BIO392-cnv-freq

Angela Topic

10/01/2021

Step 1: Install package

```
if (!require(devtools)){
    install.packages('devtools')
}
```

```
## Loading required package: devtools
```

```
## Loading required package: usethis
```

```
devtools::install_github('progenetix/pgxRpi')
```

```
## Skipping install of 'pgxRpi' from a github remote, the SHA1 (ad1cf8a1) has not changed since last ins
##   Use `force = TRUE` to force installation
```

```
library(pgxRpi)
```

Analysis for different kind of cancer

For Glioblastoma (NCIT:C3058)

```
freq <- pgxLoader(type='frequency', output='pgxseg',filters='NCIT:C3058',
                  codematches=T)
```

```
##
##  accessing IntervalFrequencies service from Progenetix
```

The retreived data is an object contaning two slots `meta` and `data`.

```
freq$meta
```

```
##         code        label sample_count
## 1 NCIT:C3058 Glioblastoma         4305
```

```
names(freq$data)
```

```
## [1] "NCIT:C3058" "total"
```

```
head(freq$data$`NCIT:C3058`)
```

```
##      filters reference_name   start     end gain_frequency loss_frequency index
## 1 NCIT:C3058              1       0 1000000         10.337          6.527     0
## 2 NCIT:C3058              1 1000000 2000000         11.638          6.620     1
## 3 NCIT:C3058              1 2000000 3000000         12.474         13.287     2
## 4 NCIT:C3058              1 3000000 4000000         16.400         16.330     3
```

```
## 5 NCIT:C3058              1 4000000 5000000          11.661          18.513      4
## 6 NCIT:C3058              1 5000000 6000000          10.848          18.931      5
```

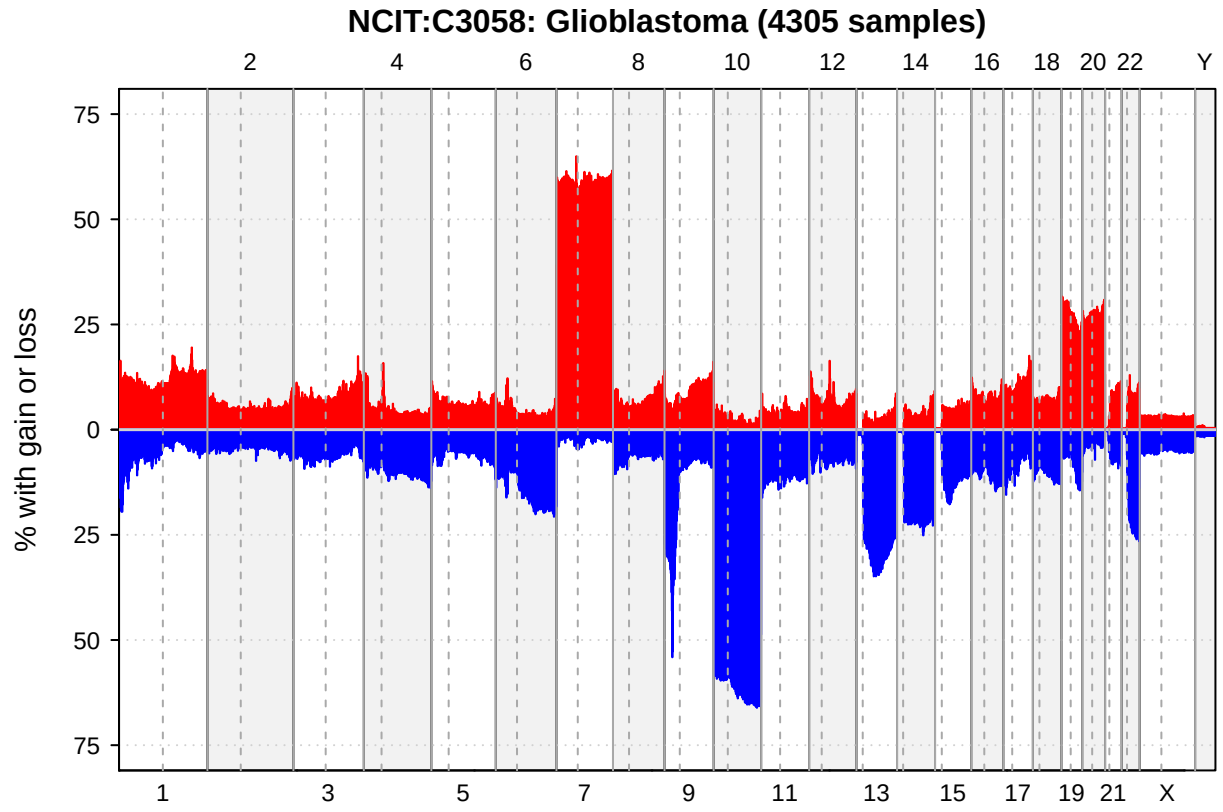
```
dim(freq$data$`NCIT:C3058`)
```

```
## [1] 3102    7
```

