

# R을 활용한 의생명 데이터 분석 - Part 3. R을 활용한 TCGA 암유전체데이터 분석

*Jinho Kim*

*Thursday May 30 2019*

## Contents

<b>1</b>	<b>CDGS R</b>	<b>1</b>
1.1	Install cgdsr package . . . . .	1
1.2	Loading the package . . . . .	1
1.3	Downloading data . . . . .	1
<b>2</b>	<b>Plots</b>	<b>8</b>
<b>3</b>	<b>ggplot2</b>	<b>12</b>
<b>4</b>	<b>Heatmap</b>	<b>16</b>

## 1 CDGS R

### 1.1 Install cgdsr package

```
# install.packages('cgdsr')
```

### 1.2 Loading the package

```
library(knitr)
library(cgdsr)
```

### 1.3 Downloading data

```
## Create CGDS object
mycgds = CGDS("http://www.cbioportal.org/")

## Select a cancer study View(getCancerStudies(mycgds)) #View table of
## studies list
cancer.studies <- getCancerStudies(mycgds)
kable(cancer.studies[, c(1, 2)])
```

cancer_study_id	name
paac_jhu_2014	Acinar Cell Carcinoma of the Pancreas (JHU, J Pathol 2014)
mel_tsam_liang_2017	Acral Melanoma (TGEN, Genome Res 2017)
all_stjude_2015	Acute Lymphoblastic Leukemia (St Jude, Nat Genet 2015)
all_stjude_2016	Acute Lymphoblastic Leukemia (St Jude, Nat Genet 2016)

cancer_study_id	name
laml_tcga_pub	Acute Myeloid Leukemia (TCGA, NEJM 2013)
laml_tcga_pan_can_atlas_2018	Acute Myeloid Leukemia (TCGA, PanCancer Atlas)
laml_tcga	Acute Myeloid Leukemia (TCGA, Provisional)
acyc_fmi_2014	Adenoid Cystic Carcinoma (FMI, Am J Surg Pathl. 2014)
acyc_jhu_2016	Adenoid Cystic Carcinoma (JHU, Cancer Prev Res 2016)
acyc_mda_2015	Adenoid Cystic Carcinoma (MDA, Clin Cancer Res 2015)
acyc_mgh_2016	Adenoid Cystic Carcinoma (MGH, Nat Gen 2016)
acyc_mskcc_2013	Adenoid Cystic Carcinoma (MSKCC, Nat Genet 2013)
acyc_sanger_2013	Adenoid Cystic Carcinoma (Sanger/MDA, JCI 2013)
acbc_mskcc_2015	Adenoid Cystic Carcinoma of the Breast (MSKCC, J Pathol. 2015)
acc_tcga_pan_can_atlas_2018	Adrenocortical Carcinoma (TCGA, PanCancer Atlas)
acc_tcga	Adrenocortical Carcinoma (TCGA, Provisional)
sarc_tcga_pub	Adult Soft Tissue Sarcomas (TCGA, Cell 2017)
ampca_bcm_2016	Ampullary Carcinoma (Baylor College of Medicine, Cell Reports 2016)
odg_msk_2017	Anaplastic Oligodendroglioma and Anaplastic Oligoastrocytoma (MSKCC, Neuro
blca_mskcc_solit_2014	Bladder Cancer (MSKCC, Eur Urol 2014)
blca_mskcc_solit_2012	Bladder Cancer (MSKCC, J Clin Onco 2013)
blca_plasmacytoid_mskcc_2016	Bladder Cancer (MSKCC, Nat Genet 2016)
blca_tcga_pub_2017	Bladder Cancer (TCGA, Cell 2017)
blca_bgi	Bladder Urothelial Carcinoma (BGI, Nat Genet 2013)
blca_dfarber_mskcc_2014	Bladder Urothelial Carcinoma (DFCI/MSKCC, Cancer Discov 2014)
blca_tcga_pub	Bladder Urothelial Carcinoma (TCGA, Nature 2014)
blca_tcga_pan_can_atlas_2018	Bladder Urothelial Carcinoma (TCGA, PanCancer Atlas)
blca_tcga	Bladder Urothelial Carcinoma (TCGA, Provisional)
lgg_tcga_pan_can_atlas_2018	Brain Lower Grade Glioma (TCGA, PanCancer Atlas)
lgg_tcga	Brain Lower Grade Glioma (TCGA, Provisional)
brca_metabric	Breast Cancer (METABRIC, Nature 2012 & Nat Commun 2016)
breast_msk_2018	Breast Cancer (MSK, Cancer Cell 2018)
brca_bccrc_xenograft_2014	Breast Cancer Xenografts (British Columbia, Nature 2015)
bfn_duke_nus_2015	Breast Fibroepithelial Tumors (Duke-NUS, Nat Genet 2015)
brca_bccrc	Breast Invasive Carcinoma (British Columbia, Nature 2012)
brca_broad	Breast Invasive Carcinoma (Broad, Nature 2012)
brca_sanger	Breast Invasive Carcinoma (Sanger, Nature 2012)
brca_tcga_pub2015	Breast Invasive Carcinoma (TCGA, Cell 2015)
brca_tcga_pub	Breast Invasive Carcinoma (TCGA, Nature 2012)
brca_tcga_pan_can_atlas_2018	Breast Invasive Carcinoma (TCGA, PanCancer Atlas)
brca_tcga	Breast Invasive Carcinoma (TCGA, Provisional)
cellline_ccle_broad	Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012)
cesc_tcga_pan_can_atlas_2018	Cervical Squamous Cell Carcinoma (TCGA, PanCancer Atlas)
cesc_tcga	Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (TCGA, Pr
chol_msk_2018	Cholangiocarcinoma (MSK, Clin Cancer Res 2018)
chol_nccs_2013	Cholangiocarcinoma (National Cancer Centre of Singapore, Nat Genet 2013)
chol_nus_2012	Cholangiocarcinoma (National University of Singapore, Nat Genet 2012)
chol_tcga_pan_can_atlas_2018	Cholangiocarcinoma (TCGA, PanCancer Atlas)
chol_tcga	Cholangiocarcinoma (TCGA, Provisional)
lccl_broad_2013	Chronic Lymphocytic Leukemia (Broad, Cell 2013)
cll_iuopa_2015	Chronic Lymphocytic Leukemia (IUOPA, Nature 2015)
cllsl_icgc_2011	Chronic lymphocytic leukemia (ICGC, Nature Genetics 2011)
coadread_dfci_2016	Colorectal Adenocarcinoma (DFCI, Cell Reports 2016)
coadread_genentech	Colorectal Adenocarcinoma (Genentech, Nature 2012)
coadread_tcga_pub	Colorectal Adenocarcinoma (TCGA, Nature 2012)
coadread_tcga_pan_can_atlas_2018	Colorectal Adenocarcinoma (TCGA, PanCancer Atlas)

