

# Summary of TMB runs

Lena Morrill

24/05/2021

```
## Warning in .recacheSubclasses(def@class, def, env): undefined subclass
## "numericVector" of class "Mnumeric"; definition not updated

## Loading required package: viridisLite

## Loading required package: coda

## Loading required package: MASS

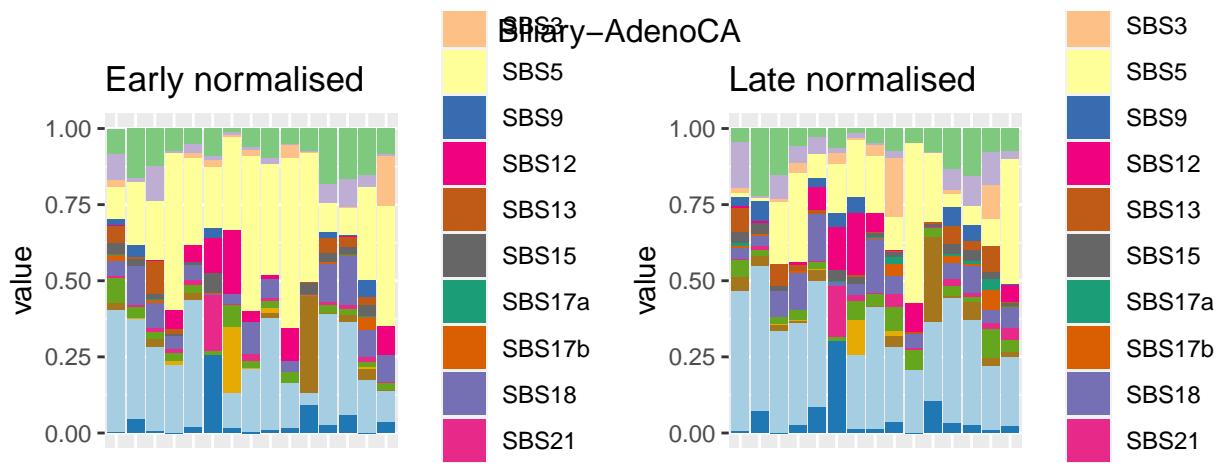
## Warning in .recacheSubclasses(def@class, def, env): undefined subclass
## "numericVector" of class "Mnumeric"; definition not updated

## ##
## ## Markov Chain Monte Carlo Package (MCMCpack)

## ## Copyright (C) 2003-2021 Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park

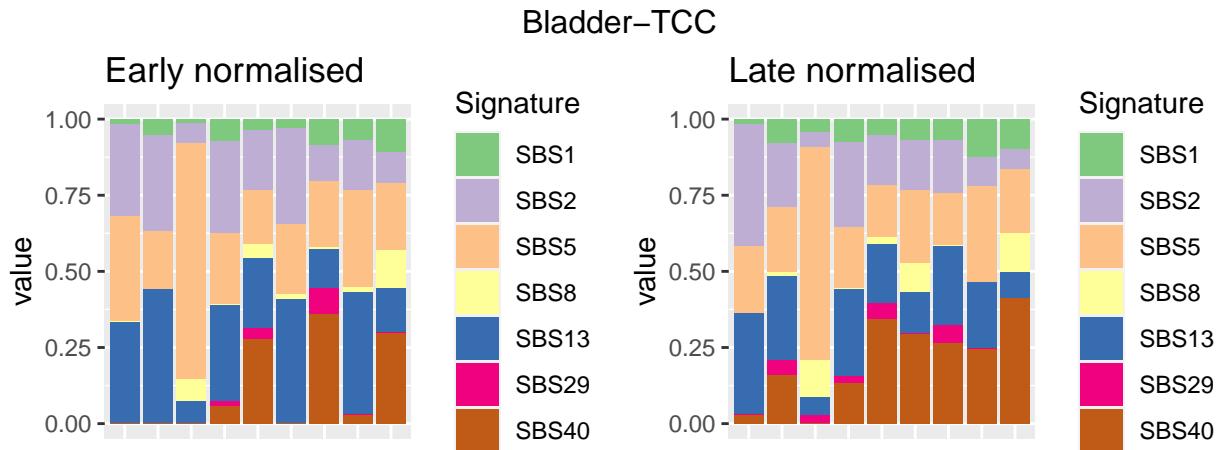
## ##
## ## Support provided by the U.S. National Science Foundation
## ## (Grants SES-0350646 and SES-0350613)
## ##
```

## Using subset of active signatures from the PCAWG paper

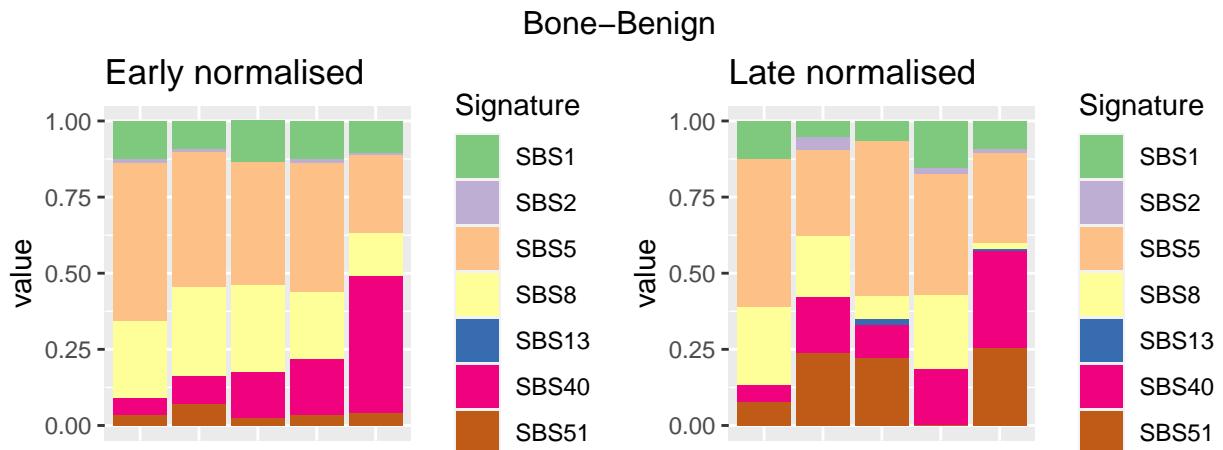


```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
```

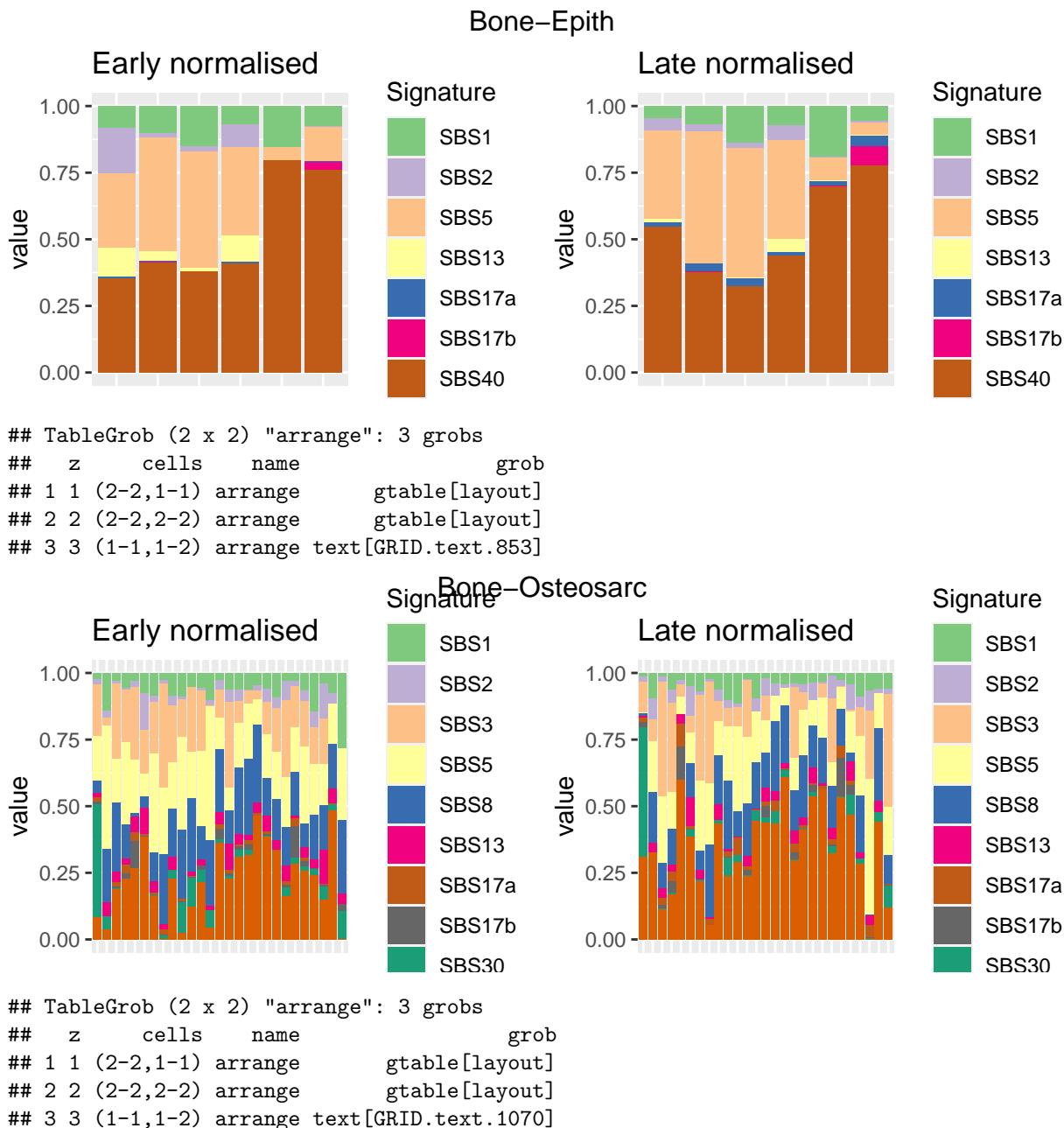
```
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.328]
```

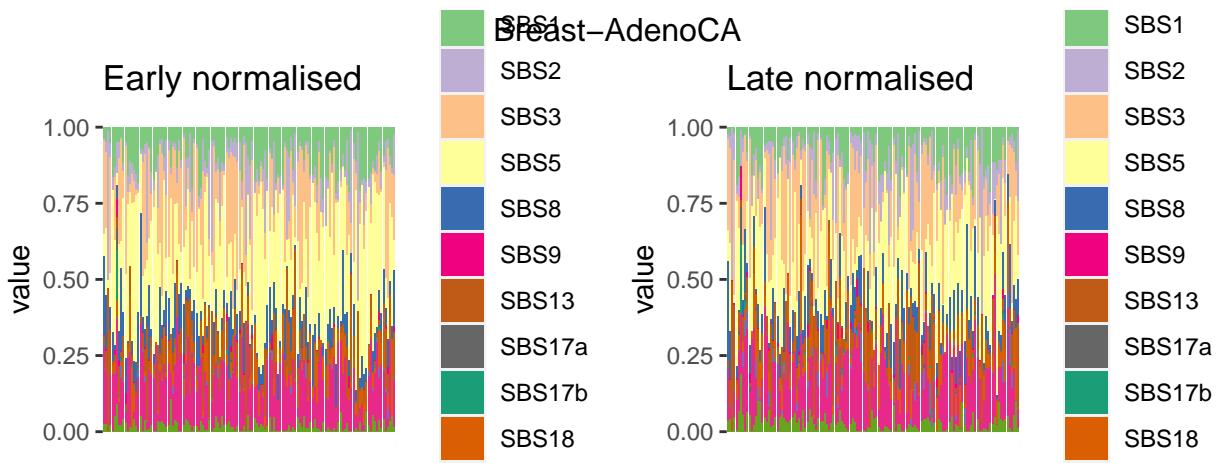


```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.503]
```



```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z      cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.678]
```





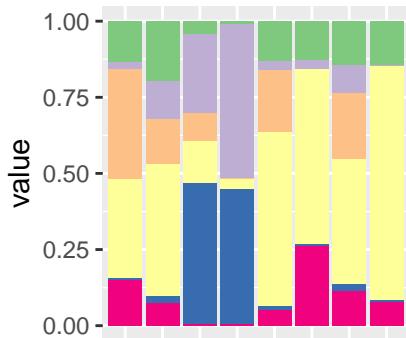
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1329]
```



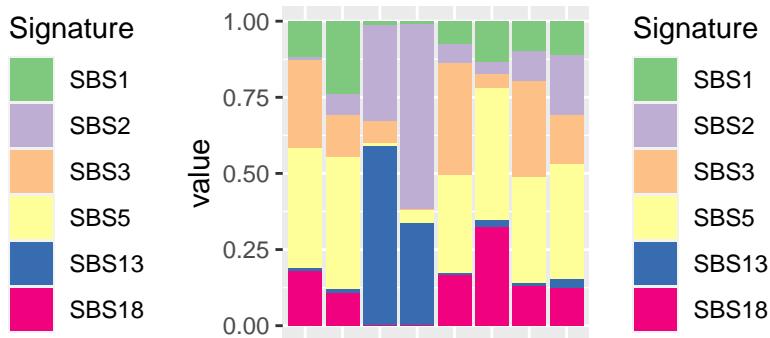
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1448]
```

### Breast–LobularCA

Early normalised



Late normalised



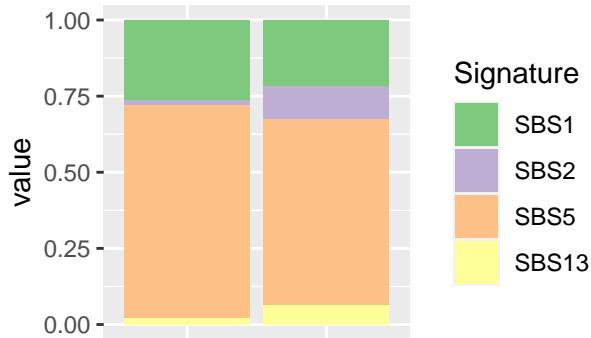
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1609]
```

### Cervix–AdenoCA

Early normalised



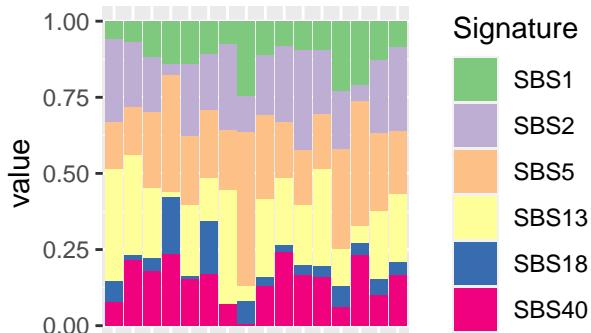
Late normalised



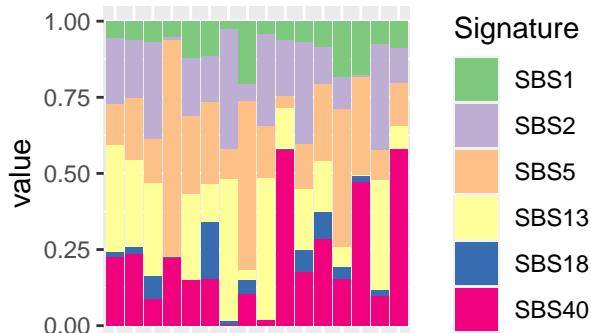
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1742]
```

## Cervix–SCC

Early normalised



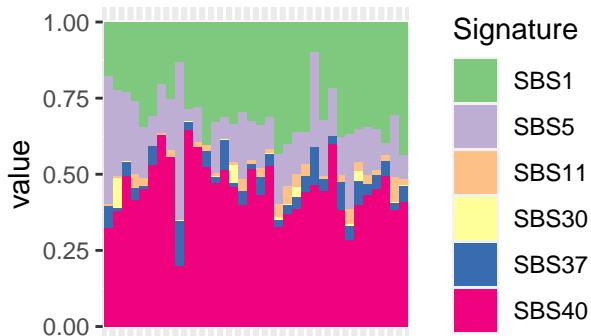
Late normalised



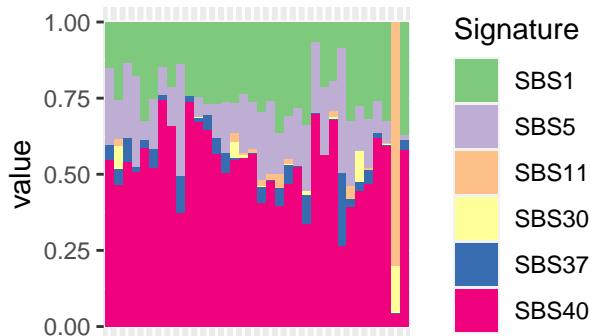
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1903]
```

## CNS–GBM

Early normalised



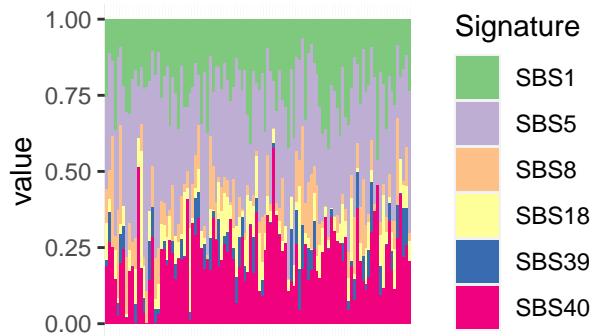
Late normalised



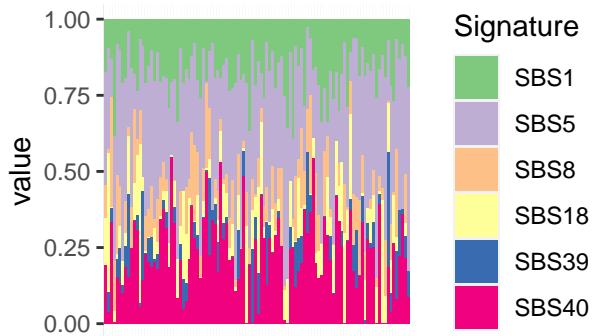
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2064]
```

### CNS-Medullo

Early normalised



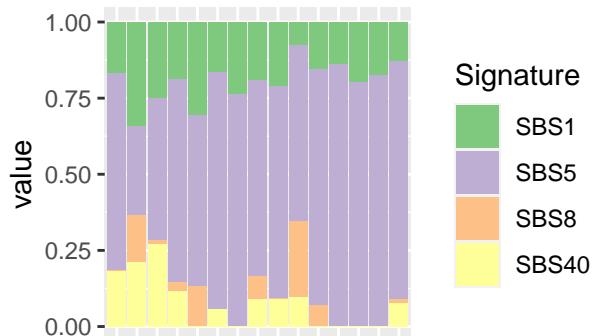
Late normalised



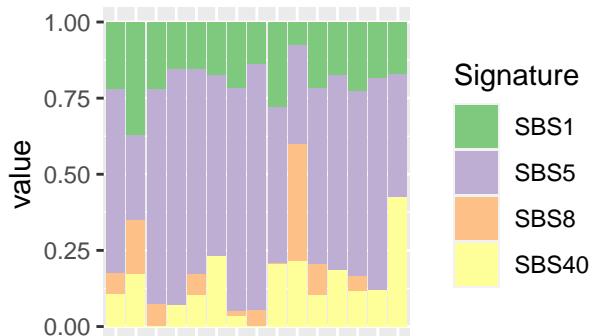
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2225]
```

### CNS-Oligo

Early normalised

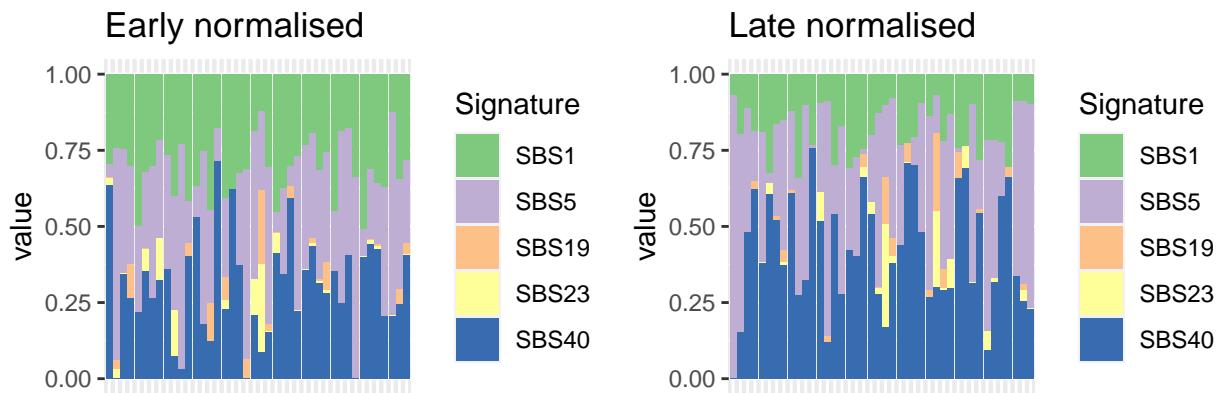


Late normalised

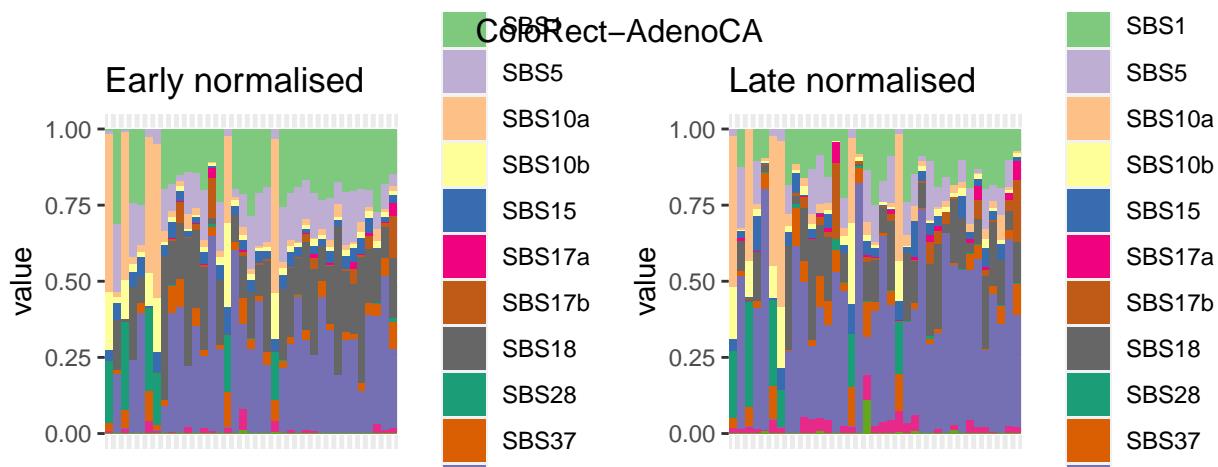


```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2358]
```

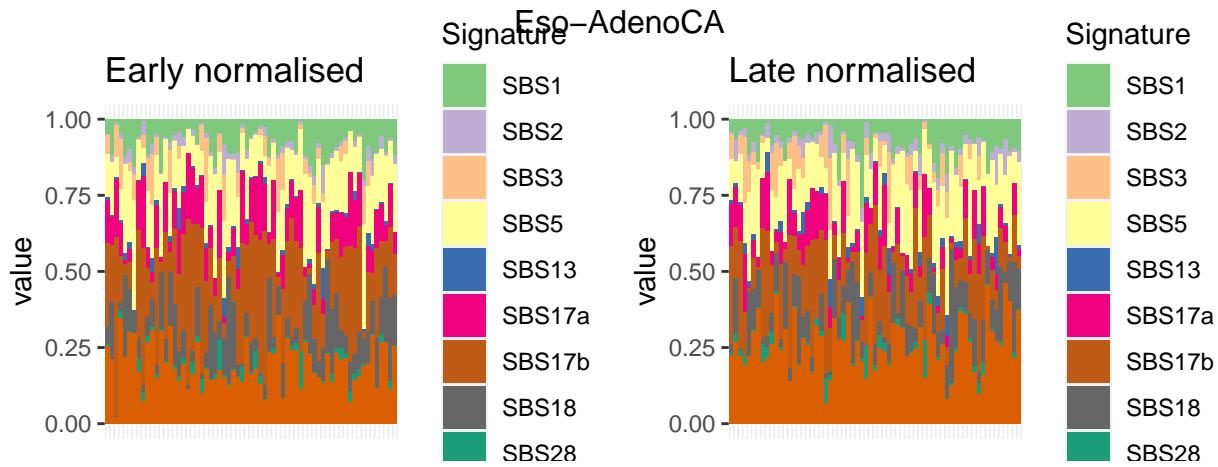
## CNS–PiloAstro



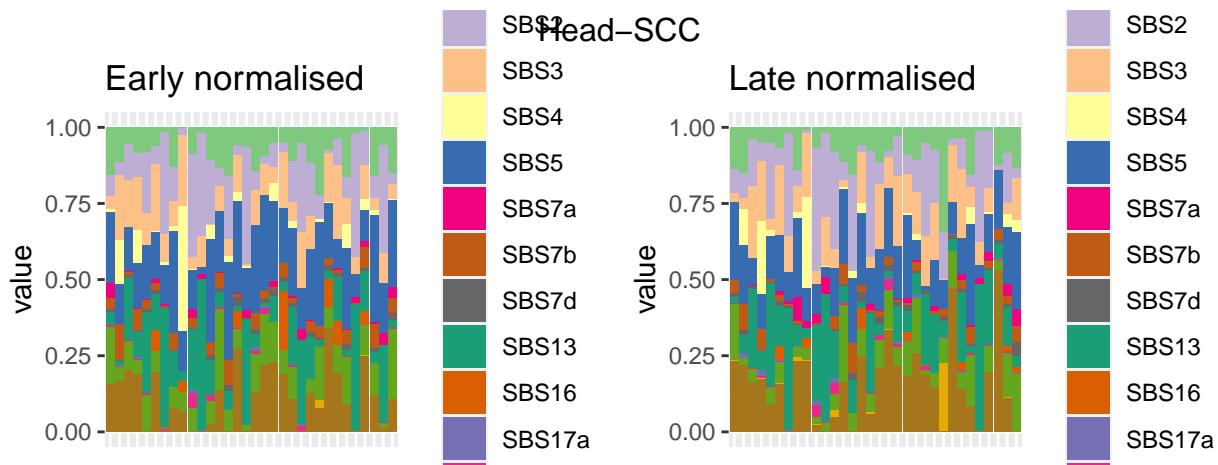
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2505]
```



```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2764]
```

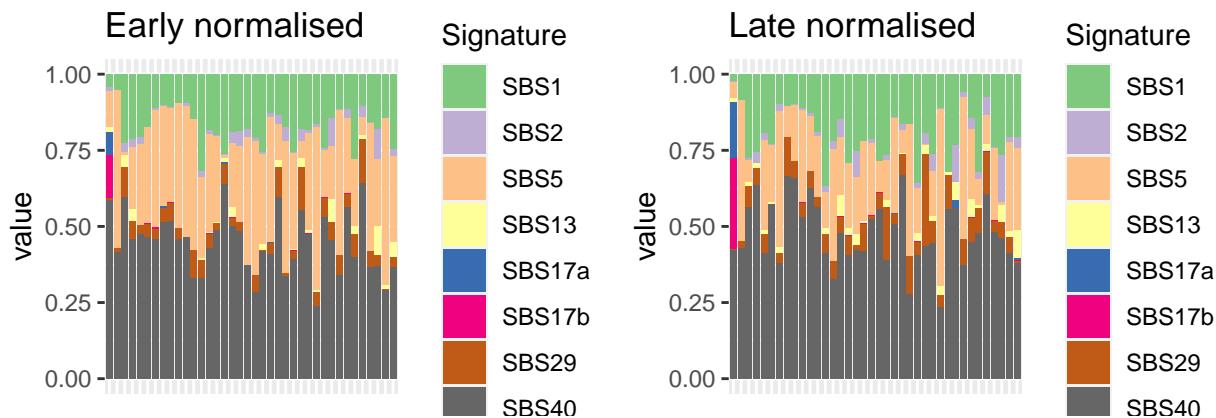


```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2981]
```



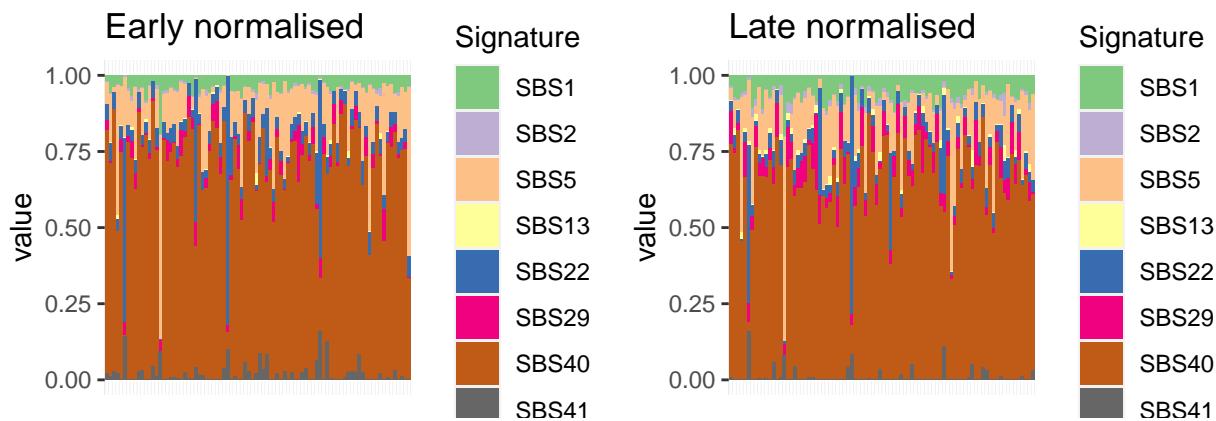
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3268]
```

### Kidney–ChRCC



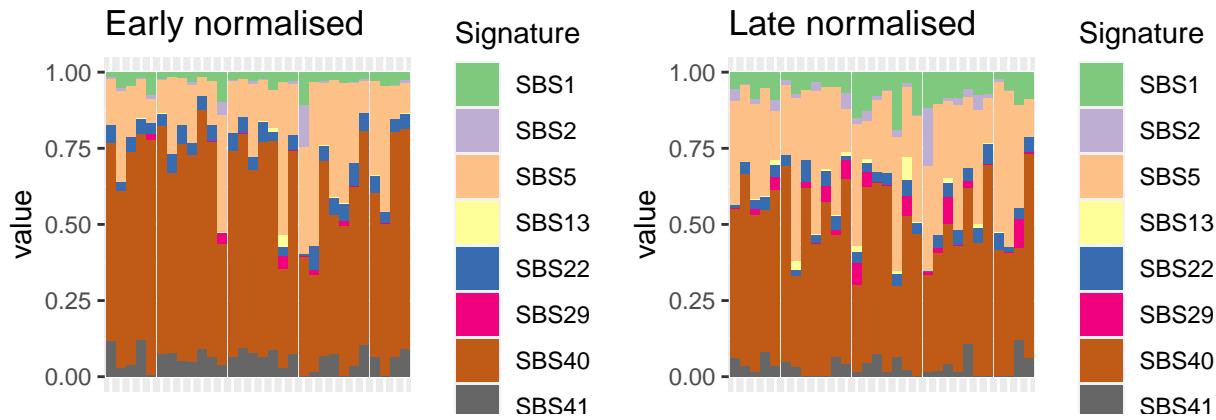
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3457]
```

