Federal Statistical Office, it is still

the fourth most common cause of death among women [3].

Early detection plays a crucial role in improving prognosis and increasing the chances of recovery. Understanding the sequence of genetic events that occur during tumor evolution can provide valuable insights into early tumorigenesis, potential therapeutic targets, and personalized treatment strategies. The MHN History Tree Web App supports this goal by visualizing reconstructed genetic event sequences, thereby facilitating the identification of relevant patterns in breast cancer research.

4.2 Uploading Input Files

As previously mentioned, a JSON file is needed to visualize the History Tree. Optionally, the corresponding CSV file can also be uploaded to support the interpretation of the History Tree. Both files are generated in a 2-step process by the MHN algorithm and are typically provided by bioinformaticians.

In this example, the breast cancer dataset is used to demonstrate the application's functionality (`BREAST_orders.json` and `BREAST_oMHN.csv`). In a regular workflow, users can upload their files via the respective upload buttons. Upon clicking, a window opens, allowing users to select the desired file from their local system.
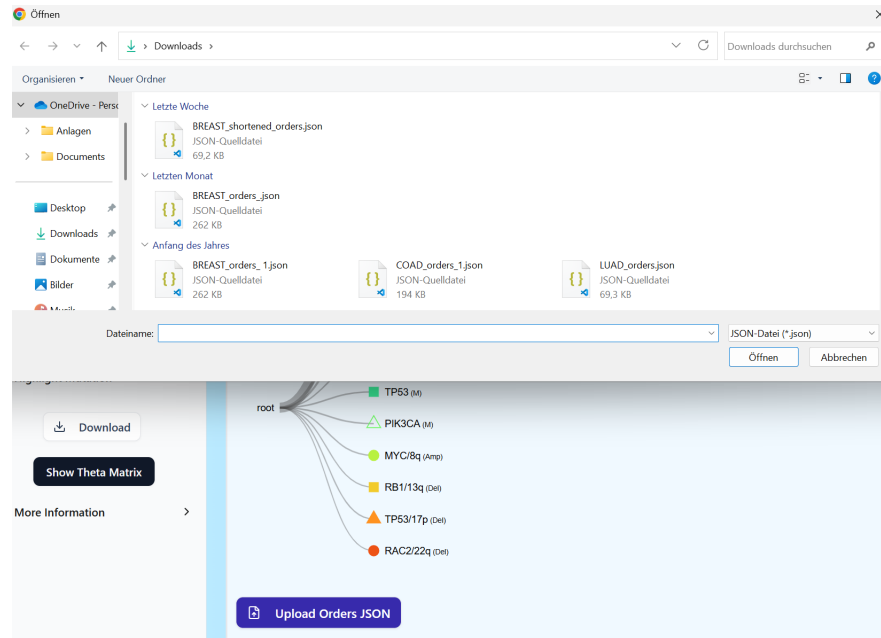

Figure 3: User interface after clicking on JSON upload button and opening a file selection dialog.

Both files can be uploaded independently. Once selected, their file names appear above the History Tree visualization to ensure that users always know which data they are currently viewing (see Figure 5). The upload button for the Theta Matrix is revealed by clicking on the `Show Theta Matrix` button.
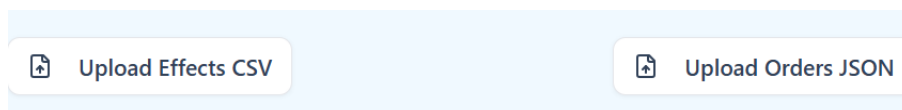

Figure 4: Upload buttons for the JSON and CSV file in the user interface after revealing the Theta Matrix by clicking the `Show Theta Matrix` button.

The following section provides an overview of the user interface and explains its main components.