cancer_study_id	name
coadread_tcga	Colorectal Adenocarcinoma (TCGA, Provisional)
coadread_mskcc	Colorectal Adenocarcinoma Triplets (MSKCC, Genome Biol 2014)
csc_c_dfarber_2015	Cutaneous Squamous Cell Carcinoma (DFCI, Clin Cancer Res 2015)
csc_c_hgsc_bcm_2014	Cutaneous Squamous Cell Carcinoma (MD Anderson, Clin Cancer Res 2014)
ctcl_columbia_2015	Cutaneous T Cell Lymphoma (Columbia U, Nat Genet 2015)
pact_jhu_2011	Cystic Tumor of the Pancreas (Johns Hopkins, PNAS 2011)
desm_broad_2015	Desmoplastic Melanoma (Broad Institute, Nat Genet 2015)
dlbcl_dfc_2018	Diffuse Large B cell Lymphoma (DFCI, Nat Med 2018)
dlbc_broad_2012	Diffuse Large B-Cell Lymphoma (Broad, PNAS 2012)
dlbcl_duke_2017	Diffuse Large B-Cell Lymphoma (Duke, Cell 2017)
dlbc_tcga_pan_can_atlas_2018	Diffuse Large B-Cell Lymphoma (TCGA, PanCancer Atlas)
nhl_bcgsc_2013	Diffuse Large B-cell Lymphoma (BCGSC, Blood 2013)
ucec_msk_2018	Endometrial Cancer (MSK, 2018)
esca_broad	Esophageal Adenocarcinoma (DFCI, Nat Genet 2013)
esca_tcga_pan_can_atlas_2018	Esophageal Adenocarcinoma (TCGA, PanCancer Atlas)
esca_tcga	Esophageal Carcinoma (TCGA, Provisional)
esc_c_icgc	Esophageal Squamous Cell Carcinoma (ICGC, Nature 2014)
esc_c_ucla_2014	Esophageal Squamous Cell Carcinoma (UCLA, Nat Genet 2014)
es_iocurie_2014	Ewing Sarcoma (Institut Curie, Cancer Discov 2014)
gbc_msk_2018	Gallbladder Cancer (MSK, Cancer 2018)
gbc_shanghai_2014	Gallbladder Carcinoma (Shanghai, Nat Genet 2014)
egc_tmucih_2015	Gastric Adenocarcinoma (TMUCIH, PNAS 2015)
gct_msk_2016	Germ Cell Tumors (MSKCC, J Clin Oncol 2016)
gbm_tcga_pub2013	Glioblastoma (TCGA, Cell 2013)
gbm_tcga_pub	Glioblastoma (TCGA, Nature 2008)
gbm_tcga_pan_can_atlas_2018	Glioblastoma Multiforme (TCGA, PanCancer Atlas)
gbm_tcga	Glioblastoma Multiforme (TCGA, Provisional)
glioma_msk_2018	Glioma (MSK, 2018)
hnsk_broad	Head and Neck Squamous Cell Carcinoma (Broad, Science 2011)
hnsk_jhu	Head and Neck Squamous Cell Carcinoma (Johns Hopkins, Science 2011)
hnsk_tcga_pub	Head and Neck Squamous Cell Carcinoma (TCGA, Nature 2015)
hnsk_tcga_pan_can_atlas_2018	Head and Neck Squamous Cell Carcinoma (TCGA, PanCancer Atlas)
hnsk_tcga	Head and Neck Squamous Cell Carcinoma (TCGA, Provisional)
liad_inserm_fr_2014	Hepatocellular Adenoma (INSERM, Cancer Cell 2014)
hcc_mskimpact_2018	Hepatocellular Carcinoma (MSK, Clin Cancer Res 2018)
hcc_inserm_fr_2015	Hepatocellular Carcinomas (INSERM, Nat Genet 2015)
histiocytosis_cobi_msk_2019	Histiocytosis Cobimetinib (MSK, Nature 2019)
all_stjude_2013	Hypodiploid Acute Lymphoid Leukemia (St Jude, Nat Genet 2013)
panet_shanghai_2013	Insulinoma (Shanghai, Nat Commun 2013)
chol_jhu_2013	Intrahepatic Cholangiocarcinoma (JHU, Nat Genet 2013)
kich_tcga_pub	Kidney Chromophobe (TCGA, Cancer Cell 2014)
kich_tcga_pan_can_atlas_2018	Kidney Chromophobe (TCGA, PanCancer Atlas)
kich_tcga	Kidney Chromophobe (TCGA, Provisional)
kirc_bgi	Kidney Renal Clear Cell Carcinoma (BGI, Nat Genet 2012)
ccrcc_irc_2014	Kidney Renal Clear Cell Carcinoma (IRC, Nat Genet 2014)
kirc_tcga_pub	Kidney Renal Clear Cell Carcinoma (TCGA, Nature 2013)
kirc_tcga_pan_can_atlas_2018	Kidney Renal Clear Cell Carcinoma (TCGA, PanCancer Atlas)
kirc_tcga	Kidney Renal Clear Cell Carcinoma (TCGA, Provisional)
kirp_tcga_pan_can_atlas_2018	Kidney Renal Papillary Cell Carcinoma (TCGA, PanCancer Atlas)
kirp_tcga	Kidney Renal Papillary Cell Carcinoma (TCGA, Provisional)
hcc_msk_venturaa_2018	Liver Hepatocellular Adenoma and Carcinomas (MSK, PLOS One 2018)
lihc_amc_prv	Liver Hepatocellular Carcinoma (AMC, Hepatology 2014)

cancer_study_id	name
lihc_riken	Liver Hepatocellular Carcinoma (RIKEN, Nat Genet 2012)
lihc_tcga_pan_can_atlas_2018	Liver Hepatocellular Carcinoma (TCGA, PanCancer Atlas)
lihc_tcga	Liver Hepatocellular Carcinoma (TCGA, Provisional)
lgg_ucsf_2014	Low-Grade Gliomas (UCSF, Science 2014)
luad_broad	Lung Adenocarcinoma (Broad, Cell 2012)
luad_mskcc_2015	Lung Adenocarcinoma (MSKCC, Science 2015)
luad_tcga_pub	Lung Adenocarcinoma (TCGA, Nature 2014)
luad_tcga_pan_can_atlas_2018	Lung Adenocarcinoma (TCGA, PanCancer Atlas)
luad_tcga	Lung Adenocarcinoma (TCGA, Provisional)
luad_tsp	Lung Adenocarcinoma (TSP, Nature 2008)
lusc_tcga_pub	Lung Squamous Cell Carcinoma (TCGA, Nature 2012)
lusc_tcga_pan_can_atlas_2018	Lung Squamous Cell Carcinoma (TCGA, PanCancer Atlas)
lusc_tcga	Lung Squamous Cell Carcinoma (TCGA, Provisional)
dlbc_tcga	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (TCGA, Provisional)
msk_impact_2017	MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)
mixed_allen_2018	MSS Mixed Solid Tumors (Broad/Dana-Farber, Nat Genet 2018)
mpnst_mskcc	Malignant Peripheral Nerve Sheath Tumor (MSKCC, Nat Genet 2014)
mcl_idibips_2013	Mantle Cell Lymphoma (IDIBIPS, PNAS 2013)
mbn_mdacc_2013	Mature B-cell malignancies (MD Anderson Cancer Center)
mb_l_broad_2012	Medulloblastoma (Broad, Nature 2012)
mb_l_icgc	Medulloblastoma (ICGC, Nature 2012)
mb_l_pcg	Medulloblastoma (PCGP, Nature 2012)
mb_l_sickkids_2016	Medulloblastoma (Sickkids, Nature 2016)
skcm_broad_dfarber	Melanoma (Broad/Dana Farber, Nature 2012)
lgggbm_tcga_pub	Merged Cohort of LGG and GBM (TCGA, Cell 2016)
meso_tcga_pan_can_atlas_2018	Mesothelioma (TCGA, PanCancer Atlas)
meso_tcga	Mesothelioma (TCGA, Provisional)
brca_igr_2015	Metastatic Breast Cancer (INSERM, PLoS Med 2016)
crc_msk_2017	Metastatic Colorectal Cancer (MSKCC, Cancer Cell 2018)
egc_msk_2017	Metastatic Esophagogastric Cancer (MSKCC, Cancer Discovery 2017)
skcm_vanderbilt_mskcc_2015	Metastatic Melanoma (MSKCC, JCO Precis Oncol 2017)
skcm_ucla_2016	Metastatic Melanoma (UCLA, Cell 2016)
prad_mich	Metastatic Prostate Adenocarcinoma (MCTP, Nature 2012)
prad_su2c_2019	Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019)
prad_su2c_2015	Metastatic Prostate Cancer (SU2C/PCF Dream Team, Cell 2015)
mm_broad	Multiple Myeloma (Broad, Cancer Cell 2014)
mds_tokyo_2011	Myelodysplasia (UTokyo, Nature 2011)
cellline_nci60	NCI-60 Cell Lines (NCI, Cancer Res 2012)
npc_nusingapore	Nasopharyngeal Carcinoma (Singapore, Nat Genet 2014)
nbl_amc_2012	Neuroblastoma (AMC Amsterdam, Nature 2012)
nbl_broad_2013	Neuroblastoma (Broad, Nat Genet 2013)
nbl_ucologne_2015	Neuroblastoma (Broad, Nature 2015)
nepc_wcm_2016	Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016)
nhl_bcgsc_2011	Non-Hodgkin Lymphoma (BCGSC, Nature 2011)
lung_msk_2017	Non-Small Cell Cancer (MSKCC, Cancer Discov 2017)
nsclc_pdl_msk_2018	Non-Small Cell Lung Cancer (MSKCC, J Clin Oncol 2018)
nsclc_unito_2016	Non-Small Cell Lung Cancer (University of Turin, Lung Cancer 2017)
blca_nmibc_2017	Nonmuscle Invasive Bladder Cancer (MSK Eur Urol 2017)
hnsc_mdanderson_2013	Oral Squamous Cell Carcinoma (MD Anderson, Cancer Discov 2013)
ov_tcga_pub	Ovarian Serous Cystadenocarcinoma (TCGA, Nature 2011)
ov_tcga_pan_can_atlas_2018	Ovarian Serous Cystadenocarcinoma (TCGA, PanCancer Atlas)
ov_tcga	Ovarian Serous Cystadenocarcinoma (TCGA, Provisional)