### Kidney–RCC.clearcell

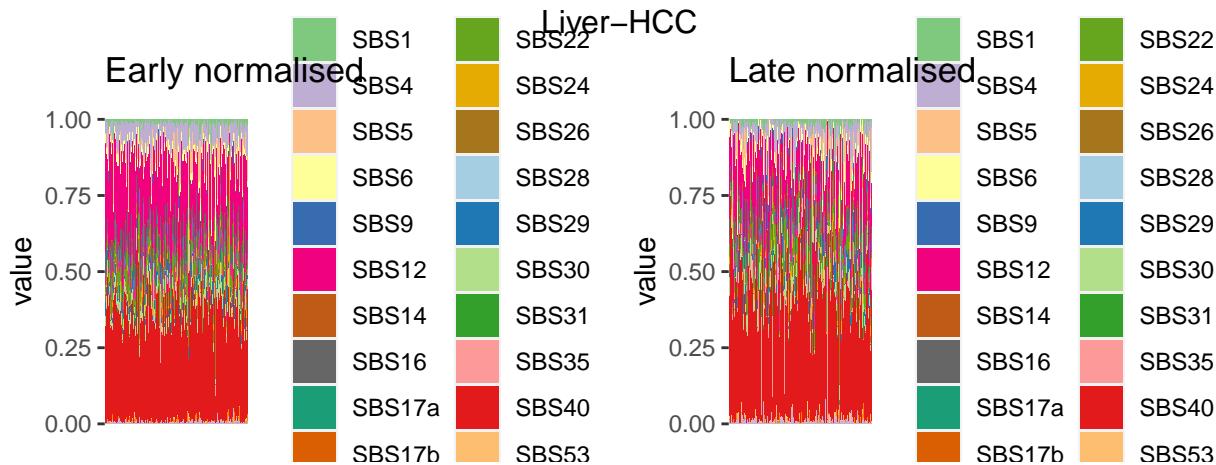


```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells  name          grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3646]
```

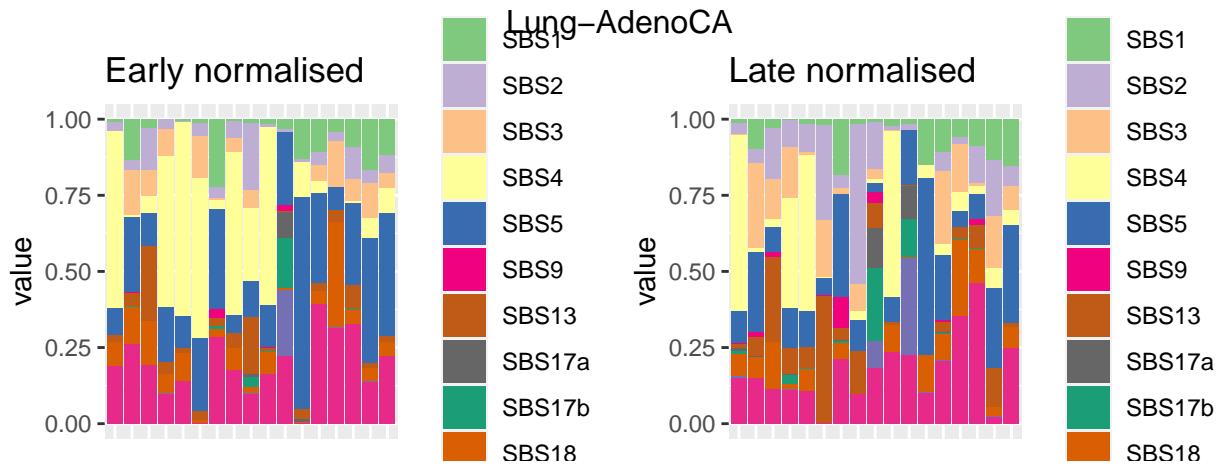
### Kidney–RCC.papillary



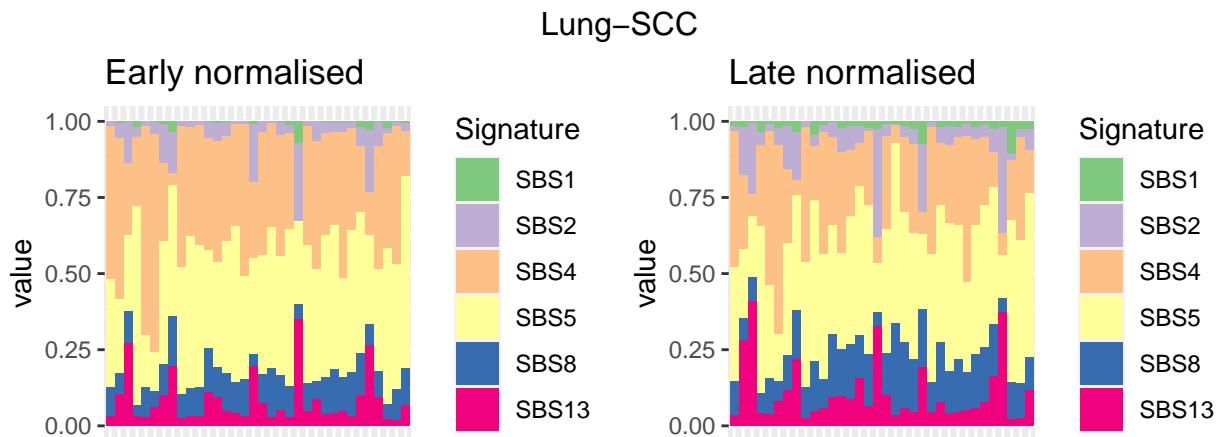
```
## TableGrob (2 x 2) "arrange": 3 grobs
##  z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3835]
```



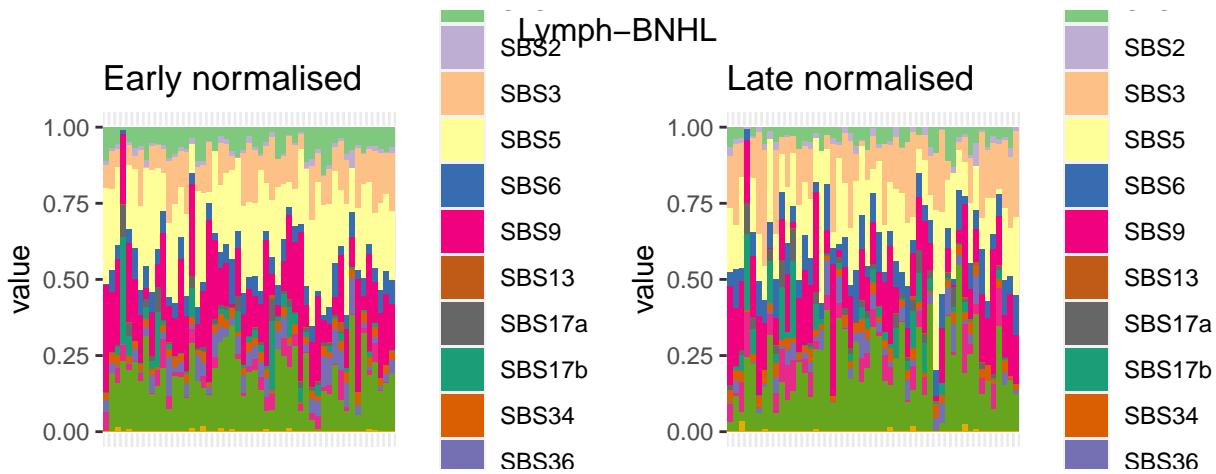
```
## TableGrob (2 x 2) "arrange": 3 grobs
##  z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4248]
```



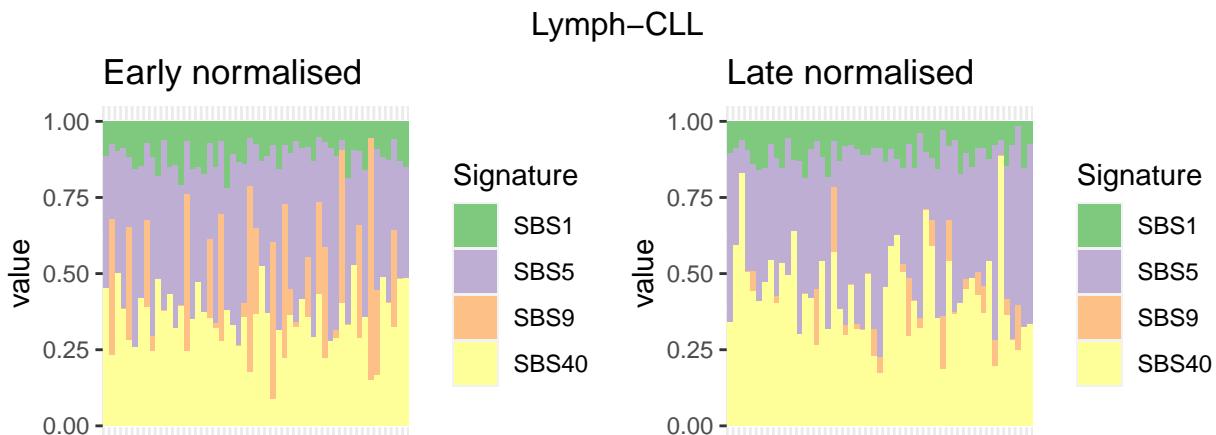
```
## TableGrob (2 x 2) "arrange": 3 grobs
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4493]
```



```
## TableGrob (2 x 2) "arrange": 3 grobs
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4654]
```



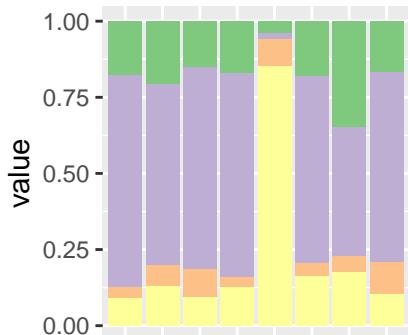
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4927]
```



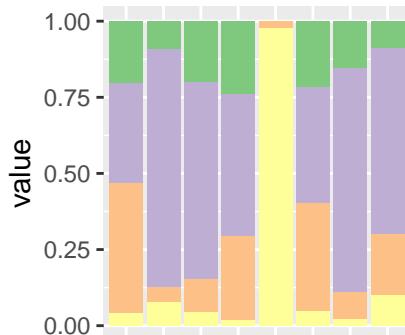
```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5060]
```

### Myeloid–AML

**Early normalised**



**Late normalised**



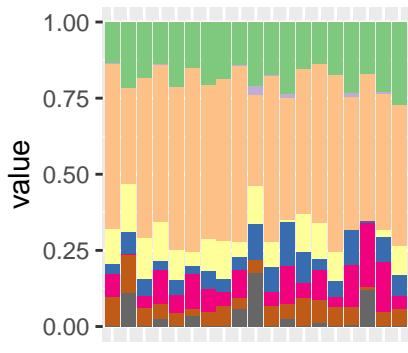
```

## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5193]

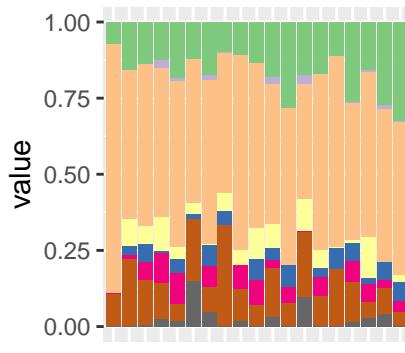
```

### Myeloid–MPN

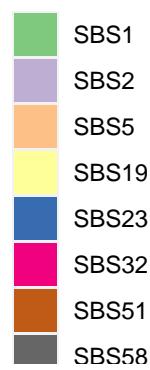
**Early normalised**



**Late normalised**



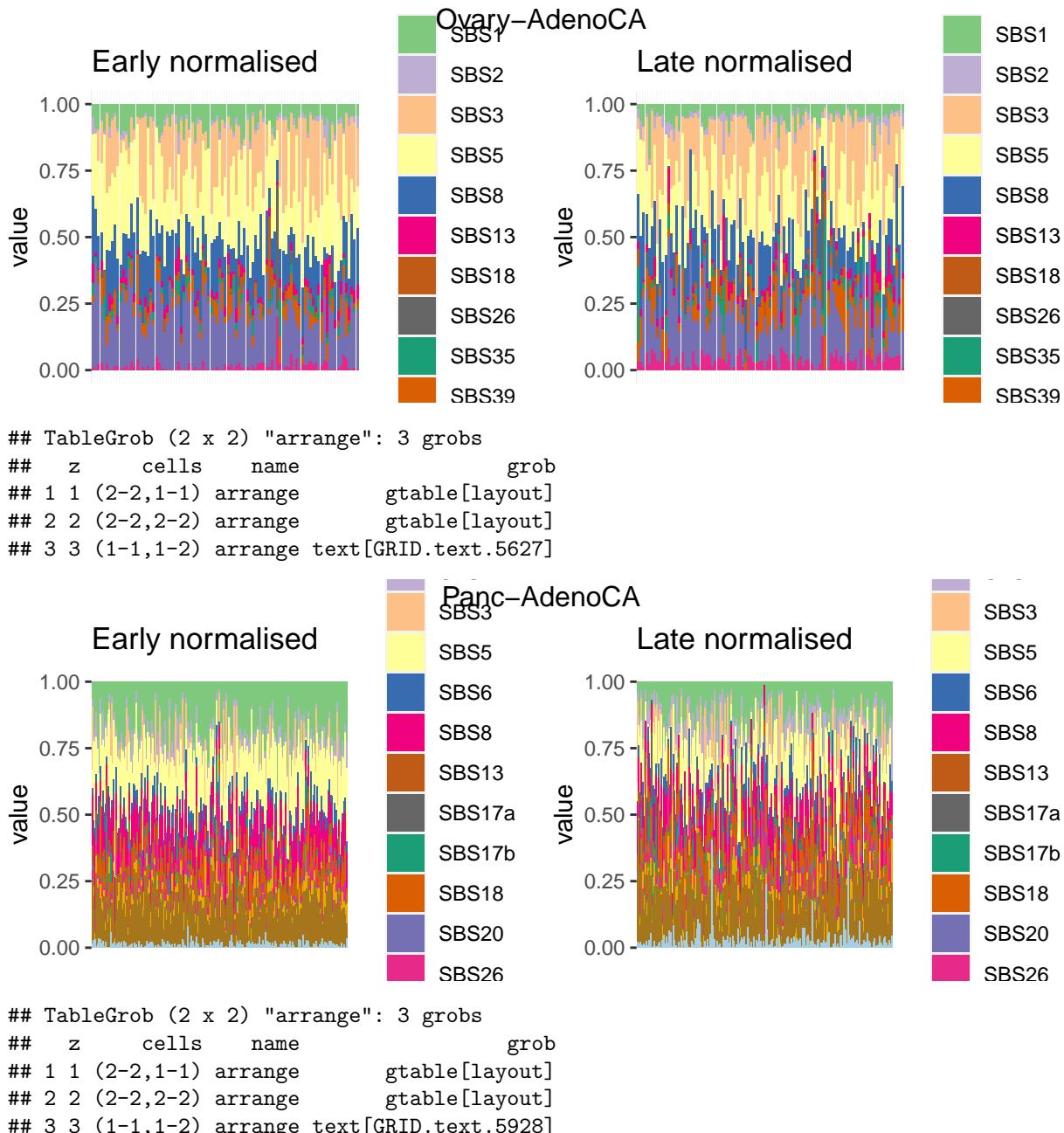
**Signature**

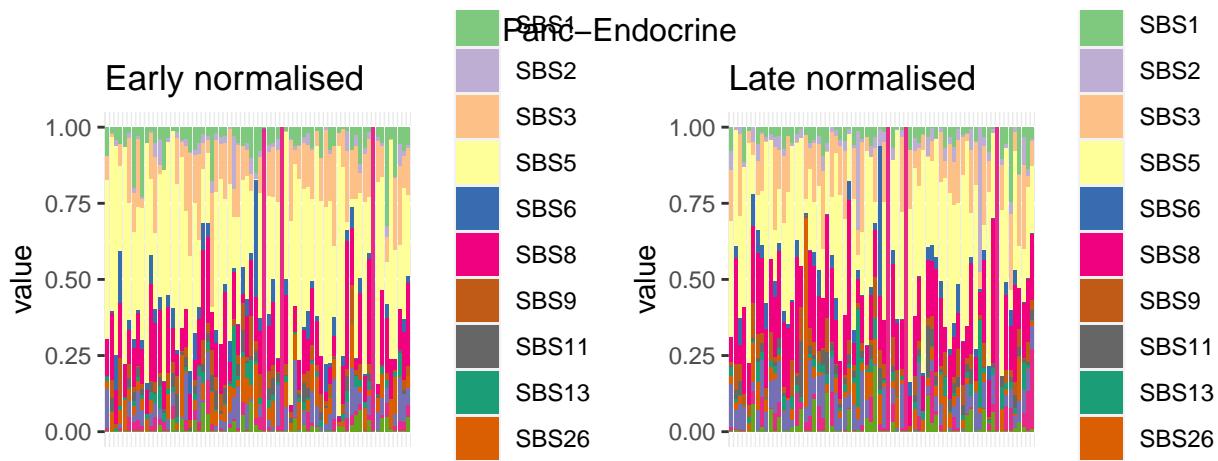


```

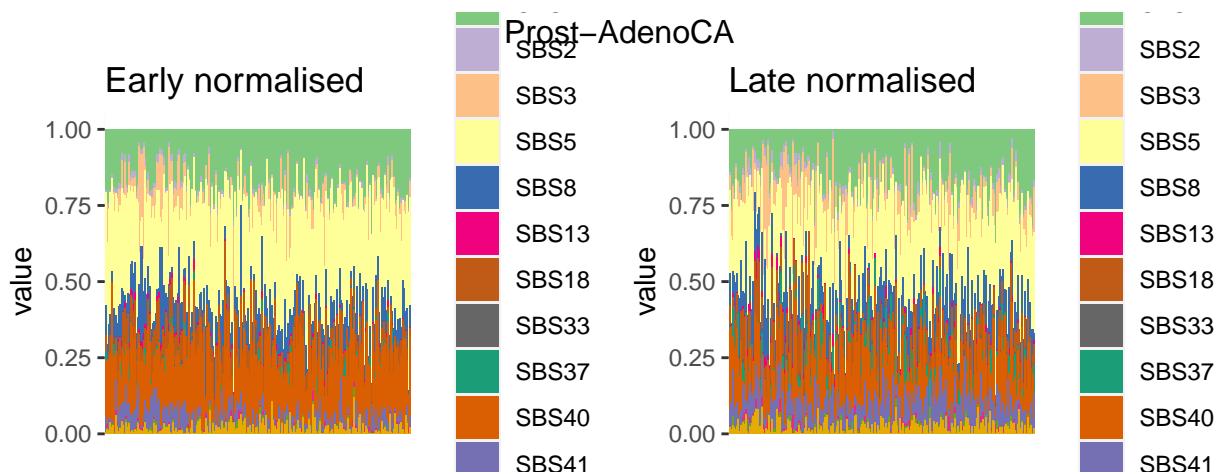
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5382]

```

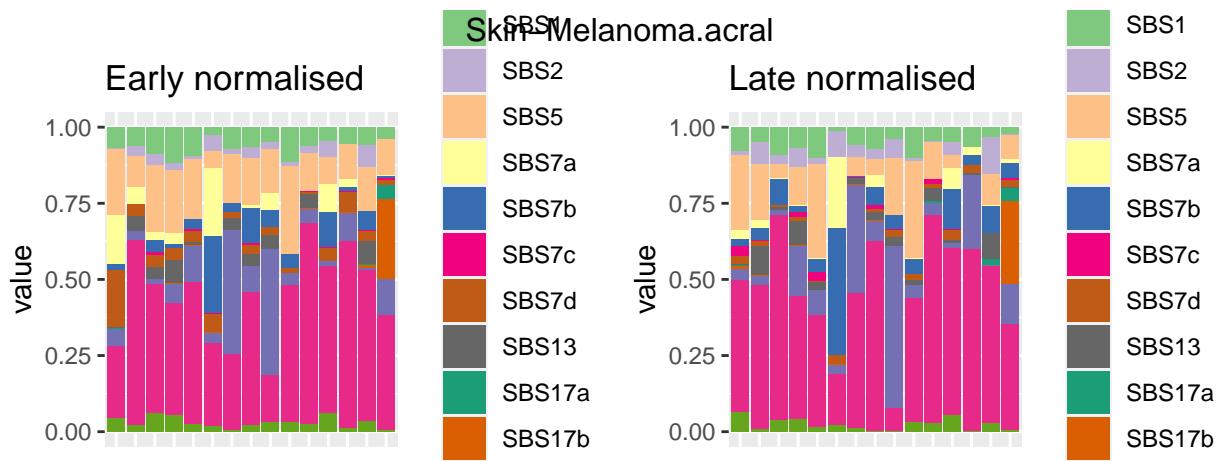




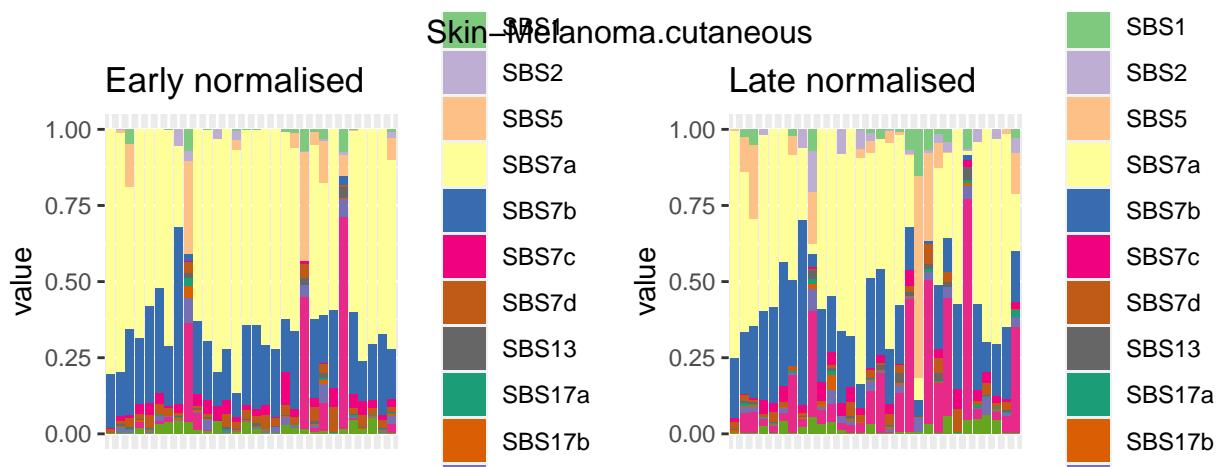
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6187]
```



```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6460]
```



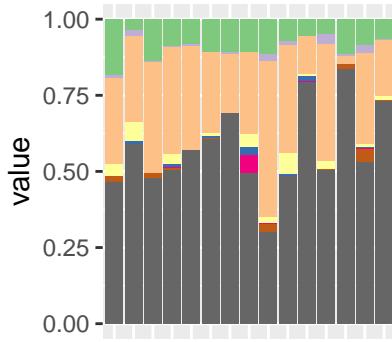
```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6719]
```



```
## # TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6978]
## Error : $ operator is invalid for atomic vectors
```

### SoftTissue–Leiomyo

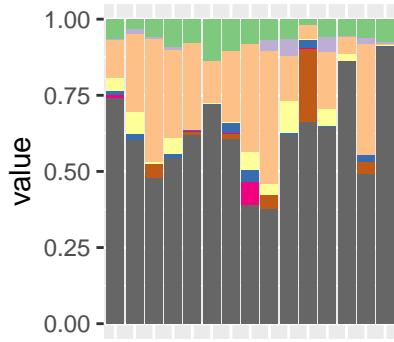
Early normalised



Signature

- SBS1
- SBS2
- SBS5
- SBS13
- SBS17a
- SBS17b
- SBS30
- SBS40

Late normalised



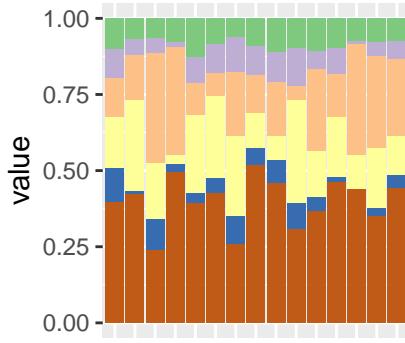
Signature

- SBS1
- SBS2
- SBS5
- SBS13
- SBS17a
- SBS17b
- SBS30
- SBS40

```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7167]
```

### SoftTissue–Liposarc

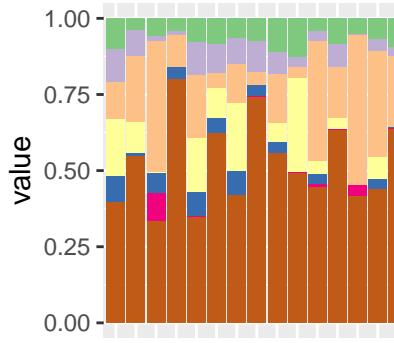
Early normalised



Signature

- SBS1
- SBS2
- SBS3
- SBS5
- SBS13
- SBS37
- SBS40

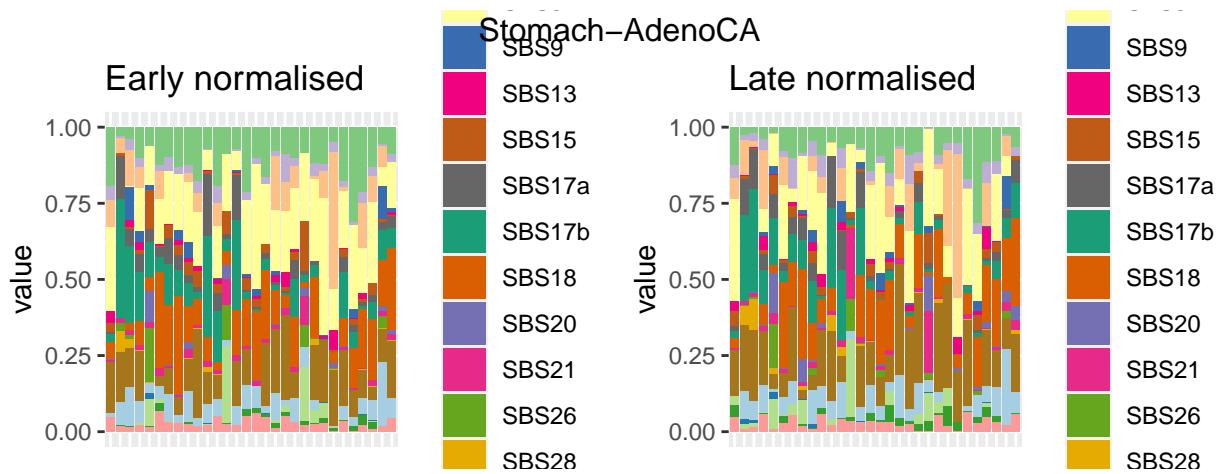
Late normalised



Signature

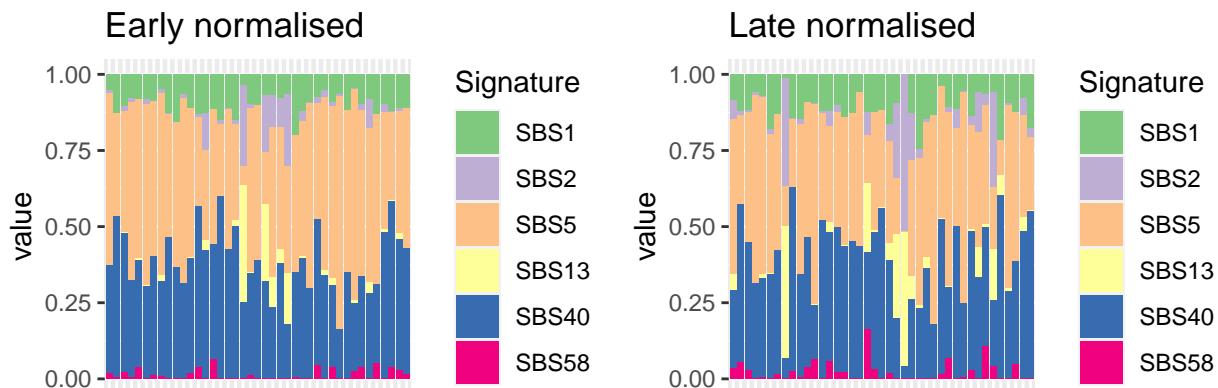
- SBS1
- SBS2
- SBS3
- SBS5
- SBS13
- SBS37
- SBS40

```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7342]
```