4.3   User Interface

After successfully starting the application in the browser and uploading a JSON file, the web application automatically renders the corresponding History Tree, presenting an intuitive interface consisting of two main sections: the sidebar on the left, highlighted in yellow, and the visualization panel on the right, highlighted in red.
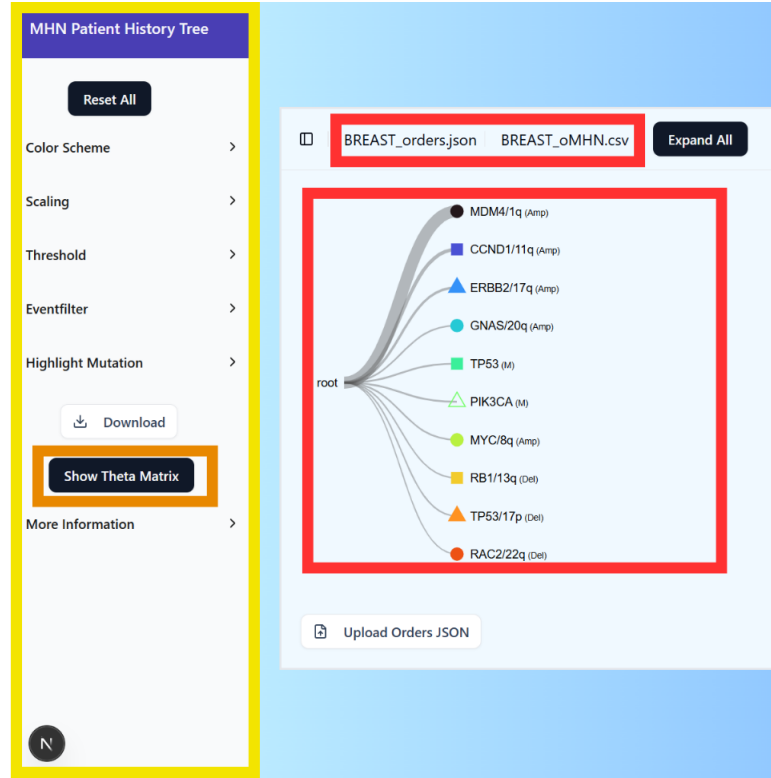

Figure 5: User interface view after uploading the input file(s). The sidebar (yellow) contains several customization features; the main panel (red) displays the History Tree and the names of the uploaded files above. The button highlighted in orange reveals the Theta Matrix.

The sidebar enables users to adjust various parameters, including the color scheme, threshold values, edge scaling, and event filtering. Clicking the `Reset` button at the top of the sidebar sets all features back to their default values. Additionally, the current History Tree view can be downloaded using the `Download` button, and the Theta Matrix (accessible via the orange-highlighted button) can be toggled on or off.

The sidebar also includes a collapsible section labeled "More Information", which provides access to the following resources:

- a link to the subpage `/info`, which contains the user guide

- a link to the PDF version of this master's thesis

- and a link to a *Miro* board (`https://miro.com/`) in the online workspace, providing an overview of the project structure and code, which was developed in the course of this project
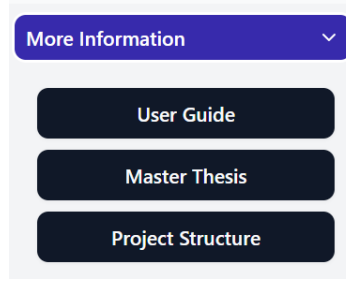
Figure 6: Collapsible section "More Information" in the sidebar, containing links to additional resources such as a user guide or the project structure on Miro.

The visualization panel represents the main part of the user interface, displaying the interactive History Tree. The name of the uploaded file (handled by clicking on the `Upload Orders JSON` button) is shown above. The `Expand All` button fully expands the History Tree (see Figure 8).