cancer_study_id	name
nsclc_tcga_broad_2016	Pan-Lung Cancer (TCGA, Nat Genet 2016)
paad_icgc	Pancreatic Adenocarcinoma (ICGC, Nature 2012)
paad_qcmg_uq_2016	Pancreatic Adenocarcinoma (QCMG, Nature 2016)
paad_tcga_pan_can_atlas_2018	Pancreatic Adenocarcinoma (TCGA, PanCancer Atlas)
paad_tcga	Pancreatic Adenocarcinoma (TCGA, Provisional)
paad_utsu_2015	Pancreatic Cancer (UTSW, Nat Commun 2015)
panet_jhu_2011	Pancreatic Neuroendocrine Tumors (Johns Hopkins University, Science 2011)
panet_arcnet_2017	Pancreatic Neuroendocrine Tumors (Multi-Institute, Nature 2017)
thca_tcga_pub	Papillary Thyroid Carcinoma (TCGA, Cell 2014)
all_phase2_target_2018_pub	Pediatric Acute Lymphoid Leukemia - Phase II (TARGET, 2018)
aml_target_2018_pub	Pediatric Acute Myeloid Leukemia (TARGET, 2018)
es_dfarber_broad_2014	Pediatric Ewing Sarcoma (DFCI, Cancer Discov 2014)
nbl_target_2018_pub	Pediatric Neuroblastoma (TARGET, 2018)
pediatric_dkfz_2017	Pediatric Pan-Cancer (DKFZ, Nature 2017)
mixed_pipseq_2017	Pediatric Pan-cancer (Columbia U, Genome Med 2016)
rt_target_2018_pub	Pediatric Rhabdoid Tumor (TARGET, 2018)
wt_target_2018_pub	Pediatric Wilms' Tumor (TARGET, 2018)
pcpg_tcga_pub	Pheochromocytoma and Paraganglioma (TCGA, Cell 2017)
pcpg_tcga_pan_can_atlas_2018	Pheochromocytoma and Paraganglioma (TCGA, PanCancer Atlas)
pcpg_tcga	Pheochromocytoma and Paraganglioma (TCGA, Provisional)
past_dkfz_heidelberg_2013	Pilocytic Astrocytoma (ICGC, Nature Genetics 2013)
plmeso_nyu_2015	Pleural Mesothelioma (NYU, Cancer Res 2015)
thyroid_mskcc_2016	Poorly-Differentiated and Anaplastic Thyroid Cancers (MSKCC, JCI 2016)
pcnsl_mayo_2015	Primary Central Nervous System Lymphoma (Mayo Clinic, Clin Cancer Res 2015)
prad_broad_2013	Prostate Adenocarcinoma (Broad/Cornell, Cell 2013)
prad_broad	Prostate Adenocarcinoma (Broad/Cornell, Nat Genet 2012)
prad_cpcg_2017	Prostate Adenocarcinoma (CPC-GENE, Nature 2017)
prad_fhrc	Prostate Adenocarcinoma (Fred Hutchinson CRC, Nat Med 2016)
prad_mskcc	Prostate Adenocarcinoma (MSKCC, Cancer Cell 2010)
prad_mskcc_2014	Prostate Adenocarcinoma (MSKCC, PNAS 2014)
prad_p1000	Prostate Adenocarcinoma (MSKCC/DFCI, Nature Genetics 2018)
prad_eurourol_2017	Prostate Adenocarcinoma (SMMU, Eur Urol 2017)
prad_tcga_pub	Prostate Adenocarcinoma (TCGA, Cell 2015)
prad_tcga_pan_can_atlas_2018	Prostate Adenocarcinoma (TCGA, PanCancer Atlas)
prad_tcga	Prostate Adenocarcinoma (TCGA, Provisional)
prad_mskcc_cheny1_organoids_2014	Prostate Adenocarcinoma Organoids (MSKCC, Cell 2014)
prad_mskcc_2017	Prostate Cancer (MSKCC, JCO Precis Oncol 2017)
hnc_mskcc_2016	Recurrent and Metastatic Head & Neck Cancer (MSKCC, JAMA Oncol 2016)
ccrcc_utokyo_2013	Renal Clear Cell Carcinoma (UTokyo, Nat Genet 2013)
nccrcc_genentech_2014	Renal Non-Clear Cell Carcinoma (Genentech, Nat Genet 2014)
mrt_bcgsc_2016	Rhabdoid Cancer (BCGSC, Cancer Cell 2016)
rms_nih_2014	Rhabdomyosarcoma (NIH, Cancer Discov 2014)
summit_2018	SUMMIT - Neratinib Basket Study (Multi-Institute, Nature 2018)
sarc_mskcc	Sarcoma (MSKCC/Broad, Nat Genet 2010)
sarc_tcga_pan_can_atlas_2018	Sarcoma (TCGA, PanCancer Atlas)
sarc_tcga	Sarcoma (TCGA, Provisional)
skcm_broad	Skin Cutaneous Melanoma (Broad, Cell 2012)
skcm_tcga_pan_can_atlas_2018	Skin Cutaneous Melanoma (TCGA, PanCancer Atlas)
skcm_tcga	Skin Cutaneous Melanoma (TCGA, Provisional)
skcm_yale	Skin Cutaneous Melanoma (Yale, Nat Genet 2012)
skcm_broad_brafresist_2012	Skin Cutaneous Melanoma (Broad, Cancer Discov 2014)
scco_mskcc	Small Cell Carcinoma of the Ovary (MSKCC, Nat Genet 2014)

cancer_study_id	name
sclc_clcgp	Small Cell Lung Cancer (CLCGP, Nat Genet 2012)
sclc_jhu	Small Cell Lung Cancer (Johns Hopkins, Nat Genet 2012)
sclc_ucologne_2015	Small Cell Lung Cancer (U Cologne, Nature 2015)
sclc_cancercell_gardner_2017	Small-Cell Lung Cancer (Multi-Institute, Cancer Cell 2017)
vsc_cuk_2018	Squamous Cell Carcinoma of the Vulva (CUK, Exp Mol Med 2018)
stad_pfizer_uhongkong	Stomach Adenocarcinoma (Pfizer and UHK, Nat Genet 2014)
stad_tcga_pub	Stomach Adenocarcinoma (TCGA, Nature 2014)
stad_tcga_pan_can_atlas_2018	Stomach Adenocarcinoma (TCGA, PanCancer Atlas)
stad_tcga	Stomach Adenocarcinoma (TCGA, Provisional)
stad_utokyo	Stomach Adenocarcinoma (U Tokyo, Nat Genet 2014)
stad_uhongkong	Stomach Adenocarcinoma (UHK, Nat Genet 2011)
stes_tcga_pub	TCGA data for Esophagus-Stomach Cancers (TCGA, Nature 2017)
tmb_mskcc_2018	TMB and Immunotherapy (MSKCC, Nat Genet 2019)
tgct_tcga	Testicular Germ Cell Cancer (TCGA, Provisional)
tgct_tcga_pan_can_atlas_2018	Testicular Germ Cell Tumors (TCGA, PanCancer Atlas)
angs_projectPainter_2018	The Angiosarcoma Project - Count Me In (Provisional, September 2018)
brca_mbcproject_wagle_2017	The Metastatic Breast Cancer Project (Provisional, October 2018)
prad_mpcproject_2018	The Metastatic Prostate Cancer Project (Provisional, December 2018)
lung_msk_pdx	Thoracic PDX (MSK, Provisional)
tet_nci_2014	Thymic Epithelial Tumors (NCI, Nat Genet 2014)
thym_tcga_pan_can_atlas_2018	Thymoma (TCGA, PanCancer Atlas)
thym_tcga	Thymoma (TCGA, Provisional)
thca_tcga_pan_can_atlas_2018	Thyroid Carcinoma (TCGA, PanCancer Atlas)
thca_tcga	Thyroid Carcinoma (TCGA, Provisional)
urcc_mskcc_2016	Unclassified Renal Cell Carcinoma (MSK, Nature 2016)
utuc_mskcc_2015	Upper Tract Urothelial Cancer (MSK, Eur Urol 2015)
utuc_cornell_baylor_mdacc_2019	Upper Tract Urothelial Carcinoma (Cornell/Baylor/MDACC, Nat Comm 2019)
blca_cornell_2016	Urothelial Carcinoma (Cornell/Trento, Nat Gen 2016)
ucs_jhu_2014	Uterine Carcinosarcoma (Johns Hopkins, Nat Commun 2014)
ucs_tcga_pan_can_atlas_2018	Uterine Carcinosarcoma (TCGA, PanCancer Atlas)
ucs_tcga	Uterine Carcinosarcoma (TCGA, Provisional)
uccc_nih_2017	Uterine Clear Cell Carcinoma (NIH, Cancer 2017)
ucec_tcga_pub	Uterine Corpus Endometrial Carcinoma (TCGA, Nature 2013)
ucec_tcga_pan_can_atlas_2018	Uterine Corpus Endometrial Carcinoma (TCGA, PanCancer Atlas)
ucec_tcga	Uterine Corpus Endometrial Carcinoma (TCGA, Provisional)
um_qimr_2016	Uveal Melanoma (QIMR, Oncotarget 2016)
uvm_tcga_pan_can_atlas_2018	Uveal Melanoma (TCGA, PanCancer Atlas)
uvm_tcga	Uveal Melanoma (TCGA, Provisional)

```
mycancerstudy = cancer.studies[38, 1] #Breast Invasive Carcinoma (TCGA, Cell 2015)
print(mycancerstudy)
```

```
## [1] "brca_tcga_pub2015"
## Get available case lists for a given cancer study
## View(getCaseLists(mycgds,mycancerstudy))
case.list <- getCaseLists(mycgds, mycancerstudy)
kable(case.list[, 1:2])
```