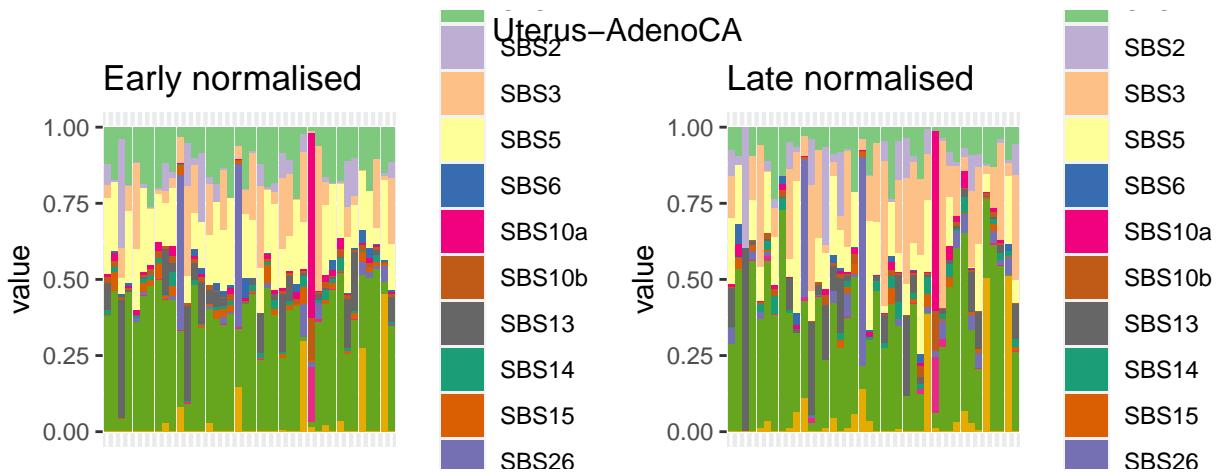


```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7699]
```

### Thy–AdenoCA



```
## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange    gtable[layout]
## 2 2 (2-2,2-2) arrange    gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7860]
```



```

## TableGrob (2 x 2) "arrange": 3 grobs
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange     gtable[layout]
## 2 2 (2-2,2-2) arrange     gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.8133]

## $`Biliary-AdenoCA`
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange     gtable[layout]
## 2 2 (2-2,2-2) arrange     gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.328]
## 
## $`Bladder-TCC`
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange     gtable[layout]
## 2 2 (2-2,2-2) arrange     gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.503]
## 
## $`Bone-Benign`
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange     gtable[layout]
## 2 2 (2-2,2-2) arrange     gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.678]
## 
## $`Bone-Epith`
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange     gtable[layout]
## 2 2 (2-2,2-2) arrange     gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.853]
## 
## $`Bone-Osteosarc`
##   z   cells   name      grob
## 1 1 (2-2,1-1) arrange     gtable[layout]
## 2 2 (2-2,2-2) arrange     gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1070]
## 
## $`Breast-AdenoCA`
```

```

##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1329]
##
## $`Breast-DCIS`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1448]
##
## $`Breast-LobularCA`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1609]
##
## $`Cervix-AdenoCA`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1742]
##
## $`Cervix-SCC`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.1903]
##
## $`CNS-GBM`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2064]
##
## $`CNS-Medullo`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2225]
##
## $`CNS-Oligo`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2358]
##
## $`CNS-PiloAstro`
##   z   cells   name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]

```

```

## 3 3 (1-1,1-2) arrange text[GRID.text.2505]
##
## $`ColoRect-AdenoCA`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2764]
##
## $`Eso-AdenoCA`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.2981]
##
## $`Head-SCC`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3268]
##
## $`Kidney-ChRCC`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3457]
##
## $`Kidney-RCC.clearcell`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3646]
##
## $`Kidney-RCC.papillary`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.3835]
##
## $`Liver-HCC`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4248]
##
## $`Lung-AdenoCA`
## z cells name grob
## 1 1 (2-2,1-1) arrange gtable[layout]
## 2 2 (2-2,2-2) arrange gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4493]
##
## $`Lung-SCC`

```

```

##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4654]
##
## $`Lymph-BNHL`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.4927]
##
## $`Lymph-CLL`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5060]
##
## $`Myeloid-AML`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5193]
##
## $`Myeloid-MPN`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5382]
##
## $`Ovary-AdenoCA`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5627]
##
## $`Panc-AdenoCA`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.5928]
##
## $`Panc-Endocrine`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6187]
##
## $`Prost-AdenoCA`
##   z    cells    name          grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]

```

```

## 3 3 (1-1,1-2) arrange text[GRID.text.6460]
##
## $`Skin-Melanoma.acral`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6719]
##
## $`Skin-Melanoma.cutaneous`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.6978]
##
## $`Skin-Melanoma.mucosal`
## [1] "Error : $ operator is invalid for atomic vectors\n"
## attr(,"class")
## [1] "try-error"
## attr(,"condition")
## <simpleError: $ operator is invalid for atomic vectors>
##
## $`SoftTissue-Leiomyo`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7167]
##
## $`SoftTissue-Liposarc`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7342]
##
## $`Stomach-AdenoCA`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7699]
##
## $`Thy-AdenoCA`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.7860]
##
## $`Uterus-AdenoCA`
##   z   cells   name           grob
## 1 1 (2-2,1-1) arrange      gtable[layout]
## 2 2 (2-2,2-2) arrange      gtable[layout]
## 3 3 (1-1,1-2) arrange text[GRID.text.8133]
##
## [1] ".../data/pcawg_robjcts_cache/tmb_results/nlminb/fullREDM"

```

```

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Breast-
## AdenoCA_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_CNS-
## GBM_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_CNS-
## Medullo_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_CNS-
## PiloAstro_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Eso-
## AdenoCA_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Kidney-
## RCC.clearcell_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Liver-
## HCC_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Lung-
## SCC_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Lymph-
## CLL_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Panc-
## AdenoCA_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

```

```

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Skin-
## Melanoma.cutaneous_signaturesPCAWG.RDS', probable reason 'No such file or
## directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMnoscalingnonexosubset_Thy-
## AdenoCA_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdab2nonexo_Bone-
## Osteosarc_signaturesPCAWG.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Breast-
## AdenoCA_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_CNS-
## GBM_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_CNS-
## Medullo_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_CNS-
## PiloAstro_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Eso-
## AdenoCA_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Kidney-
## RCC.clearcell_signaturesPCAWGSaA.RDS', probable reason 'No such file or
## directory'

## Error in gzfile(file, "rb") : cannot open the connection

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Liver-
## HCC_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'

## Error in gzfile(file, "rb") : cannot open the connection

```

```

## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Lung-
## SCC_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'
## Error in gzfile(file, "rb") : cannot open the connection
## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Lymph-
## CLL_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'
## Error in gzfile(file, "rb") : cannot open the connection
## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Panc-
## AdenoCA_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'
## Error in gzfile(file, "rb") : cannot open the connection
## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Skin-
## Melanoma.cutaneous_signaturesPCAWGSaA.RDS', probable reason 'No such file or
## directory'
## Error in gzfile(file, "rb") : cannot open the connection
## Warning in gzfile(file, "rb"): cannot open compressed file '../data/
## pcawg_robjetc_cache/tmb_results/nlminb/fullREDMonefixedlambdanonexo_Thy-
## AdenoCA_signaturesPCAWGSaA.RDS', probable reason 'No such file or directory'
## Error in gzfile(file, "rb") : cannot open the connection
pvals_diagRE_DMDL_SP <- sapply(diagRE_DMDL_SP, function(i) try(wald_TMB_wrapper(i)))
pvals_diagRE_DMDL_nonexo_SP <- sapply(diagRE_DMDL_nonexo_SP, function(i) try(wald_TMB_wrapper(i)))

## Check data - slope appears to be of length one (binomial)
pvals_fullRE_M_nonexo_SP <- sapply(fullRE_M_nonexo_SP, function(i) try(wald_TMB_wrapper(i)))

## Check data - slope appears to be of length one (binomial)
pvals_fullRE_DMSL_nonexo_SP <- sapply(fullRE_DMSL_nonexo_SP, function(i) try(wald_TMB_wrapper(i)))

## Check data - slope appears to be of length one (binomial)
pvals_fullREDMnoscaling_SP_nonexo_subsets_and_amalgamations <- sapply(fullREDMnoscaling_SP_nonexo_subsets_and_amalgamations, function(i) try(wald_TMB_wrapper(i)))

## Check data - slope appears to be of length one (binomial)
pvals_fullREDMonefixedlambdanonexo_SP <- sapply(fullREDMonefixedlambdanonexo_SP, function(i) try(wald_TMB_wrapper(i)))

## Check data - slope appears to be of length one (binomial)
pvals_fullREDMonefixedlambdanonexo_SPSaA <- sapply(fullREDMonefixedlambdanonexo_SPSaA, function(i) try(wald_TMB_wrapper(i)))

names(fullREDMonefixedlambdanonexo_SPSaA) <- names(pvals_fullREDMonefixedlambdanonexo_SP) <- names(pvals_fullREDMonefixedlambdanonexo_SP)
names(pvals_fullRE_DMSL_nonexo_SP) <- enough_samples

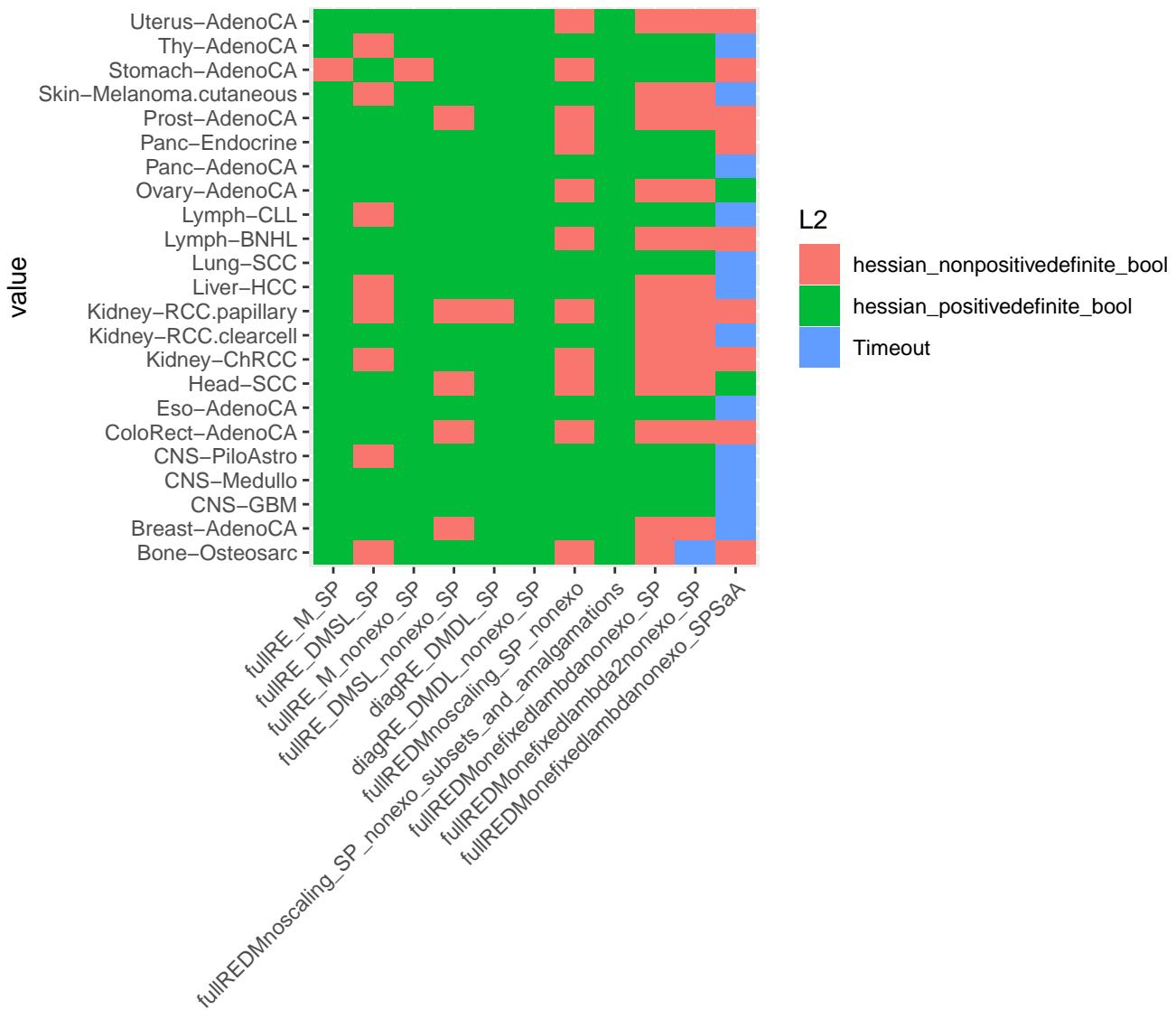
pvals_diagRE_DMDL_SP

```

```

##          Bone-Osteosarc      Breast-AdenoCA      CNS-GBM
##          1.080828e-04       2.239756e-28      3.390137e-03
##          CNS-Medullo        CNS-PiloAstro      ColoRect-AdenoCA
##          8.431463e-03       5.615238e-04      6.356131e-26
##          Eso-AdenoCA        Head-SCC          Kidney-ChRCC
##          5.329093e-21       4.975610e-05      1.562125e-09
##          Kidney-RCC.clearcell Kidney-RCC.papillary NA
##          4.027485e-18          NA                  Liver-HCC
##          Lung-SCC           Lymph-BNHL        Lymph-CLL
##          7.747310e-22       3.908637e-19      6.611927e-20
##          Ovary-AdenoCA       Panc-AdenoCA      Panc-Endocrine
##          8.965185e-38       4.096402e-119     3.987099e-10
##          Prost-AdenoCA      Skin-Melanoma.cutaneous Stomach-AdenoCA
##          6.474116e-99       9.272113e-25      1.715150e-06
##          Thy-AdenoCA        Uterus-AdenoCA
##          8.821583e-06       4.819867e-10
## $Timeout
## [1] "Breast-AdenoCA"      "CNS-GBM"
## [3] "CNS-Medullo"         "CNS-PiloAstro"
## [5] "Eso-AdenoCA"         "Kidney-RCC.clearcell"
## [7] "Liver-HCC"            "Lung-SCC"
## [9] "Lymph-CLL"            "Panc-AdenoCA"
## [11] "Skin-Melanoma.cutaneous" "Thy-AdenoCA"
##
## $hessian_nonpositivedefinite_bool
## [1] "Bone-Osteosarc"      "ColoRect-AdenoCA"      "Kidney-ChRCC"
## [4] "Kidney-RCC.papillary" "Lymph-BNHL"          "Panc-Endocrine"
## [7] "Prost-AdenoCA"        "Stomach-AdenoCA"      "Uterus-AdenoCA"
##
## $hessian_positivedefinite_bool
## [1] "Head-SCC"             "Ovary-AdenoCA"

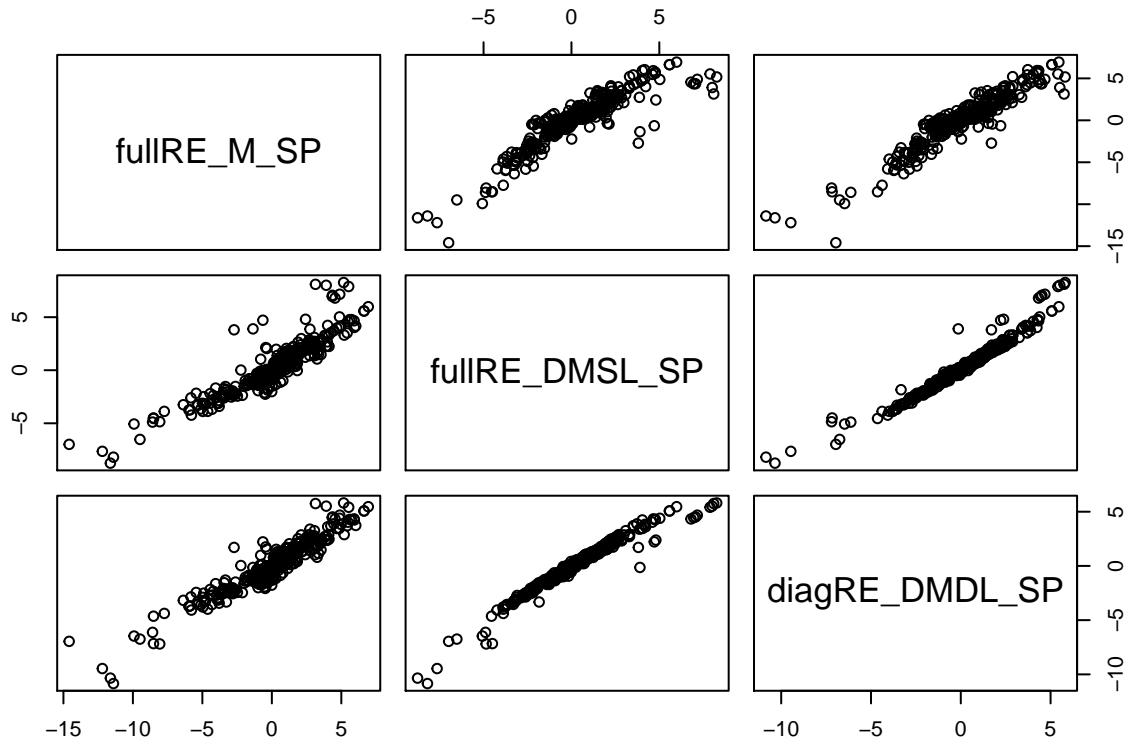
```



```

## comparison of betas
# give_betas(get(list_models_SP)[[1]][[1]])
all_betas_SP <- do.call('cbind', lapply(c( 'fullRE_M_SP', 'fullRE_DMSL_SP',
                                             'diagRE_DMDL_SP'), function(j) do.call('c', sapply(get(j), function(i) as.vector(give_b
colnames(all_betas_SP) <- c( 'fullRE_M_SP', 'fullRE_DMSL_SP',
                             'diagRE_DMDL_SP')
pairs(all_betas_SP)

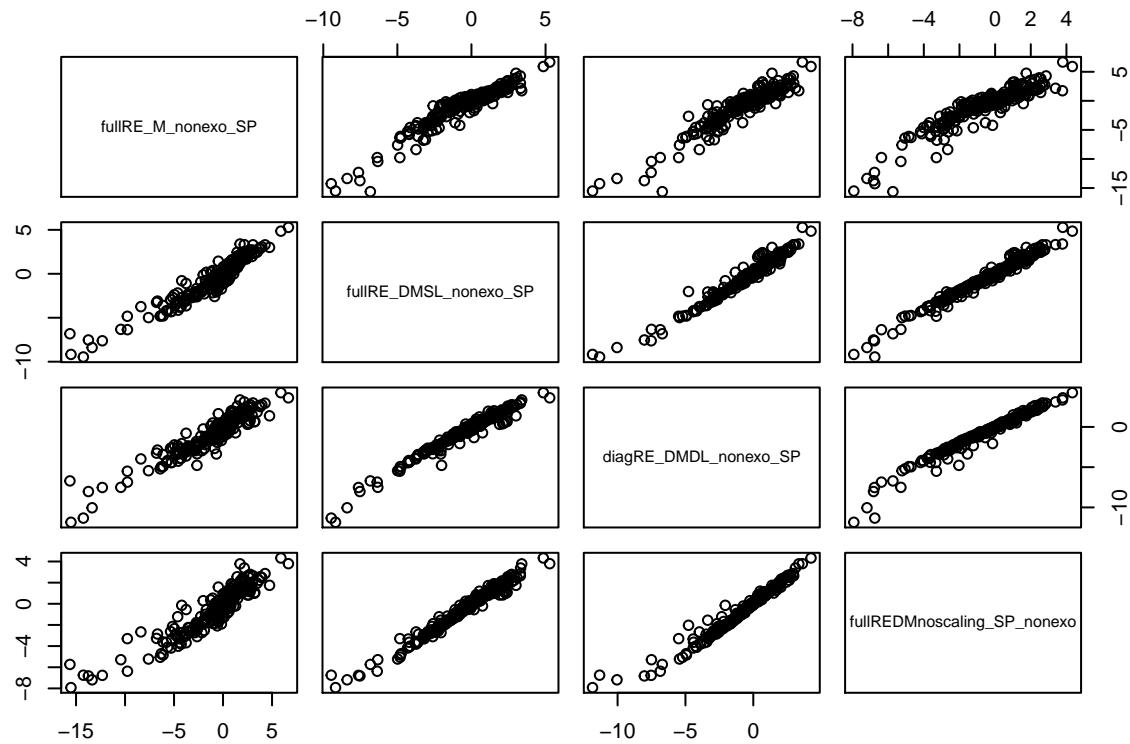
```



```

all_betas_SP_nonexo <- do.call('cbind', lapply(c('fullRE_M_nonexo_SP','fullRE_DMSL_nonexo_SP',
                                                 'diagRE_DMDL_nonexo_SP', 'fullREDMnoscaling_SP_nonexo'), function(j) do.call('c', sapply
colnames(all_betas_SP_nonexo) <- c('fullRE_M_nonexo_SP','fullRE_DMSL_nonexo_SP',
                                         'diagRE_DMDL_nonexo_SP', 'fullREDMnoscaling_SP_nonexo')
pairs(all_betas_SP_nonexo)

```



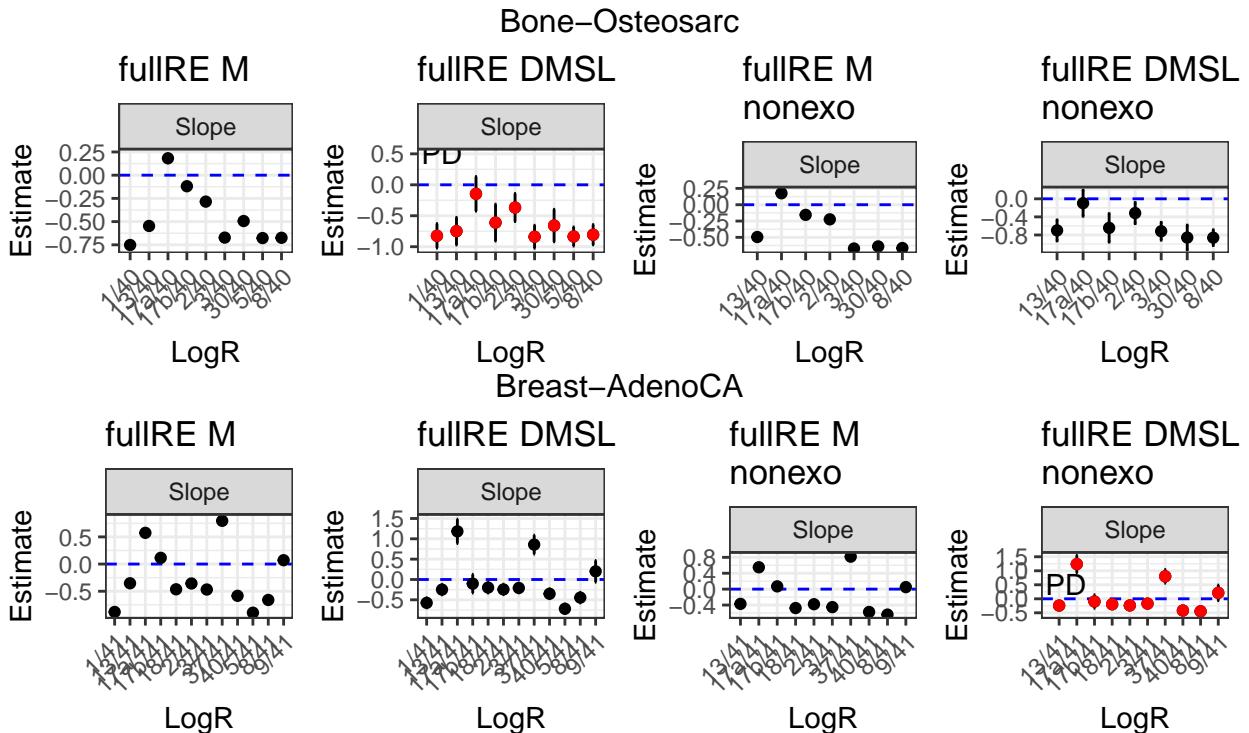
differences between signatures from sigprofiler from the paper, and the ones I get

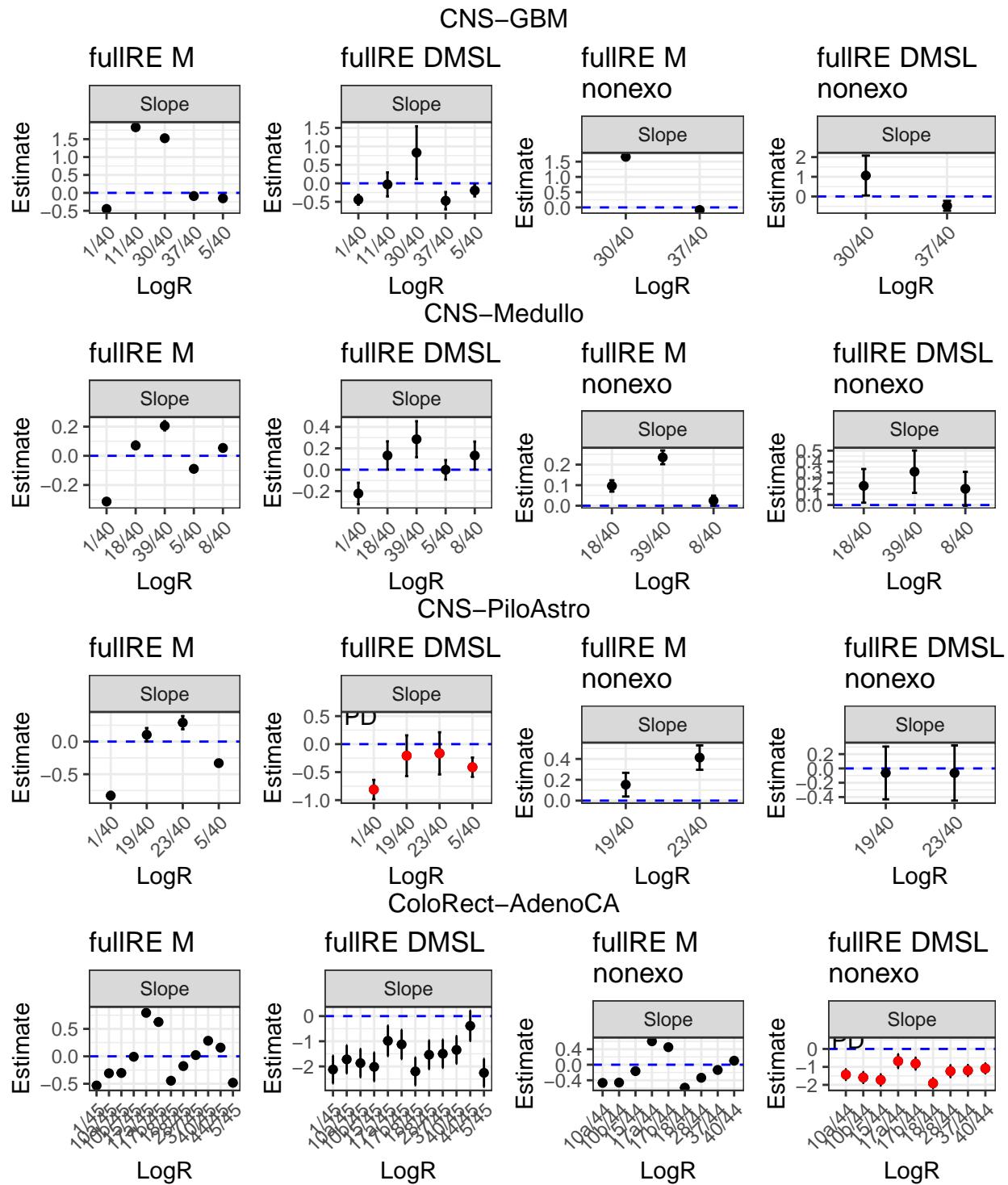
- biliary adenoca similar, though not exact
- bladder tcc: very similar
- bone benign: similar
- bone epith: extremely similar
- bone osteosarc: I have a lot of SBS8, which they don't. Other than that, similar
- breast adenocaral I have a lot of SBS9, which they don't. Other than that, similar
- breast DCIS: I have more SBS40 than they do
- breast lobularcarc: very similar
- cervix adenoca: very similar
- cervix SCC: very similar, although I have more SBS40
- CNS GBM: very similar, mine seem to be more homogeneous
- CNS medullo: very similar, mine seem to be more homogeneous
- CNS oligo: very similar, mine seem to be more homogeneous
- CNS piloastro: very similar, mine seem to be more homogeneous
- Colorect adenoca: quite similar
- eso adenoca: very similar
- head scc: very similar, mine seem to be more homogeneous
- kidney chrc: very similar
- kidney rcc.clearcell: very similar, although I have more SBS29
- kidney papillary: only I have it
- liver hcc: different. I have a lot of SBS40 and SBS12, they have mostly SBS5 \*\*\*
- lung adenoca: very similar
- lung SCC: very similar, though I have more SBS8
- lymph BNHL: very similar
- Lymph CLL: very similar, althpugh theirs are much more sparse
- myeloid AML: very similar, although I don't have any SBS60 and they seem to have