4.4 History Tree View
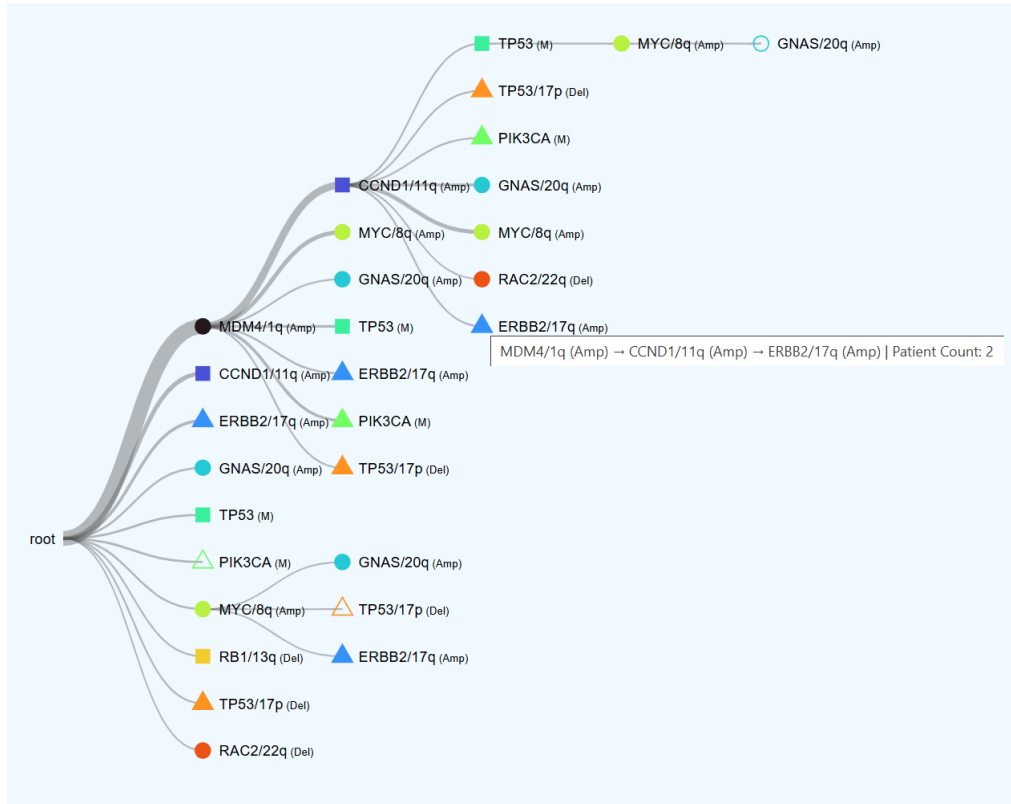
i <u>Important for interpretation</u>

Figure 7: User interface after expanding nodes in the History Tree. Filled symbols indicate parent nodes with children, while unfilled symbols represent terminal nodes with no children. A tooltip with event information, including patient count and sequence up to the selected event, appears on hover.

Each node corresponds to a genetic event, distinguishable by a unique combination of color and symbol. The letter in brackets describes the type of genetic event. "Del" stands for *deletion*, "M" for *point mutation*, and "Amp" for *amplification*. Filled symbols indicate that a node has children (further nodes), while hollow symbols (stroke only) represent terminal nodes with no children. By clicking on a node, the subordinate node(s) is/are expanded (or collapsed), accompanied by an animation. A

tooltip (a small, interactive pop-up element that appears when hovering over a node) with information about the patient count and the genetic event trajectory appears on hover, as pointed out in the Figure above. The link between two genetic events (referred to as *edge*) is determined by the patient count, providing insight into how many patients share identical tumor histories.

4.5 History Tree Features

In addition to the features displayed in the History Tree, users can access further interactive elements, primarily located in the sidebar, which are listed and described in the following subsections.

4.5.1 Expand All

The `Expand All` button expands all nodes within the History Tree, providing a comprehensive overview of the complete structure.
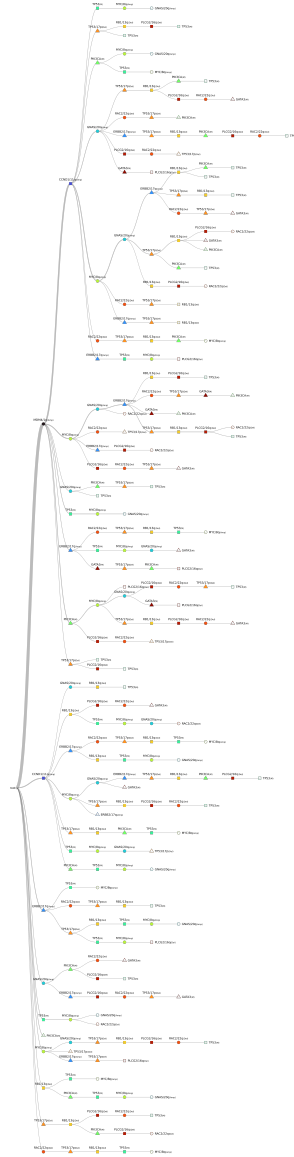

Figure 8: Fully expanded History Tree after clicking the `Expand All` button. The tree displays all nodes and branches in the dataset.