case_list_id	case_list_name
brca_tcga_pub2015_all	All samples
brca_tcga_pub2015_erpos	ER+ breast tumors

case_list_id	case_list_name
brca_tcga_pub2015_erneg	ER- breast tumors
brca_tcga_pub2015_her2pos	Her2-positive breast tumors
brca_tcga_pub2015_idc_luma	Invasive Ductal Cancer (Luminal A)
brca_tcga_pub2015_idc_basal	Invasive Ductal Cancer (PAM50 Basal-like)
brca_tcga_pub2015_idc_her2	Invasive Ductal Cancer (PAM50 Her2-enriched)
brca_tcga_pub2015_idc_lumB	Invasive Ductal Cancer (PAM50 Luminal B)
brca_tcga_pub2015_ilc	Invasive Lobular Cancer
brca_tcga_pub2015_ilc_lumA	Invasive Lobular Cancer (Luminal A)
brca_tcga_pub2015_mixed	Mixed IDC/ILC breast cancer
brca_tcga_pub2015_other	Other histologies breast cancer
brca_tcga_pub2015_cna	Samples with CNA data
brca_tcga_pub2015_methylation_all	Samples with methylation data
brca_tcga_pub2015_methylation_hm27	Samples with methylation data (HM27)
brca_tcga_pub2015_methylation_hm450	Samples with methylation data (HM450)
brca_tcga_pub2015_mrna	Samples with mRNA data (Agilent microarray)
brca_tcga_pub2015_rna_seq_v2_mrna	Samples with mRNA data (RNA Seq V2)
brca_tcga_pub2015_cnaseq	Samples with mutation and CNA data
brca_tcga_pub2015_sequenced	Samples with mutation data
brca_tcga_pub2015_rppa	Samples with protein data (RPPA)
brca_tcga_pub2015_freeze	TCGA Freeze 2015
brca_tcga_pub2015_trineg	Triple-negative breast tumors

```
mycaselist = case.list[1, 1] #All Complete Tumors
print(mycaselist)
```

```
## [1] "brca_tcga_pub2015_all"
```

```
## Get available genetic profiles
genetic.profile <- getGeneticProfiles(mycgds, mycancerstudy)
kable(genetic.profile[, 1:2])
```

genetic_profile_id	genetic_profile_name
brca_tcga_pub2015_protein_quantification	Protein expression (RPPA)
brca_tcga_pub2015_gistic	Putative copy-number alterations from GISTIC
brca_tcga_pub2015_mrna	mRNA expression (microarray)
brca_tcga_pub2015_rna_seq_v2_mrna	mRNA expression (RNA Seq V2 RSEM)
brca_tcga_pub2015_linear_CNA	Relative linear copy-number valuesdata_filename: data_linear_CNA.txt
brca_tcga_pub2015_methylation_hm27	Methylation (HM27)
brca_tcga_pub2015_methylation_hm450	Methylation (HM450)
brca_tcga_pub2015_mutations	Mutations

```
cna <- genetic.profile[5, 1] #linear_CNA
print(cna)
```

```
## [1] "brca_tcga_pub2015_linear_CNA"
```

```
mrna <- genetic.profile[4, 1] #rna_seq_v2_mrna
print(mrna)
```

```
## [1] "brca_tcga_pub2015_rna_seq_v2_mrna"
```

## 2 Plots

```
## Get data slices for a specified list of genes, genetic profile and case
## list
cnadata = getProfileData(mycgds, c("ERBB2"), cna, mycaselist) #CNA data of ERBB2
kable(head(cnadata))
```

	ERBB2
TCGA.LQ.A4E4.01	-0.787
TCGA.A2.A3KC.01	-0.007
TCGA.A2.A3KD.01	-0.016
TCGA.A7.A0D9.01	-0.741
TCGA.A7.A0DA.01	-0.292
TCGA.A7.A0CD.01	1.437

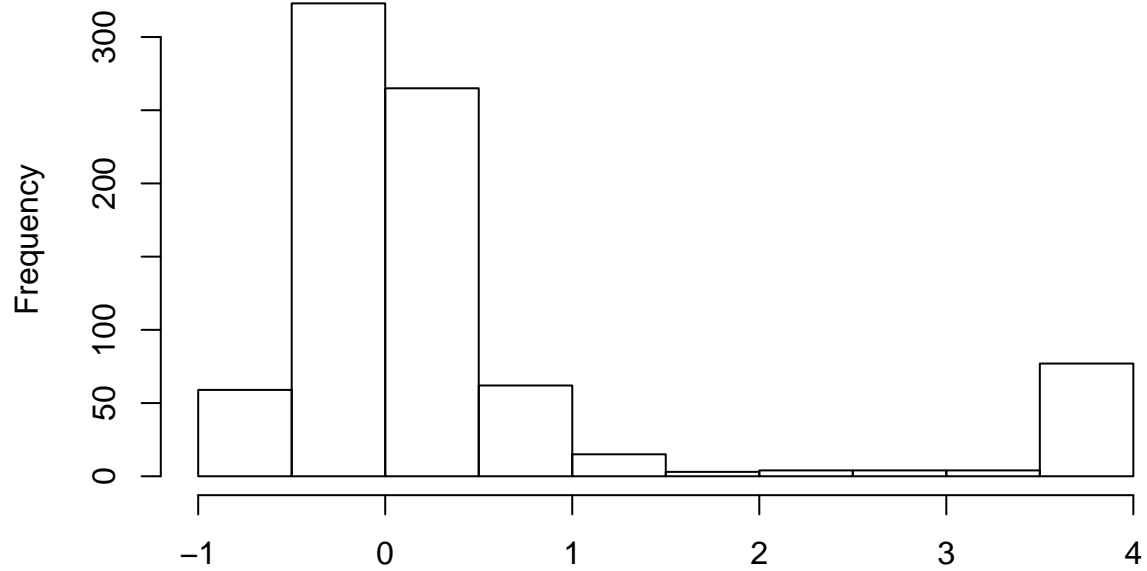
```
mrnadata = getProfileData(mycgds, c("ERBB2"), mrna, mycaselist) #mRNA data of ERBB2
kable(head(mrnadata))
```

	ERBB2
TCGA.LQ.A4E4.01	6846.946
TCGA.A2.A3KC.01	14814.131
TCGA.A2.A3KD.01	8941.431
TCGA.A7.A0D9.01	5291.478
TCGA.A7.A0DA.01	5035.810
TCGA.A7.A0CD.01	15139.034

```
## Histogram
hist(cnadata$ERBB2, main = "ERBB2", xlab = "CNA") # plot generation
```