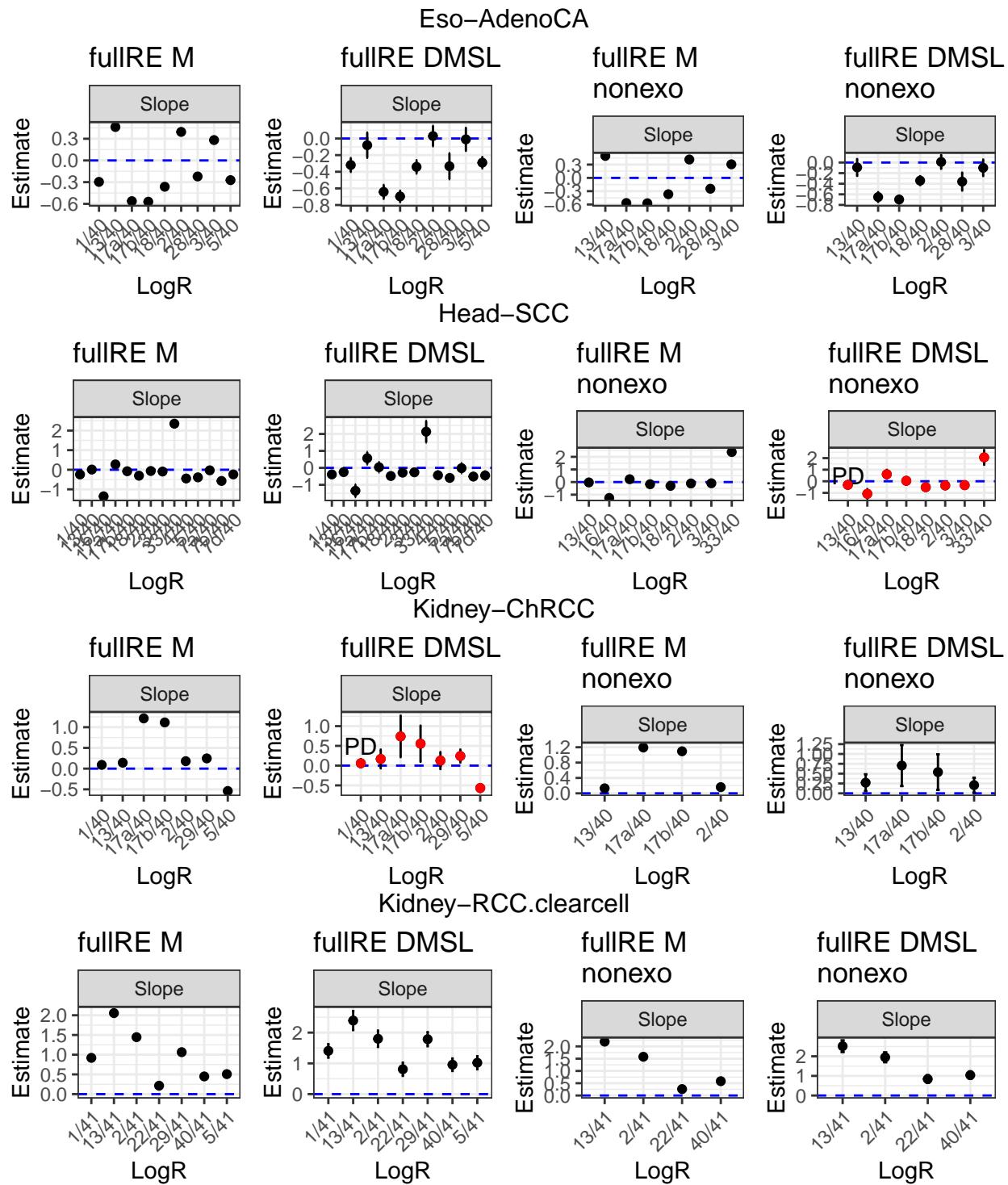
- myeloid MDS: I don't have it
- myeloid MPN: similar, although mine is much more sparse
- ovary adenoca: different. I have a lot of SBS40, which in their case is rare, and they have much more of SBS3 than I do \*\*\*
- panc-adenoca: different. I have a lot of SBS8 that they don't have. \*\*\*
- panc-endocrine: sort of similar. I have more SBS8 than they and they have more SBS5
- Prost-adenoca: sort of similar, I have more SBS8
- skin-melanoma.acral: they don't have this category. They have "skin-melanoma", which might be both together? (!!!) Similar exposures...
- softtissue-leiomyo: very similar exposures
- softtissue-liposarc: very similar exposures
- stomach adenoca: very similar, mine seem to be more homogeneous
- thy-adenoca: very similar
- uterus-adenoca: very similar

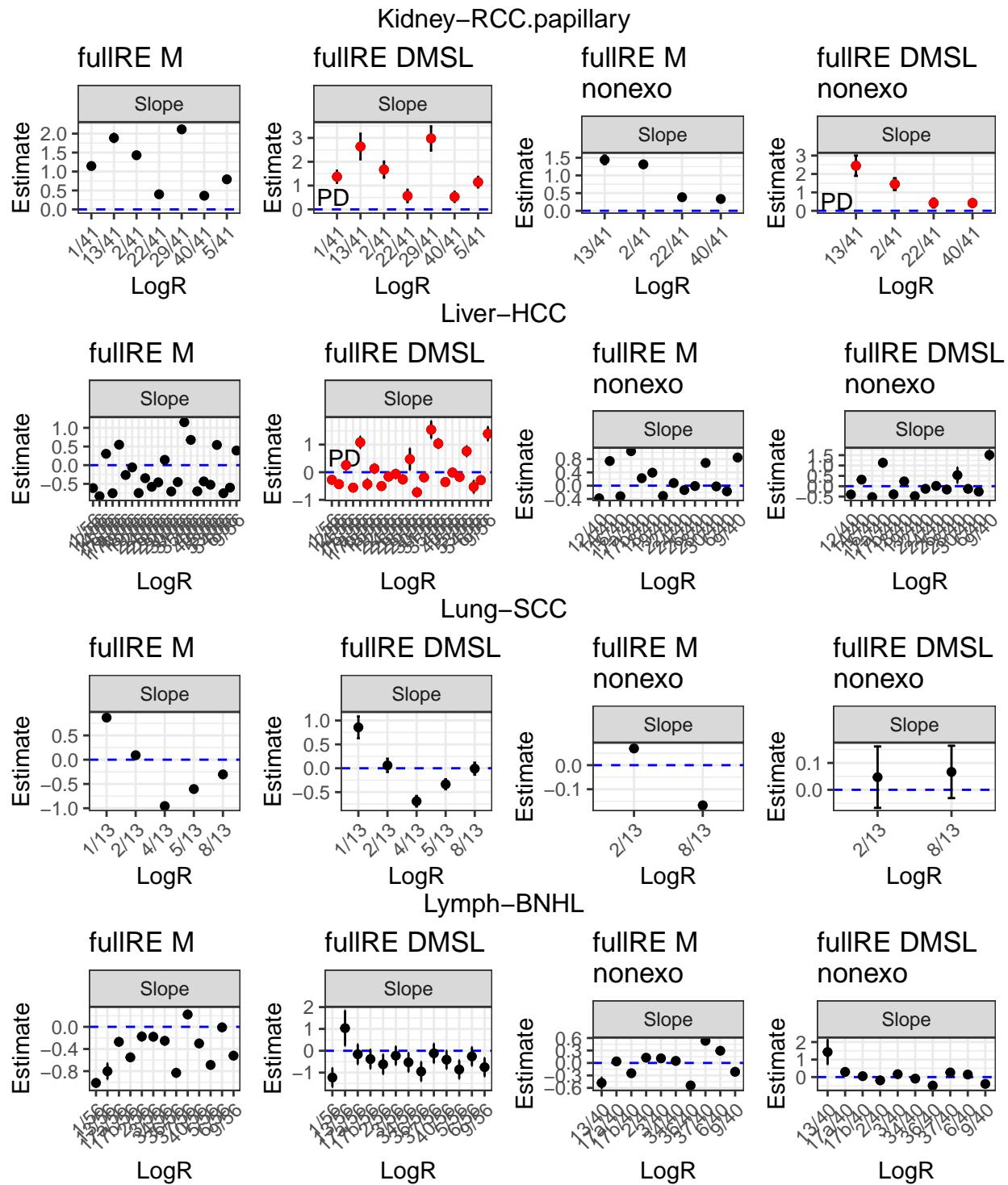
## Betas from PCAWG subset of signatures

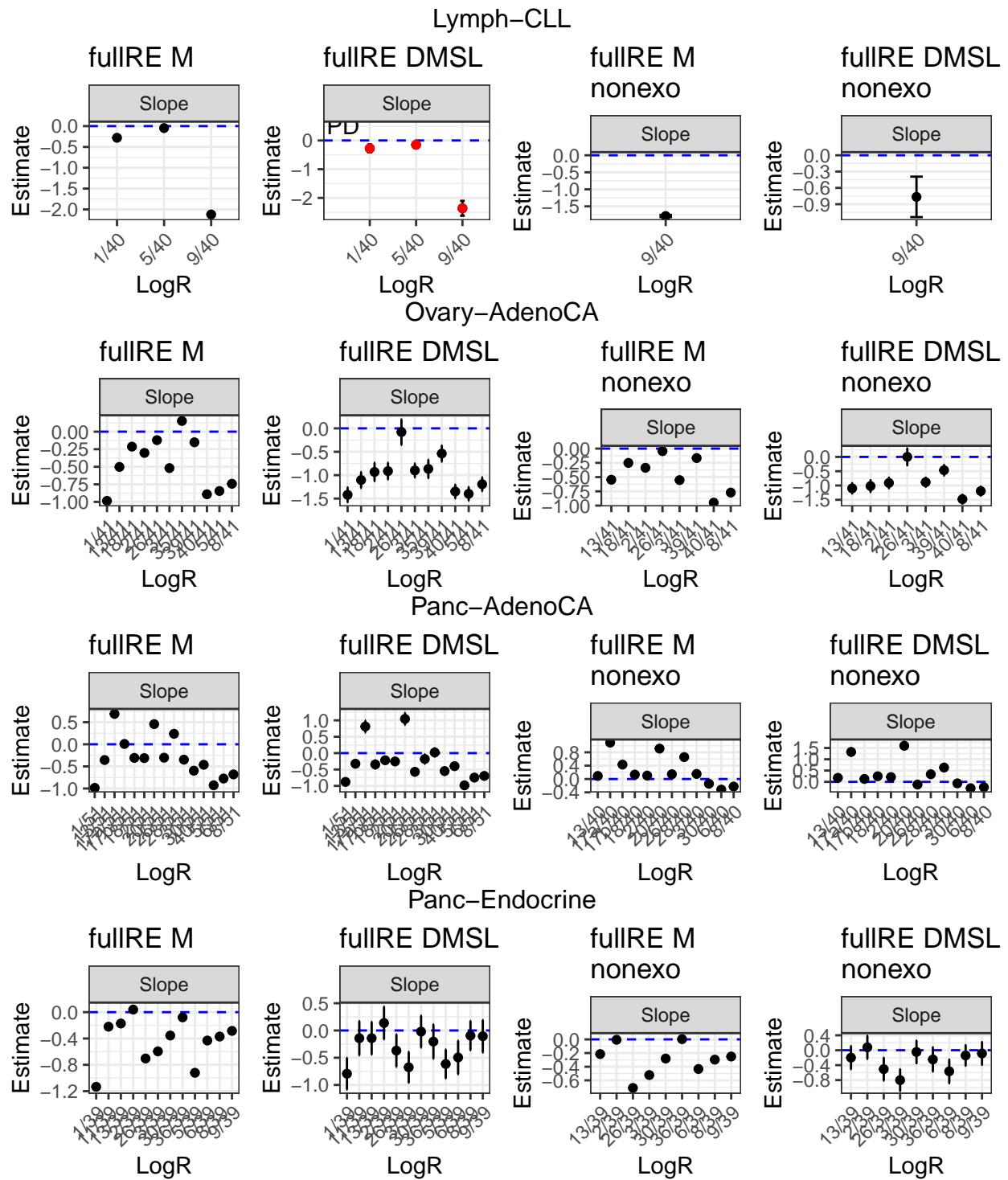
Reminder that the clonesig signatures were a subset of WES data.



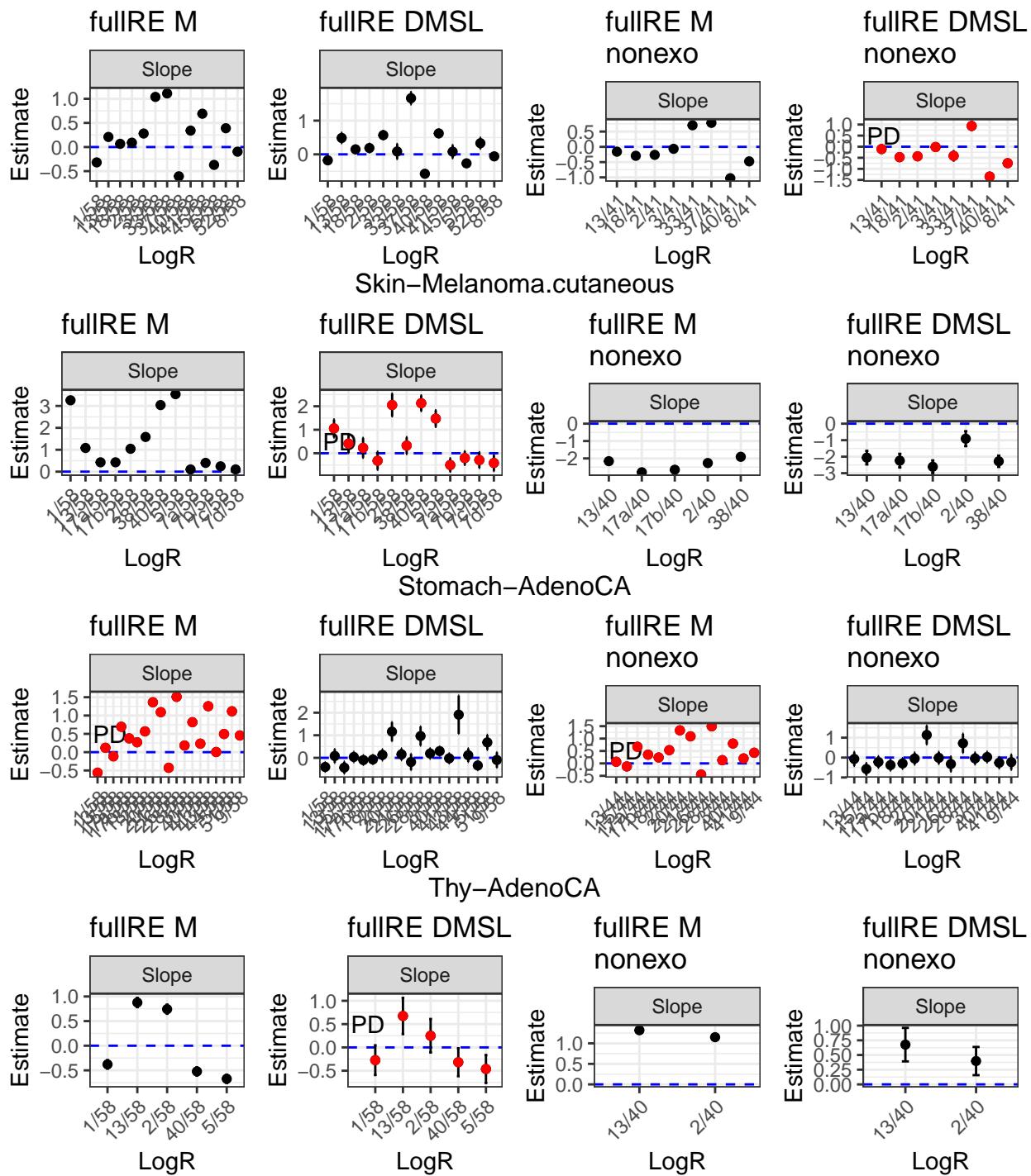


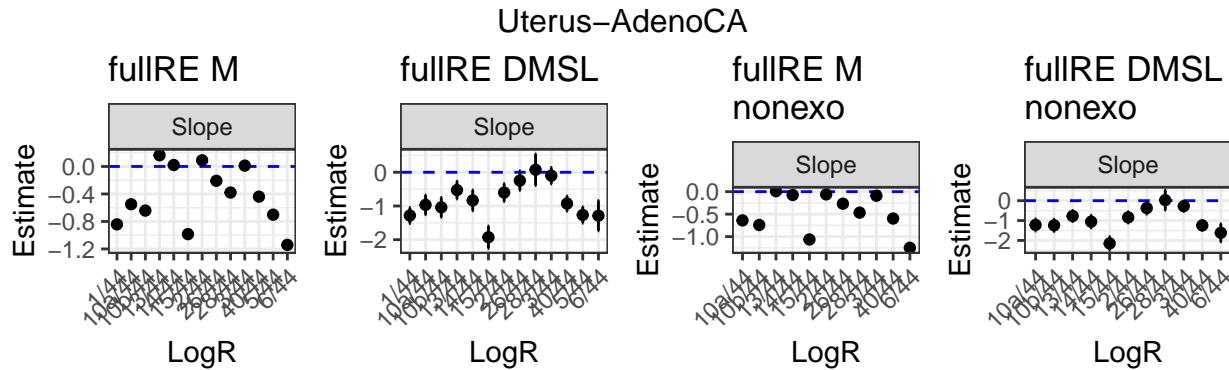






### Prost–AdenoCA



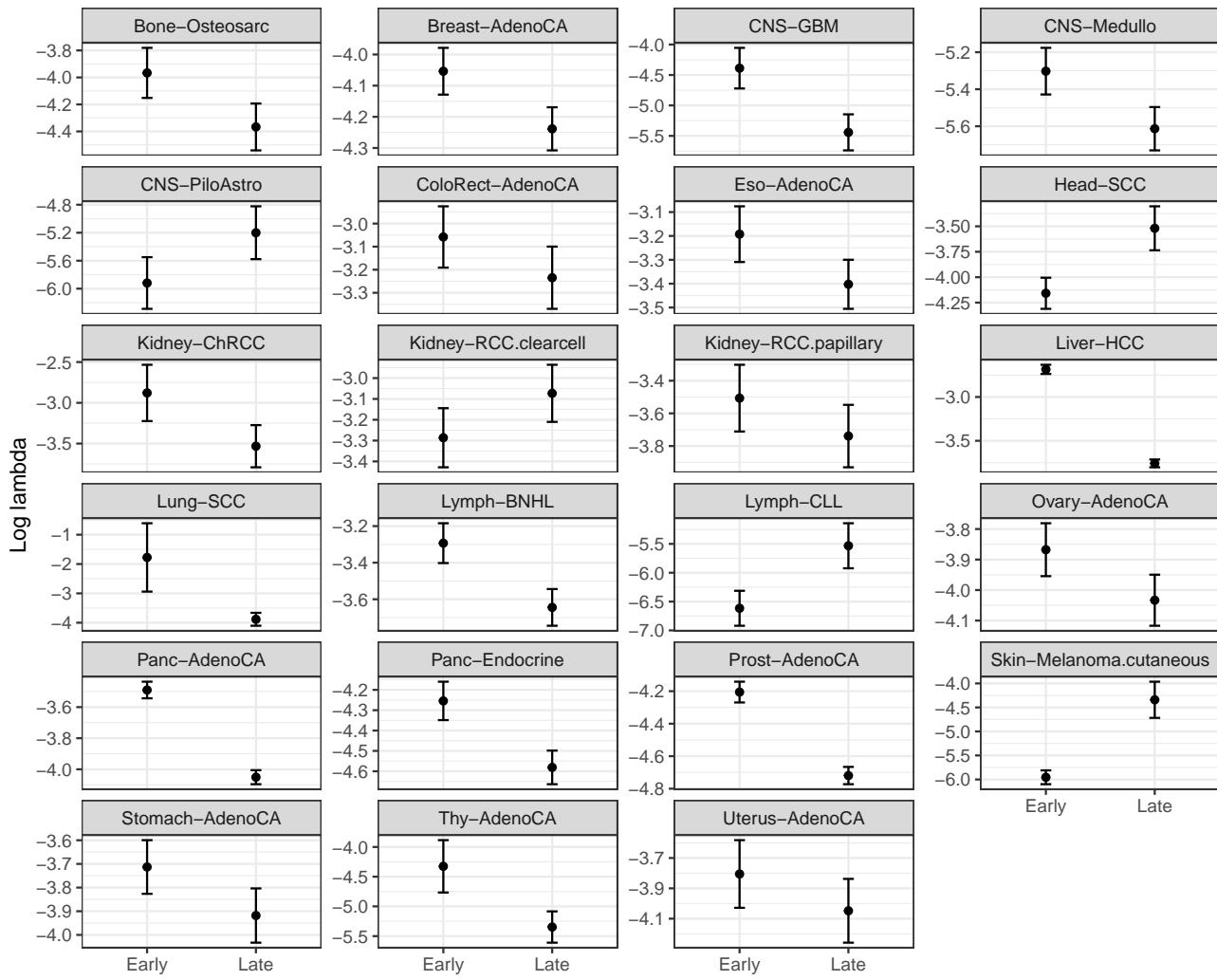


```
## pdf
## 2
```

## Overdispersion parameters in double-lambda DM

```
ovrdisp <- do.call('rbind.data.frame', lapply(1:length(diagRE_DMDL_nonexo_SP), try(function(idx){
  if(diagRE_DMDL_nonexo_SP[[idx]]$pdHess){
    cbind.data.frame( plot_lambdas(diagRE_DMDL_nonexo_SP[[idx]], return_df=T, plot=F), ct=names(diagRE_DM
  }else{
    c(NA, NA)
  }
})))
ovrdisp[ovrdisp$name == 'Lambda 1','name'] = 'Early'
ovrdisp[ovrdisp$name == 'Lambda 2','name'] = 'Late'

ggplot(ovrdisp, aes(x=name, y='Estimate'))+
  geom_point()+
  geom_errorbar(aes(ymin='Estimate'-`Std..Error`,
                     ymax='Estimate'+`Std..Error`), width=.1)+
  theme_bw()+
  facet_wrap(~ct, scales = "free_y", nrow=6)+
  labs(x='', y='Log lambda')
```



Test for differential precision (1/overdispersion) parameter

```
differential_precision <- p.adjust(sapply(diagRE_DMDL_nonexo_SP, wald_TMB_wrapper_overdisp), method = 'fdr')
names(differential_precision) <- names(diagRE_DMDL_nonexo_SP)
sort(differential_precision)
```

##	Liver-HCC	Panc-AdenoCA	Skin-Melanoma.cutaneous
##	5.160159e-58	7.096293e-18	9.923539e-09
##	Prost-AdenoCA	Lung-SCC	Head-SCC
##	1.219769e-07	1.139987e-05	8.968677e-02
##	Panc-Endocrine	CNS-GBM	Lymph-CLL
##	8.968677e-02	1.094428e-01	1.257732e-01
##	Thy-AdenoCA	Lymph-BNHL	CNS-Medullo
##	1.332616e-01	1.707526e-01	2.670091e-01
##	Breast-AdenoCA	Kidney-ChRCC	CNS-PiloAstro
##	3.645074e-01	3.883654e-01	4.524310e-01
##	Stomach-AdenoCA	Uterus-AdenoCA	Bone-Osteosarc
##	4.524310e-01	4.524310e-01	4.745719e-01
##	Eso-AdenoCA	Ovary-AdenoCA	Kidney-RCC.clearcell
##	4.815932e-01	4.979627e-01	5.431631e-01

```

##          ColoRect-AdenoCA      Kidney-RCC.papillary
##          6.589716e-01      6.806308e-01


|    | Liver-HCC     | Lung-SCC                | Panc-AdenoCA |
|----|---------------|-------------------------|--------------|
| ## | 5.160159e-58  | 1.139987e-05            | 7.096293e-18 |
| ## | Prost-AdenoCA | Skin-Melanoma.cutaneous |              |
| ## | 1.219769e-07  | 9.923539e-09            |              |



```

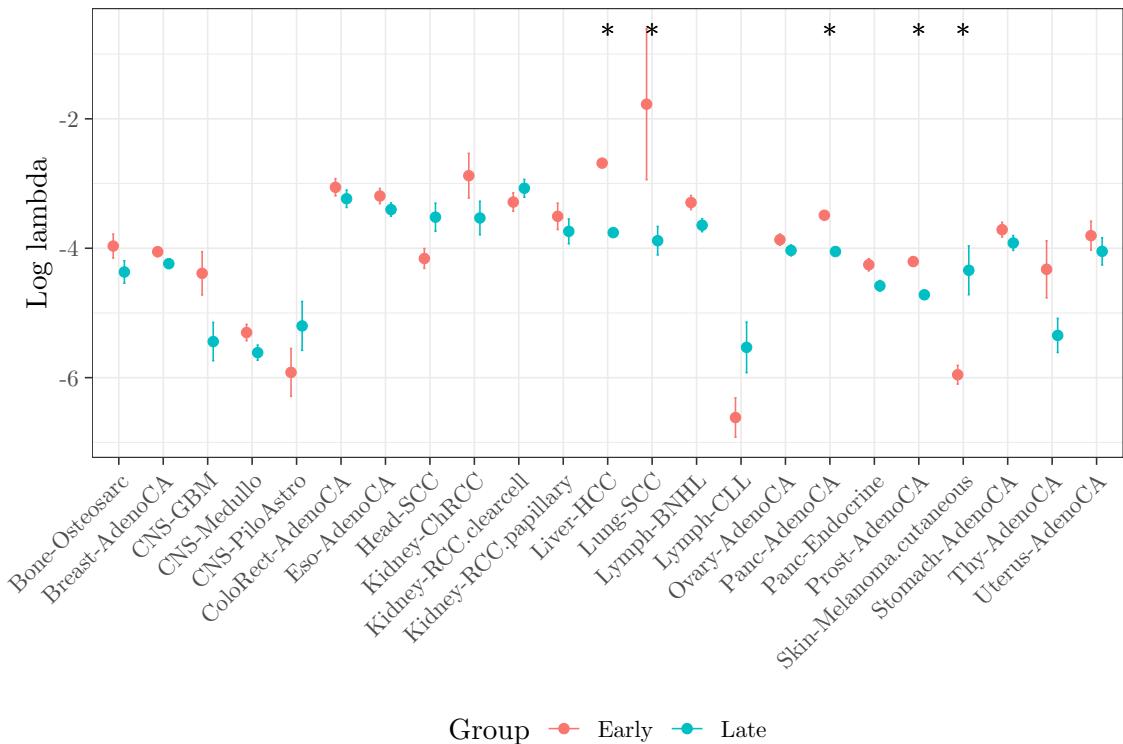
ovrdisp$differentially_abundant = ifelse(ovrdisp$ct %in% names(differential_precision[(differential_precision <= 0.05)]), differential_precision[(differential_precision <= 0.05)], 0)
ovrdisp$differentially_abundant
```



```

## [1] ""
## [20] "*"
## [39] "*"
ggplot(ovrdisp, aes(x=ct, y=Estimate, group=name, col=name))+
  geom_point(position=position_dodge(width=0.5))+
  geom_errorbar(aes(ymin=Estimate-Std..Error,
                     ymax=Estimate+Std..Error), width=.1, position=position_dodge(width=0.5))+
  theme_bw()+
  geom_text(aes(y=Inf, label=differentially_abundant, vjust=1.8), col='black')+
  theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+
  labs(x='', y='Log lambda', col='Group')+theme(legend.position = "bottom")+
  theme(
    legend.margin=margin(0,0,0,0),
    legend.box.margin=margin(-10,-10,-10,-10),
    plot.margin = unit(c(1,1,1,1), "cm"))
```


```



Group —●— Early —●— Late

Test for differential precision (1/overdispersion) parameter

```
differential_precision_2 <- p.adjust(sapply(diagRE_DMDL_nonexo_SP, ttest_TMB_wrapper_overdisp), method = names(differential_precision_2) <- names(diagRE_DMDL_nonexo_SP)
sort(differential_precision_2)
```

##	Liver-HCC	Panc-AdenoCA	Prost-AdenoCA
##	8.145123e-24	2.860615e-07	1.126237e-04
##	Skin-Melanoma.cutaneous	CNS-GBM	Head-SCC
##	1.723631e-02	2.834265e-01	2.834265e-01
##	Lymph-BNHL	Panc-Endocrine	Lung-SCC
##	2.834265e-01	2.834265e-01	3.053481e-01
##	Lymph-CLL	Thy-AdenoCA	Breast-AdenoCA
##	3.053481e-01	3.147768e-01	3.578140e-01
##	CNS-Medullo	Bone-Osteosarc	Kidney-ChRCC
##	3.578140e-01	4.331744e-01	4.331744e-01
##	CNS-PiloAstro	Eso-AdenoCA	Ovary-AdenoCA
##	4.352908e-01	4.352908e-01	4.352908e-01
##	Stomach-AdenoCA	Kidney-RCC.clearcell	ColoRect-AdenoCA
##	4.502790e-01	5.128995e-01	5.601231e-01
##	Kidney-RCC.papillary	Uterus-AdenoCA	
##	5.770837e-01	5.770837e-01	

```
table(differential_precision_2 <= 0.05)
```

