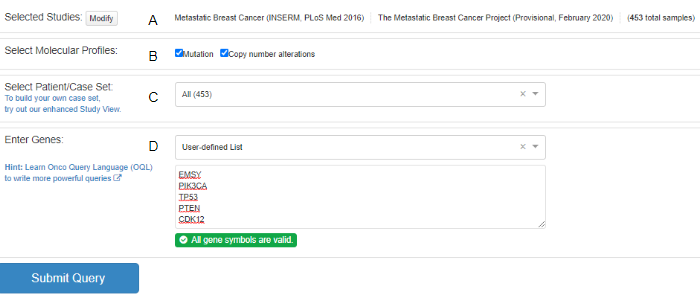
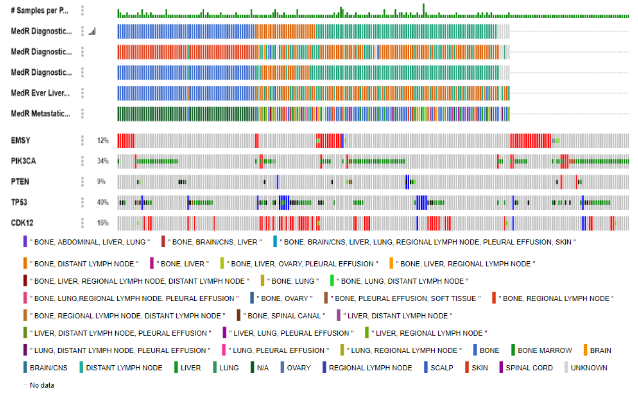
**Appendix 1**



**Appendix 1. Performing a Query**

The example shows the information used to conduct a query on genes shared in both liver and breast cancer and recurrently expressed in MBC studies. (A) select study/studies of interest, (B) select molecular profiles, (C) define patient/case set, (D) enter genes of interest. The "query by gene" pathway of analysis is accessed from the portal's homepage. Two or more genes must be entered to run a query.

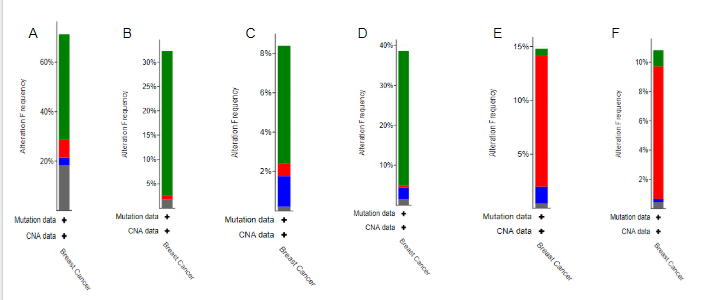
**Appendix 2.**

****

**Appendix 2. An OncoPrint for genetic alterations of EMSY, PIK3CA, PTEN, TP53, and CDK12**

The oncoprint provides a concise summary of genetic alterations expressed by the genes included in the query. Each row is a gene and each column is a sample. The alteration frequency for each gene is represented as a percentage next to each gene. Each sample is color-coded based on alteration type: amplifications (red), deletions (blue), missense mutations (green), truncating mutations (black). Additional clinical tracks included in the query are as follows: (1) MedR estrogen receptor (ER) status, (2) MedR human epidermal growth factor receptor type 2 (HER2) status, (3)MedR progesterone receptor (PR) status, (4) MedR Ever Liver Metastasis, (5) MedR Metastatic Site at Metastatic Diagnosis. Teal samples represent a positive receptor status and liver metastasis; orange is negative; blue is inconclusive. The legend shows the relationship between metastatic site(s) and the color of the sample.