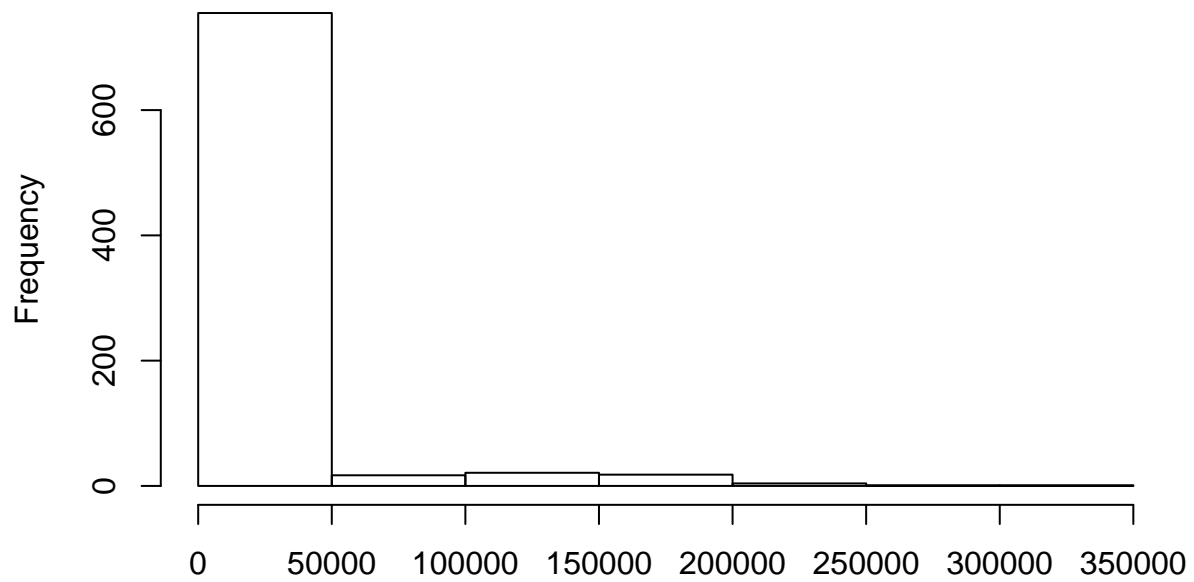
## ERBB2



CNA

```
hist(mrnadata$ERBB2, main = "ERBB2", xlab = "mRNA")
```

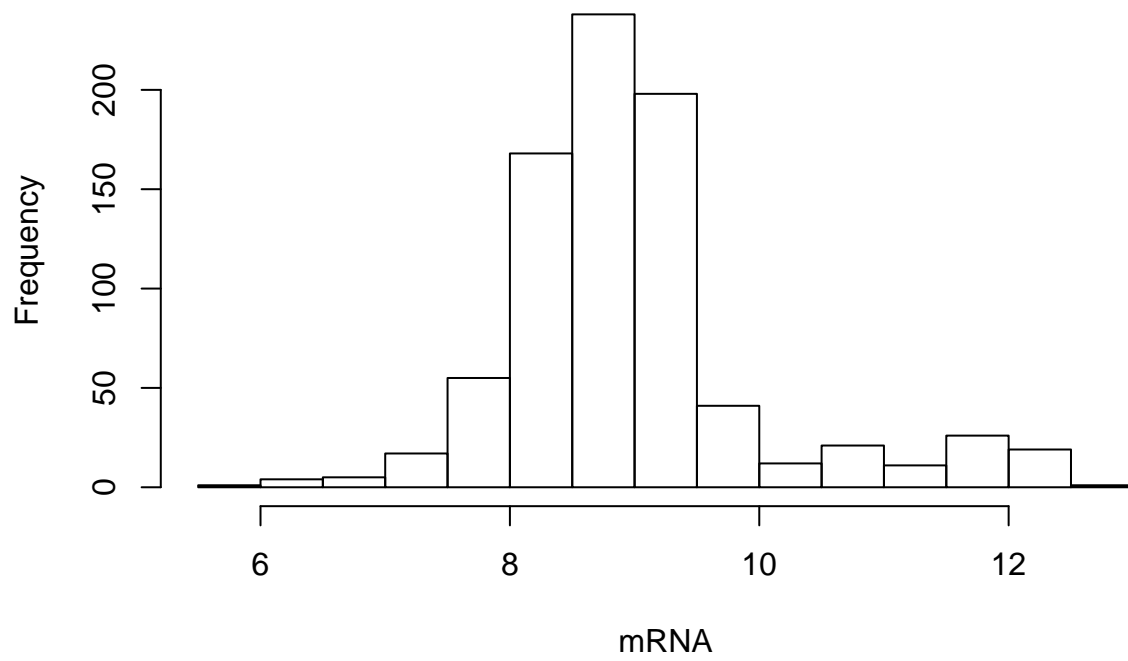
## ERBB2



mRNA

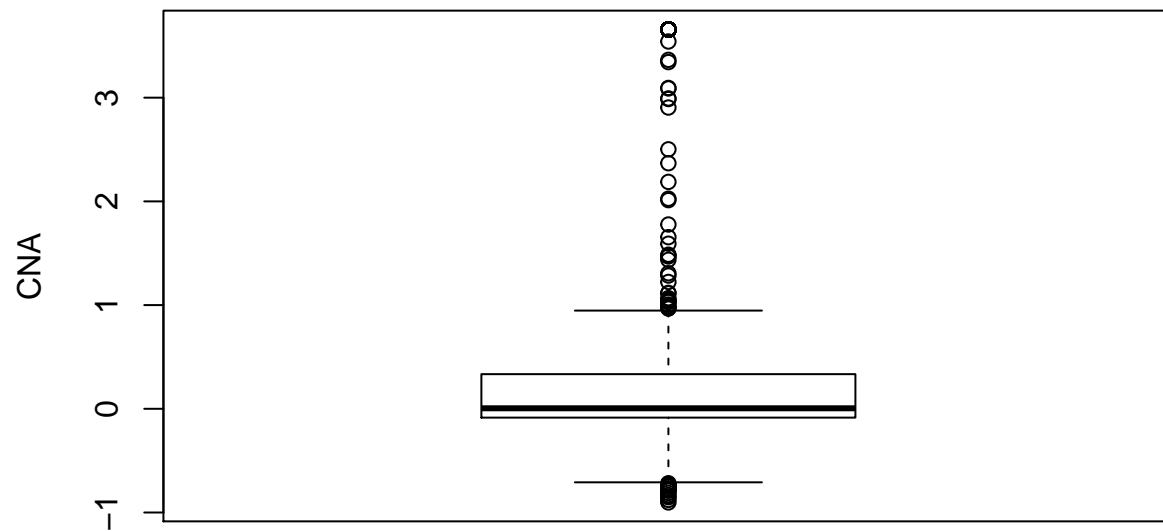
```
hist(log(mrnadata$ERBB2), main = "ERBB2", xlab = "mRNA") # log transformation
```

## ERBB2



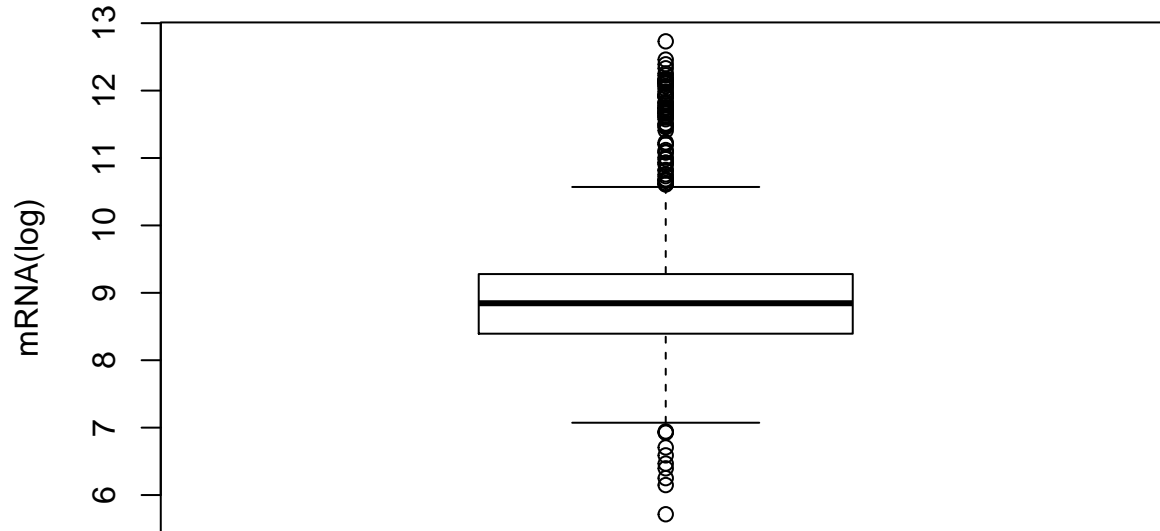
```
## box plot
boxplot(cnadata$ERBB2, main = "ERBB2", ylab = "CNA")
```