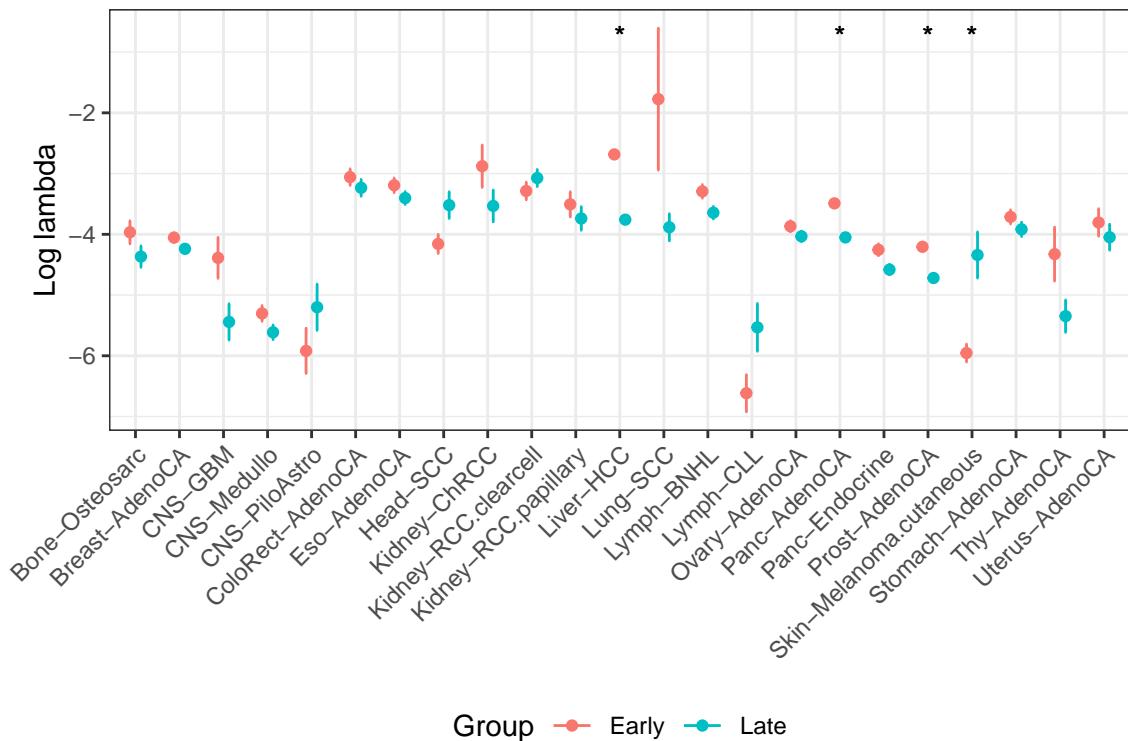
```
##
## FALSE TRUE
## 19 4
```

```
differential_precision_2[(differential_precision_2 <= 0.05)]
```

```
##          Liver-HCC      Panc-AdenoCA      Prost-AdenoCA
## 8.145123e-24 2.860615e-07 1.126237e-04
## Skin-Melanoma.cutaneous
## 1.723631e-02

ovrdisp$differential_precision_2 = ifelse(ovrdisp$ct %in% names(differential_precision_2)[(differential_precision_2 <= 0.05)], differential_precision_2, 0)

ggplot(ovrdisp, aes(x=ct, y=Estimate, group=name, col=name))+
  geom_point(position=position_dodge(width=0.5))+
  geom_errorbar(aes(ymin=Estimate-Std..Error,
                     ymax=Estimate+Std..Error), width=.1, position=position_dodge(width=0.5))+
  theme_bw()+
  geom_text(aes(y=Inf, label=differential_precision_2, vjust=1.8), col='black')+
  theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+
  labs(x='', y='Log lambda', col='Group')+theme(legend.position = "bottom")+
  theme(
    legend.margin=margin(0,0,0,0),
    legend.box.margin=margin(-10,-10,-10,-10),
    plot.margin = unit(c(1,1,1,1), "cm"))
```



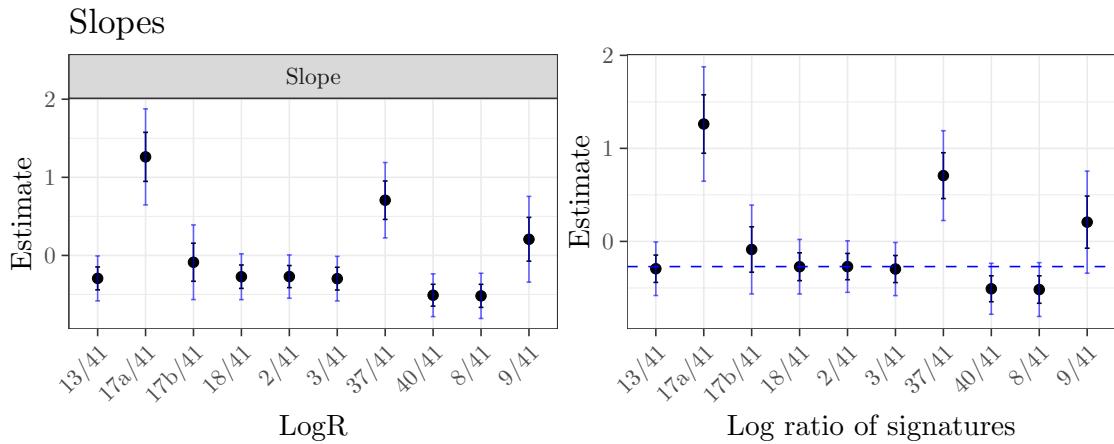
## Minimal perturbation

Note: this is using the original, non SP, signatures (hence not the best version).

```
minimalpert_L2 <- function(i){
  sum(i)/sum(i^2)
}
```

```

#
# betas_breast <- data.frame(plot_betas(diagRE_DMSL_nonexo[["Breast-AdenoCA"]], names_cats= logR_nonexo_r
#                               return_df=T, plot=F))
#
# .slopes_minpert <- betas_breast %>% dplyr::filter(type_beta == "Slope") %>% dplyr::select(Estimate) %>%
# #
# # minimalpert_L2(softmax(c(.slopes_minpert, 0)))
# median(c(.slopes_minpert, 0))
#
# aa <- plot_betas(diagRE_DMSL_nonexo[["Breast-AdenoCA"]], names_cats= logR_nonexo_notsorted[["Breast-AdenoCA"]]
#                   return_df=F, plot=F, only_slope = T, line_zero=F)
# # aa <- geom_hline(yintercept = 0)+geom_vline(xintercept = 1)+geom_hline(yintercept = median(c(.slopes_minpert, 0)))
# aa + geom_hline(yintercept = median(c(.slopes_minpert, 0)), lty='dashed', col='blue')+geom_hline(yintercept = mean(c(.slopes_minpert, 0)), lty='dashed', col='red')
# 
```



For the thesis:

```

pdf("../results/results_TMB/pcawg/minimalperturbation_SP_all.pdf", width = 4, height = 3)
for(ct in names(diagRE_DMDL_nonexo_SP)){
  .betas_ct_it <- data.frame(plot_betas(TMB_obj = diagRE_DMDL_nonexo_SP[[ct]],
                                         names_cats= logR_nonexo_notsorted_SP[[ct]],
                                         return_df=T, plot=F))
  .slopes_minpert <- .betas_ct_it %>% dplyr::filter(type_beta == "Slope") %>% dplyr::select(Estimate) %>%
  print(aaa <- plot_betas(TMB_obj = diagRE_DMDL_nonexo_SP[[ct]], names_cats= logR_nonexo_notsorted_SP[[ct]],
                           return_df=F, plot=F, only_slope = T, line_zero=F, add_confint = T, return_ggplot=TRUE)
  +
  geom_hline(yintercept = median(c(.slopes_minpert, 0)), lty='dashed', col='blue')+ggtitle(ct))
}
dev.off()

## pdf
## 2

\subsection{Minimal perturbation in diagRE_DMDL_nonexo_S}

perturbed_betas_diagRE_DMDL_nonexo_SP <- lapply(names(diagRE_DMDL_nonexo_SP), try(function(idx_sp){
  .betas_SP <- data.frame(plot_betas(diagRE_DMDL_nonexo_SP[[idx_sp]], names_cats= logR_nonexo_notsorted_SP[[idx_sp]]),
                            return_df=T, plot=F))

  .slopes_minpert_SP <- .betas_SP %>% dplyr::filter(type_beta == "Slope") %>% dplyr::select(Estimate) %>%
  
```

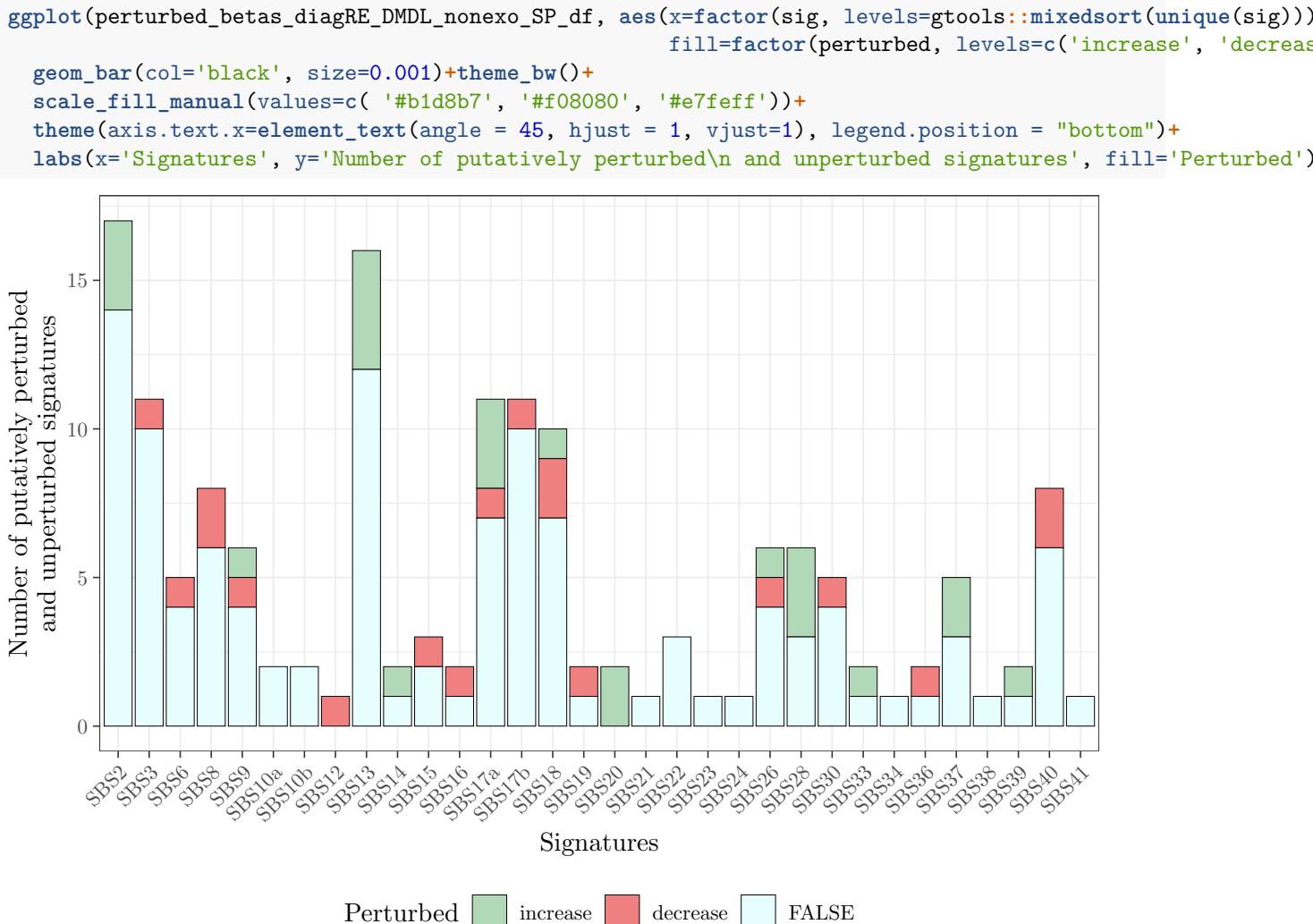
```

# print(.slopes_minpert_SP)
## check if the CI of the betas touches this median value
.summary_betas_slope_SP <- python_like_select_rownames(summary(diagRE_DMDL_nonexo_SP[[idx_sp]]), 'beta'
nrow(.summary_betas_slope_SP)

minimal_change_baseline <- median(c(.slopes_minpert_SP, 0))
# print(.summary_betas_slope_SP)
# print(logR_nonexo_notsorted_SP[[idx_sp]])
# print(dim(.summary_betas_slope_SP))
if(!is.null(dim(.summary_betas_slope_SP))){
  .params_in_ci <- give_params_in_CI(vec_est=.summary_betas_slope_SP[,1],
                                         vec_stderr=.summary_betas_slope_SP[,2],
                                         vec_true=rep(minimal_change_baseline, nrow(.summary_betas_slope_SP)))
} else{
  .params_in_ci <- give_params_in_CI(vec_est=.summary_betas_slope_SP[1],
                                         vec_stderr=.summary_betas_slope_SP[2],
                                         vec_true=minimal_change_baseline)
}
.params_in_ci <- sapply(1:length(.params_in_ci), function(i){
  ## for the ones in which there is a change, say whether it's up- or down-regulated
  if(!.params_in_ci[i]){
    ## if there is a change: not in confidence interval
    if(is.null(dim(.summary_betas_slope_SP))){ 
      ## one-dim
      if(.summary_betas_slope_SP[1] > minimal_change_baseline){
        'increase'
      } else{
        'decrease'
      }
    } else{
      ## multi-dim
      if(.summary_betas_slope_SP[i,1] > minimal_change_baseline){
        'increase'
      } else{
        'decrease'
      }
    }
  } else{
    'FALSE'
  }
})
names(.params_in_ci) <- sapply(logR_nonexo_notsorted_SP[[idx_sp]], function(i) strsplit(i, '/')[[1]][1]
.baseline <- strsplit(logR_nonexo_notsorted_SP[[idx_sp]][[1]], '/')[[1]][2]
return(list(betas_perturbed=.params_in_ci, baseline=.baseline))
}))}

perturbed_betas_diagRE_DMDL_nonexo_SP_vec <- do.call('c', sapply(perturbed_betas_diagRE_DMDL_nonexo_SP, f
perturbed_betas_diagRE_DMDL_nonexo_SP_df <- cbind.data.frame(sig=gsub("betas_perturbed.", "", 
  names(perturbed_betas_diagRE_DMDL_n
perturbed=perturbed_betas_diagRE_DMDL_nonexo_SP))

```



Minimal perturbation per signature

```

names(perturbed_betas_diagRE_DMDL_nonexo_SP) <- names(diagRE_DMDL_nonexo_SP)
ggplot(reshape2::melt(lapply(perturbed_betas_diagRE_DMDL_nonexo_SP, `[, 'betas_perturbed'])),
  aes(x=L1, fill=factor(value, levels=c('increase', 'decrease', 'FALSE')))+geom_bar(col='black', size=0.001)+scale_fill_manual(values=c( '#b1d8b7', '#f08080', '#e7feff'))+theme(legend.position = "bottom")+
  theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+labs(x='Cancer type',
  y='Number of putatively perturbed and unperturbed signatures', fill='Perturbed')

```

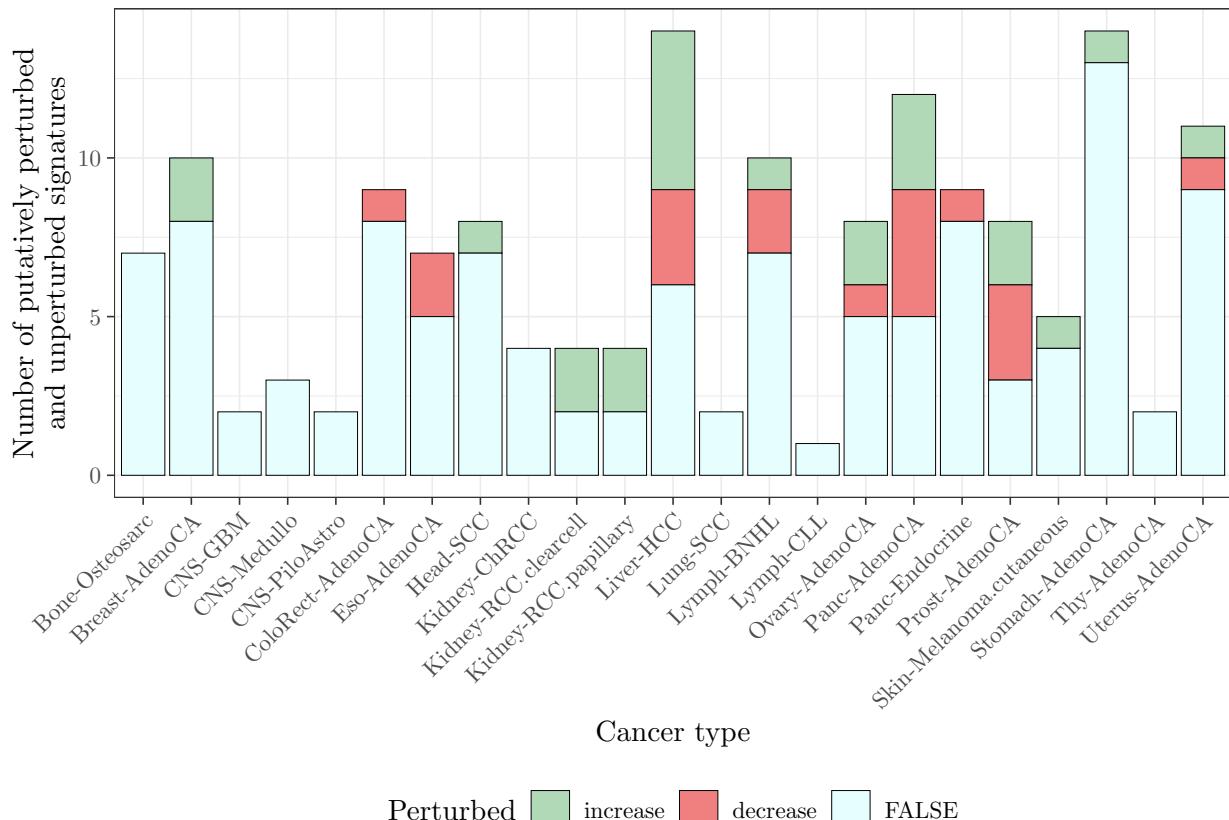


Table of perturbed signatures

```

relevel_perturbation <- cbind(c('increase', 'decrease'), c('+', '-'))

write("", ".../results/results_TMB/pcawg/minimal_perturbation_sigs.txt", append = F)
# for(i in names(perturbed_betas_diagRE_DMDL_nonexo_SP)){
#   write(paste0(i, '&', paste0(names(perturbed_betas_diagRE_DMDL_nonexo_SP[[i]]$betas_perturbed)[perturb
# })
for(i in names(perturbed_betas_diagRE_DMDL_nonexo_SP)){
  # write(x = paste0(i, '&', paste0(sapply(which(perturbed_betas_diagRE_DMDL_nonexo_SP[[i]]$betas_perturb
  #   paste0(names(perturbed_betas_diagRE_DMDL_nonexo_SP[[i]]$betas_perturbed)[j], '(', relevel_pertur
  #   }), collapse=' ', sep=''), '\n\n'), file=".../results/results_TMB/pcawg/minimal_perturbation_s
  write(x = paste0(paste0(i, '&', paste0(sapply(which(perturbed_betas_diagRE_DMDL_nonexo_SP[[i]]$betas_
    paste0(names(perturbed_betas_diagRE_DMDL_nonexo_SP[[i]]$betas_perturbed)[j], '(', relevel_perturb
  }), collapse=' ', sep=''), '\n\n')), file=".../results/results_TMB/pcawg/minimal_perturbation_s
}

## [1] 27 8
## [1] 27 8
## [1] 136 11
## [1] 136 11
## [1] 34 3
## [1] 34 3
## [1] 106 4
## [1] 106 4
## [1] 41 3

```

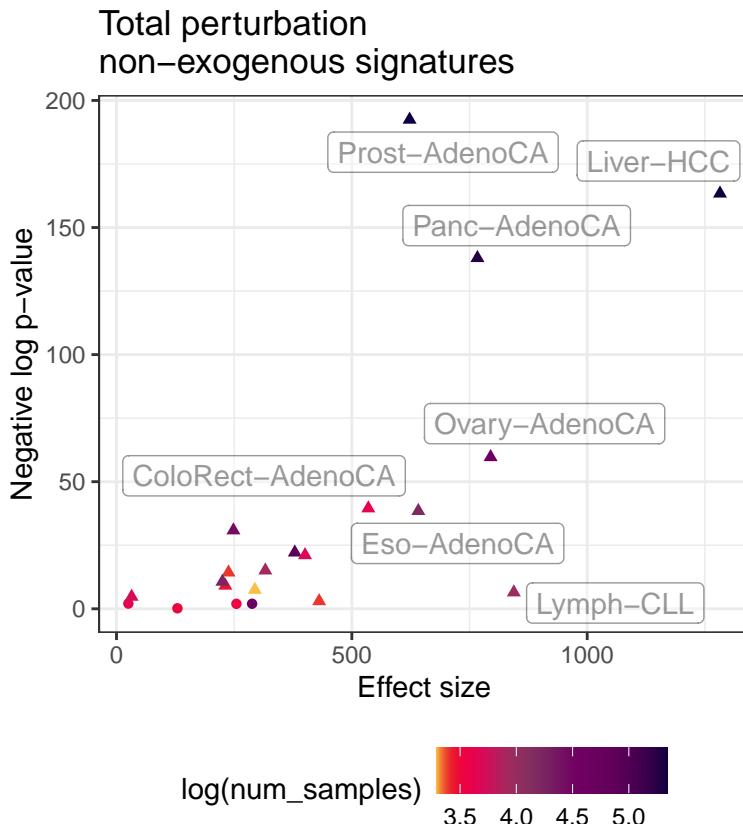
```

## [1] 41  3
## [1] 37 10
## [1] 37 10
## [1] 65  8
## [1] 65  8
## [1] 32  9
## [1] 32  9
## [1] 38  5
## [1] 38  5
## [1] 86  5
## [1] 86  5
## [1] 30  5
## [1] 30  5
## [1] 207 15
## [1] 207 15
## [1] 34  3
## [1] 34  3
## [1] 51 11
## [1] 51 11
## [1] 53  2
## [1] 53  2
## [1] 97  9
## [1] 97  9
## [1] 193 13
## [1] 193 13
## [1] 70 10
## [1] 70 10
## [1] 208  9
## [1] 208  9
## [1] 29  6
## [1] 29  6
## [1] 30 15
## [1] 30 15
## [1] 41  3
## [1] 41  3
## [1] 40 12
## [1] 40 12

##          Bone-Osteosarc      Breast-AdenoCA      CNS-GBM
##                      8                  11                  3
##          CNS-Medullo      CNS-PiloAstro      ColoRect-AdenoCA
##                      4                  3                  10
##          Eso-AdenoCA       Head-SCC        Kidney-ChRCC
##                      8                  9                  5
##          Kidney-RCC.clearcell Kidney-RCC.papillary Liver-HCC
##                      5                  5                  15
##          Lung-SCC           Lymph-BNHL       Lymph-CLL
##                      3                  11                  2
##          Ovary-AdenoCA      Panc-AdenoCA      Panc-Endocrine
##                      9                  13                  10
##          Prost-AdenoCA     Skin-Melanoma.cutaneous Stomach-AdenoCA
##                      9                  6                  15

```

```
## Thy-AdenoCA           Uterus-AdenoCA
##                         3                  12
## Error in apply(pert, 1, function(i) sqrt(sum((i - 1/(ncol(exposures_cancertype_obj$Y))))^2))) :
##   dim(X) must have a positive length
## Error in apply(pert, 1, function(i) sqrt(sum((i - 1/(ncol(exposures_cancertype_obj$Y))))^2))) :
##   dim(X) must have a positive length
## Warning: NAs introduced by coercion
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
## Warning: Removed 2 rows containing missing values (geom_point).
## Warning: Removed 2 rows containing missing values (geom_label_repel).
## Warning: ggrepel: 14 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



##	Bone-Osteosarc	Breast-AdenoCA	CNS-GBM
##	320.02704	469.29125	524.08517
##	CNS-Medullo	CNS-PiloAstro	ColoRect-AdenoCA
##	217.31161	27.02067	1031.46320
##	Eso-AdenoCA	Head-SCC	Kidney-ChRCC
##	522.66804	221.26052	103.16727
##	Kidney-RCC.clearcell	Kidney-RCC.papillary	Liver-HCC
##	247.53963	156.30958	1332.22216
##	Lung-SCC	Lymph-BNHL	Lymph-CLL

```

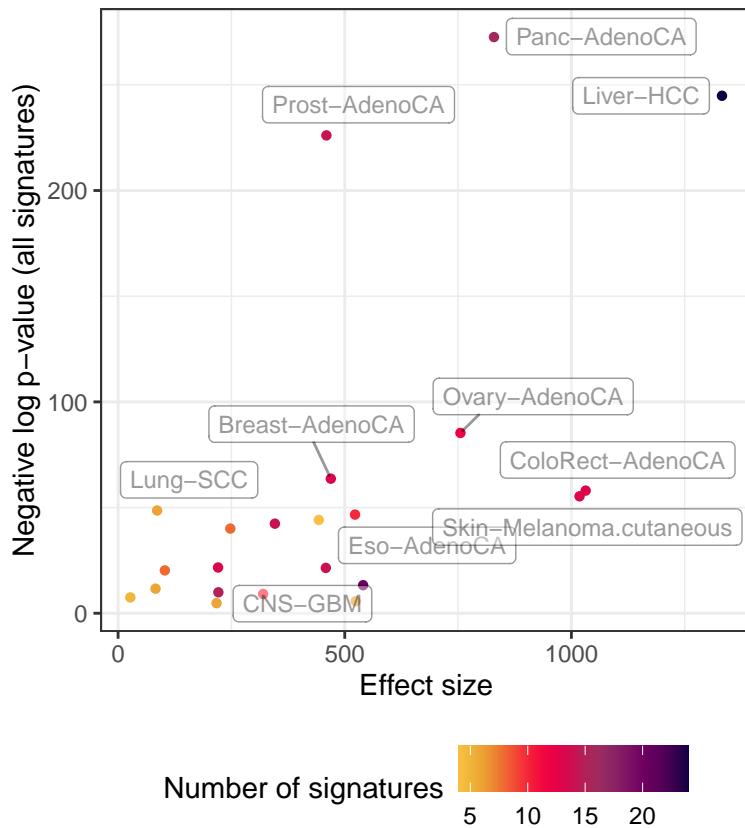
##          86.07143      345.88535      442.78303
## Ovary-AdenoCA      Panc-AdenoCA      Panc-Endocrine
##          755.21919      829.19711      220.55650
## Prost-AdenoCA Skin-Melanoma.cutaneous      Stomach-AdenoCA
##          459.55271      1017.96421      540.50633
## Thy-AdenoCA       Uterus-AdenoCA
##          82.57481      458.24059

## Bone-Osteosarc      Breast-AdenoCA      CNS-GBM
##          1.080828e-04      2.239756e-28      3.390137e-03
## CNS-Medullo      CNS-PiloAstro      ColoRect-AdenoCA
##          8.431463e-03      5.615238e-04      6.356131e-26
## Eso-AdenoCA       Head-SCC      Kidney-ChRCC
##          5.329093e-21      4.975610e-05      1.562125e-09
## Kidney-RCC.clearcell      Kidney-RCC.papillary      Liver-HCC
##          4.027485e-18      NA      4.747822e-107
## Lung-SCC      Lymph-BNHL      Lymph-CLL
##          7.747310e-22      3.908637e-19      6.611927e-20
## Ovary-AdenoCA      Panc-AdenoCA      Panc-Endocrine
##          8.965185e-38      4.096402e-119      3.987099e-10
## Prost-AdenoCA Skin-Melanoma.cutaneous      Stomach-AdenoCA
##          6.474116e-99      9.272113e-25      1.715150e-06
## Thy-AdenoCA       Uterus-AdenoCA
##          8.821583e-06      4.819867e-10

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_label_repel).
## Warning: ggrepel: 12 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps

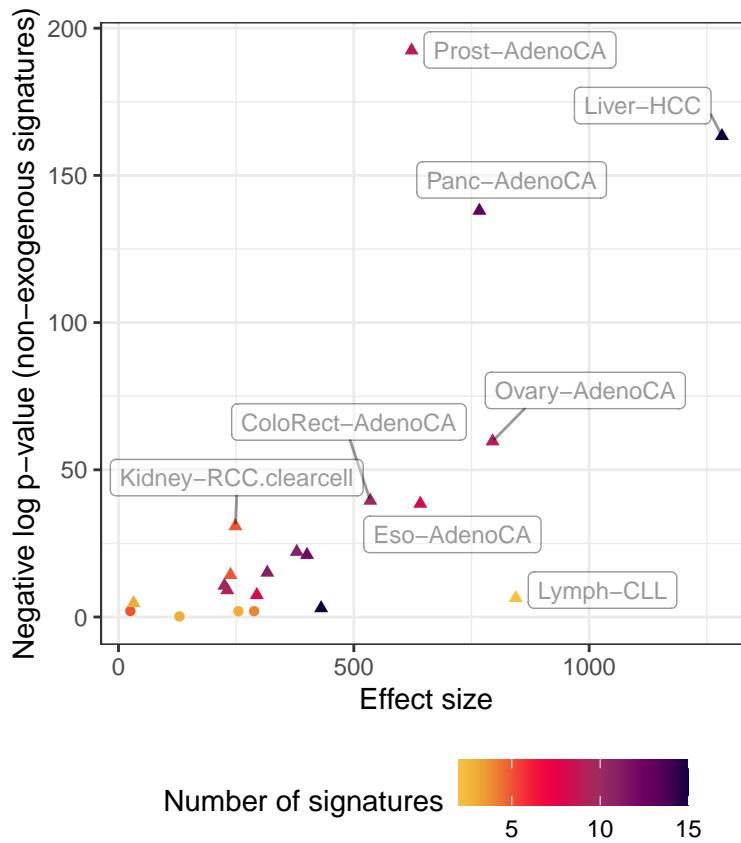
```



```

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
## Warning: Removed 2 rows containing missing values (geom_point).
## Warning: Removed 2 rows containing missing values (geom_label_repel).
## Warning: ggrepel: 13 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps

```



## HMP

Simple HMP tests to see what is differentially abundant

```

## Loading required package: HMP
## Loading required package: dirmult
##
## Attaching package: 'dirmult'
## The following object is masked from 'package:MCMCpack':
##     rdirichlet
##
## Attaching package: 'HMP'
## The following object is masked from 'package:dirmult':
##     weirMoM
pvals_diagRE_DMDL_SP_adj <- p.adjust(pvals_diagRE_DMDL_SP, 'fdr')
pvals_diagRE_DMDL_nonexo_SP_adj <- p.adjust(pvals_diagRE_DMDL_nonexo_SP, 'fdr')
pvals_fullREDMnoscaling_SP_nonexo_subsets_and_amalgamations_adj <- p.adjust(pvals_fullREDMnoscaling_SP_no
pvals_fullRE_M_nonexo_SP_adj <- p.adjust(pvals_fullRE_M_nonexo_SP, 'fdr')
pvals_fullRE_DMSL_nonexo_SP_adj <- p.adjust(pvals_fullRE_DMSL_nonexo_SP, 'fdr')
```

```

names(pvals_diagRE_DMDL_SP_adj) <- names(pvals_diagRE_DMDL_nonexo_SP_adj) <- names(pvals_fullRE_M_nonexo_)

table(diagRE_DMDL_DA=pvals_diagRE_DMDL_nonexo_SP_adj <= 0.05,
      fullRE_DMSL_DA=pvals_fullRE_DMSL_nonexo_SP_adj <= 0.05)