4.5.2 Coloring

In the first collapsible section of the sidebar, a custom color scheme can be selected. Overall, there are six different color schemes to choose from. "Turbo" is set by default.
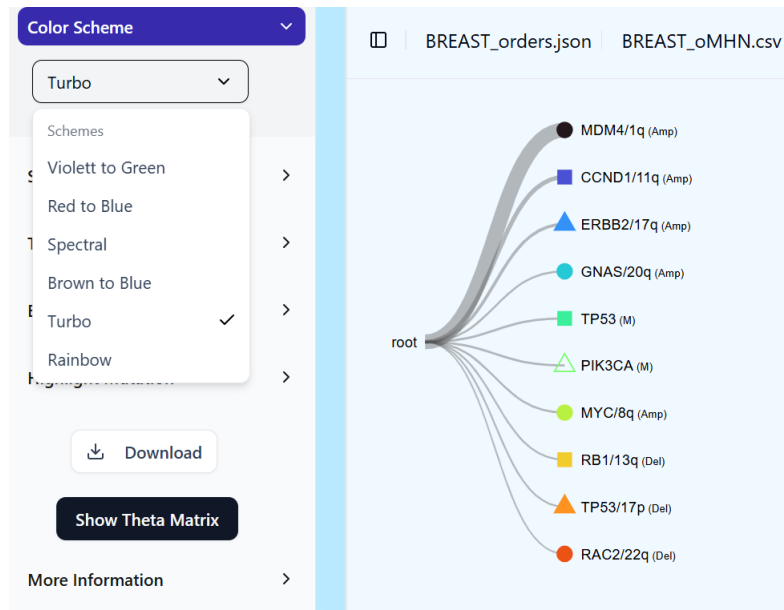

Figure 9: User interface after opening the color scheme selection menu in the sidebar. "Turbo" is selected by default.

4.5.3 Edge Scaling

ℹ Important for interpretation

The edge scaling slider linearly maps the stroke width to the patient count. In other words, the more patients share a specific tumor sequence in their tumor history, the thicker the corresponding edge is rendered. When the slider is deactivated through deselecting the checkbox, all edges are rendered with equal thickness. Reactivating the checkbox sets the slider back to its default value.
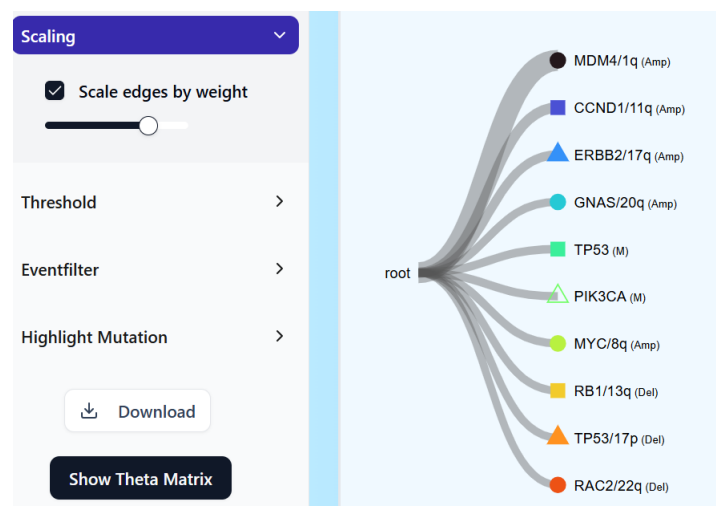

Figure 10: User interface after adjusting the slider for edge scaling. By deselecting the checkbox, all edges are displayed with the same stroke width.

4.5.4 Threshold

i Important for interpretation

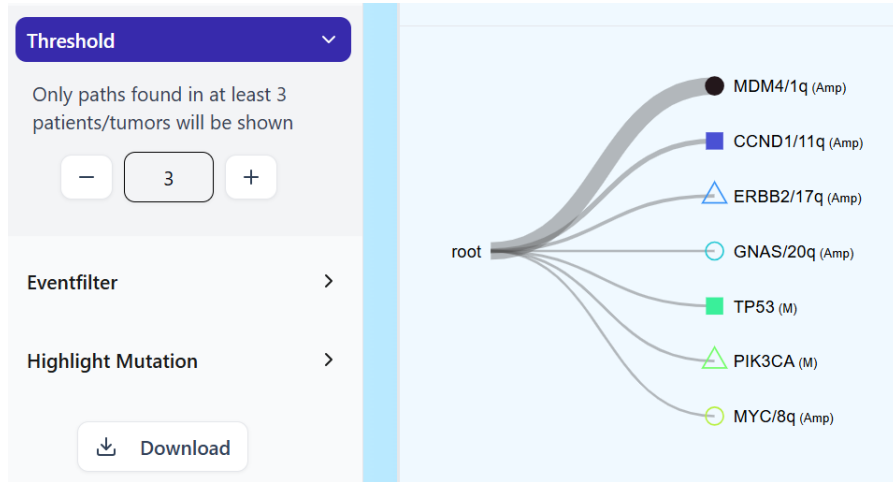The threshold input field allows users to filter data based on a minimum patient count.

