**Onco-Wiki Literature Search**

1. **Data Variability**

In order to simulate a cancer patient, we require reliable sources of gene expression levels of putative driver mutations from original sources such as publications and scientific reviews. Although many research papers will not have used RNA-Sequencing or even compared tumour vs. healthy, many papers will publish tables of upregulated genes / downregulated genes in a variety of tumours with measures of significance in p-value, fold change or log fold change formats.

1. **Literature Search**

Look for original publications in RNA-Sequencing relevant to your cancer and specific cancer subtype. Search generally for your cancer subtype with gene expression levels and also look in the supplementary materials of publications for tables or extended tables of genes that are differentially expressed or statistically significant in their analysis.

1. **Compile Resources**

After finding reliable sources, begin to translate original statistical values into fold changes where a positive fold change indicates upregulation and a negative fold change indicates downregulation so for example, a +10 fold change indicates a 10-fold upregulation. Your new fold changes do not have to be exact but serve as estimations of relevant genes. Look to adjust as many relevant genes that are believed to be mutated in one patient.