##          fullRE_DMSL_DA
## diagRE_DMDL_DA FALSE TRUE
##      FALSE      5     1
##      TRUE       1    11

names(pvals_diagRE_DMDL_nonexo_SP_adj)[which((pvals_diagRE_DMDL_nonexo_SP_adj <= 0.05) & (fullRE_DMSL_DA=0.05))

## [1] "Lymph-CLL"

names(pvals_diagRE_DMDL_nonexo_SP_adj)[which((pvals_diagRE_DMDL_nonexo_SP_adj > 0.05) & (fullRE_DMSL_DA=0.05))

## [1] "Stomach-AdenoCA"

table(diagRE_DMDL_DA=pvals_diagRE_DMDL_nonexo_SP_adj <= 0.05,
      fullRE_M_DA=pvals_fullRE_M_nonexo_SP_adj <= 0.05)

##          fullRE_M_DA
## diagRE_DMDL_DA TRUE
##      FALSE      5
##      TRUE     17

# table(diagRE_DMDL_DA=pvals_diagRE_DMDL_nonexo_SP_adj <= 0.05,
#       HMP_DA=HMP_res_sigs_pval_adj <= 0.05)

# table(diagRE_DMDL_DA=pvals_diagRE_DMDL_nonexo_SP_adj < 0.05, HMP_res_sigs_DA=HMP_res_sigs < 0.05)

p.adjust(pvals_diagRE_DMDL_SP, method = "BH")

##          Bone-Osteosarc      Breast-AdenoCA          CNS-GBM
##      1.251484e-04      9.854928e-28      3.551572e-03
##          CNS-Medullo      CNS-PiloAstro      ColoRect-AdenoCA
##      8.431463e-03      6.176762e-04      2.330581e-25
##          Eso-AdenoCA      Head-SCC      Kidney-ChRCC
##      1.302667e-20      6.081301e-05      2.291117e-09
##      Kidney-RCC.clearcell      Kidney-RCC.papillary      Liver-HCC
##      7.383723e-18                  NA      5.222605e-106
##          Lung-SCC      Lymph-BNHL      Lymph-CLL
##      2.130510e-21      7.817274e-19      1.454624e-19
##          Ovary-AdenoCA      Panc-AdenoCA      Panc-Endocrine
##      4.930852e-37      9.012085e-118      6.747398e-10
##          Prost-AdenoCA      Skin-Melanoma.cutaneous      Stomach-AdenoCA
##      4.747685e-98      2.914093e-24      2.358331e-06
##          Thy-AdenoCA      Uterus-AdenoCA
##      1.141617e-05      7.574077e-10

p.adjust(pvals_diagRE_DMDL_nonexo_SP, method = "BH")

##          Bone-Osteosarc      Breast-AdenoCA          CNS-GBM
##      8.712593e-04      6.123721e-10      1.557247e-01

```

```

##          CNS-Medullo      CNS-PiloAstro    ColoRect-AdenoCA
## 1.557247e-01      5.270878e-01   3.141298e-17
##          Eso-AdenoCA       Head-SCC        Kidney-ChRCC
## 7.328650e-17      1.788732e-04   1.557247e-01
## Kidney-RCC.clearcell Kidney-RCC.papillary   Liver-HCC
## 1.137607e-13      1.190455e-06   1.263498e-70
##          Lung-SCC         Lymph-BNHL     Lymph-CLL
## 8.301558e-01      5.881884e-07   2.256505e-03
##          Ovary-AdenoCA     Panc-AdenoCA   Panc-Endocrine
## 6.973340e-26      8.893577e-60   4.169125e-05
##          Prost-AdenoCA Skin-Melanoma.cutaneous Stomach-AdenoCA
## 6.044387e-83      9.919659e-17   6.320633e-02
##          Thy-AdenoCA        Uterus-AdenoCA
## 1.170750e-02      1.568961e-09

pcawg_palette <- pcawg.colour.palette(gsub("\\\\.*", "", enough_samples), scheme = "tumour.subtype")
names(pcawg_palette) <- enough_samples

df_pvals_DMDL_SP <- cbind.data.frame(pvals_DM=pvals_diagRE_DMDL_SP_adj,
                                         pvals_DMnonexo=pvals_diagRE_DMDL_nonexo_SP_adj,
                                         num_samples=as.numeric(num_samples_all_SP),
                                         num_sigs_nonexo=as.numeric(num_sigs_nonexo_SP),
                                         ct=enough_samples,
                                         pvals_DM_censored=sapply(-log(pvals_diagRE_DMDL_SP_adj),
                                         function(i) min(i, 25)),
                                         pvals_DMnonexo_censored=sapply(-log(pvals_diagRE_DMDL_nonexo_SP_adj),
                                         function(i) min(i, 25)),
                                         bool_censored=(( -log(pvals_diagRE_DMDL_nonexo_SP_adj) > 25 ) | ( -log(pvals_diagRE_
```

ggplot(df\_pvals\_DMDL\_SP,

  aes(x=pvals\_DM\_censored, y=pvals\_DMnonexo\_censored,

    # size=num\_samples,

    label=ct, size=bool\_censored))+geom\_point(aes (col=ct))+

  geom\_hline(yintercept = -log(0.05), lty='dashed')+geom\_vline(xintercept = -log(0.05), lty='dashed')+

  geom\_label\_repel(size=3.2, alpha=0.6, max.overlaps = 30)+ theme\_bw()+

  theme(legend.position = "bottom", legend.text=element\_text(size=8))+

  labs(x=' Log p-value all signatures', y=' Log p-value nonexogenous signatures')+

  guides(size=FALSE, col=FALSE)+ #, col=guide\_legend(ncol=4),

  scale\_color\_manual(values = pcawg\_palette)+theme(legend.position = "bottom")+

  lims(x=c(0, 30), y=c(0,30))

```

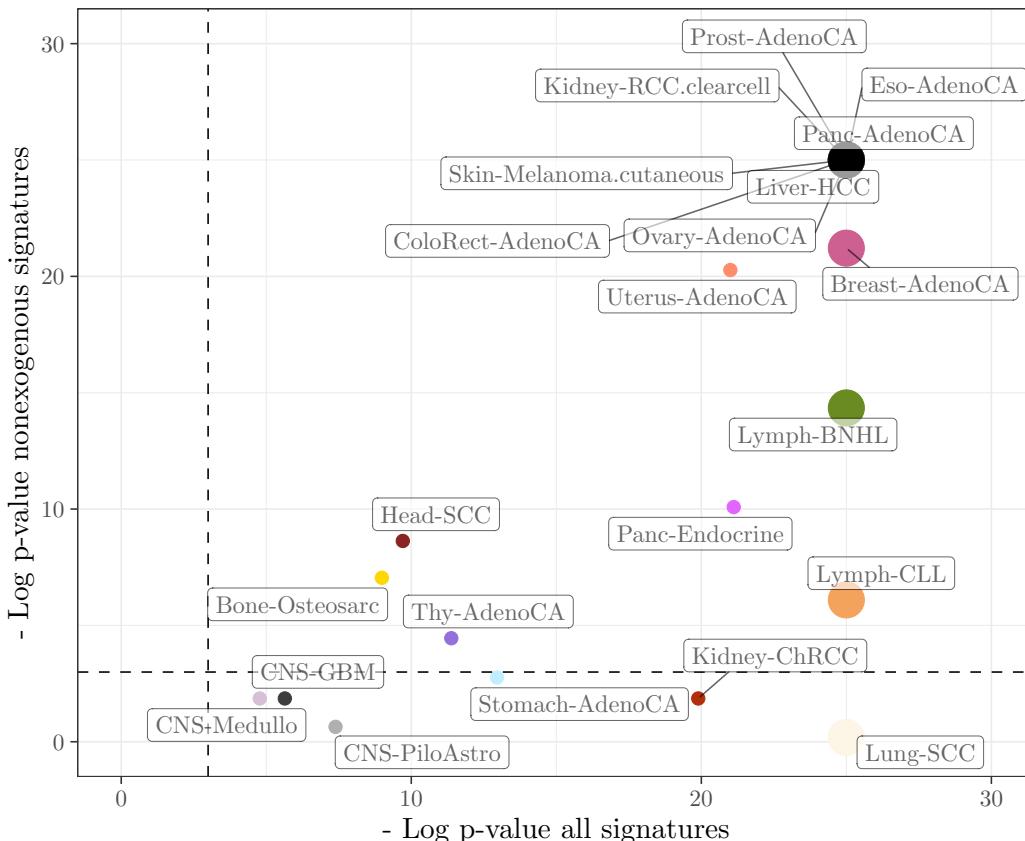
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

## Warning: Using size for a discrete variable is not advised.

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing missing values (geom_label_repel).

```



```
t.test(df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo <= 0.05, 'num_samples'],
       df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo > 0.05, 'num_samples'])
```

```
## 
## Welch Two Sample t-test
## 
## data: df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo <= 0.05, "num_samples"] and df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo > 0.05, "num_samples"]
## t = 1.8016, df = 19.213, p-value = 0.08733
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -5.663104 76.055261
## sample estimates:
## mean of x mean of y
## 82.52941 47.33333

t.test(df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo <= 0.05, 'num_sigs_nonexo'],
       df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo > 0.05, 'num_sigs_nonexo'])

## 
## Welch Two Sample t-test
## 
## data: df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo <= 0.05, "num_sigs_nonexo"] and df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo > 0.05, "num_sigs_nonexo"]
## t = 1.466, df = 7.0453, p-value = 0.1858
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.886557 8.063027
```

```

## sample estimates:
## mean of x mean of y
## 8.588235 5.500000
mean(df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo <= 0.05, 'num_sigs_nonexo'])

## [1] 8.588235
mean(df_pvals_DMDL_SP[df_pvals_DMDL_SP$pvals_DMnonexo > 0.05, 'num_sigs_nonexo'])

## [1] 5.5

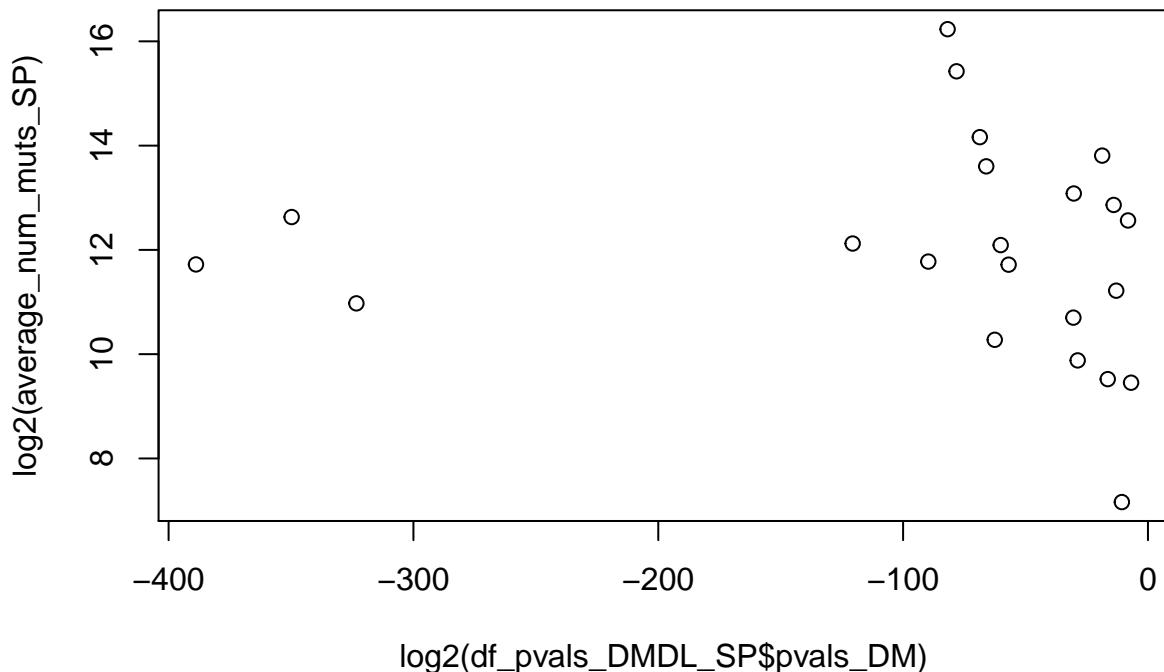
Does DA or not scale with the average number of mutations in the observed exposures (i.e. per patient and group) of the relevant ct?

average_num_muts_SP <- sapply(enough_samples, function(ct){
  .xx <- all_objects_SP[[ct]]
  try(mean(rowSums(.xx$Y)))
})

average_num_muts_SP

##          Bone-Osteosarc      Breast-AdenoCA      CNS-GBM
##                2382.1852        3507.5588       6057.6618
##          CNS-Medullo      CNS-PiloAstro  ColoRect-AdenoCA
##                  700.8915        143.3452       77035.7838
##          Eso-AdenoCA      Head-SCC           Kidney-ChRCC
##                 12439.3692       7445.4219       942.5263
## Kidney-RCC.clearcell Kidney-RCC.papillary      Liver-HCC
##                   3365.9477        2740.9833       6340.5121
##          Lung-SCC          Lymph-BNHL      Lymph-CLL
##                 18340.1471        4367.0392       1239.9057
##          Ovary-AdenoCA      Panc-AdenoCA      Panc-Endocrine
##                   4458.4691        3376.0415       1664.5857
## Prost-AdenoCA Skin-Melanoma.cutaneous      Stomach-AdenoCA
##                  2011.8966        43998.4833      14345.9667
##          Thy-AdenoCA      Uterus-AdenoCA
##                   734.8902        8670.1125
plot(log2(df_pvals_DMDL_SP$pvals_DM), log2(average_num_muts_SP))

```



`log2(average_num_mutations_SP)`

```
t.test(log2(average_num_mutations_SP)[df_pvals_DMDL_SP$pvals_DMnonexo <= 0.05],
       log2(average_num_mutations_SP)[df_pvals_DMDL_SP$pvals_DMnonexo > 0.05])
```

##

## Welch Two Sample t-test

##

## data: log2(average\_num\_mutations\_SP)[df\_pvals\_DMDL\_SP\$pvals\_DMnonexo <= 0.05] and log2(average\_num\_mutations\_SP)[df\_pvals\_DMDL\_SP\$pvals\_DMnonexo > 0.05]

## t = 0.85052, df = 6.3966, p-value = 0.4257

## alternative hypothesis: true difference in means is not equal to 0

## 95 percent confidence interval:

## -1.88287 3.93586

## sample estimates:

## mean of x mean of y

## 12.19856 11.17207

```
pcawg_palette <- pcawg.colour.palette(gsub("\\..*", "", enough_samples), scheme = "tumour.subtype")
names(pcawg_palette) <- enough_samples
```

`pvals_diagRE_DMDL_nonexo_SP_adj`

##	Bone-Osteosarc	Breast-AdenoCA	CNS-GBM
##	8.712593e-04	6.123721e-10	1.557247e-01
##	CNS-Medullo	CNS-PiloAstro	ColoRect-AdenoCA
##	1.557247e-01	5.270878e-01	3.141298e-17
##	Eso-AdenoCA	Head-SCC	Kidney-ChRCC
##	7.328650e-17	1.788732e-04	1.557247e-01
##	Kidney-RCC.clearcell	Kidney-RCC.papillary	Liver-HCC
##	1.137607e-13	1.190455e-06	1.263498e-70
##	Lung-SCC	Lymph-BNHL	Lymph-CLL
##	8.301558e-01	5.881884e-07	2.256505e-03
##	Ovary-AdenoCA	Panc-AdenoCA	Panc-Endocrine

```

##          6.973340e-26          8.893577e-60          4.169125e-05
## Prost-AdenoCA Skin-Melanoma.cutaneous      Stomach-AdenoCA
##          6.044387e-83          9.919659e-17          6.320633e-02
## Thy-AdenoCA           Uterus-AdenoCA
##          1.170750e-02          1.568961e-09

pvals_fullREDMnoscaling_SP_nonexo_subsets_and_amalgamations_adj

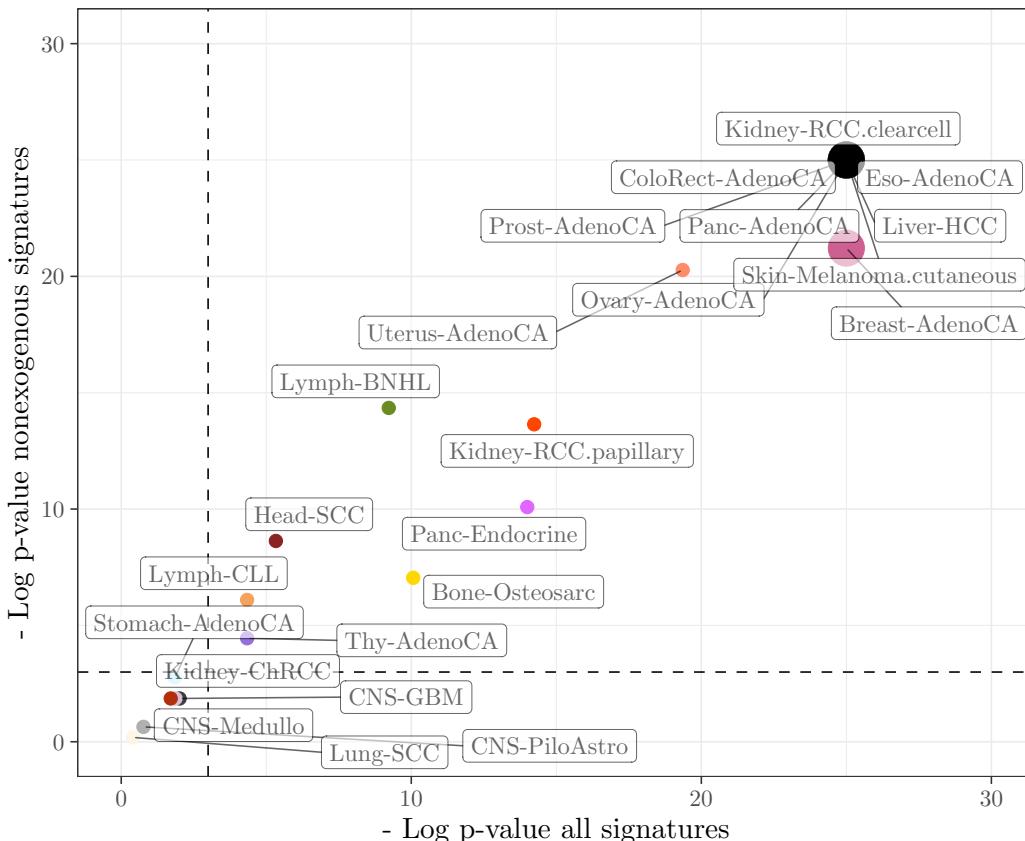
##          Bone-Osteosarc      Breast-AdenoCA          CNS-GBM
##          4.248457e-05          7.759783e-12          1.336685e-01
##          CNS-Medullo        CNS-PiloAstro       ColoRect-AdenoCA
##          1.538381e-01          4.648237e-01          1.023484e-15
##          Eso-AdenoCA          Head-SCC            Kidney-ChRCC
##          3.759670e-18          4.834276e-03          1.814391e-01
## Kidney-RCC.clearcell    Kidney-RCC.papillary      Liver-HCC
##          3.652385e-14          6.557344e-07          3.704923e-63
##          Lung-SCC             Lymph-BNHL            Lymph-CLL
##          6.693019e-01          9.868396e-05          1.303426e-02
##          Ovary-AdenoCA        Panc-AdenoCA       Panc-Endocrine
##          1.231778e-22          7.710539e-59          8.327076e-07
##          Prost-AdenoCA Skin-Melanoma.cutaneous      Stomach-AdenoCA
##          6.572758e-67          3.732926e-14          1.576047e-01
##          Thy-AdenoCA          Uterus-AdenoCA
##          1.303426e-02          3.911013e-09

df_pvals_fullREDMnoscaling_SP_nonexo_subsets_and_amalgamations <- cbind.data.frame(pvals_DMnonexo_nonscal-
                                         pvals_DMnonexo=pvals_diagRE_DMDL_nonexo_SP_adj,
                                         num_samples=as.numeric(num_samples_all_SP),
                                         num_sigs_nonexo=as.numeric(num_sigs_nonexo_SP),
                                         ct=enough_samples,
                                         pvals_DM_nonscaling_censored=sapply(-log(pvals_fullREDMnoscaling_SP_nonexo_subset
                                             function(i) min(i, 25)),
                                         pvals_DMnonexo_censored=sapply(-log(pvals_diagRE_DMDL_nonexo_SP_adj),
                                             function(i) min(i, 25)),
                                         bool_censored=(( -log(pvals_diagRE_DMDL_nonexo_SP_adj) > 25 ) | ( -log(pvals_fullRE
ggplot(df_pvals_fullREDMnoscaling_SP_nonexo_subsets_and_amalgamations,
  aes(x=pvals_DM_nonscaling_censored, y=pvals_DMnonexo_censored,
      # size=num_samples,
      label=ct, size=bool_censored))+geom_point(aes (col=ct))+
  geom_hline(yintercept = -log(0.05), lty='dashed')+geom_vline(xintercept = -log(0.05), lty='dashed')+#
  geom_label_repel(size=3.2, alpha=0.6, max.overlaps = 30)+ theme_bw()+
  theme(legend.position = "bottom", legend.text=element_text(size=8))+#
  labs(x=' - Log p-value all signatures', y=' - Log p-value nonexogenous signatures')+#
  guides(size=FALSE, col=FALSE)+ #, col=guide_legend(ncol=4),
  scale_color_manual(values = pcawg_palette)+theme(legend.position = "bottom")+
  lims(x=c(0, 30), y=c(0,30))

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

## Warning: Using size for a discrete variable is not advised.

```



See script PCAWG\_HMP\_and\_alternative\_methods.R for the analyses of PCAWG data using alternative models.

## Tracksig

```

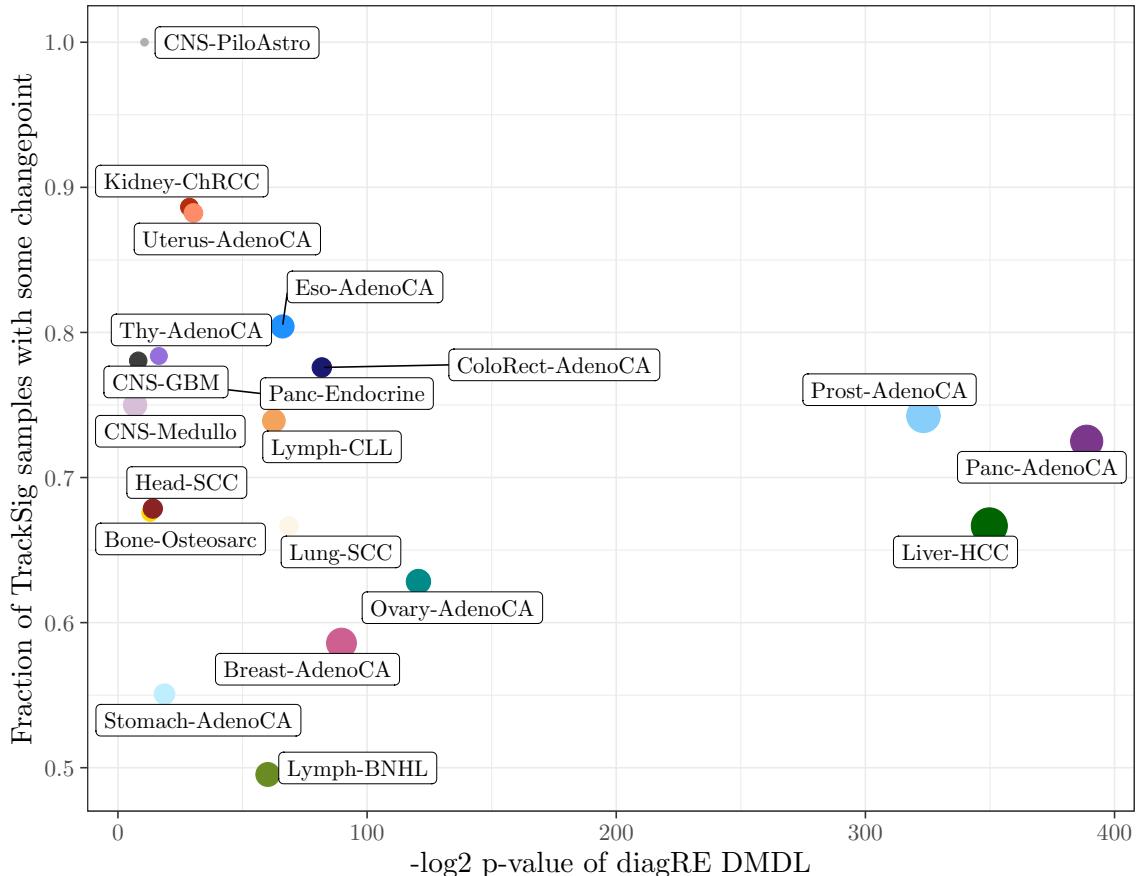
tracksig = read.csv("../data/restricted/tracksig/changepoints_stats_tracksig.csv", stringsAsFactors =
tracksig = tracksig %>% group_by(type) %>%
  dplyr::summarize(count = n(), bool_changepoints=sum(n_changepoints > 0)) %>%
  mutate(tracksig_frac= bool_changepoints/count)
tracksig = cbind.data.frame(pvals_diagRE_DMDL_SP_adj,
                           tracksig[match(names(pvals_diagRE_DMDL_SP_adj), tracksig$type),],
                           effect_size3_SP=effect_size3_SP[match(names(pvals_diagRE_DMDL_SP_adj), names(effect_size3_SP))])
tracksig$ct = rownames(tracksig)
tracksig$minpvals = -log2(tracksig$pvals_diagRE_DM)

pcawg_palette <- pcawg.colour.palette(gsub("\\\\..*", "", tracksig$ct), scheme = "tumour.subtype")
names(pcawg_palette) <- tracksig$ct

ggplot(tracksig, aes(x=-log2(pvals_diagRE_DMDL_SP_adj), y=tracksig_frac, label=ct, size=count))+geom_point()
  labs(x='-\log2 p-value of diagRE DMDL', y='Fraction of TrackSig samples with some changepoint')+ 
  scale_color_manual(values = pcawg_palette)+theme(legend.position = "bottom")+
  # scale_x_continuous(trans = "log2")+
  theme_bw() + theme(legend.position = "bottom")

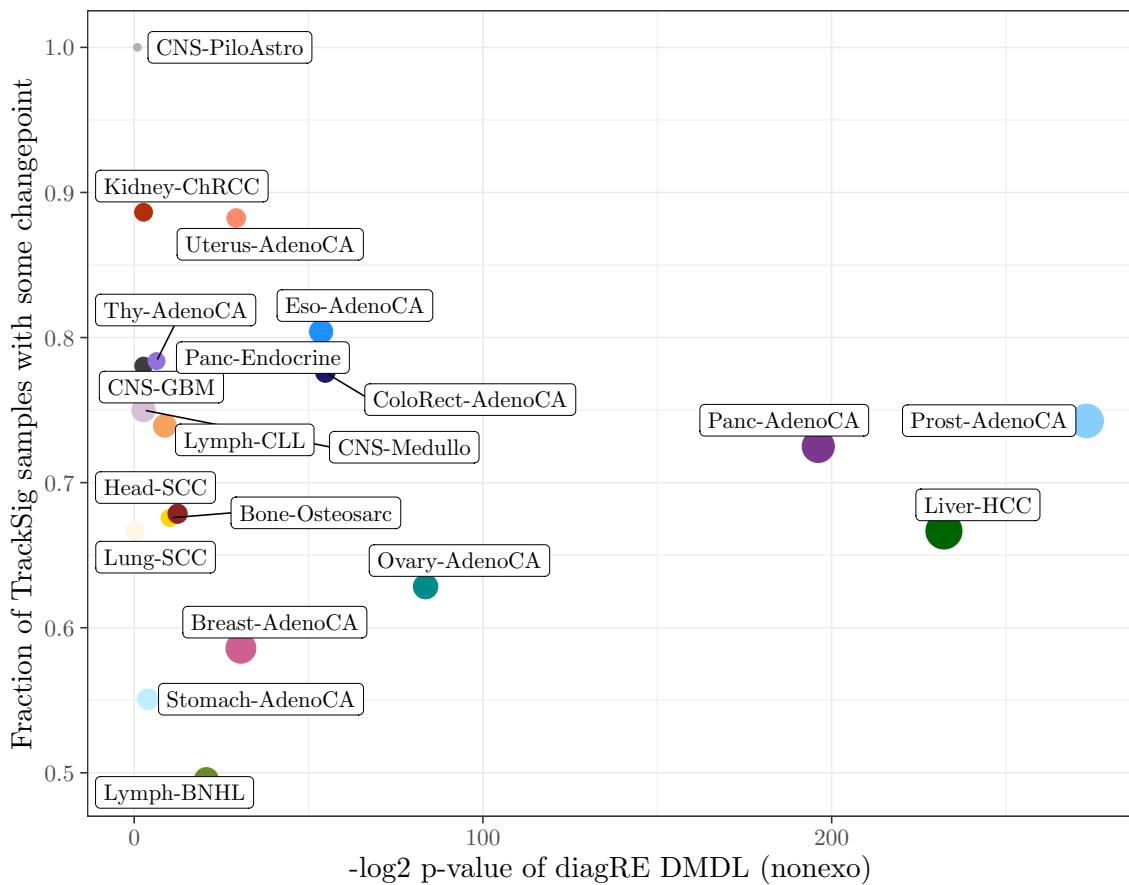
```

```
## Warning: Removed 3 rows containing missing values (geom_point).
## Warning: Removed 3 rows containing missing values (geom_label_repel).
```



