

Readme

- This GDC Publication page hosts extra data files to specifically support the paper of “Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma” published on NEJM 2018
- Special note: all controlled-access raw and summary data are split into three portions of phs001444, phs001184 and phs000178 for separate controls. Please read file #2 for more details.
- The files are described as following:

Index	File Name	Description	Addition to NEJM Paper	Access
1	Readme.pdf	Initial file to read regarding the files on the publication page.	addition	public
2	Introduction_access_control_data.pdf	Brief procedure to access controlled raw sequencing data, MAF, and Copy Number.	addition	public
3	RNAseq_gene_expression_562.txt	Normalized FPKM values in log2 scale at individual level.	addition	public
4	CopyNumber_39_TCGA-DLBC_phs000178.txt	Log ratio of copy number results at individual level derived from Affymetrix Genotyping Console. Need access permissions from dbGaP. Read file #2 for more details how to access the files.	addition	controlled
5	CopyNumber_45_CTSP-DLBCL1_phs001184.txt			
6	CopyNumber_476_NCIICCR-DLBCL_phs001444.txt			
7	MAF_TCGA-DLBC_phs000178.txt	Somatic variants (mutations /del/ins/indel) and associated annotation info. at individual level. Sub-clonal variants are included. Need access permissions via dbGaP. Read file #2 for more details how to access the files	addition	controlled
8	MAF_CTSP-DLBCL1_phs001184.txt			
9	MAF_NCIICCR-DLBCL_phs001444.txt			

10.	Agilent_HaloPlex_Custom_DLBC_370_Genes	Bed files for more detailed genomic regions targeted by HaloPlex sequencing. For the list of 370 targeted genes, see Tab S11 of file #13.	addition	public
11	SignatureDB_021418.xlsx	Database used for gene signature analysis. Also available at https://lymphochip.nih.gov/signaturedb/	addition	public
12	SignatureDB_annotation_021418.xlsx			
13	Supplementary Appendix 2.xlsx	Supplementary files from the paper. The excel contains 13 summary tables. Select individual table by scrolling 13 tabs on the bottom of excel file.	from NEJM paper	public
14	Supplementary Appendix 1.pdf			
15	MAF_Header_Explanation.pdf	Explanation of the headers in MAF file	Addition	public

- Contact
 1. For questions about the procedures or tools on dbGaP and GDC, please contact their help desks of dbGAP and/or GDC.
 2. For specific questions about data format and meanings, please contact the author of data scientist at huangdawei@mail.nih.gov
- Last edit: 04/08/2018