Figure 11: User interface after filtering out edges with counts below the input value "3".

For example, setting the threshold to "3" leads to filtering out genetic events with a patient count of "2" or less. Therefore, only paths found in **at least** three patients or tumors are displayed as illustrated in the Figure above.

4.5.5 Event Filtering

i Important for interpretation

By default, all genetic events are selected and displayed. Using the event filter, specific genetic events can be selected or deselected for visualization. The desired changes can be submitted via the provided `Submit` button.

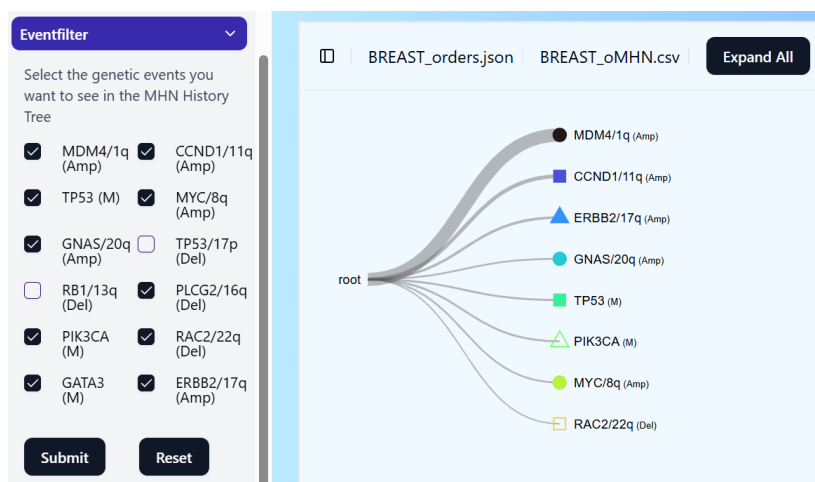Figure 12: User interface after deselecting *TP53/17p* and *RB1/13q*. The History Tree does not display these genetic events.

The `Reset` button restores the default setting (full selection).

4.5.6   Highlight Paths

i Important for interpretation

The "Highlight Paths" feature allows users to select a specific genetic event and highlight all paths in the tree that include it. However, **only the paths currently expanded in the History Tree view are taken into account**. If additional nodes are expanded afterwards, the feature must be executed again to ensure that all relevant edges are highlighted.
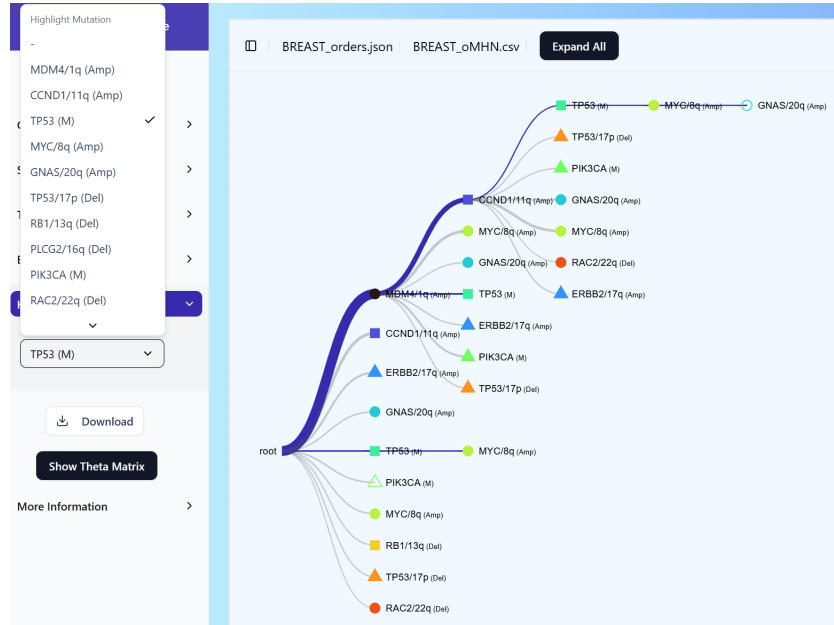

Figure 13: User interface after highlighting paths containing *TP53 (M)* in the current History Tree view.

Additionally, individual paths can be highlighted by clicking directly on an edge within the tree. When a path is selected in this way, all preceding nodes are automatically highlighted as well to provide full context.