```
  300      • Bone-Osteosarc      ● ColoRect-AdenoCA      ○ Kidney-RCC.papillary      ● Ovary-Aden
ct      ● Breast-AdenoCA      ● Eso-AdenoCA      ● Liver-HCC      ● Panc-Adeno
      ● CNS-GBM      ● Head-SCC      ● Lung-SCC      ● Panc-Endoc
      ● CNS-Medullo      ● Kidney-ChRCC      ● Lymph-BNHL      ● Prost-Adeno
      ● CNS-PiloAstro      ● Kidney-RCC.clearcell      ● Lymph-CLL      ● Skin-Melan
ggplot(tracksig, aes(x=-log2(pvals_diagRE_DMDL_nonexo_SP_adj), y=tracksig_frac, label=ct, size=count))+
  labs(x='-log2 p-value of diagRE DMDL (nonexo)', y='Fraction of TrackSig samples with some changepoint')
  scale_color_manual(values = pcawg_palette)+theme(legend.position = "bottom")+
  theme_bw()+theme(legend.position = "bottom")
```

```
## Warning: Removed 3 rows containing missing values (geom_point).
## Warning: Removed 3 rows containing missing values (geom_label_repel).
```



```

  300      ● Bone-Osteosarc   ● ColoRect-AdenoCA   ● Kidney-RCC.papillary   ● Ovary-Aden
          ● Breast-AdenoCA  ● Eso-AdenoCA       ● Liver-HCC                 ● Panc-Adeno
  ct     ● CNS-GBM           ● Head-SCC           ● Kidney-ChRCC             ● Lung-SCC
          ● CNS-Medullo      ● Kidney-ChRCC       ● Lymph-BNHL                ● Panc-Endoc
          ● CNS-PiloAstro     ● Kidney-RCC.clearcell ● Lymph-CLL                 ● Prost-Adenc
          ● CNS-PiloAstro     ● Kidney-RCC.clearcell ● Lymph-CLL                 ● Skin-Melanoc

```

```

ggplot(tracksig, aes(x=effect_size3_SP, y=tracksig_frac, label=ct, col=ct))+  

  geom_point(aes(size=minpvals))+geom_label_repel(max.overlaps = 5)+  

  labs(x='Effect size', y='Fraction of TrackSig samples with some changepoint', col="")  

  +theme_bw()  

  +scale_color_manual(values = pcawg_palette)+theme(legend.position = "bottom")  

  +guides(col=guide_legend(ncol=4), size=FALSE)  

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  

## "none")` instead.  

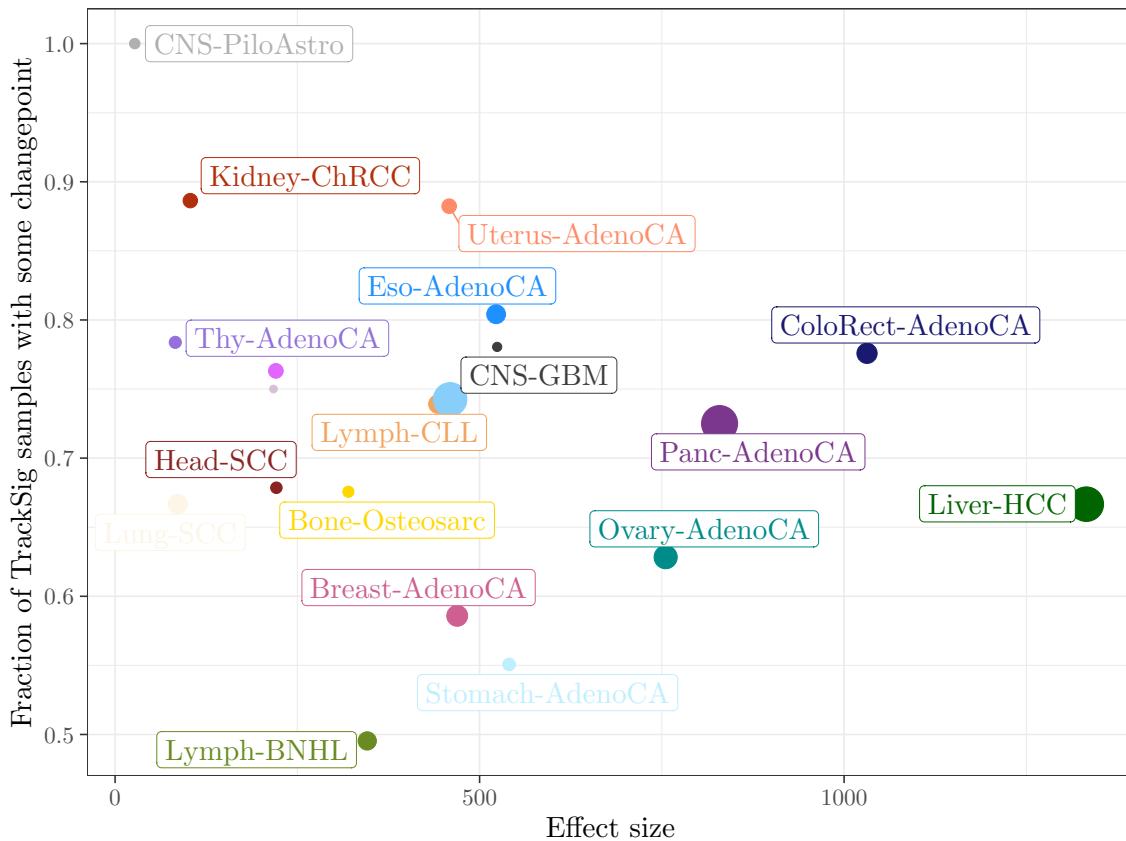
## Warning: Removed 3 rows containing missing values (geom_point).  

## Warning: Removed 3 rows containing missing values (geom_label_repel).  

## Warning: ggrepel: 3 unlabeled data points (too many overlaps). Consider  

## increasing max.overlaps

```



a	Bone-Osteosarc	a	Eso-AdenoCA	a	Lung-SCC	a	Prost-AdenoCA
a	Breast-AdenoCA	a	Head-SCC	a	Lymph-BNHL	a	Skin-Melanoma.cutaneo
a	CNS-GBM	a	Kidney-ChRCC	a	Lymph-CLL	a	Stomach-AdenoCA
a	CNS-Medullo	a	Kidney-RCC.clearcell	a	Ovary-AdenoCA	a	Thy-AdenoCA
a	CNS-PiloAstro	a	Kidney-RCC.papillary	a	Panc-AdenoCA	a	Uterus-AdenoCA
a	ColoRect-AdenoCA	a	Liver-HCC	a	Panc-Endocrine		

Same plots, but smaller, for images

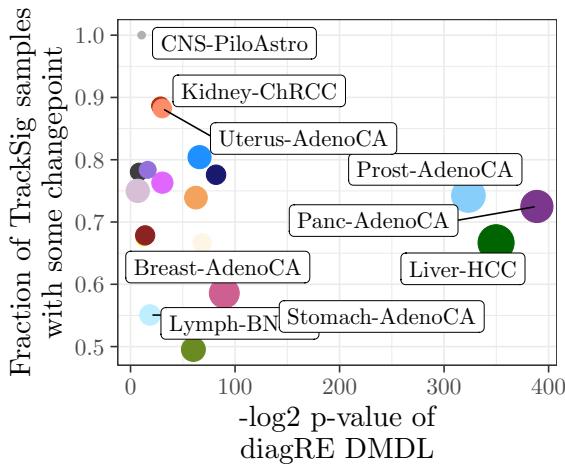
```
ggplot(tracksig, aes(x=-log2(pvals_diagRE_DMDL_SP_adj), y=tracksig_frac, label=ct, size=count))+geom_point()
  labs(x='-log2 p-value of\n diagRE DMDL', y='Fraction of TrackSig samples\n with some changepoint')+ 
  scale_color_manual(values = pcawg_palette)+theme(legend.position = "bottom")+
  # scale_x_continuous(trans = "log2")+
  theme_bw()+theme(legend.position = "bottom")+guides(col=FALSE)+labs(size='N. obs')

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

## Warning: Removed 3 rows containing missing values (geom_point).

## Warning: Removed 3 rows containing missing values (geom_label_repel).

## Warning: ggrepel: 11 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



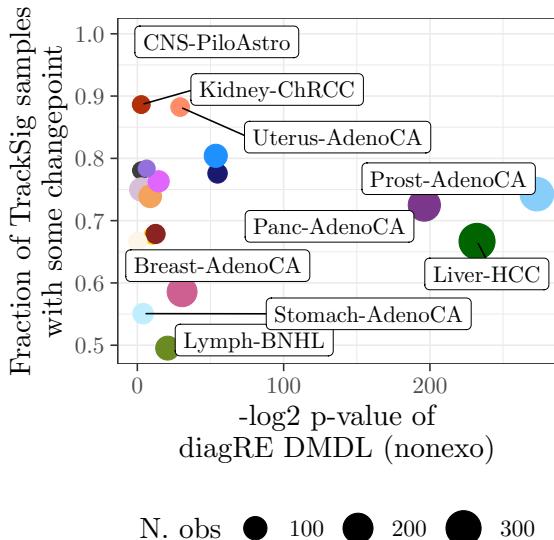
```
ggplot(tracksig, aes(x=-log2(pvals_diagRE_DMDL_nonexo_SP_adj), y=tracksig_frac, label=ct, size=count))+
  geom_label_repel(size=3, col='black', max.overlaps = 4) +
  labs(x='-\log2 p-value of\n diagRE DMDL (nonexo)', y='Fraction of TrackSig samples\n with some changepoint') +
  scale_color_manual(values = pcawg_palette) + theme(legend.position = "bottom") +
  theme_bw() + theme(legend.position = "bottom") + guides(col=FALSE) + labs(size='N. obs')
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

## Warning: Removed 3 rows containing missing values (geom_point).

## Warning: Removed 3 rows containing missing values (geom_label_repel).

## Warning: ggrepel: 11 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



```
plot_for_ct_legend <- ggplot(tracksig, aes(x=-log2(pvals_diagRE_DMDL_nonexo_SP_adj), y=tracksig_frac, label=ct, size=count))+
  geom_label_repel(size=3, col='black', max.overlaps = 2) +
```

```

labs(x='-log2 p-value of\n diagRE DMDL (nonexo)', y='Fraction of TrackSig samples\n with some changepoi
  scale_color_manual(values = pcawg_palette)+theme(legend.position = "bottom")+
theme_bw()+theme(legend.position = "bottom")+labs(size='N. observations')+guides(size=FALSE)+labs(col='')

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

legend_ct <- (cowplot::get_legend(plot_for_ct_legend))

## Warning: Removed 3 rows containing missing values (geom_point).

## Warning: Removed 3 rows containing missing values (geom_label_repel).

# pdf("../results/results_TMB/pcawg/legend_cts.pdf", height = 1, width = 6.5)
grid.newpage()
grid.draw(legend_ct)

● Bone-Osteosarc      ● ColoRect-AdenoCA      ● Kidney-RCC.papillary      ● Ovary-AdenoCA      ● Stomach-AdenoCA
● Breast-AdenoCA     ● Eso-AdenoCA        ● Liver-HCC            ● Panc-AdenoCA       ● Thy-AdenoCA
● CNS-GBM             ● Head-SCC          ● Lung-SCC           ● Panc-Endocrine    ● Uterus-AdenoCA
● CNS-Medullo         ● Kidney-ChRCC        ● Lymph-BNHL          ● Prost-AdenoCA
● CNS-PiloAstro        ● Kidney-RCC.clearcell ● Lymph-CLL           ● Skin-Melanoma.cutaneous

# dev.off()

gerstung_changing_sigs_early_late <- readxl::read_excel("/Users/morril01/Documents/PhD/GlobalDA/data/rest

## New names:
## * ` ` -> ...1

gerstung_changing_sigs_clonal_subclonal <- readxl::read_excel("/Users/morril01/Documents/PhD/GlobalDA/dat

## New names:
## * ` ` -> ...1

# gerstung_changing_sigs <- readxl::read_excel("/Users/morril01/Documents/PhD/GlobalDA/data/restricted/p
# gerstung_changing_sigs <- readxl::read_excel("/Users/morril01/Documents/PhD/GlobalDA/data/restricted/p
# gerstung_constant_sigs <- readxl::read_excel("/Users/morril01/Documents/PhD/GlobalDA/data/restricted/p
# gerstung_changing_sigs
# gerstung_constant_sigs

gerstung_changing_sigs_earlylate <- gerstung_changing_sigs_early_late #gerstung_changing_sigs[gerstung_ch
gerstung_changing_sigs_clonalsubclonal <- gerstung_changing_sigs_clonal_subclonal #gerstung_changing_sigs

gerstung_changing_sigs_earlylate$signature[(gerstung_changing_sigs_earlylate$signature == "SBS6.14.15.20.
gerstung_changing_sigs_clonalsubclonal$signature[(gerstung_changing_sigs_clonalsubclonal$signature == "SE
gerstung_changing_sigs_earlylate$signature <- gsub("_", ".", gerstung_changing_sigs_earlylate$signature)
gerstung_changing_sigs_clonalsubclonal$signature <- gsub("_", ".", gerstung_changing_sigs_clonalsubclonal

# df_changes_el <- gerstung_changing_sigs_earlylate %>% group_by(signature) %>% dplyr::summarise(median(
# df_changes_cs <- gerstung_changing_sigs_clonalsubclonal %>% group_by(signature) %>% dplyr::summarise(m
df_changes_el <- gerstung_changing_sigs_earlylate %>% group_by(signature) %>% dplyr::summarise(median(log(
df_changes_cs <- gerstung_changing_sigs_clonalsubclonal %>% group_by(signature) %>% dplyr::summarise(medi

```

```

# grid.arrange(ggplot(gerstung_changing_sigs_earlylate, aes(x=factor(signature, levels=df_changes_el$signature),
#                                         y=mean_change, group=signature, col=histology_abbreviation))+geom_boxplot()
#                                         geom_hline(yintercept = 0, lty='dashed')+guides(col=FALSE)+theme_bw())
#     theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Early vs late'),
# ggplot(gerstung_changing_sigs_clonalsubclonal, aes(x=factor(signature, levels=df_changes_cs$signature,
#                                         y=mean_change, group=signature, col=histology_abbreviation))+geom_boxplot()+
#     geom_jitter()+guides(col=FALSE)+theme_bw())
#     theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Clonal vs subclonal'),
# nrow=2)

## removing cancer types where there aren't many observations
select_self <- function(i) i[i]
gerstung_changing_sigs_earlylate <- gerstung_changing_sigs_earlylate %>% dplyr::filter(signature %in% names(select_self))
gerstung_changing_sigs_clonalsubclonal <- gerstung_changing_sigs_clonalsubclonal %>% dplyr::filter(signature %in% names(select_self))

pcawg_palette <- pcawg.colour.palette(x = gsub("-", ".", tolower(gsub("\\\\..*", "", sort(unique(gerstung_changing_sigs_earlylate$tumour_type)))))

## Warning in pcawg.colour.palette(x = gsub("-", ".", tolower(gsub("\\\\..*", :, :
## Unrecognized input value for x. Default to fill.colour.
names(pcawg_palette) <- sort(unique(gerstung_changing_sigs_earlylate$tumour_type))

unique(gerstung_changing_sigs_clonal_subclonal$signature)

## [1] "DBS2"                  "DBS5"                  "ID1"
## [4] "ID2"                   "ID8"                   "n_unassigned"
## [7] "SBS1"                   "SBS2.13"                "SBS3"
## [10] "SBS40"                 "SBS4"                  "SBS16"
## [13] "SBS4"                   "SBS5"                  "SBS6"
## [16] "DBS9"                   "DBS10"                 "DBS3"
## [19] "SBS10"                 "SBS28"                 "DBS11"
## [22] "SBS18"                 "SBS12"                 "SBS17"
## [25] "DBS7"                   "SBS37"                 "SBS9"
## [28] "SBS8"                   "SBS41"                 "SBS35"
## [31] "DBS1"                   "ID13"                  "SBS7"
## [34] "SBS54"                 "DBS8"                  "SBS38"
## [37] "SBS22"                 "SBS34"                 "SBS56"
## [40] "SBS29"                 "SBS6.14.15.20.21.26.44" "SBS36"
## [43] "SBS30"                 "SBS31"                 "SBS33"
## [46] "SBS39"                 "SBS24"                 "SBS11"
## [49] "SBS19"                 "SBS23"

gerstung_changing_sigs_earlylate$signature <- factor(gerstung_changing_sigs_earlylate$signature,
                                                       levels=df_changes_el$signature[order(df_changes_el$signature)])
gerstung_changing_sigs_clonalsubclonal$signature <- factor(gerstung_changing_sigs_clonalsubclonal$signature,
                                                       levels=df_changes_cs$signature[order(df_changes_cs$signature)])

# grid.arrange(ggplot(gerstung_changing_sigs_earlylate, aes(x=signature,
#                                         y=log2fc_earlyLate, group=signature, col=tumor_type)))

```

```

# geom_hline(yintercept = 0, lty='dashed')+  

# guides(col=FALSE)+theme_bw()+
# theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Early vs late')  

# scale_color_manual(values = pcawg_palette)+labs(x='', y='Log 2 fold-change between early  

# ggplot(gerstung_changing_sigs_clonalsubclonal, aes(x=signature,  

# y=log2fc_clonalSubclonal, group=signature)  

# geom_hline(yintercept = 0, lty='dashed')+  

# geom_boxplot()+ geom_jitter(alpha=0.2)+guides(col=FALSE)+theme_bw()+
# theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Clonal vs subclonal')  

# scale_color_manual(values = pcawg_palette)+labs(x='', y='Log 2 fold-change between clonal vs  

# nrow=2)  

ggplot(gerstung_changing_sigs_earlylate, aes(x=signature,  

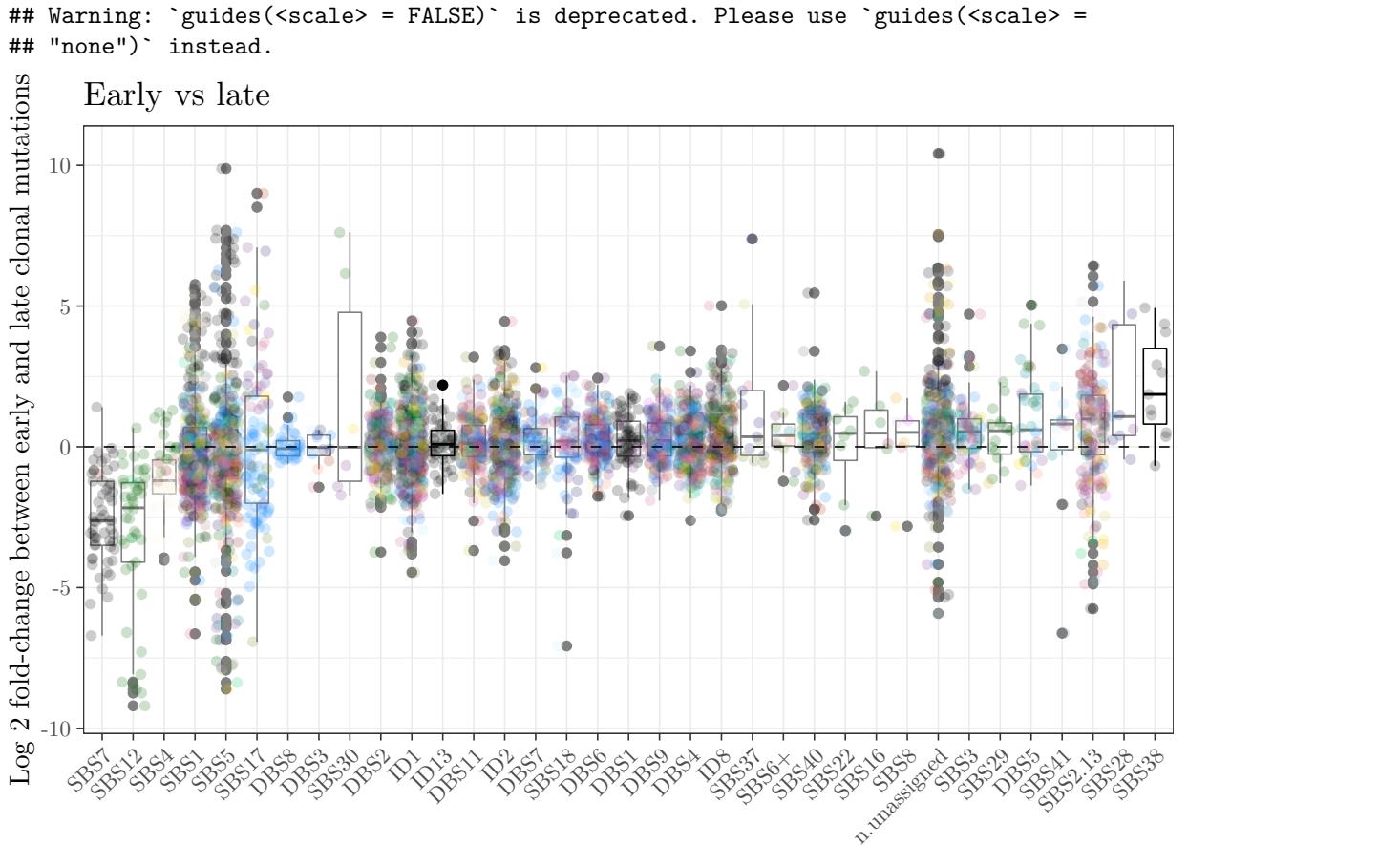
y=log2fc_earlyLate, group=signature,col=tumour)  

geom_hline(yintercept = 0, lty='dashed')+  

guides(col=FALSE)+theme_bw()+
theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Early vs late')  

scale_color_manual(values = pcawg_palette)+labs(x='signature', y='Log 2 fold-change between early and late clonal mutations')

```



```

ggplot(gerstung_changing_sigs_clonalsubclonal, aes(x=signature,  

y=log2fc_clonalSubclonal, group=signature)  

geom_hline(yintercept = 0, lty='dashed')+  

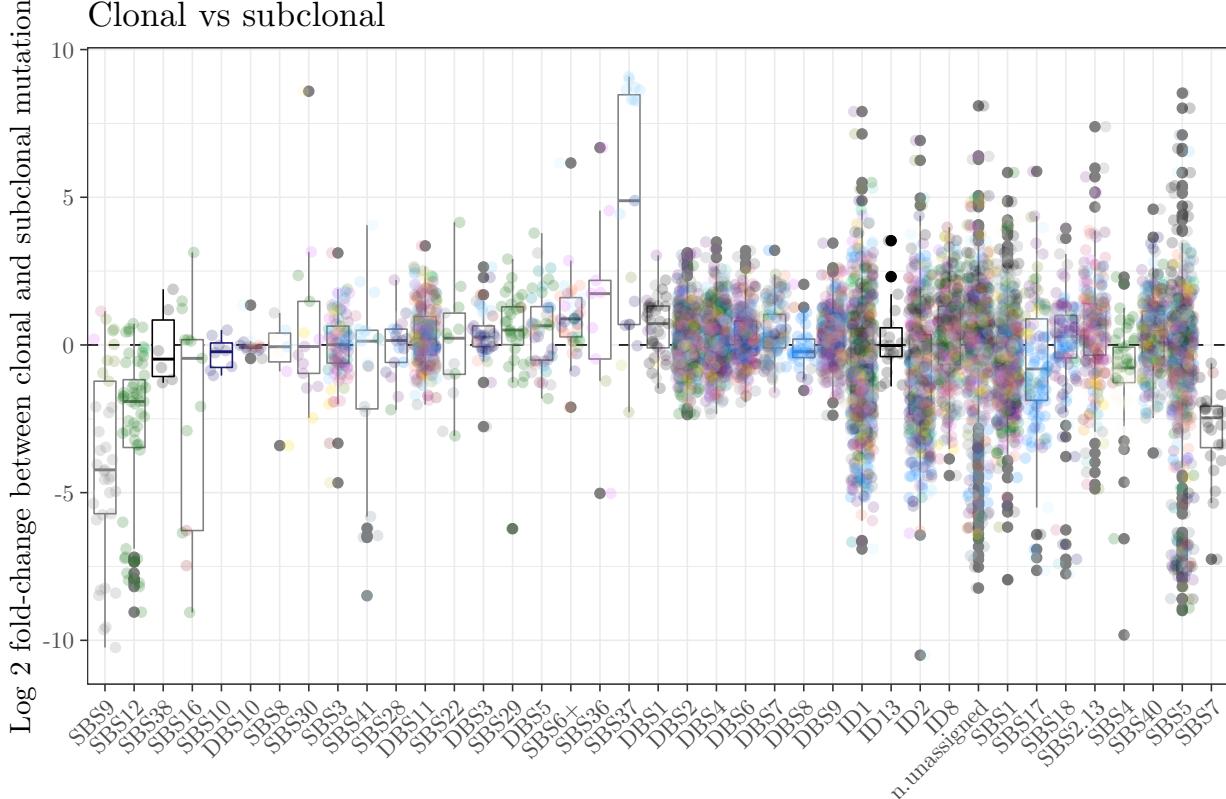
geom_boxplot()+ geom_jitter(alpha=0.2)+guides(col=FALSE)+theme_bw()+
theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Clonal vs subclonal')

```

```

  scale_color_manual(values = pcawg_palette)+labs(x=' ', y='Log 2 fold-change between clonal
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
## Warning: Removed 38 rows containing non-finite values (stat_boxplot).
## Warning: Removed 38 rows containing missing values (geom_point).

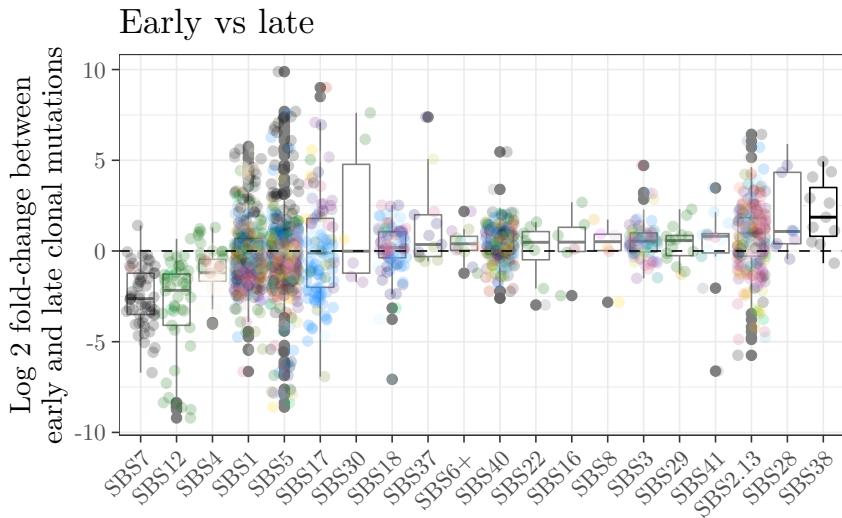
```



```

ggplot(gerstung_changing_sigs_earlylate[grepl('SBS', gerstung_changing_sigs_earlylate$signature),], aes(x=
y=log2fc_earlyLate, group=signature,col=tumour_
geom_hline(yintercept = 0, lty='dashed')+guides(col=FALSE)+theme_bw()+
theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Early vs late')+scale_color_manual(values = pcawg_palette)+labs(x=' ', y='Log 2 fold-change between \nearly
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

```

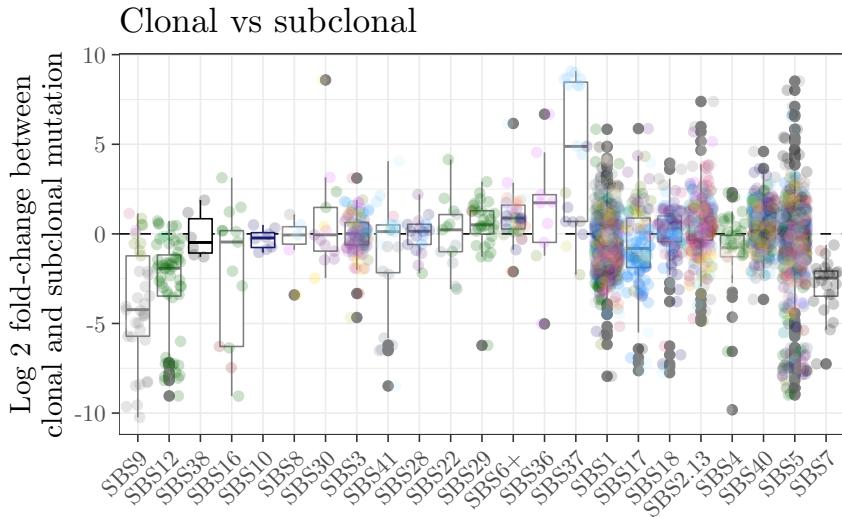


```
ggplot(gerstung_changing_sigs_clonalsubclonal[grep('SBS', gerstung_changing_sigs_clonalsubclonal$signature),
                                               y=log2fc_clonalSubclonal, group=signature]) +
  geom_hline(yintercept = 0, lty='dashed') +
  geom_boxplot() + geom_jitter(alpha=0.2)+guides(col=FALSE)+theme_bw()+
  theme(axis.text.x=element_text(angle = 45, hjust = 1, vjust=1))+ggtitle('Clonal vs subclonal')
  scale_color_manual(values = pcawg_palette)+labs(x=' ', y='Log 2 fold-change between \nclonal and subclonal mutation')

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")` instead.