Visualize data

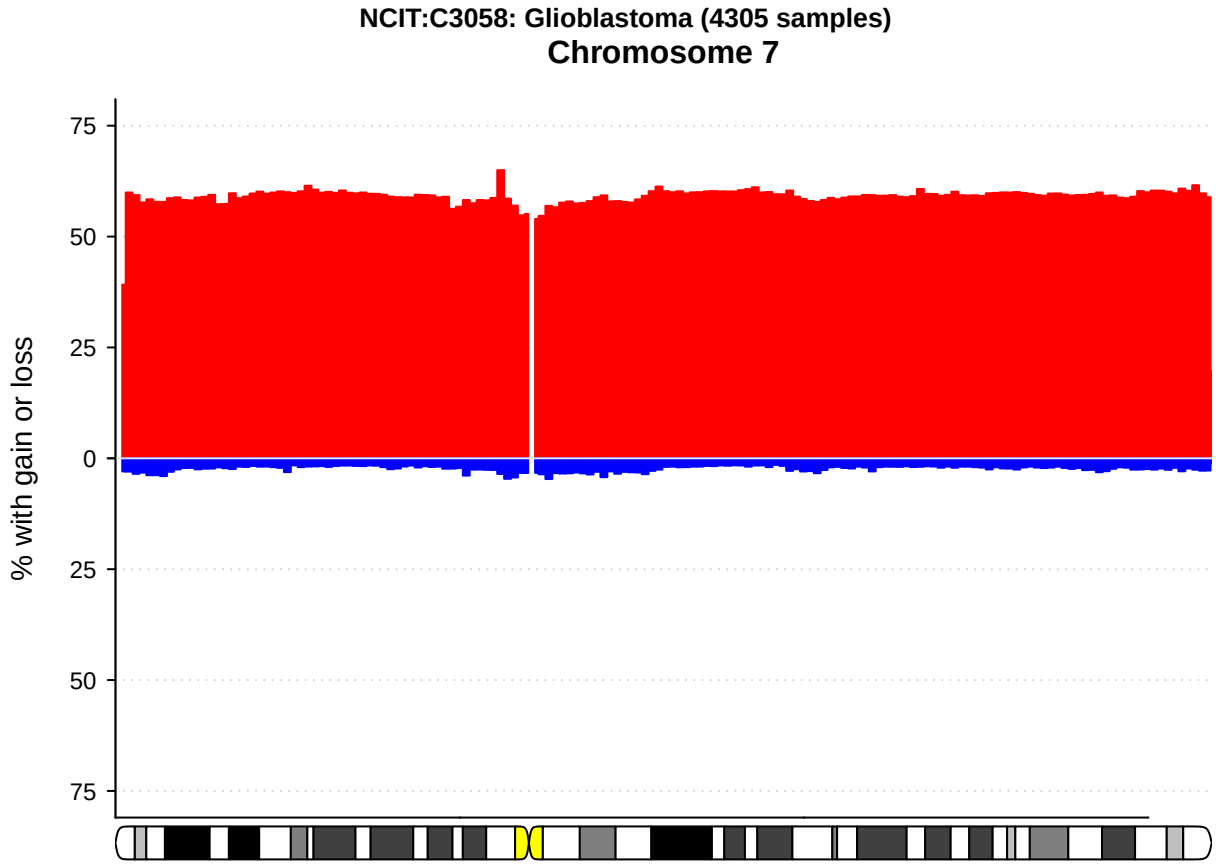
```
pgxFreqplot(freq)
```

By genome




```
pgxFreqplot(freq,chrom = 7)
```

By chromosome



Analyse the data

Gain is on 7q and 7p; Loss is on 9p; 10p and 10q; 13q;