## ERBB2



```
boxplot(log(mrnadata$ERBB2), main = "ERBB2", ylab = "mRNA(log)")
```

## ERBB2



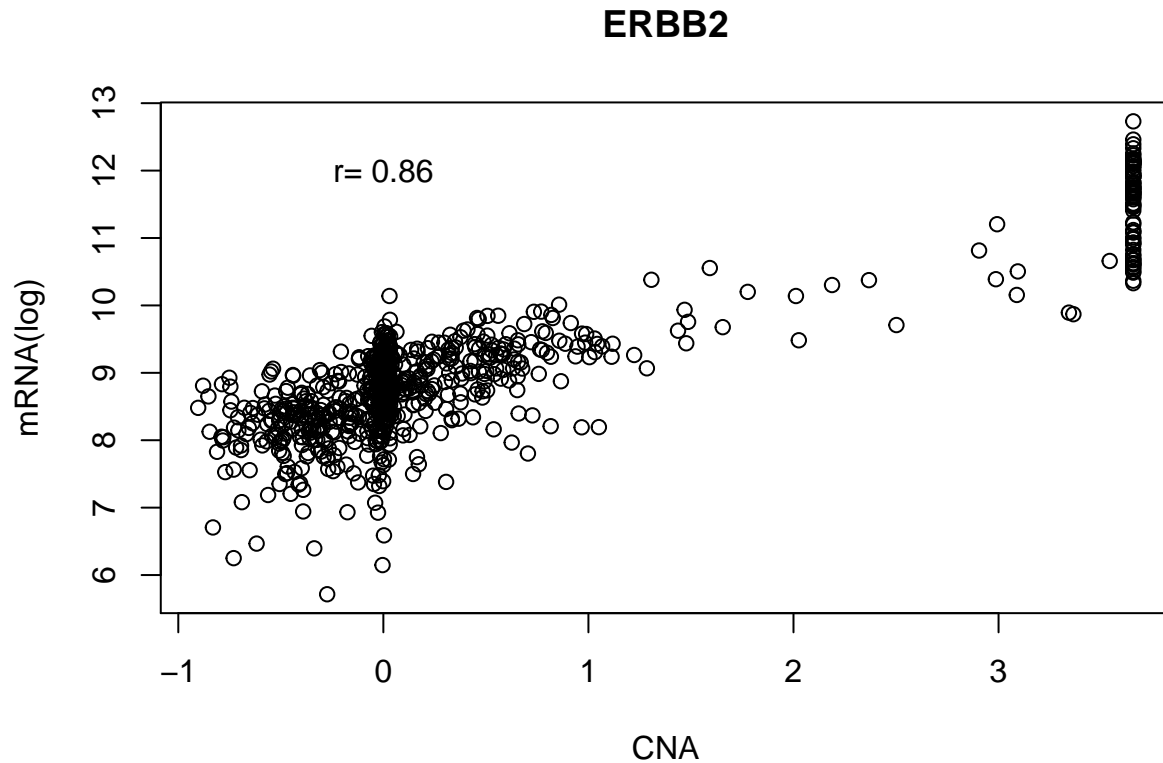
```
## Scatter plot
co = cor(cnadata$ERBB2, log(mrnadata$ERBB2), use = "na.or.complete") #pearson correlation
cor(cnadata$ERBB2, log(mrnadata$ERBB2), method = "pearson", use = "na.or.complete")

## [1] 0.8606259

cor(cnadata$ERBB2, log(mrnadata$ERBB2), method = "spearman", use = "na.or.complete")

## [1] 0.6995344

plot(cnadata$ERBB2, log(mrnadata$ERBB2), main = "ERBB2", xlab = "CNA", ylab = "mRNA(log)")
# Add text on the plot
text(0, 12, sprintf("r= %.2f", co)) # add text
```



```
pdf("TCGA_BRCA_ERBB2_correlation.pdf") #open PDF
plot(cnadata$ERBB2, log(mrnadata$ERBB2), main = "ERBB2", xlab = "CNA", ylab = "mRNA(log)")
text(0, 12, "r=0.85") # add text
dev.off() #close PDF
```

### 3 ggplot2

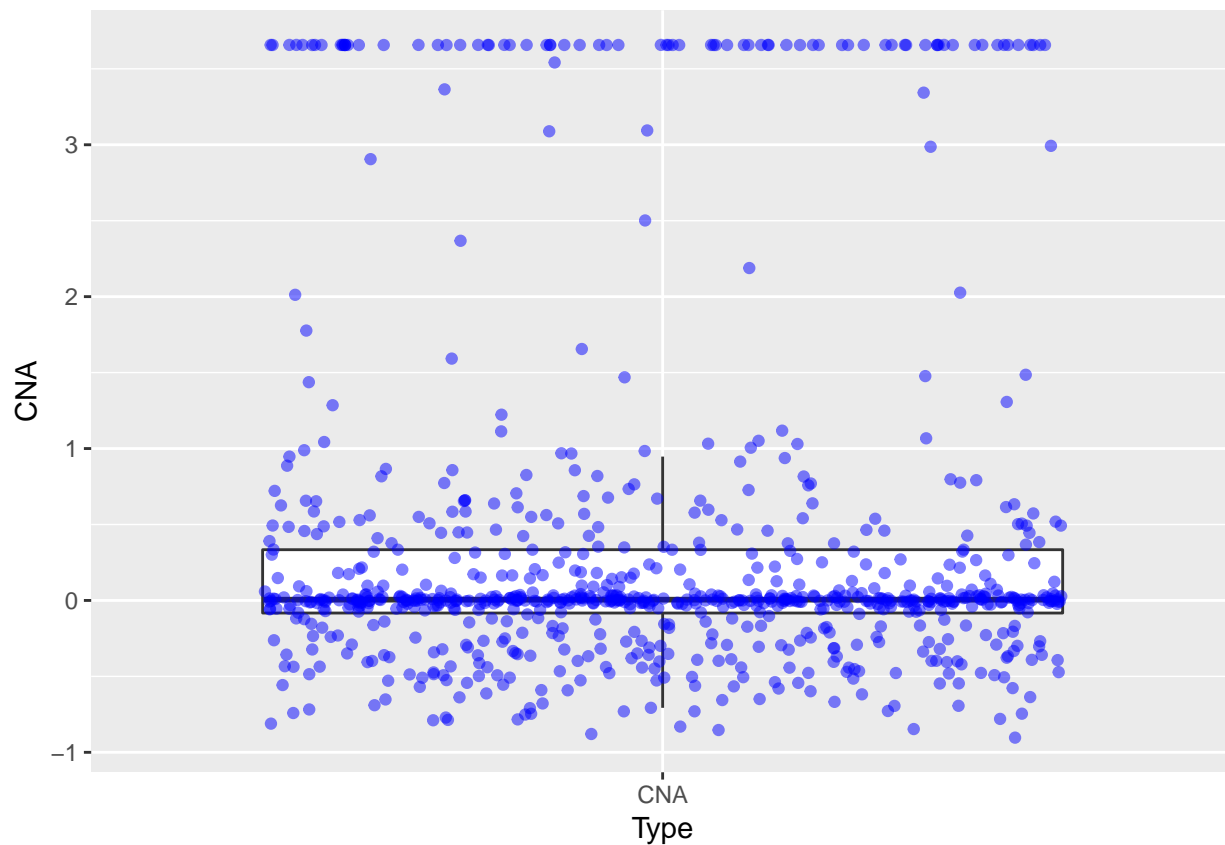
```
# install.packages('ggplot2') ## library install
```

```
library(ggplot2) ## Loading library
```

```
input.cna = data.frame(cnadata, "CNA", stringsAsFactors = F)
colnames(input.cna) = c("CNA", "Type")
```

```
plot1 = ggplot(input.cna, aes(x = Type, y = CNA)) + geom_boxplot(width = 0.7,
  outlier.size = NA) + geom_jitter(width = 0.35, colour = "blue", alpha = 0.5)
```