## Warning: Removed 13 rows containing non-finite values (stat_boxplot).

## Warning: Removed 13 rows containing missing values (geom_point).
```



```
df_changes_el_persample <- gerstung_changing_sigs_earlylate %>% group_by(samplename) %>% dplyr:::summarise()
df_changes_cs_persample <- gerstung_changing_sigs_clonalsubclonal %>% group_by(samplename) %>% dplyr:::summarise()
gerstung_changing_sigs_earlylate$samplename <- factor(gerstung_changing_sigs_earlylate$samplename,
                                                       levels=df_changes_el_persample$samplename[order])
gerstung_changing_sigs_clonalsubclonal$samplename <- factor(gerstung_changing_sigs_clonalsubclonal$samplename,
                                                               levels=df_changes_cs_persample$samplename[order])
```

```

table(is.na(gerstung_changing_sigs_earlylate$samplename))

##
## FALSE
## 5347

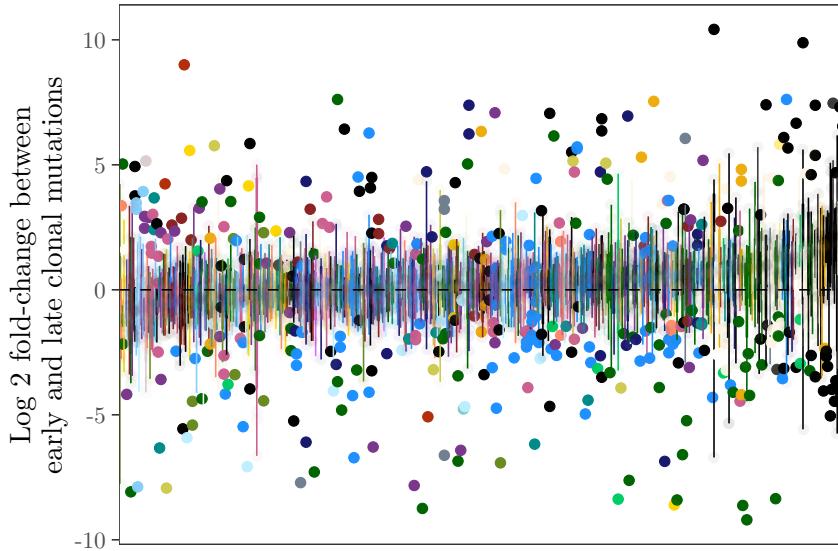
table(is.na(gerstung_changing_sigs_clonalsubclonal$samplename))

##
## FALSE
## 7216

ggplot(gerstung_changing_sigs_earlylate, aes(x=samplename,
                                               y=log2fc_earlyLate, group=samplename, col=tumour_type))+
  geom_hline(yintercept = 0, lty='dashed')+
  guides(col=FALSE)+theme_bw()+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())+
  scale_color_manual(values = pcawg_palette)+labs(x='', y='Log 2 fold-change between\nnearly and late clonal mutations')+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())

```

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")` instead.



```

ggplot(gerstung_changing_sigs_clonalsubclonal, aes(x=samplename,
                                                       y=log2fc_clonalSubclonal, group=samplename, col=tumour_type))+
  geom_hline(yintercept = 0, lty='dashed')+
  geom_boxplot()+
  geom_jitter(alpha=0.05)+guides(col=FALSE)+theme_bw()+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())+ggtitle('Clonal vs subclonal')+
  scale_color_manual(values = pcawg_palette)+labs(x='', y='Log 2 fold-change between\nnclonal and subclonal mutations')+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())

```

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")` instead.

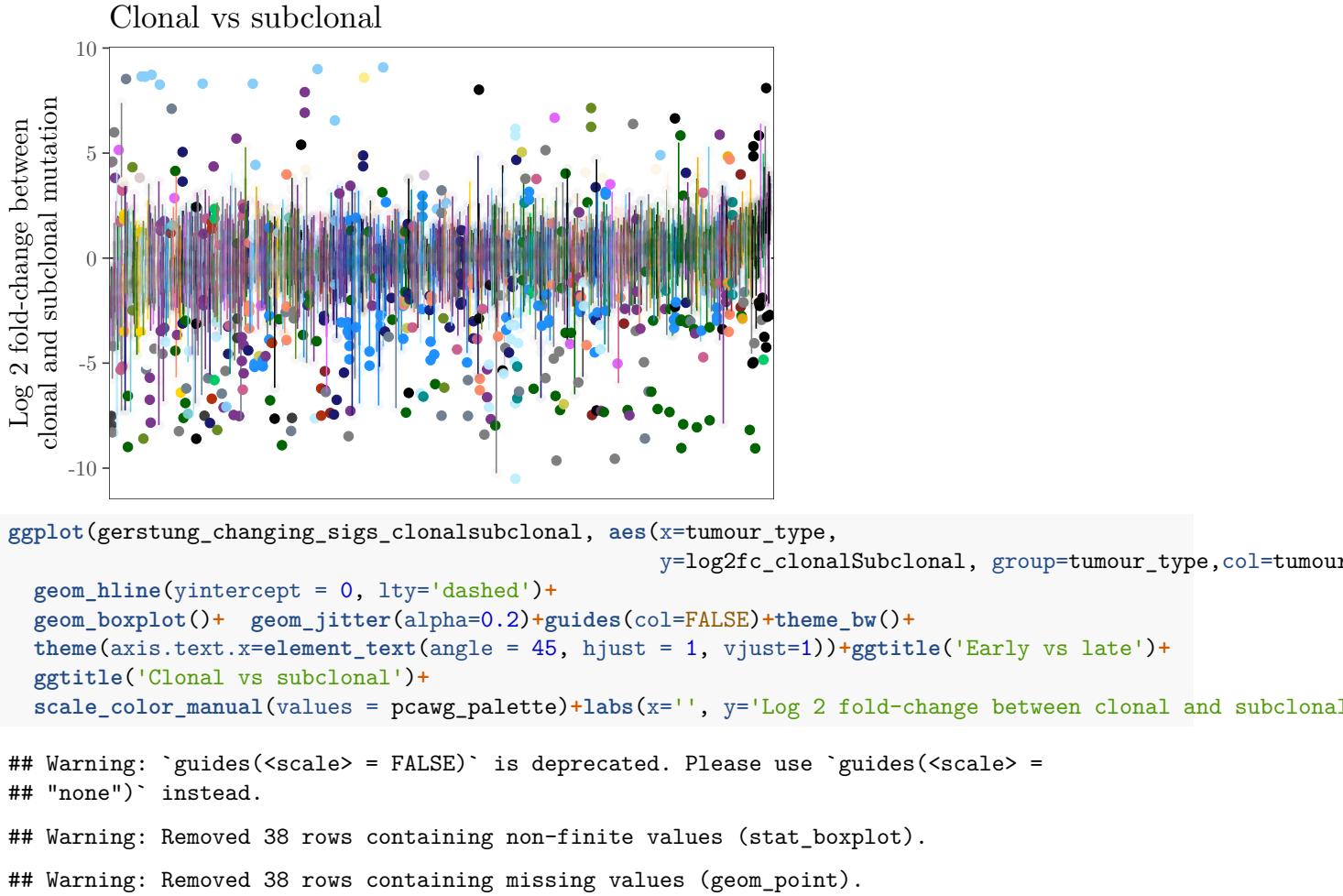
```

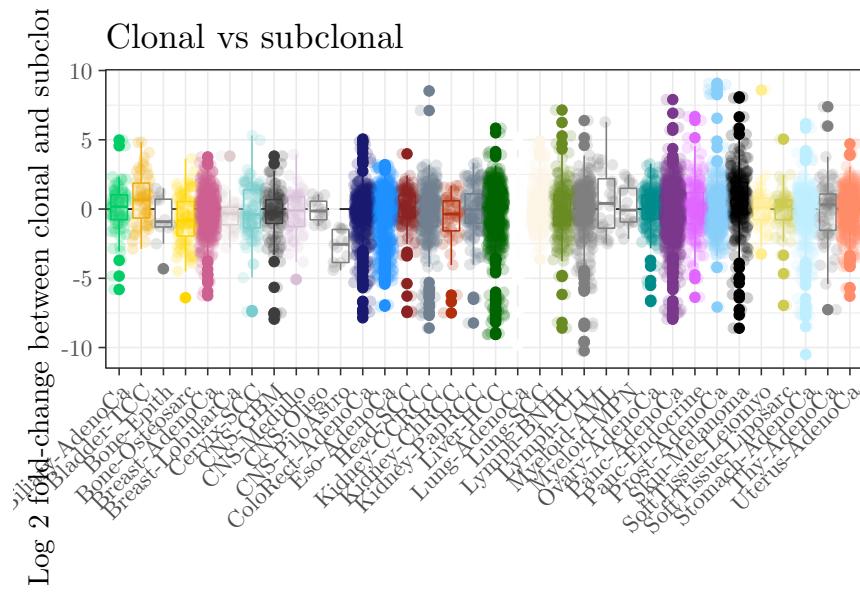
## "none")` instead.

## Warning: Removed 38 rows containing non-finite values (stat_boxplot).

## Warning: Removed 38 rows containing missing values (geom_point).

```





```

gerstung_changing_sigs_clonalsubclonal_var <- gerstung_changing_sigs_clonalsubclonal %>%
  dplyr::group_by(tumour_type) %>% summarise(varlog2=var(log2fc_clonalSubclonal), count=n())
ggplot(gerstung_changing_sigs_clonalsubclonal_var,
       aes(x=varlog2, y=count, label=tumour_type, col=tumour_type)) + geom_point() +
  scale_color_manual(values = pcawg_palette) + guides(col=FALSE) + theme_bw() + geom_label_repel(max.overlaps =

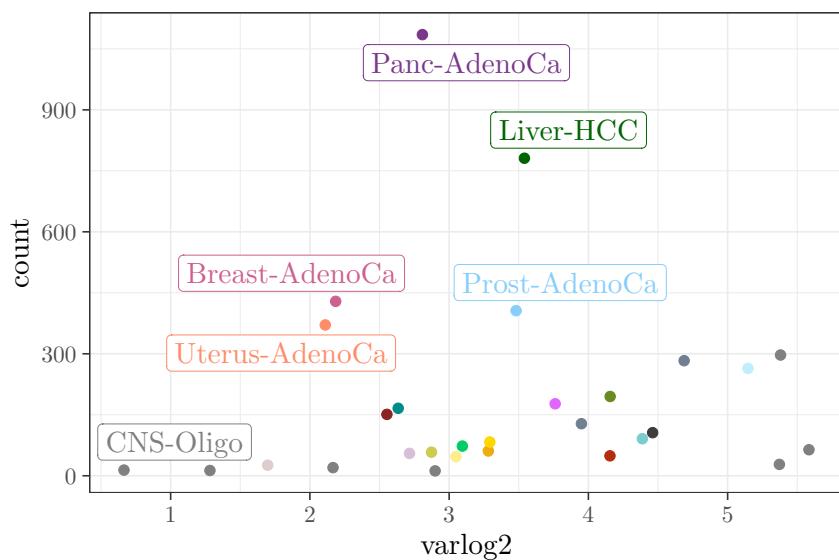
```

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")` instead.

## Warning: Removed 4 rows containing missing values (geom\_point).

## Warning: Removed 4 rows containing missing values (geom\_label\_repel).

## Warning: ggrepel: 24 unlabeled data points (too many overlaps). Consider increasing max.overlaps



```

  labs(x='Variance of log2 fold change in cancer type', y='Number of observations')

## $x
## [1] "Variance of log2 fold change in cancer type"
##
## $y
## [1] "Number of observations"
##
## attr(,"class")
## [1] "labels"

perturbed_betas_diagRE_DMDL_nonexo_SP_df_summary <- perturbed_betas_diagRE_DMDL_nonexo_SP_df %>% dplyr::g
  summarise(meanperturbed=mean( !(perturbed == 'FALSE')))

perturbed_betas_diagRE_DMDL_nonexo_SP_df_summary

## # A tibble: 32 x 2
##       sig   meanperturbed
##   <chr>      <dbl>
## 1 SBS10a      0
## 2 SBS10b      0
## 3 SBS12      1
## 4 SBS13     0.25
## 5 SBS14     0.5
## 6 SBS15    0.333
## 7 SBS16     0.5
## 8 SBS17a    0.364
## 9 SBS17b    0.0909
## 10 SBS18     0.3
## # ... with 22 more rows

comparison_with_gerstung_earlylate <- cbind.data.frame(perturbed_betas_diagRE_DMDL_nonexo_SP_df_summary,
                                                       df_changes_el[match(perturbed_betas_diagRE_DM
                                                       df_changes_el$signature),

comparison_with_gerstung_earlylate

##       sig meanperturbed signature median(log2fc_earlyLate)
## 1 SBS10a 0.000000000 <NA>          NA
## 2 SBS10b 0.000000000 <NA>          NA
## 3 SBS12 1.000000000 SBS12 -2.16839553
## 4 SBS13 0.250000000 <NA>          NA
## 5 SBS14 0.500000000 <NA>          NA
## 6 SBS15 0.333333333 <NA>          NA
## 7 SBS16 0.500000000 SBS16 0.48743401
## 8 SBS17a 0.36363636 <NA>          NA
## 9 SBS17b 0.09090909 <NA>          NA
## 10 SBS18 0.30000000 SBS18 0.19987446
## 11 SBS19 0.50000000 SBS19 -1.96353701
## 12 SBS2 0.17647059 <NA>          NA
## 13 SBS20 1.00000000 <NA>          NA
## 14 SBS21 0.00000000 <NA>          NA
## 15 SBS22 0.00000000 SBS22 0.47923554
## 16 SBS23 0.00000000 <NA>          NA

```

```

## 17 SBS24 0.00000000 SBS24 -10.24985022
## 18 SBS26 0.33333333 <NA> NA
## 19 SBS28 0.50000000 SBS28 1.07561394
## 20 SBS3 0.09090909 SBS3 0.53915013
## 21 SBS30 0.20000000 SBS30 -0.01377077
## 22 SBS33 0.50000000 <NA> NA
## 23 SBS34 0.00000000 SBS34 -0.20973256
## 24 SBS36 0.50000000 SBS36 0.99203837
## 25 SBS37 0.40000000 SBS37 0.35693398
## 26 SBS38 0.00000000 SBS38 1.86314214
## 27 SBS39 0.50000000 SBS39 -0.28583208
## 28 SBS40 0.25000000 SBS40 0.46436260
## 29 SBS41 0.00000000 SBS41 0.80987484
## 30 SBS6 0.20000000 <NA> NA
## 31 SBS8 0.25000000 SBS8 0.51226887
## 32 SBS9 0.33333333 SBS9 1.61493581

```

```

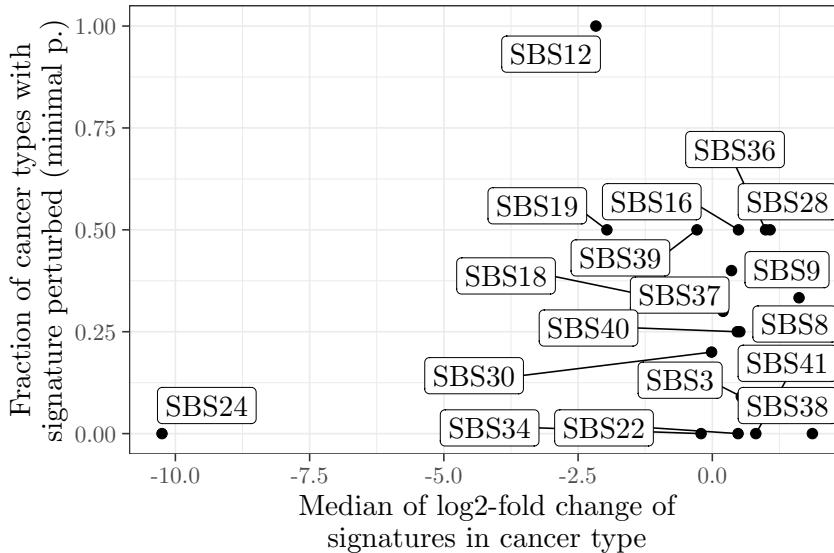
comparison_with_gerstung_earlylate$medianlog2fcearlylate = comparison_with_gerstung_earlylate$`median(log2fc)`

ggplot(comparison_with_gerstung_earlylate, aes(x=medianlog2fcearlylate, y=meanperturbed, label=signature))
  labs(x='Median of log2-fold change of\nsignatures in cancer type', y='Fraction of cancer types with\nsignatures')

## Warning: Removed 14 rows containing missing values (geom_point).

## Warning: Removed 14 rows containing missing values (geom_label_repel).

```



Barplots of cancer types with and without differential abundance

```

give_barplot_from_obj(obj = signatures_PCAWG[['CNS-Medullo']], legend_on = F,
                      nrow=1, verbose=F,
                      only_normalised=T)

```

```

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

```

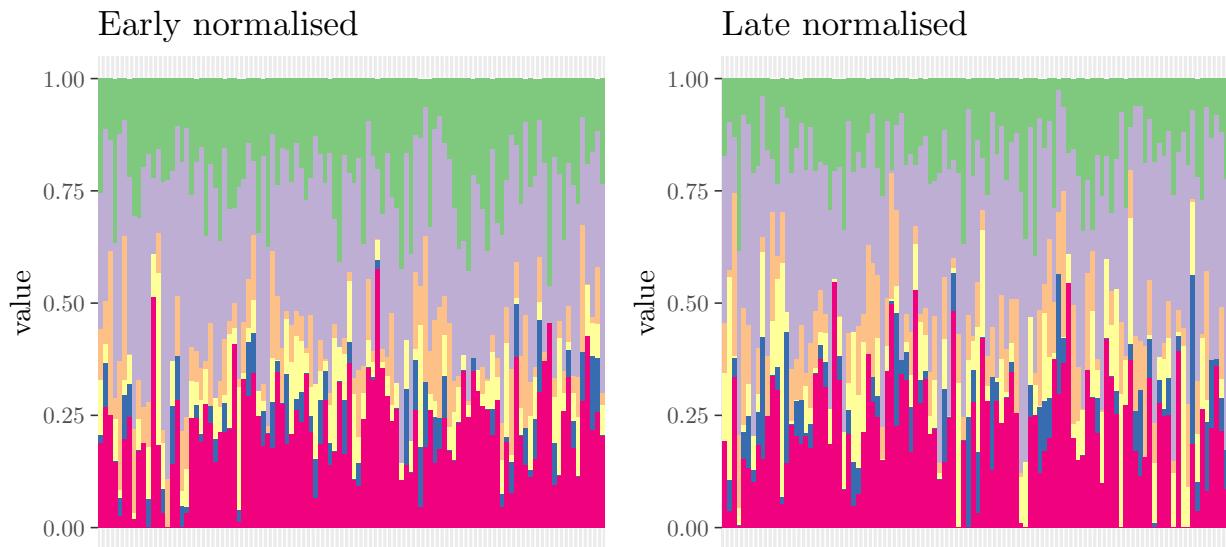
```

## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.

```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  
## "none")` instead.
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  
## "none")` instead.
```



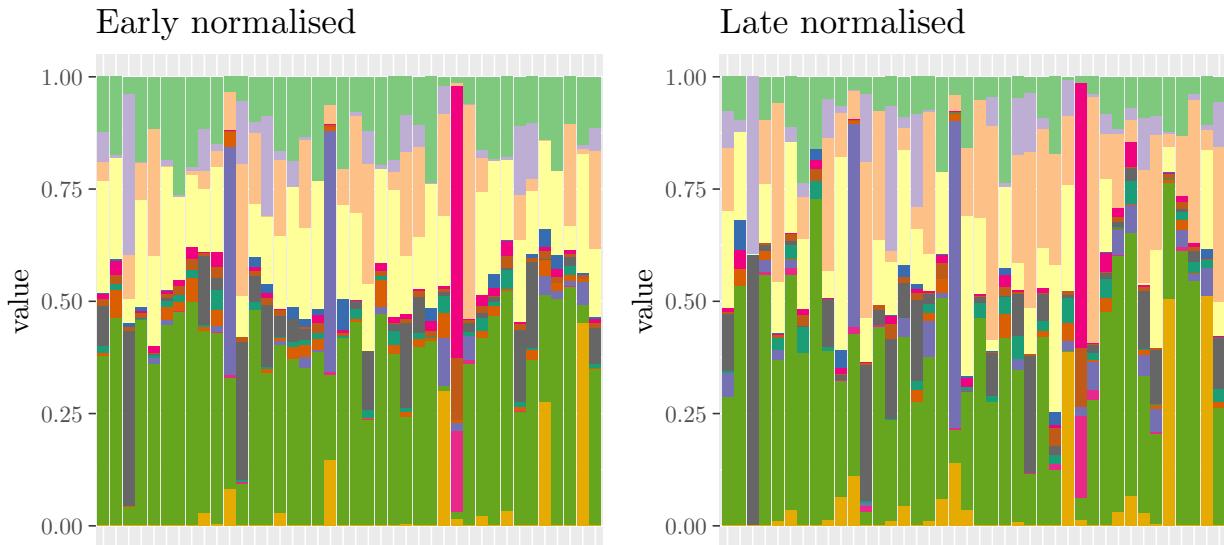
```
give_barplot_from_obj(obj = signatures_PCAWG[['Uterus-AdenoCA']], legend_on = F,  
                      nrow=1, verbose=F,  
                      only_normalised=T)
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  
## "none")` instead.
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  
## "none")` instead.
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  
## "none")` instead.
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =  
## "none")` instead.
```



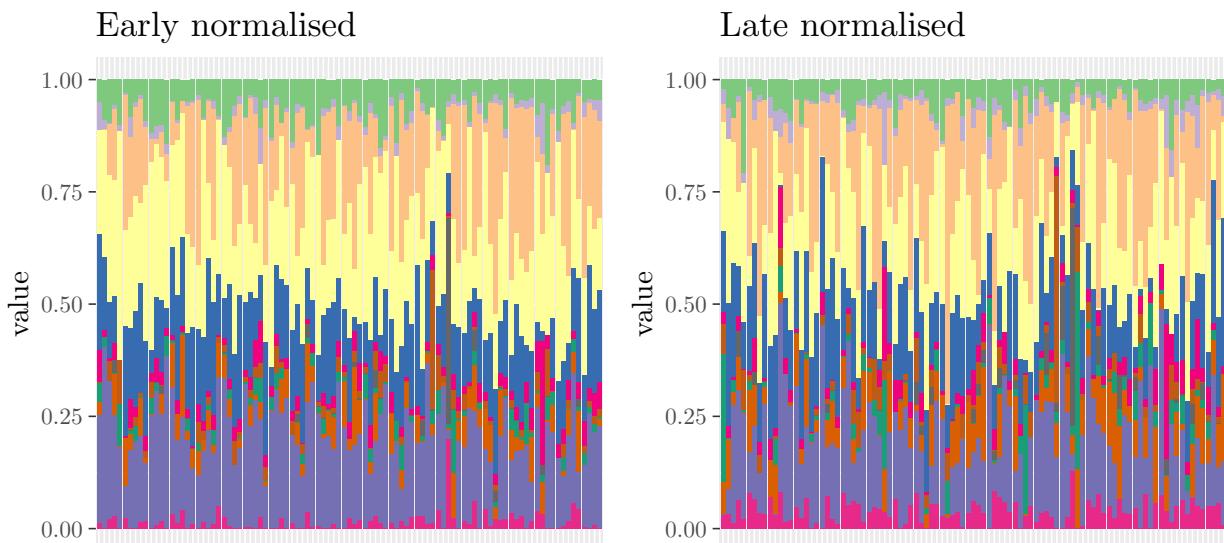
```
give_barplot_from_obj(obj = signatures_PCAWG[['Ovary-AdenoCA']], legend_on = F,
                      nrow=1, verbose=F,
                      only_normalised=T)
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```

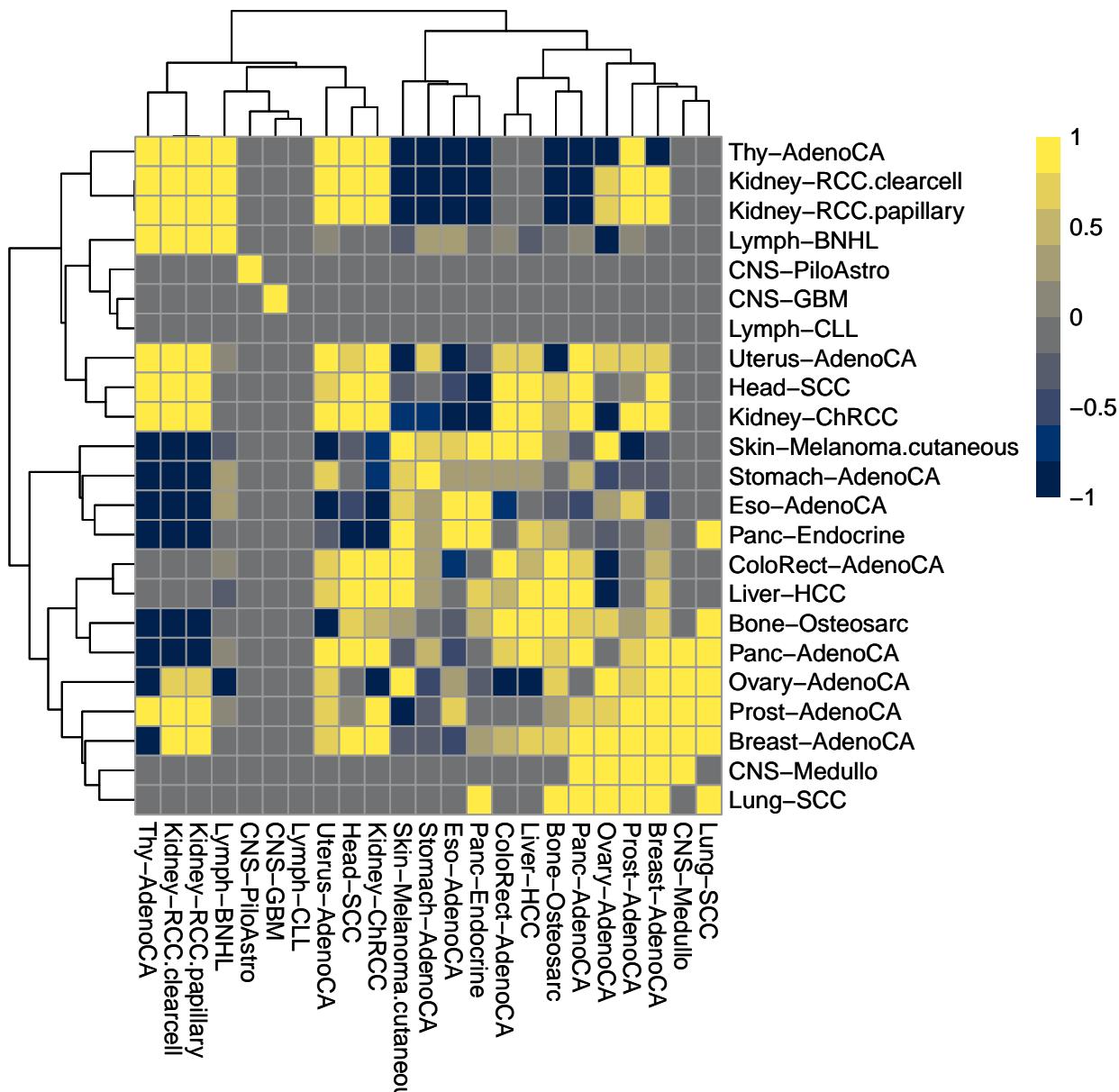
```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```



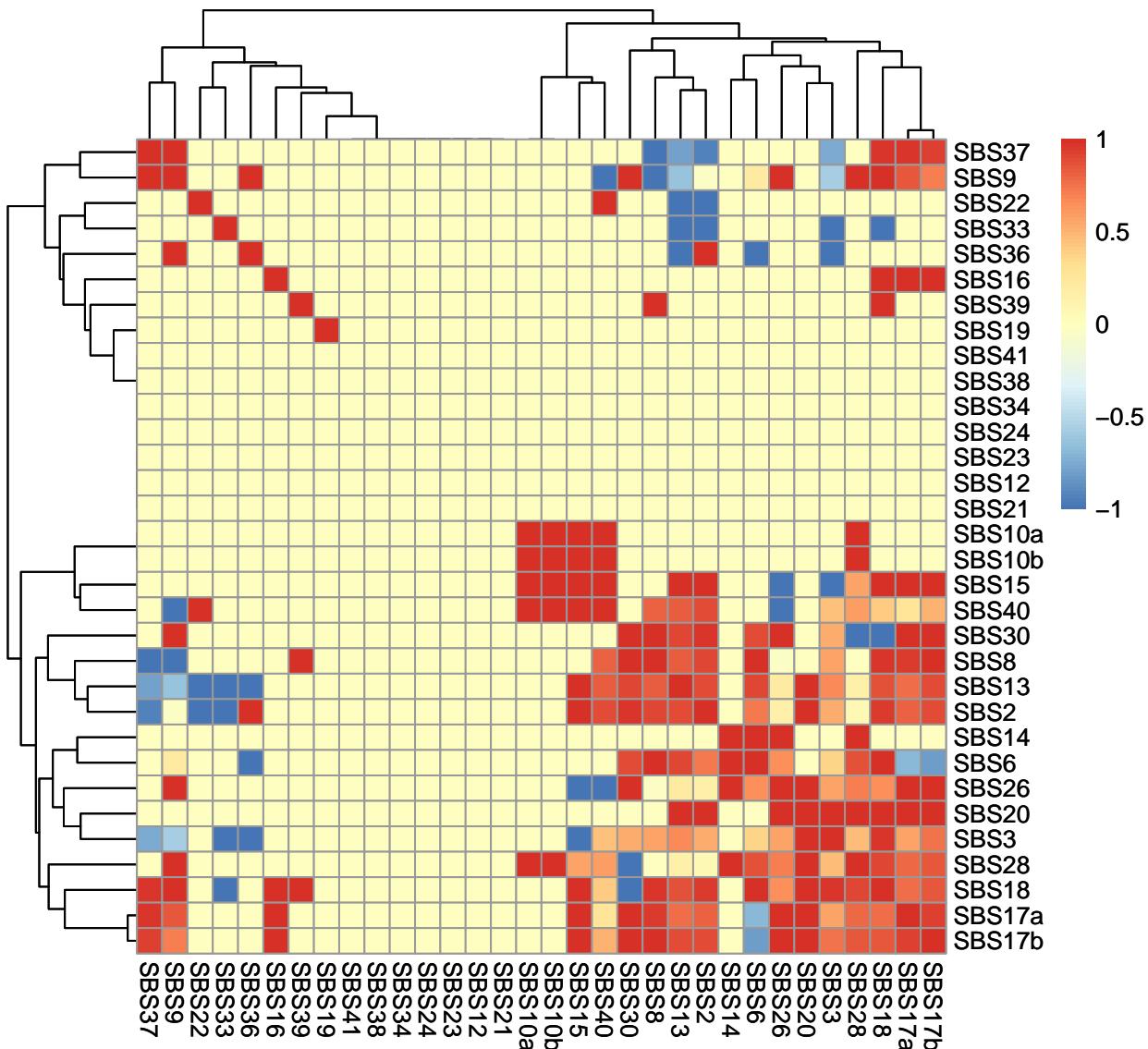
## Correlations of cancer types and of signatures based on betas



```

## null device
##      1
## null device
##      1

```



```
## null device  
## 1  
## null device  
## 1
```