

# Intro to ALL data for Bioc monograph

VJ Carey

October 10, 2026

## 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 |

|               |               |        |               |            |
|---------------|---------------|--------|---------------|------------|
| N.unique :128 | N.unique :115 | M :83  | 1st Qu.:19.00 | B3 :23     |
| N.blank : 0   | N.blank : 0   | NAs: 3 | Median :29.00 | B1 :19     |
| Min.nchar: 4  | Min.nchar: 8  |        | Mean :32.37   | T2 :15     |
| Max.nchar: 5  | Max.nchar: 10 |        | 3rd Qu.:45.50 | B4 :12     |
|               | NAs : 2       |        | Max. :58.00   | T3 :10     |
|               |               |        | NAs :5        | (Other):13 |

|           |               |               |               |               |
|-----------|---------------|---------------|---------------|---------------|
| remission | CR            | date.cr       | t(4;11)       | t(9;22)       |
| CR :99    | Length :128   | Length :128   | Mode :logical | Mode :logical |
| REF:15    | N.unique : 4  | N.unique : 92 | FALSE:86      | FALSE:67      |
| NAs:14    | N.blank : 0   | N.blank : 0   | TRUE :7       | TRUE :26      |
|           | Min.nchar: 2  | Min.nchar: 8  | NAs :35       | NAs :35       |
|           | Max.nchar: 18 | Max.nchar: 10 |               |               |
|           | NAs : 7       | NAs : 31      |               |               |

|               |               |             |                |         |
|---------------|---------------|-------------|----------------|---------|
| cyto.normal   | citog         | mol.biol    | fusion protein | mdr     |
| Mode :logical | Length :128   | ALL1/AF4:10 | p190 :17       | NEG:101 |
| FALSE:69      | N.unique : 15 | BCR/ABL :37 | p190/p210: 8   | POS: 24 |
| TRUE :24      | N.blank : 0   | E2A/PBX1: 5 | p210 : 8       | NAs: 3  |
| NAs :35       | Min.nchar: 6  | NEG :74     | NAs :95        |         |
|               | Max.nchar: 20 | NUP-98 : 1  |                |         |
|               | NAs : 35      | p15/p16 : 1 |                |         |

|            |               |               |               |               |
|------------|---------------|---------------|---------------|---------------|
| kinet      | ccr           | relapse       | transplant    | f.u           |
| dyploid:94 | Mode :logical | Mode :logical | Mode :logical | Length :128   |
| hyperd.:27 | FALSE:74      | FALSE:35      | FALSE:91      | N.unique : 11 |
| NAs : 7    | TRUE :26      | TRUE :65      | TRUE :9       | N.blank : 0   |
|            | NAs :28       | NAs :28       | NAs :28       | Min.nchar: 3  |
|            |               |               |               | Max.nchar: 17 |
|            |               |               |               | NAs : 28      |

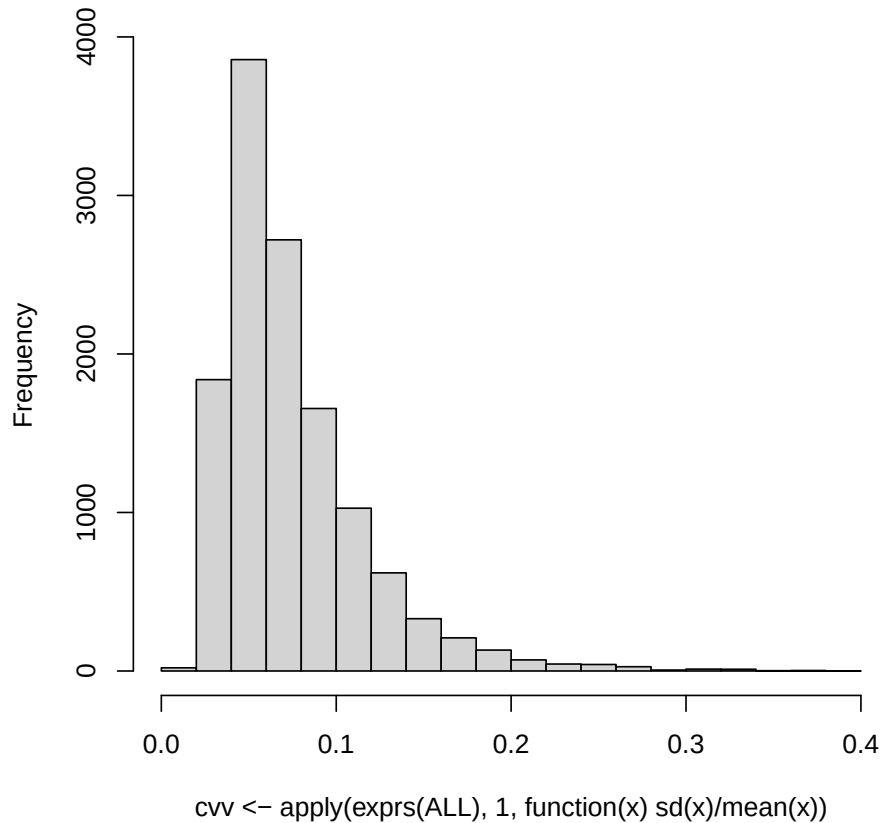
```

date last seen
Length :128
N.unique : 87
N.blank : 0
Min.nchar: 8
Max.nchar: 15
NAs : 35

```

```
> 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)

```

