

Laboratoria nr.4

Imię i nazwisko

Punkty:

Przeglądanie danych Affymetrix'owych i normalizacja

1. Wczytywanie eksperymentu 'estrogen'

Wczytanie bibliotek (ewentualnie importowanie biocLite)
limma, affy, estrogen, hgu95av2cdf, vsn

Ładowanie danych:

```
> datadir <- system.file("extdata", package = "estrogen")
> datadir
> dir(datadir) # wyświetlenie plików z eksperymentu
```

Za pomocą funkcji **ReadTargets** wczytujemy do zmiennej **targets** plik 'targLimma.txt'

Zmienna **targets** powinna wyglądać tak

```
> targets
      filename estrogen time.h
low10-1.cel low10-1.cel absent 10
low10-2.cel low10-2.cel absent 10
high10-1.cel high10-1.cel present 10
high10-2.cel high10-2.cel present 10
low48-1.cel low48-1.cel absent 48
low48-2.cel low48-2.cel absent 48
high48-1.cel high48-1.cel present 48
high48-2.cel high48-2.cel present 48
```

Wczytanie do zmiennej **a** plików z eksperymentu za pomocą funkcji **ReadAffy**, z parametrem **filenames=targets\$FileName**

Jakiego typu obiektem jest zmienna **a**?

....

Wyświetl opis dla tego obiektu i podaj następujące polecenia w R:

Wyświetlenie pierwszych pięciu wierszy dla intensywności sygnałów *perfect match*

.....

Ile jest punktów dla każdej próbki (sample), a ile jest genów?

....

Jakim poleceniem w R wyświetlić numery punktów mikromacierzy, które są przyporządkowane do każdego z genów?

.....

2. Wykresy diagnostyczne – zamieść w sprawozdaniu. Jakie widzisz problemy na wykresach?

- Obraz płytki mikromacierzowej: **image**
Co widać na obrazie mikromacierzy w pliku 'bad.cel' {`image(ReadAffy("bad.cel"))`}
- **boxplot** – jaka jest różnica w wyświetlanych wartościach na osi Y jeśli użyjemy funkcji **Mbox**, a funkcji **boxplot**? (przypomnij sobie co jest przedstawiane na wykresach MAplot)
- **hist**, jakim poleceniem w R namalować histogram dla wszystkich próbek naraz, a jak tylko dla jednej?
- **MAplot** – co obserwujesz podczas rysowania wykresu MA? Ile wykresów MAplot jest wyświetlanych? Wykorzystując metodę **par** ustaw liczbę rysowanych wykresów podając liczbę kolumn i wierszy (**mfrow**)

3. Wybór metody normalizacji:

Wyświetl metody normalizacji danych, korekcji tła i podsumowania kilku sond dla jednego genu

```
normalize.methods(a)
```

```
bgcorrect.methods()
```

```
pmcorrect.methods()
```

```
express.summary.stat.methods()
```

Wykonaj normalizację (+ korekcję tła, metodę podsumowania różnych sond dla jednego genu) trzema różnymi metodami używając funkcji np. `expresso()` lub `rma()`. Zamieść w sprawozdaniu screen z sesji R. Porównaj wykresy diagnostyczne dla nowych zbiorów (`boxplot`, `hist`, `MAplot`). Przedstaw wnioski – która metoda jest najlepsza wg Ciebie?

```
> e1 = expresso(a, bg.correct = FALSE, normalize.method = "vsn",  
+ normalize.param = list(subsample = 1000), pmcorrect.method = "pmonly",  
+ summary.method = "medianpolish")
```

```
> e2 = rma(a)
```

Przydatne materiały 'affy.pdf'

```
>??affy
```