Azizi 2018 metadata

```r
globus_metadata <- readRDS("/mnt/nmorais-nfs/marta/pC_myeinfobank/metadata/globus/processed/GSE114724_A:
```

```r
globus_metadata
```

```
## Loading required package: SeuratObject
```

```
## Loading required package: sp
```

```
## An object of class Seurat
## 12908 features across 26947 samples within 1 assay
## Active assay: RNA (12908 features, 0 variable features)
```

```r
colnames(globus_metadata)[1:5]
```

```
## [1] "s1_AAACCTGAGCAGACTG-1" "s1_AAACCTGAGGTCGGAT-1" "s1_AAACCTGAGTGTACCT-1"
## [4] "s1_AAACCTGAGTGTACTC-1" "s1_AAACCTGCAACACCCG-1"
```

```r
colnames(globus_metadata@meta.data)
```

```
## [1] "patient"   "batch"     "condition" "sex"       "cell_type" "tissue"
```

```r
table(globus_metadata$patient)
```

```
##
##  BC10  BC11   BC9
##  4580  9360 13007
```

```r
table(globus_metadata$batch)
```

```
##
## GSE114724-BC10-s3 GSE114724-BC11-s4 GSE114724-BC11-s5  GSE114724-BC9-s1
##              4580              5031              4329              6553
##  GSE114724-BC9-s2
##              6454
```

```r
table(globus_metadata$condition)
```

```
##
## tumor_unknown
##         26947
```

```r
table(globus_metadata$sex)
```

```
##
## female
##  26947
```

```r
table(globus_metadata$tissue)
```

```
##
## breast
##  26947
```

```r
metadata <- data.frame(Cell_ID = colnames(globus_metadata),
                       Patient = globus_metadata$patient,
                       Sample = globus_metadata$batch)
rownames(metadata) <- NULL
```

```r
metadata$Sample_Type <- "Fresh"
metadata$Protocol <- "10X_chromium_5'"
metadata$Technology <- "10X_Genomics"
metadata$Dissociation <- "Mechanical,Enzymatic"
metadata$CellOrNucleus <- "Cell"
metadata$Sequencing_Machine <- "Illumina_NextSeq_500"
metadata$Preprocessing <- "CellRanger_v.2.1.1"
metadata$Reference_Genome <- "GRCh38"
metadata$Gender <- "Female"
metadata$Race <- NA
metadata$Menopause <- "Post"
metadata$Parity <- NA
metadata$Tissue <- "Tumour"
metadata$Treatment_Status <- NA
metadata$Treatment_Type <- NA
```

```r
metadata$Age <- metadata$Patient
metadata$Age[metadata$Age == "BC9"] <- 65
metadata$Age[metadata$Age == "BC10"] <- 73
metadata$Age[metadata$Age == "BC11"] <- 50
```

```r
metadata$Subtype <- metadata$Patient
metadata$Subtype[metadata$Subtype == "BC9"] <- "ER+"
metadata$Subtype[metadata$Subtype == "BC10"] <- "TNBC"
metadata$Subtype[metadata$Subtype == "BC11"] <- "HER2+"
```

```r
metadata$Grade <- metadata$Patient
metadata$Grade[metadata$Grade == "BC9"] <- 2
metadata$Grade[metadata$Grade == "BC10"] <- 3
metadata$Grade[metadata$Grade == "BC11"] <- 3
```

```r
metadata$Cancer_Type <- metadata$Patient
metadata$Cancer_Type[metadata$Cancer_Type == "BC9"] <- "ILC"
metadata$Cancer_Type[metadata$Cancer_Type == "BC10"] <- "IDC"
metadata$Cancer_Type[metadata$Cancer_Type == "BC11"] <- "IDC"
```

```r
metadata[1:5,]
```

```
##               Cell_ID Patient            Sample Sample_Type       Protocol
## 1 s1_AAACCTGAGCAGACTG-1     BC9 GSE114724-BC9-s1       Fresh 10X_chromium_5'
## 2 s1_AAACCTGAGGTCGGAT-1     BC9 GSE114724-BC9-s1       Fresh 10X_chromium_5'
## 3 s1_AAACCTGAGTGTACCT-1     BC9 GSE114724-BC9-s1       Fresh 10X_chromium_5'
## 4 s1_AAACCTGAGTGTACTC-1     BC9 GSE114724-BC9-s1       Fresh 10X_chromium_5'
## 5 s1_AAACCTGCAACACCCG-1     BC9 GSE114724-BC9-s1       Fresh 10X_chromium_5'
##     Technology         Dissociation CellOrNucleus   Sequencing_Machine
## 1 10X_Genomics Mechanical,Enzymatic          Cell Illumina_NextSeq_500
## 2 10X_Genomics Mechanical,Enzymatic          Cell Illumina_NextSeq_500
## 3 10X_Genomics Mechanical,Enzymatic          Cell Illumina_NextSeq_500
## 4 10X_Genomics Mechanical,Enzymatic          Cell Illumina_NextSeq_500
## 5 10X_Genomics Mechanical,Enzymatic          Cell Illumina_NextSeq_500
```

```
##         Preprocessing Reference_Genome Gender Race Menopause Parity Tissue
## 1 CellRanger_v.2.1.1          GRCh38 Female   NA      Post     NA Tumour
## 2 CellRanger_v.2.1.1          GRCh38 Female   NA      Post     NA Tumour
## 3 CellRanger_v.2.1.1          GRCh38 Female   NA      Post     NA Tumour
## 4 CellRanger_v.2.1.1          GRCh38 Female   NA      Post     NA Tumour
## 5 CellRanger_v.2.1.1          GRCh38 Female   NA      Post     NA Tumour
##   Treatment_Status Treatment_Type Age Subtype Grade Cancer_Type
## 1               NA             NA  65     ER+     2         ILC
## 2               NA             NA  65     ER+     2         ILC
## 3               NA             NA  65     ER+     2         ILC
## 4               NA             NA  65     ER+     2         ILC
## 5               NA             NA  65     ER+     2         ILC
```

```
#saveRDS(metadata, "azizi-2018-clinical-metadata.rds")
```