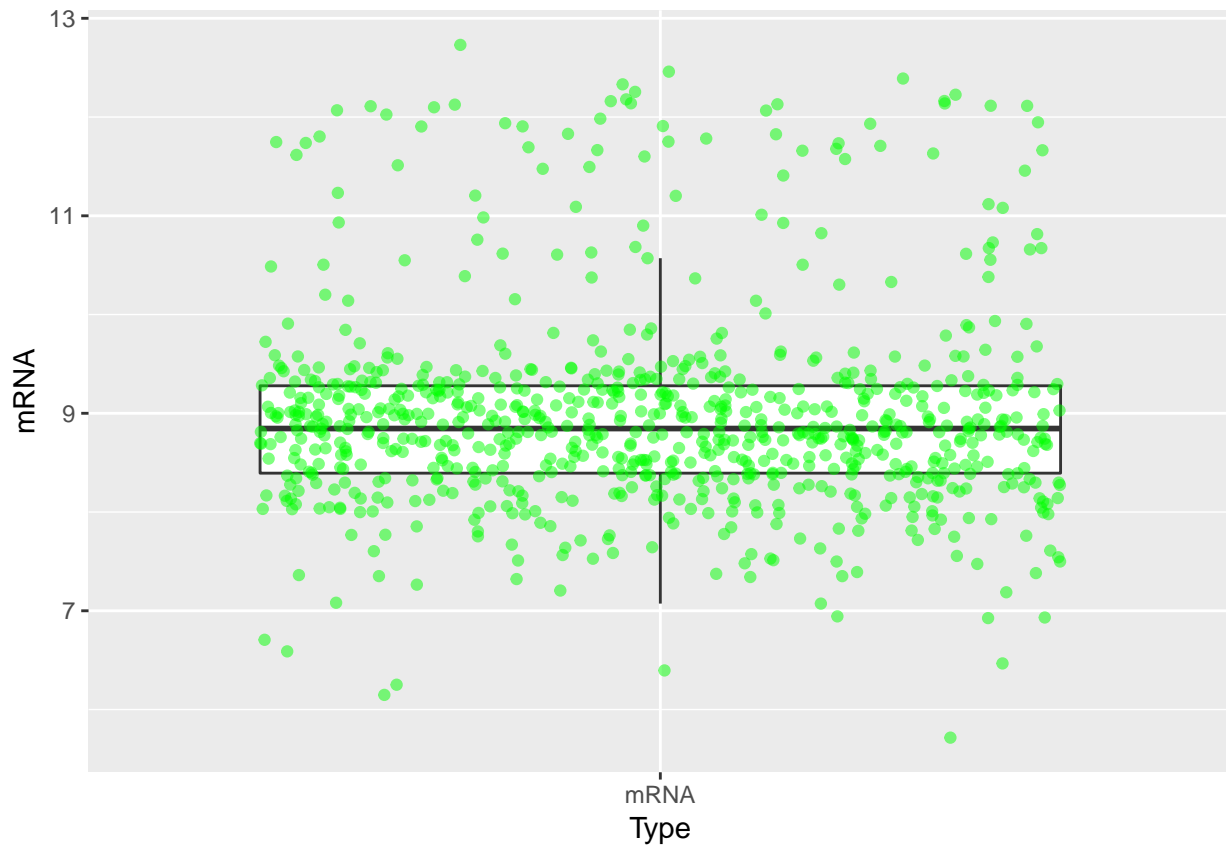
```
plot1
```



```
input.mrna = data.frame(log(mrnadata), "mRNA", stringsAsFactors = F)
colnames(input.mrna) = c("mRNA", "Type")

plot2 = ggplot(input.mrna, aes(x = Type, y = mRNA)) + geom_boxplot(width = 0.7,
  outlier.size = NA) + geom_jitter(width = 0.35, colour = "green", alpha = 0.5)

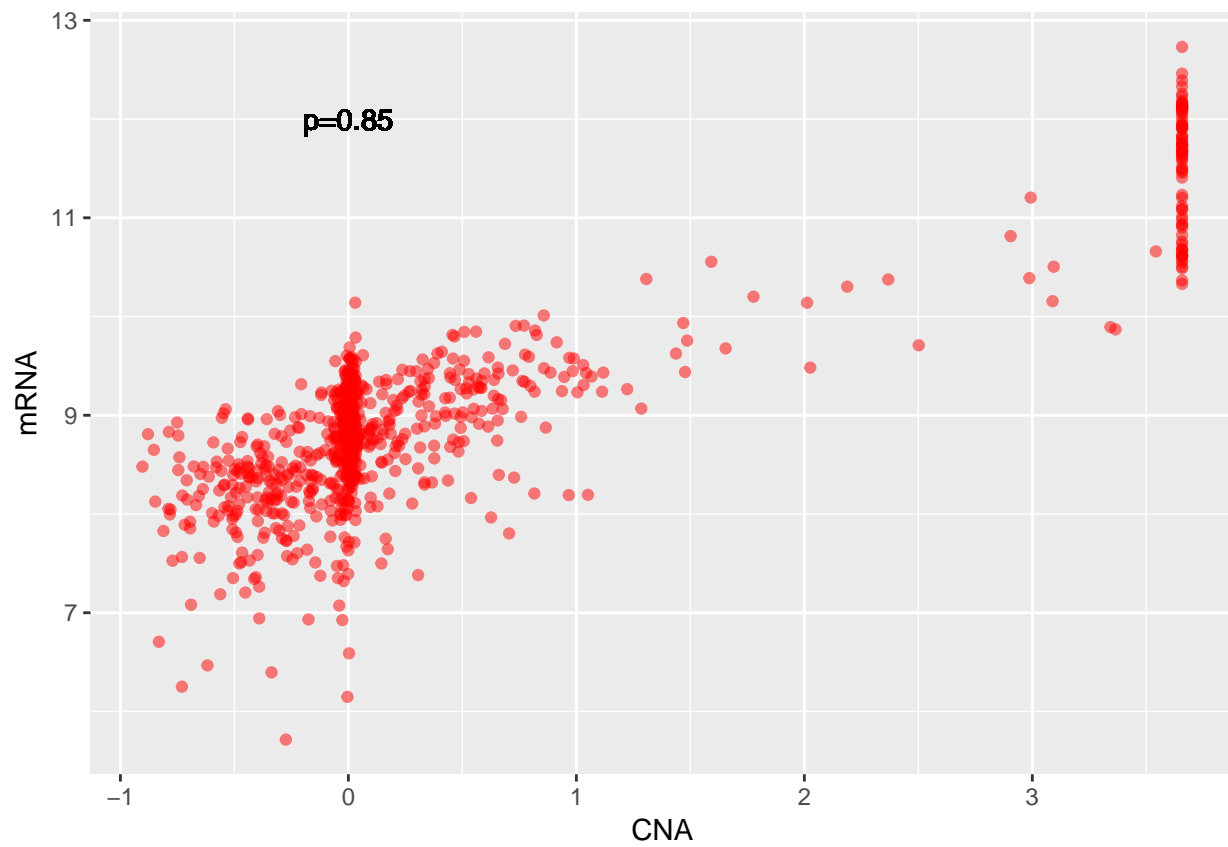
plot2
```



```
input.data = data.frame(cnadata, log(mrnadata), stringsAsFactors = F)
colnames(input.data) = c("CNA", "mRNA")

plot3 = ggplot(input.data, aes(x = CNA, y = mRNA)) + geom_point(color = "red",
  alpha = 0.5) + geom_text(aes(0, 12), label = "p=0.85")

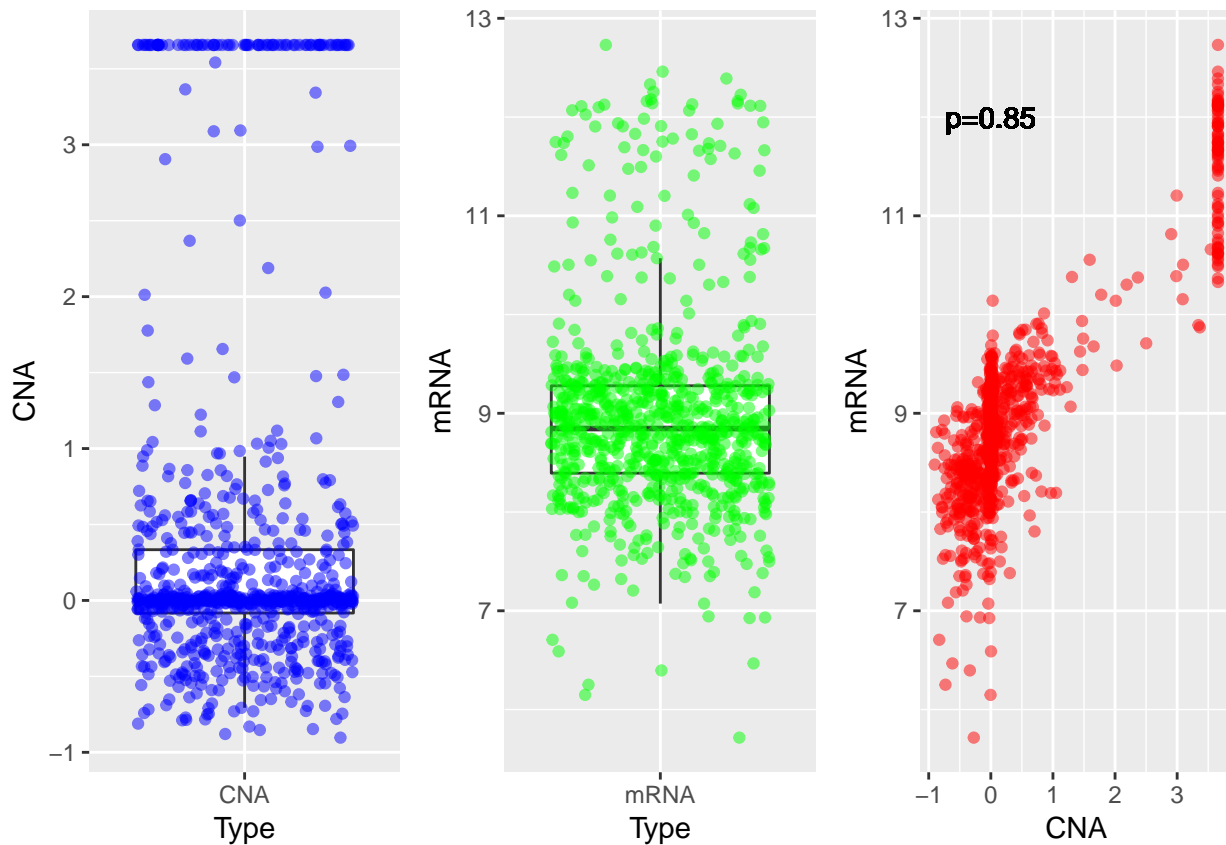
plot3
```



```
# install.packages('gridExtra')
```

```
library(gridExtra)
```

```
grid.arrange(plot1, plot2, plot3, nrow = 1, ncol = 3)
```



```
pdf("ggplot2.pdf", width = 12, height = 4)
grid.arrange(plot1, plot2, plot3, nrow = 1, ncol = 3)
dev.off()
```

