

Intro to ALL data for Bioc monograph

VJ Carey

October 24, 2025

1 Introduction

This document is for authors of the Bioc monograph, it just goes over various aspects of the ALL data. Example analyses can be added here for illustration.

2 Attachment and data list

```
> library(ALL)
> data(ALL)
> show(ALL)
```

```
ExpressionSet (storageMode: lockedEnvironment)
assayData: 12625 features, 128 samples
  element names: exprs
protocolData: none
phenoData
  sampleNames: 01005 01010 ... LAL4 (128 total)
  varLabels: cod diagnosis ... date last seen (21 total)
  varMetadata: labelDescription
featureData: none
experimentData: use 'experimentData(object)'
  pubMedIds: 14684422 16243790
Annotation: hgu95av2
```

3 Tables and graphs for phenodata

```
> print(summary(pData(ALL)))
```

cod	diagnosis	sex	age	BT
Length:128	Length:128	F :42	Min. : 5.00	B2 :36

```

Class :character   Class :character   M      :83   1st Qu.:19.00   B3      :23
Mode  :character   Mode  :character   NA's: 3   Median :29.00   B1      :19
                                         Mean  :32.37   T2      :15
                                         3rd Qu.:45.50   B4      :12
                                         Max.   :58.00   T3      :10
                                         NA's   :5       (Other):13

remission          CR                  date.cr                  t(4;11)
CR :99   Length:128                  Length:128                  Mode :logical
REF :15   Class :character            Class :character            FALSE:86
NA's:14   Mode  :character            Mode  :character            TRUE :7
                                         NA's :35

```

```

t(9;22)            cyto.normal          citog                      mol.biol
Mode :logical      Mode :logical          Length:128                  ALL1/AF4:10
FALSE:67           FALSE:69              Class :character            BCR/ABL :37
TRUE :26           TRUE :24              Mode  :character            E2A/PBX1: 5
NA's :35           NA's :35                      NEG      :74
                                         NUP-98   : 1
                                         p15/p16  : 1

```

```

fusion protein     mdr                kinet                ccr                relapse
p190      :17      NEG :101      dyploid:94      Mode :logical      Mode :logical
p190/p210: 8      POS : 24      hyperd.:27      FALSE:74          FALSE:35
p210      : 8      NA's: 3      NA's : 7      TRUE :26          TRUE :65
NA's      :95                      NA's :28          NA's :28

```

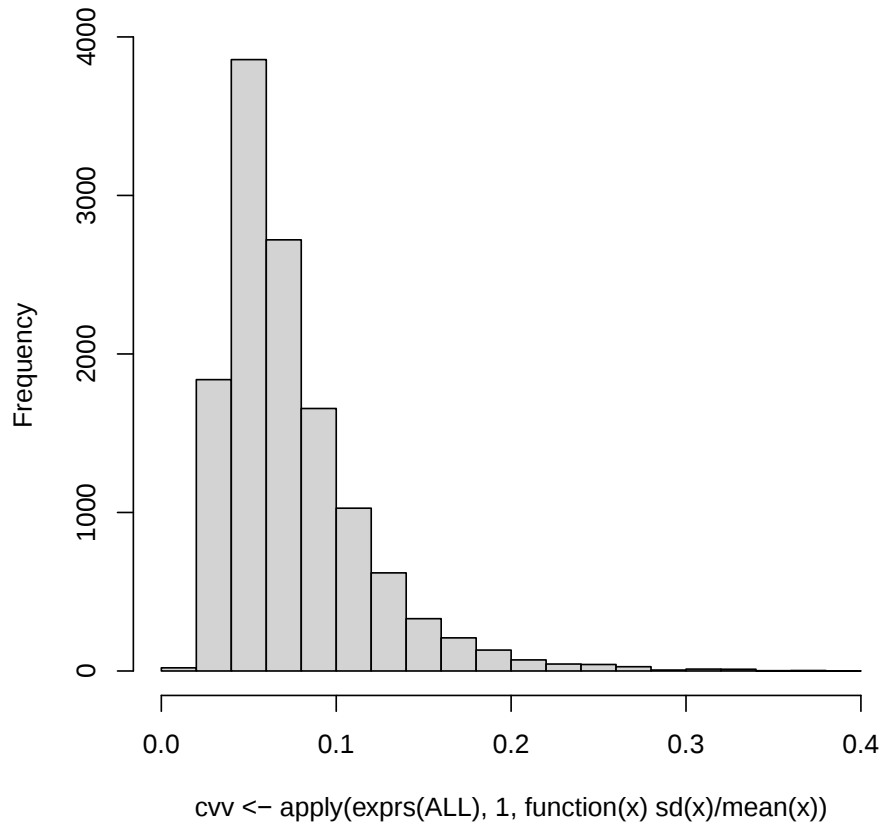
```

transplant          f.u                date last seen
Mode :logical      Length:128                  Length:128
FALSE:91           Class :character            Class :character
TRUE :9            Mode  :character            Mode  :character
NA's :28

```

```
> hist(cvv <- apply(exprs(ALL),1,function(x)sd(x)/mean(x)))
```

Histogram of `cvv <- apply(exprs(ALL), 1, function(x) sd(x)/mean`



```
> ok <- cvv > .08 & cvv < .18
> fALL <- ALL[ok,]
> show(fALL)
```

```
ExpressionSet (storageMode: lockedEnvironment)
assayData: 3841 features, 128 samples
  element names: exprs
protocolData: none
phenoData
  sampleNames: 01005 01010 ... LAL4 (128 total)
  varLabels: cod diagnosis ... date last seen (21 total)
  varMetadata: labelDescription
featureData: none
experimentData: use 'experimentData(object)'
  pubMedIds: 14684422 16243790
Annotation: hgu95av2
```

```
> allx2 <- data.frame(t(exprs(fALL)), class=ALL$BT)
```

```

> library(rpart)
> rp1 <- rpart(class~.,data=allx2)
> plot(rp1)
> text(rp1)

```