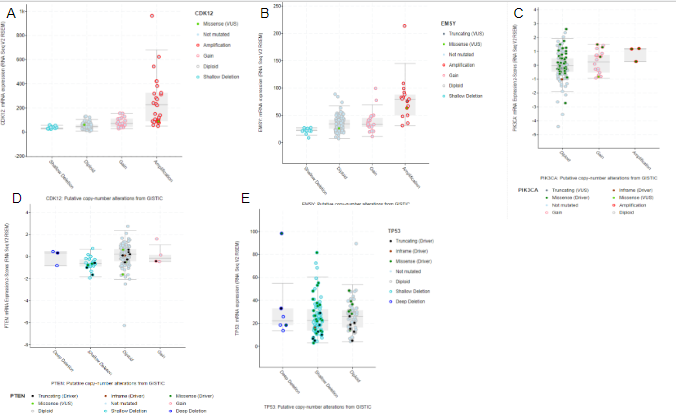
**Appendix 3.**



**Appendix 3. Cancer Types Summary for Queried Genes**

The figures above shows the alteration frequency for all genes included in the query based on cancer type: (a) all queried genes, (b) PIK3CA, (c) PTEN, (d) TP53, (e) CDK12, (f) EMSY. The data is color-coded based on alteration type: mutations (green), amplification (red), deep deletions (blue), multiple alterations (gray). Hover over an alteration to get a breakdown of the frequency of each alteration in table format. The mutation and CNA data are plotted on the x-axis, and the alteration frequency is plotted on the y-axis.

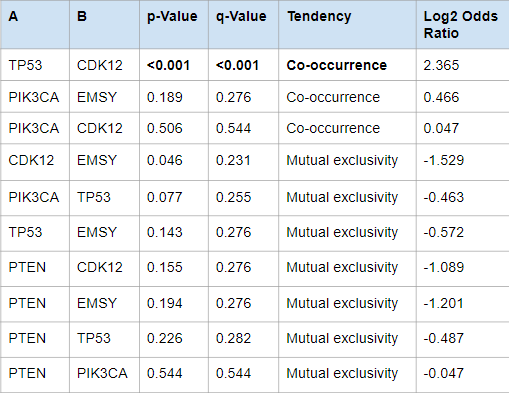
**Appendix 4**



**Appendix 4. Plots for copy-number alterations vs mRNA expression of *EMSY, PIK3CA, PTEN, TP53, CDK12***

Figures A-E show the relationship between copy-number alterations and mRNA expression for each of the queried genes: (a) CDK12, (b) EMSY, (c) PIK3CA, (d) PTEN, (e) TP53. Copy-number alteration type is plotted on the x-axis, and mRNA expression is expressed on the y-axis. Each alteration is color-coded based on type. Amplifications (red), shallow deletions (teal), deep deletions (blue), gain (pink), gray (diploid), missense mutations (green), inframe mutations (brown), truncating mutations (black). Each dot is a sample. Dots filled in light blue lack mutations.

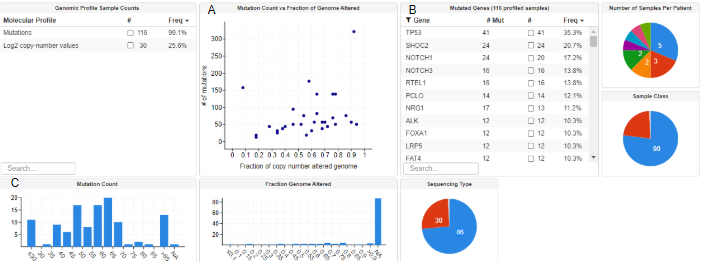
**Appendix 5**



**Appendix 5. Mutual Exclusivity for Queried Genes**

The table shows the analysis of 10 gene pairs for their tendency for mutual exclusivity or co-occurrence. A positive log2 odds ratio value indicates the tendency for co-occurrence, while a negative value indicates the tendency for mutual exclusivity. Significant tendencies are those with a p-value <0.001. Co-occurrence tendencies of gene pairs show alterations of these genes happen in multiple genes in the same cancer sample. Mutual exclusivity tendencies of gene pairs demonstrate each tumor is likely to only have one of those genetic events.