In the literature I found especially the gain in chromosome 7 and the loss in 10. paper-link

For Invasive Breast Carcinoma (NCIT:C9245)

```
freq <- pgxLoader(type='frequency', output='pgxseg',filters='NCIT:C9245',
                  codematches=T)
```

```
##
##  accessing IntervalFrequencies service from Progenetix
```

The retreived data is an object contaning two slots `meta` and `data`.

```
freq$meta
```

```
##         code                      label sample_count
## 1 NCIT:C9245 Invasive Breast Carcinoma         2745
```

```
names(freq$data)
```

```
## [1] "NCIT:C9245" "total"
```

```
head(freq$data$`NCIT:C9245`)
```

```
##      filters reference_name   start     end gain_frequency loss_frequency index
```

```
## 1 NCIT:C9245              1       0 1000000          2.441          2.404     0
## 2 NCIT:C9245              1 1000000 2000000          3.424          8.306     1
## 3 NCIT:C9245              1 2000000 3000000          4.736         38.251     2
## 4 NCIT:C9245              1 3000000 4000000          3.607         37.341     3
## 5 NCIT:C9245              1 4000000 5000000          3.679         38.834     4
## 6 NCIT:C9245              1 5000000 6000000          3.424         38.069     5
```

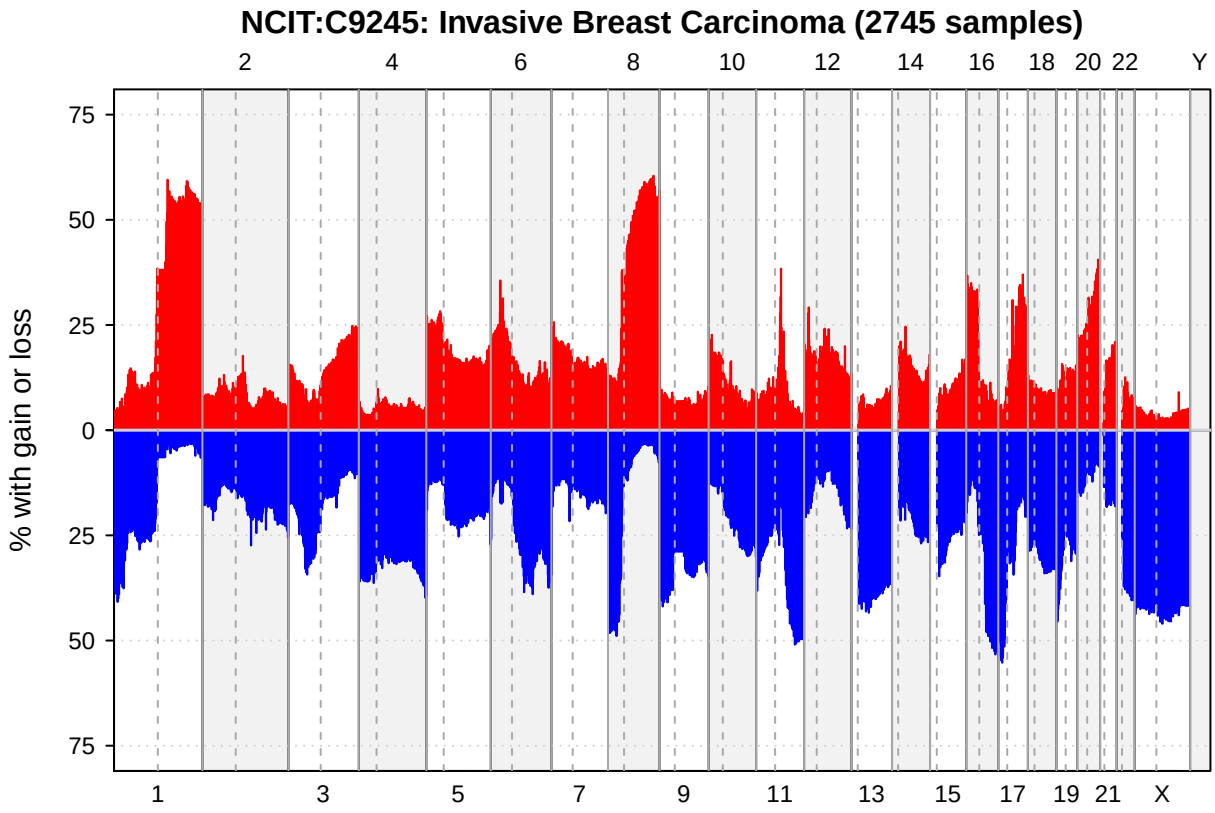
```
dim(freq$data$`NCIT:C9245`)
```

```
## [1] 3102    7
```