## 4 Heatmap

```
library(ComplexHeatmap)

nmf.options(grid.patch = TRUE) # to generate a plot without the first blank page.

## Error in eval(expr, envir, enclos): could not find function "nmf.options"
geneset = c("RUNX1", "PIK3CA", "TP53", "GATA3", "FOXA1", "SF3B1", "PTEN", "CBFB",
"CDH1", "TBX3", "MAP2K4", "MAP3K1", "ERBB2", "KMT2C", "NCOR1", "FAM86B2",
"CDKN1B", "HIST1H3B", "THEM5", "FAM86B1", "GPS2", "AQP12A", "PIK3R1", "ACTL6B",
"ZFP36L1", "RB1", "KRAS", "EPDR1", "C1QTNF5", "ZFP36L2", "CTCF", "ASB10",
"FBXW7", "RPGR", "MYB", "TBL1XR1", "CASP8", "TCP10", "WSCD2", "AARS", "FAM20C",
"HIST1H2BC", "ARID1A", "PTHLH")

geneset_cnadata = getProfileData(mycgds, geneset, cna, mycaselist) #cna data

geneset_cnadata = geneset_cnadata[c(1:44), ] #sample selection

geneset_cnadata = t(geneset_cnadata) # transpose data

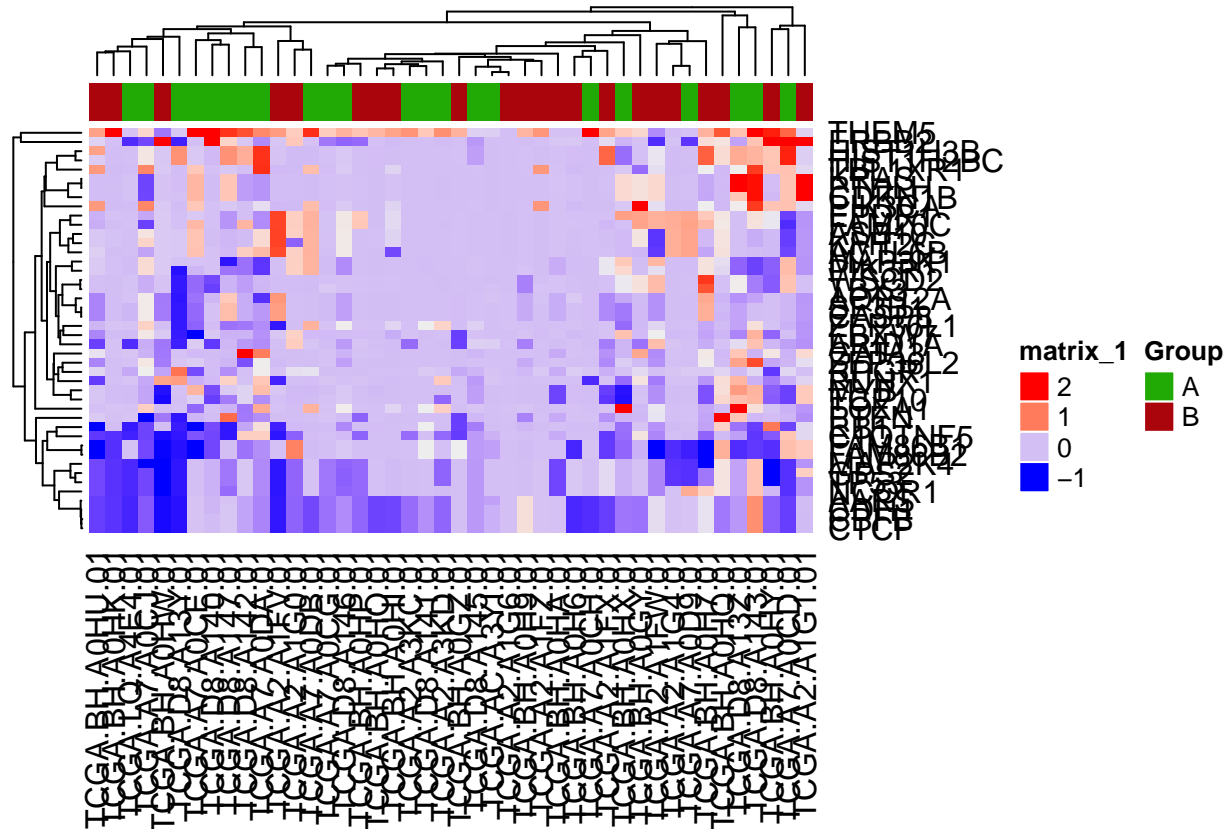
geneset_cnadata[is.na(geneset_cnadata)] <- 0
```



```
# View(geneset_cnadata)
```

```
ann_col = HeatmapAnnotation(Group = c(rep("A", 22), rep("B", 22)))
```

```
Heatmap(geneset_cnadata, top_annotation = ann_col)
```



```
pdf("heatmap.pdf", paper = "a4")
```

```
Heatmap(geneset_cnadata, top_annotation = ann_col)
```

```
dev.off()
```

```
write.table(geneset_cnadata[1:10, 1:10], "cnadata.10x10.txt")
```

```
write.table(geneset_cnadata[1:10, 1:10], "cnadata.10x10.txt", sep = "\t", quote = FALSE)
```