**Steps to access the data:**

1. Visit GDC data portal at <https://portal.gdc.cancer.gov/>.

2. Click on "Repository".

3. From the navigation area on the left of the web page, select "Cases".

4. Tick primary site from the “Primary Sites” area

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | THCA | KIRC | PRAD | LUSC |
| Primary Site | thyroid gland | kidney | prostate gland | bronchus and lung |

5. From the "Program" area, tick "TCGA"

6. Tick project from the “Project” area.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | THCA | KIRC | PRAD | LUSC |
| Project | TCGA-THCA | TCGA-KIRC | TCGA-PRAD | TCGA-LUSC |

7. Click “Files” on the top of the navigation area next to the “Cases”.

8. Tick “transcriptome profiling” under “Data Category”.

9. Tick “Gene Expression Quantification” under “Data Type”.

10. Tick “RNA-Seq” under “Experimental Strategy”.

11. Tick “STAR-Counts” under “Workflow type”.

12. Tick “tsv” under “Data Format”.

13. Click “Add All Files to Cart” located at the top right area of the web page.

14. Click “Cart” on the top right corner of the web page.

15. Click “Metadata” to download metadata.

16. Click the drop-down arrow on the “Download” button and select “Cart” to download the files in the cart.

Note: Please download each dataset separately and empty the cart after each download by clicking the red button “Remove From Cart”.