4.6 Theta Matrix

ℹ Important for interpretation

As mentioned before, the Theta Matrix is revealed by clicking on the `Show Theta Matrix` button. The matrix shows how genetic events influence each other, comprising both promoting and inhibiting effects (multiplicative effects). The Theta Matrix also includes base rates, which indicate the frequency at which an event occurs independently of other genetic events. The observation column represents the rate at which a tumor is clinically detected, based on the preceding events.
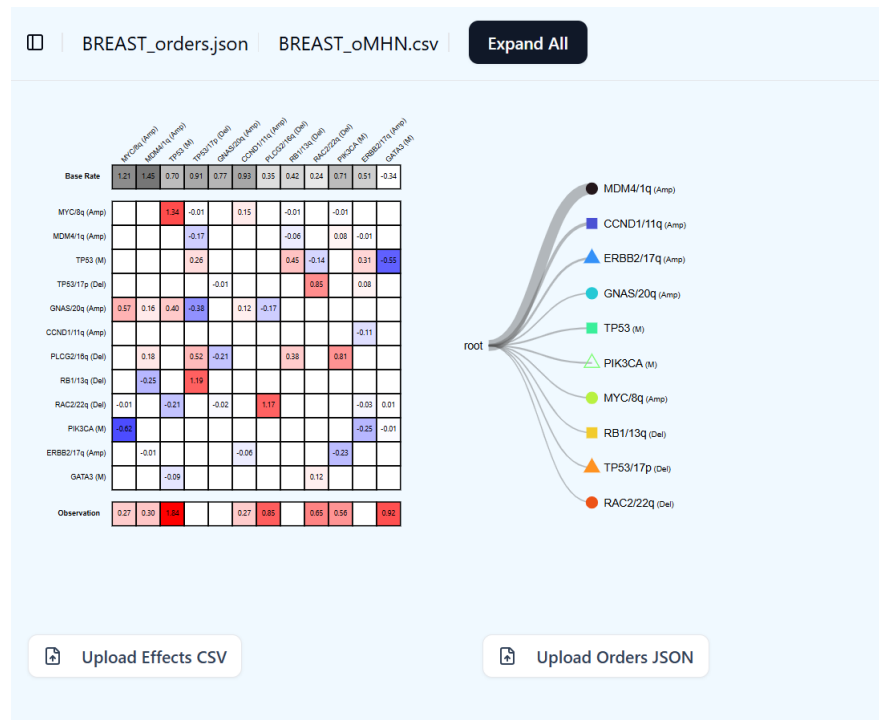

Figure 14: User interface after toggling the Theta Matrix. Rows and columns represent genetic events. Cell colors encode the strength and direction of multiplicative effects. Red for promoting effects, blue for inhibiting effects. The value zero leads to a blank cell. Base rates represent the rate at which an event occurs independently, while the observation column indicates the rate at which a tumor is clinically detected, based on the preceding events.

The Theta Matrix uses a color gradient to visualize the effect values: Blue for negative values below 0 (inhibiting effects), red for positive values above 0 (promoting effects). Blank cells represent the value 0. The intensity of the color reflects the effect size. The top row visualizes the base rates, and the bottom row shows the observation effects. The deviating values (e.g., promoting effects > 0, not > 1 as described in the introduction) can be attributed to the use of the natural logarithm on the values. Further details on the mathematical background can be found in the paper by Schill et al. [2].

ℹ Important for interpretation

The matrix is read from column to row, i.e., the value in a cell represents the influence of the event in the column on the event in the row. The correct interpretation is supported by the tooltip.

Figure 15: Theta Matrix and tooltip showing the multiplicative effect of "RB1/13q" on "MYC/8q", which equals 0.45.

Bibliography

[1] Robert Koch Institute and Association of Population-based Cancer Registries in Germany, eds. *Cancer in Germany 2019/2020*. 14th ed. Berlin: Robert Koch Institute, 2024. URL: `https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/krebs_in_deutschland_node.html`.

[2] Rudolf Schill et al. "Overcoming Observation Bias for Cancer Progression Modeling". In: *bioRxiv* (2023). DOI: `10.1101/2023.12.03.569824`. URL: `https://www.biorxiv.org/content/early/2023/12/05/2023.12.03.569824`.

[3] Statistisches Bundesamt (Destatis). *Leading causes of death among females in Germany*. Accessed July 13, 2025. 2025. URL: `https://www.destatis.de/EN/Themes/Society-Environment/Health/_Graphic/_Interactive/causes-death-leading-female.html`.