Constructing the map from 'corrected RMA' spreadsheet to CEL files

VJ Carey (following concepts of Baggerly and Coombs)

April 19, 2009

```
> library(dressCheck)
> data(corrp)  # direct read of corrected RMA xls/csv
> data(c119)  # justRMA of CELs
> c119r = c119[rownames(corrp),]  # some genes omitted in corrected  RMA, align
> badn = sampleNames(DrAsGiven)
> getMap = function(nm) {  # this will associate column in corrp to column in c119r
+   x9 = corrp[,nm]
+   an = as.numeric
+   cc = sapply(1:length(sampleNames(c119r)), function(x) cor(an(x9), an(exprs(c119r)[
+   sampleNames(c119)[which.max(cc)]
+ }
> map = sapply(badn, getMap)
> table(table(map))  # some samples don't uniquely map

  1   2
113   3

> options(digits=4)
> cor(cbind(corrp[,c("M3514", "M4161")],
+   exprs(c119r)[,c("M3514", "M4161")]))

        M3514  M4161  M3514  M4161
M3514 1.0000 0.9580 0.9972 0.9591
M4161 0.9580 1.0000 0.9595 0.9558
M3514 0.9972 0.9595 1.0000 0.9619
M4161 0.9591 0.9558 0.9619 1.0000

> # from the above we conclude M3514 maps to M3514
> cor(cbind(corrp[,c("D1837", "M359", "M1572")],
+   exprs(c119r)[,c("D1837", "M359", "M1572")]))
```

```
        D1837    M359  M1572  D1837    M359  M1572
D1837 1.0000 0.9237 0.9350 0.9102 0.9318 0.9059
M359  0.9237 1.0000 0.9520 0.9477 0.9349 0.9371
M1572 0.9350 0.9520 1.0000 0.9469 0.9898 0.9435
D1837 0.9102 0.9477 0.9469 1.0000 0.9326 0.9381
M359  0.9318 0.9349 0.9898 0.9326 1.0000 0.9334
M1572 0.9059 0.9371 0.9435 0.9381 0.9334 1.0000

> # from the above we conclude M1572 maps to M359
>
> #from the next, we conclude that M2070 maps to M4161
> cor(cbind(corrp[,c("M2070", "M444", "M4161")],
+   exprs(c119r)[,c("M2070", "M444", "M4161")]))

        M2070   M444  M4161  M2070   M444  M4161
M2070 1.0000 0.9641 0.9587 0.9581 0.9673 0.9977
M444  0.9641 1.0000 0.9553 0.9508 0.9520 0.9612
M4161 0.9587 0.9553 1.0000 0.9546 0.9532 0.9558
M2070 0.9581 0.9508 0.9546 1.0000 0.9519 0.9580
M444  0.9673 0.9520 0.9532 0.9519 1.0000 0.9671
M4161 0.9977 0.9612 0.9558 0.9580 0.9671 1.0000
```