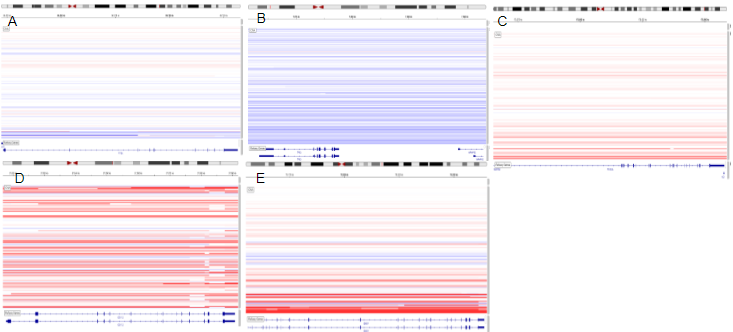
**Appendix 6**

****

**Appendix 6. Explore Select Study View**

The example shows a summary of the data obtained from a 2015 Breast Xenograft Study. (A) A plot showing the relationship between mutation count and the fraction of the genome altered for the samples in the selected cancer study. Each dot is a sample. The x-axis shows the fraction of copy number alterations in the genome, and the y-axis shows the number of mutations corresponding to the fraction of the genome altered. (B) A table of frequently mutated genes of the 116 profiled samples. #Mut, the total number of mutations exhibited by the gene; #, number of samples with one or mutations of the selected gene; Freq, the percentage of samples with one or more mutations of the selected gene. (C) A bar graph representing the mutation count for each of the samples included in the study. The mutation count is plotted on the x-axis, while the number of samples is plotted on the y-axis.

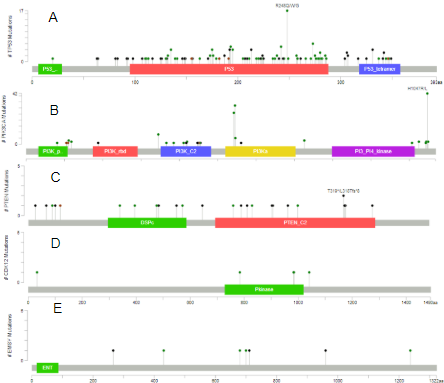
**Appendix 7.**



**Appendix 8. CN Segments for EMSY, PIK3CA, TP53, PTEN, and CDK12**

Figures A-E demonstrate the gain or loss of function of *EMSY* (E), *PIK3CA* (C), *CDK12* (D), *PTEN* (A), and *TP53* (B). Each line is a sample, and each segment represents the chromosomal location of the gene. *EMSY* is located on chromosome 11; *PIK3CA* is located on chromosome 3; *TP53* is located on chromosome 17; *CDK12* is located on chromosome 17; *PTEN* is located on chromosome 10. Deletions (blue) indicate a loss of function; gains (light red) and amplifications (red) indicate a gain of function.

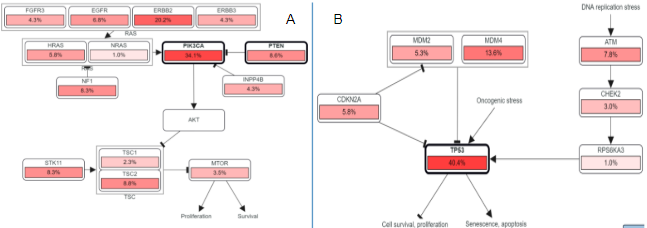
**Appendix 8.**



**Appendix 8. Mutation Charts of EMSY, PTEN, TP53, CDK12, and PIK3CA**

The mutation charts illustrate the amino acid changes resulting from each of the genetic mutations and their frequency of occurrence: (a) TP53, (b) PIK3CA, (c) PTEN, (d) CDK12, (e) EMSY. Each lollipop is a different amino acid. The height of the line from the protein domain to the dot represents the frequency of occurrence for the mutation. The tallest lollipop represents a mutational hotspot for that gene. R248Q/W/G amino acids are a hotspot for TP53 mutations, H1047R/L amino acids for PIK3CA mutations, and T319\*/L318Tfs\*8 amino acids for PTEN mutations. CDK12 and EMSY lack any mutational hotspots.

**Appendix 9**

******

**Appendix 9. Signaling Pathways for the Queried Genes**

Figures 9A and B show pathways involving *TP53*, *PIK3CA*, and *PTEN*. Figure 14A shows the interaction of *PTEN* and *PIK3CA* in the signaling of the PI3K/AKT/mTOR pathway. This pathway is responsible for proliferation and survival of cancer cells. Figure 14B illustrates the signaling pathway of *TP53*; *TP53* is shown to be activated by oncogenic and DNA replication stresses. These stresses cause *TP53* loss of function as a tumor suppressor, and the continued survival and proliferation of cancer cells.