Visualize data

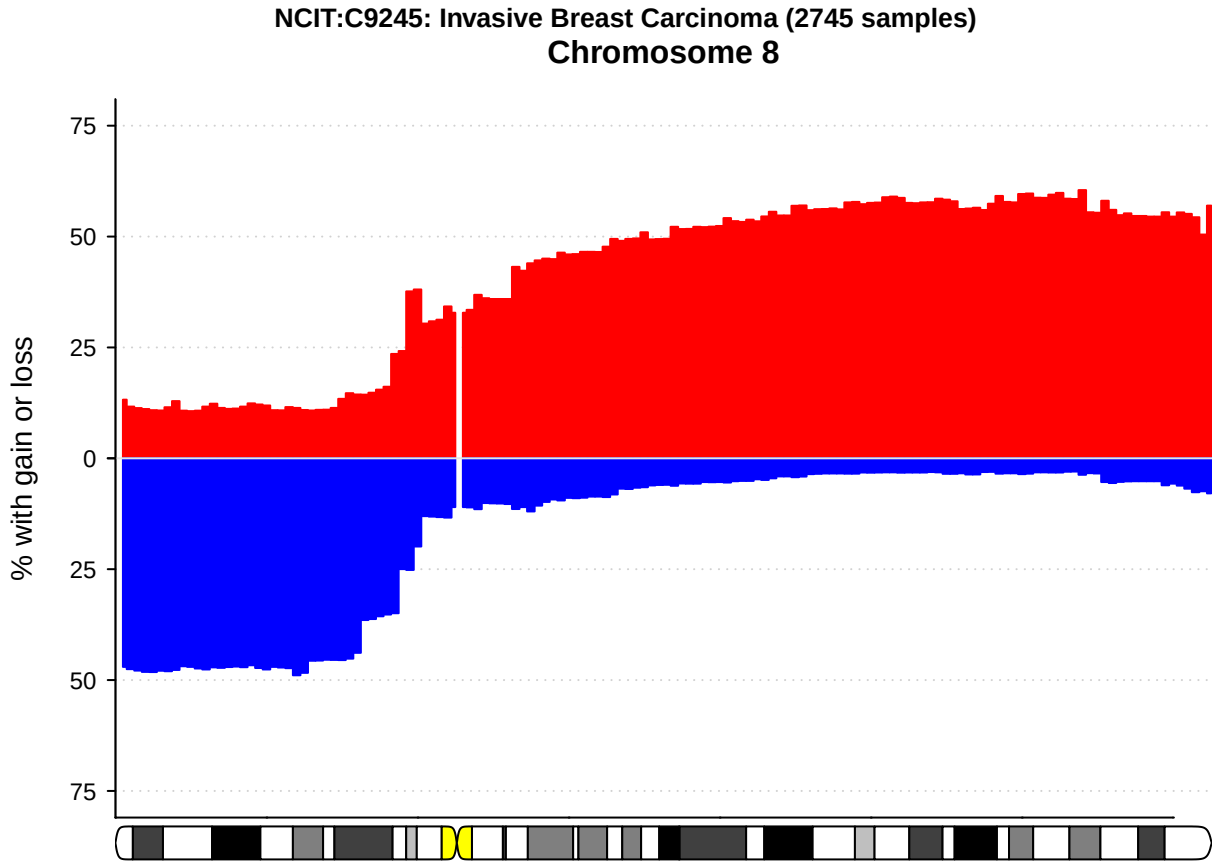
```
pgxFreqplot(freq)
```

By genome




```
pgxFreqplot(freq,chrom = 8)
```

By chromosome



Analyse the data

Gains with a frequency over 50% are in: 1q; 8q;

Losses are in a lot of chromosomes frequent with 25 -50%

What I found in literature: It is consistent what I found with the gains but with the loses there were others found, which are more frequent then others. Especially 5,6 and 8. For chromosome 5, I doesn't really find a high frequency.

paper-link